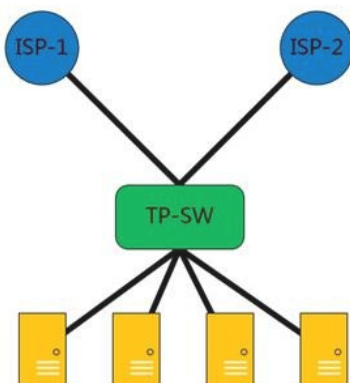


某局点S6800-54QF设备路由更新延迟问题

BGP 翁青山 2020-02-28 发表

组网及说明



我们的设备是图中的TP-SW，黄色图标是服务器，ISP是路由器。

S6800和服务器建立IBGP邻居，S6800和ISP建立EBGP邻居，4台服务器会同时发送或者撤销同一条32位明细路由。

问题描述

4台服务器会同时发送或者撤销同一条32位明细路由，发现S6800设备上立即撤销了该路由，但多台ISP设备撤销路由的时间存在较大时延，大概慢10-30S左右。因此怀疑是S6800存在问题。

目前设备都是S6800的，同样的架构没有异常的设备软件版本都是26xx的，有上述异常的都是2510的。

过程分析

1、当TP-SW运行在R2510P01版本上时，当TP-SW下连Server发布10.10.10.10/32的主机路由及特定地址的32位主机路由时TP-SW会形成对应数量的等价路由，在此以主机路由114.67.92.238/32为例进行分析，故障设备上对应114.67.92.238/32的表项情况如下所示：

路由表：

```
<HD_BX_AZTPSW_2Q-1-A_G-17/H17>dis ip routing-table 114.67.92.238
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
114.67.92.238/32	BGP	200	0	10.10.10.10	RAGG14
				10.10.10.10	RAGG13
				10.10.10.10	RAGG13

Fib表：

114.67.92.238/32	100.127.42.68	UGHR	RAGG14	Null
114.67.92.238/32	100.127.42.64	UGHR	RAGG13	Null
114.67.92.238/32	100.127.42.66	UGHR	RAGG12	Null

现网中R26xx版本的类似应用组网情况下的表项情况：

路由表：

```
43.226.164.79/32 BGP 200 0 10.10.10.10 RAGG4
```

Fib表：

43.226.164.79/32	100.127.202.86	UGHR	RAGG4	Null
43.226.164.79/32	100.127.202.90	UGHR	RAGG5	Null
43.226.164.79/32	100.127.202.94	UGHR	RAGG6	Null
43.226.164.79/32	100.127.202.98	UGHR	RAGG3	Null

R26xx主机路由情况，在设备上单条路由；

```
=====display ip routing-table=====
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
43.226.166.174/32	BGP	200	0	10.10.10.10	RAGG4

2、以上区别是由于S6800-54QF R25xx版本与R26xx版本在路由表等价路由规格处理上的区别导致，R25xx版本中相同下一跳的路由还允许形成等价，而R26xx版本中此情况则不允许形成等价，当时这两个版本中根据路由迭代结果生成对应目的地址的Fib表项的规则未变，在指导报文转发上并不会区别。现网设备进行信息收集时发现TP-SW从Server收到的路由撤销消息并不是完全同步的，由此导致TP-SW再向ISP设备发送BGP路由更新时触发了route-update-interval限制。

3、让一线同事在R25xx设备上开启了debug bgp packet，复现问题，收集相应的debug信息，分析确认：在始发者服务器撤销和发布路由的时候，由于时间上不完全同步，导致先到的路由先撤销，稍微晚到一点点的路由，由于触发了route-update-interval（默认30s）的机制，导致晚到的路由需要延迟30s再同步给其他邻居，所以对于ISP设备，会发现路由撤销有问题。

//从peer 100.127.42.66 (Server 1) 接收到路由撤销消息

```
*Aug 12 15:15:37:666 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Recv UPDATE(Withdraw) from peer 100.127.42.66 for destinations:
    114.67.92.238/32,
```

//从peer 100.127.42.68 (Server 2) 接收到路由撤销消息

```
*Aug 12 15:15:37:667 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Recv UPDATE(Withdraw) from peer 100.127.42.68 for destinations:
    114.67.92.238/32,
```

//TP-SW向ISP设备发送路由更新信息。

```
*Aug 12 15:15:37:669 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Send UPDATE to update-group 1 for following destinations:
    Origin      : IGP
    AS path     : 65030
    Next hop    : 10.10.10.10
    Community   : No-Advertise
    114.67.92.238/32,

*Aug 12 15:15:37:669 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Send UPDATE MSG to peer 100.127.96.36(IPv4-UNC) NextHop: 100.127.96.37.

*Aug 12 15:15:37:670 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Send UPDATE MSG to peer 100.127.96.38(IPv4-UNC) NextHop: 100.127.96.39.

*Aug 12 15:15:37:670 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Send UPDATE MSG to peer 100.127.96.100(IPv4-UNC) NextHop: 100.127.96.101.

*Aug 12 15:15:37:670 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Send UPDATE MSG to peer 100.127.96.102(IPv4-UNC) NextHop: 100.127.96.103.
```

//从peer 100.127.42.64 (Server 3) 接收到路由撤销消息

```
*Aug 12 15:15:37:675 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Recv UPDATE(Withdraw) from peer 100.127.42.64 for destinations:
    114.67.92.238/32,
```

//30秒后TP-SW再次向ISP设备发送路由更新信息。

```
*Aug 12 15:16:07:370 2019 HD_BX_AZTPSW_2Q-1-A_G-17/H17 BGP/7/DEBUG:
  BGP.: Send UPDATE(Withdraw) to update-group 1 for destinations:
    114.67.92.238/32,
```

解决方法

1、建议此应用场景下将版本升级至R2612P02版本，并打上R2612P02H05补丁。