



某局点S125R BGP路由聚合问题

BGP

张文宁

2023-06-29 发表

组网及说明

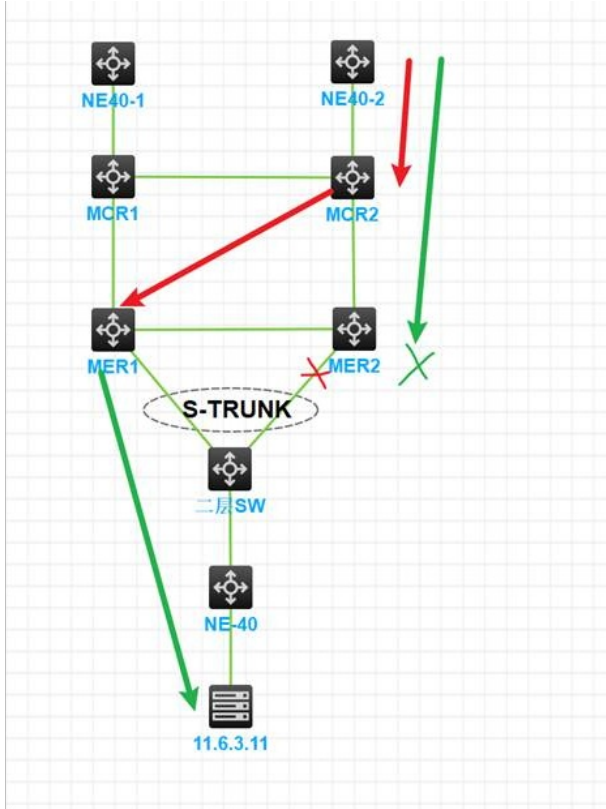
/

问题描述

设备型号及版本: S12504R R3608

问题描述:

125R设备作为MER设备, 现场想要精简路由规格, 因此在两台125R设备BGP视图下的ipv4地址族做路由聚合, 发现聚合后右侧NE40-2设备访问server不通了, 原因是MCR-2上路由指向MER2, MER2上路由指向黑洞, 所以不通; 红色为正常流量走向, 绿色为不通时走向。



初步分析为: MER设备对接大二层交换机时使用S-trunk功能, 即两台MER只有其中一台设备上存在用户路由信息, 另外一台设备不存在用户路由信息。如果做路由聚合, 则两台设备都会向MCR发送聚合后的路由信息。导致寻址出现问题。

正常路由信息为: NE40-1(软交换核心) =====>MCR-1=====》MER-1=====》大二层交换机=====》NE40(用户接入) =====》用户

NE40-2(软交换核心) =====>MCR-2=====》MER-1=====》大二层交换机=====》NE40(用户接入) =====》用户

由于S-trunk功能原因只有MER-1上存在用户路由, 所以此时寻址正常

异常路由信息为: NE40-1(软交换核心) =====>MCR-1=====》MER-1=====》大二层交换机=====》NE40(用户接入) =====》用户

NE40-2(软交换核心) =====>MCR-2=====》MER-2=====》大二层交换机=====》NE40(用户接入) =====》用户

由于做了路由聚合, 则此时MER2也向MCR宣告路由, 则寻址不正常

初步分析为: MER设备对接大二层交换机时使用S-trunk功能, 即两台MER只有其中一台设备上存在用户路由信息, 另外一台设备不存在用户路由信息。如果做路由聚合, 则两台设备都会向MCR发送聚合后的路由信息。导致寻址出现问题。

正常路由信息为: NE40-1(软交换核心) =====>MCR-1=====》MER-1=====》大二层交换机=====》NE40(用户接入) =====》用户

NE40-2(软交换核心) =====>MCR-2=====》MER-1=====》大二层交换机=====》NE40(用户接入) =====》用户

由于S-trunk功能原因只有MER-1上存在用户路由, 所以此时寻址正常

异常路由信息为: NE40-1(软交换核心) =====>MCR-1=====》MER-1=====》大二层交换机=====》NE40(用户接入) =====》用户

NE40-2(软交换核心) =====>MCR-2=====》MER-2=====》大二层交换机=====》NE40(用户接入) =====》用户

由于做了路由聚合, 则此时MER2也向MCR宣告路由, 则寻址不正常

MER2上的路由:

<SDZZ-551J-4F-S12504R-SM-E006>dis ip routing-table vpn-instance IMS_RAN 11.6.0.0

过程分析

Summary count : 1

分析发现属于现场配置问题：

现场MER2上符合路由聚合11.6.0.0/16的明细路由如下，有一部分是MER1发过来的路由（下一跳是RAGG1000.1000），还有一部分是MCR2发过来的路由（下一跳是RAGG999），此时就会有问题。如果把聚合路由改成11.6.0.0/21，这样MCR2发过来的明细路由就不会被聚合，就不会有问题

取消路由聚合后，正常网段路由指向MER1的，没有问题。

```
SDZZ-551J-4F-S12504R-SM-E006-bgp-default-ipv4-IMS_RAN#dis ip routing-table vpn-instance IMS_RAN 11.6.0.0
BGP 255 0 10.53.221.5 RAGG1000.1000
11.6.0.0/24 BGP 255 0 10.53.221.5 RAGG1000.1000
11.6.255.0/24 BGP 255 0 10.53.221.5 RAGG999
```

```
SDZZ-551J-4F-S12504R-SM-E006-bgp-default-ipv4-IMS_RAN#dis ip routing-table vpn-instance IMS_RAN 11.6.0.0
BGP 255 0 10.53.221.255 RAGG999
```

```
11.6.44.0/24 BGP 255 0 10.53.221.255 RAGG999
```

```
11.6.44.1/32 BGP 255 0 10.53.221.255 RAGG999
```

Summary count : 1

```
11.6.60.0/24 BGP 255 0 10.53.221.255 RAGG999
```

```
11.6.60.1/32 BGP 255 0 10.53.221.255 RAGG999
```

```
Destination/Mask Proto Pre Cost NextHop Interface
11.6.0.0/24 BGP 255 0 10.53.221.5 RAGG1000.1000
```

SDZZ-551J-4F-S12504R-SM-E006-bgp-default-ipv4-IMS_RAN#dis ip routing-table vpn-instance IMS_RAN 11.6.0.0

的聚合路由，但是MER2的聚合以后看发布的路由下一跳是指向自己的，而不是MER-1，这个有点奇怪

Summary count : 1

```
Destination/Mask Proto Pre Cost NextHop Interface
RAGG1000.1000 BGP 130 0 127.0.0.1 NULL0
```

Current state: DOWN

Line protocol state: DOWN

Description: to_SDZZ-XiGang-S9312-SW1_TRUNK44:(IMS_RAN)

Maximum transmission unit: 1500

Internet protocol processing: Disabled

IP packet frame type: Ethernet II, hardware address: 78aa-82c8-1c01

IPv6 packet frame type: Ethernet II, hardware address: 78aa-82c8-1c01

Last clearing of counters: Never

Last 300 second input rate: 17 bytes/sec, 136 bits/sec, 0 packets/sec

Last 300 second output rate: 141 bytes/sec, 1128 bits/sec, 1 packets/sec

Input: 44801 packets, 11748433 bytes, - drops

Output: 1924604 packets, 176780102 bytes, - drops

解决方法

聚合掩码错误导致,错误地把上行MCR过来的路由也并到一起了,导致我们认为本地都有明细路由,所以此时聚合后的16位掩码路由直接置为黑洞了.

修正配置后解决.

