

Huawei Enterprise **A Better Way**

交换机横向虚拟化：CSS

www.huawei.com

HUAWEI TECHNOLOGIES CO., LTD.



内容综述

- **通过本胶片，您能够了解到：**
 - › **什么是CSS？CSS可以为我们带来什么？**
 - » CSS的定义、用途
 - › **CSS在CE12800上面是怎样实现的？有哪些特征？**
 - » CSS的技术实现
 - » CSS对外表现的特征和能力
 - › **CSS要如何部署和应用？**
 - » CSS的网络部署方式
 - » CSS的组网性能

目录

1. CSS简介

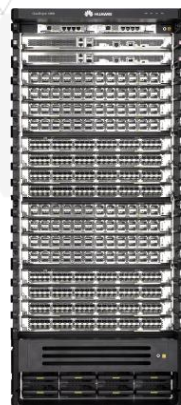
2. CSS在CE12800上的实现

3. CSS的部署和应用

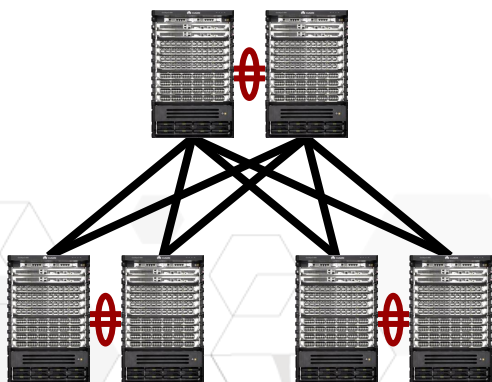
什么是CSS ?



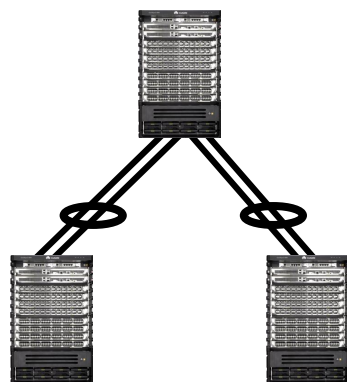
CSS设备物理形态



CSS设备逻辑形态



CSS网络物理形态



CSS网络逻辑形态

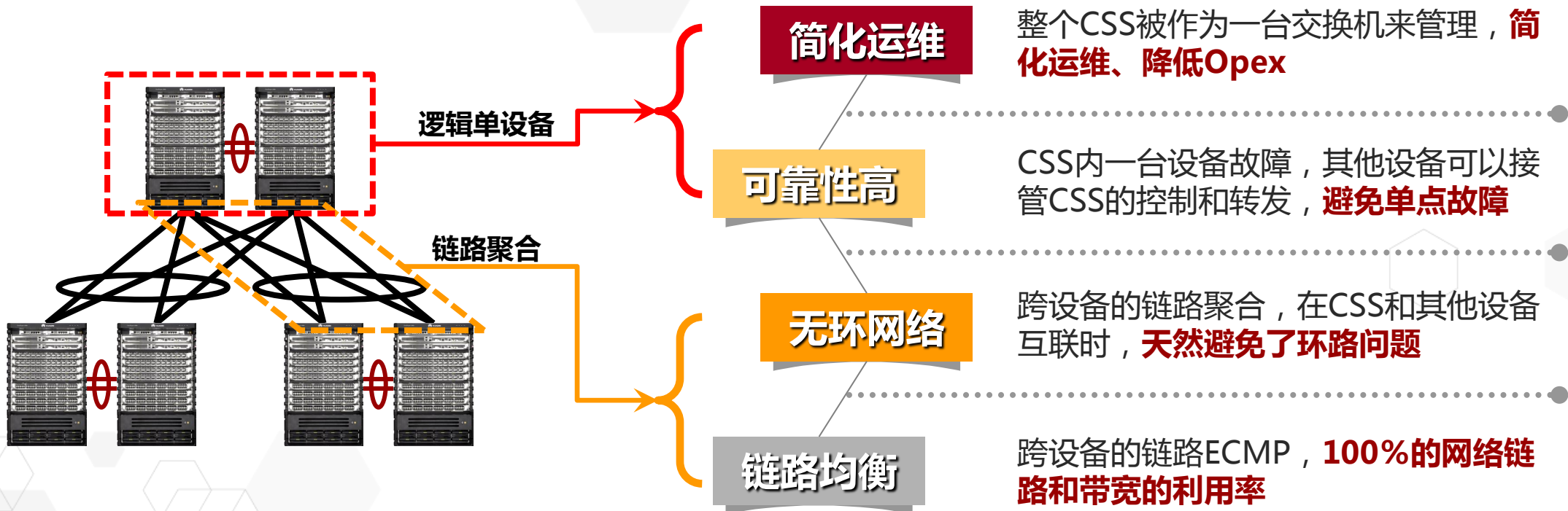
● CSS的定义

- **Cluster Switch System** : 集群交换机系统。是指把多台支持集群的交换机链接起来，组成一台更大的交换机

● CSS的特征

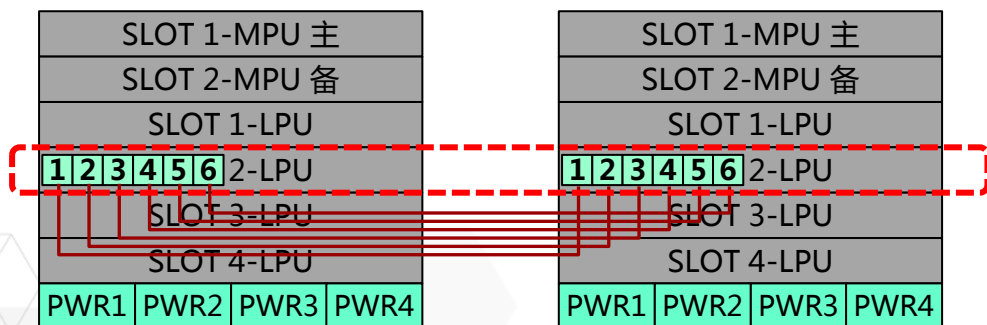
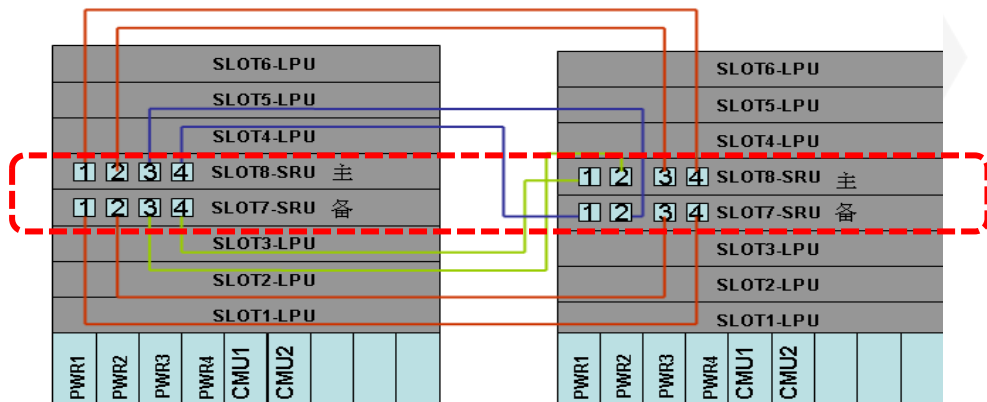
- **交换机多虚一** : CSS对外表现为一台逻辑交换机，控制平面合一，统一管理
- **转发平面合一** : CSS内物理设备转发平面合一，转发信息共享并实时同步
- **跨设备链路聚合** : 跨CSS内物理设备的链路被聚合成一个ETH-TRUNK端口，和下游设备实现互联

CSS为我们带来什么？



简化的管理、高效的网络，中小规模数据中心的首选方案

CSS的种类和特征



● 专用线卡式CSS

- **部署方式**：CSS内各交换机使用**专用线卡**、配以**专用线缆**实现互联
- **部署收益**：不占用线卡槽位、高速端口增强CSS互联带宽
- **典型产品**：**S9300**

● 业务线卡式CSS

- **部署方式**：CSS内各交换机使用**业务线卡**、配以**标准线缆**实现互联
- **部署收益**：无需采购额外部件、支持长距CSS、ETH-Trunk口提升CSS互联带宽
- **典型产品**：**CE12800**

随着设备性能的显著提升，业务线卡式CSS将大放异彩

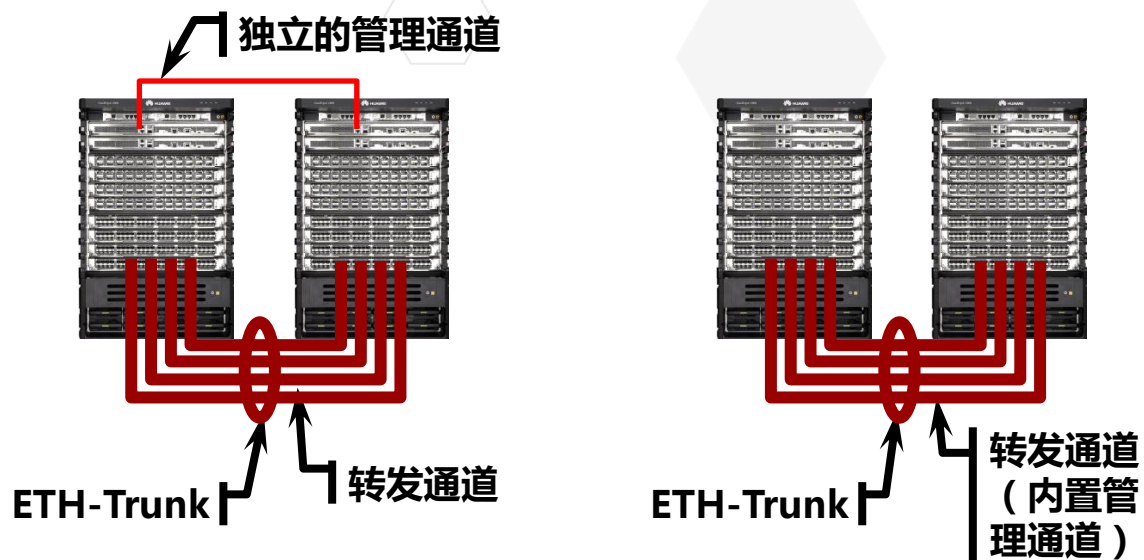
目录

1. CSS简介

2. CSS在CE12800上的实现

3. CSS的部署和应用

CSS的物理架构



类型	带外管理通道型CSS	带内管理通道型CSS
部署	主控板的GE端口互联，作为管理通道	管理通道部署在CSS转发通道内
优势	带外管理，可靠性高，不占用业务资源	简化网络部署，节省布线，便于CSS拉远

灵活的CSS架构

- 产品型号自由组合：支持不同型号的CE12800交换机组成CSS
- 两种管理通道部署：管理通道支持带内、带外两种模式
- 多种转发通道带宽：转发通道兼容10GE、40GE端口，未来可支持100GE

高速的CSS互联

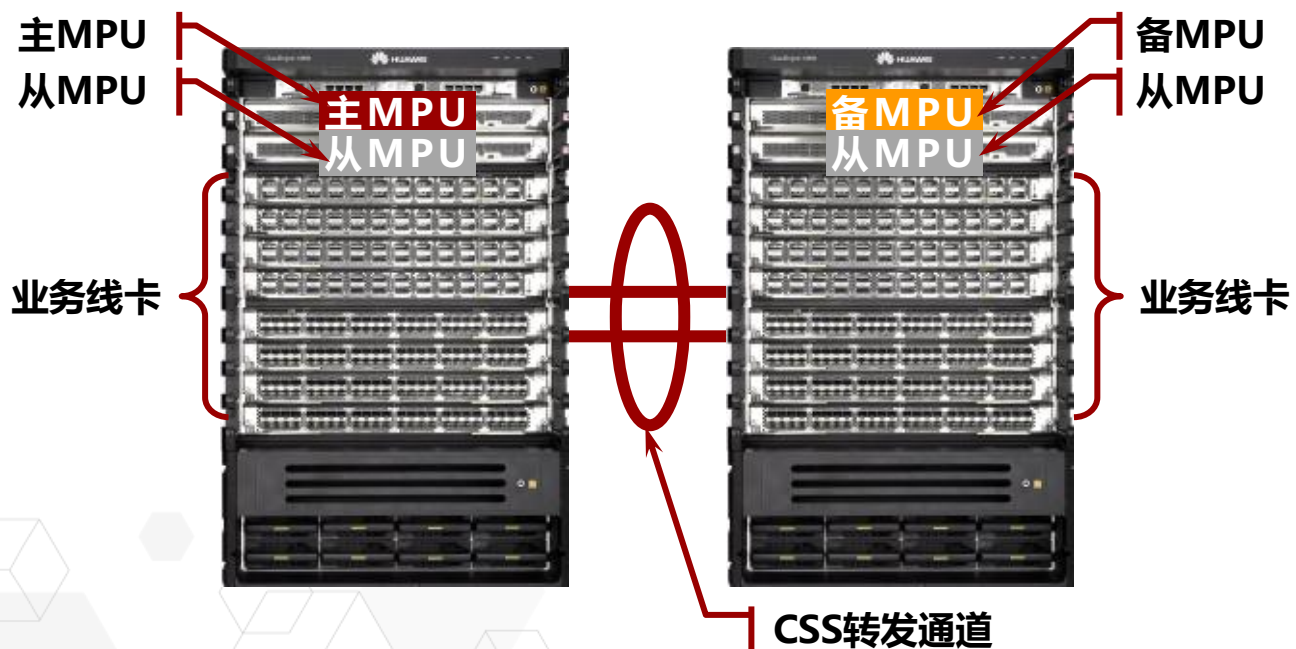
- 支持**16个端口**聚合成CSS转发通道
- 转发通道支持单端口**40GE**
- CSS互联带宽可达到**单向640G bps**

名词解释：

管理通道：CSS设备间管理和控制平面的交互通道。负责：设备管理、路由同步、CSS状态同步，等工作

转发通道：CSS设备间转发平面的交互通道。负责：报文转发、MAC表同步，等工作

CSS系统内的职责分工



● MPU

- **主MPU**：负责整个CSS系统的管理
- **备MPU**：主MPU的热备份，在主MPU故障时承担CSS的管理
- **从MPU**：作为MPU的冷备份，一旦主备MPU中的一个发生故障，从MPU将自动升级为备MPU

● CSS转发通道

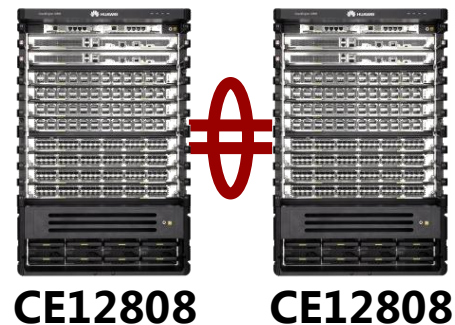
- CSS主备间的业务转发、数据同步
- 通过业务线卡互联组成，所有互联链路组成一个ETH-Trunk端口
- 支持跨子卡的端口组成转发通道

● 业务线卡

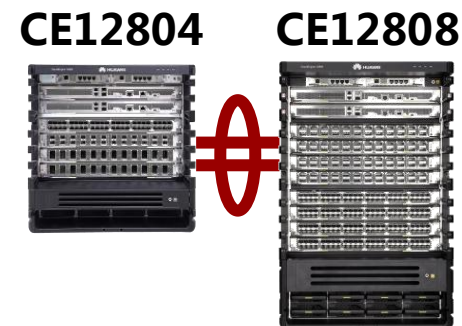
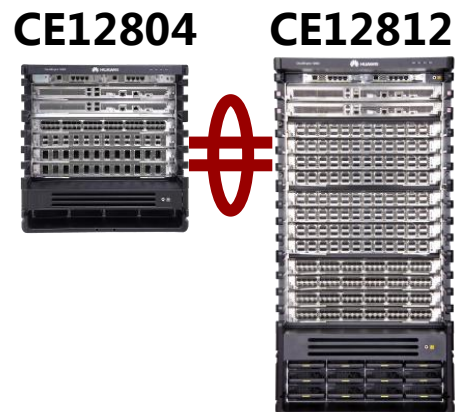
- 向CSS系统外提供业务接入和转发
- CSS转发通道的端口不能够向外提供业务

CE12800支持不同型号间CSS的自由组合

同产品组成CSS



跨产品组成CSS



CSS的一生

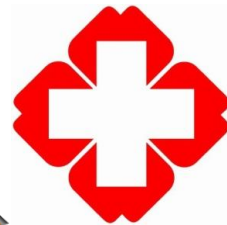
CSS定期体检

CSS内各设备间执行**双主检测**，及时发现CSS内的故障



CSS就医诊治

CSS的**分裂处理**



CSS诞生

自动**协商CSS主备、同步配置和数据**，启动CSS的运行



CSS日常生活

CSS运行过程中，处理**业务转发、实时同步配置和数据**



CSS寿终正寝

CSS的**解除**

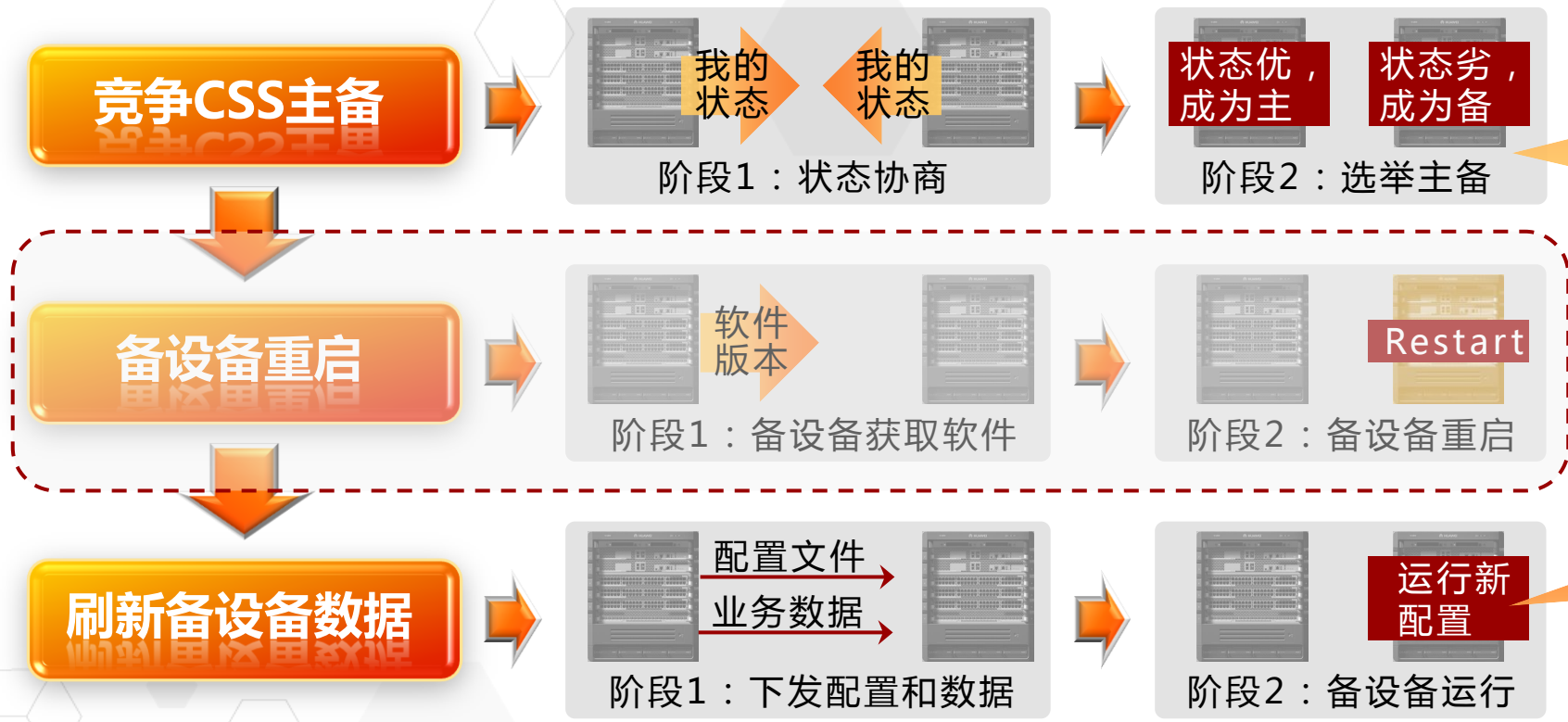


CSS的启动：主备竞争、配置同步

- 主备竞争算法**
- 运行状态----优
 - CSS优先级--高
 - 软件版本号--新
 - 设备MAC---小

- 主备版本不一致：执行此步骤；
- 版本一致：跳过此步骤，直接刷新备设备数据

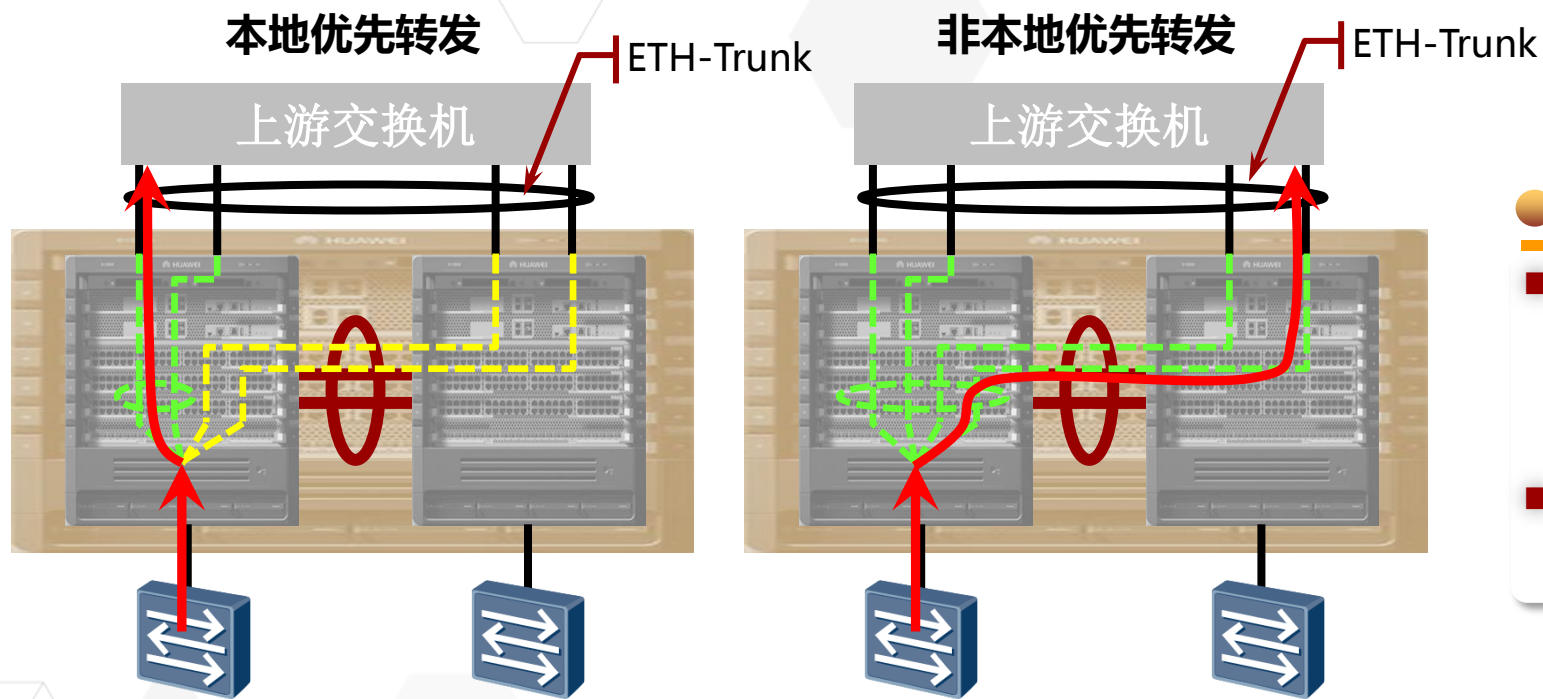
- 备设备原有配置将被删除



主备完成协商，CSS开始运行



CSS转发平面的处理：本地优先转发



本地优先转发

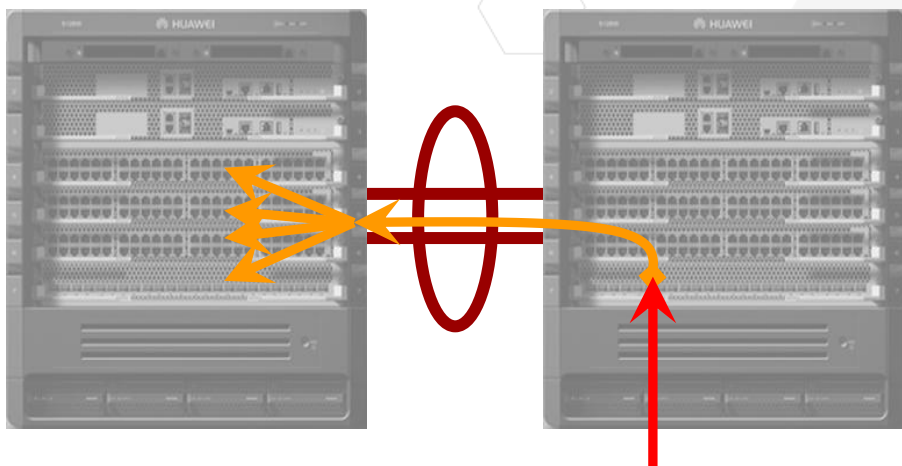
- **转发行为**：CSS会从连接上游设备的ETH-Trunk端口中，首先选择属于本物理设备的端口完成报文的转发
- **优势**：减轻CSS转发通道的负担，减少转发延迟

- 优选hash链路，优先被用于链路hash选择
- 次优hash链路，优选hash链路故障时使用

通过本地转发，实现更精细化的业务处理，高效使用系统资源

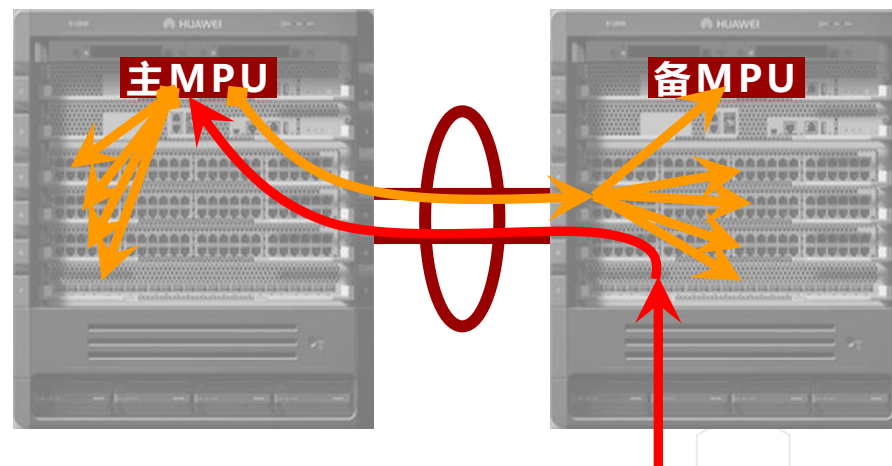


CSS运行中，需要实时同步配置和数据



● 无需控制平面参与的数据同步

- **同步动作**：线卡之间直接完成数据同步，主设备和备设备的线卡都可以发起同步。跨框同步动作，由CSS转发通道完成
- **典型业务**：MAC表的同步



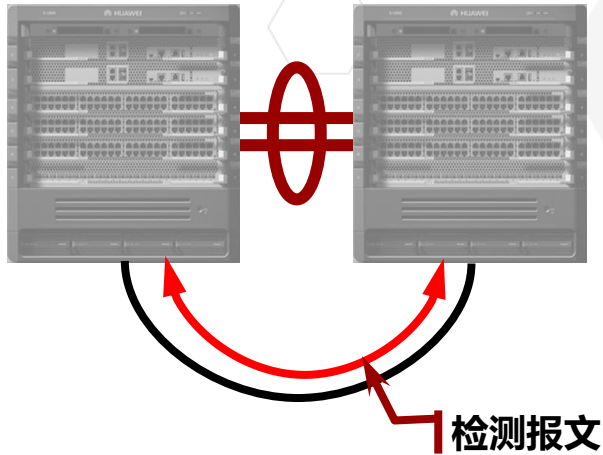
● 需要控制平面参与的数据同步

- **同步动作**：所有信息上送主设备，由主MPU执行控制平面的计算和处理，然后向所有的线卡、备MPU发送数据。跨框同步动作，由CSS管理通道完成
- **典型业务**：路由表、ARP表的同步

CSS逻辑上是一台设备，在整个系统内实时同步配置和转发信息

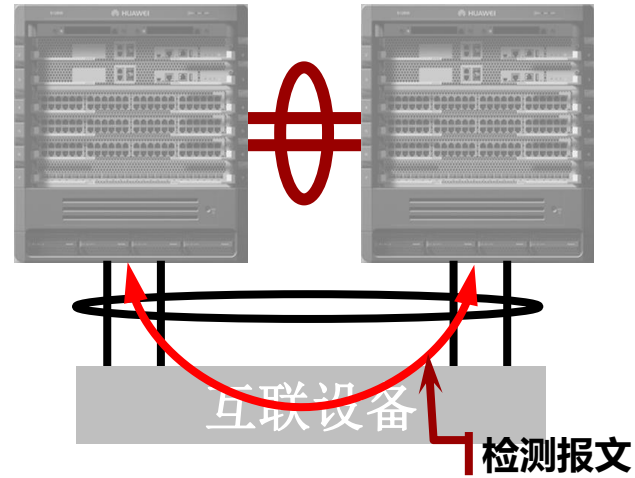


CSS健康状态的实时监控：双主检测



● 通过直连链路检测

- **报文收发**：CSS主备设备之间部署直连链路，用于收发检测报文
- **辅助需求**：直连链路上不能运行任何业务，CSS通过block端口屏蔽环路



● 通过业务链路检测

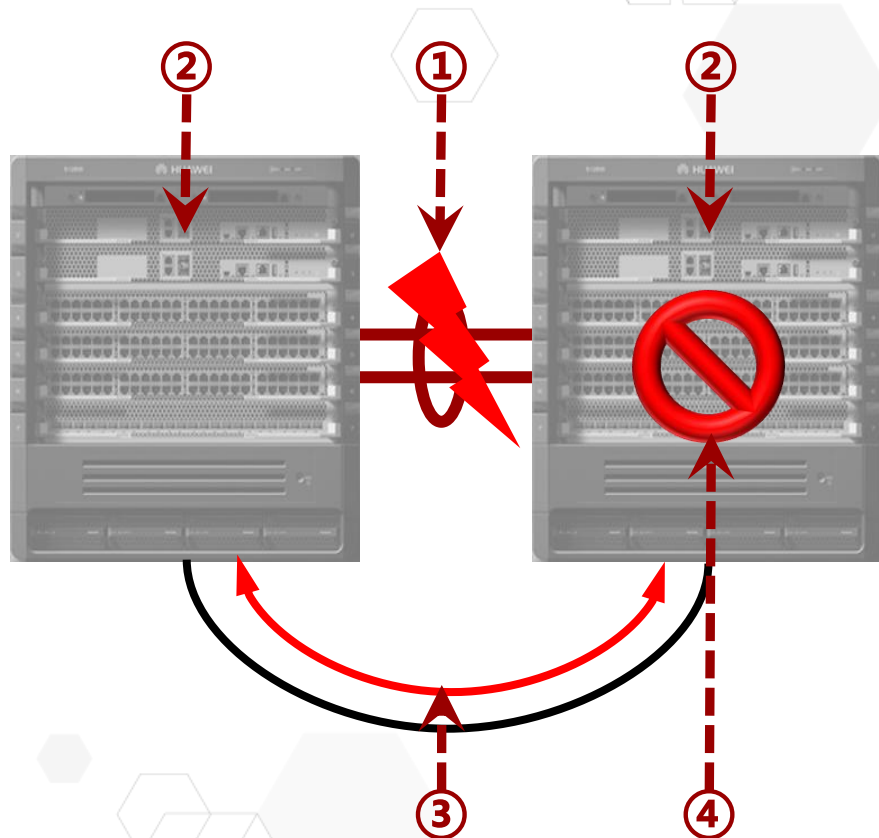
- **报文收发**：在CSS和互联设备之间的ETH-Trunk链路上发送和接收
- **辅助需求**：互联设备上启动CSS代理，把检测报文从ETH-Trunk端口发回去

备注：需要使用华为的交换机作为互联设备来启动CSS代理

通过双主检测，可以避免CSS系统的故障给整个网络带来影响



CSS的分裂



● CSS分裂的问题

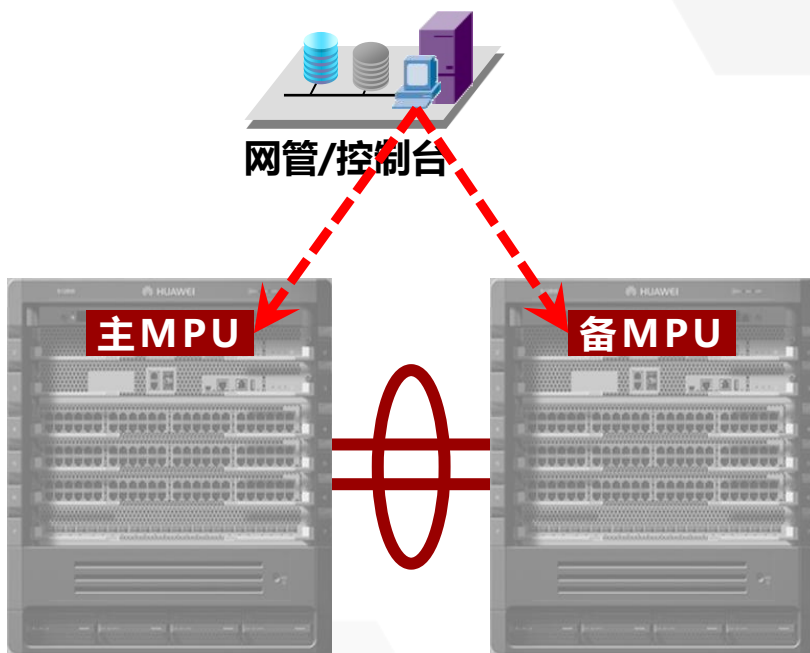
- **CSS分裂的定义**：主备设备间的CSS管理通道断裂，导致两台设备独立于网络之中
- **CSS分裂的影响**：由于两台设备配置了CSS，会使用相同的MAC、IP来处理业务。一旦分裂，会导致网络中业务的混乱

● CSS分裂的处理

- ① CSS管理通道故障，主备设备的联系中断
- ② 两台设备重新竞争主备，并且都默认自己是“主”
- ③ 通过双主检测，发现两台设备都是“主”
- ④ 优先级低的设备需要把所有的业务端口 Shutdown

为避免影响整个网络，CSS分裂后，只有一台设备允许转发业务

CSS的解除



● CSS的解除

- ① 操作员分别向主备设备下发CSS的“Disable”命令
- ② 操作员分别向主备设备下发Restart命令
- ③ 设备重启后，成为两台独立设备，此时操作员可以依据网络部署需要，重新下发配置和管理这两台设备

CSS解除后，两台设备各自独立运行，重新获取配置和执行转发

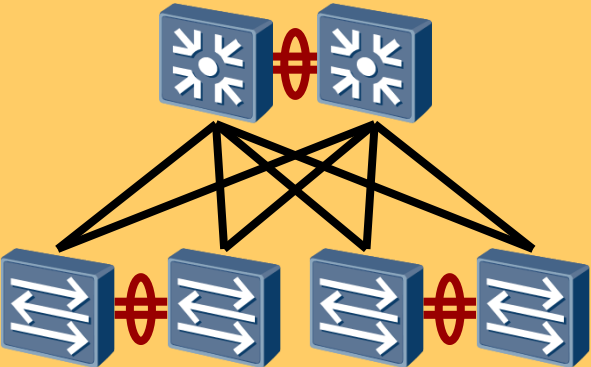
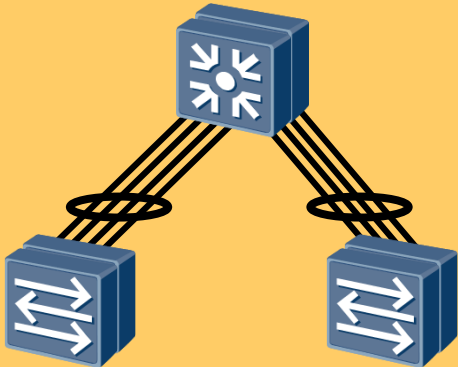
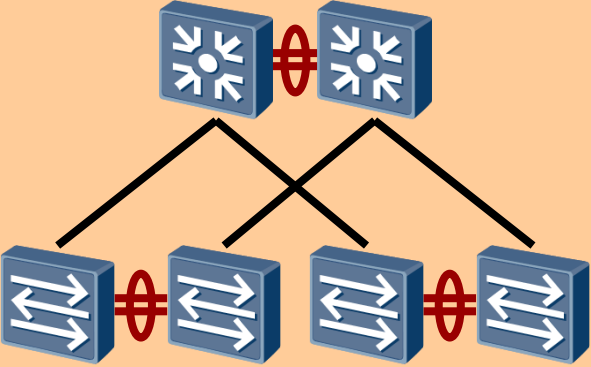
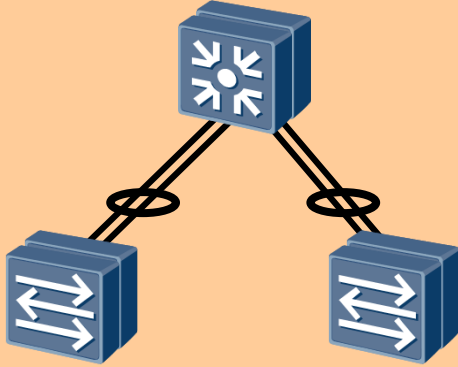
目录

1. CSS简介

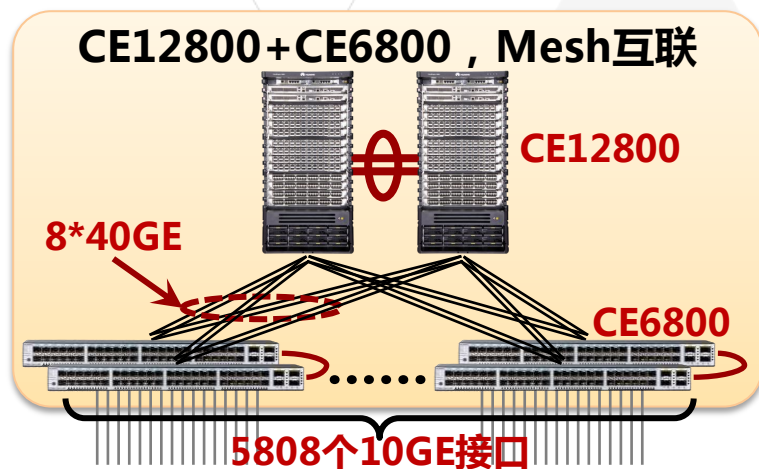
2. CSS在CE12800上的实现

3. CSS的部署和应用

CSS的两种典型部署拓扑

部署拓扑	网络物理形态	网络逻辑形态	方案说明
Mesh互联			<p>特征：上游和下游的CSS系统间，两两互联组成Mesh拓扑</p> <p>优势：更高的可靠性、更大的网络带宽</p> <p>应用：适用于链路资源充足，转发性能要求高的场景</p>
单链路互联			<p>特征：上游和下游的CSS系统间，物理设备单链路互联</p> <p>优势：在提供了可靠性保护的前提下，节省链路资源</p> <p>应用：适用于链路资源有限，或布线数量受限的场景</p>

CSS的实际部署举例



TOR的接入数量（按照物理TOR交换机计算）：
 $24 (40GE端口/槽位) * 11 (槽位/台) * 2 (台) / 4 (端口/TOR) = 132$ 台

接入能力：
10GE端口数： $44 (个10GE/TOR) * 132 (TOR) = 5808$ 个

收敛比： $(4*40GE) : (44*10GE) = 1 : 2.75$

组网描述

- CE12800和CE6800组成Mesh网络
- CE12800：使用1个业务线卡做CSS转发通道互联、上行设备接入；剩下11个业务线卡用于接入TOR交换机
- CE6800：每2台CE6800堆叠在一起，占用4个10GE端口；每个堆叠组上行8个40GE端口、下行88个10GE端口

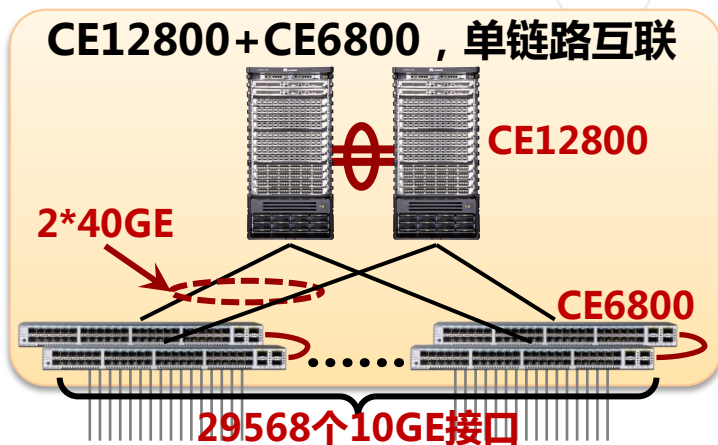
方案优势

- **高可靠的网络**：Mesh组网
- **高性能的网络**：41T的双向转发能力
- **高接入的网络**：5808个10GE端口

备注：CSS最常见的组网形式，是和TOR交换机组建无环的二层网络

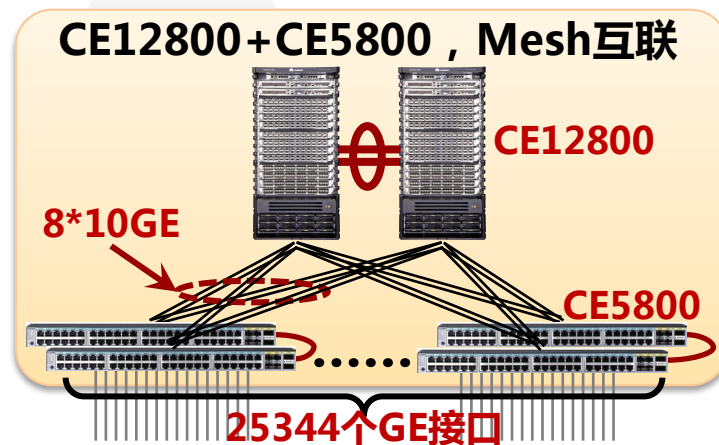
高性能的产品，带来高性能的网络、提供高质量的服务

CSS的其他部署实例



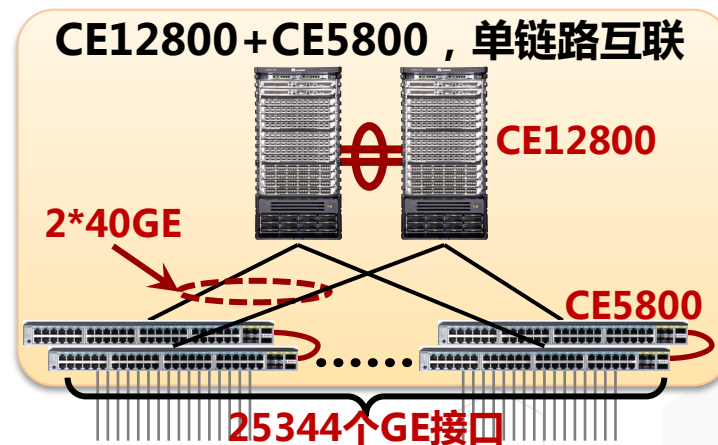
网络特点

- 单链路互联：节约端口、简化布线
- 40GE互联：提升转发性能
- TOR数量：528台
- 高密接入：29568个10GE端口
- 接入收敛比：1：14



网络特点

- Mesh互联：提升可靠性和转发性能
- TOR数量：528台
- 高密接入：25344个GE端口
- 接入收敛比：1：1.2



网络特点

- 单链路互联：节约端口、简化布线
- 40GE互联：提升转发性能
- TOR数量：528台
- 高密接入：25344个GE端口
- 接入收敛比：1：1.2

灵活的部署方案，充分满足各种实际部署需求

内容总结

- **CSS的定义**

- › Cluster Switch System，通过虚拟化，把2台物理交换机虚拟化成1台逻辑交换机。
- › CSS不但提升了单台设备的能力，并且简化了管理、降低了部署复杂度

- **CSS的实现**

- › 业务线卡式CSS，使用业务线卡组成CSS转发通道。支持ETH-Trunk绑定多个端口，扩展CSS转发通道带宽。
- › 灵活的CSS组成，相同型号交换机、不同型号交换机，都可以组成CSS系统。
- › 简单的配置、自主的运行、主动的排障、优化的转发。

- **CSS的部署**

- › 灵活的布线方案、灵活的设备选型。
- › 配合高性能的CE12800交换机，提供高性能、高可靠的CSS系统和网络。



HUAWEI

Huawei Enterprise *A Better Way*