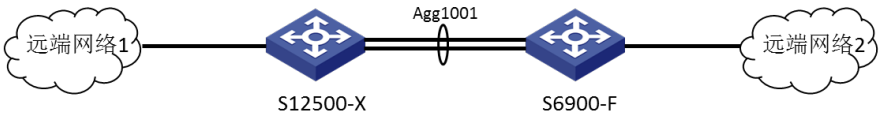
S12510-X接口下关闭stp仍透传stp报文导致对端displute问题案例

STP 张文宁 2018-02-13 发表



如图所示，远端网络1和S12500-X建立的是vxlan隧道，S12500-X是分布式网关VTEP设备，S12500-X的agg1001口是ac口，并且在接口下关闭了stp。S12500-X和S6900-F均全局开启了stp。客户在变更新增业务过程中，反馈左侧远端网络1突然无法ping通远端网络2，进一步查看S12500-X连对端S6900F的直连arp都学不到，直连ping不通，登录S6900-F查看端口agg1001发现端口被displute。现场为了恢复业务，紧急通过在S6800-F接口下也关闭stp后业务恢复。

查看S6900-F故障时间收集的诊断信息，确实发现存在displute情况：

%Jan 16 18:19:58:813 2018 GD-KXC2F-0625(0626)-H3C6900 STP/4/STP_DISPUTE: Instance 0&＃39;s port Bridge-Aggregation1001 received an inferior BPDU from a designated port which is in forwarding or learning state.

==========display stp abnormal-port==========

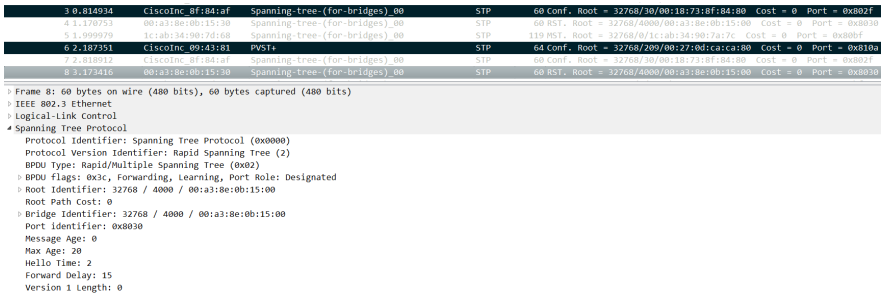
---[Bridge-Aggregation1001]---

MST ID	BlockReason	Time
0	Disputed	18:26:14 01/16/2018
0	Disputed	18:26:12 01/16/2018
0	Disputed	18:26:10 01/16/2018

查看S12500-X的接口配置，确实故障发生前端口前就已经关闭了stp：

```
interface Bridge-Aggregation1001
description to-GD-KXC2F-0625(0626)-H3C6900-bri1001
port link-type trunk
undo port trunk permit vlan 1
port trunk permit vlan 2 to 4094
undo stp enable
service-instance 321
 encapsulation s-vid 321
 xconnect vsi 7790001
service-instance 533
 encapsulation s-vid 533
```

为了坐实S12500-X从agg 1001发送报文给S6900-F，从S6900-F的agg 1001入方向抓包，发现确实有收到BPDU报文：



看源mac，有四个都是远端网络1设备的，在S12500-X上看到都是evpn发布过来的，说明是从tunnel传过来的：

```
VSI name: 755002
MAC address     Link ID/Name    Flags   Next hop
000c-29c6-bbf7  Tunnel17        B       120.80.96.63
1cab-3490-7a7c  0               DL      -
000c-296a-7fb8  Tunnel17        B       120.80.96.63
00a3-8e0b-1530  Tunnel17        B       120.80.96.63
VSI name: 755000
MAC address     Link ID/Name    Flags   Next hop
d461-fe69-c7d9  0               DL      -
1cab-3490-7a7c  0               DL      -
0027-0dca-ca80  Tunnel17        B       120.80.96.63
```

68ef-bd09-4381 Tunnel17   B   120.80.96.63
VSI name: 755001
MAC address   Link ID/Name   Flags   Next hop
d461-fe69-c7d9 0   DL   -
1cab-3490-7a7c 0   DL   -
0001-0a0a-0a02 Tunnel17   B   120.80.96.63
0001-0a0a-0a26 Tunnel17   B   120.80.96.63
0018-738f-84af Tunnel17   B   120.80.96.63
6805-ca19-373b Tunnel17   B   120.80.96.63
Tunnel17
Current state: UP
Line protocol state: UP
Description: Tunnel17 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 4434
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 120.80.172.85, destination 120.80.96.63
Tunnel protocol/transport UDP_VXLAN/IP
Last 300 seconds input rate: 8 bytes/sec, 64 bits/sec, 0 packets/sec
Last 300 seconds output rate: 0 bytes/sec, 0 bits/sec, 0 packets/sec
Input: 68255 packets, 4102968 bytes, 0 drops
Output: 53 packets, 2968 bytes, 0 drops
理论上stp报文是不会通过隧道口去到远端的，为什么会透过来呢？进一步查看上述远端VTEP设备，发现远端设备并没有开启stp！所以远端VTEP设备AC口收到BPDU报文，直接当做普通报文封装传到了S12500-X，同时S12500-X解封装后这个报文并不会再上送cpu，因此不知道是BPDU报文，虽然agg1001关闭了stp，但无法识别所以传给了S6900-F。而反方向S6900-F发给S12500-X的BPDU报文，因为S12500-X开启了stp，因此收到的报文直接丢弃了。所以出现了现场BPDU单通，进程触发displute保护。

网络应该统一，建议远端VTEP也开启stp，这样BPDU报文就不会通过隧道过来了。

组网问题，Displute就是因为一边能收到bpdu报文，另一端收不到触发的一种保护机制。排查此类问题的时候使用顺藤摸瓜思路会很有帮助。