S9850-32HEBGP邻居状态正常学习不到对端发布的路由

BGP　祁振峰本尊　2024-05-30 发表

问题描述

我们设备与友商建立EBGP对等体，如下：

```
============================================================
==============display bgp peer ipv4==============

 BGP local router ID: 192.168.48.254
 Local AS number: 65535
 Total number of peers: 4     Peers in established state: 4

 * - Dynamically created peer
 ^ - Peer created through link-local address
 Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

 192.168.48.216  3232248024       59       68    0       0 00:25:03 Established
 192.168.48.217  3232248025       59       79    0       0 00:25:05 Established
 192.168.48.218  3232248026       56       68    0       0 00:25:05 Established
 192.168.48.219  3232248027       55       72    0       0 00:25:11 Established
```

对端引入路由后，本端学不到：

对端BGP路由表（模糊不清时请点击放大）：

```
root@sd1a-rsserver0:~# gobgp neighbor 192.168.52.254 adj-in
   ID  Network            Next Hop            AS_PATH            Age       Attrs
   0   0.0.0.0/0          192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
   0   10.255.254.0/24    192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
   0   172.31.0.0/24      192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
   0   172.31.1.0/24      192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
   0   192.168.48.0/24    192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
   0   192.168.49.0/24    192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
   0   192.168.52.254/32  192.168.52.254      65535              00:00:40  [{Origin: i} {Med: 0}]
   0   192.168.54.0/24    192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
   0   192.168.55.0/24    192.168.52.254      65535              00:00:40  [{Origin: ?} {Med: 0}]
root@sd1a-rsserver0:~# gobgp neighbor 192.168.52.254 adj-out
   ID  Network            Next Hop            AS_PATH             Attrs
   1   60.217.65.246/32   192.168.49.183      3232248024 3232248247 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:1000]}]
   1   111.35.68.195/32   192.168.49.183      3232248024 3232248247 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:1000]}]
   1   192.168.50.0/24    192.168.49.162      3232248024 3232248226 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:3920403]}]
   1   192.168.50.2/32    192.168.49.183      3232248024 3232248247 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:3920403]}]
   1   192.168.50.4/32    192.168.49.183      3232248024 3232248247 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:3920403]}]
   1   192.168.51.0/24    192.168.49.183      3232248024 3232248247 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:5724808]}]
   1   192.168.51.2/32    192.168.49.162      3232248024 3232248226 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:5724808]}]
   1   192.168.51.4/32    192.168.49.183      3232248024 3232248247 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:5724808]}]
   1   192.168.51.5/32    192.168.49.172      3232248024 3232248236 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:5724808]}]
   1   192.168.51.6/32    192.168.49.172      3232248024 3232248236 [{Origin: ?} {Med: 0} {Communities: 0:65535, 1:24465} {Extcomms: [1:5724808]}]
```

本地BGP路由表：

```
<CORE-SW-S9850-E04-E05-39U>display  bgp  peer  ipv4

 BGP local router ID: 192.168.48.254
 Local AS number: 65535
 Total number of peers: 4                 Peers in established state: 4

  * - Dynamically created peer
  ^ - Peer created through link-local address
  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  192.168.48.216  3232248024       67       77    0       0 00:29:03 Established
  192.168.48.217  3232248025       67       90    0       0 00:29:05 Established
  192.168.48.218  3232248026       64       77    0       0 00:29:05 Established
  192.168.48.219  3232248027       63       81    0       0 00:29:11 Established
<CORE-SW-S9850-E04-E05-39U>display  bgp  rou
<CORE-SW-S9850-E04-E05-39U>display  bgp  routing-table ipv4

 Total number of routes: 16

 BGP local router ID is 192.168.48.254
 Status codes: * - valid, > - best, d - dampened, h - history
               s - suppressed, S - stale, i - internal, e - external
               a - additional-path
       Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop                 MED        LocPrf     PrefVal Path/Ogn

* >  0.0.0.0            172.31.0.1              0                     32768   ?
* >  10.255.254.0/24    10.255.254.254          0                     32768   ?
* >  10.255.254.254/32  127.0.0.1               0                     32768   ?
* >  172.31.0.0/24      172.31.0.254            0                     32768   ?
* >  172.31.0.254/32    127.0.0.1               0                     32768   ?
* >  172.31.1.0/24      172.31.1.254            0                     32768   ?
* >  172.31.1.254/32    127.0.0.1               0                     32768   ?
* >  192.168.48.0       192.168.48.254          0                     32768   ?
* >  192.168.48.254/32  127.0.0.1               0                     32768   ?
* >  192.168.49.0       192.168.49.254          0                     32768   ?
* >  192.168.49.254/32  127.0.0.1               0                     32768   ?
* >  192.168.52.254/32  127.0.0.1               0                     32768   i
* >  192.168.54.0       192.168.54.254          0                     32768   ?
* >  192.168.54.254/32  127.0.0.1               0                     32768   ?
* >  192.168.55.0       192.168.55.254          0                     32768   ?
* >  192.168.55.254/32  127.0.0.1               0                     32768   ?
<CORE-SW-S9850-E04-E05-39U>
```

过程分析

从BGP路由表看，对端发布的BGP路由（例如：192.168.50.0/24）本端确实没有加BGP路由表，在本端debug后发现有如下提示：

*Jan 23 00:48:45:164 2015 CORE-SW-S9850-E04-E05-39U BGP/7/DEBUG:
BGP.: Recv UPDATE(Withdraw) from peer 192.168.48.219 for destinations:
192.168.51.0/24 PathID 0 ,
*Jan 23 00:48:45:164 2015 CORE-SW-S9850-E04-E05-39U BGP/7/DEBUG:
BGP.: 192.168.48.219 MSG ignored: Invalid NEXT_HOP attribute 192.168.49.162,
while recv UPDATE.
*Jan 23 00:48:45:164 2015 CORE-SW-S9850-E04-E05-39U BGP/7/DEBUG:
BGP.: Recv UPDATE(Withdraw) from peer 192.168.48.219 for destinations:
192.168.50.0/24 PathID 0 ,
%Jan 23 00:48:51:164 2015 CORE-SW-S9850-E04-E05-39U BGP/5/BGP_STATE_CHANGED: BGP
.: 192.168.48.218 state has changed from OPENCONFIRM to ESTABLISHED.
%Jan 23 00:48:51:168 2015 CORE-SW-S9850-E04-E05-39U BGP/5/BGP_STATE_CHANGED: BGP
.: 192.168.48.217 state has changed from OPENCONFIRM to ESTABLISHED.
*Jan 23 00:48:51:168 2015 CORE-SW-S9850-E04-E05-39U BGP/7/DEBUG:
BGP.: 192.168.48.218 MSG ignored: Invalid NEXT_HOP attribute 192.168.49.183,
while recv UPDATE.

根据debug的提示，本端收到路由更新后，认为路由下一跳是无效的，原因是该下一跳地址不是本端和对端建立EBGP邻居关系时使用的地址。正常来讲，关于EBGP路由发布是有约定俗成的规范的：EBGP邻居的BGP路由的下一跳设置为自身向这个EBGP邻居发送BGP报文的源地址(即本端与EBGP邻居建立邻居所使用接口的接口地址)该故障也是对端发给我们的路由下一跳没有改，保持了对端自己路由表的下一跳导致的。

解决方法

两种方法：

1．本端和对端改用loopback口建立BGP邻居，然后指定连接接口为vlan 48口

2．对端修改bgp下一跳属性为192.168.48.x，例如：

*Jan 23 00:48:51:168 2015 CORE-SW-S9850-E04-E05-39U BGP/7/DEBUG:
BGP.: Recv UPDATE(Withdraw) from peer 192.168.48.218 for destinations:

192.168.51.0/24 PathID 0 ,
*Jan 23 00:48:51:168 2015 CORE-SW-S9850-E04-E05-39U BGP/7/DEBUG:
BGP.: 192.168.48.218 MSG ignored: Invalid NEXT_HOP attribute 192.168.49.162,
while recv UPDATE.
192.168.48.218向我们发布的去往192.168.51.0/24的路由时，应该把NEXT_HOP修改为192.168.48.218再发布而不是使用192.168.49.162

方法2需要对端做BGP属性的修改；方法1会断开BGP邻居重新建立，选用方法1，需要通过命令peer ebgp-max-hop来指定允许的最大跳数。