

配置手册

RG-S7500系列交换机

S7500_RGOS11.0(4)B16

文档版本 : V1.0

版权声明

copyright © 2017 锐捷网络

保留对本文档及本声明的一切权利。

未得到锐捷网络的书面许可，任何单位和个人不得以任何方式或形式对本文档的部分内容或全部进行复制、摘录、备份、修改、传播、翻译成其他语言、将其全部或部分用于商业用途。



以上均为锐捷网络的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

免责声明

您所购买的产品、服务或特性等应受商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，锐捷网络对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。锐捷网络保留在没有任何通知或者提示的情况下对文档内容进行修改的权利。

本手册仅作为使用指导。锐捷网络在编写本手册时已尽力保证其内容准确可靠，但并不确保手册内容完全没有错误或遗漏，本手册中的所有信息也不构成任何明示或暗示的担保。

前 言

读者对象

本书适合下列人员阅读

- 网络工程师
- 技术推广人员
- 网络管理员

技术支持

- 锐捷网络官方网站：<http://www.ruijie.com.cn/>
- 锐捷网络在线客服：<http://webchat.ruijie.com.cn>
- 锐捷网络官方网站服务与支持版块：<http://www.ruijie.com.cn/service.aspx>
- 7×24 小时技术服务热线：4008-111-000
- 锐捷网络技术论坛：<http://bbs.ruijie.com.cn/portal.php>
- 常见问题搜索：<http://www.ruijie.com.cn/service/known.aspx>
- 锐捷网络技术支持与反馈信箱：4008111000@ruijie.com.cn

本书约定

1. 命令行格式约定

命令行格式意义如下：

粗体：命令行关键字（命令中保持不变必须照输的部分）采用加粗字体表示。

斜体：命令行参数（命令中必须由实际值进行替代的部分）采用斜体表示

[]：表示用[]括起来的部分，在命令配置时是可选的。

{ x | y | ... }：表示从两个或多个选项中选取一个。

[x | y | ...]：表示从两个或多个选项中选取一个或者不选。

//：由双斜杠开始的行表示为注释行。

2. 各类标志

本书还采用各种醒目标志来表示在操作过程中应该特别注意的地方，这些标志的意义如下：

-
-  警告标志。表示用户必须严格遵守的规则。如果忽视此类信息，可能导致人身危险或设备损坏。
 -  注意标志。表示用户必须了解的重要信息。如果忽视此类信息，可能导致功能失效或性能降低。
 -  说明标志。用于提供补充、申明、提示等。如果忽视此类信息，不会导致严重后果。
 -  产品/版本支持情况标志。用于提供产品或版本支持情况的说明。
-

3. 说明

- 本手册举例说明部分的端口类型同实际可能不符，实际操作中需要按照各产品所支持的端口类型进行配置。
- 本手册部分举例的显示信息中可能含有其它产品系列的内容（如产品型号、描述等），具体显示信息请以实际使用的设备信息为准。
- 本手册中涉及的路由器及路由器产品图标，代表了一般意义下的路由器，以及运行了路由协议的三层交换机。



配置指南-系统配置

本分册介绍系统配置配置指南相关内容，包括以下章节：

1. 命令行界面
2. 基础管理
3. LINE
4. TIME RANGE
5. HTTP 服务
6. 系统日志
7. CWMP
8. CA-MONITOR
9. 软件授权管理
10. 模块热插拔
11. 管理板冗余
12. USB
13. UFT
14. PKG_MGMT

1 命令行界面

1.1 概述

命令行界面(Command Line Interface , CLI)是用户与网络设备进行文本指令交互的窗口，用户可以在命令行界面输入命令，实现对网络设备的配置和管理。

协议规范

命令行界面无对应的协议规范。

1.2 典型应用

典型应用	场景描述
通过CLI配置管理网络设备	通过在命令行界面输入命令对网络设备进行配置管理。

1.2.1 通过CLI配置管理网络设备

应用场景

以下图为例，用户通过终端登录网络设备 A，在命令行界面输入命令实现对设备的配置管理。

图 1-1

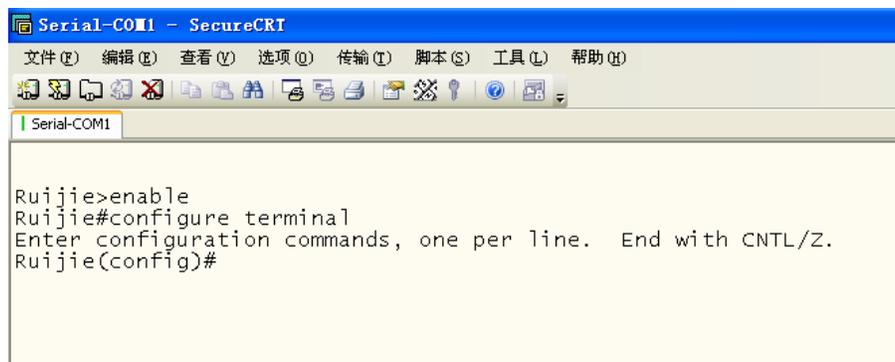


【注释】 A 为需要被管理的网络设备
PC 为用户端。

功能部署

下图列举了在 PC 上通过 Secure CRT 与网络设备 A 建立连接，并打开命令行界面配置命令。

图 1-2



```

Serial-COM1 - SecureCRT
文件(F) 编辑(E) 查看(V) 选项(O) 传输(T) 脚本(S) 工具(L) 帮助(H)
Serial-COM1

Ruijie>enable
Ruijie#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Ruijie(config)#

```

1.3 功能详解

功能特性

功能特性	作用
访问CLI	登录网络设备进行配置管理。
命令模式	命令行接口分为若干种命令模式，不同的命令模式可使用的命令不同。
系统帮助	用户在 CLI 配置过程中可获取系统的帮助信息。
缩写命令	如果输入的字符足够识别唯一的命令关键字，可以不必完整输入。
命令的no和default选项	通过 no 或 default 命令，禁止某个功能特性、执行与命令本身相反的操作或恢复缺省配置。
错误命令的提示信息	当用户输入错误命令时，会弹出相应的错误提示信息。
历史命令	用户可以通过快捷键的方式查询、调用历史命令。
编辑特性	系统提供相关快捷键便于用户编辑命令。
show命令的查找和过滤	用户可以在 show 命令输出的信息中查找或过滤指定的内容。
命令别名	配置命令的别名，可以替代命令执行配置。

1.3.1 访问CLI

在使用 CLI 之前，用户需要通过一个终端或 PC 和网络设备连接。启动网络设备，在网络设备硬件和软件初始化后就可以使用 CLI。在首次使用网络设备时，只能通过串口（Console）连接网络设备，称为带外（Out band）管理方式。在进行了相关配置后，还可以通过 Telnet 虚拟终端方式连接和管理网络设备。

1.3.2 命令模式

设备可供使用的命令非常多，为便于使用这些命令，将命令按功能进行分类。命令行接口分为若干个命令模式，所有命令都注册在某种（或几种）命令模式下。当使用某条命令时，需要先进入这个命令所在的模式。不同的命令模式之间既有联系又有区别。

当用户和网络设备管理界面建立一个新的会话连接时，用户首先处于用户模式（User EXEC 模式）。在此模式下，只可以使用少量命令，并且命令的功能也受到一些限制，例如像 show 命令等。用户模式的命令的操作结果不会被保存。

要使用更多的命令，首先须进入特权模式（Privileged EXEC 模式）。通常，在进入特权模式时必须输入特权模式的口令。在特权模式下，用户可以使用所有的特权命令，并且能够由此进入全局配置模式。

使用配置模式（全局配置模式、接口配置模式等）的命令，会对当前运行的配置产生影响。如果用户保存了配置信息，这些命令将被保存下来，并在系统重新启动时再次执行。要进入各种配置模式，首先必须进入全局配置模式。在全局配置模式下配置，可以进入接口配置模式等各种配置子模式。

各个命令模式概要如下（假定网络设备的名字为缺省的“Ruijie”）：

命令模式	访问方法	提示符	离开或访问下一模式	关于该模式
User EXEC (用户模式)	访问网络设备时默认进入该模式。	Ruijie>	输入 exit 命令离开该模式。 要进入特权模式，输入 enable 命令。	使用该模式来进行基本测试、显示系统信息。
Privileged EXEC (特权模式)	在用户模式下，使用 enable 命令进入该模式。	Ruijie#	要返回到用户模式，输入 disable 命令。 要进入全局配置模式，输入 configure 命令。	使用该模式来验证设置命令的结果。该模式是具有口令保护的。
Global configuration (全局配置模式)	在特权模式下，使用 configure 命令进入该模式。	Ruijie(config)#	要返回到特权模式，输入 exit 命令或 end 命令，或者键入 Ctrl+C 组合键。 要进入接口配置模式，输入 interface 命令。在 interface 命令中必须指明要进入哪一个接口配置子模式。 要进入 VLAN 配置模式，输入 vlan vlan_id 命令。	使用该模式的命令来配置影响整个网络设备的全局参数。
Interface configuration (接口配置模式)	在全局配置模式下，使用 interface 命令进入该模式。	Ruijie(config-if)#	要返回到特权模式，输入 end 命令，或键入 Ctrl+C 组合键。 要返回到全局配置模式，输入 exit 命令。在 interface 命令中必须指明要进入哪一个接口配置子模式。	使用该模式配置网络设备的各种接口。
Config-vlan (VLAN 配置模式)	在全局配置模式下，使用 vlan vlan_id 命令进入该模式。	Ruijie(config-vlan)#	要返回到特权模式，输入 end 命令，或键入 Ctrl+C 组合键。 要返回到全局配置模式，输入 exit 命令。	使用该模式配置 VLAN 参数。

1.3.3 系统帮助

用户在输入命令行的过程中，可以通过如下方式获取系统帮助。

1. 在任意模式的命令提示符下，输入问号（?）列出当前命令模式支持的命令及其描述信息。

例如：

```
Ruijie>?  
Exec commands:  
<1-99>      Session number to resume  
disable     Turn off privileged commands  
disconnect  Disconnect an existing network connection  
enable      Turn on privileged commands  
exit        Exit from the EXEC  
help        Description of the interactive help system  
lock        Lock the terminal  
ping        Send echo messages  
show        Show running system information  
telnet      Open a telnet connection  
traceroute  Trace route to destination
```

2. 在一条命令的关键字后空格并输入问号（?），可以列出该关键字关联的下一个关键字或变量。

例如：

```
Ruijie(config)#interface ?  
Aggregateport  Aggregate port interface  
Dialer         Dialer interface  
GigabitEthernet Gigabit Ethernet interface  
Loopback       Loopback interface  
Multilink      Multilink-group interface  
Null           Null interface  
Tunnel         Tunnel interface  
Virtual-ppp    Virtual PPP interface  
Virtual-template Virtual Template interface  
Vlan           Vlan interface  
range         Interface range command
```



如果该关键字后带的是一个参数值，则列出该参数的取值范围及其描述信息，如下所示：

```
Ruijie(config)#interface vlan ?  
<1-4094> Vlan port number
```

3. 在输入不完整的命令关键字后输入问号（?），可以列出以该字符串开头的所有命令关键字。

例如：

```
Ruijie#d?  
debug delete diagnostic dir disable disconnect
```

4. 在输入不完整的命令关键字后，如果该关键字后缀唯一，可以键入<Tab>键生成完整关键字。

例如：

```
Ruijie# show conf<Tab>
```

```
Ruijie# show configuration
```

5. 在任何命令模式下，还可以通过 **help** 命令获取帮助系统的摘要描述信息。

例如：

```
Ruijie(config)#help
Help may be requested at any point in a command by entering
a question mark '?'. If nothing matches, the help list will
be empty and you must backup until entering a '?' shows the
available options.
Two styles of help are provided:
1. Full help is available when you are ready to enter a
   command argument (e.g. 'show ?') and describes each possible
   argument.
2. Partial help is provided when an abbreviated argument is entered
   and you want to know what arguments match the input
   (e.g. 'show pr?'.)
```

1.3.4 简写命令

如果命令比较长，想简写命令，只需要输入命令关键字的一部分字符，且这部分字符足够识别唯一的命令关键字即可。

例如进入 GigabitEthernet 0/1 接口配置模式的命令 “**interface gigabitEthernet 0/1**” 可以简写成：

```
Ruijie(config)#int g0/1
Ruijie(config-if-GigabitEthernet 0/1)#
```

1.3.5 命令的no和default选项

大部分命令有 **no** 选项。通常，使用 **no** 选项来禁止某个特性或功能，或者执行与命令本身相反的操作。例如接口配置命令 **no shutdown** 执行关闭接口命令 **shutdown** 的相反操作，即打开接口。使用不带 **no** 选项的关键字，打开被关闭的特性或者打开缺省是关闭的特性。

配置命令大多有 **default** 选项，命令的 **default** 选项将命令的设置恢复为缺省值。大多数命令的缺省值是禁止该功能，因此在许多情况下 **default** 选项的作用和 **no** 选项是相同的。然而部分命令的缺省值是允许该功能，在这种情况下，**default** 选项和 **no** 选项的作用是相反的。这时 **default** 选项打开该命令的功能，并将变量设置为缺省的允许状态。

 各命令的 **no** 或 **default** 选项作用请参见相应的命令手册。

1.3.6 错误命令的提示信息

当用户输入错误命令时，会弹出相应的错误提示信息。

常见的 CLI 错误信息：

错误信息	含义	如何获取帮助
% Ambiguous command: "show c"	用户没有输入足够的字符，网络设备无法识别唯一的命令。	重新输入命令，紧接着发生歧义的单词输入一个问号。可能输入的关键字将被显示出来。
% Incomplete command.	用户没有输入该命令的必需的关键字或者变量参数。	重新输入命令，输入空格再输入一个问号。可能输入的关键字或者变量参数将被显示出来。
% Invalid input detected at '^' marker.	用户输入命令错误，符号 (^) 指明了产生错误的单词的位置。	在所在地命令模式提示符下输入一个问号，该模式允许的命令的关键字将被显示出来。

1.3.7 历史命令

系统能够自动保存用户最近输入的历史命令，用户可以通过快捷键的方式查询、调用历史命令。

操作方法如下：

操作	结果
Ctrl-P 或上方向键	在历史命令表中浏览前一条命令。从最近的一条记录开始，重复使用该操作可以查询更早的记录。
Ctrl-N 或下方向键	在使用了 Ctrl-P 或上方向键操作之后，使用该操作在历史命令表中回到更近的一条命令。重复使用该操作可以查询更近的记录。

1.3.8 编辑特性

用户在进行命令行编辑时，可以使用如下按键或快捷键：

功能	按键、快捷键	说明
在编辑行内移动光标。	左方向键或 Ctrl-B	光标移到左边一个字符。
	右方向键或 Ctrl-F	光标移到右边一个字符。
	Ctrl-A	光标移到命令行的首部。
	Ctrl-E	光标移到命令行的尾部。
删除输入的字符。	Backspace 键	删除光标左边的一个字符。
	Delete 键	删除光标右边的一个字符。
输出时屏幕滚动一行或一页。	Return 键	在显示内容时用回车键将输出的内容向上滚动一行，显示下一行的内容，仅在输出内容未结束时使用。
	Space 键	在显示内容时用空格键将输出的内容向上滚动一页，显示下一页内容，仅在输出内容未结束时使用。

当编辑的光标接近右边界时，命令行会整体向左移动 20 个字符，命令行前部被隐藏的部分被符号 (\$) 代替，可以使用相关按键或快捷键将光标移到前面的字符或者回到命令行的首部。

例如配置模式的命令 **access-list** 的输入可能超过一个屏幕的宽度。当光标第一次接近行尾时，命令行整体向左移动 20 个字符，命令行前部被隐藏的部分被符号 (\$) 代替。每次接近右边界时都会向左移动 20 个字符长度。

```
access-list 199 permit ip host 192.168.180.220 host
$ost 192.168.180.220 host 202.101.99.12
$0.220 host 202.101.99.12 time-range tr
```

可以使用 Ctrl-A 快捷键回到命令行的首部，这时命令行尾部被隐藏的部分将被符号 (\$) 代替：

```
access-list 199 permit ip host 192.168.180.220 host 202.101.99.$
```

i 默认的终端行宽是 80 个字符。

1.3.9 show 命令的查找和过滤

要在 **show** 命令输出的信息中查找指定的内容，可以在使用以下命令：

命令	作用
show any-command begin regular-expression	在 show 命令的输出内容中查找指定的内容，将第一个包含该内容的行以及该行以后的全部信息输出。

i 支持在任意模式下执行 **show** 命令。

i 查找的信息内容需要区分大小写，以下相同。

要在 **show** 命令的输出信息中过滤指定的内容，可以使用以下命令：

命令	作用
show any-command exclude regular-expression	在 show 命令的输出内容中进行过滤，除了包含指定内容的行以外，输出其他的信息内容。
show any-command include regular-expression	在 show 命令的输出内容中进行过滤，仅输出包含指定内容的行，其他信息将被过滤。

要在 **show** 命令的输出内容中进行查找和过滤，需要输入管道符号 (竖线, "|")。在管道字符之后，可以选择查找和过滤的规则和查找和过滤的内容 (字符或字符串)，并且查找和过滤的内容需要区分大小写：

```
Ruijie#show running-config | include interface
interface GigabitEthernet 0/0
interface GigabitEthernet 0/1
interface GigabitEthernet 0/2
interface GigabitEthernet 0/3
interface GigabitEthernet 0/4
interface GigabitEthernet 0/5
interface GigabitEthernet 0/6
interface GigabitEthernet 0/7
interface Mgmt 0
```

1.3.10 命令别名

用户可以指定任意单词作为命令的别名，来简化命令行字符串的输入。

配置效果

1. 一个单词代替一条命令。

例如：将 “**ip route 0.0.0.0 0.0.0.0 192.1.1.1**” 配置别名 “mygateway”，执行该命令只要输入 “mygateway” 即可。

2. 一个单词代替一条命令的前半部分，再输入后半部分。

例如：将 “**ip address**” 配置别名 “ia”，执行 IP 地址配置可以先输入 “ia”，再输入指定的 IP 地址及掩码。

配置方法

系统默认别名

- 在普通或特权用户模式下，部分命令存在默认的别名，可以通过 **show aliases** 命令查看：

```
Ruijie(config)#show aliases
Exec mode alias:
  h                help
  p                ping
  s                show
  u                undebug
  un               undebug
```

i 这些默认的别名不能删除。

配置命令别名

- 相关命令如下：

【命令格式】 **alias mode command-alias original-command**

【参数说明】 **mode**：别名所代表的命令所处的命令模式。

command-alias：命令别名。

original-command：别名所代表的实际命令。

【命令模式】 全局模式

【使用指导】 在全局配置模式下，输入 **alias ?**可以列出当前可以配置别名的全部命令模式。

查看命令别名设置

使用 **show aliases** 命令可以查看系统中的别名设置。

注意事项

- 别名替代的命令必须是命令行的第一个字符开始。
- 别名替代的命令必须是一个完整的输入形式。
- 命令别名在使用时必须完整输入，否则不能被识别。

配置举例

📌 定义一个别名替代整条命令

【配置方法】 在全局配置模式下，配置命令别名“ir”代表默认路由设置“**ip route 0.0.0.0 0.0.0.0 192.168.1.1**”

```
Ruijie#configure terminal
Ruijie(config)#alias config ir ip route 0.0.0.0 0.0.0.0 192.168.1.1
```

【检验方法】 ● 通过 **show alias** 查看别名是否设置成功。

```
Ruijie(config)#show alias
Exec mode alias:
  h             help
  p             ping
  s             show
  u             undebug
  un            undebug
Global configuration mode alias:
  ir            ip route 0.0.0.0 0.0.0.0 192.168.1.1
```

● 使用设置好别名执行命令，通过 **show running-config** 查看是否配置成功。

```
Ruijie(config)#ir
Ruijie(config)#show running-config

Building configuration...
!
alias config ir ip route 0.0.0.0 0.0.0.0 192.168.1.1 //配置别名
...
ip route 0.0.0.0 0.0.0.0 192.168.1.1 //输入别名“ir”的配置结果
!
```

📌 定义一个别名替代一个命令的前半部分

【配置方法】 在全局配置模式下，配置命令别名“ir”代表默认路由设置的“**ip route**”

```
Ruijie#configure terminal
Ruijie(config)#alias config ir ip route
```

【检验方法】 ● 通过 **show alias** 查看别名是否设置成功。

```
Ruijie(config)#show alias
Exec mode alias:
  h             help
  p             ping
  s             show
  u             undebug
```

```

un                undebug
Global configuration mode alias:
ir                ip route

```

- 输入别名“ir”，再配置后半部分命令“0.0.0.0 0.0.0.0 192.168.1.1”。
- 通过 **show running-config** 查看是否配置成功。

```

Ruijie(config)#ir 0.0.0.0 0.0.0.0 192.168.1.1
Ruijie(config)#show running

Building configuration...
!
alias config ir ip route //配置别名
!
ip route 0.0.0.0 0.0.0.0 192.168.1.1 //输入别名“ir”及后半部分命令的配置结果
!

```

命令别名支持的系统帮助

1. 命令别名支持帮助信息，在别名前面会显示一个星号(*)，格式如下：

```
*command-alias=original-command
```

例如，在 EXEC 模式下，默认的命令别名“s”表示“show”关键字。输入“s?”，可以获取's'开头的关键字和别名的帮助信息：

```

Ruijie#s?
*s=show show start-chat start-terminal-service

```

2. 如果别名所代表的命令不止一个单词，在帮助信息中将携带引号显示。

例如，在 EXEC 模式下配置别名“sv”代替命令 **show version**，输入“s?”，可以获取's'开头的关键字和别名的帮助信息：

```

Ruijie#s?
*s=show *sv="show version" show start-chat
start-terminal-service

```

3. 获取系统帮助时，命令别名可以获取与该命令相关的帮助信息。

例如，配置接口模式下的命令别名“ia”代表“ip address”，在接口模式下输入“ia?”，可获取等同“ip address?”的帮助信息，并且将别名替换成实际的命令：

```

Ruijie(config-if)#ia ?
A.B.C.D IP address
dhcp IP Address via DHCP
Ruijie(config-if)#ip address

```

! 如果在命令之前输入了空格，将无法获取该别名表示的命令

2 基础管理

2.1 概述

基础管理为首次接触网络设备管理的入门手册，介绍一些常用的网络设备管理、监控和维护的功能。

协议规范

无

2.2 典型应用

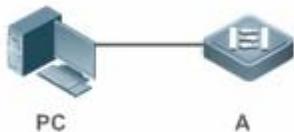
典型应用	场景描述
网络设备的管理	用户通过终端登录网络设备，在命令行界面输入命令实现对设备的配置管理。

2.2.1 网络设备的管理

应用场景

在本文档中，所涉及的管理都是通过命令行界面进行的，用户通过终端登录网络设备 A，在命令行界面输入命令实现对设备的配置管理。如下图所示：

图 2-1



2.3 功能详解

基本概念

▾ TFTP

TFTP (Trivial File Transfer Protocol,简单 文件传输协议) 是TCP/IP协议族中的一个用于客户机与 服务器之间进行简单文件传输的协议。

AAA

AAA (Authentication Authorization Accounting , 认证授权计帐)。

Authentication认证：验证用户的身份与可使用的 网络服务。

Authorization 授权：依据认证结果开放网络服务给用户。

Accounting计帐：记录用户对各种网络服务的用量，并提供给 计费系统。整个系统在 网络管理与安全问题中十分有效。

RADIUS

RADIUS (Remote Authentication Dial In User Service , 远程用户拨号认证系统) 是目前应用最广泛的 AAA协议。

Telnet

Telnet是位于OSI模型的第7层---应用层上的一种协议，是一个通过创建 虚拟终端提供连接到远程 主机 终端仿真的TCP/IP协议。这一协议需要通过用户名和口令进行认证，是Internet远程登陆服务的标准协议。应用Telnet协议能够把 本地用户所使用的计算机变成远程 主机系统的一个 终端。

系统信息

系统信息主要包括系统描述，系统上电时间，系统的硬件版本，系统的软件版本，系统的 Ctrl 层软件版本，系统的 Boot 层软件版本。

硬件信息

硬件信息主要包括物理设备信息及设备上的插槽和模块信息。设备本身信息包括：设备的描述，设备拥有的插槽的数量。插槽信息：插槽在设备上的编号，插槽上模块的描述（如果插槽没有插模块，则描述为空），插槽所插入模块包括物理端口数，插槽最多可能包含的端口的最大个数（所插模块包括的端口数）。

功能特性

功能特性	作用
控制用户访问	通过使用口令保护和划分特权级别来控制网络上的终端访问网络设备。
控制登录认证	启用 AAA 的模式下，用户登录网络设备进行管理的时候可以通过一些服务器来根据用户名和密码进行用户的管理权限的认证。
系统基本参数	系统的各项参数，例如时钟，标题，控制台速率等。
查看配置信息	查看系统配置信息主要包括查看系统正在运行的配置信息，以及查看存储在 NVRAM (非易失性随机存取存储器) 上设备的配置等。
配置回滚	配置回滚功能允许用户将当前的配置制作一个检查点 (checkpoint)，并且在不需要重启设备的情况下，将检查点的配置重新应用到设备上。
使用Telnet	Telnet 属于 TCP/IP 协议族的应用层协议，它给出通过网络提供远程登录和虚拟终端通讯功能的规范。
重启	介绍系统重启。

2.3.1 控制用户访问

通过使用口令保护和划分特权级别来控制网络上的终端访问网络设备。

工作原理

▾ 授权级别

网络设备的命令行界面针对用户划分 0-15 共 16 个授权级别，不同级别的用户可以执行的命令是不同的。数字小的级别权限较小，其中 0 级为最低级别，只能执行少数几条命令；15 级为最高级别，可以执行所有的命令。0-1 级一般称为普通用户级别，不允许对设备进行配置（默认不允许进入全局配置模式），2-15 级一般称为特权用户级别，可以对设备进行配置。

▾ 口令类别

口令分为 password 和 security 两种。password 为简单加密的口令，只能设置为 15 级口令。security 口令为安全加密口令，可以为 0~15 级设置口令。如果系统中同级别同时存在以上两种口令，则 password 口令不生效。如果设置非 15 级的 password 口令，则会给出警告提示，并自动转为 security 口令；如果设置 15 级的 password 口令和 security 口令完全相同，则会给出警告提示；口令必须以加密形式保存，password 口令使用简单加密，security 口令使用安全加密。

▾ 口令保护

在网络设备上为每个特权级别设置口令，当用户想升高权限级别时，需要输入目的级别对应的口令，口令校验通过以后才允许升高权限级别。用户降低级别则不需要通过口令校验。

缺省时系统只有两个受口令保护的授权级别：普通用户级别（1 级）和特权用户级别（15 级）。但是用户可以为每个模式的命令划分 16 个授权级别。通过给不同的级别设置口令，就可以通过不同的授权级别使用不同的命令集合。

在特权用户级别口令没有设置的情况下，进入特权级别亦不需要口令校验。为了安全起见，我们提醒您最好为特权用户级别设置口令。

▾ 命令授权

每一条命令都有最低执行级别的要求，如果用户的权限级别达不到要求是无法执行该命令的。此时可以通过命令授权操作，将命令执行权限授予某个特权级别，将允许权限达到（大于或等于）该级别的用户执行该命令。

相关配置

▾ 设置 password 口令

- 使用 **enable password** 命令设置 password 口令。

▾ 设置 secret 口令

- 使用 **enable secret** 命令设置安全口令。
- 需要在切换用户级别时进行 secret 口令校验，可以配置此项。功能与 password 口令相同，但使用了更好的口令加密算法。为了安全起见，建议使用 secret 口令。

📌 设置命令的级别

- 使用 **privilege** 命令设置命令的级别。
- 如果想让更多的授权级别使用某一条命令，则可以将该命令设置较低的用户级别；而如果想让命令的使用范围小一些，则可以将该命令设置较高的用户级别。
- 使用 **login privilege log** 命令打开级别提升的日志信息提示功能。

📌 升高/降低用户级别

- 使用 **enable / disable** 命令升高/降低用户级别。
- 已经登录网络设备的用户，可以通过改变当前的用户级别，以访问不同级别的命令。

📌 启用 line 线路口令保护

- 对远程登录（如 TELNET）进行口令验证，要配置 **line** 口令保护。
- 应先使用 **password[0 | 7] line** 命令配置 **line** 线路口令，然后执行 **login** 命令启动口令保护。
- 终端在缺省情况下不支持 **lock** 命令。

2.3.2 控制登录认证

在未启用 AAA 模式下，用户登录网络设备进行管理的时候，如果线路上设置了登陆认证（login），需要通过线路上所配置的口令进行校验，通过校验的用户才允许登录。如果线路上设置了本地认证（login local），则需要通过本地用户数据库来根据用户名和密码进行用户的管理权限的认证。

在启用 AAA 模式下，用户登录网络设备进行管理的时候，可以利用一些服务器根据用户名和密码进行用户的管理权限的认证，通过认证的用户才允许登录。

例如，利用 RADIUS 服务器，根据用户登录时的用户名和密码，控制用户对网络设备的管理权限。通过这种方式，网络设备不再用本地保存的密码信息进行认证，而是将加密后的用户信息发送到 RADIUS 服务器上验证。服务器统一配置用户的用户名、用户密码、共享密码和访问策略等信息，便于管理和控制用户访问，提高用户信息的安全性。

工作原理

📌 线路口令

配置线路（line）口令的目的，是为了在未启用 AAA 模式的情况下，用于终端登录时的口令校验。启用了 AAA 模式以后，线路上的口令校验将不生效。

📌 本地认证

配置本地认证的目的，是为了在未启用 AAA 模式的情况下，通过本地用户数据库来根据用户名和密码进行用户的管理权限的认证。启用了 AAA 模式以后，线路上的本地认证设置将不生效。

📌 AAA 模式

AAA 是认证、授权和记账 (Authentication, Authorization and Accounting) 的简称, AAA 是一种体系结构框架, 它提供包括认证、授权和记账在内三个互相独立的安全功能。启用了 AAA 模式以后, 终端登录时候需要根据 AAA 所设置的登录认证方法列表的要求, 通过一些服务器 (或本地用户数据库) 来根据用户名和密码进行用户的管理权限的认证。AAA 功能详解参见 AAA 配置指南。

相关配置

配置本地用户

- 使用 **username** 命令配置用于本地身份认证和授权的账号信息, 包括用户名、密码以及可选的授权信息。

线路登录进行本地认证

- 使用 **login local** 命令在 AAA 关闭时, LINE 线路登录认证时走本地用户认证。
- 应在每台设备上配置。

线路登录进行 AAA 认证

- AAA 打开的情况下, 默认使用 **default** 认证方法。
- 使用 **login authentication** 命令在 LINE 线路上配置登录认证方法列表。
- AAA 设置为采用本地认证方法时需要配置。

设置连接超时时间

- 缺省的超时时间为 10 分钟。
- 使用 **exec-timeout** 命令设置连接超时时间。当前已接受的连接, 在指定时间内, 没有任何输入时, 将中断此连接。
- 在需要延长或缩短这段等待时间时, 应执行此配置项。

设置会话超时时间

- 缺省的超时时间为 0 min, 代表永不超时。
- 使用 **session-timeout** 命令设置会话超时时间。
- 当前 LINE 上已经建立的会话, 在指定时间内, 没有任何输入信息, 将中断当前连接到远程终端的会话。并且恢复终端为空闲状态。在需要延长或缩短这段等待时间时, 应执行此配置项。

会话锁定

- 终端在缺省情况下不支持 **lock** 命令。
- 使用 **lockable** 命令允许用户锁住当前线路所连接的终端。
- 要使用会话锁定功能, 需要在 line 配置模式下启用锁住 line 终端的功能, 并在相应终端的 EXEC 模式下, 通过使用 **lock** 命令锁住终端。

2.3.3 系统基本参数

系统时间

网络设备的系统时钟主要用于系统日志等需要记录事件发生时间的地方。该时钟提供具体日期(年、月、日)和时间(时、分、秒)以及星期等信息。

对于一台网络设备，当第一次使用时你需要首先手工配置网络设备系统时钟为当前的日期和时间。

配置系统名称和命令提示符

为了管理的方便，可以为一台网络设备配置系统名称(System Name)来标识它。默认系统名为“Ruijie”，如果系统名称超过32个字符，则截取其前32个字符。默认情况下，系统名称作为默认的命令提示符，提示符将随着系统名称的变化而变化。

标题

标题可以提供一些常规的登录提示信息。可以创建的标题（banner）类型有两种：每日通知和登录标题。

- 每日通知针对所有连接到网络设备的用户，当用户登录网络设备时，通知消息将首先显示在终端上。利用每日通知，你可以发送一些较为紧迫的消息（比如系统即将关闭等）给用户。
- 登录标题显示在每日通知之后，它的主要作用是提供一些常规的登录提示信息。

配置控制台速率

通过配置控制台接口可以对网络设备进行管理。当网络设备第一次使用的时候，必须采用通过控制台口方式对其进行配置。使用时可以根据实际需求，改变网络设备串口的速率。需要注意的是，用来管理网络设备的终端的速率设置必须和网络设备的控制台的速率一致。

设置连接超时

配置设备的连接超时时间，控制该设备已经建立的连接（包括已接受连接，以及该设备到远程终端的会话）。当空闲时间超过设置值，没有任何输入输出信息时，中断此连接。

相关配置

设置系统的日期和时钟

- 使用 **clock set** 命令通过手工的方式来设置网络设备上的时间。当你设置了网络设备的时钟后，网络设备的时钟将以你设置的时间为准一直运行下去，即使网络设备下电，网络设备的时钟仍然继续运行。

更新硬件时钟

- 如果硬件时钟和软件时钟不同步，使用 **clock update-calendar** 命令可以通过软件时钟的日期和时间复制给硬件时钟。

设置系统名称

- 使用 **hostname** 命令可以修改默认的系统名称。
- 缺省的主机名为 Ruijie。

↘ 设置命令提示符

- 通过 **prompt** 命令可以设置用户命令接口的提示符。

↘ 设置每日通知

- 缺省没有每日通知。
- 使用 **banner motd** 命令配置每日通知信息。
- 每日通知针对所有连接到网络设备的用户，当用户登录网络设备时，通知消息将首先显示在终端上。利用每日通知，你可以发送一些较为紧迫的消息（比如系统即将关闭等）给用户。

↘ 配置登录标题

- 缺省没有登录标题。
- 使用 **banner login** 命令设置登录标题，用于提供一些常规的登录提示信息。

↘ 设置控制台的传输速率

- 使用 **speed** 命令配置终端设备的速率。
- 缺省的速率是 9600。

2.3.4 查看配置信息

查看系统配置信息主要包括查看系统正在运行的配置信息，以及查看存储在 NVRAM（非易失性随机存取存储器）上设备的配置等。

工作原理

↘ 系统正在运行的配置信息

系统正在运行的配置信息即 **running-config** 是系统上所有的组件模块当前运行的配置的总和，具有实时性的特点。在查看的时候，先需要向所有的运行组件请求搜集配置，并经过一定的编排组合后显示给用户。正因为实时性的特点，只有运行中的组件才可能提供此配置信息，如果组件未加载则不会显示其配置。这样，在系统启动、组件进程重启、以及运行热补丁的过程中，组件处于不稳定状态情况下所收集的系统运行配置会有一些的差异。例如在某一时段收集的信息中缺乏某个组件的配置，过一段时间后再收集就有了。

↘ 系统的启动配置信息

存储在 NVRAM（非易失性随机存取存储器）上设备的配置即 **startup-config** 为设备启动时执行的配置，在系统重新启动后将会导入 **startup-config** 成为新的运行配置。查看永久配置的过程就是读取设备 NVRAM 上的 **startup-config** 文件信息并显示。

相关配置

↘ 查看系统正在运行的配置信息

执行 **show running-config [interface interface]** 命令查看系统正在运行的配置信息或某个接口下的配置信息。

查看设备的启动配置信息

执行 **show startup-config** 命令查看设备的启动配置信息。

保存设备的启动配置信息

执行 **write** 命令或者 **copy running-config startup-config** 将设备的当前正在运行的配置信息，保存成为新的启动配置信息。

2.3.5 配置回滚

配置回滚功能允许用户将当前的配置制作一个快照，或称为检查点（checkpoint），并且在不需要重启设备的情况下，将检查点的配置重新应用到设备上。

工作原理

任何授权的用户，可以在任何时候创建一个检查点，就是当前运行配置的一个副本。在系统中，检查点将被保存为一个文本文件，用户可以在将来某个时候，用这个文件来恢复创建检查点时的运行配置。RGOS 系统支持用户创建多个检查点，用来保存不同版本的运行配置。

 在设备上，同时只能有一个用户在进行创建检查点、回滚配置的操作。

相关配置

创建检查点

执行 **checkpoint [cp-name] [description description]** 创建一个检查点，可以指定检测点名称、描述。

查看检查点信息

执行 **show checkpoint { cp-name [all] | summary }** 命令查看单个或全部检查点的信息，以及显示所有检查点的概要信息。

回滚配置

执行 **rollback running-config checkpoint cp-name** 命令回滚一个指定的检查点。

清除检查点数据

执行 **clear checkpoint database** 命令删除系统中所有的检查点数据，以及删除检查点对应的配置文件。

2.3.6 使用Telnet

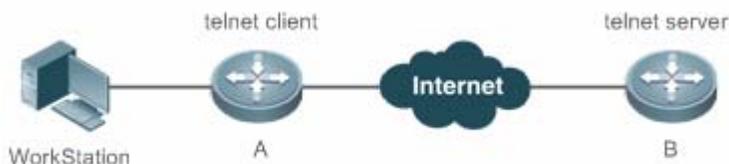
工作原理

Telnet 属于 TCP/IP 协议族的应用层协议，它给出通过网络提供远程登录和虚拟终端通讯功能的规范。

Telnet Client 服务为已登录到本网络设备上的本地用户或远程用户提供使用本网络设备的 Telnet Client 程序访问网上其他远程系统资源的服务。如下图所示用户在微机上通过终端仿真程序或 Telnet 程序建立与网络设备 A 的连接后，可通过输入 telnet 命令再登录设备 B，并对其进行配置管理。

锐捷网络的 Telnet 程序同时支持使用 IPV4 地址进行通讯。作为 Telnet Server，可以同时接受 IPV4 的 Telnet 连接请求。作为 Telnet Client，可以向 IPV4 地址的主机发起连接请求。

图 2-1



相关配置

使用 telnet client

- 使用 **telnet** 命令通过 telnet 登录到远程设备。

恢复已建立的 Telnet Client 会话连接

- 执行 **<1-99>** 命令恢复已建立的 Telnet Client 会话连接。

断开挂起的 Telnet Client 连接

- 执行 **disconnect session-id** 命令断开指定的 Telnet Client 连接。

使用 Telnet Server

- 使用 **enable service telnet-server** 命令打开 Telnet Server 服务。
- 需要使用 Telnet 登录本地设备时，需要打开该服务。

2.3.7 重启

定时重启功能，它在某些场合下(比如出于测试目的或其它需要)可以为用户提供操作上的便利。

- 指定系统在经过一定时间间隔后重启。这里的时间间隔由 *mmm* 或 *hhh:mm* 决定，以分钟为单位，用户可以任选一种格式输入。用户可以在这里为这个计划起一个助记名，以便能直观地反映该重启的用途。
- 指定系统在将来的某个时间点重启。输入的时间值必须是将来的某个时间点。

! 如果用户要使用 **at** 选项，则要求当前系统必须支持时钟功能。建议使用之前先配置好系统的时钟，以便更切合您的用途。如果用户之前已经设置了重启计划，则后面再设置的计划将覆盖前面的设置。如果用户已经设置了重启计划，假如在该计划生效前用户重启了系统，则该计划将丢失。

 重启计划中的时间与当前时间的跨度不能超过 31 天并且要大于当前系统时间。同时用户在设置了重启计划之后最好不要再修改系统时钟,否则有可能会造成设置失效,比如将系统时间调到重启时间之后。

相关配置

设置重启

- 使用 `reload` 命令设置重启策略。
- 使用该命令可以指定设备在指定的时刻启动,方便进行管理。

2.4 配置详解

配置口令与权限	 可选配置。设置口令与命令级别划分。	
	<code>enable password</code>	设置 password 口令
	<code>enable secret</code>	设置 secret 口令
	<code>enable</code>	升高用户级别
	<code>login privilege log</code>	在终端用户权限级别提升的情况下,输出相应的日志信息
	<code>disable</code>	降低用户级别
	<code>privilege</code>	设置命令的级别划分
	<code>password</code>	指定 line 线路口令
	<code>login</code>	启用 line 线路口令保护
配置登录与认证	 可选配置。配置不同登录方式及认证。	
	<code>username</code>	配置本地用户账号以及可选的授权信息
	<code>login local</code>	线路登录进行本地认证
	<code>loginauthentication</code>	线路登录进行 AAA 认证
	<code>telnet</code>	使用 Telnet Client
	<code>enable service telnet-server</code>	使用 Telnet Server
	<code>exec-timeout</code>	配置连接超时时间
	<code>session-timeout</code>	配置会话超时时间
	<code>lockable</code>	启用锁住 line 终端的功能
<code>lock</code>	锁住当前 line 终端	
设置系统基本参数	 可选配置。设置系统基本参数。	
	<code>clock set</code>	设置系统的日期和时钟
	<code>clock update-calendar</code>	更新硬件时钟
	<code>hostname</code>	设置系统名称
	<code>prompt</code>	设置命令提示符

	banner motd	设置每日通知
	bannerlogin	配置登录标题
	speed	设置控制台的传输速率
打开或关闭指定的服务	 可选配置。打开与关闭指定的服务。	
	enable service	打开某项服务。
回滚系统配置	 可选配置。回滚系统配置。	
	checkpoint [<i>cp-name</i>] [description <i>description</i>]	创建一个配置检查点
	rollback running-config checkpoint <i>cp-name</i>	回滚配置
	clear checkpoint database	清除检查点数据
设置重启策略	 可选配置。设置系统重启时的策略。	
	reload	重启设备。

2.4.1 配置口令与权限

配置效果

- 设置用户的口令，可以控制对网络设备的访问。
- 对命令使用权限进行分级，对于特定级别的命令，只有达到或高于这个级别的用户才可以使用。
- 将命令的使用权授予较低的用户级别，让更多的授权级别使用该条命令。
- 该命令的使用权授予较高的用户级别，则该命令的使用范围会缩小。

注意事项

- 在设置口令中，如果使用带 **level** 关键字时，则为指定特权级别定义口令。设置了特定级别的口令后，给定的口令只适用于那些需要访问该级别的用户。
- 缺省没有设置任何级别的 **password** 或 **secret** 口令，如果没有指定 **level**，则缺省的级别是 15 级。
- 如果设置非 15 级的 **password** 口令，系统将自动转换为 **secret** 口令，并给出提示信息。
- 如果同时设置了 **password** 口令和 **secret** 口令，则系统将选择使用 **secret** 口令。

配置方法

📌 设置 password 口令

- 可选配置。需要在切换用户级别时进行 **password** 口令校验，可以配置此项。

- 使用 **enable password** 命令设置 password 口令。

📌 设置 secret 口令

- 可选配置。需要在切换用户级别时进行 secret 口令校验，可以配置此项。
- 使用 **enable secret** 命令设置安全口令。
- 功能与 password 口令相同，但使用了更好的口令加密算法。为了安全起见，建议使用 secret 口令。

📌 设置命令的级别

- 可选配置。
- 如果想让更多的授权级别使用某一条命令，则可以将该命令设置较低的用户级别；而如果想让命令的使用范围小一些，则可以将该命令设置较高的用户级别。

📌 升高/降低用户级别

- 已经登录网络设备的用户，可以通过改变当前的用户级别，以访问不同级别的命令。
- 使用 **enable / disable** 命令升高/降低用户级别。
- 使用 **login privilege log** 命令打开级别提升的日志信息提示功能。

📌 启用 line 线路口令保护

- 可选配置。对远程登录（如 TELNET）进行口令验证，要配置 line 口令保护。
- 应先使用 **password[0 | 7] line** 命令配置 line 线路口令，然后执行 **login** 命令启动口令保护。

⚠️ 如果没有配置登录认证，即使配置了 line 口令，登录时，也不会提示用户输入口令进行认证。

检验方法

- 可以使用 **show privilege** 命令查看当前用户级别。
- 可以使用 **show running-config** 命令查看配置。

相关命令

📌 设置 password 口令

【命令格式】 **enable password [level level] { password | [0 | 7] encrypted-password }**

【参数说明】 *level* : 用户的级别。

password : 用户进入特权 EXEC 配置层的口令。

0 : 表示输入的口令字符串为明文字符串。

7 : 表示输入的口令字符串为密文字符串。

encrypted-password : 口令文本。必须包含 1 到 26 个大小写字母和数字字符。

📘 口令前面可以有前导空格，但被忽略。中间及结尾的空格则作为口令的一部分。

- 【命令模式】 全局模式
- 【使用指导】 目前只能设置 15 级用户的口令，并且只能在未设置 security 口令的情况下有效。
如果设置非 15 级的口令，系统将会给出一个提示，并自动转为 security 口令。
如果设置的 15 级 password 口令和 15 级安全口令完全相同，系统将会给出一个警告信息。

 如果指定了加密类型，然后输入一条明文口令，则不能重新进入特权 EXEC 模式。不能恢复用任意方法加密的已丢失口令。只能重新配置设备口令。

设置 secret 口令

- 【命令格式】 **enable secret [level level] {secret | [0 | 5] encrypted-secret }**
- 【参数说明】 *level* : 用户的级别。
secret : 用户进入特权 EXEC 配置层的口令。
0 | 5 : 口令的加密类型，0 无加密，5 安全加密。
encrypted-password : 口令文本。
- 【命令模式】 全局配置模式
- 【使用指导】 配置不同权限级别的安全的口令。

升高用户级别

- 【命令格式】 **enable [privilege-level]**
- 【参数说明】 *privilege-level* : 权限等级。
- 【命令模式】 特权用户模式
- 【使用指导】 从权限较低的级别切换到权限较高的级别需要输入相应级别的口令。

降低用户级别

- 【命令格式】 **disable [privilege-level]**
- 【参数说明】 *privilege-level* : 权限等级
- 【命令模式】 特权用户模式
- 【使用指导】 从权限较高的级别切换到权限较低的级别需要输入相应级别的口令。

使用该命令从特权用户模式退到普通用户模式。如果加上权限等级，则将当前权限等级降低到指定的权限等级。

 **disable** 命令后面所跟权限等级必须小于当前权限等级。

提升用户权限级别的日志信息提示功能

- 【命令格式】 **login privilege log**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 如果需要监控终端用户权限级别提升，要配置这个功能，该功能对所有终端都有效。

设置命令的级别划分

- 【命令格式】 **privilege mode [all] {level level | reset} command-string**
- 【参数说明】 *mode* : 要授权的命令所属的 CLI 命令模式，例如 :config 表示全局配置模式，exec 表示特权命令模式，interface

表示接口配置模式等等。

all：将指定命令的所有子命令的权限，变为相同的权限级别。

level level：授权级别，范围从 0 到 15。

reset：将命令的执行权限恢复为默认级别。

command-string：要授权的命令。

【命令模式】 全局模式

【使用指导】 可以在全局配置模式下使用 **no privilege mode [all]level level command** 命令，恢复一条已知的命令授权。

指定 line 线路口令

【命令格式】 **password[0 | 7] line**

【参数说明】 **0**：以明文方式配置口令。

7：以密文方式配置口令。

line：配置的口令字符串。

【命令模式】 **line** 配置模式

【使用指导】 -

启用 line 线路口令保护

【命令格式】 **login**

【参数说明】 -

【配置模式】 **line** 配置模式

【使用指导】 -

配置举例

配置命令授权

【网络环境】 将 **reload** 命令及其子命令授予级别 1 并且设置级别 1 为有效级别（通过设置口令为“test”）。

【配置方法】 ● 将 **reload** 命令及其子命令授予级别 1

```
Ruijie# configure terminal
Ruijie(config)# privilege exec all level 1 reload
Ruijie(config)# enable secret level 1 0 test
Ruijie(config)# end
```

【检验方法】 ● 进入 1 级，查看 **reload** 命令及子命令是否存在。

```
Ruijie# disable 1
Ruijie> reload ?
at                reload at a specific time/date
cancel            cancel pending reload scheme
in                reload after a time interval
<cr>
```

常见错误

- 无

2.4.2 配置登录与认证

配置效果

- 建立线路登录身份认证。
- 通过网络设备上的 telnet 命令登录到远程设备上去。
- 当前已接受的连接，在指定时间内，没有任何输入信息，服务器端将中断此连接。
- 当前 LINE 上已经建立的会话，在指定时间内，没有任何输入信息，将中断当前连接到远程终端的会话。并且恢复终端为空闲状态。
- 使用锁住会话终端的功能，以防止访问。终端被锁定后，在终端下输入任何字符，系统都会提示输入解锁口令，口令认证成功后，系统自动解锁。

注意事项

- 无

配置方法

▾ 配置本地用户

- 必选配置。
- 使用 **username** 命令配置用于本地身份认证和授权的账号信息，包括用户名、密码以及可选的授权信息
- 应在每台设备上配置本地身份认证的账号信息
-

▾ 线路登录进行本地认证

- 必选配置。
- 在 AAA 关闭时，LINE 线路登录认证时走本地用户认证。
- 应在每台设备上配置。

▾ 线路登录进行 AAA 认证

- 可选配置。AAA 设置为采用本地认证方法时需要配置。
- AAA 认证模式打开时，设置线路登录进行 AAA 认证。

- 应在每台设备上配置。

↘ 使用 telnet client

- 通过 telnet 登录到远程设备。

↘ 恢复已建立的 Telnet Client 会话连接

- 可选配置。Telnet Client 会话连接暂时退出后，如果需要恢复该连接，可以使用本命令恢复。

↘ 断开挂起的 Telnet Client 连接

- 可选配置。如果需要断开指定的 Telnet Client 连接，可以在 Telnet Client 设备上执行该配置项。

↘ 使用 Telnet Server

- 可选配置。需要使用 Telnet 登录本地设备时，需要打开该服务。
- 打开 Telnet Server 服务。

↘ 设置连接超时时间

- 可选配置。
- 当前已接受的连接，在指定时间内，没有任何输入时，将中断此连接。
- 在需要延长或缩短这段等待时间时，应执行此配置项。

↘ 设置会话超时时间

- 可选配置。
- 当前 LINE 上已经建立的会话，在指定时间内，没有任何输入信息，将中断当前连接到远程终端的会话。并且恢复终端为空闲状态。
- 在需要延长或缩短这段等待时间时，应执行此配置项。

↘ 会话锁定

- 可选配置。在已建立会话后需要临时离开设备时，在设备上执行会话锁定功能。
- 要使用会话锁定功能，需要在 line 配置模式下启用锁住 line 终端的功能，并在相应终端的 EXEC 模式下，通过使用 lock 命令锁住终端。

检验方法

- 使用 **show running-config** 命令可以查看配置。
- 在 AAA 关闭时，配置了本地用户以后，并在线路上设置采用本地认证。用户登录时将提示输入用户名和口令，认证通过后才允许进入命令行界面。
- 在 AAA 打开时，配置了本地用户后，并在 AAA 的登录认证方法中指定采用本地方法。用户登录时将提示输入用户名和口令，认证通过后才允许进入命令行界面。

- 已经登录进入命令行界面的用户，可以使用 **show user** 命令查看当前登录的用户信息。
- 在本地设备上开启 Telnet Server 后，用户可以使用 Telnet 客户端连接本地设备。
- 用户在被锁住的界面上输入回车后，会提示输入口令，只有口令与之前所设置的相符，才会解锁这个终端会话。
- 使用 **show sessions** 命令，可以查看已经建立的 Telnet Client 实例的每个实例信息。

相关命令

▾ 配置本地用户

【命令格式】 **username name [login mode { console | ssh | telnet }] [online amount number] [permission oper-mode path] [privilege privilege-level] [reject remote-login] [web-auth] [pwd-modify] [nopassword | password [0 | 7] text-string]**

【参数说明】 *name*：用户名。

login mode：配置账号的登录方式限制。

console：限制账号的登录方式为 console。

ssh：限制账号的登录方式为 ssh。

telnet：限制账号的的登录方式为 telnet。

online amount number：配置账号的同时在线数量。

permission oper-mode path：配置账号对指定文件的操作权限，*op-mode* 表示操作模式，*path* 表示作用的文件或者目录的路径。

privilege privilege-level：配置账号的权限级别，取值范围 0 到 15。

reject remote-login：限制使用该账号进行远程登录。

web-auth：此账号只能用于 web 认证。

pwd-modify：允许使用该账号的 web 认证用户修改密码，该选项只有在配置了 **web-auth** 之后才可用。

nopassword：该账号不配置密码。

password [0 | 7] text-string：配置账号的密码，0 表示输入明文密码，7 表示输入密文密码，默认为输入明文密码。

【命令模式】 全局配置模式

【使用指导】 用于建立本地用户数据库，供认证使用。

如果指定加密类型为 7，则输入的合法密文长度必须为偶数。

通常无须指定加密类型为 7。一般情况下，只有当复制并粘贴已经加密过的口令时，才需要指定加密类型为 7。

▾ 线路登录进行本地认证

【命令格式】 **login local**

【参数说明】 -

【命令模式】 line 配置模式

【使用指导】 如果没有启用 AAA 安全服务，则该命令用于配置 LINE 线路登录认证时走本地用户认证。这里的本地用户是指通过 **username** 命令配置的用户信息。

▾ 线路登录进行 AAA 认证

【命令格式】 **Login authentication { default | list-name }**

- 【参数说明】 **default** : 默认的认证方法列表名。
list-name : 可选的方法列表名。
- 【配置模式】 line 配置模式
- 【使用指导】 AAA 认证模式打开时, 设置线路登录进行 AAA 认证。认证时使用 AAA 方法列表中的认证方法, 包括 Radius 认证、本地认证、无认证等。

▾ 使用 Telnet Client

- 【命令格式】 **telnet [oob] host [port] [/source { ip A.B.C.D | interface interface-name }] [via mgmt-name]**
- 【参数说明】 **Oob**: 通过带外通信 (一般指通过 MGMT 接口) 远程连接到 Telnet 服务器, 只有在设备具备 MGMT 管理口的时候才会有该选项。
Host : Telnet 服务器的 IPV4 地址或者主机名。
Port : Telnet 服务器的 TCP 端口号, 默认值为 23。
/source:指定 Telnet 客户端使用的源 IP 或者源接口。
ip A.B.C.D : 指定 Telnet 客户端使用的源 IPV4 地址。
interface interface-name : 指定 Telnet 客户端使用的源接口。
via mgmt-name : 指定 Telnet 客户端在 oob 选项时使用的 MGMT 口
- 【命令模式】 特权用户模式
- 【使用指导】 通过 telnet 登录到远程设备, 可以是 IPV4 主机名、IPV4 地址。

▾ 恢复已建立的 Telnet Client 会话连接

- 【命令格式】 <1-99>
- 【参数说明】 -
- 【命令模式】 普通用户模式
- 【使用指导】 该命令用于恢复使用已经建立的 Telnet Client 会话连接。当使用 **telnet** 命令发起 Telnet Client 会话连接时, 可以使用热键 (ctrl+shift+6 x) 暂时退出该连接。如果需要恢复该连接, 可以使用<1-99>命令进行恢复。同时, 如果连接已建立, 可以使用 **show sessions** 命令查看已建立的连接信息。

▾ 断开挂起的 Telnet Client 连接

- 【命令格式】 **disconnect session-id**
- 【参数说明】 *session-id* : 挂起的 Telnet Client 连接会话号。
- 【命令模式】 普通用户模式
- 【使用指导】 通过输入指定的 Telnet Client 连接会话号, 断开指定的 Telnet Client 连接。

▾ 使用 Telnet Server

- 【命令格式】 **enable service telnet-server**
- 【参数说明】 -
- 【配置模式】 全局模式
- 【使用指导】 打开 Telnet Server 服务;

▾ 配置连接超时时间

- 【命令格式】 **exec-timeout minutes [seconds]**

- 【参数说明】 *minutes* : 指定的超时时间的分钟数。
seconds : 指定的超时时间的秒数。
- 【命令模式】 line 配置模式
- 【使用指导】 配置 LINE 上, 已接受连接的超时时间, 当超过配置时间, 没有任何输入时, 将中断此连接。
在 LINE 配置模式下使用 **no exec-timeout** 命令, 取消 LINE 下连接的超时设置。

配置会话超时时间

- 【命令格式】 **session-timeout** *minutes* [**output**]
- 【参数说明】 *minutes* : 指定的超时时间的分钟数。
output : 是否将输出数据也作为输入, 来判断是否超时。
- 【命令模式】 line 配置模式
- 【使用指导】 配置 LINE 上, 连接到远程终端的会话超时时间, 在指定时间内, 没有任何输入时, 将中断此会话。
在 LINE 配置模式下使用 **no session-timeout** 命令, 取消 LINE 下到远程终端的会话超时时间设置。

启用锁住 line 终端的功能

- 【命令格式】 **lockable**
- 【参数说明】 -
- 【命令模式】 line 配置模式
- 【使用指导】 -

锁住当前 line 终端

- 【命令格式】 **lock**
- 【参数说明】 -
- 【配置模式】 line 配置模式
- 【使用指导】 -

配置举例

建立与远程网络设备的 Telnet 会话

- 【配置方法】
- 建立与远程网络设备的 Telnet 会话, 远程网络设备的 IP 地址是 192.168.65.119。

```
Ruijie# telnet 192.168.65.119
Trying 192.168.65.119 ... Open
User Access Verification
Password:
Ruijie# telnet 2AAA:BBBB::CCCC
Trying 2AAA:BBBB::CCCC ... Open
User Access Verification
Password:
```

- 【检验方法】
- 如果能正常与远程设备建立会话, 则配置成功。

连接超时

- 【配置方法】
- 设置超时时间为 20min

```
Ruijie# configure terminal//进入全局配置模式
Ruijie# line vty 0 //进入 LINE 配置模式
Ruijie(config-line)#exec-timeout 20 //设置超时时间为 20min
```

- 【检验方法】
- 连接到本地设备的终端，在这段时间内容没有任何输入，将断开连接并退出。

设置超时时间为 20min

- 【配置方法】
- 设置超时时间为 20min

```
Ruijie# configure terminal//进入全局配置模式
Ruijie(config)# line vty 0 //进入 LINE 配置模式
Ruijie(config-line)#session-timeout 20//设置超时时间为 20min
```

- 【检验方法】
- 连接到远程设备的终端，在这段时间内容没有任何输入，将断开连接并退出。

常见配置错误

- 无

2.4.3 设置系统基本参数

配置效果

- 设置系统的基本参数。

注意事项

- 无

配置方法

设置系统的日期和时钟

- 必选配置。
- 通过手工的方式来设置网络设备上的时间。当你设置了网络设备的时钟后，网络设备的时钟将以你设置的时间为准一直运行下去，即使网络设备下电，网络设备的时钟仍然继续运行。

 但是对于没有提供硬件时钟的网络设备，手工设置网络设备上的时间实际上只是设置软件时钟，它仅对本次运行有效，当网络设备下电后，手工设置的时间将失效。

更新硬件时钟

- 可选配置。
- 如果硬件时钟和软件时钟不同步，需要通过软件时钟的日期和时间复制给硬件时钟时，执行此配置项。

设置系统名称

- 可选配置。可以修改默认的系统名称。

设置命令提示符

- 可选配置。可以修改默认的命令提示符名称。

设置每日通知

- 可选配置。在希望告知使用者一些重要提示或警告信息时，可以选择在系统上设置每日通知。
- 你可以创建包含一行或多行信息的通知信息，当用户登录网络设备时，这些信息将会被显示。

配置登录标题

- 可选配置。如果希望对使用者在登录或退出作一些重要信息的提示，可以选择配置此项。

设置控制台的传输速率

- 可选配置。可以修改默认的控制台速率。

检验方法

- 使用 **show clock** 命令来显示系统时间信息。
- 标题的信息将在你登录网络设备时显示。
- 使用 **show version** 命令查看系统、版本信息。

相关命令

设置系统的日期和时钟

【命令格式】 **clock set** *hh:mm:ss month day year*

【参数说明】 *hh:mm:ss*：当前时间，格式为小时（24小时制）：分钟：秒。
day：日期（1-31），一个月中的日期。
month：月份（1-12），一年中的月份。
year：公元年（1993-2035），不能使用缩写。

【命令模式】 特权用户模式

【使用指导】 使用该命令设置系统时间，方便管理。

对于没有提供硬件时钟的网络设备，使用 **clock set** 设置网络设备上的时间仅对本次运行有效，当网络设备下电后，手工设置的时间将失效。

更新硬件时钟

- 【命令格式】 **clock update-calendar**
- 【参数说明】 -
- 【命令模式】 特权用户模式
- 【使用指导】 软件时钟就会覆盖硬件时钟的值。

设置系统名称

- 【命令格式】 **hostname name**
- 【参数说明】 *name*：系统名称，名称必须由可打印字符组成，长度不能超过 63 个字节。
- 【命令模式】 全局模式
- 【使用指导】 在全局配置模式下使用 **no hostname** 来将系统名称恢复缺省值。

设置命令提示符

- 【命令格式】 **prompt string**
- 【参数说明】 *string*：名称必须由可打印字符组成，如果长度超过 32 个字符，则截取其前 32 个字符。
- 【命令模式】 特权用户模式
- 【使用指导】 在全局配置模式下使用 **no prompt** 来将命令提示符恢复为缺省值。

设置每日通知

- 【命令格式】 **banner motd c message c**
- 【参数说明】 *c*：分界符，这个分界符可以是任何字符(比如'&'等字符)。
- 【命令模式】 全局配置模式
- 【使用指导】 输入分界符后，然后按回车键，即可以开始输入文本，需要在键入分界符并按回车键来结束文本的输入。需要注意的是，如果键入结束的分界符后仍然输入字符，则这些字符将被系统丢弃。通知信息的文本中不应该出现作为分界符的字母，文本的长度不能超过 255 个字节。

配置登录标题

- 【命令格式】 **banner login c message c**
- 【参数说明】 *c*：分界符，这个分界符可以是任何字符(比如'&'等字符)。
- 【命令模式】 全局配置模式
- 【使用指导】 输入分界符后，然后按回车键，即可以开始输入文本，需要在键入分界符并按回车键来结束文本的输入，需要注意的是，如果键入结束的分界符后仍然输入字符，则这些字符将被系统丢弃。登录标题的文本中不应该出现作为分界符的字母，文本的长度不能超过 255 个字节。
在全局配置模式下使用 **no banner login** 来删除登录标题。

设置控制台的传输速率

- 【命令格式】 **speed speed**
- 【参数说明】 *speed*：，单位是 bps。对于串行接口。只能将传输速率设置为 9600、19200、38400、57600、115200 中的一个，缺省的速率是 9600。
- 【命令模式】 line 配置模式
- 【使用指导】 用户可以根据需要来设置异步线路的波特率。命令 **speed** 将同时设置异步线路的接收速率以及发送速率。

配置举例

配置系统时间

- 【配置方法】
- 把系统时间改成 2003-6-20，10:10:12

```
Ruijie# clock set 10:10:12 6 20 2003 //设置系统时间和日期
```

- 【检验方法】
- 在特权模式下使用 **show clock** 命令来显示系统时间信息

```
Ruijie# show clock //确认修改系统时间生效  
clock: 2003-6-20 10:10:54
```

配置每日通知

- 【配置方法】
- 使用(#)作为分界符，每日通知的文本信息为“Notice: system will shutdown on July 6th.”

```
Ruijie(config)# banner motd #//开始分界符  
Enter TEXT message. End with the character '#'.  
Notice: system will shutdown on July 6th. # //结束分界符  
Ruijie(config)#
```

- 【检验方法】
- 使用 **show running-config** 命令查看配置。
 - 使用控制台、Telnet 或 SSH 连接本地设备，进入命令行界面之前时将显示每日通知信息。

```
C:\>telnet 192.168.65.236  
Notice: system will shutdown on July 6th.  
Access for authorized users only. Please enter your password.  
User Access Verification  
Password:
```

配置登录标题

- 【配置方法】
- 使用(#)作为分界符，登录标题的文本为“Access for authorized users only. Please enter your password.”

```
Ruijie(config)# banner login #//开始分界符  
Enter TEXT message. End with the character '#'.  
Access for authorized users only. Please enter your password.  
# //结束分界符  
Ruijie(config)#
```

- 【检验方法】
- 使用 **show running-config** 命令查看配置。
 - 使用控制台、Telnet 或 SSH 连接本地设备，进入命令行界面之前时将显示登录标题信息。

```
C:\>telnet 192.168.65.236  
Notice: system will shutdown on July 6th.  
Access for authorized users only. Please enter your password.  
User Access Verification  
Password:
```

如何将串口速率设置为 57600 bps

【配置方法】

- 将串口速率设置为 57600 bps

```
Ruijie# configure terminal //进入全局配置模式
Ruijie(config)# line console 0 //进入控制台线路配置模式
Ruijie(config-line)# speed 57600 //设置控制台速率为 57600
Ruijie(config-line)# end //回到特权模式
```

【检验方法】

- 使用 **show** 命令查看。

```
Ruijie# show line console 0 //查看控制台配置
CON      Type      speed  Overruns
* 0      CON      57600  0
Line 0, Location: "", Type: "vt100"
Length: 25 lines, Width: 80 columns
Special Chars: Escape Disconnect Activation
                ^x      none      ^M
Timeouts:      Idle EXEC      Idle Session
                never      never
History is enabled, history size is 10.
Total input: 22 bytes
Total output: 115 bytes
Data overflow: 0 bytes
stop rx interrupt: 0 times
Modem: READY
```

常见配置错误

- 无

2.4.4 打开或关闭指定的服务

配置效果

- 在系统运行过程中，可以动态地调整系统所提供的服务，打开与关闭指定的服务（SNMP Server / SSH Server / Telnet Server）。

注意事项

-

配置方法

▾ 打开 SNMP Server / SSH Server / Telnet Server

- 可选配置。在需要使用这些服务时执行此配置项。

检验方法

- 使用 **show running-config** 命令查看配置。
- 使用 **show services** 命令查看服务的开启状态。

相关命令

▾ 打开 SSH-Server/telnet-server/snmp-agent

【命令格式】 **enable service { ssh-server | telnet-server | snmp-agent }**

【参数说明】 **ssh-server** : 打开与关闭 SSH Server。
telnet-server : 打开与关闭 Telnet Server。
snmp-agent : 打开与关闭 Snmp Agent。

【命令模式】 全局模式

【使用指导】 该命令用于打开与关闭指定的服务。

配置举例

▾ 打开 SSH Server

【配置方法】 ● 打开 SSH Server

```
Ruijie# configure terminal //进入全局配置模式
Ruijie(config)#enable service ssh-server //打开 SSH Server
```

【检验方法】 ● 使用 **show running-config** 命令查看配置。
● 使用 **show ip ssh** 命令查看 SSH 服务配置和运行状况。

常见配置错误

无

2.4.5 回滚系统配置

配置效果

设备创建检查点之后，就保存了一份当时运行配置的副本，如果在某种情况下需要恢复该副本配置，就可以使用配置回滚将系统配置恢复。

注意事项

无

配置方法

创建检查点

- 可选配置，在需要创建配置检查点的时候设置此项。
- 用户可以在特权模式下执行 **checkpoint** 命令来创建一个检查点，就是创建一个当前运行配置的副本。
- 用户在创建检查点之后，可以在特权模式执行 **show checkpoint** 命令来查看所创建的检查点信息。

回滚配置

- 可选配置，在需要将设备配置回滚到某个检查点的时候设置此项。
- 用户可以在特权模式下执行 **rollback** 命令，将该检查点所对应的配置副本，重新应用为当前的运行配置。

清除检查点数据

- 可选配置，在想清除所有的配置检查的时候设置此项。
- 用户可以在特权模式下执行 **clear checkpoint database** 命令删除所有的检查点。

检验方法

- 使用 **show running-config** 命令查看当前的运行配置。
- 使用 **show checkpoint** 命令查看要回滚的检查点信息，并使用 **more** 命令打开检查点所对应的配置文件，查看检查点对应的配置副本信息。
- 使用 **rollback** 命令进行配置回滚操作，完成之后再次使用 **show running-config** 命令查看配置，可以检查配置副本是否被重新应用称为当前的运行配置。

相关命令

创建检查点

- 【命令格式】 **checkpoint** [*cp-name*] [**description** *description*]
- 【参数说明】 *cp-name* : 检查点名字，1-80 个字符。
description *description* : 检查点描述，最长 80 个字符。
- 【缺省配置】 -
- 【命令模式】 特权用户模式

【使用指导】 使用该命令可以创建一个检查点，并可以指定检查点名称、描述。

↘ 回滚配置

【命令格式】 **rollback running-config checkpoint** *cp-name*

【参数说明】 *cp-name* : 检查点名字，1-80 个字符。

【缺省配置】 -

【命令模式】 特权用户模式

【使用指导】 使用该命令可以回滚一个指定的检查点。

↘ 清除检查点数据

【命令格式】 **clear checkpoint database**

【参数说明】 -

【命令模式】 特权用户模式

【使用指导】 使用该命令可以删除系统中所有的检查点，以及删除检查点对应的配置数据信息。

常见错误

无

2.4.6 设置重启策略

配置效果

设备在某些情况下需要重启，设置重启策略能使设备按照预设的方式进行重启。

注意事项

无

配置方法

↘ 直接重启

表示立即重启设备，用户可以在特权模式下直接键入 **reload** 命令来重启系统。

↘ 定时重启

```
reload in mmm | hhh:mm [string]
```

指定系统在经过一定时间间隔后重启。这里的时间间隔由 *mmm* 或 *hhh:mm* 决定，以分钟为单位，用户可以任选一种格式输入。参数 *string* 是一个帮助提示，用户可以在这里为这个计划起一个助记名，以便能直观地反映该重启的用途。

```
reload at hh:mm month day year [string]
```

指定系统在将来的某个时间点重启。输入的时间值必须是将来的某个时间点。参数 `year` 是可选的,如果用户没有输入,则默认的年份是系统时钟的年份,由于我们限制了时间跨度不能超过 31 天,所以一般地,如果当前系统日期是在 1 月 1 日到 11 月 30 日之间,则用户就没有必要输入年份。但是,如果当前系统月份为 12 月份,这时用户指定的重启时间就有可能并且允许是明年 1 月的某个时间,在这种情况下,用户就需要输入年份来通知系统重启时间是明年 1 月份而不是今年 1 月份,如果没有输入,则默认的会被认为是今年 1 月份,从而导致设置失败。

! 如果用户要使用 `at` 选项,则要求当前系统必须支持时钟功能。建议使用之前先配置好系统的时钟,以便更切合您的用途。如果用户之前已经设置了重启计划,则后面再设置的计划将覆盖前面的设置。如果用户已经设置了重启计划,假如在该计划生效前用户重启了系统,则该计划将丢失。

! 重启计划中的时间与当前时间的跨度不能超过 31 天并且要大于当前系统时间。同时用户在设置了重启计划之后最好不要再修改系统时钟,否则有可能会導致设置失效,比如将系统时间调到重启时间之后。

删除已设置的重启策略

使用 `reload cancel` 命令删除已设置的重启策略。

检验方法

-

相关命令

重启设备

【命令格式】 `reload [text | in [hh:] mm [text] | at hh:mm [month day year] [text] | cancel]`

【参数说明】 `text`: 重启的原因, 1-255 个字节。

`in [hh:] mm`: 在指定时间间隔后重启系统, 最大时间为 24 天。

`at hh:mm`: 在指定的时刻重启系统。

`month`: 月份的字符表示, 如 3 月份, Mar。

`day`: 日期, 从 1 到 31。

`year`: 公元年 (1993-2035), 不能使用缩写。

`cancel`: 取消重启计划。

【命令模式】 特权用户模式

【使用指导】 使用该命令可以指定设备在指定的时刻启动, 方便进行管理。

常见错误

无

2.5 监视与维护

查看运行情况

作用	命令
show checkpoint { <i>cp-name</i> [all] summary }	显示检查点信息，或所有检查点的摘要信息。
show clock	显示当前系统时间。
show line { console <i>line-num</i> vty <i>line-num</i> <i>line-num</i> }	查看线路的配置信息。
show reload	查看系统的重新启动设置。
show running-config [interface <i>interface</i>]	查看当前设备系统正在运行的配置信息或某个接口下的配置信息。
show startup-config	查看存储在 NVRAM（非易失性随机存取存储器）上设备的配置。
show this	查看系统当前模式下生效的配置信息。
show version [devices module slots]	查看一些系统的信息。
show sessions	显示已经建立 Telnet Client 实例的每个实例信息。

3 LINE

3.1 概述

在网络设备上一般都具有多种类型的终端线路（line），并针对这些终端按类进行分组管理，对这些终端进行的配置称为线路（line）配置。在网络设备上，终端线路类型分为 CTY、VTY 等。

协议规范

- 无

3.2 典型应用

典型应用	场景描述
通过控制台访问设备	通过控制台进入网络设备的命令行界面。
通过 VTY 访问设备	通过 Telnet 或 SSH 进入网络设备的命令行界面。

3.2.1 通过控制台访问设备

应用场景

图 3-1



【注释】 A 为需要被管理的网络设备。
PC 为网络管理站。

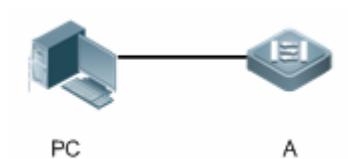
功能部属

网络管理站使用串口线连接被管理的网络设备的控制台端口，用户在网络管理站上，通过控制台软件（超级终端或其他终端仿真软件）连接网络设备上的控制台并进入命令行界面，对网络设备进行配置和管理。

3.2.2 通过VTY访问设备

应用场景

图 3-2



【注释】 A 为需要被管理的网络设备。
PC 为网络管理站。

功能部属

网络管理站和被管理的网络设备通过网络连接，用户在网络管理站上，通过 VTY 客户端软件（例如 Putty），使用 Telnet 或 SSH 连接网络设备上并进入命令行界面，对网络设备进行配置和管理。

3.3 功能详解

基本概念

▾ CTY

CTY 线路类型指的是控制台端口（Console Port），大多数网络设备都会具有一个控制台端口，用户可以使用控制台终端，通过这个端口访问本地系统。

▾ VTY

VTY 线路类型是虚拟终端线路，并没有与之对应的硬件，虚拟终端线路用于 Telnet 或 SSH 连接。

功能特性

功能特性	作用
基本功能	配置终端，显示、清除终端连接信息等。

3.3.1 基本功能

工作原理

无

相关配置

配置终端线路

在全局配置模式下，使用 **line** 命令可以进入指定的终端配置模式。

可以对终端的各项属性进行配置。

清除终端连接

当用户终端已经与设备连接时，对应的终端线路就处于占用状态，此时使用 **show user** 命令可以查看这些终端线路的连接状态。如果要使用户终端断开与网络设备的连接，可以使用 **clear line** 命令指定清除一个终端。被清除的终端线路上如果有关联的通讯协议（例如 Telnet、SSH 等）将会断开，已经进入的命令界面也会退出。清除后的终端线路将恢复为非占用的状态，用户可以重新建立起连接。

设置 VTY 终端数目

使用 **line vty** 命令不仅可以进入 VTY 线路配置模式，还可以指定 VTY 终端的数目。

默认的 VTY 终端数目为 5 个，编号为 0~5。可以将终端数目最多扩展到 36 个，扩展的编号为 5~35。扩展的终端可以被删除，但默认的终端不可删除。

3.4 配置详解

配置项	配置建议 & 相关命令	
进入 line 模式	 必选配置。用于进入 line 模式。	
	line [console vty] first-line [last-line]	进入到指定的 LINE 模式
	line vty line-number	增加或减少当前可以使用的 VTY 连接数目

3.4.1 进入line模式

配置效果

进入 line 模式进行其他功能项的配置。

注意事项

无

配置方法

进入 LINE 模式

- 必选配置。
- 若无特殊情况，应在每台设备上进入 line 模式进行功能配置。

增加/减少 LINE VTY 数目

- 可选配置。
- 在需要增加或减少 LINE 线路时应使用此配置项。

检验方法

使用 **show line** 命令查看线路的配置信息。

相关命令

进入 LINE 模式

【命令格式】 **line** [console | vty] *first-line* [*last-line*]

【参数说明】 **console** : 控制台口。

vty : 虚终端线路，适用于 Telnet 或 SSH 连接。

first-line : 要进入的 *first-line* 编号。

last-line : 要进入的 *last-line* 编号。

【命令模式】 全局配置模式

【使用指导】 -

增加/减少 LINE VTY 数目

【命令格式】 **line vty** *line-number*

【参数说明】 *line-number* : VTY 连接数目，范围：0~35。

【命令模式】 全局配置模式

【使用指导】 使用 **no line vty** *line-number* 命令减少当前可以使用的 VTY 连接数目。

查看线路配置信息

【命令格式】 **show line** { **console** *line-num* | **vty** *line-num* | *line-num* }

【参数说明】 **console** : 控制台口。

vty : 虚终端线路，适用于 Telnet 或 SSH 连接。

line-num : 查看的 line 线路。

【命令模式】 特权配置模式

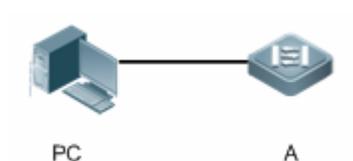
【使用指导】 -

配置举例

配置 VTY 终端扩展

【网络环境】

图 3-3



- 【配置方法】
- PC 使用控制台线连接网络设备 A，通过控制台终端进入命令行界面。
 - 执行 **show user** 查看终端线路连接状态。
 - 执行 **show line console 0** 查看控制台线路状态。
 - 进入全局配置模式，使用 **line vty** 命令将 VTY 终端数目扩展至 36 个。

A

```
Ruijie#show user
Line          User          Host(s)          Idle          Location
-----
* 0 con 0    ---          idle            00:00:00    ---

Ruijie#show line console 0

CON  Type  speed  Overruns
* 0   CON   9600   0
Line 0, Location: "", Type: "vt100"
Length: 24 lines, Width: 79 columns
Special Chars: Escape Disconnect Activation
                ^x      ^D      ^M
Timeouts:      Idle EXEC  Idle Session
                00:10:00  never
History is enabled, history size is 10.
Total input: 490 bytes
Total output: 59366 bytes
Data overflow: 0 bytes
stop rx interrupt: 0 times

Ruijie#show line vty ?
<0-5>  Line number
```

```
Ruijie#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#line vty 35
Ruijie(config-line)#
*Oct 31 18:56:43: %SYS-5-CONFIG_I: Configured from console by console
```

- 【检验方法】
- 输入 **show line** 命令，获取帮助时可以发现终端数量已经被扩展。
 - 执行 **show running-config** 命令查看配置。

A

```
Ruijie#show line vty ?
  <0-35> Line number

Ruijie#show running-config

Building configuration...
Current configuration : 761 bytes

version 11.0(1C2B1) (10/16/13 04:23:54 CST -ngcf78)
ip tcp not-send-rst
vlan 1
!
interface GigabitEthernet 0/0
!
interface GigabitEthernet 0/1
 ip address 192.168.23.164 255.255.255.0
!
interface GigabitEthernet 0/2
!
interface GigabitEthernet 0/3
!
interface GigabitEthernet 0/4
!
interface GigabitEthernet 0/5
!
interface GigabitEthernet 0/6
!
interface GigabitEthernet 0/7
!
interface Mgmt 0
!
line con 0
```

```
line vty 0 35
 login
 !
end
```

常见错误

无

3.5 监视与维护

清除各类信息



在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除线路的连接状态。	clear line { console <i>line-num</i> vtty <i>line-num</i> <i>line-num</i> }

查看运行情况

作用	命令
查看线路的配置信息。	show line { console <i>line-num</i> vtty <i>line-num</i> <i>line-num</i> }

4 TIME RANGE

4.1 概述

Time range 是一个时间控制服务，它提供给某些应用进行时间控制。例如，如果想要让 ACL 在一个星期的某些时间段内生效，可以配置一个 time range 并让 ACL 和这个 time range 相关联。

4.2 典型应用

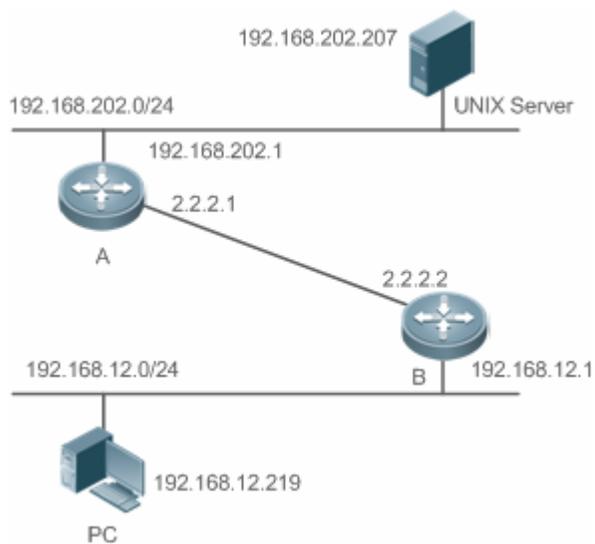
典型应用	场景描述
ACL中的time range应用	应用在 ACL 模块，满足 ACL 基于时间生效的需求。

4.2.1 ACL中的time range应用

应用场景

某单位限制只能在正常上班时间访问远程 UNIX 主机 TELNET 服务：

图 4-1



【注释】 要求通过在设备 B 上配置访问列表，实现以下安全功能：

192.168.12.0/24 网段的主机只能在正常上班时间访问远程 UNIX 主机 TELNET 服务。

功能部属

- 在网络设备 B 上使用 ACL 对来自 192.168.12.0/24 网段的 TELNET 访问进行控制，ACL 应用关联一个 time range，只有在工作时间才允许其访问 Unix 主机。

4.3 功能详解

基本概念

▾ 绝对时间区间

绝对时间区间是指可以为 time range 设置一个起始时间以及结束时间的区间。典型的绝对时间区间例如[2000 年 1 月 1 日 12 : 00 , 2001 年 1 月 1 日 12 : 00]。Time range 应用关联到这个 time range 之后，可以在该时间区间之内使某项功能起作用。

▾ 周期时间

周期时间是指可以为 time range 设置一个周期性的时间。典型的周期时间如“每周一 8 : 00 到每周五 17 : 00”。Time range 应用关联到这个 time range 之后，可以周期性地每周一到每周五使某项功能起作用。

功能特性

功能特性	作用
使用绝对时间区间	设置绝对时间区间允许 time range 应用在这个绝对时间区间之内使某项功能生效。
使用周期时间	设置周期时间允许 time range 应用在某个周期性的时间之内使某项功能生效。

4.3.1 使用绝对时间区间

工作原理

基于 time range 的应用在开启某项功能时，会判断当前的时间是否处于绝对时间区间之内，如果在其中，则可以让该功能在当前时间生效或者在当前时间不生效。

相关配置

▾ 配置 time range

缺省情况下，没有配置 time range。

使用 `time-range time-range-name` 命令来配置一个 time range。

▾ 配置绝对时间区间

缺省情况下，绝对时间区间为[0 年 1 月 1 日 00 : 00 , 9999 年 12 月 31 日 23 : 59]。

使用 **absolute** { [start time date] [end time date] }命令来配置绝对时间区间。

4.3.2 使用周期时间

工作原理

基于 time range 的应用在开启某项功能时，会判断当前的时间是否处于周期时间之内，如果在其中，则可以让该功能在当前时间生效或者在当前时间不生效。

相关配置

配置 time range

缺省情况下，没有配置 time range。

使用 **time-range** *time-range-name* 命令来配置一个 time range。

配置周期时间

缺省情况下，没有配置周期时间。

使用 **periodic** *day-of-the-week* *time* **to** [*day-of-the-week*] *time* 命令来配置周期时间。

4.4 配置详解

配置项	配置建议 & 相关命令
配置time range	 必须配置。如果要使用 time range 功能，必须配置 time range。
	time-range <i>time-range-name</i> 配置 time range。
	 可选配置。配置分类参数。
	absolute { [start time date] [end time date] } 配置绝对时间区间。
	periodic <i>day-of-the-week</i> <i>time</i> to [<i>day-of-the-week</i>] <i>time</i> 配置周期时间。

4.4.1 配置time range

配置效果

- 配置 time range，配置其绝对时间区间或周期时间，以便让 time range 应用在对的时间区间内使某项功能生效。

配置方法

配置 time range

- 必须配置。
- 在需要应用 time range 的设备上配置。

配置绝对时间区间

- 可选配置。

配置周期时间

- 可选配置。

检验方法

- 使用 **show time-range** [*time-range-name*]命令，可以查看所配置的 time range 信息。

相关命令

配置 time range

【命令格式】 **time-range** *time-range-name*

【参数说明】 *time-range-name*：要创建的 time range 的名字。

【命令模式】 全局模式

【使用指导】 有些应用（例如 ACL）可能基于时间运行，比如让 ACL 在一个星期的某些时间段内生效等。为了达到这个要求，必须首先配置一个 Time-Range。创建完 time range 之后，可以在 time range 模式中配置相应的时间控制。

配置绝对时间区间

【命令格式】 **absolute** { [*start time date*] [*end time date*] }

【参数说明】 **start time date**：区间的开始时间。

end time date：区间的结束时间。

【命令模式】 time-range 模式

【使用指导】 如果想要让某个功能在一个绝对时间区间内生效，可以使用 **absolute** 命令配置一个开始和结束的时间区间。

配置周期时间

【命令格式】 **periodic** *day-of-the-week time to* [*day-of-the-week*] *time*

【参数说明】 *day-of-the-week*：周期时间开始或者结束是在星期几

time：周期时间开始或者结束是在几点几分

【命令模式】 time-range 模式

【使用指导】 如果想要让某个功能在一个周期时间内生效，可以使用 **periodic** 命令配置一个周期时间。

4.5 监视与维护

查看运行情况

作用	命令
显示 time range。	show time-range [<i>time-range-name</i>]

5 HTTP 服务

5.1 概述

HTTP (Hypertext Transfer Protocol , 超文本传输协议) 用来在 Internet 上传递 Web 页面信息。HTTP 位于 TCP/IP 协议栈的应用层 , 传输层采用面向连接的 TCP。

HTTPS (Hypertext Transfer Protocol Secure) 是支持 SSL (Secure Sockets Layer , 安全套接层) 协议的 HTTP 协议。主要作用是在不安全的网路上创建一个安全的通道 , 保证信息很难被监听以及对中间人攻击提供一定的合理保护。HTTPS 目前已被广泛用于互联网上安全敏感的通讯 , 如电子交易支付等。

协议规范

- RFC1945 : Hypertext Transfer Protocol -- HTTP/1.0
- RFC2616 : Hypertext Transfer Protocol -- HTTP/1.1
- RFC2818 : Hypertext Transfer Protocol Over TLS -- HTTPS

5.2 典型应用

典型应用	场景描述
HTTP应用服务	用户通过 Web 管理设备。

5.2.1 HTTP应用服务

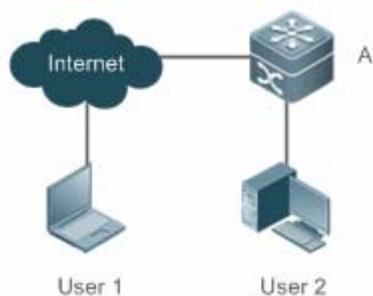
应用场景

设备开启 HTTP 服务后 , 用户只需在 PC 机浏览器中输入 http://+设备的 IP 地址 , 认证通过后就可以登陆到 Web 管理界面。在 Web 界面中 , 用户可以进行设备状态信息监控、配置设备、上传和下载文件等操作。

以下图为例 , 用户进行 Web 管理。

- 用户可以通过 Internet 进行远程访问设备 , 也可以在本地局域网中通过登陆 Web 服务器对设备进行配置管理。
- 用户可以根据实际情况 , 选择在设备上单独启用 HTTPS 服务或者 HTTP 服务 , 也可以同时启用 HTTPS 和 HTTP 服务。
- 用户还可以再浏览器上设置使用 HTTP/1.0 还是 HTTP/1.1 协议来访问设备的 HTTP 服务。

图 5-1



- 【注释】 A 为锐捷设备。
用户 User1 通过 Internet 网络访问设备。
用户 User2 通过局域网访问设备。

功能部署

- 设备运行 HTTP 协议，用户在 PC 浏览器中通过 `http://设备的 ip 地址`，访问设备。
- 设备运行 HTTPS 协议，用户在 PC 浏览器中通过 `https://设备的 ip 地址`，访问设备。

5.3 功能详解

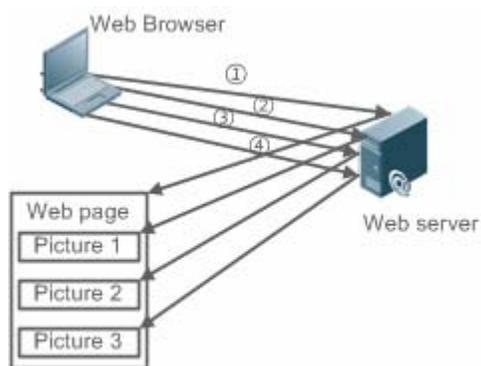
基本概念

HTTP 服务

HTTP 服务是指在 Internet 上利用 HTTP 协议传递 Web 页面信息。HTTP/1.0 是目前业界使用最广泛的 HTTP 协议版本，由于一个 Web 服务器每天可能有上万甚至上百万的访问量，为了便于连接管理，HTTP/1.0 采用短连接方式。一个请求创建一个 TCP 连接，请求完成后释放连接，服务器不需要记录和跟踪过去的请求。HTTP/1.0 虽然简化了连接管理，但是却引入了性能缺陷。

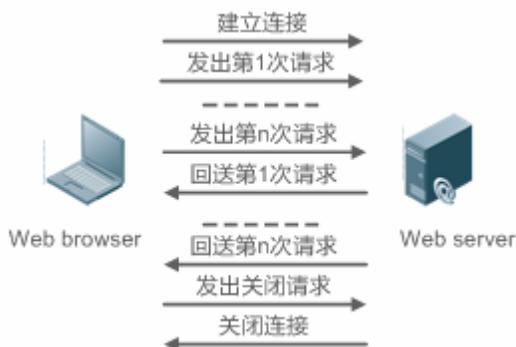
例如一个网页中可能需要很多图片，网页中包括的不是真正的图片内容，而是它们的 URL 连接地址，这样浏览器在访问过程中会发出多次请求，每次请求都要建立一个单独的连接，每次连接都是完全隔离的。建立和释放连接是一个相对费劲的过程，从而严重影响了客户机和服务器的性能，如下图所示。

图 5-2



HTTP/1.1 克服了这个缺陷。该版本支持持久连接，即一个连接可以传输多个请求和响应，这样客户机可以不用等待上一次请求完成就可以发送第二个请求，减少了网络时延，提高性能，如下图所示。

图 5-3



目前，锐捷设备支持 HTTP/1.0 和 HTTP/1.1 两种协议版本。

i 设备使用哪种协议版本由 Web 浏览器决定。

HTTPS 服务

HTTPS 服务就是在 HTTP 基础上加入 SSL 层，其安全基础是 SSL。要使协议能够正常运行，服务器必需有 PKI (Public Key Infrastructure, 公钥基础设施) 证书，而客户端则不一定。SSL 协议提供的服务主要有：

- 认证用户和服务器，确保数据发送到正确的客户机和服务器；
- 加密数据以防止数据中途被窃取；
- 维护数据的完整性，确保数据在传输过程中不被改变。

图 5-4



HTTP 升级服务

本地升级时，设备作为 HTTP 服务器，用户通过 Web 浏览器登陆到设备，将需要升级的文件上传到设备，实现设备上文件的升级。

功能特性

功能特性	作用
HTTP服务	用户通过 Web 界面登陆到设备中进行配置与管理。
HTTP本地升级服务	将需要升级的文件上传到设备，实现设备上文件的升级。

5.3.1 HTTP服务

HTTP 是为 Web 管理提供服务，用户通过 Web 界面登录到设备中进行配置与管理。

工作原理

Web 管理包括 Web 客户端和 Web 服务器，同理 HTTP 服务也采用客户端/服务器模式。HTTP 客户端内嵌在 Web 管理客户端的 Web 浏览器中，能够发送 HTTP 报文和接收处理 HTTP 响应报文。而 Web 服务器(即 HTTP 服务器)则内嵌于设备中。客户端和服务器之间的信息交互过程如下：

- 在客户端与服务器之间建立 TCP 连接，HTTP 服务默认端口号是 80，HTTPS 服务默认端口号是 443。
- 客户端向服务器发送请求消息。
- 服务器解析客户端发送的请求，请求内容包括获取 web 页面、执行 cli 命令，上传文件等。
- 服务器执行完请求内容后，将响应发送回给客户端。

相关配置

使能 HTTP 服务

缺省情况下，HTTP 服务功能关闭。

使用 **enable service web-server** 命令可以使能 HTTP 服务功能，包括 HTTP 服务和 HTTPS 服务。

必须使能 HTTP 服务功能，用户才能通过 Web 界面登录到设备中进行配置与管理。

配置 HTTP 认证信息

缺省情况下，系统默认创建两个账号，admin 账号和 guest 账号，这两个账号不可被删除，只可修改密码。其中 guest 账号默认对应 level 2 权限，在 web 端只拥有查看系统首页的权限。管理员账号为 admin 对应 level 0 权限，在 web 端管理员账号拥有所有的功能，并且可以编辑其他管理账号并授权该账号可访问的页面，新添加的账号对应 level 1 权限。

使用 **webmaster level** 命令可以配置认证的用户名和密码。

通过配置该命令，用户需要输入所配置的用户名和密码进行认证才能登陆 Web 页面。

配置 HTTP 服务端口

缺省情况下，HTTP 服务端口为 80。

使用 **http port** 命令可以配置 HTTP 服务端口号，取值范围是 80 及 1025-65535。

通过配置 HTTP 服务端口号，可以减少非法用户对 HTTP 服务的攻击。

配置 HTTPS 服务端口

缺省情况下，HTTPS 服务端口为 443。

使用 **http secure-port** 命令可以配置 HTTPS 服务端口号，取值范围是 443 及 1025-65535。

通过配置 HTTPS 服务端口号，可以减少非法用户对 HTTPS 服务的攻击。

5.3.2 HTTP本地升级服务

设备作为 HTTP 服务器，用户通过 Web 浏览器登陆到设备，将需要升级的文件（包括组件包、web 包）上传到设备或者直接通过 tftp 上传文件到设备中。

工作原理

- 通过 Web 的“本地升级”功能上传组件包或 web 包
- 设备接收文件信息，接收成功后进行版本与合法性校验
- 文件校验成功后，如果是 web 包，直接升级；如果是组件包，用户在浏览器端通过是否重启设备来决定升级与否。

相关配置

更新 Web 包

使用 **upgrade web download** 命令从 TFTP 服务器上下载 WEB 包。

通过配置该命令，从 TFTP 服务器上下载 WEB 包，合法性校验通过后，WEB 包直接进行升级，无需重启设备。

也可以使用 **upgrade web** 命令直接升级本地存在的 WEB 包。

更新子系统组件

缺省情况下，不管是通过浏览器还是 TFTP 上传子系统组件，设备默认都是不升级的。

用户要升级子系统组件，都必须重启设备。

5.4 配置详解

配置项	配置建议 & 相关命令	
配置HTTP服务	 必须配置。用于启动 HTTP 服务。	
	enable service web-server	使能 HTTP 服务
	webmaster level	配置 HTTP 认证信息
	http port	配置 HTTP 服务端口
	http secure-port	配置 HTTPS 服务端口
配置HTTP本地升级	 必须配置。用于实现 HTTP 本地升级。	
	upgrade web	升级设备存放的 WEB 包。
	upgrade web download	自动从服务器上下载 WEB 包，并自动升级

5.4.1 配置HTTP服务

配置效果

设备开启 HTTP 服务，用户通过认证后可以登陆到 Web 管理界面，进行设备状态信息监控、配置设备、上传和下载文件等操作。

配置方法

使能 HTTP 服务

- 必须配置。
- 若无特殊要求，应在每台锐捷设备上使能 HTTP 服务，否则 web 服务不可访问。

配置 HTTP 认证信息

- 默认情况下，已经配置用户名 admin、guest，对应的密码是 admin、guest。
- 若无特殊要求，用户可以使用默认的用户名登陆 web 页面，直接通过 web 浏览器来更新认证信息；如果一直使用默认账户，会存在安全隐患，因为 IP 一旦泄露，非授权人员就可以通过 web 获取到设备的配置信息等。

配置 HTTP 服务端口

- 如果要求改变 HTTP 服务端口，则必须配置 HTTP 服务端口。
- 若无特殊要求，可以使用默认的 HTTP 服务端口 80 进行访问。

配置 HTTPS 服务端口

- 如果要求改变 HTTPS 服务端口，则必须配置 HTTPS 服务端口。
- 若无特殊要求，可以使用默认的 HTTPS 服务端口 443 进行访问。

检验方法

- 用户在浏览器上输入 `http://设备的ip:服务端口`，验证浏览器是否会跳转到认证页面。
- 用户在浏览器上输入 `https://设备的ip:服务端口`，验证浏览器是否会跳转到认证页面。

相关命令

使能 HTTP 服务

【命令格式】 **enable service web-server [http | https | all]**

【参数说明】 **http | https | all** :打开相应的服务。**http** 为打开 HTTP 服务，**https** 为打开 HTTPS 服务，**all** 为同时打开 HTTP 和 HTTPS 服务。缺省为同时打开 HTTP 和 HTTPS 服务。

【命令模式】 全局模式

【使用指导】 如果执行该命令时后面不跟任何关键字或者跟 **all**，则表示同时打开 HTTP 服务和 HTTPS 服务；如果跟 **http** 关键字，则表示只打开 HTTP 服务；如果跟 **https** 关键字，则表示只打开 HTTPS 服务。

使用 **no enable service web-server** 或者 **default enable service web-server** 用于关闭相应的 HTTP 服务。如果该 **no 命令或 default 命令**后面不跟任何关键字，则表示关闭 HTTP 服务和 HTTPS 服务。

配置 HTTP 认证信息

【命令格式】 **webmaster level *privilege-level* username *name* password { *password* | [0 | 7] *encrypted-password* }**

【参数说明】 *privilege-level* : 用户绑定权限等级。

name : 用户名。

password : 用户口令。

0 | 7 : 口令的加密类型, 0 无加密, 7 简单加密。缺省为 0。

encrypted-password : 口令文本。

【命令模式】 全局模式

【使用指导】 在使用 HTTP Server 的时候, 需要进行登陆认证才能进入 Web 页面。使用该命令配置 Web 登陆认证的用户名和密码。

执行 **no webmaster level *privilege-level*** 删除指定权限等级的所有用户名与密码。

执行 **no webmaster level *privilege-level* username *name*** 删除指定用户名和密码。

 用户名和密码有三个权限等级; 每个权限等级最多可以配置 10 个用户名和密码。

 系统默认创建两个账号, admin 账号和 guest 账号, 这两个账号不可被删除, 只可修改密码。其中 guest 账号默认对应 level 2 权限, 在 web 端只拥有查看系统首页的权限。管理员账号为 admin 对应 level 0 权限, 在 web 端管理员账号拥有所有的功能, 并且可以编辑其他管理账号并授权该账号可访问的页面, 新添加的账号对应 level 1 权限。

配置 HTTP 服务端

【命令格式】 **http port *port-number***

【参数说明】 *port-number* : 设置 HTTP 服务端, 取值范围为 80 及 1025-65535。

【命令模式】 全局模式

【使用指导】 使用该命令设置 HTTP 服务的端口。

配置 HTTPS 服务端

【命令格式】 **http secure-port *port-number***

【参数说明】 *port-number* : 设置 HTTPS 服务端, 取值范围为 443 及 1025-65535。

【命令模式】 全局模式

【使用指导】 使用该命令设置 HTTPS 服务的端口。

配置举例

使用 Web 管理一台锐捷设备, 通过 Web 浏览器登陆到锐捷中进行相关功能的配置

- 使用默认配置的 admin 认证信息进行登录。
- 为了提高安全性, 要求 Web 浏览器即可以通过 HTTP 协议访问, 也可以通过 HTTPS 协议访问。
- 要求自己配置 HTTP 服务端, 减少非法用户对 HTTP 的攻击。

【网络环境】

图 5-5



【配置方法】

- 需要配置同时打开 HTTP 和 HTTPS 服务。
- 可以配置 HTTP 服务端口号为 8080；HTTPS 服务端口号为 4430。

A

```
A#configure terminal
A(config)# enable service web-server
A(config)# http port 8080
A(config)# http secure-port 4430
```

【检验方法】

查看 HTTP 配置信息。

A

```
A# show web-server status
http server status: enabled
http server port: 8080
https server status:enabled
https server port: 4430
```

常见错误

- 如果 HTTP 服务端口不是默认的 80 与 443，用户在浏览器中必须输入配置的具体服务端口，否则 web 端无法访问设备。

5.4.2 配置HTTP本地升级

配置效果

用户可以通过浏览器或者 upgrade web 命令升级。

注意事项

- 通过浏览器上传 Web 包，只要上传成功，并且版本校验通过，设备默认会直接升级为最新的 Web 包。
- 通过 **upgrade web download** 命令，自动从 tftp 服务器下载文件，并自动升级。
- 通过 **upgrade web** 命令，自动升级本地文件系统的 WEB 包。

配置方法

无

检验方法

- 用户直接通过浏览器访问，通过查看最新的 WEB 页面。

相关命令

从 TFTP 服务器下载 Web 文件包

【命令格式】 **upgrade web download { oob_tftp: path | tftp: path }**

【参数说明】 oob_tftp：表示通过 MGMT 口连接 tftp 服务器下载 WEB 包，仅支持 MGMT 口的设备支持该参数
tftp：表示通过普通数据口连接 tftp 服务器下载 WEB 包
path：tftp 服务器上 WEB 包的路径

【命令模式】 特权模式

【使用指导】 该命令是从 tftp server 端中下载 WEB 包，并自动升级。

从升级设备的 Web 文件包

【命令格式】 **upgrade web uri**

【参数说明】 uri：WEB 包存放的本地路径。

【命令模式】 特权模式

【使用指导】 该命令用于升级设备内存放的 WEB 包，并自动升级。

配置举例

用户通过官网获取到最新的 Web 包，希望设备运行最新的 Web 包

【网络环境】

图 5-6



- 【配置方法】
- 与本地 PC 机相连，PC 机的 IP 地址是 10.10.10.13；给设备配置一个同网段的 IP 地址 10.10.10.131。
 - 通过 web 登陆到设备中，并上传最新的 WEB 包到设备中。

A

```
A#configure terminal
A(config)# vlan 1
A(config-vlan)# exit
A(config)# interface vlan 1
A(config-VLAN 1)# ip address 10.10.10.131 255.255.255.0
A(config-VLAN 1)# exit
A(config)# enable service web-server
```

在 PC 机中，使用 WEB 页面的“本地升级”功能上传 WEB 包升级

【检验方法】 在 PC 机中，重新进行 Web 认证登陆，验证是否显示最新的 Web 页面。

通过 upgrade web download 方式升级 WEB 包

【网络环境】

图 5-7



【配置方法】

- 与本地 PC 机相连，PC 机的 IP 地址是 10.10.10.13；给设备配置一个同网段的 IP 地址 10.10.10.131。
- 打开 tftp 服务器。

A

```
A#configure terminal
A(config)# vlan 1
A(config-vlan)# exit
A(config)# interface vlan 1
A(config-VLAN 1)# ip address 10.10.10.131 255.255.255.0
A(config-VLAN 1)# end
A#upgrade web download tftp:// 10.10.10.13/web.upd
Press Ctrl+C to quit
!!!!!!!!!!
download 3896704 bytes
Begin to upgrade the web package...
Web package upgrade successfully.
```

【检验方法】 在 PC 机中，重新进行 Web 认证登陆，验证是否显示最新的 Web 页面。

通过 upgrade web 方式升级 WEB 包

【网络环境】

图 5-8



【配置方法】

- 与本地 PC 机相连，PC 机的 IP 地址是 10.10.10.13；给设备配置一个同网段的 IP 地址 10.10.10.131。
- 打开 tftp 服务器。

A

```
A#configure terminal
A(config)# vlan 1
A(config-vlan)# exit
A(config)# interface vlan 1
A(config-VLAN 1)# ip address 10.10.10.131 255.255.255.0
A(config-VLAN 1)# end
A#copy tftp://10.10.10.13/web.upd flash:/web.upd
Press Ctrl+C to quit
!!!!!!!!!!
Accessing tftp:// 10.10.10.13/web.upd finished, 3896704 bytes prepared
Flushing data to flash:/web.upd...
```

```
Flush data done
A #upgrade web flash:/web.upd
Web package upgrade successfully.
A #
```

【检验方法】 在 PC 机中，重新进行 Web 认证登陆，验证是否显示最新的 Web 页面。

常见配置错误

- 通过浏览器访问，发现没有更新到新的 WEB 包，可能是本地浏览器有缓存；将浏览器的缓存清空，在重新访问一次。

5.5 监视与维护

查看运行情况

作用	命令
查看 Web 服务配置信息和状态。	show web-server status

6 系统日志

6.1 概述

设备在运行过程中，会发生各种状态变化如链路状态 UP、DOWN 等，也会遇到一些事件如收到异常报文、处理异常等。锐捷产品系统日志提供一种机制，在状态变化或发生事件时，就自动生成固定格式的消息（日志报文），这些消息可以被显示在相关窗口（控制台、监视终端等）上或被记录在相关媒介（内存缓冲区、日志文件）上或发送到网络上的一组日志服务器上，供管理员分析网络情况和定位问题。同时为了方便管理员对日志报文的读取和管理，这些日志报文可以被打上时间戳和序号，并按日志信息的优先级进行分级。

i 下文仅介绍系统日志的相关内容。

协议规范

- RFC 3164 : The BSD syslog Protocol

6.2 典型应用

典型应用	场景描述
系统日志输出到控制台	通过控制台监控系统日志信息。
系统日志发送到日志服务器	通过服务器监控系统日志信息。

6.2.1 系统日志输出到控制台

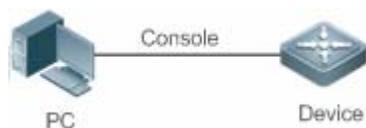
应用场景

可以将系统日志输出到控制台，方便管理员监控系统的运行状态，网络部署要求如下：

- 1、信息级别高于等于 informational（6 级）的日志信息允许输出到控制台。
- 2、只允许 ARP 模块和 IP 模块的日志信息输出到控制台。

组网环境如下所示：

图 6-1 系统日志输出到控制台组网图



功能部属

设备端的配置要点如下：

- 1、 设置允许输出到控制台的日志信息级别为 informational（6 级）。
- 2、 设置日志信息的过滤方向为：terminal（终端方向）。
- 3、 设置日志信息的过滤方式为：contains-only（“只包含”过滤方式）。
- 4、 设置日志信息的过滤规则为：“单个匹配”规则，模块名包含 ARP 或 IP。

6.2.2 系统日志发送到日志服务器

应用场景

可以将系统日志发送到日志服务器，方便管理员在服务器上统一监控设备的日志信息，假设网络中存在如下部署要求：

- 1、 系统日志信息发送到日志服务器上，日志服务器的 IP 地址为：10.1.1.1。
- 2、 信息级别高于等于 debugging（7 级）的所有模块的日志信息允许发送到日志服务器上。
- 3、 系统日志信息发送到日志服务器的报文源接口为 Loopback 0。

组网环境如下所示：

图 6-2 系统日志发送到日志服务器组网图



功能部属

设备端的配置要点如下：

- 1、 设置日志服务器 IPv4 地址：10.1.1.1。
- 2、 设置允许发送到服务器的日志信息级别为 debugging（7 级）。
- 3、 设置发往服务器的日志信息的源接口为 Loopback 0。

6.3 功能详解

基本概念

系统日志的分类

系统日志可以分为如下两类：

- log 类，日志类信息
- debug 类，调试类信息

系统日志的级别

系统日志按严重性划分为 8 个等级，严重性由高到底依次为：emergencies、alerts、critical、errors、warnings、notifications、informational 和 debugging，并分别对应于 0~7 这 8 个数值，值越小代表级别越高。

根据日志级别输出信息时，将会输出日志级别高于或等于所设置级别的日志，比如，输出规则中指定允许级别为 informational 的信息输出，则级别为 emergencies ~ informational 的信息均会输出。

相关日志级别的说明如下表所示：

关键字	级别	描述
emergencies	0	系统不能正常运行的信息
alerts	1	需要立即采取措施改正的信息
critical	2	严重情况
errors	3	错误信息
warnings	4	警告信息
notifications	5	普通但是需要关注的信息
informational	6	说明性的信息
debugging	7	调试信息

系统日志的输出方向

系统日志的输出方向，可以分为 5 类，分别为：console、monitor、server、buffer、file，各个输出方向上的缺省输出级别和输出的日志分类各不相同，用户在使用过程中，可以对不同的输出方向配置不同的过滤规则。

相关日志输出方向的说明如下表所示：

输出方向的名称	缺省输出方向	缺省输出级别	描述
console	控制台	debugging (7 级)	可以输出 log、debug 信息
monitor	监视终端	debugging (7 级)	可以输出 log、debug 信息，便于远程维护
server	日志服务器	informational (6 级)	可以输出 log、debug 信息
buffer	日志缓冲区	debugging (7 级)	可以输出 log、debug 信息，是设备运行过程中的一块缓存，用于记录系统日志
file	日志文件	informational (6 级)	可以输出 log、debug 信息，定时将缓存中的日志信息写入到文件中

RFC3164 日志格式

按照系统日志的输出方向不同，系统日志可能有不同格式。

- 当输出方向为非日志服务器（控制台、监视终端、日志缓冲区和日志文件）时，系统日志格式为：

```
seq no: *timestamp: sysname %module-level-mnemonic: content
```

对应的格式中文件说明如下：

序列号：*时间戳：系统名称 %模块名-级别-助记符：日志文本

例如，用户退出配置模式时，在控制台可以看到格式如下的日志：

```
001233: *May 22 09:44:36: Ruijie %SYS-5-CONFIG_I: Configured from console by console
```

● 当输出方向为日志服务器，系统日志格式为：

```
<priority>seq no: *timestamp: sysname %module-level-mnemonic: content
```

对应的格式中文件说明如下：

<优先级>序列号：*时间戳：系统名称 %模块名-级别-助记符：日志文本

例如，用户退出配置模式时，在日志服务器可以看到格式如下的日志：

```
<189>001233: *May 22 09:44:36: Ruijie %SYS-5-CONFIG_I: Configured from console by console
```

下面对每一个字段做详细说明：

4. priority (优先级)

本字段只有在向日志服务器输出日志时才有效。

优先级的计算按如下公式： $facility * 8 + level$ 。其中： $level$ 表示日志信息的级别； $facility$ 表示设备值，在设置日志信息的设备值时可以设置，默认值为 $local7 (23)$ ，参数取值范围如下表所示：

numerical code (标号)	facility keyword (设备值关键字)	facility description (设备值描述)
0	kern	kernel messages
1	user	user-level messages
2	mail	mail system
3	daemon	system daemons
4	auth1	security/authorization messages
5	syslog	messages generated internally by syslogd
6	lpr	line printer subsystem
7	news	network news subsystem
8	uucp	UUCP subsystem
9	clock1	clock daemon
10	auth2	security/authorization messages
11	ftp	FTP daemon
12	ntp	NTP subsystem
13	logaudit	log audit
14	logalert	log alert
15	clock2	clock daemon
16	local0	local use 0 (local0)
17	local1	local use 1 (local1)
18	local2	local use 2 (local2)
19	local3	local use 3 (local3)

20	local4	local use 4 (local4)
21	local5	local use 5 (local5)
22	local6	local use 6 (local6)
23	local7	local use 7 (local7)

5. seq no (序列号)

系统日志的序列号为 6 位整型数，并按系统日志产生的条目逐条递增，缺省情况下，该字段信息不会显示出来，可以通过命令开启或关闭该字段信息的输出。

6. timestamp (时间戳)

时间戳记录了系统日志产生的时间，方便用户查看和定位系统事件。锐捷设备的系统日志时间戳格式有两种，分别为 :datetime 格式和 uptime 格式。

- i** 如果当前设备不存在 RTC 时钟（一种用于记录系统绝对时间的硬件装置），缺省采用设备启动时间（uptime 格式）作为系统日志的时间戳。如果设备存在 RTC 时钟，则缺省采用设备绝对时间（datetime 格式）作为日志信息时间戳。

下面将对这两种时间戳格式进行详细说明：

- datetime 格式：

datetime 格式时间戳完整格式如下所示：

```
Mmm dd yyyy hh:mm:ss.msec
```

各个参数字段的说明如下表所示：

时间戳参数	参数名称	描述
Mmm	月份	Mmm 代表月份的英文缩写，1~12 月份依次为 :Jan、Feb、Mar、Apr、May、Jun、Jul、Aug、Sep、Oct、Nov、Dec
dd	天数	dd 代表当前月份对应的天数
yyyy	年份	yyyy 代表对应的年份，缺省情况下没有打开
hh	小时	hh 代表当前对应的小时数
mm	分钟	mm 代表当前对应的分钟数
ss	秒	ss 代表当前对应秒数
msec	毫秒	msec 代表当前对应的毫秒数

缺省情况下，系统输出的日志信息 datetime 格式时间戳不带年份和毫秒信息，用户可以通过命令开启或关闭系统日志的 datetime 格式时间戳是否携带年份和毫秒信息。

- uptime 格式：

uptime 格式时间戳完整格式如下所示：

```
dd:hh:mm:ss
```

整个时间戳字符串代表：系统自启机以来运行的天数：小时数：分钟数：秒数

7. sysname (系统名称)

该字段记录了生成该日志的设备名称，便于日志服务器标识该日志从哪个主机发送过来。缺省情况下，该字段信息不会显示出来，可以通过命令开启或关闭该字段信息的输出。

8. module (模块名)

该字段表示产生此日志的功能模块的名称，为一个 2~20 个字符的大写字符串（可包含大写字母、数字、下划线）。log 类的日志信息默认都要携带 module 字段，debug 类的日志信息有可能没有携带 module 字段。

9. level (日志级别)

系统日志的级别共分为 8 级，分别为 0~7 级。各模块生成的系统日志的级别在开发阶段已经确定，用户不能修改。

10. mnemonic (助记符)

该字段表示产生此日志的信息摘要，为一个 4~32 个字符的大写字符串（可包含大写字母、数字、下划线）。log 类的日志信息默认都要携带 mnemonic 字段，debug 类的日志信息有可能没有携带 mnemonic 字段。

11. content (日志文本)

该字段表示该系统日志的具体内容。

功能特性

功能特性	作用
系统日志功能开关	用于设置系统日志功能的打开与否。
系统日志格式设置	用于设置系统日志的显示格式。
系统日志信息设置	用于设置系统日志发往各个方向的参数信息。
系统日志过滤功能设置	用于设置系统日志过滤功能的参数信息。
系统日志上送功能设置	用于设置系统日志上送功能的参数信息
系统日志监控功能设置	用于设置系统日志监控功能的参数信息。

6.3.1 系统日志功能开关

用于设置系统日志功能的打开与否，主要包括：日志开关、重定向日志开关、日志信息统计功能开关。

相关配置

📌 打开日志开关

缺省情况下，日志开关是打开的。

使用 **logging on** 命令在全局配置模式下打开日志开关，打开日志开关后，系统产生的日志信息才能往各个输出方向输出，并用于监视系统的运行状态。

📌 重定向日志开关

VSU 环境下面，重定向日志开关默认是开启的。

使用 **logging rd on** 命令在全局配置模式下打开重定向日志开关，打开重定向日志开关后，在 VSU 环境下面，从机或从管理板的日志信息可以重定向到主机或主管理板输出，方便管理员进行日志信息统一管理。

📌 启用日志信息统计功能开关

缺省情况下，日志信息统计功能是关闭的。

使用 **logging count** 命令在全局配置模式下开启日志信息统计功能，打开日志信息统计功能后，系统将记录各模块产生的日志信息的次数，以及最后产生此日志信息的时间等。

6.3.2 系统日志格式设置

用于设置系统日志的显示格式，主要包括：RFC5424 日志格式、日志时间戳格式、日志系统名称、日志序列号等。

相关配置

▾ 启用日志信息时间戳开关

缺省情况下，系统日志使用的格式为 `datetime` 格式，且 `datetime` 时间戳格式没有携带年份和毫秒信息。

使用 **service timestamps** 命令在全局配置模式下打开系统日志的 `datetime` 格式的时间戳的年份和毫秒信息，或者将系统日志的格式修订成 `uptime` 格式。

▾ 启用日志信息系统名称开关

缺省情况下，系统输出的日志信息没有携带 `sysname`（系统名称）。

使用 **service sysname** 命令在全局配置模式下开启系统日志的 `sysname`（系统名称）。

▾ 启用日志信息序列号开关

缺省情况下，系统输出的日志信息没有携带序列号。

使用 **service sequence-numbers** 命令在全局配置模式下开启日志信息的序列号。

▾ 启用标准日志格式显示开关

缺省情况下，设备上面的日志信息显示格式如下：

```
*timestamp: %module-level-mnemonic: content
```

依次为：

```
*时间戳: %模块名-级别-助记符: 日志文本。
```

使用 **service standard-syslog** 命令在全局配置模式下开启标准日志格式显示开关，开启标准日志格式显示开关后，设备输出的日志信息显示格式如下：

```
timestamp %module-level-mnemonic: content
```

与缺省情况相比，标准日志格式的时间戳中前面少了一个 ‘ * ’、后面少了一个 ‘ : ’

▾ 启用私有日志格式显示开关

缺省情况下，设备上面的日志信息显示格式如下：

```
*timestamp: %module-level-mnemonic: content
```

依次是：

*时间戳: %模块名-级别-助记符: 日志文本。

使用 **service private-syslog** 命令在全局配置模式下开启私有日志格式显示开关, 开启私有日志格式显示开关后, 设备输出的日志信息显示格式如下:

```
timestamp module-level-mnemonic: content
```

与缺省情况相比, 私有日志格式的时间戳中前面少了一个 ' * '、后面少了一个 ' : ' , 模块名前面少了一个 '%'

6.3.3 系统日志信息设置

用于设置日志信息输出各个方向的参数信息, 主要包括: 日志信息输出控制台参数信息、日志信息输出监视终端参数信息、日志信息写入内存缓冲区参数信息、日志信息发往日志服务器参数信息、日志信息写入日志文件参数信息等。

相关配置

设置用户输入与日志信息输出同步

缺省情况下, 用户输入与日志信息输出功能是关闭的。

使用 **logging synchronous** 命令在线路配置模式下设置用户输入与日志信息输出同步功能, 防止用户正在输入字符时被打断。

设置日志信息速率控制功能

缺省情况下, 日志信息不进行速率限制。

使用 **logging rate-limit { number | all number | console { number | all number } } [except [severity]]**命令在全局配置模式下设置日志信息速率限制功能, 限制每秒内允许输出的日志信息。

设置重定向日志信息速率控制功能

缺省情况下, VSU 环境中, 限制从机重定向到主机的日志信息每秒最多 200 条。

使用 **logging rd rate-limit number [except severity]**命令在全局配置模式下设置重定向日志信息速率限制功能, 限制每秒内允许从机重定向到主机、从管理板重定向到主管理板的日志信息条目。

设置日志信息输出控制台的级别

缺省情况下, 日志信息输出到控制台的级别为 debugging (7 级)。

使用命令 **logging console [level]**命令在全局配置模式下设置允许在控制台上显示的日志信息级别。

设备允许日志信息输出到监视终端

缺省情况下, 日志信息不允许输出到监视终端。

使用命令 **terminal monitor** 命令在特权模式下设置允许将日志信息输出到监视终端。

设置日志信息输出到监视终端的级别

缺省情况下, 日志信息输出到监视终端的级别为 debugging (7 级)。

使用命令 **logging monitor** [*level*]命令在全局配置模式下设置允许在监视终端上输出的日志信息级别。

✎ 设置日志信息写入到内存缓冲区的参数

缺省情况下，日志信息默认会写入到内存缓冲区，且默认级别为 `debugging`（7级）。

使用 **logging buffered** [*buffer-size*] [*level*]命令在全局配置模式下设置日志写入的内存缓冲区的参数（包括缓冲区大小、日志信息等级）。

✎ 设置日志信息发送往日志服务器

缺省情况下，日志信息不会发往日志服务器。

使用 **logging server** [*oob*] { *ip-address* } [*via mgmt-name*] [**udp-port** *port*]命令在全局配置模式下设置日志发往指定的日志服务器。

✎ 设置日志信息发往日志服务器的级别

缺省情况下，日志信息发往日志服务器的级别为 `informational`（6级）。

使用命令 **logging trap** [*level*]命令在全局配置模式下设置允许发往日志服务器的日志信息级别。

✎ 设置日志信息发往日志服务器的设备值

在没有开启 RFC5424 日志格式的情况下，日志信息发往服务器的系统设备值默认为 `local7`（23）；在开启 RFC5424 日志格式的情况下，日志信息发往服务器的系统设备值默认为 `local0`（16）。

使用 **logging facility** *facility-type* 命令在全局配置模式下设置发往日志服务器的日志信息的系统设备值。

✎ 设置发往日志服务器的日志报文源地址

缺省情况下，发往 Syslog Server 的日志报文源地址为发送报文接口的 IP 地址。

使用 **logging source** [*interface*] *interface-type interface-number* 命令设置日志报文的源接口。倘若设备上未配置该源接口、或该源接口上未配置 IP 地址，则日志报文源地址也仍为发送报文接口的 IP 地址。

使用 **logging source** { *ip ip-address* }命令设置日志报文的源 IP 地址。倘若设备上未配置该 IP 地址，则日志报文源 IP 地址仍为发送报文接口的 IP 地址。

✎ 设置日志信息写入到日志文件参数

缺省情况下，日志信息不会写入日志文件中，开启日志信息写文件功能后，默认的级别为 `informational`（6级）。

使用 **logging file** { *flash:filename* | *usb0:filename* | *usb1:filename* } [*max-file-size*] [*level*]命令在全局配置模式下设置日志信息写入的日志文件参数（包括文件存储的设备类型、文件名称、文件大小、日志信息等级）。

✎ 设置日志信息写入到日志文件的时间间隔

缺省情况下，日志信息写入日志文件的时间间隔为 3600 秒（1小时）。

使用 **logging flash interval** *seconds* 命令在全局配置模式下设置日志信息写入日志文件的时间间隔。

✎ 设置日志信息写入到日志文件的保存时间

缺省情况下，系统对日志文件的保存时间是没有限制的。

使用 **logging life-time level level days** 命令在全局配置模式下设置日志信息的保存时间，方便管理员针对不同级别的日志信息指定不同的保存天数。

📌 设置将缓冲区当中的日志信息立即写入到日志文件中

缺省情况下，设备产生的日志信息会先缓存在系统日志缓冲区中，只有当缓冲区满或定时器到期后，才会将缓冲区中的日志信息写入到日志文件中。

使用 **logging flash flush** 命令在全局配置模式下将系统缓冲区中的日志信息立即写入到日志文件中，方便用户进行日志信息收集。

6.3.4 系统日志过滤功能设置

缺省情况下，系统打出来的日志信息都可以输出到各个方向，当某些情况下，用户可能不关心某些日志信息或者只关心某些日志信息，则可以使用日志过滤功能，对该日志信息进行过滤。

工作原理

📌 过滤方向

日志过滤方向主要分为以下四类：

- **buffer**：代表过滤掉去向日志缓冲区的日志信息（即 **show logging** 显示出来的日志信息）；
- **file**：代表过滤掉去向日志文件的日志信息；
- **server**：代表过滤掉去向日志服务器的日志信息；
- **terminal**：代表过滤掉去向控制台和监视终端（包括 Telnet/SSH 等）的日志信息；

以上四类过滤方向为或（|）关系，即可以联合使用（对往多个方向的日志信息进行过滤），也可以单独使用（只对往某一方向的日志信息进行过滤）。

📌 过滤方式

日志过滤方式主要分为以下两种：

- **contains-only**：代表“只包含”，意思是：只输出包含在过滤规则里面的关键字的日志信息，其它没有包含在过滤规则里面的关键字的日志信息不会输出。某些情况下，用户可能只关心某些日志信息是否产生，则可以在设备上面应用“只包含”这一日志过滤类型，让包含了此规则的日志信息才输出到终端界面，方便用于观察某些事件是否有发生。
- **filter-only**：代表“只过滤”，意思是：将过滤掉包含在过滤规则里面的关键字的日志信息，不会输出这些过滤掉的日志信息。某些情况下，当遇到某一个模块打出来的日志信息太多，可能会引起终端界面出现刷屏，且用户又不关心此类日志信息的时候，可以在设备上面应用“只过滤”这一日志过滤类型，并配置对应的过滤规则，将刷屏的日志信息过滤掉。

以上两种过滤方式为互斥关系，即同一时刻只能配置一种过滤方式。

📌 过滤规则

日志过滤规则主要分为以下两种：

- **exact-match** : 代表精确匹配, 若选择精确匹配, 则后面的三个过滤选项 (日志模块名、日志等级、日志助记符) 都需要选上。某些情况下, 用户可能只想过滤掉某一特定的日志信息, 则可以使用 “精确匹配” 规则。
- **single-match** : 代表单个匹配, 若选择单个匹配, 则后面的三个过滤选项 (日志模块名、日志等级、日志助记符) 只需要选择其中的一个。某些情况下, 用户可能想过滤掉某一类型的日志信息, 则可以使用 “单个匹配” 规则。

当用户配置的日志信息过滤规则中, 若 “单个匹配” 规则和 “精确匹配” 规则中同时配置了一样的模块名、助记符或信息等级, 则单个匹配规避的优先级高于精确匹配。

相关配置

📌 设置日志信息的过滤方向

缺省情况下, 日志信息的过滤方向为 `all`, 即过滤去往所有方向的日志信息。

使用 `logging filter direction { all | buffer | file | server | terminal }` 命令在全局配置模式下设置日志信息的过滤方向, 指定过滤去往哪几个方向的日志信息。

📌 设置日志信息的过滤方式

缺省情况下, 日志信息的过滤方式为 “只过滤”。

使用 `logging filter type { contains-only | filter-only }` 命令在全局配置模式下设置日志信息的过滤方式。

📌 设置日志信息的过滤规则

缺省情况下, 设备上面没有配置日志信息的过滤规则, 不对日志信息进行过滤。

使用 `logging filter rule exact-match module module-name mnemonic mnemonic-name level level` 命令在全局配置模式下设置日志信息的 “精确匹配” 过滤规则。

使用 `logging filter rule single-match { level level | mnemonic mnemonic-name | module module-name }` 命令在全局配置模式下设置日志信息的 “单个匹配” 过滤规则。

6.3.5 系统日志监控功能设置

打开日志监控功能后, 系统将对外界连接到设备的行为进行监控, 并记录对应的 LOG 信息。

工作原理

在设备上面开启记录用户登录或退出 LOG 信息后, 系统将对外界连接到设备的行为进行记录, 记录的信息包括: 登录的用户名、登录的源地址等。

在设备上面开启记录用户操作的 LOG 信息, 系统将对修改设备配置的行为进行记录, 记录的信息包括: 操作的用户名、操作的源地址、操作的内容。

相关配置

设置用户登录或退出 LOG 信息

缺省情况下，用户登录或退出设备的时候，设备是不会记录相关的 Log 信息。

使用 **logging userinfo** 命令在全局配置模式下设置用户登录/退出的 Log 信息。设置此功能后，当外界通过 Telnet、SSH、HTTP 等方式连接到设备时，设备将打出对应的 Log 信息，方便管理员监控设备的连接情况。

设置用户操作的 LOG 信息

缺省情况下，用户修订设备配置的时候，设备是不会记录相关的操作 Log 信息。

使用 **logging userinfo command-log** 命令在全局配置模式下设置用户操作的 Log 信息。设置此功能后，当有用户修改设备配置时，系统就会打出相应的 Log 信息提醒设备管理员。

6.4 产品说明



在卡式设备的 VSU 环境中，打开重定向日志开关后，从管理板或备份管理板的日志信息将重定向到主管理板进行输出，输出时，会在日志信息内容的最前面添加上对应的角色标志串“(*设备号/管理板名称)”，用于标识该日志信息是重定向日志信息。在 VSU 环境下面，若同时存在四块管理板，角色标志串可以形成四种形式：(*1/M1)、(*1/M2)、(*2/M1)、(*2/M2)。



logging buffered 命令 *buffer-size* 参数核心交换机取值范围从 4K 到 10M Bytes，默认值为 1M Bytes。

6.5 配置详解

配置项	配置建议&相关命令	
配置系统日志的显示格式	可选配置，用于设置系统日志的显示格式	
	service timestamps [<i>message-type</i> [<i>uptime</i> <i>datetime</i> [<i>msec</i>] [<i>year</i>]]]	设置系统日志的时间戳格式
	service sysname	设置系统日志格式中添加系统名称
	service sequence-numbers	设置系统日志格式中添加系列号
	service standard-syslog	设备系统日志格式为标准日志格式
	service private-syslog	设备系统日志格式为私有日志格式
配置系统日志输出到控制台	可选配置，用于设置系统日志输出到控制台的参数信息	
	logging on	打开日志开关
	logging count	打开日志信息统计功能
	logging console [<i>level</i>]	设置日志信息允许输出到控制台的级别
	logging rate-limit { <i>number</i> all number console { <i>number</i> all number } } [except [<i>severity</i>]]	设置日志信息速率限制功能

配置系统日志输出到监视终端	 可选配置，用于设置系统日志输出到监视终端的参数信息	
	terminal monitor	允许在当前监视终端上显示日志信息
	logging monitor [level]	设置日志信息允许输出到监视终端的级别
配置系统日志写入到内存缓冲区	 可选择配置，用于设置系统日志写入内存缓冲区的参数信息	
	logging buffered [buffer-size] [level]	设置日志写入的内存缓冲区的参数(包括缓冲区大小、日志信息等级)
配置系统日志发送给日志服务器	 可选配置，用于设置系统日志发送到日志服务器的参数信息	
	logging server [oob] { ip-address } [via mgmt-name] [udp-port port]	设置日志发往指定的日志服务器
	logging trap [level]	设置允许发往日志服务器的日志级别
	logging facility facility-type	设置发往服务器的日志信息的系统设备值
	logging source [interface] interface-type interface-number	设置发往服务器的日志信息的源接口
	logging source { ip ip-address }	设置发往服务器的日志信息的源地址
配置系统日志写入到日志文件	 可选配置，用于设置系统日志写入文件的参数信息	
	logging file { flash:filename usb0:filename usb1:filename } [max-file-size] [level]	设置日志信息写入的文件参数(包括文件存储的类型、文件名称、文件大小、日志信息等级)
	logging flash interval seconds	设置日志信息写入文件的频率,缺省值为 3600
	logging life-time level level days	设置日志信息写入文件的保存时间
配置系统日志过滤功能	 可选配置，用于设置系统日志的过滤功能参数信息	
	logging filter direction { all buffer file server terminal }	设置日志信息的过滤方向
	logging filter type { contains-only filter-only }	设置日志信息的过滤方式
	logging filter rule exact-match module module-name mnemonic mnemonic-name level level	设置日志信息的“精确匹配”过滤规则
配置系统日志重定向功能	 可选配置，用于设置系统日志的监控功能参数信息	
	logging rd on	打开日志重定向功能
	logging rd rate-limit number [except severity]	设置重定向日志信息速率限制功能
配置系统日志监控功能	 可选配置，用于设置系统日志的监控功能参数信息	

	logging userinfo	开启记录用户登录/退出的日志信息
	logging userinfo command-log	开启记录用户操作的日志信息
配置用户输入与日志信息同步输出功能	 可选配置，用于设置用户输入与日志信息同步输出功能	
	logging synchronous	设置用户输入与日志信息输出同步功能

6.5.1 配置系统日志的显示格式

配置效果

- 调整系统日志的显示格式。

注意事项

▾ RFC3164 日志格式

- 如果当前设备不存在 RTC 时钟（一种用于记录系统绝对时间的硬件装置），系统缺省采用设备启动时间（**uptime** 格式）作为日志信息时间戳，此时配置设备时间无效，如果设备存在 RTC 时钟，则缺省采用设备时间（**datetime** 格式）作为日志信息时间戳。
- 日志序列号是一个长整型数值，每产生一条日志，序列号就递增，但是由于日志序列号只显示 6 位整数，故当序列号从 1 开始每增加到 1000000 或序列号到达 2^{32} 时候就会发生一次翻转，即序列号又从 000000 开始显示。

配置方法

▾ 设置系统日志的时间戳格式

- 可选配置，缺省情况下系统日志的时间戳采用 **datetime** 格式。
- 若无特殊要求，在需要设置系统日志时间戳格式的设备上面配置。

▾ 设置系统日志格式中添加系统名称

- 可选配置，缺省情况下系统日志的格式中没有添加系统名称。
- 若无特殊要求，在需要为日志格式中添加系统名称的设备上面配置。

▾ 设置系统日志格式中添加系列号

- 可选配置，缺省情况下系统日志的格式中没有添加系列号。
- 若无特殊要求，在需要为日志格式添加系列号的设备上面配置。

▾ 设置系统日志格式为标准日志格式

- 可选配置，缺省情况下系统日志的格式中为默认格式。
- 若无特殊要求，在需要使用标准日志格式的设备上面配置。

设置系统日志格式为私有日志格式

- 可选配置，缺省情况下系统日志的格式中为默认格式。
- 若无特殊要求，在需要使用私有日志格式的设备上面配置。

检验方法

- 通过触发系统产生一条日志信息，用于查看设置后的系统日志的显示格式。

相关命令

设置系统日志的时间戳格式

【命令格式】 **service timestamps** [*message-type* [**uptime** | **datetime** [**msec**] [**year**]]]

【参数说明】 *message-type*：日志类型，有两种 log 和 debug

uptime：设备启动时间，格式：*天*小时*分*秒，例：07:00:10:41

datetime：当前设备日期，格式：月 日期 时：分：秒，例：Jul 27 16:53:07

msec：当前设备日期支持毫秒显示

year：当前设备日期支持年份显示

【命令模式】 全局配置模式

【使用指导】 系统日志的时间戳格式有两种：设备启动时间(**uptime**)格式或者设备日期(**datetime**)格式，用户可以根据需要选择不同类型的时间戳格式。

设置系统日志格式中添加系统名称

【命令格式】 **service sysname**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 可以在日志信息中系统名称，加上系统名称以后，系统日志发送到服务器后，在服务器上，可以清楚地知道日志信息来自哪个设备。

设置系统日志格式中添加序列号

【命令格式】 **service sequence-numbers**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 可以在日志信息中加上序列号，序列号从 1 开始。加上序号以后，就可以非常清楚地知道日志信息有没有丢失，以及日志产生的先后顺序。

设置系统日志格式为标准日志格式

【命令格式】 **service standard-syslog**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 默认情况下，设备上面的日志信息显示格式如下（默认格式）：

```
*timestamp: %module-level-mnemonic: content
```

依次是：

*时间戳: %模块名-级别-助记符: 日志文本。

若打开标准日志格式显示功能，设备上面的日志信息显示格式如下：

```
timestamp %module-level-mnemonic: content
```

与缺省情况相比，标准日志格式的时间戳中前面少了一个 ‘ * ’、后面少了一个 ‘ : ’

设置系统日志格式为私有日志格式

【命令格式】 **service private-syslog**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 默认情况下，设备上面的日志信息显示格式如下（默认格式）：

```
*timestamp: %module-level-mnemonic: content
```

依次是：

*时间戳: %模块名-级别-助记符: 日志文本。

若打开标准日志格式显示功能，设备上面的日志信息显示格式如下：

```
timestamp module-level-mnemonic: content
```

与缺省情况相比，私有日志格式的时间戳中前面少了一个‘*’、后面少了一个‘:’，模块名前面少了一个‘%’

配置举例

配置 RFC3164 日志显示格式

【网络环境】 假设网络环境中，有以下日志时间戳格式设置要求：

- 1、切换日志格式为 RFC3164 格式；
- 2、日志时间戳格式调整为 **datetime** 格式，并且开启毫秒信息和年份信息的显示；
- 3、日志时间戳格式中要求添加系统名称；
- 4、日志时间戳格式中要求添加序列号。

【配置方法】 ● 在设备上面配置系统日志的显示格式

```
Ruijie# configure terminal
Ruijie(config)# no service log-format rfc5424
Ruijie(config)# service timestamps log datetime year msec
Ruijie(config)# service timestamps debug datetime year msec
Ruijie(config)# service sysname
Ruijie(config)# service sequence-numbers
```

【检验方法】 用户设置了日志时间戳格式后，在系统新产生日志信息的时候，将会依据所设置的时间戳格式进行日志信息的构造和输出。

- 通过 **show logging config** 命令可以查看用户配置的相关参数信息。
- 通过进入/退出全局配置模式触发产生一条新的日志信息，可以观察新产生的日志信息的时间戳格式。

```
Ruijie(config)#exit
```

```
001302: *Jun 14 2013 19:01:40.293: Ruijie %SYS-5-CONFIG_I: Configured from console by admin on console
Ruijie#show logging config
Syslog logging: enabled
  Console logging: level informational, 1306 messages logged
  Monitor logging: level informational, 0 messages logged
  Buffer logging: level informational, 1306 messages logged
  File logging: level informational, 121 messages logged
  File name:syslog_test.txt, size 128 Kbytes, have written 5 files
  Standard format:false
  Timestamp debug messages: datetime
  Timestamp log messages: datetime
  Sequence-number log messages: enable
  Sysname log messages: enable
  Count log messages: enable
  Trap logging: level informational, 121 message lines logged,0 fail
```

6.5.2 配置系统日志输出到控制台

配置效果

- 可以将系统产生的日志信息输出到控制台，方便管理员监控系统的运行状态。

注意事项

- 如果系统产生的日志信息太多，则可以通过限制日志信息的速率来减少输出到控制台日志信息。

配置方法

▾ 打开日志开关

- 可选配置，缺省情况下系统日志开关已经打开。

▾ 打开日志信息统计功能

- 可选配置，缺省情况下系统日志信息统计功能是关闭的。
- 若无特殊要求，在需要打开日志信息统计功能的设备上面配置。

▾ 设置日志信息允许输出控制台的级别

- 可选配置，缺省级别为 debugging（7级）。
- 若无特殊要求，在需要设置日志信息允许输出控制台级别的设备上面配置。

设置日志信息速率限制功能

- 可选配置，缺省情况下不进行速率限制。
- 若无特殊要求，在需要设置日志信息速率限制功能的设备上配置。

检验方法

- 通过 `show logging config` 命令可以查看设置的允许输出控制台的日志级别参数。

相关命令

打开日志开关

- 【命令格式】 `logging on`
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 缺省情况下，系统日志开关是打开的，一般情况下，不要关闭日志开关，如果觉得打印的信息太多，可以通过设置不同设备日志信息的显示级别来减少日志信息的打印。

打开日志信息统计功能

- 【命令格式】 `logging count`
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 缺省情况下，系统日志信息统计功能是关闭的。启用了日志报文统计功能后，从命令打开时将系统中输出的日志信息进行分类统计，主要记录日志信息的产生次数，以及最后产生的时间等。

设置日志信息允许输出到控制台的级别

- 【命令格式】 `logging console [level]`
- 【参数说明】 `level`：日志信息的级别
- 【命令模式】 全局配置模式
- 【使用指导】 控制台默认允许显示的日志信息级别为 `debugging`（7级）。可以通过特权命令 `show logging config` 来查看允许在控制台上显示的日志信息级别。

设置日志信息速率限制功能

- 【命令格式】 `logging rate-limit { number | all number | console { number | all number } } [except [severity]]`
- 【参数说明】
 - `number`：每秒钟内允许处理的日志信息，范围为 1~10000。
 - `all`：设置对所有的日志信息进行速率控制，包括 0~7 级所有日志信息。
 - `console`：设置每秒钟内允许在控制台上显示的日志信息数。
 - `except severity`：小于等于此严重性级别的日志信息，不进行速率控制；默认级别为 `error(3)`，对小于等于 `error` 级别的日志信息不进行速率控制。
- 【命令模式】 全局配置模式
- 【使用指导】 默认情况下，不对日志信息进行速率限制。

配置举例

配置系统日志输出到控制台

- 【网络环境】 假设网络环境中，有以下日志输出控制台格式要求：
- 1、打开日志信息统计功能；
 - 2、设置允许输出到控制台的日志信息级别为 informational（6级）；
 - 3、设置日志信息输出到控制台的速率为每秒 50 条；

- 【配置方法】 ● 在设备上面配置系统日志输出到控制台

```
Ruijie# configure terminal
Ruijie(config)# logging count
Ruijie(config)# logging console informational
Ruijie(config)# logging rate-limit console 50
```

- 【检验方法】 ● 通过 **show logging config** 命令可以查看用户配置的相关参数信息。

```
Ruijie(config)#show logging config
Syslog logging: enabled
  Console logging: level informational, 1303 messages logged
  Monitor logging: level debugging, 0 messages logged
  Buffer logging: level debugging, 1303 messages logged
  File logging: level informational, 118 messages logged
  File name:syslog_test.txt, size 128 Kbytes, have written 5 files
  Standard format:false
  Timestamp debug messages: datetime
  Timestamp log messages: datetime
  Sequence-number log messages: enable
  Sysname log messages: enable
  Count log messages: enable
  Trap logging: level informational, 118 message lines logged,0 fail
```

6.5.3 配置系统日志输出到监视终端

配置效果

- 可以将系统产生的日志信息输出到远程监视终端，方便管理员监控系统的运行状态。

注意事项

- 如果系统产生的日志信息太多，则可以通过限制日志信息的速率来减少输出到监视终端的日志信息。

- 默认情况下，用户远程连接到设备后，当前监视终端上不允许输出日志信息。需要手动输入 **terminal monitor** 命令开启当前终端的日志信息输出功能。

配置方法

允许在当前监视终端上显示日志信息

- 必选配置，缺省情况下不允许在监视终端上显示日志信息。
- 若无特殊要求，应在每个连接到设备的监视终端配置。

设置日志信息允许输出到监视终端的级别

- 可选配置，缺省级别为 debugging（7 级）。
- 若无特殊要求，在需要设置日志信息允许输出到监视终端级别的设备上面配置。

检验方法

- 通过 **show logging config** 命令可以查看设置的允许输出到监视终端的日志级别参数。

相关命令

允许在当前监视终端上显示日志信息

【命令格式】 **terminal monitor**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 默认情况下，用户远程连接到设备后，当前监视终端上不允许输出日志信息。需要手动输入 **terminal monitor** 命令开启当前终端的日志信息输出功能。

设置日志信息允许输出到监视终端的级别

【命令格式】 **logging monitor [level]**

【参数说明】 *level*：日志信息的级别

【命令模式】 全局配置模式

【使用指导】 监视终端默认允许显示的日志信息级别为 debugging（7 级）。
可以通过特权命令 **show logging config** 来查看允许在监视终端上显示的日志信息级别。

配置举例

配置系统日志输出到监视终端

- 【网络环境】 假设网络环境中，有以下日志信息输出到监视终端设置要求：
- 1、设置允许在监视终端上显示日志信息；
 - 2、设置允许输出到控制台的日志信息级别为 informational（6 级）。

- 【配置方法】
- 在设备上面配置系统日志输出到监视终端

```
Ruijie# configure terminal
Ruijie(config)# logging monitor informational
Ruijie(config)# line vty 0 4
Ruijie(config-line)# monitor
```

- 【检验方法】
- 通过 **show logging config** 命令可以查看用户配置的相关参数信息。

```
Ruijie#show logging config
Syslog logging: enabled
  Console logging: level informational, 1304 messages logged
  Monitor logging: level informational, 0 messages logged
  Buffer logging: level debugging, 1304 messages logged
  File logging: level informational, 119 messages logged
  File name:syslog_test.txt, size 128 Kbytes, have written 5 files
  Standard format:false
  Timestamp debug messages: datetime
  Timestamp log messages: datetime
  Sequence-number log messages: enable
  Sysname log messages: enable
  Count log messages: enable
  Trap logging: level informational, 119 message lines logged,0 fail
```

常见错误

- 若要取消当前终端的日志信息输出功能，需要使用的命令是：**terminal no monitor**，而不是 **no terminal monitor**。

6.5.4 配置系统日志写入到内存缓冲区

配置效果

- 可以将系统产生的日志信息写入到内存缓冲区，方便管理员通过 **show logging** 命令查看近期系统产生的日志信息。

注意事项

- 系统日志写入内存缓冲区后，当缓冲区满时，将循环覆盖重写。

配置方法

- ▾ 设置日志写入的内存缓冲区的参数

- 可选配置，缺省情况下系统会将日志信息写入到内存缓冲区，且默认级别为 debugging（7 级）。
- 若无特殊要求，在需要设置日志写入内存缓冲区级别的设备上配置。

检验方法

- 通过 **show logging config** 命令可以查看设置的允许写入内存缓冲区的日志级别参数。
- 通过 **show logging** 命令可以查看系统写入内存缓冲区的日志信息。

相关命令

设置日志写入的内存缓冲区的参数

【命令格式】 **logging buffered** [*buffer-size*] [*level*]

【参数说明】 *buffer-size* : 内存缓冲的大小

level : 允许写入到内存缓冲区的信息级别

【命令模式】 全局配置模式

【使用指导】 默认写入内存缓冲区的日志信息级别为 debugging（7 级）。

可以通过特权命令 **show logging** 来查看允许写入内存缓冲区的日志信息级别和缓冲的大小等参数信息。

配置举例

配置系统日志写入到内存缓冲区的参数

【网络环境】 假设网络环境中，有以下日志信息写入到内存缓冲区设置要求：

- 1、设置日志内存缓冲区的大小为 128K（131072 字节）；
- 2、设置允许写入到内存缓冲区的日志信息级别为 informational（6 级）。

【配置方法】 ● 在设备上配置系统日志写入到内存缓冲区参数信息

```
Ruijie# configure terminal
Ruijie(config)# logging buffered 131072 informational
```

【检验方法】 ● 通过 **show logging** 命令可以查看用户配置的相关参数信息及系统最近产生的日志信息。

```
Ruijie#show logging
Syslog logging: enabled
  Console logging: level informational, 1306 messages logged
  Monitor logging: level informational, 0 messages logged
  Buffer logging: level informational, 1306 messages logged
  File logging: level informational, 121 messages logged
  File name:syslog_test.txt, size 128 Kbytes, have written 5 files
  Standard format:false
  Timestamp debug messages: datetime
```

- 【网络环境】 假设网络环境中，有以下日志信息写入到内存缓冲区设置要求：
- 1、设置日志内存缓冲区的大小为 128K (131072 字节) ；
 - 2、设置允许写入到内存缓冲区的日志信息级别为 informational (6 级) 。

- 【配置方法】 ● 在设备上面配置系统日志写入到内存缓冲区参数信息

```
Ruijie# configure terminal
Ruijie(config)# logging buffered 131072 informational
```

- 【检验方法】 ● 通过 **show logging** 命令可以查看用户配置的相关参数信息及系统最近产生的日志信息。

```
Timestamp log messages: datetime
Sequence-number log messages: enable
Sysname log messages: enable
Count log messages: enable
Trap logging: level informational, 121 message lines logged, 0 fail
Log Buffer (Total 131072 Bytes): have written 4200
001301: *Jun 14 2013 19:01:09.488: Ruijie %SYS-5-CONFIG_I: Configured from console by admin on console
001302: *Jun 14 2013 19:01:40.293: Ruijie %SYS-5-CONFIG_I: Configured from console by admin on console
// 这里省略其它日志信息，客户 show logging 时以实际为准。
```

6.5.5 配置系统日志发送往日志服务器

配置效果

- 可以将系统产生的日志信息发送往日志服务器，方便管理员在服务器上统一监控设备的日志信息。

注意事项

- 如果设备上面具有 MGMT 接口，且网络环境当中设备是通过 MGMT 口连接到日志服务器，则在配置 logging server 时，需要添加 oob 选项（代表 Syslog 报文走 MGMT 接口发送到日志服务器）。
- 要将日志信息发送给日志服务器，必须打开日志信息的时间戳开关或序列号开关，否则日志信息将不会发给日志服务器。

配置方法

📌 设置日志发往指定的日志服务器

- 必选配置，缺省情况下系统产生的日志信息不会发送日志服务器。
- 若无特殊要求，应在每台设备上面配置。

设置日志信息允许发往日志服务器的级别

- 可选配置，缺省情况下系统发往日志服务器的日志级别为 informational（6 级）。
- 若无特殊要求，在需要设置日志信息允许发往日志服务器级别的设备上面配置。

设置发往服务器的日志信息的系统设备值

- 可选配置，在没有开启 RFC5424 日志格式的情况下，日志信息发往服务器的系统设备值默认为 local7（23）；在开启 RFC5424 日志格式的情况下，日志信息发往服务器的系统设备值默认为 local0（16）。
- 若无特殊要求，在需要设置发往服务器的日志信息的系统设备值的设备上面配置。

设置发往服务器的日志信息的源接口

- 可选配置，缺省情况下发往日志服务器的日志报文源地址为发送报文接口的 IP 地址。
- 若无特殊要求，在需要设备发往服务器的日志信息的源接口的设备上面配置。

设置发往服务器的日志信息的源地址

- 可选配置，缺省情况下发往日志服务器的日志报文源地址为发送报文接口的 IP 地址。
- 若无特殊要求，在需要设置发往服务器的日志信息的源地址的设备上面配置。

检验方法

- 通过 `show logging config` 命令可以查看设置的日志服务器参数信息

相关命令

设置日志发往指定的日志服务器

【命令格式】 `logging server [oob] { ip-address } [via mgmt-name] [udp-prot port]`
或 `logging { ip-address } [udp-prot port]`

【参数说明】 `oob`：将日志服务器指定为带外通信（一般指通过 MGMT 口发往日志服务器）

`ip-address`：接收日志信息的主机 IP 地址

`via mgmt-name`：指定日志主机在 `oob` 选项时使用的 MGMT 口

`udp-prot port`：指定日志主机的端口号（默认端口号为 514）

【命令模式】 全局配置模式

【使用指导】 该命令用于指定接收日志信息的日志服务器地址，可以同时指定多个日志服务器，日志信息将被同时分给配置的所有的日志服务器。

 在命令中启用 `oob` 参数时，`via` 参数才可以被指定使用。

 锐捷产品允许配置最多 5 个日志服务器。

设置日志信息允许发往日志服务器的级别

【命令格式】 `logging trap [level]`

- 【参数说明】 *level* : 日志信息的级别
- 【命令模式】 全局配置模式
- 【使用指导】 默认发送往日志服务器的日志信息级别为 informational (6 级)。
可以通过特权命令 **show logging config** 来查看允许发送往日志服务器的级别。

📌 设置发往服务器的日志信息的系统设备值

- 【命令格式】 **logging facility** *facility-type*
- 【参数说明】 *facility-type* : 日志信息设备值
- 【命令模式】 全局配置模式
- 【使用指导】 在没有开启 RFC5424 日志格式的情况下, 日志信息发往服务器的系统设备值默认为 local7 (23); 在开启 RFC5424 日志格式的情况下, 日志信息发往服务器的系统设备值默认为 local0 (16)。

📌 设置发往服务器的日志信息的源接口

- 【命令格式】 **logging source [interface]** *interface-type interface-number*
- 【参数说明】 *interface-type* : 接口类型
interface-number : 接口编号
- 【命令模式】 全局配置模式
- 【使用指导】 默认情况下, 发送给服务器的日志报文源 IP 地址是报文发送接口的 IP 地址。
为了便于跟踪管理, 可以使用该命令将所有日志报文的源 IP 地址固定为某个接口的 IP 地址, 这样管理员就通过唯一地址识别从哪台设备发送出来的日志报文, 倘若设备上未配置该源接口或源接口上未配置 IP 地址, 则日志报文源 IP 地址仍为报文发送接口的 IP 地址。

📌 设置发往服务器的日志信息的源地址

- 【命令格式】 **logging source { ip** *ip-address* }
- 【参数说明】 **ip** *ip-address* : 指定向 IPV4 日志主机发送日志报文的源 IPV4 地址
- 【命令模式】 全局配置模式
- 【使用指导】 默认情况下, 发送给 Syslog Server 的日志报文源 IP 地址是报文发送接口的 IP 地址。
为了便于跟踪管理, 可以使用该命令将所有日志报文的源 IP 地址固定为某个 IP 地址, 这样管理员就通过唯一地址识别从哪台设备发送出来的日志报文, 倘若设备上未配置该 IP 地址, 则日志报文源 IP 地址仍为报文发送接口的 IP 地址。

配置举例

📌 配置系统日志发送往日志服务器

- 【网络环境】 假设网络环境中, 有以下日志信息发送往日志服务器设置要求:
- 1、设置日志服务器 IPv4 地址: 10.1.1.100;
 - 2、设置允许发送到日志服务器的日志信息级别为 debugging (7 级);
 - 3、设置发往日志服务器的日志信息的源接口为 Loopback 0。
- 【配置方法】
- 在设备上面配置系统日志发送往日志服务器

```
Ruijie# configure terminal
Ruijie(config)# logging server 10.1.1.100
Ruijie(config)# logging trap debugging
Ruijie(config)# logging source interface Loopback 0
```

- 【检验方法】
- 通过 **show logging config** 命令可以查看用户配置的相关参数信息。

```
Ruijie#show logging config
Syslog logging: enabled
  Console logging: level informational, 1307 messages logged
  Monitor logging: level informational, 0 messages logged
  Buffer logging: level informational, 1307 messages logged
  File logging: level informational, 122 messages logged
  File name:syslog_test.txt, size 128 Kbytes, have written 5 files
  Standard format:false
  Timestamp debug messages: datetime
  Timestamp log messages: datetime
  Sequence-number log messages: enable
  Sysname log messages: enable
  Count log messages: enable
  Trap logging: level debugging, 122 message lines logged,0 fail
    logging to 10.1.1.100
```

6.5.6 配置系统日志写入到日志文件

配置效果

- 可以将系统产生的日志信息按指定的频率写入到日志文件，便于管理员在设备本地随时查看历史日志信息。

注意事项

- 系统产生的日志信息是先缓冲到内存缓冲区当中，然后当缓冲区的时候或定时（默认为间隔 1 小时）写入到日志文件的，并不是产生日志信息的时候就立即写入到日志文件当中。

配置方法

📄 设置日志信息写入的日志文件参数

- 必选配置，缺省情况下系统产生的日志信息不会写入日志文件中。
- 若无特殊要求，应在每台设备上配置。

📄 设置日志信息写入文件的时间间隔

- 可选配置，缺省情况下系统日志写入到文件的时间间隔为每小时写一次。
- 若无特殊要求，在需要设置日志信息写入文件的时间间隔的设备上面配置。

设置日志信息写入文件的保存时间

- 可选配置，缺省情况下系统对日志文件的保存时间是没有限制的。
- 若无特殊要求，在需要设备日志信息写入文件的保存时间的设备上面配置。

设置将缓冲区当中的日志信息立即写入到日志文件中

- 可选配置，缺省情况下设备产生的日志信息会先缓存在系统日志缓冲区中，只有当缓冲区满或定时器到期后，才会将缓冲区中的日志信息写入到日志文件中。
- 若无特殊要求，应在用户收集日志文件的时候进行配置，且该命令配置一次作用一次，配置后立即将存在缓冲区中的日志信息写入到日志文件中。

检验方法

- 通过 `show logging config` 命令可以查看设置的日志服务器参数信息

相关命令

设置日志信息写入的日志文件参数

【命令格式】 `logging file { flash:filename | usb0:filename | usb1:filename } [max-file-size] [level]`

【参数说明】 `flash`：日志文件选择保存在扩展 FLASH 当中。

`usb0`：日志文件选择保存在 USB0 当中，此选项需要设备具有 1 个 USB 接口时才支持，并插入扩展的 USB 设备。

`usb1`：日志文件选择保存在 USB1 当中，此选项需要设备具有 2 个 USB 接口时才支持，并插入扩展的 USB 设备。

`max-file-size`：日志文件的最大值。从 128K 到 6M bytes，缺省大小为 128K。

`level`：允许写入到日志文件的信息级别。

【命令模式】 全局配置模式

【使用指导】 该命令将在指定的文件存储设备上根据指定的文件名创建文件用于储存日志，文件大小会随日志增加而增加，但其上限以配置的 `max-file-size` 为准，若没有指定 `max-file-size`，则日志文件的大小默认为 128K。

配置该命令后，系统将日志信息保存到文件中，日志文件名不要带文件类型的后缀名。日志文件后缀为固定为 `txt` 类型，配置文件后缀名将被拒绝。

配置了日志写文件功能后，日志信息将间隔 1 小时，写入到文件当中，而日志文件的名称（假设此次已经配置：`logging file flash:syslog`）依次为 `syslog.txt`、`syslog_1.txt`、`syslog_2.txt`..... `syslog_14.txt`、`syslog_15.txt` 总共 16 个日志文件。这 16 个日志文件循环重写 比如 写完 `syslog.txt` 后 写 `syslog_1.txt` 直至 `syslog_15.txt`，然后再返回来写 `syslog.txt`，这样子循环重写。

设置日志信息写入文件的时间间隔

【命令格式】 `logging flash interval seconds`

- 【参数说明】 *seconds* : 日志信息写入到 FLASH 文件的时间间隔, 范围: 1~51840, 单位: 秒
- 【命令模式】 全局配置模式
- 【使用指导】 通过此命令设置日志信息保存到文件中的时间间隔, 且从命令配置后开始计时。

设置日志信息写入文件的保存时间

- 【命令格式】 **logging life-time level level days**
- 【参数说明】 *level* : 日志信息的级别。
days : 日志信息保存时间。单位: 天。保存时间不小于 7 天。
- 【命令模式】 全局配置模式
- 【使用指导】 用户开启了基于时间的日志保存功能, 系统针对同一级别、同一天内产生的日志信息, 写入到同一个日志文件中, 日志文件的名称形如 “yyyy-mm-dd_filename_level.txt”, 其中: yyyy-mm-dd 为日志信息产生的当天绝对时间; filename 为 **logging file flash** 命令配置的日志文件名称, level 为对应的日志信息级别。
用户对某个等级的日志信息进行保存时间限制后, 当对应级别的日志信息超过日志保存时间限制后, 将进行删除。为了网管的方便, 目前系统要求日志信息最少可以保存 7 天, 最长可以保存 365 天。
为了兼容以前的配置命令, 用户在没有开启基于时间的日志保存功能时, 日志仍然基于文件大小进行日志信息的保存。

设置将缓冲区当中的日志信息立即写入到日志文件中

- 【命令格式】 **logging flash flush**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 在系统开启日志信息写日志文件功能后, 设备产生的日志信息会先缓存在系统日志缓冲区中, 只有当缓冲区满或定时器到期后, 才会将缓冲区中的日志信息写入到日志文件中, 可以通过该命令设置将系统缓冲区中的日志信息立即写入到日志文件中。

 用户配置 **logging flash flush** 命令时, 配置一次作用一次, 配置后立即将存在缓冲区中的日志信息写入到日志文件中

配置举例

配置系统日志写入到日志文件

- 【网络环境】 假设网络环境中, 有以下日志信息写入到日志文件设置要求:
- 1、设置日志文件名称为 *syslog* ;
 - 2、设置允许输出到控制台的日志信息级别为 *debugging* (7 级) ;
 - 3、设备日志信息写入到文件的时间间隔为 10 分钟 (600 秒) 。

- 【配置方法】 ● 在设备上面配置系统日志写入到日志文件

```
Ruijie# configure terminal
Ruijie(config)# logging file flash:syslog debugging
Ruijie(config)# logging flash interval 600
```

- 【检验方法】
- 通过 **show logging config** 命令可以查看用户配置的相关参数信息。

```
Ruijie(config)#show logging config
Syslog logging: enabled
  Console logging: level informational, 1307 messages logged
  Monitor logging: level informational, 0 messages logged
  Buffer logging: level informational, 1307 messages logged
  File logging: level debugging, 122 messages logged
  File name:syslog.txt, size 128 Kbytes, have written 1 files
  Standard format:false
  Timestamp debug messages: datetime
  Timestamp log messages: datetime
  Sequence-number log messages: enable
  Sysname log messages: enable
  Count log messages: enable
  Trap logging: level debugging, 122 message lines logged,0 fail
    logging to 10.1.1.100
```

6.5.7 配置系统日志过滤功能

配置效果

- 在某些情况下，管理员可能不想让某些日志信息显示出来，则可以通过此功能过滤系统产生的日志信息。
- 默认情况下，各个模块打出来的日志信息都可以显示到控制台或其它终端上面。设置日志信息过滤原则可以让某些日志信息打出到某些终端中，或者只想让某些日志信息打出到某些终端中。

注意事项

- 日志信息的两种过滤类型，分为：“只包含”和“只过滤”，某一时刻只能配置其中的一种类型。
- 当用户配置的日志信息过滤规则中，若单个匹配规则和精确匹配规则中同时配置了一样的模块名、助记符或信息等级，则单个匹配规则的优先级高于精确匹配。

配置方法

▾ 设置日志信息的过滤方向

- 可选配置，缺省情况下过滤方向为 all（即过滤所有方向的日志信息）。
- 若无特殊要求，在需要设置日志信息的过滤方向的设备上配置。

▾ 设置日志信息的过滤方式

- 可选配置，缺省情况下日志过滤方式为“只过滤”。

- 若无特殊要求，在需要设置日志信息的过滤方式的设备上面配置。

设置日志信息的过滤规则

- 必选配置，缺省情况下，系统没有设置任何过滤规则，不对日志信息进行过滤。
- 若无特殊要求，在需要设置日志信息的过滤规则的设备上面配置。

检验方法

- 通过 `show running` 命令可以查看设置的日志过滤功能参数信息

相关命令

设置日志信息的过滤方向

【命令格式】 `logging filter direction { all | buffer | file | server | terminal }`

【参数说明】 `all`：代表过滤往所有方向的日志信息。

`buffer`：代表过滤往日志缓冲区的日志信息（即 `show logging` 显示出来的日志信息）；

`file`：代表只过滤往日志文件的日志信息；

`server`：代表只过滤往日志服务器的日志信息；

`terminal`：代表过滤往控制台和 VTY 终端（包括 Telnet/SSH 等）的日志信息。

【命令模式】 全局配置模式

【使用指导】 默认为 `all`，即过滤所有方向的日志信息。

`default logging filter direction` 命令恢复日志信息的过滤方向为 `all`。

设置日志信息的过滤方式

【命令格式】 `logging filter type { contains-only | filter-only }`

【参数说明】 `contains-only` 代表“只包含”，意思是：只输出包含了过滤规则里面的关键字的日志信息，其它没有包含过滤规则里面的关键字的日志信息不会输出；

`filter-only` 代表“只过滤”，意思是：将过滤掉包含了过滤规则里面的关键字的日志信息，不会输出这些过滤掉的日志信息。

【命令模式】 全局配置模式

【使用指导】 日志过滤方式分为“只包含”和“只过滤”两种方式。默认为 `filter-only`，即“只过滤”。

设置日志信息的过滤规则

【命令格式】 `logging filter rule { exact-match module module-name mnemonic mnemonic-name level level | single-match { level level | mnemonic mnemonic-name | module module-name } }`

【参数说明】 `exact-match`：代表精确匹配，若选择精确匹配，则后面的三个过滤选项都需要选上。

`single-match`：代表单个匹配，若选择单个匹配，则后面的三个过滤选项只需要选择其中的一个。

`module module-name`：模块名，即填写要过滤的模块名称。

`mnemonic mnemonic-name`：助记符名称，即填写要过滤的日志信息助记符名称。

`level level`：日志信息级别，即填写要过滤的日志信息等级。

【命令模式】 全局配置模式

【使用指导】 日志过滤规则分为“精确匹配”和“单个匹配”两种过滤规则。

no logging filter rule exact-match [module *module-name* mnemonic *mnemonic-name* level *level*]命令删除日志信息的“精确匹配”过滤规则。支持一次性删除所有的“精确匹配”过滤规则，也可以逐条进行删除。

no logging filter rule single-match [level *level* | mnemonic *mnemonic-name* | module *module-name*]命令删除日志信息的“单个匹配”过滤规则。支持一次性删除所有的“单个匹配”过滤规则，也可以逐条进行删除。

配置举例

配置系统日志过滤功能

【网络环境】 假设网络环境中，有以下日志信息过滤功能设置要求：

- 1、设置日志信息的过滤方向为 **terminal**、**server** 两个方向；
- 2、设置日志信息的过滤方式为“只过滤”；
- 3、设备日志信息的过滤规则为“单个匹配”，并且模块名包含 SYS 的日志信息过滤掉。

【配置方法】 ● 在设备上面配置系统日志的过滤功能

```
Ruijie# configure terminal
Ruijie(config)# logging filter direction server
Ruijie(config)# logging filter direction terminal
Ruijie(config)# logging filter type filter-only
Ruijie(config)# logging filter rule single-match module SYS
```

【检验方法】 ● 通过 **show running-config | include logging** 命令可以查看用户配置的相关参数信息。

● 通过进入/退出全局配置模式，观察系统是否会输出日志信息。

```
Ruijie#configure
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#exit
Ruijie#
Ruijie#show running-config | include logging
logging filter direction server
logging filter direction terminal
logging filter rule single-match module SYS
```

6.5.8 配置系统日志重定向功能

配置效果

- 在 VSU 环境当中，从机或备机上面的日志信息不仅可以显示在从机或备机的 Console 窗口上，也可以重定向到主机上面进行输出，包括输出到主机的 Console 窗口、VTY 窗口上，也可以记录在主机的内存缓冲区、扩展 FLASH 和 Syslog Server 上。
- 在盒式设备的 VSU 环境中，打开重定向日志开关后，从机或备机的日志信息将重定向到主机进行输出，输出时，会在日志信息内容的最前面添加上对应的角色标志串“(*设备号)”，用于标识该日志信息是重定向日志信息。在 VSU 环境下面，假设同时存在四个设备，主机设备号为 1，从机设备号为 2，备机设备号分别为 3 和 4，则主机自身产生的日志不会添加角色标志串，从机重定向到主机的日志将添加角色标志串：(*2)，备机重定向到主机的日志将分别添加角色标志串：(*3) 和(*4)。
- 在卡式设备的 VSU 环境中，打开重定向日志开关后，从管理板或备份管理板的日志信息将重定向到主管理板进行输出，输出时，会在日志信息内容的最前面添加上对应的角色标志串“(*设备号/管理板名称)”，用于标识该日志信息是重定向日志信息。在 VSU 环境下面，若同时存在四块管理板，角色标志串可以形成四种形式：(*1/M1)、(*1/M2)、(*2/M1)、(*2/M2)。

注意事项

- 此功能只在 VSU 主从环境当中才存在，其它单机环境不存在。
- 对重定向到主机的日志信息进行速率限制，防止从机或备机上面出现大量日志信息的情况对系统造成负担。

配置方法

📌 打开日志重定向功能

- 可选配置，缺省情况下 VSU 环境中，日志重定向功能是开启的。
- 若无特殊要求，应在 VSU 主机或主管理板上配置。

📌 设置重定向日志信息速率限制功能

- 可选配置，缺省情况下 VSU 环境中，限制从机重定向到主机的日志信息每秒最多 200 条。
- 若无特殊要求，应在 VSU 主机或主管理板上配置。

检验方法

- 通过 **show running** 命令可以查看设置的日志重定向参数信息

相关命令

📌 打开日志重定向功能

- 【命令格式】 **logging rd on**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 VSU 环境下面，重定向日志开关默认是开启的。

设置重定向日志信息速率限制功能

【命令格式】 **logging rd rate-limit number [except level]**

【参数说明】 **rate-limit number** : 每秒钟内允许重定向的日志信息, 范围为 1~10000

except level : 小于等于此严重性级别 (0 ~ level) 的日志信息, 不进行速率控制; 默认级别为 error(3), 对小于等于 error 级别的日志信息不进行速率控制

【命令模式】 全局配置模式

【使用指导】 系统默认情况下, 限制从机重定向到主机的日志信息每秒最多 200 条。

配置举例

配置系统日志重定向功能

【网络环境】 假设在 VSU 环境当中, 有以下日志信息重定向功能设置要求:

- 1、打开日志重定向功能;
- 2、限制大于 critical (2 级) 的日志信息重定向速率为每秒 100 条。

【配置方法】

- 在设备上面配置日志重定向功能

```
Ruijie# configure terminal
Ruijie(config)# logging rd on
Ruijie(config)# logging rd rate-limit 100 except critical
```

【检验方法】

- 通过 **show running-config | include logging** 命令可以查看用户配置的相关参数信息。
- 通过在从机上面触发产生一条日志信息, 并在主机界面上面观察重定向到主机的日志信息。

```
Ruijie#show running-config | include logging
logging rd rate-limit 100 except critical
```

6.5.9 配置系统日志监控功能

配置效果

- 记录用户登录/退出的日志信息。开启记录用户登录/退出的日志信息后, 当外界通过 Telnet/SSH 连接到设备时, 设备将打出对应的 Log 信息, 方便管理员监控设备的连接情况。
- 记录用户修订设备配置的日志信息。开启记录用户操作的日志信息后, 当用户修订设备配置的时候, 设备将打出对应的 Log 信息, 方便管理员监控设备的配置修订情况。

注意事项

- 若设备上面同时配置 **logging userinfo** 和 **logging userinfo command-log**, 则进行 **show running-config** 查看时, 只会显示 **logging userinfo command-log**。

配置方法

▾ 开启记录用户登录/退出日志信息

- 可选配置，缺省情况下用户输入与日志信息输出同步功能是关闭的。
- 若无特殊要求，应在设备各个线路上面配置。

▾ 开启记录用户操作的日志信息

- 可选配置，缺省情况下用户输入与日志信息输出同步功能是关闭的。
- 若无特殊要求，应在设备各个线路上面配置。

检验方法

- 通过 **show running** 命令可以查看设置的用户输入同步输出功能参数信息

相关命令

▾ 开启记录用户登录/退出日志信息

【命令格式】 **logging userinfo**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 默认情况下，用户登录/退出设备的时候，设备是不会记录相关的 Log 信息。

▾ 开启记录用户操作的日志信息

【命令格式】 **logging userinfo command-log**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 设置执行配置命令时，记录用户操作的 Log 信息。默认情况下，用户修订设备配置的时候，设备是不会记录相关的操作 Log 信息。

配置举例

▾ 配置系统日志监控功能

【网络环境】 假设在网络环境当中，有以下日志信息监控功能设置要求：

- 1、开启记录用户登录/退出日志信息；
- 2、开启记录用户操作的日志信息。

【配置方法】 ● 在设备上配置日志监控功能

```
Ruijie# configure terminal
Ruijie(config)# logging userinfo
Ruijie(config)# logging userinfo command-log
```

- 【检验方法】
- 通过 **show running-config | include logging** 命令可以查看用户配置的相关参数信息。
 - 通过在设备全局配置模式里面配置一条命令，触发系统产生用户操作的日志信息。

```
Ruijie#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#interface gigabitEthernet 0/0
*Jun 16 15:03:43: %CLI-5-EXEC_CMD: Configured from console by admin command: interface
GigabitEthernet 0/0
Ruijie#show running-config | include logging
logging userinfo command-log
```

6.5.10 配置用户输入与日志信息同步输出功能

配置效果

- 默认情况下，用户输入与日志信息输出不同步。配置输入同步功能后，即使在用户输入的过程中打印日志，在打印结束后仍然会将用户之前的输入显示出来，从而保证输入的完整性和连贯性。

注意事项

- 该配置命令需要在线路配置模式下面进行配置，并且在每个需要开启此功能的线路上面均要进行配置。

配置方法

▾ 设置用户输入与日志信息输出同步功能

- 可选配置，缺省情况下用户输入与日志信息输出同步功能是关闭的。
- 若无特殊要求，应在设备各个需要开启此功能的线路上面配置。

检验方法

- 通过 **show running** 命令可以查看设置的用户输入同步输出功能参数信息

相关命令

▾ 设置用户输入与日志信息输出同步功能

- 【命令格式】 **logging synchronous**
- 【参数说明】 -
- 【命令模式】 线路配置模式
- 【使用指导】 此命令打开用户输入与日志信息输出同步功能，可以防止用户正在输入的字符时被打断。

配置举例

配置用户输入与日志信息输出同步功能

- 【网络环境】 假设在网络环境当中，有以下用户输入同步输出功能设置要求：
- 1、设置用户输入与日志信息同步输出功能。

- 【配置方法】 ● 在设备上面配置用户

```
Ruijie# configure terminal
Ruijie(config)# line console 0
Ruijie(config-line)# logging synchronous
```

- 【检验方法】 ● 通过 **show running-config | begin line** 命令可以查看用户配置的相关参数信息。

```
Ruijie#show running-config | begin line
line con 0
logging synchronous
login local
```

如下所示，当用户敲入“vlan”后接口 0/1 发生状态改变，打印日志，打印结束后日志模块会自动把用户已经输入的“vlan”打印出来，使得用户可以继续输入：

```
Ruijie(config)#vlan
*Aug 20 10:05:19: %LINK-5-CHANGED: Interface GigabitEthernet 0/1, changed state to up
*Aug 20 10:05:19: %LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet 0/1, changed state to up
Ruijie(config)#vlan
```

6.6 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除内存缓冲区中的日志信息	clear logging

查看运行情况

作用	命令
查看内存缓冲区中的日志报文，以及日志相关统计信息，日志信息按时间戳从旧到新的顺序显示	show logging
查看系统日志配置的参数、统计信息	show logging config
查看系统中各模块日志信息统计情况	show logging count

7 CWMP

7.1 概述

CWMP 协议(CPE WAN Management Protocol 即 CPE 广域网管理协议)提供了设备统一管理的通用框架、消息规范、管理方法和数据模型,解决了用户侧设备数量繁多,部署分散,不易统一管理和维护的问题,提高了问题响应效率,节约了运维成本。

CWMP 协议主要提供下面一些功能:

- 自动配置和动态服务提供,用户侧设备启动初次接入网络时自动从管理服务器处获取配置,用户侧设备在运行过程中,管理服务器可以动态的改变其配置和状态;
- 主程序 / 配置文件管理,提供主程序和配置文件的升级及配置文件的上传;
- 软件模块功能的管理,通过各软件模块实现的数据模型对各软件模块进行管理;
- 状态行为监控,用户侧设备运行时状态及配置变化通告给管理服务器,通过这些实时变化的通告实现对用户侧设备的监控;
- 故障诊断,管理服务器通过用户侧设备提供的信息诊断或解决连接性问题及其他服务性问题,同时可以执行一些预定义的诊断行为。

 下文仅介绍 CWMP 的相关内容。

协议规范

TR069 的协议规范详见官方论坛：<http://www.broadband-forum.org/technical/trlist.php>。以下是主要的几份规范：

- 《TR-069_Amendment-4.pdf》，CWMP 协议标准。
- 《TR-098_Amendment-2.pdf》，CWMP 协议网关产品数据模型规范。
- 《TR-106_Amendment-6.pdf》，CWMP 协议 CPE 数据模型标准。
- 《TR-181_Issue-2_Amendment-5.pdf》，CPE 数据模型 2 规范。
- 《tr-098-1-4-full.xml》，CWMP 协议网关产品数据模型定义。
- 《tr-181-2-4-full.xml》，CWMP 协议 CPE 数据模型 2 定义。

7.2 典型应用

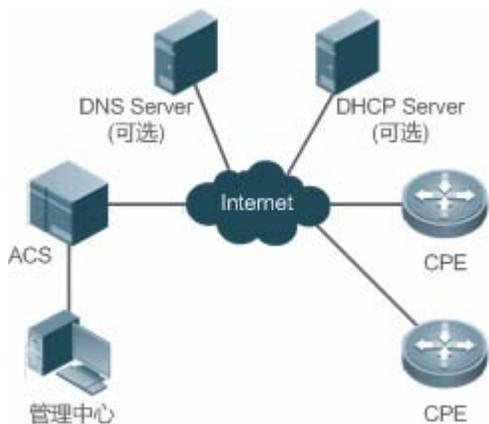
典型应用	场景描述
CWMP网络应用场景	通过配置 CPE 设备与 ACS 服务器建立连接,以实现 CPE 设备主程序升级,配置文件上传,恢复等功能。

7.2.1 CWMP网络应用场景

应用场景

CWMP 的网络结构中主要包括 CPE、ACS、管理中心、DHCP 服务器和 DNS 服务器。大量的 CPE 接受 ACS 的管理，管理中心通过控制 ACS 服务器，实现对 CPE 设备的管理控制，一般控制中心为 WEB 浏览器，通过 WEB 浏览器控制 ACS 服务器

图 7-1



- 【注释】
- DHCP 服务器用于动态获取 ACS 的 URL，如果使用静态配置 ACS 的 URL，DHCP 服务器在该网络中为可选元素
 - DNS 服务器用于解析 ACS 或 CPE 的域名，如果 ACS 和 CPE 的 URL 中直接使用 IP 地址而不是域名，DNS 服务器在该网络中为可选元素

功能部属

- CPE 和 ACS 设备要运行 HTTP 协议

7.3 功能详解

基本概念

常用术语

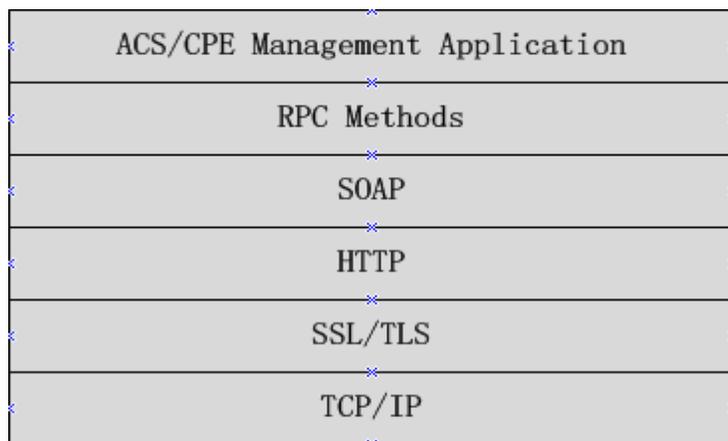
- CPE : Customer Premises Equipment (顾客预置设备)。
- ACS : Auto-Configuration Server (自动配置服务器)。

- RPC：Remote Procedure Call（远程过程调用）。
- DM：Data Model，数据模型。

协议结构

CWMP 的协议结构如下图所示：

图 7-2 CWMP 的协议结构



如图所示，协议规范将 CWMP 协议的工作分为 6 个层次，各层次的功能及作用说明如下：

- ACS/CPE Management Application

运用程序管理层，该层并非 CWMP 协议本身的范畴，它是指 CPE/ACS 的各功能模块为了支持 CWMP 的管理进行的开发，如同 SNMP 一样，各功能模块的 MIB 管理本身即不属于 SNMP 的协议范畴。

- RPC Methods

RPC 方法管理层，该层提供了 ACS 与 CPE 之间交互的各种 RPC 方法，实现各种 RPC 方法的操作。

- SOAP

简单对象访问协议层，该层提供了 CWMP 协议的 XML 形式封装与解封装，CWMP 消息格式必须符合 SOAP 的封装语法。

- HTTP

所有的 CWMP 消息最终通过 HTTP 协议进行传输，ACS 和 CPE 同时支持 HTTP 客户端和服务器端功能，服务器端用于监控对端的反响连接。

- SSL/TLS

该层提供 CWMP 协议的安全性保证，包括数据完整性，机密性及认证的保护。

- TCP/IP

TCP/IP 协议栈。

RPC 方法管理

ACS 对 CPE 的管理监控主要是通过 RPC 方法进行的，主要包括如下的一些方法：

- GET 系列方法

该系列方法主要用于 ACS 远程获取 CPE 支持的 RPC 方法、CPE 支持的数据模型参数名、数据模型参数的值和数据模型参数的属性。

- SET 系列方法

该系列方法主要用于 ACS 远程设置 CPE 支持的数据模型参数的值和数据模型参数的属性。

- INFORM 方法

INFORM 方法用于 CPE 向 ACS 通告自己的设备标识、参数信息及所发生的事件。INFORM 方法为 ACS 与 CPE 建立会话时交互的第一个方法。

- DownLoad 方法

DownLoad 方法实现 ACS 远程控制 CPE 下载文件的管理，包括 CPE 主程序升级的控制、配置文件升级的控制和 WEB 包升级的控制。

- UpLoad 方法

UpLoad 方法实现 ACS 远程控制 CPE 上传文件的管理，包括 CPE 配置文件上传的控制、日志文件上传的控制。

- Reboot 方法

Reboot 方法用于 ACS 远程控制 CPE 的重启行为。

▾ 会话管理

CWMP 协议工作的基础是 CWMP 协议会话，CWMP 的交互就是 CWMP 的会话交互，CWMP 协议在 ACS 与 CPE 之间的所有交互都以其会话为基础，通过会话传输、管理、维护其操作，实现 ACS 与 CPE 之间的有效交互，实现 ACS 对 CPE 的管理和监控。ACS 与 CPE 的一次会话过程即为两者建立 TCP 连接，Inform 协商开始到当前所有交互完成 TCP 连接断开为止，这个过程称之为一次会话过程。根据会话发起方角色的不同将其分为 CPE 主动发起的会话和 ACS 请求的会话两种，下面就这两种运用场景进行说明。

▾ 数据模型管理

CWMP 数据模型是 CWMP 工作的依据，CWMP 对所有功能模块的管理都是对 CWMP 数据模型的操作，各功能模块注册并实现自己支持数据模型，如果 SNMP 中各功能模块实现的 MIB 一样。

CWMP 数据模型以字符串的形式表示，为了区分数据模型的层次关系，以“.”分隔符区分上下级数据模型节点之间的关系，如 InternetGatewayDevice.LANDevice 的数据模型表示中，InternetGatewayDevice 为 LANDevice 的父数据模型节点，而 LANDevice 为 InternetGatewayDevice 子数据模型节点。

数据模型节点分为两类，一类为对象节点（object），一类为参数节点（parameter），也叫叶子节点。对象节点是指那些其下还有子节点的节点。而参数节点即没有子节点的叶子节点。对象节点分单实例对象节点和多实例对象节点，单实例对象节点指只存在单个实例对象的节点，多实例对象节点指存在多个实例对象的节点。数据模型节点分为可读节点和可读可写节点，可读节点只能读取其参数的值，不能进行修改，可读可写节点除了可读该节点参数值外还能对其进行修改。

数据模型节点存在两种属性，一种为是否通告的属性，即该数据模型对应参数值发生变化（非 CWMP 协议引起的变化）时是否将其通告给 ACS 服务器；一种是模型节点参数可被其他管理方式（非 ACS）写操作的标识，即其他管理方式如 Telnet 等是否可对该参数值进行修改。ACS 可以通过 RPC 方法修改数据模型的属性。

CWMP 协议对数据模型的管理通过对应的 RPC 方法进行。

▾ 事件管理

在 CPE 设备上,当一些 ACS 感兴趣或关心的事件发生时,CPE 需要将这些事件通告给 ACS,ACS 通过监控这些事件来监控 CPE 的工作状态,CWMP 的事件如同 SNMP 中的 TRAP 和产品日志功能中的日志信息。ACS 可以通过 RPC 方法控制和调整自己关心的事件,过滤掉不关心的事件类型。CWMP 中的事件总体分为两类,单量事件类型和增量事件类型,单量事件类型的事件是指同一个事件第二次发生时,该事件不再有量的变化而是丢弃老的,保留新的,增量事件类型是指同一事件后续多次发生时,老的不能丢弃,新产生的事件作为一个完整的事件保留,量上加 1。

CPE 产生的所有事件通过 INFORM 方法向 ACS 通告。

功能特性

功能特性	作用
主程序升级	ACS 通过 DownLoad 方法控制 CPE 的主程序升级
配置文件升级	ACS 通过 DownLoad 方法控制 CPE 的配置文件升级
配置文件上传	ACS 通过 UpLoad 方法控制 CPE 的配置文件上传
CPE备份恢复	当设备出现脱管状态时,远程设备恢复到脱管前的状态

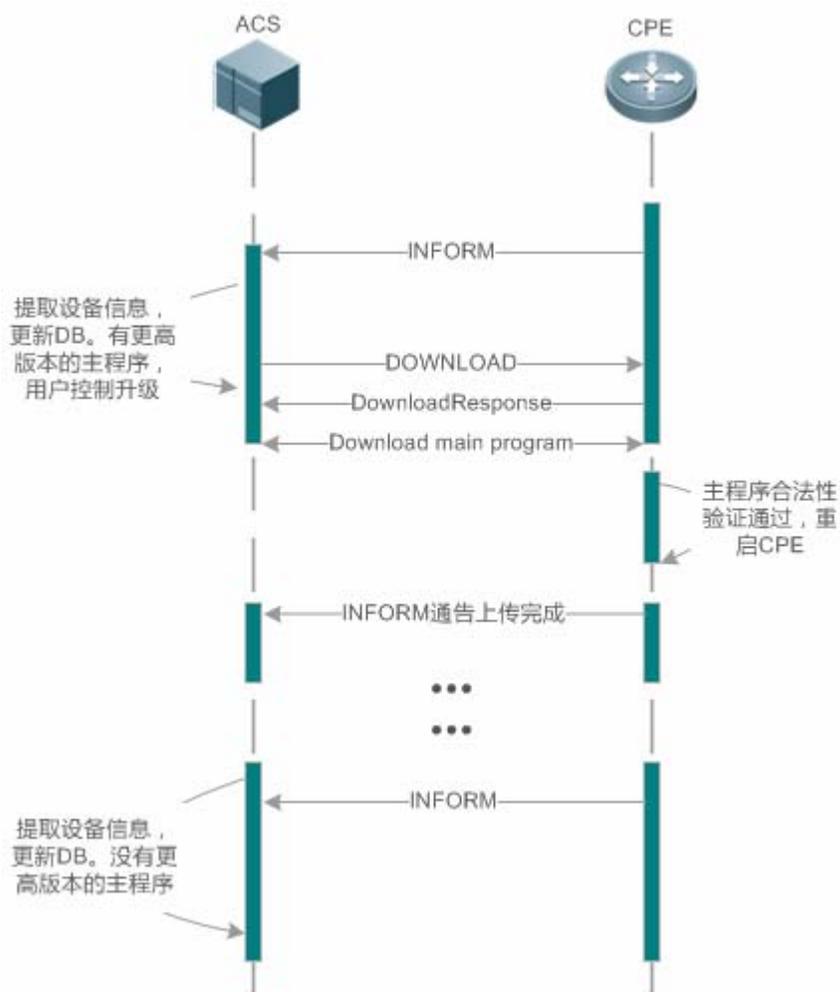
7.3.1 主程序升级

主程序升级,是指对各网元设备的主程序进行更新,通过主程序的升级来达到对设备版本的升级或更新换代

工作原理

主程序升级时序图

图 7-3



用户指定 CPE 升级主程序，ACS 向 CPE 下发升级主程序的 DownLoad 方法。CPE 从 DownLoad 方法中指定的文件服务器下载主程序，升级自身的主程序，更新完主程序后重启 CPE 设备，完成主程序的升级。CPE 重启后向 ACS 通告主程序管理升级完成。

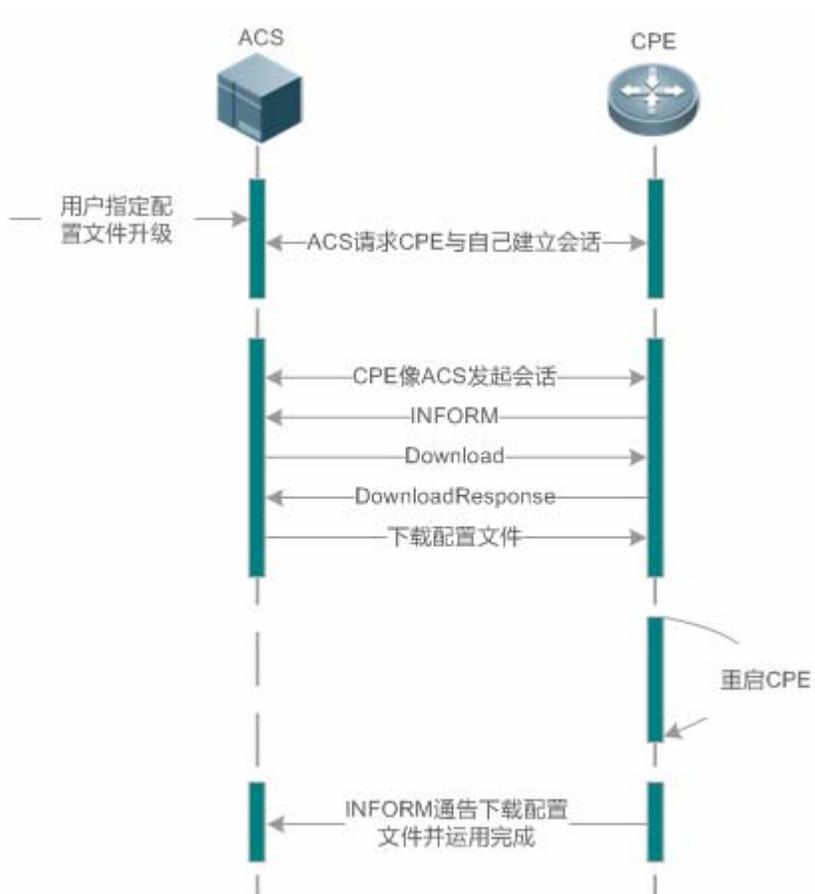
i ACS 可以同时作为文件服务器，文件服务器也可以作为单独的服务器部署。

7.3.2 配置文件升级

配置文件升级，是指将整个设备的当前的配置文件替换为指定的配置，设备复位后，系统将运行全新的配置

工作原理

图 7-4



用户指定 CPE 升级配置文件，ACS 向 CPE 下发升级配置文件的 DownLoand 方法。CPE 从 DownLoad 方法中指定的文件服务器下载配置文件，升级自身的配置文件，更新完配置文件后重启 CPE 设备，完成配置文件的升级。CPE 重启后向 ACS 通告配置文件升级完成。

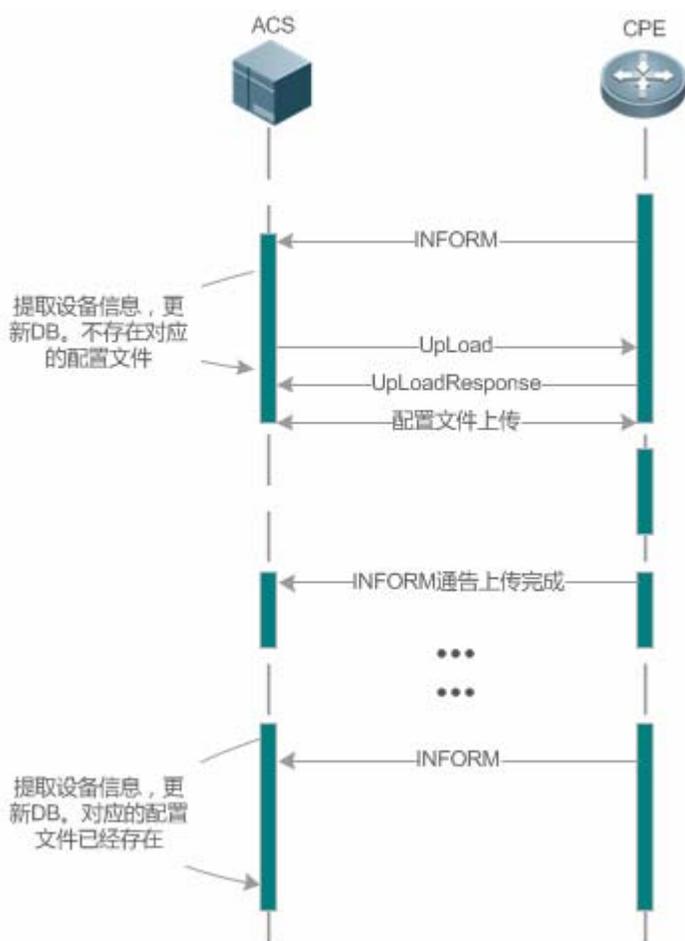
i ACS 可以同时作为文件服务器，文件服务器也可以作为单独的服务器部署。

7.3.3 配置文件上传

配置文件上传，ACS 控制 CPE 的配置文件是通过 UpLoad 方法上传其配置文件

工作原理

图 7-5



CPE 初次上线 ACS，ACS 需要学习 CPE 的配置文件，学习过程如下：

- ACS 初次收到 CPE 的 `INFORM` 消息，根据 `INFORM` 消息中的设备信息，找到或建立对应的 CPE DB 信息；
- ACS DB 中还不存在当前 CPE 的配置文件，ACS 向该 CPE 下发上传配置文件的 `UpLoad` 方法；
- CPE 将当前的配置文件上传给 ACS 服务器；
- CPE 通告 ACS 配置文件上传完成。

7.3.4 CPE备份恢复

CPE 备份恢复，是指管理端由于一些异常的操作导致远程设备脱管，当这种情况发生时，需要远程设备恢复到脱管前的状态，恢复对远程设备的管理，重新对远程设备执行正确的管理与操作。

工作原理

在设备上配置 CPE 主程序/配置在异常情况下的恢复功能，指当 CPE 进行主程序/配置升级后无法连接 ACS，出现脱管现象时，能及时的恢复到脱管前的主程序和配置，恢复 ACS 对 CPE 的管理，这种情况的出现一般是下发了错误的主程序或配置所导致。

CPE 在每次接收主程序升级以及配置下发时 先备份当前的配置与主程序 并提供机制用于判断是否出现上述场景描述的问题，若出现，则将系统恢复到之前的可管理状态

7.4 配置详解

配置项	配置建议 & 相关命令	
建立CWMP基本连接	 必须配置。配置 CPE 连接 ACS 时用于认证的用户名和密码及配置 ACS 连接 CPE 时用于认证的用户名和密码	
	cwmp	使能 CWMP 并进入 CWMP 配置模式
	acs username	配置 CPE 连接 ACS 时用于认证的用户名
	acs password	配置 CPE 连接 ACS 时用于认证的密码
	cpe username	配置 ACS 连接 CPE 时用于认证的用户名
	cpe password	配置 ACS 连接 CPE 时用于认证的密码
	 可选配置。配置 CPE 及 ACS 设备的 URL	
	acs url	配置 CPE 连接 ACS 的 URL。
cpe url	配置 ACS 连接 CPE 的 URL	
配置CWMP相关属性	 可选配置。配置 CPE 设备的基本功能(如 CPE 主程序/配置备份恢复功能，不向 ACS 上传配置文件和日志文件的管理等)	
	cpe inform	配置 CPE 周期性 INFORM 通告功能
	cpe back-up	配置 CPE 主程序/配置备份恢复功能
	disable download	配置关闭从 ACS 下载主程序和配置文件的功能
	disable upload	配置关闭向 ACS 上传配置文件和日志文件的功能
	timer cpe- timeout	配置 ACS 无响应 CPE 超时时间

7.4.1 建立CWMP基本连接

配置效果

- 实现 ACS 设备与 CPE 设备会话连接的建立

注意事项

- 无

配置方法

▾ 使能 CWMP 并进入 CWMP 配置模式

- 默认开启 CWMP 功能。
- CPE 设备上必须配置。

【命令格式】 **cwmp**
【参数说明】 -
【缺省配置】 CWMP 功能开启
【命令模式】 全局模式
【使用指导】 -

▾ 配置 CPE 连接 ACS 时用于认证的用户名

- ACS 设备上必须配置。
- 只能配置一个 ACS 用户名，多次配置 ACS 的用户名时，最新的配置生效。

【命令格式】 **acs username *username***
【参数说明】 **username *username*** : 配置 CPE 连接 ACS 用于认证的用户名。
【缺省配置】 无默认 ACS 用户名
【命令模式】 cwmp 配置模式
【使用指导】 -

▾ 配置 CPE 连接 ACS 时用于认证的密码

- ACS 设备上必须配置。
- ACS 用户密码可以为明文和密文形式，只能配置一个 ACS 用户密码，多次配置 ACS 的用户密码时，最新的配置生效。

【命令格式】 **acs password {*password* | *encryption-type* *encrypted-password*}**
【参数说明】 ***password***: CPE 连接 ACS 用于认证的密码。
encryption-type: 可配置为 0 或 7，为 0 表示无加密，为 7 表示简单加密
encrypted-password: 密码文本
【缺省配置】 ***encryption-type*** 默认为 0，***encrypted-password*** 默认为空。
【命令模式】 cwmp 配置模式
【使用指导】 -

▾ 配置 ACS 连接 CPE 时用于认证的用户名

- CPE 设备上必须配置。
- 只能配置一个 CPE 用户名，多次配置 CPE 的用户名时，最新的配置生效。

【命令格式】 **cpe username *username***
【参数说明】 ***Username***: 配置 ACS 连接 CPE 用于认证的用户名
【缺省配置】 无默认 CPE 用户名
【命令模式】 cwmp 配置模式
【使用指导】 -

▾ 配置 ACS 连接 CPE 时用于认证的密码

- CPE 设备上必须配置。
- CPE 用户密码可以为明文和密文形式，只能配置一个 CPE 用户密码，多次配置 CPE 的用户密码时，最新的配置生效。

【命令格式】 **cpe password** {*password* | *encryption-type* *encrypted-password*}

【参数说明】 *password*: ACS 连接 CPE 用于认证的密码。

encryption-type: 可配置为 0 或 7，为 0 表示无加密，为 7 表示简单加密

encrypted-password: 密码文本

【缺省配置】 *encryption-type* 默认为 0，*encrypted-password* 默认为空。

【命令模式】 cwmp 配置模式

【使用指导】 配置 ACS 连接 CPE 用于认证的密码，通常无须输入加密类型。一般情况下，只有当复制并粘贴已经加密过后该命令的密码时，才需要输入加密类型。有效密码的格式要求如下：

- 必须包含 1 到 26 个大小写字母和数字字符。
- 密码前面可以有前导空格，但被忽略。中间及结尾的空格则作为密码的一部分。
- *encryption-type* 为 7 时，输入的合法字符只能是数字 0~9、字符 a~f、A~F。

配置 CPE 连接 ACS 的 URL

- 可选配置，在 CPE 设备上配置。
- 只能配置一个 ACS URL，多次配置 ACS 的 URL 时，最新的配置生效，ACS 的 URL 必须是 HTTP 的形式。

【命令格式】 **acs url** *url*

【参数说明】 *url*: ACS 的 URL。

【缺省配置】 无默认 ACS URL

【命令模式】 CWMP 配置模式

【使用指导】 配置 CPE 连接 ACS 的 URL，在没有手动配置 ACS URL 的情况下，如果使用了 DHCP 获取到了动态的 ACS URL，将使用动态获取到的 ACS URL 向 ACS 发起连接。对 ACS 的 URL 格式要求如下：

- ACS 的 URL 格式必须为：`http://host[:port]/path`（或是 `https://host[:port]/path`）的格式。
- ACS URL 的最大长度为 256 个字符。

配置 ACS 连接 CPE 的 URL

- 可选配置，CPE 设备上配置。
- 只能配置一个 CPE URL，多次配置 CPE 的 URL 时，最新的配置生效。CPE 的 URL 必须是 HTTP 的形式，不支持域名的形式配置 CPE 的 URL。

【命令格式】 **cpe url** *url*

【参数说明】 *url*: cpe 的 URL。

【缺省配置】 无默认 CPE URL

【命令模式】 cwmp 配置模式

【使用指导】 配置 ACS 连接 CPE 的 URL，在没有手动配置的情况下，CPE 将根据 ACS 的 URL 自动选取 CPE 的 URL，CPE 的 URL 格式要求如下：

- CPE 的 URL 必须是 `http://ip [: port]/` 的格式。
- CPE URL 的最大长度为 256 个字符。

检验方法

- 通过 `show cwmp configuration` 命令查看。

【命令格式】 **show cwmp configuration**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 -

【命令展示】 1：显示 CWMP 功能的当前配置。

```
Ruijie(config-cwmp)#show cwmp configuration
CWMP Status           : enable
ACS URL               : http://www.ruijie.com.cn/acs
ACS username         : admin
ACS password         : *****
CPE URL              : http://10.10.10.2:7547/
CPE username         : ruijie
CPE password         : *****
CPE inform status    : disable
CPE inform interval  : 60s
CPE inform start time : 0:0:0 0 0 0
CPE wait timeout     : 50s
CPE download status  : enable
CPE upload status    : enable
CPE back up status   : enable
CPE back up delay time : 60s
```

配置举例

i 以下配置举例，仅介绍 CWMP 相关的配置。

在 CPE 设备上配置用户名和密码

【网络环境】

图 7-6



【配置方法】

- 使能 CWMP
- 在 CPE 设备上配置 CPE 连接 ACS 时用于认证的用户名和密码
- 在 CPE 设备上配置 ACS 连接 CPE 时用于认证的用户名和密码

CPE

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
```

```
Ruijie(config)# cwmp
Ruijie(config-cwmp)# acs username USERB
Ruijie(config-cwmp)# acs password PASSWORDB
Ruijie(config-cwmp)# cpe username USERB
Ruijie(config-cwmp)# cpe password PASSWORDB
```

【检验方法】 ● 通过在 CPE 设备上 **show** 命令查看命令是否配置成功

CPE

```
Ruijie # show cwmp configuration

CWMP Status           : enable
ACS URL                : http://10.10.10.1:7547/acs
ACS username          : USERA
ACS password           : *****
CPE URL                : http://10.10.10.2:7547/
CPE username           : USERB
CPE password           : *****
```

配置 ACS 和 CPE 的 URL 连接

【网络环境】 同图 7-6

【配置方法】 ● 配置 ACS 设备的 URL 地址
● 配置 CPE 设备的 URL 地址

CPE

```
Ruijie# configure terminal
Ruijie(config)# cwmp
Ruijie(config-cwmp)# acs url http://10.10.10.1:7547/acs
Ruijie(config-cwmp)# cpe url http://10.10.10.1:7547/
```

【检验方法】 通过在 CPE 设备上 **show** 命令查看命令是否配置成功

CPE

```
Ruijie #show cwmp configuration

CWMP Status           : enable
ACS URL                : http://10.10.10.1:7547/acs
ACS username          : USERA
ACS password           : *****
CPE URL                : http://10.10.10.2:7547/
```

常见错误

- 如果用户输入密码类型为密文，密码长度超过 254 或者长度不为偶数；
- 如果用户输入密码类型为明文，密码长度超过 100；
- 如果用户输入密码类型为明文，并且包含非法字符；

- 如果用户输入密码类型为密文，并且包含非法字符（合法字符只能是：数字 0~9、字符 a~f、A~F）；
- 如果 ACS URL 地址为 NULL，则提示错误；
- 如果 CPE URL 地址为 NULL，则提示错误。

7.4.2 配置CWMP相关属性

配置效果

- 用于实现 CPE 设备常用功能的配置(如主程序/配置备份恢复，是否接受 ACS 下发主程序及配置文件，是否向 ACS 上传配置文件和日志文件等)

配置方法

配置 CPE 周期性通告功能

- 可选配置，单位为秒，取值范围 30~3600，缺省值 600s。
- 当 CPE 设备需要重新设定周期性通告时间的时候配置该功能。

【命令格式】 **cpe inform [interval seconds] [starttime time]**

【参数说明】 **seconds**：配置 CPE 周期性 INFORM 通告时间间隔。单位为秒，取值范围 30~3600，缺省值 600。
time：开始周期性 INFORM 的日期时间，格式为 yyyy-mm-ddThh:mm:ss

【缺省配置】 CPE 的 INFORM 通告时间间隔为 600 秒

【命令模式】 cwmp 配置模式

【使用指导】 配置 CPE 周期性 INFORM 通告功能。

- 在没有配置 INFORM 开始时间的情况下，周期性 INFORM 从开启该功能开启，每经过一个 INFORM 周期通告一次。
- 在配置了 INFORM 开始日期时间的情况下，周期性 INFORM 的开始时间为该指定时间。如配置 INFORM 周期为 60 秒，开始时间为明天中午 12 点，则周期性 INFORM 通告从明天中午 12 点才开始，且每经过 60 秒 INFORM 通告一次。

配置关闭从 ACS 下载主程序和配置文件的功能

- 可选配置，默认开启下载主程序和配置文件的功能。
- 当 CPE 设备要关闭下载主程序和配置文件功能的时候，配置该功能。

【命令格式】 **disable download**

【参数说明】 -

【缺省配置】 CPE 文件下载功能为默认允许接收

【命令模式】 cwmp 配置模式

【使用指导】 配置关闭从 ACS 下载主程序和配置文件的功能。

- 这个命令对配置脚本文件不起作用，配置 disable 的情况下，配置脚本可以执行。

配置关闭向 ACS 上传配置文件和日志文件的功能

- 可选配置，默认为开启上传配置和日志文件功能。
- 当 CPE 设备要关闭上传配置文件和日志文件功能的时候，配置该功能。

【命令格式】 **disable upload**

【参数说明】

【缺省配置】 CPE 文件上传功能为开启

【命令模式】 cwmp 配置模式

【使用指导】 配置关闭向 ACS 上传配置和日志文件的功能。

配置 CPE 主程序/配置备份恢复功能

- 可选配置，默认开启 CPE 主程序/配置备份恢复功能，单位为秒，取值范围 30-10000，默认备份恢复时间为 60s。
- 恢复时间设置越大，CPE 启动恢复延迟时间越久。
- 当 CPE 设备需要修改 CPE 主程序/配置备份恢复功能的时间时，配置该功能。

【命令格式】 **cpe back-up [delay-time seconds]**

【参数说明】 *seconds*: CPE 主程序/配置备份恢复的延迟时间

【缺省配置】 CPE 备份恢复功能为开启，默认恢复时间为 60s

【命令模式】 cwmp 配置模式

【使用指导】 -

配置 ACS 无响应 CPE 超时时间

- 可选配置，单位为秒，取值范围 10~600，默认值为 30s。
- CPE 设备上配置，当 CPE 设备需要修改 ACS 无响应 CPE 超时时间时，配置该功能。

【命令格式】 **timer cpe- timeout seconds**

【参数说明】 *seconds*: 超时时间，单位为秒，取值范围 10~600。

【缺省配置】 默认值为 30 秒。

【命令模式】 cwmp 配置模式

【使用指导】 -

检验方法

- 通过 **show cwmp configuration** 命令查看。

【命令格式】 **show cwmp configuration**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 -

【命令展示】 1: 显示 CWMP 功能的当前配置。

```
Ruijie(config-cwmp)#show cwmp configuration
CWMP Status           : enable
```

```
ACS URL : http://www.ruijie.com.cn/acs
ACS username : admin
ACS password : *****
CPE URL : http://10.10.10.2:7547/
CPE username : ruijie
CPE password : *****
CPE inform status : disable
CPE inform interval : 60s
CPE inform start time : 0:0:0 0 0 0
CPE wait timeout : 50s
CPE download status : enable
CPE upload status : enable
CPE back up status : enable
CPE back up delay time : 60s
```

配置举例

配置 CPE 周期 INFORM 通告时间间隔

【网络环境】 同图 7-6

- 【配置方法】
- 开启 CWMP 功能并进入 CWMP 配置模式
 - 配置 CPE 周期 INFORM 通告时间间隔为 60 秒

```
CPE
Ruijie#config
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#cwmp
Ruijie(config-cwmp)#cpe inform interval 60
```

【检验方法】 通过在 CPE 设备上 **show** 命令查看命令是否配置成功

```
CPE
Ruijie #show cwmp configuration
CWMP Status : enable
.....
CPE inform interval : 60s
```

配置关闭从 ACS 下载主程序和配置文件的功能

【网络环境】 同图 7-6

- 【配置方法】
- 开启 CWMP 功能并进入 CWMP 配置模式
 - 配置关闭从 ACS 下载主程序和配置文件的功能

```
CPE
Ruijie#config
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#cwmp
Ruijie(config-cwmp)#disable download
```

【检验方法】 通过在 CPE 设备上 **show** 命令查看命令是否配置成功

```
CPE
Ruijie #show cwmp configuration
CWMP Status                : enable
.....
CPE download status        : disable
```

▾ 配置关闭向 ACS 上传配置文件和日志文件的功能

【网络环境】 同图 7-6

【配置方法】

- 开启 CWMP 功能并进入 CWMP 配置模式
- 配置关闭向 ACS 上传配置文件和日志文件的功能

```
CPE
Ruijie#config
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#cwmp
Ruijie(config-cwmp)# disable upload
```

【检验方法】 通过在 CPE 设备上 **show** 命令查看命令是否配置成功

```
CPE
Ruijie #show cwmp configuration
CWMP Status                : enable
.....
CPE upload status          : disable
```

▾ 配置备份恢复的延迟时间

【网络环境】 同图 7-6

【配置方法】

- 开启 CWMP 功能并进入 CWMP 配置模式
- 配置备份恢复的延迟时间为 30s

```
CPE
Ruijie#config
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#cwmp
Ruijie(config-cwmp)# cpe back-up Seconds 30
```

【检验方法】

- 通过在 CPE 设备上 **show** 命令查看命令是否配置成功

```
CPE Ruijie #show cwmp configuration
CWMP Status           : enable
.....
CPE back up delay time : 30s
```

配置 CPE 无数据超时时间

【网络环境】 同图 7-6

- 【配置方法】
- 开启 CWMP 功能并进入 CWMP 配置模式
 - 配置备份恢复的延迟时间为 100s

```
CPE Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# cwmp
Ruijie(config-cwmp)# timer cpe-timeout 100
```

- 【检验方法】
- 通过在 CPE 设备上 show 命令查看命令是否配置成功

```
CPE Ruijie#show cwmp configuration
CWMP Status           : enable
.....
CPE wait timeout      : 100s
```

常见配置错误

无

7.5 监视与维护

查看运行情况

作用	命令
显示 CWMP 功能的当前配置	show cwmp configuration
显示 CWMP 的当前运行状态	show cwmp status

8 CA-MONITOR

8.1 概述

CA 智能监控，硬件智能管理功能，包括智能电源管理、智能风扇调速和智能温度监控三个部分。智能监控功能主要负责机箱电源功率的监控、板卡上下电处理和功率监控、根据环境温度变化自动调整风扇转速、实时监控板卡的温度变化给用户以提示、在板卡温度超过设备所能承受的最高温度时对板卡自动下电以保护设备等功能。

CA 智能监控在设备上电运行之后默认开启，自动进入智能监控模式，用户无需任何配置即可使用智能监控功能。除了缺省配置外，系统还提供了一系列配置方法，用户可以随时根据自己的需要改变系统的缺省配置。

协议规范

无

8.2 典型应用

典型应用	场景描述
电源供电配置	配置板卡的供电优先级，控制单板上下电，保存优先级配置
风扇工作模式配置	配置风扇的调速模式，支持配置的模式包括：标准模式（缺省模式）、静音模式、用户自定义模式
智能温控配置	配置板卡温度探测点的监控温度门限值

8.2.1 电源供电配置

应用场景

- 设备采用智能电源进行供电，可以配置板卡的供电优先级，控制单板上下电，读取电源工作温度、输入电压、风扇状态等信息。
- 板卡供电是有优先级之分的，优先级越高的板卡越早上电且越迟下电。默认的板卡供电优先级为：管理板 > FE 卡 > VSL 卡 > 线卡，相同类型的板卡，则槽位号越小的优先级越高。用户可以通过 CLI 命令手动更改线卡的供电优先级。

功能部属

无

8.2.2 风扇工作模式配置

应用场景

- 设备使用智能风扇进行系统散热，风扇转速可以随着环境温度的变化而自动调整，从而达到最佳的散热效果，而且在环境温度较低的情况下，维持较低的风扇转速，还能降低功耗和噪音。
- 支持 3 种风扇调速模式：
- 标准模式：设备出厂工作模式。风扇起始转速维持在中间转速，之后随着环境温度变化而自动降低或提高转速，以达到最佳的散热效果。
- 静音模式：风扇起始转速维持在相对较低的转速，之后随着环境温度变化而自动降低或提高转速。该模式的风扇转速相对比标准模式小一些，因此可以进一步降低噪音。
- 用户自定义模式：这种模式下，系统中风扇的转速将由用户自行设置，且所有风扇盘的转速保持一致（用户自定义的值）。用户自定义模式支持 7 个级别的风扇转速设置，如表 1-1 所示。设置的级别越高，则风扇转速越高。当用户将风扇工作模式切换到自定义模式时，若未指定风扇转速级别，则默认工作在第 3 级别。

表 1-1 用户自定义模式下的风扇转速级别定义

风扇转速级别	风扇转速（相对于风扇最高转速的百分比）
1	40%
2	50%
3	60%
4	70%
5	80%
6	90%
7	100%

 用户自定义模式的转速是固定的，只会根据用户设定的级数设置风扇转速，不会随着温度变化自动调整。因此，一般推荐用户使用标准模式或静音模式，以便风扇转速随着温度的变化自动调整，从而最大限度地保护设备不因过温而出现故障。

功能部属

无

8.2.3 智能温控配置

应用场景

- 设备的每张板卡，都包含 5 个温度探测点：入风口、出风口、CPU、2 个板上最热点。部分板卡（如多业务卡）包含 2 个 CPU，2 个 CPU 都会监控。除此之外，线卡和 FE 卡上还有 MAC 温度需要监控。这些温度点，每个温度点都有两个温度阈值（即门限值），分别是告警温度和危险温度。两个温度阈值的定义如下：
- 告警温度：板卡温度超过告警温度，主管理板打印 syslog 提示，且面板 Alarm 亮黄灯。
- 危险温度：即下电温度，板卡温度超过该温度，会自动下电，同时主管理板打印 syslog 提示，且面板 Alarm 亮红灯。

 CPU 温度和 MAC 温度未定义告警温度，仅包含危险温度。

- 设备监控的所有温度点，一共可以分为三类：
- 主板温度：包含板卡入风口、出风口、CPU、板上最热点的温度
- CPU 温度：板卡 CPU 的温度，每张板卡至少一个
- MAC 温度：引擎无，线卡和 FE 卡每张卡至少一个
- 用户可以针对上述 3 类温度分别设置阈值，阈值的设置只针对机箱进行设置，同一个机箱内所有板卡共用一套温度阈值。系统默认分配了一套阈值，用户可以使用缺省配置，也可以根据自己的需要更改系统默认配置。但是，用户配置的阈值，必须满足如下规则：告警温度配置值必须低于当前危险温度；主板温度危险温度配置值不能超过 90 度，CPU 温度和 MAC 温度危险温度配置值不能超过 110 度。

 系统默认阈值如下：主板温度的告警温度为 56 度，危险温度为 80 度；CPU 温度和 MAC 温度无告警温度，危险温度为 100 度。

- 板卡插入后，主管理板会自动对该板卡进行智能温度监控，一旦监控到该板卡某个温度点超过当前告警温度，会打印 syslog 提示，并给主管理板面板 Alarm 灯亮黄灯。假如监控到该板卡某个温度点超过当前危险温度，则会打印 syslog 提示且 Alarm 亮红灯，同时给该板卡下电。

 用户配置温度阈值时，系统不会对该阈值的下限进行判断，因此用户配置的阈值不宜过低，否则可能会频繁打印 syslog 或给板卡下电，影响用户对设备的使用。一般推荐主板告警温度以 50 度为下限，主板危险温度以 80 度为下限，CPU 温度和 MAC 温度危险温度以 90 度为下限。

功能部属

无

8.3 功能详解

基本概念

电源工作模式

电源工作在两种工作模式下：非冗余模式、N+M 冗余模式。

- 非冗余模式：又称叠加模式，电源默认工作于该模式下。在该模式下，系统总功率为所有电源功率之和。

- N+M 冗余模式：系统中总共有 (N+M) 个电源，用户配置冗余电源个数为 M，系统中当前可用电源个数为 N，系统总功率为 N 个电源输出功率之和。

供电优先级

板卡的供电优先级，即板卡上下电优先级，指示了板卡的上电和下电的优先级。优先级越高的板卡，越早被上电，越迟被下电。在系统功率不足需要对线卡下电时，优先级最低的线卡最先被下电。默认的板卡供电优先级为：管理板 > FE 卡 > VSL 卡 > 线卡，相同类型的板卡，则槽位号越小的优先级越高。用户只能改变线卡的供电优先级，范围为 1-16，其中 1 为最低优先级，16 为最高优先级。

风扇工作模式

产品上，风扇盘的正常工作模式，一共有 3 种：

- 标准模式：系统缺省模式。风扇起始转速维持在中间转速，之后随着环境温度变化而自动降低或提高转速，以达到最佳的散热效果。
- 静音模式：风扇起始转速维持在相对较低的转速，之后随着环境温度变化而自动降低或提高转速。该模式的风扇转速相对比标准模式小一些，因此可以进一步降低噪音。
- 用户自定义模式：这种模式下，系统中风扇的转速将由用户自行设置，且所有风扇盘的转速保持一致（用户自定义的值）。系统最多支持用户设置 7 个级别的风扇转速。

i 在 VSU 模式下，两个机箱的风扇工作模式保持一致。

温度阈值

温度阈值，又称温度门限，指的是在温度监控过程中，当所监控的温度值达到该设定值时，系统需要采取相应的动作。例如，假如温度到达告警温度阈值，系统需要打印 syslog 告警。如果是达到危险温度阈值，则相应板卡要被下电。

功能特性

功能特性	作用
线卡自动下电功能	在系统功率不足时，按照线卡当前供电优先级，自动下电部分板卡以满足系统功率分配需求。
板卡手动上下电控制	用户随时可以通过命令手动给板卡上下电。
风扇智能调速	风扇转速随着温度变化自动调节，以满足系统散热需求。
温度智能监控	系统自动监控每个插槽上插入板卡的温度，在超过阈值时自动告警或给板卡下电处理。

8.3.1 保存电源供电优先级

用户保存电源供电优先级后，在起机阶段，若系统功率不足以支持所有板卡上电，则按照优先级高低，优先保证高优先级板卡上电。

工作原理

在系统运行阶段，用户可以配置电源供电优先级，此功能用于下次起机上电时，保证高优先级的板卡优先上电。

用户可以通告命令 **power priority [switch devid] slot slotid prio** 来配置板卡优先级，然后，使用命令 **power priority save** 将优先级保存，重启设备后，生效

相关配置

保存电源优先级

用户可以通过 **power priority save** 命令保存电源优先级。

power priority save

 用户从外部导入配置文件后，必须保存电源优先级。

8.3.2 板卡手动上下电控制

用户可以通过 **power** 命令手动给指定板卡上下电。

 机箱主管理板自身不允许手动下电。当指定板卡为机箱的主管理板时，将提示错误。在 VSU 模式下，两个机箱的本地主管理板，都不允许手动下电。

工作原理

板卡处于上电状态时，用户可以通过 **power off** 命令将板卡下电；

板卡处于下电状态时，用户可以通过 **power on** 命令将板卡上电；

用户还可以使用 **power cycle** 命令将板卡下电一段时间后自动上电。

相关配置

无

8.3.3 风扇智能调速

当风扇工作在标准模式、静音模式下，随着环境温度的变化，风扇会自动调节转速，以达到系统散热目的。

工作原理

系统根据当前风扇的工作模式，自动为风扇指定一个默认的起始转速，之后随着环境温度的升高和下降，风扇会自动提高或降低转速，以达到散热目的，同时确保噪音不会太大。

 系统还提供了一个自定义模式，可以使所有风扇自动按照用户配置的转速运转。该模式下，风扇转速是固定的，不会随着环境温度的变化而自动调整。因此，一般推荐用户使用标准模式或静音模式，以便风扇转速随着温度的变化自动调整，从而最大限度地保护设备不因过温而出现故障。

相关配置

更改风扇工作模式

缺省情况下，风扇工作于标准模式下。用户可以通过 `fan mode` 命令更改风扇的工作模式，使风扇在标准模式、静音模式、用户自定义模式之间切换。

8.3.4 温度智能监控

正常运行过程中，主管理板会随时监控所有板卡的温度变化情况，并及时通告给用户知晓。

工作原理

系统定义了三种温度类型：主板温度、CPU 温度、MAC 温度。同时定义了两种温度阈值：告警温度和危险温度。其中，CPU 温度和 MAC 温度无告警温度，只有危险温度。进入智能监控后，系统会每隔 2 分钟轮询一次所有板卡的所有温度点，在监控到某张板卡的某个温度点达到或超过告警温度时，主管理板会产生 `syslog` 提示，同时面板 Alarm 亮黄灯。如果监控的某个温度点达到或超过危险温度，则主管理板会将该板卡下电，并产生 `syslog` 提示，同时面板 Alarm 亮红灯。

! 机箱主管理板或 VSU 模式下的全局主管理板在监控到自身温度超过危险温度阈值时，只会产生 `syslog` 并给面板 Alarm 亮红灯，但是不会自动下电。此时用户需要立即采取有效降温措施，或手动拔出主管理板散热。

相关配置

更改温度阈值配置

缺省情况下，主板温度的告警温度为 56 度，危险温度为 80 度；CPU 温度和 MAC 温度的危险温度为 100 度。

用户可以通过如下命令更改上述温度阈值：

```
threshold set temperature [switch devic] {board | cpu | mac} {warning | shutdown} temp
```

! 主板温度告警温度配置值必须低于当前危险温度，否则将配置失败。

! 主板温度危险温度配置值不能超过 90 度，CPU 温度和 MAC 温度危险温度配置值不能超过 110 度，否则将配置失败。

8.4 产品说明



- 线卡默认优先级按以下先后关系决定：

- 1) VSL 卡优先级比其它线卡高；
- 2) 线卡槽号越小，优先级越高。



- S7508 产品的板卡供电优先级，缺省如下所示：
管理板优先级最高；

FE 卡优先级比 VSL 卡和线卡高；
VSL 卡优先级比其他线卡高；
线卡和 FE 卡槽号越小，优先级越高。

- 当某槽位插入非智能卡时，优先级配置对该槽位不生效，该槽位板卡将会自动上电，无法由用户配置控制。



show power 命令在各产品上显示的信息有差异：

- 在 S7508 产品上可以读取到如下信息：电源冗余模式；每个槽位电源的型号、在位状态、额定功率、输出功率、输出电流、输入/输出电压、Fail/告警状态；系统可用的总功率、已经分配使用的功率、剩余可用的功率；每个槽位板卡的名称、需求功率、已分配功率、每个槽位的供电状态。



- S7508 产品包含的温度点有：板卡入风口、板卡出风口、板上最热点（每张卡 2 个）、CPU 温度（多业务卡 2 个）、MAC 温度（引擎无 MAC，显示为“N/A”，部分板卡有多个 MAC 芯片，每个 MAC 芯片显示一个 MAC 温度值）。
- **show temperature** 命令在 S7508 产品则支持告警温度和危险温度的定义和显示。



电源节能：

- 在 S7508 产品上不支持电源节能功能。



风扇节能：

- 在 S7508 产品上不支持电源节能功能。



取消板卡功率不足自动下电策略：

- 在 S7508 产品上不支持当功率不足的时候，对板卡进行自动下电，仅会 SYSLOG 告警提示用户。

8.5 配置详解

配置项	配置建议 & 相关命令	
配置线卡上下电	可选配置。用于执行线卡上下电相关配置和功能。	
	power priority	配置线卡的供电优先级。
	power	给指定板卡执行上下电操作。
	power cycle	给指定板卡先下电再上电。用户可以指定下电后再次上电的时间间隔，缺省为 1 秒。
配置风扇工作模式	可选配置，用于更改风扇的工作模式。	
	fan mode	配置风扇工作于标准模式、静音模式或用户自定义模式。

配置温度阈值	 可选配置，用于更改温度阈值。	
	threshold set temperature	配置主板温度的告警温度、危险温度或 CPU、MAC 温度的危险温度。

8.5.1 配置线卡上下电

配置效果

- 更改线卡上下电优先级。
- 给指定板卡执行上下电操作。
- 给指定板卡执行先下电再上电的操作。

注意事项

- 给指定板卡下电时，不能指定为机箱主管理板。在 VSU 模式下，包括两个机箱的本地主管理板。

配置方法

- 可选配置。
- 线卡上下电优先级可按默认优先级，也可以通过该配置项更改。
- 给指定板卡上下电的操作，非配置命令，在特权模式下执行。

检验方法

- **show power priority** 命令可以查看线卡上下电优先级，以及线卡自动下电功能是否打开。
- **show power** 命令可以查看每张板卡当前的供电状态。

相关命令

配置线卡的供电优先级

【命令格式】 **power priority [switch devid] slot slotid prio**

【参数说明】 **switch devid**：仅在 VSU 模式下支持，用于指定需要配置上下电优先级的板卡槽位所属机箱号。缺省为本机箱。

slot slotid：指定要配置的线卡槽位号。线卡槽位号的范围，按照机箱类型来区分，分别为 3 槽、5 槽、8 槽和 12 槽。

prio：指定要配置的线卡优先级，范围为 1-16，1 为最低优先级，16 为最高优先级。

【命令模式】 全局模式

【使用指导】 此命令用于改变线卡和 VSL 卡的默认供电优先级，FE 卡的优先级只能使用默认优先级，无法改变。

给指定板卡执行上下电操作

【命令格式】 **power {on | off} [switch *devid*] slot *slotid***

【参数说明】 **on**：给指定板卡上电。

off：给指定板卡下电。

switch *devid*：仅在 VSU 模式下支持，用于指定需要上下电的板卡槽位所属的机箱号。缺省为本机箱。

slot *slotid*：指定要执行上下电操作的板卡槽位号。管理板槽位号为 M1 和 M2、FE 卡槽位号为 FE1、FE2、FE3、FE4。

【配置模式】 全局模式

【使用指导】 -

给指定板卡先下电再上电

【命令格式】 **power cycle [switch *devid*] slot *slotid* [interval *seconds*]**

【参数说明】 **switch *devid***：仅在 VSU 模式下支持，用于指定需要上下电的板卡槽位所属的机箱号。缺省为本机箱。

slot *slotid*：指定要执行上下电操作的板卡槽位号。管理板槽位号为 M1 和 M2、FE 卡槽位号为 FE1、FE2、FE3、FE4。

interval *seconds*：指定下电后再上电的时间间隔。缺省为 1 秒钟。

【配置模式】 全局模式

【使用指导】 -

常见配置错误

无

8.5.2 配置风扇工作模式

配置效果

- 风扇工作模式可以在标准模式、静音模式、用户自定义模式之间切换，系统默认为标准模式。

注意事项

-

配置方法

- 可选配置。
- 风扇默认工作在标准模式下，可切换为静音模式或用户自定义模式。

检验方法

- **show fan** 命令可以查看所有风扇盘的工作模式。
- **show fan detail** 命令可以查看每个风扇盘内部小风扇的实际转速。

相关命令

配置风扇工作模式

【命令格式】 **fan mode {normal | quiet | {defined [speed-level level]}}**

【参数说明】 **normal**：配置风扇工作于标准模式（Normal Mode，缺省工作模式）。

quiet：配置风扇工作于静音模式（Quiet Mode）。

defined：配置风扇工作于用户自定义模式（User Defined Mode）。在自定义模式下，机箱风扇盘中每个风扇的转速都保持一致，不会因为系统温度的变化而智能调整。因此，这种模式不推荐用户使用。

speed-level level：在用户自定义模式下，指定风扇的转速级别。支持用户配置的风扇等级为 7 级（level 取值为 1-7），缺省为配置到第 3 级。

【配置模式】 全局模式

【使用指导】 -

配置举例

配置风扇工作在静音模式

- 【配置方法】
- 配置风扇工作模式为静音模式

```
Ruijie(config)#fan mode quiet
```

- 【检验方法】 **show fan** 命令可以查看风扇当前工作模式。

- 检查风扇工作模式切换是否生效。

```
Ruijie#show fan
Chassis-type: RG_S8605E
Fan-id: 1
  Fan-type:      M05_FAN
  Serial Number: 1234567890123
  Energy-saving: off

fan-id  status  mode      speed-level
```

- 【配置方法】
- 配置风扇工作模式为静音模式

```
Ruijie(config)#fan mode quiet
```

【检验方法】 **show fan** 命令可以查看风扇当前工作模式。

- 检查风扇工作模式切换是否生效。

```
-----  
1      ok      quiet      N/A
```

常见配置错误

- 用户在温度变化较大的环境中配置了自定义模式，结果风扇无法智能调速，设备散热效果不佳。

8.5.3 配置温度阈值

配置效果

- 更改默认的告警温度和危险温度阈值。

注意事项

- 温度阈值按照机箱配置，同一个机箱内所有板卡使用同一套阈值。

配置方法

- 可选配置。
- 用户可以使用系统默认温度阈值，也可以使用此命令更改。

检验方法

- **show temperature** 命令可以查看每张板卡的温度阈值以及当前温度值。

相关命令

▾ 配置板卡温度阈值

【命令格式】 **threshold set temperature [switch devid] {board | cpu | mac} {warning | shutdown} temp**

【参数说明】 **switch devid**：仅在 VSU 模式下支持，用于指定需要设置温度门限值的板卡槽位所属机箱号。缺省表示本机。
board：指定配置主板温度（包括入风口、出风口、板上最热点的温度）的温度阈值。所有板卡的主板温度阈

值保持一致。

cpu：指定配置 CPU 温度阈值。所有板卡 CPU 温度阈值保持一致。

mac：指定配置 MAC 温度阈值。所有板卡 MAC 温度阈值保持一致。

warning：指定设置板卡的告警温度阈值。当配置的温度点类型为 **cpu** 或 **mac** 时，该关键字不可见。

shutdown：指定设置板卡的危险温度（即下电温度）阈值。

temp：指定要设置的温度阈值。

【配置模式】 全局模式

【使用指导】 -

配置举例

配置 VSU 设备主板温度的告警温度为 75 度

【配置方法】

- 配置机箱 1 主板温度的告警温度阈值
- 配置机箱 2 主板温度的告警温度阈值

```
Ruijie(config)#threshold set temperature switch 1 board warning 75
```

```
Ruijie(config)#threshold set temperature switch 2 board warning 75
```

【检验方法】

使用 **show temperature** 命令，可以查看当前板卡的告警温度和危险温度阈值。

- 检查主板温度的告警温度，确认告警温度设置是否生效。

```
Ruijie#show temperature
```

```
Switch 1: RG_N18010
```

```
Switch 2: RG_N18010
```

slot	card_type	warning (C)	shutdown (C)	current (C)
1/1	N/A	N/A	N/A	N/A
1/2	M18000-48GT-CB	75	80 100 100	30 31 37 40 (N/A) (N/A)
1/3	N/A	N/A	N/A	N/A
1/M1	M18010-CM	75	80 100 100	22 31 25 31 (35) (N/A)
1/M2	M18010-CM	75	80 100 100	23 29 25 29 (34) (N/A)
2/1	M18000-10QXS-CB	75	80 100 100	24 26 28 30 (35) (64)

.....

常见配置错误

- 设置的温度阈值超过系统允许的范围，导致设置失败。
- 设置的告警温度阈值太低，系统频繁打印告警日志。

8.6 监视与维护

查看运行情况

作用	命令
查看电源信息、板卡供电优先级等。	show power [priority version]
查看风扇信息。	show fan [{{{<i>dev</i>id} <i>fan</i>id} detail} version]
查看板卡当前温度和温度阈值等信息。	show temperature

9 软件授权管理

9.1 概述

软件授权是用户使用设备某些扩展功能的中间环节，用户需要在安装上正确的授权文件之后才能使用相应的扩展功能。RGOS 提供的一些扩展功能，包括 VSD、TRILL、FCOE、SSLVPN 支持的并发用户数量。需要经过授权方可使用其功能或使用其扩展功能。

- i** 通常，RGOS 所有特性均在出厂时均已安装于设备之中，但某些特性在获取授权前不能使用。
- i** 各项特性是否需要授权、授权类型，在各自的配置指南中说明。未加特别说明时，系统的基础功能，是不需要授权的。

9.2 典型应用

9.2.1 VSD

应用场景

VSD 是一种网络系统虚拟化技术，通过虚拟化将一台物理设备划分为多台逻辑设备。最大限度地利用现有资源，降低了网络运营成本。

在未授权前，只能使用 VSD0（就是普通的 CLI）；授权之后，才能创建其他 VSD。

功能部署

授权模块只作用于 VSD0 下，其他 VSD 上不能按照授权文件。

9.2.2 FCoE

应用场景

FCoE 技术可以将光纤通道映射到以太网，将光纤通道信息插入以太网信息包内，从而让服务器-SAN 存储设备的光纤通道请求和数据可以通过以太网连接来传输，而无需专门的光纤通道结构，从而可以在以太网上传输 SAN 数据。

设备未安装 FCoE 授权时，FCoE 相应的功能不能使用。安装 FCoE 授权后，用户方能正常使用该功能。

功能部署

9.3 功能详解

基本概念

授权特性

通过授权文件、硬件实体或法律合同等手段获得某项特殊特性的使用许可，该许可规定了允许使用的用户数量、允许使用的实例数目以及允许使用的时间等。

授权软件

授权后才能使用的软件功能。

主机号

每个设备区别于其他设备的唯一的序列号。

 在 VSU 环境下，授权可使用任何一台机箱的主机号，但当这台机箱退出 VSU 环境时，相应的授权特性自然随着相应机箱的退出而失效。

按产品授权

一个授权只能用于一台设备，匹配主机号，不支持授权的迁移。

购买凭证

购买授权时的凭证。凭证内包括产品授权码及下载授权的网址。

产品授权码 (PAK)

购买凭证的合法所有者登录到下载授权的网址（购买凭证上有列举），利用 PAK 进行注册，之后网页就会提供相应的授权文件供下载或直接发送到注册的邮箱。

授权文件

授权文件用以在设备上安装，设备上安装授权文件后，用户即能使用相应的功能。每个授权文件包含数字签名，以防止篡改。授权文件是按产品授权的。

叠加授权

各项授权可叠加在某设备上使用。例如，设备同时有 FCoE 及 TRILL 功能，用户可以同时购买 FCoE 及 TRILL 的授权在同一台设备上使用。

临时授权

该授权有时效性，一段时间后就失效。

评估授权

临时授权的一种，通常是在出厂时已经安装于设备内，主要是提供给用户试用某特性的功能。该授权不依赖于主机号。

永久授权

该授权的使用没有时间限制。

单实例授权

设备上一个特性同一时刻最多只能安装一份授权文件。

! 目前，VSD、TRILL、FCoE、SSLVPN 支持的并发用户数量等特性均是单实例授权。

多实例授权

设备上一个特性可以同时安装多份授权文件。

! 目前只有 AP 支持的用户数量这一特性是多实例授权。

友好周期警告

友好周期警告是在授权失效数日前通过日志及 TRAP 形式进行通告。设置友好周期警告能防止具有时效的授权在超期后突然停止工作，影响网络运行。

功能特性

功能特性	作用
授权获取及使用	介绍授权的获取及使用。
备份、更新、卸载授权	介绍备份授权、更新授权、卸载授权。

9.3.1 授权获取及使用

授权使用

正式授权需要通过锐捷的网站和销售途径来获取授权，授权是基于每台设备的。依照购买凭证上的网址登陆，按照提示提供产品授权码和主机号，便能获得授权，授权通过直接下载或邮件的形式授予用户。得到授权之后，用户需要对授权文件进行安装。安装成功之后，用户就可以使用该授权所对应的特性。

RGOS 授权方式包括永久授权和临时授权（评估授权是其中一种），当用户第一次启用临时授权使用某特性时，便开始进行授权计时，授权时效过后，特性将被禁用。若想继续使用，用户可以登录锐捷的网站或销售途径通过购买手段获得授权（永久授权或临时授权）。

! 在 VSU 环境下，多台设备构成一台“虚拟设备”运行，这种运行场景下，只要有一台设备获取了某项特性的授权，整个 VSU 系统就获得了该项特性的授权，但在这些设备再次分开使用时，授权检查仍是基于设备的。

! RGOS 提供的 VSD 功能，能够将一台 RGOS 设备划分为多个“虚拟设备”，对于外部用户而言，每台虚拟设备看起来就像一个独立设备一样。在这种环境下，授权的管理是全局的，即一旦某项特性经过授权，所有 VSD 域的该特性都获得授权；并且，授权只能在 VSD 的缺省 VSD 中安装及管理。

授权检查

设备运行之后，各个授权特性要进行授权的验证检查。若对应的授权被正确安装，那么该特性的功能便处于就绪状态，用户可以使用；否则，该特性的功能便失效，用户无法使用。

 各种特性的授权检查时机有所不同，有些只在启动时检查、有些是实时检查的。

丢失授权文件

授权文件存储于设备内/data 中，在软件升级中不会丢失。

在产品使用过程中，若出现存储器件损坏或文件系统损坏等情况，在系统修复后，如果此前备份了系统中的授权文件，重新安装这些备份的授权文件即可。如果授权文件未进行备份，则还可以登录锐捷的授权网站，重新获取授权，这一过程与新获得授权过程没有差异（不需要重新购买）。

授权是基于设备的，一项授权给指定设备，授权文件只能在该设备上验证通过。在维修或调换设备时，可能出现主机号发生变化，不再识别此前获取的授权文件，如果出现这种情况，需联系锐捷的产品售后服务人员。

9.3.2 备份、更新、卸载授权

备份授权

当设备发生文件系统存储介质损坏等故障并修理恢复之后，设备上的授权可能已经丢失。所以提前备份系统里的授权文件，以便在故障之后能够重新安装之前的授权。

更新授权

若系统已有的授权不能满足用户对该特性的使用需求，那么用户可以自行登录锐捷的网站购买满足需求的授权，之后在本地进行授权的更新。

卸载授权

当用户不需要某项特性的时候，为了提高内存等资源的利用率，可以选择卸载该特性授权。卸载后如果需要重新使用，再次安装授权文件即可。建议妥善保存授权文件。

9.4 配置详解

配置项	配置建议 & 相关命令	
软件授权的基本功能	 必须配置。用于授权功能的使用。	
	<code>show license hostid</code>	获取设备主机号。
	<code>license install</code>	安装授权。
备份授权文件	 可选配置。用于将授权文件进行备份。	
	<code>license copy</code>	备份授权文件
友好周期警告	 可选配置。用于设置授权过期的提前告警时间。	
	<code>license grace-peroid</code>	设置友好周期警告

更新授权	 可选配置。用于更新在设备已存在的授权。	
	license update	更新授权
卸载授权	 可选配置。用于卸载设备中已有的授权。	
	license uninstall	卸载授权
解除绑定授权	可选配置。用于解除绑定设备中已有的授权。	
	license unbind	解除绑定授权

9.4.1 软件授权的基本功能

配置效果

安装授权文件，使对应功能生效。

注意事项

- 在指定的网址下载到授权文件后，授权文件需要上传到设备上（也可以放于 usb 设备中）进行安装。安装的过程，设备无需联网。
- 不同设备不能共用授权文件。
- 授权文件在设备上安装后，系统内部自动在/data 下进行了备份（授权文件是以“.lic”后缀结尾）。卸载授权时，该备份文件也会删除。

配置方法

📌 获取设备的 host-id

- 必须配置。
- host-id 是主机号。
- 若未授权情况下使用某个授权特性的功能，CLI 会进行提示，说明该特性还未经授权，不能使用，并给出下载授权文件的网址。

【命令格式】 **show license hostid**

【参数说明】 -

【命令模式】 特权模式

【使用指导】

 授权文件名不可修改。

 此命令不需要授权。

▾ 获取产品授权码

- 必须配置。
- 产品授权码通常在用户购买授权后，锐捷公司会提供纸质的购买凭证。

▾ 通过 RUIJIE WEB URL，获取产品授权文件

- 必须配置。
- 在登录至授权网站后，按照提示操作，即可获取授权文件。

▾ 将授权文件复制到设备的文件系统中

- 必须配置。
- 采用常规的文件系统操作命令，例如，采用 TFTP 下载、U 盘复制等方式。

▾ 安装授权文件

- 可选配置。使用 **license install** 安装授权文件。

! 在 VSU 环境下使用 **license install** 命令，只要有一台设备获取了某项特性的授权，整个 VSU 系统就获得了该项特性的授权，但在这些设备再次分开使用时，授权仍是基于设备的。

! VSU 环境下，CM 管理板上，使用 **license install** 命令，那么其他各板都安装了授权；非 CM 管理板，使用 **license install** 命令，授权只在自己本地安装。

! VSD 环境下，授权的管理是全局的，即一旦某项特性经过授权，所有 VSD 域的该特性都获得授权；并且，授权只能在 VSD 的缺省 VSD 中安装及管理。

- 【命令格式】** **license install { flash: | usb0: } filename**
- 【参数说明】** **flash:** 指定授权文件在内部 flash 文件系统的位置
usb0: 指定授权文件在 USB 文件系统的位置
filename: 指定授权文件名字
- 【缺省配置】** -
- 【命令模式】** 特权模式
- 【使用指导】** -

检验方法

- 使用 **show license all-license** 命令，查看展示信息中是否有对应的授权功能名来确定是否正确安装。

- 【命令格式】** **show license { all-license | file [license] }**
- 【参数说明】** **all-license:** 显示设备中所有已安装授权文件的列表。
file filename: 显示一个指定授权名的信息。
- 【命令模式】** 特权模式
- 【使用指导】** 用户可以通过查看特性名称，查看该授权文件是否安装成功。
- 【命令展示】**

```
Ruijie#show license all-license
Searching license in the system...
```

```
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCoE00000012268889
2. Service name: LIC-VSD
   Attribute: Temporary, Single_instance, Releasable
   Left days: 362
   Licensed serial number: LIC-VSD00000012268888
```

配置举例

非 VSU 环境安装 VSD 功能授权文件

【网络环境】 要开通 VSD 功能。

- 【配置方法】
- 通过 **show license hostid** 命令获取设备主机号；
 - 登录授权网站注册，按照提示操作，获得特性 VSD 的授权文件 vsd.lic（这是需要用到设备主机号、产品授权码等），本例省略；
 - 将 vsd.lic 放于 usb 中，usb 连入设备；
 - 安装 vsd.lic

```
Ruijie#show license hostid
8708EH5F00042
Ruijie#license install usb0:vsd.lic
License file install success, service name: LIC-VSD.
```

【检验方法】 使用 **show license all-license**，查看展示信息中是否有对应的授权功能名来确定是否正确安装。

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-VSD
   Attribute: Temporary, Single_instance, Releasable
   Left days: 362
   Licensed serial number: LIC-VSD00000012268888
```

常见错误

- 授权文件不是本设备的
- 环境中没有匹配的设备
- 授权文件已安装
- 系统已安装有更新的授权文件

9.4.2 备份授权文件

配置效果

- 备份系统上的一个或所有特性的授权文件。

注意事项

- 不能备份评估版本的授权
- 必须有足够的存储空间放可备份授权文件。

i 一般一个授权文件占用空间 4K-10K 不等。

配置方法

▾ 备份系统的授权文件

- 备份出的授权文件是普通的常规文件
- 当备份系统中所有授权文件时，生成的是 tar 文件。

【命令格式】 **license { copy-all | copy-file filename } { flash: | usb0: } [target-filename]**

【参数说明】 **copy-all** : 备份系统上的所有永久授权

copy-file filename : 备份系统上由 *filename* 指定的授权。*filename* 可以是系统已安装的授权文件名，也可以是特性名，当 *filename* 为特性名是，就备份该特性下所有已安装的授权文件。

flash: 指定授权文件在内部 flash 文件系统的位置

usb0: 指定授权文件在 USB 文件系统的位置

filename : 指定授权文件名字

【缺省配置】 -

【命令模式】 特权模式

【使用指导】 -

▾ 将授权文件上传至其它系统保存

- 可使用文件系统管理命令将授权文件进一步保存至其它存储设备，例如 USB。

检验方法

- 可使用 **dir** 确定已生成的授权文件，并与 **show license all-license** 中各个永久授权特性的 **installed license** 字段显示的授权文件名对比，查看备份是否正确。

【命令格式】 **dir [filesystem:] [file-url]**

【参数说明】 **Filesystem** : 文件系统的 url，后面加一个冒号，可能的文件系统包括：**flash:**、**usb:**、**tmp:**。

file-url : 路径名，如果是以 "/" 开头，则表示绝对路径，否则表示相对路径。

【命令模式】 特权模式

【使用指导】 此命令不需要授权

【命令展示】

```
Ruijie#dir usb0:rg-license-lics
Directory of usb0:/rg-license-lics
  1 drwx          4096  Fri Jun 20 12:29:32 2014  .
  2 drwx          4096  Fri Jun 20 12:28:37 2014  ..
  3 -rwx          8704  Fri Jun 20 12:29:12 2014  lics.tar
1 file, 2 directories
536870912 bytes total (740,687,872 bytes free)
```

【命令格式】 **show license { all-license | file [license] }**

【参数说明】 **all-license** : 显示设备中所有已安装授权文件的列表。

file filename : 显示一个指定授权名的信息。

【命令模式】 特权模式

【使用指导】 用户可以通过查看特性名称，查看该授权文件是否安装成功。

 命令展示中 installed license 字段只有多实例授权才有，多实例授权系统内部备份的文件名由多实例授权的 id 号决定。单实例授权由于同一时刻系统内最多只有一个授权文件，所以内部授权文件备份直接以特性名命名。

【命令展示】

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCoE00000012268889
2. Service name: LIC-VSD
   Attribute: Temporary, Single_instance, Releasable
   Left days: 362
   Licensed serial number: LIC-VSD00000012268888
3. Service name: LIC-AP-64
   Attribute: Permanent, Multiple_instance, Releasable
   [Installed licenses]      [Licensed serial number]
19881021.lic                LIC-AP-6400000012264966
19881023.lic                LIC-AP-6400000012264988
```

配置举例

备份系统所有授权文件

【网络环境】 备份系统中所有授权文件，并命名为 lics.tar

- 【配置方法】
- 将 usb 连入设备；
 - 将设备上所有永久授权备份到 usb 中，命名为 lics.tar；

```
Ruijie#lic copy-all usb0:rg-license-lics/lics.tar
Success to copy all permanent license.
```

- 【检验方法】
- 可使用 **dir** 确定已生成的授权文件压缩包，解压后，与 **show license all-license** 中各个永久授权特性的 **installed license** 字段显示的授权文件名对比，查看备份是否正确。

⚠ **installed license** 字段只有多实例授权才有，多实例授权系统内部备份的文件名由多实例授权的 id 号决定。单实例授权由于同一时刻系统内最多只有一个授权文件，所以内部授权文件备份直接以特性命名。

⚠ 本例中 19881021.lic, 19881023.lic 该 id 号是内嵌在授权文件中，由于打包时是根据特性分文件夹存放，所以用户仍然能够区分某个授权文件对应哪个特性。

```
Ruijie#dir usb0:rg-license-lics
Directory of usb0:/rg-license-lics
  1 drwx          4096  Fri Jun 20 12:29:32 2014  .
  2 drwx          4096  Fri Jun 20 12:28:37 2014  ..
  3 -rwx          8704  Fri Jun 20 12:29:12 2014  lics.tar
1 file, 2 directories
536870912 bytes total (740,687,872 bytes free)
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCOE00000012268889
2. Service name: LIC-VSD
   Attribute: Temporary, Single_instance, Releasable
   Left days: 362
   Licensed serial number: LIC-VSD00000012268888
3. Service name: LIC-AP-64
   Attribute: Permanent, Multiple_instance, Releasable
   [Installed licenses]      [Licensed serial number]
19881021.lic                LIC-AP-6400000012264966
19881023.lic                LIC-AP-6400000012264988
```

常见配置错误

- 系统上不存在指定授权或指定文件
- 指定要备份的授权是临时的，不能备份

9.4.3 友好周期警告

配置效果

- 评估授权过期以前，通过日志形式给出告警，使用户可以提前采取应对

注意事项

- 每个授权特性单独设置
- 该项设置会受到设备时间调整的影响
- 永久性授权不需要设置友好周期警告。

配置方法

📄 设置一个授权的特性的友好周期警告

【命令格式】 **license grace-period filename days**

【参数说明】 **filename** : 授权特性名字
days : 提前告警的时间。

【缺省配置】 在 120 天、评估授权时间的一半两个值中，缺省值为其中的较小值。

【命令模式】 特权模式

【使用指导】 -

检验方法

- 方法 1：在授权期限小于设置的告警时间时，每隔一段时间便会出现警告的提示信息。
- 方法 2：使用 **show license** 命令查看每个授权对应的 Alarm starting point 字段的告警时间，是否与自己设置的一致。

【命令格式】 **show license { all-license | file [license] }**

【参数说明】 **all-license** : 显示设备中所有已安装授权文件的列表。
file filename : 显示一个指定授权名的信息。

【命令模式】 特权模式

【使用指导】 此命令不需要授权

【命令展示】

```
Ruijie#show license file LIC-VSD
Service name: LIC-VSD
Attribute: Temporary, Single instance, Releasable
Left days: 362
Licensed serial number: LIC-VSD00000012268888
```

配置举例

↘ 更改友好周期警告时间

【网络环境】 设备上已安装有特性 vsd 的临时授权，设置友好周期警告的天数为 100 天。

【配置方法】 ● 设置友好周期警告的天数为 100 天；

```
Ruijie#lic grace-period LIC-VSD 100
RG_LICENSE: success to set alarm starting point of license LIC-VSD.
```

【检验方法】 特性 vsd 的授权期限小于 100 天之后，每隔一定的时间便会出现友好周期警告的提示信息。

```
Ruijie#*Jun 18 10:06:36: %RG_LICENSE-4-LICENSE_DEADLINE_INFO: Service LIC-VSD will be disabled 80
days behind. In order to avoid the inconvenience to you, please log on website
http://192.168.5.227:8080/login.jsf to get a new license.
```

常见配置错误

- 设置永久授权的友好周期警告
- 预设置的特性系统中还未安装授权

9.4.4 更新授权

配置效果

- 更新系统上一个特性的授权，通常用于将评估授权更新为临时授权。

注意事项

- 一个正式的永久授权不需要更新。
 - 不能将一个新版本的授权更新为旧版本的授权
- i** 授权文件中有一个生成时间的字段，该字段的值以授权文件从网站生成的时间为准，该字段的值更接近当前时间，则该授权文件的版本更新。

配置方法

↘ 更新授权文件

- 可以使用 **license update** 来更新一个特性授权。

- 更新授权，设备无需联网。

- 【命令格式】 **license update** { **flash:** | **usb0:** } *filename*
- 【参数说明】 **flash:** 指定授权文件在内部 flash 文件系统的位置
usb0: 指定授权文件在 USB 文件系统的位置
filename : 指定授权文件名字
- 【缺省配置】 -
- 【命令模式】 特权模式
- 【使用指导】 -

检验方法

- 可使用 **show license** 命令，查看 Attribute 字段授权属性是否为永久来确定对应属性是否有更新。

- 【命令格式】 **show license** { **all-license** | **file** [*license*] }
- 【参数说明】 **all-license** : 显示设备中所有已安装授权文件的列表。
file filename : 显示一个指定授权名的信息。
- 【命令模式】 特权模式
- 【使用指导】 此命令不需要授权
- 【命令展示】

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCoE00000012268889
2. Service name: LIC-VSD
   Attribute: Temporary, Single_instance, Releasable
   Left days: 362
   Licensed serial number: LIC-VSD00000012268888
```

配置举例

更新 VSD 授权

【网络环境】 将系统上 VSD 的临时授权更新为永久授权。

- 【配置方法】
- 通过购买等手段获得 VSD 的永久授权文件 vsd_perm.lic，将 vsd_perm.lic 放于 usb 中，usb 连入设备；
 - 更新特性 VSD 的授权；

```
Ruijie#license update usb0:vsd_perm.lic
License file update success, temporary license LIC-VSD changes into permanent.
```

- 【检验方法】
- 可使用 **show license** 命令，查看 Attribute 字段授权属性是否为永久来确定对应属性是否有更新。

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCOE00000012268889
2. Service name: LIC-VSD
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-VSD00000012266666
```

常见配置错误

- 授权文件不是本设备的
- 用旧版本授权替代新版本授权。
- 同一个授权文件重复安装
- 试图用临时授权代替永久授权
- 更新前系统未安装此特性

9.4.5 卸载授权

配置效果

- 卸载系统上一个或所有特性授权。

i 当用户不需要某项特性的时候，为了提高内存等资源的利用率，可以选择卸载该特性的授权。卸载后，该特性的功能将失效。

注意事项

- 如果一个特性正在使用，授权卸载后，将在下一次特性使能或启动重启后发生作用。
- 授权卸载后，需要有授权文件方可重新安装，建议卸载前对授权文件进行备份。

配置方法

▾ 卸载系统中的一个授权

- 使用 **license uninstall** 命令卸载系统中的一个授权

【命令格式】 **license uninstall** { **all** | *license [filename]* }

【参数说明】 **all**: 卸载系统内所有授权文件

license : 要卸载的授权名

filename : 要卸载的文件名

- 【缺省配置】 -
- 【命令模式】 特权模式
- 【使用指导】 当授权相关的特性正在运行时，授权卸载后不会立即生效
授权卸载后不能恢复，建议卸载前先备份授权文件
此命令不需要授权

检验方法

- 可使用 **show license all-license** 命令，查看 **Service name** 字段，如果未出现已卸载的功能名称，则卸载成功。

- 【命令格式】 **show license { all-license | file [license] }**
- 【参数说明】 **all-license** : 显示设备中所有已安装授权文件的列表。
file filename : 显示一个指定授权名的信息。

- 【命令模式】 特权模式
- 【使用指导】 此命令不需要授权

【命令展示】

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCoE00000012268889
```

配置举例

▾ 卸载 VSD 授权

- 【网络环境】 卸载系统上的 VSD 授权。

- 【配置方法】 ● 卸载特性 VSD 的授权；

```
Ruijie#license uninstall LIC-VSD
License file uninstall LIC-VSD success.
```

- 【检验方法】 ● 使用 **show license all-license** 命令，查看 **Service name** 字段，如果未出现已卸载的功能名称，则卸载成功。

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCoE00000012268889
```

常见配置错误

- 设备上还没安装该项授权
- 在设备上不存在指定的授权文件

9.4.6 解除绑定授权

配置效果

- 解除绑定并同时卸载一个已有授权

i 当用户希望在授权网站解除一个授权码和原有设备的绑定关系，需要先在设备上解除绑定该授权码。

注意事项

- 设备上成功解除绑定之后，会获得一串验证码。该验证码在授权网站解除绑定时需要。
- 授权码在设备上解除绑定之后，对应的授权文件将不能再重新安装。

配置方法

▾ 解除绑定系统中的一个授权

- 使用 `license unbind` 命令解除绑定系统中的一个授权

【命令格式】 `license unbind pak`

【参数说明】 `pak`：要解除绑定的授权码

【缺省配置】 -

【命令模式】 特权模式

【使用指导】 已安装了哪些授权码，可通过 `show license all-license` 命令知晓
此命令不需要授权

检验方法

- 可使用 `show license all-license` 命令，查看 `Licensed serial number` 字段，如果未出现已解除绑定的授权码，说明解除绑定成功。

【命令格式】 `show license { all-license | file [license] }`

【参数说明】 `all-license`：显示设备中所有已安装授权文件的列表。

`file filename`：显示一个指定授权名的信息。

【命令模式】 特权模式

【使用指导】 此命令不需要授权

【命令展示】

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-FCoE
   Attribute: Permanent, Single_instance, Releasable
   Licensed serial number: LIC-FCoE00000012268889
```

配置举例

▾ 解除绑定授权码 LIC-VSD00000012268888

【网络环境】 已安装有授权码 LIC-VSD00000012268888 对应的授权。

【配置方法】

- 解除绑定授权码 LIC-VSD00000012268888 ；

```
Ruijie#license unbind LIC-VSD00000012268888
Success to unbind license LIC-VSD00000012268888.
The verification string is: 775719468737BA269825589557F558657575B5D5D5D785782598859765A8355855
```

【检验方法】

- 使用 **show license all-license** 命令，查看 Licensed serial number 字段，如果未出现已解除绑定的授权码，说明解绑定成功。

解绑定前：

```
Ruijie#show license all-license
Searching license in the system...
1. Service name: LIC-AP-64
   Attribute: Releasable
   [Permanent licenses]   [Licensed serial number]
   19880966.lic           LIC-AP-6400000012264966
   19880988.lic           LIC-AP-6400000012264988

   [Temporary license]   [Licensed serial number]
   19880900.lic           LIC-AP-6400000012264900
   (63 days left)

2. Service name: LIC-VSD
   Attribute: Permanent, Releasable
   Licensed serial number: LIC-VSD00000012268888
```

解绑定后：

```
Ruijie#show license all-license
Searching license in the system...
 1. Service name: LIC-AP-64
    Attribute: Releasable
    [Permanent licenses]    [Licensed serial number]
    19880966.lic             LIC-AP-6400000012264966
    19880988.lic             LIC-AP-6400000012264988

    [Temporary license]    [Licensed serial number]
    19880900.lic             LIC-AP-6400000012264900
    (63 days left)
```

常见配置错误

- 系统中没有匹配的授权码

9.5 监视与维护

验证授权配置

作用	命令
查看授权配置	show license { all-license file <i>[license]</i>}
查看正在使用的授权	show license usage
显示用于授权的设备序列号	show license hostid
显示本设备已解绑定授权的验证码	show license unbind-code

10 模块热插拔

10.1 概述

模块热插拔是机箱式设备常用的一种维护功能。

模块热插拔功能可以处理机箱式设备上各种可插拔模块（包括管理板、线卡、交换网板、多业务卡）插入机箱后的自动安装、卸载、复位、及模块信息的查看。

协议规范

10.2 典型应用

典型应用	场景描述
复位在线模块	在日常维护过程中，发现某个模块运行存在异常，复位该模块以排除故障。
清除模块配置	日常维护过程中，用户将一个插槽上的模块更换为另一种类型的模块。
清除 VSU 成员设备配置	日常维护过程中，用户将一台 VSU 成员设备上所有模块的配置清除，重新进行配置。
删除配置文件中保留的 MAC 信息	日常维护中，用户不希望继续使用 VSU 设备的 MAC 地址，需要重新进行 MAC 选举。
修改配置文件中保留的 MAC 信息	网关模式下，在使用新的交换机替换旧交换机时，使用该命令配置新交换机的 MAC 地址与旧交换机的 MAC 地址相同，从而下游设备无需修改绑定网关 MAC 配置。

10.2.1 复位在线模块

应用场景

在日常维护过程中，用户发现运行在某个槽位上模块运行存在异常，需要复位该模块以排除故障。

功能部属

用户在控制台上输入复位模块命令，复位该模块。

10.2.2 清除模块配置

应用场景

机箱式设备中，某个插槽上原来插的是 A 类型模块。日常维护过程中，用户发现 A 类型模块已经无法满足其使用需求，需要更换为 B 类型模块，又不想影响到其他模块的使用。

功能部属

用户按照以下步骤操作：

- 把该槽位上原有模块拔出。
- 在设备上输入清除模块配置命令，清除该模块配置。
- 在该槽位上插入新模块。

10.2.3 清除VSU成员设备配置

应用场景

VSU 模式下，因为业务变更导致某台成员设备的所有配置都要清除重新配置。这时用户不需逐个输入命令清除成员设备内各模块的配置，而是通过一条命令完成此项功能。

功能部属

用户按照以下步骤操作：

- 在设备上输入清除 VSU 成员设备配置命令。
- 保存配置。
- 重启 VSU，该设备的配置已经被清除。

10.2.4 删除配置文件中保留的MAC信息

应用场景

一般情况下，系统使用的 MAC 地址烧写在管理板或者机箱的 FLASH 中，但在 VSU 模式下，为避免 MAC 地址变化导致断流，系统会自动将使用的 MAC 地址保存在配置文件中，重新启动后，如果配置文件中存在有效的 MAC 地址则会优先使用。通过 `no sysmac` 命令可以删除配置文件中的 MAC 信息，重新使用默认烧写在 FLASH 中的 MAC 地址。

功能部属

用户按照以下步骤操作：

- 在设备上输入删除 MAC 地址配置命令。

- 保存配置。
- 重启 VSU，该设备的 MAC 地址已经被重新选举。

10.2.5 修改配置文件中保留的MAC信息

应用场景

在网关模式下（指系统配置了 `auth-mode gateway` 命令），一些周边设备配置了绑定网关 MAC 地址，如果更换网关设备，用户可以使用 `sysmac` 命令指定网关设备的 MAC 地址为被替换的网关设备的 MAC 地址，从而使下游绑定网关的设备不必修改配置。该命令仅在系统配置为网关模式时生效。

功能部属

用户按照以下步骤操作：

- 网关模式下，在设备上输入修改 MAC 地址配置命令。
- 保存配置。
- 重启设备，该设备的 MAC 地址已经被修改。

10.3 功能详解

功能特性

功能特性	作用
插入模块自动安装	用户插入新的模块到机箱式设备上，设备的管理软件自动安装该模块的驱动程序。
复位在线模块	复位在线的模块。

10.3.1 插入模块自动安装

用户可以在设备正在运行的情况下插入模块或拔出模块（热插拔），其他模块的工作不会受到影响。模块插入插槽后，设备的管理软件会试着自动安装（Install）这个模块的驱动程序。模块拔出后，相关的配置会被保留，可以继续对该模块进行设置。如果模块被再次插入，该模块将会自动启用，所有的配置都将自动生效。

- ❗ 这里的模块是管理板、线卡、交换网板和多业务卡的统称，管理板只允许插在管理板插槽（M1 或 M2）内工作，线卡和多业务卡允许插在线卡插槽内工作，交换网板只能插在交换网板插槽内工作。

工作原理

模块插入后，设备的管理软件会试着自动安装（Install）这个模块的驱动程序，并把这个模块的信息（如模块上的端口数量、端口的类型等）保存在设备内，供用户配置时使用。模块拔出后，模块的信息并不会被设备的管理软件清除，用户仍可利用这些信息对该模块进行配置。当模块被再次插入时，设备的管理软件会将用户对该模块的配置集中下发到该模块上，并使之自动生效。

10.3.2 复位在线模块

设备的管理软件提供复位在线模块功能，用户可以复位在线模块以修复该模块的软件故障。

 复位在线模块，可能导致设备上的部分业务中断。

工作原理

设备的管理软件接收到用户输入的复位在线模块命令以后，通过设备内的硬件或软件接口，令目标模块上的软件重新运行，硬件芯片状态恢复为刚刚上电的状态。模块复位以后，一些该模块的软件故障得以恢复。

10.4 产品说明



支持多张交换矩阵卡。交换矩阵卡用于机箱内各线卡之间的数据交换，如果设备内没有插交换矩阵卡，设备将不能工作。工作过程中，设备内所有交换矩阵卡都被拔出时，设备将自动复位，以便令网络中的其他设备将流量切换到备用链路上。

10.5 配置详解

 模块热插拔的功能是自动实现的，无需用户配置。

配置项	配置建议 & 相关命令	
清除模块与设备配置	 可选配置，以下命令用于清除配置，运行于配置模式下，用户输入命令以后，需要保存以便下次系统启动后仍可生效。	
	<code>remove configuration module [device-id]/slot-num</code>	清除模块配置。
	<code>remove configuration device device-id</code>	清除 VSU 成员设备配置。
	<code>no sysmac</code>	删除配置文件中保留的 MAC 信息。

10.5.1 清除模块与设备配置

配置效果

- 清除模块配置；
- 清除 VSU 成员设备配置；
- 删除配置文件中保留的 MAC 信息。

注意事项

无

配置方法

清除模块配置

- 可选配置。需要清除模块配置出现在：用户曾经在设备上某个槽位上插过一张板卡，但后来不使用了，拔出该板卡，删除该板卡及所属端口的配置。

【命令格式】 **remove configuration module** *[device-id/slot-num]*

【参数说明】 *device-id* : 机箱号（VSU 模式下需要输入模块所在机箱号，单机模式下不需要）
slot-num : 插槽号

【缺省配置】 -

【命令模式】 全局配置模式

【使用指导】 如果需要清除某个模块的已有配置，可以使用本命令。该命令主要用于清除不在位板卡的配置信息。

❗ 为了防止在线板卡配置中的一些防环路配置被清除后导致网络环路，对于在线板卡，禁止使用该命令。

清除 VSU 成员设备配置

- 可选配置。需要清除 VSU 成员设备配置出现在：用户曾经在 VSU 内使用了某台成员设备，后来从该 VSU 内将该设备移除，删除 VSU 内对该成员设备的所有配置。

【命令格式】 **remove configuration device** *device-id*

【参数说明】 *device-id* : 机箱号

【缺省配置】 -

【命令模式】 全局配置模式

【使用指导】 执行本命令可以清除 VSU 成员设备配置。

删除配置文件中保留的 MAC 信息

- 可选配置。需要清除配置文件中保留的 MAC 信息出现在：用户希望改变系统的 MAC 地址，使用系统重新选举的 MAC 地址信息。
- 一般情况下，系统使用的 MAC 地址烧写在管理板或者机箱的 FLASH 中，但在 VSU 模式下，为避免 MAC 地址变化导致断流，系统会自动将使用的 MAC 地址保存在配置文件中，重新启动后，如果配置文件中存在有效的 MAC 地址则会优先使用。

【命令格式】 **no sysmac**

【参数说明】 -

- 【缺省配置】 -
- 【命令模式】 全局配置模式
- 【使用指导】 执行本命令可以删除配置文件中的 MAC 信息，重新使用默认烧写在 FLASH 中的 MAC 地址。

修改配置文件中保留的 MAC 信息

- 可选配置。在需要修改设备的系统 MAC 地址时使用。
- 在网关模式下（指系统配置了 **auth-mode gateway** 命令），一些周边设备配置了绑定网关 MAC 地址，如果更换网关设备，用户可以使用 **sysmac** 命令指定网关设备的 MAC 地址为被替换的网关设备的 MAC 地址，从而使下游绑定网关的设备不必修改配置。该命令仅在系统配置为网关模式时生效。

- 【命令格式】 **sysmac** *mac-address*
- 【参数说明】 *mac-address* : 设置的 MAC 地址
- 【缺省配置】 -
- 【命令模式】 全局配置模式
- 【使用指导】 为系统指定 MAC 地址，需保存配置并重启后生效。

检验方法

使用 **show version slot** 命令查看线卡的安装信息。

- 【命令格式】 **show version slots** [*device-id* / *slot-num*]
- 【参数说明】 *device-id* : 机箱号，可选（VSU 模式下，如果要输入插槽号，必须同时输入模块所在机箱号）
slot-num : 插槽号，可选
- 【命令模式】 特权用户模式
- 【使用指导】 执行本命令可以查看模块的在线状态信息。其中 Configured Module 一列表示安装的模块信息。执行清除模块配置命令后，被清除的模块安装信息在此列将不会存在。

【命令展示】

查看模块在线状态信息

```
Ruijie# show version slots
```

Dev	Slot	Port	Configured Module	Online Module	Software Status
1	1	0	none	none	none
1	2	24	M8606-24SFP/12GT	M8606-24SFP/12GT	none
1	3	2	M8606-2XFP	M8606-2XFP	cannot startup
1	4	24	M8606-24GT/12SFP	M8606-24GT/12SFP	ok
1	M1	0	N/A	M8606-CM	master
1	M2	0	N/A	none	none

配置举例

清除不在线模块配置

【网络环境】 ● 用户环境中曾经在槽位 1 上使用过一张板卡，但后来由于组网变动，不再使用该板卡了。在整理设备配置时，希望删除该板卡所属端口的所有配置，使设备的配置文件更加简洁。

【配置方法】 ● 使用清除模块配置命令删除该板卡配置

```
Ruijie(config)# remove configuration module 1
```

【检验方法】 ● 输入 **show version slots** 命令，发现槽位 1 上板卡的配置被清除了。

```
Ruijie# show version slots
```

Dev	Slot	Port	Configured Module	Online Module	Software Status
1	1	0	none	none	none
1	2	24	M8606-24SFP/12GT	M8606-24SFP/12GT	none
1	3	2	M8606-2XFP	M8606-2XFP	cannot startup
1	4	24	M8606-24GT/12SFP	M8606-24GT/12SFP	ok
1	M1	0	N/A	M8606-CM	master
1	M2	0	N/A	none	none

常见错误

10.6 监视与维护

清除各类信息

 在设备运行过程中执行 **reset module** 命令，可能因为模块复位而导致业务中断。

作用	命令
复位某个模块	reset module slot-num reset module device-id / slot-num (VSU 模式)

查看运行情况

作用	命令
查看模块详细信息	show version module detail [slot-num] show version module detail [device-id/slot-num] (VSU 模式)
查看模块在线信息	show version slots [slot-num] show version slots [device-id/slot-num] (VSU 模式)

11 管理板冗余

11.1 概述

管理板冗余是一种通过管理板业务运行状态的实时备份（也称为热备）以提高设备可用性的机制。

在采用控制面与转发面分离结构的网络设备上，控制面运行在管理板上，转发面运行在线卡上。在设备运行过程中，主管理板的控制面信息实时备份到从管理板上，当主管理板出现计划内停机（如软件升级）或计划外停机（如软硬件异常）时，设备能够自动地快速切换到从管理板上工作，不丢失用户的相应配置，从而保证网络能够正常运行。在切换过程中，转发面能够继续进行转发工作，并且在控制面重新启动过程中，不会出现转发停止或者拓扑波动。

管理板冗余技术的实现可为网络服务提供以下便利：

1. 提高了网络的可用性

管理板冗余技术在切换过程中维持了数据转发及用户会话的状态信息。

2. 避免邻居检测到 link flap

在切换过程中转发面并未重新启动，因此邻居不会检测到链路先 Down 后 UP 的状态变化。

3. 避免 routing flaps

在切换过程中转发面维持转发通信，并且控制面快速构造新的转发表，没有明显的新旧转发表替换过程，从而避免出现 routing flaps。

4. 用户会话(user sessions)不会丢失

在切换前已建立的用户会话由于状态进行了实时同步而不会丢失。

协议规范

- 私有协议

11.2 典型应用

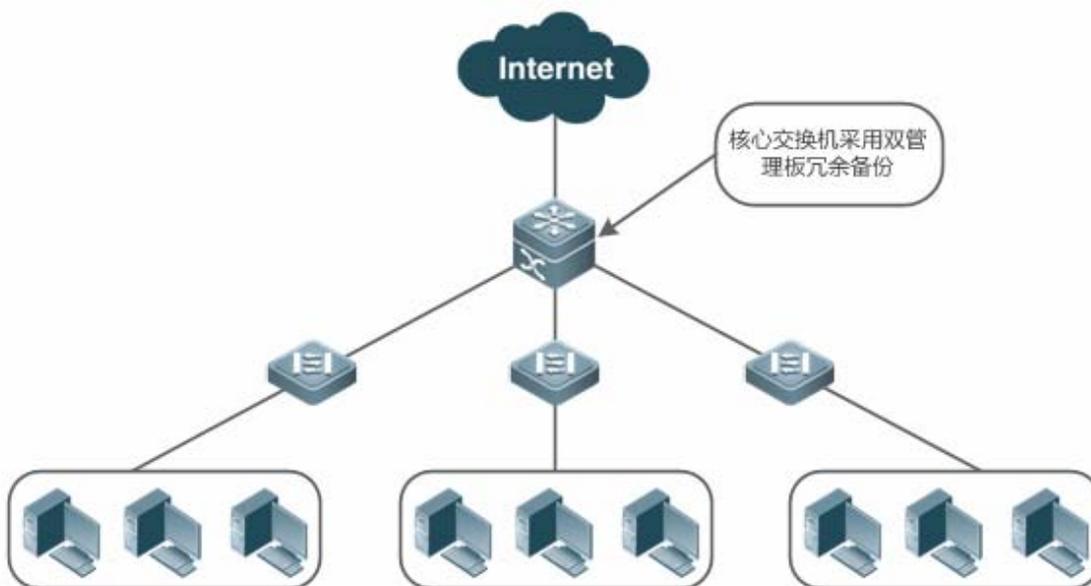
典型应用	场景描述
管理板冗余备份	在安装有管理板的核心交换机中，通过冗余备份技术，可提高网络的稳定及系统的可用性。

11.2.1 管理板冗余备份

应用场景

如下图，在该网络拓扑中，如果核心交换机出现故障，将会导致下联的各个网络瘫痪，为了提高网络的稳定性，要求核心交换机配置有两张管理板进行冗余备份。主管理板管理整个系统，从管理板实时备份主管理板的业务运行状态信息，当进行手工切换或者主管理板故障时进行的强制切换，从管理板立即接替主管理板运行，此时转发面仍然可以继续数据进行数据转发，提高了系统的可用性。

图 11-1



功能部属

对于机箱式设备，系统自带主从备份机制，只要符合冗余条件，即插即用。

对于盒式设备，每台设备相当于一张管理板和一张线卡，多台盒式设备组成的 VSU 系统也存在主从备份机制。

11.3 功能详解

基本概念

主管理板、从管理板

在安装两块管理板的设备上，系统选举其中一张管理板作为当前正常使用的管理板，称为“主管理板”；另一块作为备用的管理板，在主管理板发生故障或主动要求切换时接替成为新的主管理板，称为“从管理板”。在一般状况下，从管理板不参与交换机的管理，而是监控主管理板的运行状态。

全局主、全局从、全局候选

在两台或更多台机箱式设备组成的 VSU 系统中，对于每台机箱存在两张管理板，其中一张作为主管理板管理整个机箱，另一张作为本机箱的备份管理板；对于整个 VSU 系统，存在两张或更多以上的管理板，除选举出一块主管理板管理整个 VSU 系统和一块从管理板作为 VSU 系统备份外，其余的管理板作为候选管理板，在主从管理板失效时代替主从管理板，以主管理板或

从管理板角色来运行，称为“候选管理板”。在一般情况下，候选管理板不参与备份。为区分机箱设备内部的主从管理板和 VSU 系统中的主从管理板，我们把 VSU 系统中的主、从和候选也称为“全局主”、“全局从”和“全局候选”，管理板冗余机制作用在全局主和全局从上，因此，在下文 VSU 环境下提到的主从概念指的就是全局的主从概念。

在两台或更多台盒式设备组成的 VSU 系统中，每台盒式设备相当于一张管理板和一张线卡组成，系统也会选举一台设备作为全局主，一台设备作为全局从，其余的作为全局候选。

实现管理板冗余的前提要求

设备系统内的所有管理板必须在软件和硬件都兼容的情况下才能保持管理板冗余正常工作；

在启动过程中，主/从管理板之间需要先进行批量同步，以使两管理板达到状态一致，在此之前管理板冗余不能完全发挥作用。

管理板冗余的状态

主管理板在进行主从备份的过程中会有如下状态变化：

- alone，即单独状态。在这种状态下，系统中只有一张管理板在运行，或者是主从切换还未完成，新的主管理板和新的从管理板之间还未建立冗余备份；
- batch，即批量备份状态。主从管理板之间建立冗余备份，正在进行批量备份；
- realtime，即实时备份状态。主从管理板批量备份结束后，进入此状态，主从之间进行实时备份，只有处于该状态才能够执行手工切换。

功能特性

功能特性	作用
主从管理板的选举方式	设备可以根据系统当前的情况自动选择主从管理板，或由用户手工选择。
管理板的信息同步	在管理板冗余环境下，主管理板会实时同步状态信息和配置文件。

11.3.1 主从管理板的选举方式

工作原理

机箱式设备自动选择主从管理板

- 用户可以在设备的运行中对管理板进行插入和拔出操作。设备将根据系统当前的情况，自动选择相应的引擎运行，同时不影响正常的的数据交换，在使用过程中可能遇到以下情况，主管理板将做相应的选择：
- 如果在设备启动的时候，只插一块管理板，无论是插在 M1 槽还是 M2 槽，设备都将选择该管理板作为主管理板。
- 如果在设备启动的时候，插两块管理板，在缺省情况下将选择 M1 槽的管理板作为主管理板，M2 槽的管理板为从管理板，起备份冗余作用，并输出相应提示信息。
- 如果在设备启动的时候，只插一块管理板，在运行过程中，插入另一块管理板，无论是插在 M1 槽还是 M2 槽，后插入的管理板都将做为从管理板，起备份冗余作用，并输出相应提示信息。

- 如果在设备启动的时候，插两块管理板，在运行过程中，拔出其中一块管理板（或者其中一块管理板工作异常），如果该管理板在拔出（或者异常）之前为从管理板，则拔出（或者异常）后设备仅仅提示从管理板被拔出（或者不能运行）；如果该管理板在拔出（或者异常）之前为主管理板，那么另外一块管理板将从从管理板状态升级到主管理板状态，并输出相应提示信息。

手动选择主从管理板

- 用户可以通过手动配置选择主从管理板，基于不同的环境，采用的方式不同，具体如下：
- 在单机模式下，用户可手工执行主从切换，管理板复位后即可生效。
- 在 VSU 模式下，用户可手工执行主从切换，将全局从升为全局主，如果 VSU 系统中只有两张管理板，则原全局主复位后作为新的全局从，如果存在两张以上的管理板，则选举一张全局候选升为新的全局从，原全局主复位后作为全局候选。

相关配置

手工执行主从切换

- 缺省情况下，设备可自动选择主管理板。
- 在单机和 VSU 模式下，用户都可通过 **redundancy forcesswitch** 进行手工切换。

11.3.2 管理板信息同步

工作原理

- 状态同步

主管理板将其运行状态实时同步至从管理板，以使从管理板能够在任意时刻接替主管理板的功能，而不至产生可觉察的变化。

- 配置同步

设备运行过程中，存在两种系统的配置文件：一种是运行过程中动态生成的配置文件，会随着业务的配置而变化，称为 **running-config**；另一种是设备起机时导入的系统配置文件，称为 **startup-config**，可以通过 **write** 命令将 **running-config** 写入到 **startup-config** 中，或者通过 **copy** 命令进行复制。

对于一些与不间断转发没有直接关联的功能，只需要依靠系统配置文件的同步，就能保证用户配置在切换过程中保持一致。

在双管理板冗余下，主管理板会定时地将 **startup-config** 和 **running-config** 配置文件同步到从管理板和所有候选管理板中。在以下操作中，也会触发配置的同步：

- 1) 在用户配置从全局模式退出到特权模式时会进行 **running-config** 的同步；
- 2) 当用户使用 **write** 或 **copy** 命令保存配置时会进行 **startup-config** 的同步；
- 3) 采用 SNMP 进行的配置不会进行自动同步，需要由 CLI 配置方式触发 **running-config** 的同步。

相关配置

- 缺省情况下，**startup-config** 和 **running-config** 每小时自动同步一次。

- 通过 `auto-sync time-period` 命令可调整主管理板同步配置文件的时间间隔。

11.4 配置详解

配置项	配置建议 & 相关命令	
配置手工主从切换	⚠️ 可选配置。	
	<code>show redundancy states</code>	查看热备状态。
	<code>redundancy forceswitch</code>	手工执行主从切换。
配置自动同步周期	⚠️ 可选配置	
	<code>redundancy</code>	进入冗余配置模式。
	<code>auto-sync time-period</code>	配置双管理板冗余时自动同步配置文件的时间间隔。
复位管理板	⚠️ 可选配置	
	<code>redundancy reload</code>	复位从管理板或同时复位主从管理板。

11.4.1 配置手工主从切换

配置效果

原主管理板将被复位，从管理板成为新的主管理板。

如果系统中存在多于两张管理板的情况下，原从管理板升为主管理板，并从候选管理板中选举出一张作为新的从管理板，原主管理板复位后作为候选管理板。

注意事项

为了确保切换过程中，数据转发不受影响，主/从管理板之间需要先进行批量同步，以使两管理板达到状态一致，即管理板冗余处于实时备份状态时，才能进行手工切换。另一方面，为确保配置文件同步的完整性，在同步期间，业务模块会临时禁止手工主从切换。因此，执行手工切换需要同时具备以下条件：

- 在主管理板上执行，且存在从管理板；
- 系统中所有 VSD 下的热备都处于实时状态；
- 系统中所有 VSD 下的热备切换都未被业务模块临时禁止。

如果设备虚拟成多个 VSD 时，要求所有 VSD 的管理板冗余处于可切换状态时，才能够成功执行手工切换。

配置方法

- 可选配置。
- 在主管理板上配置。

检验方法

- 通过 `show redundancy states` 命令可查看主从管理板是否切换。

相关命令

查看热备状态

【命令格式】 **show redundancy states**

【参数说明】 -

【命令模式】 特权或全局配置模式

【使用指导】 -

手工执行主从切换

【命令格式】 **redundancy forceswitch**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 -

配置举例

手工主从切换

【配置方法】 在已配置有一个名字为 `staff` 的 VSD 环境下，进行主从切换

```
Ruijie> enable
Ruijie# show redundancy states
Redundancy role: master
Redundancy state: realtime
Auto-sync time-period: 3600 s

VSD staff redundancy state: realtime
Ruijie# redundancy forceswitch
This operation will reload the master unit and force switchover to the slave unit. Are you sure to
continue? [N/y] y
```

【检验方法】 在原从管理板上，通过 **show redundancy states** 命令查看

```
Ruijie# show redundancy states
Redundancy role: master
Redundancy state: realtime
Auto-sync time-period: 3600 s

VSD staff redundancy state: realtime
```

常见错误

-

11.4.2 配置自动同步周期

配置效果

- 改变配置文件 startup-config 和 running-config 的自动同步周期，较短的同步周期可以使得用户改变配置时，较频繁的同步到其余管理板上，避免出现主管理板故障强制切换到从管理板时产生的配置丢失。

注意事项

-

配置方法

- 可选配置，在需要改变同步周期时进行配置。
- 在主管理板上配置。

检验方法

- 通过查看时打印的 syslog 确认是否有进行定时同步

相关命令

↘ 进入冗余配置模式

- 【命令格式】 **redundancy**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

↘ 配置配置文件自动同步时间周期

- 【命令格式】 **auto-sync time-period value**
- 【参数说明】 **time-period value**：自动同步的周期间隔，单位为秒；时间范围从 1 秒到 1 个月（2678400 秒）。
- 【命令模式】 冗余配置模式
- 【使用指导】 配置双管理板冗余时自动同步 startup-config 和 running-config 配置文件的时间间隔

配置举例

配置自动同步周期

【配置方法】 在主管理板的冗余配置模式下配置自动同步周期为 60 秒：

```
Ruijie(config)# redundancy
Ruijie(config-red)# auto-sync time-period 60
Redundancy auto-sync time-period: enabled (60 seconds).
Ruijie(config-red)# exit
```

【检验方法】 通过 **show redundancy states** 命令查看

```
Ruijie# show redundancy states
Redundancy role: master
Redundancy state: realtime
Auto-sync time-period: 3600 s
```

常见错误

11.4.3 复位管理板

配置效果

- 仅复位从管理板不会影响数据转发，在从管理板复位期间不会产生转发中断或用户会话信息的丢失。
- 在单机模式下，执行 **redundancy reload shelf** 将会导致整机箱管理板和线卡同时复位，在 VSU 模式下，将会复位指定编号的设备，如果系统中存在 2 台或更多的设备，且复位的是全局主所在的设备，则系统会进行主从切换。

注意事项

- 在 VSU 模式下，如果系统热备没有进入实时备份状态，复位全局主所在的设备，将会导致整个 VSU 系统复位。

配置方法

- 可选配置，在发现管理板或设备运行异常时进行复位。

检验方法

相关命令

- 【命令格式】 **redundancy reload {peer | shelf [switchid]}**
- 【参数说明】 **peer** : 仅复位从管理板
shelf [switchid] : 单机模式下对主从管理板都进行复位；VSU 模式下需指定待复位的设备号
- 【命令模式】 特权模式
- 【使用指导】 单机模式下复位设备的命令格式为 **redundancy reload shelf**，即复位整个设备；VSU 模式下复位设备命令格式为 **redundancy reload shelf switchid**，即复位指定设备号的设备。

配置举例

VSU 模式下复位设备

- 【配置方法】 在全局主的特权模式下执行复位编号为 2 的设备

```
Ruijie# redundancy reload shelf 2
This operation will reload the device 2. Are you sure to continue? [N/y] y
Preparing to reload device 2!
```

- 【检验方法】 观察相应的管理板或设备是否重启

常见错误

11.5 监视与维护

清除各类信息

查看运行情况

作用	命令
查看当前双管理板冗余状态。	show redundancy states

查看调试信息

12 USB

12.1 概述

USB (Universal Serial Bus , 通用串行总线) 是一种外部总线标准 , 这里指的是遵循 USB 标准的外围设备 , 如 U 盘。

USB 是一种热拔插产品 , 用户可以使用它 , 将通信设备内的文件 (例如配置文件、日志文件等) 数据方便地拷贝复制出来 , 也可以将外部的数据 (例如系统升级文件) 拷贝至设备内部存储器。

USB 的具体应用场景请参见不同功能的配置指南 , 这里只介绍 USB 产品的识别、查看、使用、卸载等。

12.2 典型应用

典型应用	场景描述
利用USB升级设备	U 盘中存放着升级文件 , 设备上电后 , 检测到 U 盘 , 执行升级命令 , 从中加载升级文件 , 加载成功后 , 设备复位 , 运行升级的新版本。

12.2.1 利用USB升级设备

应用场景

U 盘中存放着升级文件 , 设备上电后 , 检测到 U 盘 , 执行升级命令 , 从中加载升级文件 , 加载成功后 , 设备复位 , 运行升级的新版本。如 :

```
upgrade usb0:/s12k-ppc_11.0(1B2)_20131025_main_install.bin。
```

执行上面的命令 , 若文件有效 , 命令执行成功 , 设备会自动复位 , 运行新的版本。

功能部属

- 用 `usb0:/` 前缀 , 访问 USB 0 设备 , `show usb` 可以看到 id 为 0 的设备相关信息。
- `upgrade` 命令是升级功能相关命令。

12.3 功能详解

▾ USB 的使用

将 USB 产品插入 USB 插槽 , 系统会自动查找 USB 产品 , 找到后驱动模块会自动对它进行驱动初始化 , 成功初始化 USB 产品之后 , 自动加载里面的文件系统 , 后面就可以读写这个 USB 产品了。

- 系统找到 USB 产品并加载驱动成功的话，会打印出如下的提示信息：

```
*Jan 1 00:09:42: %USB-5-USB_DISK_FOUND: USB Disk <Mass Storage> has been inserted to USB port 0!
*Jan 1 00:09:42: %USB-5-USB_DISK_PARTITION_MOUNT: Mount usb0 (type:FAT32),size : 1050673152B(1002MB)
```

- **i** Mass Storage 是找到的设备的名称，“usb0:”是指第 1 个 USB 设备。Size 是分区的大小，如上，该 u 盘有 1002MB 的空间。

▾ USB 的卸载

使用 CLI 命令先将 USB 产品进行移除，以防系统正在使用设备，导致出错。

- 如果 USB 产品移除成功，会打印如下提示信息：

```
OK, now you can pull out the device 0.
```

在打印出如上面的提示信息后，方可拔出 USB 产品。

12.4 配置详解

配置项	配置建议 & 相关命令	
使用 USB	 必须配置。	
	无	
卸载 USB	 必须配置。卸载 USB 产品。	
	usb remove	移除 USB 产品

12.4.1 使用USB

配置效果

USB 产品加载到系统后，可以直接使用文件系统的命令（**dir**，**copy**，**del** 等）对 USB 产品进行操作。

注意事项

- RGOS 系统只能使用支持标准 SCSI 指令的产品（一般通用的 U 盘），非标准 SCSI 指令的产品在系统中无法使用（比如 USB 上网卡附带的 U 盘，附带虚拟 USB 光驱的 U 盘），另外有些产品配置了 USB 转串口功能。
- USB 产品仅支持 FAT 文件系统，其他文件系统的 USB 产品需要在 PC 上格式成 FAT 文件系统后，才能在设备上使用。
- RGOS 系统支持 HUB，U 盘接在 HUB 端口时，设备访问的路径不同。U 盘直接插入设备 USB 端口，设备访问路径是 `usbX:/`，其中 X 表示 device id，`show usb` 可以查看到；U 盘插在 HUB 端口，HUB 插在设备 USB 端口，设备访问路径是 `usbX-Y:/`，

其中 X 表示 device id，Y 表示 HUB 端口号，比如 `usb0-3:/` 表示设备 0 号 USB 端口上的 HUB 所属的 3 号端口。

配置方法

▾ 识别 USB

USB 产品的插入无需命令行操作，直接将其插入 USB 插槽即可。

▾ 使用 USB

如下操作是将 USB 产品中的文件复制到 flash 中。

- 使用 **cd** 命令进入 USB 产品分区
- 使用 **copy** 命令将 USB 产品中的文件复制到设备的 flash 中。
- 使用 **dir** 命令查看文件是否已添加在设备中。

-
- ❗ 当 USB 产品上有多个分区的时候，在设备上仅能访问第一个 FAT 分区。
 - ❗ USB 产品的路径不存在上层目录的概念。通过 “cd usbX:\” 进入 USB 产品后，通过 “cd flash:\” 返回 flash 文件系统。
-

检验方法

使用 **show usb** 命令查看已插入的 USB 产品的信息。

相关命令

无

配置举例

▾ 使用 USB

【网络环境】 单机环境

- 【配置方法】
- 将 U 盘插入设备 USB 插槽。
 - 在设备控制台执行 **show usb** 命令。
 - 将 U 盘里的 config.txt 文件拷贝至设备 flash。

```
Ruijie#show usb
Device: Mass Storage
ID: 0
URL prefix: usb0
Disk Partitions:
usb0(type:vfat)
```

```

Size:15789711360B(15789.7MB)

Available size:15789686784B(15789.6MB)

Ruijie#

Ruijie#

Ruijie#dir usb0:/

Directory of usb0:/

  1 -rwx          4 Tue Jan  1 00:00:00 1980  fac_test
  2 -rwx          1 Mon Sep 30 13:15:48 2013  config.txt

2 files, 0 directories

15,789,711,360 bytes total (15,789,686,784 bytes free)

Ruijie#

Ruijie#

Ruijie#copy usb0:/config.txt flash:/

Copying: !

Accessing usb0:/config.txt finished, 1 bytes prepared

Flushing data to flash:/config.txt...

Flush data done

Ruijie#

Ruijie#

```

【检验方法】

- 查看 flash 中是否有 config.txt 文件

```

Ruijie#

Ruijie#dir flash:/

Directory of flash:/

  1 drw-          160 Wed Mar 31 08:40:01 2010  at
  2 drwx          160 Thu Jan  1 00:00:11 1970  dm
  3 drwx          160 Thu Jan  1 00:00:05 1970  rep
  4 drwx          160 Mon Apr 26 03:42:00 2010  scc
  5 drwx          160 Wed Mar 31 08:39:52 2010  ssh
  6 drwx          224 Thu Jan  1 00:00:06 1970  var
  7 d---          288 Sat May 29 06:07:45 2010  web
  8 drwx          160 Thu Jan  1 00:00:11 1970  addr
  9 drwx          160 Sat May 29 06:07:44 2010  cwmp
 10 drwx          784 Sat May 29 06:07:47 2010  sync
 11 --w-          92 Tue Feb  2 01:06:55 2010  config_vsu.dat
 12 -rw-          244 Sat Apr  3 04:56:52 2010  config.text

```

```
13 -rwx          1 Thu Jan  1 00:00:30 1970  .issu_state
14 -rw-          0 Tue Feb  2 01:07:03 2010  ss_ds_debug.txt
15 -rw-        8448 Thu Jan  1 00:01:41 1970  .shadow
16 -rwx         268 Thu Jan  1 00:01:41 1970  .pswdinfo
17 -rw-          4 Tue May 25 09:12:01 2010  reload
18 drwx         232 Wed Mar 31 08:40:00 2010  snpv4
19 drwx        6104 Sat May 29 06:10:45 2010  .config
20 ----          1 Thu Jan  1 00:04:51 1970  config.txt
21 d---         160 Thu Jan  1 00:00:12 1970  syslog
22 drwx         160 Tue May 25 03:05:01 2010  upgrade_ram
23 drwx         160 Tue Feb  2 01:06:54 2010  dm_vdu
24 -rwx         16 Thu Jan  1 00:01:41 1970  .username.data
9 files, 15 directories
5,095,424 bytes total (4,960,256 bytes free)
Ruijie#
```

常见错误

- 使用了非标准 SCSI 指令的产品插入设备中。
- USB 产品为非 FAT 文件系统，则系统无法识别。

12.4.2 卸载USB

配置效果

卸载 USB 产品，并使 USB 产品及设备保持完好。

注意事项

- 一定要先执行移除命令移除设备后再拔出，以免系统出现错误。

配置方法

📌 移除 USB

- 必须配置。
- 应在每次卸载 USB 产品时先对其执行移除命令。

📌 拔出 USB

在成功执行移除命令后，直接将 USB 产品拔出即可。

检验方法

使用 **show usb** 命令查看已插入的 USB 设备的信息。

相关命令

📄 移除 USB 产品

【命令格式】 **usb remove** *device-id*

【参数说明】 *device-id* : 设备上 USB 接口号, 它是 USB 显示信息中的 ID 号。该号可以通过 **show usb** 来获取。

【命令模式】 特权模式

【使用指导】 在拔出 USB 产品之前, 需要用命令移除该产品, 以防系统正在使用而导致错误。若移除成功, 系统会打印出提示, 此时方可拔出设备。如果移除失败, 说明系统正在使用该 USB 产品, 请在系统未使用 USB 产品时再执行移除操作。

配置举例

📄 卸载 USB

【网络环境】 单机环境

- 【配置方法】
- 执行 **show usb** 命令查看 USB 设备的 ID 信息。
 - 执行 **usb remove** 命令卸载 USB 设备。

```
Ruijie#show usb
Device: Mass Storage
ID: 0
URL prefix: usb0
Disk Partitions:
usb0(type:vfat)
Size:15789711360B(15789.7MB)
Available size:15789686784B(15789.6MB)
Ruijie#
Ruijie#
Ruijie#usb remove 0
OK, now you can pull out the device 0.
```

- 【检验方法】
- 再次执行 **show usb** 命令, 查看 USB 设备是否已经卸载。 **show usb** 命令不显示 id 为 0 的设备。

```
Ruijie#show usb
Ruijie#
```

常见错误

无

12.5 监视与维护

查看运行情况

作用	命令
查看已插入的 USB 设备信息。	show usb

13 UFT

13.1 概述

UFT (Unified Forwarding Table 统一转发表) 提供了动态分配交换机硬件转发表项资源的功能。

协议规范

-

13.2 典型应用

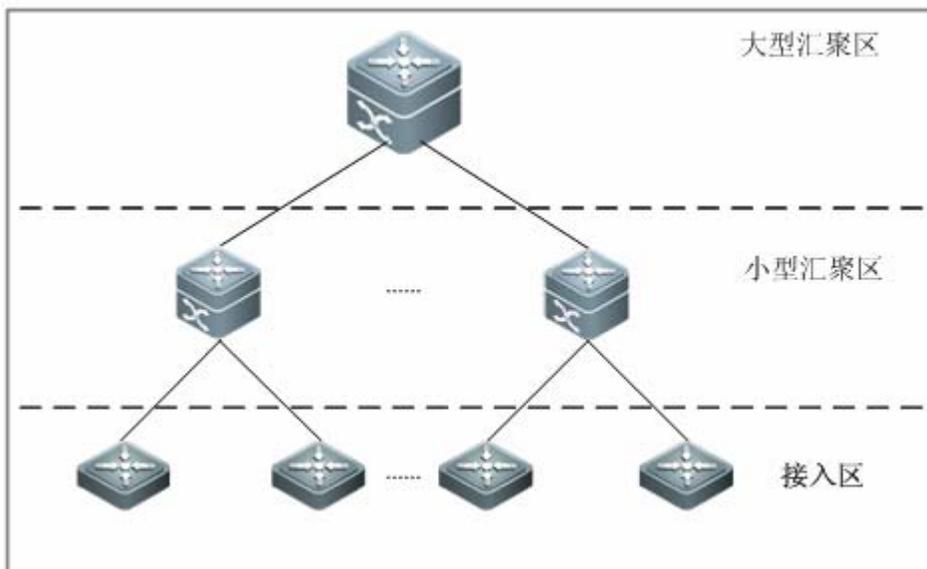
典型应用	场景描述
动态分配表项资源	当设备工作在普通路由场景下，不需要使用 MPLS 标签转发，其对应的表项容量就不会被使用到，如果这些资源能够用于其他表项，比如用于 arp/nd 表项，就可以学习到更多的 arp/nd。

13.2.1 动态分配表项资源

应用场景

下图为常见的园区网简化拓扑，核心设备可能部署在小型汇聚区作为接入设备的小汇聚，主要开启的功能为二层；或者部署在大型汇聚区域，作为大汇聚设备，主要作为网关。在小汇聚的位置上，需要有足够大的 MAC 地址表来支持。

所以另外一个主要场景为大汇聚设备，即大网关，此处设备的主要接入能力表现为产品的 ARP 与 ND 容量，即能够接入的 IPv4 与 IPv6 终端设备数量。以 Windows 7 操作系统为例，其支持 ipv4 和 ipv6 双栈，每个终端设备接入时会需要占用 ARP 与 ND 表项各一条。因此该场景下，对 arp、nd 表项容量要求非常大。



功能部属

- 将交换机配置为 UFT 模式中的 Bridge 模式，可调节转发表项支持纯 2 层模式，增大 mac 地址表容量。
- 将交换机配置为 UFT 模式中的 Gateway 模式，可调节转发表项支持大网关模式，增大 arp、nd 表项容量。

13.3 功能详解

基本概念

功能特性

功能特性	作用
UFT 工作模式	UFT 提供一种机制让用户根据不同的场景选择相应的工作模式来适配场景需求。

13.3.1 UFT工作模式

工作原理

UFT 提供一种机制让用户根据不同的场景选择相应的工作模式来适配场景需求。总共支持 8 种工作模式，都需要保存配置重启生效。

📄 Default

交换机的 UFT 模式默认配置为 Default 模式，该模式下交换机的各硬件表项能适用于大部分用户场景。

↘ Bridge

将交换机的 UFT 模式配置为 Bridge 模式，即二层转发模式，适用于纯二层业务为主的场景。该模式大量减少 ARP/ND 和 MPLS 容量，将大部分容量分配给 MAC 地址表。

↘ Gateway

将交换机的 UFT 模式配置为 Gateway 模式，即网关模式。网关模式下又细分为网关模式、大网关模式、ND 网关模式。网关模式，即 gateway 模式，用于三层业务为主的场景。大网关模式，即 gateway-max 模式、适用于大量终端的场景。ND 网关模式，即 gateway-ndmax 模式，适用于大量 v6 终端的场景。

↘ Route

将交换机的 UFT 模式配置为 Route 模式，即网络路由模式，适用于大量路由转发为主的场景。

路由模式下又细分成 route-v4max 模式，这两种模式分别将网络路由容量最大限量的分配给 IPv4。

↘ Label

将交换机的 UFT 模式配置为 Label 模式，即 MPLS 模式。该模式减少 MAC 地址表和 ARP/ND 容量，增加 MPLS 容量。

13.4 产品说明



- S7508 产品 UFT 只支持三种模式：default、gateway-max、route-v4max。
- UFT 模式配置为 gateway-ndmax 时，mac 静态地址和过滤地址的容量只支持 2K。

13.5 配置详解

配置项	配置建议 & 相关命令	
配置UFT工作模式	可选配置，用于交换机切换当前 UFT 工作模式。	
	switch-mode mode_type slot slot_num	单机模式下切换 UFT 工作模式。
	switch-mode mode_type switch switch_num slot slot_num	vsu 模式下切换 UFT 工作模式。

13.5.1 配置UFT工作模式

配置效果

- 通过配置 Bridge 模式，可以扩大二层表项的容量，主要应用在二层业务为主的情况下。

- 通过配置 Gateway 模式，可以扩大 ARP 和 ND 容量，主要应用与三层业务为主的情况下。
- 通过配置 Route 模式，该模式主要应用在大量路由转发的场景。

注意事项

- 配置完成后，需要保存配置，且需要整机重启后，配置才会生效。
- UFT 功能可能导致线卡自动重启 1 次。

配置方法

📄 单机模式下切换 UFT 模式

必须配置。

可以通过 `switch-mode mode_type slot slot_num` 命令对交换机 UFT 模式进行切换。

【命令格式】 `switch-mode mode_type slot slot_num`

【参数说明】 `mode_type`：UFT 工作模式类型

`slot_num`：对机箱设备，用于表示对应机箱的对应线卡

【缺省配置】 Default 模式

【命令模式】 全局配置模式

【使用指导】

 在单机模式下可以配置线卡工作模式如下：

default：默认模式，适用于大多数用户场景

bridge：桥模式，适用于纯二层业务为主的场景

gateway：网关模式，适用于三层业务为主的场景

gateway-max：大网关模式，适用于大量终端的场景

gateway-ndmax：ND 网关模式，适用于大量 v6 终端的场景

label：标签模式，适用于大量 MPLS 的场景

route-v4max：v4 网络路由模式，适用于大量 IPv4 路由的场景

📖 VSU 模式下切换 UFT 模式

必须配置。

可以通过 `switch-mode mode_type switch switch_num slot slot_num` 命令对交换机 UFT 模式进行切换。

【命令格式】 `switch-mode mode_type switch switch_num slot slot_num`

【参数说明】 `mode_type`：UFT 模式类型中的一种

`switch_num`：在单机情况下 **switch** 关键字不可见；在 VSU 情况下，**switch** 用于表示对应机箱或者盒式设备。

`slot_num`：对机箱设备，用于表示对应机箱的对应线卡

【缺省配置】 Default 模式

【命令模式】 全局配置模式

【使用指导】

 在 VSU 模式下可以配置线卡工作模式如下：

default：默认模式，适用于大多数用户场景

- bridge** : 桥模式, 适用于纯二层业务为主的场景
- gateway** : 网关模式, 适用于三层业务为主的场景
- gateway-max** : 大网关模式, 适用于大量终端的场景
- gateway-ndmax** : ND 网关模式, 适用于大量 v6 终端的场景
- label** : 标签模式, 适用于大量 MPLS 的场景
- route-v4max** : v4 网络路由模式, 适用于大量 IPv4 路由的场景

5. 检验方法

- 整机重启后, 通过调用 **show run** 命令查看当前线卡状态, 查看配置是否生效。
- 使用 **show switch-mode status** 命令查看 UFT 模式状态。

【命令格式】 **show switch-mode status**

【参数说明】 无

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 -

【命令展示】

```
1. Ruijie(config)#show switch-mode status
2. Slot No          Switch-Mode
3. switch 1 slot 3  bridge
```

6. 配置举例

在单机环境下切换交换机 UFT 工作模式

【网络环境】 -

【配置方法】 将交换机槽位 3 的线卡的 UFT 工作模式切换为 Bridge 模式

```
4. Ruijie(config)#switch-mode bridge slot 3
5. Please save current config and restart your device!
6. Ruijie(config)#show run
7. Building configuration...
8. Current configuration : 1366 bytes
9. version 11.0(1B2)
10. !
11. cwmp
12. !
13. install 3 M8600E-24XS4QXS-DB
14. !
15. sysmac 1414.4b34.5624
16. !
17. nfpp
18. !
19. switch-mode bridge slot 3
```

【检验方法】 通过 **show switch-mode status** 可以查看到配置信息。

```
20. Ruijie(config)#show switch-mode status
21. Slot No          Switch-Mode
22. 3               bridge
```

常见错误

-

13.6 监视与维护

清除各类信息

-

查看运行情况

作用	命令
显示交换机 UFT 模式信息	show switch-mode status

查看调试信息

-

-  上述监视和维护命令兼容机箱设备和盒式设备，以及单机和 VSU。
-  针对单机情况，**device** 关键字不可见。机箱设备通过 **slot** 可指定某张线卡；

14 PKG_MGMT

14.1 概述

Package Management 是 RGOS 系统的包管理及升级模块，负责对设备内各个组件安装、升降级、查询、维护，其中升级是主要功能。通过对设备的软件进行升级，用户可以在系统上安装更加稳定的或含有更多的特性的软件版本，RGOS 系统采用模块化的构成方式，系统既可以进行整体升级和子系统的升级，也可以进行各个功能包的独立升级，还能以热补丁方式升级。

- ✔ 本文描述的组件升级涵盖了盒式设备和机架设备的组件升级，且本文只针对 11.0 以后的各项目平台，不涉及 11.0 以前项目升级到 11.0 以后项目。

协议规范

无

14.2 典型应用

典型应用	场景描述
升降级子系统组件	升降级盒式和机架设备的 boot, kernel, rootfs 等子系统组件。
升降级单个功能组件包	升降级盒式和机架设备单个功能组件包。
安装热补丁包	安装热补丁，对功能组件的某一部分进行修补。
自动同步升级	配置系统的自动同步升级策略

14.2.1 升降级子系统组件

应用场景

升级子系统组件包，完成升级后设备内原先的系统软件全部被更新，整体软件功能得到增强。机架设备子系统组件包称为“机架包”。

该升级方式的主要特点是：升级完成后设备内所有软件都将更新，所有已知软件 bug 都将得到完整解决，但升级过程较长。

功能部署

机架设备升级时，该机架包较大，设备自身的存储空间不足以存放该包，要求一定要将机架包存放在 U 盘内再完成升级。

14.2.2 一键升级子系统组件

应用场景

也是升降级子系统，一般常用于 VSU 双机系统，在特权模式下通过 upgrade auto 命令执行自动升级，升级过程中用户无须干预，且业务仍可不间断运行。对于单台设备（不论配置为 VSU 模式还是单机模式），执行该命令相当于执行 upgrade 命令后自动整机重启，无法做到业务不间断。

功能部署

如升降级子系统组件。

14.2.3 升降级单个功能组件包

应用场景

设备软件由若干功能组件组成，每个功能组件都是一个独立的功能模块。升级独立的功能组件包，在完成升级后仅该安装包对应功能模块的缺陷得到了修订、或者功能组件得到了增强，其它功能组件保持不变。

该升级场景的特点是：功能组件包通常较小，升级速度较快，升级完成后仅对应的功能模块得到改善，其它功能模块保持不变。

功能部署

升级功能组件包前，可以将其存放在 TFTP 服务器的根目录，通过网络下载安装到本地完成升级。也可将包存放在 U 盘内，插入设备完成升级。

14.2.4 安装热补丁包

应用场景

如果需要在不重启设备的条件下完成软件缺陷的修复，可安装热补丁包。该包仅适用于对特定软件版本的修复。通常只有当用户环境不能重启设备时，才会针对该软件版本发布专门的热补丁包用于缺陷修复。

热补丁升级最显著的特点是：升级完成后，设备无需重启即可修复缺陷。

功能部署

升级热补丁包前，可以将其存放在 TFTP 服务器的根目录，通过网络下载安装到本地完成升级，也可将包存放在 U 盘内，插入设备完成热补丁升级。

14.2.5 自动同步升级

应用场景

在分布式构成设备中，如机架式设备、VSU 系统，自动同步升级功能保证在软件升级及新加入的成员（新板卡或新机箱）时，按照预置的自动同步策略，对匹配成员的软件版本进行升级，以保证整个系统的协作运行。

功能部署

- 配置系统的自动同步升级策略。
- 配置系统的自动同步升级所需要的系统安装包文件路径。

14.3 功能详解

基本概念

↳ 子系统

子系统以映像的方式存储于设备，RGOS 的子系统包括：

- boot：设备上电启动首先加载 boot 运行，它负责设备的基础初始化，加载并运行系统映像。
- kernel：它是系统的 OS 核心部分，负责屏蔽系统的硬件构成、给应用程序提供抽象的运行环境。
- rootfs：它是系统中应用程序的集合。

📄 机架包

- 机架设备子系统组件升级使用机架包，该包包含了管理板和所有线卡的主安装包，使用这种包可以一次性的升级机架设备上的所有板卡。

📄 RGOS 的功能组件包

- RGOS 的功能组件包则是指实现某个功能的集合，在设备出厂时，所有已支持的功能均已包含在 rootfs 子系统中，通过升级单个功能组件包可以只更新系统内特定功能或特性。

📄 热补丁包

- 热补丁包包含了若干功能组件的热补丁，升级该包可以依次为各功能组件包打上补丁，并且无需重启设备立刻具有新的功能特性。

⚠️ 本文中的“安装包”均指包含子系统或功能模块的安装文件。

功能特性

功能特性	作用
子系统组件升降级及管理	升降级子系统，查询当前有哪些子系统组件可用。
功能组件升降级及管理	升降级功能组件或安装热补丁包，查询当前设备可用的子系统组件，并激活选定的子系统组件。
热补丁包的升降级及管理	方便用户查询设备中存在的功能组件包及其版本和安装信息。
自动同步升级	自动同步升级可以保证整个系统在组成成员发生变化时软件版本如预期

14.3.1 子系统组件升降级及管理

子系统的升降级就是将安装包内的子系统组件替换设备内的子系统组件，达到软件功能更新的目的。因为存在子系统冗余设计，所以升降级时往往并不是直接覆盖设备内当前正在使用的子系统，而是在设备内新增子系统然后再激活新增子系统。

工作原理

📄 升降级

各子系统在设备内存在的形式各有不同，因此对子系统的升降级方式也各有差别：

- boot：该子系统一般以映像形式存在于 norflash 设备内，所以该子系统的升降级就是将映像写入 norflash 设备。
- kernel：该子系统以文件形式存在于特定分区，所以该子系统的升降级就是文件的写入。
- rootfs：该子系统一般以映像形式存在于 nandflash 设备内，所以该子系统的升降级就是将映像写入 nandflash 设备。

📄 管理

查询当前有哪些子系统组件可用，之后依据实际需求，有选择性的加载子系统组件。

各子系统组件都包含冗余设计，在升降级过程中：

- boot：始终存在主、从两个 boot，升级只涉及主 boot，从 boot 始终冗余。
- kernel：至少存在一个冗余备份，若空间足够可存在多个冗余。
- rootfs：始终存在一个冗余备份。

对于 boot 组件因为较为特殊，并不将该组件纳入子系统管理的范畴。在升级 kernel 或 rootfs 子系统组件时升降级模块总是在配置文件记录当前使用的子系统组件和冗余的子系统组件以及各种版本管理信息。

相关配置

📄 升级

- 将升级文件存放在设备本地后，使用 **upgrade** 命令升级。

14.3.2 功能组件升降级及管理

工作原理

功能组件升级的原理实际上就是组件文件的替换过程，即包内的组件文件替换设备中的组件文件。

功能组件及热补丁的管理是利用数据库记录功能组件和热补丁的信息。安装组件，显示组件信息，卸载组件实际上就是数据库添加，查询，删除的结果。

另外，需要注意当系统进行过补丁升级后，就不能再进行组件升级。

相关配置

📄 升级

- 将升级文件存放在设备本地后，使用 **upgrade** 命令升级。

14.3.3 热补丁包的升降级及管理

工作原理

功能组件升级的原理实际上就是组件文件的替换过程，即包内的组件文件替换设备中的组件文件。

热补丁包升级原理类似，不同之处在于它只替换需要修订的文件，并且文件替换完成后，新文件自动生效。

另外，需要注意当系统进行过组件升级后，就不能再进行补丁升级。

📄 管理

热补丁的管理同功能组件一样包含查询、安装、卸载等，这些操作对应着数据库的插入，查询、删除等操作。

热补丁和功能组件的管理是基于同一技术原理实现的，但是热补丁的不同之处在于，热补丁包含未安装，已安装，已激活这三种状态，其中：

已安装仅仅表明设备内存在热补丁，但是该补丁功能并没有真正生效，

已激活状态的热补丁才真正有效。

相关配置

升级

- 将升级文件存放在设备本地文件系统中后，使用 **upgrade** 命令升级。

热补丁的激活

- 使用 **patch active** 命令临时激活已安装的补丁，设备重启后补丁作用失效，需重新激活；
- 或使用 **patch running** 命令永久激活已安装的补丁，设备重启后仍然生效。
- 未激活的补丁不会真正生效。

热补丁的失效

- 如果需要使已激活的补丁失效，可通过 **patch deactivate** 命令完成。

卸载热补丁

- **patch delete** 用于卸载热补丁。

14.3.4 自动同步升级

工作原理

自动同步升级的目的是为了保证一个系统内的多个子系统（指板卡、机箱）的协同工作：在进行软件升级时，自动将升级包发送给类型匹配的成员；或者，在新加入成员时，能够按照预定的升级策略将新员工的软件版本升级到指定版本上，以保证整个系统的软件版本在预期之中。

自动同步升级策略可选择其中之一：

- 不对新成员进行同步升级；
- 以版本前后判断为依据，如果新成员的软件版本低于系统安装包的，则进行升级；否则，不进行升级；
- 以系统升级包为基准，无论新成员的当前版本如何，都升级到基准版本上；

自动同步检查的时机包括：

- 各类升级过程中，未指定特定目标方时，由升级功能将升级包发送给所有匹配升级包的成员（包括板卡、机箱）
- 设备启动时，对每一个成员进行版本检查，按照同步策略及版本检查结果，如有必要，则进行升级；
- 新成员加入时，对其进行版本检查，按照同步策略及检查结果，如果必要，则进行升级。

管理

自动同步升级时，需要预先配置自动同步升级策略，并且预先安装好升级包。

相关配置

配置自动同步升级策略

- 为保证新成员加入时，能够进行预期的升级，需提前检查自动同步升级的相关配置，如安装包准备情况。

- 如果由于系统未打开自动同步升级策略，而错过了对某块板卡的检查，此时仍然可以手工升级的方式，对特定成员进行升级。

14.4 配置详解

配置项	配置建议 & 相关命令	
安装包升降级	 必须配置，用于子系统组件包，功能组件包及热补丁包的安装，升降级。	
	<code>upgrade [slot { num M1 M2 all }] url [force]</code>	安装，升级本地文件系统内安装包。
	<code>upgrade download tftp://path</code>	自动完成安装文件拷贝操作并升级。
	<code>upgrade rollback</code>	子系统组件回滚
	<code>patch active</code>	临时激活已安装的补丁。
	<code>patch running</code>	永久激活已安装的补丁。
热补丁的失效与卸载	 可选配置，使已激活的热补丁失效或者卸载热补丁。	
	<code>patch deactivate</code>	使补丁作用失效。
	<code>patch delete</code>	清除设备上已存在的热补丁包。
自动同步升级	 可选配置，用于配置分布式构成系统的自动同步升级策略。	
	<code>upgrade auto-sync policy</code>	设置系统的自动同步策略。
	<code>upgrade auto-sync range</code>	设置系统的自动同步升级的范围。
	<code>upgrade auto-sync package</code>	设备系统的自动同步升级的系统安装包路径。

14.4.1 安装包升降级

配置效果

可用安装包包括机架包，各功能组件包，热补丁包。

- 升级机架包，完成升级后机架设备内原先的系统软件全部被更新，整体软件功能得到增强。
 - 升级独立的功能组件包，在完成升级后仅该安装包对应功能模块的 bug 得到了修订，功能组件得到了增强，其它功能组件保持不变。
 - 升级热补丁包，即在不重启设备的条件下完成软件缺陷修复，该包仅适用于对特定软件版本的升级。
-  通常发布机架包来升级机架式设备。

注意事项

-

配置方法

升级机架包

- 可选配置。设备内原先的系统软件全部需要被更新时，选择此配置项。
- 由于机架包较大，使用机架包升级子系统组件时要把机架包存放在设备自带的 U 盘内，使用 **upgrade** 命令升级。

✔ 通常发布机架包来升级机架式设备

一键升级设备对应的安装包

- 可选配置，组成 vsu 的双机设备原先的系统软件全部需要被更新时，且要求业务仍可不间断运行时，选择此配置项。
- 升级前需要将安装包下载到设备本地，使用 **upgrade auto** 命令升级。

i 通常发机架包包来升级组成 vsu 的设备。

升级各功能组件包

- 可选配置。如果仅需要对某功能模块的缺陷进行修复，或增强该功能模块的性能的话，选择此配置项。
- 升级前需要将安装包下载到设备本地，使用 **upgrade** 命令升级。

升级热补丁包

- 可选配置。如果需要在不重启设备的条件下完成软件 bug 修订，选择此配置项。
- 升级前需要将安装包下载到设备本地。使用 **upgrade** 命令升级。
- 升级后需要激活补丁方能使用，该步骤是必选配置。激活方式有两种：使用 **patch active** 命令临时激活已安装的补丁；或使用 **patch running** 命令永久激活已安装的补丁。

⚠ 用户场景下通常要求使用 **patch running** 命令永久激活补丁。仅当用户打算临时验证补丁功能时可使用 **patch active** 命令激活补丁。

子系统回滚

- 可选配置。如果需要将子系统回滚到升级前的状态，选择此配置，作用是使系统回归到升级前的状态。
- 使用 **upgrade** 命令升级子系统组件后（如机架包）该配置才能生效。

⚠ 用户场景下使用 **upgrade** 命令升级子系统组件成功后，回滚命令仅能生效一次，不可连续回滚。

检验方法

- 完成升级子系统组件后可执行 **show upgrade history** 命令查看是否升级成功。
- 完成升级功能组件后可执行 **show component** 命令查看是否升级成功。
- 完成升级热补丁包后可执行 **show patch** 命令查看是否升级成功。

相关命令

升级机架包

【命令格式】 **upgrade [slot { num | M1 | M2 | all }] url [force]**

【参数说明】 **slot** 表明对指定槽号的设备执行该命令；

num 指定的线卡设备槽号；

M1，**M2** 指定管理板；

all 指定所有设备。

【命令模式】 特权模式

【使用指导】 该命令支持各子系统组件安装包，机架设备安装包，功能组件安装包，热补丁包。使用该命令前，先使用 `copy` 命令将功能包拷入设备文件系统中。

当使用该命令、不指定范围参数时，升级功能将按照自动同步配置，对匹配到的系统部件进行升级。

【命令格式】 **upgrade download tftp:/path**

【参数说明】 *path*：为 tftp 服务器上安装包的路径，该命令自动从服务器上下载安装包，并自动升级。

【命令模式】 特权模式

【使用指导】 该命令支持各子系统组件安装包，机架设备安装包，功能组件安装包，热补丁包。使用该命令自动完成安装文件拷贝操作并升级。

查看设备内存放的安装包文件信息

【命令格式】 **show upgrade file url**

【参数说明】 *url* 设备文件系统中安装包存放路径

【命令模式】 特权模式

【使用指导】 -

查看设备升级过程

【命令格式】 **show upgrade status**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 仅机架设备存在该命令，用于显示各个线卡的升级状态

显示当前系统升级历史信息

【命令格式】 **show upgrade history**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 -

子系统组件回滚

【命令格式】 **upgrade rollback [slot { *num* | **M1** | **M2** | **all** }]**

【参数说明】 **slot** 表明对指定槽号的设备执行该命令，*num* 指定的线卡设备槽号；**M1**，**M2** 指定管理板；**all** 指定所有设备

【命令模式】 特权模式

【使用指导】 该命令撤销上一次子系统升级操作，使子系统返回升级前的状态。只有上次升级的是子系统且子系统升级成功才能进行回滚操作，回滚命令无法连续执行。

显示已安装的功能组件

【命令格式】 **show component [slot { *num* | **M1** | **M2** | **all** }] [*component_name*]**

【参数说明】 [*component_name*] 组件名称。

当不存在此参数值时：命令用于显示设备中所有已安装的组件及各组件的基本信息。

当存在此参数值时：命令用于显示对应组件的详细信息，并校验组件内容是否完整，检测该组件能否正常工作。

slot 表明对指定槽号的设备执行该命令，*num* 指定的线卡设备槽号；**M1**，**M2** 指定管理板；**all** 指定所有设备

【命令模式】 特权模式

【使用指导】 -

▾ 显示安装的补丁包信息

【命令格式】 **show patch** [slot { num | M1 | M2 | all }] [package_name]

【参数说明】 slot 表明对指定槽号的设备执行该命令，num 指定的线卡设备槽号；M1，M2 指定管理板；all 指定所有设备

【命令模式】 特权模式

【使用指导】 -

▾ 临时激活已安装的补丁

【命令格式】 **patch active** [slot { num | M1 | M2 | all }]

【参数说明】 slot 表明对指定槽号的设备执行该命令，num 指定的线卡设备槽号；M1，M2 指定管理板；all 指定所有设备

【命令模式】 特权模式

【使用指导】 只能对已安装补丁的设备执行该操作。该命令只能临时激活补丁，设备重启后补丁作用失效。

▾ 永久激活已安装的补丁

【命令格式】 **patch running** [slot { num | M1 | M2 | all }]

【参数说明】 slot 表明对指定槽号的设备执行该命令，num 指定的线卡设备槽号；M1，M2 指定管理板；all 指定所有设备

【命令模式】 特权模式

【使用指导】 只能对已安装补丁的设备执行该操作。该命令可永久激活补丁。

配置举例

▾ VSU 双机系统，使用一键升级命令举例

【网络环境】 升级前必须将安装包拷入设备内，升级模块提供以下几种解决方案。

- 用户首先使用 **copy tftp**，**copy xmodem** 等文件系统命令将服务器上的安装包拷入设备文件系统，再使用 **upgrade auto url** 升级本地文件系统内的安装包；
- 将安装包拷入 U 盘内并插入设备，使用 **upgrade auto url** 命令升级位于 U 盘内的安装包。

【配置方法】 ● 执行 **upgrade auto** 命令

- 升级后主从设备会依次重启

```
2015-04-09_09-56-23 Ruijie#upgrade auto usb0:S6220_RGOS11.0(5)BI_install.bin
2015-04-09_09-56-24 Ruijie#*Jan 1 00:23:40: %7:
2015-04-09_09-56-24 *Jan 1 00:23:40: %7: [Slot 1/0]:Upgrade processing is 10%
2015-04-09_09-56-26 Ruijie#show upgrade status
2015-04-09_09-56-26 [Slot 1/0]
2015-04-09_09-56-26 dev_type: s6k
2015-04-09_09-56-26 status : upgrading
2015-04-09_09-56-26 [Slot 2/0]
2015-04-09_09-56-26 dev_type: s6k
2015-04-09_09-56-26 status : transmission
2015-04-09_09-58-20 *Jan 1 00:25:36: %7: [Slot 2/0]:Upgrade processing is 10%
2015-04-09_09-58-30 Ruijie#show upgrade status
2015-04-09_09-58-30 [Slot 1/0]
2015-04-09_09-58-30 dev_type: s6k
```

```
2015-04-09_09-58-30      status  : upgrading
2015-04-09_09-58-30 [Slot 2/0]
2015-04-09_09-58-30      dev_type: s6k
2015-04-09_09-58-30      status  : upgrading
2015-04-09_09-58-39 *Jan 1 00:25:56: %7:
2015-04-09_09-58-39 *Jan 1 00:25:56: %7: [Slot 2/0]:Upgrade processing is 60%
2015-04-09_09-59-19 *Jan 1 00:26:35: %7:
2015-04-09_09-59-19 *Jan 1 00:26:35: %7: [Slot 2/0]:Upgrade processing is 90%
2015-04-09_09-59-19 *Jan 1 00:26:35: %7:
2015-04-09_09-59-19 *Jan 1 00:26:35: %7: [Slot 2/0]:
2015-04-09_09-59-19 *Jan 1 00:26:35: %7: Upgrade info [OK]
2015-04-09_09-59-19 *Jan 1 00:26:36: %7:   Kernel
version[2.6.32.6b311610a8eb91->2.6.32.6b31161115502c]
2015-04-09_09-59-19 *Jan 1 00:26:36: %7:   Rootfs version[1.0.0.eb75cd01->1.0.0.3d978b6c]
2015-04-09_09-59-19 *Jan 1 00:26:36: %7:
2015-04-09_09-59-19 *Jan 1 00:26:36: %7: [Slot 2/0]:Reload system to take effect!
2015-04-09_09-59-21 *Jan 1 00:26:37: %7:
2015-04-09_09-59-21 *Jan 1 00:26:37: %7: [Slot 2/0]:Upgrade processing is 100%
2015-04-09_10-00-28 Ruijie#show upgrade status
2015-04-09_10-00-28 [Slot 1/0]
2015-04-09_10-00-28      dev_type: s6k
2015-04-09_10-00-28      status  : upgrading
2015-04-09_10-00-28 [Slot 2/0]
2015-04-09_10-00-28      dev_type: s6k
2015-04-09_10-00-28      status  : success
2015-04-09_10-01-39 *Jan 1 00:28:56: %7:
2015-04-09_10-01-39 *Jan 1 00:28:56: %7: [Slot 1/0]:Upgrade processing is 60%
2015-04-09_10-02-17 *Jan 1 00:29:33: %7:
2015-04-09_10-02-17 *Jan 1 00:29:33: %7: [Slot 1/0]:Upgrade processing is 90%
2015-04-09_10-02-17 *Jan 1 00:29:33: %7:
2015-04-09_10-02-17 *Jan 1 00:29:33: %7: [Slot 1/0]:
2015-04-09_10-02-17 *Jan 1 00:29:34: %7: Upgrade info [OK]
2015-04-09_10-02-17 *Jan 1 00:29:34: %7:   Kernel
version[2.6.32.6b311610a8eb91->2.6.32.6b31161115502c]
2015-04-09_10-02-17 *Jan 1 00:29:34: %7:   Rootfs version[1.0.0.eb75cd01->1.0.0.3d978b6c]
2015-04-09_10-02-17 *Jan 1 00:29:34: %7:
2015-04-09_10-02-18 *Jan 1 00:29:34: %7: [Slot 1/0]:Reload system to take effect!
2015-04-09_10-02-19 *Jan 1 00:29:35: %7:
2015-04-09_10-02-19 *Jan 1 00:29:35: %7: [Slot 1/0]:Upgrade processing is 100%
2015-04-09_10-02-19 *Jan 1 00:29:36: %7: %PKG_MGMT:auto-sync config synchronization, Please wait
for a moment....
```

```

2015-04-09_10-02-20 *Jan 1 00:29:36: %7:
2015-04-09_10-02-20 [ 1784.116069] rtc-pcf8563 6-0051: retrieved date/time is not valid.
2015-04-09_10-02-20 *Jan 1 00:29:36: %7: [Slot 2/0]:auto sync config: space not enough left
57229312, need 114597815
2015-04-09_10-02-20 *Jan 1 00:29:36: %7:
2015-04-09_10-02-20 *Jan 1 00:29:36: %7: [Slot 2/0]:auto sync package config err
2015-04-09_10-02-20 *Jan 1 00:29:37: %7: [Slot 1/0]
2015-04-09_10-02-21 *Jan 1 00:29:37: %7: device_name: s6k
2015-04-09_10-02-21 *Jan 1 00:29:37: %7: status: SUCCESS
2015-04-09_10-02-21 *Jan 1 00:29:37: %7: [Slot 2/0]
2015-04-09_10-02-21 *Jan 1 00:29:37: %7: device_name: s6k
2015-04-09_10-02-21 *Jan 1 00:29:37: %7: status: SUCCESS
2015-04-09_10-02-21 *Jan 1 00:29:38: %7: %Do with dtm callback...
2015-04-09_10-02-21 *Jan 1 00:29:38: %VSU-5-DTM_AUTO_UPGRADE: Upgrading the system, wait a moment
please.

```

【检验方法】

- 查看当前设备运行的版本信息，若版本信息发生变换说明升级成功

```

Ruijie-VSU#show version detail
System description      : Ruijie Full 10G Routing Switch(S6220-48XS4QXS-II) By Ruijie Networks
System start time      : 1970-02-07 00:33:24
System uptime          : 0:00:47:28
System hardware version : 2.00
System software version : S6220_RGOS 11.0(5)B1
System patch number    : NA
System software number  : M01181304092015
System serial number    : 1234942570019
System boot version     : 1.2.9.f23dcbe(150317)
System core version     : 2.6.32.6b31161115502c
Module information:
  Slot 2/0 : RG-S6220-48XS4QXS-II
    Hardware version    : 2.00
    System start time   : 1970-02-07 00:33:24
    Boot version        : 1.2.9.f23dcbe(150317)
    Software version    : S6220_RGOS 11.0(5)B1
    Software number     : M01181304092015
    Serial number       : 1234942570019
  Slot 1/0 : RG-S6220-48XS4QXS-II
    Hardware version    : 2.00
    System start time   : 1970-02-07 00:33:47
    Boot version        : 1.2.9.f23dcbe(150317)
    Software version    : S6220_RGOS 11.0(5)B1

```

```
Software number   : M01181304092015
Serial number     : 1234942570018
```

📌 机架设备子系统安装包升级举例

【网络环境】 机架设备上通常自带 U 盘，升级机架包时需要先将机架包放入 U 盘内，升级模块提供以下几种解决方案。

- 用户首先使用 **copy tftp**，**copy xmodem** 等文件系统命令将服务器上的安装包拷入设备文件系统，再使用 **upgrade url** 升级本地文件系统内的安装包。
- 将安装包拷入 U 盘内并插入设备，使用 **upgrade url** 命令升级位于 U 盘内的安装包。
-

【配置方法】

- 执行升级命令
- 升级过程中可查看板卡状态
- 完成子系统升级后设备重启生效

```
Ruijie# upgrade usb0:/ca-octeon_11.0(1B2)_20131106_main_install.bin
[Slot M1]:Upgrade processing is 10%

[Slot 1]:Upgrade processing is 10%

[Slot M1]:Upgrade processing is 60%

[Slot 1]:Upgrade processing is 60%

[Slot M1]:Upgrade processing is 90%

[Slot M1]:
Upgrade info [OK]
  Kernel version[2. 6. 32. abb2b41f170c81->2. 6. 32. abb2b415749f40]
  Rootfs version[1. 0. 0. d5f0de03->1. 0. 0. 660e0085]

[Slot M1]:Reload system to take effect!

[Slot M1]:Upgrade processing is 100%
[Slot 1]:Upgrade processing is 90%

[Slot 1]:
Upgrade info [OK]
  Kernel version[2. 6. 32. 9f8b56f1d45ab2 ->2. 6. 32. 0f48cb9f170c81]
  Rootfs version[1. 0. 0. 2ad02537->1. 0. 0. 1bcc12e8]

[Slot 1]:Restart to take effect !

[Slot 1]:Upgrade processing is 100%
[slot: M1]
```

```

device_name: ca-octeon-cm
Status:      SUCCESS
[slot: 1]
device_name: ca-octeon-lc
status:      SUCCESS

```

```

N18000#show upgrade status
[slot: M1]
    dev_type: ca-octeon-cm
    status   : upgrading
[slot: 1]
    dev_type: ca-octeon-lc
    status   : transmission

```

```

Reload system?(Y/N)y
Restarting system.

```

- 【检验方法】
- 查看当前设备运行的版本信息，若版本信息发生变换说明升级成功

```

Ruijie#show upgrade history
Last Upgrade Information:
    Time:          2014-08-31 12:15:03
    Method:        LOCAL
Package Name: N18000_RGOS11.0(1)B1_CM_01200616_install.bin
Package Type: Distribution

```

📌 机架设备功能包升级举例

- 【网络环境】 机架设备上通常自带 U 盘，升级机架包时需要先将机架包放入 U 盘内，升级模块提供以下几种解决方案。

- 用户首先使用 **copy tftp**，**copy xmodem** 等文件系统命令将服务器上的安装包拷入设备文件系统，再使用 **upgrade url** 升级本地文件系统内的安装包。
- 将安装包拷入 U 盘内并插入设备，使用 **upgrade url** 命令升级位于 U 盘内的安装包。

- 【配置方法】
- 执行升级命令
 - 依照升级后的提示确定是否需要设备重启

```

Ruijie#upgrade sata0://ca-octeon-cm_bridge-1.0.0.05151504-1311060616.mips.rpm
[Slot M1]:Upgrade processing is 10%

[Slot M1]:Upgrade processing is 60%

[Slot M1]:
Upgrade info [OK]
    bridge version[1.0.0.97521231->1.0.0.05151504]

[Slot M1]:Restart to take effect !

```

```
[Slot M1]:Upgrade processing is 100%
[slot: M1]
    device_name: ca-octeon-cm
    status:      SUCCESS
[slot: M2]
    device_name: ca-octeon-lc
    status:      NOT SUPPORT
Reload system?(Y/N)y
[ 1586.114348] Restarting system.
```

- 【检验方法】
- 查看当前设备功能组件版本信息，若版本信息发生变换说明升级成功

```
Ruijie# show component slot M1
[Slot M1]:
Package : bridge
    Version: 1.0.0.05151504      Build time: Wed May 15 07:05:06 2013
    Size: 11      Install time: Thu Jan  1 00:48:09 1970
    Description: bridge component
    Required packages: None
-----
.....
```

📌 机架设备补丁包安装举例

- 【网络环境】 机架设备上通常自带 U 盘，升级机架包时需要先将机架包放入 U 盘内，升级模块提供以下几种解决方案。
- 用户首先使用 **copy tftp**，**copy xmodem** 等文件系统命令将服务器上的安装包拷入设备文件系统，再使用 **upgrade url** 升级本地文件系统内的安装包。
 - 将安装包拷入 U 盘内并插入设备，使用 **upgrade url** 命令升级位于 U 盘内的安装包。

- 【配置方法】
- 执行升级命令
 - 激活热补丁

```
Ruijie#upgrade usb0:/ca-octeon-cm_RGOS11.0(1C2)_20131008_patch.bin
[Slot M1]: Upgrade processing is 10%
[Slot M1]: Upgrade processing is 60%

[Slot M1]:
Upgrade info [OK]
    patch_bridge version[1.0.0.1952]
Upgrade processing is 90%

[Slot M1]:
Upgrade info [OK]
    patch_install version[1.0.0.192e35a]
[slot: M1]
```

```

    device_name: ca-octeon-cm
    status:      SUCCESS
[slot: M2]
    device_name: ca-octeon-lc
    status:      NOT SUPPORT
Ruijie#patch running slot M1
The patch on the system now is in running status

```

【检验方法】 ● 查看当前设备安装热补丁信息

```

Ruijie# show patch slot M1
[Slot M1]:
Patch package patch_install installed in the system, version:pal
Package : patch_bridge
Status : running
Version: pal      Build time: Mon May 13 09:03:07 2013
      size: 277    Install time: Tue May 21 03:07:17 2013
Description: a patch for bridge
Required packages: None

```

▾ 机架设备子系统升级回滚配置举例

 只有当上次升级的是子系统，子系统升级成功才能进行子系统回滚操作，否则回滚命令无法继续执行。

 机架设备可指定对哪块板卡设备进行回滚操作。

【网络环境】 机架设备上通常自带 U 盘，升级机架包时需要先将机架包放入 U 盘内，升级模块提供以下几种解决方案。

- 用户首先使用 **copy tftp**，**copy xmodem** 等文件系统命令将服务器上的安装包拷入设备文件系统，再使用 **upgrade url** 升级本地文件系统内的安装包。
- 将安装包拷入 U 盘内并插入设备，使用 **upgrade url** 命令升级位于 U 盘内的安装包。
-

【配置方法】 ● 执行子系统回滚命令
● 设备重启后回滚生效

```

Ruijie#upgrade rollback slot M1
kernel rollback version[2.6.32.abb2b415749f40->2.6.32.abb2b41f170c81][OK]
rootfs rollback version[1.0.0.660e0085->1.0.0.d5f0de03][OK]
Rollback success!
Reload system to take effect!
Reload system?(Y/N)y
Restarting system.

```

【检验方法】 ● 查看当前设备运行的版本信息，若版本回复到升级前状态则子系统回滚成功。

```
Ruijie#show upgrade history
```

```
Last Upgrade Information:
  Time:          2014-08-31 12:15:03
  Method:       LOCAL
Package Name: N18000_RGOS11.0(1)B1_CM_01200616_install.bin
Package Type: Distribution
```

常见错误

若升级过程中出现错误，升级模块会加以提示例如：

```
Upgrade info [ERR]
  Reason:creat config file err(217)
```

- 常见错误提示有以下几种：
- 安装包无效：可能的原因是该安装包已经被损坏或者根本不是一个安装包。该错误的处理方式要求用户重新获取安装包，再执行升级操作。
- 设备不支持安装包：可能的原因是误用了其它设备的安装包。该错误的处理方式要求用户重新获取并核对安装包后在执行升级操作。
- 设备空间不足：通常出现在机架设备中。该错误的处理方式是要求用户检查设备是否存在 U 盘，按要求这些设备往往附带 U 盘。

14.4.2 热补丁的失效与卸载

配置效果

使已激活的热补丁失效或者卸载热补丁。

注意事项

- 未激活的热补丁不生效，所以未激活的热补丁不能使其失效。

配置方法

▾ 使已激活的补丁失效

- 可选配置。如果需要使已激活的补丁失效，可通过 **patch deactivate** 命令完成。

▾ 卸载热补丁

- 可选配置。如果需要卸载已安装的热补丁，可使用 **patch delete** 命令完成。

检验方法

- 可使用 **show patch** 命令检测补丁是否被激活或已经被卸载。

相关命令

使已激活的补丁作用失效

【命令格式】 **patch deactivate** [slot { num | M1 | M2 | all }]

【参数说明】 **slot** : 表明对指定槽号的设备执行该命令 ;

num : 指定的线卡设备槽号 ;

M1, M2 : 指定管理板 ;

all : 指定所有设备。

【命令模式】 特权模式

【使用指导】 只有对处于激活状态的补丁才能执行该操作，该操作使对应补丁失效。

卸载热补丁

【命令格式】 **patch delete** [slot { num | M1 | M2 | all }]

【参数说明】 **slot num** : 机架设备中使用，根据槽号指定对应的线卡。

slot all : 机架设备中使用，指定所有设备。

slot M1 : 机架设备中使用，指定对管理板 M1 操作。

slot M2 : 机架设备中使用，指定对管理板 M2 操作。

【命令模式】 特权模式

【使用指导】 用于清除设备上已存在的热补丁包。

配置举例

机架设备使补丁作用失效并卸载补丁

- 【配置方法】
- 执行补丁失效命令
 - 执行补丁卸载命令

```
Ruijie#patch deactivate slot M1
[Slot M1]:
Deactivate the patch package success
Ruijie# patch delete slot M1
[Slot M1]:
Clear the patch patch_bridge success
Clear the patch success
```

- 【检验方法】
- 查看补丁状态信息

```
Ruijie#show patch slot M1
[Slot M1]:
No patch package installed in the system
```

常见配置错误

- 补丁未处于激活状态时就执行 **patch deactive** 命令。解决方法确认补丁所处的状态，只有当提示 status : running 时，才能执行 **patch deactive** 命令。

14.4.3 自动同步升级

配置效果

配置分布式构成系统的自动同步升级策略。

注意事项

- -

配置方法

▾ 配置自动同步升级策略

可使用 **upgrade auto-sync policy** 来配置自动同步升级的策略，有 3 种可供选择：

- none：不进行自动同步升级；
- compatible：以版本前后为依据来判断是否需要自动同步
- coordinate：以管理板存储的系统升级包为依据，将所有其它成员的软件同步至此版本

▾ 配置自动同步升级的范围

可使用 **upgrade auto-sync range** 来配置自动同步升级的策略，有 2 种范围可供选择：

- chassis：自动同步范围在 1 台机箱内；
- vsd：自动同步范围在 VSU 系统内；
- 配置自动同步升级的范围

▾ 配置自动同步升级的升级包路径

在系统中有新成员（板卡或机箱）加入时，升级功能将使用该路径来确定升级包的位置。通常，在每次进行系统升级时，升级功能会自动记录本次升级所使用升级包路径，用于自动同步升级。用户也可以使用 **upgrade auto-sync package** 手工进行设定。

检验方法

- 可使用 **upgrade auto-sync** 命令上述配置是否是预期的。

相关命令

配置自动同步升级的策略

【命令格式】 **upgrade auto-sync policy [none | compatible | coordinate]**

【参数说明】 **none**: 不进行自动同步升级；
compatible: 以版本前后为依据来判断是否需要自动同步；
coordinate: 以管理板存储的系统升级包为依据，同步至此版本。

【命令模式】 特权模式

【使用指导】 建议配置为 **coordinate**

✔ 各参数仅适用于机架式设备及 VSU 系统。

配置自动同步升级的范围

【命令格式】 **upgrade auto-sync range [chassis | vsu]**

【参数说明】 **chassis**: 在机箱范围内进行版本自动同步。
vsu : vsu 系统范围内进行版本自动同步升级。

【命令模式】 特权模式

【使用指导】 建议设置为 vsu，以在最大范围内保证版本一致。

✔ 各参数仅适用于机架式设备及 vsu 系统。

配置自动同步升级的系统升级包路径

【命令格式】 **upgrade auto-sync package url**

【参数说明】 **url**: 升级包所在路径。

【命令模式】 特权模式

【使用指导】 通常不需要设置。

✔ 各参数仅适用于机架式设备及 vsu 系统。

配置举例

配置自动同步升级策略为以管理板为基准

【配置方法】 ● 执行配置自动同步升级策略命令

```
Ruijie# upgrade auto-sync policy coordinate
Upgrade auto-sync policy is set as coordinate
```

【检验方法】 ● 查看当前自动同步升级策略

```
Ruijie#show upgrade auto-sync
auto-sync policy: coordinate
auto-sync range: vsu
auto-sync package: flash:/eg1000m_main_1.0.0.0f328e91.bin
```

配置自动同步升级范围为 vsu 系统

【配置方法】 ● 执行配置自动同步升级策略命令

```
Ruijie# upgrade auto-sync range vsu
Upgrade auto-sync range is set as vsu.
```

【检验方法】 ● 查看当前自动同步升级的升级包路径

```
Ruijie#show upgrade auto-sync
auto-sync policy: coordinate
auto-sync range: vsu
auto-sync package: flash:/eg1000m_main_1.0.0.0f328e91.bin
```

常见配置错误

- 自动同步升级的 `url` 未指向有效的升级包。

14.5 监视与维护

清除各类信息

作用	命令
删除已安装的热补丁包。	patch delete [slot { num M1 M2 all }]
清除本地安装包。	clear storage [url]

查看运行情况

作用	命令
显示当前设备已安装所有组件及各组件信息。	show component [slot { num M1 M2 all }] [component_name]
显示设备中已安装的热补丁包的相关信息。	show patch [slot { num M1 M2 all }] [patch_name]
显示机架式设备各板卡的升级状态。	show upgrade status
显示升级历史信息	show upgrade history



配置指南-以太网交换

本分册介绍以太网交换配置指南相关内容，包括以下章节：

1. 接口
2. MAC 地址
3. Aggregate Port
4. VLAN
5. SUPER-VLAN
6. Protocol VLAN 配置
7. PRIVATE VLAN
8. MSTP
9. GVRP
10. LLDP
11. QINQ
12. MGMT/管理用以太网接口命令

1 接口

1.1 概述

接口是网络设备上能够实现数据交换功能的重要部件。我司网络设备上支持两种类型的接口：物理接口和逻辑接口。物理接口意味着该接口在设备上有对应的、实际存在的硬件接口，如：百兆以太网接口、千兆以太网接口等。逻辑接口意味着该接口在路由器上没有对应的、实际存在的硬件接口，逻辑接口可以与物理接口关联，也可以独立于物理接口存在，如：Loopback 接口和 Tunnel 接口等等。实际上对于网络协议而言，无论是物理接口还是逻辑接口，都是一样对待的。

i 下文仅介绍接口的相关内容。

协议规范

- 无。

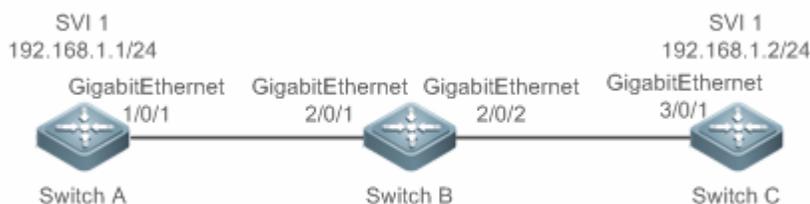
1.2 典型应用

典型应用	场景描述
以太网物理接口实现二层数据交换	通过二层以太网物理接口实现网络设备的二层数据通信。
以太网物理接口实现三层数据路由	通过三层以太网物理接口实现网络设备的三层数据通信。

1.2.1 以太网物理接口二层数据交换

应用场景

图 1-1



上图中，三台交换机设备 Switch A、Switch B 和 Switch C 组成了一个简单的二层数据交换网络。

【注释】 -

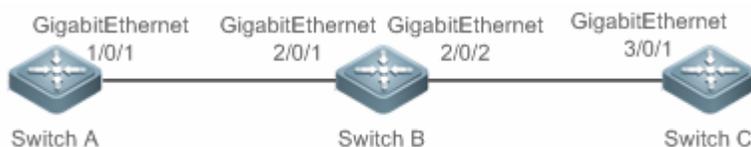
功能部属

- Switch A 和 Switch B 分别通过千兆以太网物理接口 GigabitEthernet 1/0/1 和 GigabitEthernet 2/0/1 进行相连。
- Switch B 和 Switch C 分别通过千兆以太网物理接口 GigabitEthernet 2/0/2 和 GigabitEthernet 3/0/1 进行相连。
- 将接口 GigabitEthernet 1/0/1、GigabitEthernet 2/0/1、GigabitEthernet 2/0/2 和 GigabitEthernet 3/0/1 配置为 Trunk 口。
- 分别在 Switch A 和 Switch C 上创建一个交换虚拟接口(Switch Virtual Interface, SVI) SVI 1，并给 SVI 1 接口配置相同网段的 IP 地址，其中，Switch A 的 SVI 1 接口的 IP 地址配置为 192.168.1.1/24，Switch C 的 SVI 1 接口的 IP 地址配置为 192.168.1.2/24。
- 在 Switch A 和 Switch C 上分别执行 ping 192.168.1.2 和 ping 192.168.1.1 操作，可以实现设备 B 上的二层数据交换功能。

1.2.2 以太网物理接口三层路由通信

应用场景

图 1-2



上图中，三台交换机设备 Switch A、Switch B 和 Switch C 组成了一个简单的三层数据通信网络。

【注释】 -

功能部属

- Switch A 和 Switch B 分别通过千兆以太网物理接口 GigabitEthernet 1/0/1 和 GigabitEthernet 2/0/1 进行相连。
- Switch B 和 Switch C 分别通过千兆以太网物理接口 GigabitEthernet 2/0/2 和 GigabitEthernet 3/0/1 进行相连。
- 将接口 GigabitEthernet 1/0/1、GigabitEthernet 2/0/1、GigabitEthernet 2/0/2 和 GigabitEthernet 3/0/1 配置为三层路由口。
- 分别给 GigabitEthernet 1/0/1 和 GigabitEthernet 2/0/1 配置相同网段的 IP 地址，其中，GigabitEthernet 1/0/1 的 IP 地址配置为 192.168.1.1/24，GigabitEthernet 2/0/1 的 IP 地址配置为 192.168.1.2/24。
- 分别给 GigabitEthernet 2/0/2 和 GigabitEthernet 3/0/1 配置相同网段的 IP 地址，其中，GigabitEthernet 2/0/2 的 IP 地址配置为 192.168.2.1/24，GigabitEthernet 3/0/1 的 IP 地址配置为 192.168.2.2/24。
- 在 Switch C 上配置一条静态路由表项使其能够三层直通 192.168.1.0/24 网段。
- 在 Switch A 和 Switch C 上分别执行 ping 192.168.2.2 和 ping 192.168.1.1 操作，可以实现设备 B 上的三层路由通信功能。

1.3 功能详解

基本概念

▾ 接口类型分类

锐捷设备的接口类型可分为以下两大类：

- 二层接口(L2 interface)
- 三层接口(L3 interface) (三层设备支持)
- FC 接口(某些数据中心产品支持)

1. 常见的二层接口可分为以下几种类型：

- 交换端口 (Switch Port)
- 二层聚合端口 (L2 Aggregate Port)

2. 常见的三层接口可分为以下几种类型

- 路由端口 (Routed Port)
- 三层聚合端口 (L3 Aggregate Port)
- SVI 口
- Loopback 接口
- Tunnel 接口

3. FC 接口类型

- FC 接口
- FC 聚合口

▾ 交换端口

交换端口由设备上的单个物理端口构成，只有二层交换功能。交换端口被用于管理物理接口和与之相关的第二层协议。

▾ 二层聚合端口

聚合端口是由多个物理成员端口聚合而成的。我们可以把多个物理链接捆绑在一起形成一个简单的逻辑链接，这个逻辑链接我们称之为一个聚合端口（以下简称聚合端口）。

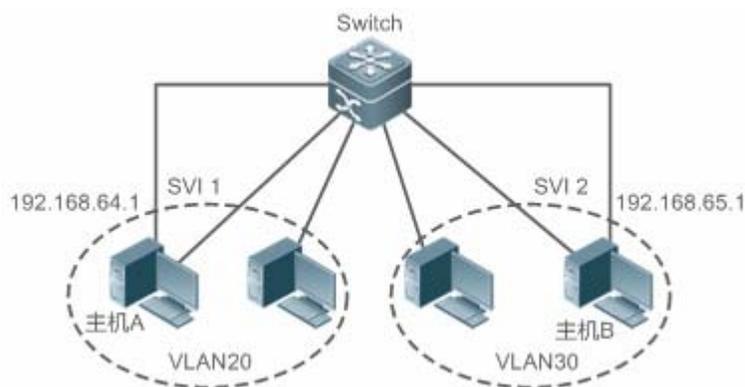
对于二层交换来说聚合端口就好像一个高带宽的交换端口，它可以把多个端口的带宽叠加起来使用，扩展了链路带宽。此外，通过二层聚合端口发送的帧还将在二层聚合端口的成员端口上进行流量平衡，如果聚合端口中的一条成员链路失效，二层聚合端口会自动将这个链路上的流量转移到其他有效的成员链路上，提高了连接的可靠性。

▾ SVI 口

SVI 接口可以做为本机的管理接口，通过该管理接口管理员可管理设备。用户也可以创建 SVI 接口为一个网关接口，就相当于是对应各个 VLAN 的虚拟接口，可用于三层设备中跨 VLAN 之间的路由。创建一个交换虚拟接口很简单，用户可通过 **interface vlan** 接口配置命令来创建 SVI 接口，然后给交换虚拟接口分配 IP 地址来建立 VLAN 之间的路由。

如图所示，VLAN 20 的主机可直接互相通讯，无需通过三层设备的路由，若 VLAN 20 内的主机 A 想和 VLAN 30 内的主机 B 通讯必须通过 VLAN 20 对应的 SVI 1 和 VLAN30 对应的 SVI 2 才能实现。

图 1-3



路由端口

在三层设备上,可以把单个物理端口设置为路由端口,作为三层交换的网关接口。一个路由端口与一个特定的 VLAN 没有关系,而是作为一个访问端口。路由端口不具备二层交换的功能。用户可通过 **no switchport** 命令将一个交换端口转变为路由端口,然后给路由端口分配 IP 地址来建立路由。注意的是,当使用 **no switchport** 接口配置命令时,将删除该端口的所有二层特性。

- i** 当一个端口是二层聚合端口的成员端口或者是未认证成功的 DOT1X 认证口时,是不能用 **switchport** 或者 **no switchport** 命令进行层次切换的。

三层聚合端口

三层聚合端口同二层聚合端口一样,也是由多个物理成员端口汇聚构成的一个逻辑上的聚合端口组。汇聚的端口必须为同类型的三层接口。对于三层交换来说,聚合端口作为三层交换的网关接口,它相当于把同一聚合组内的多条物理链路视为一条逻辑链路,是链路带宽扩展的一个重要途径。此外,通过三层聚合端口发送的帧同样能在三层聚合端口的成员端口上进行流量平衡,当聚合端口中的一条成员链路失效后,三层聚合端口会自动将这个链路上的流量转移到其它有效的成员链路上,提高了连接的可靠性。

三层聚合端口不具备二层交换的功能。用户可通过 **no switchport** 命令将一个无成员的二层聚合端口转变为三层聚合端口,接着将多个路由端口加入此三层聚合端口,然后给三层聚合端口分配 IP 地址来建立路由。

Loopback 口

Loopback 接口是完全软件模拟的本地三层逻辑接口,它永远都处于 UP 状态。发往 Loopback 接口的数据包将会在设备本地处理,包括路由信息。Loopback 接口的 IP 地址可以用来作为 OSPF 路由协议的设备标识、实施发向 Telnet 或者作为远程 Telnet 访问的网络接口等等。配置一个 Loopback 接口类似于配置一个以太网接口,可以把它看作一个虚拟的以太网接口。

Tunnel 口

Tunnel 接口来实现隧道功能,允许利用传输协议(如 IP)来传送任意协议的网络数据包。同其它逻辑接口一样,Tunnel 接口也是系统虚拟的接口。Tunnel 接口并不特别指定传输协议或者负载协议,它提供的是一个用来实现标准的点对点的传输模式。由于 Tunnel 实现的是点对点的传输链路,所以对于每一个单独的链路都必须设置一个 Tunnel 接口。

FC 口

FC 接口(Fibre Channel interface , FC 接口)是物理光纤通道接口,用于支撑 FC 存储网络通讯。通过配置 FC 接口的不同工作模式 (E、F、NP) , 允许和原有的 FC SAN 网络或者新建 FC SAN 网络建立丰富的连接, 从而实现融合网络的组网。

FC 聚合口

FC 聚合口也称 FC AP , 同二层聚合口以及三层聚合口类似。FC 聚合口就是把多个工作于 E 模式的 FC 物理端口捆绑在一起形成一个虚拟逻辑端口。理论上, FC 聚合口带宽是其所有成员接口的各带宽总和。因此, 使用 FC 聚合功能用于满足更高的带宽需求。

功能特性

功能特性	作用
接口配置命令的使用	进入接口配置模式, 在接口配置模式下用户可配置接口的相关属性。对于逻辑口, 用户进入接口模式时, 如果该接口不存在, 将会首先创建出该接口。
接口的描述和管理状态	用户可以为一个接口起一个专门的名字来标识这个接口, 有助于用户记住一个接口的功能; 用户可以设置接口的管理状态。
接口的MTU	用户可以通过设置端口的 MTU 来控制该端口允许收发的最大帧长。
配置接口带宽	用户可以基于接口配置接口的带宽。
配置接口的Load-interval	用户可以指定每隔多少时间计算报文输入输出的负载情况。
配置接口载波时延	用户可以调整接口的载波时延来调整接口状态从 Down 状态到 Up 状态或者从 Up 状态到 Down 状态的时间延时。
接口的LinkTrap策略	在设备中可以基于接口配置是否发送该接口的 LinkTrap 信息。
接口索引永久化功能	接口索引永久化功能, 即设备重启后接口索引不变。
配置路由口	在三层设备上, 用户可以把物理端口设置为路由端口, 作为三层交换的网关接口。
配置三层AP口	在三层设备上, 可以把 AP 端口设置为三层 AP 端口, 作为三层交换的网关接口。
选择接口介质类型	光电复用端口, 用户可以根据需要选择使用光口还是电口。
接口的速率, 双工、流控和自协商因子模式	用户可以调整接口的速率, 双工模式、流控模式和自协商因子模式。
模块自动检测	在配置接口速率为自动协商模式的情况下, 能够根据插入的模块类型自动调节接口的速率。
保护口	用户可以通过将某些端口设置为保护口来实现端口之间不能互相通信。同时还可以通过配置操作来设置保护口之间不能进行路由。
端口违例恢复	当端口因发生违例而被关闭之后, 用户可以在全局模式下使用端口违例恢复命令来将所有违例接口从错误状态中恢复过来, 重新复位使能该接口。
端口震荡保护	用户配置端口震荡保护功能, 当端口发生震荡时, 系统自动 shutdown 端口, 用于保护端口

1.3.1 接口配置命令的使用

用户可在全局配置模式下使用 **interface** 命令进入接口配置模式。在接口配置模式下用户可配置接口的相关属性。

工作原理

在全局配置模式下输入 `interface` 命令，进入接口配置模式。对于逻辑口，用户进入接口模式时，如果该接口不存在，将会首先创建出该接口。用户也可以在全局配置模式下使用 `interface range` 或 `interface range macro` 命令配置一定范围的接口（接口的编号）。但是定义在一个范围内的接口必须是相同类型和具有相同特性的。

对于逻辑口，可以在全局配置模式下通过执行 `no interface` 命令删除指定的逻辑接口。

接口编号规则

对于物理端口，在单机模式下编号由两部分组成：插槽号和端口在插槽上的编号，例如端口所在的插槽编号为 2，端口在插槽上的编号为 3，则端口对应的接口编号为 2/3；在 VSU 模式或者堆叠模式下编号由三部分组成：设备号，插槽号和端口在插槽上的编号，例如设备号为 1，端口所在的插槽编号为 2，端口在插槽上的编号为 3，则端口对应的接口编号为 1/2/3。

设备号是从 1 到支持的成员设备的最大数量。

插槽的编号规则：静态插槽的编号固定为 0，动态插槽（可插拔模块或线卡）的编号是从 1 - 插槽的个数。动态插槽的编号规则是：面对设备的面板，插槽按照从前至后，从左至右，从上至下的顺序一次排列，对应的插槽号从 1 开始依次增加。

插槽上的端口编号是从 1 - 插槽上的端口数，编号顺序是从左到右。

对于可以选择介质类型的设备，端口包括两种介质：光口和电口，称为光电复用端口，无论使用那种介质，都使用相同的端口编号。

对于聚合端口，其编号的范围为 1 - 设备支持的聚合端口个数。

对于交换虚拟接口，其编号就是这个交换虚拟接口对应的 VLAN 的 VID。

配置一定范围的接口

用户可以使用全局配置模式下的 `interface range` 命令同时配置多个接口。当进入 `interface range` 配置模式时，此时设置的属性适用于所选范围内的所有接口。

输入一定范围的接口。

`interface range` 命令可以指定若干范围段。

`macro` 参数可以使用范围段的宏定义，参见配置和使用端口范围的宏定义。

每个范围段可以使用逗号（,）隔开。

同一条命令中的所有范围段中的接口必须属于相同类型。

当使用 `interface range` 命令时，请注意 `range` 参数的格式：

常见的有效的接口范围格式：

- **FastEthernet** device/slot/{第一个 port} - {最后一个 port}；
- **GigabitEthernet** device/slot/{第一个 port} - {最后一个 port}；
- **TenGigabitEthernet** device/slot/{第一个 port} - {最后一个 port}；
- **FortyGigabitEthernet** device/slot/{第一个 port} - {最后一个 port}；
- **AggregatePort** Aggregate-port 号- Aggregate-port 号，范围是 1 ~ 设备支持的最大聚合端口数量；
- **vlan** vlan-ID-vlan-ID, VLAN ID 范围 1 ~ 4094；
- **Loopback** loopback-ID-loopback-ID, 范围是 1 ~ 2147483647；
- **Tunnel** tunnel-ID-tunnel-ID, 范围是 0 ~ 设备支持的最大 Tunnel 端口数量减一。

在一个 **interface range** 中的接口必须是相同类型的，即或者全是 FastEthernet、GigabitEthernet、AggregatePort，或者全是 SVI 接口等。

配置和使用端口范围的宏定义

用户可以自行定义一些宏来取代端口范围的输入。但在用户使用 **interface range** 命令中的 **macro** 关键字之前，必须先在全局配置模式下使用 **define interface-range** 命令定义这些宏。

在全局配置模式下使用 **no define interface-range macro_name** 命令来删除设置的宏定义。

1.3.2 接口的描述和管理状态

用户可以为一个接口起一个专门的名字来标识这个接口，有助于用户记住一个接口的功能。

用户可以进入接口模式对接口进行关闭和打开管理。

工作原理

接口的描述

用户可以根据要表达的含义来设置接口的具体名称，比如，用户想将 GigabitEthernet 1/1 分配给用户 A 专门使用，用户就可以将这个接口的描述设置为“Port for User A”。

接口的管理状态

在某些情况下，用户可能需要禁用某个接口。用户可以通过设置接口的管理状态来直接关闭一个接口。如果关闭一个接口，则这个接口上将不会接收和发送任何帧，这个接口将丧失这个接口对应的所有功能。用户也可以通过设置管理状态来重新打开一个已经关闭的接口。接口的管理状态有两种：Up 和 Down，当端口被关闭时，端口的管理状态为 Down，否则为 Up。

1.3.3 接口的MTU

用户可以通过设置端口的 MTU 来控制该端口允许收发的最大帧长。

工作原理

当端口进行大吞吐量数据交换时，可能会遇到大于以太网标准帧长度的帧，这种帧被称为 jumbo 帧。MTU 是指帧中有效数据段的长度，不包括以太网封装的开销。

端口收到或者转发的帧，如果长度超过设置的 MTU 将被丢弃。

MTU 允许设置的范围为 64~9216 字节，粒度为 4 字节，缺省一般为 1500 字节。

 此配置命令只对物理端口和 AP 口有效。

1.3.4 配置接口带宽

工作原理

主要用于一些路由协议(如 OSPF 路由协议)计算路由量度和 RSVP 计算保留带宽。修改接口带宽不会影响物理接口的数据传输速率。

i 接口的带宽命令不能实际影响某个接口的带宽，它只是个路由参数，不会影响物理链路的接口的真正带宽。

1.3.5 配置接口的Load-interval

工作原理

接口的 load-interval 可以指定每隔多少时间计算报文输入输出的负载情况，一般是每隔 10 秒钟计算一次每秒中输入输出的报文数和比特数。

1.3.6 配置接口载波时延

工作原理

接口的载波时延 Carry-delay 是指接口链路的载波检测信号 DCD 从 Down 状态到 Up 状态或者从 Up 状态到 Down 状态的延时，如果 DCD 在延时之内发生变化，那么系统将忽略这种状态的变化而不至于上层的数据链路层重新协商。如果参数设置的比较大，那么几乎每次瞬间的 DCD 变化将无法被检测到；相反，如果参数设置成 0，那么每次微小的 DCD 信号的跳变都将被系统检测到，这样系统也就将增加不稳定性。

i 如果 DCD 载波中断时间比较长，那么将该参数设长些，可以尽快加速拓扑收敛和路由汇聚，以便网络拓扑或者路由表可以较快的收敛。如果相反，DCD 载波中断时间小于网络拓扑或者路由汇聚所花的时间，那么应该将该参数设置相对的大些，以免造成没有必要的网络拓扑振荡或者路由振荡。

1.3.7 接口的LinkTrap策略

在设备中，用户可以基于接口配置选择是否发送该接口的 LinkTrap 状态变化信息。

工作原理

当接口的 LinkTrap 发送功能打开时，如果该接口的 Link 状态变化，SNMP 将发出 LinkTrap 信息，反之则不发。

1.3.8 接口索引永久化功能

和接口的名字一样，接口索引也可以用于标识一个接口，接口索引是一个接口的“身份 ID”，每个接口创建时，系统会自动为每个接口分配不重复的接口索引值，而当设备重启后，一个接口的索引值可能会和重启前不一致。接口索引永久化功能，即设备重启后接口索引不变。

工作原理

当配置了该功能，设备重启后相同接口的接口索引值保持不变。

1.3.9 配置路由口

工作原理

在三层设备上，可以把物理端口设置为路由端口，作为三层交换的网关接口。路由端口不具备二层交换的功能。用户可通过 **no switchport** 命令将一个交换端口转变为路由端口，然后给路由端口分配 IP 地址来建立路由。注意的是，当使用 **no switchport** 接口配置命令时，将删除该端口的所有二层特性。

1.3.10 配置三层AP口

工作原理

在三层设备上，类似三层路由口一样，用户可通过 **no switchport** 命令将一个二层 AP 端口转变为三层 AP 端口，然后给该 AP 端口分配 IP 地址来建立路由。注意的是，当使用 **no switchport** 接口配置命令时，将删除该 AP 端口的所有二层特性。

- ❗ 当 AP 口中含有成员口时，不允许将二层 AP 口配置为三层 AP 口，反之，也不允许将带有成员口的三层 AP 口转变为二层 AP 口。

1.3.11 接口的速率、双工、流控和自协商因子模式

对于以太网物理接口和 AP 口，用户可以配置管理接口的速率、双工、流控和自协商因子模式。

工作原理

▾ 接口的速率

通常情况下，以太网物理接口速率是通过和对端设备自协商决定的。协商得到的速率可以是接口速率能力范围内的任意一个速率。用户也可以通过配置接口能力范围内的任意一个具体速率值让以太网物理接口工作在该指定速率值上。

对于 AP 口，当用户设置 AP 口的速率时，实际上是生效到该 AP 口的所有成员口上(这些成员口都是以太网物理接口)的。

▾ 接口的双工

- 以太网物理接口和 AP 口的双工模式时存在三种情况：
- 可以将接口设置为全双工属性实现接口在发送数据包的同时可以接收数据包；
- 可以将接口设置为半双工属性控制接口同一时刻只能发送数据包或接收数据包时；
- 当设置接口的双工属性为自协商模式时，接口的双工状态由本接口和对端接口自动协商而定。
- 对于 AP 口，当用户设置 AP 口的双工模式时，实际上是生效到该 AP 口的所有成员口上(这些成员口都是以太网物理接口)的。

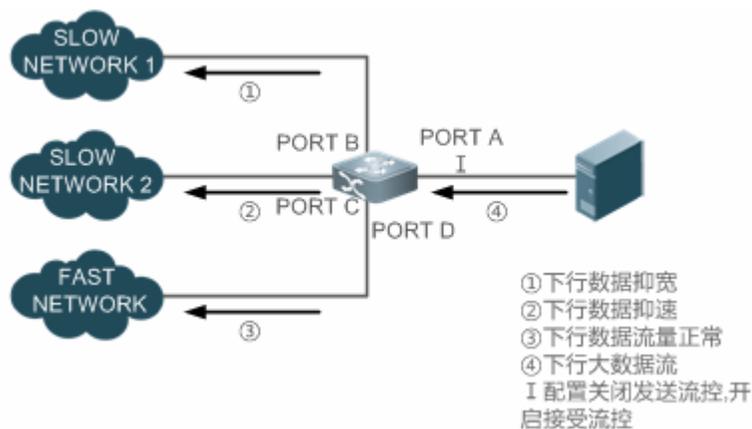
▾ 接口的流控

接口的流控模式分为非对称流控模式 and 对称流控模式：

- 对称流控模式，即在一般情况下，接口开启流控模式后，接口上将会处理接收到的流控帧，并在接口出现拥塞时发送流控帧，接收和发送流控帧的处理是一致的，这就是对称流控模式。
- 非对称流控模式，即在一些情况下，设备希望某个接口能够处理接收到的流控帧保证报文不会因为拥塞而丢弃，又不想发出流控帧而导致整个网络速率下降，这个时候，就要通过配置非对称流控，将接收流控帧和发送流控帧的处理步调分开。
- 对于 AP 口，当用户设置 AP 口的流控模式时，实际上是生效到该 AP 口的所有成员口上(这些成员口都是以太网物理接口)的。

如**错误！未找到引用源。**所示，设备的端口 A 为上联口，端口 B-D 为下联端口，其中端口 B 和 C 对应的是一个慢速网络，假如端口 A 上开启了接收流控和发送流控功能，由于端口 B 和 C 对应的是一个慢速网络，在发送端口 B 的数据流过大，导致端口 B 和 C 拥塞，进而导致端口 A 上的入口拥塞，端口 A 上就会发送流控帧，当上联设备响应流控帧时，就会降低往端口 A 的数据流，间接导致端口 D 上的网速下降。这个时候，可以配置端口 A 的发送流控功能关闭，来保障整个网络带宽利用率。

图 1-4



▾ 接口的自协商因子模式

- 接口的自协商因子模式有 on 和 off 两种。接口的自协商状态和接口的自协商因子模式并不完全等同，接口的自协商状态通常由接口的速率、双工、流控和自协商因子模式共同决定。

- 对于 AP 口，当用户设置 AP 口的自协商因子模式时，实际上是生效到该 AP 口的所有成员口上(这些成员口都是以以太网物理接口)的。

- ❗ 一般情况下，只要接口的速率、双工和流控中的一种属性为 auto 模式，或者接口的自协商模式为 on 模式，那么接口的自协商工作状态就是 on 的，即接口的自协商功能是打开的；反之，当接口的速率、双工和流控中的属性全部为非 auto 模式，并且接口的自协商模式为 off 模式时，那么接口的自协商工作状态就是 off 的，即接口的自协商功能是关闭的。
- ❗ 对于百兆光口，接口的自协商功能永远都是关闭的，即百兆光口的自协商工作状态永远都是 off 的；对于千兆电口，接口的自协商功能永远都是开启的，即千兆电口的自协商工作状态永远都是 on 的。

1.3.12 模块自动检测

在配置接口速率为自动协商模式的情况下，能够根据插入的模块类型自动调节接口的速率。

工作原理

目前支持的模块有 SFP 和 SFP+两种模块，其中 SFP 为千兆模块，SFP+为万兆模块，若插入的是 SFP 模块，则接口工作在千兆模式，若插入的是 SFP+模块，则接口工作在万兆模式。

- ❗ 模块的自动检测功能只在速率配置为自动协商时才能生效。

1.3.13 保护口

有些应用环境下，要求交换机上的部分端口间不能互相通讯，可以通过将某些端口设置为保护口(Protected Port)来达到目的。同时还可以通过配置操作来设置保护口之间不能进行路由。

工作原理

▾ 保护口

当端口设为保护口之后，保护口之间互相无法通讯，保护口与非保护口之间可以正常通讯。

保护口有两种模式，一种是阻断保护口之间的二层交换，但允许保护口之间进行路由，第二种是同时阻断保护口之间的二层交换和阻断路由；在两种模式都支持的情况下，第一种模式将作为缺省配置模式。

当两个保护口设为一个镜像会话端口对时，该镜像会话的源端口发送或接收的帧依然能够镜像到该镜像会话的目的端口上。

目前只支持在以太网物理接口和 AP 口上设置保护口。当一个 AP 口被设置为保护口时，该 AP 所有成员口都被设置为保护口。

▾ 保护口之间三层路由阻断

缺省情况下，保护口之间的三层路由并没有被阻断，这个时候可以通过设置保护口之间不能进行路由的功能来实现保护口之间的路由阻断功能。

1.3.14 端口违例恢复

某些协议具备设置端口违例（关闭端口）的功能，用以保证网络的安全性和稳定性。比如端口安全协议，当用户配置开启端口安全，并配置端口上最大安全地址数量，当端口下学习到的地址数超过最大安全地址数时，将产生端口违例事件。另外生成树协议、DOT1X 协议、REUP 协议等也都具备类似的功能，违例的端口会自动关闭该接口，以保证安全性。

工作原理

当端口因发生违例而被关闭之后，可以在全局模式下使用端口违例恢复命令来将所有违例接口从错误状态中恢复过来，重新复位使能该接口。可以选择手动恢复，也可以选择定时自动恢复。

1.3.15 端口并行检测配置

开启并行检测以后，在千兆光口开启自协商时，无论对端的千兆光口的协商能力是强制还是自协商，都可以正常使用

 仅工作在 1000M 的电口和光电复用口支持并行检测功能。

1.3.16 端口震荡保护

当发生接口震荡时，会产生大量硬件中断，从而消耗大量 CPU 资源，另一方面频繁的接口震荡容易损害接口，用户可以配置接口震荡保护功能来保护接口。

工作原理

接口链路震荡保护功能由用户自行决定是否开启，默认情况下为开启保护功能。当接口发生震荡时，接口每 2s 或 10 秒都会检测一次震荡，如果检测到接口 2s 内震荡 6 次，则打印提示信息，连续打印 10 次提示信息（也就是 20s 内连续检测到接口震荡），则关闭接口；对于 10s 检测一次震荡，如果检测到接口 10s 发生 10 次震荡，则打印提示信息，不关闭接口。

1.3.17 接口 Syslog

用户可以通过配置打开或关闭 Syslog 开关来决定是否查看接口状态发生改变或异常的信息。

工作原理

接口 Syslog 开关由用户自行决定是否开启，默认情况下为开启。当接口发生异常情况，比如接口状态发生改变，接口收到错误帧或接口发生震荡时，系统将打印提示信息告之用户。

1.4 产品说明



S7508 所有系列线卡的内联口仅支持配置 speed auto 及 speed 10G。



- 对于万兆光口，接口的自协商功能永远都是关闭的，即万兆光口的自协商工作状态永远都是 off 的；
- 对于 40G 光口，插入光模块的时，自协商状态是关闭的，插入铜缆时，自协商状态是开启的；



- S7508 的所有线卡的 mtu 配置即对入口生效，也对出口生效
- S7508 所有系列线卡 mtu 的生效值比配置值多 26 字节(包含 14 字节以太网头部、4 个字节的 FCS、2 个 TAG)；



配置此功能时需要注意保证三层接口的 IP MTU 与链路 MTU 的合理性，IP MTU 不大于接口 MTU，三层接口包括路由口、三层 AP 口和 SVI。

1.5 配置详解

配置项	配置建议 & 相关命令	
接口配置管理	可选配置。主要用于进行接口的创建、删除、接口描述管理等管理配置。	
	interface	创建一个接口，并进入指定接口配置模式，或者直接进入该接口的接口配置模式。
	interface range	输入一定范围的接口，当这些接口未被创建时，同时进行接口创建，并进入接口批量配置模式。
	define interface-range	将批量操作的接口定义成宏定义形式。
	snmp-server if-index persist	开启接口索引永久化功能，即设备重启后接口索引不变。
	description	在接口配置模式下，使用该命令设置接口的描述，最多 80 字符。
	snmp trap link-status	基于接口配置是否发送该接口的 LinkTrap 信息。
	shutdown	在接口配置模式下，使用该命令关闭接口。
	physical-port dither protect	在全局配置模式下，配置接口震荡保护功能
	logging [link-updown error-frame link-dither]	在全局配置模式下，配置接口开启打印接口状态信息
配置接口属性	可选配置。主要用于进行接口的属性等管理配置。	
	bandwidth	在接口配置模式下，使用该命令设置接口的带宽参数。
	carrier-delay	在接口配置模式下，使用该命令设置接口载波时延。
	load-interval	在接口配置模式下，使用该命令设置接口的负载计算的间隔时间
	duplex	设置接口的双工模式。

flowcontrol	打开或关闭接口的流量控制。
mtu	设置接口的 MTU。
negotiation mode	设置接口的自协商因子模式。
speed	设置接口的速率。
switchport	在接口配置模式下使用不带任何参数的 <code>switchport</code> 命令，将一个接口设置为二层接口模式，使用 <code>no switchport</code> 命令将一个接口设置为三层接口模式。
switchport protected	设置接口为保护口模式。
protected-ports route-deny	在全局配置模式下，使用该命令来设置保护口之间的三层路由阻断功能。
errdisable recovery	在全局配置模式下，使用改命令来恢复违例端口
parallel detect disable	在接口配置模式下，关闭端口并行检测状态

1.5.1 接口配置管理

配置效果

- 能够创建出指定的单个逻辑口，并进入接口的配置模式，或者对于已经存在的物理接口或者逻辑接口，可以进入接口的配置模式。
- 能够批量创建出指定的逻辑口，并进入接口批量操作的配置模式，或者对于已经存在的物理接口或者逻辑接口，可以进入接口批量操作的配置模式。
- 能够实现相同接口在设备重启前后接口索引保持不变。
- 设置接口的描述符，对该接口直观、形象化的理解。
- 能够启用或者关闭接口的 LinkTrap 功能。
- 配置接口管理状态，关闭或者打开接口。

注意事项

- 对于逻辑接口，可以使用该命令的 `no` 命令形式删除接口或者将指定范围接口的批量删除，但不可以使用该命令的 `no` 命令形式删除指定的物理接口或批量删除指定范围的物理接口。
- 可以使用该命令的 `default` 命令形式将指定物理接口或者逻辑接口或者指定范围的接口在接口模式下的相关配置恢复到缺省配置。

配置方法

📌 配置单个指定的接口

- 可选配置。
- 可以用于需要创建某个不存在的逻辑接口或者进入已经存在的物理接口和逻辑接口的接口配置模式以进行接口相关的配置时，需要配置该命令。

【命令格式】 **interface** *interface-type* *interface-number*

【参数说明】 *interface-type interface-number* :即接口的类型和接口编号,可以是以太网物理接口、AP 口、SVI 口、Loopback 口等。

【缺省配置】 无

【命令模式】 全局配置模式

- 【使用指导】
- 对于逻辑接口，如果该接口未被创建，则首先创建出该接口并进入接口的配置模式。
 - 对于物理接口或者已经创建的逻辑接口，直接进入该接口的配置模式。
 - 使用 **no** 命令形式删除指定的逻辑接口。
 - 使用 **default** 命令形式将该接口的接口模式下配置恢复到缺省配置。

▾ 配置一定范围的接口

- 可选配置。
- 可以用于需要批量创建不存在的逻辑接口或者进入已经存在的物理接口和逻辑接口的接口批量配置模式以进行接口相关的配置时，需要配置该命令。

【命令格式】 **interface range** { *port-range* | **macro** *macro_name* }

【参数说明】 *port-range* : 即批量操作的接口类型和接口编号范围,可以是以太网物理接口、AP 口、SVI 口、Loopback 口等。

macro_name : 即一定范围接口类型的宏定义名。

【缺省配置】 无

【命令模式】 全局配置模式

- 【使用指导】
- 对于逻辑接口，如果接口未被创建，则首先创建出接口然后再进入接口的批量配置模式。
 - 对于物理接口或者已经创建的逻辑接口，直接进入接口的批量配置模式。
 - 使用 **default** 命令形式批量将接口模式下配置恢复到缺省配置。
 - 使用宏定义的时候，需要在全局配置模式下，先将一定范围的接口类型通过 **define interface-range** 命令进行宏定义成 *macro_name*，然后再通过 **interface range macro macro_name** 进行接口的批量配置管理。

▾ 配置接口的索引永久化

- 可选配置。
- 可以用于需要保持接口索引在系统重启前后一致时使用。

【命令格式】 **snmp-server if-index persist**

【参数说明】 -

【缺省配置】 该功能关闭。

【命令模式】 全局配置模式

- 【使用指导】 执行该命令后，保存配置时将会把当前所有接口的索引保存起来，重启后接口使用重启前分配的接口索引。可以使用该命令的 **no** 命令或者 **default** 命令形式关闭该功能。

配置接口的描述符

- 可选配置。
- 可以用于为接口设置描述信息时使用。

【命令格式】 **description** *string*

【参数说明】 *string*：最长 80 个字符

【缺省配置】 缺省无接口描述符

【命令模式】 接口配置模式

【使用指导】 该命令配置接口的描述符。可以使用该命令的 **no** 命令或者 **default** 命令形式取消配置接口的描述符。 -

配置接口的 LinkTrap

- 可选配置。
- 可以用于通过 SNMP 获取接口状态变化的 Trap 信息。

【命令格式】 **snmp trap link-status**

【参数说明】 -

【缺省配置】 缺省情况下，该功能打开

【命令模式】 接口配置模式

【使用指导】 该命令配置是否发送该接口的 LinkTrap，当功能打开时，如果接口的 Link 状态变化，SNMP 将发出 LinkTrap，反之则不发。可以使用该命令的 **no** 命令或者 **default** 命令形式关闭该功能。

配置接口的管理状态

- 可选配置。
- 用于关闭或者打开接口。
- 接口关闭后将无法收发报文。

【命令格式】 **shutdown**

【参数说明】 -

【缺省配置】 接口的管理状态是 UP

【命令模式】 接口配置模式

【使用指导】 对接口执行 **shutdown** 操作时，即关闭该接口，执行 **no shutdown** 操作将重新打开该接口。注意有些情况下，不允许将端口执行 **no shutdown** 操作，比如该端口处于端口违例状态，那么该端口将无法执行 **no shutdown** 操作。可以使用该命令的 **no** 命令或者 **default** 命令形式重新打开该接口。

配置接口震荡保护功能

- 可选配置。
- 用于保护发生震荡的接口。

【命令格式】 **physical-port dither protect**

【参数说明】

【缺省配置】 默认开启接口震荡保护功能

【命令模式】 全局配置模式

【使用指导】

📌 配置打印接口 Syslog 信息功能

- 可选配置。
- 用于开启或关闭打印接口状态信息功能。

【命令格式】 **[no] logging [link-updown | error-frame | link-dither]**

【参数说明】

【缺省配置】 开启打印接口 Syslog 信息

【命令模式】 全局配置模式

【使用指导】

检验方法

📌 配置单个指定的接口

- 执行 **interface** 操作，如果能够正常进入接口模式，即说明配置是成功的。
- 对于逻辑接口，如果是执行 **no interface** 操作，也可以通过 **show running** 命令或者 **show interfaces** 命令查看对应的接口是否存在，如果不存在，则该逻辑接口是被正常删除的。
- 执行 **default interface** 操作，通过 **show running** 命令查看对应的接口下的配置是否都恢复到了缺省配置，如果是，则说明该操作是成功的。

📌 配置一定范围的接口

- 执行 **interface range** 操作，如果能够正常进入接口批量配置模式，即说明配置是成功的。
- 执行 **default interface range** 操作，通过 **show running** 命令查看对应的接口下的配置是否都恢复到了缺省配置，如果是，则说明该操作是成功的。

📌 配置接口索引永久化

- 配置完该命令后，执行 **write** 保存配置操作，重启设备后，通过 **show interface** 命令查看接口的接口索引值，如果对于同一个接口的索引值在设备重启后保持一致，那么说明接口的索引永久化功能是正常的。

📌 配置接口的 LinkTrap

- 选择一个物理端口，进行网线插拔，同时打开 SNMP 服务器，如果在网线插拔的时候，SNMP 服务器能够正常收到该接口的 Link 状态变化的 Trap 信息，则说明该功能是正常的。
- 执行 **no** 命令形式操作，如果验证到在一个物理端口，进行网线插拔，同时打开 SNMP 服务器，如果在网线插拔的时候，SNMP 服务器无法收到该接口的 Link 状态变化的 Trap 信息，则说明已经正常关闭了接口的 LinkTrap 功能。

📌 配置接口的管理状态

- 选择一个物理端口，插上网线，让端口 Up 起来，对该端口执行 **shutdown** 关闭接口的操作，用户在控制台上能够看到端口状态变成管理 Down 的 Syslog 信息，同时该端口上的指示灯灭掉，则关闭端口的功能是正常的，

并且通过 **show interfaces** 命令可以看到接口状态显示为 **administratively down**。在此基础上，执行 **no shutdown** 重新打开该接口，用户在控制台上能够看到端口 Up 的 Syslog 信息，同时该端口上的指示灯重新亮起来，则打开端口的功能是正常的。

配置接口震荡保护功能

- 在全局配置模式下配置命令，如 **physical-port dither protect**。选择一个物理端口，频繁插拔网线模拟端口发生震荡的情况，在控制台上可以看到系统打印端口发生震荡的信息，经过连续打印若干次之后，系统会提示端口不稳定将 **shutdown** 接口。

配置打印接口 Syslog 信息功能

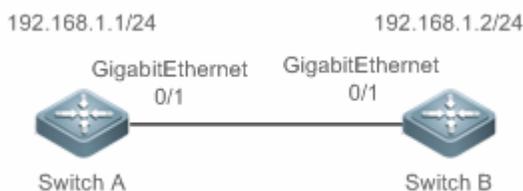
- 在全局配置模式下配置命令，如 **logging link-updown** 查看接口状态信息。选择一个物理端口，插拔网线，接口将发生两次状态改变，用户可以在控制台上看到接口 link 状态从 up 变为 down 的信息，又从 down 变为 up 的信息；配置 **no logging link-updown** 后，再次插拔网线，控制台上看不到信息，说明该功能是正常的。

配置举例

配置接口基本属性

【网络环境】

图 1-5



【配置方法】

- 将 2 台设备通过交换端口进行连接。
- 分别给 2 台设备配置一个 SVI 口，并配置相同网段的 IP 地址。
- 在 2 台设备上分别配置接口索引永久化。
- 在 2 台设备上分别启用 LinkTrap 功能。
- 在两台设备上配置接口管理状态。

```

A
A# configure terminal
A(config)# snmp-server if-index persist
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip address 192.168.1.1 255.255.255.0
A(config-if-VLAN 1)# exit
A(config)# interface gigabitethernet 0/1
A(config-if-GigabitEthernet 0/1)# snmp trap link-status
A(config-if-GigabitEthernet 0/1)# shutdown
A(config-if-GigabitEthernet 0/1)# end
A# write
  
```

```

B
B# configure terminal
B(config)# snmp-server if-index persist
  
```

```

B(config)# interface vlan 1
B(config-if-VLAN 1)# ip address 192.168.1.2 255.255.255.0
B(config-if-VLAN 1)# exit
B(config)# interface gigabitethernet 0/1
B(config-if-GigabitEthernet 0/1)# snmp trap link-status
B(config-if-GigabitEthernet 0/1)# shutdown
B(config-if-GigabitEthernet 0/1)# end
B# write

```

【检验方法】 在 A、B 设备上分别进行如下检验：

- 检查设备上的 GigabitEthernet 0/1 和 SVI 1 在接口 GigabitEthernet 0/1 在 **shutdown** 操作后的接口状态是否正确
- 检查接口 GigabitEthernet 0/1 **shutdown** 操作后，是否有发出该接口 Link Down 的 Trap 信息
- 重启设备后，接口 GigabitEthernet 0/1 的接口索引值是否和重启前的一致

A

```

A# show interfaces gigabitEthernet 0/1
Index(dec):1 (hex):1
GigabitEthernet 0/1 is administratively down , line protocol is DOWN
Hardware is GigabitEthernet, address is 00d0.f865.de9b (bia 00d0.f865.de9b)
Interface address is: no ip address
  MTU 1500 bytes, BW 1000000 Kbit
  Encapsulation protocol is Bridge, loopback not set
  Keepalive interval is 10 sec , set
  Carrier delay is 2 sec
  Rxload is 1/255, Txload is 1/255

```

Queue	Transmitted packets	Transmitted bytes	Dropped packets	Dropped bytes
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	4	440	0	0

```

Switchport attributes:
  interface's description:""
  lastchange time:0 Day:20 Hour:15 Minute:22 Second
  Priority is 0
  admin medium-type is Copper, oper medium-type is Copper   admin duplex mode is AUTO, oper duplex
is Unknown
  admin speed is AUTO, oper speed is Unknown
  flow control admin status is OFF, flow control oper status is Unknown

```

```

admin negotiation mode is OFF, oper negotiation state is ON
Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
Port-type: access
Vlan id: 1
10 seconds input rate 0 bits/sec, 0 packets/sec
10 seconds output rate 0 bits/sec, 0 packets/sec
4 packets input, 408 bytes, 0 no buffer, 0 dropped
Received 0 broadcasts, 0 runts, 0 giants
0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
4 packets output, 408 bytes, 0 underruns , 0 dropped
0 output errors, 0 collisions, 0 interface resets
A# show interfaces vlan 1
Index(dec):4097 (hex):1001
VLAN 1 is UP , line protocol is DOWN
Hardware is VLAN, address is 00d0.f822.33af (bia 00d0.f822.33af)
Interface address is: 192.168.1.1/24
ARP type: ARPA, ARP Timeout: 3600 seconds
MTU 1500 bytes, BW 1000000 Kbit
Encapsulation protocol is Ethernet-II, loopback not set
Keepalive interval is 10 sec , set
Carrier delay is 2 sec
Rxload is 0/255, Txload is 0/255

```

B

```

B# show interfaces gigabitEthernet 0/1
Index(dec):1 (hex):1
GigabitEthernet 0/1 is administratively down , line protocol is DOWN
Hardware is GigabitEthernet
Interface address is: no ip address, address is 00d0.f865.de9b (bia 00d0.f865.de9b)
MTU 1500 bytes, BW 1000000 Kbit
Encapsulation protocol is Bridge, loopback not set
Keepalive interval is 10 sec , set
Carrier delay is 2 sec
Rxload is 1/255, Txload is 1/255

```

Queue	Transmitted packets	Transmitted bytes	Dropped packets	Dropped bytes
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	4	440	0	0

```
Switchport attributes:
  interface's description:""
  lastchange time:0 Day:20 Hour:15 Minute:22 Second
  Priority is 0
  admin medium-type is Copper, oper medium-type is Copper
  admin duplex mode is AUTO, oper duplex is Unknown
  admin speed is AUTO, oper speed is Unknown
  flow control admin status is OFF, flow control oper status is Unknown
  admin negotiation mode is OFF, oper negotiation state is ON
  Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
Port-type: access
  Vlan id: 1
  10 seconds input rate 0 bits/sec, 0 packets/sec
  10 seconds output rate 0 bits/sec, 0 packets/sec
  4 packets input, 408 bytes, 0 no buffer, 0 dropped
  Received 0 broadcasts, 0 runts, 0 giants
  0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
  4 packets output, 408 bytes, 0 underruns , 0 dropped
  0 output errors, 0 collisions, 0 interface resets
B# show interfaces vlan 1
Index(dec):4097 (hex):1001
VLAN 1 is UP , line protocol is DOWN
Hardware is VLAN, address is 00d0.f822.33af (bia 00d0.f822.33af)
Interface address is: 192.168.1.2/24
ARP type: ARPA, ARP Timeout: 3600 seconds
  MTU 1500 bytes, BW 1000000 Kbit
  Encapsulation protocol is Ethernet-II, loopback not set
  Keepalive interval is 10 sec , set
  Carrier delay is 2 sec
  Rxload is 0/255, Txload is 0/255
```

常见错误

- 无。

1.5.2 配置接口属性

配置效果

- 将设备通过交换端口或者路由端口进行连接和数据通信。

- 在设备上分别调整各种接口属性。

注意事项

- 无。

配置方法

配置路由端口

- 可选配置。
- 可以用于需要将接口转换为三层路由口时使用。
- 配置成三层路由口后将使该接口上运行的二层协议失效。
- 支持二层交换口上配置。

【命令格式】 **no switchport**

【参数说明】 -

【缺省配置】 交换机上的以太网物理端口缺省为二层交换

【命令模式】 接口配置模式

【使用指导】 在三层交换机设备上，使用该命令可以将一个二层交换口配置成三层路由口。使用 **switchport** 命令可以将一个三层路由口转换成二层交换口。

配置三层 AP 口

- 可选配置。
- 可以在接口配置模式下，执行 **no switchport** 命令将一个二层 AP 口配置成三层 AP 口。使用 **switchport** 命令时，可以将一个三层 AP 口配置成二层 AP 口。
- 配置成三层路由口后将使该接口上运行的二层协议失效。
- 支持二层聚合口上配置。

【命令格式】 **no switchport**

【参数说明】 -

【缺省配置】 缺省情况下，交换机上的 AP 端口缺省为二层 AP 口

【命令模式】 接口配置模式

【使用指导】 在三层交换机设备上，进入二层 AP 口的接口模式后，使用该命令可以将一个二层 AP 口配置成三层 AP 口。进入三层 AP 口的接口模式后，使用 **switchport** 命令可以将一个三层 AP 口转换成二层 AP 口。

配置接口介质类型

- 可选配置。
- 对于光电复用口，接口的缺省介质为电口。
- 配置的端口介质类型变化时，可能会引起端口状态震荡。

- 支持以太网物理端口上及聚合口上配置。

【命令格式】 **medium-type { auto-select [prefer [fiber | copper]] | fiber | copper }**

【参数说明】 **auto-select**：自动选择介质类型

prefer [fiber | copper]：自动选择的时候优先选择某种介质类型

fiber：强制选择光口

copper：强制选择电口

【缺省配置】 缺省情况下，端口上选择的介质类型为电口

【命令模式】 接口配置模式

【使用指导】 如果同一端口可以选择光口和电口两种介质类型，用户只能使用其中之一。一旦确定介质类型之后，在配置端口的属性，例如状态、双工、流量控制和速率等，都是指端口当前选择类型的属性。改变端口类型后，新类型对应的端口的属性为其默认属性，用户可以根据需要重新配置。

如果用户配置端口介质自动选择，在端口只有一种介质连接上时，设备使用当前连接的介质；在端口的两种介质都连接上时，设备将使用用户配置的优先介质。介质自动选择优先介质默认为电口，用户可以通过配置 `medium-type auto-select prefer fiber` 来设置优先介质为光口。在自动选择模式下，端口的速率、双工、流控等属性将使用默认值。

配置接口速率

- 可选配置。
- 配置的端口速率模式变化时，可能会引起端口震荡。
- 支持以太网物理端口上及聚合口上配置。

【命令格式】 **speed [10 | 100 | 1000 | 10G | auto]**

【参数说明】 **10**：表示接口的速率为 10Mbps。

100：表示接口的速率为 100Mbps。

1000：表示接口的速率为 1000Mbps。

10G：表示接口的速率为 10Gbps。

auto：表示接口的速率为自适应的。

【缺省配置】 缺省情况下，接口的速率是自协商模式，即接口的速率配置缺省为 auto 模式

【命令模式】 接口模式

【使用指导】 如果接口是聚合端口的成员，则该接口的速率由聚合端口的速率决定。接口退出聚合端口时使用自己的速率设置。使用 `show interfaces` 命令查看设置。接口类型不同，允许设置的速率类型也会有所不同，如 SFP 类型的接口就不允许把速率设为 10M。

配置接口双工模式

- 可选配置。
- 配置的端口双工模式变化时，可能会引起端口震荡。
- 支持以太网物理端口上及聚合口上配置。

【命令格式】 **duplex { auto | full | half }**

【参数说明】 **auto**：表示全双工和半双工自适应。

full：表示全双工。

half：表示半双工。

【缺省配置】 缺省情况下，接口的双工是自协商模式，即接口的双工配置缺省为 auto 模式

【命令模式】 接口模式

【使用指导】 接口的双工属性与接口的类型有关。可以使用 **show interfaces** 命令查看接口双工的设置。

配置接口流控模式

- 可选配置。
- 一般情况下，接口的流控模式缺省为 off 模式。部分产品的缺省模式为 on 模式。
- 接口开启流控模式后，在接口上出现拥塞时，将接收或者发送流控帧调整网络数据流量。
- 配置的端口流控模式变化时，可能会引起端口震荡。
- 支持以太网物理端口上及聚合口上配置。

【命令格式】 **flowcontrol { auto | off | on }**

【参数说明】 **auto**：自协商流量控制。

off：关闭流量控制。

on：打开流量控制。

receive：非对称流量控制接收方向。

send：非对称流量控制发送方向。

【缺省配置】 缺省情况下，接口的流控一般是 off 模式，即接口的流控功能缺省是关闭的

【命令模式】 接口配置模式

【使用指导】 部分产品不支持非对称流控，不支持 **send** 和 **receive** 关键字。使用 **show interfaces** 查看接口流量控制和实际的流量控制。

配置接口自协商因子模式

- 可选配置。
- 配置的端口自协商因子变化时，可能会引起端口震荡。
- 支持以太网物理端口上及聚合口上配置。

【命令格式】 **negotiation mode { on | off }**

【参数说明】 **on**：自协商因子模式为 on 模式。

off：自协商因子模式为 off 模式。

【缺省配置】 缺省情况下，接口的自协商因子是 off 模式

【命令模式】 接口配置模式

【使用指导】 -

配置接口 MTU

- 可选配置。
- 可以通过设置端口的 MTU 来控制端口允许收发的最大帧长。
- 支持以太网物理端口及 SVI 口设置。

【命令格式】 **mtu num**

- 【参数说明】 *num* : 64 - 9216
- 【缺省配置】 缺省情况下，接口的 MTU 值一般为 1500 字节
- 【命令模式】 接口模式
- 【使用指导】 设置接口所支持的 MTU，即链路层数据部分的最大长度。目前只支持设置物理端口和包含成员口的 AP 口的 MTU。

配置接口带宽

- 可选配置。
- 一般情况下，接口的带宽值和接口支持的速率值相同。

- 【命令格式】 **bandwidth kilobits**
- 【参数说明】 *kilobits* : 以每秒 K 比特为单位，范围为 1 到我司设备目前支持的最大以太网速率能力值。
- 【缺省配置】 缺省情况下，接口带宽值一般和接口类型相匹配，比如对于千兆以太网物理端口，该接口的缺省带宽值为 1000000，万兆以太网物理端口则为 10000000
- 【命令模式】 接口配置模式
- 【使用指导】 -

-

配置接口载波时延

- 可选配置。
- 配置的载波时延时间较长时，接口物理状态变化时会较晚引起协议状态的变化，若配置为 0 秒时，接口物理状态变化则立刻引起协议状态变化。

- 【命令格式】 **carrier-delay {[milliseconds] num | up [milliseconds] num down [milliseconds] num}**
- 【参数说明】 *num* : 默认以秒为单位，范围 0~60 秒。
milliseconds : 配置以毫秒为单位的载波延迟，范围 0~60000 毫秒。
Up : 设置载波检测信号 DCD 从 Down 状态到 Up 状态的时间延时。
Down : 设置载波检测信号 DCD 从 Up 状态到 Down 状态的时间延时。
- 【缺省配置】 缺省情况下，接口的 Carry-delay 值为 2 秒
- 【命令模式】 接口配置模式
- 【使用指导】 -以毫秒为单位设置载波延迟必须是 100 毫秒的整数倍

配置接口 Load-interval

- 可选配置。
- 配置的报文采样时间影响接口报文平均速率的计算，配置的时间较短时，报文平均速率能较快反映报文实时流量的变化。

- 【命令格式】 **load-interval seconds**
- 【参数说明】 *seconds* : 以秒为单位，范围 5-600 秒。
- 【缺省配置】 缺省情况下，接口的 load-interval 值为 10 秒
- 【命令模式】 接口配置模式
- 【使用指导】 -

设置保护口

- 可选配置。
- 配置为保护口的端口之间无法进行二层报文转发。
- 支持以太网物理端口上及聚合口上配置。

【命令格式】 **switchport protected**

【参数说明】 -

【缺省配置】 缺省情况下，接口不是一个保护口

【命令模式】 接口配置模式

【使用指导】 -

保护口之间三层路由阻断

- 可选配置。
- 配置了该命令后，配置了保护口命令的端口之间无法进行三层路由转发。

【命令格式】 **protected-ports route-deny**

【参数说明】 -

【缺省配置】 缺省情况下，该功能关闭

【命令模式】 全局配置模式

【使用指导】 缺省情况下，保护口之间的三层路由并没有被阻断，这个时候可以通过设置保护口之间不能进行路由的功能来实现保护口之间的路由阻断功能。

端口违例恢复

- 可选配置。
- 端口违例发生后，端口被关闭，缺省情况下不会恢复。配置了端口违例恢复后，违例的端口会被恢复，端口会被打开。

【命令格式】 **errdisable recovery [interval time]**

【参数说明】 *time*：自动恢复定时时间，取值范围为 30-86400，单位是秒。

【缺省配置】 缺省情况下，没有该功能

【命令模式】 全局配置模式

【使用指导】 缺省情况下，端口违例不会恢复，这个时候可以使用此命令手动恢复或者配置自动恢复。

检验方法

- 可以通过 **show interfaces** 命令查看接口的属性配置是否正常。

【命令格式】 **show interfaces [interface-type interface-number] [description | switchport | trunk]**

【参数说明】 *interface-type interface-number*：接口类型和接口编号

description：接口的描述符信息，包括 link 状态

switchport：二层接口信息，只对二层接口有效

trunk：Trunk 端口信息，对物理端口和聚合端口有效

【命令模式】 特权模式

【使用指导】 如果不加参数，则显示接口的基本信息

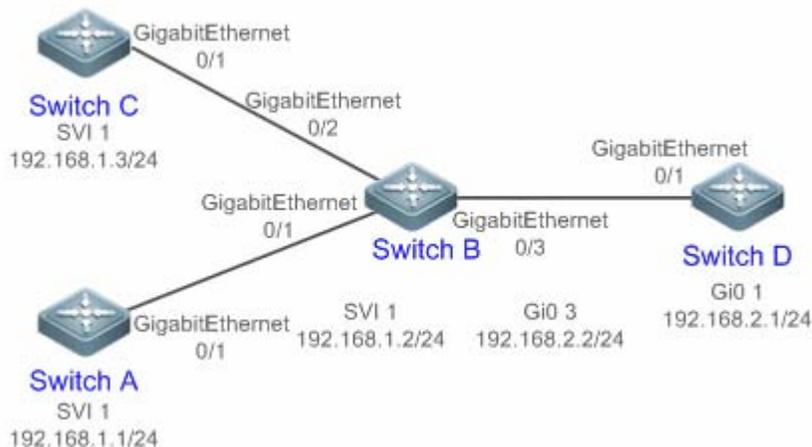
【命令展示】

```
SwitchA#show interfaces GigabitEthernet 0/1
Index(dec):1 (hex):1
GigabitEthernet 0/1 is DOWN , line protocol is DOWN
  Hardware is Broadcom 5464 GigabitEthernet, address is 00d0.f865.de9b (bia 00d0.f865.de9b)
  Interface address is: no ip address
  MTU 1500 bytes, BW 1000000 Kbit
  Encapsulation protocol is Ethernet-II, loopback not set
  Keepalive interval is 10 sec , set
  Carrier delay is 2 sec
  Ethernet attributes:
    Last link state change time: 2012-12-22 14:00:48
    Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds
    Priority is 0
    Medium-type is Copper
    Admin duplex mode is AUTO, oper duplex is Unknown
    Admin speed is AUTO, oper speed is Unknown
    Flow receive control admin status is OFF,flow send control admin status is OFF
    Flow receive control oper status is Unknown,flow send control oper status is Unknown
    Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
  Bridge attributes:
    Port-type: trunk
    Native vlan:1
    Allowed vlan lists:1-4094 //Trunk 口的许可 VLAN 列表
    Active vlan lists:1, 3-4 //实际生效的 vlan ( 即该设备上仅创建了 VLAN1、3 和 4 )
  Queueing strategy: FIFO
    Output queue 0/0, 0 drops;
    Input queue 0/75, 0 drops
  Rxload is 1/255,Txload is 1/255
  5 minutes input rate 0 bits/sec, 0 packets/sec
  5 minutes output rate 0 bits/sec, 0 packets/sec
    0 packets input, 0 bytes, 0 no buffer, 0 dropped
    Received 0 broadcasts, 0 runts, 0 giants
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
    0 packets output, 0 bytes, 0 underruns , 0 dropped
    0 output errors, 0 collisions, 0 interface resets
```

配置举例

配置接口属性

【网络环境】

图
1-6

【配置方法】

- 在 Switch A 上配置 GigabitEthernet 0/1 为 Access 模式交换端口，缺省 VLAN ID 为 1，配置 SVI 1，并为 SVI1 配置 IP 以及到 Switch D 的路由。
- 在 Switch B 上配置 GigabitEthernet 0/1 和 GigabitEthernet 0/2 为 Trunk 模式交换端口，Native VLAN ID 为 1，并配置 SVI 1，并为 SVI 1 配置 IP，配置 GigabitEthernet 0/3 为路由口并为该端口配置另一个网段的 IP。
- 在 Switch C 上配置 GigabitEthernet 0/1 为 Access 模式交换端口，缺省 VLAN ID 为 1，配置 SVI 1，并为 SVI 1 配置 IP。
- 在 Switch D 上配置 GigabitEthernet 0/1 为路由端口，并为该端口配置 IP 以及到 Switch A 的路由。

A

```
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# switchport mode access
A(config-if-GigabitEthernet 0/1)# switchport access vlan 1
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip address 192.168.1.1 255.255.255.0
A(config-if-VLAN 1)# exit
A(config)# ip route 192.168.2.0 255.255.255.0 VLAN 1 192.168.1.2
```

B

```
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# switchport mode trunk
B(config-if-GigabitEthernet 0/1)# exit
B(config)# interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/2)# switchport mode trunk
B(config-if-GigabitEthernet 0/2)# exit
B(config)# interface vlan 1
B(config-if-VLAN 1)# ip address 192.168.1.2 255.255.255.0
B(config-if-VLAN 1)# exit
B(config)# interface GigabitEthernet 0/3
```

```
B(config-if-GigabitEthernet 0/3)# no switchport
B(config-if-GigabitEthernet 0/3)# ip address 192.168.2.2 255.255.255.0
B(config-if-GigabitEthernet 0/3)# exit
```

C

```
C# configure terminal
C(config)# interface GigabitEthernet 0/1
C(config-if-GigabitEthernet 0/1)# port-group 1
C(config-if-GigabitEthernet 0/1)# exit
C(config)# interface aggregateport 1
C(config-if-AggregatePort 1)# switchport mode access
C(config-if-AggregatePort 1)# switchport access vlan 1
C(config-if-AggregatePort 1)# exit
C(config)# interface vlan 1
C(config-if-VLAN 1)# ip address 192.168.1.3 255.255.255.0
C(config-if-VLAN 1)# exit
```

D

```
D# configure terminal
D(config)# interface GigabitEthernet 0/1
D(config-if-GigabitEthernet 0/1)# no switchport
D(config-if-GigabitEthernet 0/1)# ip address 192.168.2.1 255.255.255.0
D(config-if-GigabitEthernet 0/1)# exit
A(config)# ip route 192.168.1.0 255.255.255.0 GigabitEthernet 0/1 192.168.2.2
```

【检验方法】 在 A、B、C、D 四台设备上分别进行如下检验：

- A Ping 其它 3 台设备的接口 IP，两两之间可以相互访问。
- B 和 D 互 Ping 能通。
- 检查接口状态是否正确。

A

```
A# show interfaces gigabitEthernet 0/1
Index(dec):1 (hex):1
GigabitEthernet 0/1 is UP , line protocol is UP
Hardware is GigabitEthernet, address is 00d0.f865.de90 (bia 00d0.f865.de90)
Interface address is: no ip address
  MTU 1500 bytes, BW 100000 Kbit
  Encapsulation protocol is Ethernet-II, loopback not set
  Keepalive interval is 10 sec , set
  Carrier delay is 2 sec
Ethernet attributes:

  Last link state change time: 2012-12-22 14:00:48

  Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds

  Priority is 0

  Admin medium-type is Copper, oper medium-type is Copper
  Admin duplex mode is AUTO, oper duplex is Full
```

```
Admin speed is AUTO, oper speed is 100M
Flow control admin status is OFF, flow control oper status is OFF
Admin negotiation mode is OFF, oper negotiation state is ON
Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
```

Bridge attributes:

Port-type: access

Vlan id: 1

```
Rxload is 1/255, Txload is 1/255
10 seconds input rate 0 bits/sec, 0 packets/sec
10 seconds output rate 67 bits/sec, 0 packets/sec
 362 packets input, 87760 bytes, 0 no buffer, 0 dropped
Received 0 broadcasts, 0 runts, 0 giants
 0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
 363 packets output, 82260 bytes, 0 underruns , 0 dropped
 0 output errors, 0 collisions, 0 interface resets
```

B

```
B# show interfaces gigabitEthernet 0/1
Index(dec):1 (hex):1
GigabitEthernet 0/1 is UP , line protocol is UP
Hardware is GigabitEthernet, address is 00d0.f865.de91 (bia 00d0.f865.de91)
Interface address is: no ip address
  MTU 1500 bytes, BW 100000 Kbit
  Encapsulation protocol is Ethernet-II, loopback not set
  Keepalive interval is 10 sec , set
  Carrier delay is 2 sec
Ethernet attributes:

  Last link state change time: 2012-12-22 14:00:48

  Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds

  Priority is 0

  Admin medium-type is Copper, oper medium-type is Copper
  Admin duplex mode is AUTO, oper duplex is Full
  Admin speed is AUTO, oper speed is 100M
  Flow control admin status is OFF, flow control oper status is OFF
  Admin negotiation mode is OFF, oper negotiation state is ON
  Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF

Bridge attributes:

  Port-type: trunk
  Native vlan: 1
  Allowed vlan lists: 1-4094
  Active vlan lists: 1

Rxload is 1/255, Txload is 1/255
```

```
10 seconds input rate 0 bits/sec, 0 packets/sec
10 seconds output rate 67 bits/sec, 0 packets/sec
 362 packets input, 87760 bytes, 0 no buffer, 0 dropped
Received 0 broadcasts, 0 runts, 0 giants
 0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
 363 packets output, 82260 bytes, 0 underruns , 0 dropped
 0 output errors, 0 collisions, 0 interface resets
```

C

```
C# show interfaces gigabitEthernet 0/1
Index(dec):1 (hex):1
GigabitEthernet 0/1 is UP , line protocol is UP
Hardware is GigabitEthernet, address is 00d0.f865.de92 (bia 00d0.f865.de92)
Interface address is: no ip address
  MTU 1500 bytes, BW 100000 Kbit
  Encapsulation protocol is Ethernet-II, loopback not set
  Keepalive interval is 10 sec , set
  Carrier delay is 2 sec
Ethernet attributes:

  Last link state change time: 2012-12-22 14:00:48
  Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds
  Priority is 0
  Admin medium-type is Copper, oper medium-type is Copper
  Admin duplex mode is AUTO, oper duplex is Full
  Admin speed is AUTO, oper speed is 100M
  Flow control admin status is OFF, flow control oper status is OFF
  Admin negotiation mode is OFF, oper negotiation state is ON
  Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF
Rxload is 1/255, Txload is 1/255
10 seconds input rate 0 bits/sec, 0 packets/sec
10 seconds output rate 67 bits/sec, 0 packets/sec
 362 packets input, 87760 bytes, 0 no buffer, 0 dropped
Received 0 broadcasts, 0 runts, 0 giants
 0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
 363 packets output, 82260 bytes, 0 underruns , 0 dropped
 0 output errors, 0 collisions, 0 interface resets
```

D

```
D# show interfaces gigabitEthernet 0/1
Index(dec):1 (hex):1
GigabitEthernet 0/1 is UP , line protocol is UP
Hardware is GigabitEthernet, address is 00d0.f865.de93 (bia 00d0.f865.de93)
Interface address is: 192.168.2.1/24
  MTU 1500 bytes, BW 100000 Kbit
```

```

Encapsulation protocol is Ethernet-II, loopback not set
Keepalive interval is 10 sec , set
Carrier delay is 2 sec
Ethernet attributes:

  Last link state change time: 2012-12-22 14:00:48

  Time duration since last link state change: 3 days, 2 hours, 50 minutes, 50 seconds

  Priority is 0

  Admin medium-type is Copper, oper medium-type is Copper
  Admin duplex mode is AUTO, oper duplex is Full
  Admin speed is AUTO, oper speed is 100M

  Flow control admin status is OFF, flow control oper status is OFF
  Admin negotiation mode is OFF, oper negotiation state is ON

  Storm Control: Broadcast is OFF, Multicast is OFF, Unicast is OFF

Rxload is 1/255, Txload is 1/255
10 seconds input rate 0 bits/sec, 0 packets/sec
10 seconds output rate 67 bits/sec, 0 packets/sec
  362 packets input, 87760 bytes, 0 no buffer, 0 dropped
  Received 0 broadcasts, 0 runts, 0 giants
  0 input errors, 0 CRC, 0 frame, 0 overrun, 0 abort
  363 packets output, 82260 bytes, 0 underruns , 0 dropped
  0 output errors, 0 collisions, 0 interface resets

```

常见错误

- 无。

1.6 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除接口的统计值。	clear counters [<i>interface-type interface-number</i>]
接口硬件复位。	clear interface <i>interface-type interface-number</i>

查看运行情况

📄 显示接口配置和状态

作用	命令
显示指定接口的全部状态和配置信息。	show interfaces [<i>interface-type interface-number</i>]
显示接口的状态。	show interfaces [<i>interface-type interface-number</i>] status
显示接口违例状态。	show interfaces [<i>interface-type interface-number</i>] status err-disable
查看端口链路状态变化时间和次数。	show interfaces [<i>interface-type interface-number</i>] link-state-change statistics
显示可交换接口（非路由接口）的 administrative 和 operational 状态信息。	show interfaces [<i>interface-type interface-number</i>] switchport
显示指定接口的描述配置和接口状态。	show interfaces [<i>interface-type interface-number</i>] description
显示指定端口的统计值信息 其中速率显示可能有 0.5% 内的误差。	show interfaces [<i>interface-type interface-number</i>] counters
显示上一个采样时间间隔内增加的报文统计值。	show interfaces [<i>interface-type interface-number</i>] counters increment
显示错误报文统计值。	show interfaces [<i>interface-type interface-number</i>] counters error
显示接口报文收发速率	show interfaces [<i>interface-type interface-number</i>] counters rate
显示接口简要统计值	show interfaces [<i>interface-type interface-number</i>] counters summary
线缆检测状态显示。在线缆处于短路或断路等异常状态时 线缆检测有助于正确判断线缆的工作状况。	show interfaces [<i>interface-type interface-number</i>] line-detect
显示接口带宽利用率	show interfaces [<i>interface-type interface-number</i>] usage

显示光模块信息

作用	命令
显示指定接口的光模块基本信息。	show interfaces [<i>interface-type interface-number</i>] transceiver
显示指定接口的光模块当前故障告警信息，当没有故障时显示 “None”。	show interfaces [<i>interface-type interface-number</i>] transceiver alarm
显示指定接口的光模块诊断参数的当前测量值。	show interfaces [<i>interface-type interface-number</i>] transceiver diagnosis

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

无。

线缆检测

管理员可以通过线缆检测命令来检测线缆的工作状况。在线缆处于短路或断路等异常状态时，线缆检测有助于正确判断线缆的工作状况。

 只有电介质的物理口才支持线缆检测，光介质物理口、聚合端口不支持线缆检测。

 在正常连接的接口执行线缆检测，会导致连接暂时断掉，然后再重新建立连接。

作用	命令
在接口模式下执行线缆检测。在线缆处于短路或断路等异常状态时，线缆检测有助于正确判断线缆的工作状况。	line-detect

2 MAC 地址

2.1 概述

MAC 地址表记录了与该设备相连的设备的 MAC 地址、接口号以及所属的 VLAN ID。

设备在转发报文时通过报文的目的 MAC 地址以及报文所属的 VLAN ID 的信息在 MAC 地址表中查找相应的转发输出端口。

根据 [mac 地址](#) 查找到转发出口后就可以采取单播、组播或广播的方式转发报文。

i 本文只涉及动态地址、静态地址与过滤地址的管理，组播地址的管理不在本文内描述，请参看《IGMP Snooping 配置指南》。

协议规范

- IEEE 802.3 : Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications
- IEEE 802.1Q : Virtual Bridged Local Area Networks

2.2 典型应用

典型应用	场景描述
动态地址学习	通过动态地址学习，实现报单播转发
MAC地址变化通知	通过 MAC 地址添加删除通知，监控网络设备下用户变化。

2.2.1 动态地址学习

应用场景

通常情况下 MAC 地址表的维护都是通过动态地址学习的方式进行，其工作原理如下：

设备的 MAC 地址表为空的情况下，UserA 要与 UserB 进行通讯，UserA 首先发送报文到交换机的端口 GigabitEthernet 0/2，此时设备将 UserA 的 MAC 地址学习到 MAC 地址表中。

由于地址表中没有 UserB 的源 MAC 地址，因此设备以广播的方式将报文发送到除了 UserA 以外的所有端口，包括 User B 与 User C 的端口，此时 UserC 能够收到 UserA 所发出的不属于它的报文。

图 2-1 动态地址学习步骤一

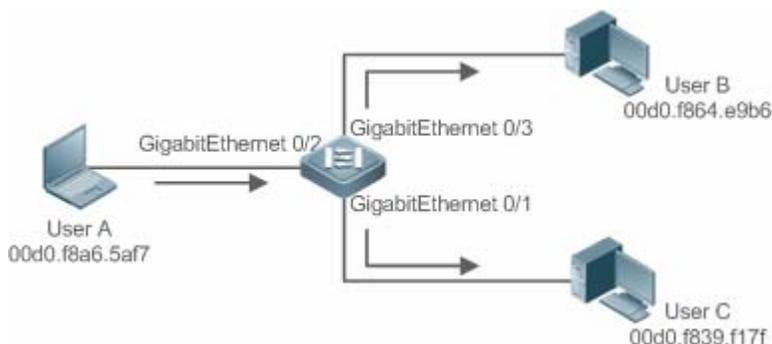


图 2-2 以太网交换 MAC 地址表一

Status	VLAN	MAC地址	端口
动态	1	00d0.f8a6.5af7	GigabitEthernet 0/2

UserB 收到报文后将回应报文通过设备的端口 GigabitEthernet 0/3 发送 UserA，此时设备的 MAC 地址表中已存在 UserA 的 MAC 地址，所以报文被以单播的方式转发到 GigabitEthernet 0/2 端口，同时设备将学习 UserB 的 MAC 地址，与步骤 1 中所不同的是 UserC 此时接收不到 UserB 发送给 UserA 的报文。

图 2-3 动态地址学习步骤二

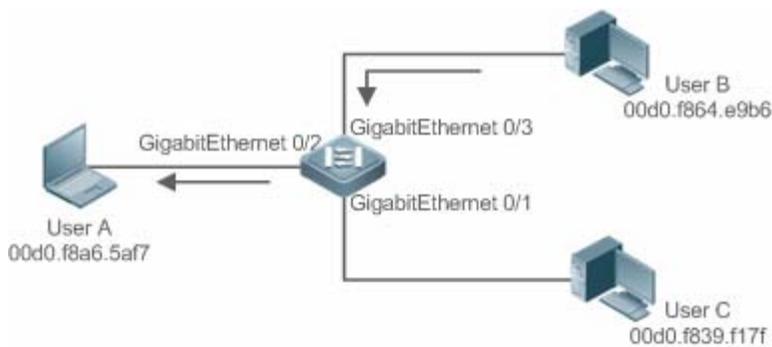


图 2-4 设备 MAC 地址表二

Status	VLAN	MAC地址	端口
动态	1	00d0.f8a6.5af7	GigabitEthernet 0/2
动态	1	00d0.f864.e9b6	GigabitEthernet 0/3

通过 UserA 与 UserB 的一次交互过程后，设备学习到了 UserA 与 UserB 的源 MAC 地址，之后 UserA 与 UserB 之间的报文交互则采用单播的方式进行转发，此后 UserC 将不再接收到 UserA 与 UserB 之间的交互报文。

功能部属

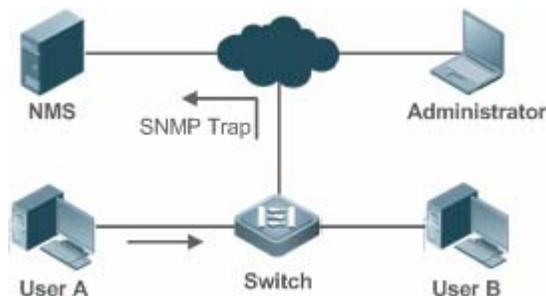
- 二层交换设备通过动态地址学习，实现报文单播转发，减少广播报文，减轻网络不必要的负荷。

2.2.2 MAC地址变化通知

设备的 MAC 地址通知功能通过与网络管理工作站（NMS）的协作为网络管理提供了监控网络设备下用户变化的机制。

应用场景

图 2-5 MAC 地址通知



打开 MAC 地址通知的功能后，当设备学习到一个新的 MAC 地址或老化掉一个已学习到的 MAC 地址时，一个反映 MAC 地址变化的通知信息就会产生，并以 SNMP Trap 的方式将通知信息发送给指定的 NMS(网络管理工作站)。

当一个 MAC 地址增加的通知产生，就可以知道一个由此 MAC 地址标识的新用户开始使用网络，当一个 MAC 地址删除的通知产生，则表示一个用户在地址老化时间内没有新的报文发送，通常可以认为此用户已经停止使用网络了。

当使用设备下接的用户较多时，可能会出现短时间内会有大量的 MAC 地址变化产生，导致网络流量增加。为了减轻网络负担，可以设置发送 MAC 地址通知的时间间隔。在达到配置的时间间隔之后，系统将这个时间内的所有通知信息进行打包封装，此时在每条地址通知信息中，包含了若干个 MAC 地址变化的信息，从而可以有效地减少网络流量。

当 MAC 地址通知产生时，通知信息同时会记录到 MAC 地址通知历史记录表中。此时即便没有配置接收 Trap 的 NMS，管理员也可以通过查看 MAC 地址通知历史记录表来了解最近 MAC 地址变化的消息。

i MAC 地址通知仅对动态地址有效，对于配置的静态地址与过滤地址的变化将不会产生通知信息。

功能部属

- 二层交换设备开启 MAC 地址变化通知，实现监控网络设备下的用户变化。

2.3 功能详解

基本概念

▾ 动态地址

通过设备的自动地址学习过程产生的 MAC 地址表项被称为动态地址。

地址老化

设备的 MAC 地址表是有容量限制的，设备采用地址表老化机制进行不活跃的地址表项淘汰。

设备在学习到一个新的地址的同时启动该地址的老化记时，在达到老化记时前，如果设备没有再一次收到以该地址为源 MAC 地址的报文，则该地址在达到老化时间后会从 MAC 地址表中删除。

单播转发

- 设备能够在 MAC 地址表中查到与报文的源 MAC 地址和 VLAN ID 相对应的表项并且表项中的输出端口是唯一的，报文直接从表项对应的端口输出。

广播转发

- 设备收到目的地址为 ffff.ffff.ffff 的报文或者在 MAC 地址表中查找不到对应的表项时，报文被送到所属的 VLAN 中除报文输入端口外的其他所有端口输出。

功能特性

功能特性	作用
VLAN 的动态地址个数限制	用户可规划各个 VLAN 内可学习的动态地址数
接口的动态地址个数限制	用户可规划各个接口下可学习的动态地址数

2.3.1 VLAN 的动态地址个数限制

工作原理

设备的 MAC 地址表的容量是有限制的并且所有的 VLAN 共享整个 MAC 地址表容量，为避免一个 VLAN 内大量动态地址将整个 MAC 地址表所占用的而使其他 VLAN 无法学习动态地址，导致其他 VLAN 的报文都采用广播方式转发，锐捷设备提供了 VLAN 的动态地址个数限制功能，用户可以规划各个 VLAN 内可学习的动态地址数，为每个 VLAN 配置可动态学习的地址的个数上限。配置了 VLAN 的动态地址个数限制功能的 VLAN 只能学到用户所指定个数的 MAC 地址，对超出用户配置上限部分的地址将不再学习，以这些地址为目的地址的报文将以广播方式转发。

- 如果配置 VLAN 的动态地址学习个数限制的上限小于当前 VLAN 中已学习到的动态地址数，此时设备不再学习该 VLAN 中的地址，直到 VLAN 内的地址数通过地址老化删除到小于上限后，设备才会重新学习。
- MAC 地址复制功能，复制到指定 VLAN 的 MAC 地址表项，不受该 VLAN 下动态 MAC 地址学习个数的限制。

2.3.2 接口的动态地址个数限制

工作原理

配置了接口的动态地址个数限制功能的接口只能学到用户所指定个数的 MAC 地址，对超出用户配置上限部份的地址将不再学习，以这些地址为目的地址的报文将以广播方式转发。

- i** 如果配置接口的动态地址学习个数限制的上限小于当前接口下已学习到的动态地址数，此时设备不再学习该接口下的地址，直到接口下的地址数通过地址老化删除到小于上限后，设备才会重新学习。

2.4 产品说明



S7508 产品：源 MAC 为全 0，目的 MAC 为全 0 的报文，不学习且不转发。



S7508 线卡：配置一条过滤地址，需占用 2 条地址容量。如全局过滤 MAC 地址数目为 10K，整机可用地址容量减少 20K。

2.5 配置详解

配置项	配置建议 & 相关命令	
配置动态地址	可选配置。用于实现动态地址学习。	
	mac-address-learning	配置全局或接口 MAC 地址学习能力
	mac-address-table aging-time	配置动态地址老化时间
配置静态地址	可选配置。用于绑定设备下接的网络设备的 MAC 地址与端口关系。	
	mac-address-table static	配置静态地址
配置过滤地址	可选配置。用于过滤报文。	
	mac-address-table filtering	配置过滤地址
配置MAC地址变化通知	可选配置。用于监控网络设备下的用户变化。	
	mac-address-table notification	配置全局 MAC 地址变化通知功能
	snmp trap mac-notification	配置接口 MAC 地址变化通知功能
配置 MAC 地址漂移打印 Syslog告警功能	可选配置。用于配置 MAC 地址漂移被检测到后，打印 Syslog 的功能。	
	mac-address-table flapping-logging	开启 MAC 地址漂移打印 Syslog 告警功能

2.5.1 配置动态地址

配置效果

实现动态地址学习，报文正常单播转发。

配置方法

配置全局 MAC 地址学习能力

- 可选配置。
- 如果需要关闭全局 MAC 地址学习能力，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **mac-address-learning { enable | disable }**

【参数说明】 **enable**：开启全局 MAC 地址学习能力

disable：关闭全局 MAC 地址学习能力

【缺省配置】 全局地址学习能力开启

【命令模式】 全局模式

【使用指导】 -

- ❗ 全局 MAC 地址学习能力缺省开启。当全局 MAC 地址学习能力关闭时，全局无法进行 MAC 地址学习；当全局 MAC 地址学习能力开启时，按端口的 MAC 地址学习能力生效。

配置接口 MAC 地址学习能力

- 可选配置。。
- 如果需要关闭接口 MAC 地址学习能力，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **mac-address-learning**

【参数说明】 -

【缺省配置】 地址学习能力开启

【命令模式】 接口模式

【使用指导】 接口必须是二层接口，包括交换口、AP 口。

- ❗ MAC 地址学习能力缺省开启，如果端口上配置了 DOT1X、IP SOURCE GUARD 绑定，端口安全功能，端口的学习能力不能开启；同样，关闭端口学习能力的端口不能开启接入控制功能。

配置动态地址老化时间

- 可选配置。
- 如果需要修改动态地址老化时间，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **mac-address-table aging-time value**

【参数说明】 **value**：老化时间。取值范围{ 0 | 10 - 1000000 }，缺省值 300 秒。

【缺省配置】 缺省值是 300 秒

【命令模式】 全局模式

【使用指导】 当设置该值为 0 时，地址老化功能将被关闭，学习到的地址将不会被老化。

- ❗ 地址表的实际老化时间会与设定值存在一定偏差，但不会超过设定值的 2 倍。

检验方法

- 检查设备是否能正常学习动态地址。
- 通过 `show mac-address-table dynamic` 命令查看动态地址信息。
- 通过 `show mac-address-table aging-time` 命令查看动态地址老化时间。

【命令格式】 **show mac-address-table dynamic** [**address** *mac-address*] [**interface** *interface-id*] [**vlan** *vlan-id*]

【参数说明】 **address** *mac-address* : 查看设备上特定动态 MAC 地址信息。

interface *interface-id* : 指定的物理接口或是 Aggregate Port。

vlan *vlan-id* : 查看特定的 VLAN 中的动态地址。

【命令模式】 特权模式，全局模式，接口模式

【使用指导】 -

【命令展示】

```
Ruijie# show mac-address-table dynamic
Vlan      MAC Address      Type      Interface
-----
1         0000.0000.0001   DYNAMIC   GigabitEthernet 1/1
1         0001.960c.a740   DYNAMIC   GigabitEthernet 1/1
1         0007.95c7.dff9   DYNAMIC   GigabitEthernet 1/1
1         0007.95cf.eee0   DYNAMIC   GigabitEthernet 1/1
1         0007.95cf.f41f   DYNAMIC   GigabitEthernet 1/1
1         0009.b715.d400   DYNAMIC   GigabitEthernet 1/1
1         0050.bade.63c4   DYNAMIC   GigabitEthernet 1/1
```

字段解释：

字段	说明
Vlan	MAC 地址所在的 VLAN
MAC Address	MAC 地址
Type	MAC 地址类型
Interface	MAC 地址所在的接口

【命令格式】 **show mac-address-table aging-time**

【参数说明】 -

【命令模式】 特权模式，全局模式，接口模式

【使用指导】 -

【命令展示】

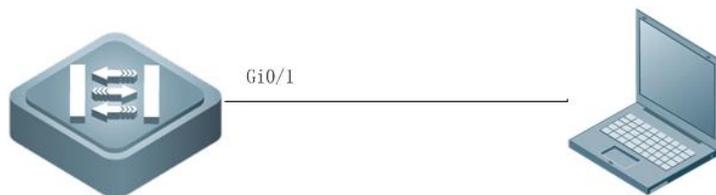
```
Ruijie# show mac-address-table aging-time
Aging time      : 300
```

配置举例

配置动态地址

【网络环境】

图 2-6



【配置方法】

- 打开接口 MAC 地址学习能力
- 配置动态地址老化时间为 180 秒
- 删除接口 GigabitEthernet 0/1 下 VLAN 1 中的所有动态地址

```
Ruijie# configure terminal
Ruijie(config-if-GigabitEthernet 0/1)# mac-address-learning
Ruijie(config-if-GigabitEthernet 0/1)# exit
Ruijie(config)# mac aging-time 180
Ruijie# clear mac-address-table dynamic interface GigabitEthernet 0/1 vlan 1
```

【检验方法】

- 查看接口 MAC 地址学习能力
- 查询动态地址老化时间
- 查看接口 GigabitEthernet 0/1 下 VLAN 1 中的所有动态地址

```
Ruijie# show mac-address-learning
GigabitEthernet 0/1      learning ability: enable
Ruijie# show mac aging-time
Aging time      : 180 seconds
Ruijie# show mac-address-table dynamic interface GigabitEthernet 0/1 vlan 1
```

Vlan	MAC Address	Type	Interface
1	00d0.f800.1001	STATIC	GigabitEthernet 1/1

常见错误

配置接口地址学习能力时，接口没有先配置成二层接口，包括交换口、AP 口。

2.5.2 配置静态地址

配置效果

- 配置静态地址，绑定设备下接的网络设备的 MAC 地址与端口关系。

配置方法

▾ 配置静态地址

- 可选配置。。
- 如果需要绑定设备下接的网络设备的 MAC 地址与端口关系，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **mac-address-table static** *mac-address* **vlan** *vlan-id* **interface** *interface-id*

【参数说明】 **address** *mac-address* : 指定要删除的 MAC 地址

vlan *vlan-id* : 指定要删除的 MAC 地址所在的 VLAN。

interface *interface-id* : 指定的物理接口或是 Aggregate Port。

【缺省配置】 缺省没有设置任何静态地址

【命令模式】 全局模式

【使用指导】 当设备在 *vlan-id* 指定的 VLAN 上接收到以 *mac-address* 为目的地址的报文时，这个报文将被转发到 *interface-id* 所指定的接口上。

检验方法

- 通过命令 `show mac-address-table static` 显示静态地址信息是否正确。

【命令格式】 **show mac-address-table static** [**address** *mac-address*] [**interface** *interface-id*] [**vlan** *vlan-id*]

【参数说明】 **address** *mac-address* : 查看设备上特定静态 MAC 地址信息。

interface *interface-id* : 指定的物理接口或是 Aggregate Port。

vlan *vlan-id* : 查看特定的 VLAN 中的静态地址。

【命令模式】 特权模式，全局模式，接口模式

【使用指导】 -

【命令展示】

```
Ruijie# show mac-address-table static
Vlan    MAC Address      Type      Interface
-----
1       00d0.f800.1001  STATIC   GigabitEthernet 1/1
1       00d0.f800.1002  STATIC   GigabitEthernet 1/1
1       00d0.f800.1003  STATIC   GigabitEthernet 1/1
```

配置举例

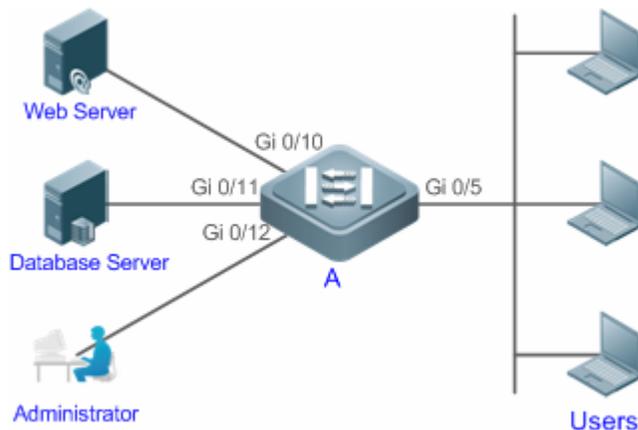
配置静态地址

本例的 MAC 地址同 VLAN、接口对应关系如下表所示：

角色	MAC 地址	VLAN ID	接口 ID
Web 服务器	00d0.3232.0001	VLAN2	Gi0/10
信息服务器	00d0.3232.0002	VLAN2	Gi0/11
网络管理员	00d0.3232.1000	VLAN2	Gi0/12

【网络环境】

图 2-7



【配置方法】

- 指定表项对应的目的 MAC 地址 (Mac-address)
- 指定该地址所属的 VLAN (vlan-id)
- 接口 ID (Interface-id)

A

```
A# configure terminal
A(config)# mac-address-table static 00d0.f800.3232.0001 vlan 2 interface gigabitEthernet 0/10
A(config)# mac-address-table static 00d0.f800.3232.0002 vlan 2 interface gigabitEthernet 0/11
A(config)# mac-address-table static 00d0.f800.3232.1000 vlan 2 interface gigabitEthernet 0/12
```

【检验方法】

在交换机上查看配置的静态 MAC 地址

A

```
A# show mac-address-table static
```

Vlan	MAC Address	Type	Interface
2	00d0.f800.3232.0001	STATIC	GigabitEthernet 0/10
2	00d0.f800.3232.0002	STATIC	GigabitEthernet 0/11
2	00d0.f800.3232.1000	STATIC	GigabitEthernet 0/12

常见错误

- 配置静态地址时，指定接口没有先配置成二层接口，包括交换口、AP 口。

2.5.3 配置过滤地址

配置效果

- 配置过滤地址，当在对应 VLAN 中接收到源 MAC 或目的 MAC 为过滤地址的报文时，将丢弃此报文。

配置方法

配置过滤地址

- 可选配置。
- 如果需要过滤报文，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **mac-address-table filtering** *mac-address* **vlan** *vlan-id*

【参数说明】 **address** *mac-address*：指定要删除的 MAC 地址
vlan *vlan-id*：指定要删除的 MAC 地址所在的 VLAN。

【缺省配置】 缺省没有设置任何过滤地址

【命令模式】 全局模式

【使用指导】 当设备在 *vlan-id* 指定的 VLAN 上接收到以 *mac-address* 指定的地址为源地址或目的地址的报文将被丢弃。

检验方法

- 通过命令 `show mac-address-table filter` 显示过滤地址信息。

【命令格式】 **show mac-address-table filter** [**address** *mac-address*] [**vlan** *vlan-id*]

【参数说明】 **address** *mac-address*：查看设备上特定过滤 MAC 地址信息。
vlan *vlan-id*：查看特定的 VLAN 中的过滤地址。

【命令模式】 特权模式，全局模式，接口模式

【使用指导】 -

【命令展示】

```
Ruijie# show mac-address-table filtering
Vlan      MAC Address      Type      Interface
-----
1         0000.2222.2222  FILTER
```

配置举例

配置过滤地址

- 【配置方法】
- 指定过滤地址对应的目的 MAC 地址 (Mac-address)
 - 指定过滤地址所属的 VLAN (vlan-id)

```
Ruijie# configure terminal
Ruijie(config)# mac-address-table static 00d0.f800.3232.0001 vlan 1
```

【检验方法】 在交换机上查看配置的过滤 MAC 地址

```
Ruijie# show mac-address-table filter
Vlan      MAC Address      Type      Interface
-----
1         00d0.f800.3232.0001  FILTER
```

常见错误

无。

2.5.4 配置MAC地址变化通知

配置效果

- 配置 MAC 地址变化通知，监控网络设备下的用户变化。

配置方法

配置接收 MAC 地址通知的 NMS

- 可选配置。
- 如果需要接收 MAC 地址通知，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **snmp-server host** *host-addr* **traps** [**version** { 1 | 2c | 3 [**auth** | **noauth** | **priv**] }] *community-string*

【参数说明】 **host** *host-addr* : 指明接收者的 IP。

version { 1 | 2c | 3 [**auth** | **noauth** | **priv**] } : 指明发送哪种版本的 snmp trap 报文，对 v3 版本还可以指定是否认证以及安全等级参数。

community-string : 认证名

【缺省配置】 缺省不需要配置

【命令模式】 全局模式

【使用指导】 -

配置使能发送 Trap 功能

- 可选配置。
- 如果需要发送 Trap，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **snmp-server enable traps**

【参数说明】 -

【缺省配置】 缺省不需要配置

【命令模式】 全局模式

【使用指导】 -

配置全局 MAC 地址通知开关

- 可选配置。
- 全局开关被关闭，所有接口的 MAC 地址通知功能也均被关闭。
- 交换机设备上配置。

【命令格式】 **mac-address-table notification**

- 【参数说明】 -
- 【缺省配置】 缺省全局 MAC 地址变化通知开关关闭
- 【命令模式】 全局模式
- 【使用指导】 -

配置接口 MAC 地址通知开关

- 可选配置
- 如果需要接收接口 MAC 地址变化通知，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **snmp trap mac-notification { added | removed }**

【参数说明】 **added**：当地址增加时通知。
removed：当地址被删除时通知。

【缺省配置】 缺省接口 MAC 地址变化通知开关关闭

【命令模式】 接口模式

【使用指导】 -

配置 MAC 地址通知的时间间隔与历史记录容量

- 可选配置。
- 如果需要修改 MAC 地址通知的时间间隔或历史记录容量，则应该执行此配置项。
- 交换机设备上配置。

【命令格式】 **mac-address-table notification { interval value | history-size value }**

【参数说明】 **interval value**：设置产生 MAC 地址通知的时间间隔(可选)。时间间隔的单位为秒，范围为 1 - 3600，缺省为 1 秒。

history-size value：MAC 通知历史记录表中记录的最大个数，范围 1 - 200，缺省为 50。

【缺省配置】 时间间隔缺省为 1 秒，表项默认通告的最大通告个数为 50。

【命令模式】 全局模式

【使用指导】 -

检验方法

- 通过命令 **show mac-address-table notification** 检查 NMS 是否能正常接收 MAC 地址变化通知。

【命令格式】 **show mac-address-table notification [interface [interface-id] | history]**

【参数说明】 **interface**：显示全部接口上的 MAC 通知功能设置。
interface-id：查看接口的 MAC 地址变化通知的使能状况。
history：查看 MAC 地址变化通知信息的历史记录表。

【命令模式】 特权模式，全局模式，接口模式

【使用指导】 -

【使用指导】 1、查看 MAC 地址通告功能的全局配置信息

```
Ruijie#show mac-address-table notification
MAC Notification Feature : Enabled
Interval(Sec): 300
```

Maximum History Size : 50

Current History Size : 0

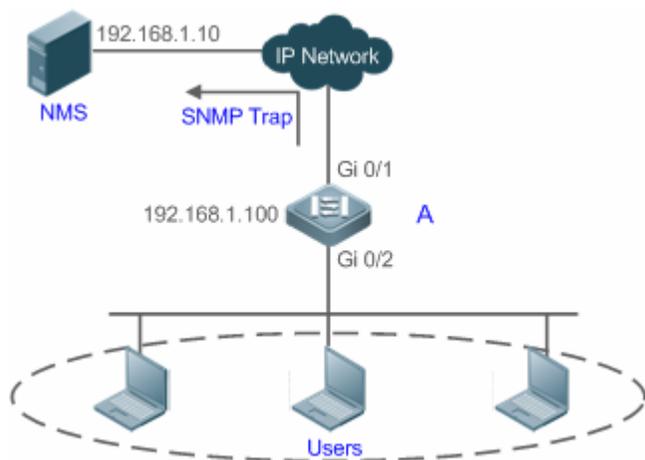
字段解释：

字段	说明
Interval(Sec)	通告 MAC 地址的时间间隔
Maximum History Size	MAC 地址通告历史记录表的最大表项个数
Current History Size	当前记录条目数

配置举例

【网络环境】

图 2-8



图为某企业内部网络示意图。下联用户通过 Gi0/2 口连接到交换机。

为了便于管理员对下联用户使用网络情况信息的掌控，希望通过配置达到以下目的：

- 当交换机下联用户的接口学习到一个新的 MAC 地址或老化掉一个已学习到的地址时，将地址变化信息记录到 MAC 地址通知历史记录表中，供管理员了解最近的 MAC 地址变化信息。
- 同时，交换机能将 MAC 地址变化通知以 SNMP Trap 的方式将通知信息发送给指定的 NMS(网络管理工作站)
- 当交换机下联用户较多时，能尽量避免短时间内产生大量的 MAC 地址变化信息，减轻网络的负担。

【配置方法】

- 打开交换机全局 MAC 地址通知开关，在 Gi0/2 接口上配置 MAC 地址通知功能。
- 配置 NMS 主机地址，使能交换机主动发送 SNMP Trap 通知。交换机到 NMS（网络管理工作站）的路由可达。
- 设置交换机发送 MAC 地址通知的时间间隔为 300 秒（默认时间间隔为 1 秒）。

A

```
Ruijie# configure terminal
Ruijie(config)# mac-address-table notification
Ruijie(config)# interface gigabitEthernet 0/2
Ruijie(config-if-GigabitEthernet 0/2)# snmp trap mac-notification added
Ruijie(config-if-GigabitEthernet 0/2)# snmp trap mac-notification removed
Ruijie(config-if-GigabitEthernet 0/2)# exit
```

```
Ruijie(config)# snmp-server host 192.168.1.10 traps version 2c comefrom2
Ruijie(config)# snmp-server enable traps
Ruijie(config)# mac-address-table notification interval 300
```

- 【检验方法】
- 查看 MAC 地址通知功能的全局配置信息。
 - 查看接口的 MAC 地址变化通知的使能状况。
 - 查看接口 MAC 地址表，并使用使用 clear mac-address-table dynamic 命令模拟动态地址的老化。
 - 查看 MAC 地址通知功能的全局配置信息。
 - 查看 MAC 地址变化通知信息的历史记录表。

```
A
Ruijie# show mac-address-table notification
MAC Notification Feature : Enabled
Interval(Sec): 300
Maximum History Size : 50
Current History Size : 0
Ruijie# show mac-address-table notification interface GigabitEthernet 0/2
Interface          MAC Added Trap    MAC Removed Trap
-----
GigabitEthernet 0/2  Enabled          Enabled
Ruijie# show mac-address-table interface GigabitEthernet 0/2
Vlan      MAC Address      Type      Interface
-----
1         00d0.3232.0001  DYNAMIC  GigabitEthernet 0/2
Ruijie# show mac-address-table notification
MAC Notification Feature : Enabled
Interval(Sec): 300
Maximum History Size : 50
Current History Size : 1
Ruijie# show mac-address-table notification history
History Index : 0
Entry Timestamp: 221683
MAC Changed Message :
Operation:DEL Vlan:1 MAC Addr: 00d0.3232.0003 GigabitEthernet 0/2
```

常见错误

无。

2.5.5 配置MAC地址漂移打印Syslog功能

配置效果

- 当同一个 VLAN 中，两个端口上同一个 MAC 地址在短时间内发生漂移（即之前在一个端口上学习到，然后在另一个端口学习到），则打印一条 Syslog 告警，。

注意事项

- 无。

配置方法

配置 MAC 地址漂移打印 Syslog 告警

- 可选配置。
- 如果需要在同一 VLAN 内检测到同一个 MAC 地址在不同端口上短时间内漂移时，打印 Syslog 告警，则应配置此项。
- 交换机设备上配置。

【命令格式】 **mac-address-table flapping-logging**

【参数说明】

【缺省配置】 缺省该功能关闭。

【命令模式】 全局模式

【使用指导】

检验方法

- 可以通过 show run 查看配置结果。

配置举例

配置告警间隔

【配置方法】 ● 配置 MAC 地址漂移打印 Syslog 功能开启

```
Ruijie# configure terminal
Ruijie(config)# mac-address-table flapping-logging
```

【检验方法】 在交换机上通过 **show running** 命令可以查看到相应配置。

常见错误

无。

2.6 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除动态地址表项。	clear mac-address-table dynamic [address <i>mac-address</i>] [interface <i>interface-id</i>] [vlan <i>vlan-id</i>]

查看运行情况

作用	命令
查看 MAC 地址表。	show mac-address-table { dynamic static filter } [address <i>mac-address</i>] [interface <i>interface-id</i>] [vlan <i>vlan-id</i>]
查看动态地址老化时间	show mac-address-table aging-time
查看动态地址个数限制情况	show mac-address-table max-dynamic-mac-count
查看地址变化通知配置及历史记录表	show mac-address-table notification [interface [<i>interface-id</i>] history]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 MAC 运行情况的调试开关。	debug bridge mac

3 Aggregate Port

3.1 概述

Aggregate Port (简称 AP)是将多个物理链接捆绑在一起形成一个逻辑链接,可以用于扩展链路带宽,提供更高的连接可靠性。AP 支持流量平衡,可以把流量均匀地分配给各成员链路。AP 还实现了链路备份,当 AP 中的一条成员链路断开时,系统会将该成员链路的流量自动地分配到 AP 中的其它有效成员链路上。AP 中一条成员链路收到的广播或者多播报文,将不会被转发到其它成员链路上。

比如两台设备之间,单个端口相连最多为 1000M(假定两台设备的端口都为 1000M),当该链路上承载的业务流量超过 1000M 时,超过的部分就会被丢弃,而端口聚合将可以解决这一问题。例如,使用若干根网线连接这两台设备,再将这若干个端口进行聚合绑定,这样这些端口就逻辑捆绑形成了 1000M * n 的最大流量。

又比如,如果两台设备是通过单个网线相连接,当这两个端口之间出现链路断开时,这条线路上承载的业务就会断掉,而如果将多个互连的端口进行聚合绑定,只要有一条链路没有出现链路断开,那么在那些端口上承载的业务就不会断掉。

i 下文仅介绍 AP 的相关内容。

协议规范

- IEEE 802.3ad

3.2 典型应用

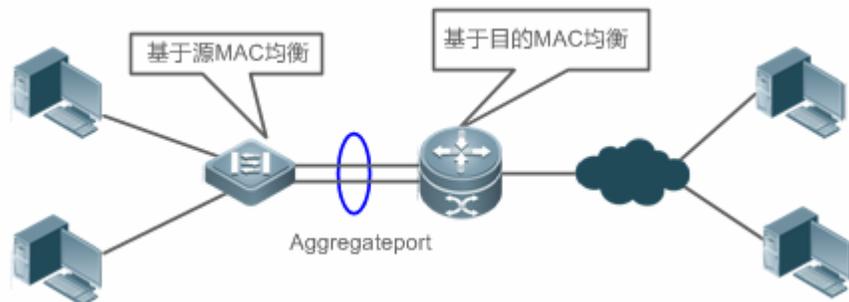
典型应用	场景描述
AP链路聚合及流量平衡	汇聚和核心设备之间通常存在大量报文流,需要更大的端口带宽来支撑,这时候就可以把设备上多条物理链路聚合成一条逻辑链路,增大链路带宽,并通过配置适当的流量平衡算法,使聚合口上的报文尽可能平衡到每一条物理链路,以提高带宽利用率。

3.2.1 AP链路聚合及流量平衡

应用场景

在下图中,左边的交换机设备通过 AP 与右边的路由器进行通讯,所有内网中的设备(如图中的左边 2 台 PC 机)以路由器为网关,所有外网(如图中的右边 2 台 PC 机)经路由器发出的报文的源 MAC 都是网关的 MAC 地址,为了让路由器与其他主机之间的通讯流量能由其他链路来分担,应设置为根据目的 MAC 地址进行流量平衡;而在交换机处,则需要设置为根据源 MAC 地址进行流量平衡。

图 3-1 AP 链路聚合及流量平衡示意图



【注释】 -

功能部属

- 把交换机与路由器之间的直连端口配置成一个静态 AP 或 LACP
- 在交换机上设置基于源 MAC 的流量平衡算法
- 在路由器上设置基于目的 MAC 的流量平衡算法

3.3 功能详解

基本概念

▾ 静态 AP

静态 AP 模式是一种利用手工配置模式直接将物理端口加入到 AP 聚合组中，在物理端口的链路状态和协议状态准备好的情况下，就能进行数据报文转发的一种聚合模式。

静态 AP 模式下的 AP 接口，称为静态 AP 口，对应的成员口称为静态 AP 成员口。

▾ LACP

LACP 是一个关于动态链路聚合的协议，它通过协议报文 LACPDU(Link Aggregation Control Protocol Data Unit，链路聚合控制协议数据单元)和相连的设备交互信息。

LACP 模式下的 AP 接口，称为 LACP AP 口，对应的成员口称为 LACP AP 成员口。

▾ AP 成员端口模式

AP 成员端口有 3 种聚合模式：主动(Active)模式、被动模式(Passive)和静态模式。

其中主动模式的端口会主动发起 LACP 报文协商，被动模式的端口则只会对收到的 LACP 报文做应答，静态模式不会发出 LACP 报文进行协商，这种模式只会静态 AP 模式下生效。各个聚合模式的相邻端口聚合模式要求如下：

端口模式	相邻端口聚合模式要求
主动模式	主动模式或者被动模式
被动模式	主动模式

静态模式	静态模式
------	------

AP 成员端口状态

静态 AP 成员端口的状态主要有以下两种：

- 当成员端口的链路处于 Down 状态，端口不能转发任何数据报文，显示为“Down”状态；
- 当成员端口链路处于 Up 状态，且链路协议准备好后，端口可以参与转发数据报文，显示为“Up”状态。

LACP 成员端口可能处于以下三种状态：

- 当端口的链路处于 Down 状态，端口不能转发任何数据报文，显示为“down”状态；
- 端口链路处于 Up 状态，并经过 LACP 协商后，端口被置于聚合状态(端口被作为一个聚合组的一个成员，参与聚合组的数据报文转发)，显示为“bndl”状态；
- 当端口链路处于 UP 状态，但是由于对端没有启用 LACP，或者因为端口属性和主端口不一致等一些因素导致经过报文协商端口被置于挂起状态（处于挂起状态的端口不参与数据报文转发），显示为“susp”状态。

i 只有全双工的端口才能进行 LACP 聚合。

i 成员端口的速率、流控、介质类型以及成员端口的二、三层属性必须一致才能进行 LACP 聚合绑定。

i LACP 成员端口聚合后修改端口的上述属性将导致同聚合组内的其他端口也无法进行 LACP 聚合绑定。

! 已经启用禁止成员口加入或者退出 AP 功能的端口不能将端口加入静态 AP 或者 LACP AP，或者从静态 AP 或者 LACP AP 中退出。

AP 容量模式

由于系统中总的成员口数量有限，系统总支持成员口数 = 系统支持的最大 AP 口数量 * 单个 AP 口支持最大成员口数。因此当希望系统中最大 AP 口数量大一点，那么单个 AP 口下的最大成员口数就会小一点，反过来单个 AP 最大成员口数大一点，全局最大 AP 数量就小一点。某些特定的场景有这种需求，这就引出了 AP 容量模式的概念，在某些产品设备上支持 AP 容量模式可配置，比如系统支持 16384 个成员口，那么容量模式可以选择 1024*16、512*32 等等（最大 AP 数*单个 AP 下最大成员口数）。

LACP 的系统 ID

每台设备仅能配置一个 LACP 聚合系统。聚合系统有一个系统 ID 来标示这个系统的优劣，同时存在一个系统优先级，这是一个可配置的数值。系统 ID 由 LACP 的系统优先级和设备 MAC 地址组成。系统优先级越小，系统 ID 的优先级越高；在系统优先级相同的情况下，比较设备的 MAC 地址，设备 MAC 地址越小，系统 ID 的优先级越高。系统 ID 优先级较高的系统决定端口状态，低优先级系统的端口状态随高优先级系统的端口状态变化而变化。

LACP 的端口 ID

每个端口有独立的 LACP 端口优先级，这是一个可配置的数值。端口 ID 由 LACP 的端口优先级和端口号组成。端口优先级数值越小，端口 ID 的优先级越高；在端口优先级相同的情况下，端口号越小，端口 ID 的优先级越高。

LACP 的主端口

当有动态成员处于 Up 状态时，LACP 会根据端口的速率，双工速率等关系，并综合聚合组内端口 ID 优先级、聚合组内已经 Up 的成员口的绑定状态等信息，选择其中的一个成员口端口作为主端口。只有和主端口属性相同的端口才能处于聚合状态，

参与聚合组的数据转发。当端口的属性变化时，LACP 会重新选择主端口；当新的主端口不处于聚合状态时，LACP 会把同一个聚合组内的成员解聚合，重新聚合。

AP 优选口

通常用在 AP 口同服务器双系统对接的场景下。通过指定 AP 的某个成员为优选口，使得特定报文（管理 vlan 的报文）经优选口转发至服务器，而不会被流量均衡到其他成员口，保障了同服务器间的正常通信。

 请将连接服务器管理网卡的端口设置为 AP 优选口。

在某些 Linux 服务器上存在双系统，比如 HP 服务器存在主系统和远程管理系统，主系统也就是 Linux 系统，远程管理系统，即 ILO(Integrated Light-Out)，提供硬件级的远程管理功能。ILO 即使在主系统重启的过程中，依然能对服务器进行远程管理。主系统双网卡绑定成聚合口，用于主系统业务处理；管理系统使用其中一张网卡做远程管理，也就是两个系统复用了一张网卡，但业务由不同 VLAN 隔离开，管理系统所用的 VLAN 我们称为管理 VLAN。这样对于交换机设备来说，和服务器双网卡对接的也会是一个聚合口，这个聚合口上管理 VLAN 的报文就需要往同服务器网卡相连的那个成员口发出，这样才能保证与服务器上远程管理系统正常通信。为此，可以通过配置 AP 优选口来指定管理 VLAN 报文的转发。

 若服务器双网卡使用 LACP 绑定，当主系统重启时，LACP 协议还未运行，设备上 LACP 协议处于协商失败的状态，聚合口会处于 Down 状态，这时 AP 优选口会自动降为静态成员口，直接绑定到聚合口，以便聚合口能够服务器远程管理系统保持通信，直到 Linux 系统重启完成，LACP 协议正常运行之后，设备上 AP 优选口重新启用 LACP 协议进行协商。

功能特性

功能特性	作用
链路聚合	将物理链路通过静态或动态的方式聚合，以达到扩展带宽、链路备份的作用。
流量平衡	通过不同的流量均衡模式，可以灵活地对聚合组内流量进行负载均衡。

3.3.1 链路聚合

工作原理

AP 链路聚合方式分为两种，一种是通过手工配置，即静态 AP；另一种是通过 LACP 协议动态聚合。

- 静态 AP

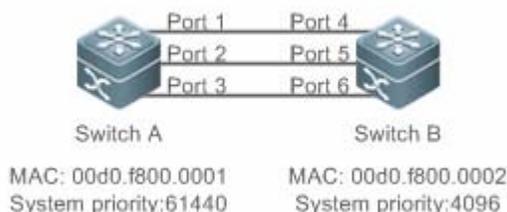
静态 AP 实现简单，用户只要将指定的物理端口通过配置命令加入到同一个聚合组 AP 中，就可以实现多条物理链路的聚合。成员端口一旦加入聚合组后，即可参与 AP 聚合组的数据收发功能，并参与聚合组的流量均衡。

- 动态 AP(LACP)

如果端口启用 LACP 协议，端口会发送 LACPDU 来通告自己的系统优先级、系统 MAC、端口的优先级、端口号和操作 key 等。相连设备收到对端的 LACP 报文后，根据报文中的系统 ID 比较两端的系统优先级。在系统 ID 优先级较高的一端，将按照端口 ID 优先级从高到低的顺序，设置聚合组内端口处于聚合状态，并发出更新后的 LACP 报文，对端设备收到报文后，也会把相应的端口设置成聚合状态，从而使双方在端口退出或者加入聚合组上达到一致。只有双方的端口都完成动态聚合绑定操作后，该物理链路才能进行数据报文的转发。

LACP 成员口链路绑定之后，还会进行周期性的 LACP 报文交互，在一段时间没有收到 LACP 报文时，就认为收包超时，成员口链路解除绑定，端口重新处于不可转发状态。这里的超时时间有两种模式：长超时模式和短超时模式。在长超时模式下，端口间隔 30 秒发送一个报文，若 90 秒没有收到对端报文，就处于收包超时；在短超时模式下，端口间隔 1 秒发送一个报文，若 3 秒钟没有收到对端报文，就处于收包超时。

图 3-2 LACP 协商



如上图所示，交换机 A 和交换机 B 通过 3 个端口连接在一起。设置交换机 A 的系统优先级为 61440，设置交换机 B 的系统优先级为 4096。在交换机 A 和 B 的 3 个直连端口上打开 LACP 链路聚合，设置 3 个端口的聚合模式为主动模式，设置 3 个端口的端口优先级为默认优先级 32768。

在收到对端的 LACP 报文后，交换机 B 发现自己的系统 ID 优先级比较高(交换机 B 的系统优先级比交换机 A 高)，于是按照端口 ID 优先级的顺序(端口优先级相同的情况下，按照端口号从小到大的顺序)设置端口 4、5、6 处于聚合状态。交换机 A 收到交换机 B 更新后的 LACP 报文后，发现对端的系统 ID 优先级比较高，并且把端口设置成聚合状态了，也把端口 1、2、3 设置成聚合状态了。

3.3.2 流量平衡

工作原理

AP 可以根据报文的源 MAC 地址、目的 MAC 地址、源 IP 地址、目的 IP 地址、L4 层源端口、L4 层目的端口号等报文特征信息，进行一种或几种组合模式算法对报文流进行区分，将属于同一报文流从同一条成员链路通过，不同的报文流则平均分配到各个成员链路中。例如，采用源 MAC 地址流量平衡模式，会根据报文的源 MAC 地址将报文分配到 AP 的各个成员链路上。不同源 MAC 的报文，根据源 MAC 地址在各成员链路间平衡分配；相同源 MAC 的报文，固定从同一个成员链路转发。

目前可支持的 AP 流量平衡模式如下：

- 源 MAC 或目的 MAC 地址
- 源 MAC+目的 MAC 地址
- 源 IP 地址或目的 IP 地址
- 源 IP 地址+目的 IP 地址
- L4 层源端口或 L4 层目的端口
- L4 层源端口+L4 层目的端口
- 源 IP+L4 层源端口
- 源 IP+L4 层目的端口
- 目的 IP+L4 层源端口

- 目的 IP+L4 层目的端口
- 源 IP+L4 层源端口+L4 层目的端口
- 目的 IP+L4 层源端口+L4 层目的端口
- 源 IP+目的 IP+L4 层源端口
- 源 IP+目的 IP+L4 层目的端口
- 源 IP+目的 IP+L4 层源端口+L4 层目的端口
- 输入报文的面板端口
- MPLS 报文的 Label 均衡
- 聚合链路成员口轮询均衡
- 增强模式

i 根据报文的 IP 地址或端口号进行流量平衡的模式仅适用于三层报文，如果在此流量平衡模式下收到二层报文，则自动根据设备的默认方式进行流量平衡。

i 各种流量平衡模式都是利用流量算法（哈希算法）、根据该模式采用的输入参数（源 MAC、目的 MAC、源 MAC+目的 MAC、源 IP、目的 IP、源 IP+目的 IP、源 ip+目的 ip 和 L4 端口号等）计算特定报文应选择的成员链路，来实现流量均衡。这种算法能够保证输入参数不同的报文被大致均衡地分配给各成员链路，但并不意味着，输入参数不同的报文就一定选择不同的成员链路。比如，对 IP 模式而言，两个具有不同源 IP+目的 IP 地址的报文，通过计算可能分配到同一个 AP 的成员链路。

i 不同产品，流量均衡支持度可能存在差异。

▾ 增强模式流量均衡

增强模式允许用户将不同报文类型的多个字段进行组合以达到流量均衡，包含 L2 报文对应的字段 src-mac、dst-mac、l2-protocol、vlan、src-port、dst-port，IPv4 报文的对应的字段 src-ip、dst-ip、protocol、l4-src-port、l4-dst-port、vlan、src-port、dst-port、l2-etype、src-mac、dst-mac。

在增强模式下，设备先判断发送报文的类型，然后根据指定报文的字段进行流量均衡。比如，源 IP 变化的 IPv4 报文要从 AP 口输出，那么 AP 会根据用户指定的 IPv4 报文字段 src-ip 进行流量均衡。

i 以上所有流量平衡模式都适用于二层 AP 和三层 AP，用户应根据不同的网络环境设置合适的流量分配方式，以便能把流量较均匀地分配到各个链路上，充分利用网络的带宽。

i 增强模式中 L2 均衡包含 src-mac、dst-mac、vlan，IPv4 均衡包含 src-ip，如果输入的报文是源 MAC 变化的 IPv4 报文，那么均衡算法不生效，因为该模式先检查报文类型为 IPv4，那么只会根据 IPv4 对应设置的字段 src-ip 均衡。

i 增强模式中 MPLS 均衡算法只对 MPLS L3 VPN 报文生效，对 MPLS L2 VPN 报文无效。

3.4 产品说明



- 缺省情况下，对于 S7508 产品，每个 AP 口最多包含的成员口数量为 8 个，设备支持最大 AP 口数量为 256 个。
- 对于 S7508 产品，AP 的容量模式支持可配置，支持配置 256*8、128*16、64*32 和 32*64 模式。上述模式对应地支持每个 AP 口最大成员口数量为 8、16、32、64，最大 AP 口数量为 256、128、64、32 个。



- 对于 S7508 产品,当采用基于源 MAC 地址、目的 MAC 地址或源 MAC 地址+目的 MAC 地址流量平衡这三种模式时,设备会默认将单播报文的以太网类型字段和 VLAN 字段也作为均衡因子。
- 对于 S7508 产品,流量均衡采用非增强模式,IGMP SNOOPING 或者组播路由打开后的组播报文均衡的关键字为 src-ip, dst-ip 或 src-ip+dst-ip,其他多播,未知名单播,广播的均衡的关键字为 src-mac、dst-mac 或 src-mac+dst-mac。比如三层报文(未知单播、组播、广播)走二层无法根据 src-ip,dst-ip 负载均衡,而采用增强模式可以处理,因为增强模式是根据报文类型做流量均衡。
- Src-dst-ip-l4port,该模式下, S7508 产品对于 L4port 的变化只对单播报文有效。
- S7508 产品增强型均衡模板可配置的域如下:

L2 模板: src-mac dst-mac vlan l2-protocol src-port

- IPV4 模板:src-ip dst-ip protocol vlan l4-src-port l4-dst-port src-port

3.5 配置详解

配置项	配置建议 & 相关命令	
配置静态AP	⚠️ 必须配置。用于手工设置链路聚合。	
	interface aggregateport	创建一个以太网 AP 口。
	port-group	配置以太网静态 AP 成员口。
配置LACP	⚠️ 必须配置。用于动态设置链路聚合。	
	port-group mode	配置 LACP 成员口。
	lacp system-priority	配置 LACP 系统的优先级。
	lacp port-priority	配置端口的优先级。
	lacp short-timeout	配置端口为短超时模式
配置AP的LinkTrap功能	⚠️ 可选配置。用于打开接口的 LinkTrap 通告功能。	
	snmp trap link-status	打开发送 AP 口 LinkTrap 通告功能。
	aggregateport member linktrap	打开发送 AP 成员口 LinkTrap 通告功能。
配置流量平衡模式	⚠️ 可选配置。用于指定当前聚合链路的流量均衡模式。	
	aggregateport load-balance	设置 AP 的全局或单个 AP 口流量平衡算法。
	⚠️ 可选配置。用于设置增强模式的模板。	
	load-balance-profile	创建增强模式模板。
	l2 field	配置二层报文的负载均衡方式。
	ipv4 field	配置 Ipv4 报文的负载均衡方式。
配置AP的容量模式	⚠️ 可选配置。用于指定当前系统的 AP 容量模式。	
	aggregateport capacity mode	设置全局 AP 容量模式。

3.5.1 配置静态AP

配置效果

- 通过手工添加 AP 口成员，将多个物理端口绑定，以实现链路聚合。
- 聚合后的逻辑链路带宽是成员链路带宽的总和。
- 当 AP 中的一条成员链路断开时，系统会将该成员链路的流量自动地分配到 AP 中的其它有效成员链路上。

注意事项

- 只有物理端口才允许加入 AP 口。
 - 不同介质类型或者不同端口类型的接口不允许加入同一个 AP 口。
 - 二层端口只能加入二层 AP，三层端口只能加入三层 AP；包含成员口的 AP 口不允许改变二层/三层属性。
 - 一个端口加入 AP，端口的属性将被 AP 的属性所取代。
 - 一个端口从 AP 中删除，则端口的属性将恢复为加入 AP 前的属性。
- ❗ 当一个端口加入 AP 后，该端口的属性取代为 AP 口的属性，所以一般情况下不允许在 AP 成员口上进行配置，或者将配置单独生效到 AP 成员口上。但一些少数的命令或者功能，如 `shutdown` 和 `no shutdown` 配置命令等，这些仍然可以支持在 AP 成员口上配置，且配置能生效。所以用户在使用 AP 成员口的时候，需要根据具体的功能要求来确定是否支持单独在 AP 成员口上生效，并进行正确配置。

配置方法

创建以太网 AP 口

- 必须配置。
- 在支持 AP 功能的设备上配置。以太网口使用聚合功能时需要创建对应的以太网 AP 口。
 - 【命令格式】 `interface aggregateport ap-number`
 - 【参数说明】 `ap-number`：AP 接口编号
 - 【缺省配置】 缺省情况下，AP 口未被创建。
 - 【命令模式】 全局配置模式
 - 【使用指导】 在全局配置模式下，用户可以通过 `interfaces aggregateport` 配置命令创建一个以太网 AP 口。用户可以在全局配置模式下，通过 `no interfaces aggregateport ap-number` 删除指定的以太网 AP 口。

- ❗ 用户可以通过在指定以太网端口的接口模式下，执行 `port-group` 命令将物理端口加入一个静态 AP；如果该 AP 不存在，则同时自动创建这个 AP 口。
- ❗ 用户也可以通过在指定物理端口的接口模式下，执行 `port-group mode` 命令将物理端口加入一个 LACP AP；如果该 AP 不存在，则同时自动创建这个 AP 口。
- ❗ 配置 AP 功能时，需要在链路两端的设备上都配置，且需要配置相同的 AP 类型(静态 AP 或者 LACP)。

配置以太网静态 AP 成员口

- 必须配置。

- 在支持 AP 功能的设备上配置。使用静态聚合功能时需要配置对应的静态 AP 成员口。

【命令格式】 **port-group ap-number**

【参数说明】 **port-group ap-number**: AP 接口编号

【缺省配置】 以太网端口不属于任何静态 AP 的成员口

【命令模式】 以太网接口配置模式

【使用指导】 在接口模式下，用户可以通过 **port-group** 配置命令向 AP 口中添加成员口。在接口配置模式下使用 **no port-group** 命令将此成员口退出 AP。

- ❗ 为保证链路聚合功能正常，在链路两端的设备上需要对称配置静态 AP 成员口。
- ❗ 将普通端口加入某个 AP 口后，当该端口再次从 AP 口退出时，普通端口上的原先相关的配置可能会恢复为缺省的配置。不同功能对 AP 口的成员的原有配置的处理方式有所不同，因此建议在端口从 AP 口退出后，应查看并确认端口的配置。
- ❗ AP 成员端口从 AP 口退出变成普通端口后，该端口会被 **shutdown** 以防止出现环路等问题，用户需要在确认拓扑无异常之后再接口模式下执行 **no shutdown** 命令重新使能该接口。

📌 二层 AP 与三层 AP 的转化

- 为可选配置。
- 如果需要启用 AP 口的三层路由等功能，比如需要在 AP 口上配置 IP 地址，或者配置静态路由表项等，需要先将二层 AP 口转化为三层 AP 口，再在三层 AP 口上启用路由等功能。
- 该功能可在三层交换机或者无线 AC 等支持二、三层功能和 AP 功能的设备上配置。

【命令格式】 **no switchport**

【参数说明】 -

【缺省配置】 在支持二、三层功能和接口二、三层转换功能的设备上，AP 口缺省为二层口。

【命令模式】 AP 接口配置模式

【使用指导】 L3 AP 是三层设备才支持的功能，所有二层设备均不支持。

- ❗ 对于三层设备，如果该设备不支持二层功能，AP 口被创建时，则是一个三层 AP 口，否则 AP 口被创建时是一个二层 AP 口。

📌 创建以太网 AP 子接口

- 可选配置。
- 如果设备支持配置子接口的功能，则也同时支持在 AP 口上通过 **interface aggregateport sub-ap-number** 创建相应的子接口。
- 该功能可在三层交换机等支持三层功能和 AP 功能的设备上配置。

【命令格式】 **interface aggregateport sub-ap-number**

【参数说明】 **sub-ap-number**: AP 子接口编号

【缺省配置】 缺省 AP 口没有任何子接口。

【命令模式】 AP 接口配置模式

【使用指导】 在支持二、三层功能和接口二、三层转换功能的设备上，AP 口主接口需要先转换为三层口，才允许该 AP 口主接口创建子接口。

检验方法

- 通过 **show running** 命令查看相应的配置。
- 通过 **show aggregateport summary** 命令查看 AP 口配置情况。

【命令格式】 **show aggregateport aggregate-port-number [load-balance | summary]**

【参数说明】 **aggregate-port-number** : AP 接口号

load-balance : 显示 AP 的流量平衡算法

summary : 显示 AP 中的每条链路的摘要信息

【命令模式】 各模式均可执行

【使用指导】 如果没有指定 AP 接口号, 则所有 AP 的信息将被显示出来

【命令展示】

```
Ruijie# show aggregateport 1 summary

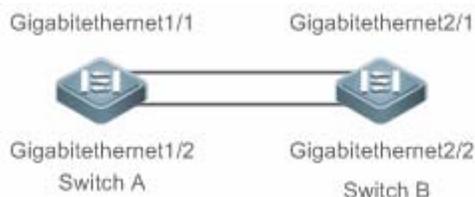
AggregatePort  MaxPorts      SwitchPort Mode    Load balance      Ports
-----
-----
Ag1             8             Enabled  ACCESS  dst-mac             Gi0/2
```

配置举例

配置以太网静态 AP

【网络环境】

图 3-3



【配置方法】

- 将 SwitchA 上的端口 GigabitEthernet 1/1 和 GigabitEthernet 1/2 加入到静态 AP 3 中。
- 将 SwitchB 上的端口 GigabitEthernet 2/1 和 GigabitEthernet 2/2 加入到静态 AP 3 中。

SwitchA

```
SwitchA# configure terminal
SwitchA(config)# interface range GigabitEthernet 1/1-2
SwitchA(config-if-range)# port-group 3
```

SwitchB

```
SwitchB# configure terminal
SwitchB(config)# interface range GigabitEthernet 2/1-2
SwitchB(config-if-range)# port-group 3
```

【检验方法】

- 通过 **show aggregateport summary** 查看 AP 口和成员口的对应关系是否正确。

SwitchA

```
SwitchA# show aggregateport summary
AggregatePort MaxPorts SwitchPort Mode Ports
-----
-----
Ag3           8           Enabled  ACCESS  Gi1/1, Gi1/2
```

SwitchB

```
SwitchB# show aggregateport summary
```

AggregatePort	MaxPorts	SwitchPort	Mode	Ports
Ag3	8	Enabled	ACCESS	Gi2/1, Gi2/2

常见错误

3.5.2 配置LACP

配置效果

- 相连设备根据 LACP 自协商，动态聚合链路。
- 聚合后的逻辑链路带宽是成员链路带宽的总和。
- 当 AP 中的一条成员链路断开时，系统会将该成员链路的流量自动地分配到 AP 中的其它有效成员链路上。
- 长超时模式时，链路故障后 90 秒才能感知到；配置短超时模式时，3 秒钟就能感知到。

注意事项

- 将普通端口加入某个 LACP AP 口后，当该端口再次从 LACP AP 口退出时，普通端口上的原先相关的配置可能会恢复为缺省的配置。不同功能对 LACP AP 口的成员的原有配置的处理方式有所不同，因此建议在端口从 LACP AP 口退出后，应查看并确认端口的配置。
- 改变 LACP 的系统优先级可能引起 LACP 的成员端口出现解聚合再聚合现象。
- 改变 LACP 成员口的端口优先级可能引起该 LACP 成员口对应的聚合组所有端口出现解聚合再聚合现象。

配置方法

配置 LACP 成员口

- 必须配置。
- 将指定的物理端口配置为 LACP 成员口。在支持 LACP 功能的设备上配置。使用 LACP 功能时需要配置对应的 LACP 成员口。

【命令格式】 **port-group** *key-number* **mode** { **active** | **passive** }

【参数说明】 *Key-number* :为聚合组的管理 *key* , *Key-number* 取值范围根据不同产品支持的聚合组数量不同而变，这个 *Key-number* 值就是对应的 LACP AP 口的端口号。

active:表示端口以主动模式加入动态聚合组

passive:模式表示端口以被动模式加入聚合组

【缺省配置】 物理端口不属于任何 LACP 的成员口

【命令模式】 物理接口配置模式

【使用指导】 在接口模式下，用户可以通过下面的配置命令向 LACP AP 口中添加成员口。

i 为保证 LACP 功能正常，在链路两端的设备上需要对称配置 LACP 成员口。

配置 LACP 的系统优先级

- 为可选配置。
- 在需要调整该设备系统 ID 优先级时进行配置，配置值越小，系统 ID 优先级越高，系统 ID 优先级高的设备优先选择聚合端口。
- 可在支持 LACP 功能的设备上配置该功能。

【命令格式】 **lacp system-priority system-priority**

【参数说明】 *system-priority* : LACP 系统的优先级，可选范围为 0-65535，默认优先级为 32768。

【缺省配置】 LACP 的系统优先级为 32768

【命令模式】 全局配置模式

【使用指导】 在全局模式下，用户可以通过下面的配置命令配置 LACP 的系统优先级。一台设备的所有的动态链路组只能有一个 LACP 系统优先级，修改这个值会影响到交换机上的所有聚合组。在接口配置模式下使用 **no lacp system-priority** 命令将 LACP 的系统优先级恢复到缺省值。

配置 LACP 成员口的端口优先级

- 可选配置。
- 在需要调整端口 ID 优先级时进行配置，配置值越小，端口 ID 优先级越高，端口 ID 优先级高的端口会被优选为主端口。
- 可在支持 LACP 功能的设备上配置该功能。

【命令格式】 **lacp port-priority port-priority**

【参数说明】 *port-priority* : 端口的优先级，可选范围为 0-65535，默认优先级为 32768。

【缺省配置】 LACP 成员口的端口优先级为 32768

【命令模式】 物理接口配置模式

【使用指导】 在全局模式下，用户可以通过下面的配置命令配置 LACP 的系统优先级。在接口配置模式下使用 **no lacp port-priority** 命令将 LACP 的系统优先级恢复到缺省值。

配置 LACP 成员口的超时模式

- 可选配置。
- 在需要更实时感知链路故障的场景下，需要配置成短超时模式。配置短超时模式时，端口 3 秒收包超时，长超时模式，端口 90 秒收包超时。
- 可在支持 LACP 功能的设备上配置该功能，比如交换机产品等。

【命令格式】 **lacp short-timeout**

【参数说明】 -

【缺省配置】 LACP 成员口的端口超时模式为长超时

【命令模式】 接口配置模式

【使用指导】 仅在物理口上支持。

在接口配置模式下使用 **no lacp short-timeout** 命令将 LACP 超时模式恢复为缺省值。

检验方法

- 通过 **show running** 命令查看相应的配置。
- 通过 **show lacp summary** 命令查看 LACP 链路状态。

【命令格式】 **show lacp summary** [*key-number*]

【参数说明】 *key-name* : 指定的 LACP AP 接口号

【命令模式】 各模式均可执行

【使用指导】 如果没有指定 *key-number* , 则所有 LACP AP 的链路聚合状态信息将被显示出来。

【命令展示】

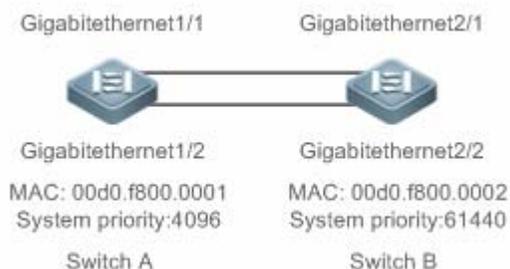
```
Ruijie(config)# show lacp summary 3
System Id:32768, 00d0.f8fb.0002
Flags: S - Device is requesting Slow LACPDUs
F - Device is requesting Fast LACPDUs.
A - Device is in active mode.      P - Device is in passive mode.
Aggregate port 3:
Local information:
LACP port      Oper   Port   Port
Port   Flags   State  Priority   Key    Number  State
-----
Gi0/1   SA     bndl   4096      0x3    0x1     0x3d
Gi0/2   SA     bndl   4096      0x3    0x2     0x3d
Gi0/3   SA     bndl   4096      0x3    0x3     0x3d
Partner information:
          LACP port
Port   Flags   Priority  Dev ID   Oper   Port   Port
-----
Gi0/1   SA     61440    00d0.f800.0001  0x3    0x1     0x3d
Gi0/2   SA     61440    00d0.f800.0001  0x3    0x2     0x3d
Gi0/3   SA     61440    00d0.f800.0001  0x3    0x3     0x3d
```

配置举例

配置 LACP

【网络环境】

图 3-4



【配置方法】

- 在 SwitchA 上设置 LACP 系统优先级为 4096。
- 在 SwitchA 上的端口 GigabitEthernet1/1 和 GigabitEthernet1/2 上启用动态链路聚合协议，将其加入到 LACP 3 中。
- 在 SwitchB 上设置 LACP 系统优先级为 61440。
- 在 SwitchB 上的端口 GigabitEthernet2/1 和 GigabitEthernet2/2 启用动态链路聚合协议，将其加入到 LACP 3 中。

SwitchA

```
SwitchA# configure terminal
SwitchA(config)# lacp system-priority 4096
SwitchA(config)# interface range GigabitEthernet 1/1-2
SwitchA(config-if-range)# port-group 3 mode active
SwitchA(config-if-range)# end
```

SwitchB

```
SwitchB# configure terminal
SwitchB(config)# lacp system-priority 61440
SwitchB(config)# interface range GigabitEthernet 2/1-2
SwitchB(config-if-range)# port-group 3 mode active
SwitchB(config-if-range)# end
```

【检验方法】

- 通过 **show lacp summary 3** 查看 LACP 和成员口的对应关系是否正确。

SwitchA

```
SwitchA# show LACP summary 3
System Id:32768, 00d0.f8fb.0001
Flags: S - Device is requesting Slow LACPDUs
      F - Device is requesting Fast LACPDUs.
      A - Device is in active mode.          P - Device is in passive mode.
Aggregate port 3:
Local information:
LACP port      Oper   Port      Port
Port          Flags  State     Priority  Key     Number   State
-----
Gi1/1         SA     bnd1     32768    0x3     0x1     0x3d
Gi1/2         SA     bnd1     32768    0x3     0x2     0x3d
Partner information:
LACP port      Oper   Port      Port
```

SwitchB

```

Port      Flags    Priority  Dev ID      Key    Number  State
-----
Gi2/1     SA       32768    00d0.f800.0002  0x3    0x1     0x3d
Gi2/2     SA       32768    00d0.f800.0002  0x3    0x2     0x3d

SwitchB# show LACP summary 3
System Id:32768, 00d0.f8fb.0002
Flags: S - Device is requesting Slow LACPDUs
      F - Device is requesting Fast LACPDUs.
      A - Device is in active mode.          P - Device is in passive mode.
Aggregate port 3:
Local information:
LACP port      Oper  Port      Port
Port      Flags    State    Priority    Key    Number  State
-----
Gi2/1     SA       bndl     32768      0x3    0x1     0x3d
Gi2/2     SA       bndl     32768      0x3    0x2     0x3d

Partner information:
          LACP port
Port      Flags    Priority  Dev ID      Oper  Port  Port
          Key    Number  State
-----
Gi1/1     SA       32768    00d0.f800.0001  0x3    0x1     0x3d
Gi1/2     SA       32768    00d0.f800.0001  0x3    0x2     0x3d

```

常见错误

-

3.5.3 配置AP的LinkTrap功能**配置效果**

当聚合链路发生变化时，系统会发出相应的 LinkTrap 通告。

注意事项

-

配置方法**配置 AP 口的 LinkTrap**

- 在接口模式下配置。为可选配置。AP 口的 LinkTrap 通告功能默认开启，在此情况下，AP 口的链路状态或者协议状态发生变化时，设备会发出 LinkTrap 通告；当不需要要该 AP 口的 LinkTrap 通告时，配置关闭该功能。
- 可在所有支持 AP 功能的设备上配置该功能。

【命令格式】 **snmp trap link-status**

【参数说明】 -

【缺省配置】 LinkTrap 通告默认开启

【命令模式】 AP 接口配置模式

【使用指导】 在接口模式下，用户可以对指定的 AP 口设置是否发送 LinkTrap 通告功能。当该功能打开，AP 口发生 Link 状态变化时将发出 LinkTrap 通告，反之则不发。缺省情况下，该功能是打开的。用户可以在指定 AP 口的接口模式下，通过配置 **no snmp trap link-status** 命令关闭指定 AP 口的 LinkTrap 通告功能。

AP 成员口不支持在端口模式下打开 LinkTrap 通告功能。需要通过下面的配置，即在全局模式下配置 **aggregateport member linktrap** 命令来打开 AP 成员口的 LinkTrap 通告功能。

配置 AP 成员口的 LinkTrap

- 为可选配置。成员口 LinkTrap 默认关闭，当需要使能成员口的 LinkTrap 通告功能时，配置开启。
- 可在所有支持 AP 功能的设备上配置该功能。

【命令格式】 **aggregateport member linktrap**

【参数说明】 -

【缺省配置】 缺省情况下，AP 成员口的 LinkTrap 通告功能是关闭的。

【命令模式】 全局配置模式

【使用指导】 用户可以在全局配置模式下，通过配置 **aggregateport member linktrap** 命令打开所有 AP 成员口的 LinkTrap 通告功能。默认情况下，AP 成员口不发送 LinkTrap 通告。用户可以在全局配置模式下，通过配置 **no aggregateport member linktrap** 命令关闭所有 AP 成员口的 LinkTrap 通告功能。

检验方法

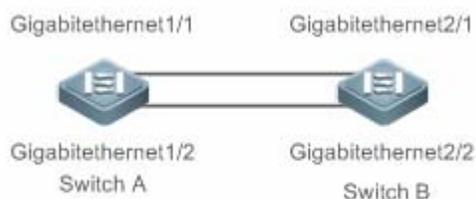
- 通过 **show running** 命令查看相应的配置。
- 打开 LinkTrap 通告的情况下，通过 MIB 软件可以监控到 AP 口或成员口的 LinkTrap 通告。

配置举例

配置 AP 的 LinkTrap 功能

【网络环境】

图 3-5



- 【配置方法】
- 将 SwitchA 上的端口 GigabitEthernet 1/1 和 GigabitEthernet 1/2 加入到静态 AP 3 中。
 - 将 SwitchB 上的端口 GigabitEthernet 2/1 和 GigabitEthernet 2/2 加入到静态 AP 3 中。
 - 在 SwitchA 上配置关闭 AP 3 的 LinkTrap 功能，同时打开成员口的 LinkTrap 功能。
 - 在 SwitchB 上配置关闭 AP 3 的 LinkTrap 功能，同时打开成员口的 LinkTrap 功能。

SwitchA

```
SwitchA# configure terminal
SwitchA(config)# interface range GigabitEthernet 1/1-2
SwitchA(config-if-range)# port-group 3
SwitchA(config-if-range)# exit
SwitchA(config)# aggregateport member linktrap
SwitchA(config)# interface Aggregateport 3
SwitchA(config-if-AggregatePort 3)# no snmp trap link-status
```

SwitchB

```
SwitchB# configure terminal
SwitchB(config)# interface range GigabitEthernet 2/1-2
SwitchB(config-if-range)# port-group 3
SwitchB(config-if-range)# exit
SwitchB(config)# aggregateport member linktrap
SwitchB(config)# interface Aggregateport 3
SwitchB(config-if-AggregatePort 3)# no snmp trap link-status
```

- 【检验方法】
- 通过 **show running** 查看 AP 的流量均衡算法配置是否正确。

SwitchA

```
SwitchA# show run | include AggregatePort 3
Building configuration...
Current configuration: 54 bytes
interface AggregatePort 3
  no snmp trap link-status
SwitchA# show run | include AggregatePort
aggregateport member linktrap
```

SwitchB

```
SwitchB# show run | include AggregatePort 3
Building configuration...
Current configuration: 54 bytes
interface AggregatePort 3
  no snmp trap link-status
SwitchB# show run | include AggregatePort
aggregateport member linktrap
```

常见错误

3.5.4 配置流量平衡模式

配置效果

系统会根据指定的流量平衡算法，对输入报文进行流量分配。同一报文流将固定通过同一条链路输出，不同报文流将平均分配到各个链路。在增强模式下，设备先判断发送报文的类型，然后根据指定报文的字段进行流量均衡。比如，源 IP 变化的 IPv4 报文要从 AP 口输出，那么 AP 会根据用户指定的 IPv4 报文字段 `src-ip` 进行流量均衡。

注意事项

配置方法

设置 AP 的全局流量平衡算法

- 为可选配置，当需要改变 AP 的流量平衡算法以实现更好的流量均衡时，需要配置该功能。
- 可在所有支持 AP 功能的设备上配置该功能。

【命令格式】 `aggregateport load-balance { dst-mac | src-mac | src-dst-mac | dst-ip | src-ip | src-dst-ip | src-dst-ip-l4port | enhanced profile profile-name }`

【参数说明】 `dst-mac`：根据输入报文的目的 MAC 地址进行流量分配。

`src-mac`：根据输入报文的源 MAC 地址进行流量分配。

`src-dst-ip`：根据源 IP 与目的 IP 进行流量分配。

`dst-ip`：根据输入报文的目的 IP 地址进行流量分配。

`src-ip`：根据输入报文的源 IP 地址进行流量分配。

`src-dst-mac`：根据源 MAC 与目的 MAC 进行流量分配。

`src-dst-ip-l4port`：根据源 IP 与目的 IP 和 L4 源端口号与 L4 目的端口号进行流量分配。

`enhanced profile profile-name`：根据增强模式模板 `profile-name` 设置对应的报文类型字段进行流量分配。

【缺省配置】 AP 的流量均衡模式为基于源和目的 MAC(如交换机产品系列)或者基于源和目的 IP(如网关产品系列)的流量均衡方式。

【命令模式】 全局配置模式

【使用指导】 要将 AP 的流量平衡设置恢复到缺省值，可以在全局配置模式下使用 `no aggregateport load-balance` 命令。在某些支持基于指定 AP 口配置流量平衡算法的产品上，上述的流量平衡算法配置命令也可以进入 AP 口的接口模式下进行配置，配置生效后，该 AP 口上就会以新配置的流量平衡算法进行工作。同样的，在这些产品下面，用户可以在 AP 口的接口模式下使用 `no aggregateport load-balance` 命令使该 AP 口下配置的流量平衡算法失效，进而生效为当前设备上生效的 AP 全局流量平衡算法。

i 在支持基于 AP 口配置流量均衡的产品上，`aggregateport load-balance` 还支持在 AP 口接口模式下进行配置。

创建增强模式模板

- 如果选择增强模式均衡流量，必须创建增强模式模板，否则会导致无法将 AP 的流量均衡模式设置为增强型模板模式。其他情况下，该配置为可选配置。
- 可在汇聚或者核心交换机等支持增强型流量均衡功能的设备上配置该功能。

【命令格式】 **load-balance-profile** *profile-name*

【参数说明】 *profile-name*：模板名称。支持最多 31 个字符。

【缺省配置】 系统中没有增强式模板配置

【命令模式】 全局配置模式

【使用指导】 要删除增强模式模板，可以在全局配置模式下使用 **no load-balance-profile** *profile-name* 命令。
在全局模式下使用 **load-balance-profile** *profile-name* 命令创建模板名，创建成功的时候就保存了一份缺省的模板配置。
全局只支持一个模板，使用 **show load-balance-profile** 查看目前的配置。

配置二层报文流量均衡模式

- 为可选配置。当采用增强型模板配置作为 AP 的流量均衡方式时，可以根据网络的流量特征配置适当配置该功能，以实现更优的流量均衡。
- 可在汇聚或者核心交换机等支持增强型流量均衡功能的设备上配置该功能。

```
l2 field { [ src-mac ] [ dst-mac ] [ l2-protocol ] [ vlan ] [ src-port ] }
```

```
{
<
|
}
;
```

【参数说明】 **src-mac**：根据输入二层的报文的源 MAC 地址进行流量分配。

dst-mac：根据输入二层的报文的目的 MAC 地址进行流量分配。

l2-protocol：根据输入二层的报文的二层协议类型进行流量分配。

vlan：根据输入二层的报文的 vlan 值进行流量分配。

src-port：根据输入二层报文的端口进行流量分配。

【缺省配置】 缺省模式下，二层报文负载均衡方式为 **src-mac**、**dst-mac**、**vlan**。

【命令模式】 **profile** 配置模式

【使用指导】 配置指定增强模板中二层报文的负载均衡方式。缺省为 **src-mac**、**dst-mac**、**vlan**。
要将二层报文的流量平衡设置恢复到缺省值，可以在该模式下使用 **no l2 field** 命令。

配置 IPv4 报文流量均衡模式

- 为可选配置。
- 需要修改 IPv4 报文均衡方式时，在增强型模板配置模式下进行配置。
- 可在汇聚或者核心交换机等支持增强型流量均衡功能的设备上配置该功能。

【命令格式】 **ipv4 field** { [**src-ip**] [**dst-ip**] [**protocol**] [**l4-src-port**] [**l4-dst-port**] [**vlan**] [**src-port**] }

- 【参数说明】 **src-ip**：根据输入 IPv4 报文的源 IP 地址进行流量分配。
dst-ip：根据输入 IPv4 报文的目的 IP 地址进行流量分配。
protocol：根据输入的 IPv4 报文的协议类型进行流量分配。
I4-src-port：根据输入的 IPv4 报文的 L4 层的源端口号进行流量分配。
I4-dst-port：根据输入的 IPv4 报文的 L4 层的目的端口号进行流量分配。
vlan：根据输入 IPv4 报文的 vlan 值进行流量分配。
src-port：根据输入 IPv4 报文的面板端口进行流量分配。

【缺省配置】 缺省模式下，IPv4 报文负载均衡方式为 **src-ip**、**dst-ip**。

【命令模式】 **profile** 配置模式

【使用指导】 配置指定增强模板中 IPv4 报文的负载均衡方式。缺省为 **src-ip**、**dst-ip**。
 要将 IPv4 报文的流量平衡设置恢复到缺省值，可以在该模式下使用 **no ipv4 field** 命令。

【命令格式】 **show load-balance-profile [profile-name]**

【参数说明】 *profile-name*：模板名称

【命令模式】 各模式均可执行

【使用指导】 如果没有指定 *profile-name*，则所有增强模式模板的信息将被显示出来。

【命令展示】 Ruijie# show load-balance-profile module0

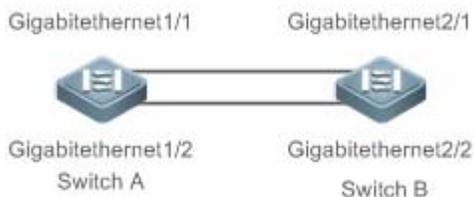
```
Load-balance-profile: module0
Packet Hash Field:
IPv4: src-ip dst-ip
L2 : src-mac dst-mac vlan
```

配置举例

配置流量平衡模式

【网络环境】

图 3-6



- 【配置方法】
- 将 SwitchA 上的端口 GigabitEthernet 1/1 和 GigabitEthernet 1/2 加入到静态 AP 3 中。
 - 将 SwitchB 上的端口 GigabitEthernet 2/1 和 GigabitEthernet 2/2 加入到静态 AP 3 中。
 - 在 SwitchA 上配置全局的 AP 流量均衡模式为基于源 MAC 地址的流量均衡方式。
 - 在 SwitchB 上配置全局的 AP 流量均衡模式为基于目的 MAC 地址的流量均衡方式。

SwitchA

```
SwitchA# configure terminal
SwitchA(config)# interface range GigabitEthernet 1/1-2
```

```
SwitchA(config-if-range)# port-group 3
SwitchA(config-if-range)# exit
SwitchA(config)# aggregateport load-balance src-mac
```

SwitchB

```
SwitchB# configure terminal
SwitchB(config)# interface range GigabitEthernet 2/1-2
SwitchB(config-if-range)# port-group 3
SwitchB(config-if-range)# exit
SwitchB(config)# aggregateport load-balance dst-mac
```

【检验方法】

- 通过 `show aggregateport load-balance` 查看 AP 的流量均衡算法配置是否正确。

SwitchA

```
SwitchA# show aggregatePort load-balance
Load-balance   : Source MAC
```

SwitchB

```
SwitchB# show aggregatePort load-balance
Load-balance   : Destination MAC
```

常见错误

3.5.5 配置AP的容量模式

配置效果

- 改变当前系统支持的最大可配置 AP 口数和单个 AP 口下最大可配置成员口数。

注意事项

- 默认配置下，系统有一个默认的 AP 容量模式，可以通过 `show aggregateport capacity` 命令查看当前容量模式。
- 配置容量模式时，当系统中已经存在的最大 AP 号或者某个 AP 下成员口数量超过了要配置的容量值，则容量模式配置会失败。

配置方法

📌 配置 AP 容量模式

- 为可选配置，当需要改变当前系统 AP 的容量值时配置，以适应网络部署中 AP 的个数或者每个 AP 口允许聚合的成员口个数的变化需求。
- 可在核心交换机等支持改变 AP 容量功能的设备上配置该功能。

【命令格式】 `aggregateport capacity mode capacity-mode`

【参数说明】 `capacity-mode`：模式选项

- 【缺省配置】 缺省情况下，AP 的容量模式随着不同的产品系列而不同，比如有 256*16(其中，256 代表设备支持的最大 AP 口个数，16 代表每个 AP 口支持的最大成员口个数)等容量模式。
- 【命令模式】 全局配置模式
- 【使用指导】 在支持容量模式配置的产品中，系统会提供几种可配置的容量模式供用户选择，在全局配置模式下，用户可以通过 **aggregateport capacity mode capacity-mode** 配置命令来选择需要的容量模式。用户可以在全局配置模式下，通过 **no aggregateport capacity mode** 将容量模式恢复为默认值。

检验方法

- 通过 **show running** 命令查看相应的配置。
- 通过 **show aggregateport capacity** 命令查看当前 AP 容量模式以及 AP 口容量使用情况。

【命令格式】 **show aggregateport capacity**

【参数说明】 -

【命令模式】 各模式均可执行

【使用指导】 -

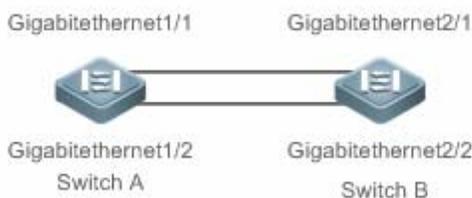
```
Ruijie# show aggregateport capacity
AggregatePort Capacity Information:
Configuration Capacity Mode: 128*16.
Effective Capacity Mode      : 256*8.
Available Capacity           : 128*8.
Total Number: 128, Used: 1, Available: 127.
```

配置举例

配置 AP 的容量模式

【网络环境】

图 3-7



- 【配置方法】
- 将 SwitchA 上的端口 GigabitEthernet 1/1 和 GigabitEthernet 1/2 加入到静态 AP 3 中。
 - 将 SwitchB 上的端口 GigabitEthernet 2/1 和 GigabitEthernet 2/2 加入到静态 AP 3 中。
 - 将 SwitchA 上的 AP 容量模式配置为 128*128 模式。
 - 将 SwitchB 上的 AP 容量模式配置为 256*64 模式。

SwitchA

```
SwitchA# configure terminal
SwitchA(config)# interface range GigabitEthernet 1/1-2
```

SwitchB

```
SwitchA(config-if-range)# port-group 3
SwitchA(config-if-range)# exit
SwitchA(config)# aggregateport capacity mode 128*128

SwitchB# configure terminal
SwitchB(config)# interface range GigabitEthernet 2/1-2
SwitchB(config-if-range)# port-group 3
SwitchB(config-if-range)# exit
SwitchB(config)# aggregateport capacity mode 256*64
```

【检验方法】

- 通过 show aggregateport capacity 查看 AP 的容量模式是否正确。

SwitchA

```
SwitchA# show aggregatePort capacity
AggregatePort Capacity Information:
Configuration Capacity Mode: 128*128.
Effective Capacity Mode      : 128*128.
Available Capacity Mode     : 128*128.
Total Number : 128, Used: 1, Available: 127.
```

SwitchB

```
SwitchB# show aggregatePort capacity
AggregatePort Capacity Information:
Configuration Capacity Mode: 256*64.
Effective Capacity Mode     : 256*64.
Available Capacity Mode     : 256*64.
Total Number : 256, Used: 1, Available: 255.
```

常见错误

-

3.6 监视与维护

清除各类信息

-

查看运行情况

作用	命令
显示增强模式模板的设置。	show load-balance-profile [<i>profile-name</i>]
查看 LACP 的链路聚合状态，可指定显示特定聚合组的信息，参数 <i>key-numebr</i> 表示 LACP 聚合组的 ID。	show lacp summary [<i>key-numebr</i>]

显示 AP 口摘要信息或流量平衡算法。	show aggregateport [<i>ap-number</i>] { load-balance summary }
显示 AP 当前容量模式以及 AP 口容量使用情况	show aggregateport capacity

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 AP 的调试开关。	debug lsm ap
打开 LACP 的调试开关。	debug lacp { packet event database ha realtime stm timer all }

4 VLAN

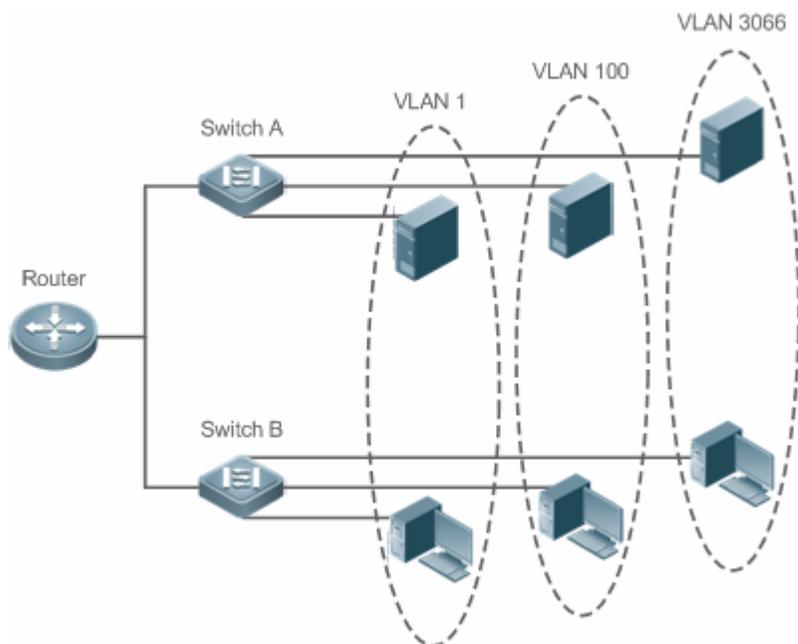
4.1 概述

VLAN 是虚拟局域网 (Virtual Local Area Network) 的简称,它是在一个物理网络上划分出来的逻辑网络。这个网络对应于 ISO 模型的第二层网络。

VLAN 有着和普通物理网络同样的属性,除了没有物理位置的限制,它和普通局域网一样。第二层的单播、广播和多播帧在一个 VLAN 内转发、扩散,而不会直接进入其他的 VLAN 之中。

可以把一个端口定义为一个 VLAN 的成员,所有连接到这个特定端口的终端都是虚拟网络的一部分,并且整个网络可以支持多个 VLAN。当在 VLAN 中增加、删除和修改用户的时候,不必从物理上调整网络配置。VLAN 之间的通讯必须通过三层设备,见下图。

图 4-1



协议规范

- IEEE 802.1Q

4.2 典型应用

典型应用	场景描述
VLAN间二层隔离、三层互连	用户内网被划分为多个 VLAN, 实现相互间的 2 层隔离, VLAN 间通过 3 层核心交换

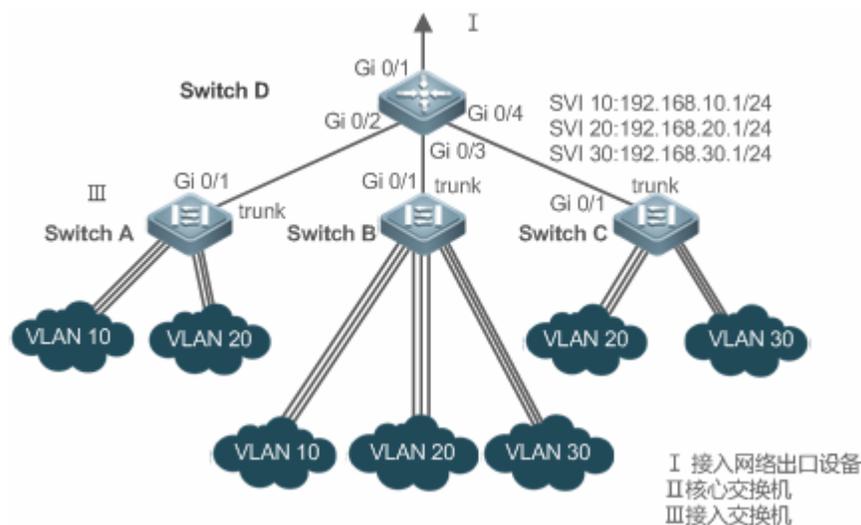
机的 IP 转发能力实现子网互连。

4.2.1 VLAN间二层隔离、三层互连

应用场景

某用户内网被划分为 VLAN 10、VLAN 20、VLAN 30，以实现相互间的 2 层隔离；3 个 VLAN 对应的 IP 子网分别为 192.168.10.0/24、192.168.20.0/24、192.168.30.0/24，3 个 VLAN 通过 3 层核心交换机的 IP 转发能力实现子网互连。

图 4-2



【注释】 Switch A、Switch B、Switch C 为接入交换机。

在核心交换机配置 3 个 VLAN，配置下连接入交换机的端口为 trunk 口，并指定许可 vlan 列表，实现 2 层隔离；在核心交换机配置 3 个 SVI 口，分别作为 3 个 VLAN 对应 IP 子网的网关接口，配置对应的 IP 地址；分别在 3 台接入交换机创建 VLAN，为各 VLAN 分配 Access 口，指定上连核心交换机的 trunk 口。

功能部属

- 在 Intranet 中通过划分多个 VLAN，实现 VLAN 间的二层隔离。
- 三层交换设备中配置 SVI 接口，实现 VLAN 之间的三层通信。

4.3 功能详解

基本概念

▾ VLAN

VLAN 是虚拟局域网 (Virtual Local Area Network) 的简称，它是在一个物理网络上划分出来的逻辑网络。VLAN 有着和普通物理网络同样的属性，除了没有物理位置的限制，它和普通局域网一样。第二层的单播、广播和多播帧在一个 VLAN 内转发、扩散，而不会直接进入其他的 VLAN 之中。

- i 产品支持的 VLAN 遵循 IEEE802.1Q 标准，最多支持 4094 个 VLAN(VLAN ID 1-4094)，其中 VLAN 1 是不可删除的默认 VLAN。
- i 许可配置的 VLAN ID 范围为 1-4094。
- i 当硬件资源不足的情况下，系统将返回创建 VLAN 失败信息。

📌 VLAN 成员类型

可以通过配置一个端口的 VLAN 成员类型，来确定这个端口能通过怎样的帧，以及这个端口可以属于多少个 VLAN。关于 VLAN 成员类型的详细说明，请看下表：

端口类型	作用
Access 端口	一个 Access 端口，只能属于一个 VLAN，并且是通过手工设置指定 VLAN 的。
Trunk 端口 (802.1Q)	一个 Trunk 口，在缺省情况下是属于本设备所有 VLAN 的，它能够转发所有 VLAN 的帧，也可以通过设置许可 VLAN 列表(Allowed-VLANs)来加以限制。
Uplink 端口	一个 Uplink 口，在缺省情况下是属于本设备所有 VLAN 的，它能够转发所有 VLAN 的帧，并且以 tag 方式转发 native-vlan 的帧。
Hybrid 端口	一个 Hybrid 口，在缺省情况下是属于本设备所有 VLAN 的，它能够转发所有 VLAN 的帧，并且允许以 untag 方式转发多个 VLAN 的帧，也可以通过设置许可 VLAN 列表 (Allowed-VLANs)来加以限制。

功能特性

功能特性	作用
VLAN	划分的 VLAN 间二层隔离

4.3.1 VLAN

VLAN 是虚拟局域网的简称，每个 VLAN 具备 VLAN 的独立广播域，不同的 VLAN 之间是二层隔离的。

工作原理

每个 VLAN 具备 VLAN 的独立广播域，不同的 VLAN 之间是二层隔离的。

VLAN 的二层隔离：如果 VLAN 没有配置 SVI，各个 VLAN 之间是二层隔离的，即 VLAN 间的用户之间不能通信；

VLAN 的三层互连：三层交换设备中如果 VLAN 配置 SVI，各个 VLAN 间能三层互连通信；

4.4 配置详解

配置项	配置建议 & 相关命令
-----	-------------

配置基本VLAN	 必选配置。用于创建 VLAN，加入 ACCESS 模式接口。
	vlan 输入一个 VLAN ID。
	 可选配置。配置 ACCESS 口，用于传输单个 VLAN 的信息。
	switchport mode access 定义该接口的类型为二层 Access 口
	switchport access vlan 将这个接口分配给一个 vlan
	add interface 向当前 VLAN 中添加一个或一组 Access 口
	 可选配置，用于 VLAN 重命名。
name 为 VLAN 取一个名字。	
配置TRUNK	 必选配置。配置接口模式为 TRUNK 口。
	switchport mode trunk 定义该接口的类型为二层 Trunk 口
	 可选配置。配置 TRUNK 口，用于传输多个 VLAN。
	switchport trunk allowed vlan 配置这个 Trunk 口的许可 VLAN 列表。
	switchport trunk native vlan 为这个口指定一个 Native VLAN
配置UPLINK	 必选配置。配置接口模式为 UPLINK 口。
	switchport mode uplink 配置为端口为 Uplink 口
	 可选配置，用于恢复接口模式。
	no switchport mode 删除端口模式
配置HYBRID	 必选配置。配置接口模式为 HYBRID 口。
	switchport mode hybrid 配置为端口为 Hybrid 口
	 可选配置。用于转发多个 VLAN 的帧，并且允许以 UNTAG 方式转发多个 VLAN 的帧。
	no switchport mode 删除端口模式
	switchport hybrid allowed vlan 设置端口的输出规则
	switchport hybrid native vlan 设置 Hybrid 口的默认 VLAN

4.4.1 配置基本VLAN

配置效果

- 一个 VLAN 是以 VLAN ID 来标识的。在设备中，您可以添加、删除、修改 VLAN 2-4094，而 VLAN 1 是由设备自动创建，并且不可被删除。可以在接口配置模式下配置一个端口的 VLAN 成员类型或加入、移出一个 VLAN。

注意事项

- 无

配置方法

📌 创建、修改一个 vlan

- 必须配置。
- 当硬件资源不足的情况下，系统将返回创建 VLAN 失败信息。

- 使用 **vlan** *vlan-id* 命令添加一个新的 VLAN 或者进入 VLAN 模式。
- 交换机设备上配置。

【命令格式】 **vlan** *vlan-id*

【参数说明】 *vlan-id*: VLAN vid, 范围为 1-4094

【缺省配置】 VLAN 1 由设备自动创建, 并且不可被删除

【命令模式】 全局配置模式

【使用指导】 如果输入的是一个新的 VLAN ID, 则设备会创建一个 VLAN, 如果输入的是已经存在的 VLAN ID, 则修改相应的 VLAN。使用 **no vlan** *vlan-id* 命令可以删除 vlan, 其中不允许删除的 VLAN 有: 默认 VLAN1、配置 SVI 的 VLAN、SUBVLAN 等。

📌 vlan 重命名

- 可选配置。
- 用户不能将 VLAN 重命名为其他 VLAN 的缺省名字。
- 交换机设备上配置。

【命令格式】 **name** *vlan-name*

【参数说明】 *vlan-name*: 要重新命名的 VLAN 名字

【缺省配置】 缺省情况下, VLAN 的名称为该 VLAN 的 VLAN ID。比如, VLAN 0004 就是 VLAN 4 的缺省名字。

【命令模式】 VLAN 配置模式

【使用指导】 如果想把 VLAN 的名字改回缺省名字, 只需输入 **no name** 命令即可

📌 将当前 ACCESS 口加入到指定 VLAN

- 可选配置。
- 通过 **switchport mode access** 命令指定二层接口 (switch port) 的模式为 access 口。
- 通过 **switchport access vlan** *vlan-id* 命令将一个 access port 加入指定 VLAN, 可传输该 VLAN 流量。
- 交换机设备上配置。

【命令格式】 **switchport mode access**

【参数说明】 -

【缺省配置】 switch port 缺省模式为 access

【命令模式】 接口配置模式

【使用指导】 -

【命令格式】 **switchport access vlan** *vlan-id*

【参数说明】 *vlan-id*: VLAN vid :

【缺省配置】 Access 口缺省仅加入 VLAN 1

【命令模式】 接口配置模式

【使用指导】 如果把一个接口分配给一个不存在的 VLAN, 那么这个 VLAN 将自动被创建。

📌 向当前 VLAN 添加 ACCESS 口

- 可选配置。
- 该命令只对 Access 口有效, VLAN 添加 Access 口后, 接口可传输该 VLAN 数据。
- 交换机设备上配置。

- 【命令格式】 **add interface** { *interface-id* | **range** *interface-range* }
- 【参数说明】 *interface-id* : 单个接口
interface-range : 多个接口
- 【缺省配置】 缺省情况下，所有二层以太网口都属于 VLAN1
- 【命令模式】 VLAN 配置模式
- 【使用指导】 在 VLAN 配置模式下，将指定的 Access 口加入该 VLAN。该命令的配置效果同在接口模式下指定该接口所属 VLAN 的命令（即 **switchport access vlan *vlan-id***）效果一致。

i 对于两种形式的接口加入 VLAN 命令，配置生效的原则是后配置的命令覆盖前面配置的命令

检验方法

- 往 ACCESS 口发送 untag 报文，报文在该 VLAN 内广播。
- 使用命令 **show vlan** 和 **show interface switchport** 查看配置显示是否生效。

【命令格式】 **show vlan** [*id* *vlan-id*]

【参数说明】 *vlan-id* : VLAN ID 号

【命令模式】 所有模式

【使用指导】 -

【命令展示】 Ruijie(config-vlan)#show vlan id 20

VLAN Name	Status	Ports
20 VLAN0020	STATIC	Gi0/1

配置举例

基本 VLAN 与 access 口配置

以下配置举例，仅介绍 VLAN 相关的配置。

- 【配置方法】
- 创建一个新 VLAN，并且重命名
 - 将一个 ACCESS 口加入加入 VLAN，两种方式。

```
Ruijie# configure terminal
Ruijie(config)# vlan 888
Ruijie(config-vlan)# name test888
Ruijie# configure terminal
Ruijie(config)# interface GigabitEthernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# switchport mode access
Ruijie(config-if-GigabitEthernet 0/3)# switchport access vlan 20
```

或者用如下方式：把 Access 口 (GigabitEthernet 0/3) 添加到 VLAN20：

```
Ruijie# configure terminal
SwitchA(config)#vlan 20
SwitchA(config-vlan)#add interface GigabitEthernet 0/3
```

【检验方法】 show 显示是否正确

```
Ruijie(config-vlan)#show vlan
VLAN Name                Status   Ports
-----
 1 VLAN0001                STATIC
20 VLAN0020                STATIC   Gi0/3
888 test888                STATIC
Ruijie(config-vlan)#

Ruijie# show interface GigabitEthernet 0/3 switchport
Interface                Switchport Mode   Access Native Protected VLAN lists
-----
GigabitEthernet 0/3     enabled  ACCESS  20    1    Disabled ALL
```

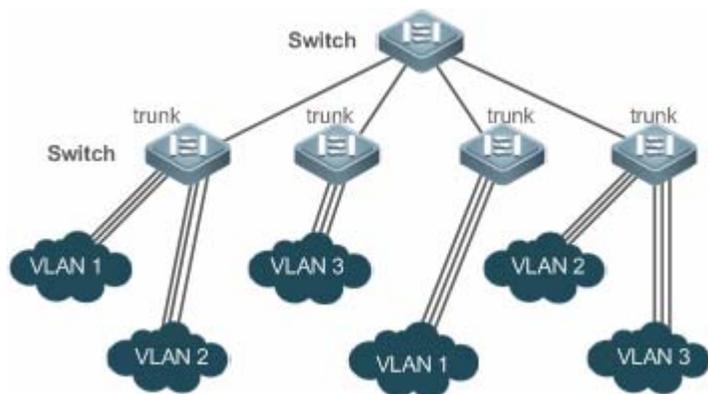
4.4.2 配置TRUNK

配置效果

一个 Trunk 是将一个或多个以太网交换接口和其他的网络设备（如路由器或交换机）进行连接的点对点链路，一条 Trunk 链路可以传输属于多个 VLAN 的流量。

锐捷设备的 Trunk 采用 802.1Q 标准封装。下图显示了一个采用 Trunk 连接的网络。

图 4-3



您可以把一个普通的以太网端口，或者一个 Aggregate Port 设为一个 Trunk 口（关于 Aggregate Port 的详细说明，请见配置 Aggregate Port）。

必须为 Trunk 口指定一个 Native VLAN。所谓 Native VLAN，就是指在这个接口上收发的 UNTAG 报文，都被认为是属于这个 VLAN 的。显然，这个接口的缺省 VLAN ID（即 IEEE 802.1Q 中的 PVID）就是 Native VLAN 的 VLAN ID。同时，在 Trunk 上发送属于 Native VLAN 的帧，则必然采用 UNTAG 的方式。每个 Trunk 口的缺省 Native VLAN 是 VLAN 1。

在配置 Trunk 链路时，请确认连接链路两端的 Trunk 口使用相同的 Native VLAN。

配置方法

配置一个 TRUNK 口

- 必须配置。
- 将接口配置成 trunk 可传输多个 VLAN 的流量。
- 交换机设备上配置。

【命令格式】 **switchport mode trunk**

【参数说明】 -

【缺省配置】 缺省模式是 ACCESS 模式，可配置成 TRUNK 模式

【命令模式】 接口配置模式

【使用指导】 如果想把一个 Trunk 口的所有 Trunk 相关属性都复位成缺省值，请使用 **no switchport mode** 配置命令。

定义 Trunk 口的许可 VLAN 列表

- 可选配置。
- 一个 Trunk 口缺省可以传输本设备支持的所有 VLAN (1 - 4094) 的流量。也可以通过设置 Trunk 口的许可 VLAN 列表来限制某些 VLAN 的流量不能通过这个 Trunk 口。
- 交换机设备上配置。

【命令格式】 **switchport trunk allowed vlan {all | [add | remove | except | only] } vlan-list**

【参数说明】 参数 vlan-list 可以是一个 VLAN，也可以是一系列 VLAN，VLAN ID 按顺序排列，中间用“-”号连接。如：10-20。

all 的含义是许可 VLAN 列表包含所有支持的 VLAN；

add 表示将指定 VLAN 列表加入许可 VLAN 列表；

remove 表示将指定 VLAN 列表从许可 VLAN 列表中删除；

except 表示将除列出的 VLAN 列表外的所有 VLAN 加入许可 VLAN 列表；

only 表示将列出的 VLAN 列表加入许可 VLAN 列表，其他 VLAN 从许可列表中删除；

【缺省配置】 trunk 口和 uplink 口属于所有 VLAN

【命令模式】 接口配置模式

【使用指导】 如果想把 Trunk 的许可 VLAN 列表改为缺省的许可所有 VLAN 的状态，请使用 **no switchport trunk allowed vlan** 接口配置命令

配置 Native VLAN

- 可选配置。
- 一个 Trunk 口能够收发 TAG 或者 UNTAG 的 802.1Q 帧。其中 UNTAG 帧用来传输 Native VLAN 的流量。缺省的 Native VLAN 是 VLAN 1。
- 如果一个帧带有 Native VLAN 的 VLAN ID，在通过这个 Trunk 口转发时，会自动被剥去 TAG。
- 交换机设备上配置。

【命令格式】 **switchport trunk native vlan vlan-id**

【参数说明】 *vlan-id*: VLAN vid

【缺省配置】 trunk/uplink 的默认 VLAN 为 VLAN 1

【命令模式】 接口配置模式

【使用指导】 如果想把 Trunk 的 Native VLAN 列表改回缺省的 VLAN 1，请使用 **no switchport trunk native vlan** 接口配置命令。

- 把一个接口的 Native VLAN 设置为一个不存在的 VLAN 时，设备不会自动创建此 VLAN。此外，一个接口的 Native VLAN 可以不在接口的许可 VLAN 列表中。此时，Native VLAN 的流量不能通过该接口。

检验方法

- 往 TRUNK 口发送 tag 报文，报文在指定 VLAN 内广播。
- 使用命令 **show vlan** 和 **show interface switchport** 查看配置显示是否生效。

【命令格式】 **show vlan [id vlan-id]**

【参数说明】 *vlan-id* : VLAN ID 号

【命令模式】 所有模式

【使用指导】 -

【命令展示】 Ruijie(config-vlan)#show vlan id 20

VLAN Name	Status	Ports
20 VLAN0020	STATIC	Gi0/1

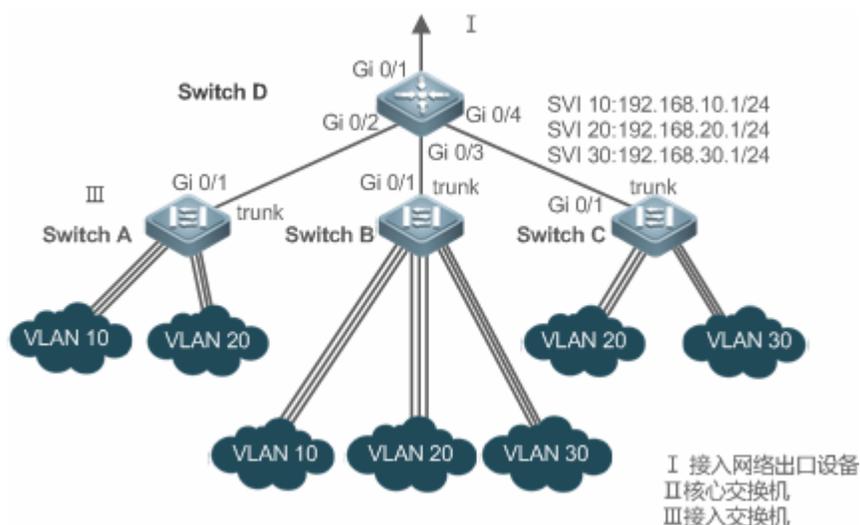
配置举例

⚠ 以下配置举例，仅介绍 TRUNK 相关的配置。

配置基本 VLAN，实现二层隔离、三层互连

【网络环境】

图 4-4



【配置方法】

组网需求

如上图所示，某用户内网被划分为 VLAN 10、VLAN 20、VLAN 30，以实现相互间的 2 层隔离；3 个 VLAN 对应的 IP 子网分别为 192.168.10.0/24、192.168.20.0/24、192.168.30.0/24，3 个 VLAN 通过 3 层核心交换机的 IP 转发能力实现子网互连。

配置要点

本例以核心交换机和 1 台接入交换机为例说明配置过程。要点如下：

- 在核心交换机配置 3 个 VLAN，配置下连接入交换机的端口为 trunk 口，并指定许可 vlan 列表，实现 2 层隔离；
- 在核心交换机配置 3 个 SVI 口，分别作为 3 个 VLAN 对应 IP 子网的网关接口，配置对应的 IP 地址；
- 分别在 3 台接入交换机创建 VLAN，为各 VLAN 分配 Access 口，指定上连核心交换机的 trunk 口。本例以接入交换机 Switch A 为例说明配置步骤。

D

```
D#configure terminal
D(config)#vlan 10
D(config-vlan)#vlan 20
D(config-vlan)#vlan 30
D(config-vlan)#exit
D(config)#interface range GigabitEthernet 0/2-4
D(config-if-range)#switchport mode trunk
D(config-if-range)#exit
D(config)#interface GigabitEthernet 0/2
D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan remove 1-4094
D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan add 10,20
D(config-if-GigabitEthernet 0/2)#interface GigabitEthernet 0/3
D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan remove 1-4094
D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan add 10,20,30
D(config-if-GigabitEthernet 0/2)#interface GigabitEthernet 0/4
D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan remove 1-4094
D(config-if-GigabitEthernet 0/2)#switchport trunk allowed vlan add 20,30
D#configure terminal
D(config)#interface vlan 10
D(config-if-VLAN 10)#ip address 192.168.10.1 255.255.255.0
D(config-if-VLAN 10)#interface vlan 20
D(config-if-VLAN 20)#ip address 192.168.20.1 255.255.255.0
D(config-if-VLAN 20)#interface vlan 30
D(config-if-VLAN 30)#ip address 192.168.30.1 255.255.255.0
D(config-if-VLAN 30)#exit
```

A

```
A#configure terminal
A(config)#vlan 10
A(config-vlan)#vlan 20
A(config-vlan)#exit
A(config)#interface range GigabitEthernet 0/2-12
A(config-if-range)#switchport mode access
A(config-if-range)#switchport access vlan 10
A(config-if-range)#interface range GigabitEthernet 0/13-24
A(config-if-range)#switchport mode access
A(config-if-range)#switchport access vlan 20
A(config-if-range)#exit
```

```
A(config)#interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)#switchport mode trunk
```

【检验方法】 在核心交换机上查看 vlan 配置

- 查看 vlan 信息，包括 vlan id、名称、状态、包括的端口
- 查看端口 Gi 0/2、Gi 0/3、Gi 0/4 的 vlan 状态

D

```
D#show vlan
VLAN Name          Status            Ports
-----
  1 VLAN0001    STATIC          Gi0/1, Gi0/5, Gi0/6, Gi0/7
                               Gi0/8, Gi0/9, Gi0/10, Gi0/11
                               Gi0/12, Gi0/13, Gi0/14, Gi0/15
                               Gi0/16, Gi0/17, Gi0/18, Gi0/19
                               Gi0/20, Gi0/21, Gi0/22, Gi0/23
                               Gi0/24
 10 VLAN0010    STATIC          Gi0/2, Gi0/3
 20 VLAN0020    STATIC          Gi0/2, Gi0/3, Gi0/4
 30 VLAN0030    STATIC          Gi0/3, Gi0/4

D#show interface GigabitEthernet 0/2 switchport
Interface                Switchport Mode    Access Native Protected VLAN lists
-----
GigabitEthernet 0/2      enabled  TRUNK    1    1    Disabled 10,20

D#show interface GigabitEthernet 0/3 switchport
Interface                Switchport Mode    Access Native Protected VLAN lists
-----
GigabitEthernet 0/3      enabled  TRUNK    1    1    Disabled 10,20,30

D#show interface GigabitEthernet 0/4 switchport
Interface                Switchport Mode    Access Native Protected VLAN lists
-----
GigabitEthernet 0/4      enabled  TRUNK    1    1    Disabled 20,30
```

常见错误

- 无

4.4.3 配置UPLINK

配置效果

- UPLINK 端口一般用于 QinQ(出自标准 IEEE 802.1ad)环境中,它和 TRUNK 端口的功能很相似,不同之处在于 UPLINK 端口只发送 TAG 帧,而 TRUNK 端口缺省 VLAN 的帧以 UNTAG 形式发送。

配置方法

配置一个 UPLINK 口

- 必须配置。
- 将接口配置成 uplink 口，可传输多个 vlan 的流量，但只能发送 TAG 帧。
- 交换机设备上配置。

【命令格式】 **switchport mode uplink**

【参数说明】 -

【缺省配置】 缺省模式是 ACCESS 模式，可配置成 ACCESS 模式 UPLINK 模式

【命令模式】 接口配置模式

【使用指导】 如果想把一个 UPLINK 口的所有 UPLINK 相关属性都复位成缺省值，请使用 **no switchport mode** 配置命令。

定义 UPLINK 口的许可 VLAN 列表

- 可选配置。
- 可以通过设置 UPLINK 口的许可 VLAN 列表来限制某些 VLAN 的流量不能通过这个 UPLINK 口。
- 交换机设备上配置。

【命令格式】 **switchport trunk allowed vlan {all | [add | remove | except | only] } vlan-list**

【参数说明】 参数 vlan-list 可以是一个 VLAN，也可以是一系列 VLAN，VLAN ID 按顺序排列，中间用“-”号连接。如：
10-20。

all 的含义是许可 VLAN 列表包含所有支持的 VLAN；

add 表示将指定 VLAN 列表加入许可 VLAN 列表；

remove 表示将指定 VLAN 列表从许可 VLAN 列表中删除；

except 表示将除列出的 VLAN 列表外的所有 VLAN 加入许可 VLAN 列表；

only 表示将列出的 VLAN 列表加入许可 VLAN 列表，其他 VLAN 从许可列表中删除；

【命令模式】 接口配置模式

【使用指导】 如果想把 UPLINK 的许可 VLAN 列表改为缺省的许可所有 VLAN 的状态，请使用 **no switchport trunk allowed vlan** 接口配置命令

配置 Native VLAN

- 可选配置。
- 如果一个帧带有 Native VLAN 的 VLAN ID，在通过这个 UPLINK 口转发时，不会被剥去 TAG。这与 TRUNK 相反。
- 交换机设备上配置。

【命令格式】 **switchport trunk native vlan vlan-id**

【参数说明】 *vlan-id*: VLAN vid

【命令模式】 接口配置模式

【使用指导】 如果想把 UPLINK 的 Native VLAN 列表改回缺省的 VLAN 1，请使用 **no switchport trunk native vlan** 接口配置命令。

检验方法

- 往 UPLINK 口发送 tag 报文，报文在指定 VLAN 内广播。
- 使用命令 **show vlan** 和 **show interface switchport** 查看配置显示是否生效。

【命令格式】 **show vlan [id vlan-id]**

【参数说明】 *vlan-id* : VLAN ID 号

【命令模式】 所有模式

【使用指导】 -

【命令展示】 Ruijie(config-vlan)#show vlan id 20

VLAN Name	Status	Ports
20 VLAN0020	STATIC	Gi0/1

配置举例

配置一个 uplink 口

 以下配置举例，仅介绍 UPLINK 相关的配置。

【配置方法】 下面是一个把端口 Gi0/1 变成 UPLINK 的例子：

```
Ruijie# configure terminal
Ruijie(config)# interface gi 0/1
Ruijie(config-if-GigabitEthernet 0/1)# switchport mode uplink
Ruijie(config-if-GigabitEthernet 0/1)# end
```

【检验方法】 show 显示是否正确

```
Ruijie# show interfaces GigabitEthernet 0/1 switchport
```

Interface	Switchport Mode	Access	Native	Protected	VLAN lists
GigabitEthernet 0/1	enabled	UPLINK	1	1	disabled ALL

4.4.4 配置HYBRID

配置效果

- HYBRID 端口一般用于 SHARE VLAN 的环境中。HYBRID 端口在缺省情况下与 TRUNK 端口相同，不同是它可以设置除了缺省 VLAN 外的其它 VLAN 的帧以 UNTAG 形式发送

配置方法

配置一个 HYBRID 口

- 必须配置。
- 将接口配置成 hybrid 口，可传输多个 VLAN 的流量。

- 交换机设备上配置。

【命令格式】 **switchport mode hybrid**

【参数说明】 -

【缺省配置】 缺省模式是 ACCESS 模式，可配置成 HYBRID 模式

【命令模式】 接口配置模式

【使用指导】 如果想把一个 HYBRID 口的所有 HYBRID 相关属性都复位成缺省值，请使用 **no switchport mode** 配置命令。

▾ 定义 HYBRID 口的许可 VLAN 列表

- 可选配置。
- 一个 HYBRID 口缺省可以传输本设备支持的所有 VLAN (1 - 4094) 的流量。也可以通过设置 HYBRID 口的许可 VLAN 列表来限制某些 VLAN 的流量不能通过这个 HYBRID 口。
- 交换机设备上配置。

【命令格式】 **switchport hybrid allowed vlan [[add | only] tagged | [add] untagged | remove] vlan_list**

【参数说明】 *vlan-id*: VLAN vid

【缺省配置】 默认 hybrid 口属于所有 VLAN，端口以 Tag 形式加入所有除了默认 VLAN 以外的其它 VLAN，默认 VLAN 以 UNTag 形式加入

【命令模式】 接口配置模式

【使用指导】 -

▾ 配置 Native VLAN

- 可选配置。
- 如果一个帧带有 Native VLAN 的 VLAN ID，在通过这个 HYBRID 口转发时，会自动被剥去 TAG。
- 交换机设备上配置。

【命令格式】 **switchport hybrid native vlan vlan_id**

【参数说明】 *vlan-id*: VLAN vid

【缺省配置】 缺省的 Native VLAN 是 VLAN 1

【命令模式】 接口配置模式

【使用指导】 如果想把 HYBRID 的 Native VLAN 列表改回缺省的 VLAN 1，请使用 **no switchport hybrid native vlan** 接口配置命令。

检验方法

- 往 HYBRID 口发送 tag 报文，报文在指定 VLAN 内广播。
- 使用命令 **show vlan** 和 **show interface switchport** 查看配置显示是否生效。

【命令格式】 **show vlan [id vlan-id]**

【参数说明】 *vlan-id* : AP VLAN ID 号

【命令模式】 所有模式

【使用指导】 -

【命令展示】 Ruijie(config-vlan)#show vlan id 20

VLAN Name	Status	Ports
20 VLAN0020	STATIC	Gi0/1

配置举例

配置一个 hybrid 口

 以下配置举例，仅介绍 HYBRID 相关的配置。

【配置方法】 下面是一个端口 Gi0/1 关于 HYBRID 配置的例子：

```
Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# switchport mode hybrid
Ruijie(config-if-GigabitEthernet 0/1)# switchport hybrid native vlan 3
Ruijie(config-if-GigabitEthernet 0/1)# switchport hybrid allowed vlan untagged 20-30
Ruijie(config-if-GigabitEthernet 0/1)# end
```

【检验方法】 show run 显示是否正确

```
Ruijie(config-if-GigabitEthernet 0/1)#show run interface gigabitEthernet 0/1

Building configuration...
Current configuration : 166 bytes

interface GigabitEthernet 0/1
 switchport
 switchport mode hybrid
 switchport hybrid native vlan 3
 switchport hybrid allowed vlan add untagged 20-30
```

4.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
查看 VLAN 配置	show vlan
查看交换口配置	show interface switchport

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
----	----

打开 VLAN 的调试开关。

```
debug bridge vlan
```

5 Super VLAN

5.1 概述

Super VLAN 是 VLAN 划分的一种方式。Super VLAN 又称为 VLAN 聚合，是一种专门优化 IP 地址的管理技术。

采用 Super VLAN 技术可以极大的节省 IP 地址，它只需对包含多个 Sub VLAN 的 Super VLAN 分配一个 IP 地址，既节省地址又方便网络管理。

i 下文仅介绍 Super VLAN 的相关内容。

5.2 典型应用

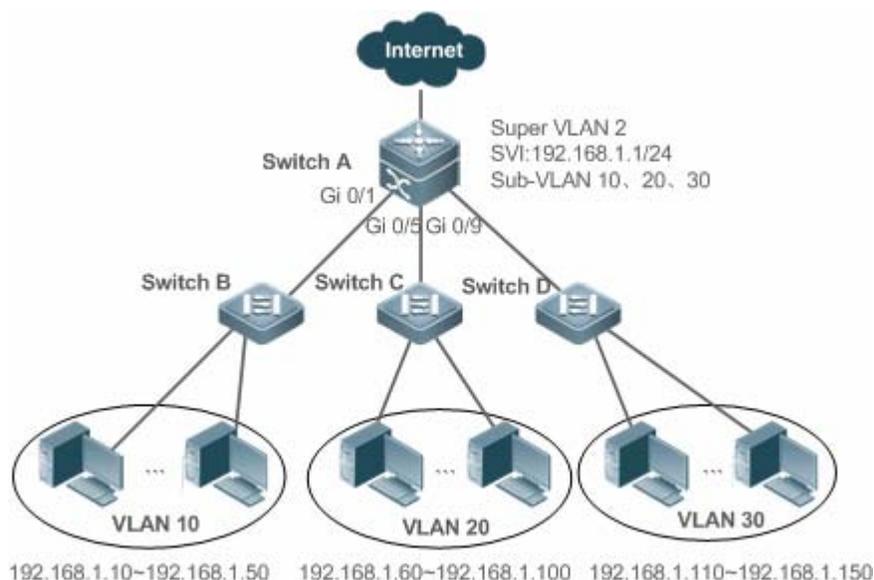
典型应用	场景描述
多个VLAN共享一个IP网关	接入用户通过划分 VLAN 实现二层隔离，所有 VLAN 用户共享一个 IP 网关，实现三层通信以及与外网通信。

5.2.1 多个VLAN共享一个IP网关

应用场景

在一台三层设备上实现多个 VLAN 的二层隔离，但是这些 VLAN 的用户可以在一个网段并进行三层通信。

图 5-1



【注释】 Switch A 为网关设备或核心交换机。

Switch B、Switch C、Switch D 为接入交换机。

Switch A 上配置 Super VLAN 和 Sub VLAN，并为 Super VLAN 配置三层口和三层口的 IP 地址。

Switch B、Switch C、Switch D 上分别配置 VLAN 10、VLAN 20、VLAN 30，将公司的不同部门分别划分在这些 VLAN 内。

功能部属

在 Intranet 中通过 Super VLAN 实现多个子 VLAN 的共享一个 IP 网关，又能保证 VLAN 间的二层隔离。

子 VLAN 内的用户间可以通过 Super VLAN 的网关进行三层通信。

5.3 功能详解

基本概念

Super VLAN

Super VLAN 又称为 VLAN 聚合，是一种专门优化 IP 地址的管理技术，将多个 VLAN 聚合到一个 IP 网段。Super VLAN 不能加入任何物理口，主要通过其 SVI 口来管理 Sub VLAN 的跨 VLAN 通信，不能当做正常的 802.1Q VLAN 来使用。可以将 Super VLAN 看成 Sub VLAN 的主 VLAN。

Sub VLAN

Sub VLAN 又称子 VLAN，每一个 Sub VLAN 都是一个独立的广播域，它们之间是二层隔离。同一 Super VLAN 的 Sub VLAN 或不同 Super VLAN 的 Sub VLAN 的用户之间通信需要依靠各自 Super VLAN 三层口 SVI 实现。

ARP 代理

只有 Super VLAN 才能创建三层口 SVI，子 VLAN 不能创建 SVI。子 VLAN 依靠其主 VLAN (Super VLAN) 的三层口通过 ARP 代理，实现同一 Super VLAN 不同 Sub VLAN 之间以及不同网段用户之间的通信。Sub VLAN 的用户向其他 VLAN 的用户发送 ARP 请求时，主 VLAN 的网关用其 MAC 地址代替其发送和回应 ARP 请求，这个过程称为 ARP 代理。

Sub VLAN IP 地址范围

每个 Sub VLAN 可以根据主 VLAN 配置的网关 IP 地址来配置一个子 IP 地址范围，可以限制 Sub VLAN 内用户所在的 IP 范围。

功能特性

功能特性	作用
Super VLAN	创建三层口，作为一个虚接口通过 ARP 代理来实现其所有 Sub VLAN 共用一个 IP 网段

5.3.1 Super VLAN

Super VLAN 可以使其所有的 Sub VLAN 内用户都划分在同一个 IP 范围内，并通用一个 IP 网关，用户通过这个网关可以跨 VLAN 通信，而不用每个 VLAN 划分一个网关，从而节省了 IP 地址。

工作原理

Super VLAN 的工作原理是将一个网段的 IP 地址分给不同的子 VLAN (Sub VLAN)，这些 Sub VLAN 同属于一个 Super VLAN。每个 Sub VLAN 具备 VLAN 的独立广播域，不同的 Sub VLAN 之间是二层隔离的。当 Sub VLAN 内的用户需要进行三层通信时，使用 Super VLAN 的虚接口的 IP 地址作为网关地址，这样多个 VLAN 共享一个 IP 网关，不用每个 VLAN 配置一个网关。同时，为了实现不同的 Sub VLAN 间的三层互通及 Sub VLAN 与其它网段互通，需要利用 ARP 代理功能，通过 ARP 代理可以进行 ARP 请求和响应报文的转发和处理，从而实现三层通信。

Sub VLAN 的二层通信：如果 Super VLAN 没有配置 SVI，Super VLAN 内的各个 Sub VLAN 之间是二层隔离的，即 Sub VLAN 内的用户之间不能通信；如果 Super VLAN 配置了 SVI，通过 Super VLAN 的网关作为 ARP 代理，同一 Super VLAN 内的 Sub VLAN 之间可以通信，因为这些 Sub VLAN 用户的 IP 是同一个网段，认为还是二层通信。

Sub VLAN 的三层通信：Sub VLAN 内的用户要跨网段进行三层通信时，其所属的 Super VLAN 的网关作为 ARP 代理，代替 Sub VLAN 回应 ARP 请求。

相关配置

配置 Super VLAN

缺省情况下，VLAN 都是普通 VLAN。

使用 **supervlan** 命令将一个普通 VLAN 变成 Super VLAN。

普通 VLAN 变成 Super VLAN 后，加入该 VLAN 的端口都将从这个 VLAN 退出，这是因为 Super VLAN 内不能有任何物理端口。

配置 Sub VLAN

缺省情况下，VLAN 都是普通 VLAN。

使用 **subvlan *vlan-id-list*** 命令将普通 VLAN 变成 Super VLAN 的 Sub VLAN，这些 VLAN 内可以有物理口。

Sub VLAN 内的用户通信由 Super VLAN 来管理。

使能 ARP 代理

缺省情况下，Super VLAN 和 Sub VLAN 的 ARP 代理功能是打开的。

Super VLAN 和 Sub VLAN 均开启该功能，Sub VLAN 内的用户才能和其它 VLAN 的用户通信。

5.4 配置详解

配置项	配置建议 & 相关命令	
配置 Super VLAN 基本功能	 必须配置。	
	supervlan	配置 Super VLAN
	subvlan <i>vlan-id-list</i>	配置 Sub VLAN
	proxy-arp	ARP 代理使能
	interface vlan <i>vlan-id</i>	创建 Super VLAN 的虚拟接口
	ip address <i>ip mask</i>	设置 Super VLAN 虚拟接口的 IP 地址
	 可选配置。	
subvlan-address-range <i>start-ip end-ip</i>	指定 Sub VLAN 内用户的 IP 地址范围	

5.4.1 配置 Super VLAN 基本功能

配置效果

启动 Super VLAN 功能，给 Super VLAN 配置 SVI，实现 Sub VLAN 跨 VLAN 的二三层通信。

同一 Super VLAN 的所有 Sub VLAN 内用户共用一个 IP 网关，不用每个 VLAN 指定一个网段，从而节省 IP 地址。

注意事项

 因为 Super VLAN 不属于任何物理口，所以配置 Super VLAN 的设备不能处理 Tag 为 Super VLAN 的报文。

 必须同时使能 Super VLAN 和 Sub VLAN 的 ARP 代理功能。

 必须为 Super VLAN 配置 SVI 和 IP 地址，作为其所有 Sub VLAN 通信的虚接口，Sub VLAN 内的用户才能跨 VLAN 通信。

配置方法

配置 Super VLAN

必须配置。

该 VLAN 内不包括任何物理口。

必须使能 ARP 代理功能，默认是打开的。

 必须为 Super VLAN 配置 Sub VLAN，Super VLAN 才有意义

 VLAN 1 不能配置为 Super VLAN。

 Super VLAN 不能配置为其它 Super VLAN 的 Sub VLAN，反之亦然。

配置 Super VLAN 的虚拟口

必须配置。

Super VLAN 不能加入任何物理口，通过配置的 VLAN 三层口 SVI 作为虚拟口，通过配置的 IP 网关来代理 Sub VLAN 内用户的 IP 回应其它用户的 ARP 请求。

 Super VLAN 配置 SVI 时，会同时为其所有 Sub VLAN 分配一个对用户不可见的三层口，如果因为资源不足不能为 Sub VLAN 分配三层口，此时会将该 Sub VLAN 恢复为普通 VLAN。

配置 Sub VLAN

必须配置。

Sub VLAN 内可以加入任何物理口，同一 Super VLAN 的 Sub VLAN 共享 Super VLAN 的网关地址，使用同一网段。

必须使能 ARP 代理功能，默认是打开的。

 Sub VLAN 不能直接通过 `no vlan` 命令删除，必须先恢复为普通 VLAN 后才能被删除。

 不同 Super VLAN 的 Sub VLAN 不能有交叠。

ARP 代理

必须配置，默认是打开的。

只有 Super VLAN 和 Sub VLAN 同时使能 ARP 代理功能，Sub VLAN 内的用户才能通过 Super VLAN 的网关代理实现跨 VLAN 二三层通信。

 必须在 Super VLAN 和 Sub VLAN 上使能 ARP 代理，否则 ARP 代理功能不起作用。

Sub VLAN IP 地址范围

如果需要划分用户的 IP 地址使用范围，可以为每个 Sub VLAN 划分一个 IP 地址范围，Sub VLAN 内用户的 IP 地址在规定的范围内才能与其它 VLAN 的用户正常通信，否则不能跨 VLAN 通信。

若无特殊要求，可以不用划分 IP 地址范围。

 这里划分了 IP 地址范围，并不能保证 DHCP 给用户动态 IP 地址也在该范围内，如果 DHCP 分配的 IP 地址不在规定的范围内，用户不能对外通信，请慎用。

 必须保证 Sub VLAN 的 IP 地址范围在 Super VLAN 的网关范围内，否则 Sub VLAN 内的用户无法通信。

 Sub VLAN 内的用户的 IP 地址必须在 Sub VLAN 的 IP 地址范围内，否则 Sub VLAN 内的用户无法通信。

检验方法

各个 Sub VLAN 关联网关后，Sub VLAN 内用户之间互 ping 能通。

相关命令

配置 Super VLAN

- 【命令格式】 **supervlan**
- 【参数说明】 -
- 【命令模式】 VLAN 模式
- 【使用指导】 缺省情况下，Super VLAN 功能是关闭的。
Super VLAN 不能加入任何物理口。
一旦 VLAN 不再是 Super VLAN，其所属的所有 Sub VLAN 都恢复成普通静态 VLAN。

配置 Super VLAN 的虚拟口

- 【命令格式】 **interface vlan *vlan-id***
- 【参数说明】 *vlan-id* : Super VLAN 的 id。
- 【命令模式】 全局模式
- 【使用指导】 为 Super VLAN 配置三层口，作为 Super VLAN 的虚拟接口，必须配置。

Super VLAN 网关

- 【命令格式】 **ip address *ip mask***
- 【参数说明】 *ip* : IP 地址，Super VLAN 虚拟接口的网关地址。
Mask : 掩码。
- 【命令模式】 接口模式
- 【使用指导】 为 Super VLAN 配置网关，Super VLAN 所有 Sub VLAN 的用户都属于这个网关。

Sub VLAN

- 【命令格式】 **subvlan *vlan-id-list***
- 【参数说明】 *vlan-id-list* : 指定若干 VLAN 作为某个 Super VLAN 的 Sub VLAN。
- 【命令模式】 VLAN 模式
- 【使用指导】 用户连接的接口可以加入 Sub VLAN。
不能使用 `no vlan [id]` 来删除 Sub VLAN，必须先转化成普通 VLAN 后才能删除。
不能为 Sub VLAN 配置 VLAN 三层口 SVI。

 如果 Super VLAN 配置了 VLAN 三层口 SVI，再增加 Sub VLAN 时可能会因为资源不足而导致 Sub VLAN 配置失败。

 如果 Super VLAN 配置了 Sub VLAN，再配置 VLAN 三层口 SVI，这时可能会因为资源不足将部分 Sub VLAN 恢复为普通 VLAN。

使能 ARP 代理

【命令格式】 **proxy-arp**

【参数说明】 -

【命令模式】 VLAN 模式

【使用指导】 默认是打开的。

此命令用来使能 Super VLAN 和 Sub VLAN 的 ARP 代理。

只有 Super VLAN 和对应的 Sub VLAN 都使能，Sub VLAN 的用户才能跨 VLAN 进行二三层通信。

配置 Sub VLAN IP 地址范围

【命令格式】 **subvlan-address-range start-ip end-ip**

【参数说明】 *start-ip* : Sub VLAN 的起始 IP 地址。

end-ip : Sub VLAN 的最大 IP 地址。

【命令模式】 VLAN 模式

【使用指导】 可选配置。

此命令用来划分 Sub VLAN 内用户 IP 地址使用范围。

同一个 Super VLAN 的 Sub VLAN 的 IP 地址范围不可以有交叠。

 Sub VLAN 的 IP 地址范围必须在其所属 Super VLAN 的 IP 地址范围内，否则 Sub VLAN 内的用户无法通信。

 Sub VLAN 内的用户 IP 地址(无论静态分配还是 DHCP 静态分配)必须在这个范围内才能与其它 VLAN 的用户进行通信。

 不能保证 DHCP 分配的 IP 地址都在这个范围内，这样就可能造成用户不能通信，所以该命令需要慎用。

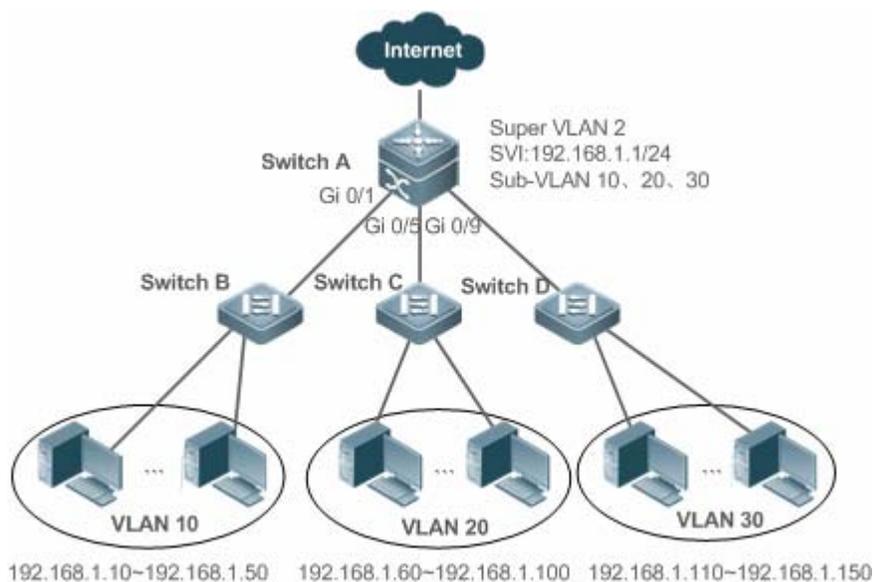
配置举例

以下配置举例，仅介绍 Super VLAN 相关的配置。

在网络中配置 Super VLAN，使其 Sub VLAN 的用户使用同一网段，共享一个 IP 网关，节省 IP 地址。

【网络环境】

图 5-2



【配置方法】

在核心交换机上配置 Super VLAN 必须配置部分。略
在接入交换机上配置对应核心交换机上 Sub VLAN 的普通 VLAN。

A

```
SwitchA#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
SwitchA(config)#vlan 2
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 10
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 20
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 30
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 2
SwitchA(config-vlan)#supervlan
SwitchA(config-vlan)#subvlan 10,20,30
SwitchA(config-vlan)#exit
SwitchA(config)#interface vlan 2
SwitchA(config-if-VLAN 2)#ip address 192.168.1.1 255.255.255.0
SwitchA(config)#vlan 10
SwitchA(config-vlan)#subvlan-address-range 192.168.1.10 192.168.1.50
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 20
SwitchA(config-vlan)#subvlan-address-range 192.168.1.60 192.168.1.100
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 30
SwitchA(config-vlan)#subvlan-address-range 192.168.1.110 192.168.1.150
```

```
SwitchA(config)#interface range gigabitEthernet 0/1,0/5,0/9
SwitchA(config-if-range)#switchport mode trunk
```

【检验方法】 使 Source (192.168.1.10) 与 Dest (192.168.1.60) 之间互 ping 能通。

A

```
SwitchA(config-if-range)# show supervlan
supervlan id  supervlan arp-proxy  subvlan id  subvlan arp-proxy  subvlan ip range
-----
                2                ON                10                ON                192.168.1.10 - 192.168.1.50
                20                ON                192.168.1.60 - 192.168.1.100
                30                ON                192.168.1.110 - 192.168.1.150
```

常见错误

Super VLAN 没有配置 SVI 和 IP 网关，导致 Sub VLAN 之间、Sub VLAN 与其它 VLAN 之间不能通行。

关闭 Super VLAN 或者 Sub VLAN 的 ARP 代理功能，导致 Sub VLAN 的用户间不能跨 VLAN 通信。

配置了 Sub VLAN 的 IP 地址范围，但是给用户分配的 IP 地址不在该范围内。

5.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
Super VLAN 配置	show supervlan

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 Super VLAN 调试开关。	debug bridge svlan

6 Protocol VLAN 配置

6.1 概述

Protocol VLAN技术就是基于报文协议类型的VLAN分类技术，其可以将某一协议类型的空VLAN ID报文都划分到同一个VLAN。即交换机可以根据端口接收到的报文所属的协议类型以及封装格式，将收到的不携带VLAN 标记的报文，与用户设定的协议模板相匹配，匹配成功的自动分发到相应的VLAN中传输。Protocol VLAN共有两种类型：基于IP地址的VLAN分类和端口上的基于报文类型和以太网类型的VLAN分类两种VLAN分类技术，后续说明文中将基于协议类型的Protocol VLAN简称为协议VLAN，基于IP地址类型的Protocol VLAN简称为子网VLAN。

i 下文仅介绍 Protocol VLAN 的相关内容。

i 协议只作用于 Trunk、Hybrid 模式下端口。

协议规范

IEEE standard 802.1Q

6.2 典型应用

典型应用	场景描述
协议VLAN配置应用	实现不同协议报文的用户二层通信隔离，减少网络流量
子网VLAN配置应用	实现根据用户报文所属的 IP 网段确定其 VLAN 范围

6.2.1 协议VLAN配置应用

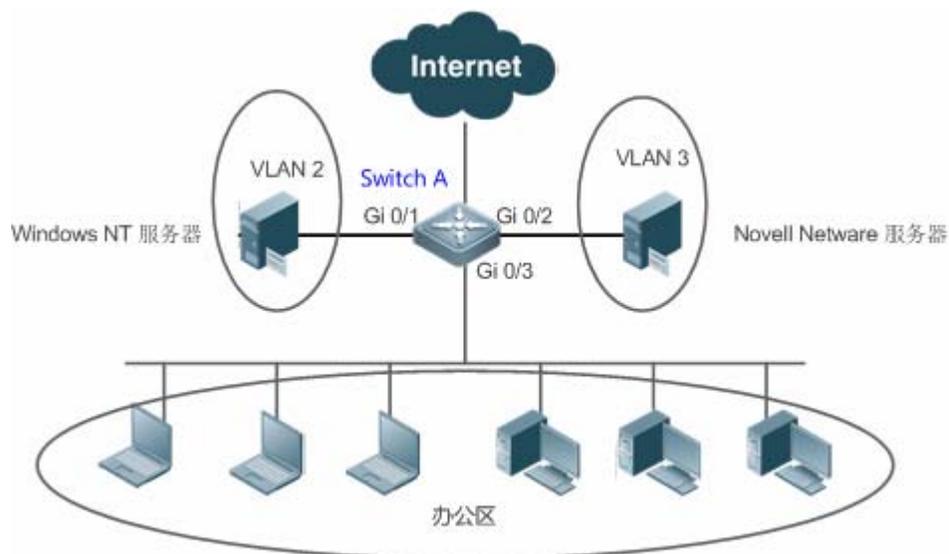
应用场景

如下图所示，由Windows NT和Novell Netware操作系统互联的网络结构，办公区通过HUB与三层设备Switch A相连。在办公区内分散着不同的PC用户，一部分采用Windows NT操作系统，支持IP协议；一部分采用Novell Netware操作系统，支持IPX协议。整个办公区通过上链口Gi 0/3与外网以及服务器通信。

主要需求如下：

- 实现Windows NT和Novell Netware操作系统的PC用户二层通信隔离，减少网络流量。

图 6-1



【注释】 Switch A为交换机设备，端口 Gi 0/3 为 Hybrid 口。Gi0/1 为 Access 口，所属 VLAN 2；Gi0/2 也为 Access 口，所属 VLAN3。

功能部署

- 配置报文类型和以太网类型的 profile（本例将支持 IP 协议的报文对应 Profile 1，支持 IPX 协议的报文对应 Profile 2）
- 将 Profile 应用到上链口（本例对应为 Gi 0/3）上，并与 VLAN 关联（本例将 Profile 1 关联 VLAN 2，Profile 2 关联 VLAN 3）。

! 配置协议 VLAN 的端口仅针对 Trunk 口和 Hybrid 口生效。

6.2.2 子网VLAN配置应用

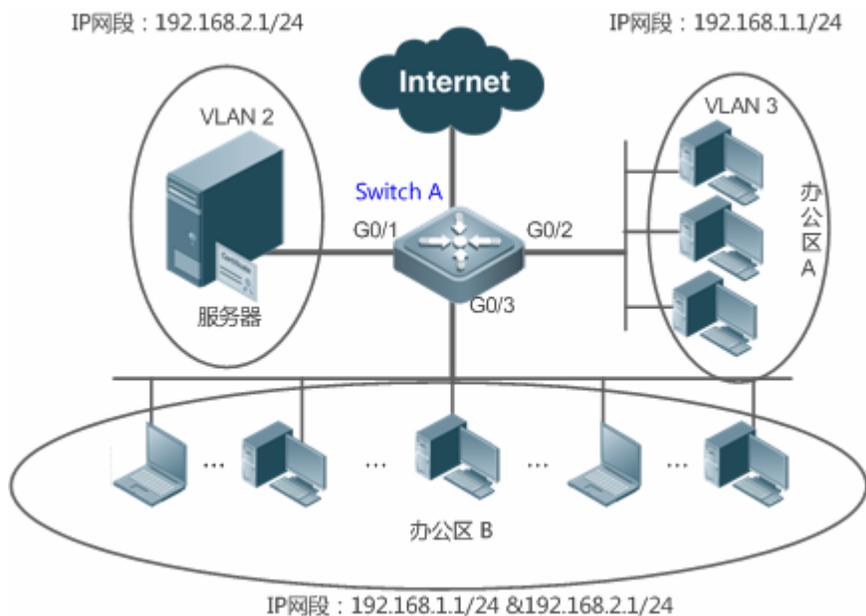
应用场景

如下图所示，办公区A和办公区B通过HUB与三层设备Switch A相连：办公区A内，分布着固定网段的办公用户，统一基于端口划分属于一个VLAN；办公区B内，分布着两个网段的办公用户，无法基于固定端口划分VLAN。

主要需求如下：

对于办公区 B 内的 PC 用户，Switch A 可以根据报文所属的 IP 网段确定其 VLAN 范围。

图 6-2



【注释】 Switch A为交换机设备,端口 G0/1 为 Access 口,所属 VLAN 2 ;G0/2 为 Access 口,所属 VLAN 3 ;G0/3 为 Hybrid 口。

功能部署

- 全局配置子网 VLAN(本例将 IP 网段为 192.168.1.1/24 划分属于 VLAN 3 ,IP 网段为 192.168.2.1/24 划分属于 VLAN 2) , 并在上链口 (本例为 Gi 0/3) 使能子网功能

⚠ 配置子网 VLAN 的端口仅针对 Trunk 口和 Hybrid 口生效。

6.3 功能详解

基本概念

Protocol VLAN

Protocol VLAN 技术就是基于报文协议类型的 VLAN 分类技术 其可以将某一协议类型的空 VLAN ID 报文都划分到同一 VLAN。

设备端口接收到的报文, 都需要进行 VLAN 分类, 使报文属于唯一的一个 VLAN, 有以下三种可能:

- 如果报文是空 VLAN ID 报文 (UNTAG 或 Priority 报文) , 而设备仅支持基于端口的 VLAN 分类的话, 报文所添加 TAG 的 VLAN ID 将是输入端口的 PVID。
- 如果报文是空 VLAN ID 报文 (UNTAG 或 Priority 报文) , 而设备支持基于报文协议类型的 VLAN 分类的话, 报文所添加 TAG 的 VLAN ID 将会从输入端口上的协议组配置相对应的 VLAN ID 集中选取, 而如果报文的协议类型与输入端口上的所有协议组配置都不相符的话, 将按照基于端口的 VLAN 分类来分配 VLAN ID。
- 如果报文是 TAG 报文, 其所属 VLAN 分类由 TAG 中的 VLAN ID 决定。

其中子网 VLAN 只有全局配置 , 即端口上配置只有开启/关闭 Protocol VLAN 功能。协议 VLAN 全局配置报文类型, 接口上配

置对应报文类型分配 VLAN。如下所示：

- 如果输入报文为空 VLAN ID 报文，且输入报文的 IP 地址匹配用户配置的 IP 地址的话，该报文将被划分到用户配置的用户网 VLAN 内。
- 如果输入报文为空 VLAN ID 报文，且输入报文的报文类型和以太网类型，匹配用户配置在输入端口上的报文类型和以太网类型的话，该报文将被划分到用户配置的协议 VLAN 内。

Protocol VLAN 优先级

子网 VLAN 优先级高于协议 VLAN，即同时配置了子网 VLAN 和协议 VLAN，且输入报文同时符合两者的话，将是子网 VLAN 分配起作用。

功能特性

功能特性	作用
根据报文类型自动划分VLAN	将网络中提供的服务类型与 VLAN 相绑定，或将指定 IP 网段发出的报文在指定的 VLAN 中传送，方便管理和维护。

6.3.1 根据报文类型自动划分VLAN

工作原理

- 设置规则到硬件，并在端口使能规则，且只有在端口上使能后，才真正生效；这些规则包括报文类型，报文的 IP 地址；当端口上收到符合规则的，不带 VLAN 标记的数据流报文，将其自动划分到规则中指定的 VLAN 中传送。端口关闭功能则不带 VLAN 标记的数据流报文遵守端口上配置，都划分到 native VLAN 中。

相关配置

6.4 配置详解

配置项	配置建议 & 相关命令
配置协议VLAN功能	 必须配置。用于使能 Protocol-VLAN 基于报文类型和以太网类型分类功能
	protocol-vlan profile num frame-type [type] ether-type [type] 配置报文类型和以太网类型的 profile
	protocol-vlan profile num ether-type [type] (某些型号不支持 frame 识别)配置以太网类型的 profile
	protocol-vlan profile num vlan vid (接口模式下)端口上应用协议 VLAN
配置子网VLAN功能	 必须配置。用于使能 Protocol-VLAN 基于 ip 地址 VLAN 分类功能

protocol-vlan ipv4 address mask address vlan vid	配置 IP 地址、子网掩码以及 VLAN 分类
protocol-vlan ipv4	(接口模式下)端口上使能子网 VLAN

6.4.1 配置协议VLAN功能

配置效果

将网络中提供的服务类型与 VLAN 相绑定，方便管理和维护。

注意事项

- 用户最好在配置好 VLAN、端口的 Trunk、Hybrid、Access 和 AP 属性后，再配置 Protocol VLAN；
- 如果用户在 Trunk 或 Hybrid 口上配置了 Protocol VLAN，那么用户需要报文 Trunk 和 Hybrid 口的许可 VLAN 列表包含 Protocol VLAN 相关的所有 VLAN。

配置方法

↘ 全局配置协议 VLAN

- 必须配置。
- 只有全局配置上，接口上才能应用对应协议 VLAN。

【命令格式】 **protocol-vlan profile num frame-type type ether-type type**

【参数说明】 *num* : profile 索引

type : 报文类型和以太网类型

【缺省配置】 默认关闭

【命令模式】 全局模式

【使用指导】 只有在 Protocol-VLAN 全局配置存在的情况，Protocol-VLAN 接口上才能配置协议 VLAN。删除全局配置时，会删除对应索引的所有接口协议 VLAN 配置。

↘ 切换端口模式为 trunk / hybrid 模式

- 必须配置，协议 VLAN 功能只有在 trunk / hybrid 模式的端口上才生效。

↘ 端口上使能协议 VLAN

- 必须配置，默认为关闭应用。
- 只有接口上应用，才真正使能协议 VLAN。

【命令格式】 **protocol-vlan profile num vlan vid**

【参数说明】 *num* : profile 索引

vid : VLAN ID, 1-产品支持的最大 VLAN

【缺省配置】 默认关闭

- 【命令模式】 接口模式
 【使用指导】 接口必须为 trunk / hybrid 模式。

检验方法

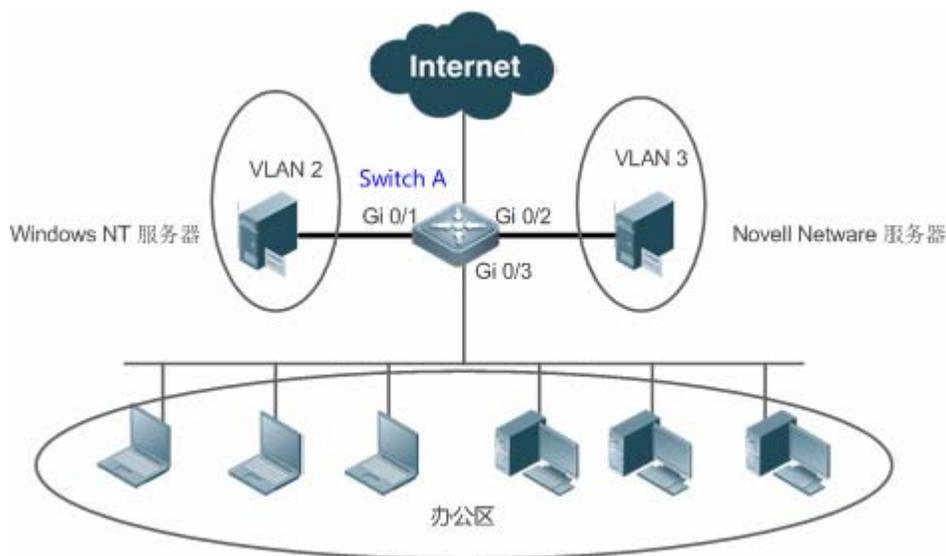
show protocol-vlan profile 查看配置信息。

配置举例

在拓扑环境中开启协议 VLAN 功能

【网络环境】 ●

图 6-3



-
- 【配置方法】 ●
- 在交换机 A 配置用户通信的 VLAN 2-3。
 - 在交换机 A 全局配置协议 VLAN（本例将支持 IP 协议的报文对应 Profile 1，支持 IPX 协议的报文对应 Profile 2）；并在上链口（本例为 Gi 0/3）使能协议 VLAN 功能，完成协议与 VLAN 关联（本例将 Profile 1 关联 VLAN 2，Profile 2 关联 VLAN 3）。
 - 端口 Gi 0/1 为 Access 口，所属 VLAN 2；Gi 0/2 为 Access 口，所属 VLAN 3；Gi 0/3 为 Hybrid 口。必须保证 Hybrid 口的 untagged VLAN 许可列表包含用户通信的 VLAN。

A

1：创建用户网络通信的 VLAN 2-3。

```
A# configure terminal
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
A(config)# vlan range 2-3
```

2：配置端口模式

```
A(config)#interface gigabitEthernet 0/1
```

```
A(config-if-GigabitEthernet 0/1)#switchport
```

```

A(config-if-GigabitEthernet 0/1)#switchport access vlan 2
A(config-if-GigabitEthernet 0/1)#exit
A(config)#interface gigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)#switchport
A(config-if-GigabitEthernet 0/2)#switchport access vlan 3
A(config-if-GigabitEthernet 0/2)#exit
A(config)# interface gigabitEthernet 0/3
A(config-if-GigabitEthernet 0/3)#switchport
A(config-if-GigabitEthernet 0/3)# switchport mode hybrid
A(config-if-GigabitEthernet 0/3)# switchport hybrid allowed vlan untagged 2-3

```

3：全局配置协议 VLAN。

IP、IPX 协议配置相应的 Profile 1、2。（此处假设报文使用 EthernetII 封装，IP 和 IPX 分别对应的以太网类型为 0X0800、0X8137）

```

A(config)#protocol-vlan profile 1 frame-type ETHERII ether-type 0x0800
A(config)#protocol-vlan profile 2 frame-type ETHERII ether-type 0x8137

```

4：将 Profile 1、2 应用到端口 Gi 0/3 上，划分为 VLAN 2 和 VLAN 3。

```

A(config)# interface gigabitEthernet 0/3
A(config-if-GigabitEthernet 0/3) #protocol-vlan profile 1 vlan 2
A(config-if-GigabitEthernet 0/3) #protocol-vlan profile 2 vlan 3

```

【检验方法】

- 查看设备上 Protocol VLAN 配置是否正确。

A

```
A(config)#show protocol-vlan profile
```

profile	frame-type	ether-type/DSAP+SSAP	interface	vlan
1	ETHERII	0x0800	Gi0/3	2
2	ETHERII	0x8137	Gi0/3	3

常见配置错误

- 设备连接的端口不是 Trunk/Hybrid 模式。
- 设备连接的端口许可 VLAN 列表不包含用户通信的 VLAN。
- 端口未使能协议 VLAN 功能

6.4.2 配置子网VLAN功能

配置效果

将指定网段或 IP 地址发出的报文在指定的 VLAN 中传送。

注意事项

- 用户最好在配置好 VLAN、端口的 Trunk、Hybrid 、 Access 和 AP 属性后，再配置 Protocol VLAN ；
- 如果用户在 Trunk 或 Hybrid 口上配置了 Protocol VLAN ， 那么用户需要报文 Trunk 和 Hybrid 口的许可 VLAN 列表包含 Protocol VLAN 相关的所有 VLAN。

配置方法

↘ 全局配置子网 VLAN

- 必须配置。
- 只有全局配置上，接口上才能应用对应子网 VLAN。

【命令格式】 **protocol-vlan ipv4 address mask address vlan vid**

【参数说明】 *address* : ip 地址

vid : VLAN ID , 1-产品支持的最大 VLAN

【缺省配置】 默认关闭

【命令模式】 全局模式

【使用指导】 在Protocol-VLAN未全局使能的状态下，接口也可配置使能，但只有在Protocol-VLAN全局配置存在的情况，Protocol-VLAN配置子网VLAN才有作用。

↘ 切换端口模式为 trunk / hybrid 模式

- 必须配置，子网 VLAN 功能只有在 trunk / hybrid 模式的端口上才生效。

↘ 端口上使能子网 VLAN

- 必须配置，默认为关闭应用。
- 只有接口上应用，才真正使能子网 VLAN 功能。

【命令格式】 **protocol-vlan ipv4**

【参数说明】 -

【缺省配置】 默认关闭

【命令模式】 接口模式

【使用指导】 接口必须为 trunk / hybrid 模式。

检验方法

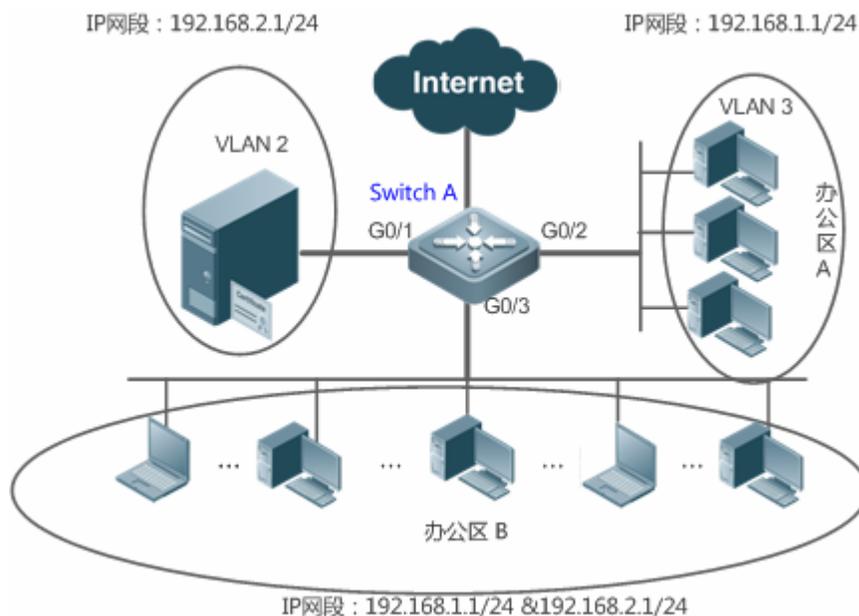
show protocol-vlan ipv4 查看配置信息。

配置举例

↘ 在拓扑环境中开启子网 VLAN 功能

【网络环境】

图 6-4



【配置方法】

- 在交换机 A 配置用户通信的 VLAN 2-3
- 在交换机 A 全局配置子网 VLAN (本例将 IP 网段为 192.168.1.1/24 划分属于 VLAN 3, IP 网段为 192.168.2.1/24 划分属于 VLAN 2), 并在上链口 (本例为 Gi 0/3) 使能子网 VLAN 功能
- 端口 Gi 0/1 为 Access 口, 所属 VLAN 2; Gi 0/2 为 Access 口, 所属 VLAN 3; Gi 0/3 为 Hybrid 口。必须保证 Hybrid 口的 untagged VLAN 许可列表包含用户通信的 VLAN。

A

1: 创建用户网络通信的 VLAN 2-3。

```
A# configure terminal
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
A(config)# vlan range 2-3
```

2: 配置端口模式

```
A(config)#interface gigabitEthernet 0/1
```

```
A(config-if-GigabitEthernet 0/1)#switchport
```

```
A(config-if-GigabitEthernet 0/1)#switchport access vlan 2
```

```
A(config-if-GigabitEthernet 0/1)#exit
```

```
A(config)#interface gigabitEthernet 0/2
```

```
A(config-if-GigabitEthernet 0/2)#switchport
```

```
A(config-if-GigabitEthernet 0/2)#switchport access vlan 3
```

```
A(config-if-GigabitEthernet 0/2)#exit
```

```
A(config)# interface gigabitEthernet 0/3
```

```
A(config-if-GigabitEthernet 0/3)#switchport
```

```
A(config-if-GigabitEthernet 0/3)# switchport mode hybrid
```

```
A(config-if-GigabitEthernet 0/3)# switchport hybrid allowed vlan untagged 2-3
```

2: 全局配置子网 VLAN。

```
A(config)# protocol-vlan ipv4 192.168.1.0 mask 255.255.255.0 vlan 3
```

```
A(config)# protocol-vlan ipv4 192.168.2.0 mask 255.255.255.0 vlan 2
4: 接口上使能子网 VLAN, 默认关闭。
A(config-if-GigabitEthernet 0/1)# protocol-vlan ipv4
```

【检验方法】 ● 查看设备上 Protocol VLAN 配置是否正确。

A

```
A# show protocol-vlan ipv4
ip          mask          vlan
-----
192.168.1.0 255.255.255.0 3
192.168.2.0 255.255.255.0 2

interface   ipv4 status
-----
Gi0/3       enable
```

常见配置错误

- 设备连接的端口不是 Trunk/Hybrid 模式。
- 设备连接的端口许可 VLAN 列表不包含用户通信的 VLAN。
- 端口未使能子网 VLAN

6.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
显示 Protocol VLAN 的内容	show protocol-vlan

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 Protocol VLAN debug 开关	debug bridge protvlan

7 Private VLAN

7.1 概述

私有 VLAN(Private VLAN)将一个 VLAN 的二层广播域划分成多个子域,每个子域都由一个私有 VLAN 对组成:主 VLAN(Primary VLAN)和辅助 VLAN(Secondary VLAN)。

一个私有 VLAN 域可以有多个私有 VLAN 对,每一个私有 VLAN 对代表一个子域。在一个私有 VLAN 域中所有的私有 VLAN 对共享同一个主 VLAN。每个子域的辅助 VLAN ID 不同。

服务提供商如果给每个用户一个 VLAN,则由于一台设备支持的 VLAN 数最大只有 4096 而限制了服务提供商能支持的用户数;在三层设备上,每个 VLAN 被分配一个子网地址或一系列地址,这种情况导致 IP 地址的浪费。Private VLAN 技术可以很好的同时解决以上两种问题,后续说明文中将 Private VLAN 简称为 PVLAN。

7.2 典型应用

典型应用	场景描述
PVLAN跨设备二层应用	企业内用户之间可以进行通信,企业间用户通信隔离。
PVLAN单台设备三层应用	所有企业用户共享一个网关地址,可以与外网通信。

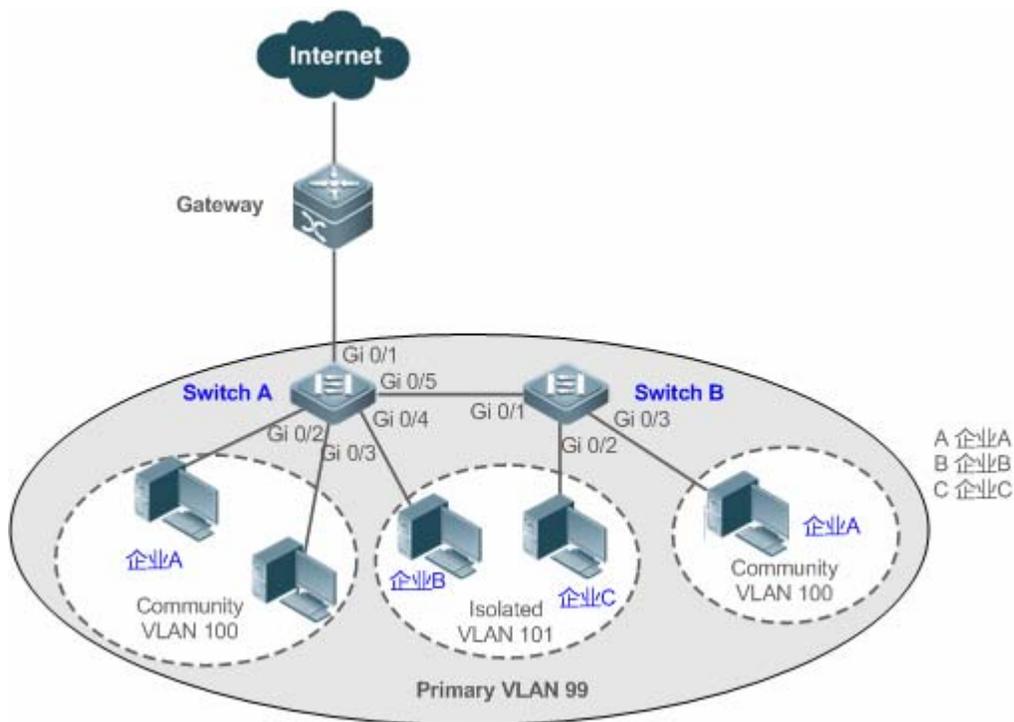
7.2.1 PVLAN跨设备二层应用

应用场景

如下图所示,在主机托管业务运营网络中,各企业用户通过设备 Switch A、Switch B 接入网络。主要需求如下:

- 企业内用户之间可以进行通信,企业间用户通信隔离
- 所有企业用户共享一个网关地址,可以与外网通信。

图 7-1



【注释】 Switch A、B 为接入交换机。

跨设备运行 PVLAN，需要将相连的端口配置为 Trunk Port，将 A 的 Gi 0/5 和 B 的 Gi 0/1 均配置为 Trunk Port。

与网关相连的 A 的 Gi 0/1 需要配置为 Promiscuous Port；

网关设备 Gi 0/1 口可以配置为 Trunk Port 或者 Hybrid Port，且 Native VLAN 是 PVLAN 的 Primary VLAN。

功能部属

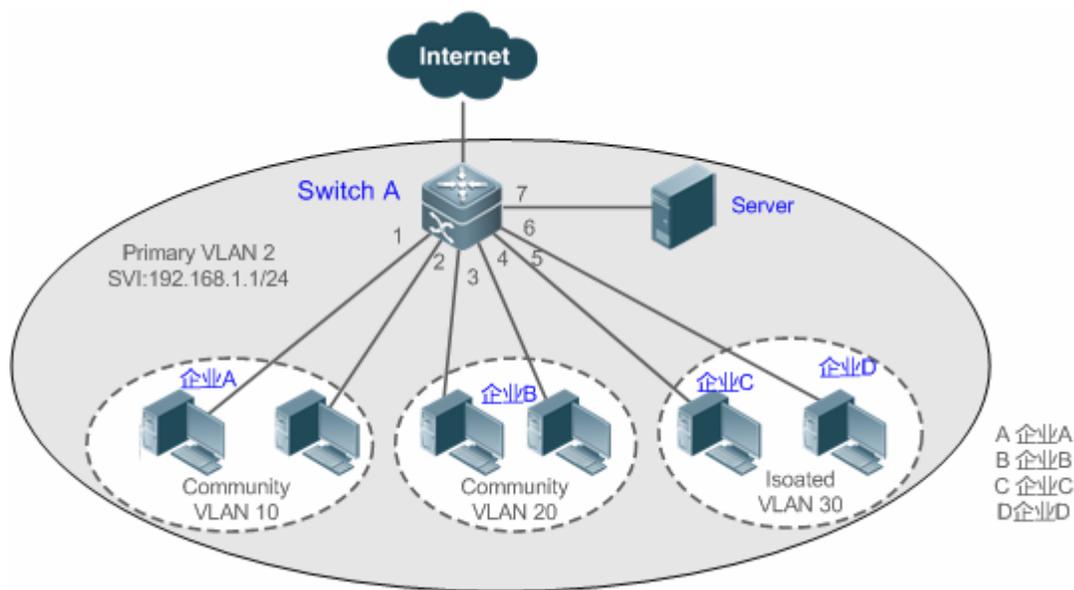
- 将所有企业配置属于同一个 PVLAN（本例为 Primary VLAN 99），所有企业用户均通过该 VLAN 共享一个三层接口，实现外网通信。
- 如果企业内有多用户，可以将各企业划分属于不同的 Community VLAN，即将相关企业用户连接的端口配置为 Community VLAN 的 host Port，实现企业内用户互相通信，企业间用户通信隔离。
- 如果企业内仅有一个用户，可以将这些企业的该用户划分属于一个 Isolated VLAN 的 host Port，实现企业间用户通信隔离。

7.2.2 PVLAN单台设备三层应用

如下图所示，在主机托管业务运营网络中，各企业用户通过三层设备 Switch A 接入网络。主要需求如下：

- 企业内用户之间可以通信，企业间用户通信隔离。
- 所有企业用户都可以访问服务器
- 所有企业用户共享一个网关地址，可以与外网通信。

图 7-2



【注释】 A 为网关交换机。

单台设备连接时，为了各企业用户能与服务器通信，连接服务器的端口 Gi 0/7 配置为 Promiscuous Port。为了用户能与外网通信，需要将 Primary VLAN 和 Secondary VLAN 进行三层映射。

功能部属

- 将直连服务器的端口设置为 Promiscuous Port，所有企业用户都可以通过 Promiscuous Port 和服务器通信。
- 在三层设备(本例为 Switch A)配置 PVLAN 的网关地址(本例配置 VLAN 2 的 SVI 为 192.168.1.1/24)，并配置 Primary VLAN 和 Secondary VLAN 的三层接口映射关系，所有企业用户可以通过这个网关地址与外网通信。

7.3 功能详解

基本概念

▾ PVLAN

PVLAN 中包含了三种类型的 VLAN：主 VLAN(Primary vlan)、隔离 VLAN(Isolated VLAN)、群体 VLAN(Community VLAN)。一个私有 VLAN 域中只有一个主 VLAN，辅助 VLAN 实现同一个私有 VLAN 域中的二层隔离，有两种类型的辅助 VLAN。

▾ 隔离 VLAN

- 隔离 VLAN(Isolated VLAN)：同一个隔离 VLAN 中的端口不能互相进行二层通信。一个私有 VLAN 域中只有一个隔离 VLAN。

▾ 群体 VLAN

- 群体 VLAN(Community VLAN)：同一个群体 VLAN 中的端口可以互相进行二层通信，但不能与其它群体 VLAN 中的端口进行二层通信。一个私有 VLAN 域中可以有多个群体 VLAN。

↘ PVLAN 的二层关联

PVLAN 的三种 VLAN 必须进行二层关联才能构成 PVLAN 对，主 VLAN 才有了指定的辅助 VLAN，辅助 VLAN 才有了指定的主 VLAN。主 VLAN 和辅助 VLAN 之间是一对多的关系。

↘ PVLAN 的三层关联

PVLAN 中只有主 VLAN 可以创建三层口 SVI，辅助 VLAN 必须通过与主 VLAN 三层关联后，辅助 VLAN 内的用户才能进行三层通信，否则仅能二层通信。

↘ 隔离端口(Isolated Port)

隔离 VLAN 中的端口，只能和混杂口通信。隔离端口接受到的报文可允许转发到 Trunk Port，但 Trunk Port 接收到 vid 是 Isolated VLAN 的报文不能向隔离端口转发。

↘ 群体端口(Community Port)

属于 Community VLAN 的端口，同一个 Community VLAN 的群体端口可以互相通讯，也可以与混杂口通讯。不能与其它群体 VLAN 中的群体端口及隔离 VLAN 中的隔离端口通讯。

↘ 混杂端口 (Promiscuous Port)

属于主 VLAN 的端口，可以和任意端口通信，包括同一个 PVLAN 域中辅助 VLAN 的隔离端口和群体端口。

- PVLAN 中，只有主 VLAN 可以创建 SVI 接口，辅助 VLAN 不可以创建 SVI。
- PVLAN 中的端口可以作为镜像源端口，不可以作为镜像目的端口。

功能特性

功能特性	作用
PVLAN 二层隔离和节省 IP 地址	通过配置各种 PVLAN 类型的端口，实现 VLAN 中间用户的互通和隔离。 主 VLAN 和辅助 VLAN 二层映射后，只能支持一些二层通信。如果要进行三层通信，辅助 VLAN 的用户需要借助主 VLAN 的 SVI 口来进行三层通信。

7.3.1 PVLAN 二层隔离和节省 IP 地址

通过将用户加入 PVLAN 的各个子域，可以隔离企业间、企业用户间的通信。

工作原理

通过配置 PVLAN、PVLAN 的主 VLAN 和子 VLAN 二三层关联、以及用户、外网设备、服务器等连接的端口设置为 PVLAN 的各种端口，实现各子域的划分，各子域用户与外网和服务器的通信。

↘ 各种端口类型间的报文转发关系

输出端口 \ 输入端口	混杂端口	隔离端口	群体端口	TRUNK 端口 (同 VLAN 内)
混杂端口	通	通	通	通

隔离端口	通	不通	不通	通
群体端口	通	不通	通	通
TRUNK 端口 (同 VLAN 内)	通	不通	通	通

各种端口类型间的报文转发后 VLAN TAG 变化关系

输出端口	混杂端口	隔离端口	群体端口	TRUNK 端口 (同 VLAN 内)
输入端口				
混杂端口	不变	不变	不变	加上主 VLAN ID TAG
隔离端口	不变	NA	NA	加上隔离 VLAN ID TAG
群体端口	不变	NA	不变	加上群体 VLAN ID TAG
TRUNK 端口 (同 VLAN 内)	去掉 VLAN TAG	NA	去掉 VLAN TAG	不变
交换机 CPU	Untag	Untag	Untag	加上主 VLAN ID TAG

相关配置

配置 PVLAN 及 PVLAN 子域

- 缺省情况下，所有 VLAN 都是普通 VLAN。
- 使用 `private-vlan { community | isolated | primary }` 命令可以将 VLAN 配置为 PVLAN 的主 VLAN 和辅助 VLAN。

PVLAN 二层关联

- 缺省情况下，配置各种 PVLAN 后，主 VLAN 和辅助 VLAN 之间没有任何关系，都是一个个单独的个体，只有进行二层关联后主 VLAN 才有辅助 VLAN，辅助 VLAN 才有主 VLAN。
- 使用 `private-vlan association { svlist | add svlist | remove svlist }` 可以增加或取消 PVLAN 的主 VLAN 和辅助 VLAN 的二层关联，二层关联后才能构成一个 PVLAN 子域，一旦取消二层关联对应的子域也就不存在了。另外不进行二层关联，后续的各种隔离端口和混杂口配置关联 PVLAN 对时会失败或取消已经配置的端口关联 VLAN。

PVLAN 三层关联

- 缺省情况下，仅 PVLAN 的主 VLAN 可以配置三层口 SVI，辅助 VLAN 不能进行三层通信。
- 如果 PVLAN 辅助 VLAN 内的用户要进行三层通信，需要借助于主 VLAN 的 SVI 来收发包。

- 使用 **private-vlan mapping** { *svlist* | **add** *svlist* | **remove** *svlist* } 命令可以增加或取消 PVLAN 的主 VLAN 和辅助 VLAN 之间的三层关联。三层关联后，辅助 VLAN 内的用户才可以与外网进行三层通信。关闭后辅助 VLAN 内的用户不能进行三层通信。

7.4 配置详解

配置项	配置建议 & 相关命令		
配置PVLAN基本功能	<ul style="list-style-type: none"> ● 必须配置。用于实现主 VLAN 和辅助 VLAN 的配置。 		
	<table border="1"> <tr> <td>private-vlan { community isolated primary }</td> <td>配置 PVLAN 类型</td> </tr> </table>	private-vlan { community isolated primary }	配置 PVLAN 类型
	private-vlan { community isolated primary }	配置 PVLAN 类型	
	<ul style="list-style-type: none"> ● 必选配置，用于实现 PVLAN 的主 VLAN 和辅助 VLAN 的二层关联，构成 PVLAN 对 		
	<table border="1"> <tr> <td>private-vlan association { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }</td> <td>主 VLAN 和辅助 VLAN 二层关联，这样才构成 PVLAN 对。</td> </tr> </table>	private-vlan association { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }	主 VLAN 和辅助 VLAN 二层关联，这样才构成 PVLAN 对。
	private-vlan association { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }	主 VLAN 和辅助 VLAN 二层关联，这样才构成 PVLAN 对。	
	<ul style="list-style-type: none"> ● 可选配置，用于将用户划分到隔离 VLAN 或者群体 VLAN 		
	<table border="1"> <tr> <td>switchport mode private-vlan host</td> <td>配置成私有 VLAN host 端口</td> </tr> </table>	switchport mode private-vlan host	配置成私有 VLAN host 端口
	switchport mode private-vlan host	配置成私有 VLAN host 端口	
	<table border="1"> <tr> <td>switchport private-vlan host-association <i>p_vid</i> <i>s_vid</i></td> <td>关联二层端口与 PVLAN，将端口划分到相关子域</td> </tr> </table>	switchport private-vlan host-association <i>p_vid</i> <i>s_vid</i>	关联二层端口与 PVLAN，将端口划分到相关子域
switchport private-vlan host-association <i>p_vid</i> <i>s_vid</i>	关联二层端口与 PVLAN，将端口划分到相关子域		
<ul style="list-style-type: none"> ● 可选配置，将端口配置为混杂端口 			
<table border="1"> <tr> <td>Switchport mode private-vlan promiscuous</td> <td>配置成私有 VLAN 混杂端口</td> </tr> </table>	Switchport mode private-vlan promiscuous	配置成私有 VLAN 混杂端口	
Switchport mode private-vlan promiscuous	配置成私有 VLAN 混杂端口		
<table border="1"> <tr> <td>switchport private-vlan mapping <i>p_vid</i> { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }</td> <td>配置私有 VLAN 混杂端口所在的 Primary VLAN 以 Secondary VLAN 列表。配置之后对应 PVLAN 的报文才能通过这个端口通信。</td> </tr> </table>	switchport private-vlan mapping <i>p_vid</i> { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }	配置私有 VLAN 混杂端口所在的 Primary VLAN 以 Secondary VLAN 列表。配置之后对应 PVLAN 的报文才能通过这个端口通信。	
switchport private-vlan mapping <i>p_vid</i> { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }	配置私有 VLAN 混杂端口所在的 Primary VLAN 以 Secondary VLAN 列表。配置之后对应 PVLAN 的报文才能通过这个端口通信。		
<ul style="list-style-type: none"> ● 可选配置。用于实现辅助 VLAN 的用户的三层通信。 			
<table border="1"> <tr> <td>private-vlan mapping { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }</td> <td>创建 PVLAN 并进行二层关联后，配置主 VLAN 的 SVI，并将主 VLAN 和辅助 VLAN 进行三层关联，子 VLAN 可以借助主 VLAN 的 SVI 进行三层通信。</td> </tr> </table>	private-vlan mapping { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }	创建 PVLAN 并进行二层关联后，配置主 VLAN 的 SVI，并将主 VLAN 和辅助 VLAN 进行三层关联，子 VLAN 可以借助主 VLAN 的 SVI 进行三层通信。	
private-vlan mapping { <i>svlist</i> add <i>svlist</i> remove <i>svlist</i> }	创建 PVLAN 并进行二层关联后，配置主 VLAN 的 SVI，并将主 VLAN 和辅助 VLAN 进行三层关联，子 VLAN 可以借助主 VLAN 的 SVI 进行三层通信。		

7.4.1 配置PVLAN基本功能

配置效果

- 构成 PVLAN 子域，实现企业和企业用户间的隔离；
- 多个辅助 VLAN 三层映射到主 VLAN，多个 VLAN 利用同一个 IP 网关，节省 IP 地址。

注意事项

- 在配置主 VLAN 和辅助 VLAN 后，必须进行二层关联才能构成 PVLAN 子域。
- 必须将用户连接的端口配置为特定 PVLAN 端口类型，用户才能加入对应的子域，才能实现真正的用户隔离。
- 同时连接外网和服务器的接口必须配置成混杂口，上下行报文转发才能正常。
- 辅助 VLAN 只有与主 VLAN 进行三层映射后才能借助主 VLAN 的 SVI 进行三层通信。

配置方法

配置 PVLAN

- 必须配置。
- 需要配置主 VLAN 和辅助 VLAN，这两种 VLAN 不能独立存在。

PVLAN 的二层关联

- 必须配置。
- PVLAN 的主 VLAN 和辅助 VLAN 进行二层关联后才能构成 PVLAN 子域，才能配置相应的隔离口、群体口以及三层关联等。

PVLAN 的三层关联

- 如果辅助 VLAN 域内的用户需要进行三层通信，需要给主 VLAN 配置一个三层口 SVI，然后在 SVI 口上配置主 VLAN 和辅助 VLAN 的三层。

隔离端口和群体端口

- 配置 PVLAN 的主 VLAN 和辅助 VLAN 并进行二层关联后，还需要对用户连接的设备端口进行划分，才能真正对用户所在的子域进行划分。
- 如果企业内仅有一个用户，可以考虑将用户连接的端口设置成隔离端口。
- 如果企业内有多个用户，可以将用户连接的端口设置成群体端口。

混杂端口

- 从功能详解章节各种端口收发报文的规则表格可见 PVLAN 的单一端口类型不能保证上下行报文转发对称，为了保证用户能正常访问外网和服务器，一般需要将连接外网和服务器的端口配置成混杂口。

检验方法

使 PVLAN 端口内的用户能按照 PVLAN 的端口转发规则进行收发报文，达到隔离作用；通过三层关联，同一个 PVLAN 内的主 VLAN 和辅助 VLAN 共用一个网关 IP 进行三层通信。

相关命令

配置 PVLAN

【命令格式】 **private-vlan { community | isolated | primary }**

【参数说明】 **community** : 指定 VLAN 类型为群体 VLAN。
isolated : 指定 VLAN 类型为隔离 VLAN。
primary : 指定 VLAN 类型为 PVLAN 对的主 VLAN。

【命令模式】 VLAN 模式

【使用指导】 此命令用来指定 PVLAN 的主 VLAN 和辅助 VLAN。

主 VLAN 和辅助 VLAN 二层关联

【命令格式】 **private-vlan association { svlist | add svlist | remove svlist }**

【参数说明】 **svlist** : 指定需要关联或解关联的辅助 VLAN 列表。
add svlist : 增加关联的辅助 VLAN。
remove svlist : 解除 **svlist** 与主 VLAN 的关联。

【命令模式】 PVLAN 的主 VLAN 模式

【使用指导】 此命令用来进行主 VLAN 和辅助 VLAN 的二层关联，构成 PVLAN 对。
每个主 VLAN 只能关联一个 Isolated VLAN，但可以关联多个 Community VLAN。

配置隔离端口和群体端口

【命令格式】 **switchport mode private-vlan host**

switchport private-vlan host-association p_vid s_vid

【参数说明】 **p_vid** : PVLAN 对中的主 VLAN id。
s_vid : PVLAN 对中的辅助 VLAN id ,如果为 Isolated VLAN ,则该端口为隔离端口 ;如果为 Community VLAN ,则该端口为群体端口。

【命令模式】 两个命令都在接口模式

【使用指导】 需要上面两条命令来完成，并且配置为隔离端口或混杂端口前，端口的模式必须先配置成 host 口模式。
是配置成隔离端口还是群体端口，视 **s_vid** 参数而定。
p_vid 和 **s_vid** 必须是有二层关联的 PVLAN 对。
一个 host 口仅能关联一对 PVLAN 对。

混杂端口

【命令格式】 **switchport mode private-vlan promiscuous**

switchport private-vlan mapping p_vid { svlist | add svlist | remove svlist }

【参数说明】 **p_vid** : PVLAN 对的主 VLAN。
svlist : 混杂端口关联的辅助 VLAN，必须和 **p_vid** 是二层关联的。
add svlist : 增加端口关联的辅助 VLAN。
remove svlist : 取消端口关联的辅助 VLAN

【命令模式】 接口模式

【使用指导】 必须向将端口模式配置成混杂模式。
端口必须与 PVLAN 对进行关联，否则不起作用。
一个混杂口可以关联一个主 VLAN 内的过个 PVLAN 对，但不能关联多个主 VLAN。

主 VLAN 和辅助 VLAN 的三层关联

【命令格式】 **private-vlan mapping** { *svlist* | **add** *svlist* | **remove** *svlist* }

【参数说明】 *svlist* : 三层映射的辅助 VLAN 列表。

add *svlist* : 增加三层口关联的辅助 VLAN。

remove *svlist* : 取消三层口关联的辅助 VLAN

【命令模式】 主 VLAN 的接口模式

【使用指导】 必须先为主 VLAN 配置三层口 SVI。

仅主 VLAN 可以配置三层口。

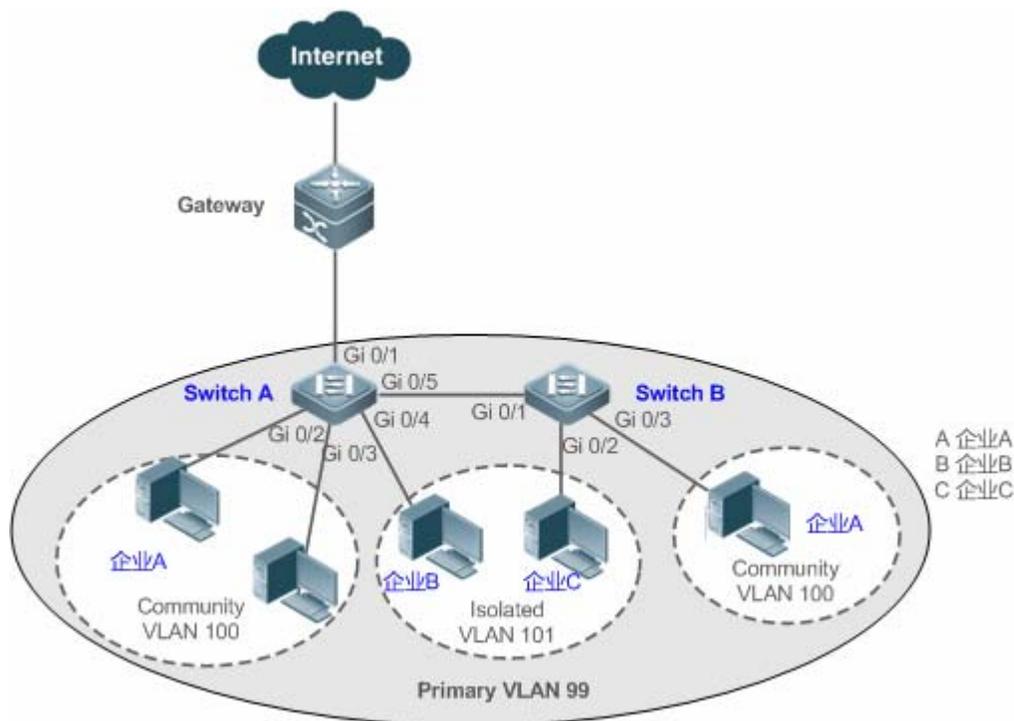
关联的辅助 VLAN 必须和主 VLAN 是二层关联的。

配置举例

i 以下配置举例，介绍与跨设备二层应用的配置。

在跨二层设备上应用 PVLAN

【图 7-3】



- 【配置方法】
- 将所有企业配置属于同一个 PVLAN (本例为 Primary VLAN 99)，所有企业用户均通过该 VLAN 共享一个三层接口，实现外网通信。
 - 如果企业内有多个用户，可以将各企业划分属于不同的 Community VLAN (本例将企业 A 划分属于 Community VLAN 100)，实现企业内用户互相通信，企业间用户通信隔离。
 - 如果企业内仅有一个用户，可以将这些企业划分属于同一个 Isolated VLAN (本例将企业 B 和 C 划分属于 Isolated VLAN 101)，实现企业间用户通信隔离。

A

```
SwitchA#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
```

```
SwitchA(config)#vlan 99
SwitchA(config-vlan)#private-vlan primary
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 100
SwitchA(config-vlan)#private-vlan community
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 101
SwitchA(config-vlan)#private-vlan isolated
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 99
SwitchA(config-vlan)#private-vlan association 100-101
SwitchA(config-vlan)#exit
SwitchA(config)#interface range gigabitEthernet 0/2-3
SwitchA(config-if-range)#switchport mode private-vlan host
SwitchA(config-if-range)#switchport private-vlan host-association 99 100
SwitchA(config-if-range)#exit
SwitchA(config)#interface gigabitEthernet 0/4
SwitchA(config-if-GigabitEthernet 0/4)#switchport mode private-vlan host
SwitchA(config-if-GigabitEthernet 0/4)#switchport private-vlan host-association 99 101
SwitchA(config)#interface gigabitEthernet 0/5
SwitchA(config-if-GigabitEthernet 0/5)#switchport mode trunk
SwitchA(config-if-GigabitEthernet 0/5)#exit
```

B

```
SwitchB#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
SwitchB(config)#vlan 99
SwitchB(config-vlan)#private-vlan primary
SwitchB(config-vlan)#exit
SwitchB(config)#vlan 100
SwitchB(config-vlan)#private-vlan community
SwitchB(config-vlan)#exit
SwitchB(config)#vlan 101
SwitchB(config-vlan)#private-vlan isolated
SwitchB(config-vlan)#exit
SwitchB(config)#vlan 99
SwitchB(config-vlan)#private-vlan association 100-101
SwitchB(config-vlan)#exit
SwitchB(config)#interface gigabitEthernet 0/2
SwitchB(config-if-GigabitEthernet 0/2)#switchport mode private-vlan host
SwitchB(config-if-GigabitEthernet 0/2)# switchport private-vlan host-association 99 101
SwitchB(config-if-GigabitEthernet 0/2)#exit
SwitchB(config)#interface gigabitEthernet 0/3
```

```
SwitchB(config-if-GigabitEthernet 0/3)#switchport mode private-vlan host
SwitchB(config-if-GigabitEthernet 0/3)# switchport private-vlan host-association 99 100
SwitchB(config-if-GigabitEthernet 0/3)#exit
SwitchB(config)#interface gigabitEthernet 0/1
SwitchB(config-if-GigabitEthernet 0/1)#switchport mode trunk
SwitchB(config-if-GigabitEthernet 0/1)#exit
```

【检验方法】 检查 VLAN 和端口上配置是否正确。根据功能详解中的转发规则查看报文转发是否正确。

A

```
SwitchA#show running-config
!
vlan 99
  private-vlan primary
  private-vlan association add 100-101
!
vlan 100
  private-vlan community
!
vlan 101
  private-vlan isolated
!
interface GigabitEthernet 0/1
  switchport mode private-vlan promiscuous
  switchport private-vlan mapping 99 add 100-101
!
interface GigabitEthernet 0/2
  switchport mode private-vlan host
  switchport private-vlan host-association 99 100
!
interface GigabitEthernet 0/3
  switchport mode private-vlan host
  switchport private-vlan host-association 99 100
!
interface GigabitEthernet 0/4
  switchport mode private-vlan host
  switchport private-vlan host-association 99 101
!
interface GigabitEthernet 0/5
  switchport mode trunk
!
SwitchA# show vlan private-vlan
```

VLAN	Type	Status	Routed	Ports	Associated VLANs
99	primary				100-101
100	community				
101	isolated				

```
-----  
99   primary   active   Disabled Gi0/1, Gi0/5       100-101  
100  community active   Disabled Gi0/2, Gi0/3, Gi0/5   99  
101  isolated  active   Disabled Gi0/4, Gi0/5       99
```

```
• • • • •
```

B

```
SwitchB#show running-config  
!  
vlan 99  
  private-vlan primary  
  private-vlan association add 100-101  
!  
vlan 100  
  private-vlan community  
!  
vlan 101  
  private-vlan isolated  
!  
interface GigabitEthernet 0/1  
  switchport mode trunk  
!  
interface GigabitEthernet 0/2  
  switchport mode private-vlan host  
  switchport private-vlan host-association 99 101  
!  
interface GigabitEthernet 0/3  
  switchport mode private-vlan host  
  switchport private-vlan host-association 99 100
```

```
• • • • •
```

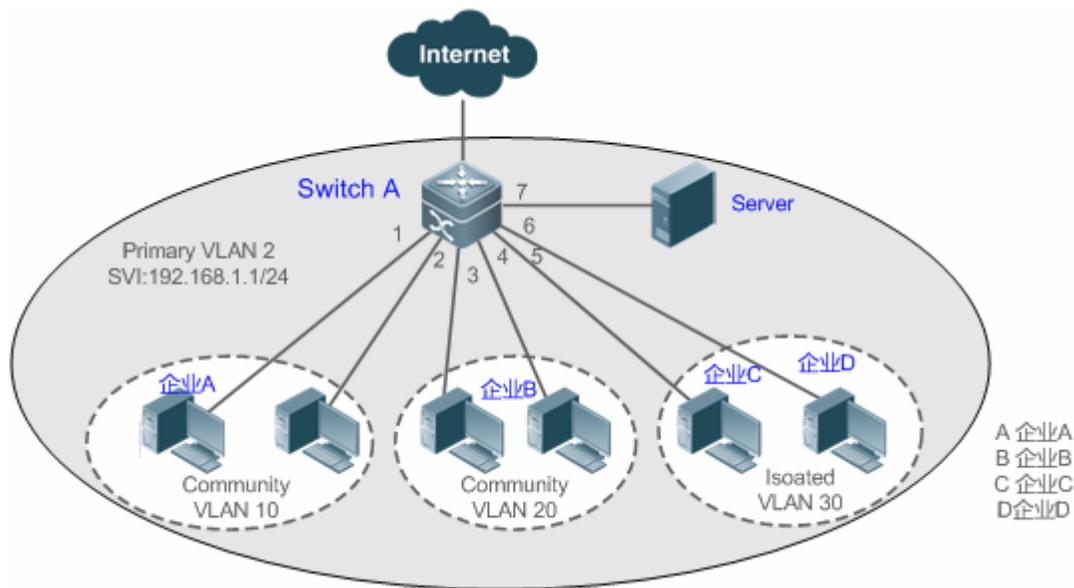
常见错误

- PVLAN 的主 VLAN 和辅助 VLAN 没有二层关联，配置隔离端口、混杂口、群体口时增加端口的 VLAN 列表失败。
- 一个 host 口关联多个 PVLAN 对时失败。

配置举例

📌 PVLAN 单台设备三层应用

【图 7-4】



- 【配置方法】
- 在设备上（本例为 Switch A）配置 PVLAN 功能，具体配置要点可参考“PVLAN 跨设备二层应用”章节的配置要点。
 - 将直连服务器的端口（本例为端口 Gi 0/7）设置为 Promiscuous Port，所有企业用户都可以通过 Promiscuous Port 和服务器通信。
 - 在三层设备（本例为 Switch A）配置 PVLAN 的网关地址（本例配置 VLAN 2 的 SVI 为 192.168.1.1/24），并配置 Primary VLAN（本例为 VLAN 2）和 Secondary VLAN（本例为 VLAN 10、20、30）的三层接口映射关系，所有企业用户可以通过这个网关地址与外网通信
-
- 跨设备运行 PVLAN，需要将跨界的设备连接端口配置为 Trunk 口。

A

```
SwitchA#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
SwitchA(config)#vlan 2
SwitchA(config-vlan)#private-vlan primary
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 10
SwitchA(config-vlan)#private-vlan community
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 20
SwitchA(config-vlan)#private-vlan community
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 30
SwitchA(config-vlan)#private-vlan isolated
SwitchA(config-vlan)#exit
SwitchA(config)#vlan 2
SwitchA(config-vlan)#private-vlan association 10, 20, 30
SwitchA(config-vlan)#exit
```

```
SwitchA(config)#interface range gigabitEthernet 0/1-2
SwitchA(config-if-range)#switchport mode private-vlan host
SwitchA(config-if-range)#switchport private-vlan host-association 2 10
SwitchA(config-if-range)#exit
SwitchA(config)#interface range gigabitEthernet 0/3-4
SwitchA(config-if-range)#switchport mode private-vlan host
SwitchA(config-if-range)#switchport private-vlan host-association 2 20
SwitchA(config-if-range)#exit
SwitchA(config)#interface range gigabitEthernet 0/5-6
SwitchA(config-if-range)#switchport mode private-vlan host
SwitchA(config-if-range)#switchport private-vlan host-association 2 30
SwitchA(config-if-range)#exit
SwitchA(config)#interface gigabitEthernet 0/7
SwitchA(config-if-GigabitEthernet 0/7)#switchport mode private-vlan promiscuous
SwitchA(config-if-GigabitEthernet 0/7)#switchport private-vlan mapping 2 10, 20, 30
SwitchA(config-if-GigabitEthernet 0/7)#exit
SwitchA(config)#interface vlan 2
SwitchA(config-if-VLAN 2)#ip address 192.168.1.1 255.255.255.0
SwitchA(config-if-VLAN 2)#private-vlan mapping 10, 20, 30
SwitchA(config-if-VLAN 2)#exit
```

【检验方法】 使各个子域内的用户 IP 来 ping 网关地址 192.168.1.1 能 ping 通。

A

```
SwitchA#show running-config
!
vlan 2
  private-vlan primary
  private-vlan association add 10, 20, 30
!
vlan 10
  private-vlan community
!
vlan 20
  private-vlan community
!
vlan 30
  private-vlan isolated
!
interface GigabitEthernet 0/1
  switchport mode private-vlan host
  switchport private-vlan host-association 2 10
```

```
!  
interface GigabitEthernet 0/2  
  switchport mode private-vlan host  
  switchport private-vlan host-association 2 10  
!  
interface GigabitEthernet 0/3  
  switchport mode private-vlan host  
  switchport private-vlan host-association 2 20  
!  
interface GigabitEthernet 0/4  
  switchport mode private-vlan host  
  switchport private-vlan host-association 2 20  
!  
interface GigabitEthernet 0/5  
  switchport mode private-vlan host  
  switchport private-vlan host-association 2 30  
!  
interface GigabitEthernet 0/6  
  switchport mode private-vlan host  
  switchport private-vlan host-association 2 30  
!  
interface GigabitEthernet 0/7  
  switchport mode private-vlan promiscuous  
  switchport private-vlan mapping 2 add 10,20,30  
!  
interface VLAN 2  
  no ip proxy-arp  
  ip address 192.168.1.1 255.255.255.0  
  private-vlan mapping add 10,20,30  
!  
SwitchA#show vlan private-vlan  
VLAN Type Status Routed Ports Associated VLANs  
-----  
2 primary active Enabled Gi0/7 10,20,30  
10 community active Enabled Gi0/1, Gi0/2 2  
20 community active Enabled Gi0/3, Gi0/4 2  
30 isolated active Enabled Gi0/5, Gi0/6 2
```

📌 常见配置错误

- PVLAN 的主 VLAN 和辅助 VLAN 没有二层关联，在三层关联时配置失败。

- 没进行三层关联就连接外网，结果无法与外网通信。
- 连接服务器和外网的接口没有配置为混杂口，导致上下行报文转发不对称。

7.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
查看 PVLAN 的配置。	<code>show vlan private-vlan</code>

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 PVLAN 调试开关。	<code>debug bridge pvlan</code>

8 MSTP

8.1 概述

生成树协议是一种二层管理协议，它通过选择性地阻塞网络中的冗余链路来消除二层环路，同时还具备链路备份的功能。

与众多协议的发展过程一样，生成树协议也是随着网络的发展而不断更新的，从最初的 STP (Spanning Tree Protocol, 生成树协议) 到 RSTP (Rapid Spanning Tree Protocol, 快速生成树协议)，再到最新的 MSTP (Multiple SpanningTree Protocol, 多生成树协议)。

对二层以太网来说，两个 LAN 间只能有一条活动着的通路，否则就会产生广播风暴。但是为了加强一个局域网的可靠性，建立冗余链路又是必要的，其中的一些通路必须处于备份状态，如果当网络发生故障，另一条链路失效时，冗余链路就必须被提升为活动状态。手工控制这样的过程显然是一项非常艰苦的工作，STP 协议就自动地完成这项工作。它能使一个局域网中的设备起以下作用：

- 发现并启动局域网的一个最佳树型拓扑结构。
- 发现故障并随之进行恢复，自动更新网络拓扑结构，使在任何时候都选择了可能的最佳树型结构。

局域网的拓扑结构是根据管理员设置的一组网桥配置参数自动进行计算的。使用这些参数能够生成最好的一棵拓扑树。只有配置得当，才能得到最佳的方案。

RSTP 协议完全向下兼容 802.1D STP 协议，除了和传统的 STP 协议一样具有避免回路、提供冗余链路的功能外，最主要的特点就是“快”。如果一个局域网内的网桥都支持 RSTP 协议且管理员配置得当，一旦网络拓扑改变而要重新生成拓扑树只需要不超过 1 秒的时间 (传统的 STP 需要大约 50 秒)。

STP 和 RSTP 存在的不足：

- STP 不能快速迁移，即使是在点对点链路或边缘端口，也必须等待两倍的 Forward Delay 的时间延迟，端口才能迁移到转发状态。
- RSTP 可以快速收敛，但和 STP 一样还存在如下缺陷：由于局域网内所有 VLAN 都共享一棵生成树，因此所有 VLAN 的报文都沿这棵生成树进行转发，不能按 VLAN 阻塞冗余链路，也无法在 VLAN 间实现数据流量的负载均衡。

MSTP(Multiple Spanning Tree Protocol, 多生成树协议), 由 IEEE 制定的 802.1s 标准定义, 它可以弥补 STP、RSTP 的缺陷, 既可以快速收敛, 也能使不同 VLAN 的流量沿各自的路径转发, 从而为冗余链路提供了更好的负载分担机制。

简单地说, STP/RSTP 是基于端口的, MSTP 是基于实例的。所谓实例就是多个 VLAN 的一个集合, 通过多个 VLAN 捆绑到一个实例的方法可以节省通信开销和资源占用率。

本设备既支持 STP 协议, 也支持 RSTP 协议与 MSTP 协议, 遵循 IEEE 802.1D、IEEE 802.1w 及 IEEE 802.1s 标准。

 下文仅介绍 MSTP 的相关内容。

协议规范

- IEEE 802.1D : Media Access Control (MAC) Bridges
- IEEE 802.1w : Part 3: Media Access Control (MAC) Bridges—Amendment 2: Rapid Reconfiguration
- IEEE 802.1s : Virtual Bridged Local Area Networks—Amendment 3: Multiple Spanning Trees

8.2 典型应用

典型应用	场景描述
MSTP+VRRP双核心拓扑	通过设计层次化的网络架构模型，使用 <u>MSTP+VRRP 协议实现冗余备份和负载均衡，提高网络系统可用性。</u>
BPDU TUNNEL应用	介绍在 QINQ 网络环境中，使用 BPDU TUNNEL 功能，实现 STP 协议报文的隧道透传。

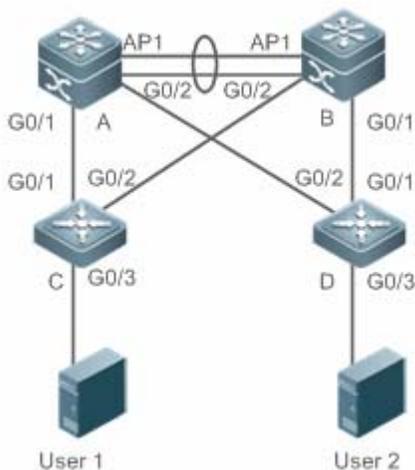
8.2.1 MSTP+VRRP双核心拓扑

应用场景

MSTP 协议典型的应用场景是 MSTP+VRRP 的双核心应用方案。该方案是提高网络系统可用性的一个比较优秀的解决方案，通常采用层次化的网络架构模型，分为三层（核心层、汇聚层和接入层）或二层（核心层和接入层）架构，共同组成交换网络系统，提供数据交换服务。

这种架构的主要优点在于层次化的结构。在层次化网络架构中，每一层次网络设备的各种容量指标、特点和功能，都针对其所在的网络位置和作用进行了优化，稳定性和可用性都得到了加强。

图 8-1 MSTP+VRRP 双核心拓扑



【注释】 上述拓扑分成两层的拓扑结构，分别为核心层（设备 A,B）和接入层（设备 C,D）

功能部署

- 核心层：MSTP 配置多实例达到负载均衡的效果。比如创建两个实例 1,2。实例 1 映射 VLAN 10，实例 2 映射 VLAN 20。设备 A 为实例 0,1 的根桥（实例 0 即 CIST 是默认存在的），设备 B 为实例 2 根桥。
- 核心层：设备 A 为 VLAN 10 的 VRRP 的主设备，设备 B 为 VLAN 20 的 VRRP 的主设备。

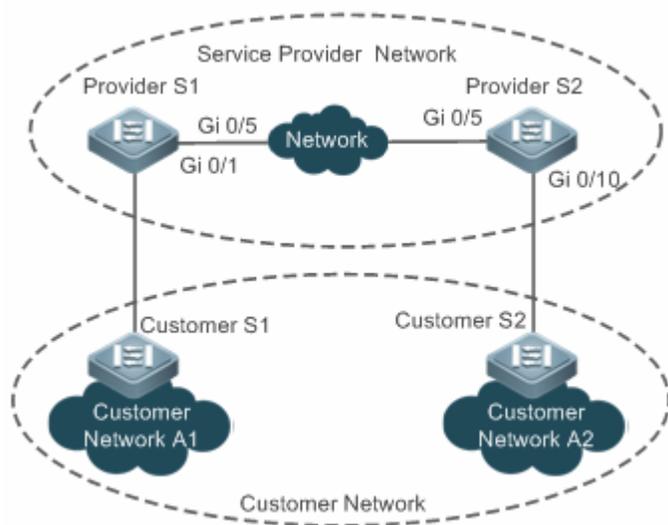
- 接入层：将直连终端（PC 或服务器）的端口配置成 Portfast 端口。同时配置 BPDU Guard 功能，防止用户私自接入非法的设备。

8.2.2 BPDU TUNNEL应用

应用场景

在 QINQ 网络中，通常分为用户网络和运营商网络。为了实现用户网络之间 STP 协议报文的传输而又不影响运营商网络产生影响，可以使用 BPDU TUNNEL 功能，以达到用户网络和运营商网络的 STP 协议分开计算，互不干扰。

图 8-2 BPDU Tunnel 应用拓扑图



【注释】 如上图所示，上部为运营商网络，下部为用户网络。其中，运营商网络包括边缘设备 Provider S1 和 Provider S2。Customer Network A1 和 Customer Network A2 为同一用户在不同地域的两个站点，Customer S1 和 Customer S2 为用户网络到运营商网络的接入设备，分别通过 Provider S1 和 Provider S2 接入运营商网络。应用 BPDU TUNNEL 功能，可以满足处于不同地域的 Customer Network A1 和 Customer Network A2 可以跨越运营商网络进行统一生成树计算，而不影响运营商网络的生成树计算。

功能部属

- 在运营商边缘设备（本例为 Provider S1/Provider S2 上开启基本 QinQ 功能，实现用户网络的数据报文在运营商网络的指定 VLAN 内传输。
- 在运营商边缘设备（本例为 Provider S1/Provider S2 上开启 STP 协议透传功能，使运营商网络可以通过 BPDU TUNNEL 对用户网络的 STP 报文进行隧道传输。

8.3 功能详解

基本概念

▾ BPDU (Bridge Protocol Data Units)

要生成一个稳定的树型拓扑网络需要依靠以下元素：

- 每个网桥拥有的唯一的桥 ID (Bridge ID)，由桥优先级和 Mac 地址组合而成。
- 网桥到根桥的路径花费 (Root Path Cost)，以下简称根路径花费。
- 每个端口 ID (Port ID)，由端口优先级和端口号组合而成。

网桥之间通过交换 BPDU (Bridge Protocol Data Units，网桥协议数据单元) 帧来获得建立最佳树形拓扑结构所需要的信息。这些帧以组播地址 01-80-C2-00-00-00 (十六进制) 为目的地址。

每个 BPDU 由以下这些要素组成：

- Root Bridge ID (本网桥所认为的根桥 ID)。
- Root Path Cost (本网桥的根路径花费)。
- Bridge ID (本网桥的桥 ID)。
- Message Age (报文已存活的时间)
- Port ID (发送该报文端口的 ID)。

Forward-Delay Time、Hello Time、Max-Age Time 三个协议规定的时间参数。

其他一些诸如表示发现网络拓扑变化、本端口状态的标志位。

当网桥的一个端口收到高优先级的 BPDU (更小的 Bridge ID，更小的 Root Path Cost 等)，就在该端口保存这些信息，同时向所有端口更新并传播这些信息。如果收到比自己低优先级的 BPDU，网桥就丢弃该信息。

这样的机制就使高优先级的信息在整个网络中传播开，BPDU 的交流就有了下面的结果：

- 网络中选择了一个网桥为根桥 (Root Bridge)。
- 除根桥外的每个网桥都有一个根口 (Root Port)，即提供最短路径到 Root Bridge 的端口。
- 每个网桥都计算出了到根桥 (Root Bridge) 的最短路径。
- 每个 LAN 都有了指派网桥 (Designated Bridge)，位于该 LAN 与根桥之间的最短路径中。指派网桥和 LAN 相连的端口称为指派端口 (Designated Port)。
- 根口 (Root port) 和指派端口 (Designated Port) 进入 Forwarding 状态。

▾ Bridge ID

按 IEEE 802.1W 标准规定，每个网桥都要有单一的网桥标识 (Bridge ID)，生成树算法中就是以它为标准来选出根桥 (Root Bridge) 的。Bridge ID 由 8 个字节组成，后 6 个字节为该网桥的 mac 地址，前 2 个字节如下表所示。前 4 bit 表示优先级 (Priority)，后 8 bit 表示 System ID，为以后扩展协议而用，在 RSTP 中该值为 0，因此给网桥配置优先级就要是 4096 的倍数。

	Bit 位	值
Priority value	16	32768
	15	16384

	14	8192
	13	4096
System ID	12	2048
	11	1024
	10	512
	9	256
	8	128
	7	64
	6	32
	5	16
	4	8
	3	4
	2	2
1	1	

📌 Spanning-Tree Timers (生成树的定时器)

以下描述影响到整个生成树性能的三个定时器。

- Hello timer：定时发送 BPDU 报文的时间间隔。
- Forward-Delay timer：端口状态改变的时间间隔。当 RSTP 协议以兼容 STP 协议模式运行时，端口从 Listening 转变向 Learning，或者从 Learning 转向 Forwarding 状态的时间间隔。
- Max-Age timer：BPDU 报文消息生存的最长时间。当超出这个时间，报文消息将被丢弃。

📌 Port Roles and Port States

每个端口都在网络中有扮演一个角色（Port Role），用来体现在网络拓扑中的不同作用。

- Root port：提供最短路径到根桥（Root Bridge）的端口。
- Designated port：每个 LAN 的通过该口连接到根桥。
- Alternate port：根口的替换口，一旦根口失效，该口就立该变为根口。
- Backup port：Designated Port 的备份口，当一个网桥有两个端口都连在一个 LAN 上，那么高优先级的端口为 Designated Port，低优先级的端口为 Backup Port。
- Disable port：当前不处于活动状态的口，即 Operation State 为 Down 的端口都被分配了这个角色。

以下为各个端口角色的示意图 1、2、3：

R = Root Port D = Designated Port A = Alternate Port B = Backup Port

在没有特别说明情况下，端口优先级从左到右递减。

图 8-3

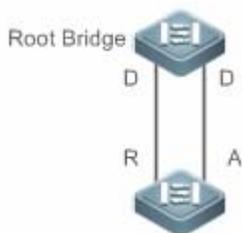


图 8-4

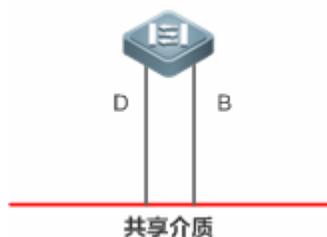
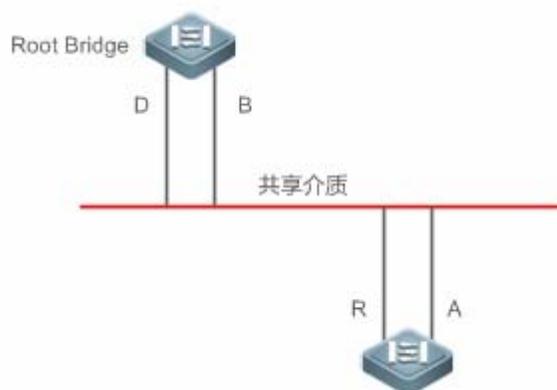


图 8-5



每个端口有三个状态（Port State）来表示是否转发数据包，从而控制着整个生成树拓扑结构。

- Discarding：既不对收到的帧进行转发，也不进行源 Mac 地址学习。
- Learning：不对收到的帧进行转发，但进行源 Mac 地址学习，这是个过渡状态。
- Forwarding：既对收到的帧进行转发，也进行源 Mac 地址的学习。

对一个已经稳定的网络拓扑，只有 Root Port 和 Designated Port 才会进入 Forwarding 状态，其它端口都只能处于 Discarding 状态。

📌 Hop Count

IST 和 MSTI 已经不用 Message Age 和 Max Age 来计算 BPDU 信息是否超时，而是用类似于 IP 报文 TTL 的机制来计算，它就是 Hop Count。

可以用 **spanning-tree max-hops** 全局配置命令来设置。在 Region 内，从 Region Root Bridge 开始，每经过一个设备，Hop Count 就会减 1，直到为 0 则表示该 BPDU 信息超时，设备收到 Hops 值为 0 的 BPDU 就要丢弃它。

为了和 Region 外的 STP、RSTP 兼容，MSTP 依然保留了 Message Age 和 Max Age 的机制。

功能特性

功能特性	作用
STP协议	STP (Spanning Tree Protocol , 生成树协议) , 由 IEEE 制定的 802.1D 标准定义 , 用于在局域网中消除数据链路层物理环路的协议。
RSTP协议	RSTP (Rapid Spanning Tree Protocol , 快速生成树协议) , 由 IEEE 制定的 802.1w 标准定义 , 它在 STP 基础上进行了改进 , 实现了网络拓扑的快速收敛。
MSTP协议	MSTP (Multiple Spanning Tree Protocol , 多生成树协议) , 由 IEEE 制定的 802.1s 标准定义 , 它可以弥补 STP、RSTP 和 PVST 的缺陷 , 既可以快速收敛 , 也能使不同 VLAN 的流量沿各自的路径转发 , 从而为冗余链路提供了更好的负载分担机制。
MSTP的可选特性	包括以下功能 : Port Fast 特性、BPDU Guard、BPDU Filter、Tc-protection、TC Guard、TC 过滤、BPDU 源 MAC 检查、BPDU 非法长度过滤、边缘口的自动识别、ROOT Guard 功能及 LOOP Guard 功能。

8.3.1 STP

STP 协议是用来避免链路环路产生的广播风暴、并提供链路冗余备份的协议。

工作原理

对二层以太网来说，两个 LAN 间只能有一条活动着的通路，否则就会产生广播风暴。但是为了加强一个局域网的可靠性，建立冗余链路又是必要的，其中的一些通路必须处于备份状态，如果当网络发生故障，另一条链路失效时，冗余链路就必须被提升为活动状态。手工控制这样的过程显然是一项非常艰苦的工作，STP 协议就自动地完成这项工作。它能使一个局域网中的设备起以下作用：

- 发现并启动局域网的一个最佳树型拓扑结构。
- 发现故障并随之进行恢复，自动更新网络拓扑结构，使在任何时候都选择了可能的最佳树型结构。

局域网的拓扑结构是根据管理员设置的一组网桥配置参数自动进行计算的。使用这些参数能够生成最好的一棵拓扑树。只有配置得当，才能得到最佳的方案。

相关配置

▾ 打开 spanning-tree 功能

缺省情况下，spanning-tree 功能是关闭的。

使用 **spanning-tree [forward-time seconds | hello-time seconds | max-age seconds]** 命令可以打开 STP，所带参数可在打开 STP 的同时，设置全局的基本设置。

forward-time 取值范围是 <4-30>，hello-time 取值范围是 <1-10>，max-age 取值范围是 <6-40>。

! 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。forward-time、hello-time、max-age 三个值的范围是相关的，修改了其中一个会影响到其他两个的值范围。这三个值之间有一个制约关系： $2 * (\text{Hello Time} + 1.0 \text{ second}) \leq \text{Max-Age Time} \leq 2 * (\text{Forward-Delay} - 1.0 \text{ second})$ ，不符合这个条件的值也会设置不成功。

8.3.2 RSTP

RSTP 协议完全向下兼容 802.1D STP 协议，除了和传统的 STP 协议一样具有避免回路、提供冗余链路的功能外，最主要的特点就是“快”。如果一个局域网内的网桥都支持 RSTP 协议且管理员配置得当，一旦网络拓扑改变而要重新生成拓扑树只需要不超过 1 秒的时间（传统的 STP 需要大约 50 秒）。

工作原理

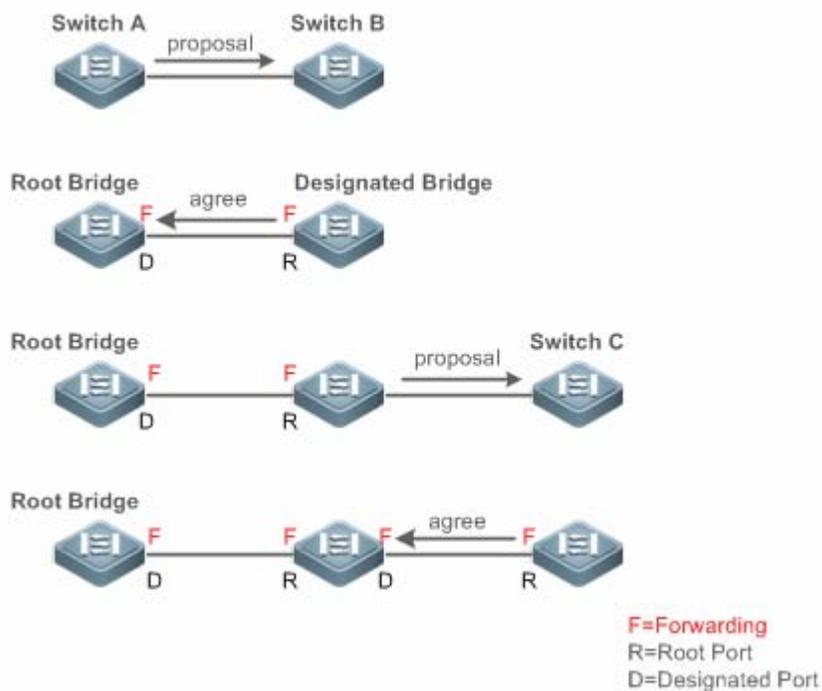
▾ RSTP 的快速收敛

现在开始介绍 RSTP 所特有的功能，即能让端口“快速”的 Forwarding。

STP 协议是选好端口角色（Port Role）后等待 30 秒（为 2 倍的 Forward-Delay Time，Forward-Delay Time 可配置，默认为 15 秒）再 Forwarding 的，而且每当拓扑发生变化后，每个网桥重新选出的 Root Port 和 Designated Port 都要经过 30 秒再 Forwarding，因此要等整个网络拓扑稳定为一个树型结构就大约需要 50 秒。

而 RSTP 端口的 Forwarding 过程就大不一样了，如下图所示，Switch A 发送 RSTP 特有“Proposal”报文，Switch B 发现 Switch A 的优先级比自身高，就选 Switch A 为根桥，收到报文的端口为 Root Port，立即 Forwarding，然后从 Root Port 向 Switch A 发送“Agree”报文。Switch A 的 Designated Port 得到“同意”，也就 Forwarding 了。然后 Switch B 的 Designated Port 又发送“Proposal”报文依次将生成树展开。因此在理论上，RSTP 是能够在网络拓扑发生变化的一瞬间恢复网络树型结构，达到快速收敛。

图 8-6



i 以上的“握手”过程是有条件的，就是端口间必须是“Point-to-point Connect (点对点连接)”。为了让设备发挥最大的功效，最好不要使设备间为非点对点连接。

以下列出了“非点对点连接”的范例图。

非点对点连接范例：

图 8-7

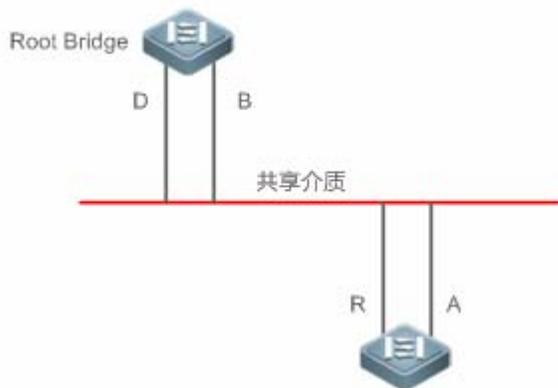
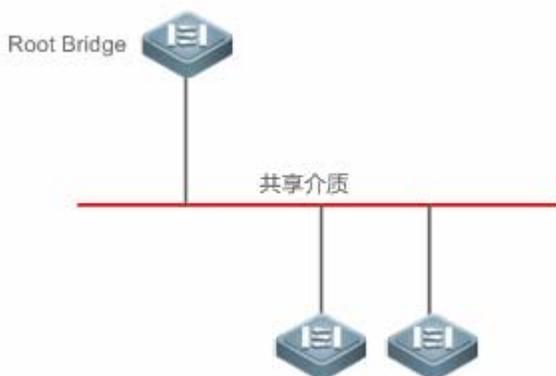
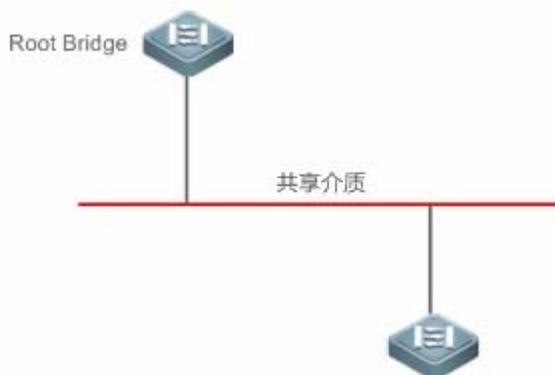


图 8-8



下图为“点对点”连接，请用户注意区分

图 8-9



↘ RSTP 与 STP 的兼容

RSTP 协议可以与 STP 协议完全兼容，RSTP 协议会根据收到的 BPDU 版本号来自动判断与之相连的网桥是支持 STP 协议还是支持 RSTP 协议，如果是与 STP 网桥互连就只能按 STP 的 Forwarding 方法，过 30 秒再 Forwarding，无法发挥 RSTP 的最大功效。

另外，RSTP 和 STP 混用还会遇到这样一个问题。如下图所示，Switch A 是支持 RSTP 协议的，Switch B 只支持 STP 协议，它们俩互连，Switch A 发现与它相连的是 STP 桥，就发 STP 的 BPDU 来兼容它。但后来如果换了台 Switch C，它支持 RSTP 协议，但 Switch A 却依然在发 STP 的 BPDU，这样使 Switch C 也认为与之互连的是 STP 桥了，结果两台支持 RSTP 的设备却以 STP 协议来运行，大大降低了效率。

为此 RSTP 协议提供了 Protocol-migration 功能来强制发 RSTP BPDU (这种情况下，对端网桥必须支持 RSTP)，这样 Switch A 强制发了 RSTP BPDU，Switch C 就发现与之互连的网桥是支持 RSTP 的，于是两台设备就都以 RSTP 协议运行了，如图 13。

图 8-10

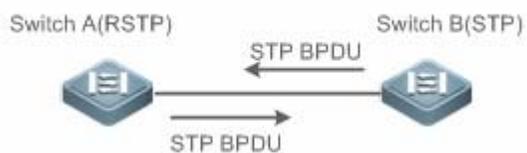
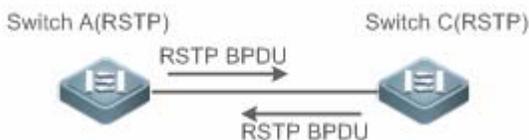


图 8-11



相关配置

配置 Protocol Migration 处理

使用 `clear spanning-tree detected-protocols [interface interface-id]` 命令可以让该端口强制进行版本检查。相关说明请参看 RSTP 与 STP 的兼容。

8.3.3 MSTP 协议

MSTP (Multiple Spanning Tree Protocol), 多生成树协议, 它可以弥补 STP、RSTP 的缺陷, 既可以快速收敛, 也能使不同 VLAN 的流量沿各自的路径转发, 从而为冗余链路提供了更好的负载分担机制。

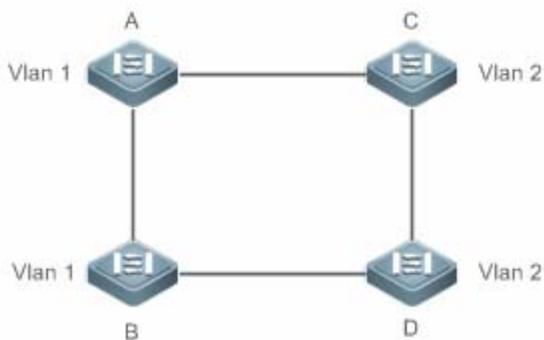
工作原理

本设备支持 MSTP, MSTP 是在传统的 STP、RSTP 的基础上发展而来的新的生成树协议, 本身就包含了 RSTP 的快速 FORWARDING 机制。

由于传统的生成树协议与 Vlan 没有任何联系, 因此在特定网络拓扑下就会产生以下问题:

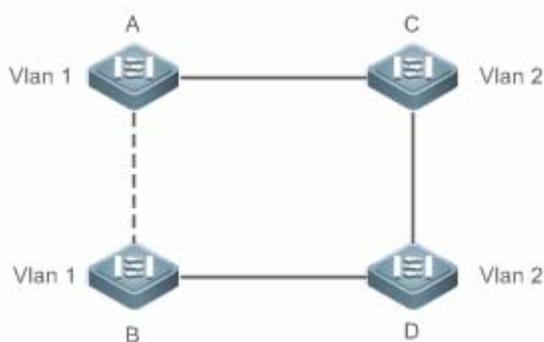
如下图所示, 设备 A、B 在 Vlan1 内, 设备 C、D 在 Vlan2 内, 然后连成环路。

图 8-12



若从设备 A 依次通过设备 C、D 到达 B 的链路花费比从设备 A 直接到 B 的链路花费更少的情况下, 会造成把设备 A 和 B 间的链路给 DISCARDING (如图 15 所示)。由于设备 C、D 不包含 Vlan1, 无法转发 Vlan1 的数据包, 这样设备 A 的 Vlan1 就无法与设备 B 的 Vlan1 进行通讯。

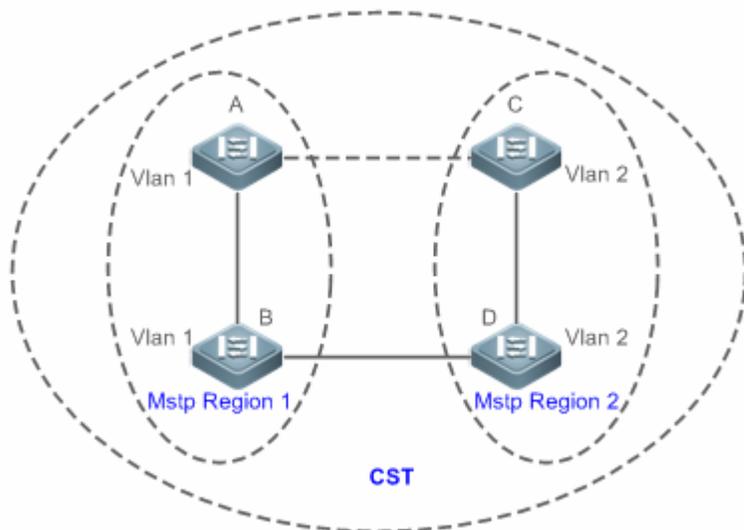
图 8-13



为了解决这个问题，MSTP 就产生了，它可以把一台设备的一个或多个 Vlan 划分为一个 Instance，有着相同 Instance 配置的设备就组成一个域(MST Region)，运行独立的生成树(这个内部的生成树称为 IST ,Internal Spanning-tree)；这个 MST region 组合就相当于一个大的设备整体，与其他 MST Region 再进行生成树算法运算，得出一个整体的生成树，称为 CST (Common Spanning Tree)。

按这种算法，以上网络就可以在 MSTP 算法下形成图 16 的拓扑：设备 A 和 B 都在 MSTP Region 1 内，MSTP Region 1 没能环路产生，所以没有链路 DISCARDING，同理 MSTP Region 2 的情况也是一样的。然后 Region 1 和 Region 2 就分别相当于两个大的设备，这两台“设备”间有环路，因此根据相关配置选择一条链路 DISCARDING。

图 8-14



这样，既避免了环路的产生，也能让相同 Vlan 间的通讯不受影响。

划分 MSTP Region

根据以上描述，很明显，要让 MSTP 产生应有的作用，首先就要合理地划分 MSTP Region，相同 MSTP Region 内的设备“MST 配置信息”一定要相同。

MST 配置信息包括：

- MST 配置名称 (Name)：最长可用 32 个字节长的字符串来标识 MSTP Region。

- MST Revision Number：用一个 16bit 长的修正值来标识 MSTP Region。
- MST Instance—vlan 的对应表：每台设备都最多可以创建 64 个 Instance（id 从 1 到 64），Instance 0 是强制存在的，所以系统最多可以支持 65 个 Instance。用户还可以按需要分配 1-4094 个 Vlan 属于不同的 Instance（0 - 64），未分配的 Vlan 缺省就属于 Instance 0。这样，每个 MSTI（MST Instance）就是一个“Vlan 组”，根据 BPDU 里的 MSTI 信息进行 MSTI 内部的生成树算法，不受 CIST 和其他 MSTI 的影响。

可在用 spanning-tree mst configuration 全局配置命令进入“MST 配置模式”配置以上信息。

MSTP BPDU 里附带以上信息，如果一台设备收到的 BPDU 里的 MST 配置信息和自身的一样，就会认为该端口上连着的设备和自己是属于同一个 MST Region，否则就认为是从另外一个 Region 来的。

i 建议在关闭 MSTP 模式后配置 Instance—vlan 的对应表，配置好后再打开 MSTP，以保证网络拓扑的稳定和收敛。

📌 IST (MSTP region 内的生成树)

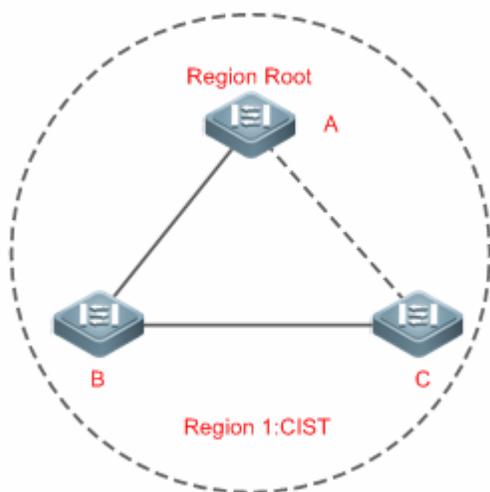
划分好 MSTP Region 后，每个 Region 里就按各个 Instance 所设置的 Bridge Priority、Port Priority 等参数选出各个 Instance 独立的 Root Bridge，以及每台设备上各个端口的 Port Role，然后就 Port Role 指定该端口在该 Instance 内是 FORWARDING 还是 DISCARDING 的。

这样 经过 MSTP BPDU 的交流，IST(Internal Spanning Tree) 就生成了，而各个 Instance 也独立的有了自己的生成树(MSTI)，其中 Instance 0 所对应的生成树与 CST 共同称为 CIST (Common Instance Spanning Tree)。也就是说，每个 Instance 都为各自的“vlan 组”提供了一条单一的、不含环路的网络拓扑。

如下图所示，在 Region 1 内，设备 A、B、C 组成环路。

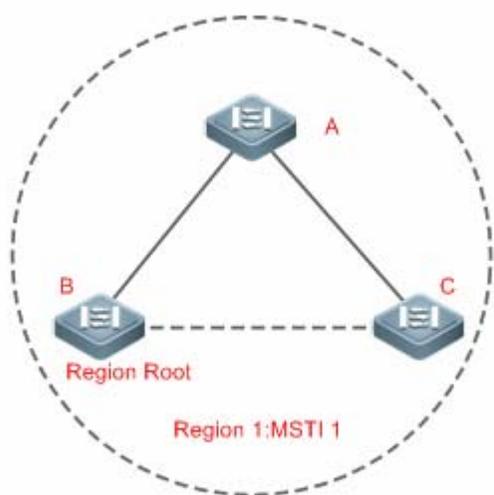
在 CIST (Instance 0) 中，如图 17，因 A 的优先级最高，被选为 Region Root，再根据其他参数，把 A 和 C 间的链路给 DISCARDING。因此，对 Instance 0 的“Vlan 组”来说，只有 A 到 B、B 到 C 的链路可用，打断了这个“Vlan 组”的环路。

图 8-15



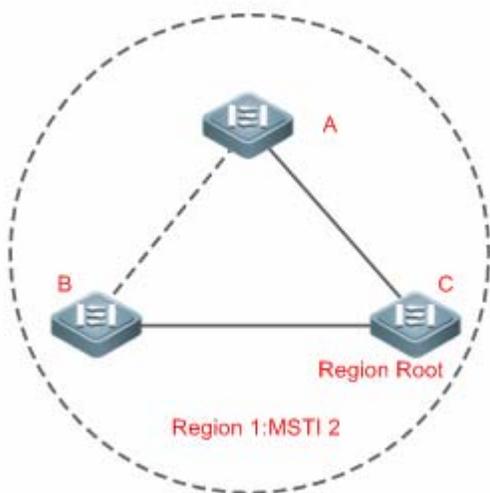
而对 MSTI 1 (Instance 1) 来说，如图 18，B 的优先级最高，被选为 Region Root，再根据其他参数，把 B 和 C 间的链路给 DISCARDING。因此，对 Instance 1 的“Vlan 组”来说，只有 A 到 B、A 到 C 的链路可用，打断了这个“Vlan 组”的环路。

图 8-16



而对 MSTI 2 (Instance 2) 来说，图 19，C 的优先级最高，被选为 Region Root，再根据其他参数，把 A 和 B 间的链路给 DISCARDING。因此，对 Instance 2 的“Vlan 组”来说，只有 B 到 C、A 到 C 的链路可用，打断了这个“Vlan 组”的环路。

图 8-17

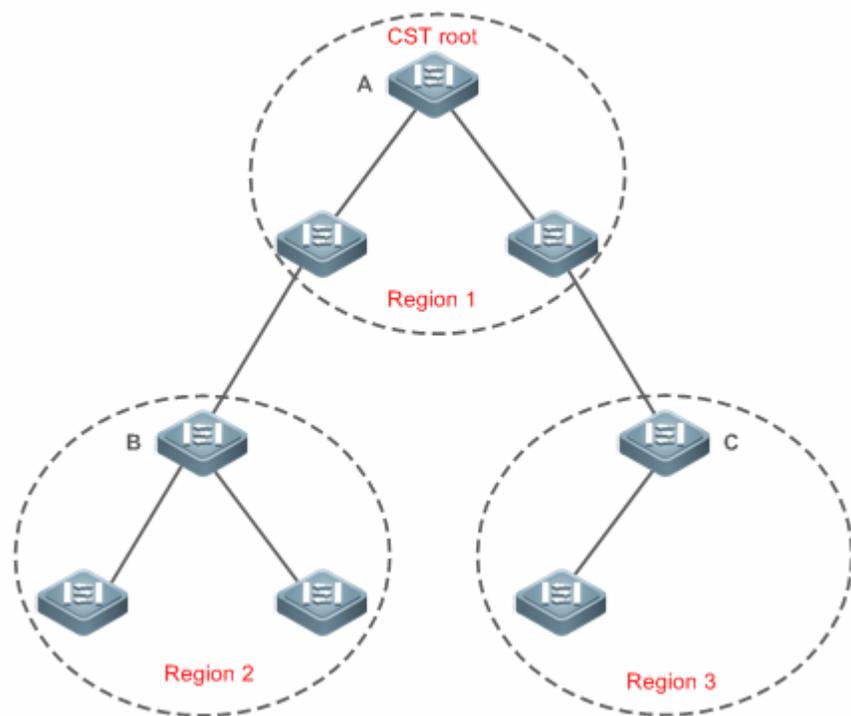


用户在这里要注意的是 MSTP 协议本身不关心一个端口属于哪个 Vlan，所以用户应该根据实际的 Vlan 配置情况来为相关端口配置对应的 Path Cost 和 Priority，以防 MSTP 协议打断了不该打断的环路。

↘ CST (MSTP region 间的生成树)

个 MSTP region 对 CST 来说可以相当于一个大的设备整体，不同的 MSTP Region 也生成一个大的网络拓扑树，称为 CST(Common Spanning Tree)。如图 20 所示，对 CST 来说，Bridge ID 最小的设备 A 被选为整个 CST 的根(CST Root)，同时也是这个 Region 内的 CIST Regional Root。在 Region 2 中，由于设备 B 到 CST Root 的 Root Path Cost 最短，所以被选为这个 Region 内的 CIST Regional Root。同理，Region 3 选设备 C 为 CIST Regional Root。

图 8-18



CIST Regional Root 不一定是该 Region 内 Bridge ID 最小的那台设备，它是指该 Region 内到 CST Root 的 Root Path Cost 最小的设备。

同时，CIST Regional Root 的 Root Port 对 MSTI 来说有了个新的 Port Role，为“Master port”，作为所有 Instance 对外的“出口”，它对所有 Instance 都是 FORWARDING 的。为了使拓扑更稳定，我们建议每个 Region 对 CST Root 的“出口”尽量只在该 Region 的一台设备上！

📌 MSTP 和 RSTP、STP 协议的兼容

对 STP 协议来说，MSTP 会像 RSTP 那样发 STP BPDU 来兼容它，详细情况请参考“RSTP 与 STP 的兼容”章节内容。而对 RSTP 协议来说，本身会处理 MSTP BPDU 中 CIST 的部分，因此 MSTP 不必专门的发 RSTP BPDU 以兼容它。每台运行 STP 或 RSTP 协议的设备都是单独的一个 Region，不与任何一个设备组成同一个 Region。

相关配置

📌 配置 STP 的模式

缺省情况下，STP 模式是 MSTP。

使用 `spanning-tree mode [stp | rstp | mstp]` 命令可以修改 STP 模式。

8.3.4 MSTP的可选特性

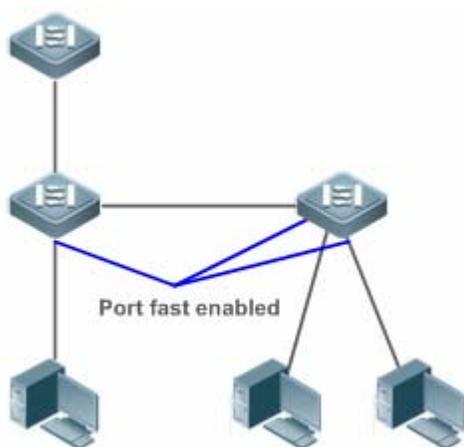
MSTP 的可选特性，主要包括 Port Fast 端口设置、BPDU Guard 设置、BPDU Filter 设置、TC Guard 和 Guard 模式设置等。主要用来在 MSTP 的组网应用中，能够根据网络的拓扑结构和应用特点，针对性地进行 MSTP 的配置部署，增加 MSTP 协议运行的稳定性、健壮性和抗攻击性，满足 MSTP 协议在不同用户场景的应用需求。

工作原理

Port Fast

如果设备的端口直连着网络终端，那么就可以设置该端口为 Port Fast，端口直接 Forwarding，这样可免去端口等待 Forwarding 的过程（如果不配置 Port Fast 的端口，就要等待 30 秒 Forwarding）。下图表示了一个设备的哪些端口可以配置为 Port Fast enable。

图 8-19



如果在设了 Port Fast 的端口中还收到 BPDU，则它的 Port Fast Operational State 为 Disabled。这时该端口会按正常的 STP 算法进行 Forwarding。

BPDU Guard

BPDU Guard 既能全局的 enable，也能针对单个 Interface 进行 enable。这两者有些细小的差别。

可以在全局模式中用 spanning-tree portfast bpduguard default 命令打开全局的 BPDU Guard enabled 状态，在这种状态下，如果某个 Interface 打开了 Port Fast 或该接口自动识别为边缘口，而该 Interface 收到了 BPDU，该端口就会进入 Error-disabled 状态，以示配置错误；同时整个端口被关闭，表示网络中可能被非法用户增加了一台网络设备，使网络拓扑发生改变。

也可以在 Interface 配置模式下用 spanning-tree bpduguard enable 命令来打开单个 Interface 的 BPDU Guard（与该端口是否打开 Port Fast 无关）。在这个情况下如果该 Interface 收到了 BPDU，就进入 Error-disabled 状态。

BPDU Filter

BPDU Filter 既能全局的 enable，也能针对单个 Interface 进行 enable。这两者有些细小的差别。

可以在全局模式中用 `spanning-tree portfast bpdufilter default` 命令打开全局的 BPDU Filter enabled 状态,在这种状态下,Port Fast enabled 的 Interface 将既不收 BPDU,也不发 BPDU,这样,直连 Port Fast enabled 端口的主机就收不到 BPDU。而如果 Port Fast enabled 的 Interface 因收到 BPDU 而使 Port Fast Operational 状态 disabled, BPDU Filter 也就自动失效。

也可以在 Interface 配置模式下用 `spanning-tree bpdufilter enable` 命令设置了单个 Interface 的 BPDU Filter enable (与该端口是否打开 Port Fast 无关)。在这个情况下该 Interface 既不收 BPDU,也不发 BPDU,并且是直接 Forwarding 的。

▾ Tc-protection

TC-BPDU 报文是指携带 TC 标志的 BPDU 报文,交换机收到这类报文表示网络拓扑发生了变化,会进行 MAC 地址表的删除操作,对三层交换机,还会引发快转模块的重新打通操作,并改变 ARP 表项的端口状态。为避免交换机受到伪造 TC-BPDU 报文的恶意攻击时频繁进行以上操作,负荷过重,影响网络稳定,可以使用 TC-protection 功能进行保护。

Tc-protection 只能全局的打开和关闭,缺省情况下为关闭此功能。

在打开相应功能时,收到 TC-BPDU 报文后的一定时间内(一般为 4 秒),只进行一次删除操作,同时监控该时间段内是否收到 TC-BPDU 报文。如果在该时间段内收到了 TC-BPDU 报文,则设备在该时间超时后再进行一次删除操作。这样可以避免频繁的删除 MAC 地址表项和 ARP 表项。

▾ TC Guard

Tc-Protection 功能可以保证网络产生大量 tc 报文时减少动态 MAC 地址和 ARP 的删除,但在遇到 TC 报文攻击的时候还是会产生很多的删除操作,并且 TC 报文是可扩散的,将影响整个网络。使用 TC Guard 功能,我们允许用户在全局或者端口上禁止 TC 报文的扩散。当一个端口收到 TC 报文的时候,如果全局配置了 TC Guard 或者是端口上配置了 TC Guard,则该端口将屏蔽掉该端口接收或者是自己产生的 TC 报文,使得 TC 报文不会扩散到其它端口,这样能有效控制网络中可能存在的 TC 攻击,保持网络的稳定,尤其是在三层设备上,该功能能有效避免接入层设备的振荡引起核心路由中断的问题。

❗ 错误的使用 tc-guard 功能会使网络之间的通讯中断。

❗ 建议在确认网络当中有非法的 tc 报文攻击的情况下再打开此功能。

❗ 打开全局的 tc-guard,则所有端口都不会对外扩散 tc 报文。适用于桌面接入设备上开启。

❗ 打开接口的 tc-guard,则对于该接口产生的拓扑变化以及收到的 tc 报文,将不向其它端口扩散。适合在上链口,尤其是汇聚接核心的端口开启该功能。

▾ TC 过滤

配置 TC Guard 功能,端口将不扩散 TC 报文到本设备上其它参与生成树计算的端口,这里的不扩散包括了两种情况:一种是端口收到的 TC 报文不扩散,一种是端口自己产生的 TC 报文不扩散。端口自己产生的 TC 报文是指当端口转发状态发生变化时(例如从 block 到 forwarding 的转变),端口会产生 TC 报文,表示拓扑可能发生了变化。

这样,可能引发的问题时,由于 TC Guard 阻止了 TC 报文的扩散,导致当发生拓扑变化的时候,设备没有清除相应端口的 MAC 地址,转发数据出错。

因此,引入了 TC 过滤的概念。TC 过滤是指对于端口收到的 TC 报文不处理,而正常的拓扑变化的情况,能够处理。这样,解决了未配置 Portfast 的端口频繁地 UP/DOWN 引起的清地址和核心路由中断的问题,又能保证发生拓扑变化时,核心路由表项能够得到及时地更新。

❗ TC 过滤功能缺省关闭。

▾ BPDU 源 MAC 检查

BPDU 源 MAC 检查是为了防止通过人为发送 BPDU 报文来恶意攻击交换机而使 MSTP 工作不正常。当确定了某端口点对点链路对端相连的交换机时，可通过配置 BPDU 源 MAC 检查来达到只接收对端交换机发送的 BPDU 帧，丢弃所有其他 BPDU 帧，从而达到防止恶意攻击。你可以在 interface 模式下来为特定的端口配置相应的 BPDU 源 MAC 检查 MAC 地址，一个端口只允许配置一个过滤 MAC 地址，通过 no bpdu src-mac-check 来禁止 BPDU 源 MAC 检查，此时端口接收任何 BPDU 帧。

▾ BPDU 非法长度过滤

BPDU 的以太网长度字段超过 1500 时，该 BPDU 帧将被丢弃，以防止收到非法 BPDU 报文。

▾ 边缘口的自动识别

指派口在一定的时间内(为 3 秒)，如果收不到下游端口发送的 BPDU，则认为该端口相连的是一台网络设备，从而设置该端口为边缘端口，直接进入 Forwarding 状态。自动标识为边缘口的端口因收到 BPDU 而自动识别为非边缘口。

可以通过 spanning-tree autoedge disabled 命令取消边缘口的自动识别功能。

该功能是缺省打开的。

- ⚠ 边缘口的自动标识功能与手工的 Port Fast 冲突时，以手工配置的为准。
- ⚠ 该功能作用于指派口与下游端口进行快速协商转发的过程中，所以 STP 协议不支持该功能。同时如果指派口已经处于转发状态，对该端口进行 Autoedge 的配置不会生效，只有在重新快速协商的过程中才生效，如拔插网线。
- ⚠ 端口如果先打开了 BPDU Filter，则该端口直接 Forwarding，不会自动识别为边缘口。
- ⚠ 该功能只适用与指派口。

▾ ROOT Guard 功能

在网络设计中常常将根桥和备份根桥划分在同一个域内，由于维护人员的错误配置或网络中的恶意攻击，根桥有可能收到优先级更高的配置信息，从而失去当前根桥的位置，引起网络拓扑的错误的变动。Root Guard 功能就是为了防止这种情况的出现。

接口打开 Root Guard 功能时，强制其在所有实例上的端口角色为指定端口，一旦该端口收到优先级更高的配置信息时，Root Guard 功能会将该接口置为 root-inconsistent (blocked) 状态，在足够长的时间内没有收到更优的配置信息时，端口会恢复成原来的正常状态。

当端口因 Root Guard 而处于 blocked 状态时，可以通过手动恢复为正常状态，即关闭端口的 ROOT Guard 功能或关闭接口的保护功能（在接口模式下配置 spanning-tree guard none）。

- ⚠ 错误的使用 ROOT Guard 特性会导致网络链路的断开。
- ⚠ 在非指派口上打开 ROOT Guard 功能会强制其为指派口，同时端口会进入 BKN 状态，该状态表示端口因 Root 不一致而进入 blocked 状态。
- ⚠ 如果端口在 MST0 因收到更优的配置消息而进入 BKN 状态，会强制端口在其它所有的实例中处于 BKN 状态。
- ⚠ 端口的 ROOT Guard 和 LOOP Guard 同一时刻只能有一个生效。

▾ LOOP Guard 功能

由于单向链路的故障，根口或备份口由于收不到 BPDU 会变成指派口进入转发状态，从而导致了网络中环路产生，LOOP Guard 功能防止了这种情况的发生。

对于配置了环路保护的端口，如果收不到 BPDU，会进行端口角色的迁移，但端口状态将一直被设成 discarding 状态。直到重新收到 BPDU 而进行生成树的重计算。

⚠ 可以基于全局或接口打开 LOOP Guard 特性。

⚠ 端口的 ROOT Guard 和 LOOP Guard 同一时刻只能有一个生效。

⚠ MSTP 进程重启前，端口进入环路保护的 block 状态，而 MSTP 进程重启后，如果端口仍然接收不到 BPDU，则端口将转变成指派口并进入 forward 状态。因此，建议在重启 MSTP 进程前，检查端口进入环路保护的 block 状态的原因并及时解决，避免进程重启后生成树拓扑仍然出现异常。

📌 BPDU 透传

在 IEEE 802.1Q 标准中，BPDU 的目的 MAC 地址 01-80-C2-00-00-00 是作为保留地址使用的，即遵循 IEEE 802.1Q 标准的设备，对于接收到的 BPDU 帧是不转发的。然而，在实际的网络布署中，可能需要设备能够支持透传 BPDU 帧。例如，设备未开启 STP 协议时，需要透传 BPDU 帧，使得与之互连的设备之间的生成树计算正常。

⚠ BPDU 透传默认关闭。

⚠ BPDU 透传功能只在 STP 协议关闭时才起作用。当 STP 协议打开时，设备不透传 BPDU 帧。

📌 BPDU TUNNEL

在 QINQ 网络中，通常分为用户网络和运营商网络。QinQ 的基本原理是在用户报文进入运营商网络之前封装上一个运营商网络的 VLAN Tag，而把用户报文中原有的 VLAN Tag 当做数据，使报文带着两层 VLAN Tag 穿越运营商网络。在运营商网络中，报文只根据外层 VLAN Tag 传播，当用户报文离开运营商网络时，剥去外层 VLAN Tag。

为了实现用户网络之间 STP 协议报文的传输而又不对运营商网络产生影响，可以使用 STP 报文透传功能，即 BPDU TUNNEL 功能。当用户网络中 STP 协议报文进入边缘设备后，将目的 mac 地址改成私有地址在运营商网络中转发，到了另外一端边缘设备后，再将目的 mac 地址改成公有地址回到另一端用户网络，以达到 STP 协议报文在运营商网络透传的效果，从而使用户网络和运营商网络的 STP 协议分开计算，互不干扰。

相关配置

📌 配置接口的 Portfast 开关

缺省情况下，接口上的 Port Fast 开关是关闭的。

在全局配置模式下，使用 **spanning-tree portfast default** 命令可以打开所有接口的 Portfast 开关；使用 **no spanning-tree portfast default** 命令关闭所有接口的 portfast 开关。

在接口配置模式下使用 **spanning-tree portfast** 命令可以打开某个接口的 Portfast 开关；使用 **spanning-tree portfast disabled** 命令关闭某个接口的 portfast 开关。

📌 配置接口的 BPDU guard 开关

缺省情况下，接口上的 BPDU guard 开关是关闭的。

在全局配置模式下，使用 **spanning-tree portfast bpduguard default** 命令可以打开所有接口的 BPDU guard 开关；使用 **no spanning-tree portfast bpduguard default** 命令关闭所有接口的 BPDU guard 开关。

在接口配置模式下使用 **spanning-tree bpduguard enabled** 命令可以打开某个接口的 BPDU guard 开关，使用 **spanning-tree bpduguard disabled** 命令关闭某个接口的 BPDU guard 开关。

配置接口的 BPDU Filter 开关

缺省情况下，接口上的 BPDU Filter 开关是关闭的。

在全局配置模式下，使用 **spanning-tree portfast bpdufilter default** 命令可以打开所有接口的 BPDU Filter 开关；使用 **no spanning-tree portfast bpdufilter default** 命令关闭所有接口的 BPDU Filter 开关。

在接口配置模式下使用 **spanning-tree bpdufilter enabled** 命令可以打开某个接口的 BPDU Filter 开关；使用 **spanning-tree bpdufilter disabled** 命令关闭某个接口的 BPDU Filter 开关。

配置 Tc-protection 开关

缺省情况下，Tc-protection 开关是关闭的。

在全局配置模式下，使用 **spanning-tree tc-protection** 命令可以打开所有接口的 Tc-protection 开关；使用 **no spanning-tree tc-protection** 命令关闭所有接口的 Tc-protection 开关。

Tc-protection 只能全局的打开和关闭。

配置接口的 TC Guard 开关

缺省情况下，接口上的 tc guard 开关是关闭的。

在全局配置模式下，使用 **spanning-tree tc-protection tc-guard** 命令可以打开所有接口的 tc guard 开关；使用 **no spanning-tree tc-protection tc-guard** 命令关闭所有接口的 tc guard 开关。

在接口配置模式下使用 **spanning-tree tc-guard** 命令可以打开某个接口的 tc guard 开关；使用 **no spanning-tree tc-guard** 命令关闭某个接口的 tc guard 开关。

配置接口的 TC 过滤开关

缺省情况下，接口上的 TC 过滤功能是关闭的。

在接口配置模式下使用 **spanning-tree ignore tc** 命令打开某个接口的 TC 过滤功能；使用 **no spanning-tree ignore tc** 命令关闭某个接口的 TC 过滤功能。

配置接口的 BPDU 源 MAC 检查

缺省情况下，接口上的 BPDU 源 MAC 检查功能是关闭的。

在接口配置模式下使用 **bpdu src-mac-check H.H.H** 命令打开某个接口的 BPDU 源 MAC 检查功能；使用 **no bpdu src-mac-check** 命令关闭某个接口的 BPDU 源 MAC 检查功能。

配置接口的边缘口自动识别功能

缺省情况下，接口上的边缘口自动识别功能是关闭的。

在接口配置模式下使用 **spanning-tree autoedge** 命令打开某个接口的边缘口自动识别功能；使用 **spanning-tree autoedge disabled** 命令关闭某个接口的边缘口自动识别功能。

配置接口的 Root Guard 功能

缺省情况下，接口上的 Root Guard 功能是关闭的。

在接口配置模式下使用 **spanning-tree guard root** 命令打开某个接口的 Root Guard 功能；使用 **no spanning-tree guard root** 命令关闭某个接口的 Root Guard 功能。

配置接口的 Loop Guard 功能

缺省情况下，接口上的 Loop Guard 功能是关闭的。

在全局配置模式下，使用 **spanning-tree loopguard default** 命令打开所有接口的 Loop Guard 功能；使用 **no spanning-tree loopguard default** 命令关闭所有接口的 Loop Guard 功能。

在接口配置模式下使用 **spanning-tree guard loop** 命令打开某个接口的 Loop Guard 功能；使用 **no spanning-tree guard loop** 命令关闭某个接口的 Loop Guard 功能。

配置 BPDU 透传功能

缺省情况下，BPDU 透传功能是关闭的。

在全局配置模式下，使用 **bridge-frame forwarding protocol bpdu** 命令打开 BPDU 透传功能；使用 **no bridge-frame forwarding protocol bpdu** 命令关闭 BPDU 透传。

BPDU 透传功能只在 STP 协议关闭时才起作用。当 STP 协议打开时，设备不透传 BPDU 帧。

配置 BPDU TUNNEL

缺省情况下，BPDU TUNNEL 功能是关闭的。

在全局配置模式下，使用 **I2protocol-tunnel stp** 命令使能全局的 BPDU TUNNEL 功能；使用 **no I2protocol-tunnel stp** 命令关闭全局的 BPDU TUNNEL 功能。

在接口模式下，使用 **I2protocol-tunnel stp enable** 命令使能接口的 BPDU TUNNEL 功能；使用 **no I2protocol-tunnel stp enable** 命令关闭接口的 BPDU TUNNEL 功能。

BPDU TUNNEL 功能只在全局和接口同时使能的情况下才起作用。

8.4 配置详解

配置项	配置建议 & 相关命令	
打开生成树协议	 必须配置。用于打开生成树协议。	
	spanning-tree	打开生成树协议，并配置基本属性
	spanning-tree mode	配置生成树模式
配置生成树的兼容性	 可选配置。用于兼容友商设备。	
	spanning-tree compatible enable	打开接口的兼容模式
	clear spanning-tree detected-protocols	对 BPDU 进行强制版本检查
配置MSTP Region	 可选配置。用于配置 MSTP Region。	
	spanning-tree mst configuration	进入 MSTP Region 配置模式

配置RSTP快速收敛	 可选配置。用于配置端口的连接类型是不是“点对点连接”。	
	spanning-tree link-type	配置 link type
配置优先级	 可选配置。用于配置设备优先级或者端口优先级。	
	spanning-tree priority	配置设备优先级
	spanning-tree port-priority	配置端口优先级
配置接口的路径花费	 可选配置。用于配置端口的路径花费或路径花费缺省计算方法。	
	spanning-tree cost	配置端口的路径花费
	spanning-tree pathcost method	配置路径花费的缺省计算方法
配置BPDU帧的最大跳数	 可选配置。用于配置 BPDU 帧的最大跳数。	
	spanning-tree max-hops	配置 BPDU 帧的最大跳数。
配置接口port fast的相关特性	 可选配置。用于配置 port fast 特性。	
	spanning-tree portfast	打开 port fast 特性
	spanning-tree portfast bpduguard default	打开所有接口的 BPDU Guard
	spanning-tree bpduguard enabled	打开某个接口的 BPDU Guard
	spanning-tree portfast bpdufilter default	打开所有接口的 BPDU Filter
	spanning-tree bpdufilter enabled	打开某个接口的 BPDU Filter
配置TC相关的特性	 可选配置。用于配置 TC 特性。	
	spanning-tree tc-protection	打开 tc protection
	spanning-tree tc-protection tc-guard	打开所有接口的 tc guard
	spanning-tree tc-guard	打开某个接口的 tc guard
	spanning-tree ignore tc	打开某个接口的 tc 过滤
配置BPDU源MAC检查	 可选配置。用于配置 BPDU 源 MAC 检查功能。	
	bpdu src-mac-check	打开某个接口的 BPDU 源 MAC 检查
配置边缘口的自动识别	 可选配置。用于配置边缘口的自动识别功能。	
	spanning-tree autoedge	打开某个接口的边缘口自动识别 ,缺省是打开的。
配置接口保护相关的特性	 可选配置。用于配置接口保护相关的功能。	
	spanning-tree guard root	打开某个接口的 root guard
	spanning-tree loopguard default	打开所有接口的 loop guard
	spanning-tree guard loop	打开某个接口的 loop guard
	spanning-tree guard none	关闭某个接口的 guard 特性
配置BPDU透传功能	 可选配置。用于配置 BPDU 透传功能。	
	bridge-frame forwarding protocol bpdu	打开 BPDU 透传功能

配置BPDU TUNNEL	 可选配置。用于配置 BPDU TUNNEL 功能。	
	<code>l2protocol-tunnel stp</code>	全局使能 BPDU TUNNEL 功能
	<code>l2protocol-tunnel stp enable</code>	接口使能 BPDU TUNNEL 功能
	<code>l2protocol-tunnel stp tunnel-dmac</code>	配置 BPDU TUNNEL 的透传地址

8.4.1 打开生成树协议

配置效果

- 打开全局 Spanning Tree 协议，同时设置全局的基本设置
- 配置 Spanning Tree 模式

注意事项

- 缺省情况下，Spanning Tree 协议是关闭的；当打开 Spanning Tree 协议时，设备即开始运行生成树协议，本设备缺省运行的是 MSTP 协议。
- Spanning Tree 协议的缺省模式是 MSTP 模式。
- Spanning Tree 协议与数据中心的 TRILL 协议功能互斥。

配置方法

📌 打开 Spanning Tree 协议

- 必须配置。
- 若无特殊要求，应在每台设备上启动 Spanning Tree 协议。

📌 配置 Spanning Tree 模式

- 可选配置
- 按 802.1 相关协议标准，STP、RSTP、MSTP 这三个版本的 Spanning Tree 协议本来就无须管理员再多做设置，版本间自然会互相兼容。但考虑到有些厂家不完全按标准实现，可能会导致一些兼容性的问题。因此我们提供这么一条命令配置，以供管理员在发现其他厂家的设备与本设备不兼容时，能够切换到低版本的 Spanning Tree 模式，以兼容之。

检验方法

- 显示验证

相关命令

📌 打开 Spanning Tree 协议

【命令格式】 **spanning-tree** [**forward-time** *seconds* | **hello-time** *seconds* | **max-age** *seconds* | **tx-hold-count** *numbers*]

【参数说明】 **forward-time** *seconds* : 端口状态改变的时间间隔, 取值范围为 4-30 秒, 缺省值为 15 秒。
hello-time *seconds* : 设备定时发送 BPDU 报文的时间间隔, 取值范围为 1-10 秒, 缺省值为 2 秒。
max-age *second* : BPDU 报文消息生存的最长时间, 取值范围为 6-40 秒, 缺省值为 20 秒。
tx-hold-count *numbers* : 配置每秒最多发送 BPDU 个数, 取值范围为 1-10 个, 缺省值为 3 个。

【命令模式】 全局配置模式

【使用指导】 **forward-time**、**hello-time**、**max-age** 三个值的范围是相关的, 修改了其中一个会影响到其他两个的值范围。这三个值之间有一个制约关系:

$$2 * (\text{Hello Time} + 1.0 \text{ second}) \leq \text{Max-Age Time} \leq 2 * (\text{Forward-Delay} - 1.0 \text{ second})$$

您配置的这三个参数必须满足这个条件, 否则有可能导致拓扑不稳定, 也会设置不成功。

配置 Spanning Tree 模式

【命令格式】 **spanning-tree mode** [**stp** | **rstp** | **mstp**]

【参数说明】 **stp** : Spanning tree protocol(IEEE 802.1d)
rstp : Rapid spanning tree protocol(IEEE 802.1w)
mstp : Multiple spanning tree protocol(IEEE 802.1s)

【命令模式】 全局配置模式

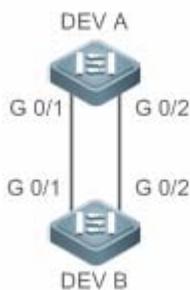
【使用指导】 有些友商产品不完全按标准实现, 可能会导致一些兼容性的问题。在管理员发现其他厂家的设备与本设备不兼容时, 使用此命令可以切换到低版本的 Spanning Tree 模式, 以兼容之。

配置举例

配置 Spanning Tree 协议和定时器参数

【网络环境】

图 8-20



- 【配置方法】
- 设备开启生成树协议, 同时配置生成树协议模式为 STP 协议。
 - 配置根桥 DEV A 的定时器参数为: Hello Time=4s, Max Age=25s, Forward Delay=18s。

DEV A

第一步, 开启生成树协议, 同时配置生成树协议模式为 STP 协议。

```
Ruijie#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#spanning-tree
Ruijie(config)#spanning-tree mode stp
```

第二步，配置根桥 DEV A 的定时器参数

```
Ruijie(config)#spanning-tree hello-time 4
Ruijie(config)#spanning-tree max-age 25
Ruijie(config)#spanning-tree forward-time 18
```

DEV B**第一步，开启生成树协议，同时配置生成树协议模式为 STP 协议。**

```
Ruijie#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#spanning-tree
Ruijie(config)#spanning-tree mode stp
```

【检验方法】

- 通过 `show spanning-tree summary` 查看生成树拓扑和协议配置参数。

DEV A

```
Ruijie#show spanning-tree summary

Spanning tree enabled protocol stp
  Root ID    Priority    0
             Address    00d0.f822.3344
             this bridge is root
             Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec

  Bridge ID  Priority    0
             Address    00d0.f822.3344
             Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec

Interface      Role Sts Cost      Prio  OperEdge Type
-----
Gi0/2          Desg FWD 20000    128   False  P2p
Gi0/1          Desg FWD 20000    128   False  P2p
```

DEV B

```
Ruijie#show spanning-tree summary

Spanning tree enabled protocol stp
  Root ID    Priority    0
             Address    00d0.f822.3344
             this bridge is root
             Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec

  Bridge ID  Priority    32768
             Address    001a.a917.78cc
             Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

Interface      Role Sts Cost      Prio  OperEdge Type
```

Gi0/2	Altn BLK 20000	128	False	P2p Bound (STP)
Gi0/1	Root FWD 20000	128	False	P2p Bound (STP)

常见错误

- 配置生成树协议定时器相关参数，只有在设备选举为生成树的根桥时才起作用。即非根桥的定时器参数是以根桥的定时器参数为准。

8.4.2 配置生成树的兼容性

配置效果

- 配置接口的兼容性模式，可以实现与其它产商之间的互连。
- 配置Protocol Migration进行强制版本检查会影响 RSTP与STP的兼容。

注意事项

- 配置接口的兼容性模式，可以使该端口发送 BPDU 时根据当前端口的属性有选择的携带不同的 MSTI 的信息，以实现与其它产商之间的互连。

配置方法

配置接口的兼容性模式

- 可选配置

配置 Protocol Migration

- 可选配置
- 管理员发现对端设备可支持 RSTP 协议时，可将本设备设置为强制版本检查，强制两对接设备运行 RSTP 协议。

检验方法

- 显示验证。

相关命令

配置接口的兼容性模式

【命令格式】 **spanning-tree compatible enable**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 打开接口的兼容模式，可以使当前端口的接口属性信息有选择性的携带 MSTI 的信息进行发送，以实现与其它产商之间的互连。

配置 Protocol Migration

【命令格式】 **clear spanning-tree detected-protocols** [**interface** *interface-id*]

【参数说明】 **interface** interface-id：对应的接口

【命令模式】 特权模式

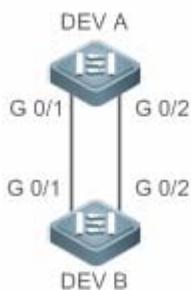
【使用指导】 此命令用来强制接口发送 RSTP BPDU 帧，对 BPDU 帧执行强制检查。

配置举例

配置 Spanning Tree 协议兼容模式

【网络环境】

图 8-21



- 【配置方法】
- 设备 A, B 配置实例 1, 2。实例 1 关联 VLAN 10，实例 2 关联 VLAN 20。
 - 端口 gi 0/1 属于 VLAN 10，gi 0/2 属于 VLAN 20，配置端口的生成树兼容模式。

DEV A

第一步，创建实例 1, 2。实例 1 关联 VLAN 10，实例 2 关联 VLAN 20。

```
Ruijie#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#spanning-tree mst configuration
Ruijie(config-mst)#instance 1 vlan 10
Ruijie(config-mst)#instance 2 vlan 20
```

第二步，配置端口所属的 VLAN，同时开启端口的生成树兼容模式。

```
Ruijie(config)#int gi 0/1
Ruijie(config-if-GigabitEthernet 0/1)#switchport access vlan 10
Ruijie(config-if-GigabitEthernet 0/1)#spanning-tree compatible enable
Ruijie(config-if-GigabitEthernet 0/1)#int gi 0/2
Ruijie(config-if-GigabitEthernet 0/2)#switchport access vlan 20
Ruijie(config-if-GigabitEthernet 0/2)#spanning-tree compatible enable
```

DEV B

同 DEV A。

- 【检验方法】 ● 通过 show spanning-tree summary 查看生成树拓扑计算是否正确。

DEV A

```
Ruijie#show spanning-tree summary

Spanning tree enabled protocol mstp
MST 0 vlans map : 1-9, 11-19, 21-4094
  Root ID   Priority   32768
           Address   001a.a917.78cc
           this bridge is root
           Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

  Bridge ID Priority   32768
           Address   001a.a917.78cc
           Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

Interface      Role Sts Cost      Prio   OperEdge Type
-----
Gi0/2          Desg FWD 20000    128    False   P2p
Gi0/1          Desg FWD 20000    128    False   P2p

MST 1 vlans map : 10
  Region Root Priority   32768
           Address   001a.a917.78cc
           this bridge is region root

  Bridge ID Priority   32768
           Address   001a.a917.78cc

Interface      Role Sts Cost      Prio   OperEdge Type
-----
Gi0/1          Desg FWD 20000    128    False   P2p

MST 2 vlans map : 20
  Region Root Priority   32768
           Address   001a.a917.78cc
           this bridge is region root

  Bridge ID Priority   32768
           Address   001a.a917.78cc

Interface      Role Sts Cost      Prio   OperEdge Type
```

DEV B

```

-----
Gi0/2          Desg FWD 20000    128    False   P2p
Ruijie#show spanning-tree summary

Spanning tree enabled protocol mstp
MST 0 vlans map : 1-9, 11-19, 21-4094
  Root ID      Priority    32768
                Address     001a.a917.78cc
                this bridge is root
                Hello Time  2 sec  Forward Delay 15 sec  Max Age 20 sec

  Bridge ID    Priority    32768
                Address     00d0.f822.3344
                Hello Time  4 sec  Forward Delay 18 sec  Max Age 25 sec

Interface      Role Sts Cost      Prio    OperEdge Type
-----
Gi0/2          Altn BLK 20000    128    False   P2p
Gi0/1          Root FWD 20000    128    False   P2p

MST 1 vlans map : 10
  Region Root Priority    32768
                Address     001a.a917.78cc
                this bridge is region root

  Bridge ID    Priority    32768
                Address     00d0.f822.3344

Interface      Role Sts Cost      Prio    OperEdge Type
-----
Gi0/1          Root FWD 20000    128    False   P2p

MST 2 vlans map : 20
  Region Root Priority    32768
                Address     001a.a917.78cc
                this bridge is region root

  Bridge ID    Priority    32768
                Address     00d0.f822.3344

Interface      Role Sts Cost      Prio    OperEdge Type
-----

```

Gi0/2	Root FWD 20000	128	False	P2p
-------	----------------	-----	-------	-----

常见错误

- 配置端口的兼容模式，需要关注端口的 VLAN 裁剪信息。建议链路两端的端口 VLAN 列表配置一致。

8.4.3 配置MSTP Region

配置效果

- 配置 MSTP Region 可以改变哪些设备处于同一个 MSTP Region 内，从而影响网络拓扑。

注意事项

- 要让多台设备处于同一个 MSTP Region，就要让这几台设备有相同的名称（Name）、相同的 Revision Number、相同的 Instance—Vlan 对应表。
- 可以配置 0 - 64 号 Instance 包含哪些 Vlan，剩下的 Vlan 就自动分配给 Instance 0。一个 Vlan 只能属于一个 Instance。
- 建议您在关闭 STP 的模式下配置 Instance—Vlan 的对应表，配置好后再打开 MSTP，以保证网络拓扑的稳定和收敛。

配置方法

▾ 配置 MSTP Region

- 可选配置
- 要让多台设备处于同一个 MSTP Region 时配置。

检验方法

- 显示验证。

相关命令

▾ 进入 MSTP Region 配置模式

【命令格式】 **spanning-tree mst configuration**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 进入 MST 配置模式后，

▾ 配置 MST Instance 与 Vlan 的对应关系

【命令格式】 **instance** *instance-id* **vlan** *vlan-range*

【参数说明】 *instance-id* : MST Instance ID , 范围为 0 - 64。

vlan-range : VLAN ID , 范围为 1 - 4094。

【命令模式】 MST 配置模式

【使用指导】 把 vlan 组添加到一个 MST instance 中使用此命令。

举例来说：

instance 1 vlan 2-200 就是把 vlan 2 到 vlan 200 都添加到 instance 1 中。

instance 1 vlan 2,20,200 就是把 vlan 2、vlan 20 ,vlan 200 添加到 instance 1 中。

同样，您可以用 no 命令把 vlan 从 instance 中删除，删除的 vlan 自动转入 instance 0。

配置 MST 名称

【命令格式】 **name** *name*

【参数说明】 *name* : MST 配置名称，该字符串最多可以有 32 个字节。

【命令模式】 MST 配置模式

【使用指导】 -

配置 MST 版本号

【命令格式】 **revision** *version*

【参数说明】 *version* : 指定 MST revision number , 范围为 0 - 65535。缺省值为 0

【命令模式】 MST 配置模式

【使用指导】 -

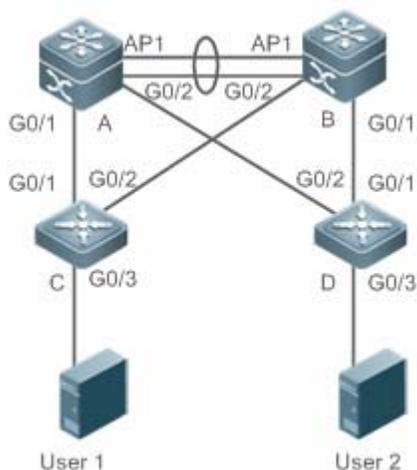
配置举例

i 以下配置举例，仅介绍与 MSTP 和 VRRP 相关的配置。

在 MSTP+VRRP 拓扑中，配置 MSTP 协议，实现 VLAN 的负载均衡

【网络环境】

图 8-22



- 【配置方法】**
- 在交换机 A, B, C, D 上, 打开 MSTP 协议, 创建实例 1, 2。
 - 配置交换机 A 为 MSTP 的实例 0 和 1 的根桥, 交换机 B 为实例 2 的根桥。
 - 配置交换机 A 为 VLAN 1,10 的 VRRP 的 Master 设备, 交换机 B 为 VLAN 20 的 VRRP 的 Master 设备。

A

第一步, 配置 VLAN 10, 20, 同时设备互联端口配置成 Trunk 口

```
A(config)#vlan 10
A(config-vlan)#vlan 20
A(config-vlan)#exit
A(config)#int range gi 0/1-2
A(config-if-range)#switchport mode trunk
A(config-if-range)#int ag 1
A(config-if-AggregatePort 1)# switchport mode trunk
```

第二步, 打开 MSTP, 同时创建实例 1, 2

```
A(config)#spanning-tree
A(config)# spanning-tree mst configuration
A(config-mst)#instance 1 vlan 10
A(config-mst)#instance 2 vlan 20
A(config-mst)#exit
```

第三步, 配置设备 A 为实例 0 和 1 的根桥

```
A(config)#spanning-tree mst 0 priority 4096
A(config)#spanning-tree mst 1 priority 4096
A(config)#spanning-tree mst 2 priority 8192
```

第四步, 配置 VRRP 的优先级, 使设备 A 为 VLAN 10 的 VRRP Master 设备, 同时配置 VRRP 虚网关 IP 地址

```
A(config)#interface vlan 10
A(config-if-VLAN 10)ip address 192.168.10.2 255.255.255.0
A(config-if-VLAN 10) vrrp 1 priority 120
A(config-if-VLAN 10) vrrp 1 ip 192.168.10.1
```

第五步, VRRP 的默认优先级为 100, 使设备 A 为 VLAN 20 的 VRRP Backup 设备

```
A(config)#interface vlan 20
A(config-if-VLAN 20)ip address 192.168.20.2 255.255.255.0
A(config-if-VLAN 20) vrrp 1 ip 192.168.20.1
```

B

第一步, 配置 VLAN 10, 20, 同时设备互联端口配置成 Trunk 口

```
B(config)#vlan 10
B(config-vlan)#vlan 20
B(config-vlan)#exit
B(config)#int range gi 0/1-2
B(config-if-range)#switchport mode trunk
B(config-if-range)#int ag 1
```

```
B(config-if-AggregatePort 1)# switchport mode trunk
```

第二步，打开 MSTP，同时创建实例 1，2

```
B(config)#spanning-tree
B(config)# spanning-tree mst configuration
B(config-mst)#instance 1 vlan 10
B(config-mst)#instance 2 vlan 20
B(config-mst)#exit
```

第三步，配置设备 A 为实例 2 的根桥

```
B(config)#spanning-tree mst 0 priority 8192
B(config)#spanning-tree mst 1 priority 8192
B(config)#spanning-tree mst 2 priority 4096
```

第四步，配置 VRRP 虚网关 IP 地址

```
B(config)#interface vlan 10
B(config-if-VLAN 10)ip address 192.168.10.3 255.255.255.0
B(config-if-VLAN 10) vrrp 1 ip 192.168.10.1
```

第五步，配置 VRRP 的优先级为 120，使设备 B 为 VLAN 20 的 VRRP Master 设备

```
B(config)#interface vlan 20
B(config-if-VLAN 20)vrrp 1 priority 120
B(config-if-VLAN 20)ip address 192.168.20.3 255.255.255.0
B(config-if-VLAN 20) vrrp 1 ip 192.168.20.1
```

C

第一步，配置 VLAN 10, 20，同时设备互联端口配置成 Trunk 口

```
C(config)#vlan 10
C(config-vlan)#vlan 20
C(config-vlan)#exit
C(config)#int range gi 0/1-2
C(config-if-range)#switchport mode trunk
```

第二步，打开 MSTP，同时创建实例 1，2

```
C(config)#spanning-tree
C(config)# spanning-tree mst configuration
C(config-mst)#instance 1 vlan 10
C(config-mst)#instance 2 vlan 20
C(config-mst)#exit
```

第三步，配置设备 C 直接用户的端口为 Portfast 口，同时启用 BPDU Guard。

```
C(config)#int gi 0/3
C(config-if-GigabitEthernet 0/3)#spanning-tree portfast
```

```
C(config-if-GigabitEthernet 0/3)#spanning-tree bpduguard enable
```

D

同设备 C。

【检验方法】

- 通过 show spanning-tree summary 查看生成树拓扑计算的正确性。
- 通过 show vrrp brief 查看 VRRP 主备是否建立成功。

A

```
Ruijie#show spanning-tree summary
```

```
Spanning tree enabled protocol mstp
```

```
MST 0 vlans map : 1-9, 11-19, 21-4094
```

```
  Root ID   Priority   4096
           Address   00d0.f822.3344
           this bridge is root
           Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec
```

```
  Bridge ID Priority   4096
           Address   00d0.f822.3344
           Hello Time 4 sec Forward Delay 18 sec Max Age 25 sec
```

Interface	Role	Sts	Cost	Prio	OperEdge	Type
Ag1	Desg	FWD	19000	128	False	P2p
Gi0/1	Desg	FWD	200000	128	False	P2p
Gi0/2	Desg	FWD	200000	128	False	P2p

```
MST 1 vlans map : 10
```

```
  Region Root Priority   4096
           Address   00d0.f822.3344
           this bridge is region root
```

```
  Bridge ID Priority   4096
           Address   00d0.f822.3344
```

Interface	Role	Sts	Cost	Prio	OperEdge	Type
Ag1	Desg	FWD	19000	128	False	P2p
Gi0/1	Desg	FWD	200000	128	False	P2p
Gi0/2	Desg	FWD	200000	128	False	P2p

```
MST 2 vlans map : 20
```

```
  Region Root Priority   4096
           Address   001a.a917.78cc
```

```

        this bridge is region root

    Bridge ID Priority    8192
           Address    00d0.f822.3344

Interface      Role Sts Cost      Prio    OperEdge Type
-----
Ag1             Root FWD 19000    128     False   P2p
Gi0/1          Desg FWD 200000   128     False   P2p
Gi0/2          Desg FWD 200000   128     False   P2p

```

B Ruijie#show spanning-tree summary

```

Spanning tree enabled protocol mstp
MST 0 vlans map : 1-9, 11-19, 21-4094
  Root ID   Priority    4096
           Address    00d0.f822.3344
           this bridge is root
           Hello Time  4 sec  Forward Delay 18 sec  Max Age 25 sec

  Bridge ID Priority    8192
           Address    001a.a917.78cc
           Hello Time  2 sec  Forward Delay 15 sec  Max Age 20 sec

Interface      Role Sts Cost      Prio    OperEdge Type
-----
Ag1             Root FWD 19000    128     False   P2p
Gi0/1          Desg FWD 200000   128     False   P2p
Gi0/2          Desg FWD 200000   128     False   P2p

MST 1 vlans map : 10
  Region Root Priority    4096
           Address    00d0.f822.3344
           this bridge is region root

  Bridge ID Priority    8192
           Address    001a.a917.78cc

Interface      Role Sts Cost      Prio    OperEdge Type
-----
Ag1             Root FWD 19000    128     False   P2p
Gi0/1          Desg FWD 200000   128     False   P2p

```

```
Gi0/2          Desg FWD 200000    128    False   P2p
```

```
MST 2 vlans map : 20
```

```
  Region Root Priority  4096
                Address  001a.a917.78cc
                this bridge is region root
```

```
  Bridge ID Priority  4096
                Address  001a.a917.78cc
```

```
Interface      Role Sts Cost      Prio    OperEdge Type
-----
```

```
Ag1            Desg FWD 19000    128    False   P2p
```

```
Gi0/1          Desg FWD 200000    128    False   P2p
```

```
Gi0/2          Desg FWD 200000    128    False   P2p
```

C

```
Ruijie#show spanning-tree summary
```

```
Spanning tree enabled protocol mstp
```

```
MST 0 vlans map : 1-9, 11-19, 21-4094
```

```
  Root ID  Priority  4096
            Address  00d0.f822.3344
            this bridge is root
            Hello Time  4 sec Forward Delay 18 sec Max Age 25 sec
```

```
  Bridge ID Priority  32768
            Address  001a.a979.00ea
            Hello Time  2 sec Forward Delay 15 sec Max Age 20 sec
```

```
Interface      Role Sts Cost      Prio    Type  OperEdge
-----
```

```
Fa0/2          Altn BLK 200000    128    P2p   False
```

```
Fa0/1          Root FWD 200000    128    P2p   False
```

```
MST 1 vlans map : 10
```

```
  Region Root Priority  4096
                Address  00d0.f822.3344
                this bridge is region root
```

```
  Bridge ID Priority  32768
                Address  001a.a979.00ea
```

```

Interface          Role Sts Cost      Prio   Type  OperEdge
-----
Fa0/2              Altn BLK 200000   128    P2p   False
Fa0/1              Root FWD 200000   128    P2p   False

```

```
MST 2 vlans map : 20
```

```

Region Root Priority 4096
Address 001a.a917.78cc
this bridge is region root

```

```

Bridge ID Priority 32768
Address 001a.a979.00ea

```

```

Interface          Role Sts Cost      Prio   Type  OperEdge
-----
Fa0/2              Root FWD 200000   128    P2p   False
Fa0/1              Altn BLK 200000   128    P2p   False

```

D 略

常见错误

- MSTP 拓扑中，MST 域的配置建议配置一致。
- 配置实例和 VLAN 的映射关系时，VLAN 没有创建。
- 在 MSTP+VRRP 拓扑中，设备如果运行 STP 或 RSTP 协议，则该设备是按照不同 MST 域的算法进行生成树计算。

8.4.4 配置RSTP快速收敛

配置效果

- 配置 link-type 关系到 RSTP 是否能快速的收敛。

注意事项

- 配置该端口的连接类型是不是“点对点连接”，这一点关系到RSTP是否能快速的收敛。请参照“RSTP的快速收敛”。当您不设置该值时，设备会根据端口的“双工”状态来自动设置的，全双工的端口就设link type为point-to-point，半双工就设为shared。您也可以强制设置link type来决定端口的连接是不是“点对点连接”。

配置方法

配置 link-type

- 可选配置

检验方法

- 显示验证。
- 使用 **show spanning-tree [mst instance-id] interface interface-id** 命令查看生成树接口的配置信息。

相关命令

配置 link-type

【命令格式】 **spanning-tree link-type [point-to-point | shared]**

【参数说明】 **point-to-point**：强制设置该接口的连接类型为 point-to-point
shared：强制设置该接口的连接类型为 shared

【命令模式】 接口配置模式

【使用指导】 配置该端口的连接类型是不是“点对点连接”，这一点关系到 RSTP 是否能快速的收敛。当用户不设置该值时，设备会根据端口的“双工”状态来自动设置的。

配置举例

配置 RSTP 快速收敛

【配置方法】 ● 配置端口的连接类型为点对点网络。

```
Ruijie(config)#int gi 0/1
Ruijie(config-if-GigabitEthernet 0/1)#spanning-tree link-type point-to-point
```

【检验方法】 ● 通过 **show spanning-tree summary** 查看端口连接类型。

```
Ruijie#show spanning-tree summary

Spanning tree enabled protocol mstp
MST 0 vlans map : ALL
  Root ID    Priority    32768
             Address    001a. a917. 78cc
             this bridge is root
             Hello Time  2 sec  Forward Delay 15 sec  Max Age 20 sec

  Bridge ID  Priority    32768
             Address    00d0. f822. 3344
             Hello Time  2 sec  Forward Delay 15 sec  Max Age 20 sec

Interface          Role Sts Cost          Prio   OperEdge Type
-----
-----
```

Gi0/1	Root FWD 20000	128	False	P2p
-------	----------------	-----	-------	-----

常见错误

- 端口的连接类型和速率、双工有关。如果是半双工，则连接类型为 shared。

8.4.5 配置优先级

配置效果

- 设置设备优先级（Switch Priority）关系着到底哪个设备为整个网络的根，同时也关系到整个网络的拓扑结构。
- 设置端口的优先级（Port Priority）关系着到底哪个端口进入 Forwarding 状态。

注意事项

- 建议管理员把核心设备的优先级设得高些（数值小），这样有利于整个网络的稳定。可以给不同的 Instance 分配不同的设备优先级，各个 Instance 可根据这些值运行独立的生成树协议。对于不同 Region 间的设备，它们只关心 CIST(Instance 0) 的优先级。如 Bridge ID 所讲，优先级的设置值有 16 个，都为 4096 的倍数，分别是 0, 4096, 8192, 12288, 16384, 20480, 24576, 28672, 32768, 36864, 40960, 45056, 49152, 53248, 57344, 61440。缺省值为 32768。
- 当有两个端口都连在一个共享介质上，设备会选择一个高优先级（数值小）的端口进入 Forwarding 状态，低优先级（数值大）的端口进入 Discarding 状态。如果两个端口的优先级一样，就选端口号小的那个进入 Forwarding 状态。您可以在一个端口上给不同的 Instance 分配不同的端口优先级，各个 Instance 可根据这些值运行独立的生成树协议。
- 端口优先级和设备优先级一样，可配置的优先级值也有 16 个，都为 16 的倍数，分别是 0, 16, 32, 48, 64, 80, 96, 112, 128, 144, 160, 176, 192, 208, 224, 240。缺省值为 128。

配置方法

配置设备优先级

- 可选配置
- 在管理员需要改变网络的根或者拓扑结构时需要配置设备优先级。

配置端口优先级

- 可选配置
- 在管理员需要改变哪个端口优先进入 Forwarding 状态时配置。

检验方法

- 显示验证
- 使用 `show spanning-tree [mst instance-id] interface interface-id` 命令查看生成树接口的配置信息。

相关命令

配置设备优先级

【命令格式】 **spanning-tree [mst instance-id] priority priority**

【参数说明】 **mst instance-id** : Instance 号, 范围为 0 - 64

priority priority : 设备优先级, 可选用 0, 4096, 8192, 12288, 16384, 20480, 24576, 28672, 32768, 36864, 40960, 45056, 49152, 53248, 57344 和 61440。共 16 个整数, 均为 4096 的倍数。

【命令模式】 全局配置模式

【使用指导】 设置设备的优先级关系到哪个设备为整个网络的根, 同时也关系到整个网络的拓扑结构。

配置端口优先级

【命令格式】 **spanning-tree [mst instance-id] port-priority priority**

【参数说明】 **mst instance-id** : Instance 号, 范围为 0 - 64。

port-priority priority : 端口优先级, 可选用 0, 16, 32, 48, 64, 80, 96, 112, 128, 144, 160, 176, 192, 208, 224, 240, 共 16 个整数, 均为 16 的倍数。

【命令模式】 接口配置模式

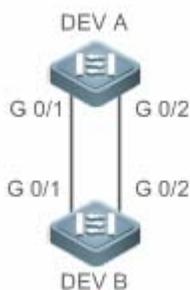
【使用指导】 在 Region 内形成环路时, 优先选择高优先级的端口处于发送状态。优先级相同时, 以选用接口号较小的端口。使用此命令, 这将影响到 Region 内形成环路中的哪个端口会处于发送状态。

配置举例

配置端口优先级

【网络环境】

图 8-23



【配置方法】

- 配置网桥优先级, 使 DEV A 为生成树根桥。
- 配置 DEV A 的端口 gi 0/2 的端口优先级为 16, 使 DEV B 的端口 gi 0/2 选举为根端口。

DEV A

第一步, 打开生成树协议, 配置网桥优先级。

```
Ruijie(config)#spanning-tree
Ruijie(config)#spanning-tree mst 0 priority 0
```

第二步, 配置端口 Gi 0/2 的端口优先级。

```
Ruijie(config)# int gi 0/2
```

```
Ruijie(config-if-GigabitEthernet 0/2)#spanning-tree mst 0 port-priority 16
```

DEV B

```
Ruijie(config)#spanning-tree
```

【检验方法】

- 通过 **show spanning-tree summary** 查看生成树拓扑计算结果。

DEV A

```
Ruijie# Ruijie#show spanning-tree summary
```

```
Spanning tree enabled protocol mstp
```

```
MST 0 vlans map : ALL
```

```

Root ID   Priority   0
          Address   00d0.f822.3344
          this bridge is root
          Hello Time   2 sec   Forward Delay 15 sec   Max Age 20 sec

```

```

Bridge ID Priority   0
          Address   00d0.f822.3344
          Hello Time   2 sec   Forward Delay 15 sec   Max Age 20 sec

```

```
Interface      Role Sts Cost      Prio    OperEdge Type
-----
```

```

Gi0/2          Desg FWD 20000    16      False   P2p
Gi0/1          Desg FWD 20000    128     False   P2p

```

DEV B

```
Ruijie#show spanning-tree summary
```

```
Spanning tree enabled protocol mstp
```

```
MST 0 vlans map : ALL
```

```

Root ID   Priority   0
          Address   00d0.f822.3344
          this bridge is root
          Hello Time   2 sec   Forward Delay 15 sec   Max Age 20 sec

```

```

Bridge ID Priority   32768
          Address   001a.a917.78cc
          Hello Time   2 sec   Forward Delay 15 sec   Max Age 20 sec

```

```
Interface      Role Sts Cost      Prio    OperEdge Type
-----
```

```

Gi0/2          Root FWD 20000    128     False   P2p
Gi0/1          Altn BLK 20000    128     False   P2p

```

常见错误

- 端口优先级只有在指派端口修改才起作用。

8.4.6 配置接口的路径花费

配置效果

- 端口的路径花费 (Path Cost) 会影响端口的转发状态，及影响整个网络的拓扑结构。
- 当某端口 Path Cost 为缺省值时，配置路径花费的计算方法会影响端口的路径花费计算结果。

注意事项

- 设备是根据哪个端口到根桥 (Root Bridge) 的 Path Cost 总和最小而选定 Root Port 的，因此 Port Path Cost 的设置关系到本设备 Root Port。它的缺省值是按 Interface 的链路速率 (The Media Speed) 自动计算的，速率高的花费小，如果管理员没有特别需要可不必更改它，因为这样算出的 Path Cost 最科学。您可以在一个端口上针对不同的 Instance 分配不同的路径花费，各个 Instance 可根据这些值运行独立的生成树协议。
- 当该端口 Path Cost 为缺省值时，设备会自动根据端口速率计算出该端口的 Path Cost。但 IEEE 802.1d-1998 和 IEEE 802.1t 对相同的链路速率规定了不同 Path Cost 值，802.1d-1998 的取值范围是短整型 (short) (1—65535)，802.1t 的取值范围是长整型 (long) (1—200,000,000)。其中对于 AP 的 Cost 值有两个方案：我司的私有方案固定为物理口的 Cost 值*95%；标准推荐的方案为 20,000,000,000/(AP 的实际链路带宽)，其中 AP 的实际链路带宽为成员口的带宽*UP 成员口个数。请管理员一定要统一好整个网络内 Path Cost 的标准。缺省模式为私有长整型模式。
- 下表列出两种方法对不同链路速率自动设置的 Path Cost。

端口速率	Interface	IEEE 802.1d (short)	IEEE 802.1t (long)	IEEE 802.1t (long standard)
10M	普通端口	100	2000000	2000000
	Aggregate Link	95	1900000	2000000 ÷ linkupcnt
100M	普通端口	19	200000	200000
	Aggregate Link	18	190000	200000 ÷ linkupcnt
1000M	普通端口	4	20000	20000
	Aggregate Link	3	19000	20000 ÷ linkupcnt
10000M	普通端口	2	2000	2000
	Aggregate Link	1	1900	20000 ÷ linkupcnt

- 默认采用我司的私有长整型模式。修改成标准推荐方案的 path cost 方案后，AP 的 cost 会随着 UP 成员口数量的变化而变化，而端口 cost 值变化会导致网络拓扑发生变化。
- AP 为静态 AP 时，表格中的 linkupcnt 为 UP 成员口个数；AP 为 LACP AP 时，表格中的 linkupcnt 为参与 AP 数据转发的成员口个数；当 AP 内没有任何成员口 linkup 时，linkupcnt 为 1。具体 AP 和 LACP 的配置，请参见 AP 章节的说明。

配置方法

▾ 配置端口的路径花费

- 可选配置

- 在管理员需要数据报文优先走哪个端口或哪条路径时配置。

配置 Path Cost 的缺省计算方法

- 可选配置
- 在管理员需要修改路径花费计算方式时配置。

检验方法

- 显示验证。
- 使用 **show spanning-tree [mst instance-id] interface interface-id** 命令查看生成树接口的配置信息。

相关命令

配置端口的路径花费

- 【命令格式】 **spanning-tree [mst instance-id] cost cost**
- 【参数说明】 **mst instance-id** : Instance 号, 范围为 0 - 64
cost cost : 路径花费值, 范围为 1 - 200, 000, 000
- 【命令模式】 接口配置模式
- 【使用指导】 cost 值越大表明路径花费越高。

配置 Path Cost 的缺省计算方法

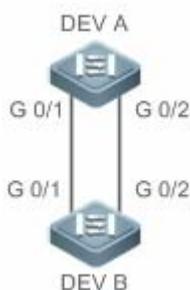
- 【命令格式】 **spanning-tree pathcost method { long [standard] | short }**
- 【参数说明】 **long** : 采用 802.1t 标准设定 path-cost 的值。
standard : standard 表示按照标准推荐的公式计算 cost 值。
short : 采用 802.1d 标准设定 path-cost 的值。
- 【命令模式】 全局配置模式
- 【使用指导】 当该端口 Path Cost 为缺省值时, 设备会自动根据端口速率计算出该端口的 Path Cost。

配置举例

配置端口的路径花费

【网络环境】

图 8-24



【配置方法】

- 配置网桥优先级，使 DEV A 为生成树根桥。
- 配置 DEV B 的端口 gi 0/2 的端口路径花费为 1，使端口 gi 0/2 选举为根端口。

DEV A

```
Ruijie(config)#spanning-tree
Ruijie(config)#spanning-tree mst 0 priority 0
```

DEV B

```
Ruijie(config)#spanning-tree
Ruijie(config)# int gi 0/2
Ruijie(config-if-GigabitEthernet 0/2)# spanning-tree cost 1
```

【检验方法】

- 通过 **show spanning-tree summary** 查看生成树拓扑计算结果。

DEV A

```
Ruijie# Ruijie#show spanning-tree summary

Spanning tree enabled protocol mstp
MST 0 vlans map : ALL
  Root ID   Priority   0
           Address   00d0.f822.3344
           this bridge is root
           Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

  Bridge ID Priority   0
           Address   00d0.f822.3344
           Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

Interface      Role Sts Cost      Prio   OperEdge Type
-----
Gi0/2          Desg FWD 20000    128    False   P2p
Gi0/1          Desg FWD 20000    128    False   P2p
```

DEV B

```
Ruijie#show spanning-tree summary

Spanning tree enabled protocol mstp
MST 0 vlans map : ALL
  Root ID   Priority   0
           Address   00d0.f822.3344
           this bridge is root
           Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

  Bridge ID Priority   32768
           Address   001a.a917.78cc
           Hello Time 2 sec Forward Delay 15 sec Max Age 20 sec

Interface      Role Sts Cost      Prio   OperEdge Type
```

Gi0/2	Root FWD 1	128	False	P2p
Gi0/1	Altn BLK 20000	128	False	P2p

常见错误

- 修改端口路径花费，只有在接收端口配置才起作用。

8.4.7 配置BPDU帧的最大跳数

配置效果

- 配置 BPDU 帧的最大跳数（Maximum-Hop Count），会影响 BPDU 的生命期，从而影响网络拓扑。

注意事项

- BPDU 帧最大跳数的缺省值是 20，一般不需要进行修改。

配置方法

配置 Maximum-Hop Count

- 可选配置。如果网络拓扑规模较大，使得 BPDU 帧的传递超过了默认的 20 跳，则建议更改 max-hops 配置。

检验方法

- 显示验证。

相关命令

设置 BPDU 帧的最大跳数

【命令格式】 **spanning-tree max-hops** *hop-count*

【参数说明】 *hop-count*：BPDU 在被丢弃之前可以经过设备的次数，范围为 1 - 40

【命令模式】 全局配置模式

【使用指导】 在 Region 内，Root Bridge 发送的 BPDU 包含一个 Hop Count 项，从 Root Bridge 开始，每经过一个设备，Hop Count 就会减 1，直到为 0 则表示该 BPDU 信息超时，设备收到 Hops 值为 0 的 BPDU 就要丢弃它。此命令指定了 BPDU 在一个 Region 内经过多少台设备后被丢弃。改变 max-hops 将影响到所有 Instance。

配置举例

设置 BPDU 帧的最大跳数

- 【配置方法】
- 配置 BPDU 帧的最大跳数为 25。

```
Ruijie(config)# spanning-tree max-hops 25
```

- 【检验方法】
- 通过 show spanning-tree 命令查看配置。

```
Ruijie# show spanning-tree
StpVersion : MSTP
SysStpStatus : ENABLED
MaxAge : 20
HelloTime : 2
ForwardDelay : 15
BridgeMaxAge : 20
BridgeHelloTime : 2
BridgeForwardDelay : 15
MaxHops: 25
TxHoldCount : 3
PathCostMethod : Long
BPDUGuard : Disabled
BPDUFilter : Disabled
LoopGuardDef : Disabled

##### mst 0 vlans map : ALL
BridgeAddr : 00d0.f822.3344
Priority: 0
TimeSinceTopologyChange : 2d:0h:46m:4s
TopologyChanges : 25
DesignatedRoot : 0.001a.a917.78cc
RootCost : 0
RootPort : GigabitEthernet 0/1
CistRegionRoot : 0.001a.a917.78cc
CistPathCost : 20000
```

常见错误

无

8.4.8 配置接口port fast的相关特性

配置效果

- 打开 Port Fast 后该端口会直接 Forwarding。但会因为收到 BPDU 而使 Port Fast Operational State 为 disabled，从而正常的参与 STP 算法而 Forwarding。
- 端口打开 BPDU Guard 后，如果在该端口上收到 BPDU，则会进入 Error-disabled 状态。
- 打开 BPDU Filter 后，相应端口会既不发，也不收 BPDU。

注意事项

- 打开某接口的 port fast，全局的 BPDU guard 配置才生效。
- 打开全局的 BPDU Filter enabled 状态下，Port Fast enabled 的 Interface 将既不收 BPDU，也不发 BPDU，这样，直连 Port Fast enabled 端口的主机就收不到 BPDU。而如果 Port Fast enabled 的 Interface 因收到 BPDU 而使 Port Fast Operational 状态 disabled，BPDU Filter 也就自动失效。
- 打开某接口的 port fast，全局的 BPDU filter 配置才生效。

配置方法

配置 port fast

- 可选配置
- 如果设备的端口直连着网络终端，那么就可以设置该端口为 Port Fast。

打开 BPDU Guard

- 可选配置
- 如果设备的端口直连着网络终端，为了防止受到 BPDU 攻击导致生成树拓扑发生异常，可以在这些端口上配置 BPDU Guard 功能。开启 BPDU Guard 的端口收到 BPDU，端口会进入 Error-disabled 状态。
- 如果设备的端口直连着网络终端，为了防止端口下连出现环路，也可以配置 BPDU Guard 功能防止环路。该应用依赖下连设备（比如 HUB）能够转发 BPDU 帧。

打开 BPDU Filter

- 可选配置
- 为了防止异常的 BPDU 报文对生成树拓扑的影响，可以在端口配置 BPDU Filter 功能过滤掉这些异常的 BPDU。

检验方法

- 显示验证。
- 使用 **show spanning-tree [mst instance-id] interface interface-id** 命令查看生成树接口的配置信息。

相关命令

配置接口的 Port Fast

【命令格式】 **spanning-tree portfast**

【参数说明】 -

- 【命令模式】 接口配置模式
- 【使用指导】 打开 Port Fast 后该端口会直接 Forwarding。但因为收到 BPDU 而使 Port Fast Operational State 为 disabled，从而正常的参与 STP 算法而 Forwarding。

📌 配置所有接口的 BPDU Guard

- 【命令格式】 **spanning-tree portfast bpduguard default**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 打开 BPDU guard，如果在该端口上收到 BPDU，则会进入 error-disabled 状态。使用 show spanning-tree 命令查看设置。

📌 配置某个接口的 BPDU Guard

- 【命令格式】 **spanning-tree bpduguard enabled**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 打开单个接口的 BPDU Guard 的情况下，如果该接口收到了 BPDU，就进入 Error-disabled 状态。

📌 配置所有接口的 BPDU Filter

- 【命令格式】 **spanning-tree portfast bpdufilter default**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 打开 BPDU Filter 后，相应端口会既不发也不收 BPDU。

📌 配置某个接口的 BPDU Filter

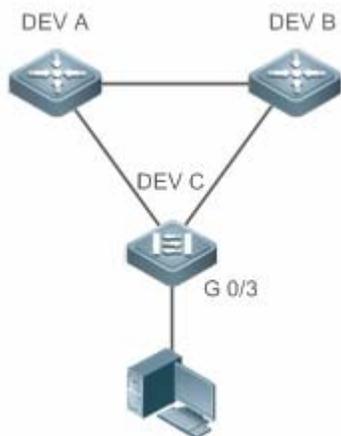
- 【命令格式】 **spanning-tree bpdufilter enabled**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 打开 BPDU Filter 后，相应端口会既不发 BPDU，也不收 BPDU。

配置举例

📌 配置端口的 Port Fast 特性

【网络环境】

图 8-25



【配置方法】

- 配置 DEV C 的端口 gi 0/3 为 Port Fast 端口，同时开启 BPDU Guard 功能。

DEV C

```

Ruijie(config)# int gi 0/3
Ruijie(config-if-GigabitEthernet 0/3)# spanning-tree portfast
%Warning: portfast should only be enabled on ports connected to a single
host. Connecting hubs, switches, bridges to this interface when portfast is
enabled, can cause temporary loops.
Ruijie(config-if-GigabitEthernet 0/3)#spanning-tree bpduguard enable
  
```

【检验方法】

- 通过 **show spanning-tree interface** 命令查看端口的配置信息。

DEV C

```

Ruijie#show spanning-tree int gi 0/3

PortAdminPortFast : Enabled
PortOperPortFast : Enabled
PortAdminAutoEdge : Enabled
PortOperAutoEdge : Enabled
PortAdminLinkType : auto
PortOperLinkType : point-to-point
PortBPDUGuard : Enabled
PortBPDUFilter : Disabled
PortGuardmode : None

##### MST 0 vlans mapped :ALL
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 0.00d0.f822.3344
PortDesignatedCost : 0
PortDesignatedBridge :0.00d0.f822.3344
PortDesignatedPortPriority : 128
  
```

```
PortDesignatedPort : 4
PortForwardTransitions : 1
PortAdminPathCost : 20000
PortOperPathCost : 20000
Inconsistent states : normal
PortRole : designatedPort
```

常见错误

无

8.4.9 配置TC相关的特性

配置效果

- 打开 TC Protection 功能时，收到 TC-BPDU 报文后的一定时间内（一般为 4 秒），只进行一次删除操作。这样可以避免频繁的删除 MAC 地址表项和 ARP 表项。
- 打开 TC Guard 后，当一个端口收到 TC 报文的时候，该端口将屏蔽掉该端口接收或者是自己产生的 TC 报文，使得 TC 报文不会扩散到其它端口，这样能有效控制网络中可能存在的 TC 攻击，保持网络的稳定。
- TC 过滤是指对于端口收到的 TC 报文不处理，而正常的拓扑变化的情况，能够处理。

注意事项

- 建议在确认网络当中有非法的 tc 报文攻击的情况下再打开 TC Guard 功能。

配置方法

📌 打开 TC Protection 功能

- 可选配置
- 缺省是关闭的。

📌 打开 TC Guard 功能

- 可选配置
- 缺省是关闭的。
- 需要过滤掉端口收到的 TC 报文或端口因拓扑变化自己产生的 TC 报文时，可以配置端口的 TC Guard 功能。

📌 打开 TC 过滤功能

- 可选配置
- 缺省是关闭的。

- 只需要过滤掉端口收到的 TC 报文时，可以配置端口的 TC 过滤功能。

检验方法

- 显示验证。

相关命令

▾ 打开 tc protection

- 【命令格式】 **spanning-tree tc-protection**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 -

▾ 配置所有接口的 tc guard

- 【命令格式】 **spanning-tree tc-protection tc-guard**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 启用 tc-guard 功能，能防止 tc 报文的扩散。

▾ 配置某个接口的 tc guard

- 【命令格式】 **spanning-tree tc-guard**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 启用 tc-guard 功能，能防止 tc 报文的扩散。

▾ 配置某个接口的 tc 过滤

- 【命令格式】 **spanning-tree ignore tc**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 启用 tc 过滤功能，则端口收到的 TC 报文将不处理。

配置举例

▾ 配置端口的 TC Guard 功能

- 【配置方法】 ● 配置端口的 TC Guard 功能

```
Ruijie(config)#int gi 0/1
Ruijie(config-if-GigabitEthernet 0/1)#spanning-tree tc-guard
```

- 【检验方法】 ● 通过 **show run interface** 命令查看端口的 TC Guard 配置。

```
Ruijie#show run int gi 0/1
```

```
Building configuration...
Current configuration : 134 bytes

interface GigabitEthernet 0/1
 switchport mode trunk
 spanning-tree tc-guard
```

常见错误

- 错误地配置 TC Guard 或 TC 过滤功能，可能会导致网络设备报文转发出错。比如在拓扑发生变化的情况下，没有及时清除 MAC 地址导致报文转发出错。

8.4.10 配置BPDU源MAC检查

配置效果

- 打开 BPDU 源 MAC 检查开关，将只接受源 MAC 地址为指定 MAC 的 BPDU 帧，过滤掉其它所有接收的 BPDU 帧。

注意事项

- 当确定了某端口点对点链路对端相连的交换机时，可以配置 BPDU 源 MAC 检查来达到只接收对端交换机发送的 BPDU 帧。

配置方法

📌 打开 BPDU 源 MAC 检查

- 可选配置
- 缺省是关闭的。
- 为了防止恶意的 BPDU 攻击，可以配置 BPDU 源 MAC 检查功能。

检验方法

- 显示验证。

相关命令

📌 打开某个接口的 bpdu 源 mac 检查

【命令格式】 **bpdu src-mac-check H.H.H**

- 【参数说明】 *H.H.H* : 表示只接收源 mac 地址为该地址的 bpdu 帧。
- 【命令模式】 接口模式
- 【使用指导】 使用 BPDU 源 MAC 检查是为了防止通过人为发送 BPDU 报文来恶意攻击交换机而使 MSTP 工作不正常。当确定了某端口点对点链路对端相连的交换机时,可通过配置 BPDU 源 MAC 检查来达到只接收对端交换机发送的 BPDU 帧,丢弃所有其他 BPDU 帧,从而达到防止恶意攻击。
- 可以在 interface 模式下来为特定的端口配置相应的 BPDU 源 MAC 检查 MAC 地址,且一个端口只允许配置一个过滤 MAC 地址。

配置举例

配置端口的 BPDU 源 MAC 检查功能

- 【配置方法】 ● 配置端口的 BPDU 源 MAC 检查

```
Ruijie(config)#int gi 0/1
Ruijie(config-if-GigabitEthernet 0/1)#bpdu src-mac-check 00d0.f800.1234
```

- 【检验方法】 ● 通过 **show run interface** 命令查看端口的 Spanning Tree 配置。

```
Ruijie#show run int gi 0/1

Building configuration...
Current configuration : 170 bytes

interface GigabitEthernet 0/1
 switchport mode trunk
 bpdu src-mac-check 00d0.f800.1234
 spanning-tree link-type point-to-point
```

常见错误

- 配置 BPDU 源 MAC 检查,是只接收以配置的 MAC 为源 MAC 的 BPDU 帧,而丢弃其它所有 BPDU 帧。

8.4.11 配置边缘口的自动识别

配置效果

- 打开边缘口自动识别功能时,如果在一定的时间范围内(为 3 秒),指派口没有收到 BPDU,则自动识别为边缘口。但会因为收到 BPDU 而使 Port Fast Operational State 为 disabled。

注意事项

- 一般情况下不需要关闭边缘口自动识别功能。

配置方法

打开边缘口的自动识别

- 可选配置
- 缺省是打开的。

检验方法

- 显示验证。

相关命令

打开边缘口的自动识别

【命令格式】 **spanning-tree autoedge**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 指派口在一定的时间内(为 3 秒),如果收不到下游端口发送的 BPDU,则认为该端口相连的是一台网络设备,从而设置该端口为边缘端口,直接进入 Forwarding 状态。自动标识为边缘口的端口因收到 BPDU 而自动识别为非边缘口。

可以通过 **spanning-tree autoedge disabled** 命令取消边缘口的自动识别功能。

配置举例

关闭端口的 Auto Edge 功能

- 【配置方法】
- 关闭端口的 Auto Edge 功能

```
Ruijie(config)#int gi 0/1
Ruijie(config-if-GigabitEthernet 0/1)#spanning-tree autoedge disabled
```

- 【检验方法】
- 通过 **show spanning-tree interface** 命令查看端口的 Spanning Tree 配置。

```
Ruijie#show spanning-tree interface gi 0/1

PortAdminPortFast : Disabled
PortOperPortFast : Disabled
PortAdminAutoEdge : Disabled
PortOperAutoEdge : Disabled
PortAdminLinkType : point-to-point
PortOperLinkType : point-to-point
PortBPDUGuard : Disabled
PortBPDUFilter : Disabled
```

```
PortGuardmode : None

##### MST 0 vlans mapped :ALL
PortState : forwarding
PortPriority : 128
PortDesignatedRoot : 0.00d0.f822.3344
PortDesignatedCost : 0
PortDesignatedBridge :0.00d0.f822.3344
PortDesignatedPortPriority : 128
PortDesignatedPort : 2
PortForwardTransitions : 6
PortAdminPathCost : 20000
PortOperPathCost : 20000
Inconsistent states : normal
PortRole : designatedPort
```

常见错误

- 边缘端口的自动识别功能，默认指派口 3 秒内未接收到 BPDU 就将端口识别成边缘端口并立即 Forward。如果网络环境存在丢包或收发报文延迟现象，建议将端口的自动识别功能关闭。

8.4.12 配置接口保护相关的特性

配置效果

- 接口打开 Root Guard 功能时，强制其在所有实例上的端口角色为指定端口，一旦该端口收到优先级更高的配置信息时，Root Guard 功能会将该接口置为 root-inconsistent (blocked)状态,在足够长的时间内没有收到更优的配置信息时，端口会恢复成原来的正常状态。
- 由于单向链路的故障，根口或备份口由于收不到 BPDU 会变成指派口进入转发状态，从而导致了网络中环路产生，LOOP Guard 功能防止了这种情况的发生。

注意事项

- 端口的 ROOT Guard 和 LOOP Guard 同一时刻只能有一个生效。

配置方法

📌 打开 ROOT Guard 特性

- 可选配置。

- 为了防止因维护人员的错误配置或网络中的恶意攻击，根桥可能收到优先级更高的配置信息，从而失去当前根桥的位置，引起网络拓扑的错误的变动，可以在设备的指派端口上配置 ROOT Guard 功能。

📄 打开 LOOP Guard 特性

- 可选配置。
- 为了防止接收端口（根端口、Master 端口或 Alternate 端口）因接收不到指派网桥发送的 BPDU 而使网络拓扑发生变化，从而引起可能的环路，可以在上述接收端口上配置 LOOP Guard 功能，提高设备的稳定性。

📄 关闭 Guard 特性

- 可选配置。
- 缺省是关闭的。

检验方法

- 显示验证。

相关命令

📄 打开某个接口的 root guard 特性

- 【命令格式】 **spanning-tree guard root**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 启用 root guard 功能，能防止因错误配置或非法报文的攻击导致当前根桥地位的变化。

📄 打开所有接口的 loop guard 特性

- 【命令格式】 **spanning-tree loopguard default**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 启用 loop guard 功能，能防止根端口或备份口因收不到 bpdu 而产生的可能的环路。

📄 打开某个接口的 loop guard 特性

- 【命令格式】 **spanning-tree guard loop**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 启用 loop guard 功能，能防止根端口或备份口因收不到 bpdu 而产生的可能的环路。

📄 关闭某个接口的 guard 特性

- 【命令格式】 **spanning-tree guard none**
- 【参数说明】 -
- 【命令模式】 接口配置模式

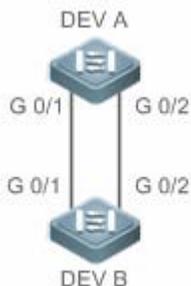
【使用指导】 缺省是关闭 guard 功能。

配置举例

配置端口的 Loop Guard 特性

【网络环境】

图 8-26



- 【配置方法】
- 配置 DEV A 为生成树根桥，DEV B 为非根桥。
 - 配置 DEV B 的端口 gi 0/1 和 gi 0/2 的 LOOP Guard 特性。

DEV A

```
Ruijie(config)#spanning-tree
Ruijie(config)#spanning-tree mst 0 priority 0
```

DEV B

```
Ruijie(config)#spanning-tree
Ruijie(config)# int range gi 0/1-2
Ruijie(config-if-range)#spanning-tree guard loop
```

- 【检验方法】
- 通过 **show spanning-tree interface** 命令查看端口的 Spanning Tree 配置。

DEV A 略

DEV B

```
Ruijie#show spanning-tree int gi 0/1
```

```
PortAdminPortFast : Disabled
PortOperPortFast : Disabled
PortAdminAutoEdge : Enabled
PortOperAutoEdge : Disabled
PortAdminLinkType : auto
PortOperLinkType : point-to-point
PortBPDUGuard : Disabled
PortBPDUFilter : Disabled
PortGuardmode : Guard loop

##### MST 0 vlans mapped :ALL
PortState : forwarding
PortPriority : 128
```

```
PortDesignatedRoot : 0.001a.a917.78cc
PortDesignatedCost : 0
PortDesignatedBridge :0.001a.a917.78cc
PortDesignatedPortPriority : 128
PortDesignatedPort : 17
PortForwardTransitions : 1
PortAdminPathCost : 20000
PortOperPathCost : 20000
Inconsistent states : normal
PortRole : rootPort

Ruijie#show spanning-tree int gi 0/2

PortAdminPortFast : Disabled
PortOperPortFast : Disabled
PortAdminAutoEdge : Enabled
PortOperAutoEdge : Disabled
PortAdminLinkType : auto
PortOperLinkType : point-to-point
PortBPDUGuard : Disabled
PortBPDUFilter : Disabled
PortGuardmode : Guard loop

##### MST 0 vlans mapped :ALL
PortState : discarding
PortPriority : 128
PortDesignatedRoot : 0.001a.a917.78cc
PortDesignatedCost : 0
PortDesignatedBridge :0.001a.a917.78cc
PortDesignatedPortPriority : 128
PortDesignatedPort : 18
PortForwardTransitions : 1
PortAdminPathCost : 20000
PortOperPathCost : 20000
Inconsistent states : normal
PortRole : alternatePort
```

常见错误

- 将 ROOT Guard 功能配置在根端口、Master 端口或 Alternate 端口，可能会错误地将端口 BLOCK。

8.4.13 配置BPDU透传功能

配置效果

- 设备未开启 STP 协议时，需要透传 BPDU 帧，使得与之互连的设备之间的生成树计算正常。

注意事项

- BPDU 透传功能只在 STP 协议关闭时才起作用。当 STP 协议打开时，设备不透传 BPDU 帧。

配置方法

▾ 配置 BPDU 透传功能

- 可选配置
- 设备未开启 STP 协议时，如里需要透传 BPDU 帧，则需要配置 BPDU 透传功能。

检验方法

- 显示验证。

相关命令

▾ 配置 BPDU 透传功能

【命令格式】 **bridge-frame forwarding protocol bpdu**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 在 IEEE 802.1Q 标准中，BPDU 的目的 MAC 地址 01-80-C2-00-00-00 是作为保留地址使用的，即遵循 IEEE 802.1Q 标准的设备，对于接收到的 BPDU 帧是不转发的。然而，在实际的网络布署中，可能需要设备能够支持透传 BPDU 帧。例如，设备未开启 STP 协议时，需要透传 BPDU 帧，使得与之互连的设备之间的生成树计算正常。

BPDU 透传功能只在 STP 协议关闭时才起作用。当 STP 协议打开时，设备不透传 BPDU 帧。

配置举例

▾ 配置 BPDU 透传功能

【网络环境】

图 8-27



- DEV A, C 上开启生成树协议, DEV B 未开启生成树协议。

【配置方法】

- DEV B 上配置 BPDU 透传功能, 使得 DEV A, C 之间的 STP 协议能够正确计算。

DEV B

```
Ruijie(config)#bridge-frame forwarding protocol bpdu
```

【检验方法】

- 通过 show run 查看 BPDU 透传功能是否开启。

DEV B

```
Ruijie#show run
```

```
Building configuration...
```

```
Current configuration : 694 bytes
```

```
bridge-frame forwarding protocol bpdu
```

常见错误

无

8.4.14 配置BPDU TUNNEL

配置效果

- 配置 BPDU TUNNEL 功能, 使用户网络的 STP 协议报文能够通过运营商网络进行隧道透传, 用户网络之间的 STP 协议报文的传输对运营商网络不会产生影响, 从而使得用户网络和运营商网络的 STP 协议分开计算, 互不干扰。

注意事项

- 需要全局和接口同时开启 BPDU TUNNEL 功能后, BPDU TUNNEL 功能才生效。

配置方法

▾ 配置 BPDU 透传功能

- 可选配置。在 QINQ 网络中, 如果需要用户网络和运营商网络的 STP 协议分开计算, 互不干扰, 可以通过配置 BPDU TUNNEL 达到效果。

检验方法

- 通过 show l2protocol-tunnel stp 命令查看 BPDU TUNNEL 配置。

相关命令

配置全局使能 BPDU TUNNEL 功能

- 【命令格式】 **I2protocol-tunnel stp**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 BPDU TUNNEL 功能只在全局和接口同时使能的情况下才起作用。

配置接口使能 BPDU TUNNEL 功能

- 【命令格式】 **I2protocol-tunnel stp enable**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 BPDU TUNNEL 功能只在全局和接口同时使能的情况下才起作用。

配置 BPDU TUNNEL 的透传地址

- 【命令格式】 **I2protocol-tunnel stp tunnel-dmac mac-address**
- 【参数说明】 *mac-address* : 需要透传的 STP 协议地址
- 【命令模式】 全局配置模式
- 【使用指导】 BPDU TUNNEL 应用中, 当用户网络的 STP 协议报文进入运营商网络的边缘设备后, 将目的 mac 地址改成私有地址在运营商网络中转发, 到了另外一端边缘设备后, 再将目的 mac 地址改成公有地址回到另一端用户网络, 以达到 STP 协议报文在运营商网络透传的效果。这个私有地址, 即为 BPDU TUNNEL 的透传地址。

 STP 报文可选透传地址范围: 01d0.f800.0005、011a.a900.0005、010f.e200.0003、0100.0ccd.cdd0、0100.0ccd.cdd1、0100.0ccd.cdd2。

 当未配置透传地址时, BPDU TUNNEL 缺省使用的地址为 01d0.f800.0005。

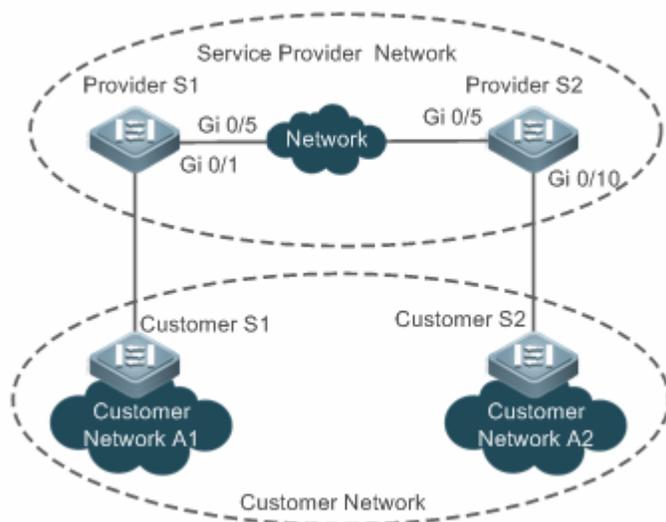
配置举例

 以下配置举例, 仅介绍与 MSTP 和 QINQ 相关的配置。

配置 BPDU TUNNEL 功能

【网络环境】

图 8-28



【配置方法】

- 在运营商网络边缘设备（本例为 Provider S1/Provider S2 上开启基本 QinQ 功能，使用户网络的数据报文在运营商网络的 VLAN 200 中传输。
- 在运营商网络边缘设备（本例为 Provider S1/Provider S2 上开启 STP 协议透传功能，使运营商网络可以通过 BPDU TUNNEL 对用户网络的 STP 报文进行隧道传输。

Provider S1

第一步，创建服务商 VLAN 200

```
Ruijie#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)#vlan 200
Ruijie(config-vlan)#exit
```

第二步，在连接用户网络的接口上开启基本 QinQ 功能，使用 VLAN 200 对用户网络的数据进行隧道传输

```
Ruijie(config)#interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#switchport mode dot1q-tunnel
Ruijie(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel native vlan 200
Ruijie(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel allowed vlan add untagged 200
```

第三步，在连接用户网络的接口上开启 STP 协议透传功能

```
Ruijie(config-if-GigabitEthernet 0/1)#l2protocol-tunnel stp enable
Ruijie(config-if-GigabitEthernet 0/1)#exit
```

第四步，全局开启 STP 协议透传功能

```
Ruijie(config)#l2protocol-tunnel stp
```

第五步，配置 uplink port

```
Ruijie(config)# interface gigabitEthernet 0/5
Ruijie(config-if-GigabitEthernet 0/5)#switchport mode uplink
```

Provider S2

Provider S2 设备上的配置同 Provider S1 配置类似，请参考上文 Provider S1 的配置。此处不再重复说明。

- 【检验方法】
- 查看 BPDU TUNNEL 配置是否正确。
 - 查看 Tunnel 口的配置是否正确，关注点：接口类型是否为 dot1q-tunnel，外层 Tag VLAN 是否为 Native VLAN 且其是否已加入接口的许可 VLAN 列表，运营商网络边缘设备上链口的类型是否为 Uplink。

Provider S1 1：查看 BPDU TUNNEL 配置是否正确：

```
Ruijie#show l2protocol-tunnel stp

L2protocol-tunnel: stp Enable
L2protocol-tunnel destination mac address: 01d0.f800.0005
GigabitEthernet 0/1 l2protocol-tunnel stp enable
```

2：查看 QINQ 配置是否正确：

```
Ruijie#show running-config
interface GigabitEthernet 0/1
  switchport mode dot1q-tunnel
  switchport dot1q-tunnel allowed vlan add untagged 200
  switchport dot1q-tunnel native vlan 200
  l2protocol-tunnel stp enable
  spanning-tree bpdupfilter enable
!
interface GigabitEthernet 0/5
  switchport mode uplink
```

Provider S2 同 Provider S1

常见错误

- 运营商网络中，配置的 BPDU TUNNEL 透传地址要一致，才能正确透传 BPDU 帧。

8.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除端口的收发包统计信息	clear spanning-tree counters [interface <i>interface-id</i>]
清除 STP 的拓扑改变信息	clear spanning-tree mst <i>instance-id</i> topochange record

查看运行情况

作用	命令
----	----

显示 MSTP 的各项参数信息及生成树的拓扑信息	show spanning-tree
显示 MSTP 的收发包统计信息	show spanning-tree counters [interface <i>interface-id</i>]
显示 MSTP 的各 instance 的信息及其端口转发状态信息	show spanning-tree summary
显示因根保护或环路保护而 block 的端口	show spanning-tree inconsistentports
显示 MST 域的配置信息	show spanning-tree mst configuration
显示该 instance 的 MSTP 信息	show spanning-tree mst <i>instance-id</i>
显示指定 interface 的对应 instance 的 MSTP 信息	show spanning-tree mst <i>instance-id</i> interface <i>interface-id</i>
显示指定实例中的接口的拓扑改变信息	show spanning-tree mst <i>instance-id</i> topochange record
显示指定 interface 的所有 instance 的 MSTP 信息	show spanning-tree interface <i>interface-id</i>
显示 forward-time	show spanning-tree forward-time
显示 Hello time	show spanning-tree hello time
显示 max-hops	show spanning-tree max-hops
显示 tx-hold-count	show spanning-tree tx-hold-count
显示 pathcost method	show spanning-tree pathcost method
显示 BPDU TUNNEL 信息	show l2protocol-tunnel stp

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开生成树所有的调试开关	debug mstp all
打开生成树 GR 的调试开关	debug mstp gr
打开接收 BPDU 报文的调试开关	debug mstp rx
打开发送 BPDU 报文的调试开关	debug mstp tx
打开生成树事件调试开关	debug mstp event
打开生成树 Loop Guard 特性调试开关	debug mstp loopguard
打开生成树 Root Guard 特性调试开关	debug mstp rootguard
打开 Bridge Detect 状态机调试开关	debug mstp bridgedetect
打开 Port Information 状态机调试开关	debug mstp portinfo
打开 Port Protocol Migration 状态机调试开关	debug mstp protomigrat
打开生成树拓扑变化的调试开关	debug mstp topochange
打开生成树接收状态机调试开关	debug mstp receive
打开 Port Role Transitions 状态机调试开关	debug mstp roletran
打开 Port State Transition 状态机调试开关	debug mstp statetran
打开生成树发送状态机调试开关	debug mstp transmit

9 GVRP

9.1 概述

GVRP (GARP VLAN Registration Protocol , GARP VLAN 注册协议) 是一种动态配置和扩散VLAN成员关系的GARP (Generic Attribute Registration Protocol , 通用属性注册协议) 应用。

通过GVRP功能,可以简化VLAN配置管理,减少了用户手动配置VLAN和端口加入VLAN的工作,减少因配置不一致导致网络不通的问题的可能性。而且能动态维护VLAN的创建和端口加入/退出VLAN,保证拓扑内VLAN的连通性。

i 下文仅介绍 GVRP 的相关内容。

协议规范

IEEE standard 802.1D

IEEE standard 802.1Q

9.2 典型应用

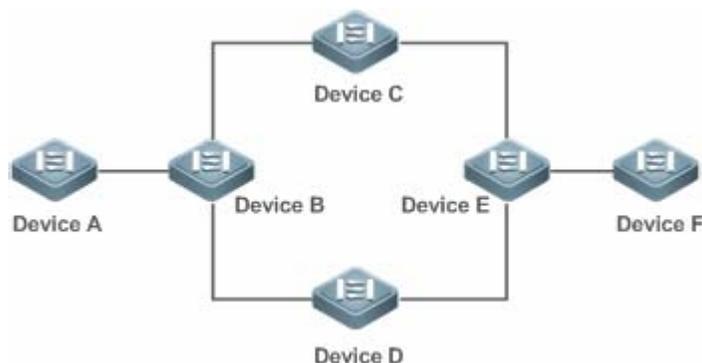
典型应用	场景描述
局域网内配置GVRP	两台交换机局域网内相连,实现 vlan 的同步

9.2.1 局域网内配置GVRP

应用场景

通过启用 GVRP 功能,并配置GVRP 的注册模式为Normal 模式,来实现Device A 和Device F 之间所有动态和静态VLAN的注册和注销。

图 9-1



- 【注释】 Device A-F 为交换机设备，设备之间相连的端口均为 Trunk 口。
Device A 和 Device F 两台交换机配置需要用来通信的静态 VLAN。
交换机 Device A-F 设备上均开启 GVRP 功能。

功能部署

- 所有设备均开启 GVRP 功能，且使能动态创建 vlan 功能，确保中间设备都能创建动态 VLAN。
- 在交换机 Device A 和 Device F 配置需要用来通信的静态 VLAN，交换机 Device B-E 设备上通过 GVRP 协议动态学习这些 VLAN。

 为防止用户拓扑产生环路，建议开启 STP（Spanning Tree Protocol，生成树协议）功能

9.3 功能详解

基本概念

▾ GVRP

GVRP（GARP VLAN Registration Protocol，GARP VLAN 注册协议）就是 GARP 的应用之一，用于注册和注销 VLAN 属性。GVRP 协议实现 VLAN 属性注册和注销的方式如下：

- 当端口收到一个 VLAN 属性的声明时，该端口将注册该声明中所包含的 VLAN 属性（即该端口加入到该 VLAN 中）。
- 当端口收到一个 VLAN 属性的回收声明时，该端口将注销该声明中所包含的 VLAN 属性（即该端口退出该 VLAN）。

图 9-2



▾ 动态 VLAN

可以动态创建和删除，无需用户手动配置的 VLAN 称为动态 VLAN。

可以通过手动配置，将 VLAN 从动态模式切换为静态模式，但无法从静态模式切换为动态模式。

由 GVRP 协议功能创建的动态 VLAN 加入端口由协议状态机控制，即 GVRP 协议创建的动态 VLAN 仅能加入收到 GVRP 注册该 VLAN 的 Trunk 口，而不是所有的 Trunk 口，同样不能手动将端口加入动态 VLAN。

▾ 消息类型

(1) Join 消息

当一个 GARP 应用实体希望其它 GARP 实体注册自己的属性信息时，它会发送 Join 消息；当收到来自其它实体的 Join 消息或由于本实体静态配置了某些属性而需要其它实体进行注册时，它也会发送 Join 消息。Join 消息又分为 JoinEmpty 和 JoinIn 两种，二者的区别如下：

- JoinEmpty：用于声明一个本身没有注册的属性。

- JoinIn：用于声明一个本身已经注册的属性。

(2) Leave 消息

当一个GARP 应用实体希望其它GARP 实体注销自己的属性信息时，它会发送Leave 消息；当收到来自其它实体的Leave 消息或由于本实体静态注销了某些属性而需要其它实体进行注销时，它也会发送Leave 消息。Leave 消息又分为LeaveEmpty 和 LeaveIn 两种，二者的区别如下：

- LeaveEmpty：用于注销一个本身没有注册的属性。
- LeaveIn：用于注销一个本身已经注册的属性。

(3) LeaveAll 消息

每个 GARP 应用实体启动时都会启动各自的LeaveAll 定时器，当该定时器超时后，它就会发送LeaveAll 消息来注销所有的属性，从而使其它GARP 实体重新注册属性信息；当收到来自其它实体的LeaveAll 消息时，它也会发送LeaveAll 消息。在发送LeaveAll 消息同时重新启动LeaveAll定时器，开始新一轮循环。

📌 定时器类型

GARP 定义了四种定时器，用于控制各种GARP 消息的发送。

(1) Hold 定时器

Hold 定时器用来控制GARP 消息（包括Join 消息和Leave 消息）的发送。当GARP 应用实体的属性改变或收到来自其它实体的GARP 消息时，不会立即将该消息发送出去，而是在Hold 定时器超时后，将此时段内待发送的所有GARP 消息封装成尽可能少的报文发送出去，这样就减少了报文的发送数量，从而节省了带宽资源。

(2) Join 定时器

Join 定时器用来控制Join 消息的发送。为了保证Join 消息能够可靠地传输到其它实体，GARP应用实体在发出Join 消息后将等待一个Join 定时器的时间间隔：如果在该定时器超时前收到了其它实体发来的JoinIn 消息，它便不会重发该Join 消息；否则，它将重发一次该Join 消息。并非每个属性都有自己的 Join 定时器，而是每个GARP 应用实体共用一个。

(3) Leave 定时器

Leave 定时器用来控制属性的注销。当GARP 应用实体希望其它实体注销自己的某属性信息时会发送Leave 消息，收到该消息的实体将启动Leave 定时器，只有在该定时器超时前没有收到该属性信息的Join 消息，该属性信息才会被注销。

(4) LeaveAll 定时器

每个 GARP 应用实体启动时都会启动各自的LeaveAll 定时器，当该定时器超时后，GARP 应用实体就会对外发送LeaveAll 消息，从而使其它实体重新注册属性信息。随后再重新启动LeaveAll定时器，开始新一轮的循环。

📌 GVRP 通告模式

GVRP通告模式指的是交换机设备告诉其它互连的设备自己有哪些VLAN，对端设备可能需要创建哪些VLAN并将收发GVRP报文的端口加入相关VLAN。

GVRP的通告模式有两种：

- normal 模式：对外通告本设备上的 VLAN 信息，包括动态和静态 VLAN 信息。
- non- applicant 模式：不对外通告本设备上的 VLAN 信息。

📌 GVRP 注册模式

GVRP注册模式指的是交换机设备收到GVRP报文后，是否处理报文内的VLAN信息，如动态创建不存在的VLAN并将收报文的端口加入VLAN等。

GVRP的通告模式有两种：

- normal 模式：收到 GVRP 报文后，处理报文内的 VLAN 信息。

- disabled 模式：收到 GVRP 报文后，不处理报文内的 VLAN 信息。

功能特性

功能特性	作用
同步拓扑内VLAN信息	同步拓扑内 VLAN 信息，动态创建 VLAN 并将端口动态加入/退出 VLAN，减少用户手动配置工作和减少用户因配置遗漏导致 VLAN 内网络不通的概率。

9.3.1 同步拓扑内vlan信息

工作原理

- GVRP 是 GARP 应用的一种，它基于 GARP 的工作机制来维护设备中的 VLAN 动态注册信息，并将该信息向其它设备传播：当设备启动了 GVRP 之后，就能够接收来自其它设备的 VLAN 注册信息，并动态更新本地的 VLAN 注册信息；此外，设备还能够将本地的 VLAN 注册信息向其它设备传播，从而使同一局域网内所有设备的 VLAN 信息都达成一致。GVRP 传播的 VLAN 注册信息既包括本地手工配置的静态注册信息，也包括来自其它设备的动态注册信息。

↘ 对外通告 VLAN 信息

开启 GVRP 功能的设备上的 Trunk 口会定时收集 Trunk 口内的 VLAN 信息，告诉对端设备本 Trunk 口加入了哪些 VLAN 或者退出哪些 VLAN，通过将这些 VLAN 信息封装在 GVRP 报文内发送给对端设备，对端设备连接的 Trunk 口收到 GVRP 报文后会解析 VLAN 信息，动态的创建 VLAN 并将端口加入 VLAN 或者将端口退出 VLAN。具体会有哪些 VLAN 信息可以见上面将的 GVRP 消息类型。

相关配置

- 缺省功能关闭。
- 使用命令[no] gvrp enable 进行功能开关闭。
- 打开则开启 GVRP 功能，对外发送含 VLAN 信息的 GVRP 报文；关闭则不对外发送含 VLAN 信息的 GVRP 报文，且不处理 GVRP 报文。

↘ 注册/注销 VLAN

交换机设备在收到 GVRP 报文后会根据端口的注册模式选择是否处理 VLAN 信息。具体行为见上面的 GVRP 注册模式解释。

相关配置

- gvrp 功能开启时，trunk 模式端口默认开启注册动态 VLAN 功能。
- 接口上使用命令 gvrp registration mode normal 开启注册动态 VLAN 功能，gvrp register mode disable 关闭注册动态 VLAN 功能

- 打开则收到对端 VLAN 信息后创建动态 VLAN；关闭则收到 GVRP 报文时，不进行添加动态 VLAN 动作。

9.4 配置详解

配置项	配置建议 & 相关命令	
配置GVRP基本功能，同步VLAN信息	⚠ 必须配置。用于使能 GVRP 及允许动态创建 vlan 功能	
	gvrp enable	启动 GVRP 功能
	gvrp dynamic-vlan-creation enable	启动允许动态创建 vlan 功能
	switchport mode trunk	(端口模式下)切换端口模式为 trunk, trunk 模式下的端口 GVRP 功能才生效
	switchport trunk allowed vlan all	允许全部 vlan 通过
	gvrp applicant state	设置端口的通告模式，normal 模式表示对外通告 VLAN 信息，发送 GVRP 报文。否则不对外通告 VLAN 信息。
	gvrp registration mode	设置端口的登记模式，normal 模式表示收到 GVRP 报文后，会处理相关 VLAN 信息，如动态创建 VLAN，将端口加入 VLAN。否则不关心报文内容。
配置GVRP PDU透传功能	⚠ 可选配置。用于设置定时器，端口的登记模式及通告模式。	
	gvrp timer	设置定时器
配置GVRP PDU透传功能	⚠ 可选配置。用于配置 GVRP PDUs 透传功能。	
	bridge-frame forwarding protocol gvrp	打开 GVRP PDUs 透传功能

9.4.1 配置GVRP基本功能，同步VLAN信息

配置效果

- 可以动态创建/删除 VLAN，并将端口动态加入/退出 VLAN。
- 设备之间同步各自的 VLAN 信息，拓扑内通信正常。
- 减少用户手动配置工作，方便 VLAN 管理。

注意事项

- 相互连接进行通信的两台设备都应启动 GVRP，GVRP 信息只在 Trunk Links 中传播，但传播的信息包括当前设备的所有 VLAN 信息，不管 VLAN 是动态学习的，或是手工设置的。
- 在运行 STP (Spanning-tree Protocol) 的情况下，只有状态为 Forwarding 的端口才会参与 GVRP 的运行，如接收、发送 GVRP PDU，只有状态为 Forwarding 的端口的 VLAN 信息会被 GVRP 扩散。

- 所有由 GVRP 添加的 VLAN Port 都是 Tagged Port。
- 所有由 GVRP 动态学习的 VLAN 信息都未保存在系统中，当设备复位时，这些信息将全部丢失。用户也不可以保存这些动态学习到的 VLAN 信息。
- 网络中所有需要交换 GVRP 信息的设备的 GVRP Timers (Join , Leave , Leaveall) 必须保持一致。
- 在未运行 STP (Spanning-tree Protocol , 生成树协议) 的环境，所有可用端口都可以参与 GVRP 的运行。在运行 SST (Single Spanning-tree , 单生成树) 的环境中，只有在当前 SST Context 中处于 Forwarding 状态的端口才参与 GVRP 的运行。在运行 MST (Multi Spanning-tree , 多生成树) 的环境中，GVRP 可在 VLAN 1 所属的 Spanning-tree Context 中运行，用户不能指定其它 Spanning-tree Context。

配置方法

▾ 使能 GVRP 功能

- 必须配置。
- 只有开启该功能，设备才能处理 GVRP 报文

▾ 使能动态创建 VLAN 的功能

- 必须配置。
- 只有开启了该功能，设备在收到 GVRP 的 join 类型报文时，才会动态创建 VLAN。

▾ 切换端口模式为 trunk 模式

必须配置，GVRP 功能只有在 trunk 模式的端口上才生效。

▾ 配置定时器

- 可选配置。
- GVRP 功能有三个定时器，Join timer、Leave timer 和 Leaveall timer，用来控制各种类型报文的发送间隔。
- 三个定时器之间的大小关系：Leave timer 必须大于等于三倍的 Join timer；Leaveall timer 必须大于 Leave timer。
- 三个定时器由 GVRP 状态机控制，并且相互之间可以触发。

▾ 配置端口的通告模式

- 可选配置。
- GVRP 的通告模式有 normal 和 non- applicant 两种，默认为 normal 模式。
- Normal 模式：表示对外通告本设备的 VLAN 信息，如果对端设备连接端口的。
- Non- applicant 模式：表示不对外通告本设备的

▾ 配置端口的登记模式

- 可选配置。
- GVRP 的登记模式有 normal 和 disabled 两种。

▾ 切换端口模式为 trunk 模式

- 必须配置，GVRP 功能只有在 trunk 模式的端口上才生效。
- GVRP 功能仅在 Trunk 口上生效。

检验方法

- **show gvrp configuration** 查看配置信息。
- 查看是否有创建动态 VLAN，并将对应端口加入 VLAN。

相关命令

启动 GVRP

- 【命令格式】 **gvrp enable**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 只有在全局使能允许的情况下GVRP才会启动。在GVRP未全局使能的状态下，其它GVRP参数可以进行配置，但只有在GVRP开始运行时，这些GVRP选项设置才能发生作用。

控制动态 VLAN 的创建

- 【命令格式】 **gvrp dynamic-vlan-creation enable**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 当一个端口接收到的信息（仅限于Joinin Joinempty）中所指示的VLAN在本地设备不存在时，GVRP可能会创建这个VLAN。是否允许动态创建VLAN由用户控制。

i 用户不能修改由 GVRP 创建的动态 VLAN 的参数。

配置端口的登记模式

- 【命令格式】 **gvrp registration mode { normal | disabled }**
- 【参数说明】 **normal**：端口允许加入动态 VLAN
disabled：端口不允许加入动态 VLAN
- 【命令模式】 接口模式
- 【使用指导】 设置端口的 GVRP 登记模式

i 这两种登记模式不会影响端口上的静态 VLAN，用户创建的静态 VLAN 永远都是 Fixed Registrar。

配置端口的通告模式

- 【命令格式】 **gvrp applicant state { normal | non-applicant }**
- 【参数说明】 **normal**：端口对外通告 VLAN 消息
non-applicant：端口不对外通告 VLAN 消息
- 【命令模式】 接口模式
- 【使用指导】 设置端口的 GVRP 通告模式

配置 GVRP 定时器

- 【命令格式】 **gvrp timer { join timer-value | leave timer-value | leaveall timer-value }**

- 【参数说明】 *timer-value* : 1-2147483647ms
- 【命令模式】 全局模式
- 【使用指导】 Leave timer 时间必须大于或等于 jointimer 的 3 倍。
Leaveall timer 的值必须比 leave timer 大。
时间单位为毫秒。
在实际组网中，建议用户将 GVRP 定时器配置为以下的推荐值：
Join Timer : 6000ms (6 秒钟) ；
Leave Timer : 30000ms (30 秒钟) ；
LeaveAll Timer : 120000ms (2 分钟) 。

i 要保证所有互联的 GVRP 设备中的 GVRP Timer 设置保持一致，否则 GVRP 可能工作异常。

配置举例

在拓扑环境中开启 GVRP 功能，动态维护 VLAN 以及 VLAN 和端口关系

【网络环境】

图 9-3



【配置方法】

- 在交换机 A 和 C 上配置用户通信的 VLAN。
- 在交换机 A、B 和 C 上开启 GVRP 功能，开启动态 VLAN 创建的开关。
- 交换机之间互连的接口设置为 Trunk 口，必须保证这些 Trunk 口的 VLAN 许可列表包含用户通信的 VLAN，默认 Trunk 口是允许所有 VLAN。
- 环境中最好开启 STP 协议，以免造成环路。

A

```

1：创建用户网络通信的 VLAN 1-200。
A# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
A(config)# vlan range 1-200
2：开启 GVRP 和动态 VLAN 创建的功能。
A(config)# gvrp enable
A(config)# gvrp dynamic-vlan-creation enable
3：将设备连接端口配置为 Trunk 口，Trunk 口默认允许所有 VLAN。
A(config)# interface gigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# switchport mode trunk
4：配置 Trunk 口的通告模式和注册模式，默认为 normal，可以不用手动配置。
A(config-if-GigabitEthernet 0/1)# gvrp applicant state normal
A(config-if-GigabitEthernet 0/1)# gvrp registration mode normal
A(config-if-GigabitEthernet 0/1)# end
  
```

C

- 配置同交换机 A，这里不再赘述。

B

```

1：设备上开启 GVRP 和动态 VLAN 创建的功能。
B# configure terminal
  
```

```

B(config)# gvrp enable
B(config)# gvrp dynamic-vlan-creation enable
2 : 将互连设备的端口设置为 trunk 口。
B(config)# interface range GigabitEthernet 0/2-3
B(config-if-GigabitEthernet 0/2)# switchport mode trunk

```

- 【检验方法】**
- 查看各个设备上 GVRP 配置是否正确。查看交换机 B 上是否有动态创建 VLAN 2-100。并查看交换机 B 上 G 0/2 和 G0/3 口是否有加入这些动态 VLAN。

A

```

A# show gvrp configuration
Global GVRP Configuration:
GVRP Feature:enabled
GVRP dynamic VLAN creation:enabled
Join Timers(ms):200
Leave Timers(ms):600
Leaveall Timers(ms):1000
Port based GVRP Configuration:

```

PORT	Applicant Status	Registration Mode
GigabitEthernet 0/1	normal	normal

B

```

B# show gvrp configuration
Global GVRP Configuration:
GVRP Feature:enabled
GVRP dynamic VLAN creation:enabled
Join Timers(ms):200
Leave Timers(ms):600
Leaveall Timers(ms):1000
Port based GVRP Configuration:

```

PORT	Applicant Status	Registration Mode
GigabitEthernet 0/2	normal	normal
GigabitEthernet 0/3	normal	normal

C

```

C# show gvrp configuration
Global GVRP Configuration:

```

```
GVRP Feature:enabled
GVRP dynamic VLAN creation:enabled
Join Timers(ms):200
Leave Timers(ms):600
Leaveall Timers(ms):1000
Port based GVRP Configuration:
      PORT                Applicant Status      Registration Mode
-----
GigabitEthernet 0/1      normal                normal
```

常见配置错误

- 设备连接的端口不是 Trunk 模式。
- 设备连接的端口许可 VLAN 列表不包含用户通信的 VLAN。
- Trunk 口的 GVRP 通告模式和注册模式不是 normal 模式。

9.4.2 配置GVRP PDUs 透传功能

配置效果

设备未开启 GVRP 协议时，需要透传 GVRP PDUs 帧，使得与之互连的设备之间的 GVRP 计算正常。

注意事项

GVRP PDUs 透传功能只在 GVRP 协议关闭时才起作用。当 GVRP 协议打开时，设备不透传 GVRP PDUs 帧。

配置方法

配置 GVRP PDUs 透传功能

- 可选配置
- 设备未开启 GVRP 协议时，如里需要透传 GVRP PDUs 帧，则需要配置 GVRP PDUs 透传功能。

检验方法

显示验证。

相关命令

配置 GVRP PDU 透传功能

【命令格式】 **bridge-frame forwarding protocol gvrp**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 在 IEEE 802.1Q 标准中，GVRP PDU 的目的 MAC 地址 01-80-C2-00-00-06 是作为保留地址使用的，即遵循 IEEE 802.1Q 标准的设备，对于接收到的 GVRP PDU 帧是不转发的。然而，在实际的网络布署中，可能需要设备能够支持透传 GVRP PDU 帧。例如，设备未开启 GVRP 协议时，需要透传 GVRP PDU 帧，使得与之互连的设备之间的 GVRP 计算正常。

GVRP PDU 透传功能只在 GVRP 协议关闭时才起作用。当 GVRP 协议打开时，设备不透传 GVRP PDU 帧。

配置举例

配置 GVRP PDU 透传功能

【网络环境】

图 9-4



DEV A, C 上开启 GVRP 协议，DEV B 未开启 GVRP 协议。

【配置方法】 DEV B 上配置 GVRP PDU 透传功能，使得 DEV A, C 之间的 GVRP 协议能够正确计算。

DEV B Ruijie(config)#bridge-frame forwarding protocol gvrp

【检验方法】 通过 **show run** 查看 GVRP PDU 透传功能是否开启。

DEV B Ruijie#show run

```
Building configuration...
Current configuration : 694 bytes
bridge-frame forwarding protocol gvrp
```

常见错误

无。

9.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除端口的统计值	clear gvrp statistics { <i>interface-id</i> all }

查看运行情况

作用	命令
显示端口的统计值	show gvrp statistics { <i>interface-id</i> all }
显示当前 GVRP 的运行状态	show gvrp status
显示当前 GVRP 的配置状态	show gvrp configuration

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 GVRP 事件 debug 开关	debug gvrp event
打开 GVRP 定时器 debug 开关	debug gvrp timer

10 LLDP

10.1 概述

LLDP (Link Layer Discovery Protocol, 链路层发现协议) 是由 IEEE 802.1AB 定义的一种链路层发现协议。通过 LLDP 协议能够进行拓扑的发现及掌握拓扑的变化情况。LLDP 将设备的本地信息组织成 TLV 的格式 (Type/Length/Value, 类型/长度/值) 封装在 LLDPDU (LLDP data unit, 链路层发现协议数据单元) 中发送给邻居设备, 同时它将邻居设备发送的 LLDPDU 以 MIB (Management Information Base, 管理信息库) 的形式存储起来, 提供给网络管理系统访问。

通过 LLDP, 网络管理系统可以掌握拓扑的连接情况, 比如设备的哪些端口与其它设备相连接, 链路连接两端的端口的速率、双工是否匹配等, 管理员可以根据这些信息快速地定位及排查故障。

一台支持 LLDP 协议的锐捷交换机产品, 当对端设备是支持 LLDP 协议的锐捷交换机产品, 或支持 LLDP-MED 协议的终端设备的时候, 该产品可以发现邻居信息。

- 支持 LLDP 协议的锐捷交换机产品。
- 支持 LLDP-MED 协议的终端设备。

协议规范

- IEEE 802.1AB 2005 : Station and Media Access Control Connectivity Discovery
- ANSI/TIA-1057 : Link Layer Discovery Protocol for Media Endpoint Devices

10.2 典型应用

典型应用	场景描述
利用LLDP查看拓扑连接情况	网络拓扑中有若干交换机设备、MED 设备、NMS 设备。
利用LLDP进行错误检测	网络拓扑中有直连的两台交换机设备, 错误配置信息将显示。

10.2.1 利用LLDP查看拓扑连接情况

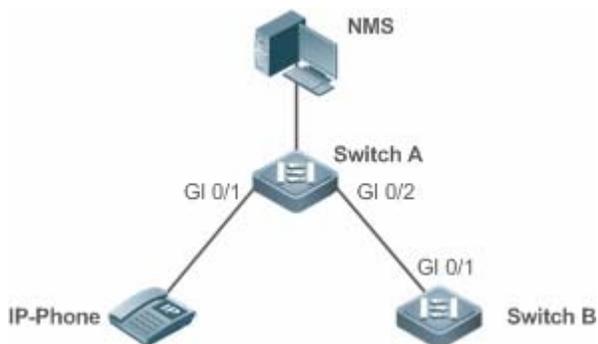
应用场景

网络拓扑中有若干交换机设备、MED 设备、NMS 设备。

以下图为例, LLDP 功能默认打开, 不需要再进行配置。

- Switch A 和 Switch B 可以互相发现对方是自己的邻居设备。
- Switch A 在端口 Gi 0/1 上可以发现邻居 MED 设备 IP-Phone。
- NMS (Network Management System, 网络管理系统) 能够访问 Switch A 的邻居设备信息。

图 10-1



- 【注释】 锐捷交换机产品 Switch A 和 Switch B、IP-Phone 都支持 LLDP 和 LLDP-MED。
交换机端口上 LLDP 的工作模式为 TxRx。
LLDP 报文的发送时间参数采用缺省值，即发送时间间隔为 30 秒、传输 LLDP 报文的延迟时间为 2 秒。

功能部属

- 在交换机中运行 LLDP 协议，实现邻居发现。
- 在交换机中运行 SNMP 协议，实现网络管理系统获取和设置交换机中的 LLDP 相关信息。

10.2.2 利用LLDP进行错误检测

应用场景

网络拓扑中有直连的两台交换机设备，错误配置信息将显示。

以下图为例，LLDP 功能默认打开，LLDP 错误检测功能缺省打开，不需要再进行配置。

- 管理员在对 Switch A 进行 VLAN 配置、端口速率双工配置、聚合端口配置和端口 MTU 配置时，如果配置的信息与相连接的邻居设备 Switch B 的配置不匹配，将提示相应的错误信息。反之亦然。

图 10-2



- 【注释】 两台锐捷交换机产品 Switch A 和 Switch B 都支持 LLDP 协议。
交换机端口上 LLDP 的工作模式为 TxRx。
LLDP 报文的发送时间参数采用缺省值，即发送时间间隔为 30 秒、传输 LLDP 报文的延迟时间为 2 秒。

功能部属

- 在交换机中运行 LLDP 协议，实现邻居发现，并检测两端的交换机直接接口的配置信息是否错误。

10.3 功能详解

基本概念

LLDPDU

LLDPDU 是指封装在 LLDP 报文中的协议数据单元，它由一系列的 TLV 封装而成。这些 TLV 集合包括了三个固定的 TLV 加上一系列可选的 TLVs 和一个 End Of TLV 组成。LLDPDU 的具体格式如图所示：

图 10-3 LLDPDU 格式



其中：

- M 表示是固定的 TLV。
- 在 LLDPDU 中，Chassis ID TLV、Port ID TLV、Time To Live TLV 和 End Of LLDPDU TLV 是必须携带的，而其它类型的 TLV 是可选携带。

LLDP 报文封装格式

LLDP 报文支持两种封装格式：Ethernet II 和 SNAP（Subnetwork Access Protocols，子网访问协议）。

其中 Ethernet II 格式封装的 LLDP 报文如图所示：

图 10-4 Ethernet II 格式封装的 LLDP 报文

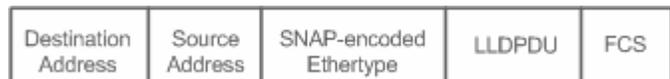


其中：

- Destination Address：目的 MAC 地址，为 LLDP 的组播地址 01-80-C2-00-00-0E。
- Source Address：源 MAC 地址，为设备的端口 MAC 地址。
- Ethertype：以太网类型，为 0x88CC。
- LLDPDU：LLDP 协议数据单元。
- FCS：帧校验序列。

SNAP 格式封装的 LLDP 报文如图所示：

图 10-5 SNAP 格式封装的 LLDP 报文



其中：

- Destination Address：目的 MAC 地址，为 LLDP 的组播地址 01-80-C2-00-00-0E。
- Source Address：源 MAC 地址，为设备的端口 MAC 地址。

- SNAP-encoded Ethertype：SNAP 封装的以太网类型，为 AA-AA-03-00-00-00-88-CC。
- LLDPDU：LLDP 协议数据单元。
- FCS：帧校验序列。

TLV

LLDPDU 中封装的 TLV 可以分成二个大类：

- 基本管理 TLV
- 组织定义 TLV

基本管理 TLV 是一组用于网络管理的基础 TLV 集合。组织定义 TLV 是由标准组织和其它机构定义的 TLV，比如 IEEE 802.1 组织、IEEE 802.3 组织分别定义了各自的 TLV 集合。

1. 基本管理 TLV

基本管理 TLV 集合包含了两种类型的 TLV：固定 TLV 和可选 TLV。固定 TLV 是指该 TLV 信息必须包含在 LLDPDU 中发布，可选 TLV 是指根据需要确定 TLV 是否包含在 LLDPDU 中发布。

基本管理 TLV 的内容见表：

TLV 类型	TLV 说明	在 LLDPDU 中用法
End Of LLDPDU TLV	LLDPDU 的结束标志，占用 2 个字节	固定
Chassis ID TLV	用于标识设备，通常用 MAC 地址表示	固定
Port ID TLV	用于标识发送 LLDPDU 的端口	固定
Time To Live TLV	本地信息在邻居设备上的存活时间，当收到 TTL 为 0 的 TLV 时，此时需要删除掉对应的邻居信息。	固定
Port Description TLV	发送 LLDPDU 的端口描述符	可选
System Name TLV	描述设备的名称	可选
System Description TLV	设备描述信息，包括硬件/软件版本、操作系统等信息	可选
System Capabilities TLV	描述设备的主要功能，例如桥接、路由、中继等功能	可选
Management Address TLV	管理地址，同时包含了接口号和 OID (Object Identifier, 对象标识)。	可选

- ✔ 锐捷交换机系列产品 LLDP 协议支持基本管理 TLV 的发布。

2. 组织定义 TLV

不同的组织 (例如 IEEE 802.1、IEEE 802.3、IETF 或者设备供应商) 定义特定的 TLV 信息去通告设备的特定信息。TLV 格式中通过 OUI (Organizationally Unique Identifier, 组织唯一标识符) 字段来区分不同的组织。

- 组织定义 TLV 属于可选的 TLV 集合，根据用户的实际需要在 LLDPDU 中发布。目前比较常见的组织定义 TLV 有以下三种：IEEE 802.1 组织定义的 TLV

IEEE 802.1 组织定义的 TLV 见表：

TLV 类型	TLV 说明
Port VLAN ID TLV	端口的 VLAN 标识符
Port And Protocol VLAN ID TLV	端口的协议 VLAN 标识符

VLAN Name TLV	端口的 VLAN 名称
Protocol Identity TLV	端口支持的协议类型

✔ 锐捷交换机系列产品 LLDP 协议，不支持发送 Protocol Identity TLV，但支持接收该类型的 TLV。

- IEEE 802.3 组织定义的 TLV

IEEE 802.3 组织定义的 TLV 见表：

TLV 类型	TLV 说明
MAC/PHY Configuration//Status TLV	端口的速率双工状态、是否支持并使能自动协商功能
Power Via MDI TLV	端口的供电能力
Link Aggregation TLV	端口的链路聚合能力及当前的聚合状态
Maximum Frame Size TLV	端口所能传输的最大的帧的大小

✔ 锐捷交换机系列产品 LLDP 协议支持 IEEE 802.3 组织定义的 TLV 的发布。

- LLDP-MED TLV

LLDP-MED 以 IEEE 802.1AB LLDP 协议为基础，它扩展了 LLDP，使用户能够更方便地部署 VoIP（Voice Over IP，基于 IP 的语音传输）网络及进行故障检测。它提供了网络配置策略、设备发现、以太网供电管理和目录管理等应用，满足了节约成本、有效地管理和易于部署方面的需求，简化了语音设备地部署。

LLDP-MED 定义的 TLV 见表：

TLV 类型	TLV 说明
LLDP-MED Capabilities TLV	设备是否支持 LLDP-MED、LLDPDU 中封装的 LLDP-MED TLV 类型以及当前设备的类型（网络连接设备或终端）
Network Policy TLV	通告端口的 VLAN 的配置、支持的应用类型（如语音或视频）、二层的优先级信息等
Location Identification TLV	定位标识终端设备。在网络拓扑收集等应用中能够精确地定位出终端设备
Extended Power-via-MDI TLV	提供了更高级的供电管理
Inventory – Hardware Revision TLV	MED 设备的硬件版本
Inventory – Firmware Revision TLV	MED 设备的固件版本
Inventory – Software Revision TLV	MED 设备的软件版本
Inventory – Serial Number TLV	MED 设备的序列号
Inventory – Manufacturer Name TLV	MED 设备的制造商的名称
Inventory – Model Name TLV	MED 设备的模块名称
Inventory – Asset ID TLV	MED 设备的资产标识符，用于目录管理和资产跟踪

✔ 锐捷交换机系列产品 LLDP 协议支持 LLDP-MED 定义的 TLV 的发布。

功能特性

功能特性	作用
------	----

LLDP工作模式	配置 LLDP 报文收发的模式。
LLDP报文的传输机制	直连支持 LLDP 协议的交换机设备可发送 LLDP 报文给对方。
LLDP报文的接收机制	直连支持 LLDP 协议的交换机设备可接收对方发送的 LLDP 报文。

10.3.1 LLDP工作模式

配置 LLDP 工作模式，能够使交换机收发 LLDP 报文的方式发生变化。

工作原理

LLDP 提供了三种工作模式：

- TxRx：既发送也接收 LLDPDU。
- Rx Only：只接收不发送 LLDPDU。
- Tx Only：只发送不接收 LLDPDU。

当端口的 LLDP 工作模式发生变化时，端口将对协议状态机进行初始化操作，通过配置端口初始化的延迟时间，可以避免由于工作模式频繁改变而导致端口不断地进行初始化操作。

相关配置

配置 LLDP 工作模式

缺省情况下，接口上的工作模式为 TxRx。

使用 `lldp mode` 命令可以改变接口上的工作模式。

必须在接口上配置工作模式为 TxRx 才能使 LLDP 协议报文收发功能正常。若接口工作模式配置为 Rx Only，那么设备只能接收 LLDP 报文，但无法发送 LLDP 报文；若接口工作模式配置为 Tx Only，那么设备只能发送 LLDP 报文，但无法接收 LLDP 报文；若接口工作模式关闭，将不再收发 LLDP 报文。

10.3.2 LLDP报文的传输机制

LLDP 报文的传输能让对端设备发现其邻居设备的存在，当取消 LLDP 传输模式或端口被管理 Shutdown 的时候，能够通告给对端设备其邻居信息不再有效。

工作原理

LLDP 工作在 TxRx 或 Tx Only 模式时，会周期性的发送 LLDP 报文。当本地设备的信息发生变化时，会立即发送 LLDP 报文。为了避免本地信息的频繁变化引起的频繁发送 LLDP 报文，在发送完一个 LLDP 报文后需要延迟一定的时间后再发往下一个 LLDP 报文。该延迟时间可以手工配置。

LLDP 提供了两种报文类型：

- 标准 LLDP 报文：包含了本地设备的管理和配置信息。

- Shutdown 通告报文：当取消了 LLDP 的传输模式或者端口被管理 Shutdown 时，将触发 LLDP Shutdown 通告报文的发送。Shutdown 通告报文由 Chassis ID TLV、Port ID TLV、Time To Live TLV 和 End OF LLDP TLV 组成。其中 Time To Live TLV 中 TTL 等于 0。当设备收到 LLDP Shutdown 通告报文时，将认为邻居信息已经不再有效并立即删除邻居信息。当 LLDP 工作模式由关闭或 Rx 转变为 TxRx 或 Tx，或者发现新邻居时（即收到新的 LLDP 报文且本地尚未保存该邻居信息），为了让邻居设备尽快学习到本设备的信息，将启动快速发送机制。快速发送机制调整 LLDP 报文的发送周期为 1 秒，并连续发送一定数量的 LLDP 报文。

相关配置

配置 LLDP 工作模式

缺省情况下，接口上的工作模式为 TxRx。

使用 `lldp mode txrx` 和 `lldp mode tx` 命令可以使 LLDP 报文传输功能打开，使用 `lldp mode rx` 和 `no lldp mode` 命令可以使 LLDP 报文传输功能关闭。

必须在接口上配置工作模式为 TxRx 或 Tx Only 才能使 LLDP 的报文传输功能正常。若接口工作模式配置为 Rx Only，那么设备只能接收 LLDP 报文，但无法发送 LLDP 报文。

配置 LLDP 报文的发送延迟时间

缺省情况下，LLDP 报文的发送延迟时间为 2 秒。

使用 `lldp timer tx-delay` 命令可以修改 LLDP 报文的发送延迟时间。

延迟时间配置过小，本地信息的频繁变化引起的频繁发送 LLDP 报文；配置值太大，本地信息的变化可能不能使发送 LLDP 报文。

配置 LLDP 报文的发送时间间隔

缺省情况下，LLDP 报文的发送时间间隔为 30 秒。

使用 `lldp timer tx-interval` 命令可以修改 LLDP 报文的发送时间间隔。

配置值太小，则会使 LLDP 发送频率过高；配置值太大，则可能会使对端设备不能及时发现本地设备。

配置允许发布的 TLV 类型

缺省情况下，接口上允许发布除 Location Identification TLV 之外的所有类型的 TLV。

使用 `lldp tlv-enable` 命令可以改变允许发布的 TLV 类型。

增加或减少发送的 LLDP 报文中 LLDPDU 的对应 TLV 字段。

配置 LLDP 快速发送报文的个数

缺省情况下，LLDP 快速发送报文的个数为 3 个。

使用 `lldp fast-count` 命令可以改变 LLDP 快速发送报文的个数。

改变快速发送机制下快速发送报文的个数。

10.3.3 LLDP报文的接收机制

LLDP 报文的接收能够发现邻居设备的存在以及何时应该老化邻居信息。

工作原理

LLDP 工作在 TxRx 或 RxOnly 模式时，能够接收 LLDP 报文。当设备收到 LLDP 报文时，会进行有效性检查。通过报文校验后，判断是新的邻居信息还是已经存在的邻居信息更新，并将邻居信息保存在本地设备。同时根据报文中 TTL TLV 的值设置邻居信息在本地设备的存活时间。如果收到 TTL TLV 的值为 0，表示需要立即老化掉该邻居信息。

相关配置

配置 LLDP 工作模式

缺省情况下，接口上的工作模式为 TxRx。

使用 `lldp mode txrx` 和 `lldp mode rx` 命令可以使 LLDP 报文接收功能打开，使用 `lldp mode tx` 和 `no lldp mode` 命令可以使 LLDP 报文接收功能关闭。

必须在接口上配置工作模式为 TxRx 或 Rx Only 才能使 LLDP 的报文接收功能正常。若接口工作模式配置为 Tx Only 或关闭，那么设备只能发送 LLDP 报文，但无法接收 LLDP 报文。

10.4 配置详解

配置项	配置建议 & 相关命令	
配置LLDP功能	⚠️ 可选配置。用于打开或关闭全局和接口的 LLDP 功能。	
	<code>lldp enable</code>	打开 LLDP 功能
	<code>no lldp enable</code>	关闭 LLDP 功能
配置LLDP工作模式	⚠️ 可选配置。用于配置 LLDP 报文收发模式。	
	<code>lldp mode {rx tx txrx }</code>	配置 LLDP 工作模式
	<code>no lldp mode</code>	关闭 LLDP 工作模式
配置允许发布的TLV类型	⚠️ 可选配置。用于配置允许发布的 TLV 类型。	
	<code>lldp tlv-enable</code>	配置允许发布的 TLV 类型
	<code>no lldp tlv-enable</code>	取消发布指定的 TLV 类型
配置LLDP报文中发布管理地址	⚠️ 可选配置。用于配置 LLDP 报文中发布。	
	<code>lldp management-address-tlv [ip-address]</code>	配置 LLDP 报文中发布管理地址
	<code>no lldp management-address-tlv</code>	取消管理地址的发布
配置快速发送LLDP报文的个	⚠️ 可选配置。用于配置快速发送 LLDP 报文的个数。	

数	lldp fast-count <i>value</i>	配置快速发送 LLDP 报文的个数
	no lldp fast-count	恢复缺省快速发送 LLDP 报文个数
配置TTL乘数和LLDP报文发送时间间隔	 可选配置。用于配置 TTL 乘数和 LLDP 报文发送时间间隔。	
	lldp hold-multiplier <i>value</i>	配置 TTL 乘数
	no lldp hold-multiplier	恢复缺省 TTL 乘数
	lldp timer tx-interval <i>seconds</i>	配置 LLDP 报文发送时间间隔
配置LLDP报文的发送延迟时间	 可选配置。用于配置 LLDP 报文的发送延迟时间。	
	lldp timer tx-delay <i>seconds</i>	配置 LLDP 报文的发送延迟时间
	no lldp timer tx-delay	恢复缺省 LLDP 报文的发送延迟时间
配置端口初始化的延迟时间	 可选配置。用于配置端口初始化的延迟时间。	
	lldp timer reinit-delay <i>seconds</i>	配置端口初始化的延迟时间
	no lldp timer reinit-delay	恢复缺省端口初始化的延迟时间
配置LLDP Trap功能	 可选配置。用于配置 LLDP Trap 功能。	
	lldp notification remote-change enable	打开 LLDP Trap 功能
	no lldp notification remote-change enable	关闭 LLDP Trap 功能
	lldp timer notification-interval	配置发送 LLDP Trap 信息的时间间隔
配置LLDP错误检测功能	 可选配置。用于配置 LLDP 错误检测功能。	
	lldp error-detect	打开 LLDP 错误检测功能
	no lldp error-detect	关闭 LLDP 错误检测功能
配置LLDP报文封装格式	 可选配置。用于配置 LLDP 报文封装格式。	
	lldp encapsulation snap	配置 LLDP 报文的封装格式为 SNAP
	no lldp encapsulation snap	配置 LLDP 报文的封装格式为 Ethernet II
配置LLDP Network Policy策略	 可选配置。用于配置 LLDP Network Policy 策略。	
	lldp network-policy profile <i>profile-num</i>	配置 LLDP Network Profile 策略
	no lldp network-policy profile <i>profile-num</i>	删除 LLDP Network Profile 策略
配置设备的普通地址信息	 可选配置。用于配置设备的普通地址信息。	

	<pre>{ country state county city division neighborhood street-group leading-street-dir trailing-street-suffix street-suffix number street-number-suffix landmark additional-location-information name postal-code building unit floor room type-of-place postal-community-name post-office-box additional-code } ca-word</pre>	配置设备的普通地址信息
	<pre>no { country state county city division neighborhood street-group leading-street-dir trailing-street-suffix street-suffix number street-number-suffix landmark additional-location-information name postal-code building unit floor room type-of-place postal-community-name post-office-box additional-code } ca-word</pre>	删除设备的普通地址信息
配置设备的紧急电话号码信息	 可选配置。用于配置设备的紧急电话号码信息。	
	<pre>lldp location elin identifier id elin-location tel-number</pre>	配置设备的紧急电话号码信息
	<pre>no lldp location elin identifier id</pre>	删除设备的紧急电话号码信息

10.4.1 配置LLDP功能

配置效果

- 打开或关闭 LLDP 的功能。

注意事项

- 如果要求接口上 LLDP 功能生效，则要同时开启全局和该接口上的 LLDP 功能。

配置方法

- 可选配置。
- 可对全局或接口下配置 LLDP 功能。

检验方法

显示 LLDP 的状态信息。

- 检查全局 LLDP 功能是否开启。
- 检查接口下 LLDP 功能是否开启。

相关命令

打开 LLDP 功能

- 【命令格式】 **lldp enable**
- 【参数说明】 -
- 【命令模式】 全局模式、接口模式
- 【使用指导】 需要全局打开 LLDP 开关，接口的 LLDP 功能才生效。

关闭 LLDP 功能

- 【命令格式】 **no lldp enable**
- 【参数说明】 -
- 【命令模式】 全局模式、接口模式
- 【使用指导】 -

配置举例

关闭 LLDP 功能

- 【配置方法】 ● 关闭全局 LLDP 功能。

```
Ruijie(config)#no lldp enable
```

- 【检验方法】 ● 显示 LLDP 全局状态信息。

```
Ruijie(config)#show lldp status  
Global status of LLDP: Disable
```

常见错误

- 接口已开启 LLDP 功能，但是全局没有开启 LLDP 功能，此时接口下的 LLDP 功能还是不能生效。
- 端口学习到的邻居个数限制在 5 个，即端口最多只能学习到 5 个邻居。
- 如果邻居设备不支持 LLDP，但是邻居设备下连的设备支持 LLDP，由于邻居设备可能会转发 LLDP 的报文，这样，端口可能会学习到非直连的设备的信息。

10.4.2 配置LLDP工作模式

配置效果

- 配置接口的 LLDP 的工作模式为 TxRx，则该接口可发送和接收报文。
- 配置接口的 LLDP 的工作模式为 Tx，则该接口只能发送报文，不能接收报文。
- 配置接口的 LLDP 的工作模式为 Rx，则该接口只能接收报文，不能发送报文。
- 关闭接口的 LLDP 工作模式，则该接口不能接收和发送报文。

注意事项

- LLDP 运行在实际的物理接口上（对于 AP 口，则实际是运行在 AP 成员口上）。堆叠口，VSL 口不支持 LLDP。

配置方法

- 可选配置。
- 用户可根据实际需要在工作模式修改为 Tx 或 Rx 模式。

检验方法

显示接口下 LLDP 的状态信息。

- 检查接口下 LLDP 的工作模式是否和配置的不同。

相关命令

配置 LLDP 工作模式

【命令格式】 **lldp mode { rx | tx | txrx }**

【参数说明】 rx：表示只接收不发送 LLDPDU

tx：表示只发送不接收 LLDPDU

txrx：表示即发送又接收 LLDPDU

【命令模式】 接口模式

【使用指导】 接口 LLDP 功能生效的前提是全局使能了 LLDP 且接口 LLDP 的工作模式处于 tx、rx 或 txrx。

关闭 LLDP 工作模式

【命令格式】 **no lldp mode**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 关闭接口的 LLDP 工作模式，此时接口不再发送和接收 LLDP 报文。

配置举例

配置 LLDP 工作模式

【配置方法】 接口下配置 LLDP 的工作模式为 Tx 模式。

```
Ruijie(config)#interface gigabitethernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#lldp mode tx
```

【检验方法】 显示 LLDP 在接口下的状态信息。

```
Ruijie(config-if-GigabitEthernet 0/1)#show lldp status interface gigabitethernet 0/1
Port [GigabitEthernet 0/1]
Port status of LLDP          : Enable
Port state                   : UP
Port encapsulation          : Ethernet II
Operational mode            : TxOnly
Notification enable         : NO
Error detect enable         : YES
Number of neighbors         : 0
Number of MED neighbors     : 0
```

常见配置错误

- -

10.4.3 配置允许发布的TLV类型

配置效果

- 用户可以通过配置运行发布的 TLV 类型，使发送 LLDP 报文中 LLDPDU 的内容改变。

注意事项

- 配置基本管理 TLV、IEEE 802.1 组织定义 TLV、IEEE 802.3 组织定义 TLV 时，如果指定 **all** 参数，将发布该类型的所有可选 TLV。
- 配置 LLDP-MED TLV 时，如果指定 **all** 参数，将发布除 Location Identification TLV 之外的所有类型的 LLDP-MED TLV。
- 配置允许发布 LLDP-MED Capability TLV 时，需要先配置允许发布 LLDP 802.3 MAC/PHY TLV；取消发布 LLDP 802.3 MAC/PHY TLV 时，需要先取消发布 LLDP-MED Capability TLV
- 配置 LLDP-MED TLV 时，必须配置允许发布 LLDP-MED Capability TLV，才可以配置允许发布 LLDP-MED 其它类型的 TLV。取消发布 LLDP-MED TLV，必须先取消发布 LLDP-MED 其它类型的 TLV，才允许取消发布 LLDP-MED Capability TLV。当设备下联 IP 电话，若 IP 电话支持 LLDP-MED，则可以通过配置 network policy TLV 下发策略给 IP 电话
- 如果设备缺省支持 DCBX 功能，缺省情况下端口上不允许发布 IEEE 802.3 TLV 及 LLDP-MED TLV

配置方法

- 可选配置。

- 用户可根据实际需要在某接口下配置允许发布的 TLV 类型。

检验方法

显示端口上可发布的 TLV 配置信息。

- 检查接口下允许发布的 TLV 是否和配置的一致。

相关命令

配置 LLDP 允许发布的 TLV

【命令格式】 `lldp tlv-enable { basic-tlv { all | port-description | system-capability | system-description | system-name } | dot1-tlv { all | port-vlan-id | protocol-vlan-id [vlan-id] | vlan-name [vlan-id] } | dot3-tlv { all | link-aggregation | mac-physic | max-frame-size | power } | med-tlv { all | capability | inventory | location { civic-location | elin } | identifier id | network-policy profile [profile-num] | power-over-ethernet }`

【参数说明】 **basic-tlv**：基本管理 TLV

port-description：表示 Port Description TLV

system-capability：表示 System Capabilities TLV

system-description：表示 System Description TLV

system-name：表示 System Name TLV

dot1-tlv：802.1 组织定义的 TLV

port-vlan-id：表示 Port VLAN ID TLV

protocol-vlan-id：表示 Port And Protocol VLAN ID TLV

vlan-id：表示端口协议 VLAN ID，配置范围为：1-4094

vlan-name：表示 VLAN Name TLV

vlan-id：表示指定 VLAN 名称对应的 VLAN ID，配置范围为：1-4094

dot3-tlv：802.3 组织定义的 TLV

link-aggregation：表示 Link Aggregation TLV

mac-physic：表示 MAC/PHY Configuration/Status TLV

max-frame-size：表示 Maximum Frame Size TLV

power：表示 Power Via MDI TLV

med-tlv：LLDP MED TLV

capability：表示 LLDP-MED Capabilities TLV

inventory：表示目录管理 TLV，包括硬件版本、固件版本、软件版本、序列号、制造产商名称、模块名称和资产标识符等

location：表示 Location Identification TLV

civic-location：表示封装网络连接设备的普通地址信息

elin：表示封装紧急电话号码信息

id：表示配置的策略 ID，配置范围为：1-1024

network-policy：表示 Network Policy TLV

profile-num : Network Policy 策略 ID , 配置范围为 : 1-1024

power-over-ethernet : 表示 Extended Power-via-MDI TLV

【命令模式】 接口模式

【使用指导】



取消发布指定的 TLV 类型

【命令格式】 **no lldp tlv-enable {basic-tlv { all | port-description | system-capability | system-description | system-name } | dot1-tlv { all | port-vlan-id | protocol-vlan-id | vlan-name } | dot3-tlv { all | link-aggregation | mac-physic | max-frame-size | power } | med-tlv { all | capability | inventory | location { civic-location | elin } identifier id | network-policy profile [profile-num] | power-over-ethernet }**

【参数说明】 **basic-tlv** : 基本管理 TLV

port-description : 表示 Port Description TLV

system-capability : 表示 System Capabilities TLV

system-description : 表示 System Description TLV

system-name : 表示 System Name TLV

dot1-tlv : 802.1 组织定义的 TLV

port-vlan-id : 表示 Port VLAN ID TLV

protocol-vlan-id : 表示 Port And Protocol VLAN ID TLV

vlan-name : 表示 VLAN Name TLV

dot3-tlv : 802.3 组织定义的 TLV

link-aggregation : 表示 Link Aggregation TLV

mac-physic : 表示 MAC/PHY Configuration/Status TLV

max-frame-size : 表示 Maximum Frame Size TLV

power : 表示 Power Via MDI TLV

med-tlv : LLDP MED TLV

capability : 表示 LLDP-MED Capabilities TLV

inventory : 表示目录管理 TLV , 包括硬件版本、固件版本、软件版本、序列号、制造产商名称、模块名称和资产标识符等

location : 表示 Location Identification TLV

civic-location : 表示封装网络连接设备的普通地址信息

elin : 表示封装紧急电话号码信息

id : 表示配置的策略 ID , 配置范围为 : 1-1024

network-policy : 表示 Network Policy TLV

profile-num : Network Policy 策略 ID , 配置范围为 : 1-1024

power-over-ethernet : 表示 Extended Power-via-MDI TLV

【命令模式】 接口模式

【使用指导】



配置举例

配置 LLDP 允许发布的 TLV

【配置方法】 配置取消发布 IEEE 802.1 组织定义的 Port And Protocol VLAN ID TLV

```
Ruijie(config)#interface gigabitethernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#no lldp tlv-enable dot1-tlv protocol-vlan-id
```

【检验方法】 显示 LLDP 在接口下的 TLV 配置信息。

```
Ruijie(config-if-GigabitEthernet 0/1)#show lldp tlv-config interface gigabitethernet 0/1
LLDP tlv-config of port [GigabitEthernet 0/1]

```

NAME	STATUS	DEFAULT

Basic optional TLV:		
Port Description TLV	YES	YES
System Name TLV	YES	YES
System Description TLV	YES	YES
System Capabilities TLV	YES	YES
Management Address TLV	YES	YES
IEEE 802.1 extend TLV:		
Port VLAN ID TLV	YES	YES
Port And Protocol VLAN ID TLV	NO	YES
VLAN Name TLV	YES	YES
IEEE 802.3 extend TLV:		
MAC-Physic TLV	YES	YES
Power via MDI TLV	YES	YES
Link Aggregation TLV	YES	YES
Maximum Frame Size TLV	YES	YES
LLDP-MED extend TLV:		
Capabilities TLV	YES	YES
Network Policy TLV	YES	YES
Location Identification TLV	NO	NO
Extended Power via MDI TLV	YES	YES
Inventory TLV	YES	YES

常见配置错误

● -

10.4.4 配置LLDP报文中发布管理地址

配置效果

- 配置接口下 LLDP 报文中的发布管理地址，可使管理地址 TLV 发生改变。
- 取消管理地址发布将使 LLDP 报文中的管理地址按缺省情况下选取。

注意事项

- LLDP 运行在实际的物理接口上（对于 AP 口，则实际是运行在 AP 成员口上）。堆叠口，VSL 口不支持 LLDP。

配置方法

- 可选配置。
- 在接口下配置 LLDP 报文发布的管理地址。

检验方法

显示本地设备接口下的 LLDP 信息。

- 检查本地设备接口下的 LLDP 信息是否和配置的相同。

相关命令

配置 LLDP 报文中发布的管理地址

【命令格式】 **lldp management-address-tlv** [*ip-address*]

【参数说明】 *ip-address* : LLDP 报文中发布的管理地址

【命令模式】 接口模式

【使用指导】 缺省情况下，LLDP 报文发布管理地址。发布的管理地址为端口允许通过的最小 VLAN 的 IPv4 地址，如果该 VLAN 未配置 IPv4 地址，则继续查找下一个允许通过的最小 VLAN，直到找到 IPv4 地址为止。

取消管理地址的发布

【命令格式】 **no lldp management-address-tlv**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 缺省情况下，LLDP 报文发布管理地址。发布的管理地址为端口允许通过的最小 VLAN 的 IPv4 地址，如果该 VLAN 未配置 IPv4 地址，则继续查找下一个允许通过的最小 VLAN，直到找到 IPv4 地址为止。

配置举例

配置 LLDP 报文中发布的管理地址

【配置方法】 在接口下配置 LLDP 报文发布的管理地址为 192.168.1.1

```
Ruijie(config)#interface gigabitethernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#lldp management-address-tlv 192.168.1.1
```

【检验方法】 查看对应接口下相应的配置信息

```
Ruijie(config-if-GigabitEthernet 0/1)#show lldp local-information interface GigabitEthernet 0/1
Lldp local-information of port [GigabitEthernet 0/1]
  Port ID type           : Interface name
  Port id                : GigabitEthernet 0/1
  Port description       : GigabitEthernet 0/1

  Management address subtype : ipv4
  Management address       : 192.168.1.1
  Interface numbering subtype : ifIndex
  Interface number         : 1
  Object identifier        :

  802.1 organizationally information
  Port VLAN ID             : 1
  Port and protocol VLAN ID (PPVID) : 1
    PPVID Supported        : YES
    PPVID Enabled          : NO
  VLAN name of VLAN 1     : VLAN0001
  Protocol Identity        :

  802.3 organizationally information
  Auto-negotiation supported : YES
  Auto-negotiation enabled   : YES
  PMD auto-negotiation advertised : 1000BASE-T full duplex mode, 100BASE-TX full duplex mode,
100BASE-TX half duplex mode, 10BASE-T full duplex mode, 10BASE-T half duplex mode
  Operational MAU type      : speed(100)/duplex(Full)
  PoE support                : NO
  Link aggregation supported : YES
  Link aggregation enabled   : NO
  Aggregation port ID       : 0
  Maximum frame Size        : 1500

  LLDP-MED organizationally information
  Power-via-MDI device type : PD
  Power-via-MDI power source : Local
```

```
Power-via-MDI power priority      :  
Power-via-MDI power value        :  
Model name                        : Model name
```

常见配置错误

- -

10.4.5 配置快速发送LLDP报文的个数

配置效果

- 改变快速发送机制下 LLDP 报文发送的个数。

注意事项

- -

配置方法

- 可选配置。
- 在全局配置模式下配置快速发送 LLDP 报文个数。

检验方法

显示全局 LLDP 的状态信息。

- 检查 LLDP 快速发送个数是否和配置的不同。

相关命令

▾ 配置快速发送 LLDP 报文的个数

【命令格式】 **lldp fast-count value**

【参数说明】 value : LLDP 快速发送报文的个数，缺省为 3 个，可配置的范围为 1-10

【命令模式】 全局模式

【使用指导】 -

▾ 恢复缺省快速发送 LLDP 报文个数

【命令格式】 **no lldp fast-count**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

配置举例

配置快速发送 LLDP 报文的个数

【配置方法】 全局配置模式下配置快速发送 LLDP 报文的个数为 5 个

```
Ruijie(config)#lldp fast-count 5
```

【检验方法】 显示全局 LLDP 的状态信息。

```
Ruijie(config)#show lldp status
Global status of LLDP           : Enable
Neighbor information last changed time :
Transmit interval               : 30s
Hold multiplier                 : 4
Reinit delay                    : 2s
Transmit delay                  : 2s
Notification interval          : 5s
Fast start counts               : 5
```

常见配置错误

- -

10.4.6 配置TTL乘数和LLDP报文发送时间间隔

配置效果

- 改变 TTL 乘数的值。
- 改变 LLDP 报文发送时间间隔。

注意事项

- -

配置方法

- 可选配置。
- 全局配置模式下进行配置。

检验方法

显示接口下 LLDP 的状态信息。

- 检查接口下 LLDP 的工作模式是否和配置的不同。

相关命令

配置 TTL 乘数

【命令格式】 **lldp hold-multiplier** *value*

【参数说明】 *value* : TTL 乘数, 缺省为 4, 配置范围为 2-10

【命令模式】 全局模式

【使用指导】 LLDP 报文中 Time To Live TLV 的值=TTL 乘数×报文发送时间间隔+1。因此, 通过调整 TTL 乘数可以控制本设备信息在邻居设备的存活时间。

恢复缺省 TTL 乘数

【命令格式】 **no lldp hold-multiplier**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 LLDP 报文中 Time To Live TLV 的值=TTL 乘数×报文发送时间间隔+1。因此, 通过调整 TTL 乘数可以控制本设备信息在邻居设备的存活时间。

配置 LLDP 报文发送时间间隔

【命令格式】 **lldp timer tx-interval** *seconds*

【参数说明】 *seconds* : LLDP 报文的发送时间间隔, 可配置范围为 5-32768

【命令模式】 全局模式

【使用指导】 -

恢复缺省 LLDP 报文发送时间间隔

【命令格式】 **no lldp timer tx-interval**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

配置举例

配置 LLDP 工作模式

【配置方法】 配置 TTL 乘数为 3, LLDP 报文的发送间隔为 20 秒, 此时, 本地设备信息在邻居设备的存活时间为 61 秒

```
Ruijie(config)#lldp hold-multiplier 3
```

```
Ruijie(config)#lldp timer tx-interval 20
```

【检验方法】 显示全局 LLDP 状态信息。

```
Ruijie(config)#lldp hold-multiplier 3
Ruijie(config)#lldp timer tx-interval 20
Ruijie(config)#show lldp status
Global status of LLDP           : Enable
Neighbor information last changed time :
Transmit interval               : 20s
Hold multiplier                 : 3
Reinit delay                    : 2s
Transmit delay                  : 2s
Notification interval          : 5s
Fast start counts               : 3
```

常见配置错误

- -

10.4.7 配置LLDP报文的发送延迟时间

配置效果

- 改变 LLDP 报文的发送延迟时间。

注意事项

- -

配置方法

- 可选配置。
- 用户可根据实际需要在全局配置模式下进行配置。

检验方法

显示全局 LLDP 的状态信息。

- 检查 LLDP 报文的发送延迟时间是否和配置的相同。

相关命令

配置 LLDP 报文的发送延迟时间

- 【命令格式】 **lldp timer tx-delay seconds**
- 【参数说明】 seconds : LLDP 报文的发送延迟时间, 可配置范围为 1-8192
- 【命令模式】 全局模式
- 【使用指导】 当本地信息发生变化时, 会立即向邻居设备发送 LLDP 报文。为了避免本地信息频繁变化引起的频繁地发送 LLDP 报文, 可以配置 LLDP 报文的发送延迟时间来限制 LLDP 报文的频繁发送。

恢复缺省的 LLDP 报文的发送延迟时间

- 【命令格式】 **no lldp timer tx-delay**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 当本地信息发生变化时, 会立即向邻居设备发送 LLDP 报文。为了避免本地信息频繁变化引起的频繁地发送 LLDP 报文, 可以配置 LLDP 报文的发送延迟时间来限制 LLDP 报文的频繁发送。

配置举例

配置 LLDP 报文的发送延迟时间

- 【配置方法】 配置发送 LLDP 报文的延迟时间为 3 秒
- ```
Ruijie(config)#lldp timer tx-delay 3
```

- 【检验方法】 查看全局 LLDP 状态信息
- ```
Ruijie(config)#show lldp status
Global status of LLDP           : Enable
Neighbor information last changed time :
Transmit interval               : 30s
Hold multiplier                 : 4
Reinit delay                    : 2s
Transmit delay                  : 3s
Notification interval          : 5s
Fast start counts               : 3
```

常见配置错误

-

10.4.8 配置端口初始化的延迟时间

配置效果

- 改变端口初始化的延迟时间。

注意事项

- -

配置方法

- 可选配置。
- 用户可根据实际需要对接口状态机初始化的延迟时间进行配置。

检验方法

显示全局 LLDP 的状态信息。

- 检查全局 LLDP 的端口初始化的延迟时间是否和配置的相同。

相关命令

配置端口初始化的延迟时间

【命令格式】 **lldp timer reinit-delay** *seconds*

【参数说明】 *seconds*：端口初始化的延迟时间，配置范围为 1-10 秒

【命令模式】 全局模式

【使用指导】 为了避免端口的工作模式的频繁变化引起的频繁地初始化状态机，可以配置端口初始化的延迟时间。

恢复缺省端口初始化的延迟时间

【命令格式】 **no lldp timer reinit-delay**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 为了避免端口的工作模式的频繁变化引起的频繁地初始化状态机，可以配置端口初始化的延迟时间。

配置举例

配置端口初始化的延迟时间

【配置方法】 配置端口初始化的延迟时间为 3 秒，并显示 LLDP 的状态信息。

```
Ruijie(config)#lldp timer reinit-delay 3
```

【检验方法】 显示全局 LLDP 的状态信息。

```
Ruijie(config)#show lldp status
```

```
Global status of LLDP : Enable
```

```
Neighbor information last changed time :
Transmit interval                      : 30s
Hold multiplier                        : 4
Reinit delay                          : 3s
Transmit delay                         : 2s
Notification interval                 : 5s
Fast start counts                      : 3
```

常见配置错误

- -

10.4.9 配置LLDP Trap功能

配置效果

- 改变发送 LLDP Trap 信息的时间间隔。

注意事项

- -

配置方法

📄 打开 LLDP Trap 功能

- 可选配置。
- 接口配置模式下进行配置。

📄 配置发送 LLDP Trap 信息的时间间隔

- 可选配置。
- 全局配置模式下进行配置。

检验方法

显示 LLDP 的状态信息。

- 检查 LLDP Trap 功能是否打开。
- 检查发送 LLDP Trap 信息的时间间隔和配置的不同。

相关命令

▾ 打开 LLDP Trap 功能

- 【命令格式】 **lldp notification remote-change enable**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 通过配置 Trap 功能，可以将本地设备的 LLDP 信息（例如发现新邻居、检测到与邻居的通信链路故障等信息）发送给网管服务器，管理员可以根据此信息监控网络的运行状况。

▾ 关闭 LLDP Trap 功能

- 【命令格式】 **no lldp notification remote-change enable**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 通过配置 Trap 功能，可以将本地设备的 LLDP 信息（例如发现新邻居、检测到与邻居的通信链路故障等信息）发送给网管服务器，管理员可以根据此信息监控网络的运行状况。

▾ 配置发送 LLDP Trap 信息的时间间隔

- 【命令格式】 **lldp timer notification-interval seconds**
- 【参数说明】 *seconds*：配置发送 LLDP Trap 信息的时间间隔，缺省的时间间隔是 5 秒，可配置的范围是 5-3600
- 【命令模式】 全局模式
- 【使用指导】 为了防止 LLDP Trap 信息的频繁发送，可以配置发送 LLDP Trap 的时间间隔。在这段时间间隔内，检测到 LLDP 信息变化，将发送 Trap 给网管服务器。

▾ 恢复缺省的发送 LLDP Trap 信息的时间间隔

- 【命令格式】 **no lldp timer notification-interval**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 为了防止 LLDP Trap 信息的频繁发送，可以配置发送 LLDP Trap 的时间间隔。在这段时间间隔内，检测到 LLDP 信息变化，将发送 Trap 给网管服务器。

配置举例

▾ 打开 LLDP Trap 功能及配置发送 LLDP Trap 信息的时间间隔

- 【配置方法】 使能 LLDP Trap 功能，并配置 LLDP Trap 信息的发送时间间隔为 10 秒。

```
Ruijie(config)#lldp timer notification-interval 10
Ruijie(config)#interface gigabitethernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#lldp notification remote-change enable
```

- 【检验方法】 显示 LLDP 的状态信息。

```
Ruijie(config-if-GigabitEthernet 0/1)#show lldp status
Global status of LLDP                : Enable
Neighbor information last changed time :
```

```
Transmit interval          : 30s
Hold multiplier            : 4
Reinit delay              : 2s
Transmit delay            : 2s
Notification interval     : 10s
Fast start counts         : 3
```

Port [GigabitEthernet 0/1]

```
Port status of LLDP       : Enable
Port state                : UP
Port encapsulation       : Ethernet II
Operational mode         : RxAndTx
Notification enable      : YES
Error detect enable      : YES
Number of neighbors      : 0
Number of MED neighbors  : 0
```

常见配置错误

- -

10.4.10 配置LLDP错误检测功能

配置效果

- LLDP 错误检测功能打开，当 LLDP 检测到错误时，将打印 LOG 信息提示管理员。
- 配置 LLDP 错误检测功能，错误检测包括链路两端的 VLAN 配置检测、端口状态检测、端口聚合配置检测、MTU 配置检测及环路检测

注意事项

- -

配置方法

- 可选配置。
- 用户可根据实际需要在接口模式下进行配置，打开或关闭 LLDP 错误检测功能。

检验方法

显示接口下 LLDP 的状态信息。

- 检查接口下 LLDP 错误检测功能是打开还是关闭，与实际配置是否一致。

相关命令

打开 LLDP 错误检测功能

【命令格式】 **lldp error-detect**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 LLDP 错误检测功能是依靠链路两端的设备交互 LLDP 报文中的特定的 TLV 信息进行的，为了保证检测功能的正确运行，需要设备发布正确的 TLV 信息。

关闭 LLDP 错误检测功能

【命令格式】 **no lldp error-detect**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 LLDP 错误检测功能是依靠链路两端的设备交互 LLDP 报文中的特定的 TLV 信息进行的，为了保证检测功能的正确运行，需要设备发布正确的 TLV 信息。

配置举例

打开 LLDP 错误检测功能

【配置方法】 打开 LLDP 在接口 GI 0/1 下的错误检测功能。

```
Ruijie(config)#interface gigabitethernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#lldp error-detect
```

【检验方法】 显示 LLDP 在接口下的状态信息。

```
Ruijie(config-if-GigabitEthernet 0/1)#show lldp status interface gigabitethernet 0/1
Port [GigabitEthernet 0/1]
Port status of LLDP           : Enable
Port state                    : UP
Port encapsulation            : Ethernet II
Operational mode              : RxAndTx
Notification enable           : NO
Error detect enable           : YES
Number of neighbors           : 0
Number of MED neighbors       : 0
```

常见配置错误

- -

10.4.11 配置LLDP报文封装格式

配置效果

- 改变 LLDP 报文的封装格式。

注意事项

- -

配置方法

- 可选配置。
- 用户可根据实际需要在接口下改变 LLDP 报文的封装格式。

检验方法

显示接口下 LLDP 的状态信息。

- 检查接口下 LLDP 报文封装格式是否和配置的相同。

相关命令

▾ 配置 LLDP 报文的封装格式为 SNAP

【命令格式】 **lldp encapsulation snap**

【参数说明】 -

【命令模式】 接口模式

【使用指导】



为了保证本地设备和邻居设备的正常通信，需要将 LLDP 报文配置成相同的封装格式。

▾ 恢复缺省的 LLDP 报文的封装格式，即为 Ethernet II

【命令格式】 **no lldp encapsulation snap**

【参数说明】 -

【命令模式】 接口模式

【使用指导】



为了保证本地设备和邻居设备的正常通信，需要将 LLDP 报文配置成相同的封装格式。

配置举例

配置 LLDP 报文的封装格式为 SNAP

【配置方法】 配置 LLDP 报文的封装格式为 SNAP。

```
Ruijie(config)#interface gigabitethernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#lldp encapsulation snap
```

【检验方法】 显示 LLDP 在接口下的状态信息。

```
Ruijie(config-if-GigabitEthernet 0/1)#show lldp status interface gigabitethernet 0/1
Port [GigabitEthernet 0/1]
Port status of LLDP          : Enable
Port state                   : UP
Port encapsulation          : Snap
Operational mode             : RxAndTx
Notification enable         : NO
Error detect enable         : YES
Number of neighbors         : 0
Number of MED neighbors     : 0
```

常见配置错误

- -

10.4.12 配置LLDP Network Policy策略

配置效果

- 改变 LLDP Network Policy 策略。
- 当设备下联 IP 电话，若 IP 电话支持 LLDP-MED，则可以通过配置 Network Policy TLV 下发策略给 IP 电话，由 IP 电话修改语音流 Tag 和 QOS。在设备上，除配置上述策外，还需要配置步骤为：1.使能 Voice VLAN 功能，把连接 IP 电话的端口静态加入 Voice VLAN；2.把连接 IP 电话的端口配置为 QOS 信任口（推荐使用信任 DSCP 模式）；3.如果在此端口上同时开启了 1X 认证，则还需要配置一条安全通道，允许 Voice VLAN 内的报文通过。若 IP 电话不支持 LLDP-MED，则必须使能 Voice VLAN 功能，并将话机 MAC 地址手动配置到 Voice VLAN OUI 列表中。
- QOS 信任模式的配置方法请参见《IP QOS》章节；Voice VLAN 的配置方法请参见《Voice VLAN》章节；安全通道的配置方法请参见《ACL》章节。

注意事项

- -

配置方法

- 可选配置。
- 用户可根据实际需要配置 LLDP Network Policy 策略。

检验方法

显示本地设备的 LLDP network-policy 配置策略信息。

- 检查 LLDP Network Policy 策略是否和配置的相同。

相关命令

配置 LLDP Network Profile 策略

【命令格式】 **lldp network-policy profile** *profile-num*

【参数说明】 *profile-num* : LLDP network-policy 策略的标识, 范围为 : 1-1024

【命令模式】 全局模式

【使用指导】 使用此命令进入 LLDP network-policy 配置模式, 使用此命令时需要指定策略 ID。
进入 LLDP network-policy 配置模式后, 可使用{ voice | voice-signaling } vlan 命令配置具体的 network-policy 策略。

删除 LLDP Network Profile 策略

【命令格式】 **no lldp network-policy profile** *profile-num*

【参数说明】 *profile-num* : LLDP network-policy 策略的标识, 范围为 : 1-1024

【命令模式】 接口模式

【使用指导】 使用此命令进入 LLDP network-policy 配置模式, 使用此命令时需要指定策略 ID。
进入 LLDP network-policy 配置模式后, 可使用{ voice | voice-signaling } vlan 命令配置具体的 network-policy 策略。

配置举例

配置 LLDP Network Profile 策略

【配置方法】 配置接口 1 发布的 LLDP 报文中 Network Policy TLV 策略为 1 : voice 应用类型 vlan id 是 3 , cos 是 4 , dscp 是 6。

```
Ruijie#config
Ruijie(config)#lldp network-policy profile 1
Ruijie(config-lldp-network-policy)# voice vlan 3 cos 4
Ruijie(config-lldp-network-policy)# voice vlan 3 dscp 6
Ruijie(config-lldp-network-policy)#exit
Ruijie(config)# interface gigabitethernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# lldp tlv-enable med-tlv network-policy profile 1
```

【检验方法】 显示本地设备的 LLDP network-policy 配置策略信息。

```
network-policy information:
-----
network policy profile :1
voice vlan 3 cos 4
voice vlan 3 dscp 6
```

常见配置错误

- -

10.4.13 配置设备的普通地址信息

配置效果

- 设备的地址信息发生变化。

注意事项

- -

配置方法

- 可选配置。
- 用户可根据实际需要配置设备的普通地址信息。

检验方法

显示本地设备的 LLDP 普通地址信息。

- 检查 LLDP 普通地址信息是否和配置的不同。

相关命令

▾ 配置设备的普通地址信息

【命令格式】 配置 LLDP 普通地址信息。用户可以使用 no 选项删除地址信息。

```
{ country | state | county | city | division | neighborhood | street-group | leading-street-dir |
trailing-street-suffix | street-suffix | number | street-number-suffix | landmark |
additional-location-information | name | postal-code | building | unit | floor | room | type-of-place |
postal-community-name | post-office-box | additional-code } ca-word
```

【参数说明】 **country** : 国家代码 , 2 个字符。china : CH

state : 地址信息 CA 类型为 1
county : CA 类型为 2
city : CA 类型为 3
division : CA 类型为 4
neighborhood : CA 类型为 5
street-group : CA 类型为 6
leading-street-dir : CA 类型为 16
trailing-street-suffix : CA 类型为 17
street-suffix : CA 类型为 18
number : CA 类型为 19
street-number-suffix : CA 类型为 20
landmark : CA 类型为 21
additional-location-information : CA 类型为 22
name : CA 类型为 23
postal-code : CA 类型为 24
building : CA 类型为 25
unit : CA 类型为 26
floor : CA 类型为 27
room : CA 类型为 28
type-of-place : CA 类型为 29
postal-community-name : CA 类型为 30
post-office-box : CA 类型为 31
additional-code : CA 类型为 32
ca-word : 地址信息

【命令模式】 LLDP Civic Address 配置模式

【使用指导】 进入 LLDP Civic Address 配置模式后，配置 LLDP 普通地址信息。

📌 删除设备的普通地址信息

【命令格式】 **no** { **country** | **state** | **county** | **city** | **division** | **neighborhood** | **street-group** | **leading-street-dir** | **trailing-street-suffix** | **street-suffix** | **number** | **street-number-suffix** | **landmark** | **additional-location-information** | **name** | **postal-code** | **building** | **unit** | **floor** | **room** | **type-of-place** | **postal-community-name** | **post-office-box** | **additional-code** }

【参数说明】 -

【命令模式】 LLDP Civic Address 配置模式

【使用指导】 进入 LLDP Civic Address 配置模式后，配置 LLDP 普通地址信息。

📌 配置设备类型信息

【命令格式】 **device-type** *device-type*

【参数说明】 *device-type* : 设备类型，缺省为 1，取值范围为 0-2

0 表示设备类型为 DHCP Server

1 表示设备类型为 Switch

2 表示设备类型为 LLDP MED 终端

【命令模式】 LLDP Civic Address 配置模式

【使用指导】 进入 LLDP Civic Address 配置模式后，配置 LLDP 普通地址中设备类型信息。

↘ 恢复设备类型信息

【命令格式】 **no device-type**

【参数说明】 -

【命令模式】 LLDP Civic Address 配置模式

【使用指导】 进入 LLDP Civic Address 配置模式后，恢复 LLDP 普通地址中设备类型信息为缺省值。

配置举例

↘ 配置设备的普通地址信息

【配置方法】 配置设备接口 1 的地址为：交换机设备，地址是国家：CH，城市：Fuzhou，邮编：350000。

```
Ruijie#config
Ruijie(config)#lldp location civic-location identifier 1
Ruijie(config-lldp-civic)# country CH
Ruijie(config-lldp-civic)# city Fuzhou
Ruijie(config-lldp-civic)# postal-code 350000
```

【检验方法】 显示设备接口 1 的 LLDP 普通地址信息。

```
civic location information:
-----
Identifier           :1
country              :CH
device type          :1
city                 :Fuzhou
postal-code          :350000
```

常见配置错误

- -

10.4.14 配置设备的紧急电话号码信息

配置效果

- 更改设备的紧急电话号码信息。

注意事项

- -

配置方法

- 可选配置。
- 用户可根据实际需要配置设备的紧急电话号码信息。

检验方法

显示本地设备的紧急电话号码信息。

- 检查本地设备的紧急电话号码信息是否和配置的相同。

相关命令

▾ 配置设备的紧急电话号码信息

- 【命令格式】 `lldp location elin identifier id elin-location tel-number`
- 【参数说明】 `id` : 表示紧急电话号码信息的配置标识号, 范围为: 1-1024
`tel-number` : 表示紧急电话号码, 范围: 10 – 25 字节
- 【命令模式】 全局模式
- 【使用指导】 使用此命令来配置紧急电话号码信息。

▾ 删除设备的紧急电话号码信息

- 【命令格式】 `no lldp location elin identifier id`
- 【参数说明】 `id` : 表示紧急电话号码信息的配置标识号, 范围为: 1-1024
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

▾ 配置设备的紧急电话号码信息

- 【配置方法】 配置设备接口 1 的紧急电话号码为: 085285555556。

```
Ruijie#config
Ruijie(config)#lldp location elin identifier 1 elin-location 085283671111
```

- 【检验方法】 显示设备接口 1 的紧急电话号码信息。

```
elin location information:
-----
```

```
Identifier          :1
elin number        :085283671111
```

常见配置错误

-

10.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除 LLDP 的统计信息。	clear lldp statistics [interface <i>interface-name</i>]
清除 LLDP 的邻居信息。	clear lldp table [interface <i>interface-name</i>]

查看运行情况

作用	命令
显示本地设备的 LLDP 信息，这些信息将被组织成 TLV 发送给邻居设备。	show lldp local-information [global interface <i>interface-name</i>]
显示本地设备的 LLDP 普通地址信息或者紧急电话号码信息。	show lldp location { civic-location elin-location } { identifier <i>id</i> interface <i>interface-name</i> static }
显示邻居设备的 LLDP 信息。	show lldp neighbors [interface <i>interface-name</i>] [detail]
显示本地设备的 LLDP network-policy 配置策略信息	show lldp network-policy { profile [<i>profile-num</i>] interface <i>interface-name</i> }
显示 LLDP 的统计信息。	show lldp statistics [global interface <i>interface-name</i>]
显示 LLDP 的状态信息。	show lldp status [interface <i>interface-name</i>]
显示端口上可发布的 TLV 配置信息。	show lldp tlv-config [interface <i>interface-name</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 LLDP 错误处理的调试开关。	debug lldp error
打开 LLDP 事件处理的调试开关。	debug lldp event
打开 LLDP 热备份处理的调试开关。	debug lldp ha
打开 LLDP 报文接收的调试开关。	debug lldp packet
打开 LLDP 状态机相关的调试开关。	debug lldp stm

11 QINQ

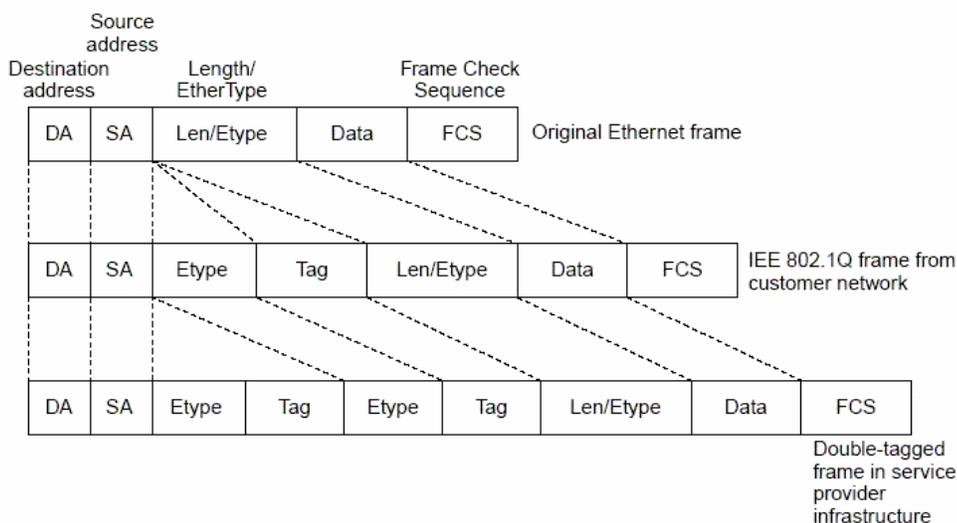
11.1 概述

QinQ 技术是指在用户报文进入服务提供商网络之前封装上一个服务提供商网络的公网 VLAN Tag，而把用户报文中的私网用户 VLAN Tag 当做数据，使报文带着两层 VLAN Tag 穿越服务提供商网络。

在城域网中需要大量的 VLAN 来隔离用户，IEEE 802.1Q 协议仅支持的 4094 个 VLAN 远远不能满足需求。通过 QinQ 技术双层 Tag 封装，在服务提供商网络中报文只根据公网上分配的唯一外层 VLAN Tag 传播，这样不同的私网用户 VLAN 可以重复使用，实际上扩大了用户可利用的 VLAN Tag 数量，同时提供一种简单的二层 VPN 功能。

下图显示了双 Tag 添加的过程：边界设备的入口称为 dot1q-tunnel port 或简称 tunnel port，所有进入边界设备的帧都被当作是 Untagged 帧，而不管它实际上是 Untagged 还是已经带 802.1Q Tag 头的帧，都被封装上服务商的 Tag，VLAN 号为 Tunnel port 的缺省 VLAN。

图 11-1 外 Tag 封装



协议规范

- IEEE 802.1ad

11.2 典型应用

典型应用	场景描述
基于端口的基本QinQ实现二层VPN	企业 A、B 的数据在传送至对端时可以保留原有 VLAN 信息，同时两个企业相同 VLAN 编号的数据在服务提供商网络中传输时不会产生冲突。
基于 C-TAG 的灵活QinQ实现二层	可以根据不同业务 VLAN 分配更灵活的外 Tag，一方面可以实现二层 VPN；另一方面

VPN和业务流管理	实现带宽上网、IPTV 等多种业务的有效区分和实行不同的 QOS 服务策略 比基本 QinQ 更为灵活。
基于ACL的灵活QinQ实现二层VPN和业务流管理	根据 ACL 对下连用户的带宽上网、IPTV 等多种业务进行区分，通过灵活 QinQ 以便针对不同业务流实行不同的 QOS 服务策略。
基于QinQ实现二层协议透传：BPDU和GVRP	处于不同地域的客户网络 A 和 B 可以跨越服务提供商网络进行 MSTP 统一生成树计算或 VLAN 部署，而不影响服务商网络。

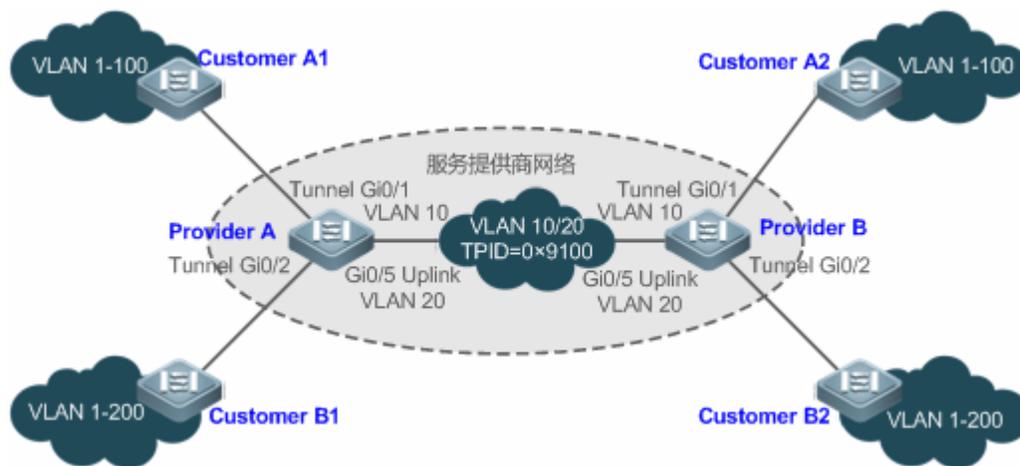
11.2.1 基于端口实现二层VPN业务

应用场景

服务提供商为企业 A 和企业 B 提供 VPN：

- 在公网上企业 A 和企业 B 属于不同的 VLAN，各自通过所属的公网 VLAN 通信。
- 企业 A 和企业 B 内的 VLAN 对公网来说是透明的，企业 A 和企业 B 内的用户 VLAN 可以重复使用并且不冲突。
- Tunnel 会对用户数据报文再封装一层 Native VLAN 的 VLAN Tag。在公网中用户数据报文以 Native VLAN 传播，不影响不同企业用户网络的 VLAN 使用，并实现简单的二层 VPN。

图 11-2



【注释】 Customer A1 和 Customer A2、Customer B1 和 Customer B2 分别为企业用户 A、企业用户 B 所在网络的边缘设备。

Provider A 和 Provider B 为服务提供商网络边缘设备，企业 A 和企业 B 通过提供商边缘设备接入公网。

企业 A 使用的办公网络 VLAN 范围为 VLAN 1-100。

企业 B 使用的办公网络 VLAN 范围为 VLAN 1-200。

功能部属

- 下连用户网络的数据无需区分，在服务提供商边缘设备上启用基本 QinQ 即能二层 VPN 的需求。
- 交换机(包括锐捷交换机)普遍的 TPID 值是 0x8100，但存在部分厂商的交换机 TPID 值采用的不是 0x8100，这时需在在服务提供商网络边缘设备 Uplink 接口上将 TPID 值调整为与第三方设备一样的值。

- 在提供商网络边缘设备 Tunnel 口上设置 cos 的优先级复制和优先级映射功能，并设置 cos 的 QOS 策略(详见 QOS 策略配置文档)，使用户数据报文享用不同的 QOS 策略。

11.2.2 基于C-Tag的灵活QinQ实现二层VPN和业务流管理

应用场景

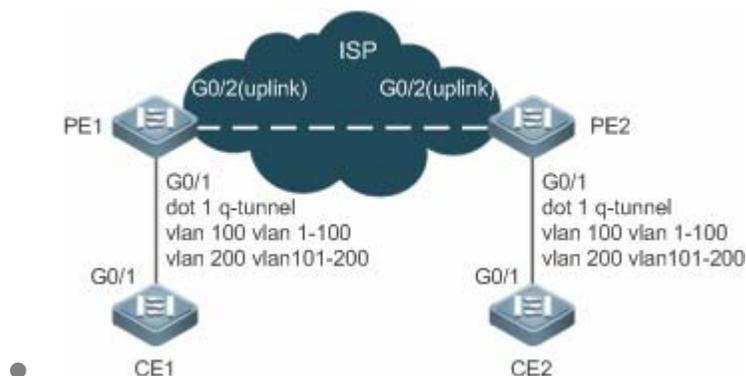
基本 QinQ 只能将用户数据报文封装一层 Native VLAN 的外 Tag，即外层 Tag 的封装依赖于 Tunnel 口的 Native VLAN。灵活 QinQ 提供根据用户报文的 Tag(即 C-Tag)来灵活封装服务提供商(ISP)的外 Tag(即 S-Tag)，以便更灵活实现 VPN 透传和业务流 QOS 策略。

- 宽带上网、IPTV 业务都是城域网的所承载业务的重要部分，城域网服务商网络针对不同业务流划分 VLAN 来区分管理，并提供针对这些 VLAN 或 cos 设置了 QOS 策略服务。可以在服务商边缘设备上运用基于 C-Tag 的 QinQ 将用户的业务流封装相关的 VLAN，在透传的同时利用服务商网络的 QOS 策略进行保障性传输。
- 企业分公司之间实现了统一的 VLAN 规划，重要业务和一般业务分别在不同的 VLAN 范围内，企业网可以利用基于 C-Tag 的灵活 QinQ 透传公司内部的业务，又能利用服务商网络的 QOS 策略优先保障重要业务的数据传输。

如下图所示，城域网内用户端设备通过小区的楼道交换机汇聚，宽带上网、IPTV 业务通过分配不同的 VLAN 进行区分，分别享用不同的 QOS 服务策略。

- 在公网中，宽带上网和 IPTV 的不同业务流以不同的 VLAN 传播，实现用户业务的透传。
- ISP 网络针对 VLAN 或者 cos 设置了 QOS 策略，在服务商边缘设备上可以针对用户业务封装对应的 VLAN 或设置 cos，使得用户业务在 ISP 网络中优先传输。
- 可以通过优先级映射或优先级复制灵活改变用户业务报文的 cos 值，灵活运用服务商网络中的 QOS 服务策略。

图 11-3



- 【注释】 CE1 和 CE2 为连接用户网络的边缘设备，PE1 和 PE2 为提供商服务网络边缘设备。CE1 和 CE2 设备上 VLAN 1-100 为用户宽带上网业务流，VLAN 101-200 为用户 IPTV 业务流。PE1 和 PE2 设备上配置 Tunnel 口和 VLAN 映射以区分不同的业务数据。

功能部属

- 在服务商网络边缘设备 PE1 和 PE2 连接用户网络设备的接口（如本例 PE1 和 PE2 的 G0/1）上配置基于 C-Tag 的灵活 QinQ，实现业务流的划分和透传。
- 如果 ISP 网络基于 VLAN 或 cos 设定了 QOS 策略，在 PE1 和 PE2 设备上将用户业务流映射到相关的 VLAN、或通过优先级映射和优先级复制修改报文的 cos 值，以使用户网络的业务流能使用 ISP 网络的 QOS 策略保障传输。

11.2.3 基于ACL的灵活QinQ实现二层VPN和业务流管理

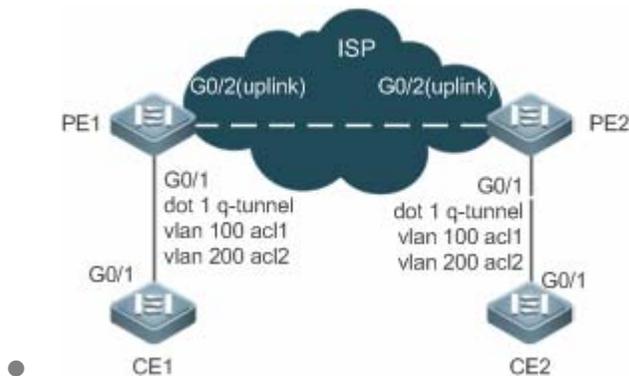
应用场景

用户网络不是根据 VLAN 而是根据 MAC、IP、或协议类型等来划分业务流、或者用户网络存在大量老式的低端网络接入设备，无法通过 VLAN ID 对业务流进行有效的区分时，这时无法根据 C-Tag 为用户网络封装外层 Tag 用以透传和实施 QOS 策略。ACL 可以根据 MAC、IP、协议类型等来划分业务流，灵活 QinQ 借助于 ACL 来区分不同业务添加和修改外层 Tag，从而针对不同业务数据实现 VPN 和 QOS 服务策略。

以下图为例，PE1 和 PE2 根据 ACL 划分的业务流分配不同的 VLAN 实现用户业务透传，如果 ISP 网络中针对不同业务提供了不同的 QOS 服务策略，则可以保障一些业务的优先传输。

- 不同业务数据被封装不同的外层 VLAN Tag，企业数据可以实现透传，企业分公司之间能够互访。
- 通过封装不同的 VLAN 或设置报文的 cos 值，利用 ISP 网络针对 VLAN 或 cos 实施的 QOS 策略优先保障相应业务的数据传输。

图 11-4



- **【注释】** CE1 和 CE2 为用户网络的边缘设备，PE1 和 PE2 为提供商服务网络边缘设备。
PE1 和 PE2 设备上基于 ACL 进行流分类：acl1 识别 PPPOE 协议类型业务流，acl2 识别 IPTV(IPOE)协议类型业务流。
PE1 和 PE2 设备上配置 Tunnel 口，并针对不同的 ACL 识别的业务流设置外 Tag 策略。

功能部署

- 服务商网络边缘设备（PE1 和 PE2）上配置 ACL，用于区分和规划不同的业务数据。
- 在服务商网络边缘设备连接用户网络的的接口（本例为 PE1 和 PE2G0/1 口）上，配置基于 ACL 的灵活 QinQ 功能，对用户业务进行区分及分流。

- 如果 ISP 网络基于 VLAN 或 cos 设定了 QOS 策略，在 PE1 和 PE2 设备上将用户业务流映射到相关的 VLAN、或通过优先级映射和优先级复制修改报文的 cos 值，以使用户网络的业务流能使用 ISP 网络的 QOS 策略保障传输。

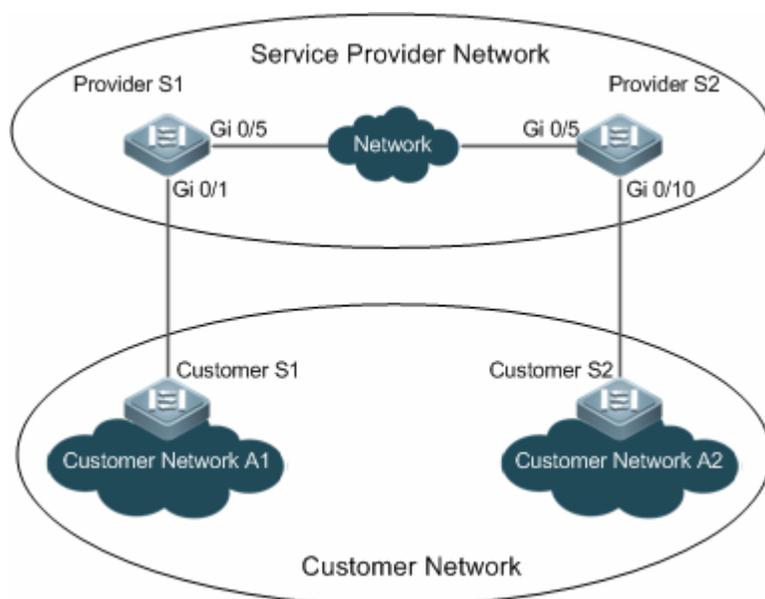
11.2.4 基于QinQ实现二层协议透传：BPDU和GVRP

应用场景

在一个网络中，用户网络之间的二层协议报文不对运营商网络产生影响。

- 用户网络中的二层协议报文对运营商网络是透明的，可以从用户一端网络传到用户另一端网络，对运营商网络不产生影响。

图 11-5



- **【注释】** Customer S1 和 Customer S2 为用户的两端网络边缘设备。
Provider S1 和 Provider S1 为运营商网络的边缘设备。
Provider S1 和 Provider S1 设备上配置全局二层协议透传功能，并在端口 Gi 0/1 和 Gi 0/10 开启二层协议透传功能。

功能部属

- 在服务提供商边缘设备（本例为 Provider S1/Provider S2）连接用户网络的边缘设备的接口上，配置用户所需要透传的二层协议功能，实现用户网络的二层协议报文透明传输，而不影响运营商的网络配置和拓扑。
- 根据用户需要配置 stp 协议透传功能，实现用户网络的 BPDU 报文在运营商网络进行隧道传输，使跨地域的用户网络跨越服务提供商网络统一生成树计算。
- 根据用户需要配置 GVRP 协议透传功能，实现用户网络的 GVRP 报文透传，实现跨网络的用户网络的动态 VLAN 配置。

11.3 功能详解

基本概念

基本 QinQ

基本 QinQ 功能，将接口模式设置为 dot1q-tunnel 并配置一个 Native VLAN，报文从这个接口进来会被封装一层外层 Tag 为 Native VLAN 的 Tag。基本 QinQ 基于端口为报文封装 Native VLAN，不能区分从端口进来的各种流，也不能灵活选择 VLAN 来封装，不够灵活。

灵活 QinQ

灵活 QinQ 主要有两种：基于 C-Tag (Client VLAN Tag) 的灵活 QinQ 和基于 ACL 的灵活 QinQ。

基于 C-Tag 的灵活 QinQ，根据用户的 VLAN 来封装外层 Tag，用以区分不同类型的流和实现透传。

基于 ACL 的灵活 QinQ，ACL 策略能区分不同的业务流，基于 ACL 可以针对不同的流封装不同的 VLAN 实现透传。

TPID

以太网帧 Tag 包含四个字段：TPID(Tag Protocol Identifier，标签协议标志)、 User Priority、CFI、VLAN ID。

对于 TPID 值，缺省采用 IEEE802.1Q 协议规定的 0x8100。也有部分厂商设备的 TPID 值为 0x9100 或其他值，为了和第三方设备兼容，提供 TPID 值的设置，使得报文转发出去时 TPID 值与第三方设备兼容。

优先级映射和优先级复制

以太网帧 Tag 的 User Priority 默认为 0，为普通流。为了保证一些报文优先处理和传输，用户可以设置这个字段。这个字段值对应 QOS 策略中的 cos 值，可以通过配置基于 cos 的 QOS 策略保证业务的优先先。

优先级复制：如果用户报文的 VLAN Tag 设置了用户优先级 cos 值，运营商网络中针对这个 cos 值设置了 QOS 优先级策略，那么可以将用户报文 VLAN Tag 的 cos 值复制给外层 Tag 的 cos，保证用户报文在运营商网络中沿用用户 VLAN 的优先级传输。

优先级映射：运营商网络针对多个的 cos 值设置了不同的 QOS 策略，对应不同的业务流服务。为了保证用户业务在运营商网络享有优先传输的特性，可以根据用户报文 VLAN Tag 的 cos 值设置外层 Tag 的 cos 值。

二层协议透传

STP 报文和 GVRP 报文进入运营商网络可能会影响网络的拓扑，而跨服务商网络的用户网络希望在统一内部拓扑的同时又不影响服务商网络的拓扑。为了不影响运营商网络的拓扑，将用户网络的 STP 报文和 GVRP 报文在运营商网络透传，运营商网络就不会被用户网络影响。

功能特性

功能特性	作用
基本QINQ	配置端口为 dot1q-tunnel 口，指明端口输出报文是否需要带 Tag。
灵活QINQ	根据规则为不同的数据流打上不同的外层 VLAN Tag。
TPID设置	对于 TPID 值，缺省采用 IEEE802.1Q 协议规定的 0x8100。而某些厂商的设备将报文外层 Tag 的 TPID 值设置为 0x9100 或其他值。为了和这些设备兼容，提供了基于端口的报文 TPID 可配置功能。

MAC地址复制	采用基于 ACL 的灵活 QinQ 时，交换机学到的 mac 的 vid 是 native vlan 的。故当采用基于数据流的 vlan 转换时，当报文从对端回来时候，就会发生无法查询到 mac 地址而泛洪的情况，此时可将 native vlan 的 mac 地址复制到外层 Tag 所在的 vlan 中。
二层协议透传	实现用户网络之间二层协议报文的传输而又不影响运营商网络。
优先级复制	服务商网络中基于用户 VLAN Tag 的 User Priority 设置了 QOS 优先级策略，可以通过优先级复制，使外层 VLAN Tag 沿用用户 VLAN 的优先级策略。
优先级映射	添加外层 VLAN Tag 时，可以通过优先级映射，根据内层 VLAN Tag 的 User Priority 设置外层 Tag 的 User Priority，使报文封装外层 Tag 后能利用网络的 QOS 优先级策略。

11.3.1 基本QINQ

基本 QinQ 能简单实现二层 VPN 功能。实现简单，但外层 VLAN Tag 封装方式不够灵活。

工作原理

端口配置成 tunnel 口后，当该端口接收到报文，设备会为该报文打上 tunnel 口缺省 VLAN 的 VLAN Tag。如果接收到的已经是带有 VLAN Tag 的报文，该报文就封装成为双 Tag 的报文；如果接收到的是不带 VLAN Tag 的报文，该报文就封装成带有端口缺省 VLAN Tag 的报文。

相关配置

配置端口模式为 dot1q-tunnel

缺省情况下，基本 QINQ 功能关闭。

在端口模式下使用 `switchport mode dot1q-tunnel` 命令让端口变成 dot1q-tunnel 口。

添加 dot1q-tunnel 口的许可 vlan，并指明输出是否需要带 Tag

在端口模式下使用 `switchport dot1q-tunnel allowed vlan { [add] tagged vlist | [add] untagged vlist | remove vlist }` 命令配置。端口收到响应 VLAN 的报文是会根据设置情况加 Tag 和剥离 Tag。

配置端口的 Native-vlan

在端口模式下使用 `switchport dot1q-tunnel native vlan VID` 命令，设置 dot1q-tunnel 口的缺省 vlan。

如果 native vlan 是以 untag 形式加入许可列表，端口的输出报文是不带 Tag；如果 native vlan 是以 Tag 形式加入许可列表，端口的输出报文会打上 native vlan 的 Tag。所以为了保证上下行报文的通信，必须将端口的 native vlan 以 untag 形式加入端口的 VLAN 许可列表中。

11.3.2 灵活QINQ

灵活 QinQ 可以为不同的数据流打上不同的外层 VLAN Tag，外层 VLAN Tag 封装方式灵活。

工作原理

灵活 QINQ 根据用户 VLAN Tag、MAC 地址、IP 协议、源地址、目的地址、优先级、或应用程序的端口号等信息的不同，封装不同的外层 Tag。借助以上各种分类方法，实现不同用户、不同业务、不同优先级的报文进行外层 VLAN Tag 封装。

当前，在配置上可以使用的具体策略有：

- 根据内层 VLAN Tag 添加外层 VLAN Tag；
- 根据外层 VLAN Tag 修改外层 VLAN Tag；
- 根据内层 VLAN Tag 修改外层 VLAN Tag；
- 根据内、外层 VLAN Tag 修改外层 VLAN Tag；
- 利用 ACL，根据 ACL 添加外层 VLAN Tag；
- 利用 ACL，根据 ACL 修改外层 VLAN Tag；
- 利用 ACL，根据 ACL 修改内层 VLAN Tag；

相关配置

灵活 QINQ 依据不同的匹配规则，相关配置较多，详见配置详解。

11.3.3 TPID设置

工作原理

以太网帧 Tag 包含四个字段：TPID(Tag Protocol Identifier，标签协议标志)、User Priority、CFI、VLAN ID。对于 TPID 值，缺省采用 IEEE802.1Q 协议规定的 0x8100。而某些厂商的设备将报文外层 Tag 的 TPID 值设置为 0x9100 或其他值。为了和这些设备兼容，提供了基于端口的报文 TPID 可配置功能。用户可自行配置端口的 TPID 值，那么这些端口在转发报文时，会将报文的外层 Vlan Tag 中的 TPID 替换为用户设定的值，以达到不同厂商之间 TPID 兼容。

相关配置

TPID 设置功能，缺省情况下是关闭的。

修改 TPID 值

TPID 设置支持基于端口和基于全局配置，以基于端口配置为例：

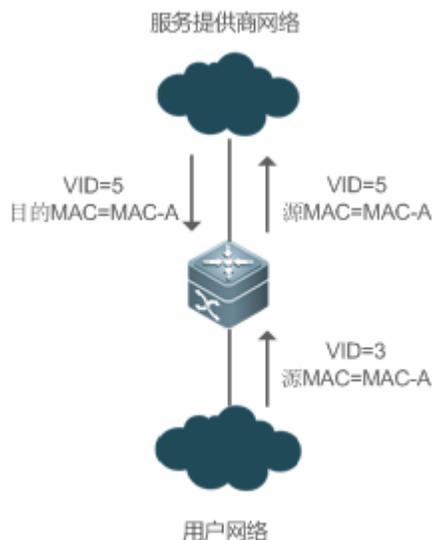
在端口模式下配置 frame-tag tpid 0x9100，可修改 TPID 值为 0x9100。TPID 值的限制见 1.4.5 节。

11.3.4 MAC地址复制

工作原理

采用基于 ACL 的灵活 QinQ 时，交换机学到的 mac 的 vid 是 native vlan 的。使用该策略封装外层 VLAN Tag 时，当报文从对端回来时带的 VLAN 是外 Tag 的 VLAN，就会发生在外层 VLAN 内无法查询到 mac 地址而泛洪的情况。

图 11-6



如上图所示，交换机和用户网络相连端口是 dot1q-tunnel port，且在该端口上配置 Native vlan 是 4，并且配置基于 ACL 策略的 QinQ 对源 MAC 为 A 的报文封装外层 Tag 为 VLAN 5。当交换机收到用户网络 vlan 3 且源 MAC 为 A 的报文后，为其添加 VLAN 5 的外层 Tag，同时由于接收端口的 native VLAN 是 VLAN4，MAC-A 将被学习到 VLAN 4 中。当响应报文返回时，由于其外层 Tag 的 VLAN 是 VLAN 5，将在 VLAN 5 中查找 MAC-A，但是 MAC-A 并没有被学习到 VLAN 5 中，此报文将被泛洪。为此，可以将 Native VLAN 的 mac 地址复制到外层 Tag 所在的 VLAN 中，解决从公网返回的报文被持续泛洪的问题。同样，也可以将外层 Tag 所在 VLAN 的 mac 地址复制到 Native VLAN，解决从用户网络发往公网的报文被持续泛洪的问题。

相关配置

MAC 地址复制功能，缺省情况下是关闭的。

✎ 开启 MAC 地址复制

在 TRUNK 下配置 `mac-address-mapping <1-8> source-vlan src-vlan-list destination-vlan dst-vlan-id`，可开启 MAC 地址复制功能。src-vlan-list、dst-vlan-id 为 VLAN 的取值范围。

11.3.5 二层协议透传

工作原理

为了实现用户网络之间二层协议报文的传输而不对运营商网络产生影响，可以使用二层报文透传功能。当用户网络中二层协议报文进入提供商网络边缘设备时，将目的 MAC 地址改成私有地址在运营商网络中转发，到了另外一端边缘设备后，再将目的 mac 地址改成公有地址回到另一端用户网络，以达到二层协议报文在运营商网络透传的效果。

相关配置

二层协议透传功能，缺省情况下是关闭的。

✎ 开启 STP 透传

在全局模式下，全局开始 STP 透传：l2protocol-tunnel stp

在端口模式下，开启端口 STP 透传：l2protocol-tunnel stp enable

▾ 开启 GVRP 透传

在全局模式下，全局开始 STP 透传：l2protocol-tunnel gvrp

在端口模式下，开启端口 STP 透传：l2protocol-tunnel gvrp enable

11.3.6 优先级复制

工作原理

如果用户 VLAN Tag 的 User Priority（即 cos）设置了值，而服务商网络中针对该 cos 设置了 QOS 优先级策略，可以在封装外层 Tag 的时候把外层 VLAN Tag 的 cos 值设置为和内层 VLAN Tag 一样的值，这样就可以使用户的报文既能透传又能使用服务商网络提供的 QOS 优先级策略。

相关配置

优先级复制功能，缺省情况下是关闭的。

▾ 开启优先级复制

在 dot1q-tunnel 口下配置 inner-priority-trust enable，可开启优先级复制功能。

11.3.7 优先级映射

工作原理

服务商网络中针对一些 cos 设置了 QOS 优先级策略，在封装外层 Tag 的时候按照用户报文的重要性和优先级性，根据用户的内层 VLAN Tag 的 cos 指定外层 VLAN Tag 的 cos 值，使报文封装外层 VLAN Tag 后能享用服务商网络的 QOS 策略。

相关配置

优先级映射功能，缺省情况下是关闭的。

▾ 开启优先级映射

在 dot1q-tunnel 口下配置 dot1q-Tunnel cos inner-cos-value remark-cos outer-cos-value，可开启优先级映射功能。

inner-cos-value、outer-cos-value 取值范围为 0 - 7。

11.4 产品说明



- S7508 产品支持全局配置 4 个 tpid 值，除 0x8100 外其他 3 个可配置为任意值。

11.5 配置详解

配置项	配置建议 & 相关命令	
基本QinQ配置	必须配置。	
	switchport mode dot1q-tunnel	接口设置成 dot1q-tunnel port。
	switchport dot1q-tunnel allowed vlan { [add] tagged vlist [add] untagged vlist remove vlist }	添加 Tunnel 口的许可 vlan，可以指定外层 VLAN 和用户 VLAN 以 Tag 或 UnTag 形式加入 Tunnel 口
	switchport dot1q-tunnel native vlan VID	设置 dot1q-tunnel 口的缺省 vlan。
配置基于C-Tag的灵活QinQ功能	启用该功能必须配置，此功能必须依赖基本 QinQ 配置，灵活 QinQ 功能比基本 QinQ 功能优先级高	
	dot1q outer-vid VID register inner-vid v_list	配置基于内部Tag添加外部Tag的VID策略。
配置基于ACL的灵活QinQ功能	启用该功能必须配置，此功能必须依赖基本 QinQ 配置，灵活 QinQ 功能比基本 QinQ 功能优先级高	
	traffic-redirect access-group acl nested-vlan VID in	配置基于数据流添加外部Tag的VID策略。
配置TPID	可选配置，用于兼容不同厂商的 TPID	
	frame-tag tpid tpid	设置帧 Tag 中的 TPID。如果想设置为 0x9100。那么直接输入 frame-tag tpid 9100。注意默认是 16 进制。该功能需要在出口配置才能生效。
配置MAC地址复制	可选配置，用于解决基于 ACL 中 MAC 地址策略的 QinQ 功能防报文泛洪	
	mac-address-mapping x source-vlan src-vlan-list destination-vlan dst-vlan-id	将接口在源 vlan 中学习到的动态地址复制到目的 vlan 中。
配置外层和内层VLAN Tag修改策略	可选配置，用于根据网络拓扑适当调整运营商网络中数据报的外层 Tag 和内层 Tag	
	dot1q relay-vid VID translate local-vid v_list	配置基于外部Tag修改外层Tag的VID策略。
	dot1q relay-vid VID translate inner-vid v_list	配置基于内部Tag修改外层Tag的VID策略。

	dot1q new-outer-vlan VID translate old-outer-vlan vid inner-vlan v_list	配置基于外层 Tag + 内层 Tag 修改外部 Tag 的修改外层 VID 策略。
	traffic-redirect access-group acl outer-vlan VID in	配置基于 ACL 修改外层 Tag 的 VID 策略。
	traffic-redirect access-group acl inner-vlan VID out	配置基于 ACL 修改内层 Tag 的 VID 策略。
配置优先级复制和优先级映射	 可选配置，用于沿用用户网络中用户数据的优先级策略	
	inner-priority-trust enable	复制内层 Tag(C-Tag)的 priority 字段值到外层 Tag 的 priority 字段值 (S-Tag)。
	 可选配置，用于根据外层 Tag 设置用户数据的优先级策略	
	dot1q-Tunnel cos inner-cos-value remark-cos outer-cos-value	根据内层 Tag(C-Tag)的 priority 字段值设置外层 Tag 的 priority 字段值 (S-Tag)。
配置二层协议透传	 可选配置，用于透传 MSTP 和 GVRP 协议报文，满足用户网络的拓扑而不影响运营商网络的拓扑	
	I2protocol-tunnel stp	配置全局使能 STP 协议报文透传功能。
	I2protocol-tunnel stp enable	在接口上使能 STP 协议报文透传功能。
	I2protocol-tunnel gvrp	配置全局使能 GVRP 协议报文透传功能。
	I2protocol-tunnel gvrp enable	在接口上使能 GVRP 协议报文透传功能。
	I2protocol-tunnel{STP GVRP}tunnel-dmac mac-address	配置相应协议的透传地址。

 Qinq 配置有如下限制:

-  路由口不能设置为 Tunnel Port。
-  配置为 Tunnel 的端口不能再启用 802.1x 功能。
-  配置为 Tunnel 的端口不能再启用端口安全功能。
-  配置 Tunnel Port 作为 RSPAN 的源口时，外部 TAG 中的 VID 等于 RSPAN VLAN 的报文视为监控数据流。
-  对于应用在 Tunnel Port 上的 ACL，若要匹配用户 TAG 中的 VID，需要使用 inner 关键词。
-  请与服务商网络相连的用户网络的出口也配置为 Uplink 口，如果在用户网络中配置了 Qinq 端口的服务商 Tag 的 TPID 值，那么用户网络出口的 Uplink 口的服务商 Tag 的 TPID 值也需要配置为相同值。
-  接口的 MTU 值默认为 1500 字节。当为报文加上外层 VLAN Tag 后，报文的长度会增加 4 个字节，建议用户适当增加运营商网络中各接口的 MTU 值，至少为 1504 字节。
-  在设备上配置了 Qinq 口之后，若要在设备开启 igmp snooping，需要使用 SVGL 共享模式，否则 igmp snooping 在 Qinq 口上将无法正常工作。

 若某个报文同时匹配两个或两个以上数据流策略的规则，且数据流策略未指明优先级时，只执行其中一条策略。建议指定优先级。

11.5.1 基本QINQ配置

配置效果

- 基于端口的 QINQ 策略，实现二层 VPN 功能。

注意事项

- 建议不要将服务商网络中 trunk 口的 native vlan 设置为 tunnel 口的缺省 vlan，因为 trunk 口会剥去 VID 为其 native vlan 的 Tag。

配置方法

配置端口模式为 dot1q-tunnel 口

- 必须配置，把端口模式设置为 dot1q-tunnel 口。

设置 dot1q-tunnel 口的 native vlan

- 必须配置。
- 配置端口的 native vlan 为供应商网络的 vlan。
- 配置 native vlan 后，该端口并没有真正加入端口。还必须将 native vlan 必须以 untag 形式加入该端口的许可列表。

添加 dot1q-tunnel 口的许可 vlan

- 必须配置。
- 配置 native vlan 后，必须将 native vlan 以 untag 形式加入端口的 vlan 的许可列表。
- 对于基于端口的 QinQ 功能，用户网络的 VLAN 无需加入 Tunnel 口的 VLAN 许可列表。
- 灵活 QinQ 需要根据实际情况，将用户网络的 VLAN 以 Tag 或者 UNTag 形式加入端口的 VLAN 许可列表。

检验方法

查看配置为 dot1q-tunnel 端口信息：

- 检查设备端口 dot1q-tunnel 是否打开和 dot1q-tunnel 端口配置信息是否正确。

相关命令

配置端口模式为 dot1q-tunne 口

【命令格式】 **switchport mode dot1q-tunnel**

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 -

配置 dot1q-tunnel 口的 native vlan

【命令格式】 **switchport dot1q-tunnel native vlan VID**

【参数说明】 VID：指定的 native vlan，范围为：1-4094，缺省为 1

【命令模式】 接口配置模式

【使用指导】 Native vlan 设置成供应商网络的 vlan。

添加 dot1q-tunnel 口的许可 vlan，并指明输出相应许可 vlan 的报文时，是否需要带 Tag

【命令格式】 **switchport dot1q-tunnel allowed vlan { [add] tagged vlist | [add] untagged vlist | remove vlist }**

【参数说明】 v_list：端口允许通过的 vlan 列表

【命令模式】 接口配置模式

【使用指导】 通过配置该命令，可增加/删除 dot1q-tunnel 口的许可 vlan，并指明输出是带 Tag 还是 untag。

对于基本 QINQ 功能，仅需要将端口的 native vlan 以 untag 形式加入端口的许可列表。

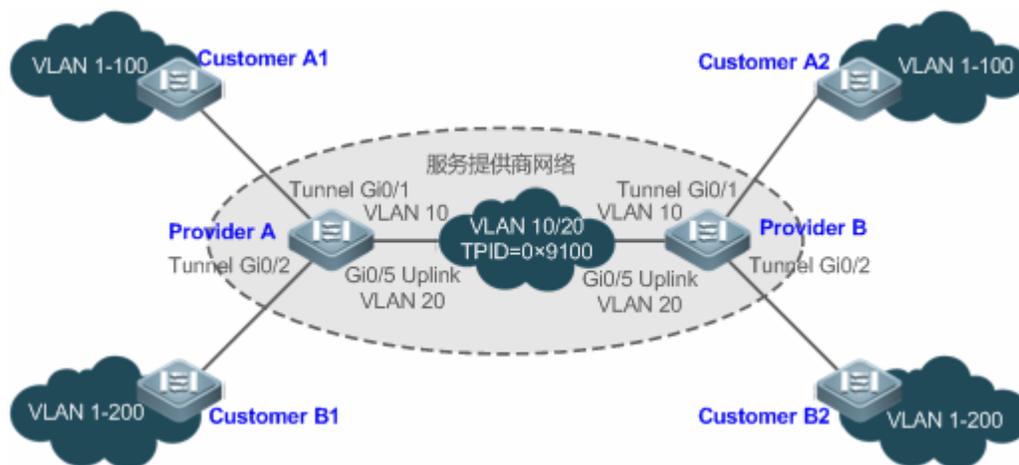
配置举例

i 以下配置举例，仅介绍 QINQ 相关的配置。

配置基本 QinQ，实现二层 VPN

【网络环境】

图 11-7



- 【配置方法】
- 在服务提供商网络边缘设备上配置 tunnel 口，将用户网络的边缘设备连接在该端口上。
 - 设置 tunnel 口的 native vlan，并将 native vlan 以 untag 形式加入端口的 VLAN 许可列表。
 - 在用户网络根据用户需求配置用户 VLAN。

i 开启 QinQ 功能的设备会为用户报文封装其他 VLAN 的外层 Tag，不会按报文中原始的 VLAN 进行转发，因此不需要在服务提供商网络设备上配置用户 VLAN

i 交换机的 TPID 一般采用 IEEE802.1Q 中规定的 0x8100，存在部分第三方设备的 TPID 采用其它值，网

络中存在该类型设备时，需要在连接该设备的出口上设置 TPID 值，以便兼容。

- !** 当边缘设备连接服务商网络的上链口或服务提供商设备之间相互连接的接口为 Trunk port、Hybrid port 的时候，请避免将 Trunk port 或 Hybrid port 的 Native vlan 设置为 tunnel 口的缺省 vlan。因为当报文从 Trunk port 或 Hybrid port 输出时，会被剥去 VID 为其 Native vlan 的 Tag。

Provider A

第一步，创建服务商 VLAN 10、20 用于区别两个企业的数据库

```
ProviderA#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
ProviderA(config)#vlan 10
ProviderA(config-vlan)#exit
ProviderA(config)#vlan 20
ProviderA(config-vlan)#exit
```

第二步，在连接企业 A 网络的接口上启用基本 QinQ 功能，使用 VLAN10 对企业 A 网络的数据进行隧道传输。

```
ProviderA(config)#interface gigabitEthernet 0/1
ProviderA(config-if-GigabitEthernet 0/1)#switchport mode dot1q-tunnel
ProviderA(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel native vlan 10
ProviderA(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel allowed vlan add untagged 10
```

第三步，在连接企业 B 网络的接口上启用基本 QinQ 功能，使用 VLAN20 对企业 B 网络的数据进行隧道传输。

```
ProviderA(config)#interface gigabitEthernet 0/2
ProviderA(config-if-GigabitEthernet 0/2)#switchport mode dot1q-tunnel
ProviderA(config-if-GigabitEthernet 0/2)#switchport dot1q-tunnel native vlan 20
ProviderA(config-if-GigabitEthernet 0/2)#switchport dot1q-tunnel allowed vlan add untagged 20
```

第四步，配置 Uplink 口

```
ProviderA(config)# interface gigabitEthernet 0/5
ProviderA(config-if-GigabitEthernet 0/5)#switchport mode uplink
```

第五步，在 Uplink 口上调整输出报文的 TPID 值为第三方设备可识别的值，如 0x9100

```
ProviderA(config-if-GigabitEthernet 0/5)#frame-tag tpid 9100
```

第六步，配置 Provider B，与 Provider A 类似，这里不再赘述。

【检验方法】

- 从 Customer A1 发送一个 VLAN 100、目的 MAC 为 Customer A2 用户 2 的 MAC 的报文，从 Provider A 出来
- 报文被打上 tunnel 口的外层 Tag，到达 Customer A2 的报文是用户原有的 VLAN 100。
- 查看 Tunnel 口的配置是否正确。
- 查看需要的 TPID 值设置是否正确。

Provider A

```
ProviderA#show running-config
interface GigabitEthernet 0/1
switchport mode dot1q-tunnel
switchport dot1q-tunnel allowed vlan add untagged 10
switchport dot1q-tunnel native vlan 10
spanning-tree bpdupfilter enable
!
```

```
interface GigabitEthernet 0/2
  switchport mode dot1q-tunnel
  switchport dot1q-tunnel allowed vlan add untagged 20
  switchport dot1q-tunnel native vlan 20
  spanning-tree bpdufilter enable
!
interface GigabitEthernet 0/5
  switchport mode uplink
  frame-tag tpid 0x9100
ProviderA#show interfaces dot1q-tunnel

=====Interface Gi0/1=====
Native vlan: 10
Allowed vlan list:1,10,
Tagged vlan list:

=====Interface Gi0/2=====
Native vlan: 20
Allowed vlan list:1,20,
Tagged vlan list:
ProviderA#show frame-tag tpid
Ports      Tpid
-----
Gi0/5      0x9100
```

Provider B

Provider B 设备上的配置验证同 Provider A 的类似，请参考上文 Provider A 的配置验证过程，此处不再重复描述。

常见错误

- 没有将 tunnel 口的 native vlan 以 untag 形式加入端口的 VLAN 许可列表里面。
- 存在第三方设备的 TPID 值不是默认的 0x8100，在连接第三方设备的出口上未设置 TPID，导致报文无法被第三方设备识别。

11.5.2 配置基于C-Tag的灵活QinQ功能

配置效果

- 根据用户 VLAN (C-Tag) 为用户报文灵活封装外层 VLAN Tag (S-Tag) ，实现企业用户数据的二层 VPN 和业务流优先传输管理。

注意事项

- 必须依赖基本 QinQ 配置。
- 灵活 QINQ 的一些规则，由于有些芯片的限制，会出现一些产品的支持情况和限制。
- 如果需要沿用用户网络 VLAN Tag 优先级，可以通过配置优先级复制功能，使用户报文封装外层 Tag 后仍沿用用户 Tag 的优先级。
- 如果运营商网络中需要用户数据包采用外层 VLAN Tag 的优先级和优先级传输，还需要配置优先级复制功能将外层 Tag 的 cos 设置为指定值。

配置方法

配置基于内部 Tag 添加外部 Tag 的 VID 策略

- 必须执行此配置项。
- 网络环境中，对 dot1q-tunnel 端口上收到的报文，需要根据内层 Tag 的 VID 修改外层 Tag 的 VID。按照该命令功能，可以通过指定内部 VLAN 添加与内部 VID 相同的外部 VID，并将出口加入该 VLAN 的 UnTag 端口集中，可以实现从出口输出的为原始内部 Tag 报文。

i 基于数据流的 VID 变更策略表优先于基于端口和 C-Tag 的 VID 变更策略表生效。

i 当 AP 口添加或者删除成员口时，AP 上配置的 VID 添加策略或修改策略会被删除，需要重新配置。建议用户配置完 AP 成员后，再在 AP 上进行 VID 策略配置。

! 在 Tunnel port 上必须允许外层 Tag VLAN(包括 Native VLAN)通过，同时接入公网的接口也必须允许这些 VLAN 报文通过。

检验方法

企业分公司之间规划 VLAN 内用户能互通。

- 企业分公司的 VLAN 内的用户能实现二层 VPN。
- 能通过添加外 Tag、优先级复制或优先级映射等实现不同业务的优先级传输策略。

相关命令

配置基于内 Tag 添加外 Tag 的策略

【命令格式】 `dot1q outer-vid VID register inner-vid v_list`

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 -

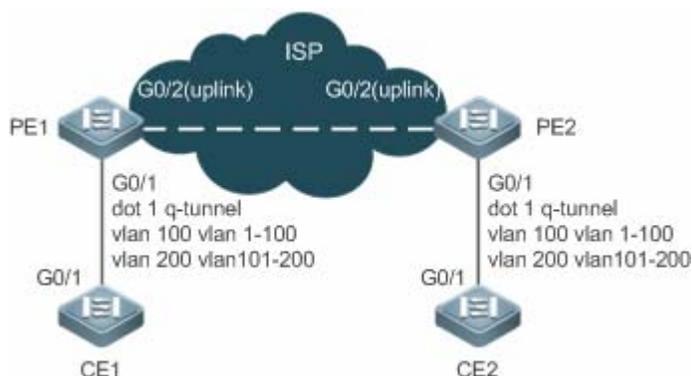
配置举例

i 下面配置举例，仅介绍与灵活 QinQ 相关配置

基于 C-Tag 灵活实现二层 VPN 及业务流管理

【网络环境】

图 11-8



【配置方法】

- 在服务商网络边缘设备 PE1 和 PE2 连接用户网络边缘设备的接口配置为 Tunnel 口。
- 根据用户网络中的业务数据 VLAN 在 Tunnel 口上配置外层 Tag 添加策略。
- 如果 ISP 网络提供了基于 VLAN 的 QOS 优先级策略，可以将重要业务或质量要求较高的业务流的外层 Tag 封装为相关 QOS 策略对应的 VLAN。
- 如果 ISP 网络提供了基于 cos 的 QOS 优先级策略，如果 QOS 策略对应的 cos 是用户 VLAN Tag 的 cos 值，可以通过优先级复制将用户 VLAN Tag 的 cos 复制给外层 VLAN Tag 的 cos，使报文封装外层 Tag 后能沿用户 VLAN Tag 的优先级策略。
- 如果 ISP 网络提供了基于 cos 的 QOS 优先级策略，还可以通过优先级映射，根据用户 VLAN Tag 的 cos 值设定外层 VLAN Tag 的 cos 值为 QOS 策略对应的 cos 值，使得报文封装外层 Tag 后能使用优先级策略。

PE1

第一步，配置透传 VLAN

```
PE1#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
PE1(config)#vlan 100
PE1(config-vlan)#exit
PE1(config)#vlan 200
PE1(config-vlan)#exit
```

第二步，在接入交换机的下联口上配置基于 C-Tag 添加外层 VLAN Tag 的灵活 QinQ 功能

配置 Gi 0/1 接口类型为 Tunnel port

```
PE1(config)#interface gigabitEthernet 0/1
PE1(config-if)# switchport mode dot1q-tunnel
```

将服务商 VLAN101, 201 加入到 Tunnel port 许可 VLAN 列表，并配置对端报文返回至 Tunnel port 输出时剥离服务商 Tag

```
PE1(config-if)# switchport dot1q-tunnel allowed vlan add untagged 100,200
```

配置从 Tunnel port 输入的 vlan1-100 (C-tag) 的数据帧在服务商网络中传输打上 vlan 100(S-tag) 的标签

```
PE1(config-if)# dot1q outer-vid 100 register inner-vid 1-100
```

配置从 Tunnel port 输入的 vlan101-200 (C-tag) 的数据帧在服务商网络中传输打上 vlan 200(S-tag) 的标签

```
PE1(config-if)# dot1q outer-vid 200 register inner-vid 101-200
```

第三步，将连接服务商网络的接入公网的接口配置 Uplink 口

```
PE1(config)# interface gigabitEthernet 0/2
```

```
PE1(config-if-GigabitEthernet 0/2)#switchport mode uplink
```

PE2

和 PE1 上配置类似，这里就不再赘述

【检验方法】

- 确认配置是否正确，关注点：下接口类型是否为 dot1q-tunnel，外层 Tag VLAN 是否已加入接口的许可 VLAN
- 列表，接口上的映射策略是否正确，uplink 口是否设置正确。
- 确认 VLAN 映射策略是否正确。

PE1

```
PE1#show running-config interface gigabitEthernet 0/1
interface GigabitEthernet 0/1
  switchport mode dot1q-tunnel
  switchport dot1q-tunnel allowed vlan add untagged 100,200
  dot1q outer-vid 100 register inner-vid 1-200
  dot1q outer-vid 200 register inner-vid 101-200
  spanning-tree bpdupfilter enable
```

!

第二步，确认基于 C-Tag 添加 TAG 的映射策略；关注点：内外层 VLAN 标签映射关系是否正确。

```
PE1#show registration-table
```

Ports	Type	Outer-VID	Inner-VID-list
Gi0/1	Add-outer	100	1-200
Gi0/1	Add-outer	200	101-200

11.5.3 配置基于ACL的灵活QinQ功能

配置效果

- 根据用户网络的流特征基于 ACL 进行分类，对用户报文灵活封装外层 VLAN Tag (S-Tag)，便于运营商网络不同用户业务的管理。

注意事项

- 必须依赖 QinQ 基本配置。
- 灵活 QINQ 的一些规则，由于有些芯片的限制，会出现一些产品的支持情况和限制。
- 如果需要沿用用户网络 VLAN Tag 优先级，可以通过配置优先级复制功能，使用户报文封装外层 Tag 后仍沿用用户 Tag 的优先级。

- 如果运营商网络中需要用户数据包采用外层 VLAN Tag 的优先级和优先级传输，还需要配置优先级复制功能将外层 Tag 的 cos 设置为指定值。
-
- ❗ 基于 ACL 的 VID 变更策略比基于端口和 C-Tag 的 VID 变更策略优先级高。
 - ❗ 当 ACL 被删除时，与此 ACL 相关的策略会被自动删除。
 - ❗ 当 dot1q-tunnel 口收到 ≥ 2 层 Tag 的报文时，无法采用基于流的匹配规则来添加外层 Tag。
 - ❗ 若某个报文同时匹配两个或两个以上数据流添加 VID 策略，且数据流策略未指明优先级时，只执行其中一条策略。建议指定优先级。
- ⚠ 在 Tunnel port 上必须允许外层 Tag VLAN(包括 Native VLAN)通过，同时接入公网的接口也必须允许这些 VLAN 报文通过。
-

配置方法

配置基于数据流添加外部 Tag 的 VID 策略。

- 必须执行此配置项。
- 网络环境中，对 dot1q-tunnel 端口的输入报文，有时需要能够由报文内容的不同，转发时为报文添加外部 Tag 指定不同的 VID。

检验方法

企业分公司之间业务内用户能互通，并且还可以通过 VPLS 相关配置能使特定业务数据得到较高的传输优先级

- 企业分公司相同业务的用户能实现二层 VPN。
- 能通过添加外 Tag、优先级复制或优先级映射等实现不同业务的优先级传输策略。

相关命令

配置基于数据流添加外部 Tag 的 VID 策略。

- 【命令格式】 `traffic-redirect access-group acl nested-vlan VID in`
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 -

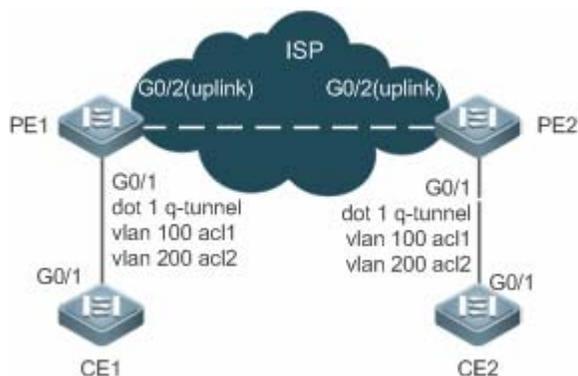
配置举例

- ❗ 下面配置举例，仅介绍与灵活 QinQ 相关配置

基于 ACL 灵活实现二层 VPN 及业务流管理

【网络环境】

图 11-9



【配置方法】

- 在服务商网络边缘设备 PE1 和 PE2 连接用户网络边缘设备的接口配置为 Tunnel 口。
- 在 PE1 和 PE2 设备上配置 ACL 策略，区分用户网络不同的业务流。
- 根据 ACL 区分的业务流在 Tunnel 口上配置外层 Tag 添加策略。
- 如果 ISP 网络提供了基于 VLAN 的 QoS 优先级策略，可以将重要业务或质量要求较高的业务流的外层 Tag 封装为相关 QoS 策略对应的 VLAN。
- 如果 ISP 网络提供了基于 cos 的 QoS 优先级策略，并且用户网络的报文带 Tag，那么可以通过优先级映射和优先级复制设定外层 Tag 的 cos 值，使用户报文封装外层 Tag 后使用 QoS 优先级策略。

PE1

第一步，创建用与区分流的 ACL，PPPOE 协议类型 0x8863/0x8864，匹配 IPOE 协议类型 0x0800

```
PE1#configure terminal
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
PE1(config)# expert access-list extended acl1
```

```
PE1(config-exp-nacl)# permit 0x8863 any any
```

```
PE1(config-exp-nacl)# permit 0x8864 any any
```

```
PE1(config-exp-nacl)#exit
```

```
PE1(config)# expert access-list extended acl2
```

```
PE1(config-exp-nacl)#permit 0x0800 any any
```

第二步，创建服务商 VLAN 100、200，用于区分用户数据。

```
PE#configure terminal
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
PE1(config)#vlan 100
```

```
PE1(config-vlan)#exit
```

```
PE1(config)#vlan 200
```

```
PE1(config-vlan)#exit
```

第三步，在接入交换机的下联口上配置基于 ACL 添加外层 VLAN Tag 的灵活 QinQ 功能

配置 Gi 0/1 接口类型为 Tunnel port

```
PE1(config)#interface gigabitEthernet 0/1
```

```
PE1(config-if)# switchport mode dot1q-tunnel
```

将服务商 VLAN100, 200 加入到 Tunnel port 许可 VLAN 列表，并配置当对端报文返回至 Tunnel port 输出时剥离服务商网络 Tag

```
PE1(config-if)#switchport dot1q-tunnel allowed vlan add untagged 100,200
```

配置从 Tunnel port 输入的匹配 ACL1 的数据帧在服务商网络中传输打上 vlan 101(S-tag) 的标签

```
PE1(config-if)# traffic-redirect access-group acl1 nested-vlan 100 in
```

配置从 Tunnel port 输入的匹配 ACL2 的数据帧在服务商网络中传输打上 vlan200(S-tag) 的标签

```
PE1(config-if)# traffic-redirect access-group acl1 nested-vlan 200 in
```

第三步，将连接服务商网络的接入公网的接口配置 Uplink 口

```
PE1(config)# interface gigabitEthernet 0/2
```

```
PE1(config-if-GigabitEthernet 0/2)#switchport mode uplink
```

【检验方法】 企业分公司之间业务内用户能互通，用户相关业务能得到优先级保障。

- 企业分公司的 VLAN 内的用户能实现二层 VPN。
- 确认 ACL 的配置是否正确。
- 业务的优先级正确。
- 确认配置是否正确：下连接口类型是否为 dot1q-tunnel，外层 Tag VLAN 是否已加入接口的许可 VLAN 列表，接口上的映射策略是否正确等。

PE1

第一步, 查看 Tunnel 配置是否正确

```
Ruijie#show running-config interface gigabitEthernet 0/1
```

```
interface GigabitEthernet 0/1
  switchport mode dot1q-tunnel
  switchport dot1q-tunnel allowed vlan add untagged 100,200
  traffic-redirect access-group acl1 nested-vlan 100 in
  traffic-redirect access-group acl2 nested-vlan 200 in
  spanning-tree bpdufilter enable
!
```

第二步，确认基于 C-Tag 添加 TAG 的映射策略；关注点：内外层 VLAN 标签映射关系是否正确。

```
PE1#show traffic-redirect
```

Ports	Type	VID	Match-filter
Gi0/1	Nested-vid	101	acl1
Gi0/1	Nested-vid	201	acl2

常见错误

- ACL 策略没有配置。
- ACL 策略是根据 MAC 地址来划分流，没有配置 MAC 地址复制功能，存在报文泛洪情况。

11.5.4 配置TPID

配置效果

- 实现服务商网络设备 Tag 中的 TPID 值，兼容第三方设备的不同 TPID 值。

注意事项

- 如果服务商网络边缘设备接入服务商网络的接口连接的第三方设备的 TPID 值不是 IEEE 802.1Q 默认的 0x8100，则需要在该接口上设置 TPID 值。

 不允许将 `tpid` 指定为如下知名类型 0x0806(ARP)、0x0200(PUP)、0x8035(RARP)、0x0800(IP)、0x86DD(IPv6)、0x8863/0x8864(PPPoE)、0x8847/0x8848(MPLS)、0x8137(IPX/SPX)、0x8000(IS-IS)、0x8809(LACP)、0x888E(802.1x)、0x88A7(集群)、0x0789(我司保留)。

配置方法

- 如果存在第三方设备的 TPID 值不是 0x8100，必须在连接第三方设备的接口上设置 TPID。

检验方法

查看 TPID 是否设置成功。

相关命令

配置端口的 TPID

【命令格式】 `frame-tag tpid tpid`

【参数说明】 `tpid`:新的 TPID 值

【命令模式】 接口配置模式

【使用指导】 第三方设备的 TPID 值不是 0x8100 时，在连接第三方设备的接口上设置该值。

配置举例

配置端口的 TPID 值。

- 【配置方法】
- 配置端口的 TPID 值。

```
Ruijie(config)# interface gigabitethernet 0/1
Ruijie(config-if)# frame-tag tpid 9100
```

- 【检验方法】
- 查看接口上的 TPID 值。

```
Ruijie# show frame-tag tpid interfaces gigabitethernet 0/1
Port      tpid
-----
Gi0/1     0x9100
```

常见错误

- 无

11.5.5 配置 MAC地址复制

配置效果

- 当一个端口学习到动态地址时，将其从一个 vlan 复制到另一个 vlan 中。
- 在基于 ACL 的灵活 QinQ 划分业务流时，如果 ACL 规则是根据 MAC 地址来划分流，避免可能存在的报文泛洪的问题。

注意事项

- i** 关闭 VLAN 间 MAC 地址复制功能后，系统将删除目的 VLAN 中通过该功能学到的其他 VLAN 的所有 MAC 地址表项。
- !** VLAN 间 MAC 地址复制功能在一个端口下对某个目的 VLAN 只能配置一次。如果需要修改，必须先删除之前的配置，然后再重新进行配置。
- !** 不能和 share vlan 共用；不能将地址复制到动态 VLAN 中。
- !** 每个端口最多只能配置 8 个目的 VLAN。即使端口不在指定的目的 VLAN 中，该功能也会生效。
- !** 地址复制不能在 host/promiscuous 口上配置，不能在打开端口安全，镜像的目的口，打开 1x 功能的端口上配置。
- !** 只复制动态地址；静态地址不复制；地址表已满时，不复制；开启功能前，源地址已经存在的 mac 地址不复制。
- !** 复制地址的优先级比动态地址高，比其它类型地址的优先级低。
- !** 当 MAC 地址老化时候，由其复制出来的地址也要相应老化；当 MAC 地址被删除时，由其复制出来的地址也将被自动删除。
- !** 不支持热备，故当发生主从切换后，建议用户关闭复制功能，再重新打开。
- i** 用户不能手工删除通过 VLAN 间 MAC 地址复制功能获得的 MAC 地址表项，如果确实需要删除该表项，可以通过关闭 VLAN 间 MAC 地址复制功能实现。

配置方法

配置 MAC 地址复制

- 如果需要避免报文泛洪，则应该执行此配置项，将 MAC 地址从一个 vlan 复制到另一个 vlan。

检验方法

- 查看指定 VLAN 的 MAC 地址是否正确的复制到另一个 VLAN。

相关命令

配置 MAC 地址复制

【命令格式】 **mac-address-mapping** *x source-vlan src-vlan-list destination-vlan dst-vlan-id*

【参数说明】 *x* : MAC 地址复制索引号, 只能配置<1-8>。

src-vlan-list : 源 vlan 列表。

dst-vlan-id : 目的 vlan 列表。

【命令模式】 接口配置模式

【使用指导】 -

配置举例

配置 MAC 地址复制功能。

- 【配置方法】
- 配置 MAC 地址复制功能。

```
Ruijie(config)# interface gigabitethernet 0/1
Ruijie(config-if)# switchport mode trunk
Ruijie(config-if)#mac-address-mapping 1 source-vlan 1-3 destination-vlan 5
```

- 【检验方法】
- 查看端口上配置是否生效。
 - 发送源 VLAN 的报文, 查看设备上该报文的源 MAC 也有一份学在目的 VLAN 上。

```
Ruijie# show interfaces mac-address-mapping
Ports      destination-VID  Source-VID-list
-----
Gi0/1      5                1-3
```

常见错误

- 见注意事项。

11.5.6 配置外层和内层VLAN Tag修改策略

配置效果

- 根据实际组网需求修改外层 Tag 或内层 Tag。

注意事项

- i** 基于 ACL 的 VID 变更策略比基于端口和 C-Tag 的 VID 变更策略优先级高。
- i** 当 ACL 被删除时，与此 ACL 相关的策略会被自动删除。
- i** 修改策略只能在 Access、Trunk、Hybrid、Uplink 口上起作用。
- i** 修改策略主要是针对服务商网络中需要调整内外层 Tag 的需求。
- i** 若某个报文同时匹配两个或两个以上数据流添加 VID 策略，且数据流策略未指明优先级时，只执行其中一条策略。建议指定优先级。

配置方法

配置基于内层 Tag 修改外部 Tag 的 VID 策略

- 可选配置。
- 如果需要灵活的根据内层 Tag 的 VID 修改外部 Tag 的 VID，则必须执行此配置项。
- 对 Access，Trunk，Hybrid，Uplink 端口的输入报文，有时需要根据报文内部 Tag 中的不同 VID，将外部 Tag 的 VID 修改为不同的 VID 值。

配置基于外层 Tag + 内层 Tag 的 VID 修改外层 VID 策略

- 可选配置。
- 如果需要灵活的根据内外层 Tag 的 VID 修改外部 Tag 的 VID，则必须执行此配置项。
- 对 Access，Trunk，Hybrid，Uplink 端口的输入报文，有时需要根据报文内部 Tag 的 VID + 报文外部 Tag 的 VID，将外部 Tag 的 VID 修改为不同的值。

配置基于外层 Tag 修改外层 Tag 的 VID 策略

- 可选配置。
- 如果需要灵活的根据外层 Tag 的 VID 修改外部 Tag 的 VID，则必须执行此配置项。
- 网络环境中，对 Access，Trunk，Hybrid，Uplink 端口的输入报文，有时需要能够由报文外部 Tag 中的不同 VID，指定在转发时修改为不同的外部 Tag 的 VID。

配置基于 ACL 的内部 vid 修改策略表

- 可选配置。
- 对 Access，Trunk，Hybrid，Uplink 端口的输出报文，有时需要能够由报文内容修改内层 Tag 的 VID。
- 需要先配置 ACL 设置区分数据流

配置基于 ACL 的内部 vid 修改策略表

- 可选配置。
- 对 Access，Trunk，Hybrid，Uplink 端口的输出报文，有时需要能够由报文内容修改外层 Tag 的 VID。
- 需要先配置 ACL 设置区分数据流。

检验方法

查看端口上 Tag 修改策略是否生效，端口收到报文后是否依据策略修改报文的 Tag。

相关命令

配置基于外层 Tag 修改外层 Tag 的 VID 策略

- 【命令格式】 **dot1q relay-vid VID translate local-vid v_list**
- 【参数说明】 VID : 修改后的外部 tag 中的 vid。
v_list : 输入报文外层 vid 列表。
- 【配置模式】 接口配置模式
- 【使用指导】 -

配置基于内层 Tag 修改外层 Tag 的 VID 策略

- 【命令格式】 **dot1q relay-vid VID translate inner-vid v_list**
- 【参数说明】 VID : 修改后的外部 tag 中的 vid。
v_list : 输入报文内层 vid 列表。
- 【命令模式】 接口模式
- 【使用指导】 -

配置基于基于内层 Tag + 外层 Tag 修改外层 Tag VID 的策略

- 【命令格式】 **dot1q new-outer-vlan new-vid translate old-outer-vlan vid inner-vlan v_list**
- 【参数说明】 new-vid : 输入报文修改后的新外层 Tag VID
vid : 输入报文修改前的外层 Tag VID。
v_list : 输入报文内层 vid 列表。
- 【命令模式】 接口模式
- 【使用指导】 -

配置基于 ACL 的内层 vid 修改策略表

- 【命令格式】 **traffic-redirect access-group acl inner-vlan vid out**
- 【参数说明】 acl : 用于匹配流的 acl。
vid : 修改后的报文内层 vid。
- 【命令模式】 接口模式
- 【使用指导】 -

配置基于 ACL 的内层 vid 修改策略表

- 【命令格式】 **traffic-redirect access-group acl outer-vlan vid in**
- 【参数说明】 acl : 用于匹配流的 acl。
vid : 修改后的外层 vid。
- 【命令模式】 接口模式
- 【使用指导】 -

配置举例

配置基于外层 Tag 修改外层 Tag 的 VID 策略

- 【配置方法】
- 根据实际组网需求在接口上配置内层 Tag 和外层 Tag 修改策略。
 - 这里仅针对基于 Tag 和基于流各列举一中 VID 的修改策略，其它策略配置类就不一一列举了。具体言之和功能作用详见上。

第一种：基于外 VLAN Tag 修改外 VLAN Tag。

```
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if)# switchport mode trunk
Ruijie(config-if)# dot1q relay-vid 100 translate local-vid 10-20
```

第二种：基于流的外层 VLAN tag 修改策略。

```
Ruijie# configure terminal
Ruijie(config)# ip access-list standard 2
Ruijie(config-acl-std)# permit host 1.1.1.1
Ruijie(config-acl-std)# exit
Ruijie(config)# interface gigabitEthernet 0/2
Ruijie(config-if)# switchport mode trunk
Ruijie(config-if)# traffic-redirect access-group 2 outer-vlan 3 in
```

- 【检验方法】
- 查看接口上配置是否生效。
 - 报文的 Tag 是否是依据配置策略进行修改。

常见配置错误

- 无

11.5.7 配置优先级映射和优先级复制

配置效果

- 如果服务商网络基于用户的 VLAN Tag 的 User Priority 值设置了 QOS 策略，通过配置优先级复制，可以使外层 Tag 具有与内层 Tag 一样的优先级策略。
- 如果服务商网络基于用户的 VLAN Tag 的 User Priority 值设置了 QOS 策略，通过配置优先级映射，可以设置外层 Tag 为服务商提供的 User Priority 值。

注意事项

 只有 dot1q-tunnel 端口允许配置用户 Tag 的优先级复制，其优先级高于信任模式的 QOS，低于基于数据流的 QOS。

 优先级复制与优先级映射功能不能在同一接口上同时打开。

- ⚠️ 只有 dot1q-tunnel 端口允许配置用户 Tag 的优先级映射，其优先于 QOS 生效。
- ⚠️ 如果没有配置信任模式，即 trust none，则优先级映射的配置不生效；如果配置的信任模式和配置的映射不匹配，也不生效。

配置方法

- 两种配置都必须依赖 Tunnel 口。
- 如果需要利用服务商网络为用户 VLAN Tag 提供的 QOS 策略，则优先级复制功能必须配置。
- 如果需要根据用户的 VLAN Tag 设置外层 VLAN Tag 的 User Priority，灵活应用 QOS 优先级策略，则必须配置优先级映射策略。

📘 在未配置优先级映射的情况下，使用以下优先级映射：

inner pri	0	1	2	3	4	5	6	7

outer pri	0	1	2	3	4	5	6	7

检验方法

- 查看端口上优先级映射或优先级复制功能是否生效。
- show inner-priority-trust interfaces *type intf-id* 和 show interfaces *type intf-id* remark

相关命令

配置优先级映射

- 【命令格式】 **dot1q-Tunnel cos inner-cos-value remark-cos outer-cos-value**
- 【参数说明】 *inner-cos-value*：内部 Tag 的 cos 值。
outer-cos-value：外部 Tag 的 cos 值。
- 【命令模式】 接口配置模式
- 【使用指导】 -

配置优先级复制

- 【命令格式】 **inner-priority-trust enable**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 -

配置举例

配置优先级映射、优先级复制。

- 【配置方法】
- 为了能够维持报文的优先级，需要在 Tunnel 口上，将用户报文的内层 Tag 的优先级复制到外层 Tag 中。
 - 为了能够在 Tunnel 口上实现对报文优先级的灵活控制，可以根据报文的内层优先级的不同，为报文封装 Tag 时候标记上不同优先级的外层标签。

配置优先级映射

```
Ruijie(config)# interface gigabitethernet 0/1
Ruijie(config-if)# mls qos trust cos
Ruijie(config-if)# inner-priority-trust enable
Ruijie(config)# end
```

配置优先级复制

```
Ruijie(config)# interface gigabitethernet 0/2
Ruijie(config-if)# dot1q-tunnel cos 3 remark-cos 5
```

- 【检验方法】
- 查看端口上的优先级配置

查看 Tunnel 口上的优先级复制关系：

查看 Tunnel 口上优先级映射关系：

```
Ruijie# show interface gigabitethernet 0/1 remark
Ports          Type          From value  To value
```

查看 Tunnel 口上优先级复制开关是否打开：

```
Ruijie# show inner-priority-trust interfaces gigabitethernet 0/1
Port          inner-priority-trust
-----
Gi0/1         enable
```

查看 Tunnel 口上优先级映射关系：

```
Ruijie# show interfaces gigabitethernet 0/1 remark
Ports          Type          From value  To value
-----
Gi0/1          Cos-To-Cos   3           5
```

常见配置错误

- 见注意事项。

11.5.8 配置二层协议透传

配置效果

- 实现二层协议透传，保证用户网络的拓扑并且对服务商网络不产生影响。

注意事项

- ⚠ 未启 STP 协议时，还需再配置 bridge-frame forwarding protocol bpdu 才能透传 STP 协议报文。
- ⚠ 当全局使能协议透传后，接口上使能协议透传时候才有效。当接口上协议透传功能生效时候，该接口不参与该协议计算。若透传口收到目的 MAC 为特殊组播地址的报文，则表示组网出现错误，将直接丢弃该报文。

配置方法

STP 协议报文透传设置

- 需要透传 STP 协议的 BPDU 报文，必须配置。
- 必须全局和接口下均开启 STP 协议透传功能。

GVRP 协议报文透传设置

- 需要透传 GVRP 协议报文，必须配置。
- 必须全局和接口下均开启 GVRP 协议透传功能。

配置透传地址

- 可选配置。
- 配置对应协议的透传地址。

检验方法

查看透传相关配置是否正确：

- `show l2protocol-tunnel stp`、`show l2protocol-tunnel gvrp` 查看配置信息是否正确。

相关命令

接口上开启 STP 协议报文透传

- 【命令格式】 `l2protocol-tunnel stp`
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

接口上开启 STP 协议报文透传

- 【命令格式】 `l2protocol-tunnel stp enable`
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 -

全局开启 STP 协议报文透传功能

【命令格式】 `l2protocol-tunnel gvrp`

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

接口上开启 GVRP 协议报文透传功能

【命令格式】 `l2protocol-tunnel gvrp enable`

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 -

配置透传地址

【命令格式】 `l2protocol-tunnel { stp | gvrp } tunnel-dmac mac-address`

【参数说明】 *mac-address* : 设置的协议报文的透传地址。

【命令模式】 接口配置模式

【使用指导】

① 其中 STP 报文可选地址范围：01d0.f800.0005、011a.a900.0005、010f.e200.0003、0100.0ccd.cdd0、0100.0ccd.cdd1、0100.0ccd.cdd2；其中 GVRP 报文可选地址范围为：01d0.f800.0006、011a.a900.0006。

② 当未配置透传地址时，缺省使用的地址前三字节为 01d0f8，后 3 字节为 (stp: 000005, gvrp: 000006)。

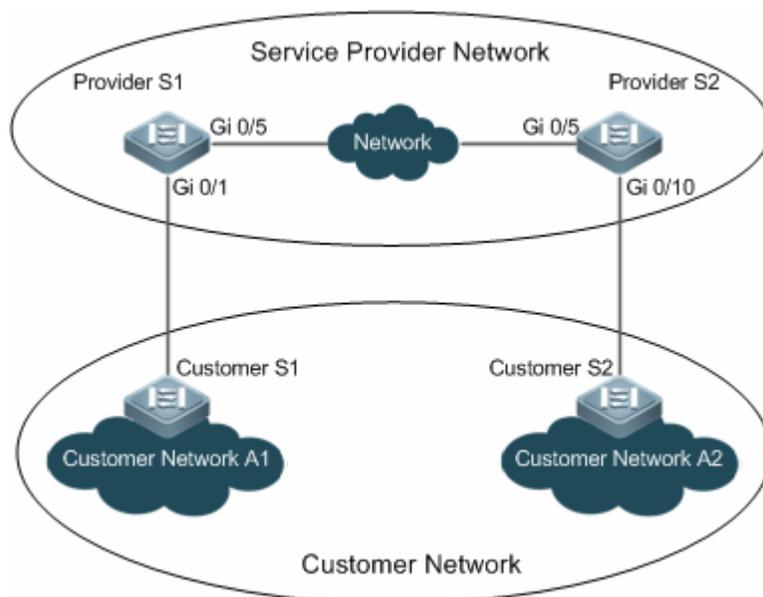
配置举例

这里仅举例配置 STP 协议报文透传配置，GVRP 协议透传配置类似就不再赘述。

透传 STP 协议的 BPDU 报文

【网络环境】

图 11-10



- 【配置方法】**
- 在服务器边缘设备上全局和接口下均开启 STP 协议透传功能。
 - STP 协议透传功能必须以设备能转发 STP 协议报文为前提，所以必须全局开启 STP 协议转发功能。
 -

Provider S1

第一步，开启 STP 协议转发功能。

```
bridge-frame forwarding protocol bpdu
```

第二步，创建用来透传协议报文的 VLAN

```
ProviderS1#configure terminal
```

Enter configuration commands, one per line. End with CNTL/Z.

```
ProviderS1(config)#vlan 200
```

```
ProviderS1(config-vlan)#exit
```

第三步，在连接用户网络的接口上开启基本 QinQ 功能，使用 VLAN200 对用户网络的数据进行隧道传输

```
ProviderS1(config)#interface gigabitEthernet 0/1
```

```
ProviderS1(config-if-GigabitEthernet 0/1)#switchport mode dot1q-tunnel
```

```
ProviderS1(config-if-GigabitEthernet 0/1)#switchport dot1q-tunnel native vlan 200
```

第四步，在连接用户网络的接口上开启 STP 协议透传功能

```
ProviderS1(config-if-GigabitEthernet 0/1)#l2protocol-tunnel stp enable
```

```
ProviderS1(config-if-GigabitEthernet 0/1)#exit
```

第五步，全局开启 STP 协议透传功能

```
ProviderS1(config)#l2protocol-tunnel stp
```

第六步，配置 uplink port

```
ProviderS1(config)# interface gigabitEthernet 0/5
```

```
ProviderS1(config-if-GigabitEthernet 0/5)#switchport mode uplink
```

Provider S2

Provider S2 设备上的配置验证同 Provider S1 的类似，请参考上文 Provider S1 的配置验证过程，此处不再重复说明。

【检验方法】

第一步，验证 STP 协议透传功能是否全局使能并在接口上开启。

```
ProviderS1#show l2protocol-tunnel stp
```

```
L2protocol-tunnel: Stp Enable
```

```
GigabitEthernet 0/1 l2protocol-tunnel stp enable
```

第二步，确认 Tunnel 口的配置是否正确，关注点：接口类型是否为 dot1q-tunnel，外层 Tag VLAN 是否为 Native VLAN 且其是否已加入接口的许可 VLAN 列表，服务商网络边缘设备上链口的类型是否为 Uplink。

```
ProviderS1#show running-config
```

```
interface GigabitEthernet 0/1
```

```
switchport mode dot1q-tunnel
```

```
switchport dot1q-tunnel allowed vlan add untagged 200
```

```
switchport dot1q-tunnel native vlan 200
```

```
l2protocol-tunnel stp enable
```

```
spanning-tree bpdufilter enable
```

```
!
```

```
interface GigabitEthernet 0/5
```

```
switchport mode uplink
```

常见错误

- 需要透传 STP 协议报文时，未全局开启 STP 协议转发功能，导致 STP 协议无法透传。
- 透传协议报文时，没有在全局和接口上开启全局透传使能，导致无法透传。

11.6 监视与维护

清除各类信息

无

查看运行情况

作用	命令
显示接口的 dot1q-tunnel 是否打开	<code>show dot1q-tunnel [interfaces <i>intf-id</i>]</code>
显示 dot1q-tunnel 口配置	<code>show interfaces dot1q-tunnel</code>
显示基于协议的 dot1q-tunnel 端口 vid 添加策略表	<code>show registration-table [interfaces <i>intf-id</i>]</code>
显示基于协议 access,trunk,hybrid 端口 vid 修改策略表	<code>show translation-table [interfaces <i>intf-id</i>]</code>
显示基于数据流的 vid 变更或添加策略表	<code>show traffic-redirect [interfaces <i>intf-id</i>]</code>
显示接口 tpid 的配置	<code>show frame-tag tpid interfaces [<i>intf-id</i>]</code>
显示优先级复制配置	<code>show inner-priority-trust</code>
显示优先级映射的配置	<code>show interface <i>intf-name</i> remark</code>
显示 MAC 地址复制配置	<code>show mac-address-mapping</code>
显示二层透传配置	<code>show l2protocol-tunnel { gvrp stp }</code>

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 QINQ 的调试开关。	<code>debug bridge qinq</code>

12 MGMT

12.1 概述

在我司产品面板上的以太网接口，由于内部构成的限制，这些接口实际上与设备内部的转发面部件是分离的，不参与设备转发面及控制面的功能，相应的这些接口上的通信与设备上运行的业务通信是分开的，称为“带外通信”。利用这个接口可以对设备进行管理，这种方式与从 Console 接口上登录的管理方式类似。管理用以太网接口只用于管理设备，不支持通信转发功能，习惯上称为 MGMT (Management) 接口。

采用 MGMT 接口，用户可以将用于管理的网络与业务网络分离，这样管理就不会受到业务网络流量及通信状态的干扰，从而提高管理的可靠性，特别是当业务网络出现故障时，仍可以通过管理网络对设备进行管理，这种优点是业务网络的带内管理方式无法比拟的。

另外，与 Console 接口相比，MGMT 接口带宽较高(如 100M vs 115200bps)，并且在具备日志服务器管理网络中，通过 MGMT 接口可以将日志发往日志服务器，这样日志的发送与存储也不受业务网络通信状况的影响。

i 由于硬件构成的差异，MGMT 接口可能是百兆以太网接口或千兆以太网接口。

i 下文仅介绍管理用以太网接口配置的相关内容。

协议规范

- 无

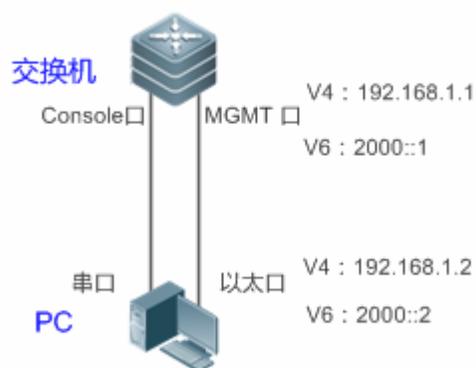
12.2 典型应用

典型应用	场景描述
网络管理工具	通过 MGMT 口进行网络通信的管理与调试。
文件管理	通过 MGMT 口进行管理网络与设备之间的文件复制。
网络登陆管理	通过 MGMT 口实现本设备远程登陆到其它设备或主机上。
MIB管理	通过 MGMT 口向 NMS 服务器发送 SNMP trap 信息。
LOG管理	通过 MGMT 口向 SYSLOG 服务器发送 LOG 信息。

12.2.1 网络管理工具

应用场景

图 12-1网络管理工具



上图中，通过 PC 的串口连接交换机的 Console 口对 MGMT 接口配置三层接口属性及二层接口属性；探测 MGMT 口可达的主机及对这些可达的主机进行路径跟踪。

【注释】 -

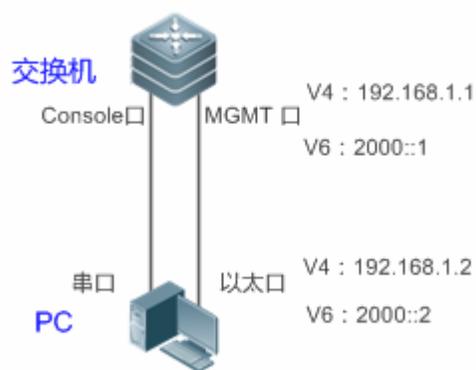
功能部属

- PC 的串口和交换机的 Console 口进行互联。
- PC 的以太网口和交换机的 MGMT 口进行互联。
- 通过 PC 的串口配置交换机的 MGMT 口。
- 通过 PC 的串口下达探测 MGMT 可达的主机的指令。
- 通过 PC 的串口下达跟踪 MGMT 可达主机的路径的指令。

12.2.2 文件管理

应用场景

图 12-2文件管理



上图中，通过 PC 的串口连接交换机的 Console 口对 MGMT 接口配置三层接口属性及二层接口属性；交换机经由 MGMT 口从文件服务器复制文件。

【注释】 -

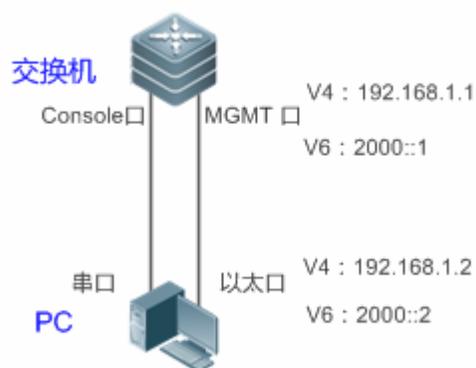
功能部属

- PC 的串口和交换机的 Console 口进行互联。
- PC 的以太网口和交换机的 MGMT 口进行互联。
- 通过 PC 的串口配置交换机的 MGMT 口。
- PC 开启文件服务器功能。
- 通过 PC 的串口下达交换机经由 MGMT 口从文件服务器复制文件的指令。

12.2.3 网络登陆管理

应用场景

图 12-3 网络登录管理



上图中，通过 PC 的串口连接交换机的 Console 口对 MGMT 接口配置三层接口属性及二层接口属性；交换机经由 MGMT 口登录到 PC 的 Telnet 服务器。

【注释】 -

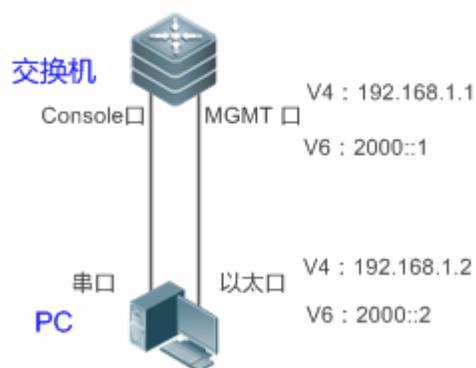
功能部属

- PC 的串口和交换机的 Console 口进行互联。
- PC 的以太网口和交换机的 MGMT 口进行互联。
- 通过 PC 的串口配置交换机的 MGMT 口。
- PC 开启 Telnet 服务器功能。
- 通过 PC 的串口下达交换机经由 MGMT 口登录到 PC 的 Telnet 服务器的指令。

12.2.4 MIB管理

应用场景

图 12-4 MIB 管理



上图中，通过 PC 的串口连接交换机的 Console 口对 MGMT 接口配置三层接口属性及二层接口属性；配置交换机经由 MGMT 口向 PC 的 NMS 服务器发送 SNMP trap 信息。

【注释】 -

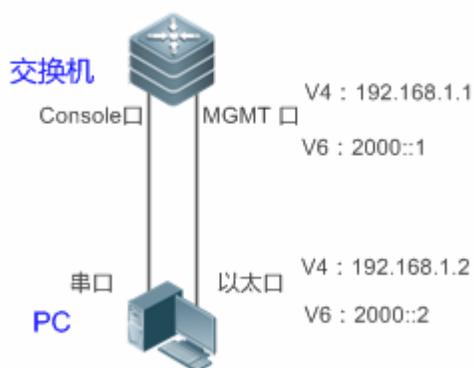
功能部属

- PC 的串口和交换机的 Console 口进行互联。
- PC 的以太网口和交换机的 MGMT 口进行互联。
- 通过 PC 的串口配置交换机的 MGMT 口。
- PC 开启 NMS 服务器功能。
- 通过 PC 的串口下达交换机经由 MGMT 口向 PC 的 NMS 服务器发送 SNMP trap 信息的指令。

12.2.5 LOG管理

应用场景

图 12-5 LOG 管理



上图中，通过 PC 的串口连接交换机的 Console 口对 MGMT 接口配置三层接口属性及二层接口属性；配置交换机经由 MGMT 口向 PC 的 SYSLOG 服务器发送 LOG 信息。

【注释】 -

功能部属

- PC 的串口和交换机的 Console 口进行互联。
- PC 的以太网口和交换机的 MGMT 口进行互联。
- 通过 PC 的串口配置交换机的 MGMT 口。
- PC 开启 SYSLOG 服务器功能。
- 通过 PC 的串口下达交换机经由 MGMT 口向 PC 的 SYSLOG 服务器发送 LOG 信息的指令。

12.3 功能详解

基本概念

-

功能特性

功能特性	作用
接口属性管理	从网络通信的角度来看，MGMT 接口与其它 LAN 接口没有本质区别，只不过由于它不支持通信转发功能，因此它的可配置功能项要比其它 LAN 接口要少一些，对于一些命令需要专门指明它们用于带外通信。
网络管理工具	为方便管理与调试网络的通信，系统提供了一些通过 MGMT 口管理网络的命令工具。
文件管理	系统提供了通过 MGMT 口实现管理网络与设备之间的文件复制功能。
网络登录管理	系统提供了通过本设备的 MGMT 口远程登录到其它设备或主机上的功能。

MIB管理	为方便 MIB 管理，系统提供了通过 MGMT 口能够向 NMS 服务器发送 SNMP trap 信息的功能。
LOG管理	为方便 LOG 管理，系统提供了通过 MGMT 口能够向 SYSLOG 服务器发送 LOG 信息的功能。

12.3.1 接口属性管理

工作原理

从网络通信的角度来看，MGMT 接口与其它 LAN 接口没有本质区别，因此，通过配置 MGMT 口的接口属性，可以实现 MGMT 口同普通 LAN 接口类似的网络通信功能。需要注意的是，MGMT 接口不支持通信转发功能。

相关配置

配置 MGMT 接口的 IPv4 地址

缺省情况下，MGMT 口上没有 IPv4 地址。用户在 MGMT 口的接口模式下，通过下面命令配置 MGMT 口上的 IPv4 地址。

```
ip address address mask
```

其中，*address* 为 IPv4 地址，*mask* 为 IPv4 地址掩码。

配置 MGMT 接口的 IPv4 网关

缺省情况下，MGMT 口上没有 IPv4 网关配置。用户在 MGMT 口的接口模式下，通过下面命令配置 MGMT 口的 IPv4 网关地址。

```
gateway A.B.C.D
```

其中，*A.B.C.D* 为 IPv4 网关地址。

配置 MGMT 接口的 MTU

缺省情况下，MGMT 口的 MTU 值为 1500，通过下面命令配置 MGMT 口上的 MTU。

```
MTU mtu-value
```

其中，*mtu-value* 为 MTU 值，可配置的范围为 64-设备支持的最大 MTU 值。

配置 MGMT 接口的速率

缺省情况下，MGMT 口的速率模式为 auto，通过下面命令配置 MGMT 口上的速率模式。

```
speed {10 | 100 | 1000 | auto}
```

配置 MGMT 接口的双工

缺省情况下，MGMT 口的双工模式为 auto，通过下面命令配置 MGMT 口上的双工模式。

```
duplex {full | half | auto}
```

配置 MGMT 接口的描述符

缺省情况下，MGMT 口没有接口描述符配置，通过下面命令配置 MGMT 口的接口描述符。

```
description text
```

其中，*text* 为接口描述符。

📌 关闭 MGMT 接口

缺省情况下，MGMT 口是打开的，通过下面命令配置关闭 MGMT 口。

```
shutdown
```

12.3.2 网络管理工具

为方便管理与调试网络的通信，系统提供了一些通过 MGMT 口实现管理网络的命令工具。

工作原理

- 通过 MGMT 接口使用 ping 报文检测管理网络上设备节点/主机的 IPv4 地址的可达性。
- 通过 MGMT 接口使用 traceroute 报文检测管理网络上设备节点/主机的 IPv4 路由。

相关配置

📌 探测 IPv4 地址可达性

特权模式下，下面命令用于通过 MGMT 口探测设备节点/主机的 IPv4 地址可达性。

```
ping oob address via mgmt-name
```

其中，*address* 为被探测的设备节点/主机 IPv4 地址，*mgmt-name* 为 oob 模式下报文的出口管理口。

📌 跟踪 IPv4 路由

特权模式下，下面命令用于通过 MGMT 口跟踪设备节点/主机的 IPv4 路由。

```
traceroute oob address via mgmt-name
```

其中，*address* 为被跟踪的设备节点/主机 IPv4 地址，*mgmt-name* 为 oob 模式下报文的出口管理口。

12.3.3 文件管理

系统提供了通过 MGMT 口实现管理网络与设备之间的文件复制功能。

工作原理

通过 MGMT 口将制定的文件从源 URL 处复制到目的 URL 处。

相关配置

↘ 文件复制

特权模式下，下面命令用于通过 MGMT 口将制定的文件从源 URL 处复制到目的 URL 处。

↘ `copy oob_ftp://source-url destination-url`

其中，source-url 为文件源地址，destination-url 为文件目的地址。

12.3.4 网络登录管理

系统提供了用于通过本设备 MGMT 口远程登陆到其它设备或主机上的功能。

工作原理

通过 MGMT 口登录到指定的设备节点/主机上实现远程操控这个设备节点。

相关配置

↘ 网络登录管理

特权模式下，下面命令用于通过 MGMT 口登录到指定的设备节点/主机上。

```
telnet oob ip-address
```

其中，ip-address 为设备节点/主机的 IPv4 地址，ip-address 为设备节点。

12.3.5 MIB管理

为方便 MIB 管理，系统提供了专用于 SNMP 指定 MGMT 口向 NMS 服务器发送 trap 信息的功能。

工作原理

通过 MGMT 口和 NSM 服务器的 IPv4 地址向 NSM 服务器发送 SNMP trap 信息。

相关配置

↘ 向 NMS 服务器的 IPv4 地址发送 trap 信息

全局配置模式下，下面命令用于通过 MGMT 口向 NMS 服务器的 IPv4 地址发送 trap 信息。缺省该功能是关闭的。

```
snmp-server host oob ip-address
```

其中，ip-address 为 NMS 服务器的 IPv4 地址。

12.3.6 LOG管理

为方便 LOG 管理，系统提供了专用于指定 MGMT 口向 SYSLOG 服务器其发送 LOG 信息的功能。

工作原理

通过 MGMT 口和 SYSLOG 服务器的 IPv4 地址向 SYSLOG 服务器发送 LOG 信息。

相关配置

通过 MGMT 口和 SYSLOG 服务器的 IPv4 地址发送 LOG 信息

全局配置模式下，下面命令用于通过 MGMT 口和 SYSLOG 服务器的 IPv4 地址向 SYSLOG 服务器发送 LOG 信息。缺省该功能关闭的。

```
logging server oob ip-address
```

其中，*ip-address* 为 SYSLOG 服务器的 IPv4 地址。

12.4 配置详解

配置项	配置建议 & 相关命令	
配置MGMT接口属性	 MGMT 口下的 IPv4 地址为必须配置，通过配置 IPv4 地址，可以实现通过 MGMT 口来管理这台设备。	
	<code>ip address address mask</code>	配置接口的 IPv4 地址与子网掩码
	<code>gateway A.B.C.D</code>	配置 IPv4 管理网络的网关
	 可选配置，通过配置管理使其能够根据网络部署的需要调整 MGMT 口工作在最佳的工作状态下。	
	<code>mtu mtu-value</code>	接口的最大传输单元 (MTU)
	<code>speed { 10 100 1000 auto }</code>	设置接口的 speed，缺省值是 auto
	<code>duplex { full half auto }</code>	设置接口的 duplex，缺省值是 auto
	<code>shutdown</code>	关闭 MGMT 接口
	<code>description text</code>	配置描述符
网络管理工具	 可选配置。通过这些配置可以实现通过 MGMT 口进行网络管理，例如进行 ping 操作、跟踪网络路由等，来检测网络主机的可达性和路由信息。	
	<code>ping oob address</code>	使用 ICMP echo request 检测管理网络上主机的可达性
	<code>tracert oob address</code>	检测到管理网络内主机的路由
文件管理	 可选配置。通过这些配置实现通过 MGMT 口进行管理网络与设备之间的文件复制功能。	

	copy oob_tftp://source-url destination-url	将文件由 source-url 指定的位置复制到 destination-url 指定的位置
网络登陆管理	 可选配置。通过这些配置实现通过 MGMT 口远程登陆到其它设备或主机上。	
	telnet oob ip-address	使用该命令，指定在设备上执行 telnet 命令时，通过 MGMT 口进行数据交互，配置中不需要指定 ip 等参数来指定是使用 ipv4 协议，系统会自行判断输入的地址是一个合法的 v4 地址。
MIB管理	 可选配置。通过这些配置实现通过 MGMT 口向 NMS 服务器发送 SNMP trap 信息功能。	
	snmp-server host oob ip-address	使用该命令配置 snmp agent 指定通过 MGMT 口向 NMS 服务器的 ipv4 地址发送 trap 信息。
LOG管理	 可选配置。通过这些配置实现通过 MGMT 口向 SYSLOG 服务器发送 LOG 信息功能。	
	logging server oob ip-address	使用该命令配置 syslog 指定通过 MGMT 口向 syslog server 服务器的 ipv4 地址发送 log 信息。

12.4.1 接口属性管理

配置效果

- 配置 MGMT 接口三层地址。
- 配置管理网络的网关地址。
- 配置 MGMT 接口的物理属性。
- 配置完毕后，MGMT 能够用于设备管理。

注意事项

- MGMT 不支持通信转发功能。

配置方法

▾ 配置 MGMT 接口三层地址

- 进入 MGMT 接口配置模式。
- 配置 MGMT 接口三层地址。

▾ 配置管理网络的网关地址

- 进入 MGMT 接口配置模式。
- 配置管理网络的网关地址。

检验方法

- 通过 **show running** 命令查看相应的配置。

相关命令

配置 MGMT 接口 IPv4 地址

- 【命令格式】 **ip address address mask**
- 【参数说明】 *address* : IPv4 地址
mask : IPv4 地址掩码
- 【命令模式】 MGMT 接口配置模式
- 【使用指导】 -

配置管理网络的 IPv4 网关

- 【命令格式】 **gateway A.B.C.D**
- 【参数说明】 *A.B.C.D* : IPv4 网关地址
- 【命令模式】 MGMT 接口配置模式
- 【使用指导】 -

配置接口的最大传输单元 (MTU)

- 【命令格式】 **mtu mtu-value**
- 【参数说明】 *mtu-value* : 接口 MTU 值
- 【命令模式】 MGMT 接口配置模式
- 【使用指导】 -

配置 MGMT 接口的速率

- 【命令格式】 **speed {10 | 100 | 1000 | auto}**
- 【参数说明】 缺省值是 auto
- 【命令模式】 MGMT 接口配置模式
- 【使用指导】 -

配置 MGMT 接口的双工

- 【命令格式】 **duplex {full | half | auto}**
- 【参数说明】 缺省值是 auto
- 【命令模式】 MGMT 接口配置模式
- 【使用指导】 -

配置 MGMT 接口的描述符

- 【命令格式】 **description** *text*
- 【参数说明】 *text* : 接口描述符, 缺省无
- 【命令模式】 MGMT 接口配置模式
- 【使用指导】 -

关闭 MGMT 接口

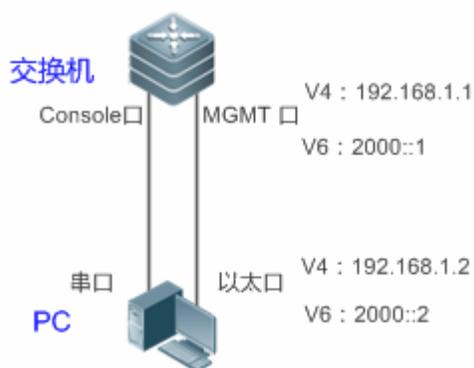
- 【命令格式】 **shutdown**
- 【参数说明】 缺省为 no shutdown
- 【命令模式】 MGMT 接口配置模式
- 【使用指导】 -

配置举例

配置 MGMT 接口

【网络环境】

图 12-6



- 【配置方法】
- PC 的串口跟交换机的 Console 口互联
 - 配置交换机上 MGMT 接口三层 IPv4 地址为 192.168.1.1
 - 配置管理网络的 v4 网关为 192.168.1.2
 - 配置交换机上 MGMT 接口的速率为 1000M
 - 关闭交换机上 MGMT 接口

交换机

```
Ruijie# configure
Ruijie(config)# interface mgmt 0
Ruijie(config-if-Mgmt 0)# ip address 192.168.1.1 255.255.255.0
Ruijie(config-if-Mgmt 0)# gateway 192.168.1.1
Ruijie(config-if-Mgmt 0)# gateway 2000::2
Ruijie(config-if-Mgmt 0)# speed 1000
Ruijie(config-if-Mgmt 0)# shutdown
```

- 【检验方法】
- 通过 show running 查看交换机上述配置。

交换机

```
Ruijie# show run int mgmt 0
Building configuration...
Current configuration : 168 bytes
```

```
!  
interface MGMT 0  
no switchport  
speed 1000  
no ip proxy-arp  
ip address 192.168.1.1 255.255.255.0  
gateway 192.168.1.1  
shutdown
```

常见错误

- -

12.4.2 网络管理工具

配置效果

- 通过 MGMT 口探测设备节点/主机的 IPv4 地址可达性。
- 通过 MGMT 口跟踪设备节点/主机的 IPv4 路由。

注意事项

- -

配置方法

📄 探测 IPv4 地址可达性

- 进入特权模式。
- 通过 MGMT 口探测设备节点/主机的 IPv4 地址可达性。

📄 跟踪 IPv4 路由

- 进入特权模式。
- 通过 MGMT 口跟踪设备节点/主机的 IPv4 路由。

检验方法

- 查看即时过程。

相关命令

探测 IPv4 地址可达性

- 【命令格式】 **ping oob address via mgmt-name**
- 【参数说明】 *address* : IPv4 地址
mgmt-name : 指定在 oob 模式下报文的出口管理口。
- 【命令模式】 特权配置模式
- 【使用指导】 -

跟踪 IPv4 路由

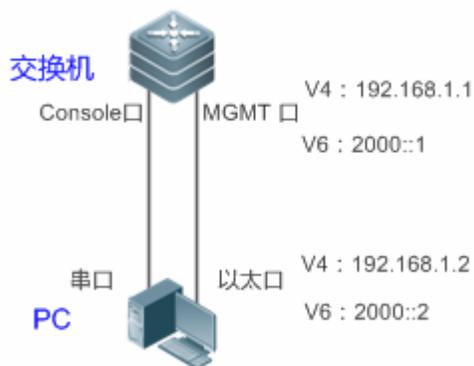
- 【命令格式】 **tracroute oob address via mgmt-name**
- 【参数说明】 *address* : IPv4 地址
mgmt-name : 指定在 oob 模式下报文的出口管理口。
- 【命令模式】 特权配置模式
- 【使用指导】 -

配置举例

网络管理工具

【网络环境】

图 12-7



- 【配置方法】
- PC 的串口跟交换机的 Console 口互联
 - 配置交换机上 MGMT 接口三层 IPv4 地址为 192.168.1.1
 - PC 的以太网口跟交换机的 MGMT 口互联

交换机

```
Ruijie# configure
Ruijie(config)# int mgmt 0
Ruijie(config-if-Mgmt 0)# ip address 192.168.1.1 255.255.255.0
Ruijie# ping oob 192.168.1.2
Ruijie# tracroute oob 192.168.1.2
```

- 【检验方法】
- 查看即时过程，可以看到可以 ping 通管理网络内的主机和 tracroute 到管理网络内主机的路由。

交换机

```
Ruijie# ping oob 192.168.1.2
Sending 5, 100-byte ICMP Echoes to 192.168.1.2, timeout is 2 seconds:
!!!!
Success rate is 100 percent, round-trip min/avg/max = 4/4/4 ms
Ruijie# traceroute oob 192.168.1.2
Tracing route to 192.168.1.2 over a maximum of 10 hops
 1  <10 ms  <10 ms  <10 ms  192.168.1.2
```

常见错误

- -

12.4.3 文件管理

配置效果

- 通过 MGMT 口将文件由源 URL 指定的位置复制到目的 URL 指定的位置。

注意事项

- -

配置方法

📄 文件管理

- 进入特权模式。
- 将文件由源 URL 指定的位置复制到目的 URL 指定的位置。

检验方法

- 查看即时过程。

相关命令

📄 文件管理

【命令格式】 **copy oob_tftp://source-url destination-url**

【参数说明】 *source-url* : 文件源地址

destination-url : 文件目的地址

【命令模式】 特权模式

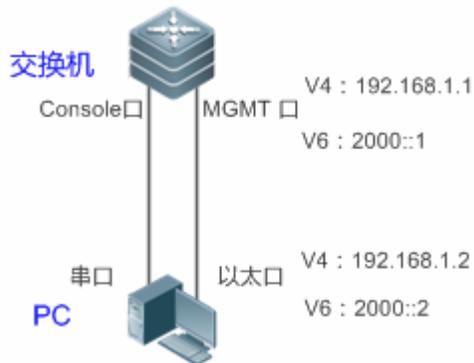
【使用指导】 -

配置举例

文件管理

【网络环境】

图 12-8



- 【配置方法】
- PC的串口跟交换机的Console口互联
 - 配置交换机上MGMT接口三层IPv4地址为192.168.1.1
 - PC的以太网口跟交换机的MGMT口互联
 - PC的以太网口配置IPv4地址192.168.1.2
 - PC基于IPv4开启tftp服务器
 - 从管理网络内的IPv4主机下载一个文件到flash的文件系统中

交换机

```
Ruijie# configure
Ruijie(config)# int mgmt 0
Ruijie(config-if-Mgmt 0)# ip address 192.168.1.1 255.255.255.0
Ruijie# copy oob_tftp://192.168.1.2/ngsa-compress.bin
Ruijie# copy oob_tftp://[2000::2]/ngsa-compress.bin
```

- 【检验方法】
- 查看即时过程，即从管理网络内的IPv4主机下载了一个文件到flash的文件系统中。

交换机

```
Ruijie# copy oob_tftp://192.168.1.2/ngsa-compress.bin
flash:file.bin
Accessing tftp://192.168.1.2/ngsa-compress.bin...
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Success : Transmission success, file length 1183856 bytes
Ruijie# copy oob_tftp://[2000::2]/ngsa-compress.bin
flash:file.bin
Accessing tftp://192.168.1.2/ngsa-compress.bin...
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Success : Transmission success, file length 1183856 bytes
```

常见错误

- -

12.4.4 网络登录管理

配置效果

- 通过 MGMT 口登录到其它的设备或主机。

注意事项

- -

配置方法

▾ 网络登录管理

- 进入特权模式。
- 通过 MGMT 口登录到其它的设备或主机。

检验方法

- 查看即时过程。

相关命令

▾ 网络登录管理

【命令格式】 **telnet oob ip-address**

【参数说明】 *ip-address* : IPv4 地址

【命令模式】 特权模式

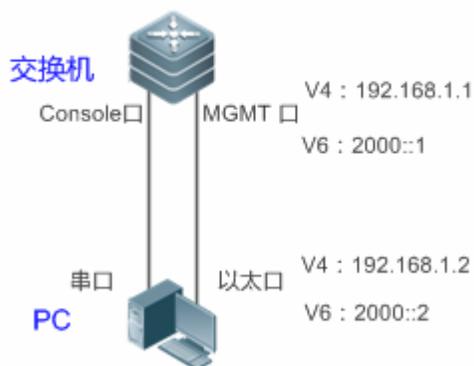
【使用指导】 使用该命令，指定在设备上执行 telnet 命令时，通过 MGMT 口进行数据交互，配置中不需要指定 ip 等参数来指定是使用 IPv4 协议，系统会自行判断输入的地址是一个合法的 IPv4 地址。

配置举例

▾ 网络登录管理

【网络环境】

图 12-9



【配置方法】

- PC 的串口跟交换机的 Console 口互联
- 配置交换机上 MGMT 接口三层 IPv4 地址为 192.168.1.1
- PC 的以太网口跟交换机的 MGMT 口互联
- PC 的以太网口配置 IPv4 地址 192.168.1.2
- PC 基于 IPv4 开启 telnet 服务器
- 交换机 A 通过 MGMT 口登录到 PC

交换机

```
Ruijie A# configure
Ruijie A(config)# int mgmt 0
Ruijie A(config-if-Mgmt 0)# ip address 192.168.1.1 255.255.255.0
Ruijie A# telnet oob 192.168.1.2
Ruijie A# telnet oob 2000::2
```

【检验方法】

- 查看即时过程，即交换机可以登录到 PC 上。

交换机 A

```
Ruijie A# telnet oob 192.168.1.2
User Access Verification
Password:
Ruijie A# telnet oob 2000::2
User Access Verification
Password:
```

常见错误

- -

12.4.5 MIB管理

配置效果

- 指定通过 MGMT 口和 NMS 服务器的 IPv4 地址向 NSM 服务器发送 trap 信息。

注意事项

- -

配置方法

▾ MIB 管理

- 进入全局模式。
- 配置指定通过 MGMT 口和 NMS 服务器的 IPv4 地址发送 trap 信息。

检验方法

- 通过 `show running` 命令查看。

相关命令

▾ 指定通过 MGMT 口和 NMS 服务器的 IPv4 地址发送 trap 信息

【命令格式】 `snmp-server host oob ip-address`

【参数说明】 `ip-address` : IPv4 地址

【命令模式】 全局配置模式

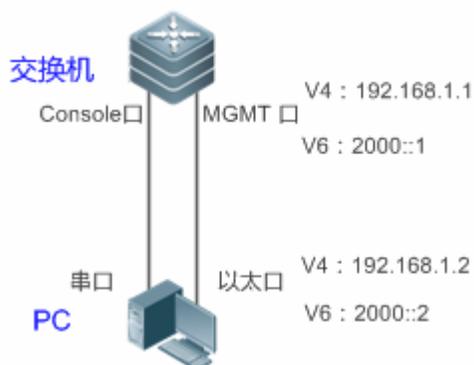
【使用指导】 -

配置举例

▾ 配置 MGMT 接口 MIB 管理功能

【网络环境】

图 12-10



- 【配置方法】
- PC 的串口跟交换机的 Console 口互联
 - 配置交换机 MGMT 接口三层 IPv4 地址为 192.168.1.1

- PC 的以太网口跟交换机的 MGMT 口互联
- PC 的以太网口配置 IPv4 地址 192.168.1.2
- PC 基于 IPv4 开启 NMS 服务器
- 指定通过 MGMT 口和 NMS 服务器的 IPv4 地址发送 trap 信息
- 指定通过 MGMT 口和 NMS 服务器的 IPv4 地址发送 trap 信息

交换机

```
Ruijie# configure
Ruijie(config)# int mgmt 0
Ruijie(config-if-Mgmt 0)# ip address 192.168.1.1 255.255.255.0
Ruijie(config)# snmp-server host oob 192.168.1.2
```

【检验方法】

- 通过 show running 查看交换机上述配置。

交换机

```
Ruijie# show running | include snmp-server
snmp-server host oob 192.168.1.2
```

常见错误

- -

12.4.6 LOG管理

配置效果

- 指定通过 MGMT 口和 SYSLOG 服务器的 IPv4 地址发送 LOG 信息。

注意事项

- -

配置方法

📄 LOG 管理

- 进入全局模式。
- 配置指定通过 MGMT 口和 SYSLOG 服务器的 IPv4 地址发送 LOG 信息。

检验方法

- 通过 show running 命令查看。

相关命令

配置指定通过 MGMT 口和 SYSLOG 服务器的 IPv4 地址发送 LOG 信息

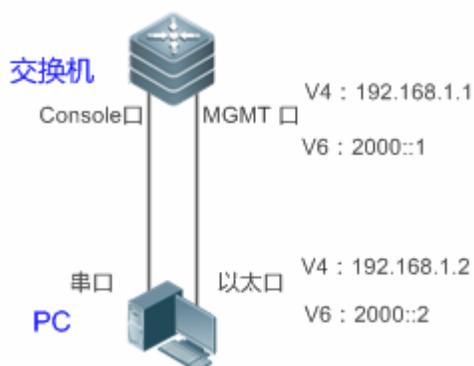
- 【命令格式】 **logging server oob** *ip-address*
- 【参数说明】 *ip-address* : IPv4 地址
- 【命令模式】 全局配置模式
- 【使用指导】 -

配置举例

配置 MGMT 接口 LOG 管理功能

【网络环境】

图 12-11



- 【配置方法】
- PC 的串口跟交换机的 Console 口互联
 - 配置交换机 MGMT 接口三层 IPv4 地址为 192.168.1.1
 - PC 的以太网口跟交换机的 MGMT 口互联
 - PC 基于 IPv4 开启 SYSLOG 服务器

交换机

```
Ruijie# configure
Ruijie(config)# int mgmt 0
Ruijie(config-if-Mgmt 0)# ip address 192.168.1.1 255.255.255.0
Ruijie(config)# logging server oob 192.168.1.2
```

- 【检验方法】
- 通过 show running 查看上述配置

交换机

```
Ruijie# show running | include logging
logging server oob 192.168.1.2
```

常见错误

- -

12.5 监视与维护

清除各类信息

-

查看运行情况

作用	命令
查看虚拟 mgmt 口的成员状态和统计信息。	show mgmt virtual

查看调试信息

-



配置指南-IP 地址及应用

本分册介绍 IP 地址及应用配置指南相关内容，包括以下章节：

1. IP 地址与服务
2. ARP
3. DHCP
4. DNS
5. FTP Server
6. FTP Client
7. 网络通信检测工具
8. TCP
9. 软件 IPv4 快转

1 IP 地址与服务

1.1 概述

因特网协议（Internet Protocol，IP）使用逻辑虚拟的地址将数据包从源方发送到目的方，即 IP 地址。在网络层，路由设备使用 IP 地址完成数据包转发。

i 以下仅针对 IPv4 地址进行介绍。

协议规范

- RFC 1918 : Address Allocation for Private Internets
- RFC 1166 : Internet Numbers

1.2 典型应用

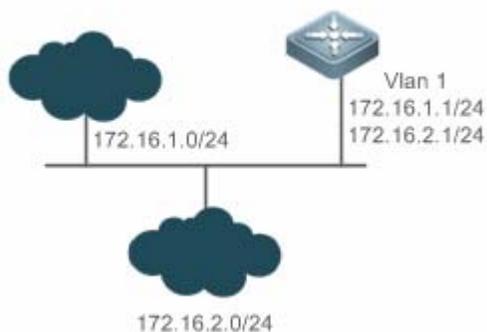
典型应用	场景描述
配置IP地址通信	两个网络使用同一个交换机接口进行通信

1.2.1 配置IP地址通信

应用场景

交换机连接一个局域网，局域网分为两个网段：172.16.1.0/24 和 172.16.2.0/24。要求两个网段的计算机都可以通过交换机和因特网通信，并且两个网段的计算机之间可以互相通信。

图 1-1 IP 地址配置范例



功能部属

- 在 vlan1 口上配置两个 ip 地址，一个主 ip 地址，一个从 ip 地址。
- 在 172.16.1.0/24 网段中的主机上配置网关为 172.16.1.1，在 172.16.2.0/24 网段中的主机上配置网关为 172.16.2.1。

1.3 功能详解

基本概念

IP 地址

IP 地址由 32 位二进制组成，为了书写和描述方便，一般用十进制表示。十进制表示时，分为四组，每组 8 位，范围从 0~255，组之间用“.”号隔开，比如“192.168.1.1”就是用十进制表示的 IP 地址。

IP 地址顾名思义，自然是 IP 层协议的互连地址。32 位的 IP 地址由两个部分组成：1) 网络部分；2) 本地地址部分。根据网络部分的头几个比特位的值，目前使用中的 IP 地址可以划分成四大类。

A 类地址，最高比特位为“0”，有 7 个比特位表示网络号，24 个比特位表示本地地址。这样总共有 128 个 A 类网络。

图 1-2



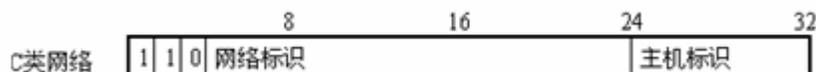
B 类地址，前两个最高比特位为“10”，有 14 个比特位表示网络号，16 个比特位表示本地地址。这样总共有 16,384 个 B 类网络。

图 1-3



C 类地址，前三个最高比特位为“110”，有 21 个比特位表示网络号，8 个比特位表示本地地址。这样总共有 2,097,152 个 C 类网络。

图 1-4



D 类地址，前四个最高比特位为“1110”，其余比特位为组播地址。

图 1-5



i 前四个最高比特位为“1111”的地址是不允许分配的，这些地址称为 E 类地址，属于保留地址。

在建设网络过程中，进行 IP 地址规划时，一定要根据建设网络的性质进行 IP 地址分配。如果建设的网络需要与互联网连接，则需要到相应的机构申请分配 IP 地址。中国地区可以向中国互联网信息中心（CNNIC）申请，负责 IP 地址分配的最终机构为国际互联网名字与编号分配公司（ICANN, Internet Corporation for Assigned Names and Numbers）。如果建设的网络为内部私有网络，就不需要申请 IP 地址，但是也不能随便分配，最好分配专门的私有网络地址。

下表为保留与可用的地址列表：

类别	地址空间	状态
A 类网络	0.0.0.0~0.255.255.255	保留
	1.0.0.0~126.255.255.255	可用
	127.0.0.0~127.255.255.255	保留
B 类网络	128.0.0.0~191.254.255.255	可用
	191.255.0.0~191.255.255.255	保留
C 类网络	192.0.0.0~192.0.0.255	保留
	192.0.1.0~223.255.254.255	可用
	223.255.255.0~223.255.255.255	保留
D 类网络	224.0.0.0~239.255.255.255	组播地址
E 类网络	240.0.0.0~255.255.255.254	保留
	255.255.255.255	广播地址

其中专门有三个地址块提供给私有网络，这些地址是不会在互联网中使用的，如果分配了这些地址的网络需要连接互联网，则需要将这些 IP 地址转换成有效的互联网地址。下表为私有网络地址空间，私有网络地址由 RFC 1918 文档定义：

类别	地址空间	状态
A 类网络	10.0.0.0~10.255.255.255	1 个 A 类网络
B 类网络	172.16.0.0~172.31.255.255	16 个 B 类网络
C 类网络	192.168.0.0~192.168.255.255	256 个 C 类网络

关于 IP 地址、TCP/UDP 端口及其它编码的分配情况，请参考 RFC 1166 文档。

子网掩码

网络掩码也是一个 32 比特的数值，标识着该 IP 地址的哪几个比特为网络部分。网络掩码中，值为“1”的比特对应的 IP 地址比特位就是网络部分，值为“0”的比特对应的 IP 地址比特位就是主机地址部分。如 A 类网络对应的网络掩码为“255.0.0.0”。您可以利用网络掩码对一个网络进行子网划分，子网划分就是将主机地址部分的一些比特位也作为网络部分，缩小主机容量，增加网络的数量，这时的网络掩码就称为子网掩码。

广播报文

广播报文是指目标地址为某个物理网络上所有主机的数据包。锐捷产品支持两种类型广播报文：1) 定向广播，是指数据包接收者为一个指定网络的所有主机，目标地址的主机部分全为“1”；2) 淹没广播，是指数据包接收者为所有网络的主机，目标地址 32 比特位全为“1”。

ICMP 报文

ICMP 是 (Internet Control Message Protocol) Internet 控制报文协议。它是 TCP/IP 协议族的一个子协议，用于在 IP 主机、网络设备之间传递控制消息，主要用于网络出现异常的时候通知相应设备。

📌 TTL

TTL (Time-To-Live)，生存时间。指定数据包被路由器丢弃之前允许通过的网段数量。它是IP协议报文中的一个值，它告诉网络，数据包在网络中的时间是否太长而应被丢弃。

功能特性

功能特性	作用
IP地址	用于配置接口 IP 地址，该接口才允许运行 IP 协议。
广播报文处理	设置 IP 广播地址，转发处理定向广播报文。
发送ICMP报文	控制 ICMP 协议报文的收发。
控制ICMP差错报文的发送速率	防止拒绝服务攻击。
IP MTU	用于配置接口 IP 报文的最大传输单元。
IP TTL	用于配置单播报文和广播报文的 TTL。
IP源路由	用于对接收报文的源路由进行检查。

1.3.1 IP地址

接口获取 IP 地址有以下方式：

- (1) 手工配置 IP 地址。
- (2) 利用 DHCP 协议获取 IP 地址。
- (3) 通过 PPP 协商获得 IP 地址。
- (4) 借用其它接口的 IP 地址。

这几种方式是互斥的，配置新的获取 IP 地址方式时会覆盖通过原有方式获取的 IP 地址。

i 利用 DHCP 协议获取 IP 地址请参见“DHCP”章节，以下仅介绍其他三种获取 IP 地址的方式。

📌 配置接口 IP 地址

一个设备只有配置了 IP 地址，才可以接收和发送 IP 数据包，接口配置了 IP 地址，说明该接口允许运行 IP 协议。

📌 接口配置多个 IP 地址

锐捷产品可以支持一个接口配置多个 IP 地址，其中一个为主 IP 地址，其余全部为次 IP 地址。次 IP 地址的配置理论上没有数目限制，但是次 IP 地址与主 IP 以及次 IP 地址之间必须属于不同网络。在网络建设中，会经常使用到次 IP 地址，通常在以下情况下应该考虑使用次 IP 地址：

- 一个网络没有足够多的主机地址。例如，现在一般局域网需要一个 C 类网络，可分配 254 台主机。但是当局域网主机超过 254 台时，一个 C 类网络将不够分配，有必要分配另一个 C 类网络地址。这样设备就需要连接两个网络，所以就配置多个 IP 地址。

- 许多旧的网络是基于第二层的桥接网络，没有进行子网的划分。次 IP 地址的使用可以使该网络很容易升级到基于 IP 层的路由网络。对于每个子网，设备都配置一个 IP 地址。
- 一个网络的两个子网被另外一个网络隔离开，可以创建一个被隔离网络的子网，通过配置次 IP 地址的方式，将隔离的子网连接起来。一个子网不能在设备的两个或两个以上接口出现。

i 配置次 IP 地址之前，需要确定已经配置了主 IP 地址。如果网络上的一台设备配置了次 IP 地址，则其它设备也必须配置同一网络的次 IP 地址。当然如果其它设备原先没有分配 IP 地址，可以配置为主地址。

配置接口借用 IP 地址

所谓“借用 IP 地址”，是指一个接口上没有配置 IP 地址，但为了使该接口能正常使用，就向同一设备上其它有 IP 地址的接口借用一个 IP 地址。

- i** 以太网接口、隧道接口和环回接口的 IP 地址可以被其它接口借用，但它们不能借用其它接口的 IP 地址。
- i** 被借用接口的 IP 地址不能是借用其它接口的 IP 地址。
- i** 如果被借用接口有多个 IP 地址，只有主 IP 地址被借用。
- i** 一个接口的 IP 地址可以借给多个接口。
- i** 借用接口的 IP 地址始终和被借用接口的 IP 地址保持一致，随着被借用接口的 IP 地址变化而变化。

相关配置

配置接口一个或多个 IP 地址

- 缺省情况接口没有配置 IP 地址。
- 通过 `ip address` 命令配置接口 IP 地址。
- 配置后根据冲突检测即可使用该 IP 地址进行通信。
- 通过 `ip address ip-address mask secondary` 可以配置多个次 IP 地址。

配置接口借用 IP 地址

- 缺省情况接口没有配置 IP 地址。
- 通过 `ip unnumbered` 命令可以向其他接口借用 IP 地址。

1.3.2 广播报文处理

工作原理

广播分两种，全广播，即 IP 地址为 255.255.255.255，由于会被路由器禁止传输，所以也叫本地网络广播。另一种是所有的主机位都为 1 的广播，例如：192.168.1.255/24，这种广播，通过配置是可以被转发的。

如果 IP 网络设备转发淹没广播（一般指目标 IP 地址为全“1”的广播报文），可能会引起网络的超负载，严重影响网络的运行，这种情况称为广播风暴。设备提供了一些办法能够将广播风暴限制在本地网络，阻止其继续扩张。但对于桥和交换机等基于二层网络设备，将转发和传播广播风暴。

解决广播风暴最好的办法就是给每个网络指定一个广播地址，这就是定向广播，这要求使用广播报文的 IP 协议尽可能应用定向广播而不是淹没广播进行数据传播。

关于广播问题的详细描述，请参见 RFC 919 和 RFC 922。

IP 定向广播报文是指目标地址为某个 IP 子网广播地址的 IP 报文，如目标地址为 172.16.16.255 的报文就称为定向广播报文。但是产生该报文的节点又不是目标子网的成员。

没有与目标子网直连的设备接收到 IP 定向广播报文，跟转发单播报文一样处理定向广播报文。当定向广播报文到达直连该子网的设备后，设备将把定向广播报文转换为淹没广播报文（一般指目标 IP 地址为全“1”的广播报文），然后以链路层广播方式发送给目标子网上的所有主机。

相关配置

配置 IP 广播地址

- 缺省情况下接口 IP 广播地址为 255.255.255.255。
- 如果需要定义其它地址的广播报文，可以在接口下配置 `ip broadcast-address` 命令。

允许转发定向广播

- 缺省情况接口不允许转发定向广播。
- 用户可以在指定的接口上，通过 `ip directed-broadcast` 命令配置接口允许转发定向广播，这样该接口就可以转发到直连网络的定向广播了。该命令只影响定向广播报文在目标子网的传输，而不影响其它定向广播报文的正常转发。
- 在接口上，用户还可以通过定义访问控制列表来控制转发某些定向广播。当定义了访问列表时，只有符合访问列表中定义的定向广播才会被转发。

1.3.3 发送ICMP报文

工作原理

ICMP 协议不可达消息

当设备接收到目标为自己的非广播报文，但是该数据包中采用了设备不能处理的 IP 协议，设备就向源地址发送 ICMP 协议不可达消息。另外，如果设备由于不知道路由而不能转发数据包时，也会发送 ICMP 主机不可达消息。

ICMP 重定向消息

路由有时会不够优化，使得设备从一个接口接收到的数据包，还要从该接口发送出去。如果设备将数据包从接收接口重新发送出去，设备就会给数据源发送一个 ICMP 重定向消息，告诉数据源到该目标地址的网关为同一子网上的另外一台设备。这样数据源就会将后续的数据包按照最佳的路径进行发送。

📌 ICMP 掩码应答消息

网络设备有时需要知道互联网上某个子网的子网掩码,为了获取该信息,网络设备可以发送 ICMP 掩码请求消息,接收到 ICMP 掩码请求消息的网络设备就会发送掩码应答消息。

相关配置

📌 启用 ICMP 协议不可达消息

- 缺省情况接口启用 ICMP 协议不可达消息功能。
- 可通过[no] ip unreachable 命令关闭或启用该功能。

📌 启用 ICMP 重定向消息

- 缺省情况接口启用 ICMP 协议重定向消息功能。
- 可通过[no] ip redirects 命令关闭或启用该功能。

📌 启用 ICMP 掩码应答消息

- 缺省情况接口启用 ICMP 掩码应答消息功能。
- 可通过[no] ip mask-reply 命令关闭或启用该功能。

1.3.4 控制ICMP差错报文的发送速率

工作原理

为了防止拒绝服务攻击,对 ICMP 差错报文的发送速率进行限制,采用令牌桶算法。

如果 IP 报文需要分片,但是 IP 首部的不可分片位被设置了,设备会向源 IP 地址发送编号为 4 的 ICMP 目的不可达报文,这种 ICMP 差错报文的主要用途是路径 MTU 发现。为了防止其它 ICMP 差错报文太多导致发不出编号为 4 的 ICMP 目的不可达报文,从而导致路径 MTU 发现功能失效,对编号为 4 的 ICMP 目的不可达报文和其它 ICMP 差错报文分别限速。

相关配置

📌 配置 IP 首部不可分片位触发的 ICMP 目的不可达报文的发送速率

- 缺省速率是 100 毫秒 10 个。
- 可通过 ip icmp error-interval DF 配置发送速率。

📌 配置其它 ICMP 差错报文的发送速率

- 缺省速率是 100 毫秒 10 个。
- 可通过 ip icmp error-interval 配置发送速率。

1.3.5 IP MTU

工作原理

如果一个 IP 报文超过 IP MTU 的大小，RGOS 软件就会对报文进行拆分。所有在同一物理网段上的设备，其互联接口的 IP MTU 一定要一致。锐捷产品允许调整接口的链路 MTU 值，而且接口的链路 MTU 的变化会引起接口的 IP MTU 的变化，接口的 IP MTU 会自动与接口的链路 MTU 保持一致。但是反之不行，如果调整了接口的 IP MTU 值，接口的链路 MTU 不会跟着改变。

相关配置

▾ 设置 IP MTU

- 缺省情况接口 IP MTU 为 1500。
- 可通过 `ip mtu` 设置 IP 包最大传输单元(MTU)。

1.3.6 IP TTL

工作原理

IP 数据包从源地址向目的地址经过路由器间传播，设置一个 TTL 数值，每过一个路由器 TTL 值就减一，当减到零的时候，路由器就把这个包丢掉，这样可以防止无用的包在网络上无限传播下去，浪费网络带宽。

相关配置

▾ 设置 IP TTL

- 缺省情况接口 IP TTL 为 64。
- 可通过 `ip ttl` 设置接口的 IP TTL 值。

1.3.7 IP源路由

工作原理

锐捷产品支持 IP 源路由。当设备接收到 IP 数据包时，会对 IP 报头的严格源路由、宽松源路由和记录路由等选项进行检查，这些选项在 RFC 791 中有详细描述。如果检测到该数据包启用了其中一个选项，就会执行响应的动作；如果检测到无效的选项，就会给数据源发送一个 ICMP 参数问题消息，然后丢弃该数据包。

开启 IP 源路由，在 IP 数据报选项中增加源路由选项，可用于测试某特定网络的吞吐率，也可以是数据报绕开出错的网络。然而，可能会导致诸如源地址欺骗(Source Address Spoofing)、IP 欺骗(IP Spoofing)等的网络攻击。

相关配置

配置 IP 源路由

- 缺省情况开启 IP 源路由功能。
- 可通过 `ip source-route` 开启或关闭该功能。

1.4 产品说明



配置此功能时需要注意保证三层接口的 IP MTU 与链路 MTU 的合理性，IP MTU 不大于接口 MTU，三层接口包括路由口、三层 AP 口和 SVI。

1.5 配置详解

配置项	配置建议 & 相关命令	
配置接口IP地址	⚠ 必须配置。用于配置 ip 地址，允许接口运行 IP 协议。	
	<code>ip address</code>	手工配置接口 IP 地址
	<code>ip address negotiate</code>	配置通过 PPP 协商获取 ip 地址
	<code>ip unnumbered</code>	配置接口借用 IP 地址
配置广播报文处理方式	⚠ 可选配置。用于设置 IP 广播地址，允许转发定向广播报文。	
	<code>ip broadcast-address</code>	配置 IP 广播地址
	<code>ip directed-broadcast</code>	允许转发定向广播
配置发送ICMP报文	⚠ 可选配置。用于控制 ICMP 协议报文的收发。	
	<code>ip unreachable</code>	启用 ICMP 协议不可达和主机不可达消息
	<code>ip redirects</code>	启用 ICMP 重定向消息
	<code>ip mask-reply</code>	启用掩码应答消息
配置ICMP差错报文的发送速率	⚠ 可选配置。	
	<code>ip icmp error-interval DF</code>	配置 IP 首部不可分片位触发的 ICMP 目的不可达报文的发送速率
	<code>ip icmp error-interval</code>	配置其它 ICMP 差错报文和 ICMP 重定向报文的发送速率
设置IP MTU	⚠ 可选配置。用于配置接口 IP 报文的最大传输单元。	
	<code>ip mtu</code>	设置 MTU 值
设置IP TTL	⚠ 可选配置。用于配置单播报文和广播报文的 TTL。	
	<code>ip ttl</code>	设置 TTL 值
配置IP源路由	⚠ 可选配置。用于配置对接收报文的源路由进行检查。	
	<code>ip source-route</code>	启用 IP 源路由

1.5.1 配置接口IP地址

配置效果

通过配置接口 IP 地址实现 IP 网络通信。

注意事项

-

配置方法

✎ 手工配置接口 IP 地址

- 必须配置。
- 在三层接口模式下配置。

✎ 配置通过 PPP 协商获取接口 IP 地址

- 可选配置。
- 如果点对点接口上没有配置 IP 地址，且需要通过 PPP 协商获取 IP 地址时配置。
- 在三层接口模式下配置。

✎ 配置接口借用 IP 地址

- 可选配置。
- 如果接口上没有配置 IP 地址，且需要向其它接口借用 IP 地址时配置。
- 在三层接口模式下配置。

检验方法

通过 **show ip interface** 可以看到配置的地址生效

相关命令

✎ 手工配置接口 IP 地址

【命令格式】 **ip address** *ip-address network-mask* [**secondary**]

【参数说明】 *ip-addr0065ss* : 32 个比特位 IP 地址，8 位一组，以十进制方式表示，组之间用点隔开。
network-mask : 32 个比特位网络掩码，“1”表示掩码位，“0”表示主机位。每 8 位一组，以十进制方式表示，组之间用点隔开。

secondary : 表示配置的次 IP 地址。

【命令模式】 接口模式

【使用指导】 -

配置接口借用 IP 地址

【命令格式】 **ip unnumbered** *interface-type interface-number*

【参数说明】 *interface-type* : 关联接口类型。

interface-number : 关联接口编号。

【命令模式】 接口模式

【使用指导】 无编号接口就是只在接口启动 IP 协议，但是不分配 IP 地址，无编号接口需要关联一个具有 IP 地址的接口。无编号接口产生的 IP 数据包，该数据包的源 IP 地址为关联接口的 IP 地址。另外路由协议进程也根据关联接口的 IP 地址，决定是否往无编号接口发送路由更新报文。应用无编号接口，需要注意以下限制：

以太网接口不能配置成无编号接口。

当串行口封装 SLIP、HDLC、PPP、LAPB、Frame-relay 时，可以配置成无编号接口。但是封装帧中继时，只有点到点子接口才允许配置无编号接口。X.25 封装是不允许配置无编号接口的。

不能用 ping 命令来检测无编号接口是否工作正常，因为无编号接口没有 IP 地址。但是通过 SNMP 可以远程监测到无编号接口状态。

不能通过无编号接口进行网络启动。

配置举例

给接口配置 IP 地址

【配置方法】 在接口 GigabitEthernet 0/0 配置 ip 地址 192.168.23.110 255.255.255.0

```
Ruijie#configure terminal
Ruijie(config)#interface gigabitEthernet 0/0
Ruijie(config-if-GigabitEthernet 0/0)# no switchport
Ruijie(config-if-GigabitEthernet 0/0)#ip address 192.168.23.110 255.255.255.0
```

【检验方法】 使用 **show ip interface** 可以看到接口 GigabitEthernet 0/0 添加地址成功

```
Ruijie# show ip interface gigabitEthernet 0/0
GigabitEthernet 0/0
  IP interface state is: UP
  IP interface type is: BROADCAST
  IP interface MTU is: 1500
  IP address is:
    192.168.23.110/24 (primary)
```

给点对点接口配置通过 PPP 协商获取 IP 地址

【配置方法】 在接口点对点接口上配置通过协商获取 ip 地址

```
Ruijie(config)#int virtual-ppp 1
```

```
Ruijie(config-if-Virtual-ppp 1)#ip address negotiate
```

【检验方法】 使用 **show run** 可以看到点对点接口相关配置

```
Ruijie#show run interface virtual-ppp 1
```

```
Building configuration...
```

```
Current configuration: 48 bytes
```

```
interface Virtual-ppp 1
```

```
 ip address negotiate
```

1.5.2 配置广播报文处理方式

配置效果

配置接口广播地址为 0.0.0.0，并允许转发定向广播报文。

注意事项

-

配置方法

▾ 配置 IP 广播地址

- 可选配置，有些老的主机可能只认 0.0.0.0 的广播地址，此时需要配置接口的广播地址为 0.0.0.0。
- 在三层接口模式下配置。

▾ 允许转发定向广播

- 可选配置，向处在一个广播域的全部主机发送广播，但是发送者并不处在这个广播域内，此时需要配置允许转发定向广播。
- 在三层接口模式下配置。

检验方法

通过 **show running-config interface** 可以看到配置生效

相关命令

▾ 配置 IP 广播地址

【命令格式】 **ip broadcast-address ip-address**

- 【参数说明】 *ip-address* : IP 网络的广播地址。
- 【命令模式】 接口模式
- 【使用指导】 目前 IP 广播报文的目标地址一般为全 “1” , 表示为 255.255.255.255。RGOS 软件可以通过定义产生其它 IP 地址的广播报文, 而且可以同时接收全 “1” 以及自己定义的广播包。

允许转发定向广播

- 【命令格式】 **ip directed-broadcast** [*access-list-number*]
- 【参数说明】 *access-list-number* : 访问列表号, 范围从 1-199, 1300 - 2699。如果定义了访问列表号, 只有匹配该访问列表的 IP 定向广播报文才转换。
- 【命令模式】 接口模式
- 【使用指导】 如果在接口上配置了 **no ip directed-broadcast**, RGOS 将丢弃接收到的直连网络的定向广播报文。

配置举例

- 【配置方法】 在设备端口 gigabitEthernet 0/1 配置 IP 广播报文的目标地址为 0.0.0.0, 启用定向广播的转发。

```
Ruijie#configure terminal
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# no switchport
Ruijie(config-if-GigabitEthernet 0/1)# ip broadcast-address 0.0.0.0
Ruijie(config-if-GigabitEthernet 0/1)# ip directed-broadcast
```

- 【检验方法】 使用 **show ip interface** 可以看到接口 GigabitEthernet 0/1 配置成功

```
Ruijie#show running-config interface gigabitEthernet 0/1
ip directed-broadcast
ip broadcast-address 0.0.0.0
```

1.5.3 配置发送ICMP报文

配置效果

启用接口 ICMP 协议不可达消息, ICMP 重定向消息以及掩码应答消息。

注意事项

-

配置方法

启用 ICMP 协议不可达消息

- 缺省开启 ICMP 协议不可达消息。

- 可选配置，通过 **no ip unreachable**s 禁止该功能。
- 在三层接口模式下配置。

▾ 启用 ICMP 重定向消息

- 缺省开启 ICMP 重定向消息。
- 可选配置，通过 **no ip redirects** 禁止该功能。
- 在三层接口模式下配置。

▾ 启用 ICMP 掩码应答消息

- 缺省开启 ICMP 掩码应答消息。
- 可选配置，通过 **no ip mask-reply** 禁止该功能。
- 在三层接口模式下配置。

检验方法

通过 **show ip interface** 可以看到配置生效。

相关命令

▾ 启用 ICMP 协议不可达消息

- 【命令格式】 **ip unreachable**s
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 -

▾ 启用 ICMP 重定向消息

- 【命令格式】 **ip redirects**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 -

▾ 启用 ICMP 掩码应答消息

- 【命令格式】 **ip mask-reply**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 -

配置举例

【配置方法】 在设备端口 gigabitEthernet 0/1 启用 ICMP 协议不可达消息 ,ICMP 重定向消息以及 ICMP 掩码应答消息功能。

```
Ruijie#configure terminal
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# no switchport
Ruijie(config-if-GigabitEthernet 0/1)# ip unreachable
Ruijie(config-if-GigabitEthernet 0/1)# ip redirects
Ruijie(config-if-GigabitEthernet 0/1)# ip mask-reply
```

【检验方法】 使用 **show ip interface** 可以看到接口 GigabitEthernet 0/1 配置成功

```
Ruijie#show ip interface gigabitEthernet 0/1
GigabitEthernet 0/1
  ICMP mask reply is: ON
  Send ICMP redirect is: ON
  Send ICMP unreachable is: ON
```

1.5.4 配置ICMP报文差错报文的发送速率

配置效果

配置 ICMP 差错报文的发送速率。

注意事项

-

配置方法

▾ 配置 IP 首部不可分片位触发的 ICMP 目的不可达报文的发送速率

- 可选配置。
- 在全局模式下配置。

▾ 配置其它 ICMP 差错报文的发送速率

- 可选配置。
- 在全局模式下配置。

检验方法

执行 **show running-config** 可以看到配置生效。

相关命令

配置 IP 首部不可分片位触发的 ICMP 目的不可达报文的发送速率

【命令格式】 **ip icmp error-interval DF milliseconds [bucket-size]**

【参数说明】 *milliseconds* : 令牌桶的刷新周期, 取值范围 0~2147483647, 缺省值为 100, 单位为毫秒。取值为 0 时, 表示不限制 ICMP 差错报文的发送速率。

bucket-size : 令牌桶中容纳的令牌数, 取值范围 1~200, 缺省值为 10。

【命令模式】 全局模式

【使用指导】 为了防止拒绝服务攻击, 对 ICMP 差错报文的发送速率进行限制, 采用令牌桶算法。

如果 IP 报文需要分片, 但是 IP 首部的不可分片位被设置了, 设备会向源 IP 地址发送编号为 4 的 ICMP 目的不可达报文, 这种 ICMP 差错报文的主要用途是路径 MTU 发现。为了防止其它 ICMP 差错报文太多导致发不出编号为 4 的 ICMP 目的不可达报文, 从而导致路径 MTU 发现功能失效, 对编号为 4 的 ICMP 目的不可达报文和其它 ICMP 差错报文分别限速。

因为定时器的精度是 10 毫秒, 建议用户把令牌桶的刷新周期配置成 10 毫秒的整数倍。如果令牌桶的刷新周期大于 0 小于 10, 实际生效的刷新周期是 10 毫秒, 例如配置 5 毫秒 1 个, 实际效果是 10 毫秒 2 个; 如果令牌桶的刷新周期不是 10 毫秒的整数倍, 实际生效的刷新周期自动换算成 10 毫秒的整数倍, 例如配置 15 毫秒 3 个, 实际效果是 10 毫秒 2 个。

配置其它 ICMP 差错报文的发送速率

【命令格式】 **ip icmp error-interval milliseconds [bucket-size]**

【参数说明】 *milliseconds* : 令牌桶的刷新周期, 取值范围 0~2147483647, 缺省值为 100, 单位为毫秒。取值为 0 时, 表示不限制 ICMP 差错报文的发送速率。

bucket-size : 令牌桶中容纳的令牌数, 取值范围 1~200, 缺省值为 10。

【命令模式】 全局模式

【使用指导】 为了防止拒绝服务攻击, 对 ICMP 差错报文的发送速率进行限制, 采用令牌桶算法。

因为定时器的精度是 10 毫秒, 建议用户把令牌桶的刷新周期配置成 10 毫秒的整数倍。如果令牌桶的刷新周期大于 0 小于 10, 实际生效的刷新周期是 10 毫秒, 例如配置 5 毫秒 1 个, 实际效果是 10 毫秒 2 个; 如果令牌桶的刷新周期不是 10 毫秒的整数倍, 实际生效的刷新周期自动换算成 10 毫秒的整数倍, 例如配置 15 毫秒 3 个, 实际效果是 10 毫秒 2 个。

配置举例

【配置方法】 配置 IP 首部不可分片位触发的 ICMP 目的不可达报文的发送速率为 1 秒 100 个, 配置其它 ICMP 差错报文的发送速率为 1 秒 10 个。

```
Ruijie(config)# ip icmp error-interval DF 1000 100
```

```
Ruijie(config)# ip icmp error-interval 1000 10
```

【检验方法】 执行 **show running-config** 可以看到配置生效

```
Ruijie#show running-config | include ip icmp error-interval
```

```
ip icmp error-interval 1000 10
```

```
ip icmp error-interval DF 1000 100
```

1.5.5 配置IP MTU

配置效果

调整 IP 包最大传输单元。

注意事项

-

配置方法

- 可选配置，所有在同一物理网段上的设备，当互联接口的 IP MTU 不一致时需要配置为一致。
- 在三层接口模式下配置。

检验方法

通过 **show ip interface** 可以看到配置生效

相关命令

▾ 配置 IP MTU

【命令格式】 **ip mtu bytes**

【参数说明】 *bytes* : IP 包最大传输单元，以字节为单位，范围 68~1500。

【命令模式】 接口模式

【使用指导】 -

配置举例

【配置方法】 将 gigabitEthernet 0/1 接口的 IP MTU 值设为 512 字节

```
Ruijie#configure terminal
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# no switchport
Ruijie(config-if-GigabitEthernet 0/1)# ip mtu 512
```

【检验方法】 使用 **show ip interface** 可以看到接口 GigabitEthernet 0/1 配置成功

```
Ruijie# show ip interface gigabitEthernet 0/1
IP interface MTU is: 512
```

1.5.6 配置IP TTL

配置效果

修改接口的 IP TTL 值。

注意事项

-

配置方法

- 可选配置。
- 在三层接口模式下配置。

检验方法

通过 **show run-config** 可以看到配置生效

相关命令

▾ 配置 IP TTL

- 【命令格式】 **ip ttl value**
- 【参数说明】 *value* : TTL 值, 取值范围是 0~255。
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

- 【配置方法】 配置本机发送的单播报文的缺省 TTL 值为 100。

```
Ruijie#configure terminal
Ruijie(config)#ip ttl 100
```

- 【检验方法】 通过 **show run-config** 可以看到配置生效

```
Ruijie#show running-config
ip ttl 100
```

1.5.7 配置IP源路由

配置效果

开启或关闭 IP 源路由信息的处理功能。

注意事项

-

配置方法

- 缺省情况下开启 IP 源路由功能。
- 可选配置，通过 **no ip source-route** 可关闭 IP 源路由功能。

检验方法

通过 **show run-config** 可以看到配置生效。

相关命令

▾ 配置 IP 源路由

【命令格式】 **ip source-route**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

配置举例

【配置方法】 关闭了 IP 源路由信息的处理功能。

```
Ruijie#configure terminal
Ruijie(config)# no ip source-route
```

【检验方法】 通过 **show run-config** 可以看到配置生效

```
Ruijie#show running-config
no ip source-route
```

1.6 监视与维护

清除各类信息

-

查看运行情况

作用	命令
显示接口 IP 信息	show ip interface [<i>interface-type interface-number</i> brief]
显示 IP 报文统计值	show ip packet statistics [total <i>interface-name</i>]

查看调试信息

-

2 ARP

2.1 概述

在局域网中，每个 IP 网络设备都有两个地址：1) 本地地址，由于它包含在数据链路层的帧头中，更准确地说应该是数据链路层地址，但实际上对本地地址进行处理的是数据链路层中的 MAC 子层，因此习惯上称为 MAC 地址，MAC 地址在局域网上代表着 IP 网络设备；2) 网络地址，在互联网上代表着 IP 网络设备，同时它也说明了该设备所属的网络。

局域网上两台 IP 设备之间需要通信，必须要知道对方的 48 比特的 MAC 地址。根据 IP 地址来获知 MAC 地址的过程称为地址解析。地址解析的方式有两类：1) 地址解析协议 (ARP)；2) 代理地址解析协议 (Proxy ARP)。关于 ARP、Proxy ARP，分别在 RFC 826，RFC 1027 文档中描述。

ARP(Address Resolution Protocol，地址解析协议)是用来绑定 MAC 地址和 IP 地址的，以 IP 地址作为输入，ARP 能够知道其关联的 MAC 地址。一旦知道了 MAC 地址，IP 地址与 MAC 地址对应关系就会保存在设备的 ARP 缓存中。有了 MAC 地址，IP 设备就可以封装链路层的帧，然后将数据帧发送到局域网上去。缺省配置下，以太网上 IP 和 ARP 的封装为 Ethernet II 类型。

协议规范

- RFC826 : An Ethernet Address Resolution Protocol
- RFC1027 : Using ARP to implement transparent subnet gateways

2.2 典型应用

典型应用	场景描述
在局域网内提供地址解析协议服务	在同一网段中，主机学习其他设备的 MAC 地址，需要用到地址解析协议。
使用代理ARP实现透明的子网网关	通过代理地址解析服务，允许主机在不知道另一个网络是否存在的情况下和另一网络内的主机直接通讯。

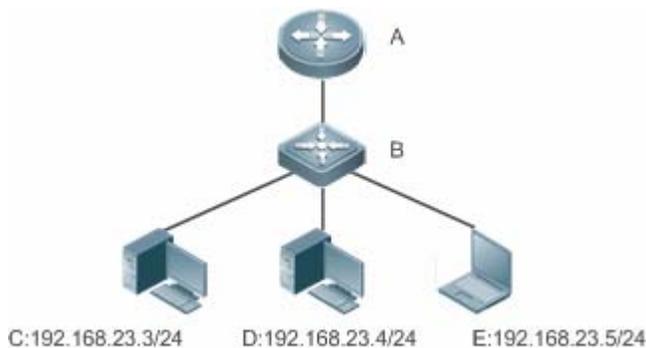
2.2.1 在局域网内提供地址解析协议服务

应用场景

在所有 IPv4 局域网内，都需要用到 ARP 协议。

- 主机需要通过 ARP 协议来学习其他设备的 MAC 地址，只有学到 MAC 地址后，主机才可以和其他设备通信。

图 2-1



- 【注释】 A 为路由器
B 为交换机，作为用户主机网段的网关。
C、D、E 为用户主机

功能部属

- 在局域网内运行 ARP 协议，实现 IP 地址和 MAC 地址的映射。

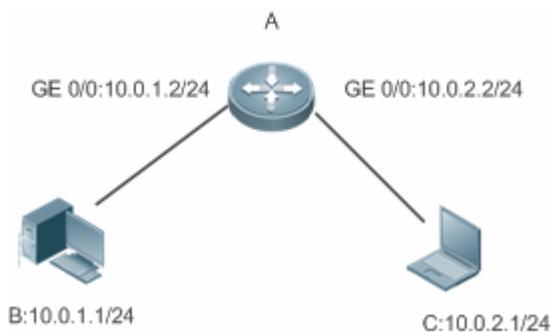
2.2.2 使用代理ARP实现透明的子网网关

应用场景

在不同的 IPv4 局域网内，实现透明的子网网关。

- 通过在设备上配置代理 ARP 的功能，实现不同网段内主机的直接通讯。

图 2-2



- 【注释】 A 为路由器，连接两个局域网
B、C 为用户主机，不配置默认网关，在不同的子网

功能部属

- 在子网网关上运行代理 ARP 功能，可以帮助没有路由信息的主机获得其它子网 IP 地址的 MAC 地址。

2.3 功能详解

功能特性

功能特性	作用
静态ARP	用户手工指定 IP 地址和 MAC 地址的映射，防止设备学到错误的 ARP 表项而影响网络。
ARP属性设置	用户指定 ARP 表项的超时时间、ARP 请求重传次数和间隔、未解析 ARP 表项数上限。
可信任ARP	防止 ARP 欺骗。
免费ARP	检测 IP 地址冲突，以及让外围设备更新本机的 ARP。
代理ARP	代理应答请求其他设备的 ARP 请求。
本地代理ARP	代理应答一台设备请求子网内的其他设备的 ARP 请求
ARP可信检测	通过 NDU（邻居不可达探测），保证学习的 ARP 表项正确。
ARP防IP报文攻击	通过设置触发 ARP 丢弃表项的 IP 报文个数，触发设置丢弃表项到硬件，来防止未知名单播报文大量送 CPU 对 CPU 造成冲击。
抑制往认证VLAN发送ARP请求	通过设置抑制往认证 vlan 发送广播 arp 请求报文，可以减少网络中 arp 广播报文的数量，改善网络环境

2.3.1 静态ARP

静态 ARP 包括手工配置的静态 ARP 和认证下发的静态 ARP。手工配置的静态 ARP 优先级大于认证下发的静态 ARP。静态 ARP 能够防止设备学到错误的 ARP 表项而影响网络。

工作原理

静态 ARP，设备不会再去主动更新 ARP 表项，并且永久存在。

设备转发三层报文时，以太头部的目的 MAC 地址将采用静态配置的 MAC 地址来封装。

相关配置

配置静态 ARP

手工配置的静态 ARP，在全局模式下，使用 `arp ip-address mac-address type` 命令配置静态 ARP 表项。缺省情况下用户没有配置任何静态 ARP 表项。ARP 封装只支持 Ethernet II 类型，用 arpa 表示。

2.3.2 ARP属性设置

用户指定 ARP 表项的超时时间、ARP 请求重传次数和间隔、未解析 ARP 表项数上限、接口 ARP 学习数量限制。

工作原理

↘ ARP 超时设置

ARP 超时设置只对动态学习到的 IP 地址和 MAC 地址映射起作用。当一个 ARP 表项超时后，设备会发送单播 ARP 请求报文探测对方是否在线，假如能收到对方的 ARP 应答，则说明对方仍在线，该 ARP 表项不会删除，否则会删除该 ARP 表项。

超时时间设置得越短，ARP 缓冲中保存的映射表就越真实，但是 ARP 消耗网络带宽也越多。

↘ ARP 请求重传时间间隔和次数

IP 地址解析成 MAC 地址时连续发送 ARP 请求的时间间隔和次数。时间间隔越短，解析速率更快。次数越多，解析成功率更大，但是 ARP 消耗网络带宽也越多。

↘ 未解析 ARP 表项的数量限制

在局域网中可能存在对网关的攻击，扫描网段，使网关生成大量未解析的 ARP 表项，从而使网关无法正常学习主机的 MAC 地址。为了防止这种攻击，用户可以配置未解析 ARP 表项的数量限制。

↘ 接口 ARP 学习数量限制

改成通过配置指定接口的用户 ARP 表项个数，灵活控制 ARP 表项资源的按需分配，防止表项资源浪费。

相关配置

↘ ARP 超时设置

在接口模式下，使用命令 **arp timeout seconds** 配置 ARP 的超时时间。默认情况下超时时间为 3600 秒，用户可以根据实际情况重新调整。

↘ ARP 请求重传时间间隔和次数

- 在全局模式下，使用命令 **arp retry interval seconds** 配置 ARP 的重传时间间隔。默认情况下超时时间为 1 秒，用户可以根据实际情况重新调整。
- 在全局模式下，使用命令 **arp retry times number** 配置 ARP 的重传次数。默认情况下可以连续发送 5 次，用户可以根据实际情况重新调整。

↘ 未解析 ARP 表项的数量限制

在全局模式下，使用命令 **arp unresolve number** 配置 ARP 的未解析表项数。默认为 arp 容量的最大值，用户可以根据实际情况重新调整。

↘ 接口 ARP 学习数量限制

在接口模式下，使用命令 **arp cache interface-limit limit** 配置接口 ARP 的学习数量限制。默认不限制接口上 ARP 学习的数量，用户可以根据实际情况重新调整。此数量限制包含静态 ARP。

2.3.3 可信任ARP

工作原理

可信任 ARP 作为一类特殊 ARP，添加在交换机端的 ARP 表中，用于防止 ARP 欺骗。可信任 ARP 同时具有静态 ARP 和动态 ARP 两者的特征，其优先级高于动态 ARP 表项、并且低于静态 ARP 表项。可信任 ARP 具有类似于动态 ARP 的老化机制，在 ARP 老化时主动发送 ARP 请求报文探测主机是否存在，如果主机有应答则代表主机还是活动的，那么就更新 ARP 的老化时间，否则删除 ARP 表项。可信任 ARP 具有静态 ARP 的相关特征，即不会通过学习 ARP 报文动态更新 ARP 表项的 MAC、接口等相关字段。

可信任 ARP 是 GSN 客户端用户认证上线时，认证服务端通过接入交换机获取用户真实的 IP-MAC 关联信息，并根据用户的网关信息，在网关交换机上添加的。该过程对于网络管理员来说是透明的，不会对网络管理员的原有网络管理产生任何影响。

综上所述，因为可信任 ARP 来源真实有效，且不会被 ARP 报文动态更新，所以可以有效的防止针对网关的 ARP 欺骗。

相关配置

配置可信任 ARP 功能

- 全局模式下，使用命令 **service trustedarp** 打开可信任 arp 功能，缺省情况下该功能是关闭的。
- 全局模式下，使用命令 **arp trusted user-vlan vid1 translated-vlan vid2** 实现 VLAN 转换，缺省情况下没有任何 VLAN 转换。如果服务器下发的 VLAN 和可信任 ARP 表项生效的 VLAN 不同，则用户需要配置 VLAN 转换。
- 全局模式下，使用命令 **arp trusted aging** 允许可信任 ARP 老化。缺省情况下可信任 ARP 表项不允许老化。
- 全局模式下，使用命令 **arp trusted number** 设置可信任 ARP 表项的容量。缺省情况下为总容量的一半，用户可以根据实际情况更改容量。

2.3.4 免费ARP

工作原理

免费 ARP 报文是一种特殊的 ARP 报文，该报文的发送端 IP 地址和目标 IP 地址都是本机 IP 地址。免费 ARP 的主要用途有：

1. IP 地址冲突检测。当设备收到免费 ARP 报文后，如果发现报文中的 IP 地址和自己的 IP 地址相同，向发送免费 ARP 报文的设备返回一个 ARP 应答，告诉该设备 IP 地址冲突。
2. 当接口的 MAC 地址变化时，发送免费 ARP 通知其它设备更新 ARP 表项。

设备具有免费 ARP 报文学习功能。当设备收到免费 ARP 报文时，设备判断是否存在和免费 ARP 报文源 IP 地址对应的动态 ARP 表项，如果存在，根据免费 ARP 报文中携带的信息更新 ARP 表项。

相关配置

配置免费 ARP

接口模式下，使用命令 `arp gratuitous-send interval seconds [number]` 允许接口定时发送免费 ARP 请求报文。缺省情况下接口上该功能是关闭的。一般在该接口充当下联设备网关时，需要开启这个功能，定时更新使下联设备的网关 mac，防止他人冒充网关。

2.3.5 代理ARP

工作原理

设备的代理 ARP 功能可以帮助没有路由信息的主机，获得其它子网 IP 地址的 MAC 地址。比如设备接收到一个 ARP 请求，ARP 请求的发送者 IP 地址与目标 IP 地址不属于同一网段，而设备又知道所请求 IP 地址的路由，设备就会发送 ARP 响应，响应的 MAC 地址为设备自身的以太网 MAC 地址，这个过程就是代理 ARP 的功能。

相关配置

配置代理 ARP

- 接口模式下，使用命令 `ip proxy-arp` 开启代理 ARP 功能。
- 缺省情况下路由器上开启了代理 ARP 功能，而交换机上关闭了代理 ARP 功能。

2.3.6 本地代理ARP

工作原理

本地 ARP 代理指的是同一个 VLAN 内的代理，这里头的 VLAN 指的是普通 VLAN 或者 Sub VLAN。

设备开启本地代理 ARP 功能后，可以帮助主机获得子网内其他主机的 MAC 地址。比如在设备开启端口保护的情况下，不同端口下的用户被二层隔离，在开启本地 ARP 代理功能后，设备接收到一个 ARP 请求，就会代理发送 ARP 响应，响应的 MAC 地址为设备自身的以太网 MAC 地址，这样不同的主机之间的通信靠三层路由来实现。这个过程就是本地代理 ARP 的功能。

相关配置

配置本地代理 ARP

- 接口模式下，使用命令 `local-proxy-arp` 开启本地代理 ARP 功能。
- 缺省情况下关闭本地代理 ARP 功能。
- 该命令只在 SVI 口上支持。

2.3.7 ARP可信检测

工作原理

该命令用于防止 arp 欺骗导致无用的 arp 表项过多占用设备资源。在三层接口开启 arp 可信检测功能后，从该接口上收到 arp 请求报文：

1. 如果对应表项不存在，则创建动态 arp 表项，并经过 1 到 5 秒的一个随机时间后进入 NUD（邻居不可达探测），即将新学习的 arp 表项设置为老化状态并单播 arp 请求，在老化时间内收到对端 arp 更新，则保存表项，否则直接删除该表项。
2. 如果对应 arp 表项已经存在，则不进行 NUD 探测逻辑。
3. 如果已有的动态 arp 表项的 MAC 地址被更新，也走 NUD 探测逻辑。

该功能由于在 ARP 学习过程中增加了一个严格确认的过程，所以开启该功能会影响到 ARP 的学习性能。

关闭该功能后，arp 表项的学习和更新不再走 NUD 逻辑。

相关配置

配置 ARP 可信检测

接口模式下，使用命令 **arp trust-monitor enable** 命令开启 ARP 可信检查功能，缺省情况下没有开启该功能。

2.3.8 ARP防IP报文攻击

工作原理

在收到未解析的 IP 报文时，交换机设备不能够进行硬件转发，需要把报文送 CPU 进行地址解析，如果此类报文大量送 CPU，就会对 CPU 造成冲击，影响交换机其它业务的运行。

开启 ARP 防 IP 报文攻击后，在 ARP 请求期间，交换机 CPU 会统计收到的目的 IP 命中该 ARP 表项的报文个数，当这个个数等于配置的个数时，会设置一个丢弃表项到硬件，后续硬件收到所有该目的 IP 的报文都不会送 CPU；在地址解析完成时，更新上述表项为转发状态，使得交换机能够对该目的 IP 的报文进行硬件转发。

相关配置

配置 ARP 防 IP 报文攻击

- 全局模式下，使用命令 **arp anti-ip-attack** 配置触发 ARP 设丢弃表项的 IP 报文个数。
- 缺省情况下，在 3 个目的 IP 地址相同的未知名单播报文送 CPU 后，就会设置丢弃表项。

2.3.9 抑制往认证VLAN发送ARP请求

工作原理

在网关认证模式下，SuperVLAN 下的所有子 VLAN 默认都是认证 VLAN，认证 VLAN 下的认证用户需要在认证后才能上网。用户认证后会在设备上生成静态 ARP 表项，因此设备访问认证用户时，不需要往认证 VLAN 发送 ARP 请求。若设备需要访问免认证 VLAN 下的用户时，只需要往免认证 VLAN 发送 ARP 请求。

在网关认证模式下，设备默认开启了抑制往认证 VLAN 发送 ARP 请求的功能。如果设备需要访问认证 VLAN 下的免认证用户，需要关闭该功能。

相关配置

配置抑制往认证 VLAN 发送 ARP 请求

- 接口模式下，使用命令 `arp suppress-auth-vlan-req` 开启抑制往认证 VLAN 发送 ARP 请求功能。
- 缺省情况下开启抑制往认证 VLAN 发送 ARP 请求功能。

2.4 配置详解

配置项	配置建议 & 相关命令	
配置静态ARP	 可选配置，用于 IP 地址和 MAC 地址的静态绑定。	
	<code>arp</code>	定义静态 ARP
配置ARP属性	 可选配置，用于指定 ARP 表项的超时时间、ARP 请求重传次数和间隔、未解析 ARP 表项数上限、接口 ARP 学习数量限制。	
	<code>arp timeout</code>	配置 ARP 超时时间
	<code>arp retry interval</code>	配置 ARP 请求重传时间间隔
	<code>arp unresolve</code>	配置未解析 ARP 表项的数量限制
	<code>arp cache interface-limit</code>	配置接口 ARP 学习数量限制
配置可信任ARP	 可选配置，用于防止 ARP 欺骗。	
	<code>service trustedarp</code>	启用可信任 ARP 功能
	<code>arp trusted user-vlan</code>	添加可信任 ARP 时进行 VLAN 转换
	<code>arp trusted aging</code>	允许可信任 ARP 老化
	<code>arp trusted</code>	调整可信任 ARP 的容量
配置免费ARP	 可选配置，用于检测 IP 地址冲突，以及让外围设备更新本机的 ARP。	
	<code>arp gratuitous-send interval</code>	开启定时发送免费 ARP 的功能

配置代理ARP	 可选配置，用于代理应答请求不同子网内其他设备的 ARP 请求。
	ip proxy-arp 开启代理 ARP 功能。
配置本地代理ARP	 可选配置，用于代理应答请求子网内其他设备的 ARP 请求。
	local-proxy-arp 开启本地代理 ARP 功能。
配置 ARP 可信检测	 可选配置，用于发送单播 ARP 请求确认，以保证学习 ARP 表项正确性。
	arp trusted-monitor enable 开启 ARP 可信检测功能
配置ARP防IP报文攻击	 可选配置，防止 IP 报文大量送 CPU 对 CPU 造成冲击。
	arp anti-ip-attack 配置触发 ARP 设丢弃表项的 IP 报文个数。
配置抑制往认证VLAN发送ARP请求	 可选配置，用于抑制往认证 VLAN 发送 ARP 请求。
	arp suppress-auth-vlan-req 开启抑制往认证 VLAN 发 ARP 请求功能。

2.4.1 配置静态ARP

配置效果

用户手工指定 IP 地址和 MAC 地址的映射，防止设备学到错误的 ARP 表项而影响网络。

注意事项

对于三层交换机，配置完静态 ARP 表项后，交换机必须在学习到该静态 ARP 表项的 MAC 地址对应的物理端口后才能进行正常的三层路由。

配置方法

配置静态 ARP

- 可选配置
- 在汇聚设备上，可以通过静态绑定上联设备的 IP 和 MAC 地址的映射，防止设备因受到 ARP 攻击而更改掉上联设备的 ARP 表项的 MAC 地址，导致网络异常。
- 在全局模式下配置

检验方法

使用命令 **show running-config** 查看命令是否生效，或使用命令 **show arp static** 查看是否成功创建了静态 ARP 缓存表。

相关命令

配置静态 ARP

【命令格式】 **arp [oob] ip-address mac-address type**

【参数说明】 **oob**: 为 MGMT 口配置静态 ARP。

ip-address: 与 MAC 地址对应的 IP 地址, 分为四组十进制表示的数值, 组之间用点隔开。

mac-address: 数据链路层地址, 48 个比特位组成。

type: ARP 封装类型。对于以太网接口, 关键字为 arpa。

【命令模式】 全局模式

【使用指导】 RGOS 使用 ARP 缓冲表, 根据 32 个比特位 IP 地址查找 48 个比特位的 MAC 地址。

由于大多数主机支持动态 ARP 解析, 所以通常不需要配置静态 ARP 映射。利用 **clear arp-cache** 命令可以删除动态学习到的 ARP 映射。

配置举例

【网络环境】 网络拓扑如图 2-1 所示

【配置方法】 在设备 B 上配置静态 ARP 表项, 静态绑定设备 A 的 IP 和 MAC 地址映射。

```
Ruijie(config)#arp 192.168.23.1 00D0.F822.334B arpa
```

【检验方法】 通过 **show arp static** 命令可查看静态 ARP 表项:

```
Ruijie(config)#show arp static
Protocol Address      Age(min) Hardware      Type  Interface
Internet 192.168.23.1    <static> 00D0.F822.334B arpa
1 static arp entries exist.
```

常见配置错误

- 静态绑定的 MAC 地址错误。

2.4.2 配置ARP属性

配置效果

用户指定 ARP 表项的超时时间、ARP 请求重传次数和间隔、未解析 ARP 表项数上限、接口 ARP 学习数量限制。

注意事项

无

配置方法

配置 ARP 超时设置

- 可选配置
- 局域网中如果用户上下线较频繁，则可以将 ARP 超时时间设置小一点，可以将无效的 ARP 表项尽早删除。
- 在接口模式下配置

↘ ARP 请求重传时间间隔和次数

- 可选配置
- 在网络带宽资源不足时，可以将重传时间间隔配大，次数配小，以减少网络带宽的消耗。
- 在全局模式下配置

↘ 未解析 ARP 表项的数量限制

- 可选配置
- 在网络带宽资源不足时，可以将未解析 ARP 表项的数量配小，以减少网络带宽的消耗。
- 在全局模式下配置

↘ 接口 ARP 学习数量限制

- 可选配置
- 在接口模式下配置

检验方法

使用命令 **show arp timeout** 可以查看所有接口的老化超时时间。

使用命令 **show running-config** 查看 ARP 请求重传时间间隔和次数、未解析 ARP 表项是数量限制、接口 ARP 学习数量限制命令是否生效。

相关命令

↘ ARP 超时设置

- 【命令格式】 **arp timeout seconds**
- 【参数说明】 *seconds*：超时时间，以秒为计算单位，默认值为 3600，范围 0-2147483。
- 【命令模式】 接口模式
- 【使用指导】 ARP 超时设置只对动态学习到的 IP 地址和 MAC 地址映射起作用。超时时间设置得越短，ARP 缓存中保存的映射表就越真实，但是 ARP 消耗网络带宽也越多，所以需要权衡利弊。除非有特别的需要，否则一般不需要配置 ARP 超时时间。

↘ ARP 请求重传时间间隔和次数

- 【命令格式】 **arp retry interval seconds**
- 【参数说明】 *seconds*：<1-3600>,ARP 请求的重传时间可以设置为 1~3600 秒，默认值为 1 秒。
- 【配置模式】 全局模式

- 【使用指导】 当发现本设备有频繁的向外发送 ARP 请求，引起网络繁忙等其它问题时，可以将 ARP 请求的重传时间设置长一点，一般不要超过动态 ARP 表项的老化时间。

未解析 ARP 表项的数量限制

- 【命令格式】 **arp unresolve number**
- 【参数说明】 *number*：未解析 ARP 表项的最大个数，取值范围为 < 1-8192 >。默认值为 8192。
- 【配置模式】 全局模式
- 【使用指导】 当发现 ARP 缓存表中出现大量未解析表项，并且一段时间后还没有消失时，可以用此命令限制未解析表项的个数。

接口 ARP 学习数量限制

- 【命令格式】 **arp cache interface-limit limit**
- 【参数说明】 *limit*：指定接口所能学习的 ARP 数量最大限制，包括静态配置和动态学习的 ARP，取值范围为 0-设备支持的 ARP 表项容量，0 表示不限制接口 ARP 学习数量。
- 【配置模式】 接口模式模式
- 【使用指导】 限制接口的 ARP 学习数量，可防止恶意的 ARP 攻击，让设备生成大量的 ARP 表项，占用过多的表项资源。配置的值必须不小于当前接口已经学习到的 ARP 表项数量，否则配置不生效。该限制受限于设备支持的 ARP 容量。

配置举例

【网络环境】 网络拓扑如图 2-1 所示

- 【配置方法】
- 配置接口 GigabitEthernet 0/1 下的 ARP 超时时间为 60 秒
 - 配置接口 GigabitEthernet 0/1 下的 ARP 学习数量限制为 300
 - 配置 ARP 请求重传时间间隔为 3 秒
 - 配置 ARP 请求重传次数为 4 次
 - 配置未解析 ARP 表项数量限制为 4096

```
Ruijie(config)#interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#arp timeout 60
Ruijie(config-if-GigabitEthernet 0/1)#arp cache interface-limit 300
Ruijie(config-if-GigabitEthernet 0/1)#exit
Ruijie(config)#arp retry interval 3
Ruijie(config)#arp retry times 4
Ruijie(config)#arp unresolve 4096
```

- 【检验方法】
- 通过 **show arp timeout** 查看接口的老化时间
 - 通过 **show running-config** 查看 ARP 请求重传时间间隔和次数、未解析 ARP 表项是数量限制、接口 ARP 学习数量限制

```
Ruijie#show arp timeout
Interface                arp timeout(sec)
-----
GigabitEthernet 0/1     60
```

```
GigabitEthernet 0/2    3600
GigabitEthernet 0/4    3600
GigabitEthernet 0/5    3600
GigabitEthernet 0/7    3600
VLAN 100                3600
VLAN 111                3600
Mgmt 0                  3600

Ruijie(config)# show running-config
arp unresolve 4096
arp retry times 4
arp retry interval 3
!
interface GigabitEthernet 0/1
  arp cache interface-limit 300
```

常见配置错误

无

2.4.3 配置可信任ARP

配置效果

可以有效防止针对网关的 ARP 欺骗。

注意事项

可信任 ARP 只在交换机上支持。

配置方法

- 如果需要部署 GSN 方案，则应该执行此配置项。
- 部署 GSN 全局安全网络解决方案时，需要配置开启可信任 ARP 功能。
- 在全局模式下配置

检验方法

使用 **show arp trusted** 命令查看可信 ARP 表项；

使用 **show running** 命令查看可信任 ARP 的相关配置是否生效。

相关命令

▾ 启用可信任 ARP 功能

【命令格式】 **service trustedarp**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 设备的可信任 ARP 功能是一种防止 ARP 欺骗的功能，作为 GSN 方案的一部分 需要和 GSN 方案一起使用。

▾ 添加可信任 ARP 时进行 VLAN 转换

【命令格式】 **arp trusted user-vlan vid1 translated-vlan vid2**

【参数说明】 *vid1*：服务器设置的 VID

vid2：转换后的 VID

【配置模式】 全局模式

【使用指导】 要使此命令生效，首先启用可信任 ARP 功能。只有在服务器下发的 VLAN 和可信任 ARP 生效的 VLAN 不同时，才需要配置此命令。

▾ 查看交换机上的可信任 ARP

【命令格式】 **show arp trusted [ip [mask]]**

【参数说明】 *ip*：IP 地址，显示指定 IP 地址的 ARP 表项；如果指定 **trusted** 关键字，则只显示可信任 ARP 表项，否则显示非可信任 ARP 表项。

mask：显示 IP 子网内的 ARP 表项；如果指定 **trusted** 关键字，则只显示可信任 ARP 表项，否则显示非可信任 ARP 表项。

【配置模式】 特权模式

【使用指导】 -

▾ 删除交换机上的可信任 ARP

【命令格式】 **clear arp trusted [ip [mask]]**

【参数说明】 *ip*：IP 地址，显示指定 IP 地址的 ARP 表项；如果指定 **trusted** 关键字，则只显示可信任 ARP 表项，否则显示非可信任 ARP 表项。

mask：显示 IP 子网内的 ARP 表项；如果指定 **trusted** 关键字，则只显示可信任 ARP 表项，否则显示非可信任 ARP 表项。

【配置模式】 特权模式

【使用指导】 执行 **clear arp trusted** 会删除交换机上的所有的可信 ARP，可能导致用户不能上网。

一般情况下使用 **clear arp trusted ip** 删除指定的可信任 ARP 表项。

▾ 允许可信任 ARP 老化

【命令格式】 **arp trusted aging**

【参数说明】 -

【配置模式】 全局模式

- 【使用指导】 使用该命令后可信任 ARP 开始老化，老化时间和动态 ARP 老化时间相同。老化时间可以通过接口模式下 **arp timeout** 命令设置。

调整可信任 ARP 的容量

- 【命令格式】 **arp trusted number**
- 【参数说明】 *number*：取值范围最小为 10，最大为对应产品 arp 容量减去 1024，缺省可信 arp 的最大表项数为 arp 总容量的一半。
- 【配置模式】 特权模式
- 【使用指导】 要使此命令生效，首先启用可信任 ARP 功能。可信任 ARP 表项和其它表项共享内存，如果可信任表项占用过多，可能导致动态 ARP 表项空间不够。一般按需设置，不要设置得太大。

配置举例

【网络环境】 网络拓扑如图 2-1 所示

- 【配置方法】
- 开启可信任 ARP 功能
 - 配置 VLAN 转换
 - 配置可信任 ARP 表项运行老化
 - 配置可信任 ARP 表项的容量为 1024

```
Ruijie(config)#service trustedarp
Ruijie(config)#arp trusted user-vlan 2-9 translated-vlan 10
Ruijie(config)#arp trusted aging
Ruijie(config)#arp trusted 1024
```

- 【检验方法】
- 通过 **show running-config** 查看上面的配置是否生效

```
Ruijie(config)# show running-config
service trustedarp
arp trusted user-vlan 2-9 translated-vlan 10
arp trusted aging
arp trusted 1024
```

常见配置错误

- 可信任 ARP 功能未开启，导致 ARP 表项下发失败

2.4.4 配置免费ARP

配置效果

接口定时发送免费 ARP 报文。

注意事项

无

配置方法

- 可选配置
- 设备做用户网关时，为了防止因为 ARP 欺骗导致其他用户学习到错误的网关 MAC 后会一直上不了网，需要在接口上开启免费 ARP 功能。
- 在接口模式下配置

检验方法

使用 **show running-config interface <name>** 查看是否配置成功。

相关命令

▾ 开启定时发送免费 ARP 的功能

【命令格式】 **arp gratuitous-send interval seconds [number]**

【参数说明】 *seconds*：发送免费 ARP 请求的时间间隔，单位秒，取值范围<1-3600>。

number：发送免费 ARP 请求的数量，缺省值是 1，取值范围<1-100>。

【命令模式】 接口模式

【使用指导】 当设备的网络接口作为下联设备的网关时，如果下联设备中有冒充网关的行为，则可以在此接口配置定时发送免费 ARP 请求，公告自己才是真正的网关。

配置举例

【网络环境】 网络拓扑如图 2-1 所示

【配置方法】 配置 GigabitEthernet 0/0 口发送免费 ARP 功能，频率为每 5 秒发送一个免费 ARP 请求报文。

```
Ruijie(config-if-GigabitEthernet 0/0)#arp gratuitous-send interval 5
```

【检验方法】 使用 **show running-config interface** 命令查看配置是否生效

```
Ruijie#sh running-config interface gigabitEthernet 0/0
```

```
Building configuration...
```

```
Current configuration : 127 bytes
```

```
!
```

```
interface GigabitEthernet 0/0
```

```
duplex auto
```

```
speed auto
```

```
ip address 30.1.1.1 255.255.255.0
```

```
arp gratuitous-send interval 5
```

常见配置错误

无

2.4.5 配置代理ARP

配置效果

设备代理应答非本机的 ARP 请求报文。

注意事项

三层交换机缺省关闭代理 ARP 功能，路由器缺省开启代理 ARP 功能。

配置方法

- 可选配置。
- 没有路由信息的主机需要获得其它子网 IP 地址的 MAC 地址，设备需要开启代理 ARP 功能，代理应答 ARP。
- 在接口模式下配置

检验方法

使用 **show ip interface <name>**命令查看是否配置成功。

相关命令

▾ 开启代理 ARP 功能

- 【命令格式】 **ip proxy-arp**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 -

配置举例

- 【网络环境】 网络拓扑如图 2-1 所示
- 【配置方法】 配置 GigabitEthernet 0/0 口开启代理 ARP 功能

```
Ruijie(config-if-GigabitEthernet 0/0)#ip proxy-arp
```
- 【检验方法】 使用 **show ip interface** 命令查看是否配置成功

```
Ruijie#show ip interface gigabitEthernet 0/0
GigabitEthernet 0/0
  IP interface state is: DOWN
  IP interface type is: BROADCAST
  IP interface MTU is: 1500
  IP address is:
    No address configured
  IP address negotiate is: OFF
  Forward direct-broadcast is: OFF
  ICMP mask reply is: ON
  Send ICMP redirect is: ON
  Send ICMP unreachable is: ON
  DHCP relay is: OFF
  Fast switch is: ON
  Help address is: 0.0.0.0
  Proxy ARP is: ON
ARP packet input number: 0
  Request packet      : 0
  Reply packet       : 0
  Unknown packet     : 0
TTL invalid packet number: 0
ICMP packet input number: 0
  Echo request       : 0
  Echo reply         : 0
  Unreachable        : 0
  Source quench      : 0
  Routing redirect   : 0
```

常见配置错误

无

2.4.6 配置本地代理ARP

配置效果

设备代理应答子网内非本机的 ARP 请求报文。

注意事项

只在 SVI 口下支持。

配置方法

- 可选配置。
- 在开启端口保护时，如果有需要 VLAN 内的主机通信，则需要配置本地 ARP 代理。
- 在接口模式下配置

检验方法

使用 **show run interface <name>**命令查看是否配置成功。

相关命令

▾ 开启本地代理 ARP 功能

- 【命令格式】 **local-proxy-arp**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 -

配置举例

【网络环境】 网络拓扑如图 2-1 所示

【配置方法】 配置 VLAN 1 口开启代理 ARP 功能

```
Ruijie(config-if-VLAN 1)#local-proxy-arp
```

【检验方法】 使用 **show ip interface** 命令查看是否配置成功

```
Ruijie#show running-config interface vlan 1
```

```
Building configuration...
```

```
Current configuration : 53 bytes
```

```
interface VLAN 1
 ip address 192.168.1.2 255.255.255.0
 local-proxy-arp
```

常见配置错误

无

2.4.7 配置ARP可信检测

配置效果

开启 arp 可信检测功能，在收到 arp 请求报文后，如果对应表项不存在，进入 NUD（邻居不可达探测）。如果已有的动态 arp 表项的 MAC 地址被更新，马上走 NUD 探测逻辑，起到防止 arp 攻击的作用。

注意事项

该功能由于在 ARP 学习过程中增加了一个严格确认的过程，所以开启该功能会影响到 ARP 的学习性能。

配置方法

- 可选配置。
- 如果有要求严格学习 ARP 表项的需求时，设备上可以开启 arp 可信功能，设备在收到 arp 请求报文后，如果之前不存在对应 arp 表项，则需要发送单播 ARP 请求报文，在确认对端真实存在后才学习 ARP 表项，否则不学习 ARP 表项。在 arp 表项的 mac 地址发生了变化后，马上走 NUD 探测，防止 arp 欺骗。
- 在接口模式下配置

检验方法

使用 `show running-config interface <name>` 查看是否配置成功。

相关命令

📌 开启 ARP 可信检测功能

【命令格式】 `arp trust-monitor enable`

【参数说明】 -

【命令模式】 接口模式

【使用指导】

- ❗ 开启该功能，如果对应 arp 表项已经存在，且 mac 地址没发生更新，则不进行 NUD 探测逻辑。
- ❗ 开启该功能，如果已有的动态 arp 表项的 mac 地址被更新，则马上走 NUD 探测逻辑。
- ❗ 关闭该功能后，arp 表项的学习和更新不需要 NUD 过程。

配置举例

【网络环境】 网络拓扑如图 2-1 所示

【配置方法】 配置 GigabitEthernet 0/0 口开启 ARP 可信检测功能

```
Ruijie(config-if-GigabitEthernet 0/0)#arp trust-monitor enable
```

【检验方法】 使用 **show running-config interface** 查看是否配置是否生效

```
Ruijie#show running-config interface gigabitEthernet 0/0
```

```
Building configuration...
Current configuration : 184 bytes
!
interface GigabitEthernet 0/0
 duplex auto
 speed auto
 ip address 30.1.1.1 255.255.255.0
 arp trust-monitor enable
```

常见配置错误

无

2.4.8 配置 ARP防IP报文攻击

配置效果

交换机 CPU 收到配置个数的目的 IP 命中该 ARP 表项的报文时，后续所有该目的 IP 的报文都不会送 CPU。

注意事项

只在交换机产品上支持。

配置方法

- 可选配置。
- 在交换机产品上，默认情况下，在 3 个未知名单播报文送 CPU 后设置丢弃表项。通过过此命令用户可以针对具体网络环境调整这个参数，也可以关闭该功能。
- 在全局模式下配置。

检验方法

使用 **show run** 命令查看是否配置成功。

相关命令

配置 ARP 防 IP 报文攻击

【命令格式】 **arp anti-ip-attack num**

【参数说明】 设置触发 ARP 丢弃表项的 IP 报文个数，取值范围<0-100>。
0 表示关闭 ARP 防 IP 报文攻击功能。缺省值为 3。

【命令模式】 接口模式

【使用指导】  如果硬件路由资源比较充分，**arp anti-ip-attack num** 可以设置得小一些。在硬件路由资源比较紧张的情况下，要优先满足正常路由的使用，可以将 **arp anti-ip-attack num** 设置得比较大，或者关闭该功能。

配置举例

【网络环境】 网络拓扑如图 2-1 所示

【配置方法】 在设备B上配置 ARP防IP报文攻击。

```
Ruijie(config)#arp anti-ip-attack 10
```

【检验方法】 使用 **show running-config** 查看配置是否生效

```
Ruijie#show running-config
```

```
Building configuration...
```

```
Current configuration : 53 bytes
```

```
arp anti-ip-attack 10
```

常见配置错误

无

2.4.9 配置抑制往认证VLAN发送ARP请求

配置效果

设备不往认证 VLAN 发送 ARP 请求报文。

注意事项

只在 SVI 口下支持。

配置方法

- 可选配置。

- 在开启网关认证模式下，设备默认不往认证 VLAN 发送 ARP 请求报文。若需要往认证 VLAN 发送 ARP 请求，使用该命令的 no 形式取消该功能。
- 在接口模式下配置

检验方法

使用 **show run interface <name>**命令查看是否配置成功。

相关命令

▾ 抑制往认证 VLAN 发送 ARP 请求

- 【命令格式】 **arp suppress-auth-vlan-req**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 -

配置举例

【网络环境】 网络拓扑如图 2-1 所示

【配置方法】 配置 VLAN 2 口关闭抑制往认证 VLAN 发送 ARP 请求的功能。

```
Ruijie(config-if-VLAN 2)#no arp suppress-auth-vlan-req
```

【检验方法】 使用 **show running-config interface <name>**查看配置是否生效

```
Ruijie#show running-config interface vlan 2
```

```
Building configuration...
```

```
Current configuration : 53 bytes
```

```
interface VLAN 2
```

```
ip address 192.168.1.2 255.255.255.0
```

```
no arp suppress-auth-vlan-req
```

常见配置错误

无

2.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除动态 ARP 表项。在网关认证模式下，不会删除认证 VLAN 下的动态 ARP 表项。	clear arp-cache

查看运行情况

作用	命令
显示 ARP 表。	show arp [detail] [interface-type interface-number [ip [mask] mac-address static complete incomplete] subvlan {subvlan-number min-max min_value max_value}]
显示 ARP 表	show ip arp
显示可信任 ARP 表	show arp [detail] trusted [ip [mask]]
显示 ARP 表项相应计数	show arp counter
显示动态 ARP 表项的老化时间	show arp timeout

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
显示 ARP 报文的收发情况	debug arp
显示 ARP 表项的创建删除情况	debug arp event

3 DHCP

3.1 概述

DHCP (Dynamic Host Configuration Protocol , 动态主机设置协议) 是一个 局域网的 网络协议, 使用UDP协议工作, 被广泛用来动态分配可重用的网络资源, 如IP地址。

DHCP 是基于 Client/Server 工作模式, DHCP 客户端通过发送请求消息向 DHCP 服务器获取 IP 地址, 等其他配置信息。当 DHCP 客户端与服务器不在同一个子网上, 必须有 DHCP 中继代理 (DHCP Relay) 来转发 DHCP 请求和应答消息。

协议规范

- RFC2131 : Dynamic Host Configuration Protocol
- RFC2132 : DHCP Options and BOOTP Vendor Extensions
- RFC3046 : DHCP Relay Agent Information Option

3.2 典型应用

典型应用	场景描述
在局域网内提供DHCP服务	为局域网内下游用户分配地址。
设备启动DHCP Client功能	局域网内下游多设备启动 DHCP Client 功能。
MPLS环境下的DHCP Server典型应用	L2vpn/L3vpn 环境下 DHCP Server 的应用。
AM规则在DHCP-server中的典型应用	Supervlan 场景下 DHCP Server 的应用。
有线场景中DHCP Relay典型应用	有线场景中跨网段的用户申请 IP 上网。
AM规则在DHCP Relay中的典型应用	Supervlan 场景中跨网段的用户申请 IP 上网。

3.2.1 在局域网内提供DHCP服务

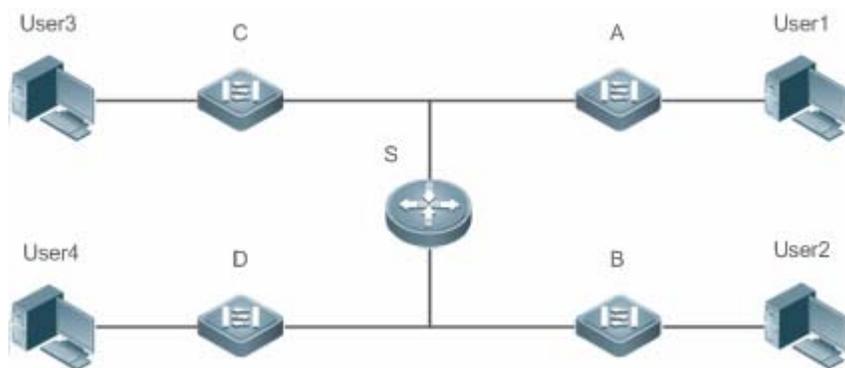
应用场景

在一个局域网内, 为四个用户分配 IP 地址。

以下图为例, 为 User1、User2、User3 、User4 分配 IP 地址。

- User1、User2、User3 、User4 通过 A、B、C、D 与 Server 相连

图 3-1



【注释】 S为出口网关设备，作 DHCP-Server。
 A、B、C、D 为接入交换机，作二层透传
 User1、User2、User3 、User4 为用户

功能部属

- Server(S)上运行 DHCP-Server 服务
- 在 A、B、C、D 上实行二层 VLAN 透传功能
- User1、User2、User3 、User4 上主动发起 DHCP-Client 请求

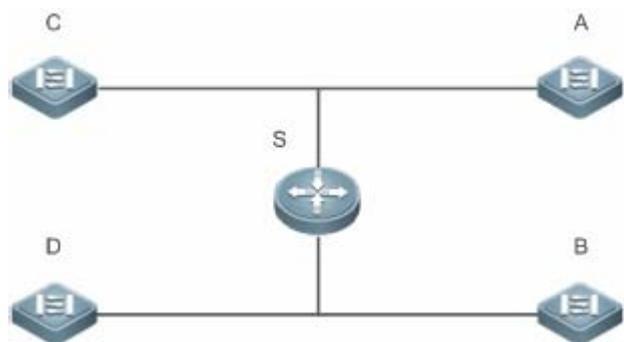
3.2.2 设备启动DHCP Client功能

应用场景

在一个局域网内，A、B、C、D 四个接入设备向 S 请求地址

以下图为例，A、B、C、D 接口上开启 DHCP-Client 功能，请求 IP 地址。

图 3-2



- 【注释】 S 为出口网关设备，作 DHCP-Server。
A、B、C、D 为接入交换机，接口启动 DHCP-Client 功能

功能部属

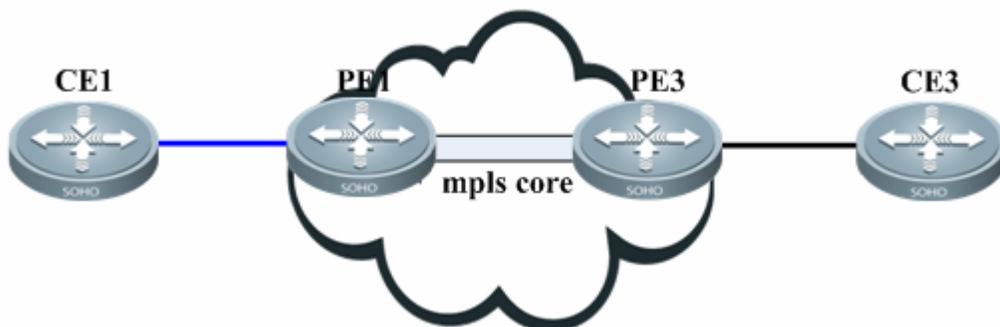
- Server(S)上运行 DHCP-Server 服务
- 在 A、B、C、D 在接口上开启 DHCP-Client 功能

3.2.3 MPLS环境下的DHCP Server典型应用

应用场景

如下图所示，如果在 L3VPN 环境下，PE1 为 DHCP-Server，CE1 为 DHCP-Client，这种场景下无影响；如果在 L2VPN 环境下，PE1 为 DHCP-Server，CE1 为 DHCP-Client，这种场景下有影响，PE1 的 AC 侧接口上无法配置 IP 地址，导致 DHCP 报文无法到达 Server 模块

图 3-3 MPLS 组网拓扑图(L2VPN/L3VPN)



- 【注释】 在一个 MPLS core 域内，PE1，PE3 组成一个 VPLS 网络。CE1，CE3 分别被 PE1，PE3 接入到同一个 VPLS，CE1 多归(Multi-Homing)到 PE1 和 PE3 并且接入同一个 VPLS 域。PE1 上启动 DHCP-Server

功能部属

- 配置 PE1、PE3 为 L3VPN 的场景，并在 PE1 启动 DHCP-Server 功能。
- CE1 上启动 DHCP-Client。

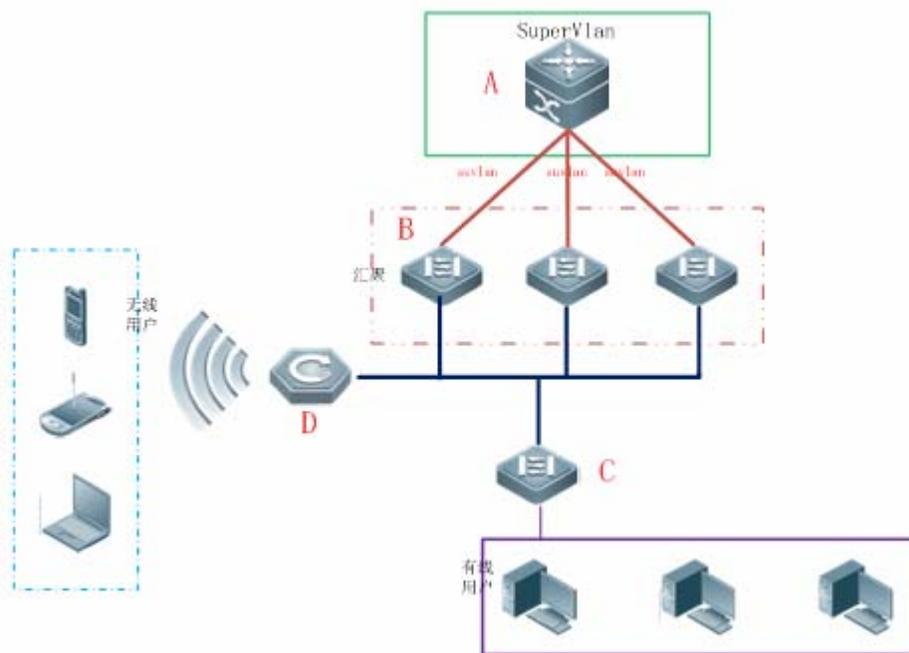
3.2.4 AM规则在DHCP-Server中的典型应用

应用场景

如下图所示，设备 A 作为核心交换设备，配置 Supervlan 场景、AM 过滤规则及启动 Dhcp-Server，B 作为汇聚交换设备层，C 用作接入交换设备，D 作为无线接入交换设备。主要需求如下：

- 基于 vlan+port 的 AM 规则进行动态地址分配
- 基于 vlan 的 AM 规则进行静态地址分配
- 基于缺省 AM 规则进行动态地址分配

图 3-4 AM 规则在 DHCP-Server 中的组网拓扑图



- 【注释】
- A 作为核心设备。
 - B 作为汇聚设备。
 - C 作为有线接入设备。
 - D 作为无线接入设备。

功能部署

- 在 A 上配置 AM 规则、启动 Dhcp-Server 服务、创建 Supervlan。
- 在 B、C 上创建 Vlan，对有线用户 DHCP 报文透传至设备 A，进行地址获取。
- 在 D 上启动无线功能，将无线用户 DHCP 报文透传至设备 A，进行地址获取。

3.2.5 有线场景中DHCP Relay典型应用

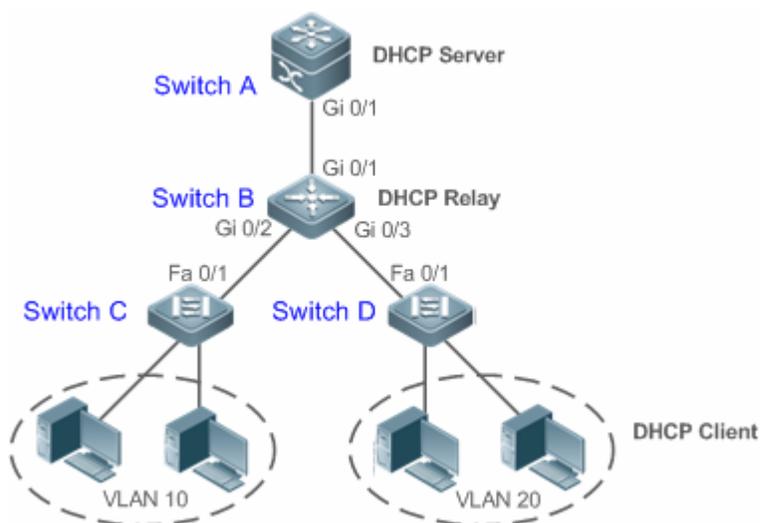
应用场景

如下图所示，Switch C 和 Switch D 作为接入设备，分布着 VLAN 10 和 VLAN 20 的 PC 用户，Switch B 作为网关设备，Switch A 作为核心设备。主要需求如下：

Switch A 可以充当 DHCP Server，为不同 VLAN 用户动态分配不同网段的 IP 地址。

Switch C 和 Switch D 下的接入用户可以跨网段动态获取 IP 地址。。

图 3-5DHCP Relay 组网拓扑图



- 【注释】
- Switch C 与 Switch D 作为接入设备。
 - Switch B 作为网关设备。
 - Switch A 作为核心设备。

功能部属

- 配置 Switch B 和 Switch C、D 之间的二层通信。
- 在 Switch B 上，指定 DHCP 服务器地址，并开启 DHCP Relay 功能。
- 在 Switch A 上，分别为 VLAN 10 和 VLAN 20 的用户创建 DHCP 地址池，开启 DHCP Server 功能。

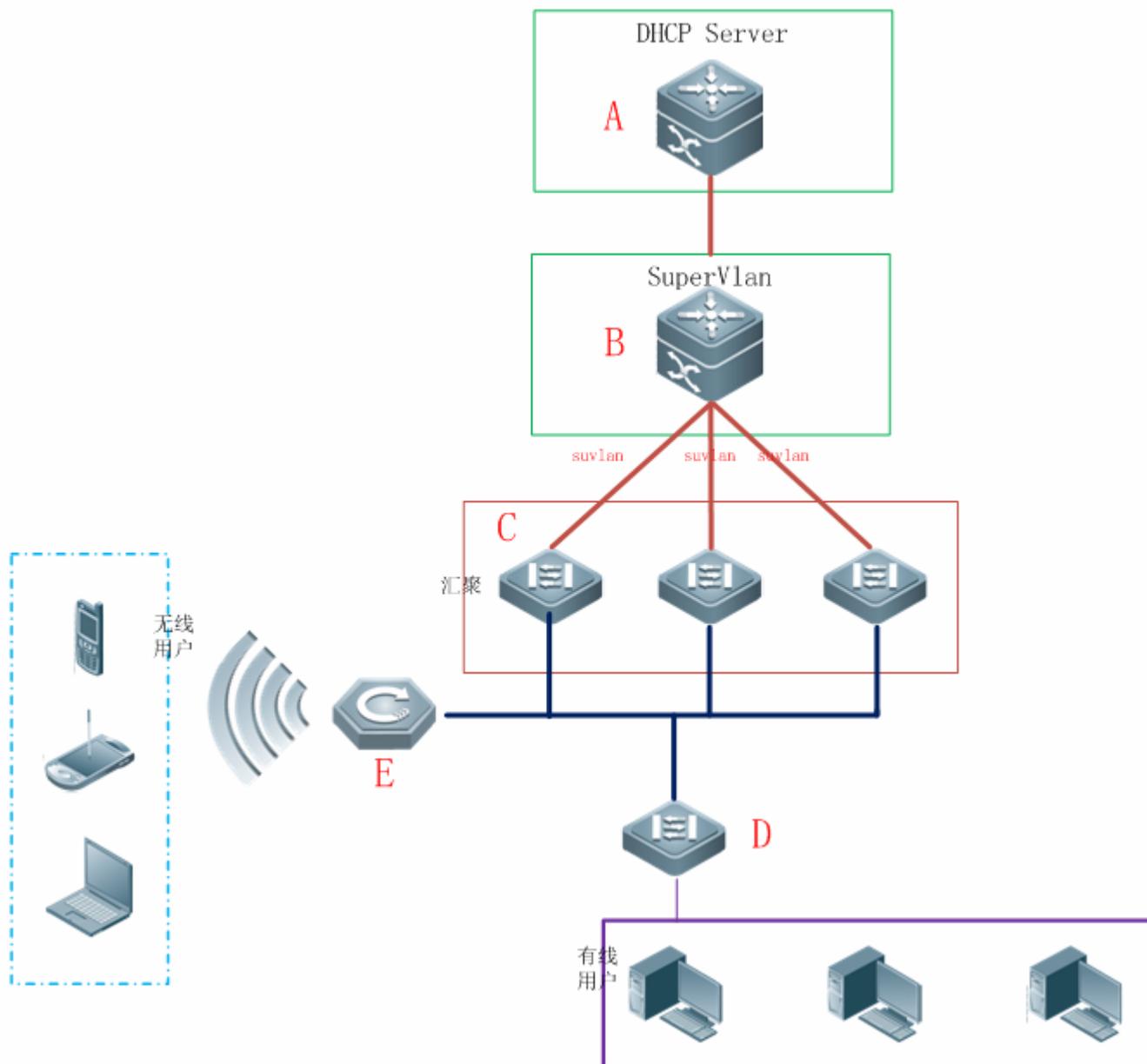
3.2.6 AM规则在DHCP Relay中的典型应用

应用场景

如下图所示,设备A作为DHCP Server设备,设备B作为核心交换设备,配置Supervlan场景、AM过滤规则及启动Dhcp Relay, C作为汇聚交换设备层,D用作接入交换设备,E作为无线接入交换设备。主要需求如下:

- 基于vlan+port的AM规则进行选择客户端子网作为中继报文Giaddress,并转发报文给DHCP Server分配对应客户端子网的地址
- 基于缺省AM规则进行选择客户端子网作为中继报文Giaddress,并转发报文给DHCP Server分配对应客户端子网的地址

图 3-6 AM 规则在 DHCP Relay 中的组网拓扑图



【注释】 A作为核心设备。

- B 作为核心设备。
- C 作为汇聚设备。
- D 作为有线接入设备。
- E 作为无线接入设备。

功能部署

- 在 A 上启动 Dhcp Server 服务。
- 在 B 上配置 AM 规则、启动 Dhcp Relay 服务、创建 Supervlan。
- 在 C、D 上创建 Vlan，对有线用户 DHCP 报文透传至设备 B，进行地址获取。
- 在 E 上启动无线功能，将无线用户 DHCP 报文透传至设备 B，进行地址获取。

3.3 功能详解

基本概念

↘ DHCP 服务器

锐捷产品的 DHCP 服务器完全根据 RFC 2131 来实现的，主要功能就是为主机分配和管理 IP 地址。

↘ DHCP 客户端

DHCP 客户端可以让设备自动地从 DHCP 服务器获得 IP 地址以及其它配置参数。

↘ DHCP 中继

当 DHCP 客户端与服务器不在同一个子网上，就必须有 DHCP 中继代理来转发 DHCP 请求和应答消息。

↘ 租约

租约是客户机可使用指派的 IP 地址期间 DHCP 服务器指定的时间长度。租用给客户时，租约是活动的。在租约过期之前，客户机一般需要通过服务器更新其地址租约指派。当租约期满或在服务器上删除时，租约是非活动的。租约期限决定租约何时期满以及客户需要用服务器更新它的次数。

↘ 排除地址

排除地址是指从 DHCP 服务器中排除指定的一些 IP 地址序列，排除地址作用是为了确保在这些地址都不会是由 DHCP 服务器提供给 DHCP 客户机。

↘ 地址池

地址池是指 DHCP 服务器可分配给用户的地址集合，所有分配给用户的地址都从管理员配置的池中取出的。

↘ 选项类型

选项类型是 DHCP 服务器在向 DHCP 客户机提供租约服务时指派的配置参数。例如,某些公用选项包括默认网关(路由器)、WINS 服务器和 DNS 服务器的 IP 地址。DHCP-Server 还允许配置其它选项。虽然大多数选项都是在 RFC 2132 中预定义的,但若需要的话,可添加自定义选项类型。

功能特性

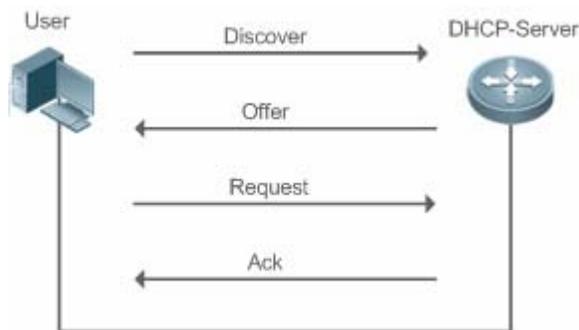
功能特性	作用
DHCP服务器	设备启用 DHCP Server 功能,可以为主机动态分配 IP 地址和提供主机配置参数。
DHCP中继	设备启用 DHCP Relayr 功能,可以在不同网段之间转发 DHCP 请求和应答消息。
DHCP客户端	设备启用 DHCP Client 功能,可以自动从 DHCP 服务器获取 IP 地址以及其它配置参数。

3.3.1 DHCP服务器

工作原理

▾ DHCP 工作的基本流程

图 3-7



DHCP 请求 IP 地址的过程如下：

3. 主机发送 DHCPDISCOVER 广播包在网络上寻找 DHCP 服务器；
4. DHCP 服务器向主机发送 DHCPOFFER 单播/广播(依据主机报文相关属性确定)数据包,包含 IP 地址、MAC 地址、域名信息以及地址租期；
5. 主机发送 DHCPREQUEST 广播包,正式向服务器请求分配已提供的 IP 地址；
6. DHCP 服务器向主机发送 DHCPACK 单播包,确认主机的请求。

i DHCP 客户端可以接收到多个 DHCP 服务器的 DHCPOFFER 数据包,然后可能接受任何一个 DHCPOFFER 数据包,但客户端通常只接受收到的第一个 DHCPOFFER 数据包。另外, DHCP 服务器 DHCPOFFER 中指定的地址不一定为最终分配的地址,通常情况下, DHCP 服务器会保留该地址直到客户端发出正式请求。

正式请求 DHCP 服务器分配地址 DHCPREQUEST 采用广播包,是为了让其它所有发送 DHCPOFFER 数据包的 DHCP 服务器也能够接收到该数据包,然后释放已经 OFFER (预分配)给客户端的 IP 地址。

如果发送给 DHCP 客户端的 DHCP OFFER 信息包中包含无效的配置参数，客户端会向服务器发送 DHCP DECLINE 信息包拒绝接受已经分配的配置信息。

在协商过程中，如果 DHCP 客户端没有及时响应 DHCP OFFER 信息包，DHCP 服务器会发送 DHCP NAK 消息给 DHCP 客户端，导致客户端重新发起地址请求过程。

在网络建设中，应用锐捷产品 DHCP 服务器，可以带来以下好处：

- 降低网络接入成本。一般采用静态地址分配的接入费用比较昂贵，应用动态地址分配的接入成本较低。
- 简化配置任务，降低网络建设成本。采用动态地址分配，大大简化了设备配置，对于在没有专业技术人员的地方部署设备，更是降低了部署成本。
- 集中化管理。在对多个子网进行配置管理时，有任何配置参数的变动，只需要修改和更新 DHCP 服务器的配置即可。

▾ 地址池

Server 收到来自 Client 请求报文，首先选择出一个合法有效地址池，并在该池中通过 PING 机制确认一个可用的地址，接着下发该池相关配置信息与地址至客户端，同时本地保存该租约信息在，以供该客户续租时检查有效性使用；由此完成整个租约分配流程。

地址池中可以带有各种配置参数，以下列举几个常用的：

- 地址池范围，可以分配给用户的地址范围
- 网关地址，通告用户网关地址，最多可以有八个
- DNS 地址，通告用户 DNS 地址，最多可以有八个
- 租约周期，通告用户地址何时老化，用户何时该请求续租

▾ VRRP 监控功能

在 VRRP (Virtual Router Redundancy Protocol，虚拟路由冗余协议)应用场景下，DHCP 提供配置命令来决定是否监控当前接口的 VRRP 状态。对于配置了 VRRP 地址的接口，当配置监控 VRRP 状态后，DHCP 服务器仅对处于 Master 状态的接口上来的 DHCP 客户端请求报文进行处理，处于备份(Backup)状态的接口请求报文将被丢弃。而对于没配置 VRRP 地址的接口，DHCP 服务器不再监控 VRRP 状态，所有 DHCP 请求报文都会得到处理。VRRP 监控命令只能在三层口上配置，默认情况下 VRRP 监控功能关闭，即只有主机处理 DHCP 业务备机不处理。

▾ 基于 vlan+端口+ip-range 地址分配功能

在部署地址池的环境下，为每个 vlan+端口号来分配指定 ip-range 的地址功能(在满足正常动态地址分配逻辑后，才能从本配置中选择有效地址)。主要有三种应用场景：1.只有全局默认配置；2.只有基于 vlan+端口+ip-range 的配置；3.上述两种配置均有；场景 1 有全局配置，默认分配全局配置的区间地址；场景 2 来自指定 vlan+端口的用户分配指定区间的地址，其余则用户无地址分配；场景 3 满足场景 2 的分配指定区间地址，其余用户分配全局默认配置地址。

相关配置

▾ 全局启动 DHCP-Server 服务

- 缺省情况下，该服务关闭。

- 全局使用 **service dhcp** 开启该服务。
- 必须在全局使用 **service dhcp** 功能，才能进行 DHCP 服务。

配置地址池

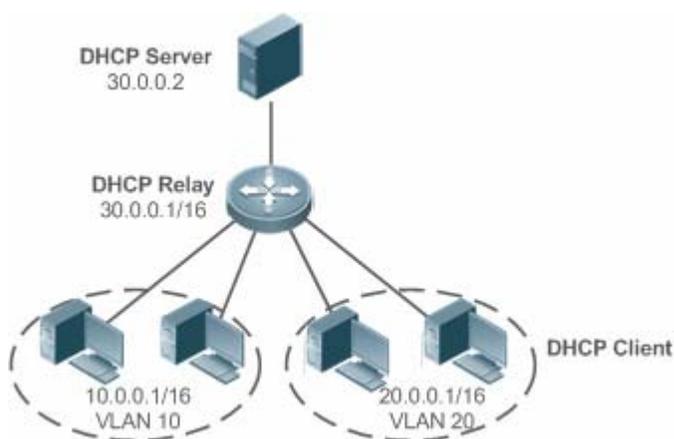
- 缺省情况下，无地址池。
- 使用 **ip dhcp pool** 命令可以进入到地址池配置模式，进行地址范围、网关地址、DNS 等信息配置。
- 不配置地址池范围将无地址可分配，无法下发任何地址。

3.3.2 DHCP中继代理

工作原理

DHCP 请求报文的目的 IP 地址为 255.255.255.255，这种类型报文的转发局限于子网内。为了实现跨网段的动态 IP 地址分配，DHCP 中继就产生了。DHCP 中继将收到的 DHCP 请求报文以单播方式转发给 DHCP 服务器，同时将收到的 DHCP 响应报文转发给 DHCP 客户端。DHCP 中继相当于一个转发站，负责沟通位于不同网段的 DHCP 客户端和 DHCP 服务器，即转发客户端 DHCP 请求报文、转发服务端 DHCP 应答报文。这样就实现了只要安装一个 DHCP 服务器，就可以实现对多个网段的动态 IP 管理，即 Client—Relay—Server 模式的 DHCP 动态 IP 管理。如图所示：

图 3-8DHCP Relay 应用场景



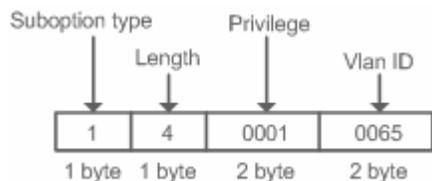
VLAN 10 和 VLAN 20 分别对应 10.0.0.1/16 和 20.0.0.1/16 的网络，而 DHCP 服务器在 30.0.0.1/16 的网络上，30.0.0.2 的 DHCP 服务器要对 10.0.0.1/16 和 20.0.0.1/16 的网络进行动态 IP 管理，只要在作为网关的设备上打开 DHCP 中继功能，并配置 30.0.0.2 为 DHCP 服务器的 IP 地址。

配置 DHCP Relay Agent Information(option 82)

根据 RFC3046 的定义，中继设备进行 DHCP Relay 时，可以通过添加 option 的方式来详细的标明 DHCP 客户端的一些网络信息，从而使服务器可以根据更精确的信息给用户分配不同权限的 IP，根据 RFC3046 的定义，所使用 option 选项的选项号为 82，故也被称作 option 82。锐捷实现的 Relay agent information 目前存在三种应用方案，下面分别对三种应用方案进行说明：

1. Relay agent information option dot1x：此种应用方案需要结合 802.1x认证以及锐捷产品RG-SAM。DHCP中继根据 RG-SAM在 802.1x认证过程中下发的IP权限，以及DHCP客户端所属vid，组合构成Circuit ID子选项。选项格式如 图 3-9 所示：

图 3-9 选项格式



2. Relay agent information option82：此种option的应用不需要结合其他协议模块的运行。DHCP中继根据接收DHCP请求报文的实体端口，以及设备自身的物理地址信息，组合构成option82 选项。选项格式如 图 3-10、图 3-11 所示：

图 3-10 Agent Circuit ID

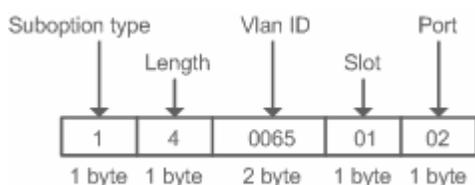
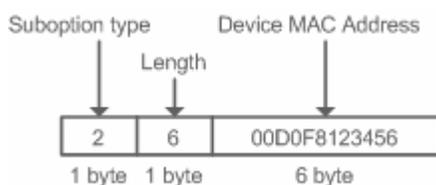
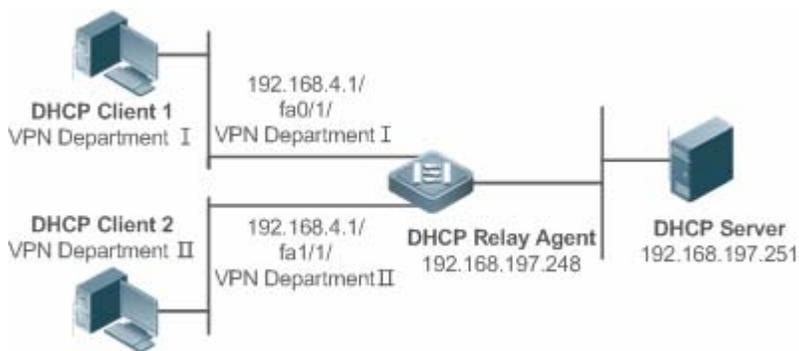


图 3-11 Agent Remote ID



3. Relay agent information option vpn：此种 option 的应用需要结合 MPLS VPN 相关功能。

图 3-12 MPLS VPN 环境中应用

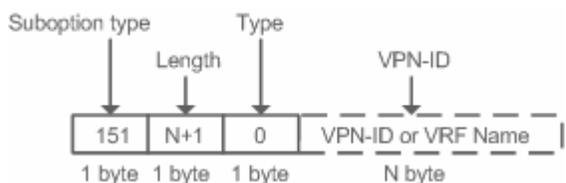


如图 3-12 所示，在MPLS VPN环境中，DHCP Client 1 和DHCP中继上的fa0/1 口相连，DHCP Client 2 和DHCP中继上的fa1/1 相连，接口fa0/1 和接口fa1/1 分别属于不同的VRF，DHCP Client 1 和DHCP Client 2 通过DHCP获取地址。按照网络规划，VPN DepartmentI和VPN DepartmentII使用重叠网段 192.168.4.0/24，在该应用环境下，传统的DHCP应用根本无法支持该部

署。为了实现在MPLS VPN环境下对DHCP中继的支持，在DHCP中继中引入了option vpn选项，该选项包括VPN-ID、Subnet-Selection以及Server-Identifier-Override三个子选项，简单说明一下这三个子选项的意义：

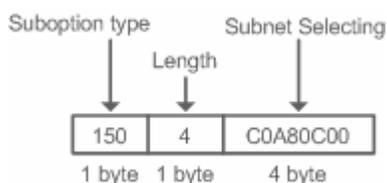
- VPN-ID：在接收到 DHCP 请求报文时，将 DHCP 客户端所属的 VPN 信息，以选项形式加入 DHCP 请求报文中。DHCP 服务器发送响应报文时，将该选项信息原样保留，DHCP 中继根据该选项，将 DHCP 响应报文转发到正确的 VRF 中。选项格式如 2 所示：

图 3-13 VPN-ID



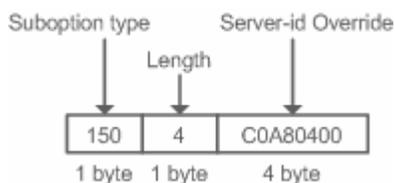
- Subnet-Selection：在传统的 DHCP 中继环境中，通过 gateway address[giaddr]字段表示客户端所在的网络信息以及 DHCP 服务器与 DHCP 中继的通讯地址。在 MPLS VPN 环境中，将 giaddr 修改为 DHCP 中继连接 DHCP 服务器的接口 IP，使 DHCP 服务器可以与 DHCP 中继直接通讯。但是客户端的子网信息必须通过新的选项 Subnet-Selection 来表示。选项格式如 3 所示：

图 3-14 Subnet-Selection



- Server-Identifier-Override：在 MPLS VPN 环境下，DHCP 客户端后续的请求报文都无法直接发送到 DHCP 服务器。DHCP 中继使用该选项携带 DHCP 中继与 DHCP 客户端直连的接口地址信息，DHCP 服务器发送响应报文的时候，用该选项覆盖 Server-identifier 选项信息。从而使 DHCP 客户端在与 DHCP 服务器交互的过程中，能够将报文送往 DHCP 中继，然后由 DHCP 中继将报文转发到 DHCP 服务器。选项格式如 4 所示：

图 3-15 Server-Identifier-Override



▾ DHCP Relay Check Server-id 功能

在 DHCP 应用环境中，通常会为每一个网络配备多个 DHCP 服务器，从而进行备份，防止因为一台服务器的工作不正常影响网络的正常使用。在 DHCP 获取的四个交互过程中，当 DHCP 客户端在发送 DHCP REQUEST 时已经选定了服务器，此时会在请求的报文中携带一个 server-id 的 option 选项，在某些特定的应用环境中为了减轻网络服务器压力，需要我们 Relay 能够启用此选项，只把请求报文发给此选项里的 DHCP 服务器，而不是发送给每一个配置的 DHCP 服务器，上述就是 DHCP Relay check server-id 功能。

📌 DHCP Relay suppression 功能

在指定接口上配置命令 `ip DHCP Relay suppression` 后，将屏蔽该接口上收到的 DHCP 请求报文；而对于其他接口上收到的 DHCP 请求报文，则正常转发。

相关配置

📌 启动设备上的 DHCP Relay 功能

- 缺省情况下，设备上的 DHCP Relay 功能关闭。
- 使用 `service dhcp` 命令可以启动设备上的 DHCP Relay 功能。
- 必须在设备上启用 DHCP Relay 功能，DHCP Relay 才能正常工作。

📌 配置 DHCP 服务器的 IP 地址

- 缺省情况下，无 DHCP 服务器的 IP 地址表项。
- 使用 `ip helper-address` 命令可以添加 DHCP 服务器地址表项，DHCP 服务器地址可以全局配置，也可以在三层接口上配置。全局或者每个三层接口上最多可以配置 20 个 DHCP 服务器地址。
- 在接口上收到 DHCP 请求报文时，首先使用接口上的 DHCP 服务器列表；如果接口上面没有配置 DHCP 服务器列表，则使用全局配置的 DHCP 服务器列表。

📌 启动 DHCP option 82 功能

- 缺省情况下，设备上的 DHCP option 82 功能关闭。
- 使用 `ip dhcp relay information option82` 命令可以启动设备上的 DHCP option 82 功能。

📌 启动 DHCP Relay check server-id 功能

- 缺省情况下，设备上的 DHCP Relay check server-id 功能关闭。
- 使用 `ip dhcp relay check server-id` 命令可以启动设备上的 DHCP Relay check server-id 功能。

📌 启动 DHCP Relay suppression 功能

- 缺省情况下，所有接口上 DHCP Relay suppression 功能关闭。
- 使用 `ip dhcp relay suppression` 命令可以启动对应接口上的 DHCP Relay suppression 功能。

3.3.3 DHCP客户端

工作原理

Client 状态机进入 Init 状态，主动发出广播 Discover 报文，之后 Client 有可能收到多份 Offer，进入 Offer 选择阶段选择一份最优的 Offer 后给予该服务器响应，此后在地址的老化 1/2、4/5 周期内还会发出续租等报文请求对地址的继续使用。

相关配置

接口上启动 DHCP-Client 功能

- 缺省情况下，该服务关闭。
- 接口模式下使用 `ip address dhcp` 开启功能。
- 必须开启客户端功能，才能进行 DHCP 服务。
- 该功能只在三层接口上有效，如 SVI、Router Port 等；

3.3.4 AM规则

工作原理

AM 规则用于规划不同 vlan + port/vlan 上来的 DHCP 客户端请求的 IP 范围，可快速定位出问题的 DHCP 客户端所属的 vlan + port/vlan，也可以更有效地分配地址池的地址。使用 AM 规则后，所有来自配置 vlan + port/vlan 的 DHCP 客户端能够正常获得地址；反之，若 DHCP 客户端来源未配置 vlan + port/vlan 时：如果配置了缺省 AM 规则，DHCP 客户端将获得缺省区间中的地址，如果未配置缺省 AM 规则，DHCP 客户端无法获得地址。

相关配置

在全局配置模式下进入 AM 规则配置模式

- 全局配置模式下使用 `address-manage` 进入 AM 配置模式；
- 使用 `match ip default` 命令配置缺省 AM 规则；
- 使用 `match ip` 命令配置基于 vlan+port/vlan 的 AM 规则；

3.4 配置详解

配置 DHCP 服务器

配置项	配置建议 & 相关命令	
配置DHCP服务器动态分配IP地址	 必须配置，用于启用 DHCP 服务器实现动态 IP 地址分配。	
	<code>service dhcp</code>	启动 DHCP-SERVER 功能
	<code>ip dhcp pool</code>	配置地址池
	<code>network</code>	配置 DHCP 地址池的网络号和掩码
	 可选配置，用于设置地址池相关属性。	
	<code>default-router</code>	配置客户端缺省网关
	<code>lease</code>	配置地址租期
	<code>next-server</code>	配置客户端启动的下载服务器地址
	<code>bootfile</code>	配置客户端启动文件
	<code>domain-name</code>	配置客户端的域名

	dns-server	配置域名服务器
	netbios-name-server	配置 NetBIOS WINS 服务器
	netbios-node-type	配置客户端 NetBIOS 节点类型
	lease-threshold	配置地址池告警门限值
	option	配置自定义选项
	pool-status	配置地址池启用或关闭
配置DHCP服务器手工地址绑定	 可选配置，用于为客户静态配置 IP 地址。	
	ip dhcp pool	配置地址池名并进入地址池配置模式
	host	配置客户端主机的 IP 地址和网络掩码
	hardware-address	配置客户端的硬件地址
	client-identifier	配置客户端的唯一标识
	client-name	配置客户端的名字
配置DHCP服务器全局属性	 可选配置，用于设置 DHCP 服务器相关属性。	
	ip dhcp excluded-address	配置排除地址
	ip dhcp force-send-nak	配置 DHCP 服务器强制回复 NAK
	ip dhcp monitor-vrrp-state	配置监控 VRRP 状态
	ip dhcp ping packets	配置 Ping 包次数
	ip dhcp ping timeout	配置 Ping 包超时时间

配置 DHCP 中继代理

配置项	配置建议 & 相关命令	
配置DHCP Relay基本功能	 必须配置。用于建立 DHCP Relay 服务。	
	service dhcp	启动 DHCP Relay 功能
	ip helper-address	配置 DHCP 服务器的 IP 地址
配置DHCP Relay option 82 功能	 可选配置。结合设备自身物理接口信息,给用户分配不同权限 IP。该功能与 dhcp option dot1x 不可以同时使用。	
	ip dhcp relay information option82	启用 DHCP option82 功能
配置DHCP Relay check server-id 功能	 可选配置。DHCP Relay 仅将 DHCP 请求报文转发到 option server-id 中指定的服务器。	
	ip dhcp relay check server-id	启用 DHCP Relay check server-di 功能
配置DHCP Relay suppression功能	 可选配置。屏蔽对应接口地址上 DHCP 请求报文。	
	ip dhcp relay suppression	启用 DHCP Relay suppression 功能

配置 DHCP 客户端

配置项	配置建议 & 相关命令
-----	-------------

配置DHCP客户端	 必须配置，用于启用 DHCP 客户端	
	ip address dhcp	使得以太网或者 PPP、HDLC、FR 封装的接口能够通过 DHCP 获得 IP 地址信息

3.4.1 配置DHCP服务器动态分配IP地址

配置效果

向所有 dhcp-client 提供 dhcp 服务，包括地址、网关等信息下发

注意事项

DHCP 服务器和 DHCP 中继共用 `service dhcp` 这条命令，但是这两个功能是互斥的，两者之间的切换依赖于是否配置了 DHCP 地址池。

配置方法

启动 DHCP-SERVER 功能

- 实现动态分配地址功能，为必选配置。
- 在配置模式下执行 `service dhcp` 命令。

配置地址池

- 创建地址池，为必选配置。
- 在配置模式下执行 `ip dhcp pool` 命令。

配置 DHCP 地址池的网络号和掩码

- 动态分配地址范围，为必选配置。
- 在地址池模式下执行 `network` 命令。

配置客户端缺省网关

- 用于通告客户端网关地址，为可选配置。
- 在地址池模式下执行 `default-router` 命令。

配置地址租期

- 用于通告客户端租约老化周期，默认值为 24h，为可选配置。
- 在地址池模式下执行 `lease` 命令。

配置客户端启动的下载服务器地址

- 用于通告客户端 TFTP 服务器地址，为可选配置。

- 在地址池模式下执行 **next-server** 命令。

↘ 配置客户端的域名

- 用于通告客户端的域名，为可选配置。
- 在地址池模式下执行 **domain-name** 命令。

↘ 配置域名服务器

- 用于通告客户端 dns 地址，为可选配置。
- 在地址池模式下执行 **dns** 命令。

↘ 配置 NetBIOS WINS 服务器

- 用于通告 windows 客户端 dns 地址，为可选配置。
- 在地址池模式下执行 **netbios-name-server** 命令。

↘ 配置客户端 NetBIOS 节点类型

- 用于通告 windows 客户端节点类型，为可选配置。
- 在地址池模式下执行 **netbios-name-type** 命令。

↘ 配置地址池告警门限值

- 用于管理租约数量，达到限制时打印警告，默认为 90%，为可选配置。
- 在地址池模式下执行 **lease-threshold** 命令。

↘ 配置自定义选项

- 用于通告客户端相当配置信息，为可选配置。
- 在地址池模式下执行 **option** 命令。

↘ 配置地址池启用或关闭

- 用于配置地址池是否可用，默认为开启，为可选配置。
- 在地址池模式下执行 **pool-status** 命令。

检验方法

利用 DHCP 客户端与 DHCP 服务器进行连接

- 检查客户端是否能取到服务器上配置的相关信息

相关命令

↘ 启动 DHCP-SERVER 功能

【命令格式】 **service dhcp**

【参数说明】 -

- 【命令模式】 全局模式
- 【使用指导】 启用 DHCP 服务器和 DHCP 中继代理功能，DHCP 服务器和 DHCP 中继共用 **service dhcp** 这条命令，两功能可以同时存在，但是报文是通过 Relay 转发还是直接由 Server 处理，取决于设备上是否配置了合法有效的地址池，如果存在地址池则由 Server 处理，不存在由 Relay 转发。

配置地址池

- 【命令格式】 **ip dhcp pool** *dhcp-pool*
- 【参数说明】 *pool-name* : 地址池名称
- 【命令模式】 全局模式
- 【使用指导】 要给用户下发地址，首先要配置地址池名并进入地址池配置模式

配置 DHCP 地址池的网络号和掩码

- 【命令格式】 **network** *network-number mask [low-ip-address high-ip-address]*
- 【参数说明】 *network-number*: DHCP 地址池的 IP 地址网络号
mask: DHCP 地址池的 IP 地址网络掩码。如果没有定义掩码，缺省为自然网络掩码
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】 进行动态地址绑定的配置，必须配置新建地址池的子网及其掩码，为 DHCP 服务器提供了一个可分配给客户端的地址空间。DHCP 在分配地址池中的地址，是按顺序进行的，如果该地址已经在 DHCP 绑定表中或者检测到该地址已经在该网段中存在，就检查下一个地址，直到分配一个有效的地址。
因此，为了避免上述问题，要求网络管理员在构建网络的时候，通过其它的方式，如调整物理链路或者网络路径，来避免这种客户端到服务器的路径冗余

配置客户端缺省网关

- 【命令格式】 **default-router** *address [address2...address8]*
- 【参数说明】 *address* : 定义客户端默认网关的 IP 地址。要求至少配置一个
ip-address2...ip-address8 : (可选) 最多可以配置 8 个网关
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】 配置客户端默认网关，这个将作为服务器分配给客户端的默认网关参数。缺省网关的 IP 地址必须与 DHCP 客户端的 IP 地址在同一网络

配置地址租期

- 【命令格式】 **lease** {*days [hours] [minutes]* | **infinite**}
- 【参数说明】 *days* : 定义租期的时间，以天为单位
hours: (可选) 定义租期的时间，以小时为单位。定义小时数前必须定义天数
minutes: (可选) 定义租期的时间，以分钟为单位。定义分钟前必须定义天数和小时数
infinite: 定义没有限制的租期
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】 DHCP 服务器给客户端分配的地址，缺省情况下租期为 1 天。当租期快到时客户端需要请求续租，否则过期后就不能使用该地址

配置客户端启动文件

- 【命令格式】 **bootfile** *filename*
- 【参数说明】 *file-name* : 定义用于启动的文件名
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】 客户端启动文件是客户端启动时要用到的启动映像文件。启动映像文件通常是 DHCP 客户端需要下载的操作系统

配置客户端的域名

- 【命令格式】 **domain-name** *domain*
- 【参数说明】 *domain-name*: 定义 DHCP 客户端的后缀域名字符串
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】 可以指定客户端的域名, 这样当客户端通过主机名访问网络资源时, 不完整的主机名会自动加上域名后缀形成完整的主机名

配置域名服务器

- 【命令格式】 **dns-server** *ip-address* [*ip-address2...ip-address8*]
- 【参数说明】 *ip-address*: 定义 DNS 服务器的 IP 地址。要求至少配置一个
ip-address2...ip-address8: (可选) 最多可以配置 8 个 DNS 服务器
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】 当客户端通过主机名访问网络资源时, 需要指定 DNS 服务器进行域名解析。要配置 DHCP 客户端可使用的域名服务器

配置 NetBIOS WINS 服务器

- 【命令格式】 **netbios-name-server** *address* [*address2...address8*]
- 【参数说明】 *address*: 定义 WINS 服务器的 IP 地址。要求至少配置一个
ip-address2...ip-address8: (可选) 最多可以配置 8 个 WINS 服务器
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】 WINS 是微软 TCP/IP 网络解析 NetBIOS 名字到 IP 地址的一种域名解析服务。WINS 服务器是一个运行在 Windows NT 下的服务器。当 WINS 服务器启动后, 会接收从 WINS 客户端发送的注册请求, WINS 客户端关闭时, 会向 WINS 服务器发送名字释放消息, 这样 WINS 数据库中与网络上可用的计算机就可以保持一致了

配置客户端 NetBIOS 节点类型

- 【命令格式】 **netbios-node-type** *type*
- 【参数说明】 *type*: 定义 NetBIOS 节点类型, 有两种方式
- 数字定义, 范围从 0~FF, 十六进制数, 但只能取以下值:
- 代表 b-node
 - 代表 p-node
 - 代表 m-node
 - 8, 代表 h-node
- 字符串定义:
- b-node, 广播型节点
 - p-node, 对等型节点

- m-node, 混合型节点
- h-node, 复合型节点

【命令模式】 DHCP 地址池配置模式

【使用指导】 微软 DHCP 客户端 NetBIOS 节点类型有四种：1) Broadcast, 广播型节点, 通过广播方式进行 NetBIOS 名字解析；2) Peer-to-peer, 对等型节点, 通过直接请求 WINS 服务器进行 NetBIOS 名字解析；3) Mixed, 混合型节点, 先通过广播方式请求名字解析, 后通过与 WINS 服务器连接进行名字解析；4) Hybrid, 复合型节点, 首先直接请求 WINS 服务器进行 NetBIOS 名字解析, 如果没有得到应答, 就通过广播方式进行 NetBIOS 名字解析。

缺省情况下, 微软操作系统的节点类型为广播型或者复合型。如果没有配置 WINS 服务器, 就为广播型节点；如果配置了 WINS 服务器, 就为复合型节点

配置自定义选项

【命令格式】 **option code { ascii string | hex string | ip ip-address }**

【参数说明】 **code:** 定义 DHCP 选项代码

ascii string: 定义一个 ASCII 字符串

hex string: 定义十六进制字符串

ip ip-address: 定义 IP 地址列表

【命令模式】 DHCP 地址池配置模式

【使用指导】 DHCP 提供了一个机制, 允许在 TCP/IP 网络中将配置信息传送给主机。DHCP 报文专门有 option 字段, 该部分内容为可变化内容, 用户可以根据实际情况进行定义, DHCP 客户端必须能够接收携带至少 312 字节 option 信息的 DHCP 报文。另外 DHCP 报文中的固定数据字段也称为一个选项

配置地址池启用或关闭

【命令格式】 **pool-status { enable | disable }**

【参数说明】 **enable:** 启用地址池

disable: 关闭地址池

默认为开启

【命令模式】 DHCP 地址池配置模式

【使用指导】 -

配置举例

配置地址池

- 【配置方法】
- 定义了一个地址池 net172
 - 地址池网段为 172.16.1.0/24
 - 缺省网关为 172.16.16.254
 - 地址租期为 1 天
 - 排除 172.16.1.2~172.16.1.100 地址

```
Ruijie(config)# ip dhcp excluded-address 172.16.1.2 172.16.1.100
Ruijie(dhcp-config)# ip dhcp pool net172
Ruijie(dhcp-config)# network 172.16.1.0 255.255.255.0
```

```
Ruijie(dhcp-config)# default-router 172.16.1.254
Ruijie(dhcp-config)# lease 1
```

【检验方法】 1.show run 查看

```
Ruijie(config)#show run | begin ip dhcp
ip dhcp excluded-address 172.16.1.2 172.16.1.100
ip dhcp pool net172
default-router 172.16.1.254
domain-name rg.com
dns-server 172.16.1.253
netbios-name-server 172.16.1.252
netbios-node-type h-node
lease 30
```

3.4.2 配置DHCP服务器手工地址绑定

配置效果

向某些特定的 dhcp-client 下发特定的 ip 地址及其它配置信息

注意事项

无

配置方法

▾ 配置地址池名并进入地址池配置模式

- 创建地址池，为必选配置。
- 在配置模式下执行 **ip dhcp pool** 命令。

▾ 配置客户端主机的 IP 地址和网络掩码

- 配置静态 ip 地址及网络掩码，必选配置。
- 在地址池模式下执行 **host** 命令。

▾ 配置客户端的硬件地址

- 配置静态 mac 地址，可选配置。
- 在地址池模式下执行 **hardware** 命令。

▾ 配置客户端的唯一标识

- 配置静态用户 `uid`，可选配置。
- 在地址池配置下执行 `client-identifier` 命令。

配置客户端的名字

- 配置静态用户名字，可选配置。
- 在地址池模式下执行 `host-name` 命令。

检验方法

对应的用户上线，判断是否能取到相应地址。

相关命令

配置地址池

- 【命令格式】 `ip dhcp pool dhcp-pool`
- 【参数说明】 `pool-name`：地址池名称
- 【命令模式】 全局模式
- 【使用指导】 要给用户下发地址，首先要配置地址池名并进入地址池配置模式

手工地址绑定

- 【命令格式】
`host ip-address [netmask]`
`client-identifier unique-identifier`
`client-name name`
- 【参数说明】
`ip-address`: 定义 DHCP 客户端主机的 IP 地址
`netmask`: 定义 DHCP 客户端主机的网络掩码
`unique-identifier`: 定义客户端硬件地址，如 `aabb.bbbb.bb88`;定义客户端的标识，如 `01aa.bbbb.bbbb.88`
`name`: (可选)用标准的 ASCII 字符定义客户端的名字，名字不要包括域名。如定义 `mary` 主机名，不可定义成 `mary.rg.com`
- 【命令模式】 DHCP 地址池配置模式
- 【使用指导】
地址绑定是指 IP 地址和客户端 MAC 地址的映射关系。地址绑定有两种：1) 手工绑定，就是在 DHCP 服务器数据库中，通过手工定义将 IP 地址和 MAC 地址进行静态映射，手工绑定其实是一个特殊地址池；2) 动态绑定，DHCP 服务器接收到 DHCP 请求时，动态地从地址池中分配 IP 地址给客户端，而形成的 IP 地址和 MAC 地址映射。
要定义手工地址绑定，首先需要为每一个手动绑定定义一个主机地址池，然后定义 DHCP 客户端的 IP 地址和硬件地址或客户端标识。硬件地址就是 MAC 地址。客户端标识，微软客户端一般定义客户端标识，而不定义 MAC 地址，客户端标识包含了网络媒介类型和 MAC 地址。关于媒介类型的编码，请参见 RFC 1700 中关于“Address Resolution Protocol Parameters”部分内容。以太网类型为“01”

配置举例

动态地址池

- 【配置方法】
- 地址池 vlan1 20.1.1.0 255.255.255.0
 - 缺省网关为 20.1.1.1
 - 租约时间为 1 天

```
Ruijie(config)# ip dhcp pool vlan1
Ruijie(dhcp-config)# network 20.1.1.0 255.255.255.0
Ruijie(dhcp-config)# default-router 172.16.1.254
Ruijie(dhcp-config)# lease 1 0 0
```

【检验方法】

1. show run 查看

```
Ruijie(config)#show run | begin ip dhcp
ip dhcp pool vlan1
network 20.1.1.0 255.255.255.0
hardware-address 00d0.df34.32a3 ethernet
default-router 20.1.1.1
lease 1 0 0
```

手工绑定配置

- 【配置方法】
- 主机地址 172.16.1.101，掩码为 255.255.255.0
 - 主机名 Billy.rg.com
 - 缺省网关为 172.16.16.254
 - MAC 地址为 00d0.df34.32a3

```
Ruijie(config)# ip dhcp pool Billy
Ruijie(dhcp-config)# host 172.16.1.101 255.255.255.0
Ruijie(dhcp-config)# client-name Billy
Ruijie(dhcp-config)# hardware-address 00d0.df34.32a3 ethernet
Ruijie(dhcp-config)# default-router 172.16.1.254
Ruijie(dhcp-config)# dns-server 172.16.1.253
Ruijie(dhcp-config)# dns-server 172.16.1.253
```

【检验方法】

1. show run 查看

```
Ruijie(config)#show run | begin ip dhcp
ip dhcp pool Billy
host 172.16.1.101 255.255.255.0
client-name Billy
hardware-address 00d0.df34.32a3 ethernet
default-router 172.16.1.254
dns-server 172.16.1.253
```

3.4.3 配置基于vlan / port地址分配

配置效果

配置该命令后，可依据端口+VLAN 按区间进行地址分配

注意事项

锐捷产品目前版本支持以太网接口、千兆口以及 FR、PPP、HDLC 接口上的配置。

配置方法

在 config 模式下执行 address-manage

在 address-manage 模式下执行 match ip 命令

检验方法

查看不同 vlan、端口下的用户是否取到有效地址

相关命令

▾ 配置缺省区间

【命令格式】 **match ip default** *ip-address netmask*

【参数说明】 *ip-address*: 网络地址

netmask: 地址掩码

【命令模式】 address-manage 模式下

【使用指导】 配置该命令后所有来自未配置 vlan + port 的 DHCP 客户端将取得缺省区间内的地址，若无该配置命令同时也无任何其它 vlan + port 配置，则按正常流程分配地址。

配置基于 vlan/port 规则下的动态地址分配

【命令格式】 **match ip** *ip-address netmask interface* [**add/remove**] **vlan** *vlan-list*

【参数说明】 *ip-address*: 网络地址

netmask: 地址掩码

interface: 接口名称

add/remove: 添加或删除指定 vlan

vlan-list: *vlan* 索引

【命令模式】 address-manage 模式下

【使用指导】 配置该命令后来自指定 vlan + port 的 DHCP 客户端将取得配置区内地址。

配置基于 vlan 规则下的静态地址分配

【命令格式】 **match ip** *ip-address netmask* [**add/remove**] **vlan** *vlan-list*

【参数说明】 *ip-address*: 网络地址

netmask: 地址掩码

add/remove: 添加或删除指定 vlan

vlan-list: *vlan* 索引

【命令模式】 address-manage 模式下

【使用指导】 在 supervlan 场景下，满足 Dhcp 静态地址池配置的用户，无论在哪个 subvlan 下都只分配该静态地址；此时 AM 无需基于所有 subvlan/port 对该地址进行配置，只需要配置该地址在对应的 vlan 区间生效即可。该规则当前只对静态地址分配生效，动态地址不生效。

配置举例

AM 规则配置

- 【配置方法】
- 配置缺省规则规则
 - 配置指定 vlan+port+地址区间规则
 - 配置指定 vlan+地址区间规则

```
Ruijie(config)# address-manage
Ruijie(config-address-manage)# match ip default 172.50.128.0 255.255.128.0
Ruijie(config-address-manage)# match ip 10.1.5.0 255.255.255.0 Gi5/3 vlan 1005
Ruijie(config-address-manage)# match ip 10.1.6.0 255.255.255.0 vlan 1006
```

【检验方法】 1 : **show run** 查看

```
address-manage
match ip default 172.50.128.0 255.255.128.0
match ip 10.1.5.0 255.255.255.0 Gi5/3 vlan 1005
```

```
match ip 10.1.6.0 255.255.255.0 vlan 1006
```

3.4.4 配置DHCP服务器全局属性

配置效果

开启服务器一些特定的功能，如 ping 机制、强制 nak 等。

注意事项

Nak 命令的配置可能引起网络中其它服务器的功能异常。

配置方法

配置排除地址

- 配置某些地址或地址段不可用，为可选配置。
- 在配置模式下执行 **ip dhcp excluded-address** 命令

配置 DHCP 服务器强制回复 NAK

- 针对某些用户的错误地址请求，服务器回复 nak 报文，可选配置。
- 在配置模式下执行 **ip dhcp force-send-nak** 命令。

配置监控 VRRP 状态

- 启动该功能后，主机 server 处理 DHCP 相关报文，备机 server 则不处理 DHCP 相关报文，可选配置。
- 在配置模式下执行 **ip dhcp monitor-vrrp-state** 命令。

配置 Ping 包次数

- 检查地址的可达性，执行 ping 操作，默认值为 2，可选配置。
- 在配置模式下执行 **ip dhcp ping packet** 命令。

配置 Ping 包超时时间

- 检查地址的可达性，设置 ping 返回时长，默认值为 500ms，可选配置。
- 在配置模式下执行 **ip dhcp ping timeout** 命令。

检验方法

启动 **dhcp-server** 下发地址过程中可检验。

相关命令

配置排除地址

【命令格式】 **ip dhcp excluded-address** *low-ip-address* [*high-ip-address*]

【参数说明】 *low-ip-address*: 排斥 IP 地址范围的起始 IP 地址
high-ip-address : 排斥地址范围的结束 IP 地址

【命令模式】 全局模式

【使用指导】 如果没有特别配置，DHCP 服务器会试图将在地址池中定义的所有子网地址分配给 DHCP 客户端。因此，如果想保留一些地址不分配，比如已经分配给服务器或者设备了，必须明确定义这些地址是不允许分配给客户端的；配置 DHCP 服务器，一个好的习惯是将所有已明确分配的地址全部不允许 DHCP 分配，这样可以带来两个好处：1) 不会发生地址冲突；2) DHCP 分配地址时，减少了检测时间，从而提高 DHCP 分配效率

配置 DHCP 服务器强制回复 NAK

【命令格式】 **ip dhcp force-send-nak**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 -

配置 Ping 包次数

【命令格式】 **ip dhcp ping packets** [*number*]

【参数说明】 *Number* : (可选) 范围从 0 到 10，0 表示关闭 ping 操作。缺省 ping 两个包

【命令模式】 全局配置模式

【使用指导】 缺省情况，当 DHCP 服务器试图从地址池中分配一个 IP 地址时，会对该地址执行两次 Ping 命令(一次一个数据包)。如果 Ping 没有应答，DHCP 服务器认为该地址为空闲地址，就将该地址分配给 DHCP 客户端；如果 Ping 有应答，DHCP 服务器认为该地址已经在使用，就试图分配另外一个地址给 DHCP 客户端，直到分配成功

配置 Ping 包超时时间

【命令格式】 **ip dhcp ping timeout** *milliseconds*

【参数说明】 *milli-seconds* : DHCP 服务器等待 ping 应答的时间(以毫秒计)。取值范围为 100 到 10000

【命令模式】 全局配置模式

【使用指导】 缺省情况下，DHCP 服务器 Ping 操作如果 500 毫秒没有应答，就认为没有该 IP 地址主机存在。可以通过调整 Ping 包超时时间，改变服务器 Ping 等待应答的时间

配置举例

配置 ping 机制

- 【配置方法】
- 配置 ping 次数为 5
 - 配置 ping 超时时长为 800ms

```
Ruijie(config)# ip dhcp ping packet 5
Ruijie(config)# ip dhcp ping timeout 800
```

【检验方法】 1.show run 查看

```
Ruijie(config)#show run | begin ip dhcp
ip dhcp ping packet 5
ip dhcp ping timeout 800
```

▾ 配置排除地址**【配置方法】 ● 排除 192.168.0.0 – 192.168.255.255 的所有地址**

```
Ruijie(config)# ip dhcp excluded-address 192.168.0.0 192.168.255.255
```

【检验方法】 1.show run 查看

```
Ruijie(config)#show run | begin ip dhcp
ip dhcp excluded-address 192.168.0.0 192.168.255.255
```

3.4.5 配置DHCP Relay基本功能

配置效果

- 建立 Client—Relay—Server 模式的 DHCP 动态 IP 管理，解决 DHCP 客户端与 DHCP 服务器不在同一网段时 DHCP 客户端与在其他网段的 DHCP 服务器通讯问题。

注意事项

- DHCP Relay 需要借助网络中现有的单播路由。因此，网络中必须配置 IPv4 单播路由。

配置方法

▾ 启动 DHCP Relay 功能

- 必须配置。
- 若无特殊要求，应在设备上启动 DHCP Relay 功能。

▾ 配置 DHCP 服务器的 IP 地址

- 必须配置。
- 应在设备上启动 DHCP 服务器的 IP 地址。

检验方法

- 检查用户主机能否通过 DHCP Relay 成功获取到 IP 地址。

相关命令

启动 DHCP Relay 功能

- 【命令格式】 **service dhcp**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置 DHCP 服务器的 IP 地址

- 【命令格式】 **ip helper-address { cycle-mode [A.B.C.D]}**
- 【参数说明】 *cycle-mode* : 开启 dhcp 请求报文转发所有 dhcp 服务器
A.B.C.D: Server 的 ip 地址
- 【命令模式】 全局模式、接口模式
- 【使用指导】 配置接口必须是三层接口, 包括: 路由口、L3AP、SVI、loopback 接口。
所有配置接口应 IPv4 单播路由可达。

配置举例

i 以下配置举例, 仅介绍与 DHCP Relay 相关的配置。

有线场景中 DHCP Relay 配置

【网络环境】

图 3-16



- 【配置方法】
- 用户设备启动通过 DHCP 获取地址的功能。
 - 在作为 DHCP Relay Agent 的网络设备中启动 DHCP Relay 功能。
 - 配置 DHCP Server。

A 用户设备启动 DHCP 获取地址的功能。

B # 启用 DHCP 中继代理

```
Ruijie(config)# service dhcp
```

```
# 添加一个全局的 DHCP 服务器的地址
```

```
Ruijie(config)# ip helper-address 172.2.2.1
```

```
# 配置与用户设备连接的端口的 IP 地址
```

```
Ruijie(config)# interface gigabitEthernet 0/1
```

```
Ruijie(config-if)# ip address 192.1.1.1 255.255.255.0
```

```
# 配置与 Server 设备连接的端口的 IP 地址
```

```
Ruijie(config)# interface gigabitEthernet 0/2
```

```
Ruijie(config-if-gigabitEthernet 0/2)# ip address 172.2.2.2 255.255.255.0
```

C # 启用 DHCP SERVER 功能

```
Ruijie(config)# service dhcp
```

添加一个客户端地址池

```
Ruijie(config)# ip dhcp pool relay
```

```
Ruijie (dhcp-config)#network 192.1.1.0 255.255.255.0
```

```
Ruijie (dhcp-config)#default-router 192.1.1.1
```

配置与 relay 设备连接的端口的 IP 地址

```
Ruijie(config)# interface gigabitEthernet 0/1
```

```
Ruijie(config-if-gigabitEthernet 0/2)# ip address 172.2.2.1 255.255.255.0
```

【检验方法】 查看用户是否能获取到 IP 地址。

- 检查用户设备是否能获取到 IP 地址。
- 检查 DHCP Relay 配置是否正确。

A 用户设备能获取到 IP 地址

B 登录到 DHCP Relay Agent 设备后在特权模式下用 **show running-config** 命令显示 DHCP Relay 配置

```
Ruijie# show running-config
service dhcp
ip helper-address 172.2.2.1
!
interface GigabitEthernet 0/1
ip address 192.1.1.1 255.255.255.0
!
interface GigabitEthernet 0/2
ip address 172.2.2.2 255.255.255.0
!
```

常见错误

- IPv4 单播路由配置错误。
- 没有启动 DHCP Relay 功能。
- 没有配置 DHCP Relay 与 DHCP Service 之间的路由。
- 没有配置 DHCP 服务器 IP 地址。

3.4.6 配置DHCP Relay option 82 功能

配置效果

- 中继设备进行 DHCP Relay 时，可以通过添加 option 的方式来详细的标明 DHCP 客户端的一些网络信息，从而使服务器可以根据更精确的信息给用户分配不同权限的 IP。

注意事项

- 必须配置 DHCP Relay 基本功能。

配置方法

启动 DHCP Relay 基本功能

- 必须配置。
- 若无特殊要求，应在设备上启动 DHCP Relay 基本功能。

启动 DHCP option82 功能

- 缺省情况下，设备上的 DHCP option 82 功能关闭。
- 使用 `ip dhcp relay information option82` 命令可以启动或关闭设备上的 DHCP option 82 功能。

检验方法

- 检查客户端获取到的 IP 地址，是否是根据 option 82 规则分配。。

相关命令

配置 DHCP option82 功能

- 【命令格式】 `ip dhcp relay information option82`
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

启动 DHCP option 82 功能。

- 【配置方法】 ● 启动 DHCP option 82 功能

```
Ruijie(config)# ip dhcp relay information option82
```

- 【检验方法】 登录到 DHCP Relay Agent 设备后在特权模式下用 `show running-config` 命令显示 DHCP Relay 配置。

```
Ruijie#show ru | incl ip dhcp relay  
ip dhcp relay information option82
```

常见配置错误

- DHCP Relay 基本功能没有配置，或配置失败。

3.4.7 配置DHCP Relay check server-id功能

配置效果

- 当配置命令 **ip dhcp relay check server-id** 后，DHCP Relay 仅将 DHCP 请求报文转发到 option server-id 中指定的服务器。如果没有配置该命令，则向所有配置的 DHCP 服务器转发 DHCP 请求报文。

注意事项

- 必须配置 DHCP Relay 基本功能。

配置方法

▾ 启动 DHCP Relay check server-id 功能

- 缺省情况下，设备上的 DHCP Relay check server-id 功能关闭。
- 使用 **ip dhcp relay check server-id** 命令可以启动设备上的 DHCP Relay check server-id 功能。

检验方法

DHCP Relay 是否仅将 DHCP 请求报文转发到 option server-id 中指定的服务器。

相关命令

▾ 配置 DHCP Relay check server-id 功能

- 【命令格式】 **ip dhcp relay check server-id**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

▾ 配置 DHCP Relay check server-id 功能。

- 【配置方法】
 - 配置 DHCP Relay 基本功能。略
 - 在对应接口上配置 DHCP Relay check server-id 功能。

```
Ruijie# configure terminal
Ruijie(config)# ip dhcp relay check server-id
```

- 【检验方法】 登录到 DHCP Relay Agent 设备后在特权模式下用 **show running-config** 命令显示 DHCP Relay 配置。

```
Ruijie# show running-config | include check server-id
ip dhcp relay check server-id
Ruijie#
```

常见配置错误

- DHCP Relay 基本功能没有配置，或配置失败。

3.4.8 配置DHCP Relay suppression功能

配置效果

- 在指定接口上配置命令 **ip dhcp relay suppression** 后，将屏蔽该接口上收到的 DHCP 请求报文；而对于其他接口上收到的 DHCP 请求报文，则正常转发。

注意事项

- 必须配置 DHCP Relay 基本功能。

配置方法

▾ 启动 DHCP Relay suppression 功能

缺省情况下，设备上所有接口的 DHCP Relay suppression 功能关闭。

使用 **ip dhcp relay suppression** 命令可以启动设备上的 DHCP Relay suppression 功能。

检验方法

- 接口上收到的 DHCP 请求报文是否被过滤。

相关命令

▾ 配置 DHCP Relay suppression 功能

- 【命令格式】 **ip dhcp relay suppression**
【参数说明】 -
【命令模式】 接口模式
【使用指导】 -

配置举例

配置 DHCP Relay suppression 功能。

- 【配置方法】
- 配置 DHCP Relay 基本功能。略
 - 在对应接口上配置 DHCP Relay suppression 功能。

```
Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# ip dhcp relay suppression
Ruijie(config-if-GigabitEthernet 0/1)#end
Ruijie#
```

- 【检验方法】 登录到 DHCP Relay Agent 设备后在特权模式下用 **show running-config** 命令显示 DHCP Relay 配置。

```
Ruijie# show running-config | include relay suppression
ip dhcp relay suppression
Ruijie#
```

常见配置错误

DHCP Relay 基本功能没有配置，或配置失败。

3.4.9 配置DHCP客户端

配置效果

设备启动 dhcp-client，可动态取得地址及其它需求配置。

注意事项

锐捷产品目前版本支持以太网接口以及 FR、PPP、HDLC 接口上的 DHCP 客户端。

配置方法

在接口上执行 **ip address dhcp** 命令

检验方法

查看接口是否取到 ip 地址

相关命令

配置 DHCP 客户端

【命令格式】 **ip address dhcp**

【参数说明】 -

【命令模式】 接口配置模式

- 【使用指导】
- 锐捷产品支持以太网端口通过 DHCP 获得动态分配的 IP 地址
 - 锐捷产品支持 ppp 封装的端口通过 DHCP 获得动态分配的 IP 地址
 - 锐捷产品支持 FR 封装的端口通过 DHCP 获得动态分配的 IP 地址
 - 锐捷产品支持 HDLC 封装的端口通过 DHCP 获得动态分配的 IP 地址

配置举例

DHCP 客户端配置

【配置方法】 1：为设备接口 FastEthernet 0/0 配置 DHCP 自动分配地址

```
Ruijie(config)# interface FastEthernet0/0
Ruijie(config-if-FastEthernet 0/0)#ip address dhcp
```

【检验方法】 1：**show run** 查看

```
Ruijie(config)#show run | begin ip address dhcp
ip address dhcp
```

3.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除 DHCP 地址绑定	clear ip dhcp binding { address * }
清除 DHCP 地址冲突	clear ip dhcp conflict { address * }
清除 DHCP 服务器统计状态	clear ip dhcp server statistics
清除 DHCP 中继统计状态	clear ip dhcp relay statistics
清除 DHCP 服务器性能统计信息	clear ip dhcp server rate

查看运行情况

作用	命令
----	----

显示 DHCP 租约信息	show dhcp lease
显示手工配置的地址	show dhcp manual
显示 dhcp 用的套接字	show ip dhcp socket
显示已经分配的地址	show ip dhcp binding
显示创建的地址池	show ip dhcp pool
显示 dhcp-server 统计信息	show ip dhcp server statistic
显示 dhcp-relay 统计信息	show ip dhcp relay statistic
显示冲突地址	show ip dhcp conflict
显示 DHCP 服务器地址池的 ID 和地址使用情况	show ip dhcp identifier

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
DHCPagent 调试开关	debug ip dhcp server agent
DHCP 热备调试开关	debug ip dhcp server ha
DHCP 地址池调试开关	debug ip dhcp server pool
DHCP VRRP 调试开关	debug ip dhcp server vrrp
DHCP 打开所有调试开关	debug ip dhcp server all
DHCP 报文调试开关	debug ip dhcp client
DHCP Relay 事件调试开关。	debug ip dhcp relay

4 DNS

4.1 概述

DNS (Domain Name System , 域名系统) , 因特网上作为域名和IP地址相互映射的一个 分布式数据库 , 能够使用户更方便的访问 互联网 , 而不用去记住能够被机器直接读取的IP数串。通过主机名 , 最终得到该主机名对应的IP地址的过程叫做域名解析 (或主机名解析) 。

 下文仅介绍 DNS 的相关内容。

协议规范

- RFC1034 : DOMAIN NAMES - CONCEPTS AND FACILITIES
- RFC1035 : DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION

4.2 典型应用

典型应用	场景描述
静态域名解析	直接在本设备上根据预设的域名/IP 对应表进行域名解析
动态域名解析	从网络上的 DNS 服务器动态获取域名对应的地址

4.2.1 静态域名解析

应用场景

- 在设备上预设置域名和 IP 的对应表
- 设备上的一些应用 (比如 Ping , Telnet 等) 进行域名操作时 , 直接在设备上就能解析到预设的 IP , 无需连到网络上的服务器。

功能部属

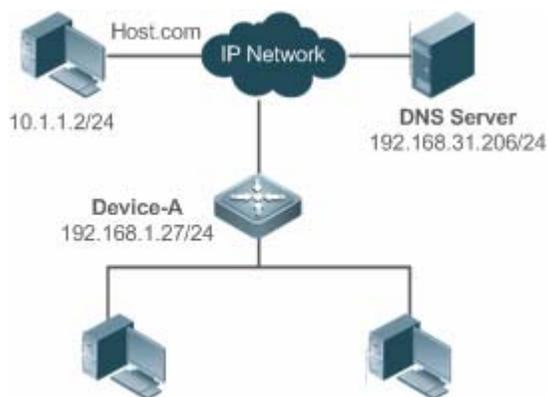
- 在设备上预设置域名和 IP 的对应关系

4.2.2 动态域名解析

应用场景

- “DNS Server” 部署在网络上，对外提供域名服务
- “host.com” 部署在网络上，使用域名(host.com)对外提供服务
- “Device-A”设备指定 “DNS Server” 作为 DNS 服务器，从 “DNS Server” 上获取到 “host.com”的地址

图 4-1 动态域名解析配置组网图



功能部属

- 将 DNS Server 部署为“Device-A”的 DNS 服务器

4.3 功能详解

基本概念

DNS

DNS 由解析器和域名服务器组成。域名服务器是指保存有网络中所有主机的域名和 IP 地址的对应关系，并提供将域名和 IP 互转的服务器。DNS 的 TCP 和 UDP 端口号都是 53，通常使用 UDP。

功能特性

功能特性	作用
域名解析	根据域名从域名服务器或本地数据库获取对应的 IP 地址

4.3.1 域名解析

工作原理

▾ 静态域名解析

静态域名解析，就是用户在设备上预先设置好域名和 IP 的对应关系，当用户使用某些应用(比如 Ping、Telnet 等等)进行域名操作时，系统从本设备上解析出域名对应的 IP，而不需要到网络上的 DNS 服务器获取域名对应的 IP。

▾ 动态域名解析

动态域名解析，就是当用户使用某些应用进行域名操作时，系统 DNS 解析器查询外部的 DNS 服务器，获取到域名对应的 IP。

动态域名解析过程：

4. 用户应用(Ping、Telnet 等)向系统 DNS 解析器请求域名对应的 IP
5. 系统 DNS 解析器先查找动态缓存，如果动态缓存的域名未过期则返回给应用程序
6. 如果不存在未过期的域名，DNS 解析器向外部的 DNS 服务器发起域名转 IP 的请求
7. DNS 解析器接收到 DNS 服务器的应答，缓存并转发给应用程序

相关配置

▾ 开启域名解析功能

- 缺省情况下，设备是开启域名解析功能。
- 通过 `ip domain-lookup` 命令开启或关闭域名解析功能。

▾ 配置静态域名对应的 IP

- 缺省情况下，没有域名/IP 的静态配置。
- 通过 `ip host` 命令指定域名对应的 IPv4 地址

▾ 配置域名服务器

- 缺省情况下，未配置域名服务器。
- 通过 `ip name-server` 命令配置域名服务器。

4.4 配置详解

配置项	配置建议 & 相关命令
配置静态域名解析	 可选配置

	ip domain-lookup	开启域名解析功能
	ip host	配置域名对应的 IPv4 地址
配置动态域名解析	 可选配置	
	ip domain-lookup	开启域名解析功能
	ip name-server	配置域名服务器

4.4.1 配置静态域名解析

配置效果

系统解析器从设备本地解析域名对应的 IP

配置方法

▾ 开启域名解析功能

- 缺省已开启域名解析功能
- 如果关闭该功能，静态域名解析不生效。

▾ 配置静态域名对应的 IPv4 地址

- 必须配置，用户使用到的域名必须配置对应的 IP。

检验方法

- 通过 **show run** 查看配置信息。
- 通过 **show hosts** 当前的域名和 IP 对应关系

相关命令

▾ 配置域名对应的 IPv4 地址

【命令格式】 **ip host** *host-name ip-address*

【参数说明】 *host-name* : 域名

ip-address : 对应的 IPv4 地址

【命令模式】 全局模式

【使用指导】 -

配置举例

配置静态域名解析

- 【配置方法】
- 在设备上静态配置域名 `www.test.com` 的 IP 地址为 `192.168.1.1`

```
Ruijie#configure terminal
Ruijie(config)# ip host www.test.com 192.168.1.1
Ruijie(config)# exit
```

- 【检验方法】 通过 **show hosts** 查看是否有所配置的静态域名表项

```
Ruijie#show hosts
Name servers are:

Host                type    Address      TTL(sec)
www.test.com        static  192.168.1.1  ---
```

4.4.2 配置动态域名解析

配置效果

系统解析器从 DNS 服务器解析域名对应的 IP

配置方法

开启域名解析功能

- 缺省已开启域名解析功能
- 如果关闭该功能，动态域名解析不生效。

配置 DNS 服务器

- 必须配置，使用动态域名解析必须配置外部的 DNS 服务器。

检验方法

- 通过 **show run** 查看配置信息

相关命令

配置域名服务器

【命令格式】 **ip name-server** [oob] { ip-address } [via mgmt-name]

【参数说明】 *ip-address* : DNS 服务器的 IPv4 地址

oob : DNS 服务器支持带外管理接口 (interface of mgmt)。

via : 配置 mgmt 出口

mgmt-name : 指定在 oob 模式下报文的出口管理口

【命令模式】 全局模式

【使用指导】 -

配置举例

配置动态域名解析

【网络环境】

图 4-2



DEVICE : 从网络上的 DNS 服务器(192.168.10.1)解析域名

【配置方法】 在设备上配置 DNS 服务器地址为 192.168.10.1

```
DEVICE#configure terminal
DEVICE(config)# ip name-server 192.168.10.1
DEVICE(config)# exit
```

【检验方法】 通过 **show hosts** 查看是否配置指定 DNS 服务器

```
Ruijie(config)#show hosts
Name servers are:
192.168.10.1 static

Host          type      Address                               TTL(sec)
```

4.5 监视与维护

清除各类信息

! 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除动态主机名缓存表。	clear host [<i>host-name</i>]

查看运行情况

作用	命令
查看 DNS 的相关参数	show hosts [<i>host-name</i>]

查看调试信息

! 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开调试功能	debug ip dns

5 FTP-Server

5.1 概述

FTP Server 功能可以将一台设备配置为 FTP 服务器。这样可以通过 FTP 客户端与之连接，通过 FTP 协议往设备上传或下载文件。

用户可以利用 FTP Server 功能方便地获取设备中的文件，如 syslog 日志文件等；也可以通过 FTP 直接往设备的文件系统拷贝文件。

i 下文仅介绍 FTP 的相关内容。

协议规范

- RFC959 : FILE TRANSFER PROTOCOL (FTP)
- RFC3659 : Extensions to FTP
- RFC2228: FTP Security Extensions
- RFC2428: FTP Extensions for IPv6 and NATs
- RFC1635: How to Use Anonymous FTP

5.2 典型应用

典型应用	场景描述
局域网内提供FTP服务	在一个局域网内为同一个用户提供上传与下载服务

5.2.1 局域网内提供FTP服务

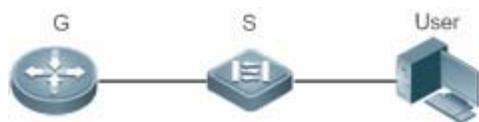
应用场景

在一个局域网内为同一个用户提供上传与下载服务

以下图为例，仅在局域网内开启 FTP-Server 服务

- G 开启 FTP Server 服务，S 二层透传功能
- User 发起 FTP 上传与下载请求

图 5-1



【注释】 G 为出口网关设备。
S 为接入设备

功能部属

- G 启动 FTP Server
- S 当作二层交换机，起到二层透传的作用

5.3 功能详解

基本概念

FTP 协议

FTP (File Transfer Protocol) 是 IETF Network Working Group 所制定的一套标准协议，属于网络协议组的应用层，FTP 基于 TCP 传输控制协议(Transmission Control Protocol)实现文件传输。FTP 使用户能在两个联网的计算机之间传输文件，它是 Internet 传递文件最主要的方法。使用匿名 FTP，用户可以免费获取 Internet 丰富的资源。除此之外，FTP 还提供登录、目录查询、文件操作及其它会话控制等功能。FTP 协议在 TCP/IP 协议族中属于应用层协议，使用 TCP 端口 20 和 21 进行传输。端口 20 用于传输数据，端口 21 用于传输控制消息。FTP 协议基本操作在 RFC959 中进行了描述。

用户授权

FTP Client要连上 FTP Server，必须要有该 FTP服务器授权的帐号，只有拥有一个用户标识和一个口令后才能登陆FTP服务器，享受FTP服务器提供的服务。

FTP 文件传输模式

FTP 有两种文件传输模式：

- 文本传输方式：也称为 ASCII 模式，用于传输文本格式的文件（比如后缀名为.txt、.bat 和.cfg 的文件），与 Binary 模式的区别是回车换行的处理，ASCII 模式将回车换行转换为本机的回车字符，比如 Unix 下是\n，Windows 下是\r\n，Mac 下是\r。假定用户正在拷贝的文件包含 ASCII 码文本，如果在远程机器上运行的不是 UNIX，当文件传输时 FTP 通常会自动地调整文件的内容以便于把文件解释成对端计算机存储文本文件的格式。
- 二进制传输模式：也称为 Binary 模式，用于传输程序文件（比如后缀名为.app、.bin 和.btm 的文件），用来传送可执行文件，压缩文件，和图片文件，不对数据进行任何处理，比文本模式更快，可以传输所有 ASCII 值，保证不出错。

FTP 工作方式

FTP 的两种工作方式：

图 5-2

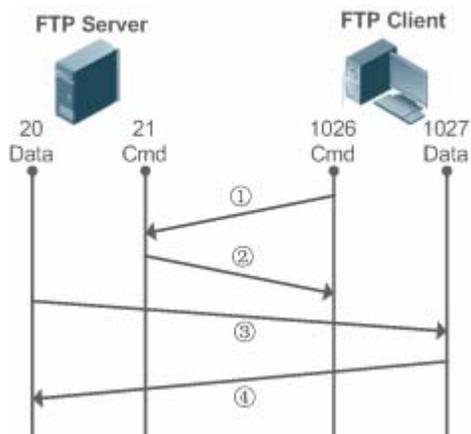
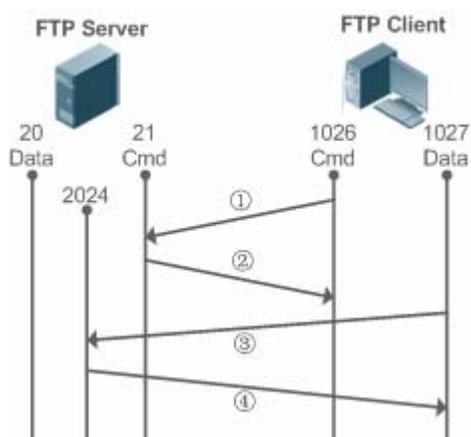


图 5-3



- PORT模式见 图 5-2 : FTP 客户端首先通过端口(1026)和FTP服务器的端口(21)建立连接,通过这个通道发送命令,客户端需要接收数据的时候在这个通道上发送PORT命令。PORT命令包含了客户端数据通道端口(1027)来接收数据。在传送数据的时候,服务器端通过自己的端口(20)连接至客户端的端口(1027)建立数据通道,实现数据收发;FTP Server必须和客户端建立一个新的连接用来传送数据。
- PASV模式见 图 5-3 :在建立控制通道的时候与PORT模式类似,但建立连接后发送的不是Port命令,而是PASV命令。FTP服务器收到PASV命令后,随机打开一个高端端口(2024)并且通知客户端在该端口上传送数据,客户端用端口(1027)连接FTP服务器该端口,之后便可以在通道上进行数据收发,这个时候FTP Server不再需要建立一个新的和客户端之间的连接。

支持的 FTP 命令

当设备收到 FTP 连接请求时, FTP 服务器将要求客户端提供登录用户名和密码以进行身份认证。

如果客户端通过身份认证,即可执行 FTP 客户端命令进行操作。目前的 FTP 服务器并没有支持所有的 FTP 命令,具体支持的 FTP 客户端命令如下:

ascii	delete	mdelete	mput	quit	send
bin	dir	mdir	nlist	recv	size
bye		mget		rename	system

cd	get	mkdir	passive		type
cdup		mls	put	rmdir	user
close	ls		pwd		

以上 FTP 客户端命令的用法请参考您所使用的 FTP 客户端软件的文档。另外不少 FTP 客户端工具（如 CuteFTP、FlashFXP 等）均提供了图形化的操作界面，使用此类工具可以无需再通过 FTP 命令进行操作。

功能特性

功能特性	作用
开启FTP Server服务	为 FTP-Client 提供上传、下载、显示文件、创建文件、删除文件等功能

5.3.1 开启FTP Server服务

工作原理

基本工作原理如上一章所述，我司设备需要配置用户名、密码、顶层目录即可为用户提供 FTP 服务。

相关配置

✚ 全局使能 FTP Server

缺省情况下，全局不开启 FTP 服务器

使用 **ftp-server enable** 开启

必须在全局开启 FTP 服务器功能，否则无法使用

✚ 配置用户名密码及顶层目录

缺省情况下，无用户授权及顶层目录

使用 **ftp-server password**、**ftp-server username**、**ftp-server topdir** 来设置授权与顶层目录

以上三项必须配置无配置无法启动 FTP 服务器功能

5.4 配置详解

配置项	配置建议&相关命令
配置FTP Server基本功能	 必须配置，用于启动 FTP 服务器。
	ftp-server enable 启动 FTP 服务器功能
	ftp-server login timeout 配置 FTP 登陆有效时长
	ftp-server login times 配置 FTP 登陆有效次数

<code>ftp-server topdir</code> <i>directory</i>	配置 FTP 服务器顶层目录
<code>ftp-server username</code> <i>username</i>	设置用户名为 <i>username</i>
<code>ftp-server password</code> [<i>type</i>] <i>password</i>	指定的 FTP 用户名相关联的缺省口令
 可选配置	
<code>ftp-server timeout</code> <i>time</i>	配置 FTP 会话的空闲时限

5.4.1 配置FTP Server基本功能

配置效果

- 建立 FTP Server，向 FTP Client 提供 FTP 服务

注意事项

- 需要配置用户名、密码及顶层访问目录
- 如果需要服务器在有限时间内关闭异常的会话，需要配置会话空闲时限

配置方法

启动 FTP Server 功能

- 必须配置
- 若无特殊要求，应在每台路由器上启动 FTP Server 功能

配置顶层目录

- 必须配置
- 若无特殊要求，应每台路由器上配置顶层目录为根目录

配置登录用户名和密码

- 必须配置
- 注意用户名和密码的长度有限制

配置会话空闲时限

- 可先配置
- 当前 FTP 服务器只支持一个用户在线，如果该用户连接异常中断或用户非正常中断连接，FTP 服务器可能无法知道用户断开而将继续保持连接，导致与服务器的连接被长期占用使服务器无法响应其他用户的登录请求，因此可以配置该选项保证异常发生时在一定时间段内让其它用户可连接上

检验方法

利用 FTP 客户端与服务器进行连接

- 检查客户端是否能连接成功
- 检查客户端相关操作是否正常

相关命令

启动 FTP Server 功能

【命令格式】 **ftp-server enable**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 在正确配置服务器的顶层目录、登录用户名和密码之前客户端仍然无法访问 FTP 服务器，因此建议在首次启动服务之前先参考后面的章节完成服务器顶层目录、登录用户名与密码的配置

配置会话登陆有效次数

【命令格式】 **ftp-server login times times**

【参数说明】 *times* : 有效次数 (范围 : 1-10)

【命令模式】 全局模式

【使用指导】 会话的有效次数是指在一个 FTP 会话在登陆过程中，用户最多可以进行账号密码认证的次数。默认设置为 3 次，即在累计三次输入错误的用户名或密码时，会话被中止，从而允许其他用户上线。

配置会话登陆有效时长

【命令格式】 **ftp-server login timeout timeout**

【参数说明】 *timeout* : 登陆有效时间 (单位 : 分钟 ; 范围 : 1-30)

【命令模式】 全局模式

【使用指导】 登陆有效时间是指用户建立链接后，每次认证用户账号和密码的最长在线时间。在该有效时间内用户若未再次进行用户密码认证将被中止会话，从而保证其他用户能够登陆。

配置服务器顶层目录

【命令格式】 **ftp-server topdir directory**

【参数说明】 *directory*: 指定用户访问路径

【命令模式】 全局模式

【使用指导】 如可以指定服务器的顶层目录为 “/syslog” 目录，则 FTP 客户端登录后将仅能访问设备上 “/syslog” 目录下的文件和文件夹，客户端由于顶层目录的限制将无法退到 “/syslog” 目录的上级目录中

配置服务器登录用户名

【命令格式】 **ftp-server username***username*

【参数说明】 *username* : 用户名

【命令模式】 全局模式

【使用指导】 FTP 服务器不支持匿名用户，因此需要配置用户名

用户名最大长度为 64 个字符，中间不允许有空格。用户名可以由英文字母、半角数字和半角符号组成

配置服务器登录密码

【命令格式】 **ftp-server password** [*type*] *password*

【参数说明】 *type* : 0 或 7, 0 代表密码未加密（明文），7 代表密码为加密过的密文

password : 密码

【命令模式】 全局模式

【使用指导】 最多只能配置一个密码

密码必须为字母或数字，密码前后可以有空格，但将被忽略；密码中间的空格作为密码的一部分

明文密码的最小长度为 1 个字符、最大长度为 25 个字符；密文密码的最小长度为 4 个字符、最大长度为 52 个字符

配置会话空闲时限

【命令格式】 **ftp-Server timeout** *time*

【参数说明】 *time* : 空闲时限（单位：分钟；范围：1-3600）

【命令模式】 全局模式

【使用指导】 会话的空闲时间是指在一个 FTP 会话中从上次 FTP 操作完成后到下次 FTP 操作开始之间的时间。服务器在响应完一个 FTP Client 命令操作后（如一个文件全部传输完毕后）重新开始计算会话空闲时间；在下一个 FTP Client 命令操作到来的时停止计算会话空闲时间。因此会话空闲时限的配置并不会对某些耗时的文件传输操作带来任何影响

查看服务器的状态信息

【命令格式】 **show ftp-server**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 显示 FTP 服务器的相关状态信息

打开服务器的调试信息

【命令格式】 **debug ftp-server pro/err**

【参数说明】 -

【命令模式】 特权模式

【使用指导】 打开 FTP 服务器的过程/错误调试信息输出

配置举例

i 以下配置举例，仅介绍与 FTP Server 相关的配置。

在 IPv4 网络上建立 FTP Server 服务

- 【配置方法】
- 开启 FTP Server 服务
 - 配置顶层目录/syslog
 - 配置用户名为 user、密码为 password
 - 配置会话空闲时限为 5 分钟

```
Ruijie(config)#ftp-server username user
Ruijie(config)#ftp-server password password
Ruijie(config)#ftp-server timeout 300
Ruijie(config)#ftp-server topdir /
Ruijie(config)#ftp-server enable
```

【检验方法】 **1.show ftp-server 查看**

```
Ruijie#sho ftp-server
ftp-server information
=====
enable : Y
topdir : /
timeout: 30min
username config : Y
password config : Y
transfer type: ASCII
control connection : N
port data connection : N
passive data connection : N
Ruijie#
```

常见错误

- 未配置用户名
- 未配置密码
- 未配置顶层目录

5.5 监视与维护

清除各类信息

查看运行情况

作用	命令
查看 FTP Server 配置	show ftp-server

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 FTP Server 错误事件的调试开关。	debug ftp-server err
打开 FTP Server 消息事件的调试开关。	debug ftp-server pro

6 FTP Client

6.1 概述

FTP (File Transfer Protocol, 文件传输协议), 是 TCP/IP 的一种具体应用, 通过在 FTP 客户端和服务器之间建立面向连接的, 可靠的 TCP 连接, 用户可以访问一个运行有 FTP 服务器程序的远程计算机。

FTP Client 为用户提供在设备上通过 FTP 协议与远程 FTP 服务器进行文件传输的功能。用户通过客户端向服务器发出命令, 服务器响应命令并把执行结果返回客户端, 通过这种命令交互, 用户可以察看服务器目录下的文件, 并把文件从远程计算机上拷到本地, 或把本地的文件传送到远程计算机去。

FTP 主要是作用是: 促进程序/数据文件的共享; 鼓励 (通过程序) 使用远程计算机; 使用户不必面对不同主机上不同文件系统的差异; 对数据进行高效可靠的传输。适用于远程安全的文件传输。

锐捷 FTP Client 并不像标准 FTP 客户端一样实现交互式命令, 其控制连接相关的 open、user、pass 指令由 CLI 输入 copy 命令自动完成, 在控制连接建立完成后, 则进入文件传输过程, 建立数据连接, 实现文件的上传或下载。

i 用于原来的设备支持 TFTP, 但是 TFTP 是用于小文件传输, FTP 协议支持大文件传输, 在设备上实现文件传输协议 FTP, 使设备可以同其它客户机或服务器进行文件传输。

协议规范

- RFC959 : FILE TRANSFER PROTOCOL (FTP)

6.2 典型应用

典型应用	场景描述
从本地上传一个文件到远程服务上	本地与远程的文件需要共享, 如需要从本地上传一个文件到 远程服务上
从远程服务器中下载一个文件到本地设备	本地与远程的文件需要共享, 如需要从远程服务器中下载一个文件到本地设备。

6.2.1 从本地上传一个文件到 远程服务上

应用场景

本地与远程的文件需要共享, 如需要从本地上传一个文件到 远程服务上。

以下图为例, 仅在 Intranet 提供共享资源作用。

图 6-1



功能部属

- 在 Intranet 中只实现通信。
- FTP Client 打开 FTP Client 文件上传功能。
- FTP Server 打开 FTP Server 文件上传功能。

6.2.2 从远程服务器中下载一个文件到本地设备

应用场景

本地与远程的文件需要共享，如需要从远程服务器中下载一个文件到本地设备。

以下图为例，仅在 Intranet 提供共享资源作用。

图 6-2



功能部属

- 在 Intranet 中只实现通信。
- FTP Client 打开 FTP Client 文件下载功能。
- FTP Server 打开 FTP Server 文件下载功能。

6.3 功能详解

基本概念

📄 FTP 文件上传

从 FTP Client 上把文件上传到 FTP Server 上。

📄 FTP 文件下载

把 FTP Server 上的文件下载到 FTP Client 上。

📌 FTP 连接模式

FTP Client 与 FTP Server 的连接方式，有主动连接和被动连接之分。

📌 FTP 传输模式

FTP Client 与 FTP Server 的之间的传输数据的方式，FTP 的传输有两种方式：文本（ASCII）传输模式和二进制（BINARY）数据传输模式。

📌 FTP 传输指定源接口 IP

FTP Client 可以对与服务端进行通信的客户端源 IP 地址进行绑定。

功能特性

功能特性	作用
FTP文件上传	从 FTP Client 上把文件上传到 FTP Server 上
FTP文件下载	将 FTP Server 上的文件下载到 FTP Client 上
FTP连接模式	FTP Client 与 FTP Server 的连接方式
FTP传输模式	FTP Client 与 FTP Server 的之间的传输数据的方式
FTP传输指定源接口IP	FTP Client 可以对与服务端进行通信的客户端源 IP 地址进行绑定

6.3.1 FTP文件上传

FTP 具有文件上传的功能。进行 FTP 文件上传文件需要 FTP Client 与 FTP Server 两个设备同时打开，从 FTP Client 上把文件上传到 FTP Server 上。

6.3.2 FTP文件下载

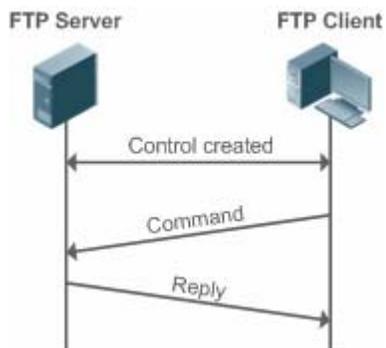
FTP 具有文件下载的功能。进行 FTP 文件下载文件需要 FTP Client 与 FTP Server 两个设备同时打开，把 FTP Server 上的文件下载到 FTP Client 上。

6.3.3 FTP连接模式

FTP 协议要用到两个 TCP 连接，一个是控制链路（也称命令链路），用来在 FTP 客户端与服务器之间传递命令；另一个是数据链路，用来上传或下载数据。

1. 控制连接：对于一些比较简单的连接只需要建立控制连接，客户端向服务器发送命令，服务器接收到命令则进行命令响应，其过程如下：

图 6-3 控制连接



2. 控制连接与数据连接：当客户端发出的命令需要上传或下载数据时，这时不仅要建立控制连接还需要建立数据连接。

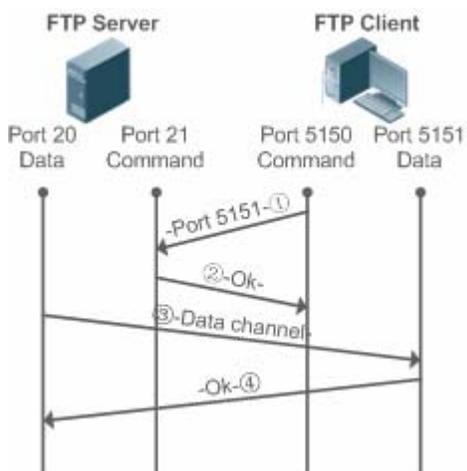
FTP 协议有两种数据连接方式：主动（PORT）方式和被动（PASV）方式。这两种工作模式主要区别在于数据连接建立方式不同，控制连接基本是一样的。

● 主动方式

该模式下 FTP server 在数据连接时是主动去连接 FTP client，所以被称为主动连接，其主要执行如下四个步骤：

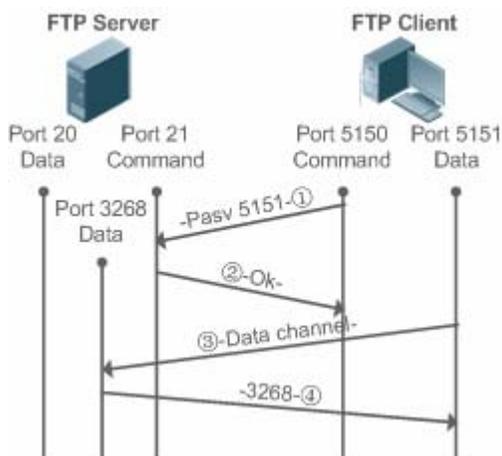
1. 客户端使用图例中的源端口 5150 与 server 端的 21 端口通信，请求建立连接，告诉服务器将用的端口是 5151。
2. server 收到后，发送应答信息，OK(ACK)，client and server 通过控制端口交换控制信令。
3. 服务器打开 20 端口作为数据发送的源端口，向客户端的 5151 端口发送。
4. 客户端应答，传输过程结束。

图 6-4 PORT（主动）模式



● 被动方式

图 6-5 PORT（主动）模式



该模式一般通过 passive 命令进行设置，由于 FTP server 在数据连接时是被动连接 FTP Client，所以称为被动连接，其主要执行如下四个步骤：

1. 被动模式下，客户端初始化控制信令连接，使用图例中 5150 源端口与服务器的 21 端口建立连接，并使用 passive 命令请求进入被动模式。
2. 服务器同意进入 PASV 模式，并随机选择一个大于 1024 的端口号，告知客户端。
3. 客户端接收到此信息后，使用图例中的 5151 端口与刚才服务器提供的 3268 端口进行数据通信，这里 5151 是源端口，3268 是目的端口。
4. 服务器收到信息，回传数据并发送应答 ACK (OK)。

当客户端和服务器建立数据连接后，就可以进行 FTP 最基本的上传和下载功能，并且在客户端可以对服务器进行一些相关文件操作。

i 用于传输命令和反馈信息的传输的控制连接始终存在，而数据连接只在需要的时候建立；PASV 和 PORT 模式的设置选择权仅在 FTP Client，由 FTP Client 发出命令建立不同的数据连接模式，我司 FTP Client 默认方式为被动模式

6.3.4 FTP传输模式

FTP 的传输有两种方式：文本 (ASCII) 传输模式和二进制 (BINARY) 数据传输模式。我司产品 FTP Client 目前支持 ASCII 和 BINARY 两种传输模式，默认情况下为 BINARY 传输模式。

- 文本模式

ASCII 模式和 BINARY 模式的区别是回车换行的处理，ASCII 模式将回车换行转换为本机的回车字符，比如 Unix 下是 \n，Windows 下是 \r\n，Mac 下是 \r。

- 二进制模式

BINARY 模式可用来传送可执行文件，压缩文件和图片文件，不对数据进行任何处理。以 Unix 传送文本文件到 Windows 为例，使用 BINARY 模式时，不会对 Unix 下的换行符进行从 \r 到 \r\n 的转换，因此在 windows 上看这个文件是没有换行的，里面是一个个的黑方块。由于不进行回车换行的处理，因此 BINARY 模式比文本模式更快，可以传输所有 ASCII 值，保证不出错。

6.3.5 FTP传输指定源接口IP

FTP Client 可以对与服务端进行通信的客户端源 IP 地址进行绑定，这样可以用指定的源 IP 与 FTP Server 进行连接和传输文件。

6.4 配置详解

配置项	配置建议 & 相关命令	
配置FTP Client基本功能	 必须配置。配置 FTP CLEINT 功能	
	copy flash	文件上传
	copy ftp	文件下载
配置FTP Client可选功能	 可选配置。配置 FTP CLEINT 功能的工作模式	
	ftp-client port	设置 FTP 为主动连接模式
	ftp-client ascii	设置 FTP 为文本传输模式
	ftp-client source-address	配置进行 FTP 连接的客户端源 IP 地址
	default ftp-client	恢复 FTP Client 为缺省配置，数据连接为被动方式，文件传输为二进制模式，清除源 IP 绑定

6.4.1 配置FTP Client基本功能

配置效果

- 实现文件上传与下载。

注意事项

- 文件上传与下载的格式。

配置方法

📄 文件上传

- 需要实现文件上传时，为必选配置。
- 在特权模式下的 copy 下的目的地址上配置 ftp 相关的 url。

📄 文件下载。

- 需要实现文件下载时，为必选配置。
- 在特权模式下的 copy 下的源地址上配置 ftp 相关的 url。

检验方法

- 在 FTP Server 的目录中看所上传的文件是否存在。
- 在目的地址上查看下载的文件是否存在。

相关命令

文件上传

【命令格式】 **copy flash:[local-directory/]local-file**

ftp: //username:password@dest-address[/remote-directory]/remote-file

【参数说明】 *local-directory* : 指定设备目录，如果未指定，则表示当前工作目录。

local-file : 表示要操作的本地文件名

username : 指定访问 FTP Server 的用户名，最长不超过三十二个字节，不可包含 “:”、“@”、“/” 和空格等字符，不可省略。

Password : 指定访问 FTP Server 的密码，最长不超过三十二个字节，不可包含 “:”、“@”、“/” 和空格等字符，不可省略。

dest-address : 指定 FTP Server 的 IP 地址

remote-directory : 指定 Server 端的目录路径

remote-file : 指定要操作的 Server 端文件名

 如果包含 *local-directory* 字段，则必须保证设备中已创建了该目录，此下载命令不支持目录的自动创建。

【命令模式】 全局模式

【使用指导】 使用该命令从本地设备的 flash 上上传一个文件到 FTP SERVER 上去。

文件下载

【命令格式】 **copy ftp://username:password@dest-address[/remote-directory]/remote-file**

flash:[local-directory/]local-file

【参数说明】 *username* : 指定访问 FTP Server 的用户名，最长不超过三十二个字节，不可包含 “:”、“@”、“/” 和空格等字符，不可省略。

password : 指定访问 FTP Server 的密码，最长不超过三十二个字节，不可包含 “:”、“@”、“/” 和空格等字符，不可省略。

dest-address : 指定 FTP Server 的 IP 地址。

remote-directory : 指定 Server 端的目录路径。

remote-file : 指定要操作的 Server 端文件名。

local-directory : 指定设备目录，如果未指定，则表示当前工作目录。

local-file : 表示要操作的本地文件名。

 如果包含 *local-directory* 字段，则必须保证设备中已创建了该目录，此下载命令不支持目录的自动创建。

【命令模式】 全局模式

【使用指导】 使用该命令从 FTP SERVER 下载一个文件到本地设备的 flash 上去。

配置举例

 以下配置举例，仅介绍与 FTP Client 上传下载相关的配置。

上传文件示例

【配置方法】 将设备 home 目录中的 local-file 文件上传到用户名为 user，密码为 pass，IP 地址为 192.168.23.69 的 FTP Server 的 root 目录下，文件命名为 remote-file。

```
Ruijie# copy flash: home/local-file ftp://user:pass@192.168.23.69/root/remote-file
```

【检验方法】 在 FTP SERVER 上查看 remote-file 是否存在。

下载文件示例

【配置方法】 从用户名为 user，密码为 pass，IP 地址为 192.168.23.69 的 FTP Server 的 root 目录下载文件名为 remote-file 的文件到设备上的 home 目录中，存储的文件名为 local-file

```
Ruijie# copy ftp://user:pass@192.168.23.69/root/remote-file flash: home/local-file
```

【检验方法】 在 flash 的 home 目录下查看 remote-file 是否存在。

常见配置错误

- 上传下载输入的格式错误。
- 用户名或密码错误。

6.4.2 配置FTP Client可选功能

配置效果

- 根据配置能让 FTP 工作在指定连接、传输模式及指定的 IP 地址下进行文件上传与下载。

注意事项

- -

配置方法

📌 设置 FTP 为主动连接方式。

- 可选配置。
- 配置 FTP 的连接模式。

📌 设置 FTP 为文本传输模式。

- 可选配置。
- 配置 FTP 的传输模式。

📌 设置进行 FTP 连接的客户端源 IP 地址。

- 可选配置。
- 配置进行 FTP 连接的客户端源 IP 地址。

📌 恢复 FTP Client 为缺省配置。

- 可选配置。
- 恢复 FTP Client 为缺省配置。

检验方法

通过 **show run** 查看

相关命令

📌 配置 FTP 为主动连接模式

【命令格式】 **ftp-client port**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 使用该命令可以将连接模式设置为主动方式，主动方式下，服务器主动去连接客户端。默认情况下 FTP 连接为被动 (PASV) 方式。

📌 配置进行 FTP 连接的客户端源 IP 地址

【命令格式】 **ftp-client source-address {ip-address }**

【参数说明】 *ip-address* : 本地接口的 ipv4 地址。

【命令模式】 全局配置模式

【使用指导】 使用该命令可以绑定客户端不同的接口 IP 地址，使客户端使用此 IP 地址连接服务器。默认情况下客户端不进行本地 IP 绑定，由路由进行选择。

📌 设置 FTP 为文本传输模式

- 【命令格式】 **ftp-client ascii**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 使用该命令可以将文件传输方式设置为文本（ASCII）方式。默认情况下 FTP 传输模式为二进制（BINARY）方式。

▾ 恢复 FTP Client 为缺省配置

- 【命令格式】 **default ftp-client**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 恢复 FTP Client 为缺省配置，数据连接为被动方式，文件传输为二进制模式，清除源 IP 绑定。

配置举例

 以下配置举例，仅介绍与 FTP Client 可选项相关的配置。

▾ 可选项配置

- 【配置方法】
- 配置 FTP 连接模式为：port
 - 配置传输模式为：ASCII
 - 配置源 IP 为 192.168.23.167
 -

```
Ruijie# configure terminal
Ruijie(config)# ftp-client ascii
Ruijie(config)# ftp-client port
Ruijie(config)# ftp-client source-address 192.168.23.167
Ruijie(config)# end
```

- 【检验方法】 在设备上进行 **show run**，能看到以下信息

```
Ruijie# show run

!
ftp-client ascii
ftp-client port
ftp-client source-address 192.168.23.167
!
```

常见配置错误

- 源 IP 不是本地 IP。

6.5 监视与维护

清除各类信息

无

查看配置情况

作用	命令
查看 FTP Client 的配置	show run

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 FTP Client 调试开关。	debug ftp-client

7 网络通信检测工具

7.1 概述

网络通信检测工具可以用于检查网络是否能够连通，用好网络通信监测工具可以很好地帮助我们分析判定网络故障。网络通信检测工具包括 PING（Packet Internet Groper，因特网包探索器）和 Traceroute（路由侦测）。PING 工具主要用于检测网络通与不通，以及网络的时延，时延值越大，则表示网络速度越慢。Traceroute 工具则可以帮助用户了解网络的物理与逻辑连接的拓扑情况以及数据传输的效率。在网络设备上，这两个工具所对应的命令为 ping 和 traceroute。

协议规范

- RFC792：Internet Control Message Protocol
- RFC4443：Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

7.2 典型应用

典型应用	场景描述
端对端连通性检查	网络设备与目标主机都连接在 IP 网络上，都配置有 IP 地址。
主机路由检查	网络设备与目标主机都连接在 IP 网络上，都配置有 IP 地址。

7.2.1 端对端连通性检查

应用场景

图 7-1 网络设备 A 与目标主机 B 都连接在 IP 网络上。

网络设备与目标主机都连接在 IP 网络上，端对端连通性检查就是判定 IP 报文能否在二者之间传输。目标主机可以是网络设备本身，这种情况一般用于检查设备自身网络接口和 TCP/IP 协议配置的正确性。



功能部属

通过在网络设备上运行 Ping 功能。

7.2.2 主机路由检查

应用场景

图 7-2 网络设备 A 与目标主机 B 都连接在 IP 网络上。

网络设备与目标主机都连接在 IP 网络上，主机路由检查就是判定 IP 报文在二者之间传输，究竟需要经过多少网关（路由器）。目标主机通常不是网络设备本身，并且通常与网络设备不在同一个 IP 网段。



功能部属

通过在网络设备上运行 Traceroute 功能。

7.3 功能详解

功能特性

功能特性	作用
Ping连通性测试	检测指定 IPv4 地址是否可达，并输出相关信息。
Traceroute连通性测试	显示 IPv4 数据包从源地址到目的地址所经过的网关。

7.3.1 Ping连通性测试

工作原理

PING 工具向目标 IP 地址发送一个 ICMP 请求（ICMP Request）数据包，要求对方返回一个 ICMP 回声（ICMP Echo）数据包，来确定两台网络机器是否连接相通，时延是多少。

相关配置

- 通过 ping 命令进行配置

7.3.2 Traceroute连通性测试

工作原理

Traceroute 工具利用 ICMP 及 IP 报文头部的 TTL (Time To Live) 字段。首先, 网络设备的 Traceroute 工具送出一个 TTL 是 1 的 ICMP Request 到目的主机, 当路径上的第一个路由器收到这个报文时, 它将 TTL 减 1。此时 TTL 变为 0 了, 所以该路由器会将此报文丢弃, 并送回一个 ICMP 超时 (ICMP time exceeded) 消息, Traceroute 工具收到这个消息后, 便知道这个路由器存在于这个路径上, 接着再送出另一个 TTL 是 2 的报文, 发现第 2 个路由器。Traceroute 工具每次将送出的报文的 TTL 加 1 来发现另一个路由器, 这个重复的动作一直持续到某个数据报文到达目的主机。当报文到达目的主机后, 该主机不会送回 ICMP time exceeded 消息, 而是送回 ICMP Echo, Traceroute 工具结束探测并显示从网络设备到目的主机的路径信息。

相关配置

- 通过 `traceroute` 命令进行配置

7.4 配置详解

配置项	配置建议 & 相关命令	
Ping连通性测试	 可选配置, 用于检测 IPv4 地址是否可达。	
	<code>ping</code>	运行 Ping 功能。
Traceroute连通性测试	 可选配置, 显示 IPv4 数据包从源地址到目的地址所经过的网关。	
	<code>traceroute</code>	运行 Traceroute 功能。

7.4.1 Ping连通性测试

配置效果

在网络设备上采用 Ping 连通性测试, 可以得知该网络设备和目的主机之间是否保持连通, 报文是否可以在网络设备和目的主机之间传输。

注意事项

执行 PING 操作的网络设备本身需要配置 IP 地址。

配置方法

- 如果需要检测 IPv4 地址是否可达, 可通过 `Ping IPv4` 命令。

检验方法

输入 **ping** 命令，即可在 CLI 界面显示相关信息。

相关命令

📄 Ping IPv4

【命令格式】 **ping** [**oob** | **ip**] [*address* [**via** *mgmt-name*] [**length** *length*] [**ntimes** *times*] [**timeout** *seconds*] [**data** *data*] [**source** *source*] [**df-bit**] [**validate**] [**detail**]

【参数说明】 **oob**：设置使用带外通道。当指定 MGMT 口作为源接口时，必须设置该参数。

address：指定目的 IPv4 地址或域名。

mgmt-name：指定在 oob 模式下报文的出口管理口。

length：指定发送数据包数据填充段的长度，范围：36~18024，默认填充长度为 100。

times：指定发送数据包的个数，范围：1~ 4294967295。

seconds：指定超时的时间，范围：1~10（秒）。

data：指定报文填充数据，格式为 1-255 长度的字符串，默认填充为 abcd。

source：指定报文源 IPv4 地址或源接口。其中，环回接口地址（例如 127.0.0.1）不允许作为源地址。

df-bit：设置 IP 的 DF 标识位，当 DF 位被设置为 1 时，表示不对数据包进行分段处理，默认 DF 位为 0。

validate：设置是否校验响应报文。

detail：设置回显是否显示详细信息，默认只显示 ‘!’ 和 ‘.’ 。

【命令模式】 在普通用户模式下，只能运行基本的 **ping** 功能；在特权用户模式下，还可以运行 **ping** 的扩展功能。

在其他模式下，可以通过 do 命令执行 **ping** 的扩展功能，具体配置请参考 do 命令说明。

【使用指导】 运行 **ping** 功能，如果有应答，则显示出应答的相关信息，最后输出一个统计信息。在扩展 **ping** 中，可以指定发送数据包的个数、长度、超时的时间等等，和基本的 **ping** 功能一样，最后也输出一个统计信息。要使用域名功能，则要先配置域名服务器，具体配置请参考 DNS 配置部分。

配置举例

📄 运行普通 Ping 功能

【配置方法】 在特权模式下输入 Ping IPv4 地址 192.168.21.26

```
常规 ping
Ruijie# ping 192.168.21.26
Sending 5, 100-byte ICMP Echoes to 192.168.21.26, timeout is 2 seconds:
< press Ctrl+C to break >
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/10 ms
显示 detail 的 ping
Ruijie#ping 192.168.21.26 detail
Sending 5, 100-byte ICMP Echoes to 192.168.21.26, timeout is 2 seconds:
```

```

< press Ctrl+C to break >
Reply from 192.168.21.26: bytes=100 time=4ms TTL=64
Reply from 192.168.21.26: bytes=100 time=3ms TTL=64
Reply from 192.168.21.26: bytes=100 time=1ms TTL=64
Reply from 192.168.21.26: bytes=100 time=1ms TTL=64
Reply from 192.168.21.26: bytes=100 time=1ms TTL=64

Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/4 ms.

```

【检验方法】 缺省将 5 个数据段长度为 100Byte 的数据包发送到指定的 IP 地址，在指定的时间（缺省为 2 秒）内，显示相应的探测信息，最后输出一个统计信息。

运行扩展 Ping 功能

【配置方法】 在特权模式下输入 Ping IPv4 地址 192.168.21.26，并指定发送数据包的长度、个数、超时的时间等。

```

常规 ping
Ruijie# ping 192.168.21.26 length 1500 ntimes 100 data ffff source 192.168.21.99 timeout 3
Sending 100, 1500-byte ICMP Echoes to 192.168.21.26, timeout is 3 seconds:
< press Ctrl+C to break >
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!
Success rate is 100 percent (100/100), round-trip min/avg/max = 2/2/3 ms

显示 detail 的 ping
ping 192.168.21.26 length 1500 ntimes 20 data ffff source 192.168.21.99 timeout 3 detail
Sending 20, 1500-byte ICMP Echoes to 192.168.21.26, timeout is 3 seconds:
< press Ctrl+C to break >
Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64
Reply from 192.168.21.26: bytes=1500 time=2ms TTL=64
Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64

```

```
Reply from 192.168.21.26: bytes=1500 time=3ms TTL=64
Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64
Reply from 192.168.21.26: bytes=1500 time=1ms TTL=64

Success rate is 100 percent (20/20), round-trip min/avg/max = 1/1/3 ms.
```

【检验方法】 将 20 个长度为 1500Byte 的数据包发送到指定的 IP 地址，在指定的时间（3 秒）内，如果有应答，显示相应的探测信息，最后输出一个统计信息。

7.4.2 Traceroute 连通性测试

配置效果

在网络设备上采用 Traceroute 连通性测试，可以得知该网络设备和目的主机之间的路由拓扑信息，报文从网络设备到目的主机经过了多少个网关。

注意事项

执行 Traceroute 操作的网络设备本身需要配置 IP 地址。

配置方法

- 如果需要跟踪 IPv4 数据包到达目的主机经过哪些网关，可通过配置 Traceroute IPv4 命令。

检验方法

输入 `traceroute` 命令，即可在 CLI 界面显示相关信息。

相关命令

Traceroute IPv4

【命令格式】 `traceroute [oob | ip] [address [via mgmt-name] [probe number] [source source] [timeout seconds] [ttl minimum maximum]]`

【参数说明】 `oob`：设置使用带外通道。当指定 MGMT 口作为源接口时，必须设置该参数。

`address`：指定目的 IPv4 地址或域名。

`mgmt-name`：指定在 oob 模式下报文的出口管理口。

`number`：指定发送的探测的数量，范围：1~255。

`source`：指定报文源 IPv4 地址或源接口。其中，环回接口地址（例如 127.0.0.1）不允许作为源地址

`seconds`：指定超时的时间，范围：1~10（秒）。

minimum maximum : 指定最小和最大 TTL 值, 范围: 1~255。

- 【命令模式】 在普通用户模式下, 只能运行基本的 **tracert** 功能; 在特权用户模式下, 还可以运行 **tracert** 的扩展功能。
- 【使用指导】 **Tracert** 命令主要用于检查网络的连通性, 并在网络故障发生时, 准确的定位故障发生的位置。要使用域名功能, 则要先配置域名服务器, 具体配置请参考 DNS 配置部分。

配置举例

网络畅通的 Tracert 举例

- 【配置方法】 在特权模式下, 输入 Tracert IPv4 地址 61.154.22.36。

```
Ruijie# tracert 61.154.22.36
< press Ctrl+C to break >
Tracing the route to 61.154.22.36
 0  192.168.12.1          0 msec  0 msec  0 msec
 1  192.168.9.2           4 msec  4 msec  4 msec
 2  192.168.9.1           8 msec  8 msec  4 msec
 3  192.168.0.10          4 msec  28 msec 12 msec
 4  202.101.143.130       4 msec  16 msec  8 msec
 5  202.101.143.154      12 msec  8 msec  24 msec
 6  61.154.22.36         12 msec  8 msec  22 msec
```

从上面的结果可以清楚地看到, 从源地址要访问 IP 地址为 61.154.22.36 的主机, 网络数据包都经过了哪些网关 (1 - 6), 同时给出了到达该网关所花费的时间。

网络中某些网关不通的 Tracert 举例

- 【配置方法】 在特权模式下, 输入 Tracert IPv4 地址 202.108.37.42。

```
Ruijie# traceroute 202.108.37.42
< press Ctrl+C to break >
Tracing the route to 202.108.37.42
 1  192.168.12.1      0 msec  0 msec  0 msec
 2  192.168.9.2       0 msec  4 msec  4 msec
 3  192.168.110.1    16 msec 12 msec 16 msec
 4  * * *
 5  61.154.8.129     12 msec 28 msec 12 msec
 6  61.154.8.17      8 msec 12 msec 16 msec
 7  61.154.8.250     12 msec 12 msec 12 msec
 8  218.85.157.222   12 msec 12 msec 12 msec
 9  218.85.157.130   16 msec 16 msec 16 msec
10  218.85.157.77    16 msec 48 msec 16 msec
11  202.97.40.65     76 msec 24 msec 24 msec
12  202.97.37.65     32 msec 24 msec 24 msec
13  202.97.38.162    52 msec 52 msec 224 msec
14  202.96.12.38     84 msec 52 msec 52 msec
15  202.106.192.226  88 msec 52 msec 52 msec
16  202.106.192.174  52 msec 52 msec 88 msec
17  210.74.176.158   100 msec 52 msec 84 msec
18  202.108.37.42    48 msec 48 msec 52 msec
```

从上面的结果可以清楚地看到，从源地址要访问 IP 地址为 202.108.37.42 的主机，网络数据包都经过了哪些网关（1 - 17），并且网关 4 出现了故障。

8 TCP

8.1 概述

TCP 协议为应用层提供了一个可靠的、有连接的基于 IP 的传输层协议。

应用层向 TCP 层发送用于网间传输的、用 8 位字节表示的数据流，然后 TCP 把数据流分割成适当长度的报文段，最大分段大小 (MSS) 通常受该计算机连接的网路的数据链路层的最大传送单元 (MTU) 限制。之后 TCP 把报文传给 IP 层，由它来通过网络将报文传送给接收端实体的 TCP 层。

TCP 为了保证不发生丢包，就给每个字节一个序号，同时序号也保证了传送到接收端实体的包的按序接收。然后接收端实体对已成功收到的字节发回一个相应的确认 (ACK)；如果发送端实体在合理的往返时延 (RTT) 内未收到确认，那么对应的数据 (假设丢失了) 将会被重传。

- 在数据正确性与合法性上，TCP 用一个校验和函数来检验数据是否有错误，在发送和接收时都要计算校验和；同时可以使用 MD5 认证对数据进行校验。
- 在保证可靠性上，采用超时重传和捎带确认机制。
- 在流量控制上，采用滑动窗口协议，协议中规定，对于窗口内未经确认的分组需要重传。

协议规范

- RFC 793 : Transmission Control Protocol
- RFC 1122 : Requirements for Internet Hosts -- Communication Layers
- RFC 1191 : Path MTU Discovery
- RFC 1213 : Management Information Base for Network Management of TCP/IP-based internets:MIB-II
- RFC 2385 : Protection of BGP Sessions via the TCP MD5 Signature Option
- RFC 4022 : Management Information Base for the Transmission Control Protocol (TCP)

8.2 典型应用

典型应用	场景描述
TCP性能优化	TCP 传输路径上某一段链路的 MTU 比较小，为了避免 TCP 报文分片，可以开启 TCP 的路径 MTU 发现功能。
TCP连接异常检测	TCP 探测对端是否还在正常工作。

8.2.1 TCP性能优化

应用场景

以下图为例，A 和 D 建立 TCP 连接，A 和 B 之间链路的 MTU 是 1500 字节，B 和 C 之间链路的 MTU 是 1300 字节，C 和 D 之间链路的 MTU 是 1500 字节，为了使 TCP 传输性能达到最优，需要避免 TCP 报文在设备 B 和设备 C 上分片。

图 8-1



【注释】 A、B、C 和 D 为路由器。

功能部署

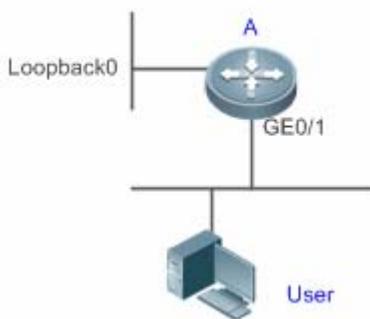
- 在 A 和 D 上开启 TCP 的路径 MTU 发现功能。

8.2.2 TCP连接异常检测

应用场景

以下图为例，用户远程登录到设备 A，用户异常关机，如果设备 A 等待 TCP 重传超时，会导致用户的 TCP 连接残留比较长的一段时间，可以利用 TCP 保活功能快速检测出用户的 TCP 连接异常。

图 8-2



【注释】 A 是路由器。

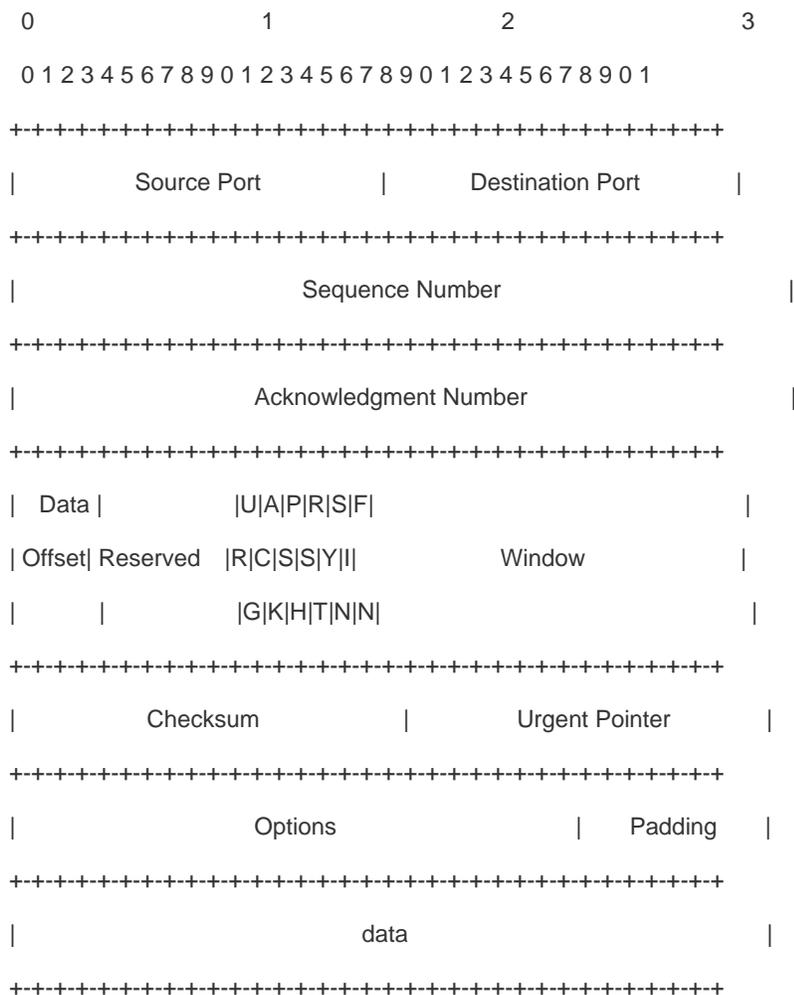
功能部署

- 在设备 A 上开启 TCP 保活功能。

8.3 功能详解

基本概念

▾ TCP 首部格式



- Source Port 是源端口，16 位。
- Destination Port 是目的端口，16 位。
- Sequence Number 是序列号，32 位。
- Acknowledgment Number 是确认序列号，32 位。

- Data Offset 是数据偏移，4 位，该字段的值是 TCP 首部（包括选项）长度除以 4。
- 标志位：6 位，URG 表示 Urgent Pointer 字段有意义，ACK 表示 Acknowledgment Number 字段有意义，PSH 表示 Push 功能，RST 表示复位 TCP 连接，SYN 表示 SYN 报文（在建立 TCP 连接的时候使用），FIN 表示发送方没有数据需要发送了（在关闭 TCP 连接的时候使用）。
- Window 表示接收缓冲区的空闲空间，16 位，用来告诉 TCP 连接对端自己能够接收的最大数据长度。
- Checksum 是校验和，16 位。
- Urgent Pointers 是紧急指针，16 位，只有 URG 标志位被设置时该字段才有意义，表示紧急数据相对序列号（Sequence Number 字段的值）的偏移。

✚ TCP 三次握手

- TCP 三次握手的过程如下：
 - (1) 客户端发送 SYN 报文给服务器端。
 - (2) 服务器端收到 SYN 报文，回应一个 SYN ACK 报文。
 - (3) 客户端收到服务器端的 SYN 报文，回应一个 ACK 报文。
- 三次握手完成，TCP 客户端和服务器端成功地建立连接，可以开始传输数据了。

功能特性

功能特性	作用
配置SYN超时	配置 TCP 发送 SYN 报文或者 SYN ACK 报文后等待应答报文的超时
配置窗口大小	配置窗口大小
配置端口不可达时是否发送 reset 报文	配置在收到端口不可达的 TCP 报文时是否发送 reset 报文
配置 MSS	配置 TCP 连接的 MSS
路径MTU发现功能	探测 TCP 传输路径上的最小 MTU，根据最小 MTU 调整发送的 TCP 报文的大小，避免分片
TCP保活功能	探测 TCP 连接对端是否还在正常工作

8.3.1 配置SYN超时

工作原理

建立 TCP 连接需要经过三次握手：发起方先发送 SYN 报文，响应方回应 SYN+ACK 报文，然后发起方再回应 ACK。

- 在发起方发送 SYN 报文后，如果响应方一直不回应 SYN+ACK 报文，发起方会不断的重传 SYN 报文直到超过一定的重传次数或超时时间。

- 在发起方发送 SYN 报文后，响应方回应 SYN+ACK 报文，但发起方不再回复 ACK，响应方也会一直重传直到超过一定的重传次数或超时时间。（SYN 报文攻击会出现这种情况。）

相关配置

设置 TCP SYN 超时时间

- TCP SYN 超时时间的缺省值是 20 秒。
- 用户可以在全局配置模式下使用“`ip tcp synwait-time seconds`”命令设置 SYN 超时时间，取值范围是 5 到 300，单位是秒。
- 如果网络中存在 SYN 攻击，减少 SYN 超时时间可以防止一些资源消耗，但对连续的 SYN 攻击达不到效果。在设备主动与外界请求建立连接时，减少 SYN 超时时间可以减少用户等待时间，如 telnet。如果网络比较差也可以适当增加超时时间。

8.3.2 配置窗口大小

工作原理

TCP 的接收缓冲区用来缓存从对端接收到的数据，这些数据后续会被应用程序读取。一般情况下，TCP 的窗口值反映接收缓冲区的空闲空间的大小。对于带宽比较大、有大量数据的连接，增大窗口可以显著提高 TCP 传输性能。

相关配置

设置窗口大小

- 用户可以在全局配置模式下使用“`ip tcp window-size size`”命令设置窗口大小，单位是字节，取值范围是 128 到(65535<<14)，缺省值是 65535。如果窗口大于 65535 字节，自动开启窗口扩大功能。
- 实际通告给对端的窗口大小是从配置的窗口大小和接收缓冲区的剩余空间取较小值。

8.3.3 配置端口不可达时是否发送 reset 报文

工作原理

TCP 协议在分发 TCP 报文给应用程序时，如果找不到该报文所属的 TCP 连接会主动回复一个 reset 报文以终止对端的 TCP 连接。攻击者可能利用大量的端口不可达的 TCP 报文对设备进行攻击。

相关配置

配置端口不可达时是否发送 reset 报文

收到端口不可达的 TCP 报文时，默认发送 reset 报文。

用户可以在全局配置模式下使用 “no ip tcp send-reset” 命令禁止发送 reset 报文。

如果允许发送 reset 报文，攻击者可能利用大量的端口不可达的 TCP 报文对设备进行攻击。

8.3.4 配置MSS

工作原理

最大分段大小 (Maximum Segment Size, MSS)，指一个 TCP 报文的数据载荷的最大长度，不包括 TCP 选项。

在 TCP 建立连接的三次握手中需要进行 MSS 协商，连接的双方都在 SYN 报文中增加 MSS 选项，其选项值表示本端最大能接收的段大小，即对端最大能发送的段大小。连接的双方取本端发送的 MSS 值和接收对端的 MSS 值的较小者作为本连接最大传输段大小。

发送 SYN 报文时 MSS 选项值的计算方法如下：

- IPv4 TCP：MSS = 对端 IP 地址对应的出口的 IP MTU - 20 字节 IP 首部 - 20 字节 TCP 首部。

i 实际生效的 MSS 是从根据 MTU 计算得到的 MSS 和用户配置的 MSS 取较小值。

i 如果该连接支持某些选项，那么 MSS 还要减去选项 4 字节对齐后的长度值。如 MD5 选项要减去 20 字节，MD5 选项长度 18 字节，对齐后 20 字节。

相关配置

设置 MSS

- 用户可以在全局配置模式下使用 “ip tcp mss max-segment-size” 命令设置 TCP 连接的 MSS，单位是字节，取值范围是 68 到 10000，默认使用根据 MTU 计算得到的 MSS。如果用户配置了 MSS，实际生效的 MSS 是从根据 MTU 计算得到的 MSS 和用户配置的 MSS 取较小值。
- MSS 太小会降低传输性能，增加 MSS 可以提高传输性能，但不是越大越好，选择 MSS 值可以参考接口的 MTU，如果 MSS 大于接口的 MTU，TCP 报文需要分片重组，会降低传输性能。

8.3.5 路径MTU发现功能

工作原理

RFC1191 规定的 TCP 连接的路径 MTU 发现功能，用来发现 TCP 报文传输路径的最小 MTU，避免分片重组，可以提高网络带宽的利用率。IPv4 TCP 路径 MTU 发现的过程如下：

- (1) TCP 源端将发送的 TCP 报文的外层 IP 首部设置不可分片标志位。

- (2) 如果 TCP 路径上某路由器的出口 MTU 值小于该 IP 报文长度，则会丢弃报文，并向 TCP 源端发送 ICMP 差错报文，报文中会携带该出口 MTU 值。
- (3) TCP 源端通过解析该 ICMP 差错报文，可知 TCP 路径上当前最小的 MTU 值，即路径 MTU。
- (4) 后续 TCP 源端发送数据段的长度不超过 MSS， $MSS = \text{路径 MTU} - \text{IP 头部长度} - \text{TCP 头部长度}$ 。

相关配置

✎ 启用路径 MTU 发现功能

TCP 缺省关闭路径 MTU 发现功能。

用户可以在全局配置模式下使用“`ip tcp path-mtu-discovery`”命令开启路径 MTU 发现功能。

 从 11.0 版本开始只对 IPv4 TCP 生效。

8.3.6 TCP保活功能

工作原理

如果 TCP 希望知道对端是否还在正常工作，可以开启保活功能。当 TCP 对端在一段时间内（称为空闲时间）没有发送过报文给本端，本端开始发送保活报文，连续发送若干次，如果没有收到一个应答报文，就认为对端异常，关闭 TCP 连接。

相关配置

✎ 启用保活功能

- TCP 缺省关闭保活功能。
- 用户可以在全局配置模式下使用“`ip tcp keepalive [interval num1] [times num2] [idle-period num3]`”命令开启保活功能。interval 是时间间隔，默认值是 75 秒；times 是发送保活报文的最大次数，默认值是 6 次；idle-period 是空闲时间，默认值是 15 分钟。

 该命令不再区分服务器端和客户端，对所有的 TCP 连接都生效。

8.4 配置详解

配置项	配置建议 & 相关命令	
TCP性能优化	 可选配置，用于优化 TCP 连接的性能。	
	<code>ip tcp synwait-time</code>	配置建立 TCP 连接的超时时间。
	<code>ip tcp window-size</code>	配置 TCP 窗口大小。

	ip tcp send-reset	配置收到端口不可达的 TCP 报文时是否发送 reset 报文。
	ip tcp mss	配置 TCP 连接的 MSS。
	ip tcp path-mtu-discovery	开启路径 MTU 发现功能。
TCP连接异常检测	 可选配置，用于检测 TCP 对端是否正常工作。	
	ip tcp keepalive	开启 TCP 保活功能。

8.4.1 TCP性能优化

配置效果

- TCP 连接的传输性能达到最优，避免分片。

注意事项

-

配置方法

配置 SYN 超时

- 可选配置。
- 在 TCP 连接的两端配置。

配置 TCP 窗口大小

- 可选配置。
- 在 TCP 连接的两端配置。

配置端口不可达时是否发送 reset 报文

- 可选配置。
- 在 TCP 连接的两端配置。

配置 MSS

- 可选配置。
- 在 TCP 连接的两端配置。

配置 TCP 的路径 MTU 发现功能

- 可选配置。

- 在 TCP 连接的两端配置。

检验方法

-

相关命令

配置 SYN 超时

【命令格式】 **ip tcp synwait-time seconds**

【参数说明】 *seconds* : SYN 报文超时时间。单位为秒，取值范围是 5 到 300，缺省值是 20。

【命令模式】 全局模式

【使用指导】 如果网络中存在 SYN 攻击，减少 SYN 超时时间可以防止一些资源消耗，但对连续的 SYN 攻击达不到效果。在设备主动与外界请求建立连接时，减少 SYN 超时时间可以减少用户等待时间，如 telnet。如果网络比较差也可以适当增加超时时间。

配置 TCP 窗口大小

【命令格式】 **ip tcp window-size size**

【参数说明】 *size* : 单位是字节，取值范围是 128 到(65535 << 14)，缺省值是 65535。

【命令模式】 全局模式

【使用指导】 -

配置端口不可达时是否发送 reset 报文

【命令格式】 **ip tcp send-reset**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 收到端口不可达的 TCP 报文时，默认发送 reset 报文。

配置 MSS

【命令格式】 **ip tcp mss max-segment-size**

【参数说明】 *max-segment-size* : MSS 的上限值。单位为字节，取值范围是 68 到 10000，默认使用根据 MTU 计算得到的 MSS。

【命令模式】 全局模式

【使用指导】 **ip tcp mss** 的作用就是限制即将建立的 TCP 连接的 MSS 的最大值。任何新建立的连接协商的 MSS 值不能超过配置的值。如果要减小连接的最大 MSS 值，可以配置该命令，一般情况下不需要配置。

配置路径 MTU 发现功能

【命令格式】 **ip tcp path-mtu-discovery [age-timer minutes | age-timer infinite]**

【参数说明】 **age-timer minutes** : TCP 在发现路径 MTU 后，重新进行探测的时间间隔。单位是分钟，取值范围是 10 到 30。缺省值是 10。

age-timer infinite : TCP 在发现路径 MTU 后, 不重新探测。

【命令模式】 全局模式

【使用指导】 TCP 的路径 MTU 发现功能是按 RFC1191 实现的, 这个功能可以提高网络带宽的利用率。当用户使用 TCP 来批量传输大块数据时, 该功能可以使传输性能得到明显提升。

按 RFC1191 的描述, TCP 在发现路径 MTU 后, 隔一段时间可以使用更大的 MSS 来探测新的路径 MTU。这个时间间隔就是使用参数 **age-timer** 来指定。当设备发现的路径 MTU 比 TCP 连接两端协商出来的 MSS 小时, 设备就会按上述配置时间间隔, 去尝试发现更大的路径 MTU。直到路径 MTU 达到 MSS 的值, 或者用户停止这个定时器, 这个探测过程才会停止。停止这个定时器, 使用 **age-timer infinite** 参数。

配置举例

✎ 开启 TCP 的路径 MTU 发现功能。

【配置方法】 在设备上开启 TCP 的路径 MTU 发现功能, 重新探测的时间间隔取缺省值。

```
Ruijie# configure terminal
Ruijie(config)# ip tcp path-mtu-discovery
Ruijie(config)# end
```

【检验方法】 用户可以执行命令 **show tcp pmtu** 查看 IPv4 TCP 连接的路径 MTU。

```
Ruijie# show tcp pmtu
```

Number	Local Address	Foreign Address	PMTU
1	192.168.195.212.23	192.168.195.112.13560	1440

常见错误

-

8.4.2 TCP连接异常检测

配置效果

- TCP 探测对端是否还在正常工作。

注意事项

-

配置方法

▾ 开启保活功能

- 可选配置。

检验方法

相关命令

▾ 开启保活功能

【命令格式】 **ip tcp keepalive [interval num1] [times num2] [idle-period num3]**

【参数说明】 **interval num1**：发送保活报文的时间间隔，单位是秒，取值范围是 1 到 120，缺省值是 75 秒。

times num2：发送保活报文的最大次数，取值范围是 1 到 10，缺省值是 6。

idle-period num3：空闲时间，即对端没有向本端发送过报文的时间长度，单位是秒，取值范围是 60 到 1800，缺省值是 15 分钟。

【命令模式】 全局模式

【使用指导】 如果 TCP 希望知道对端是否还在正常工作，可以开启保活功能，默认关闭。

假设用户开启保活功能，时间间隔，次数和空闲时间都使用缺省值，TCP 在 15 分钟内没有收到过对端发送的报文，开始发送保活报文，每隔 75 秒发送一次，连续发送 6 次，如果没有收到对方发送的任何 TCP 报文，就认为 TCP 连接无效，自动释放 TCP 连接。

配置举例

▾ 开启 TCP 保活功能。

【配置方法】 在设备上开启 TCP 保活功能，空闲时间是 3 分钟，发送保活报文的时间间隔是 60 秒，如果连续发送 4 次保活报文，没有收到对方发送的任何 TCP 报文，就认为 TCP 连接无效。

```
Ruijie# configure terminal
Ruijie(config)# ip tcp keepalive interval 60 times 4 idle-period 180
Ruijie(config)# end
```

【检验方法】 用户远程登录到设备，然后用户异常关机，在设备上执行 show tcp connect 观察用户的 IPv4 TCP 连接被删除的时间。

常见错误

8.5 监视与维护

清除各类信息

-

查看运行情况

作用	命令
显示 IPv4 TCP 连接的基本信息	show tcp connect [local-ip <i>a.b.c.d</i>] [local-port <i>num</i>] [peer-ip <i>a.b.c.d</i>] [peer-port <i>num</i>]
显示 IPv4 TCP 连接的统计信息	show tcp connect statistics
显示 IPv4 TCP 路径 MTU 的信息	show tcp pmtu [local-ip <i>a.b.c.d</i>] [local-port <i>num</i>] [peer-ip <i>a.b.c.d</i>] [peer-port <i>num</i>]
显示 IPv4 TCP 端口使用情况	show tcp port [<i>num</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
查看 IPv4 TCP 报文的调试信息	debug ip tcp packet [in out] [local-ip <i>a.b.c.d</i>] [peer-ip <i>a.b.c.d</i>] [global] [local-port <i>num</i>] [peer-port <i>num</i>] [deeply]
查看 IPv4 TCP 连接的调试信息	debug ip tcp transactions [local-ip <i>a.b.c.d</i>] [peer-ip <i>a.b.c.d</i>] [local-port <i>num</i>] [peer-port <i>num</i>]

9 软件 IPv4 快转

9.1 概述

在不支持硬件转发的产品上，由软件转发 IPv4 报文，为了使软件转发性能达到最优，我司实现了软件 IPv4 快转。

快转主要维护两张表：转发表和邻接表。转发表用来存放路由；邻接表用来存放下一跳的链路层信息，相当于 ARP 表和 IPv6 邻居表。

快转可以主动解析下一跳，还可以实现流量负载均衡。

i 下文仅介绍软件 IPv4 的相关内容。

协议规范

9.2 典型应用

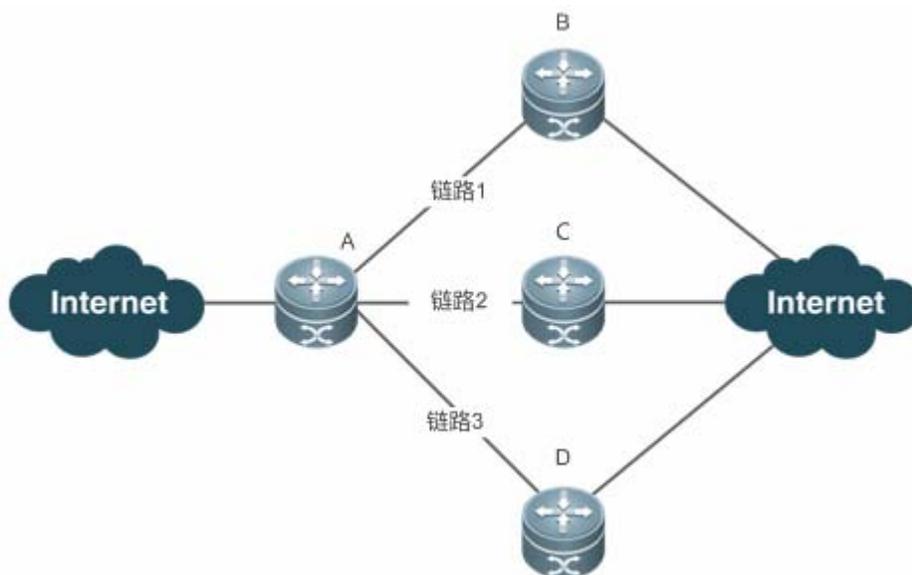
典型应用	场景描述
流量负载均衡	在网络路由中，当路由前缀关联到多个下一跳时，快转可以在多个下一跳中实现流量负载均衡。

9.2.1 流量负载均衡

应用场景

以下图为例，路由器 A 上，对于某条路由前缀关联 3 个下一跳，即链路 1、链路 2 和链路 3。缺省情况下，快转使用目的 IP 地址进行负载均衡，还可以根据源 IP 地址和目的 IP 地址进行负载均衡。

图 9-1



【注释】 A 为运行软件快转的路由器。
B、C、D 可以为其它转发设备。

功能部属

- 路由器 A 上运行软件快转。

9.3 功能详解

基本概念

IPv4 快转主要涉及以下基本概念：

📌 路由表

IPv4 路由表中存储着指向特定网络地址的路径，同时含有网络周边的拓扑信息。在报文转发的过程中 IPv4 快转根据路由表选择报文的传输路径。

📌 邻接节点

邻接节点包含了被路由报文的输出接口信息。例如下一跳列表、下一个处理部件、链路层输出封装等信息。当报文与该邻接节点匹配时，直接对报文进行封装，而后调用该节点的发送函数即可实现转发。为了便于检索和更新，邻接节点构成的表一般组织成哈希表的形式；为了支持路由负载均衡，邻接节点的下一条列表信息被组织为负载均衡表的形式；邻接节点中也可以不包含下一跳信息，也可以包含下一个处理部件的索引号（例如其它线卡，多业务卡）。

📌 主动解析

快转支持主动解析下一跳。对于以太网接口上的下一跳，如果不知道 MAC 地址，快转将主动解析下一跳。IPv4 快转请求 ARP 模块解析下一跳。

▾ 报文转发路径

报文的路由转发是根据报文的 IPv4 地址，所以如果指定了报文源 IPv4 地址和目的 IPv4 地址，则该报文的转发路径将是确定的。

9.3.1 快转负载均衡策略

快转负载均衡就是利用多个网络设备通道均衡分担流量。

工作原理

快转支持报文的负载均衡处理，目前交换机缺省支持基于报文目的 IP 地址的负载均衡策略。在快转模型中，当路由前缀关联到多个下一跳时，即多径路由，该路由将关联到一个负载均衡表，并依路由权重实现负载均衡。当 IPv4 报文依最长前缀匹配到该均衡表时，快转根据报文的 IPv4 地址进行散列，选中其中的一条路径转发报文。

9.4 配置详解

交换机缺省支持基于报文的目的 IP 地址的负载均衡策略，无需配置，可以通过以下命令进行监视与维护。

统计快转报文信息

快转报文统计信息即快转所处理的报文统计信息，包括了转发的报文数目，以及各种原因丢弃的报文数目等。快转提供配置信息查看和清除当前的统计信息，以供判断报文的转发行为是否和预期相同。

命令	作用
show ip ref packet statistics	显示 IPv4 快转当前的报文统计信息
clear ip ref packet statistics	清除 IPv4 快转当前的报文统计信息

查看邻接信息

用户可以通过以下命令来查看当前的邻接信息：

命令	作用
show ip ref adjacency [glean local ip-address {interface interface_type interface_number} discard statistics]	可以指定显示 IPv4 快转的集合邻接、本地邻接、指定 IP 对应邻接、指定接口关联邻接及所有邻接节点相关信息。

查看主动解析信息

用户可以通过以下命令来查看需要主动解析的下一跳：

命令	作用
show ip ref resolve-list	查看 IPv4 快转主动解析的下一跳。

查看报文转发路径信息

报文的路由转发是根据报文的 IPv4 地址，所以如果指定了报文源 IPv4 地址和目的 IPv4 地址，则该报文的转发路径将是确定的。执行下面的命令，并指定报文的源 IPv4 地址与目的 IPv4 地址，将会显示该报文的实际转发路径，比如报文丢弃、提交 CPU 或转发，进一步还可以知道从哪个接口转发等等。

命令	作用
show ip ref exact-route [oob] source-ipaddress dest_ipaddress	显示某特定报文的实际转发路径。oob 表示带外，即管理口所属的管理网络。

查看快转表路由信息

通过下面的命令可以查看快转表的路由信息：

命令	作用
show ip ref route [oob] [default {ip mask}] statistics]	显示当前 IPv4 快转表中的路由信息，参数 default 表示显示缺省路由。oob 表示带外，即管理口所属的管理网络。



配置指南-IP 路由

本分册介绍 IP 路由配置指南相关内容，包括以下章节：

1. 路由策略
2. 密钥
3. RIP
4. OSPFv2
5. 策略路由
6. 路由管理

1 路由策略

1.1 概述

路由策略是一种改变报文转发路径或路由信息的策略集，通常通过过滤列表和路由图进行实现。路由策略的应用灵活广泛，主要有以下几种方式：

- 在路由协议中引用过滤列表，可以过滤路由信息或修改路由信息。
- 在路由协议中引用路由图，可以过滤路由信息或修改路由信息，其中路由图可以进一步引用过滤列表。
- 在策略路由中引用路由图，可以控制报文转发或修改报文字段。

协议规范

1.2 典型应用

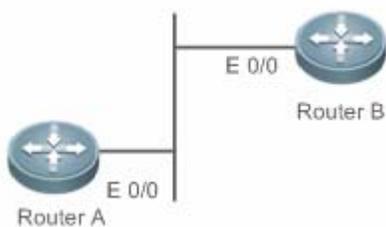
典型应用	场景描述
路由过滤	在路由协议中引用过滤列表，过滤其发送或接收的路由信息。
路由重分布	在路由协议中引用路由图，可以过滤路由信息或修改路由信息，将 RIP 路由重分布到 OSPF 中，要求只重分布跳数为 4 的 RIP 路由。
策略路由	在策略路由中引用路由图，可以控制报文转发或修改报文字段，为来自不同子网的报文，分别指定优选出接口。

1.2.1 路由过滤

缺省情况下，路由协议将发布和学习所有路由信息，使用过滤列表，路由协议只通告必需的路由，或只接收必需的路由信息。

应用场景

图 1-1



如上图所示，路由设备 A 有 3 个网络的路由：10.0.0.0、20.0.0.0、30.0.0.0

在设备上配置过滤列表，实现如下目的：

- 在路由设备 A 上过滤发送的路由信息，过滤其不需要发送的路由。
- 在路由设备 B 上过滤接收的路由信息，过滤其不需要学习的路由。

功能部属

- 在路由设备 A 上过滤发送的路由信息 30.0.0.0。
- 在路由设备 B 上过滤接收的路由信息 20.0.0.0，则设备 B 上仅学习到路由信息 10.0.0.0。

1.2.2 路由重分布

缺省情况下，路由重分布会将一个路由协议中所有路由信息，全部重分布到另一个路由协议中，路由的属性也将完全继承过来。使用路由图，可以对两个路由协议之间的重分布进行有条件的控制，包括如下。

- 允许指定路由重分发的范围，只重分发符合规则条件的路由信息。
- 允许设置重分发生成路由的路由属性。

应用场景

图 1-2



如上图所示，在设备上配置路由重分布，实现如下目的：

- 将 RIP 路由重分布到 OSPF 中，要求只重分布跳数为 4 的 RIP 路由。
- 在 OSPF 路由域中，该路由的初始量度值为 40，路由类型为外部路由 type-1，路由标记值设为 40。

功能部属

- 配置路由图 rip_to_ospf：match 跳数为 4 的路由；set 路由的初始量度值为 40，路由类型为外部路由 type-1，路由标记值设为 40。
- 配置路由重分布，将 RIP 路由重分发到 OSPF 中，引用路由图 rip_to_ospf。

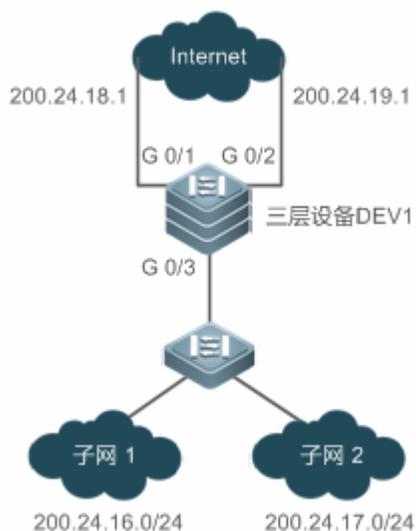
1.2.3 策略路由

策略路由（Policy-Based Routing）就是基于策略的路由。将包含策略的路由图应用在接口和设备上，即可实现策略路由。

与静态路由相似，策略路由也是手工配置的路由，不能随网络变化自动更新，仅在本地有效。与静态路由、动态路由相比，策略路由更加灵活。静态路由、动态路由只能依据目的地址转发报文。策略路由可以依据源地址、目的地址、报文长度、入端口等转发报文。

应用场景

图 1-3



在三层设备 DEV1 上配置策略路由，实现如下目的：

- 来自子网 1 (200.24.16.0/24) 的报文，优先从 GE0/1 接口发出。
- 来自子网 2 (200.24.17.0/24) 的报文，优先从 GE0/2 接口发出。

功能部属

- 配置两个不同的 ACL，分别用来匹配子网 1 和子网 2 的报文。
- 配置路由图 RM_FOR_PBR：策略 10 实现“来自子网 1 的报文，优先从 GE0/1 接口发出”；策略 20 实现“来自子网 2 的报文，优先从 GE0/2 接口发出”。
- 对 GE0/3 接口收到的报文执行策略路由，引用路由图 RM_FOR_PBR。

1.3 功能详解

功能特性

功能特性	作用
过滤列表	根据某种路由属性定义一组列表，供路由协议进行路由过滤
路由图	策略定义了“如果匹配某些条件，则执行某些处理动作”。

1.3.1 过滤列表

过滤列表是根据某种路由属性定义的一组列表，是路由策略进行过滤的工具，单独配置过滤列表没有任何作用，只有在路由协议中应用这些过滤列表，才能达到路由过滤的效果。

工作原理

根据不同的路由属性，过滤列表主要包括以下几种：

访问控制列表 (ACL)

访问控制列表 (ACL) 包括 IPv4 的 ACL，用户定义访问控制列表时可以指定 IPv4 地址和掩码，用于匹配路由信息的目的网段或下一跳地址。

访问控制列表的相关内容请参见“ACL 配置指导”。

地址前缀列表 (prefix-list)

地址前缀列表的作用类似于访问控制列表，包括 IPv4 地址前缀列表，过滤路由时用于匹配路由信息的目的网段。

相关配置

创建访问控制列表

缺省情况下，未配置访问控制列表，未设置策略。

在全局模式下，使用 **ip access-list { extended | standard } { id | name }** 创建一个 IPv4 的 ACL。

可以在 ACL 中设置多条策略，策略可按序列号排列，策略有 permit 和 deny 两种工作模式。

创建地址前缀列表

缺省情况下，未配置地址前缀列表，未设置表项。

在全局模式下，使用 **ip prefix-list prefix-list-name [seq seq-number] { deny | permit } ip-prefix [ge minimum-prefix-length] [le maximum-prefix-length]** 创建一个 IPv4 前缀列表并增加一条前缀列表表项。

可以在前缀列表中设置多条表项，可按序号排列，表项分为 permit 和 deny 两种工作模式。

使用 **ip prefix-list prefix-list-name description description-text** 为前缀列表添加文字描述。

使用 **ip prefix-list sequence-number** 为前缀列表启用排序功能。

创建自治域路径过滤列表

缺省情况下，未配置自治域路径过滤列表，未设置表项。

在全局模式下，使用 **ip as-path access-list path-list-num { permit | deny } regular-expression** 创建一个自治域路径过滤列表并增加一条表项。

可以在自治域路径过滤列表中设置多条表项，表项分为 permit 和 deny 两种工作模式。

创建团体属性过滤列表

缺省情况下，未配置团体属性过滤列表，为配置表项。

在全局模式下，使用 **ip community-list { { standard | expanded } community-list-name | community-list-number } { permit | deny } [community-number..]** 创建一个自治域路径过滤列表并增加一条表项。

可以在团体属性列表中设置多条表项，表项分为 permit 和 deny 两种工作模式。

创建扩展团体属性过滤列表

缺省情况下，未配置扩展团体属性列表，未设置表项。

在全局模式下，使用 **ip extcommunity-list { standard-list | standard list-name } { permit | deny } [rt value] [soo value]** 创建一个标准类型的扩展团体属性列表并增加一条表项。

或者使用 **ip extcommunity-list { expanded-list | expanded list-name } { permit | deny } [regular-expression]** 创建一个扩展类型的扩展团体属性列表并增加一条表项。

也可以使用 **ip extcommunity-list { expanded-list | expanded list-name | standard-list | standard list-name }** 创建扩展团体属性列表，并进入 **ip extcommunity-list** 的配置模式添加表项。

可以在扩展团体属性列表中设置多条表项，表项分为 permit 和 deny 两种工作模式。

1.3.2 路由图

一条策略就是一个 match ..., set... 语句，表示“如果匹配某些条件，则执行某些处理动作”。

工作原理

策略的执行

一个路由图中可以包含多条策略。每条策略都有对应的序号（Sequence），序号越小优先级越高。策略按序号顺序依次执行，一旦满足某条策略的匹配条件，就执行该策略对应的处理动作，并退出路由图。如果不能满足任何一条策略的匹配条件，则不做任何处理。

策略的工作模式

策略有以下两种工作模式：

- permit（允许）：当满足策略的匹配条件时，执行该策略对应的处理动作，并退出路由图。
- deny（拒绝）：当满足策略的匹配条件时，不执行该策略对应的处理动作，并退出路由图。

策略的匹配条件

策略的匹配条件包含 0 条、1 条或多条 match 规则。

- 如果包含 0 条 match 规则，则无条件匹配。
- 如果包含 1 条或多条 match 规则，则必须同时匹配所有 match 规则，才能满足策略的匹配条件。

策略的处理动作

策略的处理动作包含 0 条、1 条或多条 set 规则。

- 如果包含 0 条 set 规则，则不做任何处理，直接退出路由图。
- 如果包含 1 条或多条 set 规则，则执行所有处理动作后，退出路由图。

 如果 set 规则有优先级关系，则仅优先级最高者生效。

相关配置

创建路由图（策略）

缺省情况下，未配置路由图，未设置策略。

在全局模式下，使用 **route-map route-map-name [permit | deny] [sequence-number]** 命令创建路由图，并添加一条策略。可以在一个路由图中设置多条策略。每条策略使用不同的序列号。

设置策略的匹配条件

缺省情况下，未设置 match 规则（即：策略的匹配条件包括 0 条 match 规则）。

在路由图模式下，使用 **match** 命令设置 match 规则。一条 **match** 命令对应一条 match 规则。

RGOS 提供了丰富的 **match** 命令，可以设置灵活的匹配条件。

命令	说明
match interface	以路由的出接口为匹配条件。
match ip address	以路由目的 IPv4 地址为匹配条件。
match ip next-hop	以下一跳 IPv4 地址为匹配条件。
match ip route-source	以路由源 IPv4 地址为匹配条件。
match metric	以路由的度量值为匹配条件。
match mpls-label	以路由是否带标签信息为匹配条件。
match origin	以路由的来源为匹配条件。
match route-type	以路由的类型为匹配条件。
match tag	路由的标记值为匹配条件。

设置策略的处理动作

缺省情况下，未设置 set 规则（即：策略的处理动作包括 0 条 set 规则）。

在路由图模式下，使用 **set** 命令设置 set 规则。一条 **set** 命令对应一条 set 规则。

RGOS 提供了丰富的 **set** 命令，可以设置灵活的处理动作。

命令	说明
set aggregator as	修改路由的聚合者的 AS 属性值。
set as-path prepend	增加指定的 as-path 属性值。
set atomic-aggregate	设置路由 atomic-aggregate 属性。
set comm-list delete	对匹配 match 规则的路由删除团体属性列表中的所有的团体属性值。
set community	设置路由的团体属性值。
set dampening	设置路由的路由振荡参数。
set extcomm-list delete	对匹配 match 规则的路由删除扩展团体属性列表中的所有的扩展团体属性值。
set extcommunity	设置路由的扩展团体属性值。
set fast-reroute	设置快速重路由的备份出口和备份下一跳。
set ip default nexthop	指定路由的缺省下一跳。优先级比普通路由低。
set ip dscp	修改 IP 报文的 dscp 字段。
set ip global next-hop	指定路由的下一跳。
set ip global default next-hop	指定路由的缺省下一跳。
set ip nexthop	指定路由的下一跳。
set ip next-hop verify-availability	指定路由的下一跳 IP 地址，并且通过 BFD 检测该下一跳的可达性。
set ip precedence	修改 IP 报文的 precedence 字段。
set ip tos	修改 IP 报文的 tos 字段。
set level	设置路由要导入的目标区域类型。
set local-preference	设置路由的 local-preference 属性值。
set metric	修改路由的度量值。
set metric-type	设置路由的度量值类型。
set mpls-label	设置 MPLS 标签。
set next-hop	设置路由的下一跳 IP 地址。
set origin	设置路由的来源属性。
set originator-id	设置路由的源发者 IP 地址。
set tag	设置路由的标记值。
set weight	设置路由的权重值。

1.4 配置详解

配置项	配置建议 & 相关命令	
配置路由图	 可选配置。用于定义策略。	
	route-map	创建策略（路由图）。
	match	设置策略的匹配条件。
	set	设置策略的处理动作。
配置过滤列表	 可选配置。用于定义过滤列表。	
	ip as-path	定义 AS 路径过滤规则
	ip community-list	定义团体属性列表
	ip extcommunity-list	定义扩展团体属性列表
	ip prefix-list	创建一个前缀列表

	ip prefix-list description	为一个前缀列表添加文字描述
	ip prefix-list sequence-number	为前缀列表启用排序功能

1.4.1 配置路由图

配置效果

- 定义一组路由策略，供路由协议或策略路由使用。

注意事项

- 如果 **match** 命令中引用 ACL 来定义报文匹配条件，则必须配置此 ACL。
- 不能同时配置的 **match** 命令有：

如下 match 命令	不能与如下 match 命令同时配置
match ip address	match ip prefix-list
match ip next-hop	match ip next-hop prefix-list
match ip route-source	match ip route-source prefix-list

- 不能同时配置的 **set** 命令有：

如下 set 命令	不能与如下 set 命令同时配置
set ip next-hop	set ip next-hop verify-availability
set ip dscp	set ip tos
set ip dscp	set ip precedence

配置方法

📌 创建策略（路由图）

- 必须配置。
- 在需要应用策略的设备上配置。

📌 设置策略的匹配条件

- 可选配置。
- 如果未配置 **match** 规则，则无条件匹配。
- 如果配置了多条 **match** 规则，则必须同时匹配全部 **match** 规则。
- 在需要应用策略的设备上配置。

📌 设置策略的处理动作

- 可选配置。
- 如果未配置 **set** 规则，则不做任何处理。
- 如果配置了多条 **set** 规则，则需执行全部 **set** 规则（如果 **set** 规则有优先级关系，则仅优先级最高者生效）。
- 在需要应用策略的设备上配置。

检验方法

- 检查路由图配置信息。

相关命令

创建策略（路由图）

- 【命令格式】 **route-map** *route-map-name* [{ **permit** | **deny** } *sequence*]
- 【参数说明】 *route-map-name* : 路由图名称，不允许超过 32 个字符。
permit : 指定该策略的工作模式为允许模式，缺省为允许模式。
deny : 指定该策略的工作模式为拒绝模式，缺省为允许模式。
sequence : 指定该策略的序列号，值越小优先级越高，缺省为 10。
- 【命令模式】 全局模式
- 【使用指导】 如果此路由图尚不存在，则此命令创建路由图，并添加一条策略。
如果此路由图已经存在，则此命令只添加一条策略。

设置策略的匹配条件

- 【命令格式】 **match interface** *interface-type interface-number* [...*interface-type interface-number*]
- 【参数说明】 *interface-type interface-number* : 接口类型和接口编号。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 **match** 规则用来匹配路由或报文的下一跳出口
- 【命令格式】 **match ip address** { *access-list-number* [*access-list-number...* | *access-list-name...*] | *access-list-name* [*access-list-number...* | *access-list-name*] | **prefix-list** *prefix-list-name* [*prefix-list-name...*] }
- 【参数说明】 *access-list-number* : 访问列表号。标准访问列表号范围：1-99，1300-1999；扩展访问列表号范围：100-199，2000-2699。
access-list-name : 访问列表名字。
prefix-list *prefix-list-name* : 要匹配的前缀列表名字。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 **match** 规则用访问控制列表或前缀列表规则来匹配报文或路由的目的 IPv4 地址
匹配 ACL 与匹配前缀列表不能同时配置。
- 【命令格式】 **match ip next-hop** { *access-list-number* [*access-list-number...* | *access-list-name...*] | *access-list-name* [*access-list-number...* | *access-list-name*] | **prefix-list** *prefix-list-name* [*prefix-list-name...*] }
- 【参数说明】 *access-list-number* : 访问列表号。标准访问列表号范围：1-99，1300-1999；扩展访问列表号范围：100-199，2000-2699。
access-list-name : 访问列表名字。
prefix-list *prefix-list-name* : 要匹配的前缀列表名字。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 **match** 规则用访问控制列表或前缀列表规则来匹配路由下一跳 IPv4 地址匹配 ACL 与匹配前缀列表不能同时配置。
- 【命令格式】 **match ip route-source** { *access-list-number* [*access-list-number...* | *access-list-name...*] | *access-list-name* [*access-list-number...* | *access-list-name*] | **prefix-list** *prefix-list-name* [*prefix-list-name...*] }
- 【参数说明】 *access-list-number* : 访问列表号。标准访问列表号范围：1-99，1300-1999；扩展访问列表号范围：100-199，2000-2699。

access-list-name : 访问列表名字。

prefix-list *prefix-list-name* : 要匹配的前缀列表名字。

【命令模式】 路由图配置模式

【使用指导】 此 match 规则用控制列表或前缀列表规则来匹配路由源 IPv4 地址匹配 ACL 与匹配前缀列表不能同时配置。

【命令格式】 **match metric** *metric*

【参数说明】 *metric* : 路由量度值, 取值范围: 0-4294967295。

【命令模式】 路由图配置模式

【使用指导】 此 match 规则用来匹配路由的度量值

【命令格式】 **match mpls-label**

【参数说明】 -

【命令模式】 路由图配置模式

【使用指导】 此 match 规则用来匹配带标签的路由信息

【命令格式】 **match origin { egp | igp | incomplete }**

【参数说明】 **egp** : 来源是远端的 EGP。

igp : 来源是本地的 IGP。

incomplete : 来源是不完整类型。

【命令模式】 路由图配置模式

【使用指导】 此 match 规则用来匹配路由来源

【命令格式】 **match route-type { local | internal | external [type-1 | type-2] }**

【参数说明】 **local** : 本地生成的路由。

Internal : OSPF 的内部路由。

external : 外部路由(BGP 或 OSPF 的外部路由)。

type-1 | type-2 : OSPF 的类型 1 或类型 2 外部路由。

【命令模式】 路由图配置模式

【使用指导】 此 match 规则用来匹配路由的路由类型

【命令格式】 **match tag** *tag* [...*tag*]

【参数说明】 *tag* : 路由的标记值。

【命令模式】 路由图配置模式

【使用指导】 此 match 规则用来匹配路由的标记值

▾ 设置策略的处理动作

【命令格式】 **set aggregator as** *as-number ip-address*

【参数说明】 *as-number* : 聚合者的 AS 域号, AS 号范围为 1..4294967295, 用圆点模式表示为 1..65535.65535。

ip-address : 聚合者的地址。

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来修改路由的聚合者的 AS 属性值

- 【命令格式】 **set as-path prepend** *as-number*
- 【参数说明】 *as-number* : 准备加入 AS_PATH 属性的 AS 号, AS 号范围为 1..4294967295, 用圆点模式表示为 1..65535.65535。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来添加指定的 as-path 属性值
- 【命令格式】 **set atomic-aggregate**
- 【参数说明】 -
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来设置路由 atomic-aggregate 属性
- 【命令格式】 **set comm-list** { *community-list-number* | *community-list-name* } **delete**
- 【参数说明】 *community-list-number* : 团体列表号。标准团体列表 : 编号范围为 1-99 ; 扩展团体列表 : 标号范围为 100-199。
community-list-name : 团体列表名字, 不超过 80 字符。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来对匹配 match 规则的路由删除团体属性列表中的所有的团体属性值
- 【命令格式】 **set community** { *community-number* [*community-number* ...] **additive** | **none** }
- 【参数说明】 *community-number* : 团体属性值。
additive : 在原有 community 属性的基础上增加。
none : 设置团体属性为空。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来设置路由的团体属性值
- 【命令格式】 **set dampening** *half-life reuse suppress max-suppress-time*
- 【参数说明】 *half-life* : 路由可达或不可达时的半衰期。范围 1..45(分钟), 缺省值为 15 分钟。
reuse : 当路由的惩罚值低于该值时, 路由抑制解除。范围 1..20000, 缺省值为 750。
suppress : 当路由的惩罚值超过该值时, 路由被抑制。范围 1..20000, 缺省 2000。
max-suppress-time : 路由能被抑制的最长时间。范围 1..255(分钟), 缺省值为 4* half-life。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来设置路由的路由振荡参数
- 【命令格式】 **set extcomm-list** { *extcommunity-list-number* | *extcommunity-list-name* } **delete**
- 【参数说明】 *extcommunity-list-number* : 扩展团体列表号, 扩展标准团体列表 : 编号范围为 1-99 ; 扩展团体列表 : 标号范围为 100-199。
extcommunity-list-name : 扩展团体列表名字, 不超过 80 字符。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来对匹配 match 规则的路由删除扩展团体属性列表中的所有的扩展团体属性值
- 【命令格式】 **set extcommunity** { *rt extend-community-value* | **soo** *extend-community-value* }

- 【参数说明】 **rt** : 设置路由的 RT 属性值。
soo : 设置路由的 SOO 属性值。
extend-community-value : 扩展团体的值。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来设置路由的扩展团体属性值
- 【命令格式】 **set fast-reroute backup-interface** *interface-type interface-number* [**backup-nexthop** *ip-address*]
- 【参数说明】 *interface-type interface-number* : 指定备份出口。
backup-nexthop *ip-address* : 指定备份下一跳, 非点到点接口必须指定备份下一跳。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来设置快速重路由的备份出口和备份下一跳
- 【命令格式】 **set ip default next-hop** *ip-address* [*weight*] [...*ip-address* [*weight*]]
- 【参数说明】 *ip-address* : 下一跳 IP 地址。
weight : 本下一跳的权重。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来指定路由的缺省下一跳
- 【命令格式】 **set ip dscp** *dscp_value*
- 【参数说明】 *dscp_value* : 设置 IP 报文中 IP 头 DSCP 值。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来修改 IP 报文的 dscp 字段
- 【命令格式】 **set ip next-hop** *ip-address* [*weight*] [...*ip-address* [*weight*]]
- 【参数说明】 *ip-address* : 下一跳 IP 地址。
weight : 本下一跳的权重。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来指定路由的下一跳
- 【命令格式】 **set ip next-hop verify-availability** *ip-address bfd interface-type interface-number gateway*
- 【参数说明】 *ip-address* : 下一跳 IP 地址。
bfd : 通过 BFD 方式进行邻居检测
interface-type : 配置接口类型
interface-number : 接口编号
gateway : 配置网关 IP, 即为 BFD 的邻居 IP。配置下一跳为该邻居的将通过 BFD 进行检测该转发路径的可达性。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来指定路由的下一跳, 并通过 BFD 快速检测下一跳有效性
- 【命令格式】 **set ip precedence** { *number* | **critical** | **flash** | **flash-override** | **immediate** | **internet** | **network** | **priority** | **routine** }

- 【参数说明】 *number* : 用数字表示 IP 头的优先级, 取值范围 0-7。
- 7 : critical
6 : flash
5 : flash-override
4 : immediate
3 : internet
2 : network
1 : priority
0 : routine
- critical | flash | flash-override | immediate | internet | network | priority | routine** : IP 头的优先级。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来修改 IP 报文头的优先级 **precedence** 字段
- 【命令格式】 **set ip tos { *number* | max-reliability | max-throughput | min-delay | min-monetary-cost | normal }**
- 【参数说明】 *number* : 用数字表示 IP 头的 TOS, 取值范围 0-15。
- 2 : **max-reliability**
4 : **max-throughput**
8 : **min-delay**
1 : **min-monetary-cost**
0 : **normal**
- max-reliability | max-throughput | min-delay | min-monetary-cost | normal** : IP 头的优先级。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来修改 IP 报文的 tos 字段
- 【命令格式】 **set level { stub-area | backbone }**
- 【参数说明】 **stub-area** : 指定重分发路由通告到 OSPF Stub Area。
backbone : 指定重分发路由通告到 OSPF 骨干区。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来设置路由要导入的目标区域类型
- 【命令格式】 **set local-preference *number***
- 【参数说明】 *number* : 本地优先度的度量值。取值范围 0-4294967295, 值越大优先级越高。
- 【命令模式】 路由图配置模式
- 【使用指导】 此 set 规则用来设置路由的 local-preference 属性值
- 【命令格式】 **set metric [+ *metric-value* | - *metric-value* | *metric-value*]**
- 【参数说明】 + : 增加 (在原路由的 metric 上增加)。
- : 减少 (在原路由的 metric 上减少)。
metric-value : 设置重分布路由的度量, 值越大优先级越低。
- 【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来修改路由的量度值

【命令格式】 **set metric-type type**

【参数说明】 *type* : 设置重分布路由的类型。OSPF 重分布路由的类型缺省为 type-2。

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来设置度量值类型

【命令格式】 **set mpls-label**

【参数说明】 -

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来设置 MPLS 标签

【命令格式】 **set next-hop ip-address**

【参数说明】 *ip-address* : 下一跳 IP 地址。

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来设置下一跳 IP 地址

【命令格式】 **set origin { egp | igp | incomplete }**

【参数说明】 **egp** : 来源是远端的 EGP。

igp : 来源是本地的 IGP。

incomplete : 来源是不完整类型, 一般指由重分发产生的路由。

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来设置路由来源属性

【命令格式】 **set originator-id ip-address**

【参数说明】 *ip-address* : 源发者的地址。

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来设置路由源发者 IP 地址

【命令格式】 **set tag tag**

【参数说明】 *tag* : 设置重分布路由的标记。

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来设置路由标记值

【命令格式】 **set weight number**

【参数说明】 *number* : 设置路由的权重值。取值范围 : 0-65535, 值越大优先级越高。

【命令模式】 路由图配置模式

【使用指导】 此 set 规则用来设置路由权重值

查看路由图配置信息

【命令格式】 **show route-map [name]**

【参数说明】 *name* : 指定路由图。

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 使用 **show route-map** 命令查看路由图的配置信息。

如果配置路由图时引用了 ACL，则可以使用 **show access-list** 命令查看 ACL 配置信息。

配置举例

i 仅介绍与路由图相关的配置。

路由重分布中使用路由图，过滤并修改路由信息。

【网络环境】 图示：设备上同时连接着 OSPF 路由域和 RIP 路由域。

图 1-4



【配置要求】

- 将 RIP 路由重分布到 OSPF 中，要求只重分布跳数为 4 的 RIP 路由。在 OSPF 路由域中，路由类型为外部路由 type-1，路由标记值设为 40。

- 将 OSPF 路由重分布到 RIP 中，要求只重分布标记为 10 的 OSPF 路由。在 RIP 路由域中，该路由的初始量度值设为 10。

【配置方法】

- 配置路由图 redrip：**match** 跳数为 4 的路由；**set** 路由的初始量度值为 40，路由类型为外部路由 type-1，路由标记值设为 40。

- 配置路由图 redospf：**match** 标记为 10 的路由；**set** 路由的初始量度值为 10。

- 配置将 RIP 路由重分布到 OSPF 中，应用路由图 redrip。

- 配置将 OSPF 路由重分布到 RIP 中，应用路由图 redospf。

```
Ruijie(config)# route-map redrip permit 10
Ruijie(config-route-map)# match metric 4
Ruijie(config-route-map)# set metric-type type-1
Ruijie(config-route-map)# set tag 40
Ruijie(config-route-map)# exit
Ruijie(config)# route-map redospf permit 10
Ruijie(config-route-map)# match tag 10
Ruijie(config-route-map)# set metric 10
Ruijie(config-route-map)# exit
Ruijie(config)# router ospf 1
Ruijie(config-router)# redistribute rip subnets route-map redrip
Ruijie(config-router)# exit
Ruijie(config)# router rip
Ruijie(config-router)# redistribute ospf 1 route-map redospf
Ruijie(config-router)# exit
```

【检验方法】

- 查看路由图配置信息，确认策略规则。

- 查看 OSPF 路由信息库，确认符合策略规则的路由被重分布。

```
Ruijie# show route-map
route-map redrip, permit, sequence 10
```

```
Match clauses:
  metric 4
Set clauses:
  metric 40
  metric-type type-1
  tag 40
route-map redospf, permit, sequence 10
Match clauses:
  tag 10
Set clauses:
  metric 10

Ruijie# show ip ospf database external

          OSPF Router with ID (192.100.1.9) (Process ID 1)

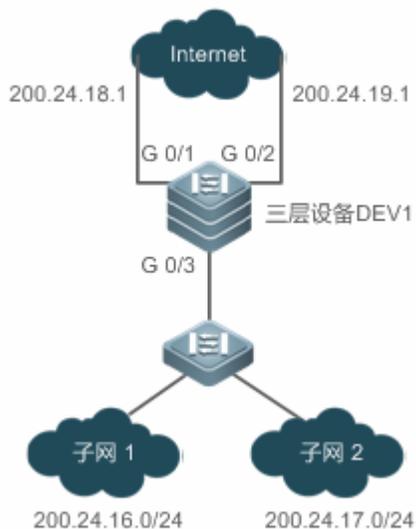
          AS External Link States

LS age: 5
Options: 0x2 (-|-|-|-|-|E|-)
LS Type: AS-external-LSA
Link State ID: 192.168.199.0 (External Network Number)
Advertising Router: 192.100.1.9
LS Seq Number: 80000001
Checksum: 0x554d
Length: 36
Network Mask: /24
  Metric Type: 1
  TOS: 0
  Metric: 4
  Forward Address: 0.0.0.0
  External Route Tag: 40
```

↘ 在策略路由中应用路由图

【组网环境】

图 1-5



【配置要求】 在设备 DEV1 上配置策略路由，实现如下目的：

- 来自子网 1 (200.24.16.0/24) 的报文，优先从 GE0/1 接口发出。
- 来自子网 2 (200.24.17.0/24) 的报文，优先从 GE0/2 接口发出。
- 两条出口链路互为备份。

【配置方法】

- 配置两个不同的 ACL，分别用来匹配子网 1 和子网 2
- 配置路由图 RM_FOR_PBR：策略 10 实现“来自子网 1 的报文，优先从 GE0/1 接口发出”；策略 20 实现“来自子网 2 的报文，优先从 GE0/2 接口发出”。
- 配置策略路由，对 GE0/3 接口收到的报文应用路由图 RM_FOR_PBR。
- 设置策略路由在多个下一跳之间实现冗余备份。

! 在冗余备份模式下，多个 set 下一跳的配置顺序即为优先生效的顺序。

```
Ruijie(config)# access-list 1 permit 200.24.16.0 0.0.0.255
Ruijie(config)# access-list 2 permit 200.24.17.0 0.0.0.255
Ruijie(config)# route-map RM_FOR_PBR 10
Ruijie(config-route-map)# match ip address 1
Ruijie(config-route-map)# set ip next-hop 200.24.18.1
Ruijie(config-route-map)# set ip next-hop 200.24.19.1
Ruijie(config-route-map)# exit
Ruijie(config)# route-map RM_FOR_PBR 20
Ruijie(config-route-map)# match ip address 2
Ruijie(config-route-map)# set ip next-hop 200.24.19.1
Ruijie(config-route-map)# set ip next-hop 200.24.18.1
Ruijie(config-route-map)# exit
Ruijie(config)# interface GigabitEthernet 0/3
Ruijie(config-if)# ip policy route-map RM_FOR_PBR
Ruijie(config)# ip policy redundance
```

- 【检验方法】
- 查看策略路由配置信息，确认接口上应用了路由图。
 - 查看路由图配置信息，确认策略规则。
 - 查看 ACL 配置信息，确认报文过滤规则。

```
Ruijie# show ip policy
Balance mode: redundance

Interface                               Route map
GigabitEthernet 0/3                     RM_FOR_PBR
!
Ruijie# show route-map
route-map RM_FOR_PBR, permit, sequence 10
  Match clauses:
    ip address 1
  Set clauses:
    ip next-hop 200.24.18.1
    ip next-hop 200.24.19.1
route-map RM_FOR_PBR, permit, sequence 20
  Match clauses:
    ip address 2
  Set clauses:
    ip next-hop 200.24.19.1
    ip next-hop 200.24.18.1
Ruijie# show access-lists
ip access-list standard 1
 10 permit 200.24.16.0 0.0.0.255
 10 permit 200.24.16.0 0.0.0.255
ip access-list standard 2
 10 permit 200.24.17.0 0.0.0.255
```

常见错误

- 配置 match 匹配 ACL、前缀列表等策略后，没有定义相应的 ACL、前缀列表。

1.4.2 配置过滤列表

配置效果

- 定义一组路由过滤规则，供路由协议使用。

注意事项

- 过滤列表的配置必须关联路由协议应用才能够实际生效。

配置方法

▾ 配置地址前缀列表

- 如果需要过滤地址前缀，则应该执行此配置项。
- 若无特殊要求，应在需要进行地址前缀列表过滤的路由器上配置。

▾ 配置自治域路径过滤列表

- 如果需要过滤地址前缀，则应该执行此配置项。
- 若无特殊要求，应在需要进行自治域路径过滤的路由器上配置。

▾ 配置团体属性列表

- 如果需要过滤团体属性，则应该执行此配置项。
- 若无特殊要求，应在需要过滤团体属性的路由器上配置。

▾ 配置扩展团体属性列表

- 如果需要过滤扩展团体属性，则应该执行此配置项。
- 若无特殊要求，应在需要过滤扩展团体属性的路由器上配置。

检验方法

- 过滤列表配置正确。
- 查看路由表，路由能够被正确过滤。

相关命令

▾ 定义 AS 路径过滤规则

【命令格式】 **ip as-path access-list** *path-list-num* { **permit** | **deny** } *regular-expression*

【参数说明】 *path-list-num* 基于正则表达式的 AS 路径控制列表名，AS 路径列表标识符，范围：1---500。

permit：允许访问

deny：拒绝访问

regular-expression 正则表达式。范围：1..255 字符。

【命令模式】 全局模式

【使用指导】 -

▾ 定义团体属性列表

【命令格式】 **ip community-list** { { **standard** | **expanded** } *community-list-name* | *community-list-number* } { **permit** | **deny** } [*community-number..*]

【参数说明】 **standard**：标准团体列表

expanded：扩展团体列表

community-list-name：团体列表名字，团体列表名不超过 80 字符

community-list-number：团体列表号，标准团体列表：编号范围为 1-99，扩展团体列表：编号范围为 100-199

permit：允许访问

deny：拒绝访问

community-number：团体属性值。

【配置模式】 全局模式

【使用指导】 使用该命令来定义用于 BGP 的团体列表。

▾ 定义扩展团体属性列表

【命令格式】 **ip extcommunity-list** { *expanded-list* | **expanded** *list-name* } { **permit** | **deny** } [*regular-expression*]

【参数说明】 *expand-list*：取值范围[100~199]，用来标识扩展的 extcommunity 列表，一个 extcommunity 列表可包含多个规则。

standard-list : 取值范围[1~99]，用来标识标准的 extcommunity 列表，一个 extcommunity 表可以包含多个规则。

expanded list-name : 扩展 extcommunity 的名称，使用该参数将进入扩展团体属性列表的配置模式。名字长度不超过 32 个字符。

standard list-name : 标准 extcommunity 的名称，用来标识标准的 extcommunity 表的名称，使用该参数将进入扩展团体属性列表的配置模式。名字长度不超过 32 个字符。

permit : 定义一个允许的 extcommunity 规则。

deny : 定义一个拒绝的 extcommunity 规则。

regular-expression : (可选) 定义匹配的模板，用来匹配 extcommunity。

sequence-number : (可选) 定义该规则的顺序序列号，范围从 1 到 2147483647。如果不指定，缺省情况下每增加一个规则其序列号自动以 10 递增，起始从 10 开始编号。

rt : (可选) 设置 RT 属性值，该命令只能用在标准 extcommunity 配置，不能用于扩展 extcommunity 配置

soo : (可选) 设置 SOO 属性值，该命令只能用在标准 extcommunity 配置，不能用于扩展 extcommunity 配置

value : 扩展团体的值 (extend_community_value)

【配置模式】 全局配置模式、ip extcommunity-list 配置模式

【使用指导】 -

📌 创建一个前缀列表

【命令格式】 **ip prefix-list** *prefix-list-name* [**seq** *seq-number*] { **deny** | **permit** } *ip-prefix* [**ge** *minimum-prefix-length*] [**le** *maximum-prefix-length*]

【参数说明】 *prefix-list-name* : 前缀列表名字

seq-number : 给一条前缀列表表项赋以序列号。其范围是 1 到 2147483647。如果命令中没有包含这个序列号，则系统会为这条前缀列表表项分配一个默认的序列号。第一个表项的默认序列号是 5，后续每一个未赋值的表项的默认序列号数值是比前一个序列号大的第一个 5 的倍数的数值。

deny : 对于匹配情况拒绝访问

permit : 对于匹配情况允许访问

ip-prefix : 配置网络地址以及掩码。掩码长度从 0 到 32。

minimum-prefix-length : 指定的最小范围值(即范围的起始长度)

maximum-prefix-length : 指定的最大范围值(即范围的终止长度)

【配置模式】 全局模式

【使用指导】 -

📌 为一个前缀列表添加文字描述

【命令格式】 **ip prefix-list** *prefix-list-name* **description** *descripton-text*

【参数说明】 *prefix-list-name* : 前缀列表名字

descripton-text : 前缀列表文字描述

【配置模式】 全局模式

【使用指导】 -

📌 为前缀列表启用排序功能

【命令格式】 **ip prefix-list** **sequence-number**

【参数说明】 -

【配置模式】 全局模式

【使用指导】 -

常见配置错误

- 配置了过滤列表，但未正确在路由协议中引用，导致过滤列表无法生效

1.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
-	-

查看运行情况

作用	命令
查看路由图的配置信息	show route-map [<i>route-map-name</i>]
查看访问控制列表的配置信息	show access-lists [<i>id</i> <i>name</i>]
查看 IPv4 地址前缀列表的配置信息	show ip prefix-list [<i>prefix-name</i>]
查看自治域路径过滤列表列表的配置信息	show ip as-path-access-list [<i>num</i>]
查看团体属性过滤列表的配置信息	show ip community-list [<i>community-list-number</i> <i>community-list-name</i>]
查看扩展团体属性过滤列表的配置信息	show ip extcommunity-list [<i>extcommunity-list-num</i> <i>extcommunity-list-name</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
-	-

2 密钥

2.1 概述

密钥是一种参数，在明文转换为密文或将密文转换为明文的算法中使用。路由协议的报文认证功能中，支持明文认证和密文认证，需要引用密钥。

i 目前仅 RIP 和 ISIS 的报文认证功能会引用密钥。

协议规范

-

2.2 典型应用

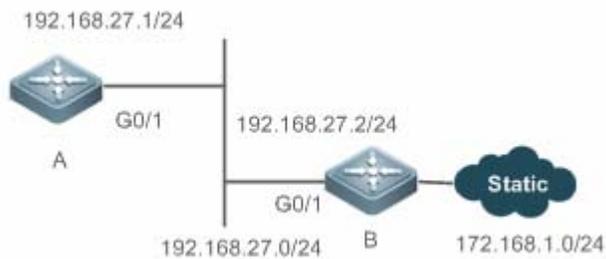
典型应用	场景描述
RIP 认证	RIP 协议使用密钥进行报文认证

2.2.1 RIP认证

应用场景

网络中的设备运行 RIP 协议，并采用 MD5 认证方式提升协议安全。

图 2-1



功能部属

- 在 A 上配置密钥链。配置 RIP 协议，启用报文认证功能，引用密钥链。
- 在 B 上配置密钥链。配置 RIP 协议，启用报文认证功能，引用密钥链。

2.3 功能详解

功能特性

功能特性	作用
密钥链	为路由协议中的认证机制提供工具。

2.3.1 密钥链

工作原理

一个密钥链中可以包含多个不同的密钥，每个密钥包含以下属性：

- 密钥 ID：用于标识每个密钥，在当前密钥链中，每个密钥和 ID 一一对应。
- 认证串：密钥用以认证的一组字符，用于路由协议校验认证串配置的一致性。
- 生存期：指定当前密钥发送或接收的生存期，在不同时期可以使用不同的认证密钥。

相关配置

创建密钥链和密钥

在全局模式下，使用 **key chain** *key-chain-name* 定义密钥链，并进入密钥链配置模式。

在密钥链配置模式，使用 **key** *key-id* 定义一个密钥，并进入密钥链密钥配置模式。

配置认证串

在密钥链密钥配置模式下，使用 **key-string** [0|7] *text* 指定认证串。

- 默认配置明文认证串，0 代表配置明文认证串。
- 7 代表配置密文认证串。
- 加密认证服务默认关闭，使用 **service password-encryption** 开启加密服务，可将明文认证串强制转成密文。

配置生存期

在密钥链密钥配置模式下，可配置密钥链接收方向和发送方向的生存期

- **accept-lifetime** *start-time* { **infinite** | *end-time* | **duration** *seconds* }，配置密钥链接收方向的生存期。
- **send-lifetime** *start-time* { **infinite** | *end-time* | **duration** *seconds* }，配置密钥链发送方向的生存期。

2.4 配置详解

配置项	配置建议 & 相关命令	
配置密钥链	 必须配置。用于创建密钥。	
	key chain	创建密钥链
	key	配置密钥 ID
	key-string	配置密钥串
	accept-lifetime	配置接收方向的生存期
	send-lifetime	配置发送方向的生存期

2.4.1 配置密钥链

配置效果

- 定义一个密钥链，供路由协议使用。

注意事项

- 密钥链必须关联到路由协议中，才能实际生效。

配置方法

创建密钥链

- 如果需要使用密钥链则必须配置。
- 若无特殊要求，应在每台需要进行路由协议认证的路由器上配置。

配置密钥 ID

- 如果需要使用密钥链则必须配置。
- 若无特殊要求，应在每台需要进行路由协议认证的路由器上配置。

配置密钥串

- 如果需要使用密钥链则必须配置。
- 若无特殊要求，应在每台需要进行路由协议认证的路由器上配置。

配置接收方向的生存期

- 可选配置。
- 如果未配置发送方向的生存期，则密钥链一直有效。

配置发送方向的生存期

- 可选配置。
- 如果未配置发送方向的生存期，则密钥链一直有效。

检验方法

- 在路由协议中引用密钥，并观察此路由协议建立的邻居关系。密钥不一致则不能建立邻居。

相关命令

配置密钥链

- 【命令格式】 **key chain** *key-chain-name*
- 【参数说明】 *key-chain-name*：密钥链名。
- 【命令模式】 全局模式
- 【使用指导】 要使密钥链生效，必须配置至少一个密钥。

配置密钥 ID

- 【命令格式】 **key** *key-id*
- 【参数说明】 *key-id*：密钥链中认证密钥的 ID，取值范围：0-2147483647
- 【命令模式】 密钥链配置模式
- 【使用指导】 -

配置密钥认证串

- 【命令格式】 **key-string** [0|7] *text*
- 【参数说明】 0：指定密钥以明文显示
7：指定密钥以密文显示
text：指定的认证串字符
- 【命令模式】 密钥链密钥配置模式
- 【使用指导】 -

配置发送生存期

【命令格式】 **send-lifetime** *start-time* {**infinite** | *end-time* | **duration** *seconds*}

【参数说明】 *start-time* : 指定生存期的开始时间

infinite : 该参数表示该密钥从开始时间一直有效。

end-time : 密钥有效期结束时间, 必须晚于 *start-time*

duration *seconds* : 指定有效期从 *start-time* 开始的持续时间, 取值范围: 1-2147483646

【命令模式】 密钥链密钥配置模式

【使用指导】 使用该命令定义密钥发送方向的生存期。

配置接收生存期

【命令格式】 **accept-lifetime** *start-time* {**infinite** | *end-time* | **duration** *seconds*}

【参数说明】 *start-time* : 指定生存期的开始时间

infinite : 该参数表示该密钥从开始时间一直有效。

end-time : 密钥有效期结束时间, 必须晚于 *start-time*

duration *seconds* : 指定有效期从 *start-time* 开始的持续时间, 取值范围: 1-2147483646

【命令模式】 密钥链密钥配置模式

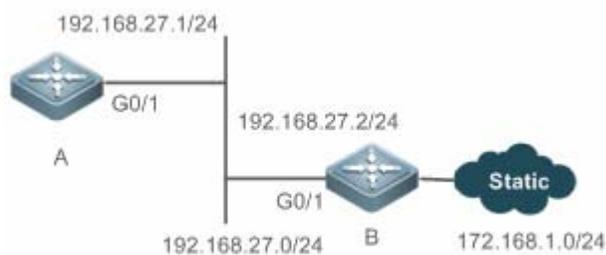
【使用指导】 使用该命令定义密钥接收方向的生存期。

配置举例

配置密钥链, 并在 RIP 报文认证中引用

【网络环境】

图 2-2



- 【配置方法】
- 在所有路由设备上配置密钥
 - 在所有路由设备上配置 RIP
 - 在所有路由设备上启用 RIP 认证

A

```
A>enable
A#configure terminal
A(config)#key chain ripchain
A(config-keychain)#key 1
A(config-keychain-key)#key-string Hello
A(config-keychain-key)#accept-lifetime 16:30:00 Oct 1 2013 duration 43200
A(config-keychain-key)#send-lifetime 16:30:00 Oct 1 2013 duration 43200
A(config-keychain-key)#exit
A(config-keychain)#key 2
A(config-keychain-key)#key-string World
```

```
A(config-keychain-key)#accept-lifetime 04:00:00 Oct 2 2013 infinite
A(config-keychain-key)#send-lifetime 04:00:00 Oct 2 2013 infinite
A(config-keychain-key)#exit
A(config)#interface gigabitEthernet 0/1
A(config-if)#ip address 192.168.27.1 255.255.255.0
A(config-if)#ip rip authentication key-chain ripchain
A(config-if)#ip rip authentication mode md5
A(config-if)#exit
A(config)#router rip
A(config-router)#version 2
A(config-router)#network 192.168.27.0
```

B

```
B>enable
B#configure terminal
B(config)#key chain ripchain
B(config-keychain)#key 1
B(config-keychain-key)#key-string Hello
B(config-keychain-key)#accept-lifetime 16:30:00 Oct 1 2013 duration 43200
B(config-keychain-key)#send-lifetime 16:30:00 Oct 1 2013 duration 43200
B(config-keychain-key)#exit
B(config-keychain)#key 2
B(config-keychain-key)#key-string World
B(config-keychain-key)#accept-lifetime 04:00:00 Oct 2 2013 infinite
B(config-keychain-key)#send-lifetime 04:00:00 Oct 2 2013 infinite
B(config-keychain-key)#exit
B(config)#interface gigabitEthernet 0/1
B(config-if)#ip address 192.168.27.2 255.255.255.0
B(config-if)#ip rip authentication key-chain ripchain
B(config-if)#ip rip authentication mode md5
B(config-if)#exit
B(config)#router rip
B(config-router)#version 2
B(config-router)#network 192.168.27.0
B(config-router)#redistribute static
```

【检验方法】 通过 show ip route rip 看 A 设备是否能收到 B 发出的 RIP 路由

A

```
A(config)#show ip route rip
R    172.168.0.0/16 [120/1] via 192.168.27.2, 00:05:16, GigabitEthernet 0/1
```

常见错误

- 密钥未正确关联到路由协议，导致认证不生效
- 多台设备上配置的密钥不一致，导致认证失败

2.5 监视与维护

查看运行情况

作用	命令
查看密钥链的配置信息	show key chain [<i>key-chain-name</i>]

3 RIP

3.1 概述

RIP (Routing Information Protocol , 路由信息协议) 是一种单播路由协议 , 应用于 IPv4 网络。运行 RIP 的路由器之间通过交换路由信息 , 获得到达远程网络的路由。

RIP 属于 IGP (Interior Gateway Protocol , 内部网关协议) , 只能在 AS 内运行 , 适用于规模较小的网络 (不大于 16 跳) 。

协议规范

- RFC1058 : 定义了 RIPv1
- RFC2453 : 定义了 RIPv2

3.2 典型应用

典型应用	场景描述
RIP 基本应用	在一个小型网络中 , 通过 RIP 自动维护路由信息。

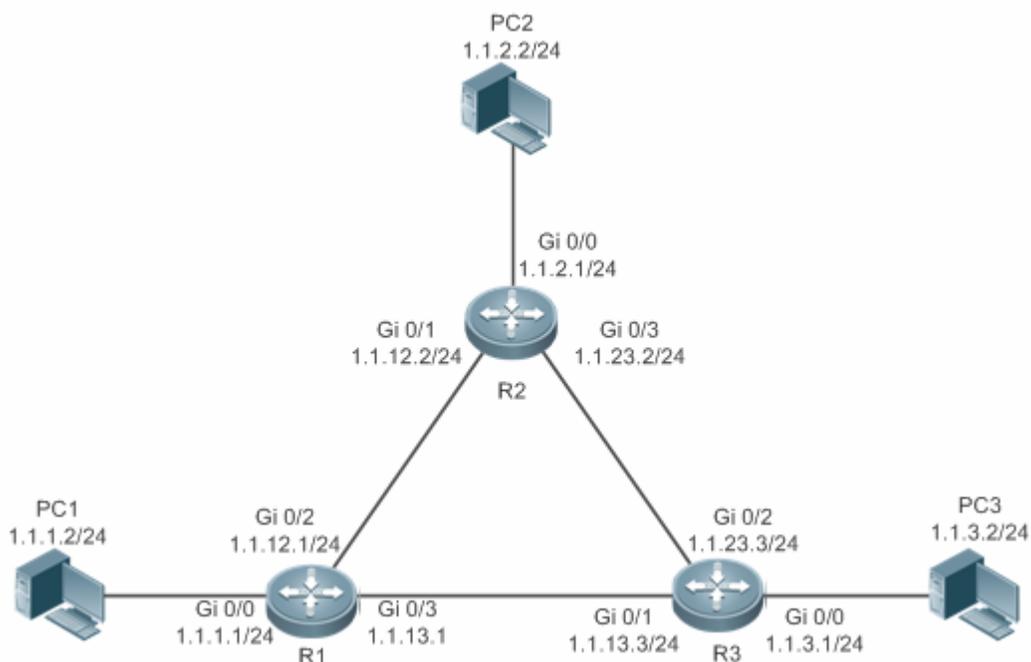
3.2.1 RIP基本应用

应用场景

在组网结构比较简单的网络中 , 可以通过配置 RIP 实现网络互通 , 相比 OSPF 等 IGP 协议 RIP 的配置更为简单 , 相比静态路由 RIP 能够动态适应网络结构变化 , 更容易维护。

以下图为例 , 为了使得 PC1 , PC2 , PC3 互通 , 可以在 R1 , R2 , R3 上配置 RIP 路由。

图 3-1



功能部属

- 在三台 PC 上配置 IP 和网关；
- 在三台路由设备上配置 IP 和掩码
- 在三台路由设备上配置 RIP 协议

3.3 功能详解

基本概念

▾ IGP、EGP

内部网关协议（Interior Gateway Protocol，简称 IGP）：在一个自治系统内部运行。比如：RIP。

外部网关协议（Exterior Gateway Protocol，简称 EGP）：运行于不同自治系统之间。比如：BGP。

▾ 有类路由协议、无类路由协议

从支持有类路由还是无类路由的角度来分类：

有类路由协议（Classful Routing Protocol）：支持有类路由。比如：RIPv1。

无类路由协议（Classless Routing Protocol）：支持无类路由。比如：RIPv2。

功能特性

功能特性	说明
RIPv1 与 RIPv2	RIP 有两个版本：RIPv1 和 RIPv2。
交换路由信息	通过在设备间交换路由信息，RIP 可以自动获得远程网络的路由，并保持实时更新。
路由算法	RIP 是一种基于距离矢量（Distance-Vector）算法的协议，运用矢量叠加的方法来计算路由信息。
避免路由环路	RIP 使用“水平分割”和“毒性逆转”等功能来避免路由环路。
安全措施	RIP 使用“认证”、“源地址校验”等功能来保障协议安全。
可靠性措施	RIP 使用“与 BFD 联动”、“快速重路由”等功能来增强协议可靠性。
多实例	RIP 支持多实例，支持 VPN 应用。

3.3.1 RIPv1 与 RIPv2

RIP 有两个版本：RIPv1、RIPv2。

工作原理

▾ RIPv1

RIPv1 以广播方式发送 RIPv1 报文，广播地址 255.255.255.255，UDP 端口号 520。RIPv1 不能识别子网掩码，仅支持有类别路由。

▾ RIPv2

RIPv2 以组播方式发送 RIPv2 报文，组播地址 224.0.0.9，UDP 端口号 520。RIPv2 能够识别子网掩码，支持无类别路由、路由汇聚、超网路由。RIPv2 支持认证，提供明文认证和 MD5 认证两种方式。

相关配置

▾ 启动 RIP 进程

缺省情况下，未启动 RIP 进程。

使用 **router rip** 命令，启动 RIP 进程。

必须在设备上启动 RIP 进程，否则与 RIP 相关的任何功能都无法生效。

↳ 在接口上运行 RIP

缺省情况下，接口上未运行 RIP。

使用 **network** 命令，定义一个地址范围，则在此地址范围内的接口上运行 RIP。

在接口上运行 RIP 后，接口上才能交互 RIP 报文，RIP 才能学到其直连网段的路由。

↳ 定义 RIP 版本

缺省情况下，接口接收 RIPv1 和 RIPv2 报文，发送 RIPv1 报文。

使用 **version** 命令，定义所有接口上接收和发送的 RIP 报文的版本号。

使用 **ip rip send version** 命令，在接口上定义发送的 RIP 报文的版本号。

使用 **ip rip receive version** 命令，在接口上定义接收的 RIP 报文的版本号。

 如果相邻路由器上运行的 RIP 版本不一致，运行 RIPv1 的路由器会学习到错误的路由。

↳ 限制接口上的报文交互

缺省情况下，运行 RIP 的接口允许接收、发送 RIP 报文。

使用 **no ip rip receive enable** 命令，禁止接口接收 RIP 报文。

使用 **no ip rip send enable** 命令，禁止接口发送 RIP 报文。

使用 **passive-interface** 命令，禁止接口以广播或组播的方式发送 RIP 报文。

↳ 发送 RIP 报文的其他方式

缺省情况下，以广播方式发送 RIPv1 报文，以组播方式发送 RIPv2 报文。

使用 **ip rip v2-broadcast** 命令，在接口上以广播方式发送 RIPv2 报文。

使用 **neighbor** 命令，以单播方式向指定的邻居路由器发送 RIP 报文。

3.3.2 交换路由信息

与静态路由相比，动态路由协议的一大优点：通过在设备间交换路由信息，设备可以自动获得远程网络的路由，并保持实时更新。

工作原理

↳ 初始化

路由器启动 RIP 后，会向相邻路由器发送请求报文（Request message），请求全部路由信息（即路由表）。邻居收到请求报文后，回送包含本地路由表的响应报文（Response message）。路由器收到响应报文后，更新本地路由表，同时向相邻路由器发送更新报文（Update message），通告路由更新信息。相邻路由器收到更新报文后，更新本地路由表，并向其它相邻路由器发送更新报文。在一连串的更新通告后，各路由器都能获得并保持最新的路由信息。

↳ 定时更新

缺省情况下，RIP 启用定时更新。相邻路由器之间每隔 30 秒（update timer）进行完全的路由信息交换（即将整个路由表发送给邻居）。1 份更新报文只能包含 25 条路由，发送整个路由表可能需要大量更新报文。通过设置更新报文间的发送延时，可以避免路由信息丢失。

 对于每条非本地路由，如果 180 秒内（invalid timer）得不到刷新，则 metric 值改为 16（不可达）；如果随后的 120 秒内（flush timer）得不到刷新，则从路由表中删除。

↳ 触发更新

启用触发更新后，定时更新自动关闭。当路由信息发生改变时，立即向邻居发送与改变有关的路由（而非完整的路由表），并采用确认+重发机制来保障邻居成功接收。与定时更新相比，触发更新有助于减少泛洪，加速路由收敛。

能够触发更新的事件包括：路由设备启动、接口状态改变、路由信息发生改变（比如 metric 值）、接收到请求报文。

↳ 路由汇聚

RIP 向邻居发送路由信息时，会将属于同一个有类网络的子网路由汇聚成一条路由，发送给邻居。比如 将 80.1.1.0/24 (metric=2) 和 80.1.2.0/24 (metric=3) 汇聚成 80.0.0.0/8 (metric=2) ，汇聚后的路由选用最优的 metric 值。

仅 RIPv2 支持路由汇聚。聚合路由可以缩减路由表，提高交换路由信息的效率。

↳ 超网路由

当路由的掩码长度小于其自然掩码长度时，这条路由称为超网路由。比如：80.0.0.0/6。根据 IP 地址分类，80.0.0.0 属于 A 类网络，其自然掩码长度为 8，所以 80.0.0.0/6 是超网路由。

仅 RIPv2 支持超网路由。

↳ 缺省路由

在路由表中，目的网络为 0.0.0.0/0 的路由为缺省路由。

可以从邻居学到缺省路由，也可以向邻居发送缺省路由。

↳ 路由重分布

对于 RIP 来说，其他类型的路由（比如直连路由、静态路由、其他路由协议的路由）为外部路由。

可以将设备上的外部路由（不包括缺省路由）重分布到 RIP 中，通告给邻居。

↳ 路由过滤

可以通过设置过滤规则，限制相邻设备间交换的路由信息。只有被过滤规则允许的路由信息，才能被发送或接收。

相关配置

↳ 更新报文间的发送延时

缺省情况下，连续发送无延时。

使用 **output-delay** 命令，设置更新报文间的发送延时。

↳ RIP 时钟

缺省情况下，update timer 为 30 秒，invalid timer 为 180 秒，flush 为 120 秒。

使用 **timers basic** 命令，修改 RIP 时钟的超时时间。

增加超时时间，有助于减少路由振荡。缩短超时时间，有助于加快路由收敛。

相邻设备上的 RIP 时钟值必须一致。若没有特殊要求，不建议修改 RIP 时钟。

↳ 触发更新

缺省情况下，启用定时更新。

使用 **ip rip triggered** 命令，在接口上启动触发更新，定时更新关闭。

使用 **ip rip triggered retransmit-timer** 命令，修改更新报文的重新发送间隔，缺省值 5 秒。

使用 **ip rip triggered retransmit-count** 命令，修改更新报文的最大重新发送次数，缺省值 36 次。

↳ 路由汇聚

缺省情况下，如果允许接口发送 RIPv2 报文，则自动进行路由汇聚。

使用 **no auto-summary** 命令，关闭自动汇聚。

使用 **ip rip summary-address** 命令，在接口上配置路由汇聚。

↳ 超网路由

缺省情况下，如果允许接口发送 RIPv2 报文，则允许发送超网路由。

使用 **no ip rip send supernet-routes** 命令，禁止发送超网路由。

↳ 缺省路由

使用 **ip rip default-information** 命令，在接口上向邻居通告缺省路由。

使用 **default-information originate** 命令，从所有接口向邻居通告缺省路由。

↘ 路由重分布

使用 **redistribute** 命令，将外部路由（不包括缺省路由）重分布到 RIP，通告给邻居。

↘ 路由过滤

使用 **distribute-list out** 命令，设置过滤规则，限制设备发出的路由信息。

使用 **distribute-list in** 命令，设置过滤规则，限制设备接收的路由信息。

3.3.3 路由算法

RIP 是一种基于距离矢量（Distance-Vector）算法的协议，运用矢量叠加的方法来计算路由信息。

工作原理

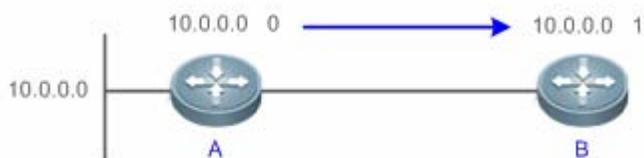
↘ 距离矢量算法

RIP 是一种基于距离矢量算法的协议。“距离矢量”就是将一条路由信息看作一个由目的网络和距离（metric）组成的矢量。路由器从邻居处获得一条路由信息，并在这条路由信息上叠加从自己到这个邻居的距离矢量，从而形成自己的路由信息。

RIP 使用跳数（Hop Count）来衡量到达目的网络的距离（metric）。缺省情况下，路由器到与它直接相连网络的跳数为 0，通过一个路由器可达的网络的跳数为 1，其余依此类推。也就是说，metric 等于从本网络到达目的网络间的路由器数量。为限制收敛时间，RIP 规定 metric 取 0~15 之间的整数，大于或等于 16 表示目的网络或主机不可达。由于这个限制，使得 RIP 不可能在大型网络中得到应用。

如下图所示，A 连接网络 10.0.0.0。B 从 A 处获得路由信息（10.0.0.0，0），则在此基础上叠加 metric 1，得到自己的路由信息（10.0.0.0，1），下一跳指向 A。

图 3-2

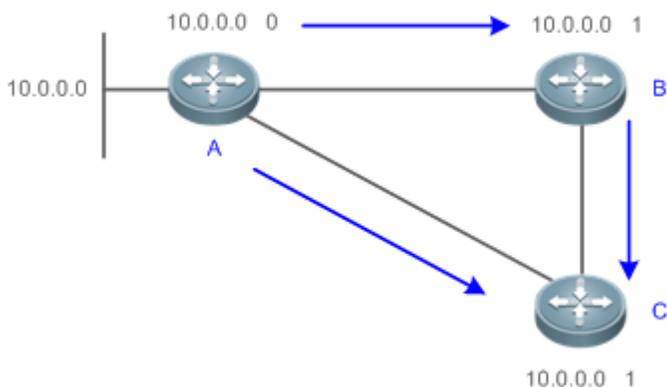


↘ 选择最优路由

RIP 选择最优路由的原则：如果路由器获得了多条到达同一目的网络的路由信息，优先选用 metric 值较小的路由信息。

如下图所示，A 连接网络 10.0.0.0。C 从 A 处获得路由信息（10.0.0.0，0），从 B 处获得路由信息（10.0.0.0，1），则 C 会选用来自 A 的路由信息，并在此基础上叠加 metric 1，得到自己的路由信息（10.0.0.0，1），下一跳指向 A。

图 3-3



i 当路由器上存在多种来源的路由时，优先选用 **distance** 值较小的路由。

路由来源	缺省 distance 值
直连网络	0
静态路由	1
OSPF 路由	110
RIP 路由	120
不可达路由	255

相关配置

修改 distance 值

缺省情况下，RIP 路由的 distance 值为 120。

使用 **distance** 命令，可以修改 RIP 路由的 distance 值。

修改 metric 值

对于设备主动发现的 RIP 路由，缺省 metric 值等于从本地到目的网络的跳数。对于人为配置的 RIP 路由（缺省路由、重分布路由），缺省 metric 值为 1。

使用 **offset-list in** 命令，可以增加收到的 RIP 路由的 metric 值。

使用 **offset-list out** 命令，可以增加发出的 RIP 路由的 metric 值。

使用 **default-metric** 命令，可以修改重分布路由的缺省 metric 值。

使用 **redistribute** 命令，可以在重分布路由时修改其 metric 值。

使用 **default-information originate** 命令，可以在引入缺省路由时修改其 metric 值。

使用 **ip rip default-information** 命令，可以在创建缺省路由时修改其 metric 值。

3.3.4 避免路由环路

RIP 使用“水平分割”、“毒性逆转”等功能来避免路由环路。

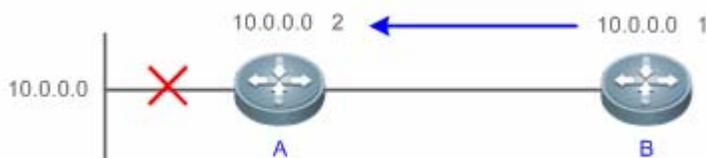
工作原理

路由环路

由于距离矢量算法的固有缺陷，RIP 会发生路由环路。

如下图所示，A 连接网络 10.0.0.0，每隔 30 秒发送一次 Update 报文。因此 B 每 30 秒会收到从 A 发送来的到 10.0.0.0 的路由。如果 A 到 10.0.0.0 的连接中断，到 10.0.0.0 的路由就会从 A 的路由表中消失，下次 A 发出的 Update 报文将不再包含此路由。因为 B 没有收到关于 10.0.0.0 的 Update 报文，在 180 秒内认为到 10.0.0.0 的路由有效，并使用 Update 报文将此路由发送给 A。因为 A 上不存在到 10.0.0.0 的路由，因此将从 B 学到的路由加入路由表。于是，B 认为通过 A 能够到达网络 10.0.0.0，A 认为通过 B 能够到达网络 10.0.0.0，形成路由环路。

图 3-4

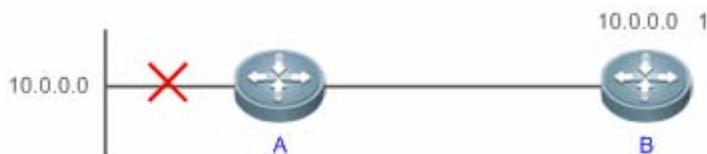


水平分割 (Split Horizon)

水平分割可以防止路由环路。在接口上启动水平分割功能后，从该接口接收到的路由，不会再从该接口发送出去。

如下图所示，在 A 与 B 相连的接口上启动水平分割后，B 不会再把到 10.0.0.0 路由发回给 A，180 秒后 B 会知道 10.0.0.0 不可达。

图 3-5



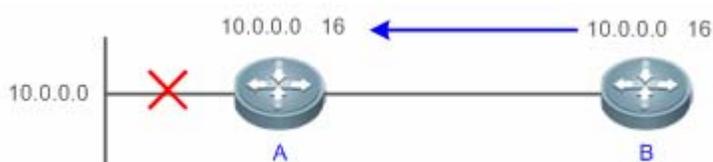
毒性逆转 (Poison Reverse)

毒性逆转也可以防止路由环路。相比水平分割，毒性逆转更加可靠，但会带来更多的协议报文，加重网络拥塞。

在接口上启动毒性逆转后，从该接口接收到的路由，会再从该接口发送出去，且度量值被改为 16（即不可达）。

如下图所示，B 从 A 学习到路由 10.0.0.0 之后，就将该路由的度量设置为 16 定时发回 A。当该路由失效后，B 向 A 通告度量为 16 的 10.0.0.0 将加快该路由信息的清除。

图 3-6



相关配置

水平分割

缺省情况下，启动水平分割。

使用 `no ip rip split-horizon` 命令，关闭水平分割。

毒性逆转

缺省情况下，关闭毒性逆转。

使用 `ip rip split-horizon poisoned-reverse` 命令，启动毒性逆转（水平分割自动关闭）。

3.3.5 安全措施

RIP 使用“认证”、“源地址校验”等功能来保障协议安全。

工作原理

认证

RIPv1 不支持认证，RIPv2 支持认证。

在接口上启动认证功能后，如果认证失败，则相邻设备间无法交换路由信息。认证功能用来阻止非法设备接入 RIP 路由域。

两种认证方式：明文认证和 MD5 认证。

源地址校验

当运行 RIP 的设备收到更新报文时，会校验报文的源 IP 地址是否与入接口 IP 地址在同一网段。如果不在同一网段，则丢弃该报文。源地址校验功能用来保障 RIP 只在相邻的路由设备间交换路由信息。

- i 如果是无编号 IP 地址接口，则不进行源地址校验（不可配置）。
- i 如果启动了触发更新，则自动启用源地址校验（不可配置）。
- i 如果关闭了水平分割，则自动启用源地址校验（不可配置）。

相关配置

认证

缺省情况下，认证功能关闭。

使用 `ip rip authentication mode text` 命令，在接口上启动明文认证。

使用 `ip rip authentication mode md5` 命令，在接口上启动 MD5 认证。

使用 `ip rip authentication text-password` 命令，在接口上设置明文认证的字符串。

使用 `ip rip authentication key-chain` 命令，在接口上引用 KEY 作为认证的密钥。

源地址校验

缺省情况下，启用源地址校验功能。

使用 `no validate-update-source` 命令，关闭源地址校验功能。

3.3.6 可靠性措施

RIP 使用“与 BFD 联动”、“快速重路由”等功能来增强协议可靠性。

工作原理

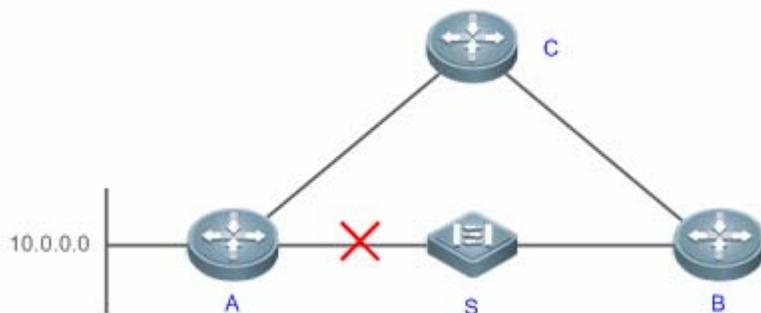
与 BFD 联动、快速重路由

当网络中发生链路故障（或设备故障）时，经过此路由传输的报文将会丢失，直到路由重新收敛。

如下图所示，A 与 S 之间的链路通信发生故障后，B 可能需要等待 180 秒才能发现路由（目的网络 10.0.0.0，下一跳 A）失效。

而后 B 可能还需要等待 30 秒才能获得来自 C 的路由信息，重新获得路由（目的网络 10.0.0.0，下一跳 C）。流量中断 210 秒。

图 3-7



快速发现路由失效，或者快速切换备份路由，都有助于缩短流量中断时间。

- 在 A、B 之间建立 BFD 会话，并与 RIP 联动。BFD 可以快速检测相邻路由器之间的连通状态。一旦发生链路故障，RIP 在 1 秒内就能发现路由失效。
- 使用“快速重路由”功能，预先在 B 上配置一条备份路由（目的网络 10.0.0.0，下一跳 C），一旦 RIP 发现路由失效，将立即启用备份路由。

相关配置

与 BFD 联动

缺省情况下，RIP 与 BFD 之间不存在联动关系。

使用 `bfd all-interfaces` 命令，建立 RIP 与 BFD 之间的联动关系，对所有接口有效。

使用 `ip rip bfd` 命令，在当前接口上建立 RIP 与 BFD 之间的联动关系。

快速重路由

缺省情况下，未启动快速重路由。

使用 `fast-reroute route-map` 命令，启动快速重路由，引用路由图。

使用 `set fast-reroute backup-interface backup-nexthop` 命令，在路由图中配置备份路由。

3.4 配置详解

配置项	配置建议 & 相关命令	
配置 RIP 基本功能	⚠️ 必须配置。用于构建 RIP 路由域。	
	<code>router rip</code>	启动 RIP 路由进程，进入路由进程模式。
	<code>network</code>	在指定地址范围的接口上运行 RIP。
	<code>version</code>	定义 RIP 版本号。
	<code>ip rip split-horizon</code>	在接口上启动水平分割、毒性反转。
	<code>passive-interface</code>	配置被动接口。
控制 RIP 报文交互	⚠️ 可选配置。若希望改变 RIP 协议的报文发送、接收的默认机制，需要配置。	
	<code>neighbor</code>	以单播方式向指定的邻居发送 RIP 报文。
	<code>ip rip v2-broadcast</code>	在接口上以广播方式发送 RIPv2 报文。
	<code>ip rip receive enable</code>	允许接口接收 RIP 报文。
	<code>ip rip send enable</code>	允许接口发送 RIP 报文。
	<code>ip rip send version</code>	定义接口发送 RIP 报文的版本号。
	<code>ip rip receive version</code>	定义接口接收 RIP 报文的版本号。
启动触发更新	⚠️ 可选配置。	
	<code>ip rip triggered</code>	在接口上启动触发更新。
启动源地址校验	⚠️ 可选配置。	
	<code>validate-update-source</code>	启动源地址校验。
启动认证	⚠️ 可选配置。仅 RIPv2 支持认证。	
	<code>ip rip authentication mode</code>	在接口上启动认证，设置认证模式。
	<code>ip rip authentication text-password</code>	在接口上配置明文认证的字符串。
	<code>ip rip authentication key-chain</code>	在接口上配置认证密钥串。
启动路由汇聚	⚠️ 可选配置。仅 RIPv2 支持路由汇聚。	
	<code>auto-summary</code>	允许 RIP 路由自动汇聚。
	<code>ip rip summary-address</code>	在接口上配置路由汇聚。
启动超网路由	⚠️ 可选配置。仅 RIPv2 支持超网路由。	
	<code>ip rip send supernet-routes</code>	在接口上允许通告 RIP 超网路由。
通告缺省路由/外部路由	⚠️ 可选配置。	
	<code>ip rip default-information</code>	在接口上向邻居通告缺省路由。
	<code>default-information originate</code>	向邻居通告缺省路由。
	<code>redistribute</code>	路由重分布，向邻居通告外部路由。
设置路由过滤规则	⚠️ 可选配置。	

	distribute-list in	对收到的 RIP 路由信息进行过滤。
	distribute-list out	对发出的 RIP 路由信息进行过滤。
修改路由选择参数	 可选配置。	
	distance	修改 RIP 路由的管理距离。
	offset-list	增加接收或发送的 RIP 路由的 metric。
	default-metric	配置重分布到 RIP 中的外部路由的缺省 metric。
修改时钟	 可选配置。	
	timers basic	修改 RIP 路由更新、失效、清除的时间。
	output-delay	设置 RIP 路由更新报文的发送延时。
启动与 BFD 联动	 可选配置。	
	bfd all-interfaces	全局启动 RIP 与 BFD 联动。
	ip rip bfd	在接口上启动 RIP 与 BFD 联动。
启动快速重路由	 可选配置。	
	fast-reroute route-map	启动快速重路由，引用路由图。
	set fast-reroute backup-interface backup-nexthop	在路由图中配置快速重路由的备份接口和备份下一跳。

3.4.1 配置RIP基本功能

配置效果

- 在网络中建立一个 RIP 路由域。
- 域内的路由器通过 RIP 获得远端网络的路由。

注意事项

- 必须配置 IPv4 地址。
- 必须启动 IPv4 单播路由。

配置方法

启动 RIP 路由进程

- 必须配置。
- 若无特殊要求，应在 RIP 路由域内的每台路由器上配置。

关联本地网络

- 必须配置。
- 若无特殊要求，应在 RIP 路由域内的每台路由器上配置。
- 若无特殊要求，RIP 关联的本地网络应覆盖所有三层接口所在网段。

定义 RIP 版本

- 如果需要使用 RIPv2 版本的功能（比如：可变长子网掩码、认证。），则必须启动 RIPv2 版本。
- 若无特殊要求，应在每台路由器上定义相同的 RIP 版本。

启动水平分割、毒性逆转

- 缺省启动水平分割，关闭毒性逆转。

- 若无特殊要求，应在每个连接广播类型网络（比如：以太网）的接口上启动水平分割。（缺省即可）。
- 若无特殊要求，应在每个连接在点到点类型网络（比如：PPP、HDLC）的接口上启动水平分割。（缺省即可）。
- 建议在连接非广播多路访问类型网络（比如：FR、X.25）的接口上关闭水平分割和毒性逆转，否则可能导致部分设备学不到全部的路由信息。
- 如果连接非广播网络的接口上配置了次 IP 地址，建议关闭水平分割和毒性逆转。

▾ 设置被动接口

- 如果要在某个 RIP 接口抑制更新报文，则需要设置被动接口。
- 使用被动接口设置 RIP 路由域的边界。被动接口所在网段属于 RIP 路由域，但被动接口上不能发送 RIP 报文。
- RIP 路由域内的需要交互 RIP 路由的接口（比如：路由器互联接口）不能设置为被动接口。

检验方法

- 检查路由器上的路由表，确认通过 RIP 获得远端网络的路由。

相关命令

▾ 启动 RIP 路由进程

- 【命令格式】 **router rip**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 本命令用来创建 RIP 路由进程，并进入路由进程模式。

▾ 关联本地网络

- 【命令格式】 **network network-number [wildcard]**
- 【参数说明】 **network-number**：网络号。
wildcard：IP 地址比较比特位，0 表示精确匹配，1 表不做比较。
- 【命令模式】 路由进程模式
- 【使用指导】 只有被 **network** 覆盖的接口才能运行 RIP，才能被 RIP 学到直连路由，才能交互 RIP 报文。
配置 **network 0.0.0.0 255.255.255.255**，即可覆盖所有接口。
如果未配置 **wildcard** 参数，默认按照有类地址范围来处理，使落在该有类地址范围内的接口地址参与 RIP 运行。

▾ 定义 RIP 版本

- 【命令格式】 **version { 1 | 2 }**
- 【参数说明】 **1**：RIPv1。
2：RIPv2。
- 【命令模式】 全局模式
- 【使用指导】 此命令对整机生效，定义所有接口上发送和接收的 RIP 报文的版本号。

▾ 启动水平分割

- 【命令格式】 **ip rip split-horizon [poisoned-reverse]**
- 【参数说明】 **poisoned-reverse**：毒性逆转
- 【命令模式】 接口模式
- 【使用指导】 启动毒性逆转后，水平分割自动关闭。

▾ 配置被动接口

- 【命令格式】 **passive-interface { default | interface-type interface-num }**

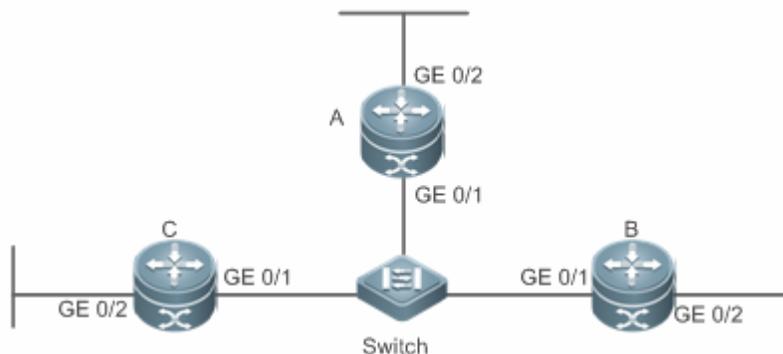
- 【参数说明】 **default** : 所有接口。
interface-type interface-num : 指定接口。
- 【命令模式】 路由进程模式
- 【使用指导】 先配置 **passive-interface default** , 将所有接口设置为被动接口。
 再配置 **no passive-interface interface-type interface-num** 取消与域内路由器互联的接口。

配置举例

建立 RIP 路由域

【网络环境】

图 3-8



【注释】 接口 IP 地址 :

A : GE0/1 110.11.2.1/24 GE0/2 155.10.1.1/24
 B : GE0/1 110.11.2.2/24 GE0/2 196.38.165.1/24
 C : GE0/1 110.11.2.3/24 GE0/2 117.102.0.1/16

【配置方法】

- 所有路由器配置接口 IP 地址。
- 所有路由器配置 RIP 基本功能。

A

```
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip address 110.11.2.1 255.255.255.0
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)# ip address 155.10.1.1 255.255.255.0
A(config)# router rip
A(config-router)# version 2
A(config-router)# network 0.0.0.0 255.255.255.255
A(config-router)# passive-interface default
A(config-router)# no passive-interface GigabitEthernet 0/1
```

B

```
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip address 110.11.2.2 255.255.255.0
B(config-if-GigabitEthernet 0/1)# exit
B(config)# interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/2)# ip address 196.38.165.1 255.255.255.0
B(config-if-GigabitEthernet 0/2)# exit
```

C

```
B(config)# router rip
B(config-router)# version 2
B(config-router)# network 0.0.0.0 255.255.255.255
B(config-router)# passive-interface default
B(config-router)# no passive-interface GigabitEthernet 0/1

C# configure terminal
C(config)# interface GigabitEthernet 0/1
C(config-if-GigabitEthernet 0/1)# ip address 110.11.2.3 255.255.255.0
C(config-if-GigabitEthernet 0/1)# exit
C(config)# interface GigabitEthernet 0/2
C(config-if-GigabitEthernet 0/2)# ip address 117.102.0.1 255.255.0.0
C(config-if-GigabitEthernet 0/2)# exit
C(config)# router rip
C(config-router)# version 2
C(config-router)#no auto-summary
C(config-router)# network 0.0.0.0 255.255.255.255
C(config-router)# passive-interface default
C(config-router)# no passive-interface GigabitEthernet 0/1
```

【检验方法】 检查 A、B、C 上的 IP 路由表，应看到 RIP 学到了远端网络的路由（蓝色字段）。

A

```
A# show ip route

Codes: C - connected, S - static, R - RIP, B - BGP
       0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
C    110.11.2.0/24 is directly connected, GigabitEthernet 0/1
C    110.11.2.1/32 is local host.
R    117.0.0.0/8 [120/1] via 110.11.2.2, 00:00:47, GigabitEthernet 0/1
C    155.10.1.0/24 is directly connected, GigabitEthernet 0/2
C    155.10.1.1/32 is local host.
C    192.168.217.0/24 is directly connected, VLAN 1
C    192.168.217.233/32 is local host.
R    196.38.165.0/24 [120/1] via 110.11.2.3, 00:19:18, GigabitEthernet 0/1
```

B

```
B# show ip route

Codes: C - connected, S - static, R - RIP, B - BGP
```

```

0 - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
C 110.11.2.0/24 is directly connected, GigabitEthernet 0/1
C 110.11.2.2/32 is local host.
R 155.10.0.0/16 [120/1] via 110.11.2.1, 00:15:21, GigabitEthernet 0/1
C 196.38.165.0/24 is directly connected, GigabitEthernet 0/2
C 196.38.165.1/32 is local host.
R 117.0.0.0/8 [120/1] via 110.11.2.2, 00:00:47, GigabitEthernet 0/1

```

C

```

C# show ip route

Codes: C - connected, S - static, R - RIP, B - BGP
0 - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set
C 110.11.2.0/24 is directly connected, GigabitEthernet 0/1
C 110.11.2.3/32 is local host.
C 117.102.0.0/16 is directly connected, GigabitEthernet 0/2
C 117.102.0.1/32 is local host.
R 155.10.0.0/16 [120/1] via 110.11.2.1, 00:20:55, GigabitEthernet 0/1
R 196.38.165.0/24 [120/1] via 110.11.2.3, 00:19:18, GigabitEthernet 0/1

```

常见错误

- 接口上未配置 IPv4 地址。
- 某台设备上没有定义 RIP 版本，或 RIP 版本号与其他路由器不一致。
- **network** 命令配置的地址范围未覆盖某接口。
- **network** 命令参数“IP 地址比较比特位”配置错误。0 表示精确匹配，1 表示不做比较。
- 设备互联接口被设置为被动接口。

3.4.2 控制RIP报文交互

配置效果

通过配置，改变 RIP 协议的默认运行机制，人为控制 RIP 报文的交互方式。

包括：

- 允许或禁止某个接口以单播方式向指定的邻居发送 RIP 报文。
- 允许或禁止某个接口以广播而非多播的方式发送 RIP version 2 报文。
- 允许或禁止在某个接口上接收 RIP 报文。
- 允许或禁止在某个接口上发送 RIP 报文。
- 允许或禁止在某个接口上接收特定版本的 RIP 报文。
- 允许或禁止在某个接口上发送特定版本的 RIP 报文。

注意事项

- 必须配置 RIP 基本功能。
- 邻居设备的连接接口上，配置的发送和接收版本需要保持一致。

配置方法

以单播方式向指定的邻居发送 RIP 路由更新

- 如果不希望接口所连接的设备均能收到路由信息更新，而只希望其中部分设备收到路由信息，则必须配置该功能。
- RIPv1 缺省使用 IP 广播地址（255.255.255.255）通告路由信息，RIPv2 缺省使用组播地址（224.0.0.9）通告路由信息。如果不希望广播网或非广播多路访问网上的全部设备，均可接收到路由信息，可以配置相应接口设置为被动接口，然后只定义某些邻居可以接收到路由信息。该命令不会影响 RIP 信息报文的接收。在接口上以广播方式发送 RIPv2 报文
- 若无特殊要求，应在以单播方式更新路由信息的路由设备上启用该功能。

在接口上以广播方式发送 RIPv2 报文

- 如果邻居设备不支持以组播方式接收 RIPv2 报文，则必须配置该功能。
- 若无特殊要求，应在每台路由器上，以广播方式发送 RIPv2 报文的接口上配置该功能。

在接口上允许接收 RIP 报文

- 该功能缺省启用，如果希望某接口不接收 RIP 报文，则应关闭该功能。
- 若无特殊要求，应在每台路由器上，不允许接收 RIP 报文的接口上配置该功能。

在接口上允许发送 RIP 报文。

- 该功能缺省启用，如果希望某接口不发送 RIP 报文，则应关闭该功能。
- 若无特殊要求，应在每台路由器上，不允许发送 RIP 报文的接口上配置该功能。

在接口上允许发送指定版本的 RIP 报文

- 如果某接口上允许发送的 RIP 报文版本需要设置为与全局不同，则必须配置该功能。
- 若无特殊要求，应在每台路由器上，需要设置允许发送的 RIP 报文版本的接口上配置该功能。

在接口上允许接收指定版本的 RIP 报文

- 如果某接口上允许接收的 RIP 报文版本需要设置为与全局不同，则必须配置该功能。
- 若无特殊要求，应在每台路由器上，需要设置允许接收的 RIP 报文版本的接口上配置该功能。

检验方法

打开调试开关 `debug ip rip packet`，查看报文是否发送以及报文的类型。

相关命令

以单播方式向指定的邻居发送 RIP 路由更新

【命令格式】 `neighbor ip-address`

【参数说明】 `ip-address`：邻居的 IP 地址。应该是本地设备直连网络地址。

- 【命令模式】 路由进程模式
- 【使用指导】 一般可以先用路由进程配置命令 **passive-interface** 将相应接口设置为被动接口，然后只定义某些邻居可以接收到路由信息。该命令不会影响 RIP 信息报文的接收。设置了 **passive** 属性的接口，重启后也不会发送请求报文。

在接口上以广播方式发送 RIPv2 报文

- 【命令格式】 **ip rip v2-broadcast**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 缺省行为取决于 **version** 命令的配置。该命令的配置结果，可以覆盖用 **version** 命令的缺省配置。该命令只影响本接口发送 RIP 报文的的行为，可以允许该接口同时发送 RIPv1 和 RIPv2 的数据包。此命令配置时不带参数则接收行为取决于 **version** 的配置。

在接口上允许接收 RIP 报文

- 【命令格式】 **ip rip receive enable**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 如果要禁止某接口接收 RIP 数据包，在该接口的接口模式下使用该命令的 **no** 形式配置。命令仅作用于配置了此命令的接口。可使用该命令的 **default** 形式恢复接口接收 RIP 数据包。

在接口上允许发送 RIP 报文。

- 【命令格式】 **ip rip send enable**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 如果要禁止某接口发送 RIP 数据包，在该接口的接口模式下使用该命令的 **no** 形式配置。命令仅作用于配置了此命令的接口。可使用该命令的 **default** 形式恢复接口发送 RIP 数据包。

在接口上允许发送指定版本的 RIP 报文

- 【命令格式】 **ip rip send version [1] [2]**
- 【参数说明】
 - 1：只发送 RIPv1 的数据包。
 - 2：只发送 RIPv2 的数据包。
- 【命令模式】 接口模式
- 【使用指导】 缺省行为取决于 **version** 命令的配置。该命令的配置结果，可以覆盖用 **version** 命令的缺省配置。该命令只影响本接口发送 RIP 报文的的行为，可以允许该接口同时发送 RIPv1 和 RIPv2 的数据包。此命令配置时不带参数则接收行为取决于 **version** 的配置。

在接口上允许接收指定版本的 RIP 报文

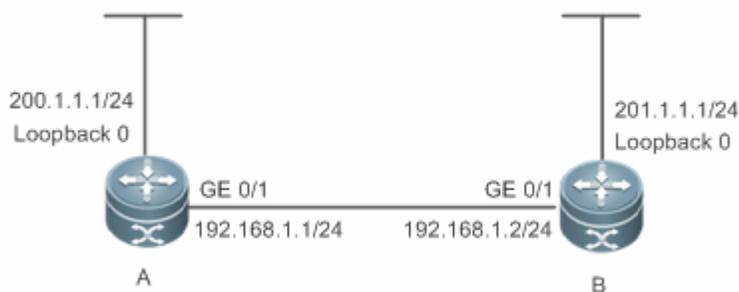
- 【命令格式】 **ip rip receive version [1] [2]**
- 【参数说明】
 - 1：只接收 RIPv1 的数据包。
 - 2：只接收 RIPv2 的数据包。
- 【命令模式】 接口模式
- 【使用指导】 缺省行为取决于 **version** 命令的配置。该命令的配置结果，可以覆盖用 **version** 命令的缺省配置。该命令只影响本接口接收 RIP 报文的的行为，可以允许该接口同时接收 RIPv1 和 RIPv2 的数据包。此命令配置时不带参数则接收行为取决于 **version** 的配置。

配置举例

在接口上禁止发送 RIP 报文

【网络环境】

图 3-9



- 【配置方法】
- 所有路由器配置接口 IP 地址。略
 - 所有路由器配置 RIP 基本功能。略
 - 路由器 A 配置禁止 RIP 报文发送。

```
A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# no ip rip send enable
```

【检验方法】 路由器 A 打开 **debug ip rip packet send**，报文无法发送。

```
A
A# debug ip rip packet recv
*Nov 4 08:19:31: %RIP-7-DEBUG: [RIP] Prepare to send BROADCAST response...
*Nov 4 08:19:31: %RIP-7-DEBUG: [RIP] Building update entries on GigabitEthernet 0/1
*Nov 4 08:19:31: %RIP-7-DEBUG: 117.0.0.0/8 via 0.0.0.0 metric 1 tag 0
*Nov 4 08:19:31: %RIP-7-DEBUG: [RIP] Interface GigabitEthernet 0/1 is disabled to send RIP packet!
```

常见错误

邻居配置的 RIP 协议版本不一致，无法兼容。

3.4.3 启动触发更新

配置效果

- 开启 RIP 的触发扩展功能，RIP 协议不再定时发送路由更新。

注意事项

- 必须配置 RIP 基本功能。
- 建议同时启用带毒性逆转的水平分割功能，否则可能会出现无效路由信息残留等现象。
- 该功能不能与“RIP 与 BFD 联动”功能同时启用。
- 必须保证连接在同一链路上的每台路由由设备都启动了触发更新，否则无法正确交互路由信息。

配置方法

启动触发更新

- 如果广域网接口为按需链路，则必须启用触发更新功能
- 允许在以下两种情况下启用该功能：A.接口仅存在一个邻居；B.存在多个邻居但是以单播方式交互。

- 建议在广域网接口上（运行 PPP、帧中继、X.25 等链路层协议）启动触发更新，以满足按需链路的要求。
- 如果接口启用了该功能，则无论是否启用源地址校验功能（`validate-update-source`），均执行源地址校验。
- 若若无特殊要求，应在每台路由器上的按需链路上启用触发更新功能。

检验方法

当启用了 RIP 触发扩展功能，RIP 协议不再定时发送路由更新，只在以下事件发生的时候 RIP 才会向广域网接口发送路由更新：

- 当接收到路由请求报文。
- RIP 路由信息发生变化。
- 当接口状态变化。
- 当路由设备启动。

相关命令

启动触发更新

【命令格式】 `ip rip triggered { retransmit-timer timer | retransmit-count count }`

【参数说明】 `retransmit-timer timer`：配置重传 Update Request、Update Response 报文的时间间隔，缺省为 5 秒，取值范围：1-3600

`retransmit-count count`：配置 Update Request、Update Response 报文的最大重传次数，缺省为 36 次，取值范围：1-3600。

【命令模式】 接口模式

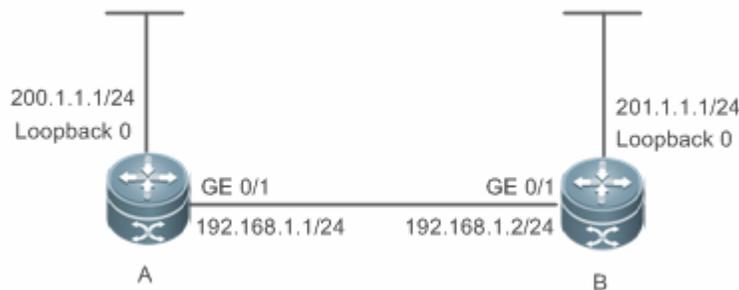
【使用指导】 配置 `ip rip triggered` 可以直接启用 RIP 触发扩展功能
触发扩展功能取消了 RIP 定时更新，因此需要采用确认和重发机制来保证更新报文在广域网上的成功发送和接收。可以使用 `retransmit-timer` 和 `retransmit-count` 指定请求和更新报文的重新间隔和最大重传次数。

配置举例

启动触发更新

【网络环境】

图 3-10



- 【配置方法】
- 所有路由器配置接口 IP 地址。略
 - 所有路由器配置 RIP 基本功能。略
 - 路由器 A 启用 RIP 触发扩展功能，并指定 Update Request、Update Response 报文的重新间隔为 10s，最大重传次数为 18。

```
A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# encapsulation ppp
```

```
A(config-if-GigabitEthernet 0/1)# ip rip triggered
A(config-if-GigabitEthernet 0/1)# ip rip triggered retransmit-timer 10
A(config-if-GigabitEthernet 0/1)# ip rip triggered retransmit-count 18
A(config-if-GigabitEthernet 0/1)# ip rip split-horizon poisoned-reverse
A(config)# router rip
A(config-router)# network 192.168.1.0
A(config-router)# network 200.1.1.0
B
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# encapsulation ppp
B(config-if-GigabitEthernet 0/1)# ip rip triggered
B(config-if-GigabitEthernet 0/1)# ip rip split-horizon poisoned-reverse
B(config)# router rip
B(config-router)# network 192.168.1.0
B(config-router)# network 201.1.1.0
```

【检验方法】 路由器 A、B 检查 RIP 的数据库，对应的路由属性是永久的。

```
A
A# sho ip rip database
201.1.1.0/24 auto-summary
201.1.1.0/24
    [1] via 192.168.12.2 GigabitEthernet 0/1 06:25 permanent
B
B# sho ip rip database
200.1.1.0/24 auto-summary
200.1.1.0/24
    [1] via 192.168.12.1 GigabitEthernet 0/1 06:25 permanent
```

常见错误

- 链路两端的 RIP 配置一致，但确启用了该功能。
- 与 BFD 同时启用
- 未在同链路所有路由器上启用该功能

3.4.4 启动源地址校验

配置效果

- 对接收到的 RIP 路由更新报文的源地址，进行有效性验证。

注意事项

- 必须配置 RIP 基本功能。

配置方法

▾ 启动源地址校验

- 该功能缺省启用，如果不需要进行源地址校验，则必须配置关闭该功能。
- 在接口上关闭水平分割，将引起 RIP 路由进程执行更新报文源地址验证，不管是否执行了路由进程配置命令 **validate-update-source**。
- 对于 ip unnumbered 接口，RIP 路由进程将不执行更新报文源地址验证，不管是否执行了路由进程配置命令 **validate-update-source**。
- 若无特殊要求，应在不需要进行源地址校验的每台路由器上关闭该功能。

检验方法

只接收来自同一个 IP 子网邻居的路由更新。

相关命令

【命令格式】 **validate-update-source**

【参数说明】 -

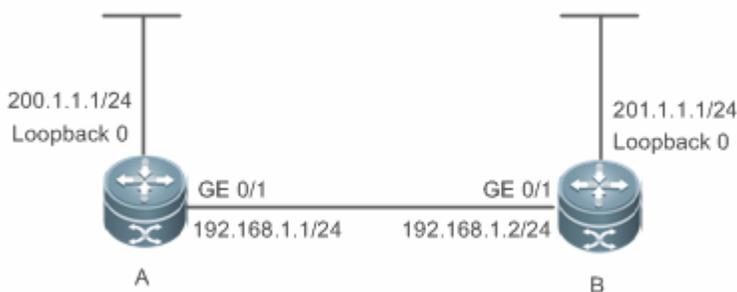
【命令模式】 路由进程模式

【使用指导】 缺省启用更新报文源地址验证。开启后可以对 RIP 路由更新报文的源地址，进行有效性验证。验证的目的是为了保证 RIP 路由进程，只接收来自同一个 IP 子网邻居的路由更新。

配置举例

【网络环境】

图 3-11



- 【配置方法】
- 所有路由器配置接口 IP 地址。略
 - 所有路由器配置 RIP 基本功能。略。
 - 所有路由器关闭更新报文源地址验证。

```
A
A# configure terminal
A(config)# router rip
A(config-router)# no validate-update-source
```

```
B
B# configure terminal
B(config)# router rip
B(config-router)# no validate-update-source
```

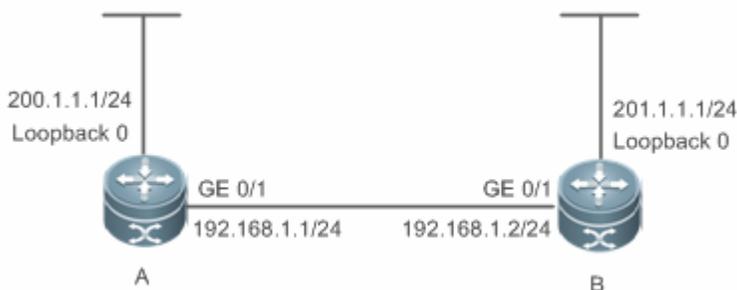
- 【检验方法】
- 检查路由器 A 的路由表，已加载 201.1.1.0/24 表项。
 - 检查路由器 B 的路由表，已加载 200.1.1.0/24 表项。

```
A
A# show ip route rip
R 201.1.1.0/24 [120/1] via 192.168.2.2, 00:06:11, GigabitEthernet 0/1
```

```
B
B# show ip route rip
```

【网络环境】

图 3-11



【配置方法】

- 所有路由器配置接口 IP 地址。略
- 所有路由器配置 RIP 基本功能。略。
- 所有路由器关闭更新报文源地址验证。

A

```
A# configure terminal
A(config)# router rip
A(config-router)# no validate-update-source
```

B

```
B# configure terminal
B(config)# router rip
B(config-router)# no validate-update-source
```

【检验方法】

- 检查路由器 A 的路由表，已加载 201.1.1.0/24 表项。
- 检查路由器 B 的路由表，已加载 200.1.1.0/24 表项。

```
R 200.1.1.0/24 [120/1] via 192.168.1.1, 00:06:11, GigabitEthernet 0/1
```

常见错误

3.4.5 启动认证

配置效果

- 防止学到非认证、无效路由，避免通告有效路由到非认证设备，保证路由系统的稳定性和抗入侵性。

注意事项

- 必须配置 RIP 基本功能。
- RIPv1 不支持 RIP 数据包认证，只有 RIPv2 支持。

配置方法

▾ 启用并指定 RIP 认证所使用的密钥串

- 如果需要启动认证，则必须执行该配置。
- 如果密钥串在接口配置中指定了，请使用 key chain 全局配置命令进行该密钥串的定义，否则将可能导致 RIP 数据包认证失败。
- 若无特殊要求，应在每台需要启动认证的路由器上执行该配置。

▾ 定义 RIP 认证模式

- 如果需要启动认证，则必须执行次配置。
- 当配置 RIP 认证时，所有需要直接交换 RIP 路由信息的设备，所配置的 RIP 认证模式必须一致，否则 RIP 数据包交换将失败。
- 如果采用明文认证，但未配置明文认证字符串，或者未配置关联密钥串，此时并不会发生认证行为。同样，如果采用 MD5 认证，但未配置关联密钥串，也不会有认证行为发生。
- 若无特殊要求，应在每台需要启动认证的路由器上执行该配置。

▾ 启用并设置 RIP 明文认证字符串

- 如果需要启动认证，则必须执行次配置。
- 如果要启用 RIP 明文认证功能，可以直接使用本命令配置明文认证字符串，也可以通过关联密钥串获取明文认证字符串，后者优先级高于前者。
- 若无特殊要求，应在每台需要启动认证的路由器上执行该配置。

检验方法

- RIP 明文认证方式只提供有限的安全，因为报文中传送的密码是可直观显示的。
- RIP MD5 加密认证方式则可以提供较高的安全性，因为报文中传送的密码是经过 MD5 算法加密的。
- 配置正确的认证，路由可以正确学习。
- 配置错误的认证，路由无法学习。

相关命令

▾ 启动源地址校验

- 【命令格式】 **ip rip authentication key-chain** *name-of-keychain*
- 【参数说明】 *name-of-keychain* 密钥串名字，指定 RIP 认证所使用的密钥串。
- 【命令模式】 接口模式
- 【使用指导】 所指定的密钥串需要使用 key chain 全局配置命令进行该密钥串的定义。

▾ 定义 RIP 认证模式

- 【命令格式】 **ip rip authentication mode** { **text** | **md5** }
- 【参数说明】 **text** RIP 认证模式为明文认证。
md5 RIP 认证模式为 MD5 认证。
- 【命令模式】 接口模式
- 【使用指导】 所有需要直接交换 RIP 路由信息的设备，所配置的 RIP 认证模式必须一致。

▾ 启用并设置 RIP 明文认证字符串

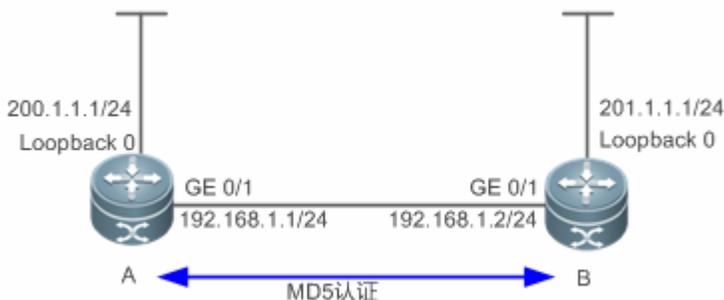
- 【命令格式】 **ip rip authentication text-password** [**0** | **7**] *password-string*
- 【参数说明】 **0** 指定密钥以明文显示
7 指定密钥以密文显示
password-string 明文认证的字符串，长度 1 ~ 16 字节。
- 【命令模式】 接口模式
- 【使用指导】 该命令仅在明文认证模式下生效。

配置举例

▾ 配置 RIP 基本功能并且启用 MD5 认证

【配置方法】

图 3-12



【配置方法】

- 所有路由器配置接口 IP 地址。略
- 所有路由器配置 RIP 基本功能。略。
- 所有路由器配置认证类型和 MD5 认证密钥。

A

```
A# configure terminal
A(config)# key chain hello
A(config-keychain)# key 1
A(config-keychain-key)# key-string world
A(config-keychain-key)# exit
A(config-keychain)# exit
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip rip authentication mode md5
A(config-if-GigabitEthernet 0/1)# ip rip authentication key-chain hello
```

B

```
B# configure terminal
B(config)# key chain hello
B(config-keychain)# key 1
B(config-keychain-key)# key-string world
B(config-keychain-key)# exit
B(config-keychain)# exit
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip rip authentication mode md5
B(config-if-GigabitEthernet 0/1)# ip rip authentication key-chain hello
```

【检验方法】

- 检查路由器 A 的路由表，已加载 201.1.1.0/24 表项。
- 检查路由器 B 的路由表，已加载 200.1.1.0/24 表项。

A

```
A# show ip route rip
R 201.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1
```

B

```
B# show ip route rip
R 200.1.1.0/24 [120/1] via 192.168.1.1, 00:06:11, GigabitEthernet 0/1
```

常见错误

- 需要交互 RIP 信息的路由器配置的密码不一致

- 需要交互 RIP 信息的路由器配置认证方式不同

3.4.6 启动路由汇聚

配置效果

缩小了路由表的规模，提高路由效率，并在一定程度上规避了路由震荡的影响，提高了网络的伸缩性和有效性。

- ① 如果有汇总路由存在，在路由表中将看不到包含在汇总路由内的子路由，这样可以大大缩小路由表的规模。
- ② 通告汇总路由会比通告单独的每条路由将更有效率，主要有以下因素：1) 当查找 RIP 数据库时，汇总路由会得到优先处理；2) 当查找 RIP 数据库时，任何子路由将被忽略，减少了处理时间。

注意事项

- 必须配置 RIP 基本功能
- 超网路由范围比有类网络更大，路由自动汇总功能对超网路由无效。
- RIPv1 总是进行自动汇总，如果需要通告明细路由，请设置协议版本为 RIPv2。

配置方法

▾ 启用路由自动汇总

- 该功能缺省启用。
- 如果希望学到具体的子网路由，而不只是汇总后的网络路由，则需要关闭路由自动汇总功能。
- 只有 RIPv2 允许关闭路由自动汇总功能，RIPv1 总是应用路由自动汇总。

▾ 配置 RIP 接口级汇聚

- 如果需要进行有类子网的汇聚，则必须配置该功能。
- 命令 `ip rip summary-address` 用来在某个指定接口下汇聚一个地址或子网。RIP 自动汇聚到有类网络边界。只有接口汇聚能配置每个有类子网。
- 该命令所配置的汇总范围不能是超类网络，即配置的掩码长度不能小于网络的自然掩码长度。
- 若无特殊要求，应在需要进行有类子网汇聚的路由器上配置该功能。

检验方法

查看对端路由表路由是否有汇聚

相关命令

▾ 启用路由汇总总功能

- 【命令格式】 `auto-summary`
- 【参数说明】 -
- 【命令模式】 路由进程模式
- 【使用指导】 RIPv1 和 RIPv2 中此命令默认开启。
只有配置 RIPv2 时才可以关闭，RIPv1 总是应用。

▾ 配置 RIP 接口级汇聚

- 【命令格式】 `ip rip summary-address ip-address ip-network-mask`
- 【参数说明】 `ip-address` 要被汇聚的 IP 地址
`ip-network-mask` 为指定 IP 地址进行路由汇聚的子网掩码

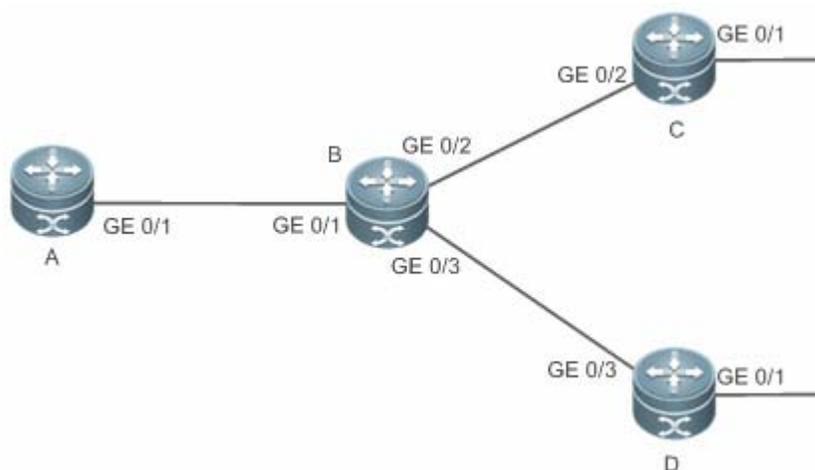
- 【命令模式】 接口模式
 【使用指导】 该用来在某个指定接口下汇聚一个地址或子网

配置举例

配置路由汇聚功能

【网络环境】

图 3-13



【注释】 接口 IP 地址：

- A : GE0/1 192.168.1.1
 B : GE0/1 192.168.1.2 GE0/2 172.16.2.1 GE0/3 172.16.3.1
 C : GE0/2 172.16.2.2 GE0/3 172.16.4.2
 D : GE0/2 172.16.3.2 GE0/3 172.16.5.2

- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 RIP 基本功能。略。
 - 路由器 B 上配置路由汇聚。

```

B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip rip summary-address 172.16.0.0 255.255.0.0
B(config)# router rip
B(config-router)# version 2
B(config-router)# no auto-summary
  
```

【检验方法】 检查路由器 A 路由表，生成 172.16.0.0/16 表项。

```

A# show ip route rip
R 172.16.0.0/16 [120/2] via 192.168.1.2, 00:01:04, GigabitEthernet 0/1
  
```

常见错误

- 未配置 RIP 基本功能，或配置失败。

3.4.7 启动超网路由

配置效果

- 允许 RIP 在指定接口上发送 RIP 超网路由。

注意事项

- 必须配置 RIP 基本功能。

配置方法

▾ 启用超网络路由

- 当 RIPv1 路由设备监听 RIPv2 路由响应报文时，如果监听到超网路由信息，由于 RIPv1 会将报文中路由信息的子网掩码忽略，从而学习到错误的路由。所以，这种情况下，需要在 RIPv2 路由设备使用该命令的 no 形式禁止在相应接口通告超网路由。命令仅作用于配置了此命令的接口。
- 该命令仅在接口发送 RIPv2 报文时有效，用于控制超网路由的发送。

检验方法

在邻居设备上观察是否学习到超网路由。

相关命令

- 【命令格式】 `ip rip send supernet-routes`
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 缺省允许接口发送 RIP 超网路由

配置举例

▾ 关闭超网络路由

【网络环境】

图 3-14



- 【配置方法】
- 所有路由器配置接口 IP 地址。略
- 所有路由器配置 RIP 基本功能。略。
- 在 B 设备 GigabitEthernet 0/1 接口禁止发送 RIP 的超网路由。

```
B# configure terminal
B(config)# ip route 207.0.0.0 255.0.0.0 Null 0
B(config)# ip route 208.1.1.0 255.255.255.0 Null 0
B(config)# router rip
B(config-router)# redistribute static
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# no ip rip send supernet-routes
```

【检验方法】 检查路由器 A 的路由表，只能学习到非超网路由 208.1.1.0/24，不能学习到超网路由 207.0.0.0/8

```
A#show ip route rip
R    208.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1
```

常见错误

3.4.8 通告缺省路由/外部路由

配置效果

- 在 RIP 路由域内引入其他 AS 的单播路由，为 RIP 路由域内的用户提供到其他 AS 的单播路由服务。
- 往 RIP 路由域内注入一条通往其他 AS 的默认路由，为 RIP 路由域内的用户提供到其他 AS 的单播路由服务。

注意事项

- 必须配置 RIP 基本功能。
- 路由重分布，不能将其他协议的缺省路由引入到 RIP 中。

配置方法

▾ 向邻居通告缺省路由

如果需要向邻居通告缺省路由，则必须启用此功能。

缺省没有指定产生缺省路由；产生的缺省路由的 metric 值缺省为 1。

如果 RIP 进程能够通过本命令生成缺省路由，那么 RIP 将不再学习邻居通告的缺省路由。

若无特殊要求，应在需要通告缺省路由的路由器上配置该功能。

▾ 在接口上向邻居通告缺省路由

如果需要在指定接口向邻居通告缺省路由，则必须启用此功能。

默认没有缺省路由配置。metric 默认为 1。

在某接口配置了该命令后，就会产生一条缺省路由，并从该接口通告此路由。

若无特殊要求，应在需要通告缺省路由的路由器上配置该功能。

▾ 路由重分布，向邻居通告外部路由

如果需要重分布其它协议的路由，则必须启用此功能。

缺省情况下，

- 若配置重分布 OSPF，则重分布该实例的所有子类型的路由；
- 其他情况下，重分布该类型的所有路由；

- 重分布后的路由的 metric 缺省为 1；
- 缺省不关联 route-map。

路由重分布时，将一个路由协议的量度转换成另一种路由协议的量度是没有必要的，因为不同路由协议所采用的量度计算方法是完全不同的。RIP 量度计算是基于跳数，OSPF 是基于带宽，因此它们计算出来的量度是没有可比性的。但是路由重分布时，又必须要设置一个象征性的量度，否则路由重分布将失败。

若无特殊要求，应在需要重分布路由的路由器上配置该功能。

检验方法

- 在邻居设备上观察 RIP 路由表中是否存在缺省路由。
- 在本机和邻居设备上观察 RIP 路由表中是否存在外部路由（到其他 AS 的路由）。

相关命令

向邻居通告缺省路由

【命令格式】 **default-information originate** [**always**] [**metric** *metric-value*] [**route-map** *map-name*]

【参数说明】 **always** 该关键字使 RIP 无条件产生缺省路由，不管本地是否存在缺省路由。
metric *metric-value* 缺省路由由初始的量度值，*metric-value* 取值范围：1-15。
route-map *map-name* 关联的 route-map 的名字，缺省没有关联 route-map

【命令模式】 路由进程模式

【使用指导】 如果路由设备的路由表中存在缺省路由，RIP 默认不会向外通告缺省路由，需要使用 **default-information originate** 路由进程配置命令进行配置，才能向邻居通告缺省路由。

如果选择 **always** 参数，RIP 路由进程不管是否存在缺省路由，都会向邻居通告一条缺省路由，但该缺省路由并不会在本地路由表中显示。要确认是否产生缺省路由，可以用 **show ip rip database** 观察 RIP 路由信息数据库。

如果要对 RIP 通告缺省路由进行更多的控制，可以使用 **route-map** 参数，如：使用 **set metric** 规则对缺省路由的度量值进行设置。

可以使用 **metric** 参数对通告缺省路由的度量值进行设置，但优先级低于 **route-map** 的 **set metric** 规则；如果没有配置 **metric** 参数，则缺省路由使用 RIP 配置的默认度量值。

对于 **ip default-network** 生成的缺省路由，仍然需要使用 **default-information originate** 命令才能够被引入到 RIP。

在接口上向邻居通告缺省路由

【命令格式】 **ip rip default-information** { **only** | **originate** } [**metric** *metric-value*]

【参数说明】 **only** 通告缺省路由，不通告其它路由
originate 通告缺省路由，也通告其它路由
metric *metric-value* 缺省路由的量度值，取值范围：1-15

【命令模式】 接口模式

【使用指导】 如果同时配置了接口下的 **ip rip default-information** 和 RIP 进程的 **default-information originate**，则只通告接口配置的缺省路由。

只要有一个接口配置了 **ip rip default-information**，RIP 就不再学习邻居通告的缺省路由。

路由重分布，向邻居通告外部路由

【命令格式】 **redistribute** { **connected** [*area-tag*] | **ospf** *process-id* | **static** } [{ **level-1** | **level-1-2** | **level-2** }] [**match** { **internal** | **external** [**1** | **2**] | **nssa-external** [**1** | **2**] }] [**metric** *metric-value*] [**route-map** *route-map-name*]

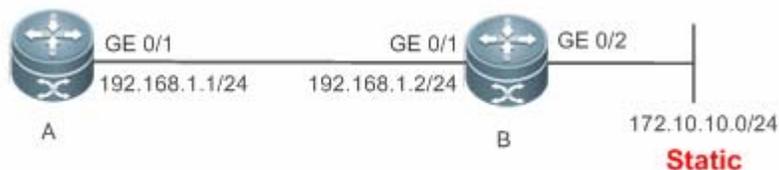
- 【参数说明】 **connected** : 从直连路由重分布
ospf process-id : 从 ospf 重分布, *process-id* 为 ospf 进程号, 取值范围 1-65535
static : 从静态路由重分布
level-1 | level-1-2 | level-2 : 针对重分布 IS-IS 路由, 用于指定特定 level 的路由进行重分布
match : 针对重分布 OSPF 路由, 用于过滤特定路由进行重分布
metric metric-value : 设置重分布的路由的 metric, 取值范围 1-16
route-map route-map-name : 设置重分布过滤规则
- 【命令模式】 路由进程模式
- 【使用指导】 -

配置举例

路由重分布, 向邻居通告外部路由

【网络环境】

图 3-15



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 RIP 基本功能。略。
 - 路由器 B 配置重分布静态路由。

B

```
B# configure terminal
B(config)# router rip
B(config-router)# redistribute static
```

- 【检验方法】 路由器 A 上检查路由表: 加载了 172.10.10.0/24 表项。

```
A# show ip route rip
R    172.10.10.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1
```

常见错误

-

3.4.9 设置路由过滤规则

配置效果

- 被过滤掉的路由不能被加载到路由表或被传播给邻居, 从而阻止网络内的用户访问指定的目的网络。

注意事项

- 必须配置 RIP 基本功能。
- 对于发出的过滤规则, 必须先配置重分布路由, 针对重分布路由实现过滤。

配置方法

对收到的 RIP 路由信息进行过滤

- 如果需要接收的路由信息进行过滤，则必须配置该功能。
- 为了拒绝接收某些指定路由，可以通过配置路由分布控制列表，对所有接收到的路由更新报文进行处理。如果没指定接口，就对所有的接口接收的路由更新报文进行处理。
- 若无特殊要求，应在需要过滤的路由器上配置该功能。

对发出的 RIP 路由信息进行过滤

- 如果需要发送的重分布路由信息进行过滤，则必须配置该功能。
- 如果该命令不跟任何可选参数，路由更新通告控制对所有接口起作用；如果跟接口选项，路由更新通告控制只对指定接口起作用；如果跟其它路由进程参数，则路由更新通告控制只对指定路由进程起作用。
- 若无特殊要求，应在需要过滤的路由器上配置该功能。

检验方法

- 通过 `show ip route rip` 确认被过滤的路由没有加载到路由表

相关命令

对收到的 RIP 路由信息进行过滤

【命令格式】 **distribute-list** { [*access-list-number* | *name*] | **prefix** *prefix-list-name* [**gateway** *prefix-list-name*] } **in** [*interface-type* *interface-number*]

【参数说明】 *access-list-number* | *name*：指定访问列表。只有访问列表允许的路由，才可以被接收。
prefix *prefix-list-name*：使用前缀列表来过滤路由。
gateway *prefix-list-name*：使用前缀列表来过滤路由的源。
interface-type *interface-number*：该分布列表，只应用在指定接口。

【命令模式】 路由进程模式

【使用指导】 -

对发出的 RIP 路由信息进行过滤

【命令格式】 **distribute-list** { [*access-list-number* | *name*] | **prefix** *prefix-list-name* } **out** [*interface* | [**connected** [*area-tag*]] | **ospf** *process-id* | **rip** | **static**]]

【参数说明】 *access-list-number* | *name*：指定访问列表。只有访问列表允许的路由，才可以被发送。
prefix *prefix-list-name*：使用前缀列表来过滤路由。
Interface：路由更新通告控制只应用在指定接口。
connected：路由更新通告控制只应用在从直连引入的路由。
ospf *process-id*：路由更新通告控制只应用在从 ospf 引入的路由，*process-id* 指定具体某个 ospf 实例。
rip：路由更新通告控制只应用在 rip 路由。
static：路由更新通告控制只应用在从静态引入的路由。

【命令模式】 路由进程模式

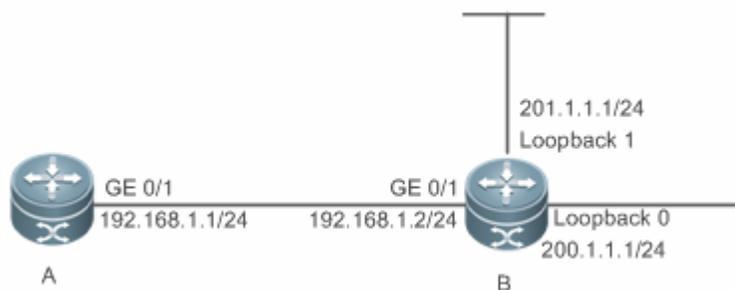
【使用指导】 -

配置举例

对收到的 RIP 路由信息进行过滤

【网络环境】

图 3-16



【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 RIP 基本功能。略。
- RIP 对从 GigabitEthernet 0/1 端口接收的路由进行控制处理，只允许接收 200.1.1.0 的路由。

A

```
A# configure terminal
A(config)# router rip
A(config-router)# distribute-list 10 in GigabitEthernet 0/1
A(config-router)# no auto-summary
A(config)# access-list 10 permit 200.1.1.0 0.0.0.255
```

【检验方法】

检查路由器 A 的路由表，只有 200.1.1.0/24 表项。

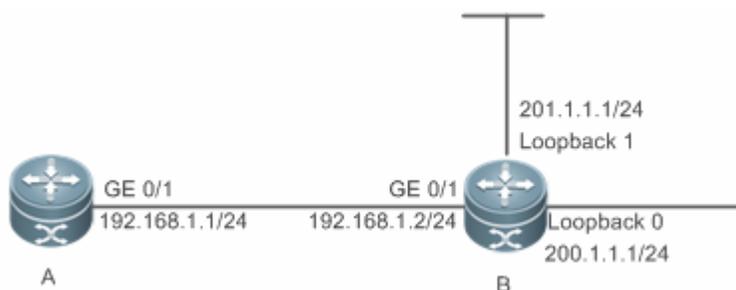
A

```
A# show ip route rip
R    200.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1
```

对发出的 RIP 路由信息进行过滤

【网络环境】

图 3-17



【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 RIP 基本功能。略。
- RIP 路由进程只对外通告 200.1.1.0/24 路由。

B

```
B# configure terminal
B(config)# router rip
B(config-router)# redistribute connected
B(config-router)# distribute-list 10 out
B(config-router)# version 2
B(config)# access-list 10 permit 200.1.1.0 0.0.0.255
```

【检验方法】 查看 A 路由表，只有 200.1.1.0 网段的路由。

```
A
A# show ip route rip
R    200.1.1.0/24 [120/1] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1
```

常见配置错误

- 列表的过滤规则错误，导致过滤失败

3.4.10 修改路由选择参数

配置效果

- 改变 RIP 路由，使流量经过或避开某些节点。
- 改变路由器选择各类路由的顺序，使 RIP 路由的优先级提高或降低。

注意事项

- 必须配置 RIP 基本功能。

配置方法

修改 RIP 路由的管理距离

- 可选配置。
- 在同时运行多种单播路由协议的路由器上，若希望改变 RIP 路由的优先级，必须配置。

增加接受或发送的 RIP 路由的 metric

- 可选配置。
- 若无特殊要求，应在需要调整路由 metric 的路由器上配置该功能。

配置重分布到 RIP 中的路由的缺省 metric

- 可选配置。
- 若无特殊要求，应在引入域外路由的 AS 边界路由器上配置。

检验方法

使用 show ip rip 查看当前所配置的 distance 值。可以使用 show ip rip data 查看重分发进来的路由的度量，确定配置是否生效。

相关命令

修改 RIP 路由的管理距离

【命令格式】 **distance** *distance* [*ip-address wildcard*]

【参数说明】 *distance*：设置 RIP 路由管理距离，1-255 之间的整数

ip-address：路由来源 IP 地址前缀

wildcard：定义 IP 地址比较比特位，0 表示精确匹配，1 表示不做比较。

【命令模式】 路由进程模式

【使用指导】 使用该命令设置 RIP 路由的管理距离。

增加接受或发送的 RIP 路由的 metric

【命令格式】 **offset-list** { *access-list-number* | *name* } { *in* | *out* } *offset* [*interface-type interface-number*]

【参数说明】 *access-list-number* | *name*：指定访问列表

In : 应用 acl 对接收的路由的 metric 进行修改

out : 应用 acl 对发送的路由的 metric 进行修改

offset : 修改的 metric 的变动值, 取值范围: 0-16

interface-type : 应用 acl 到指定接口

interface-number : 指定接口编号

【命令模式】 路由进程模式

【使用指导】 增加接收或发送的 RIP 路由的 metric 值, 若指定接口则针对对应接口, 若未指定则全局成立。

配置重分布到 RIP 中的路由的缺省 metric

【命令格式】 **default-metric metric-value**

【参数说明】 **metric-value** : 缺省的量度值, 有效值为 1 到 16。如果 **metric-value** 值大于等于 16, RGOS 将认为该路由不可达。

【命令模式】 路由进程模式

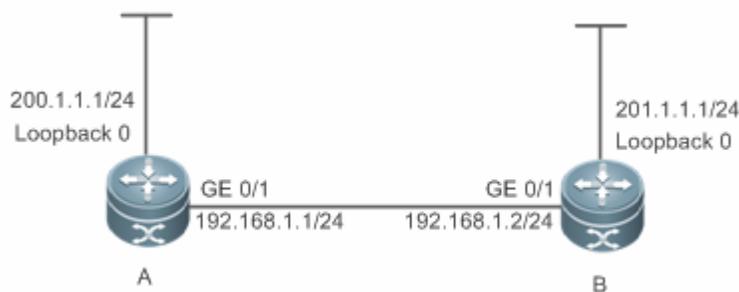
【使用指导】 该命令需要与路由协议配置命令 **redistribute** 配合使用。

配置举例

增加接受或发送的 RIP 路由的 metric

【网络环境】

图 3-18



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 RIP 基本功能。略。
 - 在 acl 7 指定的范围内的 RIP 路由才会对 metric 增加 7。
 - 对学习的在 acl 8 指定的范围内的 RIP 路由才会对 metric 增加 7。

```
A
A# configure terminal
A(config)# access-list 7 permit host 200.1.1.0
A(config)# access-list 8 permit host 201.1.1.0
A(config)# router rip
A(config-router)# offset-list 7 out 7
A(config-router)# offset-list 8 in 7
```

【检验方法】 检查路由器 A、B 的路由表, RIP 路由的 metric 为 8。

```
A
A# show ip route rip
R   201.1.1.0/24 [120/8] via 192.168.1.2, 00:06:11, GigabitEthernet 0/1

B
B# show ip route rip
R   200.1.1.0/24 [120/8] via 192.168.1.1, 00:06:11, GigabitEthernet 0/1
```

常见错误

3.4.11 修改时钟

配置效果

- 改变控制协议运行的定时器，加快或延缓协议状态的改变，加快或延缓事件的发生。

注意事项

- 必须配置 RIP 基本功能。
- 改变协议控制参数，有可能导致协议运行失败。不建议执行此项配置。

配置方法

✎ 修改 RIP 路由更新、失效、清除的时间

如果需要调整 RIP 时钟，则必须配置该功能。

通过调整时钟，可能会加快路由协议的收敛时间以及故障恢复时间。连接在同一网络上的设备，RIP 时钟值一定要一致。一般不建议对 RIP 时钟进行调整，除非有明确的需求。

在低速链路上将时钟设得很短，将会带来一定风险，因为大量的更新报文会消耗带宽。一般只有在以太网或者 2Mbps 以上的线路，可以将时钟设置得短一点，这样可以缩短网络路由的收敛时间。

若无特殊要求，应在需要修改 RIP 时钟的路由器上配置该功能。

✎ 设置 RIP 路由更新报文的发送延时

如果需要调整 RIP 更新报文的发送延时，则必须配置该功能。

使用 `output-delay` 命令增加高速设备的报文间的发送延时，使低速设备能够接收处理全部更新报文。

若无特殊要求，应在需要调整发送延时的路由器上配置该功能。

检验方法

可以用 `show ip rip` 命令进行查看当前的 RIP 时钟参数设置。

相关命令

✎ 修改 RIP 路由更新、失效、清除的时间

【命令格式】 `timers basic update invalid flush`

【参数说明】 `update`：路由更新时间，以秒计。`update` 定义了设备发送路由更新报文的周期，每接收到更新报文，`invalid` 和 `flush` 时钟就复位。缺省每隔 30 秒发送一次路由更新报文。

`invalid`：路由无效时间，以秒计，从最近一次有效更新报文开始计时。`invalid` 定义了路由表中路由因没有更新而变为无效的时间。路由无效时间至少应该为路由更新时间的三倍，如果在路由无效时间内没有接收到任何更新报文，相应的路由将变为无效，进入 `invalid` 状态；`invalid` 时间内接收到路由更新报文，时钟会复位。`Invalid` 缺省时间为 180 秒。

`flush`：路由清除时间，以秒计，从 RIP 路由进入 `invalid` 状态开始计时。`flush` 时间到期，处于 `invalid` 状态路由将被清除出路由表。`Flush` 缺省时间为 120 秒。

【命令模式】 路由进程模式

【使用指导】 缺省情况下，更新时间为 30 秒，无效时间为 180 秒，清除时间为 120 秒。

设置 RIP 路由更新报文的发送延时

【命令格式】 **output-delay delay**

【参数说明】 *delay*：设置报文间发送延时，单位为毫秒，范围是<8-50>。

【命令模式】 接口模式

【使用指导】 正常情况下，RIP 更新报文大小为 512 字节，可以包含 25 条路由，如果更新的路由条目大于 25，就会通过多个路由更新报文发送，正常情况下是尽可能快的发送。

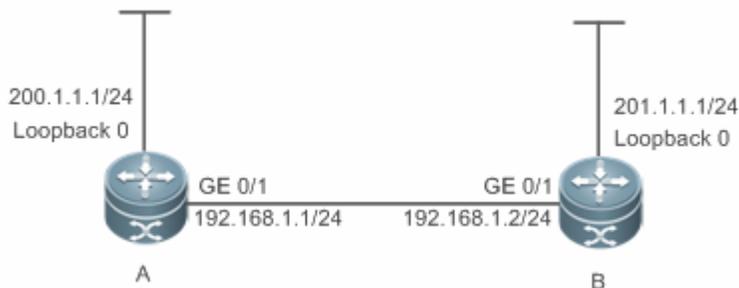
但是，当高速设备向低速设备发送大量更新报文时，低速设备可能来不及处理全部的更新报文，造成路由信息丢失。此时，需要使用 **output-delay** 命令增加高速设备的报文间的发送延时，使低速设备能够接收处理全部更新报文。

配置举例

设置 RIP 路由更新报文的发送延时

【网络环境】

图 3-19



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 RIP 基本功能。略。
 - 路由器 A 配置更新报文的延时时间。

```
A
A# configure terminal
A(config)# router rip
A(config-router)# output-delay 30
```

【检验方法】 抓包检查路由器 A 的更新报文发送时间，与原来相比延时了 30 毫秒。

常见错误

连接在同一网络上的设备，RIP 时钟值设置不一致。

3.4.12 启动与 BFD 联动

配置效果

- 一旦发生链路故障，RIP 可以快速发现路由失效。此项配置有助于缩短流量中断时间。

注意事项

- 必须配置 RIP 基本功能。

- 接口上配置的联动 BFD 优先级要高于全局配置。

配置方法

全局启动 RIP 与 BFD 联动

- 如果需要启动联动 BFD，则必须配置该功能。
- 在 RIP 上启动 BFD 检测功能后，将会为 RIP 路由信息源(RIP 路由更新报文的源地址)建立 BFD 会话，一旦 BFD 邻居失效，RIP 路由信息将直接进入失效状态，不再参与路由转发。
- 您也可以通过接口配置模式命令 `ip rip bfd [disable]` 启动或者关闭指定接口的 BFD 检测功能，接口配置的优先级高于进程配置模式下的 `bfd all-interfaces` 命令。
- 若无特殊要求，应在每台路由器上配置该功能。

在接口上启动 RIP 与 BFD 联动

- 如果需要针对部分接口启用或禁用联动 BFD，则必须配置该功能。
- 接口配置的优先级高于进程配置模式下的 `bfd all-interfaces` 命令。
- 用户可以根据实际环境，通过 `ip rip bfd` 命令选择指定的接口使用 BFD 进行链路检测，也可以使用 RIP 进程配置模式下的命令 `bfd all-interfaces` 配置参与 RIP 运行的所有接口通过 BFD 进行链路检测，使用 `ip rip bfd disable` 选择关闭指定接口的 BFD 检测功能。
- 若无特殊要求，应在需要单独配置 BFD 联动的路由器接口上设置该功能。

检验方法

- RIP 联动的 BFD 会话正确建立。
- 链路失效后，RIP 路由能快速收敛。

相关命令

全局启动 RIP 与 BFD 联动

- 【命令格式】 `bfd all-interfaces`
- 【参数说明】 -
- 【命令模式】 路由进程模式
- 【使用指导】 -

在接口上启动 RIP 与 BFD 联动

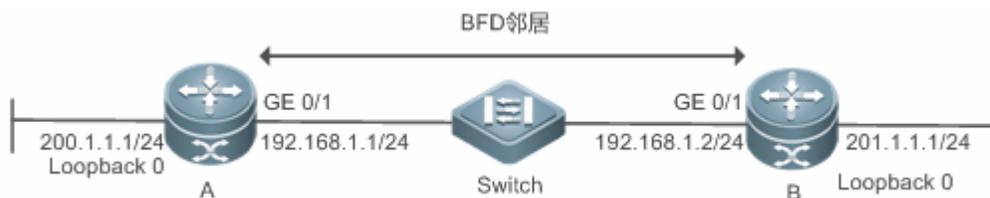
- 【命令格式】 `ip rip bfd [disable]`
- 【参数说明】 `disable`：关闭运行 RIP 的指定接口使用 BFD 机制进行链路检测。
- 【命令模式】 接口模式
- 【使用指导】 缺省没有配置，以 RIP 进程配置模式下的 BFD 配置为准。

配置举例

启动 RIP 与 BFD 联动

【网络环境】

图 3-20



【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 RIP 基本功能。略。
- 所有路由器配置接口 BFD 参数。
- 所有路由器配置 RIP 联动 BFD。

A

```
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# bfd interval 200 min_rx 200 multiplier 5
A(config)# router rip
A(config-router)# bfd all-interfaces
```

B

```
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# bfd interval 200 min_rx 200 multiplier 5
B(config)# router rip
B(config-router)# bfd all-interfaces
```

【检验方法】

- 路由器 A、B 上查看 RIP 联动 BFD 处于 UP 状态。
- 断开路由器 B 与 Switch 之间的链路，路由器 A 上的 RIP 路由删除。

A

```
A# show ip rip peer
Peer 192.168.1.2:
  Local address: 192.168.1.1
  Input interface: GigabitEthernet 0/1
  Peer version: RIPv2
  Received bad packets: 0
  Received bad routes: 0
  BFD session state up
```

B

```
A# show ip rip peer
Peer 192.168.1.1:
  Local address: 192.168.1.2
  Input interface: GigabitEthernet 0/1
  Peer version: RIPv2
  Received bad packets: 0
  Received bad routes: 0
  BFD session state up
```

常见错误

- 未开启相对应的 BFD 功能，就在 RIP 中启用以上两项命令。

3.4.13 启动快速重路由

配置效果

- 一旦 RIP 发现路由失效，路由可以立即切换到次优路径。此项配置有助于缩短流量中断时间。

注意事项

- 必须配置 RIP 基本功能。
- 必须配置路由图，并设置备份下一跳规则
- 如果要加速收敛，请将接口的 carrier-delay 配置为 0，并启用“RIP 与 BFD 联动”

配置方法

启动快速重路由，引用路由图

如果需要启用快速重路由，则需要配置该功能

配置 route-map，可以通过路由图来匹配成功的路由指定备份路径。

使用 RIP 快速重路由功能建议同时使能 BFD 链路检测功能，设备能够更快检测到链路故障，从而降低转发中断时间。针对接口 up/down 的场景，为了降低 RIP 快速重路由的转发中断时间，可以在接口模式下配置 carrier-delay 0，以达到最快的切换速度。

若无特殊要求，应在每台路由器上启用该功能。

检验方法

- 能够正确计算并生成备份路由
- 当主链路失效，数据转发能够迅速切换到备份链路

相关命令

启动快速重路由，引用路由图

【命令格式】 **fast-reroute route-map** *route-map-name*

【参数说明】 *route-map-name*：通过路由图指定备份路径

【命令模式】 路由进程模式

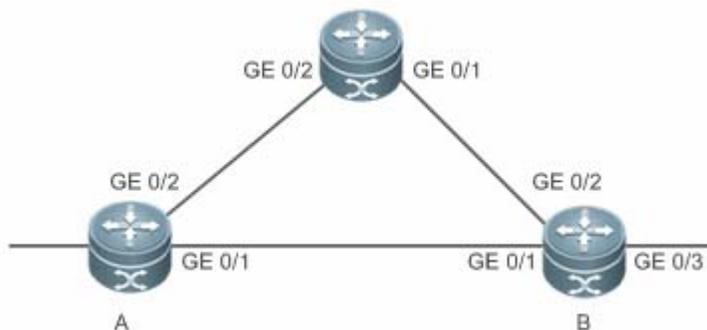
【使用指导】 目前 RIP 快速重路由功能存在以下限制：A、每条路由仅能生成一条备份下一跳；B、不能为 ECMP 生成备份下一跳；

配置举例

启动快速重路由，引用路由图

【网络环境】

图 3-21



【注释】 接口 IP 地址：

A : GE0/1 192.168.1.1 GE0/2 192.168.2.1]

B : GE0/1 192.168.1.2 GE0/2 192.168.3.1 GE0/3 192.168.4.1

C : GE0/1 192.168.3.2 GE 0/2 192.168.2.2

【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 RIP 基本功能。略。
- 路由器 A 配置快速重路由。
- 路由器 A 配置接口 carrier-delay 0。

A

```
A# configure terminal
A(config)# route-map fast-reroute
A(config-route-map)# match interface GigabitEthernet 0/2
A(config-route-map)# set fast-reroute backup-interface GigabitEthernet 0/1 backup-nextthop
192.168.1.1
A(config)# router rip
A(config-router)# fast-reroute route-map fast-reroute
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# carrier-delay 0
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)# carrier-delay 0
```

【检验方法】

检查路由器 A 路由表，192.168.4.0/24 表项存在备份路由。

A

```
A# show ip route fast-reroute | begin 192.168.4.0
R 192.168.4.0/24 [ma] via 192.168.1.2, 00:39:28, GigabitEthernet 0/1
[b] via 192.168.2.2, 00:39:28, GigabitEthernet 0/2
```

常见错误

- 路由图未正确配置备份下一跳规则。
- 未配置接口 carrier-delay 或联动 BFD，导致数据转发收敛较慢。

3.5 监视与维护

查看运行情况

作用	命令
查看 RIP 进程的基本信息	show ip rip
查看 RIP 路由表	show ip rip database [<i>network-number network-mask</i>] [count]
查看 RIP 重分布的外部路由信息	show ip rip external [connected [<i>process-id</i>] ospf process-id static]
查看 RIP 接口信息	show ip rip interface [<i>interface-type interface-number</i>]
查看 RIP 邻居信息	show ip rip peer [<i>ip-address</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
查看 RIP 进程运行过程中发生的事件	debug ip rip event
查看与 NSM 进程的交互信息	debug ip rip nsm
查看 RIP 的收发包信息	debug ip rip packet [<i>interface interface-type interface-number</i> recv send]
查看 RIP 进程路由变化信息	debug ip rip route

4 OSPFv2

4.1 概述

OSPF (Open Shortest Path First , 开放式最短路径优先协议) 是一种内部网关协议 (Interior Gateway Protocol , 简称 IGP) , 应用在 AS (Autonomous System , 自治系统) 内部 , 使路由器获得远端网络的路由。

OSPF 具备如下特点 :

- 适用范围广 : 适用于更大规模的网络 (最大可支持几百台路由器) 。
- 快速收敛 : 一旦网络拓扑发生变化 , 路由器之间能够快速通告信息 , 更新路由。
- 无自环 : 路由器之间仅同步链路状态信息 , 每台路由器独立计算路由 , 不会生成自环。
- 区域划分 : 将一个大的路由域划分为多个较小的区域 , 可以节省系统资源和网络带宽 , 使路由稳定可靠。
- 路由分类 : 将路由分为多种类型区别对待 , 支持灵活的控制策略。
- 等价路由 : 支持等价路由。
- 支持认证 : 支持报文认证 , 保障协议交互过程的安全。
- 组播发送 : 支持以组播地址发送协议报文 , 可以避免干扰无关者、节省系统资源。

i 本章使用“路由器”指代支持路由功能的网络设备。这些网络设备可以是三层交换机、路由器、防火墙等。

i 若无特殊说明,下文以“OSPF”来指代“OSPFVersion 2”。

协议规范

- RFC2328** This memo documents version 2 of the OSPF protocol. OSPF is a link-state routing protocol.
- RFC 2370** This memo defines enhancements to the OSPF protocol to support a new class of link-state advertisements (LSA) called Opaque LSAs. Opaque LSAs provide a generalized mechanism to allow for the future extensibility of OSPF.
- RFC3137** This memo describes a backward-compatible technique that may be used by OSPF (Open Shortest Path First) implementations to advertise unavailability to forward transit traffic or to lower the preference level for the paths through such a router.
- RFC3623** This memo documents an enhancement to the OSPF routing protocol, whereby an OSPF router can stay on the forwarding path even as its OSPF software is restarted.
- RFC3630** This document describes extensions to the OSPF protocol version 2 to support intra-area Traffic Engineering (TE), using Opaque Link State Advertisements.
- RFC3682** The use of a packet's Time to Live (TTL) (IPv4) or Hop Limit (IPv6) to protect a protocol stack from CPU-utilization based attacks has been proposed in many settings.

- RFC3906** This document describes how conventional hop-by-hop link-state routing protocols interact with new Traffic Engineering capabilities to create Interior Gateway Protocol (IGP) shortcuts.
- RFC4576** This document specifies the necessary procedure, using one of the options bits in the LSA (Link State Advertisements) to indicate that an LSA has already been forwarded by a PE and should be ignored by any other PEs that see it.
- RFC4577** This document extends that specification by allowing the routing protocol on the PE/CE interface to be the Open Shortest Path First (OSPF) protocol.
- RFC4750** This memo defines a portion of the Management Information Base (MIB) for use with network management protocols in TCP/IP-based internets. In particular, it defines objects for managing version 2 of the Open Shortest Path First Routing Protocol. Version 2 of the OSPF protocol is specific to the IPv4 address family.

4.2 典型应用

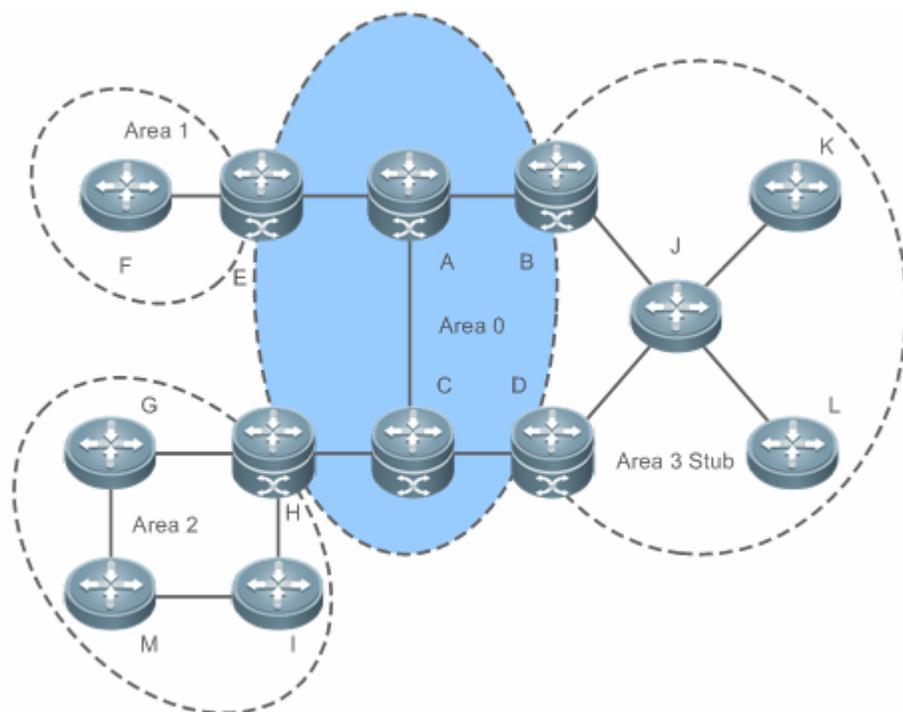
典型应用	场景描述
域内互联	在 AS 内部运行 OSPF，划分多个区域。

4.2.1 域内互联

应用场景

在 AS 内部运行 OSPF，如果路由设备超过 40 台，建议划分多个区域。通常，骨干区域应该采用性能可靠且处理速度较快的高端设备，非骨干区域可以采用性能相对较低的中低端设备。所有非骨干区域都必须与骨干区域相连。建议将位于末梢的非骨干区域设置为 Stub 区域。以图 1-1 为例，将该网络划分成 4 个区域，不同区域间的通信都经过骨干区域（即区域 0）。

图 4-10 OSPF 区域划分



【注释】 A、B、C、D、E、H 处于骨干区域，是骨干路由器。
区域 3 配置成 Stub 区域。

功能部属

- AS 内部，所有路由器运行 OSPF 协议，实现单播路由。

4.3 功能详解

基本概念

↳ 路由域

一个 AS 中的所有路由器必须相互连接，运行相同的路由协议。因此，AS 也被称为路由域（Routing Domain）。运行 OSPF 的 AS 也被称为 OSPF 路由域，简称 OSPF 域。

↳ OSPF 进程

OSPF 支持多实例（Instance），每个实例对应一个 OSPF 进程（Process）。

一台路由器上可以启用一个 OSPF 进程，也可以启用多个 OSPF 进程。每个 OSPF 进程独立执行 OSPF 协议，相互隔离。Process ID 只有本地意义，不影响相邻接口上的 OSPF 报文交互。

RouterID

OSPF 域内，使用 Router ID 唯一标识一台路由器。任意两台路由器的 Router ID 不能相同。

如果路由器上存在多个 OSPF 进程，则每个 OSPF 进程使用一个 Router ID。任意两个 OSPF 进程的 Router ID 不能相同。

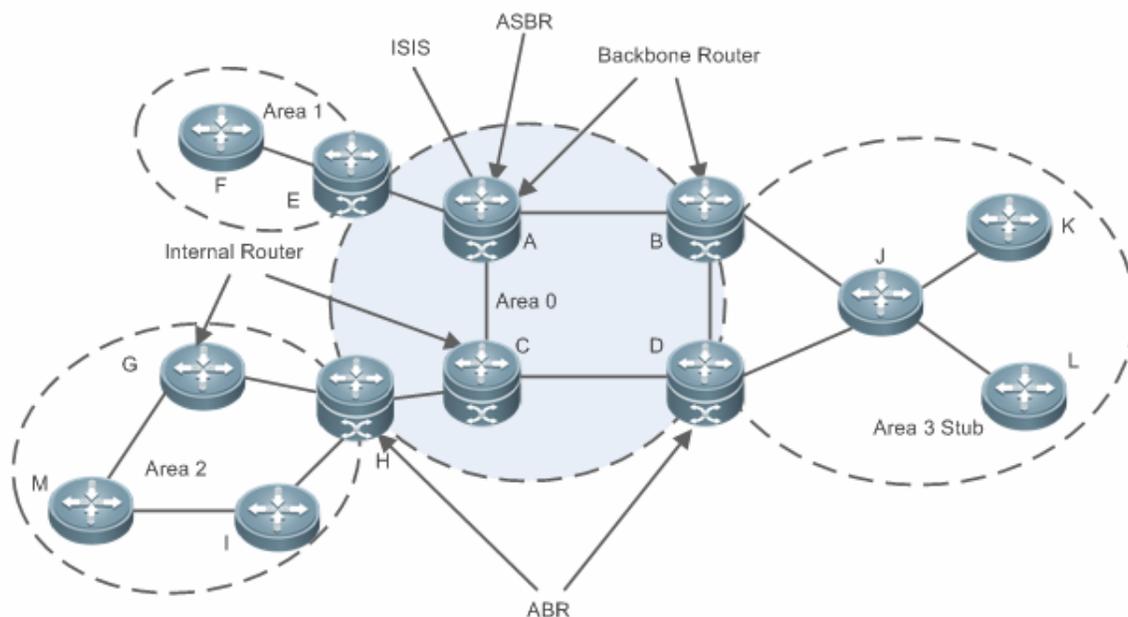
区域

OSPF 支持多区域（Area）。将一个 OSPF 域划分为多个区域，有助于缓解大规模网络的计算压力。

区域是从逻辑上将路由器划分为不同的组，每个组用区域号（Area ID）来标识。区域的边界是路由器，一台路由器可以只属于一个区域，也可以属于多个区域。一个网段（链路）只能属于一个区域，或者说每个运行 OSPF 的接口必须指明属于哪一个区域。

区域 0 为骨干区域（Backbone Area），其他区域为常规区域（Normal Area）。常规区域必须与骨干区域直接相连。

图 4-3 OSPF 区域划分



OSPF 路由器

在 OSPF 中定义了几种路由器类型，并为每种类型的路由器赋予不同的职责。

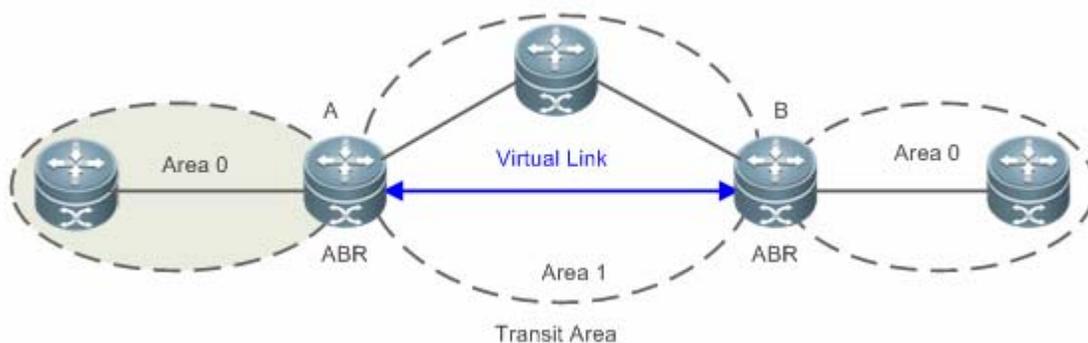
- 区域内路由器（Internal Router）
该类路由器的所有接口都属于同一个 OSPF 区域。如上图中的 A、C、F、G、I、M、J、K、L。
- 区域边界路由器（Area Border Router，简称 ABR）
该类路由器用来连接骨干区域和常规区域。ABR 同时属于两个以上的区域，其中一个必须是骨干区域。如上图中的 B、D、E、H。

- 骨干路由器（Backbone Router）
该类路由器至少有一个接口属于骨干区域。所有的 ABR 和 Area0 的区域内路由器都是骨干路由器。如上图中的 A、B、C、D、E、H。
- 自治系统边界路由器（AS Boundary Router，简称 ASBR）
该类路由器用来与其他 AS 交换路由信息。ASBR 并不一定位于 AS 的边界，它可能是区域内路由器，也可能是 ABR。如上图中的 A。

虚链路

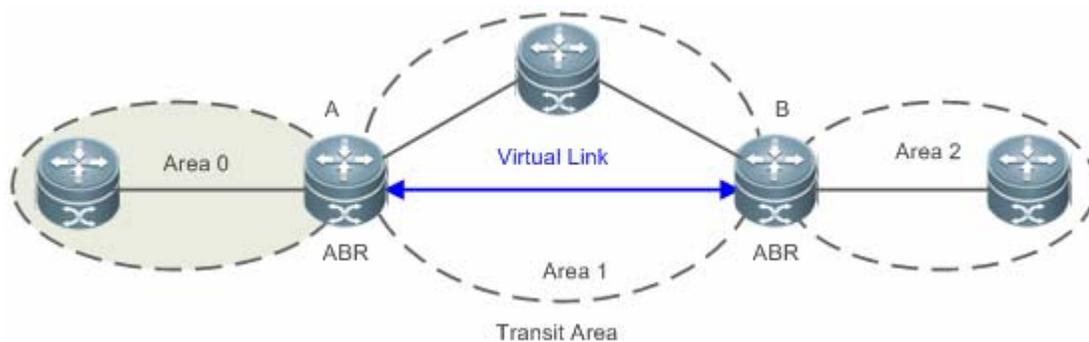
OSPF 支持虚链路（Virtual link）。虚链路是一条逻辑链路，属于骨干区域，用于解决在物理网络上骨干区域不连续或骨干区域与常规区域不能直接相连的问题。虚链路只能穿越一个常规区域，该区域被称为“传输区域”（Transit Area）。虚链路两端的路由器均为 ABR。

图 4-4 在物理网络上，骨干区域不连续



如上图所示：在 A、B 之间建立虚链路，将 Area0 连接起来；Area1 是传输区域；A、B 是 Area1 的 ABR。

图 4-5 在物理网络上，骨干区域与常规区域不能直接相连



如上图所示：在 A、B 之间建立虚链路，将 Area0 延伸到 B，则 Area0 与 Area2 在 B 上直接相连；Area1 是传输区域；A 是 Area1 的 ABR；B 是 Area1、Area2 的 ABR。

LSA

OSPF 以 LSA（Link State Advertisement，链路状态通告）的形式来描述路由信息。

LSA 类型	说明
Router-LSA (Type1)	由每个路由器始发，描述了路由器的链路状态和代价，在始发路由器所在区域内传播。

Network-LSA (Type2)	由 NBMA 类型网络上的 DR 始发, 描述本网段的链路状态, 在 DR 所在区域内传播。
Network-summary-LSA (Type3)	由 ABR 始发, 描述到其它区域的路由, 通告给非 Totally STUB、NSSA 区域。
ASBR-summary-LSA (Type4)	由 ABR 始发, 描述到 ASBR 的路由, 通告给除 ASBR 所在区域的其他区域。
AS-external-LSA (Type5)	由 ASBR 始发, 描述到 AS 外部的路由, 通告到所有区域 (除了 STUB、NSSA 区域) 。
NSSA LSA (Type7)	由 ASBR 始发, 描述到 AS 外部的路由, 仅在 NSSA 区域内传播。
Opaque LSA (Type9/Type10/Type11)	Opaque LSA 提供用于 OSPF 的扩展的通用机制。其中： <ul style="list-style-type: none"> ● Type9 LSA 仅在接口所在网段范围内传播。用于支持 GR 的 Grace LSA 就是 Type9 LSA 的一种。 ● Type10 LSA 在区域内传播。用于支持 TE 的 LSA 就是 Type10 LSA 的一种。 ● Type11 LSA 在自治域内传播, 目前还没有实际应用的例子。

i STUB 区域、NSSA 区域、Totally STUB 区域、Totally NSSA 区域, 是常规区域的几种特殊形态, 有助于减轻路由器负担, 增强 OSPF 路由稳定性。

OSPF 报文

OSPF 使用以下几种类型的协议报文。这些 OSPF 报文被封装在 IP 报文中, 以组播或单播方式发送。

报文类型	说明
Hello	周期性发送, 用来发现 OSPF 邻居, 维持邻接关系。
DD (Database Description)	描述本地 LSDB (Link-State Database, 链路状态数据库) 的摘要信息, 用于在 OSPF 邻居之间进行数据库同步。
LSR (Link State Request)	用于向对方请求所需的 LSA。只有在 OSPF 邻居双方成功交换 DD 报文后, 才会向对方发出 LSR 报文。
LSU (Link State Update)	用于向对方发送其所需要的 LSA。
LSAck (Link State Acknowledgment)	用来对收到的 LSA 进行确认。

功能特性

功能特性	作用
链路状态路由协议	在路由器上运行 OSPF, 即可获得到达网络各个位置的路由。
OSPF 路由管理	通过人为配置, 合理规划或优化 OSPF 路由, 实现管理的目的。
增强安全性、可靠性	认证、与 BFD 联动等功能使 OSPF 更安全、更稳定可靠。
支持网管功能	MIB、SYSLOG 等功能使 OSPF 更易于管理。

4.3.1 链路状态路由协议

OSPF 属于链路状态路由协议 (Link-State Routing Protocols), 工作过程概述：

- 邻居发现→双向通信
在相连的路由器之间建立 OSPF 邻居关系, 并保持双向通信。
- 数据库同步→完全邻接
路由器使用 LSA 发布自己的全部链路状态。邻居之间交换 LSA, 同步 LSDB, 达到完全邻接状态。

- 计算 SPT →形成路由表
路由器依据 LSDB 计算出到达每个目标网络的最短路径，形成 OSPF 路由表。

工作原理

↘ 邻居发现→双向通信

路由器从所有启动 OSPF 的接口上（或虚链路）发出 Hello 报文。如果两台路由器之间能够交互 Hello 报文，并且能够成功协商 Hello 报文中携带的参数，则它们就成为了邻居（Neighbor）。当互为邻居的路由器在彼此的 Hello 报文中找到自己的 Router ID 时，双向通信就建立了。

Hello 报文中包含（但不限于）如下信息：

- 始发路由器的 Router ID
- 始发路由器接口（或虚链路）的 Area ID
- 始发路由器接口（或虚链路）的网络掩码
- 始发路由器接口（或虚链路）的认证信息
- 始发路由器接口（或虚链路）的 Hello 时间间隔
- 始发路由器接口（或虚链路）的邻居死亡时间
- 始发路由器接口的优先级（用于 DR/BDR 竞选）
- DR（Designated Router，指定路由器），BDR（Backup Designated Router，备份指定路由器）的 IP 地址
- 始发路由器的邻居的 Router ID

↘ 数据库同步→完全邻接

邻居路由器之间建立双向通信后，使用 DD、LSR、LSU、LSAck 报文交换 LSA，建立邻接关系。简要过程如下：

- 路由器产生 LSA，用来描述该设备上的全部链路状态。
- 邻居之间交换 LSA。当路由器接收到邻居发送的 LSA 时，将拷贝一份到本地 LSDB，然后再传播给其它邻居。
- 当路由器及其邻居都获得了相同的 LSDB 时，达到完全邻接状态。

i 如果没有链路代价、网络增删变化，OSPF 将会十分安静。如果发生变化，则通告变化的链路状态，快速同步 LSDB。

↘ 计算最短路径→形成路由表

路由器上获得完整的 LSDB 后，将运行 Dijkstra 算法生成一棵从本地到每个目标网络的最短路径树（Shortest Path Tree，简称 SPT）。SPT 中记录了：目标网络、下一跳地址、代价。OSPF 依据 SPT 形成路由表。

如果发生链路代价、网络增删变化，则会引起 LSDB 的更新。设备将再次运行 Dijkstra 算法，生成新的 SPT，并更新路由表。

i Dijkstra 算法用来在一个有权重的有向图中，找到从一个顶点到其他顶点的最短路径。

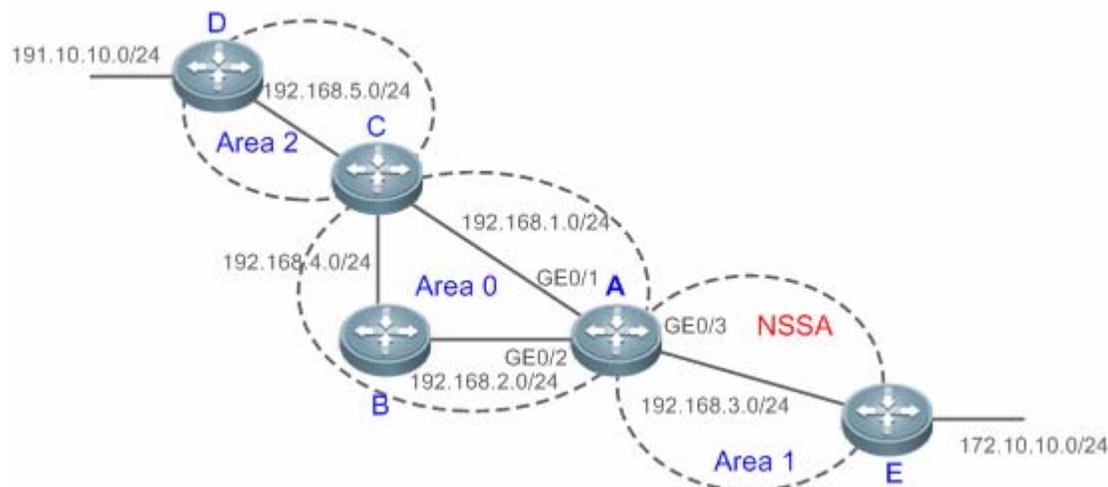
↘ OSPF 网络类型

路由器不一定需要与它的每一个邻居交换 LSA、建立邻接关系。为提高效率，OSPF 将各种链路层协议的网络归类为 5 种类型，以不同的方式交换 LSA、建立邻接关系。

- 广播 (Broadcast)
 - 自动发现邻居，选举 DR、BDR。
 - DR (或 BDR) 与其他所有路由器交换 LSA、建立邻接关系。DR、BDR 之外的路由器之间不交换 LSA、不建立邻接关系。
 - Ethernet、FDDI 默认属于此类型。
- 非广播 (Non-broadcast multiple access, 简称 NBMA)
 - 手工配置邻居，选举 DR、BDR。
 - DR (或 BDR) 与其他所有路由器交换 LSA、建立邻接关系。DR、BDR 之外的路由器之间不交换 LSA、不建立邻接关系。
 - X.25、帧中继、ATM 默认属于此类型。
- 点到点 (point-to-point, 简称 P2P)
 - 自动发现邻居，不选举 DR/BDR。
 - 链路两端路由器之间交换 LSA、建立邻接关系。
 - PPP、HDLC、LAPB 默认属于此类型。
- 点到多点 (Point-to-Multipoint, 简称 P2MP)
 - 自动发现邻居，不选举 DR/BDR。
 - 任意两台路由器之间都要交换 LSA、建立邻接关系。
 - 没有任何链路层协议的网络默认为此类型。
- 点到多点非广播
 - 手动配置邻居，不选举 DR/BDR。
 - 任意两台路由器之间都要交换 LSA、建立邻接关系。
 - 没有任何链路层协议的网络默认为此类型。

OSPF 路由类型

图 4-6



在 A 上查看路由表中的 OSPF 路由 (如下红色字体)。

```
A#show ip route

Codes: C - connected, S - static, R - RIP, B - BGP
O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default

Gateway of last resort is no set

O N2 172.10.10.0/24 [110/20] via 192.168.3.2, 00:01:00, GigabitEthernet 0/3
O E2 191.10.10.0/24 [110/20] via 192.168.1.2, 01:11:26, GigabitEthernet 0/1
C 192.168.1.0/24 is directly connected, GigabitEthernet 0/1
C 192.168.1.1/32 is local host.
C 192.168.2.0/24 is directly connected, GigabitEthernet 0/2
C 192.168.2.1/32 is local host.
C 192.168.3.0/24 is directly connected, GigabitEthernet 0/3
C 192.168.3.1/32 is local host.
O 192.168.4.0/24 [110/2] via 192.168.2.2, 00:00:02, GigabitEthernet 0/2
O IA 192.168.5.0/24 [110/3] via 192.168.1.2, 00:01:02, GigabitEthernet 0/1
```

每条 OSPF 路由前都带有一个标识，用来说明此路由的类型。OSPF 路由共有 6 种类型。

- O：区域内路由
此类路由描述了如何到达本区域内的目标网络。路由代价=本地到目标网络的代价。
- IA：区域间路由
此类路由描述了如何到达其他区域的目标网络。路由代价=本地到目标网络的代价。
- E1：第一类外部路由
此类路由描述了如何到达 AS 以外的目标网络。路由代价=本地到 ASBR 的代价+ASBR 到目标网络的代价。Stub、NSSA 区域内的路由器上不存在此类路由。
- E2：第二类外部路由
此类路由描述了如何到达 AS 以外的目标网络。路由代价=ASBR 到目标网络的代价。Stub、NSSA 区域内的路由器上不存在此类路由。
- N1：NSSA 区域的第一类外部路由
此类路由仅描述了如何经 NSSA 区域的 ASBR 到达 AS 以外的目标网络。路由代价=本地到 ASBR 的代价+ASBR 到目标网络的代价。仅 NSSA 区域的路由器上存在此类路由。
- N2：NSSA 区域的第二类外部路由
此类路由描述了如何经 NSSA 区域的 ASBR 到达 AS 以外的目标网络。路由代价=ASBR 到目标网络的代价。仅 NSSA 区域的路由器上存在此类路由。

 E2、N2 这两类路由的可信度比较低，OSPF 认为从 ASBR 到自治系统之外的代价远远大于在自治系统之内到达 ASBR 的代价。所以，计算路由代价时只考虑 ASBR 到自治系统之外的代价。

相关配置

▾ 启动 OSPF

缺省情况下，未启动 OSPF。

使用 **router ospf 1** 命令，在路由器上创建一个 OSPF 进程。

使用 **network area** 命令，在接口上启动 OSPF，指定 Area ID。

使用 **area virtual-link** 命令，在路由器上创建一个虚链接。可将虚链接视为一个逻辑接口。

▾ Router ID

缺省情况下，OSPF 进程选举所有 Loopback 接口中 IP 地址最大的作为 Router ID。如果没有配置了 IP 地址的 Loopback 接口，则选取所有物理接口中 IP 地址最大的作为路由设备标识。

也可以使用 **router-id** 命令，人为指定 Router ID。

▾ 协议控制参数

使用 **ip ospf hello-interval** 命令，修改接口上的 Hello 消息发送间隔，缺省值是 10 秒（NBMA 类型 30 秒）。

使用 **ip ospf dead-interval** 命令，修改接口上的邻居死亡时间，缺省值是 Hello 报文发送间隔的四倍。

使用 **neighbor** 命令的 **poll-interval** 参数，修改 NBMA 类型接口上的邻居轮询间隔，缺省值是 120 秒。

使用 **ip ospf transmit-delay** 命令，修改接口上的 LSU 报文传输延时，缺省值是 1 秒。

使用 **ip ospf retransmit-interval** 命令，修改接口上的 LSU 报文重传间隔，缺省值是 5 秒。

使用 **areavirtual-link** 命令的 **hello-interval** 参数，修改虚链路上的 Hello 消息发送间隔，缺省值是 10 秒。

使用 **areavirtual-link** 命令的 **dead-interval** 参数，修改虚链路上的邻居死亡时间，缺省值是 Hello 报文发送间隔的四倍。

使用 **areavirtual-link** 命令的 **transmit-delay** 参数，修改虚链路上的 LSU 报文传输延时，缺省值是 1 秒。

使用 **areavirtual-link** 命令的 **retransmit-interval** 参数，修改虚链路上的 LSU 报文重传间隔，缺省值是 5 秒。

使用 **timers throttle lsa all** 命令，修改生成 LSA 的指数退避算法参数，缺省值是 0 毫秒，5000 毫秒，5000 毫秒。

使用 **timers spacing lsa-group** 命令，修改 LSA 分组刷新的时间，缺省值是 30 秒。

使用 **timers pacing lsa-transmit** 命令，修改（LSA 分组刷新的）LS-UPD 报文发送间隔和个数，缺省值是 40 毫秒、1 个。

使用 **timers lsa arrival** 命令，修改接收相同 LSA 的延时，缺省值是 1000 毫秒。

使用 **timers throttlespf** 命令，修改计算 SPT 的延时，两次计算的最小间隔、最大间隔。缺省值是 1000 毫秒、5000 毫秒、10000 毫秒。

▾ OSPF 网络类型

缺省情况下，Ethernet、FDDI 为广播类型；X.25、帧中继、ATM 为 NBMA 类型；PPP、HDLC、LAPB 为 P2P 类型。

使用 **ip ospf network** 命令，手工指定接口的网络类型。

使用 **neighbor** 命令，手工指定邻居。NBMA、点到多点非广播类型必须手工指定邻居。

使用 `ip ospf priority` 命令，调整接口的优先级（用于 DR/BDR 竞选）。广播、NBMA 类型需要竞选 DR/BDR。优先级最大者获胜，优先级为 0 者弃权。缺省值为 1。

4.3.2 OSPF路由管理

通过人为配置，可以合理规划或优化 OSPF 路由，实现管理的目的。

工作原理

↳ (Totally) Stub、NSSA 区域

(Totally) Stub、NSSA 区域有助于减轻协议交互负担、减小路由表的规模。

- 将适当的区域设置为 (Totally) Stub、NSSA 区域，则可以免除大量的 LSAType 5、3 在区域内传播。

区域	LSAType1、2	LSAType 3	LSAType 4	LSAType 5	LSAType 7
非 (Totally) Stub、NSSA 区域	允许	允许	允许	允许	不允许
Stub 区域	允许	允许，包括一条缺省路由	不允许	不允许	不允许
Totally Stub 区域	允许	只允许一条缺省路由	不允许	不允许	不允许
NSSA 区域	允许	允许，包括一条缺省路由	允许	不允许	允许
Totally NSSA 区域	允许	只允许一条缺省路由	允许	不允许	允许

- **i** ABR 使用 LSAType 3 向 (Totally) Stub、NSSA 区域发布一条缺省路由。
- **i** ABR 将 (Totally) NSSA 区域内的 LSAType 7 转换成 LSAType 5，发布到骨干区域。
- 将适当的区域设置为 (Totally) Stub、NSSA 区域，则区域内路由器的路由表中可以免除大量 E1、E2、IA 类型的路由。

区域	区域内路由器的路由表中存在哪些路由
非 (Totally) Stub、NSSA 区域	O：到本区域某目标网段的路由 IA：到其他区域某目标网段的路由 E1 或 E2：（经过 AS 内任意 ASBR）到 AS 外部某目标网段的路由、缺省路由
Stub 区域	O：到本区域某目标网段的路由 IA：到其他区域某目标网段的路由、缺省路由
Totally Stub 区域	O：到本区域某目标网段的路由 IA：缺省路由
NSSA 区域	O：到本区域某目标网段的路由 IA：到其他区域某目标网段的路由、缺省路由 N1 或 N2：（经过本区域内的 ASBR）到 AS 外部某目标网段的路由、缺省路由
Totally NSSA 区域	O：到本区域某目标网段的路由 IA：缺省路由 N1 或 N2：（经过本区域内的 ASBR）到 AS 外部某目标网段的路由、缺省路由

↳ 路由重分布

路由重分布是指：将设备上存在的“其他路由协议的路由、其他 OSPF 进程的路由、静态路由、直连路由”引入 OSPF 进程，以 LSA Type 5、7 的方式发送给邻居。路由重分布不能引入缺省路由。

路由重分布常用于 AS 间互联。通过在 ASBR 上配置路由重分布，可以将 AS 外部的路由信息传播到 AS 内部，也可以将 AS 内部的路由信息传播到 AS 外部。

引入缺省路由

通过在 ASBR 上配置命令，可以将一条缺省路由引入 OSPF 进程，以 LSA Type 5、7 的方式发送给邻居。

“引入缺省路由”常用于 AS 间互联，以一条缺省路由代替所有 AS 外部的路由信息。

路由汇聚

路由汇聚是指将具有相同前缀的路由信息汇聚为一条路由，并将汇聚后的路由（替代大量琐碎路由）发布给邻居。路由汇聚有助于减轻协议交互负担、减小路由表的规模。

缺省情况下，对于区域间的路由信息，ABR 以网段为单位发布 LSA Type 3；对于重分布进来的路由信息，ASBR 逐条发布 LSA Type 5、7。如果其中存在一些连续的网段，则建议配置路由汇聚。

配置路由汇聚时，汇聚范围有可能超出实际路由的网络范围。如果数据发往汇聚范围内不存在的网络，将有可能发生路由环路或加重路由器的处理负担。为阻止这种情况的发生，ABR 或 ASBR 会自动添加一条 discard 路由到路由表。该路由不会被传播。

路由过滤

OSPF 允许在学习、交互、使用路由信息的过程中进行过滤，以达到安全可控的目的。

通过配置命令，可以在以下几处进行路由过滤：

- 针对接口。禁止接口发出路由信息（任意 LSA）、或禁止接口与邻居交互路由信息（任意 LSA）。
- 针对区域间发布的路由信息。只有满足过滤条件的路由信息才能被发布到另一区域（LSA Type 3）。
- 针对 AS 外部的路由信息。只有满足过滤条件的路由信息才能被重分布到 OSPF 进程（LSA Type 5、7）。
- 针对路由器收到的 LSA。OSPF 路由表中，只有由满足过滤条件的 LSA 计算出的路由，才能执行报文转发。

路由代价

如果网络中存在冗余链路或设备，则从本地到目的网络之间可能存在多条路径。OSPF 选择其中总代价最小的路径形成 OSPF 路由。一条路径的总代价等于沿途链路上的代价之和。通过修改沿途链路上的代价，使一条路径的总代价最小，即可使 OSPF 选择此路径形成路由。

通过配置命令，可以修改如下链路代价：

- 从接口到直连网段的代价、从接口到邻居的代价
- 从 ABR 到区间汇聚网段的代价、从 ABR 到缺省网段的代价
- 从 ASBR 到外部网段的代价、从 ASBR 到缺省网段的代价

i cost、metric 都表示代价，没有实质上的差别。

OSPF 管理距离

管理距离 (administrative distance, 缩写为 AD) 用于评价路由的可信度, 取值范围 0~255 (整数)。AD 值越小, 则表示此路由更值得信任。到达同一目的地存在多条路由时, 路由器优先选择 AD 值较小的路由。AD 值较大者, 成为浮动路由 (作为最优路由的备份)。

缺省情况下, 每种来源的路由对应一个 AD 值。AD 值为本地概念, 修改 AD 值只会影响当前路由器上的路由选择。

路由来源	直连网络	静态路由	EBGP 路由	OSPF 路由	ISIS 路由	RIP 路由	IBGP 路由	不可达路由
缺省 AD	0	1	20	110	115	120	200	255

相关配置

▾ Stub、NSSA 区域

缺省情况下, 没有 Stub、NSSA 区域。

使用 **area stub** 命令, 将指定区域设置为 Stub 区域。

使用 **area nssa** 命令, 将指定区域设置为 NSSA 区域。

- i 骨干区域不能设置为 Stub (或 NSSA) 区域。
- i 传输区域 (虚链路穿过的区域) 不能设置为 Stub (或 NSSA) 区域。
- i 包含 ASBR 的区域不能设置为 Stub 区域。

▾ 路由重分布、引入缺省路由

缺省情况下, 不进行路由重分布, 不引入缺省路由。

使用 **redistribute** 命令, 配置路由重分布。

使用 **default-information originate** 命令, 引入缺省路由。

配置路由重分布、引入缺省路由后, 路由器自动成为 ASBR。

▾ 路由汇聚

缺省情况下, 不进行路由汇聚。如果配置路由汇聚, 会自动添加 discard 路由。

使用 **arearange** 命令, 在 ABR 上汇聚区域间发布的路由 (LSA Type 3)。

使用 **summary-address** 命令, 在 ASBR 上汇聚重分布路由 (LSA Type 5、7)。

使用 **discard-route** 命令, 向路由表中添加一条 discard 路由。

▾ 路由过滤

缺省情况下, 不进行路由过滤。

使用 **passive-interface** 命令, 设置被动接口。被动接口上不能交互路由信息 (任意 LSA)。

使用 **ip ospfdatabase-filter all out** 命令, 禁止接口发出路由信息 (任意 LSA)。

使用 **area filter-list** 命令, 在 ABR 上过滤区域间发布的路由信息。只有满足条件的路由信息才能被发布到另一区域 (LSA Type 3)。

使用 **redistribute** 命令的 **route-map** 参数，或使用 **distribute-list out** 命令，在 ASBR 上过滤 AS 外部的路由信息。只有满足过滤条件的路由信息才能被重分布到 OSPF 进程（LSA Type 5、7）。

使用 **distribute-list in** 命令，过滤路由器收到的 LSA。OSPF 路由表中，只有由满足过滤条件的 LSA 计算出的路由才能执行报文转发。

路由代价

- 从接口到直连网段的代价（即接口上的 cost 值）
缺省值为自动代价。自动代价 = 参考带宽/接口带宽。
使用 **auto-costreference-bandwidth** 命令，设置自动代价的参考带宽，缺省值为 100Mbps。
使用 **ip ospf cost** 命令，人为设置接口上的 cost 值。此项配置优先级高于自动代价。
- 从接口到指定邻居的代价（即从本地到指定邻居的 cost 值）
缺省值为自动代价。
使用 **neighbor** 命令中的 **cost** 参数，修改接口到指定邻居的代价。此项配置优先级高于接口上的 cost 值。
只有 P2MP 类型的接口允许此项配置。
- 从 ABR 到区间汇聚网段的代价（区域间路由汇聚的 cost 值）
如果兼容 RFC1583，缺省值为被汇总的链路代价的最小值，否则为被汇总的链路代价的最大值。
使用 **compatible rfc1583** 命令，设置 OSPF 选路规则兼容 RFC1583。缺省兼容 RFC1583。
使用 **area range** 命令的 **cost** 参数，修改区域间路由汇聚的 cost 值。
- 从 ABR 到缺省网段的代价（ABR 自动向 Stub、NSSA 区域发布的缺省路由的 cost 值）
缺省值为 1。
使用 **area default-cost** 命令，修改 ABR 自动向 Stub、NSSA 区域发布的缺省路由的 cost 值。
- 从 ASBR 到外部网段的代价（外部路由的 metric 值）
缺省情况下，重分布 BGP 路由的 metric 为 1，重分布其它路由的 metric 为 20，路由类型 Type2 External。
使用 **default-metric** 命令，修改外部路由的 metric 值的缺省值。
使用 **redistribute** 命令中的 **metric**、**metric-type**、**route-map** 参数，修改外部路由的 metric 值、路由类型。
- 从 ASBR 到缺省网段的代价（人为引入的缺省路由的 metric 值）
缺省情况下，metric 为 1，路由类型 Type2 External。
使用 **default-information originate** 命令的 **metric**、**metric-type**、**route-map** 参数，修改人为引入的缺省路由的 metric 值、路由类型。
使用 **area nssa** 命令的 **default-information-originate** 参数的 **metric**、**metric-type** 参数，修改 NSSA 区域内人为引入的缺省路由的 metric 值、路由类型。
- 此外，使用 **max-metric router-lsa** 命令，可以将路由器上发布的所有链路代价均为最大值。则途径此路由器的任何路径，总代价都会变得很大，几乎不可能成为最短路径。

OSPF 管理距离

缺省情况下，OSPF 管理距离为 110。

使用 **distance** 命令，设置 OSPF 路由的管理距离值。

4.3.3 增强安全性、可靠性

认证、与 BFD 联动等功能使 OSPF 更安全、更稳定可靠。

工作原理

▾ 认证

认证是为了阻止非法接入网络的路由器和伪造 OSPF 报文的主机参与 OSPF 协议过程。在 OSPF 接口（或虚链路两端）对收到的 OSPF 报文进行认证。如果认证失败，则丢弃报文，邻接关系将不能被建立。

启动认证功能后，可以避免学到非认证、无效路由，避免通告有效路由到非认证设备。在广播类型网络中，认证还可以避免非认证设备成为指定设备的可能性，保证了路由系统的稳定性和抗入侵性。

▾ MTU 校验

OSPF 在收到 DD 时会校验邻居的接口的 MTU 和自己接口的 MTU 是否相同，如果收到的 DD 报文中指示的接口的 MTU 大于接收接口的 MTU，则邻接关系将不能被建立。关闭 MTU 的校验，可以避免此问题。

▾ 源地址校验

通常情况下，OSPF 收到报文的源地址和接收接口会在同一个网段。不过在点到点链路上，链路两端的地址是独立设置的，可以不要求在同一网段，在这种场景下，由于点到点链路协商过程中会告知对端的地址信息，因此 OSPF 会检查报文的源地址是否就是协商时对端通告的地址，若不是则认为报文非法并丢弃该报文。特别的，对于 unnumbered 接口，OSPF 始终不进行地址校验。

但在某些应用场景中，可能会出现源地址不如上面所述，导致 OSPF 校验失败，比如点到点链路上，协商的对端地址无法获取。在这种场景中，需要关闭源地址校验功能，以保证 OSPF 邻居能正常建立。

▾ 双向维持

OSPF 路由器之间通过周期性发送 Hello 报文来维持邻接关系。在较大的网络中，可能会出现大量的报文收发，占用较高的 CPU 和内存资源，造成部分部分报文的延迟或丢弃。如果 Hello 报文的处理超出了死亡间隔，将导致邻接关系断开。

如果启用双向维持功能，则除了 Hello 报文外，DD、LSU、LSR、LSAck 报文也能用来维持邻居之间的双向通信，从而使邻接关系更加稳定。

▾ 并发邻居交互限制

当路由器同时与多个邻居交换数据时，其性能会受到影响。限制 OSPF 进程可并发交互（发起或接受）的邻居最大数目，使路由器可以分批与邻居进行交互，有助于保障数据转发或其他关键业务。

▾ OVERFLOW

OSPF 要求同一个区域中的路由器保存相同的 LSDB。随着网络上路由数量不断增加，一些路由器由于系统资源有限，不能承载如此多的路由信息，有可能导致路由器系统资源耗尽而失效的故障。

OVERFLOW 是通过限制 LSDB 中外部路由的数量，达到控制 LSDB 的规模的目的。

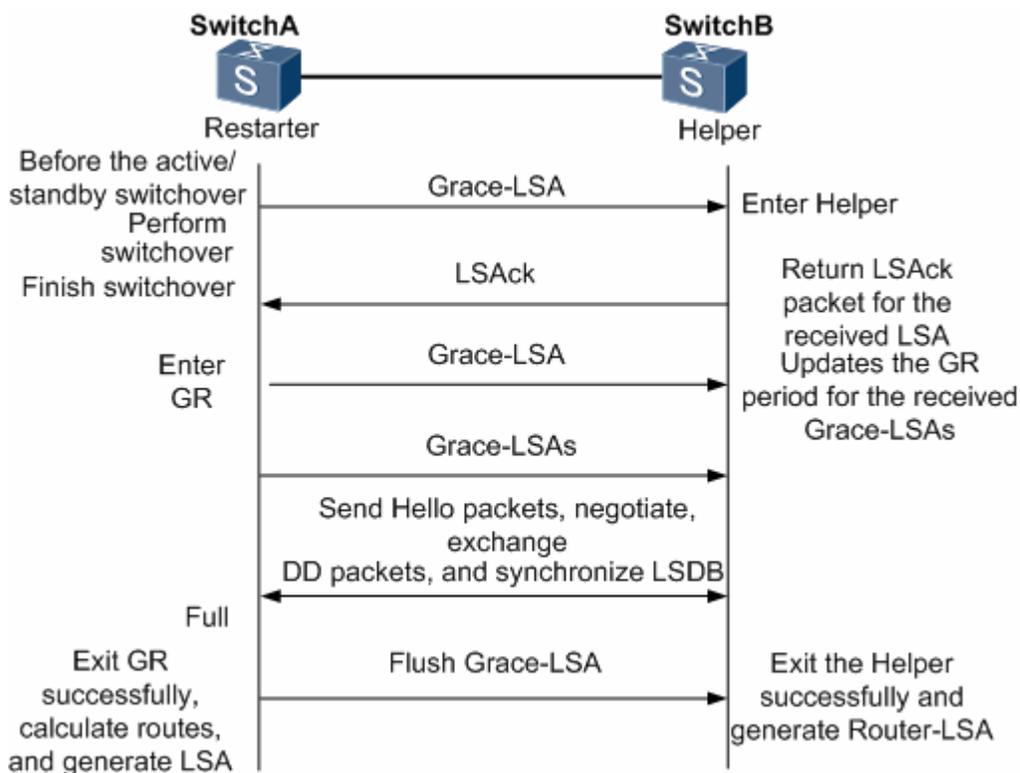
当路由器上的外部路由数量达到上限时，进入 Overflow 状态：删除 LSDB 中自己产生的外部路由、不产生新的外部路由、丢弃新收到的外部路由。Overflow 状态超时后（5 秒），如果外部路由数量低于上限，则恢复正常状态。

📌 GR

路由器普遍采用了控制和转发分离的技术。在网络拓扑保持稳定的情况下，启用 GR（Graceful Restart，优雅重启）功能的路由器在控制层面的重启过程中转发层面能够继续指导数据的转发，同时控制层面邻居关系的重建以及路由计算等动作不会影响转发层面的功能，从而避免了路由震荡引发的业务中断，提高了整网的可靠性。

目前，GR 功能只在主备切换和系统升级过程中使用。

图 4-7 正常的 OSPF GR 过程



- GR 过程需要 Restarter 和 Helper 的协作。Restarter 是发生 GR 的路由器。Helper 是 Restarter 的邻居。
- Restarter 在开始和退出 GR 时向邻居发送 Grace-LSA，通知邻居进入和退出 Helper 状态。
- Restarter 与 Helper 邻居关系达到 Full 状态，则 GR 可成功退出。

📌 与 BFD 联动、快速重路由

当发生链路故障后，OSPF 需要一段时间（约 40 秒）才能感知邻居死亡。而后再将此信息传播出去、并重新计算 SPT。这期间会发生流量中断。

- BFD（Bidirectional Forwarding Detection，双向转发检测）用于检测设备之间的连通性，最短能够在 150 毫秒内发现链路故障。将 OSPF 与 BFD 联动后，一旦发生链路故障，OSPF 最短可以在 150 毫秒内感知到邻居死亡。从而大大加快路由收敛，避免流量中断。

- 快速重路由会为 OSPF 准备一条备份路由，一旦 OSPF 感知到邻居死亡，立即将流量切换到备份路由，从而避免流量中断。

📄 iSPF 特性

- OSPF 的拓扑以区域为单位，每个区域通过运行 SPF 算法独立计算。对于标准的 SPF 算法，每次计算需要重新计算整个区域的拓扑，即使区域拓扑只有叶子节点变化。
- iSPF (增量 SPF) 在计算网络拓扑时，只修正最短路径树上受拓扑变化影响的节点，而不是重建整棵最短路径树。在大型网络中，这将有效缓解路由器处理器的压力，尤其是在网络不稳定的情况下。

相关配置

📄 OSPF 报文认证

缺省情况下，关闭认证功能。

- 使用 `area authentication` 命令，在整个区域上启动认证功能，在该区域内所有接口上生效。如果在区域 0 上启动认证功能，则也对虚链路生效。
- 使用 `ip ospf authentication` 命令，在接口上启动认证功能，优先级高于区域上的配置。
- 使用 `ip ospf authentication-key` 命令，在接口上设置文本认证的密码。
- 使用 `ip ospf message-digest-key` 命令，在接口上设置 MD5 认证的密码。
- 使用 `area virtual-link` 命令的 `authentication` 参数，在虚链路两端启动认证功能，优先级高于区域上的配置。
- 使用 `area virtual-link` 命令的 `authentication-key` 参数，在虚链路两端设置文本认证的密码。
- 使用 `area virtual-link` 命令的 `message-digest-key` 参数，在虚链路两端设置 MD5 认证的密码。

📄 MTU 校验

缺省情况下，关闭 MTU 校验。

使用 `ip ospf mtu-ignore` 命令，在接口上关闭 MTU 校验。

📄 源地址校验

缺省情况下，点到点类型的接口上启动源地址校验。

使用 `ip ospf source-check-ignore` 命令，在接口上关闭源地址校验。

📄 双向维持

缺省情况下，启动双向维持功能。

使用 `two-way-maintain` 命令，启动双向维持功能。

📄 并发邻居数量限制

使用 `max-concurrent-dd` 命令，修改当前 OSPF 进程并发交互的邻居数量限制。缺省值 5。

📄 OVERFLOW

使用 **overflow memory-lack** 命令，允许内存不足时进入 OVERFLOW 状态。缺省允许。

使用 **overflow database** 命令，允许 LSA 数量过多时进入 OVERFLOW 状态。缺省禁止。

使用 **overflow database external** 命令，允许 external-LSA 数量过多时进入 OVERFLOW 状态。缺省禁止。

▾ GR

缺省情况下，关闭 Restarter 功能，启动 Helper 功能。

使用 **graceful-restart** 命令，配置 Restarter 功能。

使用 **graceful-restart helper** 命令，配置 Helper 功能。

▾ OSPF 与 BFD 联动

缺省情况下，OSPF 不与 BFD 联动。

使用 **bfd interval min_rx multiplier** 命令，设置 BFD 参数。

使用 **bfd all-interfaces** 命令，所有接口启用 OSPF 与 BFD 联动。

使用 **ip ospf bfd** 命令，当前接口启动 OSPF 与 BFD 联动。

▾ 快速重路由

缺省情况下，关闭快速重路由。

使用 **fast-reroute route-map** 命令，在 OSPF 进程上启动快速重路由（引用路由图中定义的备份路由）。

使用 **fast-reroute lfa** 命令，在 OSPF 进程上启动快速重路由（使用无环路备份路径方式自动计算备份路由）。

使用 **fast-reroute lfdownstream-paths** 命令，在 OSPF 进程上启动快速重路由（使用下游路径方式自动计算备份路由）。

使用 **set fast-reroute backup-interfacebackup-nexthop** 命令，在路由图中定义备份路由。

使用 **ip ospf fast-reroute protection** 命令，指定接口的 LFA 保护方式。

使用 **ip ospf fast-reroute no-eligible-backup** 命令，禁止接口作为备份接口。

▾ iSPF 特性

缺省情况下，关闭 iSPF 特性。

使用 **ispf enable** 命令，在 OSPF 进程上启动 iSPF 特性。

4.3.4 支持网管功能

MIB、SYSLOG 等功能使 OSPF 更易于管理。

工作原理

▾ MIB

MIB（管理信息库）是设备端维护的设备状态信息集。管理程序可以对 MIB 节点进行查阅和设置。

路由器上可以同时启动多个 OSPF 进程，但 OSPF MIB 只能与一个 OSPF 进程绑定。

↳ Trap

告警是系统检测到故障而产生的通知，告警中携带对应的故障信息。

启动 Trap 功能，则路由器可以主动将告警信息发送到网管设备。

↳ SYSLOG

日志记录了用户对路由器的操作（比如命令配置）以及特定事件（比如网络连接失效）等信息。

允许日志记录邻接状态变化，则网络管理员可以通过查看日志了解 OSPF 邻居之间建立和维持邻接关系的全过程。

相关配置

↳ MIB

缺省情况下，MIB 绑定在进程号最小的 OSPF 进程上。

使用 **enable mib-binding** 命令，将 MIB 绑定在当前 OSPF 进程上。

↳ Trap

缺省情况下，所有 TRAP 开关关闭，且不允许设备发送 OSPFTrap。

使用 **enable traps** 命令，在 OSPF 进程下打开指定的 Trap 开关。

使用 **snmp-server enable traps ospf** 命令，允许设备发送 OSPF Trap。

↳ SYSLOG

缺省情况下，允许日志记录邻接状态变化。

使用 **log-adj-changes** 命令，允许日志记录邻接状态变化。

4.4 配置详解

配置项	配置建议&相关命令	
配置 OSPF 基本功能	 必须配置。用于构建 OSPF 路由域。	
	router ospf	创建 OSPF 进程。
	router-id	设置 Router ID。
	network area	在接口上启动 OSPF，指定 Area ID。
	area virtual-link	创建虚链路。
设置网络类型	 可选配置。如果物理网络为 X.25、帧中继、ATM，则必须配置此项。	
	ip ospf network	定义网络类型。
	neighbor	指定邻居。
	ip ospf priority	配置 DR 优先级。

配置路由重分布、缺省路由	 可选配置。如果 OSPF 路由域连接外网，则建议配置此项。	
	redistribute	路由重分布。
	default-information originate	引入缺省路由。
设置 Stub、NSSA 区域	 可选配置。有助于减少路由信息的交互，缩小路由表规模，增强路由稳定性。	
	area stub	设置 Stub 区域。
	area nssa	设置 NSSA 区域。
配置路由汇聚	 可选配置。有助于减少路由信息的交互，缩小路由表规模，增强路由稳定性。	
	area range	汇聚区域间发布的路由信息。
	summary-address	汇聚重分布引入的路由信息。
	discard-route	向路由表中添加一条 discard 路由。
配置路由过滤	 可选配置。人为控制路由信息交互过程、筛选可用的 OSPF 路由。	
	passive-interface	设置被动接口。
	ip ospf database-filter all out	禁止接口发出 LSA。
	area filter-list	对区域间发布的路由信息进行过滤。
	distribute-list out	对重分布引入的路由进行过滤。
	distribute-list in	对收到的 LSA 计算出的路由进行过滤。
修改路由代价、管理距离	 可选配置。人为控制 OSPF 计算的最短路径、是否优先选择 OSPF 路由。	
	auto-cost reference-bandwidth	修改自动代价的参考带宽。
	ip ospf cost	修改接口出方向的 cost 值。
	area default-cost	修改 Stub/NSSA 区域缺省路由的 cost 值。
	default-metric	修改重分布路由的缺省 metric 值。
	max-metric router-lsa	设置最大度量值。
	compatible rfc1583	设置选路规则兼容 RFC1583。
	distance	修改 OSPF 管理距离。
启动认证功能	 可选配置。阻止非法接入网络的路由器和伪造 OSPF 报文的主机参与 OSPF 协议。	
	area authentication	在区域上启动认证、设置认证方式
	ip ospf authentication	在接口上启动认证、设置认证方式
	ip ospf authentication-key	在接口上设置文本认证的密码
	ip ospf message-digest-key md5	在接口上设置 MD5 认证的密码
启动溢出功能	 可选配置。避免由于过度消耗内存导致 OSPF 进程停止运行的问题。	
	overflow memory-lack	允许内存不足时进入 OVERFLOW 状态。
	overflow database	允许 LSA 数量超限后进入 OVERFLOW 状态。
	overflow database external	允许 external LSA 数量超限后进入 OVERFLOW 状态。

修改并发邻居数量限制	⚠️ 可选配置。避免由于过度消耗 CPU 导致设备性能下降的问题。	
	max-concurrent-dd	修改当前 OSPF 进程上并发邻居数量限制。
	router ospf max-concurrent-dd	修改所有 OSPF 进程上并发邻居总数限制。
关闭源地址校验	⚠️ 可选配置。避免由于无法获取对端地址导致无法建立邻接关系的问题。	
	ip ospf source-check-ignore	在接口上关闭源地址校验功能。
关闭 MTU 校验	⚠️ 可选配置。避免由于邻居接口 MTU 不一致导致无法建立邻接关系的问题。	
	ip ospf mtu-ignore	在接口上关闭 MTU 校验功能。
启动双向维持	⚠️ 可选配置。避免由于 Hello 报文延迟或丢失导致邻接关系断开的问题。	
	two-way-maintain	启动双向维持功能。
启动 GR 功能	⚠️ 可选配置。在 OSPF 进程重启或主备切换过程中，维持 OSPF 路由转发，避免流量中断。	
	graceful-restart	配置 Restarter 功能。
	graceful-restart helper	配置 Helper 功能。
启动 OSPF 与 BFD 联动	⚠️ 可选配置。在链路故障时，快速发现邻居死亡，避免流量中断。	
	bfd interval min_rx multiplier	设置 BFD 参数。
	bfd all-interfaces	在所有接口上启动 OSPF 与 BFD 联动。
启动快速重路由	⚠️ 可选配置。在链路故障时，快速切换备份路由，避免流量中断。	
	fast-reroute route-map	在 OSPF 进程上启动快速重路由（引用路由图中定义的备份路由）。
	fast-reroute lfa	在 OSPF 进程上启动快速重路由（使用无环路备份路径方式自动计算备份路由）
	fast-reroute lfdownstream-paths	在 OSPF 进程上启动快速重路由（使用下游路径方式自动计算备份路由）
	set fast-reroute backup-interface backup-nexthop	在路由图中定义备份路由。
	ip ospf fast-reroute protection	指定接口的 LFA 保护方式。
	ip ospf fast-reroute no-eligible-backup	禁止接口作为备份接口。
启动 iSPF	⚠️ 可选配置。启用增量拓扑计算，缓解处理器压力。	
	ispf enable	在 OSPF 进程下启用 iSPF
配置网管功能	⚠️ 可选配置。使用户可以通过 SNMP 网管软件来管理 OSPF。	
	enable mib-binding	在 OSPF 进程下绑定 MIB。
	enable traps	在 OSPF 进程下打开指定的 Trap 开关。
	snmp-server enable traps ospf	允许设备发送 OSPF Trap。
	log-adj-changes	允许日志记录邻接状态变化。

修改协议控制参数	 可选配置。除非必要，不建议修改协议控制参数。	
	<code>ip ospf hello-interval</code>	修改 hello 报文发送间隔。
	<code>ip ospf dead-interval</code>	修改邻居死亡时间。
	<code>timers throttle lsa all</code>	修改生成 LSA 的指数退避算法参数。
	<code>timers throttle route inter-area</code>	修改区间路由计算的延时时间。
	<code>timers throttle route ase</code>	修改外部路由计算的延时时间。
	<code>timers pacing lsa-group</code>	修改分组刷新 LSA 的分组时间。
	<code>timers pacing lsa-transmit</code>	修改分组刷新 LSA 的发送间隔。
	<code>ip ospf transmit-delay</code>	修改 LSU 报文传输延时。
	<code>ip ospf retransmit-interval</code>	修改 LSU 报文重传间隔。
	<code>timers lsa arrival</code>	修改接收相同 LSA 的延时。
<code>timers throttle spf</code>	修改计算 SPT 的计时器。	

4.4.1 配置OSPF基本功能

配置效果

- 在网络中建立一个 OSPF 路由域，为网络内用户提供 IPv4 单播路由服务。

注意事项

- 需要确保 IP 单播路由功能启用，即 `ip routing` 开关未被关闭，才能够启用 OSPF。
- 强烈建议手工配置 Router ID。
- 配置 `ip ospf disable all` 后，即使接口属于该 network，也不再接收和发送任何 OSPF 报文，也不参与 OSPF 的计算。

配置方法

▾ 创建 OSPF 进程

- 必须配置。
- 每台路由器都需要配置。

▾ 设置 Router ID

- 可选配置，但强烈建议手工配置 Router ID。
- 不配置 Router ID 时，OSPF 会选择接口 IP 地址，若没有接口配置 IP 地址或者配置的 IP 地址已经被其他 OSPF 实例使用，则必须手工配置 Router ID。

▾ 在接口上启动 OSPF，指定 Area ID

- 必须配置。
- 每台路由器都需要配置。

检验方法

- 通过 `show ip route ospf` 命令查看 OSPF 路由表项加载正确。
- 通过 `ping` 命令查看 IPv4 单播业务正确。

相关命令

创建 OSPF 进程

- 【命令格式】 `router ospf process-id`
- 【参数说明】 `process-id` : OSPF 进程号。未配置进程号时, 表示配置进程 1。
- 【命令模式】 全局配置模式。
- 【使用指导】 不同 OSPF 进程间相互独立, 可以近似认为是两个独立运行的路由协议。

设置 Router ID

- 【命令格式】 `router-id router-id`
- 【参数说明】 `router-id` : 要设置的路由设备的 ID, 以 IP 地址形式表示。
- 【命令模式】 OSPF 路由进程配置模式。
- 【使用指导】 不同 OSPF 进程间相互独立, 可以近似认为是两个独立运行的路由协议。
每个 OSPF 进程使用不同的 Router ID。

在接口上启动 OSPF, 指定 Area ID

- 【命令格式】 `network ip-address wildcard area area-id`
- 【参数说明】 `ip-address` : 接口对应的 IP 地址。
`wildcard` : 定义 IP 地址比较比特位, 0 表示精确匹配, 1 表示不做比较。
`area-id` : OSPF 区域标识。一个 OSPF 区域总是关联一个地址范围, 为了便于管理, 也可以用子网作为 OSPF 区域标识。
- 【命令模式】 OSPF 路由进程配置模式。
- 【使用指导】 `ip-address` 和 `wildcard` 两个参数的定义, 允许用一条命令就可以将多个接口关联到一个 OSPF 区域。要在一个接口上运行 OSPF, 必须将该接口的主 IP 地址包括在 `network area` 定义的 IP 地址范围内。如果 `network area` 定义的 IP 地址范围只包括接口的次 IP 地址, 该接口将不会运行 OSPF。当接口地址同时与多个 OSPF 进程 `network` 命令定义的 IP 地址范围相匹配时, 按照最优匹配方式, 确定接口参与的 OSPF 进程。

设置虚链路

- 【命令格式】 `area area-id virtual-link router-id [authentication [message-digest | null]] [dead-interval seconds] [hello-interval seconds] [retransmit-interval seconds] [transmit-delay seconds] [authentication-key [0 | 7] key] [message-digest-key key-id md5 [0 | 7] key]`
- 【参数说明】 `area-id` : OSPF 过渡区域号。区域号可以是一个十进制整数值, 也可以是一个 IP 地址。

router-id : 虚拟链路邻居的路由设备标识。。

dead-interval seconds : 定义宣告邻居丢失时间, 以秒为单位, 取值范围: 0-2147483647。该参数必须与邻居保持一致。

hello-interval seconds : OSPF 往虚拟链路发送 Hello 报文的时间间隔, 以秒为单位, 取值范围: 1-65535。该参数必须与邻居保持一致。

retransmit-interval seconds : OSPF 链路状态描述 (LSA) 重传时间, 以秒为单位, 取值范围: 0-65535。

transmit-delay seconds : OSPF 发送链路状态描述 (LSA) 的时延, 以秒为单位, 取值范围: 0-65535。

authentication-key [0 | 7] key : 定义 OSPF 明文认证的密钥。

message-digest-key key-id md5 [0 | 7] key : 定义 OSPF MD5 认证的密钥标识符和密钥。

authentication : 设置认证类型, 使用明文认证。

message-digest : 设置认证类型为 MD5 认证。

null : 设置为不认证。

【命令模式】 OSPF 路由进程配置模式。

【使用指导】 在 OSPF 路由域中, 所有的区域都必须与骨干域连接, 如果骨干域断接, 就需要配置虚拟链接将骨干域接续起来, 否则网络通讯将出现问题。虚拟链接需要在两个区域边界路由设备 (ABR) 之间创建, 两个 ABR 共同所属的区域成为过渡区域。残域 (Stub Area) 和 NSSA 是不能作为过渡区域的。虚拟链接也可以用于接续其它非骨干区域。

router-id 为 OSPF 邻居路由设备标识符, 如果不确定 **router-id** 的值, 请用 **show ip ospf neighbor** 命令进行确认。可以通过配置 Loopback 地址做路由标识符。

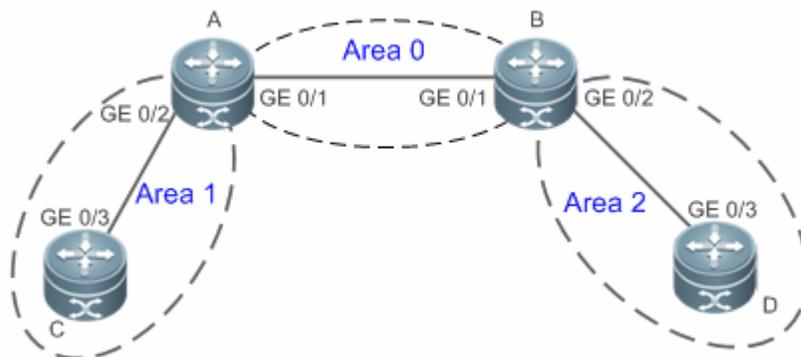
area virtual-link 命令只定义虚拟链路的认证密钥, 要启用虚拟链路所连接区域的 OSPF 报文认证, 用路由进程配置命令 **area authentication**。

虚链路两端的设备死亡间隔始终要保持一致, **hello-multiplier** 值不必设置为一致, 只要保证在死亡间隔内能够接收到至少 1 个 Hello 报文即可。

使用 **show ip ospf virtual-links** 命令监控虚链路配置的死亡间隔。

配置举例

【网络环境】



【注释】 接口 IP 地址 :

```
A : GE 0/1 192.168.1.1   GE 0/2 192.168.2.1
B : GE 0/1 192.168.1.2   GE 0/2 192.168.3.1
C : GE 0/3 192.168.2.2
D : GE 0/3 192.168.3.2
```

【配置方法】

- 在所有路由器上配置接口 IP 地址。
- 在所有路由器上启动 IPv4 单播路由功能（该功能默认已开启）。
- 在所有路由器上配置 OSPF 实例、Router ID
- 在所有路由器上配置接口使能 OSPF

```
A A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip address 192.168.1.1 255.255.255.0
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)# ip address 192.168.2.1 255.255.255.0
A(config-if-GigabitEthernet 0/2)# exit
A(config)# router ospf 1
A(config-router)# router-id 192.168.1.1
A(config-router)# network 192.168.1.0 0.0.0.255 area 0
A(config-router)# network 192.168.2.0 0.0.0.255 area 1

B B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip address 192.168.1.2 255.255.255.0
B(config-if-GigabitEthernet 0/1)# exit
B(config)# interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/2)# ip address 192.168.3.1 255.255.255.0
B(config-if-GigabitEthernet 0/2)# exit
B(config)# router ospf 1
B(config-router)# router-id 192.168.1.2
B(config-router)# network 192.168.1.0 0.0.0.255 area 0
B(config-router)# network 192.168.3.0 0.0.0.255 area 2

C C# configure terminal
C(config)# interface GigabitEthernet 0/3
C(config-if-GigabitEthernet 0/3)# ip address 192.168.2.2 255.255.255.0
C(config-if-GigabitEthernet 0/3)# exit
C(config)# router ospf 1
C(config-router)# router-id 192.168.2.2
C(config-router)# network 192.168.2.0 0.0.0.255 area 1

D D# configure terminal
D(config)# interface GigabitEthernet 0/3
D(config-if-GigabitEthernet 0/3)# ip address 192.168.3.2 255.255.255.0
```

```
D(config-if-GigabitEthernet 0/3)# exit
D(config)# router ospf 1
D(config-router)# router-id 192.168.3.2
D(config-router)# network 192.168.3.0 0.0.0.255 area 2
```

【检验方法】

- 所有路由器上检查 OSPF 邻居正确。
- 所有路由器上检查路由表加载正确。
- D 上 ping 192.168.2.2 正确。

A

```
A# show ip ospf neighbor
OSPF process 1, 2 Neighbors, 2 is Full:
Neighbor ID  Pri  State      Dead Time  Address      Interface
192.168.1.2  1    Full/DR    00:00:40   192.168.1.2  GigabitEthernet 0/1
192.168.2.2  1    Full/BDR   00:00:34   192.168.2.2  GigabitEthernet 0/2
```

```
A# show ip route ospf
0 IA 192.168.3.0/24 [110/2] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
```

B

```
B# show ip ospf neighbor
OSPF process 1, 2 Neighbors, 2 is Full:
Neighbor ID  Pri  State      Dead Time  Address      Interface
192.168.1.1  1    Full/BDR   00:00:32   192.168.1.1  GigabitEthernet 0/1
192.168.3.2  1    Full/BDR   00:00:30   192.168.3.2  GigabitEthernet 0/2
```

```
B# show ip route ospf
0 IA 192.168.2.0/24 [110/2] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
```

C

```
C# show ip ospf neighbor
OSPF process 1, 1 Neighbors, 1 is Full:
Neighbor ID  Pri  State      Dead Time  Address      Interface
192.168.1.1  1    Full/BDR   00:00:32   192.168.2.1  GigabitEthernet 0/3
```

```
C# show ip route ospf
0 IA 192.168.1.0/24 [110/2] via 192.168.2.1, 00:19:05, GigabitEthernet 0/3
0 IA 192.168.3.0/24 [110/3] via 192.168.2.1, 00:19:05, GigabitEthernet 0/3
```

D

```
D# show ip ospf neighbor
OSPF process 1, 1 Neighbors, 1 is Full:
Neighbor ID  Pri  State      Dead Time  Address      Interface
192.168.1.2  1    Full/BDR   00:00:30   192.168.3.1  GigabitEthernet 0/3
```

```
D# show ip route ospf
0 IA 192.168.1.0/24 [110/2] via 192.168.3.1, 00:19:05, GigabitEthernet 0/3
0 IA 192.168.2.0/24 [110/3] via 192.168.3.1, 00:19:05, GigabitEthernet 0/3
```

```
D# ping 192.168.2.2
Sending 5, 100-byte ICMP Echoes to 192.168.2.2, timeout is 2 seconds :
< press Ctrl+C to break >
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms.
```

常见错误

- IP 单播路由功能被关闭，OSPF 协议无法启用。
- **network** 命令配置的网段范围未包含接口 IP 地址。
- 相连接口使能的区域号不一致。
- 多台路由器上配置了相同的 Router ID，导致 Router ID 冲突。
- 多台路由器上配置了相同的接口 IP 地址，导致 OSPF 网络运行错误。

4.4.2 设置网络类型

配置效果

- 物理网络为 X.25、帧中继、ATM 时，也可以运行 OSPF 协议，提供 IPv4 单播路由服务。

注意事项

- 必须配置 OSPF 基本功能
- 广播网络以组播方式发送 OSPF 报文，自动发现邻居，需要选举 DR 和 BDR。
- P2P 网络以组播方式发送 OSPF 报文，自动发现邻居。
- NBMA 网络以单播方式发送 OSPF 报文，必须手动指定邻居，需要选举 DR 和 BDR。
- P2MP 网络（不携带 **non-broadcast** 参数）以组播方式发送 OSPF 报文，自动发现邻居。
- P2MP 网络（携带 **non-broadcast** 参数）以单播方式发送 OSPF 报文，必须手动指定邻居。

配置方法

📄 配置接口网络类型

- 可选配置。
- 链路两端路由器都需要配置。

📄 配置邻居

- 可选配置。若接口的网络类型配置为 NBMA 或者 P2MP（携带 **non-broadcast** 参数），则必须配置。
- 在 NBMA 或者 P2MP（携带 **non-broadcast** 参数）网络两端路由器上配置。

配置接口优先级

- 可选配置。若需要指定路由器做为 DR 或者禁止路由器做为 DR，需要配置。
- 在被指定为 DR 或者被禁止做为 DR 的路由器上配置。

检验方法

- 通过 `show ip ospf interface` 查看接口的网络类型正确。

相关命令

配置接口网络类型

【命令格式】 **ip ospf network { broadcast | non-broadcast | point-to-multipoint [non-broadcast] | point-to-point }**

【参数说明】

broadcast：将接口设定为广播类型。

non-broadcast：将接口设定为 NBMA 类型。

point-to-multipoint [non-broadcast]：将接口设定为 P2MP 类型，如果接口不具备广播能力，需携带 **non-broadcast** 参数。

point-to-point：将接口设定为 P2P 类型。

【命令模式】 接口配置模式

【使用指导】 广播类型要求接口具有广播能力。

P2P 类型要求接口 1 对 1 互联。

NBMA 类型要求全网状连接，所有互联的路由设备必须能够直接通讯。

P2MP 类型无要求。

配置邻居

【命令格式】 **neighbor ip-address [poll-interval seconds] [priority priority] [cost cost]**

【参数说明】

ip-address：邻居接口的 IP 地址。

poll-interval seconds：邻居轮询时间间隔，以秒为单位，取值范围：0-2147483647。只有 NBMA 类型支持该参数。

priority priority：邻居的优先权值，取值范围：0-255。只有 NBMA 类型支持该参数。

cost cost：到每个邻居所要的费用，无缺省值，取值范围：0-65535。只有 P2MP 类型支持该选项。

【命令模式】 OSPF 路由进程配置模式

【使用指导】 必须为 NBMA 类型接口和 P2MP 类型非广播接口指定邻居。邻居的 IP 地址，必须为该邻居接口的主 IP 地址。在 NBMA 网络上如果邻居路由设备变为不活跃，即在路由设备死亡时间内还没有收到 Hello 报文，OSPF 还会往该邻居发送 Hello 报文，发送 Hello 报文的时间间隔称为轮询间隔。当 OSPF 第一次开始运行时，只向优先权值不为 0 的邻居发送 Hello 报文，这样优先权值设为 0 的邻居就不会参与 DR/BDR 的竞选。等 DR/BDR

产生以后，DR/BDR 会向所有的邻居发送 Hello 报文建立邻居关系。

在 P2MP 非广播网络上，因为不具备广播能力所以无法动态的发现邻居，因此必须使用该命令手工为其配置邻居。此外 P2MP 类型的网络可以通过 cost 选项指定到达每个邻居的花费。

配置接口优先级

【命令格式】 **ip ospf priority priority**

【参数说明】 *priority*：设置接口的 OSPF 优先权值，取值范围：0-255。

【命令模式】 接口配置模式

【使用指导】 OSPF 接口的优先权值包含在 hello 报文中。当 OSPF 广播类型网络发生 DR/BDR（指定路由设备 / 备份指定路由设备）竞选时，高优先权值的路由设备，将成为 DR 或 BDR。如果优先权值一样，路由设备标识符高的路由设备将成为 DR 或 BDR。优先权值为 0 的路由设备，不参与 DR/BDR 竞选。

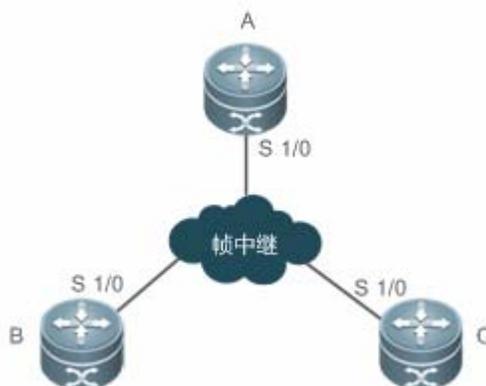
该命令只对 OSPF 广播类型和 NBMA 类型有效。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见 1.4.1.配置 OSPF 基本功能的介绍。

配置接口网络类型为 P2MP。

【网络环境】



【注释】 接口 IP 地址：

A：S1/0 192.168.1.2

B：S1/0 192.168.1.3

C：S1/0 192.168.1.4

- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 所有路由器上配置接口的网络类型为 P2MP。

A

```
A# configure terminal
A(config)# interface Serial 1/0
A(config-Serial1/0)# encapsulation frame-relay
A(config-Serial1/0)# ip ospf network point-to-multipoint
```

```
B# configure terminal
B(config)# interface Serial 1/0
B(config-Serial1/0)# encapsulation frame-relay
B(config-Serial1/0)# ip ospf network point-to-multipoint

C# configure terminal
C(config)# interface Serial 1/0
C(config-Serial1/0)# encapsulation frame-relay
C(config-Serial1/0)# ip ospf network point-to-multipoint
```

【检验方法】 查看接口的网络类型为 P2MP。

```
A# show ip ospf interface Serial 1/0
Serial 1/0 is up, line protocol is up
  Internet Address 192.168.1.2/24, Ifindex 2, Area 0.0.0.1, MTU 1500
  Matching network config: 192.168.1.0/24
  Process ID 1, Router ID 192.168.1.2, Network Type POINTMULTIPOINT, Cost: 1
  Transmit Delay is 1 sec, State Point-To-Point
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:02
  Neighbor Count is 1, Adjacent neighbor count is 0
  Crypt Sequence Number is 4787
  Hello received 465 sent 466, DD received 8 sent 8
  LS-Req received 2 sent 2, LS-Upd received 8 sent 21
  LS-Ack received 14 sent 7, Discarded 3
```

常见配置错误

- 两端接口配置的网络类型不一致，导致路由学习不正常。
- 网络类型配置为 NBMA，P2MP（携带 **non-broadcast** 参数），但没有指定邻居。

4.4.3 配置路由重分布、缺省路由

配置效果

- 在 OSPF 域内引入其他 AS 域的单播路由，为 OSPF 域内的用户提供到其他 AS 域的单播路由服务。
- 往 OSPF 域内注入一条通往其他 AS 域的默认路由，为 OSPF 域内的用户提供到其他 AS 域的单播路由服务。

注意事项

- 必须配置 OSPF 基本功能

配置方法

配置重分发外部路由

- 可选配置。若在 AS 边界路由器上需要引入 OSPF 域外路由时，需要配置
- 在 AS 边界路由器上配置。

配置生成缺省路由

- 可选配置。若在 AS 边界路由器上需要引入默认路由，使得 OSPF 域内访问其他 AS 域默认通过该边界路由器，则需要配置。
- 在 AS 边界路由器上配置。

检验方法

- OSPF 域内路由器通过 **show ip route** 查看加载了其他 AS 域的单播路由。
- OSPF 域内路由器通过 **show ip route** 查看加载了到达 ASBR 的默认路由。
- 通过 **ping** 命令查看通往其他 AS 域的 IPv4 单播业务正确。

相关命令

配置重分发路由

【命令格式】 **redistribute { connected [area-tag] | ospf process-id | rip | static } [{ level-1 | level-1-2 | level-2 }] [match { internal | external [1 | 2] | nssa-external [1 | 2] }] [metric metric-value] [metric-type { 1 | 2 }] [route-map route-map-name] [subnets] [tag tag-value]**

【参数说明】 **connected** : 从直连路由重分布

ospf process-id : 从 ospf 重分布，通过 process-id 指定具体某个 ospf 实例，取值范围：1-65535

rip : 从 rip 重分布

static : 从静态路由重分布

match : 只针对配置重分布 OSPF 路由时使用，用于过滤特定路由进行重分布，默认状态允许所有的 OSPF 路由进行重分布。

metric metric-value : 设置 OSPF external LSA 的 metric，通过 metric-value 指定 metric 大小，取值范围：0-16777214。

metric-type { 1 | 2 } : 设置外部路由类型为 E-1 或 E-2。

route-map route-map-name : 设置重分布过滤规则。

subnets : 可重分布非标准类网络。

tag tag-value : 设置重分布到 OSPF 内路由的 tag 值，取值范围：0-4294967295。

【配置模式】 OSPF 路由进程配置模式

【使用指导】 -

配置导入缺省路由

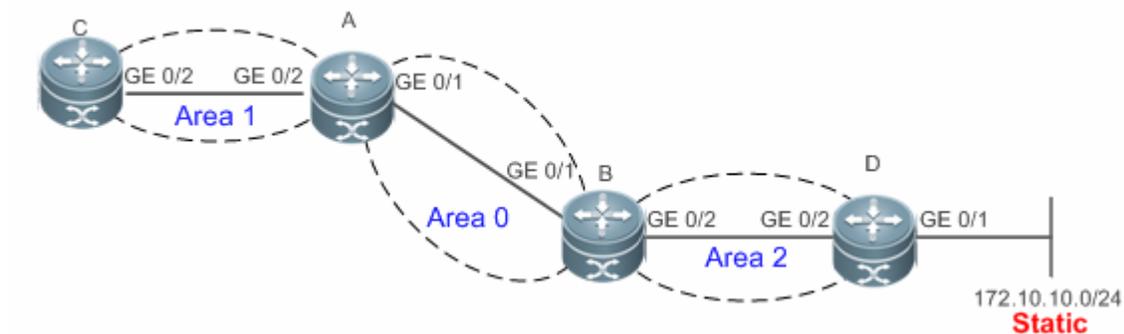
- 【命令格式】** `default-information originate [always] [metric metric] [metric-type type] [route-map map-name]`
- 【参数说明】**
- always** : 该关键字使得 OSPF 无条件产生缺省路由, 不管本地是否存在缺省路由。
 - metric *metric*** : 缺省路由初始的量度值, 取值范围: 0-16777214。
 - metric-type *type*** : 缺省路由的类型。OSPF 外部路由有两种类型: 类型 1, 不同路由设备上看到的量度值不一样; 类型 2, 所有路由设备看到的量度值都一样。类型 1 的外部路由比类型 2 可信。
 - route-map *map-name*** : 关联的 route-map 的名字, 缺省没有关联 route-map。
- 【配置模式】** OSPF 路由进程配置模式
- 【使用指导】** 当执行 **redistribute** 或者 **default-information** 命令时, OSPF 路由设备会自动变为 ASBR (自治域边界路由设备)。但是 ASBR 不会自动产生缺省路由, 并通告给 OSPF 路由域中的所有路由设备。ASBR 产生缺省路由, 必须用 **default-information originate** 路由进程配置命令进行配置。
- 如果选择 **always** 参数, OSPF 路由进程不管是否存在缺省路由, 都会向邻居通告一条外部缺省路由。但是本地路由设备不会显示该缺省路由, 要确认是否产生缺省路由, 可以用 **show ip ospf database** 观察 OSPF 链路状态数据库, 链路标识为 0.0.0.0 的外部链路描述了缺省路由。OSPF 的邻居通过执行 **show ip route** 命令, 是可以看到缺省路由的。
- 外部缺省路由的路由量度值, 只能由 **default-information originate** 命令进行定义, 不能通过命令 **default-metric** 设置。
- OSPF 有两种外部路由, 类型 1 外部路由的路由量度会改变, 而类型 2 外部路由的路由量度是固定不变。到一个目标网络存在两条并行路径, 同样的路由量度, 类型 1 会比类型 2 优先级高, 所以 **show ip route** 只显示类型 1 那条路由。
- Stub 区域中的路由设备是不可以产生外部缺省路由的。
- 所关联的 route-map 的 set metric 取值范围为 0-16777214, 如果超出该范围, 将无法引入路由。

配置举例

 以下配置举例, 需要基于 OSPF 基本功能进行配置, OSPF 基本功能请参见 1.4.1.配置 OSPF 基本功能的介绍。

配置重分发静态路由

【网络环境】



【注释】 接口 IP 地址：

A : GE 0/1 192.168.1.1 GE 0/2 192.168.2.1
 B : GE 0/1 192.168.1.2 GE 0/2 192.168.3.1
 C : GE 0/2 192.168.2.2
 D : GE 0/1 192.168.6.2 GE 0/2 192.168.3.2

【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 路由器 D 上配置引入外部静态路由。

D

```
D# configure terminal
D(config)# ip route 172.10.10.0 255.255.255.0 192.168.6.3
D(config)# router ospf 1
D(config-router)# redistribute static subnets
```

【检验方法】

- 路由器 D 上使用 **show ip ospf database external brief** 查看有外部路由对应的 LSA 生成。
- 路由器 C 上使用 **show ip route ospf** 查看已经加载引入的外部静态路由。

D

```
D# show ip ospf database external brief
      OSPF Router with ID (192.168.22.30) (Process ID 1)
      AS External Link States

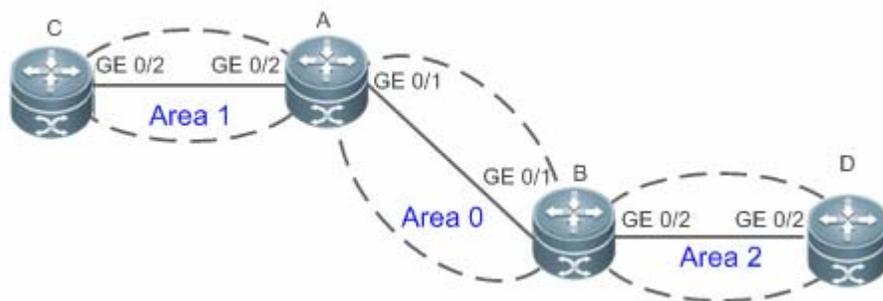
Link ID      ADV Router   Age  Seq#       CkSum  Route                Tag
-----
172.10.10.0  192.168.22.30  11  0x80000001  0xa4bb  E2 172.10.10.0/24    0
```

C

```
C# show ip route ospf
0 E2 172.10.10.0/24 [110/20] via 192.168.2.1, 00:18:03, GigabitEthernet 0/2
```

配置缺省路由

【网络环境】



【注释】 接口 IP 地址：

A : GE 0/1 192.168.1.1 GE 0/2 192.168.2.1

B : GE 0/1 192.168.1.2 GE 0/2 192.168.3.1

C : GE 0/2 192.168.2.2

D : GE 0/2 192.168.3.2

- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 路由器 D 上配置缺省路由。

```
D
D# configure terminal
D(config)# router ospf 1
D(config-router)# default-information originate always
```

- 【检验方法】
- 在路由器 D 上使用 **show ip ospf database external brief** 查看有缺省路由对应的 LSA。
 - 在路由器 C 上使用 **show ip route ospf** 查看有 ospf 缺省路由。

```
D
D# show ip ospf database external brief
      OSPF Router with ID (192.168.22.30) (Process ID 1)
      AS External Link States

Link ID      ADV Router   Age  Seq#           CkSum  Route           Tag
0.0.0.0      192.168.22.30 565  0x80000002    0xa190  E2 0.0.0.0/0    1

C
C# show ip route ospf
0 E2 0.0.0.0/0 [110/20] via 192.168.2.1, 00:18:03, GigabitEthernet 0/2
```

常见配置错误

- **redistribute** 命令的 **subnets** 参数未配置，导致子网路由未被引入。
- 多台设备配置了 **default-information originate always**，造成路由环路。
- 在 Stub 区域配置路由重分布，无法引入路由。

4.4.4 设置Stub、NSSA区域

配置效果

- 将末梢区域设置为 Stub 区域，有助于减少路由信息的交互，缩小路由表规模，增强路由稳定性。

注意事项

- 必须配置 OSPF 基本功能
- 骨干区域、传输区域不能设置为 Stub 区域、NSSA 区域。
- Stub 区域内路由器不能引入外部路由，NSSA 区域内路由器可以引入外部路由。

配置方法

配置 Stub 区域

- 可选配置。若希望减小区域内路由器的路由表规模，需要配置。
- 必须在该区域内的所有路由器上配置该区域为 Stub 区域。

配置 Nssa 区域

- 可选配置。若希望减小区域内路由器的路由表规模，又希望本区域可以引入 OSPF 域外路由，需要配置。
- 必须在该区域内的所有路由器上配置该区域为 NSSA 区域。

检验方法

检验 Stub 区域

- 在 Stub 区域内部的路由器上，通过 **show ip route** 查看路由器没有加载外部明细路由。

检验 Nssa 区域

- 在 NSSA 区域内部的路由器上，通过 **show ip ospf database** 查看引入的外部路由生成了 7 类 LSA。
- 在骨干区域内部的路由器上，通过 **show ip route** 查看路由器加载了从 NSSA 区域引入的外部路由。

相关命令

配置 Stub 区域

【命令格式】 **area area-id stub [no-summary]**

【参数说明】 **area-id** : Stub 区域的区域号。

no-summary : 阻止 ABR 往残域发送网络摘要链路通告，此时残域可称为全残域。只有 ABR 才需要配置该

参数。

【配置模式】 OSPF 路由进程配置模式

【使用指导】 OSPF 残域内所有的路由设备都必须执行 `area stub` 命令。ABR 只向残域发送三种链路状态描述(LSA)：1) 类型 1，路由设备 LSA；2) 类型 2，网络 LSA；3) 类型 3，网络摘要 LSA。从路由表角度看，残域中的路由设备只能学习到 OSPF 路由域内部的路由，包括由 ABR 产生的内部缺省路由。残域的路由设备不能学习到 OSPF 路由域外部的路由。

如果需要配置全残域，只需在 ABR 上执行 `area stub` 命令时，选择 `no-summary` 关键字。全残域的路由设备只能学到本区域内的路由，以及 ABR 产生的内部缺省路由。

将一个 OSPF 区域设置为残域有两条命令 `area stub` 和 `area default-cost`。所有连接到残域的路由设备上，都必须配置 `area stub` 命令，但 `area default-cost` 命令只需要在 Stub 区域的 ABR 上配置。`area default-cost` 命令定义了内部缺省路由的初始花费（即量度）。

配置 Nssa 区域

【命令格式】 `area area-id nssa [no-redistribution] [default-information-originate [metric value] [metric-type type]] [no-summary] [translator [stability-interval seconds | always]]`

【参数说明】 `area-id`：NSSA 区域的区域号。

no-redistribution：当该路由设备是一个 NSSA ABR 时，如果用户只想用重分布命令将路由信息导入普通区间而不让它进入 NSSA 区间时，可以选上该选项。

default-information-originate：产生缺省的类型 7 的 LSA 进入 NSSA 区间，该选项只在 NSSA ABR 或者 ASBR 上有效。

metric value：设置产生的缺省 LSA 的度量值。metric 取值范围 0-16777214，缺省值 1。

metric-type type：设置产生的缺省 LSA 的路由类型。type 取值 1 或 2。1 表示 N-1；2 表示 N-2。缺省值 2。

no-summary：阻止 NSSA 的边界路由设备(ABR)发送 summary LSAs (Type-3 LSA)。

translator：配置 NSSA 的边界路由设备(ABR)转换者。

stability-interval seconds：配置 NSSA 的边界路由设备(ABR)由转换者变为非转换者后保持的稳定间隔。单位：秒。缺省值为 40。取值范围：0- 2147483647。

always：配置当前 NSSA 的边界路由设备(ABR)总是作为转换者角色。缺省为备选转换者。

【配置模式】 OSPF 路由进程配置模式

【使用指导】 参数 **default-information-originate** 用来产生默认的 Type-7 LSA，该选项在 NSSA 的 ABR 和 ASBR 上有些差别，在 ABR 上无论路由表中是否存在缺省路由，都会产生 Type-7 LSA 缺省路由，在 ASBR 上(同时也不是 ABR)当路由表中存在缺省路由，才会产生 Type-7 LSA 缺省路由。

参数 `no-redistribution` 在 ASBR 上使得 OSPF 通过 **redistribute** 命令引入的其它外部路由不发布到 NSSA 区。该选项通常用于 NSSA 的路由设备既是 ASBR 又是 ABR 的时候，它可以阻止外部路由信息进入 NSSA。

为了进一步减少发送到 NSSA 区域中的链路状态广播(LSA)的数量，可以在 ABR 上配置 `no-summary` 属性，禁止 ABR 向 NSSA 区域内发送 summary LSAs (Type-3 LSA)。

另外 **area default-cost** 用在连接在该 NSSA 区域的 ABR/ASBR 上。该命令配置 ABR/ASBR 发送到 NSSA 区域的缺省路由的花费值。缺省情况下，发送到 NSSA 缺省路由的花费值为 1。

如果一个 NSSA 区域有两个（或者两个以上）ABR，缺省情况下，选举路由设备标识最大的 ABR 设备作为 Type-7 转 Type-5 的转换者；如果希望当前设备总是作为 Type-7 转 Type-5 的转换者 ABR，使用 `translator always` 参数。

如果当前设备的转换者角色被其它 ABR 取代，在 `stability-interval` 时间内将继续保持转换能力，如果在这段时

间内没有重新成为转换者，stability-interval 到期后由 Type-7 转换为 Type-5 的 LSA 将从自治域中被清除。为防止路由环路，由 Type-7 汇聚生成的 Type-5 LSA 将在当前设备失去转换者角色后立即被清除，而无需等待 stability-interval 到期。

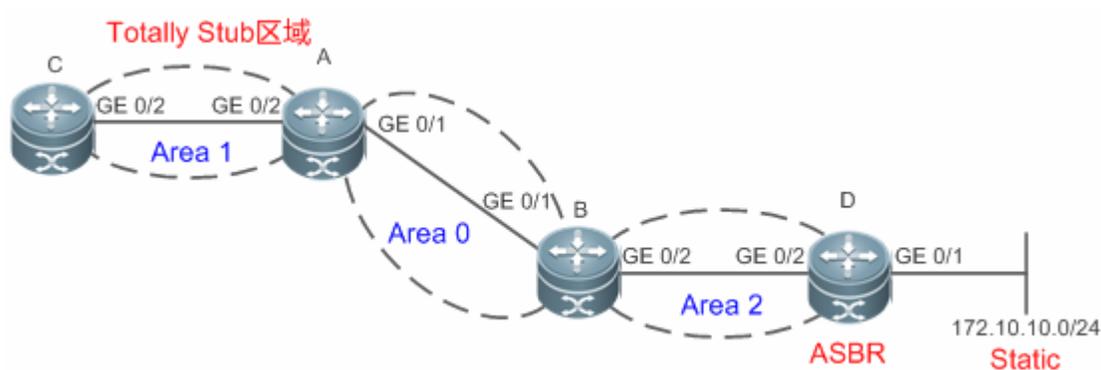
同一 NSSA 区域中，建议仅在一台 ABR 配置 translator always 参数。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见“配置 OSPF 基本功能”的介绍。

配置 Stub 区域

【网络环境】



【注释】 接口 IP 地址：

A : GE 0/1 192.168.1.1 GE 0/2 192.168.2.1
 B : GE 0/1 192.168.1.2 GE 0/2 192.168.3.1
 C : GE 0/2 192.168.2.2
 D : GE 0/1 192.168.6.2 GE 0/2 192.168.3.2

- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 路由器 D 上配置引入外部静态路由。
 - 路由器 A、C 上区域 1 配置成 Stub 类型。

```
D# configure terminal
D(config)# router ospf 1
D(config-router)# redistribute static subnets

A# configure terminal
A(config)# router ospf 1
A(config-router)# area 1 stub no-summary

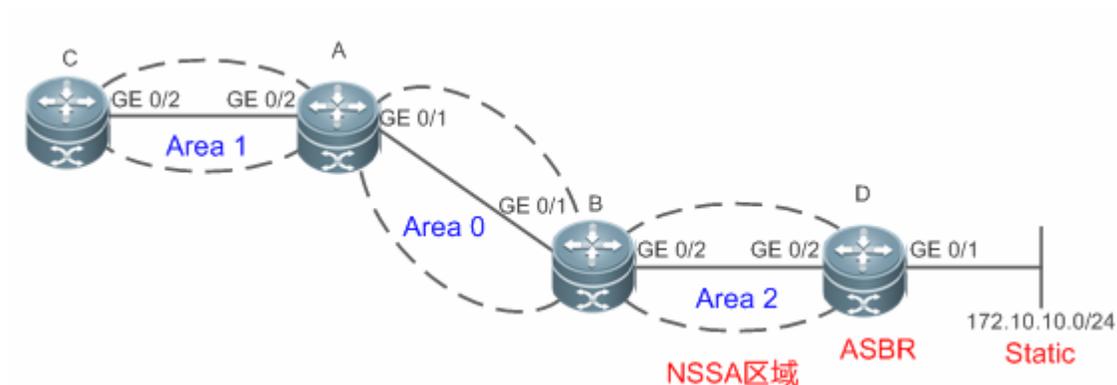
C# configure terminal
C(config)# router ospf 1
C(config-router)# area 1 stub
```

【检验方法】 在路由器 C 上使用 **show ip route ospf** 查看路由表，只有一条默认区间路由，没有 D 引入的外部静态路由

```
C# show ip route ospf
0*IA 0.0.0.0/0 [110/2] via 192.168.2.1, 00:30:53, GigabitEthernet 0/2
```

配置 NSSA 区域。

【网络环境】



【注释】 接口 IP 地址：

A : GE 0/1 192.168.1.1 GE 0/2 192.168.2.1
 B : GE 0/1 192.168.1.2 GE 0/2 192.168.3.1
 C : GE 0/2 192.168.2.2
 D : GE 0/1 192.168.6.2 GE 0/2 192.168.3.2

【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 路由器 D 上配置引入外部静态路由。
- 路由器 B、D 上区域 2 配置成 NSSA。

B

```
B# configure terminal
B(config)# router ospf 1
B(config-router)# area 2 nssa
```

D

```
D# configure terminal
D(config)# ip route 172.10.10.0 255.255.255.0 192.168.6.2
D(config)# router ospf 1
D(config-router)# redistribute static subnets
D(config-router)# area 2 nssa
```

【检验方法】

- 路由器 D 上查看生成 172.10.10.0/24 的 7 类 LSA
- 路由器 B 上查看 172.10.10.0/24 同时存在 5 类和 7 类 LSA
- 路由器 B 上查看生成了 172.10.10.0/24 的 N2 类型的路由。

D

```
D# show ip ospf database nssa-external
      OSPF Router with ID (192.168.6.2) (Process ID 1)
          NSSA-external Link States (Area 0.0.0.1 [NSSA])

LS age: 61
```

```
Options: 0x8 (-|-|-|-|N/P|-|-|-)
LS Type: AS-NSSA-LSA
Link State ID: 172.10.10.0 (External Network Number For NSSA)
Advertising Router: 192.168.6.2
LS Seq Number: 80000001
Checksum: 0xc8f8
Length: 36
Network Mask: /24
    Metric Type: 2 (Larger than any link state path)
    TOS: 0
    Metric: 20
    NSSA: Forward Address: 192.168.6.2
    External Route Tag: 0
```

```
B B# show ip ospf database nssa-external
      OSPF Router with ID (192.168.3.1) (Process ID 1)
      NSSA-external Link States (Area 0.0.0.1 [NSSA])
LS age: 314
Options: 0x8 (-|-|-|-|N/P|-|-|-)
LS Type: AS-NSSA-LSA
Link State ID: 172.10.10.0 (External Network Number For NSSA)
Advertising Router: 192.168.6.2
LS Seq Number: 80000001
Checksum: 0xc8f8
Length: 36
Network Mask: /24
    Metric Type: 2 (Larger than any link state path)
    TOS: 0
    Metric: 20
    NSSA: Forward Address: 192.168.6.2
    External Route Tag: 0
B# show ip ospf database external
      OSPF Router with ID (192.168.3.1) (Process ID 1)
      AS External Link States
LS age: 875
Options: 0x2 (-|-|-|-|-|E|-)
LS Type: AS-external-LSA
Link State ID: 172.10.10.0 (External Network Number)
Advertising Router: 192.168.3.1
LS Seq Number: 80000001
Checksum: 0xd0d3
Length: 36
```

```
Network Mask: /24
Metric Type: 2 (Larger than any link state path)
TOS: 0
Metric: 20
Forward Address: 192.168.6.2
External Route Tag: 0

B# show ip route ospf
0 N2 172.10.10.0/24 [110/20] via 192.168.3.2, 00:06:53, GigabitEthernet 0/2
```

常见配置错误

- 同一区域中的设备，区域类型配置不一致
- 在 Stub 区域中配置路由重分布，将无法加入外部路由

4.4.5 配置路由汇聚

配置效果

- 明细路由汇聚，减少路由信息的交互，缩小路由表规模，增强路由稳定性。
- 屏蔽或者过滤路由。

注意事项

- 必须配置 OSPF 基本功能
- 汇总路由的地址范围有可能超出路由表中实际的网络范围，如果数据发往汇总范围内不存在的网络，将有可能发生路由环路或加重路由设备的处理负担。为阻止这种情况的发生，需要向路由表中添加一条 discard 路由屏蔽或者过滤路由。

配置方法

配置区间路由汇聚

- 可选配置。对 OSPF 区域的路由进行汇聚时，需要配置。
- 若无特殊要求，应在被汇聚路由所在区域的 ABR 上配置。

配置外部路由汇聚

- 可选配置。对域外的路由进行汇聚时，需要配置。
- 若无特殊要求，应在引入被汇聚路由的 ASBR 上配置。

检验方法

通过 `show ip route ospf` 查看明细路由不存在，只有汇聚后的路由。

相关命令

配置区间路由汇聚

【命令格式】 `area area-id range ip-address net-mask [advertise | not-advertise] [cost cost]`

【参数说明】 `area-id`：指定要注入汇聚路由的 OSPF 区域号。区域号可以是一个十进制整数，也可以是一个 IP 地址。

`ip-address net-mask`：定义汇聚路由的网段。

`advertise | not-advertise`：定义是否要公布该汇聚路由。

`cost cost`：设置汇聚路由的度量值，取值范围：0-16777215。

【命令模式】 OSPF 路由进程配置模式

【使用指导】 该命令只能在 ABR 上执行有效，用来将一个区域的多个路由合并或汇聚成一条路由，然后通告到其它区域去。路由信息的合并行为只发生在区域边界，区域内部路由设备上看到的都是具体的路由信息，但其它区域的设备只能看到一条汇聚路由。并且可以通过设置 `advertise` 和 `not-advertise` 选项来决定是否公布这个汇聚范围，起到屏蔽和过滤的作用，缺省是向外公布。可以通过 `cost` 选项来设置汇聚路由的度量值。可以定义多个区域路由汇聚命令，这样整个 OSPF 路由域的路由将得到简化，特别是网络规模大时，会提高网络转发性能。

当配置了多条汇总，并且项目直接有包含关系的时候，路由是根据最长匹配原则确定被汇总的 `area range` 的。

配置外部路由汇聚

【命令格式】 `summary-address ip-address net-mask [not-advertise | tag value]`

【参数说明】 `ip-address`：汇聚路由的 IP 地址。

`net-mask`：汇聚路由的网络掩码。

`not-advertise`：不公告汇聚的路由，若未配置则为公告。

`tag value`：设置汇聚路由的 tag 值，取值范围：0-4294967295。

【命令模式】 OSPF 路由进程配置模式

【使用指导】 当路由从其它路由进程重新分布，注入 OSPF 路由进程时，每条路由均分别以一个外部链路状态的方式通告给 OSPF 路由设备。如果注入的路由是一个连续的地址空间，自治域边界路由设备可以只通告一个汇聚路由，从而大大减小路由表的规模。

`area range` 汇聚的是 OSPF 区域之间路由，而 `summary-address` 汇聚的是 OSPF 路由域的外部路由。

`summary-address` 命令配置在 NSSA 的 ABR 转换者上时，对重分布的路由及 LSA 从 7 类转 5 类的路由进行汇聚，而配置在 ASBR（非 NSSA 的 ABR 转换者）时，只对重分布的路由进行汇聚。

配置 Discard Route

【命令格式】 `discard-route { internal | external }`

【参数说明】 `internal`：允许添加区间路由汇聚命令 `area range` 生成的 discard 路由。

`external`：允许添加外部路由汇聚命令 `summary-address` 生成的 discard 路由。

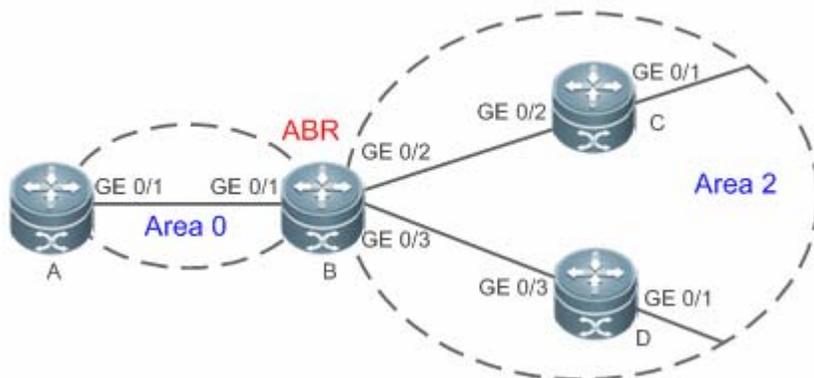
【命令模式】 OSPF 路由进程配置模式

【使用指导】 在路由汇聚时，汇聚后的范围有可能超出路由表中实际的网络范围，如果数据发往汇聚范围内不存在的网络，将有可能发生路由环路或加重路由设备的处理负担。为阻止这种情况的发生，需要在 ABR 或 ASBR 添加一条 discard 路由到路由表。该路由自动生成，并且不会被传播。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

【配置方法】



【注释】 接口 IP 地址：

A : GE0/1 192.168.1.1

B : GE0/1 192.168.1.2 GE0/2 172.16.2.1 GE0/3 172.16.3.1

C : GE0/2 172.16.2.2 GE0/1 172.16.4.2

D : GE0/2 172.16.3.2 GE0/1 172.16.5.2

- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 路由器 B 上配置将区域 2 路由汇聚。

B

```
B# configure terminal
B(config)# router ospf 1
B(config-router)# area 2 range 172.16.0.0 255.255.0.0
```

【检验方法】 路由器 A 查看路由表，生成 172.16.0.0/16 表项。

A

```
A# show ip route ospf
0 IA 172.16.0.0/16 [110/2] via 192.168.1.2, 00:01:04, GigabitEthernet 0/1
```

常见配置错误

- 在非 ABR 设备上配置 **area range**，无法进行区间路由汇聚。

4.4.6 配置路由过滤

配置效果

- 被过滤掉的路由不能被加载到路由表或被传播给邻居，网络内用户将无法访问指定的目的网络。

注意事项

- 必须配置 OSPF 基本功能。
- 通过 **distribute-list in** 过滤只影响本地路由转发，不影响 LSA 计算路由。因此，如果确定需要在 ABR 上配置过滤，则由于 LSA 仍然可以计算出路由，仍会生产 Type-3 LSA 往其他区域通告，这将导致黑洞路由存在。这时可以在 ABR 上使用 **area filter-list** 或 **area range** 带 **not-advertise** 选项来避免产生黑洞路由。

配置方法

配置区间路由过滤

- 可选配置。若限制用户访问 OSPF 某个区域的网络，建议配置。
- 若无特殊要求，应在被过滤路由所在区域的 ABR 上配置。

配置重分布路由过滤

- 可选配置。ASBR 引入外部路由时，如果需要对引入的外部路由进行过滤，必须配置。
- 若无特殊要求，应在引入被过滤外部路由的 ASBR 上配置。

配置对学习的路由过滤

- 可选配置。若需要限制用户访问指定的目的网络，必须配置。
- 若无特殊要求，应在需要过滤的路由器上配置该功能。

检验方法

- 通过 **show ip route** 查看路由器没有加载被过滤的路由。
- 通过 **ping** 命令确认指定的目的网络无法访问。

相关命令

配置被动接口

【命令格式】 **passive-interface** { **default** | *interface-type interface-number* | *interface-type interface-number ip-address* }

【参数说明】

interface-type interface-number : 要配置为被动接口的接口。

default : 该选项表示将所有接口设置为被动接口。

interface-type interface-number ip-address : 指定接口下的地址为被动地址。

【命令模式】

OSPF 路由进程配置模式

【使用指导】

为了防止网络中的其他路由设备动态的学习到本路由设备的路由信息,可以将本路由设备的指定网络接口设为被动接口或者指定网络接口下的 IP 地址为被动地址。

配置lsa更新报文过滤功能

【命令格式】

ip ospf database-filter all out

【参数说明】

-

【命令模式】

接口配置模式

【使用指导】

要阻止在某个接口上发送 LSA 的更新报文时,可以在该接口上打开该功能,打开该功能后,本设备将不向邻居通告 LSA 更新报文,但仍然向邻居建立邻接关系并接收邻居发来的 LSA 信息。

配置区间路由过滤

【命令格式】

area area-id filter-list {accessacl-name| prefix prefix-name} {in | out}

【参数说明】

area-id : 区域号。

access acl-name : 关联 ACL。

prefix prefix-name : 关联 prefix list。

in | out : 设置该条件应用于进入|发出该区的路由。

【命令模式】

OSPF 路由进程配置模式

【使用指导】

该命令只能在 ABR (区域边界路由设备) 上进行配置。

当需要在该 ABR 上配置不同区域之间路由学习的过滤条件时,使用本命令。

配置重分布路由过滤

【命令格式】

distribute-list { [access-list-number | name] | prefix prefix-list-name } out [connected [area-tag] | ospf process-id | rip | static]

【参数说明】

access-list-number | name : 使用 acl 过滤规则。

prefix prefix-list-name : 使用 prefix-list 过滤规则。

connected [area-tag] | ospf process-id | rip | static : 待过滤路由的来源。

【命令模式】

OSPF 路由进程配置模式

【使用指导】

distribute-list out 与 **redistribute route-map** 命令类似,对从其他协议重分布到 OSPF 中的路由进行过滤,但它本身不执行路由重分布,一般与 **redistribute** 命令配合使用。ACL 过滤规则和 prefix-list 过滤规则不能在该配置中共存,即若对某一来源的路由使用 ACL 过滤规则,则不能再配置 prefix-list 进行过滤。

配置对学习的路由过滤

【命令格式】

distribute-list { [access-list-number | name] | prefix prefix-list-name [gateway prefix-list-name] | route-map route-map-name } in [interface-type interface-number]

【参数说明】

access-list-number | name : 使用 acl 过滤规则。

gateway prefix-list-name : 使用 gateway 过滤规则。

prefix *prefix-list-name* : 使用 prefix-list 过滤规则。

route-map *route-map-name* : 使用 route-map 过滤规则。

interface-type *interface-number* : 配置只针对某个接口进行 LSA 路由过滤。

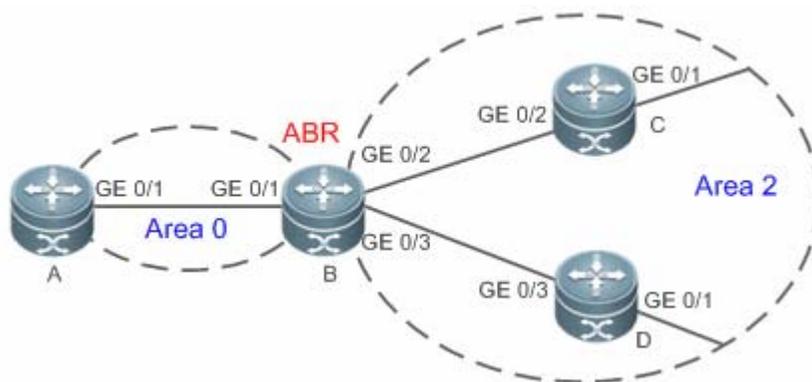
【命令模式】 OSPF 路由进程配置模式

【使用指导】 对收到的 LSA 计算出的路由进行过滤，满足过滤条件的路由才能用于转发。它并不会影响链路状态数据库，对邻居的路由表也不会产生任何影响。ACL 过滤规则、prefix-list 过滤规则和 route-map 过滤规则不能在该配置中共存，即若对指定接口使用 ACL 过滤规则，则不能再对该接口配置 prefix-list 进行过滤或 route-map 过滤规则。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见 1.4.1.配置 OSPF 基本功能的介绍。

【网络环境】



【注释】 接口 IP 地址：

A : GE0/1 192.168.1.1

B : GE0/1 192.168.1.2 GE0/2 172.16.2.1 GE0/3 172.16.3.1

C : GE0/2 172.16.2.2 GE0/3 172.16.4.2

D : GE0/2 172.16.3.2 GE0/3 172.16.5.2

【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 路由器 A 配置路由过滤。

```
A
A# configure terminal
A(config)# access-list 3 permit host 172.16.5.0
A(config)# router ospf 1
A(config-router)# distribute-list 3 in GigabitEthernet 0/1
```

【检验方法】

- 路由器 A 检查路由表，只有加载 172.16.5.0/24 表项。

```
A
A# show ip route ospf
0 172.16.5.0/24 [110/2] via 192.168.1.2, 10:39:40, GigabitEthernet 0/1
```

常见配置错误

- 通过 **distribute-list in** 过滤只影响本地路由转发，不影响 LSA 计算路由。因此，如果确定需要在 ABR 上配置过滤，则由于 LSA 仍然可以计算出路由，仍会生产 Type-3 LSA 往其他区域通告，这将导致黑洞路由存在。

4.4.7 修改路由代价、管理距离

配置效果

- 改变 OSPF 路由，使流量经过或避开某些节点。
- 改变路由器选择各类路由的顺序，使 OSPF 路由的优先级提高或降低。

注意事项

- 必须配置 OSPF 基本功能
- 如果用命令 **ip ospf cost** 设置了接口的 cost，则将覆盖自动代价算出的 cost。

配置方法

配置参考带宽

- 可选配置。
- 路由器连接不同带宽的线路，若希望优先选择带宽大的线路，建议配置。

配置接口 cost 值

- 可选配置。
- 路由器连接多条线路，若希望手动指定优先的线路，建议配置。

配置重分布缺省度量

- 可选配置。
- 在 ASBR 上引入域外路由时，若需要指定域外路由的代价，必须配置。

配置最大度量值

- 可选配置。。
- 路由器重启过程中以及重启刚完成的一段时间内，路由器可能还不稳定，用户的数据流量转发不希望经过该路由器，此时建议配置。

配置管理距离

- 可选配置。
- 在同时运行多种单播路由协议的路由器上，若希望改变 OSPF 路由的优先级，必须配置。

检验方法

- 通过 **show ip ospf interface** 查看接口的代价正确
- 通过 **show ip route** 查看 ASBR 引入的外部路由的代价正确
- 重启路由器，在指定的一段时间内，数据流量转发不经过重启的路由器

相关命令

配置参考带宽

【命令格式】 **auto-costreference-bandwidth** *ref-bw*

【参数说明】

ref-bw : 参考带宽值。以 Mbps 为单位， 范围: 1-4294967。

【命令模式】 OSPF 路由进程配置模式

【使用指导】 缺省情况下，OSPF 接口 cost 值等于自动代价参考值/接口带宽。

使用 **auto-cost** 命令自动代价参考值。缺省值 100Mbps。

使用 **bandwidth** 命令设置接口带宽。

几种典型线路的 OSPF 接口 cost 缺省值为：

64K 串行线路，cost 为 1562；

E1 线路，cost 为 48；

10M 以太网，cost 为 10；

100M 以太网，cost 为 1。

如果用命令 **ip ospf cost** 设置了接口的 cost，则将覆盖自动代价算出的 cost。

配置接口 cost 值

【命令格式】 **ip ospf cost** *cost*

【参数说明】 *cost* : OSPF 接口的花费值，取值范围：0-65535。

【命令模式】 接口配置模式

【使用指导】 缺省情况下，OSPF 接口 cost 值等于自动代价参考值/接口带宽。

使用 **auto-cost** 命令自动代价参考值。缺省值 100Mbps。

使用 **bandwidth** 命令设置接口带宽。

几种典型线路的 OSPF 接口 cost 缺省值为：

64K 串行线路，cost 为 1562；

E1 线路，cost 为 48；

10M 以太网，cost 为 10；

100M 以太网，cost 为 1。

如果用命令 **ip ospf cost** 设置了接口的 cost，则将覆盖自动代价算出的 cost

配置 Stub/NSSA 区域缺省路由的 cost 值

【命令格式】 **area** *area-id* **default-cost** *cost*

- 【参数说明】 *area-id* : Stub 区域或 NSSA 的区域号。
cost : 注入 Stub 区域或者 NSSA 的缺省汇聚路由的花费值, 取值范围: 0-16777215。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 该命令只有在 Stub 区域的 ABR 或 NSSA 区域的 ABR/ASBR 上配置才能生效。
 Stub 区域的 ABR 或 NSSA 区域的 ABR/ASBR 设备允许在 Stub 或者 NSSA 区域中通告一个表示缺省路由的 LSA, 用户可以通过 **area default-cost** 命令修改通告的 LSA 的花费值。

配置重分布缺省度量

- 【命令格式】 **default-metric** *metric*
- 【参数说明】 *metric* : 定义 OSPF 重分布路由的缺省度量值, 取值范围: 1-16777214。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 **default-metric** 命令必须与 **redistribute** 路由进程配置命令配合使用, 可以修改所有重分布路由的初始化度量值。
default-metric 命令的设置, 对以 **default-information originate** 方式注入 OSPF 路由域的外部路由不起作用。

配置最大度量值

- 【命令格式】 **max-metric router-lsa** [**external-lsa** [*max-metric-value*]] [**include-stub**] [**on-startup** [*seconds*]] [**summary-lsa** [*max-metric-value*]]
- 【参数说明】 **router-lsa** : 配置 Router LSA 中的非 stub links 的度量为最大值 (0xFFFF)。
external-lsa : 配置路由设备使用最大度量值取代外部 LSA (包括 Type-5 和 Type-7) 的度量值。
max-metric-value : LSA 的最大度量值。缺省值为 16711680。取值范围: 1-16777215。
include-stub : 配置路由设备通告 Router LSA 中的 stub links 为最大度量值。
on-startup : 配置路由设备在启动时通告最大度量。
seconds : 通告最大度量的指定时间间隔。缺省值为 600s。取值范围: 5-86400。
summary-lsa : 配置路由设备使用最大度量值取代汇总 LSA (包括 Type-3 和 Type-4) 的度量值。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 使能命令 **max-metric router-lsa** 后, 路由设备生成的 Router LSA 中的非 stub links 的度量将被设置为最大值 (0xFFFF)。在用户取消该配置, 或者定时器到期后, 恢复链路的正常度量。
 缺省情况下, 当使能该命令, stub links 仍将通告普通度量值, 即出接口代价值, 如果配置了 **include-stub** 参数, stub links 将通告为最大度量值;
 作为 ABR 设备, 如果不希望传输区间流量则使用 **summary-lsa** 参数设置汇总 LSA 为最大度量值;
 作为 ASBR 设备, 如果不希望传输外部流量则使用 **external-lsa** 参数设置外部 LSA 为最大度量值。
max-metric router-lsa 命令通常在以下场景使用:
 重启设备。当设备重启, IGP 协议通常更快收敛, 其它设备试图通过新启动的设备转发流量。
 将设备加入网络但不用于传输流量。将设备加入网络, 如果存在备选路径, 则当前设备不用于传输流量; 如果不存在备选路径, 则当前设备仍然用于传输流量。
 将设备优雅的从网络中移除。通过使能该命令, 当前设备通告所有最大度量值, 这样就能够设备关闭前, 使网络中的其它设备选择备份路径传输流量。
 在旧版本 OSPF 实现 (RFC 1247 或更早的版本) 中, LSA 中最大度量 (0xFFFF) 的 links 不参与 SPF 计算,

即没有流量发往生成这些 LSAs 的路由设备。

配置兼容 RFC1583

【命令格式】 **commpatible rfc1583**

【参数说明】 -

【命令模式】 OSPF 路由进程配置模式

【使用指导】 当存在多条到达 ASBR 或外部路由转发地址的路径时，RFC1583 与 RFC2328 定义了不同的选路规则。配置兼容 RFC1583 时，优先选择骨干区域的区域内路径或区域间路径；不兼容 RFC1583 时，优先选择非骨干区域的区域内路径。

配置管理距离

【命令格式】 **distance { distance | ospf { [intra-area distance] [inter-area distance] [external distance] }**

【参数说明】 *distance*：设置路由管理距离，取值范围：1-255。

intra-area distance：设置区内路由管理距离，取值范围：1-255。

inter-area distance：设置区间路由管理距离，取值范围：1-255。

external distance：设置外部路由管理距离，取值范围：1-255。

【命令模式】 OSPF 路由进程配置模式

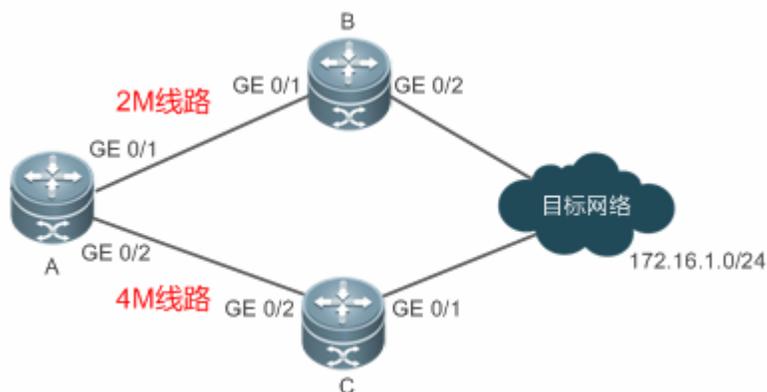
【使用指导】 若要指定 OSPF 不同类型路由对应不同管理距离时，用此命令。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

配置接口 cost 值

【网络环境】



【注释】 接口 IP 地址：

A : GE0/1 192.168.1.1 GE0/2 192.168.2.1

B : GE0/1 192.168.1.2 GE0/2 192.168.3.2

C : GE0/1 192.168.4.2 GE0/2 192.168.2.2

- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 路由器 A 配置接口代价。

```
A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip ospf cost 10
A(config)# interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)# ip ospf cost 20
```

【检验方法】 路由器 A 检查路由表，172.16.1.0/24 的最优路径下一跳是路由器 B。

```
A
A# show ip route ospf
0 E2 172.16.1.0/0 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
```

常见配置错误

- ip ospf cost 配置为 0，路由计算可能会出错，比如计算出路由环路；

4.4.8 启动认证功能

配置效果

- 所有接入 OSPF 网络的路由器都必须经过认证，保证 OSPF 协议的稳定性和抗入侵性。

注意事项

- 必须配置 OSPF 基本功能
- 在区域上配置认证，则在属于此区域的所有接口上生效。
- 区域和接口同时配置认证，接口配置优先生效。

配置方法

配置区域认证类型

- 可选配置。如果要求同一区域内所有接口启用相同的认证类型，建议配置。
- 若路由器接入需要认证的网络，需要配置。

配置接口认证类型

- 可选配置。如果要求同一区域内不同接口启用不同的认证类型，必须配置。
- 若路由器接入需要认证的网络，需要配置。

配置接口明文认证密钥

- 可选配置。

- 若路由器接入需要明文认证的网络，需要配置。

配置接口 md5 认证密钥

- 可选配置。MD5 认证安全性较高，建议用户使用 MD5 认证。明文认证与 MD5 认证必须配置一个。
- 若路由器接入需要 MD5 认证的网络，需要配置。

检验方法

- 路由器配置不同的认证密钥，通过 **show ip ospf neighbor** 查看 OSPF 邻居为空
- 路由器配置相同的认证密钥后，通过 **show ip ospf neighbor** 查看 OSPF 邻居正确

相关命令

配置区域认证类型

【命令格式】 **area area-id authentication [message-digest]**

【参数说明】 **area-id**：指定要启用 OSPF 认证的区域号。区域号可以是一个十进制整数值，也可以是一个 IP 地址。
message-digest：采用 MD5 (message digest 5) 认证模式。

【配置模式】 OSPF 路由配置模式

【使用指导】 RGOS 软件可以支持三种认证类型：

- 1) 类型 0，不要求认证，当没有用该命令启用 OSPF 认证，OSPF 数据包中的认证类型为 0；
- 2) 类型 1，为明文认证模式，配置该命令时，没有使用 message-digest 选项；
- 3) 类型 2，为 MD5 认证模式，配置该命令时，使用 message-digest 选项。

一个 OSPF 区域中所有的路由设备，必须采用同一种认证类型。如果启用了认证，还必须在有连接邻居的接口配置认证口令。接口配置命令 **ip ospf authentication-key** 可以配置明文认证口令；接口配置命令 **ip ospf message-digest-key** 可以配置 MD5 认证口令。

配置接口认证类型

【命令格式】 **ip ospf authentication [message-digest | null]**

【参数说明】 **message-digest**：该选项表示在该接口上进行 MD5 加密认证。
null：该选项表示不进行认证。

【配置模式】 接口配置模式

【使用指导】 如果命令后面不跟任何选项表示进行明文认证。注意在配置认证方式时使用 **no** 选项将认证方式恢复缺省值时，实际是否认证是由该接口所在区间上设置的认证方式决定的，如果将认证方式设置为 **null**，则强制设置为不认证，当接口上的认证方式和接口所在的区域的认证方式都设置时，优先采用接口上的认证方式。

配置接口明文认证密钥

【命令格式】 **ip ospf authentication-key [0 | 7] key**

【参数说明】 **0**：指定密钥以明文显示。

7：指定密钥以密文显示。

key：密钥，最多可以由 8 个字符组成。

【配置模式】 接口配置模式

【使用指导】 **ip ospf authentication-key** 命令配置的密钥，将被插入所有 OSPF 报头。因此密钥不一致，两台直接连接的设备是不能建立 OSPF 邻居关系，当然也不能进行路由信息的交换。

不同接口的密钥可以配置不一样，但所有连接在同一物理网段上的路由设备，必须配置一样的密钥。

一个 OSPF 区域是否启用认证，用路由进程配置命令 **area authentication** 进行配置。

在接口上也可以单独启用认证，使用接口模式下的配置命令 **ip ospf authentication** 进行配置，当区域认证和接口认证都配置时优先取接口的认证类型。

配置接口 md5 认证密钥

【命令格式】 **ip ospf message-digest-key** *key-id* **md5** [0 | 7] *key*

【参数说明】 *key-id* : 密钥标识符，从 1 到 255。

0 : 指定密钥以明文显示。

7 : 指定密钥以密文显示。

Key : 密钥，最多可以由 16 个字符组成。

【配置模式】 接口配置模式

【使用指导】 **ip ospf message-digest-key** 命令配置的密钥，将被插入所有 OSPF 报头。因此密钥不一致，两台直接连接的设备是不能建立 OSPF 邻居关系的，当然也不能进行路由信息的交换。

不同接口的密钥可以配置不一样，但所有连接在同一物理网段上的路由设备，必须配置一样的密钥。邻居路由设备相同的密钥标识对应的密钥必须一样。

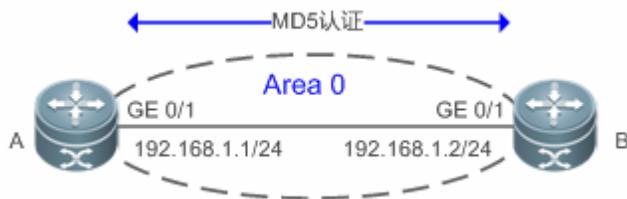
一个 OSPF 区域是否启用认证，使用路由进程配置命令 **area authentication** 进行配置。在接口上也可以单独启用认证，使用接口模式下的配置命令 **ip ospf authentication** 进行配置，当区域认证和接口认证都配置时优先取接口的认证类型。

RGOS 软件支持平滑地进行 MD5 认证密钥修改，修改应该遵循一个先加后删除的原则。当增加一个路由设备的 OSPF MD5 认证密钥时，该路由设备会认为其它路由设备还没有用新的密钥，因此会分别用不同的密钥发送多份 OSPF 报文，直到确认邻居已经配置了新的密钥。等所有路由设备配置了新的密钥后，就可以删除旧的密钥了。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见 1.4.1.配置 OSPF 基本功能的介绍。

【网络环境】



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 所有路由器配置认证类型和 MD5 认证密钥。

```
A      A# configure terminal
        A(config)# router ospf 1
        A(config-router)# area 0 authentication message-digest
        A(config-router)# exit
        A(config)# interface GigabitEthernet 0/1
        A(config-if-GigabitEthernet 0/1)# ip ospf message-digest-key 1 md5 hello

B      B# configure terminal
        B(config)# router ospf 1
        B(config-router)# area 0 authentication message-digest
        B(config-router)# exit
        B(config)# interface GigabitEthernet 0/3
        B(config-if-GigabitEthernet 0/3)# ip ospf message-digest-key 1 md5 hello
```

【检验方法】 路由器 A、B 上查看 OSPF 邻居状态正确。

```
A      A# show ip ospf neighbor
        OSPF process 1, 1 Neighbors, 1 is Full:
        Neighbor ID    Pri  State           Dead Time  Address      Interface
        192.168.1.2    1    Full/DR         00:00:32  192.168.1.2 GigabitEthernet 0/1

B      A# show ip ospf neighbor
        OSPF process 1, 1 Neighbors, 1 is Full:
        Neighbor ID    Pri  State           Dead Time  Address      Interface
        192.168.1.1    1    Full/DR         00:00:32  192.168.1.1 GigabitEthernet 0/1
```

常见配置错误

- 配置的认证方式不一致。
- 配置的认证密钥不一致。

4.4.9 启动溢出功能

配置效果

- 内存不足时路由器不再加载新的路由。
- 数据库达到上限后路由器不再加载新的路由。

注意事项

- 必须配置 OSPF 基本功能

- 进入 OVERFLOW 状态后,可以通过使用 `clear ip ospf process` 命令,或者关闭再重新启动 OSPF 协议的方式退出 OVERFLOW 状态。

配置方法

配置内存 OVERFLOW

- 可选配置。
- 若域内有大量路由,可能导致路由器内存不足,建议配置。

配置数据库 OVERFLOW

- 可选配置。
- 若域内有大量路由,可能导致路由器内存不足,建议配置。

配置外部 Isa 数据库 OVERFLOW

- 可选配置。
- 若 ASBR 引入了大量外部路由,可能导致路由器内存不足,建议配置

检验方法

- 内存不足后,网络增加新的路由,通过 `show ip route` 查看新的路由没有加载
- 数据库达到上限后,网络增加新的外部路由,通过 `show ip route` 查看新的路由没有加载

相关命令

配置内存 OVERFLOW

【命令格式】 **overflow memory-lack**

【参数说明】 -

【配置模式】 OSPF 路由配置模式

【使用指导】 OSPF 进入 OVERFLOW 状态主要是丢弃新学到外部路由,该行为能够有效保证内存不再增长。

开启该功能,在 OSPF 进入 OVERFLOW 状态丢弃新学到的外部路由,可能会引起整网出现路由环路的现象。为了减少出现这种现象的产生,OSPF 会生成一条指向 NULL 口的缺省路由,从而降低这种现象产生的概率,该路由在 OVERFLOW 状态下将一直存在。

用户可以使用 `clear ip ospf process` 命令重新复位 OSPF 运行,清除 OSPF OVERFLOW 状态。使用 `no` 选项不允许 OSPF 在内存不足时进入 OVERFLOW 状态,这可能会造成内存资源继续消耗,在内存进一步消耗到一定程度,OSPF 实例将停止运行,并删除所有学到的路由。

配置数据库 OVERFLOW

【命令格式】 **overflow database number [hard | soft]**

【参数说明】 *number*: LSA 最大数目,取值范围<1-4294967294>。

hard : LSA 超过该数目时，关闭该 OSPF 实例。

soft : LSA 超过该数目时，发出警告。

【配置模式】 OSPF 路由配置模式

【使用指导】 若超过指定数目需要关闭 OSPF 时，用 hard 参数，若不需关闭 OSPF，只需警告时，用 soft 参数。

配置外部 Lsa 数据库 OVERFLOW

【命令格式】 **overflow database external max-dbsize wait-time**

【参数说明】 **max-dbsize** : external Lsa 的最大数目（同一 AS 内所有路由设备该值需相同），取值范围：0-2147483647。

wait-time : 路由设备从 overflow 状态尝试回复普通状态所需等待的时间，取值范围：0-2147483647。

【配置模式】 OSPF 路由配置模式

【使用指导】 当设备的 external-LSA 数目超过配置的 **max-dbsize** 时，会进入溢出状态，该状态下，不再加载 external-LSA，并将清除本地生成的 external-LSA。当经过 **wait-time** 时间后，设备恢复正常状态，并重新加载 external-LSA。在使用该功能时必须保证 OSPF 主干区域和普通区域的所有设备配置相同的 **max-dbsize** 值，否则可能引起如下现象：

整网的链路状态数据库无法达成一致，邻居无法达到 Full 状态；

错误的路由，包括环路的情况；

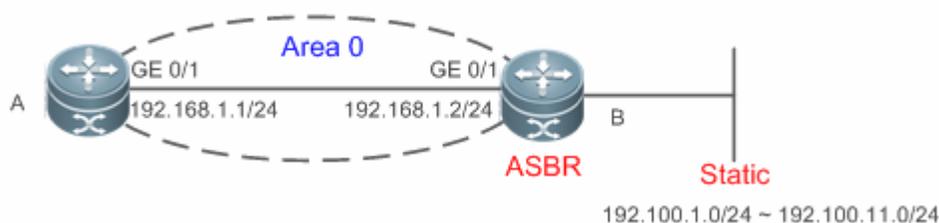
频繁的 AS-External-LSA 重传；

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

配置外部 LSA 数据库 OVERFLOW

【网络环境】



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 路由器 B 配置重分发引入外部静态路由。
 - 路由器 A 配置外部 LSA 最大值。

```

B
B# configure terminal
B(config)# router ospf 1
B(config-router)# redistribute static subnets

A
A# configure terminal
A(config)# router ospf 1
A(config-router)# overflow database external 10 3

```

【检验方法】 路由器 B 配置 11 条静态路由（192.100.1.0/24 ~ 192.100.11.0/24），路由器 A 上查看只加载了 10 条。

A

```
A# show ip route ospf
0 E2 192.100.1.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.2.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.3.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.4.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.5.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.6.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.7.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.8.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.9.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
0 E2 192.100.10.0/24 [110/20] via 192.168.1.2, 00:18:03, GigabitEthernet 0/1
```

常见配置错误

- 路由器配置的数据库上限数量不一致，导致 OSPF 邻居状态异常。

4.4.10 修改并发邻居数量限制

配置效果

- 控制 OSPF 邻居的并发建立，缓解设备压力

注意事项

- 必须配置 OSPF 基本功能

配置方法

配置进程并发邻居数

- 可选配置。若路由器连接了大量其他路由器，且希望 OSPF 邻居建立加快，建议配置。
- 在核心路由器上配置。

检验方法

- 通过 `show ip ospf neighbor` 查看正在并发建立的邻居个数

相关命令

配置当前进程并发邻居数

- 【命令格式】 **max-concurrent-dd** *number*
- 【参数说明】 *number*:指定并发交互的邻居最大数目，范围 1-65535。
- 【配置模式】 OSPF 路由配置模式
- 【使用指导】 当路由设备与多个邻居同时交换数据，其性能受到影响时，配置该命令，可限制每个 OSPF 实例允许同时并发交互（发起或接受）的邻居最大数目。

配置所有进程并发邻居数

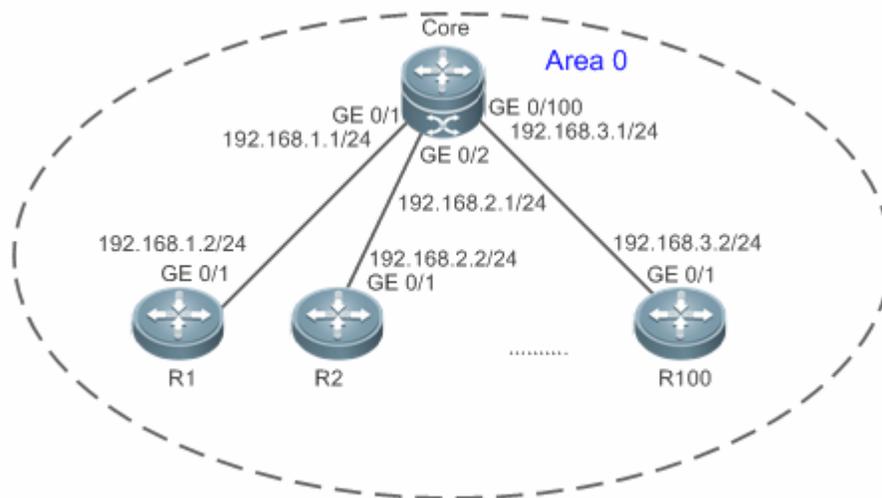
- 【命令格式】 **router ospf max-concurrent-dd** *number*
- 【参数说明】 *number*:指定并发交互的邻居最大数目，范围 1-65535
- 【配置模式】 全局配置模式
- 【使用指导】 当路由设备与多个邻居同时交换数据，其性能受到影响时，配置该命令，可限制所有 OSPF 实例允许同时并发交互（发起或接受）的邻居最大数目。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

配置进程并发邻居数

【网络环境】



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 核心路由器 Core 配置最大邻居并发数为 4。

Core

```
Core# configure terminal
Core(config)# router ospf max-concurrent-dd 4
```

【检验方法】 路由器 Core 上查看邻居状态，同时在建立的邻居最多 8 个。

常见配置错误

-

4.4.11 关闭源地址校验

配置效果

- 相邻路由器的接口 IP 地址不同网段，也能提供单播路由服务

注意事项

- 必须配置 OSPF 基本功能
- 广播网络和 NBMA 网络不能关闭源地址校验

配置方法

▾ 配置关闭源地址校验

- 可选配置。若希望两端接口 IP 地址不同网段也能建立邻接关系，必须配置。
- 在两端接口 IP 地址不同网段的路由器上配置。

检验方法

- 不同网段邻居能建立。

相关命令

▾ 配置关闭源地址校验

【命令格式】 **ip ospf source-check-ignore**

【参数说明】 -

【命令模式】 接口配置模式

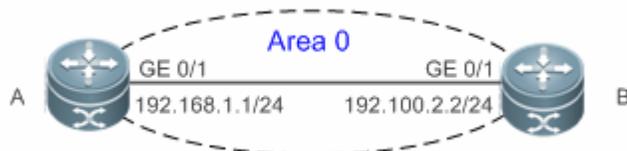
【使用指导】 通常情况下，OSPF 收到报文的源地址和接收接口会在同一个网段。不过在点对点链路上，链路两端的地址是独立设置的，可以不要要求在同一个网段，在这种场景下，由于点对点链路协商过程中会告知对端的地址信息，因此 OSPF 会检查报文的源地址是否就是协商时对端通告的地址，若不是则认为报文非法并丢弃该报文。特别的，对于 unnumbered 接口，OSPF 始终不进行地址校验。但在某些应用场景中，可能会出现源地址不如上面所述，导致 OSPF 校验失败，比如点对点链路上，协商的对端地址无法获取。在这种场景中，需要关闭源地址校验功能，以保证 OSPF 邻居能正常建立。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

配置关闭源地址校验

【网络环境】



【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 所有路由器接口配置网络类型配置为 P2P。
- 所有路由器配置关闭源地址校验。

```

A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip ospf network point-to-point
A(config-if-GigabitEthernet 0/1)# ip ospf source-check-ignore

B
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip ospf network point-to-point
B(config-if-GigabitEthernet 0/1)# ip ospf source-check-ignore

```

【检验方法】 路由器 A 上查看 OSPF 邻居正确。

```

A
A# show ip ospf neighbor
OSPF process 1, 1 Neighbors, 1 is Full:
Neighbor ID    Pri   State           Dead Time   Address        Interface
192.100.2.2    1     Full/-          00:00:34   192.100.2.2   GigabitEthernet 0/1

```

常见配置错误

4.4.12 关闭MTU校验

配置效果

- 相邻路由器的接口 MTU 不同，也能提供单播路由服务

注意事项

- 须配置 OSPF 基本功能

配置方法

配置关闭 MTU 校验

- 可选配置。默认已关闭 MTU 校验，建议保持默认配置。
- 在两端接口 MTU 不同的路由器上配置。

检验方法

MTU 不一致的邻居也能建立。

相关命令

配置关闭 mtu 校验

【命令格式】 `ip ospf mtu-ignore`

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 OSPF 在收到数据库描述报文时会校验邻居的接口的 MTU 和自己接口的 MTU 是否相同，如果收到的数据库描述报文中指示的接口的 MTU 大于接收接口的 MTU，那么邻接关系将不能被建立，此时可以通过关闭 MTU 的校验来解决。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

【网络环境】



【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 配置两边接口 MTU 不一致。
- 所有路由器配置关闭 MTU 校验（该功能默认已经开启）。

```
A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip mtu 1400
A(config-if-GigabitEthernet 0/1)# ip ospf mtu-ignore
```

```
B
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip mtu 1600
B(config-if-GigabitEthernet 0/1)# ip ospf mtu-ignore
```

【检验方法】

- 路由器 A 上查看 OSPF 邻居正确。

```
A
A# show ip ospf neighbor
OSPF process 1, 1 Neighbors, 1 is Full:
Neighbor ID    Pri   State           Dead Time   Address        Interface
192.168.1.2    1     Full/DR         00:00:34   192.168.1.2   GigabitEthernet 0/1
```

常见配置错误

4.4.13 启动双向维持

配置效果

- 非 hello 报文也能维持邻居状态

注意事项

- 必须配置 OSPF 基本功能

配置方法

▾ 配置双向维持

- 可选配置。该功能默认已开启，建议保持默认配置。
- 所有路由器上配置。

检验方法

非 hello 报文也能维持邻居。

相关命令

配置双向维持

【命令格式】 **two-way-maintain**

【参数说明】 -

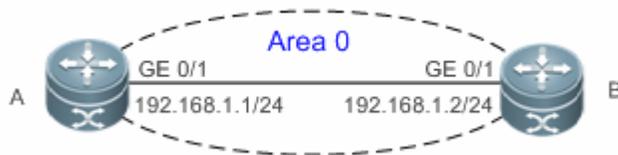
【命令模式】 OSPF 路由配置模式

【使用指导】 在较大的网络中，可能会出现大量的报文收发，占用较高的 CPU 和内存资源，造成部分部分报文的延迟或丢弃，如果 Hello 报文的处理超出了死亡间隔，将导致对应的邻接超时断开。这时启用双向维持功能，如果网络中存在大量的报文，除了 Hello 报文外，来自某个邻居的 DD、LSU、LSR、LSAck 报文也能够用于维持该邻居的双向关系，避免因 Hello 报文的延迟或丢弃而导致邻居的失效。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

【网络环境】



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 路由器 A 配置启动双向维持（该功能默认已经启动）。

```
A
A# configure terminal
A(config)# router ospf 1
A(config-router)# two-way-maintain
```

【检验方法】 建邻居过程中，路由器 A 检查邻居到期时间，并没有等路由器 B 发送 Hello 报文后才更新。

```
A
A# show ip ospf neighbor
OSPF process 1, 1 Neighbors, 1 is Full:
Neighbor ID    Pri   State           Dead Time   Address        Interface
192.168.1.2    1     Full/BDR        00:00:40   192.168.1.2   GigabitEthernet 0/1
```

常见配置错误

-

4.4.14 启动GR功能

配置效果

- 分布式路由器从主板切换到备板过程中，用户数据流量始终保持转发不断流
- OSPF 协议进程重新启动期间，用户数据流量始终保持转发不断流

注意事项

- 必须配置 OSPF 基本功能
- 邻居路由器必须支持 GR helper。
- GR 时间不能小于邻居路由器的邻居维持时间。

配置方法

配置 OSPF GR 功能

- 可选配置。该功能默认已开启，建议保持开启状态。
- 在所有路由器上配置。

配置 OSPF GR helper 功能

- 可选配置。该功能默认已开启，建议保持开启状态。
- 在所有路由器上配置。

检验方法

- 分布式路由器热备切换到备板，在此期间数据流量始终保持转发不断流。
- OSPF 协议进程重启，在此期间数据流量始终保持转发不断流。

相关命令

配置 OSPF 功能

【命令格式】 **graceful-restart [grace-period *grace-period* | inconsistent-lsa-checking]**

【参数说明】 **grace-period *grace-period***：配置 GR 重启周期，重启周期为从 OSPF 失效到 OSPF 重新启动并完成优雅重启的最长时间。*grace-period* 取值范围 1-1800 秒，缺省值为 120s。
inconsistent-lsa-checking：拓扑变化检测。检测到拓扑改变，提前退出 GR 完成收敛。开启 GR 后，默认开启检测。

【配置模式】 OSPF 路由配置模式

【使用指导】 优雅重启功能基于 OSPF 实例配置，不同的实例可以根据情况配置不同的参数。该命令用于配置设备的 GR Restarter 能力。重启周期为 GR 从开始到结束的最长时间，在这段时间内执行链路状态重建工作，使 OSPF 恢复到原有状态。当重启周期到期后，OSPF 将退出 GR 状态，执行普通的 OSPF 操作。

graceful-restart 设置 GR 的重启周期为 120 秒 **graceful-restart grace-period** 允许用户显式修改重启周期。GR 成功执行，保证转发不间断的前提是拓扑保持稳定。若拓扑发生改变，OSPF 将尽快收敛，不再等待 GR

的继续执行，避免长时间的转发黑洞。

关闭拓扑检测：对于热备过程中的拓扑变化不能及时收敛，有可能造成长时间的转发黑洞；

使能拓扑检测：有可能转发间断，但间断时间远小于关闭拓扑检测；

对于拓扑检测，大部分情况下不建议关闭。在特殊场景下，热备后拓扑会变化，但可保证不会出现长时间的转发黑洞，此时可关闭拓扑检测，最大程度减少热备过程中转发间断的时间。

配置 OSPF GR helper 功能

【命令格式】 `graceful-restart helper { disable | strict-lsa-checking | internal-lsa-checking }`

【参数说明】 **disable**：禁止设备作为辅助设备辅助其他设备执行优雅重启。

strict-lsa-checking：GR Helper 期间通过检查 types 1-5, 7 的 LSA 的变化判断网络是否发生变化，如果网络发生变化，将退出 GR Helper。

internal-lsa-checking：GR Helper 期间通过检查 types 1-3 的 LSA 的变化判断网络是否发生变化，如果网络发生变化，将退出 GR Helper。

【配置模式】 OSPF 路由配置模式

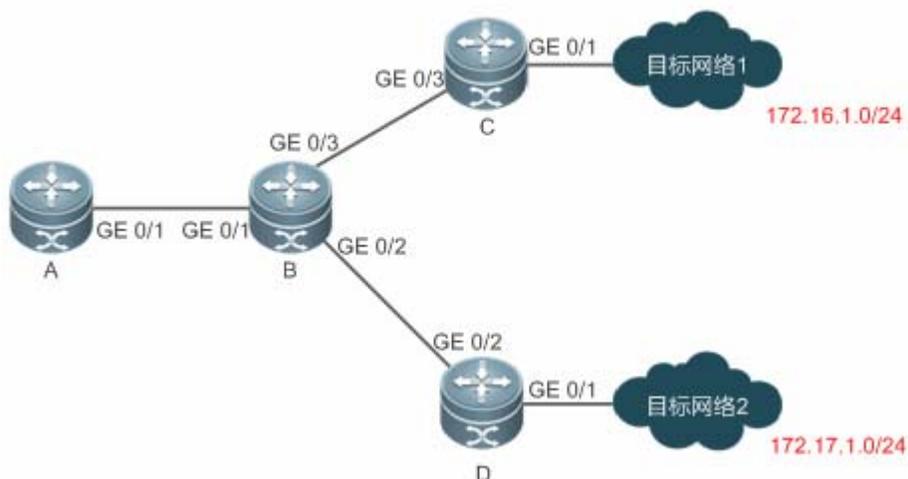
【使用指导】 该命令用于配置设备的 GR Help 能力，当邻居设备实施优雅重启时，将发送 Grace-LSA 通告其所有的邻居设备。如果本设备使能 GR Help 能力，在收到 Grace-LSA 时，会成为 GR Helper 辅助邻居完成优雅重启。disable 选项表示禁止对任何执行优雅重启的设备进行 GR 辅助。

设备成为 GR Helper 后，默认情况下不对网络变化进行检测，如果网络发生变化，也将等到 GR 完成之后才会重新进行收敛。如果用户希望在 GR 过程中，对于变化的网络能够快速探测，可以通过配置 strict-lsa-checking 或者 internal-lsa-checking 选项启动检测，其中前者对任何表示网络信息的 LSA(types 1-5, 7)进行检测，后者对表示 AS 域内路由的 LSA(types 1-3)进行检测。在网络规模较大的情况下，不建议用户启动 LSA 检测选项，因为局部网络的变化会触动 GR 结束，从而导致整网的收敛降低。

配置举例

 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

【网络环境】



【注释】 接口 IP 地址：

A : GE 0/1 192.168.1.1

B : GE 0/1 192.168.1.1 GE 0/2 192.168.2.1 GE 0/3 192.168.3.1

C : GE 0/1 192.168.4.2 GE 0/3 192.168.3.2

D : GE 0/1 192.168.5.2 GE 0/2 192.168.2.2

【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 路由器 A、C、D 开启 GR Helper 功能（该功能默认已经开启）。
- 路由器 B 开启 GR 功能。

```

B
B# configure terminal
B(config)# router ospf 1
B(config-router)# graceful-restart

```

【检验方法】

- 路由器 B 执行热备切换，期间在路由器 A 上查看到目标网络 1 和目标网络 2 的路由表保持不变。
- 路由器 B 执行热备切换，期间在路由器 A 上 ping 目标网络 1 不间断。

常见配置错误

- 配置的 GR 时间小于邻居路由器的邻居维持时间，导致 GR 过程中转发中断。

4.4.15 启动OSPF与BFD联动

配置效果

- 一旦发生链路故障，OSPF 可以快速发现路由失效。此项配置有助于缩短流量中断时间。

注意事项

- 必须配置 OSPF 基本功能
- 接口必须先配置 BFD 参数。
- 进程和接口同时配置 BFD 时，接口下的配置优先生效。

配置方法

配置与 BFD 联动

- 可选配置。若希望加快 OSPF 网络收敛，需要配置。
- 链路两端的路由器都需要配置。

检验方法

- 通过 `show bfd neighbor` 查看 BFD 邻居正常

相关命令

配置 OSPF 接口与 BFD 联动

【命令格式】 `ip ospf bfd [disable]`

【参数说明】 `disable`：关闭运行 OSPF 的指定接口使用 BFD 机制进行链路检测。

【命令模式】 接口配置模式

【使用指导】 接口配置的优先级高于进程配置模式下的 `bfd all-interfaces` 命令。

用户可以根据实际环境，通过 `ip ospf bfd` 命令选择指定的接口使用 BFD 进行链路检测，也可以使用 OSPF 进程配置模式下的命令 `bfd all-interfaces` 配置参与 OSPF 运行的所有接口通过 BFD 进行链路检测，使用 `ip ospf bfd disable` 选择关闭指定接口的 BFD 检测功能。

配置 OSPF 进程与 BFD 联动

【命令格式】 `bfd all-interfaces`

【参数说明】 -

【命令模式】 OSPF 进程配置模式

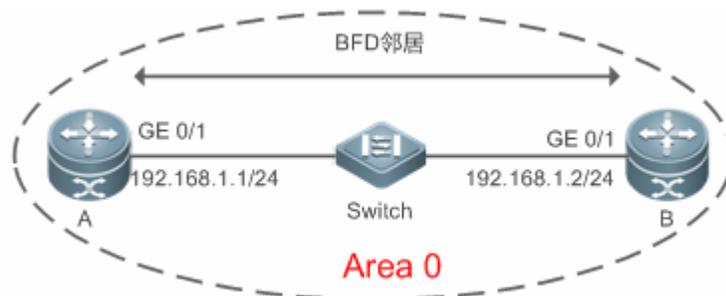
【使用指导】 OSPF 协议通过 Hello 报文动态发现邻居，当 OSPF 启动 BFD 检测功能后，将会为达到 FULL 关系的邻居建立 BFD 会话，通过 BFD 机制检测邻居状态，一旦 BFD 邻居失效，OSPF 会立刻进行网络收敛。

用户也可以通过接口配置模式命令 `ip ospf bfd [disable]` 启动或者关闭指定接口的 BFD 检测功能，接口配置的优先级高于进程配置模式下的 `bfd all-interfaces` 命令。

配置举例

 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

【网络环境】



【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 所有路由器配置接口 BFD 参数。
- 所有路由器配置 OSPF 联动 BFD。

A

```
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# bfd interval 200 min_rx 200 multiplier 5
A(config)# router ospf 1
A(config-router)# bfd all-interfaces
```

B

```
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 2/1)# bfd interval 200 min_rx 200 multiplier 5
B(config)# router ospf 1
B(config-router)# bfd all-interfaces
```

【检验方法】

- 路由器 A、B 上查看 OSPF 联动 BFD 处于 UP 状态
- 断开路由器 B 与 Switch 之间的链路，路由器 A 上 BFD 邻居断开，并且 OSPF 路由删除。

A

```
A# show ip ospf neighbor
OSPF process 1, 1 Neighbors, 1 is Full:
Neighbor ID    Pri  State           BFD State  Dead Time  Address      Interface
192.168.1.2    1    Full/BDR        Up          00:00:40   192.168.1.2 GigabitEthernet 0/1
```

B

```
B# show ip ospf neighbor
OSPF process 1, 1 Neighbors, 1 is Full:
Neighbor ID    Pri  State           BFD State  Dead Time  Address      Interface
192.168.1.1    1    Full/BDR        Up          00:00:40   192.168.1.1 GigabitEthernet 0/1
```

常见配置错误

4.4.16 启动快速重路由

配置效果

- 一旦 OSPF 发现路由失效，路由可以立即切换到次优路径。此项配置有助于缩短流量中断时间。

注意事项

- 必须配置 OSPF 基本功能
- 快速重路由 LFA 配置与虚链路配置互斥。
- 接口需要配置 carrier-delay 值为 0

配置方法

配置快速重路由

- 可选配置。若希望 OSPF 网络收敛提升到毫秒级，需要配置。
- 在到达目的网络存在多条路径的路由器上配置。

配置接口禁用备份

- 可选配置。若用户希望在最优路径失效后，数据流量不切换到指定的路径上，必须配置。此时最优路径失效后，流量将切换到其他的次优路径上，但最终 OSPF 协议重新收敛后，将根据接口代价选择新的最优路径。
- 在配置快速重路由的设备上配置。

检验方法

通过 `show ip route fast-reroute` 查看路由不仅有最优路径，还有次优路径。

相关命令

配置快速重路由

【命令格式】 `fast-reroute { lfa [downstream-paths] | route-map route-map-name }`

【参数说明】 `lfa`：启用无环路备份路径计算。

`downstream-paths`：启用下游路径计算。

`route-map route-map-name`：通过路由图指定备份路径。

【命令模式】 OSPF 路由配置模式

【使用指导】 配置 `lfa` 参数，则启用无环路备份路径计算，这时可以通过接口模式命令指令接口的路径保护方式。

网络中存在以下情况不建议开启 `lfa` 自动计算备份路径：

1. 存在虚链路

- 2.存在 Alternative ABR
- 3.当一个 ASBR 自身又是 ABR 角色
- 4.存在多个 ASBR 通告相同外部路由

在 `lfa` 基础上配 `downstream-paths`，则启用下游路径计算。

配置 `route-map`，可以通过路由图为匹配成功的路由指定备份路径。

使用 OSPF 快速重路由功能建议同时使能 BFD 链路检测功能，设备能够更快检测到链路故障，从而降低转发中断时间。针对接口 up/down 的场景，为了降低 OSPF 快速重路由的转发中断时间，可以在三层接口的接口模式下配置 `carrier-delay 0`，以达到最快的切换速度。

配置接口 LFA 保护

【命令格式】 `ip ospf fast-reroute protection { node | link-node | disable }`

【参数说明】 `node`：启用 LFA 节点保护。

`link-node`：启用 LFA 链路节点保护。

`disable`：禁用 LFA 保护。

【命令模式】 接口配置模式

【使用指导】 在 OSPF 路由进程配置模式下启用 `fast-reroutelfa` 命令，将使能 OSPF 快速重路由计算功能，并根据接口配置模式下指定的 LFA 保护方式，为主路由生成备份路由，缺省启用每个 OSPF 接口的链路保护，这种保护形式下主链路的失败不会影响备份路由的转发。

使用 `node` 参数指定启用该接口的节点保护，即主链路对应的邻居节点失败不会影响备份路由的转发。

使用 `link-node` 参数将同时对主路由对应的链路和邻居节点进行保护。

使用 `disable` 参数，将禁用该接口的 LFA 保护功能，即不为下一跳经过该接口的路由生成备份表项。

该命令在 `fast-reroute route-map` 配置下无效。

配置接口禁用备份

【命令格式】 `ip ospf fast-reroute no-eligible-backup`

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 如果某个接口剩余带宽较小或与其主接口可能同时发生失败，则该接口不适合作为备份接口，所以需要在该接口的接口配置模式下启用该命令，在 OSPF 快速重路由计算时，将该接口排除作为备份接口，启用该命令后，将从其它接口中选取备份接口。

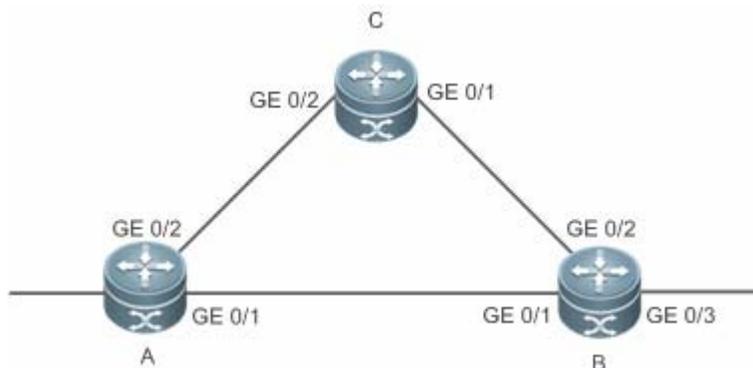
该命令在 `fast-reroute route-map` 配置下无效。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

配置快速重路由

【网络环境】



【注释】 接口 IP 地址：

A : GE0/1 192.168.1.1 GE0/2 192.168.2.1

B : GE0/1 192.168.1.2 GE0/2 192.168.3.1 GE0/3 192.168.4.1

C : GE0/1 192.168.3.2 GE 0/2 192.168.2.2

【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 路由器 A 配置快速重路由。
- 路由器 A 配置接口 carrier-delay 0。

A

```

A# configure terminal
A(config)# router ospf 1
A(config-router)# fast-reroute lfa
A(config-router)# exit
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# carrier-delay 0
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)# carrier-delay 0
  
```

【检验方法】

路由器 A 查看路由表，192.168.4.0/24 表项存在备份路由。

```

A# show ip route fast-reroute | begin 192.168.4.0
0 192.168.4.0/24 [ma] via 192.168.1.2, 00:39:28, GigabitEthernet 0/1
    [b] via 192.168.2.2, 00:39:28, GigabitEthernet 0/2
  
```

常见配置错误

4.4.17 启动iSPF特性

配置效果

- OSPF 采用 iSPF(增量 SPF)算法来计算网络拓扑。

注意事项

- 必须配置 OSPF 基本功能。

配置方法

配置 iSPF 特性

- 可选配置。若单个区域超过 100 台路由器，且希望加快收敛，建议配置
- 在区域的所有路由器上配置。

检验方法

通过 `show ip ospf` 查看 iSPF 特性已经开启。

相关命令

配置 iSPF 特性

【命令格式】 `ispf enable`

【参数说明】 -

【命令模式】 OSPF 路由配置模式

【使用指导】 启用 iSPF 特性后，OSPF 将采用增量 SPF 算法来计算网络拓扑，即网络拓扑变化后，只修正受拓扑变化影响的节点，而不是重建整棵最短路径树。

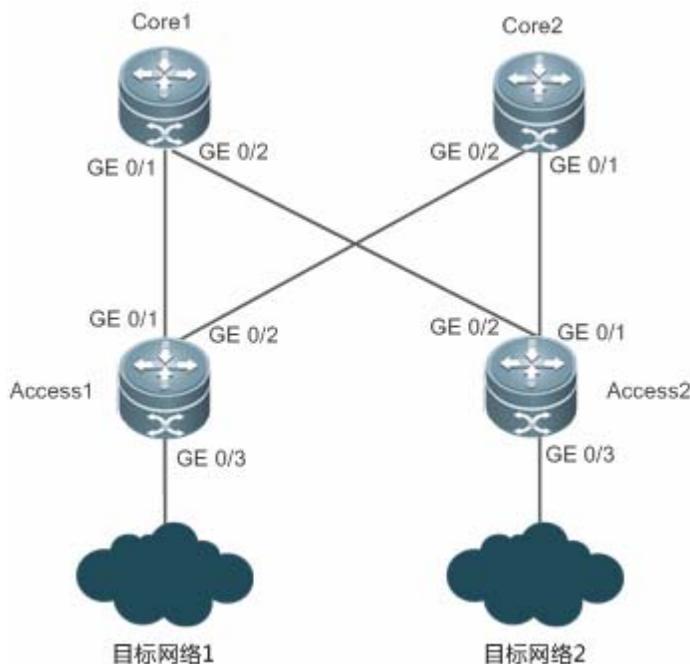
iSPF 特性通常在大型网络中用于缓解路由器处理器的压力。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

配置 iSPF 特性

【网络环境】



【注释】 接口 IP 地址：

Core1 : GE0/1 192.168.1.1 GE0/2 192.168.2.1
 Core2 : GE0/1 192.168.3.1 GE0/2 192.168.4.1
 Access1 : GE0/1 192.168.1.2 GE 0/2 192.168.3.2
 Access2 : GE0/1 192.168.4.2 GE 0/2 192.168.2.2

【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 所有路由器配置 iSPF 特性。

Core1

```
Core1# configure terminal
Core1(config)# router ospf 1
Core1(config-router)# ispf enable
```

Core2

```
Core2# configure terminal
Core2(config)# router ospf 1
Core2(config-router)# ispf enable
```

Access1

```
Access1# configure terminal
Access1(config)# router ospf 1
Access1(config-router)# ispf enable
```

Access2

```
Access2# configure terminal
Access2(config)# router ospf 1
Access2(config-router)# ispf enable
```

【检验方法】

路由器 Core1 查看 iSPF 特性是否已经开启。

```
Core1# show ip ospf
```

```
Routing Process "ospf 1" with ID 1.1.1.1
  Process uptime is 17 hours 48 minutes
  Process bound to VRF default
  Memory Overflow is enabled.
  Router is not in overflow state now.
  Conforms to RFC2328, and RFC1583Compatibility flag is enabled
  Supports only single TOS(TOS0) routes
  Supports opaque LSA
  Enable two-way-maintain
  Enable ispf
  Initial SPF schedule delay 1000 msec
  Minimum hold time between two consecutive SPF 5000 msec
  Maximum wait time between two consecutive SPF 10000 msec
  Initial LSA throttle delay 0 msec
  Minimum hold time for LSA throttle 5000 msec
  Maximum wait time for LSA throttle 5000 msec
  Lsa Transmit Pacing timer 40 msec, 1 LS-Upd
  Minimum LSA arrival 1000 msec
  Pacing lsa-group: 30 sec
  Number of incoming current DD exchange neighbors 0/5
  Number of outgoing current DD exchange neighbors 0/5
  Number of external LSA 0. Checksum 0x000000
  Number of opaque AS LSA 0. Checksum 0x000000
  Number of non-default external LSA 0
  External LSA database is unlimited.
  Number of LSA originated 2
  Number of LSA received 93
  Log Neighbor Adjacency Changes : Enabled
  Graceful-restart disabled
  Graceful-restart helper support enabled
  Number of areas attached to this router: 1: 1 normal 0 stub 0 nssa
    Area 1
      Number of interfaces in this area is 1(1)
      Number of fully adjacent neighbors in this area is 0
      Number of fully adjacent virtual neighbors through this area is 0
      Area has no authentication
      SPF algorithm executed 0 times
      iSPF algorithm last executed 00:04:14.534 ago
      iSPF algorithm executed 12 times
      Number of LSA 1. Checksum 0x0029b3
```

常见配置错误

-

4.4.18 配置网管功能

配置效果

- 通过网管软件管理 OSPF 协议的参数及监控协议运行状态。

注意事项

- 必须配置 OSPF 基本功能
- 要开启 OSPF 的 MIB 功能，必须开启 SNMP-Server 的 MIB 功能。
- 要开启 OSPF 的 Trap 功能，必须开启 SNMP-Server 的 Trap 功能。
- 要输出 OSPF 的日记信息，必须开启设备的日记开关。

配置方法

▾ 配置 MIB 绑定

- 可选配置。若希望通过网管软件管理指定的 OSPF 实例的参数，需要配置。
- 所有路由器上配置。

▾ 配置使能 Trap

- 可选配置。若希望通过网管软件监控协议的运行状态，需要配置。
- 所有路由器上配置。

▾ 配置日记输出

- 可选配置。该功能默认已开启，建议保持默认配置，若希望减少日记输出，可关闭该功能。
- 所有路由器上配置。

检验方法

- 通过网管软件可管理协议的参数。
- 通过网管软件可监控协议的运行状态。

相关命令

配置 MIB 绑定

- 【命令格式】 **enable mib-binding**
- 【参数说明】 -
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 由于 OSPFv2 MIB 没有 OSPFv2 进程信息，所以用户通过 SNMP 操作 OSPFv2 进程，只能操作唯一的进程。缺省情况下，OSPFv2 MIB 绑定在进程号最小的 OSPFv2 进程上，用户的操作都对该进程生效。如果用户希望能够通过 SNMP 操作指定的 OSPFv2 进程，可以通过本命令将 MIB 绑定到该进程上。

配置使能 Trap

- 【命令格式】 **enable traps [error [IfAuthFailure | IfConfigError | IfRxBadPacket | VirtIfAuthFailure | VirtIfConfigError | VirtIfRxBadPacket] | lsa [LsdbApproachOverflow | LsdbOverflow | MaxAgeLsa | OriginateLsa] | retransmit [IfTxRetransmit | VirtIfTxRetransmit] | state-change [IfStateChange | NbrRestartHelperStatusChange | NbrStateChange | NssaTranslatorStatusChange | RestartStatusChange | VirtIfStateChange | VirtNbrRestartHelperStatusChange | VirtNbrStateChange]]**
- 【参数说明】
- IfAuthFailure**：接口认证错误。
 - IfConfigError**：接口参数配置错误。
 - IfRxBadPacket**：接口接收到错误报文。
 - IfRxBadPacket**：接口接收到错误报文。
 - VirtIfAuthFailure**：虚拟接口认证错误。
 - VirtIfConfigError**：虚拟接口参数配置错误。
 - VirtIfRxBadPacket**：虚拟接口接收到错误报文。
 - LsdbApproachOverflow**：外部 LSA 的数量达到了上限值的 90%。
 - LsdbOverflow**：外部 LSA 的数量达到了上限值。
 - MaxAgeLsa**：LSA 达到老化时间。
 - OriginateLsa**：产生新的 LSA。
 - IfTxRetransmit**：接口上发生报文重传。
 - VirtIfTxRetransmit**：虚接口上发生报文重传。
 - IfStateChange**：接口状态变化。
 - NbrRestartHelperStatusChange**：邻居 GR 过程状态变化。
 - NbrStateChange**：邻居状态变化。
 - NssaTranslatorStatusChange**：NSSA 转换者状态变化。
 - RestartStatusChange**：本机 GR 状态变化。
 - VirtIfStateChange**：虚拟接口状态变化。
 - VirtNbrRestartHelperStatusChange**：虚拟邻居 GR 过程状态变化。
 - VirtNbrStateChange**：虚拟邻居状态变化。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 本命令对应功能受 **snmp-server** 命令的限制，必须先打开配置开关 **snmp-server enable traps ospf**，再配

置 enable traps 以后，对应的 ospf trap 才能够正确发送。
本命令不受进程绑定 MIB 的限制，允许不同进程同时打开 TRAP 开关。

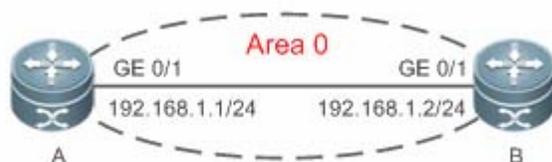
配置日记输出

- 【命令格式】 **log-adj-changes [detail]**
- 【参数说明】 **detail**：记录全部状态变化信息。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 -

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

【网络环境】



- 【配置方法】
- 所有路由器配置接口 IP 地址。略。
 - 所有路由器配置 OSPF 基本功能。略。
 - 路由器 A 配置 MIB 绑定。
 - 路由器 A 配置 Trap 使能。

```
A
A# configure terminal
A(config)# snmp-server host 192.168.2.2 traps version 2c public
A(config)# snmp-server community public rw
A(config)# snmp-server enable traps
A(config)# router ospf 10
A(config-router)# enable mib-binding
A(config-router)# enable traps
```

【检验方法】 通过 MIB 工具读取、设置 OSPF 的参数信息，并查看 OSPF 协议运行状态。

常见配置错误

SNMP-Server 配置错误：没有开启 MIB、Trap 功能等。

4.4.19 修改协议控制参数

配置效果

修改协议运行参数，改变协议运行状态。

注意事项

- 必须配置 OSPF 基本功能
- 邻居维持时间不得小于 Hello 间隔时间

配置方法

配置 hello 时间

- 可选配置，不建议修改默认配置。
- 在链路两端的路由器上配置。

配置 dead 时间

- 可选配置，不建议修改默认配置。若用户希望链路失效时加快协议收敛，可调整该配置。
- 在链路两端的路由器上配置。

配置 Isu 重传间隔

- 可选配置。若用户环境有大量路由并且网络拥塞严重，建议调整该配置。

配置 Isa 生成时间

- 可选配置，不建议修改默认配置。

配置 Isa 分组刷新时间

- 可选配置，不建议修改默认配置。若用户环境有大量路由，可调整该配置。
- 在 ASBR 或者 ABR 上配置。

配置 Isa 重复接收延迟

- 可选配置，不建议修改默认配置。

配置 SPF 计算时间

- 可选配置。若网络环境频繁震荡，可调整该配置。

配置区间路由计算延时时间

- 可选配置。不建议修改默认配置。
- 在所有路由器上配置。

配置外部路由计算延时时间

- 可选配置。不建议修改默认配置。

- 在所有路由器上配置。

检验方法

通过 `show ip ospf`、`show ip ospf neighbor` 查看协议运行参数及状态

相关命令

配置 hello 时间

【命令格式】 `ip ospf hello-interval seconds`

【参数说明】 `seconds`：设置 OSPF 发送 hello 报文的时间间隔，以秒为单位，取值范围 1-65535。

【命令模式】 接口配置模式

【使用指导】 hello 报文时间间隔包含在 hello 报文中。该时间越短，OSPF 检测到拓扑变化的时间也将越快，但是将增加网络流量。同一网段上的所有路由设备的 hello 报文时间间隔必须一致。如果还想手工修改判断邻居死亡时间，必须保证 hello 报文时间间隔不能大于邻居死亡时间。

配置 dead 时间

【命令格式】 `ip ospf dead-interval seconds`

【参数说明】 `seconds`：定义判断邻居死亡的时间，以秒为单位，取值范围：0-2147483647。

【命令模式】 接口配置模式

【使用指导】 OSPF 的死亡时间包含在 Hello 报文内。如果 OSPF 在死亡时间内，没有收到邻居发送的 Hello 报文，就宣告邻居无效并从邻居表中删除该邻居记录。死亡时间缺省为 hello 报文时间间隔的 4 倍，如果修改了 hello 时间间隔，会自动修改死亡时间。

用该命令可以手工修改 OSPF 判断邻居死亡时间，但是必须十分谨慎。有两点必须注意：

第一点：死亡时间间隔，不能小于 hello 报文时间间隔；

第二点：同一网段上的所有路由设备的死亡时间间隔必须一致。

配置 lsu 发送延迟

【命令格式】 `ip ospf transmit-delay seconds`

【参数说明】 `seconds`：定义 OSPF 接口传输 LSU 报文的延时，以秒为单位，取值范围：0-65535。

【命令模式】 接口配置模式

【使用指导】 在传输 LSU 报文之前，该报文中所有 LSAs（链路状态描述）的 Age 字段将增加，增加量根据该接口配置命令 `ip ospf transmit-delay` 的定义。设置该参数必须考虑到该接口的发送和线路传播时延，对于低速线路，接口的传输延时应该设置得大些。虚拟链路的 LSU 报文传输延时，用 `area virtual-link` 命令跟随 `transmit-delay` 关键字进行定义。

对 Age 达到 3600 的 LSA，将进行重传或请求重传，如果没有得到及时刷新，超时的 LSA 将从链路状态数据库清除。

配置 lsu 重传间隔

【命令格式】 `ip ospf retransmit-interval seconds`

- 【参数说明】 *seconds* : LSU 重传时间间隔, 以秒为单位, 取值范围: 0-65535。该时间必须大于两个邻居之间的数据包往返传输延时。
- 【命令模式】 接口配置模式
- 【使用指导】 当路由设备发送完一个 LSU 报文, 该 LSU 报文还保留在发送缓冲队列中, 如果在 **ip ospf retransmit-interval** 命令定义的时间内, 没有得到邻居的确认, 将重新发送 LSU。
为了避免不必要的重传, 在串行线路或虚拟链路上, 可以将重传时间设置得大一点。虚拟链路的 LSU 报文重传时间间隔用 **area virtual-link** 命令跟随 **retransmit-interval** 关键字进行定义。

配置 Lsa 生成时间

- 【命令格式】 **timers throttle lsa all delay-time hold-time max-wait-time**
- 【参数说明】 *delay-time* : LSA 生成的最小延迟时间, 数据库中的第一份 LSA 总是立即生成的。取值范围: 0-600000, 单位: 毫秒。
hold-time : 配置首次与第 2 次刷新 LSA 时, 最短的时间间隔。取值范围: 1-600000, 单位: 毫秒。
max-wait-time : 连续刷新 LSA 时, 两次刷新 LSA 的时间间隔最大值。同时, 这个时间间隔也是判断 LSA 是否连续刷新的值。取值范围: 1-600000, 单位: 毫秒。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 对链路变化时收敛性要求较高时, *delay-time* 可以配置为较小值。若希望减少 CPU 消耗, 可以把几个值适当调高。
配置命令时, *hold-time* 不允许小于 *delay-time*, *max-wait-time* 也不允许小于 *hold-time*。

设置 Lsa 分组刷新时间

- 【命令格式】 **timers pacing lsa-group seconds**
- 【参数说明】 *seconds* : LSA 组步调更新的步调时间, 取值范围: 10-1800, 单位: 秒。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 每个 LSA 都有自己的生存时间(LSA age), 当生存时间达到 1800 秒时, 需要进行重刷新以防止正常的 LSA 通告达到最大生存时间而被清除掉。如果对每个 LSA 都独立进行 LSA 的更新和老化计算, 将消耗大量的 CPU 资源。为了有效利用 CPU 资源, 将设备内 LSA 分组进行刷新操作。这里称一个分组刷新操作的时间间隔为组步调间隔, 分组刷新操作即将一个组步调间隔内需要刷新的 LSA 编成一组统一刷新。
在 LSA 数量一定的情况下, 步调时间间隔越大, 定时器期满后处理的 LSA 数量越多。为了保持 CPU 的稳定, 建议每次期满处理的 LSA 数量不宜太多, 所以如果 LSA 数量较大, 建议降低步调间隔。比如, 数据库中有 10000 条 LSA, 可以降低步调时间间隔, 如果只有 40~100 条, 则可以调整到 10~20 分钟。

配置 Lsa 分组刷新间隔

- 【命令格式】 **timers pacing lsa-transmit transmit-time transmit-count**
- 【参数说明】 *transmit-time* : 配置 LSA 分组发送时间间隔。取值范围: 10-1000, 单位: 毫秒。
transmit-count : 配置每组 LS-UPD 报文的个数。取值范围: 1-200, 单位: 个。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 环境中 LSA 数量很多, 设备本身负荷太大时, 配置适当的 **transmit-time** 和 **transmit-count** 可以抑制网络中洪泛的 LS-UPD 报文数量。
CPU 负荷不高, 网络带宽负荷不大时, 减小 **transmit-time** 增加 **transmit-count** 可以加快环境的收敛。

配置 Lsa 重复接收延迟

- 【命令格式】 **timers lsa arrival arrival-time**
- 【参数说明】 *arrival-time*：配置接收相同 LSA 时延时。取值范围：0-600000，单位：毫秒。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 配置指定时间之内接收到相同 LSA 时，不做处理。

配置区间路由计算延时时间

- 【命令格式】 **timers throttle route inter-area ia-delay**
- 【参数说明】 *ia-delay*：定义区间路由计算的延时时间，以毫秒为单位，取值范围：0-600000。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 对网络收敛时间有严格要求时，不能更改该时间。

配置外部路由计算延时时间

- 【命令格式】 **timers throttle route ase ase-delay**
- 【参数说明】 *ase-delay*：定义外部路由计算的延时时间，以毫秒为单位，取值范围：0-600000。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 对网络收敛时间有严格要求时，不能更改该时间。

配置 SPF 计算时间

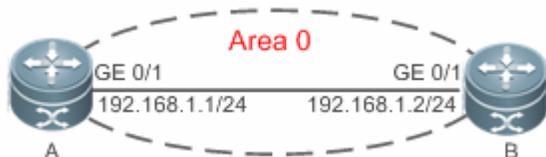
- 【命令格式】 **timers throttle spf spf-delay spf-holdtime spf-max-waittime**
- 【参数说明】 *spf-delay*：定义 SPF 计算的等待时间，以毫秒为单位，取值范围：1-600000。OSPF 路由进程接收到拓扑变化，触发的 SPF 计算至少要在 *spf-delay* 时间间隔后执行。
- spf-holdtime*：定义两次 SPF 运算之间的最小时间间隔，以毫秒为单位，取值范围：1-600000。
- spf-max-waittime*：定义两次 SPF 运算之间的最大时间间隔，以毫秒为单位，取值范围：1-600000。
- number*：聚合路由的度量值。
- 【命令模式】 OSPF 路由进程配置模式
- 【使用指导】 *spf-delay* 表示从拓扑发生变化到 SPF 开始计算，至少需要延迟的时间。第一次触发 SPF 计算到第二次触发 SPF 计算的最小时间间隔为 *spf-holdtime*，此后，连续触发 SPF 计算的时间间隔至少为上一次时间间隔的两倍，当时间间隔达到 *spf-max-waittime* 之后，将不再增加。如果两次 SPF 计算的时间间隔已经超过了要求的最小值，那么 SPF 计算时间间隔将重新从 *spf-holdtime* 开始计算。
- spf-delay* 和 *spf-holdtime* 可以设置小些以加快拓扑的收敛；*spf-max-waittime* 可以设置大些以减少 SPF 计算，可以根据实际网络拓扑稳定性采取灵活设置。
- 与 **timers spf** 命令相比，该命令能够进行更灵活的设置，既可以加快 SPF 计算收敛速度，又可以在拓扑连续变化时更好的减少 SPF 计算的系统资源消耗，所以推荐使用 **timers throttle spf** 命令进行配置。
1. *spf-holdtime* 的值不能小于 *spf-delay*，否则 *spf-holdtime* 的值将被自动设置为与 *spf-delay* 相等；
 2. *spf-max-waittime* 的值不能小于 *spf-holdtime*，否则 *spf-max-waittime* 的值将被自动设置为与 *spf-holdtime* 相等；
 3. **timers throttle spf** 配置与 **timers spf** 配置会互相覆盖；
 4. 当 **timers spf** 与 **timers throttle spf** 都未配置时，缺省值参见 **timers throttle spf** 配置的缺省值。

配置举例

i 以下配置举例，需要基于 OSPF 基本功能进行配置，OSPF 基本功能请参见前文。

配置 hello 时间和 dead 时间

【网络环境】



【配置方法】

- 所有路由器配置接口 IP 地址。略。
- 所有路由器配置 OSPF 基本功能。略。
- 所有路由器配置 Hello 时间和 Dead 时间。

```
A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip ospf hello-interval 15
A(config-if-GigabitEthernet 0/1)# ip ospf dead-interval 50

B
B# configure terminal
B(config)# interface GigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)# ip ospf hello-interval 15
A(config-if-GigabitEthernet 0/1)# ip ospf dead-interval 50
```

【检验方法】

路由器 A 查看接口参数：Hello 时间为 10S，Dead 时间为 50S。

```
A
A# show ip ospf interface
GigabitEthernet 0/1 is up, line protocol is up
  Internet Address 192.168.1.1/24, Ifindex 2, Area 0.0.0.0, MTU 1500
  Matching network config: 192.168.1.0/24
  Process ID 1, Router ID 192.168.1.2, Network Type POINTOMULTIPOINT, Cost: 1
  Transmit Delay is 1 sec, State Point-To-Point
  Timer intervals configured, Hello 15, Dead 50, Wait 40, Retransmit 5
    Hello due in 00:00:02
  Neighbor Count is 1, Adjacent neighbor count is 0
  Crypt Sequence Number is 4787
  Hello received 465 sent 466, DD received 8 sent 8
  LS-Req received 2 sent 2, LS-Upd received 8 sent 21
  LS-Ack received 14 sent 7, Discarded 3
```

常见配置错误

- 配置的邻居维持时间小于 Hello 时间间隔。

4.5 监视与维护

清除各类信息

 在设备运行过程中执行 clear 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除并重启 OSPF 进程。	clear ip ospf [process-id] process

查看运行情况

作用	命令
查看 OSPF 进程的配置信息。	show ip ospf [process-id]
查看 OSPF 内部路由表（到 ABR、ASBR 的路由）。	show ip ospf [process-id] border-routers
查看 OSPF 链路状态数据库。	show ip ospf [process-id area-id] database [{ asbr-summary external network nssa-external opaque-area opaque-as opaque-link router summary }] [{ adv-router ip-address self-originate } link-state-id brief] [database-summary max-age detail]
查看运行 OSPF 的接口。	show ip ospf [process-id] interface [interface-type interface-number brief]
查看 OSPF 的邻居表。	show ip ospf [process-id] neighbor [detail] [interface-type interface-number] [neighbor-id]
查看 OSPF 路由表。	show ip ospf [process-id] route [count]
查看 OSPF 区域的 SPT 计算次数。	show ip ospf [process-id] spf
查看 OSPF 重分布路由的汇聚路由。	show ip ospf [process-id] summary-address
查看 OSPF 网络拓扑信息。	show ip ospf [process-id [area-id]] topology [adv-router adv-router-id [router-id] self-originate [router-id]]
查看 OSPF 的虚链路	show ip ospf [process-id] virtual-links [ip-address]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 OSPF 事件的调试开关。	debug ip ospf events [abr asbr lsa nssa os restart router slink vlink]
打开 OSPF 接口的调试开关。	debug ip ospf ifsm [events status timers]
打开 OSPF 邻居的调试开关。	debug ip ospf nfm [events status timers]
打开 OSPF NSM 的调试开关。	debug ip ospf nsm [interface redistribute route]
打开 OSPF LSA 的调试开关。	debug ip ospf lsa [flooding generate install maxage refresh]
打开 OSPF 报文的调试开关。	debug ip ospf packet [dd detail hello ls-ack ls-request ls-update recv send]

打开 OSPF 路由的调试开关。

```
debug ip ospf route [ ase | ia | install | spf | time ]
```

5 策略路由

5.1 概述

策略路由（Policy-Based Routing）就是基于策略的路由。将包含策略的路由图应用在接口和设备上，即可实现策略路由。

与静态路由相似，策略路由也是手工配置的路由，不能随网络变化自动更新，仅在本地有效。与静态路由、动态路由相比，策略路由更加灵活。静态路由、动态路由只能依据目的地址转发报文。策略路由可以依据源地址、目的地址、报文长度、入端口等转发报文。

协议规范

无

5.2 典型应用

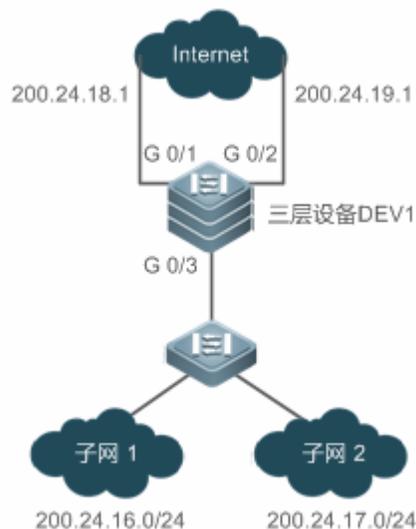
典型应用	场景描述
使用策略路由选择 ISP	为来自不同子网的报文，分别指定优选出接口。
使用策略路由实现流量分类	为来自不同子网的报文，分别指定 QoS 值

5.2.1 使用策略路由选择ISP

现有用户网络，常常会出现使用到多个 ISP（Internet Server Provider，Internet 服务提供商）资源的情形。由于不同 ISP 申请到的带宽不一，或为了保护某些重点用户的网络资源，需要使用策略路由。通过有选择的控制数据报文的转发，不仅能够充分利用 ISP 资源，又能满足灵活、多样的应用。

应用场景

图 5-1



某局域网有两个出口连接到因特网，在三层设备 DEV1 上配置策略路由，使这两个出口能够负载分担并互为备份。

具体要求如下：

- 来自子网 1 的数据流，从 GE 0/1 接口发出。
- 来自子网 2 的数据流，从 GE 0/2 接口发出。
- 如果 GE 0/1 链路断开，则将 GE 0/1 上的数据流切换到 GE 0/2 上。反之亦然。

功能部属

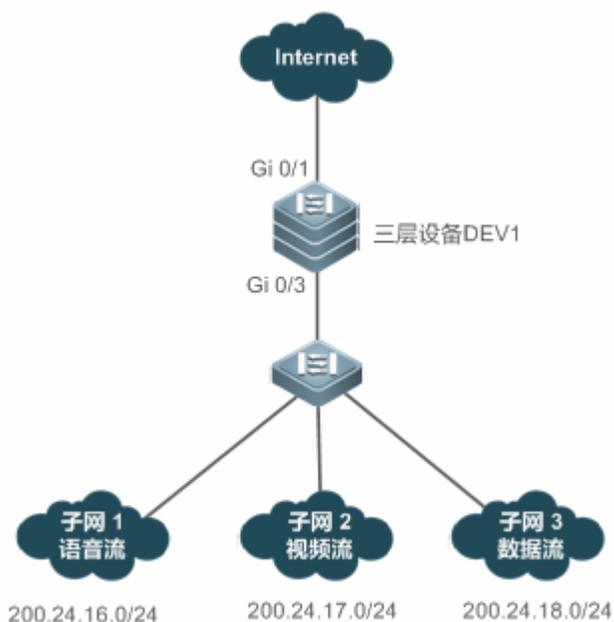
- 在三层设备 DEV1 上，配置两个不同的 ACL：
ACL1：源地址属于子网 1。
ACL2：源地址属于子网 2。
- 在三层设备 DEV1 上，在路由图中配置两条策略：
策略 1：匹配 ACL1 的报文，设置下一跳为 GE0/1、GE0/2。（按配置顺序，GE0/1 优先生效，GE0/2 为备份）
策略 2：匹配 ACL2 的报文，设置下一跳为 GE0/2、GE0/1。（按配置顺序，GE0/2 优先生效，GE0/1 为备份）
- 在 GE0/3 上配置策略路由（引用路由图），则该接口收到的报文将依照策略转发。

5.2.2 使用策略路由实现流量分类

应用场景

中小企业组网的主要特点是网络结构相对简单，各个分支节点通过运营商专线或者 Internet VPN 方式与中心节点互联，企业组网往往需要实现三网合一（语音、视频、数据），以最大程度的对现有 IP 网络进行利用并节约成本。由于所有的流量都从单一出口流出，所以在网络出口有必要对 QoS 策略进行调整，从而为带宽、时延敏感的应用提供优先通信质量保障。

图 5-2



某局域网有一个出口连接到因特网，在三层设备 DEV1 上配置策略路由，改变来自不同网络报文的 QoS 值。

具体要求如下：

- 来自子网 1 的数据流，表示语音流，DSCP 值设置为 56。
- 来自子网 2 的数据流，表示视频流，DSCP 值设置为 40。
- 来自子网 3 的数据流，表示数据流，DSCP 值设置为 24。

功能部属

- 在三层设备 DEV1 上，配置三个不同的 ACL：
 - ACL1：源地址属于子网 1。
 - ACL2：源地址属于子网 2。
 - ACL3：源地址属于子网 3。
- 在三层设备 DEV1 上，在路由图中配置三条策略：
 - 策略 1：匹配 ACL1 的报文，设置 DSCP 值为 56。
 - 策略 2：匹配 ACL2 的报文，设置 DSCP 值为 40。
 - 策略 3：匹配 ACL3 的报文，设置 DSCP 值为 24。
- 在 GE0/3 上配置策略路由（引用路由图），则该接口收到的报文将依照策略改变 DSCP 值。

5.3 功能详解

功能特性	作用
策略	配置策略路由，需要首先在路由图中配置策略。

策略路由	将包含策略的路由图应用在接口和设备上，即可实现策略路由。
------	------------------------------

5.3.1 策略

一条策略就是一个 `match ... , set...` 语句，表示“如果匹配某些条件，则执行某些处理动作”。

 关于策略的详细介绍，请参见“路由图”章节。

策略的执行

在全局模式下，使用 `route-map route-map-name [permit | deny] [sequence-number]` 命令在路由图中创建一条策略。

一个路由图中可以包含多条策略。每条策略都有对应的序号，序号越小优先级越高。策略按序号顺序依次执行，一旦满足某条策略的匹配条件，就执行该策略对应的处理动作，并退出路由图。如果不能满足任何一条策略的匹配条件，则不做任何处理。

策略有以下两种工作模式：

- `permit`：当满足策略的匹配条件时，执行该策略的处理动作，并退出路由图。
- `deny`：当满足策略的匹配条件时，不执行该策略的处理动作，并退出路由图。

策略的匹配条件

策略的匹配条件包含 0 条、1 条或多条 `match` 规则。

- 如果包含 0 条 `match` 规则，则无条件匹配。
- 如果包含 1 条或多条 `match` 规则，则必须同时匹配所有 `match` 规则，才能满足策略的匹配条件。

在路由图模式下，使用 `match` 命令设置 `match` 规则。一条 `match` 命令对应一条 `match` 规则。

策略路由支持以下 `match` 命令：

	命令	说明
IPv4 策略路由	<code>match ip address</code>	以源 IPv4 地址（和目的 IPv4 地址）作为匹配条件。  一条策略中允许配置多条 <code>match ip address</code> 命令。

-  IPv4 策略路由使用 IP 标准或者扩展 ACL 定义报文的源 IP 地址（和目的 IP 地址）范围。
-  IPv4 策略路由接口策略转发支持专家级及 MAC 名字 ACL，本地策略转发不支持专家级及 MAC 名字 ACL。
-  策略路由如果引用了 ACL，如果该 ACL 不存在，那么不匹配这条路由于图，从下一条路由于图继续匹配，当所有的路由于图都不匹配时，就走普通路由转发。如果仅配置了 ACL，但未配置 ACE，策略路由转发行为同上述 ACL 不存在的场景。

策略的处理动作

策略的处理动作包含 0 条、1 条或多条 `set` 规则。

- 如果包含 0 条 `set` 规则，则不做任何处理，直接退出路由图。

- 如果包含 1 条或多条 set 规则，则执行所有处理动作后，退出路由图。

 如果 set 规则有优先级关系，则仅优先级最高者生效。

在路由图模式下，使用 **set** 命令设置 set 规则。一条 **set** 命令对应一条 set 规则。

策略路由支持以下 **set** 命令：

命令	说明
IPv4 策略路由	set ip tos 修改 IPv4 报文的 tos 字段。  set ip tos、precedence、dscp 三者互斥，均不能与其他 2 个命令共存
	set ip precedence 修改 IPv4 报文的 precedence 字段。  set ip tos、precedence、dscp 三者互斥，均不能与其他 2 个命令共存
	set ip dscp 修改 IPv4 报文的 dscp 字段。  set ip tos、precedence、dscp 三者互斥，均不能与其他 2 个命令共存
	set ip next-hop 配置 IPv4 报文转发的下一跳。下一跳必须是直连的，否则此命令无效。 匹配 match 规则的报文将优先转发到 set ip next-hop 所指定的下一跳，而不管该报文在路由表中的实际选路结果是否和策略路由指定的下一跳一致。  交换机上，PBR 支持的下一跳对应的出接口类型为 SVI 口，路由口和 3 层 AP 口。
set ip default next-hop 配置 IPv4 报文转发的缺省下一跳。 匹配 match 规则的报文，如果该报文在路由表中选路失败或者选到缺省路由，那么报文将转发到该命令指定的缺省下一跳。	

 优先级顺序：**set ip next-hop** > 普通路由 > **set ip default next-hop** > 缺省路由。以上 set 命令允许同时配置，但只有优先级最高的生效。

5.3.2 策略路由

策略路由

将包含策略的路由图应用在接口或设备上，即可实现策略路由。

- 将路由图应用在接口上，则接口收到的报文遵循策略进行路由。
此类策略路由常用来控制设备收到的用户数据报文。它只对转发的报文起作用，对本地始发的报文不起作用。
- 将路由图应用在设备上，则本地始发的报文遵循策略进行路由。
此类策略路由常用来控制设备之间交互的协议报文（比如本地发出的 ping 报文）。它只对本地始发的报文起作用，对转发的报文不起作用。

 缺省情况下，设备上不存在策略路由，报文转发参照路由表执行。

冗余备份或负载均衡

可以在一条策略中设置多个下一跳。多个下一跳之间可以是冗余备份，也可以是负载均衡。缺省为冗余备份。

i 冗余备份或负载均衡仅对 **set ip next-hop**、**set ip default next-hop** 命令设置的下一跳生效，仅在同一 **set** 规则的多个下一跳之间生效。

- 冗余备份

按配置顺序，第一个可达的下一跳生效。当前生效的下一跳（R1）发生故障后，流量会自动切换到下一个可达的下一跳（R2）。当 R1 恢复可达时，流量会自动切回 R1。

新添加的下一跳排序在最后。假设多个下一跳的原有顺序为 R1/R2/R3，删除 R1 后再添加 R1，顺序变为 R2/R3/R1。

如果所有的下一跳都不可达，则丢弃报文。

- 负载均衡

多个可达的下一跳同时生效，支持 WCMP(Weighted Cost Multiple Path 加权多路径)和 ECMP(Equal Cost Multiple Path，等价多路径)。某一个可达的下一跳失效后，流量在其它可达的下一跳上负载均衡。

与 BFD 联动

策略路由与 BFD 联动，仅对 **set ip next-hop** 命令设置的下一跳生效。

在 **set ip next-hop** 命令中携带参数 **verify-availability**、**bfd interface-type interface-number gateway**，可以将策略路由与 BFD 会话建立联动，监控下一跳的可达性。

策略路由与 BFD 联动，有助于增强策略路由对网络环境变化的感知能力。当 BFD 发现当前用于转发到下一跳不可达时，会立即通知策略路由将流量切换到下一个可达的下一跳（冗余备份）或者其余所有可达的下一跳（负载均衡）。

i 策略路由与 BFD 联动的配置与相关命令，请参见“BFD”章节。

5.4 产品说明



- 在 S7508 产品上：
- 策略路由如果引用了 ACL，但是该 ACL 不存在，则所有报文都匹配，按普通路由查找转发；如果引用了 ACL，但是其中没有任何 ACE，则从下一个策略的 ACL 继续匹配。ACE 的 deny 选项行为执行普通转发。deny any any 行为是走普通路由转发。

5.5 配置详解

配置项	配置建议 & 相关命令
配置策略路由基本功能	 必须配置。用于应用策略路由对报文进行转发。

	ip policy route-map	对接口收到的 IPv4 报文应用策略路由。
	ip local policy route-map	对本地始发的 IPv4 报文应用策略路由。
设置冗余备份或负载均衡	 可选配置。用于设置策略路由在多个下一跳之间是冗余备份还是负载均衡。	
	ip policy { redundance load-balance }	设置 IPv4 策略路由在多个下一跳之间冗余备份或负载均衡。缺省为冗余备份。

5.5.1 配置策略路由基本功能

配置效果

通过制定灵活的策略，对用户数据流进行个性化的路由管理。

通过制定灵活的策略，对协议交互和网络拓扑进行个性化的管理。

注意事项

- 配置策略路由时必须引用路由图。因此，必须在设备上配置路由图。
- 如果配置路由图时引用了 ACL，则必须在设备上配置 ACL。

配置方法

对接口收到的 IPv4 报文应用策略路由

- 如果需要对流经设备的 IPv4 用户数据流进行个性化的路由管理，则应执行此配置。
- 在用户数据流的入接口上配置。
- 使用 **ip policy route-map** 命令将路由图应用在接口上，则接口收到的 IPv4 报文执行策略路由。

【命令格式】 **ip policy route-map** *route-map-name*

【参数说明】 *route-map-name*：路由图的名称

【缺省配置】设备上不存在策略路由，报文转发参照路由表执行

【命令模式】接口模式

【使用指导】一个接口上只能配置 1 条 **ip policy route-map** 命令。如果在同一个接口上多次配置 **ip policy route-map** 命令，仅最后一次配置生效。

如果此策略路由中引用的路由图实际不存在，则此策略路由不生效。

 路由器不支持在 DAILER 口, loopback 口, tunnel 口上配置策略路由，配置后不会生效。

对本地始发的 IPv4 报文应用策略路由

- 如果需要对 IPv4 协议交互和 IPv4 网络拓扑进行个性化的管理，则应执行此配置。
- 使用 **ip local policy route-map** 命令将路由图应用在设备上，则本地始发的 IPv4 报文执行策略路由。

- 【命令格式】 **ip local policy route-map** *route-map-name*
- 【参数说明】 *route-map-name* : 路由图的名称
- 【缺省配置】 设备上不存在策略路由，报文转发参照路由表执行
- 【命令模式】 全局模式
- 【使用指导】 一个设备上只能配置 1 条 **ip local policy route-map** 命令。
如果此策略路由中引用的路由图实际不存在，则此策略路由不生效。

检验方法

- 查看策略路由的配置信息。
- 查看策略路由引用的路由图的配置信息。
- 如果配置路由图时引用了 ACL，则应查看 ACL 的配置信息。

查看 IPv4 策略路由的配置信息

- 【命令格式】 **show ip policy** [*route-map-name*]
- 【参数说明】 *route-map-name* : 路由图的名称
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 从输出信息中观察有哪些接口上配置了 IPv4 策略路由，引用的路由图叫什么名字。

```
Ruijie# show ip policy
Banlance mode: redundance
Interface          Route map
local              RM_for_PBR_1
GigabitEthernet 0/1  RM_for_PBR_2
```

Local 表示对本地始发的 IPv4 报文应用策略路由。

查看路由图的配置信息

- 【命令格式】 **show route-map** [*route-map-name*]
- 【参数说明】 *route-map-name* : 路由图的名称
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 设备上可能存在多份路由图。关注策略路由中引用的路由图，查看其策略设置。

```
Ruijie# show route-map
route-map RM_FOR_PBR, permit, sequence 10
  Match clauses:
    ip address acl1
  Set clauses:
    ip next-hop 200.24.18.1
route-map RM_FOR_PBR, permit, sequence 20
  Match clauses:
    ip address acl2
  Set clauses:
```

```
ip next-hop 200.24.19.1
```

查看 ACL 的配置信息

【命令格式】 **show access-lists** [*acl-id* | *acl-name*]

【参数说明】 *acl-id* : ACL 的编号

acl-name : ACL 的名称

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 设备上可能存在多个 ACL。关注路由图中引用的那个 ACL，查看其配置信息。

```
Ruijie# show access-lists 1
ip access-list standard 1
 10 permit 200.24.16.0 0.0.0.255
ip access-list standard 2
 10 permit 200.24.17.0 0.0.0.255
```

查看 IPv4 策略路由的路由信息

【命令格式】 **show ip pbr route** [*interface if-name* | *local*]

【参数说明】 *if-name* : 接口名称

local : 本地

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 指定接口或设备本地，查看 IPv4 策略路由的路由信息。

```
Ruijie# show ip pbr route
PBR IPv4 Route Summay : 1
Interface      : GigabitEthernet 0/1
Sequence      : 10
Min Length    : None
Max Length    : None
VRF ID       : 0
Route Flags   :
Route Type    : PBR
Direct       : Permit
Priority      : High
Tos_Dscp     : None
Precedence   : None
Tos_Dscp     : 0
Precedence   : 0
Mode         : redundance
NextHop Count : 1
NextHop[0]   : 192.168.8.100
Weight[0]    : 1
Ifindex[0]   : 2
```

查看 IPv4 策略路由引用的路由图

- 【命令格式】 **show ip pbr route-map** *rmap-name*
- 【参数说明】 *rmap-name* : 路由图名称
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 指定路由图，查看 IPv4 策略路由引用的路由图。

```
Ruijie# show ip pbr route-map rm
PBR VRF: GLOBAL, ID: 0
Forward Mode: redundance
Forwarding: On

Route-map rm
Route-map index: Sequence 10, permit
Match rule:
ACL ID : 2900, CLS: 1, Name: acl1
Set rule:
IPv4 nexthop: 192.168.8.100, (VRF name: , ID: 0), Weight: 0
PBR state info ifx: 2, Connected: True, Track state: Up
```

查看 IPv4 策略路由转发的报文个数统计值

- 【命令格式】 **show ip pbr statistics [interface *if-name* | local]**
- 【参数说明】 *if-name* : 接口名称
- local** : 本地
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】

```
Ruijie# show ip pbr statistics
IPv4 Policy-based route statistic
gigabitEthernet 0/1
statistics : 10
```

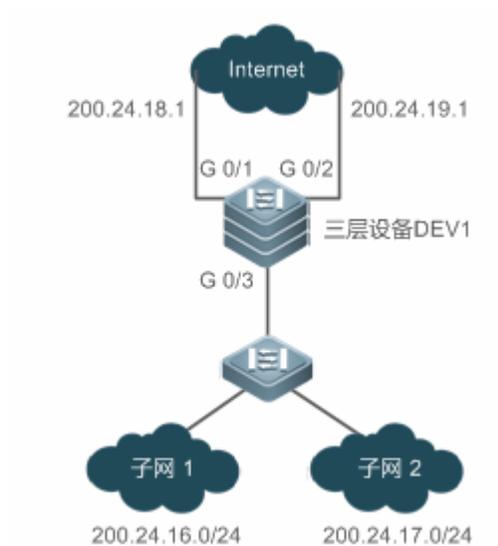
配置举例

i 以下配置举例，仅介绍与策略路由相关的配置。

配置 IPv4 策略路由，根据报文源地址选择出口链路

【网络环境】

图 5-3



三层设备 DEV1 通过 GE0/3 连接子网 1 和子网 2。子网 1 的网段为 200.24.16.0/24，子网 2 的网段为 200.24.17.0/24。

三层设备 DEV1 通过 GE0/1 和 GE0/2 连接因特网，其下一跳为 200.24.18.1、200.24.19.1。

【配置要求】

该局域网有两个出口连接到因特网，要求：

- 来自子网 1 的访问因特网的数据流走 GE 0/1。
- 来自子网 2 的访问因特网的数据流走 GE 0/2。
- 如果 GE 0/1 链路断开，则将 GE 0/1 上的数据流切换到 GE 0/2 上。反之亦然。

【配置方法】

- 配置两个 ACL，分别用来匹配来自子网 1 的报文和来自子网 2 的报文。
- 设置策略，针对来自子网 1 的报文，设置下一跳 GE0/1、GE0/2。（注意配置顺序）
- 设置策略，针对来自子网 2 的报文，设置下一跳 GE0/2、GE0/1。（注意配置顺序）
- 将策略应用在 GE 0/3 上。
- 设置策略路由在多个下一跳之间冗余备份。（缺省即为冗余备份）

i 冗余备份时，按配置顺序，第一个下一跳优先生效。

```
DEV1(config)# access-list 1 permit 200.24.16.0 0.0.0.255
DEV1(config)# access-list 2 permit 200.24.17.0 0.0.0.255
DEV1(config)# route-map RM_FOR_PBR 10
DEV1(config-route-map)# match ip address 1
DEV1(config-route-map)# set ip next-hop 200.24.18.1
DEV1(config-route-map)# set ip next-hop 200.24.19.1
DEV1(config-route-map)# exit
DEV1(config)# route-map RM_FOR_PBR 20
DEV1(config-route-map)# match ip address 2
DEV1(config-route-map)# set ip next-hop 200.24.19.1
DEV1(config-route-map)# set ip next-hop 200.24.18.1
```

```
DEV1(config-route-map)# exit
DEV1(config)# interface GigabitEthernet 0/3
DEV1(config-if-GigabitEthernet 0/3)# ip policy route-map RM_FOR_PBR
DEV1(config-if-GigabitEthernet 0/3)# exit
DEV1(config)# ip policy redundancy
```

- 【检验方法】
- 查看 IPv4 策略路由的配置信息
 - 查看路由图的配置信息
 - 查看 ACL 的配置信息

```
DEV1# show ip policy
Interface                Route map
GigabitEthernet 0/3      RM_FOR_PBR

DEV1# show route-map
route-map RM_FOR_PBR, permit, sequence 10
  Match clauses:
    ip address 1
  Set clauses:
    ip next-hop 200.24.18.1 200.24.19.1
route-map RM_FOR_PBR, permit, sequence 20
  Match clauses:
    ip address 2
  Set clauses:
    ip next-hop 200.24.19.1 200.24.18.1

DEV1# show access-lists
ip access-list standard 1
 10 permit 200.24.16.0 0.0.0.255
ip access-list standard 2
 10 permit 200.24.17.0 0.0.0.255
```

常见错误

- 配置策略路由时引用了路由图，但此路由图实际并不存在。
- 配置路由图时引用了 ACL，但此 ACL 实际并不存在。

5.5.2 设置冗余备份或负载均衡

配置效果

- 使多个下一跳互为备份，可以增强网络的可靠性。

- 在多个下一跳之间负载均衡，可以扩展网络带宽。

注意事项

- 必须配置策略路由基本功能。
- 冗余备份和负载均衡仅对如下 **set** 命令设置的下一跳生效，且仅在使用相同的 **set** 命令设置的多个下一跳之间生效。

命令	说明
set ip next-hop	配置 IPv4 报文的下一跳。此命令携带参数 <i>weight</i> ，用来设定 WCMP 的权重，缺省值 1。
set ip default next-hop	配置 IPv4 报文的缺省下一跳。此命令携带参数 <i>weight</i> ，用来设定 WCMP 的权重。缺省值 1。

i WCMP 最多可以设置 8 个下一跳，ECMP 最多可以设置 32 个下一跳。

配置方法

设置 IPv4 策略路由在多个下一跳之间冗余备份或负载均衡

- 如果需要在多个下一跳之间进行负载均衡，则需要执行此配置。
- 如果当前配置为负载均衡，需要恢复成冗余备份，则也需要执行此配置。
- 此项配置对设备上配置的所有策略路由生效。

【命令格式】 **ip policy { redundancy | load-balance }**

【参数说明】 redundancy：冗余备份
load-balance：负载均衡

【缺省配置】 冗余备份

【命令模式】 全局模式

【使用指导】 选择冗余备份，则按配置顺序，只有第一个下一跳生效。
选择负载均衡，则所有下一跳同时生效，且按权重分担流量。

检验方法

- 查看多个下一跳之间是冗余备份还是负载均衡。

查看 IPv4 策略路由多个下一跳之间是冗余备份或负载均衡

【命令格式】 **show ip policy [route-map-name]**

【参数说明】 *route-map-name*：指定路由图。

【配置模式】 特权模式、全局模式、接口模式

【使用指导】 如下示例，关注红色字段。

```
PE# show ip policy
Banlance mode: redundancy
Interface          Route map
local              test
```

GigabitEthernet 0/3 test

配置举例

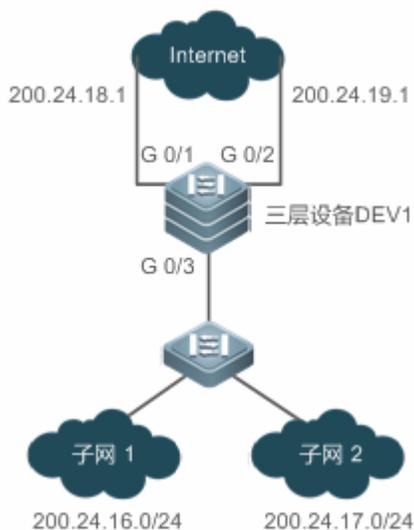
配置 IPv4 策略路由，多个下一跳之间冗余备份

参考上文举例: [配置IPv4 策略路由，根据报文源地址选择出口链路](#)

配置 IPv4 策略路由，多个下一跳之间负载均衡

【网络环境】

图 5-4



三层设备 DEV1 通过 GE0/3 连接子网 1 和子网 2。子网 1 的网段为 200.24.16.0/24，子网 2 的网段为 200.24.17.0/24。

三层设备 DEV1 通过 GE0/1 和 GE0/2 连接因特网，其下一跳为 200.24.18.1、200.24.19.1。

【配置要求】 该局域网有两个出口连接到因特网，要求：GE0/1 口与 GE0/2 口各分担一半的流量。

- 【配置方法】
- 配置策略路由基本功能。指定多个下一跳。
 - 设置负载均衡模式。

```
DEV1(config)# route-map RM_LOAD_PBR 10
DEV1(config-route-map)# set ip next-hop 200.24.18.1
DEV1(config-route-map)# set ip next-hop 200.24.19.1
DEV1(config-route-map)# exit
DEV1(config)# interface GigabitEthernet 0/3
DEV1(config-if-GigabitEthernet 0/3)# ip policy route-map RM_LOAD_PBR
DEV1(config-if-GigabitEthernet 0/3)# exit
DEV1(config)# ip policy load-balance
```

- 【检验方法】
- 查看 IPv4 策略路由的配置信息
 - 查看路由图的配置信息

```

DEV1# show ip policy
Balance mode: load-balance
Interface                Route map
GigabitEthernet 0/3      RM_LOAD_PBR
DEV1# show route-map
route-map PBR-VRF-Selection, permit, sequence 10
Match clauses:
Set clauses:
ip next-hop 200.24.18.1 8
ip next-hop 200.24.19.1 8

```

常见配置错误

无

5.6 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除 IPv4 策略路由转发的报文数统计值	clear ip pbr statistics [interface <i>if-name</i> local]

查看运行情况

作用	命令
查看 IPv4 策略路由的配置信息。	show ip policy
查看路由图的配置信息。	show route-map [<i>name</i>]
查看 ACL 的配置信息。	show access-list
查看 IPv4 策略路由与 BFD 联动信息。	show ip pbr bfd
查看 IPv4 策略路由的路由信息。	show ip pbr route [interface <i>if-name</i> local]
查看 IPv4 策略路由引用的路由图。	show ip pbr route-map <i>rmap-name</i>
查看 IPv4 策略路由的统计信息。	show ip pbr statistics [interface <i>if-name</i> local]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开策略路由的错误调试开关	debug pbr error

打开策略路由的事件相关调试开关	debug pbr events
打开策略路由支持多业务卡相关调试开关	debug pbr ms
打开策略路由的消息通讯调试开关	debug pbr msg
打开策略路由与 NSM 交互的调试开关	debug pbr nsm
打开策略路由的报文转发相关的调试开关	debug pbr packet
打开策略路由的 GR 调试开关	debug pbr restart

6 路由管理

6.1 概述

路由管理负责管理路由表，整合各种路由协议下发的路由，进行优选，并下发给转发表。路由表中保存了各种路由协议发现的路由，根据来源不同，通常分为以下三类：

- 直连路由：链路层协议发现的路由，也称为接口路由。
- 静态路由：网络管理员手工配置的。静态路由配置方便，对系统要求低，适用于拓扑结构简单并且稳定的小型网络。其缺点是每当网络拓扑结构发生变化，都需要手工重新配置，不能自动适应。
- 动态路由：动态路由协议发现的路由。

i 下文仅介绍路由表管理和静态路由的相关内容。

协议规范

无

6.2 典型应用

典型应用	场景描述
静态路由基本功能	手工方式配置路由
静态浮动路由	多路径情况下，配置备份路由
静态负载分担路由	多路径情况下，配置负载分担
静态路由与 BFD 联动	使用 BFD 检测静态路由下一跳是否可达
静态路由快速重路由	场景同浮动路由，为了提高切换性能，可以考虑快速重路由功能

6.2.1 静态路由基本功能

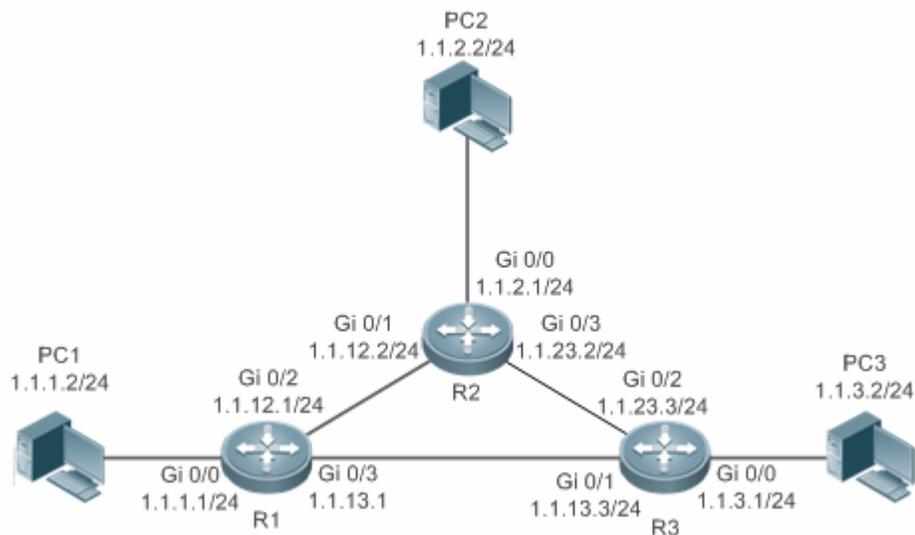
应用场景

在组网结构比较简单的网络中，只需配置静态路由就可以实现网络互通。恰当地设置和使用静态路由可以改善网络的性能，并可为重要的网络应用保证带宽。

以下图为例，为了使得 PC1，PC2，PC3 互通，可以在 R1，R2，R3 上配置静态路由。

- 在 R1 上配置到达 PC2 网段的路由走 R2，配置到达 PC3 网段的路由走 R3
- 在 R2 上配置到达 PC1 网段的路由走 R1，配置到达 PC3 网段的路由走 R3
- 在 R3 上配置到达 PC1 网段的路由走 R1，配置到达 PC2 网段的路由走 R2

图 6-1



功能部属

- 配置各接口的地址和掩码。
- 在 R1，R2，R3 上配置静态路由。

6.2.2 静态浮动路由

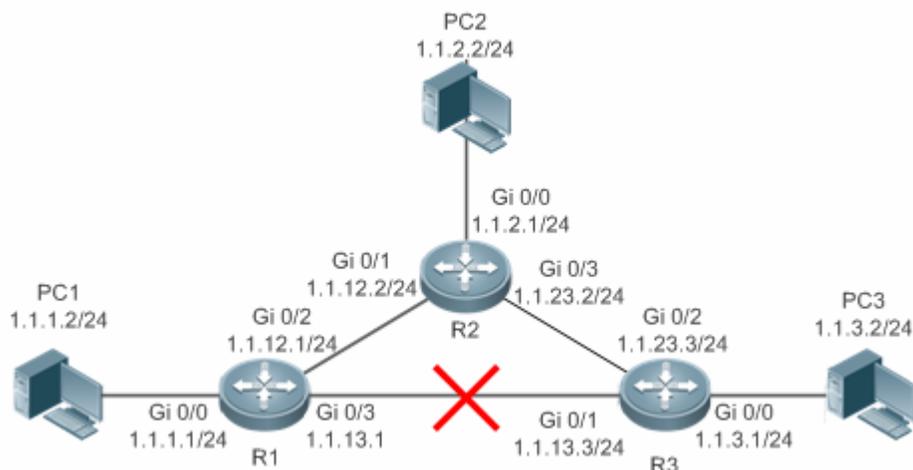
应用场景

在没有配置动态路由协议的情况下，为了避免网络线路故障导致的通信中断，可以配置静态浮动路由，实现路由的动态切换。

以下图为例，为避免 R1 与 R3 间线路故障导致的通信中断，可以在 R1 和 R3 上配置静态浮动路由。正常情况下，报文走管理距离（distance）小的路径，当该路由的链路出现故障 down 掉时，路由自动切换到管理距离大的路径。

- R1 上配置 2 条到达 PC3 网段的路由，一条走 R3（默认 distance=1），另一条走 R2（distance=2）。
- R3 上配置 2 条到达 PC1 网段的路由，一条走 R1（默认 distance=1），另一条走 R2（distance=2）。

图 6-2



功能部属

- 配置各接口的地址和掩码。
- 在 R1，R2，R3 上配置静态路由。

6.2.3 静态负载分担路由

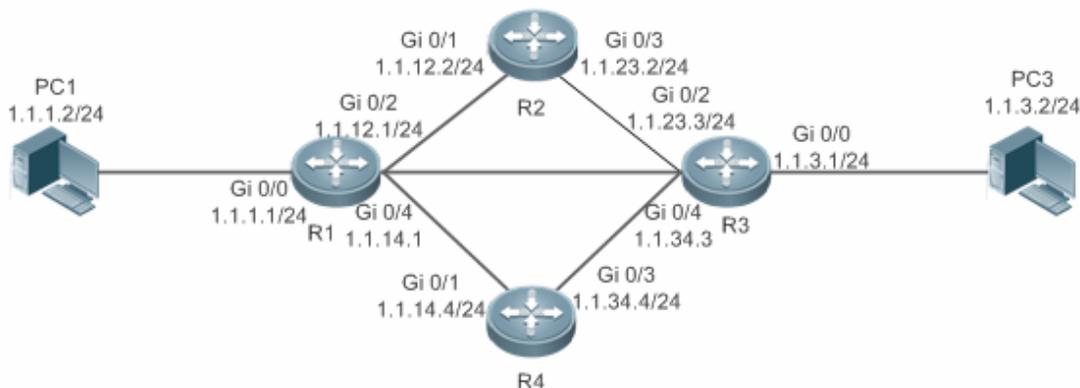
应用场景

在存在多条路径到达同一个目的的情况下，可以配置负载分担路由。与浮动路由不同的是，多条路由的管理距离（distance）相同。报文根据均衡转发策略在多条路由间分流。

以下图为例，在 R1，R3 上配置负载分担路由，使得到达 PC3 和 PC1 网段的报文，在 R2，R4 两条路径间均衡。

- 在 R1 上配置 2 条到达 PC3 网段的路由，一条走 R2，一条走 R4
- 在 R3 上配置 2 条到达 PC1 网段的路由，一条走 R2，一条走 R4

图 6-3



【注释】 交换机上配置命令 `ip ref ecmp load-balance` 指定路由的负载均衡策略
 路由器上配置命令 `ip ref load-sharing original` 指定 v4 路由的负载均衡策略

功能部属

- 配置各接口的地址和掩码。
- 在 R1, R2, R3, R4 上配置静态路由。
- 在 R1, R3 上配置负载分担策略。

6.2.4 静态路由与BFD联动

应用场景

在浮动路由/负载分担的场景下,有时出现故障的线路,接口状态正常,静态路由无法感知路由失效。为了解决这个问题,需要对静态路由下一跳的可达性进行检测,当下一跳不可达时可以切换到备份路由。

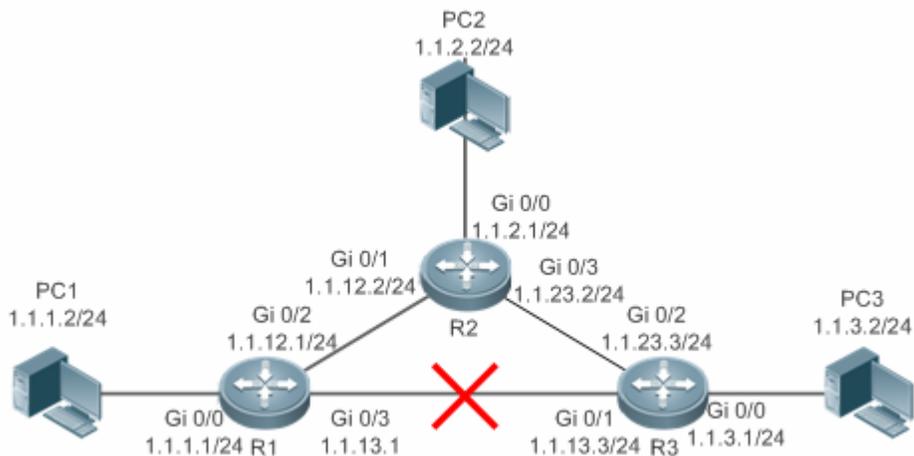
可以通过 BFD 检测静态路由下一跳是否可达。下文以 BFD 为例。

⚠️ 两种方式不能同时使用,只能使用其中的一种。

以下图为例,为避免 R1 与 R3 间线路故障导致的通信中断,在 R1 和 R3 上配置静态浮动路由,并将静态路由关联到 BFD 检测协议。

- R1 上配置 2 条到达 PC3 网段的路由,一条走 R3 (默认 `distance=1`),联动 BFD 检测 1.1.13.3 的可达性,另一条走 R2 (`distance=2`),联动 BFD 检测 1.1.12.2 的可达性。
- R3 上配置 2 条到达 PC1 网段的路由,一条走 R1 (默认 `distance=1`),联动 BFD 检测 1.1.13.1 的可达性,另一条走 R2 (`distance=2`)。联动 BFD 检测 1.1.23.2 的可达性。

图 6-4



功能部属

- 配置各接口的地址和掩码。
- 在各接口上配置 BFD 检测参数。
- 在 R1, R2, R3 上配置静态路由, 与 BFD 联动。

6.2.5 静态快速重路由

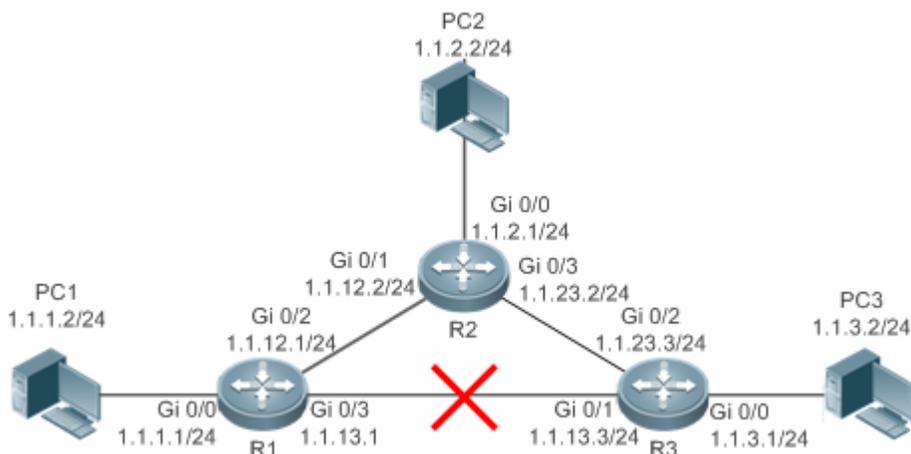
应用场景

在没有配置动态路由协议的情况下, 为了加快路由切换的速度, 减少通信中断时间, 除了可以通过静态路由联动 BFD 的方式进行下一跳检测以外, 还可以通过配置快速重路由, 进一步提升收敛性能。

以下图为例, 为避免 R1 与 R3 间线路故障导致的通信中断, 可以在 R1 和 R3 上配置静态快速重路由。正常情况下, 报文走 R1 与 R3 间的路径, 当该路由的链路出现故障 down 掉时, 路由自动快速重路由到 R2 上。

- 在 R1 上配置: 出口 Gi0/3, 下一跳 1.1.13.3 的路由, 备份出口为 Gi0/2, 备份下一跳为 1.1.12.2
- 在 R3 上配置: 出口 Gi0/1, 下一跳 1.1.13.1 的路由, 备份出口为 Gi0/2, 备份下一跳为 1.1.23.2

图 6-5



功能部属

- 配置各接口的地址和掩码。
- 在 R1，R2，R3 上配置静态路由。
- 在 R1，R3 上配置静态快速重路由。

6.3 功能详解

功能特性	作用
路由计算	在设备上产生可效的路由。
路由优选	在设备选择最优路由，以供报文转发。
缺省路由	使所有报文得以转发，且有助于缩小路由表规模。
路由可靠性	快速发现路由失效，并快速恢复通信。

6.3.1 路由计算

路由功能

路由功能分 IPv4 路由功能。若关闭了路由功能，则设备相当于一台主机，不具备路由转发功能。

动态路由

动态路由协议以邻居间交换路由的方式学习远方的路由、并保持动态更新。如果邻居失效，则下一跳为此邻居的路由随之失效。

静态路由

在组网结构比较简单的网络中，只需配置静态路由就可以实现网络互通。恰当地设置和使用静态路由可以改善网络的性能，并可为重要的网络应用保证带宽。

静态路由根据本地接口的状态计算路由的活动性。当静态路由的出口处于三层 up 状态（链路状态为 up，且配置有 IP 地址）时，该路由为活动的，可以指导转发。

6.3.2 路由优选

管理距离

当多个路由协议产生了到达同一个目的地址的路由时，根据管理距离判断这些路由的优先级。管理距离越小，优先级越高。

等价路由

到达同一个目的地址，下一跳不同，管理距离相同的多条路由，则形成等价路由。报文根据均衡转发策略在多条路由间分流，从而实现负载分担。

具体设备上，对等价路由中包括的路由条目数是有限制的，超出限制的路由不会参与转发。

浮动路由

到达同一目的地址，下一跳不同，管理距离不同的多条路由，形成浮动路由。管理距离小的优先被选择参与转发，若管理距离小的路由失效，则管理距离大的路由替代管理距离小的路由参与转发，从而达到避免网络线路故障导致的通信中断。

6.3.3 缺省路由

在转发路由表中，目的网段 0.0.0.0 掩码 0.0.0.0 的路由，就是缺省路由。无法被其他路由转发的报文，可以被缺省路由转发出去。缺省路由可以静态配置，也可以由动态路由协议生成。

静态缺省路由

三层交换机通过配置网段 0.0.0.0 掩码 0.0.0.0 的静态路由来生成缺省路由。

缺省网络

配置缺省网络的目的是为了产生缺省路由，当在设备上使用 **ip default-network** 指定一个网络（必须为 A 类，B 类，C 类的有类网络）时，这个网络如果在路由表中存在，则路由设备会将该网络作为缺省网络，该网络的下一跳成为缺省网关。因为 **ip default-network** 是有类的，如果使用该命令标记一个主类网络的某个子网，路由设备会自动生成一条主类网络的静态路由而不会产生任何缺省路由。

6.3.4 路由可靠性

当网络中的一台设备发生故障时，会使某些路由不可达，从而造成流量中断。如果能够对路由下一跳的可达性进行实时检测，就可以在故障发生时立即重新计算路由，或者切换到备份路由。

与 BFD 联动

BFD(Bidirectional Forwarding Detection，双向转发检测)协议提供一种轻负载、快速检测两台邻接路由器之间转发路径连通状态的方法。将动态路由协议或静态路由与 BFD 联动，可以使动态路由协议或静态路由快速感知下一跳不可达，从而快速反应。

快速重路由

快速重路由，就是提供了一条备份路由。当动态路由协议或静态路由感知到下一跳不可达时，可以立即切换备份路由，从而快速恢复通信。

6.4 产品说明



S7508 支持 64 组等价路径，每组等价路径最大支持 32 条等价路径条目数。等价路径组是指一个前缀的多个等价的下一跳。

6.5 配置详解

配置项	配置建议 & 相关命令	
配置静态路由	⚠ 必须配置。用于配置静态路由条目。	
	<code>ip route</code>	配置静态 IPv4 路由
配置缺省路由	⚠ 可选配置。用于配置缺省网关。	
	<code>ip route 0.0.0.0 0.0.0.0 gateway</code>	三层设备配置 IPv4 缺省网关
	<code>ip default network</code>	三层设备配置 IPv4 缺省网络
配置路由限制	⚠ 可选配置。用于限制等价路由的条数，静态路由的条数和限制路由。	
	<code>maximum-paths</code>	配置等价路由条数限制
	<code>ip static route-limit</code>	配置静态 IPv4 路由限制
	<code>no ip routing</code>	配置禁止 IPv4 路由
配置静态路由联动 BFD	⚠ 可选配置。用于静态路由联动 BFD。	
	<code>ip route static bfd</code>	配置静态 IPv4 路由联动 BFD
配置静态快速重路由	⚠ 可选配置。用于配置静态快速重路由。	

	route-map	配置路由图
	set fast-reroute backup-nexthop	配置快速重路由的备份接口和备份下一跳
	ip fast-reroute	配置静态快速重路由

6.5.1 配置静态路由

配置效果

路由表中生成一条静态路由。使用静态路由，转发去远端网络的报文。

注意事项

- 三层交换机若配置了 **no ip routing**，则不能配置 IPv4 静态路由，之前已经存在的 IPv4 静态路由也会被删除。在未重启的情况下，重新配置 **ip routing**，可以恢复被清空的 IPv4 静态路由。重启过后，则无法恢复这些 IPv4 静态路由。

配置方法

配置 IPv4 静态路由

在支持 IPv4 的路由器上，配置如下命令。

【命令格式】 `ip route network net-mask { ip-address | interface [ip-address] } [distance] [tag tag] [permanent] [weight number] [description description-text] [disabled | enabled] [global]`

【参数说明】	<i>network</i>	目标网络的网络地址
	<i>net-mask</i>	目标网络的掩码
	<i>ip-address</i>	(可选) 静态路由的下一跳地址。ip-address 和 interface 至少要指定一个，或者两者都指定。当未指定 ip-address 时，表示配置静态直连路由。
	<i>interface</i>	(可选) 静态路由的下一跳出口。ip-address 和 interface 至少要指定一个，或者两者都指定。当未指定 interface 时，表示配置静态递归路由，出口由下一跳在路由表中选路获得。
	<i>distance</i>	(可选) 静态路由的管理距离，缺省为 1。
	<i>tag</i>	(可选) 静态路由的 Tag 值，缺省为 0。
	permanent	(可选) 永久路由标识，缺省为非永久路由。
	weight number	(可选) 静态路由的权重值，缺省为 1。
	description description-text	(可选) 静态路由描述信息，缺省无描述信息，description-text 为 1~60 个字符的字符串。
	disabled/enabled	(可选) 静态路由的使能标识，缺省为 enabled。

【缺省配置】 没配置静态路由

【命令模式】 全局模式

【使用指导】 此命令的最简配置：`ip route network net-mask ip-address`

检验方法

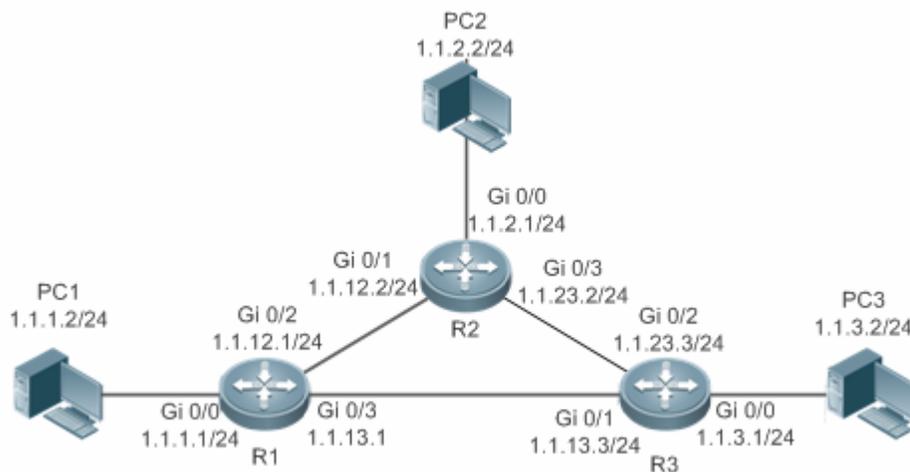
- 使用 **show ip route** 命令查看 IPv4 路由表，检查之前配置的 IPv4 静态路由是否生效。

配置举例

在 IPv4 网络上，配置静态路由使网络联通

【网络环境】

图 6-6



【配置方法】

- 在设备各接口上配置地址

R1

```
R1# configure terminal
R1(config)#interface gigabitEthernet 0/0
R1(config-if-GigabitEthernet 0/0)# ip address 1.1.1.1 255.255.255.0
R1(config-if-GigabitEthernet 0/0)# exit
R1(config)#interface gigabitEthernet 0/2
R1(config-if-GigabitEthernet 0/2)# ip address 1.1.12.1 255.255.255.0
R1(config-if-GigabitEthernet 0/0)# exit
R1(config)#interface gigabitEthernet 0/3
R1(config-if-GigabitEthernet 0/3)# ip address 1.1.13.1 255.255.255.0
```

R2

```
R2# configure terminal
R2(config)#interface gigabitEthernet 0/0
R2(config-if-GigabitEthernet 0/0)# ip address 1.1.2.1 255.255.255.0
R2(config-if-GigabitEthernet 0/0)# exit
R2(config)#interface gigabitEthernet 0/1
R2(config-if-GigabitEthernet 0/1)# ip address 1.1.12.2 255.255.255.0
R2(config-if-GigabitEthernet 0/0)# exit
R2(config)#interface gigabitEthernet 0/3
R2(config-if-GigabitEthernet 0/3)# ip address 1.1.23.2 255.255.255.0
```

R3

```
R3# configure terminal
R3(config)#interface gigabitEthernet 0/0
```

```
R3(config-if-GigabitEthernet 0/0)# ip address 1.1.3.1 255.255.255.0
R3(config-if-GigabitEthernet 0/0)# exit
R3(config)#interface gigabitEthernet 0/1
R3(config-if-GigabitEthernet 0/1)# ip address 1.1.13.3 255.255.255.0
R3(config-if-GigabitEthernet 0/0)# exit
R3(config)#interface gigabitEthernet 0/2
R3(config-if-GigabitEthernet 0/2)# ip address 1.1.23.3 255.255.255.0
```

- 在设备上配置静态路由

```
R1 R1# configure terminal
R1(config)# ip route 1.1.2.0 255.255.255.0 GigabitEthernet 0/2 1.1.12.2
R1(config)# ip route 1.1.3.0 255.255.255.0 GigabitEthernet 0/3 1.1.13.3

R2 R2# configure terminal
R2(config)# ip route 1.1.1.0 255.255.255.0 GigabitEthernet 0/1 1.1.12.1
R2(config)# ip route 1.1.3.0 255.255.255.0 GigabitEthernet 0/3 1.1.23.3

R3 R3# configure terminal
R3(config)# ip route 1.1.2.0 255.255.255.0 GigabitEthernet 0/2 1.1.23.2
R3(config)# ip route 1.1.1.0 255.255.255.0 GigabitEthernet 0/1 1.1.13.1
```

【检验方法】 显示路由表

```
R1 R1# show ip route
Codes: C - Connected, L - Local, S - Static
       R - RIP, O - OSPF, B - BGP, I - IS-IS
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       IA - Inter area, * - candidate default

Gateway of last resort is no set
C    1.1.1.0/24 is directly connected, GigabitEthernet 0/0
C    1.1.1.1/32 is local host.
S    1.1.2.0/24 [1/0] via 1.1.12.2, GigabitEthernet 0/2
S    1.1.3.0/24 [1/0] via 1.1.13.3, GigabitEthernet 0/2
C    1.1.12.0/24 is directly connected, GigabitEthernet 0/2
C    1.1.12.1/32 is local host.
C    1.1.13.0/24 is directly connected, GigabitEthernet 0/3
C    1.1.13.1/32 is local host.

R2 R2# show ip route
Codes: C - Connected, L - Local, S - Static
       R - RIP, O - OSPF, B - BGP, I - IS-IS
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
```

```
SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
IA - Inter area, * - candidate default
```

```
Gateway of last resort is no set
```

```
S 1.1.1.0/24 [1/0] via 1.1.12.1, GigabitEthernet 0/0
C 1.1.2.0/24 is directly connected, GigabitEthernet 0/0
C 1.1.2.1/32 is local host.
S 1.1.3.0/24 [1/0] via 1.1.23.3, GigabitEthernet 0/3
C 1.1.12.0/24 is directly connected, GigabitEthernet 0/1
C 1.1.12.2/32 is local host.
C 1.1.23.0/24 is directly connected, GigabitEthernet 0/3
C 1.1.23.2/32 is local host.
```

R3

```
R3# show ip route
```

```
Codes: C - Connected, L - Local, S - Static
```

```
R - RIP, O - OSPF, B - BGP, I - IS-IS
```

```
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
```

```
E1 - OSPF external type 1, E2 - OSPF external type 2
```

```
SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
```

```
IA - Inter area, * - candidate default
```

```
Gateway of last resort is no set
```

```
S 1.1.1.0/24 [1/0] via 1.1.13.1, GigabitEthernet 0/2
S 1.1.2.0/24 [1/0] via 1.1.23.2, GigabitEthernet 0/2
C 1.1.3.0/24 is directly connected, GigabitEthernet 0/0
C 1.1.3.1/32 is local host.
C 1.1.13.0/24 is directly connected, GigabitEthernet 0/1
C 1.1.13.3/32 is local host.
C 1.1.23.0/24 is directly connected, GigabitEthernet 0/2
C 1.1.23.3/32 is local host.
```

常见错误

- 接口链路没有 up
- 接口没有配置地址

6.5.2 静态配置缺省路由

配置效果

路由表中生成一条缺省路由。不能被其他路由转发的报文，使用缺省路由转发。

注意事项

- 三层交换机可以通过 `ip route 0.0.0.0 0.0.0.0 gateway` 命令配置网关。
- 三层交换机若配置了 `no ip routing`，则可通过 `ip default gateway` 命令配置网关。

配置方法

三层交换机配置 IPv4 缺省网关

【命令格式】 `ip route 0.0.0.0 0.0.0.0 { ip-address | interface [ip-address] } [distance] [tag tag] [permanent] [weight number] [description description-text] [disabled | enabled] [global]`

【参数说明】	0.0.0.0	目标网络的网络地址
	0.0.0.0	目标网络的掩码
	<i>ip-address</i>	(可选) 静态路由的下一跳地址。ip-address 和 interface 至少要指定一个，或者两者都指定。当未指定 ip-address 时，表示配置静态直连路由。
	<i>interface</i>	(可选) 静态路由的下一跳出口。ip-address 和 interface 至少要指定一个，或者两者都指定。当未指定 interface 时，表示配置静态递归路由，出口由下一跳在路由表中选路获得。
	<i>distance</i>	(可选) 静态路由的管理距离，缺省为 1。
	<i>tag</i>	(可选) 静态路由的 Tag 值，缺省为 0。
	permanent	(可选) 永久路由标识，缺省为非永久路由。
	weight number	(可选) 静态路由的权重值，缺省为 1。
	description description-text	(可选) 静态路由描述信息，缺省无描述信息， <i>description-text</i> 为 1~60 个字符的字符串。
	disabled / enabled	(可选) 静态路由的使能标识，缺省为 enabled。

【缺省配置】 无静态缺省路由

【命令模式】 全局模式

【使用指导】 此命令的最简配置：`ip route 0.0.0.0 0.0.0.0 ip-address`

三层交换机配置 IPv4 缺省网络

【命令格式】 `ip default-network network`

【参数说明】 *network* | 网络地址 (必须为 A 类, B 类, C 类的有类网络)

【缺省配置】 无缺省网络

【命令模式】 全局模式

【使用指导】 如果 `ip default-network` 指定的网络在路由表中存在，则生成一条缺省路由，以到达该网络的下一跳为缺省网关。否则不产生缺省路由。

检验方法

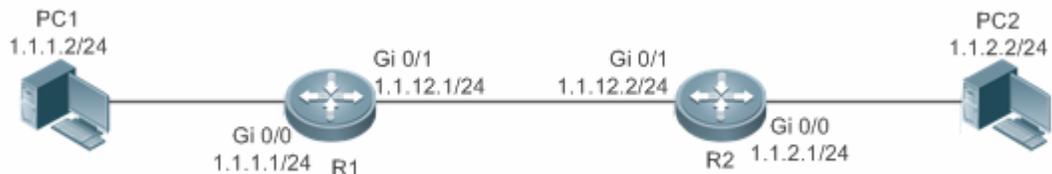
- 在 (未关闭路由功能的) 三层交换机上，使用 `show ip route` 命令查看缺省路由。

配置举例

三层交换机，配置 IPv4 缺省路由，使网络连通

【网络环境】

图 6-7



【配置方法】

● 在三层设备上配置 IP 地址

R1

```
R1# configure terminal
R1(config)#interface gigabitEthernet 0/0
R1(config-if-GigabitEthernet 0/0)# ip address 1.1.1.1 255.255.255.0
R1(config-if-GigabitEthernet 0/0)# exit
R1(config)#interface gigabitEthernet 0/1
R1(config-if-GigabitEthernet 0/1)# ip address 1.1.12.1 255.255.255.0
R1(config-if-GigabitEthernet 0/1)# exit
```

R2

```
R2# configure terminal
R2(config)#interface gigabitEthernet 0/0
R2(config-if-GigabitEthernet 0/0)# ip address 1.1.2.1 255.255.255.0
R2(config-if-GigabitEthernet 0/0)# exit
R2(config)#interface gigabitEthernet 0/1
R2(config-if-GigabitEthernet 0/1)# ip address 1.1.12.2 255.255.255.0
R2(config-if-GigabitEthernet 0/1)# exit
```

R1

● 在三层设备 R1 上配置缺省网关

```
R1# configure terminal
R1(config)# ip route 0.0.0.0 0.0.0.0 GigabitEthernet 0/1 1.1.12.2
```

R2# configure terminal

R2

```
R2(config)# ip route 0.0.0.0 0.0.0.0 GigabitEthernet 0/1 1.1.12.1
```

【检验方法】

显示路由表

R1

```
R1# show ip route
Codes: C - Connected, L - Local, S - Static
       R - RIP, O - OSPF, B - BGP, I - IS-IS
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       IA - Inter area, * - candidate default
```

```
Gateway of last resort is 1.1.12.2
S*   0.0.0.0/0 [1/0] via 1.1.12.2, GigabitEthernet 0/1
C    1.1.1.0/24 is directly connected, GigabitEthernet 0/0
C    1.1.1.1/32 is local host.
C    1.1.12.0/24 is directly connected, GigabitEthernet 0/1
C    1.1.12.1/32 is local host.
```

6.5.3 配置路由限制

配置效果

限制等价路由的条数，静态路由的条数或限制路由转发

注意事项

-

配置方法

配置等价路由的条数

【命令格式】 **maximum-paths** *number*

【参数说明】 *number* | 等价路由条数，范围 1-32，缺省值 32

【缺省配置】 等价路由条数为 32 条

【命令模式】 全局模式

【使用指导】 通过该命令限制等价路由中下一跳的数目。配置等价路由条数后，在负载均衡模式下，负载均衡的分路数不会超过配置的等价路由数。

配置静态 IPv4 路由限制

【命令格式】 **ip static route-limit** *number*

【参数说明】 *number* | 路由上限，范围 1-10000，缺省值 1024

【缺省配置】 允许配置的静态路由条数最大值为 IP 路由 1024 条

【命令模式】 全局模式

【使用指导】 使用此命令配置 IPv4 静态路由最大条数。超过了最大条数值后，IPv4 静态路由配置不成功。

配置禁止 IPv4 路由转发

【命令格式】 **no ip routing**

【参数说明】 -

【缺省配置】 IP 路由功能开启

【命令模式】 全局模式

【使用指导】 使用此命令关闭 IPv4 路由。当设备只作为桥接设备，或者只作为 VOIP 网关设备时，可以不需要 RGOS 软件的 IPv4 路由转发功能。这时可以关闭 RGOS 的 IPv4 路由功能。

检验方法

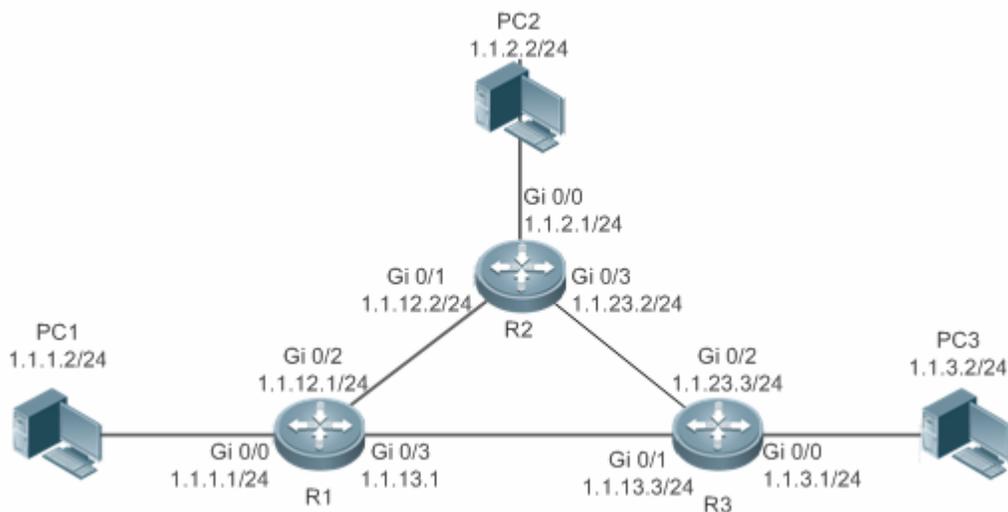
使用 **show run** 命令查看配置文件，确认存在以上配置命令。

配置举例

配置静态路由限制，不超过 2 条

【网络环境】

图 6-8



【配置方法】 在设备 R1 上配置 ip 地址、静态路由、静态路由数量限制。

```
R1# configure terminal
R1(config)#interface gigabitEthernet 0/0
R1(config-if-GigabitEthernet 0/0)# ip address 1.1.1.1 255.255.255.0
R1(config-if-GigabitEthernet 0/0)# exit
R1(config)#interface gigabitEthernet 0/2
R1(config-if-GigabitEthernet 0/2)# ip address 1.1.12.1 255.255.255.0
R1(config-if-GigabitEthernet 0/0)# exit
R1(config)#interface gigabitEthernet 0/3
R1(config-if-GigabitEthernet 0/3)# ip address 1.1.13.1 255.255.255.0
R1(config-if-GigabitEthernet 0/3)# exit
R1(config)#ip route 1.1.3.0 255.255.255.0 1.1.13.3
R1(config)#ip route 1.1.4.0 255.255.255.0 1.1.12.2
R1(config)#ip route 1.1.5.0 255.255.255.0 1.1.12.2
R1(config)# ip static route-limit 2
```

```
% Exceeding maximum static routes limit.
```

【检验方法】 查看路由表中实际生效的静态路由。

```
R1(config)# show ip route
Codes: C - Connected, L - Local, S - Static
       R - RIP, O - OSPF, B - BGP, I - IS-IS
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       SU - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       IA - Inter area, * - candidate default

Gateway of last resort is no set
C    1.1.1.0/24 is directly connected, GigabitEthernet 0/0
C    1.1.1.1/32 is local host.
S    1.1.3.0/24 [1/0] via 1.1.13.3
S    1.1.4.0/24 [1/0] via 1.1.12.2
C    1.1.12.0/24 is directly connected, GigabitEthernet 0/2
C    1.1.12.1/32 is local host.
C    1.1.13.0/24 is directly connected, GigabitEthernet 0/3
C    1.1.13.1/32 is local host.
```

常见错误

6.5.4 配置静态路由联动BFD

配置效果

静态路由在 BFD 的帮助下，能够快速发现路由失效。

注意事项

- 二层交换机不能配置
- 必须配置静态路由
- 必须配置 BFD 会话参数。使用 `bfd interval x min_rx x multiplier x` 命令。

配置方法

配置 IPv4 静态路由与 BFD 联动

【命令格式】 **ip route static bfd** *interface-type interface-number gateway* [**source ip-address**]

【参数说明】	<i>interface-type</i>	配置接口类型
	<i>interface-number</i>	配置接口编号
	<i>gateway</i>	配置网关 IP，即为 BFD 的邻居 IP。静态路由配置下一跳为该邻居的将通过 BFD 进行检测该转发路径的可达性。
	source ip-address	(可选)配置 BFD 会话所采用的源 IP 地址,如果邻居 IP 为多跳的情况下,需要配置该参数。缺省不指定源 IP 地址。

【缺省配置】 没配置静态路由联动 BFD

【命令模式】 全局模式

【使用指导】 使用此命令配置 IPV4 静态路由与 BFD 联动。若 BFD 会话检测出 down 状态，则 IPV4 静态路由不活动，不参与路由转发。

检验方法

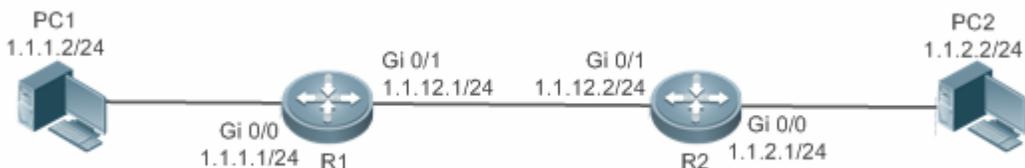
- 使用 **show bfd neighbors** 命令查看 BFD 邻居。
- 通过 **show ip route static bfd** 查看静态路由联动 BFD 的情况

配置举例

配置 IPv4 静态路由联动 BFD

【网络环境】

图 6-9



- 【配置方法】
- 在 R1、R2 互联接口上配置 BFD 会话。
 - 在 R1、R2 上配置静态路由，指定出接口/下一跳为互联接口。
 - 在 R1、R2 上配置静态路由与 BDF 联动，检测静态路由下一跳的连通性。

R1

```

R1# configure terminal
R1(config)#interface gigabitEthernet 0/1
R1(config-if-GigabitEthernet 0/1)# no switchport
R1(config-if-GigabitEthernet 0/1)# ip address 1.1.12.1 255.255.255.0
R1(config-if-GigabitEthernet 0/1)# bfd interval 50 min_rx 50 multiplier 3
R1(config-if-GigabitEthernet 0/1)# exit
R1(config)# ip route 1.1.2.0 255.0.0.0 FastEthernet 0/1 1.1.12.2
R1(config)# ip route static bfd gigabitEthernet 0/1 1.1.12.2
  
```

R2

```

R2# configure terminal
R1(config)#interface gigabitEthernet 0/1
  
```

```
R1(config-if-GigabitEthernet 0/1)# no switchport
R1(config-if-GigabitEthernet 0/1)# ip address 1.1.12.2 255.255.255.0
R1(config-if-GigabitEthernet 0/1)# bfd interval 50 min_rx 50 multiplier 3
R1(config-if-GigabitEthernet 0/1)# exit
R1(config)# ip route 1.1.1.0 255.0.0.0 FastEthernet 0/1 1.1.12.1
R1(config)# ip route static bfd gigabitEthernet 0/1 1.1.12.1
```

- 【检验方法】
- 查看 BFD 邻居状态
 - 查看与 BFD 联动的静态路由

R1

```
R1# show bfd neighbors
OurAddr      NeighAddr    LD/RD      RH/RS      Holddown(mult)  State  Int
1.1.12.1     1.1.12.2    8192/0     Up         0(3 )          Up     GigabitEthernet 0/1

R1#show ip route static bfd
S      1.1.2.0/24 via 1.1.12.2, GigabitEthernet 0/1, BFD state is Up
```

常见错误

- 接口链路没有 up
- 接口没有配置 IP 地址
- 没有配置 BFD 会话参数
- 没有配置静态路由

6.5.5 配置静态快速重路由

配置效果

配置启用静态快速重路由

注意事项

- 二层交换机不能配置
- 必须配置静态路由
- 必须配置路由图

配置方法

- ↳ 在路由图中定义备份路由

【命令格式】 **set fast-reroute backup-nexthop** interface ip-address

【参数说明】 interface 备份的出接口。

ip-address 备份的下一跳。

【缺省配置】 无

【命令模式】 全局模式

【使用指导】 使用 **route-map name [permit | deny] sequence** 命令创建一个路由图。

使用 **match** 命令定义匹配条件。

使用 **set fast-reroute backup-nexthop interface ip-address** 命令定义备份的出接口和下一跳。

若某条路由满足匹配条件，则为该路由生成一条备份路由。若不配置 **march** 命令，则无条件匹配。

配置静态快速重路由，引用路由图

【命令格式】 **ip fast-reroute static route-map** route-map-name

【参数说明】 route-map-name 定义了备份路由的路由图

【缺省配置】 没配置静态路由快速重路由

【命令模式】 全局模式

【使用指导】 使用此命令配置静态快速重路由，引用路由图。

检验方法

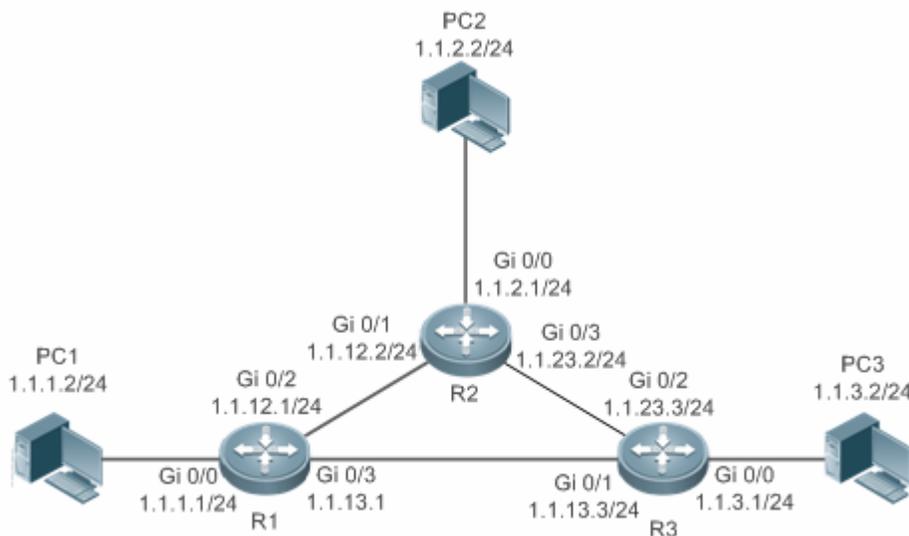
通过 **show ip route fast-reroute** 查看生效的主备路由。

配置举例

配置快速重路由

【网络环境】

图 6-10



【配置方法】 在 R1 上配置到达 PC3 网段的静态路由，出接口下一跳指向 R3。

在 R1 上配置静态快速重路由，备份路由由出接口下一跳指向 R2。

```
R1# configure terminal
R1(config)#interface gigabitEthernet 0/0
R1(config-if-GigabitEthernet 0/0)# ip address 1.1.1.1 255.255.255.0
R1(config-if-GigabitEthernet 0/0)# exit
R1(config)#interface gigabitEthernet 0/2
R1(config-if-GigabitEthernet 0/2)# ip address 1.1.12.1 255.255.255.0
R1(config-if-GigabitEthernet 0/2)# exit
R1(config)#interface gigabitEthernet 0/3
R1(config-if-GigabitEthernet 0/3)# ip address 1.1.13.1 255.255.255.0
R1(config-if-GigabitEthernet 0/3)# exit
R1(config)# ip route 1.1.3.0 255.255.255.0 GigabitEthernet 0/3 1.1.13.3
R1(config)# route-map fast-reroute
R1(config-route-map)# set fast-reroute backup-interface GigabitEthernet 0/2 backup-nexthop
1.1.12.2
R1(config-route-map)# exit
R1(config)# ip fast-reroute static route-map fast-reroute
```

【检验方法】

显示 R1 上的主备路由

```
R1# show ip route fast-reroute
Codes: C - connected, S - static, R - RIP, B - BGP
       O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default
       Status codes: m - main entry, b - backup entry, a - active entry

Gateway of last resort is no set

S    1.1.3.0 /24 [ma] via 1.1.13.3, GigabitEthernet 0/3
      [b] via 1.1.12.2, GigabitEthernet 0/2
```

常见错误

- 接口链路没有 up
- 未配置静态路由
- 路由图中未设置 **match** 匹配条件，或设置错误

6.6 监视与维护

查看运行情况

作用	命令
查看 IPv4 路由表。	show ip route

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 v4 路由管理的调试开关	debug nsm kernel ucast- v4
打开快速重路由管理的调试开关	debug nsm kernel frr
打开缺省网络管理的调试开关	debug nsm kernel default-network
打开路由管理内部事件调试开关	debug nsm events
打开路由管理与路由协议消息发送开关	debug nsm packet send
打开路由管理与路由协议消息接收开关	debug nsm packet recv



配置指南-组播

本分册介绍组播配置指南相关内容，包括以下章节：

1. IP 组播
2. IGMP
3. PIM-DM
4. PIM-SM
5. IGMP Snooping

1 IP 组播

1.1 概述

IP 组播 (IP Multicasting) 是对硬件组播的抽象，是对标准 IP 网络层协议的扩展内组播路由协议。

传统的 IP 传输只允许一台主机向单个主机 (单播通信) 或者所有主机 (广播通信) 发送报文，组播技术则提供第三种选择：允许一台主机向某些主机发送报文。

IP 组播技术适用于“一对多”的多媒体应用。

协议规范

1.2 典型应用

典型应用	场景描述
PIM-DM典型应用	同一网络内部提供 PIM-DM 组播服务。
PIM-SM典型应用	同一网络内部提供 PIM-SM 组播服务。

1.2.1 PIM-DM典型应用

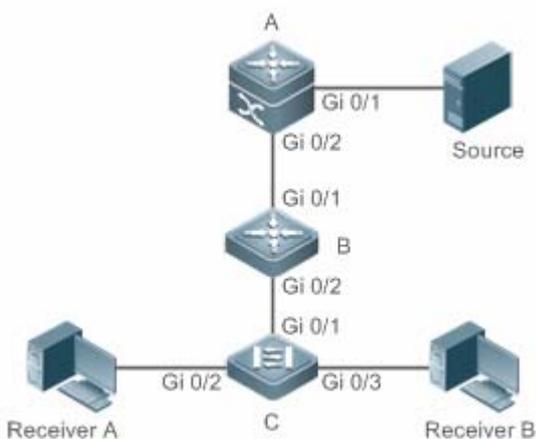
应用场景

同一网络内部提供 PIM-DM 组播服务。

以下图为例

- 组播源发出的组播报文，同一网络内的 Receiver A 与 Receiver B 接收组播源发出的组播报文。

图 1-1



【注释】 A 和 B 为三层设备, C 为二层接入设备。

将 Source 接入 A 的 Gi0/1, 将 Receiver A, Receiver B 接入 C 的 Gi0/2, Gi0/3。

功能部属

- 在同一网络中运行 OSPF 协议, 实现单播路由。
- 在同一网络中运行 PIM-DM 协议, 实现组播路由。
- 在用户主机网段运行 IGMP 协议, 实现组成员管理。

1.2.2 PIM-SM典型应用

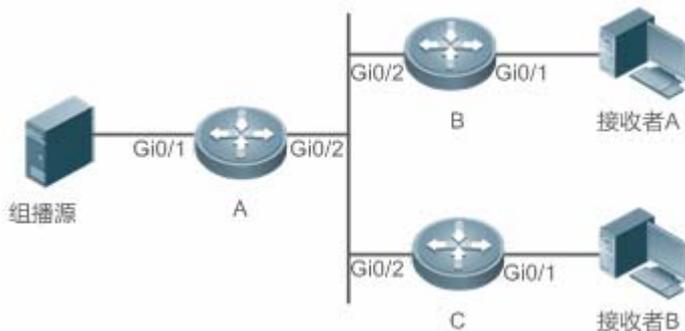
应用场景

同一网络内部提供 PIM-SM 组播服务。

以下图为例

- 组播源发出的组播报文, 同一网络内的 Receiver A 与 Receiver B 接收组播源发出的组播报文。

图 1-2



【注释】 A、B、C 为三层路由设备。

将组播源接入 A 的 Gi0/1, 将接收者 A 接入 B 的 Gi0/1, 将接收者 B 接入 C 的 Gi0/1。

功能部属

- 在同一网络中运行 OSPF 协议, 实现单播路由。
- 在同一网络中运行 PIM-SM 协议, 实现组播路由。
- 在用户主机网段运行 IGMP 协议, 实现组成员管理。

1.3 功能详解

基本概念

▾ PIM 路由器、PIM 接口

启用了 PIM 协议的路由器称为 PIM 路由器。使能了 PIM 协议的接口称为 PIM 接口。

在 PIM 路由器上执行组播报文的转发。接收组播报文的 PIM 接口称为上游接口；发出组播报文的 PIM 接口称为下游接口。上游接口所在网段称为上游网段。下游接口所在网段称为下游网段。

▾ PIM 网络、PIM 域

PIM 路由器之间通过 PIM 接口相连，组成 PIM 网络。

在某些 PIM 接口上设置“边界”，可以将一个大的 PIM 网络划分为多个 PIM 域。“边界”可以拒绝特定组播报文通过，或者限制 PIM 消息的传输。

▾ 组播分发树、DR、RP

组播报文从一点传输到多点，转发路径呈现树型结构。此转发路径被称为组播分发树（MDT，Multicast Distribution Tree），有两种类型：

- RPT（RP Tree，共享树）：以 RP（Rendezvous Point）为根，以连接组成员的 DR 为叶子。
- SPT（Shortest Path Tree，最短路径树）：以连接组播源的 DR（Designated Router）为根，以 RP 或连接组成员的 DR 为叶子。

DR、RP 都是 PIM 路由器的功能角色。

- RP 负责收集网络中的组播源信息，组成员信息。
- 连接组播源的 DR 负责向 RP 汇报组播源的信息。连接组成员的 DR 负责向 RP 汇报组成员的信息。

▾ (*,G)、(S,G)

- (*,G)：任意源发往组 G 的报文、此类报文对应的路由表项、此类报文对应的转发路径（RPT）。
- (S,G)：源 S 发往组 G 的报文、此类报文对应的路由表项、此类报文对应的转发路径（SPT）。

▾ ASM、SSM

PIM-SM 支持两种组播服务模型——ASM 和 SSM，适用不同的组播地址段。

- ASM（Any-Source Multicast，任意源组播）：在 ASM 模型中，用户主机不能选择组播源。用户主机加入组，接收所有源发往该组的报文。
- SSM（Source-Specific Multicast，指定源组播）：在 SSM 模型中，用户主机可以选择组播源。用户主机加入组时指定源地址，仅接收指定源发往该组的报文。

i SSM 模型要求：必须通过其他方式的网络服务，使用户主机能够预先知道组播源的位置，用户主机才有可能对组播源进行选择。

功能特性

功能特性	作用
配置 IP 组播基本功能	建立 PIM 网络，向网络内的数据源和用户终端提供 IPv4 组播服务。
配置 TTL 阈值	配置接口的 TTL 阈值（接口上允许通过的组播报文 TTL 的最小值）。
配置能加入到 IP 组播路由表中的条目数量	限制能够加入到组播路由表中的条目数量。

配置IP组播边界	设置接口成为特定组范围的组播边界。
配置IP组播静态路由	组播静态路由允许组播转发路径不同于单播路径。
配置组播流二层流向控制	对于某一条组播流可以配置多条命令，也就是说配置多个被允许转发的端口。一旦为某组播流配置了流向控制，该组播流只可能由这些已配置的端口转发出去。其他未被允许的端口将不允许转发组播流。
配置按照最长匹配选择RPF路由	遵循 RPF 规则分别从组播静态路由表、MBGP 路由表、单播路由表中选出一条最优路由。从这三条路由中选出掩码最长匹配的那条路由，作为 RPF 路由。
配置组播不间断转发参数	正常运行状态下，SSP 将硬件组播转发表实时同步到从管理板。当管理板切换后，原从管理板组播控制面配置命令加载，组播协议（如 PIM-SM，IGMP Snooping 等）重新收敛。组播不间断转发（Non-Stop Forwarding）功能保证在组播协议重新收敛的这段时间内，组播数据流的转发不间断。
配置组播硬件表项溢出覆盖机制	创建组播转发表项时，如果硬件转发表项资源已满，将删除最早创建的硬件表项资源，添加新增加的表项。

1.3.1 配置IP组播基本功能

建立 PIM 网络，向网络内的数据源和用户终端提供 IPv4 组播服务。

工作原理

设备通过执行组播路由协议（例如 PIM-DM，PIM-SM 等）来维护转发组播报文的路由表，通过 IGMP 协议来学习在直连网段上组成员的状态。主机通过发送 IGMP Report 消息来加入特定的 IGMP 组。

相关配置

启动 IPv4 组播路由功能

缺省情况下，IPv4 组播路由功能关闭。

使用 `ip multicast-routing` 命令可以启动 IPv4 组播路由功能。

在接口上配置 IP 组播协议

缺省情况下，接口上的 IP 组播协议关闭。

使用 `ip pim sparse-mode`、`ip pim dense-mode` 等命令开启接口上的 IPv4 组播协议。

1.3.2 配置TTL阈值

配置接口的 TTL 阈值（接口上允许通过的组播报文 TTL 的最小值）。

工作原理

通过配置接口 TTL 阈值，检查组播报文的 TTL。如果组播报文的 TTL 大于接口阈值，则可以转发报文，否则丢弃该报文。

相关配置

配置 TTL 阈值

缺省情况下，接口上的 TTL 阈值为 0。

使用 `ip multicast ttl-threshold ttl-value` 命令可以调整接口上的 TTL 阈值，取值范围是 <0-255>。

ttl-value 值越大，允许转发的组播报文的 TTL 值需要越大。*littl-value* 值越小，允许转发的组播报文的 TTL 值需要越小。

1.3.3 配置能加入到IP组播路由表中的条目数量

设备上每个收到的组播数据报文都维护相应的 IP 组播路由转发表项。但是，过量的组播路由表项可能会耗尽设备内存，从而导致设备性能下降。用户可以根据实际组网情况和业务性能要求对 IP 组播路由表中的表项数量进行限制。

工作原理

根据实际组网情况及业务性能要求，限制 IP 组播路由表中表项数量，保证设备性能。

相关配置

配置能加入到 IP 组播路由表中的条目数量

缺省情况下，能够加入到 IP 组播路由表中的表项数量为 1024。

使用 `ip multicast route-limit limit [threshold]` 命令可以调整能够加入到 IP 组播路由表中的表项数量，取值范围是 <1-65536>。

limit 值越大，允许加入到 IP 组播路由表中的表项数量越大。*limit* 值越小，允许加入到 IP 组播路由表中的表项数量越小。

1.3.4 配置IP组播边界

配置 IP 组播边界，确定组播报文的传递范围。

工作原理

通过设定组播边界，指定组播传递范围。当在一个接口上配置了组播转发边界后，将不能从该接口转发组播报文（包括本机发出的组播报文），也不能从该接口接收组播报文。

相关配置

配置 IP 组播边界

缺省情况下，没设置组播边界。

使用 `ip multicast boundary access-list [in | out]` 命令可以设置组播边界。

1.3.5 配置IP组播静态路由

通过配置 IP 组播静态路由，可以为来自特定组播源的组播报文指定 RPF 接口或 RPF 邻居。

工作原理

组播报文转发时都会进行 RPF 检查，通过配置 IP 组播静态路由，可以为来自特定组播源的组播报文指定 RPF 接口或 RPF 邻居。

相关配置

配置 IP 组播静态路由

缺省情况下，无组播静态路由配置。

使用 `ip mroute source-address mask { [ospf | rip | static] { v4rpf-address | interface-type interface-number } } [distance]` 命令可以配置 IP 组播静态路由。

1.3.6 配置组播流二层流向控制

通过配置组播流二层流向控制，控制组播流在端口上的转发行为。

工作原理

通过配置组播流二层流向控制，配置允许转发出口，组播流只可能由已配置的端口转发出去，达到控制组播流二层转发。

相关配置

配置组播流二层流向控制

缺省情况下，组播流二层流向控制关闭。

使用 `ip multicast static source-address group-address interface-type interface-number` 命令可以配置组播流二层流向控制。

1.3.7 配置按照最长匹配选择RPF路由

在组播静态路由表、MBGP 路由表、单播路由表中选取的最优路由中选取掩码最长匹配的路由，作为 RPF 路由。。

工作原理

遵循 RPF 规则分别从组播静态路由表、MBGP 路由表、单播路由表中选取可用于 RPF 检查的组播静态路由、MBGP 路由和单播路由。

- 如果设置了按照最长匹配选择路由，则从这三条路由中选出掩码最长匹配的那条路由，作为 RPF 路由。如果这三条路由的掩码一样，则选择其中优先级最高的那条路由；如果它们的优先级也相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。
- 否则，从这三条路由中选出优先级最高的那条路由；如果它们的优先级相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。

相关配置

配置按照最长匹配选择 RPF 路由

缺省情况下，选择优先级最高的路由作为 RPF 路由。如果优先级相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。。

使用 `ip multicast rpf longest-match` 命令可以配置按照最长匹配选择 RPF 路由。

1.3.8 配置组播不间断转发参数

组播不间断转发（Non-Stop Forwarding）功能保证在组播协议重新收敛的这段时间内，组播数据流的转发不间断。

工作原理

正常运行状态下，SSP 将硬件组播转发表实时同步到从管理板。当管理板切换后，原从管理板组播控制面配置命令加载，组播协议（如 PIM-SM，IGMP Snooping 等）重新收敛。组播不间断转发（Non-Stop Forwarding）功能保证在组播协议重新收敛的这段时间内，组播数据流的转发不间断。

当设置的协议收敛时间超时后，在协议收敛时间内未被更新过的所有组播转发表项将被删除。

相关配置

配置等待组播协议收敛所需的最大时间

缺省情况下，组播协议收敛所需的最大时间为 140S。

使用 `msf nsf convergence-time time` 命令可以配置等待组播协议收敛所需的最大时间，取值范围是<0-3600>。

time 值越大，等待组播协议收敛所需的最大时间越长。*time* 值越小，等待组播协议收敛所需的最大时间越短

配置组播报文泄漏的时间

缺省情况下，组播报文泄漏的时间为 150S。

使用 `msf nsf leak interval` 命令可以配置组播报文泄漏的时间，取值范围是<0-3600>。

interval 值越大，组播报文泄漏的时间越长。*interval* 值越小，组播报文泄漏的时间越短。

1.3.9 配置组播硬件表项溢出覆盖机制

创建组播转发表项时，如果硬件转发表项资源已满，将删除最早创建的硬件表项资源，添加新增加的表项。

工作原理

创建组播转发表项时，如果硬件转发表项资源已满，将删除最早创建的硬件表项资源，添加新增加的表项。

相关配置

配置组播硬件表项溢出覆盖机制

缺省情况下，组播硬件表项溢出覆盖机制关闭。

使用 `msf ipmc-overflow override` 命令配置组播硬件表项溢出覆盖机制。

1.4 产品说明



在 S7508 产品上，PIM 协议报文的 CPP 限速默认值为 1000pps，IGMP 协议报文的 CPP 限速默认值为 1000pps，超过该限速值的报文将被丢弃。当需要通过 PIM 协议建立超过 1000 个组播组转发表项时，建议调整 PIM 协议报文的 CPP 值和表项个数相等。当需要通过 IGMP 协议建立超过 1000 个组播组转发表项时，建议调整 IGMP 协议报文的 CPP 值和表项个数相等。

1.5 配置详解

配置项	配置建议 & 相关命令	
配置IP组播基本功能	必须配置。用于建立组播服务。	
	<code>ip multicast-routing</code>	启动 IPv4 组播路由功能
配置TTL阈值	可选配置。	
	<code>ip multicast ttl-threshold ttl-value</code>	配置接口的 TTL 阈值
配置能加入到IP组播路由表中的条目数量	<code>ip multicast route-limit limit [threshold]</code>	限制能够加入到组播路由表中的条目数量。
配置IP组播边界	<code>ip multicast boundary access-list [in out]</code>	设置接口成为特定组范围的组播边界。
配置IP组播静态路由	<code>ip mroute source-address mask { [ospf rip static] { v4rpf-address interface-type interface-number } } [distance]</code>	配置组播静态路由。
配置组播流二层流向控制	<code>ip multicast static source-address group-address interface-type interface-number</code>	控制二层端口的数据流流向。
配置按照最长匹配选择RPF路由	<code>ip multicast rpf longest-match</code>	配置按照最长匹配选择 RPF 路由
配置组播不间断转发参数	<code>msf nsf convergence-time time</code>	配置等待组播协议收敛所需的最大时间
	<code>msf nsf leak time</code>	配置组播报文泄漏的时间
配置组播硬件表项溢出覆盖机制	<code>msf ipmc-overflow override</code>	配置组播硬件表项溢出覆盖机制

1.5.1 配置IP组播基本功能

配置效果

- 建立 PIM 网络，向网络内的数据源和用户终端提供 IPv4 组播服务。

注意事项

- PIM 网络需要借助网络中现有的单播路由。因此，网络中必须配置 IPv4 单播路由。

配置方法

启动 IPv4 组播路由功能

- 必须配置。
- 若无特殊要求，应在每台路由器上启动 IPv4 组播路由功能。

启动接口 IP 组播协议

- 必须配置。
- 若无特殊要求，应在以下接口上启动 IP 组播协议功能

检验方法

使网络中的组播源发送组播报文，用户主机加入这些组。

- 检查用户主机能否成功接收每个组的报文。

相关命令

启动 IPv4 组播路由功能

- 【命令格式】 **ip multicast-routing**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置 IP 组播协议

 关于 IGMP 的配置方法请参见“IGMP”章节。

 关于 PIM-DM 的配置方法请参见“PIM-DM”章节。

 关于 PIM-SM 的配置方法请参见“PIM-SM”章节。

 当在 Private VLAN 和 Super VLAN 上开启三层组播后，当子 VLAN 内存在组播源时，由于组播转发时需要进行入口合法性检查，此时需要多复制一条表项，入口为组播流进入的子 VLAN。造成的影响是多占用了一条组播硬件表项，表现出来的组播容量减少 1。

显示组播转发表信息

- 【命令格式】 **show ip mroute [group-or-source-address [group-or-source-address]] [dense | sparse] [summary | count]**
- 【参数说明】 *group-or-source-address*：组地址或源地址。

group-or-source-address : 组地址或源地址。

dense : 显示 PIM-DM 组播核心表项。

sparse : 显示 PIM-SM 组播核心表项。

summary : 显示组播路由表项的摘要信息。。

count : 显示组播路由表项的计数信息。

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 3 个参数都为可选的，其中源地址和组地址必须同时指定。

当不指定源地址、组地址时，将显示所有的 mfc 表项。

当只指定源地址、组地址时，将显示源地址、组地址的 mfc 表项。

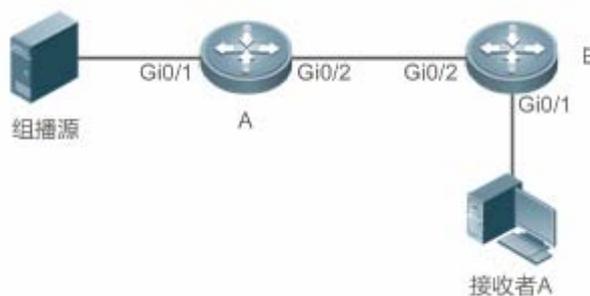
配置举例

i 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，支持 PIM-DM

【网络环境】

图 1-3



- 【配置方法】
- 在路由器上配置 IPv4 单播路由协议（如 OSPF）
 - 在所有路由器上启动 IPv4 组播路由功能。
 - 在设备互联接口上、连接用户主机和组播源的接口上启动 PIM-DM 功能。

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if)# ip pim dense-mode
A(config-if)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if)# ip pim dense-mode
A(config-if)# exit
```

B

```
B# configure terminal
B(config)# ip multicast-routing
B(config)# interface GigabitEthernet 0/1
B(config-if)# ip pim dense-mode
B(config-if)# exit
B(config)# interface GigabitEthernet 0/2
B(config-if)# ip pim dense-mode
```

```
B(config-if)# exit
```

【检验方法】 使组播源 (192.168.1.100) 向 G (233.3.3.3) 发送报文。使接收者 A 加入 G。

- 检接收者 A 收到的组播报文，接收者 A 应该能够收到 G 的组播报文。
- 检查 A、B 上的组播转发表信息。

A

```
A# show ip mroute

IP Multicast Routing Table
Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed,
       R - RPT, S - SPT, s - SSM Group
Timers: Uptime/Stat Expiry
Interface State: Interface (TTL)

(192.168.1.100, 233.3.3.3), uptime 00:01:55, stat expires 00:02:19
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/1
  Outgoing interface list:
    GigabitEthernet 0/2 (1)
```

B

```
B# show ip mroute

IP Multicast Routing Table
Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed,
       R - RPT, S - SPT, s - SSM Group
Timers: Uptime/Stat Expiry
Interface State: Interface (TTL)

(192.168.1.100, 233.3.3.3), uptime 00:00:35, stat expires 00:02:55
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/2
  Outgoing interface list:
    GigabitEthernet 0/1 (1)
```

常见错误

- IPv4 单播路由配置错误。
- 某台路由器没有使能 IPv4 组播路由。
- 某个接口没有使能 IP 组播协议

1.5.2 配置TTL阈值

配置效果

- 配置接口 TTL 阈值，检查组播报文的 TTL。如果组播报文的 TTL 大于接口阈值，则可以转发报文，否则丢弃该报文。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

- 若无特殊要求，可以在各个 PIM 路由器接口上配置 TTL 阈值

检验方法

使网络中的组播源发送组播报文，用户主机加入这些组。

- 在与用户主机相连的 PIM 路由器接口上设置 TTL 阈值，值大于组播报文的 TTL，查看用户是否能接收到组播报文。

相关命令

配置 TTL 阈值

【命令格式】 **ip multicast ttl-threshold *ttl-value***

【参数说明】 *ttl-value*：设置接口的 TTL 阈值。取值范围 0~255。缺省值 0。

【命令模式】 接口模式

【使用指导】 打开组播的设备可以为每个接口保持一个 TTL 阈值。如果组播报文的 TTL 大于接口阈值，则可以转发报文，否则丢弃该报文。TTL 阈值仅对组播播帧有效，同时必须在三层接口上配置。

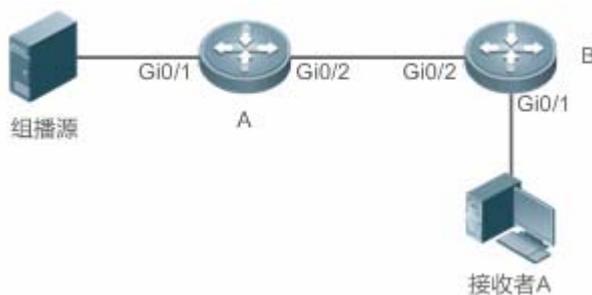
配置举例

i 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，配置 TTL 阈值

【网络环境】

图 1-4



- 【配置方法】
- 配置 IP 组播基本功能(略)
 - 在 A 设备的 Gi0/2 口配置 TTL 阈值为 100。

A

```
A# configure terminal
A(config)#int gigabitEthernet 0/2
A(config-if-GigabitEthernet 0/2)#ip multicast ttl-threshold 100
A(config-if-GigabitEthernet 0/2)# exit
```

【检验方法】 使组播源 (192.168.1.100) 向 G (233.3.3.3) 发送报文。使接收者 A 加入 G。

- 在 A 设备 Gi0/2 口配置 TTL 阈值为 100，大于组播报文的 TTL 值
- 查看配置 TTL 阈值前后，路由转发表项差别。

查看配置 TTL
阈值前

```
A# show ip mroute

IP Multicast Routing Table
Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed,
       R - RPT, S - SPT, s - SSM Group
Timers: Uptime/Stat Expiry
Interface State: Interface (TTL)

(192.168.1.100, 233.3.3.3), uptime 00:00:08, stat expires 00:03:29
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/1
  Outgoing interface list:
    GigabitEthernet 0/2 (1)
```

查看配置 TTL
阈值后

```
A# show ip mroute

IP Multicast Routing Table
Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed,
       R - RPT, S - SPT, s - SSM Group
Timers: Uptime/Stat Expiry
Interface State: Interface (TTL)

(192.168.1.100, 233.3.3.3), uptime 00:00:01, stat expires 00:03:29
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/1
  Outgoing interface list:
    GigabitEthernet 0/2 (100)
```

常见错误

1.5.3 配置能加入到IP组播路由表中的条目数量

配置效果

- 设备上每个收到的组播数据报文都维护相应的 IP 组播路由转发表项。但是，过量的组播路由表项可能会耗尽设备内存，从而导致设备性能下降。用户可以根据实际组网情况和业务性能要求对 IP 组播路由表中的表项数量进行限制。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

- 可以根据实际组网情况和业务性能要求对 IP 组播路由表中的表项数量进行限制

检验方法

网络中的组播源发送 N 组不同的组播报文，用户主机加入这些组，在设备上配置能加入到 IP 组播路由表中的条目数量为 N-1，查看用户主机接收到的组播报文为 N-1 组组播报文。

相关命令

配置能加入到 IP 组播路由表中的条目数量

【命令格式】 **ip multicast route-limit** *limit* [*threshold*]

【参数说明】 *limit*：组播路由条目数。取值范围 1~65536。缺省值为 1024。
threshold：触发产生警示消息的组播路由数量。缺省值为 65536。

【命令模式】 全局模式

【使用指导】 因为受到硬件资源的限制，超过硬件允许范围的路由表项只能由软件转发，使性能下降。

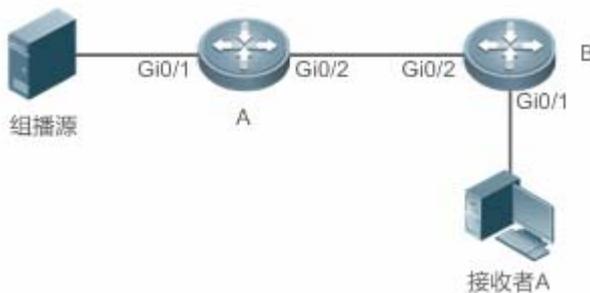
配置举例

i 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，配置能加入到 IP 组播路由表中的条目数量

【网络环境】

图 1-5



【配置方法】 ● 配置 IP 组播基本功能(略)

● B 设备配置能加入到 IP 组播路由表中的条目数量为 2

B

```
B# configure terminal
B(config)# ip multicast route-limit 2
```

【检验方法】 使组播源 (192.168.1.100) 向 G1 (233.3.3.1) , G2(233.3.3.2) , G3(233.3.3.3)发送报文。使接收者 A 加入 G1, G2 , G3。

● 检接收者 A 收到的组播报文，接收者 A 应该能够收到 G1 , G2 , G3 中的两组的组播报文。

● 检查 A、B 上组播路由表项。

● 当组播路由表中的条目数量达到上限的时候会打印提示信息。

A

```
A# show ip mroute
```

```
IP Multicast Routing Table
Flags: I – Immediate Stat, T – Timed Stat, F – Forwarder installed,
       R – RPT, S – SPT, s – SSM Group
Timers: Uptime/Stat Expiry
Interface State: Interface (TTL)

(192.168.1.100, 233.3.3.1), uptime 00:00:06, stat expires 00:03:24
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/1
  Outgoing interface list:
    GigabitEthernet 0/2 (1)

(192.168.1.100, 233.3.3.2), uptime 00:00:05, stat expires 00:03:25
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/1
  Outgoing interface list:
    GigabitEthernet 0/2 (1)

(192.168.1.100, 233.3.3.3), uptime 00:00:00, stat expires 00:03:30
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/1
  Outgoing interface list:
    GigabitEthernet 0/2 (1)
```

B

```
B# show ip mroute

IP Multicast Routing Table
Flags: I – Immediate Stat, T – Timed Stat, F – Forwarder installed,
       R – RPT, S – SPT, s – SSM Group
Timers: Uptime/Stat Expiry
Interface State: Interface (TTL)

(192.168.1.100, 233.3.3.1), uptime 00:01:13, stat expires 00:03:23
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/2
  Outgoing interface list:
    GigabitEthernet 0/1 (1)

(192.168.1.100, 233.3.3.3), uptime 00:06:08, stat expires 00:03:23
Owner PIMDM, Flags: TFS
  Incoming interface: GigabitEthernet 0/2
  Outgoing interface list:
    GigabitEthernet 0/1 (1)
```

组播路由表中的条目数量达到上限的时候会打印提示信息。

```
B#*Dec 26 10:43:07: %MROUTE-4-ROU TELIMIT: IPv4 Multicast route limit 2 exceeded - VRF default.
```

常见错误

- IPv4 单播路由配置错误。

1.5.4 配置IP组播边界

配置效果

- 配置 IP 组播边界，确定组播报文的传递范围。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

- 若无特殊要求，可以在各个 PIM 路由器接口上配置 IP 组播边界

检验方法

使网络中的组播源发送组播报文，用户主机加入这些组。在与用户主机相连 PIM 路由器接口上配置 IP 组播边界，查看用户是否能接收到组播报文。

相关命令

启动 IPv4 组播路由功能

【命令格式】 **ip multicast boundary access-list [in | out]**

【参数说明】 *access-list*：使用 ACL 定义的组地址范围。

in：表示组播边界作用在组播流的流入方向。

out：表示组播边界作用在组播流的流出方向。

【命令模式】 接口模式

【使用指导】 配置此命令后，在此接口上对与该组范围内的 IGMP、PIM-SM 报文进行过滤，组播数据流不会从该接口流入和流出。

此命令关联的 ACL 可以是标准的 ACL，也可以是扩展 ACL。对于扩展的 ACL 只匹配目的地址，不匹配源地址。

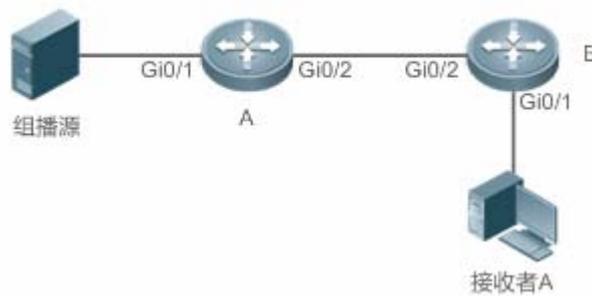
配置举例

 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，配置 IP 组播边界

【网络环境】

图 1-6



【配置方法】

- 配置 IP 组播基本功能(略)
- 在 A 设备上配置 ACL
- 在 A 设备接口 Gi0/1 上配置 IP 组播边界。

A

```
A# configure terminal
A(config)#ip access-list standard ip_multicast
A(config-std-nacl)#deny any
A(config-std-nacl)#exit
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)#ip multicast boundary ip_multicast A
A(config-if-GigabitEthernet 0/1)# exit
```

【检验方法】

使组播源 (192.168.1.100) 向 G (233.3.3.3) 发送报文。使接收者 A 加入 G。

- 打开 **debug ip pim sparse-mode events**

A

```
A# debug ip pim sparse-mode events
Jan 1 20:58:34: %7: VRF(0): No cache message: src 192.168.1.100 for 233.3.3.3 vif 2
*Jan 1 20:58:34: %7: VRF(0): Ignore No cache message: src 192.168.1.100 for 233.3.3.3 vif 2
in PIM_BOUNDARY_FLT_BOTH range
```

常见错误

- IPv4 单播路由配置错误。

1.5.5 配置IP组播静态路由

配置效果

- 通过配置 IP 组播静态路由，可以为来自特定组播源的组播报文指定 RPF 接口或 RPF 邻居。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

- 若无特殊要求，可在每台设备上配置 IP 组播静态路由。

检验方法

配置 IP 组播静态路由，可以通过 **show ip rpf { source-address [group-address] [rd route-distinguisher] } [metric]** 命令检查特定源的 RPF 信息。

相关命令

配置 IP 组播静态路由

【命令格式】 **ip mroute source-address mask { [ospf | rip | static] { v4rpf-address | interface-type interface-number } } [distance]**

【参数说明】 **source-address**：组播源地址。
mask：组播源地址的掩码。
protocol：当前正在使用的单播路由协议。
rpf-address：RPF 邻居（到组播源的下一跳）的地址。
interface-type interface-number：RPF 接口（到组播源的出接口）。
distance：路由管理距离。取值范围 0-255，缺省值 0。

【命令模式】 全局模式

【使用指导】 静态组播路由仅用于 RPF 检查。

如果要指定静态组播路由的出接口而非下一跳 IP 地址，那么该出接口必须为点对点类型。

显示特定源地址的 RPF 信息

【命令格式】 **show ip rpf { source-address [group-address] [rd route-distinguisher] } [metric]**

【参数说明】 **source-address**：指定的源地址。
group-address：指定的组地址。
rd route-distinguisher：使用 RD 的 proxy 查找。
metric：显示 MDT-SAFI 路由的 metric。

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 3 个参数都为可选的，其中源地址和组地址必须同时指定。

当不指定源地址、组地址时，将显示所有的 mfc 表项。

当只指定源地址、组地址时，将显示源地址、组地址的 mfc 表项。

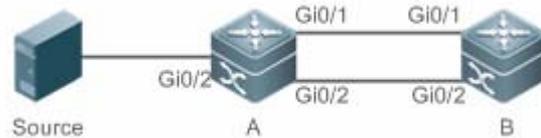
配置举例

i 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，支持 PIM-DM

【网络环境】

图 1-7



【配置方法】

- 配置 IP 组播基本功能(略)
- 在 B 设备上配置到接收者的静态路由。

A

```
B# configure terminal
B(config)# ip mroute 10.10.10.10 255.255.255.255 ospf 192.168.1.1 1
```

【检验方法】

使用 **show ip rpf** 命令查看配置前与配置后的到接收者的 RPF 信息。

配置前

```
B# show ip rpf 10.10.10.10
RPF information for 10.10.10.10
  RPF interface: GigabitEthernet 0/2
  RPF neighbor: 192.168.2.1
  RPF route: 10.10.10.10/32
  RPF type: unicast (ospf)
  RPF recursion count: 0
  Doing distance-preferred lookups across tables
  Distance: 110
  Metric: 1
```

配置后

```
B# show ip rpf 10.10.10.10
RPF information for 10.10.10.10
  RPF interface: GigabitEthernet 0/0
  RPF neighbor: 192.168.1.1
  RPF route: 10.10.10.10/32
  RPF type: static
  RPF recursion count: 0
  Doing distance-preferred lookups across tables
  Distance: 1
  Metric: 0
```

常见错误

- IPv4 单播路由配置错误。
- 某台路由器没有使能 IPv4 组播路由。

1.5.6 配置组播流二层流向控制

配置效果

通过配置组播流二层流向控制，控制组播流在端口上的转发行为。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

- 若无特殊要求，可以在含二层设备上配置组播流二层流向控制。

检验方法

在含二层的设备 A 的网络中发送组播报文，二层设备上 A 的 vlan 1 上连接有多个用户主机接收这些组，在二层设备 A 上配置组播流二层流向控制，组播报文只向配置的二层端口发送。

相关命令

配置组播流二层流向控制

【命令格式】 **ip multicast static** *source-address group-address interface-type interface-number*

【参数说明】 *source-address*：组播源地址。

group-address：组播组地址。

interface-type interface-number：允许转发该组播流的二层端口。

【命令模式】 全局模式

【使用指导】 对于某一条组播流可以配置多条命令，也就是说配置多个被允许转发的端口。一旦为某组播流配置了流向控制，该组播流只可能由这些已配置的端口转发出去。其他未被允许的端口将不允许转发组播流。该命令只控制组播流在端口上的转发行为，不直接影响组播协议对协议报文的处理。但是由于组播协议（如：PIM-DM 或 PIM-SM）的某些特性是依赖组播数据流驱动的，所以组播路由协议的行为仍然可能受影响。

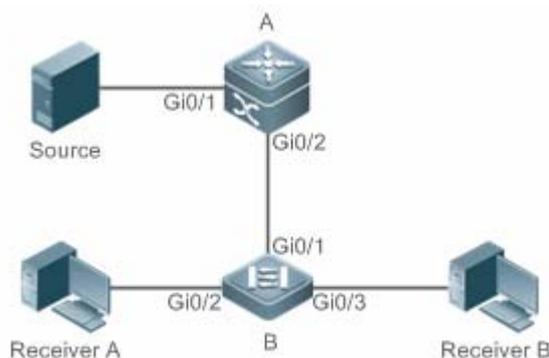
配置举例

i 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，配置组播流二层流向控制

【网络环境】

图 1-8



- 【配置方法】
- 配置 IP 组播基本功能(略)
 - 在设备 B 配置组播流二层流向控制，流只向 Gi0/2 口发送。

B

```
A# configure terminal
A(config)# ip multicast static 192.168.1.100 233.3.3.3 gigabitEthernet0/2
```

- 【检验方法】 使组播源 (192.168.1.100) 向 G (233.3.3.1) 发送报文。使接收者 A, B 加入 G。
- 检接收者 A 收到的组播报文，接收者 B 不能够收到 G 组播报文。

常见错误

- IPv4 单播路由配置错误。

1.5.7 配置按照最长匹配选择RPF路由

配置效果

在组播静态路由表、MBGP 路由表、单播路由表中选取的最优路由中选取掩码最长匹配的路由，作为 RPF 路由。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

- 若无特殊要求，可在每台设备上配置按照最长匹配选择 RPF 路由

检验方法

配置优先级相同的组播静态路由、单播静态路由，配置单播静态路由具有的掩码长度最长。

- 通过 **show ip rpf { source-address [group-address] [rd route-distinguisher] } [metric]**命令检查特定源的 RPF 信息。

相关命令

配置按照最长匹配选择 RPF 路由

【命令格式】 **ip multicast rpf longest-match**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 选择 RPF 路由的步骤：

- 1，从组播静态路由表、MBGP 路由表、单播路由表中各选出一条可用于 RPF 检查的最优路由。
- 2，从这三条路由中选择一条作为 RPF 路由。

如果配置了按照最长匹配原则选择路由命令，则从这三条路由中选出最长匹配的那条路由。如果这三

条路由的掩码一样，则选择其中优先级最高的那条路由。如果它们的优先级也相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。

如果没有配置按照最长匹配选择路由，则从这三条路由中选出优先级最高的那条路由。如果它们的优先级相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。

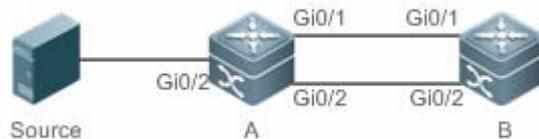
配置举例

i 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，配置按照最长匹配选择 RPF 路由

【网络环境】

图 1-9



【配置方法】

- 配置 IP 组播基本功能(略)
- 在设备 B 上配置 IP 组播静态路由，路由的掩码长度小于单播路由。
- 在设备 B 上配置按照最长匹配选择 RPF 路由。

B

```

B# configure terminal
B(config)# ip multicast-routing
B(config)# ip mroute 10.10.10.10 255.255.0.0 ospf 192.168.1.1
B(config)# ip multicast rpf longest-match
  
```

【检验方法】

使用 **show ip rpf** 查看配置按照最长匹配选择 RPF 路由前后到组播源的 RPF 信息。

配置前

```

B#show ip rpf 10.10.10.10
RPF information for 10.10.10.10
  RPF interface: GigabitEthernet 0/0
  RPF neighbor: 192.168.1.1
  RPF route: 10.10.0.0/16
  RPF type: static
  RPF recursion count: 0
  Doing distance-preferred lookups across tables
  Distance: 0
  Metric: 0
  
```

配置后

```

B# show ip rpf 10.10.10.10
RPF information for 10.10.10.10
  RPF interface: GigabitEthernet 0/2
  RPF neighbor: 192.168.2.1
  RPF route: 10.10.10.10/32
  RPF type: unicast (ospf)
  RPF recursion count: 0
  Doing prefix-length-preferred lookups across tables
  Distance: 110
  
```

Metric: 1

常见错误

- IPv4 单播路由配置错误。
- 某台路由器没有使能 IPv4 组播路由。

1.5.8 配置组播不间断转发参数

配置效果

- 组播不间断转发（Non-Stop Forwarding）功能保证在组播协议重新收敛的这段时间内，组播数据流的转发不间断。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

配置等待组播协议收敛所需的最大时间

- 若无特殊要求，可在每台设备上配置等待组播协议收敛所需的最大时间。

配置组播报文泄漏的时间

- 若无特殊要求，可在每台设备上配置组播报文泄漏的时间。

检验方法

使用 `show msf nsf` 命令，可以查看配置的组播不间断转发的参数。

相关命令

配置等待组播协议收敛所需的最大时间

【命令格式】 `msf nsf convergence-time time`

【参数说明】 `convergence-time time`：等待组播协议收敛的最大时间，取值范围 0~3600 秒，缺省值 140 秒。

【命令模式】 全局模式

【使用指导】 -

配置组播报文泄漏的时间

【命令格式】 `msf nsf leak interval`

【参数说明】 `leak intervald`：报文组播泄漏的时间，取值范围 0~3600 秒，缺省值 150 秒。

【命令模式】 全局模式

【使用指导】 -

显示组播不间断转发配置

【命令格式】 **show msf nsf**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 -

配置举例

i 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，支持 PIM-DM

【网络环境】 IP 组播服务基本环境(图略)

【配置方法】

- 配置 IP 组播基本功能(略)
- 配置等待组播协议收敛所需的最大时间。
- 配置组播报文泄漏的时间

```
A# configure terminal
A(config)# msf nsf convergence-time 200
A(config)# msf nsf leak 300
```

【检验方法】 通过使用 **show msf nsf** 命令，可以显示组播不间断转发配置。

```
A# show msf nsf
Multicast HA Parameters
-----+-----+
protocol convergence timeout          200 secs
flow leak interval                     300 secs
```

常见错误

1.5.9 配置组播硬件表项溢出覆盖机制

配置效果

- 创建组播转发表项时，如果硬件转发表项资源已满，将删除最早创建的硬件表项资源，添加新增加的表项。

注意事项

- 必须配置 IP 组播基本功能。

配置方法

- 若无特殊要求，可在每台设备上配置组播硬件表项溢出覆盖机制

检验方法

通过使用 **show running-config** 命令，可以查看是否配置组播硬件表项溢出覆盖机制。

相关命令

配置组播硬件表项溢出覆盖机制

- 【命令格式】 **msf ipmc-overflow override**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

 以下配置举例，仅介绍与 IP 组播相关的配置。

在 IPv4 网络上建立 IP 组播服务，配置组播硬件表项溢出覆盖机制

【网络环境】 IP 组播服务基本环境(图略)

- 【配置方法】
- 配置 IP 组播基本功能(略)
 - 配置组播硬件表项溢出覆盖机制。

```
A# configure terminal
A(config)#msf ipmc-overflow override
```

【检验方法】 通过使用 **show running-config** 命令，可以查看是否配置组播硬件表项溢出覆盖机制。

```
A# show running-config
...
msf ipmc-overflow override
...
```

常见错误

1.6 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除 v4 组播转发表。	<code>clear ip mroute { * v4group-address [v4source-address] }</code>
复位 v4 组播转发表统计信息。	<code>clear ip mroute statistics { * v4group-address [v4source-address] }</code>

查看运行情况

作用	命令
查看各组播模块的帮助引导信息。	<code>multicast help</code>
查看 v4 组播模块状态信息。	<code>view multicast</code>
查看 v4 组播转发表。	<code>show ip mroute [group-or-source-address [group-or-source-address]] [dense sparse] [summary count]</code>
查看 v4 静态组播路由信息。	<code>show ip mroute static</code>
查看特定 v4 源地址的 RPF 信息。	<code>show ip rpf { source-address [group-address] [rd route-distinguisher] } [metric]</code>
查看 v4 组播接口的信息。	<code>show ip mvif [interface-type interface-number]</code>
查看 IPv4 三层组播转发表。	<code>show ip mrf mfc</code>
查看 IPv4 多层组播转发表。	<code>show msf msc</code>
查看 IPv4 组播不间断转配置。	<code>show msf nsf</code>

查看调试信息



输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
查看组播核心的运行过程。	<code>debug nsm mcast all</code>
查看 IPv4 组播核心对协议模块进行通信用过程。	<code>debug nsm mcast fib-msg</code>
查看 IPv4 组播核心关于接口运行过程。	<code>debug nsm mcast vif</code>
查看 IPv4 组播核心关于接口和表项统计信息的处理过程。	<code>debug nsm mcast stats</code>
查看 IPv4 三层组播报文转发的处理过程。	<code>debug ip mrf forwarding</code>
查看 IPv4 下对三层组播转发表项的操作过程。	<code>debug ip mrf mfc</code>
查看 IPv4 下关于三层组播转发事件的处理过程。	<code>debug ip mrf event</code>
查看 IPv4 多层组播报文转发的处理过程。	<code>debug msf forwarding</code>
查看 IPv4 下对多层组播转发表项的操作过程。	<code>debug msf mfc</code>
查看 IPv4 多层组播转发操作底层硬件的处理过程。	<code>debug msf ssp</code>
查看 IPv4 多层组播转发提供的 API 接口被调用的处理过程	<code>debug msf api</code>
查看 IPv4 下关于多层组播转发事件的处理过程	<code>debug msf event</code>

2 IGMP

2.1 概述

IGMP 是 Internet Group Management Protocol (互联网组管理协议) 的简称。

它是 TCP/IP 协议族中负责 IP 组播成员管理的协议，用来在 IP 主机和与其直接相邻的组播路由器之间建立、维护组播组成员关系。IGMP 协议的行为可以分成两个部分：主机行为和设备行为。

目前 IGMP 有三个版本：IGMPv1，IGMPv2，IGMPv3。

- 所有版本的 IGMP 都支持 ASM (Any-Source Multicast，任意信源组播) 模型。
- IGMPv3 可以直接应用于 SSM (Source-Specific Multicast，指定信源组播) 模型。
- IGMPv1 和 IGMPv2 则需要在 IGMP SSM Mapping 技术的支持下才能应用于 SSM 模型。

协议规范

- RFC 1112 : Host Extensions for IP Multicasting
- RFC 2236 : Internet Group Management Protocol, Version 2
- RFC 3376 : Internet Group Management Protocol, Version 3
- RFC 4605 : Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")

2.2 典型应用

典型应用	场景描述
本地网络内 IGMP 服务	本地网络内实现 IGMP 服务。
IGMP 代理服务	简单树形网络拓扑中，不用 PIM 服务，采用 IGMP 代理服务。

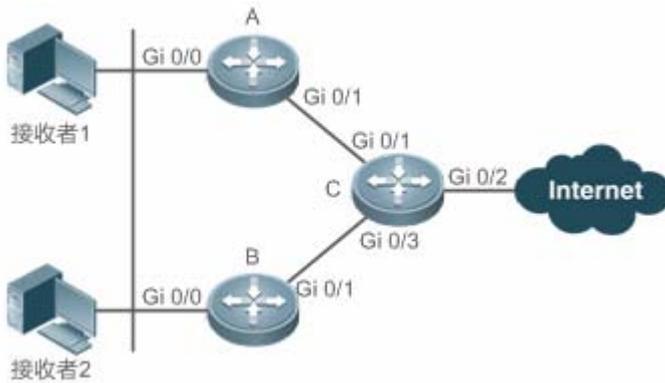
2.2.1 本地网络内IGMP服务

应用场景

以下图为例，接收者 1、接收者 2、路由器 A、路由器 B 组成了本地网络。

路由器 A 或路由器 B 发送的 Query 报文，是在本地网络有效，接收者 1 和接收者 2 发送的 Report 报文也是在本地网络有效。

图 2-1



【注释】 C 为出口网关设备。
A、B 为核心路由器。

功能部属

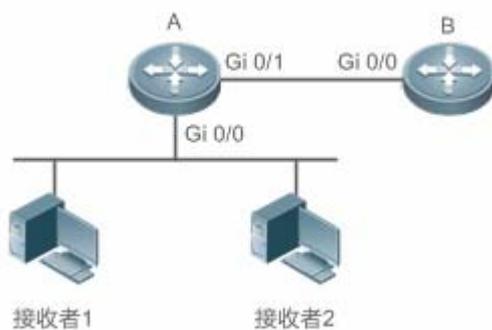
- A、B、C 路由器运行 OSPF 协议。
- A、B、C 的各接口运行组播协议（PIM-SM 或 PIM-DM）。

2.2.2 IGMP代理服务

应用场景

以下图为例，路由器 A 上实现代理功能，路由器 A 当作主机，它与路由器 B 组成本地网络组的管理。路由器 A 转发接收者 1 与接收者 2 的 Report 报文。

图 2-2



【注释】 A 路由器为代理功能。
B 路由器提供 PIM 服务。

功能部属

- A 与 B 路由器运行 OSPF 协议。

- A 与 B 的各接口运行组播协议 (PIM-SM 或 PIM-DM)。
- 在 A 路由器上 Gi0/0 和 Gi0/1 上实现组播代理。

2.3 功能详解

基本概念

主机行为、设备行为

运行组播管理协议的三层组播设备简称为设备，它的行为称为设备行为。

运行组播管理协议的 PC 机或模拟 PC 机简称为主机，它的行为称为主机行为。

查询者

设备之间交互竞争，通过比较 IP 地址，低 IP 地址成为查询者，查询者定期发送查询报文。

IGMP PROXY-SERVICE 接口

该接口执行主机行为，也称为上行口，接收上游设备发送过来的 Query 报文，发送路由器代理收集回来的 Report 信息。

IGMP MROUTE-PROXY 接口

该接口执行路由器功能，也称为下行口，发送代理服务接口接收的报文，收集主机信息并发送给代理服务接口。

IGMP 的 SSM-MAP

组播特定源的映射，IGMPv1 与 IGMPv2 不支持 SSM 模型，运行了 SSM-MAP 功能，支持 SSM 模型。

功能特性

功能特性	作用
IGMP 路由器参数配置	发送查询报文，获取本地成员信息。
支持IGMP组的过滤	对组成员进行过滤，组员个数进行限制。
支持IGMP配置静态组	本路由器存在静态组信息，而不需要主机发送 Report 获取。
模拟主机设置组信息	模拟主机行为直接设置加入组信息。
支持IGMP代理功能	在简单树形网络拓扑中，并不需要运行复杂的组播路由协议（如 PIM），可以采用些功能。
支持SSM-MAP功能	为 IGMPv1 和 IGMPv2 提供 SSM 模型的支持。主机加入组的时候，可以指定特定的源，节省带宽，避免不需要的、非法的组播数据流占用网络带宽，这尤其在多个组播源共用一个组播地址的网络环境中表现明显
支持 IGMP 报文 Router Alert 选项	支持 IGMP 报文的 Router_Alert 选项进行检查，丢弃未携带 Router_Alert 选项的 IGMP 报文。 支持发送 IGMP 报文时携带 Router_Alert 选项。

2.3.1 IGMP路由器参数配置

IGMP 版本以及查询器参数设置

工作原理

。

在运行 IGMP 协议的组播网络中，组播设备定期发送 IGMP 查询报文，并根据响应情况来确认本地组成员信息。

同一网段只有一台组播设备负责发送 IGMP 查询报文，此设备称为查询器。查询器是通过选举决定的，在初始状态下所有组播设备都处于查询器状态，当设备收到低 IP 地址设备的成员关系查询时，他们从接收者状态转变成非查询者状态。因此最终仅有一个设备处于查询状态。这个设备为网络上所有组播设备中 IP 地址最低者。

通过 IGMP 版本的设置，查询器发送不同版本的 IGMP 查询报文，通过对查询器参数的设置，调整查询器发送 IGMP 查询报文的频率、最后成员的查询次数和查询间隔、IGMP 查询报文的最大响应时间，已经查询器的存活时间。

相关配置

▾ 启动 IGMP 协议

缺省情况下，接口关闭 IGMP 协议。

使用 `ip pim { sparse-mode | dense-mode }` 命令，启动或关闭接口 IGMP 协议功能。

必须在接口配置 SM 或 DM 协议，才能启动 IGMP 协议。

▾ 配置 IGMP 版本

缺省情况下，启动 IGMP 协议是 2。

使用 `ip igmp version { 1 | 2 | 3 }` 命令，配置或恢复接口的 IGMP 版本。

▾ 配置最后成员查询间隔

缺省情况下，启动发送查询报文间隔是 1S。

使用 `ip igmp last-member-query-interval interval` 命令，配置或恢复接口查询报文间隔。

设置值越大发送间隔越大，设置值越小发送间隔越短。

▾ 配置最后成员查询次数

缺省情况下，最后成员查询次数是 2 次。

使用 `ip igmp last-member-query-count count` 命令，配置或恢复最后成员发送次数。

设置值越大最后成员查询次数就越多，设置值越小最后成员查询次数就越小。

▾ 配置普通成员查询间隔

缺省情况下，最后成员查询间隔是 125S。

使用 `ip igmp query-interval seconds` 命令，设置或恢复普通成员查询间隔。

设置值越大普通查询间隔就越大，设置值越小普通查询间隔就越小。

配置最大响应时间

缺省情况下，最大的响应时间是 10S。

使用 `ip igmp query-max-response-time seconds` 命令，设置或恢复最大响应时间。

设置值越大最大响应时间就越大，设置值越小最大响应时间就越小。

配置查询者的存活时间

缺省情况下，查询者的存活时间是 255S。

使用 `ip igmp query-timeout seconds` 命令，设置或恢复查询者的存活时间。

设置值越大存活时间越长，设置值越小存活时间越小。

2.3.2 支持IGMP组的过滤

对组成员进行过滤，组员个数进行限制。

工作原理

如果不希望接口所在网段上的主机加入某些组播组，可在该接口上配置 ACL 规则作为过滤器，接口将按照该规则对收到的 IGMP 成员关系报告报文进行过滤，只为该规则所允许的组播组维护组成员关系。设置路由器成员上限。

相关配置

配置组播组访问控制

缺省情况下，取消控制，允许加入任何组。

使用 `ip igmp access-group access-list-name` 命令，设置或恢复组播组访问控制。

设置后路由器只能接收设置的访问链表。

配置 IGMP 组成员数量限制

缺省情况下，组成员数量的限制是 1024。

使用 `ip igmp limit number` 命令，设置或恢复组播成员数量。

设置值越大，成员数量越大，设置值越小，成员数量越小。

2.3.3 支持IGMP配置静态组

本路由器存在静态组信息，而不需要主机发送 Report 获取。可以直接向 PIM 路由器交互组信息。

工作原理

通过人工设置静态组信息。

相关配置

配置静态组 static-group

缺省情况下，没有设置静态组信息。

使用 `ip igmp static-group group-address` 命令，设置或取消静态组信息。

2.3.4 模拟主机设置组信息

模拟主机行为直接设置加入组信息。

相关配置

配置主机组 Join-Group

缺省情况下，没有设置加入组信息。

使用 `ip igmp join-group group-address` 命令，设置或取消静态组信息。

2.3.5 支持IGMP代理功能

在简单树形网络拓扑中，并不需要运行复杂的组播路由协议（如 PIM），可以采用些功能，使其代理下游主机来发送 IGMP 报文及维护成员关系。

工作原理

设置成 IGMP 代理服务接口，对于上游路由器来说，相当于主机功能，可以接收上游路由器发送过来的 Query 报文，也可以转发下游的主机的组信息。设置成 IGMP 组播代理接口，对于下游路由器来说，相当于路由器功能，可以转发上游路由器的 Query 报文，也可以接收下游路由器发送过来的 Report 报文。

相关配置

配置 IGMP PROXY-SERVICE

缺省情况下，取消接口服务代理功能。

使用 `ip igmp proxy-service` 命令，设置或取消接口服务代理功能。

在使用代理时候，必须配置。

配置 IGMP MROUTE-PROXY

缺省情况下，取消接口组播代理功能。

使用 `ip igmp mroute-proxy interfacename` 命令，设置或取消接口组播代理功能。

在使用代理时候，必须配置。

2.3.6 支持SSM-MAP功能

为 IGMPv1 和 IGMPv2 提供 SSM 模型的支持。主机加入组的时候，可以指定特定的源，节省带宽，避免不需要的、非法的组播数据流占用网络带宽，这尤其在多个组播源共用一个组播地址的网络环境中表现明显。

工作原理

在 IGMP v1/v2 的基础上，IGMPv3 提供了额外的源过滤组播功能。在 IGMPv1/v2 中，主机只根据组地址来决定加入某个组并从任何一个源接收发给该组地址的组播流。而使用 IGMPv3 的主机通告该主机所希望加入的组播组，同时还通告该主机所希望接收的组播源的地址。IGMPv1 和 IGMPv2 也可以实现某种意义上的“源地址过滤”，但它是在组播流接收端上做的，使能 SSM-MAP 功能，设置 SSM-MAP 静态组功能。

相关配置

启动 IGMP SSM-MAP

缺省情况下，关闭 SSM-MAP 功能。

使用 `ip igmp ssm-map enable` 命令，设置或关闭 SSM-MAP 功能。

使用 SSM-MAP 功能必须设置。

配置 IGMP SSM-MAP STATIC

缺省情况下，没有设置 SSM-MAP 静态链表。

使用 `ip igmp ssm-map static access-list-num A.B.C.D` 命令，设置或关闭 SSM-MAP 静态链表。

2.3.7 支持IGMP报文Router Alert选项

支持 IGMP 报文的 Router_Alert 选项进行检查，丢弃未携带 Router_Alert 选项的 IGMP 报文。

支持发送 IGMP 报文时携带 Router_Alert 选项。

工作原理

Router Alert 选项，如果数据报文中含有这种选项，路由器就需要对数据报文的进行深入检查，并对相应的控制数据进行更新，如果数据报文没有这一项，路由器就不用费心检查，只要路由就好了。

配置对 IGMP 报文的 Router Alert 选项检查，如果接收到的 IGMP 报文未携带 Router Alert 选项，丢弃接收到的 IGMP 报文。

配置携带 Router Alert 选项，在发送 IGMP 报文时将封装 Router Alert 选项。

相关配置

配置对 IGMP 报文 Router Alert 选项检查

缺省情况下，不检查 IGMP 报文的 Router Alert 选项。

使用 `ip igmp enforce-router-alert` 命令，配置对 IGMP 报文的 Router Alert 选项检查。

配置发送 IGMP 报文携带 Router Alert 选项

缺省情况下，发送的 IGMP 报文不携带 Router Alert 选项。

使用 `ip igmp send-router-alert` 命令，配置发送 IGMP 报文时携带 Router Alert 选项。

2.4 配置详解

配置项	配置建议 & 相关命令	
配置IGMP基本功能	 必须配置。用于建立组播服务。	
	<code>ip multicast-routing</code>	启动 IPv4 组播路由功能
	<code>ip pim { sparse-mode dense-mode }</code>	启动 PIM-SM 或 PIM-DM 功能
IGMP路由器参数配置	<code>ip igmp version { 1 2 3 }</code>	配置 IGMP 的版本
	<code>ip igmp last-member-query-interval interval</code>	配置最后成员查询间隔
	<code>ip igmp last-member-query-count count</code>	配置最后成员查询次数
	<code>ip igmp query-interval seconds</code>	配置普通成员查询间隔
	<code>ip igmp query-max-response-time seconds</code>	配置最大响应间隔
支持IGMP组的过滤	<code>ip igmp query-timeout seconds</code>	配置其他查询者存活时间
	<code>ip igmp access-group access-list</code>	配置组播组访问控制
支持IGMP代理功能	<code>ip igmp limit number [except access-list]</code>	IGMP 组成员数量限制
	<code>ip igmp proxy-service</code>	配置 IGMP PROXY-SERVICE
支持SSM-MAP功能	<code>ip igmp mroute-proxy interface-type interface-number</code>	配置 IGMP MROUTE-PROXY
	<code>ip igmp ssm-map enable</code>	启动 SSM-MAP 功能
支持 IGMP 报文 Router Alert 选项	<code>ip igmp ssm-map static access-list source-address</code>	配置 IGMP SSM-MAP STATIC
	<code>ip igmp enforce-router-alert</code>	配置 IGMP Router Alert 选项检查
	<code>ip igmp send-router-alert</code>	配置发送的 IGMP 报文携带 Router Alert 选项

2.4.1 配置IGMP基本功能

配置效果

- 启动本地网络的组播路由功能，收集本地网络的组信息。

注意事项

- 接口必须启动 PIM-SM 或 PIM-DM 功能。

配置方法

启动 IPv4 组播路由功能

- 必须配置。
- 若无特殊要求，应在本地网络，每台路由器启动 IPv4 组播路由功能。

启动 PIM-SM 或 PIM-DM 功能

- 必须配置。
- 若无特殊要求，应在直接本地网络的接口上启动 PIM-SM 或 PIM-DM 功能。

检验方法

通过 `show ip igmp interface interface-type interface-number` 命令，查看接口是否启动了 igmp 功能。

相关命令

启动 IPv4 组播路由功能

【命令格式】 `ip multicast-routing`

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

启动 PIM-SM 或 PIM-DM 功能

【命令格式】 `ip pim { sparse-mode| dense-mode }`

【参数说明】 -

【命令模式】 接口模式

【使用指导】 PIM 接口必须是三层接口，包括：路由口、L3AP、SVI、loopback 接口。
所有 PIM 接口应 IPv4 单播路由可达。

配置举例

启动本地网络的 IGMP 功能

- 【配置方法】
- 在路由器上配置 IPv4 单播路由协议（如 OSPF），且 loopback 接口单播路由可达。略
 - 在所有路由器上启动 IPv4 组播路由功能。
 - 在设备互联接口上、连接用户主机和组播源的接口上启动 PIM-SM 或 PIM-DM 功能。

```
VSU(config)#ip multicast-routing
VSU(config)#int gi 0/5
```

```
VSU(config-if-GigabitEthernet 0/5)#ip add 192.168.1.90 255.255.255.0
VSU(config-if-GigabitEthernet 0/5)#ip pim sparse-mode
```

【检验方法】 通过 **show ip igmp interface interface-type interface-number** 命令，可以直接查看接口的 igmp 功能。

```
VSU#show ip igmp interface gigabitEthernet 0/5
Interface GigabitEthernet 0/5 (Index 5)
IGMP Active, Querier, Version 2 (default)
Internet address is 192.168.1.90
IGMP interface limit is 1024
IGMP interface has 1 group-record states
IGMP interface has 0 static-group records
IGMP activity: 3 joins, 0 leaves
IGMP query interval is 125 seconds
IGMP querier timeout is 255 seconds
IGMP max query response time is 10 seconds
Last member query response interval is 10
Last member query count is 2
Group Membership interval is 260 seconds
Robustness Variable is 2
```

常见错误

- 网络中路由器没有启动组播路由。
- 网络中接口没有组播接口。

2.4.2 IGMP路由器参数配置

配置效果

- 修改 IGMP 路由器参数功能，影响发送报文类型、方式以及查询器存活时间。

注意事项

- 必须配置 IGMP 基本功能

配置方法

配置 IGMP 版本

可选配置

若无特殊要求，可在所有的直连本地网络的路由器接口上配置。

配置最后成员查询间隔

可选配置

若无特殊要求，可在所有的直连本地网络的路由器接口上配置。

配置最后成员查询次数

可选配置

若无特殊要求，可在所有的直连本地网络的路由器接口上配置。

配置普通成员查询间隔

可选配置

若无特殊要求，可在所有的直连本地网络的路由器接口上配置。

配置最大响应间隔

可选配置

若无特殊要求，可在所有的直连本地网络的路由器接口上配置。

检验方法

通过 `show ip igmp interface interface-type interface-number` 命令，查看配置的信息。

相关命令

配置 IGMP 的版本

【命令格式】 `ip igmp version { 1 | 2 | 3 }`

【参数说明】 1：指的是版本 1
2：指的是版本 2
3：指的是版本 3

【命令模式】 接口配置模式

【使用指导】 配置此命令后，IGMP 会自动重启。

配置最后成员查询间隔

【命令格式】 `ip igmp last-member-query-interval interval`

【参数说明】 *Interval*：特定组查询报文的发送间隔。单位 0.1S，取值范围 1-255，缺省值 10（即 1S）。

【命令模式】 接口配置模式

【使用指导】 该命令只适用于 IGMPv2 或 IGMPv3。

当接口收到 Leave 报文后，会连续发送特定组查询报文，并等待主机回应。超时后，认为直连网段内已不存在该组成员，在 IGMP 组成员记录中删除该接口。超时时间等于 last-member-query-interval 与 last-member-query-count 的乘积再加上 1/2 的 query-max-response-time。

配置最后成员查询次数

【命令格式】 `ip igmp last-member-query-count count`

【参数说明】 *count*：特定组查询报文的发送次数。取值范围 2-7，缺省值 2。

【命令模式】 接口配置模式

【使用指导】 该命令只适用于 IGMPv2 或 IGMPv3。

当接口收到 Leave 报文后，会连续发送特定组查询报文，并等待主机回应。超时后，认为直连网段内已不

存在该组成员，在 IGMP 组成员记录中删除该接口。超时时间等于 last-member-query-interval 与 last-member-query-count 的乘积再加上 1/2 的 query-max-response-time。

配置普通成员查询间隔

【命令格式】 **ip igmp query-interval seconds**

【参数说明】 *seconds*：普通成员查询间隔。单位 S，取值范围 1~18000，缺省值 125。

【命令模式】 接口配置模式

【使用指导】 -

配置最大响应间隔

【命令格式】 **ip igmp query-max-response-time seconds**

【参数说明】 *seconds*：最大响应时间。单位 S，取值范围 1~25，缺省值 10。

【命令模式】 接口配置模式

【使用指导】 接口在发出查询报文后，等待响应。超时则认为直连网段内不存在组成员，删除组信息。

配置其他查询者存活时间

【命令格式】 **ip igmp query-timeout seconds**

【参数说明】 *seconds*：其它查询者存活时间，单位为 S，取值范围 60~300，缺省值 255。

【命令模式】 接口配置模式

【使用指导】 接口发出查询报文后，等待收到其他设备发出的查询报文。如果超时，则认为直连网段内只存在自己一个查询者。

配置举例

配置基本的路由器参数

- 【配置方法】
- 配置 IGMP 基本功能。略
 - 配置 IGMP 版本 3。
 - 配置最后成员的查询间隔是 15 (1.5S)。
 - 配置最后成员的查询次数是 3 次。
 - 配置普通成员查询间隔是 130S。
 - 配置最大响应时间是 15S。
 - 配置接口查询者的存活时间是 280S。

```
VSU(config-if-GigabitEthernet 0/5)#ip igmp version 3
VSU(config-if-GigabitEthernet 0/5)#ip igmp last-member-query-count 3
VSU(config-if-GigabitEthernet 0/5)#ip igmp last-member-query-interval 15
VSU(config-if-GigabitEthernet 0/5)#ip igmp query-interval 130
VSU(config-if-GigabitEthernet 0/5)#ip igmp query-max-response-time 15
VSU(config-if-GigabitEthernet 0/5)#ip igmp query-timeout 280
```

【检验方法】 通过 **show ip igmp interface interface-type interface-number** 命令，可以直接查看接口的 igmp 功能。

- 【配置方法】
- 配置 IGMP 基本功能。略
 - 配置 IGMP 版本 3。
 - 配置最后成员的查询间隔是 15 (1.5S)。
 - 配置最后成员的查询次数是 3 次。
 - 配置普通成员查询间隔是 130S。
 - 配置最大响应时间是 15S。
 - 配置接口查询者的存活时间是 280S。

```
VSU(config-if-GigabitEthernet 0/5)#ip igmp version 3
VSU(config-if-GigabitEthernet 0/5)#ip igmp last-member-query-count 3
VSU(config-if-GigabitEthernet 0/5)#ip igmp last-member-query-interval 15
VSU(config-if-GigabitEthernet 0/5)#ip igmp query-interval 130
VSU(config-if-GigabitEthernet 0/5)#ip igmp query-max-response-time 15
VSU(config-if-GigabitEthernet 0/5)#ip igmp query-timeout 280
VSU#show ip igmp interface gigabitEthernet 0/5
Interface GigabitEthernet 0/5 (Index 5)
  IGMP Enabled, Active, Querier, Version 3
  Internet address is 192.168.1.90
  IGMP interface limit is 1024
  IGMP interface has 1 group-record states
  IGMP interface has 0 static-group records
  IGMP activity: 3 joins, 0 leaves
  IGMP query interval is 130 seconds
  IGMP querier timeout is 280 seconds
  IGMP max query response time is 15 seconds
  Last member query response interval is 15
  Last member query count is 3
  Group Membership interval is 275 seconds
  Robustness Variable is 2
VSU#
```

常见错误

- 没有启动 IGMP 基本功能

2.4.3 支持IGMP组的过滤

配置效果

- 路由器对 IGMP 组信息过滤。

注意事项

- 必须配置 IGMP 基本功能。

配置方法

配置组播组访问控制

可选配置

若无特殊要求，可在所有的直连本地网络的路由器的接口上配置。

配置 IGMP 组成员数量限制

可选配置

若无特殊要求，可在所有的直连本地网络的路由器的接口上配置。

检验方法

IGMP 组的过滤

在接口上，设置只允许链表 1 的组加入，链表 1 的访问地址是 (225.0.0.1 ~ 225.0.0.255)。

设置接口加入一个组信息 225.0.0.5。

设置接口加入一个组信息 236.0.0.5。

查看当前接口的组信息。

IGMP 组成员数量限制

在接口上，设置成员数量为 5。

设置接口加入一个组信息 (225.0.0.5 ~ 225.0.0.10)。

查看接口的组信息。

相关命令

配置组播组访问控制

【命令格式】 **ip igmp access-group access-list**

【参数说明】 access-list：引用 IP 标准 ACL 或 IP 扩展 ACL 定义组地址范围，取值范围：1-199、1300-2699、WORD。

【命令模式】 接口配置模式

【使用指导】 通过在接口上配置此命令后，控制直连网段内的主机能够加入的组。引用 ACL 限定组地址范围。如果接收到被 ACL Deny 的 Report 报文，将丢弃此报文。

在启用 IGMPv3 时，此命令支持扩展 ACL。当接收到的 IGMP Report 信息是 (S1,S2,S3...Sn,G) 时，此命令将对(0,G)信息使用相应的 ACL 进行匹配检查。所以要正常使用该命令，必须对扩展 ACL 显式配置一条(0, G)，这样才可以正常的对 (S1,S2,S3...Sn,G) 进行过滤。

配置 IGMP 组成员数量限制

【命令格式】 **ip igmp limit number [except access-list]**

【参数说明】 number：IGMP 组成员记录的最大数量。取值范围和具体设备有关。接口缺省值 1024，全局缺省值 65536。

except access-list : access-list 中的组不参与计数。

access-list 为 IP 标准 ACL，取值范围：1-99、1300-1999、WORD。

【命令模式】 接口配置模式

【使用指导】 全局模式：限制整机 IGMP 组成员记录的数量。

接口模式：限制包含该接口的 IGMP 组成员记录的数量。

如果组成员记录超出了接口或全局的限制，此后接收到的 Report 报文将被忽略。

如果配置了 Except 例外列表，则可以正常受理指定范围内的 Report 报文，因此形成的组成员记录不参与计数。

接口和全局可以独立配置。如果全局配置的个数限制比接口小，则以全局配置为准。

配置举例

配置组过滤

- 【配置方法】
- 配置 IGMP 基本功能。略
 - 配置链表 1 的访问地址是 (225.0.0.1 ~ 225.0.0.255)。
 - 设置加入组为 225.0.0.5。
 - 设置加入组为 236.0.0.5。

```
VSU(config)#access-list 1 permit 225.0.0.1 225.0.0.255
VSU(config-if-GigabitEthernet 0/5)#ip igmp access-group 1
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.5
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 236.0.0.5
```

- 【检验方法】 通过 **show ip igmp groups [interface-type interface-number] [group-address] [detail]** 命令，可以直接查看接口的组信息。

```
VSU(config-if-GigabitEthernet 0/5)#show ip igmp groups
IGMP Connected Group Membership
Group Address      Interface          Uptime    Expires    Last Reporter
225.0.0.5          GigabitEthernet 0/5 00:14:00 00:02:45 192.168.1.90
```

IGMP 组成员数量限制

- 【配置方法】
- 配置 IGMP 基本功能。略
 - 配置接口组成员的限制为 5
 - 添加组信息 (225.0.0.5 ~ 225.0.0.12)。
 - 查看组信息

```
VSU(config-if-GigabitEthernet 0/5)#ip igmp limit 5
VSU(config-if-GigabitEthernet 0/5)#
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.5
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.6
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.7
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.8
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.9
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.10
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.11
VSU(config-if-GigabitEthernet 0/5)#ip igmp join-group 225.0.0.12
```

- 【检验方法】 通过 **show ip igmp groups [interface-type interface-number] [group-address] [detail]** 命令，可以直

接查看接口的组信息。

```
VSU(config-if-GigabitEthernet 0/5)#show ip igmp groups
IGMP Connected Group Membership
Group Address      Interface          Uptime    Expires    Last Reporter
225.0.0.5         GigabitEthernet 0/5  00:20:15  00:03:09  192.168.1.90
225.0.0.6         GigabitEthernet 0/5  00:20:24  00:02:58  192.168.1.90
225.0.0.7         GigabitEthernet 0/5  00:00:15  00:04:29  192.168.1.90
225.0.0.8         GigabitEthernet 0/5  00:00:13  00:04:34  192.168.1.90
225.0.0.9         GigabitEthernet 0/5  00:00:11  00:04:33  192.168.1.90
```

常见错误

- 没有启动 IGMP 基本功能

2.4.4 支持IGMP代理功能

配置效果

- 配置路由器代理功能，收集本地成员信息。

注意事项

- 必须配置 IGMP 基本功能。

配置方法

配置 IGMP PROXY-SERVICE

可选配置

若无特殊要求，可与上游直连的路由器的接口上配置。

配置 IGMP MROUTE-PROXY

可选配置

若无特殊要求，可与下游直连的主机的接口上配置。

检验方法

- 设置接口 7 与上游路由器直连的接口，设置成组播代理服务器。
- 设置接口 1 与下游主机直连的接口，设置成组播代理。
- 设置接口 1 有组 225.0.0.6，225.5.5.5 加入。
- 查看当前组信息。

相关命令

配置 IGMP PROXY-SERVICE

【命令格式】 **ip igmp proxy-service**

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 使用 **ip igmp proxy-service** 命令，将上联口设置为 Proxy-Service 接口。

使用 **ip igmp mroute-proxy** 命令，将下联口设置为 Mroute-Proxy 接口。

由 Proxy-Service 接口向 Mroute-Proxy 接口转发 IGMP 查询报文。由 Mroute-Proxy 接口向 Proxy-Service 接口转发 IGMP Report 报文。

设备最多只允许配置 32 个 Proxy-Service 接口。当 Proxy-Service 接口接收到 IGMP 查询报文后 根据 IGMP 组成员记录作出应答。

如果对 Proxy-Service 接口执行了 Switchport 命令 则 Mroute-Proxy 接口上所配置的 **ip igmp mroute-proxy** 命令将被自动删除。

配置 IGMP MROUTE-PROXY

【命令格式】 **ip igmp mroute-proxy interface-type interface-number**

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 使用 **ip igmp proxy-service** 命令，将上联口设置为 Proxy-Service 接口。

使用 **ip igmp mroute-proxy** 命令，将下联口设置为 Mroute-Proxy 接口。

由 Proxy-Service 接口向 Mroute-Proxy 接口转发 IGMP 查询报文。由 Mroute-Proxy 接口向 Proxy-Service 接口转发 IGMP Report 报文。

配置举例

- 【配置方法】
- 配置 IGMP 基本功能。略
 - 配置接口 7 为代理服务器。
 - 配置接口 1 为组播代理。
 - 接口 1 上设置加入组为 225.0.0.6 , 225.5.5.5。

```
VSU(config-if-GigabitEthernet 0/7)#ip igmp proxy-service
VSU(config-if-GigabitEthernet 0/7)#exit
VSU(config)#int gi 0/1
VSU(config-if-GigabitEthernet 0/1)#ip igmp mroute-proxy gigabitEthernet 0/7
VSU(config-if-GigabitEthernet 0/1)#ip igmp join-group 225.0.0.6
VSU(config-if-GigabitEthernet 0/1)#ip igmp join-group 225.5.5.5
```

- 【检验方法】 通过 **show ip igmp groups [interface-type interface-number] [group-address] [detail]**命令，可以直接查看接口的组信息。

```
VSU(config-if-GigabitEthernet 0/1)#show ip igmp groups
IGMP Connected Group Membership
Group Address      Interface          Uptime    Expires    Last Reporter
225.0.0.6          GigabitEthernet 0/1 00:23:05 00:02:40 192.168.36.90
225.5.5.5          GigabitEthernet 0/1 00:22:06 00:02:41 192.168.36.90

IGMP Proxy-server Connected Group Membership
Group Address      Interface          Uptime
```

```
225.0.0.6      GigabitEthernet 0/7    00:23:05
225.5.5.5      GigabitEthernet 0/7    00:22:06

VSU(config-if-GigabitEthernet 0/1)#
```

常见错误

- 没有启动 IGMP 基本功能

2.4.5 支持SSM-MAP功能

配置效果

- IGMPv3 支持源过滤功能，但 IGMPv1 和 IGMPv2 不支持源过滤功能，但提供了 SSM-MAP 功能，也解决了源过滤功能。

注意事项

- 必须配置 IGMP 基本功能。

配置方法

启动 SSM-MAP 功能

启动 SSM-MAP 功能，必须配置
启动 SSM-MAP 功能的路由器上。

配置 SSM-MAP STATIC 链表

可选配置
启动 SSM-MAP 功能的路由器上。

检验方法

通过 **show ip igmp ssm-mapping** [*group-address*] 命令，显示 SSM-MAP 信息。

相关命令

启动 SSM-MAP 功能

- 【命令格式】 **ip igmp ssm-map enable**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 使用 **ip igmp ssm-map enable** 命令，启动 SSM-MAP 功能。

使用 **ip igmp ssm-map static** 命令，设置静态映射表项。
接口运行 IGMPv3。当收到的 IGMPv1 或 IGMPv2 的 Report 报文时，添加静态映射的源地址。

配置 SSM-MAP STATIC 链表

- 【命令格式】 **ip igmp ssm-map static** *access-list source-address*
- 【参数说明】 *access-list* : 由 IP 标准 ACL 设置的组地址范围，取值范围：1-99、1300-1999、WORD
source-address : 源地址
- 【命令模式】 全局配置模式
- 【使用指导】 使用 **ip igmp ssm-map enable** 命令，启动 SSM-MAP 功能。
使用 **ip igmp ssm-map static** 命令，设置静态映射表项。
接口运行 IGMPv3。当收到的 IGMPv1 或 IGMPv2 的 Report 报文时，添加静态映射的源地址。

配置举例

- 【配置方法】
- 配置 IGMP 基本功能。略
 - 配置 SSM-MAP 使能。
 - 配置 SSM-MAP 静态链表 1。

```
VSU(config)#ip igmp ssm-map enable
VSU(config)#ip igmp ssm-map static 1 192.168.5.9
```

- 【检验方法】 通过 **show ip igmp ssm-mapping [group-address]**命令，可以直接查 SSM-MAP 信息。

```
VSU#show ip igmp ssm-mapping
SSM Mapping : Enabled
Database    : Static mappings configured
```

常见错误

- 没有启动 IGMP 基本功能

2.4.6 支持IGMP报文Router Alert选项

配置效果

- 支持对 IGMP 报文的 Router_Alert 选项进行检查，丢弃未携带 Router_Alert 选项的 IGMP 报文。
- 支持发送 IGMP 报文时携带 Router_Alert 选项。

注意事项

- 必须配置 IGMP 基本功能。

配置方法

配置对 IGMP 报文 Router Alert 选项检查

可选配置

若无特殊要求，可在配置 IGMP 基本功能的路由器上配置。

配置发送 IGMP 报文携带 Router Alert 选项

可选配置

若无特殊要求，可在配置 IGMP 基本功能的路由器上配置。

检验方法

配置对 IGMP 报文 Router Alert 选项检查

- 配置 IGMP 基本功能。
- 配置对 IGMP 报文 Router Alert 选项检查。
- 在配置 IGMP 基本功能的接口上只处理携带 Router Alert 选项的 IGMP 报文，未携带 Router_Alert 选项的 IGMP 报文被丢弃。

配置发送 IGMP 报文携带 Router Alert 选项

- 配置 IGMP 基本功能。
- 配置发送 IGMP 报文携带 Router Alert 选项。
- 在配置 IGMP 基本功能的接口上可以获取到携带 Router Alert 选项的 IGMP 报文。

相关命令

配置对 IGMP 报文 Router Alert 选项检查

【命令格式】 **ip igmp enforce-router-alert**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 使用 **ip igmp enforce-router-alert**，配置对 IGMP 报文 Router Alert 选项检查。
使用 **no ip igmp enforce-router-alert**，关闭对 IGMP 报文 Router Alert 选项检查。

配置发送 IGMP 报文携带 Router Alert 选项

【命令格式】 **ip igmp send-router-alert**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 使用 **ip igmp send-router-alert**，配置发送 IGMP 报文时携带 Router Alert 选项。
使用 **no ip igmp send-router-alert**，关闭发送 IGMP 报文时携带 Router Alert 选项。

配置举例

配置对 IGMP 报文 Router Alert 选项检查

【配置方法】 ● 配置 IGMP 基本功能。略

- 配置对 IGMP 报文 Router Alert 选项检查。

```
VSU(config)#ip igmp enforce-router-alert
```

【检验方法】 向配置了 IGMP 基本功能的接口发送携带 Router Alert 选项组地址为 225.1.1.1 的 IGMP report 报文，IGMP 能对该 IGMP report 报文进行处理，show ip igmp groups 能查看到 225.1.1.1 的组记录；向配置了 IGMP 基本功能的接口发送未携带 Router Alert 选项组地址为 225.1.1.2 的 IGMP report 报文，该 IGMP report 报文被丢弃，show ip igmp groups 不能查看到对应的 225.1.1.2 组记录

配置发送 IGMP 报文携带 Router Alert 选项

- 【配置方法】
- 配置 IGMP 基本功能。略
 - 配置发送 IGMP 报文携带 Router Alert 选项。

```
VSU(config)#ip igmp send-router-alert
```

【检验方法】 在配置 IGMP 基本功能的接口上可以获取到携带 Router Alert 选项的 IGMP 报文。

常见错误

- 没有启动 IGMP 基本功能

2.5 监视与维护

清除各类信息

作用	命令
清除 IGMP 缓存中通过响应消息得来的动态组成员的信息。	clear ip igmp group
清除 IGMP 缓存中接口的信息。	clear ip igmp interface <i>interface-type interface-number</i>

查看运行情况

作用	命令
显示直连子网中所有组的状况。	show ip igmp groups
显示直连子网中所有组的详细信息。	show ip igmp groups detail
显示直连子网中指定组的状况。	show ip igmp groups <i>A.B.C.D</i>
显示直连子网中指定组的详细信息。	show ip igmp groups <i>A.B.C.D detail</i>
显示直连子网中指定接口的 IGMP 配置信息。	show ip igmp interface <i>interface-type interface-number</i>
显示直连子网中指定接口的所有组的详细信息。	show ip igmp groups <i>interface-type interface-number detail</i>
显示直连子网中指定接口的特定组的信息。	show ip igmp groups <i>interface-type interface-number A.B.C.D</i>
显示直连子网中指定接口的特定组的详细信息。	show ip igmp groups <i>interface-type interface-number A.B.C.D detail</i>

显示 IGMP 接口的配置信息。	show ip igmp interface [<i>interface-type interface-number</i>]
显示 IGMP 所有接口的配置信息。	show ip igmp interface
显示 IGMP SSM-MAP 的配置信息。	show ip igmp ssm-mapping
显示 IGMP SSM-MAP 对组 A.B.C.D 的映射信息。	show ip igmp ssm-mapping A.B.C.D

查看调试信息

作用	命令
显示 IGMP 调试开关的打开情况。	show debugging
打开 IGMP 所有调试信息开关。	debug ip igmp all
打开 IGMP 调试报文解析开关。	debug ip igmp decode
打开 IGMP 调试报文编码开关。	debug ip igmp encode
打开 IGMP 调试事件信息开关。	debug ip igmp events
打开 IGMP 调试事件信息开关。	debug ip igmp fsm
打开 IGMP 调试状态机信息开关。	debug ip igmp tib
打开 IGMP 调试警告开关。	debug ip igmp warning

3 PIM-DM

3.1 概述

PIM (Protocol Independent Multicast , 协议无关组播) 是一种域内组播路由协议。

组播源向组地址发送报文。报文经过网络设备逐跳转发，最终到达组成员。在三层网络设备上，可以使用 PIM 创建并维护组播路由表项，从而支持组播转发。

PIM 有两种模式 :PIM-SM(Protocol Independent Multicast Sparse Mode , 协议无关组播 - 稀疏模式)、PIM-DM(Protocol Independent Multicast Dense Mode , 协议无关组播 - 密集模式)。

- PIM-SM 适用于组成员分布相对分散、范围较广的大规模网络。
- PIM-DM 适用于组成员分布相对密集的小型网络。

协议规范

- RFC3973 : Protocol Independent Multicast - Dense Mode (PIM-DM)
- RFC2715 : Interoperability Rules for Multicast Routing Protocols

3.2 典型应用

典型应用	场景描述
在同一网络内提供组播服务	在同一网络内提供组播服务。

3.2.1 在同一网络内提供组播服务

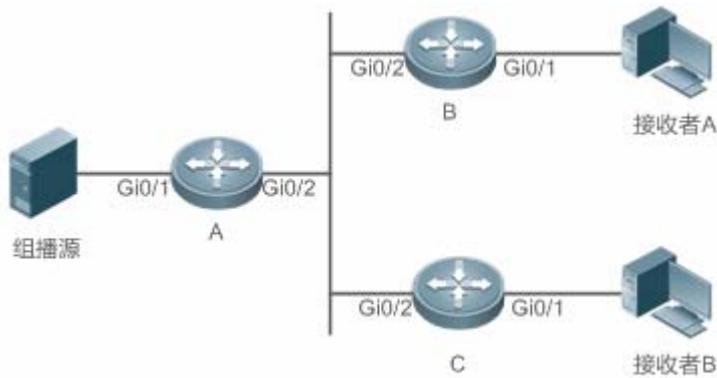
应用场景

在同一网络内部提供组播服务。

以下图为例

- 组播源发出的组播报文，同一网络内的接收者 A 与接收者 B 接收组播源发出的组播报文。

图 3-1



【注释】 A、B、C 为三层路由设备。

将组播源接入 A 的 Gi0/1，将接收者 B 接入 B 的 Gi0/1，将接收者 B 接入 C 的 Gi0/1。

功能部属

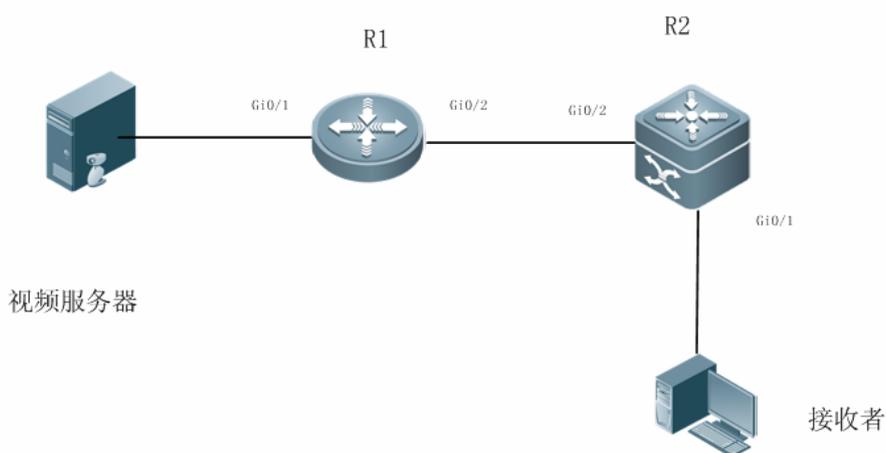
- 在同一网络中运行 OSPF 协议，实现单播路由。
- 在同一网络中运行 PIM-DM 协议，实现组播路由。
- 在用户主机网段运行 IGMP 协议，实现组成员管理。

3.2.2 热备环境中使用PIM-DM

应用场景

在热备环境中，运行 PIM-DM，设备进行热备切换，保证不断流。

图 3-2



【注释】 R1 与视频服务器相连，R2 与接收者直接相连，R2 为热备环境

R1 与 R2 上运行三层组播

功能部属

- R1 与 R2 上运行 OSPF，实现单播路由
- R1 与 R2 上运行组播 PIM-DM，实现组播路由
- R2 上双机热备运行

【注释】

在环境中，R2 有可能会执行热备切换，此时需要调整 pim hello 报文的查询间隔(默认值为 30s)，因为在执行热备切换过程中，R1 的 pim hello 的邻居的保活定时器可能已经超时了(默认为 3.5 倍的查询间隔，即 105s)，所以需要调整 R2 的 pim hello 查询间隔。因为目前组播需要依赖于单播，单播收敛完毕，组播才开始收敛，例如，目前单播默认 gr 收敛时间为 120s，建议配置 pim hello 间隔为 60s，这样 pim hello 邻居的保活时间为 210s。在此场景中，配置 pim hello 间隔需要参照单播 gr 收敛时间，配置完要保证 3.5 倍的 pim hello 查询间隔大于单播 gr 收敛时间。在热备环境中，建议配置 pim hello 间隔要大于默认值 30s，否则在热备切换过程中，可能导致对端 pim hello 邻居超时。

3.3 功能详解

基本概念

▾ PIM 路由器、PIM 接口

启用了 PIM 协议的路由器称为 PIM 路由器。使能了 PIM 协议的接口称为 PIM 接口。

在 PIM 路由器上执行组播报文的转发。接收组播报文的 PIM 接口称为上游接口；发出组播报文的 PIM 接口称为下游接口。上游接口所在网段称为上游网段。下游接口所在网段称为下游网段。

▾ PIM 网络、PIM 域

PIM 路由器之间通过 PIM 接口相连，组成 PIM 网络。

在某些 PIM 接口上设置“边界”，可以将一个大的 PIM 网络划分为多个 PIM 域。“边界”可以拒绝特定组播报文通过，或者限制 PIM 消息的传输。

▾ 组播分发树

组播报文从一点传输到多点，转发路径呈现树型结构。此转发路径被称为组播分发树（MDT，Multicast Distribution Tree

▾ (*,G)、(S,G)

- (*,G)：任意源发往组 G 的报文、此类报文对应的路由表项、此类报文对应的转发路径（RPT）。
- (S,G)：源 S 发往组 G 的报文、此类报文对应的路由表项、此类报文对应的转发路径（SPT）。

功能特性

功能特性	作用
建立PIM邻居	PIM 路由器之间建立邻居关系，从而形成 PIM 网络。
PIM状态更新	PIM-DM 使用状态更新消息（State Refresh Message）来更新网络的状态信息。

3.3.1 建立PIM邻居

PIM 路由器之间建立邻居关系，从而形成 PIM 网络。PIM 路由器之间必须首先建立邻居关系，才能交互其他 PIM 控制消息或转发组播报文。

工作原理

PIM 接口对外发送 Hello 消息。封装 Hello 消息的组播报文的目地地址为 224.0.0.13（表示同一网段中所有 PIM 路由器）、源地址为接口的 IP 地址、TTL 数值为 1。封装 Hello 消息的 IPv6 组播报文的目地地址为 ff02::d。

Hello 消息的作用：发现邻居、协调各项协议参数、维持邻居关系。

📌 发现 PIM 邻居

同一网段中的 PIM 路由器都会收到目的地址为 224.0.0.13 或 ff02::d 的组播报文，从而获得邻居信息，建立邻居关系。

当接口刚启动或是接口检测到新邻居时，会使用 Triggered-Hello-Delay 来产生一个随机时间，在这个随机时间内，接口会往外发送 hello 报文。

📌 协调各项协议参数

Hello 消息中携带多项协议参数，介绍如下：

- DR_Priority：表示各路由器接口竞选 DR 的优先级，优先级越高越容易获胜。
- Holdtime：表示保持邻居为可达状态的超时时间。
- LAN_Delay：表示共享网段内传输 Prune 消息的延迟时间。
- Override-Interval：表示 Hello 消息中携带的否决剪枝时间。

当 PIM 路由器从上游接口接收到 Prune 消息时，说明该共享网段内还存在其他下游。如果这台 PIM 路由器仍然需要接收组播数据，则必须在 Override-Interval 时间内向上游发送剪枝否决消息。

$LAN_Delay + Override-Interval = PPT$ (Prune-Pending Timer)。PIM 路由器接收到下游的 Prune 消息后，并不立即执行剪枝，而是等待 PPT 时间，超时则剪枝。在 PPT 时间内如果收到下游发来的剪枝否决消息，则取消剪枝。

📌 维持邻居关系

PIM 路由器之间周期性地发送 Hello 消息。如果 Holdtime 超时还没有收到该 PIM 邻居发出的新的 Hello 报文，则认为该邻居不可达，将其从邻居列表中清除。PIM 邻居的变化将导致网络中组播拓扑的变化。如果组播分发树上的某上游邻居或下游邻居不可达，将导致组播路由重新收敛，组播分发树迁移。

相关配置

📌 启动接口上的 PIM-DM 功能

缺省情况下，接口上的 PIM-DM 功能关闭。

使用 `ip pim dense-mode` 命令可以启动或关闭接口上的 PIM-DM 功能。

必须在接口上启用 PIM-DM 功能，接口才能参与 PIM 协议工作。

📌 调整接口上的 Hello 消息发送间隔

缺省情况下，Hello 消息发送间隔为 30s。

使用 `ip pim query-interval interval-seconds` 命令可以调整接口上的 Hello 消息发送间隔，取值范围是<1-65535>。

`interval-seconds` 值越大，Hello 消息发送间隔越大。`interval-seconds` 值越小，Hello 消息发送间隔越小。

3.3.2 PIM组播树建立

PIM-DM 使用扩散、剪枝、嫁接方式建立组播树。

工作原理

组播源开始发送组播数据报文时，系统会往组播邻居和本地成员的出口转发组播流，从设备上游接口接收到的报文都要经过 RPF (Reverse Path Forwarding, 反向路径转发) 检查，没有通过 RPF 检查的报文将被丢弃。对于通过了 RPF 检查的组播报文，存在出口继续转发，不存在出口，向上游设备发送剪枝报文，上游接口接收到剪枝报文以后，把发送该剪枝报文的接口记为剪枝状态 (Pruned)，并设置一个剪枝状态计时器。这样就建立了一棵以组播源为根的组播转发树。

系统收到本地成员加入，如果处于剪枝状态的下游设备向上游设备发送嫁接消息。上游设备收到这条嫁接消息以后，回应一条嫁接确认 (Graft-Ack) 消息，并重新向该设备接口转发组播数据报文。

- i** 当在环境部署中，设备间通过多条链路建立多个 PIM-DM 邻居，并且下游设备无接收需求或者有少量接收需求的时候，可能导致 cpu 偏高。在这种场景下建议使用 PIM-SM 模式进行环境部署。

相关配置

配置接口的剪枝否决时间

缺省情况下，剪枝否决时间为 500 毫秒。

使用 `ip pim override-interval interval-milliseconds` 命令，修改剪枝否决时间。

3.3.3 PIM状态更新

PIM-DM 使用状态更新消息 (State Refresh Message) 来更新网络的状态信息。

工作原理

与组播源直接相连的设备定期向下游设备发送状态更新消息，以通告网络的拓扑变化情况。收到状态更新消息的设备通过修改消息中的某些字段把本机的拓扑状态信息也加入到消息中，然后发送给下游设备。到达叶子设备时，整个网络的状态信息从上到下都得到了更新。

相关配置

禁止处理和转发状态更新消息

缺省情况下，允许处理和转发状态更新消息。

使用 `ip pim state-refresh disable` 命令可以禁止处理和转发状态更新消息。

- ❗ 关闭状态刷新功能可能会导致已经收敛的 PIM-DM 组播转发树重新收敛，造成不必要的带宽浪费和组播路由表振荡，所以一般情况下最好都不要关闭状态刷新功能。

调整 PIM 状态更新消息发送间隔

缺省情况下，状态更新消息发送间隔为 60s。

使用 `ip pim state-refresh origination-interval interval-seconds` 命令可以调整接口上的状态更新消息发送间隔，取值范围是 <1-100>。

`interval-seconds` 值越大，状态更新消息发送间隔越大。`interval-seconds` 值越小，状态更新消息发送间隔越小。

- ❗ 只有与组播源直接相连的设备才会周期性地往下游接口发送 PIM 状态更新消息，所以如果设备不是与组播源直接相连的，那么在它的下游接口上配置 PIM 状态更新消息发送间隔是无效的。

3.3.4 MIB功能

SNMP 管理器在与其他代理连接时，使用 MIB 中的信息，直接管理 PIM-DM 功能。

工作原理

管理信息库 MIB 指明了网络元素所维持的变量（即能够被管理进程查询和设置的信息），直接管理 PIM-DM 功能。

相关配置

调整 PIM-DM 的 mib 功能开启

缺省情况下，PIM-SM 的 mib 功能起作用。

使用 `ip pim mib dense-mode` 命令，启动 PIM-DM 的 mib 功能。

3.4 配置详解

配置项	配置建议 & 相关命令	
配置PIM-DM基本功能	⚠ 必须配置。用于建立组播服务。	
	<code>ip multicast-routing</code>	启动 IPv4 组播路由功能
	<code>ip pim dense-mode</code>	启动 PIM-DM 功能
配置PIM邻居	⚠ 可选配置。用于限制 ASM 模型下的合法组播报文的(S,G)地址范围。	
	<code>ip pim query-interval interval-seconds</code>	在接口上设置的 Hello 消息发送间隔。
	<code>ip pim propagation-delay interval-milliseconds</code>	在接口上设置发送剪枝报文的延迟时间。
	<code>ip pim override-interval interval-milliseconds</code>	在接口上设置剪枝否决时间。
配置PIM状态更新功能	<code>ip pim neighbor-filter access-list</code>	在接口上启动邻居过滤功能。
	<code>ip pim state-refresh disable</code>	禁止处理和转发状态更新消息。

	ip pim state-refresh origination-interval <i>interval-seconds</i>	在接口上设置状态更新消息发送间隔。
MIB 功能	ip pim mib dense-mode	启动 PIM-DM 的 mib 功能。

3.4.1 配置PIM-DM基本功能

配置效果

- 建立 PIM 网络，向网络内的数据源和用户终端提供 IPv4 组播服务。

注意事项

- PIM-DM 需要借助网络中现有的单播路由。因此，网络中必须配置 IPv4 单播路由。

配置方法

启动 IPv4 组播路由功能

- 必须配置。
- 若无特殊要求，应在每台路由器上启动 IPv4 组播路由功能。

启动 PIM-DM 功能

- 必须配置。
- 若无特殊要求，应在以下接口上启动 PIM-DM 功能：路由器互联接口，连接组播源、用户主机的接口。

检验方法

使网络中的组播源发送组播报文，用户主机加入这些组。

- 检查用户主机能否成功接收每个组的报文。
- 检查路由器上是否建立了正确的 PIM-DM 路由表项。

相关命令

启动 IPv4 组播路由功能

【命令格式】 **ip multicast-routing**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

启动 PIM-DM 功能

【命令格式】 **ip pim dense-mode**

【参数说明】 -

- 【命令模式】 接口模式
- 【使用指导】 PIM 接口必须是三层接口，包括：路由口、L3AP、SVI、loopback 接口。
所有 PIM 接口应 IPv4 单播路由可达。

查看 PIM-SM 路由表

- 【命令格式】 **show ip pim dense-mode mroute** [*group-or-source-address* [*group-or-source-address*]] [**summary**]
- 【参数说明】 *group-or-source-address*：组地址或源地址。
group-or-source-address：组地址或源地址（两个地址不能同时都是组地址或者同时都是源地址）
summary：显示路由表项的概要信息
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 查看是否具备足够的路由表项。
查看路由表项的上游接口、下游接口列表，确认构建了正确的生成树。

配置举例

i 以下配置举例，仅介绍与 PIM-DM 相关的配置。

在 IPv4 网络上建立 IP 组播服务

【网络环境】

图 3-3



- 【配置方法】
- 在路由器上配置 IPv4 单播路由协议（如 OSPF）
 - 在所有路由器上启动 IPv4 组播路由功能。
 - 在设备互联接口上、连接用户主机和组播源的接口上启动 PIM-DM 功能。

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if)# ip pim dense-mode
A(config-if)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if)# ip pim dense-mode
A(config-if)# exit
```

B

```
B# configure terminal
B(config)# ip multicast-routing
B(config)# interface GigabitEthernet 0/1
B(config-if)# ip pim dense-mode
```

```
B(config-if)# exit
B(config)# interface GigabitEthernet 0/2
B(config-if)# ip pim dense-mode
B(config-if)# exit
```

【检验方法】 使组播源 (192.168.1.10) 向 G (229.1.1.1) 发送报文。使接收者 A 加入 G。

- 检接收者 A 收到的组播报文，接收者 A 应该能够收到 G 的组播报文。
- 检查 A、B 上的 PIM-DM 路由表。

A

```
A# show ip pim dense-mode mroute
PIM-DM Multicast Routing Table
(192.168.1.10, 229.1.1.1)
  MRT lifetime expires in 182 seconds
  Source directly connected on GigabitEthernet 0/1
  State-Refresh Originator State: Originator
  SRT:57, SAT:147
  Upstream IF: GigabitEthernet 0/1
  Upstream State: Forwarding
  Assert State: NoInfo
  Downstream IF List:
  GigabitEthernet 0/2, in 'olist':
    Downstream State: NoInfo
    Assert State: NoInfo
```

B

```
B# show ip pim dense-mode mroute
PIM-DM Multicast Routing Table
(192.168.1.10, 229.1.1.1)
  MRT lifetime expires in 130 seconds
  RPF Neighbor: 192.168.2.1, Nexthop: 192.168.2.1, GigabitEthernet 0/2
  Upstream IF: GigabitEthernet 0/2
  Upstream State: Forwarding
  Assert State: Loser, AT:125
  Downstream IF List:
  GigabitEthernet 0/1, in 'olist':
    Downstream State: NoInfo
    Assert State: NoInfo
```

常见错误

- IPv4 单播路由配置错误。
- 某台路由器没有使能 IPv4 组播路由。
- 某个接口没有使能 PIM-DM。

3.4.2 配置PIM邻居

配置效果

- 协调各项协议参数，调整 Hello 报文参数。
- 设置邻居过滤，以提高网络的安全性。

注意事项

- 必须配置 PIM-DM 基本功能。

配置方法

- 若无特殊要求，可以在各个 PIM 路由器接口上设置参数。

检验方法

- 设置接口的 Hello 报文参数，通过 `debug ip pim dense-mode encode` 命令，查看构造的 Hello 报文的参数。
- 设置邻居过滤，可以通过 `show ip pim dense-mode decode` 命令，查看邻居过滤信息。
- 可以通过 `show running-config interface [interface-type interface-number]` 命令，查看接口上配置。

相关命令

配置 Hello 消息发送间隔

【命令格式】 `ip pim query-interval interval-seconds`

【参数说明】 `interval-seconds`：取值范围为<1-65535>，单位为秒。

【命令模式】 接口模式

【使用指导】 如果设置了 `hello interval`，`hello holdtime` 值会更新为 `hello interval` 的 3.5 倍。

i 每当 Hello 消息发送间隔被更新时，Hello 消息保持时间 (Hello hold time) 会自动更新为 Hello 消息发送间隔的 3.5 倍。如果 Hello 消息发送间隔 * 3.5 > 65535，则 Hello 消息保持时间被更新为 65535。

配置 Hello 消息中的传输延迟时间

【命令格式】 `ip pim propagation-delay interval-milliseconds`

【参数说明】 `interval-milliseconds`：取值范围<1-32767>，单位为毫秒

【命令模式】 接口模式

【使用指导】 配置接口的 `propagation-delay` 即配置接口的传播延迟时间。

配置 Hello 消息中的剪枝否决时间

【命令格式】 `ip pim override-interval interval-milliseconds`

【参数说明】 `interval-milliseconds`：取值范围<1-32767>，单位为毫秒

【命令模式】 接口模式

【使用指导】 配置接口的 `override-interval` 即配置接口的剪枝否决时间。

配置 PIM 邻居过滤

【命令格式】 **ip pim neighbor-filter** *access-list*

【参数说明】 *access-list* : 支持数字 ACL 范围为 1-99 ; 也支持命名 ACL。

【命令模式】 接口模式

【使用指导】 只有符合 ACL 过滤条件的邻居地址才能够作为当前接口的 PIM 邻居, 被 ACL 过滤的邻居地址将无法作为当前接口的 PIM 邻居。

所谓的邻接 (peering) 关系, 指的是 PIM 邻居之间的协议报文交互。如果中止了与某个 PIM 设备的邻接关系, 则不会与该设备形成邻居关系, 也不会接收来自该设备的 PIM 协议报文。

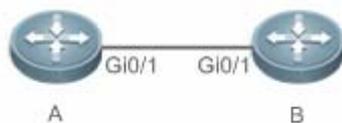
配置举例

i 以下配置举例, 仅介绍与 PIM-DM 相关的配置。

在 IPv4 网络上配置 PIM 邻居参数

【网络环境】

图 3-4



- 【配置方法】
- 配置 PIM-DM 基本功能(略)。
 - 在设备 A Gi0/1 接口上配置 Hello 报文各项协议参数。

```
A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if)# ip pim query-interval 60
A(config-if)# ip pim propagation-delay 800
A(config-if)# ip pim override-interval 1000
A(config-if)# exit
```

- 【检验方法】
- 通过 **show running-config interface [interface-type interface-number]** 命令, 查看接口上配置。
 - 通过 **debug ip pim dense-mode encode** 命令, 查看构造的 Hello 报文的参数。

```
A
A# (config)#show running-config interface gigabitEthernet 0/1

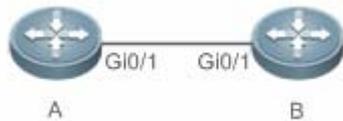
Building configuration...
Current configuration : 245 bytes
!
interface GigabitEthernet 0/1
 ip pim dense-mode
 ip pim query-interval 60
 ip pim propagation-delay 800
 ip pim override-interval 1000
A# debug ip pim dense-mode encode
```

```
*Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello Hold-Time 210
*Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello Gen-ID 1362200073
*Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello PD=800 ms, OI=1000 ms
*Dec 22 15:00:58: %7: [ENCODE] Enc Hello: Hello SR-Interval 60
*Dec 22 15:00:58: %7: [ENCODE] Enc Msg Hdr: Hello Checksum=65396, MsgLen=34 Assert State:
Loser, AT:125
```

在 IPv4 网络上配置 PIM 邻居过滤

【网络环境】

图 3-5



【配置方法】

- 配置 PIM-DM 基本功能(略)。
- 在设备 A 上配置 ACL。
- 在设备 A Gi0/1 接口上配置 PIM 邻居过滤。

```
A
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if)# ip pim query-interval 60
A(config-if)# ip pim propagation-delay 800
A(config-if)# ip pim override-interval 1000
A(config-if)# exit
```

【检验方法】

- 通过 **show running-config interface [interface-type interface-number]** 命令，查看接口上配置。
- 通过 **debug ip pim dense-mode decode** 命令，查看构造的 Hello 报文的参数。

```
A
A#show running-config interface gigabitEthernet 0/2

Building configuration...
Current configuration : 187 bytes
!
interface GigabitEthernet 0/1
 ip pim dense-mode
 ip pim neighbor-filter pim-dm

A# debug ip pim dense-mode decode
Dec 22 15:15:47: %7: [DECODE] Dec Msg: PIM Hello message, version 2
Dec 22 15:09:47: %7: [DECODE] Dec Msg: Neighbor 192.168.2.2/32 on GigabitEthernet 0/1 denied
by access-list pim-dm
```

常见错误

- IPv4 单播路由配置错误。
- 某台路由器没有使能 IPv4 组播路由。
- 某个接口没有使能 PIM-DM。

3.4.3 配置PIM状态更新功能

配置效果

- PIM 状态更新功能开启与关闭
- 调整 PIM 状态更新消息发送间隔。

注意事项

- 必须配置 PIM-DM 基本功能。

配置方法

- 在与组播源直接相连 PIM 路由口配置 PIM 状态更新消息发送间隔参数有效。

检验方法

- 配置 PIM 状态更新功能，可通过 **show running-config** 命令，查看是否配置 PIM 状态更新功能。
- 可通过 **show ip pim dense-mode track** 命令，查看 PIM 状态更新消息发送次数。
- 可以通过 **show running-config interface [interface-type interface-number]**命令，查看接口上配置。

相关命令

禁止接口处理和转发 PIM-DM 状态刷新消息功能

【命令格式】 **ip pim state-refresh disable**

【参数说明】 -。

【命令模式】 全局模式

【使用指导】 在关闭状态刷新功能时，将不会处理和转发状态刷新报文；在发送的 Hello 报文中将不包含状态刷新选项，在收到 Hello 报文时不会处理报文中的 SR Cap 字段。

关闭状态刷新功能可能会导致已经收敛的 PIM-DM 组播转发树重新收敛，造成不必要的带宽浪费和组播路由表振荡，所以一般情况下最好都不要关闭状态刷新功能。

设置 PIM-DM 的状态刷新消息的发起时间间隔

【命令格式】 **ip pim state-refresh origination-interval interval-seconds**

【参数说明】 *interval-seconds*：取值范围<1-100>，单位为秒

【命令模式】 接口模式

【使用指导】 -

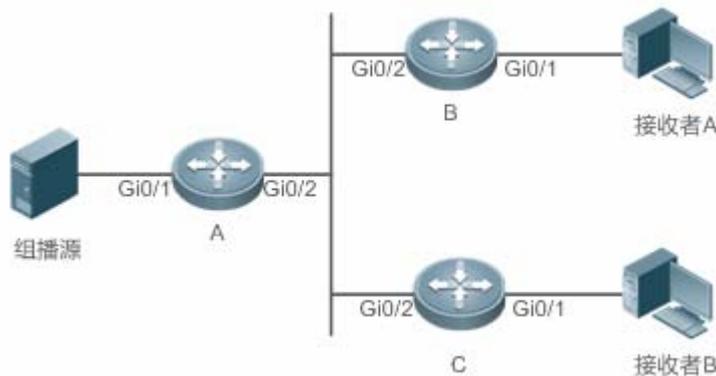
配置举例

 以下配置举例，仅介绍与 PIM-DM 相关的配置。

在 IPv4 网络上配置禁止接口处理和转发 PIM-DM 状态刷新消息功能

【网络环境】

图 3-6



【配置方法】

- 配置 PIM-DM 基本功能(略)。
- 在设备 A 上配置禁止接口处理和转发 PIM-DM 状态刷新消息功能。

A

```
A# configure terminal
A(config)# ip pim state-refresh disable
```

【检验方法】

- 通过 **show running-config** 命令，查看配置。

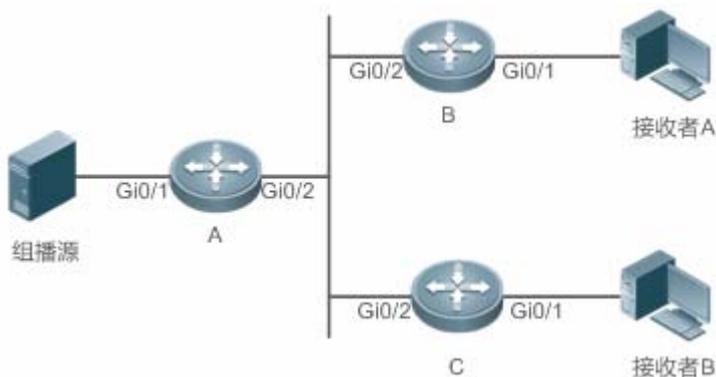
A

```
A# (config)# show running-config
...
!
ip pim state-refresh disable
!
...
```

在 IPv4 网络上配置 PIM-DM 的状态刷新消息的发起时间间隔

【网络环境】

图 3-7



【配置方法】

- 配置 PIM-DM 基本功能(略)。
- 在设备 A Gi0/1 接口上配置 PIM-DM 的状态刷新消息的发起时间间隔。

A

```
A# configure terminal
A(config)# interface GigabitEthernet 0/1
A(config-if)# ip pim state-refresh origination-interval 5
A(config-if)# exit
```

- 【检验方法】
- 通过 **show running-config interface** [*interface-type interface-number*]命令，查看接口上配置。
 - 可通过 **show ip pim dense-mode track** 命令，查看 PIM 状态更新消息发送次数。

A

```
A#show running-config interface gigabitEthernet 0/1

Building configuration...
Current configuration : 201 bytes
!
interface GigabitEthernet 0/1
 ip pim dense-mode
 ip pim state-refresh origination-interval 5

A #show ip pim dense-mode track
PIM packet counters
Elapsed time since counters cleared: 00:18:54

                received                sent
Valid PIMDM packets:    38                102
Hello:                   38                76
Join/Prune:              0                 0
Graft:                   0                 0
Graft-Ack:               0                 0
Assert:                  0                 0
State-Refresh:           0                 26
PIM-SM-Register:        0
PIM-SM-Register-Stop:   0
PIM-SM-BSM:             0
PIM-SM-C-RP-ADV:        0
Unknown Type:           0

Errors:
Malformed packets:      0
Bad checksums:          0
Unknown PIM version:    0
Send errors:            0
```

常见错误

- IPv4 单播路由配置错误。
- 某台路由器没有使能 IPv4 组播路由。
- 某个接口没有使能 PIM-DM。

3.4.4 MIB功能

配置效果

- PIM-DM 的 MIB 起作用。

检验方法

- 配置 PIM-SM 的 MIB 功能，可通过 `show running-config` 命令，查看是否配置 PIM-SM 功能。

相关命令

▾ PIM-DM 的 MIB 功能起作用

- 【命令格式】 `ip pim mib dense-mode`
- 【参数说明】 -。
- 【命令模式】 全局模式
- 【使用指导】 -

3.5 监视与维护

清除各类信息

作用	命令
重新设置报文统计的开始时间点，并清除 PIM 报文的计数器。	<code>clear ip pim dense-mode track</code>

查看运行情况

作用	命令
查看以 ip pim 为关键字的命令的帮助信息。	<code>ip pim help</code>
查看配置案例的帮助信息。	<code>pim-dm help</code>
查看接口的 PIM-DM 信息。	<code>show ip pim dense-mode interface [interface-type interface-number] [detail]</code>
查看 PIM-DM 邻居信息。	<code>show ip pim dense-mode neighbor [interface-type interface-number]</code>
查看 PIM-DM 的下一跳信息。	<code>show ip pim dense-mode nexthop</code>
查看 PIM-DM 的路由表信息。	<code>show ip pim dense-mode mroute [group-or-source-address [group-or-source-address]] [summary]</code>
查看从统计的开始时间点到现在所发送和接收的 PIM 报文的个数。	<code>show ip pim dense-mode track</code>
查看 PIM-DM 的运行状态。	<code>view pim-dm</code>

4 PIM-SM

4.1 概述

PIM(Protocol Independent Multicast，协议无关组播)是一种域内组播路由协议。

组播源向组地址发送报文。报文经过网络设备逐跳转发，最终到达组成员。在三层网络设备上，可以使用 PIM 创建并维护组播路由表项，从而支持组播转发。

PIM 定义了两种模式：PIM-SM (Protocol Independent Multicast Sparse Mode，协议无关组播 - 稀疏模式)、PIM-DM (Protocol Independent Multicast Dense Mode，协议无关组播 - 密集模式)。

- PIM-SM 适用于组成员分布相对分散、范围较广的大规模网络。
- PIM-DM 适用于组成员分布相对密集的小型网络。

 下文仅介绍 PIM-SM 的相关内容。

协议规范

- RFC4601 : Protocol Independent Multicast -Sparse Mode (PIM-SM)
- RFC5059 : Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)
- RFC3962 : Protocol Independent Multicast - Dense Mode protocol
- RFC4607 : Source-Specific Multicast for IP

4.2 典型应用

典型应用	场景描述
用PIM-SM实现ASM	接收者接收任意组播源。
用PIM-SM实现SSM	接收者对组播源进行选择。

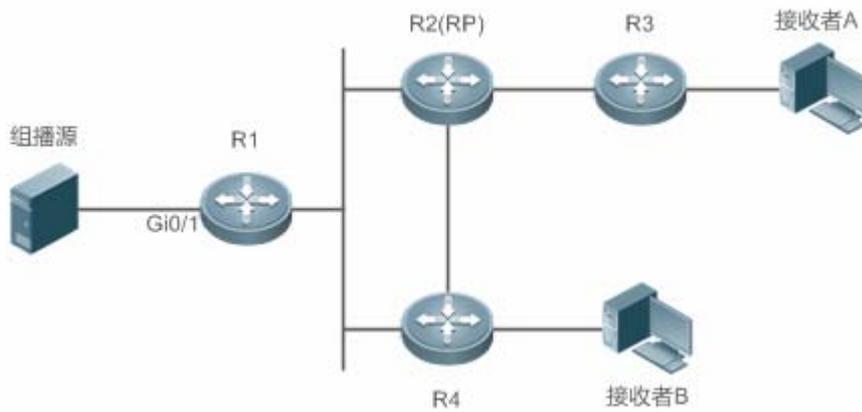
4.2.1 用PIM-SM实现ASM

应用场景

只在一个域内提供组播服务。

以下图为例，接收者接收任意组播源。

图 4-1



- 【注释】 R1 与组播源直接相连。
 R2 将被设置为 RP。
 R3 与接收者 A 直接相连。
 R4 与接收者 B 直接相连。

功能部属

- 在域内运行 OSPF 协议，实现单播路由。
- 在域内运行 PIM-SM 协议，实现组播路由。
- 在用户主机网段运行 IGMP 协议，实现组成员管理。

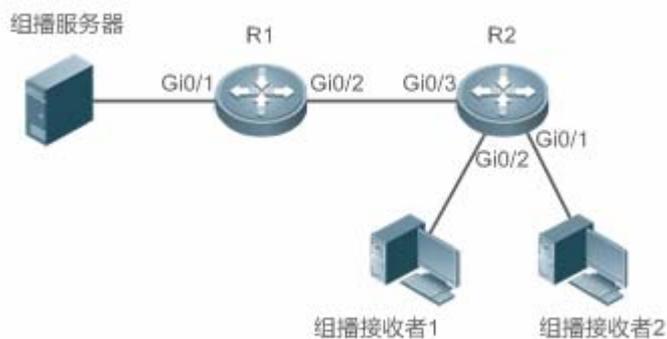
4.2.2 用PIM-SM实现SSM

应用场景

只在一个域内提供组播服务。

以下图为例，接收者接收特定组播源。

图 4-2



- 【注释】 R1 与组播源直接相连。

- R2 将被设置为 RP。
- R2 与接收者 A 直接相连。
- R2 与接收者 B 直接相连。

功能部属

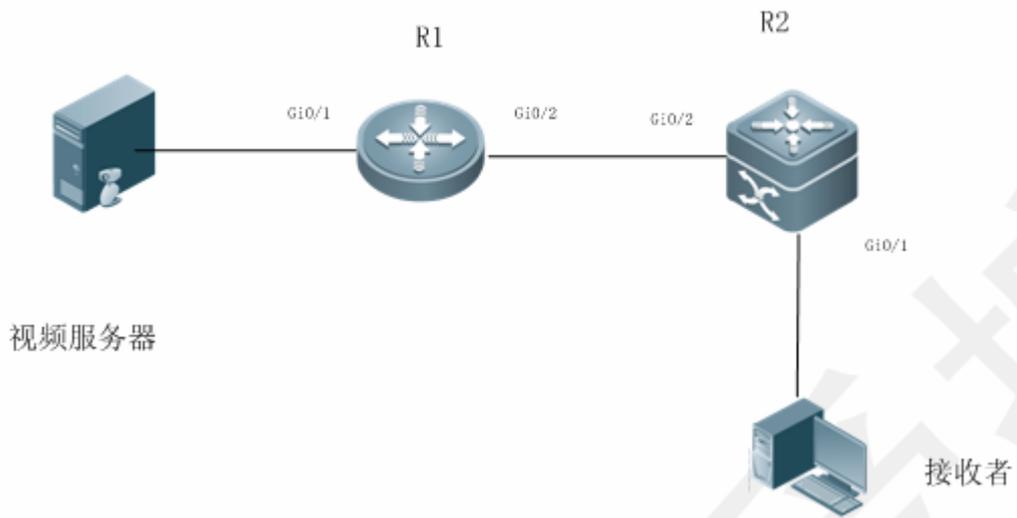
- 在域内运行 OSPF 协议，实现单播路由。
- 在域内运行 PIM-SM 协议，实现组播路由。
- 在域内运行 PIM-SM 协议的 SSM 功能，实现 SSM 功能。
- 在用户主机网段运行 IGMPv3 协议，实现组成员管理。

4.2.3 热备环境中使用PIM-SM

应用场景

在热备环境中，运行 PIM-SM，设备进行热备切换，保证不断流。

图 4-3



【注释】 R1 与视频服务器相连，R2 与接收者直接相连，R2 为热备环境
R1 与 R2 上运行三层组播

功能部属

- R1 与 R2 上运行 OSPF，实现单播路由
- R1 与 R2 上运行组播 PIM-SM，实现组播路由
- R2 上双机热备运行

【注释】

在环境中，R2 有可能会执行热备切换，此时需要调整 pim join/prune 报文发送间隔和候选 rp 通告的周期。因为目前组播需要依赖于单播，单播收敛完毕，组播才开始收敛。例如，目前单播默认 gr 收敛时间为 120s，pim join/prune 报文发送间隔也需要调整，目前默认的发送间隔为 60s，因为 pim join/prune 报文的保活时间为 3.5 倍的 pim join/prune 发送间隔。目前默认的最保活时间为 210s，参照单播收敛时间进行调整。如果 R2 配置为动态 rp，候选 rp 发送的候选 rp 通告的周期也需要调整，目前默认的发送间隔为 60s，目前保活时间为 2.5 倍的候选 rp 通告间隔，同样参照单播收敛时间设置，比如如果单播收敛超过了 150s，需要调整候选 rp 通告间隔。

需要注意，修改 pim join/prune 报文发送间隔，可能会带来这样的影响，例如 R2 发送 pim join/prune 报文丢失，R1 形成不了转发表项，需要等待 R2 重新发送 pim join/prune 报文，等待时间就是 pim join/prune 报文发送间隔。

4.3 功能详解

基本概念

▾ PIM 路由器、PIM 接口

启用了 PIM 协议的路由器称为 PIM 路由器。使能了 PIM 协议的接口称为 PIM 接口。

在 PIM 路由器上执行组播报文的转发。接收组播报文的 PIM 接口称为上游接口；发出组播报文的 PIM 接口称为下游接口。上游接口所在网段称为上游网段。下游接口所在网段称为下游网段。

▾ PIM 网络、PIM 域

PIM 路由器之间通过 PIM 接口相连，组成 PIM 网络。

在某些 PIM 接口上设置“边界”，可以将一个大的 PIM 网络划分为多个 PIM 域。“边界”可以拒绝特定组播报文通过，或者限制 PIM 消息的传输。

▾ 组播分发树、DR、RP

组播报文从一点传输到多点，转发路径呈现树型结构。此转发路径被称为组播分发树（MDT，Multicast Distribution Tree），有两种类型：

- RPT（RP Tree，共享树）：以 RP（Rendezvous Point）为根，以连接组成员的 DR 为叶子。
- SPT（Shortest Path Tree，最短路径树）：以连接组播源的 DR（Designated Router）为根，以 RP 或连接组成员的 DR 为叶子。

DR、RP 都是 PIM 路由器的功能角色。

- RP 负责收集网络中的组播源信息，组成员信息。
- 连接组播源的 DR 负责向 RP 汇报组播源的信息。连接组成员的 DR 负责向 RP 汇报组成员的信息。

▾ (*,G)、(S,G)

- (*,G)：任意源发往组 G 的报文、此类报文对应的路由表项、此类报文对应的转发路径（RPT）。
- (S,G)：源 S 发往组 G 的报文、此类报文对应的路由表项、此类报文对应的转发路径（SPT）。

▾ ASM、SSM

PIM-SM 支持两种组播服务模型——ASM 和 SSM，适用不同的组播地址段。

- ASM (Any-Source Multicast , 任意源组播) : 在 ASM 模型中, 用户主机不能选择组播源。用户主机加入组, 接收所有源发往该组的报文。
 - SSM (Source-Specific Multicast , 指定源组播) : 在 SSM 模型中, 用户主机可以选择组播源。用户主机加入组时指定源地址, 仅接收指定源发往该组的报文。
- i** SSM 模型要求: 必须通过其他方式的网络服务, 使用户主机能够预先知道组播源的位置, 用户主机才有可能对组播源进行选择。

功能特性

功能特性	作用
建立PIM邻居	PIM 路由器之间建立邻居关系, 从而形成 PIM 网络。
DR竞选	在连接组成员主机的共享网段上, PIM 邻居之间进行 DR 竞选, 获得连接组成员的 DR。 在连接组播源的共享网段上, PIM 邻居之间进行 DR 竞选, 获得连接组播源的 DR。
BSR机制	在 PIM 网络中, BSR 定期生成包括一系列候选 RP 以及相应的组地址的“自举”消息。
PR机制	在 PIM 网络中, 通过静态配置 RP 或动态 RP 竞选, 使每个 PIM 路由器获知 RP 的位置。
组播源注册信息	当网络中出现组播源时, 由连接组播源的 DR 向 RP 发送注册报文, 使 RP 获得源信息和组播报文。
构建RPT	当网络中出现组成员时, 由连接组成员的 DR 向 RP 的方向发送加入报文, 构建 RPT。如果这时网络中存在组播源, 则已送达 RP 的组播报文, 可以沿 RPT 到达组成员。
构建SPT	当数据报文到达连接组成员的 DR 时, 由连接组成员的 DR 向组播源的方向发送加入报文 构建 SPT。而后组播报文沿 SPT 转发。
ASM、SSM	PIM 路由器同时提供 ASM 模型和 SSM 模型的组播服务。对待 SSM 地址范围内的组, 使用 SSM 模型。对待其他组, 使用 ASM 模型。

4.3.1 建立PIM邻居

PIM 路由器之间建立邻居关系, 从而形成 PIM 网络。PIM 路由器之间必须首先建立邻居关系, 才能交互其他 PIM 控制消息或转发组播报文。

工作原理

PIM 接口对外发送 Hello 消息。封装 Hello 消息的组播报文的目的地址为 224.0.0.13 (表示同一网段中所有 PIM 路由器)、源地址为接口的 IP 地址、TTL 数值为 1。封装 Hello 消息的 IPv6 组播报文的目的地址为 ff02::d。

Hello 消息的作用: 发现邻居、协调各项协议参数、维持邻居关系。

📌 发现 PIM 邻居

同一网段中的 PIM 路由器都会收到目的地址为 224.0.0.13 的组播报文, 从而获得邻居信息, 建立邻居关系。

当接口刚启动或是接口检测到新邻居时, 会使用 Triggered-Hello-Delay 来产生一个随机时间, 在这个随机时间内, 接口会往外发送 hello 报文。

📌 协调各项协议参数

Hello 消息中携带多项协议参数, 介绍如下:

- DR_Priority: 表示各路由器接口竞选 DR 的优先级, 优先级越高越容易获胜。

- Holdtime：表示保持邻居为可达状态的超时时间。
- LAN_Delay：表示共享网段内传输 Prune 消息的延迟时间。
- Override-Interval：表示 Hello 消息中携带的否决剪枝时间。

当 PIM 路由器从上游接口接收到 Prune 消息时，说明该共享网段内还存在其他下游。如果这台 PIM 路由器仍然需要接收组播数据，则必须在 Override-Interval 时间内向上游发送剪枝否决消息。

$LAN_Delay + Override-Interval = PPT$ (Prune-Pending Timer)。PIM 路由器接收到下游的 Prune 消息后，并不立即执行剪枝，而是等待 PPT 时间，超时则剪枝。在 PPT 时间内如果收到下游发来的剪枝否决消息，则取消剪枝。

✎ 维持邻居关系

PIM 路由器之间周期性地发送 Hello 消息。如果 Holdtime 超时还没有收到该 PIM 邻居发出的新的 Hello 报文，则认为该邻居不可达，将其从邻居列表中清除。PIM 邻居的变化将导致网络中组播拓扑的变化。如果组播分发树上的某上游邻居或下游邻居不可达，将导致组播路由重新收敛，组播分发树迁移。

相关配置

✎ 启动接口上的 PIM-SM 功能

缺省情况下，接口上的 PIM-SM 功能关闭。

使用 `ip pim sparse-mode` 命令可以启动或关闭接口上的 PIM-SM 功能。

必须在接口上启用 PIM-SM 功能，接口才能参与 PIM 协议工作。如果作为 DR、静态 RP、C-RP 或 C-BSR 的接口未启动 PIM-SM 功能，则对应的协议角色无法工作。

✎ 调整接口上的 Hello 消息发送间隔

缺省情况下，Hello 消息发送间隔为 30S。

使用 `ip pim query-interval interval-seconds` 命令可以调整接口上的 Hello 消息发送间隔，取值范围是 <1-65535>。

Hello 消息发送间隔越大，*interval-seconds* 越大。Hello 消息发送间隔越小，*interval-seconds* 越小。

4.3.2 DR竞选

在连接组成员主机的共享网段上，PIM 邻居之间进行 DR 竞选，获得连接组成员的 DR。

在连接组播源的共享网段上，PIM 邻居之间进行 DR 竞选，获得连接组播源的 DR。

DR 负责为其直连组成员朝着组播分发树根节点方向发送“加入/剪枝”消息，或是将直连组播源的数据发向组播分发树。

工作原理

在建立 PIM 邻居过程中，通过邻居的 Hello 报文，获取邻居的 IP 地址、DR 的优先级，产生 DR。

DR 竞选的关键是：接口 DR 优先级、接口 IP 地址。

✎ 接口 DR 优先级

DR 竞选过程中，接口 DR 优先级越高，PIM 路由器竞选成功，成为 DR。

↘ 接口 IP 地址

DR 竞选过程中，接口 DR 优先级相等，就比较邻居 IP 地址，IP 地址越高，PIM 路由器竞选成功，成为 DR。

相关配置

↘ 启动接口上的 PIM-SM 功能

缺省情况下，接口上的 PIM-SM 功能关闭。

使用 **ip pim sparse-mode** 命令可以启动或关闭接口上的 PIM-SM 功能。

必须在接口上启用 PIM-SM 功能，接口才能参与 PIM 协议工作。如果作为 DR、静态 RP、C-RP 或 C-BSR 的接口未启动 PIM-SM 功能，则对应的协议角色无法工作。

↘ 调整接口上的 DR 优先级

缺省情况下，DR 优先权为 1。

使用 **ip pim dr-priority priority-value** 命令，可以调整接口 DR 优先权，优先权取值范围是<0-4294967294>。

接口 DR 优先权值用于在接口直连网段上选举 DR。值越大，越优先。

4.3.3 BSR 机制

在 PIM 网络中，BSR 定期生成包括一系列候选 RP 以及相应的组地址的“自举”消息。“自举”消息在整个域中逐跳发送。整个网络中的总有 PIM 路由器都会收到“自举”消息，并记录下候选 RP 以及相应的组地址。

工作原理

PIM-SM 域内配置了一个或多个候选自举设备(Candidate-BSR)，应用一定的规则从中选出自举设备(BSR)。

相关配置

↘ 配置候选 BSR

缺省情况下，取消候选 BSR 的配置。

使用 **ip pim bsr-candidate interface-type interface-number [hash-mask-length [priority-value]** 命令，可以或取消配置候选 BSR。

候选 BSR 通过 BSM 消息学习和竞争，产生 PIM-SM 域内全局唯一的 BSR。

↘ 配置 BSR 边界

缺省情况下，不配置 BSR 边界。

使用 **ip pim bsr-border** 命令，可以或取消配置 BSR 边界。

配置这条命令，那么此接口收到 BSM 后会立即丢弃，也不会从该接口转发 BSM，限制 BSM 的泛洪。没有配置此命令，不配置 BSR 边界。

↘ 配置对合法 BSR 范围的限定

缺省情况下，取消对 BSR 的 BSM 报文进行过滤。

使用 `ip pim accept-bsr list { <1-99> | <1300-1999> | WORD }` 命令，可以设置是否对 BSR 的 BSM 报文，进行限定。

配置这条命令，对合法 BSR 范围进行限定。不配置这条命令，那么开启了 PIM-SM 功能的设备将接收所有外部来到的 BSM 报文。

配置竞选 BSR 对合法的 C-RP 地址范围及其所服务的组播组范围进行限制

缺省情况下，取消竞选 BSR 对候选 RP 通告进行过滤。

使用 `ip pim accept-crp list { <100-199> | <2000-2699> | WORD }` 命令，可以设置是否对候选 RP 通告进行过滤。

配置这条命令，竞选 BSR 对合法的 C-RP 地址范围及其所服务的组播组范围进行限制。不配置这条命令，那么竞选 BSR 将接收所有的外部来到的候选 RP 通告报文。

配置竞选 BSR 接收 prefix-count 为 0 的 C-RP-ADV 报文

缺省情况下，竞选 BSR 不能接收 prefix-count 为 0 的 C-RP-ADV 报文。

使用 `ip pim accept-crp-with-null-group` 命令，可以设置是否要接收 prefix-count 为 0 的 C-RP-ADV 报文。

配置这条命令，竞选 BSR 能接收 prefix-count 为 0 的 C-RP-ADV 报文。不配置这条命令，那么竞选 BSR 将不处理 prefix-count 为 0 的 C-RP-ADV 报文。

4.3.4 RP 机制

在 PIM 网络中，通过静态配置 RP 或动态 RP 竞选，使每个 PIM 路由器获知 RP 的位置。RP(汇聚点)作为共享树(Shared Tree，即 RPT)树的根，RPT 树的建立、RPT 数据流的转发都必须以 RP 为转发点，进行数据建立或转发。

工作原理

在一个 PIM 域中的所有 PIM 路由器必须能够以一个特定的组播组地址映射到相同的 RP。在 PIM 网络中，RP 配置可以分为静态 RP、动态 RP。

静态 RP

静态 RP 通过各个 PIM 路由器上直接设置 RP 地址，整个 PIM 网络知道 RP 地址。

动态 RP

PIM-SM 域中还配置了候选 RP 设备(Candidate-RP)，这些候选 RP 将包含了它们的地址及可以服务的组播组等信息的数据报传单播至 BSR 设备。BSR 定期生成包括一系列候选 RP 以及相应的组地址的“自举”消息。“自举”消息在整个域中逐跳发送。设备接收并保存这些“自举”消息。接收端 DR 将使用一个 hash 算法将组地址映射至一个可以为该组服务的候选 RP，那么组地址所对应的 RP 就可以确定。

相关配置

设置静态 RP 地址

缺省情况下，没有配置 RP 地址。

使用 `ip pim rp-address rp-address [access-list]` 命令，可以设置 PIM 路由器的静态 RP 地址。

配置静态 RP 来使用 PIM-SM ,要求 PIM-SM 域内的所有设备静态 RP 配置必须一致 ,保证 PIM-SM 的组播路由没有歧义性。

📌 设置候选 C-RP 地址

缺省情况下 , 没有配置 C-RP 地址。

使用 `ip pim rp-candidate interface-type interface-number [priority priority-value] [interval interval-seconds] [group-list access-list]`命令 , 可以设置或取消 PIM 路由器为候选 C-RP 地址。

配置候选 RP 可以周期性地向 BSR 发送候选 RP 通告 , 这些候选 RP 通告所带的信息会扩散到域内所有 PIM-SM 设备 , 保证 RP 映射的唯一。

📌 忽略 RP-SET 中的 RP 的优先级

缺省情况下 , 优先选择优先级较高的 C-RP。

使用 `ip pim ignore-rp-set-priority` 命令 , 在选择组对应的 RP 时 , 可以选择或忽略 RP 的优先级。

当要对一个组播地址选择一个 RP 时 , 如果有多个 RP 都可以服务这个组播地址 , 此时在比较两个 RP 时要忽略掉 RP 的优先级 , 可以使用这条命令。如果不配置这条命令 , 那么在比较两个 RP 时 , RP 的优先级会被考虑。

4.3.5 组播源注册信息

当网络中出现组播源时 , 由连接组播源的 DR 向 RP 发送注册报文 , 使 RP 获得源信息和组播报文。

工作原理

数据源端 DR 从直连主机收到组播数据包 , 然后数据源端 DR 将组播数据封装在注册消息中单播到 RP , 生成(S, G)表项。

如果 RP 有转发表项出口 , 封装数据报文 , 转发到出口。

如果 RP 没有当前组的转发表项 , 生成(S, G)表项启动定时器 , 定时器超时 , 向 DR 发送注册停止 , 删除表项。数据源端 DR 收到注册停止后 , 定时超时以前 , 发送探测报文。

如果没有收到注册停止 , 超时数据源端 DR 将组播数据封装在注册消息中单播到 RP。

如果收到注册停止 , 重新延时 , 延时以前重新发送探测报文。

相关配置

📌 RP 注册报文可达性检测

缺省情况下 , 没有对 RP 可达性 , 进行检验。

使用 `ip pim register-rp-reachability` 命令 , 设置或取消对 RP 可达性进行检验。

从 DR 上要往 RP 发送注册报文 , 如果要检验 RP 是否可达 , 可以配置此命令。配置此命令之后 , DR 发送注册报文之前将检验 RP 是否可达 , 即查询单播路由和静态组播路由表获知是否存在到达 RP 的路由 , 如果不存在到达 RP 的路由则不发送注册报文。

📌 配置 RP 对注册报文的地址过滤

缺省情况下 , RP 允许每一个到达的注册报文。

使用 **ip pim accept-register { list access-list [route-map map-name] | route-map map-name [list access-list]}** 命令，设置或取消 RP 对到达注册报文的源地址进行过滤。

在 RP 上，要对到达注册报文的源地址进行过滤，可以配置此命令。如果不配置这条命令，那么在 RP 上允许每一个到达的注册报文。如果配了这条命令，那么注册报文中源地址和组地址被 ACL 允许的才被继续处理，否则注册报文被过滤掉。

配置注册报文发送的速度限制

缺省情况下，不限速。

使用 **ip pim register-rate-limit rate** 命令，是否对注册报文进行限速。

配置 DR 发送注册报文的速率，如果使用 no 形式，表示不限制速率。它是对每个 (S, G) 状态的注册报文发送速率进行配置，并不是对整个系统的注册报文发送速率进行配置。

配置注册报文的校验和的计算方式为计算整个注册报文

缺省情况下，注册报文中校验和的计算方式为协议规定的默认方式。

使用 **ip pim register-checksum-wholepkt [group-list access-list]** 命令，可以对计算校验和长度进行设置。

对注册报文进行校验和计算时，若要计算整个 pim 协议报文含其中封装的组播数据报文，请使用这条命令。不配置这条命令，那么注册报文中校验和的计算方式为协议规定的默认方式。

配置 RP 对 register 报文解封装之后，能够向下游口转发其中的组播数据报文

缺省情况下，取消对注册报文进行解封装并转发其中的组播数据报文。

使用 **ip pim register-decapsulate-forward** 命令，可以设置是否向下游口，转发数据报文。

若要对注册报文进行解封装并转发其中的组播数据报文，请配置这条命令。不配置这条命令，那么就不会对注册报文中的组播数据报文进行转发。

配置注册报文的源地址

缺省情况下，注册报文的源地址将会使用与组播源相连的 DR 接口地址。

使用 **ip pim register-source { local_address | Interface-type interface-number }** 命令，可以设置源地址。

如果要配置 DR 发送出来的注册报文的源地址，可以使用此命令。不配置此命令或是使用这条命令的 no 形式，注册报文的源地址将会使用与组播源相连的 DR 接口地址。如果使用该命令的地址参数，配置的地址必须是单播路由可达的。如果使用该命令的接口参数，一般来说它是回环接口，但也可以是其它类型接口，这个接口的地址必须是被单播路由公告过的。

配置注册报文抑制时间

缺省情况下，注册报文的抑制时间为 60S。

使用 **ip pim register-suppression seconds** 命令，可以设置抑制时间。

使用此命令配置注册报文抑制时间，在 DR 上配置此值将修改定义在 DR 上的注册报文抑制时间。如果没有配置 **ip pim rp-register-kat** 命令，在 RP 上定义此值将修改 RP keepalive 的周期。

配置 NULL 注册报文的探测时间

缺省情况下，探测时间为 5S。

使用 **ip pim probe-interval interval-seconds** 命令，可以设置探测时间。

在注册抑制时间超时前的一段时间间隔，源 DR 可以向 RP 发送空注册报文即 NULL-Register 报文。这段时间间隔就是探测时间，默认的探测时间是 5S。

配置 RP KAT 定时器的时间值

缺省情况下，使用 KAT 的缺省值，KAT 缺省值为注册抑制时间 * 3 + 注册探测时间。

使用 `ip pim rp-register-kat seconds` 命令，可以设置控制时间。

如果要配置 RP 上注册报文所创建(S,G)状态的存活时间，可以使用此命令。

4.3.6 构建RPT

当网络中出现组成员时，由连接组成员的 DR 向 RP 的方向发送加入报文，构建 RPT。如果这时网络中存在组播源，则已送达 RP 的组播报文，可以沿 RPT 到达组成员。

工作原理

构建 RPT，需要如下过程：

- 当接收端的 DR 收到一个来自接收端的 IGMP (*,G)Include report 报文；
- 如果接收端的 DR 不是这个组 G 的 RP，接收端的 DR 就会向 RP 方向发送一个(*,G)join 报文，收到这个(*,G)join 报文的的上游设备，又会向 RP 方向发送(*,G)join 报文，这样(*,G)join 报文逐跳发送，直到组 G 的 RP 收到了(*,G)join 报文，表示加入共享树；
- 数据源主机向组发送组播数据时，源数据被封装在注册消息内，并由数据源端的 DR 单播至 RP，RP 再将注册报文明解封，取出其中的数据报文，然后沿着共享树转发到各个组成员；
- RP 朝着数据源端的 DR 发送 (S, G) join 报文，用以加入此源的最短路径树；
- 这样当 RP 到数据源端的 DR 的 SPT 建好后，数据源的数据包将沿着其最短路径树不加封装地发送到 RP；
- 当第一个组播数据沿 SPT 到达时，RP 向源的 DR 发送注册-停止消息，以使 DR 停止注册封装过程。当数据源端的 DR 收到注册-停止消息后，不再封装注册数据报文，而是只沿着源的最短路径树发送到 RP，再由 RP 将其沿着共享树转发各个组成员。

相关配置

配置 join/prune 报文的发送间隔

缺省情况下，将 join/prune 报文发送间隔还原为默认值，即 60S

使用 `ip pim jp-timer seconds` 命令，设置 join/prune 报文发送间隔。

join/prune 报文的默认发送间隔是 60S，如果要修改这个发送间隔，可以配置这条命令。不配置这条命令，则 join/prune 的默认发送间隔为 60S。

4.3.7 构建SPT

当数据报文到达连接组成员的 DR 时，由连接组成员的 DR 向组播源的方向发送加入报文，构建 SPT。而后组播报文沿 SPT 转发。减少 RPT 中的 RP 压力，数据源端的 DR 到达接收端的跳数减少。

工作原理

形成 SPT 树，需要如下过程：

- 接收端的 DR 沿着最短路径树，向源端的 DR 发送(*,G)join 报文。这样(*,G)join 报文逐跳发送，直到源端的 DR 收到(*,G)join 报文，SPT 树形成；

相关配置

缺省情况，不启动 SPT 切换。

使用 `ip pim spt-threshold [group-list access-list]` 命令，设置是否启动 SPT 切换。

当配置了这条命令，接收到第一个(S,G)报文，一个 PIM 的加入信息被触发，并构造一棵源树；如果 group-list 关键字被定义，所有的这个特定组会切换到源树；使用这条命令的 no 形式并且不携带 group-list 关键字，关闭从共享树切换到最短路径树功能，该设备将重新转向共享树并向源发送出一个剪枝报文；如果使用这条命令的 no 形式携带 group-list 关键字，并且携带的 ACL 列表为之前已经配置的 ACL 列表，则表示取消 group-list 关联的 ACL 列表限制，允许所有的组从共享树切换到最短路径树。

4.3.8 ASM、SSM

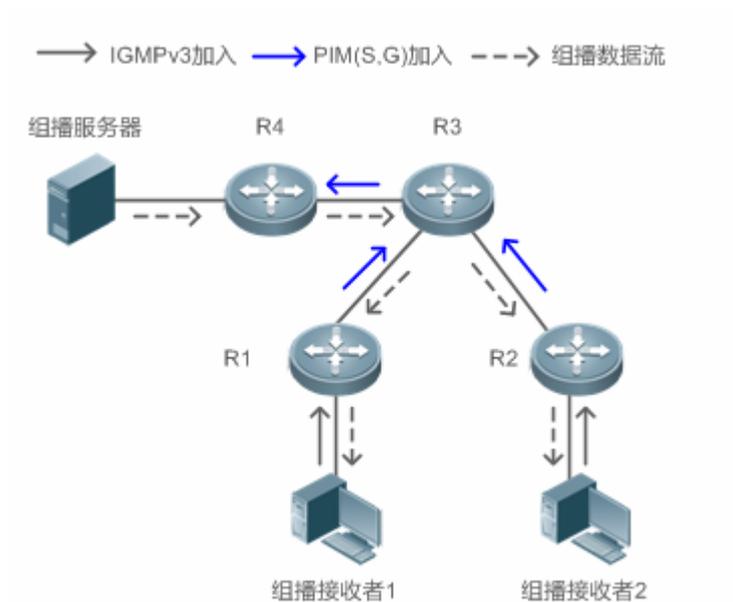
PIM 路由器同时提供 ASM 模型和 SSM 模型的组播服务。对待 SSM 地址范围内的组，使用 SSM 模型。对待其他组，使用 ASM 模型。在 ASM 模型中，组播数据接收者只指定加入某个组播组 G，而没有指定组播源 S。在 SSM 模型中，组播数据接收者可以同时指定组播源 S 和组播组 G。

工作原理

 在 IPv4 上实现 SSM 模型，需要借助 IGMPv3 来管理主机和设备的成员关系，借助 PIM-SM 来连接设备。

在 SSM 模型中，组播接收者事先已经通过某种渠道(比如：访问服务器、接受广告等)了解到组播源的信息(S,G)；于是，当组播接收者要点播某个组播服务时，组播接收者可以直接往最后一跳设备发送 IGMP(S,G)加入，如图 2 中的组播接收者 1 发送 IGMP(S,G)加入报告以点播组播服务(S,G)；最后一跳设备接收到来自组播接收者的 IGMP(S,G)加入后，往组播源逐跳发送 PIM(S,G)加入，如图 2 的 R1 接收到来自组播接收者 1 的 IGMP(S,G)加入报告后，向 R3 发送 PIM(S,G)加入，接着，R3 向 R4 发送 PIM(S,G)加入；于是，就建立一棵从组播接收者到组播源的最短路径树。如下图所示。

图 4-4 SSM 模型



实现 SSM 模型需要满足：

- 组播接收者通过某种渠道，事先了解到组播源(S,G)的信息；组播接收者对感兴趣的组播服务，发起 IGMP(S,G)加入。
- 最后一跳设备连接组播接收者的接口上，必须启动 IGMPv3。IGMPv1/IGMPv2 不支持 SSM。
- 在组播接收者到组播源的途径设备上，必须启动 PIM-SM、SSM。

! 开启 SSM 功能后，SSM 的缺省组范围是：232/8。可以通过命令修改 SSM 的组范围。

SSM 具有如下的特点：

- 在 SSM 模型中，组播接收者能够通过某些渠道(比如：广告、访问指定服务器等)，事先获取到组播源的信息。
- SSM 模型是 PIM-SM 的一个特定子集，只处理 PIM(S,G)加入和 PIM(S,G)剪枝消息；丢弃落在 SSM 范围的与 RPT 相关的消息，比如：PIM(*,G)加入/剪枝消息；对于落在 SSM 范围内的注册报文，会立即应答注册停止报文。
- 在 SSM 模型中，不需要 RP，不需要进行 RP 信息的选举和分发；在 SSM 中，构建的组播分发树都是最短路径树即 SPT。

相关配置

缺省情况，使用 ASM 模式。

使用 `ip pim ssm { default | range access-list }`命令，设置是否启动 SSM。

SSM 模式下，可以直接从组播源接收组播数据报文，而不需要走 RP 树。

4.4 配置详解

配置项	配置建议 & 相关命令
配置PIM-SM基本功能	! 必须配置。用于建立组播服务。

	ip multicast-routing	启动 IPv4 组播路由功能
	ip pim sparse-mode	启动 PIM-SM 功能
	ip pim rp-address	配置静态 RP
	ip pim rp-candidate	配置 C-RP
	ip pim bsr-candidate	配置 C-BSR
	ip pim ssm	启动 SSM
配置PIM邻居参数	 可选配置。用于邻居间 Hello 报文的发送与接收。	
	ip pim query-interval <i>interval-seconds</i>	配置 Hello 消息发送间隔
	ip pim propagation-delay <i>milliseconds</i>	配置 Hello 消息中的传输延迟时间
	ip pim override-interval <i>milliseconds</i>	配置 Hello 消息中的剪枝否决时间
	ip pim neighbor-tracking	配置 Hello 消息中的接口的加入抑制能力
	ip pim triggered-hello-delay <i>interval-seconds</i>	配置 Hello 消息的发送报文延时
	ip pim dr-priority <i>priority-value</i>	配置 Hello 消息中的 DR 优先权
	ip pim neighbor-filter <i>access_list</i>	配置邻居过滤
配置BSR参数	 可选配置。用于 BSR 配置。	
	ip pim bsr-border	配置 BSR 边界
	ip pim accept-bsr list { <1-99> <1300-1999> <i>WORD</i> }	配置 PIM 路由器对 BSM 报文限定
	ip pim accept-crp list <i>access-list</i>	配置 C-BSR 可对 C-PR 范围限定检验
配置RP、DR参数	 可选配置。用于 RP 与 DR 使用。	
	ip pim ignore-rp-set-priority	忽略 C-RP 优先级
	ip pim register-rp-reachability	源直连的 DR 对 RP 可达性检测
	ip pim accept-register list <i>access-list</i>	限制源注册的(S,G)地址范围
	ip pim register-rate-limit <i>rate</i>	配置 DR 发送注册报文速度限制的命令
	ip pim register-checksum-wholepkt [group-list <i>access-list</i>]	配置注册报文的校验和必须计算整个报文
	ip pim register-decapsulate-forward	允许 RP 对注册报文进行解封并转发其中的组播数据报文
	ip pim register-source { <i>local_address</i> <i>Interface-type interface-number</i> }	配置注册报文的源地址
	ip pim register-suppression <i>seconds</i>	设置注册报文抑制时间
	ip pim probe-interval <i>seconds</i>	设置注册探测时间
	ip pim rp-register-kat <i>seconds</i>	配置 RP 上的 KAT 间隔
配置join/prune报文的发送间隔	 可选配置。用于 join/prune 报文间隔。	
	ip pim jp-timer <i>seconds</i>	join/prune 报文的发送间隔
最后一跳设备由从共享树切	 可选配置。用于 spt 向 rpt 切换。	

换到最短路径树	<code>ip pim spt-threshold [group-list access-list]</code>	启动 SPT 切换功能
-------------------------	--	-------------

4.4.1 配置PIM-SM基本功能

配置效果

- 建立 PIM 网络，向网络内的数据源和用户终端提供 IPv4 组播服务。
- 两种模型的组播服务（ASM、SSM），可以仅支持其一，也可以两者都支持。

注意事项

- PIM-SM 需要借助网络中现有的单播路由。因此，网络中必须配置 IPv4 单播路由。
- 如果要求 PIM 网络支持 SSM 模型的组播服务，则还需要配置 IGMPv3 或 SSM Mapping。

配置方法

启动 IPv4 组播路由功能

- 必须配置。
- 若无特殊要求，应在每台路由器上启动 IPv4 组播路由功能。

启动 PIM-SM 功能

- 必须配置。
- 若无特殊要求，应在以下接口上启动 PIM-SM 功能：路由器互联接口，作为静态 RP、C-RP 和 C-BSR 的接口，连接组播源、用户主机的接口。

配置 RP

- 如果要求 PIM 网络支持 ASM 模型的组播服务，则必须配置 RP。
- 配置 RP 的三种方法：仅配置静态 RP、仅配置动态 RP、同时配置静态 RP 和动态 RP。如果同时配置了静态 RP 和动态 RP，则动态 RP 优先。
- 配置静态 RP：若无特殊要求，应在每台路由器上配置静态 RP。
- 配置动态 RP：若无特殊要求，应在一台或多台路由器上配置 C-RP 和 C-BSR。

启动 SSM

- 如果要求 PIM 网络支持 SSM 模型的组播服务，则必须启动 SSM。
- 若无特殊要求，应在每台路由器上启动 SSM。

检验方法

使网络中的组播源分别向 ASM、SSM 范围内的组发送报文，用户主机加入这些组。

- 检查用户主机能否成功接收每个组的报文。
- 检查路由器上是否建立了正确的 PIM-SM 路由表项。

相关命令

启动 IPv4 组播路由功能

- 【命令格式】 **ip multicast-routing**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

启动 PIM-SM 功能

- 【命令格式】 **ip pim sparse-mode**
- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 PIM 接口必须是三层接口，包括：路由口、L3AP、SVI、loopback 接口。
所有 PIM 接口应 IPv4 单播路由可达。

配置静态 RP

- 【命令格式】 **ip pim rp-address** *rp-address* [*access_list*]
- 【参数说明】 *rp-address*：静态 RP 的地址。
access_list：使用 ACL 限定该静态 RP 服务的组地址范围。缺省为所有组服务。
- 【命令模式】 全局模式
- 【使用指导】 此命令用来指出静态 RP 在哪里。
应选择性能较强的路由器作为静态 RP。建议以 loopback 接口地址作为静态 RP 地址。建议以 loopback 接口地址作为静态 RP 地址。
所有路由器上配置的静态 RP（包括 RP 地址及其服务的组地址范围）应完全相同。建议以 loopback 接口地址作为静态 RP 地址。
如果配置多个静态 RP，对应不同的组地址范围，可以实现负载分担。建议以 loopback 接口地址作为静态 RP 地址。

配置 C-RP

- 【命令格式】 **ip pim rp-candidate** *interface-type interface-number* [**priority** *priority-value*] [**interval** *seconds*] [**group-list** *access_list*]
- 【参数说明】 *interface-type interface-number*：以此接口的地址作为 C-RP 的地址。
priority *priority-value*：竞选 RP 的优先级。取值越大，优先级越高。取值范围 0-255，缺省值 192。
interval *seconds*：向 BSR 发送 C-RP 消息的周期。单位秒，取值范围 1-16383，缺省值 60。
group-list *access_list*：引用 ACL 限定该 C-RP 服务的组地址范围。缺省为所有组服务。
- 【命令模式】 全局模式
- 【使用指导】 此命令用来使路由器成为 C-RP。
应选择性能较强的路由器作为 C-RP。C-RP 与 C-BSR 可以位于同一台路由器，也可以位于不同的路由器。建议以 loopback 接口地址作为 C-RP 地址。
如果配置多个 C-RP 为相同的组服务，可以实现冗余备份。

如果配置多个 C-RP 为不同的组服务，可以实现负载分担。

配置 C-BSR

【命令格式】 **ip pim bsr-candidate** *interface-type interface-number* [*hash-mask-length* [*priority-value*]]

【参数说明】 *interface-type interface-number*：以此接口的地址作为 C-BSR 的地址。

hash-mask-length：竞选 RP 时使用的 HASH 掩码长度。取值范围 0-32，缺省值 10。

priority-value：竞选 BSR 的优先级。取值越大，优先级越高。取值范围 0-255，缺省值 64。

【命令模式】 全局模式

【使用指导】 此命令用来使路由器成为 C-BSR。

应选择性能较强的路由器作为 C-BSR。C-RP 与 C-BSR 可以位于同一台路由器，也可以位于不同的路由器。

建议以 loopback 接口地址作为 C-BSR 地址。

如果配置多个 C-BSR，可以实现冗余备份。

启动 SSM

【命令格式】 **ip pim ssm** { **default** / **range** *access_list* }

【参数说明】 **default**：缺省的 SSM 组地址范围，即 232.0.0.0/8。

range *access_list*：引用 ACL 限定 SSM 组地址范围。

【命令模式】 全局模式

【使用指导】 所有路由器上配置的 SSM 组地址范围应完全相同。

查看 PIM-SM 路由表

【命令格式】 **show ip pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]] [**proxy**]

【参数说明】 *group-or-source-address*：组地址或源地址（两个地址不能同时都是组地址或者同时都是源地址）

proxy：显示表项携带的 RPF Vector 信息

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 查看是否具备足够的路由表项。

查看路由表项的上游接口、下游接口列表，确认构建了正确的生成树。

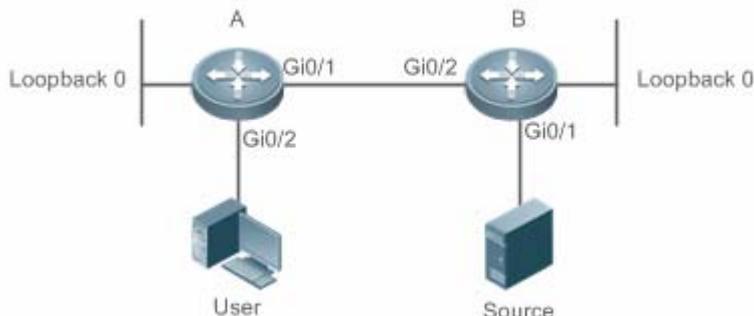
配置举例

i 以下配置举例，仅介绍与 PIM-SM 相关的配置。

在 IPv4 网络上建立 IP 组播服务，支持 ASM 和 SSM

【网络环境】

图 4-5



【配置方法】

- 在路由器上配置 IPv4 单播路由协议（如 OSPF），且 loopback 接口单播路由可达。略
- 在所有路由器上启动 IPv4 组播路由功能。
- 在设备互联接口上、连接用户主机和组播源的接口上启动 PIM-SM 功能。
- 在 A、B 的 loopback 接口上配置 C-RP 和 C-BSR。该 loopback 接口也需要启动 PIM-SM 功能。
- 在所有路由器上启动 SSM。
- 在路由器连接用户终端的接口上启动 IGMPv3。略

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# ip pim ssm default
A(config)# interface GigabitEthernet 0/1
A(config-if)# ip pim sparse-mode
A(config-if)# exit
A(config)# interface GigabitEthernet 0/2
A(config-if)# ip pim sparse-mode
A(config-if)# exit
A(config)# interface loopback 0
A(config-if)# ip pim sparse-mode
A(config-if)# exit
A(config)# ip pim rp-candidate loopback 0
```

B

```
B# configure terminal
B(config)# ip multicast-routing
B(config)# ip pim ssm default
B(config)# interface GigabitEthernet 0/1
B(config-if)# ip pim sparse-mode
B(config-if)# exit
B(config)# interface GigabitEthernet 0/2
B(config-if)# ip pim sparse-mode
B(config-if)# exit
B(config)# interface loopback 0
B(config-if)# ip pim sparse-mode
B(config-if)# exit
B(config)# ip pim bsr-candidate loopback 0
```

【检验方法】

使 Source (192.168.1.10) 向 G1 (229.1.1.1)、G2 (232.1.1.1) 发送报文。使 User 加入 G1、G2，加入 G2 时指定源地址 Source。

- 检查 User 收到的组播报文。User 应该能够收到 G1、G2 的组播报文。
- 检查 A、B 上的 PIM-SM 路由表。应该存在表项 (* , 229.1.1.1)、(192.168.1.10 , 229.1.1.1)、(192.168.1.10 , 232.1.1.1)。

A

```
switch#show ip pim sparse-mode mroute
IP Multicast Routing Table
```

```

(*,*,RP) Entries: 0
(*,G) Entries: 3
(S,G) Entries: 2
(S,G,rpt) Entries: 2
FCR Entries: 0
REG Entries: 0

(*, 229.1.1.1)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . .
i . . . . .
1 . . . . .
.
Joined
0 . . . . .
.
1 . . . . .
.
Asserted
0 . . . . .
.
1 . . . . .
.
FCR:

(192.168.1.10, 229.1.1.1)
RPF nbr: 192.168.2.1
RPF idx: GigabitEthernet 0/2
SPT bit: 1
Upstream State: JOINED
jt_timer expires in 8 seconds
kat expires in 207 seconds
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
1 . . . . .
.
Joined

```

```

0 . . . . .
.
1 . . . . .
.
Asserted
0 . . . . .
.
1 . . . . .
.
Outgoing
0 . . .
o . . . . .
1 . . . . .
.

(192.168.1.10, 229.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: PRUNED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
1 . . . . .
.
Pruned
0 . . . . .
.
1 . . . . .
.
Outgoing
0 . . .
o . . . . .
1 . . . . .
.

(*, 232.1.1.1)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local

```

```

0 . . .
i . . . . .
1 . . . . .
.
Joined
0 . . . . .
.
1 . . . . .
.
Asserted
0 . . . . .
.
1 . . . . .
.
FCR:
(192.168.1.10, 232.1.1.1)
RPF nbr: 192.168.2.1
RPF idx: GigabitEthernet 0/2
SPT bit: 1
Upstream State: JOINED
jt_timer expires in 8 seconds
kat expires in 207 seconds
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
1 . . . . .
.
Joined
0 . . . . .
.
1 . . . . .
.
Asserted
0 . . . . .
.
1 . . . . .
.
Outgoing
0 . . .
o . . . . .
1 . . . . .
.

```

```

(192.168.1.10, 232.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: PRUNED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
1 . . . . .
.
Pruned
0 . . . . .
.
1 . . . . .
.
Outgoing
0 . . .
o . . . . .
1 . . . . .
.
(*, 239.255.255.250)
RP: 192.168.10.10
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . .
i . . . . .
1 . . . . .
.
Joined
0 .
j . . . . .
1 . . . . .
.
Asserted
0 . . . . .
.
1 . . . . .
.
FCR:

```

B

```

A#
B#show ip pim sparse-mode mroute
IP Multicast Routing Table

(*,*,RP) Entries: 0
(*,G) Entries: 1
(S,G) Entries: 1
(S,G,rpt) Entries: 1
FCR Entries: 0
REG Entries: 1

(192.168.1.10, 229.1.1.1)
RPF nbr: 0.0.0.0
RPF idx: None
SPT bit: 1
Upstream State: JOINED
kat expires in 38 seconds
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
Joined
0 . .
j . . . . .
Asserted
0 . . . . .
.
Outgoing
0 . .
o . . . . .

(192.168.1.10, 229.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 192.168.2.2
RPF idx: GigabitEthernet 0/2
Upstream State: RPT NOT JOINED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
Pruned
0 . . . . .
.
Outgoing
    
```

```
0 . . . . .
.

(192.168.1.10, 232.1.1.1)
RPF nbr: 0.0.0.0
RPF idx: None
SPT bit: 1
Upstream State: JOINED
kat expires in 38 seconds
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
Joined
0 . .
j . . . . .
Asserted
0 . . . . .
.
Outgoing
0 . .
o . . . . .

(192.168.1.10, 232.1.1.1, rpt)
RP: 192.168.10.10
RPF nbr: 192.168.2.2
RPF idx: GigabitEthernet 0/2
Upstream State: RPT NOT JOINED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 . . . . .
.
Pruned
0 . . . . .
.
Outgoing
0 . . . . .
.
(*, 239.255.255.250)
RP: 192.168.10.10
RPF nbr: 192.168.2.2
RPF idx: GigabitEthernet 0/2
Upstream State: JOINED
jt_timer expires in 15 seconds
```

```

00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Local
0 .
i . . . . .
Joined
0 . . . . .
.
Asserted
0 . . . . .
.
FCR:

```

常见错误

- IPv4 单播路由配置错误。
- 某台路由器没有使能 IPv4 组播路由。
- 某台路由器没有使能 SSM，或 SSM 组地址范围与其他路由器不一致。
- 某个接口没有使能 PIM-SM（比如：被指定为 C-RP 或 C-BSR 的接口、作为用户主机或组播源的网关的接口）。
- 某个连接用户主机的接口没有使能 IGMPv3。
- 网络中没有配置 RP。
- 某台路由器上没有配置静态 RP，或配置的静态 RP 与其他路由器不一致。
- 网络中配置了 C-RP，没有配置 C-BSR。
- 静态 RP、C-RP 或 C-BSR 单播路由不可达。

4.4.2 配置PIM邻居参数

配置效果

- 协调各项协议参数，调整 Hello 报文参数。
- 各个 PIM 路由器，发现邻居、协调各项协议参数、维持邻居关系。
- 保护邻居之间关系，可对邻居进行限制。

注意事项

- 必须配置 PIM-SM 基本功能。

配置方法

- 若无特殊要求，可以在各个 PIM 路由器接口上设置参数。

检验方法

设置接口的 Hello 报文参数，通过 **debug ip pim sparse-mode packet** 命令，看 Hello 报文的参数。

设置邻居过滤，可以采用 **show ip pim sparse-mode neighbor** 命令，查看邻居之间关系。

相关命令

配置 Hello 消息发送间隔

【命令格式】 **ip pim query-interval interval-seconds**

【参数说明】 Hello 报文发送周期
单位为 S，取值范围 1-65535，缺省值 30。

【命令模式】 接口模式

【使用指导】 每当 Hello 消息发送间隔被更新时，Hello 消息保持时间也会按照以下规则随之更新：Hello 消息保持时间被更新为 Hello 消息发送间隔的 3.5 倍；如果 Hello 消息发送间隔 * 3.5 > 65535，并且 Hello 消息发送间隔被强制更新为 18725。

配置 Hello 消息中的传输延迟时间

【命令格式】 **ip pim propagation-delay milliseconds**

【参数说明】 *milliseconds*：单位为 MS，取值范围 1-32767，缺省值 500

【命令模式】 接口模式

【使用指导】 修改传播延迟或是修改剪枝否决延迟，都会影响到 J/P- override-interval。
协议上要求 J/P- override-interval 必须小于 Join-Prune 报文的 HoldTime，否则将有可能造成短暂的断流。
这一点，需由网络管理者维护保证。

配置 Hello 消息中的剪枝否决时间

【命令格式】 **ip pim override-interval milliseconds**

【参数说明】 *milliseconds*：单位为 MS，取值范围 1-65535，缺省值 2500

【命令模式】 接口模式

【使用指导】 修改传播延迟或是修改剪枝否决延迟，都会影响到 J/P-override-interval。
协议上要求 J/P- override-interval 必须小于 Join-Prune 报文的 HoldTime，否则将有可能造成短暂的断流。
这一点，需由网络管理者维护保证。

配置 Hello 消息中的接口的加入抑制能力

【命令格式】 **ip pim neighbor-tracking**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 接口的加入抑制能力被启动时，当本路由器想往上游邻居发送 Join 报文时，如果收到邻居往上游路由器发送的 Join 报文，那么，本路由器的 Join 报文会被抑制即不会发出；但是如果接口的加入抑制能力被禁止，那么此时本路由器可以发送 Join 报文。当下游接收者的加入抑制能力都被禁止时，上游邻居就可以通过接收到的 Join 报文，进而确切知道下游邻居有几个接收者，即邻居跟踪。

配置 Hello 消息的发送报文延时

【命令格式】 **ip pim triggered-hello-delay interval-seconds**

- 【参数说明】 *Seconds*：单位为 S，取值范围 1-5，缺省值 5
- 【命令模式】 接口模式
- 【使用指导】 当接口刚启动时或是接口检测到新邻居时，此时会使用 *triggered-hello-delay* 来产生一个随机时间，在这个随机时间内，接口会往外发送 Hello 报文。

配置 Hello 消息中的 DR 优先权

- 【命令格式】 **ip pim dr-priority *priority-value***
- 【参数说明】 *priority-value*：优先级，值越高，优先级越大。
取值范围 0-4294967294，缺省值 1。
- 【命令模式】 接口模式
- 【使用指导】 当选择一个 DR 时，下面几种情况将被应用：
同一局域网中的设备的 Hello 报文都有设置优先级参数，那么在选择 DR 时，先比较优先级，谁的优先级高，谁就成为 DR；若有多个设备拥有相同的优先级，那么，谁的 IP 地址高，谁就是 DR。
当同一局域网中的某个设备在 Hello 消息报文中没有设置优先级参数时，那么这个局域网中 DR 的选举中，谁 IP 地址高，谁就是 DR。

配置邻居过滤

- 【命令格式】 **ip pim neighbor-filter *access_list***
- 【参数说明】 *access_list*：使用 IP 标准 ACL 限制邻居地址范围，取值范围 1-99、字符串
- 【命令模式】 接口模式
- 【使用指导】 启用邻居过滤功能，可以加强 PIM 网络的安全性，限制邻居的合法地址范围。只要某邻居被过滤访问列表拒绝，则 PIM-SM 将不会与该邻居建立对等（Peering）关系，或者中止与该邻居已经建立的对等关系。

显示接口的邻居信息

- 【命令格式】 **show ip pim sparse-mode neighbor [detail]**
- 【参数说明】 **detail**：显示详细信息
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 -

配置举例

- 【配置方法】
- 配置 PIM-SM 基本功能。略
 - 配置 Hello 报文的发送间隔 50S
 - 配置 Hello 报文的传输延时 400MS
 - 配置 Hello 报文的裁剪否决时间 3000MS
 - 配置 Hello 消息中的接口加入抑制能力
 - 配置 Hello 消息的发送报文延时 3S
 - 配置 Hello 消息的 DR 的优先权 5

```
Ruijie# configure terminal
Ruijie (config)#int gi 0/1
Ruijie (config-if-GigabitEthernet 0/1)#ip pim query-interval 50
Ruijie (config-if-GigabitEthernet 0/1)#ip pim propagation-delay 400
Ruijie (config-if-GigabitEthernet 0/1)#ip pim override-interval 3000
Ruijie (config-if-GigabitEthernet 0/1)#ip pim triggered-hello-delay 3
Ruijie (config-if-GigabitEthernet 0/1)#ip pim neighbor-tracking
```

【检验方法】 通过 **debug ip pim sparse-mode packet** 命令，查看 Hello 报文参数。

```
Ruijie# debug ip pim sparse-mode packet
00:01:49:43: %7: VRF(0): Hello send to GigabitEthernet 0/1
00:01:49:43: %7: Send Hello message
00:01:49:43: %7: Holdtime: 175
00:01:49:43: %7: T-bit: on
00:01:49:43: %7: Propagation delay: 400
00:01:49:43: %7: Override interval: 3000
00:01:49:43: %7: DR priority: 5
00:01:49:43: %7: Gen ID: 355154648
00:01:49:43: %7: RPF Vector capable
```

【配置方法】 配置接口对邻居过滤，过滤地址（192.168.1.0~192.168.1.255）

```
Ruijie# configure terminal
Ruijie (config)#int gi 0/1
Ruijie (config-if-GigabitEthernet 0/1)# ip pim neighbor-filter 1
% access-list 1 not exist
Ruijie(config)# access-list 1 permit 192.168.1.0 0.0.0.255
Ruijie(config)#
```

【检验方法】 在配置邻居过滤以前，显示邻居信息。

```
Ruijie# show ip pim sparse-mode neighbor
Neighbor          Interface          Uptime/Expires    Ver  DR
Address                                     Priority/Mode
192.168.36.89     GigabitEthernet 0/1  01:12:13/00:01:32 v2   1 / P
```

在配置邻居过滤以后，显示邻居信息。

```
Ruijie# show ip pim sparse-mode neighbor
```

常见配置错误

- PIM-SM 基本功能没有配置，或配置失败。

4.4.3 配置BSR参数

配置效果

- 为了 BSM 报文的范围进行限制。

注意事项

- 必须配置 PIM-SM 基本功能。
- 必须配置 C-RP、C-BSR 路由器。
- 必须在域之间的接口进行边界配置。

配置方法

▾ 边界配置

- 如果存在多个域，则必须配置。
- 在两个域接口之间，进行配置。

▾ 配置 PIM 路由器对 BSM 报文限定

- 可选配置
- 若无特殊要求，可以在所有 PIM 路由器上进行此项配置。

▾ 配置 C-BSR 可对 C-PR 范围限定

- 可选配置
- 若无特殊要求，可以在所有 C-BSR 路由器上进行此项配置。

▾ 配置 C-BSR 接收 prefix-count 为 0 的 C-RP-ADV 报文

- 可选配置
- 若无特殊要求，可以在所有 C-BSR 路由器上进行此项配置。

检验方法

▾ 边界检验

启动 PIM-SM 基本功能，设置两个路由器为不同的域，路由器 B 设置 C-BSR，路由器 A 可以正常接收 BSM 报文，在路由器 A 与路由器 B 的交界处，设置为边界口，路由器 A 收不到 BSM 报文。

▾ 配置 PIM 路由器对 BSM 报文限定检验

启动 PIM-SM 基本功能，路由器 B 设置 C-BSR，路由器 A 可以正常接收 BSM 报文。路由器 A 对 C-BSR 范围进行限定，路由器 A 收不到 BSM 报文。

▾ 配置 C-BSR 可对 C-PR 范围限定检验

启动 PIM-SM 基本功能，路由器 B 设置 C-BSR，路由器 A 设置 C-RP，C-BSR 对 C-RP 范围限定，路由器 B 收不到 C-RP 发送过来的报文。

相关命令

配置 BSR 边界

【命令格式】 **ip pim bsr-border**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 为了限制 BSM 报文的泛洪,可以在接口上配置 BSR 边界,这样当这个接口收到 BSM 后会立即丢弃,BSM 也不会从这个接口转发。

配置 PIM 路由器对 BSM 报文限定

【命令格式】 **ip pim accept-bsr list { <1-99> | <1300-1999> | WORD }**

【参数说明】 **list access-list**: 引用 IP 标准 ACL 限定 BSR 地址范围,取值范围 1-99、1300-1999、Word

【命令模式】 全局模式

【使用指导】 配置此命令后,PIM-SM 路由器只接收合法 BSR 发出的 BSM 报文。

配置 C-BSR 可对 C-PR 范围限定检验

【命令格式】 **ip pim accept-crp list access-list**

【参数说明】 **list access-list**: 引用 IP 扩展 ACL 限定 C-RP 地址范围及其服务的组地址范围,取值范围:100-199、2000-2699、Word

【命令模式】 全局模式

【使用指导】 在 C-BSR 上配置此命令,当该 C-BSR 成为 BSR 时,能够对合法的 C-RP 地址范围及其所服务的组播组范围进行限制。

查看 BSM 报文

【命令格式】 **show ip pim sparse-mode bsr-router**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 -

查看本机上所有的 RP 及其服务的组报文

【命令格式】 **show ip pim sparse-mode rp mapping**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 -

配置举例

配置 BSR 边界

- 【配置方法】
- 配置 PIM-SM 基本功能。略
 - 路由器 A 上设置 C-RP 地址是 192.168.8.8。
 - 路由器 B 上设置 C-RP 地址是 192.168.5.5, C-BSR 地址是 192.168.6.6。
 - 路由器 B 上与路由器 A 交界的接口处配置 BSR 边界。

```
Ruijie# configure terminal
Ruijie(config)# int GigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# ip pim bsr-border
Ruijie(config)# end
```

【检验方法】 在配置边界以前，查看路由器 A 的 BSM 信息如下：

```
Ruijie# show ip pim sparse-mode bsr-router
PIMv2 Bootstrap information
This system is the Bootstrap Router (BSR)
  BSR address: 192.168.6.6
  Uptime:      01:14:25, BSR Priority: 64, Hash mask length: 10
  Next bootstrap message in 00:00:52
  Role: Candidate BSR  Priority: 64, Hash mask length: 10
  State: Elected BSR
  Candidate RP: 192.168.8.8(Loopback 0)
    Advertisement interval 60 seconds
    Next Cand_RP_advertisement in 00:00:06
```

 **Candidate RP:**表示是本路由器设置的所有的 C-RP，不包含其它路由器的。

在配置边界后，查看路由器 A 的 BSM 信息如下：

```
Ruijie# show ip pim sparse-mode bsr-router
```

配置 PIM 路由器对 BSM 报文限定，配置 BSM 的源地址过滤，限定源地址范围 (192.168.1.1 ~ 192.168.1.255)

【配置方法】

- 配置 PIM-SM 基本功能。略
- 路由器 A 上设置 C-RP 地址是 192.168.8.8。
- 路由器 B 上设置 C-RP 地址是 192.168.5.5，C-BSR 地址是 192.168.6.6。
- 配置 PIM 路由器 A 对 BSM 报文限定，限定的源地址范围 (192.168.1.1 ~ 192.168.1.255)

```
Ruijie# configure terminal
Ruijie(config)# ip pim accept-bsr list 1
% access-list 1 not exist
Ruijie(config)# access-list 1 permit 192.168.1.0 0.0.0.255
Ruijie(config)#
```

【检验方法】 配置 BSM 报文限定以前，查看路由器 A 的 BSM 信息如下：

```
Ruijie#show ip pim sparse-mode bsr-router
PIMv2 Bootstrap information
  BSR address: 192.168.6.6
  Uptime:      00:00:11, BSR Priority: 64, Hash mask length: 10
  Expires:     00:01:59
  Role: Non-candidate BSR  Priority: 0, Hash mask length: 10
  State: Accept Preferred
  Candidate RP: 192.168.8.8(Loopback 0)
    Advertisement interval 60 seconds
    Next Cand_RP_advertisement in 00:00:06
```

在配置 BSM 报文限定以后，查看路由器 A 的 BSM 信息如下：

```
Ruijie# show ip pim sparse-mode bsr-router
Candidate RP: 192.168.8.8(Loopback 0)
Advertisement interval 60 seconds
Next Cand_RP_advertisement in 00:00:06
```

配置 C-BSR 可对 C-PR 范围限定，配置 C-BSR 对 C-RP 范围过滤，过滤地址 (192.168.1.1 ~ 192.168.1.255)

【配置方法】

- 配置 PIM-SM 基本功能。略
- 路由器 A 上设置 C-RP 地址是 192.168.8.8。
- 路由器 B 上设置 C-RP 地址是 192.168.5.5，C-BSR 地址是 192.168.6.6。
- 路由器 B 对 C-RP 报文限定，限定的源地址范围 (192.168.1.1 ~ 192.168.1.255)

```
Ruijie# configure terminal
Ruijie(config)# ip pim accept-crp list 100
% access-list 1 not exist
Ruijie(config)# access-list 1 permit 192.168.1.0 0.0.0.255
Ruijie(config)#
```

【检验方法】

配置对 C-RP 过滤以前，在路由器 B 上查看所有的 RP 组的信息如下：

```
Ruijie#show ip pim sparse-mode rp mapping
PIM Group-to-RP Mappings
This system is the Bootstrap Router (v2)
Group(s): 224.0.0.0/4
  RP: 192.168.8.8(Not self)
    Info source: 192.168.8.8, via bootstrap, priority 192
    Uptime: 00:15:16, expires: 00:02:18
  RP: 192.168.5.5(Self)
    Info source: 192.168.6.6, via bootstrap, priority 192
    Uptime: 18:52:30, expires: 00:02:00
```

配置对 C-RP 过滤以后，在路由器 B 上查看所有的 RP 组的信息如下：

```
Ruijie#show ip pim sparse-mode rp mapping
PIM Group-to-RP Mappings
This system is the Bootstrap Router (v2)
Group(s): 224.0.0.0/4
  RP: 192.168.5.5(Self)
    Info source: 192.168.6.6, via bootstrap, priority 192
    Uptime: 21:38:20, expires: 00:02:10
```



路由器对 C-RP 过滤，只对其他路由路发送过来的 C-RP 报文过滤，不会对本路由器的 C-RP 过滤。

常见配置错误

- PIM-SM 基本功能没有配置，或配置失败。
- C-BSR 没有配置。

- BSR 边界配置，没有在不同域之间的接口上。

4.4.4 配置RP、DR参数

配置效果

- 配置忽略 C-RP 优先级，重新选择 RP。
- 配置源数据的 DR 可对 RP 可达性进行检测。
- 限制源注册的(S,G)地址范围，使 ASM 模型仅为允许范围内的组播报文提供组播服务。
- 配置源数据的 DR 发送注册报文速度进行限制。
- 配置注册报文的校验和的长度进行配置。
- 配置 RP 对注册报文解封装，在向下游接口进行分发。
- 配置注册报文的源地址。
- 配置注册报文的抑制时间。
- 配置 NULL 报文的探测时间。
- 配置 RP 收到注册报文，（S,G）生存时间。

注意事项

- 必须配置 PIM-SM 基本功能。

配置方法

配置忽略 C-RP 优先级，重新选择 RP

- 可选配置。
- 若无特殊要求，可在每台路由器上启动忽略 C-RP 优先级。

配置源数据的 DR 可对 RP 可达性进行检测

- 可选配置。
- 若无特殊要求，可在直连数据源的 DR 上启动可达性检测。

限制源注册的(S,G)地址范围

- 可选 配置
- 若无特殊要求，可在所有作为 C-RP 或静态 RP 的路由器上配置源注册过滤。

限制源的 DR 发送注册报文速率

- 可选 配置
- 若无特殊要求，可在直连数据源的 DR 上启动注册报文速度限制。

↘ 对注册报文检验和的长度进行配置

- 可选配置
- 若无特殊要求，可在所有的 C-RP 或静态的 RP 上进行配置。

↘ 对 RP 解封装注册报文后，是否分发数据进行配置

- 可选配置
- 若无特殊要求，可在所有的 C-RP 或静态的 RP 上进行配置。

↘ 配置注册报文的源地址

- 可选配置
- 若无特殊要求，可在直连数据源的 DR 上设置源地址。

↘ 配置注册报文的抑制时间。

- 可选配置
- 若无特殊要求，可在直连数据源的 DR 上设置抑制时间。

↘ 配置 NULL 报文的探测时间。

- 可选配置
- 若无特殊要求，可在直连数据源的 DR 上设置 NULL 报文的探测时间。

↘ 配置 RP 收到注册报文，(S,G) 生存时间。

- 可选配置
- 若无特殊要求，可在所有作为 C-RP 或静态 RP 的路由器上配置 (S,G) 生存时间。

检验方法

↘ 忽略 C-RP 优先级检验

路由器 A 上设置 C-RP 地址 192.168.8.8，默认优先级 192。路由器 B 上设置 C-RP 的地址为 192.168.5.5，优先级为 200，路由器 B 上设置 C-BSR 的地址是 192.168.6.6。

- 使用 `show ip pim sparse-mode rp` 233.3.3.3 命令，查看当前组所对应的 RP。

↘ 源数据的 DR 对 RP 可达性进行检测检验

路由器 A 上设置 C-RP 地址 192.168.8.8，默认优先级 192。路由器 B 上设置 C-RP 的地址为 192.168.5.5，默认优先级 192，路由器 B 上设置 C-BSR 的地址是 192.168.6.6。在路由器 B 上配置 RP 可达性检测。

- 使用 `show running-config` 命令，查看是否配置

↘ 限制源注册的(S,G)地址范围

路由器 A 上设置 C-RP 地址 192.168.8.8，默认优先级 192。路由器 B 上设置 C-BSR 的地址是 192.168.6.6。设置源地址是 192.168.1.100，发送的组播组是 233.3.3.3。路由器 A 上设置只接收源地址为 (192.168.2.0 ~ 192.168.2.255) 组播源。

- 使用 `show ip pim sparse-mode mroute` 命令，查看 (S , G) 表项。

限制源 DR 发送注册报文的速率

设置路由器 B 发送注册报文速率，然后通过 **show ip pim sparse-mode track** 命令，查看注册报文的发送个数，进行确认。

对注册报文检验和的长度进行配置

设置路由器 A 上注册报文校验为整个报文，而不是只计算报文头部和注册报文头部。通过 **show running-config** 命令，查看。

对 RP 注册报文解封装后，分发注册报文

设置路由器 A 上解封装注册报文，在分发报文，通过 **show running-config** 命令，通过 **show running-config** 命令，查看。

配置注册报文的源地址

设置路由器 B 上注册报文源地址，通过路由器 B 的 **show running-config** 命令，查看。

配置注册报文的抑制时间和探测报文时间

设置路由器 B 上注册报文的抑制时间和探测时间，然后通过 **show ip pim sparse-mode track** 命令，查看。

配置 RP 收到注册报文，(S,G) 生存时间。

设置路由器 A 上收到注册报文，(S,G) 生存时间，然后通过 **show ip pim sparse-mode mroute** 命令，查看 (S,G) 生存最大时间。

相关命令

忽略 C-RP 优先级

- 【命令格式】 **ip pim ignore-rp-set-priority**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

显示组所对应的 RP

- 【命令格式】 **show ip pim sparse-mode rp-hash group-address**
- 【参数说明】 *group-address*: 被解析的组地址
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 -

源直连的 DR 对 RP 可达性检测

- 【命令格式】 **ip pim register-rp-reachability**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 配置此命令后，会在发送注册报文前，判断 RP 是否可达。若可达则发送，若不可达则不发送。

限制源注册的(S,G)地址范围

- 【命令格式】 **ip pim accept-register** { **list** *access-list* [**route-map** *map-name*] | **route-map** *map-name* [**list** *access-list*] }
- 【参数说明】 **list** *access-list* : 引用 IP 扩展 ACL 限定(S,G)地址范围。取值范围：100-199、2000-2699、Word
route-map *map-name* : 使用 route map 限定(S,G)地址范围
- 【命令模式】 全局模式
- 【使用指导】 在静态 RP 或 C-RP 上配置此命令，限制特定源地址组地址的注册报文

查看组播路由表项命令

- 【命令格式】 **show ip pim sparse-mode mroute** [*group-or-source-address* [*group-or-source-address*]] [**proxy**]
- 【参数说明】 *group-or-source-address* : 组地址或源地址，两个地址不能同时都是组地址或者同时都是源地址。
Proxy : 显示表项携带的 RPF Vector 信息。
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 每次只能指定一个组地址，或是一个源地址，或是一个组地址和一个源地址，或是不指定具体的组地址或源地址。不能同时指定两个组地址或是两个源地址。

配置 DR 发送注册报文速度限制的命令

- 【命令格式】 **ip pim register-rate-limit** *rate*
- 【参数说明】 *Rate* : 每 S 最多允许发送多少个 Register 报文，取值范围<1-65535>
- 【命令模式】 全局模式
- 【使用指导】 它是对每个 (S , G) 状态的注册报文发送速度进行配置，并不是对整个系统的注册报文发送速度进行配置。启动此命令将降低源 DR 和 RP 的负担，不超过限制的注册报文才能被发送。

查看 PIM 报文的统计个数

- 【命令格式】 **show ip pim sparse-mode track**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 当系统启动时，会第一次设置统计的开始时间点，每次调用 `clear ip pim sparse-mode track` 时，会重新设置统计的开始时间点并清除 PIM 报文的计数器。

配置注册报文的校验和必须计算整个报文

- 【命令格式】 **ip pim register-checksum-wholepkt** [**group-list** *access-list*]
- 【参数说明】 **group-list** *access-list* : 使用 ACL 限制适用此配置的组地址。
access-list : 支持数字<1, 99>、<1300, 1999>，支持命名 ACL
- 【命令模式】 全局模式
- 【使用指导】 对注册报文校验和的计算方式为计算整个 pim 协议报文含其中封装的组播数据报文，而非单独的注册报文的 pim 头部。
如果携带参数 **group-list** *access-list*，则对所有组地址应用此配置。

允许 RP 对注册报文进行解封并转发其中的组播数据报文

- 【命令格式】 **ip pim register-decapsulate-forward**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 在静态 RP 或 C-RP 上配置此命令，对接收到的带组播数据报文的注册报文进行解封并转发其中的组播数据报文。

由于是软件对注册报文进行解封装并转发其中的组播数据报文，因此如果需要解封装并转发的注册报文较多的情况下，配置了该功能，会导致 cpu 处理较繁忙；不建议配置该命令。

配置注册报文的源地址

【命令格式】 **ip pim register-source** { *local_address* | *Interface-type interface-number* }

【参数说明】 *local_address*: 指定注册报文的源 IP 地址

interface-type interface-number: 指定这个接口的 IP 作为注册报文的源 IP 地址

【命令模式】 全局模式

【使用指导】 配置的地址必须是可达的。当 RP 发送出正确的 Register-Stop 信息时，它要作出相应的反应。所以推荐它是接口的回环地址，但也可以是其它的物理地址。

此配置不要求 PIM 已经启动。

设置注册报文抑制时间

【命令格式】 **ip pim register-suppression** *seconds*

【参数说明】 *Seconds*: 注册抑制时间，单位为 S，取值范围 1-65535，缺省值 60

【命令模式】 全局模式

【使用指导】 在 DR 上配置此值将修改定义在 DR 上的注册报文抑制时间；如果 ip pim rp-register-kat 命令没有配置，在 RP 上定义此值将修改 RPkeepalive 的周期。

设置注册探测时间

【命令格式】 **ip pim probe-interval** *seconds*

【参数说明】 *Seconds*: 注册探测时间，单位为 S，取值范围 1-65535，缺省值 5

【命令模式】 全局模式

【使用指导】 注册探测时间，即在注册抑制时间超时前，允许源 DR 向 RP 发送空注册报文即 NULL-Register 的时间间隔。

注册探测时间间隔不能超过注册抑制时间的一半，否则，配置不会成功。且会输出一个提醒即 Warning。而且 3*注册抑制时间+注册探测时间不能超过 65535，否则，会输出一个提醒即 Warning。

配置 RP 上的 KAT 间隔

【命令格式】 **ip pim rp-register-kat** *seconds*

【参数说明】 *Seconds*: KAT 定时器的时间。

单位为 S，取值范围 1-65535，缺省值 210。

【命令模式】 全局模式

【使用指导】 -

配置举例

配置有没有考虑 C-RP 优先级时，组地址所对应的 RP。

【配置方法】

- 配置 PIM-SM 基本功能。略
- 设置路由器 A 上的 C-RP 地址是 192.168.8.8。
- 设置路由器 B 上的 C-RP 地址是 192.168.5.5，优先级是 200，C-BSR 地址是 192.168.6.6。
- 查看 233.3.3.3 所对应的组。
- 在路由器 B 设置忽略 C-RP 的优先。

```
Ruijie# configure terminal
Ruijie(config)# ip pim ignore-rp-set-priority
```

【检验方法】

在配置忽略 C-RP 的优先级以前，显示信息如下。

```
Ruijie# show ip pim sparse-mode rp-hash 233.3.3.3
RP: 192.168.8.8
Info source: 192.168.8.8, via bootstrap

PIMv2 Hash Value 10(mask 255.192.0.0)
RP 192.168.8.8, via bootstrap, priority 192, hash value 1084558102
RP 192.168.5.5, via bootstrap, priority 200, hash value 1094656709
```

在配置忽略 C-RP 的优先级以后，显示信息如下。

```
Ruijie# show ip pim sparse-mode rp-hash 233.3.3.3
RP: 192.168.5.5
Info source: 192.168.6.6, via bootstrap

PIMv2 Hash Value 10(mask 255.192.0.0)
RP 192.168.8.8, via bootstrap, priority 192, hash value 1084558102
RP 192.168.5.5, via bootstrap, priority 200, hash value 1094656709
```

▾ 配置源直连的 RP 的可达性检测

【配置方法】

- 配置 PIM-SM 基本功能。略
- 配置源直连的 RP 的可达性检测。

```
Ruijie(config)# ip pim register-rp-reachability
```

【检验方法】

使用 show running-config 命令，查看出现如下显示信息：

```
Ruijie(config)#show running-config
ip pim register-rp-reachability
```

▾ 限制源注册的(S,G)地址范围

【配置方法】

- 配置 PIM-SM 基本功能。略
- 设置路由器 A 对源地址进行过滤，只接收源地址为 (192.168.2.0 ~ 192.168.2.255)。

```
Ruijie#show ip pim sparse-mode mroute
Ruijie(config)#ip pim accept-register list 101
% access-list 101 not exist
Ruijie(config)#access-list 101 permit ip 192.168.2.0 0.0.0.255 any
Ruijie#show ip pim sparse-mode mroute
```

【检验方法】

在配置源注册过滤以前，使用 show ip pim sparse-mode mroute 命令，查看组播表项，存在 (S , G) 表

项和 (S , G , RPT) 表项。

```

Ruijie#show ip pim sparse-mode mroute
IP Multicast Routing Table

(*,*,RP) Entries: 0
(*,G) Entries: 1
(S,G) Entries: 1
(S,G,rpt) Entries: 1
FCR Entries: 0
REG Entries: 0

(192.168.1.100, 233.3.3.3)
RPF nbr: 192.168.36.90
RPF idx: VLAN 1
SPT bit: 0
Upstream State: NOT JOINED
kat expires in 187 seconds
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .
Joined
0 . . . . .
. .
Asserted
0 . . . . .
. .
Outgoing
0 . . . . .
. .

(192.168.1.100, 233.3.3.3, rpt)
RP: 192.168.8.8
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: RPT NOT JOINED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .
Pruned
0 . . . . .

```

```

. .
Outgoing
0 . . . . .
. .

(*, 239.255.255.250)
RP: 192.168.8.8
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .
Joined
0 .
j . . . . .
Asserted
0 . . . . .
. .
FCR:

```

在配置源注册过滤以后，使用 show ip pim sparse-mode mroute 命令，查看组播表项，存在 (S , G) 表项和 (S , G , RPT) 表项。

```

Ruijie#show ip pim sparse-mode mroute
IP Multicast Routing Table

(*,*,RP) Entries: 0
(*,G) Entries: 1
(S,G) Entries: 0
(S,G,rpt) Entries: 0
FCR Entries: 0
REG Entries: 0

(*, 239.255.255.250)
RP: 192.168.8.8
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .

```

```

Joined
0 .
j . . . . .
Asserted
0 . . . . .
. .
FCR:

```

限制源 DR 发送注册报文的速率

【配置方法】

- 配置 PIM-SM 基本功能。略
- 查看路由器 B 发送 PIM 报文个数。
- 1 秒后查看路由器 B 发送 PIM 报文个数。
- 设置路由器 B 发送注册报文的速率。
- 1 秒后查看路由器 B 发送 PIM 报文个数。

```
Ruijie (config)#ip pim register-rate-limit 1
```

【检验方法】

设置速率限制以前，查看 DR 路由器 PIM 报文个数,提示信息如下：

```

Ruijie#show ip pim sparse-mode track
PIM packet counters track
Elapsed time since counters cleared: 04d01h01m

                received                sent
Valid PIM packets:    18754                29771
Hello:                11149                17842
Join-Prune:           0                 3234
Register:             0                 3211
Register-Stop:       3192                 0
Assert:               0                 0
BSM:                  0                 5484
C-RP-ADV:             4413                 0
PIMDM-Graft:         0
PIMDM-Graft-Ack:     0
PIMDM-State-Refresh: 0
Unknown PIM Type:    0

Errors:
Malformed packets:    0
Bad checksums:        0
Send errors:          0
Packets received with unknown PIM version: 0
Ruijie#

```

设置速率限制以前，1 秒后，查看 DR 路由器 PIM 报文个数,提示信息如下：

```

Ruijie #show ip pim sparse-mode track
PIM packet counters track

```

```

Elapsed time since counters cleared: 04d01h04ms

                received                sent
Valid PIM packets:    18765                29789
Hello:                11154                17852
Join-Prune:           0                  3236
Register:             0                  3214
Register-Stop:       3195                 0
Assert:              0                  0
BSM:                 0                  5487
C-RP-ADV:            4416                 0
PIMDM-Graft:         0
PIMDM-Graft-Ack:     0
PIMDM-State-Refresh: 0
Unknown PIM Type:    0

Errors:
Malformed packets:           0
Bad checksums:               0
Send errors:                 0
Packets received with unknown PIM version: 0
Ruijie#

```

设置速率限制以后，查看 DR 路由器 PIM 报文个数,提示信息如下：

```

Ruijie#show ip pim sparse-mode track
PIM packet counters track
Elapsed time since counters cleared: 04d01h06m

                received                sent
Valid PIM packets:    18777                29808
Hello:                11159                17862
Join-Prune:           0                  3239
Register:             0                  3215
Register-Stop:       3196                 0
Assert:              0                  0
BSM:                 0                  5489
C-RP-ADV:            4419                 0
PIMDM-Graft:         0
PIMDM-Graft-Ack:     0
PIMDM-State-Refresh: 0
Unknown PIM Type:    0

Errors:
Malformed packets:           0
Bad checksums:               0

```

```
Send errors: 0
Packets received with unknown PIM version: 0
Ruijie#
```

▾ 对注册报文检验和长度进行配置

【配置方法】

- 配置 PIM-SM 基本功能。略
- 设置路由器 A 注册报文的校验和必须计算整个报文。
- 通过 show running-config 命令，查看是否配置。

```
Ruijie(config)#ip pim register-checksum-wholepkt
```

【检验方法】

查看路由器 A 上配置，如下：

```
Ruijie#show running-config
. . . . .
!
ip pim register-checksum-wholepkt
ip pim rp-candidate Loopback 0
!
. . . . .
```

▾ RP 对注册报文进行解封并转发

【配置方法】

- 配置 PIM-SM 基本功能。略
- 设置路由器 A 转发注册报文。
- 通过 show running-config 命令，查看是否配置。

```
Ruijie(config)#ip pim register-decapsulate-forward
```

【检验方法】

查看路由器 A 的配置，如下：

```
Ruijie#show running-config
. . . . .
!
!
ip pim register-decapsulate-forward
ip pim register-checksum-wholepkt
ip pim rp-candidate Loopback 0
!
!
!
. . . . .
```

▾ 配置注册报文的源地址

【配置方法】

- 配置 PIM-SM 基本功能。略
- 设置 Loop 2 口的原地址是 192.168.2.2。
- 设置路由器 B 的注册报文的源地址为 Loop 2 口。
- 通过 show running-config 命令，查看是否配置。

【检验方法】 查看路由器 B 的配置，如下：

```
Ruijie#show running-config
!
!
!
ip pim register-source Loopback 1
ip pim bsr-candidate Loopback 0
!
!
!
!
```

配置注册报文的抑制时间和探测报文时间

【配置方法】

- 配置 PIM-SM 基本功能。略
- 设置路由器 B 的抑制时间为 20S。
- 设置路由器 B 的探测时间为 2S。
- 通过 show ip pim sparse-mode track 命令，查看注册报文个数。

```
Ruijie(config)#ip pim register-suppression 20
Ruijie(config)#ip pim probe-interval 2
```

【检验方法】 在路由器 B 上，显示 PIM 报文个数。提示信息如下：

```
Ruijie#show ip pim sparse-mode track
PIM packet counters track
Elapsed time since counters cleared: 04d23h15m

                received                sent
Valid PIM packets:    23788                43249
Hello:                13817                23178
Join-Prune:           0                    4568
Register:             0                    8684
Register-Stop:       4223                   0
Assert:              0                    0
BSM:                  0                    6819
C-RP-ADV:            5748                   0
PIMDM-Graft:         0
PIMDM-Graft-Ack:     0
PIMDM-State-Refresh: 0
Unknown PIM Type:    0

Errors:
Malformed packets:    0
Bad checksums:        0
Send errors:          0
```

```
Packets received with unknown PIM version: 0
```

```
Ruijie#
```

```
Ruijie#
```

18S 后，在路由器 B 上，显示 PIM 报文个数。提示信息如下：

```
Ruijie#show ip pim sparse-mode track
```

```
PIM packet counters track
```

```
Elapsed time since counters cleared: 04d23h17m
```

	received	sent
Valid PIM packets:	23798	43263
Hello:	13820	23184
Join-Prune:	0	4569
Register:	0	8685
Register-Stop:	4224	0
Assert:	0	0
BSM:	0	6820
C-RP-ADV:	5749	0
PIMDM-Graft:	0	
PIMDM-Graft-Ack:	0	
PIMDM-State-Refresh:	0	
Unknown PIM Type:	0	

```
Errors:
```

```
Malformed packets: 0
```

```
Bad checksums: 0
```

```
Send errors: 0
```

```
Packets received with unknown PIM version: 0
```

```
Ruijie#
```

配置 RP 收到注册报文，(S,G) 生存时间。

【配置方法】

- 配置 PIM-SM 基本功能。略
- 设置路由器 A 收到注册报文，(S,G) 生存时间 60S。
- 通过 show ip pim sparse-mode mroute 命令，查看注册报文个数。

```
Ruijie(config)#ip pim rp-register-kat 60
```

【检验方法】

设置生存时间后，查看路由器 A 上的 (S,G) 生存时间，不会超过 60S。

```
Ruijie(config)#show ip pim sparse-mode mroute
```

```
IP Multicast Routing Table
```

```
(*,* ,RP) Entries: 0
```

```
(* ,G) Entries: 1
```

```
(S,G) Entries: 1
```

```
(S,G,rpt) Entries: 1
```

```
FCR Entries: 0
```

```

REG Entries: 0

(192.168.1.100, 233.3.3.3)
RPF nbr: 192.168.36.90
RPF idx: VLAN 1
SPT bit: 0
Upstream State: NOT JOINED
kat expires in 49 seconds
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .
Joined
0 . . . . .
. .
Asserted
0 . . . . .
. .
Outgoing
0 . . . . .
. .

(192.168.1.100, 233.3.3.3, rpt)
RP: 192.168.8.8
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: RPT NOT JOINED
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .
Pruned
0 . . . . .
. .
Outgoing
0 . . . . .
. .

(*, 239.255.255.250)
RP: 192.168.8.8
RPF nbr: 0.0.0.0
RPF idx: None

```

```

Upstream State: JOINED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .
Joined
0 .
j . . . . .
Asserted
0 . . . . .
. .
FCR:

Ruijie(config)#
Ruijie(config)#show ip pi

```

常见配置错误

- PIM-SM 基本功能没有配置，或配置失败。
- 某 C-RP 或静态 RP 设备上没有限制源注册(S,G)地址范围，或配置失败。
- 限制源注册(S,G)地址范围时引用的 ACL 没有配置，或此 ACL 允许的源/组地址范围配置错误。
- 各 C-RP 或静态 RP 设备上允许注册的源/组地址范围不一致。

4.4.5 配置join/prune报文的发送间隔

配置效果

- 修改 join/prune 报文的发送间隔，形成 RPT 树或 SPT 树。

注意事项

- 必须配置 PIM-SM 基本功能。

配置方法

- 配置 join/prune 报文的发送间隔

检验方法

路由器 B 上设置 join/prune 报文的发送间隔 120S，通过 **show ip pim sparse-mode mroute** 命令，查看表项生存时间。

相关命令

↘ join/prune 报文的发送间隔

- 【命令格式】 **ip pim jp-timer seconds**
- 【参数说明】 *Seconds* : Join/Prune 报文的发送周期。
单位为秒，取值范围<1-65535>，缺省值 60
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

↘ 配置路由器的 join/prune 报文的发送间隔

- 【配置方法】
- 配置 PIM-SM 基本功能。略
 - 配置路由器的 join/prune 报文的发送间隔

```
Ruijie(config)#ip pim jp-timer 120
```

- 【检验方法】 通过 **show ip pim sparse-mode mroute** 命令，查看表项 join/prune 最大等待超时时间。

```
Ruijie(config)#show ip pim sparse-mode mroute
```

```
IP Multicast Routing Table
```

```
(*,* ,RP) Entries: 0
```

```
(* ,G) Entries: 1
```

```
(S ,G) Entries: 1
```

```
(S ,G ,rpt) Entries: 1
```

```
FCR Entries: 0
```

```
REG Entries: 1
```

```
(192.168.1.100, 233.3.3.3)
```

```
RPF nbr: 0.0.0.0
```

```
RPF idx: None
```

```
SPT bit: 1
```

```
Upstream State: JOINED
```

```
jt_timer expires in 96 seconds
```

```
kat expires in 92 seconds
```

```
  00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
```

```
Local
```

```
0 . . . . .
. .
```

```
1 . . . . .
. .
```

```
Joined
```

```
0 . . . . .
. .
```

```

1 . . . . .
. .
Asserted
0 . . . . .
. .
1 . . . . .
. .
Outgoing
0 . . . . .
. .
1 . .
0 . . . . .
.

(192.168.1.100, 233.3.3.3, rpt)
RP: 192.168.8.8
RPF nbr: 192.168.36.89
RPF idx: GigabitEthernet 0/1
Upstream State: RPT NOT JOINED
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
31
Local
0 . . . . .
. .
1 . . . . .
. .
Pruned
0 . . . . .
. .
1 . . . . .
. .
Outgoing
0 . . . . .
. .
1 . . . . .
. .

```

```

(*, 239.255.255.250)
RP: 192.168.8.8
RPF nbr: 192.168.36.89
RPF idx: GigabitEthernet 0/1
Upstream State: JOINED
jt_timer expires in 119 seconds
 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

```

```

Local
0 . . . . . i . . . . .
1 . . . . .
Joined
0 . . . . .
1 . . . . .
Asserted
0 . . . . .
1 . . . . .
FCR:

VSU(config)#

```

常见配置错误

- PIM-SM 基本功能没有配置，或配置失败。

4.4.6 最后一跳设备由从共享树切换到最短路径树

配置效果

- 将 RPT 树切换到 SPT 树。

注意事项

- 必须配置 PIM-SM 基本功能。

配置方法

- 最后一跳设备由从共享树切换到最短路径树

检验方法

先配置 PIM-SM 基本功能，源的 DR 发送数据流（*，233.3.3.3）的组，接收端强制加入 233.3.3.3 的组，开成 RPT 组，接收端的 DR 强制进行 RPT 到 SPT 的切换，查看 **show running-config** 配置命令。

相关命令

▾ 启动 SPT 切换功能

【命令格式】 **ip pim spt-threshold [group-list access-list]**

【参数说明】 **group-list access-list**：引用 ACL 限定允许 SPT 切换的组地址范围。

access-list：支持数字 ACL 范围为 1-99、1300-1999；也支持命名 ACL。

- 【命令模式】 全局模式
- 【使用指导】 如果不携带参数 `group-list access-list`，允许所有组进行 SPT 切换。

配置举例

最后一跳设备由从共享树切换到最短路径树

- 【配置方法】
- 配置 PIM-SM 基本功能。略
 - 数据源的 DR 发送组为 233.3.3.3 的码流。
 - 接收端的 DR 接收组为 233.3.3.3 的码流。
 - 接收端的 DR 设置最后一跳设备由从共享树切换到最短路径树

```
Ruijie(config)#ip pim spt-threshold
```

- 【检验方法】 通过 `show running-config` 命令，查看配置信息。

```
!
!
ip pim jp-timer 120
ip pim spt-threshold
ip pim rp-candidate Loopback 0
!
!
!
```

4.5 监视与维护

清除各类信息

 在设备运行过程中执行 `clear` 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除组播路由表项。	<code>clear ip mroute { * group-address [source-address] }</code>
清除组播路由的统计信息。	<code>clear ip mroute statistics { * group-address [source-address] }</code>
清除动态 RP 信息。	<code>clear ip pim sparse-mode bsr rp-set *</code>
清除 PIM-SM 协议报文的统计信息。	<code>clear ip pim sparse-mode track</code>

查看运行情况

作用	命令
使用此命令显示 BSR 的详细信息。	<code>show ip pim sparse-mode bsr-router</code>
显示接口的 PIM-SM 信息。	<code>show ip pim sparse-mode interface [interface-type interface-number] [detail]</code>
显示 PIM-SM 的接口的本地 IGMP 信息。	<code>show ip pim sparse-mode local-members [interface-type interface-number]</code>

显示 PIM-SM 组播路由表信息 ,proxy 参数显示 PIM-SM 表项的 RPF Vector 信息。	show ip pim sparse-mode mroute [<i>group-or-source-address</i>] [<i>group-or-source-address</i>] [proxy]
显示 PIM-SM 邻居信息。	show ip pim sparse-mode neighbor [detail]
显示来自 NSM 的 PIM-SM 下一跳信息。	show ip pim sparse-mode nexthop
使用此命令显示出某个组地址 <i>group-address</i> 对应的 RP 的信息。	show ip pim sparse-mode rp-hash <i>group-address</i>
使用此命令查看当前所有 RP 以及它们所服务的组的信息。	show ip pim sparse-mode rp mapping
查看从统计的开始时间点到现在,发送和接收的 PIM 报文的个数	show ip pim sparse-mode track

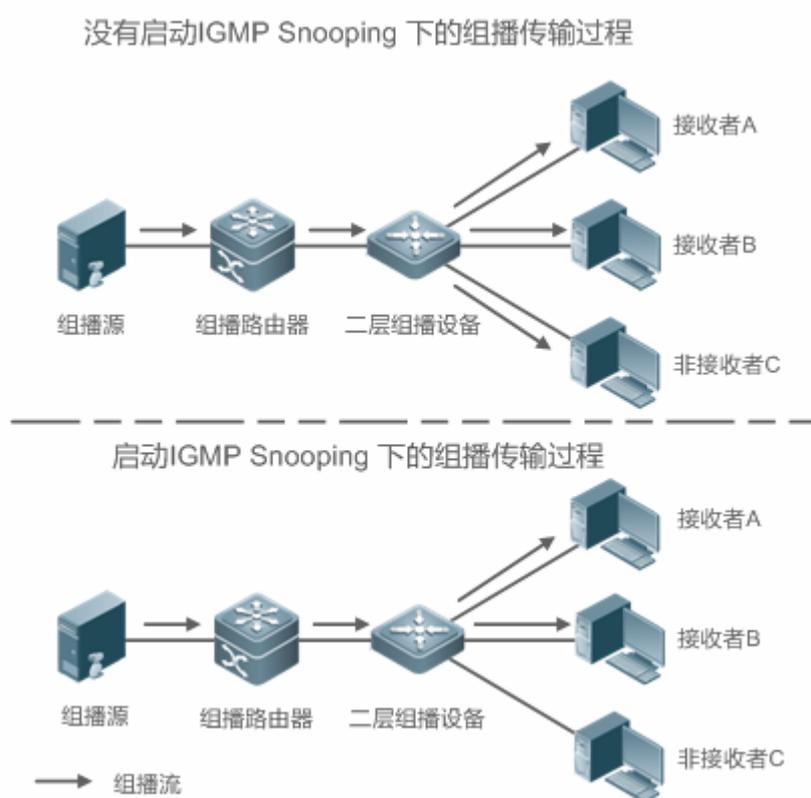
5 IGMP Snooping

5.1 概述

IGMP Snooping (Internet Group Management Protocol Snooping, 组播侦听器发现协议窥探) 是运行在 VLAN 上的 IP 组播窥探机制, 用于管理和控制 IP 组播流在 VLAN 内的转发, 实现二层组播功能。

如下图所示, 当二层组播设备没有运行 IGMP Snooping 时, IP 组播报文在 VLAN 内被广播; 当二层组播设备运行了 IGMP Snooping 后, IP 组播报文只发给组成员接收者。

图 5-1 IP 组播流在 VLAN 内的转发 (二层组播设备启动 IGMP Snooping 前后的对比)



协议规范

- RFC4541 : Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches

5.2 典型应用

典型应用	场景描述
二层组播控制	二层组播精确转发, 避免组播报文在二层泛洪。
公共组播服务 (组播VLAN)	多个 VLAN 的用户共享同一 VLAN 的组播流。

收费频道与预览

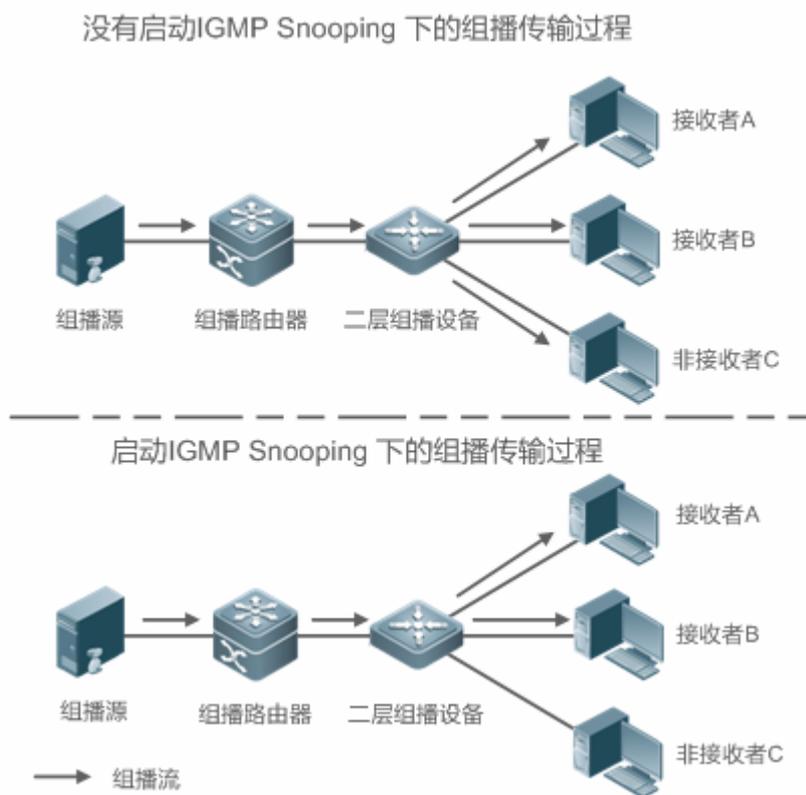
控制用户点播的组播范围，对禁止用户点播的组提供预览。

5.2.1 二层组播控制

应用场景

组播报文需要通过二层交换机设备转发给用户，如下图所示，在未启用二层组播控制的情况下，报文是通过泛洪发送给所有的用户，即使没有该组播接收需求的用户也会接收到该组播流。当二层组播设备运行了 IGMP Snooping，进行二层组播控制后，IP 组播组的组播数据报文不会在 VLAN 内被广播，而是发给指定的接收者。

图 5-2 二层组播控制（组播 VLAN）



功能部属

配置 IGMP Snooping 基本功能

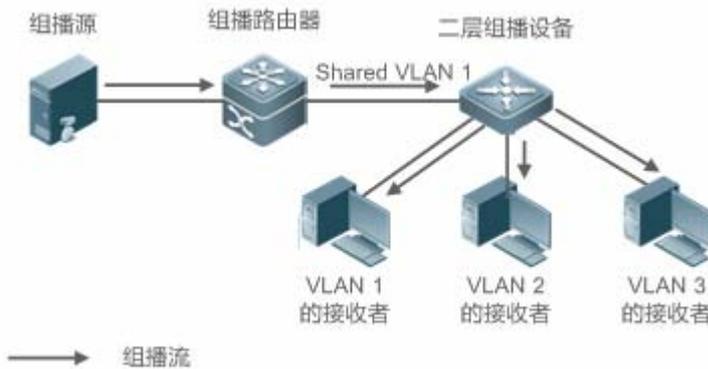
5.2.2 公共组播服务（组播VLAN）

应用场景

在 SVGL 模式 (或 IVGL-SVGL 模式) 下, 运行 IGMP Snooping 的设备可以为用户 VLAN 提供公共组播服务 (也称为 “组播 VLAN”)。公共组播服务常用于为多个 VLAN 的用户提供相同的视频点播。

如下图所示, 二层组播设备运行在 IGMP Snooping 的 SVGL 模式下。组播路由器向 VLAN1 发送了一份组播报文, 二层组播设备自动将此报文分发到 VLAN1、VLAN2、VLAN3。VLAN1 中的组播服务被 VLAN3 共享。

图 5-3 公共组播服务 (组播 VLAN)



i 如果二层组播设备运行在 IGMP Snooping 的 IVGL 模式下, 则组播路由器必须为每个 VLAN 发送一份组播报文。这样既造成了带宽的浪费, 也给设备增加了额外的负担。

功能部属

- 配置 IGMP Snooping 基本功能 (SVGL 模式 或 IVGL-SVGL 模式)。

5.2.3 收费频道与预览

应用场景

在视频点播应用中, 通过限定用户主机能够加入的组的地址范围, 使未付费的用户不能观看收费频道的节目。此时存在一种需求: 用户想先了解收费频道的节目, 再决定是否付费观看。

可以针对收费频道启用预览功能, 使未付费的用户在点播收费频道的节目时, 能够预览一段时间 (如 1 分钟), 而后节目停止播放。

功能部属

- 配置 IGMP Snooping 基本功能 (任意工作模式)。
- 控制用户点播的组的地址范围。
- 对禁止用户点播的组启动预览功能。

5.3 功能详解

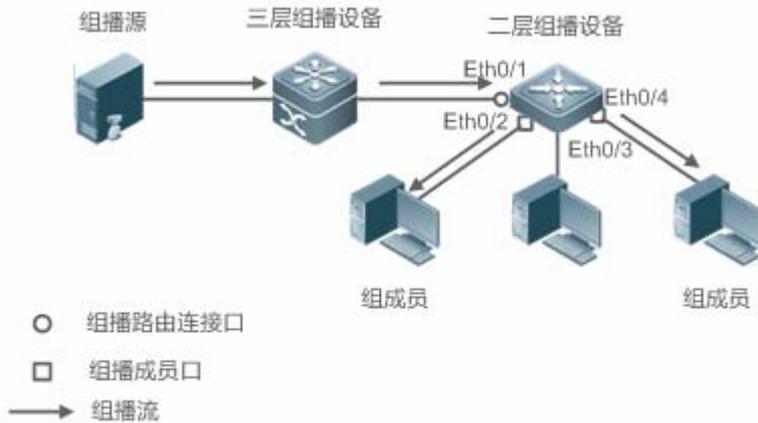
基本概念

路由连接口、成员口

i IGMP Snooping 功能是基于 VLAN 进行的，相关的端口也是指 VLAN 内部的成员口。

运行 IGMP Snooping 的设备将 VLAN 内的接口标识为路由连接口或成员口，从而能够管理和控制 IP 组播流在 VLAN 内的转发。如下图所示，在二层组播设备上运行 IGMP Snooping，组播流从路由连接口进入，从成员口发出。

图 5-4 IGMP Snooping 的两类端口



- 路由连接口 (Multicast Router Port) : 二层组播设备上连接组播路由器 (三层组播设备) 的端口，指示组播源的方向。通过监听 IGMP 报文，二层组播设备可以自动发现并维护动态路由连接口。也允许用户配置静态路由连接口。
- 成员口 (Member Port) : 二层组播设备上连接组成员主机的端口，指示组成员的方向。又称侦听器端口 (Listener Port)。通过监听 IGMP 报文，二层组播设备可以自动发现并维护动态成员口。也允许用户配置静态成员口。

IGMP Snooping 转发表项

运行 IGMP Snooping 的设备按照 IGMP Snooping 转发表项转发 IP 组播报文。

IGMP Snooping 转发表项中包含如下信息：源地址 (S)、组地址 (G)、VLAN 号 (VLAN_id)、路由连接口、成员口。表示：符合 (S,G,VLAN_id) 特征的报文，应该从路由连接口进入、从成员口发出。一个 IGMP Snooping 转发表项用一组 (S,G,VLAN_id) 来标识。

通过 **show ip igmp snooping gda-table** 命令查看 IGMP Snooping 转发表项。

```
Ruijie# show ip igmp snooping gda-table
Multicast Switching Cache Table
D: DYNAMIC //动态成员口
S: STATIC //静态成员口
M: MROUTE //路由连接口，不区分静态或动态
(*, 233.3.6.29, 1): //(源地址任意,组地址 233.3.6.29,VLAN1)
VLAN(1) 3 OPORTS:
GigabitEthernet 0/3(S)
GigabitEthernet 0/2(M)
GigabitEthernet 0/1(D)

(*, 233.3.6.30, 1): //(源地址任意,组地址 233.3.6.30,VLAN1)
VLAN(1) 2 OPORTS:
GigabitEthernet 0/2(M)
```

GigabitEthernet 0/1(D)

功能特性

功能特性	作用
监听IGMP报文	发现并识别路由接口和成员口，从而建立并维护 IGMP Snooping 转发表项。
IGMP Snooping工作模式	为用户 VLAN 提供独立组播服务或公共组播服务。
组播安全控制	控制组播服务范围和负荷，防范非法组播流。
Profile	定义允许或禁止用户点播的组地址范围，供其他功能引用。
对QinQ的处理	设置 QinQ 接口对组播报文的转发方式
IGMP查询器	在没有三层组播设备的网络中，二层组播设备自己充当 IGMP 查询器。

5.3.1 监听IGMP报文

运行 IGMP Snooping 的设备通过对收到的 IGMP 报文进行分析，发现并识别路由接口和成员口，从而建立并维护 IGMP Snooping 转发表项。

工作原理

运行 IGMP Snooping 的设备可以识别并处理如下几类 IGMP 报文。

Query 报文

i IGMP 查询者周期性发送普遍组 Query 报文。当 IGMP 查询者收到 Leave 报文时，发送指定组 Query 报文。

运行 IGMP Snooping 的设备在收到 Query 报文时，在此 VLAN 内执行如下动作：

- 将 IGMP 查询报文向所有端口（除该报文的接收端口）转发出去。
- 如果收到 IGMP 查询报文的端口是动态路由接口，则重置其老化定时器。如果定时器超时，则该端口不再作为动态路由接口。
- 如果收到 IGMP 查询报文的端口不是路由接口，则将其作为动态路由接口，并启动其老化定时器。如果定时器超时，则该端口不再作为动态路由接口。
- 如果是普遍组查询报文，则对所有动态成员口重置其老化定时器。如果定时器超时，则该端口不再作为任意组的动态成员口。默认情况下采用 IGMP 查询报文携带最大响应时间作为老化定时器的超时时间，如果配置了“查询报文最大响应时间”，则使用“查询报文最大响应时间”作为老化定时器的超时时间。
- 如果是指定组查询报文，则对指定组的动态成员口重置其老化定时器。如果定时器超时，则该端口不再作为指定组的动态成员口。默认情况下采用 IGMP 查询报文携带最大响应时间作为老化定时器的超时时间，如果配置了“查询报文最大响应时间”，则使用“查询报文最大响应时间”作为老化定时器的超时时间。
- 如果配置关闭“动态学习路由接口”功能，则 IGMP Snooping 不会学习动态路由接口。

Report 报文

i 当组成员主机收到 Query 报文后，会应答 Report 报文。如果主机要加入某个组，也会主动发送 Report 报文。

i 锐捷产品对 IGMPv3 版本的 Report 报文，只对组信息部分进行处理，不支持对携带的源信息处理。

运行 IGMP Snooping 的设备在收到 Report 报文时，在此 VLAN 内执行如下动作：

- 将 Report 报文从所有路由连接接口转发出去。如果设备上启动了“Report 报文抑制”功能，则在一个 IGMP 查询周期内，只会将每组收到的第一个 Report 报文转发出去。
- 如果收到 Report 报文的端口是动态成员口，则重置其老化定时器。如果定时器超时，则该端口不再作为指定组的动态成员口。
- 如果收到 Report 报文的端口不是成员口，则将其作为动态成员口，并启动其老化定时器。如果定时器超时，则该端口不再作为指定组的动态成员口。

📌 Leave 报文

 如果主机要离开某个组，会主动发送 Leave 报文。

运行 IGMP Snooping 的设备在收到 Leave 报文时，在此 VLAN 内执行如下动作：

- 将 Leave 报文从所有路由连接接口转发出去。
- 如果收到 Leave 报文的端口是动态成员口，且配置了“快速离开”，端口将立即从指定组的 IGMP Snooping 转发表项中删除，该端口不再作为指定组的动态成员口。
- 如果收到 Leave 报文的端口是动态成员口，且未配置“快速离开”，则该端口状态保持不变

相关配置

📌 配置静态路由连接接口

使用 `ip igmp snooping vlan mrouter interface` 命令，可静态配置接口为路由连接接口。

📌 配置静态成员口

使用 `ip igmp snooping vlan static interface` 命令，可静态配置接口为成员口。

📌 启动 Report 报文抑制功能

缺省未启用 Report 报文抑制功能

使用 `ip igmp snooping suppression enable` 命令，可配置报文抑制功能。

启用 Report 报文抑制功能后，则在一个 IGMP 查询周期内，只会将每组收到的第一个 Report 报文转发出去。转发的 Report 报文的源 MAC 地址将替换成设备本身自己的源 MAC 地址。

📌 配置快速离开功能

缺省未启用快速离开功能

使用 `ip igmp snooping fast-leave enable` 命令，可配置快速离开功能。

📌 配置动态路由连接接口学习功能

缺省启用动态路由连接接口学习功能

使用 `no ip igmp snooping mrouter learn pim-dvmrp` 命令，可配置关闭动态路由连接接口学习功能。

使用 `no ip igmp snooping vlan vid mrouter learn pim-dvmrp` 命令，可配置关闭指定 VLAN 的动态路由连接接口学习功能。

📌 配置动态路由连接接口的老化定时器

缺省老化时间为 300 秒

端口收到 Query 报文时，启动或重置动态路由接口的老化定时器，如果没有配置老化定时器的时间，将使用 Query 报文携带的最长响应时间做为动态路由接口的老化时间。

使用 `ip igmp snooping dyn-mr-aging-time` 命令，可配置动态路由接口的老化时间。

配置动态成员口的老化定时器

缺省老化时间为 260 秒

设备收到 Query 报文时，重置动态成员口的老化定时器，老化时间为 Query 报文携带的最长响应时间。

端口收到 Report 报文时，启动或重置动态成员口的老化定时器，老化时间为动态成员口老化时间。

使用 `ip igmp snooping host-aging-time` 命令，可配置动态成员口的老化时间。

配置查询报文最大响应时间

缺省无配置，使用查询报文携带的最大响应时间

使用 `ip igmp snooping query-max-response-time` 命令，可配置查询报文最大响应时间。

5.3.2 IGMP Snooping工作模式

运行 IGMP Snooping 的设备在三种工作模式（IVGL、SVGL、IVGL-SVGL）下，为用户 VLAN 提供独立组播服务或公共组播服务。

工作原理

IVGL (Independent VLAN Group Learning)

在 IVGL 模式下，运行 IGMP Snooping 的设备为每一个用户 VLAN 提供独立组播服务。

独立组播服务：组播流只能在所属 VLAN 内转发，用户主机只能在所属 VLAN 内申请组播流。

SVGL (Shared VLAN Group Learning)

在 SVGL 模式下，运行 IGMP Snooping 的设备可以为用户 VLAN 提供公共组播服务。

公共组播服务只能应用在 Shared VLAN 和 Sub VLAN 中，使用 SVGL 组地址范围。Shared VLAN 中的 SVGL 组地址范围的组播流被转发到 Sub VLAN；Sub VLAN 中的用户主机向 Shared VLAN 申请 SVGL 组地址范围的组播流。

- 在 Shared VLAN 和 Sub VLAN 中，SVGL 组地址范围的组播流适用公共组播服务，其他组播流被丢弃。
- （除 Shared VLAN 和 Sub VLAN 以外的）其他 VLAN 适用独立组播服务。

i 因此，将用户 VLAN 设置为 Shared VLAN 或 Sub VLAN，则享受公共组播服务；将用户 VLAN 设置为其他 VLAN，则享受独立组播服务。

IVGL-SVGL

IVGL-SVGL 模式也称为混杂模式。在 IVGL-SVGL 模式下，运行 IGMP Snooping 的设备可以为用户 VLAN 同时提供公共组播服务和独立组播服务。

- 在 Shared VLAN 和 Sub VLAN 中，SVGL 组地址范围的组播流适用公共组播服务，其他组播流适用独立组播服务。

- (除 Shared VLAN 和 Sub VLAN 以外的) 其他 VLAN 适用独立组播服务。

 因此, 将用户 VLAN 设置为 Shared VLAN 或 Sub VLAN, 则同时享受公共组播服务和独立组播服务; 将用户 VLAN 设置为其他 VLAN, 则享受独立组播服务。

相关配置

▾ 启动 IGMP Snooping, 选择工作模式

缺省未启动 IGMP Snooping。

使用 `ip igmp snooping ivgl` 命令, 启动 IGMP Snooping, 且运行 IVGL 模式。

使用 `ip igmp snooping svgl` 命令, 启动 IGMP Snooping, 且运行 SVGL 模式。

使用 `ip igmp snooping ivgl-svgl` 命令, 启动 IGMP Snooping, 且运行 IVGL-SVGL 模式。

必须在启动 IGMP Snooping 的同时指定其工作模式。三种工作模式, 必选其一。

▾ 配置 Shared VLAN

缺省为 VLAN1。

使用 `ip igmp snooping svgl vlan` 命令, 指定某 VLAN 为 Shared VLAN。

在 SVGL 和 IVGL-SVGL 模式下, 只能有一个 VLAN 充当 Shared VLAN。

▾ 配置 Sub VLAN

缺省为除 Shared VLAN 外的所有 VLAN。

使用 `ip igmp snooping svgl subvlan` 命令, 指定某 VLAN 为 Sub VLAN。

在 SVGL 和 IVGL-SVGL 模式下, Sub VLAN 数量不限。

▾ 配置 SVGL 组地址范围

无缺省值。

使用 `ip igmp snooping svgl profile profile_num` 命令, 配置 SVGL 组地址范围 (引用 profile 定义组地址范围)。

 在 SVGL 和 IVGL-SVGL 模式下, 必须配置 SVGL 组地址范围。否则无法提供公共组播服务。

5.3.3 组播安全控制

运行 IGMP Snooping 的设备, 能够控制组播服务范围 and 负荷, 并能够有效的防范非法组播流。

工作原理

▾ 控制用户点播的组的地址范围

控制用户点播的组的地址范围, 可以为用户定制组播服务范围、保障运营商利益, 防范非法组播流。

预先使用 Profile 定义一个允许用户加入的组地址范围。

- 将 Profile 应用在 VLAN 上, 可以限制 VLAN 内允许用户加入的组。

- 将 Profile 应用在端口上，可以限制端口下允许用户加入的组。

组播预览

对于一些组播视频流，如果用户没有开通点播此视频流的权限，而服务提供商又想能够让用户进行预览，达到预览时间后，再停止组播流。要实现这种功能，需要设备能够提供基于用户的组播预览功能。

组播预览功能主要和组播权限控制相结合使用，如在视频应用中，管理员通过“配置 VLAN 过滤器”或“配置端口过滤器”控制某些收费频道，此时未申请开通的用户无法直接进行点播观看。此时用户存在一种需求，想先了解频道的内容，再决定是否进行付费观看，为此通过组播预览功能，设定某些组播组（及收费频道），用户点播时可以免费预览一段时间（如 1 分钟）。在预览时间过后，则不允许再点播观看。

控制用户同时点播的组的个数

如果用户同时点播的组播流过多，会对设备性能产生很大的压力。限制同时点播的组的个数，可以有效保障带宽。

- 可以在全局上限制同时加入的组的个数。
- 可以在端口上限制同时动态加入的组的个数。

控制组播流的入端口

通过启用“源端口检查”功能，严格限定组播流的入端口，防范非法组播流。

- 当“源端口检查”功能打开时，只有从路由连接口进入的组播流才是合法的；而从非路由连接口进入的组播流被视为是非合法的，将被丢弃。
- 当“源端口检查”功能关闭时，从任何端口进入的组播流均是合法的。

控制组播流的源 IP 地址

通过启用“源 IP 检查”功能，严格限定组播流的源 IP 地址，防范非法组播流。

源 IP 检查包含特定组源 IP 和默认源 IP 两种配置：

- 默认源 IP 检测（又称 source-check default-server）：指为所有的 VLAN 的所有组播组指定源 IP，只有和设置的源 IP 相同的组播流才是合法的
- 特定组源 IP 检测（又称 limit -ipmc）：指为特定 VLAN 的特定组播组指定源 IP，该 VLAN 内接收到的特定组播组的数据流，只有和指定的源 IP 相同，才认为是合法的，组播设备才会对该数据流进行转发，否则进行丢弃。

相关配置

限制允许用户加入的组

缺省情况下不对用户加入的组进行过滤

通过在接口或全局配置模式使用命令 **ip igmp snooping filter** 配置组播组过滤。

启用预览功能

缺省情况下不启用预览功能。

通过 **ip igmp snooping preview** 命令启用组播预览功能，并限制可进行组播预览的组范围。

通过 **ip igmp snooping preview interval** 命令可设置组播预览时间。

限制接口下允许同时点播的组的个数

缺省情况下不限制组个数。

通过在接口模式使用命令 **ip igmp snooping max-groups** 限制点播的组个数。

↘ 限制全局组播组的个数

缺省情况下全局允许加入的组播组个数为 65536。

通过在全局模式使用命令 **ip igmp snooping l2-entry-limit** 全局允许加入的组个数。

↘ 启用源端口检查

缺省情况下不启用源端口检查功能。

通过 **ip igmp snooping source-check port** 命令可启用源端口检查功能。

↘ 启用源 IP 检查

缺省情况下，关闭源 IP 检查。

- 使用 **ip igmp snooping source-check default-server address** 命令，启动源 IP 检查，指定默认源地址（适用于任意 VLAN 任意组）。
- （可选）使用 **ip igmp snooping limit-ipmc vlan vid address group-address server source-address** 命令，为特定 VLAN 特定组指定源地址（适用于特定 VLAN 特定组）。

必须先“启动源 IP 检查，指定默认源地址”，“为特定 VLAN 特定组指定源地址”才能生效。如果配置了“为特定 VLAN 特定组指定源地址”，则特定 VLAN 特定组的组播流依据此命令指定的源地址执行检查，其他组播流依据默认源地址执行检查。

5.3.4 Profile

Profile 用于定义允许或禁止用户点播的组地址范围，供其他功能引用。

工作原理

Profile 用来定义组地址范围。

配置 SVGL 模式时，引用 profile 来定义 SVGL 组地址范围。

配置端口过滤器时，引用 rofile 来定义端口下允许/禁止用户点播的组地址范围。

配置 VLAN 过滤器时，引用 rofile 来定义 VLAN 中允许/禁止用户点播的组地址范围。

配置预览功能时，引用 rofile 来定义提供预览的组地址范围。

配置相关

↘ 配置 profile

缺省情况：

- 创建 profile，默认为 deny 所有组范围。

配置方法：

- 使用 **ip igmp profile profile-number** 命令，创建一个 Profile。

- 使用 `range low-address high_address` 命令，定义一个组地址范围。一个 Profile 中配置多条 `range` 命令。
- （可选）使用 `permit`（或 `deny`）命令，允许（或禁止）用户点播。缺省为禁止。一个 Profile 中最多只能配置一条 `permit`（或 `deny`）命令。

5.3.5 对QinQ的处理

工作原理

当在设备开启 IGMP Snooping 和配置了 dot1q-tunnel 端口之后，对于从 dot1q-tunnel 端口收到的 IGMP 报文，IGMP Snooping 可以有以下的两种处理方式。

- 方式一：将组播表项建立在 IGMP 报文所在的 VLAN 上。在 IGMP 报文所在的 VLAN 上转发 IGMP 报文，称之为报文透传功能。比如：假设备开启 IGMP Snooping，设备的端口 A 为 dot1q-tunnel 端口，端口 A 的缺省 VLAN 为 VLAN 1，且端口 A 允许 VLAN 1 和 VLAN 10 的报文通过。当向端口 A 发送 VLAN 10 的组播请求报文时，IGMP Snooping 将建立 VLAN 10 的组播表项，并将组播请求报文转发到 VLAN 10 的路由连接口上。
- 方式二：将组播表项建立在 dot1q-tunnel 端口的缺省 VLAN 上。为组播报文添加一层 dot1q-tunnel 端口缺省 VLAN 的 VLAN Tag 之后，再将其在 dot1q-tunnel 端口的缺省 VLAN 内转发。比如：假设备开启 IGMP Snooping，设备的端口 A 为 dot1q-tunnel 端口，端口 A 的缺省 VLAN 为 VLAN 1，且端口 A 允许 VLAN 1 和 VLAN 10 的报文通过。当向端口 A 发送 VLAN 10 的组播请求报文时，IGMP Snooping 将建立 VLAN 1 的组播表项，并为组播请求报文添加一层 VLAN 1 的 VLAN Tag 之后，再将其转发到 VLAN 1 的路由连接口上。

配置相关

配置 QinQ 处理方式

缺省情况下，工作于方式二。

通过命令 `ip igmp snooping tunnel` 可启用设备工作于方式一。

5.3.6 IGMP查询器

在一个存在三层组播设备的网络中，由三层组播设备充当 IGMP 查询器。二层组播设备只需要监听 IGMP 报文，即可建立并维护转发表项，实现二层组播。

在一个没有三层组播设备的网络中，无法由三层组播设备充当 IGMP 查询器。为了使二层组播设备能够监听 IGMP 报文，必须在二层设备上配置 IGMP 查询器功能。二层组播设备即要充当 IGMP 查询器，又要监听 IGMP 报文，才能建立并维护转发表项，实现二层组播。

工作原理

IGMP 查询器功能由二层设备扮演 IGMP 路由查询的角色，定时发送 IGMP 查询报文，并对用户应答的 IGMP Report 报文进行侦听维护，建立二层组播的转发表项。IGMP 查询发送的查询报文的相关参数可由用户通过配置进行调整。

当设备接收到 PIM、DVMRP 协议报文时，认为网络中存在组播路由设备，为了避免影响组播路由设备的 IGMP 路由功能，设备的查询器功能将处于失效状态，由网络中的组播路由设备充当 IGMP 查询器功能。

当设备接收到其它设备的 IGMP 查询报文时，将进行 IGMP 查询器的竞选。

启动查询器

用户可配置在指定 VLAN 或所有 VLAN 上启用查询器功能。

全局查询器功能启用的情况下，指定 VLAN 的查询器功能才能生效。

指定查询器运行的 IGMP 版本

指定查询器发送的查询报文所使用的 IGMP 版本，可配置为 IGMPv1 或 IGMPv2 版本。

配置查询器的源 IP

配置查询器发送的查询报文所携带的源 IP 地址，可基于 VLAN 配置查询器发送的查询报文的源 IP 地址。

未配置查询器源 IP 的情况下，查询器功能不会生效。

配置查询器的查询间隔

配置全局查询器发送的查询报文的时间间隔，可基于 VLAN 配置不同 VLAN 的查询器的查询间隔。

配置查询报文的最大响应时间

配置查询器发送的查询报文中携带的最大响应时间，由于 IGMPv1 版本的协议不支持报文携带最大响应时间，所有该配置对于查询器运行 IGMPv1 版本时不生效。同时可基于不同 VLAN 的查询器配置不同的最大响应时间。

配置查询器的老化时间

当网络中还存在其它 IGMP 查询器设备时，当前设备将和其它设备进行查询器竞选，如果当前设备竞选失败处于非查询器状态时，将启动查询器老化定时器，在定时器超期后，认为网络中的其它查询器设备失效，当前设备重新恢复为查询器设备。

相关配置

启动查询器

缺省情况下，设备不允许查询器功能。

通过 `ip igmp snooping querier` 命令可控制全局的查询器功能。

通过 `ip igmp snooping vlan num querier` 命令可控制指定 VLAN 的查询器功能

指定查询器运行的 IGMP 版本

缺省情况下，查询器使用 IGMPv2 版本。

通过 `ip igmp snooping querier version` 命令可控制全局的查询器的版本。

通过 `ip igmp snooping vlan querier version` 命令可控制指定 VLAN 的查询器版本

配置查询器的源 IP

缺省情况下，查询器的源 IP 为 0

通过 `ip igmp snooping querier address` 命令可设置全局的查询器源 IP。

通过 `ip igmp snooping vlan querier address` 命令可设置指定 VLAN 的查询器源 IP

配置查询器的查询间隔

缺省情况下，查询器的查询间隔为 60 秒

通过 `ip igmp snooping querier query-interval` 命令可设置全局的查询器查询间隔。

通过 `ip igmp snooping vlan querier query-interval` 命令可设置指定 VLAN 的查询器查询间隔

配置查询报文的最大响应时间

缺省情况下，查询器的查询间隔为 10 秒

通过 `ip igmp snooping querier max-response-time` 命令可设置全局查询器的查询报文最大响应时间。

通过 `p igmp snooping vlan querier max-response-time` 命令可设置指定 VLAN 查询器的查询报文最大响应时间

配置查询器的老化时间

缺省情况下，查询器老化时间间隔为 125 秒

通过 `ip igmp snooping querier max-response-time` 命令可设置全局的查询器老化时间间隔。

通过 `p igmp snooping vlan querier max-response-time` 命令可设置指定 VLAN 的查询器老化时间间隔

5.4 配置详解

配置项	配置建议 & 相关命令	
配置IGMP Snooping基本功能 (IVGL模式)	 “IVGL 模式” “SVGL 模式” “IVGL-SVGL 模式” 必选其一。 本配置项用于启动 IGMP Snooping 功能，运行 IVGL 模式。	
	<code>ip igmp snooping ivgl</code>	全局启动 IGMP Snooping 功能，运行 IVGL 模式。
	<code>no ip igmp snooping vlan num</code>	在 VLAN 上关闭 IGMP Snooping 功能。
配置IGMP Snooping基本功能 (SVGL模式)	 “IVGL 模式” “SVGL 模式” “IVGL-SVGL 模式” 必选其一。 本配置项用于启动 IGMP Snooping 功能，运行 SVGL 模式。	
	<code>ip igmp snooping svgl</code>	启动 IGMP Snooping 功能，运行 SVGL 模式。
	<code>no ip igmp snooping vlan num</code>	在 VLAN 上关闭 IGMP Snooping 功能。
	<code>ip igmp snooping svgl profile profile_num</code>	指定 SVGL 组地址范围（引用 profile）。
	<code>ip igmp snooping svgl vlan</code>	指定 Shared VLAN。
	<code>ip igmp snooping svgl subvlan</code>	指定 Sub VLAN。
配置IGMP Snooping基本功能 (IVGL-SVGL模式)	 “IVGL 模式” “SVGL 模式” “IVGL-SVGL 模式” 必选其一。 本配置项用于启动 IGMP Snooping 功能，运行 IVGL-SVGL 模式。	
	<code>ip igmp snooping ivgl-svgl</code>	全局启动 IGMP Snooping 功能，运行 IVGL-SVGL 模式。
	<code>no ip igmp snooping vlan num</code>	在 VLAN 上关闭 IGMP Snooping 功能。
	<code>ip igmp snooping svgl profile profile_num</code>	指定 SVGL 组地址范围（引用 profile）。
	<code>ip igmp snooping svgl vlan</code>	指定 Shared VLAN
	<code>ip igmp snooping svgl subvlan</code>	指定 Sub VLAN

配置协议报文处理	 可选配置。用于调整协议报文处理相关配置。	
	ip igmp snooping vlan <i>vlan-id</i> mrouter interface <i>interface-id</i>	配置静态路由连接接口。
	p igmp snooping vlan <i>vid</i> static group-address interface <i>interface-type</i> <i>interface-number</i>	配置静态成员口。
	ip igmp snooping vlan <i>vlan-id</i> mrouter learn pim-dvmrp	启动动态学习路由连接接口的功能。
	ip igmp snooping dyn-mr-aging-time <i>time</i>	配置动态路由连接接口的老化时间。
	ip igmp snooping host-aging-time <i>time</i>	配置动态成员口老化时间。
	ip igmp snooping fast-leave enable	启动动态成员口“快速离开”功能。
	ip igmp snooping query-max-response-time <i>time</i>	配置 IGMP Query 报文的最大响应时间。
	ip igmp snooping suppression enable	启动 IGMP Report 报文抑制功能。
配置组播安全控制	 可选配置。用于对用户点播进行安全控制。	
	ip igmp snooping filter <i>profile-number</i>	在端口上限制允许用户加入的组的地址范围。
	ip igmp snooping vlan <i>num</i> filter <i>profile-number</i>	在 VLAN 上限制允许用户加入的组的地址范围。
	ip igmp snooping l2-entry-limit <i>number</i>	全局限制用户加入的组的个数。
	ip igmp snooping max-groups <i>number</i>	在端口上限制用户动态加入的组个数。
	ip igmp snooping source-check port	启动源端口检查，从路由连接接口进入的组播流为合法组播流。
	ip igmp snooping source-check default-server <i>address</i>	启动源 IP 检查。 任意组播流中，匹配指定源地址的组播流为合法组播流。
	ip igmp snooping limit-ipmc vlan <i>vid</i> address <i>group-address</i> server <i>source-address</i>	特定 VLAN-ID，组地址的组播流中，匹配指定源地址的组播流为合法组播流。
	ip igmp snooping preview <i>profile-number</i>	对指定地址范围的组启动预览功能（引用 profile）。
ip igmp snooping preview interval <i>num</i>	配置预览时间。	
配置Profile	 可选配置。用于定义一个允许/禁止用户主机加入的组地址范围。	
	ip igmp profile <i>profile-number</i>	创建一个 profile。
	range <i>low-address</i> <i>high_address</i>	指定组地址范围。
	permit	允许用户主机加入。
	deny	禁止用户主机加入。
配置对QinQ的处理	 可选配置。用于设置 QinQ 接口使用报文携带的 vid 进行组播报文转发。	
	ip igmp snooping tunnel	配置 QinQ 透传 IGMP 报文。
配置IGMP查询器	 可选配置。用于在没有三层组播设备的网络上提供 IGMP 查询器的功能。	

<code>ip igmp snooping querier</code>	全局启动查询器功能。
<code>ip igmp snooping vlan num querier</code>	在 VLAN 上启动查询器功能。
<code>ip igmp snooping querier version num</code>	全局配置查询器运行的 IGMP 版本。
<code>ip igmp snooping vlan num querier version num</code>	在 VLAN 上配置查询器运行的 IGMP 版本。
<code>ip igmp snooping querier address a.b.c.d</code>	全局配置查询器的源 IP 地址。
<code>ip igmp snooping vlan num querier address a.b.c.d</code>	在 VLAN 上配置查询器的源 IP 地址。
<code>ip igmp snooping querier query-interval num</code>	全局设置查询间隔。
<code>ip igmp snooping vlan num querier query-interval num</code>	在 VLAN 上设置查询间隔。
<code>ip igmp snooping querier max-response-time num</code>	全局设置查询报文的最大响应时间。
<code>ip igmp snooping vlan num querier max-response-time num</code>	在 VLAN 上设置查询报文的最大响应时间。
<code>ip igmp snooping querier timer expiry num</code>	全局配置查询器老化定时器。
<code>ip igmp snooping vlan num querier timer expiry num</code>	在 VLAN 上配置查询器老化定时器。

5.4.1 配置IGMP Snooping基本功能 (IVGL模式)

配置效果

- 启动 IGMP Snooping，实现二层组播。
- 每个 VLAN 享有独立的组播服务。

注意事项

- SVGL 模式与 IP 组播功能相冲突。如果设备上要求同时启动 IP 组播功能，则必须选择 IGMP Snooping 的 IVGL 模式。

配置方法

✎ 全局启动 IGMP Snooping 功能，运行 IVGL 模式

必须配置。

全局启动 IGMP Snooping 功能后，则所有 VLAN 上均启动了 IGMP Snooping 功能。

若无特殊要求，建议在所有连接用户主机的二层接入设备上全局启动 IGMP Snooping 功能。

✎ 在指定 VLAN 上关闭 IGMP Snooping 功能

可选配置。如果希望在某些 VLAN 上不启用 IGMP Snooping 功能，则需要配置此命令。

必须首先全局启动 IGMP Snooping 功能，才能在指定 VLAN 上关闭 IGMP Snooping 功能。

在 IVGL 模式下，每个 VLAN 享受独立的组播服务，关闭任何一个 VLAN 的组播服务，都不会影响其他 VLAN 的组播服务。

检验方法

- 使用 `show ip igmp snooping gda-table` 命令，查看 IGMP Snooping 转发表项，确认成员口仅包括连接组成员主机的端口。
- 使用 `show ip igmp snooping` 命令，查看 IGMP Snooping 基本信息，确认 IGMP Snooping 工作在 IVGL 模式下。

相关命令

全局启动 IGMP Snooping 功能，运行 IVGL 模式

- 【命令格式】 `ip igmp snooping ivgl`
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 配置此命令后，所有 VLAN 启动 IGMP Snooping 功能。
缺省情况下，IGMP Snooping 处于关闭状态。

在指定 VLAN 上关闭 IGMP Snooping 功能

- 【命令格式】 `no ip igmp snooping vlan num`
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 必须首先全局启动 IGMP Snooping，才能在部分 VLAN 上关闭 IGMP Snooping 功能。
在 IVGL 模式下，可以在任意 VLAN 上关闭 IGMP Snooping 功能。

查看 IGMP Snooping 转发表项

- 【命令格式】 `show ip igmp snooping gda-table`
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 确认成员口仅包括连接组成员主机的端口。

查看 IGMP Snooping 工作模式

- 【命令格式】 `show ip igmp snooping`
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果 IGMP Snooping 运行在 IVGL 模式下，则可以看到如下信息：

```
IGMP Snooping running mode: IVGL
```

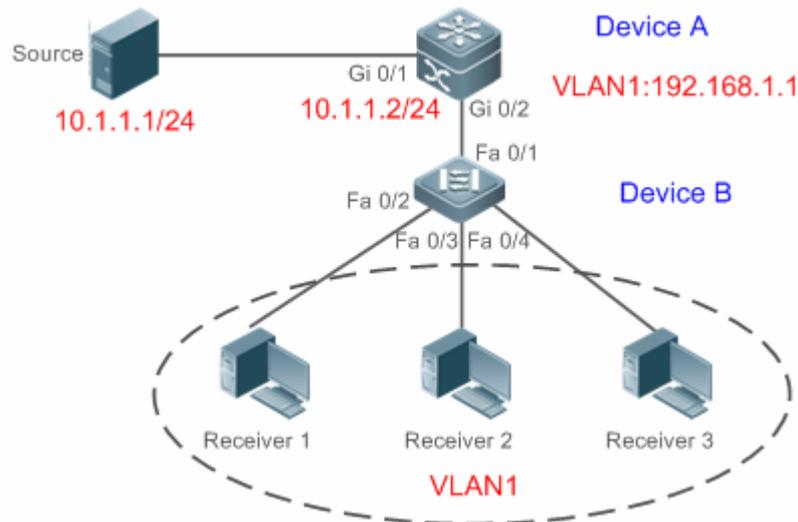
配置举例

i 以下配置举例，仅介绍与 IGMP Snooping 相关的配置。

在用户主机所在网段提供二层组播服务。

【网络环境】

图 5-5



A 作为组播路由设备，直连组播源。

B 作为二层接入设备，直连用户主机。

Receiver1、Receiver2、Receiver3 属于 VLAN1。

【配置方法】

- 在网络中配置 IP 地址、VLAN。（略）
- 在 A 上开启组播路由功能，并在三层接口上（Gi 0/1 和 VLAN 1）启动组播路由协议。
- 在 B 上开启 IGMP Snooping 功能，运行 IVGL 模式。

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip pim sparse-mode
A(config-if-VLAN 1)# exit
```

B

```
B# configure terminal
B(config)# ip igmp snooping ivgl
```

【检验方法】

使 Source (10.1.1.1) 向 G (229.1.1.1) 发送报文。使 Receiver1 加入 G。

- 确认 Receiver1 收到 (10.1.1.1, 229.1.1.1) 报文。
- 查看 B 上的 IGMP Snooping 转发表项，确认 (10.1.1.1, 229.1.1.1, 1) 的成员口仅包括 Fa 0/2。
- 确认 B 上的 IGMP Snooping 工作模式为 IVGL。

B

```
B# show ip igmp snooping gda-table
Multicast Switching Cache Table
  D: DYNAMIC
  S: STATIC
  M: MROUTE
(*, 224.1.1.1, 1):
  VLAN(1) 2 OPORTS:
    FastEthernet 0/1(M)
```

FastEthernet 0/2(D)

```
B# show ip igmp snooping
IGMP Snooping running mode: IVGL
IGMP Snooping L2-entry-limit: 65536
Source port check: Disable
Source ip check: Disable
IGMP Fast-Leave: Disable
IGMP Report suppress: Disable
IGMP Globle Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
Dynamic Host Aging Time : 260(Seconds)

vlan 1
-----
IGMP Snooping state: Enable
Multicast router learning mode: pim-dvmrp
IGMP Fast-Leave: Disabled
IGMP VLAN querier: Disable
IGMP VLAN Mode: STATIC
```

常见错误

- 未能正确选择 IGMP Snooping 的工作模式。

5.4.2 配置IGMP Snooping基本功能 (SVGL模式)

配置效果

- 启动 IGMP Snooping SVGL 模式，实现二层组播。
- 可共享 VLAN 的组播服务。

配置方法

▾ 全局启动 IGMP Snooping 功能，运行 SVGL 模式

必须配置。

全局启动 IGMP Snooping SVGL 功能。

配置 SVGL 关联的组范围。

指定 Shared VLAN

可选配置。默认情况下 VLAN 1 将作为 SVGL 的共享 VLAN，可通过该配置进行调整。

指定 SVGL 子 VLAN

可选配置。默认情况下所有 VLAN 都作为 SVGL 的子 VLAN，共享 Share VLAN 的组播服务，可通过该配置进行调整。

检验方法

- 使用 **show ip igmp snooping** 命令，查看 IGMP Snooping 基本信息，确认 IGMP Snooping 工作在 SVGL 模式下。
- 通过 **show ip igmp snooping gda-table** 命令，查看是否正常跨 VLAN 的组播表项。

相关命令

全局启动 IGMP Snooping 功能，运行 SVGL 模式

- 【命令格式】 **ip igmp snooping svgl**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 缺省情况下，IGMP Snooping 处于关闭状态。
配置完 SVGL 模式后，需要关联 SVGL 模式下的组范围

配置 SVGL 模式组范围

- 【命令格式】 **ip igmp snooping svgl profile profile_num**
- 【参数说明】 *profile_num* : 配置 SVGL 关联一条 profile
- 【命令模式】 全局模式
- 【使用指导】 缺省情况下，所有组都不属于 SVGL 模式范围

指定 Shared VLAN

- 【命令格式】 **ip igmp snooping svgl vlan vid**
- 【参数说明】 *vid* VLAN 编号
- 【命令模式】 接口模式
- 【使用指导】 默认情况下 VLAN 1 将作为 Shared VLAN

指定 SVGL 子 VLAN

- 【命令格式】 **ip igmp snooping svgl subvlan vid-range**
- 【参数说明】 *vid-range* VLAN ID 或者 VLAN ID 范围
- 【命令模式】 接口模式
- 【使用指导】 缺省除 Shared VLAN 外，所有 VLAN 都是子 VLAN

查看 IGMP Snooping 工作模式

- 【命令格式】 **show ip igmp snooping**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果 IGMP Snooping 运行在 SVGL 模式下，则可以看到如下信息：

IGMP Snooping running mode: SVGL

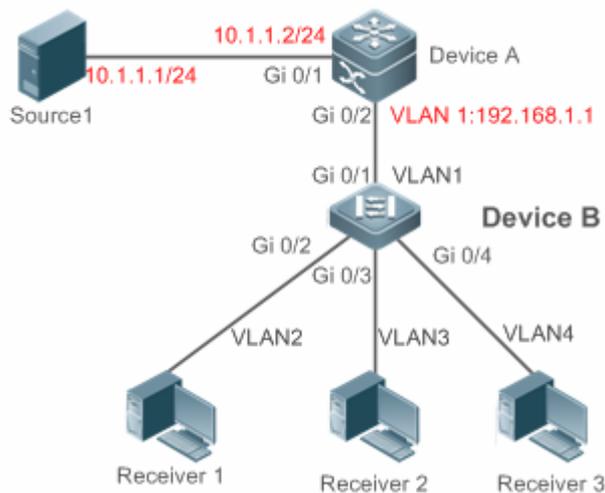
配置举例

i 以下配置举例，仅介绍与 IGMP Snooping 相关的配置。

在用户接入设备上实现 IGMP Snooping SVGL 功能

【网络环境】

图 5-6



A 作为组播路由设备，直连组播源。

B 作为二层接入设备，直连用户主机。

Receiver1 连接到 VLAN2、Receiver2 连接到 VLAN3、Receiver3 连接到 VLAN4。

【配置方法】

- 在网络中配置 IP 地址、VLAN。（略）
- 在 A 上开启组播路由功能，并在三层接口上（Gi 0/1 和 VLAN 1）启动组播路由协议。
- 在 B 上开启 IGMP Snooping 功能，运行 SVGL 模式。
- 在 B 上配置 IGMP Snooping SVGL 模式关联的组范围

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip pim sparse-mode
A(config-if-VLAN 1)# exit
```

B

```
B# configure terminal
B(config)#ip igmp profile 1
B(config-profile)#permit
B(config-profile)#range 224.1.1.1 238.1.1.1
B(config-profile)#exit
B(config)#ip igmp snooping svgl
B(config)#ip igmp snooping svgl profile 1
```

- 【检验方法】** 使 Source 发送源地址 (10.1.1.1) 目的地址 G (224.1.1.1) 的组播报文。使 Receiver1、Receiver2、Receiver3 都加入组 G。
- 确认 Receiver1、Receiver2、Receiver3 能够收到 (10.1.1.1, 224.1.1.1) 报文。
 - 查看 B 上的 IGMP Snooping 转发表项, 确认 (*, 224.1.1.1, 1) 的成员口包含 Gi0/2、Gi0/3、Gi0/4
 - 确认 B 上的 IGMP Snooping 工作模式为 SVGL。

```
B# show ip igmp snooping gda-table
Multicast Switching Cache Table
  D: DYNAMIC
  S: STATIC
  M: MROUTE
(*, 224.1.1.1, 1):
  VLAN(2) 1 OPORTS:
    GigabitEthernet 0/2(D)
  VLAN(3) 1 OPORTS:
    GigabitEthernet 0/3(D)
  VLAN(4) 1 OPORTS:
    GigabitEthernet 0/4(D)
```

```
B# show ip igmp snooping
IGMP Snooping running mode: SVGL
IGMP Snooping L2-entry-limit: 65536
SVGL vlan: 1
SVGL profile number: 1
Source port check: Disable
Source ip check: Disable
IGMP Fast-Leave: Disable
IGMP Report suppress: Disable
IGMP Globle Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
Dynamic Host Aging Time : 260(Seconds)
```

常见错误

- 未配置 SVGL 模式允许的组范围
- 发送的组播流不在 SVGL 允许的组范围内

5.4.3 配置IGMP Snooping基本功能 (IVGL-SVGL模式)

配置效果

- 启动 IGMP Snooping IVGL-SVGL 模式，实现二层组播。
- 处于 SVGL 模式组范围的组可共享 VLAN 的组播服务。
- 非 SVGL 模式组范围的组则运行在 IVGL 模式下

配置方法

全局启动 IGMP Snooping 功能，运行 IVGL-SVGL 模式

必须配置。

全局启动 IGMP Snooping IVGL-SVGL 功能。

配置 SVGL 关联的组范围。

指定 Shared VLAN

可选配置。默认情况下 VLAN 1 将作为 SVGL 的共享 VLAN，可通过该配置进行调整。

指定 SVGL 子 VLAN

可选配置。默认情况下所有 VLAN 都作为 SVGL 的子 VLAN，共享 Share VLAN 的组播服务，可通过该配置进行调整。

检验方法

- 使用 **show ip igmp snooping** 命令，查看 IGMP Snooping 基本信息，确认 IGMP Snooping 工作在 IVGL-SVGL 模式下。
- 通过 **show ip igmp snooping gda-table** 命令，查看属于 SVGL 范围的组播组是否能正常形成跨 VLAN 的组播表项。
- 通过 **show ip igmp snooping gda-table** 命令，查看不属于 SVGL 范围的组播组是否能正常形成 VLAN 内的组播表项。

相关命令

全局启动 IGMP Snooping 功能，运行 IVGL-SVGL 模式

【命令格式】 **ip igmp snooping ivgl-svgl**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 缺省情况下，IGMP Snooping 处于关闭状态。
配置完 IVGL-SVGL 模式后，需要关联 SVGL 模式下的组范围

配置 SVGL 模式组范围

【命令格式】 **ip igmp snooping svgl profile profile_num**

【参数说明】 *profile_num* : 配置 SVGL 关联一条 profile

【命令模式】 全局模式

【使用指导】 缺省情况下，所有组都不属于 SVGL 模式范围

指定 Shared VLAN

- 【命令格式】 **ip igmp snooping svgl vlan vid**
- 【参数说明】 vid VLAN 编号
- 【命令模式】 接口模式
- 【使用指导】 默认情况下 VLAN 1 将作为 Shared VLAN

指定 SVGL 子 VLAN

- 【命令格式】 **ip igmp snooping svgl subvlan vid-range**
- 【参数说明】 vid-range VLAN ID 或者 VLAN ID 范围
- 【命令模式】 接口模式
- 【使用指导】 缺省除 Shared VLAN 外，所有 VLAN 都是子 VLAN

查看 IGMP Snooping 工作模式

- 【命令格式】 **show ip igmp snooping**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果 IGMP Snooping 运行在 SVGL 模式下，则可以看到如下信息：

```
IGMP Snooping running mode: SVGL
```

查看 IGMP Snooping 工作模式

- 【命令格式】 **show ip igmp snooping**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果 IGMP Snooping 运行在 IVGL-SVGL 模式下，则可以看到如下信息：

```
IGMP Snooping running mode: IVGL-SVGL
```

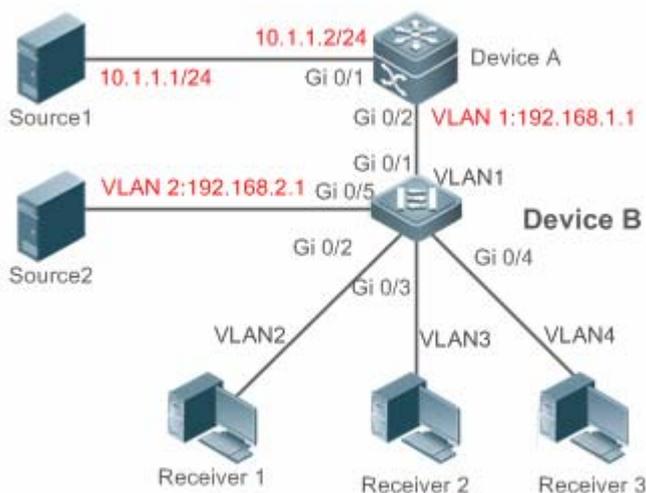
配置举例

i 以下配置举例，仅介绍与 IGMP Snooping 相关的配置。

在用户接入设备上实现 IGMP Snooping IVGL-SVGL 功能

【网络环境】

图 5-7



A 作为组播路由设备，直连组播源 Source1。

B 作为二层接入设备，直连用户主机，同时连接另一个组播源 Source2。

Receiver1 连接到 VLAN2、Receiver2 连接到 VLAN3、Receiver3 连接到 VLAN4。

【配置方法】

- 在网络中配置 IP 地址、VLAN。（略）
- 在 A 上开启组播路由功能，并在三层接口上（Gi 0/1 和 VLAN 1）启动组播路由协议。
- 在 B 上开启 IGMP Snooping 功能，运行 IVGL-SVGL 模式。
- 在 B 上配置 IGMP Snooping SVGL 模式关联的组范围

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip pim sparse-mode
A(config-if-VLAN 1)# exit
```

B

```
B# configure terminal
B(config)# ip igmp profile 1
B(config-profile)# permit
B(config-profile)# range 224.1.1.1 238.1.1.1
B(config-profile)# exit
B(config)# ip igmp snooping ivgl-svgl
B(config)# ip igmp snooping svgl profile 1
```

【检验方法】

使 Source1 发送源地址（10.1.1.1）目的地址 G（224.1.1.1）的组播报文。使 Receiver1、Receiver2、Receiver3 都加入组 G。

使 Source2 发送源地址（192.168.2.1）目的地址（239.1.1.1）的组播报文。使 Receiver1 加入组 239.1.1.1。

- 确认 Receiver1、Receiver2、Receiver3 能够收到（10.1.1.1，224.1.1.1）报文。
- 确认 Receiver1 能够收到（192.168.2.1，239.1.1.1）报文。
- 查看 B 上的 IGMP Snooping 转发表项，确认 表项（*，224.1.1.1，1）的成员口包含 Gi0/2、Gi0/3、Gi0/4；（*，239.1.1.1，1）的成员口为 Gi0/2
- 确认 B 上的 IGMP Snooping 工作模式为 IVGL-SVGL。

B

```
B# show ip igmp snooping gda-table
Multicast Switching Cache Table
  D: DYNAMIC
  S: STATIC
  M: MROUTE
(*, 224.1.1.1, 1):
  VLAN(2) 1 OPORTS:
    GigabitEthernet 0/2(D)
  VLAN(3) 1 OPORTS:
    GigabitEthernet 0/3(D)
  VLAN(4) 1 OPORTS:
```

```
GigabitEthernet 0/4(D)
(*, 239.1.1.1, 2):
VLAN(2) 1 OPORTS:
GigabitEthernet 0/2(D)
```

```
B# show ip igmp snooping
IGMP Snooping running mode: IVGL-SVGL
IGMP Snooping L2-entry-limit: 65536
SVGL vlan: 1
SVGL profile number: 0
Source port check: Disable
Source ip check: Disable
IGMP Fast-Leave: Disable
IGMP Report suppress: Disable
IGMP Globle Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
Dynamic Host Aging Time : 260(Seconds)
```

常见错误

- 未配置 SVGL 模式允许的组范围
- 发送的组播流不在 SVGL 允许的组范围内
- IVGL 模式的组播流组地址在 SVGL 范围内，导致无法进行 IVGL 转发

5.4.4 配置协议报文处理

配置效果

- 配置指定接口为静态路由接口，可接收所有组的组播流
- 配置指定接口为静态成员口，可接收指定组播组的组播流
- 配置 Report 报文抑制，在一个查询间隔内只会把第一个收到的特定 VLAN 和组的 Report 报文转发给路由接口，后继的 Report 报文将不继续向路由接口转发，这样可以减少网络中的报文数量。
- 配置快速离开，当设备某端口收到 Leave 报文时，直接从对应的转发表项的成员口中删除该端口
- 配置关闭动态路由接口学习，设备将不学习任何路由接口

- 可根据网络的负载以及组播路由设备的配置情况，调整路由接口和成员口的老化时间;调整查询报文最大响应时间；

注意事项

- 必须配置 IGMP Snooping 基本功能，相关配置功能才能生效

配置方法

配置静态路由接口

- 可选配置
- 如果要静态指定某个接口能够接收 VLAN 内的所有组播流，可在设备上执行此配置项

配置静态成员口

- 可选配置
- 如果要静态指定某个接口能够接收 VLAN 内的指定组播组的组播流，可在设备上执行此配置项

配置 Report 报文抑制

- 可选配置
- 当存在同一时间有大量的接收者需要接收同一个组播组的报文时，可通过在设备上配置 Report 报文抑制功能减少 Report 报文个数。

配置快速离开

- 可选配置
- 当接口下只存在一个接收者的时候，可通过在设备上配置快速离开，加速用户离开时的协议收敛。

配置关闭动态路由接口学习

- 可选配置
- 当组播流指需要在二层拓扑内进行转发时，不需要发送到三层路由设备，可进行此项配置。

配置动态路由接口的老化时间

- 可选配置
- 可根据连接网络负载情况进行时间调整

配置动态成员口的老化时间

- 可选配置
- 可根据连接的组播路由设备的 IGMP 查询报文发送间隔来调整该老化时间，一般设置为组播路由设备的 IGMP 查询报文发送时间间隔*2 + IGMP 最后查询响应时间

配置查询报文最大响应时间

- 可选配置
- 可根据连接网络负载情况进行时间调整

检验方法

- 使用 **show ip igmp snooping mrouter** 命令，查看已配置的静态路由连接口存在“S”标志。
- 使用 **show ip igmp snooping gda** 命令，查看已配置的静态成员口存在“S”标志。
- 使用 **show ip igmp snooping** 命令，查看 Report 报文抑制、快速离开、路由连接口学习、路由连接口老化时间、成员口老化时间、查询报文最大响应时间配置是否生效

相关命令

配置静态路由连接口

【命令格式】 **ip igmp snooping vlan vid mrouter interface interface-type interface-number**

【参数说明】 vid: VLAN 编号，取值范围 1-4094
interface-type interface-number 接口名称

【命令模式】 全局模式

【使用指导】 在 SVGL 模式下，如果未配置 Sub VLAN，则只有属于 Share VLAN 的静态路由连接口的配置可生效，其它情况可配置但不生效；如果配置了 Sub VLAN，则属于 Share VLAN 或者非 Sub VLAN 的静态路由连接口的配置可生效，其它情况可配置但不生效；
在 SVGL-IVGL 模式下，如果未配置 Sub VLAN，所有 VLAN 的静态路由连接口的配置均可生效；如果配置了 Sub VLAN，则属于 Share VLAN 或者非 Sub VLAN 的静态路由连接口的配置可生效，其它情况可配置但不生效；
在 IVGL 模式下，所有 VLAN 的静态路由连接口的配置均可生效。

配置静态成员口

【命令格式】 **ip igmp snooping vlan vid static group-address interface interface-type interface-number**

【参数说明】 vid: VLAN 编号，取值范围 1-4094
group-address: 组地址
interface-type interface-number: 接口名称

【命令模式】 全局模式

【使用指导】 默认情况下，无静态成员口配置

配置 Report 报文的抑制功能

【命令格式】 **ip igmp snooping suppression enable**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 当启用 Report 报文抑制功能后，在一个查询间隔内只会把第一个收到的特定 VLAN 和组的 Report 报文转发给路由连接口，后继的 Report 报文将不继续向路由连接口转发，这样可以减少网络中的报文数量。
只能抑制 IGMPv1/v2 的 Report 报文，对 IGMPv3 的 Report 报文无效。

配置端口快速离开功能

【命令格式】 **ip igmp snooping fast-leave enable**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 启用端口快速离开功能后，当设备某端口收到 Leave 报文时，直接从对应的转发表项的成员口中删除该端口。此后，当设备收到对应的特定组查询报文时，设备不再向该端口转发。其中，Leave 报文包括 IGMPv2 的 Leave 报文以及 IGMPv3 中 include 类型并且不带任何源地址的 Report 报文。

端口快速离开功能仅适用于设备一个端口下只连接一台主机的情况，可以节约带宽和资源。

配置动态学习路由连接口功能

- 【命令格式】 **ip igmp snooping [vlan vid] mrouter learn pim-dvmrp**
- 【参数说明】 **vlan vid** : 指定 vlan。缺省则适用于所有 VLAN
- 【命令模式】 全局模式
- 【使用指导】 路由连接口是开启 IGMP Snooping 的组播设备上与开启组播路由协议的组播邻居设备直接相连的端口。缺省情况下，启用动态学习路由连接口功能，设备自动侦听 IGMP Query/DVMRP/PIM Hello 报文，动态识别路由连接口。

配置动态路由连接口老化时间

- 【命令格式】 **ip igmp snooping dyn-mr-aging-time seconds**
- 【参数说明】 **seconds** : 动态路由口的老化时间，单位为秒，取值范围 1-3600
- 【命令模式】 全局模式
- 【使用指导】 如果某动态路由连接口在其老化时间超时前没有收到 IGMP 普遍组查询报文或 PIM Hello 报文，设备将该端口从路由连接口列表中删除。
在启用动态学习路由连接口功能的情况下，可以通过此命令调整动态路由口的老化时间。如果设置的老化时间过短可能会导致路由连接口频繁的添加和删除。

配置动态成员口老化时间

- 【命令格式】 **ip igmp snooping host-aging-time seconds**
- 【参数说明】 **seconds**: 老化时间
- 【命令模式】 全局模式
- 【使用指导】 动态成员端口老化时间是指当设备的某端口收到主机发送的加入某 IP 组播组的 IGMP 加入报文时，为这个动态成员端口设置的老化时间。
在收到 IGMP 加入报文后，会重置这个动态成员端口的老化定时器，定时器时间为 host-aging-time。如果定时器超时，则认为该端口下不存在接收组播报文的用户主机，组播设备就会将该端口从 IGMP Snooping 的成员口中删除。配置完该命令，后面收到的 IGMP 加入报文时设置的动态成员端口老化定时器的值为 host-aging-time。该配置在下次收到加入报文时生效，当前已启动的成员口的定时器不会被更新。

配置查询报文响应时间

- 【命令格式】 **ip igmp snooping query-max-response-time seconds**
- 【参数说明】 **seconds**: 响应时间
- 【命令模式】 全局模式
- 【使用指导】 在收到 IGMP 普通查询报文后，组播设备会重置所有动态成员口的老化定时器，定时器时间为 query-max-response-time。如果定时器超时，则认为该端口下不存在接收组播报文的用户主机，组播设备就会将该端口从 IGMP Snooping 的成员口中删除。
在收到 IGMP 特定组查询报文后，组播设备会重置该特定组的所有动态成员口的老化定时器，定时器时间为 query-max-response-time。如果定时器超时，则认为该端口下不存在接收组播报文的用户主机，组播设备就会将该端口从 IGMP Snooping 的成员口中删除。
该配置在下次收到查询报文时生效，当前已启动的定时器不会被更新。对于 IGMPv3 的特定组源查询报文，不做定时器的更新处理。

查看路由连接口

【命令格式】 **show ip igmp snooping mroute**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 如果路由连接口配置成功，则可以看到显示的接口信息中有“S”标志，如：

```
Ruijie(config)#show ip igmp snooping mrouter
Multicast Switching Mroute Port
  D: DYNAMIC
  S: STATIC
(*, *, 1):
  VLAN(1) 1 MROUTES:
    GigabitEthernet 0/1(S)
```

查看动态学习路由连接口

【命令格式】 **show ip igmp snooping**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 通过 **show ip igmp snooping** 命令，可查看动态路由连接口老化时间和学习状态：

```
Dynamic Mroute Aging Time : 300(Seconds)
Multicast router learning mode: pim-dvmrp
```

查看成员口

【命令格式】 **show ip igmp snooping gda-table**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 如果成员口配置成功，则可以看到显示的接口信息中有“S”标志，如：

```
Ruijie(config)#show ip igmp snooping gda-table
Multicast Switching Cache Table
  D: DYNAMIC
  S: STATIC
  M: MROUTE
(*, 224.1.1.1, 1):
  VLAN(1) 1 OPORTS:
    GigabitEthernet 0/1(S)
```

查看其它参数

【命令格式】 **show ip igmp snooping**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 通过 **show ip igmp snooping** 命令，可查看路由连接口老化时间、动态成员口老化时间、查询报文响应时间、Report 抑制和快速离开参数：

```
IGMP Fast-Leave: Enable
IGMP Report suppress: Enable
Query Max Response Time: 20(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
```

```
Dynamic Host Aging Time : 260(Seconds)
```

配置举例

静态路由连接口和静态成员口配置

- 【配置方法】
- 配置 IGMP Snooping 基本功能。略
 - 配置静态路由连接口和静态成员口

```
Ruijie# configure terminal
Ruijie(config)# ip igmp snooping vlan 1 mrouter interface GigabitEthernet 0/0
Ruijie(config)# ip igmp snooping vlan 1 static 224.1.1.1 interface GigabitEthernet 0/0
Ruijie(config)# end
```

- 【检验方法】 通过 show ip igmp snooping mrouter 和 show ip igmp snooping gda-table 查看静态配置是否成功

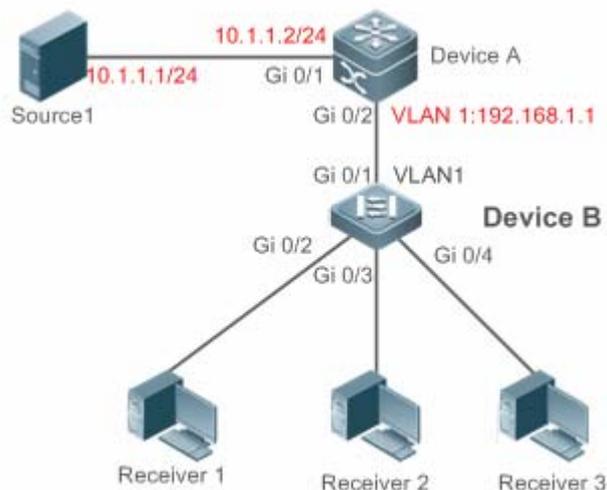
```
Ruijie#show ip igmp snooping mrouter
Multicast Switching Mroute Port
  D: DYNAMIC
  S: STATIC
(*, *, 1):
  VLAN(1) 1 MROUTES:
    GigabitEthernet 0/0(S)

Ruijie#show ip igmp snooping gda-table
Multicast Switching Cache Table
  D: DYNAMIC
  S: STATIC
  M: MROUTE
(*, 224.1.1.1, 1):
  VLAN(1) 1 OPORTS:
    GigabitEthernet 0/0(SM)
```

配置 Report 报文抑制

【网络环境】

图 5-8



A 作为组播路由设备，直连组播源 Source1。
B 作为二层接入设备，直连用户主机，同时连接另一个组播源 Source2。
Receiver1、Receiver2、Receiver3 连接到 VLAN1

- 【配置方法】
- 在网络中配置 IP 地址、VLAN。（略）
 - 在 A 上开启组播路由功能，并在三层接口上（Gi 0/1 和 VLAN 1）启动组播路由协议。
 - 在 B 上开启 IGMP Snooping 功能，运行 IVGL 模式。
 - 在 B 上配置 Report 报文抑制功能

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip pim sparse-mode
A(config-if-VLAN 1)# exit
```

B

```
B# configure terminal
B(config)# ip igmp snooping ivgl
B(config)# ip igmp snooping suppression enable
```

- 【检验方法】 Receiver1 和 Receiver2 都加入组 239.1.1.1，从设备 B 的 Gi0/1 接口上能够看到只转发出了一个组地址 239.1.1.1 的 IGMP Report 报文。

B

```
B# show ip igmp snooping
IGMP Snooping running mode: IVGL
IGMP Snooping L2-entry-limit: 65536
Source port check: Disable
Source ip check: Disable
IGMP Fast-Leave: Disable
IGMP Report suppress: Enable
IGMP Globle Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
Dynamic Host Aging Time : 260(Seconds)
```

↘ 其它参数配置

- 【配置方法】
- 配置 IGMP Snooping 基本功能。略
 - 配置快速离开
 - 配置关闭路由口学习
 - 配置路由连接口老化时间
 - 配置成员口老化时间

- 配置查询报文响应时间

```
Ruijie# configure terminal
Ruijie(config)# ip igmp snooping fast-leave enable
Ruijie(config)# no ip igmp snooping mrouter learn pim-dvmrp
Ruijie(config)# ip igmp snooping dyn-mr-aging-time 200
Ruijie(config)# ip igmp snooping host-aging-time 100
Ruijie(config)# ip igmp snooping query-max-response-time 60
Ruijie(config)# end
```

【检验方法】 通过 show ip igmp snooping 查看配置是否成功

```
Ruijie# show ip igmp snooping
IGMP Snooping running mode: IVGL
IGMP Snooping L2-entry-limit: 65536
Source port check: Disable
Source ip check: Disable
IGMP Fast-Leave: Enable
IGMP Report suppress: Enable
IGMP Globle Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
Query Max Response Time: 60(Seconds)
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 200(Seconds)
Dynamic Host Aging Time : 100(Seconds)
```

常见配置错误

- IGMP Snooping 基本功能没有配置，或配置失败。

5.4.5 配置组播安全控制

配置效果

- 控制用户能够访问的组播组范围
- 控制用户在指定的组在没有权限点播权限的情况下，可对组播频道进行预览
- 控制用户能够访问的组播组个数
- 限制用户只能接收从路由连接口进来的组播流，可用于防止终端用户发送非法组播流。
- 限制用户只能接收指定源 IP 的组流，防止非法组播流的影响

注意事项

- 必须配置 IGMP Snooping 基本功能。

配置方法

配置组地址范围限制

- 可选配置
- 如果需要限制接口上允许接收的组地址范围，通过配置接口上组地址限制实现。
- 如果需要限制某个 VLAN 允许接收的组地址范围，通过配置 VLAN 上组地址限制实现。

配置组播预览

- 可选配置
- 为不允许点播的组播组提供一段时间的预览，则需要配置组播预览功能

配置组地址个数限制

- 可选配置
- 如果需要限制某个接口允许接收的组个数，通过配置接口组个数限制实现。
- 如果需要限制全局允许接收的组个数，通过配置全局组个数限制实现。

配置源接口检查

- 可选配置
- 如果需要只允许接收从路由连接口进来的组播流，可进行该配置。

配置源 IP 检查

- 可选配置
- 指为所有的 VLAN 的所有组播组指定源 IP，只有和设置的源 IP 相同的组播流才是合法的。
- 指为特定 VLAN 的特定组播组指定源 IP，该 VLAN 内接收到的特定组播组的数据流，只有和指定的源 IP 相同，才认为是合法的，组播设备才会对该数据流进行转发，否则进行丢弃

检验方法

- 使用 `show ip igmp snooping interfaces` 命令，查看接口配置的组地址范围限制和组地址个数限制。
- 使用 `show ip igmp snooping vlan` 命令，查看 VLAN 配置的组地址限制范围。
- 使用 `show ip igmp snooping` 命令，查看配置的全局组地址个数限制、预览功能、源接口检查、源 IP 检查是否生效。

相关命令

配置接口组地址范围限制

- 【命令格式】 `ip igmp snooping filter profile-number`
- 【参数说明】 `profile-number` : profile 的编号
- 【命令模式】 接口模式
- 【使用指导】 -

配置 VLAN 组地址范围限制

- 【命令格式】 `ip igmp snooping vlan vid filter profile-number`

【参数说明】 *vid*: VLAN 编号, 取值范围 1-4094
profile-number: profile 的编号

【命令模式】 全局模式

【使用指导】 -

配置接口组个数限制

【命令格式】 **ip igmp snooping max-groups** *number*

【参数说明】 *number*: 组的个数的最大值

【命令模式】 接口模式

【使用指导】 该配置只对接口上动态加入的组进行统计, 不包含静态配置的组。接口下的组播组个数统计是基于接口所属的 VLAN 进行统计, 如接口同时属于 3 个 VLAN, 在每个 VLAN 上都接收到组播组 224.1.1.1 的请求, 即同时在 3 个 VLAN 上都接收到组播组 224.1.1.1 的点播请求, 此时该接口的组统计个数为 3 个, 不是 1 个

配置全局组个数限制

【命令格式】 **ip igmp snooping l2-entry-limit** *number*

【参数说明】 *number*: 组播组个数的最大值

【命令模式】 全局模式

【使用指导】 包含静态配置和动态学习的组播组

配置源端口检查

【命令格式】 **ip igmp snooping source-check port**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 启用源端口检查后, 若环境中不存在任何路由连接口, 设备接收到的组播数据流将被丢弃。

配置源 IP 检查

【命令格式】 **ip igmp snooping source-check default-server** *source-address*

【参数说明】 *source-address* 源 IP 地址

【命令模式】 全局模式

【使用指导】 -

配置特定组源 IP 检查

【命令格式】 **ip igmp snooping limit-ipmc vlan** *vid address group-address server source-address*

【参数说明】 *vid* vlan id

group-address 组地址

source-address 源地址

【命令模式】 全局模式

【使用指导】 -

配置预览功能

【命令格式】 **ip igmp snooping preview** *profile-number*

【参数说明】 *profile number* 范围为 1- 1024, 可进行预览的组范围

【命令模式】 全局模式

【使用指导】 -

配置预览时间

【命令格式】 **ip igmp snooping preview interval num**

【参数说明】 *num* : 配置预览时间, 范围为<1- 300>,缺省情况为 60sec

【命令模式】 全局模式

【使用指导】 对于用户无点播权限的组, 可在预览时间内接收组播组的数据, 在预览时间结束后, 无法再进行点播。需要等待 300 秒后, 才能够再次进行预览。

查看接口组地址范围限制

【命令格式】 **show ip igmp snooping interface**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 如果接口组地址限制配置成功, 则可以看到显示组范围 Profile, 如:

```
Ruijie#show ip igmp snooping interfaces gigabitEthernet 0/1
      Interface          Filter profile number      max-group
      -----
GigabitEthernet 0/1      1
```

查看 VLAN 组地址范围限制

【命令格式】 **show ip igmp snooping vlan**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 如果 VLAN 组地址限制配置成功, 则可以看到显示组范围 Profile, 如:

```
IGMP VLAN filter: 1
```

查看接口组个数限制

【命令格式】 **show ip igmp snooping interface**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 如果接口组地址限制配置成功, 则可以看到显示组范围 Profile, 如:

```
Ruijie#show ip igmp snooping interfaces gigabitEthernet 0/1
      Interface          Filter profile number      max-group
      -----
GigabitEthernet 0/1      1                          200
```

查看全局组个数限制

【命令格式】 **show ip igmp snooping vlan**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 如果 VLAN 组地址限制配置成功, 则可以看到显示组范围 Profile, 如:

```
IGMP Snooping L2-entry-limit: 65536
```

查看源端口检查

- 【命令格式】 **show ip igmp snooping vlan**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果源端口检查配置成功，则可以看到如下信息：

```
Source port check: Enable
```

查看源 IP 检查

- 【命令格式】 **show ip igmp snooping vlan**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果源 IP 检查配置成功，则可以看到如下信息：

```
Source ip check: Enable
```

查看预览功能

- 【命令格式】 **show ip igmp snooping**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果接口组地址限制配置成功，则可以看到显示预览已启用，如：

```
IGMP Preview: Enable
```

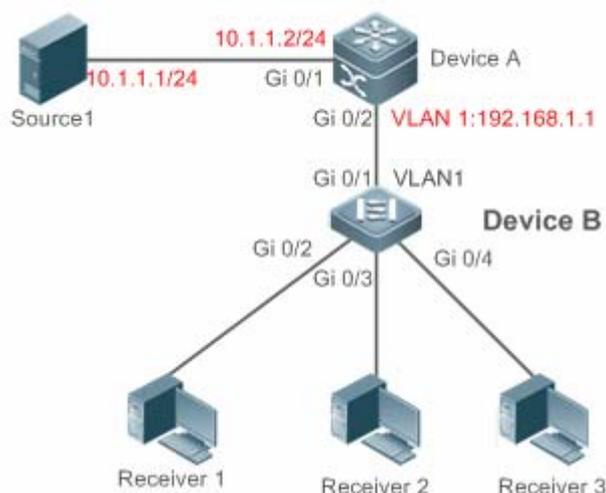
```
IGMP Preview group aging time : 60(Seconds)
```

配置举例

用户点播的组播组范围和个数限制

【网络环境】

图 5-9



A 作为组播路由设备，直连组播源 Source1。

B 作为二层接入设备，直连用户主机，同时连接另一个组播源 Source2。

Receiver1、Receiver2、Receiver3 连接到 VLAN1

通过配置限制 VLAN1 内的用户只允许接收 225.1.1.1~225.1.255.255 的组

通过配置限制 Receiver1 只能接收 225.1.1.1~225.1.1.255 的组，Receiver2 只能接收 225.1.2.1~255.1.2.255，Receiver3 只能接收 225.1.3.1~225.1.3.255

每个接口下同时最多只能加入 10 个组，全局最多接收 100 个组

【配置方法】

- 在网络中配置 IP 地址、VLAN。（略）
- 在 A 上开启组播路由功能，并在三层接口上（Gi 0/1 和 VLAN 1）启动组播路由协议。
- 在 B 上开启 IGMP Snooping 功能，运行 IVGL 模式。
- 在 B 上组播组地址范围限制和组播组个数限制

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip pim sparse-mode
A(config-if-VLAN 1)# exit
```

B

```
B# configure terminal
B(config)#ip igmp snooping ivgl
B(config)#ip igmp profile 1
B(config-profile)#permit
B(config-profile)#rang
B(config-profile)#range 225.1.1.1 225.1.255.255
B(config-profile)#exit
B(config)#ip igmp profile 2
B(config-profile)#permit
B(config-profile)#range 225.1.1.1 225.1.1.255
B(config-profile)#exit
B(config)#ip igmp profile 3
B(config-profile)#permit
B(config-profile)#range 225.1.2.1 225.1.2.255
B(config-profile)#exit
B(config)#ip igmp profile 4
B(config-profile)#permit
B(config-profile)#range
B(config-profile)#range 225.1.3.1 225.1.3.255
B(config-profile)#exit
B(config)#ip igmp snooping l2-entry-limit 100
B(config)#ip igmp snooping vlan 1 filter 1
B(config)#int gigabitEthernet 0/2
Ruijie(config-if-GigabitEthernet 0/0)#ip igmp snooping filter 2
Ruijie(config-if-GigabitEthernet 0/0)#ip igmp snooping max-groups 10
B(config)#int gigabitEthernet 0/3
Ruijie(config-if-GigabitEthernet 0/0)#ip igmp snooping filter 3
Ruijie(config-if-GigabitEthernet 0/0)#ip igmp snooping max-groups 10
B(config)#int gigabitEthernet 0/4
Ruijie(config-if-GigabitEthernet 0/0)#ip igmp snooping filter 4
```

```
Ruijie(config-if-GigabitEthernet 0/0)#ip igmp snooping max-groups 10
```

- 【检验方法】
- 通过 show ip igmp snooping interfaces 能够看到接口上配置的组范围过滤和组个数限制
 - 通过 show ip igmp snooping 能够看到全局组播组个数限制

B

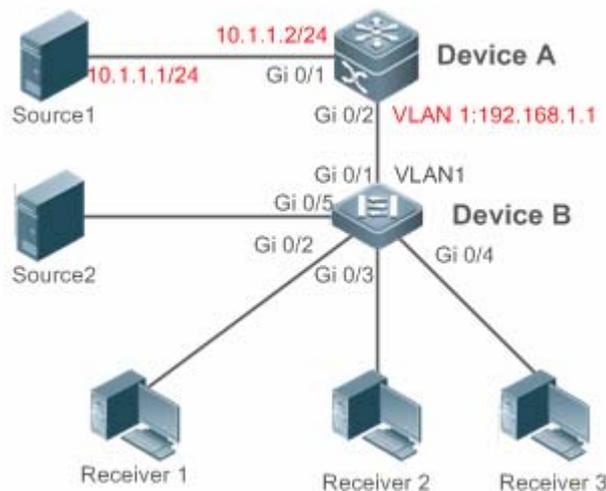
```
B#show ip igmp snooping interfaces
      Interface                Filter profile number      max-group
-----
GigabitEthernet 0/2           2                           10
GigabitEthernet 0/3           3                           10
GigabitEthernet 0/4           4                           10

B#show ip igmp snooping
IGMP Snooping running mode: IVGL
IGMP Snooping L2-entry-limit: 100
Source port check: Disable
Source ip check: Disable
IGMP Fast-Leave: Disable
IGMP Report suppress: Disable
IGMP Globle Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
Dynamic Host Aging Time : 260(Seconds)
```

源端口检查配置

【网络环境】

图 5-10



A 作为组播路由设备，直连组播源 Source1。

B 作为二层接入设备，直连用户主机，同时连接另一个组播源 Source2。

Receiver1、Receiver2、Receiver3 连接到 VLAN1

Source1 发送组播组 224.1.1.1 的组播流， Source2 发送组播组 225.1.1.1 的组播流

Receiver1 分别点播 224.1.1.1 和 225.1.1.1
启用源端口检查

- 【配置方法】**
- 在网络中配置 IP 地址、VLAN。（略）
 - 在 A 上开启组播路由功能，并在三层接口上（Gi 0/1 和 VLAN 1）启动组播路由协议。
 - 在 B 上开启 IGMP Snooping 功能，运行 IVGL 模式。
 - 在 B 上启用源端口检查功能

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip pim sparse-mode
A(config-if-VLAN 1)# exit
```

B

```
B# configure terminal
B(config)#ip igmp snooping ivgl
B(config)#ip igmp snooping source-check port
```

- 【检验方法】**
- 在设备 B 上通过 show ip igmp snooping mroute 能够看到接口 Gi0/1 被学习为路由连接口
 - Receiver1 能够点播到 224.1.1 的组播流，无法点播到 225.1.1.1 的组播流

B

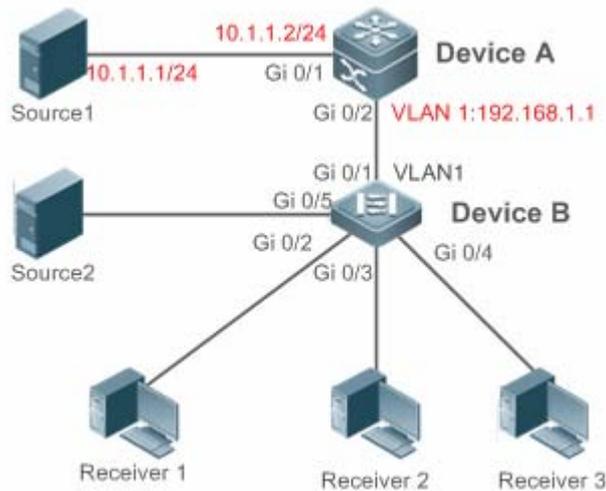
```
Multicast Switching Mroute Port
  D: DYNAMIC
  S: STATIC
(*, *, 1):
  VLAN(1) 1 MROUTES:
    GigabitEthernet 0/1(S)

B#show ip igmp snooping
IGMP Snooping L2-entry-limit: 100
Source port check: Enable
Source ip check: Disable
```

📌 源 IP 检查配置

【网络环境】

图 5-11



A 作为组播路由设备，直连组播源 Source1。

B 作为二层接入设备，直连用户主机，同时连接另一个组播源 Source2。

Receiver1、Receiver2、Receiver3 连接到 VLAN1

Source1 发送组播组 (10.1.1.1 , 224.1.1.1) 的组播流， Source2 发送组播组 (192.168.1.3 , 225.1.1.1) 和 (192.168.1.3 , 226.1.1.1) 的组播流

Receiver1 分别点播 224.1.1.1、225.1.1.1、226.1.1.1 三个组

配置默认源 IP 检查 IP 地址为 10.1.1.1

配置 limit-ipmc，组地址 225.1.1.1 的组播流，合法源地址为 192.168.1.3

【配置方法】

- 在网络中配置 IP 地址、VLAN。（略）
- 在 A 上开启组播路由功能，并在三层接口上（Gi 0/1 和 VLAN 1）启动组播路由协议。
- 在 B 上开启 IGMP Snooping 功能，运行 IVGL 模式。
- 在 B 上启用源端口检查功能

A

```
A# configure terminal
A(config)# ip multicast-routing
A(config)# interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)# ip pim sparse-mode
A(config-if-GigabitEthernet 0/1)# exit
A(config)# interface vlan 1
A(config-if-VLAN 1)# ip pim sparse-mode
A(config-if-VLAN 1)# exit
```

B

```
B# configure terminal
B(config)# ip igmp snooping ivgl
B(config)# ip igmp snooping source-check default-server 10.1.1.1
B(config)# ip igmp snooping limit-ipmc vlan 1 address 225.1.1.1 server 192.168.1.3
```

【检验方法】

- 在设备 B 上通过 show ip igmp snooping 命令可以查看到源 IP 检查已经打开
- Receiver1 能够点播到 224.1.1.1 和 225.1.1.1 的组播流，无法点播到 226.1.1.1 的组播流

B

```
B#show ip igmp snooping
IGMP Snooping running mode: IVGL
```

```
IGMP Snooping L2-entry-limit: 65536
Source port check: Disable
Source ip check: Enable
IGMP Fast-Leave: Disable
IGMP Report suppress: Disable
IGMP Globle Querier: Disable
IGMP Preview: Disable
IGMP Tunnel: Disable
IGMP Preview group aging time : 60(Seconds)
Dynamic Mroute Aging Time : 300(Seconds)
Dynamic Host Aging Time : 260(Seconds)
```

常见错误

- IGMP Snooping 基本功能没有配置，或配置失败。
- 没有学习到组播路由连接口，导致组播流无法接收。
- 源 IP 检查配置的 IP 地址和组播流 IP 地址不一致，导致无法接收到组播流。

5.4.6 配置Profile

配置效果

- 创建 IGMP 组过滤器模板

配置方法

▾ 创建 Profile

- 可选配置，创建 IGMP 组过滤模板

▾ 设置 Profile 组范围

- 可选配置，设置指定 Profile 的组地址范围

▾ 设置 Profile 过滤动作

- 可选配置，设置指定 Profile 的动作是 Permit 还是 Deny

检验方法

- 使用 `show running-config` 命令，查看配置是否生效。

相关命令

▾ 创建 Profile

- 【命令格式】 **ip igmp profile** *profile-number*
- 【参数说明】 *profile-number* profile 的编号
- 【命令模式】 全局模式
- 【使用指导】

设置 Profile 组范围

- 【命令格式】 **range** *low-ip-address* [*high-ip-address*]
- 【参数说明】 *low-ip-address* 指定范围的起始地址
high-ip-address 指定范围的结束地址。缺省则仅指定一个地址。
- 【命令模式】 profile 模式
- 【使用指导】 可同时配置多个 range 范围，若两个 range 之间存在交接，将自动对 range 进行合并

设置 Profile 过滤动作 deny

- 【命令格式】 **deny**
- 【参数说明】 -
- 【命令模式】 profile 模式
- 【使用指导】 如果只配置 deny 动作，未配置任何组范围，相当于未 deny 任何组，效果为 permit 所有组

设置 Profile 过滤动作 permit

- 【命令格式】 **permit**
- 【参数说明】 -
- 【命令模式】 profile 模式
- 【使用指导】 如果只配置 permit 动作，未配置任何组范围，相当于未 permit 任何组，效果为 deny 所有组

配置举例

创建 Profile 过滤模板

- 【配置方法】
 - 创建 Profile 过滤模板

```
B(config)#ip igmp profile 1
B(config-profile)#permit
B(config-profile)#range
B(config-profile)#range 224.1.1.1 235.1.1.1
B(config-profile)#
```

- 【检验方法】 通过 show running-config 查看配置是否成功

```
ip igmp profile 1
  permit
  range 224.1.1.1 235.1.1.1
!
```

常见配置错误

- IGMP Snooping 基本功能没有配置，或配置失败。

- 配置了 permit 的动作，未配置任何组地址范围，导致效果为 deny 所有组

5.4.7 配置对QinQ的处理

配置效果

- 将组播表项建立在 IGMP 报文所在的 VLAN 上。在 IGMP 报文所在的 VLAN 上转发 IGMP 报文，实现报文透传功能

注意事项

- 必须配置 IGMP Snooping 基本功能。

配置方法

配置 QinQ 报文透传

- 如果 QinQ 接口需要在报文携带的 VID 的 VLAN 上进行组播报文的转发，需进行 QinQ 报文透传配置

检验方法

- 使用 `show ip igmp snooping` 命令，查看配置是否生效。

相关命令

配置 QinQ 报文透传

- 【命令格式】 `ip igmp snooping tunnel`
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 启用 QinQ 透传 IGMP 报文的的功能。

查看 QinQ 配置

- 【命令格式】 `show ip igmp snooping`
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果 QinQ 配置成功，则可以看到如下信息：

```
IGMP Tunnel: Enable
```

配置举例

配置 QinQ 报文透传

- 【配置方法】
 - 配置 IGMP Snooping 基本功能。略
 - 配置启用 QinQ 报文透传功能

```
Ruijie# configure terminal
Ruijie(config)# ip igmp snooping tunnel
Ruijie(config)# Ruijie(config)# end
```

【检验方法】 通过 show ip igmp snooping 查看配置是否成功

```
IGMP Tunnel: Enable
```

常见错误

- IGMP Snooping 基本功能没有配置，或配置失败。

5.4.8 配置IGMP查询器

配置效果

- 本设备将充当 IGMP 查询器功能，定时发送 IGMP 查询报文，收集用户点播信息。

注意事项

- 必须配置 IGMP Snooping 基本功能。

配置方法

启动查询器

- 可选配置，要启用 IGMP 查询器，需要在全局上和指定 VLAN 上启用查询器功能
- 可选配置，用户可单独关闭某个 VLAN 的查询器功能。

配置查询器的源 IP

- 可选配置，配置查询器发送的查询报文所携带的源 IP 地址，可基于 VLAN 配置查询器发送的查询报文的源 IP 地址。
- 启动查询器功能后，必须为查询器指定一个源 IP 地址，否则查询器功能不生效。

配置查询报文的最大响应时间

- 可选配置，可调整 IGMP 查询报文中携带的最大响应时间值。由于 IGMPv1 版本的协议不支持报文携带最大响应时间，所有该配置对于查询器运行 IGMPv1 版本时不生效

配置查询器的查询间隔

- 可选配置，可调整 IGMP 查询器查询报文的发送间隔

配置查询器的老化时间

- 可选配置，可设置网络中其它 IGMP 查询器的老化时间

指定查询器运行的 IGMP 版本

- 可选配置，可设定 IGMP 查询的版本，默认使用 IGMP v2 版本

检验方法

- 使用 `show ip igmp snooping querier detail` 命令，查看配置是否生效。

相关命令

配置 IGMP 查询器功能

- 【命令格式】 `ip igmp snooping [vlan vid] querier`
- 【参数说明】 `vlan vid`：指定 VLAN。缺省则适用于所有 VLAN。
- 【命令模式】 全局模式
- 【使用指导】 当全局启动查询器功能后，再在 VLAN 上启用查询器功能，VLAN 上的查询器功能才能生效。如果在全局上关闭了查询器功能，所有 VLAN 上的查询器功能将全部关闭。

配置 IGMP 查询器源 IP

- 【命令格式】 `ip igmp snooping [vlan vid] querier address a.b.c.d`
- 【参数说明】 `vlan vid`：指定 VLAN。缺省则适用于所有 VLAN。
`a.b.c.d`：源 IP 地址
- 【命令模式】 全局模式
- 【使用指导】 启动查询器功能后，必须为查询器指定一个源 IP 地址，否则查询器功能不生效。如果 VLAN 上指定了查询器的源 IP 地址，优先使用 VLAN 上的配置。

配置最大响应时间

- 【命令格式】 `ip igmp snooping [vlan vid] querier max-response-time seconds`
- 【参数说明】 `vlan vid`：指定 VLAN。缺省则适用于所有 VLAN。
`seconds`：最大响应时间。单位为秒，取值范围 1-25
- 【命令模式】 全局模式
- 【使用指导】 如果 VLAN 上指定了查询间隔，优先使用 VLAN 上的配置。

配置查询间隔

- 【命令格式】 `ip igmp snooping [vlan vid] querier address a.b.c.d`
- 【参数说明】 `vlan vid`：指定 VLAN。缺省则适用于所有 VLAN。
`seconds`：查询间隔。单位为秒，取值范围 1-18000
- 【命令模式】 全局模式
- 【使用指导】 如果 VLAN 上指定了查询间隔，优先使用 VLAN 上的配置。

配置查询器超时时间

- 【命令格式】 `ip igmp snooping [vlan vid] querier timer expiry seconds`
- 【参数说明】 `vlan vid`：指定 VLAN。缺省则适用于所有 VLAN。
`seconds`：超时时间。单位为秒，取值范围 60-300
- 【命令模式】 全局模式
- 【使用指导】 启用了查询器功能后，也可能在选举中落败。如果落败者在“查询器超时时间”内没有收到当前查询器发出的查询报文，则认为当前查询器失效，发起下一轮选举。

如果相应 VLAN 上已经配置了查询器超时时间，优先使用 VLAN 上的配置值。

配置查询器版本

- 【命令格式】 **ip igmp snooping [vlan vid] querier version 1**
- 【参数说明】 **vlan vid** : 指定 VLAN。缺省则适用于所有 VLAN。
- 【命令模式】 全局模式
- 【使用指导】 查询器支持 IGMPv1 或 IGMPv2，默认情况下使用 IGMPv2，可使用命令设置查询器运行 IGMPv1 如果 VLAN 上已经配置了查询器运行的 IGMP 版本，优先使用 VLAN 上的配置。

查看 IGMP 查询器配置

- 【命令格式】 **show ip igmp snooping querier detail**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 如果 QinQ 配置成功，则可以看到如下信息：

```
Ruijie(config)#show ip igmp snooping querier detail
Vlan      IP Address      IGMP Version      Port
-----
Global IGMP switch querier status
-----
admin state           : Enable
admin version         : 2
source IP address     : 1.1.1.1
query-interval (sec) : 60
max-response-time (sec) : 10
querier-timeout (sec) : 125

Vlan 1:  IGMP switch querier status
-----
admin state           : Disable
admin version         : 2
source IP address     : 1.1.1.1
query-interval (sec) : 60
max-response-time (sec) : 10
querier-timeout (sec) : 125
operational state     : Disable
operational version   : 2
```

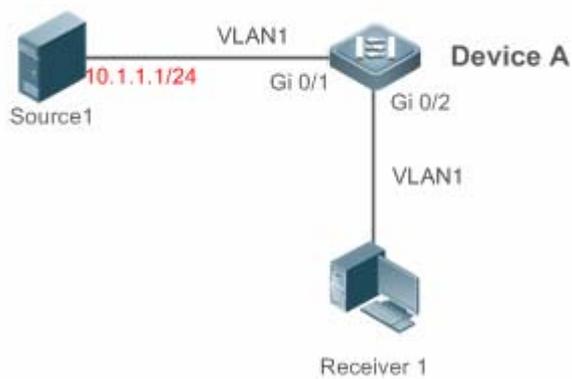
配置举例

i 以下配置举例，仅介绍与 IGMP Snooping 相关的配置。

启用 IGMP 查询器功能

【网络环境】

图 5-12



网络部署中组播流只需要在二层内进行转发，且网络中无三层组播功能的设备。
A 作为组播二层设备连接组播源和组播接收者。

【配置方法】

- 在 A 上开启 IGMP Snooping 功能，运行 IVGL 模式。
- 在 A 上配置 VLAN 1 的 IGMP 查询器功能

A

```
A(config)#ip igmp snooping ivgl
A(config)#ip igmp snooping querier
A(config)#ip igmp snooping querier address 10.1.1.1
A(config)#ip igmp snooping vlan 1 querier
```

【检验方法】

通过 show ip igmp snooping querier 命令，查看 VLAN 1 的查询器已生效

A

```
A(config)#show ip igmp snooping querier
Vlan      IP Address      IGMP Version      Port
-----
1         10.1.1.1        2                  switch

A(config)#show ip igmp snooping querier vlan 1

Vlan 1:  IGMP switch querier status
-----
elected querier is 10.1.1.1      (this switch querier)
-----

admin state           : Enable
admin version         : 2
source IP address     : 10.1.1.1
query-interval (sec)  : 60
max-response-time (sec) : 10
querier-timeout (sec) : 125
operational state     : Querier
operational version   : 2
```

常见错误

- 未配置查询器源 IP，导致查询器不生效。

5.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除 IGMP Snooping 的统计信息。	clear ip igmp snooping statistics
清除动态路由口、成员口。	clear ip igmp snooping gda-table

查看运行情况

作用	命令
查看 IGMP Snooping 基础配置。	show ip igmp snooping [vlan <i>vlan-id</i>]
查看 IGMP Snooping 统计信息。	show ip igmp snooping statistics [vlan <i>vlan-id</i>]
查看路由连接口。	show ip igmp snooping mrouter
查看 IGMP Snooping 转发表项。	show ip igmp snooping gda-table
查看 Profile。	show ip igmp profile [<i>profile-number</i>]
查看 IGMP Snooping 在端口上的配置。	show ip igmp snooping interface <i>interface-name</i>
查看 IGMP 查询器。	show ip igmp snooping querier [detail]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开与 IGMP Snooping 所有调试开关。	debug igmp-snp
打开 IGMP Snooping 事件的调试开关。	debug igmp-snp event
打开 IGMP Snooping 报文的调试开关。	debug igmp-snp packet
打开 IGMP Snooping 与 MSF 通信的调试开关。	debug igmp-snp msf
打开 IGMP Snooping 警告的调试开关。	debug igmp-snp warning



配置指南-安全

本分册介绍安全配置指南相关内容，包括以下章节：

1. AAA
2. RADIUS
3. TACACS+
4. 802.1x
5. Web 认证
6. SCC
7. 全局 IP+MAC 绑定
8. PASSWORD-POLICY
9. 端口安全
10. STORM CONTROL
11. SSH
12. URPF
13. CPP
14. DHCP Snooping
15. APP Check
16. 动态 ARP 检测
17. IP Source Guard
18. NFPP
19. DOS 保护

1 AAA

1.1 概述

AAA 是 Authentication Authorization and Accounting (认证、授权和记账) 的简称，它提供了对认证、授权和记账功能进行配置的一致性框架，锐捷网络设备产品支持使用 AAA。

AAA 以模块方式提供以下服务：

认证：验证用户是否可获得访问权，可选择使用 RADIUS 协议、TACACS+协议或 Local (本地) 等。身份认证是在允许用户访问网络和网络服务之前对其身份进行识别的一种方法。

授权：授权用户可使用哪些服务。AAA 授权通过定义一系列的属性对来实现，这些属性对描述了用户被授权执行的操作。这些属性对可以存放在网络设备上，也可以远程存放在安全服务器上。

记账：记录用户使用网络资源的情况。当 AAA 记账被启用时，网络设备便开始以统计记录的方式向安全服务器发送用户使用网络资源的情况。每个记账记录都是以属性对的方式组成，并存放在安全服务器上，这些记录可以通过专门软件进行读取分析，从而实现对用户使用情况网络资源的情况进行记账、统计、跟踪。

尽管 AAA 是最主要的访问控制方法，锐捷产品同时也提供了在 AAA 范围之外的简单控制访问，如本地用户名身份认证、线路密码身份认证等。不同之处在于它们提供对网络保护程度不一样，AAA 提供更高级别的安全保护。

使用 AAA 有以下优点：

- 灵活性和可控制性强
- 可扩充性
- 标准化认证
- 多个备用系统

协议规范

- 暂无相应规范

1.2 典型应用

典型应用	场景描述
无域环境下的认证、授权、记账	所有用户处于同一个域，进行认证、授权、记账
多域环境下的认证、授权、记账	处于不同域的用户，采用不同的方法进行认证、授权、记账

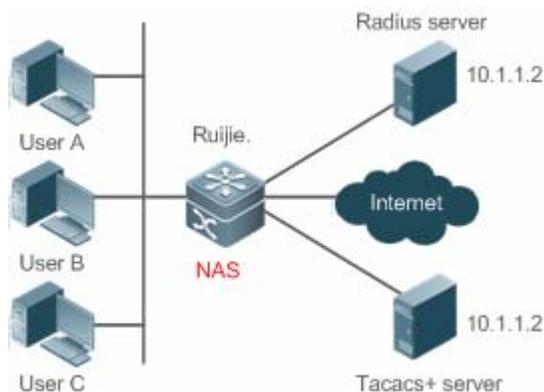
1.2.1 无域环境下的认证、授权、记账

应用场景

在图 1-1 所示的网络应用中，为了更好地对网络访问控制器设备（NAS，以下简称网络设备）进行安全管理，需要满足如下应用要求：

1. 不同的管理人员有各自的用户账号，其用户名和口令不能共享，便于帐号管理和防止泄漏。
2. 对网络设备的访问需经过认证，用户认证的实现方式可以分为本地认证和集中认证，应采用集中认证和本地认证相结合的方式，集中认证为主用、本地认证为备用。在集中认证过程中，要求先通过 RADIUS 服务器认证，若无响应再转本地认证。
3. 在认证时，不同的用户可以被限制只能访问特定的网络设备。
4. 对用户进行分权限管理：把网络管理用户分为超级用户和普通用户。其中，超级用户对网络设备拥有查看和配置的权限，普通用户对网络设备只拥有特定的查看权限。
5. 服务器端可将用户的认证信息、授权信息和网络行为记录在服务器中，以供日后查看和审计（本例采用 TACACS+ 进行记账）。

图 1-1



【注释】 UserA，UserB，UserC 直接或者通过网络和 NAS 相连接。

NAS 通常为接入交换机或者汇聚交换机。

RADIUS 服务器可以是 Windows 2000/2003 Server (IAS)、UNIX 系统所带组件，也可以是一些厂商提供的专用服务器软件。

TACACS+ 服务器可以是一些厂商提供的专用的服务器软件。

功能部属

- 在 NAS 上启用 AAA
- 在 NAS 上配置安全服务器
- 在 NAS 上配置本地用户

- 在 NAS 上配置认证
- 在 NAS 上配置授权
- 在 NAS 上配置记账

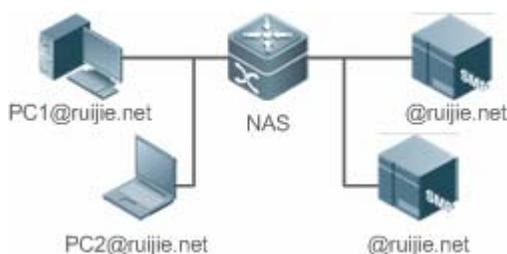
1.2.2 多域环境下的认证、授权、记账

应用场景

通过配置网络访问控制器设备实现基于域名的 AAA 服务，包括认证、授权、记账功能：

- 使用 802.1x 客户端进行登录认证，使用用户名为 PC1@ruijie.net 或 PC2@ruijie.com.cn，再输入正确的密码进行认证就可认证成功。
- 对用户进行分权限管理：把网络管理用户分为超级用户和普通用户。其中，超级用户对网络设备拥有查看和配置的权限，普通用户对网络设备只拥有特定的查看权限。
- 认证服务器端可将用户的认证信息、授权信息和网络行为记录在服务器中，以供日后查看和审计。

图 1-2



【注释】 PC1@ruijie.net，PC2@ruijie.com.cn 直接或者通过网络和 NAS 相连接。
NAS 通常为接入交换机或者汇聚交换机。
SAM 为锐捷公司提供的通用 RADIUS 服务器。

功能部属

- 在 NAS 上启用 AAA
- 在 NAS 上配置安全服务器
- 在 NAS 上配置本地用户
- 在 NAS 上定义 AAA 服务的方法列表
- 在 NAS 上打开基于域名的 AAA 服务开关
- 在 NAS 上创建域并配置域属性集

1.3 功能详解

基本概念

本地认证、远程服务器认证

对用户进行认证时，如果使用 NAS 上的用户数据库进行密码校验，就称为本地认证。

对用户进行认证时，如果使用远程服务器上的用户数据库进行密码校验，就称为远程服务器认证。目前，远程服务器认证主要是 RADIUS 服务器认证和 TACACS+服务器认证。

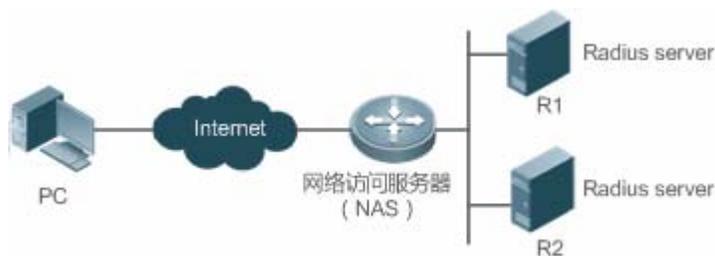
方法列表

由于对用户进行认证、授权和记账可以使用不同的安全方法，您需要使用方法列表定义一个使用不同方法对用户进行认证、授权和记账的前后顺序。方法列表可以定义一个或多个安全协议，这样可以确保在第一个方法失败时，有备用系统可用。锐捷产品使用方法列表中列出的第一个方法时，如果该方法无应答，则选择方法列表中的下一个方法。这个过程一直持续下去，直到与列出的某种安全方法成功地实现通信或用完方法列表。如果用完方法列表而还没有成功实现通信，则该安全功能宣告失败。

方法列表仅是定义将要被依次查询的、并用于认证用户身份的一系列安全方法。方法列表使您能够指定一个或多个用于身份认证的安全协议，这样确保在第一种方法失败的情况下，可以使用身份认证备份系统。我司产品使用第一种方法认证用户的身份，如果该方法无应答，将选择方法列表中的下一种方法。这个过程一直持续下去，直到与列出的某种身份认证方法成功地实现通信或用完方法列表。如果用完方法列表而还没有成功实现通信，则身份认证宣告失败。

⚠ 只有在前一种方法没有应答的情况下，锐捷产品才会尝试下一种方法。例如在身份认证过程中，某种方法拒绝了用户访问，则身份认证过程结束，不再尝试其他的身份认证方法。

图 1-3



上图说明了一个典型的 AAA 网络配置，包含两台安全服务器：R1 和 R2 是 RADIUS 服务器。以及一台网络访问服务器 (NAS)，可以作为 RADIUS 客户端。

假设系统管理员已定义了一个方法列表，在这个列表中，R1 首先被用来获取身份信息，然后是 R2，最后是访问服务器上的本地用户名数据库。如果一个远程 PC 用户试图拨号进入网络，网络访问服务器首先向 R1 查询身份认证信息，假如用户通过了 R1 的身份认证，R1 将向网络访问服务器发出一个 ACCEPT 应答，这样用户即获准访问网络。如果 R1 返回的是 REJECT 应答，则拒绝用户访问网络，断开连接。如果 R1 无应答，网络访问服务器就将它看作 TIMEOUT，并向 R2 查询身份认证信息。这个过程会一直在余下的指定方法中持续下去，直到用户通过身份认证、被拒绝或对话被中止。如果所有的方法返回 TIMEOUT，则认证失败，连接将被断开。

- ❗ REJECT 应答不同于 TIMEOUT 应答。REJECT 意味着用户不符合可用身份认证数据库中包含的标准，从而未能通过身份认证，访问请求被拒绝。TIMEOUT 则意味着安全服务器对身份认证查询未作应答，当检测到一个 TIMEOUT 时，AAA 选择身份认证方法列表中定义的下一个身份认证方法将继续进行身份认证过程。
- ❗ 在本文中，与 AAA 安全服务器相关的认证、授权和记账配置，均以 RADIUS 为例，而与 TACACS+ 有关的内容请另外参考“配置 TACACS+”。

AAA 服务器组

定义一个 AAA 服务器组，用于把一个或几个同一类型的服务器划分为同一组。配置方法列表时，引用该服务器组，则使用该方法列表进行认证、授权、记账操作时，首先向被引用服务器组中的服务器发起请求。

功能特性

功能特性	作用
AAA认证	验证是否允许用户接入网络
AAA授权	定义用户可以使用哪些服务或拥有哪些权限
AAA记账	记录用户使用网络资源的情况
AAA多域	针对不同域的 802.1x 用户，创建认证、授权和记账方案。

1.3.1 AAA 认证

在 AAA 中，认证、授权和计费是三个独立的业务过程。认证是用来验证用户是否可以获得访问权，其职责是完成各接入或服务请求的用户名、密码和用户信息的交互认证过程。在 AAA 中，可以只使用认证，而不使用授权或计费。

- ❗ 要配置 AAA 身份认证，首先得定义一个身份认证方法的命名列表，然后各个应用使用已定义列表进行认证。方法列表定义了身份认证的类型和执行顺序。对于已定义的身份认证方法，必须有特定的应用才会被执行。默认方法列表是唯一的例外。所有应用在未进行配置时使用默认方法列表。

AAA 认证方案：

- 不认证 (none)

对用户非常信任，不对其进行合法性检查。一般情况下不采用这种方法。

- 本地认证 (local)

认证过程在 NAS 设备上完成，用户信息（包括用户名、密码和各种属性）直接配置在接入设备上。当配置 local 参数使用本地数据库进行验证时，需要使用 username password 命令预先在本地创建用户数据库。

- 远程服务器组认证 (group)

认证过程在 NAS 和一个远程服务器组之间完成（一个服务器组可包含任意个相同类型的服务器），NAS 和远程服务器之间通过 RADIUS 或 TACACS+ 协议通信。用户信息集中在远程服务器上统一管理，可以实现大容量、高可靠性、支持多设备的集中式统一认证。为提防远程服务器组的服务器均无效时，可配置本地认证作为备选认证方式完成认证。

AAA 认证类型

锐捷产品目前支持以下认证类型：

- Login (登录) 认证

针对 SSH、Telnet、FTP 等终端接入用户，在用户登录到 NAS 命令行界面时进行身份认证。

- Enable 认证

针对的是用户终端登录到 NAS 上的命令行界面以后，提升命令行界面执行权限时进行认证。即对 enable (进入特权模式) 行为进行认证。

- DOT1X (IEEE802.1x) 认证

针对 IEEE802.1x 接入用户进行身份认证。

- Web-auth (二代 portal) 认证

针对使用二代 portal 服务器来进行身份认证。

相关配置

启动 AAA

缺省情况下，AAA 没有启动。

使用 `aaa new-model` 命令可以启动。

配置 AAA 认证方案

缺省情况下，没有配置任何 AAA 认证方案。

确定使用本地 (Local) 认证还是远程服务器认证。如果用户使用远程服务器认证，则需要先配置 RADIUS 或 TACACS+ 服务器。如果使用 Local 认证，则需要 NAS 上配置本地用户数据库信息。

配置 AAA 认证方法

缺省情况下，没有配置任何 AAA 认证方法。

确定要配置的接入方式，针对不同接入方式配置不同的认证方法。

1.3.2 AAA 授权

AAA 授权使管理员能够对用户可使用的服务或权限进行控制。启用 AAA 授权服务以后，网络设备通过本地或服务器中的用户配置文件信息对用户的会话进行配置。完成授权以后，该用户只能使用配置文件中允许的服务或只具备许可的权限。

AAA 授权方案

- 直接授权 (none)

对用户非常信任，直接授权用户的权限为接入设备允许用户所使用的默认权限。

- 本地授权 (local)

授权过程在 NAS 设备上完成，根据 NAS 上为本地用户配置的相关属性进行授权。

- 远程服务器授权 (group)

授权过程在 NAS 和远程服务器组之间完成。当远程服务器组的服务器均无效时，可以配置本地授权或直接授权作为备选授权方式完成授权。

▾ AAA 授权类型

- Exec 授权

针对的是用户终端登录到 NAS 上的 CLI 界面时，授予用户终端的权限级别（分为 0~15 级）。

- Config-commands 授权

对配置模式（包括全局配置模式及其子模式）下的命令进行授权。

- Console 授权

对通过控制台登录的用户所执行命令的授权。

- Command（命令）授权

用户终端登录到 NAS 上的 CLI 界面以后，针对具体命令的执行授权。

- Network（网络）授权

授予网络连接上的用户会话可用的服务。例如 PPP、SLIP 等网络连接通过 Network 授权，可以获得诸如流量、带宽、超时等服务配置。

相关配置

▾ 启动 AAA

缺省情况下，AAA 没有启动。

使用 `aaa new-model` 命令可以启动。

▾ 配置 AAA 授权方案

缺省情况下，没有配置任何 AAA 授权方案。

确定使用本地 (local) 授权还是远程服务器授权。如果用户使用远程服务器授权，则需要先配置 RADIUS 或 TACACS+ 服务器。

如果使用 Local 授权，则需要 NAS 上配置本地用户数据库信息。

▾ 配置 AAA 授权方法

缺省情况下，没有配置任何 AAA 授权方法。

确定要配置的接入方式，针对不同接入方式配置不同的认证方法。

1.3.3 AAA 记账

在 AAA 中，记账是一个和认证、授权同级别的独立流程，其职责为发送记账开始、更新和结束请求给所配置的记账服务器，由服务器记录用户使用网络资源的情况，实现对用户的活动进行计费、审计以及跟踪等功能。

在 AAA 配置中，记账方案不是必须配置的。

✎ AAA 记账方案

- 不记账 (none)

不对用户记账。

- 本地记账 (local)

记账过程在 NAS 上完成，实现了本地用户连接数的统计和限制，并没有实际的费用统计功能。

- 远程服务器组记账 (group)

记账过程在接入设备和远程的服务器之间完成。当远程服务器组失效时，可配置本地记账作为备选记账方式完成记账。

✎ AAA 记账类型

- Exec 记账

针对的是用户终端登录到 NAS 上的 CLI 界面时，在登入和登出时分别进行记账。

- Command 记账

用户终端登录到 NAS 上的 CLI 界面以后，记录其具体执行的命令。

- Network 记账

记录与网络连接用户（如 802.1X、WEB 认证等用户）会话有关的信息。

相关配置

✎ 启动 AAA

缺省情况下，AAA 没有启动。

使用 `aaa new-model` 命令可以启动。

✎ 配置 AAA 记账方案

缺省情况下，没有配置任何 AAA 记账方案。

确定使用本地 (Local) 记账还是远程服务器记账。如果用户使用远程服务器记账，则需要先配置 RADIUS 或 TACACS+ 服务器。如果使用 Local 记账，则需要先在 NAS 上配置本地用户数据库信息。

✎ 配置 AAA 记账方法

缺省情况下，没有配置任何 AAA 记账方案。

确定要配置的接入方式，针对不同接入方式配置不同的记账方法。

1.3.4 AAA 多域

在多域环境下，同一台网络访问服务器（NAS）设备可为不同域中的用户提供 AAA 服务，各域中用户的属性（例如用户名及密码、服务类型、权限等）有可能各不相同，因此有必要通过设置域的方法把它们区分开，并为每个域单独配置包括 AAA 服务方法列表（例如使用的 RADIUS）在内的属性集。

本产品支持以下几种形式的用户名

1. userid@domain-name
2. domain-name\userid
3. userid.domain-name
4. userid

对于第 4 种不带 domain-name 的形式的用户名（即以上第 4 种：userid），认为其域名称为 default，即为默认的域名。

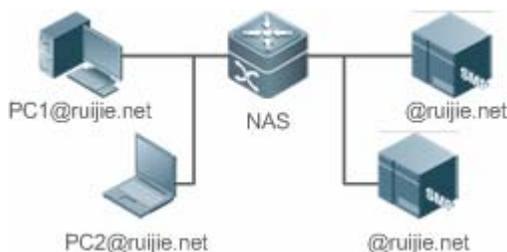
设备基于域名的 AAA 服务基本原理如下：

- 解析用户携带的域名称
- 根据域名称查找用户所配置的域
- 根据设备上域配置信息查找相应的 AAA 服务的方法列表名
- 根据方法列表名在系统中查找对应的方法列表
- 使用该方法列表提供 AAA 服务

i 上述任何一个步骤失败，用户将无法使用申请的 AAA 服务。

以下是典型的多个域环境拓扑图：

图 1-4



相关配置

启动 AAA

缺省情况下，AAA 没有启动。

使用 `aaa new-model` 命令可以启动。

定义 AAA 服务的方法列表

缺省情况下，没有配置任何 AAA 服务的方法列表。

配置方法列表请参照 5.2.1、5.2.2、5.2.3 章节

✚ 启用基于域名的 AAA 服务

缺省情况下，基于域名的 AAA 服务没有启动。

使用 `aaa domain enable` 命令可以启动基于域名的 AAA 服务。

✚ 创建域

缺省情况下，没有配置任何域。

使用 `aaa domain domain-name` 命令配置域名。

✚ 配置域属性集

缺省情况下，没有域属性集。

域属性集包括该域使用的认证、授权、记账方法列表；域的同时在线人数；是否去除用户名中的域名；域是否生效等。

✚ 查看域配置

使用 `show aaa domain` 查看域配置

 系统最多支持配置 32 个域。

1.4 配置详解

配置项	配置建议 & 相关命令	
配置AAA认证	 如果要确认用户的身份，则必须配置。	
	<code>aaa new-model</code>	开启 AAA。
	<code>aaa authentication login</code>	定义 Login 认证的认证方法列表。
	<code>aaa authentication enable</code>	定义 enable 认证的方法类型和执行顺序。
	<code>aaa authentication dot1x</code>	定义 802.1x 认证的方法类型和执行顺序。
	<code>login authentication</code>	在特定终端线路上应用 Login 认证方法。
	<code>dot1x authentication</code>	802.1x 应用认证方法。
	<code>aaa local authentication attempts</code>	设置 login 用户尝试登录次数的最大值。
<code>aaa local authentication lockout-time</code>	设置 login 用户被锁定的时间长度。	
配置AAA授权	 如果要对不同用户赋予不同的权限，限制用户可以使用服务，则必须配置。	
	<code>aaa new-model</code>	开启 AAA。
	<code>aaa authorization exec</code>	定义 exec 授权的方法类型和执行顺序。
	<code>aaa authorization commands</code>	定义 command 授权的方法类型和执行顺序。
	<code>aaa authorization network</code>	为接入用户配置授权方法列表。
	<code>authorization exec</code>	在特定终端线路上应用 exec 授权方法。

	authorization commands	在特定终端线路上应用 command 授权方法。
配置AAA记账	 如果要实现对用户使用网络资源情况的记账、统计和跟踪，则必须配置。	
	aaa new-model	开启 AAA。
	aaa accounting exec	定义 exec 记账的方法类型及方法执行顺序。
	aaa accounting commands	定义 command 记账的方法类型及方法执行顺序。
	aaa accounting network	定义 network 记账的方法类型及方法执行顺序。
	accounting exec	在特定终端线路上应用 exec 记账方法。
	accounting commands	在特定终端线路上应用 command 记账方法。
	aaa accounting update	开启记账更新功能。
	aaa accounting update periodic	设置记账更新时间间隔。
配置AAA服务器组	 如果有多台服务器且需要能灵活选择服务器进行认证、授权和记账的处理，则建议配置。	
	aaa group server	创建 AAA 自定义服务器组。
	server	添加 AAA 服务器组成员。
配置基于域名的AAA服务	 如果需要通过域来对接入的 802.1x 用户进行 AAA 管理，则必须配置。	
	aaa new-model	开启 AAA。
	aaa domain enable	开启基于域名的 AAA 服务。
	aaa domain	创建域，并进入域配置模式。
	authentication dot1x	在域中，关联 802.1X 认证方法列表。
	accounting network	在域中，关联 Network 记账方法列表。
	authorization network	在域中，关联 Network 授权方法列表。
	state	设置域的状态。
	username-format	设置是否在用户名中携带域名信息。
access-limit	设置当前域可容纳接入用户的数目限制。	

1.4.1 配置 AAA 认证

配置效果

验证用户是否可以获得访问权。

注意事项

- 如果在一个认证方案中使用多种认证方法，则认证方法的执行顺序为配置的先后顺序。只有在当前认证方法没有响应的情况下，才会采用下一种认证方法；如果当前认证方法认证失败，则不会跳转到下一个认证方案进行认证。

- 由于 none 方法使得请求接入的任何用户在所有认证方法都没有应答情况下能通过身份认证，所以仅将它作为备用的身份认证方法。
-
- i** 一般情况下，不要使用 none 身份认证。在特殊情况（如所有可能的申请接入用户都是可信任的，而且用户的工作不允许有由于系统故障造成的耽搁），可以在安全服务器无应答的情况下，将 none 作为最后一种可选的身份认证方法，建议在 none 认证方法前加上本地身份认证方法。
-
- AAA 认证开启的情况下，如果没有配置任何方法且不存在 default 认证方法时，对于控制台允许不认证直接登录；其他接入都要进行 local 认证。
 - 如果进入 CLI 界面的时候经过了 Login 身份认证（none 方法除外），将记录当前使用的用户名。此时，进行 Enable 认证的时候，将不再提示输入用户名，直接使用与 Login 认证相同的用户名进行认证，注意输入的口令要与之匹配。
 - 如果进入 CLI 界面的时候没有进行 Login 认证，或在 Login 认证的时候使用了 none 方法，将不会记录用户名信息。此时，如果进行 enable 认证，将会要求重新输入用户名。这个用户名信息不会被记录，每次进行 Enable 认证都要重新输入。

配置方法

▾ 开启 AAA

- 必须配置。
- 使用 `aaa new-model` 开启 AAA。
- 缺省情况下，没有启动 AAA。

▾ 定义 Login 认证的方法类型和执行顺序。

- 使用命令 `aaa authentication login` 配置 Login 认证的方法类型和执行顺序。
- 如果为 Login 接入用户配置认证方法列表（包括配置 default 方法列表），则必须配置此命令。
- 缺省情况下，没有配置 Login 认证方法列表。

▾ 定义 Enable 认证的方法类型和执行顺序。

- 使用命令 `aaa authentication enable` 配置 Enable 认证的方法类型和执行顺序。
- 如果为 Enable 过程配置认证方法列表（只能配置 default 方法列表），则必须配置此命令。
- 缺省情况下，没有配置 Enable 认证方法列表。

▾ 定义 802.1x 认证的方法类型和执行顺序。

- 使用命令 `aaa authentication dot1x` 配置 Login 认证的方法类型和执行顺序。
- 如果为 802.1x 接入用户配置认证方法列表（包括配置 default 方法列表），则必须配置此命令。
- 缺省情况下，没有配置 dot1x 认证方法列表。

▾ 在特定终端线路上应用 Login 认证方法。

- 使用 `login authentication(line 模式下)` 命令在特定终端线路上应用 Login 认证方法。
- 如果要在特定线路上应用指定的 Login 认证方法列表，则必须配置此命令。

- 缺省情况下，所有终端线路关联 default 方法列表。

📌 802.1x 应用认证方法

- 使用 **dot1x authentication** 命令配置 802.1x 应用认证方法。
- 如果要指定 802.1x 认证方法，则必须配置此命令。
- 缺省情况下，802.1x 认证不应用方法列表。

📌 设置 login 用户尝试登录次数的最大值。

- 可选配置。
- 缺省情况下，允许 login 用户尝试密码的失败次数为 3 次。

📌 设置 login 用户被锁定的时间长度。

- 可选配置。
- 缺省情况下，当 login 用户尝试登录的次数超过最大值，被锁定的时间为 15 分钟。

检验方法

- 使用 **show aaa method-list** 查看已配置的方法列表信息。
- 使用 **show aaa lockout** 查看用户尝试登录失败次数的最大值和用户锁定的时间长度的配置信息。
- 使用 **show running-config** 查看 Login 认证、dot1x 认证关联认证方法列表的信息。

相关命令

📌 开启 AAA

- 【命令格式】 **aaa new-model**
- 【参数说明】 无
- 【命令模式】 全局模式
- 【使用指导】 该命令是 AAA 的使能命令，如果您要使用 AAA 安全服务，就必须使用 **aaa new-model** 使能 AAA 安全服务。如果没有启用 AAA，则所有 AAA 命令将是不可配置的。

📌 定义 Login 认证的方法类型和执行顺序。

- 【命令格式】 **aaa authentication login { default | list-name } method1 [method2...]**
- 【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 Login 认证的默认方法。
list-name：定义一个 Login 认证的方法列表，可以是任何字符串。
method：必须是“local、none、group”所列关键字之一，一个方法列表最多有 4 个方法。
local：使用本地用户名数据库进行身份认证。
none：不进行身份认证。
group：使用服务器组进行身份认证，目前支持 RADIUS 和 TACACS+服务器组。
- 【命令模式】 全局模式

- 【使用指导】 如果设备启用 AAA 登录认证安全服务，用户就必须使用 AAA 进行 Login 认证协商。您必须使用 **aaa authentication login** 命令配置默认的或可选的方法列表用于 Login 认证。
只有前面的方法没有响应，才能使用后面的方法进行身份认证。
设置了 Login 认证方法后，必须将其应用在需要进行 Login 认证的终端线路上，否则将不生效。

▾ 定义 Enable 认证的方法类型和执行顺序

- 【命令格式】 **aaa authentication enable default** *method1* [*method2...*]
- 【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 Enable 认证的默认方法。
list-name：定义一个 Enable 认证的方法列表，可以是任何字符串。
method：必须是“enable、local、none、group”所列关键字之一，一个方法列表最多有 4 个方法。
enable：使用 enable 命令配置的密码进行认证。
local：使用本地用户名数据库进行身份认证。
none：不进行身份认证。
group：使用服务器组进行身份认证，目前支持 RADIUS 和 TACACS+服务器组。
- 【命令模式】 全局模式
- 【使用指导】 如果设备启用 AAA 登录认证安全服务，用户就必须使用 AAA 进行 Enable 认证协商。您必须使用 **aaa authentication enable** 命令配置默认的或可选的方法列表用于 Enable 认证。
只有前面的方法没有响应，才能使用后面的方法进行身份认证。

▾ 定义 802.1x 认证的方法类型和执行顺序。

- 【命令格式】 **aaa authentication dot1x** { **default** | *list-name* } *method1* [*method2...*]
- 【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 dot1x 认证的默认方法。
list-name：定义一个 dot1x 认证的方法列表，可以是任何字符串。
method：必须是“local、none、group”所列关键字之一，一个方法列表最多有 4 个方法。
local：使用本地用户名数据库进行身份认证。
none：不进行身份认证。
group：使用服务器组进行身份认证，目前支持 RADIUS 服务器组。
- 【命令模式】 全局模式
- 【使用指导】 如果设备启用 AAA 802.1X 安全服务，用户就必须使用 AAA 进行 802.1X 用户认证协商。您必须使用 **aaa authentication dot1x** 命令配置默认的或可选的方法列表用于 802.1X 用户认证。
只有前面的方法没有响应，才能使用后面的方法进行认证。

▾ 设置 login 用户尝试登录次数的最大值。

- 【命令格式】 **aaa local authentication attempts** *max-attempts*
- 【参数说明】 *max-attempts*：最大尝试失败次数，取值范围 1~2147483647
- 【命令模式】 全局模式
- 【使用指导】 该命令配置 Login 登录用户尝试登录失败次数。

▾ 设置 login 用户被锁定的时间长度。

- 【命令格式】 **aaa local authentication lockout-time** *lockout-time*
- 【参数说明】 *lockout-time*：锁定时间（单位：分钟），取值范围 1~2147483647

- 【命令模式】 全局模式
- 【使用指导】 配置 Login 登录用户尝试超过配置登录失败次数后被锁定的时间长度。

配置举例

i 以下配置举例，仅介绍与 AAA 认证相关的配置。

📌 AAA Login 认证配置示例。对 Login 用户先用 RADIUS 服务器进行认证，在远程服务器没有响应的情况下转本地认证。

【网络环境】

图 1-5



【配置方法】

第一步：开启 AAA。

第二步：如果用户使用远程服务器认证，则需要先配置 RADIUS 或 TACACS+服务器。如果使用 Local 认证，则需要先在 NAS 上配置本地用户数据库信息。（本例需要配置 RADIUS 服务器和本地数据库信息）

第三步：根据不同接入用户类型（本例为 Login 用户），配置 AAA 认证方法列表（本例的认证方法是先 RADIUS 认证，无响应后转 Local 认证）。

第四步：将方法应用于某个特定的接口或线路。如果使用的是 default 认证方法，则可以不必配置该步骤。

NAS

```

Ruijie#configure terminal
Ruijie(config)#username user password pass
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key ruijie
Ruijie(config)#aaa authentication login list1 group radius local
Ruijie(config)#line vty 0 20
Ruijie(config-line)#login authentication list1
Ruijie(config-line)#exit
  
```

【检验方法】

在 NAS 设备上，通过 **show aaa method-list** 命令查看配置效果。

NAS

```

Ruijie#show aaa method-list

Authentication method-list:
aaa authentication login list1 group radius local

Accounting method-list:

Authorization method-list:
  
```

以 Telnet 用户为例，用户远程登录到 NAS 设备上，CLI 界面提示输入用户名/密码。输入正确的用户名/密码，才能访问设备。

User

```
User Access Verification
```

```
Username:user
Password:pass
```

- ▾ **AAA enable 认证配置示例。对 enable 认证先使用 RADIUS 服务器进行认证，在远程服务器没有响应的情况下转本地认证，在本地认证用户名不存在的情况下转 enable 密码认证。**

【网络环境】

图 1-6



【配置方法】

第一步：开启 AAA。

第二步：如果用户使用远程服务器认证，则需要先配置 RADIUS 或 TACACS+服务器。如果使用 Local 认证，则需要 NAS 上配置本地用户数据库信息。如果使用 enable 密码认证，则需要 NAS 上配置 enable 认证密码。

第三步：根据不同接入用户类型，配置 AAA 认证方法列表。

- ① Enable 认证方法列表全局只能定义一个，因此 Enable 认证不需要定义方法列表的名称，只要配置成默认的方法列表，配置以后，会自动被应用。

NAS

```
Ruijie#configure terminal
Ruijie(config)#username user privilege 15 password pass
Ruijie(config)#enable secret w
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key ruijie
Ruijie(config)#aaa authentication enable default group radius local enable
```

【检验方法】

在 NAS 设备上，通过 **show aaa method-list** 命令查看配置效果。

NAS

```
Ruijie#show aaa method-list

Authentication method-list:
aaa authentication enable default group radius local enable

Accounting method-list:

Authorization method-list:
```

用户级别切换到 15 级，CLI 提示认证。输入正确的用户名/密码，才能访问设备。

NAS

```
Ruijie>enable
Username:user
Password:pass
```

```
Ruijie#
```

- ▾ **AAA 802.1x 认证配置示例。对 802.1x 接入用户先用 RADIUS 服务器进行认证，在远程服务器没有响应的情况下转本地认证。**

【网络环境】

图 1-7



【配置方法】

第一步：开启 AAA。

第二步：如果用户使用远程服务器认证，则需要先配置 RADIUS 服务器。如果使用 Local 认证，则需要在 NAS 上配置本地用户数据库信息。（本例需要配置 RADIUS 服务器和本地数据库信息）。目前，802.1X 认证不支持使用 TACACS+ 认证。

第三步：根据不同接入用户类型（本例为 802.1x 接入用户），配置 AAA 认证方法列表（本例的认证方法是先 RADIUS 认证，无响应后转 Local 认证）。

第四步：应用 AAA 认证方法。如果使用的是 default 认证方法，则可不配置该步骤。

第五步：接口开启 802.1x 认证功能。

NAS

```

Ruijie#configure terminal
Ruijie(config)#username user1 password pass1
Ruijie(config)#username user2 password pass2
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key ruijie
Ruijie(config)#aaa authentication dot1x default group radius local
Ruijie(config)#interface gigabitEthernet 0/1
Ruijie(config-if-gigabitEthernet 0/1)#dot1x port-control auto
Ruijie(config-if-gigabitEthernet 0/1)#exit
  
```

【检验方法】

在 NAS 设备上，通过 **show aaa method-list** 命令查看配置效果。

NAS

```

Ruijie#show aaa method-list

Authentication method-list:
aaa authentication dot1x default group radius local

Accounting method-list:

Authorization method-list:
  
```

常见错误

- 没有配置 RADIUS 服务器或者 TACACS+ 服务器。

- 没有配置本地数据库用户名和密码。

1.4.2 配置 AAA 授权

配置效果

- 定义用户可以使用哪些服务或拥有哪些权限。

注意事项

- 关于 Exec 授权：Exec 授权通常结合 Login 认证一起使用，并可以在同一个线路上同时使用 Login 认证和 Exec 授权。但是要注意，由于授权和认证可以采用不同的方法和不同的服务器，因此对于相同的用户，认证和授权可能有不同的结果。用户登录时，如果 Exec 授权失败，即使已经通过了 Login 认证，也不能进入到 CLI 界面。
- 关于授权方法：如果在一个授权方案中使用多种授权模式，则授权模式的执行顺序为配置的先后顺序。只有在当前授权模式没有响应的情况下，才会采用下一种授权模式；如果当前授权模式失败，则不会采用下一种授权模式进行授权。
- 关于 Command 授权：Command 授权功能目前仅 TACACS+协议支持。
- 关于 Console 授权：RGOS 支持区分通过控制台登录和其他终端登录的用户，可以设置控制台登录的用户，是否需要进行命令授权。如果关闭了控制台的命令授权功能，则已经应用到控制台线路的命令授权方法列表将不生效。

配置方法

▾ 开启 AAA

- 必须配置。
- 使用 `aaa new-model` 开启 AAA。
- 缺省情况下，没有启动 AAA。

▾ 定义 exec 授权的方法类型和执行顺序。

- 使用 `aaa authorization exec` 命令配置 exec 授权的方法类型和执行顺序。
- 如果要为 exec 用户配置授权方法列表（包括配置 default 方法列表），则必须配置此命令。
- 缺省情况下，没有配置授权方法。

i Exec 用户（控制台用户，可以通过 Console 口或者 Telnet 连接设备，每个连接称为一个 EXEC 用户，如 Telnet 用户、SSH 用户）的默认级别为最低权限的访问级别。

▾ 定义 command 授权的方法类型和执行顺序。

- 使用 `aaa authorization commands` 命令配置 command 授权的方法类型和执行顺序。
- 如果要为 command 授权配置授权方法列表（包括配置 default 方法列表），则必须配置此命令。

- 缺省情况下，没有配置授权方法。

▾ 为接入用户配置授权方法列表。

- 使用 **aaa authorization network** 命令为接入用户配置认证方法列表。
- 如果要为 network 用户配置授权列表（包括配置 default 方法列表），则必须配置此命令。
- 缺省情况下，没有配置授权方法。

▾ 在特定终端线路上应用 exec 授权方法。

- 使用 **authorization exec** (line 模式下)命令为特定终端线路上应用 exec 授权方法。
- 如果要在特定线路上应用指定的 exec 授权方法列表，则必须配置此命令。
- 缺省情况下，所有终端线路关联 default 授权方法列表。

▾ 在特定终端线路上应用 command 授权方法。

- 使用 **authorization commands** (line 模式下)命令为特定终端线路上应用 command 授权方法。
- 如果要在特定线路上应用指定的 command 授权方法列表，则必须配置此命令。
- 缺省情况下，所有终端线路关联 default 授权方法列表。

▾ 开启需要对配置模式下的命令进行授权。

- 使用 **aaa authorization config-commands** 命令开启需要对配置模式下的命令进行授权的功能。
- 缺省情况下，对配置模式下的命令不开启授权功能。

▾ 开启对控制台的用户执行的命令进行授权。

- 使用 **aaa authorization console** 命令开启对控制台的用户执行的命令进行授权的功能。
- 缺省情况下，不开启对控制台的用户执行的命令进行授权的功能。

检验方法

使用 **show running-config** 命令查看以上配置是否生效。

相关命令

▾ 开启 AAA。

【命令格式】 **aaa new-model**

【参数说明】 无

【命令模式】 全局模式

【使用指导】 该命令是 AAA 的使能命令，如果您要使用 AAA 安全服务，就必须使用 **aaa new-model** 使能 AAA 安全服务。如果没有启用 AAA，则所有 AAA 命令将是不可配置的。

▾ 定义 exec 授权的方法类型和执行顺序。

【命令格式】 **aaa authorization exec { default | list-name } method1 [method2...]**

【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 exec 授权的默认方法。

list-name：定义一个 exec 授权的方法列表，可以是任何字符串。

method：必须是“local、none、group”所列关键字之一，一个方法列表最多有 4 个方法。

local：使用本地用户名数据库进行 exec 授权。

none：不进行 exec 授权。

group：使用服务器组进行 exec 授权，目前支持 RADIUS 和 TACACS+服务器组。

【命令模式】 全局模式

【使用指导】 RGOS 支持对登录到 NAS 的 CLI 界面的用户进行授权，赋予其 CLI 权限级别（0~15 级）。目前对于通过了 Login 认证的用户，才进行 Exec 授权。如果 Exec 授权失败，则无法进入 CLI 界面。

配置了 Exec 授权方法后，必须将其应用在需要进行 Exec 授权的终端线路上，否则将不生效。

▾ 定义 command 授权的方法类型和执行顺序。

【命令格式】 **aaa authorization commands level { default | list-name } method1 [method2...]**

【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 command 授权的默认方法。

list-name：定义一个 command 授权的方法列表，可以是任何字符串。

method：必须是“none、group”所列关键字之一，一个方法列表最多有 4 个方法。

none：不进行 command 授权。

group：使用服务器组进行 command 授权，目前 TACACS+服务器组。

【命令模式】 全局模式

【使用指导】 RGOS 支持对用户可执行的命令进行授权，当用户输入并试图执行某条命令时，AAA 将该命令发送到安全服务器上，如果安全服务器允许执行该命令，则该命令被执行，否则该命令不执行，并会给出执行命令被拒绝的提示。

配置命令授权的时候需要指定命令的级别，这个级别是命令的默认级别（例如，某命令对于 14 级以上用户可见，则该命令的默认级别就是 14 级的）。

配置了命令授权方法后，必须将其应用在需要进行命令授权的终端线路上，否则将不生效。

▾ 为接入用户配置授权方法列表。

【命令格式】 **aaa authorization network { default | list-name } method1 [method2...]**

【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 network 授权的默认方法。

list-name：定义一个 network 授权的方法列表，可以是任何字符串。

method：必须是“none、group”所列关键字之一，一个方法列表最多有 4 个方法。

none：不进行身份认证。

group：使用服务器组进行 network 授权，目前支持 RADIUS 和 TACACS+服务器组。

【命令模式】 全局模式

【使用指导】 RGOS 支持对所有网络有关的服务请求如 PPP、SLIP 等协议进行授权。如果配置了授权，则对所有的认证用户或接口自动进行授权。

可以指定三种不同的授权方法，与身份认证一样，只有当前的授权方法没有响应，才能继续使用后面的方法进行授权，如果当前授权方法失败，则不再使用其他后继的授权方法。

RADIUS 或 TACACS+服务器是通过返回一系列的属性对来完成对认证用户的授权。所以网络授权是建立在认证的基础上的，只有认证通过了才有可能获取网络授权。

✎ 开启对配置模式（包括全局配置模式及其子模式）下的命令进行授权的功能。

【命令格式】 **aaa authorization config-commands**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 如果只对非配置模式（如特权模式）下的命令进行授权，可以使用该命令的 **no** 模式关闭配置模式的授权功能，则配置模式及其子模式下的命令不需要进行命令授权就可以执行。

✎ 开启对通过控制台登录的用户所执行的命令进行授权的功能。

【命令格式】 **aaa authorization console**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 RGOS 支持区分通过控制台登录和其他终端登录的用户，可以设置控制台登录的用户，是否需要进行命令授权。如果关闭了控制台的命令授权功能，则已经应用到控制台线路的命令授权方法列表将不生效。

配置举例

i 以下配置举例，仅介绍与 AAA 授权相关的配置。

✎ 配置 AAA exec 授权。VTY 线路 0~4 上的用户登录时采用 Login 认证，并且进行 exec 授权。其中 Login 认证采用本地认证，exec 授权先采用 RADIUS，如果没有响应可以采用本地授权。

【网络环境】

图 1-8



【配置方法】

第一步：开启 AAA。

第二步：如果用户使用远程服务器授权，则需要先配置 RADIUS 或 TACACS+ 服务器。如果使用 local 授权，则需要 NAS 上配置本地用户数据库信息。

第三步：根据不同接入方式和服务类型，配置 AAA 授权方法列表。

第四步：将方法应用于某个特定的接口或线路。如果使用的是 default 认证方法，则可以不必配置该步骤。

Exec 授权通常结合 Login 认证一起使用，并可以在同一个线路上同时使用 Login 认证和 Exec 授权。

NAS

```

Ruijie#configure terminal
Ruijie(config)#username user password pass
Ruijie(config)#username user privilege 6
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key test
Ruijie(config)#aaa authentication login list1 group local
Ruijie(config)#aaa authorization exec list2 group radius local
Ruijie(config)#line vty 0 4
  
```

```
Ruijie(config-line)#login authentication list1
Ruijie(config-line)# authorization exec list2
Ruijie(config-line)#exit
```

【检验方法】 在 NAS 设备上，通过 **show run**、**show aaa method-list** 命令查看配置效果。

NAS

```
Ruijie#show aaa method-list

Authentication method-list:
aaa authentication login list1 group local

Accounting method-list:

Authorization method-list:
aaa authorization exec list2 group radius local

Ruijie# show running-config
aaa new-model
!
aaa authorization exec list2 group local
aaa authentication login list1 group radius local
!
username user password pass
username user privilege 6
!
radius-server host 10.1.1.1
radius-server key 7 093b100133
!
line con 0
line vty 0 4
  authorization exec list2
  login authentication list1
!
End
```

👉 **配置 Command 授权。**为 Login 用户设置命令授权，应用 default 授权方法：对 15 级命令进行授权，先使用 tacacs+ 服务器授权，无响应后转 local 授权。授权同时应用于控制台登录用户和其他终端登录的用户。

【网络环境】

图 1-9



【配置方法】 第一步：开启 AAA。

第二步：如果用户使用远程服务器授权，则需要先配置 RADIUS 或 TACACS+服务器。如果使用 local 授权，则需要 NAS 上配置本地用户数据库信息。

第三步：根据不同接入方式和服务类型，配置 AAA 授权方法列表。

第四步：将方法应用于某个特定的接口或线路。如果使用的是 default 认证方法，则可以不必配置该步骤。

NAS

```
Ruijie#configure terminal
Ruijie(config)#username user1 password pass1
Ruijie(config)#username user1 privilege 15
Ruijie(config)#aaa new-model
Ruijie(config)#tacacs-server host 192.168.217.10
Ruijie(config)#tacacs-server key aaa
Ruijie(config)#aaa authentication login default local
Ruijie(config)#aaa authorization commands 15 default group tacacs+ local
Ruijie(config)#aaa authorization console
```

【检验方法】

在 NAS 设备上，通过 **show run**、**show aaa method-list** 命令查看配置效果。

NAS

```
Ruijie#show aaa method-list

Authentication method-list:
aaa authentication login default local

Accounting method-list:

Authorization method-list:
aaa authorization commands 15 default group tacacs+ local

Ruijie#show run
!
aaa new-model
!
aaa authorization console
aaa authorization commands 15 default group tacacs+ local
aaa authentication login default local
!
!
nfpp
!
vlan 1
!
username user1 password 0 pass1
username user1 privilege 15
no service password-encryption
!
tacacs-server host 192.168.217.10
```

```
tacacs-server key aaa
!
line con 0
line vty 0 4
!
!
end
```

配置 Network 授权。

【网络环境】

图 1-10



【配置方法】

第一步：开启 AAA。

第二步：如果用户使用远程服务器授权，则需要先配置 RADIUS 或 TACACS+服务器。如果使用 local 授权，则需要 NAS 上配置本地用户数据库信息。

第三步：根据不同接入方式和服务类型，配置 AAA 授权方法列表。

第四步：将方法应用于某个特定的接口或线路。如果使用的是 default 认证方法，则可以不必配置该步骤。

NAS

```
Ruijie#configure terminal
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key test
Ruijie(config)#aaa authorization network default group radius none
Ruijie(config)# end
```

【检验方法】

在 NAS 设备上，通过 **show aaa method-list** 命令查看配置效果。

NAS

```
Ruijie#show aaa method-list

Authentication method-list:

Accounting method-list:

Authorization method-list:
aaa authorization network default group radius none
```

常见配置错误

无

1.4.3 配置 AAA 记账

配置效果

- 记录用户使用网络资源的情况。
- 记录用户进行设备管理时登入登出的过程、记录执行过的命令。

注意事项

关于记账方法：

- 如果在一个记账方案中使用多种记账模式，则记账模式的执行顺序为配置的先后顺序。只有在当前记账模式没有响应的情况下，才会采用下一种记账模式；如果当前记账模式失败，则不会采用下一种记账模式进行记账。
- 默认的记账方法（default 方法）列表一旦配置，将自动应用到所有终端上。在线路上应用非默认记账方法列表，将取代默认的方法列表。如果试图应用未定义的方法列表，则会给出一个警告提示信息，该线路上的记账将不会生效，直至定义了该记账方法列表才会生效。

关于 Exec 记账：

- 只有登录到 NAS 的用户终端通过了 Login 认证，才会进行 exec 记账。如果没有设置 Login 认证，或者认证时候采用了 none 方法，则不会进行 exec 记账。针对同一个用户终端的登录，登入时如果没有进行过 Start 记账，登出时也就不会进行 Stop 记账。

关于 Command 记账：

- Command 记账功能目前仅 TACACS+协议支持。

配置方法

▾ 开启 AAA。

- 必须配置。
- 使用 `aaa new-model` 开启 AAA。
- 缺省情况下，没有启动 AAA。

▾ 定义 exec 记账的方法类型及方法执行顺序。

- 使用命令 `aaa accounting exec` 配置 exec 记账的方法类型及方法执行顺序。
- 如果要为 exec 用户配置记账方法（包括配置 default 方法列表），则必须配置此命令。
- Exec 用户（控制台用户，可以通过 Console 口或者 Telnet 连接设备，每个连接称为一个 EXEC 用户，如 Telnet 用户、SSH 用户）的默认级别为最低权限的访问级别。
- 缺省情况下，没有配置记账方法。

▾ 定义 command 记账的方法类型及方法执行顺序。

- 使用命令 **aaa accounting commands** 配置 command 记账的方法类型及方法执行顺序。
- 如果要为 command 记账配置记账方法（包括配置 default 方法列表），则必须配置此命令。
- 缺省情况下，没有配置记账方法。命令记账功能目前仅 TACACS+协议支持。

▾ 定义 network 记账的方法类型及方法执行顺序。

- 使用命令 **aaa accounting network** 配置 network 记账的方法类型及方法执行顺序。
- 如果要为 network 用户配置记账方法（包括配置 default 方法列表），则必须配置此命令。
- 缺省情况下，没有配置记账方法。

▾ 在特定终端线路上应用 exec 记账方法。

- 使用命令 **accounting exec**(line 模式下)配置在特定终端线路上应用 exec 记账方法。
- 如果要在特定线路上应用指定的 exec 记账方法列表，则必须配置此命令。
- 如果应用的是 default 方法列表，则可不配置此命令。
- 缺省情况下，所有终端线路关联 default 方法列表。

▾ 在特定终端线路上应用 command 记账方法。

- 使用命令 **accounting commands**(line 模式下)配置在特定终端线路上应用 command 记账方法。
- 如果要在特定线路上应用指定的 command 记账方法列表，则必须配置此命令。
- 如果应用的是 default 方法列表，则可不配置此命令。
- 缺省情况下，所有终端线路关联 default 方法列表。

▾ 802.1x 应用 network 记账方法

- 使用命令 **dot1x accounting network** 命令配置 802.1x 的 network 记账方法。
- 如果要指定 802.1X 记账方法，则必须配置此命令。
- 如果应用的是 default 方法列表，则可不配置此命令。
- 缺省情况下，关联 default 方法列表。

▾ 开启记账更新功能。

- 可选配置。
- 该功能有助于提高记账准确性，建议配置。
- 缺省情况下，记账更新功能关闭。

▾ 设置记账更新时间间隔。

- 可选配置。
- 除非有明确要求，否则不建议配置。

检验方法

使用 `show running-config` 命令查看配置是否生效。

相关命令

▾ 开启 AAA。

- 【命令格式】 `aaa new-model`
- 【参数说明】 无
- 【命令模式】 全局模式
- 【使用指导】 该命令是 AAA 的使能命令，如果您要使用 AAA 安全服务，就必须使用 `aaa new-model` 使能 AAA 安全服务。如果没有启用 AAA，则所有 AAA 命令将是不可配置的。

▾ 定义 exec 记账的方法类型及方法执行顺序。

- 【命令格式】 `aaa accounting exec { default | list-name } start-stop method1 [method2...]`
- 【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 exec 记账的默认方法。
list-name：定义一个 exec 记账的方法列表，可以是任何字符串。
method：必须是“none、group”所列关键字之一，一个方法列表最多有 4 个方法。
none：不进行 exec 记账。
group：使用服务器组进行 exec 记账，目前支持 RADIUS 和 TACACS+服务器组。
- 【命令模式】 全局模式
- 【使用指导】 RGOS 只有在用户通过了登录认证后，才会启用 Exec 记账功能，如果用户登录时未进行认证或认证采用的方法为 none，则不会进行 Exec 记账。
启用记账功能后，在用户登录到 NAS 的 CLI 界面时候，发送记账开始（Start）信息给安全服务器，在用户退出登录的时候，发送记账结束（Stop）信息给安全服务器。如果一个用户在登录时没有发出 Start 信息，在退出登录时也不会发出 Stop 信息。
配置了 Exec 记账方法后，必须将其应用在需要进行命令记账的终端线路上，否则将不生效。

▾ 定义 command 记账的方法类型及方法执行顺序。

- 【命令格式】 `aaa accounting commands level { default | list-name } start-stop method1 [method2...]`
- 【参数说明】 *level*：要进行记账的命令级别，范围 0~15，决定哪个级别的命令执行时，需要记录信息。
default：使用该参数，则后面定义的方法列表作为 command 记账的默认方法。
list-name：定义一个 command 记账的方法列表，可以是任何字符串。
method：必须是“none、group”所列关键字之一，一个方法列表最多有 4 个方法。
none：不进行 command 记账。
group：使用服务器组进行 command 记账，目前支持 TACACS+服务器组。
- 【命令模式】 全局模式
- 【使用指导】 RGOS 只有在用户通过了登录认证后，才会启用命令记账功能，如果用户登录时未进行认证或认证采用的方法为 none，则不会进行命令记账。启用记账功能后，在用户每次执行指定级别的命令后，将所执行的命令信息，发送给安全服务器。

配置了命令记账方法后，必须将其应用在需要进行命令记账的终端线路上，否则将不生效。

定义 network 记账的方法类型及方法执行顺序。

【命令格式】 **aaa accounting network { default | list-name } start-stop method1 [method2...]**

【参数说明】 **default**：使用该参数，则后面定义的方法列表作为 network 记账的默认方法。

list-name：定义一个 command 记账的方法列表，可以是任何字符串。

start-stop：在用户访问活动开始和结束时均发送记账报文，开始记账报文无论是否成功启用记账，都允许用户开始进行网络访问。

method：必须是“none、group”所列关键字之一，一个方法列表最多有 4 个方法。

none：不进行 network 记账。

group：使用服务器组进行 network 记账，目前支持 RADIUS 和 TACACS+服务器组。

【命令模式】 全局模式

【使用指导】 RGOS 通过给安全服务器发送记录属性对来用户活动进行记账。使用关键字 **start-stop**，制定用户记账选项。

开启记账更新功能。

【命令格式】 **aaa accounting update**

【参数说明】 无

【命令模式】 全局模式

【使用指导】 如果没有启用 AAA 安全服务，则不能使用记账更新。如果已经启用 AAA 安全服务，则该命令用设置记账更新功能。

设置记账更新时间间隔。

【命令格式】 **aaa accounting update periodic interval**

【参数说明】 **Interval**：记账更新时间间隔，以分钟为单位，最小为 1 分钟。

【命令模式】 全局模式

【使用指导】 如果没有启用 AAA 安全服务，则不能使用记账更新。如果已经启用 AAA 安全服务，则该命令用设置记账更新时间间隔。

配置举例

i 以下配置举例，仅介绍与 AAA 记账相关的配置。

配置 AAA exec 记账。VTY 线路 0~4 上的用户登录时采用 Login 认证，并且进行 exec 记账。其中 Login 认证采用本地认证，exec 记账采用 RADIUS 记账。

【网络环境】

图 1-11



【配置方法】 第一步：开启 AAA。

第二步：如果用户使用远程服务器记账，则需要先配置 RADIUS 或 TACACS+服务器。

第三步：根据不同接入方式和服务类型，配置 AAA 记账方法列表。

第四步：将方法应用于某个特定的接口或线路。如果使用的是 default 认证方法，则可以不必配置该步骤。

NAS

```
Ruijie#configure terminal
Ruijie(config)#username user password pass
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key test
Ruijie(config)#aaa authentication login list1 group local
Ruijie(config)#aaa accounting exec list3 start-stop group radius
Ruijie(config)#line vty 0 4
Ruijie(config-line)#login authentication list1
Ruijie(config-line)# accounting exec list3
Ruijie(config-line)#exit
```

【检验方法】

在 NAS 设备上，通过 **show run**、**show aaa method-list** 命令查看配置效果。

NAS

```
Ruijie#show aaa method-list

Authentication method-list:
aaa authentication login list1 group local

Accounting method-list:
aaa accounting exec list3 start-stop group radius

Authorization method-list:

Ruijie# show running-config
aaa new-model
!
aaa accounting exec list3 start-stop group radius
aaa authentication login list1 group local
!
username user password pass
!
radius-server host 10.1.1.1
radius-server key 7 093b100133
!
line con 0
line vty 0 4
    accounting exec list3
    login authentication list1
!
End
```

- ✎ 配置 command 记账。为 Login 用户设置命令记账 ,应用 default 记账方法。其中 Login 认证采用本地认证 ,使用 tacacs+ 服务器记账。

【网络环境】

图 1-12



【配置方法】

第一步：开启 AAA。

第二步：如果用户使用远程服务器记账，则需要先配置 RADIUS 或 TACACS+ 服务器。

第三步：根据不同接入方式和服务类型，配置 AAA 记账方法列表。

第四步：将方法应用于某个特定的接口或线路。如果使用的是 default 认证方法，则可以不必配置该步骤。

NAS

```

Ruijie#configure terminal
Ruijie(config)#username user1 password pass1
Ruijie(config)#username user1 privilege 15
Ruijie(config)#aaa new-model
Ruijie(config)#tacacs-server host 192.168.217.10
Ruijie(config)#tacacs-server key aaa
Ruijie(config)#aaa authentication login default local
Ruijie(config)#aaa accounting commands 15 default start-stop group tacacs+
  
```

【检验方法】

在 NAS 设备上，通过 show 命令查看配置效果。

NAS

```

Ruijie#show aaa method-list

Authentication method-list:
aaa authentication login default local

Accounting method-list:
aaa accounting commands 15 default start-stop group tacacs+
Authorization method-list:

Ruijie#show run
!
aaa new-model
!
aaa authorization config-commands
aaa accounting commands 15 default start-stop group tacacs+
aaa authentication login default local
!
!
nfpp
!
  
```

```

vlan 1
!
username user1 password 0 pass1
username user1 privilege 15
no service password-encryption
!
tacacs-server host 192.168.217.10
tacacs-server key aaa
!
line con 0
line vty 0 4
!
!
end

```

↘ 配置 network 记账。为 802.1x 用户配置记账方法列表，采用 RADIUS 远程服务器认证和记账。

【网络环境】

图 1-13



【配置方法】

第一步：开启 AAA。

第二步：如果用户使用远程服务器记账，则需要先配置 RADIUS 服务器。

第三步：根据不同接入方式和服务类型，配置 AAA 方法列表。

第四步：应用方法列表。如果使用的是 default 认证方法，则可以不必配置该步骤。

i 802.1X 用户在认证通过后才能进行记账。

NAS

```

Ruijie#configure terminal
Ruijie(config)#username user password pass
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key test
Ruijie(config)#aaa authentication dot1x autlx group radius local
Ruijie(config)#aaa accounting network acclx start-stop group radius
Ruijie(config)#dot1x authentication autlx
Ruijie(config)#dot1x accounting acclx
Ruijie(config)#interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#dot1 port-control auto
Ruijie(config-if-GigabitEthernet 0/1)#exit

```

【检验方法】 在 NAS 设备上，通过 show 命令查看配置效果。

```
NAS Ruijie#show aaa method-list

Authentication method-list:
aaa authentication dot1x autlx group radius local
Accounting method-list:
aaa accounting network acclx start-stop group radius
Authorization method-list:
```

常见配置错误

无

1.4.4 配置 AAA 服务器组

配置效果

- 创建自定义服务器组，每个服务器组可添加一台或多台服务器。
- 配置认证、授权、记账方法列表时，引用服务器组的组名作为认证、授权、记账方法，则表示在进行认证、授权、记账请求时使用该服务器组中的服务器。
- 使用自定义服务器组可以实现认证、授权、记账相分离。

注意事项

在自定义服务器组中，只能指定并应用默认服务器组中的服务器。

配置方法

📄 创建 AAA 自定义服务器组。

- 必选配置
- 在创建自定义服务组名的时候，组名尽可能有明确的含义。不可以使用预定义的关键字“radius”和“tacacs+”。

📄 添加 AAA 服务器组成员。

- 必选配置
- 使用 sever 命令添加 AAA 服务器组的成员。
- 缺省情况下，自定义组中没有添加服务器。

检验方法

使用命令 **show aaa group** 查看配置的服务器组信息。

相关命令

创建 AAA 自定义服务器组。

【命令格式】 **aaa group server {radius | tacacs+} name**

【参数说明】 *name* : 服务器组的取名, 目前不能为关键字 “radius” , “tacacs+” , 因为这是 RADIUS 和 TACACS+默认的服务器组名称。

【命令模式】 全局模式

【使用指导】 该命令配置 AAA 服务器组, 目前支持 RADIUS 和 TACACS+服务器组。

添加 AAA 服务器组成员。

【命令格式】 **server ip-addr [auth-port port1] [acct-port port2]**

【参数说明】 *ip-addr* : 服务器 ip 地址

port1 : 服务器认证端口 (仅 RADIUS 服务器组支持)

port2 : 服务器记账端口 (仅 RADIUS 服务器组支持)

【命令模式】 服务器组配置模式

【使用指导】 往指定服务器中添加服务器, 不指定端口时使用默认值。

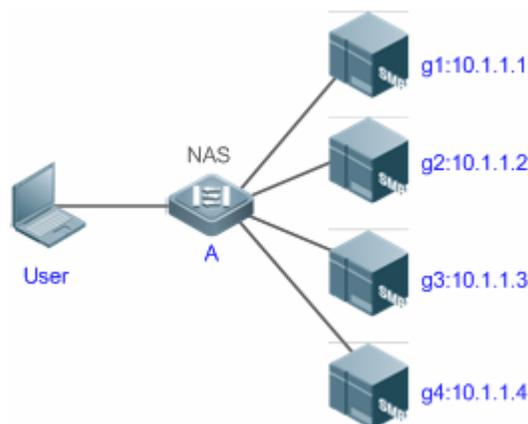
配置举例

i 以下配置举例, 仅介绍与 AAA 服务器组相关的配置。

创建 AAA 自定义服务器组。RADIUS 服务器组 g1、g2, 其中 g1 组的服务器的 IP 为 10.1.1.1 和 10.1.1.2, g2 组的服务器的 IP 为 10.1.1.3 和 10.1.1.4。

【网络环境】

图 1-14



- 【前置任务】
- 1, 网络中已经完成了接口、IP 地址、Vlan 的配置, 网络连通, NAS 设备到服务器的路由可达。
 - 2, 启用 AAA 服务。

【配置方法】 第一步：配置服务器（该服务器属于默认服务器组）

第二步：创建 AAA 自定义服务器组

第三步：在自定义服务器组中添加服务器组成员

NAS

```
Ruijie#configure terminal
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server host 10.1.1.2
Ruijie(config)#radius-server host 10.1.1.3
Ruijie(config)#radius-server host 10.1.1.4
Ruijie(config)#radius-server key secret
Ruijie(config)#aaa group server radius g1
Ruijie(config-gs-radius)#server 10.1.1.1
Ruijie(config-gs-radius)#server 10.1.1.2
Ruijie(config-gs-radius)#exit
Ruijie(config)#aaa group server radius g2
Ruijie(config-gs-radius)#server 10.1.1.3
Ruijie(config-gs-radius)#server 10.1.1.4
Ruijie(config-gs-radius)#exit
```

【检验方法】

在 NAS 设备上，通过 **show aaa group、show run** 命令查看配置效果。

NAS

```
Ruijie#show aaa group
Type      Reference  Name
-----
radius    1         radius
tacacs+   1         tacacs+
radius    1         g1
radius    1         g2

Ruijie#show run
!
radius-server host 10.1.1.1
radius-server host 10.1.1.2
radius-server host 10.1.1.3
radius-server host 10.1.1.4
radius-server key secret
!
aaa group server radius g1
  server 10.1.1.1
  server 10.1.1.2
!
aaa group server radius g2
  server 10.1.1.3
```

```
server 10.1.1.4
!  
!
```

常见配置错误

- 对于使用非默认认证、记账端口的 radius 服务器，在使用命令 server 添加服务器时要同时指定认证端口或记账端口。

1.4.5 配置基于域名的 AAA 服务

配置效果

针对不同域的 802.1x 用户，创建认证、授权和记账方案。

注意事项

关于域中引用方法列表：

- 在域配置模式下，选择 AAA 服务方法列表时，这些方法列表是在进入域配置模式前已经定义；否则在域配置模式下，允许选择 AAA 方法列表名，但提示配置不存在。
- 域选择的 AAA 服务方法列表名称必须和 AAA 服务所定义的方法列表名称必须一致。若不一致，不能够为该域中的用户提供合适的 AAA 服务。

关于缺省域：

- 缺省域（default）：在基于域名的 AAA 服务开关打开情况下，如果用户没有携带域信息，则使用缺省域。如果用户携带的域在系统中没有配置，则判定为非法用户，不提供 AAA 服务。初始时没有配置 default 域，需要手工指定创建。
- 基于域名的 AAA 服务开关打开时，默认情况下没有配置缺省域，需要手动配置完成。缺省域的名称为“default”，若配置缺省域后，用户不携带域信息时，使用缺省域进行提供 AAA 服务。若缺省域没有配置，则未携带域信息的用户不能使用 AAA 服务。

关于域名：

- 用户所携带的域名称与设备上所配置的域名的匹配采用最准确匹配。例如：设备上配置了 domain.com 和 domain.com.cn 两个域，一个用户的请求信息携带为 aaa@domain.com。则设备认为会判定该用户所属于的域为 domain.com 而不是域 domain.com.cn。
- 如果认证用户携带有域信息，而域没有在设备上配置，不能为该用户提供 AAA 服务。

配置方法

📌 开启 AAA。

- 必须配置。

- 使用 **aaa new-model** 开启 AAA。
- 缺省情况下，没有启动 AAA。
- ↳ **开启基于域名的 AAA 服务。**
 - 必选配置。
 - 使用 **aaa domain enable** 开启基于域名的 AAA 服务。
 - 缺省情况下，基于域名的 AAA 服务关闭。
- ↳ **创建域，并进入域配置模式。**
 - 必选配置。
 - 使用 **aaa domain** 命令创建域或者进入已配置的域。
 - 缺省情况下，没有配置任何域。
- ↳ **在域中，关联 802.1X 认证方法列表。**
 - 使用 **authentication dot1x** 命令关联 802.1x 认证方法列表。
 - 如果要在域中应用指定的 801.1x 认证方法，则必须配置此命令。
 - 目前基于域名的 AAA 服务，仅被应用于 802.1x 接入服务。
- ↳ **在域中，关联 Network 记账方法列表。**
 - 使用 **accounting network** 命令关联 network 记账方法列表。
 - 如果要在域中应用指定的记账方法，则必须配置此命令。
 - 如果域中没有关联方法列表，则默认使用全局的 default 方法列表进行记账。
- ↳ **在域中，关联 Network 授权方法列表。**
 - 使用 **authorization network** 命令关联 network 授权方法列表。
 - 如果要在域中应用指定的授权方法，则必须配置此命令。
 - 如果域中没有关联方法列表，则默认使用全局的 default 方法列表进行授权。
- ↳ **设置域的状态。**
 - 可选配置
 - 当域的状态为 block 时，属于该域的用户不能登录。
 - 缺省情况下，当域被创建以后，其状态为 active，即允许任何属于该域的用户请求网络服务。
- ↳ **设置是否在用户名中携带域名信息。**
 - 可选配置
 - 缺省情况下，NAS 与服务器交互时用户名中携带域信息。
- ↳ **设置当前域可容纳接入用户的数目限制。**

- 可选配置
- 缺省情况下，不对当前域可容纳的接入用户数作限制。

检验方法

使用命令 **show aaa domain** 查看配置的域信息是否生效。

相关命令

▾ 开启 AAA。

【命令格式】 **aaa new-model**

【参数说明】 无

【命令模式】 全局模式

【使用指导】 该命令是 AAA 的使能命令，如果您要使用 AAA 安全服务，就必须使用 **aaa new-model** 使能 AAA 安全服务。如果没有启用 AAA，则所有 AAA 命令将是不可配置的。

▾ 开启基于域名的 AAA 服务。

【命令格式】 **aaa domain enable**

【参数说明】 无

【命令模式】 全局模式

【使用指导】 进行基于域名的 AAA 服务配置，需要打开这个配置开关。

▾ 创建域，并进入域配置模式。

【命令格式】 **aaa domain { default | domain-name }**

【参数说明】 **default**：使用该参数，进行缺省域的配置
domain-name：指定域的名称

【命令模式】 全局模式

【使用指导】 指定基于域名的 AAA 服务配置。**default** 为缺省域配置，也就是如果用户没有携带域信息，网络设备所使用的方法列表。*domain-name* 为指定域名配置，如果用户携带该域名，则指定使用这个域所关联的方法列表。目前系统支持最多配置 32 个域。

▾ 在域中，关联 802.1X 认证方法列表。

【命令格式】 **authentication dot1x { default | list-name }**

【参数说明】 **default**：使用该参数，指定使用缺省配置方法列表
list-name：指定方法列表名称

【命令模式】 域配置模式

【使用指导】 为域指定一个 802.1x 认证方法列表。

▾ 在域中，关联 Network 记账方法列表。

【命令格式】 **accounting network { default | list-name }**

- 【参数说明】 **default** : 使用该参数, 指定使用缺省配置方法列表
list-name : 指定方法列表名称
- 【命令模式】 域配置模式
- 【使用指导】 为域指定使用的 Network 记账方法列表。

在域中, 关联 Network 授权方法列表。

- 【命令格式】 **authorization network { default | list-name }**
- 【参数说明】 **default** : 使用该参数, 指定使用缺省配置方法列表
list-name : 指定方法列表名称
- 【命令模式】 域配置模式
- 【使用指导】

设置域的状态。

- 【命令格式】 **state { block | active }**
- 【参数说明】 **block** : 配置的域无效
active : 配置的域有效
- 【命令模式】 域配置模式
- 【使用指导】 指定配置的域是否有效。

设置是否在用户名中携带域名信息。

- 【命令格式】 **username-format { without-domain | with-domain }**
- 【参数说明】 **without-domain** : 剥离域信息
with-domain : 不剥离域信息
- 【命令模式】 域配置模式
- 【使用指导】 在域配置模式下, 配置 NAS 针对指定域与服务器交互时, 用户名中是否携带域信息。

设置当前域可容纳接入用户的数目。

- 【命令格式】 **access-limit num**
- 【参数说明】 *num* : 域用户的数量限制, 只限制 802.1x 用户
- 【命令模式】 域配置模式
- 【使用指导】 使用该命令对域的用户数量进行限制。

配置举例

 以下配置举例, 仅介绍与多域 AAA 相关的配置。

- 在配置基于域的 AAA 认证记账服务。实现使用 RADIUS 服务器对通过 NAS 接入的 802.1X 域用户 (用户名为 `user@domain.com`) 进行认证和记账。NAS 向服务器发送的用户名不携带域名, 不限制接入用户数。

【网络环境】

图 1-15



【配置方法】

本例使用 RADIUS 认证和记账，需要提前配置 RADIUS 服务器。

第一步：开启 AAA

第二步：定义 AAA 服务的方法列表

第三步：开启基于域名的 AAA 服务

第四步：创建域

第五步：在指定域中关联 AAA 方法列表

第六步：设置域属性

NAS

```

Ruijie#configure terminal
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 10.1.1.1
Ruijie(config)#radius-server key test
Ruijie(config)#aaa authentication dot1x default group radius
Ruijie(config)#aaa accounting network list3 start-stop group radius
Ruijie(config)# aaa domain enable
Ruijie(config)# aaa domain domain.com
Ruijie(config-aaa-domain)# authentication dot1x default
Ruijie(config-aaa-domain)# accounting network list3
Ruijie(config-aaa-domain)# username-format without-domain
  
```

【检验方法】

在 NAS 设备上，通过 **show run**、**show aaa domain** 命令查看配置效果。

NAS

```

Ruijie#show aaa domain domain.com

=====Domain domain.com=====
State: Active
Username format: With-domain
Access limit: No limit
802.1X Access statistic: 0

Selected method list:
 authentication dot1x default
 accounting network list3

Ruijie#show run

Building configuration...
Current configuration : 1449 bytes
version RGOS 10.4(3) Release(101069) (Wed Oct 20 09:12:40 CST 2010 -ngcf67)
  
```

```
co-operate enable
!
aaa new-model
aaa domain enable
!
aaa domain domain.com
  authentication dot1x default
  accounting network list3
!
aaa accounting network list3 start-stop group radius
aaa authentication dot1x default group radius
!
nfpp
!
no service password-encryption
!
radius-server host 10.1.1.1
radius-server key test
!
line con 0
line vty 0 4
!
end
```

常见配置错误

无

1.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除被锁定的用户列表。	clear aaa local user lockout {all user-name <i>username</i> }

查看运行情况

作用	命令
----	----

显示记账更新相关的信息。	<u>show aaa accounting update</u>
显示当前所有配置域信息。	<u>show aaa domain</u>
显示当前 login 的锁定配置参数。	<u>show aaa lockout</u>
显示 AAA 配置的所有服务器组。	<u>show aaa group</u>
显示 AAA 所有的方法列表。	<u>show aaa method-lis</u>
显示 AAA 用户相关信息。	<u>show aaa user</u>

查看调试信息

无

2 RADIUS

2.1 概述

RADIUS (Remote Authentication Dial-In User Service, 远程认证拨号用户服务)是一种分布式的客户机/服务器系统。

RADIUS 与 AAA 配合对试图连接的用户进行身份认证,防止未经授权的访问。在 RGOS 的实现中,RADIUS 客户端运行在设备或网络访问服务器 (NAS) 上,并向中央 RADIUS 服务器发出身份认证请求,中央服务器包含了所有的用户身份认证和网络服务信息。除了提供认证服务之外,RADIUS 服务器还提供接入用户的授权和记账的服务。

RADIUS 常应用在既要求较高安全性、又允许远程用户访问的各种网络环境中。由于 RADIUS 是一种完全开放的协议,很多系统如 UNIX、WINDOWS 2000、WINDOWS 2008 等均将 RADIUS 服务器作为一个组件安装,因此 RADIUS 是目前应用最广泛的安全服务器。

RADIUS 动态授权扩展协议 (Dynamic Authorization Extensions to Remote Authentication Dial In User Service),在 IETF 的 RFC3576 中进行定义。该协议定义了一种针对用户下线管理方法。设备和 RADIUS 服务器之间通过 Disconnect-Messages (简称 DM)消息,将已认证通过的用户下线。该协议使得不同厂商间的设备和 RADIUS 服务器,在用户下线的处理上能够兼容。

DM 消息机制,由 RADIUS 服务器主动向设备发起用户下线请求,设备依据请求报文中携带的用户会话、用户名等信息来匹配用户并对其进行下线处理,然后将处理结果以回应报文形式返回给 RADIUS 服务器,以实现服务器对用户的下线管理功能。

协议规范

- RFC2865 : Remote Authentication Dial In User Service (RADIUS)
- RFC2866 : RADIUS Accounting
- RFC2867 : RADIUS Accounting Modifications for Tunnel Protocol Support
- RFC2868 : RADIUS Attributes for Tunnel Protocol Support
- RFC2869 : RADIUS Extensions
- RFC3576 : Dynamic Authorization Extensions to Remote Authentication Dial In User Service (RADIUS)

2.2 典型应用

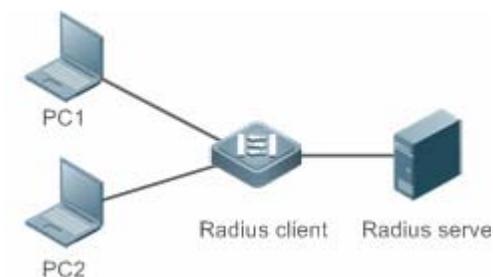
典型应用	场景描述
为接入用户提供认证、授权、记账服务	对网络中的接入用户进行认证、授权、记账,以防止未经授权的访问或操作。
已上线用户被服务器强制下线	对于已经认证的用户,服务器强制其下线

2.2.1 为接入用户提供认证、授权、记账服务

应用场景

RADIUS 的典型应用为对接入用户进行认证、授权、记账。网络设备作为 RADIUS 客户端，将用户信息发送给 RADIUS 服务器。RADIUS 服务器处理完成，给 RADIUS 客户端返回认证接受/认证拒绝/记账响应等信息。RADIUS 客户端根据 RADIUS 服务器的响应信息对接入用户进行相应处理。

图 2-1 典型的 RADIUS 网络配置



【注释】 PC1 和 PC2 作为接入用户通过有线或者无线方式和 RADIUS 客户端连接，并发起认证、记账请求。

Radius Client 通常为接入交换机或者汇聚交换机。

Radius Server 可以是 Windows 2000/2003 Server (IAS)、UNIX 系统所带组件，也可以是一些厂商提供的专用服务器软件。

功能部署

- 在 Radius Server 配置接入设备信息，包括接入设备 IP，共享密钥等。
- 在 Radius Client 配置 AAA 的认证、授权、记账方法列表。
- 在 Radius Client 配置 Radius server 信息，包括 IP，共享密钥等。
- 在 Radius Client 配置接入端口开启访问控制。
- 配置网络，使 Radius Client 和 Radius Server 之间通讯正常。

2.2.2 用户强制下线

应用场景

出于管理需要，RADIUS 服务器对于已经认证上线的用户，采取强制下线的措施。

网络配置请参考图 1-1

功能部署

- 在 1.2.1 的功能部署基础上加上以下部署：
- 在 Radius Client 使能 RADIUS 动态授权扩展功能

2.3 功能详解

基本概念

客户端/服务器模式

- 客户端：RADIUS 客户端作为 RADIUS 请求的发起端，通常运行在设备或者网络访问服务器(NAS)上，负责把用户信息发送给 RADIUS 服务器，并接受 RADIUS 服务器的返回信息，进行相应的处理。处理包括接受用户接入或者拒绝用户接入或者收集更多用户信息提供给服务器进行处理。
- 服务器：RADIUS 客户端和 RADIUS 服务器通常是多对一的关系。RADIUS 服务器维护所有的 RADIUS 客户端的 IP 和共享密钥信息，以及所有认证用户的信息。RADIUS 服务器接收 RADIUS 客户端的请求信息，并进行认证、授权、记账处理，然后返回客户端需要的认证、授权、记账信息。

RADIUS 报文结构

RADIUS 的报文结构如下图所示：

8	16	32bit
Code	Identifier	Length
Authenticator(16bytes)		
Attributes		

- Code — Code 域长度为一个字节，用于标识 RADIUS 报文的类型，取值及含义参考下表。

Code	报文类型	Code	报文类型
1	Access-Request 认证请求报文	4	Accounting-Request 记账请求报文
2	Access-Accept 认证接受报文	5	Accounting-Response 记账相应报文
3	Access-Reject 认证拒绝报文	11	Access-Challenge 认证质询报文

- Identifier — Identifier 域占用 1 个字节，用于匹配请求和响应报文。同一类型的请求报文和响应报文的 Identifier 值相同。
- Length — Length 域占用 2 个字节，标识整个 RADIUS 报文的长度，包括 Code、Identifier、Length、Authenticator、Attributes 在内。超过 Length 域的字节将被忽略。如果接收到的报文的实际长度小于 Length 的值，则丢弃该报文。
- Authenticator — Authenticator 域占用 16 个字节。RADIUS 客户端使用该域来验证服务器的回应报文。Authenticator 域也用于用户密码的加密/解密。

- Attributes — Attributes 域的长度是不定的，用于携带认证、授权、记账信息。Attributes 域通常包含多个属性。每个属性采用 TLV(Type、Length、Value)三元组的结构表示。其中，Type 为 1 个字节，表示属性的类型，下表列出了 RADIUS 认证、授权、记账常用的属性；Length 为 1 个字节，表示该属性的长度，单位为字节；Value 为该属性的信息。

属性号	属性名	属性号	属性名
1	User-Name	43	Acct-Output-Octets
2	User-Password	44	Acct-Session-Id
3	CHAP-Password	45	Acct-Authentic
4	NAS-IP-Address	46	Acct-Session-Time
5	NAS-Port	47	Acct-Input-Packets
6	Service-Type	48	Acct-Output-Packets
7	Framed-Protocol	49	Acct-Terminate-Cause
8	Framed-IP-Address	50	Acct-Multi-Session-Id
9	Framed-IP-Netmask	51	Acct-Link-Count
10	Framed-Routing	52	Acct-Input-Gigawords
11	Filter-ID	53	Acct-Output-Gigawords
12	Framed-MTU	55	Event-Timestamp
13	Framed-Compression	60	CHAP-Challenge
14	Login-IP-Host	61	NAS-Port-Type
15	Login-Service	62	Port-Limit
16	Login-TCP-Port	63	Login-LAT-Port
18	Reply-Message	64	Tunnel-Type
19	Callback-Number	65	Tunnel-Medium-Type
20	Callback-ID	66	Tunnel-Client-Endpoint
22	Framed-Route	67	Tunnel-Server-Endpoint
23	Framed-IPX-Network	68	Acct-Tunnel-Connection
24	State	69	Tunnel-Password
25	Class	70	ARAP-Password
26	Vendor-Specific	71	ARAP-Features
27	Session-Timeout	72	ARAP-Zone-Access
28	Idle-Timeout	73	ARAP-Security
29	Termination-Action	74	ARAP-Security-Data
30	Called-Station-Id	75	Password-Retry
31	Calling-Station-Id	76	Prompt
32	NAS-Identifier	77	Connect-Info
33	Proxy-State	78	Configuration-Token
34	Login-LAT-Service	79	EAP-Message
35	Login-LAT-Node	80	Message-Authenticator
36	Login-LAT-Group	81	Tunnel-Private-Group-id
37	Framed-AppleTalk-Link	82	Tunnel-Assignment-id
38	Framed-AppleTalk-Network	83	Tunnel-Preference

39	Framed-AppleTalk-Zone	84	ARAP-Challenge-Response
40	Acct-Status-Type	85	Acct-Interim-Interval
41	Acct-Delay-Time	86	Acct-Tunnel-Packets-Lost
42	Acct-Input-Octets	87	NAS-Port-Id

共享密钥

RADIUS 客户端和 RADIUS 服务器进行通讯，相互之间通过共享密钥来确定对方的身份。共享密钥不能通过网络传输。此外，在传输过程中，为保证安全性，用户密码都是加密的。

RADIUS 服务器组

RADIUS 安全协议，也称 RADIUS 方法，是以 RADIUS 服务器组为单位进行配置的。每一个 RADIUS 方法对应一个 RADIUS 服务器组，每一个 RADIUS 服务器组可配置一至多台 RADIUS 服务器（关于使用 RADIUS 方法的细节信息，请参见“AAA 配置”章节）。如果您在一个 RADIUS 服务器组中配置了多台 RADIUS 服务器，那么当设备同第一台 RADIUS 服务器通讯失败，或者第一台 RADIUS 服务器变成不可达的状态时，设备将自动尝试同第二台 RADIUS 服务器通讯，以此类推，直到成功或者全部失败为止。

RADIUS 属性类型

- 标准属性
- RFC 相关标准规定了 RADIUS 的属性号和属性的内容，但是对于某些属性类型，没有规定属性内容的格式。因此，为适应不同的 RADIUS 服务器要求，需要配置属性内容的格式。目前支持设置 RADIUS Calling-Station-ID 属性（属性号为 31）。

RADIUS Calling-Station-ID 属性用于网络设备向 RADIUS Server 发送请求报文时候，标识认证用户的身份。Calling-Station-ID 属性内容是字符串，可以有多种组成格式，由于要求必须能唯一标识一个用户，因此常选择使用用户的 MAC 地址作为其内容。例如在使用 IEEE 802.1X 认证时，选择使用安装 IEEE 802.1X 客户端所在设备的 MAC 地址。关于这 MAC 地址的格式，有以下几种：

格式	说明
ietf	IETF (RFC3580) 规定的标准格式，使用 ‘-’ 作为分隔符。例如： 00-D0-F8-33-22-AC
normal	常用的表示 MAC 地址的格式（点分十六进制格式），使用 ‘.’ 作为分隔符。例如： 00d0.f833.22ac
unformatted	无格式，没有任何分隔符，默认使用这个格式。例如： 00d0f83322ac

- 私有属性

RADIUS 协议是一个可扩展的协议。RFC2865 中定义了 26 号属性（Vendor-Specific）用于设备厂商对 RADIUS 协议进行扩展，以实现其私有的或者标准 RADIUS 没有定义的功能。锐捷公司支持的私有属性如表 1-3 所示。其中 TYPE 为锐捷产品私有属性类型的默认配置；扩展 TYPE 为扩展厂商类型的默认配置。

ID	功能	TYPE	扩展 TYPE
1	max-down-rate	1	76
2	port-priority	2	77

3	user-ip	3	3
4	vlan-id	4	4
5	last-supplicant-version	5	5
6	net-ip	6	6
7	user-name	7	7
8	password	8	8
9	file-directory	9	9
10	file-count	10	10
11	file-name-0	11	11
12	file-name-1	12	12
13	file-name-2	13	13
14	file-name-3	14	14
15	file-name-4	15	15
16	max-up-rate	16	16
17	current-supplicant-version	17	17
18	flux-max-high32	18	18
19	flux-max-low32	19	19
20	proxy-avoid	20	20
21	dailup-avoid	21	21
22	ip-privilege	22	22
23	login-privilege	42	42
27	ipv4-multicast-address	87	87
62	sdg-type	62	62
85	sdg-zone-name	85	85
103	sdg-group-name	103	103

功能特性

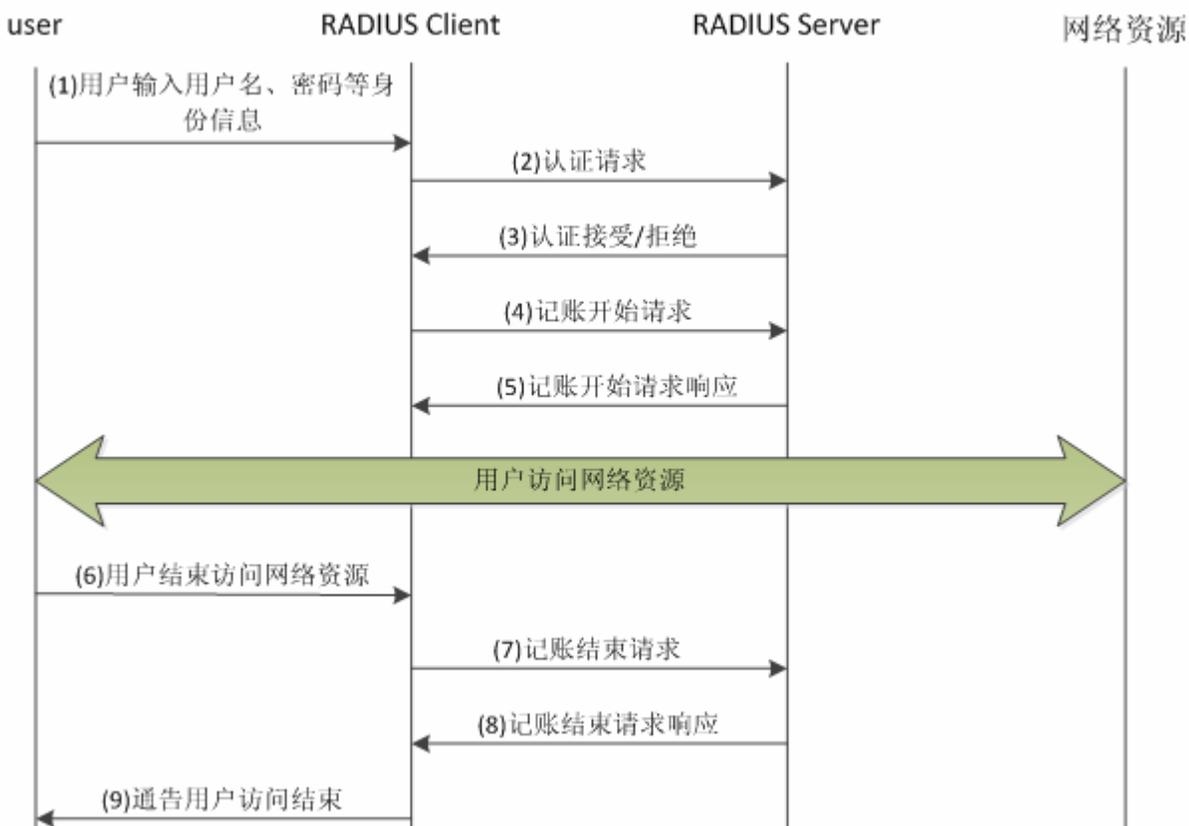
功能特性	作用
RADIUS认证、授权	对访问用户进行身份认证、记账，保护网络安全以及便于网络管理员进行管理。
指定RADIUS报文源地址	指定 RADIUS 客户端向 RADIUS 服务器传送报文时的源 IP 地址。
RADIUS超时重传	指定 RADIUS 服务器对 RADIUS 客户端传送的报文一定的时间内无响应时 RADIUS 客户端重传报文的参数。
RADIUS服务器可达性检测	RADIUS 客户端主动探测 RADIUS 服务器是否可达，并维护各 RADIUS 服务器的可达性状态。进行业务处理时，总是优先选择状态为可达的服务器，以提高 RADIUS 业务的处理性能。
RADIUS强制下线	对于已认证的用户，RADIUS 服务器主动要求其下线

2.3.1 RADIUS 认证、授权、记账

对访问用户进行身份认证、记账，保护网络安全以及便于网络管理员进行管理。

工作原理

图 2-2



RADIUS 的认证和授权流程为：

- (1) 用户输入用户名、密码等身份信息，传送给 RADIUS 客户端。
- (2) RADIUS 客户端获取用户的用户名、密码信息，向 RADIUS 传送认证请求报文。其中密码是加密的，加密方法请参照 RFC2865。
- (3) RADIUS 服务器根据用户名、密码信息，决定接受或拒绝此次认证请求。如果接受，同时下发授权信息。不同类型的访问用户，其授权信息也不相同。

RADIUS 的记账流程为：

- (4) 如果步骤(3)中 RADIUS 服务器返回认证接受，则 RADIUS 客户端紧接着发送记账开始请求报文。
- (5) RADIUS 服务器回应记账开始响应报文，开始记账。
- (6) 用户结束访问网络资源，请求 RADIUS 客户端断开连接。
- (7) RADIUS 客户端发送记账结束请求报文。
- (8) RADIUS 服务器返回记账结束响应报文，停止记账。

(9)用户断开连接，无法再访问网络资源

相关配置

配置 RADIUS 服务器参数

缺省情况下，没有配置任何 RADIUS 服务器。

使用 **radius-server host** 命令可以配置 RADIUS 服务器的相关信息。

必须至少配置一个 RADIUS 服务器，RADIUS 相关业务才能正常运转。

配置 AAA 认证方法列表

缺省情况下，没有配置任何 AAA 认证方法列表。

使用 **aaa authentication** 命令配置不同用户类型的方法列表，并且认证方法选择 group radius。

必须配置相应用户类型的 aaa 认证方法列表，才能进行 radius 认证。

配置 AAA 授权方法列表

缺省情况下，没有配置任何 AAA 授权方法列表。

使用 **aaa authorization** 命令配置不同类型的授权方法列表，并且授权方法选择 group radius。

必须配置相应类型的 aaa 授权方法列表，才能进行 radius 授权。

配置 AAA 记账方法列表

缺省情况下，没有配置任何 AAA 记账方法列表。

使用 **aaa accounting** 命令配置不同类型的记账方法列表，并且记账方法选择 group radius。

必须配置相应类型的 aaa 记账方法列表，才能进行 radius 记账。

2.3.2 指定 RADIUS 报文源地址

指定 RADIUS 客户端向 RADIUS 服务器传送报文时的源 IP 地址。

工作原理

配置 RADIUS 时，通过指定 RADIUS 客户端向 RADIUS 服务器发送 RADIUS 报文的源 IP 地址，可以减少在 RADIUS 服务器上维护大量的 NAS 信息的工作量。

相关配置

缺省配置为使用全局路由寻路，确定发送 RADIUS 报文的源地址。

使用 `ip radius source-interface` 命令指定发送 RADIUS 报文的源接口，设备将把指定接口的第一个 ip 地址作为 radius 报文的源地址。

2.3.3 RADIUS 超时重传

工作原理

RADIUS 客户端向 RADIUS 服务器传送报文后，启动定时器检测 RADIUS 服务器的响应，如果一定时间内 RADIUS 服务器没有响应，则 RADIUS 客户端重传报文。

相关配置

配置 RADIUS 服务器超时时间

缺省配置的超时时间为 5 秒。

使用命令 `radius-server timeout` 命令可以配置超时时间，时间范围为 1 到 1000 秒。

RADIUS 服务器的响应时间和其自身的性能、网络环境有关。需要根据实际情况配置合适的超时时间。

配置重传次数

缺省配置的重传次数为 3 次。

使用命令 `radius-server retransmit` 命令配置重传次数，范围 1 到 100 次。

配置记账更新是否重传

缺省配置为不会对计费更新报文进行重传。

使用命令 `radius-server account update retransmit` 命令配置二代 Web 认证用户的记账更新报文进行重传的功能。

2.3.4 RADIUS 服务器可达性检测

工作原理

RADIUS 客户端主动探测 RADIUS 服务器是否可达，并维护各 RADIUS 服务器的可达性状态。进行业务处理时，总是优先选择状态为可达的服务器，以提高 RADIUS 业务的处理性能。

相关配置

配置设备判定 RADIUS 安全服务器不可达的标准

缺省配置的判定 RADIUS 服务器不可达的标准为同时满足以下两个条件：一、设备在 60 秒内没有收到来自 RADIUS 安全服务器的正确响应报文；二、设备向同一个 RADIUS 安全服务器发送的请求报文连续超时次数达到 10 次。

使用命令 `radius-server dead-criteria` 可以配置设备判定 RADIUS 安全服务器不可达的标准。

配置主动探测 RADIUS 安全服务器的测试用户名

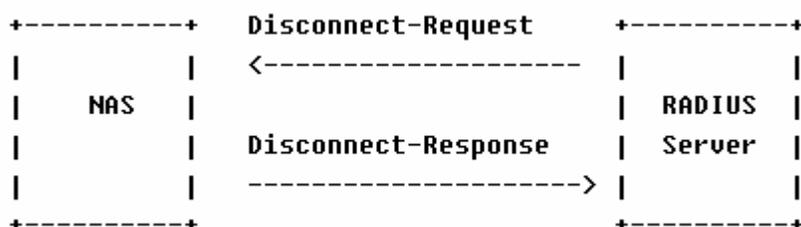
缺省配置时，不对 RADIUS 安全服务器指定主动探测的测试用户名。

使用命令 `radius-server host x.x.x.x testusername xxx` 来配置测试用户名。

2.3.5 RADIUS 强制下线

工作原理

图 2-3 RADIUS 动态授权扩展 DM 消息交互图



RADIUS 服务器和设备之间的 DM 消息交互图如上。RADIUS 服务器发送 Disconnect-Request 消息到设备的 3799 UDP 端口，设备处理结束之后，将处理结果通过 Disconnect-Response 消息返回给 RADIUS 服务器。

相关配置

2.4 配置详解

配置项	配置建议 & 相关命令	
RADIUS基本配置	必须配置。用于 radius 认证、授权、记账	
	<code>radius-server host</code>	配置远程 RADIUS 安全服务器的 IP 地址
	<code>radius-server key</code>	配置设备和 RADIUS 服务器进行通讯的共享密钥
	<code>radius-server retransmit</code>	配置设备在确认 RADIUS 无效以前发送请求的次数
	<code>radius-server timeout</code>	配置设备重传请求以前等待的时间
	<code>radius-server account update retransmit</code>	配置二代 Web 认证用户的记账更新报文进行重传的功能
	<code>ip radius source-interface</code>	配置 RADIUS 报文的源地址
配置RADIUS属性类型	可选配置。用于定义设备封装和解析 radius 报文时对属性的处理。	
	<code>radius-server attribute 31</code>	配置 RADIUS 的 31 号属性 (Calling-Station-ID) 的 MAC 地址格式。

	radius attribute	配置 RADIUS 私有属性类型
	radius set qos cos	配置设备处理服务器下发的私有属性 port-priority 为接口 cos 值。Cos 相关概念请参考“配置 Qos”
	radius support cui	配置设备支持 cui 属性
	radius vendor-specific	配置设备解析私有属性的方式
配置RADIUS可达性检测	 可选配置。用于检测 RADIUS 服务器是否可达，以及维护 RADIUS 服务器的可达性状态。	
	radius-server dead-criteria	配置全局的 RADIUS 安全服务器不可达的判定标准
	radius-server deadtime	配置设备停止向不可达状态的 RADIUS 服务器发送请求报文的时长
	radius-server host	配置远程 RADIUS 安全服务器的 IP 地址,指定认证端口和记帐端口,指定主动探测的相关参数

2.4.1 RADIUS 基本配置

配置效果

- 完成 RADIUS 基本配置，即可进行 RADIUS 认证、授权、记账。

注意事项

- 在设备上配置 RADIUS 之前，应确保 RADIUS 服务器的网络通讯良好。
- 使用命令 **ip radius source-interface** 配置 RADIUS 报文的源地址时，应确保此源地址和 RADIUS 服务器通讯良好。

配置方法

配置远程 RADIUS 安全服务器

- 必须配置。
- 配置 RADIUS 安全服务器的 IP 地址、认证端口、记账端口、共享密钥。

配置设备和 RADIUS 服务器进行通讯的共享密钥。

- 可选配置。
- 这里通过全局配置对所有未配置共享密钥选项的服务器配置一个共享密钥。

 设备上的共享密钥和 RADIUS 服务器上的共享密钥必须一致。

配置设备在确认 RADIUS 无效以前发送请求的次数

- 可选配置。
- 根据实际网络环境，配置设备确认 RADIUS 无效以前发送请求的次数。

配置设备重传请求以前等待的时间

- 可选配置。
- 根据实际网络环境，配置设备重传请求以前等待的时间。

! 在使用 RADIUS 安全协议的 802.1x 认证环境中，如果网络设备作为 802.1x 认证者，并且采用锐捷 SU 作为 802.1x 客户端软件时，建议在网络设备上设置 **radius-server timeout** 值为 3 秒（默认为 5 秒），设置 **radius-server retransmit** 值为 2 次（默认为 3 次）

配置二代 Web 认证用户的记账更新报文进行重传的功能

- 可选配置。
- 根据实际需要，决定是否开启二代 Web 认证用户的记账更新报文重传功能。

配置 RADIUS 报文的源地址

- 可选配置。
- 根据实际网络环境，配置 RADIUS 报文的源地址。

检验方法

- 配置 AAA 方法列表使用 RADIUS 方法，用户进行认证、授权、记账。
- 设备与 RADIUS 服务器进行交互，通过抓包可以看到是通过 RADIUS 协议进行通信的。

相关命令

配置远程 RADIUS 安全服务器

【命令格式】 `radius-server host [oob [via mgmt_name]] { ipv4-address } [auth-port port-number] [acct-port port-number] [test username name [idle-time time] [ignore-auth-port] [ignore-acct-port]] [key [0 | 7] text-string]`

【参数说明】

- oob** : oob 认证，即向此服务器发送报文时源接口为 mgmt 口。
- via mgmt_name**: oob 多个 mgmt 口时，指定具体的 mgmt 口
- ipv4-address** : RADIUS 安全服务器主机的 IPv4 地址。
- auth-port port-number** : RADIUS 身份认证的 UDP 端口，取值范围 0 - 65535，如果设置为 0，则该主机不进行身份认证。
- acct-port port-number** : RADIUS 记帐的 UDP 端口，取值范围 0 - 65535，如果设置为 0，则该主机不进行记帐。
- test username name** : 开启对该 RADIUS 安全服务器的主动探测功能，并指定主动探测所使用的用户名。
- idle-time time** : 配置设备向处于可达状态的 RADIUS 安全服务器发送测试报文的时间间隔。默认值为 60 分钟，可配置的范围为 1-1440 分钟（24 小时）。
- ignore-auth-port** : 关闭对 RADIUS 安全服务器的认证端口的检测，默认开启。
- ignore-acct-port** : 关闭对 RADIUS 安全服务器的记账端口的检测，默认开启。

key [0 | 7] text-string : 配置用于该服务器的共享密钥, 未配置则使用全局配置。配置的密钥可以指定加密类型, 0 为无加密, 7 简单加密, 默认为 0。

【命令模式】 全局模式。

【使用指导】 为了使用 RADIUS 实现 AAA 安全服务, 必须定义 RADIUS 安全服务器。您可以使用 **radius-server host** 命令定义一个或多个 RADIUS 安全服务器。如果没有把 RADIUS 安全服务器配置在某个 RADIUS 服务器组中, 则设备向 RADIUS 服务器发送 radius 报文时使用全局路由表。

配置设备和 RADIUS 服务器进行通讯的共享密钥。

【命令格式】 **radius-server key [0 | 7] text-string**

【参数说明】 *text-string* : 共享密钥的文本。

0 | 7 : 口令的加密类型, 0 无加密, 7 简单加密, 默认为 0。

【命令模式】 全局模式

【使用指导】 共享密钥是设备和 RADIUS 安全服务器进行正确通信的基础。为了使设备和 RADIUS 安全服务器能进行通信, 必须在设备和 RADIUS 安全服务器上定义相同的共享密钥。

配置设备在确认 RADIUS 无效以前发送请求的次数

【命令格式】 **radius-server retransmit retries**

【参数说明】 *retries* : RADIUS 尝试重发次数, 取值范围是 1-100

【命令模式】 全局模式

【使用指导】 AAA 在使用下一个方法对用户进行认证的前提是当前认证的安全服务器没有反应。设备判断安全服务器没有反应的标准是安全服务器在设备重发指定次数 RADIUS 报文期间均没有应答, 每次重发之间有超时间隔。

配置设备重传请求以前等待的时间

【命令格式】 **radius-server timeout seconds**

【参数说明】 *seconds* : 超时时间 (单位为秒)。可设置的值范围为 1-1000 秒。

【命令模式】 全局模式

【使用指导】 使用该命令对重发报文的超时时间进行调整。

配置二代 Web 认证用户的记账更新报文进行重传的功能

【命令格式】 **radius-server account update retransmit**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 配置二代 Web 认证用户的记账更新报文进行重传的功能, 默认不重传。该配置不影响其他类型的用户。

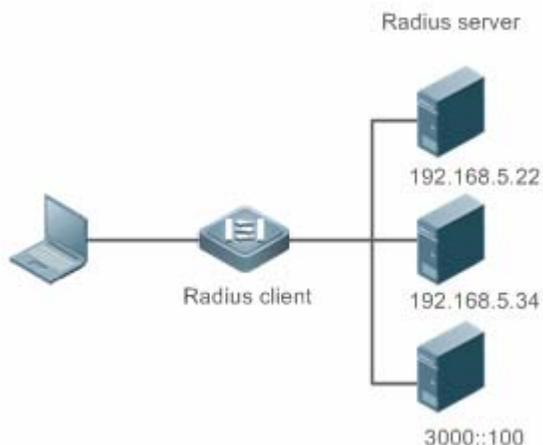
配置举例

 以下配置举例, 仅介绍与 RADIUS 相关的配置。

Login 用户使用 RADIUS 认证、授权、记账

【网络环境】

图 2-4



【配置方法】

- 配置启用 aaa。
- 配置 radius-server 信息。
- 配置使用 radius 的认证方法、授权方法、记账方法。
- 在接口上应用配置认证方法。

RADIUS Client

```

Ruijie# configure terminal
Ruijie (config)# aaa new-model
Ruijie (config)# radius-server host 192.168.5.22
Ruijie (config)# radius-server host 3000::100
Ruijie (config)# radius-server key aaa
Ruijie (config)# aaa authentication login test group radius
Ruijie (config)# aaa authorization exec test group radius
Ruijie (config)# aaa accounting exec test start-stop group radius
Ruijie (config)# line vty 0 4
Ruijie (config-line)# login authentication test
Ruijie (config-line)# authorization exec test
Ruijie (config-line)# accounting exec test
  
```

【检验方法】

在 PC 上 telnet 到设备上，要求输入用户名和密码。输入正确的用户名和密码，能够登录到设备上。并且被服务器授予一定的权限级别，仅运行执行该权限级别下的命令。在 RADIUS 服务器上可以查看到此用户的认证日志。用户对设备进行管理操作后退出登录，在 RADIUS 服务器上可以查看到此用户的记账信息。

```

Ruijie#show running-config
!
radius-server host 192.168.5.22
radius-server host 3000::100
radius-server key aaa
aaa new-model
aaa accounting exec test start-stop group radius
aaa authorization exec test group radius
  
```

```
aaa authentication login test group radius
no service password-encryption
ip tcp not-send-rst
!
vlan 1
!
line con 0
line vty 0 4
  accounting exec test
  authorization exec test
  login authentication test
!
```

常见错误

- 设备配置的 key 与服务器配置的 key 不一致。
- 没有配置方法列表。

2.4.2 配置 RADIUS 属性类型

配置效果

- 定义设备封装和解析 radius 报文时对属性的处理。

注意事项

- 设置 RADIUS 属性类型一节所涉及的私有属性均指锐捷公司的私有属性。

配置方法

📌 配置 RADIUS 的 31 号属性 (Calling-Station-ID) 的 MAC 地址格式

- 可选配置
- 根据服务器类型，配置 Calling-Station-Id 的 MAC 地址格式为服务器支持的类型。

📌 配置 RADIUS 私有属性类型

- 可选配置
- 如果服务器为锐捷公司的应用服务器，则需要配置 RADIUS 私有属性类型来适应。

📌 配置设备处理服务器下发的私有属性 port-priority 为接口的 cos 值

- 可选配置
- 根据需要，配置服务器下发的私有属性 port-priority 为接口的 cos 值。

配置设备支持 cui 属性

- 可选配置
- 根据需要，配置设备是否支持 RADIUS 的 CUI 属性。

配置设备解析私有属性的方式

- 可选配置
- 根据需要，配置设备解析锐捷私有属性时私有属性号的索引。

检验方法

- 配置 AAA 方法列表使用 RADIUS 方法，用户进行认证、授权、记账
- 设备与 RADIUS 服务器进行交互，通过抓包查看 Calling-Station-Id 的 MAC 地址格式。
- 设备与 RADIUS 服务器进行交互，通过设备 debug 信息查看锐捷公司的私有属性被设备正确的解析。
- 设备与 RADIUS 服务器进行交互，通过设备 debug 信息查看 CUI 属性被设备正确的解析。

相关命令

配置 RADIUS 的 31 号属性 (Calling-Station-ID) 的 MAC 地址格式

【命令格式】 **radius-server attribute 31 mac format { ietf | normal | unformatted }**

【参数说明】 **ietf** : 指定 ETF (RFC3580) 规定的标准格式，使用 '-' 作为分隔符。例如：00-D0-F8-33-22-AC。
normal : 指定常用的表示 MAC 地址的格式 (点分十六进制格式)，使用 '.' 作为分隔符。例如 :00d0.f833.22ac。
unformatted : 指定无格式，没有任何分隔符，默认使用这个格式。例如：00d0f83322ac。

【命令模式】 全局模式

【使用指导】 部分 RADIUS 安全服务器 (主要用于 802.1x 认证) 可能只识别 IETF 的格式，这种情况下需要将 Calling-Station-ID 属性设置为 IETF 格式类型。

配置 RADIUS 私有属性类型

【命令格式】 **radius attribute { id | down-rate-limit | dscp | mac-limit | up-rate-limit } vendor-type type**

【参数说明】 **id** : 功能 id <1-255>
type : 私有属性 type
down-rate-limit : 下行速率限制属性
dscp : dscp 属性
mac-limit : mac-limit 属性
up-rate-limit : 上行速率限制属性

【命令模式】 全局模式

【使用指导】 使用该命令配置私有属性类型值。

配置设备处理服务器下发的私有属性 port-priority 为接口的 cos 值

【命令格式】 **radius set qos cos**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 配置该命令，可以将传下的 qos 值作为 cos 值，默认时作为 dscp 值。

配置设备支持 cui 属性

【命令格式】 **radius support cui**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 配置该命令，使 radius 支持 cui 属性。

配置设备解析私有属性的方式

【命令格式】 **radius vendor-specific extend**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 使用该命令可对所有厂商 id 的属性按照配置是由类型识别。

配置举例

 以下配置举例，仅介绍与 RADIUS 相关的配置。

配置 RADIUS 属性类型

【网络环境】 单设备

- 【配置方法】
- 配置 RADIUS 的 Calling-Station-Id 的 MAC 地址格式。
 - 配置私有属性类型值。
 - 配置 radius 传下的 qos 值为接口 cos 值。
 - 配置 radius 支持 cui 属性。
 - 扩展为不区别私有厂商 id。

```
Ruijie(config)# radius-server attribute 31 mac format ietf
Ruijie(config)# radius attribute 16 vendor-type 211
Ruijie(config)# radius set qos cos
Ruijie(config)# radius support cui
Ruijie(config)# radius vendor-specific extend
```

【检验方法】 通过抓包或者设备 debug 信息查看 radius 标准属性和私有属性的封装/解析是否正确。

2.4.3 配置 RADIUS 可达性检测

配置效果

设备维护所配置的每台 RADIUS 服务器的可达性状态：可达或者不可达。设备不会向处于不可达状态的 RADIUS 服务器发送接入用户的认证、授权和记账请求，除非，该 RADIUS 服务器所在的 RADIUS 服务器组的所有服务器均为不可达状态。

设备支持对指定的 RADIUS 服务器进行主动探测，默认关闭。如果您为指定的 RADIUS 服务器开启主动探测功能，那么设备将会根据配置，定期向该 RADIUS 服务器发送探测请求（认证请求或者记账请求）。其时间间隔周期为：

- 处于可达状态的 RADIUS 服务器：该 RADIUS 服务器的可达状态的主动探测间隔时间（默认值为 60 分钟）。
- 处于不可达状态的 RADIUS 服务器：固定为 1 分钟。

注意事项

为指定的 RADIUS 服务器开启主动探测功能，需要满足如下所有条件：

- 在设备上配置了该 RADIUS 服务器的测试用户名。
- 在设备上至少配置了一个该 RADIUS 服务器的被测端口（认证端口或者记账端口）。

对于一台处于可达状态的 RADIUS 服务器，当以下两个条件均满足时，设备认为该 RADIUS 服务器进入不可达状态：

- 距离上次收到该 RADIUS 服务器的正确响应超过 `radius-server dead-criteria time seconds` 设定的时间。
- 在上次收到该 RADIUS 服务器的正确响应之后，设备发往该 RADIUS 服务器的请求而未收到正确响应的次数（包括重传），达到 `radius-server dead-criteria tries number` 设定的次数。

对于一台处于不可达状态的 RADIUS 服务器，当以下任一条件满足时，设备认为该 RADIUS 服务器进入可达状态：

- 设备收到来自该 RADIUS 服务器的正确响应。
- 该 RADIUS 服务器处于不可达状态超过 `radius-server deadtime` 设定的时间，并且该 RADIUS 服务器没有启用主动探测功能。
- 在设备上更新该 RADIUS 服务器的认证端口或者记账端口。

配置方法

配置全局的 RADIUS 安全服务器不可达的判定标准

- 必须配置
- 配置全局的 RADIUS 安全服务器不可达的判定标准是开启主动探测功能的必要条件。

配置远程 RADIUS 安全服务器的 IP 地址，指定认证端口和记帐端口，指定主动探测的相关参数

- 必须配置

- 指定 RADIUS 服务器主动探测的相关参数是开启主动探测功能的必要条件。

配置设备停止向不可达状态的 RADIUS 服务器发送请求报文的时长

- 可选配置
- RADIUS 服务器没有启用主动探测功能时,配置设备停止向不可达状态的 RADIUS 服务器发送请求报文的时长才会生效。

检验方法

- 通过 `show radius server` 命令可以查看各个 RADIUS 服务器的可达性信息。

相关命令

配置全局的 RADIUS 安全服务器不可达的判定标准

【命令格式】 `radius-server dead-criteria { time seconds [tries number] | tries number }`

【参数说明】 `time seconds`: 配置时间条件参数。设备在指定的时间内没有收到来自 RADIUS 安全服务器的正确响应报文,则认为该 RADIUS 安全服务器满足不可达的时长条件。可设置的值的范围为 1-120 秒。

`tries number`: 配置请求连续超时次数条件参数。当设备向同一个 RADIUS 安全服务器发送的请求报文连续超时次数达到所设定的次数,则认为该 RADIUS 安全服务器满足不可达的连续超时次数条件。可设置的值的范围为 1-100。

【命令模式】 全局模式

【使用指导】 如果一台 RADIUS 安全服务器同时满足时间条件和请求连续超时次数条件,则设备认为该 RADIUS 安全服务器不可达。使用该命令,用户可以对时间条件和请求连续超时次数条件的参数进行调整。

配置设备停止向不可达状态的 RADIUS 服务器发送请求报文的时长

【命令格式】 `radius-server deadtime minutes`

【参数说明】 `minutes`: 配置设备停止向处于不可达状态的 RADIUS 安全服务器发送请求的时间,单位为分钟。可设置的值的范围为 1-1440 分钟(24 小时)。

【命令模式】 全局模式

【使用指导】 如果设备对一台 RADIUS 安全服务器启用了主动探测功能,那么 `radius-server deadtime` 的时间参数对该 RADIUS 安全服务器不起作用;否则,该 RADIUS 安全服务器将在处于不可达状态的时间超过 `radius-server deadtime` 指定的时间时,被设备自动恢复为可达状态。

配置举例

 以下配置举例,仅介绍与 RADIUS 相关的配置。

配置对 RADIUS 服务器进行不可达检测

【网络环境】

图 2-5



【配置方法】

- 配置全局的 RADIUS 安全服务器不可达的判定标准。
- 配置远程 RADIUS 安全服务器的 IP 地址，指定认证端口和记帐端口，指定主动探测的相关参数。

RADIUS Client

```
Ruijie(config)# radius-server dead-criteria time 120 tries 5
Ruijie(config)# radius-server host 192.168.5.22 test username test ignore-acct-port idle-time 90
```

【检验方法】

使设备与 192.168.5.22 服务器网络通讯断开。通过设备进行 radius 认证。120 秒后，使用命令 **show radius server** 命令查看服务器状态为 dead。

```
Ruijie#show running-config
...
radius-server host 192.168.5.22 test username test ignore-acct-port idle-time 90
radius-server dead-criteria time 120 tries 5
...
```

2.5 监视与维护

清除各类信息

! 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
将 RADIUS 动态授权扩展功能的统计信息清零，重新开始统计。	clear radius dynamic-authorization-extension statistics

查看运行情况

作用	命令
显示 RADIUS 服务器全局参数。	show radius parameter
显示 RADIUS 服务器配置情况。	show radius server
显示 RADIUS 私有属性类型配置。	show radius vendor-specific
显示 RADIUS 动态授权扩展相关统计信息。	show radius dynamic-authorization-extension statistics
显示 RADIUS 认证相关统计信息	show radius auth statistics

显示 RADIUS 计费相关统计信息	show radius acct statistics
显示 RADIUS 服务器组的配置	show radius group

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 radius 事件的调试开关。	debug radius event
打开 radius 报文打印的调试开关。	debug radius detail
打开 radius 动态授权扩展功能的调试开关。	debug radius extension event
打开 radius 动态授权扩展报文打印的调试开关。	debug radius extension detail

3 TACACS+

3.1 概述

TACACS+是在 TACACS (Terminal Access Controller Access Control System , 终端访问控制器访问控制系统) 基础上进行了功能增强的安全协议。用于实现多种用户的 AAA 功能，包括认证、授权、记账。

协议规范

- RFC 1492 Terminal Access Controller Access Control System

3.2 典型应用

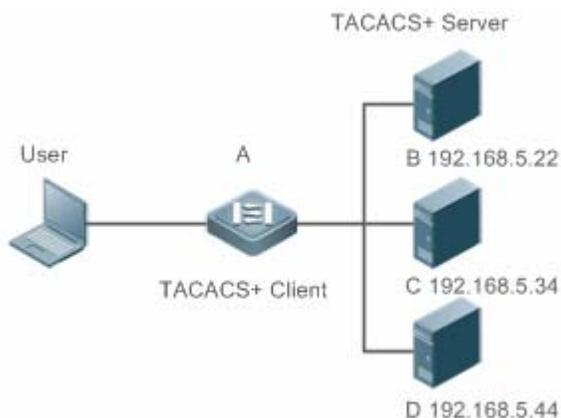
典型应用	场景描述
终端用户的登陆管理控制	对终端用户进行密码校验并进行授权。

3.2.1 终端用户的登陆管理控制

应用场景

TACACS+的典型应用为终端用户的登陆管理控制，网络设备作为 TACACS+的客户端，将用户名和密码发给 TACACS+服务器进行验证，验证通过并得到授权之后可以登录到网络设备上进行操作。下图所示：

图 3-1



- 【注释】 A 为发起 TACACS+请求的客户端。
B、C、D 为处理 TACACS+请求的服务器。

功能部署

- 在服务器 B, C, D 上启动 TACACS+ Server，并且配置接入设备（设备 A）的信息，以便能够为设备提供基于 TACACS+ 协议的 AAA 功能。
- 。在设备 A 上开启 AAA 功能，为用户登录过程启用认证过程。
- 在设备 A 上启用 TACACS+ Client 功能，并且添加 TACACS+ Server（服务器 B, C, D）的 IP 地址和对应的共享密钥，以便设备 A 能和 TACACS+ Server 进行 TACACS+ 协议通信来实现 AAA 功能。

3.3 功能详解

基本概念

▾ TACACS+的报文格式

图 3-2

4	8	16	24	32 bit
Major	Minor	Packet type	Sequence no.	Flags
Session ID				
Length				

- Major Version — 主要 TACACS+ 版本号。
- Minor Version — 次要 TACACS+ 版本号。
- Packet Type — 可能值包括：
 - TAC_PLUS_AUTHEN：= 0x01（认证）；
 - TAC_PLUS_AUTHOR：= 0x02（授权）；
 - TAC_PLUS_ACCT：= 0x03（记账）。
- Sequence Number — 当前会话中的数据包序列号。会话中的第一个 TACACS+ 数据包序列号必须为 1，其后的每个数据包序列号逐次加 1。因此客户机只发送奇序列号数据包，而 TACACS+ Daemon 只发送偶序列号数据包。
- Flags — 该字段包括各种位图格式的标志（flag）。Flag 值表明数据包是否进行加密。
- Session ID — 该 TACACS+ 会话的 ID。
- Length — TACACS+ 数据包主体长度（不包括头部），报文全部以加密形式在网络上传输。

功能特性

功能特性	作用
TACACS+认证、授权、记账	对终端用户进行认证、授权、记账

3.3.1 TACACS+认证、授权、记账

工作原理

下边以 Login 登录的基本认证、授权和记账说明 TACACS+运行中的数据报文的交互：

图 3-3



在整个过程中的基本消息交互流程可以分为三个部分：

1. 认证过程包含：
 - 1) 用户请求登录网络设备。

- 2) TACACS+客户端收到请求之后，向 TACACS+服务器发送认证开始报文。
 - 3) TACACS+服务器发送认证回应报文，请求用户名；
 - 4) TACACS+客户端向用户询问用户名。
 - 5) 用户输入登陆的用户名信息。
 - 6) TACACS+客户端收到用户名后，向 TACACS+服务器发送认证持续报文，其中包括了用户名。
 - 7) TACACS+服务器发送认证回应报文，请求登录密码；
 - 8) TACACS+客户端收到向用户询问登录密码。
 - 9) 用户输入登陆的密码信息。
 - 10) TACACS+客户端收到登录密码后，向 TACACS+服务器发送认证持续报文，其中包括了登录密码。
 - 11) TACACS+服务器发送认证回应报文，指示用户通过认证。
2. 认证通过后对用户进行授权：
- 1) TACACS+客户端向 TACACS+服务器发送授权请求报文。
 - 2) TACACS+服务器发送授权回应报文，指示用户通过授权。
 - 3) TACACS+客户端收到授权回应成功报文，向用户输出网络设备的配置界面。
3. 授权通过后，需要对登陆的用户进行记账，审计。
- 1) TACACS+客户端向 TACACS+服务器发送记账开始报文。
 - 2) TACACS+服务器发送记账回应报文，指示记账开始报文已经收到。
 - 3) 用户退出。
 - 4) TACACS+客户端向 TACACS+服务器发送记账结束报文。
 - 5) TACACS+服务器发送记账回应报文，指示记账结束报文已经收到。

3.4 配置详解

配置项	配置建议 & 相关命令	
配置TACACS+基本功能	 必须配置。用于使能 TACACS+安全服务。	
	tacacs-server host	配置 tacacs+服务器
	tacacs-server key	指定服务器与网络设备共享的密钥
	tacacs-server timeout	配置设备与 TACACS+服务器通信时，等待服务器的全局超时时间
配置TACACS+的认证、授权、记账分离处理功能	 可选配置。用于分离处理认证、授权、记账请求。	
	aaa group server tacacs+	配置 TACACS+ 服务器组，将不同的 TACACS+服务器划分到不同的组
	server	添加 TACACS+服务器组的服务器

3.4.1 配置TACACS+基本功能

配置效果

- 配置完成，TACACS+的基本功能准备就绪。配置 AAA 的方法列表时，指定使用 TACACS+方法，即可实现 TACACS+ 的认证、授权、记账。
- 进行认证、授权、记账操作时，TACACS+依据配置顺序向所配置的 TACACS+服务器发起认证、授权、记账请求。如果服务器响应超时，则依次遍历 TACACS+服务器列表。

注意事项

- TACACS+安全服务是 AAA 服务的一种，需要使用命令 **aaa new-model** 来使能安全服务。
- 配置了 TACACS+基本功能后，只是提供了一种安全服务，需要在配置 AAA 方法列表时指定使用 TACACS+服务，TACACS+的功能才会生效。

配置方法

▾ 启用 AAA

- 必须配置，启用 AAA 之后，才能配置 AAA 方法列表。TACACS+提供服务是依赖于 AAA 方法列表的。

【命令格式】 **aaa new-model**

【参数说明】 -

【缺省配置】 AAA 功能没有打开。

【命令模式】 全局模式

【使用指导】 启用 AAA 之后，才能配置 AAA 方法列表。TACACS+提供服务是依赖于 AAA 方法列表的。

▾ 配置 TACACS+服务器的 IP 地址

- 必须配置，否则设备无法和 TACACS+服务器通信来实现 AAA 功能。

【命令格式】 **tacacs-server host oob [via mgmt_name] { ipv4-address } [port integer] [timeout integer] [key [0 | 7] text-string]**

【参数说明】 *ipv4-address* : TACACS+服务器的 ipv4 地址。

oob : TACACS+通信使用 mgmt 口作为源接口，默认为使用非 mgmt 口进行通信。

via mgmt_name: oob 多个 mgmt 口时，指定具体的 mgmt 口

port integer : TACACS+通信使用的 TCP 端口，默认为 TCP 端口 49。

timeout integer : 与该 TACACS+服务器通信的超时时间，默认使用全局配置的超时时间。

key [0 | 7] text-string : 配置用于该服务器的共享密钥，未配置则使用全局配置。配置的密钥可以指定加密类型，0 为无加密，7 简单加密，默认为 0。。

【缺省配置】 没有配置任何 TACACS+服务器

【命令模式】 全局模式

- 【使用指导】
- 可以在配置 IP 地址的同时指定该服务器的共享密钥，如果没有指定，则使用 **tacacs-server key** 命令配置的全局密钥作为该服务器的共享密钥。共享密钥必须与服务器上配置的完全相同。
 - 可以在配置 IP 地址的同时指定该服务器的通信端口。
 - 可以在配置 IP 地址的同时指定与该服务器通信的超时时间。

配置 TACACS+服务器的共享密钥

- 可选配置。
- 如果没有通过该命令配置全局的通信协议，则在使用 **tacacs-server host** 命令添加服务器信息时，需要通过 **key** 关键字指定基于服务器的共享密钥，否则设备将无法和 TACACS+服务器进行通信。
- 如果在使用 **tacacs-server host** 命令添加服务器时没有通过 **key** 关键字指定共享密钥，则使用该全局密钥。

【命令格式】 **tacacs-server [key [0 | 7] text-string]**

【参数说明】 *text-string*：共享口令的文本
0 | 7：口令的加密类型，**0** 无加密，**7** 简单加密。

【缺省配置】 没有配置任何 TACACS+服务器的共享密钥。

【命令模式】 全局模式

【使用指导】 该命令配置全局使用的共享密钥服务器当我们需要为每个服务器指定不同的密钥时，我们使用 **tacacs-server host** 命令中的 **key** 选项实现。

配置 TACACS+服务器的超时时间

- 可选配置。
- 当设备和服务器之间的链路不稳定时，可以将超时时间改大。

【命令格式】 **tacacs-server timeout seconds**

【参数说明】 *seconds*：超时时间（单位为秒）。可设置的值范围为 1-1000 秒。

【缺省配置】 5 秒

【命令模式】 全局模式

【使用指导】 配置全局的服务器响应超时时间。当我们需要为每个服务器指定不同的超时时间时，使用 **tacacs-server host** 命令中的 **timeout** 选项实现。

检验方法

配置 AAA 方法列表使用 TACACS+方法，用户进行认证、授权、记账

- 设备与 TACACS+服务器进行交互，通过抓包可以查看设备与服务器间的 TACACS+协议的交互过程。
- 通过服务器的日志确认认证、授权、记账是否正常。

配置举例

Login 认证使用 TACACS+

【网络环境】

图 3-4



- 【注释】 A 为发起 TACACS+请求的客户端。
B 为处理 TACACS+请求的服务器。

- 【配置方法】
- 配置启用 AAA。
 - 配置 tacacs+ server 信息。
 - 配置使用 tacacs+的认证方法。
 - 在接口上应用配置认证方法。

```
A
Ruijie# configure terminal
Ruijie(config)# aaa new-model
Ruijie(config)# tacacs-server host 192.168.5.22
Ruijie(config)# tacacs-server key aaa
Ruijie(config)# aaa authentication login test group tacacs+
Ruijie(config)# line vty 0 4
Ruijie(config-line)# login authentication test
```

- 【检验方法】 在 PC 上 telnet 到设备上, 要求输入用户名和密码。输入正确的用户名和密码, 能够登录到设备上。在 TACACS+ 服务器上可以查看到此用户的认证日志。

常见错误

- 没有使能 AAA 安全服务。
- 设备配置的 key 与服务器配置的 key 不一致。
- 没有配置方法列表。

3.4.2 配置TACACS+的认证、授权、记账分离处理功能

配置效果

- 安全服务中的认证、授权、记账分别由不同的 TACACS+服务器处理。可以提高安全性, 并实现一定负载均衡。

注意事项

- TACACS+安全服务是 AAA 服务的一种，需要使用命令 **aaa new-model** 来使能安全服务。
- 配置了 TACACS+基本功能后，只是提供了一种安全服务，需要在配置 AAA 方法列表时指定使用 TACACS+服务，TACACS+的功能才会生效。

配置方法

配置 TACACS+服务器组

- 必须配置。默认情况下，只有 tacacs+这一个服务器组，无法实现认证、授权、记账分离处理。
- 配置三个 TACACS+服务器组分别用于认证、授权、记账处理。

【命令格式】 **aaa group server tacacs+ group-name**

【参数说明】 *group-name*：组的名称，组名称不可为“radius”和“tacacs+”（不包括引号），这两个名字为内置组名字

【缺省配置】 没有配置 TACACS+服务器组

【命令模式】 全局模式

【使用指导】 通过对 TACACS+服务器进行分组，认证、授权、计帐可以使用不同的服务器组来完成。

配置 TACACS+服务器组引用服务器

- 必须配置。如果没有配置，则服务器组内没有服务器，设备无法与 TACACS+服务器通信。
- 在服务器组配置模式下，引用已经使用 **tacacs-server host** 命令配置好的服务器。

【命令格式】 **server {ipv4-address}**

【参数说明】 *ipv4-address*：TACACS+服务器的 IPv4 地址。

【缺省配置】 无服务器配置。

【配置模式】 TACACS+服务器组配置模式

【使用指导】 配置此命令前，必须通过命令 **aaa group server tacacs+** 进入 TACACS+组配置模式。

TACACS+服务器组中配置的服务器地址，必须在全局配置模式下，通过 **tacacs-server host** 命令配置此服务器。

如果一个服务器组内引用了多个服务器时，当一个服务器没有响应时，设备会继续向组内的下一个服务器发送 TACACS+请求。

检验方法

配置 AAA 方法列表使用 TACACS+方法，用户进行认证、授权、记账。

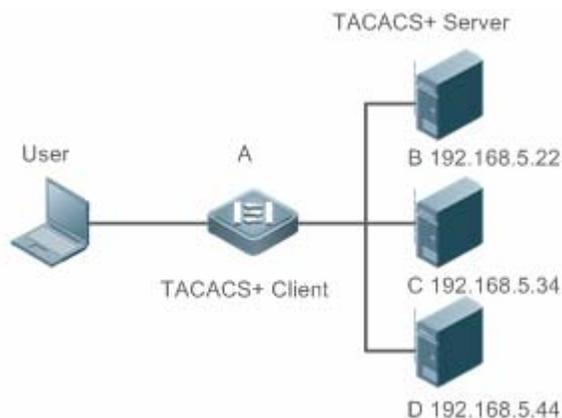
- 设备与 TACACS+服务器进行交互，通过抓包可以看到认证、授权、记账报文分别和不同的服务器交互，以及报文的源地地址。

配置举例

配置认证、授权、记账使用不同 TACACS+服务器组。

【网络环境】

图 3-5



- 【注释】
- A 为发起 TACACS+请求的客户端。
 - B 为处理 TACACS+认证请求的服务器。
 - C 为处理 TACACS+授权请求的服务器。
 - D 为处理 TACACS+记账请求的服务器。

【配置方法】

- 配置启用 AAA。
- 配置 TACACS+ server 信息。
- 配置 TACACS+服务器组。
- 向服务器组中添加服务器
- 配置使用 TACACS+的认证方法。
- 配置使用 TACACS+的授权方法。
- 配置使用 TACACS+的记账方法。
- 在接口上应用认证方法。
- 在接口上应用授权方法。
- 在接口上应用记账方法。

```
Ruijie# configure terminal
Ruijie(Ruijie(config)# aaa new-model
Ruijie(config)# tacacs-server host 192.168.5.22
Ruijie(config)# tacacs-server host 192.168.5.34
Ruijie(config)# tacacs-server host 192.168.5.44
Ruijie(config)# tacacs-server key aaa
Ruijie(config)# aaa group server tacacs+ tacgrp1
Ruijie(config-gs-tacacs)# server 192.168.5.22
Ruijie(config-gs-tacacs)# exit
Ruijie(config)# aaa group server tacacs+ tacgrp2
Ruijie(config-gs-tacacs)# server 192.168.5.34
Ruijie(config-gs-tacacs)# exit
Ruijie(config)# aaa group server tacacs+ tacgrp3
Ruijie(config-gs-tacacs)# server 192.168.5.44
Ruijie(config-gs-tacacs)# exit
Ruijie(config)# aaa authentication login test1 group tacacs+
Ruijie(config)# aaa authentication enable default group tacgrp1
Ruijie(config)# aaa authorization exec test2 group tacgrp2
Ruijie(config)# aaa accounting commands 15 test3 start-stop group tacgrp3
Ruijie(config)# line vty 0 4
Ruijie(config-line)# login authentication test1
Ruijie(config-line)# authorization exec test2
Ruijie(config-line)# accounting commands 15 test3
```

【检验方法】 在 PC 上 telnet 到设备上 ,要求输入用户名和密码。输入正确的用户名和密码 ,能够登录到设备上。输入 **enable** 命令 ,输入正确的 **enable** 密码 ,发起 **enable** 认证 ,认证通过后 ,进入特权模式。对设备进行操作后 ,用户退出设备。

在服务器 192.168.5.22 上可以查看到此用户的认证日志。

在服务器 192.168.5.22 上可以查看到此用户的 **enable** 认证日志。

在服务器 192.168.5.34 上可以查看到此用户的 **exec** 授权日志。

在服务器 192.168.5.44 上可以查看到此用户的命令记账日志。

常见配置错误

- 没有使能 AAA 安全服务。
- 设备配置的 key 与服务器配置的 key 不一致。
- 服务器组引用为定义的服务器。
- 没有配置方法列表。

3.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
显示和各 TACACS+服务器的交互运行情况。	show tacacs

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 TACACS+的调试开关。	debug tacacs+

4 802.1x

4.1 概述

IEEE802.1x (Port-Based Network Access Control) 是一个基于端口的网络访问控制标准，为 LAN 提供安全接入服务。

IEEE 802 LAN 中，用户只要能接到网络设备上，不需要经过认证和授权即可直接访问网络资源。这种不受控行为会给网络带来安全隐患。IEEE 802.1x 协议就是为了解决 802 LAN 安全问题提出来的。

802.1x 支持 Authentication , Authorization , Accounting 三种安全应用，简称 AAA。

- Authentication : 认证，用于判定用户是否可以获得访问权，限制非法用户；
- Authorization : 授权，授权用户可以使用哪些服务，控制合法用户的权限；
- Accounting : 记账，记录用户使用网络资源的情况，为收费提供依据。
- 802.1x 可以部署在对接入用户进行控制的网络中，以实现对接入用户身份验证、授权服务等

协议规范

- IEEE802.1x : Port-Based Network Access Control

4.2 典型应用

典型应用	场景描述
有线 802.1x认证	校园网安全准入，在接入交换机上部署 802.1x 认证

4.2.1 有线 802.1x认证

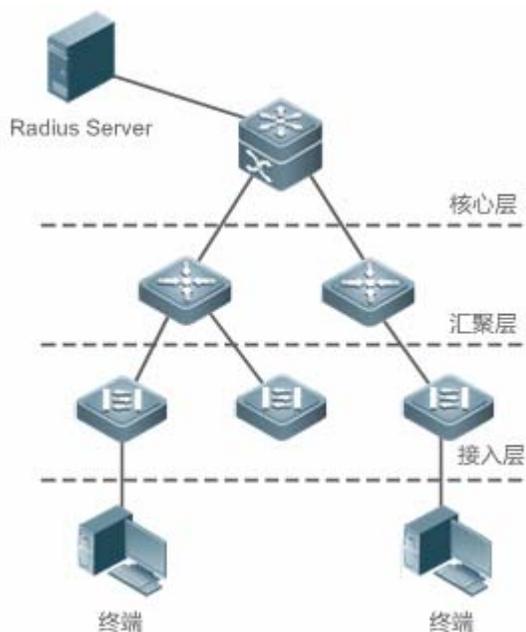
应用场景

校园网部署接入、汇聚、核心三层架构，接入交换机连接宿舍，部署 802.1x 做安全准入，宿舍用户访问校园网时需要先通过 802.1x 认证。

以下图为例：

- 用户终端上要装有 802.1x 的客户端软件（操作系统自带，或者锐捷 supplicant，或者其他符合 IEEE802.1x 标准的客户端软件）
- 接入交换机支持 IEEE 802.1x；
- 有一台（或多台）支持标准 RADIUS 的服务器作为认证服务器

图 4-1



【注释】 终端安装 supplicant 软件（也可以操作系统自带），执行 802.1x 认证。接入交换机或者汇聚交换机或者核心交换机部署 802.1x 身份认证。Radius 服务器运行 radius server 软件，执行身份校验。

功能部属

- 接入交换机连接用户的端口配置 802.1x 认证功能，实现端口受控，只有认证通过的用户才能访问网络
- 配置 AAA 方法列表，使 802.1x 可以匹配正确的方法，使用正确的认证服务器
- 配置 radius 参数,参考 RDS-SCG 的说明，确保交换机可以和服务器正常通信
- 如果采用锐捷 radius 服务器，还需要配置 snmp 参数，可以支持 radius 服务器对设备进行查询设置等操作
- 接入交换机上联口，必须是非受控口，以便设备能正常地与服务器进行通讯
- Radius 服务器创建帐号，注册接入交换机的 ip 地址，并配置 radius 相关参数，radius 服务器才会对交换机的请求作出响应

4.3 功能详解

基本概念

用户

在有线环境下，802.1x 协议是基于 LAN 的一个协议，对用户的识别不是基于账号，而是基于物理信息，在一个 LAN 里面，一个 MAC 地址+VID 的组合表示一个用户。除了这个两个信息是唯一外，其他信息都可以变，比如账号、ip 地址等。

radius

radius (remote authentication dial-in user service) 是一种远程认证协议，在 RFC2865 中定义，有着广泛的支持。利用该协议，可以实现服务器远程部署并实施认证。实际部署 802.1x 时，server 通常都是选择远程部署，设备和 server 间的 802.1x 认证信息通过 radius 传输。

📌 超时

认证过程中设备需要和终端、服务器通信，如果终端或服务器在协议指定的时间内没做出应答，则认为超时，超时会导致认证失败。实际部署时，需要注意 802.1x 协议有自己的超时时间，radius 协议也有自己的超时时间，配合使用时必须保证 802.1x 的超时时间大于 radius 的超时时间。

📌 MAB

MAB 是指使用 MAC 地址作为用户名和密码进行认证，对于一些哑终端，比如网络打印机来说，无法安装 supplicant 软件，但是有需要做安全控制，此时可以通过 MAB 实现安全准入。

📌 EAP

802.1x 协议使用 eap 协议承载认证信息，eap 协议在 rfc3748 中定义。eap 协议提供了一个通用的认证框架，在这个框架内可以嵌套多种认证方法，比如 MD5 认证、CHAP 认证、PAP 认证、TLS 认证等。锐捷 802.1x 认证支持 MD5、CHAP、PAP、PEAP-MSCHAP、TLS 等认证方法。

📌 授权

- 授权是指用户认证通过后给用户绑定一定的服务，比如绑定 ip 地址、绑定 vlan、绑定 ACL、绑定 QoS 等。

📌 计费

计费功能可以实现用户网络审计，比如使用网络的时间、产生的流量，这有利于网络运维和管理。

- 📘 有些 radius 服务器，比如锐捷 SAM 和锐捷 SMP 软件，需要依靠计费报文来判断用户的上下线状态，因此选择这些服务器软件作为 radius 服务器时，必须要配置计费功能。

功能特性

功能特性	作用
认证	提供用户的安全准入，通过认证的用户才可以访问网络
授权	通过认证的用户具备的网络访问权限，比如 ip 绑定、acl 绑定等。
计费	提供上网记录审计，比如上网时长、流量等。

4.3.1 认证

认证的目的是为了确认用户身份是否合法，避免非法用户接入网络。用户为了获得访问网络的权限，需要先通过身份认证，服务器确认账号正确后，用户才可以访问网络。在用户通过认证之前，只有 EAPOL 报文(Extensible Authentication Protocol over LAN，802.1x 协议报文)可以在网络上通行(用于认证)。

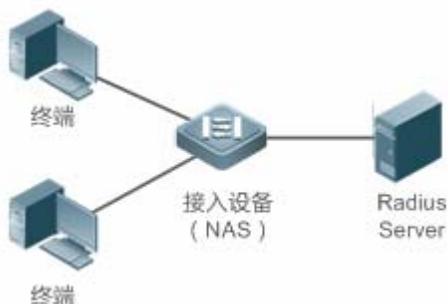
工作原理

802.1x 认证的原理比较简单，就是用户提交账号信息，设备将账号信息发给远程的 radius 服务器进行身份验证，认证通过后允许用户访问网络。

认证过程的角色

IEEE802.1x 标准认证体系由恳请者(supplicant)、认证者(authenticator)、认证服务器(server)三个角色构成，在实际应用中，三者分别对应为：终端 (Client)、接入设备(network access server , NAS)、认证服务器(最常见的是 Radius 服务器)。

图 4-2



● 恳请者

恳请者是最终用户所扮演的角色，一般是个人 PC。它请求对网络服务的访问，并对认证者的请求报文进行应答。恳请者必须运行符合 IEEE 802.1x 客户端标准的软件，目前最典型的的就是操作系统自带的 IEEE802.1x 客户端支持，另外，锐捷也已推出符合该客户端标准的 RG Supplicant 软件。

● 认证者

认证者一般为交换机或者无线访问热点等网络接入设备。该设备的职责是根据客户端当前的认证状态控制其与网络的连接状态。在客户端与服务器之间，该设备扮演着中介者的角色：从客户端要求用户名，核实从服务器端的认证信息，并且转发给客户端。因此，设备除了扮演 IEEE802.1x 的认证者的角色，还扮演 RADIUS Client 角色，因此我们把设备称作 network access server(NAS)，它要负责把从客户端收到的回应封装到 RADIUS 格式的报文并转发给 RADIUS Server，同时它要把从 RADIUS Server 收到的信息解释出来并转发给客户端。

扮演认证者角色的设备有两种类型的端口：受控端口 (controlled Port) 和非受控端口 (uncontrolled Port)。连接在受控端口的用户只有通过认证才能访问网络资源；而连接在非受控端口的用户无须经过认证便可以直接访问网络资源。我们把用户连接在受控端口上，便可以实现对用户的控制；非受控端口主要是用来连接认证服务器，以便保证服务器与设备的正常通讯。

● 认证服务器

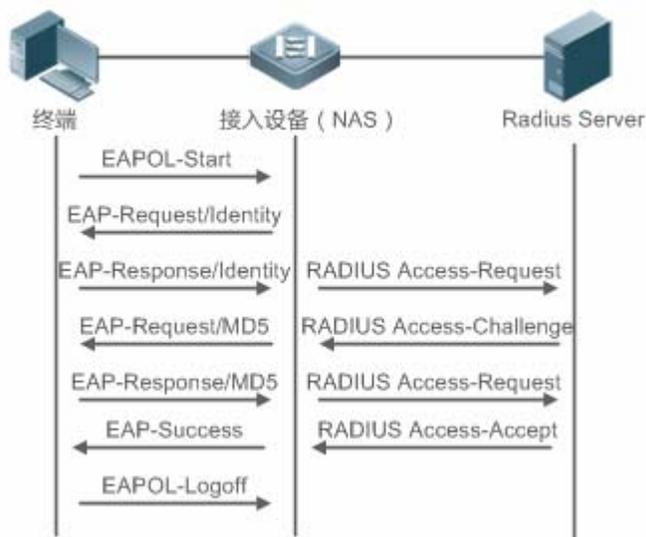
认证服务器通常为 RADIUS 服务器，认证过程中与认证者配合，为用户提供认证服务。认证服务器保存了用户名及密码，以及相应的授权信息，一台服务器可以对多台认证者提供认证服务，这样就可以实现对用户的集中管理。认证服务器还负责管理从认证者发来的记帐数据。锐捷 802.1x 兼容标准的 Radius Server，如微软 IAS/NPS、Free Radius Server、思科 ACS 等。

认证过程及报文交互

恳请者和认证者之间通过 EAPOL 协议交换信息，而认证者和认证服务器通过 RADIUS 协议交换信息，通过这种转换完成认证过程。EAPOL 协议封装于 MAC 层之上，类型为 0x888E。同时，标准为该协议申请了一个组播 MAC 地址 01-80-C2-00-00-03，用于初始认证过程中的报文传递。我司认证客户端还有可能将 01-D0-F8-00-00-03 用于认证开始的报文。

图 1-4 是有线设备一次典型的认证过程中，三个角色设备的报文交互过程：

图 4-3



该过程是一个典型的由用户发起的认证过程。在一些特殊的情形下，设备也可能主动发出认证请求，过程与该图一致，只是少了用户主动发出请求这一步。

认证用户状态

802.1x 中根据端口的认证状态来决定该端口上的用户是否允许访问网络，锐捷产品扩展了 802.1X 协议，默认是基于用户控制(有线设备上一般以 mac+vid 标识一个用户)，所以，默认是根据一个端口下的用户的认证状态来决定该用户是否允许访问网络资源。锐捷 802.1x 也支持端口模式，详细可参考后续配置章节的描述。

一个非受控端口下的所有用户均可使用网络资源，而一个受控端口下的用户只有处于已认证状态 (Authorized) 才能访问网络资源。一个用户刚发起认证时，状态处于未认证状态 (unauthorized)，这时它不能访问网络，在认证通过后，该用户的状态会变为已认证状态 (authorized)，此时该用户便可以使用网络资源。

如果工作站不支持 802.1x，而该工作站连接在受控端口下，当设备请求该用户的用户名时，由于工作站不支持导致没对该请求做出响应。这就意味着该用户仍然处于未认证状态 (unauthorized)，不能访问网络资源。

相反地，如果工作站支持 802.1x，而所连的设备不支持 802.1x。用户发出的 EAPOL-START 帧无人响应，用户在发送一定数目的 EAPOL-START 帧仍未收到回应的情形下，将认为自己所连的端口是非受控端口，而直接使用网络资源。

在支持 802.1x 的设备下，所有端口的默认设置是非受控端口，我们可以把一个端口设置成受控端口，从而要求这个端口下的所有用户都要进行认证。

当用户通过了认证（设备收到了从 RADIUS Server 服务器发来的成功报文），该用户便转变成已认证状态(authorized)，该用户可以自由使用网络资源。如果用户认证失败以至仍然处于未认证状态，可以重新发起认证。如果设备与 RADIUS server 之间的通讯有故障，那么该用户仍然处于未认证状态 (unauthorized)，网络对该用户来说仍然是不可使用的。

当用户发出 EAPOL-LOGOFF 报文后，该用户的状态由已认证(authorized)转向未认证状态(unauthorized)。

当设备的某个端口变为 LINK-DOWN 状态，该端口上的所有用户均变为未认证(unauthorized)状态。

当设备重新启动，该设备上的所有用户均变为未认证状态(unauthorized)。

如果要强制一个终端免认证，可以通过添加静态 MAC 地址或者配置 IP+MAC 绑定来实现。

▾ 搭建认证服务器

802.1x 认证使用 radius server 作为认证服务器，因此部署 802.1x 安全准入时，需要同时部署 radius server。常见的 radius server 有微软的 IAS/NPS、思科的 ACS 以及锐捷的 SAM/SMP 等。具体的部署步骤可参考对应软件的说明手册。

▾ 配置设备的认证参数

为了使用 802.1x 认证，需要在接入端口上开启 802.1x 认证功能，然后配置 aaa 的方法列表以及 radius 服务器参数。需要保证设备和 radius 服务器是可达的，需要保证 802.1x 的服务器超时时间是大于 radius 的服务器超时时间。

▾ supplicant

用户需要在终端上打开 supplicant 软件，输入账号并发起认证，如果使用的是操作系统自带的客户端，则操作系统在网络可用时会弹出对话框让用户输入账号。不同客户端软件的实现可能存在差异，界面的操作方式也可能存在差异，推荐使用锐捷 supplicant 软件作为认证客户端，如果使用其他软件，请参考相应的软件说明书。

▾ 下线

用户如果不想访问网络了，可以选择下线。下线有多种方式，包括：关机、端口网络连接、部分 supplicant 提供的下线功能。

4.3.2 授权

授权是指在用户通过认证之后，限定用户对网络使用的范围，比如 mac 绑定 ip、限制可上网时间或时段、可访问的 vlan、可享受的带宽等。

工作原理

授权是指将权限和用户绑定，根据前面的描述，用户以 mac+vid 标识，授权就是在 mac+vid 的基础上再增加一些绑定信息，比如绑定 ip、绑定 vlan 等。

▾ ip 授权

802.1x 认证标准是不支持 ip 信息识别的，锐捷 802.1x 认证扩展了 802.1x 应用，支持 mac+ip 绑定，称为 ip 授权。ip 授权有四种模式，包括：

Supplicant 授权：ip 地址由 supplicant 提供，这个模式需要锐捷 supplicant 配合才能支持；

Radius 授权：ip 地址由 radius 服务器在认证通过后下发给设备；

Dhcp 授权：用户终端认证通过后发起 dhcp 请求，获取到 ip 地址后将该 ip 和终端 mac 绑定，适用于动态 ip 环境；该环境下部分用户可能不适合使用 dhcp 获取 ip 地址，而是使用静态 ip 地址，可以单独为这些静态 ip 的用户配置 ip+mac 的绑定；

Mixed 授权：认证用户按照 Supplicant 授权、Radius 授权和 Dhcp 授权的顺序对用户进行 mac+ip 的绑定。即 supplicant 提供了 ip 地址，则优先使用该 ip 地址，如果没有提供则使用 radius 服务器提供的 ip 地址，如果 radius 服务器没有提供，则最后使用 dhcp 提供的 ip 地址；

ACL 授权

用户通过认证之后，服务器针对用户下发 acl 或者 ace，如果下发 acl，则需要事先在设备上配置好 acl，如果是下发 ace，则无需其他配置。ACL 授权基于 radius 的属性下发，支持标准属性、锐捷私有属性、思科私有属性，具体需要参考所使用的 radius 服务器的软件说明。

踢线

锐捷 802.1x 和锐捷 SAM/SMP 配合使用时，支持服务器对在线用户实施踢线，踢线后用户将无法访问网络。该功能在一些上网时段控制、上网费用实时检查的环境中可以使用。

4.3.3 计费

计费功能允许网络运营方对接入用户实施上网审计或者费用审计，通常包括时间和流量的审计等。

工作原理

设备配置计费功能，radius 服务器支持 rfc2869 定义的计费审计，用户上线时设备向服务器发送计费开始报文，服务器开始计费，用户下线时，设备向服务器发送计费结束报文，服务器完成一次审计，形成上网费用审计清单。关于计费，不同服务器可能会有不同实现，另外也不是所有服务器都支持计费功能，因此实际部署计费时需要参考服务器的使用说明。

计费开始

配置了计费功能情况下，用户通过认证后，设备会向服务器发送一个计费开始报文，报文携带用户的计费属性，比如用户名和计费 id 等，服务器收到报文后开始对用户计费。

计费更新

设备周期性的向服务器发送计费更新报文，计费更新报文可以使服务器的计费实时性特到提高。计费更新的间隔可以服务器下发，也可以设备上配置。

计费结束

用户下线后设备向服务器发送计费结束报文，携带用户的上网时长、上网消耗的流量等信息，服务器根据这些信息形成用户的上网记录。

4.4 配置详解

配置项	配置建议 & 相关命令	
配置 801.1x 基本功能	 必须配置，用于部署基本的安全认证和计费。	
	<code>aaa new-model</code>	使能 aaa
	<code>aaa authentication dot1x</code>	配置认证方法列表
	<code>aaa accounting networks</code>	配置计费方法列表
	<code>radius-server host</code>	配置 radius 服务器

	radius-server key	配置设备和 radius 服务器通信的密钥
	dot1x port-control auto	配置端口上的 802.1x 认证
配置 802.1x协议参数	 可选配置。用于调整 802.1x 协议参数。  要确保 802.1x 的服务器超时时间大于 radius 的服务器超时时间。  锐捷客户端在线检测功能仅适用于锐捷 supplicant	
	dot1x re-authentication	配置重认证功能
	dot1x timeout re-authperiod	配置重认证间隔
	dot1x timeout tx-period	配置 request/id 报文重传间隔
	dot1x reauth-max	配置 request/id 报文重传次数
	dot1x timeout supp-timeout	配置 request/challenge 报文重传间隔
	dot1x max-req	配置 request/challenge 报文重传次数
	dot1x timeout server-timeout	配置服务器超时时间
	dot1x timeout quiet-period	配置认证失败后的静默时间
	dot1x auth-mode	配置认证模式(eap/chap/pap)
	dot1x client-probe enable	配置锐捷客户端在线检测
	dot1x probe-timer interval	配置锐捷客户端检测周期
	dot1x probe-timer alive	配置锐捷客户端检测时长
配置授权	 可选配置。用于调整 802.1x 协议参数  ip 授权模式中的 supplicant 授权需配合锐捷 supplicant	
	aaa authorization ip-auth-mode	配置 ip 授权模式
	dot1x private-supplicant-only	配置过滤非锐捷客户端功能
	dot1x redirect	配置二代 su 升级功能，通过浏览器重定向到指定的资源网站下载 supplicant 软件
	snmp	配置 snmp 参数，锐捷 SAM/SMP 支持对 802.1x 在线用户实施体现，该体现通过 snmp 协议实现，使用该功能需要配置 snmp 参数
配置MAB	 可选配置，用于支持 mac 认证  802.1x 认证优先级高于 MAB 认证  MAB 认证不支持 ip 授权  单 MAB 和多 MAB 互斥  MAB 采用 PAP 认证模式，部署时需要注意服务器的配置	
	dot1x mac-auth-bypass	配置单 MAB 认证
	dot1x mac-auth-bypass multi-user	配置多 MAB 认证
	dot1x multi-mab quiet-period	配置多 MAB 认证失败后的静默时间
	dot1x mac-auth-bypass timeout-activity	配置 MAB 认证超时时间
	dot1x mac-auth-bypass violation	配置 MAB 违例

	dot1x mac-auth-bypass vlan	配置 MAB VLAN
	dot1x mac-auth-bypass valid-ip-auth	配置获取 MAB 用户 IP 之后发起认证
配置服务器失效旁路认证	 可选配置，用于配置服务器失效旁路认证	
	dot1x critical	配置服务器失效旁路认证
	dot1x critical recovery action reinitialize	配置服务器失效旁路认证恢复处理
配置扩展功能	 可选配置，用于配置端口主动请求  可选配置，配置可认证主机列表  可选配置，配置设备发送伪 mac  可选配置，配置同 mac 多帐号	
	dot1x auto-req	配置设备主动发起 802.1x 认证
	dot1x auto-req packet-num	配置设备主动发起认证请求报文的个数
	dot1x auto-req user-detect	配置主动认证检测是否有用户在认证
	dot1x auto-req req-interval	配置设备主动发起认证请求报文的间隔时间
	dot1x auth-address-table address	配置可认证主机列表
	dot1x pseudo source-mac	配置设备使用虚拟 mac 作为设备发出的 802.1x 报文的源 mac 地址
	dot1x multi-account enable	配置支持一个 mac 使用多账号认证
	dot1x valid-ip-acct enable	配置获取 IP 后开始计费功能
	dot1x valid-ip-acct timeout	配置用户认证通过之后，允许等待该用户获取 IP 的时间，超过该时间用户未获取 IP 地址将被踢下线

4.4.1 配置 802.1x 基本功能

配置效果

- 提供基本的认证和计费服务。
- 有线环境下,使用接口模式下的配置端口 802.1x 命令可以启动或关闭接口上的 802.1x 认证功能.
- 通过 radius 服务器命令配置服务器的 ip 和协议通信端口信息，配置设备和服务器间的 radius 加密密钥，确保通信安全.
- 使用全局命令 **aaa accounting update** 命令开启计费更新，计费更新间隔可以在设备上通过 **aaa accounting update interval** 命令配置参数，也可以在服务器上配置，这取决于服务器是否支持该功能。如果服务器有下发，则优先使用服务器下发的参数，如果服务器没下发，则使用本机配置参数.

注意事项

- 注意 radius 参数的配置准确性，确保基本的 radius 协议通信正常。

- 802.1x 使用的认证方法列表和计费发列表必须在 aaa 里面已经配置好了，否则会导致认证和计费出错。
- 对于交换机产品，由于芯片限制，只要有一个端口打开了 802.1x 功能，所有的端口都会将 802.1x 协议报文送到 cpu。
- 如果端口 802.1x 开启并且认证的用户数大于端口安全的最大用户数，此时无法开启端口安全。
- 端口安全和 802.1x 同时开启时，如果安全地址老化，则 802.1x 对应的用户必须重新认证才可以继续通信。
- 静态地址或者符合 ip+mac 绑定的用户无需认证即可访问网络。
- 802.1x 默认使用 default 方法类表，如果 aaa 配置的不是 default 方法列表，需要通过 dot1x authentication 和 dot1x accounting 命令重新制定 802.1x 使用的方法列表。
- 配合锐捷 SAM/SMP 软件使用时，必须配置计费功能，否则用户下线时服务器无法感知导致表项残留。

配置方法

▾ 使能 aaa

- 必须配置，使能 aaa 之后 802.1x 认证计费功能才会生效。
- 在使用 802.1x 对接入用户进行受控的设备开启

【命令格式】 **aaa new-model**

【参数说明】 -

【缺省配置】 关闭

【命令模式】 全局模式

【使用指导】 默认关闭，部署 802.1x 认证必须要配置该命令

▾ 配置 aaa 认证方法

- 必须配置。
- 需要和 801.x 使用的认证方法一致
- 在使用 802.1x 对接入用户进行受控的设备开启

【命令格式】 **aaa authentication dot1x list-name group radius**

【参数说明】 *list-name* : aaa 的 dot1x 认证方法列表

【缺省配置】 关闭

【命令模式】 全局模式

【使用指导】 默认关闭

需要 802.1x 的认证方法一致

▾ 配置 radius 服务器参数

- 必须配置，可以实现设备和 radius 服务器的正常通信。
- 在使用 802.1x 对接入用户进行受控的设备开启

【命令格式】 **radius-server host ip-address [auth-port port1] [acct-port port2]**

【参数说明】 *ip-address* : 指定服务器 ip 地址

port1 : 认证协议端口

- `port2` : 计费协议端口
- 【缺省配置】 默认无 radius 服务器参数
- 【命令模式】 全局模式
- 【使用指导】 -

配置 radius 服务器通信密钥

- 必须配置，可以实现设备和 radius 服务器的正常通信。
- 在使用 802.1x 对接入用户进行受控的交换机开启

- 【命令格式】 **radius-server key string**
- 【参数说明】 *string* : radius 通信密钥
- 【缺省配置】 默认无 radius 通信密钥
- 【命令模式】 全局模式
- 【使用指导】 设备的 ip 地址必须和服务器上注册的设备地址一致
设备和服务器的通信的 key 也必须配置一致
如果服务器更改了默认的 radius 通信端口，则配置时也需要指定协议端口

配置有线 802.1x

- 有线交换机上必须配置。
- 在有线交换机上开启

- 【命令格式】 **dot1x port-control auto**
- 【参数说明】 -
- 【缺省配置】 关闭
- 【命令模式】 接口模式
- 【使用指导】 默认关闭，部署 802.1x 必须要配置该命令
默认使用 default 方法列表，如果 aaa 中的 802.1x 的方法列表不是 default1，802.1x 的方法列表也需要匹配

检验方法

开启 supplicant 软件并发起认证，输入正确的账号并发起认证，然后通过 802.1x 的检查和 radius 的检查确认配置是否准确。

通过 show dot1x summary 查看 802.1x 是否有创建认证表项

- 【命令格式】 **show dot1x summary**
- 【参数说明】 -
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 查看认证用户表项信息，通过该信息，可以知道认证终端当前处于什么阶段，比如正在认证、已经认证或者静默。

【命令展示】

```
Ruijie#show dot1x summary
```

ID	Username	MAC	Interface	VLAN	Auth-State	Backend-state	Port-Status
----	----------	-----	-----------	------	------------	---------------	-------------

```

User-Type Time
-----
-----
16777302 ts-user b048.7a7f.f9f3 wlan 1 1 Authenticated Idle Authed
static 0days 0h 0m12s

```

通过 show aaa user all 查看 aaa 是否有用户表项

【命令格式】 **show aaa user all**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 显示 AAA 用户相关信息。

【命令展示】

```
Ruijie#show aaa user all
```

```

-----
Id ----- Name
2345687901 wwx
-----

```

-
- 通过设备和服务器间的 radius 报文检测服务器是否响应了认证，如果没有响应，则属于网络不通或者参数配置错误，如果服务器直接返回拒绝，则需要查看服务器的 log 文件，看是因为什么原因，比如服务器的认证方法配置错误等。

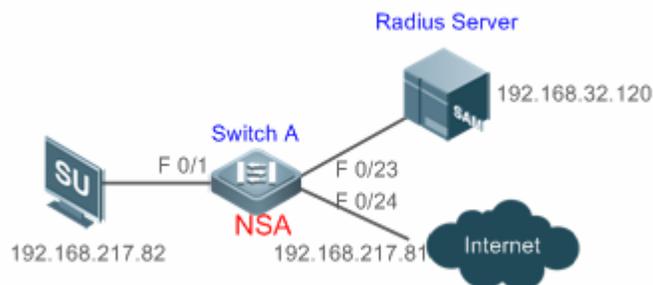
配置举例

i 以下配置举例，以锐捷 SAM 作为认证服务器。

配置有线 802.1x 认证

【网络环境】

图 4-4



- 【配置方法】
- 服务器上注册设备的 ip 信息，并配置设备和服务器的通信密钥
 - 服务器上创建账号信息
 - 设备使能 aaa
 - 设备配置 radius 参数
 - 设备接口上使能 802.1x 认证

如下为设备上的相关配置，服务器端的配置请参考具体服务器的配置指导手册：

```
ruijie# configure terminal
ruijie (config)# aaa new-model
ruijie (config)# radius-server host 192.168.32.120
ruijie (config)# radius-server key ruijie
ruijie (config)# interface FastEthernet 0/1
ruijie (config-if)# dot1x port-control auto
```

【检验方法】 测试是否可以正常认证以及认证前后的网络访问行为是否变化。

- 服务器创建账号，比如 username:tests-user,password:test。
- 终端未认证前无法 ping 通 192.168.32.120。
- 终端打开 supplicant 后输入账号并点击认证，认证成功，可 ping 通 192.168.32.120。
- 可以显示认证通过的用户信息

```
ruijie# show dot1x summary
ID      Username  MAC          Interface VLAN Auth-State  Backend-State
Port-Status User-Type Time
-----
-----
16778217 ts-user   0023.aeaa.4286 Fa0/1    2    Authenticated  Idle        Authed
static   0days 0h 0m 7s
```

常见错误

- radius 参数配置错误。
- 服务器有特殊的接入策略，比如要求 radius 报文必须携带某些属性等。
- aaa 的方法列表和 802.1x 的方法类表不一致导致无法认证

4.4.2 配置 802.1x协议参数

配置效果

- 根据网络实际情况调整协议的参数值，比如服务器性能较差的环境中，可以将服务器超时时间适当配大。

注意事项

- 802.1x 协议有自己的服务器超时参数，radius 也有自己的服务器超时参数，默认情况下，802.1x 的超时参数是 5 秒，小于 radius 的超时参数 15 秒。实际使用时，需要确保 802.1x 的服务器超时参数大于 radius 的服务器超时参数。可使用 dot1x 服务器超时配置命令将 802.1x 的超时参数配置大，radius 的超时规则请参考 radius 配置手册。

- 锐捷客户端在线检测功能仅适用于锐捷 supplicant。

配置方法

▾ 使能重认证

- 可选配置，开启后即可定期对在线用户重认证。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x re-authentication**
【参数说明】 -
【缺省配置】 默认关闭
【命令模式】 全局模式
【使用指导】 在对认证用户定时重认证时可以配置此命令

▾ 配置重认证间隔

- 可选配置，配置用户的重认证周期。
- 在设备开启 802.1x 认证之后配置，开启重认证功能后生效

【命令格式】 **dot1x timeout re-authperiod period**
【参数说明】 *period*：重认证间隔，单位秒，默认 3600 秒
【缺省配置】 默认 3600 秒
【命令模式】 全局模式
【使用指导】 根据需要来调整认证用户的重认证间隔

▾ 配置 request/id 报文重传间隔

- 可选配置，设备重传报文的周期，周期越长则报文重传时间也越长。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x timeout tx-period period**
【参数说明】 *period*：报文重传间隔，单位秒，默认 3 秒
【缺省配置】 默认 3 秒
【命令模式】 全局模式
【使用指导】 使用默认值即可，根据认证客户端响应设备请求的时间长短来调整该数值

▾ 配置 request/id 报文重传次数

- 可选配置，设备重传报文的次数，数值越大，重传次数就越大。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x reauth-max num**
【参数说明】 *num*：报文重传次数，默认 3
【缺省配置】 默认 3 次
【命令模式】 全局模式

【使用指导】 使用默认值即可，容易丢包的环境增大该值可以提高客户端收到设备报文的概率

配置 request/challenge 报文重传间隔

- 可选配置，设备重传 request/challenge 报文的间隔，数值越大，重传间隔就越大。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x timeout supp-timeout** *time*

【参数说明】 *time*：报文重传间隔，单位秒，默认 3 秒

【缺省配置】 默认 3 秒

【命令模式】 全局模式

【使用指导】 使用默认值即可，容易丢包的环境中可以增大该数值

配置 request/challenge 报文重传次数

- 可选配置，设备重传 request/challenge 报文的次数，数值越大，重传次数就越大。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x max-req** *num*

【参数说明】 *num*：报文重传次数，单位秒，默认 3

【缺省配置】 默认 3 次

【命令模式】 全局模式

【使用指导】 可选配置

使用默认值即可，容易丢包的环境中可以增大该数值

配置服务器超时时间

- 可选配置，设备服务器超时时间，数值越大，等待服务器超时的时间就越长。
- 在设备开启 802.1x 认证之后配置
- Radius 和服务器之间的通信超时必须大于 802.1x 的服务器超时时间

【命令格式】 **dot1x timeout server-timeout** *time*

【参数说明】 *time*：服务器超时时间，单位秒，默认 5 秒

【缺省配置】 默认 5 秒

【命令模式】 全局模式

【使用指导】 使用默认值即可，设备和服务器通信不稳定的环境中可以增大该数值

配置认证失败后的静默时间

- 可选配置，用户认证失败后的时间，数值越大，静默用户的时间就越长。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x timeout quiet-period** *time*

【参数说明】 *time*：认证失败后的静默时间，单位秒，默认 10 秒

【缺省配置】 默认 10 秒

【命令模式】 全局模式

【使用指导】 使用默认值即可，增大该数值可以降低认证失败的用户频繁向服务器发起认证增加服务器的负担

配置认证模式

- 可选配置，配置 802.1x 认证方式。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x auth-mode {eap | chap | pap}**

【参数说明】 eap：采用 eap 方式认证
chap：采用 chap 方式认证
pap：采用 pap 方式认证

【缺省配置】 默认 eap

【命令模式】 全局模式

【使用指导】 认证模式的选择取决于 supplicant 和认证服务器的支持情况。

配置锐捷客户端在线检测

- 可选配置，配置锐捷在线客户端检测功能，开启该功能之后，可以及时发现客户端下线，避免对用户的错误计费。
- 客户端必须是锐捷 802.1x 认证客户端
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x client-probe enable**

【参数说明】 -

【缺省配置】 默认关闭

【命令模式】 全局模式

【使用指导】 使用锐捷 supplicant 时建议使能该功能。

配置锐捷客户端在线报文发送周期

- 可选配置，配置锐捷在线客户端报文发送周期，该数值越大，客户端发送检测报文的时间间隔就越长。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x probe-timer interval time**

【参数说明】 *time*：锐捷 supplicant 向设备发送心跳报文的时间间隔，单位秒，默认 20 秒

【缺省配置】 默认 20 秒

【命令模式】 全局模式

【使用指导】 建议使用默认值即可

配置锐捷客户端在线检测时长

- 可选配置，配置锐捷在线客户端检测时长，该数值越大，判断客户端下线的的时间间隔就越长。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x probe-timer alive time**

【参数说明】 *time*：报文检测时长，单位秒，默认 250 秒

【缺省配置】 默认 250 秒

【命令模式】 全局模式

【使用指导】 可选配置

终端认证上线后，在检测时长内设备没收到终端的任何探测报文响应，则认为终端下线，建议使用默认值即可

检验方法

可以通过 show dot1x 查看参数配置是否生效。

配置举例

配置认证模式

【网络环境】 单机

【配置方法】 配置认证模式为 chap：

```
Ruijie(config)#dot1x auth-mode chap
```

【检验方法】 显示配置结果。

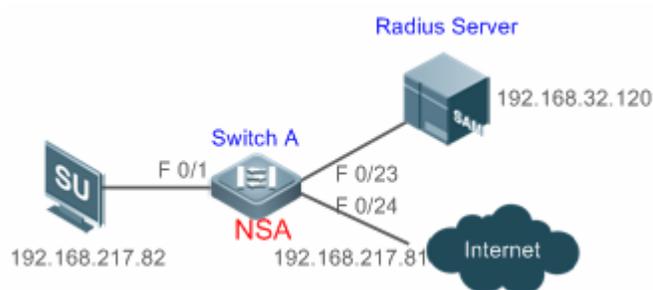
```
Ruijie(config)#show dot1x

802.1X basic information:
 802.1X Status ..... enable
 Authentication Mode ..... chap
 Authorization mode ..... disable
 Total User Number ..... 0 (exclude dynamic user)
 Authenticated User Number ..... 0 (exclude dynamic user)
 Dynamic User Number ..... 0
 Re-authentication ..... disable
 Re-authentication Period ..... 3600 seconds
 Re-authentication max ..... 3 times
 Quiet Period ..... 10 seconds
 Tx Period ..... 30 seconds
 Supplicant Timeout ..... 3 seconds
 Server Timeout ..... 5 seconds
 Maximum Request ..... 3 times
 Client Online Probe ..... disable
 Eapol Tag ..... disable
 802.1x redirect ..... disable
 Private supplicant only ..... disable
```

配置锐捷客户端在线检测

【网络环境】

图 4-5



【配置方法】 开启锐捷客户端在线检测功能：

```
Ruijie(config)#dot1x client-probe enable
```

- 由锐捷 supplicant 客户端有按时发送在线检测报文的才能持续在线
- 显示配置结果。

【检验方法】

```
Ruijie(config)#show dot1x
```

```
802.1X basic information:
```

```

802.1X Status ..... enable
Authentication Mode ..... chap
Authorization mode ..... disable
Total User Number ..... 0 (exclude dynamic user)
Authenticated User Number ..... 0 (exclude dynamic user)
Dynamic User Number ..... 0
Re-authentication ..... disable
Re-authentication Period ..... 3600 seconds
Re-authentication max ..... 3 times
Quiet Period ..... 10 seconds
Tx Period ..... 30 seconds
Supplicant Timeout ..... 3 seconds
Server Timeout ..... 5 seconds
Maximum Request ..... 3 times
Client Online Probe ..... enable
Eapol Tag ..... disable
802.1x redirect ..... disable

```

常见错误

- server-timeout 比 radius 超时参数小。
- 认证软件不是锐捷 supplicant 却配置了客户端在线检测功能。

4.4.3 配置授权

配置效果

- 配置 ip 授权可限定认证用户必须使用指定的 ip 地址访问网络，可以防止 ip 盗用等问题。ip 授权可以全局配置，也可以基于端口配置，端口配置的 ip 授权优先级高于全局配置。
- 配置过滤非锐捷客户端功能，可以限定终端必须使用锐捷 supplicant 软件认证，这样就可以享受到锐捷 supplicant 提供的相关服务，比如防代理或者短消息等功能。
- 配置 redirect 功能，可以支持二代 su 部署。所谓的二代 su 部署是指终端先通过网页下载 supplicant 软件，然后再通过 supplicant 软件认证。二代 su 部署在用户量大的环境中有利于 su 的快速部署。

注意事项

- 如果使用锐捷 SAM/SMP 软件的实时踢线功能，需要配置正确的 snmp 参数，详细可参考 snmp 配置手册。
- 环境中有多种认证客户端软件时，不能开启过滤非锐捷客户端功能。
- 更改 ip 授权模式会导致已认证用户全部下线，需要重新认证才可以上网。
- 使用混合授权模式时，用户在认证上线过程中如果更高优先级的 ip 授权出现，则使用高优先级的 ip 授权，比如原来用户使用 radius 授权，再一次认证时，如果 supplicant 提供了 ip 地址，则使用该 ip 地址进行授权。
- dhcp 授权环境下，如果有必须使用静态 ip 地址的情况，可以为这些用户配置 ip+mac 的绑定，确保在该环境下也可以正常上网，同时保证其他用户必须通过 dhcp 来授权；
- 二代 su 部署功能和 web 认证无法同时使用。
- 二代 su 部署需要配置重定向参数，具体请参考 web 认证配置手册。
- 锐捷 SAM/SMP 的踢线功能使用 snmp 协议，因此需要配置 snmp 参数，具体可参考 snmp 的配置说明

配置方法

▾ 全局配置 ip 授权模式

- supplicant 授权模式仅支持锐捷 supplicant
- radius-server 授权模式需要服务器支持通过 framed-ip 属性下发 ip 地址
- dhcp-server 授权模式需要设备同时开启 dhcp snooping
- 可选配置，配置对用户绑定 ip 和 mac 的模式。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **aaa authorization ip-auth-mode { disable | supplicant | radius-server | dhcp-server | mixed }**

【参数说明】 **disable** : 关闭 ip 授权
supplicant : supplicant 授权 ip

radius-server : radius 服务器授权 ip

dhcp-server : dhcp 服务器授权 ip

mixed : 混合授权 ip

【缺省配置】 关闭

【命令模式】 全局模式

【使用指导】 根据环境部署情况选择 ip 授权模式。

▾ 接口配置 ip 授权模式

- supplicant 授权模式仅支持锐捷 supplicant
- radius-server 授权模式需要服务器支持通过 framed-ip 属性下发 ip 地址
- dhcp-server 授权模式需要设备同时开启 dhcp snooping
- 可选配置，接口上配置对用户绑定 ip 和 mac 的模式。
- 在端口开启 802.1x 认证之后配置
- 接口配置的 ip 授权优先级高于全局配置

【命令格式】 **dot1x authorization ip-auth-mode { disable | supplicant | radius-server | dhcp-server | mixed }**

【参数说明】 **disable** : 关闭 ip 授权

supplicant : supplicant 授权 ip

radius-server : radius 服务器授权 ip

dhcp-server : dhcp 服务器授权 ip

mixed : 混合授权 ip

【缺省配置】 关闭

【命令模式】 接口模式

【使用指导】 根据环境部署情况选择 ip 授权模式。

▾ 配置静态 ip 和 mac 的绑定

- 可选配置，为 dhcp 授权模式下和混合授权模式下的 dhcp 授权情况，提供 ip+mac 绑定信息，这样的用户不必通过 dhcp 授权 ip。

【命令格式】 **dot1x address-binding mac mac-addr ip ip-address**

【参数说明】 **mac-address** : 要绑定的 mac 地址

ip-address : 要绑定的 ip 地址

【缺省配置】 关闭

【命令模式】 全局模式

【使用指导】 根据部署情况来配置

▾ 配置二代 su 部署

- 可选配置，配置第二代 su 部署，配置之后，受控口上的无 802.1x 认证客户端的用户可以通过 web 页面下载安装 802.1x 认证客户端。
- 在设备开启 802.1x 认证之后配置
- 必须配置重定向参数，具体请参考 web 认证配置手册。

- 【命令格式】 **dot1x redirect**
- 【参数说明】 -
- 【缺省配置】 关闭
- 【命令模式】 全局模式
- 【使用指导】 需要配置好重定向参数，具体参考 web 认证配置手册

配置过滤非锐捷客户端

- 可选配置，开启之后非锐捷客户端无法认证。
- 在设备开启 802.1x 认证之后配置

- 【命令格式】 **dot1x private-supplicant-only**
- 【参数说明】 -
- 【缺省配置】 关闭
- 【命令模式】 全局模式
- 【使用指导】 要限制终端必须使用锐捷 supplicant 软件认证时才可以配置该功能

检验方法

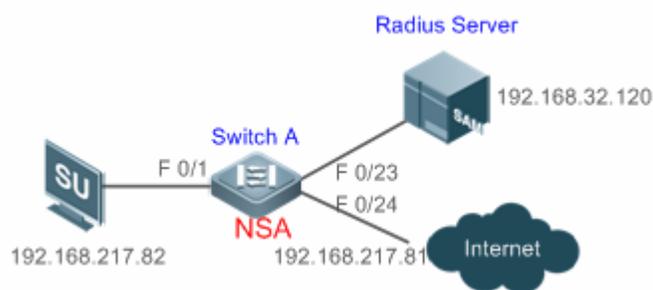
- 开启 ip 授权后，客户端先认证上线，然后更改客户端的 ip 地址，客户端无法访问网络
- 开启二代 su 部署，打开浏览器访问网址时，会自动重定向到下载页面并下载认证客户端，然后通过客户端认证通过后才可以访问网络。
- 用户认证上线后，在锐捷 SAM/SMP 上执行踢线功能，设备会将用户下线，此时用户就无法访问网络

配置举例

配置 ip 授权模式

【网络环境】

图 4-6



- 【配置方法】
- 配置 aaa
- 配置 radius
- 受控口开启 802.1x 功能
- 全局开启 supplicant 授权

```
Ruijie(config)#aaa authorization ip-auth-mode supplicant
```

- 由锐捷 supplicant 客户端发起认证，认证成功
- 该认证客户端只能使用 192.168.217.82 地址进行通信
- 显示配置结果。

【检验方法】

```
Ruijie(config)#show dot1x user name ts-user
Supplicant information:
  MAC address ..... b048.7a7f.f9f3
  Username ..... ts-user
  User ID ..... 16777303
  Type ..... static
  VLAN ..... 1
  Port ..... wlan 1
  Online duration ..... 0days 0h 0m21s
  Up average bandwidth ..... 0 kbps
  Down average bandwidth ..... 0 kbps
  Authorized VLAN ..... 1
  Authorized session time ..... 20736000 seconds
  Authorized flux ..... unlimited
  Accounting ..... No
  Proxy user ..... Permit
  Dial user ..... Permit
  IP privilege ..... 0
  Private supplicant ..... no
  Max user number on this port ..... 0
  Authorization ip address ..... 192.168.217.82
```

常见错误

- 网络中有多种认证客户端，但是开启了过滤非锐捷客户端功能，导致部分终端无法认证。
- 使用锐捷 SAM/SMP，但是设备没有配置 snmp 参数导致踢线功能失败。
- 未正确配置重定向参数导致二代 su 升级功能无法正常使用。

4.4.4 配置MAB

配置效果

- 使用接入终端的 mac 地址作为认证账号，终端无需安装认证客户端软件，适用于一些哑终端，比如网络打印机等。
- 单 MAB 适用于端口下只有一个哑终端的情况，或者只有一个终端需要认证，认证后其他终端都可以访问网络，比如端口下连了一个无线路由器，可以配置对无线路由器实时 MAB 认证，认证通过后，无线路由器下的用户均可以访问网络。
- 多 MAB 适用于端口下存在多种哑终端的情况，比如网络呼叫中心部署多台 voip 接入等。

- 多 MAB 支持和 802.1x 认证混合使用, 适用于一些混合接入环境, 比如 pc+voip 菊花链接入的方式。

注意事项

- 配置了 MAB 的端口, 每隔 tx-period 发出一个认证请求报文, 发送 reauth-max 次之后, 如果没有客户端响应, 则该端口进入 MAB 模式。进入 MAB 模式的端口可以学习 mac 地址, 并使用这些 MAC 地址为账号进行认证
- 服务器配置 MAC 账号的用户名和密码时, 必须使用不带分隔符的格式, 比如终端 mac 地址为 00-d0-f8-00-01-02, 服务器上添加帐号时需要配置为 00d0f8000102。
- 802.1x 优先级高于 MAB, 因此一个终端先 MAB 认证通过后, 如果再使用客户端软件做 802.1x 认证, MAB 的表项将被删除。
- MAB 仅支持 PAP 认证模式, 服务器上的配置需要注意。
- MAB 功能主要主动发现终端是否可以做 802.1x 认证, 通过主动认证来实现这个目的, 因此部署 MAB 的同时必须同时开启主动认证功能。

配置方法

配置单 MAB

- 可选配置
- 适用于一个端口下只有一个终端需要认证的场景。
- 在设备的 802.1x 受控口配置

【命令格式】 **dot1x mac-auth-bypass**

【参数说明】 -

【缺省配置】 关闭

【命令模式】 接口模式

【使用指导】 仅交换机平台可用适用于与端口下单个哑终端需要认证的场景, 如果要限制终端数量, 可以配置违例处理。

配置 MAB 超时时间

- 可选配置
- MAB 模式下的 MAC 地址认证上线后, 除非重认证失败、端口 down 或者因为管理策略原因下线, 比如管理员强制下线等, 否则设备将认为该 MAC 地址一直是在线的。用户可以配置许可这些认证地址的在线时间, 默认是 0, 表示允许一直在线。
- 在设备的 802.1x 受控口配置

【命令格式】 **dot1x mac-auth-bypass timeout-activity value**

【参数说明】 *value*: MAB 可在线时间, 单位为秒, 默认为 0, 表示不限制时间

【缺省配置】 默认为 0, 表示不限制时间

【命令模式】 接口模式

【使用指导】 对单 MAB 和对 MAB 均适用

配置 MAB 违例

- 可选配置
- 在设备的 802.1x 受控口配置
- 默认情况下，有一个 MAC 地址通过 MAB 认证后，该端口下的所有设备的数据都允许被转发。但是在某些安全应用下，管理员会要求一个 MAB 端口下只能有一个 MAC 地址存在，此时可以在该端口上配置 MAB 违例。配置了 MAB 违例后，一旦端口进入了 MAB 模式，如果发现该端口下有超过 1 个 MAC 地址，该端口将产生违例。

【命令格式】 **dot1x mac-auth-bypass violation**

【参数说明】 -

【缺省配置】 关闭

【命令模式】 接口模式

【使用指导】 仅交换机平台可用

要限制端口下只有一个哑终端时可以配置该命令，其它场景不能配置。

- 仅适用于单 MAB

配置多 MAB

- 可选配置
- 在设备的 802.1x 受控口配置

【命令格式】 **dot1x mac-auth-bypass multi-user**

【参数说明】 -

【缺省配置】 关闭

【命令模式】 接口模式

【使用指导】 仅交换机平台可用

端口下多哑终端需要做安全认证时配置该命令

配置多 MAB 认证失败的静默时间

- 可选配置
- 在设备的开启了多 mab 功能之后配置
- 开启多 MAB 认证功能时，为了防止接口下的非法用户对设备进行攻击，需要禁止非法用户频繁认证，以减少服务器的压力。在全局下配置多 MAB 用户的静默时间，也就是说，如果一个 mac 地址认证失败后，这个 mac 地址需要等待该时间之后才会重新发起认证，这个静默时间可以根据环境进行配置，默认配置为 0，表示一个用户认证失败后可以马上进行再认证。

【命令格式】 **dot1x multi-mab quiet-period value**

【参数说明】 *value*：认证失败后的静默时间

【缺省配置】 默认为 0s

【命令模式】 全局模式

【使用指导】 仅交换机平台可用

端口下认证的哑终端太多时以配置该命令限制认证频率。

配置 MAB VLAN

- 可选配置
- 在设备的开启了多 mab 功能之后配置
- 为了仅允许接口上部分 VLAN 内的用户进行 MAB 认证，可以将这些 VLAN 配置成 MAB VLAN，不在 MAB VLAN 之内的用户不能进行 MAB 认证。

【命令格式】 **dot1x mac-auth-bypass vlan *vlan-list***

【参数说明】 *vlan-list* : 允许进行 MAB 认证的 VLAN

【缺省配置】 关闭

【命令模式】 接口模式

【使用指导】 仅交换机平台可用

端口下仅允许相应的 VLAN 内的用户进行 MAB 认证时可以使用此命令

配置获取 MAB 用户 IP 之后发起认证

- 可选配置
- 在设备的开启了多 mab 功能之后配置
- 要求用户发起 MAB 认证时必须携带 IP 地址可以配置此命令。MAB 用户的 IP 地址一般通过 DHCP 获取，需要设备同时配置 DHCP SNOOPING；使用静态 IP 的 MAB 用户可以通过配置静态 IP+MAC 绑定来获取，具体见“配置授权”章节中的相关命令。

【命令格式】 **dot1x mac-auth-bypass valid-ip-auth**

【参数说明】 -

【缺省配置】 关闭

【命令模式】 全局模式

【使用指导】 仅交换机平台可用

要求用户发起 MAB 认证时必须携带 IP 地址可以配置此命令

检验方法

通过哑终端接入网络是否可以访问网络验证 MAB 是否生效。

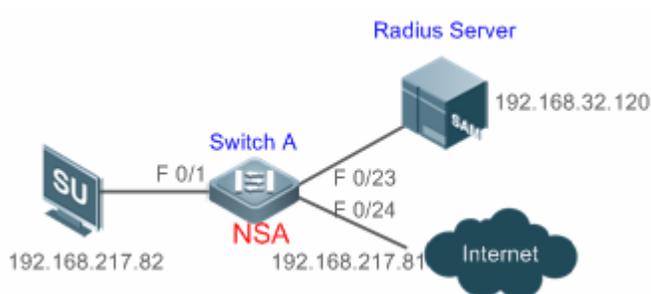
- 服务器和设备上先配置好 MAB 相关功能
- 不符合 MAC 地址账号的哑终端接入，无法访问网络
- 符合 MAC 地址账号的哑终端接入，可以访问网络

配置举例

配置有线的多 MAB 认证

【网络环境】

图 4-7



【配置方法】

- 服务器上注册设备的 ip 信息，并配置设备和服务器的通信密钥
- 服务器上创建账号信息
- 设备使能 aaa
- 设备配置 radius 参数
- 设备接口上使能 802.1x 认证和多 MAB 认证功能

如下为设备上的相关配置，服务器端的配置请参考具体服务器的配置指导手册：

```
ruijie# configure terminal
ruijie (config)# aaa new-model
ruijie (config)# radius-server host 192.168.32.120
ruijie (config)# radius-server key ruijie
ruijie (config)# interface FastEthernet 0/1
ruijie (config-if)# dot1x port-control auto
ruijie (config-if)# dot1x mac-auth-bypass multi-user
```

【检验方法】

测试是否可以正常认证以及认证前后的网络访问行为是否变化。

- 服务器创建账号，比如 username: 0023aeaa4286,password: 0023aeaa4286。
- 终端未认证前无法 ping 通 192.168.32.120。
- 终端连接上设备，认证成功，可 ping 通 192.168.32.120。
- 可以显示认证通过的用户信息

```
ruijie# show dot1x summary
ID      Username  MAC          Interface VLAN Auth-State  Backend-State
Port-Status User-Type Time
-----
-----
16778217 0023aea... 0023.aeaa.4286 Fa0/1    2    Authenticated Idle        Authed
static   0days 0h 5m 8s
```

常见错误

- 服务器上的 MAC 账号格式不准确。

4.4.5 服务器失效旁路认证

配置效果

- 配置了服务器失效旁路认证 (Inaccessible Authentication Bypass, 简称 IAB 功能) 之后, 当设备上配置的所有 RADIUS 服务器都不可达的时候, 确保新认证的用户可以访问网络。
- 使用服务器失效旁路认证恢复处理, 在 RADIUS 服务器恢复可达的时候, 可以重新验证服务器不可达期间授权用户身份的合法性。

注意事项

- RADIUS 需要配置相应的测试服务器可达的测试帐号和服务器不可达的判断标准, 具体见 RADIUS 的配置手册
- 只有 802.1x 全局配置的认证方法列表中仅存在 RADIUS 认证方法且该方法列表中的 RADIUS 服务器全部失效时, IAB 功能才生效; 如果方法列表中还存在其它认证方法(如: local、none), IAB 功能不生效。
- 全局开启 AAA 多域认证功能后, 802.1x 用户认证时, 不再使用全局配置的方法列表。由于 IAB 功能判断 802.1x 全局配置方法列表中的 RADIUS 服务器全部失效后, 直接向用户返回认证成功, 不需要输入用户名, 因此 AAA 多域认证功能在该端口上失效。
- 通过 IAB 方式认证的用户, 不再向计费服务器发起计费请求。
- 已正常认证通过的用户, 不受服务器失效的影响, 可以正常访问网络
- 全局开启 802.1x IP 授权功能时, 如果端口上已经存在认证成功的用户, 该端口上的其它用户不能通过 IAB 方式进行认证。

配置方法

▾ 服务器失效旁路认证

- 可选配置, 配置服务器失效之后对新来的认证用户进行授权。
- 在设备开启 802.1x 认证之后配置
 - 【命令格式】 **dot1x critical**
 - 【参数说明】 -
 - 【缺省配置】 关闭
 - 【命令模式】 接口模式
 - 【使用指导】 适用于需要在服务器不可达期间对新认证用户进行授权的端口。

▾ 服务器失效旁路认证恢复处理

- 可选配置, 服务器重新生效之后对之前服务器失效情况下授权的用户重新认证。
- 在设备开启 802.1x 认证之后配置
 - 【命令格式】 **dot1x critical recovery action reinitialize**

- 【参数说明】 -
- 【缺省配置】 关闭
- 【命令模式】 接口模式
- 【使用指导】 开启服务器失效旁路认证恢复处理功能之后,服务器恢复后,端口下正常认证通过的用户,可以继续访问网络,不需要重新认证;只有在服务器失效期间,通过 IAB 方式认证的用户,交换机会发起主动认证交互。

检验方法

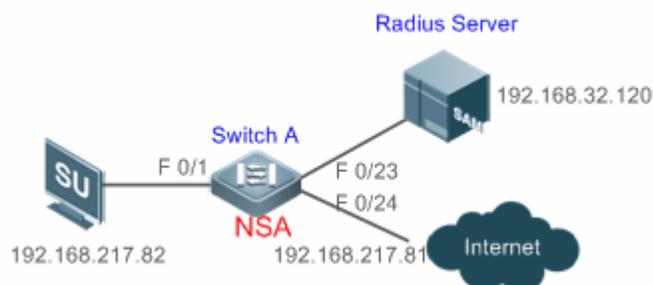
- 服务器正常可达期间,用户只能使用正确的用户名和密码认证通过才能上线
- 服务器不可达期间,认证用户具体得到授权范围网络相关命令

配置举例

配置服务器失效旁路认证

【网络环境】

图 4-8



- 【配置方法】
- 服务器上注册设备的 ip 信息,并配置设备和服务器的通信密钥
 - 服务器上创建账号信息
 - 设备使能 aaa
 - 设备配置 radius 参数,以及探测服务器可达的功能
 - 设备接口上使能 802.1x 认证和多 MAB 认证功能

如下为设备上的相关配置,服务器端的配置请参考具体服务器的配置指导手册:

```

ruijie# configure terminal
ruijie (config)# aaa new-model
ruijie (config)# radius-server host 192.168.32.120
ruijie (config)# radius-server key ruijie
ruijie (config)# interface FastEthernet 0/1
ruijie (config-if)# dot1x port-control auto
  
```

- 【检验方法】 测试是否可以正常认证以及认证前后的网络访问行为是否变化。
- 服务器创建账号,比如 username: test,password: test。
 - 认证服务器支持情况下,终端未认证前无法 ping 通 192.168.32.120。

- 认证服务器失效之后，终端连接上设备，认证成功，可 ping 通 192.168.32.120。
- 可以显示认证通过的用户信息

```
ruijie# show dot1x summary
ID          Username  MAC          Interface VLAN Auth-State  Backend-State
Port-Status User-Type Time
-----
-----
16778217   test     0023.aeea.4286 Fa0/1     2    Authenticated Idle         Authed
static     0days 0h10m20s
```

4.4.6 扩展功能配置

配置效果

- 部分终端采用操作系统自带认证客户端，这些客户端在终端接入网络后不一定会马上发起认证，影响用户使用网络，使用主动认证功能，可以促使这些终端接入网络后及时发起认证。
- 主动认证是指设备主动发出 request/id 报文，该报文可触发 supplicant 执行 802.1x 认证，因此可以利用该功能检测终端是否有使用 supplicant 软件，比如部署 MAB 时就需要使用此功能。
- 可认证主机列表可以限制接入端口下哪些终端可以认证，通过控制终端的接入物理位置来提高网络安全性。
- 多账号功能支持一个终端重认证时切换账号，对于一些特殊场景，比如 windows 的域认证，存在接入域时多次认证且认证时会变更账号，该功能适用这类场景。
- 默认情况下，设备使用设备本机的 mac 地址作为 802.1x 认证时的 eap 报文源 mac 地址。锐捷 supplicant 的部分版本，会根据 eap 报文的源 mac 地址来判断接入设备是否为锐捷设备，并实施一些私有特性，和这些 supplicant 配合做 802.1x 认证时，如果要使用相关私有特性，可以开启虚拟源 mac 地址功能。
- 802.1X 支持用户获取 IP 地址后再开始计费，这样可以满足服务器要求用户计费时必须携带 IP 地址的要求。用户先认证上线，可以从 supplicant 或者 dhcp snooping 等获用户的 IP，获取到 IP 地址后 802.1x 才会发起计费请求。为避免设备长时间没有获取到认证客户端的 IP 导致一直不发起计费，该功能配备了一个 IP 检测超时时间。如果在配置的时间内（默认 5min）没有获取到终端的 IP 地址，则将用户下线。
- 802.1x 功能支持在获取了认证客户端的终端信息之后再向服务器发起计费，这样可以将认证客户端的终端信息传递到服务器。在有线设备上为避免长时间没有获取到认证客户端的终端信息，允许配置相应的超时时间，超过该时间设备如果还没有获取到终端信息，可以将终端踢下线。
- 网关认证模式下，并且开启了 IP 授权的情况下，802.1x 支持有线用户的客户端(包括 802.1x 认证用户和 MAB 认证用户)迁移到新的 VLAN 和端口间时，进行 ARP 探测。也即设备发现用户迁移之后，在该用户原来的端口和 VLAN 上发出 ARP 请求探测该用户是否迁移。某些情况下，不适合进行这种 arp 探测，可以将此功能关闭。

注意事项

- 部署计费的环境中，不能开启多帐号功能，否则会影响计费准确性。
- MAB 认证需要使用到主动认证，因此部署 MAB 的环境必须使能主动认证功能
- 配置用户获取 IP 地址之后再开始计费时，需要注意：ipv4 环境且部署了锐捷 supplicant 客户端，由于 supplicant 具备上传终端 ipv4 地址的能力，因此这个环境下无需开启这个功能；部署静态 IP 的环境中无法使用该功能

配置方法

配置主动认证

- 可选配置，配置之后受控口发送主动请求，认证客户端收到主动请求之后就会发起 802.1x 认证。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x auto-req**

【参数说明】 -

【缺省配置】 开启

【命令模式】 全局模式

【使用指导】 主动认证报文的地址是组播报文；下联客户端有可能不会主动发起认证时，可以使用此命令让设备主动发起认证。受控口是 TRUNK 口时，开启此命令，可以基于 TRUNK 口的每一个 VLAN 发送主动请求

配置主动认证发送的报文数

- 可选配置，配置设备发送主动请求报文数量。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x auto-req packet-num num**

【参数说明】 *num*：主动认证报文数

【缺省配置】 不限报文数量

【命令模式】 全局模式

【使用指导】 在开启主动认证命令下，可以用来控制端口发出的主动认证报文数量，避免发送多余的报文。

配置主动认证用户检测

- 可选配置，配置受控口上有认证用户之后就不再发送主动请求。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x auto-req user-detect**

【参数说明】 -

【缺省配置】 开启

【命令模式】 全局模式

【使用指导】 配置该命令之后，ACCESS 的受控口上如果有认证用户则不再发出主动认证的报文；TRUNK 口上基于每一个 VLAN 来判断是有认证用户，有认证用户的则该 VLAN 内不再发出主动认证报文。

配置主动认证用发送报文的间隔

- 可选配置，配置设备发送主动请求报文时间间隔。

- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x auto-req req-interval time**
【参数说明】 *Time* : 主动认证的报文间隔时间
【缺省配置】 30s
【命令模式】 全局模式
【使用指导】

▾ 配置可认证主机列表

- 可选配置，配置受控口上的可认证主机列表，只有在列表中的客户端才允许进行 801.1x 认证。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x auth-address-table address mac-addr interface interface**
【参数说明】 *mac-addr* : 接入终端的 mac 地址
interface : 接入终端所在的端口
【缺省配置】 用户均可以认证
【命令模式】 全局模式
【使用指导】 需要限制可以在受控口上认证的终端时可以使用此命令

▾ 配置使用虚拟 mac 作为设备 802.1x 认证的源 mac

- 可选配置，设备的 mac 无法被我司 su 识别为我司设备时配置此命令。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x pseudo source-mac**
【参数说明】 -
【缺省配置】 关闭
【命令模式】 全局模式
【使用指导】 部分锐捷 supplicant 无法通过设备发出的 eapol 报文源 mac 认出该设备是锐捷设备，无法在认证过程中实施私有属性，此时可以开启此命令，则设备发出的 eapol 报文使用 00-1A-A9-17-FF-FF 作为源 mac，确保前述客户端可以将设备认出为锐捷设备。

▾ 配置多帐号认证

- 可选配置，允许同一个 mac 被多个帐号使用时可以配置此命令。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x multi-account enable**
【参数说明】 -
【缺省配置】 关闭
【命令模式】 全局模式
【使用指导】 DOT1X 认证中，某些环境中存在切换帐号认证的需求，比如部署 windows 的域认证，需要配置该功能，这样认证客户端可以在前一个帐号还没有下线的情况下，直接使用新的帐号发起认证。默认禁止切换帐号认证。

▾ 配置接口下的认证用户数限制

- 可选配置，可以限制受控口上线的用户数量，包括静态用户和动态用户。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x default-user-limit num**

【参数说明】 *num*:用户数限制

【缺省配置】 不限制端口用户

【命令模式】 接口模式

【使用指导】 默认不限制端口可以认证的用户数量，需要限制端口下可认证用户数时可配置此命令

▾ 配置获取 ip 后开始计费功能

- 可选配置，设备获取到客户端的 ip 地址之后，才会向服务器发出记账。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x valid-ip-acct enable**

【参数说明】

【缺省配置】 关闭

【命令模式】 全局模式

【使用指导】 开启此命令后，在开启记账的情况下，只有获取到了认证客户端的 IP 之后，设备才会发起记账，超时没有获取到 IP 则将此用户强制下线；没有开启记账时开启此命令，设备获取到 IP 之后不会发起记账，而超时没有获取到 IP 则同样会将此用户强制下线。

▾ 配置用户认证通过之后，允许等待该用户获取 ip 的时间

- 可选配置，开启获取 ip 开始计费功能，允许等待获取该用户 ip 的时间。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x valid-ip-acct timeout time**

【参数说明】 *time* : 超时时间，单位为分钟，默认为 5 分钟

【缺省配置】 默认为 5 分钟

【命令模式】 全局模式

【使用指导】 使用默认值即可，需要改变用户认证通过后等待获取 IP 的时间，可以使用此命令

▾ 配置获取终端信息后开始计费功能

- 可选配置，设备获取到客户端的终端地址之后，才会向服务器发出记账。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x valid-terminal-acct enable**

【参数说明】

【缺省配置】 关闭

【命令模式】 全局模式

【使用指导】 开启此命令后，在开启记账的情况下，只有获取到了认证客户端终端信息之后，设备才会发起记账，超时没有获取到终端信息则将此用户强制下线；没有开启记账时开启此命令，设备获取到终端信息之后不会发起记账，而超时没有获取到终端信息则同样会将此用户强制下线。

配置用户认证通过之后，允许等待该用户获取终端信息的时间

- 可选配置，开启获取终端信息开始计费功能，允许等待获取该用户终端信息的时间。
- 在设备开启 802.1x 认证之后配置

【命令格式】 **dot1x valid-terminal-acct timeout time**

【参数说明】 *time* : 超时时间，单位为分钟，默认为 5 分钟

【缺省配置】 默认为 5 分钟

【命令模式】 全局模式

【使用指导】 使用默认值即可，需要改变用户认证通过后等待获取终端信息的时间，可以使用此命令

检验方法

- 无。

配置举例

- 无。

4.5 监视与维护

清除各类信息

 关闭 802.1x 认证功能后，认证用户信息可以被清除。

作用	命令
清除 802.1x 认证用户信息。	no do1x port-control auto
清除 802.1x 认证用户信息	clear dot1x user
恢复 802.1x 的默认配置	dot1x default

查看运行情况

作用	命令
查看 radius 服务器参数和状态	show radius server
查看 802.1x 功能状态和协议参数	show dot1x
查看可认证主机列表	show dot1x auth-address-table
查看主动认证状态	show dot1x auto-req
查看接口受控情况	show dot1x port-control
查看客户端探测功能状态和参数	show dot1x probe-timer
查看认证用户表项信息	show dot1x summary
查看 equest/challenge 报文重传次数	show dot1x max-req

查看受控口信息	show dot1x port-control
查看过滤非锐捷客户端开关的状态	show dot1x private-supplicant-only
查看重认证开关的状态	show dot1x re-authentication
查看 request/id 报文重传次数	show dot1x reauth-max
查看认证失败之后的静默时间	show dot1x timeout quiet-period
查看重认证周期	show dot1x timeout re-authperiod
查看服务器超时时间	show dot1x timeout server-timeout
查看客户端超时时间	show dot1x timeout supptimeout
查看 request/id 报文重传间隔	show dot1x timeout tx-period
根据 id 来查看用户信息	show dot1x user id
根据用户 mac 来查看用户信息	show dot1x user mac
根据用户名来查看用户信息	show dot1x user name
查看 RIPT 的 AP 上的认证用户概要信息	show dot1x ript summary
根据 id 来 RIPT 的 AP 上的查看用户信息	show dot1x ript user id
根据用户 mac 来 RIPT 的 AP 上的查看用户信息	show dot1x ript user mac
根据用户名来 RIPT 的 AP 上的查看用户信息	show dot1x ript user name

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
AAA 调试信息 (详见 AAA 配置手册)	debug aaa
Radius 调试信息(详见 Radius 配置手册)	debug radius
打开 dot1x 事件相关的调试开关	debug dot1x event
打开 dot1x 报文处理相关的调试开关	debug dot1x packet
打开 dot1x 认证状态机相关的调试开关	debug dot1x stm
打开 dot1x 内部通信相关的调试开关	debug dot1x com
打开 dot1x 错误相关的调试开关	debug dot1x error

5 Web 认证

5.1 概述

5.1.1 Web认证概述

Web 认证是一种对用户访问网络的权限进行控制的身份认证方法，这种认证方法不需要用户安装专用的客户端认证软件，使用普通的浏览器软件就可以进行身份认证。

未认证用户使用浏览器上网时，网络设备会强制浏览器访问特定站点，也就是 Web 认证服务器，通常称为 Portal 服务器。用户无需认证即可访问 Portal 服务器上的服务，比如下载安全补丁、阅读公告信息等。当用户需要访问认证服务器以外的其它网络资源时，就必须通过浏览器在 Portal 服务器上进行身份认证，只有认证通过后才可以使用网络资源。

除了认证上的便利性之外，由于 Portal 服务器和用户的浏览器有页面交互，可以利用这个特性在 Portal 服务器页面放置一些广告、通知、业务链接等个性化的服务。

锐捷 Web 认证概述

锐捷 web 认证有 2 个版本，不同版本的 Web 认证流程不同，我们将其分别称为锐捷一代 Web 认证和锐捷二代 Web 认证。细节详见功能详解章节。

 由于一代 Web 认证和二代 Web 认证存在较大差异，配置参数也差别很大，因此在配置 Web 认证相关功能前必须仔细阅读对应章节内容，避免配置错误。

 下文仅介绍 Web 认证的相关内容。

协议规范

- HTTP : RFC1945 RFC2068
- HTTPS : RFC2818
- SNMP : RFC1157 RFC 2578
- RADIUS : RFC2865 RFC2866 RFC3576

5.2 典型应用

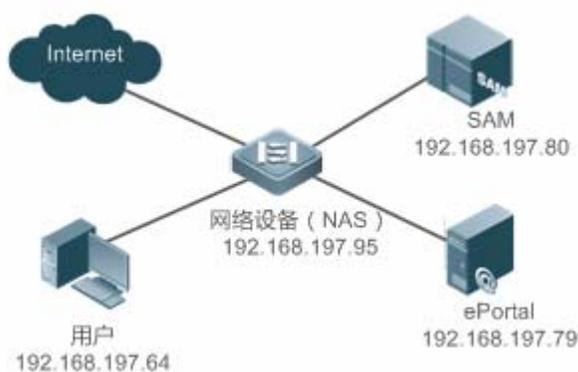
典型应用	场景描述
web认证基本场景	常见的二层基本认证场景，设备、portal 服务器、RADIUS 服务器组成认证体系，用户通过二层网络连接设备

5.2.1 web认证基本场景

应用场景

- 如图 5-1 所示
- 在网络设备上部署 web 认证方案
- 下联用户需要通过 web 认证才能访问 internet

图 5-1 Web 认证方案网络拓扑图



【注释】 web 认证方案可以用于二三层网络，此处以二层接入设备为例
SAM 为锐捷 RADIUS 服务器软件，ePortal 为锐捷 portal 服务器软件

功能部署

- 在 NAS 上用户所在接口开启 web 认证受控。
- 在 NAS 上配置 web 认证服务器和加密密钥。
- 在 NAS 上配置 web 认证服务器 SNMP 通信参数。
- 在 web 认证服务器和 SAM 服务器上配置相匹配的通信参数(仅一代认证)。
- 在 SAM 服务器上配置开通用户账户。
- 在 NAS 上配置 AAA 功能和方法列表(仅二代认证)。
- 在 NAS 上配置 RADIUS 服务器地址(仅二代认证)。
- 在 NAS 上配置 web 认证方法列表名(仅二代认证)。

5.3 功能详解

基本概念

一代 web 认证方案

一代 web 认证方案需要锐捷专有 ePortal 服务器软件的配合，用户通过 ePortal 软件提供的认证页面提交认证信息，ePortal 服务器直接向相应的 RADIUS 服务器请求认证，认证通过后将用户信息通过 SNMP 协议通告设备，由设备完成用户的准入控制。一代 web 认证通过私有 SNMP 节点进行认证通信，认证记账功能由 ePortal 服务器承载，减轻了设备的业务压力。

二代 web 认证方案

二代 web 认证方案是锐捷一代 web 认证方案的升级版，可以扩展兼容中国移动 portal 协议规范。portal 服务器单纯负责用户页面交互部分，较为简单；而 RADIUS 服务器认证交互部分由设备来实现；portal 和设备间的交互遵循《中移动 portal 协议规范》。用户通过 ePortal 软件提供的认证页面提交认证信息，ePortal 服务器将用户信息通过 portal 协议告知网络设备，网络设备利用该身份信息完成 RADIUS 服务器认证，对合法用户进行准入，同时将结果信息回应给 ePortal 服务器。

二代 web 认证方案由于主要流程都是在设备上完成，对设备的要求比较高，处理能力压力大；但是 ePortal 服务器得到了简化，同时使用业界支持度较高的中国移动 portal 规范进行交互，使得各设备厂商和服务器厂商能够开发出兼容的产品。

锐捷各类 Web 认证的对比

认证角色：

- 客户端：功能一样。
- 网络设备：锐捷一代 Web 认证里面，网络设备只做重定向，以及接收 Portal Server 下发的用户是否可上网的通知。锐捷二代 Web 认证里面，网络设备需要负责重定向、用户的认证、通告 Portal Server 认证是否成功。
- Portal Server：锐捷一代 Web 认证里面，PortalServer 负责和客户端的页面交互、用户的认证、通告网络设备用户认证是否可上网。锐捷二代 Web 认证里面，Portal Server 负责和客户端的页面交互、通告网络设备用户的认证信息、接收网络设备通告的用户是否认证成功。
- Radius Server：功能一样。

认证流程：

- 锐捷二代 Web 认证将认证记账从 Portal Server 迁移到网络设备。
- 锐捷二代 Web 认证由于认证在网络设备上，因此就无需等待来自 Portal Server 发送的用户是否可上网的通告。

下线流程：

- 锐捷一代 Web 认证中，计费结束报文是 Portal Server 发起的。锐捷二代 Web 认证中，计费结束报文是网络设备发起的。

- 实际网络中是部署一代 Web 认证，还是部署二代 Web 认证，取决于所使用的 Portal 服务器的类型。
- 各类 web 认证中有部分参数是可以共用，有些是不能共用的，本文档用章节分离的形式给出了区分，请仔细阅读，避免参数配置不当导致 Web 认证使用不正常。

功能特性

功能特性	作用
锐捷一代web认证	实现锐捷一代 web 认证方案
锐捷二代web认证	实现锐捷二代 web 认证方案

5.3.1 锐捷一代web认证

HTTP 拦截

HTTP 拦截指网络设备将原本需要转发的 HTTP 报文拦截下来，不进行转发。这些 HTTP 报文是连接在网络设备下的用户的浏览器所发出的，但目的并不是网络设备本身。例如，某用户通过 IE 浏览器访问 www.google.com，网络设备本应该将这些 HTTP 请求报文转发到网关的，但如果启动 HTTP 拦截，这些报文将不被转发。

HTTP 拦截之后，网络设备需要将用户的 HTTP 连接请求转向自己，于是网络设备和用户之间将建立起连接会话。网络设备将利用 HTTP 重定向功能，将重定向页面推送给用户，用户的浏览器上将弹出一个页面，这个页面可以是认证页面，也可以是下载软件的链接等等。

在 Web 认证功能中，哪些用户所发出的到哪个目的端口的 HTTP 报文需要进行拦截，哪些不需要，都是可以设置的。一般地，未经过认证的用户发出的 HTTP 请求报文会被拦截，已通过认证的用户将不被拦截。HTTP 拦截是 Web 认证功能的基础，一旦浏览器发出的 HTTP 报文被拦截，就会自动触发 Web 认证的过程。

HTTP 重定向

根据 HTTP 协议规定，正常情况下，用户的浏览器发出 HTTP GET 或 HEAD 请求报文后，如果接收一方能够提供资源，则以 200 报文响应，如果本地不能提供资源，则可以使用 302 报文响应。在 302 响应报文中，提供了一个新的站点路径，用户收到响应后，可以向这个新的站点重新发出 HTTP GET 或 HEAD 报文请求资源，这就是重定向。

HTTP 重定向是 Web 认证的重要环节，是发生在 HTTP 拦截之后的，利用的就是 HTTP 协议中的 302 报文的特性。HTTP 拦截过程将使得网络设备和用户之间建立起连接会话，随后用户将（本应发给其他站点的）HTTP GET 或 HEAD 报文发给网络设备，网络设备收到后，回应以 302 报文，并且在 302 报文中加入重定向页面的站点路径，这样用户将向这个站点路径重新发出请求，就会获取到重定向的页面。

工作原理

组网拓扑图同 图 5-2 Web 认证方案网络拓扑图。

Web 认证的角色：

1. 认证客户端：通常是浏览器，运行 HTTP 协议，用户通过浏览器上网时浏览器将发出 HTTP 请求。
2. 网络设备：在网络拓扑中一般是接入层设备（例如在无线网络中可以是无线 AP），一般与用户终端设备直接相连接，在设备上需要启动 Web 认证功能。

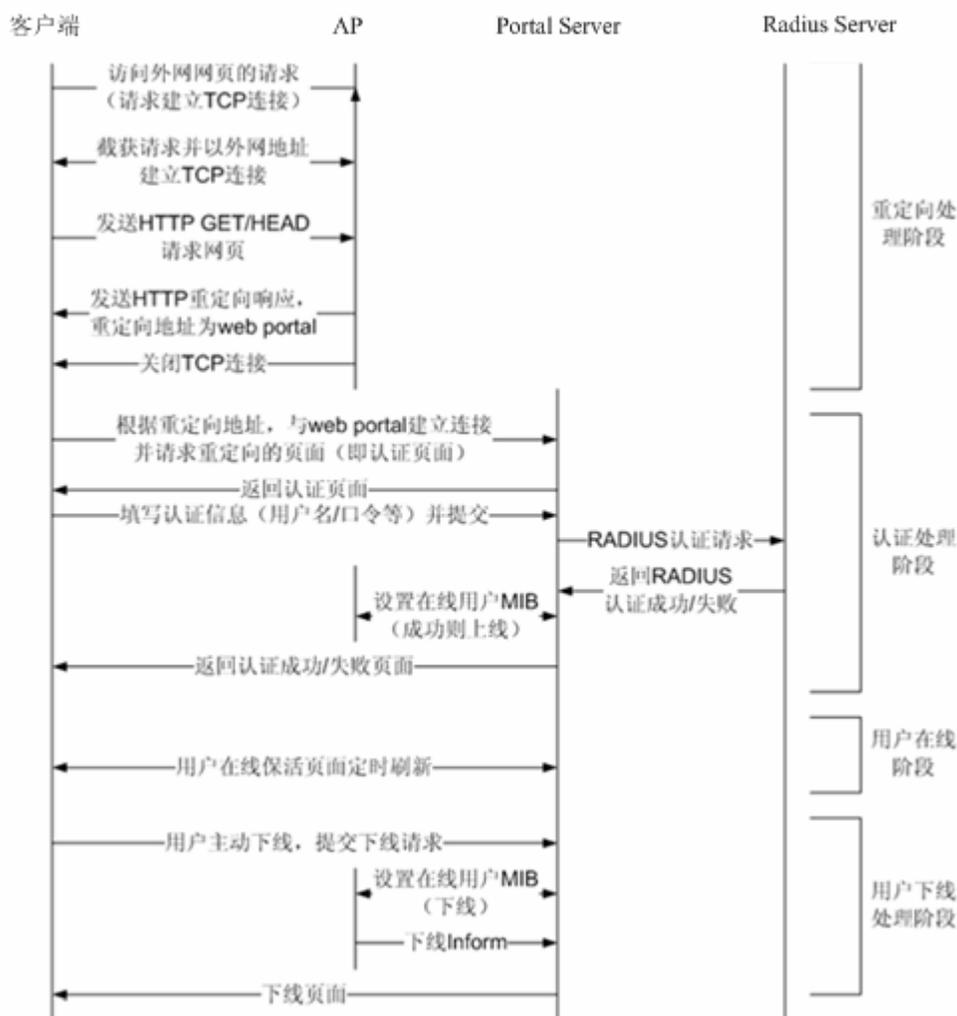
3. Portal Server：提供 Web 认证的认证界面和相关操作。Portal Server 接受认证客户端发出的基于 HTTP 的认证请求，提取其中的账号信息，将此信息发送到认证服务器进行认证，然后通告用户和网络设备认证结果。图中为锐捷 ePortal 服务器。
4. Radius 服务器：提供基于 RADIUS 协议的远程用户认证，Portal 服务器从 HTTP 中获取用户的认证账号信息，然后通过 RADIUS 协议向 Radius 服务器请求认证。Radius 服务器通过 RADIUS 协议向 Portal 服务器反馈认证结果。图中为锐捷 SAM 服务器。

Web 认证的流程：

- 在认证之前，设备将未认证用户发出的所有 HTTP 请求都拦截下来，并重定向到 Portal 服务器去，这样在用户的浏览器上将弹出一个认证页面。
- 在认证过程中，用户在认证页面上输入认证信息（用户名、口令、校验码等等）与 Portal 服务器交互，完成身份认证的功能。
- 在认证通过后，Portal 服务器将通知设备该用户已通过认证，设备将允许用户访问互联网资源。

详细的原理图如下图所示：(图中以 AP 为例)

图 5-3 锐捷一代 Web 认证流程图



用户下线流程：

用户下线分两种类型,一种是设备端检测到用户下线,比如用户时长到达,流量用完,链路断开等;一种是 portal 服务器端发现用户下线,比如用户通过下线页面触发了下线申请,保活页面失效等。

- 第一种情况,设备发现用户下线时,通告 Portal Server 用户下线,服务器设置设备删除用户信息(通过 SNMP 协议)。Portal 服务器向用户返回下线页面。
- 第二种情况,Portal Server 发现用户下线后,通告设备用户下线(通过 SNMP 协议)。Portal 服务器向用户返回下线页面。
- 2 种情况下,Portal Server 都会向 Radius Server 发起计费结束请求,通告 Radius Server 用户下线。

相关配置

创建配置模板

缺省情况下无配置

在全局配置模式下使用 **web-auth template eportalv1** 命令创建模板

可以使用该配置模板进行 web 认证功能

配置服务器 IP 地址

缺省情况下无配置

在模板配置模式使用 **ip { ip-address }**来配置服务器 IP 地址

访问服务器的请求被设备放行,并且支持对发往服务器的请求进行限速保护

配置服务器认证 URL

缺省情况下无配置

在模板配置模式使用 **url { url-string }**来配置

用户重定向到的 URL 地址,通常使用 portal 认证页面地址

配置通信加密密钥

缺省情况下无配置

在全局配置模式下使用 **web-auth portal key { string }**来配置通信加密密钥

用作 URL 参数加密,避免信息泄漏

开启 web 认证

缺省情况下该功能关闭

在接口模式下使用 **web-auth enable** 命令开启用户所在端口 web 认证受控。

开启认证后端口下未认证用户会被重定向到认证页面

配置 web 认证 SNMP 协议服务器

缺省情况下无配置

在全局配置模式下使用 **snmp-server host** { *ip-address* } **version 2c** { *community-string* } **web-auth** 来配置 web 认证 SNMP 协议服务器

设备可以通过 inform/trap 报文通告用户下线消息给配置的服务器

配置服务器 SNMP 通信团体字

缺省情况下无配置

在全局配置模式下使用 **snmp-server community** {*community-string*} **rw** 来配置服务器 SNMP 通信团体字
服务器通过该团体字来读写设备的用户表项信息

启用 SNMP TRAP/INFORM 通告功能

缺省情况下无配置

在全局配置模式下使用 **snmp-server enable traps web-auth** 来配置启用 SNMP TRAP/INFORM 通告功能
使能通告功能，用于设备向服务器通告用户下线消息

5.3.2 锐捷二代web认证

HTTP 拦截

同锐捷一代 Web 认证的 HTTP 拦截技术。

HTTP 重定向

同锐捷一代 Web 认证的 HTTP 重定向技术。

工作原理

组网拓扑图同 图 5-4 Web 认证方案网络拓扑图。

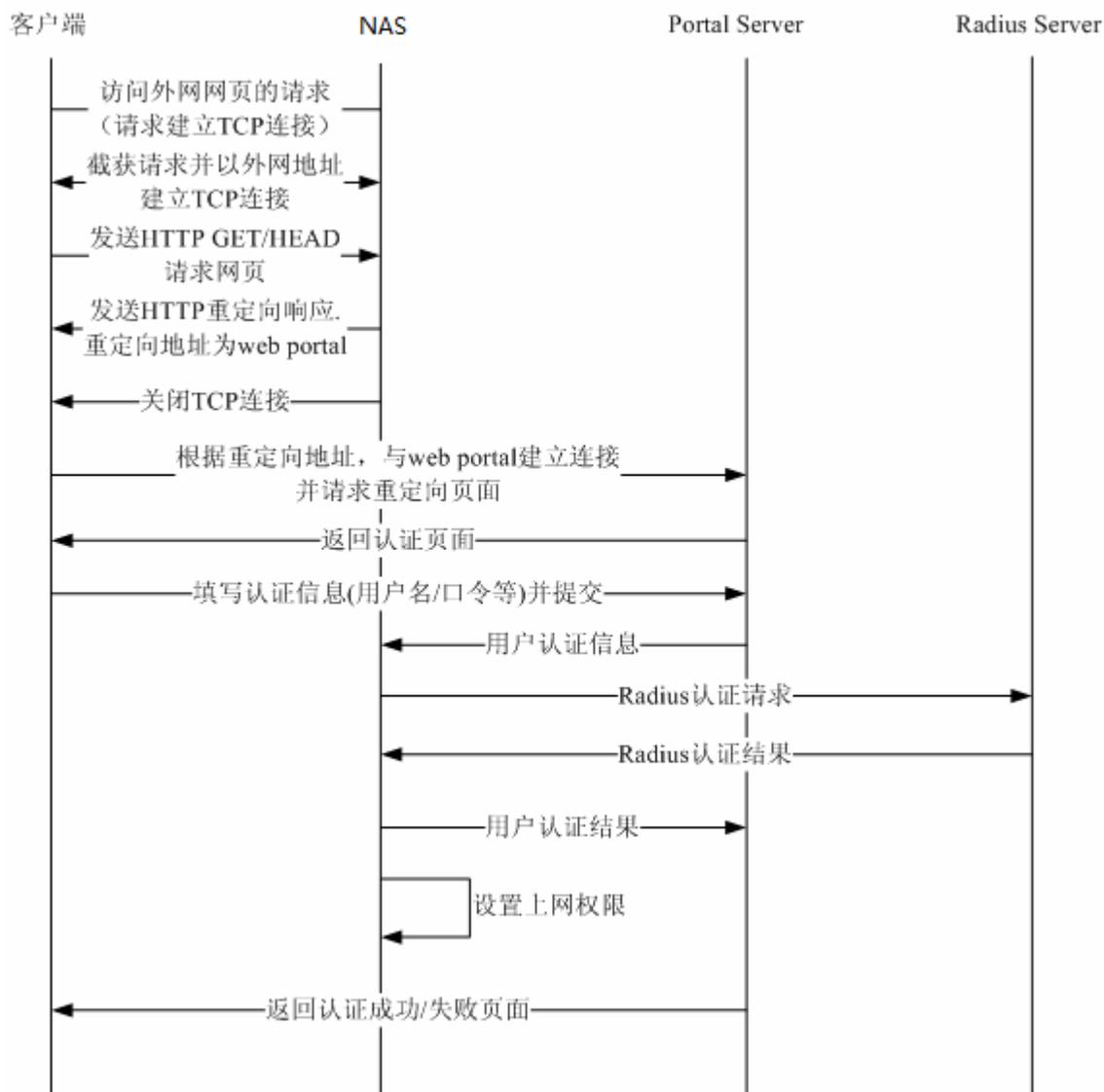
Web 认证的角色：

- 认证客户端：通常是一个浏览器，运行 HTTP 协议，用户通过浏览器上网时浏览器将发出 HTTP 请求。
- 网络设备：在网络拓扑中一般是接入层设备（例如在无线网络中可以是无线 AP），一般与用户终端设备直接相连接，在设备上需要启动 Web 认证功能。设备接收 Portal Server 发过来的用户认证信息，并向 Radius Server 发起认证请求，根据认证结果设置用户是否可以上网，同时向 Portal Server 反馈认证结果。
- Portal Server：提供 Web 认证的认证界面和相关操作。Portal Server 接受认证客户端发出的基于 HTTP 的认证请求，提取其中的账号信息，将此信息发送到网络设备，同时根据网络设备反馈的认证结果，通过页面反馈给用户。图中为锐捷 ePortal 服务器。
- Radius Server：提供基于 radius 协议的远程用户认证。图中为锐捷 SAM 服务器。

Web 认证主要流程：

- 在认证之前，设备将未认证用户发出的所有 HTTP 请求都拦截下来，并重定向到 Portal 服务器去，这样在用户的浏览器上将弹出一个认证页面。
- 在认证过程中，用户在认证页面上输入认证信息（用户名、口令、校验码等等）与 Portal 服务器交互。
- Portal 服务器将用户的认证信息发给设备
- 设备向 Radius Server 发起认证，并将认证结果反馈给 Portal 服务器。
- Portal 服务器向用户返回页面，提示认证结果（成功或失败）。

图 5-5 锐捷二代 Web 认证流程



用户下线流程：

用户下线分两种类型,一种是设备端检测到用户下线,比如用户时长到达,流量用完,链路断开等;一种是 portal 服务器端发现用户下线,比如用户通过下线页面触发了下线申请等。

- 第一种情况,设备发现用户下线时,通告 Portal Server 用户下线(通过 portal 协议),然后删除用户信息。Portal 服务器向用户返回下线页面。
- 第二种情况,Portal Server 发现用户下线后,通告设备用户下线(通过 portal 协议)。Portal 服务器向用户返回下线页面。
- 2 种情况下,设备都会向 Radius Server 发起计费结束请求,通告 Radius Server 用户下线。

相关配置

创建配置模板

缺省情况下无配置

在全局配置模式下使用 **web-auth template**{eportalv2 | *template-name* v2}命令创建二代模板

可以使用该配置模板进行 web 认证功能

配置服务器 IP 地址

缺省情况下无配置

在模板配置模式使用 **ip** { *ip-address* }来配置服务器 IP 地址

访问服务器的请求被设备放行,并且支持对发往服务器的请求进行限速保护

配置服务器认证 URL

缺省情况下无配置

在模板配置模式使用 **url** { *url-string* }来配置

用户重定向到的 URL 地址,通常使用 portal 认证页面地址

配置通信加密密钥

缺省情况下无配置

在全局配置模式下使用 **web-auth portal key** { *string* }来配置通信加密密钥

用作 URL 参数加密,避免信息泄漏

开启 web 认证

缺省情况下该功能关闭

在接口模式下使用 **web-auth enable**{eportalv2 | *template-name*}命令开启用户所在端口 web 认证受控。

开启认证后端口下未认证用户会被重定向到认证页面

配置启用 AAA 认证

缺省情况下关闭 AAA 认证

在全局配置模式下使用 **aaa new-model** 命令来开启 AAA 认证功能

二代 Web 认证功能需要依赖于 AAA 认证功能，使用时需要开启 AAA 功能

配置 RADIUS 服务器和通信密钥

缺省情况下无配置

在全局配置模式下使用 **radius-server host** 命令来配置 RADIUS 服务器和通信密钥

对应于 web 认证中的 RADIUS 服务器。用于为 Web 认证用户进行身份校验。

配置 AAA 模块 web 认证的方法列表

缺省情况下无配置

在全局配置模式下使用 **aaa authentication web-auth** 命令来配置二代 web 认证的认证方法

Web 认证功能可以配置使用该方法列表来进行认证交互

配置 AAA 模块网络记账方法列表

缺省情况下无配置

在全局配置模式下使用 **aaa accounting network** 命令来配置网络记账方法

Web 认证功能可以配置使用该方法列表来进行记账交互

配置 web 认证使用的 AAA 认证方法列表名

缺省情况下使用 default 方法

在模板配置模式下使用 **authentication** 命令来配置

Web 认证功能通过该方法列表名字向 AAA 功能发起认证请求

配置 web 认证使用的 AAA 认证记账列表名

缺省情况下使用 default 方法

在模板配置模式下使用 **accounting** 命令来配置

Web 认证功能通过该方法列表名字向 AAA 功能发起记账请求

配置 portal 服务器的通信 UDP 端口号

缺省情况下使用 50100 端口号

在模板配置模式下使用 **port** 命令来配置

设备通过发送报文到该端口来同 portal 服务器进行交互

5.4 配置详解

配置项	配置建议 & 相关命令
配置一代Web认证功能	 必须配置。用于配置一代 web 认证功能基本参数

	web-auth template eportalv1	创建一代 eportalv1 模板
	ip { <i>ip-address</i> }	配置服务器的 IP 地址
	url { <i>url-string</i> }	配置服务器的主页地址
	web-auth portal key { <i>key-string</i> }	配置服务器通信密钥
	snmp-server community { <i>community-string</i> } rw	配置 SNMP 通信团体字
	snmp-server host { <i>ip-address</i> } inform version 2c { <i>community-string</i> } web-auth	配置 SNMP 通信服务器
	snmp-server enable traps web-auth	使能 web 认证 trap/inform 通告
	web-auth enable	在接口上使能 web 认证
配置二代web认证功能	 必须配置。用于配置二代 web 认证功能基本参数	
	aaa new-model	开启 AAA 功能
	radius-server host { <i>ip-address</i> } [auth-port <i>port-number</i>] [acct-port <i>port-number</i>] key { <i>string</i> }	配置 RADIUS 服务器和密钥
	aaa authentication web-auth { default <i>list-name</i> } method1 [<i>method2...</i>]	配置 web 认证方法列表(使用 RADIUS 认证)
	aaa accounting network { default <i>list-name</i> } start-stop <i>method1</i> [<i>method2...</i>]	配置网络记账方法列表(使用 RADIUS 记账)
	web-auth template { eportalv2 <i>portal-name v2</i> }	创建二代认证模板
	ip { <i>ip-address</i> }	配置服务器的 IP 地址
	url { <i>url-string</i> }	配置服务器的主页地址
	web-auth portal key { <i>key-string</i> }	配置服务器通信密钥
web-auth enable { eportalv2 <i>template-name</i> }	在接口上使能 web 认证	
配置认证方法列表名	 可选配置。用于指定认证方法列表名，和 AAA 模块的方法列表配置保持一致	
	authentication { <i>mlist-name</i> }	配置认证方法列表名(仅二代)
配置记账方法列表名	 可选配置。用于指定记账方法列表名，和 AAA 模块的方法列表配置保持一致	
	accounting { <i>mlist-name</i> }	配置记账方法列表名(仅二代)
配置portal服务器通信端口	 可选配置。用于指定 portal 服务器通信端口，和服务器支持的通信端口保持一致	
	port { <i>port-num</i> }	配置 portal 服务器通信端口
配置绑定模式	 可选配置。用于指定用户表项绑定模式	
	bindmode { ip-mac-mode ip-only-mode }	配置模板表项绑定模式
设置重定向的HTTP端口	 可选配置。用于指定重定向 TCP 拦截端口，补充拦截环境中的特定端口报文进行重定向	

	<code>http redirect port { port-num }</code>	配置重定向 TCP 拦截端口
设置web认证模块SYSLOG功能	 可选配置。用于配置 web 认证 SYSLOG 功能	
	<code>web-auth logging enable { num }</code>	配置 web 认证 SYSLOG 输出速率
设置未认证用户的最大HTTP会话数	 可选配置。用于调整 HTTP 会话数限制，对于后台会话数较多的场景需要放宽限制	
	<code>http redirect session-limit { session-num } [port { port-session-num }]</code>	配置用户的 HTTP 会话限制数
设置维持重定向连接的超时时间	 可选配置。用于修改重定向连接超时时间，在网络环境较差时调大参数有利于完成重定向	
	<code>http redirect timeout { seconds }</code>	用于设置重定向连接超时时间
设置直通ARP资源范围	 可选配置。用于放行指定地址的 ARP，开启 ARP CHECK 时需要放行网关 ARP	
	<code>http redirect direct-arp { ip-address [ip-mask] }</code>	用于配置直通 ARP 资源范围
设置无需认证用户IP范围	 可选配置。用于配置特殊用户不用认证就能上网	
	<code>web-auth direct-host { ip-address [ip-mask] [arp] } [port interface-name]</code>	用于配置免认证用户
配置DHCP地址核查	 可选配置。用于检测认证用户的 ip 地址是否是 DHCP 服务器分配的，如果不是则拒绝认证请求。	
	<code>web-auth dhcp-check</code>	检测终端的 ip 地址是否是 DHCP 服务器分配

5.4.1 配置一代web认证功能

配置效果

未认证用户能够被重定向到认证页面并完成认证

注意事项

无

配置方法

配置 portal 服务器

- 必须配置，要成功应用 Web 认证功能，必须设置并应用 portal 服务器。

- 当接入/汇聚设备发现未认证用户试图通过 HTTP 访问网络资源时，设备将用户的访问请求重定向到指定的 portal 认证页面，通过认证页面，引导用户向认证服务器发起认证。Portal 服务器地址将被设置为一个特殊的免认证的网络资源，未认证用户才可以直接与这个地址进行 HTTP 通讯。

配置设备与认证服务器之间的通信密钥

- 必须配置，要成功应用 Web 认证功能，必须设置接入/汇聚设备与认证服务器进行通信的密钥。
- 当设备发现未认证用户在访问网络资源时，设备将通过重定向功能，向用户弹出认证页面，通过认证页面，引导用户向认证服务器发起认证。在认证过程中，设备与认证服务器间通过密钥对部分数据进行加密，以加强安全性。

设置设备与认证服务器之间的 SNMP 网管参数

- 必须配置，要成功应用 Web 认证功能，必须设置设备与认证服务器之间的 SNMP 网管通信参数。
- 接入/汇聚设备和认证服务器之间通过 SNMP/MIB 对认证用户进行管理，在设备上，使用 MIB 来管理认证用户表，认证服务器通过访问这个 MIB，可以获取用户相关的统计信息，以及进行控制用户的上线、下线的操作。当用户下线时，设备将发送 SNMP-Inform 消息给认证服务器。

在端口上开启 Web 认证功能

- 必须配置。
- 当 Web 认证功能基于端口开启时，默认情况下，端口未开启 Web 认证功能，此时这个端口下所连接的用户不进行 Web 认证。

检验方法

- 未认证用户被要求认证
- 已认证用户可以正常使用网络

相关命令

创建模板

- 【命令格式】 **web-auth template eportalv1**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 eportalv1 为默认的一代 web 认证模板

配置服务器 IP

- 【命令格式】 **ip { ip-address }**
- 【参数说明】 portal 服务器的地址
- 【命令模式】 web 认证模板配置模式
- 【使用指导】 -

配置服务器 URL

- 【命令格式】 **url { url-string }**
- 【参数说明】 portal 服务器的认证页面地址
- 【命令模式】 Web 认证模板配置模式
- 【使用指导】 以 http://或 https://开头

配置服务器加密密钥

- 【命令格式】 **web-auth portal key { key-string }**
- 【参数说明】 portal 服务器的加密密钥，配置设备与认证服务器进行通信的密钥；密钥最大长度为 255 个字符。
- 【命令模式】 全局配置模式
- 【使用指导】 -

配置服务器 SNMP 通信团体字

- 【命令格式】 **snmp-server community { community-string } rw**
- 【参数说明】 *community-string* : Community 字符串；
rw : 由于需要对 MIB 进行 Set 操作，因此需要设置成 RW，支持读写操作；
- 【命令模式】 全局配置模式
- 【使用指导】 设置 SNMP Community，认证服务器可以使用这个 Community 管理接入/汇聚设备上的在线用户。

配置服务器 SNMP 通信服务器

- 【命令格式】 **snmp-server host { ip-address } inform version 2c { community-string } web-auth**
- 【参数说明】 *ip-address* : 目的主机地址，也就是认证服务器的地址。
community-string : 通信团体字，发送 SNMP-Inform 消息使用的 Community 字符串。
- 【命令模式】 全局配置模式
- 【使用指导】 设置发送 Web 认证消息的目的主机，类型、版本、Community 等参数。
inform : 设置发送 SNMP-Inform 类型的消息。由于接入/汇聚设备在用户下线的时候向认证服务器发送消息，为了防止消息丢失，所以采用 SNMP-Inform 而不是 SNMP-Trap。
version 2c : SNMPv2 以后的版本才支持 SNMP-Inform 类型，因此这里不能设置为 SNMPv1。
web-auth : 指明发送 Web 认证消息采用上述的参数。

 SNMP 的配置命令和其他具体内容参见“SNMP 配置指南”中的相关章节。

 这里列出的 SNMP 通信参数是以 SNMPv2 的设置为例的，如果要求设备和认证服务器之间的 SNMP 网管通信有更高的安全性，可以考虑采用 SNMPv3。这样，SNMP Community 的设置需要改为 SNMP User，并且 SNMP-Inform 的版本也需要改为 SNMPv3 版本的，另外还需要设置和 SNMPv3 相关的安全参数，具体参见“SNMP 配置”中相关的章节，这里不再进行详细介绍。

配置使能 web 认证 trap/inform 通告

- 【命令格式】 **snmp-server enable traps web-auth**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 设置接入/汇聚设备允许向外发送 Web 认证的消息，消息类型包括 Trap 和 Inform。
web-auth : 即 Web 认证的消息。

配置接口上启用 web 认证

- 【命令格式】 **web-auth enable**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 -

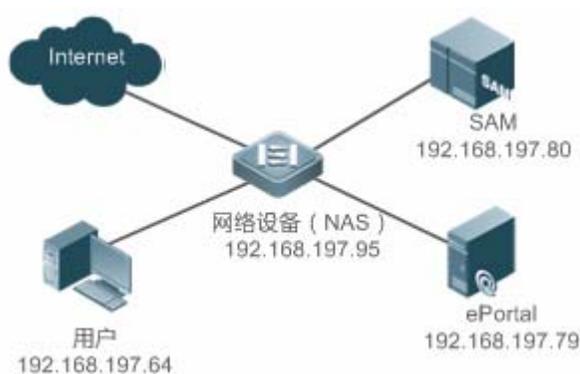
配置举例

i 以下配置举例，仅介绍与 Web 认证相关的配置。

一代 web 认证

【网络环境】

图 5-6



- 【配置方法】
- 在网络设备上设置认证服务器的 IP 地址及与认证服务器进行通信的密钥(ruijie)
 - 在网络设备上设置认证页面的主页地址
 - 设置网络设备与认证服务器之间的 SNMP 网管参数(团体字 public)
 - 在网络设备上对 GigabitEthernet 0/2、GigabitEthernet 0/3 两个端口上开启 Web 认证功能

```
Ruijie# config
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
Ruijie(config)#web-auth template eportalv1
```

```
Ruijie(config.tmpl.eportalv1)#ip 192.168.197.79
```

```
Ruijie(config.tmpl.eportalv1)#exit
```

```
Ruijie(config)# web-auth portal key ruijie
```

```
Ruijie(config)# web-auth template eportalv1
```

```
Ruijie(config.tmpl.eportalv1)#url http://192.168.197.79:8080/eportal/index.jsp
```

```
Ruijie(config.tmpl.eportalv1)#exit
```

```
Ruijie(config)# snmp-server community public rw
```

```
Ruijie(config)# snmp-server enable traps web-auth
```

```
Ruijie(config)# snmp-server host 192.168.197.79 inform version 2c public web-auth
Ruijie(config)# exit
```

```
Ruijie(config)# interface range GigabitEthernet 0/2-3
Ruijie(config-if-range)# web-auth enable
Ruijie(config-if-range)# exit
```

【检验方法】

- web 认证配置是否成功

```
Ruijie(config)#show running-config
...
snmp-server host 192.168.197.79 inform version 2c public web-auth
snmp-server enable traps web-auth
snmp-server community public rw
...
web-auth template eportalv1
  ip 192.168.197.79
  url http://192.168.197.79:8080/eportal/index.jsp
!
web-auth portal key ruijie
...
interface GigabitEthernet 0/2
  web-auth enable
!
interface GigabitEthernet 0/3
  web-auth enable
```

```
Ruijie#show web-auth control
```

Port	Control	Server Name	Online User Count
-----	-----	-----	-----
...			
GigabitEthernet 0/2	On	eportalv1	0
GigabitEthernet 0/3	On	eportalv1	0
...			

```
Ruijie#show web-auth template
Webauth Template Settings:
-----
Name:      eportalv1
Url:       http://17.17.1.21:8080/eportal/index.jsp
Ip:        17.17.1.21
BindMode:  ip-mac-mode
Type:      v1
.....
```

常见错误

无

5.4.2 配置二代web认证功能

配置效果

未认证用户能够被重定向到认证页面并完成认证

注意事项

无

配置方法

▾ 配置开启 AAA 认证

- 必须配置，要使用二代 Web 认证功能，必须开启 AAA 认证。
- 二代 web 认证向服务器发起认证的功能由设备完成，在设备 AAA 功能实现。

▾ 配置 RADIUS 服务器和密钥

- 必须配置，要成功应用二代 Web 认证功能，必须设置 RADIUS 服务器。
- 用户账户信息保存在 RADIUS 服务器上，设备需要连接 RADIUS 服务器来确认用户身份合法性。

▾ 配置 AAA 中 web 认证方法

- 必须配置，要成功应用二代 Web 认证功能，必须设置 AAA 认证方法。

- 认证方法列表将 web 认证的请求和 RADIUS 服务器关联起来，设备依据认证方法列表来选择认证方式和对应的服务器。

配置 AAA 网络记账方法

- 必须配置，要成功应用二代 Web 认证功能，必须设置 AAA 网络记账方法。
- 记账方法用于关联对应的记账方式和服务器，web 认证需要记账功能记录用户信息或费用。

配置 portal 服务器

- 必须配置，要成功应用二代 Web 认证功能，必须设置并应用 portal 服务器。
- 当接入/汇聚设备发现未认证用户试图通过 HTTP 访问网络资源时，设备将用户的访问请求重定向到指定的 portal 认证页面，通过认证页面，引导用户向认证服务器发起认证。Portal 服务器地址将被设置为一个特殊的免认证的网络资源，未认证用户才可以直接与这个地址进行 HTTP 通讯。

配置设备与认证服务器之间的通信密钥

- 必须配置，要成功应用二代 Web 认证功能，必须设置接入/汇聚设备与认证服务器进行通信的密钥。
- 当设备发现未认证用户在访问网络资源时，设备将通过重定向功能，向用户弹出认证页面，通过认证页面，引导用户向认证服务器发起认证。在认证过程中，设备与认证服务器间通过密钥对部分数据进行加密，以加强安全性。

设置全局/接口上使用的 portal 服务器

- 二代认证必须配置，要成功应用二代 Web 认证功能，必须在全局/接口上指定使用二代 portal。
- 设备会优先选择所在接口配置的 portal 服务器，如果所在接口不存在 portal 服务器配置，设备会选择全局配置的 portal 服务器，全局不存在配置时默认使用 eportalv1。设备将用户重定向到所选择的 portal 服务器。

在端口上开启 Web 认证功能

- 必须配置。
- 当 Web 认证功能基于端口开启时，默认情况下，端口未开启 Web 认证功能，此时这个端口下所连接的用户不进行 Web 认证。

检验方法

- 未认证用户被要求认证
- 已认证用户可以正常使用网络

相关命令

开启 AAA 功能

- 【命令格式】 `aaa new-model`
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 aaa 的方法列表等命令需要在功能开启后才能输入

配置 RADIUS 服务器和密钥

- 【命令格式】 **radius-server host** { *ip-address* } [**auth-port** *port-number1*] [**acct-port** *port-number2*] **key** { *string* }
- 【参数说明】
 - ip-address* : 服务器 IP 地址
 - port-number1* : 认证端口号
 - port-number2* : 记账端口号
 - string* : 密钥字符串
- 【命令模式】 全局配置模式
- 【使用指导】 认证端口默认 1812, 记账端口默认 1813

配置 AAA 中 web 认证方法列表

- 【命令格式】 **aaa authentication web-auth** { **default** | *list-name* } *method1* [*method2*...]
- 【参数说明】
 - list-name* : 方法列表名
 - method1* : 方法 1
 - method2* : 方法 2
- 【命令模式】 全局配置模式
- 【使用指导】 二代 web 认证通常使用 RADIUS 认证方法

配置网络记账方法列表

- 【命令格式】 **aaa accounting network** { **default** | *list-name* } **start-stop** *method1* [*method2*...]
- 【参数说明】
 - list-name* : 方法列表名
 - method1* : 方法 1
 - method2* : 方法 2
- 【命令模式】 全局配置模式
- 【使用指导】 二代 web 认证通常使用 RADIUS 记账方法

创建模板

- 【命令格式】 **web-auth template** { **eportalv2** | *portal-name v2* }
- 【参数说明】 自定义的 portal 服务器名
- 【命令模式】 全局配置模式
- 【使用指导】 eportalv2 为默认的二代 web 认证模板

配置服务器 IP

- 【命令格式】 **ip** { *ip-address* }
- 【参数说明】 portal 服务器的地址
- 【命令模式】 Web 认证模板配置模式
- 【使用指导】 -

配置服务器 URL

- 【命令格式】 **url** { *url-string* }
- 【参数说明】 portal 服务器的认证页面地址
- 【命令模式】 Web 认证模板配置模式

【使用指导】 以 http://或 https://开头

配置服务器加密密钥

【命令格式】 **web-auth portal key { key-string }**

【参数说明】 portal 服务器的加密密钥，。配置设备与认证服务器进行通信的密钥；密钥最大长度为 255 个字符。

【命令模式】 全局配置模式

【使用指导】 -

配置接口上启用 web 认证

【命令格式】 **web-auth enable { eportalv2 | template-name }**

【参数说明】 自定义模板名

【命令模式】 接口配置模式

【使用指导】 -

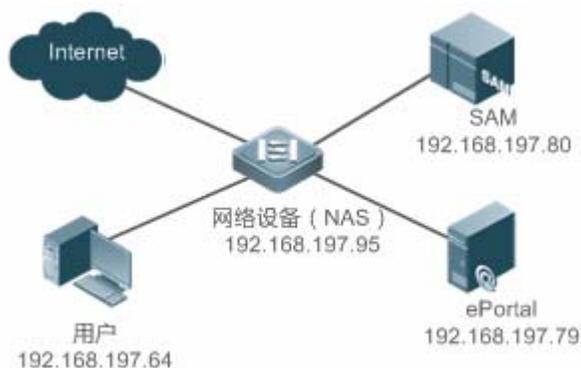
配置举例

i 以下配置举例，仅介绍与 Web 认证相关的配置。

二代 web 认证

【网络环境】

图 5-7



- 【配置方法】
- 在网络设备上开启 AAA
 - 在网络设备配置 RADIUS 服务器和密钥
 - 在网络设备配置 AAA 的 web 认证默认方法列表和默认网络记账方法列表
 - 在网络入设备上设置认证服务器的 IP 地址及与认证服务器进行通信的密钥(ruijie)
 - 在网络设备上设置认证页面的主页地址
 - 在网络设备上配置全局使用二代 portal 进行认证
 - 在网络设备上对 GigabitEthernet 0/2、GigabitEthernet 0/3 两个端口上开启 Web 认证功能

```
Ruijie#configure
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
Ruijie(config)#aaa new-model
```

```
Ruijie(config)#radius-server host 192.168.197.79 key ruijie
```

```
Ruijie(config)#aaa authentication web-auth default group radius
```

```
Ruijie(config)#aaa accounting network default start-stop group radius
```

```
Ruijie(config)#web-auth template eportalv2
```

```
Ruijie(config.tmpl.eportalv2)#ip 192.168.197.79
```

```
Ruijie(config.tmpl.eportalv2)#exit
```

```
Ruijie(config)#web-auth portal key ruijie
```

```
Ruijie(config)# web-auth template eportalv2
```

```
Ruijie(config.tmpl.eportalv2)#url http://192.168.197.79:8080/eportal/index.jsp
```

```
Ruijie(config.tmpl.eportalv2)#exit
```

```
Ruijie(config)# interface range GigabitEthernet 0/2-3
```

```
Ruijie(config-if-range)# web-auth enable eportalv2
```

```
Ruijie(config-if-range)# exit
```

【检验方法】

- web 认证配置是否成功

```
Ruijie(config)#show running-config
```

```
...
```

```
aaa new-model
```

```
aaa authentication web-auth default group radius
```

```
aaa accounting network default start-stop group radius
```

```
...
```

```
radius-server host 192.168.197.79 key ruijie
```

```
...
```

```
web-auth template eportalv2
```

```
ip 192.168.197.79
```

```
url http://192.168.197.79:8080/eportal/index.jsp
```

```
!
```

```
web-auth portal key ruijie
```

```
...
```

```
interface GigabitEthernet 0/2
```

```
web-auth enable eportalv2
```

```
!
interface GigabitEthernet 0/3
 web-auth enable eportalv2
```

```
Ruijie#show web-auth control
```

Port	Control	Server Name	Online User Count

...			
GigabitEthernet 0/2	On	eportalv2	0
GigabitEthernet 0/3	On	eportalv2	0
...			

```
Ruijie#show web-auth template
```

```
Webauth Template Settings:
```

```
-----
Name:      eportalv1
Url:       http://17.17.1.21:8080/eportal/index.jsp
Ip:        17.17.1.21
BindMode:  ip-mac-mode
Type:      v1
```

```
-----
Name:      eportalv2
Url:       http://17.17.1.21:8080/eportal/index.jsp
Ip:        17.17.1.21
BindMode:  ip-only-mode
Type:      v2
Port:      50100
Acctmlist:
Authmlist:
```

```
...
```

常见错误

无

5.4.3 配置认证方法列表名

配置效果

- 当用户提交认证信息时，portal 服务器会向设备发起认证请求，设备依据配置的认证方法列表名解析认证服务器等信息，发起认证过程。
- 配置认证方法列表名后，设备可以通过指定的方法列表名选择认证服务器进行认证。

注意事项

- 配置认证方法列表名前，必须保证在 AAA 中已经定义了该方法。对应定义命令为 **aaa authentication web-auth { default | list-name } method1 [method2...]**。

配置方法

- 可选配置
- 默认使用 default 方法，在 AAA 修改方法列表名或存在多个方法列表名时，使用本命令进行配置。

检验方法

- 在 AAA 中配置两个方法列表，方法列表 1 使用服务器 1，方法列表 2 使用服务器 2
- 在服务器 1 创建用户 a 和密码；服务器 2 创建用户 b
- 配置使用方法列表 1
- 用户使用账号 b 进行认证，认证失败
- 用户使用账户 a 进行认证，认证成功

相关命令

📄 配置认证方法列表名

- 【命令格式】 **authentication { mlist-name }**
- 【参数说明】 方法列表名
- 【命令模式】 Web 认证模板配置模式
- 【使用指导】 应与 AAA 中 web 认证方法列表名一致

配置举例

📄 配置认证方法列表名

- 【配置方法】
- 配置认证方法列表名为 mlist1

```
Ruijie(config.tmpl.eportalv2)#authentication mlist1
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie#show web-auth template
Webauth Template Settings:
-----
Name:      eportalv2
Url:       http://17.17.1.21:8080/eportal/index.jsp
Ip:        17.17.1.21
BindMode:  ip-only-mode
Type:      v2
Port:      50100
Acctmlist:
Authmlist: mlist1
```

5.4.4 配置记账方法列表名

配置效果

- 用户认证通过后，设备会自动发起记账请求，请求的对象依赖于记账方法列表的配置，通常为认证所在服务器。
- 配置记账方法列表名后，设备可以通过指定的方法列表名选择记账服务器进行记账。

注意事项

- 配置记账方法列表名前，必须保证在 AAA 中已经定义了该方法。对应定义命令为 **aaa accounting network { default | list-name } start-stop method1 [method2...]**。

配置方法

- 可选配置
- 默认使用 default 方法，在 AAA 修改方法列表名或存在多个方法列表名时，使用本命令进行配置。

检验方法

- 在 AAA 中配置两个记账方法列表，方法列表 1 使用服务器 1，方法列表 2 使用服务器 2
- 配置使用方法列表 1
- 用户使用合法账号进行认证上线
- 在服务器 1 和服务器 2 上分别查看用户记账信息；只有服务器 1 存在

相关命令

配置记账方法列表名

- 【命令格式】 **accounting** { *mlist-name* }
- 【参数说明】 方法列表名
- 【命令模式】 Web 认证模板配置模式
- 【使用指导】 应与 AAA 中网络记账方法列表名一致

配置举例

配置记账方法列表名

- 【配置方法】
 - 配置记账方法列表名为 mlist1

```
Ruijie(config.tmpl.eportalv2)#accounting mlist1
```

- 【检验方法】
 - 查看配置是否成功

```
Ruijie#show web-auth template
Webauth Template Settings:
-----
Name:      eportalv2
Url:       http://17.17.1.21:8080/eportal/index.jsp
Ip:        17.17.1.21
BindMode:  ip-only-mode
Type:      v2
Port:      50100
Acctmlist: mlist1
Authmlist: mlist1
```

5.4.5 配置portal服务器通信端口

配置效果

- 设备检测到用户下线等情况时，需要同时通告 portal 服务器用户下线；设备和 portal 服务器使用 portal 协议进行交互，协议使用约定的端口号进行报文侦听和收发。
- 当 portal 服务器侦听端口改变时，设备需要修改 portal 服务器通信端口才能进行交互。

注意事项

- 端口号配置需要和服务器实际使用端口相一致。

配置方法

- 可选配置
- 服务器不使用默认端口号时，使用本命令配置来保持一致。

检验方法

- 改变服务器侦听端口为 10000
- 通过本命令配置端口号为 10000
- 用户 web 认证上线
- 在设备端踢用户下线，刷新在线页面，提示用户下线

相关命令

配置 portal 服务器通信端口

- 【命令格式】 `port port-num`
- 【参数说明】 `port-num`：端口号
- 【命令模式】 web 认证模板配置模式
- 【使用指导】 -

配置举例

配置 portal 服务器通信端口

- 【配置方法】 ● 配置服务器通信端口为 10000

```
Ruijie(config.tmpl.t.eportalv2)#port 10000
```

【检验方法】 ● 查看配置是否成功

```
Ruijie#show web-auth template
Webauth Template Settings:
-----
Name:      eportalv2
Url:       http://17.17.1.21:8080/eportal/index.jsp
Ip:        17.17.1.21
BindMode:  ip-only-mode
Type:      v2
Port:      10000
Acctmlist:
Authmlist:
```

5.4.6 配置绑定模式

配置效果

- 用户成功上线时，需要将用户表项写到转发规则中，指定不同的绑定模式，可以改变转发规则的匹配方式，影响用户的上网规则。比如仅 IP 绑定时，只要符合该 IP 的报文都被放行，用户都能上网；而 IP+MAC 绑定时，只有同时符合该 IP 和 MAC 的用户能访问网络。

注意事项

- 在三层认证场景中，设备看到的 MAC 地址都是用户网关地址，MAC 地址都不准确，此时应该采用仅 IP 绑定模式。

配置方法

- 可选配置，缺省默认：IP+MAC 绑定。
- 依据设备能获得的用户准确信息决定选择哪种绑定模式，当用户 IP、MAC 均准确时，比如二层网络部署，优先选择 IP+MAC 绑定；否则优先选择仅 IP 绑定模式。

检验方法

- 改变绑定模式为仅 IP 模式
- 用户认证上线
- 修改用户 MAC，或者使用另一台同 IP，不同 MAC 的客户端访问网络
- 用户上网正常

相关命令

配置绑定模式

【命令格式】 **bindmode {ip-mac-mode | ip-only-mode}**

【参数说明】 **ip-mac-mode** : IP+MAC 同时绑定

ip-only-mode : 仅 IP 绑定

【命令模式】 web 认证模板配置模式

【使用指导】 -

配置举例

配置绑定模式

- 【配置方法】
- 配置绑定模式为仅 IP 模式

```
Ruijie(config.tmpl.eportalv2)#bindmode ip-only-mode
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie#show web-auth template
Webauth Template Settings:
-----
Name:      eportalv2
Url:       http://17.17.1.21:8080/eportal/index.jsp
Ip:        17.17.1.21
BindMode:  ip-only-mode
Type:      v2
Port:      10000
Acctmlist:
Authmlist:
```

5.4.7 设置重定向的HTTP端口

配置效果

- 当用户访问网络资源时（例如使用浏览器上网），此时用户会发出 HTTP 报文，接入/汇聚设备通过拦截来自用户的 HTTP 报文，来判断用户是否在访问网络资源。当设备检测到未认证的用户在访问网络资源时，将阻止用户访问网络资源，并向用户弹出认证页面。缺省情况下，网络设备通过拦截用户发出的端口号为 80 的 HTTP 报文，来检测用户是否在访问网络资源。
- 设置重定向的 HTTP 端口后，可以对用户发出的特定目的端口号的 HTTP 请求进行重定向。

注意事项

- 接入/汇聚设备上常用的管理协议端口（例如 22、23、53）以及系统内部保留的端口，不允许被设置为重定向端口。实际上，除了 80 端口外，HTTP 协议很少会使用小于 1000 的端口号。为了避免与知名 TCP 协议端口冲突，除非必要，尽量不要设置较小端口号的端口作为重定向端口。

配置方法

- 可选配置
- 在配置自动获取客户端时，如果要新增网络设备拦截用户发出的特定端口号的 HTTP 报文，可以进行该配置。

检验方法

- 配置拦截端口
- 未认证 pc，使用浏览器访问该端口的外网网络
- 用户访问请求被重定向到认证页面

相关命令

📄 设置重定向的 HTTP 端口

【命令格式】 **http redirect port** *port-num*

【参数说明】 *port-num*：端口号

【命令模式】 全局配置模式

【使用指导】 最大允许配置 10 个不同的目的端口号，端口号 80 也含在该总数量范围内。

配置举例

📄 设置重定向的 HTTP 端口

- 【配置方法】
- 配置 8080 端口为重定向的 HTTP 端口

```
Ruijie(config)#http redirect port 8080
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie(config)#show web-auth rdport  
  
Rd-Port:  
  
80 8080
```

5.4.8 设置web认证模块SYSLOG功能

配置效果

- 当用户上下线时，web 认证模块会通过 SYSLOG 将上下线的用户信息和事件呈现给管理员。缺省情况下，该 SYSLOG 信息被屏蔽。
- 设置 SYSLOG 限速功能后，可以按照一定的速率将该信息呈现出来。

注意事项

- 当认证上下线速率很高时，频繁的 SYSLOG 输出会影响设备性能，同时导致输出信息刷屏。

配置方法

- 可选配置
- 需要查看基本上下线 SYSLOG 信息时，可以配置 SYSLOG 限速功能。

检验方法

- 配置 SYSLOG 限速
- 用户按照一定速率上下线
- SYSLOG 按照限制要求打印输出

相关命令

📄 设置 SYSLOG 限速

【命令格式】 **web-auth logging enable num**

【参数说明】 *num* : SYSLOG 输出速率 (条/秒)

- 【命令模式】 全局配置模式
- 【使用指导】 0 为不限制速率，不输出受限的 SYSLOG。受限的 SYSLOG 不包括严重级 SYSLOG，及异常错误输出的 SYSLOG。

配置举例

设置 SYSLOG 限速

- 【配置方法】
- 配置不限制 SYSLOG 输出

```
Ruijie(config)#web-auth logging enable 0
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie(config)#show running-config
...
web-auth logging enable 0
...
```

5.4.9 设置未认证用户的最大HTTP会话数

配置效果

- 未认证的用户在访问网络资源时，用户 PC 会发出 HTTP 会话连接请求，HTTP 报文会被接入/汇聚设备拦截，并通过重定向要求用户进行 Web 认证。为了防止同一个未认证用户发起过多的 HTTP 连接请求，以节约网络设备的资源，需要在设备上限制未认证用户的最大 HTTP 会话数。
- 由于用户在认证时，会占用一个 HTTP 会话，而用户的其他应用程序也可能占用着 HTTP 会话，因而不建议设置未认证用户的最大 HTTP 会话数为 1。缺省情况下，全局每个未认证用户的最大 HTTP 会话数为 255，而每个端口下的未认证用户的 HTTP 会话总数最大为 300。

注意事项

- 如果一个用户在进行 Web 认证时，出现经常无法弹出认证页面的情况，则很可能是受到最大 HTTP 会话数的限制了。此时，应该建议用户暂时关闭一些可能会占用 HTTP 会话的应用程序，然后再进行 Web 认证。

配置方法

- 可选配置
- 要更改每个未认证用户的最大 HTTP 会话数及每个端口下的未认证用户的 HTTP 会话总数时，可进行该配置。

- 在配置自动获取 SU 客户端功能时，需要进行该配置。

检验方法

- 修改未认证用户最大会话数
- 未认证用户构造相同会话不间断对设备进行连接
- 未认证用户通过浏览器访问外网，访问请求不被重定向，设备提示用户会话数超过限制

相关命令

▾ 设置每个未认证用户的最大 HTTP 会话数

【命令格式】 **http redirect session-limit { session-num } [port { port-session-num }]**

【参数说明】 *session-num*：最大会话数。取值范围 1-255，默认为 255。

port-session-num：端口最大会话数。取值范围 1-65535，默认为 300。

【命令模式】 全局配置模式

【使用指导】 无

配置举例

▾ 设置每个未认证用户的最大 HTTP 会话数

- 【配置方法】 ● 设置每个未认证用户的最大 HTTP 会话数为 3

```
Ruijie(config)#http redirect session-limit 3
```

- 【检验方法】 ● 查看配置是否成功

```
Ruijie(config)#show web-auth parameter
HTTP redirection setting:
  session-limit: 3
  timeout:      3
Ruijie(config)#
```

5.4.10 设置维持重定向连接的超时时间

配置效果

- 设置维持重定向连接的超时时间。因为未认证的用户通过 HTTP 访问网络资源时，其 TCP 连接请求将被拦截，实际上是与接入/汇聚设备建立起 TCP 连接。在连接建立后，设备需要等待用户发出的 HTTP 的 GET/HEAD 报文，然后回复 HTTP

重定向报文后才能关闭连接。设置这个限制可以防止用户不发 GET/HEAD 报文，而又长时间占用 TCP 连接。缺省情况下，维持重定向连接的超时时间为 3 秒。

注意事项

- 无

配置方法

- 可选配置
- 要更改维持重定向连接的超时时间时，可进行该配置。

检验方法

- 修改超时时间配置
- 使用网络发包工具构造建立 tcp 连接
- 查看设备上该 tcp 连接状态，超过超时时间后连接被关闭

相关命令

设置维持重定向连接的超时时间

【命令格式】 **http redirect timeout { seconds }**

【参数说明】 *Seconds*：重定向连接超时时间，单位为秒。取值范围 1-10。默认为 3 秒。

【命令模式】 全局配置模式

【使用指导】 无

配置举例

设置维持重定向连接的超时时间

- 【配置方法】
- 设置维持重定向连接的超时时间 5

```
Ruijie(config)#http redirect timeout 5
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie(config)#show web-auth parameter
```

```
HTTP redirection setting:
```

```
session-limit: 255
```

```
timeout:      5
```

5.4.11 设置免认证网络资源范围

配置效果

- 在端口上启动 Web 认证或者 802.1x 认证后，未认证用户需先通过 Web 认证或者 802.1x 认证，才能访问网络资源。
- 使用此命令设置免认证的网络资源，可以允许未认证用户，也可以访问一些免认证的网络资源。
- 设置了免认证的网络资源，如果某网站属于免认证的网络资源，那么所有用户（包括未认证用户）都可以访问该网站。缺省情况下，没有设置免认证的网络资源，未认证用户不能访问网络资源。

注意事项

- 设置免认证的网络资源和设置无需认证用户共享资源，这两者各自都不能超过 50 个。此外，实际可用数量也会受其它安全功能占用表项的影响而减少。因此，如果需要设置的地址较多，请尽量使用网段的方式进行设置。
- http redirect direct-site 是配置免认证访问地址，http redirect 是配置 web 认证的服务器地址。从效果上看，用这两条命令配置的地址都是可以不用认证就直接访问的，但是实际用途是不一样的，因此实际使用时建议不要用 http redirect direct-site 来配置 web 认证服务器地址，否则会引起误解。

配置方法

- 可选配置
- 如果需要让未认证用户能够访问网络中的资源，使用本命令实现。

检验方法

- 配置免认证网络资源
- 未认证用户 PC 直接访问该资源，访问成功

相关命令

📄 设置免认证网络资源范围

【命令格式】 **http redirect direct-site** { *ip-address* [*ip-mask*] [**arp**] }

【参数说明】 *ip-address* : 免认证网络 ip 地址

ip-mask : 免认证网络掩码

【命令模式】 全局配置模式

【使用指导】 ARP 放行请优先采用直通 ARP 命令

配置举例

📄 设置免认证网络资源范围

- 【配置方法】
- 设置免认证的网络资源范围 192.168.0.0/16

```
Ruijie(config)#http redirect direct-site 192.168.0.0 255.255.0.0
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie(config)#show web-auth direct-site
```

```
Direct sites:
```

Address	Mask	ARP Binding
---------	------	-------------

192.168.0.0	255.255.0.0	Off
-------------	-------------	-----

```
Ruijie(config)#
```

5.4.12 设置直通ARP资源范围

配置效果

- 开启 ARP CHECK 或类似功能时，用户的 ARP 学习受控，导致用户无法学到网关及其他设备的 ARP，影响用户使用。此时可以通过设置直通 ARP 资源来对指定地址的 ARP 学习报文进行放行

注意事项

- 对于开启 ARP Check 情况，需要将二层接入设备下联 PC 的网关设置为直通 ARP 资源。需要注意以下问题：
对同一地址/网段同时设置直通网站和直通 ARP 时，命令会自动进行合并；如果直通网站的配置没有指定 ARP 选项，合并后该选项会自动增加。

配置方法

- 可选配置
- 如果设备启用了 ARP CHECK 功能，那么需要对免认证的网络资源范围和网关配置 ARP 直通。

检验方法

- 配置直通 ARP 资源
- 未认证用户 PC 上清空 ARP 缓存(Windows 执行命令 arp -d)
- 未认证用户 PC 执行 ping 直通 ARP 资源
- 未认证用户 PC 查看 ARP 缓存(Windows 执行命令 arp -a)，学习到直通 ARP 资源的 ARP 地址

相关命令

设置直通 ARP 资源范围

- 【命令格式】 **http redirect direct-arp** { *ip-address* [*ip-mask*] }
- 【参数说明】 *ip-address* : 免认证网络 ip 地址
ip-mask : 免认证网络掩码
- 【命令模式】 全局配置模式
- 【使用指导】 -

配置举例

设置直通 ARP 资源范围

- 【配置方法】
 - 设置直通 ARP 资源范围 192.168.0.0/16

```
Ruijie(config)#http redirect direct-arp 192.168.0.0 255.255.0.0
```

- 【检验方法】
 - 查看配置是否成功

```
Ruijie(config)#show web-auth direct-arp

Direct arps:

  Address          Mask
  -----
192.168.0.0      255.255.0.0

Ruijie(config)#
```

5.4.13 设置无需认证用户IP范围

配置效果

- 如果用户属于无需认证用户 IP 范围，那么该用户不需要通过 Web 认证，也能访问所有可达的网络资源。缺省时，没有设置无需认证用户，所有用户都必须先通过 Web 认证，才能访问网络资源。

注意事项

无

配置方法

- 可选配置
- 要设置无需认证用户时，可进行该配置。

检验方法

- 配置用户为无需认证用户
- 用户直接访问网络，访问成功

相关命令

设置无需认证用户 IP 范围

- 【命令格式】 **web-auth direct-host** { *ip-address* [*ip-mask*] [*arp*] } [**port** *interface-name*]
- 【参数说明】 *ip-address* : 免认证用户 IP 地址
ip-mask : 免认证用户掩码
interface-name : 接口名
- 【命令模式】 全局配置模式
- 【使用指导】 *arp* 字段用来设置放行 arp 报文，当开启 ARP-CHECK 功能时需要配置
配置 *port* 字段后，用户仅在所配置接口下时免认证功能生效，在其余接口上不生效

配置举例

设置无需认证用户 IP 范围

- 【配置方法】 ● 设置无需认证用户 IP 范围

```
Ruijie (config)# web-auth direct-host 192.168.197.64
```

- 【检验方法】 ● 查看配置是否成功

```
Ruijie(config)#show web-auth direct-host
Direct hosts:
  Address          Mask             Port             ARP Binding
  -----
  192.168.197.64  255.255.255.255             Off
Ruijie(config)#
```

5.4.14 配置Portal通信端口

配置效果

- 配置后设备与 Portal 服务器通信的源端口为所配置端口。

注意事项

- 端口只能配置一个。

配置方法

- 配置指定端口为 portal 通信口。

检验方法

- 开启 web 受控后，认证时服务器上抓包，认证报文源 IP 为指定端口的 IP。

相关命令

配置通信端口

【命令格式】 ip portal source-interface interface-type interface-num

【参数说明】

【命令模式】 全局模式配置模式。

【使用指导】

配置举例

配置 portal 通信端口

- 【配置方法】
- 使用聚合口当作 portal 通信口地址。

```
Ruijie(config)# ip portal source-interface Aggregateport 1
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie(config)#show running-config  
ip portal source-interface Aggregateport 1
```

5.4.15 设置在线用户信息的更新时间间隔

配置效果

- 接入/汇聚设备维护着在线用户信息，设备需要定时地更新在线用户信息，包括在线时间等，以监控在线用户使用网络资源的情况，比如：用户的在线时间大于或等于在线时限，该用户会被停止使用网络。

注意事项

- 用户更新时间必须配置为 60 的倍数，否则实际配置自动向上取最近的 60 倍数来生效。

配置方法

- 可选配置
- 要设置无需认证用户时，可进行该配置。

检验方法

- 配置用户信息更新时间
- 超过更新时间间隔后，查看在线用户信息

相关命令

📄 设置在线用户信息更新时间间隔

【命令格式】 **web-auth update-interval** { seconds }

【参数说明】 Seconds：信息更新时间间隔，单位为秒。取值范围 30-3600。默认为 180 秒。

【命令模式】 全局配置模式

【使用指导】 如果要恢复更新时间间隔为默认值，在全局配置模式下，使用 **no web-auth update-interval**

配置举例

📄 设置已认证用户信息的更新时间间隔为 60 秒

- 【配置方法】
- 设置已认证用户信息的更新时间间隔为 60 秒

```
Ruijie (config)# web-auth update-interval 60
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie(config)#show run | include web-auth update-interval
```

```
web-auth update-interval 60
```

5.4.16 配置DHCP地址核查

配置效果

- 只有通过 DHCP 分配地址的终端才允许进行认证。

注意事项

- 需要配置 DHCP SNOOPING 功能。
- 仅支持 IPV4。
- 仅支持二代 Web 认证和内置 Portal Web 认证。
- 网络部署时明确用户使用 DHCP 获得 IP 地址，不存在静态 IP 地址混用的情况，否则会影响静态 IP 地址的用户。
- 如果有少数用户需要用静态 IP 地址，可以通过配置直通地址放行，对这些用户不认证。

配置方法

- 可选配置。
- 配置 DHCP SNOOPING。
- 配置地址核查。

检验方法

- 配置 DHCP 地址核查功能。
- 终端配置静态地址，该地址未经 DHCP 服务器分配。
- 终端连接到网络，无法认证。

相关命令

▾ 设置 DHCP 地址核查

【命令格式】 **web-auth dhcp-check**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 网络部署时指定用户需要使用 DHCP 获取 IP 地址上网，配置此功能有助于屏蔽一些私设 IP 地址的用户认证上网。

配置举例

设置 DHCP 地址核查

- 【配置方法】
- 配置 DHCP 地址核查

```
Ruijie(config)# web-auth dhcp-check
```

- 【检验方法】
- 查看配置是否成功

```
Ruijie(config)#show running-config
```

```
...
```

```
web-auth dhcp-check
```

```
...
```

5.4.17 升级兼容性说明

配置效果

- 部分配置命令在 11.X 系列软件上做了优化，格式上有所变化，具体参考后面的说明。
- 10.X 系列软件可以实现平滑升级，不会出现功能丢失的情况，但是升级后部分旧命令的显示会转化为新命令。
- 在 11.X 系列软件上对这部分旧命令执行 no 操作会提示不支持，需要用新的格式执行 no 操作。

注意事项

无

配置方法

- 部分格式变化的命令建议使用新格式配置。

检验方法

- 10.X 软件版本升级到 11.X 软件版本不会出现功能丢失，同时显示以及保存的格式采用新命令。
- 新格式命令的功能效果和旧命令一致。

相关命令

本小节

配置一代 web 认证的 portal 服务器资源地址

- 【命令格式】 `http redirect homepage url`

- 【参数说明】 `url`：一代 web 认证的 eportal 服务器资源地址
- 【命令模式】 全局配置模式
- 【使用指导】 11.X 版本首先会将该命令转化为一个 eportalv1 模板，然后用模板模式下的 `url` 命令配置和显示服务器地址，具体参考 1.4.1 章节。

▾ 配置一代 web 认证的 portal 服务器资源地址

- 【命令格式】 `http redirect homepage url`
- 【参数说明】 `url`：一代 web 认证的 eportal 服务器资源地址
- 【命令模式】 全局配置模式
- 【使用指导】 11.X 版本首先会将该命令转化为一个 eportalv1 模板，然后用模板模式下的 `url` 命令配置和显示服务器地址，具体参考 1.4.1 章节。

▾ 配置 portal-server

- 【命令格式】 `portal-server [eportal1 | eportalv2]`
- 【参数说明】 `eportav1`：一代 web 认证的 portal 信息
`eportav2`：二代 web 认证的 portal 信息
- 【命令模式】 全局配置模式
- 【使用指导】 11.X 版本首先会将该命令转化为一个 eportalv1 或者 eportalv2 模板，然后用对应的信息填充，portal-server 主要参数包括服务器 ip 地址和 url 地址，会被模板中的 ip 命令和 url 命令替代。

▾ 配置接口 web 认证受控

- 【命令格式】 `web-auth port-control`
- 【参数说明】 无
- 【命令模式】 接口配置模式
- 【使用指导】 11.X 版本会将该命令转化 `web-auth enable <type>` 这里的 type 是指一代或者二代，默认是一代。

▾ 配置仅 ip 绑定模式

- 【命令格式】 `web-auth port-control ip-only-mode`
- 【参数说明】 无
- 【命令模式】 接口配置模式
- 【使用指导】 11.X 版本首先会将该命令转化为一个 eportalv1 模板或者 eportalv2 模板，取决于实际配置。然后用模板模式下的 `bindmode` 命令配置和显示服务器绑定模式，具体参考 1.4.1 和 1.4.2 章节。

▾ 显示一代 web 认证配置信息

- 【命令格式】 `show http redirect`
- 【参数说明】 无
- 【命令模式】 特权模式
- 【使用指导】 11.X 版本该旧命令不可用，改为 `show web-auth template`。

▾ 显示端口受控信息

- 【命令格式】 `show web-auth port-control`

- 【参数说明】 无
- 【命令模式】 特权模式
- 【使用指导】 11.X 版本该旧命令不可用，改为 show web-auth control。

配置举例

配置一代 web 认证

- 【配置方法】
- 产品运行 10.X 版本并且已经配置了一代 web 认证服务器 ip

```
Ruijie(config)# http redirect 192.168.197.64
```

- 产品升级到 11.X 版本
- 【检验方法】
- 升级后 show running-config，配置已更新为新命令格式

```
Ruijie#sh running-config
web-auth template eportalv1
  ip 192.168.197.64
```

```
!
```

5.5 监视与维护

清除各类信息

作用	命令
强制用户下线	clear web-auth user { all ip <i>ip-address</i> mac <i>mac-address</i> name <i>name-string</i> session-id <i>num</i> }

查看运行情况

作用	命令
查看 web 认证基本参数配置。	show web-auth parameter
查看 web 认证模板配置信息	show web-auth template
查看免 Web 认证的用户范围。	show web-auth direct-host
查看直通地址范围。	show web-auth direct-site
查看直通 ARP 范围。	show web-auth direct-arp
查看 TCP 拦截端口。	show web-auth rdport
接口上的认证配置信息。	show web-auth control
查看所有用户或是指定用户的在线信息。	show web-auth user { all ip <i>ip-address</i> mac <i>mac-address</i> name <i>name-string</i> session-id <i>num</i> }

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
web 认证 debug	debug web-auth all

6 SCC

6.1 概述

SCC (Security Control Center, 安全控制中心)为各种接入控制和网络安全业务提供了公共的配置方法和策略整合服务,从而使得各种接入控制业务以及网络安全业务能够在一台设备上共存,实现多元化的接入安全控制需求,以满足不同的接入场景需要。

典型的接入控制业务如 dot1x、web 认证、arp check、ip source guard 等;网络安全业务如 ACL、NFPP、防网关 ARP 欺骗等。当设备上同时开启两个或两个以上的上述接入控制业务或网络安全业务时,或者同时开启接入控制业务和网络安全业务时,SCC 通过相关的策略整合负责协调共存关系。

i 有关接入控制和网络安全业务相关的说明请参考相应的配置指南,下文仅介绍 SCC 的相关内容。

协议规范

无

6.2 典型应用

典型应用	场景描述
高校大二层校园网访问控制应用	在高校校园网中学生可通过 dot1x 客户端认证上网或通过 web 认证来上网,同时要防止相互间 ARP 欺骗。另外,允许某些部门(比如校长办公室)的终端设备无需认证就能上网。

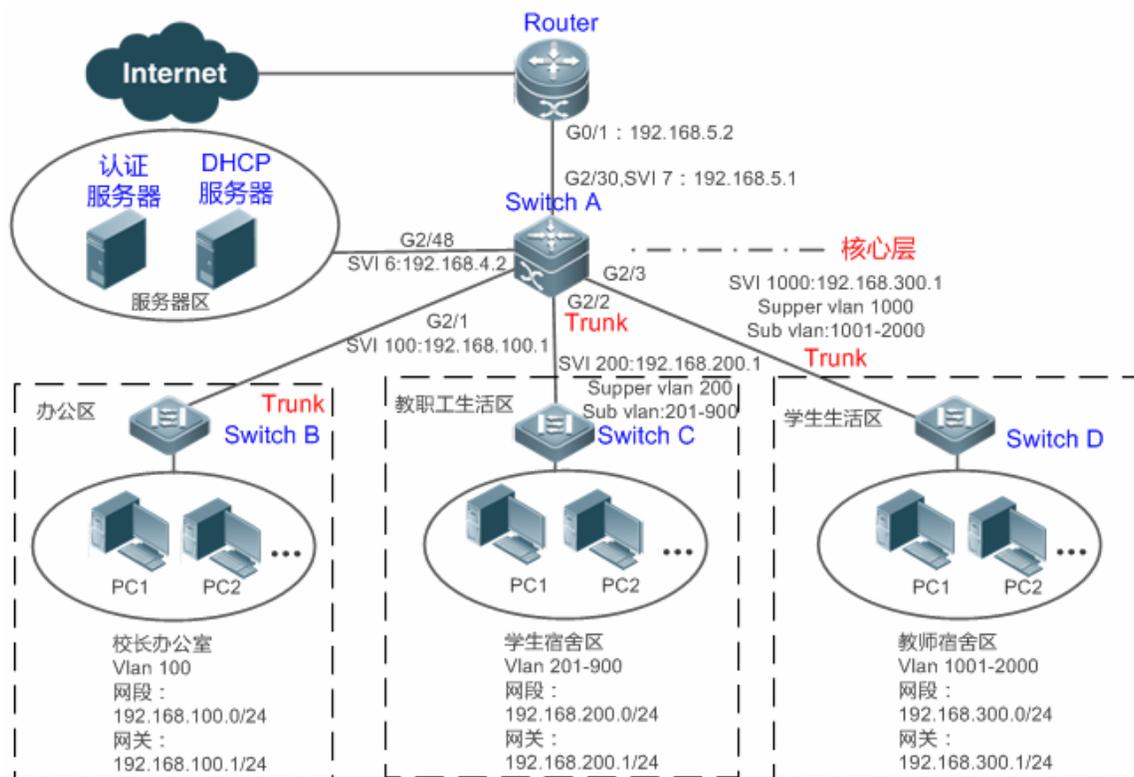
6.2.1 高校大二层校园网访问控制应用

应用场景

在高校校园网中,学生上网前一般都需要前进行 1x 认证或 web 认证,从而方便计费,以保障高校的利益:

- 学生可以通过 dot1x 客户端认证上网,也可以通过 web 认证上网。
- 防止学生相互间进行 ARP 欺骗,以保证网络的稳定性。
- 允许某些部门(如校长办公室)的终端无需认证就能上网。

图 6-1



【注释】 传统的高校校园网网络是分层设计的，有接入层、汇聚层和核心层，用户接入控制在接入层上完成；而在高校大二层校园网中，用户的接入控制是由核心设备来承担的，核心设备以下都是二层设备，中间不再有汇聚。核心设备与用户接入交换设备（如上图的 switch B、switch C 以及 switch D）之间都是 TRUNK 口。

接入层设备 B、C、D：连接各部门的 PC，各个接入端口配置成 Access 口，VLAN 与核设备对应下联端口上配置的 SUB VLAN 相对应，这样每个接入用户处于不同的 VLAN 中，防止相互间进行 arp 欺骗。

核心层设备 A 连接各种服务器 如认证服务器、DHCP 服务器等。并在下联端口上配置 Super VLAN 和 Sub VLAN。一个 Super VLAN 对应多个 Sub vlan，每个 Sub VLAN 代表一个接入用户。

功能部署

- 核心设备上通过 vlan + 端口号来区分不同的接入用户，每个接入用户（当然也可以是一组用户）一个 vlan。接入层设备下联用户的端口配置成 Access 口，并按规划为每个用户配置一个用户 vlan，核心设备上不转发 ARP 请求报文，只有被请求用户已认证才作应答 以此来达到防止用户间的 ARP 欺骗问题。核心设备 switch A 上面将用户 VLAN 作为 Sub VLAN，并配置 Super VLAN 以及将 Super VLAN 对应的 SVI 配置成用户网关。
- 通过在核心设备（本例为设备 A）下联教职工生活区和学生生活区的端口上同时开启 dot1x 认证和 web 认证功能来达到由用户自由选择使用哪种认证方式的目的。
- 对于特殊部门（本例为校长办公室）可以划到单独特定的一个 VLAN 中，通过配置这个 VLAN 为免认证 VLAN 的方式来达到不需要通过认证即可上网的目的。

6.3 功能详解

基本概念

免认证 VLAN

为简化网络管理，可以将一些特殊的部门划入免认证 VLAN 范围，实现这些部门无需认证即可访问网络资源的需求。比如，校园网中可以将校长办公室列为免认证 VLAN 的范围，从而让这些特殊的部门用户上网无需认证。

IPv4 用户容量

为了保护已上线用户的上网稳定性，同时也为了让设备能够更加稳定地运行，可以对 IPv4 接入用户数量进行限制。

- ❗ 默认情况下不会对 IPv4 接入用户数量进行限制，可以让大量用户认证上线，直到上线用户数达到设备的硬件最大容量。
- ℹ IPv4 接入用户包括 dot1x 认证产生的 IP 用户(比如 IP 授权用户)、WEB 认证用户、用户手工绑定的 IP 用户(包括 IP source guard、arp check 等)。

在线认证用户迁移

在线用户迁移指的是一个已在线的认证用户可以在不同的物理位置自由接入重新认证上线。但在校园网中，为了便于管理常常会要求学生只能在指定的位置进行认证上网，在其他接入端口上无法进行认证上网，这样的用户是不可迁移的；另一种情况，有些用户要求可以移动办公，在不同的接入点都可以认证上线，这样的用户就是可迁移的。

用户在线检测

对于计费用户来说，用户认证上线之后就会开始计费，用户离开时需要主动下线才能真正结束计费过程，但有可能用户上网结束离开时忘记了主动下线或者因终端原因无法主动下线等原因，继续产生上网费用从而导致用户的经济损失。为了更加精确地判断用户是否真的在上网，可以预设在一个时间段内用户流量低于某个值或者在一个时间段内用户无流量时就认为用户没有使用网络，直接将该用户进行下线操作。

功能特性

功能特性	作用
免认证VLAN	可以将指定 VLAN 内的用户配置成免认证用户。
IPv4 用户容量	可以对指定接口上的 IPv4 用户容量进行限制以保证容量内的用户上网稳定。
认证用户迁移	可以指定静态认证用户是否可以迁移
用户在线检测	可以指定是否对在线用户进行流量检测，在一段时间内流量低于某个值时设备主动将用户下线。

6.3.1 免认证VLAN

免认证 VLAN 主要用于有特定需求的部门，通过将各部门划入免认证 VLAN 中，从而实现无需通过 dot1x 认证或 web 认证就可以上网的目的。

工作原理

开启免认证 VLAN 的设备在检测到报文来自免认证 VLAN 列表时，直接跳过接入控制的检测，从而在实现免认证 VLAN 中的用户上网无需认证的要求。可以将免认证 VLAN 功能看作是安全通道的一种应用。

- ✔ 仅在交换机设备上支持免认证 VLAN 功能。
- ✔ 最多支持配置 100 个免认证 VLAN。
- ✔ 免认证 VLAN 会占硬件表项。在没有开启认证等接入控制功能的情况下，配置免认证 VLAN 与没有配置免认证 VLAN 效果是一样的。建议只在开启相关接入控制功能的时候，对有不需认证就能上网需求的特殊用户才进行免认证 VLAN 配置。
- ❗ 免认证 VLAN 中的报文虽然无需经过接入控制的检测，但要经过安全 ACL 的检查。如果安全 ACL 不允许该 VLAN 中的用户报文通过，用户还是不能上网。

6.3.2 IPv4 用户容量

为了让设备能够更加稳定地运行，避免非法用户的暴力冲击，可以对设备某个接口上的 IPv4 可接入用户总数进行限制。

工作原理

如果对 IPv4 的可接入用户数进行了限制，超过限制的新用户将无法接入网络，也就无法正常上网。

- ✔ 仅在交换机设备上支持 IPv4 接入用户数限制功能。
- ✔ 默认情况下，设备不对 IPv4 接入用户数进行限制，可接入用户数取决于设备的硬件容量。
- ❗ 这里的 IPv4 接入用户包括 dot1x 产生的 IPv4 授权用户、WEB 认证产生的 IPv4 用户，以及各种绑定功能产生的 IPv4 用户表项。由于 IPv4 可接入用户数限制是在接口下配置的，这个限制的范围包括在本接口上生成的 IPv4 用户，同时也包括全局生成的 IPv4 用户。比如，配置接口 Gi 0/1 的 IPv4 接入用户最大数量为 2，使用命令在接口上绑定一个 IPv4 用户，再使用命令绑定一个全局的 IPv4 用户，实际上该接口上的接入用户数已经达到最大 2，这个时候再想在该接口上绑定一个 IPv4 用户或者再想绑定一个全局的 IPv4 用户将会失败。

6.3.3 认证用户迁移

在实际网络中，接入用户并不一定都是固定在一地方上网或办公，用户在一个地方认证上线，临时要到另外一个部门或办公室办公，不主动下线，直接拔掉网线，带着移动终端到新的办公场所想上线访问网络，此时就涉及到在线认证用户的可迁移和不可迁移问题。如果不配置认证用户可迁移，一个用户在一个地方认证上线之后，没下线，跑到另外一个地方是无法认证上线的。

工作原理

认证用户迁移时，设备的 dot1x 认证或 web 认证组件检测到该用户对应的 MAC 的端口号或所在 VLAN 发生变化，会先将用户下线，用户需要用户重新认证上线。

- ✔ 仅在交换机或无线设备上支持认证用户迁移。且不支持跨交换机迁移，比如两台核心交换机上都开启了认证和迁移功能，用户从其中一台核心交换机认证上线，迁移到另一台核心交换机下，这个时候是无法迁移的。
- ❗ 认证用户迁移功能需要检查用户的 MAC 地址，对仅 IP 用户不生效。
- ❗ 是否允许认证用户迁移针对的是在一个地方上线后，在没下线情况下跑到另一个地方上线这种情况。如果在一个地方上线后又下线，或者还没上线就跑到另一个地方上线，这种情况不在认证用户迁移配置的控制范围内。
- ❗ 处理迁移时，由于是根据用户的 MAC 地址对应的 VLAN ID 或端口号是否与迁移前一致来判断用户是否发生了迁移，如果一致，表示用户没有发生迁移，否则表示用户发生了迁移。基于上述原理，如果网络中有其他用户假冒已在线用户的 MAC 地址上线，这个时候不加额外判断就会错误地把正常用户下线。为了防止这种问题的发生，在迁移时 WEB 认证和 dot1x 会做一个用户是否实际发生迁移的校验工作。对于通过 WEB 或开启了 IP 授权的 dot1x 认证上线的用户，在探测到相同 MAC 地址在其他 VLAN 或其他端口上线时，就向原位置触发 ARP 请求，如果在指定的时间内没有收到回应，就说明用户位置确实发生了变化，就允许迁移；如果在指定的时间内收到了回应，就说明该用户实际没有迁移，网络中可能存在防冒用户，就不进行迁移。ARP 请求探测默认为每秒 1 次，共 5 次，也就是 5s 后才能判断出一个用户是否发生了迁移。超时相关参数包括探测间隔和次数可通过 `arp retry times times` 和 `arp retry interval interval` 这两条 CLI 命令来调整，具体配置方法请参考《ARP-SCG.doc》。需要注意的是，对于 dot1x，只有配置了 IP 授权功能后才有办法进行这种迁移校验。

6.3.4 用户在线检测

用户在上好网之后，有可能会忘了点下线或者由于终端缘故无法主动下线，这个时候会造成持续计费而招致经济损失。在这种情况下，为了保障上网用户的利益，设备提供了判断用户是否在线即用户在线检测功能，由设备来判断用户是否真的在线，如果设备认为用户不在线，主动将该用户进行下线。

工作原理

在设备上预设一个指定的检测周期，在这个周期内如果用户流量低于某个值时就认为此时用户没有使用网络，从而直接将该用户进行下线操作。

- ✔ 仅在交换机和无线设备上支持用户在线检测功能。
- ✔ 用户在线检测功能仅针对通过 dot1x 认证或 web 认证上线的用户。
- ✔ 交换机上目前只支持无流量检测，不支持低流量检测。
- ❗ 目前交换机由于硬件芯片的限制，用户无流量下线的实际时间与 MAC 地址老化时间有关，如流量检测间隔 `interval` 被配置为 `m` 分钟，MAC 地址老化时间被配置为 `n` 分钟，则一个已认证用户不主动下线就离开网络到检测到无流量被下线之间的间隔大概在 `[m, m+n]` 分钟之间，也就是在这种配置下，一个在线用户如果没有发生任何上网流量，大概在 `[m, m+n]` 分钟后就会被下线。

6.4 配置详解

配置项	配置建议 & 相关命令
-----	-------------

配置免认证VLAN	 可选配置。用于指定哪些 VLAN 内的用户无需要认证就可上网	
	[no] direct-vlan	配置免认证 VLAN
配置IPv4 用户容量	 可选配置。用于限制某个接口上可接入的用户数	
	[no] nac-author-user maximum	配置接口下可接入的 IPv4 用户数
配置认证用户迁移	 可选配置。用于指定静态 MAC 地址的在线认证用户是否可迁移	
	[no] station-move permit	配置认证用户是否可迁移
配置用户在线检测	 可选配置。用于指定是否开启用户在线检测功能	
	offline-detect interval threshold	配置用户在线检测参数
	no offline-detect	关闭用户在线检测功能
	default offline-detect	恢复成缺省的用户在线检测方式

6.4.1 配置免认证VLAN

配置效果

通过配置免认证 VLAN，指定 VLAN 内的用户无需要通过 dot1x 认证或 web 认证就可以上网。

注意事项

免认证 VLAN 只是不作接入认证相关的检测，但必须经过安全 ACL 的检查，如果在安全 ACL 中配置了不允许指定用户或指定 VLAN 通过，用户还是无法上网。所以在配置 ACL 时需要注意，如果想让免认证 VLAN 真正免认证就能上网，不能将指定 VLAN 或指定 VLAN 内的用户拦截掉。

配置方法

配置免认证 VLAN

- 可选配置。想让指定 VLAN 下的所有用户免 dot1x 认证或 web 认证，首先必须配置免认证 VLAN。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。

【命令格式】 **[no] direct-vlan vlanlist**

【参数说明】 **no**：该选项若被配置，表示删除免认证 VLAN 配置。
vlanlist：表示所配置或删除的免认证 VLAN 列表。

【缺省配置】 无任何免认证 VLAN 配置

【命令模式】 全局模式

【使用指导】 此命令可以用来配置或删除免认证 VLAN。

检验方法

- 可以通过以下方法免认证 VLAN 的配置效果：
- 在下联用户终端的端口上开启dot1x受控。将下联用户终端的端口划入指定VLAN，并将该VLAN配置成免认证VLAN，然后打开浏览器，输入有效的外网网址（比如 www.baidu.com），可以上网则表示免认证VLAN生效，不能上网则表示免认证VLAN没生效。
- 使用 **show direct-vlan** 命令检查设备上的免认证 VLAN 配置。

【命令格式】 **show direct-vlan**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 全局模式

【使用展示】
Ruijie#show direct-vlan
direct-vlan 100

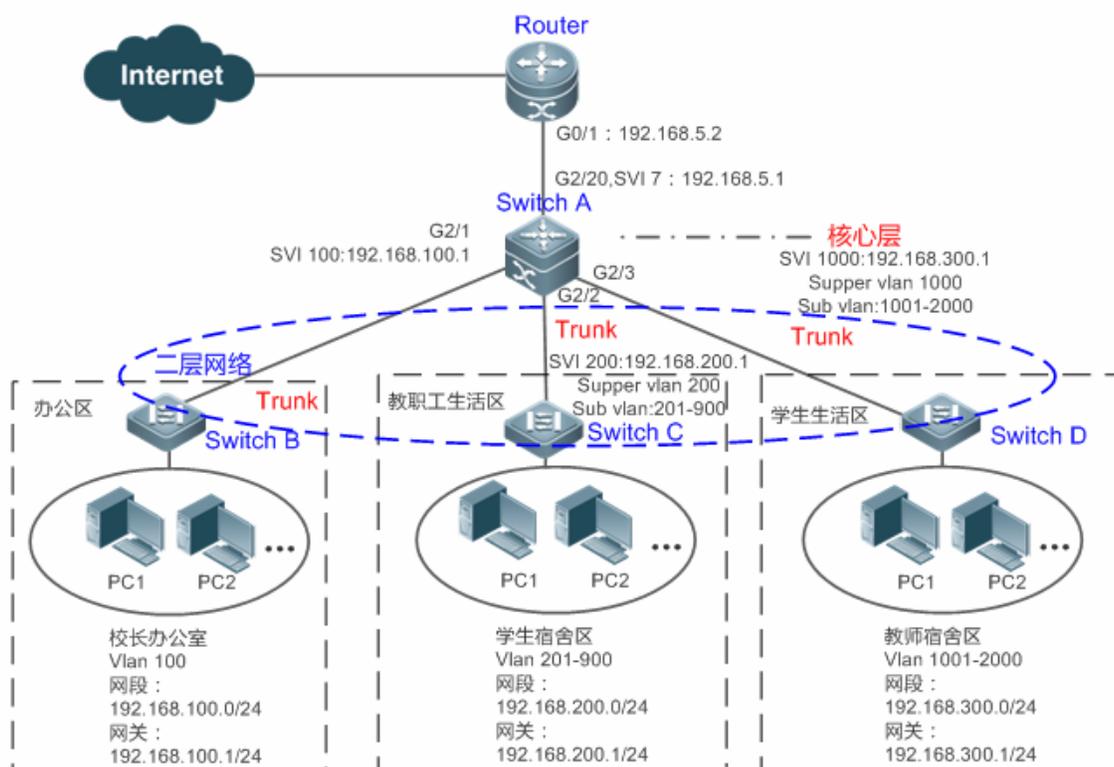
配置举例

i 以下配置举例，仅介绍与 SCC 相关的配置。

通过免认证 VLAN，实现特殊用户免认证上网

【网络环境】

图 6-2



- 【配置方法】
- 将核心网关设备 SwitchA 上的 GI 2/1 口配置成 TRUNK 口，并开启 dot1x 受控
 - 在核心网关设备 SwitchA 上将校长办公室所有 VLAN 100 配置成免认证 VLAN

Switch A

```
SwitchA(config)#vlan 100
SwitchA(config-vlan)#exit
SwitchA(config)#direct-vlan 100
SwitchA(config)#int GigabitEthernet 0/1
SwitchA(config-if-GigabitEthernet 0/1)#switchport mode trunk
SwitchA(config-if-GigabitEthernet 0/1)#dot1x port-control auto
*Oct 17 16:06:45: %DOT1X-6-ENABLE_DOT1X: Able to receive EAPOL packet and DOT1X authentication enabled.
```

- 【检验方法】
- 从校长办公室内的任意一台 PC 机上打开浏览器，输入有效的外网网址，确认可以打开网页。
 - 使用 show direct-vlan 命令可验证免认证 VLAN 是否生效。

Switch A

```
SwitchA(config)#show direct-vlan
direct-vlan 100
```

6.4.2 配置IPv4 用户容量

配置效果

通过配置 IPv4 用户容量，可以限制一个接入端口上的可接入用户数。

注意事项

无

配置方法

▾ 配置 IPv4 用户容量

- 可选配置。如果要限制一个接入端口上的最大可接入用户数，就必须配置 IPv4 的用户容量。默认没限制。如果指定接口上配置了用户容量限制，则当该接口上认证用户数量达到上限时，新用户无法认证上线，必须等到有用户下线后才可以认证上线。
- 需要配置在接入设备上，注意：接入设备有可能是网络边缘的接入交换机设备，也可能是核心的网关设备。
-

【命令格式】 **nac-author-user maximum** *max-user-num*
no nac-author-user maximum

- 【参数说明】 **no**：该选项若被配置，表示取消端口下的 IPv4 接入用户容量限制。
max-user-num：表示所配置端口下 IPv4 接入用户容量限制值，取值范围[1, 1024]。
- 【缺省配置】 不对 IPv4 接入用户数量进行限制
- 【命令模式】 接口模式
- 【使用指导】 此命令可以用来限制指定接入端口的 IPv4 接入用户数。

检验方法

- 可以通过以下方法检验端口上的 IPv4 用户容量的配置效果：
- 如果是 dot1x 认证，可以在该端口上下联的 1x 客户端认证上线达到指定的用户容量，这时再想认证上线一个用户，不能上线成功。
- 如果是 web 认证，可以在该端口上下联的客户端通过 web 认证上线达到指定的用户容量，这时再想认证上线一个用户，不能上线成功。
- 使用 **show nac-author-user [interface interface-name]**命令检查设备上的 IPv4 用户容量配置。

【命令格式】 **show nac-author-user [interface interface-name]**

【参数说明】 *interface-name*: 接口名称

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 全局模式

【使用展示】 Ruijie#show nac-author-user interface GigabitEthernet 0/1

```
Port      Cur_num  Max_num
-----  -
Gi0/1    0        4
```

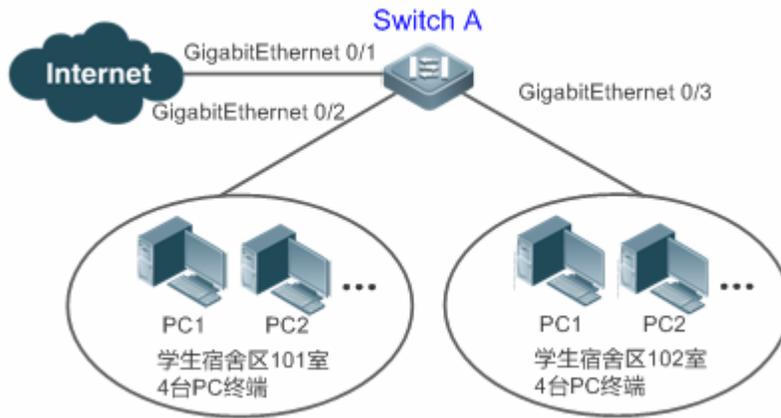
配置举例

 以下配置举例，仅介绍与 SCC 相关的配置。

👉 通过限制端口上的 IPv4 用户数，防止过多的上网终端设备冲击网络。

【网络环境】

图 6-3



- 【配置方法】
- 这里假设接入设备 Switch A 的 dot1x 认证环境都已经配置好了，dot1x 受控在 Gi 0/2 端口上开启。
 - 配置 Gi 0/2 端口下的 IPv4 接入用户最大容量为 4。

Switch A

```
SwitchA(config)#int GigabitEthernet 0/2
SwitchA(config-if-GigabitEthernet 0/2)#nac-author-user maximum 4
```

- 【检验方法】
- 将宿舍内的 4 台 PC 全部进行 dot1x 认证上线。然后额外拿一台终端接入网络，企图进行 dot1x 认证，确定无法成功认证上线。
 - 使用 show nac-author-user 命令可以查看配置是否生效

Switch A

```
SwitchA(config)#show nac-author-user
Port      Cur_num  Max_num
-----
Gi0/1     0        4
```

6.4.3 配置认证用户迁移

配置效果

默认情况下，一个认证用户在一个物理位置（接入端口加 VLAN 表示物理位置）通过 dot1x 或 web 认证上线后，未下线的情况下很快地在另一个物理位置接入试图是无法再通过 dot1x 或 web 认证上线的；当配置了认证用户可迁移后，上述这种情况下就可以。

注意事项

- 未配置认证用户可迁移时，虽然已在线用户从一个物理位置快速地切换到另一个物理位置无法认证上线，但如果用户下线后进行物理位置切换或者切换过程中用户下线了，比如用户在线检测功能将用户踢下线，这时即使不配置认证用户可迁移在另一个物理位置也还是可以正常认证上线成功。
- 在线认证用户在迁移到新物理位置时，需要重新进行 dot1x 或 web 认证才能正常上线。

配置方法

配置认证用户迁移

- 可选配置。如果要允许用户在不同的物理位置认证上线，必须开启认证用户可迁移。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。

【命令格式】 **[no] station-move permit**

【参数说明】 **no station-move permit**：该选项若被配置，表示不允许认证用户迁移。

station-move permit：该选项若被配置，允许认证用户迁移。

【缺省配置】 不允许认证用户迁移。也就是用户在网络的一个物理位置认证上线后，未下线的情况下，到另一个物理位置上想重新上线，将会失败。

【命令模式】 全局模式

【使用指导】 此命令可以用来配置是否允许认证用户在线迁移。

检验方法

- 可以通过以下方法检验认证用户可迁移配置效果：
- PC 机通过 dot1x SU 客户端在设备的一个 dot1x 受控端口上认证上线，然后不主动下线，将 PC 机移到设备的另一个受控端口下，重新进行 dot1x 认证上线，看是否能上线成功。

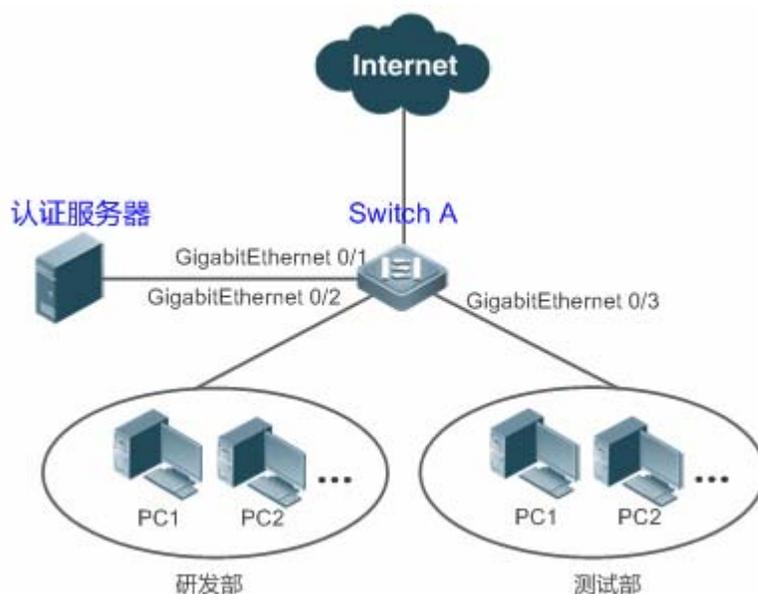
配置举例

i 以下配置举例，仅介绍与 SCC 相关的配置。

通过配置在线认证用户可迁移，允许用户在在线的情况下在不同的认证端口下重新认证上线

【网络环境】

图 6-4



- 【配置方法】
- 在接入端口 Gi 0/2 和 Gi 0/3 上开启 dot1x 受控，并配置认证所需参数，基于 MAC 认证
 - 配置在线认证用户可迁移
- Switch A**
- ```
swl(config)#station-move permit
```
- 【检验方法】
- 在研发部中的一台便携式电脑上通过 dot1x SU 客户端认证上线，直接拔掉网线，接入测试部所在局域网中，重新使用 dot1x SU 认证，确认可以成功上线。
- Switch A**
- ```
swl(config)#show running-config | include station
station-move permit
```

6.4.4 配置用户在线检测

配置效果

当配置了认证用户在线检测功能后，在指定的周期内如果流量低于一定的门限，设备会自动将用户下线，以免造成持续计费而导致用户的经济损失。

注意事项

配置如果配置无流量下线，需要注意的是，终端一般来说都会默认运行 360 安全卫士等软件，这些软件会时不时地往外发送报文，此时，只有终端关机的情况下设备才会将用户下线。

配置方法

配置用户在线检测

- 可选配置。默认为 8 小时内无流量就将用户下线。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。只对被配置的设备上有效，不会影响网络中的其他设备。

i 流量门限参数 `threshold` 如果配置成 0，则表示进行无流量检测。

【命令格式】 **offline-detect interval *interval* threshold *threshold***
no offline-detect
default offline-detect

【参数说明】 *interval*：下线检测周期，交换机设备上取值范围为 6-65535min；非交换机设备取值范围为 1-65535min。默认 8 小时，即 480min。

threshold：流量门限，取值范围为 0-4294967294Bytes。默认为 0，表示无流量检测下线。

no offline-detect：关闭用户在线检测功能。

default offline-detect：恢复成默认值，即 8 小时无流量就将已在线认证用户下线。

【缺省配置】 8 小时

【命令模式】 全局模式

【使用指导】 此命令可以用来配置用户在线活，指定在一定的时间段内在线认证用户的流量低于指定的门限时将用户下线。使用 **no offline-detect** 命令关闭用户在线检测功能，使用 **default offline-detect** 恢复成缺省的检测方式。

检验方法

- 可以通过以下方法检验认证用户在线检测的配置效果：
- 配置了在线用户检测功能后，用户上线后，将指定的已认证终端关机，然后等待指定的周期，在设备上使用 **dot1x** 或 **web** 认证提供的在线用户查询命令确认指定的用户已经下线。

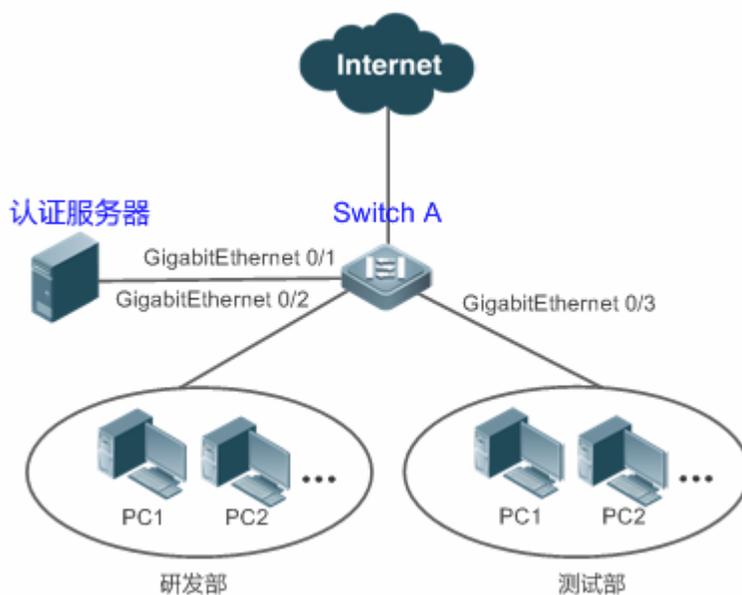
配置举例

i 以下配置举例，仅介绍与 SCC 相关的配置。

通过用户在线检测功能，指定 5min 内无流量就将用户下线。

【网络环境】

图 6-5



- 【配置方法】
- 在接入端口 Gi 0/2 上开启 dot1x 受控，并配置认证所需参数，基于 MAC 认证
 - 配置用户在线检测功能，指定 5min 内无流量就将用户下线。

Switch A

```
sw1(config)# offline-detect interval 5 threshold 0
```

- 【检验方法】
- 在研发部中的一台电脑上通过 dot1x SU 客户端认证上线 然后将电脑直接关机 等待 6min 后 在 switch1 设备上使用 dot1x 提供的在线用户查询命令确认该用户已经下线。

Switch A

```
sw1(config)#show running-config | include offline-detect  
offline-detect interval 5
```

6.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
查看免认证 VLAN 配置	show direct-vlan
显示指定接口上 IPv4 用户表项信息。	show nac-author-user [interface <i>interface-name</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
监视访问列表运行过程信息	debug scc event
调试查看当前 SCC 的相关用户表项	debug scc user [mac author mac]
调试查看当前 SCC 中保存的所有业务下发的相关 ACL 摘要信息	debug scc acl-show summary
调试查看当前 SCC 中保存的所有 ACL 信息	debug scc acl-show all

7 全局 IP+MAC 绑定

7.1 概述

通过手动配置全局 IP 和 MAC 地址绑定功能,可以对输入的报文进行 IP 地址和 MAC 地址绑定关系的验证。如果将一个指定的 IP 地址和一个 MAC 地址绑定,则设备只接收源 IP 地址和 MAC 地址均匹配这个绑定地址的 IP 报文;否则该 IP 报文将被丢弃。

利用地址绑定这个特性,可以严格控制设备的输入源的合法性。需要注意的是,通过地址绑定控制交换机的输入,将优先于 802.1X、端口安全以及 ACL 生效

 下文仅介绍全局 IP+MAC 绑定的相关内容。

协议规范

- 无

7.2 典型应用

典型应用	场景描述
全局IP+MAC地址绑定	仅指定 IP 的主机可以访问网络,主机在同一台设备下是可以移动的

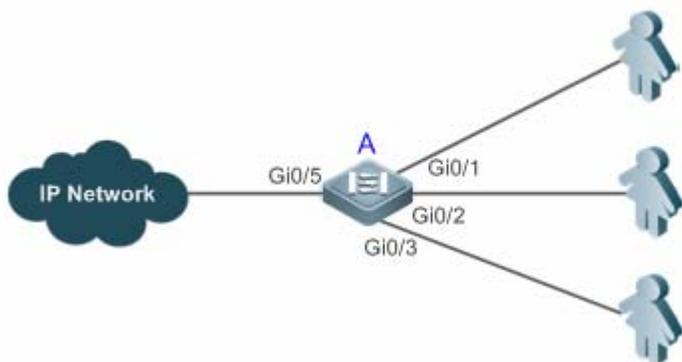
7.2.1 全局IP+MAC地址绑定

应用场景

为了方便管理,管理员为每台主机固定分配了一个 IP 地址。

- 仅指定 IP 的主机可以访问外部网络,防止非法主机盗用 IP。
- 主机可以在相同设备下可以自由移动。

图 7-1



- 【注释】 A 为接入设备。
 User 为静态分配了 IP 地址的接入主机。
 IP Network 为外部 IP 网络。

功能部属

- 手动配置全局 IP 和 MAC 地址绑定（本例列举 3 个用户）

用户	所属主机 MAC 地址	分配的 IP 地址
User1	00d0.3232.0001	192.168.1.10
User2	00d0.3232.0002	192.168.1.20
User3	00d0.3232.0003	192.168.1.30

- 全局使能 IP 和 MAC 地址绑定功能
- 将设备的上链口（本例为 Gi0/5 口）配置为例外口

7.3 功能详解

基本概念

地址绑定例外端口

IP 地址和 MAC 地址绑定功能缺省对设备上的所有端口都生效，通过配置例外端口的方式可以使绑定功能在部份端口上不生效。在应用中设备的上链端口的 IP 报文的绑定关系是不确定的，通常将设备的上链端口配置为例外口，此时上链端口则不进行 IP 地址与 MAC 地址的绑定检查。

功能特性

功能特性	作用
配置全局IP+MAC地址绑定	对 IPv4 报文进行转发控制
配置地址绑定例外端口	全局地址绑定功能在对应端口上不生效

7.3.1 配置全局IP+MAC地址绑定

工作原理

配置 IP 和 MAC 地址绑定功能，可以对输入的报文进行 IP 地址和 MAC 地址绑定关系的验证。如果将一个指定的 IP 地址和一个 MAC 地址绑定，则设备只接收源 IP 地址和 MAC 地址均匹配这个绑定地址的 IP 报文；否则该 IP 报文将被丢弃。

相关配置

▾ 配置 IP+MAC 地址绑定

全局模式下，使用 **address-bind** 命令添加或者删除 IPv4+MAC 地址绑定。

▾ 配置使得 IP+MAC 地址绑定生效

全局模式下，使用 **address-bind install** 命令配置地址绑定功能生效，默认不生效。

7.3.2 配置地址绑定例外端口

工作原理

通过配置例外端口的方式可以使绑定功能在部份端口上不生效。

相关配置

▾ 配置地址绑定例外端口

使用 **address-bind uplink** 命令可以配置例外端口，默认全部都是非例外端口。

7.4 产品说明



S7508 产品全局 IP+MAC 绑定和端口安全/DOT1X 共用时，不管安全通道是否开启，全局 IP+MAC 绑定地址和端口安全地址/DOT1X 认证用户是并集关系，即符合任意安全功能的地址可以通信。



S7508 产品全局 IP+MAC 绑定和端口安全/DOT1X 共用时，不管安全通道是否开启，全局 IP+MAC 绑定地址和端口安全地址/DOT1X 认证用户是并集关系，即符合任意安全功能的地址可以通信。

7.5 配置详解

配置项	配置建议 & 相关命令	
配置全局IP+MAC地址绑定	 必须配置。用于生成地址绑定并开启绑定功能。	
	address-bind	配置生成全局 IPv4+MAC 地址绑定
	address-bind install	开启地址绑定功能
配置地址绑定例外端口	 可选配置。用于配置部分端口的地址绑定功能不生效。	
	address-bind uplink	配置地址绑定的例外端口

7.5.1 配置全局IP+MAC地址绑定

配置效果

- 生成全局 IPv4+MAC 地址绑定
- 开启地址绑定功能对 IPv4 报文进行转发控制

注意事项

- 如果执行 **address-bind install** 之后，没有配置 IP+MAC 绑定，则所有 IP+MAC 绑定功能不生效，所有报文可以通过。

配置方法

配置全局 IP+MAC 地址绑定

- 必须配置，全局配置模式下配置。

开启地址绑定功能

- 必须配置，全局配置模式下配置。

检验方法

使用 **show run** 或者 **show address-bind** 验证配置是否生效。

相关命令

配置全局 IP+MAC 地址绑定

【命令格式】 **address-bind** { *ip-address* } *mac-address*

- 【参数说明】 *ip-address* : 绑定的 IPv4 地址
mac-address : 绑定的 MAC 地址
- 【命令模式】 全局配置模式
- 【使用指导】 配置 IP 地址和 MAC 地址的绑定关系

配置开启地址绑定功能

- 【命令格式】 **address-bind install**
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 使全局 IP 和 MAC 地址绑定生效，控制 IPv4 报文的转发。

配置举例

配置全局 IP+MAC 地址绑定并使能地址绑定功能

- 【配置方法】
- 配置生成全局 IPv4+MAC 地址绑定
 - 开启地址绑定功能

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# address-bind 192.168.5.1 00d0.f800.0001
Ruijie(config)# address-bind install
```

- 【检验方法】 查看设备上的全局 IP+MAC 地址绑定

```
Ruijie#show address-bind
Total Bind Addresses in System : 1
IP Address           Binding MAC Addr
-----
192.168.5.1         00d0.f800.0001
```

常见错误

- 无

7.5.2 配置地址绑定例外端口

配置效果

- 配置指定端口的地址绑定功能不生效，所有 IP 报文都可以转发

注意事项

- 只能在交换口或者 L2AP 口进行配置

配置方法

配置地址绑定例外端口

- 可选配置，全局配置模式下配置，需要特殊指定地址绑定功能不生效的端口时配置。

检验方法

使用 **show run** 或者 **show address-bind uplink** 验证配置是否生效。

相关命令

配置地址绑定例外端口

【命令格式】 **address-bind uplink interface-id**

【参数说明】 *interface-id* : 交换口或 L2AP 口

【命令模式】 全局配置模式

【使用指导】 -

配置举例

配置地址绑定例外端口

- 【配置方法】
- 配置生成全局 IPv4+MAC 地址绑定
 - 开启地址绑定功能
 - 配置地址绑定例外端口

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# address-bind 192.168.5.1 00d0.f800.0001
Ruijie(config)# address-bind install
Ruijie(config)# address-bind uplink GigabitEthernet 0/1
```

- 【检验方法】 查看设备上的全局 IP+MAC 地址绑定

```
Ruijie#show address-bind
Total Bind Addresses in System : 1
IP Address          Binding MAC Addr
-----
192.168.5.1        00d0.f800.0001
Ruijie#show address-bind uplink
Port      State
```

```
-----  
-----  
Gi0/1      Enabled  
Default    Disabled
```

常见错误

- 无

7.6 监视与维护

清除各类信息

无

查看运行情况

作用	命令
查看设备上的 IP 地址与 MAC 地址绑定配置	show address-bind
查看设备上的例外口信息	show address-bind uplink

查看调试信息

无

8 PASSWORD-POLICY

8.1 概述

Password Policy (口令策略) 是设备本地认证时提供的口令安全功能，它依据管理员设置的口令安全策略对用户的登录密码和用户的登录状态进行控制。

i 下文仅介绍 Password Policy 的相关内容。

协议规范

- 暂无可遵循的协议规范或标准。

8.2 功能详解

基本概念

📄 口令最小长度限制

根据系统的安全要求，管理员可设置用户口令的最小长度。当用户配置口令时，如果输入的口令长度小于限定的最小长度，系统将不允许用户设置该口令，并提示出错信息，提醒用户重新设置口令。

📄 强口令检测功能

口令的复杂度越低，其被成功破解的可能性就越大，比如与账号同名的口令、只包含字符或数字的简单口令等。出于安全性考虑，管理员可以打开强口令检测功能，确认用户设置的口令具有较高的复杂度。打开强口令检测功能后，对不符合口令强度检测策略的如下口令提示告警：

- 1、与账号同名的口令；
- 2、只包含字符或数字的简单口令。

📄 口令生存周期

口令生存周期用于限制用户口令的使用时间。当口令的使用时间超过限定值后，需要用户更换口令。

当用户登录时，如果用户输入已经过期的口令，系统将提示该口令已经过期，需要重新设置口令。在重新设置口令时，如果输入的新口令不符合要求，或者连续两次输入的新口令不一致，系统将要求用户继续重新输入。

📄 口令重复使用限制功能

当用户修改口令时，系统会要求用户设置新的口令，旧的口令将被记录下来，形成该用户的历史记录。如果用户新设置的口令以前被使用过，系统将给出错误提示，并要求用户重新设置口令。

可以配置每个用户口令历史记录的最大条数，当口令历史记录的条数超过配置的最大条数时，新的口令历史记录将覆盖该用户最老的一条口令历史记录。

📌 口令加密存储

出于安全考虑，管理员可以打开口令加密存储功能，打开此功能后，进行 **show running-config** 查看配置或 **write** 保存配置文件时，用户设置的各种口令将变成密文；如果再次关闭口令加密存储功能，已经变为密文的口令不会恢复为明文。

8.3 配置详解

配置项	配置建议 & 相关命令	
配置口令安全策略	 可选配置。用于配置口令安全相关的策略组合。	
	password policy life-cycle	设置口令生存周期。
	password policy min-size	限制用户口令的最小长度。
	password policy no-repeat-times	限制重复使用最近几次已配置过的口令。
	password policy strong	打开强口令检测功能。
	service password-encryption	设置口令加密存储。

8.3.1 配置口令安全策略基本功能

配置效果

- 为设备的本地认证提供口令安全策略，用户可以配置不同的安全策略来实现口令安全管理的目的。

注意事项

- 配置了口令安全策略后，只对全局口令（通过 **enable password**、**enable secret** 命令配置）和本地用户口令（通过 **username name password password** 命令配置），对于 Line 模式下面的口令不生效。

配置方法

📌 设置口令生存周期

- 可选配置。
- 若无特殊要求，应在每台需要设置口令生存周期的设备上面配置。

📌 限制用户口令的最小长度

- 可选配置。
- 若无特殊要求，应在每台需要限制口令最小长度的设备上面配置。

限制重复使用最近几次已配置过的口令

- 可选配置。
- 若无特殊要求，应在每台需要限制重复使用最近几次已配置过的口令的设备上面配置。

打开强口令检测功能

- 可选配置。
- 若无特殊要求，应在每台需要进行强口令检测的设备上面配置。

设置口令加密存储

- 可选配置。
- 若无特殊要求，应在每台需要设置口令加密存储的设备上面配置。

检验方法

在设备上面配置一个本地用户，并为此用户配置合法、非法口令。

- 配置合法的口令时，设备能否正确添加用户口令。
- 配置非法的口令时，设备能否提示相应的 Log 信息。

相关命令

设置口令生存周期

【命令格式】 **password policy life-cycle days**

【参数说明】 **life-cycle days**：口令生存周期，单位：天，范围：1~65535。

【命令模式】 全局配置模式

【使用指导】 口令生存周期用来限制用户口令的使用时间，当口令超过生存周期后，系统在下次用户登录时，将提示用户修改口令。

限制用户口令的最小长度

【命令格式】 **password policy min-size length**

【参数说明】 **min-size length**：指定口令最小长度，范围：1~31。

【命令模式】 全局配置模式

【使用指导】 此命令用来配置口令的最小长度限制，若没有配置口令的最小长度限制，用户设置口令时将不进行口令最小长度限制。

限制重复使用最近几次已配置过的口令

【命令格式】 **password policy no-repeat-times times**

【参数说明】 **no-repeat-times times**：最近几次已配置过的口令，范围：1~31。

【命令模式】 全局配置模式

【使用指导】 开启此功能后，用户最近几次使用过的旧口令将被记录下来，形成该用户的口令历史记录。如果用户新设置的

口令以前被使用过，系统将给出错误提示，口令更改失败。

可以配置用户口令历史记录的最大条数，当口令历史记录的条数超过配置的最大历史记录条数时，新的口令历史记录将覆盖该用户最老的一条口令历史记录。

📌 打开强口令检测功能

【命令格式】 **password policy strong**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 开启了此功能后，能够在新建用户时对不符合口令强度策略的如下口令配置提示告警：

- 1、与账号同名的口令；
- 2、只包含字符或数字的简单口令。

📌 设置口令加密存储

【命令格式】 **service password-encryption**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 没有设置口令加密存储前，用户配置过程，使用的各种口令均以明文显示和存储，除非是直接使用密文进行配置。出于安全考虑，可以打开口令加密存储功能，打开此功能后，进行 **show running-config** 或 **write** 保存时，用户设置的各​​种口令将变成密文；如果再次关闭口令加密存储功能，已经变为密文的口令不会恢复为明文。

📌 查看用户设置的口令安全策略信息

【命令格式】 **show password policy**

【参数说明】 -

【命令模式】 特权模式、全局模式、接口模式

【使用指导】 查看设备上设置的口令安全策略信息。

配置举例

 以下配置举例，介绍口令安全策略相关的配置。

📌 在设备上面

【网络环境】 假设网络环境中，有以下口令安全需求：

- 1、口令最小长度大于等于 8 个字符；
- 2、口令生存时间为 90 天；
- 3、口令使用加密存储和传输；
- 4、口令重复使用历史记录条数 3 条；
- 5、不允许口令与用户名一样或者只包含简的字符或数字。

- 【配置方法】
- 配置口令最小长度：8。
 - 配置口令生存周期：90 天。
 - 开启口令加密存储功能。

- 配置口令重复使用历史记录条数：3。
- 开启强口令检测功能。

```
Ruijie# configure terminal
Ruijie(config)# password policy min-size 8
Ruijie(config)# password policy life-cycle 90
Ruijie(config)# service password-encryption
Ruijie(config)# password policy no-repeat-times 3
Ruijie(config)# password policy strong
```

【检验方法】 用户设置了相关口令安全策略相关的配置后，在新增用户和口令的时候，将会依据口令安全策略进行相关的检测。

- 通过 **show password policy**，查看用户设置的口令安全策略信息。

```
Ruijie# show password policy

Global password policy configurations:

Password encryption:           Enabled
Password strong-check:         Enabled
Password min-size:             Enabled (8 characters)
Password life-cycle:           Enabled (90 days)
Password no-repeat-times:      Enabled (max history record: 3)
```

常见错误

- 设置口令过期前开始提醒的时间大于口令生存周期。

8.4 监视与维护

查看运行情况

作用	命令
查看用户设置的口令安全策略信息	show password policy

9 端口安全

9.1 概述

端口安全功能用于约束进入一个端口的访问。通过报文的源 MAC 地址来限定报文是否可以进入交换机的端口，用户可以静态设置特定的 MAC 地址或者动态学习限定个数的 MAC 地址来控制报文是否可以进入端口，使能端口安全功能的端口称为安全端口。

 下文仅介绍端口安全的相关内容。

协议规范

- 无

9.2 典型应用

典型应用	场景描述
仅允许特定主机使用端口	出于网络安全的考虑，设备的有些端口只允许被特定的主机使用。

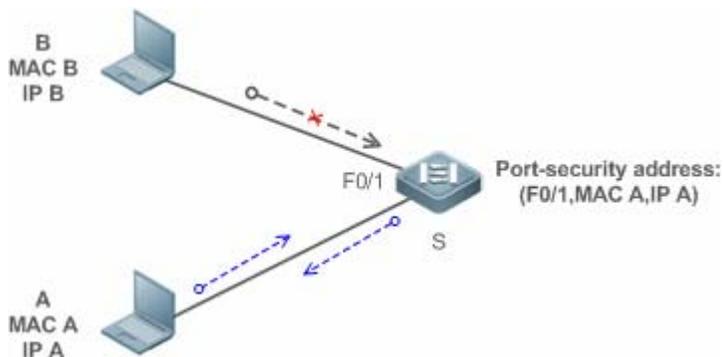
9.2.1 仅允许特定主机使用端口

应用场景

在一个对网络安全有要求的场景中，但是设备又无法做到彻底物理隔绝。需要通过在设备上配置，限制连接到设备端口上的用户 PC。

- 仅指定用户可以连接到端口，并正常使用网络。
- 其他用户即使连接上端口，也无法使用。
- 配置完成后，无需管理员经常维护。

图 9-1



- 【注释】 S为接入设备。
A为允许使用端口 F0/1 的用户 PC。
B为未知的用户 PC。

功能部署

- 在设置端口 F0/1 上的 ARP Check 处于打开模式（略）。
- 在接入设备 S 上开启端口安全功能，设置违例处理行为为 protect。
- 设置端口 F0/1 上允许的最大安全地址数为 1。
- 在端口 F0/1 上配置静态端口安全地址。

9.3 功能详解

基本概念

安全端口

配置了端口安全的端口称之为安全端口，目前锐捷设备限制安全端口不能是镜像的目的端口。

安全地址

在安全端口上绑定的地址称为安全地址，安全地址可以是二层地址，即 MAC 地址；也可以是三层地址，即 IP 或 IP+MAC 地址。当安全地址绑定 IP+MAC，同时设定静态安全 MAC 地址时，设定的静态安全 MAC 地址必须和 IP+MAC 绑定的 MAC 地址相同，否则将因为不符合绑定而无法通信；同理，设定仅 IP 绑定之后，只有符合静态设置或者学习的安全 MAC 地址，同时源 IP 为绑定 IP 的报文才可以进入设备。

动态绑定

让设备自动学习地址，并把学习到的地址转化为安全地址的方式。

静态绑定

手工命令进行安全地址绑定。

安全地址老化

定期删除安全地址记录，端口安全的安全地址支持老化配置，用户可以指定只老化动态学习的地址或同时老化静态配置与动态学习的安全地址。

Sticky MAC 地址功能

将动态学习到的安全地址转换为静态配置。地址不会老化，当保存配置后，重启不用重新学习这些动态安全地址；而如果没有启用该功能，那么动态学习到的安全 MAC 地址在设备重启后要重新学习。

安全违例事件

端口学习到的 MAC 地址超出最大安全地址个数限制时，将触发的安全违例事件。用户可以配置如下 3 种安全违例事件处理模式：

- protect：当违例产生时，对应安全端口将停止 MAC 地址学习并丢弃所有新接入用户的报文。该处理模式为默认的违例处理模式。
- restrict：当违例产生时，除了产生 protect 处理模式的行为，还将发送一个端口违例 Trap 通告。
- shutdown：当违例产生时，除了产生上述两个模式的行为，还将关闭端口。

最大安全地址数

最大安全地址数指的是静态配置与动态学习的安全地址总数，当安全端口下安全地址没有达到最大安全地址数时，安全端口能够动态学习新的动态安全地址，当安全地址数达到最大数时，安全端口将不再学习动态安全地址，如果此时有新的用户接入安全端口，将产生安全违例事件。

功能特性

功能特性	作用
开启端口安全功能	构建端口上安全地址列表
二层用户过滤	处理端口上接收到的，非安全地址的报文。
三层用户过滤	检查经过端口报文的二层地址和三层地址。
安全地址老化	定期删除安全地址。

9.3.1 开启端口安全功能

启动端口安全功能，限制能通过该端口上网的用户报文。

工作原理

当启动端口安全功能后，设备安全模块会检查接收报文的来源。只有报文来源于属于安全地址列表集合时，报文才会被正常转发，否则，报文被丢弃或是端口采用其他的违例处理行为。

端口安全和 802.1x 同时配置时，报文 MAC 字段必须满足 802.1x 或是端口安全的静态 MAC 地址配置才能进入交换机；如果端口设置了安全通道或是全局 IP+MAC 绑定，那么符合安全通道或是全局 IP+MAC 绑定的报文将绕过端口安全的检查。

相关配置

▾ 启动端口上的端口安全功能

缺省情况下，端口上的端口安全功能关闭。

使用 **switchport port-security** 命令可以启动或关闭接口上的端口安全功能。

不能在 SPAN 的目的端口上开启该功能。

▾ 设置端口最大安全地址个数

缺省情况下，一个端口下的最大安全地址个数为 128 个。

使用 **switchport port-security maximum** 命令可以调整接口上的最大安全地址个数。

安全地址个数越小，可以通过端口上网的用户越少。

▾ 设置处理违例的方式

缺省情况下，当安全地址个数满后，安全端口将丢弃未知名地址(不是该端口的安全地址中的任何一个)的包。

使用 **switchport port-security violation** 命令可以修改违例处理方式。

▾ 设置保存动态安全地址

缺省情况下，不会保存动态学习到的安全地址。

使用 **switchport port-security mac-address sticky** 命令可以将动态学习到的地址保存到配置文件，只要配置文件被保存，设备重启后不必重新学习安全地址。

9.3.2 二层用户过滤

设置端口上的安全地址，只有 MAC 地址与安全地址相同的设备才能经由该端口访问网络。

工作原理

为安全端口添加安全地址。当安全端口下安全地址没有达到最大安全地址数时，安全端口能够动态学习新的动态安全地址；当安全地址数达到最大数时，安全端口将不再学习动态安全地址。连接该端口用户的 MAC 地址必须属于安全地址集合，否则，按违例事件进行处理。

相关配置

▾ 为安全端口添加安全地址

缺省情况下，端口动态学习安全地址。若管理员有特殊的需求，可以手工配置安全地址。

使用 **switch portport-security interface** 命令可以在设备上添加/删除安全地址。

9.3.3 三层用户过滤

添加安全地址绑定，同时检查经过端口报文的二层地址和三层地址。

工作原理

三层安全地址有只绑定 IP 和绑定 IP+MAC 两种方式，并且只支持静态绑定，不能进行动态绑定。

当三层安全端口接收报文时，需要解析二层地址和三层地址。只有已经绑定地址的报文才是合法报文，否则认为是非法报文，丢弃，但不产生违例事件。

相关配置

配置安全端口上的安全地址绑定

三层安全地址绑定只能通过手工方式添加。

使用 `switchport port-security binding` 命令可以添加安全地址绑定。

若仅输入 IP 地址，则只绑定 IP；若输入 IP 地址和 MAC 地址，则绑定 IP+MAC。

9.3.4 安全地址老化

定期删除安全地址。打开这个功能，需要设置安全地址的最大个数，这样就可以让设备自动的增加和删除接口上的安全地址。

工作原理

启动老化定时器，定期查询那些老化时间到期的安全地址，并删除。

相关配置

配置安全地址的老化时间

缺省情况下，端口上的所有安全地址不进行老化。

使用 `switchport port-security aging` 命令可以启动老化时间。

使用 `static` 参数可以同时老化静态地址。

9.4 配置详解

配置项	配置建议 & 相关命令
配置安全端口及违例处理方式	 必须配置。用于启动端口安全服务。

	switchport port-security	启动端口安全功能。
	switchport port-security maximum	设置端口最大安全地址个数。
	switchport port-security violation	配置端口安全的违例处理。
	switchport port-security mac-address sticky	配置动态地址自动保存。
配置安全端口上的安全地址	 可选配置。用于配置安全过滤项。	
	switchport port-security mac-address	接口模式下配置静态安全地址。
	switchport port-security interface mac-address	全局配置模式下配置静态安全地址。
	switchport port-security binding	接口模式下配置安全地址绑定。
	switchport port-security interface binding	全局配置模式下配置安全地址绑定。
	switchport port-security aging	为一个端口上的所有安全地址配置老化时间。

9.4.1 配置安全端口及违例处理方式

配置效果

- 限制接口上面可以学习到的 MAC 地址个数。
- 过滤 MAC 地址或是 IP 地址或是 IP+MAC 组合非法的报文。

注意事项

- 一个安全端口不能是 SPAN 的目的端口。
- 无法在 DHCP Snooping 信任端口上配置端口安全功能。
- 无法在全局 IP+MAC 的例外口配置端口安全功能。
- 只能在有线的交换口、2 层 AP 口下配置开启，在接口模式下配置。
- 端口安全和其他接入控制功能如 802.1x，全局 IP+MAC 绑定，IP SOURCE GUARD 共用，当这些功能共用时，报文必须同时满足所有的安全检查才能进入交换机；如果端口设置了安全通道，那么符合安全通道的报文将绕过端口安全的检查。

配置方法

▾ 启动端口安全服务

- 必须配置。
- 若无特殊要求，在接入设备的端口上开启。

▾ 配置最大安全地址个数

- 可选配置。若希望调整安全端口上运行的最大安全地址数，可以进行配置
- 在开启了端口安全的端口上进行配置。

配置违例处理方式

- 可选配置。若希望违例情况发生时，不仅仅是丢弃报文，可以配置其他的处理方式
- 在开启了端口安全的端口上进行配置。

配置保存动态学习地址

- 可选配置。若希望设备重启后不需要重新学习安全地址，可以进行配置。
- 在开启了端口安全的端口上进行配置。

检验方法

使用设备提供的显示端口安全配置命令，可以验证配置是否生效。

相关命令

设置端口安全功能

【命令格式】 **switchport port-security**

【参数说明】 -

【命令模式】 接口配置模式

【使用指导】 利用端口安全这个特性，可以通过限制允许访问设备上某个接口的 MAC 地址以及 IP(可选)来实现严格控制对该接口的输入。

设置端口最大安全地址个数

【命令格式】 **switchport port-security maximum value**

【参数说明】 *value* : 安全地址个数 1-128

【命令模式】 接口配置模式

【使用指导】 如果将最大个数设置为 1 并且为该端口配置一个安全地址，则连接到这个口的工作站（其地址为配置的安全地址）将独享该端口的全部带宽。

配置端口安全的违例处理

【命令格式】 **switchport port-security violation { protect | restrict | shutdown }**

【参数说明】 **protect** : 发现违例，则丢弃违例的报文。

restrict : 发现违例，则丢弃违例的报文并且发送 trap。

shutdown : 发现违例，则丢弃报文、并关闭接口。

【命令模式】 接口配置模式

【使用指导】 -

保存动态安全地址到配置文件

- 【命令格式】 **switchport port-security mac-address sticky mac-address [vlan vlan-id]**
- 【参数说明】 *mac-address* : 静态安全地址。
vlan-id : MAC 地址的 VID。
- 【命令模式】 接口配置模式
- 【使用指导】 -

配置举例

i 以下配置举例，仅介绍与端口安全相关的配置。

使能接口 gigabitethernet 0/3 上的端口安全功能，设置最大地址个数为 8，设置违例方式为 protect

- 【配置方法】
- 开启端口安全功能
 - 设置最大安全地址
 - 修改违例处理方式。

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# interface gigabitethernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# switchport mode access
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security maximum 8
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security violation protect
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security mac-address sticky
Ruijie(config-if-GigabitEthernet 0/3)# end
```

- 【检验方法】 查看设备上的端口安全配置

```
Ruijie# show port-security interface gigabitethernet 0/3
Interface : Gi0/3
Port Security: Enabled
Port status : down
Violation mode: Protect
Maximum MAC Addresses:8
Total MAC Addresses:0
Configured MAC Addresses:0
Aging time : 0 mins
SecureStatic address aging : Disabled
```

常见错误

- 在 SPAN 口上开启端口安全。
- 在 DHCP 信任口上面开启端口安全。

- 设置的最大安全地址数小于当前已有安全地址个数。

9.4.2 配置安全端口上的安全地址

配置效果

- 允许特定用户使用端口。
- 定期更新用户的安全地址。

注意事项

- Sticky MAC 地址是一种特殊的 MAC 地址，不受老化机制影响，无论是配置了动态的老化还是静态的老化，Sticky MAC 地址均不会老化。

配置方法

▾ 配置安全地址

- 可选配置。需要手工添加安全地址是进行配置。
- 在开启了端口安全的端口上进行配置。

▾ 配置安全地址绑定

- 可选配置。需要添加三层安全地址时，进行配置。
- 在开启了端口安全的端口上进行配置。

▾ 配置老化时间

- 可选配置。
- 在开启了端口安全的端口上进行配置。

检验方法

- 使用设备提供的显示端口安全配置命令，可以验证配置是否生效。

相关命令

▾ 全局配置模式下为安全端口添加安全地址

【命令格式】 **switchport port-security interface** *interface-id* **mac-address** *mac-address* [**vlan** *vlan-id*]

【参数说明】 *interface-id*：接口 ID。

mac-address：静态安全地址。

vlan-id : MAC 地址的 VID。

【命令模式】 全局模式

【使用指导】 -

▾ 接口配置模式下为安全端口添加安全地址

【命令格式】 **switchport port-security mac-address** *mac-address* [**vlan** *vlan_id*]

【参数说明】 *mac-address* : 静态安全地址。

vlan-id : MAC 地址的 VID。

【配置模式】 接口配置模式

【使用指导】 -

▾ 全局配置模式下为安全端口添加安全地址绑定

【命令格式】 **switchport port-security interface** *interface-id* **binding** [*mac-address* **vlan** *vlan_id*] { *ipv4-address* }

【参数说明】 *interface-id* : 接口 ID。

mac-address : 绑定的源 MAC 地址。

vlan_id : 绑定源 MAC 的 VID。

ipv4-address : 绑定 Ipv4 的 Ip 地址。

【配置模式】 全局模式

【使用指导】 -

▾ 接口配置模式下为安全端口添加安全地址绑定

【命令格式】 **switchport port-security binding** [*mac-address* **vlan** *vlan_id*] { *ipv4-address* }

【参数说明】 *mac-address* : 绑定的源 MAC 地址。

vlan_id : 绑定源 MAC 的 VID。

ipv4-address : 绑定 Ipv4 的 Ip 地址。

【配置模式】 接口配置模式

【使用指导】 -

▾ 为一个接口上的所有安全地址配置老化时间

【命令格式】 **switchport port-security aging** { **static** | **time** *time* }

【参数说明】 **static** : 表示老化时间将同时应用于手工配置的安全地址和自动学习的地址, 否则则只应用于自动学习的地址。

time *time* : 表示这个端口上安全地址的老化时间, 范围是 0 - 1440, 单位是分钟。如果设置为 0, 则老化功能实际上被关闭。

【配置模式】 接口配置模式

【使用指导】 -

配置举例

▾ 为接口 gigabitethernet 0/3 配置一个安全地址, 地址 MAC 为 00d0.f800.073c。

- 【配置方法】
- 开启端口安全
 - 添加安全地址。

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# interface gigabitethernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# switchport mode access
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security mac-address 00d0.f800.073c vlan
1
Ruijie(config-if-GigabitEthernet 0/3)# end
```

【检验方法】 查看设备上的端口安全配置

```
Ruijie# show port-security address
Vlan Mac Address IP Address Type Port Remaining Age(mins)
-----
1 00d0.f800.073c 192.168.12.202 Configured Gi0/3 8
1 00d0.f800.3cc9 192.168.12.5 Configured Gi0/1 7
```

📌 为接口 gigabitethernet 0/3 配置一个安全绑定，绑定 IP 地址为 192.168.12.202。

- 【配置方法】**
- 开启端口安全
 - 添加安全地址绑定。

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# interface gigabitethernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# switchport mode access
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security binding 192.168.12.202
Ruijie(config-if-GigabitEthernet 0/3)# end
```

【检验方法】 查看设备上的端口安全配置

```
Ruijie# show port-security address
Vlan Mac Address IP Address Type Port Remaining Age(mins)
-----
1 00d0.f800.073c 192.168.12.202 Configured Gi0/3 8
1 00d0.f800.3cc9 192.168.12.5 Configured Gi0/1 7
```

📌 为接口 gigabitethernet 0/3 配置一个安全地址和一个安全绑定，安全地址 MAC 为 00d0.f800.073c，绑定 IP 地址为 0000::313b:2413:955a:38f4。

- 【配置方法】**
- 开启端口安全
 - 添加安全地址绑定。

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# interface gigabitethernet 0/3
```

```
Ruijie(config-if-GigabitEthernet 0/3)# switchport mode access
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security binding 00d0.f800.073c vlan 1
0000::313b:2413:955a:38f4
Ruijie(config-if)# end
```

【检验方法】 查看设备上的端口安全配置

```
Ruijie# show port-security address
Vlan Mac Address IP Address Type Port Remaining Age(mins)
-----
1 00d0.f800.073c 192.168.12.202 Configured Gi0/3 8
1 00d0.f800.3cc9 192.168.12.5 Configured Gi0/1 7
```

📌 **配置一个接口 gigabitethernet 0/3 上的端口安全的老化时间，老化时间设置为 8 分钟，老化时间同时应用于静态配置的安全地址。**

- 【配置方法】**
- 开启端口安全
 - 配置老化时间。

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# interface gigabitethernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security aging time 8
Ruijie(config-if-GigabitEthernet 0/3)# switchport port-security aging static
Ruijie(config-if-GigabitEthernet 0/3)# end
```

【检验方法】 查看设备上的端口安全配置

```
Ruijie# show port-security gigabitethernet 0/3
Interface : Gi0/3
Port Security: Enabled
Port status : down
Violation mode:Shutdown
Maximum MAC Addresses:8
Total MAC Addresses:0
Configured MAC Addresses:0
Aging time : 8 mins
SecureStatic address aging : Enabled
```

常见配置错误

- 无。

9.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
显示所有的安全地址 ,或者是指定接口的所有安全地址。	show port-security address [interface <i>interface-id</i>]
显示所有的安全绑定 ,或者是指定接口的所有安全绑定。	show port-security binding [interface <i>interface-id</i>]
显示所有生效的端口安全地址和端口安全绑定记录。	show port-security all
显示了接口上的端口安全配置	show port-security interface <i>interface-id</i>
显示端口安全的统计信息	show port-security

查看调试信息

无

10 STORM CONTROL

10.1 概述

当局域网中存在过量的广播、多播或未知名单播数据流时，就会导致网络变慢和报文传输超时机率大大增加。这种情况称之为局域网风暴。拓扑协议的执行错误或对网络的错误配置都有可能产生风暴。

用户可以分别针对广播、多播和未知名单播数据流进行风暴控制。当设备端口接收到的广播、多播或未知名单播数据流的速率超过所设定的带宽、每秒允许通过的报文数或者每秒允许通过的千比特数时，设备将只允许通过所设定带宽、每秒允许通过的报文数或者每秒允许通过的千比特数的数据流，超出限定范围部分的数据流将被丢弃，直到数据流恢复正常，从而避免过量的泛洪数据流进入局域网中形成风暴。

协议规范

无。

10.2 典型应用

典型应用	场景描述
网络防攻击	网络防攻击，开启风暴控制功能，防止泛洪

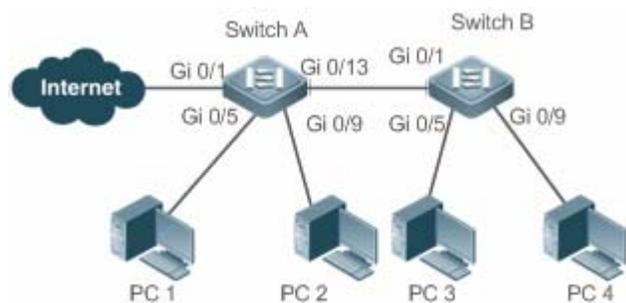
10.2.1 网络防攻击

应用场景

网络防攻击，应用需求如下：

- 防止设备受到广播、多播、未知名单播报文攻击。

图 10-1



【注释】 A、B 为接入设备

PC 1、PC 2、PC 3、PC 4 为台式机。

功能部属

- 所有接入设备（本例为 Switch A、Switch B）的端口上开启风暴控制功能。

10.3 功能详解

基本概念

▾ 风暴控制

当设备端口接收到的未知名单播数据流、组播数据流或者广播数据流的速率超过所设定的对应类型的报文数据流带宽、每秒允许通过的最大报文数或者每秒允许通过的最大千比特数时，设备将只允许通过所设定带宽、每秒最大报文数或者每秒最大千比特数的数据流，超出限定范围部分的数据流将被丢弃，直到数据流恢复正常。

▾ 基于带宽百分比的风暴控制

当设备端口接收到的数据流的速率超过所设定的带宽时，设备将只允许通过所设定带宽的数据流，超出带宽部分的数据流将被丢弃，直到数据流恢复正常。

▾ 基于每秒报文数限制的风暴控制

当设备端口接收到的数据流的速率超过所设定的每秒允许通过的最大报文数时，设备将只允许通过所设定每秒最大报文数的数据流，超出最大允许的每秒报文数部分的数据流将被丢弃，直到数据流恢复正常。

▾ 基于每秒千比特数限制的风暴控制

当设备端口接收到的数据流的速率超过所设定的每秒允许通过的最大每秒千比特数时，设备将只允许通过所设定每秒最大千比特数的数据流，超出最大允许的每秒千比特数部分的数据流将被丢弃，直到数据流恢复正常。

功能特性

功能特性	作用
单播报文风暴控制	开启单播报文风暴控制功能，可以实现对未知名单播报文的流量限制，防止泛洪
组播报文风暴控制	开启组播报文风暴控制功能，可以实现对组播报文的流量限制，防止泛洪
广播报文风暴控制	开启广播报文风暴控制功能，可以实现对广播报文的流量限制，防止泛洪

10.3.1 单播报文风暴控制

单播报文风暴控制功能用于监控设备端口接收到的未知名单播数据流的速率，以实现未知名单播报文在局域网中的流量限制，防止未知名单播报文数据流过大而出现泛洪现象。

工作原理

当设备端口接收到的未知名单播数据流的速率超过所设定的带宽、每秒允许通过的报文数或者每秒允许通过的千比特数时，设备将只允许通过所设定限定范围的未知名单播数据流，超出限定范围部分的未知名单播数据流将被丢弃，直到数据流恢复正常。

相关配置

▾ 启动接口上单播报文风暴控制之功能

缺省情况下，接口上的单播报文风暴控制功能关闭。

使用 `storm-control unicast [{ level percent | pps packets | rate-bps }]` 命令可以启动接口上的单播报文风暴控制功能。

使用 `no storm-control unicast` 或者 `default storm-control unicast` 命令可以关闭接口上的单播报文风暴控制功能。

命令默认参数由相关产品决定。

10.3.2 组播报文风暴控制

组播报文风暴控制功能用于监控设备端口接收到的组播数据流的速率，以实现对组播报文在局域网中的流量限制，防止组播报文数据流过大而出现泛洪现象。

工作原理

当设备端口接收到的组播数据流的速率超过所设定的带宽、每秒允许通过的报文数或者每秒允许通过的千比特数时，设备将只允许通过所设定限定范围的组播数据流，超出限定范围部分的组播数据流将被丢弃，直到数据流恢复正常。

相关配置

▾ 启动接口上组播报文风暴控制之功能

缺省情况下，接口上的组播报文风暴控制功能关闭。

使用 `storm-control multicast [{ level percent | pps packets | rate-bps }]` 命令可以启动接口上的组播报文风暴控制功能。

使用 `no storm-control multicast` 或者 `default storm-control multicast` 命令可以关闭接口上的组播报文风暴控制功能。

命令默认参数由相关产品决定。

10.3.3 广播报文风暴控制

广播报文风暴控制功能用于监控设备端口接收到的广播数据流的速率，以实现对广播报文在局域网中的流量限制，防止广播报文数据流过大而出现泛洪现象。

工作原理

当设备端口接收到的广播数据流的速率超过所设定的带宽、每秒允许通过的报文数或者每秒允许通过的千比特数时，设备将只允许通过所设定范围的广播数据流，超出限定范围部分的广播数据流将被丢弃，直到数据流恢复正常。

相关配置

启动接口上广播报文风暴控制之功能

缺省情况下，接口上的单播报文风暴控制功能关闭。

使用 `storm-control broadcast [{ level percent | pps packets | rate-bps }]` 命令可以启动接口上的广播报文风暴控制功能。

使用 `no storm-control broadcast` 或者 `default storm-control broadcast` 命令可以关闭接口上的广播报文风暴控制功能。

命令默认参数由相关产品决定。

10.4 产品说明



- 缺省情况下，S7508 产品，接口上的风暴控制功能关闭，风暴控制的缺省值为端口带宽的百分之一。
- 对 S7508 的所有系列线卡来说，如果同一个接口上已经配置了组播/广播报文/非单播报文基于每秒报文数的风暴控制，那么该接口就不能配置单播报文基于带宽百分比/每秒千比特数的风暴控制；如果同一个接口上已经配置了组播/广播报文/非单播报文基于每秒带宽百分比/每秒千比特数风暴控制，那么该接口就不能配置单播报文基于每秒报文数的风暴控制。



- 缺省情况下，S7508 产品，接口上的组播报文风暴控制功能关闭，组播报文风暴控制的缺省值为端口带宽的百分之一。
- 对 S7508 的所有系列线卡来说，如果同一个接口上已经配置了单播/广播报文基于每秒报文数的风暴控制，那么该接口就不能配置组播报文基于带宽百分比/每秒千比特数的风暴控制；如果同一个接口上已经配置了单播/广播报文基于每秒带宽百分比/每秒千比特数风暴控制，那么该接口就不能配置组播报文基于每秒报文数的风暴控制。



- 缺省情况下，S7508 产品，接口上的广播报文风暴控制功能关闭，广播报文风暴控制的缺省值为端口带宽的百分之一。
- 对 S7508 的所有系列线卡来说，如果同一个接口上已经配置了单播/组播报文基于每秒报文数的风暴控制，那么该接口就不能配置广播报文基于带宽百分比/每秒千比特数的风暴控制；如果同一个接口上已经配置了单播/组播报文基于每秒带宽百分比/每秒千比特数风暴控制，那么该接口就不能配置广播报文基于每秒报文数的风暴控制。

10.5 配置详解

配置项	配置建议 & 相关命令
配置风波控制基本功能	 必须配置。开启风暴控制
	<pre>storm-control { broadcast multicast unicast} [{ level percent pps packets rate-bps}]</pre> 启动风暴控制

10.5.1 配置风波控制基本功能

配置效果

- 可以基于广播、多播、单播开启风暴控制功能，功能开启之后可以实现对广播、组播、未知名单播报文进行风暴控制，防止泛洪。

注意事项

- 功能开启，配置命令可以无需带任何参数，比如命令 **storm-control unicast** 开启基于未知名单播报文的风暴控制，所有命令可选参数使用系统默认值。

配置方法

启动基于单播报文风暴控制

- 必须配置。
- 若无特殊要求，应在每台设备上启动单播报文风暴控制功能。

启动基于组播报文风暴控制

- 必须配置。
- 若无特殊要求，应在每台设备上启动组播报文风暴控制功能。

启动基于广播报文风暴控制

- 必须配置。
- 若无特殊要求，应在每台设备上启动广播报文风暴控制功能。

检验方法

- **show storm-control** 查看命令是否配置成功。

相关命令

启动基于单播报文风暴控制

【命令格式】 **storm-control unicast** [{ **level percent** | **pps packets** | **rate-bps** }]

【参数说明】 **level percent** : 指定带宽百分比。

pps packets : 指定每秒报文数

rate-bps: 指定速率

【命令模式】 端口模式

【使用指导】 必须是交换口

启动基于组播报文风暴控制

【命令格式】 **storm-control multicast** [{ **level percent** | **pps packets** | **rate-bps** }]

【参数说明】 **level percent** : 指定带宽百分比。

pps packets : 指定每秒报文数

rate-bps: 指定速率

【命令模式】 接口模式

【使用指导】 必须是交换口

启动基于广播报文风暴控制

【命令格式】 **storm-control broadcast** [{ **level percent** | **pps packets** | **rate-bps** }]

【参数说明】 **level percent** : 指定带宽百分比。

pps packets : 指定每秒报文数

rate-bps: 指定速率

【命令模式】 接口模式

【使用指导】 必须是交换口

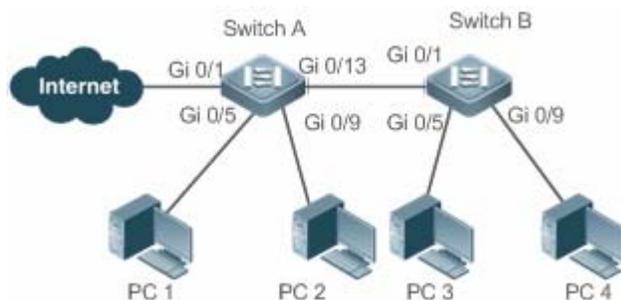
配置举例

i 以下配置举例，仅介绍与风暴控制相关的配置。

在设备上开启风暴控制功能

【网络环境】

图 10-2



【配置方法】 ● 在接入设备 Switch A、Switch B 上配置风暴控制

```

Switch A
Ruijie(config)#interface range gigabitEthernet 0/5,0/9,0/13
Ruijie(config-if-range)#storm-control broadcast
Ruijie(config-if-range)#storm-control multicast
Ruijie(config-if-range)#storm-control unicast

Switch B
Ruijie(config)#interface range gigabitEthernet 0/1,0/5,0/9
Ruijie(config-if-range)#storm-control broadcast
Ruijie(config-if-range)#storm-control multicast
Ruijie(config-if-range)#storm-control unicast

```

【检验方法】 检查 Switch A、Switch B 上是否开启风暴控制

```

Switch A
Ruijie# sho storm-control
Interface                Broadcast Control Multicast Control Unicast Control Action
-----
GigabitEthernet 0/1      Disabled           Disabled           Disabled    none
GigabitEthernet 0/5      default            default            default     none
GigabitEthernet 0/9      default            default            default     none
GigabitEthernet 0/13     default            default            default     none

Switch B
Ruijie#sho storm-control
Interface                Broadcast Control Multicast Control Unicast Control Action
-----
GigabitEthernet 0/1      default            default            default     none
GigabitEthernet 0/5      default            default            default     none
GigabitEthernet 0/9      default            default            default     none

```

常见错误

无

10.6 监视与维护

清除各类信息

无

查看运行情况

作用	命令
查看风暴控制信息。	show storm-control [<i>interface-type interface-number</i>]

查看调试信息

无

11 SSH

11.1 概述

SSH (Secure Shell , 安全外壳) 连接提供的功能类似于一个 Telnet 连接 , 与 Telnet 不同的是基于该连接所有的传输都是加密的。当用户通过一个不能保证安全的网络环境远程登录到设备时 , SSH 特性可以提供安全的信息保障和强大的认证功能 , 以保护设备不受诸如 IP 地址欺诈、明文密码截取等攻击。

设备支持 SSH 服务器功能 , 可以接受多个 SSH 客户端的连接。同时 , 设备还支持 SSH 客户端功能 , 允许用户与支持 SSH 服务器功能的设备建立 SSH 连接 , 从而实现本地设备通过 SSH 安全登录到远程设备上进行管理的功能。

i 目前 , 设备作为 SSH 服务器或 SSH 客户端时 , 支持 SSHv1 和 SSHv2 两个版本。锐捷 SSH 服务同时支持 IPv4 和 IPv6 两种协议。

i 下文仅介绍 SSH 的相关内容。如无特殊说明 , 文中的 SSH 均指 SSHv2。

协议规范

- RFC 4251 : The Secure Shell (SSH) Protocol Architecture
- RFC 4252 : The Secure Shell (SSH) Authentication Protocol
- RFC 4253 : The Secure Shell (SSH) Transport Layer Protocol
- RFC 4254 : The Secure Shell (SSH) Connection Protocol
- RFC 4419 : Diffie-Hellman Group Exchange for the Secure Shell (SSH) Transport Layer Protocol
- RFC 4716 : The Secure Shell (SSH) Public Key File Format
- RFC 4819 : Secure Shell Public Key Subsystem
- RFC 3526 : More Modular Exponential (MODP) Diffie-Hellman groups for Internet Key Exchange (IKE)
- RFC 2409 : The Internet Key Exchange (IKE)
- RFC 1950 : ZLIB Compressed Data Format Specification version 3.3
- draft-ietf-secsh-filexfer-05 : SSH File Transfer Protocol
- draft-ylonen-ssh-protocol-00 : The SSH (Secure Shell) Remote Login Protocol 其协议版本为 1.5 , Comware 实现了协议的 Server 功能 , 没有实现 Client 功能

11.2 典型应用

典型应用	场景描述
SSH设备管理	用户使用 SSH 对设备进行管理
SSH本地线路认证	采用本地线路口令认证方式进行 SSH 用户认证

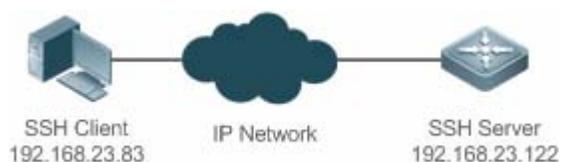
SSH的AAA认证	采用 AAA 认证方式进行 SSH 用户认证
SSH公钥认证	采用公钥认证方式进行 SSH 用户认证
SSH文件传输	用户使用客户端的 SCP 命令与 SSH 服务器端进行数据传输

11.2.1 SSH设备管理

应用场景

用户可以使用 SSH 对设备进行管理，前提是必须打开 SSH Server 功能，默认情况下是关闭该功能的。由于 Windows 自带的 Telnet 组件不支持 SSH，因此必须使用第三方客户端软件，当前兼容性较好的客户端包括：Putty，Linux，SecureCRT。下面以客户端软件 Putty 为例介绍 SSH Client 的配置，组网图如下所示。

图 11-1 SSH 设备管理组网图



功能部署

SSH Client 的配置要点如下：

- 打开 Putty 客户端工具软件。
- 在 Putty 中的 Session 选项卡中填写 SSH Server 的主机 IP、SSH 端口号 22 以及连接类型为 SSH。
- 在 Putty 中的 SSH 选项卡中选择 SSH 协议版本号为 2。
- 在 Putty 中的 SSH 选项卡中选择认证方式为 “Keyboard-interactive”。
- 点击 open 按钮连接服务器主机。
- 在用户名密码认证窗口输入正确的用户名与密码进入终端登录界面。

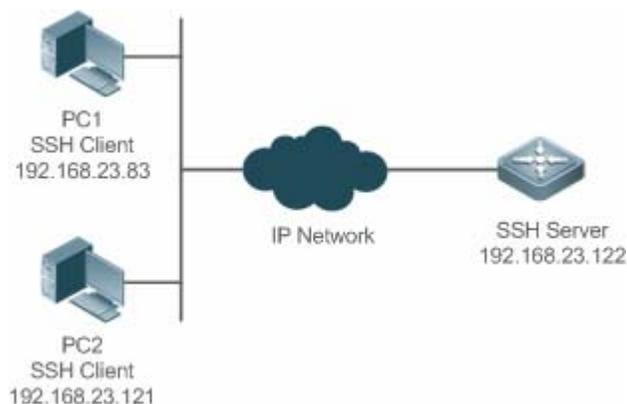
11.2.2 SSH本地线路认证

应用场景

SSH 用户可以采用本地线路口令认证方式进行用户认证，如下图所示。为了保证数据信息交换的安全，PC1、PC2 作为 SSH 客户端，采用 SSH 协议登录到打开 SSH Server 的网络设备上。具体要求如下：

- SSH 用户采用的认证方式为线路口令认证。
- 同时启用 0-4 这五条线路，其中线路 0 的登录口令为 “passzero”，其余四条线路的登录口令均为 “pass”，用户名任意。

- 图 11-2 SSH 本地线路口令保护组网图



功能部署

- SSH Server 的配置要点如下：

1. 全局打开 SSH Server。SSH Server 默认支持 SSH1 和 SSH2 两个版本。
2. 配置密钥。通过该密钥，SSH 服务器将从 SSH 客户端收到的口令密文进行解密，将解密后的明文同服务器上保存的口令进行比较，并返回认证成功或失败的消息。SSH 1 使用 RSA 密钥；SSH 2 使用 RSA 或者 DSA 密钥。
3. 配置 SSH 服务器 FastEthernet 0/1 接口的 IP 地址。SSH 客户端通过该地址连接 SSH 服务器。SSH 客户端至 SSH 服务器路由可达。

- SSH Client 的配置要点如下：

SSH 客户端软件有多种，例如 Putty、Linux、OpenSSH 等，本文中仅以客户端软件 Putty 为例，说明 SSH 客户端的配置方法。

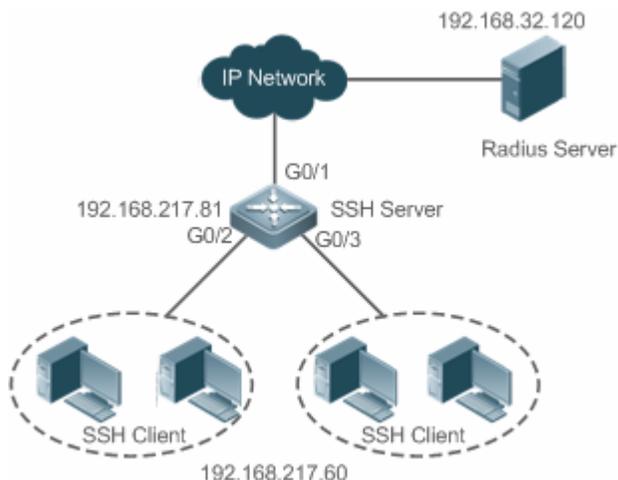
1. 打开 Putty 的连接框，选择使用 SSH1 进行认证登录，使用 SSH2 认证方法类似。
2. 设置 SSH 服务器的 IP 地址与连接端口号，由组网拓扑图可知，服务器主机 IP 为 192.168.23.122，连接端口号为 22，点击 open 按钮进行连接。由于当前认证方式不需要用户名，此处“用户名”可以任意输入，但是不能为空（本例用户名设置为 anyone）。

11.2.3 SSH的AAA认证

应用场景

SSH 用户可以采用 AAA 认证方式进行用户认证，如下图所示。为了保证数据信息交换的安全，PC 作为 SSH 客户端，采用 SSH 协议登录到打开 SSH Server 的网络设备上。为了更好地进行安全管理，SSH 客户端登录用户界面采用 AAA 认证方式；同时出于稳定性方面考虑，在 AAA 认证方法列表中配置两种认证方法：Radius 服务器认证和本地认证。优先选择 Radius 服务器，当 Radius 服务器没有响应时选择本地认证方法。

图 11-3 SSH AAA 认证组网图



功能部署

- SSH 客户端到 SSH 服务器端的路由可达，SSH 服务器到 Radius 服务器端的路由可达。
- 在网络设备上进行 SSH Server 相关配置。
- 在网络设备上进行 AAA 认证相关配置。AAA 通过创建方法列表来定义身份认证、类型，然后将这些方法列表应用于特定的服务或接口上。

11.2.4 SSH公钥认证

应用场景

SSH 用户可以采用 Public-key 认证方式进行用户认证，公钥算法为 RSA 或 DSA，如下图所示。通过配置客户端使用 SSH 协议与服务器端进行安全连接。

图 11-4 SSH 公钥认证组网图



功能部署

- 客户端公钥认证方式，首先要在客户端生成一个密钥对（RSA 或 DSA），然后将其中的公钥放置在 SSH 服务器上，并选择使用 Public-Key 认证方式。
- 在客户端生成了密钥以后，SSH 服务器端需要将客户端的公钥文件复制到 flash 中，并且与 SSH 用户名关联。每个用户可以关联一个 RSA 公钥和一个 DSA 公钥。

11.2.5 SSH文件传输

应用场景

服务器端开启 SCP 服务，客户端通过 SCP 命令与服务器端进行数据传输，如下图所示。

图 11-5 SSH 文件传输组网图



功能部署

- 服务器端开启 SCP 服务。
- 客户端使用 SCP 命令上传文件至服务器端，或从服务器端下载文件。

11.3 功能详解

基本概念

📄 用户认证机制

- password 认证

password 认证是指客户端向服务器发出用户认证请求，并将加密的用户名和密码发送给服务器；服务器将收到信息进行解密，同时将此信息与设备上保存的客户端用户名和密码进行比较，然后返回用户认证成功或失败的消息。

- public-key 认证

public-key 认证是指利用 RSA 或 DSA 等数字签名算法对客户端进行认证。客户端向服务器发送 public-key 认证请求，包括用户名、公钥和公钥算法等信息。服务器收到该信息后先检查公钥的合法性，如果不合法，则直接发送认证失败；否则，服务器对客户端进行数字签名认证，并返回用户认证成功或失败的消息。

i public-key 认证仅针对客户端版本为 SSH2.0 的用户。

📄 SSH 的通讯交互

在整个服务器端与客户端的 SSH 通讯交互过程中，为了实现安全通道，需要经历如下七个阶段：

- 建立连接阶段

服务器端在端口 22 监听，等待客户端的连接。当客户端发起 Socket 初始连接请求时，客户端与服务器端建立起了 TCP Socket 连接。

- 版本号协商阶段

如果连接成功，服务器端向客户端发送版本协商报文，客户端收到该报文后进行解析，并向服务端回应客户端决定采用的协议版本号。服务器端将对此信息进行分析，协商版本成功或失败。

- 密钥交换与算法协商阶段

如果版本号协商成功，进入密钥交换与算法协商。服务器端与客户端互相对端发送算法协商报文，并根据本端支持的算法来确定最终使用的算法。另外，服务器端与客户端利用密钥交换算法、主机密钥等相关信息，生成会话密钥以及会话 ID，利用它们进行后续的用户认证以及数据传输的加解密。

- 用户认证阶段

在加密通道建立起来之后，进入用户认证阶段。客户端向服务器端发送认证请求，服务器端对客户端进行认证，直到认证成功或者认证次数达到设定的上限，服务器端关闭连接为止。

- 会话请求阶段

认证成功后，客户端向服务器端发送会话请求，服务器等待并处理客户端的请求，请求被处理成功之后，SSH 进入会话交互阶段。

- 会话交互阶段

在会话请求成功之后，进入会话交互阶段。加密的数据在双向就可以进行传送并处理。客户端将需要执行的命令发送给服务器端，服务器接收到该命令后进行解密、解析并处理，并将执行的结果进行加密后发给客户端进行解密处理。

- 会话关闭阶段

在服务器端和客户端结束会话时，断开 socket 连接，会话被关闭。

功能特性

功能特性	作用
SSH Server	网络设备上打开 SSH Server，用户可以通过 SSH 客户端安全地连接网络设备。
SCP 服务	开启 SCP 服务后，用户可以直接下载网络设备上的文件，以及将本地文件上传至网络设备，同时所有交互数据以密文形式进行传输，具有认证和安全性等特性。
SSH Client	用户可以使用设备上的 SSH Client 与网络设备上的 SSH Server 建立安全连接

11.3.1 SSH Server

网络设备上打开 SSH Server，用户可以通过 SSH 客户端安全地连接网络设备；同时，可以关闭 SSH Server，断开全部 SSH 用户的连接。

工作原理

SSH Server 具体工作原理可参考基本概念章节中的“SSH 的通讯交互”小节。在实际应用中，开启 SSH Server 服务，同时可根据相关应用需求设置以下功能点。

设置版本号：使 SSH Server 支持 SSHv1 与 SSHv2 两种客户端的连接。

设置用户认证超时时间：使 SSH Server 从接受用户连接请求开始计时，存在两种情况：用户认证成功或认证超时断开客户端的连接。

设置用户重认证次数：使 SSH Server 从接受用户连接请求开始进行认证，如果达到重认证次数仍未认证成功，则提示认证失败。

公钥认证：公钥算法为 RSA 或 DSA，在客户端与服务端之间提供安全连接。通过将客户端的公钥文件与用户名进行关联，同时，在客户端配置公钥认证方式，而且指定对应的私钥文件。这样，在客户端登录认证时，即可实施公钥认证进行安全连接。

相关配置

SSH Server 开启

缺省情况下，SSH Server 处于关闭状态。

在全局配置模式下，执行 `[no] enable service ssh-server` 命令可以打开或关闭 SSH Server。

同时需要生成 SSH 密钥，使 SSH Server 的状态成为 ENABLE。

设置 SSH Server 支持的版本

缺省情况下，SSH Server 所支持的版本兼容 SSH1 与 SSH2，使用 SSH 1 或者 SSH2 的客户端都可以连接。

使用 `ip ssh version` 命令来配置 SSH Server 所支持的 SSH 连接的协议版本。

如果设置了版本 1 或者 2，只允许对应版本的 SSH 客户端才能够连接。

设置 SSH Server 的用户认证超时时间

缺省情况下，超时时间为 120 秒。

执行 `ip ssh time-out` 命令配置 SSH Server 用户认证的超时时间。执行 `no` 命令可以使超时时间恢复为缺省值。从接受用户连接请求开始计时，当超过指定时间没有认证成功时，则认为认证超时失败。

设置 SSH Server 的重认证次数

缺省情况下，重认证次数为 3 次。

执行 `ip ssh authentication-retries` 命令配置 SSH Server 进行用户认证的重认证次数。执行 `no` 命令可以使重认证次数恢复为缺省值。当超过 SSH Server 配置的重认证次数，仍没有认证成功，则认为用户认证失败。

设置 SSH Server 支持的加密模式

缺省情况下，SSH Server 支持的加密模式为兼容模式（支持 cbc、ctr 和其它类）。

执行 `ip ssh cipher-mode` 命令配置 SSH Server 支持的加密模式。执行 `no` 命令将 SSH Server 支持的加密模式为缺省设置。

设置 SSH Server 支持的消息认证算法

缺省情况下 SSH Server 支持的消息认证算法为：SSHv1 不支持，SSHv2 支持 4 种（MD5、SHA1、SHA1-96、MD5-96）。

执行 `ip ssh hmac-algorithm` 命令配置 SSH Server 支持的消息认证算法。执行 `no` 命令将 SSH Server 支持的消息认证算法恢复为缺省设置。

启动 SSH Server 的公钥认证

执行 `ip ssh peer` 命令开启或取消客户端的公钥文件和用户名关联，客户端登录认证时，通过用户名指定使用的公钥文件。

11.3.2 SCP服务

SSH Server 提供 SCP (Secure Copy , 安全复制) 服务，用于服务器端与客户端之间文件的安全传输。

工作原理

SCP 协议是一个支持网络文件传输的协议。它运行在 22 端口，基于 BSD RCP 协议；而 RCP 又基于 SSH 协议提供加密和认证。其中，RCP 负责文件的传输，而 SSH 协议负责认证和加密。

在服务器端开启 SCP 服务后，当用户使用 SCP 客户端进行文件上传与下载时，SCP 客户端会先解析命令行参数，然后打开一个到远程服务器的连接，再通过这个连接起另一个 SCP 进程。这个进程的运行方式可以是源模式(source)，也可以是宿模式(sink)，（前者是数据提供者；后者是数据的目的地）。前者读取文件并通过 SSH 连接发送到另一端，后者通过 SSH 连接接收文件。

相关配置

启动 SCP 服务

缺省不开启 SCP 服务器功能。

执行 `ip scp server enable` 命令在网络设备上打开或关闭 SCP 服务器功能。

11.4 配置详解

配置项	配置建议 & 相关命令	
配置SSH Server	 SSH Server 打开功能必须配置。	
	<code>enable service ssh-server</code>	配置 SSH Server 打开功能
	<code>disconnect ssh [vty] session-id</code>	断开已经建立的 SSH 连接
	<code>crypto key generate { rsa dsa }</code>	生成密钥
	<code>ip ssh version { 1 2 }</code>	配置 SSH Server 支持版本
	<code>ip ssh time-out time</code>	配置 SSH 用户认证超时时间
	<code>ip ssh authentication-retries retry times</code>	配置 SSH 重认证次数
	<code>ip ssh cipher-mode { cbc ctr others }</code>	配置 SSH Server 支持的加密模式
	<code>ip ssh hmac-algorithm { md5 md5-96 sha1 sha1-96 }</code>	配置 SSH Server 支持的消息认证算法
	<code>ip ssh peer test public-key rsa flash:rsa.pub</code>	设置用户 test 关联的 RSA 公钥文件
<code>ip ssh peer test public-key dsa flash:dsa.pub</code>	设置用户 test 关联的 DSA 公钥文件	

配置SCP服务	 必须配置。	
	<code>ip scp server enable</code>	开启 SCP 服务器功能

11.4.1 配置SSH Server

配置效果

- 在网络设备上打开 SSH Server 功能，用户可以通过 SSH 客户端安全地连接网络设备，同时所有交互信息均以密文形式进行传输，具有认证和安全性等特性。
- 用户可以采用多种认证方式进行 SSH 用户认证，包括：本地线路口令认证、AAA 认证、公钥认证等认证方式。
- 用户可以生成或删除服务器密钥。
- 用户可以配置服务器支持的 SSH 协议版本。
- 用户可以设置认证超时时间。
- 用户可以设置重认证次数。
- 用户可以设置服务器支持的加密模式。
- 用户可以设置服务器支持的消息认证算法。

注意事项

- 配置设备为 SSH Server，前提是必须保证设备所处网络环境通信正常，管理员能够连接到设备的管理界面进行相应地配置。
- 删除密钥时，不存在命令 `no crypto key generate`；而是通过命令 `crypto key zeroize` 命令删除密钥。
- SSH 模块不支持热备，因此在支持管理板热备份的产品中，管理板切换动作发生后，若新的主板上没有 SSH 密钥文件，则必须通过命令 `crypto key generate` 重新生成密钥后方可使用 SSH。

配置方法

▾ 配置 SSH Server 打开功能

- 必须配置。
- 缺省情况下，SSH Server 处于关闭状态。在全局配置模式下，打开 SSH Server，同时需要生成 SSH 密钥，使 SSH Server 的状态成为 ENABLE。

▾ 配置 SSH Server 支持版本

- 可选配置。

- 缺省情况下，SSH Server 兼容版本 1 和 2，使用 SSH 1 或者 SSH2 的客户端都可以连接。如果设置了版本 1 或者 2，只允许对应版本的 SSH 客户端才能够连接。

配置 SSH 用户认证超时时间

- 可选配置。
- 缺省情况下，SSH Server 的用户认证超时时间为 120 秒。可以根据需要配置用户认证的超时时间，取值范围为 1~120s，单位为秒。

配置 SSH 重认证次数

- 可选配置。
- 设置 SSH 用户请求连接的认证重试次数，防止恶意猜测等非法行为。缺省情况下，SSH Server 的重认证次数为 3 次，即可以允许用户尝试三次输入用户名与密码进行认证尝试。可以根据需要配置用户认证的重认证次数，取值范围为 0~5。

配置 SSH Server 支持的加密模式

- 可选配置。
- 设置 SSH Server 支持的加密模式。缺省情况下，SSH Server 支持的加密模式为兼容模式（支持 cbc、ctr 和其它类）。

配置 SSH Server 支持的消息认证算法

- 可选配置。
- 设置 SSH Server 支持的消息认证算法。缺省情况下，SSH Server 支持的消息认证算法为：SSHv1 不支持，SSHv2 支持 4 种（MD5、SHA1、SHA1-96、MD5-96）。

配置 SSH 基于公钥的认证

- 可选配置。
- 根据 SSH 协议，只有 SSHv2 才支持基于公钥的认证，SSHv1 不支持。该配置将客户端的公钥文件和用户名关联，客户端登录认证时，通过用户名指定使用的公钥文件。

检验方法

- 使用 **show ip ssh** 命令，可以查看 SSH Server 当前支持的 SSH 协议版本、用户认证的超时时间及重认证次数。
- 通过执行 **show crypto key mypubkey** 命令，查看密钥的公开部分信息是否存在来确认密钥是否已经生成。
- 在 SSH 客户端配置相应的公钥认证登录方式，并指定对应的私钥文件，观察是否可以正常登录。如果可以正常登录，表示客户端的公钥文件与用户名关联成功，公钥认证通过。

相关命令

配置 SSH Server 打开功能

【命令格式】 **enable service ssh-server**

- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 关闭 SSH Server，需要在全局配置模式下，执行 **no enable service ssh-server** 命令，使 SSH Server 的状态成为 DISABLE。

📌 断开已经建立的 SSH 连接

- 【命令格式】 **disconnect ssh [vty] session-id**
- 【参数说明】 **vtty**：已经建立的 vty 连接
session-id：已经建立的 SSH 连接会话号，取值范围为 0~35
- 【命令模式】 特权用户模式
- 【使用指导】 通过输入指定的 SSH 连接会话号，断开已经建立的 SSH 连接。或者输入指定的 VTY 连接会话号，断开指定的 SSH 连接。只能断开 SSH 类型的连接。

📌 生成密钥

- 【命令格式】 **crypto key generate { rsa | dsa }**
- 【参数说明】 **rsa**：生成 rsa 密钥
dsa：生成 dsa 密钥
- 【命令模式】 全局模式
- 【使用指导】 删除密钥时，不存在命令 **no crypto key generate**；而是通过命令 **crypto key zeroize** 命令删除密钥。根据 SSH 协议，SSH 1 使用 RSA 密钥；SSH 2 使用 RSA 或 DSA 密钥。如果已生成 RSA 密钥，则 SSH1 与 SSH2 都可用；如果仅生成 DSA 密钥，则仅有 SSH2 可以使用。

📌 配置 SSH Server 支持版本

- 【命令格式】 **ip ssh version { 1 | 2 }**
- 【参数说明】 **1**：配置 SSH Server 仅支持 SSH1 的客户端连接请求
2：配置 SSH Server 仅支持 SSH2 的客户端连接请求
- 【命令模式】 全局模式
- 【使用指导】 **no ip ssh version** 命令恢复 SSH 为缺省配置，支持 SSHv1 与 SSHv2。

📌 配置 SSH 用户认证超时时间

- 【命令格式】 **ip ssh time-out time**
- 【参数说明】 **time**：配置用户认证的超时时间，取值范围为 1~120s，单位为秒
- 【命令模式】 全局模式
- 【使用指导】 **no ip ssh time-out** 命令恢复 SSH 的缺省用户认证超时时间为 120 秒。

📌 配置 SSH 重认证次数

- 【命令格式】 **ip ssh authentication-retries retry times**
- 【参数说明】 **retry times**：配置用户认证的重认证次数，取值范围为 0~5
- 【命令模式】 全局模式
- 【使用指导】 **no ip ssh authentication-retries** 命令恢复 SSH 的重认证次数为 3 次。

📌 配置 SSH Server 支持的加密模式

【命令格式】 **ip ssh cipher-mode { cbc | ctr | others }**

【参数说明】 **cbc** : 配置 SSH Server 支持的加密模式为 cbc 模式 (密码分组链接模式 : cipher block chaining mode) , 其对应的加密算法为 : DES-CBC、3DES-CBC、AES-128-CBC、AES-192-CBC、AES-256-CBC、Blowfish-CBC
ctr : 配置 SSH Server 支持的加密模式为 ctr 模式 (计数器模式 : counter mode) , 其对应的算法为 : AES128-CTR、AES192-CTR、AES256-CTR
others : 配置 SSH Server 支持的加密模式为 others , 其对应的加密算法为 : RC4

【命令模式】 全局模式

【使用指导】 配置 SSH Server 支持的加密模式。

在锐捷设备当中, SSHv1 Server 支持的加密算法为 : DES-CBC、3DES-CBC、Blowfish-CBC ; SSHv2 Server 支持的加密算法为 : AES128-CTR、AES192-CTR、AES256-CTR、DES-CBC、3DES-CBC、AES-128-CBC、AES-192-CBC、AES-256-CBC、Blowfish-CBC、RC4。在这些加密算法当中, 可以归为三类不同的加密模式 : “cbc” (密码分组链接模式 : cipher block chaining mode) 、 “ctr” (计数器模式 : counter mode) 和 others (其它类) 。

随着密码学的不断发展, “cbc” 和 “others” 加密模式的加密算法, 都被证明可以在有限时间内进行破解。所以, 在一些安全性要求比较高的组织或单位, 可以通过配置 SSH Server 支持的加密模式为 “ctr” 模式, 来提高 SSH Server 的安全级别。

▾ 配置 SSH Server 支持的消息认证算法

【命令格式】 **ip ssh hmac-algorithm { md5 | md5-96 | sha1 | sha1-96 }**

【参数说明】 **md5** : 配置 SSH Server 支持的消息认证算法为 md5
md5-96 : 配置 SSH Server 支持的消息认证算法为 md5-96
sha1 : 配置 SSH Server 支持的消息认证算法为 sha1
sha1-96 : 配置 SSH Server 支持的消息认证算法为 sha1-96

【命令模式】 全局模式

【使用指导】 配置 SSH Server 支持的消息认证算法。

在锐捷设备当中, SSHv1 Server 不支持消息认证算法; SSHv2 Server 支持的消息认证算法为 : md5、sha1、sha1-96、md5-96。用户可以依据实际需要选择 SSH Server 支持的消息认证算法。

▾ 配置 SSH 基于 rsa 公钥的认证

【命令格式】 **ip ssh peer test public-key rsa flash:rsa.pub**

【参数说明】 **test** : 用户名
rsa : public-key 类型为 rsa
rsa.pub : 公钥文件名

【命令模式】 全局模式

【使用指导】 设置用户 test 关联的 RSA 公钥文件。

根据 SSH 协议, 只有 SSHv2 才支持基于公钥的认证, SSHv1 不支持。该命令将客户端的公钥文件和用户名关联, 客户端登录认证时, 通过用户名指定使用的公钥文件。

▾ 配置 SSH 基于 dsa 公钥的认证

【命令格式】 **ip ssh peer test public-key dsa flash:dsa.pub**

- 【参数说明】 *test* : 用户名
dsa : public-key 类型为 dsa
dsa.pub : 公钥文件名
- 【命令模式】 全局模式
- 【使用指导】 设置用户 *test* 关联的 DSA 公钥文件。
 根据 SSH 协议，只有 SSHv2 才支持基于公钥的认证，SSHv1 不支持。该命令将客户端的公钥文件和用户名关联，客户端登录认证时，通过用户名指定使用的公钥文件。

配置举例

 以下配置举例，仅介绍与 SSH 相关的配置。

生成服务器端的公共密钥

- 【配置方法】 ● 使用 **crypto key generate { rsa | dsa }** 命令生成服务器公共密钥，以 *rsa* 为例如下。

SSH Server

```
Ruijie# configure terminal
Ruijie(config)# crypto key generate rsa
Choose the size of the rsa key modulus in the range of 512 to 2048
and the size of the dsa key modulus in the range of 360 to 2048 for your
Signature Keys. Choosing a key modulus greater than 512 may take
a few minutes.

How many bits in the modulus [512]:
```

- **rsa 密钥生成成功提示信息：**

```
% Generating 512 bit RSA1 keys ... [ok]
% Generating 512 bit RSA keys ... [ok]
```

- **rsa 密钥生成失败提示信息：**

```
% Generating 512 bit RSA1 keys ... [fail]
% Generating 512 bit RSA keys ... [fail]
```

- 【检验方法】 ● 使用 **show crypto key mypubkey rsa** 命令，通过查看 *rsa* 密钥的公开部分信息是否存在来确认 *rsa* 是否已经生成。

SSH Server

```
Ruijie(config)#show crypto key mypubkey rsa
% Key pair was generated at: 1:49:47 UTC Jan 4 2013
Key name: RSA1 private
Usage: SSH Purpose Key
Key is not exportable.
Key Data:
AAAAAwEA AQAAAHJM 6izXtlpp rUSOEGZ/ UhFpRRrW nngP4BU7 mG836apf jajSYwcU
803LojHL ayJ8G4pG 7j4T4ZSf FKg09kfr 92JpRNHQ gbwaPc5/ 9UnTtX9t qFIKdj1j
0dKBcCfN tr0r/CT+ cs5t1GKV S0ICGifz oB+pYaE=
```

```
% Key pair was generated at: 1:49:47 UTC Jan 4 2013
Key name: RSA private
Usage: SSH Purpose Key
Key is not exportable.
Key Data:
AAAAAwEA AQAAAHJf LwKnz0gO F3R1KhTN /7PmQYoE v0a2VVTX 8ZCa7S1l EghLDLJc
w3T5JQXk Rr3iBD5s b1Ee0L4b 2lykZt/u UetQ0Q80 sISgIfZ9 8o5No3Zz MPM0LnQR
G4c7/28+ GOHzYkTk 4IiQuTIL HRgtbyEY XCFaaxU=
```

设置 SSH Server 的版本

- 【配置方法】 ● 使用 `ip ssh version { 1 | 2 }` 命令设置 SSH Server 的版本，以只使用版本 2 为例。

```
SSH Server Ruijie# configure terminal
Ruijie(config)# ip ssh version 2
```

- 【检验方法】 ● 使用 `show ip ssh` 命令，可以查看 SSH Server 当前支持的 SSH 协议版本。

```
SSH Server Ruijie(config)#show ip ssh
SSH Enable - version 2.0
Authentication timeout: 120 secs
Authentication retries: 3
SSH SCP Server: disabled
```

设置 SSH Server 的用户认证超时时间

- 【配置方法】 ● 使用 `ip ssh time-out time` 命令设置用户认证超时时间，以配置 100s 为例。

```
SSH Server Ruijie# configure terminal
Ruijie(config)# ip ssh time-out 100
```

- 【检验方法】 ● 使用 `show ip ssh` 命令，可以查看 SSH Server 用户认证的超时时间配置信息。

```
SSH Server Ruijie(config)#show ip ssh
SSH Enable - version 2.0
Authentication timeout: 100 secs
Authentication retries: 3
SSH SCP Server: disabled
```

设置 SSH Server 的用户重认证次数

- 【配置方法】 ● 使用 `ip ssh authentication-retries retry times` 命令设置 SSH Server 用户认证的重认证次数，以 2 次为例。

```
SSH Server Ruijie# configure terminal
Ruijie(config)# ip ssh authentication-retries 2
```

- 【检验方法】 ● 使用 `show ip ssh` 命令，可以查看 SSH Server 用户认证的重认证次数配置信息。

```
SSH Server Ruijie(config)#show ip ssh
SSH Enable - version 2.0
Authentication timeout: 100 secs
Authentication retries: 3
SSH SCP Server: disabled
```

设置 SSH Server 支持的加密模式

- 【配置方法】
- 使用 `ip ssh cipher-mode { cbc | ctr | others }` 命令来设置 SSH Server 支持的加密模式。以 ctr 加密模式为例，如下：

```
SSH Server Ruijie# configure terminal
Ruijie(config)# ip ssh cipher-mode ctr
```

- 【检验方法】
- 在 SSH 客户端也选择 ctr 相关加密模式，观察是否可以正常登录。

设置 SSH Server 支持的消息认证算法

- 【配置方法】
- 使用 `ip ssh hmac-algorithm { md5 | md5-96 | sha1 | sha1-96 }` 命令来设置 SSH Server 支持的消息认证算法。以 sha1 算法为例，如下：

```
SSH Server Ruijie# configure terminal
Ruijie(config)# ip ssh hmac-algorithm sha1
```

- 【检验方法】
- 在 SSH 客户端也选择 sha1 消息认证算法，观察是否可以正常登录。

公钥认证

- 【配置方法】
- 使用 `ip ssh peer username public-key { rsa | dsa } filename` 命令，将客户端的公钥文件和用户名关联，客户端登录认证时，通过用户名指定使用的公钥文件。以 rsa 为例如下：

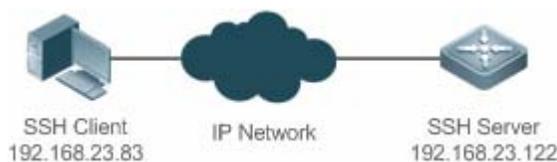
```
SSH Server Ruijie# configure terminal
Ruijie(config)# ip ssh peer test public-key rsa flash:rsa.pub
```

- 【检验方法】
- 在 SSH 客户端配置相应的公钥认证登录方式，并指定对应的私钥文件，观察是否可以正常登录。如果可以正常登录，表示客户端的公钥文件与用户名关联成功，公钥认证通过。

配置 SSH 设备管理

【网络环境】

图 11-6

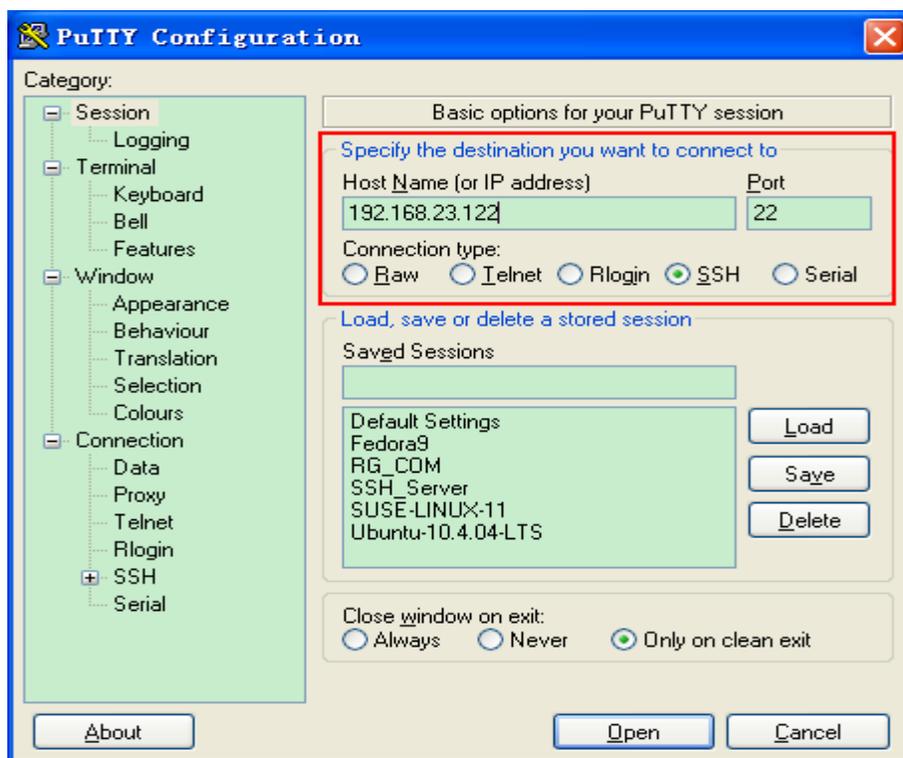


用户可以使用 SSH 对设备进行管理，前提是必须打开 SSH Server 功能，默认情况下是关闭该功能的。由于 Windows 自带的 Telnet 组件不支持 SSH，因此必须使用第三方客户端软件，当前兼容性较好的客户端包括：Putty，Linux，SecureCRT。以客户端软件 Putty 为例介绍 SSH Client 的配置。

- 【配置方法】
- 打开 Putty 客户端工具软件。

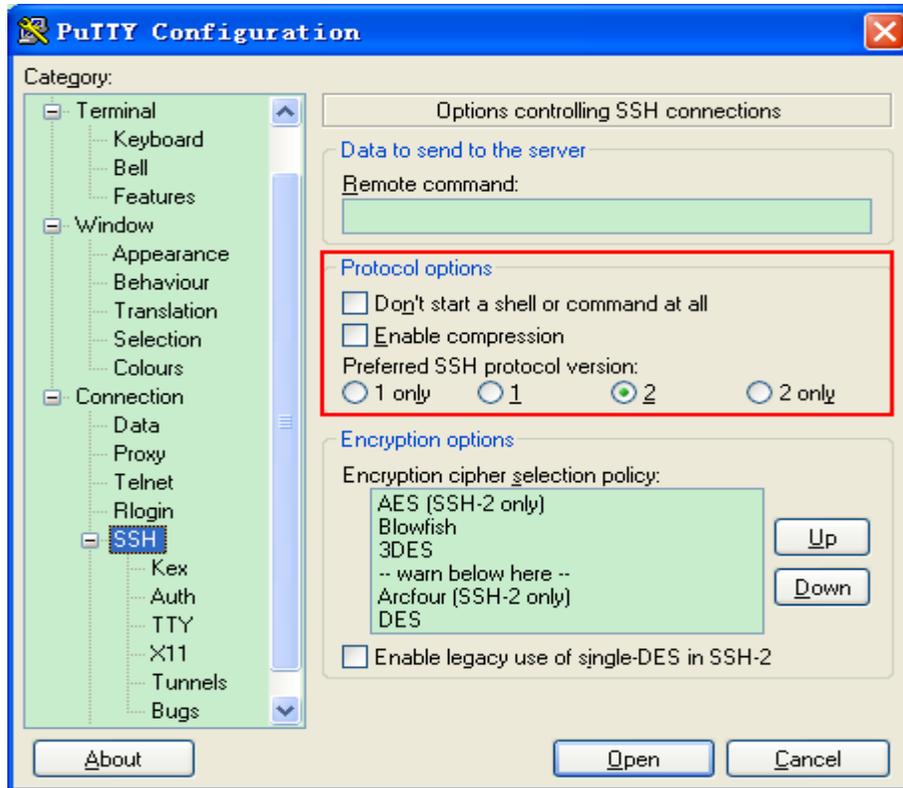
- 在 PuTTY 中的 Session 选项卡中填写 SSH Server 的主机 IP (192.168.23.122)、SSH 端口号 22 以及连接类型为 SSH。
- 在 PuTTY 中的 SSH 选项卡中选择 SSH 协议版本号为 2。
- 在 PuTTY 中的 SSH 选项卡中选择认证方式为 “Keyboard-interactive” 。
- 点击 open 按钮连接服务器主机。
- 在用户名密码认证窗口输入正确的用户名与密码进入终端登录界面。

SSH Client 图 11-7



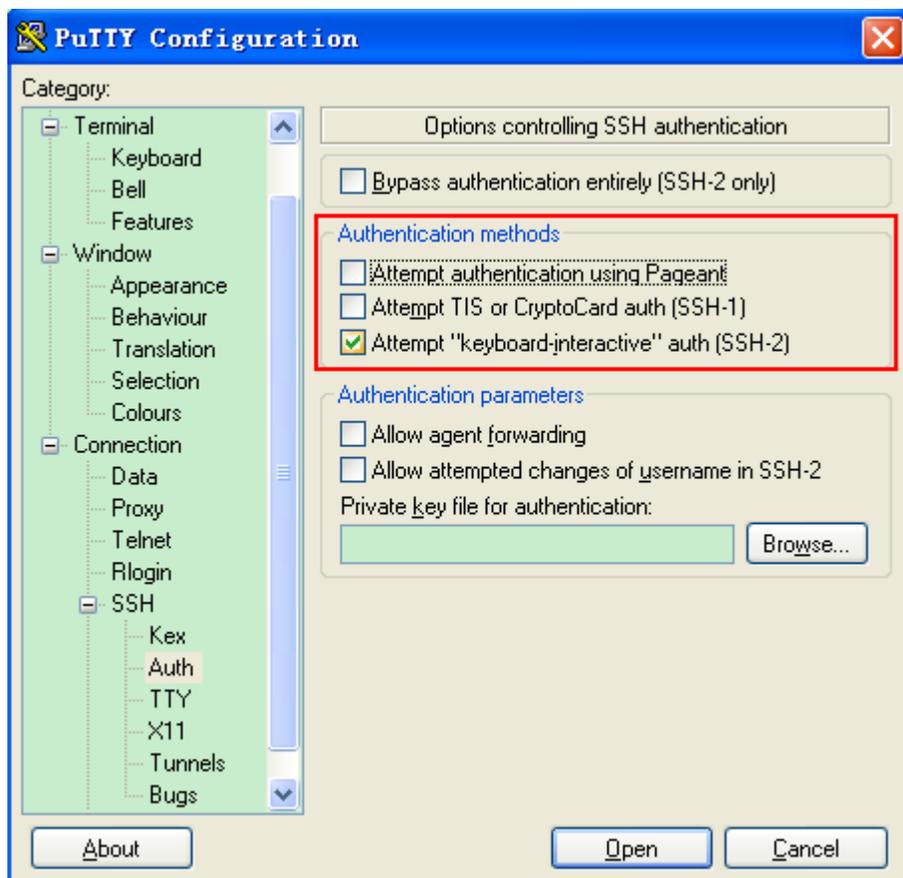
Host Name(or IP address) 就是要登陆的主机的 IP 地址 ,这里为 192.168.23.122。Port 为端口号 22 ,即 SSH 监听的默认端口号。Connection type 为连接类型 SSH。

图 11-8



如上图，使用 SSH 协议 2 进行登陆，因此在 Protocol options 中 Preferred SSH protocol version 选择 2。

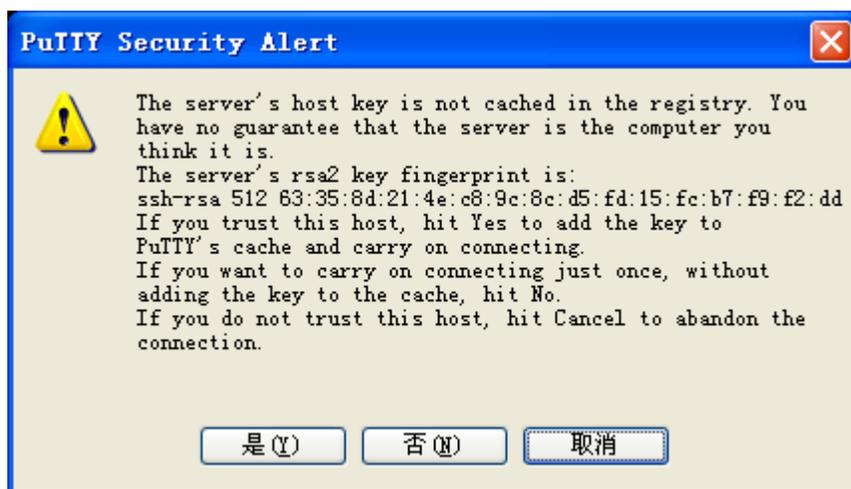
图 11-9



如上图，配置 SSH 的认证方式，在 Authentication methods 复选框中，我们选择 “keyboard-interactive” ，支持用户名密码的认证方式。

然后，点击 open 按钮，连接刚配置的服务器主机，如下图所示：

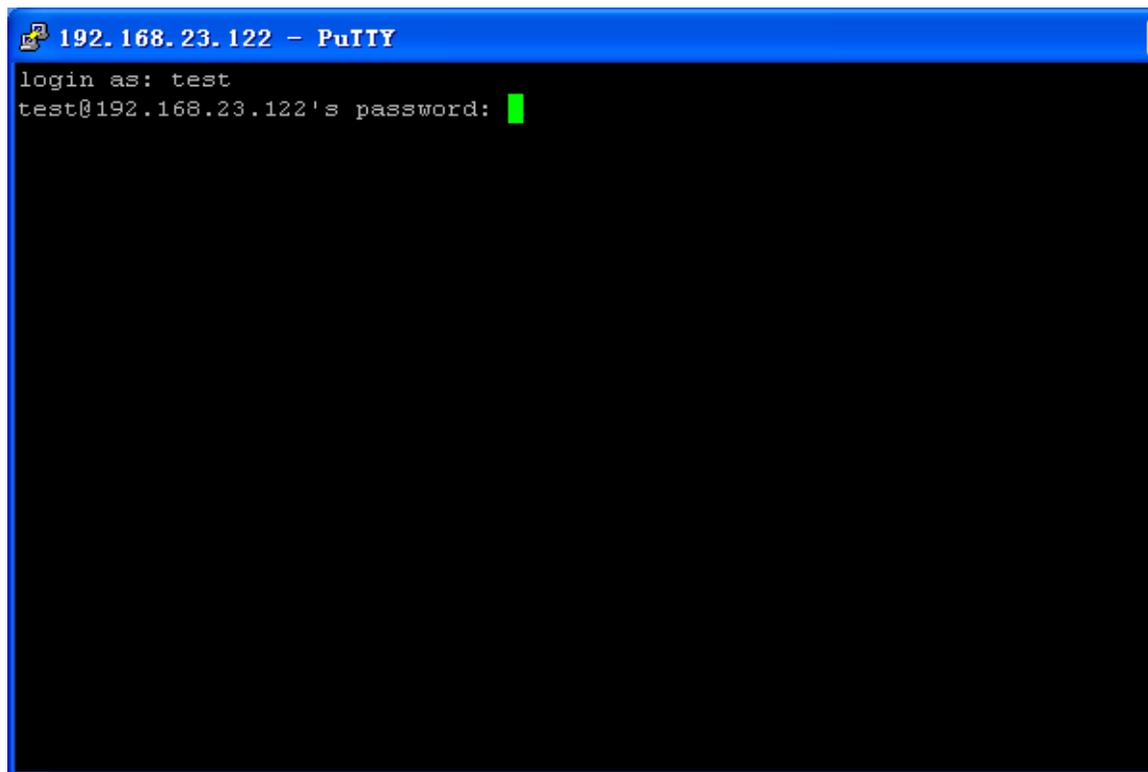
图 11-10



这里询问正在登陆服务器主机 192.168.23.122 的客户端，是否接收服务端发送过来的密钥，选择 “是”（接受而且保存），选择 “否”（只接受一次），选择 “取消”（放弃连接）。

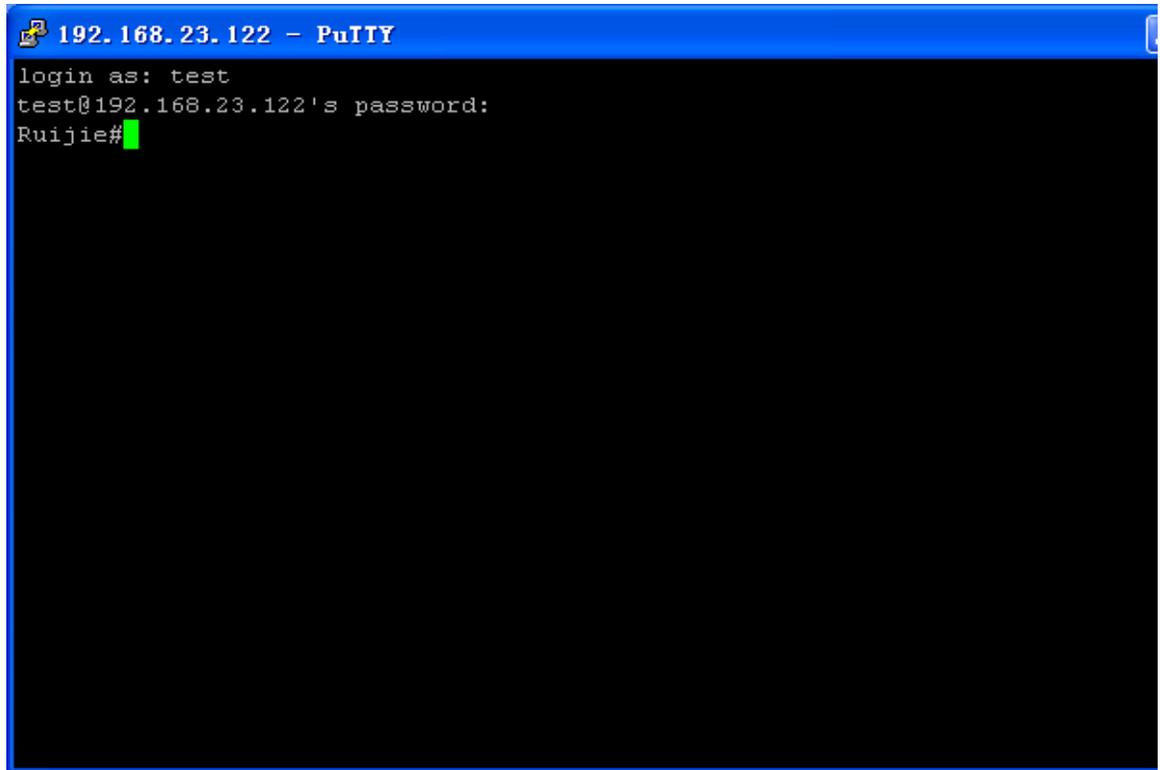
如果选择 “是”，接着，会出现下面的用户名密码登录认证对话框，如下图所示：

图 11-11



此时，输入正确的用户名和密码，即可登录 SSH 终端界面，如下图所示：

图 11-12



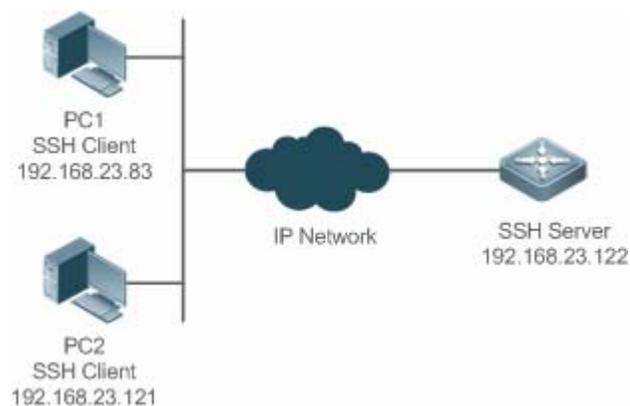
- 【检验方法】
- 通过 **show ip ssh** 命令来显示 SSH Server 当前生效的配置信息。
 - 通过 **show ssh** 命令来显示已经建立的 SSH 连接的每个连接信息。

```
Ruijie#show ip ssh
SSH Enable - version 1.99
Authentication timeout: 120 secs
Authentication retries: 3
Ruijie#show ssh
Connection Version Encryption      Hmac      State      Username
-----
0      2.0 aes256-cbc      hmac-sha1  Session started test
```

配置 SSH 本地线路认证

【网络环境】

图 11-13



SSH 用户可以采用本地线路口令认证方式进行用户认证，如图所示。为了保证数据信息交换的安全，PC1、PC2 作为 SSH 客户端，采用 SSH 协议登录到打开 SSH Server 的网络设备上。具体要求如下：

- SSH 用户采用的认证方式为线路口令认证。
- 同时启用 0-4 这五条线路，其中线路 0 的登录口令为“passzero”，其余四条线路的登录口令均为“pass”，用户名任意。

【配置方法】 SSH Server 的配置要点如下：

- 全局打开 SSH Server。SSH Server 默认支持 SSH1 和 SSH2 两个版本。
- 配置密钥。通过该密钥，SSH 服务器将从 SSH 客户端收到的口令密文进行解密，将解密后的明文同服务器上保存的口令进行比较，并返回认证成功或失败的消息。SSH 1 使用 RSA 密钥；SSH 2 使用 RSA 或者 DSA 密钥。
- 配置 SSH 服务器 FastEthernet 0/1 接口的 IP 地址。SSH 客户端通过该地址连接 SSH 服务器。SSH 客户端至 SSH 服务器路由可达。

SSH Client 的配置要点如下：

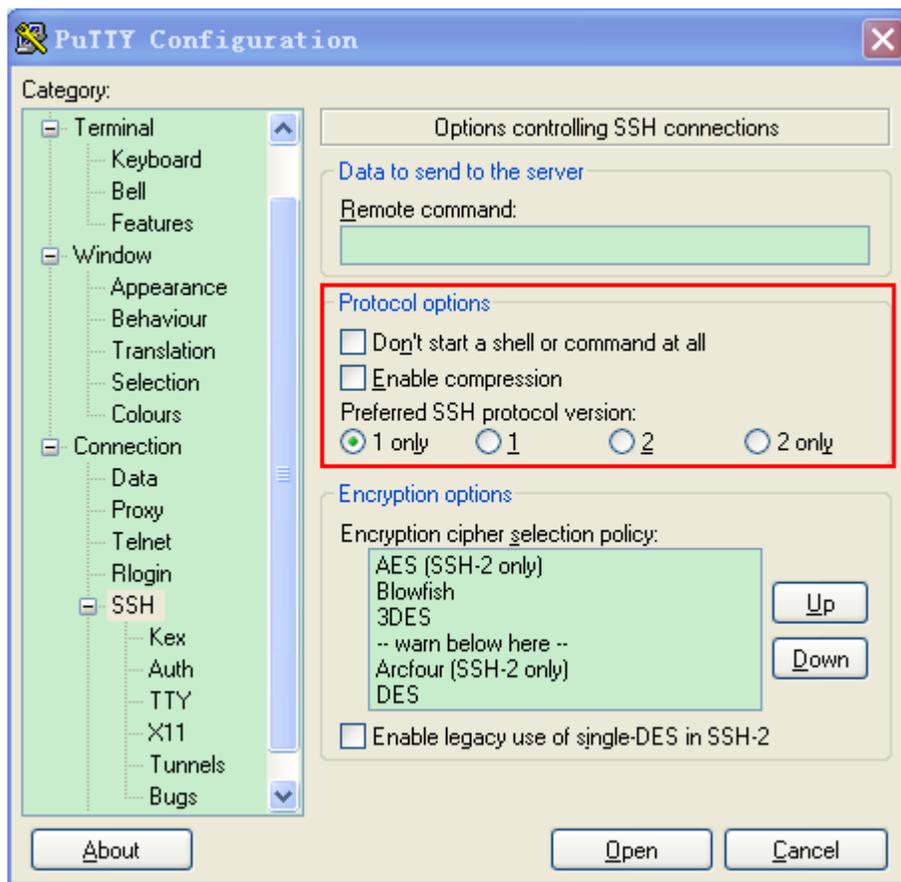
- SSH 客户端软件有多种，例如 Putty、Linux、OpenSSH 等，本文中仅以客户端软件 Putty 为例，说明 SSH 客户端的配置方法。具体配置方法请参见“配置步骤”。

SSH Server 配置 SSH 相关功能之前，请先确保 SSH 用户到 SSH 服务器所在网段的路由可达。接口 IP 配置如拓扑图所示。具体 IP 及路由配置过程此处省略。

```
Ruijie(config)# enable service ssh-server
Ruijie(config)#crypto key generate rsa
% You already have RSA keys.
% Do you really want to replace them? [yes/no]:
Choose the size of the key modulus in the range of 360 to 2048 for your
Signature Keys. Choosing a key modulus greater than 512 may take
a few minutes.
How many bits in the modulus [512]:
% Generating 512 bit RSA1 keys ... [ok]
% Generating 512 bit RSA keys ... [ok]
Ruijie(config)#interface fastEthernet 0/1
Ruijie(config-if- fastEthernet 0/1)#ip address 192.168.23.122 255.255.255.0
Ruijie(config-if- fastEthernet 0/1)#exit
Ruijie(config)#line vty 0
Ruijie(config-line)#password passzero
Ruijie(config-line)#privilege level 15
Ruijie(config-line)#login
Ruijie(config-line)#exit
Ruijie(config)#line vty 1 4
Ruijie(config-line)#password pass
Ruijie(config-line)#privilege level 15
Ruijie(config-line)#login
Ruijie(config-line)#exit
```

SSH Client
(PC1/PC2)

图 11-14



然后，设置 SSH 服务器的 IP 地址与连接端口号，由组网拓扑图可知，服务器主机 IP 为 192.168.23.122，连接端口号为 22（具体设置方式可参考《SSH 设备管理配置举例》），点击 open 按钮进行连接。由于当前认证方式不需要用户名，此处“用户名”可以任意输入，但是不能为空（本例用户名设置为 anyone）。

【检验方法】

- 通过 **show running-config** 命令来查看当前配置信息的正确性
- 验证 SSH Client 的配置

SSH Server

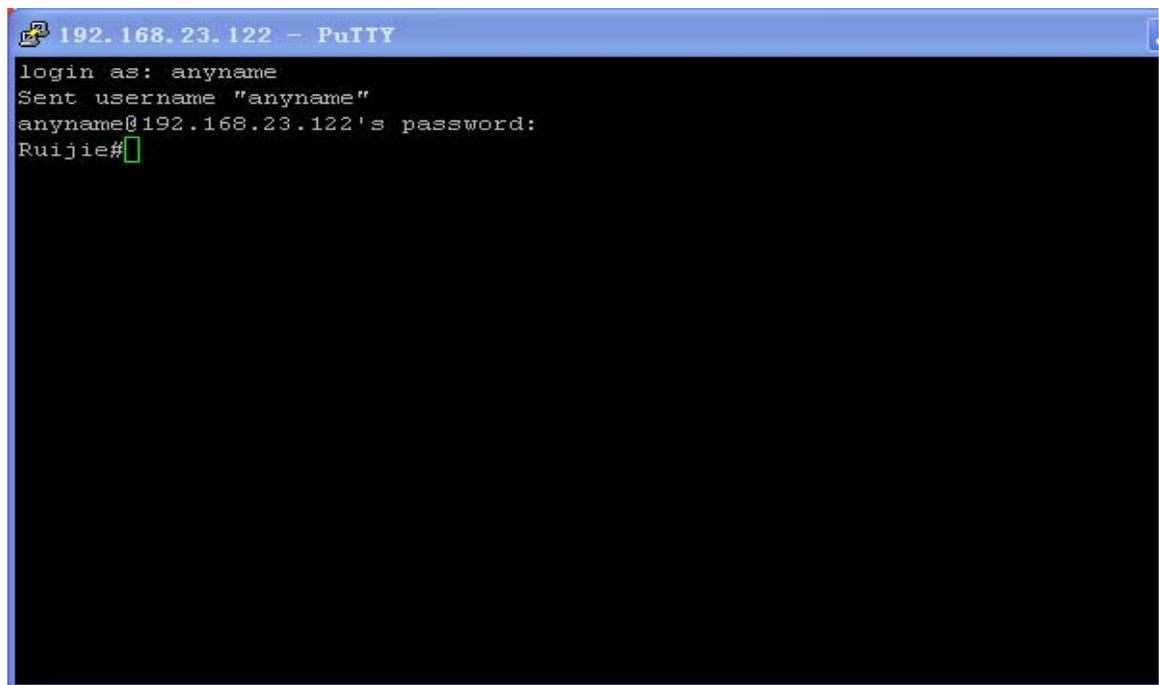
```
Ruijie#show running-config
Building configuration...
!
enable secret 5 $1$eyy2$xs28FDw4s2q0tx97
enable service ssh-server
!
interface fastEthernet 0/1
 ip address 192.168.23.122 255.255.255.0
!
line vty 0
 privilege level 15
```

```
login
password passzero
line vty 1 4
privilege level 15
login
password pass
!
end
```

SSH Client

建立连接，输入正确的口令。线路 0 的登录口令为 “passzero”，其余四条线路的登录口令均为 “pass”，即可进入 SSH Server 的操作界面。如图所示：

图 11-15



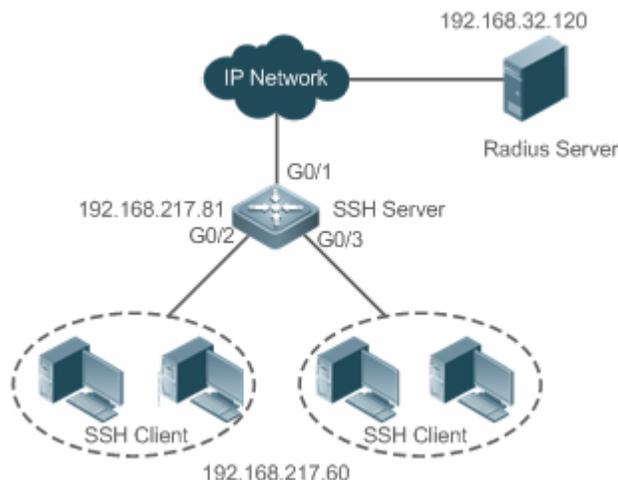
```
Ruijie#show users
```

Line	User	Host(s)	Idle	Location
* 0 con 0	---	idle	00:00:00	---
1 vty 0	---	idle	00:08:02	192.168.23.83
2 vty 1	---	idle	00:00:58	192.168.23.121

配置 SSH 的 AAA 认证

【网络环境】

图 11-16



SSH 用户可以采用 AAA 认证方式进行用户认证，如图所示。为了保证数据信息交换的安全，PC 作为 SSH 客户端，采用 SSH 协议登录到打开 SSH Server 的网络设备上。为了更好地进行安全管理，SSH 客户端登录用户界面采用 AAA 认证方式；同时出于稳定性方面考虑，在 AAA 认证方法列表中配置两种认证方法：Radius 服务器认证和本地认证。优先选择 Radius 服务器，当 Radius 服务器没有响应时选择本地认证方法。

【配置方法】

- SSH 客户端到 SSH 服务器端的路由可达，SSH 服务器到 Radius 服务器端的路由可达。
- 在网络设备上进行 SSH Server 相关配置。配置要点在上一个例子中已有描述，不再重复说明。
- 在网络设备上进行 AAA 认证相关配置。AAA 通过创建方法列表来定义身份认证、类型，然后将这些方法列表应用于特定的服务或接口上。

SSH Server

```
Ruijie(config)# enable service ssh-server
Ruijie(config)#crypto key generate rsa
% You already have RSA keys.
% Do you really want to replace them? [yes/no]:
Choose the size of the key modulus in the range of 360 to 2048 for your
Signature Keys. Choosing a key modulus greater than 512 may take
a few minutes.
How many bits in the modulus [512]:
% Generating 512 bit RSA1 keys ... [ok]
% Generating 512 bit RSA keys ... [ok]
Ruijie(config)#crypto key generate dsa
Choose the size of the key modulus in the range of 360 to 2048 for your
Signature Keys. Choosing a key modulus greater than 512 may take
a few minutes.
How many bits in the modulus [512]:
% Generating 512 bit DSA keys ... [ok]
Ruijie(config)#interface gigabitEthernet 1/1
Ruijie(config-if-gigabitEthernet 1/1)#ip address 192.168.217.81 255.255.255.0
Ruijie(config-if-gigabitEthernet 1/1)#exit
Ruijie#configure terminal
```

```
Ruijie(config)#aaa new-model
Ruijie(config)#radius-server host 192.168.32.120
Ruijie(config)#radius-server key aaradius
Ruijie(config)#aaa authentication login method group radius local
Ruijie(config)#line vty 0 4
Ruijie(config-line)#login authentication method
Ruijie(config-line)#exit
Ruijie(config)#username user1 privilege 1 password 111
Ruijie(config)#username user2 privilege 10 password 222
Ruijie(config)#username user3 privilege 15 password 333
Ruijie(config)#enable secret w
```

【检验方法】

- 通过 **show running-config** 命令来查看当前配置信息的正确性
- 在 Radius Server 上的设置。本例以 SAM 服务器为例进行说明。
- 在 PC 机上建立远程 SSH 连接。
- 查看登录用户。

```
Ruijie#show run
aaa new-model
!
aaa authentication login method group radius local
!
username user1 password 111
username user2 password 222
username user2 privilege 10
username user3 password 333
username user3 privilege 15
no service password-encryption
!
radius-server host 192.168.32.120
radius-server key aaradius
enable secret 5 $1$hbz$ArCsyqy6yyzpz03
enable service ssh-server
!
interface gigabitEthernet 1/1
 no ip proxy-arp
 ip address 192.168.217.81 255.255.255.0
!
ip route 0.0.0.0 0.0.0.0 192.168.217.1
!
line con 0
line vty 0 4
```

```
login authentication method
!
End
```

【系统管理】-【设备管理】中，添加设备 IP 地址：192.168.217.81，添加设备 Key：aaadius

【安全管理】-【设备管理权限】中，设置登录用户的权限。

【安全管理】-【设备管理员】中，添加用户名：user；口令：pass。

SSH 客户端软件设置、建立连接；SSH 客户端的创建方法请参见上例。

输入正确的口令，SSH 用户名为：user；口令为：pass。登录成功。

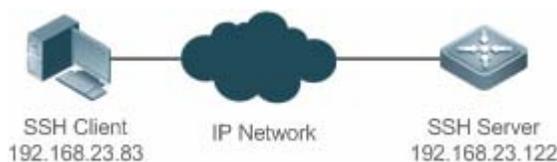
```
Ruijie#show users
```

Line	User	Host(s)	Idle	Location
0 con 0		idle	00:00:31	
* 1 vty 0	user	idle	00:00:33	192.168.217.60

配置 SSH 公钥认证

【网络环境】

图 11-17



SSH 用户可以采用 Public-key 认证方式进行用户认证，公钥算法为 RSA 或 DSA，如图所示。通过配置客户端使用 SSH 协议与服务器端进行安全连接。

【配置方法】

- 客户端公钥认证方式，首先要在客户端生成一个密钥对（这里以 RSA 密钥对为例），然后将其中的公钥放置在 SSH 服务器上，并选择使用 Public-Key 认证方式。

i 在客户端生成密钥对之后，需要将保存的公钥文件上传至服务器端，同时完成服务器的相关配置之后，才可以继续进行客户端的配置以及客户端与服务器端的连接。

- 在客户端生成了密钥以后，SSH 服务器端需要将客户端的公钥文件复制到 flash 中，并且与 SSH 用户名关联。每个用户可以关联一个 RSA 公钥和一个 DSA 公钥。

SSH Client

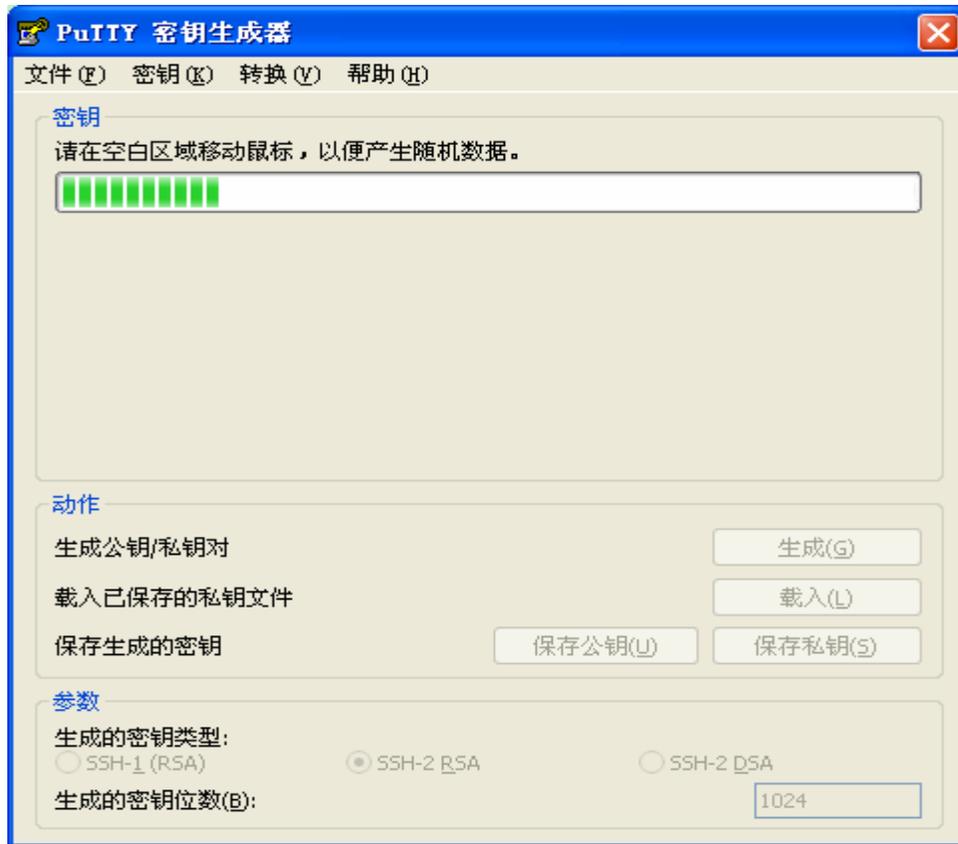
在客户端运行 puttygen.exe 软件，在参数选项栏中选择“SSH-2 RSA”，单击“生成”按钮产生密钥。如下图所示：

图 11-18



生成密钥的时候要除了绿色进度条外的地方不断晃动鼠标,否则进度条显示不动,密钥产生停止,如下图所示:

图 11-19



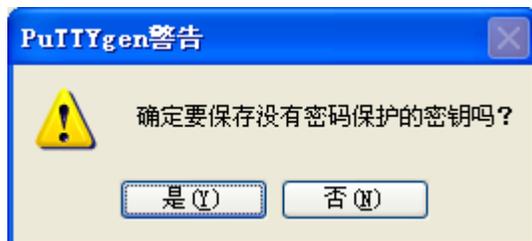
为了保证 RSA 公钥认证的安全性，生成 RSA 密钥对时，RSA 密钥对的长度必须大于或等于 768 位。这里设置为 1024：

图 11-20



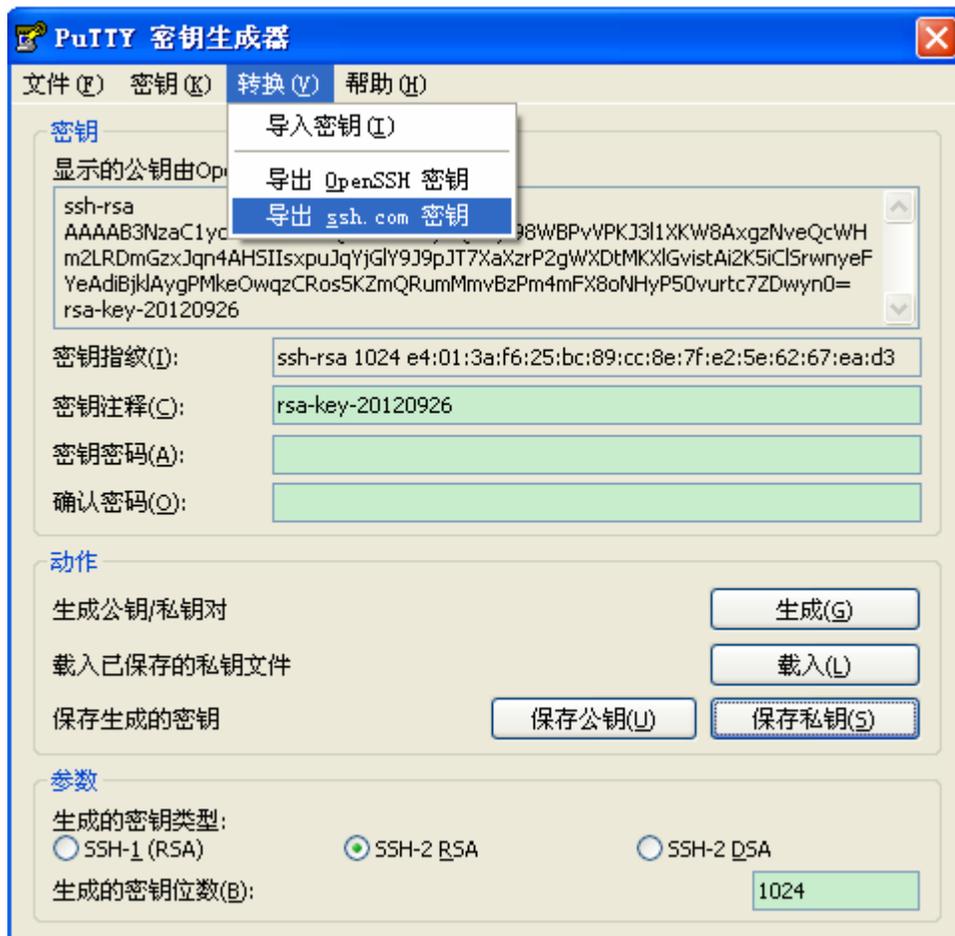
密钥对产生之后，点击“保存公钥”按钮，输入公钥名“test_key.pub”，选择保存路径，点击保存；点击“保存私钥”按钮，弹出如下图所示的警告，选择“是”，输入私钥文件名“test_private”，并点击保存。

图 11-21



一定要选择使用 OpenSSH 格式的密钥文件，否则不能使用。puttygen.exe 能生成 OpenSSH 格式的密钥对，但是 Putty 客户端却不能直接使用，还需要使用 puttygen.exe 工具把私钥转换成 Putty 格式。放在服务器上的公钥文件不需要转换，还是 OpenSSH 格式。如下图所示。

图 11-22

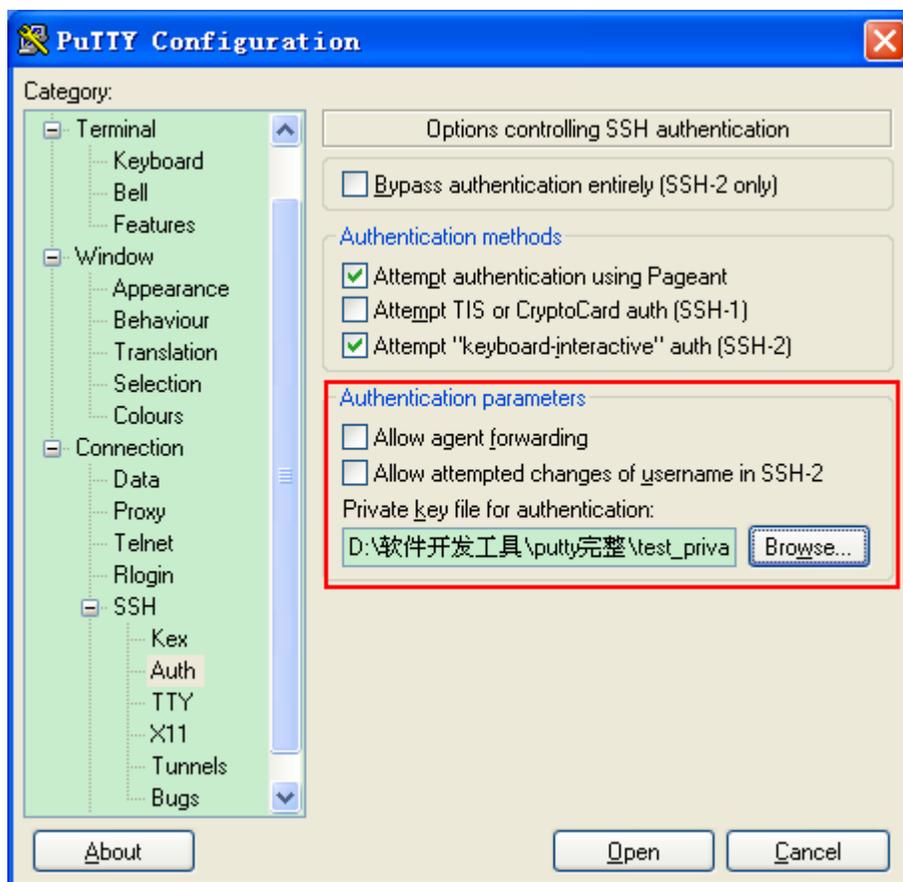
**SSH Server**

```
Ruijie# configure terminal
Ruijie(config)# ip ssh peer test public-key rsa flash:test_key.pub
```

【检验方法】

- 客户端与服务器端的基本配置完成之后，在 Putty 客户端中指定私钥文件 test_private，并设置服务器主机 IP 为 192.168.23.122，端口号为 22，建立客户端与服务器端的连接，这样，客户端就可以使用公钥认证方式登录网络设备了。

图 11-23



常见错误

- 使用命令 `no crypto key generate` 删除密钥。

11.4.2 配置SCP服务

配置效果

在网络设备上打开 SCP 服务器功能，用户可以直接对网络设备上的文件进行下载，以及将本地文件上传至网络设备，同时所有交互数据以密文形式进行传输，具有认证和安全性等特性。

注意事项

- SSH Server 已经完成配置。

配置方法

配置 SCP 服务功能

- 必须配置。
- 缺省情况下，SCP 功能处于关闭状态。在全局配置模式下，通过 **ip scp server enable** 命令开启 SCP 功能。

检验方法

使用 **show ip ssh** 命令，可以查看 SCP 服务器功能是否打开。

相关命令

配置 SCP 服务功能

- 【命令格式】 **ip scp server enable**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 该命令打开开启 SCP 服务器功能。
no ip scp server enable 命令关闭 SCP 服务器功能。

配置举例

开启 scp 功能

- 【配置方法】 ● 使用 **ip scp server enable** 开启 SCP 服务器功能。

```
Ruijie# configure terminal
Ruijie(config)# ip scp server enable
```

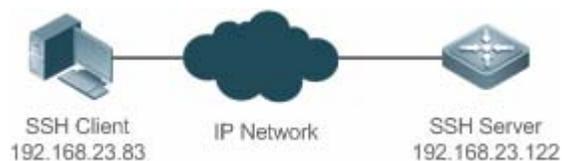
- 【检验方法】 ● 使用 **show ip ssh** 命令，可以查看 SCP 服务器功能是否打开。

```
Ruijie(config)#show ip ssh
SSH Enable - version 1.99
Authentication timeout: 120 secs
Authentication retries: 3
SSH SCP Server: enabled
```

配置 SSH 文件传输

- 【网络环境】

图 11-24



服务器端开启 SCP 服务，客户端通过 SCP 命令与服务器端进行数据传输。

- 【配置方法】 ● 服务器端开启 SCP 服务。

- ❗ SCP 服务器使用的是 SSH 线程，客户端连接网络设备进行 SCP 传输时候会占用一个 VTY 连接（通过 show user 命令查看的时候，会发现用户类型为 SSH）。

- 客户端使用 SCP 命令上传文件至服务器端，或从服务器端下载文件。

SCP 命令的语法：

```
scp [-l246BCpqr] [-c cipher] [-F ssh_config] [-i identity_file]
    [-l limit] [-o ssh_option] [-P port] [-S program]
    [[user@]host1:]file1 [...] [[user@]host2:]file2
```

部分选项说明：

- l : 使用 SSH1 版本（若不指定则默认使用 SSH2）；
- 2 : 使用 SSH2 版本（默认）；
- C : 指定使用压缩传输；
- c : 指定使用的加密算法；
- r : 指定传输整个目录；
- i : 指定使用的密钥文件；
- l : 限制传输速度（单位 Kbits）；

其他具体的参数可以查看 scp.0 文件。

- ❗ 选项大部分与客户端有关，少数是客户端和服务器都需要支持的选项，锐捷网络设备上的 SCP 服务器端不支持 -d -p -q -r 选项，使用这些选项时候会提示不支持。

SSH Server

```
Ruijie# configure terminal
Ruijie(config)# ip scp server enable
```

【检验方法】

- 文件传输举例，以在 Ubuntu 7.10 系统上操作为例：

指定用户名是 test，从 IP 为 192.168.23.122 的网络设备上，将 config.text 文件复制到本地的 /root 目录下。

```
root@dhcpd:~# scp test@192.168.23.122:/config.text /root/config.text
test@192.168.195.188's password:
config.text          100% 1506    1.5KB/s   00:00
Read from remote host 192.168.195.188: Connection reset by peer
```

常见配置错误

无

11.5 监视与维护

查看运行情况

作用	命令
显示 SSH Server 的当前生效的配置信息	show ip ssh

显示已经建立的 SSH 连接的每个连接信息	<code>show ssh</code>
显示 SSH Server 公共密钥的公开密钥部分的信息	<code>show crypto key mypubkey</code>

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 SSH 基本连接信息的调试开关	<code>debug ssh</code>

12 URPF

12.1 概述

URPF (Unicast Reverse Path Forwarding , 单播反向路径转发) 是一种阻断源地址欺骗攻击 , 保护网络的方法。

URPF 通过获取接收报文的源地址和入接口 , 根据源地址在转发表中查找转发表项 , 如果表项不存在 , 那么报文丢弃 ; 如果转发表项的出接口与报文的入接口不匹配 , 那么报文丢弃。否则 , 认为匹配 , 允许报文转发。

URPF 有两种模式 : 严格模式、宽松模式。

- 严格模式通常部署在点对点的接口上 , 并且数据流双向都需要通过该点对点接口的网络。
- 宽松模式适用于存在流量非对称问题的非对称路由和多宿主网络。

 下文仅介绍 URPF 的相关内容。

协议规范

- RFC 2827 : Network Ingress Filtering: DDOS Attacks which employ IP Source Address Spoofing
- RFC 3704 : Ingress Filtering for Multi-homed Networks

12.2 典型应用

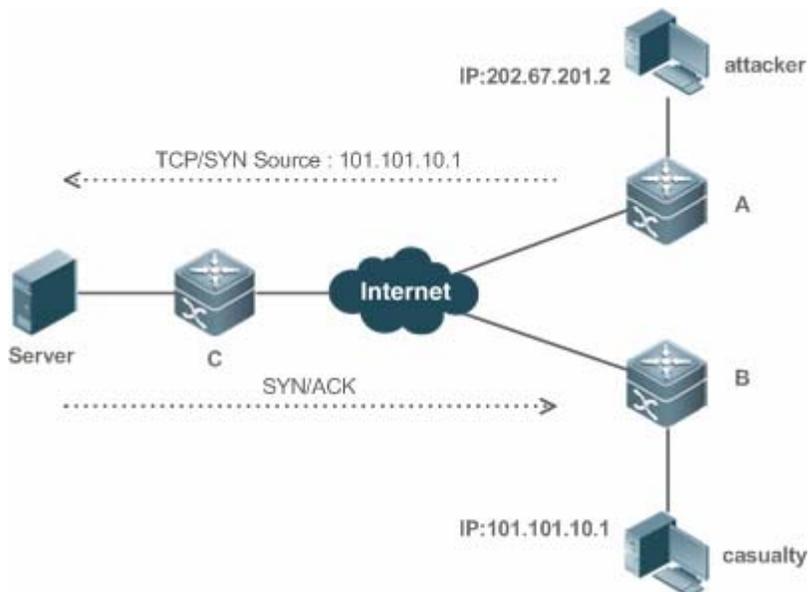
典型应用	场景描述
严格模式	为了防止用户 PC 上伪造的源地址报文传入核心层网络 , 希望能将源地址攻击报文隔离在接入层或者汇聚层 , 从而保证了在汇聚和核心网络上杜绝这类非法数据。
宽松模式	多宿主网络 , 用户网络同时连接有多个 ISP , 往返的流量不对称。为了防止非法的报文攻击用户网络内部 , 在连接 ISP 的出口上启用 URPF 宽松模式 , 从而将非法报文隔离在用户网络外。

12.2.1 严格模式

应用场景

攻击者伪造真实存在的源地址 11.0.0.1 的报文发起攻击 , 导致服务器发送大量的 SYN/ACK 报文给与攻击无关的主机上 , 使得拥有真实源地址的主机也遭受了影响。更严重的是 , 如果网络管理员判定该地址对网络产生了攻击 , 而关闭所有来自该源地址的数据流 , 则对该源地址产生了拒绝服务的现象。

图 12-1



【注释】 attacker 为入侵者，假冒 casualty 发出欺骗报文。

功能部署

- 通过在设备 A 上部署 URPF 严格模式来解决伪造源地址攻击。

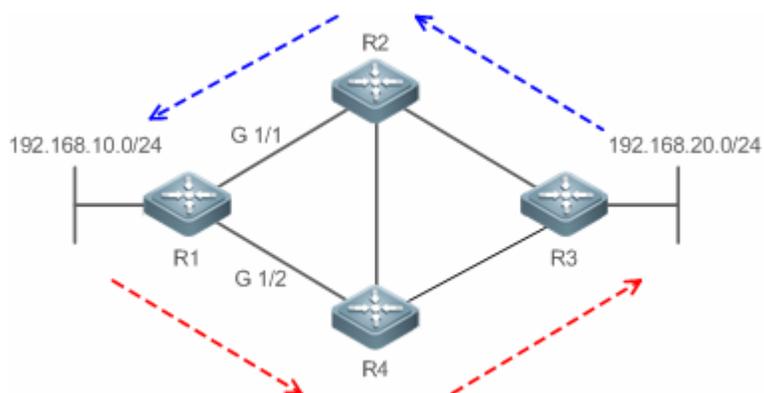
12.2.2 宽松模式

应用场景

由于控制网络流量或者路由策略需要，非对称路由也是比较常见的网络应用。

如下图所示，如果在 R1 设备的 G1/1 上启动 URPF 严格模式，那么如果在 G1/1 设备收到 192.168.20.0/24 网段的数据包时，进行 URPF 检查得到的接口为 G1/2，该报文在 URPF 检查将失败。

图 12-2



功能部属

- 根据接收报文的源 IP 进行反向路由查找，只要找到路由即可，不要求路由的下一跳出口一定是接收报文的入口。
- 通过引入 URPF 宽松模式能够解决了上述存在非对称路由的流量非对称问题，能够杜绝部分非法的数据流访问。

12.3 功能详解

基本概念

▾ URPF 严格模式

通过获取接收报文的源地址和入接口，根据源地址在转发表中查找转发表项，如果表项不存在，那么报文丢弃；如果转发表项的出接口与报文的入接口不匹配，那么报文丢弃。严格模式要求“接收报文的入口一定是到达该报文源地址的路由表项的出口”。

▾ URPF 宽松模式

根据接收报文的源 IP 进行反向路由查找，只要找到路由即可，不要求路由的下一跳出口一定是接收报文的入口，但是不能是本地主机路由。

▾ URPF 丢包率

- 特定的时间段内统计到的因 URPF 检查而丢弃的报文数量，除以该时间段值。单位：报文/秒(packets/second，简写 pps)。

▾ URPF 丢包率的计算时间间隔

- 从上一次计算丢包率到再次计算丢包率的时间间隔。

▾ URPF 丢包率的采样时间间隔

- 进行丢包率计算时，进行丢包数量采集的时间段。该值要求不小于丢包率的计算时间间隔。

▾ URPF 丢包率的水线值

能够承受最大的丢包率的数值。当丢包率高于丢包率水线值时，可以通过 Syslog 或者 Trap 向用户告警，用户可以根据网络实际情况调整丢包率水线值。

▾ URPF 丢包率的告警时间间隔

用来设置连续两次向用户告警的时间间隔。用户可以根据网络实际情况调整告警时间间隔，防止打印 Log 或者发送 Trap 过于频繁。

▾ URPF 丢包率的计算

启动 URPF 特性后，在“0~丢包率的采样时间间隔(含到达丢包率的采样时间点)”时间段内，丢包率值为当前统计到的丢包数除以启动 URPF 的时间。此后计算丢包率的方式：每隔丢包率的计算时间间隔取一次当前丢包数量，减去丢包率的采样时间间隔之前的丢包数量，然后将差值除以丢包率的采样时间间隔，得到了当前的丢包率。

功能特性

功能特性	作用
启动URPF功能	启动 URPF 功能，进行 URPF 检查，解决伪造源地址攻击。
URPF丢包率通告	为了更加方便监控启动 URPF 后报文丢弃的信息，锐捷设备提供了能够通过 Syslog 和 Trap 主动告知用户 URPF 检查丢包的信息。

12.3.1 启动URPF功能

启动 URPF 功能，对 IPv4 报文进行 URPF 检查，以解决伪造源地址攻击的问题。

工作原理

URPF 特性能够根据配置应用于 IPv4 报文，但注意，如下报文将不会被 URPF 检查：

5. URPF 检查只对目的地址为 IPv4 单播地址的报文的源地址进行检查，对目的地址为多播报文或 IPv4 广播报文将跳过检查。
6. 对于源 IP 地址为 0.0.0.0，目的 IP 地址为 255.255.255.255 的 DHCP/BOOTP 报文，跳过检查。
7. 由本机发出并在本机上环回的报文，跳过检查。

接口模式

对所配置的接口收到的报文进行 URPF 检查，包括 IPv4 URPF。

- 缺省情况下，不使用缺省路由进行 URPF 检查，如果用户需要，可以配置使用缺省路由进行检查。
 - 缺省情况下，URPF 检查失败的报文将被丢弃。如果配置了 ACL (*acl-name*)，在 URPF 检查失败的情况下将继续匹配 ACL。如果 ACL 不存在或者匹配到 deny 的 ACE，报文将被丢弃；否则，如果匹配到 permit 的 ACE，报文将被转发。
 - 启用 URPF IPv4 命令后，交换机产品将同时对 IPv4 报文启动 URPF 检查，路由器产品对 IPv4 报文启动 URPF 检查。
-
- ✔ 路由器产品 IPv4 严格模式支持匹配等价路由。
 - ✔ 交换机产品只在 Routed Port 和 L3 AP 口上支持 URPF 特性，并且存在如下限制：
-
- URPF 特性不支持关联 ACL 选项。
 - 启动 URPF 特性后，在这些接口对应的物理口上接收到的所有报文都将启动 URPF 检查，会扩大 URPF 检查报文的范围。典型的应用场景为，如果一个 Tunnel 口接收的报文从上述物理口上收到，该报文也将进行 URPF 检查，如果存在这样的应用场景，建议谨慎启动 URPF 检查。
 - 启动 URPF 特性后，设备路由转发容量将减半。
 - 启动 URPF 严格模式后，接口接收的报文在 URPF 检查时如果匹配等价路由，将转换为宽松模式处理。

相关配置

启动接口 URPF 功能

缺省情况下，接口的 URPF 功能关闭。

使用 **ip verify unicast source reachable-via { rx | any } [allow-default] [acl-name]**命令可以启动或者关闭接口的 IPv4 URPF 功能。

缺省情况下，不使用缺省路由进行 URPF 检查，如果用户需要，可以使用 **allow-default** 关键字启动该功能。

缺省情况下，URPF 检查失败的报文将被丢弃。如果配置了 ACL (*acl-name*)，在选路失败的情况下将继续匹配 ACL。如果 ACL 不存在或者匹配到 deny 的 ACE，报文将被丢弃；否则，如果匹配到 permit 的 ACE，报文将被转发。

12.3.2 URPF丢包率通告

为了更加方便监控启动 URPF 后报文丢弃的信息，锐捷设备提供了能够通过 Syslog 和 Trap 主动告知用户 URPF 检查丢包的信息。

工作原理

启动 URPF 功能后，在“0 ~ 丢包率的采样时间间隔(含到达丢包率的采样时间点)”时间段内，丢包率值为当前统计到的丢包数除以启动 URPF 的时间。此后计算丢包率的方式：每隔丢包率的计算时间间隔取一次当前丢包数量，减去丢包率的采样时间间隔之前的丢包数量，然后将差值除以丢包率的采样时间间隔，得到了当前的丢包率。

启动 URPF 丢包信息监控，能够通过 Syslog 或者 Trap 的方式主动告知用户 URPF 检查丢包的信息，方便用户对网络状况进行监控。

相关配置

配置 URPF 丢包率的计算时间间隔

缺省情况下，URPF 丢包率的计算时间间隔为 30s。如果认为计算间隔太小，

使用 **ip verify urpf drop-rate compute interval seconds** 命令可以修订 URPF 丢包率的计算时间间隔。

URPF 丢包率的计算时间间隔范围为[30,300]。

配置 URPF 丢包率的告警时间间隔

缺省情况下，URPF 丢包率的告警时间间隔为 300s。如果认为告警太频繁或者太少，使用 **ip verify urpf drop-rate notify hold-down seconds** 命令可以修订 URPF 丢包率的告警时间间隔。

URPF 丢包率的告警时间间隔单位为秒，范围[30, 300]。

配置 URPF 丢包信息监控

缺省情况下，URPF 丢包信息监控功能没有启动。

使用 **ip verify urpf drop-rate notify** 命令启动或者关闭 URPF 丢包信息监控功能。

配置 URPF 丢包率的水线值

缺省情况下，URPF 丢包率的水线值为 1000pps。如果认为丢包水线太大或者太小，使用命令 `ip verify urpf notification threshold rate-value` 命令可以修订 URPF 丢包率的水线值。

URPF 丢包率的水线值单位为 pps (packets per second) ，范围[0, 4294967295]。

12.4 产品说明



开启 urpf 功能会造成路由表容量减半，当路由表项超过容量一半时，开启 urpf 功能会造成路由表项丢失。

12.5 配置详解

配置项	配置建议 & 相关命令	
配置启动URPF功能	必须配置。用于启动 URPF 功能。	
	<code>ip verify unicast source reachable-via { rx any } [allow-default] (接口配置模式)</code>	接口启动 URPF 功能
配置URPF丢包率通告	可选配置。用于启动 URPF 丢包率信息监控。	
	<code>ip verify urpf drop-rate compute interval seconds</code>	配置 URPF 丢包率的计算时间间隔
	<code>ip verify urpf drop-rate notify</code>	配置 URPF 丢包信息监控
	<code>ip verify urpf drop-rate notify hold-down seconds</code>	配置 URPF 丢包率的告警时间间隔
	<code>ip erify urpf notification threshold rate-value</code>	配置 URPF 丢包率的水线值

12.5.1 配置启动URPF功能

配置效果

- 启动 URPF 功能，对 IPv4 报文进行 URPF 检查，以解决伪造源地址攻击的问题。
- 在接口模式下启动，接口模式下启动支持严格模式和宽松模式。

注意事项

- URPF 功能需要借助网络中现有的单播路由。因此，网络中必须配置 IPv4 单播路由。

配置方法

启动接口 IPv4 URPF 功能

- 必须配置。
- 交换机产品只在路由口和三层 AP 口上支持，其他产品在路由口上支持。

检验方法

使用源地址攻击的方法，开启 URPF 功能。

- 使用严格模式，检查是不是只有接收的 IP 报文的源地址在转发表中查找到的转发表项的出接口与该报文的入接口匹配，报文才被放行，否则报文被丢弃。
- 使用宽松模式，检查是不是只要接收的 IP 报文的源地址能够在转发表中查找到转发表项，报文被放行，否则报文被丢弃。

相关命令

启动接口 IPv4 URPF 功能

【命令格式】 **ip verify unicast source reachable-via { rx | any } [allow-default] [acl-id]**

【参数说明】 **rx**：使用严格模式进行 URPF 检查。严格模式要求通过接收的 IP 报文的源地址在转发表中查找到的转发表项的出接口要与该报文的入接口匹配。

any：使用宽松模式进行 URPF 检查。宽松模式只要求通过接收的 IP 报文的源地址能够在转发表中查找到转发表项即可。

allow-default：（可选）允许使用缺省路由进行 URPF 检查。

acl-id：（可选）ACL 号，支持：1 to 99 (IP standard access list) ，100 to 199 (IP extended access list) ，1300 to 1999 (IP standard access list, expanded range) ，2000 to 2699 (IP extended access list, expanded range)

【命令模式】 接口模式

【使用指导】 URPF 特性通过检查接收到的 IP 报文的源地址，判断在转发表中是否存在到达源地址的路由，来判断报文的合法性，如果匹配不到转发表项，则认为报文非法。

在接口配置模式下启动 URPF 特性，能对在该接口上接收到的报文启动 URPF 检查。

缺省情况下，不使用缺省路由进行 URPF 检查，如果用户需要，可以使用 **allow-default** 关键字启动该功能。

缺省情况下，URPF 检查失败的报文将被丢弃。如果配置了 ACL (*acl-name*)，在选路失败的情况下将继续匹配 ACL。如果 ACL 不存在或者匹配到 deny 的 ACE，报文将被丢弃；否则，如果匹配到 permit 的 ACE，报文将被转发。

i 启用本命令后，交换机产品将同时对 IPv4 报文启动 URPF 检查，路由器产品对 IPv4 报文启动 URPF 检查。。

i 交换机产品只在 Routed Port 和 L3 AP 口上支持 URPF 特性，并且存在如下限制：

- URPF 特性不支持关联 ACL 选项。
- 启动 URPF 特性后，在这些接口对应的物理口上接收到的所有报文都将启动 URPF 检查，会扩大 URPF 检查报文的范围。典型的应用场景为，如果一个 Tunnel 口接收的报文从上述物理口上收到，该报文也将进行 URPF 检查，如果存在这样的应用场景，建议谨慎启动 URPF 检查。

- 启动 URPF 特性后，设备路由转发容量将减半。
- 启动 URPF 严格模式后，接口接收的报文在 URPF 检查时如果匹配等价路由，将转换为宽松模式处理。

配置举例

i 以下配置举例，仅介绍与 URPF 相关的配置。

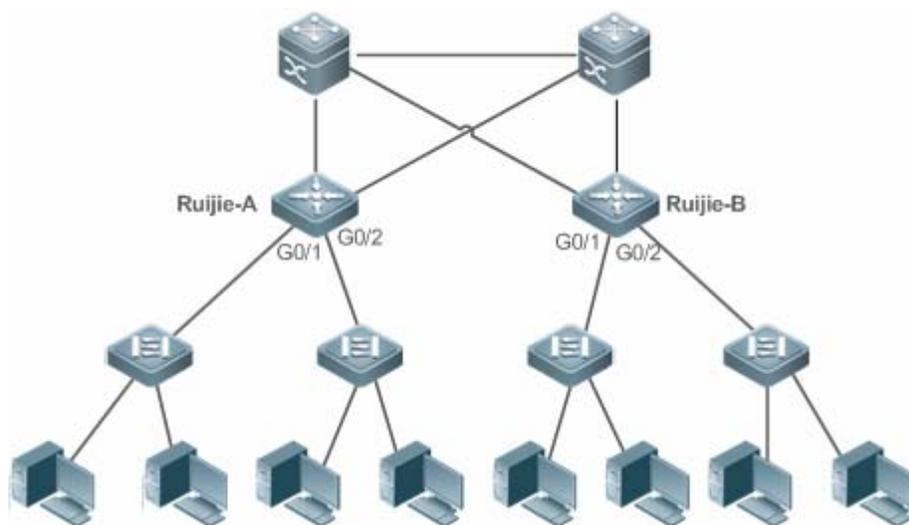
严格模式配置举例

为了防止用户 PC 上伪造的源地址报文传入核心层网络，希望能将源地址攻击报文隔离在接入层或者汇聚层，从而保证了在汇聚和核心网络上杜绝这类非法数据。

通过分析，可以在汇聚设备同接入设备连接的接口上启用 URPF 严格模式，从而满足上述需求。

【网络环境】

图 12-3



【配置方法】

如上图所示，要求在汇聚设备上启用 URPF 严格模式，即在 Ruijie-A 和 Ruijie-B 设备上启用 URPF 严格模式，配置如下所示：

Ruijie-A

```
Ruijie-A# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie-A (config)# interface gigabitEthernet0/1
Ruijie-A (config-if-GigabitEthernet 0/1)# ip address 195.52.1.1 255.255.255.0
Ruijie-A (config-if-GigabitEthernet 0/1)# ip verify unicast source reachable-via rx
Ruijie-A (config-if-GigabitEthernet 0/1)# ip verify urpf drop-rate notify
Ruijie-A (config-if-GigabitEthernet 0/1)#exit
Ruijie-A (config)# interface gigabitEthernet0/2
Ruijie-A (config-if-GigabitEthernet 0/2)# ip address 195.52.2.1 255.255.255.0
Ruijie-A (config-if-GigabitEthernet 0/2)# ip verify unicast source reachable-via rx
Ruijie-A (config-if-GigabitEthernet 0/2)# ip verify urpf drop-rate notify
Ruijie-A (config-if-GigabitEthernet 0/2)#exit
```

Ruijie-B

```
Ruijie-B# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
```

```
Ruijie-B (config)# interface gigabitEthernet0/1
Ruijie-B (config-if-GigabitEthernet 0/1)# ip address 195.52.3.1 255.255.255.0
Ruijie-B (config-if-GigabitEthernet 0/1)# ip verify unicast source reachable-via rx
Ruijie-B (config-if-GigabitEthernet 0/1)# ip verify urpf drop-rate notify
Ruijie-B (config-if-GigabitEthernet 0/1)#exit
Ruijie-B (config)# interface gigabitEthernet0/2
Ruijie-B (config-if-GigabitEthernet 0/2)# ip address 195.52.4.1 255.255.255.0
Ruijie-B (config-if-GigabitEthernet 0/2)# ip verify unicast source reachable-via rx
Ruijie-B (config-if-GigabitEthernet 0/2)# ip verify urpf drop-rate notify
Ruijie-B (config-if-GigabitEthernet 0/2)#exit
```

【检验方法】 如果网络存在源地址欺骗攻击，通过 **show ip urpf** 命令可以看到被 URPF 丢弃的欺骗报文数。

A

```
Ruijie-A#show ip urpf interface gigabitEthernet 0/1
IP verify source reachable-via RX
IP verify URPF drop-rate notify enabled
IP verify URPF notification threshold is 1000pps
Number of drop packets in this interface is 124
Number of drop-rate notification counts in this interface is 0
```

```
Ruijie-A#show ip urpf interface gigabitEthernet 0/2
IP verify source reachable-via RX
IP verify URPF drop-rate notify enabled
IP verify URPF notification threshold is 1000pps
Number of drop packets in this interface is 133
Number of drop-rate notification counts in this interface is 0
```

B

```
Ruijie-B#show ip urpf interface gigabitEthernet 0/1
IP verify source reachable-via RX
IP verify URPF drop-rate notify enabled
IP verify URPF notification threshold is 1000pps
Number of drop packets in this interface is 124
Number of drop-rate notification counts in this interface is 0
```

```
Ruijie-B#show ip urpf interface gigabitEthernet 0/2
IP verify source reachable-via RX
IP verify URPF drop-rate notify enabled
IP verify URPF notification threshold is 1000pps
Number of drop packets in this interface is 250
Number of drop-rate notification counts in this interface is 0
```

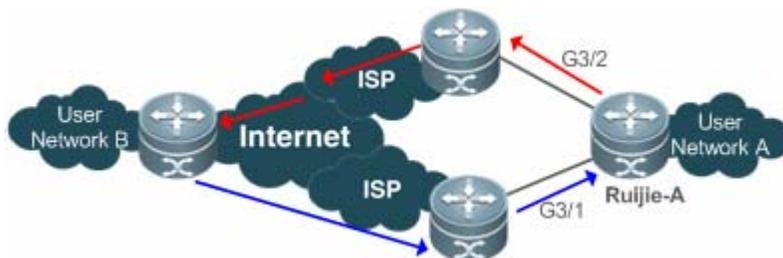
▾ 宽松模式配置举例

在用户网络 A 出口设备 Ruijie-A 上，为了防止非法的报文攻击用户网络内部，在连接两个 ISP 的出口 G3/1

和 G3/2 上启用 URPF 宽松模式，从而将非法报文隔离在用户网络外。

【网络环境】

图 12-4



【配置方法】

Ruijie-A

```
Ruijie-A# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie-A (config)# interface gigabitEthernet3/1
Ruijie-A (config-if-GigabitEthernet 3/1)# ip address 195.52.1.2 255.255.255.252
Ruijie-A (config-if-GigabitEthernet 3/1)# ip verify unicast source reachable-via any
Ruijie-A (config-if-GigabitEthernet 3/1)# ip verify urpf drop-rate notify
Ruijie-A (config-if-GigabitEthernet 3/1)# exit
Ruijie-A (config)# interface gigabitEthernet3/2
Ruijie-A (config-if-GigabitEthernet 3/2)# ip address 152.95.1.2 255.255.255.252
Ruijie-A (config-if-GigabitEthernet 3/2)# ip verify unicast source reachable-via any
Ruijie-A (config-if-GigabitEthernet 3/2)# ip verify urpf drop-rate notify
Ruijie-A (config-if-GigabitEthernet 3/2)# end
```

【检验方法】

如果网络存在源地址欺骗攻击，通过 **show ip urpf** 命令可以看到被 URPF 丢弃的欺骗报文数。

A

```
Ruijie #show ip urpf
IP verify URPF drop-rate compute interval is 300s
IP verify URPF drop-rate notify hold-down is 300s
Interface gigabitEthernet3/1
IP verify source reachable-via ANY
IP verify URPF drop-rate notify enabled
IP verify URPF notification threshold is 1000pps
Number of drop packets in this interface is 4121
Number of drop-rate notification counts in this interface is 2
Interface gigabitEthernet3/2
IP verify source reachable-via ANY
IP verify URPF drop-rate notify enabled
IP verify URPF notification threshold is 1000pps
Number of drop packets in this interface is 352
Number of drop-rate notification counts in this interface is 0
```

常见错误

● -

12.5.2 配置URPF丢包率通告

配置效果

- 启动 URPF 丢包信息监控，能够通过 Syslog 或者 Trap 的方式主动告知用户 URPF 检查丢包的信息，方便用户对网络状况进行监控。

注意事项

- 必须先配置启动 URPF 功能。

配置方法

配置 URPF 丢包率的计算时间间隔

- 可选配置。
- 全局模式。

配置 URPF 丢包率的告警时间间隔

- 可选配置。
- 全局模式。

配置 URPF 丢包信息监控

- 可选配置。
- 接口模式。

配置 URPF 丢包率的水线值

- 可选配置。
- 接口模式。

检验方法

使用源地址攻击，开启 URPF 功能，检验如下点：

- 打开告警功能后，当丢包率超过水线值后，是否能正常的告警

相关命令

配置 URPF 丢包率的计算时间间隔

- 【命令格式】 **ip verify urpf drop-rate compute interval seconds**
- 【参数说明】 **interval seconds** : URPF 丢包率的计算时间间隔。单位为秒，范围[30, 300]，缺省值 30s。
- 【命令模式】 全局模式
- 【使用指导】 URPF 丢包率的计算时间间隔基于全局配置模式进行配置，将应用于全局的 URPF 丢包率计算和基于每个启动 URPF 特性的接口的 URPF 丢包率计算。

配置 URPF 丢包率的告警时间间隔

- 【命令格式】 **ip verify urpf drop-rate notify hold-down seconds**
- 【参数说明】 **hold-down seconds** : URPF 丢包率的告警时间间隔。单位为秒，范围[30, 300]，缺省值 300s。
- 【配置模式】 全局模式
- 【使用指导】 URPF 丢包率的告警时间间隔基于全局配置模式进行配置，将应用于全局的 URPF 丢包率告警和基于每个启动 URPF 特性的接口的 URPF 丢包率告警。

配置 URPF 的丢包信息监控

- 【命令格式】 **ip verify urpf drop-rate notify**
- 【参数说明】 -
- 【配置模式】 接口模式
- 【使用指导】 通过启动 URPF 丢包信息监控，能够通过 Syslog 或者 Trap 的方式主动告知用户 URPF 检查丢包的信息，方便用户对网络状况进行监控。

配置 URPF 丢包率的水线值

- 【命令格式】 **ip verify urpf notification threshold rate-value**
- 【参数说明】 **threshold rate-value** : URPF 丢包率的水线值。单位为 pps (packets per second)，范围[0, 4294967295]，缺省值 1000pps。
- 【配置模式】 接口模式
- 【使用指导】 如果水线值为 0，表示监测到任何一个因 URPF 检查而丢弃的报文就会发送通告。用户可以根据网络实际情况调整丢包率水线值。

配置举例

配置 URPF 丢包率的计算时间间隔为 120s。

- 【配置方法】 在全局模式下配置 URPF 丢包率的计算时间间隔为 120s。

```
Ruijie# configure terminal
Ruijie(config)# ip verify urpf drop-rate compute interval 120
Ruijie(config)# end
```

- 【检验方法】 使用 show ip urpf 命令查看是否配置生效。

```
Ruijie# show ip urpf
IP verify URPF drop-rate compute interval is 120s
```

配置 URPF 丢包率的告警时间间隔为 120s

【配置方法】 在全局模式下配置 URPF 丢包率的告警时间间隔为 120s。

```
Ruijie# configure terminal
Ruijie(config)# ip verify urpf drop-rate notify hold-down 120
Ruijie(config)# end
```

【检验方法】 使用 show ip urpf 命令查看是否配置生效。

```
Ruijie# show ip urpf
IP verify URPF drop-rate compute interval is 120s
IP verify URPF drop-rate notify hold-down is 120s
```

配置接口 GigabitEthernet 0/1 启动 IPv4 URPF 丢包信息监控。

【配置方法】 在 GigabitEthernet 0/1 上配置启动 IPv4 URPF 丢包信息监控。

```
Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet0/1
Ruijie(config-if-GigabitEthernet 0/1)# ip verify unicast source reachable-via rx
Ruijie(config-if-GigabitEthernet 0/1)# ip verify urpf drop-rate notify
```

【检验方法】 使用 show ip urpf 命令查看接口是否启动了 URPF 丢包信息监控。

```
Ruijie# show ip urpf interface gigabitEthernet 0/1
IP verify source reachable-via RX
IP verify URPF drop-rate notify is enabled
IP verify URPF notification threshold is 1000pps
Number of drop packets in this interface is 0
Number of drop-rate notification counts in this interface is 0
Number of drop-rate notification counts in this interface is 0
```

配置 GigabitEthernet 0/1 IPv4 URPF 丢包率的水线值为 2000pps

【配置方法】 在 GigabitEthernet 0/1 接口上配置 IPv4 URPF 的丢包率水线值为 2000pps。

```
Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet0/1
Ruijie(config-if-GigabitEthernet 0/1)# ip verify unicast source reachable-via rx
Ruijie(config-if-GigabitEthernet 0/1)# ip verify urpf notification threshold 2000
```

【检验方法】 使用 show ip urpf 命令查看接口的 IPv4 URPF 丢包率水线值。

```
Ruijie# show ip urpf interface gigabitEthernet 0/1
IP verify source reachable-via RX
IP verify URPF drop-rate notify is enabled
IP verify URPF notification threshold is 2000pps
Number of drop packets in this interface is 0
Number of drop-rate notification counts in this interface is 0
```

常见配置错误

- -

12.6 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清除 IPv4 URPF 检查丢弃报文数量的统计信息。	clear ip urpf [interface <i>interface-name</i>]

查看运行情况

作用	命令
查看 IPv4 URPF 配置及统计信息。	show ip urpf [interface <i>interface-name</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 URPF 事件的调试开关。	debug urpf event
打开 URPF 的定时调试开关。	debug urpf timer

13 CPP

13.1 概述

- CPP (CPU Protect Policy , CPU 保护策略) 提供了交换机 CPU 保护的策略。

在网络环境中, 有各种攻击报文在网络上传播, 其会导致交换机的 CPU 利用率过高, 影响协议运行, 甚至无法正常管理交换机。针对这种情况, 必须对交换机的 CPU 进行保护, 即对送往交换机的 CPU 处理的各种报文进行流量控制和优先级处理, 保障其正常处理能力。

CPP 功能能够有效的抵御网络中的恶意攻击, 为正常的协议报文提供干净的运行环境。

CPP 缺省情况下已开启, 伴随交换机的整个运行过程, 都在发挥着其保护作用。

13.2 典型应用

典型应用	场景描述
防止恶意攻击	当网络中存在各种的恶意攻击, 例如 ARP 攻击, CPP 可以将攻击报文划分到不同的优先级队列, 使得攻击报文不会对其他报文产生影响。
CPU处理瓶颈	即使没有攻击存在, 如果正常业务流量太大, 同样会导致 CPU 处理瓶颈。通过 CPP 功能对送往 CPU 的报文进行限速, 以确保交换机正常运行。

13.2.1 防止恶意攻击

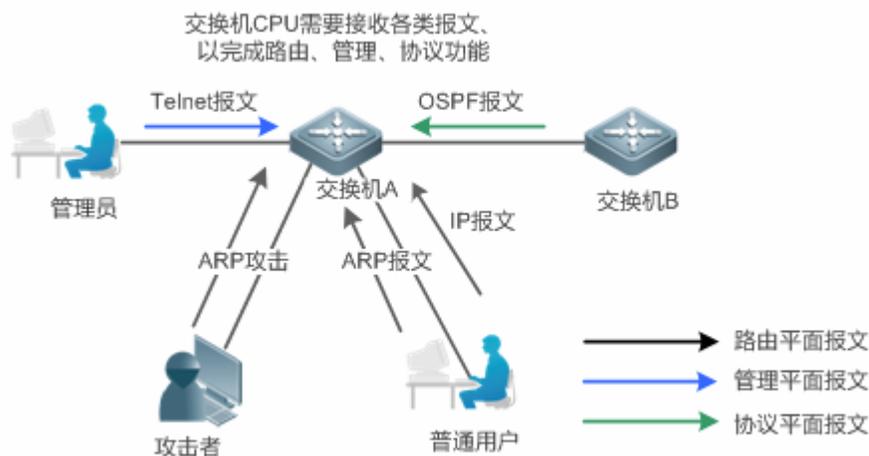
应用场景

网络中的各个层次的交换机, 都可能受到恶意的报文攻击, 典型的如 ARP 攻击。

以下图为例, 交换机 CPU 需要处理的报文分为三类: 路由平面, 为路由服务, 比如 ARP 报文, IP 路由未打通报文; 管理平面, 为管理交换机服务, 比如 Telnet 报文, Http 报文; 协议平面报文, 为协议运行服务, 比如 BPDU 报文, OSPF 报文。

当存在一个攻击者, 通过 ARP 报文进行攻击, 由于 ARP 报文会送 CPU 处理, 而 CPU 处理能力有限, 因此, 该 ARP 会挤掉其他报文 (导致其他报文被丢弃), 同时也会导致 CPU 资源耗费 (大量资源用于处理 ARP 攻击报文), 从而无法正常运行。针对图中场景, 将会导致: 普通用户无法上网; 管理员无法管理交换机; 交换机 A 与网络邻居 B 之间的 OSPF 断开, 无法进行路由学习。

图 13-1 交换机业务及攻击示意图



功能部属

- CPP 缺省配置下，将 ARP 与 Telnet 报文，IP 路由未打通报文，OSPF 报文划分到不同的优先级队列。ARP 报文的攻击不会对其他报文产生影响。
- CPP 缺省配置下，通过对 ARP 报文的限速以及对 ARP 所在优先级队列的限速，保证攻击报文不会对 CPU 资源产生太大影响。
- 与 ARP 报文在同一个优先级队列的报文会受到 ARP 攻击报文的影响，可以通过配置将其与 ARP 划分到不同优先级队列。
- 在 ARP 攻击报文存在情况下，CPP 无法保证其他正常的 ARP 报文不受到影响（CPP 只提供报文类别层面的差异性处理，同为一种报文类型，无法区分攻击报文和正常报文），这种情况下，需要 NFPP（Network Foundation Protection Policy，网络基础保护策略）功能来提供更细粒度的攻击防护。

i 针对 NFPP 相关配置说明，请参考“NFPP”章节。

13.2.2 CPU处理瓶颈

应用场景

即使不存在攻击，但在某个瞬间也可能存在大量需要送 CPU 处理的报文。

以园区网核心设备为例，接入点数以万计。最坏情况下，正常 ARP 报文流量也会达到几万 pps（packets per second，每秒报文数），如果所有报文都送到 CPU 处理，CPU 资源将无法支撑，从而可能导致协议震荡，CPU 运行异常等。

功能部属

- CPP 缺省配置下，通过对 ARP 报文的限速以及对 ARP 所在优先级队列的限速，控制 ARP 报文送到 CPU 的速率，保证 CPU 资源耗费控制在一定范围内，从而可以对其他协议进行正常处理。

- 对其他用户层面的报文，比如 Web 认证，1X 认证报文，CPP 缺省也都有对应限速配置。

13.3 功能详解

基本概念

▾ QoS , DiffServ

QoS (Quality of Service) 服务质量，是网络的一种安全机制，是用来解决网络延迟和阻塞等问题的一种技术。

DiffServ 指差分服务模型，是 QoS 实现模型的一种典型，通过对业务流进行划分，以提供差异性的服务。

▾ Bandwidth , Rate

带宽指最大可承载的数据率，在本文中主要用于指限速阈值，超过限速阈值的报文将会被丢弃。

速率指实际的数据率，当速率超过带宽时，超过部分将被丢弃。速率只能小于等于带宽。

本文中的带宽和速率的单位都为 pps (packets per second , 每秒报文数)。

▾ L2 , L3 , L4

指报文按照 TCP/IP 模型划分的层次结构。

L2 指二层头部，即以太网封装部分；L3 指三层头部，即 IP 封装部分；L4 指四层头部，一般指 TCP/UDP 的头部封装部分。

▾ 优先级队列，SP

报文在交换机内部会被缓存，输出方向即在队列中进行缓存，优先级队列是与 SP 对应的，各个队列并非对等，而是严格划分优先级。

SP(Strict Priority)严格优先级调度算法，是 QoS 调度算法的一种，当高优先级队列中有报文时，一定调度高优先级队列。调度指从队列中选择报文进行输出，本文中，指选择报文送 CPU。

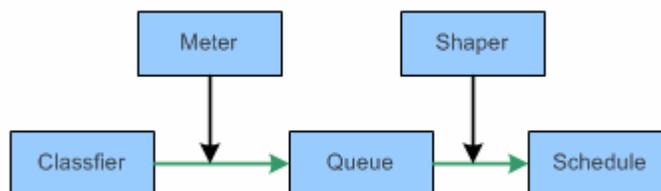
▾ CPU 端口

交换机将报文送 CPU 前，会进行缓存，报文投递 CPU 的过程，实际上类似于报文输出的过程，CPU 端口是一个虚拟端口，报文被送 CPU，即从这个虚拟端口输出，上述的优先级队列，SP 均是针对 CPU 端口的。

功能特性

CPP 功能通过标准的 QoS (服务质量) 差分服务模型 (DiffServ) 实现 CPU 保护。

图 13-2 CPP 实现模型图



功能特性	作用
Classifier	完成对报文类型的分类，为后续的 QOS 策略实施提供保证。
Meter	基于报文类型进行限速处理，控制某种报文类型的带宽。
Queue	对送往 CPU 的报文进行排队，基于报文类型可选择不同的队列进行排队。
Scheduler	选择队列进行报文调度，被调度到的报文被送往 CPU。
Shaper	完成基于优先级队列和 CPU 端口进行限速处理，控制优先级队列和 CPU 端口的带宽。

13.3.1 Classifier

工作原理

Classifier (分类器) 对每个需要送到 CPU 的报文进行分类，分类时根据报文的 L2、L3 以及 L4 信息。对报文进行分类是实施 QOS 策略的基础，在后续动作中根据分类可以实施不同策略，提供区别服务。交换机提供的分类是固定的，按照交换机支持的协议，管理功能固定划分报文类型，比如生成树协议的 BPDU 报文，网络控制报文协议的 ICMP 报文等。不支持自定义报文的分类规则。

i 由于硬件差异性，不同产品对报文的分类有差异，具体参见产品特性文档。

相关配置

- -

13.3.2 Meter

工作原理

Meter (测量器) 根据对特定分类报文按照设定的速率阈值进行限速。支持对不同的报文类型设置不同的速率阈值，当某报文类型的报文速率超过对应阈值时，超限部分将被丢弃。

通过 Meter，可以控制送 CPU 的某类报文速率不会超过某个阈值，防止特定攻击报文对 CPU 资源产生过大影响，这是 CPP 的第一级防护。

相关配置

- 在缺省情况下，所有报文类型都对应一个速率阈值(带宽)，并据此实施 Meter 策略。

- i** 某些产品不支持 Meter 功能，各产品的支持情况以及各报文类型的缺省带宽，具体参见产品特性文档。
- 在实际使用中，使用 `cpu-protect type packet-type bandwidth bandwidth-value` 命令可以对特定报文类型进行 Meter 设置。

13.3.3 Queue

工作原理

Queue（队列）完成了对报文的二级划分，可以多种不同的报文类型选择同一个队列；同时，队列完成了报文在交换机内部的缓存，为接下来的 Scheduler 和 Shaper 提供服务。

CPP 对应的队列是严格优先级（SP）队列，通过报文的入队列，确定报文的严格优先级，队列号大的优先级高。

相关配置

- 在缺省情况下，各报文类型均对应一个优先级队列。
- i** 队列数目，以及各报文类型的缺省优先级队列，具体参见产品特性文档。
- 在实际使用中，可以使用 `cpu-protect type packet-type traffic-class traffic-class-num` 命令可以对特定报文类型选择优先级队列。

13.3.4 Scheduler

工作原理

Scheduler（调度器）根据优先级队列，对报文进行严格优先级（SP）调度，即只要更高优先级队列存在报文，严格优先调度更高优先级队列中的报文。

在调度前，送 CPU 的报文缓存在队列中，当被调度时，报文被送往 CPU 处理。

- i** 调度策略不可更改，只支持严格优先级调度。

相关配置

-

13.3.5 Shaper

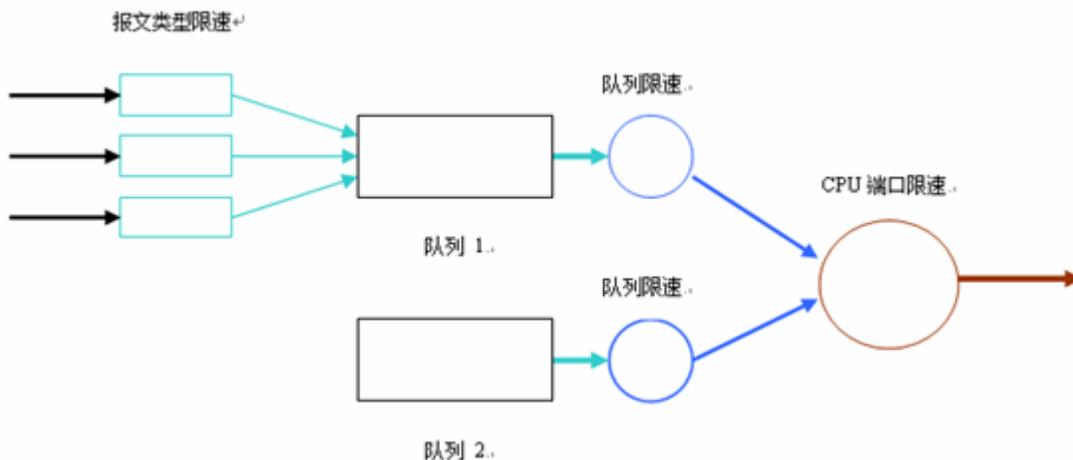
工作原理

Shaper（整形器）完成对送往 CPU 的报文进行整形，即当实际速率大于整形阈值时，报文继续缓存在队列中，不进行调度，因此，在报文速率波动情况下，通过 Shaper，可以使得送往 CPU 的报文速率是平滑的（速率小于等于整形阈值）。

在有 Shaper 的情况下，高优先级队列的报文不一定调度完成才调度低优先级队列，某个优先级队列报文速率若超过整形阈值，将会暂时停止调度。因此，通过 Shaper，可以避免低优先级队列的报文被饿死(一直调度高优先级队列报文，低优先级队列报文得不到调度)。

由于 Shaper 通过限制调度报文的速率 实际上也提供了限速的功能。针对优先级队列和所有送 CPU 的报文(CPU 端口) ,Shaper 提供了两级限速。和 Meter 功能一起，形成三级限速，为 CPU 提供三级防护。

图 13-3 CPP 的三级限速



相关配置

配置优先级队列的 Shaper

- 在缺省情况下，各优先级队列均确定了一个整形阈值（带宽）。

i 某些交换机不支持 Shaper 功能，各产品的支持情况以及各优先级队列和 CPU 端口的缺省带宽，具体参见产品特性文档。

- 在实际使用中，可使用 `cpu-protect traffic-class traffic-class-num bandwidth bandwidth_value` 命令可以对特定优先级队列进行 Shaper 配置。

配置 CPU 端口的 Shaper

- 在缺省情况下，CPU 端口确定了一个整形阈值（带宽）。

i 某些交换机不支持 Shaper 功能，各产品的支持情况以及 CPU 端口的缺省带宽，具体参见产品特性文档。

- 使用 `cpu-protect cpu bandwidth bandwidth_value` 命令可以对 CPU 端口进行 Shaper 配置。

13.4 产品说明



S7508 产品 CPU Protect 的默认值：

报文类型	带宽(pps)	报文优先级
bpd	128	6

arp	10000	1
tpp	128	6
dot1x	1500	2
gvrp	128	5
rldp	128	5
lACP	256	5
rerp	128	5
reup	128	5
lldp	768	5
cdp	768	5
dhcps	1500	2
dhcp-relay-c	1500	2
dhcp-relay-s	1500	2
option82	1500	2
tunnel-bpdu	128	2
tunnel-gvrp	128	2
unknown-v4mc	128	1
xgv-ipmc	128	2
stargv-ipmc	128	2
udp-helper	128	1
dvmrp	128	4
igmp	1000	2
icmp	4500	3
ospf	2000	4
pim	1000	4
rip	128	4
vrrp	256	6
ttl0	128	0
ttl1	2000	0
hop_limit	800	0
local-ipv4	4000	3
v4uc-route	800	1
rt-host	4096	4
mld	1000	2
nd-snp-ns-na	10000	1
nd-snp-rs	1500	1
nd-snp-ra-redirect	1500	1
erps	128	5
bgp	128	4
cfm	512	5

web-auth	2000	2
fcoe-fip	1000	4
fcoe-local	1000	4
bfd	5120	6
dldp	3200	6
other	2000	0
trill	1000	4
efm	1000	5
ip-option	800	0
mgmt	8000	-
Dns	200	2
Sdn	20000	0

各优先级队列的默认带宽值：

队列	默认带宽值
0	20000
1	20000
2	20000
3	20000
4	20000
5	20000
6	20000
7	20000

CPU 口默认带宽值：100000

13.5 配置详解

配置项	配置建议 & 相关命令	
配置CPP	 可选配置，缺省已配置。用于调整 CPP 的相关配置参数。	
	cpu-protect type <i>packet-type</i> bandwidth	配置某报文类型的带宽
	cpu-protect type <i>packet-type</i> traffic-class	配置某报文类型对应的优先级队列
	cpu-protect traffic-class <i>traffic-class-num</i> bandwidth	配置某优先级队列的带宽
	cpu-protect cpu bandwidth	配置 CPU 端口的带宽

13.5.1 配置CPP

配置效果

- 通过配置 Meter，可以对某种报文类型设置带宽，限制其速率。该报文类型的超限部分将被直接丢弃。
- 通过配置 Queue，可以对某种报文类型选择优先级队列，其中高优先级队列中的报文将优先得到调度。
- 通过配置 Shaper，可以对 CPU 端口和优先级队列设置带宽，限制其速率。超限部分将被直接丢弃。

注意事项

- 在带宽配置变小的时候要特别注意，可能会对该报文类型的正常业务流产生影响，如果需要区分用户进行 CPU 防护，需要结合 NFPP 功能。
- Meter 要与 Shaper 配置结合起来，由于是三级防护，只控制其中某一级，可能带来另外的副作用。比如需要调大某报文类型的带宽，对应优先级队列的带宽也需要一起调整，否则，该报文类型可能对同优先级队列的其他报文产生影响。

配置方法

配置报文类型的带宽(Meter)

- 可选配置，缺省已配置，且不可关闭。
- 需要进行配置修改出现在：某种报文类型未攻击却被丢弃的情况，需要将该报文类型的带宽配大；某种报文类型攻击导致 CPU 运行异常的情况，需要将该报文类型带宽配小。
- 网络环境中的各交换机均有此配置。

配置报文类型的队列(Queue)

- 可选配置，缺省已配置，且不可关闭。
- 需要进行配置修改出现在：某种报文类型攻击导致同队列的其他报文异常时，可通过将该报文类型放入一个未使用的队列中；某种报文类型不允许丢弃，但当前同其他一些使用中的报文类型在同一个队列，可以将其放入更高优先级队列中。
- 网络环境中的各交换机均有此配置。

配置优先级队列的带宽(Shaper)

- 可选配置，缺省已配置，且不可关闭。
- 需要进行配置修改出现在：某种报文类型的 Meter 配大，导致对应优先级队列中其他报文可用带宽不多，应适当对该优先级队列的带宽配大；某个优先级队列中放入攻击报文，且无其他使用中的报文，将该优先级队列的带宽配大。
- 网络环境中的各交换机均有此配置。

配置 CPU 端口的带宽(Shaper)

- 可选配置，缺省已配置，且不可关闭。
- CPU 端口的带宽值不建议进行修订。
- 网络环境中的各交换机均有此配置。

检验方法

- 通过在异常情况下进行配置改进后的系统运行情况，来检查配置是否生效。
- 通过查看对应配置及统计值可以检验配置是否生效，具体参见下面命令。

相关命令

配置报文类型的带宽

- 【命令格式】 **cpu-protect type packet-type bandwidth bandwidth_value**
- 【参数说明】 *packet-type*：指定报文类型。报文类型固定划分。
bandwidth_value：指定具体的带宽值。带宽值单位为 pps，即每秒报文数。
- 【命令模式】 全局模式
- 【使用指导】 -

配置报文类型的优先级队列

- 【命令格式】 **cpu-protect type packet-type traffic-class traffic-class-num**
- 【参数说明】 *packet-type*：指定报文类型。报文类型固定划分。
traffic-class-num：指定具体的优先级队列。
- 【命令模式】 全局模式
- 【使用指导】 -

配置优先级队列的带宽

- 【命令格式】 **cpu-protect traffic-class traffic-class-num bandwidth bandwidth_value**
- 【参数说明】 *traffic-class-num*：指定具体的优先级队列。
bandwidth_value：指定具体的带宽值。带宽值单位为 pps，即每秒报文数。
- 【命令模式】 全局模式
- 【使用指导】 -

配置 CPU 端口的带宽

- 【命令格式】 **cpu-protect cpu bandwidth bandwidth_value**
- 【参数说明】 *bandwidth_value*：指定具体的带宽值。带宽值单位为 pps，即每秒报文数。
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

通过 CPP 防止报文攻击及网络震荡

- 【网络环境】
- 系统中有 ARP 流，IP 流，OSPF 流，dot1x 流，VRRP 流，Telnet 流，ICMP 流。若假设在当前配置下，ARP 流和 dot1x 流属于同一个优先级队列 2，IP 流、ICMP 流、Telnet 流属于同一个优先级队列 4，OSPF 流属于优先级队列 3，VRRP 流属于优先级队列 6。各报文类型的带宽均为 10000PPS，各优先级队列的带宽均为 20000PPS，CPU 端口的带宽为 100000PPS。
 - 系统中存在 ARP 攻击和 IP 扫描攻击，系统运行异常，认证不上，Ping 不通，无法管理，且 OSPF 震荡。

- 【配置方法】
- 将攻击的 ARP 报文划入优先级队列 1，限制 ARP 报文或对应优先级队列的带宽。
 - 将 OSPF 报文划分更高的优先级队列 5。
 - 将攻击的 IP 未打通报文划入优先级队列 3，限制 IP 报文或优先级队列的带宽。

```
Ruijie# configure terminal
Ruijie(config)# cpu-protect type arp traffic-class 1
Ruijie(config)# cpu-protect type arp bandwidth 5000
Ruijie(config)# cpu-protect type ospf traffic-class 5
Ruijie(config)# cpu-protect type v4uc-route traffic-class 3
Ruijie(config)# cpu-protect type traffic-class 3 bandwidth 5000
Ruijie(config)# end
```

- 【检验方法】 通过 **show cpu-protect** 可以查看到配置及统计信息。

```
Ruijie# show cpu-protect
%cpu port bandwidth: 80000(pps)
Traffic-class   Bandwidth(pps)   Rate(pps)       Drop(pps)
-----
0                8000              0                0
1                8000              0                0
2                8000              0                0
3                8000              0                0
4                8000              0                0
5                8000              0                0
6                8000              0                0
7                8000              0                0

Packet Type     Traffic-class   Bandwidth(pps)   Rate(pps)       Drop(pps)       Total           Total Drop
-----
bpdu            6               128              0                0                0                0
arp             3               10000            0                0                0                0
arp-dai        3               10000            0                0                0                0
arp-proxy      3               10000            0                0                0                0
tpp            7               128              0                0                0                0
dot1x          4               128              0                0                0                0
gvrp           5               128              0                0                0                0
rldp           6               128              0                0                0                0
lACP           6               128              0                0                0                0
rerp           6               128              0                0                0                0
reup           6               128              0                0                0                0
lldp           5               128              0                0                0                0
cdp            5               128              0                0                0                0
dhcps          4               128              0                0                0                0
dhcps6         4               128              0                0                0                0
dhcp6-client   4               128              0                0                0                0
```

dhcp6-server	4	128	0	0	0	0
dhcp-relay-c	4	128	0	0	0	0
dhcp-relay-s	4	128	0	0	0	0
option82	4	128	0	0	0	0
tunnel-bpdu	5	128	0	0	0	0
tunnel-gvrp	5	128	0	0	0	0
unknown-v4mc	3	128	0	0	0	0
known-v4mc	3	128	0	0	0	0
xgv-ipmc	3	128	0	0	0	0
sgv-ipmc	3	128	0	0	0	0
udp-helper	4	128	0	0	0	0
dvmrp	5	128	0	0	0	0
igmp	4	128	0	0	0	0
icmp	4	128	0	0	0	0
ospf	5	128	0	0	0	0
ospf3	5	128	0	0	0	0
pim	6	128	0	0	0	0
rip	6	128	0	0	0	0
ripng	6	128	0	0	0	0
vrrp	6	128	0	0	0	0
vrrp6	6	128	0	0	0	0
ttl0	6	128	0	0	0	0
ttl1	6	128	0	0	0	0
err_hop_limit	1	800	0	0	0	0
local-ipv4	6	128	0	0	0	0
route-host-v4	0	4096	0	0	0	0
mld	0	1000	0	0	0	0
nd-snp-ns-na	6	128	0	0	0	0
nd-snp-rs	6	128	0	0	0	0
nd-snp-ra-redirect	6	128	0	0	0	0
nd-non-snp	6	128	0	0	0	0
cfm	0	128	0	0	0	0
fcoe-fip	6	128	0	0	0	0
fcoe-local	6	128	0	0	0	0
bfd-echo	6	5120	0	0	0	0
bfd-ctrl	6	5120	0	0	0	0
madp	7	1000	0	0	0	0
ip4-other	6	128	0	0	0	0
ip6-other	6	128	0	0	0	0
non-ip-other	6	20000	0	0	0	0
trill	2	1000	0	0	0	0

trill-oam	2	1000	0	0	0	0
efm	2	1000	0	0	0	0

常见错误

-

13.6 监视与维护

清除各类信息

作用	命令
清除 CPP 的统计信息。	clear cpu-protect counters [device <i>device_num</i>] [slot <i>slot_num</i>]
清除主设备的 CPP 的统计信息。	clear cpu-protect counters mboard

查看运行情况

作用	命令
查看报文类型的配置及统计值。	show cpu-protect type <i>packet-type</i> [device <i>device_num</i>] [slot <i>slot_num</i>]
查看优先级队列的配置及统计值。	show cpu-protect traffic-class <i>traffic-class-num</i> [device <i>device_num</i>] [slot <i>slot_num</i>]
查看 CPU 端口的配置值	show cpu-protect cpu
查看主设备上的所有配置及统计值	show cpu-protect { mboard summary }
查看 CPP 相关的所有配置及统计值	show cpu-protect [device <i>device_num</i>] [slot <i>slot_num</i>]

查看调试信息

-

- i** 上述监视和维护命令兼容机箱设备和盒式设备，以及单机和 VSU。
- i** 若 **device** 和 **slot** 关键字均不输入，针对 **clear** 命令表示清除系统内所有节点的统计信息，针对 **show** 命令表示主设备。
- i** 针对单机情况，**device** 关键字不可见。机箱设备通过 **slot** 可指定某张线卡；盒式设备的 **slot** 关键字不可见。
- i** 针对 VSU 情况，通过 **device** 可以表示对应的机箱或者盒式设备节点，若未输入 **device** 关键字，则表示主机箱或者主设备。机箱设备在输入了 **device** 关键字的情况下，必须输入 **slot** 关键字，用于标识某个机箱下的某张线卡；盒式设备的 **slot** 关键字不可见。

14 DHCP Snooping

14.1 概述

DHCP Snooping：意为 DHCP 窥探，通过对 Client 和服务器之间的 DHCP 交互报文进行窥探实现对用户 IP 地址使用情况的记录和监控，同时还可以过滤非法 DHCP 报文，包括客户端的请求报文和服务端的响应报文。DHCP Snooping 记录生成的用户数据表项可以为 IP Source Guard 等安全应用提供服务。

i 下文仅介绍 DHCP Snooping 的相关内容。

协议规范

- RFC2131：Dynamic Host Configuration Protocol
- RFC2132：DHCP Options and BOOTP Vendor Extensions

14.2 典型应用

典型应用	场景描述
DHCP服务欺骗攻击防范	在网络上存在多个 DHCP 服务器，限制 DHCP 客户端只能从合法的 DHCP 服务获取网络配置参数。
DHCP报文泛洪攻击防范	在网络上存在恶意用户，频繁的发送 DHCP 请求报文。
伪造DHCP报文攻击防范	在网络上存在恶意用户，发送伪造的 DHCP 请求报文，比如 DHCP-Release 报文。
IP/MAC欺骗攻击防范	在网络上存在恶意用户，发送伪造的 IP 报文，如篡改了报文的源地址字段。
用户私设IP限制	在网络上存在用户，不按规定从 DHCP 服务器获取 IP 地址，私设 IP 地址。
ARP入侵检测	在网络上存在恶意用户，伪造 ARP 响应报文，意图拦截正常用户之间通信的报文。

14.2.1 DHCP服务欺骗攻击防范

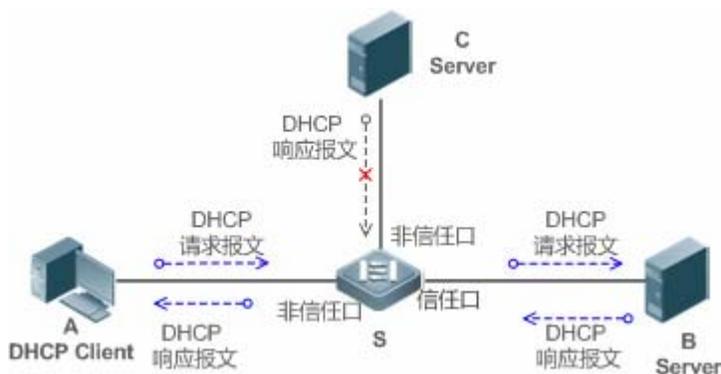
应用场景

在网络中可能存在多个 DHCP 服务器，需要保证用户 PC 只能从控制范围内的 DHCP 服务器获取网络配置参数。

以下图为例，DHCP 客户端仅与可信 DHCP 服务器通信。

- DHCP 客户端的请求报文只会传输到可信任的 DHCP 服务器。
- 只有可信任 DHCP 服务器的响应报文才会传输给客户端。

图 14-1



- 【注释】 S 为接入设备。
A 为用户 PC。
B 为控制范围内的 DHCP 服务器。
C 为不受控的 DHCP 服务器。

功能部署

- 在接入设备 S 上开启 DHCP Snooping 服务，实现 DHCP 报文监控。
- 设置接入设备 S 链接 DHCP 服务器 B 的端口为 DHCP TRUST 口，实现响应报文的转发。
- 设置接入设备 S 的其余端口为 DHCP UNTRUST 口，实现响应报文的过滤。

14.2.2 DHCP报文泛洪攻击防范

应用场景

在网络中可能存在恶意 DHCP 客户，高速率的发送 DHCP 请求报文，造成合法用户无法获得 IP、接入设备高负荷运行甚至瘫痪。需要保证网络系统运行稳定。

应用 DHCP 报文限速，DHCP 客户端仅能以低于规定的速率发送 DHCP 请求报文。

- DHCP 客户端的请求报文发送速率低于规定阈值。
- 超出限定的报文被丢弃。

功能部署

- 在接入设备 S 上开启 DHCP Snooping 服务，实现 DHCP 监控。
- 限制 UNTRUST 口的 DHCP 报文发送速率。

14.2.3 伪造DHCP报文攻击防范

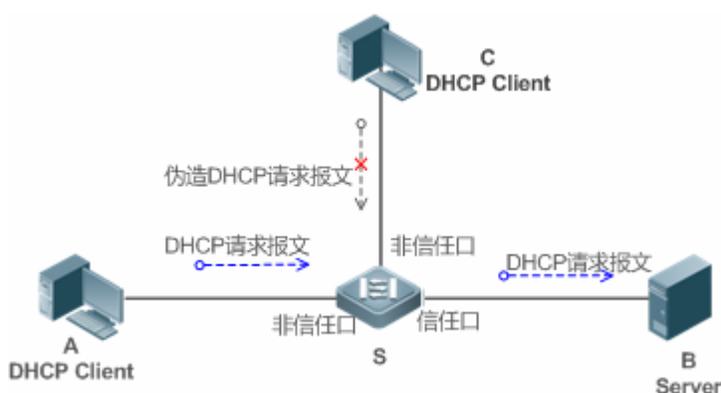
应用场景

在网络中可能存在恶意用户，伪造 DHCP 请求报文，一方面消耗了服务器的可用 IP，另一方面有可能抢夺合法用户的 IP。需要过滤掉接入网络上的非法 DHCP 报文。

以下图为例，DHCP 客户端发送的 DHCP 请求报文将被检查。

- DHCP 客户端的请求报文的源 MAC 字段与 DHCP 报文的客户硬件地址字段必须匹配。
- 客户端的 Release 报文与 Decline 报文必须与 Snooping 内部数据库的记录匹配。

图 14-2



- 【注释】 S 为接入设备。
A、C 为用户 PC。
B 为控制范围内的 DHCP 服务器。

功能部署

- 在接入设备 S 上开启 DHCP Snooping 服务，实现 DHCP 监控。
- 设置接入设备 S 链接 DHCP 服务器 B 的端口为 DHCP TRUST 口，实现响应报文的转发。
- 设置接入设备 S 的其余端口为 DHCP UNTRUST 口，实现 DHCP 报文的过滤。
- 在接入设备 S 上，所有 UNTRUST 口设置 DHCP 源 MAC 检查，过滤非法的 DHCP 报文。

14.2.4 IP/MAC欺骗攻击防范

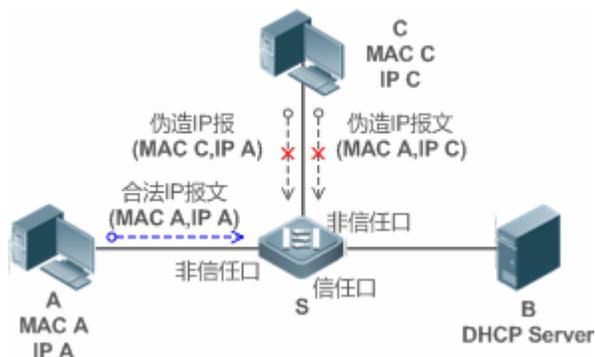
应用场景

检查来自 UNTRUST 口的 IP 报文，可以仅检查 IP 字段，也可以检查 IP+MAC 字段，过滤掉伪造的 IP 报文。

以下图为例，DHCP 客户端发送的 IP 报文将被检查。

- IP 报文的源地址字段必须和 DHCP 分配的 IP 地址匹配。
- 二层报文的源 MAC 字段必须和客户端 DHCP 请求报文中的客户硬件地址匹配。

图 14-3



- 【注释】 S 为接入设备。
A、C 为用户 PC。
B 为控制范围内的 DHCP 服务器。

功能部署

- 在接入设备 S 上开启 DHCP Snooping 服务，实现 DHCP 监控。
- 设置接入设备 S 上所有下行端口为 DHCP UNTRUST 口。
- 在接入设备 S 上，开启 IP Source Guard 功能，实现 IP 报文过滤。
- 在接入设备 S 上，设置 IP Source Guard 的匹配模式为 IP+MAC，实现对 IP 报文 MAC 字段与 IP 字段的检查。

14.2.5 用户私设IP限制

应用场景

检查来自 UNTRUST 口的 IP 报文，检查报文源地址是否和 DHCP 分配的地址一致。

若 IP 报文的源地址、连接端口、二层源 MAC 端口，与设备窥探的 DHCP 服务器分配记录不匹配，则丢弃报文。

该场景下设备的工作过程与上图一致。

功能部署

- 同场景“IP/MAC 欺骗攻击防范”。

14.2.6 ARP入侵检测

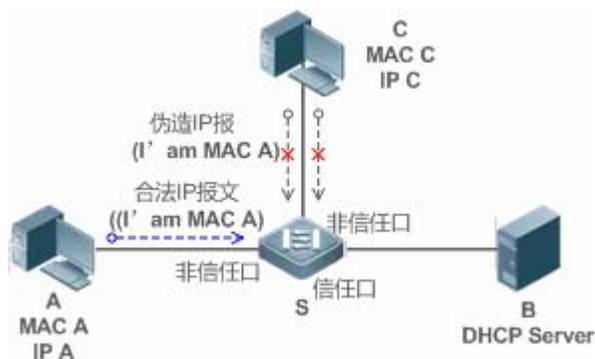
应用场景

检查来自 UNTRUST 口的 ARP 报文，过滤掉与 DHCP 服务器分配记录不匹配的 ARP 报文。

以下图为例，DHCP 客户端发送的 ARP 报文将被检查。

- 接收 ARP 报文的端口、报文的二层 MAC 地址、ARP 报文的发送者硬件地址必须与设备窥探的 DHCP 报文记录一致。

图 14-4



- 【注释】 S 为接入设备。
A、C 为用户 PC。
B 为控制范围内的 DHCP 服务器。

功能部署

- 在接入设备 S 上开启 DHCP Snooping 服务，实现 DHCP 监控。
- 设置接入设备 S 上所有下行端口为 DHCP UNTRUST 口。
- 在接入设备 S 上所有 UNTRUST 口上，开启 IP Source Guard 和 ARP Check 功能，实现 ARP 报文过滤。

! 上述所有安全控制功能仅对 DHCP UNTRUST 口生效。

14.3 功能详解

基本概念

DHCP 请求报文

DHCP 客户端发往 DHCP 服务器的报文。包括 DHCP-DISCOVER 报文、DHCP-REQUEST 报文、DHCP-DECLINE 报文、DHCP-RELEASE 报文及 DHCP-INFORM 报文。

📌 DHCP 应答报文

DHCP 服务器发往 DHCP 客户端的报文。包括 DHCP-OFFER 报文、DHCP-ACK 报文及 DHCP-NAK 报文。

📌 DHCP Snooping TRUST 口

由于 DHCP 获取 IP 的交互报文是使用广播的形式，从而存在着非法的 DHCP 服务影响用户正常 IP 的获取，更有甚者通过非法的 DHCP 服务欺骗窃取用户信息的现象，为了防止非法的 DHCP 服务的问题，DHCP Snooping 把端口分为两种类型，TRUST 口和 UNTRUST 口，设备只转发 TRUST 口收到的 DHCP 应答报文，而丢弃所有来自 UNTRUST 口的 DHCP 应答报文，这样我们把合法的 DHCP Server 连接的端口设置为 TRUST 口，则其他口为 UNTRUST 口，就可以实现对非法 DHCP Server 的屏蔽。

在交换机设备上，所有交换口或者 2 层 AP 口默认均为 UNTRUST 口，可以配置指定 TRUST 口。

📌 DHCP Snooping 报文抑制

在对个别用户禁用 DHCP 报文的情况下，需要屏蔽用户设备发出的任何 DHCP 报文，那么我们可以在 UNTRUST 口配置 DHCP 报文抑制功能，过滤掉该端口收到的所有 DHCP 报文。

📌 基于 VLAN 的 DHCP Snooping

DHCP Snooping 功能生效是以 VLAN 为单位的，默认情况下打开 DHCP Snooping 功能，会在当前设备上的所有 VLAN 上使能 DHCP Snooping 功能，可以通过配置灵活的控制 DHCP Snooping 生效的 VLAN。

📌 DHCP Snooping 绑定数据库

在 DHCP 环境的网络里经常会出现用户随意设置静态 IP 地址的问题，用户随意设置的 IP 地址不但使网络难以维护，而且会导致一些合法的使用 DHCP 获取 IP 的用户因为冲突而无法正常使用网络，DHCP Snooping 通过窥探 Client 和 Server 之间交互的报文，把用户获取到的 IP 信息以及用户 MAC、VID、PORT、租约时间等信息组成用户记录表项，从而形成 DHCP Snooping 的用户数据库，配合 ARP 检测功能或 ARP CHECK 功能的使用，进而达到控制用户合法使用 IP 地址的目的。

📌 DHCP Snooping 速率限制

DHCP Snooping 对 DHCP 报文的速率限制可以选择通过 NFPP 的速率限制命令配置，NFPP 的配置请查看 NFPP 配置指导。

📌 DHCP Option82 选项

DHCP Option82 选项又称为 DHCP 中继代理信息选项（Relay Agent Information Option），是 DHCP 报文中的一个选项。因为其选项编号为 82，故通常被简称为 Option82 选项。Option82 选项是为了增强 DHCP 服务器的安全性，改善 IP 地址的分配策略而提出的一种 DHCP 选项。该选项功能通常配置在网络接入设备的 DHCP 中继服务组件中，如 DHCP Relay、DHCP Snooping。该选项对 DHCP 客户端透明，由 DHCP 中继组件实现选项的添加与剥离。

📌 非法 DHCP 报文

DHCP Snooping 通过对经过设备的 DHCP 报文进行合法性检查，丢弃不合法的 DHCP 报文，记录用户信息并生成 DHCP Snooping 绑定数据库供其他功能（如：ARP 检测功能）查询使用。以下几种类型的报文被认为是非法的 DHCP 报文

- UNTRUST 口收到的 DHCP 应答报文，包括 DHCPACK、DHCPNACK、DHCP OFFER 等。
- UNTRUST 口收到的带有网关信息【giaddr】的 DHCP request 报文。

- 打开 mac 校验时，源 MAC 与 DHCP 报文携带的【chaddr】字段值为不同的报文。
- DHCPRELEASE 报文中的用户在 DHCP Snooping 绑定数据库中存在，但是 DHCPRELEASE 报文的 UNTRUST 口和保存在 DHCP Snooping 绑定数据库中的 UNTRUST 口不一致，那么这个 DHCPRELEASE 报文是非法的。
- DHCP 报文格式不正确或是不完整的报文。

功能特性

功能特性	作用
过滤非法DHCP报文	对交互的 DHCP 报文进行合法性检查，丢弃那些非法报文（非法报文的介绍见上节的介绍），仅向 TRUST 口转发合法的请求报文。
建立Binding数据库	窥探 DHCP 客户端与服务器的交互，生成 DHCP Snooping Binding 数据库，为其他安全过滤模块提供依据。

14.3.1 过滤非法DHCP报文

对来自 UNTRUST 口的 DHCP 报文进行合法性检查。依据上节“基本概念”中介绍的非法报文类型，进行过滤。控制报文的传播范围，防止恶意用户欺骗。

工作原理

窥探过程中，检查报文的接收端口、报文字段，达到过滤报文目的；修改报文的端口，达到控制报文传播范围的目的。

▾ 端口检查

接收到 DHCP 报文时，设备先判断接收报文的端口是否为 DHCP TRUST 口。若是 TRUST 口，跳过合法性检查、Binding 记录生成阶段，直接进入报文转发阶段。若是 UNTRUST 口，需要进行合法性检查。

▾ 检查报文封装及长度是否完整

设备检查报文是否为 UDP 报文，且目的端口为 67 或 68。检查数据包的实际长度与协议中的长度字段是否匹配。

▾ 检查 DHCP 报文字段及报文类型是否正确

依据上节“基本概念”中介绍的非法报文类型，先检查报文的【giaddr】、【chaddr】字段，再依据报文的实际类型，检查该类型特有的限制条件是否满足。

相关配置

▾ 启动全局 DHCP Snooping 功能

缺省情况下，DHCP Snooping 功能关闭。

使用 `ip dhcp snooping` 命令可以启动设备的 DHCP Snooping 功能。

必须首先开启全局 DHCP Snooping 功能，才能进一步在不同 VLAN 上启停 DHCP Snooping 功能。

设置 VLAN 上的 DHCP Snooping 功能

缺省情况下，当全局 DHCP Snooping 功能生效时，DHCP Snooping 功能对所有 VLAN 生效。

使用 `[no] ip dhcp snooping vlan` 命令可以配置 DHCP Snooping 在某个 VLAN 上生效，或将该 VLAN 从 DHCP Snooping 生效的 VLAN 范围中去除。命令参数的取值范围为实际的 VLAN 编号范围。

配置 DHCP 源 MAC 检查功能

缺省情况下，设备不对报文的二层源 MAC 及 DHCP 报文的【chaddr】字段进行校验。

使用 `ip dhcp snooping verify mac-address` 命令，设备就会对 UNTRUST 口送上来的 DHCP Request 报文进行源 MAC 和【chaddr】字段的 MAC 地址进行校验检查，丢弃 MAC 值不相同的 DHCP 请求报文。

14.3.2 建立Binding数据库

窥探 DHCP 客户端与 DHCP 服务器的交互报文，依据合法 DHCP 报文信息，生成 DHCP Snooping Binding 表项。所有这些表项作为合法用户的信息表，提供给设备的其他安全模块使用，作为网络报文过滤的依据。

工作原理

窥探过程中，依据 DHCP 报文的类型，不断更新 Binding 数据库。

生成 Binding 记录

窥探到 TRUST 口上的 DHCPACK 报文时，提取出报文中的客户端 IP 地址、客户端 MAC 地址、租约时间字段，结合设备记录的客户端所在端口 ID（有线接口索引）、客户端所属 VLAN，生成一条 Binding 记录。

删除 Binding 记录

记录的租约时间到期；或是窥探到客户端发送的合法 DHCP-RELEASE/DHCP-DECLINE 报文时；或是接收到来自 TRUST 口的 NAK 报文时；或是用户使用 clear 命令主动删除 Binding 记录时，删除对应的 Binding 记录。

相关配置

无需额外配置，只需要开启 DHCP Snooping 功能即可。

14.4 配置详解

配置项	配置建议 & 相关命令	
配置 DHCP Snooping 基本功能	 必选配置。用于建立 DHCP Snooping 服务。	
	<code>ip dhcp snooping</code>	启动 DHCP Snooping 功能
	<code>ip dhcp snooping suppression</code>	启动 DHCP 报文抑制功能

	ip dhcp snooping vlan	开关指定 VLAN 的 DHCP Snooping 功能
	ip dhcp snooping verify mac-address	配置 DHCP 源 MAC 检查功能
	ip dhcp snooping database write-delay	启动 DHCP Snooping Binding 记录定时保存功能
	ip dhcp snooping database write-to-flash	手动保存 DHCP Snooping Binding 记录
	renew ip dhcp snooping database	手动将保存在 flash 中的用户记录导入到 DHCP Snooping Binding 数据库中
	ip dhcp snooping trust	配置 DHCP Snooping TRUST 口
	ip dhcp snooping bootp	启动支持 bootp 功能
配置Option82 选项	 可选配置。用于优化 DHCP 服务器地址分配。	
	ip dhcp snooping Information option	在 DHCP 请求报文中加入 Option82 选项功能
	ip dhcp snooping information option format remote-id	设置 Option82 选项的子选项 remote-id 为自定义字符串的功能
	ip dhcp snooping vlan information option format-type circuit-id string	设置 Option82 选项的子选项 circuit-id 为自定义字符串的功能

14.4.1 配置DHCP Snooping基本功能

配置效果

- 开启 DHCP Snooping 服务。
- 生成 DHCP Snooping Binding 数据库。
- 控制 DHCP 报文的传播范围。
- 过滤非法的 DHCP 报文。

注意事项

- 设备连接可信 DHCP 服务器的端口必须设置成 DHCP TRUST 口。
- DHCP Snooping 生效的端口可以是有线的交换口、2 层 AP 口或者 2 层封装子接口，端口下的配置分为接口模式下的配置。

配置方法

📌 启动全局 DHCP Snooping 服务

- 必须配置。
- 若无特殊要求，应在接入设备上配置该功能。

按 VLAN 开关 DHCP Snooping 功能

- 如果有些 VLAN 不需要 DHCP Snooping 功能，可以关闭。
- 若无特殊要求，应在接入设备上配置该功能。

配置 DHCP TRUST 口

- 必须配置。
- 将设备连接可信 DHCP 服务器的端口设置成 DHCP TRUST 口。

启动 DHCP 源 MAC 地址检查

- 如果要求 DHCP 请求报文的【chaddr】字段必须与数据包的二层源 MAC 地址匹配，则必须配置。
- 若无特殊要求，应在接入设备的所有 UNTRUST 口上开启该功能。

启动 DHCP Snooping Binding 记录定时保存功能

- 如果要求设备重启后，之前窥探的 DHCP Snooping Binding 记录任然能够生效，需要启动该功能。
- 若无特殊要求，应在接入设备上开启该功能。

启动支持 BOOTP 功能

- 可选配置。
- 若无特殊要求，应在接入设备上开启该功能。

检验方法

用户配置设备使用 DHCP 协议获取网络配置参数。

- 检查设备上的 DHCP Snooping Binding 数据库是否生成相应用户记录。

相关命令

配置打开和关闭 DHCP Snooping

【命令格式】 [no] ip dhcp snooping

【参数说明】 -。

【命令模式】 全局配置模式

【使用指导】 打开 DHCP Snooping 全局功能后，可以使用 **show ip dhcp snooping** 命令查看 DHCP Snooping 功能是否打开。

配置 DHCP Snooping 功能生效的 VLAN

【命令格式】 [no] ip dhcp snooping vlan { vlan-rng | {vlan-min [vlan-max] } }

【参数说明】 vlan-rng : DHCP Snooping 功能生效的 vlan 范围。

vlan-min : DHCP Snooping 功能生效的 vlan 下限。

vlan-max : DHCP Snooping 功能生效的 vlan 上限。

- 【命令模式】 全局配置模式
- 【使用指导】 通过配置该命令，将在指定的 VLAN 内打开 DHCP Snooping 功能，也可关闭指定 VLAN 的 DHCP Snooping 功能。该功能必须在打开 DHCP Snooping 全局开关的基础上生效。

配置端口 DHCP 报文抑制

- 【命令格式】 `[no] ip dhcp snooping suppression`
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 通过配置该命令，可拒绝该端口下所有 DHCP 请求报文，即禁止该端口下的所有用户通过 DHCP 方式申请地址。

配置 DHCP 源 MAC 检查功能

- 【命令格式】 `[no] ip dhcp snooping verify mac-address`
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 源 MAC 地址检验功能，是对 DHCP CLIENT 发出的请求报文，检查链路层头部 MAC 地址和 DHCP 报文中的 CLIENT MAC 字段是否相同。源 MAC 地址检验失败时，报文将被丢弃。

配置定时写 DHCP Snooping 数据库信息到 flash

- 【命令格式】 `[no] ip dhcp snooping database write-delay [time]`
- 【参数说明】 *time* : 两次将 DHCP Snooping 数据库写入 FLASH 的时间间隔。
- 【命令模式】 全局配置模式
- 【使用指导】 通过配置该命令，可以将 DHCP Snooping 数据库写入 FLASH 文件。可以防止设备重新启动后，用户信息丢失，导致用户必须重新获取 IP 地址，才可以正常通讯。

手动把 DHCP Snooping 数据库信息写到 flash

- 【命令格式】 `ip dhcp snooping database write-to-flash`
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 通过执行此命令，可以实时将 DHCP Snooping 数据库中动态用户信息写入 FLASH 文件。

手动地把当前 flash 中的信息导入 DHCP Snooping 绑定数据库

- 【命令格式】 `renew ip dhcp snooping database`
- 【参数说明】 -
- 【命令模式】 特权模式
- 【使用指导】 通过执行此命令，可以实时将 flash 文件信息导入 DHCP Snooping 数据库中。

配置端口为 TRUST 口

- 【命令格式】 `[no] ip dhcp snooping trust`
- 【参数说明】 -

- 【命令模式】 接口配置模式
- 【使用指导】 通过配置该命令，将连接合法 DHCP 服务器的端口配置为 TRUST 口。TRUST 端口收到的 DHCP 响应报文被正常转发，UNTRUST 端口收到的 DHCP 响应报文将被丢弃。

配置支持 BOOTP 功能

- 【命令格式】 [no] ip dhcp snooping bootp
- 【参数说明】 -
- 【命令模式】 全局配置模式
- 【使用指导】 通过配置该命令，可支持 BOOTP 协议。

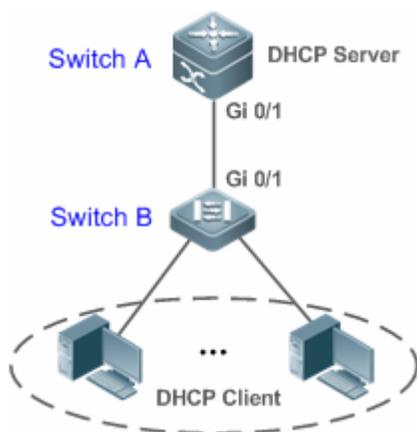
配置举例

i 以下配置举例，仅介绍与 DHCP Snooping 相关的配置。

配置 DHCP 客户端用户通过合法 DHCP 服务器动态获取 IP 地址

【网络环境】

图 14-5



- 【配置方法】
- 在接入设备（本例为 Switch B）上开启 DHCP Snooping 功能
 - 将上链口（本例为端口 Gi 0/1）设置为 TRUST 口。

B

```
B#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
B(config)#ip dhcp snooping
B(config)#interface gigabitEthernet 0/1
B(config-if-GigabitEthernet 0/1)#ip dhcp snooping trust
B(config-if-GigabitEthernet 0/1)#end
```

【检验方法】

- 确认 Switch B 的配置。
- 是否开启 DHCP Snooping 功能、配置的 DHCP Snooping TRUST 口是否为上链口。
 - 查看 Switch B 的 DHCP Snooping 配置情况，关注点为 TRUST 口是否正确。

B

```
B#show running-config
!
```

```
ip dhcp snooping
!
interface GigabitEthernet 0/1
B#show ip dhcp snooping
Switch DHCP Snooping status           :   ENABLE
DHCP Snooping Verification of hwaddr status :   DISABLE
DHCP Snooping database write-delay time  :   0 seconds
DHCP Snooping option 82 status          :   DISABLE
DHCP Snooping Support BOOTP bind status  :   DISABLE
Interface                               Trusted   Rate limit (pps)
-----
GigabitEthernet 0/1                     YES      unlimited
B#show ip dhcp snooping binding
Total number of bindings: 1
MacAddress      IPAddress      Lease(sec)   Type          VLAN   Interface
-----
0013.2049.9014  172.16.1.2   86207        dhcp-snooping 1      GigabitEthernet 0/11
```

常见错误

- 没有将上链口设置为 DHCP TRUST 口。
- 在上链口上配置了其他的接入安全选项，导致配置 DHCP TRUST 口失败。

14.4.2 配置Option82 选项

配置效果

- 让 DHCP 服务器在进行地址分配时，能够获取更多的信息，做出更佳地址分配。
- 选项对 DHCP 客户端透明，客户端无法感知到功能的开启或关闭。

注意事项

- 与 DHCP Relay 的 Option82 选项功能互斥。

配置方法

- 如果需要施加此项优化，则应该执行此配置项。
- 若无特殊要求，应在已经开启 DHCP Snooping 的接入设备上开启该功能。

检验方法

查看 DHCP Snooping 的配置选项，确保功能成功开启。

相关命令

在 DHCP 请求报文中加入 Option82 选项功能

【命令格式】 [no] ip dhcp snooping information option [standard-format]

【参数说明】 **standard-format** : Option82 选项使用标准格式。

【命令模式】 全局配置模式

【使用指导】 通过配置该命令，将在 DHCP 请求报文中添加 Option82 选项信息，DHCP 服务器根据 Option82 选项信息进行地址分配。

设置 Option82 选项的子选项 remote-id 为自定义字符串

【命令格式】 [no] ip dhcp snooping information option format remote-id { string ASCII-string | hostname }

【参数说明】 **string ASCII-string** : Option82 选项 remote-id 扩展格式内容为自定义字符串。

hostname : Option82 选项 remote-id 扩展格式内容为主机名。

【配置模式】 全局配置模式

【使用指导】 通过配置该命令，设置 DHCP 请求报文中添加 Option82 选项的 remote-id 子选项为自定义内容，DHCP 服务器根据 Option82 选项信息进行地址分配。

设置 Option82 选项的子选项 circuit-id 为自定义字符串

【命令格式】 [no] ip dhcp snooping vlan vlan-id information option format-type circuit-id string ascii-string

【参数说明】 **vlan-id** : DHCP 请求报文所在 VLAN。

ascii-string : Circuit ID 要填充的用户自定义的内容。

【配置模式】 接口配置模式

【使用指导】 通过配置该命令，设置 DHCP 请求报文中添加 Option82 选项的 circuit-id 子选项为自定义内容，DHCP 服务器根据 Option82 选项信息进行地址分配。

配置举例

下面是配置在 DHCP 请求报文中加入 Option82 选项功能的例子。

- 【配置方法】
- 配置 DHCP Snooping 基本功能。略
 - 开启添加 Option82 选项功能。

```
B Ruijie# configure terminal
Ruijie(config)# ip dhcp snooping information option
Ruijie(config)# end
```

【检验方法】 查看 DHCP Snooping 配置。

```

B#show ip dhcp snooping
Switch DHCP Snooping status           : ENABLE
DHCP Snooping Verification of hwaddr status : DISABLE
DHCP Snooping database write-delay time : 0 seconds
DHCP Snooping option 82 status         : ENABLE
DHCP Snooping Support bootp bind status : DISABLE
Interface           Trusted           Rate limit (pps)
-----
GigabitEthernet 0/1      YES           unlimited

```

常见配置错误

- 无

14.5 监视与维护

清除各类信息

 在设备运行过程中执行 **clear** 命令，可能因为重要信息丢失而导致业务中断。

作用	命令
清空 DHCP Snooping 数据库动态用户信息。	clear ip dhcp snooping binding [ip] [mac] [vlan vlan-id] [interface interface-id]

查看运行情况

作用	命令
显示 DHCP Snooping	show ip dhcp snooping
显示 DHCP Snooping 数据库信息	show ip dhcp snooping binding

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

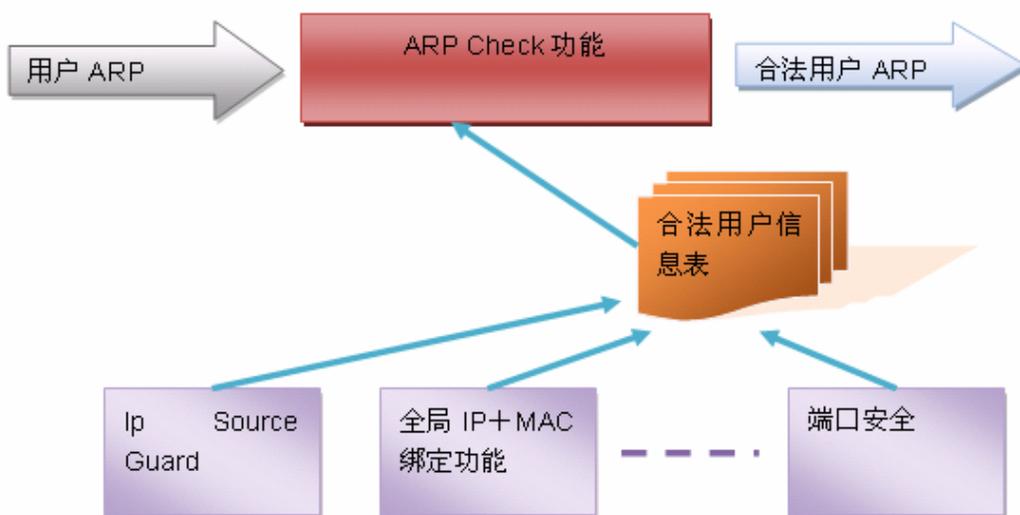
作用	命令
打开 DHCP Snooping 事件的调试开关。	debug snooping ipv4 event
关闭 DHCP Snooping 事件的调试开关。	no debug snooping ipv4 event
打开 DHCP Snooping 报文的调试开关。	debug snooping ipv4 packet
关闭 DHCP Snooping 报文的调试开关。	no debug snooping ipv4 packet

15 ARP Check

15.1 概述

ARP 报文检查 (ARP-Check) 功能，对端口下 (包括有线接入的 2 层交换口、2 层 AP 口或者 2 层封装子接口的所有的 ARP 报文进行过滤，对所有非法的 ARP 报文进行丢弃，能够有效的防止网络中 ARP 欺骗，提高网络的稳定性。在支持 ARP Check 功能的设备中，ARP Check 功能能够根据 IP Source Guard、全局 IP+MAC 绑定、802.1X 认证、GSN 绑定、WEB 认证或者端口安全等安全应用模块所生成的合法用户信息(IP 或 IP+MAC)产生相应的 ARP 过滤信息，从而实现对网络中的非法 ARP 报文的过滤。

图 15-1



如上图所示，设备安全功能模块产生的合法用户信息(仅有 IP 或 IP + MAC)，ARP Check 功能使用这些信息用于检测端口下的所有的 ARP 报文中的 Sender IP 字段或<Sender IP, Sender MAC>是否满足合法用户信息表中的匹配关系，所有不在合法用户信息表中的 ARP 报文将被丢弃。

i 下文仅介绍 ARP Check 的相关内容。

协议规范

- RFC826 : An Ethernet Address Resolution Protocol or Converting Network Protocol Addresses

15.2 典型应用

典型应用	场景描述
过滤网络上的非法ARP报文	网络上存在非法的用户，使用伪造的 ARP 报文进行攻击。

15.2.1 过滤网络上的非法ARP报文

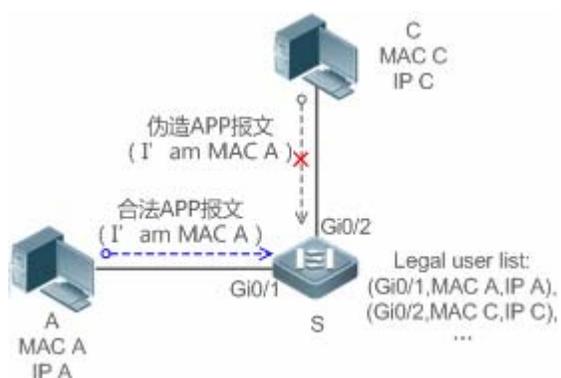
应用场景

检查来自非信任端口的 ARP 报文，过滤掉与 DHCP 服务器分配记录不匹配的 ARP 报文。

以下图为例，DHCP 客户端发送的 ARP 报文将被检查。

- 接收 ARP 报文的端口、ARP 报文的发送者 MAC 地址、ARP 报文的发送者 IP 地址必须与设备窥探的 DHCP 报文记录一致。

图 15-2



【注释】 S 为接入设备。
A、C 为用户 PC。

功能部属

- 在接入设备 S 上开启 DHCP Snooping 服务，实现 DHCP 监控。
- 设置接入设备 S 上所有下行端口为 DHCP 非信任端口。
- 在接入设备 S 上所有非信任端口上，开启 IP Source Guard 与 ARP Check 功能，实现 ARP 报文过滤。

15.3 功能详解

基本概念

↘ ARP Check 支持的安全功能模块

目前 ARP Check 支持的安全功能模块包括：

- 仅检测 IP 字段：端口安全的仅 IP 模式，Ip Source Guard 手工配置的仅 IP 模式。

- 检测 IP+MAC 字段：端口安全的 IP + MAC 绑定模式，全局 IP + MAC 绑定功能，802.1x IP 授权功能，IP Source Guard 功能，GSN 绑定功能，WEB 认证功能。

📌 ARP-Check 两种模式

ARP-Check 有 2 种模式：打开和关闭，默认为关闭。

8. 打开模式

ARP Check 功能根据如下模块提供的 IP/IP+MAC 信息对 ARP 报文的合法性进行检测。

- 全局 IP + MAC 绑定
- 802.1X 的 IP 授权
- IP Source Guard
- GSN 绑定
- 端口安全
- WEB 认证
- 端口安全 IP+Mac 或 IP 绑定

⚠️ 如果端口上仅开启 ARP-Check 功能，而没有开启上述模块提供合法用户信息，将导致来自这个端口的所有 ARP 报文被丢弃。

⚠️ 当接口开启 arp-check 功能时，如果接口同时启用了 vrrp 功能，对于接口的实地址和虚地址都能当网关，需要配置放行接口实 ip 地址和 vrrp ip 地址，否则可能导致发往网关的 arp 报文被过滤。

9. 关闭模式

不检查端口上的 ARP 报文。

功能特性

功能特性	作用
非法ARP报文过滤	检查 ARP 报文的源 IP 与源 MAC 字段，达到过滤非法 ARP 报文的目的。

15.3.1 非法ARP报文过滤

在指定端口上开启 ARP 检查功能，达到过滤非法 ARP 报文的目的。

工作原理

设备把端口下接收到 ARP 报文的源 IP 与源 MAC 字段，与设备安全数据库中的合法用户记录进行匹配。若匹配成功，则正常转发报文；若匹配失败，则丢弃报文。

相关配置

启动端口上的 ARP Check 功能

缺省情况下，端口上的 ARP Check 功能关闭。

使用 **arp-check** 命令可以启动端口上的 ARP Check 功能。

若无特殊需求，一般在接入设备的端口上设置该功能。

15.4 配置详解

配置项	配置建议 & 相关命令	
配置ARP-Check	 必须配置。用于使能 ARP-Check 服务。	
	arp-check	设置 ARP Check 功能为打开模式

15.4.1 配置ARP-Check

配置效果

- 过滤非法的 ARP 报文。

注意事项

- 打开 ARP Check 检测功能可能会使相关安全应用的策略数/用户数减少。
- 无法在镜像的目的口上配置 arp-check 功能。
- 无法在 DHCP Snooping 信任端口上配置 ARP Check 功能。
- 无法在全局 IP+MAC 的例外口配置 ARP Check 功能。
- 只能在有线的交换口、2层 AP 口、2层封装接口下配置开启，有线接入是在接口模式下配置。

配置方法

启动 ARP Check 功能

- 必选配置。功能默认关闭，如果管理员希望使用 ARP Check 功能，需要输入命令开启。

检验方法

- 使用 **show run** 命令，查看功能配置。
- 使用 **show interface { interface-type interface-number } arp-check list** 命令，查看过滤表项。

相关命令

▾ 开启 ARP 报文检查

- 【命令格式】 **arp-check**
- 【参数说明】 -
- 【命令模式】 接口配置模式
- 【使用指导】 根据安全应用模块的合法用户信息产生相应的 ARP 过滤信息，实现对网络中的非法 ARP 报文的过滤。

配置举例

i 以下配置举例，仅介绍与 ARP Check 相关的配置。

▾ 配置接口 ARP Check 模式为打开模式。

- 【配置方法】
- 开启 ARP Check 功能，限制 ARP 报文必须符合 IP Source Guard、端口安全或是全局 IP+MAC 绑定的表项。

```
Ruijie# configure terminal
Ruijie(config)#address-bind 192.168.1.3 00D0.F800.0003
Ruijie(config)#address-bind install
Ruijie(config)#ip source binding 00D0.F800.0002 vlan 1 192.168.1.4 interface gigabitEthernet 0/1
Ruijie(config)# interface GigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)#arp-check
Ruijie(config-if-GigabitEthernet 0/1)#ip verify source port-security
Ruijie(config-if-GigabitEthernet 0/1)#switchport port-security
Ruijie(config-if-GigabitEthernet 0/1)#switchport port-security binding 00D0.F800.0001 vlan 1
192.168.1.1
Ruijie(config-if-GigabitEthernet 0/1)#exit
Ruijie(config)#interface gigabitEthernet 0/4
Ruijie(config-if-GigabitEthernet 0/4)#switchport port-security
Ruijie(config-if-GigabitEthernet 0/4)#switchport port-security binding 192.168.1.5
Ruijie(config-if-GigabitEthernet 0/4)#arp-check
Ruijie(config-if-GigabitEthernet 0/4)#exit
Ruijie(config)#interface gigabitEthernet 0/5
Ruijie(config-if-GigabitEthernet 0/5)#arp-check
Ruijie(config-if-GigabitEthernet 0/5)#end
```

- 【检验方法】 使用 **show interface arp-check list** 命令，可以查看接口下实际生效的 ARP Check 表项。

```
Ruijie# show interface arp-check list
```

```
INTERFACE                SENDER MAC                SENDER IP                POLICY SOURCE
-----
```

GigabitEthernet 0/1	00d0.f800.0003	192.168.1.3	address-bind
GigabitEthernet 0/1	00d0.f800.0001	192.168.1.1	port-security
GigabitEthernet 0/1	00d0.f800.0002	192.168.1.4	DHCP snooping
GigabitEthernet 0/4	00d0.f800.0003	192.168.1.3	address-bind
GigabitEthernet 0/4		192.168.1.5	port-security
GigabitEthernet 0/5	00d0.f800.0003	192.168.1.3	address-bind

常见配置错误

- 接口需要检查 ARP 报文，但是将接口的 ARP Check 模式设置为关闭模式，导致功能无法生效。

15.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
查看端口下实际生效的 ARP Check 表项。	show interface [<i>interface-type interface-number</i>] arp-check list

查看调试信息

无

16 动态 ARP 检测

16.1 概述

DAI (Dynamic ARP Inspection , 动态 ARP 检测) 对接收到的 ARP 报文进行合法性检查。不合法的 ARP 报文会被丢弃。

DAI 确保了只有合法的 ARP 报文才会被设备转发。它主要执行以下几个步骤：

- 在打开 DAI 检查功能的 VLAN 所对应的非信任端口上拦截住所有 ARP 请求和应答报文
- 在做进一步相关处理之前，根据安全数据库的用户记录，对拦截住的 ARP 报文进行合法性检查。
- 丢弃没有通过检查的报文。
- 检查通过的报文继续做相应的处理，送给相应的目的地。
- ARP 报文是否合法的依据与 ARP Check 功能相同，具体标准请参见《ARP Check 配置指南》。
- DAI 实现的功能与 ARP Check 一致，只是 DAI 的作用范围是以 VLAN 为单位，而 ARP Check 的作用范围是以端口为单位。

协议规范

- RFC826 : An Ethernet Address Resolution Protocol or Converting Network Protocol Addresses

16.2 典型应用

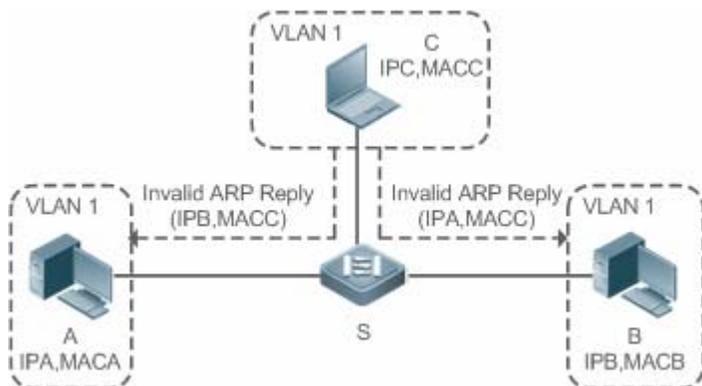
典型应用	场景描述
防范ARP欺骗攻击	防范攻击者利用 ARP 协议自身的缺陷，进行 ARP 欺骗攻击。

16.2.1 防范ARP欺骗攻击

应用场景

由于 ARP 协议本身的缺陷，ARP 协议不对收到的 ARP 报文进行合法性检查。这就造成了攻击者利用协议的漏洞轻易的进行 ARP 欺骗攻击。这其中，最典型的的就是中间人攻击。中间人攻击描述如下：

图 16-1



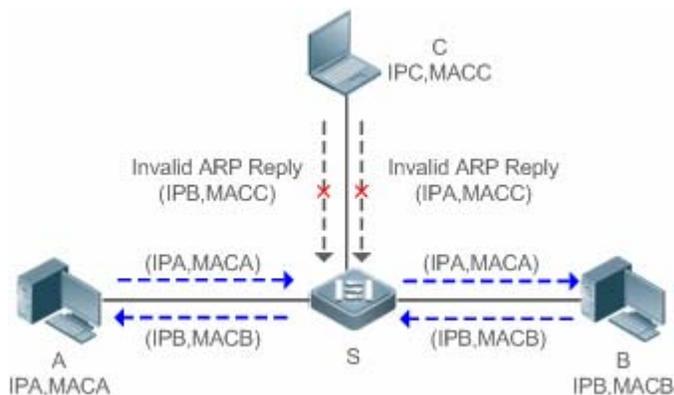
- 【注释】 S 为支持 DAI 功能的锐捷接入交换机。
 A、B 为连接在锐捷设备上的用户，并且它们位于同一个子网。
 C 为连接在锐捷设备上的恶意用户。
 IPA，MACA 分别为用户 A 的 IP 地址与 MAC 地址。
 IPB，MACB 分别为用户 B 的 IP 地址与 MAC 地址。
 IPC，MACC 分别为用户 C 的 IP 地址与 MAC 地址。

当用户 A 需要和用户 B 进行网络层通信时，用户 A 将会在子网内广播一个 ARP 请求，询问用户 B 的 MAC 值。当用户 B 接收到此 ARP 请求报文时，会更新自己的 ARP 缓存，使用的是 IPA 和 MACA，并发出 ARP 应答。用户 A 收到此应答后，会更新自己的 ARP 缓存，使用的是 IPB 和 MACB。

在这种模型下，用户 C 可以使用户 A 和用户 B 中的对应 ARP 表项对应关系不正确。使用的策略是，不断向网络中广播 ARP 应答。此应答使用的 IP 地址是 IPA 和 IPB，而 MAC 地址是 MACC，这样，用户 A 中就会存在 ARP 表项(IPB、MACC)，用户 B 中就会存在 ARP 表项(IPA、MACC)。这样，用户 A 和 B 之间的通信就变成了和用户 C 之间的通信，而用户 A、B 对此都一无所知。用户 C 充当了中间人的角色，只需要把发给自己的报文做合适的修改，转给另一方即可。这就是有名的中间人攻击。

若在设备 S 上开启 DAI 功能，只要用户 A 与用户 B 的，IP 地址和 MAC 地址是根据 1.1 节中提到的方式获得，设备会过滤掉所有伪造的 ARP 报文，达到防范 ARP 欺骗攻击的目的。DAI 工作过程如下图所示：

图 16-2



- 【注释】 S 为支持 DAI 功能的锐捷接入交换机。
 A、B 为连接在锐捷设备上的用户，并且它们位于同一个子网。
 C 为连接在锐捷设备上的恶意用户。

IPA, MACA 分别为用户 A 的 IP 地址与 MAC 地址。

IPB, MACB 分别为用户 B 的 IP 地址与 MAC 地址。

IPC, MACC 分别为用户 C 的 IP 地址与 MAC 地址。

用户 A 与用户 B 的 ARP 报文会被设备 S 正常转发,但是用户 C 伪造的 ARP 报文,由于和设备 S 内部安全数据库的记录不匹配,被丢弃。

功能部署

- 在设备上运行 DHCP Snooping 功能
- 在设备上开启 DAI 功能和 IP Source Guard 功能

16.3 功能详解

基本概念

▾ 接口信任状态和网络安全

基于设备上每一个端口的信任状态,对 ARP 报文做出相应的检查,从受信任端口接收到的报文将跳过 DAI 检查,被认为是合法的 ARP 报文;而非信任端口接收到的 ARP 报文,将严格执行 DAI 检查。

在一个典型的网络配置中,应该将连接到网络设备的二层端口设置为受信任端口,连接到主机设备的二层端口设置为非信任端口。

 将一个连接到网络设备的二层端口配置成非信任端口可能会影响到网络正常通信。

功能特性

功能特性	作用
非法ARP报文过滤	检查 ARP 报文的源 IP 与源 MAC 字段,达到过滤非法 ARP 报文的目的。
DAI信任端口	设置设备指定端口放行所有 ARP 报文。

16.3.1 非法ARP报文过滤

在指定 VLAN 上开启 ARP 报文动态检查功能,达到过滤非法 ARP 报文的目的,判断 ARP 报文是否合法的依据与 ARP Check 是相同的。

工作原理

设备把接口接收到 ARP 报文的源 IP 与源 MAC 字段,与设备安全数据库中的合法用户记录进行匹配。若匹配成功,则正常转发报文;若匹配失败,则丢弃报文。

DAI 检查依赖的合法用户记录数据来源与 ARP Check 功能一致，详细情况可参考 ARP Check 报文合法性检查的介绍。

相关配置

启动 VLAN 上的 DAI 功能

缺省情况下，VLAN 上的 DAI 功能关闭。

使用 `ip arp inspection vlan vlan-id` 命令可以启动 VLAN 上的 DAI 功能。

 开启 DAI 功能时，并非 VLAN 下所有的端口都能生效。当端口为 DHCP Snooping 信任口时，是无法在该端口上施加 DAI 检查的。

关闭 VLAN 上的 DAI 功能

缺省情况下，VLAN 上的 DAI 功能关闭。

当不再需要 DAI 功能时，使用 `no ip arp inspection vlan vlan-id` 命令可以关闭指定 VLAN 上的 DAI 功能。

 关闭 DAI 功能时，并非 VLAN 下所有的端口都能生效。当端口的 ARP Check 功能仍然生效时，设备还是会检查经过端口的 ARP 报文。

16.3.2 DAI信任端口

设置设备指定端口为 DAI 信任端口。

工作原理

如果端口是可信任的，ARP 报文将跳过进一步的检查；否则，会使用安全数据库的信息来检查当前 ARP 报文的合法性。

相关配置

设置 DAI 信任口

缺省情况下，所有端口均为非信任端口。

使用 `ip arp inspection trust` 命令可以设置端口为可信任的。

 若端口上已经开启了其他的接入安全控制命令，则无法将端口设置成信任端口。若一定要将端口设置成信任端口，需先关闭已设置的安全控制命令。

 一般将上行接口（连接网络设备的接口）设置成 DAI 信任口。

16.4 配置详解

配置项	配置建议 & 相关命令
-----	-------------

配置DAI功能	 可选配置。用于开启合法 ARP 报文检查功能。	
	ip arp inspection vlan	启动 DAI 功能
	ip arp inspection trust	设置 DAI 信任端口

16.4.1 配置DAI功能

配置效果

- 设置指定 VLAN 对输入的 ARP 报文进行合法性检查

注意事项

- VLAN 上的 DHCP Snooping 信任口无法开启 DAI 功能。

配置方法

启动 VLAN 上 ARP 报文检查功能

- 可选配置。
- 需要在 VLAN 上所有端口开启 ARP 报文检查功能时，进行配置。
- 若无特殊要求，在锐捷接入设备上配置该功能。

配置 DAI 信任端口

- 可选配置。
- 开启 DAI 功能后，建议将上行接口配置成 DAI 信任端口，否则很可能会因为上行接口是其他安全功能的信任端口，没有 DAI 需要的用户表项，导致 ARP 报文被过滤。
- 若无特殊要求，在锐捷接入设备上配置该功能。

配置 ARP 报文接收速率

- 详见 NFPP 的速率限制命令

检验方法

- 使用发包工具，构造非法 ARP 报文，检查报文是否能够经过设备。
- 使用查看命令，查看设备的配置情况

相关命令

启动 DAI 功能

【命令格式】 **ip arp inspection vlan { vlan-id | word }**

【参数说明】 *vlan-id* : 指定 VLAN 编号。
word : vlan 范围字符串, 如 1,3-5,7,9-11。

【命令模式】 全局模式

【使用指导】 -

配置 DAI 信任端口

【命令格式】 **ip arp inspection trust**

【参数说明】 -

【命令模式】 接口模式

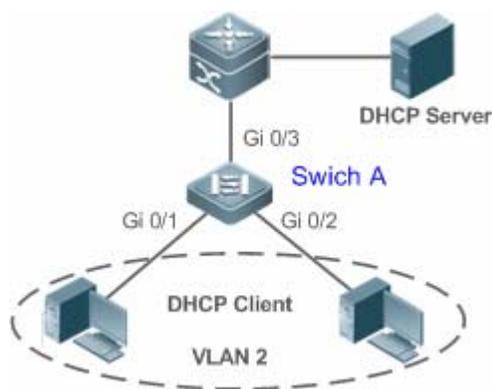
【使用指导】 当需要使某个接口收到的 ARP 报文无条件的通过 DAI 的检查时, 可以将其设置成受信任状态, 表示我们不需要检查此接口收到的 ARP 报文, 它们是合法的。

配置举例

保证用户 PC 只能使用合法 DHCP 服务器分配的地址, 防止 ARP 欺骗

【网络环境】

图 16-3



【配置方法】

- ⚠ 在接入交换机（本例为 Switch A）上启用 DHCP Snooping 并将连接合法 DHCP 服务器的上链口（本例为 GigabitEthernet 0/3）设置为信任口。
- ⚠ 接入设备上必须启用 IP Source Guard 功能
- ⚠ 在接入设备（本例为 Switch A）启用 DHCP Snooping、IP Source Guard 基础上再开启 DAI。

A

```
A#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
A(config)#vlan 2
A(config-vlan)#exit
A(config)#interface range gigabitEthernet 0/1-2
A(config-if-range)#switchport access vlan 2
A(config-if-range)#ip verify source
```

```
A(config-if-range)#exit
A(config)#ip dhcp snooping
A(config)#ip arp inspection vlan 2
A(config)#interface gigabitEthernet 0/3
A(config-if-GigabitEthernet 0/3)#switchport access vlan 2
A(config-if-GigabitEthernet 0/3)#ip dhcp snooping trust
A(config-if-GigabitEthernet 0/3)#ip arp inspection trust
```

- 【检验方法】
- 确认配置是否正确，关注点为 DHCP Snooping/IP Source Guard /DAI 是否启用，信任接口是否正确。
 - 查看 DHCP Snooping 对应的信任端口，关注点为上链口是否设置为可信任接口。
 - 查看 DAI 状态，关注点为对应的 VLAN 的使能情况和上链口是否设置为可信任接口。

```
A
A#show running-config
A#show ip dhcp snooping
A#show ip arp inspection vlan
```

常见错误

- 将已经配置了安全限制的端口设置为 DAI 信任口

16.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
显示 VLAN 是否启用 DAI 功能	show ip arp inspection vlan [vlan-id word]
显示各二层接口 DAI 配置状态	show ip arp inspection interface

查看调试信息

无

17 IP Source Guard

17.1 概述

- i** 通过 IP Source Guard 绑定功能，可以通过硬件对 IP 报文进行过滤，从而保证只有 IP 报文硬件过滤数据库中存在对应信息的用户才能正常使用网络，防止了用户私设 IP 地址及伪造 IP 报文。下文仅介绍 IP Source Guard 的相关内容。

协议规范

- 无

17.2 典型应用

典型应用	场景描述
IP/MAC欺骗攻击防范	在网络环境中，防止用户私设 IP 地址或防止用户伪造 IP 报文进行攻击。

17.2.1 IP/MAC欺骗攻击防范

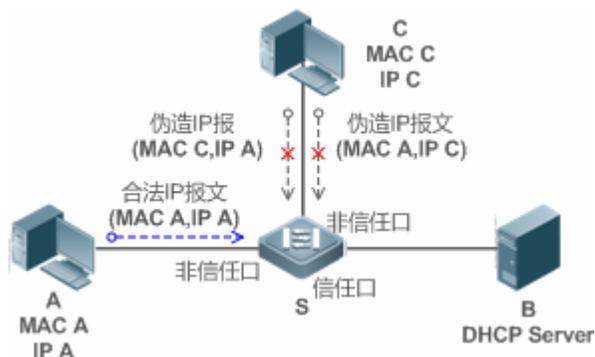
应用场景

检查来自非 DHCP 信任口的 IP 报文，可以仅检查 IP 字段，也可以检查 IP+MAC 字段，过滤掉伪造的 IP 报文。

以下图为例，DHCP 客户端发送的 IP 报文将被检查。

- IP 报文的源地址字段必须和 DHCP 分配的 IP 地址匹配。
- 二层报文的源 MAC 字段必须和客户端 DHCP 请求报文中的客户硬件地址匹配。

图 17-1



【注释】 S 为接入设备。
A、C 为用户 PC。

B 为控制范围内的 DHCP 服务器。

功能部署

- 在接入设备 S 上开启 DHCP Snooping 服务，实现 DHCP 监控。
- 设置接入设备 S 上所有下行接口为 DHCP 非信任口。
- 在接入设备 S 上，开启 IP Source Guard 功能，实现 IP 报文过滤。
- 在接入设备 S 上，设置 IP Source Guard 的匹配模式为 IP+MAC，实现对 IP 报文 MAC 字段与 IP 字段的检查。

17.3 功能详解

基本概念

源 IP

用户 IP 报文的源 IP 地址字段。

源 MAC

用户二层报文的源 MAC 地址字段。

基于源 IP 的过滤

IP 报文的过滤策略，检查所有经过该接口的 IP 报文（DHCP 报文除外），仅对报文的源 IP 地址进行检测。IP Source Guard 的缺省过滤策略。

基于源 IP + 源 MAC 的过滤

IP 报文的过滤策略，会对所有 IP 报文的源 IP + 源 MAC 进行检查，仅允许绑定用户记录数据库中存在的用户报文通过。

绑定用户记录数据库

IP Source Guard 安全控制的依据。目前，绑定用户记录数据库中数据来自两个方面。一方面数据来自 DHCP Snooping 绑定数据库，当启动 IP Source Guard 功能后，DHCP Snooping 数据库信息将同步到 IP Source Guard 的绑定用户数据库中，这样 IP Source Guard 就可以在打开 DHCP Snooping 功能的设备上对客户端的 IP 报文进行严格过滤。另一方面数据来自用户的静态配置。

例外 VLAN

默认情况下端口开启 IP Source Guard 后，会对该端口包含的所有 VLAN 生效，用户可以指定例外 VLAN 不对该 VLAN 范围内的 IP 报文进行检查和过滤，即不受 IP Source Guard 的控制，每个端口最多可以指定 32 个例外 VLAN。

功能特性

功能特性	作用
检查报文源地址字段	对经过接口的 IP 报文进行基于源 IP 过滤，或是基于源 IP + 源 MAC 的过滤。

17.3.1 检查报文源地址字段

对经过端口的 IP 报文进行基于源 IP 过滤，或是基于源 IP + 源 MAC 的过滤。防止恶意用户伪造报文进行攻击。当用户不需要检查和过滤某 VLAN 范围内的 IP 报文时，可以指定例外 VLAN 对报文进行放行。

工作原理

打开 IP Source Guard 功能后，设备对经过端口的报文进行源地址检查，端口可以是有线接入的交换口、2 层 AP 口或者 2 层封装子接口。只有源地址字段和 DHCP Snooping 生成的绑定用户记录集匹配，或是和管理员静态配置的用户集匹配的报文才能经过端口。匹配的方式有两种：

↳ 基于源 IP 地址过滤

只要报文的源 IP 字段属于绑定用户记录中的 IP 地址集合，就可以通过端口。

↳ 基于 IP+MAC 地址过滤

报文的二层源 MAC 与三层源 IP 必须和合法用户集中的某条记录完全匹配上，才能通过端口。

↳ 指定例外 VLAN

该 VLAN 范围的报文不被检查和过滤，直接通过端口。

相关配置

↳ 启动端口上的 IP Source Guard 功能

缺省情况下，端口上的 IP Source Guard 功能关闭。

使用 **ip verify source** 命令可以启动或关闭端口上的 IP Source Guard 功能。

- i** 通常 IP Source Guard 功能需要 DHCP Snooping 功能的配合，因此，还需要启动 DHCP Snooping 功能。锐捷设备对启动 DHCP Snooping 功能的时机不做限制，用户可以在 IP Source Guard 启动之前或之后启动 DHCP Snooping。

↳ 配置静态 IP Source Guard 用户

缺省情况下，IP Source Guard 检查的合法用户集全部来自 DHCP Snooping 的绑定用户。

使用 **ip source binding** 命令可以添加额外的绑定用户记录。

↳ 端口上指定 IP Source Guard 的例外 VLAN

缺省情况下，IP Source Guard 对端口上包含的所有 VLAN 生效。

使用 **ip verify source** 命令可以指定例外 VLAN 不受 IP Source Guard 的控制。

-  与端口下的 IP Source Guard 配合使用，端口下必须先启动 IP Source Guard 才可以指定例外 VLAN，端口下关闭 IP Source Guard 后会自动清除指定的例外 VLAN。
-  端口可以是有线接入的交换口、2 层 AP 口或者 2 层封装子接口。

17.4 配置详解

配置项	配置建议 & 相关命令	
配置IP Source Guard	 必须配置。用于开启 IP Source Guard 服务。	
	ip verify source	启动端口上的 IP Source Guard 功能。
	ip source binding	配置静态绑定用户。
	ip verify source exclude-vlan	端口上指定 IP Source Guard 的例外 VLAN

17.4.1 配置IP Source Guard

配置效果

- 对输入 IP 报文进行检查，过滤非法 IP 报文。

注意事项

- 打开 IP Source Guard 功能可能会影响 IP 报文的转发，一般情况下，该功能需要结合 DHCP Snooping 功能使用。
- 无法在 DHCP Snooping 信任端口上配置 IP Source Guard 功能。
- 无法在全局 IP+MAC 的例外口配置 IP Source Guard 功能。
- 只能在有线的交换口、2 层 AP 口、2 层封装子接口下配置开启，有线接入是在接口模式下配置。

配置方法

- 开启 DHCP Snooping。
- 开启 IP Source Guard。

检验方法

使用设备提供的监控命令，查看 IP Source Guard 用户过滤表项。

相关命令

打开端口上的 IP Source Guard 功能

【命令格式】 **ip verify source [port-security]**

【参数说明】 **port-security**：配置 IP Source Guard 功能进行基于 IP+MAC 检测。

【命令模式】 接口模式

【使用指导】 通过该命令打开端口的 IP Source Guard 功能，可以对用户进行基于 IP 的检测，或者进行基于 IP+MAC 的检测。

在 IP 源地址绑定数据库中添加静态用户信息

【命令格式】 **ip source binding mac-address vlan vlan-id ip-address { interface interface-id | ip-mac | ip-only }**

【参数说明】 *mac-address*：静态添加的用户的 MAC 地址。

vlan-id：静态添加的用户的 vlan id。

ip-address：静态添加的用户的 IP 地址。

interface-id：静态添加的用户所属的有线接口。

ip-mac：全局绑定的类型为 IP+MAC 绑定。

ip-only：全局绑定的类型为仅 IP 绑定。

【配置模式】 全局模式

【使用指导】 通过配置此命令可以允许部分用户通过 IP Source Guard 的检测，不需要通过 DHCP 方式进行统一控制。

接口上指定 IP Source Guard 的例外 VLAN

【命令格式】 **ip verify source exclude-vlan vlan-id**

【参数说明】 *vlan-id*：不受接口上 IP Source Guard 控制的 vlan id。

【命令模式】 接口模式

【使用指导】 启动 IP Source Guard 的接口上，通过该命令可以控制该端口的某些 VLAN 不受 IP Source Guard 控制，这些 VLAN 范围内的 IP 报文不被检查和过滤而是直接放行。

配置举例

配置打开接口 1 的 IP Source Guard 功能。

- 【配置方法】
- 打开 DHCP Snooping 功能。略
 - 开启 IP Source Guard。

```
Ruijie(config)# interface GigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# ip verify source
Ruijie(config-if-GigabitEthernet 0/1)# end
```

【检验方法】 查看 IP Source Guard 用户过滤表项

```
Ruijie# show ip verify source
```

添加一个静态绑定用户。

- 【配置方法】
- 打开 DHCP Snooping 功能。略
 - 开启 IP Source Guard。略
 - 添加静态用户

```
Ruijie# configure terminal
Ruijie(config)# ip source binding 00d0.f801.0101 vlan 1 192.168.4.243 interface GigabitEthernet 0/3
Ruijie(config)# end
```

【检验方法】 查看 IP Source Guard 用户过滤表项

```
Ruijie# show ip verify source
```

NO.	INTERFACE	FilterType	FilterStatus	IPADDRESS	MACADDRESS
VLAN TYPE					
1	GigabitEthernet 0/3 00d0.f801.0101 1 Static	UNSET	Inactive-restrict-off	192.168.4.243	
2	GigabitEthernet 0/1	IP-ONLY	Active	Deny-All	

配置打开端口的 IP Source Guard 功能并指定例外 VLAN。

- 【配置方法】
- 打开 DHCP Snooping 功能。略
 - 开启 IP Source Guard。

```
Ruijie(config)# interface GigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# ip verify source
Ruijie(config-if-GigabitEthernet 0/1)# ip verify exclude-vlan 1
Ruijie(config-if)# end
```

【检验方法】 查看接口上指定的例外 VLAN

```
Ruijie# show run
```

常见配置错误

- 在 DHCP Snooping 信任口开启 IP Source Guard。
- 未启动 IP Source Guard 就指定例外 VLAN。

17.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
查看 IP Source Guard 用户过滤表项。	show ip verify source [interface <i>interface-id</i>]
查看 IP 源地址绑定数据库的信息	show ip source binding

查看调试信息

无

18 NFPP

18.1 概述

网络基础保护策略 (Network Foundation Protection Policy), 简称 NFPP, 提供交换机防攻击功能。

在网络环境中经常发现一些恶意的攻击, 这些攻击会给交换机带来过重的负担, 引起交换机 CPU 利用率过高, 导致交换机无法正常运行。这些攻击具体表现在:

拒绝服务攻击可能导致大量消耗交换机内存、表项或者其它资源, 使系统无法继续服务。

大量的报文流砸向 CPU, 占用了整个送 CPU 的报文的带宽, 导致正常的协议流和管理流无法被 CPU 处理, 带来协议震荡或者无法管理, 从而导致数据面的转发受影响, 并引起整个网络无法正常运行。

大量的报文砸向 CPU 会消耗大量的 CPU 资源, 使 CPU 一直处于高负载状态, 从而影响管理员对设备进行管理或者设备自身无法运行。

NFPP 可以有效地防止系统受这些攻击的影响。在受攻击情况下, 保护系统各种服务的正常运行, 以及保持较低的 CPU 负载, 从而保障了整个网络的稳定运行。

18.2 典型应用

典型应用	场景描述
攻击检测限速	网络中存在各种的恶意攻击, 如 ARP 攻击, IP 扫描攻击等。导致正常的协议流和管理流无法被 CPU 处理, 带来协议震荡或者无法管理。采用 NFPP 的攻击检测限速功能, 可限速或隔离攻击流, 使网络恢复正常。
集中限速分发	正常业务流量太大, 这时给流分处理优先级。当大量的报文砸向 CPU, 使 CPU 一直处于高负载状态, 出现管理员无法对设备进行管理或者设备自身无法运行的情况。采用集中限速分发功能, 提高该部分处理优先级, 确保交换机自身稳定运行。

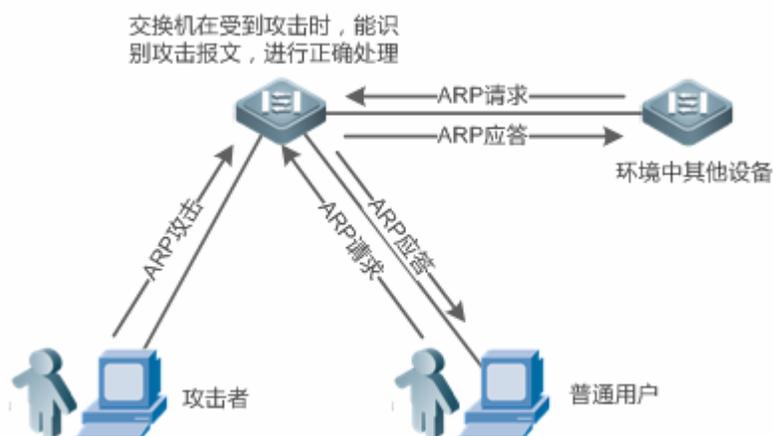
18.2.1 攻击检测限速

应用场景

NFPP 支持对多种报文的攻击检测限速, 包括 ARP、ICMP、DHCP 等; 同时还支持用户自己定义报文匹配特征, 以及对应的攻击检测限速策略。攻击检测限速功能是基于每种报文匹配生效的, 这里以 ARP 为例进行应用场景分析。

当存在一个攻击者, 进行 ARP 报文攻击, 由于 ARP 报文会送 CPU 处理, 而 CPU 处理能力有限, 因此, 该 ARP 会导致 CPU 资源耗费, 大量资源用于处理攻击者的 ARP 报文。若攻击者的 ARP 报文流量超过了交换机的 CPP(CPU Protect Policy, CPU 保护策略)的 ARP 限速带宽, 正常的 ARP 报文将出现丢包。针对图中场景, 将会导致: 普通用户无法上网, 交换机与环境中其他设备无法进行正常 ARP 应答。

图 18-1



功能部属

- 在缺省情况下，ARP 攻击检测限速功能打开，且配置了攻击检测限速策略。攻击者的 ARP 报文超过限速水线，报文将被丢弃，若超过攻击水线，则另生成监控用户，同时输出提示信息。
- 若攻击者的 ARP 报文流量很大，已超出 CPP 的限速线，影响正常 ARP 应答，用户可开启隔离功能，硬件丢弃 ARP 攻击报文，网络恢复正常。

i 针对 CPP 相关配置说明，请参考“CPP”章节。

i 为了最大限度地利用 NFPP 中抗攻击功能，请根据具体的应用环境修改 CPU Protect Policy 中各种服务的限速值，也可以使用系统提供的推荐配置，这些推荐值可以通过命令 **show cpu-protect summary** 查看。

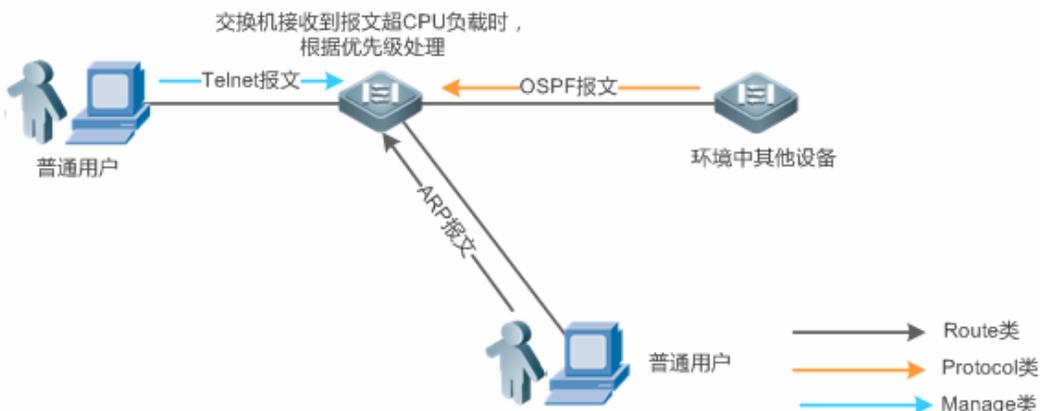
18.2.2 集中限速分发

应用场景

交换机将 CPP 中定义的各种服务按照管理类(Manage)、转发类(Route)和协议类(Protocol)的原则进行的分类，每一类都拥有独立的带宽，不同类别之间的带宽不能共享，超过带宽阈值的流将被丢弃。这样将不同的服务区分类别后，可以保证属于某类的各种服务报文在设备上得到优先处理。

在下图的应用场景中，交换机同时接收到大量的 Telnet 报文、OSPF 报文以及 ARP 报文，由于 CPU 超负载，无法全部处理所有的报文，大量报文积压在队列中，这时将出现用户 Telnet 不时断开、OSPF 协议震荡，用户 ARP 访问异常等情况。

图 18-2



功能部属

- 默认情况下，CPU 集中保护打开，为每种分类分配独立带宽与占宽比。这时 CPU 将优先处理 Telnet 报文，保证用户 Telnet 不断连；其次处理 OSPF 报文，尽量维护 OSPF 协议的稳定；最后处理 ARP 报文。
- 若在缺省配置下仍然出现上述现象，可适当调整分类带宽与占宽比参数。

18.3 功能详解

基本概念

📌 ARP 抗攻击

在局域网中，通过 ARP 协议把 IP 地址转换为 MAC 地址，ARP 协议对网络安全具有重要的意义。通过网络向网关发送大量非法的 ARP 报文，造成网关不能为正常主机提供服务，这就是基于 ARP 的拒绝服务攻击。对于这种攻击，防范措施是一方面对 ARP 报文限速，另一方面检测出攻击源，对攻击源头采取隔离措施。

📌 IP 防扫描

众所周知，许多黑客攻击、网络病毒入侵都是从扫描网络内活动的主机开始的。因此大量的扫描报文急剧占用了网络带宽，导致网络通讯无法正常进行。

为此，交换机三层设备提供了防 IP 攻击的功能，用以防止黑客扫描和类似“冲击波”病毒的攻击，还能减少三层设备的 CPU 负担。目前发现的 IP 攻击主要有两种：

目的 IP 地址变化的扫描。这种扫描是最危害网络的，不但消耗网络带宽，增加设备的负担，而且更是大部分黑客攻击手段的前奏。

向不存在目的 IP 地址高速发送 IP 报文。这种攻击主要是针对设备 CPU 的负担来设计。对三层设备来说，如果目的 IP 地址存在，则报文会被交换芯片直接转发，不会占用设备 CPU 的资源，而如果目的 IP 地址不存在，IP 报文会送到 CPU，由 CPU 发送 ARP 请求询问目的 IP 地址对应的 MAC 地址，如果送到 CPU 的报文太多，会消耗 CPU 资源。当然，这种攻击的危害比第一种小得多了。

对于“向不存在的目的 IP 地址高速发 IP 报文”这种 IP 攻击，防范措施是一方面对 IP 报文限速，另一方面检测出攻击源，对攻击源头采取隔离措施。

▾ ICMP 抗攻击

ICMP 协议作为诊断网络故障的常用手段，它的基本原理是主机发出 ICMP 回音请求报文（ICMP echo request），路由器或者交换机接收到这个请求报文后会回应一个 ICMP 回音应答（ICMP echo reply）报文。在上述这个处理过程中需要设备的 CPU 进行处理，这样就必然需要消耗 CPU 的一部分资源。如果攻击者向目标设备发送大量的 ICMP 回音请求，这样势必会导致设备的 CPU 资源被大量消耗，严重的情况可能导致设备无法正常工作，这种攻击方式也被人们命名为“ICMP 洪水”。对于这种攻击，防范措施是一方面对 ICMP 报文限速，另一方面检测出攻击源，对攻击源头采取隔离措施。

▾ DHCP 抗攻击

DHCP 协议被广泛地应用在局域网环境里来动态分配 IP 地址。DHCP 协议对网络安全起着非常重要的意义。目前，存在的最广泛的 DHCP 攻击就是称为“DHCP 耗竭”的攻击，这种攻击通过伪造的 MAC 地址来广播 DHCP 请求的方式进行。目前网络上存在多种这样的攻击工具都可以很容易地实现上述攻击。如果发出的 DHCP 请求足够多的话，网络攻击者就可以在一段时间内耗竭 DHCP 服务器所提供的地址空间，这样当一台合法的主机请求一个 DHCP IP 地址的时候无法成功，从而无法访问网络。对于这种攻击，防范措施是一方面对 DHCP 报文限速，另一方面检测出攻击源，对攻击源头采取隔离措施。

▾ ND 抗攻击

ND 的全称是 Neighbor Discovery，翻译成汉语是“邻居发现”。

ND snooping 通过监听网络中的 ND 报文，过滤非法的 ND 报文。ND snooping 要求把 ND 报文送 CPU，如果 ND 报文的速率很高，会造成对 CPU 的攻击，所以需要实现 ND guard，对 ND 报文进行限速。

▾ 自定义抗攻击

网络协议种类繁多，仅路由协议就有 OSPF、RIP 等。各种协议需要在不同设备之间进行报文交互，交互报文必须送 CPU 交由各个协议进行处理，网络设备一运行某种协议，就相当于开了一扇窗口，给了攻击者可趁之机。如果攻击者向网络设备发送大量的协议报文，将会导致设备的 CPU 资源被大量消耗，严重的情况可能导致设备无法正常工作。

考虑到网络协议多种多样，并且在持续发展中，不同的用户环境下需要使用不同的协议。为此，锐捷设备提供了自定义抗攻击的功能，允许用户自定义抗攻击的类型，灵活配置，以满足不同用户环境下的抗攻击需求。

功能特性

功能特性	作用
主机限速和识别攻击	根据主机限速水线进行限速，并识别网络中的主机用户攻击。
端口限速和识别攻击	根据端口限速水线进行限速，并识别端口攻击。
设置监控时间	在指定时间内，对主机用户攻击者进行软件监控。
配置硬件隔离	在指定时间内，对主机用户攻击者或攻击端口进行硬件隔离。
设置不监控的可信主机	对某主机用户不进行监控，即对该主机表示信任。
集中限速分发	将报文分类，区分处理优先级。

18.3.1 主机限速和识别攻击

对主机用户攻击报文进行限速，识别主机用户攻击。

识别 ARP 扫描。

识别 IP 扫描。

工作原理

识别主机有源 IP/VLAN ID/端口和链路层源 MAC/VLAN ID/端口两种识别方法。每台主机都有限速水线和攻击阈值（也称为告警水线），限速水线必须低于攻击阈值。当单台主机的攻击报文速度超过限速水线时，将丢弃这些超出限速水线的报文；如果单台主机的攻击报文速度超过攻击阈值，将识别主机用户攻击，记录到日志中，发送 TRAP。

如果在配置时间内收到超过扫描水线的 ARP 报文，链路层源 MAC 地址固定而源 IP 地址变化，或者链路层源 MAC 地址和源 IP 地址固定而目标 IP 地址不断变化，就认为有 ARP 扫描嫌疑。

如果在配置时间内收到超过扫描水线的 IP 报文，源 IP 不变，目的 IP 一直在变化，就认为有 IP 扫描嫌疑。

- i** NFPP 在检测到某种服务的某个具体报文的攻击后，可以向管理员发出告警信息，但是为了防止告警信息频繁出现，如果攻击流持续存在，NFPP 在发出告警后的连续 60 秒时间内不再重复告警。
- i** 防止频繁打印日志消耗 CPU 资源，NFPP 把攻击检测的日志信息写到缓冲区，然后以指定速率从缓冲区取出来打印。NFPP 对 TRAP 没有限速。

相关配置

以 ARP 抗攻击为例：

配置全局主机限速与攻击识别

在 nfpp 模式下：

使用 `arp-guard rate-limit {per-src-ip | per-src-mac} pps` 命令可以配置 IP/VID/端口识别主机与基于链路层源 MAC/VID/端口识别主机的限速水线。

使用 `arp-guard attack-threshold {per-src-ip | per-src-mac} pps` 命令可以配置 IP/VID/端口识别主机与基于链路层源 MAC/VID/端口识别主机的攻击水线。

使用 `arp-guard scan-threshold pkt-cnt` 命令可以配置 ARP 扫描阈值。

配置接口主机限速与攻击识别

在接口模式下：

使用 `nfpp arp-guard policy {per-src-ip | per-src-mac} rate-limit-pps attack-threshold-pps` 命令可以配置该端口上 IP/VID/端口识别主机与基于链路层源 MAC/VID/端口识别主机的限速水线和攻击水线。

使用 `nfpp arp-guard scan-threshold pkt-cnt` 命令可以配置该端口上扫描水线。

- i** 当前仅 ARP 抗攻击与 IP 防扫描支持防扫描功能。

18.3.2 端口限速和识别攻击

工作原理

每个端口都有限速水线和攻击阈值，限速水线必须低于攻击阈值。当某个端口的报文速度超过限速水线时，就丢弃报文。如果某个端口的报文速度超过攻击阈值，将记录到日志中，发送 TRAP。

相关配置

以 ARP 抗攻击为例：

配置全局端口限速与攻击识别

在 nfpp 模式下：

使用 `arp-guard rate-limit per-port pps` 命令可以配置基于端口的限速水线。

使用 `arp-guard attack-threshold per-port pps` 命令可以配置基于端口的攻击水线。

配置接口端口限速与攻击识别

接口模式下：

使用 `nfpp arp-guard policy per-port rate-limit-pps attack-threshold-pps` 命令可以配置基于端口的限速水线和攻击水线。

18.3.3 设置监控时间

工作原理

监控用户提供当前系统中存在哪些攻击者的信息。如果隔离时间为 0（即不隔离），防攻击模块将自动根据配置的监控时间对攻击者进行软件监控。当把隔离时间配置成非零值后，防攻击模块将自动对软件监控的主机采取硬件隔离。

在软件监控过程中，当隔离时间被配置为非零值时，将自动对软件监控的攻击者采取硬件隔离，并且把超时设置为隔离时间。监控时间在隔离时间为 0 的情况下才有意义。

相关配置

以 ARP 抗攻击为例：

配置全局监控时间

在 nfpp 模式下：

使用 `arp-guard monitor-period seconds` 命令可以配置监控时间。

18.3.4 配置硬件隔离

工作原理

隔离动作是在检测到攻击后交由抗攻击策略执行的。隔离利用硬件的过滤器实现，这样保证该攻击报文不会再被送到 CPU 处理，从而保证了设备正常运行。

硬件隔离支持基于主机用户的隔离以及基于端口的隔离，当前仅 ARP 抗攻击支持基于端口的硬件隔离功能。

对攻击者进行隔离，会设置一条策略到硬件中，但是硬件资源有限，当硬件资源耗尽的时候，会打印日志提醒管理员。

相关配置

以 ARP 抗攻击为例：

配置全局隔离时间

在 nfpp 模式下：

使用 `arp-guard isolate-period [seconds | permanent]` 命令可以配置隔离时间。当配置为 0 时，关闭隔离；当非 0 数值时，表示隔离时间；当为 permanent 时，表示永久隔离。

配置接口隔离时间

在接口模式下：

使用 `nfpp arp-guard isolate-period [seconds | permanent]` 命令可以配置隔离时间。当配置为 0 时，关闭隔离；当非 0 数值时，表示隔离时间；当为 permanent 时，表示永久隔离。

配置允许隔离转发功能

在 nfpp 模式下：

使用 `arp-guard isolate-forwarding enable` 命令可以配置隔离表项是否允许转发。

配置基于端口的隔离转发限速功能

在 nfpp 模式下：

使用 `arp-guard ratelimit-forwarding enable` 命令可以配置基于端口的隔离转发限速功能。

i 当前仅 ARP 抗攻击支持配置允许隔离转发功能以及配置基于端口的隔离转发限速功能。

18.3.5 设置不监控的可信主机

工作原理

如果管理员希望对某台主机不进行监控，即对该主机表示信任，则可以通过该命令配置。该可信主机发往 CPU 的报文将被允许发往 CPU。

相关配置

以 IP 防扫描为例：

配置信任用户

在 nfpp 模式下：

使用 `ip-guard trusted-host ip mask` 命令可以配置信任用户

使用 `trusted-host {mac mac_mask | ip mask}` 命令可以配置自定义抗攻击信任用户。

18.3.6 集中限速分发

工作原理

将 CPP 中定义的各种服务按照管理类(Manage)、转发类(Route)和协议类(Protocol)的原则进行的分类(具体分类如下表所列)，每一类都拥有独立的带宽，不同类别之间的带宽不能共享，超过带宽阈值的流将被丢弃。这样将不同的服务区分类别后，可以保证属于某类的各种服务报文在设备上得到优先处理。

NFPP 允许管理员根据实际的网络环境灵活分配三类报文的带宽，从而保障 protocol 类和 manage 类能得到优先处理，protocol 类的优先处理保证了协议的正确运行，而 manage 类的优先处理保证了管理员能够实施正常管理，从而保障了设备的各种重要功能的正常运行，提高设备的抗攻击能力。

经过以上的分类限速后，再将所有的分类流集中一个队列中，这样当某一类服务处理效率较低时，队列上就会堆积该服务对应的报文，并可能最终耗尽该队列资源，NFPP 允许管理员配置该队列中三类所占百分比，当某一类占用的队列长度超过总队列长度和该类所占百分比的乘积时，报文就会被丢弃。这样就有效地解决了某一类独占队列资源的问题。

三种属性分类	CPU Protect Policy 中定义服务类型
Protocol	tp-guard, dot1x, rldp, rerp, slow-packet, bpdu, gvrp, ripng, dvmrp, igmp, mpls, ospf, pim, rip, vrrp, dhcp-relay-s, dhcp-relay-c, option82
Route	unknown-ipmc, ttl1, ttl0, udp-helper, ip4-packet-other, non-ip-packet-other, arp
Manage	ip4-packet-local

 服务类型具体含义参见 CPP 配置指南

相关配置

配置每类报文允许的最大带宽

在全局模式下：

使用 `cpu-protect sub-interface { manage | protocol|route} pps pps_value` 命令设置报文对应的队列的限速水线。

配置每类报文占用队列的最大百分比

在全局模式下：

使用 `cpu-protect sub-interface { manage | protocol | route} percent percent_vaule` 命令设置报文对应的类型所占的队列的百分比。

18.4 配置详解

配置项	配置建议 & 相关命令	
配置ARP抗攻击	<code>arp-guard enable</code>	配置全局攻击检测使能
	<code>arp-guard isolate-period</code>	配置全局隔离时间
	<code>arp-guard isolate-forwarding enable</code>	配置允许隔离转发功能
	<code>arp-guard ratelimit-forwarding enable</code>	配置基于端口的隔离转发限速功能
	<code>arp-guard monitor-period</code>	配置监控时间
	<code>arp-guard monitored-host-limit</code>	配置监控主机最大数目
	<code>arp-guard rate-limit</code>	配置全局限速水线
	<code>arp-guard attack-threshold</code>	配置全局攻击水线
	<code>arp-guard scan-threshold</code>	配置全局主机扫描水线
	<code>nfpp arp-guard enable</code>	配置端口攻击检测使能
	<code>nfpp arp-guard policy</code>	配置端口限速水线、攻击水线
	<code>nfpp arp-guard scan-threshold</code>	配置端口逐级扫描水线
	<code>nfpp arp-guard isolate-period</code>	配置端口隔离时间
配置IP防扫描	<code>ip-guard enable</code>	配置全局攻击检测使能
	<code>ip-guard isolate-period</code>	配置全局隔离时间
	<code>ip-guard monitor-period</code>	配置监控时间
	<code>ip-guard monitored-host-limit</code>	配置监控主机最大数目
	<code>ip-guard rate-limit</code>	配置全局限速水线
	<code>ip-guard attack-threshold</code>	配置全局攻击水线
	<code>ip-guard scan-threshold</code>	配置全局主机扫描水线
	<code>ip-guard trusted-host</code>	配置信任用户
	<code>nfpp ip-guard enable</code>	配置端口攻击检测使能
	<code>nfpp ip-guard policy</code>	配置端口限速水线、攻击水线
	<code>nfpp ip-guard scan-threshold</code>	配置端口逐级扫描水线
	<code>nfpp ip-guard isolate-period</code>	配置端口隔离时间
配置ICMP抗攻击	<code>icmp-guard enable</code>	配置全局攻击检测使能
	<code>icmp-guard isolate-period</code>	配置全局隔离时间
	<code>icmp-guard monitor-period</code>	配置监控时间
	<code>icmp-guard monitored-host-limit</code>	配置监控主机最大数目
	<code>icmp-guard rate-limit</code>	配置全局限速水线
	<code>icmp-guard attack-threshold</code>	配置全局攻击水线

	icmp-guard trusted-host	配置信任用户
	nfpp icmp-guard enable	配置端口攻击检测使能
	nfpp icmp-guard policy	配置端口限速水线、攻击水线
	nfpp icmp-guard isolate-period	配置端口隔离时间
配置DHCP抗攻击	dhcp-guard enable	配置全局攻击检测使能
	dhcp-guard isolate-period	配置全局隔离时间
	dhcp-guard monitor-period	配置监控时间
	dhcp-guard monitored-host-limit	配置监控主机最大数目
	dhcp-guard rate-limit	配置全局限速水线
	dhcp-guard attack-threshold	配置全局攻击水线
	nfpp dhcp-guard enable	配置端口攻击检测使能
	nfpp dhcp-guard policy	配置端口限速水线、攻击水线
	nfpp dhcp-guard isolate-period	配置端口隔离时间
配置ND抗攻击	nd-guard enable	配置全局攻击检测使能
	nd-guard ratelimit-forwarding enable	配置基于端口的隔离转发限速功能
	nd-guard rate-limit per-port	配置全局限速水线
	nd-guard attack-threshold per-port	配置全局攻击水线
	nfpp nd-guard enable	配置端口攻击检测使能
	nfpp nd-guard policy per-port	配置端口限速水线、攻击水线
配置自定义抗攻击	define	配置自定义抗攻击名称
	match	配置自定义抗攻击类型需要匹配的报文字段
	global-policy	配置全局基于主机或者基于端口的限速水线和攻击水线
	monitor-period	配置自定义抗攻击的监控时间
	monitored-host-limit	配置自定义抗攻击受监控主机的最大数目
	trusted-host	配置不进行监控的主机
	define name enable	配置全局使能自定义抗攻击
	nfpp define name enable	配置端口上使能自定义抗攻击
	nfpp define	配置端口上基于主机或者基于端口的限速水线和攻击水线
NFPP日记信息	log-buffer entries	配置 NFPP 日志缓冲区大小
	log-buffer logs	配置从专用日志缓冲区取日志生成系统消息的速率
	logging vlan	指定需要记录哪些 VLAN 的日志
	logging interface	指定需要记录哪个端口的日志
	logging enable	配置日志打印到屏幕

18.4.1 配置ARP抗攻击

配置效果

- ARP 攻击识别分为基于主机和基于物理端口两个类别。基于主机又细分为基于源 IP 地址/VLAN ID/物理端口和基于链路层源 MAC 地址/ VLAN ID /物理端口。每种攻击识别都有限速水线和告警水线。当 ARP 报文速率超过限速水线时，超限报文将被丢弃。当 ARP 报文速率超过告警水线时，将打印警告信息，发送 TRAP，基于主机的攻击识别还会对攻击源头采取隔离措施。
- ARP 抗攻击还能检测出 ARP 扫描。ARP 扫描是指链路层源 MAC 地址固定而源 IP 地址变化，或者链路层源 MAC 地址和源 IP 地址固定而目标 IP 地址不断变化。由于存在误判的可能，对检测出有 ARP 扫描嫌疑的主机不进行隔离，只是提供给管理员参考。
- 配置 ARP 抗攻击隔离,针对主机用户攻击下发硬件隔离表项,攻击报文不送 CPU，不转发。

注意事项

- 对于在全局与接口上同时存在配置的命令，接口优先级高于全局。
- 隔离默认关闭，若打开隔离功能，攻击用户将占用安全模块硬件表项。
- ARP 抗攻击只是针对攻击交换机本身的 ARP 拒绝服务攻击，而不是针对 ARP 欺骗或者是解决网络中的 ARP 攻击问题。
- 对 DAI 的信任口，其 ARP 抗攻击功能不生效，避免对信任口的 ARP 流进行误判。关于 DAI 的信任口，具体参见“DAI 配置”的“例外端口”章节。

配置方法

使能攻击检测

- 必须配置，默认打开。
- 支持全局配置以及单端口独立配置。
- 当关闭 ARP 抗攻击功能时，系统将自动清除受监控的主机、扫描主机和端口隔离表项。

配置隔离时间

- 可选配置，默认关闭隔离功能。
- 在攻击用户报文流量超过 CPP 限速带宽时，可配置隔离时间，将报文直接硬件丢弃，避免占用带宽资源。
- 支持全局配置以及单端口独立配置。
- 如果把隔离时间从非零值改成零，将直接把相关端口的攻击者删除，而不是进行软件监控。

配置允许隔离转发功能

- 可选配置，默认 ARP 抗攻击丢弃转发报文。
- 在隔离生效时，希望隔离仅针对管理面生效，而不针对转发面，可配置允许隔离转发功能。
- 支持全局配置。

配置基于端口的隔离转发限速功能

- 可选配置，默认基于端口的隔离转发限速功能生效。
- 在基于端口的隔离表项生效时，希望隔离动作非全部丢弃，而是允许部分报文通过，可配置基于端口的隔离转发限速功能。
- 支持全局配置。

配置配置监控时间

- 必须配置，默认时间为 600 秒。
- 在配置了隔离时间时，攻击用户监控时间直接采用隔离时间，配置的监控时间不生效。
- 支持全局配置。

配置监控主机最大数目

- 必须配置，默认 20000 个。
- 配置监控主机最大数目，随实际监控主机数增加，处理监控用户需占用更多 CPU 资源。
- 支持全局配置
- 如果受监控主机数已经达到默认的 20000 个，此时管理员把受监控主机的最大数目设置成小于 20000，不会删除已有的受监控主机，而是打印信息 “%ERROR : The value that you configured is smaller than current monitored hosts 20000 (配置的受监控主机数) ， please clear a part of monitored hosts.” 来提醒管理员配置没有生效，需要删除部分已经被监控的主机。
- 当受监控主机表满时，打印日志 “% NFPP_ARP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 (配置的受监控主机数) monitored hosts.” 提醒管理员。

配置攻击检测水线

- 必须配置，存在默认配置参数，详细见产品特性文档。
- 为了使 ARP 抗攻击得到最佳的防攻击效果，建议管理员配置基于主机的限速水线和告警水线时遵循以下原则：基于 IP 地址的限速水线 < 基于 IP 地址的告警水线 < 基于源 MAC 地址的限速水线 < 基于源 MAC 地址的告警水线。。
- 支持全局配置以及单端口独立配置。
- 当管理员配置的限速水线大于攻击阈值时，打印命令提示信息 “%ERROR : rate limit is higher than attack threshold 500pps(配置的攻击阈值).” 提醒管理员。
- 当管理员配置的攻击阈值小于限速水线时，打印命令提示信息 “%ERROR : attack threshold is smaller than rate limit 300pps(配置的限速水线).” 提醒管理员。
- 当无法为检测到的攻击者分配内存时，打印日志 “%NFPP_ARP_GUARD-4-NO_MEMORY: Failed to alloc memory..” 提醒管理员。

- 基于 MAC 地址限速的优先级高于基于 IP 地址限速，而基于 IP 地址限速又高于基于端口限速。

配置扫描检测水线

- 必须配置，存在默认配置参数，详见产品特性文档。
- 支持全局配置以及单端口独立配置。
- ARP 扫描表只记录最新的 256 条记录。当 ARP 扫描表满了以后，最新记录将覆盖最旧记录。
- 如果 10 秒钟收到的 ARP 报文，链路层源 MAC 地址固定而源 IP 地址变化，或者链路层源 MAC 地址和源 IP 地址固定而目标 IP 地址不断变化，变化次数超过扫描水线，就认为有扫描嫌疑。

检验方法

网络主机往配置了 ARP 攻击检测限速的交换机发送 ARP 攻击报文，需确认该报文可送 CPU。

- 若超过攻击水线或扫描水线，将有攻击信息提示。
- 若攻击用户需生成隔离标项，将有用户隔离信息提示。

相关命令

使能攻击检测

【命令格式】 **arp-guard enable**

【参数说明】 -

【命令模式】 nfpp 模式下

【使用指导】 -

配置隔离时间

【命令格式】 **arp-guard isolate-period [seconds | permanent]**

【参数说明】 **seconds**：隔离时间，单位是秒，取值范围是 0 或者[30, 86400]。

permanent：永久隔离。

【命令模式】 nfpp 模式下

【使用指导】 -

配置允许隔离转发功能

【命令格式】 **arp-guard isolate-forwarding enable**

【参数说明】 -

【命令模式】 nfpp 模式下

【使用指导】 -

配置基于端口的隔离转发限速功能

【命令格式】 **arp-guard ratelimit-forwarding enable**

【参数说明】 -

【命令模式】 nfpp 模式下

【使用指导】 -

配置监控时间

【命令格式】 **arp-guard monitor-period** *seconds*

【参数说明】 *seconds* : 监控时间, 单位是秒, 取值范围是[180, 86400]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置监控主机最大数目

【命令格式】 **arp-guard monitored-host-limit** *number*

【参数说明】 *number* : 支持的最大受监控主机数, 取值范围为 1 到 4294967295。

【命令模式】 nfpp 模式下

【使用指导】 -

配置限速水线

【命令格式】 **arp-guard rate-limit** {*per-src-ip* | *per-src-mac* | *per-port*} *pps*

【参数说明】 **per-src-ip** : 对每个源 IP 地址进行限速。

per-src-mac : 对每个源 MAC 地址进行限速。

per-port : 对每个端口进行限速。

pps : 限速水线值, 取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置攻击水线

【命令格式】 **arp-guard attack-threshold** {*per-src-ip* | *per-src-mac* | *per-port*} *pps*

【参数说明】 **per-src-ip** : 配置每个源 IP 地址的攻击水线

per-src-mac : 配置每个源 MAC 地址的攻击水线

per-port : 配置每个端口的攻击水线

pps : 攻击水线, 单位是每秒报文数, 取值范围是[1,19999]

【命令模式】 nfpp 模式下

【使用指导】 攻击水线不能小于限速水线。

配置扫描水线

【命令格式】 **arp-guard scan-threshold** *pkt-cnt*

【参数说明】 *pkt-cnt* : 扫描水线值, 取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置接口上使能攻击检测

【命令格式】 **nfpp arp-guard enable**

- 【参数说明】 -
- 【命令模式】 接口模式下
- 【使用指导】 端口的 ARP 抗攻击开关优先于全局 ARP 抗攻击开关。

配置接口上隔离时间

- 【命令格式】 **nfpp arp-guard isolate-period** [*seconds* | **permanent**]
- 【参数说明】 *seconds* : 隔离时间, 单位是秒, 取值范围是 0 或者[30, 86400], 0 表示不隔离。
permanent : 永久隔离。
- 【命令模式】 接口模式下
- 【使用指导】 -

配置接口上攻击策略

- 【命令格式】 **nfpp arp-guard policy** {*per-src-ip* | *per-src-mac* | *per-port*} *rate-limit-pps* *attack-threshold-pps*
- 【参数说明】 **per-src-ip** : 配置每个源 IP 地址的限速水线和攻击水线。
per-src-mac : 配置每个源 MAC 地址的限速水线和攻击水线。
per-port : 配置每个端口的限速水线和攻击水线。
rate-limit-pps : 限速水线, 取值范围是 1 到 19999。
attack-threshold-pps : 攻击水线, 取值范围是 1 到 19999。
- 【命令模式】 接口模式下
- 【使用指导】 攻击水线不能小于限速水线。

配置接口上扫描

- 【命令格式】 **nfpp arp-guard scan-threshold** *pkt-cnt*
- 【参数说明】 *pkt-cnt* : 扫描水线值, 取值范围是[1,19999]。
- 【命令模式】 接口模式下
- 【使用指导】 -

配置举例

通过 ARP 抗攻击保护 CPU

- 【网络环境】
- 系统中带有 ARP 主机用户攻击, 导致部分用户无法正常建立 ARP 连接。
 - 系统中存在 ARP 扫描, 导致 CPU 利用率很高。
- 【配置方法】
- 配置基于主机的攻击检测水线为 5pps。
 - 配置 ARP 扫描检测水线为 10pps。
 - 配置隔离时间为 180pps。

```
Ruijie# configure terminal
Ruijie(config)# nfpp
Ruijie (config-nfpp)#arp-guard rate-limit per-src-mac 5
Ruijie (config-nfpp)#arp-guard attack-threshold per-src-mac 10
Ruijie (config-nfpp)#arp-guard isolate-period 180
```

- 【检验方法】
- 通过 **show nfpp arp-guard summary** 可以查看到配置信息。

(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.)

Interface	Status	Isolate-period	Rate-limit	Attack-threshold	Scan-threshold
Global	Disable	180	4/5/100	8/10/200	15

Maximum count of monitored hosts: 1000

Monitor period: 600s

- 通过 **show nfpp arp-guard hosts** 可以查看到监控用户。

If col_filter 1 shows '*', it means "hardware do not isolate host".

VLAN	interface	IP address	MAC address	remain-time(s)
1	Gi0/43	5.5.5.16	-	175

Total: 1 host

- 通过 **show nfpp arp-guard scan** 可以查看到扫描用户。

VLAN	interface	IP address	MAC address	timestamp
1	Gi0/5	-	001a.a9c2.4609	2013-4-30 23:50:32
1	Gi0/5	192.168.206.2	001a.a9c2.4609	2013-4-30 23:50:33
1	Gi0/5	-	001a.a9c2.4609	2013-4-30 23:51:33
1	Gi0/5	192.168.206.2	001a.a9c2.4609	2013-4-30 23:51:34

Total: 4 record(s)

常见错误

-

18.4.2 配置IP防扫描

配置效果

- IP 攻击识别分为基于主机和基于物理端口两个类别。基于主机是采用源 IP 地址/VLAN ID/物理端口三者结合识别的。每种攻击识别都有限速水线和告警水线。当 IP 报文速率超过限速水线时，超限报文将被丢弃。当 IP 报文速率超过告警水线时，将打印警告信息，发送 TRAP，基于主机的攻击识别还会对攻击源头采取隔离措施。
- IP 抗攻击还能检测出 IP 扫描。IP 防扫描针对的是目的 IP 一直发生变化，源 IP 不变，且目的 IP 地址不是本机 IP 地址的 IP 报文攻击。
- 配置 IP 抗攻击隔离，针对主机用户攻击下发硬件隔离表项，攻击报文不送 CPU，不转发。
- IP 防扫描针对的是目的 IP 地址不是本机 IP 地址的 IP 报文攻击。对于目的 IP 地址是本机 IP 地址的 IP 报文，则由 CPP (CPU Protect Policy) 限速。

注意事项

- 对于在全局与接口上同时存在配置的命令，接口优先级高于全局。
- 隔离默认关闭，若打开隔离功能，攻击用户将占用安全模块硬件表项。

配置方法

使能攻击检测

- 必须配置，默认打开。
- 支持全局配置以及单端口独立配置。
- 当关闭 IP 防扫描功能时，系统将自动清除受监控的主机。

配置隔离时间

- 可选配置，默认关闭隔离功能，。
- 在攻击用户报文流量超过 CPP 限速带宽时，可配置隔离时间，将报文直接硬件丢弃，避免占用带宽资源。
- 支持全局配置以及单端口独立配置。
- 如果把隔离时间从非零值改成零，将直接把相关端口的攻击者删除，而不是进行软件监控。

配置配置监控时间

- 必须配置，默认时间为 600 秒。
- 在配置了隔离时间时，攻击用户监控时间直接采用隔离时间，配置的监控时间不生效。
- 支持全局配置。

配置监控主机最大数目

- 必须配置，默认 20000 个。
- 提高监控主机最大数目，随实际监控主机数增加，处理监控用户需占用更多 CPU 资源。
- 支持全局配置
- 如果受监控主机数已经达到默认的 20000 个，此时管理员把受监控主机的最大数目设置成小于 20000，不会删除已有的受监控主机，而是打印信息“%ERROR: The value that you configured is smaller than current monitored hosts 20000 (配置的受监控主机数) ， please clear a part of monitored hosts.”来提醒管理员配置没有生效，需要删除部分已经被监控的主机。
- 当受监控主机表满时，打印日志 “% NFPP_IP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 (配置的受监控主机数) monitored hosts.” 提醒管理员。

配置攻击检测水线

- 必须配置，存在默认配置参数，详细见产品特性文档。
- 支持全局配置以及单端口独立配置。

- 当管理员配置的限速水线大于攻击阈值时，打印命令提示信息 “%ERROR : rate limit is higher than attack threshold 500pps(配置的攻击阈值).” 提醒管理员。
- 当管理员配置的攻击阈值小于限速水线时，打印命令提示信息 “%ERROR : attack threshold is smaller than rate limit 300pps(配置的限速水线).” 提醒管理员。
- 当无法为检测到的攻击者分配内存时，打印日志 “%NFPP_IP_GUARD-4-NO_MEMORY: Failed to alloc memory..” 提醒管理员。
- 基于源 IP 地址限速优先级高于基于端口限速。

配置扫描检测水线

- 必须配置，存在默认配置参数，详细见产品特性文档。
- 支持全局配置以及单端口独立配置。
- 如果 10 秒钟收到的 IP 报文，目的 IP 一直发生变化，源 IP 不变，且目的 IP 地址不是本机 IP 地址的 IP 报文，若变化次数超过扫描水线，就认为有扫描嫌疑。

配置不监控可信主机

- 可选配置，默认无不监控可信主机。
- IP 防扫描仅支持配置不进行监控 IP，最多可配置 500 条。
- 支持全局配置。
- 当受监控主机表中存在与可信主机相匹配的表项（IP 地址相同）时，系统将自动删除此 IP 地址对应的表项。
- 当不监控的可信主机表满时，打印提示信息 “%ERROR: Attempt to exceed limit of 500 trusted hosts.” 提醒管理员。
- 当删除可信主机失败时，打印提示信息 “%ERROR: Failed to delete trusted host 1.1.1.0 255.255.255.0 (配置的可信主机).” 提醒管理员。
- 当添加可信主机失败时，打印提示信息 “%ERROR: Failed to add trusted host 1.1.1.0 255.255.255.0 (配置的可信主机).” 提醒管理员。
- 当添加的可信主机已经存在时，打印提示信息 “%ERROR: Trusted host 1.1.1.0 255.255.255.0 (配置的可信主机) has already been configured.” 提醒管理员。
- 当要删除的可信主机不存在时，打印提示信息 “%ERROR: Trusted host 1.1.1.0 255.255.255.0(配置的可信主机) is not found.” 提醒管理员。
- 当无法为可信主机分配内存时，打印提示信息 “%ERROR: Failed to alloc memory.” 提醒管理员。

检验方法

网络主机往配置了 IP 攻击检测限速的交换机发送 IP 攻击报文，需确认该报文可送 CPU。

- 对于不满足信任用户配置的报文，若超过攻击水线或扫描水线，将有攻击信息提示。
- 若攻击用户需生成隔离标项，将有用户隔离信息提示。

相关命令

使能攻击检测

- 【命令格式】 **ip-guard enable**
- 【参数说明】 -
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

配置隔离时间

- 【命令格式】 **ip-guard isolate-period** [*seconds* | **permanent**]
- 【参数说明】 *seconds* : 隔离时间, 单位是秒, 取值范围是 0 或者[30, 86400]。
permanent : 永久隔离。
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

配置监控时间

- 【命令格式】 **ip-guard monitor-period** *seconds*
- 【参数说明】 *seconds* : 监控时间, 单位是秒, 取值范围是[180, 86400]。
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

如果把隔离时间从非零值改成零, 将直接把相关端口的攻击者删除, 而不是进行软件监控。

配置监控主机最大数目

- 【命令格式】 **ip-guard monitored-host-limit** *number*
- 【参数说明】 *number* : 支持的最大受监控主机数, 取值范围为 1 到 4294967295。
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

配置限速水线

- 【命令格式】 **ip-guard rate-limit** {**per-src-ip** | **per-port**} *pps*
- 【参数说明】 **per-src-ip** : 对每个源 IP 地址进行限速。
per-port : 对每个端口进行限速。
pps : 限速水线值, 取值范围是[1,19999]。
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

配置攻击水线

- 【命令格式】 **ip-guard attack-threshold** {**per-src-ip** | **per-port**} *pps*
- 【参数说明】 **per-src-ip** : 配置每个源 IP 地址的攻击水线。
per-port : 配置每个端口的攻击水线。

pps : 攻击水线, 单位是每秒报文数, 取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 攻击水线不能小于限速水线。

配置扫描水线

【命令格式】 **ip-guard scan-threshold *pkt-cnt***

【参数说明】 *pkt-cnt* : 扫描水线值, 取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置信任用户

【命令格式】 **ip-guard trusted-host *ip mask***

【参数说明】 *ip* : IP 地址。

mask : IP 地址的掩码。

all : 和 **no** 一起使用, 删除所有可信主机配置。

【命令模式】 nfpp 模式下

【使用指导】 如果管理员希望对某台主机不进行监控, 即对该主机表示信任, 则可以通过该命令配置。该可信主机发往 CPU 的 IP 报文将被允许发往 CPU, 不做任何的限速和告警处理。

配置接口上使能攻击检测

【命令格式】 **nfpp ip-guard enable**

【参数说明】 -

【命令模式】 接口模式下

【使用指导】 端口的 IP 防扫描开关优先于全局防扫描开关。

配置接口上隔离时间

【命令格式】 **nfpp ip-guard isolate-period [*seconds* | **permanent**]**

【参数说明】 *seconds* : 隔离时间, 单位是秒, 取值范围是 0 或者[30, 86400], 0 表示不隔离。

permanent : 永久隔离。

【命令模式】 接口模式下

【使用指导】 -

配置接口上攻击策略

【命令格式】 **nfpp ip-guard policy {*per-src-ip* | *per-port*} *rate-limit-pps attack-threshold-pps***

【参数说明】 **per-src-ip** : 配置每个源 IP 地址的攻击水线。

per-port : 配置每个端口的攻击水线。

rate-limit-pps : 限速水线, 取值范围是 1 到 19999。

attack-threshold-pps : 攻击水线, 取值范围是 1 到 19999。

【命令模式】 接口模式下

【使用指导】 攻击水线不能小于限速水线。

配置接口上扫描

- 【命令格式】 **nfpp ip-guard scan-threshold** *pkt-cnt*
- 【参数说明】 *pkt-cnt* : 扫描水位值, 取值范围是[1,19999]。
- 【命令模式】 接口模式下
- 【使用指导】 -

配置举例

通过 IP 防扫描保护 CPU

- 【网络环境】
- 系统中带有 IP 主机用户攻击, 部分用户报文无法正常路由转发。
 - 系统中存在 IP 扫描, 导致 CPU 利用率很高。
 - 系统中, 部分主机报文流量很大, 需放行
- 【配置方法】
- 配置基于主机的攻击检测水位。
 - 配置 IP 扫描检测水位。
 - 配置隔离时间为非 0。
 - 配置不监控可信主机

```
Ruijie# configure terminal
Ruijie(config)# nfpp
Ruijie (config-nfpp)#ip-guard rate-limit per-src-ip 20
Ruijie (config-nfpp)#ip-guard attack-threshold per-src-ip 30
Ruijie (config-nfpp)#ip-guard isolate-period 180
Ruijie (config-nfpp)#ip-guard trusted-host 192.168.201.46 255.255.255.255
```

- 【检验方法】
- 通过 **show nfpp ip-guard summary** 可以查看到配置信息。

```
(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.)
Interface Status Isolate-period Rate-limit Attack-threshold Scan-threshold
Global Disable 180 20/-/100 30/-/200 100

Maximum count of monitored hosts: 1000
Monitor period: 600s
```

- 通过 **show nfpp ip-guard hosts** 可以查看到监控用户。

```
If col_filter 1 shows '*', it means "hardware do not isolate host".
VLAN interface IP address Reason remain-time(s)
----
1 Gi0/5 192.168.201.47 ATTACK 160
Total: 1 host
```

- 通过 **show nfpp ip-guard trusted-host** 可以查看不监控信任主机信息。

```
IP address mask
-----
192.168.201.46 255.255.255.255
Total: 1 record(s)
```

常见错误

18.4.3 配置ICMP抗攻击

配置效果

- ICMP 攻击识别分为基于主机和基于物理端口两个类别。基于主机方式是采用源 IP 地址/虚拟局域网号/端口三者结合来识别的。每种攻击识别都有限速水线和告警水线。当 ICMP 报文速率超过限速水线时，将被丢弃。当 ICMP 报文速率超过告警水线时，将打印警告信息，发送 TRAP，基于主机的攻击识别还会对攻击源头采取隔离措施。
- 配置 ICMP 抗攻击隔离,针对主机用户攻击下发硬件隔离表项,攻击报文不送 CPU，不转发。

注意事项

- 对于在全局与接口上同时存在配置的命令，接口优先级高于全局。
- 隔离默认关闭，若打开隔离功能，攻击用户将占用安全模块硬件表项。

配置方法

▾ 使能攻击检测

- 必须配置，默认打开。
- 支持全局配置以及单端口独立配置。
- 当关闭 ICMP 抗攻击功能时，系统将自动清除受监控的主机。

▾ 配置隔离时间

- 可选配置，默认关闭隔离功能。
- 在攻击用户报文流量超过 CPP 限速带宽时，可配置隔离时间，将报文直接硬件丢弃，避免占用带宽资源。
- 支持全局配置以及单端口独立配置。
- 如果把隔离时间从非零值改成零，将直接把相关端口的攻击者删除，而不是进行软件监控。

▾ 配置配置监控时间

- 必须配置，默认时间为 600 秒。
- 在配置了隔离时间时，攻击用户监控时间直接采用隔离时间，配置的监控时间不生效。
- 支持全局配置。

▾ 配置监控主机最大数目

- 必须配置，默认 20000 个。
- 提高监控主机最大数目，随实际监控主机数增加，处理监控用户需占用更多 CPU 资源。
- 支持全局配置
- 如果受监控主机数已经达到默认的 20000 个，此时管理员把受监控主机的最大数目设置成小于 20000，不会删除已有的受监控主机，而是打印信息“%ERROR: The value that you configured is smaller than current monitored hosts 20000 (配置的受监控主机数)，please clear a part of monitored hosts.”来提醒管理员配置没有生效，需要删除部分已经被监控的主机。
- 当受监控主机满时，打印日志 “% NFPP_ICMP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 (配置的受监控主机数) monitored hosts.” 提醒管理员

配置攻击检测水线

- 必须配置，存在默认配置参数，详细见产品特性文档。
- 支持全局配置以及单端口独立配置。
- 当管理员配置的限速水线大于攻击阈值时，打印命令提示信息 “%ERROR: rate limit is higher than attack threshold 500pps(配置的攻击阈值).” 提醒管理员。
- 当管理员配置的攻击阈值小于限速水线时，打印命令提示信息 “%ERROR: attack threshold is smaller than rate limit 300pps(配置的限速水线).” 提醒管理员。
- 当无法为检测到的攻击者分配内存时，打印日志 “%NFPP_ICMP_GUARD -4-NO_MEMORY: Failed to alloc memory.” 提醒管理员。
- 基于源 IP 地址限速优先级高于基于端口限速。

配置信任用户

- 可选配置，默认无不监控可信主机。
- ICMP 防扫描仅支持配置不进行监控 IP，最多可配置 500 条。
- 支持全局配置。
- 当受监控主机表中存在与可信主机相匹配的表项（IP 地址相同）时，系统将自动删除此 IP 地址对应的表项。
- 当不监控的可信主机表满时，打印提示信息 “%ERROR: Attempt to exceed limit of 500 trusted hosts.” 提醒管理员。
- 当删除可信主机失败时，打印提示信息 “%ERROR: Failed to delete trusted host 1.1.1.0 255.255.255.0 (配置的可信主机).” 提醒管理员。
- 当添加可信主机失败时，打印提示信息 “%ERROR: Failed to add trusted host 1.1.1.0 255.255.255.0 (配置的可信主机).” 提醒管理员。
- 当添加的可信主机已经存在时，打印提示信息 “%ERROR: Trusted host 1.1.1.0 255.255.255.0 (配置的可信主机) has already been configured.” 提醒管理员。
- 当要删除的可信主机不存在时，打印提示信息 “%ERROR: Trusted host 1.1.1.0 255.255.255.0(配置的可信主机) is not found.” 提醒管理员。
- 当无法为可信主机分配内存时，打印提示信息 “%ERROR: Failed to alloc memory.” 提醒管理员。

检验方法

网络主机往配置了 ICMP 攻击检测限速的交换机发送 ICMP 攻击报文，需确认该报文可送 CPU。

- 对于不满足信任用户配置的报文，若超过攻击水位，将有攻击信息提示。
- 若攻击用户需生成隔离标项，将有用户隔离信息提示。

相关命令

使能攻击检测

【命令格式】 **icmp-guard enable**

【参数说明】 -

【命令模式】 nfpp 模式下

【使用指导】 -

配置隔离时间

【命令格式】 **icmp-guard isolate-period [seconds | permanent]**

【参数说明】 *seconds*: 隔离时间，单位是秒，取值范围是 0 或者[30, 86400]，0 表示不隔离。

permanent : 永久隔离。

【命令模式】 nfpp 模式下

【使用指导】 对攻击者的隔离时间分为全局隔离时间和基于端口的隔离时间（即局部隔离时间）。对于某个端口，如果没有配置基于端口的隔离时间，那么采用全局隔离时间；否则，采用基于端口的隔离时间。

配置监控时间

【命令格式】 **icmp-guard monitor-period seconds**

【参数说明】 *seconds* : 监控时间，单位是秒，取值范围是[180, 86400]。

【命令模式】 nfpp 模式下

【使用指导】 检测出攻击者的时候，如果隔离时间为 0，将对攻击者进行软件监控，超时为监控时间。在软件监控过程中，当隔离时间被配置为非零值时，将自动对软件监控的攻击者采取硬件隔离，并且把超时设置为隔离时间。监控时间在隔离时间为 0 的情况下才有意义。

如果把隔离时间从非零值改成零，将直接把相关端口的攻击者删除，而不是进行软件监控。

配置监控主机最大数目

【命令格式】 **icmp-guard monitored-host-limit number**

【参数说明】 *number* : 支持的最大受监控主机数，取值范围为 1 到 4294967295。

【命令模式】 nfpp 模式下

【使用指导】 如果受监控主机数已经达到默认的 20000 个，此时管理员把受监控主机的最大数目设置成小于 20000，不会删除已有的受监控主机，而是打印信息 “%ERROR : The value that you configured is smaller than current monitored hosts 20000（配置的受监控主机数），please clear a part of monitored hosts.” 来提醒管理员配置没有生效，需要删除部分已经被监控的主机。

当受监控主机满时,打印日志“% NFPP_ICMP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 (配置的受监控主机数) monitored hosts.”提醒管理员。

配置限速水线

【命令格式】 **icmp-guard rate-limit {per-src-ip | per-port} pps**

【参数说明】 **per-src-ip** : 对每个源 IP 地址进行限速。

per-port : 对每个端口进行限速。

pps : 限速水线值,取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置攻击水线

【命令格式】 **icmp-guard attack-threshold {per-src-ip | per-port} pps**

【参数说明】 **per-src-ip** : 配置每个源 IP 地址的攻击水线。

per-port : 配置每个端口的攻击水线。

pps : 攻击水线,单位是每秒报文数,取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置信任用户

【命令格式】 **icmp-guard trusted-host ip mask**

【参数说明】 *ip* : IP 地址。

mask : : IP 地址的掩码。

all : 和 **no** 一起使用,删除所有可信主机配置。

【命令模式】 nfpp 模式下

【使用指导】 如果管理员希望对某台主机不进行监控,即对该主机表示信任,则可以通过该命令配置。该可信主机发往 CPU 的 ICMP 报文将被允许发往 CPU,不做任何的限速和告警处理。通过配置掩码可以达到对某一个网段的所有主机都不进行监控。

最多支持设置 500 条可信主机。

配置接口上使能攻击检测

【命令格式】 **nfpp icmp-guard enable**

【参数说明】 -

【命令模式】 接口模式下

【使用指导】 端口的 ICMP 抗攻击开关优先于全局 ICMP 抗攻击开关。

配置接口上隔离时间

【命令格式】 **nfpp icmp-guard isolate-period [seconds | permanent]**

【参数说明】 *seconds* : 隔离时间,单位是秒,取值范围是 0 或者[30, 86400], 0 表示不隔离。

permanent : 永久隔离。

【命令模式】 接口模式下

【使用指导】 -

配置接口上攻击策略

【命令格式】 **nfpp icmp-guard policy {per-src-ip | per-port} rate-limit-pps attack-threshold-pps**

【参数说明】 **per-src-ip** : 配置每个源 IP 地址的限速水线和攻击水线。

per-port : 配置每个端口的限速水线和攻击水线。

rate-limit-pps : 限速水线, 取值范围是 1 到 19999。

attack-threshold-pps : 攻击水线, 取值范围是 1 到 19999。

【命令模式】 接口模式下

【使用指导】 攻击水线不能小于限速水线。

配置举例

通过 ICMP 抗攻击保护 CPU

【网络环境】

- 系统中带有 ICMP 主机用户攻击, 部分用户无法 ping 通。
- 系统中, 部分主机报文流量很大, 需放行

【配置方法】

- 配置基于主机的攻击检测水线。
- 配置隔离时间为非 0。
- 配置不监控可信主机

```
Ruijie# configure terminal
Ruijie(config)# nfpp
Ruijie (config-nfpp)#icmp-guard rate-limit per-src-ip 20
Ruijie (config-nfpp)#icmp-guard attack-threshold per-src-ip 30
Ruijie (config-nfpp)#icmp-guard isolate-period 180
Ruijie (config-nfpp)#icmp-guard trusted-host 192.168.201.46 255.255.255.255
```

【检验方法】

- 通过 **show nfpp icmp-guard summary** 可以查看到配置信息。

(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.)

Interface	Status	Isolate-period	Rate-limit	Attack-threshold
Global	Disable	180	20/-/400	30/-/400

Maximum count of monitored hosts: 1000

Monitor period: 600s

- 通过 **show nfpp icmp-guard hosts** 可以查看到监控用户。

If col_filter 1 shows '*', it means "hardware do not isolate host".

VLAN	interface	IP address	remain-time(s)
1	Gi0/5	192.168.201.47	160

Total: 1 host

- 通过 **show nfpp icmp-guard trusted-host** 可以查看不监控信任主机信息。

IP address	mask
------------	------

```
-----  
-----  
192.168.201.46      255.255.255.255  
Total: 1 record(s)
```

常见错误

18.4.4 配置DHCP抗攻击

配置效果

- DHCP 攻击识别分为基于主机和基于物理端口两个类别。基于主机方式是采用链路层源 MAC 地址/虚拟局域网号/端口三者结合来识别的。每种攻击识别都有限速水线和告警水线。当 DHCP 报文速率超过限速水线时，超限的 DHCP 报文将被丢弃。当 DHCP 报文速率超过告警水线时，将打印警告信息，发送 TRAP，基于主机的攻击识别还会对攻击源头采取隔离措施。
- 配置 DHCP 抗攻击隔离,针对主机用户攻击下发硬件隔离表项,攻击报文不送 CPU，不转发。

注意事项

- 对于在全局与接口上同时存在配置的命令，接口优先级高于全局。
- 隔离默认关闭，若打开隔离功能，攻击用户将占用安全模块硬件表项。
- 对 DHCP Snooping 配置的信任口，其 DHCP 抗攻击功能不生效，避免对信任口的 DHCP 流进行误判。关于 DHCP Snooping 的信任口，具体参见“DHCP Snooping 配置”的“配置 DHCP Snooping 基本功能”章节。

配置方法

▾ 使能攻击检测

- 必须配置，默认打开。
- 支持全局配置以及单端口独立配置。
- 当关闭 DHCP 抗攻击功能时，系统将自动清除受监控的主机。

▾ 配置隔离时间

- 可选配置，默认关闭隔离功能。
- 在攻击用户报文流量超过 CPP 限速带宽时，可配置隔离时间，将报文直接硬件丢弃，避免占用带宽资源。
- 支持全局配置以及单端口独立配置。
- 如果把隔离时间从非零值改成零，将直接把相关端口的攻击者删除，而不是进行软件监控。

配置配置监控时间

- 必须配置，默认打开。
- 在配置了隔离时间时，攻击用户监控时间直接采用隔离时间，配置的监控时间不生效。
- 支持全局配置。

配置监控主机最大数目

- 必须配置，默认 20000 个。
- 提高监控主机最大数目，随实际监控主机数增加，处理监控用户需占用更多 CPU 资源。
- 支持全局配置
- 如果受监控主机数已经达到默认的 20000 个，此时管理员把受监控主机的最大数目设置成小于 20000，不会删除已有的受监控主机，而是打印信息“%ERROR : The value that you configured is smaller than current monitored hosts 20000 (配置的受监控主机数) ， please clear a part of monitored hosts.”来提醒管理员配置没有生效，需要删除部分已经被监控的主机。
- 当受监控主机满时，打印日志 “% NFPP_DHCP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 (配置的受监控主机数) monitored hosts.” 提醒管理员。。

配置攻击检测水线

- 必须配置，存在默认配置参数，详细见产品特性文档。
- 支持全局配置以及单端口独立配置。
- 当管理员配置的限速水线大于攻击阈值时，打印命令提示信息 “%ERROR : rate limit is higher than attack threshold 500pps(配置的攻击阈值).” 提醒管理员。
- 当管理员配置的攻击阈值小于限速水线时，打印命令提示信息 “%ERROR : attack threshold is smaller than rate limit 300pps(配置的限速水线).” 提醒管理员。
- 当无法为检测到的攻击者分配内存时，打印日志 “%NFPP_DHCP_GUARD-4-NO_MEMORY: Failed to alloc memory.” 提醒管理员。
- 基于链路层源 MAC 地址限速优先于基于端口限速处理。

检验方法

网络主机往配置了 DHCP 攻击检测限速的交换机发送 DHCP 攻击报文，需确认该报文可送 CPU。

- 若超过攻击水线，将有攻击信息提示。
- 若攻击用户需生成隔离标项，将有用户隔离信息提示。

相关命令

使能攻击检测

【命令格式】 **dhcp-guard enable**

【参数说明】 -

【命令模式】 nfpp 模式下

【使用指导】 -

配置隔离时间

【命令格式】 **dhcp-guard isolate-period** [*seconds* | **permanent**]

【参数说明】 *seconds* : 隔离时间, 单位是秒, 取值范围是 0 或者[30, 86400], 0 表示不隔离。

permanent : 永久隔离。

【命令模式】 nfpp 模式下

【使用指导】 对攻击者的隔离时间分为全局隔离时间和基于端口的隔离时间(即局部隔离时间)。对于某个端口, 如果没有配置基于端口的隔离时间, 那么采用全局隔离时间; 否则, 采用基于端口的隔离时间。

配置监控时间

【命令格式】 **dhcp-guard monitor-period** *seconds*

【参数说明】 *seconds* : 监控时间, 单位是秒, 取值范围是[180, 86400]。

【命令模式】 nfpp 模式下

【使用指导】 检测出攻击者的时候, 如果隔离时间为 0, 将对攻击者进行软件监控, 超时为监控时间。在软件监控过程中, 当隔离时间被配置为非零值时, 将自动对软件监控的攻击者采取硬件隔离, 并且把超时设置为隔离时间。监控时间在隔离时间为 0 的情况下才有意义。

如果把隔离时间从非零值改成零, 将直接把相关端口的攻击者删除, 而不是进行软件监控。

配置监控主机最大数目

【命令格式】 **dhcp-guard monitored-host-limit** *number*

【参数说明】 *number* : 支持的最大受监控主机数, 取值范围为 1 到 4294967295。

【命令模式】 nfpp 模式下

【使用指导】 如果受监控主机数已经达到默认的 20000 个, 此时管理员把受监控主机的最大数目设置成小于 20000, 不会删除已有的受监控主机, 而是打印信息 “%ERROR : The value that you configured is smaller than current monitored hosts 20000 (配置的受监控主机数) , please clear a part of monitored hosts.” 来提醒管理员配置没有生效, 需要删除部分已经被监控的主机。

当受监控主机表满时, 打印日志 “% NFPP_DHCP_GUARD-4-SESSION_LIMIT: Attempt to exceed limit of 20000 (配置受监控主机数) monitored hosts.” 提醒管理员。

配置限速水线

【命令格式】 **dhcp-guard rate-limit** {*per-src-mac* | *per-port*} *pps*

【参数说明】 **per-src-mac** : 对每个源 MAC 地址进行限速。

per-port : 对每个端口进行限速。

pps : 限速水线值, 取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置攻击水线

【命令格式】 **dhcp-guard attack-threshold {per-src-mac | per-port} pps**

【参数说明】 **per-src-mac**：配置每个源 MAC 地址的攻击水线。
per-port：配置每个端口的攻击水线。
pps：攻击水线，单位是每秒报文数，取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

▾ 配置接口上使能攻击检测

【命令格式】 **nfpp dhcp-guard enable**

【参数说明】 -

【命令模式】 接口模式下

【使用指导】 端口的 DHCP 抗攻击开关优先于全局 DHCP 抗攻击开关。

▾ 配置接口上隔离时间

【命令格式】 **nfpp dhcp-guard isolate-period [seconds | permanent]**

【参数说明】 **seconds**：隔离时间，单位是秒，取值范围是 0 或者[30, 86400]，0 表示不隔离。
permanent：永久隔离。

【命令模式】 接口模式下

【使用指导】 -

▾ 配置接口上攻击策略

【命令格式】 **nfpp dhcp-guard policy {per-src-mac | per-port} rate-limit-pps attack-threshold-pps**

【参数说明】 **per-src-mac**：配置每个源 MAC 地址的限速水线和攻击水线。
per-port：配置每个端口的限速水线和攻击水线。
rate-limit-pps：限速水线，取值范围是 1 到 19999。
attack-threshold-pps：攻击水线，取值范围是 1 到 19999。

【命令模式】 接口模式下

【使用指导】 攻击水线不能小于限速水线。

配置举例

▾ 通过 DHCP 抗攻击保护 CPU

- 【网络环境】
- 系统中带有 DHCP 主机用户攻击，部分用户地址申请失败。
- 【配置方法】
- 配置基于主机的攻击检测水线。
 - 配置隔离时间为非 0。

```
Ruijie# configure terminal
Ruijie(config)# nfpp
Ruijie (config-nfpp)#dhcp-guard rate-limit per-src-mac 8
Ruijie (config-nfpp)#dhcp-guard attack-threshold per-src-mac 16
Ruijie (config-nfpp)#dhcp-guard isolate-period 180
```

- 【检验方法】
- 通过 show nfpp dhcp-guard summary 可以查看到配置信息。

```
(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.)
```

```
Interface Status Isolate-period Rate-limit Attack-threshold
Global Disable 180 -/8/150 -/16/300
```

```
Maximum count of monitored hosts: 1000
```

```
Monitor period: 600s
```

- 通过 **show nfpp dhcp-guard hosts** 可以查看到监控用户。

```
If col_filter 1 shows '*', it means "hardware do not isolate host".
```

```
VLAN interface MAC address remain-time(s)
----
*1 Gi0/5 001a.a9c2.4609 160
Total: 1 host
```

常见错误

18.4.5 配置ND抗攻击

配置效果

- AR ND guard 按用途把 ND 报文分成 3 类：邻居请求和邻居公告为第一类，路由器请求为第二类，路由器公告和重定向报文为第三类。第一类报文用于地址解析；第二类报文用于主机发现网关；路由器公告用于通告网关和前缀，重定向报文用于通告更优的下一跳，都和路由有关系。所以划分到第三类。
- 目前仅实现基于物理端口识别 ND 报文攻击。可以对三类报文分别配置限速水线和告警水线。当 ND 报文速率超过限速水线时，超限的 ND 报文将被丢弃。当 ND 报文速率超过告警水线时，将打印警告信息，发送 TRAP。。

注意事项

- 对于在全局与接口上同时存在配置的命令，接口优先级高于全局。
- 隔离默认关闭，若打开隔离功能，攻击用户将占用安全模块硬件表项。

配置方法

▾ 使能攻击检测

- 必须配置，默认打开。
- 支持全局配置以及单端口独立配置。

▾ 配置基于端口的隔离转发限速功能

- 可选配置，默认基于端口的隔离转发限速功能生效。
- 在基于端口的隔离表项生效时，希望隔离动作非全部丢弃，而是允许部分报文通过，可配置基于端口的隔离转发限速功能。
- 支持全局配置。

配置攻击检测水线

- 必须配置，存在默认配置参数，详细见产品特性文档。
- 支持全局配置以及单端口独立配置。
- 当管理员配置的限速水线大于攻击阈值时，打印命令提示信息 “%ERROR : rate limit is higher than attack threshold 500pps(配置的攻击阈值).” 提醒管理员。
- 当管理员配置的攻击阈值小于限速水线时，打印命令提示信息 “%ERROR : attack threshold is smaller than rate limit 300pps(配置的限速水线).” 提醒管理员。
- 当无法为检测到的攻击者分配内存时，打印日志 “%NFPP_ND_GUARD-4-NO_MEMORY: Failed to alloc memory..” 提醒管理员。

检验方法

网络主机往配置了 ND 攻击检测限速的交换机发送 ND 攻击报文，需确认该报文可送 CPU。

- 若超过攻击水线，有攻击信息示。

相关命令

使能攻击检测

- 【命令格式】 **nd-guard enable**
- 【参数说明】 -。
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

配置基于端口的隔离转发限速功能

- 【命令格式】 **nd-guard ratelimit-forwarding enable**
- 【参数说明】 -
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

配置 nd-guard 限速水线

- 【命令格式】 **nd-guard rate-limit per-port [ns-na | rs | ra-redirect] pps**
- 【参数说明】 **ns-na** : 邻居请求和邻居公告。
rs : 路由器请求。

ra-redirect : 路由器公告和重定向报文。
pps : 限速水线值, 取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 -

配置 nd-guard 攻击水线

【命令格式】 **ndguard attack-threshold per-port [ns-na | rs | ra-redirect] pps**

【参数说明】 **ns-na** : 邻居请求和邻居公告。

rs : 路由器请求。

ra-redirect : 路由器公告和重定向报文。

pps : 攻击水线, 单位是每秒报文数, 取值范围是[1,19999]。

【命令模式】 nfpp 模式下

【使用指导】 攻击水线不能小于限速水线。

配置接口上使能攻击检测

【命令格式】 **nfpp nd-guard enable**

【参数说明】 -

【命令模式】 接口模式下

【使用指导】 端口的 ND 抗攻击开关优先于全局开关。

配置接口上 nd-guard 攻击策略

【命令格式】 **nfpp nd-guard policy per-port [ns-na | rs | ra-redirect] rate-limit-pps attack-threshold-pps**

【参数说明】 **ns-na** : 邻居请求和邻居公告。

rs : 路由器请求。

ra-redirect : 路由器公告和重定向报文。

rate-limit-pps : 限速水线, 取值范围是 1 到 19999。

attack-threshold-pps : 攻击水线, 取值范围是 1 到 19999。

【命令模式】 接口模式下

【使用指导】 攻击水线不能小于限速水线。

ND snooping 把端口划分为非信任端口和信任端口, 非信任端口连接主机, 信任端口连接网关。由于通常信任端口的流量大于非信任端口的流量, 所以信任端口的限速水线应该高于非信任端口的限速水线, 开启 ND snooping 功能时, 对于信任端口, ND snooping 将通告 ND guard 把端口的三类报文限速水线都设置成每秒 800 个, 把攻击水线都设置成每秒 900 个。

ND guard 同等对待 ND snooping 设置的限速水线和管理员配置的限速水线, 后配置的值覆盖先配置的值, 并且保存到配置文件中。ND snooping 设置的攻击水线类似。

配置举例

通过 ND 抗攻击保护 CPU

【网络环境】 ● 系统中带有 ND 主机用户攻击, 部分用户邻居发现失败。

【配置方法】 ● 配置基于主机的攻击检测水线。

```
Ruijie# configure terminal
Ruijie(config)# nfpp
Ruijie (config-nfpp)# nd-guard rate-limit per-port ns-na 30
Ruijie (config-nfpp)# nd-guard attack-threshold per-port ns-na 50
```

- 【检验方法】 ● 通过 `show nfpp nd-guard summary` 可以查看到配置信息。

```
(Format of column Rate-limit and Attack-threshold is NS-NA/RS/RA-REDIRECT.)
Interface Status Rate-limit Attack-threshold
Global Disable 30/15/15
```

常见错误

18.4.6 配置自定义抗攻击

配置效果

- 通过配置自定义抗攻击类型，解决特殊应用场景中的网络攻击问题。

注意事项

- 对于在全局与接口上同时存在配置的命令，接口生效优先级高于全局。
- 隔离默认关闭，若打开隔离功能，攻击用户将占用安全模块硬件表项。
- 自定义分类检测优先级高于基本分类检测。在配置自定义分类匹配字段时，请尽量参考配置指导。

配置方法

▾ 配置自定义抗攻击类型名称

- 必须配置，配置自定义抗攻击类型名称，同时也是创建自定义抗攻击分类。
- 自定义抗攻击类型的名字不能重复，`match` 匹配的字段及值也不能完全相同，也不能与 `arp`、`icmp`、`dhcp`、`ip` 抗攻击类型相同。当配置类型重复时，将提示配置失败

▾ 配置自定义抗攻击关键字

- 必须配置。
- 自定义报文的类型可以由以太网链路层类型字段 `etype`、源 MAC 地址 `smac`、目的 MAC 地址 `dmac`、IPv4 协议号 `protocol`、源 IPv4 地址 `sip`、目的 IPv4 地址 `dip`、源传输层端口 `sport` 和目的传输层端口 `dport` 这些字段组合而成。
- `protocol` 仅在 `etype` 为 `ipv4` 时才有效；`src-ip`、`dst-ip` 仅在 `etype` 为 `ipv4` 时有效；`src-port`、`dst-port` 仅在 `protocol` 为 `tcp` 或者 `udp` 时有效

- 当自定义抗攻击类型的 match 字段及值与已存在的抗攻击类型完全一样时,打印提示信息:“%ERROR :the match type and value are the same with define name(已存在的的抗攻击类型名).”,提示管理员配置失败。
- 当 match 配置了 protocol,但 etype 不是 IPv4 时,打印提示信息:“%ERROR :protocol is valid only when etype is IPv4(0x0800).”
- 当 match 配置了 src-ip、dst-ip,但 etype 不是 IPv4 时,打印提示信息:“%ERROR :IP address is valid only when etype is IPv4(0x0800).”
- 当 match 配置了 src-port、dst-port,但 protocol 不是 TCP 或者 UDP 时,打印提示信息:“%ERROR :Port is valid only when protocol is TCP(6) or UDP(17).”
- 下面列出一些常用的网络协议对应的抗攻击策略。其中对应的限速水线与攻击水线能满足大部分网络应用场景需求,仅供参考。网络管理员应该根据实际应用场景配置有效的限速水线与攻击水线。

协议名	match	policy per-src-ip	policy per-src-mac	policy per-port
RIP	etype 0x0800 protocol 17 dst-port 520	rate-limit 100 attatch-threshold 150	不适用本策略	rate-limit 300 attatch-threshold 500
BPDU	dst-mac 0180.c200.0000	不适用本策略	rate-limit 20 attatch-threshold 40	rate-limit 100 attatch-threshold 100
RERP	dst-mac 01d0.f800.0001	不适用本策略	rate-limit 20 attatch-threshold 40	rate-limit 100 attatch-threshold 100
REUP	dst-mac 01d0.f800.0007	不适用本策略	rate-limit 20 attatch-threshold 40	rate-limit 100 attatch-threshold 100
OSPFv2	etype 0x0800 protocol 89	rate-limit 800 attatch-threshold 1200	不适用本策略	rate-limit 2000 attatch-threshold 3000
VRRP	etype 0x0800 protocol 112	rate-limit 64 attatch-threshold 100	不适用本策略	rate-limit 1024 attatch-threshold 1024
SNMP	etype 0x0800 protocol 17 dst-port 161	rate-limit 1000 attatch-threshold 1200	不适用本策略	rate-limit 2000 attatch-threshold 3000
RSVP	etype 0x0800 protocol 46	rate-limit 800 attatch-threshold 1200	不适用本策略	rate-limit 1200 attatch-threshold 1500
LDP (UDP hello)	etype 0x0800 protocol 17 dst-port 646	rate-limit 10 attatch-threshold 15	不适用本策略	rate-limit 100 attatch-threshold 150

- 自定义抗攻击为了最大限度地包含已有的协议类型,同是便于新的协议类型的扩展,开放允许用户自由组合报文的类型字段。若是配置不当,则可能导致网络出现异常。因此要求网络管理员对网络协议有较好的掌握。常用自定义抗攻击策略列出了当前已知协议的有效配置,管理员可参照进行配置。对于表中未列出的其它协议,需要谨慎配置使用

配置基于主机或基于端口的限速水线与攻击水线

- 必须配置,否则自定义抗攻击无法使能。
- per-src-ip、per-src-mac 和 per-port 三者至少要配置一个,否则策略无法生效。

- **per-src-ip** 仅在 **etype** 为 **ipv4** 时有效。
- 基于源 MAC/VID/端口的限速优先级高于基于源 IP/VID/端口的限速。
- 自定义抗攻击的端口识别主机策略要与全局的保持一致。
- 若全局没配置 **per-src-ip** 策略，端口配置 **per-src-ip** 策略时，打印提示 “%ERROR: name(自定义抗攻击名字) has not per-src-ip policy.”，提醒管理员配置失败。
- 若全局没配置 **per-src-mac** 策略，端口配置 **per-src-mac** 策略时，打印提示 “%ERROR: name(自定义抗攻击名字) has not per-src-mac policy.”，提醒管理员配置失败。
- 当无法为检测到的攻击者分配内存时，打印日志 “%NFPP_DEFINE_GUARD-4-NO_MEMORY: Failed to allocate memory.” 提醒管理员。
- 当管理员配置的限速水线大于攻击阈值时，打印命令提示信息 “%ERROR : rate limit is higher than attack threshold 500pps(配置的攻击阈值).” 提醒管理员。
- 当管理员配置的攻击阈值小于限速水线时，打印命令提示信息 “%ERROR : attack threshold is smaller than rate limit 300pps(配置的限速水线).” 提醒管理员。

配置配置监控时间

- 必须配置，默认时间为 600 秒。
- 在配置了隔离时间时，攻击用户监控时间直接采用隔离时间，配置的监控时间不生效。
- 支持全局配置。
- 检测出攻击者的时候，如果隔离时间为 0，将对攻击者进行软件监控，超时为监控时间。在软件监控过程中，当隔离时间被配置为非零值时，将自动对软件监控的攻击者采取硬件隔离，并且把超时设置为隔离时间。监控时间在隔离时间为 0 的情况下才有意义。
- 如果把隔离时间从非零值改成零，将直接把相关端口的攻击者删除，而不是进行软件监控。

配置监控主机最大数目

- 必须配置，默认 20000 个。
- 提高监控主机最大数目，随实际监控主机数增加，处理监控用户需占用更多 CPU 资源。
- 支持全局配置
- 如果受监控主机数已经达到默认的 20000 个，此时管理员把受监控主机的最大数目设置成小于 20000，不会删除已有的受监控主机，而是打印信息 “%ERROR : The value that you configured is smaller than current monitored hosts 20000 (配置的受监控主机数) ， please clear a part of monitored hosts.” 来提醒管理员配置没有生效，需要删除部分已经被监控的主机。
- 当受监控主机满时，打印日志 “% NFPP_DEFINE-4-SESSION_LIMIT: Attempt to exceed limit of name(自定义抗攻击类型名)'s 20000 (配置的受监控主机数) monitored hosts.” 提醒管理员。

配置信任用户

- 可选配置，默认无不监控可信主机。

- 自定义抗攻击支持配置不进行监控 IP 地址或 MAC 地址，最多可配置 500 条。
- 支持全局配置。
- 如果管理员希望对某台主机不进行监控，即对该主机表示信任，则可以通过该命令配置。该可信主机发往 CPU 的 ICMP 报文将被允许发往 CPU，不做任何的限速和告警处理。通过配置掩码可以达到对某一个网段的所有主机都不进行监控。
- 必须先配置 match 类型后后才能配置 trusted-host。
- 当还未配置 match 类型时，打印提示信息 “%ERROR: Please configure match rule first.”
- 当添加 IPv4 可信主机，但 match 规则的 etype 不为 IPv4 时，打印提示信息 “%ERROR: Match type can't support IPv4 trusted host.”
- 当不监控的可信主机表满时，打印提示信息 “%ERROR: Attempt to exceed limit of 500 trusted hosts.” 提醒管理员。
- 当受监控主机表中存在与可信主机相匹配的表项（IP 地址相同）时，系统将自动删除此 IP 地址对应的表项。
- 当删除可信主机失败时，打印提示信息 “%ERROR: Failed to delete trusted host 1.1.1.0 255.255.255.0 (配置的可信主机).” 提醒管理员。
- 当添加可信主机失败时，打印提示信息 “%ERROR: Failed to add trusted host 1.1.1.0 255.255.255.0 (配置的可信主机).” 提醒管理员。
- 当添加的可信主机已经存在时，打印提示信息 “%ERROR: Trusted host 1.1.1.0 255.255.255.0 (配置的可信主机) has already been configured.” 提醒管理员。
- 当要删除的可信主机不存在时，打印提示信息 “%ERROR: Trusted host 1.1.1.0 255.255.255.0(配置的可信主机) is not found.” 提醒管理员。
- 当无法为可信主机分配内存时，打印提示信息 “%ERROR: Failed to allocate memory.” 提醒管理员。

配置自定义抗攻击使能

- 必须配置。
- 基于主机或端口的自定义抗攻击策略必须至少配置一个，否则使能失败。
- 当关闭自定义抗攻击功能时，系统将自动清除受监控的主机。
- 支持全局配置以及单端口独立配置。
- 当自定义抗攻击策略还没配置完全时，无法打开自定义抗攻击功能，并提示用户缺少相应的策略配置。
- 当自定义抗攻击名字不存在时，打印提示信息：“%ERROR: The name is not exist.”。
- 当自定义抗攻击未配置 match 类型时，打印提示信息：“%ERROR: name (自定义抗攻击类型名字) doesn't match any type.”。
- 当自定义抗攻击未配置 policy 策略时，打印提示信息：“%ERROR: name (自定义抗攻击类型名字) doesn't specify any policy.”。

检验方法

网络主机往配置了 NFPP 攻击检测限速的交换机发送符合自定义抗攻击类型的报文，需确认该报文可送 CPU 处理。

- 对于不满足信任用户配置的报文，若超过攻击水线，将有攻击信息提示。
- 若攻击用户需生成隔离标项，将有用户隔离信息提示。

相关命令

配置自定义抗攻击类型名称

- 【命令格式】 **define name**
- 【参数说明】 **name**：表示自定义抗攻击名称的字符串。
- 【命令模式】 在 nfpp 模式下
- 【使用指导】 -

配置自定义抗攻击匹配报文字段

- 【命令格式】 **match [etype type] [src-mac smac [src-mac-mask smac_mask]] [dst-mac dmac [dst-mac-mask dst_mask]] [protocol protocol] [src-ip sip [src-ip-mask sip-mask]] [dst-ip dip [dst-ip-mask dip-mask]] [src-port sport] [dst-port dport]**
- 【参数说明】 *type*：以太网链路层报文类型。
smac：源 mac 地址。
smac_mask：源 mac 地址掩码。
dmac：目的 mac 地址。
dst_mask：目的 mac 地址掩码。
protocol：IPv4 报文协议号
sip：源 IPv4 地址。
sip-mask：源 IPv4 地址掩码。
dip：目的 IPv4 地址。
dip-mask：目的 IPv4 地址掩码。
sport：源传输层端口号。
dport：目的传输层端口号。
- 【配置模式】 自定义抗攻击类型模式下
- 【使用指导】 创建新的自定义抗攻击类型。指定该类型所要匹配的报文字段。

配置基于主机或基于端口的限速水线与攻击水线

- 【命令格式】 **global-policy {per-src-ip | per-src-mac | per-port} rate-limit-pps attack-threshold-pps**
- 【参数说明】 **per-src-ip**：基于源 IP/VID/端口识别主机进行速率统计。
per-src-mac：基于源 MAC/VID/端口识别主机进行速率统计。
per-port：基于每个报文接收的物理端口进行速率统计。
rate-limit-pps：限速水线。
attack-threshold-pps：攻击水线。
- 【命令模式】 自定义抗攻击类型模式下

- 【使用指导】 创建一个自定义抗攻击类型，必须为该类型指定速率统计的分类原则，即基于源 IP 识别用户、基于源 MAC 识别用户，进行基于用户的自定义报文速率统计，或者基于端口的速率统计，并且指定各分类的限速水线及攻击水线。

配置监控时间

- 【命令格式】 **monitor-period** *seconds*
- 【参数说明】 *seconds*：监控时间，单位是秒，取值范围是[180, 86400]。
- 【命令模式】 自定义抗攻击类型模式下
- 【使用指导】 -

配置监控主机最大数目

- 【命令格式】 **monitored-host-limit** *number*
- 【参数说明】 *number*：支持的最大受监控主机数，取值范围为 1 到 4294967295。
- 【命令模式】 自定义抗攻击类型模式下
- 【使用指导】 -

配置信任用户

- 【命令格式】 **trusted-host** {*mac mac_mask* | *ip mask*}
- 【参数说明】 *mac*：mac 地址。
mac_mask：mac 地址的掩码。
ip：IP 地址。
mask：IP 地址的掩码。
all：和 **no** 一起使用，删除所有可信主机配置。
- 【命令模式】 自定义抗攻击类型模式下
- 【使用指导】 -

配置自定义抗攻击使能

- 【命令格式】 **define** *name* **enable**
- 【参数说明】 *name*：自定义的抗攻击类型名称。
- 【命令模式】 nfpp 模式下
- 【使用指导】 必须配置了 **match**、**rate-count**、**rate-limit** 和 **attack-threshold** 之后，配置才能生效，否则配置失败。

配置接口上使能攻击检测

- 【命令格式】 **nfpp define** *name* **enable**
- 【参数说明】 *name*：自定义的抗攻击类型名称。
- 【命令模式】 接口模式下
- 【使用指导】 必须存在该自定义名字，并且配置了 **match**、**rate-count**、**rate-limit** 和 **attack-threshold** 之后，配置才能生效，否则配置失败。

配置接口上攻击策略

- 【命令格式】 **nfpp define** *name* **policy** {**per-src-ip** | **per-src-mac** | **per-port**} *rate-limit-pps* *attack-threshold-pps*

- 【参数说明】 *name* : 自定义抗攻击类型名称。
per-src-ip : 配置每个源 IP 地址的限速水线和攻击水线。
per-src-mac : 配置每个源 MAC 地址的限速水线和攻击水线。
per-port : 配置每个端口的限速水线和攻击水线。
rate-limit-pps : 限速水线, 取值范围是 1 到 19999。
attack-threshold-pps : 攻击水线, 取值范围是 1 到 19999。
- 【命令模式】 接口模式下
- 【使用指导】 攻击水线不能小于限速水线。

配置举例

通过自定义抗攻击保护 CPU

- 【网络环境】 ● 系统中有 rip 攻击, 使用基本抗攻击无法实现防护功能。
- 【配置方法】 ● 配置自定义抗攻击, 关键字段与 rip 报文符合。
- 配置限速水线
 - 配置隔离时间
 - 配置信任用户

```
Ruijie# configure terminal
Ruijie(config)# nfpp
Ruijie (config-nfpp)#define rip
Ruijie (config-nfpp-define)#match etype 0x0800 protocol 17 dst-port 520
Ruijie (config-nfpp-define)#global-policy per-src-ip 100 150
Ruijie (config-nfpp-define)#isolate-period 180
Ruijie (config-nfpp-define)#trusted-host 192.168.201.46 255.255.255.255
Ruijie (config-nfpp-define)#exit
Ruijie (config-nfpp)#define rip enable
```

- 【检验方法】 ● 通过 **show nfpp define summary rip** 可以查看到配置信息。

```
Define rip summary:
match etype 0x800 protocol 17 dst-port 520
Maximum count of monitored hosts: 1000
Monitor period:600s
(Format of column Rate-limit and Attack-threshold is per-src-ip/per-src-mac/per-port.)
Interface Status Isolate-period Rate-limit Attack-threshold
Global Enable 180 100/-/- 150/-/-
```

- 通过 **show nfpp define trusted-host rip** 可以查看到不监控信任主机信息。

```
Define rip:
IP trusted host number is 1:
IP address IP mask
-----
192.168.201.46 255.255.255.255
```

```
Total: 1 record(s)Global Enable 180 100/-/ 150/-/
```

- 通过 show nfpp define hosts rip 可以查看到监控用户信息。

```
If col_filter 1 shows '*', it means "hardware do not isolate host".
```

VLAN	interface	IP address	remain-time(s)
1	Gi0/5	192.168.201.47	160

```
Total: 1 host
```

常见错误

18.4.7 配置集中限速分发

配置效果

- 通过配置集中限速分发，解决网络忙情况下管理报文和协议报文优先处理问题。

注意事项

- 配置某类型所占百分比的有效值区间，必须小于等于百分之百减去其它两种类型百分比之和的差值。

配置方法

配置每类报文允许的最大带宽

- 必须配置，管理类(Manage)、转发类(Route)、协议类(Protocol)的缺省流量带宽是一样的，具体数值参见产品特性文档。

配置每类报文占用队列的最大百分比

- 必须配置，默认情况下管理类(Manage)占用 30%，转发类(Route)占用 25%、协议类(Protocol)占用 45%。

检验方法

往交换机设备发送大量协议报文，如 ospf，导致系统 CPU 高。

- 这时主机 PING 交换机设备，应保持 PING 通状态，且不丢包。

相关命令

配置每类报文允许的最大带宽

【命令格式】 **cpu-protect sub-interface { manage | protocol|route} pps pps_vaule**

【参数说明】 *pps_vaule* : 限速水线, 取值范围是 1-100000。

【命令模式】 在全局模式下

【使用指导】 -

配置每类报文占用队列的最大百分比

【命令格式】 **cpu-protect sub-interface { manage | protocol | route} percent percent_vaule**

【参数说明】 *percent_vaule* : 百分比, 取值范围是 1 到 100。

【配置模式】 在全局模式下

【使用指导】 配置某类型所占百分比的有效值区间, 必须小于等于百分之百减去其它两种类型百分比之和的差值。

配置举例

通过集中配置分发, 为送 CPU 报文分优先级

【网络环境】 ● 网络中存在多种大流量报文, 分属不同集中分类。

【配置方法】 ● 配置每类报文允许的最大带宽。

● 配置每类报文占用队列的最大百分比

```
Ruijie# configure terminal
Ruijie(config)# cpu-protect sub-interface manage pps 5000
Ruijie(config)# cpu-protect sub-interface manage percent 25
```

【检验方法】 略

常见错误

-

18.4.8 NFPP日志信息

配置效果

● NFPP 以一定速率从专用缓冲区取出日志, 生成系统消息, 并且从专用日志缓冲区清除这条日志。

注意事项

● 记录在缓冲区中日志会持续打出, 即使这时攻击已停止。

配置方法

配置日志缓冲区容量

- 必须配置。
- 若缓冲区已满，新生成日志将丢弃，并有信息提示。
- 当日志缓冲区溢出时，后续的日志将被丢弃，同时在日志缓冲区中显示一条所有属性都为“-”的表项。管理员需要增加日志缓冲区容量或者提高生成系统消息的速率。

配置生成系统消息的速率

- 必须配置。
- 由两参数决定：时间段长度以及该时间段内生成系统消息个数。
- 若两参数均为 0，表示日志立即生成系统消息，不入缓冲区。

配置日志过滤

- 可选配置，默认情况下不做过滤。
- 支持基于端口的过滤规则以及基于 vlan 的过滤规则。
- 若配置过滤，不符合过滤规则的日志丢弃。

配置日志打印到屏幕

- 必须配置，默认日志保存在缓冲区
- 若希望实时监控攻击是否存在，可配置将日志打印到屏幕，实时输出日志信息。

检验方法

通过查看日志配置信息与日志打印输出的个数与间隔时间，确认配置是否生效。

相关命令

配置日志缓冲区容量

- 【命令格式】 **log-buffer entries** *number*
- 【参数说明】 *umber*：缓冲区大小，单位是日志条数，取值范围是[0,1024]。
- 【命令模式】 nfpp 模式下
- 【使用指导】 -

配置生成系统消息的速率

- 【命令格式】 **log-buffer logs** *number_of_message interval length_in_seconds*
- 【参数说明】 *number_of_message*：范围为 0-1024，0 表示日志全部记录在专用缓冲区，不生成系统消息。
length_in_seconds：范围为 0-86400（1 天），0 表示不把日志写到缓冲区，而是立即生成系统消息。
number_of_message 和 *length_in_seconds* 都为 0 表示不把日志写到缓冲区，而是立即生成系统消息。
number_of_message /length_in_second 表示取日志生成系统消息的速率。
- 【命令模式】 Nfpp 模式下

【使用指导】

📌 配置日志基于 vlan 过滤

【命令格式】 **logging vlan *vlan-range***【参数说明】 *vlan-range* : 需要记录指定 VLAN 范围内的日志信息，输入格式如 “1-3,5”。

【命令模式】 Nfpp 模式

【使用指导】 通过该命令可以对日志进行过滤，只记录指定 VLAN 范围的日志信息。与端口范围日志过滤配置是或的关系，即只需要满足其中一条日志过滤规则，就应该记录到日志缓冲区中。

📌 配置日志基于端口过滤

【命令格式】 **logging interface *interface-id***【参数说明】 *interface-id* : 需要记录指定端口的日志信息。

【命令模式】 Nfpp 模式

【使用指导】 通过该命令可以对日志进行过滤，只记录指定端口的日志信息。与 vlan 范围日志过滤配置是或的关系，即只需要满足其中一条日志过滤规则，就应该记录到日志缓冲区中。

📌 配置日志打印到屏幕

【命令格式】 **log-buffer enable**

【参数说明】 -

【命令模式】 Nfpp 模式

【使用指导】 -。

配置举例

📌 通过 ND 抗攻击保护 CPU

【网络环境】 ● 当攻击用户过多时，日志打印影响用户界面使用，需进行限制

- 【配置方法】 ● 配置日志缓冲区容量。
-
- 配置生成系统消息的速率
-
- 配置日志基于 vlan 过滤

```
Ruijie# configure terminal
Ruijie(config)# nfpp
Ruijie (config-nfpp)#log-buffer entries 1024
Ruijie (config-nfpp)#log-buffer logs 3 interval 5
Ruijie (config-nfpp)#logging interface vlan 1
```

【检验方法】 ● 通过 **show nfpp logsummary** 可以查看到配置信息。

```
Total log buffer size : 1024
Syslog rate : 3 entry per 5 seconds
Logging:
  VLAN 1
```

- 通过
- show nfpp log buffer**
- 查看缓冲区中日志信息

Protocol	VLAN	Interface	IP address	MAC address	Reason	Timestamp
ARP	1	Gi0/5	192.168.206.2	001a.a9c2.4609	SCAN	2013-5-1 5:4:24

常见错误

18.5 监视与维护

清除各类信息

作用	命令
清除 arp-guard 扫描表。	<code>clear nfpp arp-guard scan</code>
清除 arp-guard 抗攻击受监控主机。	<code>clear nfpp arp-guard hosts</code>
清除 ip-guard 抗攻击受监控主机。	<code>clear nfpp ip-guard hosts</code>
清除 nd-guard 抗攻击受监控主机。	<code>clear nfpp nd-guard hosts</code>
清除 icmp-guard 抗攻击受监控主机。	<code>clear nfpp icmp-guard hosts</code>
清除 dhcp-guard 抗攻击受监控主机。	<code>clear nfpp dhcp-guard hosts</code>
清除自定义抗攻击受监控主机。	<code>clear nfpp define <i>name</i> hosts</code>
清除日志。	<code>clear nfpp log</code>

查看运行情况

作用	命令
查看 arp-guard 抗攻击的配置参数。	<code>show nfpp arp-guard summary</code>
查看 arp-guard 受监控主机的信息	<code>show nfpp arp-guard hosts</code>
查看 arp-guard 扫描表信息	<code>show nfpp arp-guard scan</code>
查看 ip-guard 抗攻击的配置参数。	<code>show nfpp ip-guard summary</code>
查看 ip-guard 受监控主机的信息	<code>show nfpp ip-guard hosts</code>
查看 ip-guard 扫描表信息	<code>show nfpp ip-guard trusted-host</code>
查看 icmp-guard 抗攻击的配置参数。	<code>show nfpp icmp-guard summary</code>
查看 icmp-guard 受监控主机的信息	<code>show nfpp icmp-guard hosts</code>
查看 icmp-guard 扫描表信息	<code>show nfpp icmp-guard trusted-host</code>
查看 dhcp-guard 抗攻击的配置参数。	<code>show nfpp dhcp-guard summary</code>
查看 dhcp-guard 受监控主机的信息	<code>show nfpp dhcp-guard hosts</code>
查看 nd-guard 抗攻击的配置参数。	<code>show nfpp nd-guard summary</code>
查看自定义抗攻击的配置参数	<code>show nfpp define summary [<i>name</i>]</code>
查看受监控主机的信息	<code>show nfpp define hosts <i>name</i></code>

查看不监控的可信主机	show nfpp define trusted-host <i>name</i>
查看 NFPP 日志信息配置	show nfpp log summary
显示 NFPP 的日志缓冲区	show nfpp log buffer [statistics]

19 DoS 保护

19.1 概述

DoS 是 Denial of Service 的简称，即拒绝服务，造成 DoS 的攻击行为被称为 DoS 攻击，其目的是使计算机或网络无法提供正常的服务。

DOS攻击种类繁多，具体的实现方式千变万化，但都有一个共同点，就是其根本目的是使受害 主机或网络无法及时接收并处理外界请求，或无法及时回应外界请求。特别是在二层网络中，DOS攻击报文可以在整个广播域内扩散，如果存在黑客恶意进行DOS攻击，可能会导致部分操作系统崩溃。我司产品支持以下防DOS攻击功能：

- 防 Land 攻击
- 防非法 TCP 报文攻击
- 防自身消耗攻击

协议规范

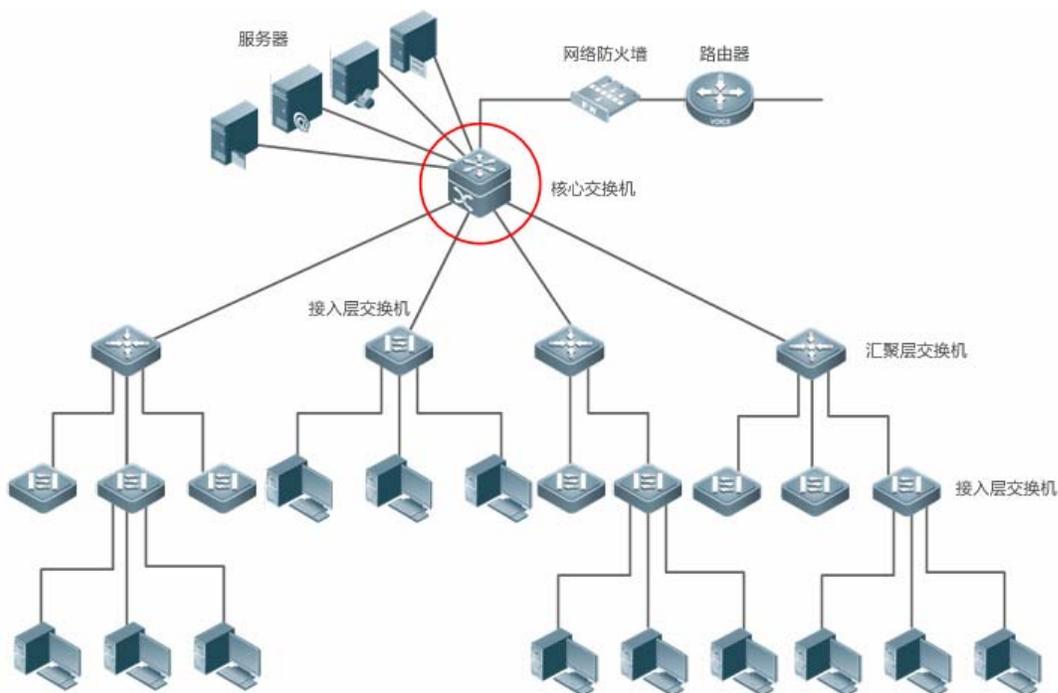
19.2 典型应用

典型应用	场景描述
防止对服务器进行DoS攻击	在园区网环境中，通过在服务器连接的设备上配置防 DOS 攻击，有效减少 DOS 攻击对服务器带来的影响。

19.2.1 防止对服务器进行DOS攻击

在下图所描述的环境中，服务器连接到其核心交换机上，通过在核心交换机上配置防 DOS 攻击功能，能防止用户的恶意 DOS 攻击，从而保障服务器正常提供服务。

图 19-1



功能部属

在核心交换机上开启防 Land 攻击功能，从而避免服务器受到 Land 攻击；

在核心交换机上开启防非法 TCP 报文攻击功能，从而避免服务器受到非法 TCP 报文攻击；

在核心交换机上开启防自身消耗攻击功能，从而避免服务器受到自身消耗攻击；

19.3 功能详解

基本概念

功能特性

功能特性	作用
防Land攻击	通过在设备上丢弃源 IP 地址和目的 IP 地址相同，四层源和目的端口号相同的报文，防止对网络中的操作系统进行攻击
防非法TCP报文攻击	通过在设备上丢弃非法 TCP 报文（具体定义见“防非法 TCP 报文攻击”章节），防止非法 TCP 报文对网络中的操作系统产生影响
防自身消耗攻击	通过在设备上丢弃四层源和目的端口号相同的报文，防止对网络中的操作系统进行攻击

19.3.1 防Land攻击

该功能用于防止 Land 攻击。

工作原理

Land 攻击主要是攻击者将一个 SYN 包的源 IP 地址和目的 IP 地址都设置为目标主机的地址，四层源和目的端口号设置为相同值，造成被攻击主机因试图与自己建立 TCP 连接而陷入死循环，甚至系统崩溃。

启动防 Land 攻击功能，设备会根据上述的 Land 报文的特征（源 IP 地址和目的 IP 地址相同，四层源和目的端口号相同）进行检查，若非法，则丢弃。

相关配置

▾ 启动防 Land 攻击功能

缺省情况下，防 Land 攻击功能关闭。

使用 `ip deny land` 命令可以启动或关闭防 Land 攻击功能。

19.3.2 防非法TCP报文攻击

该功能用于防止前文中说明的非法 TCP 报文攻击。

工作原理

在 TCP 报文的报头中，有几个标志字段：

- SYN：连接建立标志，TCP SYN 报文就是把这个标志设置为 1，来请求建立连接。
- ACK：回应标志，在一个 TCP 连接中，除了第一个报文（TCP SYN）外，所有报文都设置该字段作为对上一个报文的响应。
- FIN：结束标志，当一台主机接收到一个设置了 FIN 标志的 TCP 报文后，会拆除这个 TCP 连接。
- RST：复位标志，当 IP 协议栈接收到一个目标端口不存在的 TCP 报文的时候，会回应一个 RST 标志设置的报文。
- PSH：通知协议栈尽快把 TCP 数据提交给上层程序处理。

非法 TCP 报文攻击是通过非法设置标志字段致使主机处理的资源消耗甚至系统崩溃，例如以下几种经常设置的非法 TCP 报文：

- SYN 比特和 FIN 比特同时设置的 TCP 报文

正常情况下，SYN 标志（连接请求标志）和 FIN 标志（连接拆除标志）不能同时出现在一个 TCP 报文中，而且 RFC 也没有规定 IP 协议栈如何处理这样的畸形报文。因此各个操作系统的协议栈在收到这样的报文后的处理方式也不相同，攻击者就可以利用这个特征，通过发送 SYN 和 FIN 同时设置的报文，来判断操作系统的类型，然后针对该操作系统，进行进一步的攻击。

- 没有设置任何标志的 TCP 报文

正常情况下，任何 TCP 报文都会设置 SYN，FIN，ACK，RST，PSH 五个标志中的至少一个标志，第一个 TCP 报文（TCP 连接请求报文）设置 SYN 标志，后续报文都设置 ACK 标志。有的协议栈基于这样的假设，没有针对不设置任何标志的 TCP 报文的处理过程，因此这样的协议栈如果收到了这样的报文可能会崩溃。攻击者利用了这个特点，对目标主机进行攻击。

- 设置了 FIN 标志却没有设置 ACK 标志的 TCP 报文

正常情况下，除了第一报文（SYN 报文）外，所有的报文都设置 ACK 标志，包括 TCP 连接拆除报文（FIN 标志设置的报文）。但有的攻击者却可能向目标主机发送设置了 FIN 标志却没有设置 ACK 标志的 TCP 报文，这样可能导致目标主机崩溃。

- 设置了 SYN 标志且源端口号为 0-1023 的 TCP 报文

端口号 0-1023 是由 IANA 分配的已知端口号，并且在大多数系统中只能由系统（或根）进程或有特权的用户所执行的程序使用。这些端口（0-1023）是不能够作为客户端发送的第一个 TCP 报文（设置了 SYN 标志）的源端口号。

启动防非法 TCP 报文攻击功能，设备前会根据非 TCP 报文特征进行检查，若非法，则丢弃。

相关配置

启动防 Land 攻击功能

缺省情况下，防非法 TCP 报文攻击功能关闭。

使用 `ip deny invalid-tcp` 命令可以启动或关闭防非法 TCP 报文攻击功能。

19.3.3 防自身消耗攻击

该功能用于防止自身消耗攻击。

工作原理

- 自身消耗攻击主要是攻击者向目标主机发送与目标主机服务的 4 层端口号相同的报文，导致目标主机给自己发送 TCP 请求和连接。该攻击会使得目标主机的资源很快耗尽，甚至系统崩溃。

启动防自身消耗攻击功能，设备会检查报文的四层源端口和目的端口号，若相同，则将该报文丢弃。

相关配置

启动防自身消耗攻击功能

缺省情况下，防自身消耗攻击功能关闭。

使用 `ip deny invalid-l4port` 命令可以启动或关闭防自身消耗攻击功能。

19.4 产品说明



S7508 设备系列板卡的限制如下：

以下 4 种非法 TCP 报文：

- 1) SYN 比特和 FIN 比特同时设置的 TCP 报文
- 2) 没有设置任何标志的 TCP 报文
- 3) 设置了 FIN 标志却没有设置 ACK 标志的 TCP 报文
- 4) 设置了 SYN 标志且源端口号为 0-1023 的 TCP 报文

当开启预防非法 TCP 报文攻击时，BCM 系列产品能够完全过滤第 1、4 的非法 TCP 报文；对于第 2 种非法 TCP 报文，只能过滤没有设置任何标志且 sequence number 为 0 的 TCP 报文；对于第 3 种非法 TCP 报文，BCM 系列产品只能够过滤同时设置 FIN 标志、URG 标志和 PSH 标志且 sequence number 为 0 的 TCP 报文。

19.5 配置详解

配置项	配置建议 & 相关命令	
配置防Land攻击	 可选配置。	
	ip deny land	启动防 Land 攻击功能
配置防非法TCP报文攻击	 可选配置。	
	ip deny invalid-tcp	启动防非法 TCP 报文攻击功能
配置防自身消耗攻击	 可选配置。	
	ip deny invalid-l4port	启动防自身消耗攻击功能

19.5.1 配置防Land攻击

配置效果

启动防 Land 攻击功能，设备会根据 Land 报文的特征进行检查，若非法，则丢弃。

注意事项

-

配置方法

📌 启动防 Land 攻击功能

- 必须配置
- 一般在与服务器连接的设备上进行配置。

检验方法

- 通过 **show ip deny land** 命令可以查看防 Land 攻击状态。
- 开启该功能后，构造 Land 攻击报文，确认该报文无法转发。

相关命令

配置防 Land 攻击

- 【命令格式】 **[no] ip deny land**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

启动防 Land 攻击

- 在全局模式下启动防 Land 攻击

```
Ruijie# configure terminal
Ruijie(config)# ip deny land
Ruijie(config)# end
```

- 【检验方法】 通过 **show ip deny land** 命令可以查看防 Land 攻击状态。

下面的例子显示了如何查看防 Land 攻击的状态：

```
Ruijie# show ip deny land
          DoS Protection Mode          State
-----
protect against land attack          On
```

常见配置错误

-

19.5.2 防非法TCP报文攻击

配置效果

启动防非法 TCP 报文攻击功能，设备会根据 TCP 报文特征进行检查，若非法，则丢弃。

注意事项

配置方法

启动防非法 TCP 报文攻击功能

- 必须配置
- 一般在与服务器连接的设备上进行配置。

检验方法

- 通过 **show ip deny invalid-tcp** 命令可以查看防非法 TCP 报文攻击状态。
- 开启该功能后，构造非法 TCP 攻击报文，确认该报文无法转发。

相关命令

配置防非法 TCP 报文攻击

- 【命令格式】 **[no] ip deny invalid-tcp**
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

启动防非法 TCP 报文攻击功能

- 【配置方法】
 - 在全局模式下启动防非法 TCP 报文攻击功能

```
Ruijie# configure terminal
Ruijie(config)# ip deny invalid-tcp
Ruijie(config)# end
```

- 【检验方法】 通过 **show ip deny invalid-tcp** 命令可以查看防非法 TCP 报文攻击状态。

下面的例子显示了如何查看预防非法 TCP 报文攻击的状态：

```
Ruijie# show ip deny invalid-tcp
          DoS Protection Mode          State
-----
protect against invalid tcp attack    On
```

19.5.3 防自身消耗攻击

配置效果

启动防自身消耗攻击功能，设备会检查报文的四层源端口和目的端口号，若相同，则将该报文丢弃。打

注意事项

-

配置方法

▾ 启动防自身消耗攻击

- 必须配置
- 一般在与服务器连接的设备上进行配置。

检验方法

- 通过 `show ip deny invalid-l4port` 命令可以查看防自身消耗攻击状态。
- 开启该功能后，构造四层源端口与目的端口相同的报文，确认该报文无法转发。

相关命令

▾ 配置防自身消耗攻击

- 【命令格式】 `[no] ip deny invalid-l4port`
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

配置举例

▾ 启动防自身消耗攻击功能

- 【配置方法】
 - 在全局模式下启动防自身消耗攻击功能

```
Ruijie# configure terminal
Ruijie(config)# ip deny invalid-l4port
Ruijie(config)# end
```

【检验方法】 通过 **show ip deny invalid-l4port** 命令可以查看防自身消耗攻击状态。

下面的例子显示了如何查看预防自身消耗攻击的状态：

```
Ruijie# show ip deny invalid-l4port
      DoS Protection Mode           State
-----
protect against invalid l4port attack  On
```

19.6 监视与维护

查看运行情况

作用	命令
查看防 Land 攻击状态	show ip deny land
查看防非法 TCP 报文攻击状态	show ip deny invalid-tcp
查看防自身消耗攻击状态	show ip deny invalid-l4port
查看所有防 DOS 攻击状态	show ip deny



配置指南-ACL&QOS

本分册介绍 ACL&QOS 配置指南相关内容，包括以下章节：

1. ACL
2. QOS

1 ACL

1.1 概述

ACLs (Access Control Lists, 接入控制列表), 也称为访问列表 (Access Lists), 俗称为防火墙, 在有的文档中还称之为包过滤。通过定义一些规则对网络设备接口上的数据报文进行控制: 允许通过、丢弃。

根据使用 ACL 目的的不同可分为: 安全 ACLs 和 QoS ACLs。

- 安全 ACLs 用于控制哪些数据流允许从网络设备通过。
- QoS ACLs 对这些数据流进行优先级分类和处理。

配置访问列表的原因比较多, 最主要的主要有以下一些:

- 网络访问控制: 为了确保网络安全, 通过定义规则, 可以限制用户访问一些服务 (如只需要访问 WWW 和电子邮件服务, 其他服务如 TELNET 则禁止), 或者仅允许在给定的时间段内访问, 或只允许一些主机访问网络等等。
- 优先服务保证: 为一些重要的数据流进行优先分类处理, 这就是 QoS ACLs 作用。有关 QoS ACLs 的使用请参考 QoS 相关的配置手册。

 下文仅介绍 ACL 的相关内容。

协议规范

无

1.2 典型应用

典型应用	场景描述
企业内网访问控制应用	在企业网中根据需要对各个部门的网络访问权限进行控制和限制, 比如服务器的访问限制、QQ 和 MSN 等聊天工具的使用限制等。

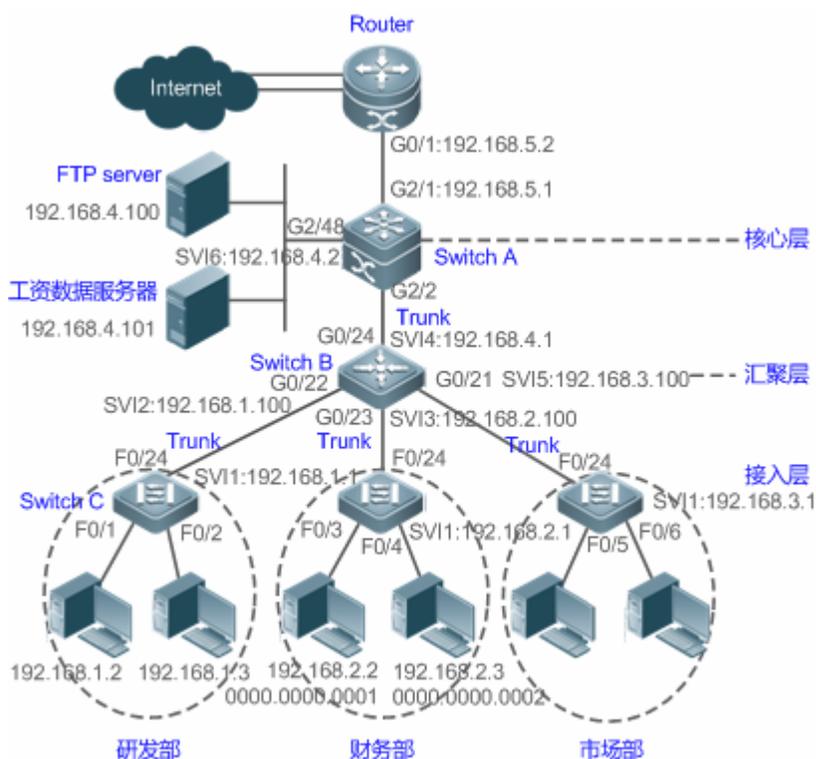
1.2.1 企业内网络访问控制应用

应用场景

Internet 病毒无处不在, 需要封堵各种病毒的常用端口, 以保障内网安全:

- 只允许内部 PC 访问服务器, 不允许外部 PC 访问服务器。
- 不允许非财务部门 PC 访问财务部 PC; 不允许非研发部门 PC 访问研发部 PC。
- 不允许研发部门人员在上班时间 (即 9:00~18:00) 使用 QQ、MSN 等聊天工具。

图 1-1



- 【注释】 接入层设备 C：连接各部门的 PC，通过千兆光纤(trunk 方式)连接汇聚层设备。
 汇聚层设备 B：划分多个 VLAN，每个部门为一个 VLAN，通过万兆光纤(trunk 方式)上连核心层设备。
 核心层设备 A：连接各种服务器，如 FTP，HTTP 服务器等，通过防火墙与 Internet 相连。

功能部属

- 通过在核心层设备（本例为设备 A）上联 Router 的端口（本例为 G2/1 口）上设置扩展 ACL 来过滤相关端口的数据包来达到防病毒的目的。
- 要求内部 PC 对服务器进行访问，不允许外部 PC 访问服务器，可以通过定义 IP 扩展 ACL 并应用到核心层设备（本例为设备 A）的下联汇聚层设备和服务器的接口（本例为 G2/2 口/SVI 2）上实现。
- 要求特定部门间不能互访，可通过定义 IP 扩展 ACL 实现（本例中分别在设备 B 的 G0/22、G0/23 上应用 IP 扩展 ACL）；
- 可通过配置时间 IP 扩展 ACL，限制研发部门在特定时间内使用 QQ/MSN 等聊天工具（本例中在设备 B 的 SVI 2 上应用时间 IP 扩展 ACL）。

1.3 功能详解

基本概念

访问列表

访问列表有：基本访问列表和动态访问列表。

用户可以根据需要选择基本访问列表或动态访问列表。一般情况下，使用基本访问列表已经能够满足安全需要。但经验丰富的黑客可能会通过一些软件假冒源地址欺骗设备，得以访问网络。而动态访问列表在用户访问网络以前，要求通过身份认证，使黑客难以攻入网络，所以在一些敏感的区域可以使用动态访问列表保证网络安全。

-
- i** 通过假冒源地址欺骗设备即电子欺骗是所有访问列表固有的问题，使用动态列表也会遭遇电子欺骗问题：黑客可能在用户通过身份认证的有效访问期间，假冒用户的地址访问网络。解决这个问题的方法有两种，一种是尽量将用户访问的空闲时间设置小些，这样可以使黑客更难以攻入网络，另一种是使用 IPSEC 加密协议对网络数据进行加密，确保进入设备时，所有的数据都是加密的。
-

访问列表一般配置在以下位置的网络设备上：

- 内部网和外部网（如 INTERNET）之间的设备
- 网络两个部分交界的设备
- 接入控制端口的设备。

访问控制列表语句的执行必须严格按照表中语句的顺序，从第一条语句开始比较，一旦一个数据包的报头跟表中的某个条件判断语句相匹配，那么后面的语句就将被忽略，不再进行检查。

📄 输入/输出 ACL、过滤域模板及规则

输入 ACL 在设备接口接收到报文时，检查报文是否与该接口输入 ACL 的某一条 ACE 相匹配；输出 ACL 在设备准备从某一个接口输出报文时，检查报文是否与该接口输出 ACL 的某一条 ACE 相匹配。

在制定不同的过滤规则时，多条规则可能同时被应用，也可能只应用其中几条。只要是符合某条 ACE，就按照该 ACE 定义的处理报文(Permit 或 Deny)。ACL 的 ACE 根据以太网报文的某些字段来标识以太网报文的，这些字段包括：

二层字段(Layer 2 Fields)：

- 48 位的源 MAC 地址(必须申明所有 48 位)
- 48 位的目的 MAC 地址(必须申明所有 48 位)
- 16 位的二层类型字段

三层字段(Layer 3 Fields)：

- 源 IP 地址字段(可以申明全部源 IP 地址值，或使用子网来定义一类流)
- 目的 IP 地址字段(可以申明全部目的 IP 地址值，或使用子网来定义一类流)
- 协议类型字段

四层字段(Layer 4 Fields)：

- 可以申明一个 TCP 的源端口、目的端口或者都申明，还可以申明源端口或目的端口的范围。
- 可以申明一个 UDP 的源端口、目的端口或者都申明，还可以申明源端口或目的端口的范围。

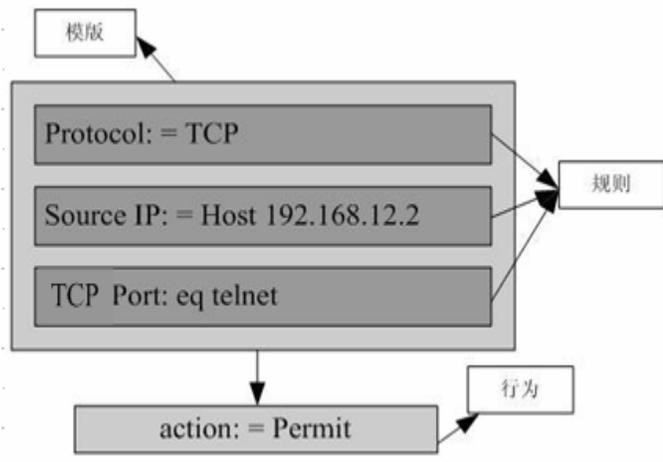
过滤域指的是，在生成一条 ACE 时，根据报文中的哪些字段用以对报文进行识别、分类。过滤域模板就是这些字段组合的定义。比如，在生成某一条 ACE 时希望根据报文的源 IP 字段对报文进行识别、分类，而在生成另一条 ACE 时，希望根据的是报文的源 IP 地址字段和 UDP 的源端口字段，这样，这两条 ACE 就使用了不同的过滤域模板。

规则(Rules)，指的是 ACE 过滤域模板对应的值。比如有一条 ACE 内容如下：

```
permit tcp host 192.168.12.2 any eq telnet
```

在这条 ACE 中，过滤域模板为以下字段的集合：源 IP 地址字段、IP 协议字段、目的 TCP 端口字段。对应的值(Rules)分别为：源 IP 地址 = Host 192.168.12.2；IP 协议 = TCP；TCP 目的端口 = Telnet。

图 1-2 对 ACE : permit tcp host 192.168.12.2 any eq telnet 的分析



- ❗ 过滤域模板可以是三层字段(Layer 3 Field)和四层字段(Layer 4 Field)字段的集合，也可以是多个二层字段(Layer 2 Field)的集合，但标准与扩展的 ACL 的过滤域模板不能是二层和三层、二层和四层、二层和三层、四层字段的集合。要使用二层、三层、四层字段集合，可以应用 Expert 扩展访问控制列表 (Expert ACLs)。
- ❗ OUT 方向 ACL 关联 SVI 的注意事项：支持 IP 标准，IP 扩展，MAC 扩展，专家级 ACL 应用。
- ❗ 当配置专家级的 ACL，并应用在接口的 out 方向时，如果该 ACL 中的某些 ACE 包含三层匹配信息（比如 IP，L4port 等），将导致从应用接口进入的非 IP 报文无法受该 ACL 的 permit 和 deny 规则控制。
- ❗ 应用 ACL 时，如果 ACL（包括 IP 访问列表和 Expert 扩展访问列表）中的 ACE 匹配了非 L2 字段，比如 SIP，DIP 时，对于带标签的 MPLS 报文匹配是无效的。

ACL logging

为了让用户更好的掌握 ACL 在设备中的运行状态，在添加规则时可以根据需要决定是否指定报文匹配日志输出选项，如果指定了该选项，则在规则匹配到报文时会输出匹配日志信息。ACL logging 信息是基于 ACE 来打印 log 信息的，也即设备周期性的打印命中报文的 ACE 信息，以及该 ACE 命中的报文数量。如下：

```
*Sep 9 16:23:06: %ACL-6-MATCH: ACL 100 ACE 10 permit icmp any any, match 78 packets.
```

为合理控制 log 输出的数量和频率，ACL logging 支持配置 log 输出间隔的配置，并且支持分别配置 IPv4 ACL 的 log 输出间隔。

- ❗ 带 ACL logging 选项的 ACE 使用更多的硬件资源，如果配置的所有 ACE 都带有 ACL logging 选项，则会导致设备的 ACE 容量减半。

- i** 默认 ACL logging 的 log 输出间隔是 0, 也即不输出 log。在为 ACE 指定了报文匹配日志输出选项后, 若要输出相应的 log, 需要配置 ACL logging 的输出间隔。
- i** 对于带 ACL logging 选项的 ACE, 如果指定的时间间隔内没有匹配到任何报文, 则不会输出与该 ACE 有关的报文匹配日志; 如果指定的时间间隔内有匹配到报文, 则时间间隔到期后, 会输出与该 ACE 有关的报文匹配日志。其中的报文中数目为该时间间隔内该 ACE 匹配到的报文总数, 即为该 ACE 上一次输出 log 到本次输出 log 之间命中的报文数。
- i** 仅在交换机设备上支持 ACL logging 功能。

ACL 报文匹配计数

由于网络管理的需要, 有时用户可能会想知道某条 ACE 有没有匹配到报文, 匹配了多少个。因此, ACL 提供了基于 ACE 的报文匹配计数, 用户可以基于 ACL 开启和关闭该 ACL 下的所有的 ACE 的报文匹配计数功能, 支持的 ACL 类型包括: IP 访问列表、MAC 访问列表、Expert 访问列表。此外, 用户可通过 ACL 的统计清除命令将 ACL 规则的报文匹配计数器清 0, 以便重新统计。

- !** 开启 ACL 的报文匹配计数功能需要更多的硬件表项, 极端情况下会使设备可以配置的 ACE 容量减半。
- i** 仅在交换机设备上支持 ACL 报文匹配计数功能。

功能特性

功能特性	作用
IP访问列表	可以根据 IPv4 报文头部的三层或四层信息对进出设备的 IPv4 报文进行控制。
MAC扩展访问列表	可以根据以太网报文的二层头部信息对进出设备的二层报文进行控制。
Expert扩展访问列表	IP 访问列表和 MAC 扩展访问列表的组合, 从而实现在同一条规则中可以实现同时根据报文的二层头部信息和报文三层或四层信息对进出设备的报文进行控制, 以决定是丢弃还是放过指定的报文。
ACL80	可以自定义匹配域和掩码, 适应固定匹配域不能满足需求的场景
ACL重定向	可以将进入设备的符合 ACL 规则的报文直接重定向到指定的出接口
全局安全ACL	可以让 ACL 在所有接口的入方向上生效, 无需在每个接口上分别应用 ACL
安全通道	可以让报文不经过 dot1x、web 认证等接入控制的检查, 以满足特定场景的需求
SVI Router ACL	可以同一 VLAN 内的用户可以正常通信
报文匹配日志	可以根据需要每隔一段时间输出 ACL 的报文匹配日志信息, 根据信息可以了解到指定规则的报文匹配情况
分片报文匹配模式	使用分片报文匹配模式可以使得访问列表对分片报文进行更精细化的控制

1.3.1 IP访问列表

IP 访问列表主要用于对进出设备的 IPv4 报文进行精细化控制, 用户可以根据实际需要阻止或允许特定的 IPv4 报文进入网络, 从而实现控制 IP 用户访问网络资源的目的。

工作原理

在 IP 访问列表中定义一系列的 IP 访问规则，然后将访问列表应用在接口的入方向或出方向上，当然还可以对 IP 访问列表进行全局应用，IPv4 报文进出设备时，设备就会通过判断报文是否与规则匹配来决定是否转发或阻断报文。

要在设备上配置访问列表，必须为协议的访问列表指定一个唯一的名称或编号，以便在协议内部能够唯一标识每个访问列表。下表列出了可以使用编号来指定访问列表的协议以及每种协议可以使用的访问列表编号范围。

协议	编号范围
标准 IP	1-99, 1300 - 1999
扩展 IP	100-199, 2000 - 2699

基本访问列表包括标准 IP 访问列表和扩展 IP 访问列表，访问列表中定义的典型规则主要包含以下匹配域：

- 源 IP 地址
- 目的 IP 地址
- IP 协议号
- 四层源端口号或 ICMP type
- 四层目的端口号或 ICMP code

标准 IP 访问列表（编号为 1 - 99, 1300 - 1999）主要是根据源 IP 地址来进行转发或阻断分组的，扩展 IP 访问列表（编号为 100 - 199, 2000 - 2699）可以对上述匹配域进行组合来控制报文的转发或阻断。

对于单一的访问列表来说，可以使用多条独立的访问列表语句来定义多种规则，其中所有的语句引用同一个编号或名字，以便将这些语句绑定到同一个访问列表。不过，使用的语句越多，阅读和理解访问列表就越来越困难。

- ✔ 路由类产品上，ACL 规则中的 ICMP code 匹配域对于 ICMP type 为 3 的 ICMP 报文无效。如果 ACL 规则中配置了要匹配 ICMP 报文的 code 字段，当 type 为 3 的 ICMP 报文进入设备执行 ACL 匹配时，匹配结果可能与预期的不一样。

📌 隐含“拒绝所有数据流”规则语句

在每个 IP 访问列表的末尾隐含着一行“拒绝所有数据流”规则语句，因此如果分组与任何规则都不匹配，将被拒绝。

如下例：

```
access-list 1 permit host 192.168.4.12
```

此列表只允许源主机为 192.168.4.12 的报文通过，其它主机都将被拒绝。因为这条访问列表最后包含了一条规则语句：
access-list 1 deny any。

又如：

```
access-list 1 deny host 192.168.4.12
```

如果列表只包含以上这一条语句，则任何主机报文通过该端口时都将被拒绝。

- ❗ 在定义访问列表的时候，要考虑到路由更新的报文。由于访问列表末尾“拒绝所有数据流”，可能导致所有的路由更新报文被阻断。

📌 输入规则语句的顺序

加入的每条规则都被追加到访问列表的最后（但在默认规则语句之前），访问列表规则语句的输入次序非常重要，它决定了该规则语句在访问列表中的优先级，设备在决定转发还是阻断报文时，是按规则语句创建的次序将进行比较的，找到匹配的规则语句后，便不再检查其他规则语句。

假设创建了一条规则语句，它允许所有的数据流通过，则后面的语句将不被检查。

如下例：

```
access-list 101 deny ip any any
access-list 101 permit tcp 192.168.12.0 0.0.0.255 eq telnet any
```

由于第一条规则语句拒绝了所有的 IP 报文，所以 192.168.12.0/24 网络的主机 Telnet 报文将被拒绝，因为设备在检查到报文和第一条规则语句匹配，便不再检查后面的规则语句。

相关配置

配置 IP 访问列表

缺省情况下，设备上无任何 IP 访问列表。

在配置模式下使用 **ip access-list { standard | extended } {acl-name | acl-id}** 命令可以创建一个标准 IP 访问列表或扩展 IP 访问列表，并进入标准或扩展 IP 访问列表模式。

配置 IP 访问列表匹配规则

缺省情况下，创建的 IP 访问列表中会有一条隐含的 deny 所有 IPv4 报文的匹配规则，这条表项对用户不可见，但当将访问列表应用在接口上时，就会生效，也就是会丢弃所有 IPv4 报文，因此，如果用户想允许某些特定的 IPv4 报文进出设备，就得往访问列表中配置一些匹配规则。

对于标准 IP 访问列表，可以通过以下方式配置匹配规则：

- 不管是命名的标准 IP 访问列表，还是数值索引的标准 IP 访问列表，都可以在标准 IP 访问列表模式下使用 **[sn] { permit | deny } { host source | any | source source-wildcard } [time-range time-range-name]** 命令为访问列表配置一条匹配规则。
- 数值索引的标准 IP 访问列表，除了可以在标准 IP 访问列表模式下使用前面提到的命令配置匹配规则外，还可以在配置模式下使用 **access-list acl-id { permit | deny } { host source | any | source source-wildcard } [time-range tm-rng-name]** 命令为标准 IP 访问列表配置一条匹配规则。

对于扩展 IP 访问列表，可以通过以下方式配置匹配规则：

- 不管是命名的扩展 IP 访问列表，还是数值索引的扩展 IP 访问列表，都可以在扩展 IP 访问列表模式下使用 **[sn] { permit | deny } protocol { host source | any | source source-wildcard } { host destination | any | destination destination-wildcard } [[precedence precedence [tos tos]] | dscp dscp] [fragment] [time-range time-range-name]** 命令为访问列表配置一条匹配规则。
- 数值索引的扩展 IP 访问列表，除了可以在扩展 IP 访问列表模式下使用前面提到的命令配置匹配规则外，还可以在配置模式下使用 **access-list acl-id { permit | deny } protocol { host source | any | source source-wildcard } { host destination | any | destination destination-wildcard } [[precedence precedence [tos tos]] | dscp dscp] [fragment] [time-range time-range-name]** 命令为标准 IP 访问列表配置一条匹配规则。

应用 IP 访问列表

缺省情况下，设备上的所有接口都没有应用 IP 访问列表，也就是说 IP 访问列表不会对进出设备的 IP 报文进行匹配过滤。

在接口模式下使用 `ip access-group { acl-id | acl-name } { in | out }` 命令可以让一个标准 IP 访问列表或扩展 IP 访问列表在指定的接口上生效。

1.3.2 MAC 扩展访问列表

MAC 扩展访问列表主要是基于报文的二层头部来对进出设备的报文进行精细化控制，用户可以根据实际需要阻止或允许特定的二层报文进入网络，从而实现控制保护网络资源不受攻击或者基于些控制用户访问网络资源的目的。

工作原理

在 MAC 扩展访问列表中定义一系列的 MAC 访问规则，将访问列表应用在接口的入方向或出方向上，报文进出设备时，设备就会通过判断报文是否与规则匹配来决定是否转发或阻断报文。

要在设备上配置 MAC 扩展访问列表，必须给访问列表指定一个唯一的名称或编号，以便唯一标识每个访问列表。下表列出可以使用编号来指定 MAC 扩展访问列表编号范围。

协议	编号范围
MAC 扩展访问列表	700-799

MAC 扩展访问列表中定义的典型规则主要有以下：

- 源 MAC 地址
- 目标 MAC 地址
- 以太网协议类型

从上面的规则字段可以看出，MAC 扩展访问列表（编号 700 -799）主要是根据源或目的 MAC 地址以及报文的以太网类型来匹配报文分组的。

对于单一的 MAC 扩展访问列表来说，可以使用多条独立的访问列表语句来定义多种规则，其中所有的语句引用同一个编号或名字，以便将这些语句绑定到同一个访问列表。不过，使用的语句越多，阅读和理解访问列表就越来越困难。

隐含“拒绝所有数据流”规则语句

在每个 MAC 扩展访问列表的末尾隐含着一条“拒绝所有数据流”规则语句，因此如果分组与任何规则都不匹配，将被拒绝。

如下例：

```
access-list 700 permit host 00d0.f800.0001 any
```

此列表只允许来自 MAC 地址为 00d0.f800.0001 的主机发出的报文通过，来自其它主机都将被拒绝。因为这条访问列表最后包含了一条规则语句：`access-list 700 deny any any`。

相关配置

配置 MAC 扩展访问列表

缺省情况下，设备上无任何 MAC 扩展访问列表。

在配置模式下使用 **mac access-list extended** {acl-name | acl-id} 命令可以创建一个 MAC 扩展访问列表，并进入 MAC 扩展访问列表模式。

配置 MAC 扩展访问列表匹配规则

缺省情况下，创建的 MAC 扩展访问列表中会有一条隐含的 deny 所有二层报文的匹配规则，这条表项对用户不可见，但当将访问列表应用在接口上时，就会生效，也就是会丢弃所有二层报文，因此，如果用户想允许某些特定的二层报文进出设备，就得往访问列表中配置一些匹配规则。

可以通过以下方式配置匹配规则：

- 不管是命名的 MAC 扩展访问列表，还是数值索引的 MAC 扩展访问列表，都可以在 MAC 扩展访问列表模式下使用[sn] { permit | deny } {any | host src-mac-addr} {any | host dst-mac-addr} [ethernet-type] [cos cos] [inner cos] [time-range tm-rng-name]命令为访问列表配置一条匹配规则。
- 数值索引的 MAC 扩展访问列表，除了可以在 MAC 扩展访问列表模式下使用前面提到的命令配置匹配规则外，还可以在配置模式下使用 **access-list** acl-id { permit | deny } {any | host src-mac-addr} {any | host dst-mac-addr} [ethernet-type] [cos cos] [inner cos] [time-range time-range-name]命令为 MAC 扩展访问列表配置一条匹配规则。

应用 IP 访问列表

缺省情况下，设备上的所有接口都没有应用 MAC 扩展访问列表，也就是说创建的 MAC 扩展访问列表不会对进出设备的二层报文进行匹配过滤。

在接口模式下使用 **mac access-group** { acl-id | acl-name } { in | out } 命令可以让一个 MAC 扩展访问列表在指定的接口上生效。

1.3.3 Expert扩展访问列表

如果用户想在同一条规则中既对报文的二层信息匹配，又对报文的三层信息进行匹配，那么就可以选择 Expert 扩展访问列表。可以将 Expert 扩展访问列表看作是 IP 访问列表和 MAC 扩展访问列表的一种结合与增强，之所以说是一种结合与增强，是因为 Expert 扩展访问列表中的规则不仅可以包含 IP 访问列表规则和 MAC 扩展访问列表规则，同时可以指定基于 VLAN ID 来匹配报文。

工作原理

在 Expert 扩展访问列表中定义一系列的访问规则，将访问列表应用在接口的入方向或出方向上，报文进出设备时，设备就会通过判断报文是否与访问规则匹配来决定是否转发或阻断报文。

要在设备上配置 Expert 扩展访问列表，必须给协议的访问列表指定一个唯一的名称或编号，以便在协议内部能够唯一标识每个访问列表。下表列出 Expert 访问列表的编号范围。

协议	编号范围
Expert 扩展访问列表	2700-2899

创建 expert 扩展访问列表时,定义的规则可以应用于所有的分组报文,通过判断分组是否与规则匹配来决定是否转发或阻断分组报文。

Expert 访问列表中定义的典型规则主要有以下：

- 基本访问列表和 MAC 扩展访问列表所有的信息
- VLAN ID

Expert 扩展访问列表 (编号 2700 -2899) 为基本访问列表和 MAC 扩展访问列表的综合体,并且能对 VLAN ID 进行过滤。

对于单一的 Expert 扩展访问列表来说,可以使用多条独立的访问列表语句来定义多种规则,其中所有的语句需引用同一个编隐含“拒绝所有数据流”规则语句

在每个 Expert 扩展访问列表的末尾隐含着一行“拒绝所有数据流”规则语句,因此如果分组与任何规则都不匹配,将被拒绝。

如下例：

```
access-list 2700 permit 0x0806 any any any any any
```

此列表只允许以太网类型为 0x0806(即 ARP)的报文通过,其他类型的报文都将被拒绝。因为这条访问列表最后包含了一条规则语句: access-list 2700 deny any any any any。

相关配置

配置 Expert 扩展访问列表

缺省情况下,设备上无任何 Expert 扩展访问列表。

在配置模式下使用 **expert access-list extended** {acl-name | acl-id} 命令可以创建一个 Expert 扩展访问列表,并进入 Expert 扩展访问列表模式。

配置 Expert 扩展访问列表匹配规则

缺省情况下,创建的 Expert 扩展访问列表中会有一条隐含的 deny 所有报文的匹配规则,这条表项对用户不可见,但当将访问列表应用在接口上时,就会生效,也就是会丢弃所有二层报文,因此,如果用户想允许某些特定的二层报文进出设备,就得往访问列表中配置一些匹配规则。

可以通过以下方式配置匹配规则：

- 不管是命名的 Expert 扩展访问列表,还是数值索引的 Expert 扩展访问列表,都可以在 Expert 扩展访问列表模式下使用[sn] { **permit** | **deny** } [protocol | [ethernet-type] [cos [out] [inner in]]] [[VID [out] [inner in]]] { source source-wildcard | host source | any } { host source-mac-address | any } { destination destination-wildcard | host destination | any } { host destination-mac-address | any } [precedence precedence] [tos tos] [fragment] [range lower upper] [time-range time-range-name] 命令为访问列表配置一条匹配规则。
- 数值索引的 MAC 扩展访问列表,除了可以在 MAC 扩展访问列表模式下使用前面提到的命令配置匹配规则外,还可以在配置模式下使用 **access-list** acl-id { **permit** | **deny** } [protocol | [ethernet-type] [cos [out] [inner in]]] [[VID [out] [inner in]]] { source source-wildcard | host source | any } { host source-mac-address | any } { destination destination-wildcard | host destination | any } { host destination-mac-address | any } [precedence precedence] [tos tos] [fragment] [range lower upper] [time-range time-range-name] 命令为 Expert 扩展访问列表配置一条匹配规则。

应用 Expert 扩展访问列表

缺省情况下，设备上的所有接口都没有应用 Expert 扩展访问列表，也就是说创建的 Expert 扩展访问列表不会对进出设备的所有二三层报文进行匹配过滤。

在接口模式下使用 `expert access-group { acl-id | acl-name } { in | out }` 命令可以让一个 Expert 扩展访问列表在指定的接口上生效。

1.3.4 ACL80

ACL80 即 Expert 高级访问列表，同时也叫自定义访问列表，针对报文的前 80 个字节进行匹配过滤。报文的 SMAC/DMAC/SIP/DIP/ETYPE 不计算在任意指定的字段中，ACL80 在匹配报文的以上这些字段之后，还能再匹配额外指定的 16 个字节内容。

工作原理

报文是由一系列的字节流组成，ACL80 可以让用户对报文的前 80 个字节中的指定 16 个字节按比特 (bit) 位进行匹配过滤。对于任意一个 16 字节字段，可以按照 bit 形式与所设置的值进行比较或不比较。也就是说，它允许我们对这 16 个字节的任意一个比特设置该值为 0 或 1。在对任何一个字节进行过滤时，有三个要素：匹配域内容、匹配域掩码以及匹配的起始位置。匹配域内容和匹配域掩码二者的比特位是一一对应的。过滤规则指明需要过滤字段值，过滤域模板指明过滤规则中对应字段是否需要过滤（1 表示匹配对应过滤规则的比特位，0 表示不匹配），所以当需要匹配某个比特时，必须将过滤域模板中对应的比特为设置为 1。如果过滤域模板比特位设置为 0，无论过滤规则中对应的比特位是什么，都不会匹配。

例如：

```
Ruijie(config)# expert access-list advanced name
Ruijie(config-exp-dacl)# permit 00d0f8123456 ffffffff 0
Ruijie(config-exp-dacl)# deny 00d0f8654321 ffffffff 6
```

用户自定义访问控制列表根据用户的定义对二层数据帧的前 80 个字节中的任意字节进行匹配，对数据报文作出相应的处理。正确使用用户自定义访问控制列表需要用户对二层数据帧的构成有深入的了解。下表为二层数据帧的前 64 个字节的示意图(每个字母代表一个 16 进制数，每两个字母代表一个字节)。

```
AA AA AA AA AA AA BB BB BB BB BB BB CC CC DD DD
DD DD EE FF GG HH HH HH II II JJ KK LL LL MM MM
NN NN OO PP QQ QQ RR RR RR RR SS SS SS SS TT TT
UU UU VV VV VV VV WW WW WW WW XY ZZ aa aa bb bb
```

在上图中，各个字母的含义及偏移量取值如下表所示：

字母	含义	偏移量	字母	含义	偏移量
A	目的 MAC	0	O	TTL 字段	34
B	源 MAC	6	P	协议号	35
C	VLAN tag 字段	12	Q	IP 校验和	36
D	数据帧长度字段	16	R	源 ip 地址	38
E	DSAP(目的服务访问点)字段	18	S	目的 ip 地址	42

F	SSAP(源服务访问点)字段	19	T	TCP 源端口	46
G	Ctrl 字段	20	U	TCP 目的端口	48
H	Org Code 字段	21	V	序列号	50
I	封装的数据类型	24	W	确认字段	54
J	IP 版本号	26	XY	IP 头长度和保留比特位	58
K	TOS 字段	27	Z	保留比特位和 flags 比特位	59
L	IP 包的长度	28	a	Windows size 字段	60
M	ID 号	30	b	其他	62
N	Flags 字段	32			

上表中各个字段的偏移量是它们在 SNAP + tag 的 802.3 数据帧中的偏移量。在用户自定义访问控制列表中，用户可以使用规则掩码和偏移量两个参数共同从数据帧中提取前 80 个字节中的任意字节，然后和用户定义的规则比较，从而过滤出匹配的数据帧，作相应的处理。用户定义的规则可以是数据的某些固定属性，比如用户要将所有的 TCP 报文过滤出来，可以将规则定义为“06”，规则掩码定义为“FF”，偏移量定义为 35，此时规则掩码和偏移量共同作用，把接收到的数据帧中的 TCP 协议号字段的内容提取出来，和规则比较，匹配出所有的 TCP 报文。

- i 仅在交换机设备上支持 ACL80;
- i ACL80 可以支持匹配以太网报文，803.3snap 报文，802.3llc 报文，如果设置匹配 DSAP 到 cntl 字段的值为 AAAA03，则表示希望匹配 803.3snap 报文，如果设置匹配 DSAP 到 cntl 字段的值为 E0E003，则表示希望匹配 803.3llc 报文。以太网报文不能设置匹配该字段;
- i ACL80 可以任意匹配的资源只有 16 个字节，如果这 16 个字节的资源已经被使用，那么就无法再匹配这 16 个字节之外的字段。

相关配置

配置 Expert 高级访问列表

缺省情况下，设备上无任何 Expert 高级访问列表。

在配置模式下使用 **expert access-list advanced *acl-name*** 命令可以创建一个 Expert 高级访问列表，并进入 Expert 高级访问列表模式。

配置 Expert 高级访问列表匹配规则

缺省情况下，创建的 Expert 高级访问列表中会有一条隐含的 deny 所有报文的匹配规则，这条表项对用户不可见，但当将访问列表应用在接口上时，就会生效，也就是会丢弃所有二层报文，因此，如果用户想允许某些特定的二层报文进出设备，就得往访问列表中配置一些匹配规则。

可以在 Expert 高级访问列表模式下使用 **[sn] { permit | deny } *hex hex-mask offset*** 命令为访问列表配置一条匹配规则。

应用 Expert 高级访问列表

缺省情况下，设备上的所有接口都没有应用 Expert 高级访问列表，也就是说创建的 Expert 高级访问列表不会对进出设备的所有报文进行匹配过滤。

在接口模式下使用 **expert access-group *acl-name* { in | out }** 命令可以让一个 Expert 高级访问列表在指定的接口上生效。

1.3.5 ACL重定向

ACL 重定向功能的作用是使得设备能够对收到的报文进行分析并且重定向到指定端口转发出去。当要分析进入设备的特定报文时，可以配置 ACL 重定向功能，把符合规则报文重定向到指定端口上，在这个端口上把报文抓下来加以分析。

工作原理

ACL 重定向通过在一个接口上绑定不同的 ACL 策略，并给每个策略指定一个输出目的口，当该接口收到报文时，将逐条查找绑定在该接口上的 ACL 策略，如果报文符合某条策略描述的特征，将从该策略所指定的目的口转发，从而达到基于流来重定向报文的效果

- ① 仅在交换机设备上支持 ACL 重定向功能；
- ② ACL 重定向功能仅在接口入方向生效。

相关配置

配置访问列表

在配置 ACL 重定向功能之前，必须先配置访问列表，访问列表的配置说明请参考相关的章节。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置 ACL 重定向

缺省情况下，设备上无任何的 ACL 重定向配置。

在接口模式下使用 **redirect destination interface** *interface-name* **acl** {*acl-id* | *acl-name*} **in** 命令配置 ACL 重定向功能。

- ❗ 只支持在以大口、聚合口、SVI 上配置 ACL 重定向功能。

1.3.6 全局安全ACL

由于安全部署上的需要，端口安全 ACL 常被配置作为病毒报文过滤及防范使用，用于过滤符合某些特征的报文，比如：TCP 攻击端口。各种病毒报文在全局网络环境中存在，且各端口下的病毒报文识别特征相同或相似，因此通常情况下会创建一个 ACL，添加匹配各种病毒特征的 deny ace 后，通过端口安全 ACL 将 ACL 应用到交换机各个端口，以达到病毒过滤的作用。

端口安全 ACL 用于病毒过滤等抗攻击场景时，存在较多不便，一是需要逐个端口配置，存在重复配置、操作性能低下及 ACL 资源过度消耗的情况；二是安全 ACL 的访问控制作用被弱化，由于被用于病毒过滤，安全 ACL 的限制路由更新、限制网络访问等基本功能无法正常使用。而全局安全 ACL 可以在不影响端口安全 ACL 的情况下，进行全局抗病毒部署及防御。全局安全 ACL 只需要一条命令就可以在所有二层接口上生效，而不需要象端口安全 ACL 那样需要在每个接口上进行重复配置。

工作原理

全局安全 ACL 在所有二层接口上生效，当全局安全 ACL 与端口安全 ACL 同时配置时，两者共同生效，对于命中全局安全 ACL 的报文将被当作病毒报文直接过滤，对于没有命中全局安全 ACL 的报文将继续受端口安全 ACL 控制；如果想让某些端口不受全局安全 ACL 的控制，可以在这些接口上独立关闭全局安全 ACL 检查功能。

- i 由于全局安全 ACL 主要用于病毒过滤，因此被关联于全局安全 ACL 的 ACL 中，只有 deny 类型的 ACE 会安装生效，permit 类型的 ACE 不会生效；
- i 与端口上应用的安全 ACL 不同，全局安全 ACL 没有默认的 deny 所有表项，即没命中规则的报文都是放过的；
- i 全局 ACL 可以在二层口上生效，也可以路由口上也生效。即可以在以下类型的端口上都生效：access、trunk、hibird、路由口、ap 口(二层或三层)。在 SVI 口上不生效。
- i 允许在物理端口和 AP 口上独立关闭全局安全 ACL 功能，不支持在 AP 成员口上关闭；
- i 全局安全 ACL 只支持关联 IP 标准 ACL、IP 扩展 ACL；
- i 全局安全 ACL 目前仅支持在接口的入方向上生效。

相关配置

配置访问列表

在配置全局安全 ACL 功能之前，必须先配置访问列表，访问列表的配置说明请参考相关的章节。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表的相关章节说明。

配置全局安全 ACL

缺省情况下，设备上无任何的全局安全 ACL 配置。

在配置模式下使用 `ip access-group acl-id { in | out }` 命令开启全局安全 ACL 功能。

配置全局安全 ACL 例外口

缺省情况下，设备上无任何的全局安全 ACL 例外口配置。

在接口模式下使用 `no global ip access-group` 命令关闭指定接口上全局安全 ACL 功能。

1.3.7 安全通道

在某些应用场景中，可能会需要保证符合某些特征的报文绕过接入控制应用的检查，比如 dot1x 认证前，要允许用户登录到指定的资源站点上下载 dot1x 认证客户端；使用安全通道可以达到这个目的。将安全 ACL 通过安全通道配置命令应用到全局或者接口，就表示该 ACL 是一条安全通道

工作原理

安全通道其实也是一个访问控制列表，可以基于全局或者接口配置。报文进入到接口时，首先进行安全通道的检查，如果满足安全通道的匹配条件，将绕过接入控制比如端口安全，web 认证、dot1x，Ip+MAC 绑定的检查直接进入交换机。应用于全局的安全通道对所有非例外口都生效。

- i 应用于安全通道的访问控制列表的 deny 行为不生效，并且不存在末尾隐含着一条“拒绝所有数据流”规则的语句，如果报文不符合安全通道的匹配条件，将按流程进行接入控制的检查；
- i 全局安全通道的例外口最多可以设置 8 个且全局安全通道的例外口不能用来设置基于接口的安全通道。
- i 如果接口上应用了安全通道，并且还在全局的安全通道，那么全局安全通道不在这个接口上生效。
- i 基于端口可迁移认证模式和安全通道共用时，安全通道不生效。
- i 仅在交换机设备上支持安全通道。

相关配置

配置访问列表

在配置安全通道功能之前，必须先配置访问列表，访问列表的配置说明请参考相关的章节。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表、MAC 扩展访问列表以及 Expert 扩展访问列表的相关章节说明。

配置接口安全通道

缺省情况下，设备上无任何的接口安全通道配置。

在接口模式下使用 **security access-group {acl-id | acl-name}** 命令配置接口安全通道。

配置全局安全通道

缺省情况下，设备上无任何的全局安全通道配置。

在配置模式下使用 **security global access-group {acl-id | acl-name}** 命令配置全局安全通道。

配置全局安全通道例外口

缺省情况下，设备上无任何的全局安全通道例外口配置。

在接口模式下使用 **security uplink enable** 命令将指定接口配置为全局安全通道例外口。

1.3.8 SVI Router ACL

默认情况下，应用在 SVI 接口上的访问列表会同时对 VLAN 内二层转发的报文及 VLAN 间的路由报文生效，从而导致同一 VLAN 内不同用户之间无法正常通信等异常现象。为此，提供了一种切换手段，可以使得应用在 SVI 接口上的访问列表仅对 VLAN 间的路由报文生效。

工作原理

缺省情况下，SVI Router ACL 功能默认关闭，SVI ACL 同时对 VLAN 间的三层转发报文及 VLAN 内的桥转发报文生效。SVI Router ACL 功能开启后，SVI ACL 仅对 VLAN 间的三层转发报文生效。

✔ 仅在交换机设备上支持 SVI Router ACL。

相关配置

配置访问列表

在配置带 SVI Router ACL 之前，一般来说要先配置访问列表应用，在应用访问列表前，必须先配置访问列表，访问列表的配置说明请参考相关的章节。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

应用访问列表

访问列表的应用配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。应用时，在 SVI 对应的接口模式应用。

配置 SVI Router ACL

全局模式使用 `svi router-acls enable` 命令开启 SVI Router ACL 功能，使得应用在 SVI 接口上的访问列表仅对三层转发的报文生效，而不对同一 VLAN 内二层转发的报文生效。

1.3.9 报文匹配日志

报文匹配日志主要用于监控访问列表规则的运行状态，为日常网络维护以及网络优化提供必要的信息。

工作原理

为了让用户更好的掌握 ACL 在设备中的运行状态，在添加规则时可以根据需要决定是否指定报文匹配日志输出选项，如果指定了该选项，则在规则匹配到报文时会输出匹配日志信息。ACL logging 信息是基于 ACE 来打印 log 信息的，也即设备周期性的打印命中报文的 ACE 信息，以及该 ACE 命中的报文数量。如下：

```
*Sep 9 16:23:06: %ACL-6-MATCH: ACL 100 ACE 10 permit icmp any any, match 78 packets.
```

为合理控制 log 输出的数量和频率，ACL logging 支持配置 log 输出间隔的配置。

❗ 带 log 选项的访问列表规则会使用更多的硬件资源，如果配置的所有规则都带有 log 选项，则会导致设备的硬件策略容量减半。

⚠ 默认报文匹配日志输出间隔是 0，即不输出上层。在配置访问列表规则时指定了 log 选项后，还需要配置输出间隔，否则不会输出匹配日志。

⚠ 对于带 `log` 选项的规则，如果指定的时间间隔内没有匹配到任何报文，则不会输出与该规则有关的报文匹配日志；如果指定的时间间隔内有匹配到报文，则时间间隔到期后，会输出与该规则有关的报文匹配日志。其中的报文命中数目为该时间间隔内该规则匹配到的报文总数，即为该规则上一次输出日志到本次输出日志之间命中的报文数。

✔ 仅在交换机设备上支持报文匹配日志功能。

相关配置

配置访问列表

在配置带 `log` 选项的访问列表规则之前，必须先配置访问列表，访问列表的配置说明请参考相关的章节。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表的相关章节说明。注意要配置 `log` 选项。

配置报文匹配日志输出间隔

在配置模式下使用 `{ip} access-list log-update interval time` 命令配置报文匹配日志输出间隔。

应用访问列表

访问列表规则的应用方法请参考 IP 访问列表的相关章节说明。

1.3.10 报文匹配计数

除了报文匹配日志外，报文匹配计数为日常的网络维护和网络优化提供了另一种选择。

工作原理

出于网络管理的需要，用户可能会想知道某条访问列表规则有没有匹配到报文，匹配了多少个。因此，ACL 提供了基于规则的报文匹配计数功能，用户可以基于 ACL 开启和关闭该 ACL 下的所有规则的报文匹配计数功能，当有报文匹配到了这条规则，对应的匹配计数就相应地增长。用户可通过 ACL 的统计清除命令将该 ACL 下所有规则的报文匹配计数器清 0，以便重新统计。

⚠ 开启 ACL 的报文匹配计数功能需要更多的硬件表项，极端情况下会使设备可以配置的硬件策略容量减半。

✔ 支持在 IP 访问列表、MAC 访问列表、Expert 扩展访问列表上开启报文匹配计数功能。

✔ 仅在交换机设备上支持 ACL 报文匹配计数功能。

相关配置

配置访问列表

在配置带 `log` 选项的访问列表规则之前，必须先配置访问列表，访问列表的配置说明请参考相关的章节。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表的相关章节说明。注意要配置 log 选项。

📌 开启报文匹配计数

如果用户想在 IP 访问列表、MAC 扩展访问列表或者 Expert 扩展访问列表上开始报文匹配计数功能,请在配置模式下使用{**mac | expert | ip**} **access-list counter** { *acl-id* | *acl-name* }命令来开启;

📌 应用访问列表

访问列表规则的应用方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

📌 清除报文匹配计数

在特权模式下使用 **clear counters access-list** [*acl-id* | *acl-name*]命令来清除。

1.3.11 分片报文匹配模式

使用分片报文匹配模式可以使得访问列表对分片报文进行更精细化的控制。

工作原理

对于 IP 报文,在网络传输时中可能会被分片。报文发生分片时,只有首片报文带有四层信息,比如 TCP 或 UDP 端口号、ICMP 类型和 ICMP 编码等,其他的分片报文都不带有这些四层信息。默认情况下,如果 ACL 规则带有 fragment 标识,则只会去匹配非首片报文;如果 ACL 规则不带有 fragment 标识,则匹配所有报文,包括首片报文和后续的所有分片报文。除了这种默认的分片报文匹配模式外,还提供了另一种新的分片报文匹配方法,用户可以根据需要在指定的 ACL 上进行切换。新的分片报文匹配模式与默认的分片报文匹配模式的区别就在于:当访问列表规则中不带有 fragment 标识时,如果报文被分片,首片报文会去匹配规则中用户定义的所有匹配域(包括三层和四层信息),而非首片报文则只会去匹配规则中的非四层信息。

- ❗ 分片报文新匹配模式下,如果 ACL 规则不带 fragment 标识,且匹配动作是 permit,这样的 ACL 规则需要占用更多的硬件表项资源,极端情况下会使得硬件策略表项容量减半;如果这样的 ACE 配置了 TCP flag 过滤控制的 established,则还会占用更多的硬件策略表项。
- ❗ 执行分片报文匹配模式切换时,会导致 ACL 的短时失效。
- ✅ 分片报文新匹配模式下,如果 ACL 规则不带 fragment 标识并且需要匹配报文的四层信息时,当匹配动作为 permit 时,ACL 规则会检查首片报文三层和四层信息,对于非首片报文只会检查报文的三层信息;当匹配动作为 deny 时,ACL 规则只会检查首片报文,不会去检查分片报文。
- ✅ 分片报文新匹配模式下,如果 ACL 规则带有 fragment 标识,不论 ACL 规则的匹配动作是 permit 还是 deny,都只检查非首片报文,而不会去检查首片报文。
- ✅ 仅在 IP 扩展 ACL 和 Expert 扩展 ACL 上支持分片报文匹配模式的切换。
- ✅ 仅在交换机设备上支持分片报文的匹配。

相关配置

📌 配置访问列表

访问列表的配置说明请参考 IP 访问列表和 Expert 扩展访问列表的相关章节说明。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表和 Expert 扩展访问列表的相关章节说明。配置时需要注意添加 fragment 选项。

切换分片报文匹配模式

在配置模式下使用 `[no] { ip | expert } access-list new-fragment-mode { acl-id | acl-name }` 命令进行分片报文匹配模式的切换。

应用访问列表

访问列表规则的应用方法请参考 IP 访问列表以及 Expert 扩展访问列表的相关章节说明。

1.4 产品说明



S7508 产品，作用在 IN 方向的 acl 不支持 TCP、UDP 报文 4 层端口的“neq”匹配，作用在 OUT 方向的 ACL 仅支持 TCP、UDP 报文 4 层端口的“eq”匹配。



S7508 产品，作用于 SVI 上的安全 ACL 同时对于 VLAN 内的桥转发报文及 VLAN 间的路由报文生效，从而导致 VLAN 内用户无法通信等异常现象。



S7508 产品，仅在 IP 扩展 ACL 和 Expert 扩展 ACL 上支持分片报文匹配模式的切换。



S7508 产品，ACL80 只支持目的 MAC、源 MAC、VID、ETYPE、IP 协议号、ipv4 源 IP、IPV4 目的 IP、目的端口号，源端口号、ICMP TYPE、ICMP CODE 这几个常规字段。



S7508 产品，开启 ACL 的报文匹配计数功能需要更多的硬件表项，极端情况下会使设备可以配置的 ACE 容量减半。

1.5 配置详解

配置项	配置建议 & 相关命令
配置IP访问列表功能	 可选配置。用于匹配过滤 IPv4 报文。
	<code>ip access-list standard</code> 配置 IP 标准访问列表
	<code>ip access-list extended</code> 配置 IP 标准访问列表
	<code>permit host any time-range log</code> 配置 permit 类型的 IP 标准访问列表规则
	<code>deny host any time-range log</code> 配置 deny 类型的 IP 标准访问列表规则

	permit host any host any tos dscp precedence fragment time-range log	配置 permit 类型的 IP 扩展访问列表规则
	deny host any host any tos dscp precedence fragment time-range log	配置 deny 类型的 IP 扩展访问列表规则
	ip access-group in out	应用 IP 标准或 IP 扩展访问列表
配置 MAC 扩展访问列表	 可选配置。用于匹配过滤二层报文	
	mac access-list extended	配置 MAC 扩展访问列表
	permit any host any host cos inner time-range	配置 permit 类型的 MAC 扩展访问列表规则
	deny any host any host cos inner time-range	配置 deny 类型的 MAC 扩展访问列表规则
	mac access-group in out	应用 MAC 扩展访问列表
配置 Expert 扩展访问列表	 可选配置。用于匹配过滤二三层报文	
	expert access-list extened	配置 Expert 扩展访问列表
	permit cos inner VID inner host any host any host any host any precedence tos fragment range time-range	配置 permit 类型的 Expert 扩展访问列表规则
	deny cos inner VID inner host any host any host any host any precedence tos fragment range time-range	配置 deny 类型的 Expert 扩展访问列表规则
	expert access-group in out	应用 Expert 扩展访问列表
配置 ACL80	 可选配置。用于自定义匹配域过滤二三层报文	
	expert access-list advanced	配置 Expert 高级访问列表
	permit	配置 permit 类型的 Expert 高级访问列表规则
	deny	配置 deny 类型的 Expert 高级访问列表规则
	expert access-group in out	应用 Expert 高级访问列表
配置 ACL 重定向	 可选配置。用于指定符合规则的报文重定向到指向的接口上	
	redirect destination interface acl in	配置 ACL 重定向
配置全局安全 ACL	 可选配置。用于让访问列表在全局上生效	
	ip access-group in out	在全局模式下配置全局安全 ACL 重定向
	no global ip access-group	在接口模式下将该接口配置为全局安全 ACL 的例外口。
配置安全通道	 可选配置。用于符合规则的报文直接跳过接入控制各种应用（比如 dot1x, web 认证）的检查。	
	security access-group	在接口模式下开启安全通道功能
	security global access-group	在配置模式下开启安全通道功能

	security uplink enable	在接口模式下将该接口配置成全局安全通道的例外口
配置访问列表注释信息	 可选配置。用于为访问列表或访问列表规则配置注释信息便于用户识别。	
	list-remark	在访问列表模式下为访问列表配置注释信息
	access-list list-remark	在全局模式为访问列表配置注释信息。
	remark	在访问列表模式下为规则配置注释信息

1.5.1 配置IP访问列表

配置效果

通过配置 IP 访问列表，并将访问列表应用到设备的接口上，就可以对在该接口进出的所有 IPv4 报文进行控制，禁止或允许特定的 IPv4 报文进入网络，从而实现控制 IP 用户访问网络资源的目的。

注意事项

无

配置方法

配置 IP 访问列表

- 必须配置。要实际针对 IPv4 用户访问网络资源的控制，首先必须配置 IP 访问列表。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。IP 访问列表只对被配置的设备上有效，不会影响网络中的其他设备。

配置 IP 访问列表规则

- 可选配置。访问列表里允许无规则，没有配置规则时，默认禁止所有 IPv4 报文进入设备。

应用 IP 访问列表

- 必须配置。要使得 IP 访问列表真正生效，就必须将 IP 访问列表应用到设备的特定接口上。
- 可以根据用户的分布，在接入、汇聚或核心设备的指定接口上应用 IP 访问列表。

检验方法

可以通过以下方法检验 IP 访问列表的配置效果：

- 通过 ping 的方式检查 IP 访问列表是否真的在指定接口上生效。比如，IP 访问列表里配置了禁止某个 IP 主机或某个 IP 范围的主机不允许访问网络，通过 ping 的方式检验是否真的 ping 不通来验证。

- 通过访问网络相关资源的方式来检验 IP 访问列表是否真的在指定接口上生效,比如访问 internet 网,或通过 ftp 访问网络上的 ftp 资源等。

相关命令

配置 IP 访问列表

【命令格式】 **ip access-list { standard | extended } {acl-name | acl-id}**

【参数说明】 **standard**: 该选项若被配置,表示要创建一个标准 IP 访问列表。

extended: 该选项若被配置,表示要创建一个扩展 IP 访问列表。

acl-name: 该选项若被配置,表示创建一个命名的标准 IP 或扩展 IP 访问列表,长度范围[1, 99]。访问列表名称不能以数字 0 - 9 开头,也不能为 “in” 或 “out”。

acl-id: 为访问列表编号,以此来唯一标识一条访问列表,该选项若被配置,表示创建一个数值索引的标准 IP 或扩展 IP 访问列表,如果创建的是标准 IP 访问列表,**acl-id**的取值范围为 1-99, 1300 – 1999,如果创建的是扩展 IP 访问列表,**acl-id**的取值范围为 100-199, 2000 – 2699。

【命令模式】 配置模式

【使用指导】 此命令可以用来配置标准 IP 或扩展 IP 访问列表,并进入标准 IP 或扩展 IP 访问列表配置模式。如果只想通过检查报文的源 IP 地址来控制用户的网络资源访问权限,那么可以配置标准 IP 访问列表;如果想通过检查报文的源 IP 地址、目的 IP 地址、报文的协议号、TCP/UDP 源或目的端口号来控制用户的网络资源访问权限,那么就需要配置扩展 IP 访问列表。

配置 IP 访问列表规则

- 为标准 IP 访问列表配置规则。

有两种方式可以为标准 IP 访问列表配置规则：

【命令格式】 **[sn] { permit | deny } { host source | any | source source-wildcard } [time-range time-range-name] [log]**

【参数说明】 **sn**: 为规则表项的序号,取值范围为[1, 2147483647]。这个序号决定了这条规则表项在该访问列表中的优先级,序号越小,优先级越大,优先级大的会优先去匹配报文。如果配置匹配规则时没有指定序号,系统会自动分配一个序号,序号的分配原则为在当前访问列表最后一条匹配规则的序列基础之上加上一个递增值,递增值默认为 10,假设当前访问列表最后一条匹配规则的序号为 100,则缺省情况下新增的这条匹配规则序号就为 11,此外,递增值是可以通过命令调整的。

permit: 该选项若被配置,表示本规则属于允许通过类的;

deny: 该选项若被配置,关键字表示本规则属于禁止通过类的;

host source: 该选项若被配置,表示要匹配源 IP 为某一台主机发出的 IP 报文;

any: 该选项若被配置,表示要匹配任意主机发出的 IP 报文;

source source-wildcard: 该选项若被配置,表示要匹配某一个 IP 网段的内主机发出的报文;

time-range time-range-name: 该选项若被配置,表示该匹配规则关联了一个时间区,只有在指定的时间区间内该规则才会生效,否则不生效,更多关于时间区的描述,请参考 time range 的配置手册

log: 该选项若被配置,表示本规则如果匹配到了报文需要定时输出匹配日志,有关匹配日志更详细描述请参考本手册的 ACL logging 一节。

【命令模式】 标准 IP 访问列表模式

【使用指导】 此命令在标准 IP 访问列表模式下为访问列表配置规则,该访问列表可以是命名访问列表,也可以是数字索引

的访问列表。

【命令格式】 `access-list acl-id { permit | deny } {host source | any | source source-wildcard } [time-range tm-rng-name] [log]`

【参数说明】 *acl-id*: 数值索引访问列表的编号，以此来唯一标识一条访问列表。取值范围为：1-99，1300 - 1999

permit: 该选项若被配置，表示本规则属于允许通过类的；

deny: 该选项若被配置，关键字表示本规则属于禁止通过类的；

host source: 该选项若被配置，表示要匹配源 IP 为某一台主机发出的 IP 报文；

any: 该选项若被配置，表示要匹配任意主机发出的 IP 报文；

source source-wildcard: 该选项若被配置，表示要匹配某一个 IP 网段的内主机发出的报文；

time-range time-range-name: 该选项若被配置，表示该匹配规则关联了一个时间区，只有在指定的时间区间内该规则才会生效，否则不生效，更多关于时间区的描述，请参考 time range 的配置手册

log: 该选项若被配置，表示本规则如果匹配到了报文需要定时输出匹配日志，有关匹配日志更详细描述请参考本手册的 ACL logging 一节。

【命令模式】 标准 IP 访问列表模式

【使用指导】 此命令在配置模式下为数字索引的 IP 访问列表配置规则。这种配置方式无法为命名的标准 IP 访问列表配置规则。

- 为扩展 IP 访问列表配置规则。

有两种方式可以为扩展 IP 访问列表配置规则：

【命令格式】 `[sn] { permit | deny } protocol {host source | any | source source-wildcard } {host destination | any | destination destination-wildcard } [[precedence precedence [tos tos]] | dscp dscp] [fragment] [time-range time-range-name] [log]`

【参数说明】 *sn*: 规则表项的序号，取值范围为[1, 2147483647]。这个序号决定了这条规则表项在该访问列表中的优先级，序号越小，优先级越大，优先级大的会优先去匹配报文。如果配置匹配规则时没有指定序号，系统会自动分配一个序号，序号的分配原则为在当前访问列表最后一条匹配规则的序列基础之上加上一个递增值，递增值默认为 10，假设当前访问列表最后一条匹配规则的序号为 100，则缺省情况下新增的这条匹配规则序号就为 11，此外，递增值是可以通过命令调整的

permit: 该选项若被配置，表示本规则属于允许通过类的；

deny: 该选项若被配置，关键字表示本规则属于禁止通过类的；

protocol: IP 协议号，取值范围[0, 255]；为方便使用，系统提供了常用 IP 协议号的简称以取代对应的 IP 协议号具体数值，包括 eigrp、gre、icmp、igmp、ip、ipinip、nos、ospf、tcp、udp。

host source: 该选项若被配置，表示要匹配源 IP 为某一台主机发出的 IP 报文；

source source-wildcard: 该选项若被配置，表示要匹配某一个 IP 网段的内主机发出的报文；

host destination: 该选项若被配置，表示要匹配目的 IP 为某一台特定主机的 IP 报文；**any** 关键字表示要匹配发往任意主机的 IP 报文。

destination destination-wildcard: 该选项若被配置，表示要匹配目标为某一个 IP 网段主机的报文。

any: 该选项若被配置，表示要匹配任意主机发出的 IP 报文或者要匹配发往任意主机的 IP 报文；

precedence precedence: 该选项若被配置，表示要匹配 IP 报文头部中的优先级域。

tos tos: 该选项若被配置，表示要匹配 IP 报文头部中的服务类型域。

dscp dscp: 该选项若被配置，表示要匹配 IP 报文头部的 dscp 域。

fragment: 该选项若被配置，表示只要匹配非首片的 IP 分片报文。

time-range time-range-name: 该选项若被配置，表示该匹配规则关联了一个时间区，只有在指定的时间区间内该规则才会生效，否则不生效，更多关于关时间区的描述，请参考 time range 的配置手册

log: 该选项若被配置，表示本规则如果匹配到了报文需要定时输出匹配日志，有关匹配日志更详细描述请参考本手册的 ACL logging 一节。

【命令模式】 扩展 IP 访问列表模式

【使用指导】 此命令在扩展 IP 访问列表模式下为访问列表配置规则，该访问列表可以是命名访问列表，也可以是数字索引的访问列表。

【命令格式】 **access-list acl-id { permit | deny } protocol { host source | any | source source-wildcard } { host destination | any | destination destination-wildcard } [[precedence precedence [tos tos]] | dscp dscp] [fragment] [time-range time-range-name] [log]**

【参数说明】 **acl-id:** 数值索引访问列表的编号，以此来唯一标识一条访问列表。取值范围为：100-199，2000 - 2699
sn: 规则表项的序号，取值范围为[1, 2147483647]。这个序号决定了这条规则表项在该访问列表中的优先级，序号越小，优先级越大，优先级大的会优先去匹配报文。如果配置匹配规则时没有指定序号，系统会自动分配一个序号，序号的分配原则为在当前访问列表最后一条匹配规则的序列基础之上加上一个递增值，递增值默认为 10，假设当前访问列表最后一条匹配规则的序号为 100，则缺省情况下新增的这条匹配规则序号就为 11，此外，递增值是可以通过命令调整的

permit: 该选项若被配置，表示本规则属于允许通过类的；

deny: 该选项若被配置，关键字表示本规则属于禁止通过类的；

protocol: IP 协议号，取值范围[0, 255]；为方便使用，系统提供了常用 IP 协议号的简称以取代对应的 IP 协议号具体数值，包括 eigrp、gre、icmp、igmp、ip、ipinip、nos、ospf、tcp、udp。

host source: 该选项若被配置，表示要匹配源 IP 为某一台主机发出的 IP 报文；

source source-wildcard: 该选项若被配置，表示要匹配某一个 IP 网段的内主机发出的报文；

host destination: 该选项若被配置，表示要匹配目的 IP 为某一台特定主机的 IP 报文；**any** 关键字表示要匹配发往任意主机的 IP 报文。

destination destination-wildcard: 该选项若被配置，表示要匹配目标为某一个 IP 网段主机的报文。

any: 该选项若被配置，表示要匹配任意主机发出的 IP 报文或者要匹配发往任意主机的 IP 报文；

precedence precedence: 该选项若被配置，表示要匹配 IP 报文头部中的优先级域。

tos tos: 该选项若被配置，表示要匹配 IP 报文头部中的服务类型域。

dscp dscp: 该选项若被配置，表示要匹配 IP 报文头部的 dscp 域。

fragment: 该选项若被配置，表示只要匹配非首片的 IP 分片报文。

time-range time-range-name: 该选项若被配置，表示该匹配规则关联了一个时间区，只有在指定的时间区间内该规则才会生效，否则不生效，更多关于关时间区的描述，请参考 time range 的配置手册

log: 该选项若被配置，表示本规则如果匹配到了报文需要定时输出匹配日志，有关匹配日志更详细描述请参考本手册的 ACL logging 一节。

【命令模式】 扩展 IP 访问列表模式

【使用指导】 此命令在配置模式下为数字索引的 IP 访问列表配置规则。这种配置方式无法为命名的标准 IP 访问列表配置规

则。

应用 IP 访问列表

【命令格式】 `ip access-group { acl-id | acl-name } { in | out }`

【参数说明】 `acl-id`: 该选项若被配置, 表示要将一个数值索引的标准 IP 或扩展 IP 访问列表应用在接口上。

`acl-name`: 该选项若被配置, 表示要将一个命名的标准 IP 或扩展 IP 访问列表应用在接口上。

`in`: 该选项若被配置, 表示这个访问列表对进入该接口的 IP 报文进行控制。

`out`: 该选项若被配置, 表示这个访问列表对从该接口发出的 IP 报文进行控制。

【命令模式】 接口模式

【使用指导】 此命令可以让 IP 访问列表在指定的接口上生效, 同时需要指定对进入设备的报文生效, 还是从设备转发出去的报文生效。

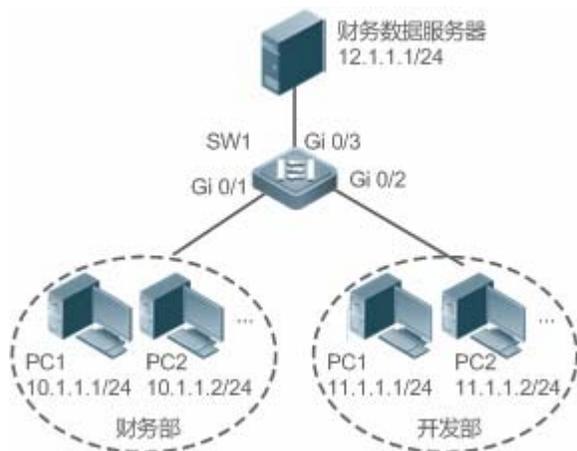
配置举例

i 以下配置举例, 仅介绍与 ACL 相关的配置。

通过 IP 访问列表, 禁止财务部以外的部门访问财务数据服务器

【网络环境】

图 1-3



- 【配置方法】**
- 配置 IP 访问列表
 - 在 IP 访问列表中添加访问规则
 - 将 IP 访问列表应用在连接财务数据服务器接口的出方向上

SW1

```
sw1(config)#ip access-list standard 1
sw1(config-std-nacl)#permit 10.1.1.0 0.0.0.255
sw1(config-std-nacl)#deny 11.1.1.1 0.0.0.255
sw1(config-std-nacl)#exit
sw1(config)#int gigabitEthernet 0/3
sw1(config-if-GigabitEthernet 0/3)#ip access-group 1 out
```

- 【检验方法】**
- 从开发部的某台 PC 机上 ping 财务数据服务器, 确认 ping 不通。

- 从财务部的某台 PC 机上 ping 财务数据服务器，确认 ping 得通

SW1

```
swl(config)#show access-lists

ip access-list standard 1
 10 permit 10.1.1.0 0.0.0.255
 20 deny 11.1.1.0 0.0.0.255

swl(config)#show access-group
ip access-group 1 out
Applied On interface GigabitEthernet 0/3
```

1.5.2 配置MAC扩展访问列表

配置效果

通过配置 MAC 扩展访问列表，并将访问列表应用到设备的接口上，就可以对在该接口进出的所有二层报文进行控制，禁止或允许特定的二层报文进入网络，从而实现基于二层报文头来控制用户访问网络资源的目的。

注意事项

无

配置方法

配置 MAC 扩展访问列表

- 必须配置。要基于二层报文头信息（比如用户 PC 的 MAC 地址）控制用户访问网络资源的权限，首先必须配置 MAC 扩展访问列表。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。MAC 扩展访问列表只在被配置的设备上有效，不会影响网络中的其他设备。

配置 MAC 扩展访问列表规则

- 可选配置。访问列表里允许无规则，没有配置规则时，默认禁止所有以太网二层报文进入设备。

应用 MAC 扩展访问列表

- 必须配置。要使得 MAC 扩展访问列表真正生效，就必须将 MAC 扩展访问列表应用到设备的特定接口上。
- 可以根据用户的分布，在接入、汇聚或核心设备的指定接口上应用 MAC 扩展访问列表。

检验方法

可以通过以下方法检验 MAC 扩展访问列表的配置效果：

- 如果 MAC 扩展访问列表希望放过或过滤某些 IP 报文，可以通过 ping 的方式检查这样的 MAC 扩展访问列表规则是否真的在指定接口上生效。比如，MAC 扩展访问列表里配置了禁止以太网类型为 0x0800 即 IP 报文从接口进入设备，可以通过 ping 的方式检验是否真的 ping 不通来验证。
- 如果 MAC 扩展访问列表希望放过或过滤某些非 IP 报文，比如 ARP 报文，这种报文也可以通过 ping 的方式检查这样的 MAC 扩展访问列表规则是否真的在指定接口上生效，比如想过滤掉 ARP 报文，可以通过 ping 的方式检验是否真的 ping 不通来验证。
- 另外，还可以通过构造符合指定特征的二层报文来检验 MAC 扩展访问列表是否真的生效。典型地可以使用两台 PC 机，一台构造二层报文并发送，另一台开启抓包软件抓包，根据访问列表规则指定的动作检查报文的转发是否如预期（转发或不转发）。

相关命令

配置 MAC 扩展访问列表

【命令格式】 **mac access-list extended** {acl-name | acl-id }

【参数说明】 **acl-name**: 该选项若被配置，表示创建一个命名的 MAC 扩展访问列表，长度范围[1, 99]。访问列表名称不能以数字 0 - 9 开头，也不能为 “in” 或 “out”。

acl-id: 为访问列表编号，以此来唯一标识一条访问列表，该选项若被配置，表示创建一个数值索引的 MAC 扩展访问列表，取值范围为 700-799。

【命令模式】 配置模式

【使用指导】 此命令可以用来配置 MAC 扩展访问列表，并进入 MAC 扩展访问列表配置模式。如果想通过检查以太网报文的二层信息来控制用户的网络资源访问权限，那么就可以配置 MAC 扩展访问列表。

配置 MAC 扩展访问列表规则

有两种方法为 MAC 扩展访问列表配置规则：

- 在 MAC 扩展访问列表模式中配置规则

【命令格式】 [sn] { **permit** | **deny** } {**any** | **host** src-mac-addr} {**any** | **host** dst-mac-addr} [ethernet-type] [**cos** cos [inner cos]] [**time-range** tm-rng-name]

【参数说明】 **sn**: 为规则表项的序号，取值范围为[1, 2147483647]。这个序号决定了这条规则表项在该访问列表中的优先级，序号越小，优先级越大，优先级大的会优先去匹配报文。如果配置匹配规则时没有指定序号，系统会自动分配一个序号，序号的分配原则为在当前访问列表最后一条匹配规则的序列基础之上加上一个递增值，递增值默认为 10，假设当前访问列表最后一条匹配规则的序号为 100，则缺省情况下新增的这条匹配规则序号就为 11，此外，递增值是可以通过命令调整的。

permit: 该选项若被配置，表示本规则属于允许通过类的；

deny: 该选项若被配置，关键字表示本规则属于禁止通过类的；

any: 该选项若被配置，表示要匹配任意主机发出的二层报文；

host src-mac-addr: 该选项若被配置，表示要匹配源 MAC 为某一台主机发出的二层报文；

any: 该选项若被配置，表示要匹配目的为任意主机发出的二层报文；

host dst-mac-addr: 该选项若被配置，表示要匹配目的 MAC 为某一台主机的二层报文；

ethernet-type: 该选项若被配置, 表示要匹配指定以太网类型的二层报文;

cos cos: 该选项若被配置, 表示要匹配二层报文里的外层 TAG 的优先级字段;

inner cos: 该选项若被配置, 表示要匹配二层报文里的内层 TAG 的优先级字段;

time-range time-range-name: 该选项若被配置, 表示该匹配规则关联了一个时间区, 只有在指定的时间区间内该规则才会生效, 否则不生效, 更多关于关时间区的描述, 请参考 time range 的配置手册

【命令模式】 MAC 扩展访问列表模式

【使用指导】 此命令在 MAC 扩展访问列表模式下为访问列表配置规则, 该访问列表可以是命名访问列表, 也可以是数字索引的访问列表。

● 在全局模式中为 MAC 扩展访问列表配置规则

【命令格式】 **access-list acl-id { permit | deny } { any | host src-mac-addr } { any | host dst-mac-addr } [ethernet-type] [cos cos [inner cos]] [time-range tm-rng-name]**

【参数说明】 *acl-id*: 数值索引访问列表的编号, 以此来唯一标识一条访问列表。取值范围为: 700-799

permit: 该选项若被配置, 表示本规则属于允许通过类的;

deny: 该选项若被配置, 关键字表示本规则属于禁止通过类的;

host src-mac-addr: 该选项若被配置, 表示要匹配源 MAC 为某一台主机发出的二层报文;

host source: 该选项若被配置, 表示要匹配源 MAC 为某一台主机发出的二层报文;

any: 该选项若被配置, 表示要匹配目的为任意主机发出的二层报文;

host dst-mac-addr: 该选项若被配置, 表示要匹配目的 MAC 为某一台主机的二层报文;

ethernet-type: 该选项若被配置, 表示要匹配指定以太网类型的二层报文;

cos cos: 该选项若被配置, 表示要匹配二层报文里的外层优先级字段;

inner cos: 该选项若被配置, 表示要匹配二层报文里的内层优先级字段;

time-range time-range-name: 该选项若被配置, 表示该匹配规则关联了一个时间区, 只有在指定的时间区间内该规则才会生效, 否则不生效, 更多关于关时间区的描述, 请参考 time range 的配置手册。

【命令模式】 全局模式

【使用指导】 此命令在配置模式下为数字索引的 MAC 扩展访问列表配置规则。这种配置方式无法为命名的 MAC 扩展访问列表配置规则。

▾ 应用 MAC 扩展访问列表

【命令格式】 **mac access-group { acl-id | acl-name } { in | out }**

【参数说明】 *acl-id*: 该选项若被配置, 表示要将一个数值索引的 MAC 扩展访问列表应用在接口上。

acl-name: 该选项若被配置, 表示要将一个命名的 MAC 扩展访问列表应用在接口上。

in: 该选项若被配置, 表示这个访问列表对进入该接口的二层报文进行控制。

out: 该选项若被配置, 表示这个访问列表对从该接口发出的二层报文进行控制。

【命令模式】 接口模式

【使用指导】 此命令可以让 MAC 扩展访问列表在指定的接口上生效, 同时需要指定对进入设备的报文生效, 还是从设备转发出去的报文生效。

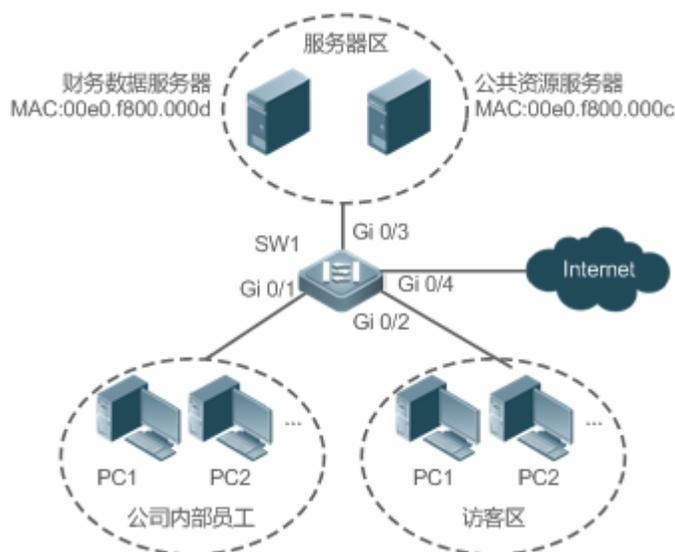
配置举例

i 以下配置举例，仅介绍与 ACL 相关的配置。

通过 MAC 扩展访问列表，限制来访客户可访问的资源

【网络环境】

图 1-4



【配置方法】

- 配置 MAC 扩展访问列表
- 在 MAC 扩展访问列表中添加访问规则
- 将 MAC 扩展访问列表应用在连接访客区接口的出方向上，允许访客 PC 访问 Internet 以及公司内部的公共服务器，但不允许访问公司的账务数据服务器，即禁止访问 MAC 地址为 00e0.f800.000d 的服务器。

SW1

```
sw1(config)#mac access-list extended 700
sw1(config-mac-nacl)#deny any host 00e0.f800.000d
sw1(config-mac-nacl)#permit any any
sw1(config-mac-nacl)#exit
sw1(config)#int gigabitEthernet 0/2
sw1(config-if-GigabitEthernet 0/2)#mac access-group 700 in
```

【检验方法】

- 从访客 PC 机上 ping 财务数据服务器，确认 ping 不通。
- 从访问 PC 机上 ping 公共资源服务器，确认可以 ping 得通。
- 在访问 PC 机上访问 Internet，比如访问百度，确认可以打开主页。

SW1

```
sw1(config)#show access-lists
mac access-list extended 700
 10 deny any host 00e0.f800.000d etype-any
 20 permit any any etype-any
sw1(config)#show access-group
mac access-group 700 in
Applied On interface GigabitEthernet 0/2
```

1.5.3 配置Expert扩展访问列表

配置效果

通过配置 Expert 扩展访问列表，并将访问列表应用到设备的接口上，可以同时基于二层和三层信息对在该接口进出的报文进行控制，禁止或允许特定的报文进入网络；另外，还可以通过配置 Expert 扩展访问列表实现基于 VLAN 来对所有二层报文进行控制，从而实现允许或拒绝某些网段的用户访问网络资源。一般来说，如果想在一条访问列表中混合使用 IP 访问规则以及 MAC 扩展访问规则时，就可以使用 Expert 扩展访问列表

注意事项

无

配置方法

配置 Expert 扩展访问列表

- 必须配置。要基于二层报文头信息（比如 VLAN ID）控制用户访问网络资源的权限，首先必须配置 Expert 扩展访问列表。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。Expert 扩展访问列表只在被配置的设备上有效，不会影响网络中的其他设备。

配置 Expert 扩展访问列表规则

- 可选配置。访问列表里允许无规则，没有配置规则时，默认禁止所有报文进入设备。

应用 Expert 扩展访问列表

- 必须配置。要使得 Expert 扩展访问列表真正生效，就必须将访问列表应用到设备的特定接口上。
- 可以根据用户的分布，在接入、汇聚或核心设备的指定接口的入或出方向上应用 Expert 扩展访问列表。

检验方法

可以通过以下方法检验 Expert 扩展访问列表的配置效果：

- 如果 Expert 扩展访问列表中配置了 IP 访问规则，放过或过滤某些 IP 报文，通过 ping 的方式来检验规则是否生效。
- 如果 Expert 扩展访问列表中配置了 MAC 访问规则，放过或过滤某些二层报文，比如 ARP 报文，这种报文也可以通过 ping 的方式检查这样的 MAC 访问列表规则是否真的在指定接口上生效，比如想过滤掉 ARP 报文，可以通过 ping 的方式检验是否真的 ping 不通来验证。
- 如果 Expert 扩展访问列表中配置了带有 VLAN ID 的访问规则，希望放过或过滤某些二层网段的报文，典型假设不想让 VLAN 1 的用户与 VLAN 2 的用户互访问，可以在 VLAN 1 所在的 PC 机上 ping VLAN 2 的 PC 机，如果 ping 不通就表示规则生效。

相关命令

配置 Expert 扩展访问列表

【命令格式】 **expert access-list extended** {acl-name | acl-id}

【参数说明】 **acl-name**: 该选项若被配置, 表示创建一个命名的 Expert 扩展访问列表, 长度范围[1, 99]。访问列表名称不能以数字 0 - 9 开头, 也不能为 “in” 或 “out” 。

acl-id: 为访问列表编号, 以此来唯一标识一条访问列表, 该选项若被配置, 表示创建一个数值索引的 Expert 扩展访问列表, 取值范围为 2700-2899。

【命令模式】 配置模式

【使用指导】 此命令可以用来配置 MAC 扩展访问列表, 并进入 Expert 扩展访问列表配置模式。

配置 Expert 扩展访问列表规则

有两种方法为 Expert 扩展访问列表配置规则：

- 在 Expert 扩展访问列表模式中配置规则

【命令格式】 [sn] { **permit** | **deny** } [protocol | [ethernet-type] [cos [out] [inner in]]] [[VID [out] [inner in]]] { source source-wildcard | host source | any } { host source-mac-address | any } { destination destination-wildcard | host destination | any } { host destination-mac-address | any } [precedence precedence] [tos tos] [fragment] [range lower upper] [time-range time-range-name]

【参数说明】 **sn**: 为规则表项的序号, 取值范围为[1, 2147483647]。这个序号决定了这条规则表项在该访问列表中的优先级, 序号越小, 优先级越大, 优先级大的会优先去匹配报文。如果配置匹配规则时没有指定序号, 系统会自动分配一个序号, 序号的分配原则为在当前访问列表最后一条匹配规则的序列基础之上加上一个递增值, 递增值默认为 10, 假设当前访问列表最后一条匹配规则的序号为 100, 则缺省情况下新增的这条匹配规则序号就为 11, 此外, 递增值是可以通过命令调整的。

permit: 该选项若被配置, 表示本规则属于允许通过类的;

deny: 该选项若被配置, 关键字表示本规则属于禁止通过类的;

protocol: IP 协议号, 取值范围[0, 255]; 为方便使用, 系统提供了常用 IP 协议号的简称以取代对应的 IP 协议号具体数值, 包括 eigrp、gre、icmp、igmp、ip、ipinip、nos、ospf、tcp、udp

ethernet-type: 该选项若被配置, 表示要匹配指定以太网类型的二层报文;

cos out: 该选项若被配置, 表示要匹指定二层报文外层 TAG 中的优先级字段;

cos inner in: 该选项若被配置, 表示要匹指定二层报文内层 TAG 中的优先级字段;

VID out: 该选项若被配置, 表示要匹指定二层报文外层 TAG 中的 VLAN ID 字段;

VID inner in: 该选项若被配置, 表示要匹指定二层报文内层 TAG 中的 VLAN ID 字段;

source source-wildcard: 该选项若被配置, 表示要匹配某一个 IP 网段的内主机发出的报文;

host source: 该选项若被配置, 表示要匹配源 IP 为某一台主机发出的 IP 报文;

any: 该选项若被配置, 表示要匹配任意主机发出的 IP 报文;

host source-mac-address: 该选项若被配置, 表示要匹配源 MAC 为某一台主机发出的二层报文;

any: 该选项若被配置, 表示要匹配目的为任意主机发出的二层报文;

destination destination-wildcard: 该选项若被配置, 表示要匹配目标为某一个 IP 网段的报文;

host destination: 该选项若被配置, 表示要匹配目的 IP 为某一台主机的 IP 报文;

any: 该选项若被配置, 表示要匹配发往任意目标的 IP 报文;

host destination-mac-address: 该选项若被配置, 表示要匹配目的 MAC 为某一台主机的二层报文;

any: 该选项若被配置, 表示要匹配目标为任意主机的二层报文;

precedence precedence: 该选项若被配置, 表示要匹配 IP 报文头部中的优先级域。

tos tos: 该选项若被配置, 表示要匹配 IP 报文头部中的服务类型域。

dscp dscp: 该选项若被配置, 表示要匹配 IP 报文头部的 dscp 域。

fragment: 该选项若被配置, 表示只要匹配非首片的 IP 分片报文。

time-range time-range-name: 该选项若被配置, 表示该匹配规则关联了一个时间区, 只有在指定的时间区间内该规则才会生效, 否则不生效, 更多关于关时间区的描述, 请参考 time range 的配置手册

【命令模式】 Expert 扩展访问列表模式

【使用指导】 此命令在 Expert 扩展访问列表模式下为访问列表配置规则, 该访问列表可以是命名访问列表, 也可以是数字索引的访问列表。

● 在全局模式下为 Expert 扩展访问列表配置规则

【命令格式】 **access-list acl-id { permit | deny } [protocol] [ethernet-type] [cos [out] [inner in]] [[VID [out] [inner in]]] { source source-wildcard | host source | any } { host source-mac-address | any } { destination destination-wildcard | host destination | any } { host destination-mac-address | any } [precedence precedence] [tos tos] [fragment] [range lower upper] [time-range time-range-name]**

【参数说明】 **acl-id**: 数值索引访问列表的编号, 以此来唯一标识一条访问列表。取值范围为: 2700-2899

permit: 该选项若被配置, 表示本规则属于允许通过类的;

deny: 该选项若被配置, 关键字表示本规则属于禁止通过类的;

protocol: IP 协议号, 取值范围[0, 255]; 为方便使用, 系统提供了常用 IP 协议号的简称以取代对应的 IP 协议号具体数值, 包括 eigrp、gre、icmp、igmp、ip、ipinip、nos、ospf、tcp、udp

ethernet-type: 该选项若被配置, 表示要匹配指定以太网类型的二层报文;

cos out: 该选项若被配置, 表示要匹指定二层报文外层 TAG 中的优先级字段;

cos inner in: 该选项若被配置, 表示要匹指定二层报文内层 TAG 中的优先级字段;

VID out: 该选项若被配置, 表示要匹指定二层报文外层 TAG 中的 VLAN ID 字段;

VID inner in: 该选项若被配置, 表示要匹指定二层报文内层 TAG 中的 VLAN ID 字段;

source source-wildcard: 该选项若被配置, 表示要匹配某一个 IP 网段的内主机发出的报文;

host source: 该选项若被配置, 表示要匹配源 IP 为某一台主机发出的 IP 报文;

any: 该选项若被配置, 表示要匹配任意主机发出的 IP 报文;

host source-mac-address: 该选项若被配置, 表示要匹配源 MAC 为某一台主机发出的二层报文;

any: 该选项若被配置, 表示要匹配目的为任意主机发出的二层报文;

destination destination-wildcard: 该选项若被配置, 表示要匹配目标为某一个 IP 网段的报文;

host destination: 该选项若被配置, 表示要匹配目的 IP 为某一台主机的 IP 报文;

any: 该选项若被配置, 表示要匹配发往任意目标的 IP 报文;

host destination-mac-address: 该选项若被配置, 表示要匹配目的 MAC 为某一台主机的二层报文;

any: 该选项若被配置, 表示要匹配目标为任意主机的二层报文;

precedence precedence: 该选项若被配置, 表示要匹配 IP 报文头部中的优先级域。

tos tos: 该选项若被配置, 表示要匹配 IP 报文头部中的服务类型域。

dscp dscp: 该选项若被配置, 表示要匹配 IP 报文头部的 dscp 域。

fragment: 该选项若被配置, 表示只要匹配非首片的 IP 分片报文。

time-range *time-range-name*: 该选项若被配置,表示该匹配规则关联了一个时间区,只有在指定的时间区间内该规则才会生效,否则不生效,更多关于关时间区的描述,请参考 time range 的配置手册。

【命令模式】 Expert 扩展访问列表模式

【使用指导】 此命令在配置模式下为数字索引的 Expert 扩展访问列表配置规则。这种配置方式无法为命名的 Expert 扩展访问列表配置规则。

应用 Expert 扩展访问列表

【命令格式】 **expert access-group** { *acl-id* | *acl-name* } { *in* | *out* }

【参数说明】 *acl-id*: 该选项若被配置,表示要将一个数值索引的 Expert 扩展访问列表应用在接口上。

acl-name: 该选项若被配置,表示要将一个命名的 Expert 扩展访问列表应用在接口上。

in: 该选项若被配置,表示这个访问列表对进入该接口的二层报文进行控制。

out: 该选项若被配置,表示这个访问列表对从该接口发出的二层报文进行控制。

【命令模式】 接口模式

【使用指导】 此命令可以让 Expert 扩展访问列表在指定的接口上生效,同时需要指定对进入设备的报文生效,还是从设备转发出去的报文生效。

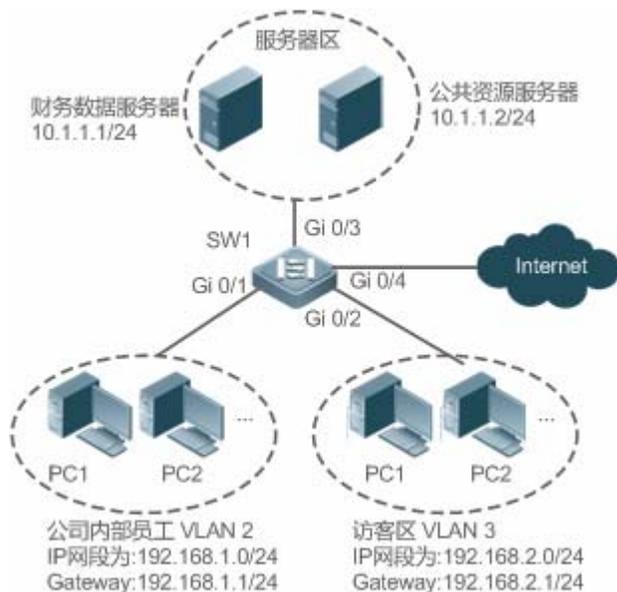
配置举例

i 以下配置举例,仅介绍与 ACL 相关的配置。

通过 Expert 扩展访问列表,限制来访客区户可访问的资源,要求访客不能与公司内部员工互访,但能访问公共资源服务器,且不能访问公司核心的账务数据服务器。

【网络环境】

图 1-5



【配置方法】

- 配置 Expert 扩展访问列表
- 在访问列表中添加规则,禁止访客区 VLAN 3 网段内主机发出目标为内部员工 VLAN2 网段的报文进入网

络。

- 在访问列表中添加规则，禁止访客访问核心账务数据服务器规则，
- 再添加一条规则，允许所有报文通过；
- 最后再将访问列表应用在与访客区相连交换机接口的入方向上。

SW1

```
swl(config)#expert access-list extended 2700
swl(config-exp-nacl)#deny ip any any 192.168.1.0 0.0.0.255 any
swl(config-exp-nacl)#deny ip any any host 10.1.1.1 any
swl(config-exp-nacl)#pemit any any any any
swl(config-exp-nacl)#exit
swl(config)#int gigabitEthernet 0/2
swl(config-if-GigabitEthernet 0/2)#expert access-group 2700 in
```

【检验方法】

- 从访客 PC 机上 ping 财务数据服务器，确认 ping 不通。
- 从访客 PC 机上 ping 公共资源服务器，确认可以 ping 得通。
- 从访客 PC 机上 ping 公司内部员工网关 192.168.1.1，确定 ping 不通。
- 在访问 PC 机上访问 Internet，比如访问百度，确认可以打开主页。

SW1

```
swl(config)#show access-lists
expert access-list extended 2700
 10 deny ip any any 192.168.1.0 0.0.0.255 any
 20 deny ip any any host 10.1.1.1 any
 30 permit ip any any any any

swl(config)#show access-group
expert access-group 2700 in
Applied On interface GigabitEthernet 0/2
```

1.5.4 配置ACL80

配置效果

当固定匹配域的 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表都无法满足要求时，那么可以通过配置 ACL80，由用户自己定义自己想匹配的报文域，从而实现自定义匹配域的目的。

注意事项

无

配置方法

配置 Expert 高级访问列表

- 必须配置。要实现 ACL80 的功能，首先就是要配置 Expert 高级访问列表，Expert 高级访问列表配置请参考相关章节的说明。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。Expert 高级访问列表只对被配置的设备上有效，不会影响网络中的其他设备。

配置 Expert 高级访问列表规则

- 必须配置。要实现自定义匹配域，必须配置自定义的访问列表规则。如果不配置访问列表规则，则默认的 deny 所有表项将会将所有报文丢弃。Expert 访问列表规则配置请参考相关章节的说明

应用 Expert 高级访问列表

- 必须配置。要使得 Expert 高级访问列表真正生效，就必须将 Expert 高级访问列表应用到设备的特定接口上。
- 可以根据用户的分布，在接入、汇聚或核心设备的指定接口上应用 Expert 高级访问列表。

检验方法

可以通过以下方法检验 Expert 高级访问列表的配置效果：

- 通过 ping 的方式来验证配置是否生效。
- 通过构造符合访问列表规则的报文来验证规则是否生效。

相关命令

配置 Expert 高级访问列表

【命令格式】 **expert access-list advanced** *acl-name*

【参数说明】 *acl-name*: Expert 高级访问列表的名称，长度范围[1, 99]。访问列表名称不能以数字 0 - 9 开头，也不能为“in”或“out”。

【命令模式】 配置模式

【使用指导】 此命令可以用来配置 MAC 扩展访问列表，并进入 Expert 扩展访问列表配置模式。

配置 Expert 高级访问列表规则

【命令格式】 [*sn*] { **permit** | **deny** } *hex hex-mask offset*

【参数说明】 *sn*: 为规则表项的序号，取值范围为[1, 2147483647]。这个序号决定了这条规则表项在该访问列表中的优先级，序号越小，优先级越大，优先级大的会优先去匹配报文。如果配置匹配规则时没有指定序号，系统会自动分配一个序号，序号的分配原则为在当前访问列表最后一条匹配规则的序列基础之上加上一个递增值，递增值默认为 10，假设当前访问列表最后一条匹配规则的序号为 100，则缺省情况下新增的这条匹配规则序号就为 11，此外，递增值是可以由命令调整的。

permit: 该选项若被配置，表示本规则属于允许通过类的；

deny: 该选项若被配置，关键字表示本规则属于禁止通过类的；

hex: 以 16 进制表示的自定义匹配内容。比如 00d0f800。

hex-mask: 匹配掩码；

offset: 匹配开始的位置，比如匹配内容为 00d0f800，匹配掩码为 00ff0000，开始位置为 6，表示要匹配报文中的目的 MAC 地址，所有目的 MAC 地址中的第二字节为 d0 的报文都能匹配到这条规则；

【命令模式】 Expert 高级访问列表模式

【使用指导】 此命令在 Expert 高级访问列表模式下为访问列表配置自定义规则。

应用 Expert 高级访问列表

【命令格式】 **expert access-group acl-n { in | out }**

【参数说明】 **acl-id**: 该选项若被配置，表示要将一个数值索引的 Expert 扩展访问列表应用在接口上。

acl-name: 该选项若被配置，表示要将一个命名的 Expert 扩展访问列表应用在接口上。

in: 该选项若被配置，表示这个访问列表对进入该接口的二层报文进行控制。

out: 该选项若被配置，表示这个访问列表对从该接口发出的二层报文进行控制。

【命令模式】 接口模式

【使用指导】 此命令可以让 Expert 扩展访问列表在指定的接口上生效，同时需要指定对进入设备的报文生效，还是从设备转发出去的报文生效。

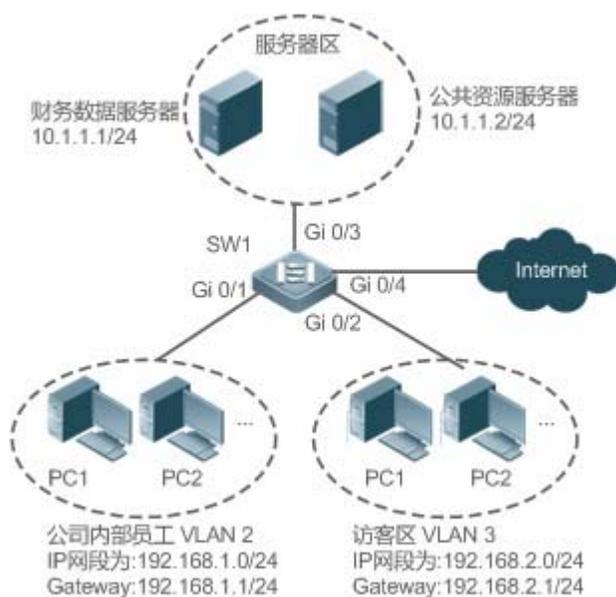
配置举例

i 以下配置举例，仅介绍与 ACL 相关的配置。

通过 ACL80 即 Expert 高级访问列表，限制来访客区可访问的资源，要求访客不能与公司内部员工互访，但能访问公共资源服务器，且不能访问公司核心的账务数据服务器。

【网络环境】

图 1-6



- 【配置方法】
- 配置 Expert 高级访问列表
 - 在访问列表中添加规则，禁止访客区 VLAN 3 网段内主机发出目标为内部员工 VLAN2 网段的报文进入网络。
 - 在访问列表中添加规则，禁止访客访问核心账务数据服务器规则，
 - 再添加一条规则，允许所有报文通过；
 - 最后再将访问列表应用在与访客区相连交换机接口的入方向上。

```
SW1 swl(config)#expert access-list advanced acl80-guest
swl(config-exp-dacl)#deny COA801 FFFFFFF 42
swl(config-exp-dacl)#deny OA010101 FFFFFFFF 42
swl(config-exp-dacl)#permit 0806 FFFF 24
swl(config-exp-dacl)#permit 0800 FFFF 24
swl(config-exp-dacl)#exit
swl(config)#int gigabitEthernet 0/2
swl(config-if-GigabitEthernet 0/2)#expert access-group acl80-guest in
```

- 【检验方法】
- 从访客 PC 机上 ping 财务数据服务器，确认 ping 不通。
 - 从访客 PC 机上 ping 公共资源服务器，确认可以 ping 得通。
 - 从访客 PC 机上 ping 公司内部员工网关 192.168.1.1，确定 ping 不通。
 - 在访问 PC 机上访问 Internet，比如访问百度，确认可以打开主页。

```
SW1 swl(config)#show access-lists
expert access-list advanced sss
 10 deny COA801 FFFFFFF 42
 20 deny OA010101 FFFFFFFF 42
 30 permit 0806 FFFF 24
 40 permit 0800 FFFF 24

expert access-group acl80-guest in
Applied On interface GigabitEthernet 0/2
```

1.5.5 配置ACL重定向

配置效果

通过在指定接口上配置 ACL 重定向功能，可以对进入在该接口的指定报文直接重定向指定端口转发出去。

注意事项

无

配置方法

配置访问列表

- 必须配置。要实现 ACL 重定向，首先就是要配置访问列表，比如 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表等，访问列表配置请参考相关章节的说明。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。

配置访问列表规则

- 可选配置。访问列表里允许无规则，没有配置规则时，相当于 ACL 重定向功能不存在。访问列表规则配置请参考相关章节的说明

配置 ACL 重定向

- 必须配置。要使得 ACL 重定向起作用，就必须在指定接口上开启重定向功能。
- 可以根据用户的分布，在接入、汇聚或核心设备的指定接口上配置 ACL 重定向功能。

检验方法

可以通过在配置 ACL 重定向所在的端口上发送符合规则的报文，然后在目标端口上使用抓包软件验证 ACL 重定向功能是否生效。

相关命令

配置访问列表

访问列表的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置 ACL 重定向

【命令格式】 **redirect destination interface** *interface-name* **acl** {*acl-id* | *acl-name*} **in**

【参数说明】 **interface** *interface-name*: 重定向目标端口名称。

acl-id: 访问列表的编号。

acl-name: 访问列表的名称。

in: 对进入接口的报文进行重定向。

【命令模式】 接口模式

【使用指导】 通过该命令，从指定接口进来的报文如果符合 ACL 规则就会被重定向到目标端口转发出去。

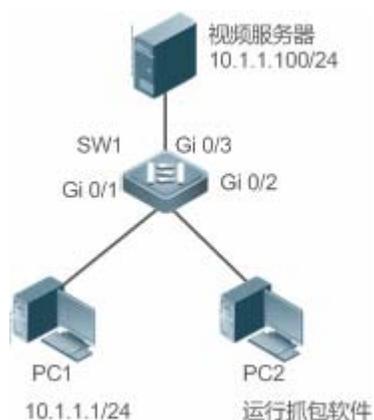
配置举例

i 以下配置举例，仅介绍与 ACL 相关的配置。

通过 ACL 重定向功能，主机 10.1.1.1 发出的报文重定向到抓包设备上进行分析

【网络环境】

图 1-7



【配置方法】

- 配置 IP 访问列表
- 在 IP 访问列表中添加允许主机 10.1.1.1 地址规则
- 在 Gi 0/1 接入上配置 ACL 重定向，目标端口为 Gi 0/2

SW1

```

sw1(config)#ip access-list standard 1
sw1 (config-std-nacl)#permit host 10.1.1.1
sw1(config-std-nacl)#exit
sw1(config)#int gigabitEthernet 0/1
sw1(config-if-GigabitEthernet 0/1)# redirect destination interface gigabitEthernet 0/2 acl 1
  
```

【检验方法】

- 在 PC2 上开启抓包，从 PC1 ping 视频服务器，在 PC2 上确认有抓到 PC1 发出的 ICMP 请求报文。

SW1

```

sw1#show access-lists
ip access-list standard 1
  10 permit host 10.1.1.1
sw1#show redirect interface gigabitEthernet 0/1
acl redirect configuration on interface gigabitEthernet 0/1
redirect destination interface gigabitEthernet 0/2 acl 1 in
  
```

1.5.6 配置全局安全ACL

配置效果

通过配置全局安全 ACL 功能，可以起到阻止企业内部访问非法网站，或者阻止病毒进入企业内部网络的目的。另外，通过配置全局安全 ACL 例外口，允许企业内部某些特殊部门可以访问外部一些站点。

注意事项

无

配置方法

配置 IP 访问列表

- 必须配置。要实现全局防护内部网络的目的，首先要配置 IP 访问列表，相关的配置方法请参考 IP 访问列表章节说明。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。配置仅在本设备上有效，不会影响网络中的其他设备。

配置 IP 访问列表规则

- 可选配置。访问列表里允许无规则，没有配置规则时，相当于全局安全 ACL 功能不存在。访问列表规则配置请参考相关章节的说明

配置全局安全 ACL

- 必须配置。要使得全局安全 ACL 起作用，就必须在开启全局安全功能。
- 可以根据用户的分布，在接入、汇聚或核心设备配置全局安全 ACL 功能。

检验方法

可以在受全局安全 ACL 防护的网络内部 ping 被规则拒绝的站点或设备来验证全局安全 ACL 是否生效。

相关命令

配置 IP 访问列表

配置方法请参考 IP 访问列表的相关章节说明。

配置 IP 访问列表规则

配置方法请参考 IP 访问列表的相关章节说明。

配置全局安全 ACL

【命令格式】 `ip access-group acl-id { in | out }`

【参数说明】 *acl-id*: IP 访问列表的编号。

in: 对进入设备的报文进行匹配过滤。

out: 对从设备转发出去的报文进行匹配过滤。

【命令模式】 配置模式

【使用指导】 通过该命令开启全局安全 ACL 功能，使得 ACL 在设备的所有二层口上生效。

! 全局安全 ACL 目前仅支持在接口的入方向上生效

配置全局安全 ACL 例外口

- 【命令格式】 **no global ip access-group**
- 【参数说明】 无
- 【命令模式】 接口模式
- 【使用指导】 通过该命令使得全局安全 ACL 在指定的接口上不生效。

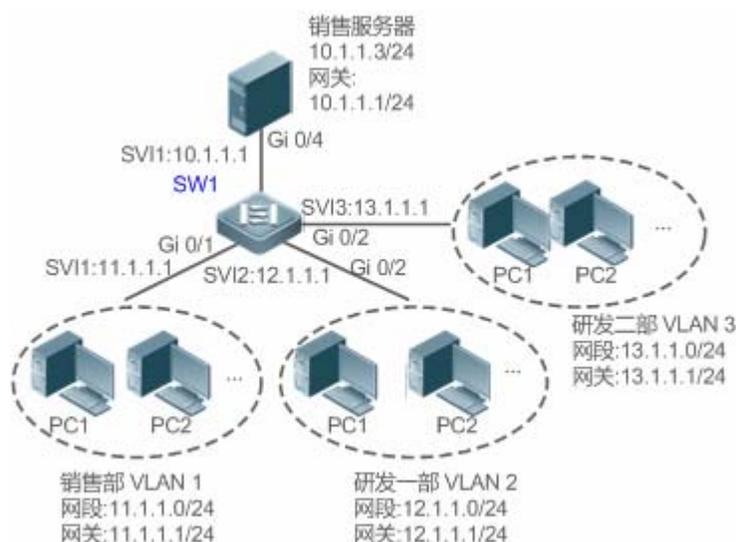
配置举例

以下配置举例，仅介绍与 ACL 相关的配置。

通过全局安全 ACL 功能，禁止研发部门访问销售服务器，但要允许销售部门访问

【网络环境】

图 1-8



- 【配置方法】
- 配置 IP 扩展访问列表 ip_ext_deny_dst_sale_server
 - 在 IP 访问列表中添加禁止目的主机 10.1.1.3/24 地址规则
 - 将访问列表 ip_ext_deny_dst_sale_server 配置为全局安全 ACL
 - 将与销售部直连的端口配置为全局安全 ACL 例外口

SW1

```
sw1(config)#ip access-list extended ip_ext_deny_dst_sale_server
sw1(config-ext-nacl)# deny ip any host 10.1.1.3
sw1(config-ext-nacl)#exit
sw1(config)#ip access-group ip_ext_deny_dst_sale_server in
sw1(config)#int gigabitEthernet 0/1
sw1(config-if-GigabitEthernet 0/1)# no global ip access-group
```

- 【检验方法】
- 在销售部内的某台 PC 机 ping 销售服务器地址，确认可以 ping 得通。
 - 在研发一部和研发二部的 PC 机上 ping 销售服务器地址，确认 ping 不通。

```
sw1#show access-lists
ip access-list extended ip_ext_deny_dst_sale_server
```

```
10 deny ip any host 10.1.1.3
sw1#show running
.....
!
ip access-group ip_ext_deny_dst_sale_server in
!
!
!
!
!
!
!
!
!
!
interface GigabitEthernet 0/1
no global ip access-group
!
.....
```

1.5.7 配置安全通道

配置效果

通过配置安全通道功能，可以使得符合安全通道规则的报文绕过接入控制相关业务。如果用户上联的设备接口上开启了某个接入控制应用比如 dot1x，但在进行 dot1x 认证前，又要允许用户登录到某个站点上下载一些资源（比如下载锐捷 SU 客户端），这种情况就可以通过配置安全通道来实现。

注意事项

无

配置方法

📌 配置访问列表

- 必须配置。要实现安全通道功能，首先要配置访问列表，访问列表的配置方法请参考相关章节说明。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。配置仅在本设备上有效，不会影响网络中的其他设备。

📌 配置访问列表规则

- 可选配置。访问列表里允许无规则，没有配置规则时，相当于安全通道功能不生效。访问列表规则配置请参考相关章节的说明。

配置接口安全通道或全局安全通道

- 如果想让安全通道在接口上生效，就在接口上配置安全通道；如果想让安全通道全局生效，就要配置全局安全通道，必须配置其中之一。
- 可以根据用户的分布，在接入、汇聚或核心设备配置安全通道功能。

配置全局安全通道例外口

- 可选配置。如果配置了全局安全通道，但又不想让安全通道在某些接口上生效，就需要将这些接口配置为全局安全通道的例外口。

配置接入控制应用

- 可选配置，为了验证安全通道功能，可以在接口上开启 dot1x 或 web 认证功能。
- 可以根据用户的分布，在接入、汇聚或核心设备配置接入控制功能。

检验方法

可以通过在受接入控制业务控制的用户 PC 机上 ping 安全通道指定放过的资源（设备或服务器）来验证安全通道。

相关命令

配置访问列表

访问列表的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置接口安全通道

【命令格式】 **security access-group** {acl-id | acl-name }

【参数说明】 *acl-id*: 该选项若被配置，表示要将指定编号的访问列表配置成安全通道。

acl-name: 该选项若被配置，表示要将指定的命名访问列表配置成安全通道

【命令模式】 接口模式

【使用指导】 通过该命令在指定接口上将指定的 ACL 配置成安全通道。

配置全局安全通道

【命令格式】 **security global access-group** {acl-id | acl-name }

【参数说明】 *acl-id*: 该选项若被配置，表示要将指定编号的访问列表配置成安全通道。

acl-name: 该选项若被配置，表示要将指定的命名访问列表配置成安全通道

【命令模式】 配置模式

【使用指导】 通过该命令将指定的 ACL 配置成全局安全通道。

配置全局安全通道例外口

- 【命令格式】 **security uplink enable**
- 【参数说明】 无
- 【命令模式】 接口模式
- 【使用指导】 通过该命令将指定的接口配置成全局安全通道例外口。

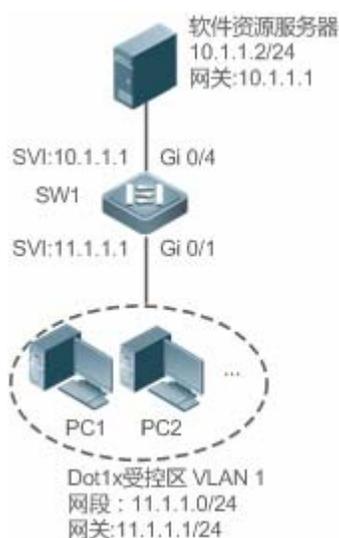
配置举例

i 以下配置举例，仅介绍与 ACL 相关的配置。

在 dot1x 认证环境中，通过安全通道，允许用户认证前从服务器上下载 SU 客户端软件

【网络环境】

图 1-9



- 【配置方法】
- 配置 Expert 扩展访问列表 exp_ext_esc
 - 在访问列表中添加允许目的主机 10.1.1.2 地址规则
 - 在访问列表中添加允许 DHCP 报文通过规则
 - 在访问列表中添加允许 ARP 报文通过规则
 - 在 dot1x 受控区接口上将访问列表 exp_ext_esc 配置为安全通道

SW1

```
swl(config)#expert access-list extended exp_ext_esc
swl(config-exp-nacl)# permit ip any any host 10.1.1.2 any
swl(config-exp-nacl)# permit 0x0806 any any any any any
swl(config-exp-nacl)# permit tcp any any any any eq 67
swl(config-exp-nacl)# permit tcp any any any any eq 68
swl(config)#int gigabitEthernet 0/1
swl(config-if-GigabitEthernet 0/1)# security access-group exp_ext_esc
```

- 【检验方法】
- 在销售部内的某台 PC 机 ping 销售服务器地址，确认可以 ping 得通。
 - 在研发一部和研发二部的 PC 机上 ping 销售服务器地址，确认 ping 不通。

```
sw1#show access-lists
expert access-list extended exp_ext_esc
 10 permit ip any any host 10.1.1.2 any
 20 permit arp any any any any any
 30 permit tcp any any any any eq 67
 40 permit tcp any any any any eq 68.....

sw1#show running-config interface gigabitEthernet 0/1

Building configuration...
Current configuration : 59 bytes

interface GigabitEthernet 0/1
 security access-group exp_ext_esc
```

1.5.8 配置基于时间区的规则

配置效果

如果想让访问列表的某些规则在指定的时间生效，或在指定的时间内失效，比如让 ACL 在一个星期的某些时间段内生效等。可以配置基于时间区的访问列表规则。

注意事项

无

配置方法

配置访问列表

- 必须配置。要实现基于时间区生效的规则，首先要配置访问列表，访问列表的配置方法请参考相关章节说明。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。配置仅在本设备上有效，不会影响网络中的其他设备。

配置带时间区的访问列表规则

- 必须配置。配置时需要带上对应的时间区选项，时间区的配置请参考时间区相关的配置手册。

应用访问列表

- 必须配置。要使得访问列表规则在指定的时间区内生效，就必须访问列表应用到设备的特定接口上。
- 可以根据用户的分布，在接入、汇聚或核心设备的指定接口上应用 IP 访问列表。

检验方法

在生效时间区内,可以通过 ping 或构造符合规则报文的方式来进行检验规则是否生效来检验;在失效时间区内,可以通过 ping 或构造符合规则报文的方式来进行检验规则是否不生效来检验。

相关命令

配置访问列表

访问列表的配置命令请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置带时间区的访问列表规则

访问列表规则的配置命令请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

应用访问列表

访问列表规则的应用命令请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

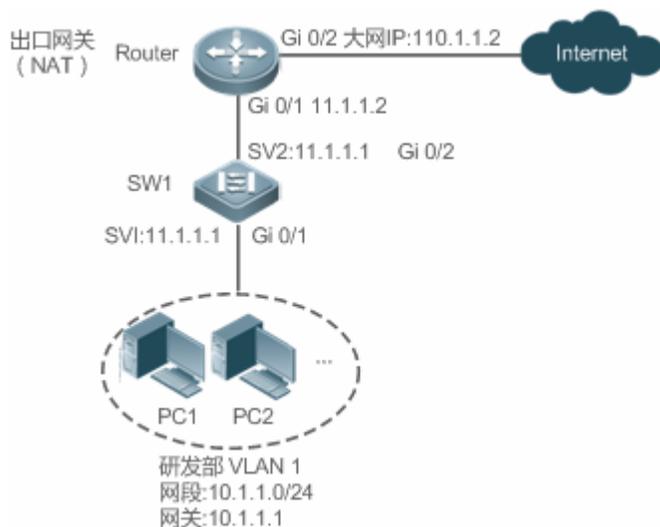
配置举例

i 以下配置举例,仅介绍与 ACL 相关的配置。

配置基于时间区的访问列表规则,只允许研发部门在每天的 12:00 到 13:30 访问 internet

【网络环境】

图 1-10



【配置方法】

- 配置名称为 access-internet 的时间区,并添加每天 12:00 到 13:30 的时间段表项。
- 配置 IP 访问列表 ip_std_internet_acl。
- 在访问列表中添加允许源 IP 网段为 10.1.1.0/24 的地址规则,关联的时间区为 access-internet。
- 在访问列表中添加禁止源 IP 网段为 10.1.1.0/24 的地址规则。表明时间区之外都不允许访问 internet
- 在访问列表中添加允许所有的地址规则

- 将访问列表应用在设备与出口网关相连接口的出方向上。

SW1

```
Ruijie(config)# time-range access-internet
Ruijie(config-time-range)# periodic daily 12:00 to 13:30
Ruijie(config-time-range)# exit
swl(config)# ip access-list standard ip_std_internet_acl
swl(config-std-nacl)# permit 10.1.1.0 0.0.0.255 time-range access-internet
swl(config-std-nacl)# deny 10.1.1.0 0.0.0.255
swl(config-std-nacl)# permit any
swl(config-std-nacl)# exit
swl(config)# int gigabitEthernet 0/2
swl(config-if-GigabitEthernet 0/2)# ip access-group ip_std_internet_acl out
```

【检验方法】

- 在时间区生效期内（12:00 至 13:30），从研发部分内的某台 PC 机访问百度主页，确认可以访问。
- 在时间区失效期（12:00 至 13:30 这个时段外），从研发部分内的某台 PC 机访问百度主页，确认不能访问。

SW1

```
swl#show time-range

time-range entry: access-internet (inactive)
  periodic Daily 12:00 to 13:30

swl#show access-lists

ip access-list standard ip_std_internet_acl
  10 permit 10.1.1.0 0.0.0.255 time-range access-internet (inactive)
  20 deny 10.1.1.0 0.0.0.255
  30 permit any

swl#show access-group

ip access-group ip_std_internet_acl out
Applied On interface GigabitEthernet 0/2
```

1.5.9 配置访问列表注释信息

配置效果

在实际的网络维护过程中，如果配置了很多访问列表且没有为这些访问列表配置注释信息，时间一长往往会难以区分这些访问列表的用途。为访问列表配置注释信息，可以方便理解 ACL 用途。

注意事项

无

配置方法

配置访问列表

- 必须配置。要实现安全通道功能，首先要配置访问列表，访问列表的配置方法请参考相关章节说明。
- 可以根据用户的分布情况，在接入、汇聚或核心设备上配置。配置仅在本设备上有效，不会影响网络中的其他设备。

配置访问列表注释信息

- 可选配置。为便于管理和理解所配置的访问列表，可以为访问列表配置注释信息。

配置访问列表规则

- 可选配置。访问列表里允许无规则，没有配置规则时，相当于安全通道功能不生效。访问列表规则配置请参考相关章节的说明。

配置访问列表规则注释信息

- 可选配置。为便于理解所配置的访问列表，除了可以为访问列表本身配置注释信息外，还可以为规则配置注释信息。

检验方法

可以通过在设备上使用 **show access-lists** 命令验证访问列表注释信息，。

相关命令

配置访问列表

访问列表的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置访问列表注释信息

有以下两种方式为访问列表配置注释信息：

【命令格式】 **list-remark comment**

【参数说明】 *comment*: 注释信息。长度[1, 100]，超过 100 个字符将被截短至 100 个字符

【命令模式】 访问列表模式

【使用指导】 通过该命令为指定的访问列表配置注释信息

【命令格式】 **access-list acl-id list-remark comment**

【参数说明】 *acl-id*: 访问列表编号

comment: 注释信息。长度[1, 100]，超过 100 个字符将被截短至 100 个字符

- 【命令模式】 配置模式
- 【使用指导】 通过该命令为指定的访问列表配置注释信息

配置访问列表规则

访问列表规则的配置方法请参考 IP 访问列表、MAC 扩展访问列表、Expert 扩展访问列表的相关章节说明。

配置访问列表规则注释信息

有以下两种方式为访问列表规则配置注释信息：

- 【命令格式】 **remark comment**
- 【参数说明】 *comment*: 注释信息。长度[1, 100]，超过 100 个字符将被截短至 100 个字符
- 【命令模式】 访问列表模式
- 【使用指导】 通过该命令为指定的访问列表规则配置注释信息

- 【命令格式】 **access-list acl-id remark comment**
- 【参数说明】 *acl-id*: 访问列表编号
comment: 注释信息。长度[1, 100]，超过 100 个字符将被截短至 100 个字符
- 【命令模式】 配置模式
- 【使用指导】 通过该命令为访问列表规则添加注释信息

配置举例

无

1.6 监视与维护

清除各类信息

作用	命令
清除访问列表报文匹配计数	clear counters access-list [acl-id acl-name]
清除访问列表 deny 报文匹配计数	clear access-list counters [acl-id acl-name]

查看运行情况

作用	命令
查看基本访问列表	show access-lists [acl-id acl-ame] [summary]

显示指定接口上绑定的重定向表项，不输入接口则显示所有接口上绑定的重定向表项。	show redirect [interface <i>interface-name</i>]
显示接口上应用的访问列表配置信息。	show access-group [interface <i>interface-name</i>]
显示接口上应用的 IP 访问列表配置信息。	show ip access-group [interface <i>interface-name</i>]
显示接口上应用的 MAC 扩展访问列表配置信息。	show mac access-group [interface <i>interface-name</i>]
显示接口上应用的 Expert 扩展访问列表配置信息。	show expert access-group [interface <i>interface-name</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
监视访问列表运行过程信息	debug acl acld event
调试查看 ACL 客户端的信息	debug acl acld client-show
调试查看所有 ACL 客户端创建的访问列表信息	debug acl acld acl-show

2 QoS

2.1 概述

QoS (Quality of Service , 服务质量) 指一个网络能够利用各种基础技术 , 为指定的网络通信提供更好的服务能力。

当网络带宽充裕的时候 , 所有的数据流都得到了较好的处理 ; 而当网络发生拥塞的时候 , 所有的数据流都有可能被丢弃 ; 为满足用户对不同应用不同服务质量的要求 , 就需要网络能根据用户的要求分配和调度资源 , 对不同的数据流提供不同的服务质量 : 对实时性强且重要的数据报文优先处理 ; 对于实时性不强的普通数据报文 , 提供较低的处理优先级 , 网络拥塞时甚至丢弃。

传统网络所采用的 “尽力而为” 的转发机制 , 已经不能满足这些需求 , QoS 应运而生。支持 QoS 功能的设备 , 能够提供传输品质服务 ; 针对某种类别的数据流 , 可以为它赋予某个级别的传输优先级 , 来标识它的相对重要性 , 并使用设备所提供的各种优先级转发策略、拥塞避免等机制为这些数据流提供特殊的传输服务。配置了 QoS 的网络环境 , 增加了网络性能的可预知性 , 并能够有效地分配网络带宽 , 更加合理地利用网络资源。

2.2 典型应用

典型应用	场景描述
端口限速+优先级重标记应用	基于校园网的不同业务需求 , 对教学楼、实验室、宿舍楼的出口流量进行限速控制及优先级处理。
优先级重标记+队列调度应用	对于企业内部访问服务器的流量进行优先级处理及带宽控制。

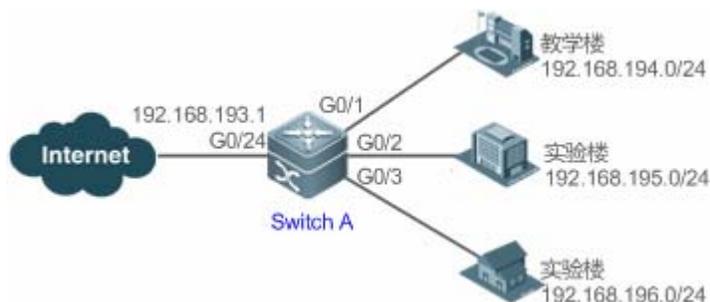
2.2.1 端口限速+优先级重标记

应用场景

某学校为了满足正常的教学业务需要 , 要求满足以下四点需求 :

- 限制该学校访问 Internet 的流量为 100M , 丢弃超出限制的报文 ;
- 限制宿舍楼的出口流量为 50M , 同样丢弃超出限制的报文 ;
- 限制实验楼发出 DSCP 优先级为 7 的报文的速率为 20M , 将速率超过 20M 的此类报文的 DSCP 优先级修改为 16 ;
- 限制教学楼的出口流量为 30M , 丢弃超出限制的报文。

图 2-1



【注释】 某学校通过 SwitchA 的 GigabitEthernet 0/24 上联 Internet，SwitchA 的 GigabitEthernet 0/1、GigabitEthernet 0/2 和 GigabitEthernet 0/3 分别下联的教学楼、实验楼和宿舍楼。

功能部署

- 在 SwitchA 连接 Internet 的 G0/24 端口配置 QoS 端口速率限制；
- 在 SwitchA 上配置对宿舍楼发出的报文进行 QoS 限速；
- 在 SwitchA 上配置对实验楼发出的 DSCP 优先级为 7 的报文限速为 20M，并将超过限速的报文的 DSCP 优先级重标记为 16；
- 在 SwitchA 上配置对教学楼发出的报文进行 QoS 限速；。

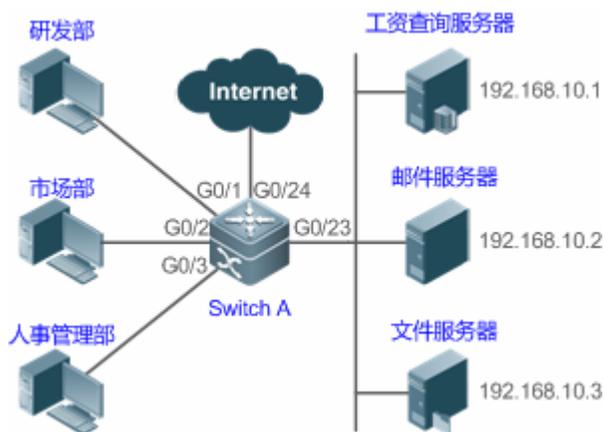
2.2.2 优先级重标记+队列调度应用

应用场景

配置优先级重标记和队列调度，实现下述需求：

- 当研发部和市场部访问服务器时，服务器报文的优先级为：邮件服务器> 文件服务器>工资查询服务器；
- 无论人事管理部访问 Internet 或访问服务器，交换机都优先处理；
- 交换机在运行过程中，时常发现网络拥塞，为了保证业务顺利运转，要求使用 WRR 队列调度，对研发部和市场部访问邮件数据库、访问文件数据库、访问工资查询数据库的 IP 数据报按照 6：2：1 的比例来调度。

图 2-2



【注释】 研发部、市场部和人事管理部分别接入 SwitchA 的端口 GigabitEthernet 0/1、GigabitEthernet 0/2 和 GigabitEthernet 0/3；工资查询服务器、邮件服务器和文件服务器连接在 SwitchA 的端口 GigabitEthernet 0/23 下。

功能部属

- 通过配置访问不同服务器数据流的 CoS 值，实现设备处理访问各种服务器报文的优先级；
- 通过配置接口的缺省 CoS 值为特定值，实现设备优先处理人事管理部发出的报文；
- 通过配置 WRR 队列调度实现按特定个数比进行数据报文传输调度。

2.3 功能详解

基本概念

差服务模型

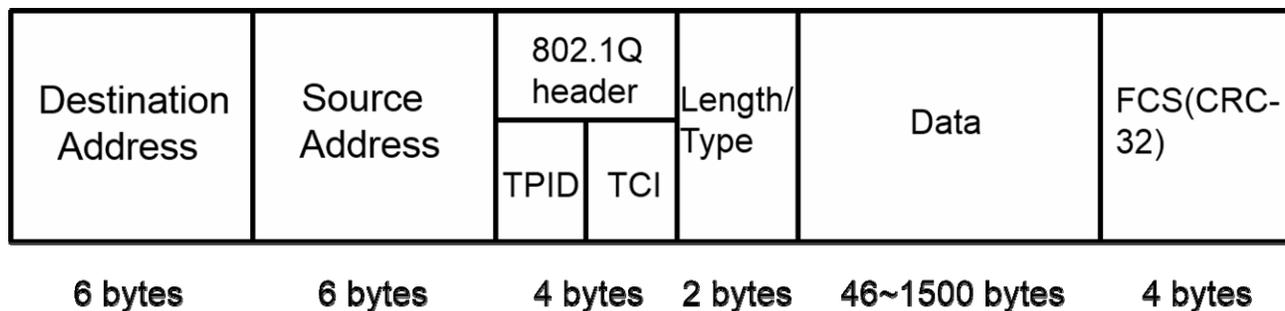
DiffServ (Differentiated Services Mode , 差分服务模型)，锐捷产品的 QoS 实现以 IETF 的 DiffServ 体系为基础。DiffServ 体系规定网络中的每一个传输报文将被划分成不同的类别，分类信息包含在二层/三层报文头中，包括：802.1P 优先级、IP 优先级、IP DSCP 优先级。

在遵循 DiffServ 体系的网络中，各设备对包含相同分类信息的报文采取相同的传输服务策略，对包含不同分类信息的报文采取不同的传输服务策略。报文的分类信息可以由网络上的主机或者其它网络设备赋予，也可以基于不同的应用策略或者基于报文内容的不同为报文赋予类别信息。设备根据报文所携带的类别信息，为各种报文流提供不同的传输优先级，或者为某种报文流预留带宽，或者适当地丢弃一些优先级较低的报文，或者采取其他一些操作等等。

802.1P(PRI)优先级

802.1 P 优先级位于带有 802.1Q 标签头的二层报文头部，适用于不需要分析三层报头，而需要在二层环境下保证 QoS 的场合，结构如下：

图 2-3

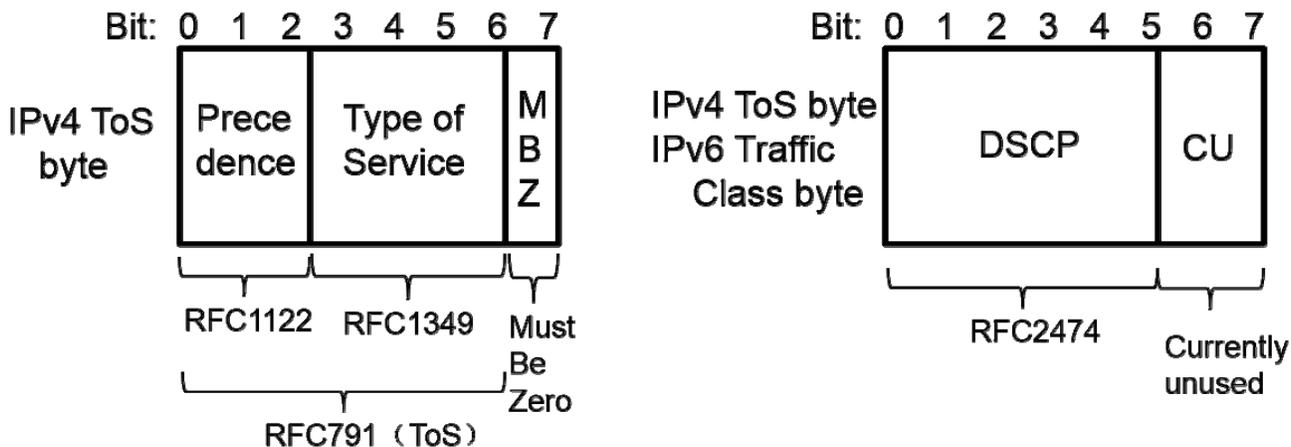


如上图所示，4个字节的802.1Q标签头包含了2个字节的TPID(Tag Protocol Identifier，标签协议标识，取值为0x8100)和2个字节的TCI(Tag Control Information，标签控制信息)，其中TCI中的前3位即802.1P优先级。

IP 优先级(IP PRE)和 DSCP 优先级

IP 报文使用 IP 优先级和 DSCP 优先级表示报文优先级。IPv4 头的 ToS (Type Of Service，服务类型) 字段有 8 个 bit，其中前 3 个 bit 表示的就是 IP PRE (IP precedence)，取值范围为 0~7。在 RFC 2474 中，重新定义了 IPv4 报文头部的 ToS 域，称之为 DS(Differentiated Services，差分服务)域，其中 DSCP (Differentiated Services Code Point) 优先级用该域的前 6 位 (0~5 位) 表示；IPv6 中使用报文头中的 Traffic Class 字段的前 6 个比特，表示 DSCP。IPv4/IPv6 报文的 IP PRE 和 DSCP 优先级位置如下图：

图 2-4



服务类别

CoS (Class of Service，服务类别) 锐捷产品中将报文优先级转化为 CoS 用于本地优先级，用于确定出端口发送时的入队列号。

功能特性

功能特性	作用
流分类	流分类采用一定的规则识别符合某类特征的报文，它是对网络业务进行区分服务的前提和基础。
优先级标记与映射	优先级标记与映射即标记报文的优先级为指定的值，并映射到相应的 CoS 值去。

流量监管	流量监管是监控进入网络的某一流量的规格,把它限制在一个合理的范围之内,丢弃超出部分的流量或者修改其优先级。
拥塞管理	拥塞管理根据数据包的优先级,来确定数据包从接口发送的顺序,在拥塞发生时,确保关键业务能够得到及时的服务。
拥塞避免	拥塞避免通过监控出端口队列的使用情况,在网络拥塞时,采取主动丢弃报文,调整网络流量的方式来解除网络过载。

2.3.1 流分类

流分类采用一定的规则识别符合某类特征的报文。它是对网络业务进行区分服务的前提和基础,通过流分类规则区分网络中不同的报文,再为不同服务等级的报文指定不同的 QoS 参数。

工作原理

流分类的规则可以是匹配 IP 报文的 PRE 或 DSCP 优先级、或者通过 ACL 识别报文的内容进行分类。用户可以通过命令定义多个流与流行为的绑定关系形成策略应用在接口上进行流分类与处理。

📌 QoS 策略

QoS 策略包含了三个要素:类、流行为、策略:

- 类
类是用来识别流的。类的要素包括:类的名称和类的规则。用户可以通过命令定义类的规则,来对报文进行分类。
- 流行为
流行为用来定义针对报文采取的 QoS 动作。流行为包括对报文进行优先级标记与流量监控。
- 策略
策略用来将指定的类和指定的流行为绑定起来。策略的要素包括:策略名称、绑定在一起的类的名称和流行为。用户通过 QoS 策略将指定的类和流行为绑定起来,然后再将策略应用到一个或多个端口上生效。

📌 QoS 逻辑端口组

可以指定一系列端口为一个 QoS 逻辑端口组(这里端口可以是 AP,也可以是以太网口),并针对这个逻辑端口组关联策略进行 QoS 处理,以流行为限速为例,对符合限速条件的报文,在同一个逻辑端口组内所有的端口共享策略所限定的带宽值。

相关配置

📌 创建类

缺省情况下,未定义任何类。

使用 `class-map` 命令,创建类并进入类配置模式。

📌 匹配 ACL

缺省情况下，类中未定义任何规则。

在类配置模式下，使用 **match access-group** 命令，定义类的规则为匹配 ACL，ACL 规则需先创建。

↘ 匹配 IP 报文的 PRE 优先级

缺省情况下，类中未定义任何规则。

在类配置模式下，使用 **match ip precedence** 命令，定义类的规则为匹配 IP 报文的 PRE 优先级，IP PRE 的取值范围为 0~7。

↘ 匹配 IP 报文的 DSCP 优先级

缺省情况下，类中未定义任何规则。

在类配置模式下，使用 **match ip dscp** 命令，定义类的规则为匹配 IP 报文的 DSCP 优先级，DSCP 优先级的取值范围为 0~63。

↘ 创建策略

缺省情况下，未定义任何策略。

使用 **policy-map** 命令，创建策略并进入策略配置模式。

↘ 关联类

缺省情况下，策略未关联任何类。

在策略配置模式下，使用 **class** 命令，关联类并进入策略类配置模式。

↘ 绑定流行为

缺省情况下，类未绑定任何流行为。

在策略类配置模式下，使用 **set** 命令对指定流修改 CoS、DSCP，其中 CoS 取值范围为 0~7，DSCP 取值范围为 0~63；使用 **police** 命令对指定流进行带宽限制及超限处理，带宽限制范围由产品决定。

↘ 配置逻辑端口组

缺省情况下，未定义任何逻辑端口组，接口也未加入任何逻辑端口组。

在全局模式下，使用 **virtual-group** 命令，来创建一个逻辑端口组；在接口配置模式下，使用 **virtual-group** 命令，将接口加入一个逻辑端口组，如果此时逻辑接口组还未创建，则创建逻辑接口组并将接口加入。可以创建 128 个逻辑端口组，取值范围为 1~128。

↘ 接口上应用策略

缺省情况下，接口上未应用任何策略。

在接口配置模式下，使用 **service-policy** 命令，在接口的输入/输出方向上应用策略。

2.3.2 优先级标记与映射

优先级用以标识报文的调度权重或者转发处理优先级的高低。根据报文类型不同，有不同的优先级类型：802.1P (PRI) 优先级、IP 优先级 (IP PRE)、DSCP 优先级。报文优先级标记与映射即标记报文的优先级为指定的值，并映射到相应的 CoS 值去。

工作原理

报文数据流进入设备端口之后，设备会根据端口配置的信任模式来分配报文的各类优先级。有如下几种形式：

- 端口信任模式为非信任时，即不信任报文中携带的优先级信息：
根据端口的默认 CoS (可配置，默认为 0) 和 COS-DSCP 映射表、DSCP-COS 映射表修改 CoS，根据最终的 CoS 入队列，如果报文出口带 802.1Q tag 那么报文优先级也会被修改成对应的 Cos。
- 端口信任模式为信任 CoS 时：
如果是带 802.1Q tag 的报文，根据报文的 PRI 值和 CoS-DSCP 映射表、DSCP-COS 映射表修改 COS，根据最终的 CoS 入队列，如果报文出口带 802.1Q tag 那么报文优先级也会被修改成对应的 Cos；
如果是不带 802.1Q tag 的报文，根据端口的默认 CoS (可配置，默认为 0)和 CoS-DSCP 映射表、DSCP-COS 映射表修改 COS，根据最终的 CoS 入队列，如果报文出口带 802.1Q tag 那么报文优先级也会被修改成对应的 Cos。
- 端口信任模式为信任 DSCP 时：
如果是非 IP 报文，按照信任 CoS 处理；
如果是 IP 报文，此时根据报文的 DSCP 值和 DSCP-CoS 映射表修改 CoS，根据最终的 CoS 入队列。
- 端口信任模式为信任 IP PRE 时：
如果是非 IPv4 报文，按照信任 CoS 处理；
如果是 IPv4 报文，此时根据报文的 IP PRE 值和 IP-PRE-DSCP 映射表，得到并修改报文的 DSCP，再根据 DSCP-CoS 映射表得到 CoS，根据最终的 CoS 入队列。
- 端口信任模式与应用于端口的策略同时作用情况下的关系：
当端口信任模式和应用于端口的策略同时作用时，端口信任模式修改 DSCP 和 CoS 的优先级低于策略修改 CoS、DSCP，并根据 DSCP-CoS 映射表得到 CoS 的优先级；
端口应用策略，但策略没有设置修改 DSCP 和 CoS 值时，按照此时端口的信任模式执行。

相关配置

▾ 配置端口的信任模式

缺省情况下，端口的信任模式为非信任。

在接口配置模式下，使用 `mls qos trust` 命令，修改信任模式，可配置的信任模式为信任 CoS、信任 DSCP、信任 IP PRE。

▾ 配置接口的缺省 CoS 值

缺省情况下，接口的 CoS 值为 0。

在接口配置模式下，使用 `mls qos cos` 命令，来修改接口缺省的 CoS 值，CoS 取值范围为 0~7。

▾ 流的优先级标记

缺省情况下，不对流的优先级进行重标记。

在策略类配置模式下，使用 `set` 命令，来修改流的 CoS、DSCP，其中 CoS 取值范围为 0~7，DSCP 取值范围为 0~63。

▾ 配置 CoS-to-DSCP Map

缺省情况下，CoS 值 0 1 2 3 4 5 6 7 分别映射到 DSCP 值 0 8 16 24 32 40 48 56。

使用 `mls qos map cos-dscp` 命令，来配置 CoS 值到 DSCP 值的映射，DSCP 取值范围为 0~63。

▾ 配置 DSCP-to-CoS Map

缺省情况下，DSCP 0~7 映射到 CoS 0，DSCP 8~15 映射到 CoS 1，DSCP 16~23 映射到 CoS 2，DSCP 24~31 映射到 CoS 3，DSCP 32~39 映射到 CoS 4，DSCP 40~47 映射到 CoS 5，DSCP 48~55 映射到 CoS 6，DSCP 56~63 映射到 CoS 7。

使用 `mls qos map dscp-cos` 命令，来配置 DSCP 值到 CoS 值的映射，其中 CoS 取值范围为 0~7，DSCP 取值范围为 0~63。

▾ 配置 IP-PRE-to-DSCP Map

缺省情况下，IP PRE 值 0 1 2 3 4 5 6 7 分别映射到 DSCP 值 0 8 16 24 32 40 48 56。

使用 `mls qos map ip-prec-dscp` 命令，来配置 IP PRE 值到 DSCP 值的映射，DSCP 取值范围为 0~63。

2.3.3 流量监管

流量监管是监控进入网络的某一流量的规格，把它限制在一个合理的范围之内，丢弃超出部分的流量或者修改报文优先级。同时能监控端口总的流量，丢弃超出部分的流量。

工作原理

流量监管监控进入网络的某一流量的规格，依据不同的评估结果，实施预先设定好的监管动作。这些动作可以是：

- 转发：对未超过流量限制的报文正常转发处理；
- 丢弃：对超过流量限制的报文进行丢弃；
- 改变优先级并转发：对超过流量限制的报文，修改其优先级后再转发。

对于超过端口总流量限制的报文，直接丢弃。

相关配置

▾ 配置流量超限后的动作

缺省情况下，未配置流量超限后动作。

在策略类配置模式下，使用 **police** 命令，来配置流量超限后的动作，可以配置为对超限流量丢弃，修改 CoS 或者 DSCP 值。流量超限范围由产品决定，流量超限后可修改的 CoS 取值范围为 0~7，DSCP 取值范围为 0~63。

配置端口总流量限制

缺省情况下，未配置端口总流量限制。

在接口配置模式下，使用 **rate-limit** 命令，来配置端口输入/输出方向总流量限制。流量限制范围由产品决定。

2.3.4 拥塞管理

当报文的接收速率超过发送速率时，在发送端口上就会出现拥塞，如果不能提供足够的缓冲区来保存这些报文，就会造成报文的丢失。拥塞管理机制根据数据包的优先权，来确定数据包发送出接口的顺序，拥塞管理功能允许对拥塞进行控制，对于一些重要的数据，提高数据报文的优先权，在拥塞发生时，优先发送，确保关键业务能够得到及时服务。

工作原理

使用队列调度机制进行拥塞管理，处理过程如下：

- 每个报文经过设备内部各个 QoS 处理环节，最终都会得到一个 CoS 值；
- 在出端口，设备会根据这个 CoS 值将报文归类到对应发送队列中；
- 出端口根据各种调度策略（SP、WRR、DRR、WFQ、SP+WRR、SP+DRR、SP+WFQ），选取其中一个队列的报文进行发送。

调度策略

队列调度策略分为 SP、WRR、DRR、WFQ、SP+WRR、SP+DRR、SP+WFQ。

- SP (Strict-Priority, 严格优先级)调度，严格按照队列 ID 进行调度。即每次发送报文之前，先检查高优先级队列中是否有报文待发送，如果有则发送；如果没有则检查下一优先级队列中是否有待发送报文，以此类推；
- WRR(Weighted Round Robin, 加权循环队列算法)调度，在队列之间进行轮流调度，保证每个队列都得到一定的服务时间。以 1 个 1000Mbps 端口 8 个输出队列为例，WRR 可为每个队列配置一个加权值（5、5、10、20、20、10、20、10，加权值表示获取资源的比重）。这样可以保证最低优先级队列至少获得 50Mbps 的带宽，避免了采用 SP 调度时低优先级队列中的报文可能长时间得不到服务的缺点；
- DRR(Dificit Round Robin, 差额循环队列算法)调度，DRR 和 WRR 类似，不过不是按照时间片，而是按照 byte 数来应用权重；
- WFQ(Weighted Fair Queueing, 加权公平队列算法)调度，提供了动态的、公平的排队方式，与 DRR 类似也是按照 byte 数来应用权重；区别在于 DRR 算法如果遇到一个空队列，会立即移到下一个队列进行传输，如果队列错过了它的传输时刻就只能等到下一个时刻才能传输，WFQ 不受这种影响，而且比 DRR 更适合于处理变长数据包；
- SP+WRR 调度，即将一个或多个发送队列配置成 SP，其他队列以 WRR 方式调度；SP 队列间，只有高优先级 SP 队列报文发送完了，才发送下一优先级 SP 队列中的报文；SP 与 WRR 调度队列间，只有所有 SP 队列报文发送完成后，才会处理 WRR 调度队列报文。

- SP+DRR 调度，即将一个或多个发送队列配置成 SP，其他队列以 DRR 方式调度；SP 队列间，只有高优先级 SP 队列报文发送完了，才发送下一优先级 SP 队列中的报文；SP 与 DRR 调度队列间，只有所有 SP 队列报文发送完成后，才会处理 DRR 调度队列报文。
- SP+WFQ 调度，即将一个或多个发送队列配置成 SP，其他队列以 WFQ 方式调度；SP 队列间，只有高优先级 SP 队列报文发送完了，才发送下一优先级 SP 队列中的报文；SP 与 WFQ 调度队列间，只有所有 SP 队列报文发送完成后，才会处理 WFQ 调度队列报文。

▾ 队列带宽

在一些产品上，允许配置队列的最小保证带宽与最大限制带宽；配置了最小保证带宽的队列可以保证此队列的带宽不小于配置值；配置了最大限制带宽的队列可以限制此队列的带宽不超过配置值，丢弃超过最大限制带宽的报文。

相关配置

▾ 配置 CoS-to-Queue Map

缺省情况下，CoS 值 0 1 2 3 4 5 6 7 分别映射到队列 1 2 3 4 5 6 7 8。

使用 `priority-queue cos-map` 命令，来配置 CoS 到队列的映射，其中 CoS 的取值范围为 0~7，队列的取值范围为 1~8。

▾ 配置输出队列调度策略

缺省情况下，输出队列的调度策略为 WRR。

使用 `mls qos scheduler` 命令，来配置队列的输出调度策略，可配置的调度策略有 SP、WRR、DRR、WFQ；或者使用 `priority-queue` 命令将调度策略配置为 SP。

▾ 配置 WRR 输出队列调度策略的轮转权重

缺省情况下，队列权重为 1:1:1:1:1:1:1:1。

使用 `wrr-queue bandwidth` 命令，来配置 WRR 输出队列调度策略的轮转权重，可配置权重范围由产品决定。

权重越大，所获得的输出时间就越多。

▾ 配置 DRR 输出队列调度策略的轮转权重

缺省情况下，队列权重为 1:1:1:1:1:1:1:1。

使用 `drr-queue bandwidth` 命令，来配置 DRR 输出队列调度策略的轮转权重，可配置权重范围由产品决定。

权重越大，所能发送的报文 bytes 就越多。

▾ 配置 WFQ 输出队列调度策略的轮转权重

缺省情况下，队列权重为 1:1:1:1:1:1:1:1。

使用 `wfq-queue bandwidth` 命令，来配置 WFQ 输出队列调度策略的轮转权重，可配置权重范围由产品决定。

权重越大，所能发送的报文 bytes 就越多。

▾ 配置队列带宽

使用 `qos queue` 来配置各个队列的最小保证带宽与最大限制带宽。队列的取值范围为 1~8，最小保证带宽、最大限制带宽的取值范围由产品决定。

2.3.5 拥塞避免

拥塞避免通过监控出端口队列的使用情况，在网络拥塞时，采取主动丢弃报文，调整网络流量的方式来解除网络过载。

工作原理

拥塞避免通过有效监控网络流量负载预期拥塞的发生，通过丢弃报文达到避免拥塞的目的，丢弃策略有尾部丢弃 (Tail-Drop)、RED (Random Early Detection, 早期随机检测) 丢弃、WRED (Weighted Random Early Detection, 加权随机早期检测) 丢弃：

▾ 尾部丢弃

传统的丢包策略采用尾部丢弃的方法。尾部丢弃对所有的流量都起作用，它并不能区分不同服务级别。在拥塞发生期间，队列尾部的数据包将被丢弃，直到拥塞解决。

▾ RED 与 WRED

运行 TCP 协议的主机会采用降低报文发送速率的方法来响应大量丢包的情况，当拥塞得到解决后，再提高数据包的发送速率。这样一来，尾部丢弃可能会引发 TCP 全局同步 (Global Synchronization) ——当队列同时丢弃多个 TCP 报文时，造成多个 TCP 连接同时进入拥塞避免和慢启动状态，同时降低并调整流量，而后又会在拥塞减少时出现流量高峰，如此反复，使网络流量忽大忽小，线路流量总在极少和饱满之间波动。当 TCP 同步发生时，连接的带宽不能充分利用，从而造成了带宽的浪费。

为了避免这种情况的发生，可以采用 RED/WRED 的报文丢弃策略，它提供了随机丢弃报文的机制，避免了 TCP 的全局同步现象。使得某个 TCP 连接的报文被丢弃，开始减速发送的时候，其他的 TCP 连接仍然有较高的发送速度。这样，无论什么时候，总有 TCP 连接在进行较快的发送，提高了线路带宽的利用率。

采用 WRED 时，用户可以设定队列的低门阈值与最大丢弃概率。当队列的长度小于低门阈值 (取值范围为 1~100) 时，不丢弃报文；当队列的长度在低门阈值和高门阈值 (固定为 100) 之间时，WRED 开始随机丢弃报文 (队列的长度越长，丢弃的概率越高，有个最大丢弃概率)；当队列的长度大于高门阈值时，以最大丢弃概率丢弃报文。

RED 与 WRED 的区别是后者引入优先权来区别丢弃策略，RED 作为 WRED 的特例，只有当接口上所有的 CoS 都映射到同一个低限和高限时，这时 WRED 就成为了 RED。

相关配置

▾ 开启 WRED 功能

缺省情况下，报文丢弃策略为尾部丢弃。

使用 `queueing wred` 命令，来开启 WRED 功能。

▾ 配置低门阈值

缺省情况下，支持 2 组低门阈值时，缺省值为 100，80 (阈值组数依产品而定)。

在接口配置模式下，使用 `wrr-queue random-detect min-threshold` 命令，来配置各个队列的 WRED 丢弃的低门阈值，队列的取值范围为 1~8，低门阈值的取值范围为 1~100。

当队列的长度小于低门阈值时，不丢弃报文；当队列的长度在低门阈值和高门阈值之间时，WRED 开始随机丢弃报文。

配置最大丢弃概率

缺省情况下，支持 2 组最大丢弃概率时，缺省值为 100，80（阈值组数依产品而定）。

在接口配置模式下，使用 `wrr-queue random-detect probability` 命令，来配置各个队列的 WRED 丢弃的最大丢弃概率，队列的取值范围为 1~8，最大丢弃概率的取值范围为 1~100。

当队列的长度在低门阈值和高门阈值之间时，WRED 开始随机丢弃报文，队列的长度越长，丢弃的概率越高，最大不超过最大丢弃概率；当队列的长度大于高门阈值时，以最大丢弃概率丢弃报文。

配置 CoS 与门阈值的映射

缺省情况下，所有的 CoS 都映射到第一组阈值（阈值组数依产品而定）。

在接口配置模式下，使用 `wrr-queue cos-map` 命令，来配置 CoS 到阈值组的映射，CoS 的取值范围为 0~7，阈值组数由产品决定。低门阈值、最大丢弃概率都可以配置多组，通过配置 CoS 到阈值组的映射，可以选择此 CoS 所对应的生效阈值组，比如 CoS 0 映射到第一组阈值，CoS 1 映射到第二组阈值，如果 CoS 0 和 1 的报文都进入队列 1 调度，此时 CoS 0 的报文使用第一组的低门阈值与最大丢弃概率进行处理；CoS 1 的报文使用第二组低门阈值与最大丢弃概率进行处理。

当接口上所有的 CoS 都映射到同一组阈值时，这时启用的 WRED 就成为了 RED。

2.4 产品说明



不支持在 SVI 口上配置 QOS 信任模式



- 对于 S7508 产品来说，逻辑端口组的成员必须在同一张线卡内。
- 加入逻辑端口组的成员必须是物理口或者是 Aggregate Port



CLASS MAP 所匹配的 ACL 表项中的 DENY 行为表项将被忽略，不会起作用



S7508 产品不支持修改 vid 的动作



- S7508 产品，应用在出口时，对于带宽超限部分的报文改写 DSCP 值，但不修改对应的 COS 值；应用在入口时，修改超限的 dscp 值，会改对应的 cos 值
- S7508 产品，支持改写带宽超限部分的报文的 CoS 值，改写 CoS 值时同步修改对应的 DSCP 值，设置 none-tos 选项后，改写 CoS 值时不修改对应的 DSCP 值。

- S7508 产品配置 set cos 时，不支持 none-tos 选项。
- 在 S7508 产品中的带宽限制指的是实际带宽，包含前导码和帧间隙所占去的负荷。（每个报文所附带的前导码和帧间隙所占的带宽为 20 字节）。
- 对于 S7508 的所有系列线卡，均支持最小限速粒度为 8Kbps，根据限速值设置的不同，最终会得到不同的限速粒度，具体的限速值和粒度值的关系大致如下表所示：

限速范围	64Kbps-2Gbps	2Gbps-4Gbps	4Gbps-8Gbps
粒度	8Kbps	16Kbps	32Kbps
限速范围	8Gbps- 6Gbps	16Gbps-32Gbps	32Gbps-40Gbps
粒度	64Kbps	128Kbps	256Kbps

- 对于 Qos 策略限速的第二个参数(突发流量)，在存在突发流量的情况下，若参数值设置过小，会导致实际速率可能过小；若参数值设置过大，会导致实际速率可能过大。用户可根据实际情况，使用如下推荐配置：
 - 1) 配置的限速值小于 1024Kbps 时，建议 burst-size 配置为 1024KByte。
 - 2) 配置的限速值小于 10240Kbps 时，建议 burst-size 配置和限速值相同或使用最大值（各产品可能允许的 burst-size 最大值低于 10240KByte）。
 - 3) 配置的限速值大于 10240Kbps 时，建议 burst-size 配置为该设备允许最大值。
- 对于 Qos 策略限速的第二个参数，当限速的速率比较大的时候，第二参数也要进行相应的调整，不然限速可能不准确。用户可以使用如下推荐配置：
 - 1) 对于万兆端口或 40G 端口，建议使用 burst-size 为 32 或 32 以上。



- S7508 交换机可以支持 policy map 应用到 out 方向。
- 对 S7508，当为 Aggregate Port 口应用 police 时，要求 AP 成员口必须满足以下条件时，设置的限制带宽才是 Aggregate Port 所有成员口的共享带宽。
 - 1) 要求 Aggregate Port 的成员口必须全部属于同一张线卡。
- S7508 由于 class map 需要关联 acl，所以 acl 配置的所有限制均适用于 qos，具体请参考 acl 配置指南。
- 不支持在 SVI 口上应用 Policy Maps。
- 支持输出方向的 Policy Maps，但不支持 AP 口。
- 输出方向的 Policy Maps 中不支持重标记报文的 CoS 值。重标记报文的 DSCP 值时，不会同时标记报文的 CoS 值。
- 目前 output 方向应用在逻辑端口组上未被支持。由于 class map 需要关联 acl，所以 acl 配置的所有限制均适用于 qos，具体请参考 acl 配置指南。



- S7508 线卡不支持 WFQ 及 SP+WFQ 调度。
- 对于 S7508 线卡，VSL 口默认采用 SP+DRR 调度算法，其中队列 7 采用 sp 调度，其他队列权重都是 1，该调度配置不会被用户配置所改变。



- 在 S7508 产品中的端口速率的限制指的是实际带宽，包含前导码和帧间隙所占去的负荷。（每个报文所附带的前导码和帧间隙所占的带宽为 20 字节）。
- 对于 S7508 的所有系列线卡，均支持最小限速粒度为 8Kbps，根据限速值设置的不同，最终会得到不同的限速粒度，具体的限速值和粒度值的关系大致如下表所示：

限速范围	64Kbps-2Gbps	2Gbps-4Gbps	4Gbps-8Gbps
粒度	8Kbps	16Kbps	32Kbps
限速范围	8Gbps-16Gbps	16Gbps-32Gbps	32Gbps-40Gbps
粒度	64Kbps	128Kbps	256Kbps

- 对于端口限速的第二个参数，当限速的速率比较大的时候，第二参数也要进行相应的调整，不然限速可能不准确。用户可以使用如下推荐配置：
 - 1) 对于万兆端口或 40G 端口，建议使用 burst-size 为 32 或 32 以上。



当前只能通过 **show run** 来查看 WRED 全局功能是否已经开启。



- 队列上配置低门阈值和最大丢弃概率称为一组 wred 配置，S7508 线卡支持 120 组 wred 配置。
- 不建议用户配置超过上述组数的 wred 配置，多配置的 wred 可以不能正常工作。
- 当低门阈值为 100%时，表示禁用 WRED 功能。



- S7508 所有系列线卡支持物理口上配置映射关系。
- 管理员可以通过配置 DSCP-CoS 和 CoS-Threshold 的映射关系来实现 DSCP 与 threshold 的映射
- 管理员可以通过配置 CoS-Threshold 和 CoS-Queue 的映射关系来实现 Queue 与 threshold 的映射



在 VSU 模式下，管理报文的优先级默认为 7，进入队列 8，不建议更改这个映射关系。

2.5 配置详解

配置项	配置建议 & 相关命令	
配置流分类	可选配置。用于创建流分类信息。	
	class-map	创建类
	match access-group	匹配 ACL 规则
	match ip precedence	匹配 IP 报文 PRE 优先级
	match ip dscp	匹配 IP 报文 DSCP 优先级
	policy-map	创建策略

	class	关联类
	police	绑定流的带宽限制与超限后的报文处理行为
	set	绑定修改流的 CoS、DSCP 的行为
	virtual-group	创建/接口加入逻辑端口组
	service-policy	在接口上应用策略
配置报文优先级标记与映射	 可选配置。用于配置接口信任模式、缺省 CoS、各种映射关系。	
	mls qos trust	修改接口的信任模式
	mls qos cos	修改接口缺省 CoS 值
	mls qos map cos-dscp	配置 CoS 到 DSCP 的映射
	mls qos map dscp-cos	配置 DSCP 到 CoS 的映射
	mls qos map ip-precedence-dscp	配置 IP PRE 到 DSCP 的映射
配置端口限速	 可选配置。配置端口的限速。	
	rate-limit	配置端口流量限制
配置拥塞管理	 可选配置。配置 CoS 到队列映射，队列调度策略与轮转权重。	
	priority-queue cos-map	配置 CoS 到队列映射
	priority-queue	配置队列的输出调度策略为 SP
	mls qos scheduler	配置队列的输出调度策略
	wrr-queue bandwidth	配置 WRR 输出队列调度策略的轮转权重
	drp-queue bandwidth	配置 DRR 输出队列调度策略的轮转权重
	wfq-queue bandwidth	配置 WFQ 输出队列调度策略的轮转权重
配置拥塞避免	 可选配置。通过设置报文丢弃的方式来避免网络拥塞。	
	queueing wred	开启 WRED 功能
	wrr-queue random-detect min-threshold	配置 WRED 丢弃的低门阈值
	wrr-queue random-detect probability	配置 WRED 丢弃的最大丢弃概率
	wrr-queue cos-map	配置 threshold 到 CoS 的映射

2.5.1 配置流分类

配置效果

- 创建类，匹配分类规则。
- 创建策略，绑定类与流行为。并关联到接口上。

注意事项

- 类与策略的名称不能超过 31 个字符。
- 接口上的配置只支持在 AP 口和以太网口上配置。

配置方法

↘ 创建类，匹配规则

- 可选配置。
- 创建类，在类配置模式下，匹配 ACL、IP PRE、DSCP 中的一个。

↘ 创建策略

- 可选配置。
- 创建策略，在策略配置模式下，绑定类与流行为。

↘ 创建并将接口加入逻辑端口组

- 可选配置。
- 创建逻辑端口组，将接口加入逻辑端口组。

↘ 配置接口上应用策略

- 可选配置。
- 将配置好的策略关联到指定的接口或逻辑端口组上。

检验方法

- 使用 **show class-map** 命令，可以查看类是否创建成功，规则是否匹配成功。
- 使用 **show policy-map** 命令，可以查看策略是否创建成功，类与流行为是否绑定成功。
- 使用 **show mls qos interface** 命令，可以查看接口上是否关联策略。
- 使用 **show virtual-group** 命令，可以查看逻辑接口组下的接口。
- 使用 **show mls qos virtual-group** 命令，可以查看逻辑接口组上是否关联策略。

相关命令

↘ 创建类

- 【命令格式】 **class-map** *class-map-name*
- 【参数说明】 *class-map-name*：要创建的类的名字，名称不能超过 31 个字符。
- 【命令模式】 全局模式
- 【使用指导】 -

↘ 匹配 ACL

- 【命令格式】 **match access-group** *access-list-number*
- 【参数说明】 *access-list-number* : 要匹配的访问控制列表编号。
- 【命令模式】 类配置模式
- 【使用指导】 -

↘ 匹配 IP 报文的 PRE

- 【命令格式】 **match ip precedence** *precedence-value... [precedence-value...]*
- 【参数说明】 *precedence -value* : 要匹配的 IP PRE , 取值范围为 0~7。
- 【命令模式】 类配置模式
- 【使用指导】 -

↘ 匹配 IP 报文的 DSCP

- 【命令格式】 **match ip dscp** *dscp-value... [dscp-value...]*
- 【参数说明】 *dscp -value* : 要匹配的 DSCP , 取值范围为 0~63。
- 【命令模式】 类配置模式
- 【使用指导】 -

↘ 创建策略

- 【命令格式】 **policy-map** *policy-map-name*
- 【参数说明】 *policy-map-name* : 要创建的策略的名字, 名称不能超过 31 个字符。
- 【命令模式】 全局模式
- 【使用指导】 -

↘ 关联类

- 【命令格式】 **class** *class-map-name*
- 【参数说明】 *class-map-name* : 要关联的类名字。
- 【命令模式】 策略配置模式
- 【使用指导】 -

↘ 绑定修改流的 CoS、DSCP 的行为

- 【命令格式】 **set { ip dscp** *new-dscp* | **cos** *new-cos* }
- 【参数说明】 **ip dscp** *new-dscp* : 修改流的 DSCP 值为 *new-dscp* , 取值范围为 0~63。
cos *new-cos* : 修改流的 CoS 值为 *new-cos* , 取值范围为 0~7。
- 【命令模式】 类配置模式
- 【使用指导】 -

↘ 绑定流的带宽限制与超限后的报文处理行为

- 【命令格式】 **police** *rate-bps* *burst-byte* [**exceed-action** { **drop** | **dscp** *new-dscp* | **cos** *new-cos* [**none-tos**] }]
- 【参数说明】 *rate-bps* : 每秒钟带宽限制量(KBits) , 由产品决定取值范围。
burst-byte : 突发流量限制值(KBytes) , 由产品决定取值范围。
drop : 丢弃带宽超限部分的报文。

dscp *new-dscp* : 修改带宽超限部分报文的 DSCP 值为 *new-dscp* , 取值范围为 0~63。

cos *new-cos* : 修改带宽超限部分报文的 CoS 值为 *new-cos* , 取值范围为 0~7。

none-tos : 改变报文 CoS 值时, 不修改报文的 DSCP 值。

【命令模式】 类配置模式

【使用指导】 -

创建逻辑端口组/接口加入逻辑接口组

【命令格式】 **virtual-group** *virtual-group-number*

【参数说明】 *virtual-group-number* : 逻辑端口组号, 取值范围 1~128。

【命令模式】 全局模式下创建逻辑端口组/接口模式将接口加入逻辑端口组, 如果逻辑接口组不存在, 则先创建逻辑接口组, 再将接口加入。

【使用指导】 -

在接口上应用策略

【命令格式】 **service-policy** { **input** | **output** } *policy-map-name*

【参数说明】 **input** : 接口的输入方向;

output : 接口的输出方向;

policy-map-name : 要应用在接口上的策略名。

【命令模式】 接口模式

【使用指导】 -

配置举例

创建 4 个流分类, 分别匹配 ACL、IP PRE、DSCP。

- 【配置方法】
- 创建 ACL 规则
 - 创建 4 个流分类, 分别匹配 ACL、DSCP、IP PRE

```
Ruijie# configure terminal
Ruijie(config)# access-list 11 permit host 192.168.23.61
Ruijie(config)# class-map cmap1
Ruijie(config-cmap)# match access-group 11
Ruijie(config-cmap)# exit
Ruijie(config)# class-map cmap2
Ruijie(config-cmap)# match ip dscp 21
Ruijie(config-cmap)# exit
Ruijie(config)# class-map cmap3
Ruijie(config-cmap)# match ip precedence 5
Ruijie(config-cmap)# exit
```

- 【检验方法】
- 检查创建的 ACL 规则、流分类规则是否成功。

```
Ruijie# show access-lists
```

```
ip access-list standard 11
10 permit host 192.168.23.61
```

```
Ruijie# show class-map
Class Map cmap1
  Match access-group 11
Class Map cmap2
  Match ip dscp 21
Class Map cmap3
  Match ip precedence 5
```

📌 创建策略，绑定类与流行为，并关联到接口上。

- 【配置方法】
- 创建流分类 cmap1，匹配 DSCP 值为 18 的报文；创建 cmap2，匹配 IP PRE 为 7 的报文；
 - 创建策略 pmap1，关联 cmap1，绑定其行为为修改流的 CoS 为 6；关联 cmap2，绑定其行为为修改流的 DSCP 为 15，并限制其每秒流量为 10000 KBits，触发流量为每秒 1024 KBits，修改超限流量的 DSCP 值为 7；
 - 将策略 pmap1 应用在接口 gigabitEthernet 0/0 的出口上；
 - 创建虚拟逻辑组 1，并将接口 gigabitEthernet 0/1、gigabitEthernet 0/2 加入，将策略 pmap1 应用在虚拟逻辑组的入口上。

```
Ruijie# configure terminal
Ruijie(config)# class-map cmap1
Ruijie(config-cmap)# match ip dscp 18
Ruijie(config-cmap)# exit
Ruijie(config)# class-map cmap2
Ruijie(config-cmap)# match ip precedence 7
Ruijie(config-cmap)# exit
Ruijie(config)# policy-map pmap1
Ruijie(config-pmap)# class cmap1
Ruijie(config-pmap-c)# set cos 6
Ruijie(config-pmap-c)# exit
Ruijie(config-cmap)# class cmap2
Ruijie(config-pmap-c)# set ip dscp 15
Ruijie(config-pmap-c)# police 10000 1024 exceed-action dscp 7
Ruijie(config-pmap-c)# exit
Ruijie(config-pmap)# exit
Ruijie(config)# interface gigabitEthernet 0/0
Ruijie(config-if-GigabitEthernet 0/0)# service-policy output pmap1
Ruijie(config-if-GigabitEthernet 0/0)# exit
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# virtual-group 1
Ruijie(config-if-GigabitEthernet 0/1)# exit
Ruijie(config)# interface gigabitEthernet 0/2
```

```
Ruijie(config-if-GigabitEthernet 0/2)# virtual-group 1
Ruijie(config-if-GigabitEthernet 0/2)# exit
Ruijie(config)# virtual-group 1
Ruijie(config-VirtualGroup)# service-policy input pmap1
Ruijie(config-VirtualGroup)# exit
```

【检验方法】

- 检查流分类规则是否创建成功；
- 检查策略是否创建成功，并成功绑定流与流行为；
- 检查策略是否应用到接口上；
- 检查逻辑接口组是否创建成功，关联上接口，并成功应用上策略。

```
Ruijie# show class-map
Class Map cmap1
  Match ip dscp 18
Class Map cmap2
  Match ip precedence 7
```

```
Ruijie# show policy-map
Policy Map pmap1
  Class cmap1
    set cos 6
  Class cmap2
    set ip dscp 15
    police 10000 1024 exceed-action dscp 7
```

```
Ruijie# show mls qos interface gigabitEthernet 0/0
Interface: GigabitEthernet 0/0
Ratelimit input:
Ratelimit output:
Attached input policy-map:
Attached output policy-map: pmap1
Default trust: none
Default cos: 0
```

```
Ruijie# show virtual-group 1
virtual-group      member
-----
1                  Gi0/1 Gi0/2
Ruijie# show mls qos virtual-group 1
Virtual-group: 1
Attached input policy-map: pmap1
```

2.5.2 配置报文优先级标记与映射

配置效果

- 配置接口的信任模式，默认 CoS 值。
- 配置 CoS-to-DSCP、DSCP-to-CoS、IP-PRE-to-DSCP 映射关系。

注意事项

- 接口上的配置只支持在 AP 口和以太网口上配置。

配置方法

▾ 配置接口的信任模式、默认 CoS 值

- 可选配置。
- 在接口模式下，可配置接口的信任模式与默认 CoS 值。

▾ 配置 CoS-to-DSCP、DSCP-to-CoS、IP-PRE-to-DSCP 映射关系

- 可选配置。
- 配置各种映射关系。

检验方法

- 使用 `show mls qos interface` 命令，可以查看接口的信任模式、默认 CoS 值。
- 使用 `show mls qos maps` 命令，可以查看 CoS-to-DSCP、DSCP-to-CoS、IP-PRE-to-DSCP 映射关系。

相关命令

▾ 配置接口的信任模式

【命令格式】 `mls qos trust { cos | ip-precedence | dscp }`

【参数说明】 `cos`：配置接口的信任模式为 CoS；
`ip-precedence`：配置接口的信任模式为 IP PRE；
`dscp`：配置接口的信任模式为 DSCP。

【命令模式】 接口模式

【使用指导】 -

▾ 配置接口的缺省 CoS 值

【命令格式】 `mls qos cos default-cos`

【参数说明】 *default-cos* : 所要配置的缺省 CoS 值, 默认值为 0, 取值范围为 0~7。

【命令模式】 接口模式

【使用指导】 -

配置 CoS-to-DSCP MAP

【命令格式】 **mls qos map cos-dscp** *dscp1...dscp8*

【参数说明】 *dscp1...dscp8* : CoS 所映射的 DSCP 值, 缺省 CoS 0~7 分别映射到 DSCP 0 8 16 24 32 40 48 56, DSCP 取值范围为 0~63。

【命令模式】 全局模式

【使用指导】 -

配置 DSCP-to-CoS MAP

【命令格式】 **mls qos map dscp-cos** *dscp-list to cos*

【参数说明】 *dscp-list* : 要映射到 CoS 的 DSCP 列表, 缺省 DSCP 0~7 映射到 CoS 0, DSCP 8~15 映射到 CoS 1, DSCP 16~23 映射到 CoS 2, DSCP 24~31 映射到 CoS 3, DSCP 32~39 映射到 CoS 4, DSCP 40~47 映射到 CoS 5, DSCP 48~55 映射到 CoS 6, DSCP 56~63 映射到 CoS 7, DSCP 取值范围为 0~63。

cos : *dscp-list* 所要映射到的 CoS, 取值范围为 0~7。

【命令模式】 全局模式

【使用指导】 -

配置 IP-PRE-to-DSCP MAP

【命令格式】 **mls qos map ip-prec-dscp** *dscp1...dscp8*

【参数说明】 *dscp1...dscp8* : IP PRE 所映射的 DSCP 值, 缺省 IP PRE 0~7 分别映射到 DSCP 0 8 16 24 32 40 48 56, DSCP 取值范围为 0~63。

【命令模式】 全局模式

【使用指导】 -

配置举例

配置接口的信任模式, 默认 CoS 值。

- 【配置方法】
- 修改接口 gigabitEthernet 0/0 的信任模式为信任 DSCP ;
 - 修改接口 gigabitEthernet 0/1 的默认 CoS 值为 7。

```
Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet 0/0
Ruijie(config-if-GigabitEthernet 0/0)# mls qos trust dscp
Ruijie(config-if-GigabitEthernet 0/0)# exit
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# mls qos cos 7
Ruijie(config-if-GigabitEthernet 0/1)# exit
```

- 【检验方法】
- 检查接口配置, 信任模式与默认 CoS 是否配置成功。

```
Ruijie# show mls qos interface gigabitEthernet 0/0
Interface: GigabitEthernet 0/0
Ratelimit input:
Ratelimit output:
Attached input policy-map:
Attached output policy-map:
Default trust: dscp
Default cos: 0
Ruijie# show mls qos interface gigabitEthernet 0/1
Interface: GigabitEthernet 0/1
Ratelimit input:
Ratelimit output:
Attached input policy-map:
Attached output policy-map:
Default trust: none
Default cos: 7
```

配置 CoS-to-DSCP、DSCP-to-CoS、IP-PRE-to-DSCP 映射关系。

- 【配置方法】
- 配置 CoS-to-DSCP 将 CoS 0 1 2 3 4 5 6 7 分别映射到 DSCP 7 14 21 28 35 42 49 56 ；
 - 配置 DSCP-to-CoS 将 DSCP 0 1 2 3 4 映射到 CoS 4 ，将 DSCP 11 12 13 14 映射到 CoS 7 ；
 - 配置 IP-PRE-to-DSCP 将 IP PRE 0 1 2 3 4 5 6 7 分别映射到 DSCP 31 26 21 15 19 45 47 61 ；

```
Ruijie# configure terminal
Ruijie(config)# mls qos map cos-dscp 7 14 21 28 35 42 49 56
Ruijie(config)# mls qos map dscp-cos 0 1 2 3 4 to 4
Ruijie(config)# mls qos map dscp-cos 11 12 13 14 to 7
Ruijie(config)# mls qos map ip-precedence-dscp 31 26 21 15 19 45 47 61
```

- 【检验方法】
- 检查各映射关系是否配置成功。

```
Ruijie# show mls qos maps cos-dscp
cos dscp
---- ----
0 7
1 14
2 21
3 28
4 35
5 42
6 49
7 56
Ruijie# show mls qos maps dscp-cos
dscp cos      dscp cos      dscp cos      dscp cos
```

```

-----
0 4      1 4      2 4      3 4
4 4      5 0      6 0      7 0
8 1      9 1     10 1     11 7
12 7     13 7     14 7     15 1
16 2     17 2     18 2     19 2
20 2     21 2     22 2     23 2
24 3     25 3     26 3     27 3
28 3     29 3     30 3     31 3
32 4     33 4     34 4     35 4
36 4     37 4     38 4     39 4
40 5     41 5     42 5     43 5
44 5     45 5     46 5     47 5
48 6     49 6     50 6     51 6
52 6     53 6     54 6     55 6
56 7     57 7     58 7     59 7
60 7     61 7     62 7     63 7

Ruijie# show mls qos maps ip-prec-dscp
ip-precedence dscp
-----
0 31
1 26
2 21
3 15
4 19
5 45
6 47
7 61

```

2.5.3 配置端口限速

配置效果

- 配置端口的流量限制。

注意事项

- 只支持在以太网口上配置。

配置方法

配置端口的流量限制

- 可选配置。
- 可以配置端口上允许通过的流量与突发流量的限制值。

检验方法

- 使用 `show mls qos rate-limit` 命令，可以查看端口的限速信息

相关命令

配置端口的流量限制

【命令格式】 `rate-limit { input | output } bps burst-size`

【参数说明】 **input**：接口的输入方向。

output：接口的输出方向。

bps：每秒钟的带宽限制量(KBits)，取值范围依产品而定。

burst-size：突发流量限制值(Kbytes)，取值范围依产品而定。

【命令模式】 接口模式

【使用指导】 -

配置举例

以典型应用---端口限速+优先级重标记应用为例。

- 【配置方法】
- 对于出口访问 Internet，在端口 G0/24 上配置端口的出口流量限制，带宽限制每秒 102400KBits，突发流量限制每秒 256Kbytes；
 - 对于宿舍楼，在端口 G0/3 上配置端口的入口流量限制，带宽限制每秒 51200KBits，突发流量限制每秒 256Kbytes；
 - 对于教学楼，在端口 G0/1 上配置端口的入口流量限制，带宽限制每秒 30720KBits，突发流量限制每秒 256Kbytes；
 - 对于实验楼，创建类 `cmap_dscp7` 匹配 DSCP 优先级 7，创建策略 `pmap_shiyan`，关联 `cmap_dscp7`，绑定流行为为将速度超过 20M 的报文的 DSCP 值改为 16；将 `pmap_shiyan` 应用在接口 G0/2 上，并配置接口信任 DSCP。

```
Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet 0/24
Ruijie(config-if-GigabitEthernet 0/24)# rate-limit output 102400 256
Ruijie(config-if-GigabitEthernet 0/24)# exit
Ruijie(config)# interface gigabitEthernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# rate-limit input 51200 256
Ruijie(config-if-GigabitEthernet 0/3)# exit
Ruijie(config)# interface gigabitEthernet 0/1
```

```
Ruijie(config-if-GigabitEthernet 0/1)# rate-limit input 30720 256
Ruijie(config-if-GigabitEthernet 0/1)# exit
Ruijie(config)# class-map cmap_dscp7
Ruijie(config-cmap)# match ip dscp 7
Ruijie(config-cmap)# exit
Ruijie(config)# policy-map pmap_shiyan
Ruijie(config-pmap)# class cmap_dscp7
Ruijie(config-pmap-c)# police 20480 128 exceed-action dscp 16
Ruijie(config-pmap-c)# exit
Ruijie(config-pmap)# exit
Ruijie(config)# interface gigabitEthernet 0/2
Ruijie(config-if-GigabitEthernet 0/2)# service-policy input pmap_shiyan
Ruijie(config-if-GigabitEthernet 0/2)# mls qos trust dscp
Ruijie(config-if-GigabitEthernet 0/2)# exit
```

- 【检验方法】
- 检查接口限速配置是否成功;
 - 检查类与策略是否创建成功，并成功应用在接口上。

```
Ruijie# show mls qos rate-limit
Interface: GigabitEthernet 0/1
  rate limit input Kbps = 30720 burst = 256
Interface: GigabitEthernet 0/3
  rate limit input Kbps = 51200 burst = 256
Interface: GigabitEthernet 0/24
  rate limit output Kbps = 102400 burst = 256
Ruijie# show class-map cmap_dscp7

Class Map cmap_dscp7
  Match ip dscp 7
Ruijie# show policy-map pmap_shiyan

Policy Map pmap_shiyan
  Class cmap_dscp7
    police 20480 128 exceed-action dscp 16
Ruijie# show mls qos interface gigabitEthernet 0/2
Interface: GigabitEthernet 0/2
Ratelimit input:
Ratelimit output:
Attached input  policy-map: pmap_shiyan
Attached output policy-map:
Default trust: dscp
Default cos: 0
```


【命令模式】 全局模式

【使用指导】 -

配置输出队列调度策略为 SP

【命令格式】 **priority-queue**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

配置输出队列调度策略

【命令格式】 **mls qos scheduler { sp | wrr | drr | wfq }**

【参数说明】 **sp**：设置输出队列调度算法为绝对优先级调度；
wrr：设置输出队列调度算法为带帧数量权重轮转调度；
drr：设置输出队列调度算法为带帧长度权重轮转调度；
wfq：设置输出队列调度算法为加权公平队列调度。

【命令模式】 全局模式

【使用指导】 -

配置输出队列调度策略轮转权重

【命令格式】 **{ drr-queue | wrr-queue | wfq-queue } bandwidth weight1...weight8**

【参数说明】 **drr-queue**：DRR 输出队列调度策略的轮转权重；
wrr-queue：WRR 输出队列调度策略的轮转权重；
wfq-queue：WFQ 输出队列调度策略的轮转权重。
weight1...weight8：输出队列 1~8 的权重，取值范围依产品而定，取值为 0 表示该队列采用 SP 调度，缺省各队列权重皆为 1:1。

【命令模式】 全局模式

【使用指导】 -

配置队列的最小保证带宽与最大限制带宽

【命令格式】 **qos queue queue-id bandwidth { minimum | maximum } bandwidth**

【参数说明】 **queue-id**：所要配置的队列 ID，取值范围为 1~8。
minimum bandwidth：最小保证带宽 Kbps，取值范围依产品而定，缺省没有配置。
maximum bandwidth：最大限制带宽 Kbps，取值范围依产品而定，缺省没有配置。

【命令模式】 接口模式

【使用指导】 -

配置举例

配置 CoS 到队列映射，修改调度策略与其轮转权重。

- 【配置方法】
- 配置 CoS 到队列的映射为 CoS 值 0 1 2 3 4 5 6 7 分别映射到队列 1 2 5 5 5 5 7 8；
 - 配置队列的输出调度策略为 DRR，轮转权重为 2:1:1:1:6:6:6:8。

```
Ruijie# configure terminal
Ruijie(config)# priority-queue cos-map 5 2 3 4 5
Ruijie(config)# mls qos scheduler drr
Ruijie(config)# drr-queue bandwidth 2 1 1 1 6 6 6 8
```

- 【检验方法】**
- 检测 CoS 到队列是否映射成功，队列输出调度策略与轮转权重是否配置成功。

```
Ruijie# show mls qos scheduler
Global Multi-Layer Switching scheduling
  Deficit Round Robin
Ruijie# show mls qos queueing
Cos-queue map:
cos  qid
----  ---
0    1
1    2
2    5
3    5
4    5
5    5
6    7
7    8

wrr bandwidth weights:
qid  weights
----  -
1    1
2    1
3    1
4    1
5    1
6    1
7    1
8    1

drr bandwidth weights:
qid  weights
----  -
1    2
2    1
3    1
4    1
```

```

5 6
6 6
7 6
8 8

wfq bandwidth weights:
qid weights
-----
1 1
2 1
3 1
4 1
5 1
6 1
7 1
8 1

```

配置队列的最小保证带宽与最大限制带宽。

- 【配置方法】
- 配置接口 gigabitEthernet 0/1 上队列 1 的最大限制带宽为 10M，最小保证带宽 5M；配置队列 2 的最小保证带宽为 2M。

```

Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# qos queue 1 bandwidth maximum 10240
Ruijie(config-if-GigabitEthernet 0/1)# qos queue 1 bandwidth minimum 5120
Ruijie(config-if-GigabitEthernet 0/1)# qos queue 2 bandwidth minimum 2048
Ruijie(config-if-GigabitEthernet 0/1)# exit

```

- 【检验方法】
- 检查接口上最小保证带宽与最大带宽限制是否配置成功。

```

Ruijie# show qos bandwidth interface gigabitEthernet 0/1
Interface: GigabitEthernet 0/1
-----
queue-id | minimum-bandwidth | maximum-bandwidth
-----
1         |          5120     |          10240
2         |          2048     |              0
3         |           0       |              0
4         |           0       |              0
5         |           0       |              0
6         |           0       |              0
7         |           0       |              0
8         |           0       |              0

```

```
-----
Total minimum-bandwidth:          7168
Total maximum-bandwidth:         10240
```

▾ 以典型应用——优先级重标记+队列调度应用为例。

【配置方法】

- 创建访问各类服务器的 ACL，并创建类匹配这些 ACL。
- 创建策略，关联各个类，为访问各类服务器的报文重新指定 CoS。并将其关联到研发部和市场部的入口上，配置端口信任 CoS。
- 配置人事管理部端口的缺省 CoS 为最高优先级 7，优先保障人事部发出的报文。
- 配置队列的输出调度策略为 WRR，轮转权重为 1:1:1:2:6:1:1:0，即对人事管理部报文实行 SP 调度，对研发部与市场部访问邮件数据库、文件数据库、工资查询数据库的报文按照 6 : 2 : 1 的比例来调度

```
Ruijie# configure terminal
Ruijie(config)# ip access-list extended salary
Ruijie(config-ext-nacl)# permit ip any host 192.168.10.1
Ruijie(config-ext-nacl)# exit
Ruijie(config)# ip access-list extended mail
Ruijie(config-ext-nacl)# permit ip any host 192.168.10.2
Ruijie(config-ext-nacl)# exit
Ruijie(config)# ip access-list extended file
Ruijie(config-ext-nacl)# permit ip any host 192.168.10.3
Ruijie(config-ext-nacl)# exit
Ruijie(config)# class-map salary
Ruijie(config-cmap)# match access-group salary
Ruijie(config-cmap)# exit
Ruijie(config)# class-map mail
Ruijie(config-cmap)# match access-group mail
Ruijie(config-cmap)# exit
Ruijie(config)# class-map file
Ruijie(config-cmap)# match access-group file
Ruijie(config)# policy-map toserver
Ruijie(config-pmap)# class mail
Ruijie(config-pmap-c)# set cos 4
Ruijie(config-pmap-c)# exit
Ruijie(config-pmap)# class file
Ruijie(config-pmap-c)# set cos 3
Ruijie(config-pmap-c)# exit
Ruijie(config-pmap)# class salary
Ruijie(config-pmap-c)# set cos 2
Ruijie(config-pmap-c)# end
Ruijie(config)# interface gigabitEthernet 0/1
Ruijie(config-if-GigabitEthernet 0/1)# service-policy input toserver
```

```
Ruijie(config-if-GigabitEthernet 0/1)# mls qos trust cos
Ruijie(config-if-GigabitEthernet 0/1)# exit
Ruijie(config)# interface gigabitEthernet 0/2
Ruijie(config-if-GigabitEthernet 0/2)# service-policy input toserver
Ruijie(config-if-GigabitEthernet 0/2)# mls qos trust cos
Ruijie(config-if-GigabitEthernet 0/2)# exit
Ruijie(config)# interface gigabitEthernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# mls qos cos 7
Ruijie(config)#wrr-queue bandwidth 1 1 1 2 6 1 1 0
Ruijie(config)#mls qos scheduler wrr
```

【检验方法】

- 检查 ACL 是否创建成功，类是否成功关联 ACL；
- 检查策略是否创建成功，类与流行为是否绑定成功，策略是否成功应用在接口上；
- 检查接口默认 CoS 是否配置成功，调度策略与轮转权重是否配置成功。

```
Ruijie# show access-lists

ip access-list extended file
 10 permit ip any host 192.168.10.3

ip access-list extended mail
 10 permit ip any host 192.168.10.2

ip access-list extended salary
 10 permit ip any host 192.168.10.1
Ruijie# show class-map

Class Map salary
  Match access-group salary
Class Map mail
  Match access-group mail
Class Map file
  Match access-group file
Ruijie# show policy-map

Policy Map toserver
  Class mail
    set cos 4
  Class file
    set cos 3
  Class salary
    set cos 2
```

```
Ruijie# show mls qos interface gigabitEthernet 0/1
Interface: GigabitEthernet 0/1
Ratelimit input:
Ratelimit output:
Attached input  policy-map: toserver
Attached output policy-map:
Default trust: cos
Default cos: 0
Ruijie# show mls qos interface gigabitEthernet 0/2
Interface: GigabitEthernet 0/2
Ratelimit input:
Ratelimit output:
Attached input  policy-map: toserver
Attached output policy-map:
Default trust: cos
Default cos: 0
Ruijie# show mls qos interface gigabitEthernet 0/3
Interface: GigabitEthernet 0/2
Ratelimit input:
Ratelimit output:
Attached input  policy-map:
Attached output policy-map:
Default trust: none
Default cos: 7
Ruijie# show mls qos scheduler
Global Multi-Layer Switching scheduling
  Weighted Round Robin
Ruijie# Ruijie#show mls qos queueing
Cos-queue map:
cos qid
----
0  1
1  2
2  3
3  4
4  5
5  6
6  7
7  8

wrr bandwidth weights:
```

```
qid weights
-----
1 1
2 1
3 1
4 2
5 6
6 1
7 1
8 0

drr bandwidth weights:
qid weights
-----
1 1
2 1
3 1
4 1
5 1
6 1
7 1
8 1

wfq bandwidth weights:
qid weights
-----
1 1
2 1
3 1
4 1
5 1
6 1
7 1
8 1
```

2.5.5 配置拥塞避免

配置效果

- 配置 WRED 的低门阈值，当队列中报文的长度小于低门阈值时，不丢弃报文。

- 配置最大丢弃概率，当队列中报文的长度在低门阈值和高门阈值之间时，随机丢弃报文，此项配置了丢弃的最大概率。
- 配置 CoS 值与门阈值的映射关系。

注意事项

- 接口上的配置只支持在 AP 口和以太网口上配置。

配置方法

▾ 开启 WRED 功能

- 可选配置。
- 如果需要 WRED 功能，则开启。

▾ 配置低门阈值

- 可选配置。
- 如需修改低门阈值，则配置。

▾ 配置最大丢弃概率

- 可选配置。
- 如需修改最大丢弃概率，则配置。

▾ 配置 CoS 与门阈值的映射

- 可选配置。
- 如需修改 CoS 与门阈值的映射关系，则配置。

检验方法

- 使用 `show queueing wred interface` 命令，查看 WRED 配置信息。

相关命令

▾ 开启 WRED 功能

- 【命令格式】 `queueing wred`
- 【参数说明】 -
- 【命令模式】 全局模式
- 【使用指导】 -

▾ 配置低门阈值

【命令格式】 **wrr-queue random-detect min-threshold** *queue_id thr1* [*thr2*]

【参数说明】 *queue_id*: 接口队列 ID, 取值范围为 1~8。
thrN: 支持 2 组低门阈值, 取值范围为 1~100。

【命令模式】 接口模式

【使用指导】 -

配置最大丢弃概率

【命令格式】 **wrr-queue random-detect probability** *queue_id prob1* [*prob2*]

【参数说明】 *queue_id*: 接口队列 ID, 取值范围为 1~8。
probN: 支持 2 组最大丢弃概率, 取值范围为 1~100。

【命令模式】 接口模式

【使用指导】 -

配置 CoS 与门阈值的映射关系

【命令格式】 **wrr-queue cos-map** *threshold_id cos1* [*cos2* [*cos3* [*cos4* [*cos5* [*cos6* [*cos7* [*cos8*]]]]]]]]]

【参数说明】 *threshold_id*: 阈值组 ID 号, 支持 2 组, 取值范围 1~2。
cos1...cos8: 要映射到阈值组上的 CoS 值, 取值范围为 0~7, 缺省情况所有的 CoS 都映射到第一组阈值。

【命令模式】 接口模式

【使用指导】 -

配置举例

开启 WRED 功能, 配置低门阈值, 最大丢弃概率, CoS 与门阈值的映射关系 (假设产品的阈值组数为 2)。

- 【配置方法】
- 开启 WRED 功能
 - 配置接口 gigabitEthernet 0/2 队列 2 的低门阈值为 10 20;
 - 配置接口 gigabitEthernet 0/2 队列 2 的最大丢弃概率 60 80;
 - 配置接口 gigabitEthernet 0/2 上 CoS 0 1 2 3 使用阈值组 2。

```
Ruijie# configure terminal
Ruijie(config)# queueing wred
Ruijie(config)# interface gigabitEthernet 0/2
Ruijie(config-if-GigabitEthernet 0/2)# wrr-queue random-detect min-threshold 2 10 20
Ruijie(config-if-GigabitEthernet 0/2)# wrr-queue random-detect probability 2 60 80
Ruijie(config-if-GigabitEthernet 0/2)# wrr-queue cos-map 2 0 1 2 3
```

- 【检验方法】
- 检测 WRED 功能是否开启, 门阈值是否配置成功, CoS 与门阈值映射是否配置成功。

```
Ruijie# show running-config

Building configuration...

Current configuration : 1654 bytes

version 11.0(1C2B1) (09/11/13 00:16:26 CST -ngcf78)
```

```

queueing wred

Ruijie# show queueing wred interface gigabitEthernet 0/2
-----
qid  max_1  min_1  prob_1  max_2  min_2  prob_2
-----
 1    0     0     0       1     1     1
 2   80    10    60      90    20    80
 3    0     0     0       1     1     1
 4    0     0     0       1     1     1
 5    0     0     0       1     1     1
 6    0     0     0       1     1     1
 7    0     0     0       1     1     1
 8    0     0     0       1     1     1

-----
cos  qid  threshold_id
-----
 0   1   2
 1   2   2
 2   5   2
 3   5   2
 4   5   1
 5   5   1
 6   7   1
 7   8   1

```

2.6 监视与维护

清除各类信息

-

查看运行情况

作用	命令
显示流分类信息	show class-map [<i>class-map-name</i>]
显示 QoS 策略信息	show policy-map [<i>policy-map-name</i> [class <i>class-map-name</i>]]
显示接口上应用的策略信息	show policy-map interface <i>interface-id</i>

显示逻辑端口组信息	show virtual-group [<i>virtual-group-number</i> summary]
显示逻辑端口组应用的策略信息	show mls qos virtual-group [<i>virtual-group-number</i> policers]
显示各类映射	show mls qos maps [cos-dscp dscp-cos ip-prec-dscp]
显示端口速度限制信息	show mls qos rate-limit [interface <i>interface-id</i>]
显示 QoS 队列、调度策略轮转权重信息	show mls qos queueing
显示输出队列调度策略信息	show mls qos scheduler
显示 WRED 的配置信息	show queueing wred interface <i>interface-id</i>
显示接口的 QoS 信息	show mls qos interface <i>interface-id</i> [policers]
显示队列带宽信息	show qos bandwidth [interfaces <i>interface-id</i>]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 QoS 库的调试开关。	debug qos lib [event message]
打开 QoS 通信服务端的调试开关	debug qos server [event message]
打开 QoS 用户命令处理的调试开关	debug qos mls
打开与 VMSUP 相关配置的调试开关	debug qos vmsup



配置指南-可靠性

本分册介绍可靠性配置指南相关内容，包括以下章节：

1. RLDP
2. DLDP
3. VRRP
4. VRRP Plus
5. BFD
6. IP Event Dampening
7. VSU
8. VSD

1 RLDP

1.1 概述

RLDP (Rapid Link Detection Protocol , 快速链路检测协议) 是一种以太网链路故障检测协议 , 用于快速检测单向链路故障、双向链路故障以及下联环路故障。如果发现故障存在 , RLDP 会根据用户配置的故障处理方式自动关闭或通知用户手工关闭相关端口 , 以避免流量的错误转发或者防止以太网二层环路。

协议规范

- 无

1.2 典型应用

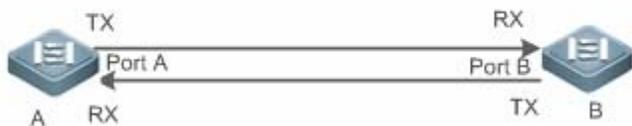
典型应用	场景描述
单向链路检测	检测链路单向故障
双向链路检测	检测链路双向故障
下联环路检测	检测链路环路故障

1.2.1 单向链路检测

应用场景

如下图所示 , 设备 A 与设备 B 之间通过光纤相连 , 图中的两条线分别表示光纤的 Tx 线与 Rx 线 , A 与 B 分别使能 RLDP 的单向链路检测功能。如果端口 A 的 Tx 与端口 B 的 Rx 或者端口 A 的 Rx 与端口 B 的 Tx 中任意一个出现故障 , 那么协议可以检测出单向故障并做出相应的处理。故障如果被恢复 , 管理员可以手工在 A 和 B 上恢复协议状态并重新开始检测。

图 1-1



- 【注释】 A、B 为二层或者三层交换机。
A 上的 Port A 的 TX 与 B 上的 Port B 的 RX 连接。
A 上的 Port A 的 RX 与 B 上的 Port A 的 TX 连接。

功能部署

- 全局配置 RLDP 使能。
- 接口下配置 RLDP 的单向链路检测功能并指定单向故障发生时的处理方式。

1.2.2 双向链路检测

应用场景

如下图所示，设备 A 与设备 B 之间通过光纤相连，图中的两条线分别表示光纤的 Tx 线与 Rx 线，A 与 B 分别使能 RLDP 的双向链路检测功能。如果端口 A 的 Tx 与端口 B 的 Rx 以及端口 A 的 Rx 与端口 B 的 Tx 同时出现故障，那么协议可以检测出双向故障并做出相应的处理。故障如果被恢复，管理员可以手工在 A 和 B 上恢复协议状态并重新开始检测。

图 1-2



- 【注释】 A、B 为二层或者三层交换机。
A 上的 Port A 的 TX 与 B 上的 Port B 的 RX 连接。
A 上的 Port A 的 RX 与 B 上的 Port A 的 TX 连接。

功能部署

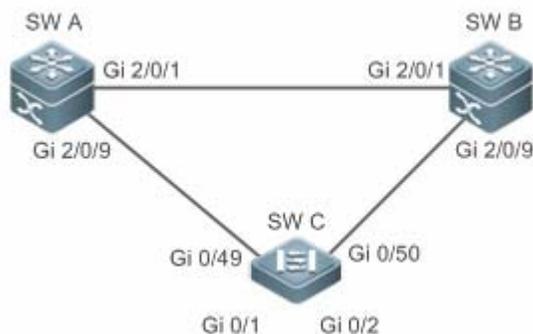
- 全局配置 RLDP 使能。
- 接口下配置 RLDP 的双向链路检测功能并指定双向故障发生时的处理方式。

1.2.3 下联环路链路检测

应用场景

如下图所示，设备 A、设备 B 以及设备 C 之间连成网络环路，A 使能 RLDP 的下联环路检测功能，协议此时可以检测出环路故障并做响应的。

图 1-3



【注释】 A、B、C 为二层或者三层交换机。
A、B、C 通过交换口两两互连。

功能部署

- A 上全局配置 RLDP 使能。
- A 与 B 的连接端口、A 与 C 的连接端口分别配置 RLDP 下联环路检测功能并指定环路故障发生时的处理方式。

1.3 功能详解

一般的以太网链路检测机制都只是利用物理连接的状态，通过物理层的自动协商来检测链路的连通性。但是这种检测机制存在一定的局限性，有些情况下物理层虽然处于连通状态并能正常工作，但是实际对应的二层链路却是无法通信或者存在异常。RLDP 协议通过与邻居设备交互探测报文、探测响应报文或者环路报文来识别邻居设备并检测链路是否存在故障。

基本概念

↘ 链路单向故障

光纤交叉连接、一条光纤未连接、一条光纤断路、双绞线中的一条线路断路或者两台设备之间的中间设备出现单向断路等情况下会出现链路单向故障，这种链路一边能通而另一边不能通会导致流量被错误转发或者环路保护协议（比如 STP）功能失效。

↘ 链路双向故障

两条光纤断路、双绞线中的两条线路断路或者两台设备之间的中间设备出现双向断路等情况下会出现链路双向故障，这种链路双向都不通会导致流量被错误的转发。

↘ 链路环路故障

设备下联被用户错误的接入其他设备形成了环路，这种会在网络中引起广播风暴。

↘ RLDP 协议报文

协议定义了三种类型的报文：探测报文（Prob）、探测响应报文（Echo）以及环路报文（Loop）。

- Prob 报文为二层组播报文，用于邻居协商、单向或者双向链路检测，报文的默认封装格式为 SNAP 类型，如果邻居发出的报文格式为 EthernetII 格式则封装方式自动变更为 EthernetII；
- Echo 报文为响应 Prob 报文的二层单播报文，用于单向或者双向链路检测，报文的默认封装格式为 SNAP 类型，如果邻居发出的报文格式为 EthernetII 格式则封装方式自动变更为 EthernetII；
- Loop 报文为二层组播报文，用于下联环路检测，这类报文只会被发送方所接收，报文的封装格式为 SNAP 封装方式。

▾ RLDP 探测间隔及最大探测次数

RLDP 可以配置探测间隔与最大探测次数。探测间隔决定了 Prob 报文与 Loop 报文的发送周期，设备在接收到 Prob 报文后会立即响应 Echo 报文。探测间隔与最大探测次数决定了单向或者双向链路探测的最大探测时间（探测间隔 × 最大探测次数 + 1），最大探测时间内如果无法正确接收到邻居的 Prob 报文或者 Echo 报文可以触发单向或者双向故障的处理。

▾ RLDP 邻居协商

配置了单向或者双向检测功能的端口可以学习到对端设备作为邻居，一个端口支持学习一个邻居，邻居可变化。协商功能启用后，端口下协商到邻居后才开始单向或者双向检测，协商过程中如果成功接收到邻居发送的 Prob 报文就认为协商成功。但是，在已存在故障的情况下才使能协议，会出现无法正常学习邻居而导致检测不能启动，建议此时先恢复链路的错误状态。

▾ RLDP 端口故障时的处理方式

- warning：只打印相关的 Syslog 说明当前的故障端口和故障类型。
- Shutdown SVI：打印 Syslog 的基础上，如果故障端口为物理交换口或者 L2 AP 成员口，那么会根据端口所属的 Access VLAN 或者 Native VLAN 查询出对应的 SVI 并执行 Shutdown 操作。
- 端口违例：打印 Syslog 的基础上，设置故障端口为违例状态，此时端口物理上会进入 Linkdown 状态。
- Block：打印 Syslog 的基础上，设置故障端口的转发状态为 Block，此时端口不对收到的报文进行转发。

▾ RLDP 端口故障后的恢复方式

- 手工执行 Reset：手工将所有故障端口恢复到初始化状态，此时会重新启动链路检测。
- 手工或者自动执行 Errdisable Recovery：手工或者定时（默认每 30s，可配置）恢复所有故障端口到初始化状态并重新启动链路检测。
- 自动恢复：单向或者双向链路检测的情况下，如果指定的故障处理方式不是端口违例，那么可以依赖与邻居交互的 Prob 报文自动恢复到初始化状态并重新启动链路检测。

▾ RLDP 端口状态

- normal：端口下配置启动检测后的状态。
- error：端口下检测出链路故障后的状态，可以是单向、双向或者环路故障导致。

▾ 功能特性

功能特性	作用
建立RLDP检测	启用单向、双向或者下联环路检测功能，永远检测单向、双向或者环路故障并进行相应的故障处理。

1.3.1 建立RLDP检测

RLDP 的链路检测模式主要包括单向链路检测、双向链路检测以及下联环路检测等。

工作原理

↘ RLDP 单向链路检测

单向链路检测启动后，端口后周期的发送 Prob 报文并接收邻居响应的 Echo 报文，同时接收邻居的 Prob 报文并及时响应 Echo 报文给邻居。在最大探测时间内，如果只能接收到邻居的 Prob 报文但无法接收到邻居的 Echo 报文或者既不能接收到邻居的 Prob 报文也不能接收到邻居的 Echo 报文，那么会触发单向故障的处理并停止检测。

↘ RLDP 双向链路检测

双向链路检测启动后，端口后周期的发送 Prob 报文并接收邻居响应的 Echo 报文，同时接收邻居的 Prob 报文并及时响应 Echo 报文给邻居。在最大探测时间内，如果既不能接收到邻居的 Prob 报文也不能接收到邻居的 Echo 报文，那么会触发双向故障的处理并停止检测。

↘ RLDP 下联环路检测

下联环路检测启动后，端口会周期的发送 Loop 报文，同一个设备的相同端口或者不同端口接收到 Loop 报文后，如果报文发送端口与接收端口为路由口或者 L3 AP 成员口并且发送口与接收后相同则触发环路故障，或者报文发送端口与接收端口为交换口或者 L2 AP 成员口并且端口的默认 VLAN 相同同时转发状态均为 Forward 则触发环路故障，故障发生后按相应的故障处理方式来处理并停止检测。

相关配置

- 配置 RLDP 检测功能

缺省情况下，检测功能不生效。

使用 RLDP 全局命令 `rldp enable` 和接口命令 `rldp port` 可以启动 RLDP 检测功能，并指定检测类型与故障处理方式。

用户可以根据实际环境通过 `rldp neighbor-negotiation` 指定邻居协商、`rldp detect-interval` 指定探测间隔、`rldp detect-max` 指定探测次数、`rldp reset` 恢复故障端口状态等。

1.4 配置详解

配置项	配置建议&相关命令
配置RLDP基本功能	 全局模式，必须配置。配置全局开启 RLDP 探测功能
	<code>rldp enable</code> 全局下启动 RLDP 检测，生效到所有端口。

 接口模式，必须配置。指定接口下的探测类型以及故障处理方式。	
rldp port	端口下启动 RLDP 检测，指定具体的检测类型以及发生故障后的处理方式。
 全局模式，可选配置。指定探测过程中的探测间隔、探测次数、是否需要邻居协商。	
rldp detect-interval	全局修改 RLDP 配置参数，包括探测间隔、最大探测次数以及邻居协商，可生效到所有端口下的 RLDP 检测。
rldp detect-max	
rldp neighbor-negotiation	
 特权模式，可选配置。	
rldp reset	特权下恢复故障端口的状态，可生效到所有端口下的 RLDP 检测。

1.4.1 配置RLDP基本功能

配置效果

- 启用 RLDP 单向、双向或者下联环路检测，用于发现单向、双向或者环路故障。

注意事项

- 对于 AP 成员口上的 RLDP 配置，如果是配置环路检测，则会同步配置到该 AP 的所有成员口，如果是配置单向链路检测和双向链路检测，则直接在 AP 成员口生效。
- 对于物理口加入 AP 的情况，新加入的 AP 成员口的环路检测配置需要和该 AP 现有的成员口的环路检测配置一致。这里分 3 种情况：1) 如果新加入的 AP 成员口没有配置环路检测，而该 AP 现有的成员口有配置环路检测，则新加入的 AP 成员口同步环路检测的配置和检测结果。2) 如果新加入的 AP 成员口有配置环路检测，而该 AP 现在的所有成员口都没有配置环路检测，则新加入的 AP 成员口清除环路检测配置并加入 AP。3) 如果新加入的 AP 成员口的环路检测配置和该 AP 现有的成员口的环路检测配置不一致，则新加入的 AP 成员口同步环路检测的配置和检测结果。
- AP 成员口配置 RLDP 时，故障处理方法只能配置为“shutdown-port”，如果故障处理方法配置为非“shutdown-port”时，将转换成“shutdown-port”的配置并生效。
- 配置了“shutdown-port”故障处理的端口在出现故障后将无法主动恢复 RLDP 探测，如果用户确认故障已经解决，则可以使用 **rldp reset** 命令或者 **errdisable recovery** 命令来恢复并重新启动检测，**errdisable recovery** 的配置可以参考 <<SWITCH-INTF-SCG.doc>>。

配置方法

▾ 全局配置使能

- 必须配置。

- 全局模式下配置，配置后各端口的检测可以启动。

↘ 全局配置邻居协商

- 可选配置。
- 全局模式下配置，配置后各端口检测的启动依赖于邻居协商的成功。

↘ 全局配置探测间隔

- 可选配置。
- 全局模式下配置，可以指定具体的时间间隔。

↘ 全局配置最大探测次数

- 可选配置。
- 全局模式下配置。
- 可以指定具体的最大探测次数。

↘ 接口下配置检测功能

- 必须配置。
- 接口模式下配置。
- 在接口下配置 RLDP 功能，可以选择单向、双向或者下联环路检测类型，同时指定对应的故障处理方式。

↘ 特权下配置恢复所有故障端口状态

- 可选配置。
- 特权模式下配置，配置后可以恢复所有故障状态端口，重新启动检测。

检验方法

- 查看设备的 RLDP 信息，包括全局、端口以及邻居的相关信息。

相关命令

↘ 全局使能 RLDP 检测功能

- 【命令格式】 `rldp enable`
- 【参数说明】 -
- 【命令模式】 全局模式。

【使用指导】 全局启用 RLDP 检测功能。

▾ 接口下启动 RLDP 检测功能

【命令格式】 `rldp port { unidirection-detect | bidirection-detect | loop-detect } { warning | shutdown-svi | shutdown-port | block }`

【参数说明】 **unidirection-detect**：单向链路检测。
bidirection-detect：双向链路检测。
loop-detect：下联环路检测。
warning：故障处理方式为告警。
shutdown-svi：故障处理方式为关闭接口所在的 SVI 口。
shutdown-port：故障处理方式为端口违例。
block：故障处理方式为关闭端口的学习转发能力。

【命令模式】 接口模式

【使用指导】 接口包括：2 层交换口、3 层路由口、L2AP 下的成员口、L3AP 下的成员口等接口。

▾ 全局修改 RLDP 检测参数

【命令格式】 `rldp {detect-interval interval | detect-max num | neighbor-negotiation }`

【参数说明】 **detect-interval interval**：探测间隔。
detect-max num：最大探测次数。
neighbor-negotiation：邻居协商。

【命令模式】 全局模式

【使用指导】 当实际环境变化，需要修改所有 RLDP 检测的参数时，对所有端口生效。

▾ 恢复 RLDP 故障端口状态

【命令格式】 `rldp reset`

【参数说明】 -

【命令模式】 特权模式

【使用指导】 恢复 RLDP 所有故障端口状态到初始状态并重新启动检测。

▾ 查看 RLDP 状态信息

【命令格式】 `show rldp [interface interface-name]`

【参数说明】 *interface-name*：指定要查看的具体接口

【命令模式】 特权模式、全局模式、接口模式

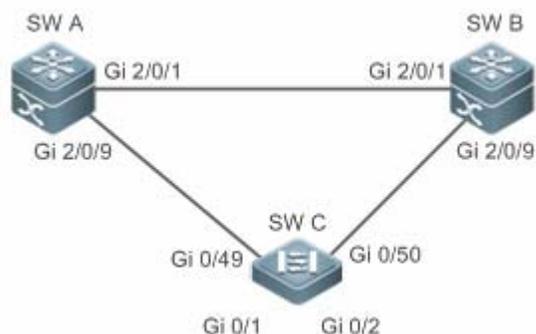
【使用指导】 查看 RLDP 状态信息。

配置举例

▾ 在环网拓扑中开启 RLDP 检测功能

【网络环境】 如下图所示，汇聚与接入为环网拓扑，环网中各设备均开启 STP 来防止环路并提供冗余保护，为了防止环路

图 1-4 中链路出现单向或者双向故障进而导致 STP 协议失效，汇聚设备与汇聚设备之间以及汇聚与接入设备之间启用 RLDP 单向和双向检测，为了防止汇聚设备下联被错误的接入而出现环路，汇聚设备与接入设备的下联端口均开启 RLDP 环路检测；为了防止接入设备下联被错误的接入而出现环路，接入设备的下联端口均开启 RLDP 环路检测



- 【配置方法】**
- SW A、SW B 作为汇聚，SW C 作为接入，SW C 下联可以接用户设备，三台设备组成环网拓扑，每台设备开启 STP，STP 配置参考相关配置指南。
 - SW A 开启 RLDP，两个端口需要配置单向和双向链路检测，下联端口需要配置开启环路检测。
 - SW B 开启 RLDP，两个端口需要配置单向和双向链路检测，下联端口需要配置开启环路检测。
 - SW C 开启 RLDP，上联两个端口需要配置单向和双向链路检测，下联两个端口需要配置开启环路检测

A

```
A#configure terminal
A(config)#rldp enable
A(config)#interface GigabitEthernet 2/0/1
A(config-if-GigabitEthernet 2/0/1)#rldp port unidirection-detect shutdown-port
A(config-if-GigabitEthernet 2/0/1)#rldp port bidirection-detect shutdown-port
A(config-if-GigabitEthernet 2/0/1)# exit
A(config)#interface GigabitEthernet 2/0/9
A(config-if-GigabitEthernet 2/0/1)#rldp port unidirection-detect shutdown-port
A(config-if-GigabitEthernet 2/0/1)#rldp port bidirection-detect shutdown-port
A(config-if-GigabitEthernet 2/0/1)#rldp port loop-detect shutdown-port
A(config-if-GigabitEthernet 2/0/1)#exit
```

B

同 A 的配置

C

```
C#configure terminal
C(config)#rldp enable
C(config)#interface GigabitEthernet 0/49
C(config-if-GigabitEthernet 0/49)#rldp port unidirection-detect shutdown-port
C(config-if-GigabitEthernet 0/49)#rldp port bidirection-detect shutdown-port
C(config-if-GigabitEthernet 0/49)# exit
C(config)#interface GigabitEthernet 0/50
C(config-if-GigabitEthernet 0/50)#rldp port unidirection-detect shutdown-port
C(config-if-GigabitEthernet 0/50)#rldp port bidirection-detect shutdown-port
C(config-if-GigabitEthernet 0/50)#exit
```

```
C(config)#interface GigabitEthernet 0/1
C(config-if-GigabitEthernet 0/1)# rldp port loop-detect shutdown-port
C(config-if-GigabitEthernet 0/1)#exit
C(config)#interface GigabitEthernet 0/2
C(config-if-GigabitEthernet 0/2)# rldp port loop-detect shutdown-port
C(config-if-GigabitEthernet 0/2)#exit
```

【检验方法】 ● 检查 A、B、C 设备的 RLDP 状态信息，以 A 为例。

A

```
A#show rldp
rldp state          : enable
rldp hello interval: 3
rldp max hello      : 2
rldp local bridge   : 00d0.f822.33aa
-----
Interface GigabitEthernet 2/0/1
port state          : normal
neighbor bridge    : 00d0.f800.51b1
neighbor port      : GigabitEthernet 2/0/1
unidirection detect information:
  action: shutdown-port
  state : normal
bidirection detect information:
  action: shutdown-port
  state : normal

Interface GigabitEthernet 2/0/9
port state          : normal
neighbor bridge    : 00d0.f800.41b0
neighbor port      : GigabitEthernet 0/49
unidirection detect information:
  action: shutdown-port
  state : normal
bidirection detect information:
  action: shutdown-port
  state : normal
loop detect information:
  action: shutdown-port
  state : normal
```

常见错误

- 与私有组播地址认证或者 TPP 等功能不可以同时开启。

- 配置单双向检测时不指定邻居协商，要求邻居设备在全局和接口下使能 RLDP，否则会被检测为单向或者双向故障。
- 配置单双向检测时如果指定了先协商邻居后开始检测，那么在已存在故障的情况下由于无法学习到邻居而导致不能正常检测，建议先恢复链路错误状态。
- 路由口下建议不要指定故障处理方式为 Shutdown SVI。
- STP 等环路保护协议使能的端口下建议不要指定故障处理方式为 Block。

1.5 监视与维护

查看运行情况

作用	命令
查看 RLDP 运行状态。	show rldp [interface <i>interface-name</i>]

2 DLDP

2.1 概述

DLDP (Data Link Detection Protocol, 数据链路检测协议), 是一种基于快速检测以太网链路故障的检测协议。

一般的以太网链路检测机制都只是利用物理连接的状态, 通过物理层的自动协商来检测链路的连通性。但是该方法对于三层的数据连通性检测, 如物理连接状态正常但三层数据通信异常的场景, 存在一定的局限性。

DLDP 在这类场景中, 为用户提供可靠的三层链路检测信息。同时, 在检测出故障后, DLDP 主动 SHUTDOWN 三层接口的逻辑状态, 促使三层协议快速收敛。

协议规范

- 无

2.2 典型应用

典型应用	场景描述
同网段DLDP检测	检测端口的源 IP 与检测 IP 属于同一网段。
跨网段DLDP检测	检测端口的源 IP 与检测 IP 属不同网段

2.2.1 同网段DLDP检测

应用场景

检测端口的源 IP 与检测 IP 属于同一网段的基本应用场景。

以下图为例, 交换机 A 上的三层口 Gi 0/1 与交换机 C 上的三层口 Gi 0/2 属同网段, 若要检测 Gi 0/1 到 Gi 0/2 三层线路的连通性, 仅需在 A 或 C 的相应三层口开启 DLDP 检测功能即可。

图 2-1



【注释】 A、C 为交换机。

Gi 0/1 与 Gi 0/2 均为三层口，且属同一网段。
B 为同网段网络。

功能部署

- 在 Gi 0/1 或 Gi 0/2 上开启 DLDP 检测即可。

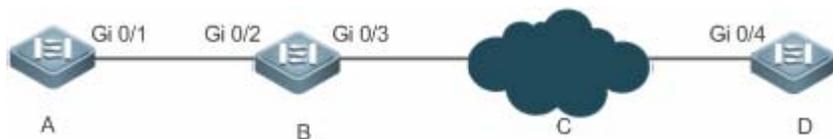
2.2.2 跨网段DLDP检测

应用场景

检测端口的源 IP 与检测 IP 在不同网段的应用场景。

以下图为例，交换机 A 上的三层口 Gi 0/1 与交换机 D 上的三层口 Gi 0/4 在不同网段，若要检测 Gi 0/1 到 Gi 0/4 三层线路的连通性，则需在 A 的 Gi 0/1 开启 DLDP 功能的同时还需再配置 DLDP 下一跳 IP 地址（交换机 B 上的 Gi 0/2 的 IP 地址）。

图 2-2



【注释】 A、B、D 为交换机。
Gi 0/1 与 Gi 0/4 均为三层口，但在不同网段。

功能部署

- 在 Gi 0/1 上开启 DLDP 并配置 DLDP 下一跳 IP 地址。

2.3 功能详解

基本概念

DLDP 的探测间隔及重传次数

探测间隔：指 DLDP 探测报文（ICMP echo）的发送间隔。

重传次数：指 DLDP 探测失效所需的发包次数。

当网络设备在“探测间隔”×“重传次数”的时间周期内没有收到对端的应答报文，则认为三层链路故障，主动 SHUTDOWN 三层接口逻辑状态（实际物理链路还是连通的）。一旦三层链路正常通讯，则恢复三层接口逻辑状态 UP。

DLDP 检测模式

DLDP 的检测模式包括：主动模式和被动模式。

主动模式：指主动发送 ICMP 探测报文的模式，缺省配置为主动模式。

被动模式：指被动接收 ICMP 探测报文的模式。

DLDP 的下一跳

下一跳：指跨网段 DLDP 检测中，检测 IP 对应的路由下一节点。

在某些情形下，DLDP 需要检测非直联网段 IP 的可达性。这时需要配置该接口的下一跳 IP，以便 DLDP 能够通过 ARP 报文获取下一跳 MAC 地址，正确的封装 ICMP 报文并发送。

但在这种情形下，需避免响应报文从其他链路回应的情况，否则会造成 DLDP 误判该接口没有收到 ICMP 应答。

DLDP 的恢复次数

恢复次数：指 DLDP 从故障中恢复需连续收到的响应报文（ICMP reply）次数。

在有些情形下，检测链路可能不太稳定，比如 PING 断了三次，通一次，又断了多次。如果按简单的逻辑，其中的 DLDP 检测就是 UP、DOWN 多次，实际可能加剧环网的不稳定。

恢复次数表示链路从 DOWN 状态为设置为 UP 状态前，需要收到连续的 DLDP 检测报文响应次数。恢复次数缺省为 3 次，即只有该链路上连续 PING 通了 3 次才会设置为 UP。在这种情况下，虽然降低了链路检测的灵敏度，但增加其稳定性，同时相关参数在实际应用中还可根据网络情况进行调整。

DLDP 的绑定 MAC

绑定 MAC：指 DLDP 检测 IP 所绑定的 MAC 地址

在复杂的网络环境下，检测链路中可能存在异常 ARP 报文（ARP 欺骗），使得 DLDP 则获取到非法的 MAC 地址，从而导致检测无法正常工作。

在此种环境下，通过配置绑定 MAC 地址，可将检测 IP(或下一跳 IP)与静态 MAC 地址进行绑定，不再受异常 ARP 报文欺骗而引起 DLDP 功能失效。

功能特性

功能特性	作用
建立DLDP检测	实现 DLDP 三层链路连通性检测，在三层链路异常时，主动 SHUTDOWN 掉对应三层口。
绑定MAC地址	当网络中存在 ARP 欺骗等异常情况，可将检测 IP 与设备 MAC 地址进行绑定，避免协议异常发生。
DLDP被动模式	当检测链路两端都开启 DLDP，其中一端可配置为被动模式，以节省带宽资源和设备 CPU 资源。

2.3.1 建立DLDP检测

DLDP 三层链路连通性检测，在检测到三层链路异常时，主动 SHUTDOWN 掉对应三层口。

工作原理

启动 DLDP 功能后，DLDP 会通过 ARP 报文获取被检测设备或者到达被检测设备的下一跳设备的 mac 地址和出接口，然后通过周期性的 IPv4 ICMP echo 报文进行通路检测。如果在指定时间内被检测设备没有回应 IPv4 ICMP reply 报文，则认为这个接口通路出现问题，将该接口设置为“三层接口 DOWN”。

相关配置

- 配置 DLDP 检测功能

缺省情况下，接口上不开启 DLDP 检测功能。

使用 dldp 命令并指定要检测的目的 IP 地址就可以启动 DLDP 检测功能。

用户可以根据实际环境选择是否配置下一跳 IP、MAC 地址、发送间隔、重传次数、恢复次数等参数。

2.3.2 绑定MAC地址

网络中存在 ARP 欺骗等异常情况，可将检测 IP（或下一跳 IP）与设备 MAC 地址进行绑定，避免协议发生异常情况。

工作原理

网络中存在 ARP 欺骗的情况下，通过配置绑定 MAC 地址，可将检测 IP(或下一跳 IP)与静态 MAC 地址进行绑定，不再受异常 ARP 报文欺骗而引起 DLDP 功能失效。

相关配置

缺省情况下开启 DLDP 检测功能时不指定 MAC 地址绑定。

通过 dldp 命令开启检测时并同时指定要绑定的 MAC 地址，若存在下一跳 IP 地址，则配置的是下一跳设备的 MAC 地址，否则配置目的检测设备的 MAC 地址。

配置开启后，DLDP 探测过程中发送的 ARP 报文与 ICMP 报文中的目的 IP 与目的 MAC 是固定的，如果接收的报文中源 IP 与源 MAC 和绑定的 IP 与 MAC 不匹配，则不会进行处理。

2.3.3 DLDP被动模式

当检测链路两端都开启 DLDP，其中一端可配置为被动模式，以节省带宽资源和设备 CPU 资源。

工作原理

一端设备使用 ICMP echo 发包，另一端用同样的检测参数来确认报文的及时可达，一样能达到双方设备检测链路通路的效果，同时也节省了带宽资源和设备 CPU 资源。

相关配置

缺省情况下不开启 DLDP 被动检测模式。

通过 `dldp passive` 命令开启被动检测。

配置开启后，DLDP 将不再主动发起 ICMP echo 报文进行探测，只需要在接收到 ICMP echo 报文后响应 ICMP Reply 报文即可，在指定时间内如果没有收到 ICMP echo 报文则认为接口通路出现问题。

2.4 配置详解

配置项	配置建议&相关命令
配置DLDP检测功能	 必须配置，接口模式，开启 DLDP 检测功能
	<code>dldp</code> 配置 DLDP 检测功能
	 可选配置，接口模式，开启被动检测功能。
	<code>dldp passive</code> 配置接口被动模式
	 可选配置，全局模式，指定探测过程中的探测间隔、重传次数以及恢复次数。
	<code>dldp interval</code> <code>dldp retry</code> <code>dldp resume</code> 全局修改 DLDP 配置参数，可生效到所有 DLDP 检测

2.4.1 配置DLDP检测功能

配置效果

- 实现 DLDP 三层链路连通性检测，在三层链路异常时，主动 SHUTDOWN 掉对应三层口。

注意事项

- 一个三层接口下，DLDP 可以配置多个 IP 检测，当所有 IP 都没有 ICMP 响应时，才认为接口 DOWN；而一旦有一个 IP 恢复通讯，则认为接口恢复 UP。
- DLDP 使用该三层接口的第一个 IP 地址作为报文的源 IP 地址进行通讯。

配置方法

启动 DLDP 检测功能

- 必须配置。
- 在接口下配置 DLDP 功能：根据实际环境选择是否配置下一跳 IP、MAC 地址、发送间隔、重传次数、恢复次数等参数。

配置 DLDP 检测模式

- 可选配置。
- 在接口下配置 DLDP 检测模式：根据实际环境选择配置为主动或被动模式。
- 如当三层链路两端都需要开启 DLDP 功能，为节省了带宽资源和设备 CPU 资源，可将其中一端的 DLDP 检测模式改为被动模式。

全局配置 DLDP 参数

- 可选配置。
- 根据需求，可在全局下修改所有 DLDP 检测的参数，包括：检测报文的发送间隔、检测报文的重新次数、恢复次数。

检验方法

- 查看设备的 DLDP 信息，包括所有 DLDP 检测的状态信息以及其统计信息。

相关命令

启动 DLDP 检测功能

【命令格式】 **dldp** *ip-address* [*next-hop-ip*] [**mac-address** *mac-addr*] [**interval** *tick*] [**retry** *retry-num*] [**resume** *resume-num*]

【参数说明】 *ip-address*：DLDP 检测 IP 地址。
next-hop-ip：下一跳 IP 地址。
mac-addr：绑定 MAC 地址，若存在下一跳 IP 地址，则配置的是下一跳设备的 MAC 地址。
tick：检测报文的发送间隔，取值范围：5-6000 tick（1 tick = 10 毫秒），缺省值 100 tick（1 秒）。
retry-num：取值范围：1-3600，缺省值 4。
resume-num：恢复次数，取值范围：1-200，缺省值 3。

【命令模式】 接口模式

【使用指导】 接口必须是三层接口，包括：路由口、L3AP、SVI 等接口。

配置 DLDP 检测模式

【命令格式】 **dldp** *passive*

- 【参数说明】 -
- 【命令模式】 接口模式
- 【使用指导】 接口下必须先开启 DLDP 检测功能后才能配置 DLDP 检测模式。

全局修改 DLDP 检测参数

- 【命令格式】 **dldp { interval tick | retry retry-num | resume resume-num }**
- 【参数说明】 *tick* : 检测报文的发送间隔, 取值范围 : 5-6000 tick (1 tick =10 毫秒), 缺省值 100 tick (1 秒)。
retry-num : 检测报文的重传次数, 取值范围 : 1-3600, 缺省值 4。
resume-num : 恢复次数, 取值范围 : 1-200, 缺省值 3。
- 【命令模式】 全局模式
- 【使用指导】 当实际环境变化, 需要修改所有 DLDP 检测的参数时, 可使用该命令快速生效。

查看 DLDP 状态信息

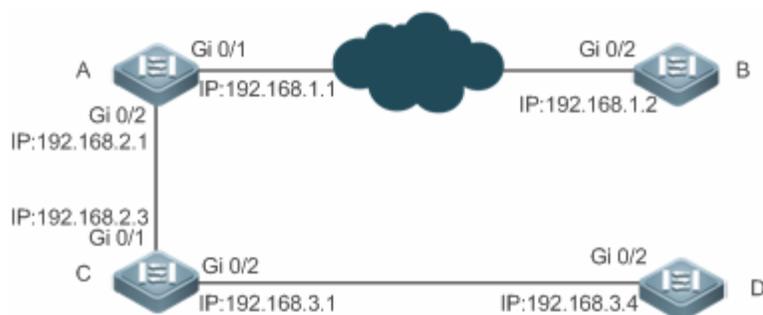
- 【命令格式】 **show dldp statistic [interface interface-name]**
- 【参数说明】 *interface-name* : 查看信息的三层接口
- 【命令模式】 特权模式、全局模式、接口模式
- 【使用指导】 查看接口下的 DLDP 工作状态信息。
查看所有 DLDP 检测的统计信息。

配置举例

在三层网络上开启 DLDP 检测功能, 分别控制 A、B 设备的三层口

【网络环境】

图 2-3



- 【配置方法】
- 在设备 A 上的路由口 (Gi 0/1、Gi 0/2) 开启 DLDP 功能, 检测 A 到 B 和 D 的三层网络的链通性。
 - 若需控制 B 设备的路由口 (Gi 0/2), 则在该接口上开启 DLDP 功能, 并配置为被动模式。

```
A
A#configure terminal
A(config)#interface GigabitEthernet 0/1
A(config-if-GigabitEthernet 0/1)#dldp 192.168.1.2
A(config-if-GigabitEthernet 0/1)# exit
A(config)#interface GigabitEthernet 0/2
A(config-if-GigabitEthernet 0/1)#dldp 192.168.3.4 192.168.2.3
```

```

B
B#configure terminal
B(config)#interface GigabitEthernet 0/2
B(config-if-GigabitEthernet 0/1)#dldp 192.168.1.1
B(config-if-GigabitEthernet 0/1)#dldp passive

```

【检验方法】 ● 检查 A、B 设备的 DLDP 状态信息，检测 DLDP 检测是正常开启并工作。

```

A
A# show dldp
Interface  Type      Ip      Next-hop  Interval  Retry  Resume  State
-----  -
Gi0/1     Active   192.168.1.2
Gi0/1     Active   192.168.3.4  192.168.2.3  100      4      3      Up

```

```

B
B# show dldp
Interface  Type      Ip      Next-hop  Interval  Retry  Resume  State
-----  -
Gi0/2     Passive  192.168.1.1

```

常见错误

- IPv4 单播路由不可达，误以为 DLDP 检测失效。
- 对端设备不支持 arp/icmp 回应导致 DLDP 功能失效。
- 跨网段检测没有配置下一跳 IP 地址。

2.5 监视与维护

清除各类信息

作用	命令
清除 DLDP 统计信息。	clear dldp [interface <i>interface-name</i> [<i>ip-address</i>]]

查看运行情况

作用	命令
查看 DLDP 运行状态。	show dldp [interface <i>interface-name</i>]
查看 DLDP 的 down/up 统计信息	show dldp statistic

3 VRRP

3.1 概述

VRRP (Virtual Router Redundancy Protocol , 虚拟路由冗余协议) 是一种路由容错协议。

VRRP 设计采用主备模式 , 以保证当主路由设备发生故障时 , 备份路由设备可以在不影响内外数据通信的前提下进行功能切换 , 且不需要再修改内部网络的参数。VRRP 组内多个路由设备都映射为一个虚拟的路由设备。VRRP 保证同时有且只有一个路由设备在代表虚拟路由设备进行包的发送 , 而主机则是把数据包发向该虚拟路由设备 , 这个转发数据包的路由设备被选择成为主路由设备。如果这个主路由设备在某个时候由于某种原因而无法工作的话 , 则处于备份状态的路由设备将被选择来代替原来的主路由设备。VRRP 使得局域网内的主机看上去只使用了一个路由设备 , 并且即使在它当前所使用的首跳路由设备失败的情况下仍能够保持路由的连通性。

- VRRP 适用于需要对局域网中设备的路由出口进行冗余备份的场景。

 下文仅介绍 VRRP 的相关内容。

协议规范

- RFC2338 : Virtual Router Redundancy Protocol
- RFC3768 : Virtual Router Redundancy Protocol (VRRP)
- RFC5798 : Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6

3.2 典型应用

典型应用	场景描述
路由冗余	局域网中路由设备配置单备份组实现简单的路由冗余。
负载均衡	局域网中路由设备配置多备份组实现流量的负载均衡。

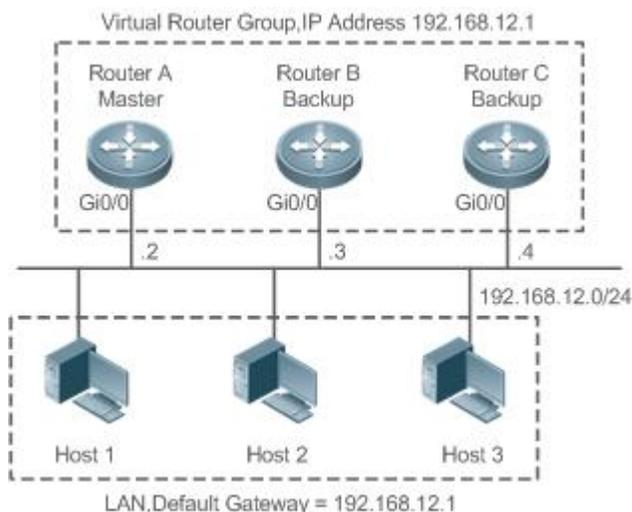
3.2.1 路由冗余

应用场景

在连接局域网的路由设备上配置单备份组 , 局域网中的主机以该备份组的虚拟 IP 作为默认网关。

- 局域网内主机 1、2 以及 3 发往其它网络的数据包将由选举出来的主路由设备(在图 3-1 中是路由设备A)进行路由转发。
- 一旦主路由设备 A 失效 , 将在路由设备 B 与 C 之间选举出新的主路由设备来承担虚拟路由设备的路由转发功能 , 由此实现了简单路由冗余。

图 3-1



功能部属

- 路由设备 A、B 以及 C 均使用以太网口与局域网连接。
- 路由设备 A、B 以及 C 在与局域网连接的以太网接口上设置了 VRRP。
- 路由设备 A、B 以及 C 在与局域网连接的以太网接口上处于同一个 VRRP 组并且该 VRRP 组的虚拟 IP 地址为 192.168.12.1。
- 局域网内的主机 1、2 以及 3 以虚拟路由设备的 IP 地址 192.168.12.1 作为网关。

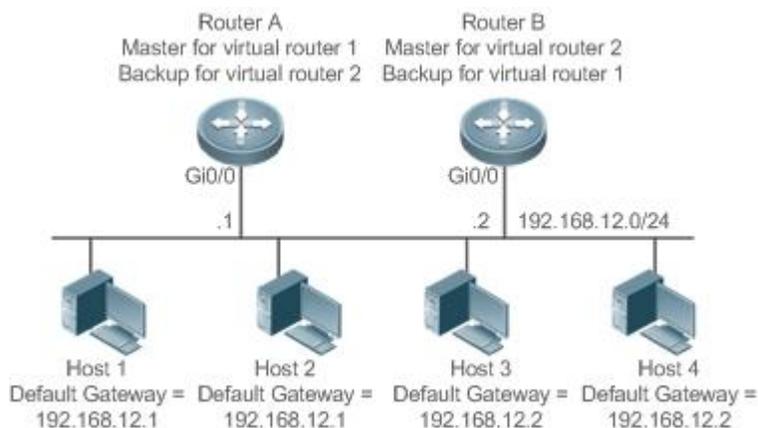
3.2.2 负载均衡

应用场景

在连接局域网的路由设备上配置多个备份组，局域网中的主机分别以各备份组的虚拟 IP 作为网关，路由设备互相作为主路由设备和备份路由设备。

- 以虚拟路由器 1 的虚拟 IP 地址作为默认网关的主机 1 和 2 发往其它网络的数据包将由主路由设备(在图 3-2 中是路由设备 A)进行路由转发。
- 以虚拟路由器 2 的虚拟 IP 地址作为默认网关的主机 3 和 4 发往其它网络的数据包将由主路由设备(在图 3-2 中是路由设备 B)进行路由转发。
- 路由设备 A 和 B 实现了路由冗余，并同时分担了来自局域网的流量即实现了负载均衡。

图 3-2



功能部属

- 路由设备 A 和 B 均使用以太网口与局域网连接。
- 路由设备 A 和 B 在与局域网连接的以太网接口上设置了两个虚拟路由设备。
- 设置了两个虚拟路由设备。对于虚拟路由设备 1，路由设备 A 使用以太网口 Gi0/0 的 IP 地址 192.168.12.1 作为虚拟路由设备的 IP 地址，这样路由设备 A 就成为主路由设备，而路由设备 B 成为备份路由设备。
- 对于虚拟路由设备 2，路由设备 B 使用以太网口 Gi0/0 的 IP 地址 192.168.12.2 作为虚拟路由设备的 IP 地址，这样路由设备 B 就成为主路由设备，而路由设备 A 成为备份路由设备。
- 在局域网内，主机 1 和主机 2 使用虚拟路由设备 1 的 IP 地址 192.168.12.1 作为默认网关，主机 3 和主机 4 使用虚拟路由设备 2 的 IP 地址 192.168.12.2 作为默认网关。

3.3 功能详解

基本概念

虚拟路由器

又称 VRRP 备份组，被当作一个共享局域网内主机的缺省网关。包括一个虚拟路由器标识符和一组虚拟 IP 地址。

虚拟 IP 地址

虚拟路由器的 IP 地址，一个虚拟路由器可以配置有一个或多个虚拟 IP 地址。

IP 地址拥有者

如果 VRRP 组的虚拟 IP 地址与所在以太网接口上的 IP 一致，那么就认为该 VRRP 组占用(Own)了以太网接口实际 IP 地址，认为该设备是 IP 地址拥有者，此时该 VRRP 组的优先级为 255，如果对应的以太网接口可用，那么该 VRRP 组将自动处于 Master 状态。IP 地址拥有者接收处理目的 IP 地址为虚拟路由器的 IP 地址的报文。

虚拟 MAC 地址

VRRP 的虚拟 MAC 地址就是协议规定的、由国际标准统一分配的：00-00-5E-00-01-{-VRID}，其中前五个字节固定，最后一个字节是备份组的组号。当虚拟路由器回应 ARP 请求时，使用虚拟 MAC 地址，而不是接口的实 MAC 地址。

Master 路由器

在一个 VRRP 备份组中，只有 Master 路由器负责 ARP 响应和转发 IP 数据包。如果该设备是 IP 地址拥有者，通常它将成为 Master 路由器。

Backup 路由器

在一个 VRRP 备份组中，Backup 路由器不承担 ARP 响应和转发 IP 数据包的任务，只负责监听 Master 路由器的状态。当 Master 路由器出现故障时，它们将有机会通过选举成为新的 Master 路由器。

抢占模式

如果 VRRP 组工作在抢占模式，一旦它发现自己的优先级高于当前 Master 的优先级，它将抢占成为该 VRRP 组的主路由设备。

功能特性

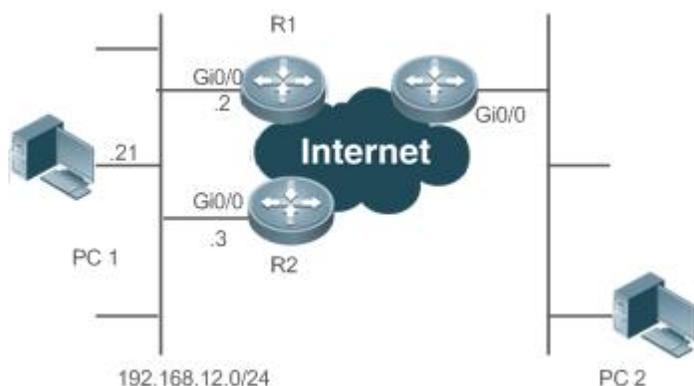
功能特性	作用
VRRP	通过启动 VRRP 功能，VRRP 对共享多存取访问介质（如以太网）上终端设备的默认网关进行冗余备份，从而在其中一台路由设备宕机时，备份路由设备及时接管转发工作，向用户提供透明的切换，提高了网络服务质量。

3.3.1 VRRP

通过启动 VRRP 功能，当主路由设备发生故障时，备份路由设备可以在不影响内外数据通信的前提下进行功能切换，且不需要再修改内部网络的参数。

工作原理

图 3-3 VRRP 工作原理



VRRP 的运作方式

RFC 2338、RFC3768、RFC5798 中定义了 VRRP 类型的 IP 报文格式及其运作机制，VRRP 报文是一类指定目的地址的组播报文，该报文由主路由设备定时发出来标志其运行正常同时该报文也用于选举主路由设备。VRRP 允许为 IP 局域网承担路由转发功能的路由设备失效后，局域网中另外一个路由设备将自动接管失效的路由设备，从而实现 IP 路由的热备份与容错，同时也保证了局域网内主机通讯的连续性和可靠性。一个 VRRP 应用组通过多台路由设备来实现冗余，但是任何时候只有一台路由设备作为主路由设备来承担路由转发功能，其他的为备份路由设备，VRRP 应用组中不同路由设备间的切换对局域网内的主机则是完全透明的。

▾ 路由设备的切换规则

RFC 规定了路由设备的切换规则：

- VRRP 协议采用简单竞选的方法选择主路由设备。首先比较同一个 VRRP 组内的各台路由设备对应接口上设置的 VRRP 优先级的大小，优先级最大的为主路由设备，它的状态变为 Master。若路由设备的优先级相同，则比较对应网络接口的主 IP 地址大小，主 IP 地址大的就成为主路由设备，由它提供实际的路由转发服务。
- 主路由设备选出后，其它路由设备作为备份路由设备(状态变为 Backup)，并通过主路由设备定时发出的 VRRP 报文监测主路由设备的状态。当正常工作时，主路由设备会每隔一段时间发送一个 VRRP 组播报文，称为通告报文，以通知备份路由设备：主路由设备处于正常工作状态。如果组内的备份路由设备在设定的时间段没有接收到来自主路由设备的报文，则将自己状态转为 Master。当组内有多台状态为 Master 路由设备时，重复 1) 的竞选过程。通过这样一个过程就会将优先级最大的路由设备选成新的主路由设备，从而实现 VRRP 的备份功能。

一旦在一个 VRRP 备份组选举出它的主路由设备，局域网内的主机将通过主路由设备进行路由转发。

▾ 通迅过程

通迅过程可以由上图来说明。在图中，路由设备 R1 和 R2 均通过以太网口 Gi0/0 与局域网 192.168.12.0/24 连接，路由设备 R1 与 R2 的 Gi 0/0 接口上设置了 VRRP，局域网内的主机都以该 VRRP 组的虚拟路由设备 IP 地址作为默认网关。对于局域网内的主机而言，它们只能感受到由 VRRP 组的虚拟路由设备，而实际承担路由转发功能的 VRRP 组的主路由设备对它们而言则是透明的。譬如，局域网内的主机 PC 1 如果要与其它网络内的主机 PC 2 通讯，PC 1 会以虚拟路由设备为默认网关来发送通向 PC 2 的网络数据包，VRRP 组中的主路由设备在接收到该数据包后会将该数据包转发给 PC 2。在这个通迅过程中，PC 1 只能感受到虚拟路由设备而不知道扮演虚拟路由设备角色的主路由设备究竟是 R1 还是 R2，在这个 VRRP 组中的主路由设备是在 R1 与 R2 之间选举产生的，一旦主路由设备失效，那么另外一台将自动成为主路由设备。

相关配置

▾ 配置启动 VRRP 功能

缺省情况下，接口上 VRRP 功能关闭。

在接口模式下，使用 `vrrp group ip ipaddress [secondary]` 命令设置备份组号和虚拟 IP 地址来启用 VRRP 功能。

必须在接口上配置启动 VRRP 功能，才能参与 VRRP 协议工作。

▾ 设置 IPv4 VRRP 的验证字符串

缺省状态下，VRRP 处于无验证模式。

使用 `vrrp group authentication string` 命令可以设置 IPv4 VRRP 的验证字符串的同时也设定该 VRRP 组处于明文密码验证模式。在明文密码验证模式下，明文密码长度不能超过 8 个字节。

VRRP 备份组成员必须处于相同的验证模式下才可能正常通讯。明文密码验证模式下,在同一个 VRRP 组中的路由设备必须设置相同的验证口令。明文验证口令不能保证安全性,它只是用来防止/提示错误的 VRRP 配置。

设置 VRRP 备份组的通告发送间隔

缺省状态下,系统默认主路由设备的 VRRP 通告发送间隔为 1 秒。

使用 `vrrp group timers advertise { advertise-interval | csec centisecond-interval }` 可以调整 VRRP 备份组的通告发送间隔。

在没有设置 VRRP 定时设备学习功能的时候,同一个 VRRP 备份组要设置相同的 VRRP 通告发送间隔,否则处于备份状态的路由设备将会丢弃接收到的 VRRP 通告。

设置路由设备在 VRRP 备份组中的抢占模式

缺省情况下,VRRP 组工作在抢占模式下。

使用 `vrrp group preempt [delay seconds]` 设置 VRRP 组工作在抢占模式下,可选参数 `delay seconds` 缺省为 0 秒。

如果 VRRP 组工作在抢占模式下,一旦它发现自己的优先级高于当前 Master 的优先级,它将抢占成为该 VRRP 组的主路由设备。如果 VRRP 组工作在非抢占模式下,即便它发现自己的优先级高于当前 Master 的优先级,它也不会抢占成为该 VRRP 组的主路由设备。VRRP 组使用以太网接口 IP 地址情况下,抢占模式是否设置意义不大,因为此时该 VRRP 组具有最大优先级,它自动成为该 VRRP 组中的主路由设备。可选参数 Delay Seconds 定义了处于备份状态的 VRRP 路由设备准备宣告自己拥有 Master 身份之前的延迟。

设置路由设备在 VRRP 备份组中的优先级

缺省状态下,VRRP 组默认其优先级为 100。

使用 `vrrp group priority level` 可以调整 VRRP 组的优先级。

工作在抢占模式下具有最高优先级并且已获得虚拟 IP 地址的路由设备将成为该备份组的活动的(或主)路由设备,同一个备份组中低于该路由设备优先级的其它路由设备将成为备份的(或监听的)路由设备。

设置 VRRP 备份组监视的接口

缺省状态下,没有设置 VRRP 备份组监视的接口。

使用 `vrrp group track { interface-type interface-number | bfd interface-type interface-number ipv4-address } [priority]`。

配置了 VRRP 备份组监视的接口后,系统将根据所监视接口的状态动态地调整本路由设备的优先级。一旦所监视的接口状态变为不可用就按照设置的数值减少本路由设备在 VRRP 备份组中的优先级,而此时同一个备份组中接口状态更稳定并且优先级更高的其它路由设备就可以成为该 VRRP 备份组的活动的(或主)路由设备。

设置 VRRP 备份组监视的 IP 地址

缺省状态下,没有设置 VRRP 备份组监视的 IP 地址。

使用 `vrrp group track ip-address [interval interval-value] [timeout timeout-value] [retry retry-value] [priority] [interval interval-value] [timeout timeout-value] [retry retry-value] [priority]` 命令可以设置监视 IP 地址。

配置了 VRRP 备份组监视的 IP 地址后,系统将根据所监视的地址是否可达来动态地调整本设备的优先级。一旦所监视的 IP 地址变为不可达,即 ping 不通,就按照设置的数值减少本设备在 VRRP 备份组中的优先级,而此时同一个备份组中优先级更高的其它路由设备就可以成为该 VRRP 备份组的活动的(或主)路由设备。

设置 VRRP 通告定时设备学习功能

缺省状态下，系统没有为 VRRP 组设置定时设备学习功能。

使用 `vrrp group timers learn` 命令设置 VRRP 组启用定时设备学习功能。

一旦启用了定时设备学习功能，如果当前路由设备是 VRRP 备份路由设备，在设置了定时设备学习功能后，它会从主路由设备的 VRRP 通告中学习 VRRP 通告发送间隔，并由此来计算 Master 路由设备失效判断间隔，而不是使用自己本地设置的 VRRP 通告发送间隔来计算。本命令可以实现 Backup 路由设备与 Master 路由设备的 VRRP 通告发送定时设备同步。

设置路由设备在 VRRP 备份组的描述字符串

缺省状态下，系统没有设置 VRRP 组描述符。

使用 `vrrp group description text` 命令可以设置 VRRP 组描述符。

为 VRRP 组设置描述符，可以便于区分 VRRP 组。

设置 VRRP 备份组延迟启动

缺省状态下，系统没有设置 VRRP 备份组延迟启动。

使用 `vrrp delay { minimum min-seconds | reload reload-seconds }` 可以设置 VRRP 备份组延迟启动，两种 VRRP 备份组延迟启动时间的取值范围均为 0~60 秒。

本命令配置某个接口上 VRRP 备份组的延迟启动时间；延迟时间有两种：系统启动时的延迟时间，与接口状态变为活动时的延迟时间，可以分别配置，也可同时配置。配置本命令后，当系统启动，或者接口状态变为活动时，该接口上的 VRRP 备份组不会立即启动；而是等待相应的延迟时间后再启动 VRRP 备份组，保证非抢占配置不会失效。如果在延迟启动 VRRP 时该接口上接收到 VRRP 报文，则会取消延迟，立即启动 VRRP 协议。配置此命令将对接口的 IPv4 VRRP 备份组同时生效。

设置 IPv4 VRRP 的 VRRP 报文发送标准

缺省状态下，IPv4 VRRP 的 VRRP 报文使用 VRRPv2 标准。

使用 `vrrp group version { 2 | 3 }` 命令设置 IPv4 VRRP 的 VRRP 报文发送标准。

参数配置为 2，则 IPv4 VRRP 的 VRRP 报文使用 VRRPv2 标准；参数配置为 3，则 IPv4 VRRP 的 VRRP 报文使用 VRRPv3 标准。

设置 Super VLAN 中 IPv4 VRRP 报文的发送方式

缺省状态下，只往 Super VLAN 的第一个 UP 的 Sub VLAN 发送。

使用 `vrrp detection-vlan first-subvlan` 命令可以设置 Super VLAN 中 IPv4 VRRP 报文只往 Super VLAN 的第一个 UP 的 Sub VLAN 发送。Super VLAN 接口如果同时启用 VRRP 和 VRRP PLUS 功能，则不支持 VRRP 协议报文只往 Super VLAN 接口第一个 UP 的 Sub VLAN 口发送。

设置 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送，可以减少 VRRP 协议报文数量，避免影响设备的性能和占用网络带宽。但是要求组成 IPv4 VRRP 备份组的设备必须能够在该 Super VLAN 的第一个 UP 的 Sub VLAN 互通。

设置 IPv4 VRRP 组和 BFD 联动

缺省状态下，接口上没有指定 IPv4 VRRP 组和 BFD 联动。

使用 `vrrp group bfd ip-address` 命令可以设置 IPv4 VRRP 组和 BFD 联动。

对于备份路由器，执行这条命令，使该 IPv4 VRRP 组和 BFD 联动，并不关心配置的 IP 地址。而对于主路由器，由于不知道备份路由器接口的主 IP 地址，所以只能由管理员指定备份路由器的 IP 地址。

配置时首先确保配置的接口配置了 IP 和 BFD 会话参数。

通过指定 IPv4 VRRP 组和 BFD 联动，当主路由器出现故障时，可以把备份路由器检出主路由器故障的时间缩短到 1 秒钟以内。

📌 设置全局 IPv4 VRRP BFD

缺省状态下，VRRP 不采用全局 IPv4 VRRP BFD 方式检测 master 是否处于活动状态。

使用 `vrrp bfd interface-type interface-number ip-address` 命令可以设置全局 VRRP BFD。

配置全局 IPv4 VRRP BFD，使得多个 IPv4 VRRP 备份组可以共用该全局 BFD 会话，实现快速检测、主备快速切换。

配置时首先确保配置的接口配置了 IP 和 BFD 会话参数。

3.4 配置详解

配置项	配置建议 & 相关命令
配置IPv4 VRRP	 必须配置。用于启动 IPV4 VRRP 备份功能。
	<code>vrrp group ip ipaddress [secondary]</code> 启用 IPv4 VRRP
	 可选配置。用于配置 IPV4 VRRP 备份组参数。
	<code>vrrp group authentication string</code> 设置 IPv4 VRRP 的验证字符串
	<code>vrrp group timers advertise { advertise-interval csec centisecond-interval }</code> 设置 IPv4 VRRP 主路由设备 VRRP 通告间隔
	<code>vrrp group preempt [delay seconds]</code> 设置 Ipv4 VRRP 备份组处于抢占模式
	<code>vrrp group priority level</code> 设置 IPv4 VRRPVRRP 备份组的优先级
	<code>vrrp group track { interface-type interface-number bfd interface-type interface-number ipaddress } [priority]</code> 设置 IPv4 VRRP 备份组监视的接口
	<code>vrrp group track ip-address [interval interval-value] [timeout timeout-value] [retry retry-value] [priority]</code> 设置 IPv4 VRRP 备份组监视的 IP 地址
	<code>vrrp group timers learn</code> 设置 IPv4 VRRP 定时设备学习功能
	<code>vrrp group description text</code> 设置 IPv4 VRRP 组描述字符串
	<code>vrrp delay { minimum min-seconds reload reload-seconds }</code> 设置接口上 VRRP 备份组延迟启动时间
	<code>vrrp group version { 2 3 }</code> 设置 IPv4 VRRP 的 VRRP 报文标准
<code>vrrp detection-vlan first-subvlan</code> 设置 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送	

	<code>vrrp group bfd ip-address</code>	设置 IPv4 VRRP 组和 BFD 联动
	<code>vrrp bfd interface-type interface-number ip-address</code>	设置全局 IPv4 VRRP BFD
配置VRRP+MSTP	 VRRP 的基本配置跟 IPv4VRRP 配置一样。	

3.4.1 配置IPv4 VRRP

配置效果

- 配置单备份组时，通过设置备份组号和虚拟 IP 地址，可以在指定的局域网段上添加一个备份组，从而启动对应接口的 VRRP 单备份功能。
- 配置多备份组时，在同一个以太网接口上配置多个 VRRP 备份组，可以实现负载均衡同时通过互相备份来提供更稳定可靠的网络服务。
- 通过设置监视接口，动态的调整备份组的优先级，实现动态故障监测，实现备份组的状态切换。

注意事项

- 同一个 VRRP 备份组内路由设备上 VRRP 需要配置相同的虚拟 IPv4 地址。
- 实现 IPv4 VRRP 多备份组相互备份，需要通过在接口上配置多个相同的 IPv4 VRRP 备份组，且通过配置不同优先级等形式形成互为主备关系，从而实现 IPv4 VRRP 多备份组功能。
- 启动 VRRP 需要在三层口上配置。

配置方法

▾ 启用 IPv4 VRRP

- 默认接口上没有启用 IPv4 VRRP，如果接口要启用 IPv4 VRRP 备份功能则必须配置。

▾ 设置 IPv4 VRRP 的验证字符串

- 默认 VRRP 处于无验证模式，如果需要 VRRP 处于明文密码验证模式则配置。

▾ 设置 IPv4 VRRP 主路由设备 VRRP 通告间隔

- 默认 VRRP 的主路由设备 VRRP 通告间隔为 1 秒，如果需要手动设置则配置。

▾ 设置 Ipv4 VRRP 备份组处于抢占模式

- 默认 VRRP 处于抢占模式，抢占延迟时间缺省为 0 秒。

▾ 设置 IPv4 VRRP 备份组的优先级

- 默认 VRRP 组优先级为 100，如果需要手动设置则配置。
- ↳ **设置 IPv4 VRRP 备份组监视的接口**
- 默认 IPv4 VRRP 组未配置监视接口，缺省优先级改变值为 10，如果需要通过监视接口来实现故障检测则配置。
- ↳ **设置 IPv4 VRRP 定时设备学习功能**
- 默认 VRRP 组未开启定时设备学习功能，如果备份路由设备需要学习主路由设备的 VRRP 通告间隔则配置。
- ↳ **设置 IPv4 VRRP 组描述字符串**
- 默认 VRRP 组未配置描述字符串，如果需要便于区分 VRRP 组则配置。
- ↳ **设置接口上 VRRP 备份组延迟启动时间**
- 默认 VRRP 组未配置延迟启动时间，如果为了保证非抢占模式不失效则配置。
- ↳ **设置 IPv4 VRRP 的 VRRP 报文标准**
- 默认 IPv4 VRRP 使用 VRRPv2 标准，如果需要手动设置则配置。
- ↳ **设置 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送**
- 默认 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送，如果需要手动设置则配置。
- ↳ **设置 IPv4 VRRP 组和 BFD 联动**
- 默认接口上没有指定 IPv4 VRRP 组和 BFD 联动，如果需要手动设置则配置。
- ↳ **设置全局 IPv4 VRRP BFD**
- 默认 VRRP 不采用全局 IPv4 VRRP BFD 方式检测 master 是否处于活动状态，如果需要手动设置则配置。

检验方法

- `show vrrp` 命令显示配置是否生效

相关命令

↳ 启用 IPv4 VRRP

【命令格式】 `vrrp group ip ipaddress [secondary]`

【参数说明】 `group`：VRRP 组号，不同产品型号取值范围不同。

`ipaddress`：虚拟设备的 IP 地址。

`secondary`：标明是该虚拟设备的次 IP 地址。

【命令模式】 接口模式

【使用指导】 如果不指定虚拟 IP 地址，路由设备就不会参与 VRRP 备份组。如果不使用 Secondary 参数，那么设置的 IP 地址将成为虚拟路由设备的主 IP 地址。

设置 IPv4 VRRP 的验证字符串

【命令格式】 **vrrp group authentication string**

【参数说明】 *group* : VRRP 组号。

string : 用于 VRRP 组验证的字符串(不能超过 8 个字节, 这里的验证口令是明文口令)。

【命令模式】 接口模式

【使用指导】 在同一个 VRRP 组中的设备必须设置相同的验证口令。明文验证口令不能保证安全性, 它只是用来防止/提示错误的 VRRP 配置。此命令只对 VRRPv2 报文适用, 对于 VRRPv3 不适用。

设置 IPv4 VRRP 主路由设备 VRRP 通告间隔

【命令格式】 **vrrp group timers advertise { advertise-interval | csec centisecond-interval }**

【参数说明】 *group* : VRRP 组号。

advertise-interval : VRRP 通告发送间隔(以秒为单位)。

csec centisecond-interval : 备份组中 Master 发送 VRRP 报文的时间间隔。整数形式, 取值范围是 50 ~ 99。单位是厘秒。无缺省值。只对 VRRPv3 生效, 如果 VRRPv2 配置了此命令, 则取默认的主设备的通告发送间隔, 即 1 秒。

【命令模式】 接口模式

【使用指导】 如果当前设备成为 VRRP 组中的主设备, 它将以设定的间隔发送 VRRP 通告来通告自己的 VRRP 状态、优先级以及其它信息。

根据 RFC 标准, 配置了组播报文发送标准为 VRRPv3 的 IPv4 VRRP 组, 其组播报文中最大的报文通告间隔为 40 秒, 所以如果配置的报文通告间隔超过 40 秒, 则取最大的通告间隔 40 秒, 但该通告间隔配置是生效的。

设置 IPv4 VRRP 备份组处于抢占模式

【命令格式】 **vrrp group preempt [delay seconds]**

【参数说明】 *group* : VRRP 组号。

delay seconds : 准备宣告自己拥有 Master 身份之前的延迟。缺省值为 0 秒。

【命令模式】 接口模式

【使用指导】 如果 VRRP 组工作在抢占模式下, 一旦它发现自己的优先级高于当前 Master 的优先级, 它将抢占成为该 VRRP 组的主设备。如果 VRRP 组工作在非抢占模式下, 即便它发现自己的优先级高于当前 Master 的优先级, 它也不会抢占成为该 VRRP 组的主设备。VRRP 组使用以太网接口 IP 地址情况下, 抢占模式是否设置意义不大, 因为此时该 VRRP 组具有最大优先级, 它自动成为该 VRRP 组中的主设备。

设置 IPv4 VRRP 备份组的优先级

【命令格式】 **vrrp group priority level**

【参数说明】 *group* : VRRP 组号。

level : VRRP 组的优先级。

【命令模式】 接口模式

【使用指导】 该命令将手动设置 VRRP 组的优先级。

设置 IPv4 VRRP 备份组监视的接口

【命令格式】 **vrrp group track { interface-type interface-number | bfd interface-type interface-number ipv4-address }**

[*priority*]

【参数说明】 *group* : VRRP 组号。

interface-type interface-number : 被监视的接口。

bfd *interface-type interface-number ipv4-address* : 通过 BFD 跟踪指定的邻居 IP。

priority : 被监视的接口状态改变时其 VRRP 优先级改变的尺度。缺省值为 10。

【命令模式】 接口模式

【使用指导】 被监视的接口只允许是三层可路由的逻辑接口(如 Routed Port、SVI、Loopback、Tunnel 等等)。

如果 VRRP 组占用(Own)了以太网接口实际 IP 地址, 此时该 VRRP 组的优先级为 255, 不能配置监视接口。

设置 IPv4 VRRP 备份组监视的 IP 地址

【命令格式】 **vrrp group track ipv4-address [interval interval-value] [timeout timeout-value] [retry retry-value] [priority]**

【参数说明】 *group* : VRRP 组号。

ipv4-address : 被监视的 IPv4 地址。

interval *interval-value* : 发送探测报文的时间间隔。单位为秒。如果不设定, 系统缺省值为 3 秒。

timeout *timeout-value* : 发送探测报文后等待应答的超时时间。如果超时时间到了, 没有收到应答, 则认为不可达。单位为秒。如果不设定, 系统缺省值为 1 秒。

retry *retry-value* : 确认不可达的次数, 如果在连续 *retry-value* 次都没有收到应答, 则认为不可达。单位为次数, 如果不设定, 系统缺省值为 3 次。

priority : 被监视的接口状态改变时其 VRRP 优先级改变的尺度。缺省值为 10。

【命令模式】 接口模式

【使用指导】 如果是监视主机, 对于 IPv4 虚拟路由器, 指定主机的 IPv4 地址。

如果 VRRP 组占用(Own)了以太网接口实际 IP 地址, 此时该 VRRP 组的优先级为 255, 不能配置监视 IP 地址。

设置 IPv4 VRRP 定时设备学习功能

【命令格式】 **vrrp group timers learn**

【参数说明】 *group* : VRRP 组号。

【命令模式】 接口模式

【使用指导】 一旦启用了定时器学习功能, 如果当前设备是 VRRP 备份设备, 在设置了定时器学习功能后, 它会从主设备的 VRRP 通告中学习 VRRP 通告发送间隔, 并由此来计算 Master 设备失效间隔, 而不是使用自己本地设置的 VRRP 通告发送间隔来计算。本命令可以实现与 Master 设备的 VRRP 通告发送定时器同步。

设置 IPv4 VRRP 组描述字符串

【命令格式】 **vrrp group description text**

【参数说明】 *group* : VRRP 组号。

text : VRRP 组描述符。

【命令模式】 接口模式

【使用指导】 为 VRRP 组设置描述符, 可以便于区分 VRRP 组。

设置接口上 VRRP 备份组延迟启动时间

- 【命令格式】 **vrrp delay { minimum *min-seconds* | reload *reload-seconds* }**
- 【参数说明】 **minimum *min-seconds*** : 接口状态变为活动时的延迟时间。
reload *reload-seconds* : 系统启动时的延迟时间。
- 【命令模式】 接口模式
- 【使用指导】 配置本命令后, 当系统启动, 或者接口状态变为活动时, 该接口上的 VRRP 备份组不会立即启动; 而是等待相应的延迟时间后再启动 VRRP 备份组, 保证非抢占配置不会失效。如果在延迟启动 VRRP 时该接口上接收到 VRRP 报文, 则会取消延迟, 立即启动 VRRP 协议。两种 VRRP 备份组延迟启动时间的取值范围均为 0~60 秒。

设置 IPv4 VRRP 的 VRRP 报文标准

- 【命令格式】 **vrrp group version { 2 | 3 }**
- 【参数说明】 **2** : 使用 VRRPv2 报文发送标准。
3 : 使用 VRRPv3 报文发送标准。
- 【命令模式】 接口模式
- 【使用指导】 对于 IPv4 VRRP 考虑到 VRRPv2 和 VRRPv3 的兼容性问题, 用户可以根据网络实际环境选择 VRRP 报文的发送标准。VRRPv2 基于 RFC3768, VRRPv3 基于 RFC 5798。
此命令只适用于 IPv4 VRRP。

设置 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送

- 【命令格式】 **vrrp detection-vlan first-subvlan**
- 【参数说明】 **first-subvlan** : 只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送
- 【命令模式】 接口模式
- 【使用指导】 本命令用于配置 Super VLAN 接口中 IPv4 VRRP 协议报文的发送方式。IPv4 VRRP 协议报文在 Super VLAN 中的发送方式包括两种 : 只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送 ; 往 Super VLAN 中的所有 Sub VLAN 发送。Super VLAN 接口如果同时启用 VRRP 和 VRRP PLUS 功能, 则不支持 VRRP 协议报文只往 Super VLAN 接口第一个 UP 的 Sub VLAN 口发送。
本命令在 VLAN 接口上配置, 只对 Super VLAN 口生效。

设置 IPv4 VRRP 组和 BFD 联动

- 【命令格式】 **vrrp group bfd *ip-address***
- 【参数说明】 **group** : VRRP 组号。
ip-address : 指定的邻居 IP。
- 【命令模式】 接口模式
- 【使用指导】 对于备份路由器, 执行这条命令, 使该 IPv4 VRRP 组和 BFD 联动, 并不关心配置的 IP 地址。而对于主路由器, 由于不知道备份路由器接口的主 IP 地址, 所以只能由管理员指定备份路由器的 IP 地址。
如果配置了全局 IPv4 VRRP BFD, 不能配置 IPv4 VRRP 组和 BFD 联动。
配置时首先确保配置的接口配置了 IP 和 BFD 会话参数。

设置全局 IPv4 VRRP BFD

- 【命令格式】 **vrrp bfd *interface-type interface-number ip-address***
- 【参数说明】 ***interface-type interface-number*** : 配置接口类型和接口编号。

ip-address : 指定的邻居 IP。

- 【命令模式】 全局配置模式
- 【使用指导】 如果配置了全局 IPv4 VRRP BFD，会删除所有配置的 IPv4 VRRP 组和 BFD 联动。
配置时首先确保配置的接口配置了 IP 和 BFD 会话参数。
全局 IPv4 VRRP BFD 会话只适用于两台设备组成的 IPv4 VRRP 虚拟路由器。

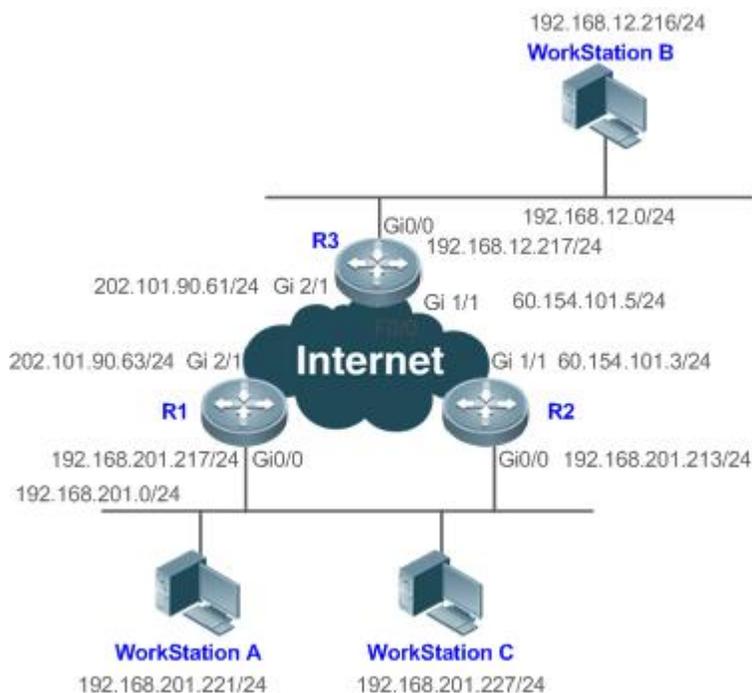
配置举例

i 以下配置举例，仅介绍与 VRRP 相关的配置。

在 IPv4 的 VRRP 单备份组和监视接口

【网络环境】

图 3-4



- 【配置方法】
- 用户工作站群(192.168.201.0/24)使用路由设备 R1 与 R2 组成的备份组，并将其网关指向该备份组设置的虚拟路由设备的 IP 地址 192.168.201.1，经由虚拟路由设备 192.168.201.1 访问远程用户工作站群(其工作网络为 192.168.12.0 /24)。
 - 路由设备 R1 中设置了 VRRP 监视接口 GigabitEthernet 2/1。
 - 路由设备 R3 上没有配置 VRRP 而只是配置了普通路由功能。

R3

```
R3#configure terminal
R3(config)#interface GigabitEthernet 0/0
// "no switchport"命令只有在交换机上才需要
R3(config-if-GigabitEthernet 0/0)#no switchport
R3(config-if-GigabitEthernet 0/0)#ip address 192.168.12.217 255.255.255.0
R3(config-if-GigabitEthernet 0/0)#exit
R3(config)#interface GigabitEthernet 1/1
```

```
// “no switchport”命令只有在交换机上才需要
R3(config-if-GigabitEthernet 1/1)#no switchport
R3(config-if-GigabitEthernet 1/1)#ip address 60.154.101.5 255.255.255.0
R3(config-if-GigabitEthernet 1/1)#exit
R3(config)#interface GigabitEthernet 2/1
// “no switchport”命令只有在交换机上才需要
R3(config-if-GigabitEthernet 2/1)#no switchport
R3(config-if-GigabitEthernet 2/1)#ip address 202.101.90.61 255.255.255.0
R3(config-if-GigabitEthernet 2/1)#exit
R3(config)#router ospf
R3(config-router)#network 202.101.90.0 0.0.0.255 area 10
R3(config-router)#network 192.168.12.0 0.0.0.255 area 10
R3(config-router)#network 60.154.101.0 0.0.0.255 area 10
```

R1

```
R1#configure terminal
R1(config)#interface GigabitEthernet 0/0
R1(config-if-GigabitEthernet 0/0)#ip address 192.168.201.217 255.255.255.0
R1(config-if-GigabitEthernet 0/0)#vrrp 1 priority 120
R1(config-if-GigabitEthernet 0/0)#vrrp 1 version 3
R1(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3
R1(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1
R1(config-if-GigabitEthernet 0/0)#vrrp 1 track GigabitEthernet 2/1 30
R1(config-if-GigabitEthernet 0/0)#exit
R1(config)#interface GigabitEthernet 2/1
R1(config-if-GigabitEthernet 2/1)#ip address 202.101.90.63 255.255.255.0
R1(config-if-GigabitEthernet 2/1)#exit
R1(config)#router ospf
R1(config-router)#network 202.101.90.0 0.0.0.255 area 10
R1(config-router)#network 192.168.201.0 0.0.0.255 area 10
```

R2

```
R2#configure terminal
R2(config)#interface GigabitEthernet 0/0
R2(config-if-GigabitEthernet 0/0)#ip address 192.168.201.213 255.255.255.0
R2(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1
R2(config-if-GigabitEthernet 0/0)#vrrp 1 version 3
R2(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3
R2(config-if-GigabitEthernet 0/0)#exit
R2(config)#interface GigabitEthernet 1/1
// “no switchport”命令只有在交换机上才需要
R2(config-if-GigabitEthernet 1/1)#no switchport
R2(config-if-GigabitEthernet 1/1)#ip address 60.154.101.3 255.255.255.0
R2(config-if-GigabitEthernet 1/1)#exit
R2(config)#router ospf
```

```
R2(config-router)#network 60.154.101.0 0.0.0.255 area 10
R2(config-router)#network 192.168.201.0 0.0.0.255 area 10
```

【检验方法】 通过 **show vrrp** 命令显示验证。

- 检查当路由设备 R1 在作为 Master 路由设备状态下发现与广域网的接口 GigabitEthernet 2/1 不可用，路由设备 R1 是否降低自己的 VRRP 备份组优先级 30 而成为 90，这样路由设备 R2 就会成为 Master 路由设备。
- 检查当路由设备 R1 发现自己的与广域网的接口 GigabitEthernet 2/1 恢复可用，是否增加自己的 VRRP 备份组优先级 30 而恢复到 120，这样路由设备 R1 将再次成为主路由设备。

R1

```
R1#show vrrp
GigabitEthernet 0/0 - Group 1
  State is Master
  Virtual IP address is 192.168.201.1 configured
  Virtual MAC address is 0000.5e00.0101
  VRRP standard version is V3
  Advertisement interval is 3 sec
  Preemption is enabled
    min delay is 0 sec
  Priority is 120
  Master Router is 192.168.201.217 (local), priority is 120
  Master Advertisement interval is 3 sec
  Master Down interval is 10.59 sec
  Tracking state of 1 interface, 1 up:
    up  GigabitEthernet 2/1 priority decrement=30
```

R2

```
R2#show vrrp
GigabitEthernet 0/0 - Group 1
  State is Backup
  Virtual IP address is 192.168.201.1 configured
  Virtual MAC address is 0000.5e00.0101
  VRRP standard version is V3
  Advertisement interval is 3 sec
  Preemption is enabled
    min delay is 0 sec
  Priority is 100
  Master Router is 192.168.201.217 , priority is 120
  Master Advertisement interval is 3 sec
  Master Down interval is 10.82 sec
```

常见错误

- 同一个 VRRP 备份组内路由设备上 VRRP 的虚拟 IP 地址不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。

- 同一个 VRRP 备份组内路由设备上 VRRP 的通告发送间隔不一致,并且未设置定时设备学习功能导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 同一个 VRRP 备份组内路由设备上 VRRP 协议报文版本不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 对于 VRRPv2,同一 VRRP 备份组内各路由设备以太网接口上 VRRP 组验证模式相同均为明文密码模式,但是验证字符串不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。

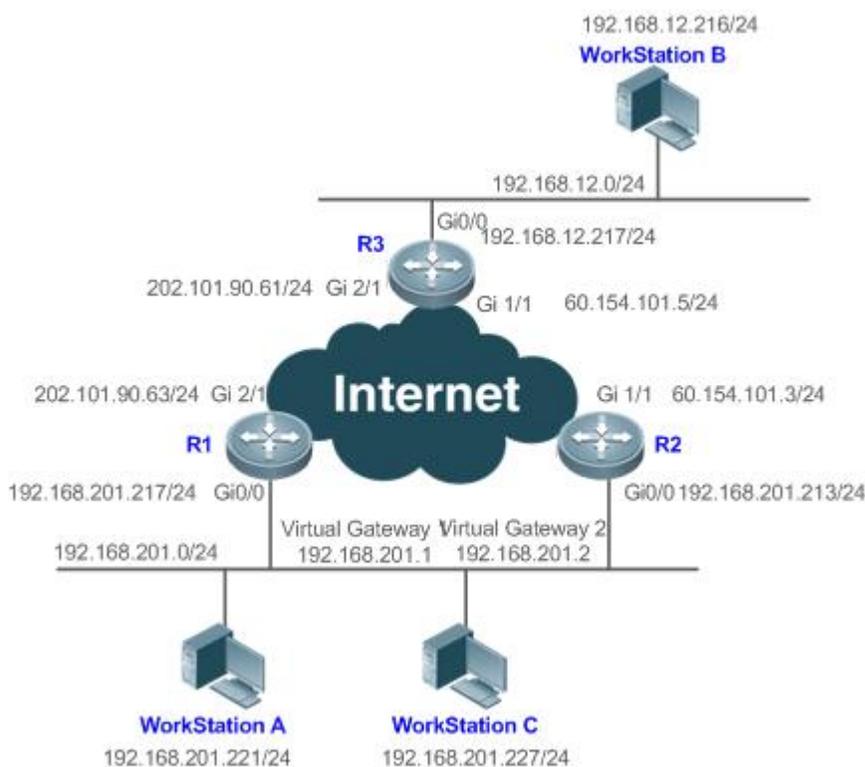
配置举例

i 以下配置举例,仅介绍与 VRRP 相关的配置。

在 IPv4 的 VRRP 多备份组

【网络环境】

图 3-5



【配置方法】

- 用户工作站群(192.168.201.0/24)使用路由设备 R1 与 R2 组成的备份组,其中部分用户工作站(如 A)将其网关指向备份组 1 的虚拟 IP 地址 192.168.201.1,部分用户工作站(如 C)则将其网关指向备份组 2 的虚拟 IP 地址 192.168.201.2。在所有路由器上启动 IPv4 组播路由功能。
- 路由设备 R1 在备份组 2 中作为主路由设备,在备份组 1 中作为备份路由设备。
- 路由设备 R2 在备份组 2 中作为备份路由设备,在备份组 1 中作为主路由设备。

R3

```
R3#configure terminal
R3(config)#interface GigabitEthernet 0/0
// "no switchport"命令只有在交换机上才需要
```

```
R3(config-if-GigabitEthernet 0/0)#no switchport
R3(config-if-GigabitEthernet 0/0)#ip address 192.168.12.217 255.255.255.0
R3(config-if-GigabitEthernet 0/0)#exit
R3(config)#interface GigabitEthernet 1/1
// “no switchport”命令只有在交换机上才需要
R3(config-if-GigabitEthernet 1/1)#no switchport
R3(config-if-GigabitEthernet 1/1)#ip address 60.154.101.5 255.255.255.0
R3(config-if-GigabitEthernet 1/1)#exit
R3(config)#interface GigabitEthernet 2/1
// “no switchport”命令只有在交换机上才需要
R3(config-if-GigabitEthernet 2/1)#no switchport
R3(config-if-GigabitEthernet 2/1)#ip address 202.101.90.61 255.255.255.0
R3(config-if-GigabitEthernet 2/1)#exit
R3(config)#router ospf
R3(config-router)#network 202.101.90.0 0.0.0.255 area 10
R3(config-router)#network 192.168.12.0 0.0.0.255 area 10
R3(config-router)#network 60.154.101.0 0.0.0.255 area 10
```

R1

```
R1#configure terminal
R1(config)#interface GigabitEthernet 0/0
R1(config-if-GigabitEthernet 0/0)#ip address 192.168.201.217 255.255.255.0
R1(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3
R1(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1
R1(config-if-GigabitEthernet 0/0)#vrrp 2 priority 120
R1(config-if-GigabitEthernet 0/0)#vrrp 2 timers advertise 3
R1(config-if-GigabitEthernet 0/0)#vrrp 2 ip 192.168.201.2
R1(config-if-GigabitEthernet 0/0)#vrrp 2 track GigabitEthernet 2/1 30
R1(config-if-GigabitEthernet 0/0)#exit
R1(config)#interface GigabitEthernet 2/1
R1(config-if-GigabitEthernet 2/1)#ip address 202.101.90.63 255.255.255.0
R1(config-if-GigabitEthernet 2/1)#exit
R1(config)#router ospf
R1(config-router)#network 202.101.90.0 0.0.0.255 area 10
R1(config-router)#network 192.168.201.0 0.0.0.255 area 10
```

R2

```
R2#configure terminal
R2(config)#interface GigabitEthernet 0/0
R2(config-if-GigabitEthernet 0/0)#ip address 192.168.201.213 255.255.255.0
R2(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.201.1
R2(config-if-GigabitEthernet 0/0)#vrrp 1 timers advertise 3
R2(config-if-GigabitEthernet 0/0)#vrrp 1 priority 120
R2(config-if-GigabitEthernet 0/0)#vrrp 2 ip 192.168.201.2
R2(config-if-GigabitEthernet 0/0)#vrrp 2 timers advertise 3
```

```
R2(config-if-GigabitEthernet 0/0)#exit
R2(config)#interface GigabitEthernet 1/1
R2(config-if-GigabitEthernet 1/1)#ip address 60.154.101.3 255.255.255.0
R2(config-if-GigabitEthernet 1/1)#exit
R2(config)#router ospf
R2(config-router)#network 60.154.101.0 0.0.0.255 area 10
R2(config-router)#network 192.168.201.0 0.0.0.255 area 10
```

【检验方法】 通过 **show vrrp** 命令显示验证。

- 检查当路由设备 R1 在 VRRP 组 2 中作为 Master 路由设备状态下发现与广域网的接口 GigabitEthernet 2/1 不可用，路由设备 R1 是否降低自己的 VRRP 备份组优先级 30 而成为 90，这样 VRRP 组 2 中路由设备 R2 就会成为 Master 路由设备。
- 检查当路由设备 R1 发现自己的与广域网的接口 GigabitEthernet 2/1 恢复可用，是否增加自己的 VRRP 备份组优先级 30 而恢复到 120，这样路由设备 R1 将再次成为 VRRP 组 2 主路由设备。

R1

```
R1#show vrrp
GigabitEthernet 0/0 - Group 1
  State is Backup
  Virtual IP address is 192.168.201.1 configured
  Virtual MAC address is 0000.5e00.0101
  Advertisement interval is 3 sec
  Preemption is enabled
    min delay is 0 sec
  Priority is 100
  Master Router is 192.168.201.213 , priority is 120
  Master Advertisement interval is 3 sec
  Master Down interval is 10.82 sec
GigabitEthernet 0/0 - Group 2
  State is Master
  Virtual IP address is 192.168.201.2 configured
  Virtual MAC address is 0000.5e00.0102
  Advertisement interval is 3 sec
  Preemption is enabled
    min delay is 0 sec
  Priority is 120
  Master Router is 192.168.201.217 (local), priority is 120
  Master Advertisement interval is 3 sec
  Master Down interval is 10.59 sec
  Tracking state of 1 interface, 1 up:
    up  GigabitEthernet 2/1 priority decrement=30
```

R2

```
R2#show vrrp
GigabitEthernet 0/0 - Group 1
```

```
State is Master
Virtual IP address is 192.168.201.1 configured
Virtual MAC address is 0000.5e00.0101
Advertisement interval is 3 sec
Preemption is enabled
  min delay is 0 sec
Priority is 120
Master Router is 192.168.201.213 (local), priority is 120
Master Advertisement interval is 3 sec
Master Down interval is 10.59 sec
GigabitEthernet 0/0 - Group 2
State is Backup
Virtual IP address is 192.168.201.2 configured
Virtual MAC address is 0000.5e00.0102
Advertisement interval is 3 sec
Preemption is enabled
  min delay is 0 sec
Priority is 100
Master Router is 192.168.201.217 , priority is 120
Master Advertisement interval is 3 sec
Master Down interval is 10.82 sec
```

常见错误

- 同一个 VRRP 备份组内路由设备上 VRRP 的虚拟 IP 地址不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 同一个 VRRP 备份组内路由设备上 VRRP 的通告发送间隔不一致,并且未设置定时设备学习功能导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 同一个 VRRP 备份组内路由设备上 VRRP 协议报文版本不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 对于 VRRPv2,同一 VRRP 备份组内各路由设备以太网接口上 VRRP 组验证模式相同均为明文密码模式,但是验证字符串不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。

3.4.2 配置VRRP+MSTP

配置效果

- MTSP 和 VRRP 同时使用可以达到链路级备份和网关级备份,极大提高网络健壮性。

注意事项

- 同一个 VRRP 备份组内路由设备上 VRRP 需要配置相同的虚拟 IPv4 地址。
- 启动 VRRP 需要在三层口上配置。

配置方法

▾ 启用 IPv4 VRRP

- 默认接口上没有启用 IPv4 VRRP，如果接口要启用 IPv4 VRRP 备份功能则必须配置。

▾ 设置 IPv4 VRRP 的验证字符串

- 默认 VRRP 处于无验证模式，如果需要 VRRP 处于明文密码验证模式则配置。

▾ 设置 IPv4 VRRP 主路由设备 VRRP 通告间隔

- 默认 VRRP 的主路由设备 VRRP 通告间隔为 1 秒，如果需要手动设置则配置。

▾ 设置 IPv4 VRRP 备份组处于抢占模式

- 默认 VRRP 处于抢占模式，抢占延迟时间缺省为 0 秒。

▾ 设置 IPv4 VRRP 备份组的优先级

- 默认 VRRP 组优先级为 100，如果需要手动设置则配置。

▾ 设置 IPv4 VRRP 备份组监视的接口

- 默认 IPv4 VRRP 组未配置监视接口，缺省优先级改变值为 10，如果需要通过监视接口来实现故障检测则配置。

▾ 设置 IPv4 VRRP 定时设备学习功能

- 默认 VRRP 组未开启定时设备学习功能，如果备份路由设备需要学习主路由设备的 VRRP 通告间隔则配置。

▾ 设置 IPv4 VRRP 组描述字符串

- 默认 VRRP 组未配置描述字符串，如果需要便于区分 VRRP 组则配置。

▾ 设置接口上 VRRP 备份组延迟启动时间

- 默认 VRRP 组未配置延迟启动时间，如果为了保证非抢占模式不失效则配置。

▾ 设置 IPv4 VRRP 的 VRRP 报文标准

- 默认 IPv4 VRRP 使用 VRRPv2 标准，如果需要手动设置则配置。

▾ 设置 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送

- 默认 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送，如果需要手动设置则配置。

▾ 设置 IPv4 VRRP 组和 BFD 联动

- 默认接口上没有指定 IPv4 VRRP 组和 BFD 联动，如果需要手动设置则配置。

设置全局 IPv4 VRRP BFD

- 默认 VRRP 不采用全局 IPv4 VRRP BFD 方式检测 master 是否处于活动状态，如果需要手动设置则配置。

检验方法

- `show vrrp` 命令显示配置是否生效

相关命令

启用 IPv4 VRRP

【命令格式】 `vrrp group ip ipaddress [secondary]`

【参数说明】 `group` : VRRP 组号，不同产品型号取值范围不同。

`ipaddress` : 虚拟设备的 IP 地址。

`secondary` : 标明是该虚拟设备的次 IP 地址。

【命令模式】 接口模式

【使用指导】 如果不指定虚拟 IP 地址，路由设备就不会参与 VRRP 备份组。如果不使用 Secondary 参数，那么设置的 IP 地址将成为虚拟路由设备的主 IP 地址。

设置 IPv4 VRRP 的验证字符串

【命令格式】 `vrrp group authentication string`

【参数说明】 `group` : VRRP 组号。

`string` : 用于 VRRP 组验证的字符串(不能超过 8 个字节，这里的验证口令是明文口令)。

【命令模式】 接口模式

【使用指导】 在同一个 VRRP 组中的设备必须设置相同的验证口令。明文验证口令不能保证安全性，它只是用来防止/提示错误的 VRRP 配置。此命令只对 VRRPv2 报文适用，对于 VRRPv3 不适用。

设置 IPv4 VRRP 主路由设备 VRRP 通告间隔

【命令格式】 `vrrp group timers advertise { advertise-interval | csec centisecond-interval }`

【参数说明】 `group` : VRRP 组号。

`advertise-interval` : VRRP 通告发送间隔(以秒为单位)。

`csec centisecond-interval` : 备份组中 Master 发送 VRRP 报文的时间间隔。整数形式，取值范围是 50 ~ 99。

单位是厘秒。无缺省值。只对 VRRPv3 生效，如果 VRRPv2 配置了此命令，则取默认的主设备的通告发送间隔，即 1 秒。

【命令模式】 接口模式

【使用指导】 如果当前设备成为 VRRP 组中的主设备，它将以设定的间隔发送 VRRP 通告来通告自己的 VRRP 状态、优先级及其它信息。

根据 RFC 标准，配置了组播报文发送标准为 VRRPv3 的 IPv4 VRRP 组，其组播报文中最大的报文通告间隔为 40 秒，所以如果配置的报文通告间隔超过 40 秒，则取最大的通告间隔 40 秒，但该通告间隔配置是生效的。

设置 IPv4 VRRP 备份组处于抢占模式

【命令格式】 **vrrp group preempt** [**delay seconds**]

【参数说明】 *group* : VRRP 组号。

delay seconds : 准备宣告自己拥有 Master 身份之前的延迟。缺省值为 0 秒。

【命令模式】 接口模式

【使用指导】 如果 VRRP 组工作在抢占模式下,一旦它发现自己的优先级高于当前 Master 的优先级,它将抢占成为该 VRRP 组的主设备。如果 VRRP 组工作在非抢占模式下,即便它发现自己的优先级高于当前 Master 的优先级,它也不会抢占成为该 VRRP 组的主设备。VRRP 组使用以太网接口 IP 地址情况下,抢占模式是否设置意义不大,因为此时该 VRRP 组具有最大优先级,它自动成为该 VRRP 组中的主设备。

设置 IPv4 VRRP 备份组的优先级

【命令格式】 **vrrp group priority level**

【参数说明】 *group* : VRRP 组号。

level : VRRP 组的优先级。

【命令模式】 接口模式

【使用指导】 该命令将手动设置 VRRP 组的优先级。

设置 IPv4 VRRP 备份组监视的接口

【命令格式】 **vrrp group track** { *interface-type interface-number* | **bfd** *interface-type interface-number ipv4-address* } [*priority*]

【参数说明】 *group* : VRRP 组号。

interface-type interface-number : 被监视的接口。

bfd *interface-type interface-number ipv4-address* : 通过 BFD 跟踪指定的邻居 IP。

priority : 被监视的接口状态改变时其 VRRP 优先级改变的尺度。缺省值为 10。

【命令模式】 接口模式

【使用指导】 被监视的接口只允许是三层可路由的逻辑接口(如 Routed Port、SVI、Loopback、Tunnel 等等)。如果 VRRP 组占用(Own)了以太网接口实际 IP 地址,此时该 VRRP 组的优先级为 255,不能配置监视接口。

设置 IPv4 VRRP 备份组监视的 IP 地址

【命令格式】 **vrrp group track ipv4-address** [**interval interval-value**] [**timeout timeout-value**] [**retry retry-value**] [*priority*]

【参数说明】 *group* : VRRP 组号。

ipv4-address : 被监视的 IPv4 地址。

interval interval-value : 发送探测报文的时间间隔。单位为秒。如果不设定,系统缺省值为 3 秒。

timeout timeout-value : 发送探测报文后等待应答的超时时间。如果超时时间到了,没有收到应答,则认为不可达。单位为秒。如果不设定,系统缺省值为 1 秒。

retry retry-value : 确认不可达的次数,如果在连续 *retry-value* 次都没有收到应答,则认为不可达。单位为次数,如果不设定,系统缺省值为 3 次。

priority : 被监视的接口状态改变时其 VRRP 优先级改变的尺度。缺省值为 10。

【命令模式】 接口模式

【使用指导】 如果是监视主机,对于 IPv4 虚拟路由器,指定主机的 IPv4 地址。

如果 VRRP 组占用(Own)了以太网接口实际 IP 地址，此时该 VRRP 组的优先级为 255，不能配置监视 IP 地址。

📌 设置 IPv4 VRRP 定时设备学习功能

【命令格式】 **vrrp group timers learn**

【参数说明】 *group* : VRRP 组号。

【命令模式】 接口模式

【使用指导】 一旦启用了定时器学习功能，如果当前设备是 VRRP 备份设备，在设置了定时器学习功能后，它会从主设备的 VRRP 通告中学习 VRRP 通告发送间隔，并由此来计算 Master 设备失效间隔，而不是使用自己本地设置的 VRRP 通告发送间隔来计算。本命令可以实现与 Master 设备的 VRRP 通告发送定时器同步。

📌 设置 IPv4 VRRP 组描述字符串

【命令格式】 **vrrp group description text**

【参数说明】 *group* : VRRP 组号。

text : VRRP 组描述符。

【命令模式】 接口模式

【使用指导】 为 VRRP 组设置描述符，可以便于区分 VRRP 组。

📌 设置接口上 VRRP 备份组延迟启动时间

【命令格式】 **vrrp delay { minimum *min-seconds* | reload *reload-seconds* }**

【参数说明】 **minimum** *min-seconds* : 接口状态变为活动时的延迟时间。

reload *reload-seconds* : 系统启动时的延迟时间。

【命令模式】 接口模式

【使用指导】 配置本命令后，当系统启动，或者接口状态变为活动时，该接口上的 VRRP 备份组不会立即启动；而是等待相应的延迟时间后再启动 VRRP 备份组，保证非抢占配置不会失效。如果在延迟启动 VRRP 时该接口上接收到 VRRP 报文，则会取消延迟，立即启动 VRRP 协议。两种 VRRP 备份组延迟启动时间的取值范围均为 0~60 秒。

📌 设置 IPv4 VRRP 的 VRRP 报文标准

【命令格式】 **vrrp group version { 2 | 3 }**

【参数说明】 **2** : 使用 VRRPv2 报文发送标准。

3 : 使用 VRRPv3 报文发送标准。

【命令模式】 接口模式

【使用指导】 对于 IPv4 VRRP 考虑到 VRRPv2 和 VRRPv3 的兼容性问题，用户可以根据网络实际环境选择 VRRP 报文的发送标准。VRRPv2 基于 RFC3768，VRRPv3 基于 RFC 5798。
此命令只适用于 IPv4 VRRP。

📌 设置 IPv4 VRRP 协议报文只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送

【命令格式】 **vrrp detection-vlan first-subvlan**

【参数说明】 **first-subvlan** : 只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送

【命令模式】 接口模式

- 【使用指导】 本命令用于配置 Super VLAN 接口中 IPv4 VRRP 协议报文的发送方式。IPv4 VRRP 协议报文在 Super VLAN 中的发送方式包括两种：只往 Super VLAN 中的第一个 UP 的 Sub VLAN 发送；往 Super VLAN 中的所有 Sub VLAN 发送。Super VLAN 接口如果同时启用 VRRP 和 VRRP PLUS 功能，则不支持 VRRP 协议报文只往 Super VLAN 接口第一个 UP 的 Sub VLAN 口发送。
- 本命令在 VLAN 接口上配置，只对 Super VLAN 口生效。

设置 IPv4 VRRP 组和 BFD 联动

- 【命令格式】 **vrrp group bfd ip-address**
- 【参数说明】 *group*：VRRP 组号。
ip-address：指定的邻居 IP。
- 【命令模式】 接口模式
- 【使用指导】 对于备份路由器，执行这条命令，使该 IPv4 VRRP 组和 BFD 联动，并不关心配置的 IP 地址。而对于主路由器，由于不知道备份路由器接口的主 IP 地址，所以只能由管理员指定备份路由器的 IP 地址。
- 如果配置了全局 IPv4 VRRP BFD，不能配置 IPv4 VRRP 组和 BFD 联动。
- 配置时首先确保配置的接口配置了 IP 和 BFD 会话参数。

设置全局 IPv4 VRRP BFD

- 【命令格式】 **vrrp bfd interface-type interface-number ip-address**
- 【参数说明】 *interface-type interface-number*：配置接口类型和接口编号。
ip-address：指定的邻居 IP。
- 【命令模式】 全局配置模式
- 【使用指导】 如果配置了全局 IPv4 VRRP BFD，会删除所有配置的 IPv4 VRRP 组和 BFD 联动。
- 配置时首先确保配置的接口配置了 IP 和 BFD 会话参数。
- 全局 IPv4 VRRP BFD 会话只适用于两台设备组成的 IPv4 VRRP 虚拟路由器。

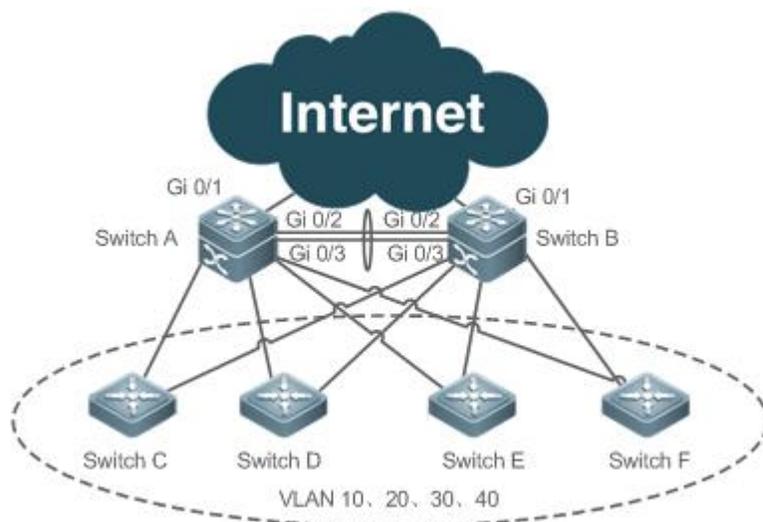
配置举例

 以下配置举例，仅介绍与 VRRP 相关的配置。

配置 VRRP+MSTP

【网络环境】

图 3-6



【配置方法】

- 在设备上（本例为 Switch A、B、C、D、E、F）使能 MSTP 功能，配置 VLAN-Instance 之间的实例映射（本例将 VLAN 10、20 对应 Instance 1，VLAN 30、40 对应 Instance 2，其余 VLAN 对应 Instance 0），并设置网关设备（本例为 Switch A 和 Switch B）为对应实例的根桥。
- 将各 VLAN 的 SVI 加入相应的 VRRP 备份组，并设置网关设备为对应备份组的 Master 路由设备和 Backup 路由设备。本例具体配置如下表所示：

网关设备	VLAN ID	SVI	备份组	虚拟 IP 地址	状态
Switch A	10	192.168.10.2	VRRP 10	192.168.10.1	Master
Switch B		192.168.10.3			Backup
Switch A	20	192.168.20.2	VRRP 20	192.168.20.1	Master
Switch B		192.168.20.3			Backup
Switch A	30	192.168.30.2	VRRP 30	192.168.30.1	Backup
Switch B		192.168.30.3			Master
Switch A	40	192.168.40.2	VRRP 40	192.168.40.1	Backup
Switch B		192.168.40.3			Master

- 将对应备份组的 Master 路由设备的上链口（本例为 Switch A 和 Switch B 的端口 Gi 0/1）设置为 Master 路由设备的监视接口。
- 第一步，在设备上创建 VLAN，在 Switch A、B 上创建 VLAN 10、20、30、40
- 第二步，配置 MST 域，在 Switch A、B 上配置 VLAN 10、20 对应 Instance 1，VLAN 20、30 对应 Instance 2，其余 VLAN 对应 Instance 0。
- 第三步，配置 Switch A 为 MST 0 和 MST 1 的根桥，Switch B 为 MST 2 的根桥。
- 第四步，使能 MSTP。
- 第五步，配置各 VLAN 的 SVI，并加入对应的备份组，同时设置备份组的虚拟 IP 地址。对应配置参见上表。

- 第六步，配置各备份组对应的 Master 路由设备和 Backup 路由设备。
- 第七步，将备份组的 Master 路由设备的上链口配置为 VRRP 组的监视接口。注意：监视接口必须为三层接口。

第八步，配置双核心设备的互联端口为 AP 口。

SwitchA

```
//在 Switch A 上创建 VLAN 10、20、30、40
SwitchA#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
SwitchA(config)#vlan range 10,20,30,40
SwitchA(config-vlan-range)#exit
//配置 VLAN 10、20 对应 Instance 1，VLAN 30、40 对应 Instance 2，其余 VLAN 对应 Instance 0。
SwitchA(config)#spanning-tree mst configuration
SwitchA(config-mst)#instance 1 vlan 10,20
%Warning:you must create vlans before configuring instance-vlan relationship
SwitchA(config-mst)#instance 2 vlan 30,40
%Warning:you must create vlans before configuring instance-vlan relationship
SwitchA(config-mst)#exit
//在 Switch A 上设置 MST 0 和 MST 1 的优先级为 4096，MST 2 的优先级为 8192。
SwitchA(config)#spanning-tree mst 0 priority 4096
SwitchA(config)#spanning-tree mst 1 priority 4096
SwitchA(config)#spanning-tree mst 2 priority 8192
//开启 MSTP
SwitchA(config)#spanning-tree
Enable spanning-tree.
//配置各 VLAN 的 SVI，并加入对应的备份组，同时设置备份组的虚拟 IP 地址。
SwitchA(config)#interface vlan 10
SwitchA(config-if-VLAN 10)#ip address 192.168.10.2 255.255.255.0
SwitchA(config-if-VLAN 10)#vrrp 10 ip 192.168.10.1
SwitchA(config-if-VLAN 10)#exit
SwitchA(config)#interface vlan 20
SwitchA(config-if-VLAN 20)#ip address 192.168.20.2 255.255.255.0
SwitchA(config-if-VLAN 20)#vrrp 20 ip 192.168.20.1
SwitchA(config-if-VLAN 20)#exit
SwitchA(config)#interface vlan 30
SwitchA(config-if-VLAN 30)#ip address 192.168.30.2 255.255.255.0
SwitchA(config-if-VLAN 30)#vrrp 30 ip 192.168.30.1
SwitchA(config-if-VLAN 30)#exit
SwitchA(config)#interface vlan 40
SwitchA(config-if-VLAN 40)#ip address 192.168.40.2 255.255.255.0
SwitchA(config-if-VLAN 40)#vrrp 40 ip 192.168.40.1
SwitchA(config-if-VLAN 40)#exit
```

SwitchB

```
//将 Switch A 的备份组 10、20 的优先级调高为 120
SwitchA(config)#interface vlan 10
SwitchA(config-if-VLAN 10)#vrrp 10 priority 120
SwitchA(config-if-VLAN 10)#exit
SwitchA(config)#interface vlan 20
SwitchA(config-if-VLAN 20)#vrrp 20 priority 120
SwitchA(config-if-VLAN 20)#exit
//配置 Switch A 的端口 Gi 0/1 为 Route Port，并设置 IP 地址为 10.10.1.1/24
SwitchA(config)#interface gigabitEthernet 0/1
SwitchA(config-if-GigabitEthernet 0/1)#no switchport
SwitchA(config-if-GigabitEthernet 0/1)#ip address 10.10.1.1 255.255.255.0
SwitchA(config-if-GigabitEthernet 0/1)#exit
//将 Switch A 的端口 Gi 0/1 配置为备份组 10、20 的监视接口，并设置 Priority decrement 为 30。
SwitchA(config)#interface vlan 10
SwitchA(config-if-VLAN 10)#vrrp 10 track gigabitEthernet 0/1 30
SwitchA(config-if-VLAN 10)#exit
SwitchA(config)#interface vlan 20
SwitchA(config-if-VLAN 20)#vrrp 20 track gigabitEthernet 0/1 30
SwitchA(config-if-VLAN 20)#exit
//配置端口 Gi 0/2 和 Gi 0/3 属于 AP 口，并设置 AP 口为 trunk 口。
SwitchA#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
SwitchA(config)#interface range gigabitEthernet 0/2-3
SwitchA(config-if-range)#port-group 1
SwitchA(config)#interface aggregateport 1
SwitchA(config-if-AggregatePort 1)#switchport mode trunk
SwitchB
//在 Switch B 上创建 VLAN 10、20、30、40
SwitchB#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
SwitchB(config)#vlan range 10,20,30,40
SwitchB(config-vlan-range)#exit
//配置 VLAN 10、20 对应 Instance 1，VLAN 30、40 对应 Instance 2，其余 VLAN 对应 Instance 0。
SwitchB(config)#spanning-tree mst configuration
SwitchB(config-mst)#instance 1 vlan 10,20
%Warning:you must create vlans before configuring instance-vlan relationship
SwitchB(config-mst)#instance 2 vlan 30,40
%Warning:you must create vlans before configuring instance-vlan relationship
SwitchB(config-mst)#exit
//在 Switch B 上设置 MST 2 的优先级为 4096，MST 0 和 MST 1 的优先级为 8192
SwitchB(config)#spanning-tree mst 2 priority 4096
SwitchB(config)#spanning-tree mst 0 priority 8192
```

```
SwitchB(config)#spanning-tree mst 1 priority 8192
//开启 MSTP
SwitchB(config)#spanning-tree
Enable spanning-tree.
//配置各 VLAN 的 SVI，并加入对应的备份组，同时设置备份组的虚拟 IP 地址。
SwitchB(config)#interface vlan 10
SwitchB(config-if-VLAN 10)#ip address 192.168.10.3 255.255.255.0
SwitchB(config-if-VLAN 10)#vrrp 10 ip 192.168.10.1
SwitchB(config-if-VLAN 10)#exit
SwitchB(config)#interface vlan 20
SwitchB(config-if-VLAN 20)#ip address 192.168.20.3 255.255.255.0
SwitchB(config-if-VLAN 20)#vrrp 20 ip 192.168.20.1
SwitchB(config-if-VLAN 20)#exit
SwitchB(config)#interface vlan 30
SwitchB(config-if-VLAN 30)#ip address 192.168.30.3 255.255.255.0
SwitchB(config-if-VLAN 30)#vrrp 30 ip 192.168.30.1
SwitchB(config-if-VLAN 30)#exit
SwitchB(config)#interface vlan 40
SwitchB(config-if-VLAN 40)#ip address 192.168.40.3 255.255.255.0
SwitchB(config-if-VLAN 40)#vrrp 40 ip 192.168.40.1
SwitchB(config-if-VLAN 40)#exit
//将 Switch B 的 VRRP 30、40 的优先级调高为 120。
SwitchB(config)#interface vlan 30
SwitchB(config-if-VLAN 30)#vrrp 30 priority 120
SwitchB(config-if-VLAN 30)#exit
SwitchB(config)#interface vlan 40
SwitchB(config-if-VLAN 40)#vrrp 40 priority 120
SwitchB(config-if-VLAN 40)#exit
//配置 Switch B 的端口 Gi 0/1 为 Route Port，并设置 IP 地址为 10.10.2.1/24
SwitchB(config)#interface gigabitEthernet 0/1
SwitchB(config-if-GigabitEthernet 0/1)#no switchport
SwitchB(config-if-GigabitEthernet 0/1)#ip address 10.10.2.1 255.255.255.0
SwitchB(config-if-GigabitEthernet 0/1)#exit
//将 Switch B 的端口 Gi 0/1 配置为备份组 30、40 的监视接口，并设置 Interface -Priority 为 30。
SwitchB(config)#interface vlan 30
SwitchB(config-if-VLAN 30)#vrrp 30 track gigabitEthernet 0/1 30
SwitchB(config-if-VLAN 30)#exit
SwitchB(config)#interface vlan 40
SwitchB(config-if-VLAN 40)#vrrp 40 track gigabitEthernet 0/1 30
SwitchB(config-if-VLAN 40)#exit
//配置端口 Gi 0/2 和 Gi 0/3 属于 AP 口，并设置 AP 口为 trunk 口。
```

```
SwitchB #configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
SwitchB (config)#interface range gigabitEthernet 0/2-3
SwitchB (config-if-range)#port-group 1
SwitchB (config)#interface aggregateport 1
SwitchB (config-if-AggregatePort 1)#switchport mode trunk
```

【检验方法】

SwitchA

```
//查看设备的配置信息。
SwitchA#show running-config
!
vlan 10
!
vlan 20
!
vlan 30
!
vlan 40
!
spanning-tree
spanning-tree mst configuration
 instance 0 vlan 1-9, 11-19, 21-29, 31-39, 41-4094
 instance 1 vlan 10, 20
 instance 2 vlan 30, 40
spanning-tree mst 0 priority 4096
spanning-tree mst 1 priority 4096
spanning-tree mst 2 priority 8192
interface GigabitEthernet 0/1
 no switchport
 no ip proxy-arp
 ip address 10.10.1.1 255.255.255.0
!
interface GigabitEthernet 0/2
 port-group 1
!
interface GigabitEthernet 0/3
 port-group 1
!
interface AggregatePort 1
 switchport mode trunk
!
interface VLAN 10
```

```
no ip proxy-arp
ip address 192.168.10.2 255.255.255.0
vrrp 10 priority 120
vrrp 10 ip 192.168.10.1
vrrp 10 track GigabitEthernet 0/1 30
!
interface VLAN 20
no ip proxy-arp
ip address 192.168.20.2 255.255.255.0
vrrp 20 priority 120
vrrp 20 ip 192.168.20.1
vrrp 20 track GigabitEthernet 0/1 30
!
interface VLAN 30
no ip proxy-arp
ip address 192.168.30.2 255.255.255.0
vrrp 30 ip 192.168.30.1
!
interface VLAN 40
no ip proxy-arp
ip address 192.168.40.2 255.255.255.0
vrrp 40 ip 192.168.40.1
//查看设备的 VRRP 状态
SwitchA#show vrrp brief
Interface Grp Pri timer Own Pre State Master addr Group addr
VLAN 10 10 120 3 - P Master 192.168.10.2 192.168.10.1
VLAN 20 20 120 3 - P Master 192.168.20.2 192.168.20.1
VLAN 30 30 100 3 - P Backup 192.168.30.3 192.168.30.1
VLAN 40 40 100 3 - P Backup 192.168.40.3 192.168.40.1
//断开 Switch A 的上行链路，查看设备的 VRRP 状态
SwitchA#show vrrp brief
Interface Grp Pri timer Own Pre State Master addr Group addr
VLAN 10 10 90 3 - P Backup 192.168.10.3 192.168.10.1
VLAN 20 20 90 3 - P Backup 192.168.20.3 192.168.20.1
VLAN 30 30 100 3 - P Backup 192.168.30.3 192.168.30.1
VLAN 40 40 100 3 - P Backup 192.168.40.3 192.168.40.1
```

SwitchB

```
//查看设备的配置信息。
SwitchB#show running-config
!
```

```
vlan 10
!
vlan 20
!
vlan 30
!
vlan 40
!
spanning-tree
spanning-tree mst configuration
  instance 0 vlan 1-9, 11-19, 21-29, 31-39, 41-4094
  instance 1 vlan 10, 20
  instance 2 vlan 30, 40
spanning-tree mst 0 priority 8192
spanning-tree mst 1 priority 8192
spanning-tree mst 2 priority 4096
interface GigabitEthernet 0/1
  no switchport
  no ip proxy-arp
  ip address 10.10.2.1 255.255.255.0
!
interface GigabitEthernet 0/2
  port-group 1!
interface GigabitEthernet 0/3
  port-group 1
!
interface AggregatePort 1
  switchport mode trunk
!
interface VLAN 10
  no ip proxy-arp
  ip address 192.168.10.3 255.255.255.0
  vrrp 10 ip 192.168.10.1
!
interface VLAN 20
  no ip proxy-arp
  ip address 192.168.20.3 255.255.255.0
  vrrp 20 ip 192.168.20.1
!
interface VLAN 30
  no ip proxy-arp
```

```
ip address 192.168.30.3 255.255.255.0
vrrp 30 priority 120
vrrp 30 ip 192.168.30.1
vrrp 30 track GigabitEthernet 0/1 30
!
interface VLAN 40
no ip proxy-arp
ip address 192.168.40.3 255.255.255.0
vrrp 40 priority 120
vrrp 40 ip 192.168.40.1
vrrp 40 track GigabitEthernet 0/1 30
//查看设备的 VRRP 状态
SwitchB#show vrrp brief
Interface Grp Pri timer Own Pre State Master addr Group addr
VLAN 10 10 100 3 - P Backup 192.168.10.2 192.168.10.1
VLAN 20 20 100 3 - P Backup 192.168.20.2 192.168.20.1
VLAN 30 30 120 3 - P Master 192.168.30.3 192.168.30.1
VLAN 40 40 120 3 - P Master 192.168.40.3 192.168.40.1
//断开 Switch B 的上行链路，查看设备的 VRRP 状态如下
SwitchB#show vrrp brief
Interface Grp Pri timer Own Pre State Master addr Group addr
VLAN 10 10 100 3 - P Master 192.168.10.3 192.168.10.1
VLAN 20 20 100 3 - P Master 192.168.20.3 192.168.20.1
VLAN 30 30 120 3 - P Master 192.168.30.3 192.168.30.1
VLAN 40 40 120 3 - P Master 192.168.40.3 192.168.40.1
```

常见错误

- 同一 VRRP 备份组内各路由设备以太网接口上 VRRP 组验证模式不同导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 对于 VRRPv2，同一 VRRP 备份组内各路由设备以太网接口上 VRRP 组验证模式相同均为明文密码模式，但是验证字符串不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 同一个 VRRP 备份组内路由设备上 VRRP 的通告发送间隔不一致，并且未设置定时设备学习功能导致同一个 VRRP 备份组内出现多个 Master 路由设备。
- 同一个 VRRP 备份组内路由设备上 VRRP 的虚拟 IP 地址不一致导致同一个 VRRP 备份组内出现多个 Master 路由设备。

3.5 监视与维护

清楚各类信息

无

查看运行情况

作用	命令
查看 IPv4 VRRP 的概况或者细节。	show vrrp [brief group]
查看指定接口上的 IPv4 VRRP 组的情况。	show vrrp interface type number [brief]
查看 VRRP 报文收发的统计信息。	show vrrp packet statistics [interface-type interface-number]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 VRRP 出错提示、VRRP 事件、VRRP 报文以及状态调试开关。	debug vrrp
打开 VRRP 出错提示调试开关。	debug vrrp errors
打开 VRRP 事件提示调试开关。	debug vrrp events
打开 VRRP 数据包调试开关。	debug vrrp packets [acl acl-id [icmp protocol] interface type number [group]]
打开 VRRP 状态调试开关。	debug vrrp state

4 VRRP Plus

4.1 概述

VRRP Plus (Virtual Router Redundancy Protocol Plus, 扩展虚拟路由冗余协议) 功能是对 VRRP 协议的扩展, 利用 VRRP 协议进行 IEEE 802.3 局域网内的网关备份和负载均衡。

VRRP 协议有个不足之处是: 处于备份状态的路由设备, 没有承担报文转发的任务。如果要利用 VRRP 实现负载均衡, 则需要手工配置多个 VRRP 组, 并将局域网内主机的网关指向不同 VRRP 组的虚拟 IP 地址。这将增大网络管理员的工作量, VRRP Plus 就是为了解决上述不足而设计的。

VRRP Plus 的好处是: 自动实现负载均衡, 即自动将不同主机的流量分配到 VRRP Plus 组成员中, 无需配置多个 VRRP 组以及设置局域网内主机网关指向不同的 VRRP 组的虚拟 IP 地址。从而大大减轻网络管理员的负担。

i 下文仅介绍 VRRP Plus 的相关内容。

协议规范

4.2 典型应用

典型应用	场景描述
在一个VRRP组内启动负载均衡转发	实现 VRRP 组的负载均衡转发, 无需配置多个组及为每台主机配置不同默认网关。

4.2.1 在一个VRRP组内启动负载均衡转发策略

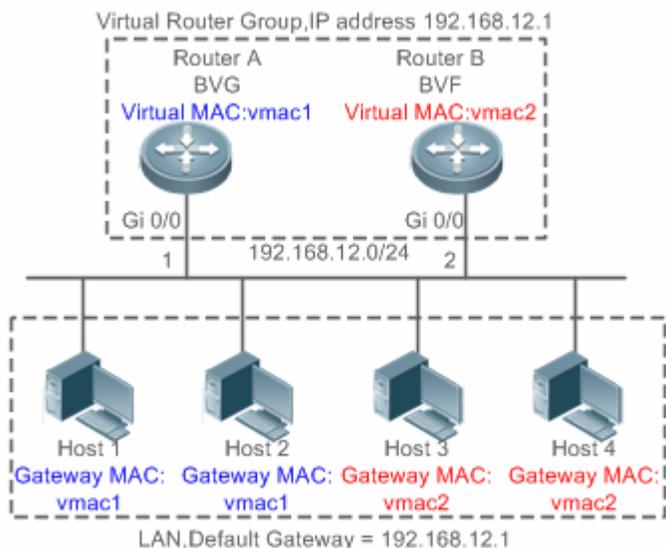
应用场景

在一个 VRRP 组内启动负载均衡转发策略, 无需配置多个 VRRP 组, 也不用为每台主机配置不同的默认网关。

以下图 4-1 为例:

- 配置由 Router A 和 Router B 组成的 VRRP 组, 并启动 VRRP Plus 功能。
- 每台主机将自己的默认网关设置为 VRRP 组的主虚 IP 地址。

图 4-1 IPv4 VRRP Plus 应用拓扑示意图



- 【注释】
1. 两台三层设备 Router A 和 Router B 组成一个 VRRP Plus 组，虚拟 IP 地址为 192.168.12.1。Router A 是 VRRP 的 Master 设备，承担 BVG 职责。Router B 是 VRRP 的 Backup 设备，承担 BVF 职责。
 2. Host1~Host4 是网段为 192.168.12.0/24 的局域网中的主机，默认网关都指向 VRRP Plus 的虚 IP：192.168.12.1。
 3. 对于不同的主机的 ARP 请求，根据在设备上配置的均衡策略进行对应的 ARP 应答。比如 Host1 和 Host2 请求网关 ARP 后，响应的 MAC 为 0000.5e00.0101，Host3 和 Host4 请求网关 ARP 后，响应的 MAC 为 001A.A916.0201。这样 Host1 和 Host2 与外网通讯的报文就发送到 Router A，Host3 和 Host4 与外网通讯的报文就发送到 Router B，从而达到了负载均衡的目的。

功能部属

- 在 Router A 和 Router B 部署 VRRP Plus 策略，实现对本地主机的负载均衡转发。

4.3 功能详解

基本概念

↘ BVG (Balancing Virtual Gateway, 均衡的虚拟网关)

负责 VRRP Plus 组成员虚拟 MAC 的分配，局域网中针对网关 ARP 的应答，同时也承担局域网中主机报文的转发。

↘ BVF (Balancing Virtual Forwarder, 均衡的虚拟转发者)

负责局域网中主机报文的转发，若分配到虚 MAC，则会参与转发，否则不会参与转发。

功能特性

功能特性	作用
VRRP Plus功能	对 VRRP 协议进行扩展，利用 VRRP 协议进行 IEEE 802.3 局域网内的网关备份和负载均衡。

4.3.1 VRRP Plus功能

自动实现负载均衡，自动将不同主机的流量分配到 VRRP Plus 组成员中，无需配置多个 VRRP 组以及设置局域内主机网关指向不同的 VRRP 组的虚拟 IP 地址。

基本原理

VRRP Plus 的基本原理是：局域网中的主机使用统一的网关 IP（即 VRRP 组的虚拟 IP），但不同主机在请求网关 ARP 时，由 BVG 应答不同的虚拟 MAC，从而将不同主机的流量分配到 VRRP Plus 的不同成员上，实现负载均衡。

▾ VRRP Plus 与 VRRP 关系

VRRP Plus 是依赖于 VRRP 协议运行的，其运行规则如下：

VRRP 中 Master 的角色，对应 VRRP Plus 中 BVG 角色，VRRP 中 Backup 的角色，对应 VRRP Plus 中 BVF 角色。局域网内主机的网关指向 VRRP 的虚拟 IP 地址。

▾ BVG 和 BVF 的 MAC 地址分配规则

BVG 负责给 BVF 分配虚拟 MAC。对于 IPv4 VRRP Plus，为和 VRRP 兼容，BVG 直接使用 VRRP 的虚拟 MAC，即 00-00-5E-00-01-VRID（其中 VRID 为 VRRP 组号）；BVF 使用的虚拟 MAC 为 00-1A-A9-16-MemberID-VRID（其中 MemberID 为 VRRP Plus 组成员的编号）。目前 VRRP Plus 支持 4 个成员，BVG 使用成员编号 01，其它三个 BVF 使用成员编号 02~04。

▾ VRRP Plus 负载均衡策略

BVG 负责响应局域网内主机的网关 ARP 请求，根据均衡策略的不同，BVG 使用不同的虚拟 MAC 进行应答。均衡策略有三种：基于主机响应，轮询响应和基于权重响应。

- 1) 基于主机响应是指对于特定的主机，使用特定的虚拟 MAC 地址进行应答；
- 2) 轮流响应是指对于收到的主机的网关地址 ARP 请求，轮流使用备份组的各个虚拟 MAC 地址进行应答；
- 3) 基于权重是指根据设备的转发能力进行 ARP 报文的应答。

如果在不同的负载均衡模式下切换，则最终都是新的模式下实现负载均衡，如之前使用的是基于轮询的负载均衡模式，之后使用基于权重的负载均衡模式，则是在新的模式下实现负载均衡，即不考虑设备之前的应答情况。如果基于权重策略，且 VRRP Plus 组内虚拟路由器总的权重值为 0，则不会应答 ARP 请求。

▾ 虚拟 MAC 地址的代理

当备份组中某台已经分配了虚拟 MAC 的设备出现故障时，此时使用该虚拟 MAC 作为网关 MAC 地址的用户的流量便会中断。

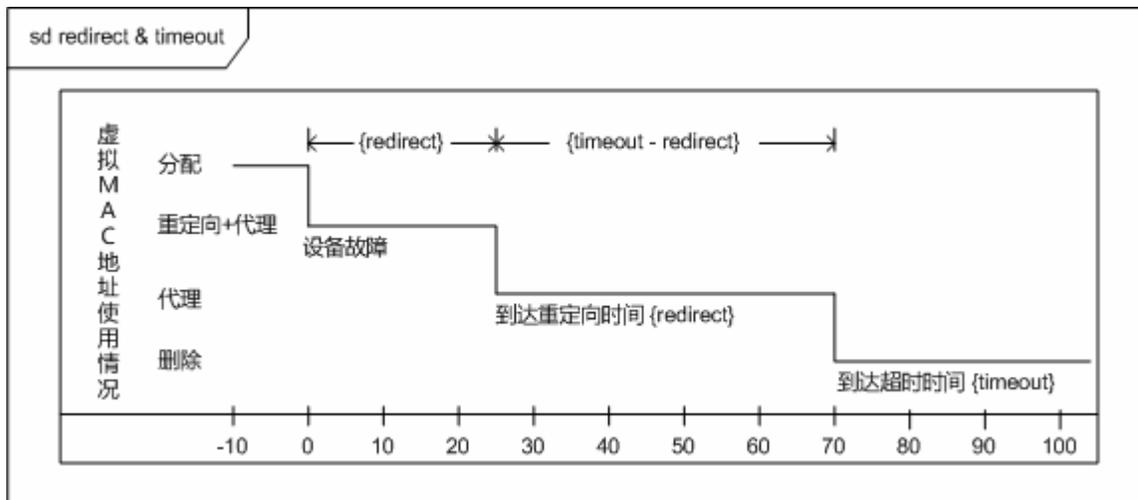
VRRP Plus 备份组中负责管理虚 MAC 地址的 BVG 设备应能够快速检测到该故障，并将该故障的 BVF 设备的虚拟 MAC 地址自动分配给备份组中的其他设备，代替故障设备对该虚 MAC 地址的报文进行转发，由该代理设备承担原用户的流量，避免用户流量中断。对于备份组中某台设备分配得到的虚 MAC 地址，可称为主虚 MAC 地址；对于由它代理的虚 MAC 地址，称为代理虚 MAC 地址。

▾ 代理虚 MAC 的重定向时间和超时时间

VRRP Plus 提供虚拟 MAC 地址的代理功能，对已经分配了虚拟 MAC 地址且出现了故障的设备进行代理转发。如果 BVF 从故障中恢复，则应该恢复其转发角色，继续承担该虚拟 MAC 地址的报文转发任务。但如果原设备一直未能恢复，则备份组应该停止重定向流量到该虚拟 MAC 地址，即再收到对网关地址的 ARP 请求时，不再使用该虚拟 MAC 地址进行应答；而且，经过足够长的时间后，我们可以认为原来使用该虚拟 MAC 地址作为网关 MAC 的用户已经更新了网关地址的 ARP 表项、流量已经由其他设备承担，此时便可以删除该虚拟 MAC 地址，发往该虚拟 MAC 地址的报文应该被丢弃。

为此，VRRP Plus 支持设置备份组的重定向时间和超时时间。当设备出现故障后，备份组将其虚拟 MAC 地址分配给其他设备进行代理。在重定向时间内，备份组继续使用该虚拟 MAC 地址应答 ARP 请求；超过重定向时间后，便不再使用该虚拟 MAC 地址进行应答。经过超时时间后，备份组应删除该虚拟 MAC 地址，停止对该虚 MAC 进行代理转发。虚拟 MAC 地址在重定向时间和超时时间内的角色变化如图 4-2 所示：

图 4-2



▾ 基于权重的转发

VRRP Plus 支持备份组的权重配置。为不同的设备设置不同的权重值，以使权重值大者分担较多流量，权重值小者分担较少流量，使不同设备的转发性能得到充分利用。当备份组中 BVF 设备的权重值低于下阈值时，将自动退出转发角色，当权重值恢复、高于上阈值时，自动申请为转发角色，能否恢复转发角色，取决于当前是否有剩余的虚 MAC 地址或者代理的虚 MAC 地址。

▾ VRRP Plus 与 BFD 联动

VRRP Plus 支持与链路检测协议联动、根据链路状态调整权重值。备份组中各台设备可以各自关联相应的链路状态，当该链路出现异常、不通时，该设备应自动降低其权重值；如果权重过低、则应该可以退出转发角色。若备份组正在使用基于权重的负

4.4.1 配置VRRP Plus

配置效果

- 启动 VRRP Plus 功能（缺省未开启接口 VRRP Plus 功能）。

注意事项

- 启动 VRRP Plus 功能，必须在相应的备份组上配置 VRRP 虚地址。

配置方法

▾ 启动 VRRP Plus 功能

- VRRP Plus 默认关闭，如果要具备 VRRP Plus 功能则必须配置。

▾ 设置 VRRP Plus 负载均衡策略

- 启动了 VRRP Plus 后，默认使用基于主机均衡的转发策略，即 host-dependent。

▾ 设置 VRRP Plus 备份组的代理虚 MAC 地址的重定向时间和超时时间

- 启动了 VRRP Plus 后，默认配置了该功能，redirect 缺省值为 300 秒，timeout 的缺省值为 14400 秒。

▾ 设置 VRRP Plus 备份组的权重和上下限值。

- 启动了 VRRP Plus 后，默认配置了该功能。备份组的权重值缺省值为 100，权重值下限缺省值为 1，上限缺省值为 100。

▾ 设置 VRRP Plus 备份组的转发抢占功能

- 启动了 VRRP Plus 后，默认打开了允许转发抢占功能。

检验方法

- 通过 **show group vrrp balance** 命令显示验证，如果该备份组有承担负载转发的任务，则显示的转发者 forwarders 中会显示“local”，且显示该备份组分配到的虚拟 MAC 地址。

相关命令

▾ 启动 VRRP Plus 功能

- 【命令格式】 **vrrp group balance**
- 【参数说明】 *group*：VRRP 组号，不同产品型号取值范围不同。
- 【命令模式】 接口配置模式
- 【使用指导】 需要配置 VRRP 组，VRRP PLUS 功能才能启动。

配置 VRRP Plus 备份组的负载均衡策略

【命令格式】 `vrrp group load-balancing { host-dependent | round-robin | weighted }`

【参数说明】 `group` : VRRP 组号。

host-dependent : 表示基于主机的均衡转发策略。

round-robin : 表示基于轮询的均衡策略。

weighted : 表示基于备份组权重的均衡。

【命令模式】 接口配置模式

【使用指导】 启动了 VRRP Plus 后，默认使用基于主机均衡的转发策略，即 **host-dependent**。整个备份组的负载均衡策略由配置在 BVG 上的均衡策略决定，若用户希望在 BVG 设备角色更换后还使用同样的均衡策略，则需要备份组内的所有设备配置相同的均衡策略。

配置 VRRP Plus 备份组的代理虚 MAC 地址的重定向时间和超时时间

【命令格式】 `vrrp group timers redirect redirect timeout`

【参数说明】 `group` : VRRP 组号。

`redirect` : 重定向时间，取值范围为 0 ~ 3600 秒，缺省值为 300 秒（即 5 分钟）。

`timeout` : 超时时间，取值范围为 (`redirect` + 600) - 64800 秒，缺省值为 14400 秒（即 4 小时）。

【命令模式】 接口配置模式

【使用指导】 启动了 VRRP Plus 后，默认配置了该功能，`redirect` 缺省值为 300 秒，`timeout` 的缺省值为 14400 秒。当设备出现故障后，备份组将其虚拟 MAC 地址分配给其他设备进行代理。在重定向时间内，备份组继续使用该虚拟 MAC 地址应答 ARP 请求；超过重定向时间后，便不再使用该虚拟 MAC 地址进行应答。经过超时时间后，备份组应删除该虚拟 MAC 地址。

配置 VRRP Plus 备份组的权重和上下限值。

【命令格式】 `vrrp group weighting maximum [lower lower] [upper upper]`

【参数说明】 `maximum` : 备份组的权重值，取值范围 2-254，缺省值为 100。

`lower lower` : 备份组的权重下限值，取值范围 1 ~ (`maximum` - 1)，缺省值为 1。

`upper upper` : 备份组的权重上限值，取值范围 `lower` ~ `maximum`，缺省值为 100。

【命令模式】 接口配置模式

【使用指导】 启动了 VRRP Plus 后，默认配置了该功能。用户通过此命令不同的设备设置不同的权重值，以使权重值大者分担较多流量，权重值小者分担较少流量。当备份组中 BVF 设备的权重值低于下限值时，将自动退出转发角色，当权重值恢复、高于上限值时，应该自动恢复转发角色。

配置 VRRP Plus 备份组的转发抢占功能

【命令格式】 `vrrp group forwarder preempt`

【参数说明】 `group` : VRRP 组号。

【命令模式】 接口模式

【使用指导】 启动了 VRRP Plus 后，默认打开了转发抢占功能。VRRP Plus 允许设置备份组的转发抢占功能，当发现处于监听状态的设备的权重值比正在负责转发的设备的权重值更高时，使监听设备自动抢占转发设备的角色，由转发能力更强的设备负责报文转发，较低的设备进行监听。

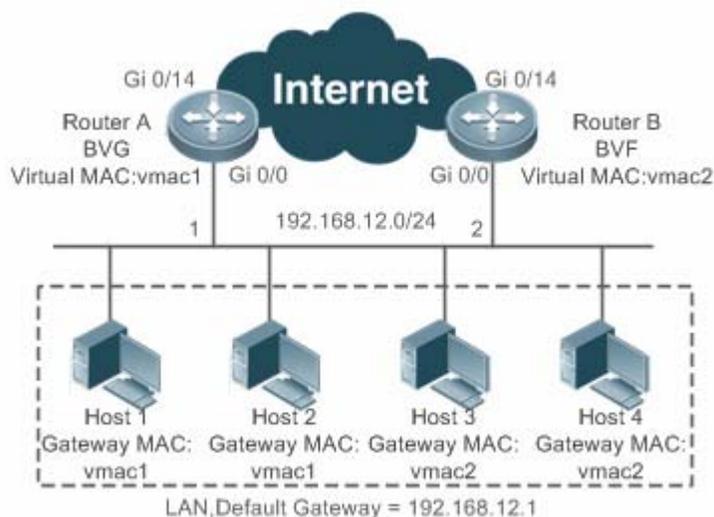
配置举例

i 以下配置举例，仅介绍与 VRRP Plus 相关的配置。

在一个 IPv4 VRRP 组内启动负载均衡转发策略

【网络环境】

图 4-3



【配置方法】

- 在 Router A 和 Router B 配置 VRRP 组，并开启 VRRP Plus 功能；通过配置本地 IP 地址使得 Router A 为 BVG (Master) 设备，Router B 为 BVF (Backup) 设备。
- 备份组权重和阈值，重定向时间和超时时间以及是否允许转发抢占功能都采用默认配置。
- 本地局域网的主机 Host1 ~ Host4 默认网关指向 VRRP 的虚 IP 地址即 192.168.12.1。

Router A

```
RuijieA#config
RuijieA(config)#interface GigabitEthernet 0/0
// 'no switchport'在交换机上使用
RuijieA(config-if-GigabitEthernet 0/0)#no switchport
RuijieA(config-if-GigabitEthernet 0/0)#ip address 192.168.12.3 255.255.255.0
RuijieA(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.12.1
RuijieA(config-if-GigabitEthernet 0/0)#vrrp 1 balance
RuijieA(config-if-GigabitEthernet 0/0)#vrrp 1 load-balancing weighted
```

Router B

```
RuijieB#config
RuijieB(config)#interface GigabitEthernet 0/0
RuijieB(config-if-GigabitEthernet 0/0)#no switchport
RuijieB(config-if-GigabitEthernet 0/0)#ip address 192.168.12.2 255.255.255.0
RuijieB(config-if-GigabitEthernet 0/0)#vrrp 1 ip 192.168.12.1
RuijieB(config-if-GigabitEthernet 0/0)#vrrp 1 balance
RuijieB(config-if-GigabitEthernet 0/0)#vrrp 1 load-balancing weighted
```

【检验方法】

通过 **show vrrp balance** 显示 VRRP PLUS 备份组的配置情况，如果该备份组有承担负载转发的任务，则显示的转发者 Forwarder 中会显示“local”，且显示该备份组分配到的虚拟 MAC 地址。

Router A

```
RuijieA# show vrrp balance interface GigabitEthernet 0/0
State is BVG
Virtual IP address is 192.168.12.1
Hello time 1 sec, hold time 3 sec
Load balancing: weighted
Redirect time 300 sec, forwarder time-out 14400 sec
Weighting 100 (configured 100), thresholds: lower 1, upper 100
  There are 2 forwarders
Forwarder 1 (local)
  MAC address:
    0000.5e00.0101
  Owner ID is 0000.0001.0006
  Preemption disabled (BVG cannot be preempted)
Forwarder 2
  MAC address:
    001a.a916.0201
  Owner ID is 00d0.f822.33a3
  Preemption enabled
```

Router B

```
RuijieB# show vrrp balance interface GigabitEthernet 0/0
State is BVF
Virtual IP address is 192.168.12.1
Hello time 1 sec, hold time 3 sec
Load balancing: weighted
Redirect time 300 sec, forwarder time-out 14400 sec
Weighting 100 (configured 100), thresholds: lower 1, upper 100
  There are 2 forwarders
Forwarder 1
  MAC address:
    0000.5e00.0101
  Owner ID is 0000.0001.0006
  Preemption disabled (BVG cannot be preempted)
Forwarder 2 (local)
  MAC address:
    001a.a916.0201
  Owner ID is 00d0.f822.33a3
  Preemption enabled
```

常见错误

- 没有在相应的组上配置 VRRP 虚地址，则配置 VRRP PLUS 功能不生效。

4.5 监视与维护

清除各类信息

无

查看运行情况

作用	命令
显示 VRRP Plus 的概况或者细节	show vrrp balance
显示指定接口上的 VRRP Plus 组的操作情况	show vrrp balance interface

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 VRRP Plus 功能所有调试开关	debug vrrp balance
打开错误信息调试开关	debug vrrp balance error
打开 VRRP Plus 组事件的调试开关	debug vrrp balance event
打开与 VRRP 模块间的消息调试开关	debug vrrp balance messages
打开 VRRP Plus 协议报文调试开关	debug vrrp balance packets
打开 VRRP Plus 组状态调试开关	debug vrrp balance state
打开 VRRP Plus 组的定时器消息调试开关	debug vrrp balance timer

5 BFD

5.1 概述

为了减小故障对业务的影响，提高网络的可用性，设备需要能够尽快检测到与相邻设备间的通信故障，以便能够及时采取措施，从而保证业务继续进行。BFD (Bidirectional Forwarding Detection, 双向转发检测)，提供一种轻负载、快速检测两台邻接路由器之间转发路径连通状态的方法。可以为各上层协议如路由协议、MPLS 等统一地快速检测两台路由器间双向转发路径的故障，加快启用备份转发路径，提升现有网络性能。

 下文仅介绍 BFD 的相关内容。

协议规范

- draft-ietf-bfd-base-09 : Bidirectional Forwarding Detection
- draft-ietf-bfd-generic-05 : Generic Application of BFD
- draft-ietf-bfd-v4v6-1hop-09 : BFD for IPv4 and IPv6 (Single Hop)
- draft-ietf-bfd-mpls-07 : BFD For MPLS LSPs

5.2 典型应用

典型应用	场景描述
OSPF与BFD联动	OSPF 利用 BFD 快速检测邻居状态
静态路由与BFD联动	静态路由利用 BFD 快速检测路由下一跳的可达性

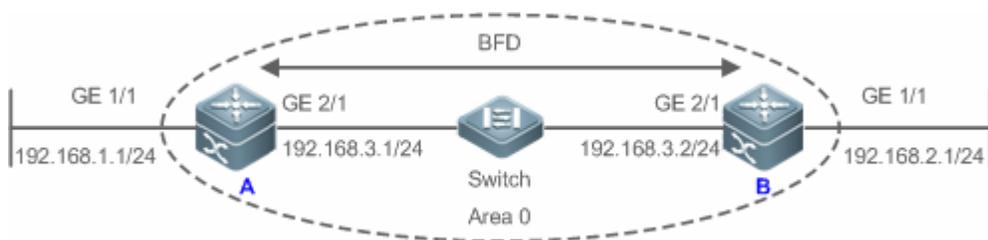
5.2.1 OSPF与BFD联动

应用场景

OSPF 协议通过 Hello 报文动态发现邻居，当 OSPF 启动 BFD 检测功能后，将会为达到 FULL 关系的邻居建立 BFD 会话，通过 BFD 机制检测邻居状态，一旦 BFD 邻居失效，OSPF 会立刻进行网络收敛。收敛的时间可以从 120 秒(缺省情况非广播型网络 OSPF hello 报文的发送间隔为 30 秒，而邻居设备失效的时间是间隔时间的 4 倍，也就是需要 120s)降到 1 秒内。

以下图为例，Router A、Router B 通过二层交换机 switch 互连，在设备上运行 OSPF 协议来建立路由，同时使能允许 OSPF 在双方接口上关联 BFD 应用。在 Router B 和二层交换机 switch 之间的链路发生故障后，BFD 能够快速检测并通告 OSPF 协议，触发协议快速收敛。

图 5-1



- 【注释】 A、B 为路由器。
 switch 为二层交换机。
 A、B 通过二层交换机 switch 互连。

功能部属

- 在 Router A 和 B 相连接口配置 IP 地址
- 在 Router A 和 B 运行 OSPF 协议
- 在 Router A 和 B 相连接口配置 BFD 参数
- 在 Router A 和 B 使能 OSPF 联动 BFD

5.2.2 静态路由与BFD联动

应用场景

静态路由与 BFD 联动，可以避免在配置的静态路由不可达的情况下，路由选路不会选择该静态路由作为转发路径。如果存在备份路由转发路径，将可以快速切换到该备份转发路径。

与动态路由协议不同，静态路由由没有发现邻居的机制，因此，当配置 BFD 与静态路由关联，静态路由的下一跳可达性将依赖于 BFD 会话状态。如果 BFD 会话检测到故障，表示静态路由的下一跳不可达，则该静态路由将不安装到 RIB 中。

以下图为例，Router A、Router B 通过二层交换机 switch 互连，在设备上配置静态路由来建立转发，同时使能允许静态路由在双方接口上关联 BFD 应用。在 Router B 和二层交换机 switch 之间的链路发生故障后，BFD 能够快速检测并通告静态路由，触发系统将该静态路由从 RIB 中删除，从而避免选路错误。

图 5-2



- 【注释】 A、B 为路由器。
 switch 为二层交换机。
 A、B 通过二层交换机 switch 互连

功能部属

- 在路由器 A 和 B 相连接口配置 IP 地址
- 在路由器 A 和 B 配置静态路由
- 在路由器 A 和 B 相连接口配置 BFD 参数
- 在路由器 A 和 B 使能静态路由联动 BFD

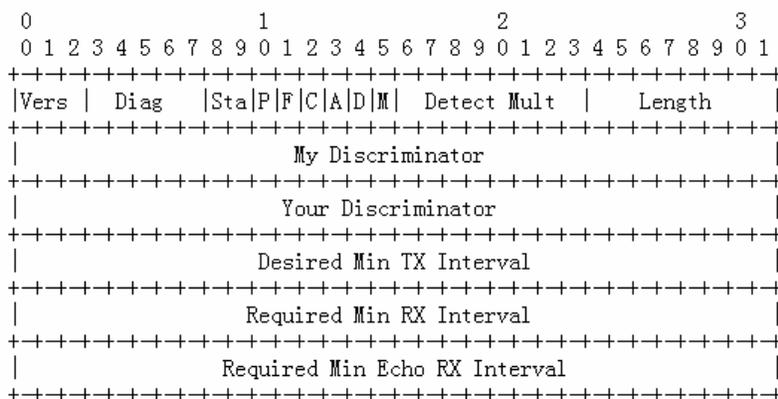
5.3 功能详解

基本概念

报文格式

BFD 发送的检测报文是 UDP 报文，有两种类型分别是控制报文和回声报文。其中回声报文只有 BFD 会话本端系统关心，远端系统不关心，因此协议没有规定其具体格式。协议只规定了控制报文的格式。目前控制报文格式有两个版本(版本 0 和版本 1)，建立 BFD 会话时缺省采用版本 1，如果收到对端系统发送的是版本 0 的报文，将自动切换到版本 0。

图 5-3



字段	说明
Vers	BFD 协议版本号，目前为 1。
Diag	给出本地最后一次从 UP 状态转到其他状态的原因，包括： 0—没有诊断信息 1—控制超时检测 2—回声功能失效 3—邻居通告会话 Down 4—转发面复位 5—通道失效

	6—连接通道失效 7—管理 Down
Sta	BFD 本地状态，包括： 0 代表 AdminDown 1 代表 Down 2 代表 Init 3 代表 Up
P	参数发生改变时，发送方在 BFD 报文中置该标志，接收方必须立即响应该报文。
F	响应 P 标志置位的回应报文中必须将 F 标志置位。
C	转发/控制分离标志，一旦置位，控制平面的变化不影响 BFD 检测，如：控制平面为 OSPF，当 OSPF 重启/GR 时，BFD 可以继续检测链路状态。
A	认证标识，置位代表会话需要进行验证。
D	查询请求，置位代表发送方期望采用查询模式对链路进行检测。
M	用于将来应用点到多点时使用，目前必须设置 0。
Detect Mult	检测超时倍数，用于检测方计算检测超时时间。
Length	报文长度。
My Discriminator	BFD 会话连接本端标识符。
Your Discriminator	BFD 会话连接远端标识符。
Desired Min Tx Interval	本地支持的最小 BFD 报文发送间隔。
Required Min RX Interval	本地支持的最小 BFD 报文接收间隔。
Required Min Echo RX Interval	本地支持的最小 Echo 报文接收间隔（如果本地不支持 Echo 功能，则设置 0）。
Auth Type	认证类型(可选)，包括： Simple Password Keyed MD5 Meticulous Keyed MD5 Keyed SHA1 Meticulous Keyed SHA1
Auth Length	认证数据长度。
Authentication Data	认证数据区。

📌 会话状态

BFD 会话有四种基本的状态，分别是 Down、Init、Up 和 AdminDown。

1. Down：会话处于 Down 状态或者刚刚创建。
2. Init：已经和对端系统通信，希望使会话进入 Up 状态。
3. Up：会话已经建立成功。
4. AdminDown：会话处于管理性 Down 状态。

BFD 根据自己的本地会话状态以及接收到的对端 BFD 报文，进行状态机迁移。

BFD 状态机的建立和拆除采用三次握手机制，以确保两端都知道状态的变化。

发送周期和检测时间

BFD 在建立过程中，两端会话会协商 BFD 参数，确定发送周期及检测时间进行会话检测。

在建立 BFD 会话后，可以动态协商 BFD 的相关参数（例如最小发送间隔、最小接收间隔等），两端协议通过发送相应的协商报文后采用新的发送周期和检测时间按，不影响会话的当前状态。

功能特性

功能特性	作用
BFD会话建立	建立 BFD 会话。
BFD会话检测	快速检测双向转发路径。
BFD与关联应用联动	快速通告 BFD 检测结果。
BFD保护策略	受攻击的情况下保护 BFD 的稳定。
BFD震荡抑制通告	线路不稳定的情况下保护关联应用的稳定。

5.3.1 BFD会话建立

BFD 检测开始于 BFD 会话的建立。

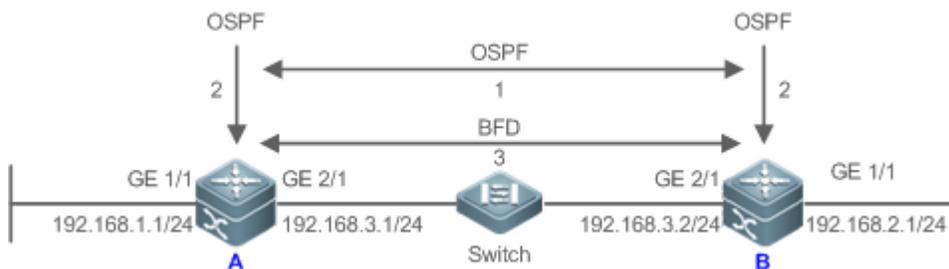
工作原理

会话建立过程

BFD 本身没有发现邻居的能力，需要上层协议通知与哪个邻居建立会话。

如下图所示，两台路由器通过一台二层交换机相连，两台路由器同时运行 OSPF 和 BFD。

图 5-4



BFD 会话建立过程：

1. OSPF 发现邻居后并与邻居建立连接。
2. OSPF 通知 BFD 与该邻居建立会话。
3. BFD 与该邻居建立起会话。

建立 BFD 会话模式

BFD 协议规定建立 BFD 会话的模式，有两种：

- 主动模式

在建立会话前不管是否收到对端发来的建立 BFD 会话的控制报文，都会主动发送建立 BFD 会话的控制报文。

- 被动模式

在建立对话前不会主动发送建立 BFD 会话的控制报文，直到收到对端发送来建立 BFD 会话的控制报文。

i 被动模式暂不支持，且不可配置。

协商 BFD 会话参数

BFD 会话建立过程，两端会进行 BFD 会话的参数协商，从而确定发送周期和检测时间，需要注意以下几点：

1. 必须在两端接口上配置 BFD 会话参数(包括 Desired Min Tx Interval , Required Min RX Interval , Detect Mult) , 否则 BFD 会话无法建立。
2. 在建立 BFD 会话的过程中，两端接口会协商 BFD 会话参数，并依据此会话参数进行会话检测。
3. 在建立 BFD 会话后，可以动态协商 BFD 的相关参数（例如最小发送间隔、最小接收间隔等），两端协议通过发送相应的协商报文后采用新的发送周期和检测时间按，不影响会话的当前状态。

5.3.2 BFD会话检测

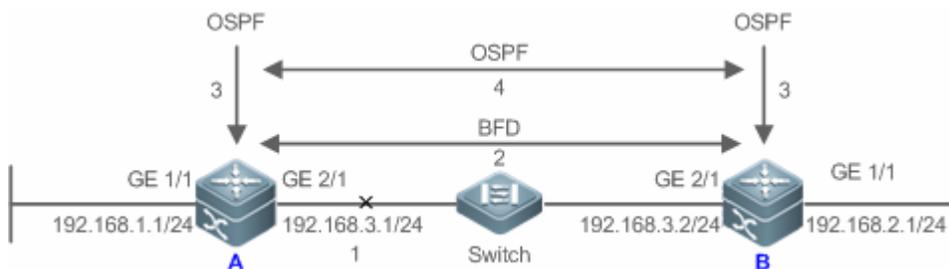
BFD 会话建立后，开始进行链路检测。周期性地发送 BFD 控制报文，如果在检测时间内未收到对端发过来的 BFD 报文，则认为会话 Down，通告联动应用，加快应用协议收敛。

工作原理

检测过程

如下图所示，两台路由器通过一台二层交换机相连，两台路由器同时运行 OSPF 和 BFD。

图 5-5



BFD 会话检测到故障后的处理过程：

1. RouterA 与 Switch 之间的链路通信发生故障。
2. RouterA 和 RouterB 之间的 BFD 会话检测到故障。

3. BFD 通知本地运行的 OSPF 到邻居的转发路径发生故障。
4. OSPF 进行邻居 Down 过程的处理，如果存在备份转发路径那么将进行协议收敛，从而启用备份转发路径。

检测模式

BFD 包含如下几种检测模式。

- 异步模式

在异步模式下，系统之间相互周期性地发送 BFD 控制报文，如果某个系统在检测时间内没有收到对端发来的 BFD 控制报文，就宣布会话为 Down。

- 查询模式

在查询模式下，假定每个系统都有一个独立的方法用来确认它连接到其他系统。这样一旦一个 BFD 会话建立起来以后，系统停止发送 BFD 控制报文，除非某个系统需要显式地验证连接性，在需要显式验证连接性的情况下，系统发送一个短序列的 BFD 控制包，如果在检测时间内没有收到返回的报文就宣布会话为 Down，如果收到对端的回应报文，表示转发路径正常。

- 回声模式

本地系统周期性的发送 BFD 回声报文，远端系统通过它的转发通道将它们环回回来。如果本地在检测周期内连续几个回声报文都没有接收到，会话就被宣布为 Down。回声功能可以和上述两种检测模式一起使用。采用回声报文的检测功能，不需要远端系统的控制面参与，报文通过远端系统的转发面转回，减少了延迟，相对于发送控制报文可以更快的检测到故障。如果在异步模式下启用回声功能，可以大大减少了控制报文的发送，因为检测工作由回声功能完成；如果在查询模式下启用回声功能，在会话建立后可以完全取消发送控制报文。BFD 会话两点必须同时启用回声功能，否则回声功能将不生效。

i 查询模式暂不支持，且不可配置。

i 只有 BFD 会话版本 1 支持 BFD 的回声模式

5.3.3 BFD与关联应用联动

关联应用通过与 BFD 联动，可以利用 BFD 快速检测故障的优点，提高关联应用协议的收敛性能。一般情况下，检测故障的时间可以缩短到 1 秒以内。

工作原理

关联应用配置联动 BFD，下发创建 BFD 会话，BFD 会话建立后进行快速故障检测。当链路出现故障时，BFD 能够快速检测到故障，通告关联应用进行处理，提高关联应用协议的收敛性能。当前 BFD 支持的关联应用有：

- 支持 RIP 联动 BFD

RIP 通过与 BFD 联动，可以利用 BFD 相对于协议自身的“HELLO”机制更快速检测故障的优点，提高协议的收敛性能。一般情况下，检测故障的时间可以缩短到 1 秒以内。

i 关于 RIP 与 BFD 联动的更多内容，请查阅“RIP”章节

- 支持 OSPF 联动 BFD

OSPF 通过与 BFD 联动，可以利用 BFD 相对于协议自身的“HELLO”机制更快速检测故障的优点，提高协议的收敛性能。一般情况下，检测故障的时间可以缩短到 1 秒以内。

i 关于 OSPF 与 BFD 联动的更多内容，请查阅“OSPF”章节

- 支持静态路由联动 BFD

静态路由与 BFD 联动，可以避免在配置的静态路由不可达的情况下，路由选路不会选择该静态路由作为转发路径。如果存在备份路由转发路径，将可以快速地切换到该备份转发路径。

与动态路由协议不同，静态路由没有发现邻居的机制。因此，当配置 BFD 与静态路由关联，静态路由的下一跳可达性将依赖于 BFD 会话状态。如果 BFD 会话检测到故障，表示静态路由的下一跳不可达，则该静态路由将不安装到 RIB 中。

如果 BFD 会话建立过程，远端系统删除 BFD 会话，将会造成 BFD 会话状态变为 Down，在这种情况下系统确保不影响静态路由的转发行为。

i 关于静态路由与 BFD 联动的更多内容，请查阅“NSM”章节

- 支持策略路由联动 BFD

策略路由与 BFD 联动，可以避免在配置的策略路由不可达的情况下，路由选路不会选择该策略路由作为转发路径。如果存在备份路由转发路径，将可以快速地切换到该备份转发路径。

与策略路由联动的方式等同于静态路由。通过 BFD 跟踪检测与指定邻居的转发路径，当会话检测到故障，将通知策略路由到达相应下一跳不可达，达到该下一跳的策略路由将不生效。

如果 BFD 会话建立过程，远端系统删除 BFD 会话，将会造成 BFD 会话状态变为 Down，在这种情况下系统确保不影响策略路由的转发行为。

i 关于策略路由与 BFD 联动的更多内容，请查阅“PBR”章节

- 支持 VRRP 联动 BFD

VRRP 与 BFD 联动可以替代 VRRP 自身的“HELLO”机制实现快速检测主备路由器的运行状态，加快了故障时主备路由器的切换，提高网络的性能。一般情况下，检测故障的时间可以缩短到 1 秒以内。

VRRP 还可以利用 BFD 来跟踪指定的邻居，如果 BFD 会话检测到与该邻居的转发路径发生故障，则自动降低一定数额的 VRRP 优先级，触发主备路由器进行切换。该配置只有在动态路由协议或者其他应用通知 BFD 与相应邻居创建会话时，才会生效。

i 关于 VRRP 与 BFD 联动的更多内容，请查阅“VRRP”章节

1. BFD 检测静态 LSP
2. BFD 检测 LDP 协议生成的 LSP
3. BFD 检测 LSP 反向路径采用 IP

i 关于 MPLS 与 BFD 联动的更多内容，请查阅“MPLS”章节

- 支持三层接口联动 BFD

BFD 支持修改三层接口状态，在配置模式下，通过 `bfd bind peer-ip` 命令来检测指定的三层接口的直连地址，该 CLI 命令所建立的 BFD 会话状态会产生对应接口的 BFD 状态，比如 BFD Down 或者 BFD Up。常用在各类型 FRR 中，通过 BFD 来检测接口状态，进行快速的 FRR 切换。

 三层接口联动 BFD，暂只支持进行 LDP FRR 切换

5.3.4 BFD保护策略

保护 BFD 由于受到攻击(比如：大量 Ping 报文攻击设备)而发生 BFD 会话震荡。

工作原理

BFD 协议是非常敏感协议，如果所启用 BFD 功能的设备受到攻击（比如：大量 Ping 报文攻击设备）而发生 BFD 会话震荡，可以通过配置启用 BFD 的保护策略进行保护。但如果启用该设备 BFD 功能的同时打开该保护策略，会导致上一跳设备发出的 BFD 报文经过该设备时，该设备会将 BFD 报文丢弃，从而影响上一跳设备与其他设备的 BFD 会话建立。

5.3.5 BFD震荡抑制通告

主要解决由于线路不稳定导致 BFD 会话在 Down 和 Up 状态之间频繁切换，从而引起关联的应用（比如静态路由）频繁的进行转发路径切换，影响业务的正常运行的问题。

工作原理

BFD 会话在 Down 和 Up 状态之间频繁切换，允许用户配置通告给关联应用会话 Up 状态前所需 Up 状态稳定的时间。在 Up 状态保持在特定时间后，才通告关联应用 BFD Up，否则通告 BFD Down。

5.4 配置详解

配置项	配置建议 & 相关命令	
配置BFD基本功能	 必须配置。用于建立 BFD 会话。	
	bfd interval	配置 BFD 参数
	-	配置关联应用联动 BFD
		 关联的应用不同，配置命令会不同，具体参见各应用相关章节。此处不一列出
	 可选配置。用于配置 BFD 的检测模式、慢速报文发送周期和 BFD 联动三层接口。	
	bfd echo	配置回声模式
	bfd slow-timer	配置慢速发送控制报文周期
	bfd bind peer-ip	配置 BFD 联动三层接口
配置BFD保护策略	 可选配置。用于保护 BFD 免受攻击的影响。	

	bfd cpp	配置 BFD 保护策略
配置BFD震荡抑制通告	 可选配置。用于关联协议免受 BFD 震荡的影响。	
	bfd up-dampening	配置 BFD 震荡抑制通告

5.4.1 配置BFD基本功能

配置效果

- 关联应用联动上 BFD。
- 建立起 BFD 会话
- BFD 会话进行链路故障检测。

注意事项

- 配置 BFD 会话参数，需要注意：
 1. 建议 BFD 会话两端的参数配置一致，这样可以确保关联 BFD 应用协议同时生效，避免由于两端配置的抑制时间不同而出现转发路径单通的情况。
 2. 配置时设置的参数需要考虑不同接口传输上的带宽差异。如果设置最小发送间隔和最小接受间隔过小，可能导致 BFD 占用过大带宽而影响本身的数据传输。
- 配置关联应用联动 BFD，需要注意：
 1. 配置时需要确保 BFD 会话邻居都启用关联应用联动 BFD，否则 BFD 会话将无法建立。但如果已经有动态路由协议或者其他应用通知 BFD 与相应邻居创建会话，那么应用关联 BFD 会话也将建立。
 2. 如果由于 IP 选路而导致 BFD 会话邻居指定的接口和实际 BFD 报文出入口不一致，或创建 BFD 会话时指定的接口和实际 BFD 回来的报文入接口不一致，则无法建立 BFD 会话。
- 配置 BFD 检测模式，需要注意：
 1. 本端的回声报文发出，在对端设备转发面处理后返回到本端，这个过程可能由于对端设备拥塞造成回声报文丢失，引发会话检测失败。在这种情况下，需要配置相应的 QOS 策略来确保回声报文优先得到处理或者关闭回声功能。
 2. BFD 的回声检测功能不支持多跳检测，所以在配置多跳时，请确认回声功能已经关闭。
 3. Echo 模式必须是 BFD 会话的两端系统都使能该模式才能生效。
 4. 在 BFD 启用 echo 模式前，需要在 BFD 会话的邻居设备上执行 **no ip redirects** 命令关闭发送 ICMP 重定向报文的功能，执行 **no ip deny land** 命令关闭 DDOS 功能(防止 Land-based 攻击)。

配置方法

配置 BFD 参数

- 必须配置。
- 若无特殊要求，应在 BFD 检测的两端路由器的 BFD 会话出口上配置 BFD 参数。
- 配置时设置的参数需要考虑不同接口传输上的带宽差异。如果设置最小发送间隔和最小接收间隔过小，可能导致 BFD 占用过大带宽而影响本身的数据传输。

【命令格式】 **bfd interval milliseconds min_rx milliseconds multiplier interval-multiplier**

【参数说明】 **interval milliseconds**：最小发送间隔，单位毫秒。

min_rx milliseconds：最小接收间隔，单位毫秒。

multiplier interval-multiplier：检测超时倍数

【缺省配置】 无 BFD 会话参数

【命令模式】 接口配置模式

【使用指导】 该命令不允许在 L3 AP 接口下进行配置。

在路由器上启用 BFD 功能前，必须先启用快转功能。

▾ 配置 BFD 检测模式

- 可选配置，端口默认运行异步模式，如果 BFD 会话要运行于回声模式，则需要配置。
- 在交换机或路由器的端口下配置。
- 只要两端路由器的其中一端配置异步模式，BFD 会话就会运行于异步模式；两端同时默认配置运行于回声模式，则 BFD 会话最终运行于回声模式。

【命令格式】 **bfd echo**

【参数说明】 -

【缺省配置】 BFD 回声模式关闭

【命令模式】 接口配置模式

【使用指导】 该命令不允许在 L3 AP 接口下配置。

缺省情况下配置了 BFD 会话参数的同时系统自动使能 echo 模式。

echo 报文的最小发送间隔和最小接收间隔采用会话配置的 Interval milliseconds 和 min_rx milliseconds 参数。

在 BFD 启用 echo 模式前，需要在 BFD 会话的邻居设备上执行 **no ip redirects** 命令关闭发送 ICMP 重定向报文的功能，执行 **no ip deny land** 命令关闭 DDOS 功能(防止 Land-based 攻击)。

▾ 配置慢速报文发送周期

- 可选配置，默认慢速报文发送周期为 3000ms，如果需要增加或减少 BFD 慢速报文的发送周期，则可以配置。
- 在交换机或路由器的全局配置模式下配置。
- BFD 运行在 ECHO 模式或者 BFD 建立过程，以该周期来发送慢速控制报文，配置周期越大，协商建立 BFD 会话的时间越长，ECHO 模式下发送的慢速 BFD 报文时间越长。

【命令格式】 **bfd slow-timer [milliseconds]**

【参数说明】 **milliseconds**：BFD 的慢速定时器时间，单位为毫秒。可配置范围从 1000 到 30000，未配置缺省值为 3000。

【缺省配置】 慢速控制报文发送周期为 3000ms

【命令模式】 全局配置模式

【使用指导】 此命令用来指定 ECHO 模式下发送慢速控制报文的周期。

配置 BFD 联动三层接口

- 可选配置。目前，BFD 关联三层接口仅在 MPLS LDP 做 FRR 快速切换时使用。
- 在交换机或路由器的端口下配置。

【命令格式】 **bfd bind peer-ip** *src-address* [**source-ip** *dst-address*] **process-pst**

【参数说明】 *src-address* : 接口对端的 ip 地址

dst-address : 接口本端 ip 地址

【缺省配置】 缺省无三层口关联 BFD 配置

【命令模式】 接口配置模式

【使用指导】 用于指定三层接口联动 BFD，可快速检测三层接口的连通性。

配置关联应用联动 BFD

- 必须配置。
- 缺省情况下关联应用联动 BFD 未开启。
- 关联的应用不同，配置命令会不同，具体参见各应用相关章节。
- 必须确保两端均配置关联应用联动 BFD，BFD 会话才能建立起来。
- 在 RIP 路由配置模式下，使用 **bfd all interfaces** 命令开启允许所有接口 RIP 关联 BFD，详细配置参考 RIP 相关章节。
- 在 OSPF 路由配置模式下，使用 **bfd all interfaces** 命令开启允许所有接口 OSPF 关联 BFD，详细配置参考 OSPF 相关章节。
- 在全局配置模式下，使用 **ip route static bfd** *interface-type interface-number gateway* [**source** *ip-address*] 命令开启静态路由联动 BFD，详细配置参考 NSM 相关章节。
- 使用 **set ip next-hop verify-availability** *next-hop-address* **bfd** *interface-type interface-number gateway* 命令开启策略路由联动 BFD，详细配置参考 PBR 相关章节。
- 使用 **vrrp bfd** *interface-type interface-number ip-address* 命令开启 VRRP 联动 BFD，详细参考见 VRRP 相关章节。
- 使用 **bfd bind static-lsp peer-ip** *ip-address* **source-ip** *ip-address* [**local-discriminator** *discr-value* **remote-discriminator** *discr-value*] [**process-state**] 命令开启静态 LSP 联动 BFD，详细配置见 MPLS 相关章节。
- 使用 **bfd bind ldp-lsp peer-ip** *ip-address* **nexthop** *ip-address* [**interface** *interface-type interface-number*] **source-ip** *ip-address* [**local-discriminator** *discr-value* **remote-discriminator** *discr-value*] [**process-state**] 命令开启 LDP LSP 联动 BFD，详细配置见 MPLS 相关章节。
- 使用 **bfd bind backward-lsp-with-ip peer-ip** *ip-address* **interface** *interface-type interface-number* [**source-ip** *ip-address*] { **local-discriminator** *discr-value* **remote-discriminator** *discr-value* } 命令开启动态 LSP 联动 BFD，详细配置见 MPLS 相关章节。

检验方法

- 关联的应用不同，其检验的方式也不尽相同，具体参见各应用相关章节。

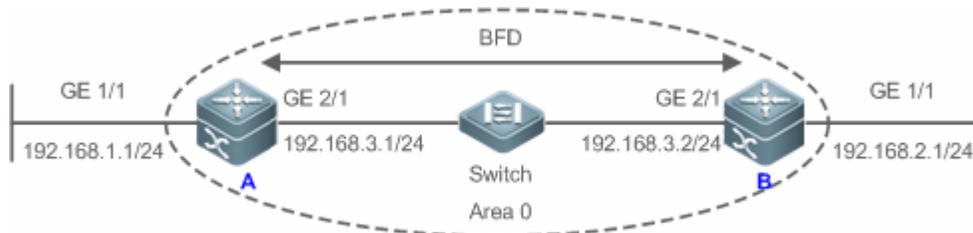
配置举例

i 以下配置举例，仅介绍与 BFD 相关的配置。

OSPF 联动 BFD

【网络环境】

图 5-1



【配置方法】

- 在 Router A 和 Router B 相连接口配置 IP 地址
- 在 Router A 和 Router B 运行 OSPF 协议
- 在 Router A 和 Router B 相连接口配置 BFD 参数
- 在 Router A 和 Router B 使能 OSPF 联动 BFD

A

```
A# configure terminal
A(config)# interface GigabitEthernet2/1
A(config-if-GigabitEthernet2/1)# no switchport           //路由器无需此配置
A(config-if-GigabitEthernet2/1)# ip address 192.168.3.1 255.255.255.0
A(config-if-GigabitEthernet2/1)# bfd interval 200 min_rx 200 multiplier 5
A(config-if-GigabitEthernet2/1)# exit
A(config)# interface GigabitEthernet1/1
A(config-if-GigabitEthernet1/1)# no switchport           //路由器无需此配置
A(config-if-GigabitEthernet1/1)# ip address 192.168.1.1 255.255.255.0
A(config-if-GigabitEthernet1/1)# exit
A(config)# router ospf 123
A(config-router)# log-adj-changes detail
A(config-router)# network 192.168.3.0 0.0.0.255 area 0
A(config-router)# network 192.168.1.0 0.0.0.255 area 0
A(config-router)# bfd all-interfaces
A(config-router)# end
```

B

```
B# configure terminal
B(config)# interface GigabitEthernet2/1
B(config-if-GigabitEthernet2/1)# no switchport           //路由器无需此配置
B(config-if-GigabitEthernet2/1)# ip address 192.168.3.2 255.255.255.0
B(config-if-GigabitEthernet2/1)# bfd interval 200 min_rx 200 multiplier 5
B(config-if-GigabitEthernet2/1)# exit
B(config)# interface GigabitEthernet1/1
B(config-if-GigabitEthernet1/1)# no switchport           //路由器无需此配置
B(config-if-GigabitEthernet1/1)# ip address 192.168.2.1 255.255.255.0
B(config-if-GigabitEthernet1/1)# exit
```

```

B(config)# router ospf 123
B(config-router)# log-adj-changes detail
B(config-router)# network 192.168.3.0 0.0.0.255 area 0
B(config-router)# network 192.168.2.0 0.0.0.255 area 0
B(config-router)# bfd all-interfaces
B(config-router)# end

```

【检验方法】 显示验证。

A

```

A# show bfd neighbors details
OurAddr      NeighAddr    LD/RD    RH/RS    Holdown(mult)    State    Int
192.168.3.1  192.168.3.2  1/2      Up       532 (3 )         Up       Ge2/1
Local Diag: 0, Demand mode: 0, Poll bit: 0
MinTxInt: 200000, MinRxInt: 200000, Multiplier: 5
Received MinRxInt: 50000, Received Multiplier: 3
Holdown (hits): 600(22), Hello (hits): 200(84453)
Rx Count: 49824, Rx Interval (ms) min/max/avg: 208/440/332
Tx Count: 84488, Tx Interval (ms) min/max/avg: 152/248/196
Registered protocols: OSPF
Uptime: 02:18:49
Last packet: Version: 1      - Diagnostic: 0
I Hear You bit: 1           - Demand bit: 0
Poll bit: 0                  - Final bit: 0
Multiplier: 3                - Length: 24
My Discr.: 2                  - Your Discr.: 1
Min tx interval: 50000       - Min rx interval: 50000
Min Echo interval: 0

```

B

```

B# show bfd neighbors details
OurAddr      NeighAddr    LD/RD    RH/RS    Holdown(mult)    State    Int
192.168.3.2  192.168.3.1  2/1      Up       532 (5 )         Up       Ge2/1
Local Diag: 0, Demand mode: 0, Poll bit: 0
MinTxInt: 50000, MinRxInt: 50000, Multiplier: 3
Received MinRxInt: 200000, Received Multiplier: 5
Holdown (hits): 600(22), Hello (hits): 200(84453)
Rx Count: 49824, Rx Interval (ms) min/max/avg: 209/440/332 last: 66 ms ago
Tx Count: 84488, Tx Interval (ms) min/max/avg: 153/249/197 last: 190 ms ago
Registered protocols: OSPF
Uptime: 02:18:49
Last packet: Version: 1      - Diagnostic: 0
I Hear You bit: 1           - Demand bit: 0
Poll bit: 0                  - Final bit: 0
Multiplier: 5                - Length: 24

```

```
My Discr.: 1          - Your Discr.: 2
Min tx interval: 200000 - Min rx interval: 200000
Min Echo interval: 0
```

常见错误

- 两端设备有一端接口未配置 BFD 参数。
- 没有使能关联应用联动 BFD。
- 两端设备只有一端使能应用联动 BFD。

5.4.2 配置BFD保护策略

配置效果

- 如果所启用 BFD 功能的设备受到攻击（比如:大量 Ping 报文攻击设备）而发生 BFD 会话震荡，可以通过配置启用 BFD 的保护策略进行保护。

注意事项

- 必须配置 BFD 基本功能
- 如果启用该设备 BFD 功能的同时打开该保护策略，会导致上一跳设备发出的 BFD 报文经过该设备时，该设备会将 BFD 报文丢弃，从而影响上一跳设备与其他设备的 BFD 会话建立。
- 该功能及限制仅对交换机生效。

配置方法

▾ 配置 BFD 保护策略

- 可选配置。
- 在交换机或路由器的全局配置模式下配置。
- BFD 保护策略是用于在设备受到攻击的场景下，提升 BFD 报文的处理优先级，保证 BFD 业务的正常运行。

【命令格式】 **bfd cpp**

【参数说明】 -

【缺省配置】 BFD 保护策略开启

【命令模式】 全局模式

【使用指导】 如果设备受攻击导致的 BFD 震荡，可开启功能进行保护。

检验方法

`show running-config` 查看对应接口下是否存在该配置。

配置举例

配置 BFD 保护策略

- 【配置方法】
- 在存在攻击行为网络中的交换机上配置。
 - 配置 BFD 保护策略

```
Ruijie# configure terminal
Ruijie(config)# bfd cpp
Ruijie(config)# end
```

- 【检验方法】 -

常见错误

无

5.4.3 配置BFD震荡抑制通告

配置效果

- 线路不稳定导致 BFD 会话在 Down 和 Up 状态之间频繁切换，从而引起关联的应用（比如静态路由）频繁的进行转发路径切换，影响业务的正常运行的问题。
- 用户配置通告给关联应用会话 Up 状态前所需 Up 状态稳定的时间。在 Up 状态保持在特定时间后，才通告关联应用 BFD Up，否则通告 BFD Down。减少线路不稳定导致的关联应用协议的震荡。

注意事项

- 必须配置 BFD 基本功能。
- 如果在 BFD 没有出现 Down 和 Up 状态频繁切换，启用 BFD 震荡抑制通告，会导致延迟通告关联应用会话 Up。

配置方法

配置 BFD 震荡抑制通告

- 可选配置，端口默认未开启 BFD 震荡抑制通告，如果在 BFD 出现 Down 和 Up 状态频繁切换，则可以启用 BFD 震荡抑制通告。
- 在交换机或路由器的端口下配置

- 开启 BFD 震荡抑制通告，可以缓解因 BFD 频繁通告状态变化而导致关联应用的频繁处理，如路由重新计算等；配置的时间越长则 BFD 需要稳定的时间越长，只有在稳定时间达到配置的时间，才会通告相应的应用模块。

【命令格式】 **bfd up-dampening** [*milliseconds*]

【参数说明】 *milliseconds* : 通告给关联应用会话 UP 状态前所需 UP 状态稳定的时间，单位为毫秒。可配置范围从 0 到 300000，配置 0,即当会话的状态从 DOWN 切换为 UP 时将立即通告给应用层.缺省值为 0。。

【缺省配置】 BFD 震荡抑制通告功能未开启

【命令模式】 接口配置模式

【使用指导】 只有在线路不稳定时，才需要开启该功能。

如果在 BFD 没有出现 Down 和 Up 状态频繁切换，配置该命令，会导致延迟通告关联应用会话 Up。

检验方法

show running-config 查看对应接口下是否存在该配置。

配置举例

配置 BFD 震荡抑制通告时间为 60,000 毫秒。

- 【配置方法】
- 在链路不够稳定会导致 BFD 频繁震荡的环境下配置。
 - 配置 BFD 震荡抑制通告时间为 60,000 毫秒

```
Ruijie# configure terminal
Ruijie(config)# interface fastEthernet 0/2
Ruijie(config)# bfd up-dampening 60000
Ruijie(config)# end
```

【检验方法】 -

常见错误

无

5.5 监视与维护

清除各类信息

无

查看运行情况

作用

命令

查看 BFD 会话信息。	<code>show bfd neighbors [client { ospf rip vrrp static-route pbr ldp-lsp static-lsp backward-lsp-with-ip pst }][ipv4 ip-address][details]</code>
--------------	---

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 BFD 事件的调试开关。	<code>debug bfd event [interface interface-type interface-number ipv4 ip-address]</code>
打开 BFD 报文的调试开关	<code>debug bfd packet [interface interface-type interface-number ipv4 ip-address]</code>

6 IP Event Dampening

6.1 概述

对三层设备来说，当三层网络接口由于人为的开启/关闭或者由于其他外在因素导致状态频繁 UP/DOWN 变化时，会造成本设备上路由表的反复震荡。若配置了路由协议，那么路由协议还可能会将这一震荡扩散到整个网络中，造成邻居路由反复进行路由的更新和重新计算，不仅浪费网络带宽，而且导致整个网络不稳定。从设备本身的角度来说，反复的路由更新和计算需要消耗大量的 CPU 资源。这些都会影响客户网络的正常运行。

IP Event Dampening (IP 事件惩罚) 功能能够检测异常的 UP/DOWN 翻转，并自动抑制接口状态频繁 UP/DOWN，使单点链路故障不会被路由协议扩散出去；当接口恢复稳定后，抑制状态将自动解除，从而减少网络震荡，降低系统的 CPU 资源消耗，增强整网的稳定性。

协议规范

- RFC2439 : BGP Route Flap Dampening

i IP Event Dampening 采用的抑制算法核心与 RFC2439 所述的 BGP 路由翻转抑制算法一致。

6.2 典型应用

典型应用	场景描述
抑制路由接口震荡	在路由器三层网络接口上监控接口状态变化，对接口的频繁震荡进行抑制。

6.2.1 抑制路由接口震荡

应用场景

在运行路由协议的网络中，两两相连的路由器之间一旦有端口频繁 UP/DOWN，将引起邻居路由反复更新计算路由，进而扩散到整个网络，导致网络震荡。通过在相连的路由口上启用 IP Event Dampening 功能，可以监控接口状态变化，对接口的频繁震荡进行抑制，减少网络震荡，降低系统的 CPU 资源消耗，增强整网的稳定性。

图 6-1



【注释】 A、B 为普通路由器。

功能部属

以上图为例，可以在路由器 A 和 B 的接口 GE0/1 上配置 IP Event Dampening 功能。

 路由器上的子接口和接口虚模板 Virtual template，不支持 dampening 功能。

6.3 功能详解

基本概念

▾ 惩罚值 (penalty)

IP Event Dampening 功能采用了一种对接口状态变化事件计算惩罚值的方式，对接口行为进行感知和控制。接口上每进行一次 up→down 翻转，惩罚值将会累加，待接口稳定时，该惩罚值又会按指数递减。通过这种惩罚值计算，使得接口能智能地感知其自身的状态行为，进行相应处理。

▾ 抑制门限 (Suppress Threshold)

为了识别接口震荡事件，IP Event Dampening 功能设置了一个抑制门限，当累积惩罚值超过该门限时，认为该接口处于震荡状态，并对接口实施抑制。

▾ 半衰期 (Half-Life Period)

半衰期是指在接口稳定情况下惩罚值衰减到原来一半所需要的时间。它定义了惩罚值按指数衰减的速度，半衰期越短，衰减的越快，探测到接口稳定的速度越快，但震荡检测的灵敏度也越低。

▾ 解除抑制门限 (Reuse Threshold)

当接口不再震荡，惩罚值衰减到一定程度（减少到抑制门限以下）时，可认为该接口恢复稳定状态，从而解除接口的抑制状态。

▾ 最大抑制时间 (Maximum Suppress Time)

为避免接口长时间震荡后，惩罚值过大，导致过久不能恢复使用，算法定义了一个最大抑制时间，即无论接口震荡多久，接口被抑制的时间不会超过最大抑制时间。

功能特性

功能特性	作用
接口震荡抑制	允许用户在接口上配置接口震荡抑制条件和参数，使得交换机设备能自动对反复震荡的接口进行识别和抑制，确保本机的路由稳定性和避免路由震荡的扩散。

6.3.1 接口震荡抑制

工作原理

当接口上配置了 Dampening 功能时，该接口上会被附上一个与之相关的惩罚值（Penalty）。每当接口 DOWN 的时候，该接口上的 Penalty 增加 1000，该 Penalty 值随时间递减，若递减过程中，接口再次出现 DOWN 事件，则接口上 Penalty 值进行累加，若累计 Penalty 值超过抑制门限（Suppress Threshold）则判断该接口处于抑制（Suppress）状态，之后无论接口的真实状态是什么，对受影响的上层协议来说，该接口始终处于 DOWN 状态。直到 Penalty 递减到解除抑制门限（Reuse Threshold），接口的抑制状态解除，还原接口的真实状态。

当三层接口上未配置 Dampening 功能，或配置了 Dampening 功能但接口未被抑制时，路由协议或其他关心三层接口状态的协议都和正常情况下的行为一样。当接口上配置了 Dampening 且接口被抑制时，上层协议认为接口状态为 down，这期间接口实质上的状态变化不会对路由表以及上层路由协议路由计算和通告造成影响，直到接口解除抑制为止。

相关配置

配置 IP Event Dampening 功能

- 缺省情况下，三层接口上的 IP Event Dampening 功能关闭。
- 使用 `dampening [half-life-period [reuse-threshold suppress-threshold max-suppress [restart [restart-penalty]]]]` 命令可以启动或关闭接口上的 IP Event Dampening 功能。

6.4 配置详解

配置项	配置建议 & 相关命令	
启动IP Event Dampening	 必须配置。用于抑制三层接口震荡。	
	<code>dampening</code>	配置 IP Event Dampening 功能

6.4.1 启动IP Event Dampening

配置效果

在配置有 IP Event Dampening 的接口上，若发现震荡情况超过用户预先设定的条件参数时，将该接口状态设置为 down。

注意事项

- 对于交换机接口层次转换（三层口转换为 2 层口），如从 routed port 转换为 switch port，该接口上配置的 `dampening` 命令将被删除。

配置方法

配置 IP Event Dampening 功能

- 必须配置
- 在三层接口模式下配置。
- 配置的时候可以指定半衰期、解除抑制门限、抑制门限、最大抑制时间和初始惩罚值，如果没有指定则使用默认值。

检验方法

可以使用以下两种方法来确认配置是否生效。

- **show running-config**。
- **show interfaces [interface-id] dampening**，查看接口 IP Event Dampening 配置情况。

相关命令

打开接口上的 Ip Event Dampening 功能

【命令格式】 **dampening** [*half-life-period* [*reuse-threshold* *suppress-threshold* *max-suppress* [**restart** [*restart-penalty*]]]]]]

【参数说明】 *half-life-period* : 半衰期，范围<1-30>，默认 5s。
reuse-threshold : 解除抑制门限，范围<1-20000>，默认 1000。
suppress-threshold : 抑制门限，范围<1-20000>，默认 2000。
max-suppress : 最大抑制时间，范围<1-255>，默认为 *half-life-period* 的 4 倍。
restart *restart-penalty* : 初始惩罚值，范围<1-20000>，默认 2000。

【命令模式】 接口模式

【使用指导】 该功能会影响直连/主机路由，静态路由，动态路由和 VRRP。当一个接口满足命令的配置条件，处于抑制状态时，以上受影响的模块认为该接口状态为 DOWN，从而会删除对应的路由，并且不会从该接口收发数据包。当在一个已经配置了 **dampening** 命令的接口上，重新配置命令，会使该接口的所有 dampening 信息清空，但接口的翻转次数仍会保留，除非使用 **clear counters** 命令清除接口的统计信息。如果配置的 *max-suppress* 太小，导致计算得出的最大惩罚值比抑制门限还要小，则该接口将永远得不到抑制。这属于配置错误，此时会打印类似如下的信息，提示用户配置失败：

```
% Maximum penalty (10) is less than suppress penalty (2000). Increase maximum suppress time
```

另外，当配置该命令时，系统内存不足以执行该配置保存，也会打印配置失败的提示：

```
% No memory, configure dampening fail!
```

配置举例

在三层接口上配置 Ip Event Dampening 功能

【网络环境】

图 6-2



【配置方法】 在路由器 A 和 B 接口 GigabitEthernet 0/1 上开启 IP Event Dampening 功能，半衰期 30 秒，解除抑制门限 1500，开始抑制门限 10000，以及最长抑制时间为 120 秒。

A
 Ruijie(config)# interface GigabitEthernet 0/1
 Ruijie(config-if-GigabitEthernet 0/1)# dampening 30 1500 10000 100

B
 Ruijie(config)# interface GigabitEthernet 0/1
 Ruijie(config-if-GigabitEthernet 0/1)# dampening 30 1500 10000 100

【检验方法】 通过 **show interfaces dampening** 命令查看接口的 dampening 的详细信息。

```
Ruijie#show interfaces dampening
GigabitEthernet 0/1
  Flaps Penalty Supp   ReuseTm HalfL   ReuseV  SuppV  MaxSTm  MaxP  Restart
  0      0      FALSE  0      30      1500   1000   100    15119  0
```

常见错误

- 在三层交换机下配置，接口必须先 **no swithport**，转为路由口后才能配置。

6.5 监视与维护

清除各类信息

作用	命令
清除接口的统计信息。	clear counters

i **clear counter** 命令相关说明请参考“接口”命令的相关章节。

查看运行情况

作用	命令
显示被惩罚的接口统计信息。	show dampening interface
显示接口 IP Event Dampening 配置情况。	show interfaces dampening

查看调试信息

! 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

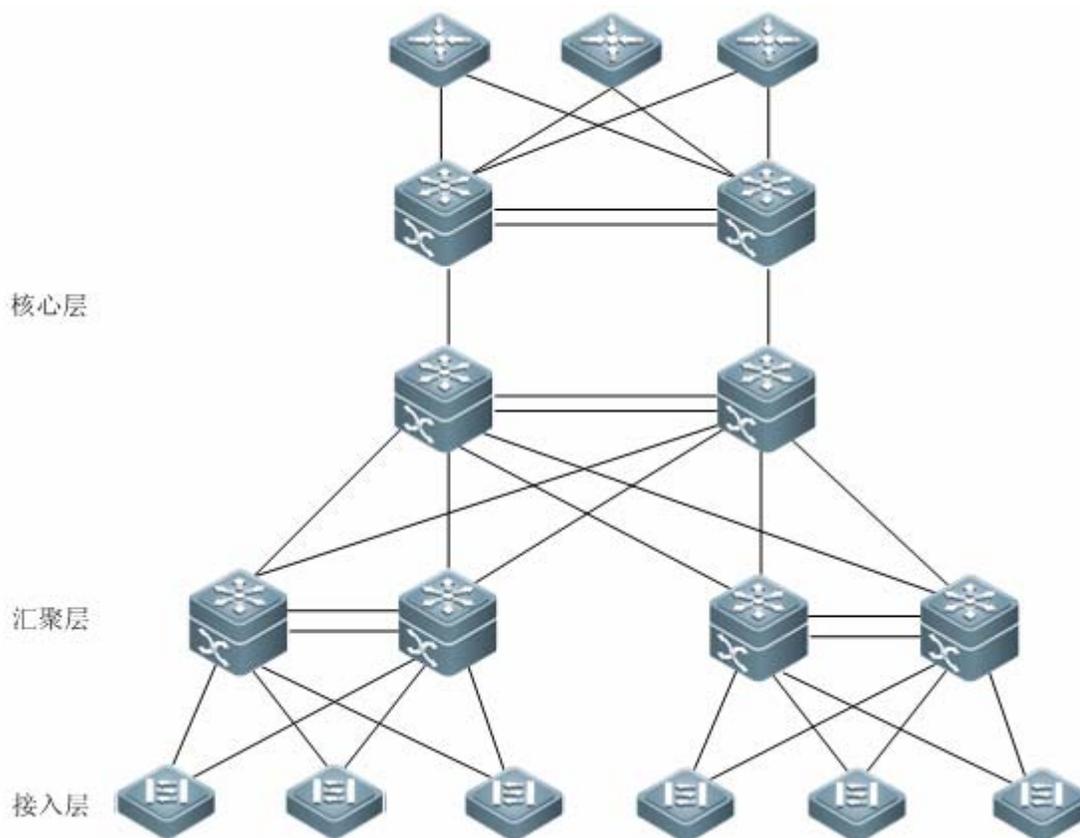
作用	命令
打开 Dampening 功能的调试开关。	debug dampening interface

7 VSU

7.1 概述

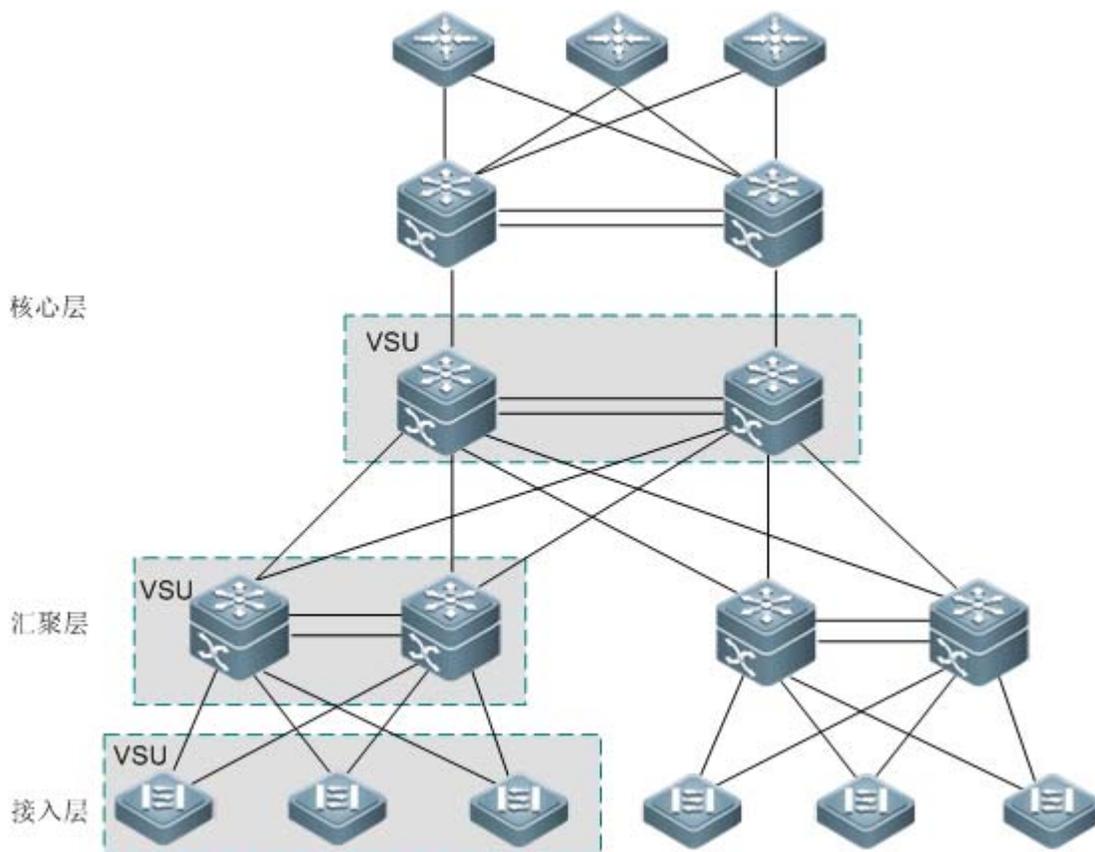
传统的网络中，为了加强网络的可靠性，一般在将核心层和汇聚层配置成双设备，起冗余备份作用，邻居设备分别连接两条链路到双设备上。下图显示的就是这样的一种典型的传统网络架构。冗余的网络架构增加了网络设计和操作的复杂性，同时大量的备份链路也降低了网络资源的利用率，减少了投资回报率。

图 7-1 传统网络结构



VSU (Virtual Switching Unit) 是一种网络系统虚拟化技术，支持将多台设备组合成单一的虚拟设备。如下图所示，接入、汇聚、核心层设备都可以组成 VSU，形成整网端到端的 VSU 组网方案。和传统的组网方式相比，这种组网可以简化网络拓扑，降低网络的管理维护成本，缩短应用恢复的时间和业务中断的时间，提高网络资源的利用率。

图 7-2 端到端的 VSU 组网方案



协议规范

- -

7.2 典型应用

典型应用	场景描述
多台设备统一管理	多台物理设备组成一台逻辑设备，统一管理。
简化网络结构	VSU 看做一台逻辑设备，简化网络结构。

7.2.1 多台设备统一管理

应用场景

当多台物理设备组成 VSU 时，可以看成一台逻辑设备。所有的配置都在全局主设备上管理。

以下图为例四台设备(设备编号从左到右，依次编号为 1、2、3、4)组成 VSU，设备 1 是全局主设备，设备 2 是全局从设备，设备 3 和设备 4 为全局候选设备。

- 对所有设备的管理只要在全局主设备上配置

图 7-3



【注释】 上图设备从左到右，编号依次为 1、2、3、4

VSL 见 1.3.1 描述

设备 1 为全局主设备

设备 2 为全局从设备

设备 3,4 为全局候选设备

功能部属

- 全局主设备负责控制整个 VSU 系统，运行控制面协议并参与数据转发；
- 全局从设备参与数据转发，并不运行控制面协议，并且作为备份当全局主设备发生故障接替全局主设备工作；
- 全局候选设备参与数据转发，并不运行控制面协议。当全局从设备发生故障时，全局候选设备可以接替全局从设备工作，此时设备角色也由此变成全局从设备，全局候选设备不能接替全局主设备工作，因此当全局主设备和全局从设备发生故障，VSU 系统会重启。

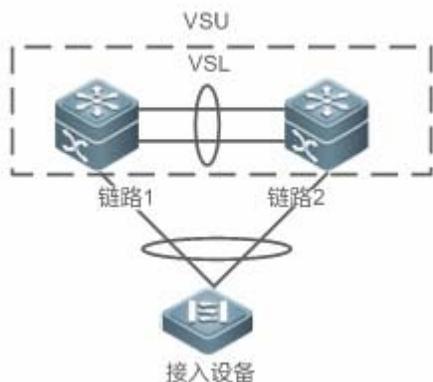
7.2.2 简化网络结构

应用场景

如图 1-1 传统网络中，为了增加网络组网的可靠性，需要增加设备和线路冗余，然而为了防止环路，需要引进许多防止环路的算法。导致网络的组网复杂。VSU 系统中，所有的设备认为是一台逻辑设备。设备之间可以互为备份，不需要引入防止环路算法，就可以简单的组网。

- 两台汇聚交换机组成 VSU。不需要配置防环路算法，两台设备可以互相冗余。
- 接入交换机，通过上联 AP 接入到汇聚交换机。
- 当 VSU 中一台设备出现故障，另一条链路还可以正常工作。

图 7-4



功能部属

- 全局主设备负责控制整个 VSU 系统，运行控制面协议并参与数据转发；
- 全局从设备参与数据转发，并不运行控制面协议，并且作为备份当全局主设备发生故障接替全局主设备工作；
- 接入设备面向用户，用于用户设备的接入。

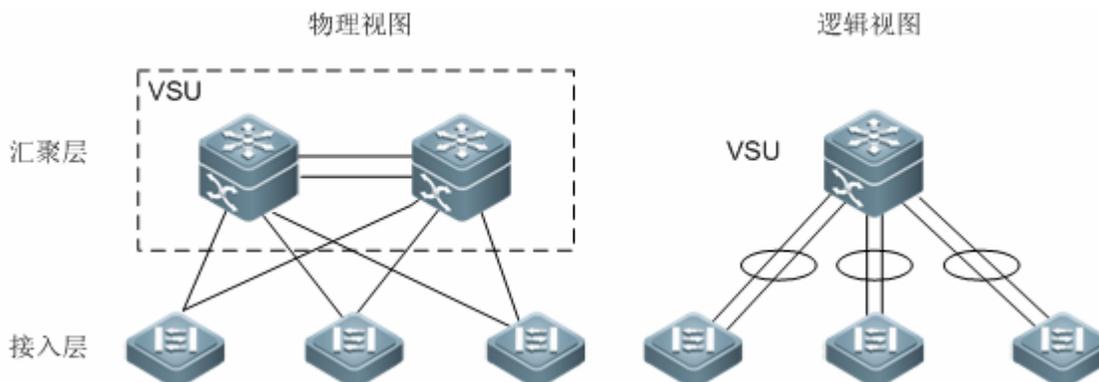
7.3 功能详解

基本概念

VSU 系统

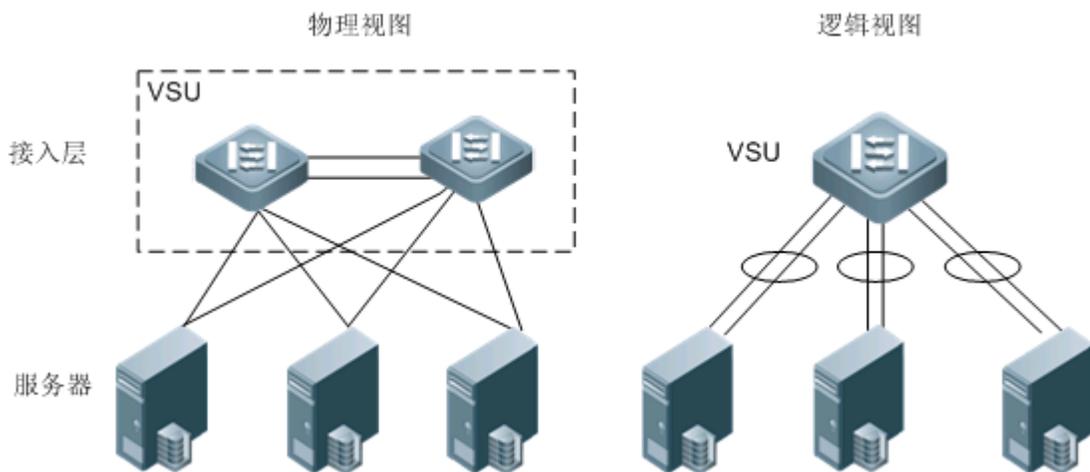
VSU 系统是由传统网络结构中的两台冗余备份的双设备组成的单一的逻辑实体，例如位于下图的汇聚层的 VSU 系统可以看作单独的一台交换机与接入层进行交互。

图 7-5 汇聚层的 VSU



上图的 VSU 网络结构中，成员之间通过内部的链路组成逻辑实体，接入层设备通过聚合链路与 VSU 建立接入到汇聚层。这样在接入层和汇聚层之间避免了二层环路。

图 7-6 接入层的 VSU



除了核心、汇聚层设备外，接入层设备也可以组成 VSU 系统。对于接入可用性要求高的服务器，一般使用“单服务器多网卡绑定为 Aggregate Port 口(简称 AP 口)”技术来与接入层设备相连。由于 AP 要求只能接入在同一台接入设备上，所以单台设备故障的风险增加了。在这种情况下，可以应用 VSU 解决这个问题。在 VSU 模式下，服务器可以使用多网卡绑定为 AP 口，连接同一个 VSU 组内不同的成员设备，这样可以防止接入设备的单点失效，或是单条链路失效导致的网络中断。

VSU 域标识

VSU 域具有唯一标识 Domain ID。只有同一个域标识的设备才能组合在一起形成同一个 VSU 系统。

成员设备编号

VSU 系统的每个成员设备都拥有唯一的编号，即 Switch ID。这个编号用于管理成员设备，以及配置成员设备上的接口等用途。用户在将设备加入 VSU 系统时需要配置该编号，并且保证成员设备编号在同一个 VSU 系统中是唯一的。VSU 系统如果发现成员设备编号冲突，根据优先级保留一台设备。

成员设备角色

VSU 系统由多台设备构成，在组建 VSU 系统时，多台设备通过一定的竞选协议选举出一台全局主设备，在支持 1:N 热备下其余为全局从设备。在支持 1:1 热备下，一台为全局主设备，一台为全局从设备，其余为候选设备。

全局主设备负责控制整个 VSU 系统，运行控制面协议并参与数据转发；其余的设备仅参与数据转发，并不运行控制面协议，所有接收到的控制面数据流都将转发给全局主设备进行处理。

全局从设备同时还实时同步接收全局主设备的状态。与全局主设备构成 1:1 或 1:N 热备份。在全局主设备失效后，全局从设备将切换成全局主角色，来管理整个 VSU 系统。

i VSU 系统的主机选举方法为：

- VSU 系统的主机选举规则如下（如果根据上一条规则不能决定主机，则根据下一条规则继续判断）：a) 当前运行的主机最优先选为主机（起机时所有设备都不是主机）。b) 优先级高的成员设备选为主机。c) 设备号小的优先为主机。d) MAC 地址小的成员设备选为主机。

- 在 1:N 热备下，选择从机的时候，优先选择与主机相近的设备为从机，这样可以尽量避免产生双主机，选择从机的条件排序为：最靠近主机/优先级/MAC 地址。
- VSU 系统支持设备的热加入。即使热加入设备的优先级比当前运行的 VSU 系统主机和从机优先级高，系统也不会进行主、从角色切换。
- 成员设备的启机顺序可能会影响主机的选举。部分成员设备可能由于启机慢（目前 VSU 系统中，在 5 分钟内没有发现邻居就直接收敛），而没有及时加入 VSU 系统。在这种情况下，该成员设备将做热加入处理，即使优先级比当前运行的 VSU 系统主机高，系统也不会发生角色切换。

功能特性

功能特性	作用
虚拟交换链路	VSU 系统内，用于连接各个设备的虚拟路径。
拓扑	介绍 VSU 系统内连接的拓扑结构。
多主机检测	避免同一个 VSU 域中存在多台主机并存的现象。
VSU 设备外部连接	介绍外部设备与 VSU 设备相连可能出现的状况。
系统管理	用于管理 VSU 系统内部的设备。

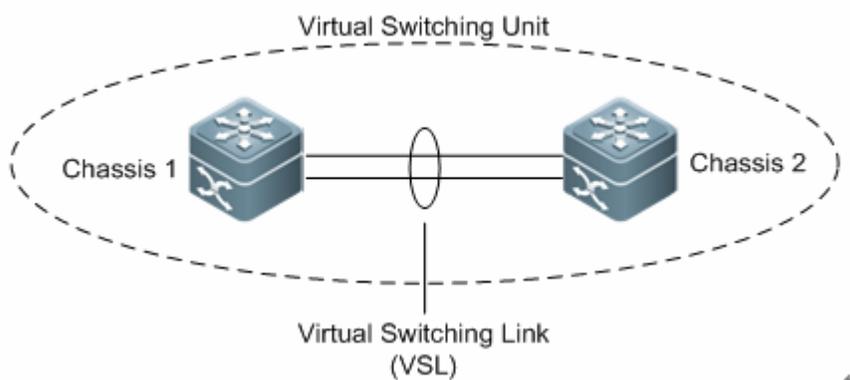
7.3.1 虚拟交换链路

工作原理

虚拟交换链路

由于 VSU 系统的多台设备作为一个网络实体，因此它们之间需要共享控制信息和部分数据流。虚拟交换链路(Virtual Switching Link，简称 VSL；为方便叙述，下文中出现的“虚拟交换链路”均以“VSL”表示)是 VSU 系统的设备间传输控制信息和数据流的特殊链路，目前支持在两台设备间通过万兆接口间建立虚拟交换链路(VSL)。虚拟交换链路在 VSU 系统内的位置如下图所示：

图 7-7 虚拟交换链路



VSL 以聚合端口组的形式存在，由 VSL 传输的数据流根据流量平衡算法在聚合端口的各个成员之间就进行负载均衡。

↘ VSL 链路流量

VSL 链路在设备间传输的控制流分为以下几种情况：

- 成员设备接收到的协议报文，需要通过 VSL 链路转发到全局主设备进行处理。
- 经过全局主设备处理的协议报文，需要通过 VSL 链路转发到其他成员设备的接口，由该接口发送该协议报文到对端设备。

VSL 链路在设备间传输的数据流分为以下几种情况：

- VLAN 内泛洪的数据流。
- 需要跨设备转发的数据流，需要通过 VSL 链路传输。

另外 VSL 链路上也传输 VSU 系统内部的管理类报文，例如热备份交换的协议信息，主机向其他成员设备下发配置信息的报文等等。

i 对于镜像(SPAN)功能，VSL 链路关联的接口既不能作为 SPAN 的源口，也不能作为 SPAN 的目的口。

↘ VSL 链路故障

如果 VSL 聚合端口组的某一成员链路发生故障，VSU 将自动调整 VSL 聚合端口的配置，使得流量不再从故障的成员链接传输。

如果 VSL 聚合端口组的所有成员链路都断开，VSU 拓扑将会发生变化。如果原先是环形拓扑，那么将会发生“环转线”，具体情况请“参考拓扑”变化章节的拓扑环线互转部分。

↘ VSL 口错帧检测

当 VSL 口上出现大量的连续的错帧的时候，需要禁用该端口，切换到其他的 VSL 口中。采用下面的检测方式：

VSL 口上会有错帧，需要进行错帧校正。默认每 5 秒检查一次 vsl 口，如果和上次比较，错帧个数大于 num 则认为是一次错帧，连续 times 次的话，则认为端口异常。在存在多条 vsl 链路的时候，如果发生错帧，vsl 链路会切换。最后一条 vsl 链路，为了防止拓扑分裂，链路不进行切换。

不同用户场景对 num 和 times 要求不一样，默认值 num 为 3，times 为 10 次。用户场景要求严格的这两个值取小值，如果比较宽松需要取大值。

7.3.2 拓扑

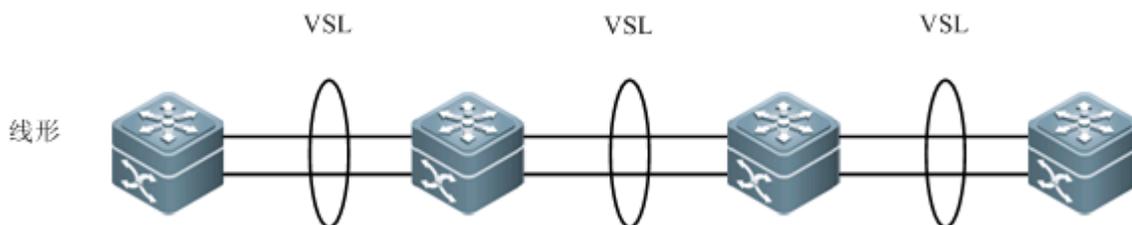
VSU 系统支持线形和环形两种拓扑结构。设备间通过 VSL 链路相连，形成一条线，所以称为线形拓扑。

工作原理

↘ 拓扑结构

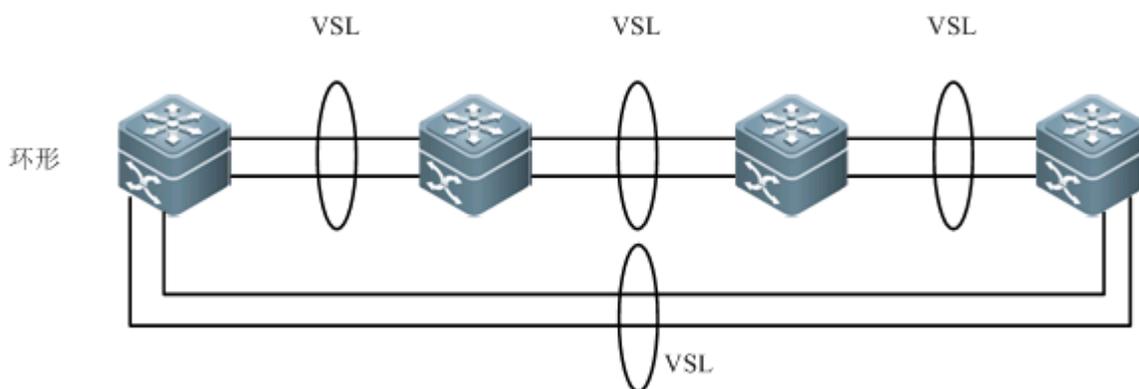
线形拓扑连接简单，使用较少的端口和线缆，但设备间只有一条通信链路，所以 VSL 链路的可靠性较低。

图 7-8 线形拓扑



除了线形拓扑外，如图所示，设备还可以组成环形拓扑，这样设备间的两条通信链路可以相互备份，形成链路冗余，提高 VSU 系统的可靠性。

图 7-9 环形拓扑



- ❗ 用户在选择 VSU 系统的拓扑时，应尽量选择环形拓扑，这样能保证任何单台设备失效、或是任何单条 VSL 链路失效都不会影响整个 VSU 系统的正常运行。
- ❗ 除了选择环形拓扑组网，建议每个 VSL-AP 中配置多根 VSL 链路，以提高单个 VSL-AP 的可靠性。建议最少配置两根链路，最大可以配置 4 根链路。合理的配置是 2 根以上 vsl 链路，并且是跨线卡。

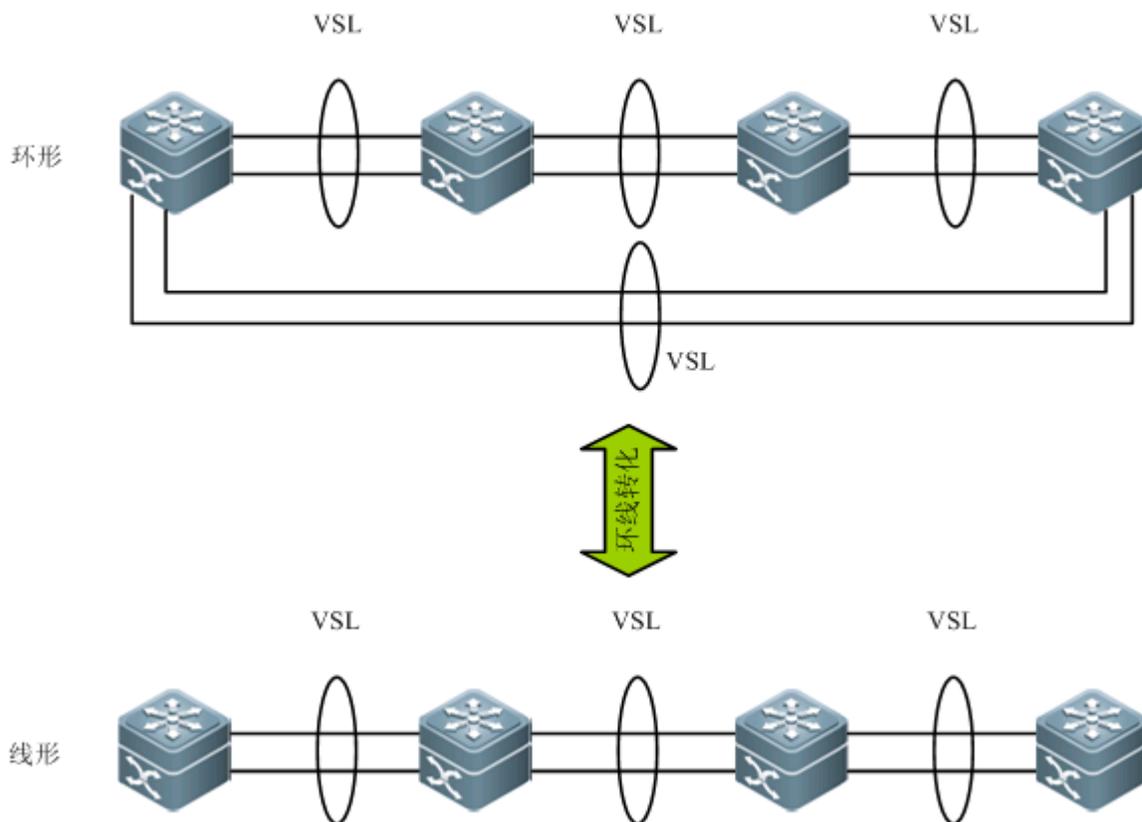
拓扑收敛

在 VSU 系统建立之前，成员设备间需要通过拓扑发现协议来发现邻居，最终确定 VSU 系统中有哪些设备，从而确定管理域的范围。然后选举出一台全局主设备来管理整个 VSU 系统，接着再选举出一台全局从设备作为主设备的备份。到此，整个 VSU 系统的拓扑已经收敛。由于不同的设备的启动时间有所不同，所以拓扑的首次收敛时间也有所不同。

拓扑环线互转

对于环形拓扑，当其中一条 VSL-AP 链路断开时，拓扑将由环形转成线形。这时整个 VSU 系统仍然能够正常工作，不会造成网络的中断。但为了避免其他的 VSL-AP 链路失效、或节点失效，此时应该要及时去排查 VSL 链路故障，将 VSL 链路恢复。VSL-AP 链路恢复后，拓扑将由线形再转回到环形。

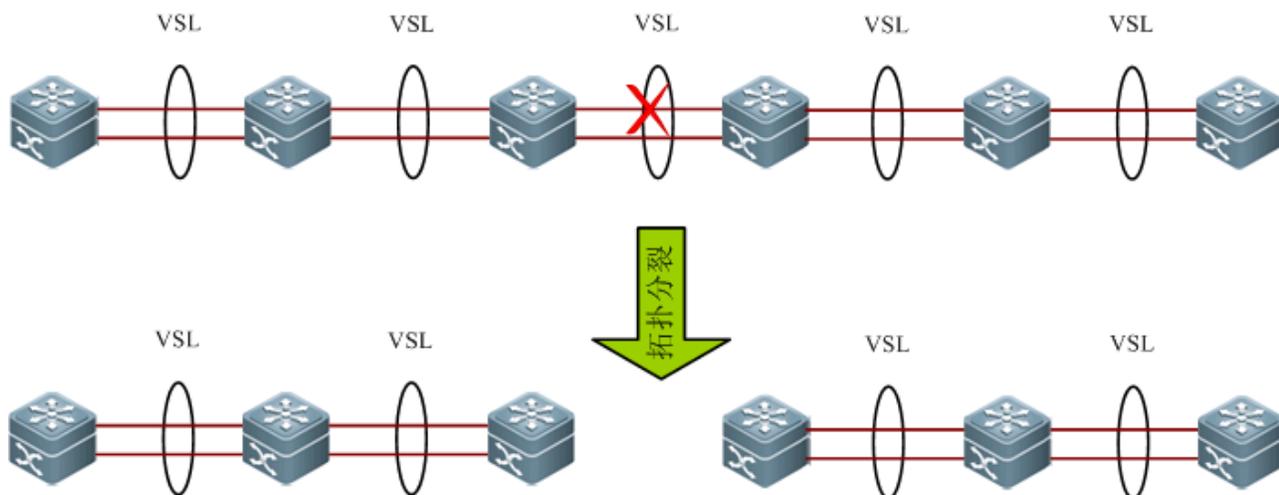
图 7-10 环转线、线转环



拓扑分裂

对于线形拓扑，如果 VSL-AP 链路断开时，拓扑将会发生分裂，如下图所示，一个 VSU 组分裂成两个 VSU 组。这种情况下，可能会导致网络中出现两台配置完全相同的设备，从而令网络无法正常工作。这种情况下需要通过部署多主机检测功能（详见 1.1.4.6 节多主机检测）来解决拓扑分裂问题。

图 7-11 拓扑分裂



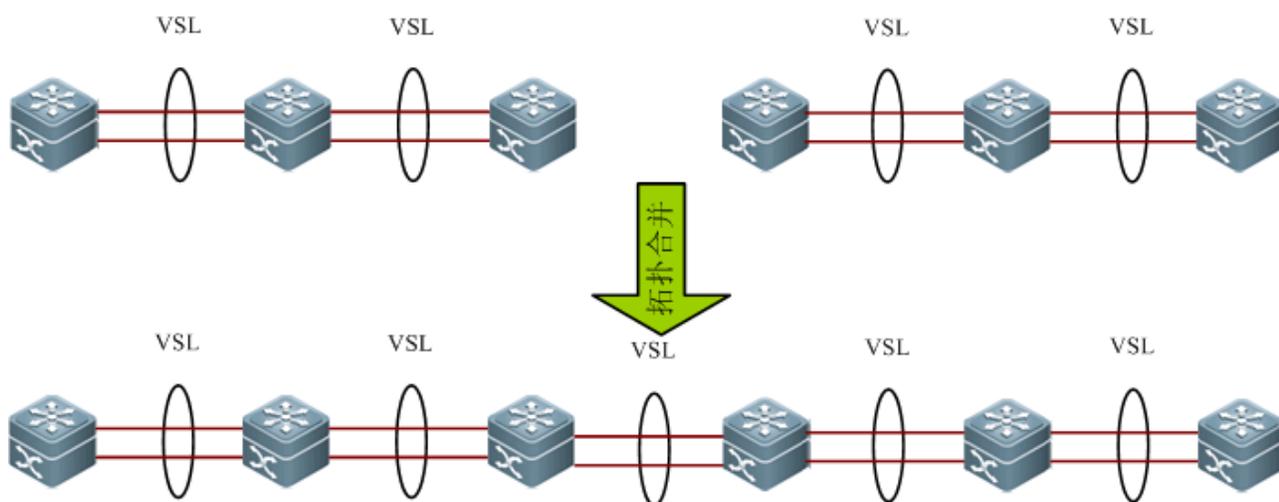
拓扑合并

domain 一致的两个 VSU 组通过 VSL-AP 链路连接，将会发生拓扑合并。在拓扑合并过程中，会重启其中一个 VSU 组，然后热加入另一个 VSU 组。

拓扑合并的原则是：最大限度的降低拓扑合并时对业务所带来的影响。其合并规则如下（从第一条开始判断，如果本条无法选出最优拓扑，继续判断下一条）：

- 用户配置为最高条件，按一堆 VSU 某台设备最高优先级高的那一堆 VSU 保留。
- 上述不能判断，swid 小(以两个全局主为准)的胜出。
- 上述不能判断，以 mac 地址小的保留(以两个全局主为准)。

图 7-12 拓扑合并



i 当两个 VSU 组进行拓扑合并时，需要进行竞选，竞选失败的一方将逐一自动重启并热加入到另一个 VSU 组。

7.3.3 多主机检测

工作原理

当 VSL 断开时，从设备切换成主设备，如果原来的主设备还在运行，那么两台设备都是主角色，由于配置完全相同，在局域网中会引起 IP 地址冲突等一系列问题。在这种情况下，VSU 系统必须检测双主机，并且采取恢复措施。VSU 支持使用两种方式进行多主机检测：

- 基于 BFD 检测
- 基于聚合口检测

多主机检测规则

- 1、一个 VSU 组中优先级高的胜出。
- 2、上述不能判断时，一个 VSU 组中物理设备台数多的胜出。

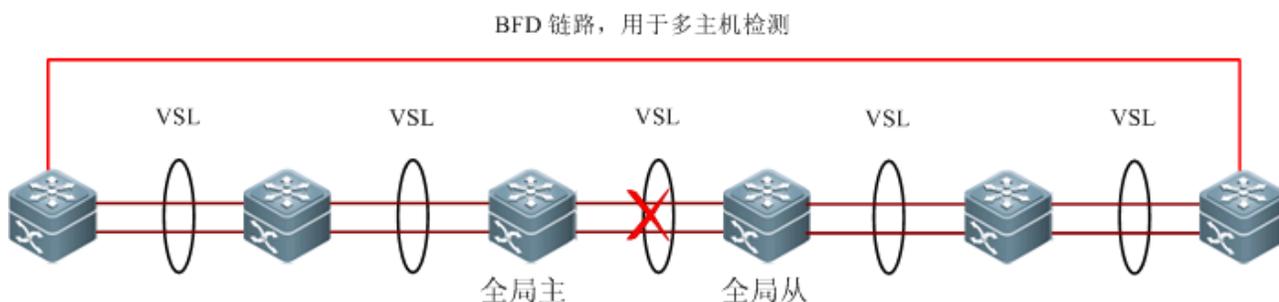
- 3、上述不能判断时，一个 VSU 组中健康度高的胜出。（健康度定义：本拓扑内 UP 的物理端口带宽总和，管理口和 VSL 口除外。）
- 4、上述不能判断时，以两个全局主中 swid 小的胜出。
- 5、上述不能判断时，两个全局主以 mac 地址小的保留。
- 6、上述不能判断时，两个全局主以起机时间大的保留。

! 如果未配置双主机检测，在拓扑分裂后会造成网络中断。

基于 BFD 检测

VSU 支持使用 BFD(Bidirectional Forwarding Detection)检测多主机情况。其拓扑连接如图所示。两个边缘设备增加一条链路，专门用于多主机检测。当全局主和全局从之间的 VSL 链路断开，此时会产生两个主机，如果配置了 BFD 双主机检测功能，则两个主机之间通过 BFD 链路互相发送 BFD 双主机检测报文，从而检测到当前有相同的两个主机存在，最后通过一定的规则（同 1.1.4.4 拓扑合并规则）将其中一个主机所在的 VSU 系统关闭，使其进入 recovery 状态，避免网络异常。

图 7-13 基于 BFD 的多主机检测



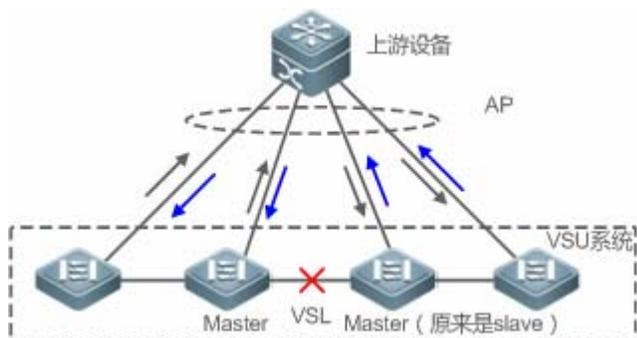
! 只有一对 BFD 检测链路时建议将检测链路部署在拓扑的两端。

! BFD 检测采用扩展 BFD，不能通过现有 BFD 的配置与显示命令配置双主机检测口。

基于聚合口检测

VSU 还支持使用聚合口检测双主机的机制，其连接拓扑如下图所示。在 VSU 系统和上游设备上，都需要支持聚合口多主机检测功能，当发生 VSL 端口断开后，产生两个主机，两个主机向聚合口的每个成员口发送检查报文，检测报文通过上游设备进行中转，到另一个主机。如下图所示，聚合口共有四个成员口，每个成员口连接在 VSU 系统的四个不同设备上，当发生分裂时，四个成员口都会发送和接受检测报文，从而检测到当前有相同的两个主机存在，最后通过一定的规则（同 1.1.4.4 拓扑合并规则）将其中一个主机所在的 VSU 系统关闭，使其进入 recovery 状态，避免网络异常。

图 7-14 聚合口的上下游方式多主机检测



✓ 以上拓扑中，上游设备必须为锐捷设备，该设备需要支持检测报文的转发。

7.3.4 VSU流量转发

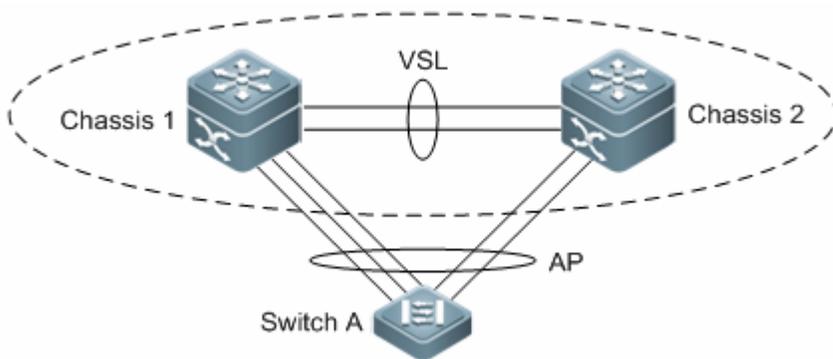
工作原理

跨设备聚合端口组

AP 把多个物理链接捆绑在一起形成一个逻辑链接。VSU 系统支持跨成员设备的 AP。

如下图所示，两台设备组成 VSU 组，外部的接入设备 Switch A 以 AP 的形式链接到 VSU，对于 Switch A 来说，图中的 AP 连接与普通的聚合端口组没有区别。

图 7-15 跨设备聚合端口



故障处理

建议配置跨设备 AP 时，外围设备与 VSU 的每台设备之间均有物理链接。一方面，这可以保留 VSL 链路的带宽(跨机箱 AP 流量优先选择同一机箱的 AP 成员作为出口，避免不必要的流量通过 VSL 链路传输)；另一方面，可以提高网络的可靠性(如果某一机箱发生故障，属于正常设备的成员接口还可以正常工作)。

以下描述跨设备 AP 可能的失败情形及导致的影响：

- 单条链路发生故障

如果跨设备 AP 的单条链路发生故障而其它链路仍然正常工作，则跨设备 AP 在剩余的正常链路之间重新分配流量。

- 全局主设备上的所有跨设备 AP 成员口链路发生故障

如果全局主设备上的所有跨设备 AP 成员口链路发生故障，则只有其它成员设备的成员口继续工作。由该 AP 进入 VSU 系统的数据流，如果数据流的转发出口在全局主设备上，则系统通过 VSL 链路转发到全局主设备对应的出口。

由于控制面协议仍然运行在全局主设备，所以进入 VSU 系统的协议报文通过 VSL 链路转发到全局主设备进行协议运算。

- 与其他成员设备的所有链路发生故障

如果跨设备 AP 与单台成员设备 A 的所有链路发生故障，则只有其它成员设备的成员口继续工作。由该 AP 进入 VSU 系统的数据流，如果数据流的转发出口在成员设备 A 上，则系统通过 VSL 链路转发到成员设备 A 对应的出口。

- 所有链路发生故障

如果跨设备 AP 的所有链路发生故障，与普通 AP 的处理相同，接口的状态变为 Link-Down。

- 全局主设备整机故障

如果主设备整机发生故障，将导致热备份切换，原来的从设备切换为主设备。同时，其它成员设备上的成员口继续工作。通过该 AP 与 VSU 相连的对端设备将检测到链路故障，调整流量平衡算法，将数据流分配到正常的链路。

- 其它成员设备整机故障

如果成员设备整机发生故障，连接在该成员设备的 AP 成员链路将断开，但其他成员链路照常工作。通过该 AP 与 VSU 相连的对端设备将检测到链路故障，调整流量平衡算法，将数据流转发路径分配到正常的链路。

▾ 流量均衡

在 VSU 系统中，流量可能有多个出口。AP 和 ECMP 有各自的流量均衡算法，比如同目的 mac 或者源 mac 等方式，具体可以参见 AP 和 ECMP 的配置手册。在本配置手册中，可以配置本地优先转发，本设备收到的报文优先在本设备转发，这样报文可以不通过 VSL 链路转发到其他设备中。

7.3.5 系统管理

工作原理

▾ 控制台访问

VSU 系统主设备的控制台同时管理系统内的多台设备。从设备、候选设备的控制台不支持命令行输入。但用户可以在主机上对指定成员设备进行 VSU 相关的配置，也可以通过从机的串口登录到主机的控制台。可以利用 session 重定向到某个设备的主管理板。

▾ 线卡命名

对于机箱式设备，在 VSU 模式下，线卡的编号命名中加入了设备编号(Switch ID)，即线卡的编号由一维变二维，如线卡 1/1，表示编号为 1 的成员设备上的 1 槽位的线卡。

▾ 接口命名

VSU 工作模式下，由于同一个插槽号可能分别出现在多台设备内，所以接口的命名方式中加入了设备编号(Switch ID)。

例如：interface gigabitEthernet 1/1/1 表示 ID 为 1 的设备插槽 0 上的千兆端口 1；interface gigabitEthernet 2/1/2 表示 ID 为 2 的设备插槽 0 的千兆端口 2。

访问文件系统

VSU 工作模式下，可以从主设备上访问其他成员设备上的文件系统。具体方式和访问本地文件系统相同。唯一不同的是使用不同的 URL 前缀。

系统升级

VSU 系统通常情况下要求成员设备主程序版本号一致，然而成员设备众多，按照单机模式逐一升级，不仅费时费力，而且容易出错。锐捷交换机提供了完善的系统升级方案，使您能够使用如下两种方法轻松完成系统升级。

- VSU 系统建立时：系统会自动匹配所有成员设备的主程序版本号，当发现主程序版本不一致时，其会选择主设备上的主程序同步到所有成员设备。
- VSU 系统建立后：可以通过 TFTP 下载的文件将自动的同步到所有成员设备。

SYSLOG

VSU 系统的所有成员设备都可以打印 SYSLOG。主机产生的 SYSLOG 直接在主机控制台上打印，且格式和单机情况下是完全一样的；其它成员设备的 SYSLOG 也在主机控制台上打印，但消息格式与单机不同，相比之下增加了设备编号信息。

例如：单机产生的 SYSLOG 信息是：“%VSU-5-DTM_TOPO_CVG: Node discovery done. Topology converged.” 那么由编号为 3 的成员设备产生的 SYSLOG 信息应该就是：“%VSU-5-DTM_TOPO_CVG:(3) Node discovery done. Topology converged.”

7.4 产品说明



每台机箱至少指定一个万兆口或 40G 端口作为 VSL 链路的成员端口。成员端口数量没有限制。万兆口和 40G 端口可以同时作为 VSL 链路的成员端口。

两台机箱的 VSL 成员端口之间的连接方式：

- 如果是万兆接口之间连接，则通过 XFP（或 SFP+）模块+光纤线缆的方式进行连接。
- 如果是 40G 接口之间连接，则可以通过 QXFP 模块+光纤或铜缆的方式进行连接。
- 万兆口和 40G 端口之间不能进行连接。

S7508 交换机最多可配置 32 个 VSL 成员端口：

- 配置 VSL 口的时候，不需要配置 VSL AP。VSL AP 自动协商
- VSL 心跳检测不允许配置。
- 40G 一分四的端口不能作为 vsl 口



S7508 交换机的线卡，任意万兆以上光口，都可以作为 VSL 口。

产品名称	堆叠模块	普通端口	线性拓扑	环形拓扑	成员设备最大数量
S7508	不支持	支持	支持	不支持	2



在 S7508 交换机，为了防止在多条 BFD 检测线时候，用户连接错误，BFD 检测口检测配置，不需要配置成一对，只需要单个配置。

以上拓扑中，上游设备必须为锐捷设备，该设备需要支持 MAD 报文的转发。目前 S7508 设备支持 MAD 报文转发功能。

7.5 配置详解

配置项	配置建议 & 相关命令		
单机模式下配置VSU参数	⚠ 必选。用于配置单机模式下的 VSU 参数。		
	switch virtual domain	配置域 ID	
	switch	配置设备在虚拟设备中的编号	
	switch priority	配置设备的优先级	
	vsl-port	进入 VSL 端口配置模式	
	port-member interface	把普通口配置到 VSL 端口池中	
	switch convert mode virtual	单机模式切换到 VSU 模式	
	⚠ 可选。用于配置 VSU 模式下的设备属性。		
	switch description	配置设备的别名	
switch crc	错帧配置		
VSU模式下配置VSU参数	⚠ 可选。用于配置 VSU 模式下的设备属性。		
	配置VSU属性	switch domain	更改机箱域 ID
		switch renumber	更改设备编号
		switch description	配置设备别名
		switch crc	错帧配置
	配置VSL链路	⚠ 可选。用于配置虚拟交换链路。	
		vsl-port	进入 VSL-PORT 模式
		port-member interface	配置 VSL-AP 成员口
	配置双主机检测	⚠ 必选。用于配置双主机检测功能。	
		dual-active detection	配置双主机检测
		dual-active bfd interface	配置 BFD 检测接口
		dual-active interface	将聚合口配置为双主机检测口
		dual-active exclude interface	配置例外端口
	配置流量平衡	⚠ 可选。用于配置 VSU 模式下的流量平衡功能。	
		switch virtual aggregateport-lbf enable	配置 AP 本地转发优先模式

		switch virtual ecmp-lff enable	配置 ECMP 本地转发优先模式
	配置从VSU模式切换到单机模式	⚠️ 可选。用于将设备从 VSU 模式切换到单机模式。	
		switch convert mode standalone	从 VSU 模式切换到单机模式
监控维护	清除各类信息	⚠️ 可选	
		remove configuration switch	清除指定设备的配置，并自动进行重启。
	查看运行情况	⚠️ 可选	
		show switch virtual	显示当前运行的 VSU 信息，拓扑形状，或当前配置的 VSU 参数
		show switch virtual dual-active	查看当前双主机配置信息
		show switch virtual link	VSU 模式下查看当前的 VSL-AP 运行信息
		session	重定向到主机或任意一台设备的控制台

7.5.1 单机模式下配置VSU参数

配置效果

将设备在单机模式下启动，配置 VSU 相关的参数。以用于组建 VSU 系统。

注意事项

-

配置方法

配置方法配置 VSU 属性

- 交换设备缺省以单机模式启动，用户需要构建 VSU 系统的两台机箱上配置相同的域 ID(domain ID)，虚拟设备号取值在局域网内域 ID 必须是唯一的。用户还需要配置每台机箱在虚拟设备中的编号。
- 首先使用命令 **switch virtual domain domain_id**，配置域 ID，该命令必选；
- 使用 **switch switch_id** 命令配置设备在虚拟设备中的编号，该命令必选；设备在 VSU 系统内编号越大，则在相同的设备优先级情况下，编号越小的优先选为全局主设备。
- 使用 **switch switch_id priority priority_num** 命令配置设备的优先级，该命令必选。
- 取值范围为 1 到 255，数值越大优先级越高。
- 使用 **switch switch_id description switch1** 命令配置设备的别名，该命令可选。默认名字为 Ruijie，为便于网络环境中设备的区分，希望标示设备别名的可选择此配置项。
- 最大 32 个字符。

- 【命令格式】 **switch virtual domain** *number*
- 【参数说明】 *number* : VSU 的虚拟域编号。
- 【缺省配置】 缺省域编号为 100。
- 【命令模式】 config-vs-domain 配置模式
- 【使用指导】 域编号相同的两台设备才能组合成一台虚拟设备，域编号在局域网内必须唯一。
-
- 【命令格式】 **switch** *switch_id*
- 【参数说明】 *switch_id* : 设备在 VSU 内的编号，1-2。
- 【缺省配置】 缺省设备编号是 1。
- 【配置模式】 config-vs-domain 配置模式
- 【使用指导】 设备编号用来在虚拟设备中标识每个成员，在 VSU 模式下，接口名称的格式从 “slot/port” 转换为 “switch/slot/port”，其中 “switch” 就是接口所属交换机的编号。
在选举主设备的过程中，如果两台设备都已经是主设备，或者都是刚启动还没有确定角色，并且两台交换机的优先级相同，那么编号小的设备成为主设备。
该命令只能在单机模式下修改交换机编号，VSU 模式下需要通过 **switch** *switch_id* **renumber** *new_switch_id* 修改交换机编号。无论是单机模式，还是 VSU 模式，修改的编号需要重新启动才能生效。
-
- 【命令格式】 **switch** *switch_id* **priority** *priority_num*
- 【参数说明】 *switch_id* : 需要配置优先级的交换机编号。
priority_num : 对应交换机的优先级，取值范围是 1 到 255。
- 【缺省配置】 *priority_num* : 缺省的优先级是 100。
- 【配置模式】 config-vs-domain 配置模式
- 【使用指导】 优先级的数值越大，表示优先级越高。在选举主设备的过程中，优先级高的设备成为主设备。
该命令在单机模式和 VSU 模式都可以使用。修改的优先级必须重启以后才会生效。
该命令不会修改 *switch_id*。单机模式下如果配置了 *switch_id* 为 1，再执行 **switch** 2 **priority** 200，则命令不会生效，除非先将设备的 *switch_id* 修改为 2，再执行 **switch** 2 **priority** 200 才会生效。VSU 模式下，*switch_id* 表示当前运行的交换机编号，如果当前不存在该编号，则配置也不生效。
-
- 【命令格式】 **switch** *switch_id* **description** *dev-name*
- 【参数说明】 *switch_id* : 需要配置别名的交换机编号。
- 【缺省配置】 -
- 【配置模式】 config-vs-domain 配置模式
- 【使用指导】 配置设备的别名，最大为 32 个字符（可选）。
该命令在单机模式和 VSU 模式都可以使用，VSU 模式下配置立即生效。

 配置优先级与别名的命令只会修改优先级，不会修改交换机编号。所以在配置时必须正确输入当前设备的编号。例如，当前已经配置交换机编号为 1，如果输入 **switch** 2 **priority** 100，则优先级配置不生效。

📌 配置 VSL 链路

- 为了组成 VSU 系统，还需要配置一些端口作为 VSL 成员端口。
- 使用 **vsl-port** 命令进入 vsl 端口配置模式，该命令必选。
- 使用 **port-member interface interface-name [copper | fiber]** 命令添加 VSL 端口，该命令必选。
- 当设备进入到 VSL-PORT 的配置模式时，可以配置或删除 VSL 口。

【命令格式】 **vsl-port**

【参数说明】 -

【缺省配置】 -

【配置模式】 config 配置模式

【使用指导】 该命令在单机模式和 VSU 模式都可以使用。

【命令格式】 **port-member interface interface-name [copper | fiber]**

【参数说明】 *interface-name* : 二维接口名，如 Tengigabitethernet 1/1, Tengigabitethernet 1/3。

copper : 电口属性。

fiber : 光口属性。

【缺省配置】 -

【配置模式】 config-vsl-port 配置模式

【使用指导】 添加 VSL-AP 链路的成员端口。*interface-name* 为单机模式下的二维端口名称，可以为万兆口，也可以为千兆口（千兆口可以为光电复用口，如果不指定介质类型，则默认为千兆电口）。对于光电复用口，必须指定其光电属性。箱式设备 VSL 口必须是万兆口。

该命令可以在 VSU 模式下，也可以在单机模式下。命令配置后需要保存配置，并重启 VSL 成员端口所在设备才能生效。

 单机模式下，VSL 端口的配置不能立即生效，需要转化为 VSU 模式重新启动后才能生效。

错帧配置

- 使用 **switch crc** 配置错帧，该命令可选。选择此命令可以修改错帧的默认检查方式。
- VSL 口上会有错帧，需要进行错帧校正。默认每 5 秒检查一次 vsl 口，如果和上次比较，错帧个数大于 3 则认为是一次错帧，连续 10 次的话，则认为端口异常。在存在多条 vsl 链路的时候，如果发生错帧，vsl 链路会切换。最后一条 vsl 链路，为了防止拓扑分裂，链路不进行切换。

【命令格式】 **switch crc errors error_num times time_num**

【参数说明】 *error_num* : 用于配置两次检查错帧递增个数（当大于这个数认为是一次错帧）

time_num : 连续多少次后，采取的动作（动作为提示或关闭端口）

【缺省配置】 errors 缺省值为 3；times 缺省值为 10

【配置模式】 config-vs-domain 配置模式

【使用指导】 默认每 5 秒检查一次 vsl 口，如果和上次比较，错帧个数大于 3 认为一次错帧，连续 10 次，可以认为端口异常。对端口异常的处理是，默认是 log 提示，可以配置成关闭端口处理，如果关闭端口，需要插拔恢复。

 不同的产品对错帧检查要求不一样，对 vsl 口的处理也不一样。11.0 版本，错帧作为可以配置。

📌 单机模式切换到 VSU 模式

- 使用 **switch convert mode virtual** 命令，将设备从单机模式切换到 VSU 模式。
- 单机模式下，执行以上命令后，软件自动进行如下动作：
 - 将单机模式下的各个 VSD 的全局配置文件 “config.text” 备份为 “vsd.standalone.text.vsd 序号”；
 - 清除各个 VSD 的全局配置文件 “config.text” 的内容；
 - 把 VSU 相关的配置写到特殊配置文件 “config_vsu.dat” 中。
- 如果交换设备上存在 “vsd.virtual_switch.text vsd 序号” 备份文件，则提示用户是否将备份文件的内容覆盖到对应 VSD 下 “config.text” (“vsd.virtual_switch.text.vsd 序号” 文件是交换设备从 VSU 模式切换到单机模式时对各 vsd 下的 “config.text” 的备份文件)，用户可选择 “yes” 或 “no”。选择 “yes” 使用 “vsd.virtual_switch.text.vsd 序号” 文件，替换对应 vsd 的 “config.text” 文件，如果选择 “no”，清空对应 vsd 的 “config.text” 文件。最后交换设备进行重启，读取 “config_vsu.dat” 中的 VSU 参数，以 VSU 模式进行启动。

【命令格式】 **switch convert mode virtual**

【参数说明】 -

【缺省配置】 缺省时设备处于单机模式。

【配置模式】 特权模式

【使用指导】 将设备从单机模式切换到 VSU 模式。

 如果当前交换设备已经在 VSU 模式，则不允许再次切换到 VSU 模式，即以上命令无效。

检验方法

通过 **show switch virtual config [switch_id]**命令查看单机模式下当前交换设备的 VSU 配置。

【命令格式】 **show switch virtual config [switch_id]**

【参数说明】 *switch_id* : 设备编号，指定这个参数可以只显示特定设备的 VSU 配置信息。

【配置模式】 特权模式

【使用指导】 显示单机或 VSU 模式下的 VSU 配置信息

 由于 VSU 相关的配置是针对单个物理设备的，其配置信息存储在特殊配置文件 config_vsu.dat 中，因此 **show running config** 看不到 VSU 相关的配置信息，只能通过 **show switch virtual config** 来查看当前 VSU 的配置。

 单机模式下，VSU 运行信息全部空，用户敲入 **show switch virtual** 等命令时，则提示当前为单机模式，无 VSU 系统运行信息。

配置举例

📌 单机配置举例

【网络环境】

图 7-16



Switch-1 及 Switch-2 组成 VSU，domain 域为 100，左边机箱配置成机箱号 1，优先级 200，别名 switch-1，上面有端口 1/1、1/2 为 VSL 口。右边机箱配置成机箱号 2，别名 switch-2，优先级 100，上面有端口 1/1、1/2 为 VSL 口。

【配置方法】

1. 在 Switch-1 机箱上配置：
 - 配置 VSU 属性、VSL 口。
 - 将单机模式转换成 VSU 模式。
2. Switch-2 机箱上配置：
 - 配置 VSU 属性、VSL 口。
 - 将单机模式转换成 VSU 模式。

Switch-1

```
Ruijie# configure terminal
Ruijie(config)# switch virtual domain 100
Ruijie(config-vs-domain)#switch 1
Ruijie(config-vs-domain)#switch 1 priority 200
Ruijie(config-vs-domain)#switch 1 description switch-1
Ruijie(config-vs-domain)# switch crc errors 10 times 20
Ruijie(config-vs-domain))#exit
Ruijie(config)#vsl-port
Ruijie(config-vsl-port)#port-member interface Tengigabitethernet 1/1
Ruijie(config-vsl-port)#port-member interface Tengigabitethernet 1/2
Ruijie(config)#exit
Ruijie#switch convert mode virtual
```

Switch-2

```
Ruijie# configure terminal
Ruijie(config)# switch virtual domain 100
Ruijie(config-vs-domain)# switch 2
Ruijie(config-vs-domain)# switch 2 priority 200
Ruijie(config-vs-domain)# switch 2 description switch-2
Ruijie(config-vs-domain)# switch crc errors 10 times 20
Ruijie(config-vs-domain))#exit
Ruijie(config)#vsl-port
```

```
Ruijie(config-vsl-port)#port-member interface Tengigabitethernet 1/1
Ruijie(config-vsl-port)#port-member interface Tengigabitethernet 1/2
Ruijie(config-vsl-port)#exit
Ruijie#switch convert mode virtual
```

【检验方法】 ● 使用 **show switch virtual config** 命令查看 Switch-1、Switch-2 的 VSU 属性。

Switch-1

```
Ruijie#show switch virtual config
switch_id: 1 (mac: 0x1201aeda0M)
!
switch virtual domain 100
!
switch 1
switch 1 priority 100
!
switch convert mode virtual
!
port-member interface Tengigabitethernet 1/1
!
port-member interface Tengigabitethernet 1/2
!
switch crc errors 10 times 20
!
```

Switch-2

```
Ruijie#show switch virtual config
switch_id: 2 (mac: 0x1201aeda0E)
!
switch virtual domain 100
!
switch 2
switch 2 priority 100
!
switch convert mode virtual
```

```
!  
port-member interface Tengigabitethernet 1/1  
  
!  
port-member interface Tengigabitethernet 1/2  
  
!  
switch crc errors 10 times 20  
  
!
```

常见配置错误

- ✓ 在箱式设备中，VSL 口必须是万兆口以上的端口。

7.5.2 配置VSU模式下的功能

7.5.2.1 配置VSU属性

配置效果

设备组成 VSU 或 VSU 系统运行过程中，如果需要修改一些参数，用户可以登录到 VSU 系统的主机控制台上进行修改，从机控制台禁止进入全局配置模式。

注意事项

- 除 **switch switch_id description switch1** 命令立即生效外，其他配置命令只有在交换设备重启后才能生效。

配置方法

↘ 进入 domain 配置模式

- 可选配置。
- VSU 模式下使用该命令进入 domain 配置模式。只有相同域编号的交换机组成 VSU。VSU 模式下，只有进入 domain 配置模式，才能修改或配置域编号、设备优先级、交换机编号。

【命令格式】 **switch virtual domain domain_id**

【参数说明】 *domain_id* : VSU 的虚拟域编号。

【缺省配置】 缺省域编号为 100。

【命令模式】 config-vs-domain 配置模式

【使用指导】 域编号相同的两台设备才能组合成一台虚拟设备，域编号在局域网内必须唯一。

↘ 更改机箱域 ID

- 可选配置。
- 如果需要修改某设备的 *domain_id*，可在 VSU 系统的主机控制台上执行此配置项。

【命令格式】 **switch** *switch_id* **domain** *new_domain_id*

【参数说明】 *switch_id*：VSU 模式下当前运行的设备编号，取值范围为 1-2。

new_domain_id：修改后的 domain id，范围为 1-255。

【缺省配置】 *domain id* 缺省值为 100。

【命令模式】 config-vs-domain 配置模式

【使用指导】 该命令只能在 VSU 模式下使用，不能在单机模式下使用，且重启后才能生效。

↘ 更改设备编号

- 可选配置
- 如果需要修改某设备的 *switch_id*，可在 VSU 系统的主机控制台上执行此配置项。

【命令格式】 **switch** *switch_id* **renumber** *new_switch_id*

【参数说明】 *switch_id*：VSU 取值范围为箱式设备为 1-2。

new_switch_id：修改后的设备编号。

【缺省配置】 -

【命令模式】 config-vs-domain 配置模式

【使用指导】 该命令只能在 VSU 模式下使用，不能在单机模式下使用，且重启后才能生效。

↘ 更改设备优先级

- 可选配置。
- 如果需要修改某设备的优先级，可在 VSU 系统的主机控制台上执行此配置项。
- 优先级的数值越大，表示优先级越高。在选举主设备的过程中，优先级高的设备成为主设备。

【命令格式】 **switch** *switch_id* **priority** *priority_num*

【参数说明】 *switch_id*：需要配置优先级的交换机编号。

priority_num：对应交换机的优先级，取值范围为 1-255。

【缺省配置】 缺省的优先级是 100。

【命令模式】 config-vs-domain 配置模式

【使用指导】 优先级的数值越大，表示优先级越高。在选举主设备的过程中，优先级高的设备成为主设备。

该命令在单机模式和 VSU 模式都可以使用。修改的优先级必须重启以后才会生效。

该命令不会修改 *switch_id*。单机模式下如果配置了 *switch_id* 为 1，再执行 `switch 2 priority 200`，则命令不会生效，除非先将 *switch_id* 修改为 2，再执行 `switch 2 priority 200` 才会生效。VSU 模式下，*switch_id* 表示当前运行的交换机编号，如果当前不存在该编号，则配置也不生效。

↘ 配置设备别名

- 可选配置。
- 如果需要配置某设备的别名，可在 VSU 系统的主机控制台上执行此配置项。
- 使用 **switch switch_id description switch1** 命令配置设备的别名，最大为 32 个字符。

【命令格式】 **switch switch_id description dev-name**

【参数说明】 *switch_id* : 需要配置优先级的交换机编号。
dev_name : 设备名称描述

【缺省配置】 -

【命令模式】 config-vs-domain 配置模式

【使用指导】 该命令在单机模式和 VSU 模式都可以使用，VSU 模式下配置立即生效。

↘ 错帧配置

- 可选配置。
- 使用 **switch crc errors error_num times time_num** 命令配置错帧触发的条件。

【命令格式】 **switch crc errors error_num times time_num**

【参数说明】 *error_num* : 用于配置两次检查错帧递增个数（当大于这个数认为是一次错帧）
time_num : 连续多少次后，采取的动作（动作为提示或关闭端口）

【缺省配置】 errors 缺省值为 3 ; times 缺省值为 10

【命令模式】 config-vs-domain 配置模式

【缺省级别】 14

【使用指导】 -

↘ 保存配置文件

使用 **exit** 命令退出虚拟设备配置模式，并使用 **write** 命令保存配置到文件 config_vs.dat 中。

检验方法

使用 **show switch virtual [topology | config]** 命令显示当前运行的 VSU 信息，拓扑形状，或当前配置的 VSU 参数。

【命令格式】 **show switch virtual [topology | config]**

【参数说明】 Topology-拓扑信息，config-VSU 配置信息

【命令模式】 特权模式

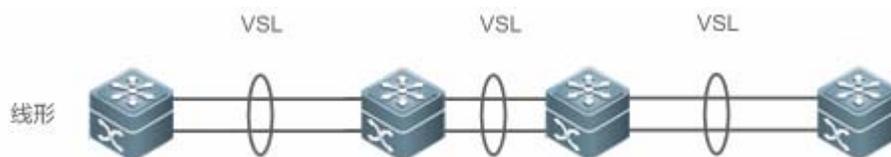
【使用指导】 查看域 ID，以及每台设备的编号、状态和角色。

配置举例

↘ 配置 VSU 属性

【网络环境】

图 7-17



Switch-1 和 Switch-2 组成 VSU，把 Switch-2 的机箱号修改为 3，优先级修改为 150。假设 Switch1 是全局主交换机，在全局主交换机上配置。

【配置方法】

- 修改 Switch-2 的配置

Switch-1

```
Ruijie#config
Ruijie(config)# switch virtual domain 100
Ruijie(config-vs-domain)# switch 2 renumber 3
Ruijie(config-vs-domain)# switch 2 priority 150
Ruijie(config-vs-domain)# switch 2 description switch-3
```

【检验方法】

- 使用命令 **show switch virtual config** 查看。

Switch-1

```
Ruijie#show switch virtual config
switch_id: 1 (mac: 0x1201aeda0M)
!
switch virtual domain 100
!
switch 1
switch 1 priority 100
!
switch convert mode virtual
!
port-member interface Tengigabitethernet 1/1
!
port-member interface Tengigabitethernet 1/2
!

switch_id: 3 (mac: 0x1201aeda0E)
!
switch virtual domain 100
```

```
!  
switch 3  
switch 3 priority 150  
!  
switch convert mode virtual  
!  
port-member interface Tengigabitethernet 1/1  
!  
port-member interface Tengigabitethernet 1/2  
!  
switch 3 description switch-3  
!
```

常见错误

-

7.5.2.2 配置VSL链路

配置效果

设备组成 VSU 或 VSU 系统运行过程中，如果需要在普通口和 VSL 端口之间互相转换，用户可以登录到 VSU 系统的主机控制台上进行修改，从机控制台禁止进入全局配置模式。

注意事项

- 可以通过串口或 telnet 登录到 VSU 系统控制台进行添加或删除 VSL 成员端口的配置。
- 为了防止实际场景连错，VSL AP 采用动态协商。先配置 vs1 口池，协商成功后，加到同一个 AP 中。和同一台设备相连的端口在同一个 AP 中。

配置方法

↘ 进入 VSL-PORT 模式

- 使用 **vs1-port** 命令进入 VSL-PORT 配置模式。该命令可选。
- 当设备进入到 VSL-PORT 的配置模式时，可以配置或删除 VSL 口。

- 【命令格式】 **vsl-port**
- 【参数说明】 -
- 【缺省配置】 -
- 【命令模式】 config 配置模式
- 【使用指导】 该命令在单机模式和 VSU 模式都可以使用

配置 VSL-AP 成员口

- **port-member interface** *interface-name* [**copper** | **fiber**]命令配置 VSL 口。该命令可选。
- 使用 **port-member interface** 命令配置或删除 VSL 口。

- 【命令格式】 **port-member interface** *interface-name* [**copper** | **fiber**]
- 【参数说明】 *interface-name* : 二维接口名, 如 GigabitEthernet 0/1, GigabitEthernet 0/3。
copper : 电口属性。
fiber : 光口属性。
- 【缺省配置】 -
- 【命令模式】 config-vsl-port 配置模式
- 【使用指导】 该命令可以在 VSU 模式下, 也可以在单机模式下。命令配置后需要保存配置, 并重启 VSL 成员端口所在设备才能生效。

i VSU 系统运行过程中, 配置的 VSL 成员链路即刻生效。所有设备上都要配置 vsl 口

i 箱式设备只能万兆以上光口做 vsl 口。

i 箱式设备上模块也必须使用万兆以上的模块。

i 40G 一分四端口不能做成 VSL 口。

! 对于 40G 端口 (无论该端口是否执行了拆分操作), 其成员口 (即 4 个 10G 口) 不允许进行转换为 VSL 成员口的操作。

! 如果端口被用户配置为 NLB 反射口必须将该配置删除, 才能进行转换为 VSL 成员口的操作。

! 为了防止 VSL 成员口退出 VSL 聚合端口的瞬间发生环路, 在执行命令将 VSL 成员口退出 VSL 聚合端口时, 系统自动将该成员口设置为 shutdown 状态。在退出 VSL 聚合端口操作完成以后, 用户可以重新连接链路并通过 no shutdown 命令重新启用该端口。配置 VSL 口时候, 系统会将端口先 shutdown, 如果配置失败, 如果想作为普口继续使用, 可以通过 no shutdown 命令重新启用该端口。添加某个成员端口编号, 必须是三维端口号。比如进入 VSL-PORT 配置模式, 执行 **port-member interface** Tengigabitethernet 1/1/1 命令, 则表示将全局三维端口 1/1/1 配置成 VSL 口。

! 从 VSL 口变为普通口, 如果导致 VSU 拓扑断裂, 不允许删除, 可以先断掉物理口, 再删除 VSL 口。

检验方法

使用 **show switch virtual link [port]**命令查看当前的 VSL-AP 运行信息。

- 【命令格式】 **show switch virtual link [port]**
- 【参数说明】 **port** : 显示 VSL 子接口的状态信息。
- 【命令模式】 特权模式
- 【使用指导】 -

配置举例

配置 VSL 链路

【网络环境】

图 7-18



【配置方法】

- 在 Switch-1 中，增加端口 1/1/3 作为 VSL 口，将 1/1/2 从 vsl 口中移除。

Switch-1

```

Ruijie#config
Ruijie(config)# vsl-port
Ruijie(config-vsl-port)# port-member interface Tengigabitethernet 1/1/3
Ruijie(config-vsl-port)# no port-member interface Tengigabitethernet 1/1/2
  
```

【检验方法】

- 使用 **show switch virtual config** 命令查看 VSL 链路的情况。假设是 Switch-1 是全局主交换机，在全局主交换机上执行命令。

Switch-1

```

Ruijie#show switch virtual config
switch_id: 1 (mac: 0x1201aeda0M)
!
switch virtual domain 100
!
switch 1
switch 1 priority 100
!
switch convert mode virtual
!
port-member interface Tengigabitethernet 1/1
!
port-member interface Tengigabitethernet 1/3
!
  
```

```
switch_id: 3 (mac: 0x1201aeda0E)
!
switch virtual domain 100
!
switch 3
switch 3 priority 150
!
switch convert mode virtual
!
port-member interface Tengigabitethernet 1/1
!
port-member interface Tengigabitethernet 1/2
!
switch 3 description switch-3
!
```

常见错误

-

7.5.2.3 配置双主机检测

配置效果

配置相关的检测机制以防止产生双主机。

注意事项

- 双主机检测只能在 VSU 模式下进行配置，单机模式下不允许配置双主机检测机制。
- 所有双主机检测配置在主从机上立即生效，且属于全局配置，**show running-config** 可以查看。
- BFD 检测的配置信息不能通过 BFD 的显示命令进行显示，只能通过双主机检测显示命令进行显示。

配置方法

配置 BFD 双主机检测

- 基于 BFD 的双主机检测，要求在两台机箱之间建立一条直连链路，链路两端的端口必须是物理路由端口。以下配置在两台机箱上均需配置。
- 首先进入检测接口的接口配置模式，将检测接口配置为路由口。
- 退出接口配置模式后通过命令 **switch virtual domain *domain_id*** 进入 config-vs-domain 配置模式。
- 在模式 config-vs-domain 下，通过命令 **dual-active detection bfd** 打开 BFD 开关。该命令可选，当需要配置 BFD 双主机检测功能时选用此命令。
- 在模式 config-vs-domain 下，通过 **dual-active bfd interface *interface-name*** 配置 BFD 检测接口。该命令可选，当配置 BFD 双主机检测功能时需使用此命令配置 BFD 检测接口。
- 删除 BFD 检测接口，如果没有剩余的 BFD 检测口，会导致 BFD 检测无法使用。

【命令格式】 **switch virtual domain *domain_id***

【参数说明】 *domain_id* : VSU 的虚拟域编号

【缺省配置】 域编号缺省为 100

【命令模式】 config-vs-domain 配置模式

【使用指导】 域编号相同的两台设备才能组合成一台虚拟设备，域编号在局域网内必须唯一。

【命令格式】 **dual-active detection { aggregateport | bfd }**

【参数说明】 **aggregateport** : 指定聚合口探测方式。

Bfd : 定 BFD 探测方式。

【缺省配置】 检测双主设备状态的功能是关闭的。

【命令模式】 config-vs-domain 配置模式

【使用指导】 该命令只能在 VSU 模式下进行配置。

【命令格式】 **dual-active bfd interface *interface-name***

【参数说明】 *interface-name* : 检测接口类型和编号

【缺省配置】 -

【命令模式】 config-vs-domain 配置模式

【使用指导】 BFD 检测接口必须是路由端口，且在不同的设备上。

- ❗ BFD 检测接口必须是直连的物理路由端口，检测端口必须在不同的设备上。
- ❗ 配置的接口类型没有限制，由于双主机检测链路只用于传输 BFD 报文，流量不大，建议使用千兆口或百兆口作为双主机检测端口。
- ❗ 当配置为双主机的三层路由口被转换为二层交换口(在该接口下执行 **switchport** 命令)后，BFD 双主机配置将自动清除。
- ❗ BFD 建议使用在直连方式，只能连接主从两台设备。
- ❗ 当 VSU 系统检测出双主机冲突并让其中一堆 VSU 进入 recovery 状态后，用户应该通过修复 VSL 故障方式来解决，而不能直接复位那堆进入 recovery 状态的 VSU，否则可能会引起网络上出现双主机冲突。

配置聚合口双主机检测

- 要配置基于聚合口检测方式，必须先配置一个 AP 聚合口，然后指定 AP 聚合口为检测口。
- 使用 **port-group ap-num** 命令将物理成员口加入到 AP 聚合口中。
- 进入 config-vs-domain 配置模式后，使用 **dual-active detection aggregateport** 命令打开聚合口方式检测开关。该命令可选。当需要配置聚合口检测功能时选用此命令。
- 使用 **dual-active interface interface-name** 命令将聚合口配置为双主机检测口。该命令可选，配置聚合口检测功能时需使用此命令将聚合口配置为双主机检测口。
- 使用 **dad relay enable** 命令打开上下游设备接口的双主机检测报文中转功能。该命令可选，当配置基于聚合口检测双主机时，需使用此命令转发 dad 报文（双主机检查报文）。
- 关闭聚合口双主机检测功能，会使聚合口双主机检测失效。
- 删除检测口，如果没有剩余的聚合口检测口，会导致聚合口检测无法使用。
- 缺省关闭基于聚合口检测双主机的转发特性。

【命令格式】 **dual-active detection { aggregateport | bfd }**

【参数说明】 **aggregateport**：指定聚合口探测方式。

bfd：定 BFD 探测方式。

【缺省配置】 检测双主设备状态的功能是关闭的。

【命令模式】 config-vs-domain 配置模式

【使用指导】 该命令只能在 VSU 模式下进行配置。

【命令格式】 **dual-active interface interface-name**

【参数说明】 **interface-name**：接口类型和接口编号，必须为 AP 类型的接口。

【缺省配置】 -

【命令模式】 config-vs-domain 配置模式

【使用指导】 基于聚合口的双主机检测口只能配置一个，在设置 AP 口为检测接口前要先创建该接口，后配置的检测口会覆盖前一次配置的检测口。

【命令格式】 **dad relay enable**

【参数说明】 -

【缺省配置】 缺省关闭基于聚合口检测双主机的转发特性。

【命令模式】 接口配置模式

【使用指导】 该命令只能在 AP 接口上使用。

i 建议加入到聚合检测口的物理接口尽量分布在不同设备上。

配置 recovery 模式的例外端口列表

- 当检测到双主机，其中一台主机必须进入 recovery 模式。Recovery 模式下，需要将所有业务口进行关闭。为了某些特殊用途业务口的正常使用（如配置一个端口远程登陆管理设备），用户可以将某些端口配置为在 recovery 模式不关闭的例外端口。
- 进入 config-vs-domain 配置模式后，通过 **dual-active exclude interface** *interface-name* 命令指定在 recovery 模式不关闭的例外端口。该命令可选。

【命令格式】 **dual-active exclude interface** *interface-name*

【参数说明】 *interface-name*：接口类型和接口编号。

【缺省配置】 -

【命令模式】 config-vs-domain 配置模式

【使用指导】 该命令只能在 VSU 模式下进行配置。例外端口必须是路由端口，不能是 VSL 端口。用户可以配置多个例外端口。

 例外端口必须是路由端口，不能是 VSL 端口。

 当例外端口由路由口被转换为交换口（在该接口下执行 switchport 命令）后，该接口关联的例外端口配置将被自动清除。

检验方法

通过命令 **show switch virtual dual-active { aggregateport | bfd | summary }** 查看当前双主机配置信息。

【命令格式】 **show switch virtual dual-active { aggregateport | bfd | summary }**

【参数说明】 **aggregateport**：查看基于聚合口检测信息。

bfd：查看基于 BFD 检测信息。

summary：显示 DAD 概要信息。

【命令模式】 特权模式

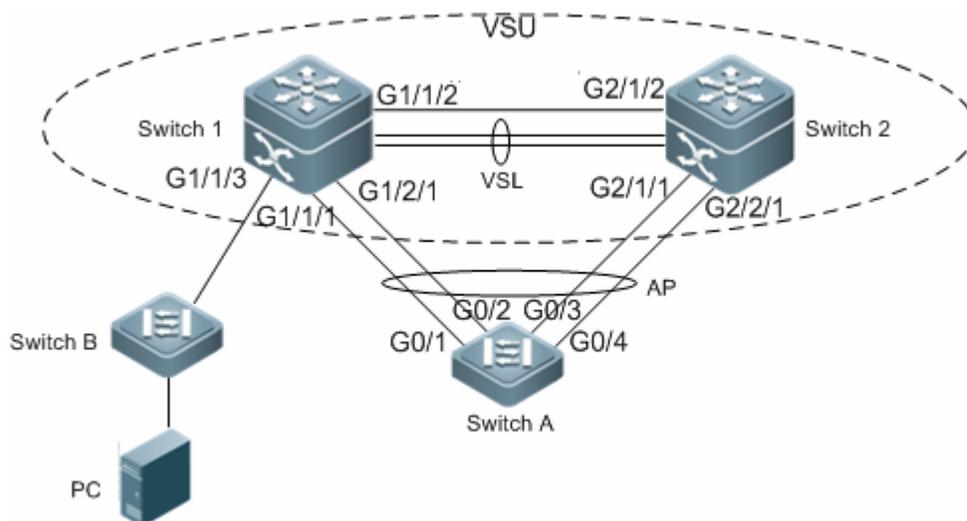
【使用指导】 -

配置举例

采用基于 BFD 来检测双主机

【网络环境】

图 7-19



- Switch 1 和 Switch 2 组成虚拟设备 VSU(domain ID 为 1), Switch 1 的优先级是 200, Switch 2 的优先级是 150。Switch 1 的 Te1/3/1, Te1/3/2 与 Switch 2 的 Te2/3/1, Te2/3/2 分别建立链接, 组成 Switch 1 和 Switch 2 之间的 VSL 链路。Switch A 的端口 G0/1, G0/2, G0/3 和 G0/4 等 4 个端口分别与 Switch 1 的 G1/1/1 和 G1/2/1, 以及 Switch 2 的 G2/1/1 和 G2/2/1 建立连接, 并构成一个包含 4 个成员链路的聚合端口组, 聚合端口组的 ID 是 1。聚合端口组 1 的所有成员均为千兆光口。G1/1/2 和 G2/1/2 均为路由口。
- G1/1/2 和 G2/1/2 是一对 BFD 双主机接口。

【配置方法】

- 将 G1/1/2 和 G2/1/2 口配置为路由口
- 开启 BFD 双主机检测功能
- 配置 G1/1/2 和 G2/1/2 为 BFD 检测接口

i 因为 Switch 1 和 Switch 2 组成虚拟设备 VSU, 所以以上配置可在 Switch 1 及 Switch 2 中任意一台设备上配置。此处以在 Switch 1 上配置为例。

Switch 1

```
Ruijie(config)# interface GigabitEthernet 1/1/2
Ruijie(config-if-GigabitEthernet 1/1/2)# no switchport
Ruijie(config)# interface GigabitEthernet 2/1/2
Ruijie(config-if-GigabitEthernet 2/1/2)# no switchport
Ruijie(config-if)# switch virtual domain 1
Ruijie(c config-vs-domain)# dual-active detection bfd
Ruijie(config-vs-domain)# dual-active bfd interface GigabitEthernet 1/1/2
Ruijie(config-vs-domain)# dual-active bfd interface GigabitEthernet 2/1/2
```

Switch A

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# interface aggregateport 1
```

```
Ruijie(config-if-aggregateport 1)# interface range GigabitEthernet 0/1-4
Ruijie(config-if-aggregateport 1)# port-group 1
Ruijie(config)# interface vlan 1
Ruijie(config-if-vlan 1)#ip address 1.1.1.2 255.255.255.0
Ruijie(config-if-vlan 1)#exit
Ruijie(config)#interface aggregateport 1
Ruijie(config-if-AggregatePort 1)# dad relay enable
Ruijie(config-if-AggregatePort 1)# exit
```

- 【检验方法】
- 查看双主机箱配置状态
 - 查看 BFD 双主机箱检测配置

Switch 1

```
Ruijie# show switch virtual dual-active summary
BFD dual-active detection enabled: No
Aggregateport dual-active detection enabled: Yes
Interfaces excluded from shutdown in recovery mode:
In dual-active recovery mode: NO
Ruijie# show switch virtual dual-active bfd
BFD dual-active detection enabled: Yes
BFD dual-active interface configured:
    GigabitEthernet 1/1/2: UP
    GigabitEthernet 2/1/2: UP
```

常见错误

- 作为 BFD 检测口，必须为路由口。
- BFD 检测和聚合口检测只能激活其中的一种。

7.5.2.4 配置流量平衡

配置效果

在 VSU 系统中，如果出口分布在多台设备中，通过该配置，优先在本台设备上转发。

注意事项

默认都是本地优先转发

配置方法

配置 AP 本地转发优先模式

- 进入 config-vs-domain 配置模式后，使用 **switch virtual aggregateport-lff enable** 命令打开 AP 本地转发优先 LFF(Local Forward First)。该命令可选。
- AP 成员口可以分布在 VSU 系统的两个机箱上。用户可以根据实际流量情况，配置 AP 的出口流量是否优先从本地成员口进行转发。
- 该功能若关闭，则流量转发根据 AP 配置规则转发流量。具体配置见 ap 配置。

【命令格式】 **switch virtual aggregateport-lff enable**

【参数说明】 -

【缺省配置】 缺省时该功能是打开的。

【命令模式】 config-vs-domain 配置模式

【使用指导】 打开 VSU 模式下 AP 口的本地优先转发特性。

配置 ECMP 本地转发优先模式

- 进入 config-vs-domain 配置模式后，使用 **switch virtual ecmp-lff enable** 命令打开 ecmp 本地转发优先 LFF(Local Forward First)。该命令可选。
- ECMP 路由出口可以分布在 VSU 系统的两个机箱上。用户可以根据实际流量情况，配置 ECMP 的出口流量是否优先从本地成员口进行转发。
- 该功能若关闭，则转发模式根据 ecmp 配置规则转发。具体配置见 ecmp 配置。

【命令格式】 **switch virtual ecmp-lff enable**

【参数说明】 -

【缺省配置】 缺省时该功能是打开的。

【命令模式】 config-vs-domain 配置模式

【使用指导】 打开 VSU 模式下 ecmp 本地成员优先转发。

 VSU 模式下，默认关闭跨机箱 AP 本地转发优先模式及 EMCP 路由口本地转发优先模式。

 三层设备若部署 VSU，建议用户配置基于 IP 的 AP 负载均衡模式（src-ip，dst-ip，src-dst-ip 等）。

检验方法

使用 **show switch virtual balance** 命令查看当前 VSU 系统流量均衡模式。

【命令格式】 **show switch virtual balance**

【参数说明】 -

【命令模式】 特权模式

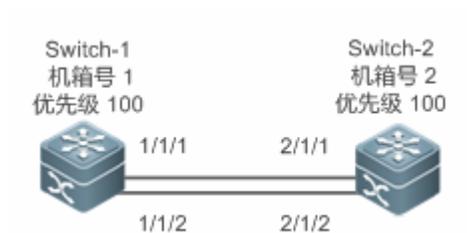
【使用指导】 显示 VSU 模式下的流量均衡模式配置。

配置举例

配置本地优先转发

【网络环境】

图 7-20



上图中 Switch-1 和 Switch-2 组成 VSU，假设 Switch-1 是全局主交换机，在 Switch-1 执行配置。

【配置方法】

- AP 本地优先转发

Switch-1

```
Ruijie#config
Ruijie(config)# switch virtual domain 100
Ruijie(config-vs-domain)# switch virtual aggregateport-lff enable
```

【检验方法】

- 使用 **show switch virtual balance** 命令查看。

Switch-1

```
Ruijie#show switch virtual balance
Aggregate port LFF : enable
Ecmp lff enable
```

常见错误

-

7.5.2.5 配置从VSU模式切换到单机模式

配置效果

将 VSU 系统转换成独立的设备，以单机模式进行工作。

注意事项

-

配置方法；

- 使用 **switch convert mode standalone** [switch_id]命令将设备切换为单机模式。该命令可选。
- 用户执行切换命令后，系统将进行以下提示：“是否将配置文件恢复为之前备份的“standalone.text” 如果选“yes”，则将配置文件恢复；如果选“no”，则清除虚拟设备模式的配置。”

【命令格式】 **switch convert mode standalone** [switch_id]

【参数说明】 switch_id 设备号

【缺省配置】 缺省时设备处于单机模式。

【命令模式】 特权模式

【使用指导】 用户执行 **switch convert mode standalone** 切换命令后，主机箱把 VSU 模式下的各个 VSD 全局配置文件备份为“vsd.virtual_switch.text.vsd 序号”，然后清除清除 VSU 模式下的各个 VSD 全局配置文件“config.text”，并提示用户是否将文件“vsd.standalone.text.vsd 序号”内容覆盖到各个 VSD 的全局配置文件“config.text”，用户选择“yes”，把“vsd.standalone.text.vsd 序号”内容覆盖到各个 VSD 的全局配置文件“config.text”；否则不恢复“config.text”。最后重启交换机。

该命令既可以在单机模式下使用，也可以在 VSU 模式下使用。如果在单机模式下使用，则切换的对象为本机；如果在 VSU 模式下使用，且加上 sw_id 参数，切换的交换机编号为 sw_id，如果没有加上 sw_id 参数，则切换的对象为主机。建议先切换从机，再切换主机。

检验方法

-

配置举例

将设备从 VSU 模式转化成单机模式

【网络环境】

图 7-21



上图中，假设 Switch-1 和 Switch-2 组成 VSU，Switch-1 是全局主设备。

- 【配置方法】
- 把交换设备 1 转化成单机模式
 - 把交换设备 2 转化成单机模式

Switch-1

```
Ruijie# switch convert mode standalone 1
```

```
Ruijie# switch convert mode standalone 2
```

【检验方法】 使用命令 **show switch virtual config** 查看设备的状态。

Switch-1

```
Ruijie#show switch virtual config  
  
switch_id: 1 (mac: 0x1201aeda0M)  
  
!  
  
switch virtual domain 100  
  
!  
  
switch 1  
  
switch 1 priority 100  
  
!  
  
switch convert mode standalone  
  
!  
  
port-member interface Tengigabitethernet 1/1  
  
!  
  
port-member interface Tengigabitethernet 1/3  
  
!
```

```
switch_id: 2 (mac: 0x1201aeda0E)  
  
!  
  
switch virtual domain 100  
  
!  
  
switch 2  
  
switch 2 priority 150  
  
!  
  
switch convert mode standalone  
  
!  
  
port-member interface Tengigabitethernet 1/1  
  
!  
  
port-member interface Tengigabitethernet 1/2  
  
!  
  
switch 2 description switch-2  
  
!
```

常见错误

7.6 监视与维护

查看运行情况

作用	命令
显示当前运行的 VSU 信息，拓扑形状，或当前配置的 VSU 参数	show switch virtual [topology config role]
查看当前双主机配置信息	show switch virtual dual-active { bfd aggregateport summary }
VSU 模式下查看当前的 VSL-AP 运行信息	show switch virtual link [port]
重定向到主机或任意一台设备的控制台	session { device <i>switch_id</i> master }
显示本设备的交换机编号	show switch id

8 VSD

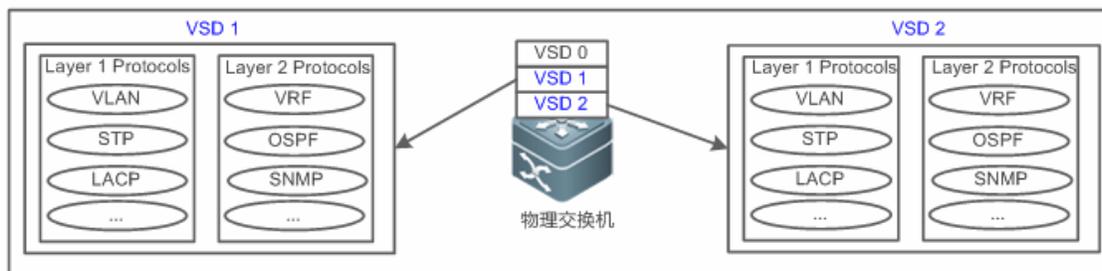
8.1 概述

随着 ICT（信息和通信技术）网络特别是数据中心网络规模不断扩张，业务种类不断丰富，一方面网络管理变得越来越复杂，另一方面对于业务的隔离、安全性、可靠性等网络属性提出的要求也越来越高。此外，随着硬件能力的迅速提升，多框、集群，分布式路由交换系统的成熟，单台物理网络设备的业务处理能力已经达到一个新的高度，如何将单台物理设备强大的业务处理能力充分利用，弹性的适应当前业务需求和未来扩展性的平滑演进已经变得刻不容缓。

VSD（Virtual Switch Device）虚拟设备单元是一种网络系统虚拟化技术，通过虚拟化将一台物理设备划分为多台逻辑设备，划分出的每台逻辑设备就称为一个 VSD。每个 VSD 拥有独立的硬件及软件资源，包括独立的接口资源、CPU 资源，独立维护路由表和转发表，有自己的管理员、配置文件。对于用户来说，每个 VSD 就是一台独立的设备。

通过 VSD 技术，可以将一台物理设备虚拟成多台逻辑意义上的设备，如下图所示。一台物理设备可以承担逻辑拓扑中的多个网络节点，最大限度地利用现有资源，降低了网络运营成本。同时，不同的 VS 可以部署不同的业务，实现业务、故障的隔离，提高了网络的安全性和可靠性。

图 8-1 VSD 示意图



上图中，VSD0 为系统默认的 VSD 即缺省 VSD，VSD1 及 VSD2 为用户创建的非缺省 VSD。

- VSD0：管理域，也称为缺省 VSD，系统默认存在，不可删除，负责非缺省 VSD 的创建和删除，可管理整个设备的物理资源，包括物理端口资源、CPU 资源、内存资源等。
- VSD1：客户域，也称为非缺省 VSD（VSD2 也属于客户域），由管理域 VSD0 创建和删除，其使用的物理资源也由 VSD0 进行分配。客户域只能使用管理域给它分配的资源，不能分配和使用其它资源。
- 物理端口资源：仅指线卡上的物理以太网端口，不包括其余任何类型的接口，比如 Console 口、管理 mgmt 口等都不能被分配。

VSD 的优点

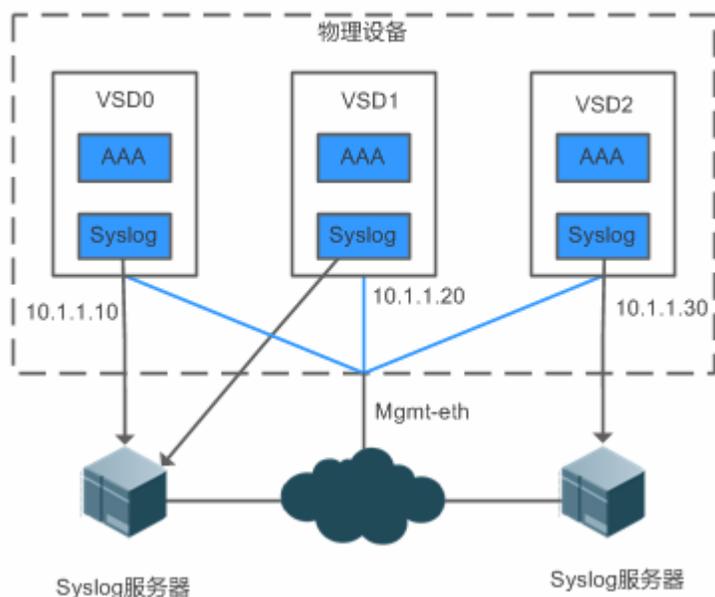
- VSD 基于业务芯片划分物理端口，可以做到业务容量叠加，即 VSD 之间独享业务规格，比如 VLAN 的业务规格为 4063 个，则 VSD1 和 VSD2 的 VLAN 业务规格都为 4063 个。
- 每个 VSD 相当于一台独立的物理设备，VSD 之间相互隔离不能直接通信，具有较高的安全性。

- 灵活调度网络设备资源，比如新的分支机构加入，管理者可以根据需要灵活划分 VSD，来节省购置新网络设备和网络设备硬件升级的开销，提高现有网络资源利用率。

VSD 的管理

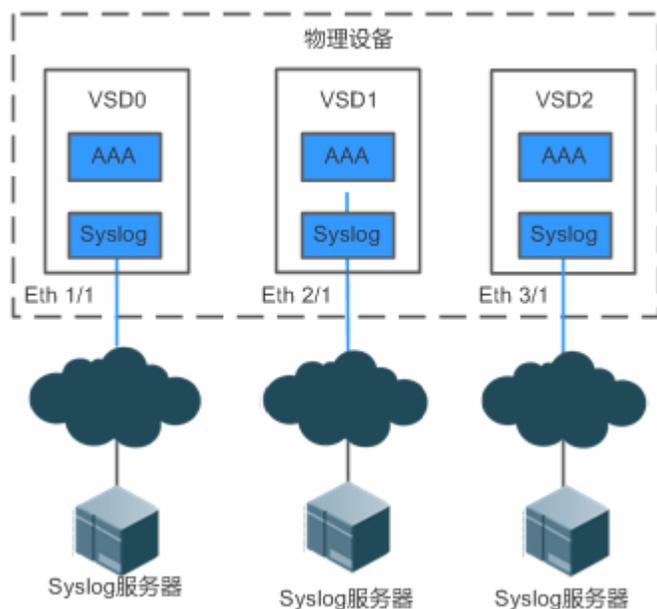
- 带外管理，即通过 mgmt 口进行管理。所有 VSD 共用管理板上的 mgmt 口，各 VSD 可以分别配置自己的 mgmt 口 IP 地址，用户使用哪个 VSD 的 mgmt 口 IP 地址登录，就登录到该 VSD 的控制台上。

图 8-2 带外管理



- 带内管理，即通过以太物理端口进行管理。用户通过哪个 VSD 所属的端口登录，就登录到哪个 VSD 的控制台上。例如，下图中通过 2/1 端口只能登录到 VSD1 的控制台上。

图 8-3 带内管理



- 串口控制台管理，即通过管理板串口进行管理。通过串口，默认登录到 VSD0 的控制台。

协议规范

- VSD 为私有协议。

8.2 典型应用

典型应用	场景描述
在一个物理设备内提供相互独立的业务	同一个物理设备虚拟出几个逻辑设备，利用虚拟出的各逻辑设备提供彼此相互隔离的业务服务。

8.2.1 在一个物理设备内提供相互独立的业务

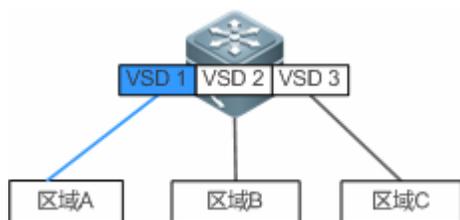
应用场景

利用 VSD 在同一个物理设备中虚拟出多个逻辑设备，虚拟出的各逻辑设备对外提供相互独立的业务服务。

以下图为例，一台汇聚设备虚拟出 3 台逻辑设备，分别为区域 A、区域 B 及区域 C 提供相互独立的服务，彼此之间业务上不会相互影响。

- VSD1 作为区域 A 的上联设备，VSD2 作为区域 B 的上联设备，VSD3 作为区域 C 的上联设备。
- VSD1、VSD2 及 VSD3 分别通过各自的链路同该区域内的设备相连。

图 8-4



功能部属

- VSD1、VSD2 及 VSD3 作为逻辑上相互独立的设备，根据区域 A、区域 B 及区域 C 的各自业务需求部署相应业务。

8.3 功能详解

基本概念

▾ 缺省 VSD

设备支持 VSD 特性后，整台物理设备默认就是一个 VSD，这个 VSD 就是缺省 VSD。用户登录某台支持 VSD 特性的物理设备实际上就是登录缺省 VSD，用户在物理设备上的配置实际上就是针对缺省 VSD 的配置。缺省 VSD 不需要创建，不能被删除。缺省 VSD 拥有对整台物理设备的所有权限，它可以使用和管理设备所有的硬件资源。缺省 VSD 可以创建、删除非缺省 VSD，为创建的非缺省 VSD 分配硬件资源，没有被分配给非缺省 VSD 的硬件资源默认由缺省 VSD 使用。缺省 VSD 的名称为 Ruijie，编号为 0。

▾ 非缺省 VSD

和缺省 VSD 相对应的就是非缺省 VSD，非缺省 VSD 需要通过缺省 VSD 创建后才能使用。非缺省 VSD 创建时没有物理端口资源（其余资源如 CPU、内存等默认和缺省 VSD 共享物理设备的相关资源），需要通过缺省 VSD 为其分配物理端口后才拥有可使用的物理端口资源。非缺省 VSD 下不能再创建/删除非缺省 VSD。

▾ 设备物理资源

设备物理资源主要包括：

- 物理端口资源：指用户物理端口资源，不包括虚拟接口、设备串口等。
- CPU 资源：指各 VSD 可以使用的 CPU 时间。
- SD 卡资源：外部存储介质 SD 卡的存储资源。
- 内存资源：物理设备内存，通常是指板卡内存。

i 物理端口资源默认情况下都属于缺省 VSD，其余物理资源默认情况下缺省 VSD 和非缺省 VSD 共享。

▾ MAC 地址

一台物理交换机通常都存在一个系统 MAC 地址，当一台物理设备被虚拟成几台逻辑设备后对应的逻辑设备也需要系统 MAC 地址，因此每一个 VSD 都会被自动分配 MAC 地址（缺省 VSD 的系统 MAC 直接继承物理设备的系统 MAC）。

📌 VSD 的主从角色

对于机箱式设备，为了支持高可用性，一般都配备有双管理引擎。正常工作时，其中一张管理引擎为主角色，另一张管理引擎为从角色，业务数据从主管理引擎备份到从管理引擎。在划分出 VSD 以后，每个 VSD 在主从管理引擎上均有分布，其中主管理引擎上分布的是主角色，从管理引擎上分布的是从角色。为了避免非缺省 VSD 用户执行主从倒换对其他 VSD 用户带来的影响，管理引擎的角色切换命令只能由 VSD0 的管理员执行。

功能特性

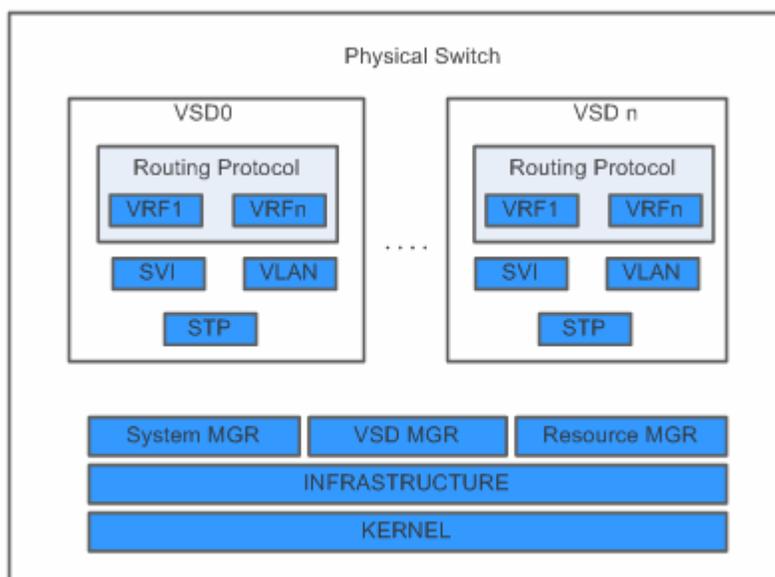
功能特性	作用
VSD	将一个物理设备虚拟成多台逻辑设备。

8.3.1 VSD

工作原理

VSD 可以认为是一种 1 : N 的虚拟化技术，它将一个物理设备虚拟成多台逻辑设备。VSD 的虚拟化基于“操作系统级别虚拟化”技术进行构造，这种技术的虚拟化方法与“虚拟机”技术不同，它是利用操作系统内核提供的系统服务来实现多个相互隔离的用户空间服务。类似采用这种虚拟化的实现的操作系统包括 Linux 系统的 LXC (Linux Containers)、OpenVZ 和 Linux-VServer，FreeBSD 系统的 jails 等。

图 8-5 VSD 的 1:N 虚拟化结构



上图中的 VSD0 即为上文提到的缺省 VSD，VSD1 到 VSDn 为用户创建的非缺省 VSD。

- **System MGR**：当一个 VSD 被创建时，系统管理进程负责启动该 VSD 启动所需的所有进程。而当某个 VSD 需要启用新服务时，系统管理进程负责启动对应的服务进程（如当 VSD2 需要启用 OSPF 服务时，系统管理进程就为 VSD2 启动一个对应的 OSPF 进程）。当一个 VSD 被删除时，系统管理进程负责删除属于该 VSD 的所有服务进程。
- **VSD MGR**：VSD 管理进程主要用于支持非缺省 VSD 管理，负责非缺省 VSD 的创建和删除。更为重要的是 VSD 管理进程提供 VSD 相关的 API 供系统管理进程或资源管理进程调用，以实现非缺省 VSD 的系统管理或资源管理。
- **Resource MGR**：资源管理进程主要负责在各 VSD 间分配资源，如 VLANs、VRFs 等。
- **INFRASTRUCTURE 层**：主要提供更上层应用与硬件资源间的接口。
- **KERNEL**：VSD 属于轻量级的虚拟化，内核依然只有一个。

8.4 配置详解

配置项	配置建议 & 相关命令	
配置VSD	 必须配置。用于创建 VSD。	
	vsd <i>vsd_name</i> [<i>id vsd_number</i>]	创建 VSD
	allocate interface <i>int_index</i>	为 VSD 分配物理端口
	allocate slot <i>slot_id</i>	为 VSD 分配多业务卡
VSD切换登陆	 可选配置。在缺省 VSD 与非缺省 VSD 间切换登陆。	
	switchto vsd <i>vsd_name</i>	从缺省 VSD 登录非缺省 VSD
	switchback	从非缺省 VSD 切换回缺省 VSD

8.4.1 配置VSD

配置效果

- 创建非缺省 VSD。
- 为新创建的非缺省 VSD 分配物理端口资源。
- 为新创建的非缺省 VSD 分配多业务卡资源。

注意事项

- 用户可以在 VSU 环境中使用 VSD 功能。为 VSD 分配端口时，端口所在线卡将会复位，如果端口所在线卡上有 VSL 链路，该链路也会因为线卡复位而断开，系统内只有一条 VSL 链路，VSU 将会发生拓扑分裂。为了保证 VSU 稳定性，建议在使用 VSD 功能的 VSU 环境上，在多张线卡上配置 VSL 端口，以避免分配 VSD 端口或者重启 VSD 等操作对 VSU 的稳定性造成的影响。
- VSD0 的管理员可以把任一 VSD 的符合作为 VSL 端口条件的普通端口设置为 VSL 端口，转换为 VSL 端口以后，该端口统一归 VSD0 管理。VSD0 的管理员可以把任一 VSL 端口转换为普通端口，转换成普通端口以后，该端口将回到该端口

所在芯片所属的 VSD 中。例如：某张线卡上所有端口都在 VSD2 中，只有一个端口被配置为 VSL 端口，管理员将该 VSL 端口转换为普通口时，普通口将回到 VSD2 中。

- VSU 环境下的双主机检测口需要配置在 VSD0 中，当 VSD0 检测到双主机状态以后，所有 VSD 都会进入 recovery 状态。

配置方法

创建 VSD

- 要使用 VSD 功能就必须配置此功能项。只能在缺省 VSD (VSD0) 创建或删除非缺省 VSD。

【命令格式】 **vsd vsd_name [id vsd_number]**

【参数说明】 *vsd_name*：非缺省 VSD 名。

vsd_number：VSD 的编号。不指定该参数时，系统会给 VSD 自动分配一个目前可用的最小的编号。

【缺省配置】 默认不创建任何非缺省 VSD。

【命令模式】 全局模式

【使用指导】 此命令用来创建一个 VSD 并进入 VSD 配置模式，如果 VSD 已经创建，则该命令用来进入 VSD 配置模式。进入指定 VSD 配置模式时，可以不输入 *vsd_number*。但如果输入，则必须和 VSD 当前的编号一致，否则会提示错误信息。

 需要获取对应授权后才能创建非缺省 VSD。VSD 的名称和设备的 hostname 为两个相互独立的名称，两者之间没有任何关联。

分配物理端口资源

- 必须配置。
- 一个物理端口被分配给某个非缺省 VSD 时该端口上的所有配置将被删除且一个物理端口只能属于一个 VSD。

- ✔ 交换机的物理端口在分配时将以端口组的形式被分配（物理上属于同一个交换芯片的所有端口称为一个端口组）。当端口组发生所属变动时（被分配或被回收）该端口组所在的业务卡将发生复位。VSU 模式下，VSL 端口所在线卡的端口被重新分配可能导致 VSU 分裂，建议在业务空闲时执行。

【命令格式】 **allocate interface int_index**

【参数说明】 *int_index*：接口索引

【缺省配置】 所有物理端口默认属于缺省 VSD，也就是 VSD0。

【命令模式】 VSD 配置模式

【使用指导】 此命令用来为非缺省 VSD 分配物理端口以及从非缺省 VSD 中回收物理端口。

 如果一个非缺省 VSD 无物理端口可分，那该 VSD 将无法使用但可以管理。允许所有的物理端口都被分配给非缺省 VSD，此时缺省 VSD 将无物理端口可用。VSL 口无法被分配。

分配多业务卡资源

- 可选配置。当用户需要在某个非缺省 VSD 中使用防火墙卡等多业务卡时，需要在该 VSD 模式下执行此配置项。

- 当某个非缺省 VSD 使用了某张多业务卡，其他非缺省 VSD 则不能使用。例如非缺省 VSD2 使用了某张防火墙卡，其他 VSD 则不能使用。

【命令格式】 **allocate slot slot_id**

【参数说明】 *slot_id* : 多业务卡槽位标识

【缺省配置】 所有多业务卡默认属于缺省 VSD，也就是 VSD0。

【命令模式】 VSD 配置模式

【使用指导】 此命令用来为非缺省 VSD 分配多业务卡资源，no 命令用来从非缺省 VSD 中回收多业务卡资源。

检验方法

新创建的 VSD 为一台独立的逻辑设备，可通过登录该 VSD 的方式进行检验配置是否正确。

- 检查能否登录该新创建的 VSD。
- 登录该 VSD 后检查能否对端口进行配置。

配置举例

在一台物理设备上创建 3 个逻辑设备进行使用

【网络环境】

图 8-6



【配置方法】

- 创建 VSDA ;
- 为 VSDA 分配物理端口 1/1-24 ;
- 为 VSDA 分配多业务卡 3
- 创建 VSDB ;
- 为 VSDB 分配物理端口 1/25-48 ;
- 为 VSDB 分配多业务卡 4
- 创建 VSDC ;
- 为 VSDC 分配物理端口 2/1-12。
- 为 VSDC 分配多业务卡 5

VSDA

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# vsd VSDA
Ruijie(config-vsd)# allocate int gi 1/1
Interface-group[1/1 ~ 1/24] and their config will be removed from vsd[0]. Are you sure to
continue(y/n)? [no]y
Ruijie(config-vsd)# allocate slot 3
Allocating slot 3 may cause some service in source vsd to stop. Are you sure to continue (y/n)? [no]y
Ruijie(config-vsd)#
```

VSDB

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# vsd VSDB
Ruijie(config-vsd)# allocate int gi 1/25
Interface-group[1/25 ~ 1/48] and their config will be removed from vsd[0]. Are you sure to
continue(y/n)? [no]y
Ruijie(config-vsd)# allocate slot 4
Allocating slot 4 may cause some service in source vsd to stop. Are you sure to continue (y/n)? [no]y
Ruijie(config-vsd)#
```

VSDC

```
Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# vsd VSDC
Ruijie(config-vsd)# allocate int gi 2/1
Interface-group[2/1 ~ 2/12] and their config will be removed from vsd[0]. Are you sure to
continue(y/n)? [no]y
Ruijie(config-vsd)# allocate slot 5
Allocating slot 5 may cause some service in source vsd to stop. Are you sure to continue (y/n)? [no]y
Ruijie(config-vsd)#
```

【检验方法】 登录到对应的 vsd 进行相关端口的配置

VSDA

```
Ruijie# swichto vsd VSDA
*****
Ruijie General Operating System Software
Copyright (c) 1998-2013s by Ruijie Networks.
```

```
All Rights Reserved.
Neither Decompiling Nor Reverse Engineering Shall Be Allowed.
*****
Ruijie-VSDA> enable
Ruijie-VSDA# show int
```

VSDB

```
Ruijie# switchto vsd VSDB
*****
Ruijie General Operating System Software
Copyright (c) 1998-2013s by Ruijie Networks.
All Rights Reserved.
Neither Decompiling Nor Reverse Engineering Shall Be Allowed.
*****
Ruijie-VSDA> enable
Ruijie-VSDA# show int
```

VSDC

```
Ruijie# switchto vsd VSDC
*****
Ruijie General Operating System Software
Copyright (c) 1998-2013s by Ruijie Networks.
All Rights Reserved.
Neither Decompiling Nor Reverse Engineering Shall Be Allowed.
*****
Ruijie-VSDA> enable
Ruijie-VSDA# show int
```

常见错误

- 没有有效授权就创建。
- 创建的非缺省 VSD 的总数已经达到该设备支持的最大值。
- 将不属于某非缺省 VSD 的端口进行回收。
- 将已经属于某非缺省 VSD 的物理端口再直接分配给另一个非缺省 VSD。
- 没有新的物理端口资源进行分配。

8.4.2 VSD切换登陆

配置效果

- 从 VSD0 登陆到其它非缺省 VSD 或从非缺省 VSD 切换回 VSD0。

注意事项

- 通过带外或带内直接登陆非缺省 VSD 的用户无法进行登陆切换。

配置方法

- 需要在 VSD 间切换登陆，则应该执行此配置项。
- 使用 **switchto vsd** 命令可以从 VSD0 登录到任何非缺省 VSD；使用 **switchback** 命令可以从任何非缺省 VSD 登录回到 VSD0。
- 非缺省 VSD 间切换必须通过缺省 VSD 进行，例如需要从非缺省 VSD1 切换到非缺省 VSD2，必须从 VSD1 切换到缺省 VSD0，再从缺省 VSD 切换到非缺省 VSD2。
- 只有 VSD0 的管理人员才能使用 **switchto vsd** 和 **switchback** 在 VSD 间进行切换。

【命令格式】 **switchto vsd vsd_name**
【参数说明】 *vsd_name*：需要登陆的 VSD 名称。
【缺省配置】 -
【命令模式】 特权模式
【使用指导】 此命令用来从缺省 VSD 登录非缺省 VSD。

【命令格式】 **switchback**
【参数说明】 -
【缺省配置】 -
【命令模式】 特权模式
【使用指导】 此命令用来从非缺省 VSD 切换回缺省 VSD。但如果是直接通过 telnet 的方式登录非缺省 VSD，则不支持该命令，即该命令只是在通过从缺省 VSD 登录到非缺省 VSD 时才生效（也就是要先有 switchto，才有 switchback）。

检验方法

登陆后查看当前 VSD 的 VSD ID。

配置举例

📌 切换到非缺省 VSD

【配置方法】

- 从 vsd0 登陆非缺省 VSD ;
- 从非缺省 VSD 登陆回 VSD0。

```
Ruijie# switchto vsd admin
*****
Ruijie General Operating System Software
Copyright (c) 1998-2013s by Ruijie Networks.
All Rights Reserved.
Neither Decompiling Nor Reverse Engineering Shall Be Allowed.
*****
admin#
```

```
Ruijie# switchto vsd admin
*****
Ruijie General Operating System Software
Copyright (c) 1998-2013s by Ruijie Networks.
All Rights Reserved.
Neither Decompiling Nor Reverse Engineering Shall Be Allowed.
*****
admin# switchback
Ruijie#
```

【检验方法】 查看当前用户所在的 VSD

```
Ruijie# sho vsd current-vsds
Current vsd is 0 - Ruijie
Ruijie #
```

常见配置错误

8.5 监视与维护

查看运行情况

作用	命令
查看 VSD 相关信息。	<code>show vsd { current-vsds membership detail all } [vsd_name]</code>



配置指南-网管和监控

本分册介绍网管和监控配置指南相关内容，包括以下章节：

1. SNMP
2. RMON
3. NTP
4. SNTP
5. SPAN-RSPAN
6. SFlow

1 SNMP

1.1 概述

SNMP 是 Simple Network Management Protocol (简单网络管理协议) 的缩写, 在 1988 年 8 月就成为一个网络管理标准 RFC1157。到目前, 因众多厂家对该协议的支持, SNMP 已成为事实上的网管标准, 适合于在多厂家系统的互连环境中使用。利用 SNMP 协议, 网络管理员可以对网络上的节点进行信息查询、网络配置、故障定位、容量规划, 网络监控和管理是 SNMP 的基本功能。

📌 SNMP 协议版本

目前 SNMP 支持以下版本:

- SNMPv1 : 简单网络管理协议的第一个正式版本, 在 RFC1157 中定义。
- SNMPv2C : 基于共同体 (Community-Based) 的 SNMPv2 管理架构, 在 RFC1901 中定义。
- SNMPv3 : 通过对数据进行鉴别和加密, 提供了以下的安全特性:
 1. 确保数据在传输过程中不被篡改;
 2. 确保数据从合法的数据源发出;
 3. 加密报文, 确保数据的机密性。

协议规范

- RFC 1157 , Simple Network Management Protocol (SNMP)
- RFC 1901 , Introduction to Community-based SNMPv2
- RFC 2578 , Structure of Management Information Version 2 (SMIv2)
- RFC 2579 , Textual Conventions for SMIv2
- RFC 3411 , An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks
- RFC 3412 , Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)
- RFC 3413 , Simple Network Management Protocol (SNMP) Applications
- RFC 3414 , User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)
- RFC 3415 , View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)
- RFC 3416 , Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)
- RFC 3417 , Transport Mappings for the Simple Network Management Protocol (SNMP)
- RFC 3418 , Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)
- RFC 3419 , Textual Conventions for Transport Addresses

1.2 典型应用

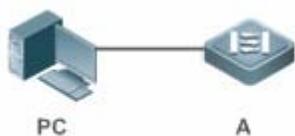
典型应用	场景描述
通过SNMP管理网络设备	通过 SNMP 网络管理器对网络设备进行管理和监控。

1.2.1 通过SNMP管理网络设备

应用场景

以下图为例，用户通过 SNMP 网络管理器，来对网络设备 A 进行管理和监控。

图 1-1



【注释】 A 为需要被管理的网络设备。
PC 为网络管理站。

功能部属

网络管理站和被管理的网络设备通过网络连接，用户在网络管理站上，通过 SNMP 网络管理器，访问网络设备上的管理信息数据库，以及接收来自网络设备主动发出的消息，来对网络设备进行管理和监控。

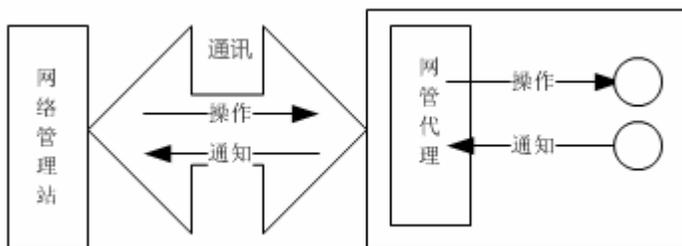
1.3 功能详解

基本概念

SNMP 是一个应用层协议，为客户机/服务器模式，包括三个部分：

- SNMP 网络管理器
- SNMP 代理
- MIB 管理信息库

图 1-2 网络管理站（NMS）与网管代理（Agent）的关系图



SNMP 网络管理器

SNMP 网络管理器，是采用 SNMP 来对网络进行控制和监控系统，也称为 NMS (Network Management System)。

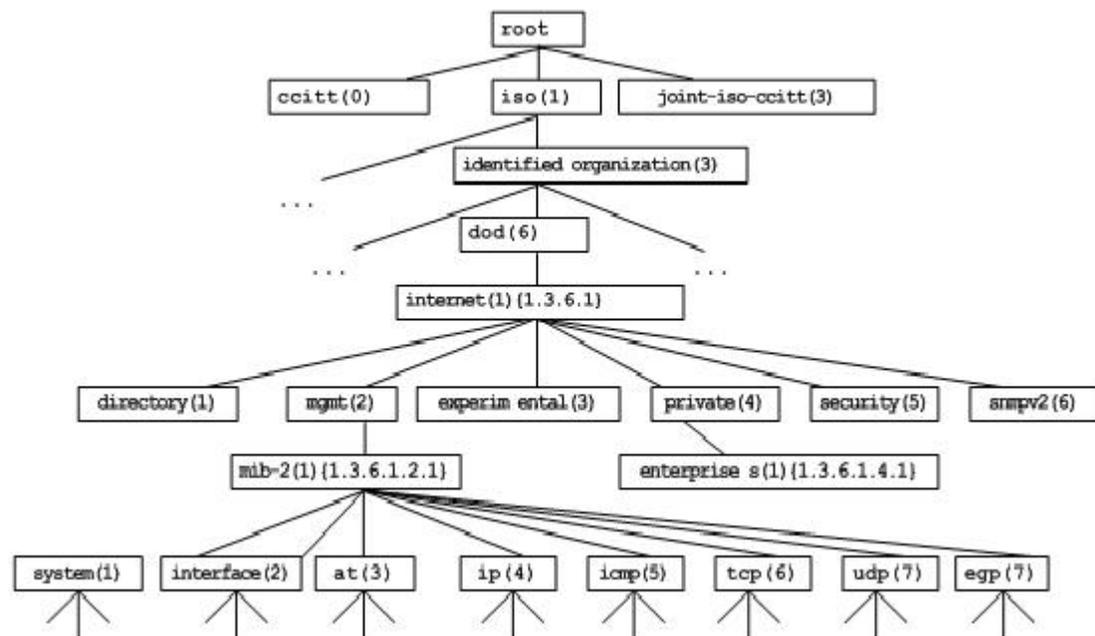
SNMP 代理

SNMP 代理 (SNMP Agent, 下文简称为 Agent) 是运行在被管理设备上的软件，负责接受、处理并且响应来自 NMS 的监控和控制报文，也可以主动发送一些消息报文给 NMS。

MIB

MIB (Management Information Base) 是一个虚拟的网络管理信息库。被管理的网络设备中包含大量信息，为了能在 SNMP 报文中唯一的标识某个特定的管理单元，MIB 采用树形层次结构来描述，树的节点表示某个特定的管理单元。为了唯一标识网络设备中的某个管理单元 System，可以采用一串的数字来表示，MIB 则是网络设备的单元标识符的集合。

图 1-3 MIB 树形层次结构



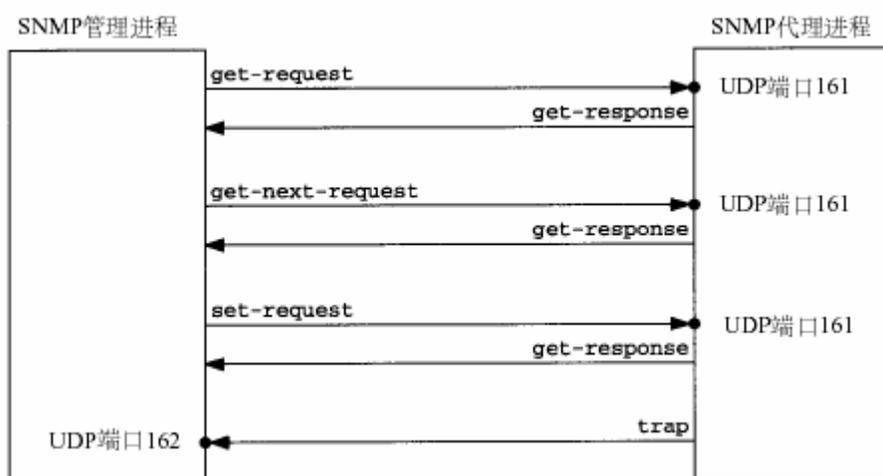
操作类型

SNMP 协议中的 NMS 和 Agent 之间的交互信息，定义了 6 种操作类型：

- Get-request 操作：NMS 从 Agent 提取一个或多个参数值。
- Get-next-request 操作：NMS 从 Agent 提取一个或多个参数的下一个参数值。
- Get-bulk 操作：NMS 从 Agent 提取批量的参数值；
- Set-request 操作：NMS 设置 Agent 的一个或多个参数值。
- Get-response 操作：Agent 返回的一个或多个参数值，是 Agent 对 NMS 前面 3 个操作的响应操作。
- Trap 操作：Agent 主动发出的报文，通知 NMS 有某些事情发生。

前面的 4 个报文是由 NMS 向 Agent 发出的，后面两个是 Agent 发给 NMS 的（注意：SNMPv1 版本不支持 Get-bulk 操作）。下图描述了这几种操作。

图 1-4 SNMP 的报文类型



NMS 向 Agent 发出的前面 3 种操作和 Agent 的应答操作采用 UDP 的 161 端口。Agent 发出的 Trap 操作采用 UDP 的 162 端口。

功能特性

功能特性	作用
SNMP基本功能	配置网络设备上的 SNMP 代理，实现对网络上的节点进行信息查询、网络配置、故障定位、容量规划等基本功能。
SNMPv1 及SNMPv2C	采用基于共同体的安全架构，包括认证名和访问权限。
SNMPv3	SNMPv3 重新定义了 SNMP 架构，主要是在安全功能上进行了增强，包括支持基于用户的安全模型，以及支持基于视图的访问控制模型等。SNMPv3 架构内已经包含了 SNMPv1 和 SNMPv2C 所有的功能。

1.3.1 SNMP基本功能

工作原理

工作过程

SNMP协议交互是应答式的（报文交互参见图 1-4 SNMP的报文类型）。NSM向Agent主动发起请求，包括Get-request、Get-next-request、Get-bulk和Set-request，Agent接收请求并完成操作后以Get-response作为应答。Agent有时候也会向NSM主动发出Trap和Inform消息，其中Trap消息不需要应答，而Inform消息则需要NSM回送一个Inform-response应答，表示收到消息，否则Agent将会重发Inform消息。

相关配置

屏蔽或关闭 SNMP 代理

缺省时启动 SNMP 功能。

使用 **no snmp-server** 命令屏蔽 SNMP 代理功能。

执行 **no enable service snmp-agent** 命令，直接关闭 SNMP 所有服务。

设置 SNMP 基本参数

缺省时系统联系方式、系统位置和设备的网元信息为空；序列号缺省值是 60FF60；缺省最大数据报文长度 1572 字节；缺省的 SNMP 服务 UDP 端口号是 161。

使用 **snmp-server contact** 命令配置或删除系统联系方式。

使用 **snmp-server location** 命令配置或删除系统位置。

使用 **snmp-server chassis-id** 命令配置系统序列码或恢复缺省值。

使用 **snmp-server packet-size** 命令配置代理最大数据报文长度或恢复缺省值。

使用 **snmp-server net-id** 命令配置或删除设备的网元信息。

使用 **snmp-server udp-port** 命令设置 SNMP 服务 UDP 端口号或恢复缺省值。

配置 SNMP 主机地址

缺省情况下，没有 SNMP 主机。

使用 **snmp-server host** 命令配置 Agent 主动发送消息的 NMS 主机地址或删除指定 SNMP 主机地址。发给主机的消息可以绑定 SNMP 的版本、接收端口、认证名或用户。该命令与 **snmp-server enable traps** 命令一起使用，主动给 NMS 发送 Trap 消息。

设置 Trap 消息参数

缺省情况下，禁止 SNMP 向 NMS 主动发送 Trap 消息；打开接口发送 Link Trap 功能；关闭发送系统重启 Trap 功能；Trap 消息缺省不带私有字段。

缺省时，SNMP 报文从哪个接口出去，就使用哪个接口的 IP 地址作为源地址。

缺省时 Trap 消息报文的队列长度为 10，发送 Trap 消息的时间间隔为 30 秒。

使用 **snmp-server enable traps** 命令配置 Agent 主动或禁止向 NMS 发送 Trap 消息。

使用 **snmp trap link-status** 命令打开或关闭接口发送 Link Trap 功能。

使用 **snmp-server trap-source** 命令指定发送消息的源地址或恢复缺省值。

使用 **snmp-server queue-length** 命令设置 Trap 消息报文的队列长度或恢复缺省值。

使用 **snmp-server trap-timeout** 命令设置发送 Trap 消息的时间间隔或恢复缺省值。

使用 **snmp-server trap-format private** 命令设置或关闭发送 Trap 消息时携带私有字段的功能。

使用 **snmp-server system-shutdown** 命令打开或关闭发送系统重启 Trap 功能。

📌 设置 SNMP 日志功能记录 NMS 对 SNMP Agent 的 GET、GET-NEXT 和 SET 操作

缺省情况下，SNMP 日志开关处于关闭状态。

使用 **snmp-server logging { get-operation | set-operation }** 命令打开记录 get、set 操作的功能，get 开关控制 GET、GET-NEXT 操作记录，set 开关控制 SET 操作记录

1.3.2 SNMPv1 及SNMPv2C

SNMPv1 和 SNMPv2C 都采用基于共同体(Community-based)的安全架构。通过定义主机地址以及认证名(Community String)来限定能够对代理的 MIB 进行操作的管理者。

工作原理

SNMPv1 和 SNMPv2 版本使用认证名来鉴别是否有权使用 MIB 对象。为了能够管理设备，网络管理系统 (NMS)的认证名必须同设备中定义的某个认证名一致。

SNMPv2C 增加了 Get-bulk 操作机制并且能够对管理工作站返回更加详细的错误信息类型。Get-bulk 操作能够一次性地获取表格中的所有信息或者获取大批量的数据，从而减少请求-响应的次数。SNMPv2C 错误处理能力的提高包括扩充错误代码以区分不同类型的错误，而在 SNMPv1 中这些错误仅有一种错误代码。现在通过错误代码可以区分错误类型。由于网络上可能同时存在支持 SNMPv1 和 SNMPv2C 的管理工作站，因此 SNMP 代理必须能够识别 SNMPv1 和 SNMPv2C 报文，并且能返回相应版本的报文。

📌 安全

一个认证名有以下属性：

- 只读(Read-only)：为被授权的管理工作站提供对所有 MIB 变量的读权限。
- 读写(Read-write)：为被授权的管理工作站提供对所有 MIB 变量的读写权限。

相关配置

设置认证名及访问权限

所有认证名的缺省访问权限为只读。

使用 `snmp-server community` 命令配置或删除认证名和访问权限。

该命令为启用设备 SNMP 代理功能的第一个重要命令，指定了团体的属性、允许访问 MIB 的 NMS 范围等等。

1.3.3 SNMPv3

SNMPv3 重新定义了 SNMP 架构，将之前的 SNMPv1 和 SNMPv2 的功能也纳入到 SNMPv3 体系中。

工作原理

网络管理系统 (NMS) 和 SNMP 代理 (SNMP Agent) 都称为 SNMP 实体。在 SNMPv3 架构中，SNMP 实体分为引擎和应用两大部分，其中 SNMP 引擎用于发送和接收信息、鉴定和加密信息以及对管理对象的控制访问。SNMP 应用指的是 SNMP 内部的应用程序，利用 SNMP 引擎提供的服务进行工作。

SNMPv3 版本使用基于用户的安全模型 (USM) 来鉴别是否有权使用 MIB 对象。为了能够管理设备，网络管理系统 (NMS) 的用户和安全级别必须同设备中定义的某个 SNMP 用户一致。

SNMPv3 版本规定 NSM 在管理设备的时候，必须先得知设备上 SNMP Agent 的引擎标识。SNMPv3 定义了 Discover 和 Report 操作机制，NSM 在不知道 Agent 引擎标识的情况下，可以先向 Agent 发送 Discover 报文，而 Agent 以 Report 响应，并在响应报文中携带了引擎标识信息。此后，NSM 和 Agent 之间的管理操作必须携带该引擎标识。

安全

- SNMPv3 通过安全模型以及安全级别来确定对数据采用哪种安全机制进行处理。目前可用的安全模型有三种类别：SNMPv1、SNMPv2C、SNMPv3。SNMPv3 将 SNMPv1 和 SNMPv2C 也纳入到安全模型中。

SNMPv1 及 SNMPv2C 安全模型和级别

安全模型	安全级别	鉴别	加密	说明
SNMPv1	noAuthNoPriv	认证名	无	通过认证名确认数据的合法性
SNMPv2c	noAuthNoPriv	认证名	无	通过认证名确认数据的合法性

SNMPv3 安全模型以及安全级别

安全模型	安全级别	鉴别	加密	说明
SNMPv3	noAuthNoPriv	用户名	无	通过用户名确认数据的合法性
SNMPv3	authNoPriv	MD5 或者 SHA	无	提供基于 HMAC-MD5 或者 HMAC-SHA 的数据鉴别机制
SNMPv3	authPriv	MD5 或者 SHA	DES	提供基于 HMAC-MD5 或者 HMAC-SHA 的数据鉴别机制提供基于 CBC-DES 的数据加密机制

引擎标识

引擎标识用于唯一标识一个 SNMP 引擎。由于每个 SNMP 实体仅包含一个 SNMP 引擎，它将在一个管理域中唯一标识一个 SNMP 实体。因此，作为一个实体的 SNMPv3 代理必须拥有一个唯一的引擎标识，即 `SnmpEngineID`。

引擎标识为一个 OCTET STRING，长度为 5~32 字节长。在 RFC3411 中定义了引擎标识的格式：

- 前 4 个字节标识厂商的私有企业号（由 IANA 分配），用 HEX 表示。
- 第 5 个字节表示剩下的字节如何标识：
- 0：保留
- 1：后面 4 个字节是一个 Ipv4 地址。
- 2：后面 16 个字节是一个 Ipv6 地址。
- 3：后面 6 个字节是一个 MAC 地址。
- 4：文本，最长 27 个字节，由厂商自行定义。
- 5：16 进制值，最长 27 个字节，由厂商自行定义。
- 6-127：保留。
- 128-255：由厂商特定的格式。

相关配置

配置 MIB 视图和组

缺省配置一个 default 视图，允许访问所有的 MIB 对象。

缺省没有配置用户组。

使用 **snmp-server view** 命令配置或删除视图；使用 **snmp-server group** 命令配置或删除用户组。

可以配置一条或者多条指令，来指定多个不同的共同体名称，使得网络设备可以供不同的权限的 NMS 的管理。

配置 SNMP 用户

缺省没有配置用户。

配置 **snmp-server user** 命令配置或删除用户。

NMS 只有使用合法的用户才能同代理进行通信。

对于 SNMPv3 用户，可以指定安全级别（是否需要认证、是否需要加密等）、认证算法（MD5 或 SHA）、认证口令、加密算法（目前只有 DES）和加密口令。

1.4 配置详解

配置项	配置建议 & 相关命令	
配置SNMP基本功能	 必须配置。使用户可以通过 NMS 访问 Agent。	
	<table border="1"> <tr> <td>enable service snmp-agent</td> <td>启动 Agent 功能。</td> </tr> </table>	enable service snmp-agent
enable service snmp-agent	启动 Agent 功能。	

	snmp-server community	配置认证名和访问权限。
	snmp-server logging	设置日志功能
	snmp-server user	配置 SNMP 用户信息。
	snmp-server view	配置 SNMP 视图。
	snmp-server group	配置 SNMP 用户组。
启用Trap功能	 可选配置。使 Agent 主动向 NMS 发送 Trap 消息。	
	snmp-server host	配置 NMS 主机地址。
	snmp-server enable traps	Agent 主动向 NMS 发送 Trap 消息。
	snmp trap link-status	打开接口发送 Link Trap 功能。
	snmp-server system-shutdown	打开发送系统重启 Trap 功能。
	snmp-server trap-source	指定发送 Trap 消息的源地址。
	snmp-server trap-format private	发送 Trap 消息时携带私有字段
屏蔽Agent功能	 可选配置。在不需要 Agent 服务的时候，屏蔽 Agent 功能。	
	no snmp-server	屏蔽 Agent 功能。
设置SNMP控制参数	 可选配置。用于设置或修改 SNMP 控制参数。	
	snmp-server contact	设置设备的联系方式。
	snmp-server location	设置设备位置。
	snmp-server chassis-id	设置设备序列码。
	snmp-server net-id	设置设备的网元信息。
	snmp-server packetsize	修改最大数据报文长度。
	snmp-server udp-port	修改 SNMP 服务 UDP 端口号。
	snmp-server queue-length	修改 Trap 消息报文的队列长度。
snmp-server trap-timeout	修改发送 Trap 消息的时间间隔。	

1.4.1 配置SNMP基本功能

配置效果

使用户可以通过 NMS 访问 Agent。

注意事项

- 网络设备上默认没有设置认证名，无法使用 SNMPv1 或 SNMPv2C 访问网络设备的 MIB。设置认证名时，如果没有指定访问权限，则默认访问权限是只读（Read-only）。

配置方法

配置 SNMP 视图

- 可选配置。
- 使用基于视图的访问控制（VACM）功能时需要进行配置。

配置 SNMP 用户组

- 可选配置。
- 使用基于视图的访问控制（VACM）功能时需要进行配置。

配置认证名和访问权限

- 必选配置。
- 使用 SNMPv1 和 SNMPv2C 管理网络设备必须在 agent 设备上设置认证名。

配置 SNMP 用户信息

- 必选配置。
- 使用 SNMPv3 管理网络设备必须设置用户。

启动 Agent 功能

- 可选配置。
- 默认开启 Agent 功能，在 Agent 功能关闭后需要再次开启时，须使用此命令。

检验方法

使用 `show snmp` 命令查看设备上的 snmp 功能。

相关命令

配置 SNMP 视图

【命令格式】 `snmp-server view view-name oid-tree { include | exclude }`

【参数说明】 `view-name`：视图名。

`oid-tree`：视图关联的 MIB 对象，是一棵 MIB 子树。

`include`：标明该 MIB 对象子树被包含在视图之内。

`exclude`：标明该 MIB 对象子树被排除在视图之外。

【命令模式】 全局配置模式

【使用指导】 指定视图的名称，用于基于视图的管理。

配置 SNMP 用户组

【命令格式】 `snmp-server group groupname { v1 | v2c | v3 { auth | noauth | priv } } [read readview] [write writeview] [access { aclnum | aclname }]`

【参数说明】 `v1 | v2c | v3`：指明 SNMP 版本。

auth : 该组的用户传输的消息需要验证但数据不需要保密, 只对 v3 有效。

noauth : 该组用户传输的消息不需要验证数据也不需要保密, 只对 v3 有效。

priv : 该组用户传输的消息需要验证同时传输的数据需要保密, 只对 v3 有效。

readview : 关联一个只读的视图。

writeview : 关联一个读写视图。

aclnum : 访问列表序列号, 关联指定的访问列表, 指定能访问 MIB 的 ipv4 NMS 地址范围。

aclname : 访问列表名称, 关联指定的访问列表, 指定能访问 MIB 的 ipv4 NMS 地址范围。

【命令模式】

全局配置模式

【使用指导】

将某些用户和一个组关联, 再将某个组与某个视图关联。一个组内的用户具有相同的访问权限。通过这种方式判定操作关联的管理对象是否在视图允许之内, 只有在视图允许之内的管理对象才被允许访问。

配置认证名和访问权限

【命令格式】

snmp-server community [0 | 7] *string* [**view** *view-name*] [[**ro** | **rw**] [**host** *ipaddr*]] [**aclnum** | **aclname**]

【参数说明】

0 : 表示输入的团体字符串为明文字符串。

7 : 表示输入的团体字符串为密文字符串。

string : 团体字符串, 相当于 NMS 和 SNMP 代理之间的通信密码。

view-name : 指定视图的名称, 用于基于视图的管理。

ro : 指定 NMS 对 MIB 的变量只能读, 不能修改。

rw : NMS 对 MIB 的变量可读可写。

aclnum : 访问列表序列号, 关联指定的访问列表, 指定能访问 MIB 的 ipv4 NMS 地址范围。

aclname : 访问列表名称, 关联指定的访问列表, 指定能访问 MIB 的 ipv4 NMS 地址范围。

【命令模式】

全局配置模式

【使用指导】

该命令为启用设备 SNMP 代理功能的第一个重要命令, 指定了团体的属性、允许访问 MIB 的 NMS 范围等等。要关闭 SNMP 代理功能, 执行 **no snmp-server** 命令即可。

配置 SNMP 用户

【命令格式】

snmp-server user *username* *groupname* { **v1** | **v2c** | **v3** [**encrypted**] [**auth** { **md5** | **sha** } *auth-password*] [**priv** **des56** *priv-password*] } [**access** { **aclnum** | **aclname** }]

【参数说明】

username : 用户名。

groupname : 该用户对应的组名。

v1 | **v2c** | **v3** : 指明 SNMP 版本。只有 v3 支持后面的安全参数。

encrypted : 指定的是密码输入的方式为密文输入。否则, 以明文输入。如果选择了以密文输入, 则需要输入连续的 16 进制数字字符表示的密钥。注意使用 MD5 的认证密钥长度为 16 字节, 而 SHA 认证协议密钥长度为 20 字节。以两个字符表示一个字节。加密表示的密钥仅对本引擎有效。

auth : 指定是否使用验证。

md5 : 指定使用 MD5 认证协议。**sha** 指定使用 SHA 认证协议。

auth-password : 配置认证协议使用的口令字符串 (不超过 32 个字符)。系统将这些口令转换成相应的认证密钥。

priv : 指定是否使用保密。**des56** 指明使用 56 位的 DES 加密协议。

priv-password : 为加密用的口令字符串 (不超过 32 个字符)。系统将这个口令转换成相应的加密密钥。

aclnum : 访问列表序列号, 关联指定的访问列表, 指定能访问 MIB 的 ipv4 NMS 地址范围。

aclname : 访问列表名称, 关联指定的访问列表, 指定能访问 MIB 的 ipv4 NMS 地址范围。

【命令模式】 全局配置模式

【使用指导】 配置用户的信息, 以使 NMS 使用合法的用户同代理进行通信。

对于 SNMPv3 用户, 可以指定安全级别、认证算法 (MD5 或 SHA)、认证口令、加密算法 (目前只有 DES) 和加密口令。

启动 Agent 功能

【命令格式】 **enable service snmp-agent**

【参数说明】

【配置模式】 特权用户模式

【使用指导】 该命令用于启动设备的 SNMP 代理功能。

设置 SNMP 日志功能记录 NMS 对 SNMP Agent 的 GET、GET-NEXT 和 SET 操作

【命令格式】 **snmp-server logging { get-operation | set-operation }**

【参数说明】 **get-operation** : 记录 GET、GET-NEXT 操作的开关

set-operation : 记录 SET 操作的开关

【命令模式】 全局配置模式

【使用指导】 使用此命令将记录 NMS 对 SNMP Agent 的 GET、GET-NEXT、SET 操作。当进行 GET 和 GET-NEXT 操作时, Agent 会记录 NMS 用户的 IP 地址, 操作类型以及操作的节点 OID。当进行 SET 操作时, Agent 会记录 NMS 用户的 IP 地址, 操作类型, 操作的节点 OID 和设置的值。

 大量的日志记录会影响设备的性能, 正常情况下, 建议关闭 SNMP 日志功能。

显示 SNMP 的状态信息

【命令格式】 **show snmp [mib | user | view | group | host]**

【参数说明】 **mib** : 显示系统中支持的 snmp mib 信息。

user : 显示 snmp 用户信息。

view : 显示 snmp 视图信息。

group : 显示 snmp 用户组信息。

host : 显示用户配置的显示信息。

【配置模式】 特权用户模式

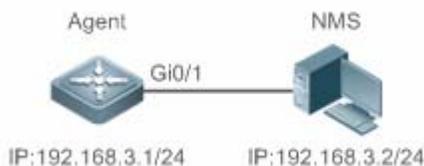
【使用指导】 -

配置举例

SNMPv3 配置举例

【网络环境】

图 1-5



- 网络工作站(NMS)基于用户的认证加密模式对网络设备(Agent)进行管理。例如 :使用用户名 “user1” , 认证方式为 MD5 , 认证密码为 123 , 加密算法为 DES56 , 加密密码为 321。
- 网络设备能够控制用户访问 MIB 对象的操作权限。例如 : 用户 “user1” 可以对 System (1.3.6.1.2.1.1) 节点下的 MIB 对象进行读操作 , 其中只能对 SysContact (1.3.6.1.2.1.1.4.0) 节点下的 MIB 对象进行写操作。
- 网络设备能够主动向网管工作站发送验证加密的消息。

【配置方法】

- 第一步, 配置 MIB 视图和组。创建一个 MIB 视图 “view1” , 包含关联的 MIB 对象 (1.3.6.1.2.1.1) ; 再创建一个 MIB 视图 “view2” , 包含关联的 MIB 对象 (1.3.6.1.2.1.1.4.0)。创建一个组 “g1” , 选择版本号为 “v3” , 配置安全级别为认证加密模式 “priv” , 并可读视图 “view1” , 可写视图 “view2” 。
- 第二步, 配置 SNMP 用户。创建用户名 “user1” , 属于组 “g1” , 选择版本号为 “v3” , 配置认证方式为 “md5” , 认证密码为 “123” , 加密方式为 “DES56” , 加密密码为 “321” 。
- 第三步, 配置 SNMP 主机地址。配置主机地址为 192.168.3.2 , 选择版本号为 “3” , 配置安全级别为认证加密模式 “priv” , 关联对应的用户名 “user1” 。使能 Agent 主动向 NMS 发送 Trap 消息。
- 第四步, 配置 Agent 的 IP 地址。配置 Gi0/1 的接口地址为 192.168.3.1/24。

Agent

```

Ruijie(config)#snmp-server view view1 1.3.6.1.2.1.1 include
Ruijie(config)#snmp-server view view2 1.3.6.1.2.1.1.4.0 include
Ruijie(config)#snmp-server group g1 v3 priv read view1 write view2
Ruijie(config)#snmp-server user user1 g1 v3 auth md5 123 priv des56 321
Ruijie(config)#snmp-server host 192.168.3.2 traps version 3 priv user1
Ruijie(config)#snmp-server enable traps
Ruijie(config)#interface gigabitEthernet 0/1
Ruijie(config-if-gigabitEthernet 0/1)#ip address 192.168.3.1 255.255.255.0
Ruijie(config-if-gigabitEthernet 0/1)#exit
  
```

【检验方法】

- 第一步, 通过 **show running-config** 命令查看设备的配置信息。
- 第二步, 通过 **show snmp user** 命令查看 SNMP 用户。
- 第三步, 通过 **show snmp view** 命令查看 SNMP 视图。
- 第四步, 通过 **show snmp group** 命令查看 SNMP 组。
- 第五步, 通过 **show snmp host** 命令查看用户配置的主机信息。
- 第六步, 安装 MIB-Browser 查询。

Agent

```

Ruijie# show running-config
!
interface gigabitEthernet 0/1
  
```

```
no ip proxy-arp
ip address 192.168.3.1 255.255.255.0
!
snmp-server view view1 1.3.6.1.2.1.1 include
snmp-server view view2 1.3.6.1.2.1.1.4.0 include
snmp-server user user1 g1 v3 encrypted auth md5 7EBD6A1287D3548E4E52CF8349CBC93D priv des56
D5CEC4884360373ABBF30AB170E42D03
snmp-server group g1 v3 priv read view1 write view2
snmp-server host 192.168.3.2 traps version 3 priv user1
snmp-server enable traps
```

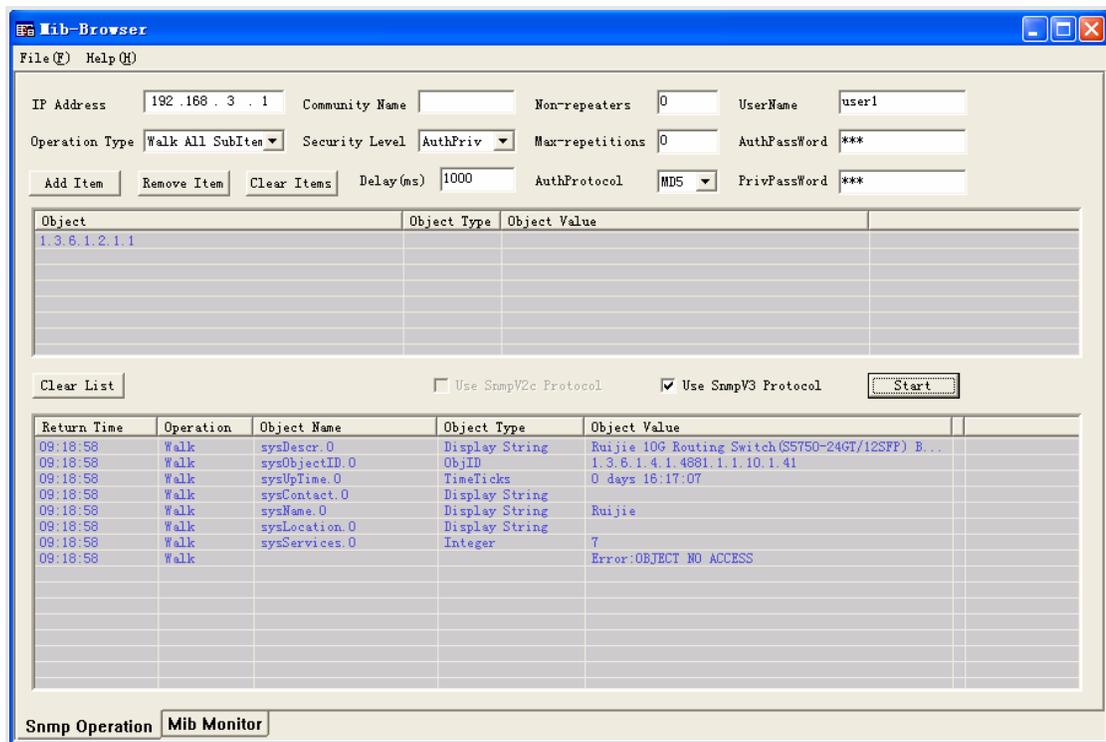
```
Ruijie# show snmp user
User name: user1
Engine ID: 800013110300d0f8221120
storage-type: permanent      active
Security level: auth priv
Auth protocol: MD5
Priv protocol: DES
Group-name: g1
```

```
Ruijie#show snmp view
view1(include) 1.3.6.1.2.1.1
view2(include) 1.3.6.1.2.1.1.4.0
default(include) 1.3.6.1
```

```
Ruijie# show snmp group
groupname: g1
securityModel: v3
securityLevel:authPriv
readview: view1
writeview: view2
notifyview:
```

```
Ruijie#show snmp host
Notification host: 192.168.3.2
udp-port: 162
type: trap
user: user1
security model: v3 authPriv
```

安装 MIB-Browser，在 IP Address 中输入设备的 IP 地址：192.168.3.1，在 UserName 中输入“user1”，在 Security Level 中选择“AuthPriv”，在 AuthPassWord 中输入“123”，在 AuthProtocol 中选择“MD5”，在 PrivPassWord 中输入“321”。点击 add item 按钮，选择要查询的 MIB 的具体管理单元，比如下图的 System。点击 Start 按钮，便开始对网络设备进行 MIB 的查询了，具体的查询结果见对话框的最下面的窗口：



常见错误

-

1.4.2 启用Trap功能

配置效果

使 Agent 主动向 NMS 发送 Trap 消息。

注意事项

-

配置方法

配置 snmp 主机地址

- 可选配置。
- 需要 Agent 主动发送消息时需要配置 NWS 的主机地址。

Agent 主动向 NMS 发送 Trap 消息

- 可选配置。
- 当需要 agent 主动向 NMS 发送 Trap 消息时，需在 agent 上配置此项。

📌 打开接口发送 Link Trap 功能

- 可选配置。
- 当需要接口发送 link trap 功能时，需在 agent 上配置接口打开此项。

📌 打开发送系统重启 Trap 功能

- 可选配置。
- 当希望 RGOS 系统在设备 reload/reboot 以前给 NMS 发送 Trap 消息通知系统重启时，需在 agent 上配置此项。

📌 指定发送 Trap 消息的源地址

- 可选配置。
- 当希望固定使用一个本地 IP 地址作为 SNMP 的源地址以便于管理时，需在 agent 上配置此项。

📌 发送 Trap 消息时携带私有字段

- 可选配置。
- 当需要 Trap 消息携带私有字段时，需在 agent 上配置此项。

检验方法

通过 `show snmp` 命令显示 SNMP 的状态信息。

通过 `show running-config` 命令查看设备的配置信息。

相关命令

📌 配置 NMS 主机地址

【命令格式】 `snmp-server host { host-addr } [traps | informs] [version { 1 | 2c | 3 { auth | noauth | priv }] community-string [udp-port port-num] [notification-type]`

【参数说明】 `host-addr`：SNMP 主机地址。
`traps | informs`：配置主机发送 trap 报文还是 inform 报文。
`Version`：选择 snmp 版本，V1、V2C、V3。
`auth | noauth | priv`：配置 V3 用户的安全级别。
`community-string`：团体字符串或用户名（V3 版本）。
`port-num`：配置 snmp 主机端口。
`notification-type`：主动发送的 Trap 类型，例如 snmp。

 如果没有指定 Trap 类型，则包括所有 Trap 类型。

【命令模式】 全局配置模式

- 【使用指导】 该命令与全局配置命令 **snmp-server enable traps** 一起使用，主动给 NMS 发送 Trap 消息。
可以配置多个不同的 SNMP 主机用于接收 Trap 消息，一个主机可以使用不同 Trap 类型组合，不同的端口，

配置 Agent 主动向 NMS 发送 Trap 消息

- 【命令格式】 **snmp-server enable traps** [*notification-type*]
- 【参数说明】 *notification-type*：启用对应事件的 Trap 通知，有以下类型：
snmp: 启动 SNMP 事件的 TRAP 通知；
bridge: 启动 BRIDGE 事件的 TRAP 通知；
mac-notification: 启动 MAC 事件的 TRAP 通知；
ospf: 启动 OSPF 事件的 TRAP 通知；
urpf: 启动 URPF 事件的 TRAP 通知；
vrrp: 启动 VRRP 事件的 TRAP 通知；
web-auth: 启动 WEB 认证事件的 TRAP 通知。
- 【命令模式】 全局配置模式
- 【使用指导】 该命令必须与全局配置命令 **snmp-server host** 一起使用，才能发送 Trap 消息。

打开接口发送 Link Trap 功能

- 【命令格式】 **snmp trap link-status**
- 【参数说明】 -
- 【配置模式】 接口配置模式
- 【使用指导】 对于接口（以太网接口、Ap 接口、SVI 接口），当功能打开时，如果接口发生 Link 状态变化，SNMP 将发出 Link Trap，反之则不发。

打开发送系统重启 Trap 功能

- 【命令格式】 **snmp-server system-shutdown**
- 【参数说明】 -
- 【配置模式】 全局配置模式
- 【使用指导】 打开 SNMP 系统重启通知功能，会在设备 **reload/reboot** 以前给 NMS 发送 Trap 消息通知系统重启。

指定发送 Trap 消息的源地址

- 【命令格式】 **snmp-server trap-source** *interface*
- 【参数说明】 *interface*：用于作为 SNMP 源地址的接口。
- 【配置模式】 全局配置模式
- 【使用指导】 缺省情况下，SNMP 报文从哪个接口出去，就使用哪个接口的 IP 地址作为源地址，为了便于管理和识别，可以使用该命令固定使用一个本地 IP 地址作为 SNMP 的源地址。

配置发送 Trap 消息时携带私有字段

- 【命令格式】 **snmp-server trap-format private**
- 【参数说明】 -
- 【配置模式】 全局配置模式
- 【使用指导】 使用该命令可配置发送 Trap 消息携带私有格式字段，包含的字段目前支持的有告警发生时间，各个字段的具

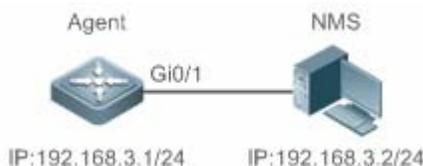
体数据类型和数据范围可参见 RUIJIE-TRAP-FORMAT-MIB.mib 文件说明。

配置举例

配置启用 trap 功能

【网络环境】

图 1-6



- 网管工作站（NMS）基于共同体认证模式对网络设备（Agent）进行管理，网络设备能够主动向网管工作站发送消息。

【配置方法】

- 第一步，配置 Agent 主动向 NMS 发送消息。配置 SNMP 主机地址为 192.168.3.2，消息格式为 Version 2c，认证名为“user1”。使能 Agent 主动发送 Trap 消息。
- 第二步，配置 Agent 的 IP 地址。配置 Gi 0/1 的接口地址为 192.168.3.1/24。

Agent

```
Ruijie(config)#snmp-server host 192.168.3.2 traps version 2c user1
Ruijie(config)#snmp-server enable traps
Ruijie(config)#interface gigabitEthernet 0/1
Ruijie(config-if-gigabitEthernet 0/1)#ip address 192.168.3.1 255.255.255.0
Ruijie(config-if-gigabitEthernet 0/1)#exit
```

【检验方法】

- 通过 **show running-config** 命令查看设备的配置信息。
- 通过 **show snmp** 命令显示 SNMP 的状态信息。

Agent

```
Ruijie# show running-config
ip access-list standard a1
 10 permit host 192.168.3.2
interface gigabitEthernet 0/1
 no ip proxy-arp
 ip address 192.168.3.1 255.255.255.0
snmp-server view v1 1.3.6.1.2.1.1 include
snmp-server location fuzhou
snmp-server host 192.168.3.2 traps version 2c user1
snmp-server enable traps
snmp-server contact ruijie.com.cn
snmp-server community user1 view v1 rw a1
snmp-server chassis-id 1234567890

Ruijie#show snmp
Chassis: 1234567890
0 SNMP packets input
```

```
0 Bad SNMP version errors
0 Unknown community name
0 Illegal operation for community name supplied
0 Encoding errors
0 Number of requested variables
0 Number of altered variables
0 Get-request PDUs
0 Get-next PDUs
0 Set-request PDUs
0 SNMP packets output
0 Too big errors (Maximum packet size 1472)
0 No such name errors
0 Bad values errors
0 General errors
0 Response PDUs
0 Trap PDUs
SNMP global trap: enabled
SNMP logging: disabled
SNMP agent: enabled
```

常见错误

-

1.4.3 无屏蔽Agent功能

配置效果

在不需要 Agent 服务的时候，屏蔽 Agent 功能。

注意事项

- 执行 **no snmp-server** 命令，可以在不需要代理服务的时候，屏蔽 SNMP 代理功能。
- 不同于屏蔽命令，执行 **no enable service snmp-agent** 命令，会直接关闭 snmp 所有服务（即 snmp 代理功能被禁用了，不接收报文、不发送响应报文及 trap），不会屏蔽代理的配置信息。

配置方法

📌 配置屏蔽设备 SNMP 代理

- 可选配置。
- 需要屏蔽所有 SNMP 代理服务配置时，可选用此项配置。

配置关闭设备 SNMP 代理

- 可选配置。
- 需要直接关闭所有服务时，应选用此配置项。

检验方法

通过 **show services** 命令查看 snmp 服务的开关状态信息。

通过 **show snmp** 命令显示 SNMP 的状态信息。

通过 **show running-config** 命令查看设备的配置信息。

相关命令

配置屏蔽设备 SNMP 代理功能

【命令格式】 **no snmp-server**

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 SNMP 代理功能服务默认关闭，在设置 SNMP 代理参数（例如 NMS 主机地址、认证名和访问权限等）时，会自动打开 SNMP 代理服务，服务开关命令 **enable service snmp-agent** 也必须同时打开，SNMP 代理服务才能生效，但只要关闭了其中的一个，SNMP 代理服务将不会生效。使用 **no snmp-server** 命令可以关闭设备支持的所有版本 SNMP 的代理服务。

使用该命令的同时，将屏蔽所有 SNMP 代理服务配置（即使用 **show running-config** 命令查看时不会显示配置，重新开启 SNMP 代理服务可以恢复），而 **enable service snmp-agent** 命令则不会屏蔽 SNMP 代理配置。

配置关闭设备 SNMP 代理功能

【命令格式】 **no enable service snmp-agent**

【参数说明】 -

【配置模式】 全局配置模式

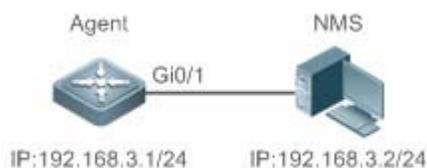
【使用指导】 关闭 SNMP 服务开关，但不会屏蔽 SNMP 代理参数。

配置举例

配置启用 snmp 服务功能

【网络环境】

图 1-7



通过设置 snmp 服务开关，以及设置 snmp 代理服务器，使得网管工作站（NMS）能通过 snmp 访问设备。

【配置方法】

- 配置启用 snmp 服务。
- 配置 snmp 代理服务器的参数，使服务生效。

A gent

```
Ruijie(config)#enable service snmp-agent
```

【检验方法】

- 通过 **show services** 命令查看 snmp 服务的开关状态信息。

Agent

```
Ruijie#show service
web-server      : disabled
web-server(https): disabled
snmp-agent      : enabled
ssh-server      : disabled
telnet-server   : enabled
```

常见错误

1.4.4 设置SNMP控制参数

配置效果

对 SNMP 的 Agent 的基本参数进行配置，包括设备的联系方式、设备位置、序列号、发送 Trap 消息的参数等，NMS 通过访问设备的这些参数，便可以得知设备的联系人，设备所在的物理位置等信息。

注意事项

配置方法

▾ 配置系统的联系方式

- 可选配置。

- 当需要修改系统的联系方式时，需在 agent 上配置此项。

↘ 配置系统位置

- 可选配置。
- 当需要修改系统的系统位置时，需在 agent 上配置此项。

↘ 配置系统序列码

- 可选配置。
- 当需要修改系统的序列码时，需在 agent 上配置此项。

↘ 配置设备的网元信息

- 可选配置。
- 当需要修改网元编码信息时，需在 agent 上配置此项。

↘ 配置 SNMP 代理最大数据报文长度

- 可选配置。
- 当需要修改 SNMP 代理最大数据报文长度时，需在 agent 上配置此项。

↘ 配置 SNMP 服务 UDP 端口号

- 可选配置。
- 当需要修改 SNMP 服务的 UDP 端口号时，需在 agent 上配置此项。

↘ 配置 Trap 消息报文的队列长度

- 可选配置。
- 当希望通过调整消息队列大小来控制消息发送速度时，需在 agent 上配置此项。

↘ 配置发送 Trap 消息的时间间隔

- 可选配置。
- 当需要修改发送 Trap 消息的时间间隔时，需在 agent 上配置此项。

检验方法

通过 `show snmp` 命令显示 SNMP 的状态信息。

通过 `show running-config` 命令查看设备的配置信息。

相关命令

↘ 配置系统的联系方式

- 【命令格式】 **snmp-server contact** *text*
- 【参数说明】 *text*：描述系统联系方式的字符串。
- 【命令模式】 全局配置模式
- 【使用指导】

配置系统位置

- 【命令格式】 **snmp-server location** *text*
- 【参数说明】 *text*：描述系统信息的字符串。
- 【配置模式】 全局配置模式
- 【使用指导】

配置系统序列码

- 【命令格式】 **snmp-server chassis-id** *text*
- 【参数说明】 *text*：系统序列号的文本，可以是数字或字符。
- 【配置模式】 全局配置模式
- 【使用指导】 SNMP 系统序列号一般使用机器的序列号，以便对设备进行识别。

配置设备的网元信息

- 【命令格式】 **snmp-server net-id** *text*
- 【参数说明】 *text*：设置设备网元编码 *text*，*text* 是长度为 1~255 的字符串，区分大小写，可包含空格。
- 【配置模式】 全局模式
- 【使用指导】 配置设备网元编码信息。

配置 SNMP 代理最大数据报文长度

- 【命令格式】 **snmp-server packet-size** *byte-count*
- 【参数说明】 *byte-count*：数据包大小，从 484 字节到 17876 字节。
- 【配置模式】 全局模式
- 【使用指导】

配置 SNMP 服务 UDP 端口号

- 【命令格式】 **snmp-server udp-port** *port-num*
- 【参数说明】 *port-num*：指定 SNMP 服务的 UDP 端口号，即接收 SNMP 报文的协议端口号。
- 【配置模式】 全局模式
- 【使用指导】 指定接收 SNMP 报文的协议端口号。

配置 Trap 消息报文的队列长度

- 【命令格式】 **snmp-server queue-length** *length*
- 【参数说明】 *length*：队列长度，大小从 1 到 1000。
- 【配置模式】 全局配置模式
- 【使用指导】 通过调整消息队列大小来控制消息发送速度。

配置发送 Trap 消息的时间间隔

- 【命令格式】 **snmp-server trap-timeout seconds**
- 【参数说明】 *seconds* : 间隔时间, 单位为秒, 取值范围: 1 – 1000。
- 【配置模式】 全局配置模式
- 【使用指导】 通过调整发送消息的时间间隔来控制消息发送速度。

配置举例

设置 SNMP 的控制参数

【网络环境】

图 1-8



- 网管工作站 (NMS) 基于共同体认证模式对网络设备 (Agent) 进行管理, 网管工作站能够获取设备的基本系统信息, 如系统的联系方式、位置、序列码。

【配置方法】

- 第一步, 配置 SNMP 代理参数。配置系统所处的位置、联系方式、序列码。
- 第二步, 配置 Agent 的 IP 地址。配置 Gi 0/1 的接口地址为 192.168.3.1/24。

Agent

```

Ruijie(config)#snmp-server location fuzhou
Ruijie(config)#snmp-server contact ruijie.com.cn
Ruijie(config)#snmp-server chassis-id 1234567890
Ruijie(config)#interface gigabitEthernet 0/1
Ruijie(config-if-gigabitEthernet 0/1)#ip address 192.168.3.1 255.255.255.0
Ruijie(config-if-gigabitEthernet 0/1)#exit
  
```

【检验方法】

- 第一步, 查看设备的配置信息。
- 第二步, 查看 SNMP 视图和组的信息。

Agent

```

Ruijie# show running-config
ip access-list standard al
 10 permit host 192.168.3.2
interface gigabitEthernet 0/1
 no ip proxy-arp
 ip address 192.168.3.1 255.255.255.0
snmp-server view v1 1.3.6.1.2.1.1 include
snmp-server location fuzhou
snmp-server host 192.168.3.2 traps version 2c user1
snmp-server enable traps
snmp-server contact ruijie.com.cn
snmp-server community user1 view v1 rw al
  
```

```
snmp-server chassis-id 1234567890
Ruijie#show snmp view
v1(include) 1.3.6.1.2.1.1
default(include) 1.3.6.1
Ruijie#show snmp group
groupname: user1
securityModel: v1
securityLevel:noAuthNoPriv
readview: v1
writeview: v1
notifyview:
groupname: user1
securityModel: v2c
securityLevel:noAuthNoPriv
readview: v1
writeview: v1
notifyview:
```

常见错误

-

1.5 监视与维护

查看运行情况

作用	命令
显示 SNMP 的状态信息	show snmp [mib user view group] host]

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
SNMP 调试开关	debug snmp

2 RMON

2.1 概述

RMON 全称是 Remote Network Monitoring，远端网络监控。

RMON 用来解决从一个中心点管理各局域分网和远程站点的问题。RMON 中，网络监视数据包含了一组统计数据 and 性能指标，这些数据可以用来监控网络利用率，以用于网络规划，性能优化和协助网络错误诊断。

RMON 适主要用于管理设备向被监控管理设备进行远程监控管理。

协议规范

STD 0059 / RFC 2819 : Remote Network Monitoring Management Information Base

RFC4502 : Remote Network Monitoring Management Information Base Version 2

RFC 3919 : Remote Network Monitoring (RMON) Protocol Identifiers for IPv6 and Multi Protocol Label Switching (MPLS)

RFC 3737 : IANA Guidelines for the Registry of Remote Monitoring (RMON) MIB Modules

RFC 3434 : Remote Monitoring MIB Extensions for High Capacity Alarms

RFC 3395 : Remote Network Monitoring MIB Protocol Identifier Reference Extensions

RFC 3287 : Remote Monitoring MIB Extensions for Differentiated Services

RFC 3273 : Remote Network Monitoring Management Information Base for High Capacity Networks

RFC 2896 : Remote Network Monitoring MIB Protocol Identifier Macros

RFC 2895 : Remote Network Monitoring MIB Protocol Identifier Reference

2.2 典型应用

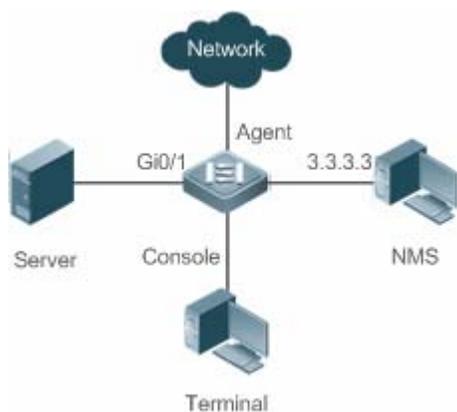
典型应用	场景描述
统计监控接口信息	应该 RMON 的四个功能于接口监管接口网络通信

2.2.1 统计监控接口信息

应用场景

用户通过 RMON 以太网统计功能监管接口的累计信息，通过历史统计功能监管接口每个监管间隔时间内的接口报文数信息，使用告警功能即时获知接口报文数异常情况。组网图如下所示：

图 2-1



功能部署

对接口 x 进行监管，分别累加统计接口报文数信息，统计接口监管时间间隔内的报文数信息以及带宽利用率，如果接口报文数异常，告警通知网管，配置要点如下：

- 在接口 x 下配置 RMON 以太网统计功能；
- 在接口 x 下配置 RMON 历史统计功能；
- 在配置模式下配置 RMON 告警表以及定义相应的 RMON 事件处理动作，告警监管的对象为接口 x 下配置的 RMON 以太网统计表的具体字段 OID 值。

2.3 功能详解

基本概念

RMON 定义了多个 RMON 组，我司产品支持其中的统计组、历史组、告警组、事件组。下面对四个组做简要的介绍：

统计组

统计组用于对以太网接口的流量信息进行监控、统计，是从创建表项起到当前阶段的累加值，统计的内容包括丢弃的数据包、广播数据包、CRC 错误、大小块、冲突等，统计结果将保存在以太网统计表中以便管理员随时查看。

历史组

历史组(History)用于定期收集网络流量信息，记录每一个周期内的网络流量信息的累加值以及带宽利用率，并保存在历史控制表中以便管理员日后处理，它包含两个小组：

- HistoryControl 组用来设置采样间隔时间、采样数据源等控制信息。
- EthernetHistory 组为管理员提供有关网段流量、错误包、广播包、利用率以及碰撞次数等统计信息的历史数据。

告警组

警报组(Alarm)用于监控指定的 MIB(Management Information Base, 管理信息库)对象，当这个 MIB 对象的值超过设定的上限值或低于设定的下限值时，会触发警报，警报被当作事件来处理。

事件组

事件组（Event）用于定义事件的处理方式。当监控的 MIB 对象达到告警条件时，就会触发事件，事件有如下四种处理方式：

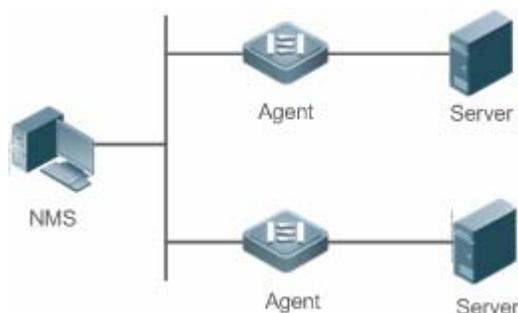
- none：不做任何动作。
- log：将事件相关信息记录在日志记录表中，以便管理员随时查看。
- snmp-trap：向网管站发送 Trap 消息告知该事件的发生。
- log-and-trap：将事件相关信息记录在日志记录表中，同时向网管站发送 Trap 消息。

工作原理

RMON 允许有多个监控者，可以通过两种方法进行数据收集：一种方法是利用专用的 RMON probe（RMON 探测仪）收集数据，NMS（网络管理系统）可以直接从 RMON probe 获取 RMON MIB 的全部信息。另一种方法是将 RMON Agent 植入网络设备，使设备具备 RMON probe 功能。NMS 用 SNMP 的基本命令与其交换数据信息，收集网络管理信息，但这种方式受设备资源的限制，一般不能获取 RMON MIB 的所有数据，一般只收集四个组信息。

下图给出了 NMS 与 RMON 代理通信的例子。通过运行在设备上的 RMON Agent，NMS 可以获取与被管网络设备接口的网段上的整体流量、错误统计和性能统计等信息，从而实现对该网络设备的远程管理。

图 2-2



功能特性

功能特性	作用
RMON以太网统计功能	对监控的以太网接口报文数、字节数等数据进行累加统计。
RMON历史统计功能	记录以太多接口在配置的间隔时间内通信的报文数、字节数等数据进行统计，并计算间隔时间内的带宽利用率。。
RMON告警功能	告警表与事件表结合使用，间隔对监控的变量的值进行采样，触及上下限就触发相关事件表做事件处理，或者不做任何处理。

2.3.1 RMON以太网统计功能

工作原理

累加统计从创建表项起到现阶段的以太网接口的网络流量信息。

相关配置

配置 RMON 统计项

- 缺省情况下，RMON 以太网统计功能关闭。
- 使用 **rmon collection stats** 命令在指定的以太网接口上创建以太网统计表项。
- 在指定接口下创建统计表项成功后，统计组就对当前接口的流量信息进行统计，它统计的是 RMON 以太网统计表定义的变量，记录的是 RMON 统计表创建起至当前阶段时间内变量的累加值。

2.3.2 RMON历史统计功能

工作原理

记录每一个周期内的以太网接口流量信息的累加统计值。

相关配置

配置 RMON 历史控制表项

- 缺省情况下，配置 RMON 历史统计功能关闭。
- 使用 **rmon collection history** 命令在以太网接口上创建历史控制表项。
- RMON 历史组统计的是 RMON 历史表定义的变量，记录的是每个周期内变量的累加值。

2.3.3 RMON告警功能

工作原理

周期性地监控告警变量的值变化，如果告警变量值触及指定的上下限值，则触发相应的事件处理，如发送 trap 信息，或者产生一条 logTable 表项记录等。但连续多次触有上限阈值或者下限阈值，只触发相应的事件处理一次，等待触发相反阈值处理。

相关配置

配置事件表

- 缺省情况下，配置 RMON 事件组功能关闭。
- 使用 **rmon event** 命令配置事件表。

配置告警表项

- 缺省情况下，配置 RMON 告警组功能关闭。
- 使用 `rmon event` 命令配置事件表、`rmon alarm` 命令配置 RMON 告警表。
- RMON 告警功能是由告警表和事件表共同实现。如果告警事件需要向管理设备发送 Trap 信息的话，则必须事先保证 SNMP Agent 已经正确配置。SNMP Agent 的配置请参见《SNMP 配置指南》。
- 如果配置的告警对象是 RMON 统计组或者历史组的某一字段节点，需要先在被监控的以太网接口下配置 RMON 统计功能或者 RMON 历史统计功能。

2.4 产品说明



我司产品当前版本只支持以太网接口的统计。



我司产品当前版本只支持以太网接口的记录

2.5 配置详解

配置项	配置建议 & 相关命令	
配置RMON以太网统计功能	⚠ 必须配置。用于累加统计以太网接口流量信息。	
	<code>rmon collection stats</code>	配置以太网统计表项。
配置RMON历史统计功能	⚠ 必须配置。用于间隔统计间隔时间内的以太网接口流量信息以及带宽使用率。	
	<code>rmon collection history</code>	配置历史控制表项。
配置RMON告警功能	⚠ 必须配置。用于监测某一变量的数据变化是否在合法范围内。	
	<code>rmon event</code>	配置事件表项。
	<code>rmon alarm</code>	配置告警表项。

2.5.1 配置RMON以太网统计功能

配置效果

可以获知被监控的以太网接口从创建表项起到现阶段的流量信息的累加统计值。

注意事项

不允许批量接口配置，即不允许在批量接口配置模式下进行配置。

配置方法

配置 RMON 统计项

- 必选配置。
- 如果需要对指定接口进行统计、监控，必须在该接口下配置以太网统计表项。

检验方法

使用 `show rmon stats` 命令可以查看以太网统计信息。

相关命令

配置 RMON 统计项

【命令格式】 `rmon collection stats index [owner ownername]`

【参数说明】 `index`：统计表项的索引号，取值范围：1~65535；

`owner ownername`：设置表项的创建者 `ownername`，`ownername` 为 1~63 个字符的字符串，区分大小写。

【命令模式】 接口模式

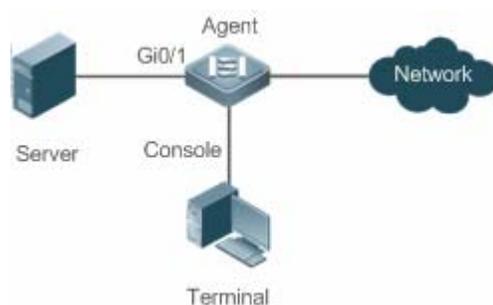
【使用指导】 不允许对已经配置的统计表项参数进行修改。

配置举例

配置 RMON 以太网统计功能

【网络环境】

图 2-3



如上图所示，RMON Agent 与 Server 服务器连接，网管需要通过 RMON 统计组来对 G0/1 接口的接收报文进行性能统计，以便随时通过查看数据了解相应接口接收报文的数据，及时对异常网络情况采取措施处理。

【配置方法】 ● 在接口 GigabitEthernet 0/3 上配置统计表实例，对该接口进行流量统计。

Agent

```
Ruijie# configure terminal
Ruijie (config)# interface gigabitEthernet 0/3
Ruijie (config-if-GigabitEthernet 0/3)# rmon collection stats 1 owner admin
```

【检验方法】 通过 **show rmon stats** 查看以太网统计信息。

Agent

```
Ruijie# show rmon stats
ether statistic table:
    index = 1
    interface = GigabitEthernet 0/1
    owner = admin
    status = 1
    dropEvents = 0
    octets = 25696
    pkts = 293
    broadcastPkts = 3
    multiPkts = 0
    crcAlignErrors = 0
    underSizePkts = 0
    overSizePkts = 0
    fragments = 0
    jabbers = 0
    collisions = 0
    packets64Octets = 3815
    packets65To127Octets = 1695
    packets128To255Octets = 365
    packets256To511Octets = 2542
    packets512To1023Octets = 152
    packets1024To15180Octets = 685
```

常见错误

重复配置或者修改已经配置的统计表表项。

2.5.2 配置RMON历史统计功能

配置效果

可以获知被监控的以太网接口的每一个周期内的流量信息累加统计值及带宽利用率。

注意事项

不允许批量接口配置，即不允许在批量接口配置模式下进行配置。

配置方法

- 必选配置。
- 如果需要对指定接口收集网络统计信息，必须在接口上配置 RMON 历史控制表项。

检验方法

使用 `show rmon history` 命令可以查看历史组统计信息。

相关命令

配置 RMON 历史控制表项

【命令格式】 `rmon collection history index [owner ownername] [buckets bucket-number] [interval seconds]`

【参数说明】 `index`：历史统计表项的索引号，取值范围：1~65535

`owner ownername`：设置表项的创建者 `ownername`，`ownername` 为 1~63 个字符的字符串，区分大小写。

`buckets bucket-number`：设置历史统计表项对应的历史表容量，即设置历史表最多可容纳的记录数 `bucket-number`，`bucket-number` 取值范围：1~65535，默认值是 10

`interval seconds`：设置统计周期值 `seconds`，单位为秒，`seconds` 取值范围：1~3600，默认值是 1800s

【命令模式】 接口模式

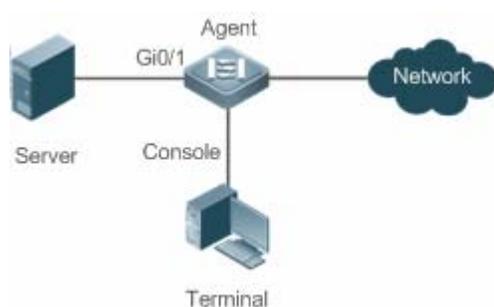
【使用指导】 不允许对已经配置的历史统计表项参数进行修改。

配置举例

配置 RMON 历史统计功能

【网络环境】

图 2-4



如上图所示，RMON Agent 与 Server 服务器连接，网管需要通过 RMON 历史组来对 G0/1 接口的接收报文进行周期性统计，周期时间为 60 秒，从而达到对网络的监控，掌握突发情况数据。

【配置方法】 ● 在接口 GigabitEthernet 0/3 上配置历史控制表，对该接口进行周期性流量统计

Agent

```
Ruijie# configure terminal
Ruijie(config)# interface gigabitEthernet 0/3
Ruijie(config-if-GigabitEthernet 0/3)# rmon collection history 1 buckets 5 interval 300 owner admin
```

【检验方法】 通过 **show rmon history** 查看历史组统计信息。

Agent

```
Ruijie# show rmon history
rmon history control table:
    index = 1
    interface = GigabitEthernet 0/1
    bucketsRequested = 5
    bucketsGranted = 5
    interval = 60
    owner = admin
    stats = 1

rmon history table:
    index = 1
    sampleIndex = 786
    intervalStart = 6d:18h:37m:38s
    dropEvents = 0
    octets = 2040
    pkts = 13
    broadcastPkts = 0
    multiPkts = 0
    crcAlignErrors = 0
    underSizePkts = 0
    overSizePkts = 0
    fragments = 0
    jabbers = 0
    collisions = 0
    utilization = 0

    index = 1
    sampleIndex = 787
    intervalStart = 6d:18h:38m:38s
    dropEvents = 0
    octets = 1791
    pkts = 16
    broadcastPkts = 1
    multiPkts = 0
    crcAlignErrors = 0
    underSizePkts = 0
    overSizePkts = 0
    fragments = 0
```

```
jabbers = 0
collisions = 0
utilization = 0

index = 1
sampleIndex = 788
intervalStart = 6d:18h:39m:38s
dropEvents = 0
octets = 432
pkts = 6
broadcastPkts = 0
multiPkts = 0
crcAlignErrors = 0
underSizePkts = 0
overSizePkts = 0
fragments = 0
jabbers = 0
collisions = 0
utilization = 0

index = 1
sampleIndex = 789
intervalStart = 6d:18h:40m:38s
dropEvents = 0
octets = 432
pkts = 6
broadcastPkts = 0
multiPkts = 0
crcAlignErrors = 0
underSizePkts = 0
overSizePkts = 0
fragments = 0
jabbers = 0
collisions = 0
utilization = 0

index = 1
sampleIndex = 790
intervalStart = 6d:18h:41m:38s
dropEvents = 0
octets = 86734
```

```
pkts = 934
broadcastPkts = 32
multiPkts = 23
crcAlignErrors = 0
underSizePkts = 0
overSizePkts = 0
fragments = 0
jabbers = 0
collisions = 0
utilization = 0
```

常见错误

重复配置或者修改已经配置的历史控制表表项。

2.5.3 配置RMON告警功能

配置效果

周期性监控告警变量的值变化是否在指定的合法范围内。

注意事项

如果触发告警事件时，需要向管理设备发送 Trap 信息的话，必须保证 SNMP Agent 已经正确配置。SNMP Agent 配置请参见 SNMP 配置指南。

如果告警变量是 RMON 统计组或者是历史组中定义的 MIB 变量时，必须在被监控的以太网接口上配置 RMON 以太网统计功能或者 RMON 历史统计功能，否则创建告警表失败。

配置方法

▾ 配置事件表项

- 必须配置。
- 在全局配置模式下配置

▾ 配置告警表项

- 必须配置。
- 在全局配置模式下配置

检验方法

- 使用 `show rmon event` 查看事件表信息。
- 使用 `show rmon alarm` 查看告警表信息。

相关命令

配置事件表

【命令格式】 `rmon event number [log] [trap community] [description description-string] [owner ownername]`

【参数说明】 `number` : 事件表的索引号, 取值范围: 1~65535。

`log` : 日志事件, 当事件被触发时, 系统会记录日志。

`trap community` : Trap 事件, 当事件被触发时, 系统会以 `community` 为团体名发送 Trap。

`description description-string` : 设置事件的描述信息 `description-string`, `description-string` 为 1~127 个字符的字符串。

`owner ownername` : 设置表项创建者 `ownername`, `ownername` 为 1~63 个字符的字符串, 区分大小写。

【命令模式】 全局配置模式

【使用指导】 允许对已经配置的事件表项参数进行修改, 包括事件类型、Trap 团体名、事件描述、事件创建者等。

配置 RMON 告警组

【命令格式】 `rmon alarm number variable interval {absolute | delta} rising-threshold value [event-number] falling-threshold value [event-number] [owner ownername]`

【参数说明】 `number` : 告警表项的索引号, 限值范围: 1~65535。

`variable` : 告警变量, 为 1~255 个字符的字符串, 并且以节点 OID 的点分格式(格式为 `entry.integer.instance`, 如 1.3.6.1.2.1.2.1.10.1) 进行表示。

`Interval` : 采样间隔时间, 单位为秒, 取值范围为 1 ~ 2147483647。

`absolute` : 采样类型为绝对值采样, 即采样时间到达时直接提取变量的值。

`delta` : 采样类型为变化值采样, 即采样时间到达时提取的是变量在采样间隔内的变化值。

`rising-threshold value` : 设置采样数量的上限参数 `value`, 取值范围: -2147483648~+2147483647。

`event-number` : 到达上下限时触发事件号为 `event-number` 的事件。

`falling-threshold value` : 设置采样数量的下限参数 `value`, 取值范围: -2147483648~+2147483647。

`owner ownername` : 设置表项的创建者 `ownername`, `ownername` 为 1~63 个字符的字符串, 区分大小写。

【命令模式】 全局配置模式

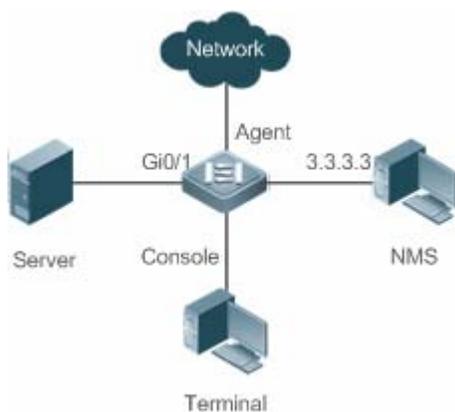
【使用指导】 允许对已经配置的告警表项参数进行修改, 包括告警变量、采样类型、表项的创建者、采样间隔时间、采样数量的上/下限值及其对应的触发事件。

配置举例

配置 RMON 告警功能

【网络环境】

图 2-5



假设 NMS 上运行 SNMPV1，访问设置时使用的团体名为 public，属性为可读写，NMS 接收 trap 的 IP 地址为 3.3.3.3。

假设监控接口 GigabitEthernet0/3 上接收到的未知协议的报文数，对应的 OID 值是 1.3.6.1.2.1.2.2.1.15.3，采样方式为相对采样，采样间隔时间为 60 秒，当相对采样值超过 100 时或者低于 10 时，分别触发事件 1 和事件 2，事件 1 发 trap 信息和 log 信息，事件 2 只生成日志记录表。

RMON Agent 通过终端 terminal 完成相关配置，与 NMS 设备连接通信，Gi0/1 跟服务器 Server 连接，现需要监控 Gi0/1 上收到未知协议的报文数。采样间隔时间 60 秒，绝对采样值小于 10 时，只记录 log，而大于 100 时，则需要记录 log 和发送 trap 给 NMS。

【配置方法】

- 配置 SNMP 主机接收告警功能发送 trap。
- 配置事件组动作来处理告警触发情况。
- 配置告警功能。

Agent

```

Ruijie# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Ruijie(config)# snmp-server community public rw
Ruijie(config)# snmp-server host 3.3.3.3 trap public
Ruijie(config)# rmon event 1 description rising-threshold-event log trap public owner admin
Ruijie(config)# rmon event 2 description falling-threshold-event log owner admin
Ruijie(config)# rmon alarm 1 1.3.6.1.2.1.2.2.1.15.3 60 delta rising-threshold 100 1
falling-threshold 10 2 owner admin
  
```

【检验方法】

- 使用 **show rmon event** 查看事件表信息。
- 使用 **show rmon alarm** 查看告警表信息。

Agent

```

Ruijie# show rmon event
rmon event table:

    index = 1
    description = rising-threshold-event
    type = 4
    community = public
    lastTimeSent = 0d:0h:0m:0s
    owner = admin
  
```

```
        status = 1

        index = 2
        description = falling-threshold-event
        type = 2
        community =
        lastTimeSent = 6d:19h:21m:48s
        owner = admin
        status = 1

rmon log table:
        eventIndex = 2
        index = 1
        logTime = 6d:19h:21m:48s
        logDescription = falling-threshold-event

Ruijie# show rmon alarm
rmon alarm table:
        index: 1,
        interval: 60,
        oid = 1.3.6.1.2.1.2.2.1.15.3
        sampleType: 2,
        alarmValue: 0,
        startupAlarm: 3,
        risingThreshold: 100,
        fallingThreshold: 10,
        risingEventIndex: 1,
        fallingEventIndex: 2,
        owner: admin,
        stauts: 1
```

常见错误

- 输入监控的对象 OID 不合理，OID 对应的变量不存在或者类型不是整型或者无符号整型。
- 上限阈值小于等于下限阈值。

2.6 监视与维护

查看运行情况

作用	命令
----	----

查看所有 RMON 配置信息	show rmon
查看以太网统计表信息	show rmon stats
查看历史控制表信息	show rmon history
查看告警表信息	show rmon alarm
查看事件表信息	show rmon event

3 NTP

3.1 概述

NTP (Network Time Protocol , 网络时间协议) , 用来使网络设备时间同步化的一种应用层协议。它可以使网络设备对其服务器或时钟源做同步化, 提供高精度度的时间校正 (LAN 上与标准时间差小于 1 毫秒, WAN 上几十毫秒) , 且可使用加密确认的方式来防止攻击。

目前我司设备支持 NTP 的客户端与服务器功能, 即设备既可以从时间服务器上同步时间, 也能够作为时间服务器对其他设备进行时间同步。在作为服务器工作时设备仅支持单播 Server 模式。

协议规范

- RFC 1305 : Network Time Protocol (Version 3)

3.2 典型应用

典型应用	场景描述
基于外部时钟参考源同步时间	设备即作为客户端从外部时钟源同步时间, 同步成功后又作为服务器向其他设备提供时间同步服务。
基于本地时钟参考源同步时间	设备将本地时钟作为 NTP 可靠参考时钟源, 作为服务器向其它设备提供时间同步服务。

3.2.1 基于外部时钟参考源同步时间

应用场景

如图所示：

- DEVICE-A 作为可靠参考时钟源对外提供时间同步服务
- DEVICE-B 指定 DEVICE-A 为 NTP 服务器, 从 DEVICE-A 同步时间。
- DEVICE-B 同步成功后向 DEVICE-C 提供时间同步服务。

图 3-1



功能部属

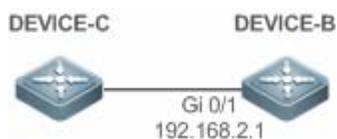
将 DEVICE-B 配置为 NTP 外部时钟参考模式

3.2.2 基于本地时钟参考源同步时间

应用场景

如图所示，DEVICE-B 将本地时钟作为 NTP 参考时钟源，向 DEVICE-C 提供时间同步服务。

图 3-2



功能部属

将 DEVICE-B 配置为 NTP 本地时钟参考模式。

3.3 功能详解

基本概念

↘ NTP 报文

根据 RFC1305 定义，NTP 采用 UDP 报文进行传输，UDP 端口号为 123。

NTP 时间同步报文格式如 图 3-3

图 3-3 NTP 时间同步报文格式

0	7	15	23	31	
LI	VN	Mode	Stratum	Poll Interval	Precision
Root Delay (32-bit)					
Root Dispersion (32-bit)					
Reference Clock Identifier (32-bit)					
Reference Timestamp (64-bit)					
Originate Timestamp (64-bit)					
Receive Timestamp (64-bit)					
Transmit Timestamp (64-bit)					
Authenticator (optional 96-bit)					

- Leap Indicator (LI) : 2 比特, 闰秒标志。

i 00-无警告信息 01-上一分钟有 61 秒 10-上一分钟有 59 秒 11-时钟未同步

- Version Number (VN) : 3 比特, NTP 版本号, 当前版本号为 3。
- Mode : 3 比特, NTP 工作模式。

i 0-未定义 1-主动对等体 2-被动对等体 3-客户端 4-服务器 5-广播 6-控制信息 7-保留

- Stratum : 8 比特, 本地时钟的层数 (0-未定义 1-主参考时钟源 其它值-次参考时钟源)。
- Poll Interval : 8 位整数, 轮询时间 (秒数)
- Precision : 8 位整数, 本地时钟的时间精度 (秒数)
- Root Delay : 32 位整数, 到主参考时钟源的往返时间
- Root Dispersion : 32 位整数, 相对于主参考时钟源的最大误差
- Reference Clock Identifier : 32 比特, 参考时钟源的标识
- Reference Timestamp : 64 位时间戳, 最后一次被设置或者被校正的时间
- Originate Timestamp : 64 位时间戳, 时间同步请求报文离开客户端的本地时间
- Receive Timestamp : 64 位时间戳, 时间同步请求报文到达服务器的本地时间
- Transmit Timestamp : 64 位时间戳, 时间同步响应报文离开服务器的本地时间
- Authenticator (可选) : 验证信息

📌 NTP 服务器

设备将本地时钟作为参考时钟源, 为网络中的其它设备提供时间同步服务。

📌 NTP 客户端

设备作为 NTP 客户端从网络中的 NTP 服务器同步时间。

层数 (stratum)

NTP 使用“层数 (stratum)”的概念来描述设备距离权威时钟源的“跳数 (hops)”。一个层数为 1 的时间服务器应当有个直连的原子钟或电波钟；层数为 2 的时间服务器就从层数为 1 的服务器获取时间；层数为 3 的服务器就从层数为 2 的获取时间.....如此递推。因此时钟层数数值更低的时钟源即被认为拥有更高的时钟精度。

硬件时钟

硬件时钟根据设备上的石英晶体振荡器频率工作，由设备的电池为其供电，设备关机后硬件时钟依然运行。在设备启动运行后，会从硬件时钟读取时间信息，作为设备的软件时间。

功能特性

功能特性	作用
NTP 时间同步	使网络设备根据其服务器或可靠时钟源进行时间同步，以实现高精度的时间校正。
NTP 安全认证	通过 NTP 报文加密认证方式，防止非可靠时钟源对设备进行时间同步干扰。
NTP 访问控制	根据访问控制列表对收到的 NTP 报文进行源过滤。

3.3.1 NTP时间同步

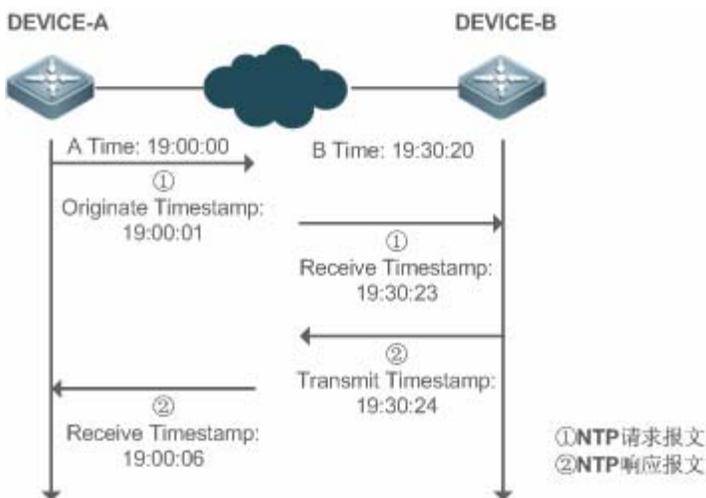
工作原理

NTP 同步时间的方式是通过客户端与服务器之间交互 NTP 报文：

- 客户端每隔 64 秒钟向所有服务器发送时间同步报文。收到服务器响应报文后，对所有服务器的响应报文进行过滤和选择，最后和优选服务器的时间进行同步。
- 服务器收时间同步请求报文时，将本地时钟作为参考源，按协议要求将本地时间信息填充到响应报文返回给客户端。

NTP时间同步报文格式如图 3-4

图 3-4 NTP 基本工作原理图



DEVICE-B (下面简称 B) 作为 NTP 参考时钟源, DEVICE-A (下面简称 A) 作为 NTP 客户端从 DEVICE-B 同步时间, 在某时刻 A 的本地时钟为 19:00:00, B 的本地时钟为 19:30:20:

4. A 发出 NTP 请求报文, 报文离开 A 的本地时间 (T0) 为 19:00:00, 填充在 Originate Timestamp
5. 经过 2 秒的网络延时, B 收到请求报文的本地时间 (T1) 为 19:30:23, 填充在 Receive Timestamp
6. B 处理 NTP 请求, 1 秒后响应 NTP 报文, 报文离开 B 的本地时间 (T2) 为 19:30:24, 填充在 Transmit Timestamp
7. 经过 2 秒的网络延时, A 接收到响应报文, 响应报文到达 A 的本地时间 (T3) 为 19:00:06

时间同步的具体算法如下:

- A 通过公式 $((T1-T0)+(T2-T3))/2$ 计算出 B 和 A 的时间差为 30 分 20 秒
- A 通过公式 $(T3-T0)-(T2-T1)$ 计算出 A 和 B 的报文往返的延时为 4 秒

▾ NTP 工作模式

- 外部时钟参考模式

在该模式下, 设备即充当服务器又充当客户端, 如果收到来自其它客户端发出的时间同步请求, 必须先从指定服务器同步时间, 同步成功后才可以向其它客户端提供时间同步服务。

- 本地时钟参考模式

在该模式下, 设备默认本地时钟即为可靠时钟源, 直接向其它客户提供时间同步服务。

相关配置

▾ 配置 NTP 服务器

- 缺省情况下, NTP 功能关闭。
- 通过 `ntp server` 命令指定 NTP 服务器 (即外部时钟参考源), 即可开启 NTP 功能。
- 配置后设备处于外部时钟参考模式。

▾ 实时同步

- 缺省情况下, 设备每隔 64 秒进行一次时间同步。

▾ 更新硬件时钟

- 缺省情况下, 设备同步完时间后不会把时间更新到硬件时钟。
- 配置 `ntp update-calendar` 命令可以使设备每次时间同步成功时会自动更新硬件时钟。

▾ 配置 NTP 主时钟

- 缺省情况下, 设备处于外部时钟参考模式。
- 通过 `ntp master` 命令可以将设备配置为本地时钟参考模式。

3.3.2 NTP安全认证

为防止对时间服务器的恶意破坏，NTP 使用了识别(Authentication)机制，检查时间同步信息是否是真正来自所宣称的服务器并检查资料的返回路径，以提供对抗干扰的保护机制。

工作原理

NTP 客户端和服务器配置相同的密钥。发送请求报文和响应报文时，设备根据指定的密钥和 NTP 报文内容采用 MD5 算法计算出报文的哈希值填充到报文的认证信息。接收设备根据认证信息判断是否报文发送端是否可信的设备或者报文是否被篡改。

相关配置

配置 NTP 全局安全认证机制

- 缺省情况下，没有开启 NTP 安全认证机制。
- 通过 `ntp authenticate` 命令可开启 NTP 安全认证机制。

配置 NTP 全局认证密钥

- 缺省情况下，没有配置全局认证密钥。
- 通过 `ntp authentication-key` 命令可开启 NTP 安全认证机制。

配置 NTP 全局信任密钥 ID

- 缺省情况下，没有配置全局信任密钥。
- 通过 `ntp trusted-key` 命令设备作为参考时钟源对外提供时间同步服务的信任密钥。

配置外部参考时钟源的信任密钥 ID

- 通过 `ntp server` 指定外部参考时钟源的同时可以指定该时钟源的信任密钥。

3.3.3 NTP访问控制

工作原理

通过 ACL 提供了一种最小限度的安全措施

相关配置

配置 NTP 服务的访问控制权限

- 缺省情况下，没有 NTP 访问控制权限。

- 通过 `ntp access-group` 可配置 NTP 的访问控制权限。

3.4 产品说明



锐捷目前的版本只支持最大 1024 认证密钥，每个服务器允许设置唯一一个密钥进行安全通信。

3.5 配置详解

配置项	配置建议&相关命令
配置NTP基本功能	必须配置，用于开启 NTP 功能，开启后设备处于外部时钟参考模式。
	<code>ntp server</code> 配置 NTP 服务器
	<code>ntp update-calendar</code> 自动更新硬件时钟
	可选配置，用于将设备配置为本地时钟参考模式。
	<code>ntp master</code> 配置 NTP 主时钟
	可选配置，用于关闭 NTP 功能。
	<code>no ntp</code> 关闭所有 NTP 功能，清空 NTP 配置。
配置NTP安全认证	<code>ntp disable</code> 禁止接收指定接口的 NTP 报文
	可选配置，用于防止非可靠时钟源对设备进行时间同步干扰。
	<code>ntp authenticate</code> 开启安全认证机制
	<code>ntp authentication-key</code> 设置安全认证全局密钥
	<code>ntp trusted-key</code> 配置时间同步服务可信密钥
配置NTP访问控制	<code>ntp server</code> 配置外部参考时钟源的可信密钥
	可选配置，用于对收到的 NTP 报文进行源过滤。
	<code>ntp access-group</code> 设置 NTP 的访问控制权限

3.5.1 配置NTP基本功能

配置效果

外部时钟参考模式

- 设备作为客户端，从外部参考时钟源同步时间到本地时钟
- 时间同步成功后，设备可作为时间同步服务器，对外提供时间同步服务

本地时钟参考模式

- 设备的本地时钟作为 NTP 参考时钟源，对外提供时间同步服务

注意事项

- 客户端/服务器模式，设备只有从外部的可靠时钟源同步成功后，才能作为时间同步服务器对外提供服务。
- 一旦配置本地时钟参考模式，系统便不会与比其时钟层数数值更高的时钟源进行同步。
- 将本地时钟设置为主时钟（尤其是指定了较低的时钟层数数值时）很有可能将真正有效时钟源覆盖。如果对同一网络中的多个设备都使用了该命令，则可能由于设备之间的时钟差异导致网络的时钟同步不稳定。
- 将本地时钟设置为主时钟前，如果系统从未与外部时钟源同步过，则有可能需要手动校准系统时钟以保证其不会有过大的偏差（关于如何手动校准系统时钟请参考配置指南中的系统时间配置部分）。

配置方法

▾ 配置 NTP 服务器

- 必须配置，至少指定一个外部参考时钟源（最多可配置 20 个不同的外部参考时钟源）。
- 如果需要关联配置 NTP 密钥，在配置 NTP 服务器前，必须先配置 NTP 安全认证。

▾ 自动更新硬件时钟

- 可选配置
- 默认情况下，时间同步成功后只更新系统时钟，不会更新硬件时钟。
- 配置此命令，时间同步成功后会自动更新硬件时钟。

▾ 配置 NTP 主时钟

- 如果需要将设备切换到本地时钟参考模式，可通过此命令。

▾ 关闭 NTP 功能

- 如果需要关闭 NTP 功能，并且清空 NTP 配置，可通过 **no ntp** 命令
- 默认情况，开启 NTP 功能后所有接口都可以接收 NTP 报文。如果需要禁止特定接口的 NTP 功能时可通过 **ntp disable** 命令。

检验方法

- 通过 **show ntp status** 查看 NTP 配置信息。
- 通过 **show clock** 查看是否完成时间同步

相关命令

▾ 配置 NTP 服务器

【命令格式】 **ntp server** [oob] { *ip-addr* | *domain* | **ip domain** } [**version** *version*] [**source** *if-name*] [**key** *keyid*] [**prefer**] [**via** *mgmt-name*]

【参数说明】 oob : 参考时钟源是否绑定 MGMT 口

ip-addr : 参考时钟源的 IPv4 地址

domain : 参考时钟源的 IPv4 域名

version : NTP 版本号, 取值为 1-3。

if-name : 接口类型, 包括 AggregatePort、Dialer、GigabitEthernet、Loopback、Multilink、Null、Tunnel、Virtual-ppp、Virtual-template、Vlan 类型。

keyid : 同参考时钟源通信采用的密钥(1-4294967295)

prefer : 参考时钟源是否高优先级

mgmt-name : 指定在 oob 模式下报文的出口管理口。

【命令模式】 全局模式

【使用指导】 在缺省情况下, 没有配置 NTP 服务器。锐捷的客户端系统支持最多同时与 20 个 NTP 服务器交互, (在全局认证以及密钥相关设置完成后) 可以为每一个服务器设置一个认证密钥, 发起与服务器的加密通信。

 如果需要设置认证密钥, 在配置 NTP 服务器前必须先配置 NTP 安全认证。

与服务器的默认通信版本为 NTP 版本 3, 同时可以配置发送 NTP 报文的源接口, 并只在发送接口上接收对应服务器的 NTP 报文。

更新硬件时钟

【命令格式】 **ntp update-calendar**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 -

设置本地参考时钟源

【命令格式】 **ntp master** [*stratum*]

【参数说明】 *stratum* : 指定本地时钟所处的层数, 范围为 1 ~ 15 ; 若不指定该参数则默认值为 8。

【命令模式】 全局模式

【使用指导】 -

关闭 NTP 功能

【命令格式】 **no ntp**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 此命令可以快速关闭 NTP 所有功能, 并且清空 NTP 所有配置

禁止接口接收 NTP 报文

【命令格式】 **ntp disable**

【参数说明】 -

【命令模式】 接口模式

【使用指导】 -

配置举例

▾ NTP 外部时钟参考模式

【网络环境】

图 3-5



- DEVICE-B：配置为 NTP 外部时钟参考模式
- DEVICE-A：作为 DEVICE-B 的参考时钟源
- DEVICE-C：从 DEVICE-B 同步时间

【配置方法】

- DEVICE-A 配置本地时钟为 NTP 参考时钟源
- DEVICE-B 配置 DEVICE-A 为参考时钟源
- DEVICE-C 配置 DEVICE-B 为参考时钟源

DEVICE-A

```
A#configure terminal
A(config)# ntp master
A(config)#exit
```

DEVICE-B

```
B#configure terminal
B(config)# ntp server 192.168.1.1
B(config)# exit
```

DEVICE-C

```
C#configure terminal
C(config)# ntp server 192.168.2.1
C(config)# exit
```

【检验方法】

- 在 DEVICE-B 上通过 **show ntp status** 查看 NTP 配置信息。
- DEVICE-B 会向 192.168.1.1 发送时间同步报文，从 DEVICE-A 同步时间。
- DEVICE-B 从 DEVICE-A 成功同步时间之后，可以响应 DEVICE-C 的时间同步请求。
- 在 DEVICE-B 和 DEVICE-C 上通过 **show clock** 命令可以查看时间是否成功同步。

▾ NTP 本地时钟参考模式

【网络环境】

图 3-6



- DEVICE-B：本地时钟为 NTP 参考时钟源
- DEVICE-C：从 DEVICE-B 同步时间

【配置方法】

- DEVICE-B 配置本地时钟为 NTP 参考时钟源
- DEVICE-C 配置 DEVICE-B 为参考时钟源

DEVICE-B

```
B#configure terminal
B(config)# ntp master
```

```
DEVICE-C B(config)# exit
          C#configure terminal
          C(config)# ntp server 192.168.2.1
          C(config)# exit
```

【检验方法】 ● 在 DEVICE-C 上通过 **show clock** 命令可以查看时间是否成功同步。

3.5.2 配置NTP安全认证

配置效果

▾ 从可信参考时钟源同步时间

设备作为客户端，只从可信任的外部参考时钟源同步时间到本地时钟

▾ 给可信设备提供时间同步服务

设备的本地时钟作为 NTP 参考时钟源，只对可信的设备提供时间同步服务

注意事项

客户端和服务器的认证密钥必须一致。

配置方法

▾ 配置 NTP 全局安全认证机制

- 必须配置
- 默认情况下设备不开启安全认证机制。

▾ 配置 NTP 全局认证密钥

- 必须配置
- 默认情况下设备没有认证密钥。

▾ 配置 NTP 全局信任密钥 ID

- 可选配置
- 给可信设备提供时间同步服务，必须通过密钥 ID 指定可信认证密钥。
- 只允许配置一个信任密钥，所指定的认证密钥必须和可信设备一致。

▾ 配置外部参考时钟源的认证密钥 ID

- 可选配置

- 从可信参考时钟源同步时间，必须通过密钥 ID 指定可信认证密钥。
- 每个可信参考时钟源分别对应一个认证密钥，认证密钥必须和可信参考时钟源的密钥一致。

检验方法

- 通过 **show run** 查看配置是否正确
- 通过 **show clock** 查看是否从可信设备同步时间

相关命令

✎ 开启安全认证机制

【命令格式】 **ntp authenticate**

【参数说明】 -

【命令模式】 全局模式

【使用指导】 缺省情况下，客户端不使用全局安全识别机制。如果未使用安全识别机制则不对通信进行加密处理。但是仅仅设置了全局安全标志，并不代表一定采用了加密方式完成服务器与客户端的通信，还必须完成其他全局密钥配置并设置服务器加密密钥才可能发起和服务器的加密通信。

✎ 设置全局认证密钥

【命令格式】 **ntp authentication-key key-id md5 key-string [enc-type]**

【参数说明】 *key-id*：认证密钥的全局 ID（1-4294967295）。

key-string：密钥字符串。

enc-type：可选。输入的密钥是否加密（0 表示无加密，7 表示简单加密，默认为无加密）。

【命令模式】 全局模式

【使用指导】 -

✎ 设置 NTP 服务的可信密钥

【命令格式】 **ntp trusted-key key-id**

【参数说明】 *key-id*：认证密钥的全局 ID（1-4294967295）。

【命令模式】 全局模式

【使用指导】 -

✎ 设置外部参考时钟源的可信密钥

参考 [“配置NTP服务器”](#)

配置举例

✎ 安全认证

【网络环境】

图 3-7



- DEVICE-B：配置为 NTP 客户端/服务器模式，给 DEVICE-C 提供需要安全认证的 NTP 服务，认证密钥为 “abcd”
 - DEVICE-A：作为 DEVICE-B 的参考时钟源
 - DEVICE-C：从 DEVICE-B 同步时间
- 【配置方法】
- DEVICE-B 配置 DEVICE-A 为参考时钟源
 - DEVICE-C 配置 DEVICE-B 为参考时钟源

DEVICE-B

```

B#configure terminal
B(config)# ntp authentication-key 1 md5 abcd
B(config)# ntp trusted-key 1
B(config)# ntp server 192.168.1.1
B(config)# exit
  
```

DEVICE-C

```

C#configure terminal
C(config)# ntp authentication-key 1 md5 abcd
C(config)# ntp server 192.168.2.1 key 1
C(config)# exit
  
```

【检验方法】

- DEVICE-B 会向 192.168.1.1 发送时间同步报文，携带认证信息，从 DEVICE-A 同步时间。
- 在 DEVICE-B 上通过 **show clock** 命令查看时间是否成功同步。

配置举例

3.5.3 配置NTP访问控制

配置效果

NTP 服务的访问控制功能提供了一种最小限度的安全措施（更安全的方法是使用 NTP 身份验证机制）。

注意事项

- 目前系统暂未支持控制查询功能 用于通过网络管理设备对 NTP 服务器进行控制（如设置闰秒标记或监控其工作状态等）。虽然是按照上述顺序进行规则匹配，但涉及到与控制查询相关的请求都无法支持。
- 如果未配置任何访问控制规则，则所有访问都是允许的。但一旦配置了访问控制规则，则仅有规则中所允许的访问才能进行。

相关配置

设置 NTP 的访问控制权限

- 可选配置
- 通过 `ntp access-group` 配置 NTP 访问控制权限及对应的 ACL

检验方法

通过 `show run` 查看 NTP 配置是否正确配置

相关命令

配置 NTP 服务的访问控制权限

【命令格式】 `ntp access-group { peer | serve | serve-only | query-only } access-list-number | access-list-name`

【参数说明】 `peer`：既允许对本地 NTP 服务进行时间请求和控制查询，也允许本地设备与远程系统同步时间（完全访问权限）。

`serve`：允许对本地 NTP 服务进行时间请求和控制查询，但不允许本地设备与远程系统同步时间。

`serve-only`：仅允许对本地 NTP 服务进行时间请求。

`query-only`：仅允许对本地 NTP 服务进行控制查询。

`access-list-number`：IP 访问控制列表标号；范围为 1~99 和 1300~1999。关于如何创建 IP 访问控制列表请参考《ACL》中的相关描述。

`access-list-name`：IP 访问控制列表名。关于如何创建 IP 访问控制列表请参考《访问控制列表配置指南》中的相关描述。

【命令模式】 全局模式

【使用指导】 配置 NTP 访问控制权限。

当一个访问请求到达时，NTP 服务按照从最小访问限制到最大访问限制的顺序依次匹配规则，以第一个匹配到的规则为准。匹配顺序为 `peer`、`serve`、`serve-only`、`query-only`。

配置举例

NTP 访问控制权限配置

【配置方法】 配置只允许 192.168.1.1 的设备对本地设备进行时间同步请求

```
Ruijie(config)# access-list 1 permit 192.168.1.1
Ruijie(config)# ntp access-group serve-only 1
```

3.6 监视与维护

查看运行情况

作用	命令
show ntp status	显示当前的 NTP 信息

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
debug ntp	打开调试功能。
no debug ntp	关闭调试功能。

4 SNTP

4.1 概述

SNTP (Simple Network Time Protocol , 简单网络时间协议) 是 NTP 的简化版本 , 主要用来同步因特网中的计算机时钟。SNTP 适用于无需完全使用 NTP 功能的情况。

NTP 算法复杂, 对系统要求较高。而 SNTP 在实现时, 计算时间用了简单的算法, 性能较高。而精确度一般也能达到 1 秒左右, 也能基本满足绝大多数场合的需要。由于 SNTP 的报文和 NTP 的报文是完全一致的, 所以设备实现的 SNTP Client 能完全兼容 NTP Server。

i 下文仅介绍 SNTP 的相关内容。

协议规范

- RFC 2030 : Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI

4.2 典型应用

典型应用	场景描述
从NTP服务器同步时间	设备作为客户端, 从 NTP 服务器同步时间

4.2.1 从NTP服务器同步时间

应用场景

如图所示, DEVICE-B 将本地时钟作为 NTP 参考时钟源, 向 DEVICE-C 提供时间同步服务。

DEVICE-C 作为 SNTP 客户端, 从 DEVICE-B 同步时间。

图 4-1



功能部属

- 指定 DEVICE-B 为 DEVICE-C 的 SNTP 服务器。

- DEVICE-C 开启 SNTP 功能

4.3 功能详解

基本概念

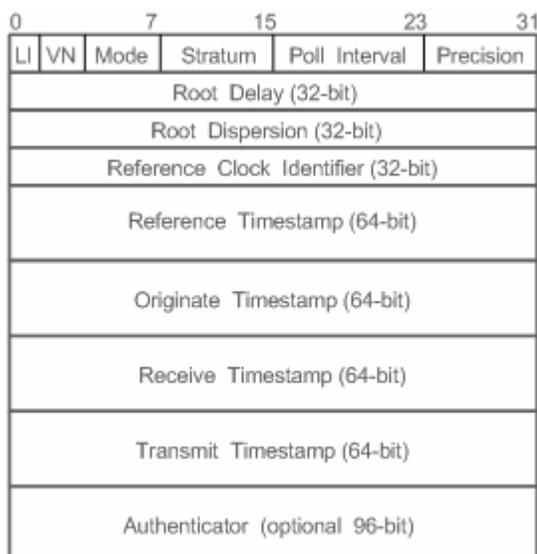
SNTP 报文

SNTPV4 是从 NTP 发展过来的，主要是简化 NTP 的功能。SNTPV4 并没有改变 NTP 规范和原有实现过程。SNTPV4 的消息格式于 RFC1305 中定义的 NTP 格式是一致的，只是某些数据域被初始化为预定的值。

同 RFC1305 定义，SNTP 采用 UDP 报文进行传输，UDP 端口号为 123。

NTP 时间同步报文格式如 图 3-3

图 4-2 SNTP 时间同步报文格式



- Leap Indicator (LI) : 2 比特，闰秒标志。
-
- i** 00-无警告信息 01-上一分钟有 61 秒 10-上一分钟有 59 秒 11-时钟未同步
-
- Version Number (VN) : 3 比特，NTP/SNTP 版本号，当前版本号为 3。
 - Mode : 3 比特，SNTP/NTP 工作模式。
-
- i** 0-未定义 1-主动对等体 2-被动对等体 3-客户端 4-服务器 5-广播 6-控制信息 7-保留
-
- Stratum : 8 比特，本地时钟的层数 (0-未定义 1-主参考时钟源 其它值-次参考时钟源)。
 - Poll Interval : 8 位整数，轮询时间 (秒数)
 - Precision : 8 位整数，本地时钟的时间精度 (秒数)
 - Root Delay : 32 位整数，到主参考时钟源的往返时间

- Root Dispersion : 32 位整数，相对于参考时钟源的最大误差
- Reference Clock Identifier : 32 比特，参考时钟源的标识
- Reference Timestamp : 64 位时间戳，最后一次被设置或者被校正的时间
- Originate Timestamp : 64 位时间戳，时间同步请求报文离开客户端的本地时间
- Receive Timestamp : 64 位时间戳，时间同步请求报文到达服务器的本地时间
- Transmit Timestamp : 64 位时间戳，时间同步响应报文离开服务器的本地时间
- Authenticator (可选) : 验证信息

功能特性

功能特性	作用
SNTP时间同步	从 SNTP/NTP 服务器同步时间到本地设备。

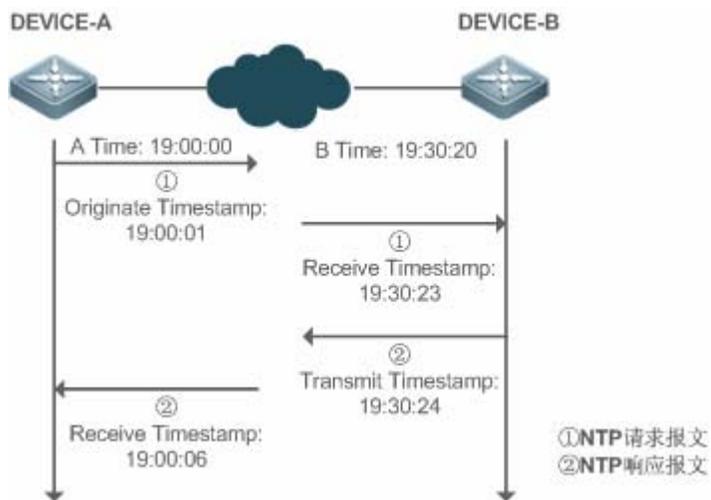
4.3.1 SNTP时间同步

工作原理

SNTP 同步时间的方式是与服务器之间交互 SNTP/NTP 报文。客户端每隔一段时间（默认是半小时）向服务器发送时间同步报文。收到服务器响应报文后进行时间同步。

SNTP时间同步报文格式如图 3-4

图 4-3 SNTP 基本工作原理图



DEVICE-B (下面简称 B) 作为 NTP 参考时钟源，DEVICE-A (下面简称 A) 作为 SNTP 客户端从 DEVICE-B 同步时间，在某一时刻 A 的本地时钟为 19:00:00，B 的本地时钟为 19:30:20：

8. A 发出 SNTP/NTP 请求报文，报文离开 A 的本地时间 (T0) 为 19:00:00，填充在 Originate Timestamp
9. 经过 2 秒的网络延时，B 收到请求报文的本地时间 (T1) 为 19:30:23，填充在 Receive Timestamp
10. B 处理 NTP 请求，1 秒后响应 NTP 报文，报文离开 B 的本地时间 (T2) 为 19:30:24，填充在 Transmit Timestamp
11. 经过 2 秒的网络延时，A 接收到响应报文，响应报文到达 A 的本地时间 (T3) 为 19:00:06

时间同步的具体算法如下：

- A 通过公式 $((T1-T0)+(T2-T3))/2$ 计算出 B 和 A 的时间差为 30 分 20 秒
- A 通过公式 $(T3-T0)-(T2-T1)$ 计算出 A 和 B 的报文往返的延时为 4 秒

相关配置

打开 SNTP

- 缺省 SNTP 状态是关闭的。
- 通过 `sntp enable` 命令开启 SNTP 功能

配置 SNTP 服务器

- 缺省情况下，没有配置 SNTP 服务器。
- 通过 `sntp server` 命令指定 SNTP 服务器。

配置 SNTP 同步时钟间隔

- 缺省情况下，SNTP 同步时钟的间隔是 1800s。
- 通过 `sntp interval` 命令指定 SNTP 服务器。

4.4 配置详解

配置项	配置建议&相关命令
配置SNTP	 必须配置，用于开启 SNTP 功能
	<code>sntp enable</code> 打开 SNTP
	<code>sntp server</code> 配置 SNTP Server 的地址
	 可选配置，用于调整 SNTP 时间同步间隔
	<code>sntp interval</code> 配置 SNTP 同步时钟的间隔

4.4.1 配置SNTP

配置效果

SNTP Client 一定的时间间隔定期访问 NTP Server，可以定时校正时钟。

注意事项

通过 SNTP 协议通讯后获取的时间都是格林威治标准时间 (GMT)，为了准确的获取本地时间，需要设置本地时区来对标准时间进行调正。

配置方法

打开 SNTP

- 必须配置，缺省 SNTP 状态是 Disable。

配置 SNTP Server 的地址

- 必须配置，缺省没有设置 SNTP/NTP 服务器

配置 SNTP 同步时钟的间隔

- 可选配置
- 默认情况下，设备每隔半小时同步一次时间

检验方法

使用 `show sntp` 命令查看 SNTP 相关参数。

相关命令

打开 SNTP

【命令格式】 `sntp enable`

【参数说明】 -

【命令模式】 全局配置模式

【使用指导】 缺省 SNTP 状态是 Disable。

`no sntp enable` 全局配置命令来关闭 SNTP。

配置 SNTP/NTP Server 的地址

【命令格式】 `sntp server [oob] ip-address [via mgmt-name]`

【参数说明】 `ip-address` : NTP/SNTP 服务器的 IP 地址。缺省没有设置任何 NTP/SNTP 服务器。

`oob` : NTP/SNTP 服务器支持带外管理接口 (interface of mgmt) 。

`mgmt-name` : 指定在 oob 模式下报文的出口管理口。

【命令模式】 全局配置模式

【使用指导】 由于 SNTP 协议和 NTP 完全兼容，所以这个 Server 完全可以配置成 internet 上公用的 NTP Server。

由于 SNTP 的报文和 NTP 的报文是完全一致的，所以 SNTP Client 能完全兼容 NTP Server。网络上存在着

较多的 NTP Server，用户可以选择一个网络延迟较少的一个作为设备上的 SNTP Server。

配置 SNTP 同步时钟的间隔

- 【命令格式】 **sntp interval seconds**
- 【参数说明】 *seconds*：定时同步的间隔，单位为“秒” 范围为 60 秒--65535 秒。缺省值为 1800s。
- 【命令模式】 全局配置模式
- 【使用指导】 该命令设置 SNTP Client 需要定时和 NTP/SNTP Server 同步时钟的时间间隔。

 这里设置的时间间隔不会立即生效，如果要立即生效，请配置完时间间隔后执行 **sntp enable** 命令。

配置举例

SNTP 时间同步

【网络环境】

图 4-4



- DEVICE-B：网络上的 NTP 服务器
- DEVICE-C：从 DEVICE-B 同步时间

【配置方法】 DEVICE-C 开启 SNTP 功能，NTP 服务器配置为 DEVICE-B

DEVICE-C

```
C#configure terminal
C(config)# sntp server 192.168.2.1
C(config)# sntp enable
C(config)# exit
```

- 【检验方法】
- 在 DEVICE-C 上通过 **show clock** 命令可以查看时间是否成功同步。
 - 在 DEVICE-C 上 **show sntp** 查看 sntp 状态和服务器是否配置成功

4.5 监视与维护

清除各类信息

查看运行情况

作用	命令
show sntp	查看 SNTP 的相关参数

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
debug sntp	打开调试功能。

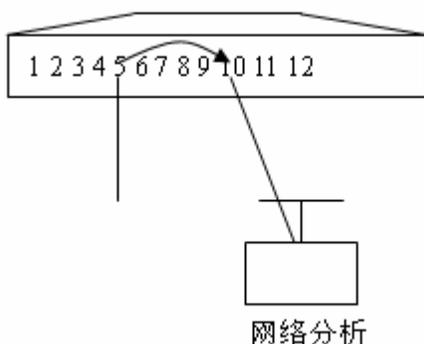
5 SPAN-RSPAN

5.1 概述

镜像(SPAN)是将指定端口的报文复制到交换机上另一个连接有网络监测设备的端口,进行网络监控与故障排除。

通过 SPAN 可以监控所有进入和从源端口输出的报文。例如,在下图中,端口 5 上的所有报文都被映射到了端口 10,连接在端口 10 上的网络分析仪虽然没有和端口 5 直接相连,但是可以接收通过端口 5 上的所有报文。

图 5-1 SPAN 配置实例



镜像功能主要应用于在网络监控和故障排查两种场景中,用于对网络信息的监控和网络故障的解决。

RSPAN(Remote SPAN, 远程镜像)是 SPAN 的扩展,能够远程监控多台设备,每个 RSPAN 会话建立于用户指定的 Remote VLAN 内。远程镜像突破了被镜像端口和镜像端口必须在同一台设备上的限制,使被镜像端口和镜像端口间可以跨越多个网络设备,这样用户就可以坐在中心机房通过分析仪观测远端被镜像端口的数据报文了。

远程镜像的应用场景和本地镜像类似,但使得用户不必呆在机房就可以对数据进行实时监控,极大地方便用户。

VSPAN 是 VLAN SPAN 的简称,是指将某些 VLAN 的数据流作为数据源镜像到目的端口,它和基于端口的镜像配置方式类似。VSPAN 具有以下特性:

- 可以指定某个 VLAN 作为镜像的数据源,这个 VLAN 不能是 Remote VLAN。
- 可以指定某些 VLAN 作为镜像的数据源,这些 VLAN 不能是 Remote VLAN。
- 配置 VLAN 做为源时只能基于 rx 方向的报文镜像。

协议规范

- 无

5.2 典型应用

典型应用	场景描述
------	------

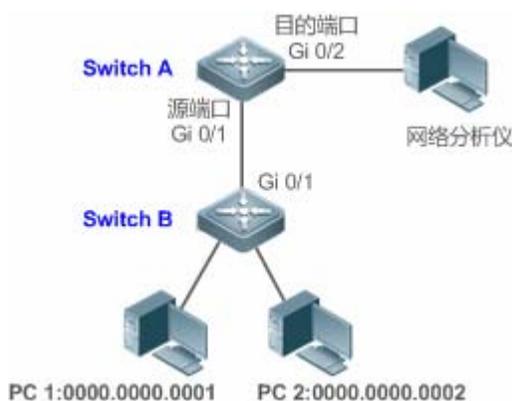
基于流的镜像	需要监控具有特定特征的数据流，比如监控指定 ACL 策略的数据流。
一对多的镜像	需要多个用户对同一端口的数据进行监控。
RSPAN基本应用	需要将镜像源设备的报文镜像到目的的设备上进行监控。

5.2.1 基于流的镜像

应用场景

如图所示，通过适当的配置，网络分析仪能够监控 Switch A 转发给 Switch B 的所有数据流，监控来自 Switch B 的特定数据流（如来自 PC1 和 PC2 的数据流）。

图 5-2 SPAN 简单应用拓扑



【注释】 0000.0000.0001 为 PC1 的 MAC 地址。
0000.0000.0002 为 PC1 的 MAC 地址。

功能部属

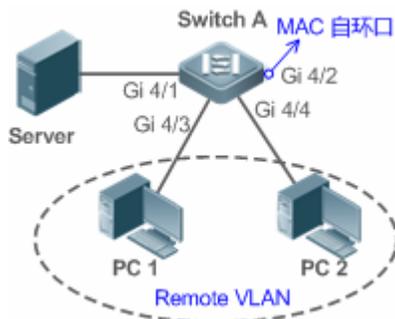
- 上图中，在连接网络分析仪的设备 Switch A 上配置 SPAN 功能，将连接 Switch B 的端口 Gi 0/1 设置为 SPAN 的源端口，将直连网络分析仪的端口 Gi 0/2 设置为 SPAN 的目的端口。
- 配置 SPAN 源端口 Gi 0/1 基于流的镜像（仅允许 PC1 和 PC2 的数据流）。

5.2.2 一对多的镜像

应用场景

如图所示，在单台设备上实现一对多镜像，即 PC1 和 PC 2 均可监控服务器相连端口的收发流量。用户可以通过适当的配置（远程 VLAN、MAC 自环口等），可以在 PC1、PC2 中对流经 Gi 4/1 的数据流进行监控，从而实现对服务器数据流的监控。

图 5-3 一对多镜像应用拓扑



【注释】 Remote VLAN：远程 VLAN。

功能部属

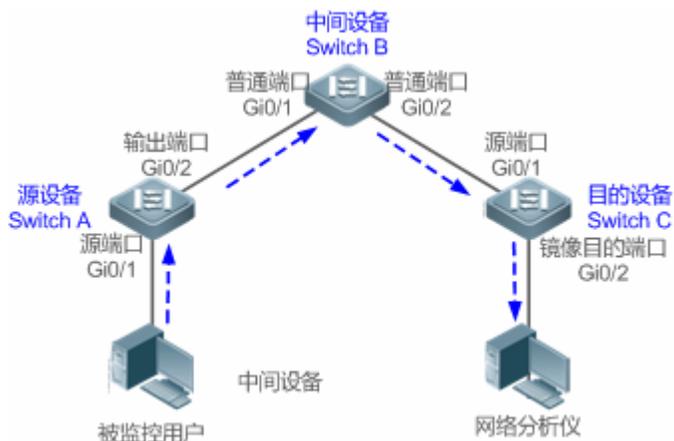
- 在设备 Switch A 上创建 Remote VLAN。
- 指定设备 Switch A 为 RSPAN 的源设备，配置直连服务器的端口 Gi4/1 为镜像源端口；选择一个 Down 状态的端口本例为 Gi 4/2 为镜像输出端口，将该端口加入 Remote VLAN，并配置 MAC 自环（可以在接口模式下通过 `mac-loopback` 命令进行配置）。
- 将直连 PC1 和 PC2 的端口加入 Remote VLAN。

5.2.3 RSPAN基本应用

应用场景

如图所示，网络分析仪可以通过远程镜像功能，实现在目的设备 Switch C 上通过中间设备 Switch B 监控连接到源设备 Switch A 上的用户。且设备之间均能正常交换数据。

图 5-4 RSPAN 基本应用拓扑



【注释】 -

功能部属

- 在源设备 Switch A、中间设备 Switch B 和目的设备 Switch C 上配置 Remote VLAN。
- 在源设备上，配置直连用户的端口 Gi 0/1 为源端口，与中间设备相连的端口 Gi 0/2 为输出端口，并配置输出端口可交换功能。
- 在中间设备上，与源设备、目的设备相连的端口 Gi 0/1 和 Gi 0/2 仅需配置为普通端口。
- 在目的设备上，与中间设备相连的端口 Gi 0/1 作为源端口，仅需配置为普通端口，与网络分析仪相连的端口 Gi 0/2 配置为镜像目的端口，并配置镜像目的端口可交换功能。

5.3 功能详解

基本概念

▾ SPAN 会话

SPAN 会话是镜像源端口与目的端口之间的数据流，可以监控单个或多个端口的输入、输出、双向的报文。Switched Port、Routed Port 和 AP(聚合端口)等类型的端口都可以配置为 SPAN 会话的源端口和目的端口。端口加入 SPAN 会话后并不影响交换机的正常操作。

用户可以在处于关闭状态的端口上配置 SPAN 会话，但是该 SPAN 会话是非活动的，只有相关的端口被打开后，SPAN 会话才会变为活动状态。另外，SPAN 会话在交换机上电后并不立即生效，直到目的端口处于可操作状态后，SPAN 会话才处于活动状态。用户可以通过 **show monitor [session session-num]** 命令查看 SPAN 会话的操作状态。

▾ 镜像数据流

SPAN 会话包含以下三种方向的数据流：

- 输入数据流：所有源端口上接收到的报文都将被复制一份到目的端口。在一个 SPAN 会话中，用户可以监控一个或多个源端口的输入报文。由于某些原因(如端口安全)，从源端口输入的报文可能被丢弃，但这不影响 SPAN 功能，该报文仍然会镜像到目的端口。
- 输出数据流：所有从源端口发送的报文都将复制一份到目的端口。在一个 SPAN 会话中，用户可以监控一个或多个源端口的输出报文。若由于某些原因，从别的端口发送到源端口的报文可能被丢弃，同样，该报文也不会发送到目的端口。由于某些原因从源端口输出的报文的格式可能改变，例如源端口输出经过路由之后的报文，报文的源 MAC、目的 MAC、VLAN ID 以及 TTL 发生变化，同样，拷贝到目的端口的报文的格式也会变化。
- 双向数据流：包括上面所说的两种数据流。在一个 SPAN 会话中，用户可监控一个或多个源端口的输入和输出方向的数据流。

▾ 源端口

源端口也被称为被监控口，在 SPAN 会话中，源端口上的数据流被监控，用于网络分析或故障排除。在单个 SPAN 会话中，用户可以监控输入、输出和双向数据流，且源端口的最大个数没有限制。

源端口具有以下特性：

- 源端口可以是 Switched Port、Routed Port 或 AP。
- 源端口不能同时作为目的端口。
- 源端口和目的端口可以属于同一 VLAN，也可以属于不同 VLAN。

目的端口

SPAN 会话有一个目的端口(也被称为监控口)，用于接收源端口的报文拷贝。

目的端口具有以下特性：

- 目的端口可以是 Switched Port、Routed Port 或 AP。
- 目的端口不能同时作为源端口。

功能特性

功能特性	作用
SPAN	同一设备上端口的镜像。
RSPAN	跨设备的端口镜像。

5.3.1 SPAN

本地镜像主要是用来监控交换机上的数据流。通过将一个端口上的帧拷贝到交换机上另一个连接有网络分析设备或 RMON 分析仪的端口上来分析该端口上的通讯。

工作原理

端口收发报文时检测用户如果有配置该端口作为镜像源时，则会将该端口收发的报文复制到目的端口一份。

配置镜像源端口

用户需要指定镜像会话 ID、源端口名字来配置镜像源端口，并通过镜像方向的可选配置项决定镜像数据流的方向或通过指定 ACL 策略镜像特定数据流。

配置镜像目的端口

用户需要指定镜像会话 ID、目的端口名字来配置镜像目的端口，并通过交换功能可选配置项决定是否在该目的镜像端口上开启交换功能和剥离 TAG 信息功能。

相关配置

系统镜像功能默认是关闭的，只有用户创建会话，并配置源和目的镜像端口才会开启镜像功能。镜像会话可以在配置镜像的源端口或者目的端口的时候进行创建。

配置镜像源端口

缺省情况下，镜像会话中没有镜像源端口。用户通过下面命令配置镜像源端口。

```
monitor session session-num source interface interface-id [ both | rx | tx ] [ acl name ]
```

其中，

session-num：镜像会话 ID，针对不同产品支持镜像会话个数会有所不同。

interface-id：待配置的镜像源端口。

rx：配置 **rx** 选项后，只监听源端口接收的报文。

tx：配置 **tx** 选项后，只监听源端口发送的报文。

both：配置 **both** 选项后，源端口收发的报文都会送到目的端口进行监听，即包含 **rx** 和 **tx**。如果用户不配置 **rx**、**tx** 和 **both** 三个选项中的任何一个则默认开启 **both** 选项。

acl：配置该选项时则需要用户指定一个 ACL 策略，即监听源端口上该策略允许的报文，默认不开启该功能。

配置镜像目的端口

缺省情况下，镜像会话中没有镜像源端口。用户通过下面命令配置镜像的目的端口。

```
monitor session session-num destination interface interface-id [ switch ]
```

其中，

switch：在配置镜像目的的口时，如果没有打开该选项，则镜像目的的口只接收镜像源的镜像报文，其它报文均丢弃。如果打开该选项，除了接收镜像源的镜像报文同时非源端口送过来的报文也不会丢弃，即不影响目的的口和外界的其他通信。

配置镜像目的端口时，如果没有配置 **switch** 选项则默认关闭相应功能。

配置基于流的镜像

缺省情况下，该功能关闭。用户通过 **monitor session session-num source interface interface-id rx acl acl-name** 命令配置基于流的镜像。

 使用过程中，用户需要特别注意以下几点：

- 镜像的目的端口参与 STP 树的计算。
- 如果改变了源端口的 VLAN 配置，配置将马上生效。
- 如果改变了目的端口的 VLAN 配置，配置马上生效。
- 如果禁用了源端口或目的端口，SPAN 将不起作用。
- 如果将源端口或目的端口加入 AP，源端口或目的端口将退出 SPAN 会话。
- 如果 VLAN (VLAN 列表) 做为镜像源时，要保证目的口有足够大的宽带能够接收整个 VLAN 的镜像数据。
- 产品的差异性，并不是所有产品都支持上述命令的所有选项。

5.3.2 RSPAN

RSPAN 能够远程监控多台设备,每个 RSPAN Session 建立于用户指定的 Remote VLAN 内。远程镜像突破了被镜像端口和镜像端口必须在同一台设备上的限制,使被镜像端口和镜像端口间可以跨越多个网络设备。

工作原理

远程镜像的原理是原设备、中间设备和目的设备通过创建一个 Remote VLAN,且所有参与会话的端口都要加入该 Remote VLAN 中,镜像报文在 Remote VLAN 内进行广播,使得镜像报文从源交换机的源端口传送到目的交换机的目的端口。

配置远程 VLAN

镜像源端口的报文就是通过远程 VLAN 进行广播来实现报文从本台交换机复制到远程交换机的。镜像源端口、输出端口、反射端口及中间设备的透传端口(中间设备的报文进入端口、输出端口)和目的交换机的目的端口及进入端口都必需位于该远程 VLAN 内。该功能需要在 VLAN 模式下将 VLAN 配置远程 VLAN。

配置远程镜像会话

远程镜像源端口和目的端口的配置和本地镜像类似,但是在配置时指定的镜像会话 ID 必需是远程镜像。

配置远程镜像源端口

配置远程镜像源端口和配置本地镜像源端口一样,只是在指定镜像会话 ID 时,需要使用远程镜像会话 ID。

配置远程镜输出端口

输出端口位于源设备中。如果用户配置的远程镜像实现的仍是一对一的镜像,则只需配置一个输出端口即可。

输出端口必需位于远程 VLAN 内,源端口被镜像的报文在该远程 VLAN 内进行广播。源设备中就是通过输出端口将报文传送到中间交换机或目的交换机中。

配置远程镜像目的端口

配置远程镜像的目的端口时必需指定远程镜像会话 ID,远程 VLAN 及端口名字,这样源端口的报文就可以通过远程 VLAN 将报文从源端口复制到目的端口。

配置基于流的远程镜像

RSPAN 是对本地 SPAN 的扩展,因此 RSPAN 同样也支持基于流的镜像,具体配置同基于流的 SPAN 配置。基于流的 RSPAN 不影响正常通讯。

用户可以在 RSPAN 源设备上配置源端口的 in 方向的 ACL,支持标准 ACL、扩展 ACL、MAC ACL、自定义 ACL。

用户可以在 RSPAN 源设备上配置源端口的 in 方向的端口 ACL,可以在 RSPAN 目的设备上配置目的端口 out 方向的端口 ACL。用户可以在 RSPAN 源交换机上基于 Remote VLAN 应用 out 方向的 ACL,在 RSPAN 目的交换机上基于 Remote VLAN 应用 in 方向的 ACL。

配置一对多的镜像

如果用户需要将同一源端口的数据流镜像到多个目的端口,可以配置一个 RSPAN 会话,该 RSPAN 的源口为一对多镜像源端口,转发口(即为源设备的输出端口)为一个其它以太网口。同时在 RSPAN 转发口的接口模式下配置 MAC 自环功

能，并把期望的镜像出口以及 RSPAN 转发口加入到 Remote VLAN 中，镜像报文在 RSPAN 转发口上进行回环，然后在 Remote VLAN 内广播，从而实现一对多镜像。

相关配置

远程镜像功能默认是关闭的，只有用户创建远程镜像会话，并配置远程 VLAN、源和目的镜像端口才会开启该功能。

配置远程 VLAN

缺省情况下，RSPAN 没有指定远程 VLAN。用户可以在 VLAN 模式下，通过 **remote-span** 命令将该 VLAN 配置为远程 VLAN。一个远程 VLAN 对应一个 RSPAN 会话。

配置远程镜像的源设备

缺省该功能关闭。用户可以在全局模式下通过 **monitor session session-num remote-source** 命令将该设备配置为指定 RSPAN 会话的远程源设备。

配置远程镜像的目的设备

缺省该功能关闭。用户可以在全局模式下通过 **monitor session session-num remote-destination** 命令将该设备配置为指定 RSPAN 会话的远程目的设备。

配置远程镜像源端口

源设备配置会话的源端口，和本地镜像配置源端口一样，只是要用远程会话 ID。缺省该功能关闭。

配置远程源镜像的输出目的端口

缺省该功能关闭。用户可以在全局模式下通过 **monitor session session-num destination remote vlan remote-vlan interface interface-name [switch]** 命令配置源镜像的输出目的端口。可选项 **switch** 配置的情况下，表示该输出目的口可以参与正常的报文交换，缺省该可选项是不配置的。注意输出端口必须加入到远程 VLAN 中。

配置远程目的设备的目的口

缺省该功能关闭。用户可以在全局模式下通过 **monitor session session-num destination remote vlan remote-vlan interface interface-name [switch]** 配置远程目的设备的目的端口。可选项 **switch** 配置的情况下，表示该输出目的口可以参与正常的报文交换，缺省该可选项是不配置的。注意目的端口必须加入到远程 VLAN 中。

 使用过程中，用户需要特别注意以下几点：

- Remote VLAN 必需在每台设备中都要进行配置，且 VLAN ID 必须一致，并且所有参与会话的端口都要加入该 VLAN 中。
- 建议不要将普通端口加入 Remote VLAN。
- 不要在与中间交换机或目的交换机相连的端口上配置镜像源端口，否则可能引起网络内的流量混乱。
- **配置 MAC 回环口**

缺省该功能关闭。用户可以在接口模式下通过 **mac-loopback** 配置端口为 MAC 回环口，实现一对多镜像。注意镜像目的端口和期望的镜像出口必须加入到远程 VLAN 中。

5.4 产品说明



S7508 产品，配置输出方向的镜像时，如果被镜像的是组播路由报文，则镜像目的端口输出的报文为路由前的组播报文。



S7508 产品，配置镜像目的出口时，需要强制开启交换功能。



S7508 产品，当在隧道源口配置输出方向上的镜像时，如果被镜像的是组播隧道报文，则镜像目的口输出的报文是隧道封装之前的报文。



S7508 产品，只能对转发数据报文镜像。



S7508 产品，指定 vlan 作为镜像数据源的功能中，不支持 tx 镜像。

5.5 配置详解

配置项	配置建议 & 相关命令	
配置SPAN基本功能	必须配置。用于创建本地镜像。	
	monitor session session-num source interface interface-id [both rx tx]	配置镜像源端口
	monitor session session-num destination interface interface-id [switch]	配置镜像目的端口
	monitor session session-num source interface interface-id rx acl acl-name	配置基于流的镜像
	monitor session session-num source vlan vlan-id [rx]	指定某个 VLAN 作为镜像的数据源
	monitor session session-num source filter vlan vlan-id-list [rx]	指定某些 VLAN 作为镜像的数据源
配置RSPAN基本功能	必须配置。用于创建远程镜像。	
	monitor session session-num remote-source	配置远程镜像会话 ID 并指定为源设备
	monitor session session-num remote-destination	配置远程镜像会话 ID 并指定为目的设备
	remote-span	配置远程 VLAN
	monitor session session-num source interface interface-id [both rx tx]	配置远程源镜像源端口
	mac-loopback	配置回环口

	<pre>monitor session <i>session-num</i> destination remote vlan <i>remote-vlan-id</i> interface <i>interface-id</i> [switch]</pre>	配置远程源镜像的输出端口或者远程目的镜像的目的端口
--	--	---------------------------

5.5.1 配置SPAN基本功能

配置效果

- 配置镜像会话的源和目的端口。
- 目的口可以监控到任何进出源端口的报文。

注意事项

- 如果将源端口或目的端口加入 AP，源端口或目的端口将退出 SPAN 会话。
- 如果镜像目的口没有开启 switch 功能，则目的口只能接收镜像报文，其它流经该端口的报文将被丢弃。开启后可以接收非镜像报文。

配置方法

↘ 镜像会话

- 全局模式。必须配置。
- 可以通过配置镜像的源端口或者目的端口时同时配置镜像会话。还可以通过配置指定某个 VLAN 或者某些 VLAN 作为镜像的数据源时配置镜像会话。

↘ 配置镜像源端口

- 全局模式。必须配置。
- 配置镜像源端口时可以选择配置的镜像方向，缺省是 both 方向，即同时监测报文的接收和发送行为。

↘ 配置镜像目的端口

全局模式。必须配置。

只有同时配置镜像的源端口或者指定 VLAN 作为镜像数据源，以及配置镜像的目的端口时，该镜像会话才真正起作用。

检验方法

- 镜像配置的校验也可以通过 **show monitor** 或者 **show running** 命令查看。也可以在镜像目的口上进行抓包分析，通过抓取的报文查看镜像功能是否生效。

相关命令

配置镜像源端口

【命令格式】 **monitor session** *session-num* **source interface** *interface-id* [**both** | **rx** | **tx**]

【参数说明】 *session-num* : 镜像会话 ID

interface-id : 接口名字

both : 同时监控输入和输出方向的报文, 为缺省值

rx : 监控输入方向的报文

tx : 监控输出方向的报文

【命令模式】 全局模式

【使用指导】 -

配置镜像目的端口

【命令格式】 **monitor session** *session-num* **destination interface** *interface-id* [**switch**]

【参数说明】 *session-num* : 镜像会话 ID

interface-id : 接口名字

switch : 支持镜像目的口交换功能, 缺省为不打开

【命令模式】 全局模式

【使用指导】 -

配置基于流的镜像

【命令格式】 **monitor session** *session-num* **source interface** *interface-id* **rx acl** *acl-name*

【参数说明】 *session-num* : 镜像会话 ID

interface-id : 接口名字

acl-name : acl 名字

【命令模式】 全局模式

【使用指导】 -

指定某个 VLAN 作为镜像的数据源

【命令格式】 **monitor session** *session-num* **source vlan** *vlan-id* [**rx**]

【参数说明】 *session-num* : 镜像会话 ID

vlan-id : 指定的 VLAN ID

rx : 监控输入方向的报文

【命令模式】 全局模式

【使用指导】 -

指定某些 VLAN 作为镜像的数据源

【命令格式】 **monitor session** *session-num* **source filter vlan** *vlan-id-list* [**rx**]

【参数说明】 *session-num* : 镜像会话 ID

vlan-id-list : 指定的某些 VLAN ID

【命令模式】 全局模式

【使用指导】 -

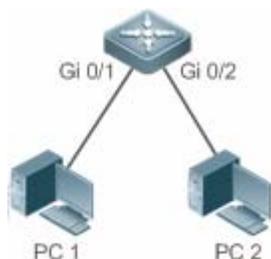
配置举例

i 以下配置举例，仅介绍与本地镜像相关的配置。

下面以本地镜像为例介绍

【网络环境】

图 5-5



- 【配置方法】
- 如图 1-5，配置设备 A 的 Gi 0/1 和 Gi 0/2 属于 VLAN 1。
 - 创建 SVI 1，并配置 SVI 1 地址为 10.10.10.10/24。
 - 配置 PC1、PC2 地址为 10.10.10.1/24、10.10.10.2/24，略。
 - 配置设备 A 的本地镜像，指定端口 Gi 0/1 和 Gi 0/2 分别为镜像的源端口和目的端口。

```

A
Ruijie# configure
Ruijie(config)# vlan 1
Ruijie(config-vlan)# exit
Ruijie(config)# interface vlan 1
Ruijie(config-if-VLAN 1)# ip address 10.10.10.10 255.255.255.0
Ruijie(config-if-VLAN 1)# exit
Ruijie(config)# monitor session 1 source interface gigabitEthernet 0/1
Ruijie(config)# monitor session 1 destination interface gigabitEthernet 0/2
  
```

【检验方法】 首先通过 **show monitor 命令**查看镜像是否正确配置，配置成功后 PC1 向 SVI 1 发送 PING 包，PC2 利用抓包工具进行监控。

```

A
Ruijie# show monitor
sess-num: 1
span-type: LOCAL_SPAN
src-intf:
GigabitEthernet 0/1      frame-type Both
dest-intf:
GigabitEthernet 0/2
  
```

常见错误

- 用户配置镜像源端口和目的端口时指定的会话 ID 不一致。
- 带宽大的端口被镜像到带宽小的端口可能会造成丢包。

5.5.2 配置RSPAN基本功能

配置效果

- 配置远程镜像会话源设备中的源端口和输出端口，配置目的设备中的目的端口。
- 远程目的设备中的目的口可以监控到任何进出源端口的报文。

注意事项

- 如果将源端口或目的端口加入 AP，源端口或目的端口将退出 SPAN 会话。
- 如果远程镜像目的口没有开启 switch 功能，则目的口只能接收镜像报文，其它流经该端口的报文将被丢弃。开启后可以接收非镜像报文。
- 所有参与镜像的报文均要加入远程 VLAN 中。
- 中间设备必需创建远程 VLAN，且透传端口要加入该 VLAN。

配置方法

▾ 远程镜像会话

- 全局模式。必须配置。
- 通过配置镜像源或者目的设备把镜像会话指定为成远程镜像会话。
- 镜像源设备和目的设备上配置的镜像 Remote VLAN ID 必须一致，保证源设备过来的镜像报文能被目的设备识别。

▾ 配置源镜像设备

- 全局模式。必须配置。
- 用于指定被远程镜像监控的设备。

▾ 配置目的镜像设备

- 全局模式。必须配置。
- 用于指定远程镜像报文输出目的设备。

▾ 配置远程镜像源端口

- 全局模式。必须配置。
- 在远程源镜像设备上配置。通过配置该功能，实现对远程镜像源端口的报文进行远程镜像监控。可以指定对经过该镜像源端口的输入方向、输出方向或者输入输出双向的 Remote VLAN 报文进行监控。

配置远程镜输出端口

- 全局模式。必须配置。
- 在远程源镜像设备上配置。通过配置该功能，实现将 Remote VLAN 接收到的镜像报文通过输出端口输出到远程镜像目的设备上。实现一对多远程镜像时需要同时配置输出端口，实现一对一远程镜像时只需配置输出端口。

配置远程镜像目的端口

- 全局模式。必须配置。
- 在远程目的镜像设备上配置。通过配置该功能，远程目的设备将 Remote VLAN 接收到的镜像报文通过目的端口转发给监控设备。

配置端口 MAC 回环

- 接口模式。可选配置。
- 在远程源镜像设备上配置。通过在远程镜像目的口配置该功能，并把其他端口加入 Remote VLAN，远程镜像报文在目的口上回环，之后在源设备上 Remote VLAN 内广播，实现一对多镜像。

检验方法

- 用户可以通过 **show monitor** 或者 **show running** 命令查看远程镜像中每台设备上面的配置是否成功。也可以在目的镜像设备上的目的镜像口抓包检查是否抓到了源镜像设备上的源端口镜像过来的报文。

相关命令

配置远程源镜像

- 【命令格式】 **monitor session session-num remote-source**
- 【参数说明】 *session-num*：远程镜像会话 ID
- 【命令模式】 全局模式
- 【使用指导】 -

配置远程目的镜像

- 【命令格式】 **monitor session session-num remote-destination**
- 【参数说明】 *session-num*：远程镜像会话 ID
- 【命令模式】 全局模式
- 【使用指导】 -

配置远程 VLAN

- 【命令格式】 **remote-span**
- 【参数说明】 -
- 【命令模式】 VLAN 模式
- 【使用指导】 -

配置源镜像源端口

【命令格式】 **monitor session** *session-num* **source interface** *interface-id* [**both** | **rx** | **tx**] [**acl** *acl-name*]

【参数说明】 *session-num* : 镜像会话 ID

interface-id : 接口名字

both : 同时监控输入和输出方向的报文，为缺省值

rx : 监控输入方向的报文

tx : 监控输出方向的报文

acl-name : ACL 名字

【命令模式】 全局模式

【使用指导】 和本地镜像配置源端口一样，但是指定的会话 ID 为远程镜像。

配置远程源镜像输出端口

【命令格式】 **monitor session** *session-num* **destination remote vlan** *remote-vlan* **interface** *interface-id* [**switch**]

【参数说明】 *session-num* : 镜像会话 ID

remote-vlan : 远程 VLAN

interface-id : 接口名字

switch : 是否参与报文交换

【命令模式】 全局模式

【使用指导】 实现一对多远程镜像时需要同时配置输出端口，实现一对一远程镜像时只需配置输出端口。

配置远程目的镜像的目的端口

【命令格式】 **monitor session** *session-num* **destination remote vlan** *remote-vlan* **interface** *interface-id* [**switch**]

【参数说明】 *session-num* : 镜像会话 ID

remote-vlan : 远程 VLAN

interface-id : 接口名字

switch : 是否参与报文交换

【命令模式】 全局模式

【使用指导】 -

配置端口 MAC 回环

【命令格式】 **mac-loopback**

【参数说明】 -

【命令模式】 接口模式

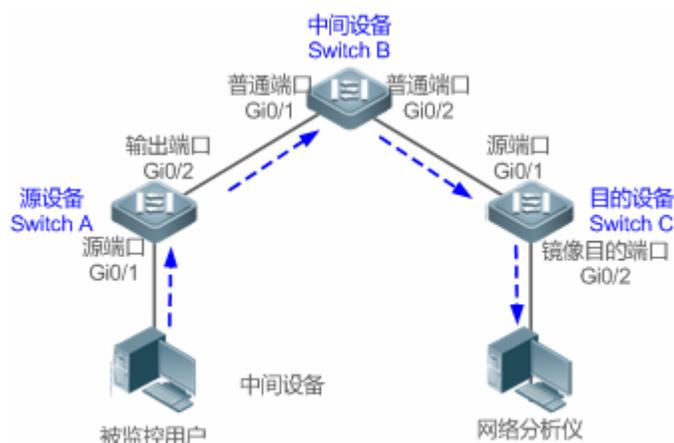
【使用指导】 实现一对多镜像时，需要打开接口 MAC 自环功能（请选择一个 DOWN 状态的端口做 MAC 自环口，且不要在该端口上添加其它配置。

配置举例

配置 RSPAN 镜像

【网络环境】

图 5-6



【配置方法】

- 如上图，源设备 A，中间设备 B 和目的设备 C 配置远程 VLAN。
- 在设备 A 中配置远程镜像源设备，并指定源端口、输出端口。
- 在设备 B 中配置端口 Gi0/1 和 Gi0/2 加入远程 VLAN。
- 设备 C 中配置远程镜像目的设备，并指定输出端口。

A	<pre>Ruijie# configure Ruijie(config)# vlan 7 Ruijie(config-vlan)# remote-span Ruijie(config-vlan)# exit Ruijie(config)# monitor session 1 remote-source Ruijie(config)# monitor session 1 source interface Gi 0/1 both Ruijie(config)# monitor session 1 destination remote vlan 7 interface Gi 0/2 switch</pre>
B	<pre>Ruijie(config)# vlan 7 Ruijie(config-vlan)# remote-span Ruijie(config-vlan)# exit Ruijie(config)# interface range Gi0/1-2 Ruijie(config-if-range)# switchport mode trunk</pre>
C	<pre>Ruijie(config)# vlan 7 Ruijie(config-vlan)# remote-span Ruijie(config-vlan)# exit Ruijie(config)# monitor session 1 remote-destination Ruijie(config)# monitor session 1 destination remote vlan 7 interface Gi 0/2</pre>

【检验方法】 分别在设备 A、设备 B 和设备 C 中执行 **show monitor** 或者 **show running** 命令查看镜像配置成功与否。

A

```
Ruijie# show monitor
sess-num: 1
span-type: SOURCE_SPAN
src-intf:
```

```
GigabitEthernet 0/1    frame-type Both
dest-intf:
GigabitEthernet 0/2
Remote vlan 7
mtp_switch on
B
!
vlan 1
  remote-span
!
!
interface GigabitEthernet 0/1
  switchport mode trunk
!
Interface GigabitEthernet 0/2
  switchport mode trunk
!
C
Ruijie# show monitor
sess-num: 1
span-type: DEST_SPAN
dest-intf:
GigabitEthernet 0/2
Remote vlan 7
mtp_switch on
```

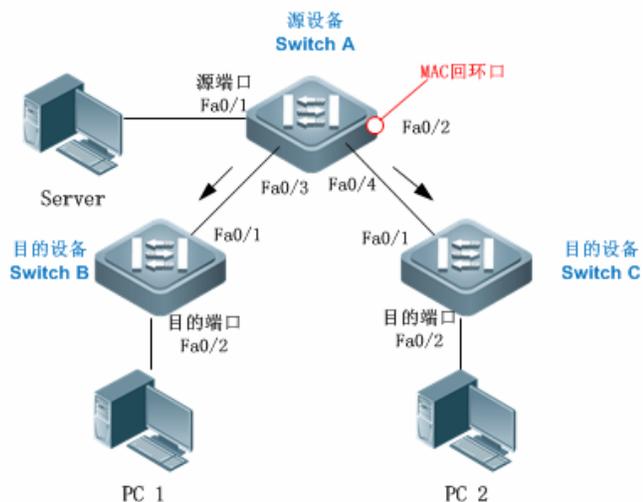
常见错误

- 源设备、中间设备、目的设备均要配置远程 VLAN 且 VID 必须一致。
- 带宽大的端口被镜像到带宽小的端口可能会造成丢包。

📌 配置一对多远程镜像

【网络环境】

图 5-7



【配置方法】

- 如上图，设备 A、设备 B 和设备 C 配置远程 VLAN。
- 在设备 A 中配置源端口、输出端口、MAC 回环口。
- 在设备 B 和设备 C 中配置目的端口。

A

```
Ruijie# configure
Ruijie(config)# vlan 7
Ruijie(config-vlan)# remote-span
Ruijie(config-vlan)# exit
Ruijie(config)# monitor session 1 remote-source
Ruijie(config)# monitor session 1 source interface fa 0/1 both
Ruijie(config)# monitor session 1 destination remote vlan 7 interface fa 0/2 switch
Ruijie(config)# interface fa0/2
Ruijie(config-if-FastEthernet 0/2)# mac-loopback
Ruijie(config-if-range)# exit
Ruijie(config)# interface range fa0/3-4
Ruijie(config-if-range)# switchport mode trunk
```

B、C

```
Ruijie(config)# vlan 7
Ruijie(config-vlan)# remote-span
Ruijie(config-vlan)# exit
Ruijie(config)# monitor session 1 remote-destination
Ruijie(config)# monitor session 1 destination remote vlan 7 interface fa 0/2
Ruijie(config)# interface fa0/1
Ruijie(config-if)# switchport mode trunk
```

【检验方法】

分别在设备 A、设备 B 和设备 C 中执行 **show monitor** 或者 **show running** 命令查看镜像配置成功与否。

A

```
Ruijie# show monitor
sess-num: 1
span-type: SOURCE_SPAN
```

```

src-intf:
FastEthernet 0/1      frame-type Both
dest-intf:
FastEthernet 0/2
Remote vlan 7
mtp_switch on
B Ruijie# show monitor
sess-num: 1
span-type: DEST_SPAN
dest-intf:
FastEthernet 0/2
Remote vlan 7
mtp_switch on
C Ruijie# show monitor
sess-num: 1
span-type: DEST_SPAN
dest-intf:
FastEthernet 0/2
Remote vlan 7
mtp_switch on

```

常见错误

- 源设备、中间设备、目的设备均要配置远程 VLAN 且 VID 必须一致。
- 带宽大的端口被镜像到带宽小的端口可能会造成丢包。
- 如果实现一对多镜像时则需要配置一个 MAC 回环口和若干输出端口。

5.6 监视与维护

清除各类信息

无。

查看运行情况

作用	命令
查看系统存在的所有镜像会话。	show monitor
查看具体的镜像会话。	show monitor session session-id

查看调试信息

 输出调试信息，会占用系统资源。使用完毕后，请立即关闭调试开关。

作用	命令
打开 SPAN 的调试开关。	debug span

6 sFlow

6.1 概述

sFlow 是由 InMon、HP 和 FoundryNetworks 于 2001 年联合开发的一种网络监测技术,目前已经完成标准化,可提供完整的第二层到第四层信息,可以适应超大网络流量环境下的流量分析,让用户详细、实时地分析网络传输流的性能、趋势和存在的问题。

sFlow 具有如下优势：

- 支持在千兆或更高速的网络上精确地监控网络流量。
- 一个 sFlow Collector 能够监控成千上百个 sFlow Agent ,具有良好的扩展性。
- sFlow Agent 内嵌在网络设备中,成本较低。

协议规范

- sFlow Version 5 : sFlow V5 协议。
- RFC 1014 : sFlow 使用的数据标准。

6.2 典型应用

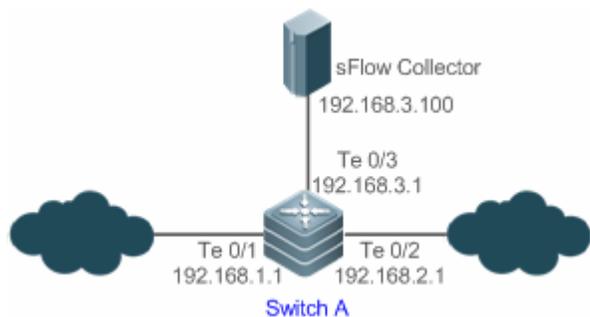
典型应用	场景描述
监控局域网流量	将设备作为 sFlow Agent ,在局域网中对接口流量进行采样 ,并将采样结果发送给 sFlow Collector 用于流量分析,以达到监控网络的目的。

6.2.1 监控局域网流量

应用场景

如图所示,启动作为 sFlow Agent 设备的交换机 SwitchA,在 Te0/1 口开启 flow 采样、counter 采样,监控 192.168.1.0 网段的流量,定时或者缓冲区满时将采样结果封装成 sFlow 报文,发送给 sFlow Collector 用于分析 sFlow Agent 监控的流量。

图 6-1



功能部属

- 在 Switch A 上配置 sFlow Agent、sFlow Collector 地址
- 在 Switch A 的 Te0/1 口开启 flow 采样、counter 采样

i 支持 sflow 的服务器软件有很多,可以在 <http://www.sflow.org/products/collectors.php> 获得,其中 sflowtrend 是免费软件。

6.3 功能详解

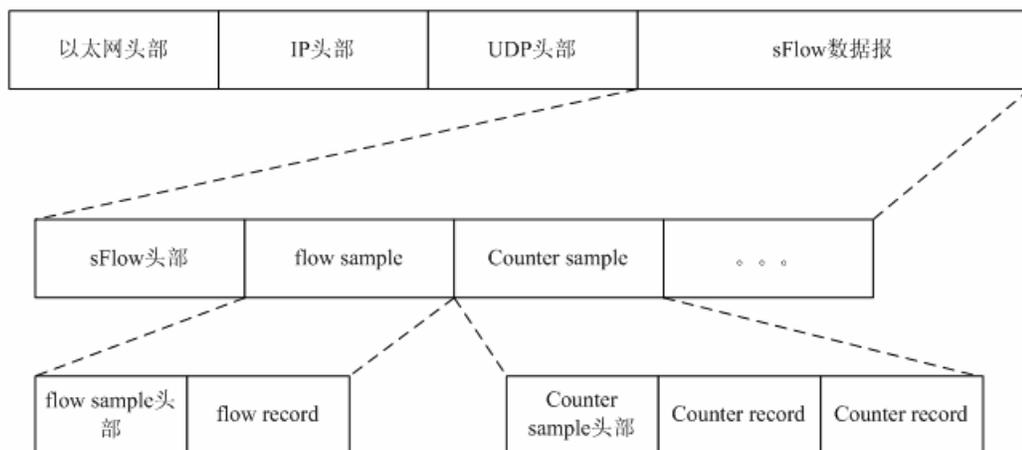
基本概念

📌 sFlow Agent

嵌入于网络设备中,通常一台网络设备可以设置成一个 sFlow Agent。sFlow Agent 可以进行 flow 采样和 counter 采样,并将采样信息封装成 sFlow 报文发送到 sFlow Collector。

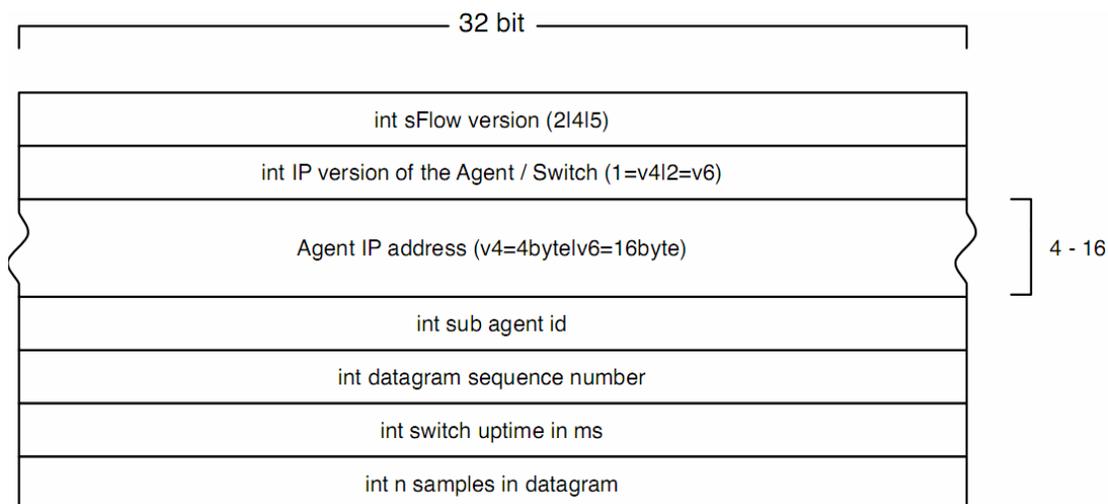
sFlow 报文采用 UDP 封装,报文格式如下图所示。**错误！未找到引用源。**

图 6-2 sFlow 报文格式



一个 sFlow 数据报可以包含一个或者多个 flow sample 和 counter sample。

图 6-3 sFlow 头部



sFlow 头部说明：

字段	说明
sFlow version	sFlow 版本号，有 2、4、5，目前我司只支持 v5
IP version of the agent/switch	SFlow Agent IP 地址的版本号
Agent IP address	SFlow Agent IP 地址
Sub agent id	Sub agent id
Datagram sequence number	sFlow 报文序列号
Switch uptime	交换机起机到当前经历了多少毫秒
n samples in datagram	报文中有多少个 samples，一个 sFlow 数据报可以包含一个或者多个 flow sample 和 counter sample

📌 sFlow Collector

接收 sFlow Agent 发送过来的 sFlow 报文，并进行分析。sFlow Collector 可以是 PC 或者服务器，在 PC 或者服务器上安装针对 sFlow 报文进行分析的软件即为一台 sFlow Collector。

📌 flow 采样

flow 采样是 sFlow Agent 设备在指定接口上按照特定采样率对报文进行采样分析，分析的内容包括：拷贝报文头部、提取以太网头部信息、提取路由信息等。

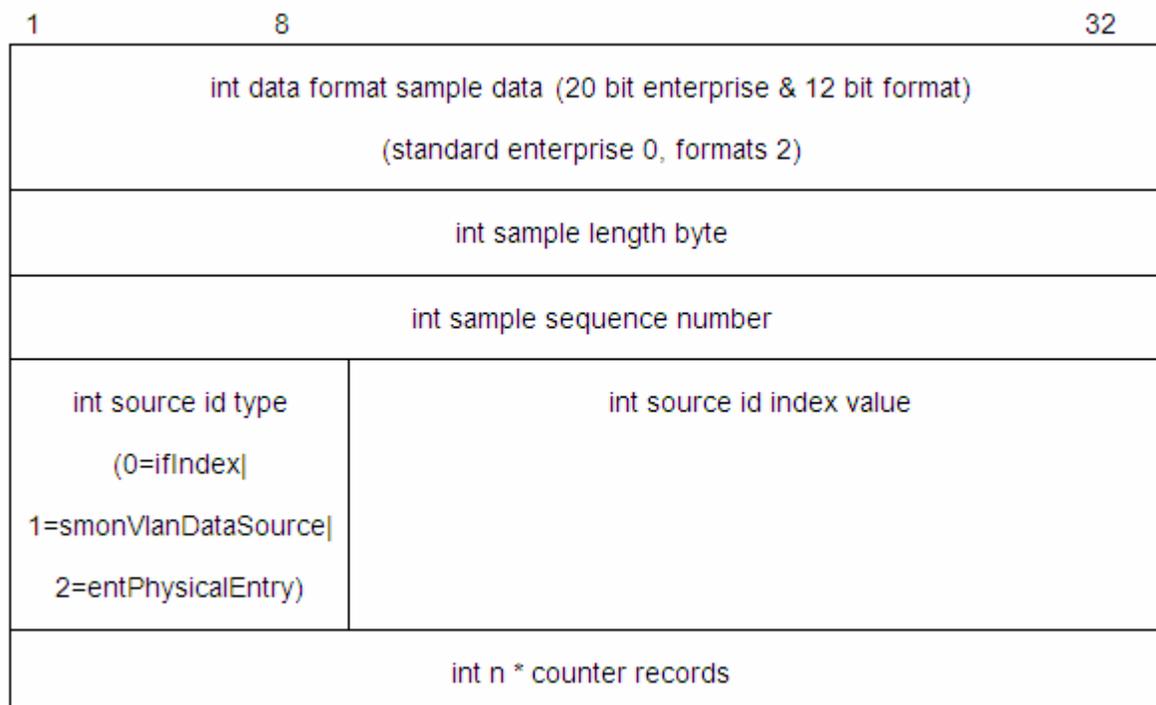
图 6-4 flow sample 头部

1	8	32
int data format sample data (20 bit enterprise & 12 bit format) (standard enterprise 0, formats 1)		
int sample length byte		
int sample sequence number		
int source id type (0=ifIndex) 1=smonVlanDataSource 2=entPhysicalEntry)	int source id index value	
int sampling rate		
int sample pool (total number of packets that could have been sampled)		
int drops (packets dropped due to a lack of resources)		
int input (SNMP ifIndex of input interface, 0 if not known)		
int output (SNMP ifIndex of output interface, 0 if not known) broadcast or multicast are handled as follows: the first bit indicates multiple destinations, the lower order bits number of interfaces		
int n * flow records		

↘ counter 采样

counter 采样是 sFlow Agent 设备周期性的获取指定接口上的统计信息、CPU 利用率。其中接口上的统计信息包括接口输入报文数、输出报文数等信息。

图 6-5 counter sample 头部



功能特性

功能特性	作用
flow 采样	对流经接口的报文进行处理，并发往 sFlow Collector 处理。
counter 采样	定时将接口的统计信息发往 sFlow Collector 处理。

6.3.1 flow采样

对流经接口的报文进行处理，并发往 sFlow Collector 处理。

工作原理

当一个报文通过某个接口时，sFlow Agent 设备根据该接口下的采样率配置对报文进行 flow 采样，包括拷贝报文的头部、提取报文的以太网头部、IP 头部、获得报文的路由信息等。最后 sFlow Agent 模块将 flow 采样结果封装成 sFlow 报文，发送到 sFlow Collector 进行分析。

6.3.2 counter采样

定时将接口的统计信息发往 sFlow Collector 处理。

工作原理

sFlow Agent 模块定时轮询接口,对于 counter 采样时间间隔到期的接口获得该接口的统计信息,然后将统计信息封装成 sFlow 报文,发送到 sFlow Collector 进行分析。

6.4 配置详解

配置项	配置建议 & 相关命令	
配置sFlow基本功能	 必须配置。用于建立 sFlow Agent 和 sFlow Collector 连接通信。	
	sflow agent {address interface}	配置 sFlow Agent 地址
	sflow collector collector-id destination	配置 sFlow Collector 地址
	 必须配置。用于开启 flow 采样和 counter 采样。	
	sflow counter collector	配置 counter 采样输出 sFlow Collector 的 ID
	sflow flow collector	配置 flow 采样输出 sFlow Collector 的 ID
	sflow enable	配置接口 sFlow 采样使能,同时开启 counter 采样和 flow 采样
配置sFlow可选参数	 可选配置。用于修改 sFlow 相关参数属性。	
	sflow collector collector-id	配置输出 sFlow 报文最大长度
	sflow max-datagram-size	配置 counter 采样时间间隔
	sflow flow max-header	配置 flow 采样拷贝报文头的最大长度
	sflow sampling-rate	配置 flow 采样的采样率

6.4.1 配置sFlow基本功能

配置效果

- sFlow Agent 设备同 sFlow Collector 之间可以通信。
- 根据缺省的采样率对流经接口的报文进行处理,并发往 sFlow Collector 处理。
- 根据缺省的采样间隔定时将接口的统计信息发往 sFlow Collector 处理。

注意事项

- 支持在物理口和聚合口下配置 flow 采样。
- 为使 sFlow Collector 可以对 flow 采样的结果进行分析,sFlow Agent 设备上必须配置 sFlow Collector 的 IP 地址。

配置方法

配置 sFlow Agent 地址

- 必须配置。
- 使用 **sflow agent { address | interface }**可配置 sFlow Agent 地址。
- sFlow Agent 地址必须是有效的地址。不能是组播、广播地址等。建议使用 sFlow Agent 设备的 IP 地址。

【命令格式】 **sflow agent { address { *ip-address* } | { interface { *interface-name* } }**

【参数说明】 **address** : 通过地址的形式配置 sFlow Agent 地址。

ip-address : sFlow Agent IPv4 地址。

interface : 通过接口的形式配置 sFlow Agent 地址

interface-name : 配置了 IPv4 地址的接口名称。

【缺省配置】 缺省未配置

【命令模式】 全局模式

【使用指导】 该命令用于配置填充在输出报文的 Agent ip address 字段，未配置报文将无法输出。地址只能为主机地址，当配置为非主机地址，比如组播地址、广播地址，将提示配置失败。建议配置的地址为 sFlow Agent 设备上的地址。

配置 sFlow Collector 地址

- 必须配置。
- 使用 **sflow collector** 命令可以配置 sFlow Collector 地址。
- sFlow Collector 地址必须是有效的地址。不能是组播、广播地址等。sFlow Collector 必须存在并且路由可达。

【命令格式】 **sflow collector collector-id destination { *ip-address* } udp-port [oob]**

【参数说明】 *collector-id* : sFlow Collector id，取值范围 1-2。

ip-address : sFlow Agent IPv4 地址，缺省未配置。

udp-port : sFlow Collector 监听端口号。

oob : 采样报文从管理口输出，缺省未配置。

【命令模式】 全局模式

【使用指导】 该命令用于配置 sFlow Collector 地址，地址只能为主机地址，当配置为非主机地址，比如组播地址、广播地址，将提示配置失败。sFlow Collector 在配置的端口号上监听 sFlow 报文。

配置 flow 采样输出 sFlow Collector 的 ID

- 必须配置。
- 使用 **sflow flow collector** 命令可以启动或关闭接口上 flow 采样的输出 sFlow Collector 功能。
- 必须在接口上启用 flow 采样输出 sFlow Collector 功能，才会将接口上的 flow 采样输出到 sFlow Collector。并且 sFlow Collector 必须是存在、可达的，sFlow Agent 设备上必须已经配置了相应 sFlow Collector 的 IP 地址。

- 【命令格式】 **sflow flow collector collector-id**
- 【参数说明】 *collector-id* : sFlow Collector id , 取值范围 1-2。
- 【缺省配置】 接口上 flow 采样的输出 sFlow Collector 功能关闭。
- 【命令模式】 接口模式
- 【使用指导】 该命令支持在物理口和聚合口下配置。
对应的 sFlow Collector 只有配置 IP 地址 , sFlow 报文才能输出。

▾ 配置 counter 采样输出 sFlow Collector 的 ID

- 必须配置。
- 使用 **sflow counter collector** 命令可以启动或关闭接口上 counter 采样的输出 sFlow Collector 功能。
- 必须在接口上启用 counter 采样输出 sFlow Collector 功能 , 才会将接口上的 counter 采样输出到 sFlow Collector。并且 sFlow Collector 必须是存在、可达的 , sFlow Agent 设备上必须已经配置了相应 sFlow Collector 的 IP 地址。

- 【命令格式】 **sflow counter collector collector-id**
- 【参数说明】 *collector-id* : sFlow Collector id , 取值范围 1-2。
- 【缺省配置】 接口上 counter 采样的输出 sFlow Collector 功能关闭。
- 【命令模式】 接口模式
- 【使用指导】 该命令支持在物理口和聚合口下配置。
对应的 sFlow Collector 只有配置 IP 地址 , sFlow 报文才能输出。

▾ 开启 counter 采样和 flow 采样

- 必须配置。
- 使用 **sflow enable** 命令可以开启接口上的 flow 采样功能以及 counter 采样功能。
- 开启 flow 采样可能影响接口的转发性能。

- 【命令格式】 **sflow enable**
- 【参数说明】 -
- 【缺省配置】 接口上 flow 采样功能关闭
- 【命令模式】 接口模式
- 【使用指导】 该命令支持在物理口和聚合口下配置。

检验方法

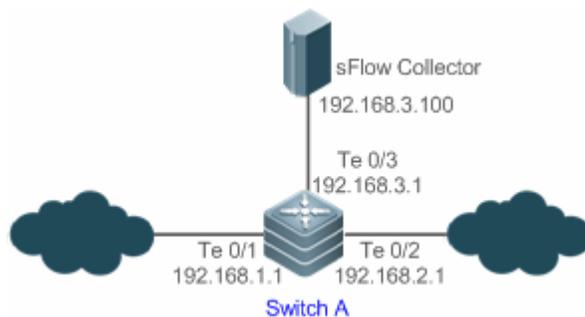
- 利用命令 **show sflow** 显示 sFlow 配置信息 , 查看显示信息是否与配置一致。

配置举例

▾ 配置 sFlow Agent 的 flow 采样和 counter 采样

【网络环境】

图 6-6



如图所示，启动作为 sFlow Agent 设备的交换机 SwitchA，在 Te0/1 口开启 flow 采样、counter 采样，监控 192.168.1.0 网段的流量，定时或者缓冲区满时将采样结果封装成 sFlow 报文，发送给 sFlow Collector 用于分析 sFlow Agent 监控的流量。

【配置方法】

- 配置 sFlow Agent 地址为 192.168.1.1。
- 配置 sFlow Collector 1 地址为 192.168.3.100，端口号为 6343。
- 在接口 TenGigabitEthernet 0/1 配置 flow 采样、counter 采样输出到 sFlow Collector 1，并使能该接口的 sFlow 采样功能。

Switch A

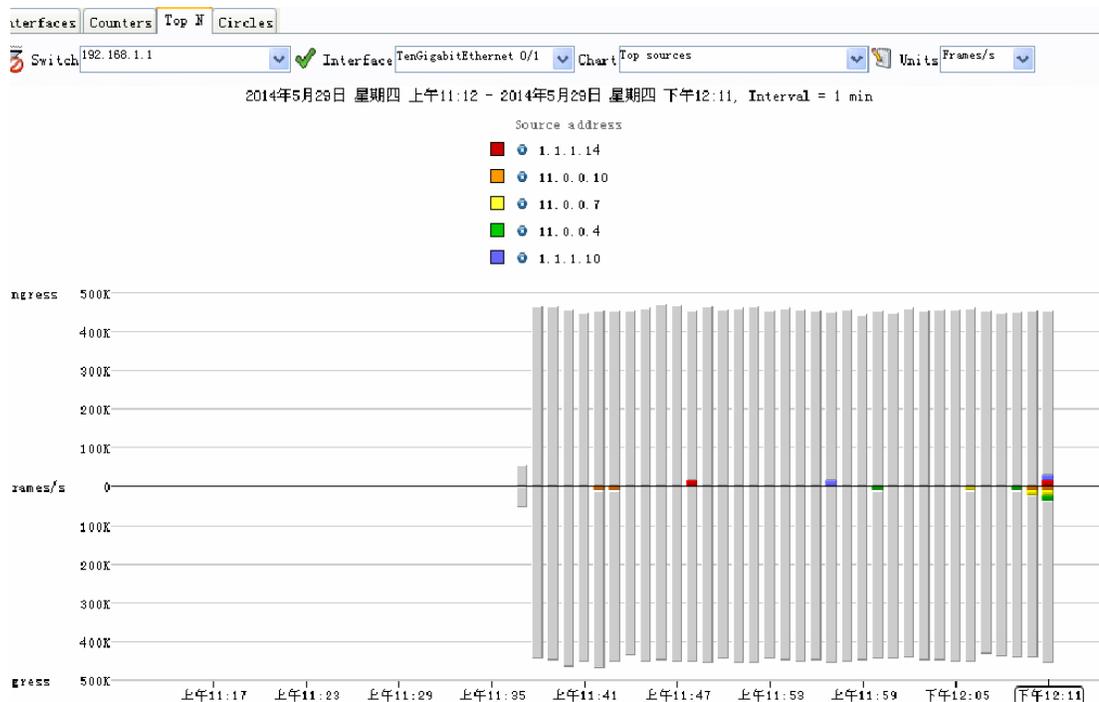
```
Ruijie# configure terminal
Ruijie(config)# sflow agent address 192.168.1.1
Ruijie(config)# sflow collector 1 destination 192.168.3.100 6343
Ruijie(config)# interface TenGigabitEthernet 0/1
Ruijie(config-if-TenGigabitEthernet 0/1)# sflow flow collector 1
Ruijie(config-if-TenGigabitEthernet 0/1)# sflow counter collector 1
Ruijie(config-if-TenGigabitEthernet 0/1)# sflow enable
Ruijie(config-if-TenGigabitEthernet 0/1)# end
```

【检验方法】

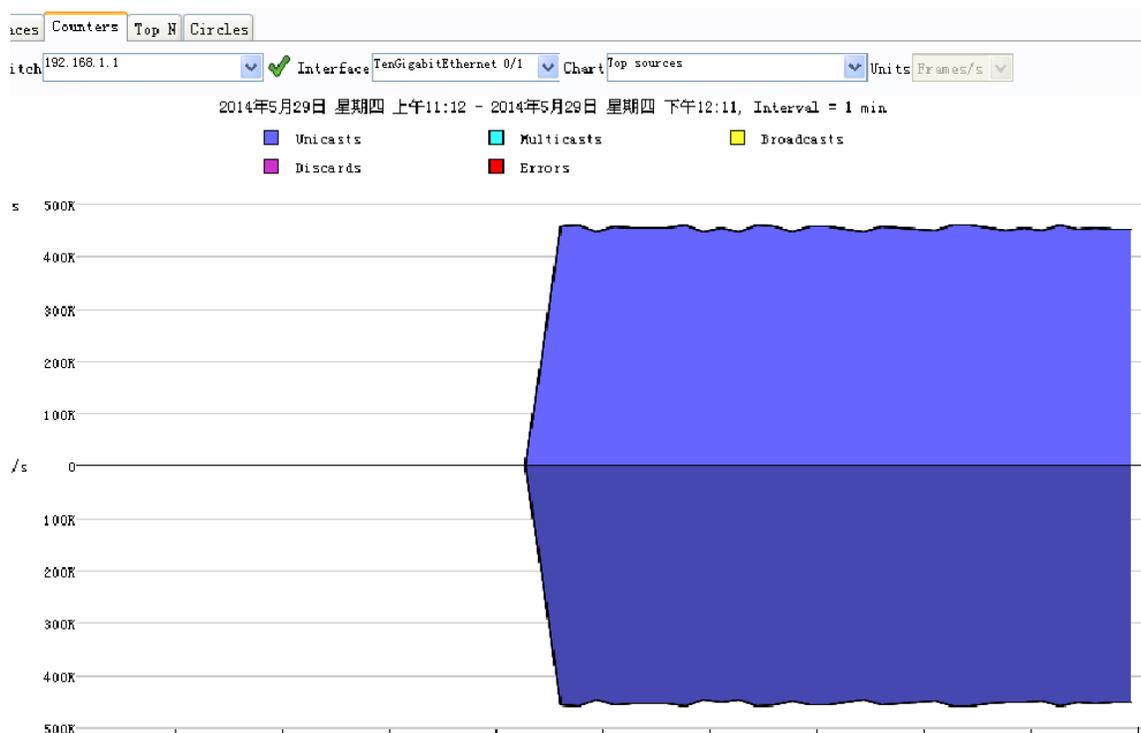
通过 **show sflow** 查看显示信息是否与配置一致。

```
Ruijie# show sflow
sFlow datagram version 5
Global information:
Agent IP: 192.168.1.1
sflow counter interval:30
sflow flow max-header:64
sflow sampling-rate:8192
Collector information:
ID  IP                Port Size VPN
1   192.168.3.100    6343 1400
2   NULL              0     1400
Port information
Interface                CID  FID  Enable
TenGigabitEthernet 0/1  1    1    Y
```

sflowtrend 上的显示



上图是 sflowtrend 的 TOP N 界面，用于显示 flow 采样结果，显示了流量最大前 5 个源 IP 地址，总流量为入方向 450KPPS 左右、出方向 450KPPS，与实际流量一致。



上图是 sflowtrend 的 counters 界面，用于显示 counter 采样结果，入方向 450KPPS，出方向 450KPPS,且所有报文均为单播报文。

常见错误

-

6.4.2 配置sFlow可选参数

配置效果

通过修改 sFlow 相关参数属性的缺省值，可调节数据采样的精确度。

注意事项

- 采样率配置太低可能影响转发性能。

配置方法

配置输出 sFlow 报文最大长度

- 可选配置。
- 使用 **sflow collector** 命令可以配置 sFlow 报文载荷的长度，不包括以太网头部、IP 头部、UDP 头部。sflow 报文中可以封装 1 个或者多个 flow 采样和 counter 采样。sflow 输出报文最大长度配置会导致处理同样数量的 flow 采样和 counter 采样输出的 sflow 报文个数可能不一样。如果配置超过 MTU，输出的 sflow 报文会被分片。

【命令格式】 **sflow collector collector-id max-datagram-size datagram-size**

【参数说明】 *collector-id* : sFlow Collector id，取值范围 1-2。

max-datagram-size datagram-size : 输出 sFlow 报文最大长度，取值范围 200-9000。

【缺省配置】 缺省值 1400

【命令模式】 全局模式

【使用指导】 -

配置 sFlow flow 采样的采样率

- 可选配置。
- 使用 **sflow sampling-rate** 命令可以配置全局 flow 采样的采样率。
- flow 采样的采样率配置可能影响到 sflow 采样的准确性，采样率越小，准确度越高，同时也越消耗 CPU，从而有可能影响到接口转发性能。

【命令格式】 **sflow sampling-rate rate**

【参数说明】 *rate* : sFlow flow 采样的采样率，即每 *rate* 个报文采样一个报文，取值范围为 4096-16777215。

【缺省配置】 全局的 flow 采样的采样率为 8192。

【命令模式】 全局模式

【使用指导】 该命令配置了 sFlow flow 采样的全局采样率，所有接口的 sFlow flow 采样都使用这个采样率。

配置 flow 采样拷贝报文头的最大长度

- 可选配置。
- 使用 **sflow flow max-header** 命令可以配置全局 flow 采样拷贝报文的长度。
- 用户可以通过该配置修改输出到 sFlow Collector 的报文信息。例如，用户关心 IP 头部，则可以配置长度为 56 字节。封装 flow 采样时将采样报文的前 56 个字节复制到 sflow 报文中。

【命令格式】 **sflow flow max-header length**

【参数说明】 *Length*：拷贝报文头最大长度，取值范围 18-256，缺省值 64，单位字节。

【缺省配置】 全局的 flow 采样拷贝报文的长度为 64 字节。

【命令模式】 全局模式

【使用指导】 配置在进行报文内容拷贝时，从原始报文的头部开始，允许拷贝的最大字节数。拷贝的内容会记录在生成的采样样本中。

配置采样时间间隔

- 可选配置。
- 使用 **sflow counter interval** 命令可以配置全局的 counter 采样时间间隔。
- 使能 counter 采样的接口每隔采样时间间隔就会将接口的统计信息发送到 sflow collector。

【命令格式】 **sflow counter interval seconds**

【参数说明】 *seconds*：时间间隔，取值范围 3-2147483647，单位为秒，缺省值 30。

【缺省配置】 全局的 counter 采样时间间隔为 30 秒。

【配置模式】 全局模式

【使用指导】 该命令配置了 sFlow counter 采样的全局时间间隔，所有接口的 sFlow counter 采样都使用这个采样间隔。

检验方法

- 在 sFlow Collector 上观察是否收到内容为 flow 采样的 sFlow 报文。
- 利用命令 **show sflow** 显示 sFlow 配置信息，查看显示信息是否与配置一致。

配置举例

配置 sflow 可选配置

【网络环境】 参见图 6-6

- 【配置方法】
- 在全局模式下配置采样率为 4096。
 - 在全局模式下配置拷贝报文头的前 128 个字节。
 - 在全局模式下配置采样间隔为 10。

```
Ruijie# configure terminal
Ruijie(config)# sflow sampling-rate 4096
Ruijie(config)# sflow flow max-header 128
Ruijie(config)# sflow counter interval 10
```

【检验方法】 使 TenGigabitEthernet 0/1 有流量经过。

- 在 sFlow Collector 1 中观察 TenGigabitEthernet 0/1 是否有流量
- 通过 **show sflow** 查看显示信息是否与配置一致

```
Ruijie# show sflow
sFlow datagram version 5
Global information:
Agent IP: 10.10.10.10
sflow counter interval:10
sflow flow max-header:128
sflow sampling-rate:4096
Collector information:
ID  IP                      Port Size VPN
1   192.168.2.100          6343 1400
2   NULL                   0    1400
Port information
Interface          CID  FID  Enable
TenGigabitEthernet 0/1  0    1    Y
```

常见错误

-

6.5 监视与维护

清除各类信息

-

查看运行情况

作用	命令
查看 sFlow 配置。	show sflow

查看调试信息
