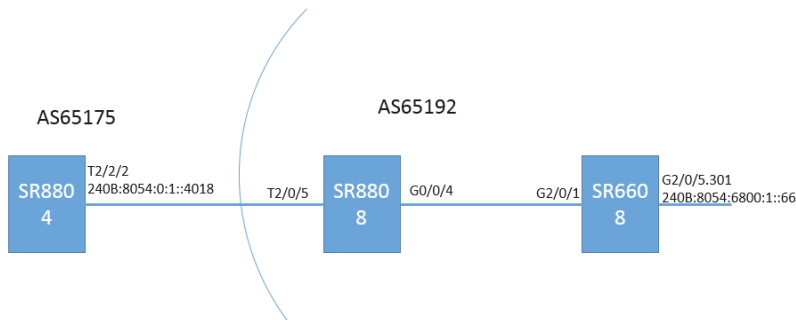


某局点SR8808-X设备关于 IPv6 BGP路由不通问题处理案例

BGP IPv6 其他硬件相关 郑标 2020-03-24 发表

组网及说明



- 1.8808上边设备8804的T2/2/2, 8808与8804是EBGP。8808是AS65192, 8804是AS65175
- 2.设备8808通过0/0/4口与6608的2/0/1口互联, 属于AS65192。8808与SR6606内部ipv6 igp 使用ospfv3建立, 业务路由发布使用bgp ipv6发布。

问题描述

6608上的2/0/5.301子接口的ipv6地址240B:8054:6800:1::66/127, 访问不了8804的T2/2/2接口地址240B:8054:0:1::4018; 6608设备接口改成126、125前缀的地址也是ping不通。
6608设备把2/0/5.301的接口地址改成前缀为128或者64的地址, 可正常ping通8804的T2/2/2接口地址240B:8054:0:1::4018。
另外, 现场反馈另外一个局点的使用的SR6616的设备, 对应使用的是MIC插卡, 同样的组网, 使用127前缀地址均可以。
该局点对应接口使用的插卡是RT-SFE-L1。

过程分析

一、在66上使用127位的地址来ping测试8804的下联口地址, 发现地址不通, 在两端查看路由表路由都有, 在8808上查看路由表和fib表项都正常

配置如下

```
[SDLY-SDZWW-YNX-SR6608-GigabitEthernet2/0/5.301] //掩码127
#
interface GigabitEthernet2/0/5.301
description shenji-ipv6
vlan-type dot1q vid 301
ipv6 address 240B:8054:6800:1::66/127
#
return
[SDLY-SDZWW-YNX-SR6608-GigabitEthernet2/0/5.301]ping ipv6 -a 240B:8054:6800:1::66
240B:8054:0:1::4018 //ping测试
Ping6(56 data bytes) 240B:8054:6800:1::66 --> 240B:8054:0:1::4018, press CTRL_C to break
Request time out
Request time out
Request time out
Request time out
Request time out
--- Ping6 statistics for 240B:8054:0:1::4018 ---
5 packet(s) transmitted, 0 packet(s) received, 100.0% packet loss
<SDLY-SDZWW-YNX-SR6608>dis ipv6 routing-table 240B:8054:0:1::4018 //从6608上查看8804的
路由
Summary count : 2
Destination: 240B:8054::/30                Protocol : BGP4+
NextHop   : 240B:8054:6800::                Preference: 255
Interface : GE2/0/1                          Cost      : 0
Destination: 240B:8054:0:1::4018/127       Protocol : BGP4+
NextHop   : 240B:8054:6800::                Preference: 255
Interface : GE2/0/1                          Cost      : 0
<SDLY-SDZWW-JQS-SR8804-B1>dis ipv6 routing-table 240B:8054:6800:1::66 //从8804上查看
6608的路由
Summary count : 2
```

```

Destination: 240B:8054::/30          Protocol : BGP4+
NextHop  : 240B:8054::1              Preference: 255
Interface : XGE2/1/1                  Cost      : 0
Destination: 240B:8054:6800::/40     Protocol : BGP4+
NextHop  : 240B:8054:0:1::4019       Preference: 255
Interface : XGE2/2/2                  Cost      : 0
[SDLY-SDZWW-XXX-GYW-SR8808-01]dis ipv6 fib 240B:8054:6800:1::66 //8808上查看表项//当66
08掩码为127为位时fib表项
Destination count: 1 FIB entry count: 1
Flag:
  U:Usable G:Gateway H:Host B:Blackhole D:Dynamic S:Static
  R:Relay F:FRR
Destination: 240B:8054:6800:1::66     Prefix length: 127
Nexthop  : FE80::DEDA:80FF:FEFB:99BD   Flags: UGR
Time stamp : 0x2e26                      Label: Null
Interface : GE0/0/4                      Token: Invalid
[SDLY-SDZWW-XXX-GYW-SR8808-01]dis ipv6 routing-table 240B:8054:6800:1::66 //8808上查看表
项//当6608掩码为127为位时路由表项
Summary count : 3
Destination: 240B:8054::/30          Protocol : BGP4+
NextHop  : 240B:8054:0:1::4018       Preference: 255
Interface : XGE2/0/5                  Cost      : 0
Destination: 240B:8054:6800::/40     Protocol : Static
NextHop  : ::                          Preference: 60
Interface : NULL0                      Cost      : 0
Destination: 240B:8054:6800:1::66/127 Protocol : BGP4+
NextHop  : 240B:8054:6800::7         Preference: 255
Interface : GE0/0/4                    Cost      : 0

```

二、在66上使用128位的地址来ping测试8804的下联口地址，发现正常，在两端查看路由表路由都有，在8808上查看路由表和fib表项都正常

配置如下

```

[SDLY-SDZWW-YNX-SR6608-GigabitEthernet2/0/5.301] //掩码128
#
interface GigabitEthernet2/0/5.301
description shenji-ipv6
vlan-type dot1q vid 301
ipv6 address 240B:8054:6800:1::66/128
#
return
[SDLY-SDZWW-YNX-SR6608-GigabitEthernet2/0/5.301]ping ipv6 -a 240B:8054:6800:1::66
240B:8054:0:1::4018 //ping测试
Ping6(56 data bytes) 240B:8054:6800:1::66 --> 240B:8054:0:1::4018, press CTRL_C to break
56 bytes from 240B:8054:0:1::4018, icmp_seq=0 hlim=63 time=8.326 ms
56 bytes from 240B:8054:0:1::4018, icmp_seq=1 hlim=63 time=7.986 ms
56 bytes from 240B:8054:0:1::4018, icmp_seq=2 hlim=63 time=7.664 ms
56 bytes from 240B:8054:0:1::4018, icmp_seq=3 hlim=63 time=8.332 ms
56 bytes from 240B:8054:0:1::4018, icmp_seq=4 hlim=63 time=8.323 ms
--- Ping6 statistics for 240B:8054:0:1::4018 ---
5 packet(s) transmitted, 5 packet(s) received, 0.0% packet loss
round-trip min/avg/max/std-dev = 7.664/8.126/8.332/0.266 ms
<SDLY-SDZWW-YNX-SR6608>dis ipv6 routing-table 240B:8054:0:1::4018 //从6608上查看8804的
路由
Summary count : 2
Destination: 240B:8054::/30          Protocol : BGP4+
NextHop  : 240B:8054:6800::          Preference: 255
Interface : GE2/0/1                  Cost      : 0
Destination: 240B:8054:0:1::4018/127 Protocol : BGP4+
NextHop  : 240B:8054:6800::          Preference: 255
Interface : GE2/0/1                  Cost      : 0
<SDLY-SDZWW-JQS-SR8804-B1>dis ipv6 routing-table 240B:8054:6800:1::66 //从8804上查看
6608的路由
Summary count : 2
Destination: 240B:8054::/30          Protocol : BGP4+

```

```

NextHop : 240B:8054::1          Preference: 255
Interface : XGE2/1/1           Cost : 0
Destination: 240B:8054:6800::/40      Protocol : BGP4+
NextHop : 240B:8054:0:1::4019       Preference: 255
Interface : XGE2/2/2           Cost : 0
[SDLY-SDZWW-XXX-GYW-SR8808-01]dis ipv6 routing-table 240B:8054:6800:1::66 //8808上查看表项//当6608掩码为128为位时路由表项

```

```

Summary count : 3
Destination: 240B:8054::/30          Protocol : BGP4+
NextHop : 240B:8054:0:1::4018       Preference: 255
Interface : XGE2/0/5               Cost : 0
Destination: 240B:8054:6800::/40      Protocol : Static
NextHop : ::                        Preference: 60
Interface : NULL0                   Cost : 0
Destination: 240B:8054:6800:1::66/128  Protocol : BGP4+
NextHop : 240B:8054:6800::7         Preference: 255
Interface : GE0/0/4                 Cost : 0

```

```

[SDLY-SDZWW-XXX-GYW-SR8808-01]
[SDLY-SDZWW-XXX-GYW-SR8808-01]dis ipv6 fib 240B:8054:6800:1::66 //8808上查看表项//当6608掩码为128为位时fib表项

```

Destination count: 1 FIB entry count: 1

Flag:

U:Usable G:Gateway H:Host B:Blackhole D:Dynamic S:Static
R:Relay F:FRR

```

Destination: 240B:8054:6800:1::66      Prefix length: 128
Nexthop : FE80::DEDA:80FF:FEFB:99BD     Flags: UGHR
Time stamp : 0x2e28                      Label: Null
Interface : GE0/0/4                       Token: Invalid

```

三、通过流统发现，当掩码为127位时，在8808这台硬转设备上丢了，丢的原因是硬件资源不足了，导致fib没有下发成功：

```
debug ipv4-drv show config slot 0
```

```
*****
```

```
- IPv4 Config Slot 0
```

```
*****
```

```
- ARP SIZE: 16384
```

```
- ArpCanNotSetToHW: NO
```

```
- IPV4 ROUTE SIZE: 65536
```

```
- ECMP SIZE: 8
```

```
- ND SIZE: 8192
```

```
- IPV6 ROUTE SIZE: 8192
```

```
- IPV6 LongPrefRT: 128-----一共128个资源，掩码65~127范围的会申请这个。
```

```
- VLAN INTF MODE: 2
```

```
- NH SIZE: 16384
```

```
debug ipv6-drv show statistics slot 0
```

```
*****
```

```
- IPv6 Statistics Slot 0
```

```
*****
```

```
- ROUTE TOTAL COUNT: 0
```

```
- ROUTE LongPref COUNT: 128-----ipv6已经用了128了，用满了。
```

```
- ROUTE DMAFAIL COUNT: 0
```

```
- ECMP ROUTE COUNT: 0
```

```
- ECMP DMAFAIL COUNT: 0
```

```
- Linklocal-ND COUNT: 0
```

```
- ND-ACL COUNT: 0
```

```
- IPV6 Plat ND Demand NUM: 1340
```

```
- IPV6 ND Successed NUM: 1305
```

```
- IPV6 Plat Route Demand NUM: 17362799
```

```
- IPV6 Route Successed NUM: 46287
```

解决方法

- 1、需要转发通的业务，路由前缀采用64位掩码或者128位掩码。
- 2、将当前板卡更换成spe板卡。

