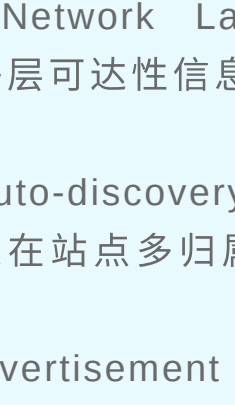


EVPN路由消息

上一期技术专题我们介绍了EVPN概述及场景，接下来我们就要开始研究EVPN技术啦~



为了支持EVPN，MP-BGP在L2VPN地址族下定义了新的子地址族—EVPN地址族，并新增了EVPN NLRI （ Network Layer Reachability Information，网络层可达性信息），即EVPN路由消息：

1. Ethernet Auto-discovery Route：以太网自动发现路由，用来在站点多归属组网中通告ES信息。
2. MAC/IP Advertisement Route：MAC/IP发布路由，用来通告MAC地址和主机路由信息（即ARP信息和ND信息）。
3. Inclusive Multicast Ethernet Tag Route：包含性组播以太网标签路由，又称为IMET路由。在EVPN VXLAN组网中用来通告VTEP及其所属VXLAN信息，以实现自动发现VTEP、自动建立VXLAN隧道和自动关联VXLAN与VXLAN隧道。
4. Ethernet Segment Route：以太网段路由，用来通告ES及其连接的VTEP信息。
5. IP Prefix advertisement route：IP前缀路由，用来以IP前缀的形式通告BGP IPv4单播路由或BGP IPv6单播路由。

在典型EVPN组网中，比较常用的是2、3、5类路由消息。



EVPN消息RT值

VTEP 在发布 EVPN 路由时，会携带 VPN Target扩展团体属性（也称为Route Target，即我们常说的RT值）。VPN target属性定义了本地发送的EVPN路由可以为哪些VTEP所接收，VTEP可以接收哪些远端VTEP发送来的EVPN路由。MP-BGP通过VPN Target属性来控制EVPN路由信息的发布与接收。EVPN中一共有三类RT值，分别如下：

1、VXLAN RT值

此RT值在VSI视图下配置。

```
vsi vpna

gateway vsi-interface 100

vxlan 100

evpn encapsulation vxlan

route-distinguisher auto

vpn-target auto export-extcommunity

vpn-target auto import-extcommunity
```

2、VPN IPv4 RT值

此RT值在VPN下的IPv4地址族中配置。

```
ip vpn-instance vpna

route-distinguisher 1:1

#

address-family ipv4

vpn-target 10:10 import-extcommunity

vpn-target 10:10 export-extcommunity

#
```

3、VPN EVPN RT值

此RT值在VPN下的EVPN地址族中。

```
ip vpn-instance vpna

route-distinguisher 1:1

#

address-family evpn

vpn-target 10:10 import-extcommunity

vpn-target 10:10 export-extcommunity

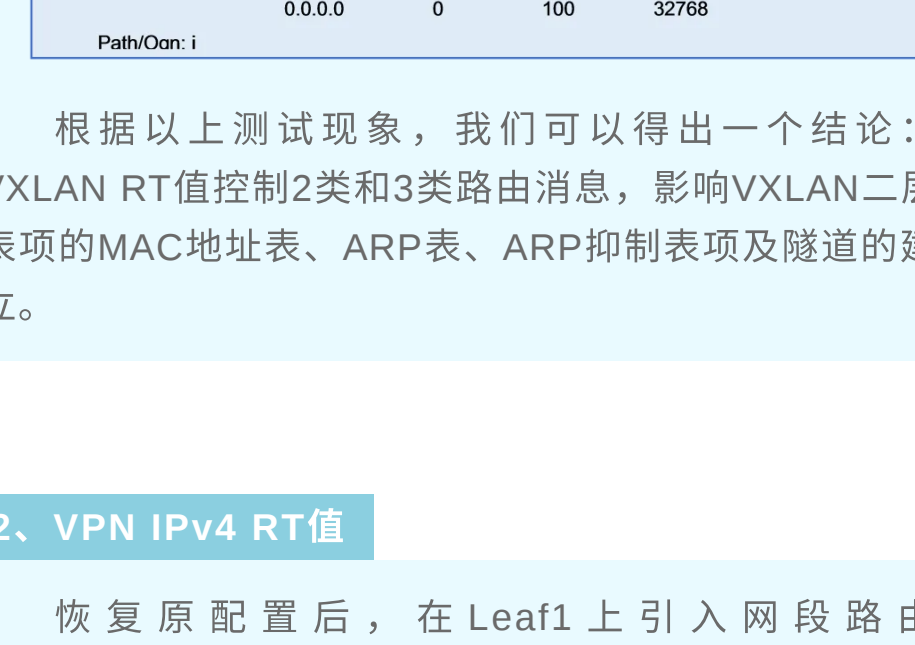
#
```

这三类RT值的作用是什么呢，下面我们来一探探索~

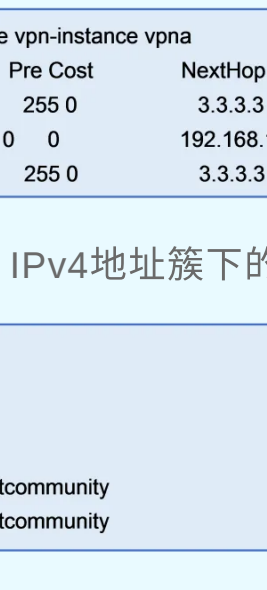


三类RT值的作用

在开始探究前，我们先搭建一个最基本的EVPN VXLAN网络。



（大家有兴趣可以在HCL上做一下这个探究实验哦~）



1、VXLAN RT值

在搭建好组网后，首先查看两边Leaf的VXLAN隧道情况，能看到自动建立的隧道：

```
<Leaf1>display vxlan tunnel

VXLAN ID: 100, VSI name: vpna, Total tunnels: 1 (1 up, 0 down, 0 defect, 0 blocked)
Tunnel name Source Destination State Type Flood proxy Out VNI
Tun0 1.1.1.1 3.3.3.3 UP Auto Disabled -

VXLAN ID: 1000, VSI name: Auto_L3VNI1000_3
```

Leaf1能学到本端和对端的2类和3类路由，Leaf2类似：

```
<Leaf1>display bgp l2vpn evpn route-type mac-ip
Route distinguisher: 1:100
Total number of routes: 3

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [2][0][48][7425-8ae3-6b7a][0][0.0.0.0]/104
3.3.3.3 0 100 0

Path/Ogn: i

* > [2][0][48][7425-8ae3-6b7a][32][100.1.1.20]/136
3.3.3.3 0 100 0

Path/Ogn: i

* > [2][0][48][7425-8ae4-444e][0][0.0.0.0]/104
0.0.0.0 0 100 32768

Path/Ogn: i

* > [2][0][48][7425-8ae4-444e][32][100.1.1.10]/136
0.0.0.0 0 100 32768

Path/Ogn: i
```

```
<Leaf1>display bgp l2vpn evpn route-type imet
Route distinguisher: 1:100
Total number of routes: 2

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [3][0][32][1.1.1.1]/80
0.0.0.0 0 100 32768

Path/Ogn: i

* > [3][0][32][3.3.3.3]/80
3.3.3.3 0 100 0

Path/Ogn: i
```

修改Leaf1的VXLAN RT值，Leaf2保持自动生成：

```
vsi vpna
gateway vsi-interface 100
vxlan 100
evpn encapsulation vxlan
route-distinguisher auto
vpn-target 100:100 export-extcommunity
vpn-target 100:100 import-extcommunity
```

此时查看两边VXLAN隧道，vsi vpna通过3类路由的自动建立的隧道消失：

```
[Leaf1]display vxlan tunnel
VXLAN ID: 100, VSI name: vpna
VXLAN ID: 1000, VSI name: Auto_L3VNI1000_3
```

查看Leaf1的2类和3类路由，vsi vpna只有本地路由，无对端Leaf路由：

```
[Leaf1]display bgp l2vpn evpn route-type mac-ip
Route distinguisher: 1:100
Total number of routes: 3

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [2][0][48][7425-8ae4-444e][0][0.0.0.0]/104
0.0.0.0 0 100 32768

Path/Ogn: i

* > [2][0][48][7425-8ae4-444e][32][100.1.1.10]/136
0.0.0.0 0 100 32768

Path/Ogn: i
```

```
[Leaf1]display bgp l2vpn evpn route-type imet
Route distinguisher: 1:100
Total number of routes: 1

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [3][0][32][1.1.1.1]/80
0.0.0.0 0 100 32768

Path/Ogn: i
```

根据以上测试现象，我们可以得出一个结论：VXLAN RT值控制2类和3类路由消息，影响VXLAN二层表项的MAC地址表、ARP表、ARP抑制表项及隧道的建立。

2、VPN IPv4 RT值

恢复原配置后，在Leaf1上引入网段路由192.168.1.0/24，在Leaf2上引入网段路由192.168.2.0/24，Leaf1上查看路由表如下，能学到32位的主机路由和远端网段路由：

```
[Leaf1]display ip routing-table vpn-instance vpna
Destination/Mask Proto Pre Cost NextHop Interface
100.1.1.20/32 BGP 255 0 3.3.3.3 Vsi3
192.168.1.0/24 Direct 0 0 192.168.1.1 Loop1
192.168.2.0/24 BGP 255 0 3.3.3.3 Vsi3
```

修改Leaf1 VPN IPv4地址簇下的RT值：

```
ip vpn-instance vpna
route-distinguisher 1:1
#
address-family ipv4
vpn-target 11:11 import-extcommunity
vpn-target 11:11 export-extcommunity
```

此时查看路由表，发现路由表中只有32位主机路由：

```
[Leaf1]display ip routing-table vpn-instance vpna
Destination/Mask Proto Pre Cost NextHop Interface
100.1.1.20/32 BGP 255 0 3.3.3.3 Vsi3
192.168.1.0/24 Direct 0 0 192.168.1.1 Loop1
```

查看EVPN路由表的5类路由，只有本地的192.168.1.0/24网段路由，没有Leaf2引入的路由：

```
[Leaf1]display bgp l2vpn evpn route-type ip-prefix
Route distinguisher: 1:1(vpna)
Total number of routes: 3

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [5][0][24][100.1.1.0]/80
0.0.0.0 0 100 32768

Path/Ogn: i

* > [5][0][24][192.168.1.0]/80
127.0.0.1 0 32768

Path/Ogn: i
```

从以上测试现象可以看出：VPN IPv4 RT可以控制5类路由消息，影响网段路由和外部引入的路由，或者向外部引出的路由。

3、VPN EVPN RT值

恢复原本的配置后，查看Leaf1上的2类路由，可以正常收到远端终端的32位主机路由消息：

```
[Leaf1]display bgp l2vpn evpn route-type mac-ip
Route distinguisher: 1:1(vpna)
Total number of routes: 2

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [2][0][48][7425-8ae3-6b7a][32][100.1.1.20]/136
3.3.3.3 0 100 0

Path/Ogn: i
Route distinguisher: 1:100
Total number of routes: 4

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [2][0][48][7425-8ae3-6b7a][0][0.0.0.0]/104
3.3.3.3 0 100 0

Path/Ogn: i
* > [2][0][48][7425-8ae3-6b7a][32][100.1.1.20]/136
3.3.3.3 0 100 0

Path/Ogn: i
* > [2][0][48][7425-8ae4-444e][0][0.0.0.0]/104
0.0.0.0 0 100 32768

Path/Ogn: i
* > [2][0][48][7425-8ae4-444e][32][100.1.1.10]/136
0.0.0.0 0 100 32768

Path/Ogn: i
```

修改Leaf1 VPN视图下EVPN地址簇下的RT值：

```
ip vpn-instance vpna
route-distinguisher 1:1
#
address-family evpn
vpn-target 11:11 import-extcommunity
vpn-target 11:11 export-extcommunity
```

查看Leaf1上VPN路由表，只有远端的网段路由，没有32位主机路由：

```
[Leaf1]display ip routing-table vpn-instance vpna
Destination/Mask Proto Pre Cost NextHop Interface
192.168.1.0/24 Direct 0 0 192.168.1.1 Loop1
192.168.2.0/24 BGP 255 0 3.3.3.3 Vsi3
```

查看Leaf1上的2类路由，没有收到远端Leaf发过来的2类路由，所以路由表中没有远端的主机路由：

```
[Leaf1]display bgp l2vpn evpn route-type mac-ip
Route distinguisher: 1:100
Total number of routes: 3

Network NextHop MED LocPrf PrefVal Path/Ogn
* > [2][0][48][7425-8ae4-444e][0][0.0.0.0]/104
0.0.0.0 0 100 32768

Path/Ogn: i
```

根据以上测试现象，VPN EVPN RT值控制2类路由消息，用于将32位主机路由导入VPN，形成主机路由。

总结

为了支持EVPN，MP-BGP在L2VPN地址族下定义了EVPN地址族，并新增了EVPN路由消息，VTEP在发布EVPN路由时，会携带RT值来控制EVPN路由信息的发布与接收。EVPN中一共有三类RT值，作用总结如下：

1)第一个RT：VXLAN RT，VSI视图下的配置，一般使用auto自动生成，由2类和3类路由消息携带用于形成基于VXLAN的MAC地址表、ARP表、ARP抑制表项及隧道建立。

Auto填充规则：自动生成的RT取值为“BGP AS:VXLAN ID”；如果是EBGP的话，使用3类消息建隧道需要注意该RT的值要手动配置。

2)第二个RT：VPN IPv4 RT，VPN下的IPv4地址族RT，用于形成VPN的IPv4路由表，5类路由携带，一般网段路由和外部引入的路由或者向外部引出路由时进行控制。

3)第三个RT：VPN EVPN RT，由ARP的TYPE2路由携带，用于将32位主机路由导入VPN，形成主机路由，在纯二层环境中ARP的2类消息中只携带第一类RT，当配置了L3-VNI，ARP的2类消息中会同时携带1和3这两个RT，两个RT均通过了才能学习路由形成主机路由。

在日常EVPN网络搭建和维护中，如果遇到路由引入、隧道建立的问题，不妨检查一下各类RT值的配置~