



以太网链路聚合与交换机堆叠、集群

主讲人：鲍婷婷

目录

1 网络可靠性需求

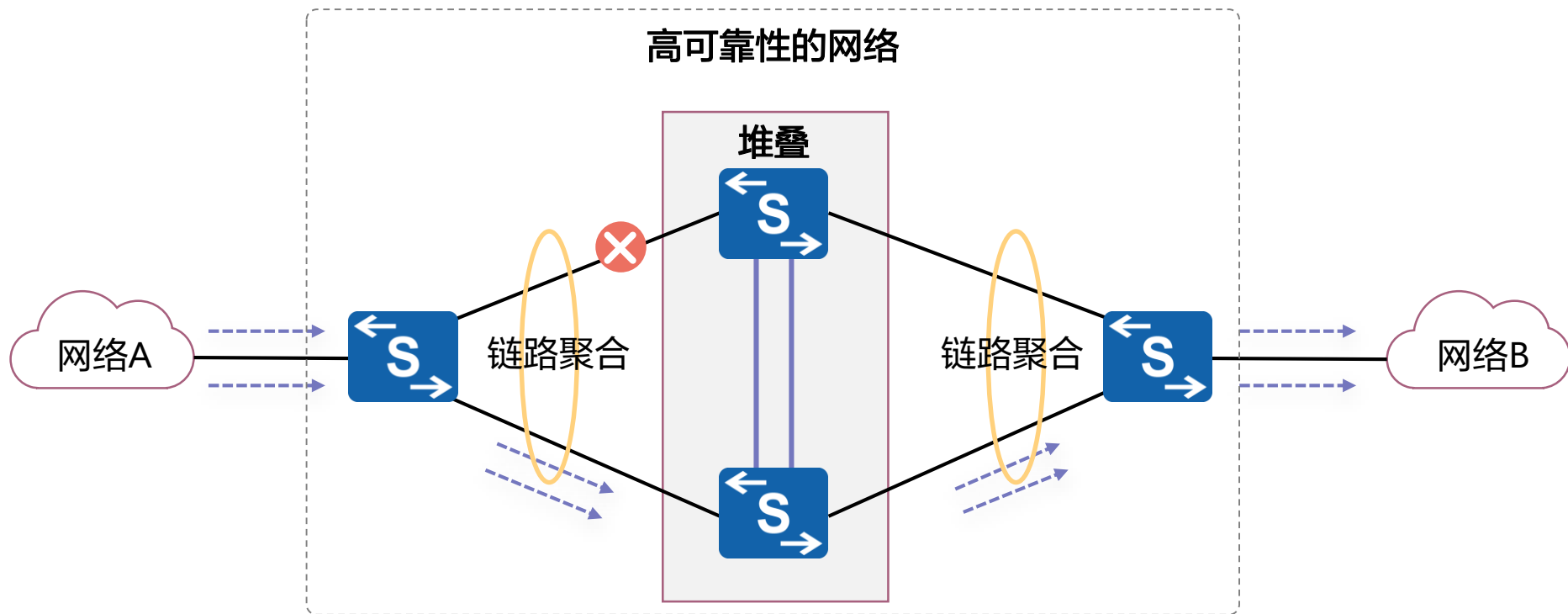
- 网络可靠性需求

2 链路聚合技术原理与配置

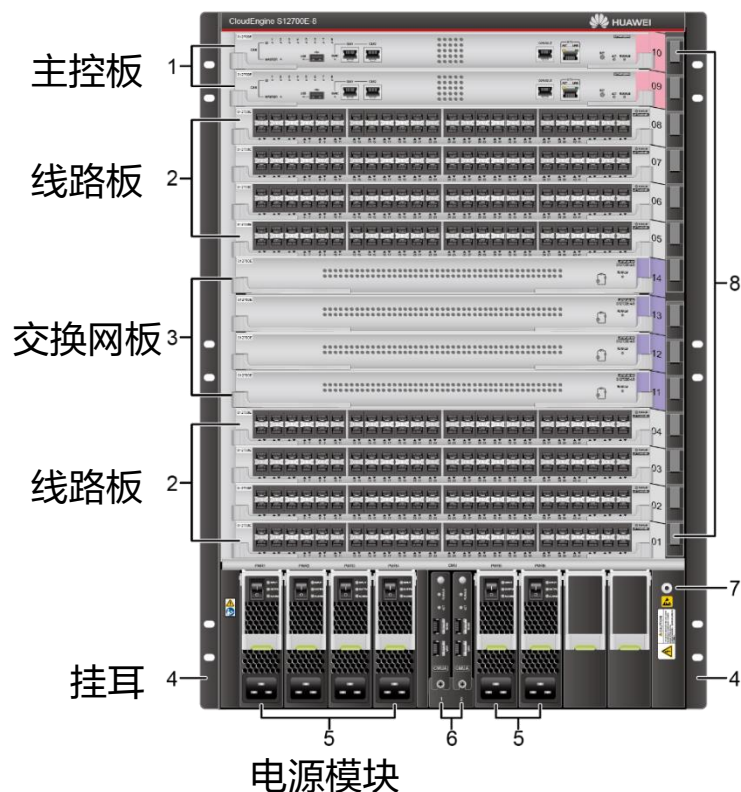
3 堆叠/集群概述

网络的可靠性

- 网络的可靠性指当设备或者链路出现单点或者多点故障时保证网络服务不间断的能力。
- 网络的可靠性可以从单板、设备、链路多个层面实现。



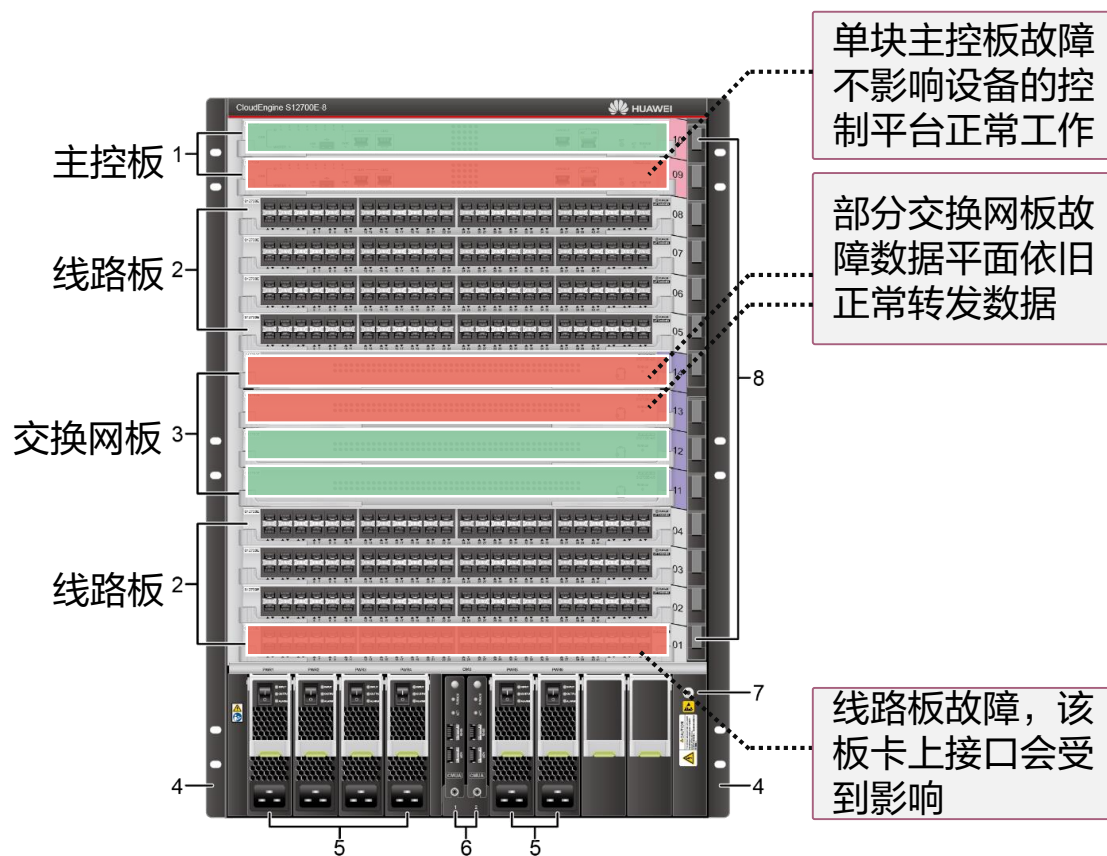
单板可靠性 (1)



S12700E-8机框正面结构

- 框式交换机由机框、电源模块、风扇模块、主控板、交换网板（SFU）、线路板（LPU）构成。
- 机框
- 电源模块
- 风扇模块
- 主控板（MPU，Main Processing Unit）
- 交换网板（SFU，Switch Fabric Unit）
- 线路板（LPU，Line Processing Unit）

单板可靠性 (2)

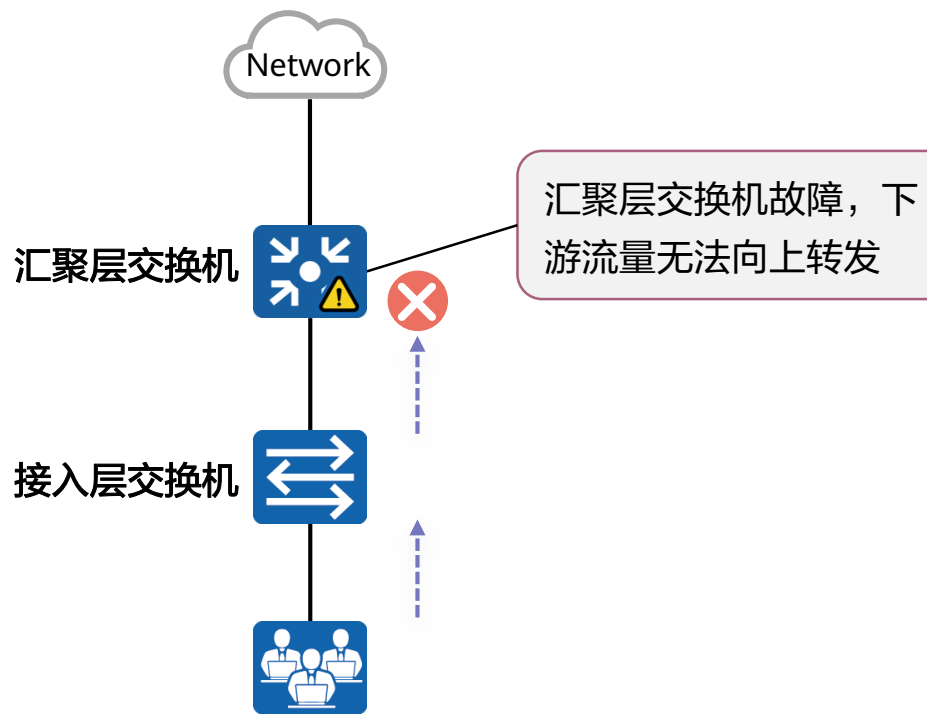


S12700E-8机框正面结构

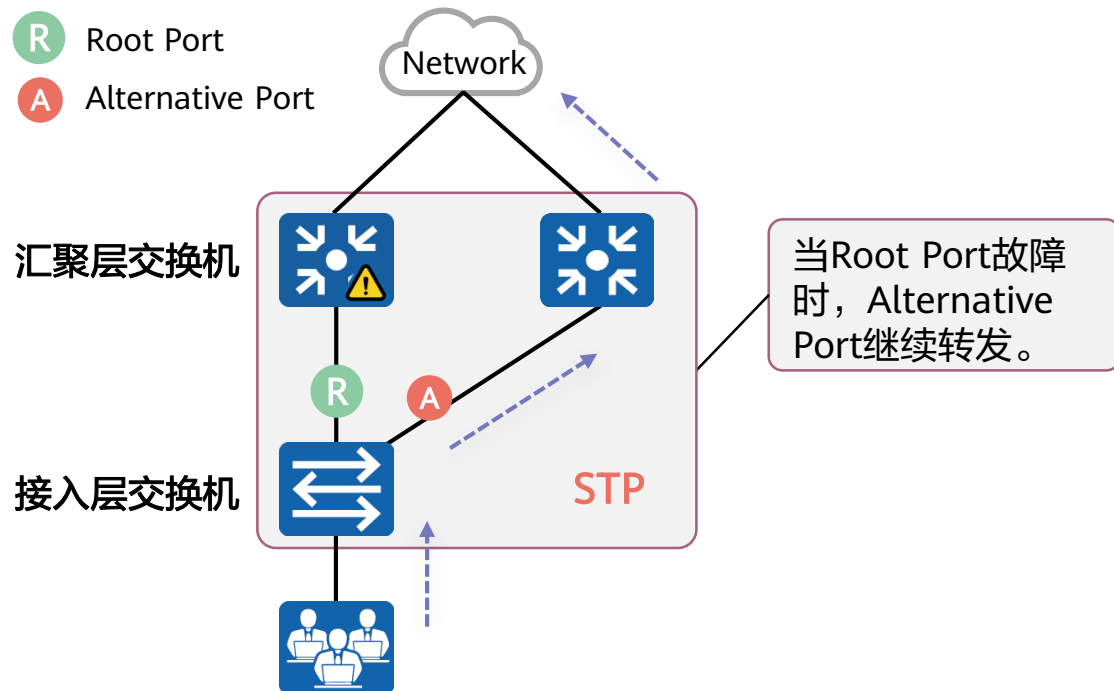
- 以S12700E-8为例, 设备提供8个线路板槽位、4个交换网板槽位、2个主控板槽位、6个电源模块槽位、4个风扇模块槽位。
- 框式交换机配置多个主控板、交换网板可保证设备自身的可靠性, 单个槽位的交换网板、主控板损坏不影响设备的正常运行。
- 框式交换机的线路板损坏后, 该板卡上的接口无法正常转发数据。

设备可靠性

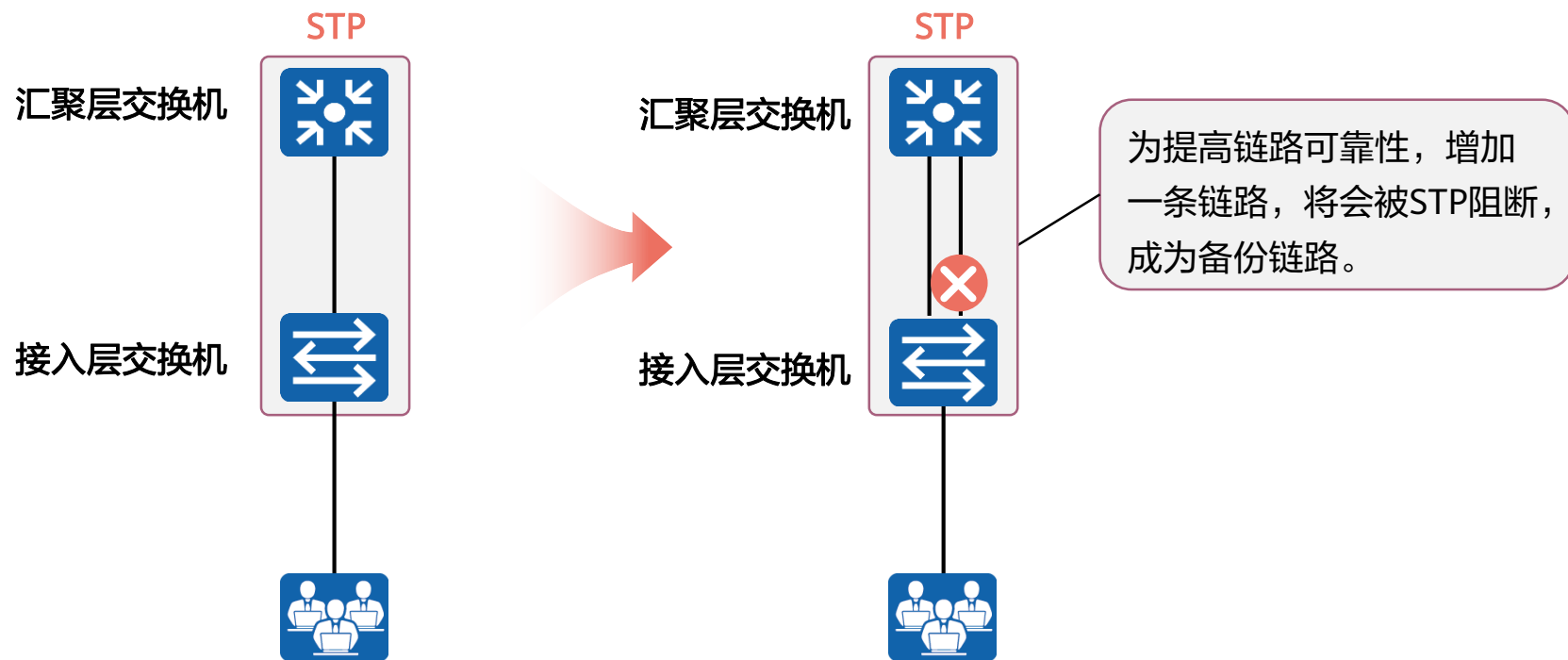
无备份



主备备份



链路可靠性



目录

1

网络可靠性需求

2

链路聚合技术原理与配置

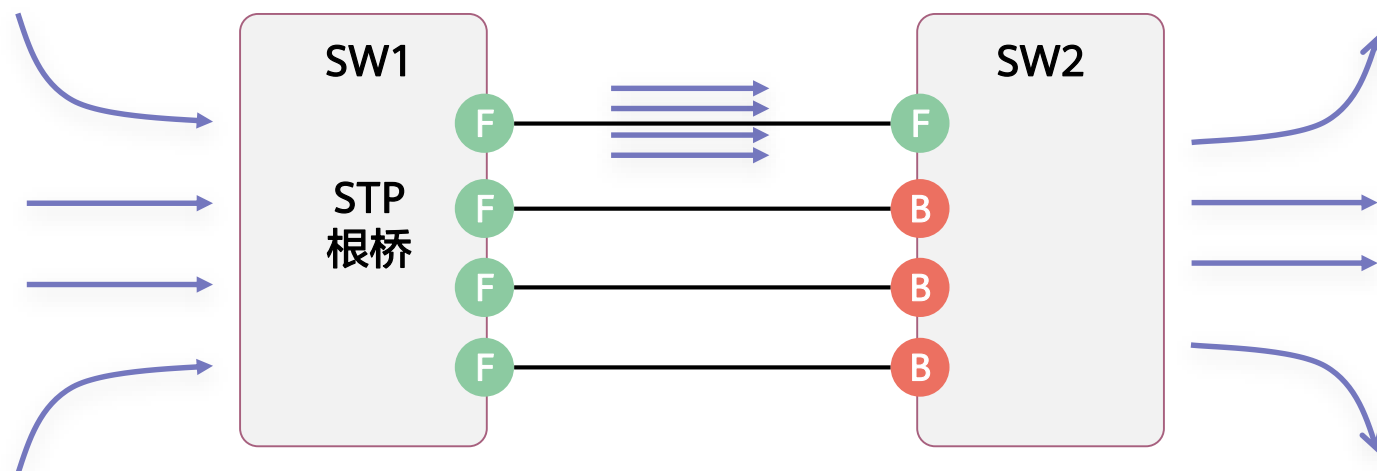
- 链路聚合基本原理
- 链路聚合手工模式
- 链路聚合LACP模式
- 链路聚合典型使用场景
- 链路聚合配置举例

3

堆叠/集群概述

提升链路带宽

- 设备之间存在多条链路时，由于STP的存在，实际只会有一条链路转发流量，设备间链路带宽无法得到提升。

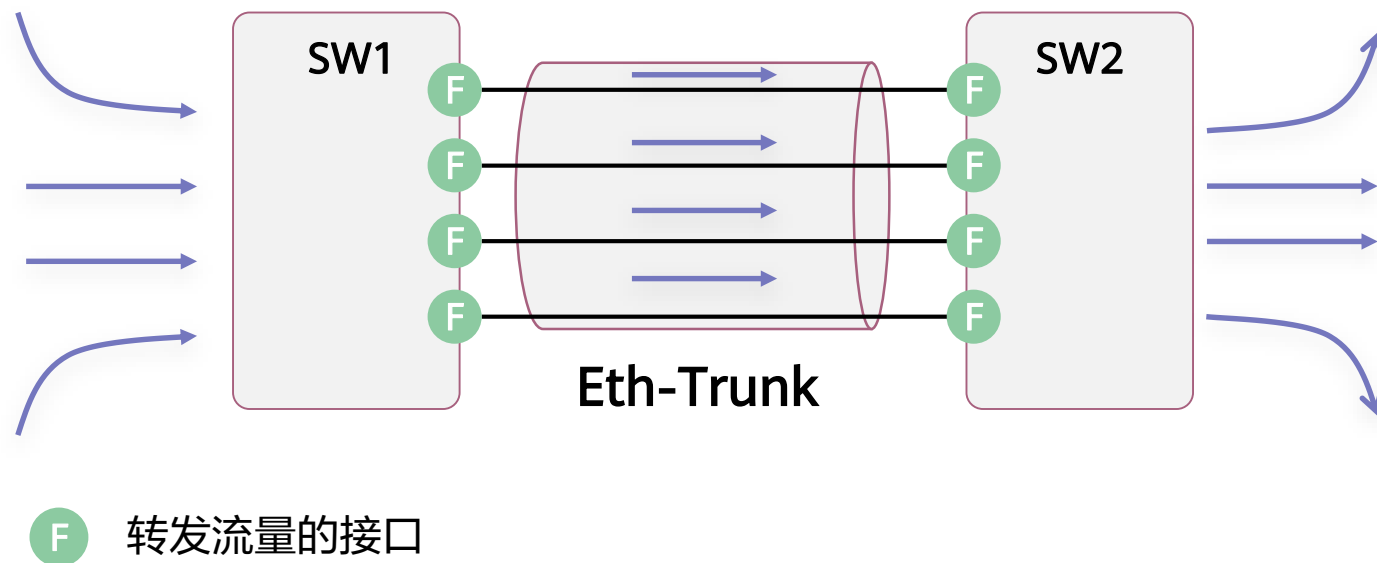


F 转发流量的接口

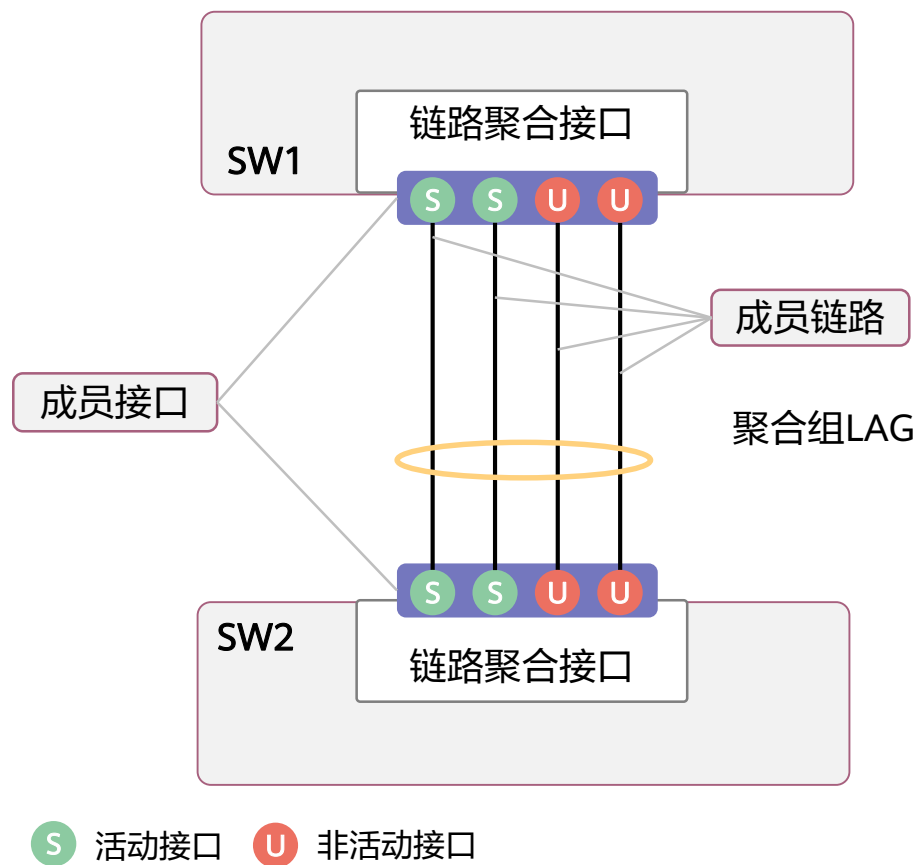
B STP阻塞端口，不转发流量

以太网链路聚合

- 以太网链路聚合Eth-Trunk：简称链路聚合，通过将多个物理接口捆绑成为一个逻辑接口，可以在不进行硬件升级的条件下，达到增加链路带宽的目的。



链路聚合基本术语/概念



- 聚合组（Link Aggregation Group，LAG）
- 成员接口和成员链路
- 活动接口和活动链路
- 非活动接口和非活动链路
- 聚合模式：手工模式和LACP模式。
- 其他概念：活动接口上限阈值和活动接口下限阈值。

目录

1

网络可靠性需求

2

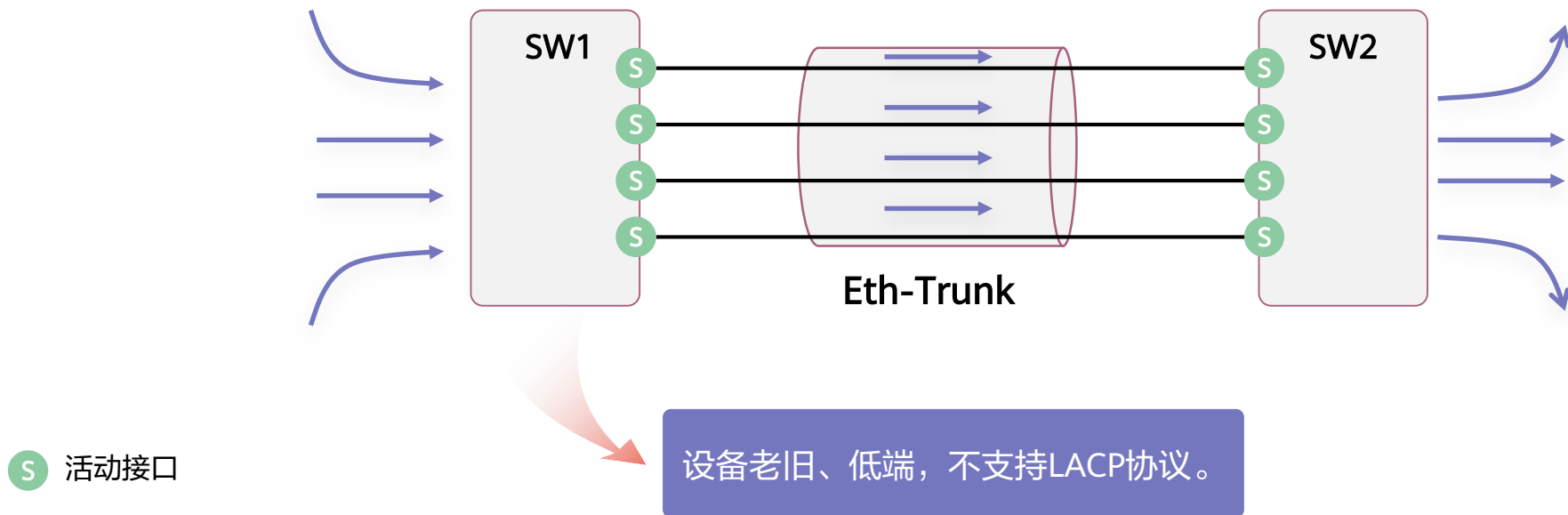
链路聚合技术原理与配置

- 链路聚合基本原理
- 链路聚合手工模式
- 链路聚合LACP模式
- 链路聚合典型使用场景
- 链路聚合配置举例

3

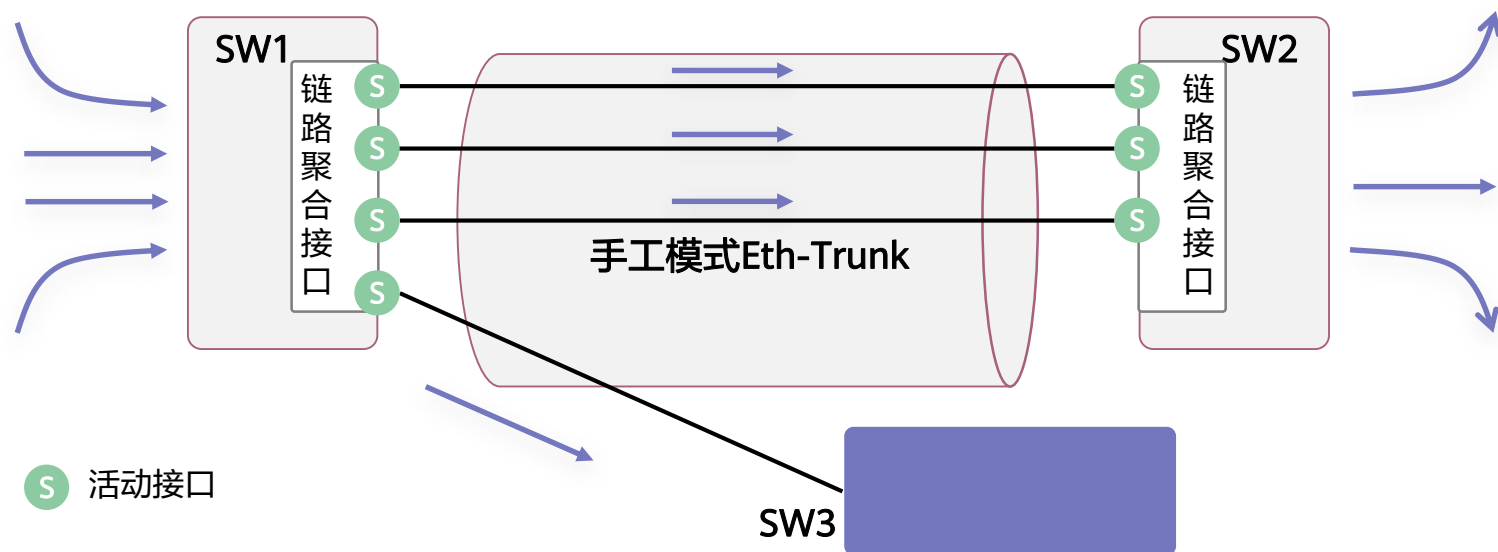
堆叠/集群概述

手工模式



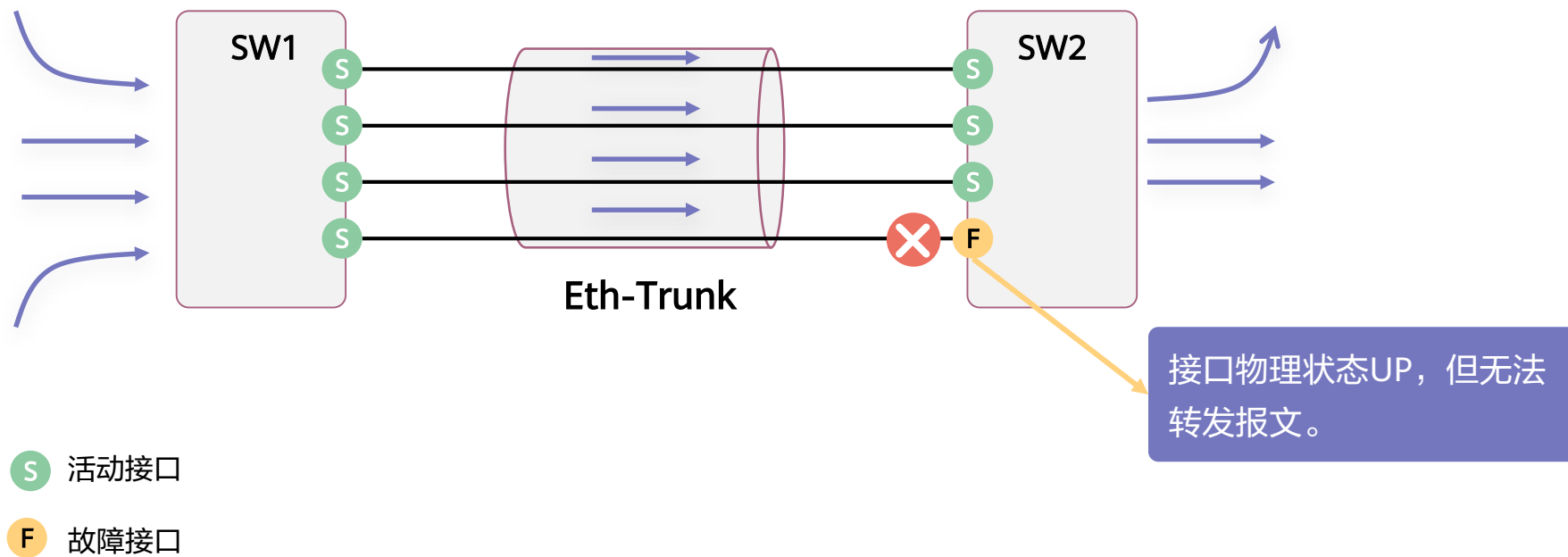
- 手工模式： Eth-Trunk的建立、成员接口的加入均由手动配置，双方系统之间不使用LACP进行协商。
- 正常情况下所有链路都是活动链路，该模式下所有活动链路都参与数据的转发，平均分担流量，如果某条活动链路故障，链路聚合组自动在剩余的活动链路中平均分担流量。
- 当聚合的两端设备中存在一个不支持LACP协议时，可以使用手工模式。

手工模式缺陷 (1)



- 为了使链路聚合接口正常工作，必须保证本端链路聚合接口中所有成员接口的对端接口：
 - 属于同一设备
 - 加入同一链路聚合接口
- 手工模式下，设备间没有报文交互，因此只能通过管理员人工确认。

手工模式缺陷 (2)



手动模式下，设备只能通过物理层状态判断对端接口是否正常工作。

目录

1

网络可靠性需求

2

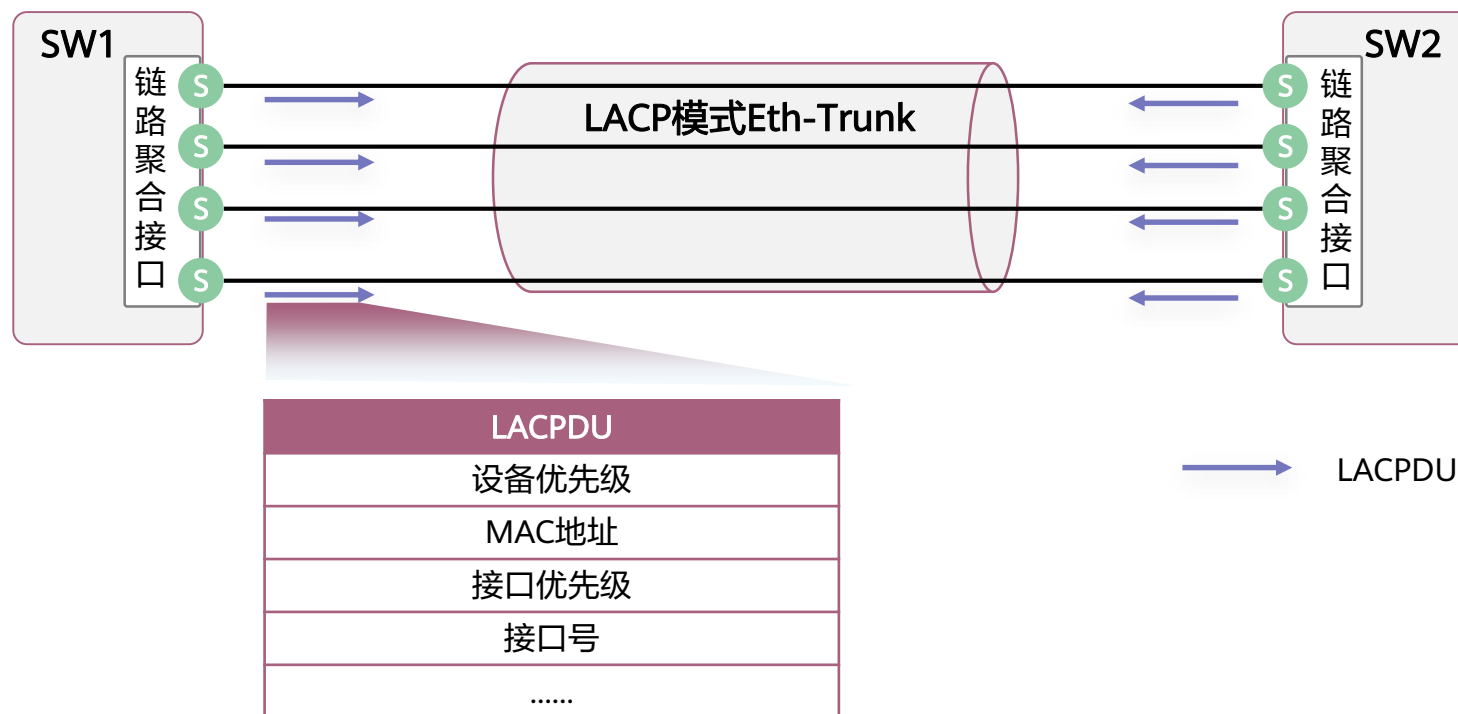
链路聚合技术原理与配置

- 链路聚合基本原理
- 链路聚合手工模式
- 链路聚合LACP模式
- 链路聚合典型使用场景
- 链路聚合配置举例

3

堆叠/集群概述

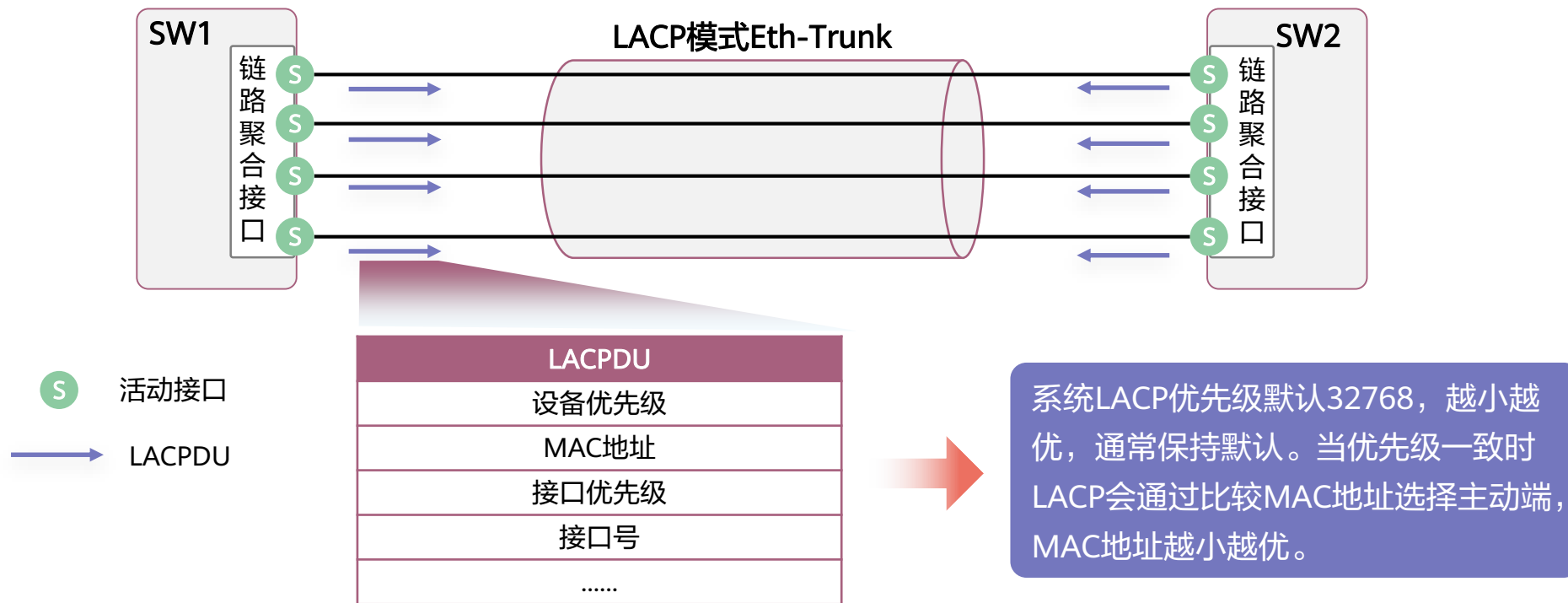
LACPDU



- LACP模式：采用LACP协议的一种链路聚合模式。设备间通过链路聚合控制协议数据单元（Link Aggregation Control Protocol Data Unit, LACPDU）进行交互，通过协议协商确保对端是同一台设备、同一个聚合接口的成员接口。
- LACPDU报文中包含设备优先级、MAC地址、接口优先级、接口号等。

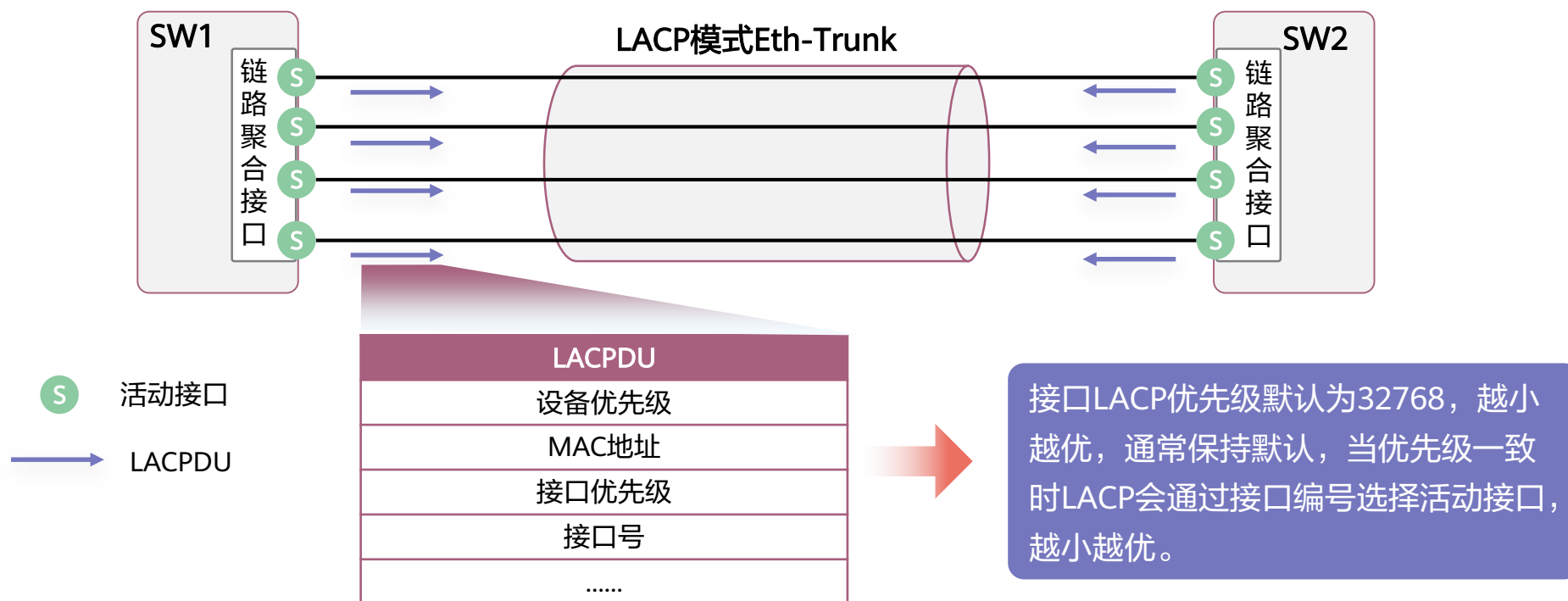
系统优先级

- LACP模式下，两端设备所选择的活跃接口数目必须保持一致，否则链路聚合组就无法建立。此时可以使其中一端成为主动端，另一端（被动端）根据主动端选择活跃接口。
- 通过系统LACP优先级确定主动端，值越小优先级越高。



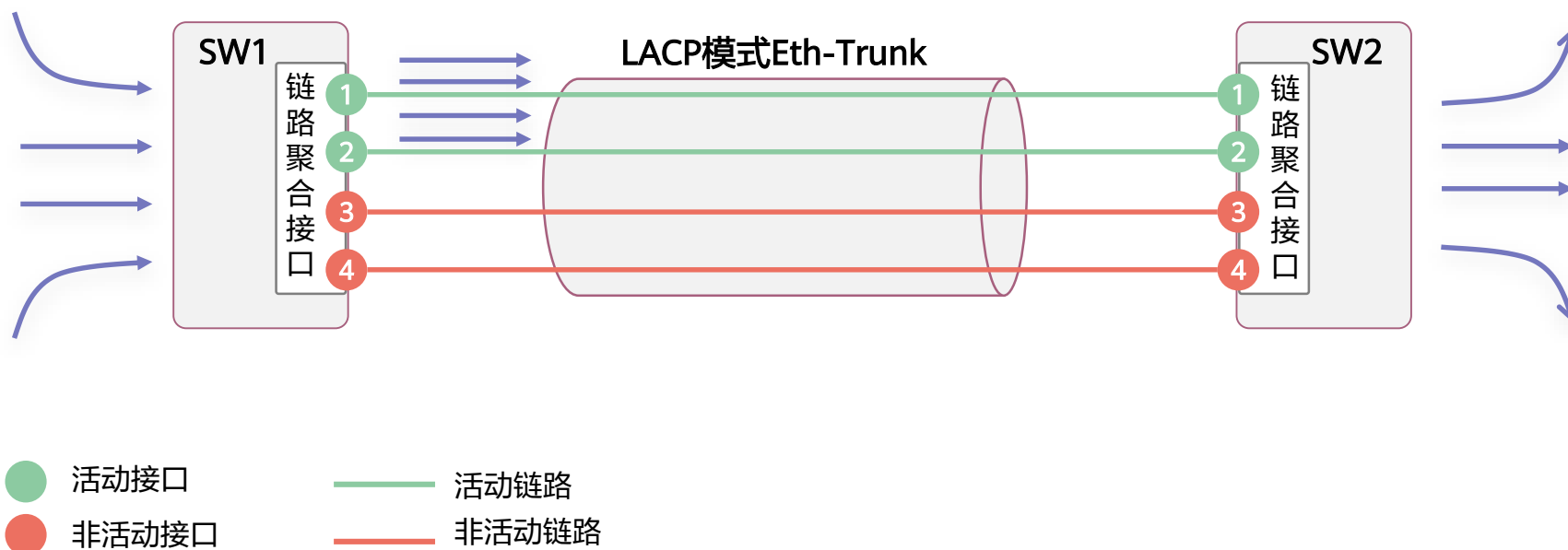
接口优先级

- 选出主动端后，两端都会以主动端的接口优先级来选择活动接口，优先级高的接口将优先被选为活动接口。接口LACP优先级值越小，优先级越高。



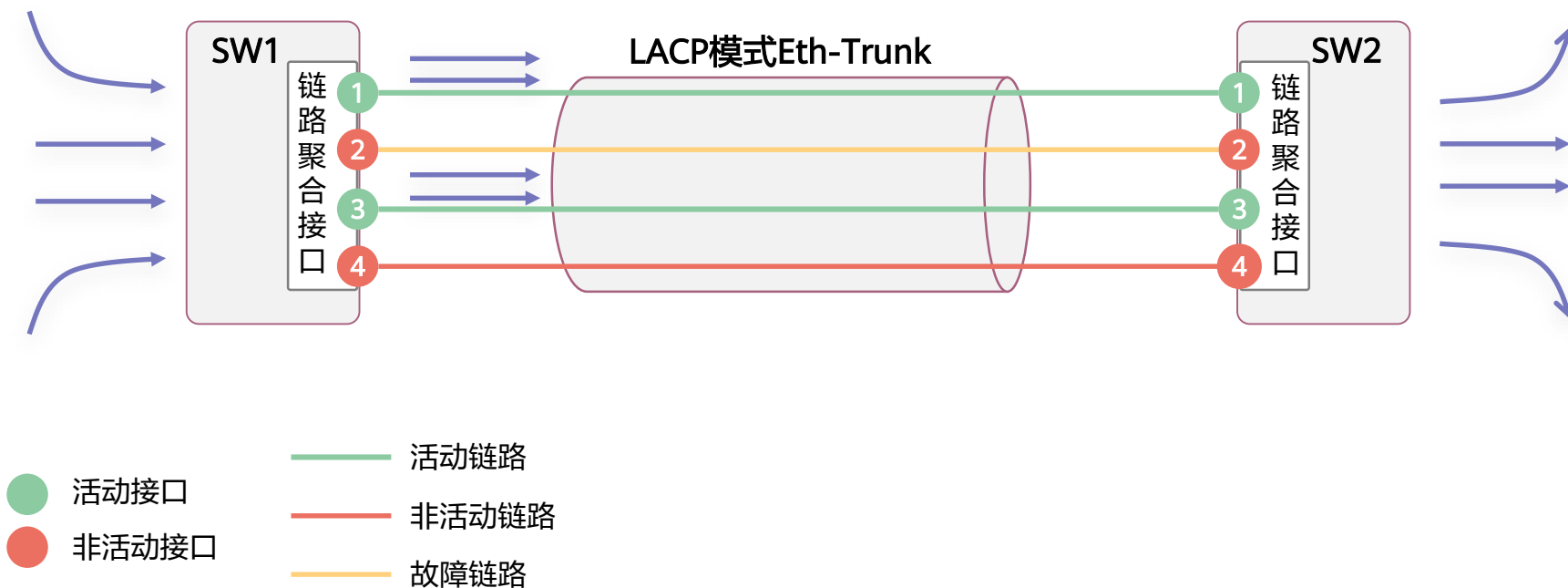
最大活动接口数 (1)

- LACP模式支持配置最大活动接口数目，当成员接口数目超过最大活动接口数目时会通过比较接口优先级、接口号选举出较优的接口成为活动接口，其余的则成为备份端口（非活动接口），同时对应的链路分别成为活动链路、非活动链路。交换机只会从活动接口中发送、接收报文。

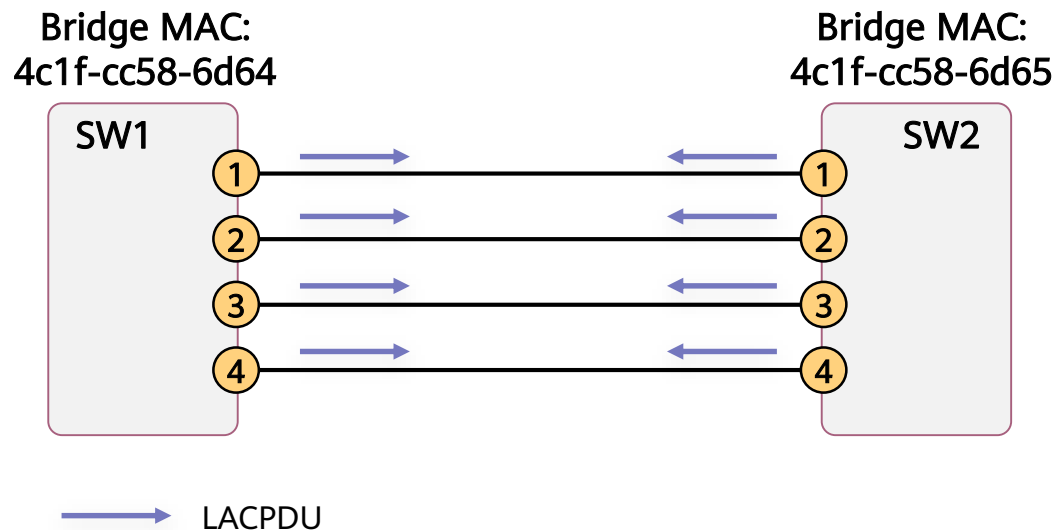


最大活动接口数 (2)

- 当活动链路中出现链路故障时，可以从非活动链路中找出一条优先级最高（接口优先级、接口编号比较）的链路替换故障链路，实现总体带宽不发生变化、业务的不间断转发。



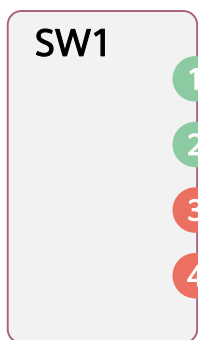
活动链路选举 (1)



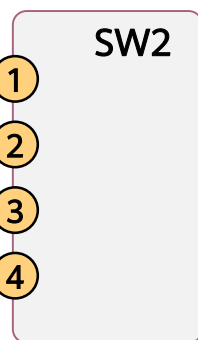
- SW1、SW2配置LACP模式的链路聚合。两端都设置最大活跃接口数为2。
- 通过LACPDU选举出优先级较高的交换机SW1，作为LACP协商过程的主动端。

活动链路选举 (2)

Bridge MAC:
4c1f-cc58-6d64



Bridge MAC:
4c1f-cc58-6d65



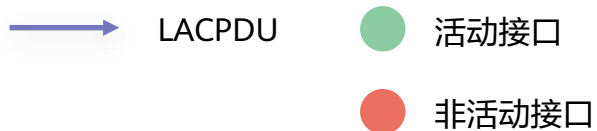
- 活动接口
- 非活动接口

- SW1在本端通过比较接口优先级、接口编号选举出活动接口，其中1、2号接口在相同的接口优先级下拥有更小的接口编号，成为活动接口。

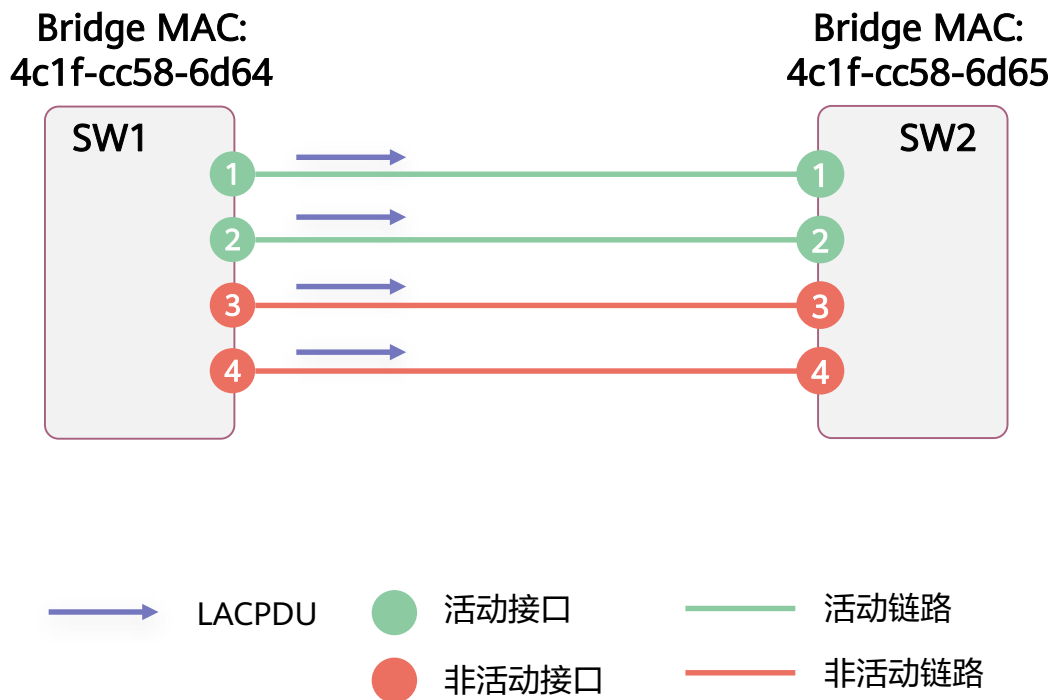
活动链路选举 (3)



- SW1通过LACPDU将本端活动端口选举结果告知对端。



活动链路选举 (4)

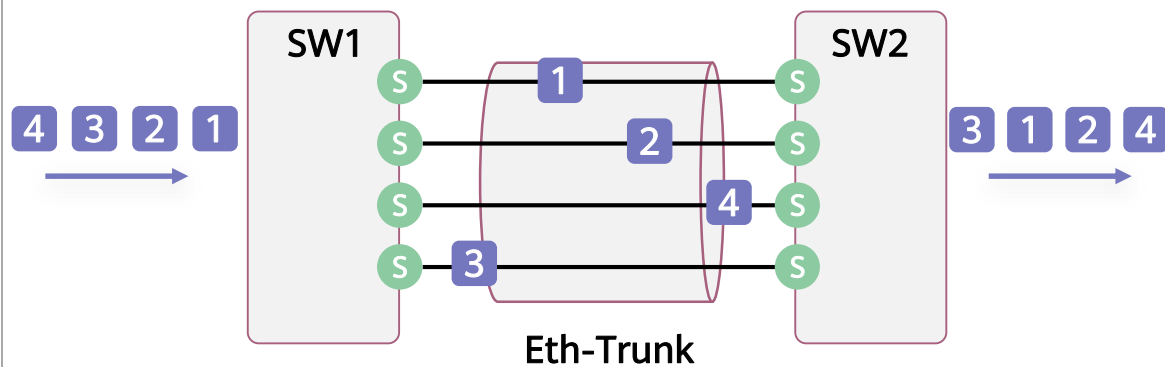


- SW2依据SW1的选举结果，明确本端的活动接口，同时对应的链路成为活动链路。
- 至此，Eth-Trunk的活动链路选举过程完成。

负载分担

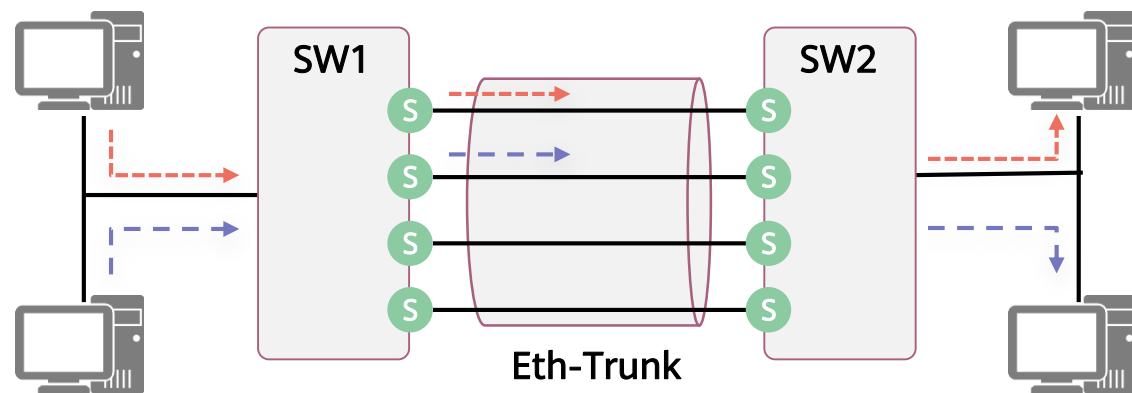
基于包的负载分担

S 活动接口



基于流的负载分担

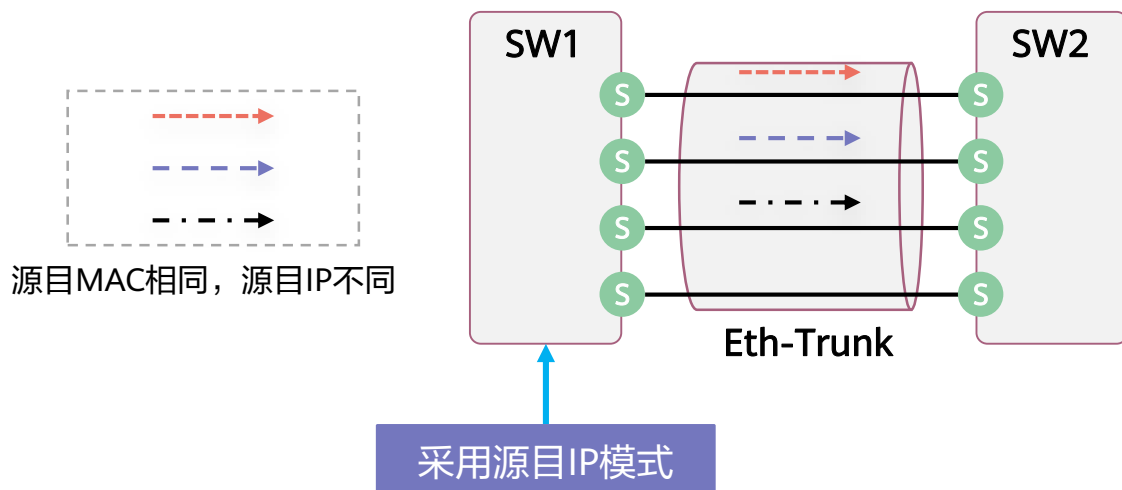
S 活动接口



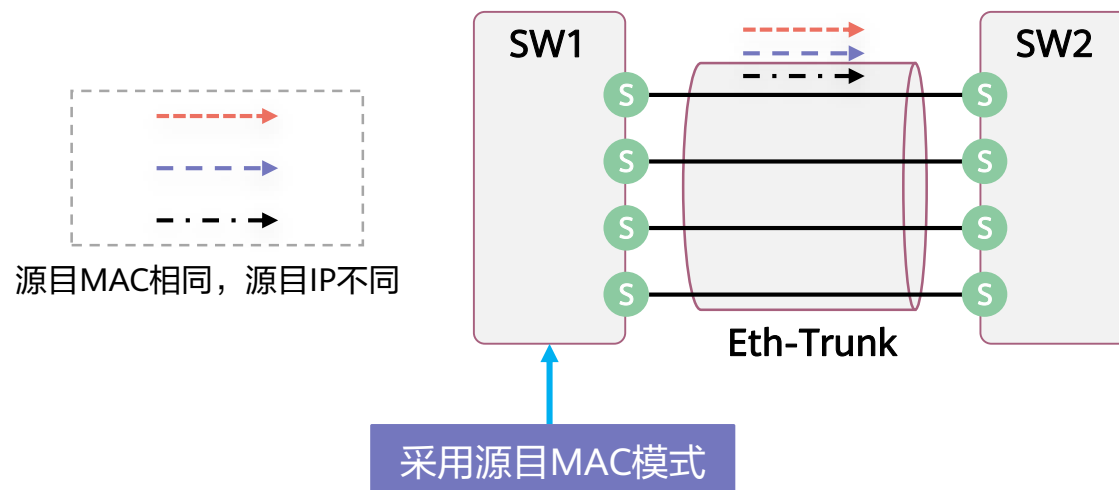
负载分担模式

- Eth-trunk支持基于报文的IP地址或MAC地址来进行负载分担，可以配置不同的模式（本地有效，对出方向报文生效）将数据流分担到不同的成员接口上。
- 常见的模式有：源IP、源MAC、目的IP、目的MAC、源目IP、源目MAC。

合适的负载分担算法



不合适的负载分担算法



目录

1

网络可靠性需求

2

链路聚合技术原理与配置

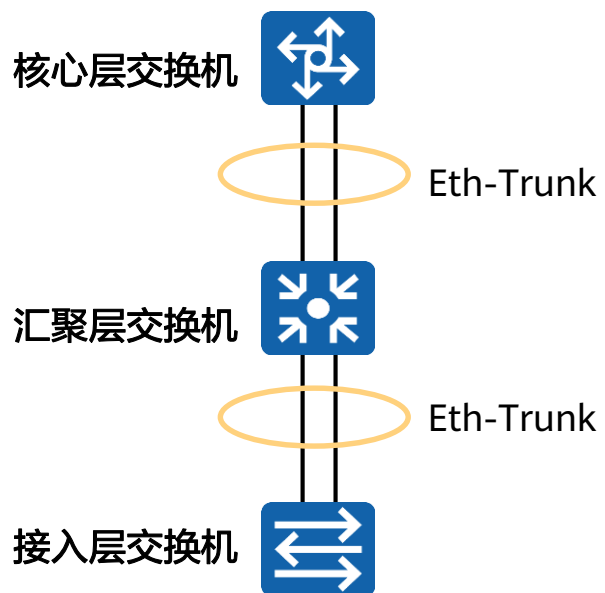
- 链路聚合基本原理
- 链路聚合手工模式
- 链路聚合LACP模式
- 链路聚合典型使用场景
- 链路聚合配置举例

3

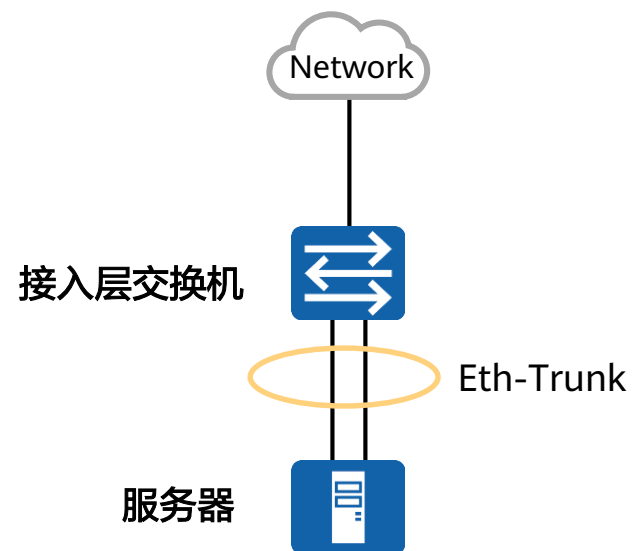
堆叠/集群概述

典型使用场景 (1)

交换机之间

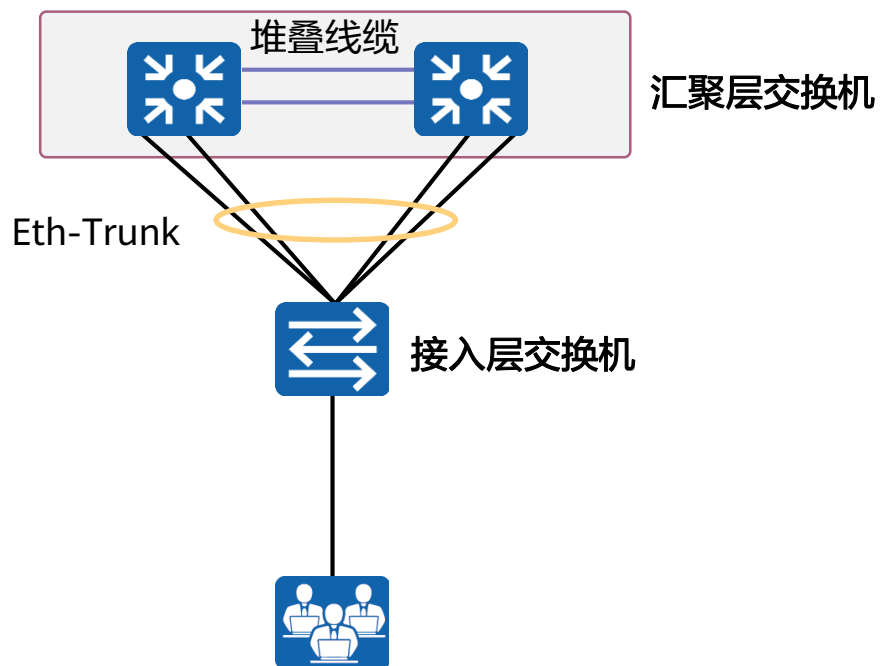


交换机与服务器之间



典型使用场景 (2)

交换机与堆叠系统



防火墙双机热备心跳线



目录

1

网络可靠性需求

2

链路聚合技术原理与配置

- 链路聚合基本原理
- 链路聚合手工模式
- 链路聚合LACP模式
- 链路聚合典型使用场景
- 链路聚合配置举例

3

堆叠/集群概述

配置命令介绍 (1)

1. 创建链路聚合组

```
[Huawei] interface eth-trunk trunk-id
```

创建Eth-Trunk接口，并进入Eth-Trunk接口视图。

2. 配置链路聚合模式

```
[Huawei-Eth-Trunk1] mode {lacp / manual load-balance }
```

Mode lacp配置链路聚合模式为lacp模式，mode manual load-balance配置链路聚合模式为手工模式。

注意：需要保持两端链路聚合模式一致。

3. 将接口加入链路聚合组中（以太网接口视图）

```
[Huawei-GigabitEthernet0/0/1] eth-trunk trunk-id
```

在接口视图下，把接口加入到Eth-Trunk中。

配置命令介绍 (2)

4. 将接口加入链路聚合组中（Eth-Trunk视图）

```
[Huawei-Eth-Trunk1] trunkport interface-type { interface-number}
```

在Eth-Trunk视图将接口加入到链路聚合组中。3、4两种方式都可以将接口加入到链路聚合组中。

5. 使能允许不同速率端口加入同一Eth-Trunk接口的功能

```
[Huawei-Eth-Trunk1] mixed-rate link enable
```

缺省情况下，设备未使能允许不同速率端口加入同一Eth-Trunk接口的功能，只能相同速率的接口加入到同一个Eth-Trunk接口中。

6. 配置系统LACP优先级

```
[Huawei] lacp priority priority
```

系统LACP优先级值越小优先级越高，缺省情况下，系统LACP优先级为32768。

配置命令介绍 (3)

7. 配置接口LACP优先级

```
[Huawei-GigabitEthernet0/0/1] lacp priority priority
```

在接口视图下配置接口LACP优先级。缺省情况下，接口的LACP优先级是32768。接口优先级取值越小，接口的LACP优先级越高。

只有在接口已经加入到链路聚合中才可以配置该命令。

8. 配置最大活动接口数

```
[Huawei-Eth-Trunk1] max active-linknumber {number}
```

配置时需注意保持本端和对端的最大活动接口数一致，只有LACP模式支持配置最大活动接口数。

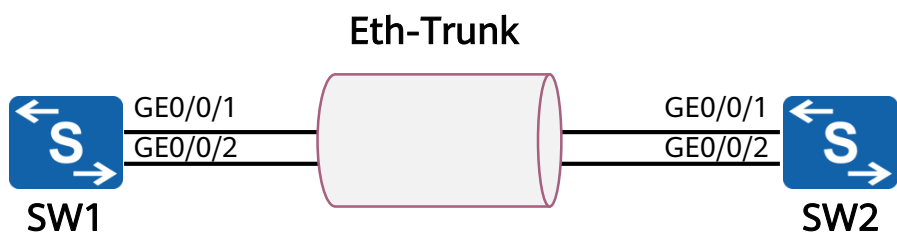
9. 配置最小活动接口数

```
[Huawei-Eth-Trunk1] least active-linknumber {number}
```

本端和对端设备的活动接口数下限阈值可以不同，手动模式、LACP模式都支持配置最小活动接口数。

配置最小活动接口数目的是为了保证最小带宽，当前活动链路数目小于下限阈值时，Eth-Trunk接口的状态转为Down。

手工模式链路聚合配置举例



SW1的配置如下:

```
[SW1] interface eth-trunk 1
[SW1-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/2
[SW1-Eth-Trunk1] port link-type trunk
[SW1-Eth-Trunk1] port trunk allow-pass vlan 10 20
```

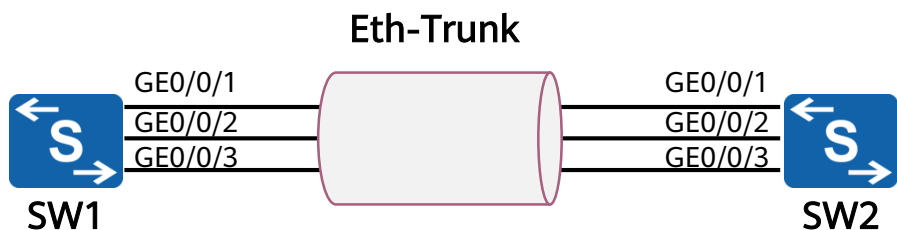
- 案例需求描述:

- SW1、SW2都连接着VLAN10、VLAN20的网络。
- SW1和SW2之间通过两根以太网链路互联，为了提供链路冗余以及保证传输可靠性，在SW1、SW2之间配置手工模式的链路聚合。

SW2的配置如下:

```
[SW2] interface eth-trunk 1
[SW2-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/2
[SW2-Eth-Trunk1] port link-type trunk
[SW2-Eth-Trunk1] port trunk allow-pass vlan 10 20
```

LACP模式链路聚合配置举例 (1)

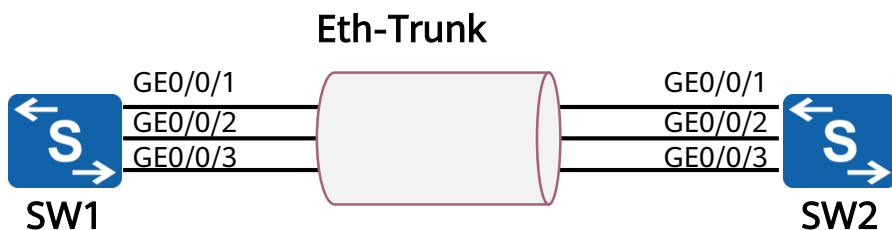


- 案例需求描述：
 - SW1、SW2都连接着VLAN10、VLAN20的网络。
 - SW1和SW2之间通过三根以太网链路互联，为了提供链路冗余以及保证传输可靠性，在SW1、SW2之间配置LACP模式的链路聚合，并且手动调整优先级让SW1成为主动端，并配置最大活跃端口为2，另外一条链路作为备份。

SW1的配置如下：

```
[SW1] interface eth-trunk 1
[SW1-Eth-Trunk1] mode lacp
[SW1-Eth-Trunk1] max active-linknumber 2
[SW1-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/3
[SW1-Eth-Trunk1] port link-type trunk
[SW1-Eth-Trunk1] port trunk allow-pass vlan 10 20
[SW1-Eth-Trunk1] quit
[SW1] lacp priority 30000
```

LACP模式链路聚合配置举例 (2)



SW2的配置如下:

```
[SW2] interface eth-trunk 1
[SW2-Eth-Trunk1] mode lacp
[SW2-Eth-Trunk1] max active-linknumber 2
[SW2-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/3
[SW2-Eth-Trunk1] port link-type trunk
[SW2-Eth-Trunk1] port trunk allow-pass vlan 10 20
[SW2-Eth-Trunk1] quit
```

- 案例需求描述:
 - SW1、SW2都连接着VLAN10、VLAN20的网络。
 - SW1和SW2之间通过三根以太网链路互联，为了提供链路冗余以及保证传输可靠性，在SW1、SW2之间配置LACP模式的链路聚合，并且手动调整优先级让SW1成为主动端，并配置最大活跃端口为2，另外一条链路作为备份。

目录

1 网络可靠性需求

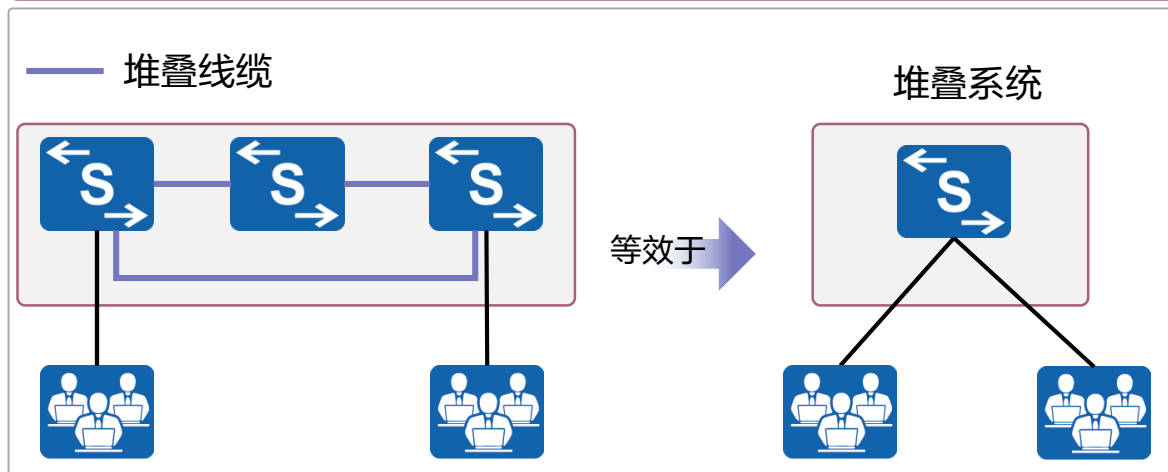
2 链路聚合技术原理与配置

3 堆叠/集群概述

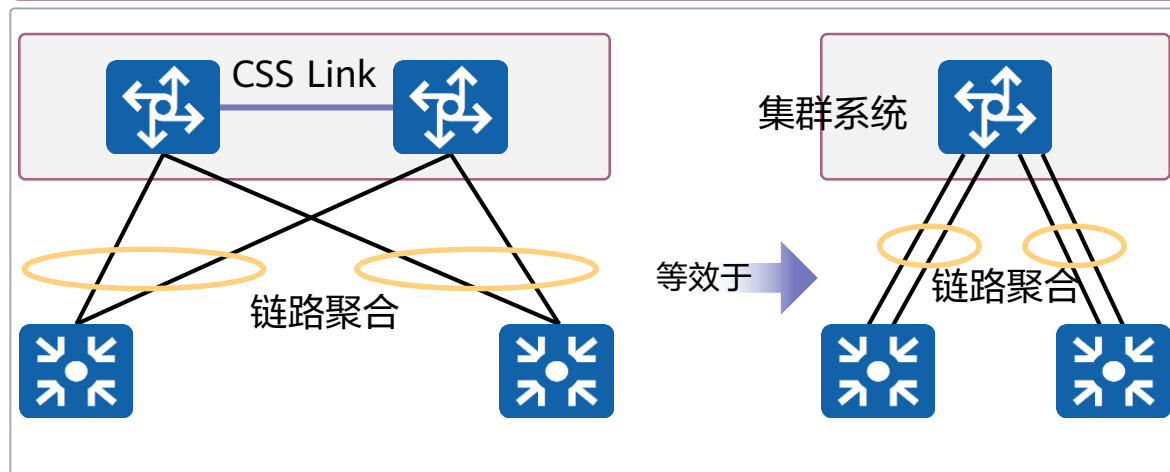
- 堆叠/集群概述

什么是堆叠、集群

堆叠-iStack

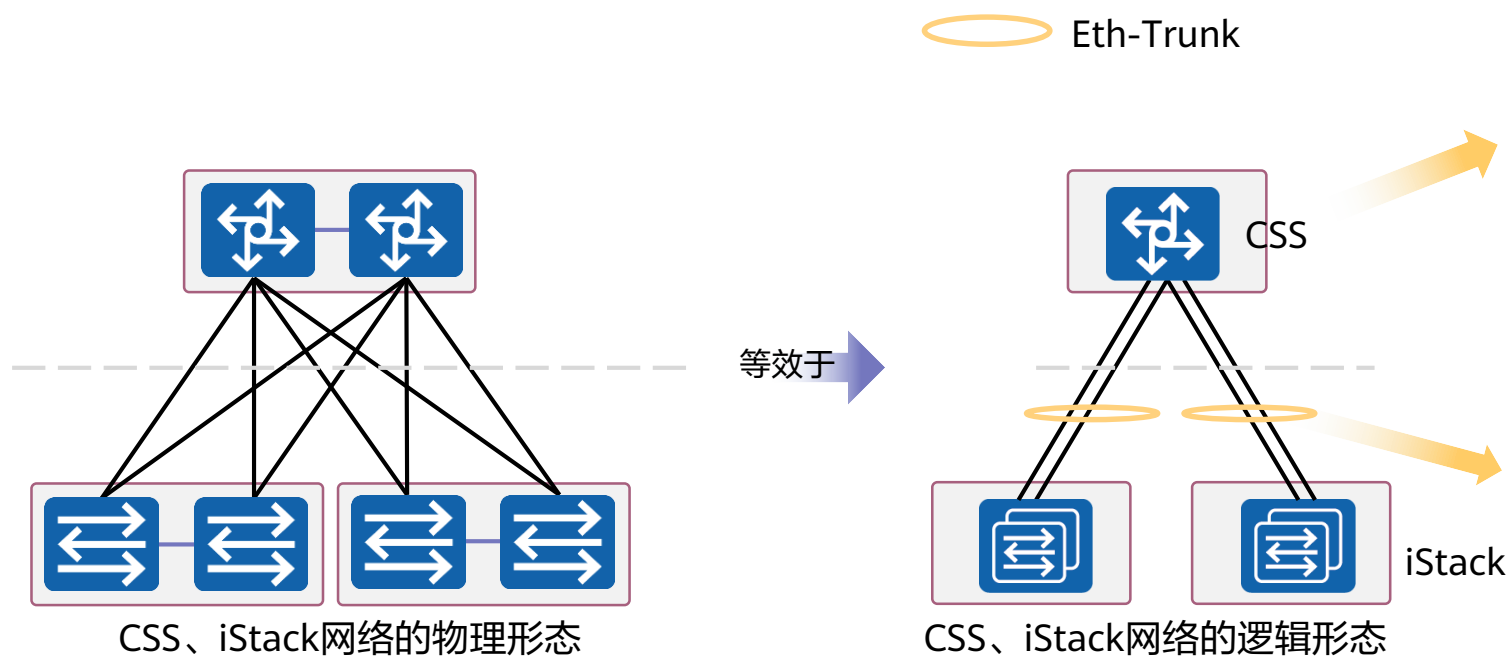


集群-CSS



- 堆叠（iStack）：多台支持堆叠特性的交换机通过堆叠线缆连接在一起，从逻辑上变成一台交换设备，作为一个整体参与数据转发。
- 集群（Cluster Switch System, CSS）：将两台支持集群特性的交换机设备组合在一起，从逻辑上组合成一台交换设备。
- 集群只支持两台设备，一般框式交换机支持CSS，盒式设备支持iStack。

堆叠、集群的优势

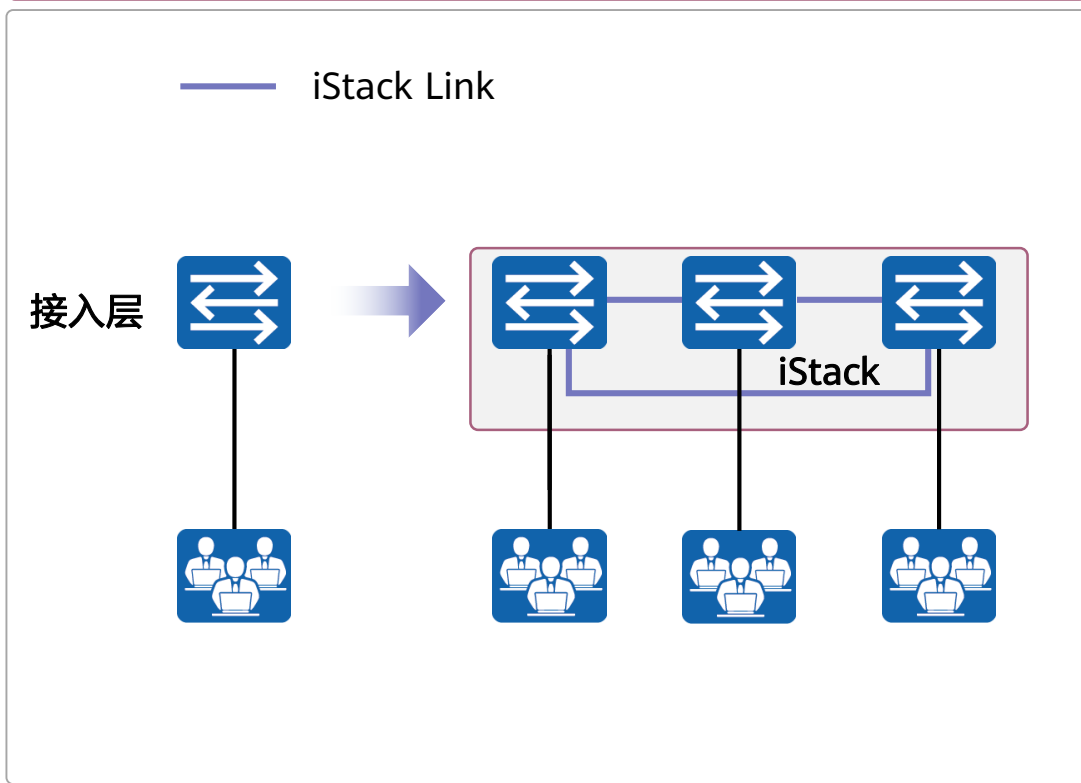


- 逻辑上一台设备，简化运维，方便管理。
- 一台物理设备故障，其他设备可以接管转发、控制平台，避免了单点故障。

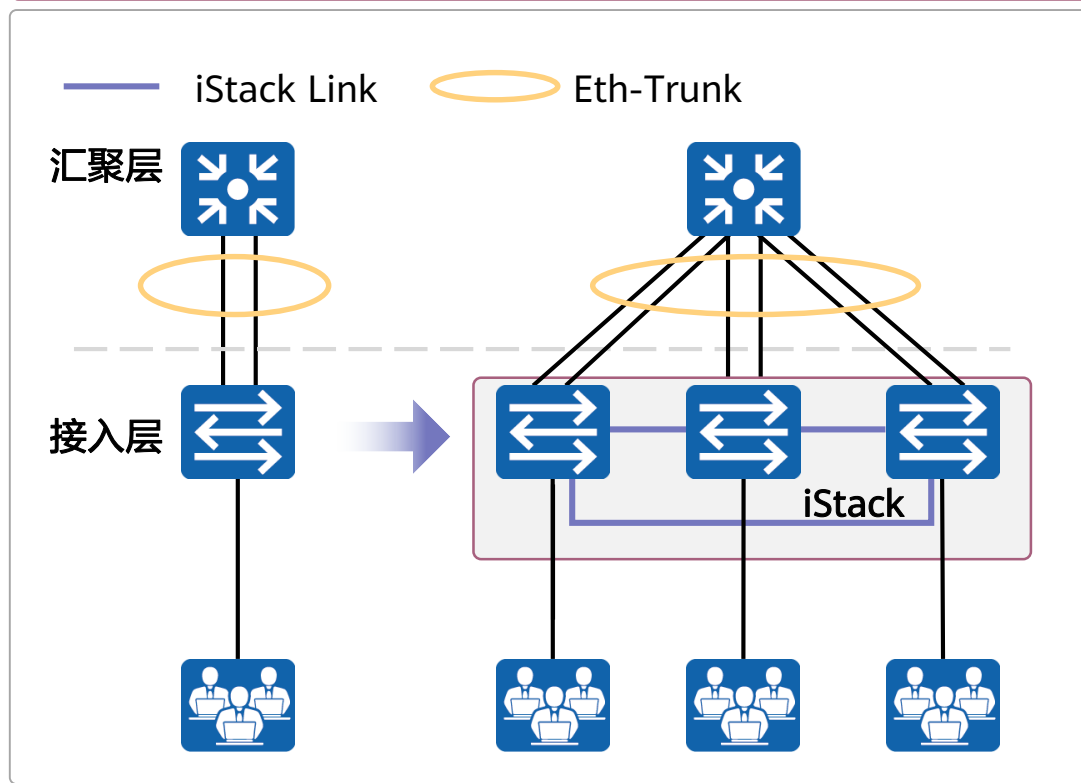
- 跨设备的链路聚合，物理上的无环网络，无需再部署STP。
- 链路聚合中的链路全部有效使用，链路利用率100%。

实际应用 (1)

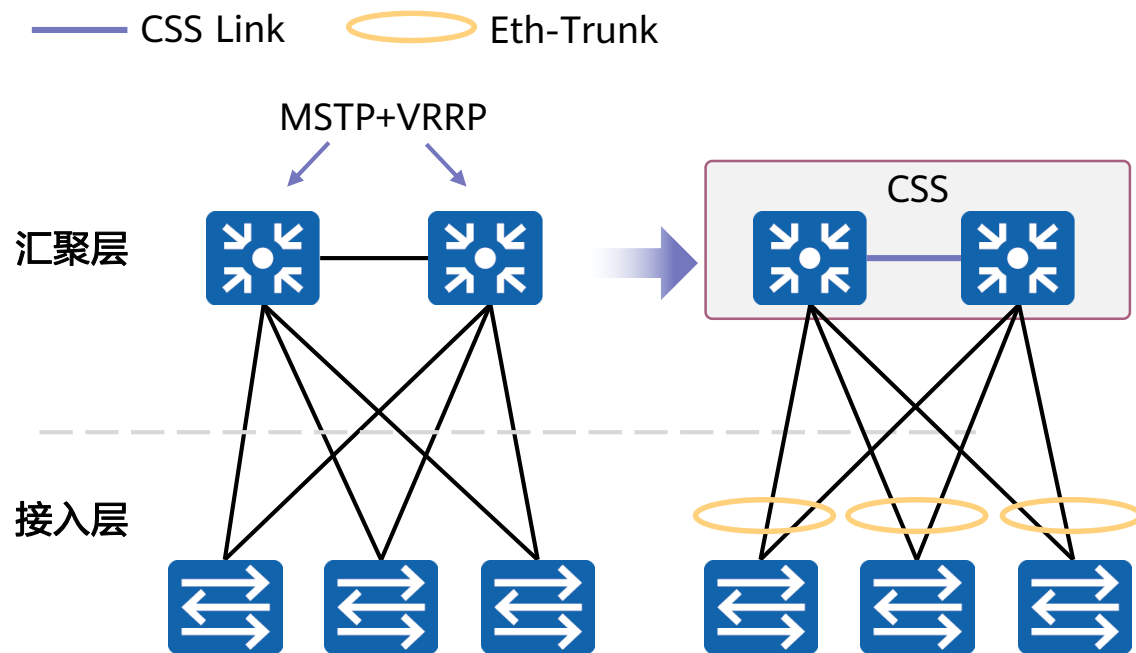
扩展端口



扩展带宽、冗余备份

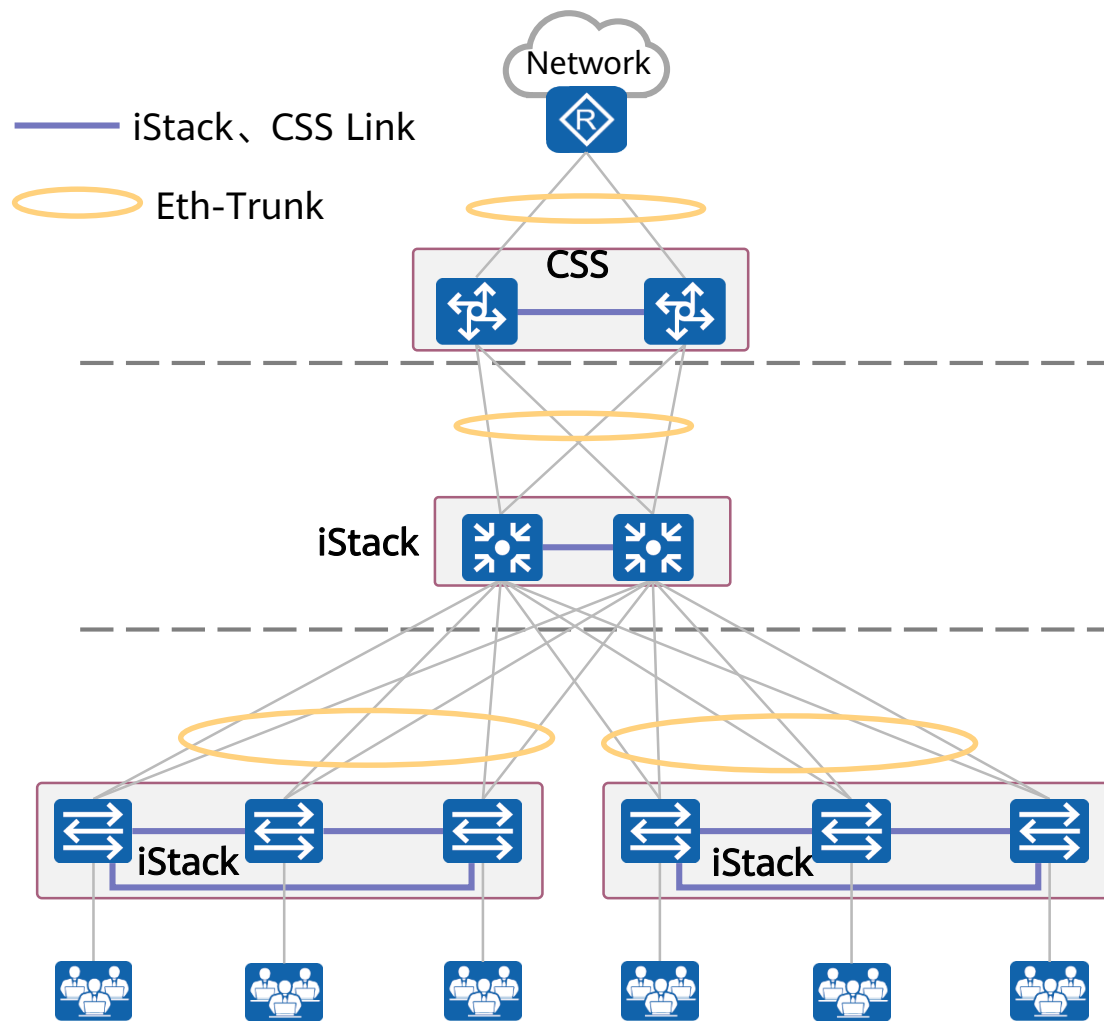


实际应用 (2)



- 两台设备组成集群，虚拟成单一的逻辑设备。简化后的组网不再需要使用MSTP、VRRP等协议，简化了网络配置，同时依靠跨设备的链路聚合，实现快速收敛，提高了可靠性。

推荐架构



核心层

- 核心使用CSS集群组网，上下行采用Eth-Trunk，构建高可靠、无环的网络。

汇聚层

- 汇聚交换机采用iStack，上下行采用Eth-Trunk，构建高可靠、无环的网络。

接入层

- 地理位置接近的接入设备（如一个楼宇内的接入交换机）使用iStack虚拟化成为一台逻辑上的设备，端口数量充足，简化了管理。
- 使用Eth-Trunk和汇聚层互联，逻辑上网络结构简单，不再需要使用STP、VRRP。具有高可靠性、高上行带宽、快速收敛的优势。

本章总结

- 链路聚合技术可以分为静态聚合和LACP模式聚合。
- LACP模式采用报文协商，可以实现活动链路的备份，在链路出现故障时将备份链路选举为活动链路继续参与转发。
- 为保证报文到达的顺序，链路聚合的负载分担采用基于流的形式。
- 使用iStack、CSS技术可以简化网络管理、简化网络结构、提高网络可靠性。