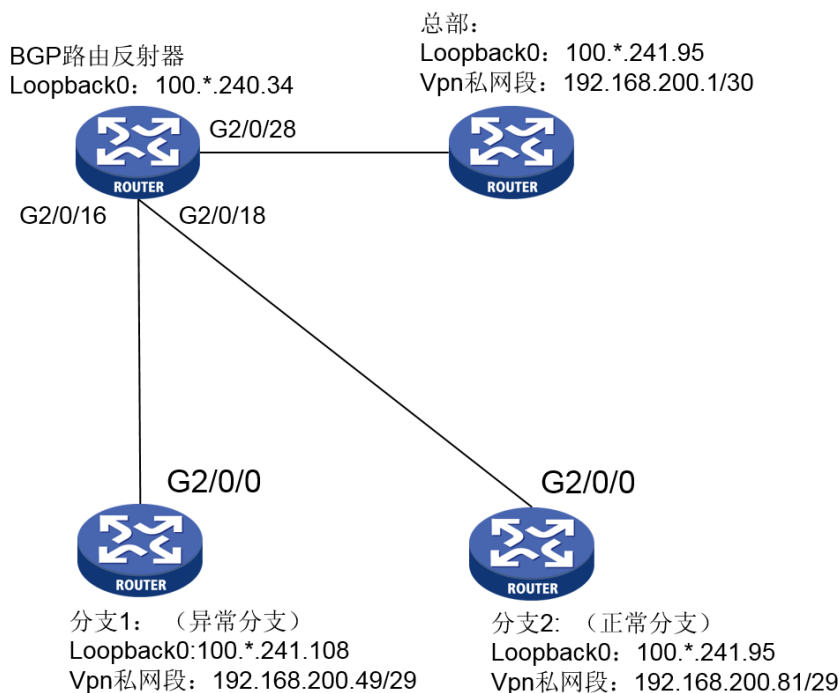


某局点MSR5620 BGP MPLS vpn分支和总部之间vpn路由无法互相学习排查经验案例

MPLS L3VPN VPN实例 徐猛 2019-01-24 发表

组网及说明

分部和总部都是MSR5620设备，多个分部和总部间建立BGP MPLS VPN，目前有一个分部和总部之间以及异常分部之间的VPN私网路由学习是正常的，另有一个异常分部和总部之间的VPN私网路由学习异常，相互之间的私网路由无法正常学习，但是该分部和其他分部之间VPN私网路由学习正常。其中母局设备SR8810作为组网中的P设备，同时配置了BGP反射器，总部和分部之间都和反射器建立BGP邻居。



问题描述

分部和总部都是MSR5620设备，多个分部和总部间建立BGP MPLS VPN，目前有一个分部和总部之间以及异常分部之间的VPN私网路由学习是正常的，另有一个异常分部和总部之间的VPN私网路由学习异常，相互之间的私网路由无法正常学习，但是该分部和其他分部之间VPN私网路由学习正常。其中母局设备SR8810作为组网中的P设备，同时配置了BGP反射器，总部和分部之间都和反射器建立BGP邻居。

(1) 异常分部的路由: 少了到总部的vpn私网192.168.200.1的路由，有到其他分部的vpn私网路由:

```
display ip routing-table vpn-instance SD_ZhengFu_SheBao
Destinations : 11    Routes : 11
Destination/Mask Proto Pre Cost NextHop Interface
0.0.0.0/32 Direct 0 0 127.0.0.1 InLoop0
127.0.0.0/8 Direct 0 0 127.0.0.1 InLoop0
127.0.0.0/32 Direct 0 0 127.0.0.1 InLoop0
127.0.0.1/32 Direct 0 0 127.0.0.1 InLoop0
127.255.255.255/32 Direct 0 0 127.0.0.1 InLoop0
192.168.200.8/29 BGP 255 0 100.78.242.180 GE2/0/0
192.168.200.49/32 Direct 0 0 127.0.0.1 InLoop0 //异常分部自己的私网路由
192.168.200.80/29 BGP 255 0 100.78.241.106 GE2/0/0 //其他分部的bgp私网路由
224.0.0.0/4 Direct 0 0 0.0.0.0 NULL0
224.0.0.0/24 Direct 0 0 0.0.0.0 NULL0
255.255.255.255/32 Direct 0 0 127.0.0.1 InLoop0
```

(2) 正常分部查看的vpn私网路由表中可以学习到异常分部的vpn私网路由以及总部的vpn私网路由:

```
display ip routing-table vpn-instance SD_ZhengFu_SheBao
Destinations : 15    Routes : 15
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface	
0.0.0.0/32	Direct	0	0	127.0.0.1	InLoop0	
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0	
127.0.0.0/32	Direct	0	0	127.0.0.1	InLoop0	
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0	
127.255.255.255/32	Direct	0	0	127.0.0.1	InLoop0	
192.168.200.0/30	BGP	255	0	100.78.241.95	GE2/0/0	//总部的bgp私网路由
192.168.200.8/29	BGP	255	0	100.78.242.180	GE2/0/0	
192.168.200.49/32	BGP	255	0	100.78.241.108	GE2/0/0	//异常分部的私网路由
192.168.200.80/29	Direct	0	0	192.168.200.81	GE2/0/1	//自己的私网路由
192.168.200.80/32	Direct	0	0	192.168.200.81	GE2/0/1	
192.168.200.81/32	Direct	0	0	127.0.0.1	InLoop0	
192.168.200.87/32	Direct	0	0	192.168.200.81	GE2/0/1	
224.0.0.0/4	Direct	0	0	0.0.0.0	NULL0	
224.0.0.0/24	Direct	0	0	0.0.0.0	NULL0	
255.255.255.255/32	Direct	0	0	127.0.0.1	InLoop0	

过程分析

(1) 首先进行异常分支和总部的联通性测试，经过ping测试，发现异常分支到SR8810反射器以及总部设备都是正常能ping通的，说明异常分支到总部的IGP路由是正常的：

(2) 由于现场是使用BGP反射器的方式来实现的BGP以及MP-BGP的路由传递，于是查看现场的BG P配置（对部分地址做了隐匿，100.*.241.95是总部的loopback接口地址，100.*.241.108是分部的loop back，100.*.240.34是反射器的loopback接口地址，三台设备均使用loopback接口地址建立邻居。）

A.异常分支侧配置：

```
interface LoopBack0
ip address 100.*.241.108 255.255.255.255
#
ip vpn-instance SD_ZhengFu_SheBao
route-distinguisher 65175:712
vpn-target 65175:712 import-extcommunity
vpn-target 65175:712 export-extcommunity
#
bgp 65175
non-stop-routing
peer 100.*.240.34 as-number 65175
peer 100.*.240.34 connect-interface LoopBack0
#
address-family ipv4 unicast
import-route direct
import-route static
peer 100.*.240.34 enable
#
address-family vpnv4
undo policy vpn-target
peer 100.*.240.34 enable
#
ip vpn-instance SD_ZhengFu_SheBao
#
address-family ipv4 unicast
import-route direct
#
```

B.反射器设备配置

```
interface LoopBack0
ip address 100.*.240.34 255.255.255.255
#
bgp 65175
non-stop-routing
group DLHJ-12-RR internal
peer DLHJ-12-RR connect-interface LoopBack0
group ESN-RR internal
peer ESN-RR connect-interface LoopBack0
```

```

peer ESN-RR password cipher $c$3$KvD8u31d0sGLzXKMI95uq2idwfmMrVM1zADybfTM=
group SSN-RR internal
peer SSN-RR connect-interface LoopBack0
peer SSN-RR password cipher $c$3$nArslvCBYC08mUaiOAc/o+MNVoF0a0Q48jbVAEo=
peer 100.*.241.95 group DLHJ-12-RR
peer 100.*.241.108 group DLHJ-12-RR
#
address-family ipv4 unicast
reflector cluster-id 100.*.240.34
peer DLHJ-12-RR enable
peer DLHJ-12-RR reflect-client
peer ESN-RR enable
peer SSN-RR enable
#
address-family vpnv4
reflector cluster-id 100.*.240.34
undo policy vpn-target
peer DLHJ-12-RR enable
peer DLHJ-12-RR reflect-client
peer ESN-RR enable
peer SSN-RR enable
#

```

C.总部侧配置查看如下:

```

interface LoopBack0
ip address 100.*.241.95 255.255.255.255
#
ip vpn-instance SD_ZhengFu_SheBao
route-distinguisher 65175:712
vpn-target 65175:712 import-extcommunity
vpn-target 65175:712 export-extcommunity
#
bgp 65175
non-stop-routing
peer 100.*.240.34 as-number 65175
peer 100.*.240.34 connect-interface LoopBack0
#
address-family ipv4 unicast
import-route direct
import-route static
peer 100.*.240.34 enable
#
address-family vpnv4
undo policy vpn-target
peer 100.*.240.34 enable
#
ip vpn-instance SD_ZhengFu_SheBao
#
address-family ipv4 unicast
import-route direct
import-route static
#

```

(3) 查看分支侧设备和总部侧设备的BGP邻居情况:

A.异常分支和反射器的BGP邻居查看正常:

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

BGP local router ID: 100.*.241.108

Local AS number: 65175

Total number of peers: 1 Peers in established state: 1

* - Dynamically created peer

Peer	AS	MsgRcvd	MsgSent	OutQ	PrefRcv	Up/Down	State
100.*.240.34	65175	765	14	0	1632	00:06:15	Established

B.总部侧设备和反射器的BGP邻居情况也正常:

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

```
BGP local router ID: 100.*.241.108
Local AS number: 65175
Total number of peers: 1          Peers in established state: 1
* - Dynamically created peer
Peer          AS  MsgRcvd  MsgSent  OutQ  PrefRcv  Up/Down  State
100.*.240.34  65175  54200   13397   0    1632  0185h09m  Established
```

但是仔细看异常分支侧设备和总部侧设备的BGP local router ID发现，异常分支设备和总部设备的BGP local router ID都是100.*.241.108。但是看总部设备配置中并未指定BGP的Router ID，后来让现场将总部侧设备强行指定了一个和异常分支设备不一样的BGP的Router ID后，两端私网路由发布正常。

解决方法

现场将总部侧设备强行指定了一个和异常分支设备不一样的BGP的Router ID后，两端私网路由发布正常。

但是现场有个疑问：100.*.241.108并非总部设备上配置的任何地址信息。按照缺省的Router ID竞选规则，不应该会选该数值作为Router ID。

缺省情况下router ID的竞选规则如下：

缺省情况下，未配置全局Router ID

如果未配置全局Router ID，则按照下面的规则进行选择：

1. 如果存在配置IP地址的Loopback接口，则选择Loopback接口地址中最大的作为Router ID
2. 如果所有Loopback接口都未配置IP地址，则从其他接口的IP地址中选择最大的作为Router ID（不考虑接口的up/down状态）

后来经了解，之前现场配置过BGP的Router ID刚好命中了设备的一条规则，规则内容如下：

· 如果是在BGP实例视图下配置的Router ID，则Router ID所在接口被删除时路由器不会重新选择Router ID，只有在BGP实例视图下使用undo router-id命令删除手工配置的Router ID后，路由器才会重新选择Router ID。