

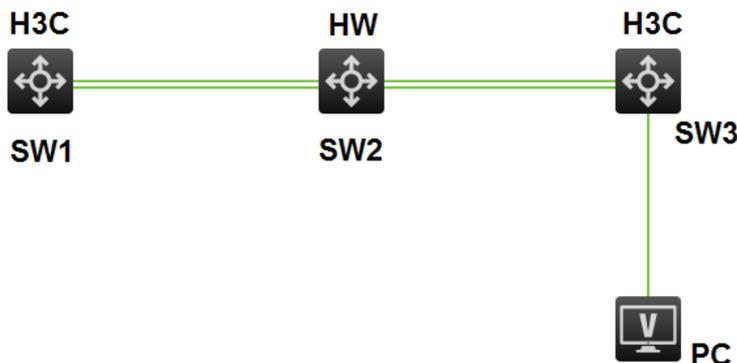
知 某局点链路聚合一条线断开导致stp重新收敛问题

PVST 二层链路聚合 STP 孙海亮 2020-01-06 发表

组网及说明

拓扑如下:

SW1和SW2、SW2和SW3分别用两个万兆口做了链路聚合, 并且都开启了STP, SW1 SW3为H3C设备, SW2为华为设备



问题描述

正常情况下业务运行没有问题, PC去ping SW1上的网关正常。

但是发现当SW1和SW2之间的聚合链路断掉一根线后, 导致SW3下面的PC业务中断30s, 然后自动恢复。

后来发现SW1、SW2、SW3三台设备都是H3C或者都是HW的, SW1和SW2之前聚合断一条线后就没有问题, 这三台无论那一台换成另外一个厂商的就会有该问题

过程分析

在SW3设备的上行聚合口和成员口开启stp no-agreement-check和 stp config-digest-snooping现象依旧

在SW3的上行口抓包发现第一个STP报文(看源mac为华为的mac, 是华为发给H3C设备的), 而华为设备发一个非标准协议的报文

```
Spanning Tree Protocol
  Protocol Identifier: Spanning Tree Protocol (0x0000)
  Protocol Version Identifier: Multiple Spanning Tree (3)
  BPDU Type: Rapid/Multiple Spanning Tree (0x02)
  B PDU flags: 0x7e, Agreement, Forwarding, Learning, Port Role: Designated, Proposal
    0... .. = Topology Change Acknowledgment: No
    .1.. .. = Agreement: Yes
    ..1. .... = Forwarding: Yes
    ...1 .... = Learning: Yes
    .... 11.. = Port Role: Designated (3)
    ..... 1. = Proposal: Yes
    .... ..0 = Topology Change: No
  Root Identifier: 8192 / 0 / dc:da:80:44:b5:80
  Root Path Cost: 2000
  Bridge Identifier: 16384 / 0 / f8:75:88:07:9e:81
  Port identifier: 0x8022
  Message Age: 1
  Max Age: 20
  Hello Time: 2
  Forward Delay: 15
  Version 1 Length: 0
  Version 3 Length: 64
  MST Extension
```

在H3C设备SW3上debug查看到

1. SW1设备聚合成员口down掉1条, cost变大, SW2设备发出的BPDU报文 cost变大, 报文如下:

SW3设备上debug显示收到的BPDU报文

```
*Jan 1 00:56:32:131 2013 test2_H3C5560 STP/7/PKT:
```

```
Port633(Bridge-Aggregation1) received MSTP-dot1s Packet(Length: 102)
```

ProtocolVersionID: 03
BPDUType : 02
CIST Root ID : 8192.dccda-8044-b580
External RPC : 2000
Reg Root ID : 16384.f875-8807-9e81
Internal RPC : 0
CIST Bridge ID : 16384.f875-8807-9e81
CIST Port ID : 128.34
(Instance)Flags : (00)Desig[AFLP]

2. SW3设备根端口收到P置位的BPDU报文，因为P/A机制，会启动同步过程阻塞非边缘端口的指定端口，从而防止环路，然后将根端口bagg1设置为转发状态，并向SW2回复A置位BPDU。。

```
*Jan 1 00:56:32:138 2013 test2_H3C5560 STP/7/PKT:  
Port633(Bridge-Aggregation1) sent MSTP-dot1s Packet(Length: 102)  
ProtocolVersionID: 03  
BPDUType : 02  
CIST Root ID : 8192.dccda-8044-b580  
External RPC : 2001  
Reg Root ID : 32768.dccda-8044-ba60  
Internal RPC : 0  
CIST Bridge ID : 32768.dccda-8044-ba60  
CIST Port ID : 128.633  
(Instance)Flags : (00)Root[ AFL ]
```

3. 接着SW3会从处于同步状态的非边缘指定端口 1/0/1发送P置位BPDU，因为没有收到A置位BPDU（PC不会回应BPDU），则该端口会切换到STP方式，需要等待30S时间才能进入转发状态

从上面信息分析看，根本原因是华为侧发出的BPDU未遵循标准协议，proposing标识应该是非forwarding 非learning位才会置位

解决方法

1. 在SW3设备上开启边缘端口
2. 手工指定SW1和SW2之间聚合端口的cost值