

RSTP 协议原理与配置

IEEE，全称是 Institute of Electrical and Electronics Engineers，电气和电子工程师协会 是一个国际性的电子技术与信息科学工程师的协会。

IEEE 于 1980 年 2 月成立了局域网标准委员会（简称 IEEE 802 委员会），专门从事局域网标准化工作，并制定了 IEEE 802 标准。

STP : IEEE 802.1d

RSTP : IEEE 802.1w

MSTP : IEEE 802.1s

```
[Huawei]dis stp
-----[CIST Global Info][Mode STP]-----
CIST Bridge           :32768.4c1f-cc96-6d5b
Config Times          :Hello 2s MaxAge 20s FwDly 15s MaxHop 20
Active Times          :Hello 2s MaxAge 20s FwDly 15s MaxHop 20
```

生成树计时器

生成树协议中用到 hello，Forward delay 和 max age 这 3 个计时器，它们会影响端口状态迁移和收敛时间，可在全局使用命令修改。

调整计时器一定要在根交换机上配置，其他交换机使用根桥交换机的计时器工作，根桥交换机的 BPDU 中的计时器优于交换机本地计时器的配置。

```
[Huawei]stp timer ?  
  forward-delay Specify forward delay  
  hello          Specify hello time interval  
  max-age        Specify max age
```

stp timer hello 300 改为 3 S
stp timer forward-delay 2000 改为 20 S
stp timer max-age 3000 改为 30 S
stp max-hops 30 改为 30 跳

hello timer : 2s root 每 2s 产生 BPDU。根交换机产生的 BPDU 的通告时间间隔

forward delay : 15s 设备状态迁移的延迟时间。
链路故障会引发网络重新进行生成树的计算，生成树的结构将发生相应的变化。不过重新计算得到的新拓扑信息无法立即传到整个网络。此时若立即将新选出的根端口和指定端口置于数据转发的状态，则可能会出现临时环路。这时要求新新选出的根端口和指定端口要经过 2 倍的 forward delay 后才能进入转发状态，这个延时足够保证新的配置消息能传遍整个网络。

max age : 20s 最大老化时间
储存 BPDU 的时间，spanning-tree 发生故障，20s 后原 blocking 状态->learning 状态

以非根桥的根接口为例，该设备将在这个接口上保存来自上游的最优 BPDU，这个 BPDU 关联着一个最大生存时间，如果在该 BPDU 到达最大生存时间之前，接口再一次收到了 BPDU，那么其最大生存时间将会被重置，而如果接口一直没有再收到 BPDU 从而导致该接口上保存的 BPDU 到达最大生存时间，那么该 BPDU 将被老化，此时设备将重新在接口上选择最优 BPDU，也就是重新进行根接口的选举。

max hop : 20

当一个 BPDU 报文到达一个桥，又被该桥转发称为一跳，当 BPDU 报文的跳数超过 max hop 时，该报文会被丢弃，该参数与网络规模相关。

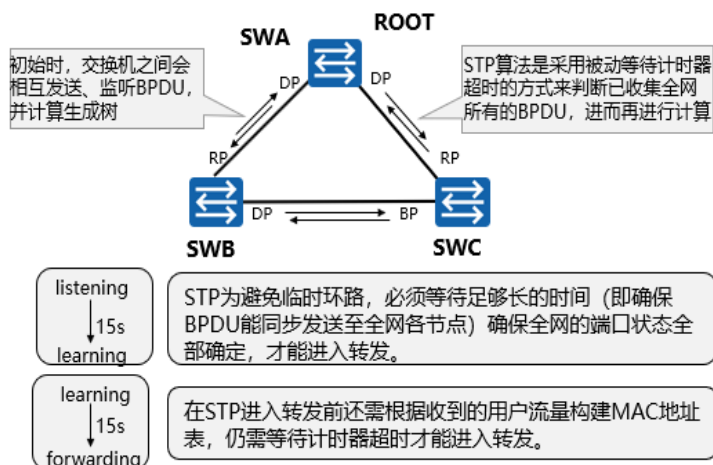
前言

- STP协议虽然能够解决环路问题，但是由于网络拓扑收敛较慢，影响了用户通信质量，而且如果网络中的拓扑结构频繁变化，网络也会随之频繁失去连通性，从而导致用户通信频繁中断，这也是用户无法忍受的。
- 由于STP的不足，IEEE于2001年发布的802.1w标准定义了RSTP。RSTP在STP基础上进行了诸多改进优化，使得协议更加清晰、规范，同时也实现了二层网络拓扑的快速收敛。那STP协议具体存在哪些不足呢？RSTP协议是如何在STP协议的基础上进行优化的呢？



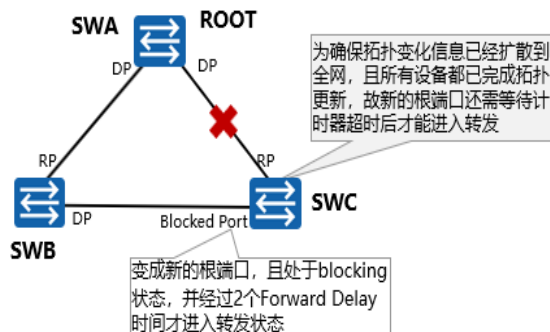
问题一：设备运行STP初始化场景

- STP从初始状态到完全收敛至少需经过30s：



问题二：交换机有BP端口，RP端口down掉场景

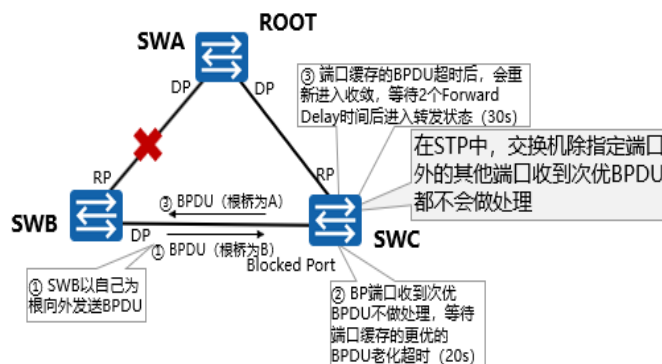
- SWC与SWA的直连链路down掉，其BP端口切换到RP端口并进入转发状态至少需要经过30s：





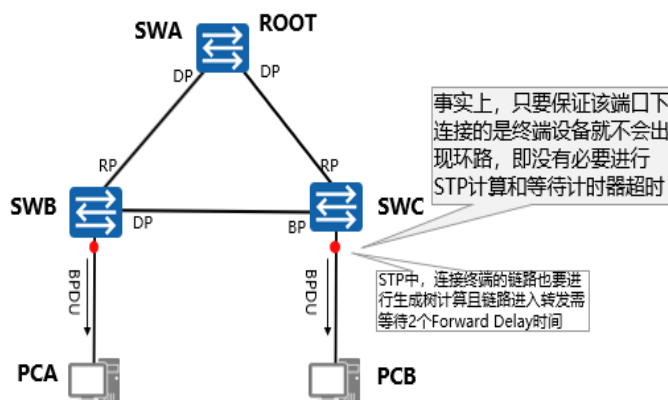
问题三：交换机无BP端口，RP端口down掉场景

- SWB与SWA的直连链路down掉，则SWC的BP端口切换成DP端口并进入转发状态大约需要50s：



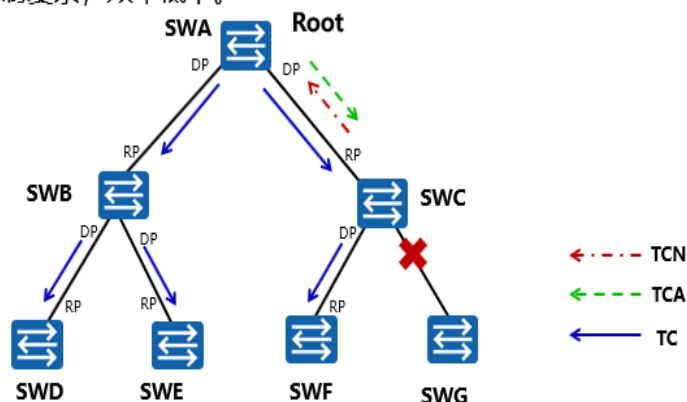
问题四：运行STP的交换机连接用户终端的场景

- 交换机连接终端的链路进入转发需要经过30s：



问题五：STP的拓扑变更机制

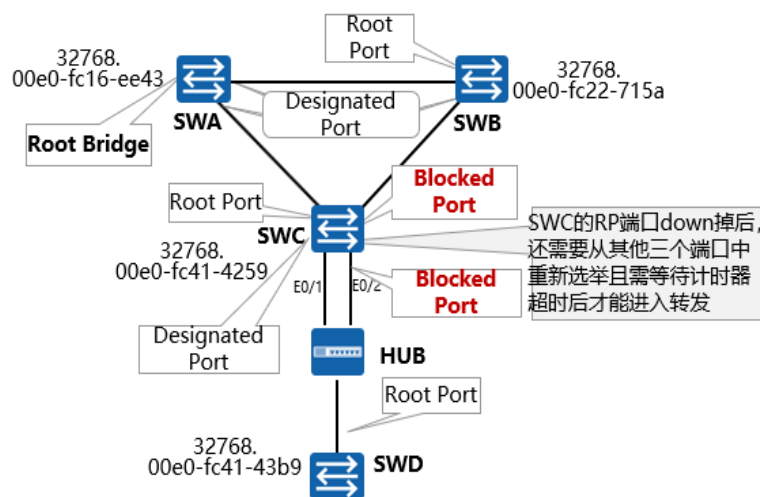
- 先由变更点朝根桥方向发送TCN消息，收到该消息的上游交换机就会回复TCA消息进行确认；最后TCN消息到达根桥后，再由根桥发送TC消息通知设备删除桥MAC地址表项，机制复杂，效率低下。



- 拓扑变更处理过程：
- 在网络拓扑发生变化后，下游设备会不间断地向上游设备发送 TCN BPDU 报文。
- 上游设备收到下游设备发来的 TCN BPDU 报文后，只有指定端口处理 TCN BPDU 报文。其它端口也有可能收到 TCN BPDU 报文，但不会处理。
- 上游设备会把配置 BPDU 报文中的 Flags 的 TCA 位设置 1，然后发送给下游设备，告知下游设备停止发送 TCN BPDU 报文。
- 上游设备复制一份 TCN BPDU 报文，向根桥方向发送。
- 重复上述步骤，直到根桥收到 TCN BPDU 报文。
- 根桥把配置 BPDU 报文中的 Flags 的 TC 位置 1 后发送，通知下游设备直接删除桥 MAC 地址表项。



STP的其他不足之处 - 端口角色



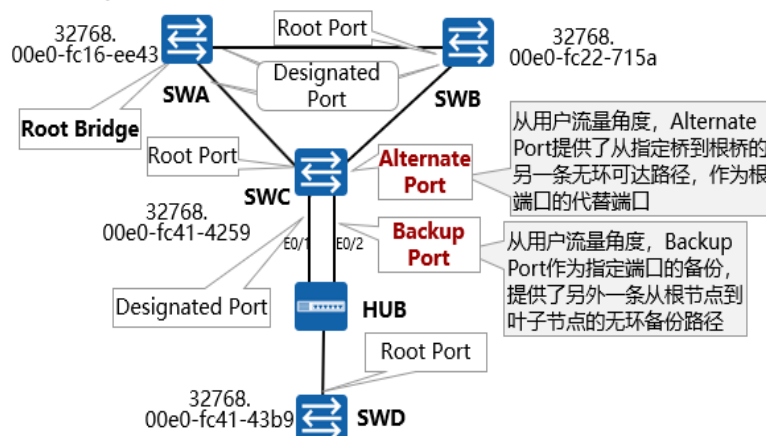
STP的其他不足之处 - 端口状态

STP端口状态	端口状态对应的行为
Disabled	不转发用户流量也不学习MAC地址
Blocking	
Listening	
Learning	不转发用户流量但是学习MAC地址
Forwarding	既转发用户流量又学习MAC地址

三种端口状态从用户使用的角度对应的行为都相同, 但呈现出不同的状态, 这样反而增加了使用难度

端口角色的重新划分

- RSTP定义了两新的端口角色：备份端口（Backup Port）和预备端口（Alternate Port）。



- 根据 STP 的不足，RSTP 新增加了两种端口角色，并且把端口属性充分地按照状态和角色解耦，使得可以更加精确地描述端口，从而使得协议状态更加简便，同时也加快了拓扑收敛。通过端口角色的增补，简化了生成树协议的理解及部署。
- 从配置 BPDU 报文发送角度来看：
- Alternate Port 就是由于学习到其它网桥发送的配置 BPDU 报文而阻塞的端口。
- Backup Port 就是由于学习到自己发送的配置 BPDU 报文而阻塞的端口。
- 从用户流量角度来看：
- Alternate Port 提供了从指定桥到根的另一条可切换路径，作为根端口的备份端口。
- Backup Port 作为指定端口的备份，提供了另外一条从根节点到叶节点的备份通路。
- 给一个 RSTP 域内所有端口分配角色的过程就是整个拓扑收敛的过程。

端口状态的重新划分

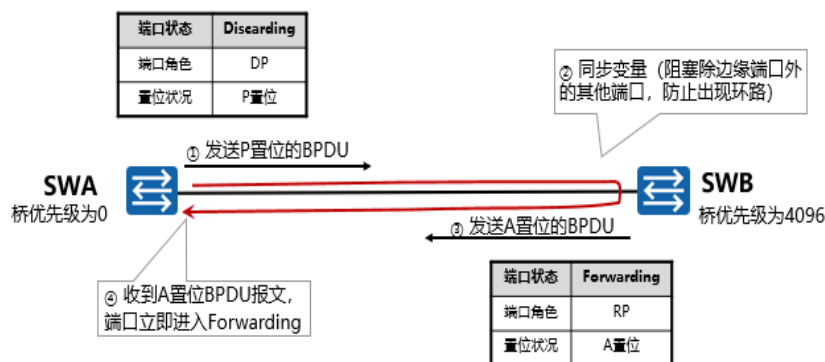
- RSTP的状态规范把原来的5种状态缩减为3种：

STP端口状态	RSTP端口状态	端口状态对应的行为
Disabled	Discarding	如果不转发用户流量也不学习MAC地址，那么端口状态就是Discarding状态。
Blocking		
Listening		
Learning	Learning	如果不转发用户流量但是学习MAC地址，那么端口状态就是Learning状态。
Forwarding	Forwarding	如果既转发用户流量又学习MAC地址，那么端口状态就是Forwarding状态。

- 从用户角度来讲，Listening、Learning 和 Blocking 状态并没有区别，都同样不转发用户流量。

针对问题一：P/A机制 (1)

- P/A机制基本原理

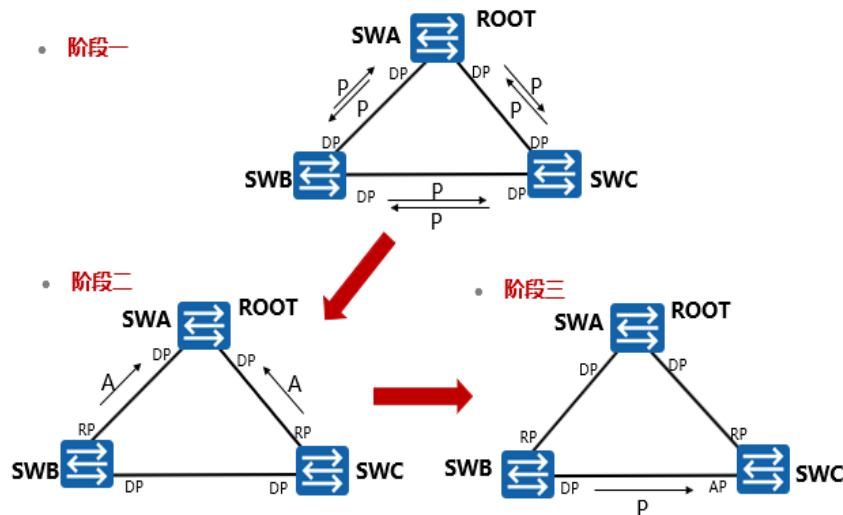


- 特点：由于有来回确认机制和同步变量机制，就无需依靠计时器来保障无环。
- Proposal/Agreement 机制，其目的是使一个指定端口尽快进入 Forwarding 状态。
- P/A 机制要求两台交换设备之间链路必须是点对点的全双工模式。一旦 P/A 协商不成功，指定端口的选择就需要等

待两个 Forward Delay，协商过程与 STP 一样。

- 事实上对于 STP，指定端口的选择可以很快完成，主要的速度瓶颈在于：为了避免环路，必须等待足够长的时间，使全网的端口状态全部确定，也就是说必须要等待至少两个 Forward Delay，所有端口才能进行转发。

针对问题一：P/A机制 (2)



- 问题一的解决方案：
- 阶段一：设备刚刚启动，RSTP 协议刚刚启用，所有交换机都认为自己是最优的根桥，向其他交换机发送 P 置位的 BPDU，并把发送 P 消息的端口变成 DP 口，同时接口处在 Discarding 状态。
- 阶段二：交换机 SWA 收到 SWB 和 SWC 的 P 消息会置之不理，因为他的桥优先级最高。交换机 SWB 和 SWC 收到 SWA 的 P 消息后，由于认同 SWA 是最优的根桥，会根据 P/A 协商流程回复 A 消息，并把发送端口变成 RP 端口，同时接口处在 Forwarding 状态。
- 阶段三：SWA 与 SWB，SWA 与 SWC 的 P/A 协商已经完成，接下来是 SWB 和 SWC 的 P/A 协商。
- SWB 和 SWC 都会发送根桥为 SWA 的 P 消息给对方。

- SWC 收到 SWB 的 P 消息后，发现 P 消息里虽然根桥和自己认可的一样，但是发送者的桥优先级比自己高(SWB>SWC)，所有马上停止发送 P 消息，但是由于已经有端口是 RP 口，也不会回 A 消息。
- SWB 收到 SWC 的 P 消息后，发现 P 消息里虽然根桥和自己认可的一样，但是发送者的桥优先级比自己低(SWB>SWC)，会不停的发送 P 消息。
- 以上状态在等待 2 个 Forward Delay 时间后，SWB 端口为 DP 端口，处在 Forwarding 状态，SWC 端口为 AP 端口，处在 Discarding 状态。
- 实际上 SWB 与 SWC 之间的协商等同于退回到 STP 的模式，但是反正是 Discarding 状态，根本不影响其他业务转发。

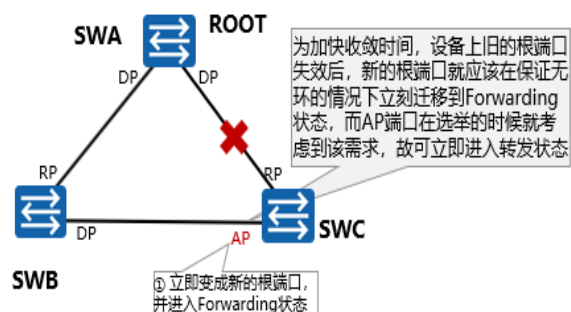
针对问题一：P/A机制 (3)

- RSTP选举原理和STP本质上相同：选举根交换机-选举非根交换机上的根端口-选举指定端口-选举预备端口和备份端口。
- 但是RSTP在选举的过程中加入了“发起请求-回复同意”（P/A机制）这种确认机制，由于每个步骤有确认就不需要依赖计时器来保证网络拓扑无环才去转发，只需要考虑BPDU发送报文并计算无环拓扑的时间（一般都是秒级）。



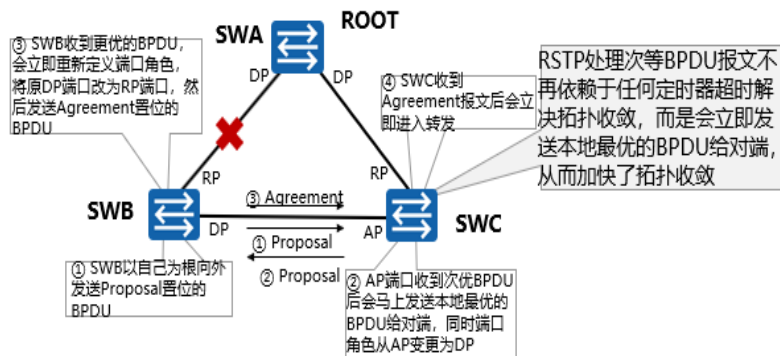
针对问题二：根端口快速切换机制

- SWC与SWA的直连链路down掉，其AP端口切换成RP端口并进入转发状态可在秒级时间内完成收敛：



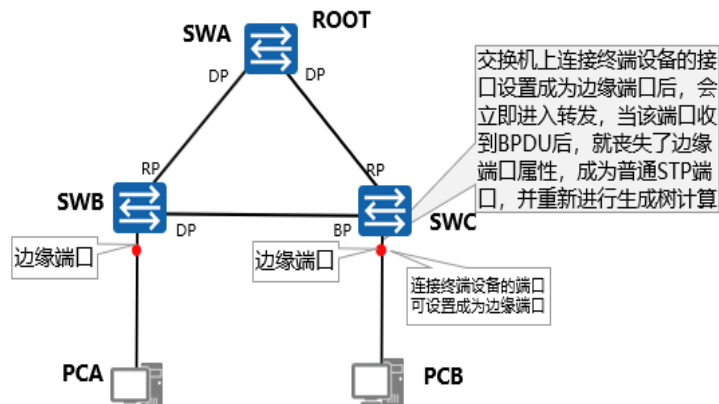
针对问题三：次等BPDU处理机制

- SWB与SWA的直连链路down掉，SWC的AP端口切换成DP端口并进入转发状态可在秒级时间内完成：



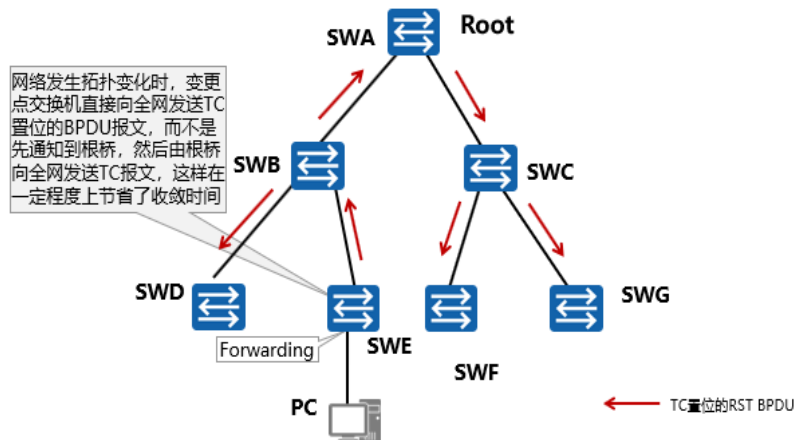
针对问题四：边缘端口的引入

- 在RSTP中，交换机连接终端的链路可立即进入转发状态：



针对问题五：拓扑变更机制的优化

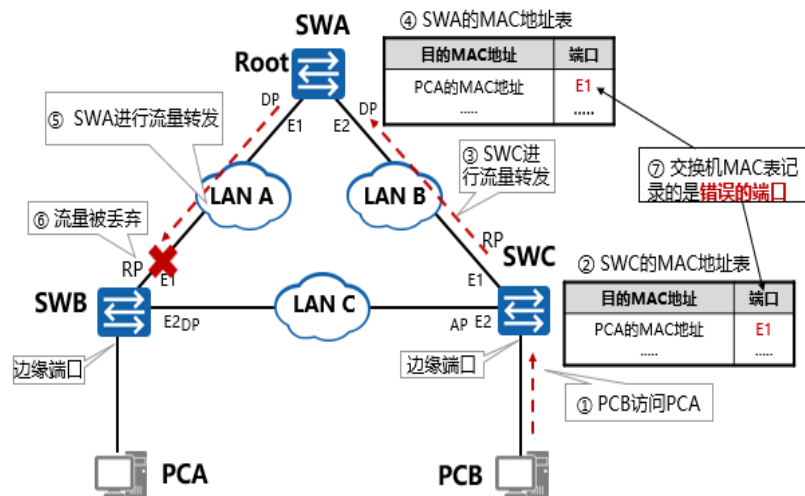
- 判断拓扑变化唯一标准：一个非边缘端口迁移到Forwarding状态。



- 一旦检测到拓扑发生变化，将进行如下处理：
- 为本交换设备的所有非边缘指定端口启动一个 TC While Timer，该计时器值是 Hello Time 的两倍。在这个时间内，清空状态发生变化的端口上学习到的 MAC 地址。同时，由这些端口向外发送 RST BPDU，其中 TC 置位。一旦 TC While Timer 超时，则停止发送 RST BPDU。

- 其他交换设备接收到 RST BPDU 后，清空所有端口学习到 MAC 地址，除了收到 RST BPDU 的端口。然后也为自己所有的非边缘指定端口和根端口启动 TC While Timer，重复上述过程。如此，网络中就会产生 RST BPDU 的泛洪。

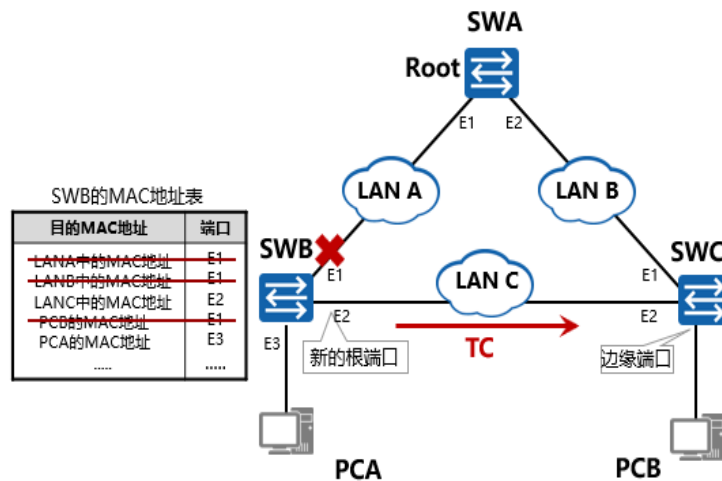
拓扑变化引发的问题



- 在 RSTP 中检测拓扑是否发生变化只有一个标准：一个非边缘端口迁移到 Forwarding 状态。
- 网络拓扑改变可能会导致交换机的 MAC 地址表产生错误。
- 如图所示，在稳定情况下，SWC 的 MAC 地址表中对应 PCA 的 MAC 地址的端口是 E1。如果 SWB 的 E1 端口发生了故障，而 SWC 的 MAC 地址表中与 PCA 的 MAC 地址对应的端口仍然是 E1，则会导致数据转发丢失的问题。



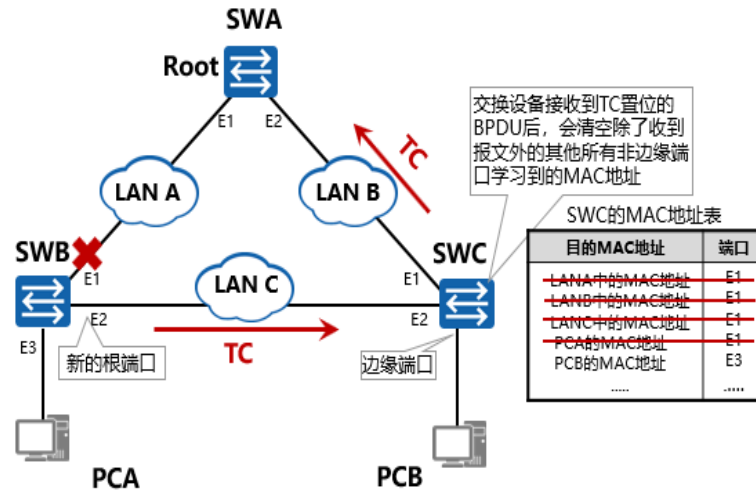
拓扑变化处理 (1)



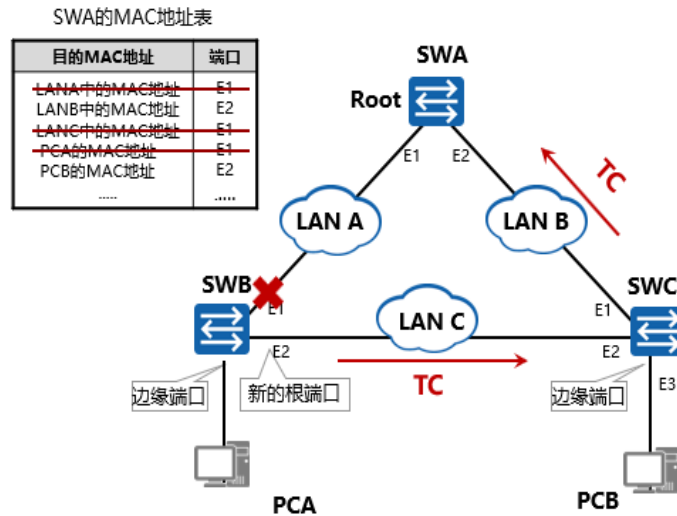
- 一旦检测到拓扑发生变化，首先将进行如下处理：
- 清空状态发生变化的端口上学习到的 MAC 地址。
- 同时，由这些端口向外发送 RST BPDU，其中 TC 置位。
- 一旦 TC While Timer 超时，则停止发送 RST BPDU。
- 如图所示，SWB 的 E1 端口出现故障之后，RSTP 的处理过程如下：
- SWB 重新计算生成树，选举 E2 为新的根端口。
- SWB 删除 MAC 地址表中 E1 端口所对应的表项。
- 生成树重新计算完成之后（需要进入转发状态的端口已经进入了转发状态），SWB 的所有非边缘端口向外发送 TC 置位的 RST BPDU。



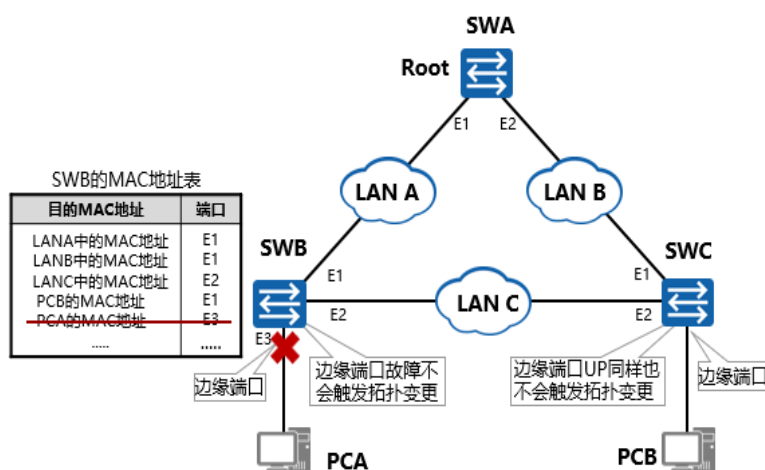
拓扑变化处理 (2)



拓扑变化处理 (3)

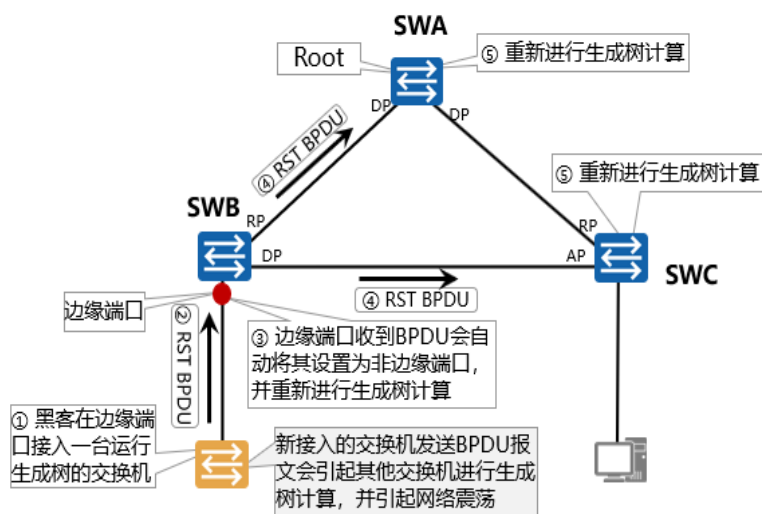


拓扑变化处理 (4)



- 边缘端口 down 掉不会触发拓扑变更，而且故障恢复后，同样也不会触发拓扑变更。

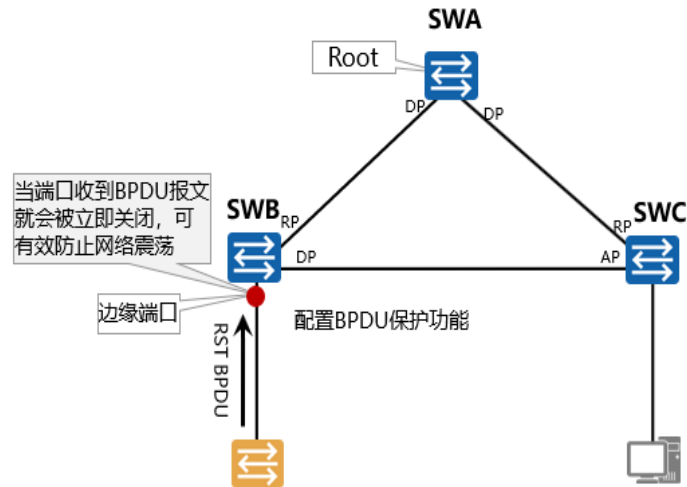
BPDU保护 (1)



- BPDU 保护
- 应用场景：防止有人伪造 RST BPDU 恶意攻击交换设备，当边缘端口接收到该报文时，会自动设置为非边缘端口，并重新进行生成树计算，引起网络震荡。

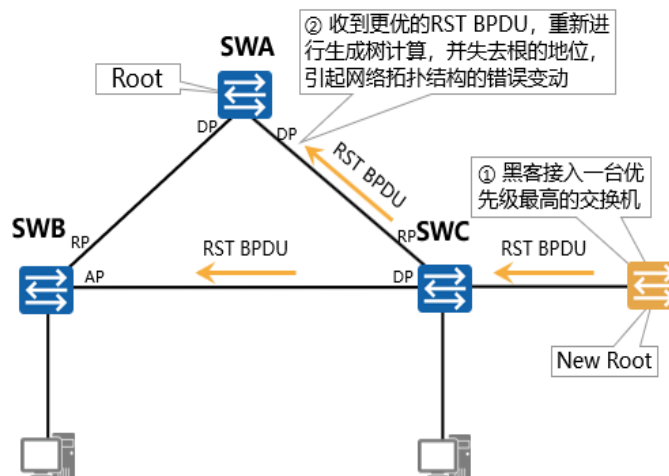
- 实现原理：配置 BPDU 保护功能后，如果边缘端口收到 BPDU 报文，边缘端口将会被立即关闭。

BPDU保护 (2)



- BPDU 保护
- 实现原理：配置 BPDU 保护功能后，如果边缘端口收到 BPDU 报文，边缘端口将会被立即关闭。

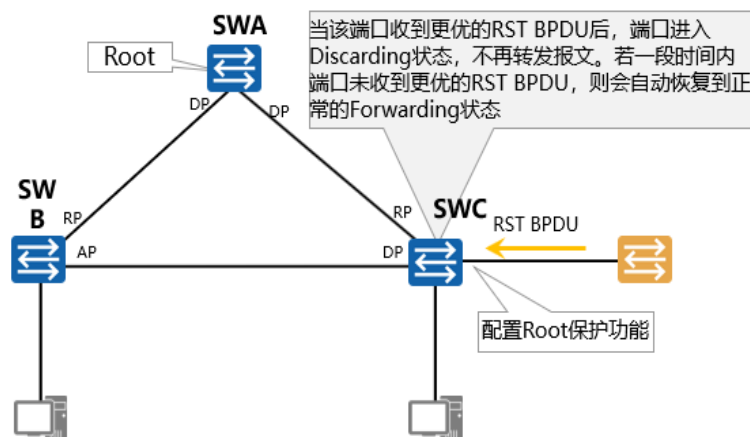
根保护 (1)



- 根保护

- 应用场景：由于维护人员的错误配置或网络中的恶意攻击，网络中合法根桥有可能会收到优先级更高的 RST BPDU，使得合法根桥失去根地位，从而引起网络拓扑结构的错误变动。
- 实现原理：一旦启用 Root 保护功能的指定端口收到优先级更高的 RST BPDU 时，端口状态将进入 Discarding 状态，不再转发报文。在经过一段时间，如果端口一直没有再收到优先级较高的 RST BPDU，端口会自动恢复到正常的 Forwarding 状态。
- Root 保护功能只能在指定端口上配置生效。

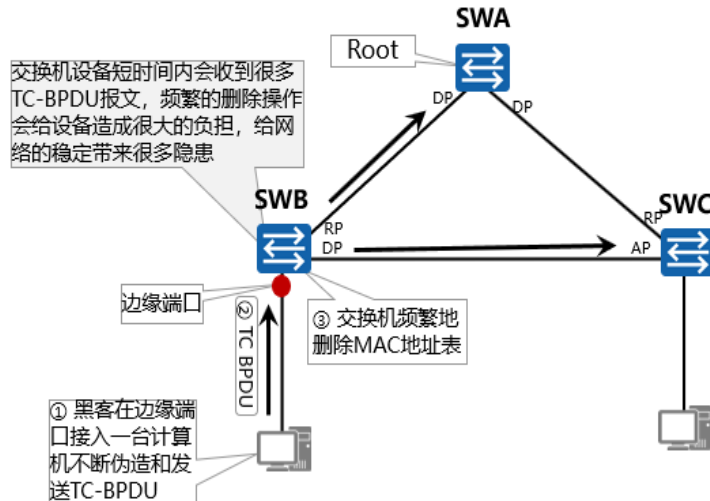
根保护 (2)



- 根保护
- 应用场景：由于维护人员的错误配置或网络中的恶意攻击，网络中合法根桥有可能会收到优先级更高的 RST BPDU，使得合法根桥失去根地位，从而引起网络拓扑结构的错误变动。
- 实现原理：一旦启用 Root 保护功能的指定端口收到优先级更高的 RST BPDU 时，端口状态将进入 Discarding 状态，不再转发报文。在经过一段时间，如果端口一直没有再收到优先级较高的 RST BPDU，端口会自动恢复到正常的 Forwarding 状态。

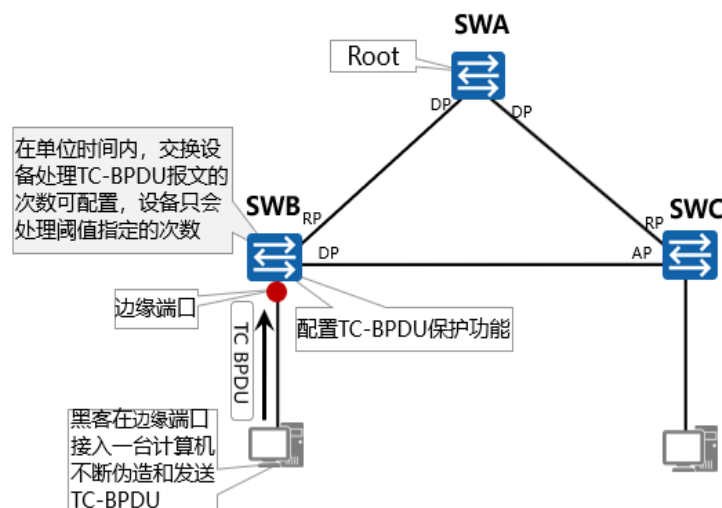
- Root 保护功能只能在指定端口上配置生效。

TC-BPDU泛洪保护 (1)



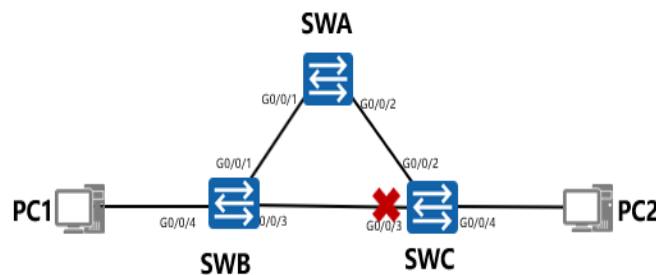
- TC-BPDU 攻击：
- 交换机在接收到 TC-BPDU 报文后，会执行 MAC 地址表的删除操作。如果有人伪造 TC-BPDU 报文恶意攻击交换机时，交换机短时间内会收到很多 TC-BPDU 报文，频繁的删除操作会给设备造成很大的负担，给网络的稳定带来很大隐患。

TC-BPDU泛洪保护 (2)



- TC-BPDU 攻击保护：
- 启用防 TC-BPDU 报文攻击功能后，在单位时间内，RSTP 进程处理 TC 类型 BPDU 报文的次数可配置（缺省的单位时间是 2 秒，缺省的处理次数是 3 次）。如果在单位时间内，RSTP 进程在收到 TC 类型 BPDU 报文数量大于配置的阈值，那么 RSTP 进程只会处理阈值指定的次数；对于其他超出阈值的 TC 类型 BPDU 报文，定时器到期后，RSTP 进程只对其统一处理一次。这样可以避免频繁的删除 MAC 地址表项，从而达到保护交换机的目的。

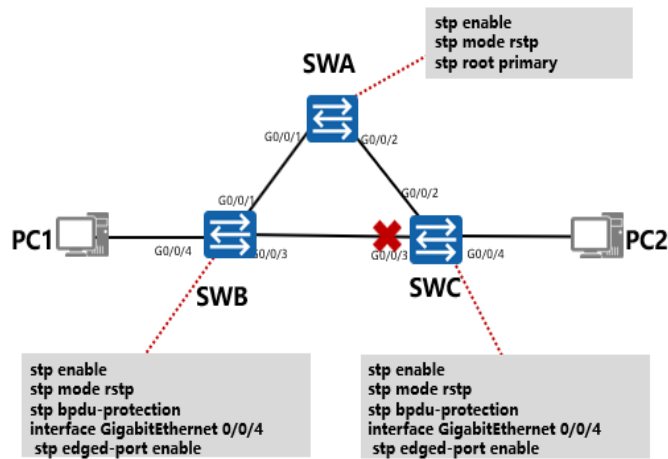
RSTP配置需求



- 如图所示，SWA、SWB和SWC组成了一个环形的交换网络，为了消除环路对网络的影响，故使交换机都运行RSTP，最终将环形网络结构修剪成无环路的树形网络结构。



RSTP配置实现



- 配置实现：
- stp enable //全局开启 STP
- stp mode rstp //配置 STP 模式为 RSTP
- stp root primary //配置 SWA 为根桥
- stp bpdu-protection //全局开启 BPDU 防护，配合边缘端口一起使用
- stp edged-port enable //配置端口为边缘端口

RSTP配置验证 (1)

- 在SWA上查看生成树信息：

```
<SWA>display stp brief
MSTID Port                Role STP State    Protection
0      GigabitEthernet0/0/1 DESI FORWARDING  NONE
0      GigabitEthernet0/0/2 DESI FORWARDING  NONE
```

```
<SWA>display stp
-----[CIST Global Info][Mode RSTP]-----
CIST Bridge      :0 .4c1f-cc5f-55e4
Config Times     :Hello 2s MaxAge 20s FwDly 15s MaxHop 20
Active Times     :Hello 2s MaxAge 20s FwDly 15s MaxHop 20
CIST Root/ERPC   :0 .4c1f-cc5f-55e4 / 0
CIST RegRoot/IRPC :0 .4c1f-cc5f-55e4 / 0
CIST RootPortId  :0.0
BPDU-Protection  :Disabled
CIST Root Type   :Primary root
```

RSTP配置验证 (2)

- 在SWB上查看生成树信息：

```
[SWB]display stp brief
MSTID Port                Role STP State    Protection
0      GigabitEthernet0/0/1 ROOT FORWARDING  NONE
0      GigabitEthernet0/0/3 DESI FORWARDING  NONE
0      GigabitEthernet0/0/4 DESI FORWARDING  BPDU
```

- 在SWC上查看生成树信息：

```
<SWC>display stp brief
MSTID Port                Role STP State    Protection
0      GigabitEthernet0/0/2 ROOT FORWARDING  NONE
0      GigabitEthernet0/0/3 ALTE DISCARDING  NONE
0      GigabitEthernet0/0/4 DESI FORWARDING  BPDU
```



思考题

1. RSTP定义了几种端口状态? ()

A. 2

B. 3

C. 4

2. RSTP定义了哪些端口角色? ()

- 答案：B。
- 答案：根端口、指定端口、备份端口、预备端口、边缘端口。