

# ARISTA

## **Arista User Guide**

**Written by Suntengfei**

**Version 1.0**

# 目录

1. Arista 部署模式.....	5
1.1 零接触部署模式.....	5
1.2 手动配置模式.....	5
2. 交换机基本配置.....	5
2.1 管理口配置.....	5
2.2 配置缺省网关路由.....	6
2.3 管理连接方式.....	6
2.3.1 console 配置.....	6
2.3.2 ssh 配置.....	6
2.3.3 telnet 配置.....	6
3. 设备设置恢复与升级.....	6
3.1 破解 Enable 密码.....	6
3.2 恢复 startup-config 出厂设置.....	7
3.3 恢复默认 EOS 与 startup-config 出厂配置.....	7
3.4 从 USB 恢复 EOS 与 startup-config 出厂配置.....	7
3.5 升级 EOS.....	8
4. 基本命令介绍.....	8
4.1 自定义命令.....	8
4.2 快捷键移动光标.....	9
4.3 快捷键历史记录查询.....	9
4.4 计划命令.....	9
4.5 事件触发命令.....	9
4.6 交换机操作平台.....	10
4.7 Linux 命令.....	11
4.8 显示命令行数.....	11
5. 管理交换机.....	12
5.1 配置域.....	12
5.2 配置时钟.....	12
5.3 配置 Banner.....	13

5.4	配置事件监视 .....	13
6.	启动交换机 .....	14
6.1	配置文件 .....	14
6.2	Boot-config .....	14
6.3	ZTP 配置 .....	15
6.4	重启配置 .....	16
6.5	状态配置 .....	16
7.	交换机端口 .....	17
7.1	转发模式配置 .....	17
7.2	QSFP 配置 .....	18
7.2.1	端口改为一个 40G 的端口 .....	19
7.2.2	端口改为四个 10G 的端口 .....	19
7.3	QSFP-SFP 端口配置 .....	19
8.	Port Channles 与 LACP .....	20
8.1	配置 MLAG .....	20
8.1.1	配置交换机间的 port-channel .....	20
8.1.2	创建 SVI .....	21
8.1.3	配置交换机对端连接 .....	22
8.1.4	级联二层交换机 .....	23
	假如有一台二层交换机接入，则相应的 MLAG 配置 .....	23
8.1.5	级联服务器 .....	23
8.2	其他命令 .....	24
8.2.1	将端口从 channel group 中移除 .....	24
8.2.2	LACP 配置命令 .....	24
9.	ACL 访问控制列表 .....	25
9.1.1	支持功能 .....	25
9.1.2	不支持功能 .....	26
10.	VRRP 与 VARP .....	26
10.1	VRRP 介绍 .....	26
10.1.1	配置 VRRP .....	27

10.2	VARP .....	29
11.	Spanning Tree Protocol .....	31
11.1	STP 协议 .....	31
11.2	Rapid Spanning Tree Protocol (RSTP) .....	32
11.3	Rapid Per-VLAN Spanning Tree Protocol (Rapid-PVST) .....	32
11.4	Multiple Spanning Tree Protocol (MSTP) .....	32
11.4.1	MTS Regions .....	32
11.4.2	MTS Instances.....	32
11.5	不同版本 SPT 相互兼容性 .....	32
11.6	Switchport interface pairs .....	33
11.7	Spanning tree instance 的结构 .....	34
11.7.1	Root bridge 与 designated birdge .....	34
11.7.2	Root port 与 designated port.....	35
11.7.3	端口的状态 .....	36
11.7.4	端口类型 .....	36
11.7.5	链路类型 .....	36
11.7.6	BPDU.....	36
11.7.7	Bridge timer .....	37
11.8	配置 Spanning tree.....	37
11.8.1	MST 配置 .....	37
11.8.2	RST .....	38
11.8.3	Rapid-PVST.....	38
11.8.4	Switchport backup 模式.....	38
11.8.5	配置 Spanning tree instance .....	39

## 1. Arista 部署模式

Arista 交换机提供两种部署模式，即零接触部署模式与手动配置模式。

查看命令

Show version 查看软件版本，内存，硬件信息，设备型号等

Show inventory 查看设备上插入的硬件配件，如电源，模块等。

Show platform 查看交换机操作平台。交换机操作平台可能会根据不同型号而改变。

### 1.1 零接触部署模式

零接触部署方式(ZTP zero touch provisioning)，默认开启（如用于企业内部核心交换机，建议关闭该模式）。该方式通过一个启动文件（startup-config）或者来自于 DHCP 的启动脚本实现自动部署。一般用于一次性进行较多台交换机的部署与配置，以减轻管理人员部署的难度以及减少部署时间。

零接触部署命令：

localhost>zerotouch cancel /关闭零接触部署模式

localhost>zerotouch /开启零接触部署模式

### 1.2 手动配置模式

手动配置即由管理员手动进行配置。在启用手动配置之前，必须先关闭零接触部署方式。

## 2. 交换机基本配置

### 2.1 管理口配置

初次配置 Arista 时，登录用户名为 admin,无密码。

Switch(config)username admin secret arista /创建用户 admin，密码为 arista

Switch(config)#interface management 1 /进入管理口 1

Switch(config)#ip address 192.0.2.8/24 /配置管理口 IP 地址

Switch(config)#end /返回特权模式

Switch#write memory /保存配置

## 2.2 配置缺省网关路由

```
Switch(config)#ip route 0.0.0.0/0 192.0.2.1
```

```
Switch(config)#end
```

```
Switch#write memory
```

## 2.3 管理连接方式

Arista 交换机提供 Console、ssh、telnet 三种连接方式。默认情况下，Arista 开启 ssh 与 console，关闭 telnet。如果需要使用 telnet，则需要手动开启（注且只有 telnet 才支持 shutdown 命令）。

### 2.3.1 console 配置

console 默认开启，可使用以下命令配置相关参数。

```
Switch(config)#management console /进入 console
```

```
switch(config-mgmt-console)#idle-timeout 180 /配置 timeout 时间（即多久时间无操作，自动退出。SSH 与 telnet 模式配置方法相同）
```

### 2.3.2 ssh 配置

```
Switch(config)#management ssh
```

```
switch(config-mgmt-ssh)#idle-timeout 180 /配置 timeout 时间
```

### 2.3.3 telnet 配置

```
switch(config)#management telnet /进入 telnet
```

```
switch(config-mgmt-telnet)#no shutdown /开启 telnet
```

```
switch(config-mgmt-telnet)#shutdown /关闭 telnet
```

```
switch(config-mgmt-telnet)#idle-timeout 180 /配置 timeout 时间
```

## 3. 设备设置恢复与升级

### 3.1 破解 Enable 密码

1. 重启物理设备
2. Ctrl+c
3. About#cd /mnt/flash 将 active directory 修改为/mnt/flash
4. 输入 about 密码，及初始时的 username admin 对应的密码

If the About password is not known, enter an empty password three times, after which the CLI displays: Type "fullrecover" and press Enter to revert /mnt/flash to factory default state, or just press Enter to reboot:

5. About#vi startup-config 在 Vi 中打开 startup-config
6. enable secret 5 \$1\$dBXo2KpF\$Pd4XYLpI0ap1ZaU7glG1w/看到 enable 密码行，将光标定位到该行前，按 d 删除该行
7. Esc 然后 “shift+: ” 输入 wq 或 z 保存并退出 vi
8. About#exit 重启设备

### 3.2 恢复 startup-config 出厂设置

1. Switch#reload 重启设备
2. Ctrl+c 进入 boot process
3. 输入 About 密码
4. About#cd /mnt/flash
5. About#mv startup-config startup-config.old / 修改 Startup-config 的名称为 startup-config.old
6. About#exit /重启设备
7. localhost>zerotouch cancel /取消 ZTP 部署模式
8. switch#delete startup-config.old 删除老的配置

### 3.3 恢复默认 EOS 与 startup-config 出厂配置

1. Switch#reload 重启设备
2. Ctrl+c 进入 boot process
3. 输入 About 密码
4. About #fullrecover
5. 输入 yes 点击回车，此操作将删除/mnt/flash 里的所有数据。并创建新的 boot-config,startup-config,EOS,swi 文件到/mnt/flash
6. About #exit /重启设备
7. localhost>zerotouch cancel /取消 ZTP 部署模式

### 3.4 从 USB 恢复 EOS 与 startup-config 出厂配置

当 console 口不可用时，可使用此种方式。首先将 EOS Image 安装到 USB

1. 用 MS-DOS 或者 FAT 文件格式化 USB
2. 在 USB 上创建 TEXT 文件，命名为 fullrecover
3. 创建 TEXT 文件，命名为 boot-config
4. 在 boot-config 文件里输入：SWI=flash:EOS.swi

5. 将 EOS image 文件拷贝到 USB 中，并重命名为 EOS.swi
6. 插入 USB 设备到交换机
7. 终端连接到 console 口查看进程信息
8. 开启设备或重启设备，交换机会删除内部 flash 的数据，并把 USB 的数据复制到内部 flash 中。
9. localhost>zerotouch cancel /取消 ZTP 部署模式

## 3.5 升级 EOS

Arista 提供多种升级模式。升级完成后需要重启设备才能生效。重启命令：switch#reload

### USB Memory

Command

```
copy usb1:/sourcefile flash:/destfile
```

Example

```
Sch#copy usb1:/EOS-4.9.2.swi flash:/EOS-4.9.2.swi
```

### FTP Server

Command

```
copy ftp:/ftp-source/sourcefile flash:/destfile
```

Example

```
Sch#copy ftp:/user:password@10.0.0.3/EOS-4.9.2.swi flash:/EOS-4.9.2.swi
```

### SCP

Command

```
copy scp://scp-source/sourcefile flash:/destfile
```

Example

```
Sch#copy scp://user:password@10.1.1.8/user/EOS-4.9.2.swi flash:/EOS-4.9.2.swi
```

### HTTP

Command

```
copy http://http-source/sourcefile flash:/destfile
```

Example

```
Sch#copy http://10.0.0.10/EOS-4.9.2.swi flash:/EOS-4.9.2.swi1.switch#copy usb1:/EOS-4.9.2.swi  
flash:/EOS-4.9.2.swi
```

## 4. 基本命令介绍

### 4.1 自定义命令

使用 Alias 命令，可创建自定义命令

#### Command Syntax 全局模式下

```
alias command_alias original_command
```

```
no alias command_alias
```



default alias command\_alias

```
Switch(config)#alias srie show running-config interface ethernet 1-5
Switch(config)#srie
interface Ethernet1
switchport access vlan 33
storm-control broadcast level 1
spanning-tree portfast
spanning-tree bpduguard enable
interface Ethernet2
switchport access vlan 33
spanning-tree portfast
interface Ethernet3
switchport access vlan 33
spanning-tree portfast
spanning-tree bpduguard enable
interface Ethernet4
interface Ethernet5
shutdown
```

## 4.2 快捷键移动光标

- Ctrl-B or the Left Arrow key: Moves the cursor back one character.
- Ctrl-F or the Right Arrow key: Moves the cursor forward one character.
- Ctrl-A: Moves the cursor to the beginning of the command line.
- Ctrl-E: Moves the cursor to the end of the command line.
- Esc-B: Moves the cursor back one word.
- Esc-F: Moves the cursor forward one word.

## 4.3 快捷键历史记录查询

Ctrl+p 或 UP 查看上一条命令

Ctrl+n 或 down 查看下一条命令

## 4.4 计划命令

可自定义自动使用某命令在某时刻。

```
switch#schedule backup-test interval 60 max-log-files 24 command show
running-config 名为 backup-test 的计划任务，每个 60 分钟，最大生成日志 24，自动使用
show running-config 命令。
```

## 4.5 事件触发命令

该命令可自定义某些事件发生时，给某个地址发送邮件。如端口状态变化，IP 地址变化，

Switch#show event-handler 查看

Switch#no event-handler onstartup 删除名为 onstartup 的事件

\$INTF interface name.

\$OPERSTATE current operational status of the specified interface.

\$IP-PRIMARY current primary IP address of the specified interface.

Trigger 触发事件条件命令

- EVENT event which will trigger the configuration mode event handler. Values include:
  - onboot triggers when the system reboots, or when you exit event-handler configuration mode. This option takes no further arguments, and passes no environment variables to the action triggered.
  - onintf INTERFACE CHANGE triggers when a change is made to the specified interface.
- INTERFACE the triggering interface. Values include:
  - ethernet number Ethernet interface specified by number.
  - loopback number loopback interface specified by number.
  - management number management interface specified by number.
  - port-channel number channel group interface specified by number.
  - vlan number VLAN interface specified by number.
- CHANGE the change being watched for in the triggering interface. Values include:
  - ip triggers when the IP address of the specified interface is changed.
  - operstatus triggers when the operational status of the specified interface changes

例如：

switch(config)#event-handler eth\_4 /创建 eth\_4 的事件

switch(config-event-eth\_4)#action bash email x@yz.com -s "Et4 \$OPERSTATE" /  
发送 Et4 \$OPERSTATE 到 emailx@yz.com

switch(config-event-eth\_4)#trigger onintf ethernet 4 operstatus /触发条件 eth4 接  
口的状态变化

switch(config-event-eth\_4)#delay 60

switch(config-event-eth\_4)#exit

注意：要发送邮件必须提前在交换机上配置邮件。

## 4.6 交换机操作平台

Switch#Show platform

- DCS-7048T Petra
- DCS-7048T-A Petra
- DCS-7050Q-16 Trident
- DCS-7050S-52 Trident
- DCS-7050S-64 Trident
- DCS-7050T-52 Trident
- DCS-7050T-64 Trident
- DCS-7120T-4S FM4000
- DCS-7124FX FM4000
- DCS-7124S FM4000
- DCS-7124SX FM4000
- DCS-7140T-4S FM4000
- DCS-7148S FM4000
- DCS-7148SX FM4000
- DCS-7504 Petra
- DCS-7508 Petra

## 4.7 Linux 命令

系统是基于 Linux 开发的，允许进入底层的 linux

Switch#bash /进入 linux

[admin@Switch ~]\$ logout /退出 linux 或使用 Ctrl+D 或 exit

在 Linux 使用的是目录结构，支持如下命令：

- delete: Delete a file or directory tree.
- copy: Copy a file.
- more: Display the file contents.
- diff: Compares the contents of files located at specified URLs.
- rename: Rename a file
- cd: Change the current working directory.
- dir: Lists directory contents, including files and subdirectories.
- mkdir: Create a directory.
- rmdir: Remove a directory.
- pwd: Display the current working directory.

## 4.8 显示命令行数

默认情况下，所有的 Console 会话，terminal length 都是 disable 的。

switch#terminal length 10 /显示 10 行

switch#no terminal length /不显示

switch(config)#terminal length 0 /不显示

switch#terminal monitor

switch#no terminal monitor

switch#default terminal monitor

## 5. 管理交换机

### 5.1 配置域

Switch(config)#hostname sales1 /配置交换机名称

Switch(config)#ip domain-name aristanetworks.com /配置域名

Switch(config)#ip name-server 10.1.1.24 10.1.1.25 172.17.0.22 /配置 DNS 服务器地址，最多支持 3 个地址

### 5.2 配置时钟

Switch#clock set 08:15:24 26 April 2010 /配置交换机时间

Switch(config)>show clock

Switch(config)#clock timezone /配置时区

交换机支持与 NTP 服务器同步事件

Switch(config)#ntp server local-NTP

Switch(config)#ntp server 172.16.0.23 Prefer

Switch(config)#ntp server 172.16.0.25

Switch(config)#show ntp status /查看 NTP 连接状况

Switch(config)#show ntp associations/查看 NTP 服务器数据配置

### 5.3 配置 Banner

交换机支持两种 Banner,一种是 login banner , 另一种是 motd banner。

Login banner

```
Switch(config)#banner login
```

This is the first line of banner text. Eof /EOF 表示完成文本输入 , 退出标语

### 5.4 配置事件监视

事件监视会监视 mac、route、arp 表。默认情况下是开启的。当新的事件被通知时 , 旧的数据将被删除 , 新的数据将会重建。

注意 : buffer 是可循环的 , 即当 buffer 满时 , 新的数据会自动覆盖旧的数据 , 从头写起。

```
Switch(config)#no event-monitor all /关闭事件监控
```

```
Switch(config)#event-monitor route /开启事件监控
```

```
Switch#event-monitor clear /清除所有事件记录 , 如果有备份也会被清除。
```

```
Switch(config)#event-monitor buffer max-size 48 /调整 buffer 大小为 48KB。
```

Buffer 范围是 6-60KB , 默认是 32KB。

```
Switch(config)#event-monitor backup path sw-event.log /备份事件监控日志到 sw-event.log
```

```
Switch(config)#dir /查看文件
```

```
Switch(config)#event-monitor backup max-size 4 /定义备份大小
```

```
Switch(config)#event-monitor sync /该命令会将 buffer 与 backup 同步到一个叫做 SQLite 的文件夹里。 /tmp/eventmon.db
```

查看日志命令

```
switch#show event-monitor mac
```

```
switch#show event-monitor arp
```

```
switch#show event-monitor route
```

## 6. 启动交换机

交换机支持 USB，但不支持 NTFS 格式的，仅支持 FAT 或 VFAT。

交换机启动进程步骤如下：

加载 EOS 镜像文件

启动交换机进程

执行自检

修复接口设定

配置其他网络参数

### 6.1 配置文件

配置文件包括 boot-config,running-config,startup-config。

### 6.2 Boot-config

包括被加载的镜像放置的位置与镜像名称。Boot-config 是 ASCII 文件，用来配置 console，定位 EOS 镜像，指定初始网络设定。

使用 SWI 命令可以指定需要加载的 EOS 镜像的位置与名称

#### **BOOT system 命令**

```
(config)#more boot-config
```

```
(config)#show boot-config
```

Switch(config)#boot system usb1:EOS1.swi /启动时通过 USB

Switch(config)# boot system flash:EOS.swi

switch(config)#boot secret 0 xr19v /设置 BOOT 密码

— 0 the password is clear text. Equivalent to the <no parameter> case.

— 5 the password is an md5 encrypted string.

switch(config)#no boot secret

switch(config)#boot console speed 4800/ 定义 console 口速率，默认是 9600，支持其他速率 1200, 2400, 4800, 9600, 19200, or 38400

### 6.3 ZTP 配置

配置启动方式

SWI=flash:EOS.swi (flash drive location)

— SWI=usb1:/EOS1.swi (usb drive location)

— SWI=file:/tmp/EOExp.swi (switch directory location)

— SWI=/mnt/flash/EOS.swi

— SWI=http://foo.com/images/EOS.swi

— SWI=ftp://foo.com/images/EOS.swi

— SWI=tftp://foo.com/EOS.swi

— SWI=nfs://foo.com/images/EOS.swi

配置网络

NETDEV=mgmt1、

NETAUTO=dhcp

NETIP=10.12.15.10

NETMASK=255.255.255.0

NETGW=10.12.15.24

NETDOMAIN=mycompany.com

NETDNS=10.12.15.13

## 6.4 重启配置

支持计划重启

switch#reload in 12:00 /12 个小时后重启

switch#reload at 12.00/ 在 12 点重启

switch#no reload/ 关闭计划重启

switch#reload reason this reload is planned /增加重启原因

switch#show reload

## 6.5 状态配置

所有的状态都有两种 action , 一个是 shutdown,一个是 ignore。

Switch(config)#environment overheat action ignore

Switch(config)#environment insufficient-fans action ignore

Switch(config)#environment fan-speed override 50 /有效值 30-100

Switch>show environment temperature /查看温度



- **Ok** All sensors report temperatures below the alert threshold.
- **Overheating** At least one sensor reports a temperature above its alert threshold.
- **Critical** At least one sensor reports a temperature above its critical threshold.
- **Unknown** The switch is initializing.
- **Sensor Failed** At least one sensor is not functioning.

Switch>show environment cooling /查看风扇

- **Ok** no more than one fan has failed or is not inserted.
- **Insufficient fans** more than one fan has failed or is not inserted. This status is also displayed if fans with different airflow directions are installed. The switch shuts down if the error is not resolved.

风扇报错

- **OK** The fan is operating normally.
- **Failed** The fan is not operating normally.
- **Unknown** The system is initializing.
- **Not Inserted** The system is unable to detect the specified fan.
- **Unsupported** The system detects a fan that the current software version does not support.

Switch>show environment power /查看电源

Switch>show environment all /查看所有

## 7. 交换机端口

x	y	z
media type or wavelength, if media type is fiber	PHY encoding type	Number of WWDM wavelengths or XAUI Lanes
C = Copper (twin axial) T = Twisted Pair S = Short (850 nm) L = Long (1310 nm) E = Extended (1550 nm) Z = Ultra extended (1550 nm)	R = LAN PHY (64B/66B) X = LAN PHY (8B/10B) W = WAN PHY(*) (64B/66B)	If omitted, value = 1 (serial) 4 = 4 WWDM wavelengths or XAUI Lanes

Table 8-1 10GBASE-xyz Interpretation

### 7.1 转发模式配置

转发模式包括 cut-through (在帧完全接收前就转发) 和 store-and-forward (等帧完全接收后再转发)

Unicast addresses: the I/G bit is 0: 1234.1111.1111 is a unicast MAC address.

- Multicast address: the I/G bit is 1: 1134.1111.1111 is a multicast MAC address.

broadcast address is always FFFF.FFFF.FFFF

switch(config)#switch forwarding-mode store-and-forward /定义转发模式为存储转发

switch(config)#show switch forwarding-mode /查看转发模式

switch(config)#interface ethernet 5-7,10 /进入端口配置

switch(config-if-Et7)#mac-address 001c.2804.17e1 /配置接口的 MAC 地址。此处配置的是逻辑 MAC 地址，并不是厂商浇筑在网卡上的地址。

交换机最多支持 384 个端口，一个 48 口的交换机有 6 个 PetraA ASIC 芯片，每个芯片控制 8 个端口。默认情况下，如下显示：

- chip 0 references ports 1 through 8
- chip 1 references ports 9 through 16
- chip 2 references ports 17 through 24
- chip 3 references ports 25 through 32
- chip 4 references ports 33 through 40
- chip 5 references ports 41 through 48

## 7.2 QSFP 配置

交换机支持 QSFP 模块。该模块可定义为一个 40G 的端口或 4 个 10G 的端口。进行 QSFP 端口修改时，端口转发会被打断，需要 1 分钟以上的时间重启端口。

QSFP+ modules are supported on these Arista switches:

- DCS-7050S-64: interfaces 49–52 (four interfaces).
- DCS-7050T-64: interfaces 49–52 (four interfaces).
- DCS-7050Q-16: interfaces 1–16 (16 interfaces).

### 7.2.1 端口改为一个 40G 的端口

```
switch(config)#interface ethernet 49/1
```

```
switch(config-if-Et49/1)#speed forced 40gfull /将端口定义为一个 40G 的口
```

如果是一个 40G 的口，则显示如下 40G port configuration: 49/1 is connected or not connected; 49/2, 49/3, and 49/4 is errdisabled.

### 7.2.2 端口改为四个 10G 的端口

```
switch(config)#interface ethernet 49/1
```

```
switch(config-if-Et49/1)#no speed forced 40gfull
```

如果是四个 10G 口，则显示如下 4x10G port configuration: 49/1, 49/2, 49/3, and 49/4 status is connected or not connected.

## 7.3 QSFP-SFP 端口配置

如果交换机要同时存在 QSFP 与 SFP 模块，要使两种模块都可以使用，则必须定义一个为 QSFP（四通道的 SFP），一组 4 个的 SFP。QSFP 是端口高密度解决方案，它可以实现一转四。将一个 40G 的口转换成四个 10G 的口。

假如要使 DCS-7050Q-16 达到高密度端口的效果（最大支持 64 个 10G 端口）。交换机本身包括 16 个 QSFP 与 8 个 SFP+ 端口。要 enable 一组 4 个的 SFP+ 端口，则必须相应的 disable 一个 QSFP 端口。为了加强这种限制，则必须定义两个 hardware port-group。下表是相应的组合选项。默认情况下是 16QSFP 端口。

Port Group 1	Port Group 2	QSFP Ports enabled	SFP Ports enabled	Default
QSFP+ enabled	QSFP+ enabled	16: Ports 1-16	none	Yes
QSFP+ enabled	SFP+ enabled	15: Ports 1-15	4: Ports 21-24	No
SFP+ enabled	QSFP+ enabled	15: Ports 1-14, 16	4: Ports 17-20	No
SFP+ enabled	SFP+ enabled	14: Ports 1-14	8: Ports 17-24	No

Table 8-2 Port Group Configuration Options

默认配置：16 个 QSFP 端口 enable，SFP 端口 disable

QSFP+ interfaces in both port groups:

```
switch(config)#hardware port-group 1 select Et15/1-4
switch(config)#hardware port-group 2 select Et16/1-4
```

组合 2: 8 个 SFP 端口 enable,1-14QSFP 端 enable,15-16QSFP 端口 disable

```
switch(config)#hardware port-group 1 select Et17-20
switch(config)#hardware port-group 2 select Et21-24
```

组合 3:4 个 SFP 端口 enable,15 个 QSFP 端口 enable.

```
switch(config)#hardware port-group 1 select Et17-20
switch(config)#hardware port-group 2 select Et16/1-4
```

## 8. Port Channles 与 LACP

Port channel 单台交换机最高支持 16 个端口链路聚合。

MLAG 即跨交换机链路聚合（两台交换机），最高支持 32 个接口，320GB 带宽。传统上的交换机冗余，通过 SPT 协议将部分端口阻塞，当网络出现故障时，则由阻塞进入转发状态。Arista 支持两台或两台以上的 active/active 模式，大大提高了数据包传输量。

MLAG 建立之前首先要将 LACP 模式开启。MLAG 建立的过程包括以每秒一次发动 LACP PDU，当同步完成后端口每 30 秒彼此交换 PDU。默认 Timeout 时间是 30 秒，失效时间为 90 秒。Fallback 模式允许端口在未接受到 LACP PDU 前，自动生成 LAG。

**Table 9-1 Valid LACP Mode Combinations**

Switch 1	Switch 2	Comments
active	active	Links aggregate when LACP negotiation is successful.
active	passive	Links aggregate when LACP negotiation is successful.
passive	passive	Links aggregate without LACP.
on	—	Links aggregate without LACP.

### 8.1 配置 MLAG

#### 8.1.1 配置交换机间的 port-channel

```
switch1#conf t
```

```
switch1(config)#int ethernet 1 -2 进入 1-2 端口
```

```
switch1(config)#channle-group 101 mode active 创建 channel-group 101，开启 LACP 协议
```

```
switch1(config)#interface port-channel 101 创建 port-channel 101
```

```
switch1(config)#switchport trunk group m1peer 创建名为 m1peer 的 trunk group
```

```
switch1(config)#exit
```

```
switch2(config)#int ethernet 1 -2 进入 1-2 端口
```

```
switch2(config)#channle-group 201 mode active 创建 channel-group 201
```

```
switch2(config)#interface port-channel 201 创建 port-channel 201
```

```
switch2(config)#switchport trunk group m1peer 创建名为 m2peer 的 trunk group
```

```
switch2(config)#exit
```

### 8.1.2 创建 SVI

```
switch1#conf t
```

```
switch1(config)#vlan 4094 创建一个不常用的 vlan 作为 MLAG 间相互通讯地址
```

```
switch1(config)#trunk group m1peer 将接口应用到该 vlan
```

```
switch1(config)#ip address 172.17.0.1/30 配置 vlan 的 IP 地址
```

```
switch1(config)#no autostate
```

```
switch1(config)#exit
```

```
switch1(config)#no spanning-tree vlan 4094
```

```
switch2#conf t
```

```
switch2(config)#vlan 4094 创建一个不常用的 vlan 作为 MLAG 间相互通讯地址
```

```
switch2(config)#trunk group m2peer 将接口应用到该 vlan
```

```
switch2(config)#ip address 172.17.0.2/30 配置 vlan 的 IP 地址
```

```
switch2(config)#no autostate
```

```
switch2(config)#exit
```

```
switch2(config)#no spanning-tree vlan 4094
```

### 8.1.3 配置交换机对端连接

```
switch1#conf t
```

```
switch1(config)#mlag configuration 进入 MLAG 配置
```

```
switch1(config)#local-interface vlan 4094 配置 MLAG 的接口是 Vlan4094
```

```
switch1(config)#peer-address 172.17.0.2 指定对端的 PEER 地址
```

```
switch1(config)#peer-link port-channel 101 指定用于连接对端的本地 port-channel
```

```
switch1(config)#domain-id mlag01 创建并加入名为 mlag01 的 MLAG 域
```

```
switch1(config)#heartbeat-interval 2500 定义心跳线更新时间 2500
```

```
switch1(config)#reload-delay 150 定义负载
```

```
switch1(config)#exit
```

```
switch2#conf t
```

```
switch2(config)#mlag configuration 进入 MLAG 配置
```

```
switch2(config)#local-interface vlan 4094 配置 MLAG 的接口是 Vlan4094
```

```
switch2(config)#peer-address 172.17.0.1 指定对端的 PEER 地址
```

```
switch2(config)#peer-link port-channel 201 指定本地的 port-channel
```

```
switch2(config)#domain-id mlag01 创建并加入名为 mlag01 的 MLAG 域（必须同在一个域）
```

```
switch2(config)#heartbeat-interval 2500 定义心跳线更新时间 2500（必须相同）
```

```
switch2(config)#reload-delay 150 定义负载
```

```
switch2(config)#exit
```

以上步骤已完成交换机间 MLAG 的创建

#### 8.1.4 级联二层交换机

假如有一台二层交换机接入，则相应的 MLAG 配置

```
switch1#conf t
```

```
switch1(config)#interface ethernet 24 进入端口
```

```
switch1(config)#channel-group 1 mode active 创建 channel-group 1
```

```
switch1(config)#interface port-channel 1 创建 port-channel 1
```

```
switch1(config)#mlag 1 打上 mlag1 的标签
```

```
switch1(config)#exit
```

```
switch2#conf t
```

```
switch2(config)#interface ethernet 23 进入端口
```

```
switch2(config)#channel-group 1 mode active 创建 channel-group 1
```

```
switch2(config)#interface port-channel 1 创建 port-channel 1
```

```
switch2(config)#mlag 1 打上 mlag1 的标签
```

```
switch2(config)#exit
```

```
switch3#conf t
```

```
switch3(config)#interface f0/18 -19
```

```
switch3(config)#channel-group 2 mode active
```

```
switch3(config)#interface port-channel 2
```

```
switch1(config)#channel-group 2 mode active
```

#### 8.1.5 级联服务器

假如有一台服务器接入

```
switch1#conf t
```

```
switch1(config)#interface ethernet 22 进入端口
switch1(config)#channel-group 1 mode active 创建 channel-group 1
switch1(config)#interface port-channel 1 创建 port-channel 1
switch1(config)#mlag 1 打上 mlag1 的标签
switch1(config)#exit
```

```
switch2#conf t
switch2(config)interface ethernet 21
switch1(config)channel-group 1 mode active
switch1(config)interface port-channel 1
switch1(config)mlag 1
```

## 8.2 其他命令

### 8.2.1 将端口从 channel group 中移除

```
switch1(config)int g0/1
switch1(config)no channel-group
```

### 8.2.2 LACP 配置命令

#### 系统优先级

系统会默认选择 LACP 系统优先级高的作为通告链路聚合交换机。LACP 的系统值越小，优先级越高。

```
Switch(config)#lacp system-priority 8192
```

#### 端口优先级

```
Switch(config-if-Et1)#lacp port-priority 4096
```

#### LACP 速率

```
Switch(config-if-Et4)#lacp rate fast
```

#### LACP fallback 模式



Fallback 模式允许端口在未接受到 LACP PDU 前，自动生成 LAG

```
Switch(config-if-Po13)#port-channel lacp fallback
```

```
Switch(config-if-Po13)#port-channel lacp fallback timeout 60 /定义 timeout 时间
```

```
switch(config-if-Po5)#port-channel min-links 4 / 定义最小支持链路条数
```

交换机在做链路负载均衡时，是通过计算一个 port-channel 内的多条链路的中数据包头中的 hash 值来选路的。Hash 值将会决定数据包有由哪一个 active 状态下的链路转发。它不仅可以做 lag 间的负载均衡，同时也保障了在一个数据流内的所有数据包经过同一个网络路径。

在 Mlags 或多路径网络中，相同的 Coat 值会导致所有的交换机采用相同的 hash 计算导致 Hash 两极化，从而引起 MLAG 或 LAG 间链路无法进行均衡的负载分配。该问题可通过定义不同的 hash 算法来解决。

Hash 算法基于 ASIC 硬件，不同的平台有不同的算法。注意：通过不同的路径选择方式来解决该问题。

#### **FM4000 平台：**

```
Switch(config)#port-channel load-balance fm4000 fields ip mac-header
```

```
Switch(config)#port-channel load-balance fm4000 fields mac dst-mac eth-type
```

#### **Trident 平台：**

```
Switch(config)#port-channel load-balance trident fields ip mac-header
```

```
Switch(config)#port-channel load-balance trident fields mac dst-mac eth-type
```

## **9. ACL 访问控制列表**

### **9.1.1 支持功能**

#### **Supported Features**

- Ingress ACLs.
- Port ACL applied on layer-2 ethernet interfaces.
- Port ACL on port-channel interfaces. Ports in a port-channel apply the port-channel's ACL.
- Filters: IPv4 protocol, source and destination address, TCP and UDP ports, TCP flags, and TTL.
- List size: 512 active rules. Diminished capacity if rules contain L4 and port range filters.
- Broadcast and Multicast storm control.

### 9.1.2 不支持功能

#### Features Not Supported

- Egress ACLs.
- Filters based on IPv6/MAC.

## 10. VRRP 与 VARP

### 10.1 VRRP 介绍

VRRP 即 Virtual Router Redundancy Protocol,使得一组路由器可以形成一个虚拟的路由器,在 active-standby 的配置环境中来提供冗余保护。该虚拟路由器在共享的 LAN 中相当于一个虚拟的网关。

一个虚拟路由器包括 VRID (virtual router identifier) 与 virtual ip address。两个虚拟路由器不能分配同一个 VRID,即使他们在不同的 VLAN。同一个 LAN 可以有多个虚拟路由器来共同分担流量。每个虚拟路由器要有唯一的 VRID。

VRRP 使用 priority ratings 来指定 Master 或 Backup。Master 路由器会在 LAN 中发送周期性 VRRP 广播信息,转发数据包。Backup 路由器处在 inactive 状态,但假如 Master 路由器故障时,则切换到 active 状态。

默认情况下的 VRRP 配置:

- Router priority (default = 100)
- Preemption option (default is enabled)
- Advertisement timer (default = one second)
- Description (optional parameter)
- Authentication (optional parameter)
- Secondary IP addresses (optional parameter)

指定 Master 与 Backup 路由器

在同一虚拟路由器组中,VRRP 通过设定优先级来选择 Master。优先级值由 254 (高优先级)到 1 (低优先级),默认优先级为 100.

Preemption mode 即抢占模式。如果 Preemption 开启,VRRP 直接选取高优先级的路由成为 Master。如果未开启,则即便是高优先级的路由器也不会直接成为 Master,除非当前的 Master 不可用。

我们可以通过设定优先级来指定 Master 路由器。

常规命令

```

switch(config-if-vl20)#vrrp 15 priority 250 /配置虚拟路由 VRID15 的优先级为 250
switch(config-if-vl20)#no vrrp 15 preempt /关闭虚拟路由 VRID15 的 preempt 模式
switch(config-if-vl20)#vrrp 30 preempt /开启虚拟路由器 VRID30 的 preempt 模式
switch(config-if-vl100)#vrrp 35 timers advertise 10 /定义 VRR35 通告时间为 10 秒
switch(config-if-vl20)#vrrp 15 description Laboratory Router /添加 VRRP15 描述

switch(config-if-vl100)#vrrp 40 authentication text 12345 /定义认证密码（同一个组里的必须使用相同的认证密码）

switch(config-if-vl20)#vrrp 15 ip 10.2.4.5 secondary /指定 secondary 地址

switch(config-if-vl20)#vrrp 15 ip 10.1.1.5 /指定 primary 地址

switch(config-if-vl20)#vrrp 24 shutdown /开始 VRRP

switch(config-if-vl20)#no vrrp 24 shutdown /关闭 VRRP

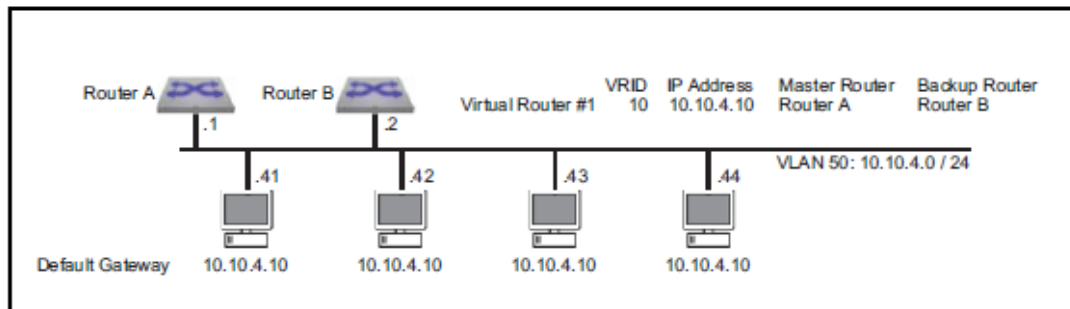
```

### 10.1.1 配置 VRRP

Example 1.

部署两个交换机在同一个虚拟路由组，以保护其中任意一台路由出现故障。

**Figure 13-2 VRRP Example 1 Network Diagram**



配置要求：配置两个交换机在同一个 VRRP 组。vlan50 中的 router-a 为 Master router,router-b 为 backup router。Router-a 优先级为 200， router-b 为默认优先级 100。通告时间为 3 秒一次。Preemption 开启（默认就是开启的）

```

Switch-A(config)#interface vlan 50
Switch-A(config-if-vl50)#ip address 10.10.4.1/24 /定义vlan ip 地址
Switch-A(config-if-vl50)#no vrrp 10 /开启preemption
Switch-A(config-if-vl50)#vrrp 10 priority 200 /定义优先级
Switch-A(config-if-vl50)#vrrp 10 timers advertise 3 /定义通告时间

```

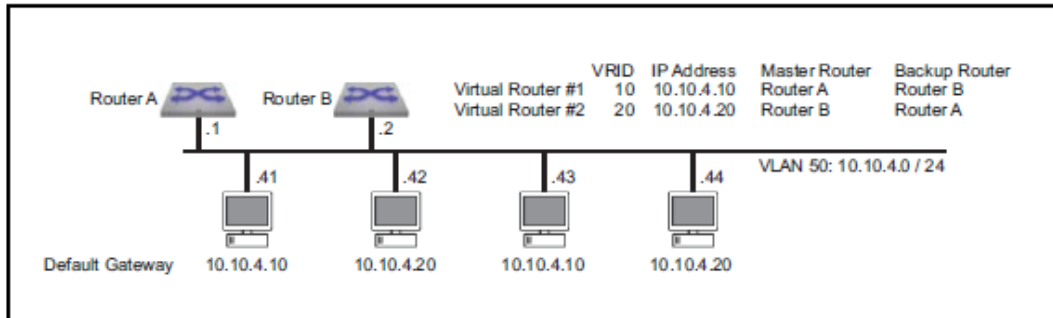
```
Switch-A(config-if-vl50)#vrrp 10 ip 10.10.4.10 /定义primary ip地址，即默认网关。  
Switch-A(config-if-vl50)#exit
```

```
Switch-B(config)#interface vlan 50  
Switch-B(config-if-vl50)#ip address 10.10.4.2/24  
Switch-B(config-if-vl50)#no vrrp 10  
Switch-B(config-if-vl50)#vrrp 10 timers advertise 3  
Switch-B(config-if-vl50)#vrrp 10 ip 10.10.4.10  
Switch-B(config-if-vl50)#exit
```

## Example 2

配置两个交换机同一个 VLAN 在两个 VRRP 组。该配置会保障 lan 中交换机故障，同时分担流量。

**Figure 13-3 VRRP Example 2 Network Diagram**



配置要求:

配置 vlan50 中虚拟路由组 10 的 router-a 为 Master，配置 router-b 为 backup。

配置 vlan50 中虚拟路由组 20 的 router-b 为 Master，配置 router-a 为 backup。

配置虚拟路由组 10 的通告间隔为 3 秒，路由组 20 的通告间隔为 5 秒。

Preemption 开启（默认就是开启的）

Switch code that implements Router A on the first switch

```
Switch-A(config)#interface vlan 50 /进入 vlan50
```

```
Switch-A(config-if-vl50)#ip address 10.10.4.1/24 /定义 vlan50 的 IP 地址
```

```
Switch-A(config-if-vl50)#no vrrp 10 /开启 preemption
```

```
Switch-A(config-if-vl50)#vrrp 10 priority 200 /定义虚拟组 10 的优先级为 200
```

```
Switch-A(config-if-vl50)#vrrp 10 timers advertise 3 /定义通告时间为 3 秒
Switch-A(config-if-vl50)#vrrp 10 ip 10.10.4.10 /定义 VRRP 虚拟 IP 地址，即默认网关
Switch-A(config-if-vl50)#no vrrp 20
Switch-A(config-if-vl50)#vrrp 20 timers advertise 5
Switch-A(config-if-vl50)#vrrp 20 ip 10.10.4.20
Switch-A(config-if-vl50)#exit
```

Switch code that implements Router B on the second switch

```
Switch-B(config)#interface vlan 50
Switch-B(config-if-vl50)#ip address 10.10.4.2/24
Switch-B(config-if-vl50)#no vrrp 10
Switch-B(config-if-vl50)#vrrp 10 timers advertise 3
Switch-B(config-if-vl50)#vrrp 10 ip 10.10.4.10
Switch-B(config-if-vl50)#no vrrp 20
Switch-B(config-if-vl50)#vrrp 20 priority 200
Switch-B(config-if-vl50)#vrrp 20 timers advertise 5
Switch-B(config-if-vl50)#vrrp 20 ip 10.10.4.20
Switch-B(config-if-vl50)#exit
```

## 10.2 VARP

VARP 即 Virtual-ARP 允许在 active-active 环境中，多个交换机同时路由相同地址的数据包。每个交换机会配置相同的 virtual IP address , virtual MAC address。在 MLAG 环境中，VARP 要优于 VRRP，因为 VARP 不需要像 VRRP 那样必须经过 Master 路由器。

VARP 最大支持 500 个虚拟地址。所有的 VLAN 端口的虚拟地址使用同样的虚拟 MAC 地址。

### 1) Virtual IP Addresses

使用 ip virtual-router address 来指定虚拟 IP 地址

```
Switch(config)#interface vlan 4094 /进入Vlan4094
Switch(config-if-Vl4094)#ip address 10.0.0.2/24 /指定primary IP地址
```

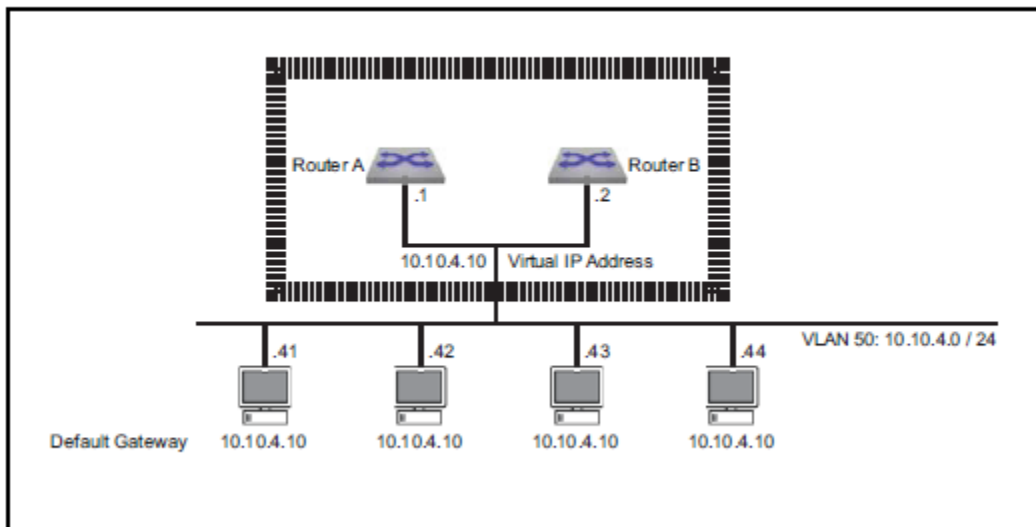
```
Switch(config-if-Vl4094)#ip virtual-router address 10.0.0.6 /指定虚拟接口的IP地址
Switch(config-if-Vl4094)#exit
```

## 2) Virtual MAC Address

使用 `IP virtual-router mac-address` 命令来指定交换机的虚拟 MAC 地址。交换机将映射所有的 virtual IP address 映射 到该 MAC 地址。该地址是只接收，不会以此地址为源发送数据包。

```
Switch(config)#ip virtual-router mac-address 001c.7300.0099 /指定虚拟 MAC 地址
switch>show ip virtual-router / 查看配置的 virtual ip address 与 virtual mac address
```

**Figure 13-1 VARP Configuration**



### Example 1

在同一个 MLAG 域里的两台交换机配置为 VARP 组。

配置要求：

配置 10.10.4.10 为 vlan50 的虚拟 IP 地址。

配置 10.24.4.1 为 vlan70 的虚拟 IP 地址。

两台交换机的虚拟 MAC 地址为 001c.7300.0999。

Switch code that implements VARP on the first switch

```
Switch-A(config)#ip virtual-router mac-address 001c.7300.0999 /配置虚拟MAC地址
```

```
Switch-A(config)#interface vlan 50 /进入vlan50
```

```
Switch-A(config-if-Vl50)#ip address 10.10.4.1/24 /配置vlan的IP地址
```

```
Switch-A(config-if-Vl50)#ip virtual-router address 10.10.4.10 /配置虚拟机IP地址
```

```
Switch-A(config-if-vl50)#interface vlan 70
Switch-A(config-if-vl70)#ip address 10.24.4.17/24
Switch-A(config-if-vl70)#ip virtual-router address 10.24.4.1
Switch-A(config-if-vl70)#exit
```

```
Switch code that implements VARP on the second switch
Switch-B(config)#ip virtual-router mac-address 001c.7300.0999
Switch-B(config)#interface vlan 50
Switch-B(config-if-vl50)#ip address 10.10.4.2/24
Switch-B(config-if-vl50)#ip virtual-router address 10.10.4.10
Switch-B(config-if-vl50)#interface vlan 70
Switch-B(config-if-vl70)#ip address 10.24.4.18/24
Switch-B(config-if-vl70)#ip virtual-router address 10.24.4.1
Switch-B(config-if-vl70)#exit
```

其他命令

```
Switch(config)#ip virtual-router mac-address advertisement-interval 60 /定义MAC地址通告时间
```

```
Switch(config)#no ip virtual-router mac address [mac_addr] /删除虚拟MAC地址
```

## 11. Spanning Tree Protocol

Spanning Tree Protocol 用来保护以太网二层不出现环路。Arista 支持 Rapid Spanning Tree, Multiple Spanning Tree, 以及 Rapid-Per VLAN Spanning Tree protocols。

一个运行正常的以太网，只允许有一条 Active 的路径存在于任意两个工作站间。Spanning tree 协议工作于二层，防止网络出现环路。STP 允许一个网络中存在冗余链路。即一个 active 链路与一个 backup 的路径，当 active 路径故障时，backup 链路自动转为 active 路径。STP 的标准是 IEEE802.1D。

Arista 支持的 STP 版本如下：

- **Rapid Spanning Tree (RSTP)**
- **Multiple Spanning Tree (MSTP)**
- **Rapid Per-VLAN Spanning Tree (Rapid-PVST)**

### 11.1 STP 协议

STP 的标准是 IEEE802.1D，作为最早版本的 STP 协议，它主要有两个限制：

1. 当网络改变时，其收敛速度较慢。
2. 整个网络在同一个 STP 里

## 11.2 Rapid Spanning Tree Protocol (RSTP)

RSTP 的标准是 802.1W。当网络出现变化时，RSTP 提供快速收敛。和 STP 一样，整个网络也在同一个 RSTP 里。

## 11.3 Rapid Per-VLAN Spanning Tree Protocol (Rapid-PVST)

PVST 支持每个 VLAN 单独形成一个 STP，即有几个 VLAN 就有几个 STP。PVST 支持在不产生环路的情况下二层 load balance。Load balance 是基于不同 VLAN 间做 Load balance。它的收敛速度较慢。

Arista 支持 Rapid-PVST。Rapid-PVST 是基于 PVST 与 RSTP 衍生的版本，其收敛速度快，且支持每个 VLAN 的 STP。可以在二层做不同 VLAN 间 Load Balance。

## 11.4 Multiple Spanning Tree Protocol (MSTP)

MST 发展于 RSTP，该协议在同一个 VLAN 里，即可以快速收敛，又可以做 load balance。MSTP 向下兼容 RSTP。默认情况下，Arista 使用的是 MSTP 协议。

MST 协议是将多个 VLAN 形成一个 Instance。该种构架支持通过多个转发路径，当一个 instance 故障时，不会影响其他 instance，因此网络冗余大大的提高。

### 11.4.1 MTS Regions

MST Regions 用于连接同样 MST 配置的桥组。每个 Region 支持最多 65 个 spanning-tree instances。MST regions 通过不同的版本号、名字、VLAN-to-instance map 来标识。所有的交换机必须配置相同。只有 MST Region members 参与到一个 Region 的 MST instance。VLAN 同一时间只允许分指定给一个 STP instance。MST 没有最大 region 数量的限制。

### 11.4.2 MTS Instances

每个 MST 都会标识为一个从与 VLAN 配套的 0-4094 的 instance 号。一个 MST region 包括保重 spanning tree instances: 一个 internal spanning tree instance(IST)，以及一个 multiple spanning tree instances(MSTI)。

IST: IST 是 MST region 中默认的 spanning tree instance，并且总是 instance 0。它为 region 标识根路由器，同时包含所有的未分配给 MST instance 的 VLAN 配置。

MSTI: MSTI 是通过 MST 配置，分配出去的连续的 VLAN。VLAN 分配给一个 MSTI 后将会从 IST instance 中移除。在 MSTI 中的 VLAN 会作为一个单独的 Spanning tree topology 运行。这是因为每个 VLAN 都属于一个单独的 Instance。MST instance 在逻辑上是独立的。

## 11.5 不同版本 SPT 相互兼容性

网络中的交换机可以运行不同的 SPT 版本。CST (common spanning tree) 对 STP, RSTP, MSTP, 以及 Rapid-PVST 来说，只是一个独立的的转发路径。

在 multi-instance 拓扑结构中，以下的 instances 可以响应 CST:



- Rapid-PVST: VLAN 1
- MST: IST (instance 0)

RSTP 与 MSTP 可以与其他的 STP 版本兼容：

- An RSTP bridge sends 802.1D (original STP) BPDUs on ports connected to an STP bridge.
- RSTP bridges operating in 802.1D mode remain in 802.1D mode even after all STP bridges are removed from their links.
- An MST bridge can detect that a port is at a region boundary when it receives an STP BPDU or an MST BPDU from a different region.
- MST ports assume they are boundary ports when the bridges to which they connect join the same region.

RSTP 向下兼容 802.1D bridges（使用 `clear spanning-tree detected-protocols` 命令可以强制 MST 端口与其邻居重新协商）：

- RSTP selectively sends 802.1D-configured BPDUs and Topology Change Notification (TCN) BPDUs on a per-port basis.
- When a port initializes, the migration delay timer starts and RSTP BPDUs are transmitted. While the migration delay timer is active, the bridge processes all BPDUs received on that port.
- If the bridge receives an 802.1D BPDU after a port's migration delay timer expires, the bridge assumes it is connected to an 802.1D bridge and starts using only 802.1D BPDUs.
- When RSTP uses 802.1D BPDUs on a port and receives an RSTP BPDU after the migration delay expires, RSTP restarts the migration delay timer and resumes using RSTP BPDUs on that port.

## 11.6 Switchport interface pairs

Switchport interface pairs 配置两个端口为 primary-backup 模式。当 primary interface 工作时，backup interface 为 standby 模式，保持在休眠状态。当 primary interface 停止工作，backup interface 接管它来控制流量。

在 primary 与 standby 接口间，流量二选一的进行平衡。如果任意一个接口关闭，另一个接管。以下是 Switchport interface pairs 使用的一些原则：

- Ethernet and Port Channels can be primary interfaces.
- Ethernet, Port Channel, Management, Loopback, and VLAN interfaces can be backup interfaces.
- The primary and backup interfaces can be different interface types.
- Interface pairs should be similarly configured to ensure consistent behavior.
- An interface can be associated with a maximum of one backup interface.
- An interface can back up a maximum of one interface.
- Any Ethernet interface configured in an interface pair cannot be a port channel member.

- STP is disabled on ports configured as primary or backup interfaces.
- Static MAC addresses should be configured after primary-backup pairs are established.

## 11.7 Spanning tree instance 的结构

二层网络包括桥与帧。当多条 active 的路径连接两个网络设备时，就会出现环路。Spanning tree 协议允许两个网络设备间只有一条 active 的路径。通过阻塞网桥的端口来实现无环。

### 11.7.1 Root bridge 与 designated birdge

根网桥是 STP 拓扑结构的中心。每个 Spanning tree instance 有一个根网桥。然后根据到根网桥的 distance 来选取路径。最短即最优。

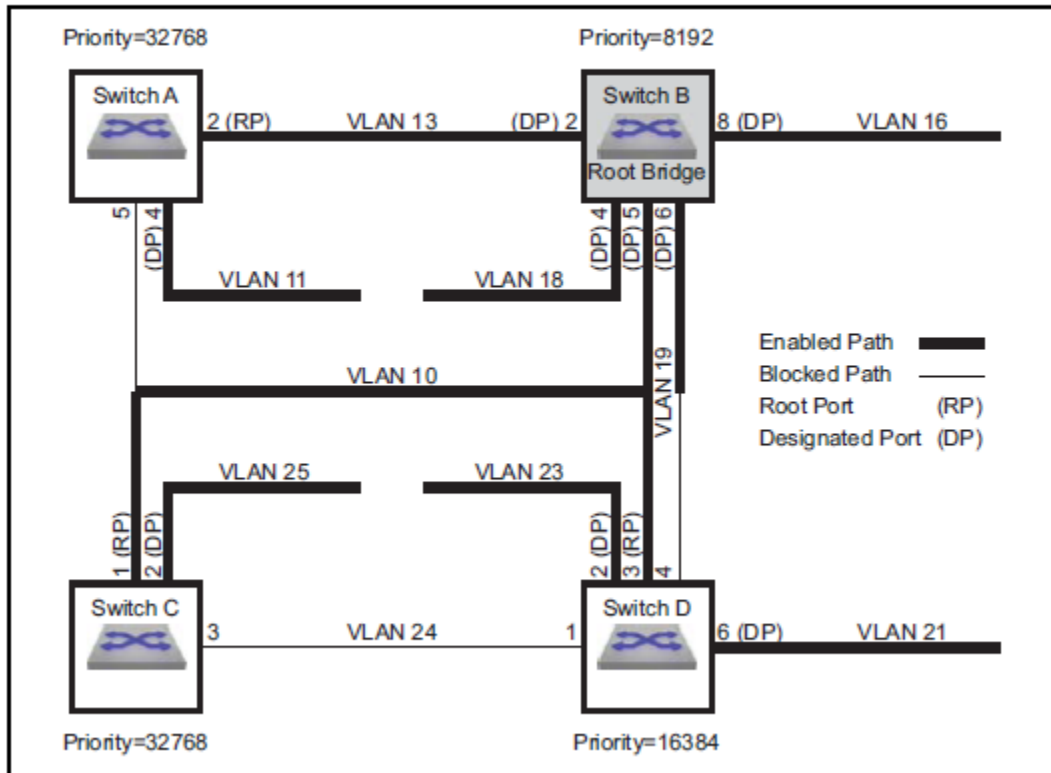
每个交换机都会分配唯一的 Bridge ID 号。网络中的所有交换机比较 Bridge ID 来选出根网桥。Bridge id 最小即成为根网桥。

- Port Priority (four bits)
- Instance number (12 bits): VLAN number (Rapid-PVST); Instance number (MST); 0 (RST)
- MAC address of switch (six bytes)

Designated bridge 是网络中到根网桥最短路径的交换机。当根网桥选举完成后，则会进行 designated bridge 的选举。

假设所有交换机的 COAT 相同，则选举如下：

**Figure 14-1 Spanning Tree Network Example**



- Switch B is the root bridge – its Bridge ID is lowest because it has the smallest port priority.
- Switch A is the designated bridge for VLAN 11.
- Switch B is the designated bridge for VLAN 10, VLAN 13, VLAN 16, VLAN 18, VLAN 19.
- Switch C is the designated bridge for VLAN 25.
- Switch D is the designated bridge for VLAN 21, VLAN 23.

### 11.7.2 Root port 与 designated port

Active 的端口有两种：root port 与 designated port，剩下的其他端口将会被阻塞。STP 同时定义了当 active 不可达时，由 backup 来接管流量。

Root port 是指到根网桥最短路径的端口。每个交换机都会比较每条可能路径的 cost 值，然后选择它的根端口。以下是上图选出的根端口

**Switch A:** port 2  
**Switch C:** port 1  
**Switch D:** port 3

Designated port: 每个 designated switch 可以作为提供多个 DP。

**Switch A:** port 4 (VLAN 11)  
**Switch B:** port 2 (VLAN 13), port 4 (VLAN 18), port 5 (VLAN 10), port 6 (VLAN 19), port 8 (VLAN 16)

**Switch C:** port 2 (VLAN 25)

**Switch D:** port 2 (VLAN 23), port 6 (VLAN 21)

**Alternate port:**提供 backup 路径。其端口处在阻塞状态，直到网络发生变化才会转化为 root port。

**Backup port:**提供到 designated bridge 的可选路径。其端口置于阻塞状态，直到网络发生变化才会转发为 designated port。

### 11.7.3 端口的状态

A port's activity state defines its current STP activity level. STP monitors BPDUs for network changes that require an activity state transition.

STP defines five port activity states:

- **Forwarding:** The port receives and sends data. Root ports and designated ports are either in, or transitioning to, this state.
- **Blocking:** The port does not receive or send data. Blocked ports receive BPDU packets. All ports except RPs and DPs are blocked, including alternate and backup ports.
- **Listening:** The first transitional post-blocking state, usually resulting from a network change that transforms a port into a root or designated port.
- **Learning:** The last transitional post-blocking state where the port prepares to forward frames by adding source addresses from inbound data packets to the switching database.
- **Disabled:** The interface does not forward frames or receive BPDU packets. Ports are manually disabled and not included in spanning tree calculations or operations.

### 11.7.4 端口类型

Port type is a configurable parameter that reflects the type of network segment that is connected to the port. Proper port type configuration results in rapid convergence after network topology changes. RSTP port types include normal, network, and edge ports. *Normal* is the default port type.

- **Normal** ports have an unspecified topology.
- **Network** ports connect only to switches or bridges.

RSTP immediately transitions network ports to the blocking state.

- **Edge** ports connect directly to end stations.

Edge ports transition directly to forwarding state, bypassing listening and learning states, because they do not create loops. An edge port becomes a normal port when it receives a BPDU.

### 11.7.5 链路类型

Link type is a configurable parameter that determines candidates for RSTP fast state transition.

- the default link type for full-duplex ports is *point-to-point*.
- the default link type for half-duplex ports is *shared*.

Fast state transitions are allowed on point-to-point links that connect bridges. Fast state transitions are not allowed on shared ports regardless of the duplex setting.

### 11.7.6 BPDU

BPDU 在交换机交换，主要信息包括 bridge ID 与 root path costs。

## BPDU 类型

- Configuration BPDU (CBPDU), used for computing Spanning Tree.
- Topology Change Notification (TCN) BPDU, announces network topology changes.
- Topology Change Notification Acknowledgment (TCA), acknowledges topology changes.

网桥会加入以下信息到出站的帧里:

默认的 hello timer 是两秒。

- source address: outbound port's MAC address.
- destination address: STP multicast address 01:80:C2:00:00:00.

### 11.7.7 Bridge timer

- **hello-time:** transmission interval between consecutive BPDU packets.
- **forward-time:** the period that ports remain in listening and learning states.
- **max-age:** the period that BPDU data remains valid after it is received.
- **max-hop:** the number of bridges in an MST region that a BPDU can traverse before it is discarded.

交换机会重新计算 spanning tree topology,当 max-age 过期时仍未收到 BPDU。当 edge ports 与 point-to-point 端口配置正确, RSTP 网络收敛并不会需要 forward-delay 与 maxAge timer。

## MSTP BPDU

MSTP BPDU 是看作单一的 instance, 同时为整个 region 提供 STP 信息。MSTP 会对标准的 IST BPDU 再次编译, 然后加入 region 信息与 MST instance 信息。

MSTP BPDU 可以与 RSTP 兼容。RSTP 网桥会被看作一个 MST region。这时将会忽略 bridge number。这是因为如下原因:

- RSTP bridges interpret MSTP BPDUs as RSTP BPDUs.
- RSTP bridges increment the *message age timer* only once while data flows through an MST region; MSTP measures *time to live* with a **remaining hops** variable, instead of the **message age timer**.

## 11.8 配置 Spanning tree

关闭 spanning tree 命令

```
switch(config)#spanning-tree mode none
```

### 11.8.1 MST 配置

默认的是 MST 版本。

#### 1) 配置 MST

```
switch(config)#spanning-tree mode mstp
```

#### 2) 配置 MST region

同一个 MST region 中的所有的交换机都要配置相同的名称，版本，以及 vlan-to-instance map。

```
switch(config-mst)#instance 8 vlans 4-7,9 / 分配 4-7, 9 到 instance 8
```

```
switch(config-mst)#no instance 10 vlans 6 /从 instance10 中移除 vlan 6
```

```
switch(config-mst)#name corporate_1 /命名 MST 为 corporate_1
```

```
switch(config-mst)#revision 3 /选择版本 3
```

```
switch(config-mst)#exit /退出 MST 并保存配置
```

```
switch(config-mst)#abort /推出 MST 但不保存配置
```

```
switch(config)#spanning-tree mode mst 4 priority 4096 /配置 MST4 的优先级为 4096
```

- spanning-tree priority
- spanning-tree root
- spanning-tree port-priority

### 11.8.2 RST

```
switch(config)#spanning-tree mode rstp /开启 RSTP 协议
```

```
switch(config)#spanning-tree priority 4096
```

```
and
```

```
switch(config)#spanning-tree mst 0 priority 4096 /mst0 的优先级为 4096
```

RST 不能使用如下的命令

```
switch(config)#spanning-tree mst 4 priority 4096
```

```
and
```

```
switch(config)#spanning-tree VLAN 3 priority 4096
```

```
switch(config)#show spanning-tree /查看 Spanning tree
```

### 11.8.3 Rapid-PVST

```
switch(config)#spanning-tree mode rapid-pvst
```

```
switch(config)#spanning-tree VLAN 4 priority 4096
```

### 11.8.4 Switchport backup 模式

当开启 switchport backup 模式时将会关闭所有的 spanning tree。

```
switch(config)#spanning-tree mode backup
```

配置端口1的backup端口为7，即1为primary port,7为backup端口。

```
switch(config)#interface ethernet 1
switch(config-if-Et1)#switchport backup interface ethernet 7
```

The *prefer* option of the **switchport backup interface** command establishes a peer relationship between the primary and backup interfaces and specifies VLAN traffic that the backup interface normally carries. If either interface goes down, the other interface carries traffic normally handled by both interfaces.

#### Example

These steps perform the following:

- configures Ethernet interface 1 as a trunk port that handles VLANs 4 through 9 traffic.
- configures Ethernet interface 2 as the backup interface.
- assigns Ethernet 2 as the preferred interface for VLANs 7 through 9.

**Step 1** Enter configuration mode for the primary interface

```
switch(config)#interface ethernet 1
```

**Step 2** Configure the primary interface as a trunk port that services VLANs 4-9

```
switch(config-if-Et1)#switchport mode trunk
switch(config-if-Et1)#switchport trunk allowed vlan 4-9
```

**Step 3** Configure the backup interface and specify the VLANs that it normally services.

```
switch(config-if-Et1)#switchport backup Ethernet 2 prefer vlan 7-9
```

### 11.8.5 配置 Spanning tree instance

A network performs these steps to set up an STP instance:

1. The bridge with the lowest ID is elected root bridge.
2. Root ports (RP) are selected on all other bridges.
3. Designated bridges are selected for each network segment.
4. Designated ports (DP) are selected on each designated bridge.
5. Networks begin forwarding data through RPs and DPs. All other ports are blocked.

#### 1) 选举根

STP 使用 bridge ID 来选举根网桥，而 bridge ID 则是由优先级，MAC 决定。

Show spanning tree root /查看根网桥

交换机提供两种命令来配置交换机的优先级。即 spanning-tree priority 与 spanning-tree root。

- spanning-tree priority options are integer multiples of 4096 between 0 and 61440.
- spanning-tree root options are primary and secondary.
  - primary assigns a priority of 8192.
  - secondary assigns a priority of 16384.

默认优先级是 32768，默认 primary 的 instance 是 0，默认是 secondary 的 instance 是 1。

以下是配置 MST instance 的命令

```
switch(config)#spanning-tree mst 0 root primary
switch(config)#spanning-tree mst 1 root secondary
switch>show spanning-tree root
Root ID Root Hello Max Fwd
Instance Priority MAC addr Cost Time Age Dly Root Port
-----
MST0 8192 001c.7301.6017 0 2 20 15 None
MST1 16385 001c.7301.6017 0 0 0 0 None
MST2 32770 001c.7301.6017 0 0 0 0 None
```

- Instance 0 root priority is 8192: primary priority plus the instance number of 0.
- Instance 1 root priority is 16385: secondary priority plus the instance number of 0.

以下是配置 Rapid-PVST Vlan 优先级

```
switch(config)#spanning-tree vlan 1 priority 8192
switch(config)#spanning-tree vlan 2 priority 16384
switch(config)#spanning-tree vlan 3 priority 8192
switch(config)#no spanning-tree vlan 4 priority
switch(config)#show spanning-tree root
Root ID Root Hello Max Fwd
Instance Priority MAC addr Cost Time Age Dly Root Port
-----
VL1 8193 001c.7301.6017 0 2 20 15 None
VL2 16386 001c.7301.6017 0 2 20 15 None
VL3 8195 001c.7301.6017 0 2 20 15 None
VL4 32788 001c.7301.6017 0 2 20 15 None
```

- VLAN 1 root priority is 8193: configured priority plus the VLAN number of 1.
- VLAN 2 root priority is 16386: configured priority plus the VLAN number of 2.
- VLAN 3 root priority is 8195: configured priority plus the VLAN number of 3.
- VLAN 4 root priority is 32788: default priority plus the VLAN number of 4.

## 2) Path Cost

Cost 值由 1-200000000。默认情况下，千兆口的 cost 是 20000。万兆口的 cost 是 2000。

配置 RSTP 与 MSTP instance 0

This cost is valid for RSTP or MSTP instance 0.

```
switch(config)#interface ethernet 5
switch(config-if-Et5)#spanning-tree cost 25000
```

配置 MSTP

```
switch(config)#interface ethernet 5
switch(config-if-Et5)#spanning-tree mst 200 cost 300000
```

配置 Rapid-PVST

```
switch(config)#interface ethernet 5
switch(config-if-Et5)#spanning-tree vlan 200-220 cost 10000
```

配置 RST 端口优先级



```
switch(config)#interface ethernet 5
switch(config-if-Et5)#spanning-tree port-priority 144
```

配置 MST 端口优先级

```
switch(config)#interface ethernet 5
switch(config-if-Et5)#spanning-tree mst 10 port-priority 144
```

端口角色配置:

- PortFast: Allows ports to skip the listening and learning states before entering forwarding state.用于连接PC终端或服务器, 打印机等。
  - Port Type and Link Type: Designates ports for rapid transitions to the forwarding state.
  - Root Guard: Prevents a port from becoming root port or blocked port.
  - Loop Guard: Prevents loops resulting from a unidirectional link failure on a point-to-point link.
  - Bridge Assurance: Prevents loops caused by unidirectional links or a malfunctioning switch.
- ```
switch(config-if-Et5)#spanning-tree portfast /开启端口5portfast
```

端口类型

RSTP只接受edge port与point-to-point links快速转发。默认情况下是normal port

- Edge ports connect to a host (end station). Configuring a port that connects to a bridge as an edge port may create a loop. Edge ports that receive a BPDU become a normal spanning tree port.
- Network ports connect only to a Layer 2 switch or bridge. Configuring a port connected to a host as a network port transitions the port to the blocking state.
- Normal ports have an unspecified topology.

```
switch(config-if-Et5)#spanning-tree portfast network
```

链路模式

- full-duplex ports are point-to-point.
  - half-duplex ports are shared.
- ```
switch(config-if-Et5)#spanning-tree link-type shared
```

root Guard 用于组织端口变为root port.

```
switch(config-if-Et5)#spanning-tree guard root
```

loop guard配置在point-to-point链路的端口, 防止环路  
当使用loop guard时, 请注意下面几点

- Do not enable loop guard on portfast-enabled ports.
- Loop guard is not functional on ports not connected to point-to-point links.
- Loop guard has no effect on disabled spanning tree instances.

```
switch(config)#spanning-tree hello-time 1000
switch(config)#spanning-tree max-hops 40
switch(config)#spanning-tree forward-time 25
switch(config)#spanning-tree max-age 25
```

```
switch(config)#spanning-tree portfast bpduguard default
switch(config)#interface ethernet 5
switch(config-if-Et5)#spanning-tree bpduguard disable
```