

## 知 某局点S6800交换机路由选择异常

EVPN IP转发 董智敏 2022-12-09 发表

### 组网及说明

终端 (192.168.2.1) ---- (bagg114) leaf (1/0/51和1/0/52) ---- (两台spine) ----- (bagg7和bagg8)  
border (bagg100) ----- 外网10.10.10.1

## 问题描述

现场反馈终端tracert到外网时的报文在border上没有安装既定的路由下一跳转发，具体如下：

在leaf上查看去10.243.121.117的路由如下，出接口l3vni没有绑定网关的vpn

```
<H3C6800-BUSI-TOR03>dis ip routing-table vpn-instance a 10.10.10.1
```

Summary count : 2

Destination/Mask	Proto	Pre Cost	NextHop	Interface
10.0.0.0/8	BGP	255 0	10.20.0.1	Vsi50000
10.10.0.0/16	BGP	255 0	10.20.0.1	Vsi40567

#

正常来说报文从leaf发出后走上上述16网段的路由带l3vni 40567的标签到border上，然后border识别40567的标记并在该l3vni绑定的vpn里面查表转发，而且这个vpn里面也有到这个目的网段的路由

#

```
interface Vsi-interface40567
```

```
ip binding vpn-instance b
```

```
ipv6 address auto link-local
```

```
l3-vni 40567
```

```
<H3C12508F-Boder>dis ip rou vpn b 10.10.10.1
```

Summary count : 1

Destination/Mask	Proto	Pre Cost	NextHop	Interface
10.10.0.0/16	Static	60 0	172.20.2.1	Vlan567

但是客户tracert发现报文走的却是l3vni50000绑定的vpn实例里的路由。

#

```
interface Vsi-interface50000
```

```
description SDN_VRF_VSI_Interface_50000
```

```
ip binding vpn-instance c
```

```
ipv6 address auto link-local
```

```
l3-vni 50000
```

#

```
<H3C12508F-Boder>dis ip rou vpn c 10.10.10.1
```

Summary count : 2

Destination/Mask	Proto	Pre Cost	NextHop	Interface
0.0.0.0/0	BGP	130 0	210.1.3.1	Vlan124
10.0.0.0/8	Static	60 0	210.1.1.2	Vlan120

## 过程分析

1.根据border上的问题现象怀疑leaf过来的报文携带的I3vni-id是50000，因此才会在I3vni 50000绑定的vpn里面查表转发，在spine上面打印了leaf出来的报文，发现报文携带的I3vni-id果然是50000

2.查看leaf上面到目的网段的路由有两条，而从leaf出去的报文带的I3vni 50000的标签，因此怀疑报文在leaf上是走了Vsi50000出去的

```
<H3C6800-BUSI-TOR03>dis ip routing-table vpn-instance a 10.10.10.1
```

Summary count : 2

Destination/Mask	Proto	Pre Cost	NextHop	Interface
10.0.0.0/8	BGP	255 0	10.20.0.1	Vsi50000
10.10.0.0/16	BGP	255 0	10.20.0.1	Vsi40567

#

3.根据上一步的分析，怀疑10.10.0.0/16这条路由没有在底层硬件下发，因为正常情况下交换机是查底层硬件表项转发的，上面查看的只是软件表项。继续查看底层硬件表项发现硬件没有学习到该路由。

```
[H3C6800-BUSI-TOR04-probe]debug ipv4-drv show route 108 10.10.0.0 16 slot 1
```

```
*****
```

```
- IPv4 Route Information      Slot 1
```

```
*****
```

```
--- UNIT: 0 ---
```

```
- Entry not found
```

4.继续查看底层local logbuffer打印的日志信息，发现里面存在很多路由下发失败的日志信息。

```
Slot01 Dec 08 2022 17:00:37:705:
```

```
LINE:4154-TASK:kfib/1-FUNC:drv_l3uc_sdk_add_ipv4_defip:
```

```
Fail to add defip:Unit=0,iRv=-6,ip=0x64588000,mask=0xffff8000,vrf=108,intf=117146.
```

5.根据上述排查怀疑是路由表资源不够导致该路由下发失败，继续查看路由资源信息，发现设备的路由表资源已经满了。现场硬件资源模式为0，路由表规格很小，只有8k路由表资源。

```
- Unit 0
```

```
L3 INTF size: 12287
```

```
L3 INTF cnt: 291
```

```
L3 TBL size: 16384
```

```
IPv4 prefix cnt: 6208
```

```
IPv6 prefix cnt: 346
```

```
L3 TBL cnt: 6900
```

```
L3 DEFIP size: 8192 /路由表规格
```

```
IPv4 route cnt: 6272
```

```
IPv6 route cnt: 1229
```

```
L3 DEFIP cnt: 8190/已使用的路由资源
```

本次故障是因为S6800交换机路由资源不足导致硬件下发失败，现场硬件资源模式为0，该模式下路由规格最小。

## 解决方法

建议修改硬件资源模式为4解决路由资源不足的问题，修改模式后需要重启设备生效。

```
hardware-resource switch-mode 4
```

#### 注意事项

修改硬件资源模式需要重启设备生效，重启设备会影响业务，建议申请操作窗口执行

