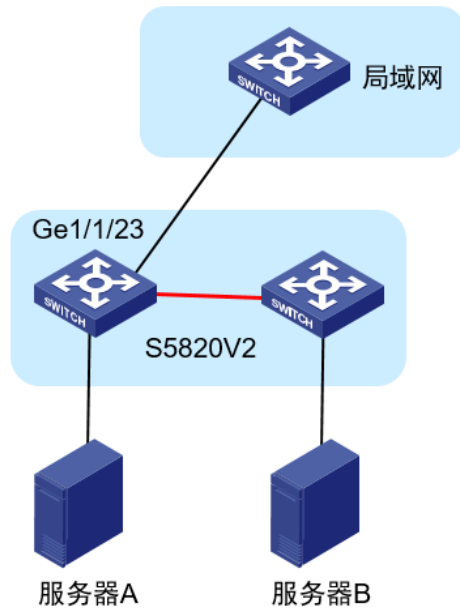


## 知 S5820V2 STP计算导致二层网络流量中断经验案例

STP 林洪宇 2018-03-19 发表



某局点组网拓扑如上所示：交换机S5820V2堆叠作为接入设备，下联第三方服务器，上联汇聚接入网络。

故障现象：

客户反馈现场故障当晚，由于上行其他设备发生故障产生大量TC报文，交换机S5820V2收到上联设备的发过来的大量TC报文，导致STP拓扑震荡，根桥不断发生变化，堆叠S5820V2上联端口的端口角色在指定端口和根端口之间变化，下联服务器业务流量出现中断十几秒的现象。

收集诊断信息，分析查看故障当时设备上的日志：

```
%Jan 26 02:16:11:354 2012 S5820V2 STP/6/STP_NOTIFIED_TC: -Slot=1; Instance 0& # 39;s port Ten-GigabitEthernet1/1/23 was notified a topology change.
```

```
%Jan 26 02:16:13:350 2012 S5820V2 STP/6/STP_NOTIFIED_TC: -Slot=1; Instance 0& # 39;s port Ten-GigabitEthernet1/1/23 was notified a topology change.
```

```
Port Ten-GigabitEthernet1/1/23
```

```
Role change      : DESI->ROOT
```

```
Time             : 2012/01/26 02:16:09
```

```
Port priority    : 0.7054-f596-f530 41000 8192.8478-ac18-47c1 0
```

```
8192.8478-ac18-47c1 128.174 128.23
```

```
Designated priority : 0.7054-f596-f530 41000 8192.8478-ac18-47c1 2
```

```
32768.487a-da7c-c34f 128.23 128.23
```

```
Port Ten-GigabitEthernet1/1/23
```

```
Role change      : ROOT->DESI
```

```
Time             : 2012/01/26 02:16:08
```

```
Port priority    : 8192.8478-ac18-47c1 0 8192.8478-ac18-47c1 0
```

```
8192.8478-ac18-47c1 128.174 128.23
```

```
Designated priority : 4096.188b-45d9-6c41 0 4096.188b-45d9-6c41 2
```

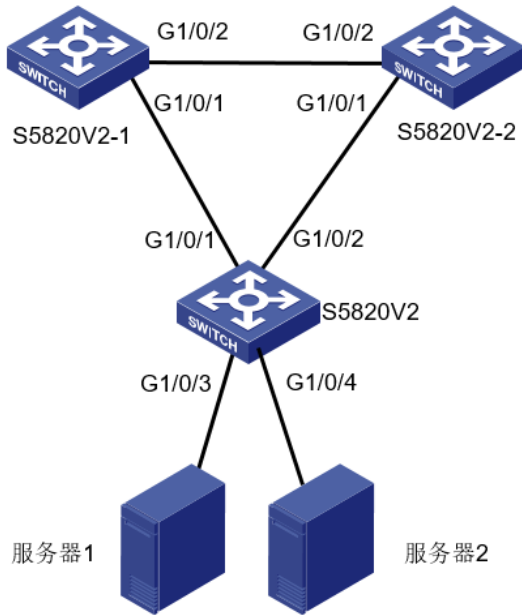
```
32768.487a-da7c-c34f 128.23 128.23
```

故障时间点设备日志上打印上联万兆端口1/1/23持续不断的收到TC报文，但是并没有端口UP/DOWN情况发生，进一步查看设备上display stp history生成树端口变化历史记录，上联端口的生成树角色一直在指定端口和根端口之间变化，生成树STP一直在震荡。

1、根据端口角色的变化信息分析，仅仅是上联万兆端口1/1/23端口的角色一直在发生转变，但是下联服务器之间的二层业务不断出现短暂中断的情况，从日志中看，下联服务的端口并没有端口角色变化的日志记录，理论上端口角色没有改变，端口一直处于转发状态，下联服务器二层互访，业务理应不会受到影响。

2、查看服务器上业务中断的情况，业务中断的时间大概15秒左右，根据业务中断时间与设备相关日志对比分析，中断时间与STP中的转发延迟（Forward Delay）时间相近，转发延迟协议默认值是15秒，当拓扑发生变化，新的配置消息要经过一定的时延才能传播到整个网络。

3、通过时间上对比分析得出业务中断时间15秒原因可能与连接服务器端口状态切换相关，但设备上没有打印下联生成树端口角色并没有切换，于是本地模拟故障环境组网进行测试分析：



4、本地测试：三台S5820V2相连组成环形组网，使能MSTP生成树协议选择性地阻塞网络中的冗余链路来消除二层环路，通过配置调整路径开销和STP优先级使接入服务器的S5820V2上行端口G1/0/1和G1/0/2端口分别为根端口和指定端口，复现了服务器1和2之间PING包会存在中断的现象，但是奇怪的是，并不是每次G1/0/1和G1/0/2根端口和指定端口切换都会导致PING丢包的现象，进而整理测试过程，进一步分析：

如上图所示：S5820V2-1为生成树的根桥，S5820V2的G1/0/1和G1/0/2端口分别为根端口和指定端口，分别处于FORWARDING的转发状态。并且G1/0/3和G1/0/4端口下的服务器进行长PING测试，测试是否会产生业务流量中断的现象。

第一次切换：

通过调整交换机的STP优先级，调整S5820V2的优先级，将其调整为根桥：

```
[H3C]#stp brief
MST ID      Port      Role STP State Protection
0           GigabitEthernet1/0/1    DESI FORWARDING NONE
0           GigabitEthernet1/0/2    DESI FORWARDING NONE

[H3C]#dis stp priority 400
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.

[H3C]#dis stp priority 0
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 detected a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.

[H3C]#dis stp brief
MST ID      Port      Role STP State Protection
0           GigabitEthernet1/0/1    ROOT FORWARDING NONE
0           GigabitEthernet1/0/2    DESI FORWARDING NONE
0           GigabitEthernet1/0/3    ALTE DISCARDING NONE
```

如上图所示：通过调整根桥制造出S5820V2上联端口根端口和指定端口切换的现象，可以看到上联端口1/0/1和1/0/2收到TC报文，与现场故障时的现象一样；但是PING包没有出现中断现象，没有复现客户现场情况；

第二次切换：

继续调整STP优先级，将S5820V2-1调整为根桥：

```
[H3C]#stp priority 800
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.

[H3C]#dis stp brief
MST ID      Port      Role STP State Protection
0           GigabitEthernet1/0/1    ROOT FORWARDING NONE
0           GigabitEthernet1/0/2    DESI FORWARDING NONE
0           GigabitEthernet1/0/3    ALTE DISCARDING NONE

[H3C]#dis stp priority 800
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.

[H3C]#dis stp brief
MST ID      Port      Role STP State Protection
0           GigabitEthernet1/0/1    ROOT FORWARDING NONE
0           GigabitEthernet1/0/2    DESI FORWARDING NONE
0           GigabitEthernet1/0/3    ALTE DISCARDING NONE
```

通过第二次切换交换机上并没有日志记录端口状态发生切换的过程，若不是复现分析，难以找到影响流量中断的原因，经过查看资料，发现交换机上默认是不开启STP端口状态变化信息显示的,需要通过命令stp port-log all开启，将此命令开启后，重新复现故障现象：

```
[H3C]#stp port-log all
[H3C]#dis stp bri
MST ID      Port      Role STP State Protection
0           GigabitEthernet1/0/1    DESI FORWARDING NONE
0           GigabitEthernet1/0/2    ROOT FORWARDING NONE
0           GigabitEthernet1/0/3    DESI FORWARDING NONE

[H3C]#Jul 14 17:05:29:830 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.
%Jul 14 17:05:29:832 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.
%Jul 14 17:05:29:832 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 has been set to discarding state.
%Jul 14 17:05:29:832 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/3 has been set to discarding state.
%Jul 14 17:05:29:877 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%Jul 14 17:05:29:877 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%Jul 14 17:05:29:877 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 has been set to forwarding state.
%Jul 14 17:05:31:859 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/2 was notified a topology change.
%Jul 14 17:05:32:344 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/1 was notified a topology change.

[H3C]#dis stp bri
MST ID      Port      Role STP State Protection
0           GigabitEthernet1/0/1    ROOT FORWARDING NONE
0           GigabitEthernet1/0/2    DESI FORWARDING NONE
0           GigabitEthernet1/0/3    DESI FORWARDING NONE

[H3C]#Jul 14 17:06:00:160 2017 H3C STP/6/STP_NOTIFICATION: Instance 0's port GigabitEthernet1/0/3 has been set to forwarding state.
[H3C]#dis stp bri
MST ID      Port      Role STP State Protection
0           GigabitEthernet1/0/1    ROOT FORWARDING NONE
0           GigabitEthernet1/0/2    DESI FORWARDING NONE
0           GigabitEthernet1/0/3    DESI FORWARDING NONE
```

如上图所示，这时发现1/0/3端口确实经历了DISCARDING状态到FORWARDING状态的STP转发延迟时间。这段时间内，端口处于阻塞状态，导致了二层流量中断。从日志中不难看出1/0/2端口也同样经历了DISCARDING状态到FORWARDING状态的变化，但是却没有经历15S的收敛时间，原因如下：RSTP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。其“快速”体现在，当一个端口被选为根端口和指定端口后，其进入转发状态的延时将大大缩短，从而缩短了网络最终达到拓扑稳定所需要的时间。在RSTP中，根端口的端口状态快速迁移的条件是：本设备上旧的根端口已经停止转发数据，而且上游指定端口已经开始转发数据。指定端口的端口状态快速迁移的条件是：指定端口是边缘端口（即该端口直接与用户终端相连，而没有连接到其他设备或共享网段上）或者指定端口与点对点链路（即两台设备直接相连的链路）相连。如果指定端口是边缘端口，则指定端口可以直接进入转发状态；如果指定端口连接着点对点链路，则设备可以通过与下游设备握手

, 得到响应后即刻进入转发状态。

5、接下来需要分析的是什么原因导致两次切换会产生两种不同的结果：

第一次切换结果

1.根桥设备由优先级4096切换为优先级为0的设备；

2.对应端口的端口指定优先级由4096切换为0, 比端口优先级更优, 因此agreed不会清除, 不会重新计算, 所以没有产生流量中断的情况。

第二次切换结果

1.根桥设备由优先级0切换为优先级为4096的设备；

2.对应端口的端口指定优先级由0切换为4096, 没有端口优先级优, 因此agreed清除, synced清除, 最端口状态会重新计算。

从协议的角度讲, 第一次切换后端口的优先级相比原来更优, 这种情况下认为没必要重新计算端口状态, 也是在协议的层面为了减少这种场景的流量中断而做的相应优化。

为了避免这种端口状态变化导致流量中断的情况发生, 端口状态实现快速迁移, 连接服务器的端口必须配置成为边缘端口。在现网设备开启STP的组网中, 一定要做好相关STP优化, 配置边缘端口, 配置根保护等, 防止由于一台设备的故障影响全网业务流量。