

知 S7500E系列交换机BGP路由聚合后部分网段无法访问的解决方法

BGP 岳斌 2012-02-02 发表

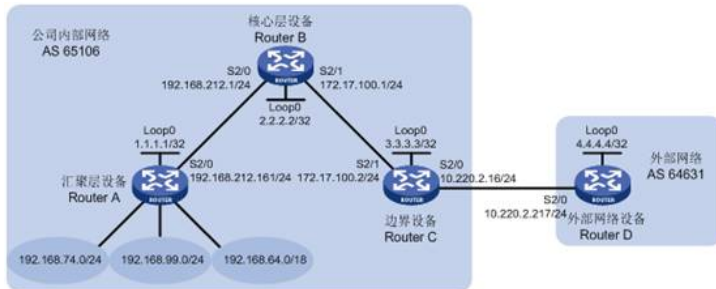
S7500E系列交换机BGP路由聚合后部分网段无法访问的解决方法

一、组网：

通过在边界设备Router C和外部网络设备Router D之间建立EBGP连接，实现公司内部网络与外部网络的互通。

在公司内部，核心层设备Router B与汇聚层设备Router A之间配置静态路由，Router B与Router C之间配置OSPF，并在OSPF路由中引入静态路由，以实现公司内部网络的互通。

公司内部网络包括三个网段：192.168.74.0/24、192.168.99.0/24和192.168.64.0/18。在Router C上配置路由聚合，将这三个网段的路由聚合为一条路由，以减少通过BGP发布的路由数量。



二、问题描述：

故障现象概述如下：

在Router C上执行如下命令配置路由聚合后，在Router D上可以ping通192.168.74.0/24和192.168.99.0/24网段内的主机，但是无法ping通192.168.64.0/18网段内的主机。

aggregate 192.168.64.0 18 detail-suppressed

具体配置及故障现象如下所述：

(1) 配置各接口的IP地址（略）

(2) 在Router A和Router B之间配置静态路由

在Router A上配置缺省路由，下一跳为Router B。

system-view

[RouterA] ip route-static 0.0.0.0 0 192.168.212.1

在Router B上配置静态路由，到达目的网络192.168.74.0/24、192.168.99.0/24和192.168.64.0/18的路由由下一跳均为Router A。

system-view

[RouterB] ip route-static 192.168.74.0 24 192.168.212.161

[RouterB] ip route-static 192.168.99.0 24 192.168.212.161

[RouterB] ip route-static 192.168.64.0 18 192.168.212.161

(3) 在Router B和Router C之间配置OSPF，并引入静态路由

在Router B上配置OSPF发布本地网段路由，并引入静态路由。

[RouterB] ospf

[RouterB-ospf-1] area 0

[RouterB-ospf-1-area-0.0.0.0] network 172.17.100.0 0.0.0.255

[RouterB-ospf-1-area-0.0.0.0] quit

[RouterB-ospf-1] import-route static

[RouterB-ospf-1] quit

在Router C上配置OSPF发布本地网段路由。

[RouterC] ospf

[RouterC-ospf-1] area 0

[RouterC-ospf-1-area-0.0.0.0] network 172.17.100.0 0.0.0.255

[RouterC-ospf-1-area-0.0.0.0] network 10.220.2.0 0.0.0.255

[RouterC-ospf-1-area-0.0.0.0] quit

[RouterC-ospf-1] quit

在Router C上查看路由表信息，可以看到Router C通过OSPF学习到了到达192.168.74.0/24、192.168.99.0/24和192.168.64.0/18网段的路由。

[RouterC] display ip routing-table

Routing Tables: Public

Destinations : 10 Routes : 10

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
3.3.3.3/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.16	S2/0
10.220.2.16/32	Direct	0	0	127.0.0.1	InLoop0

```
127.0.0.0/8      Direct 0 0      127.0.0.1      InLoop0
127.0.0.1/32     Direct 0 0      127.0.0.1      InLoop0
172.17.100.0/24   Direct 0 0      172.17.100.2    S2/1
172.17.100.2/32   Direct 0 0      127.0.0.1      InLoop0
192.168.64.0/18   O_ASE 150 1      172.17.100.1    S2/1
192.168.74.0/24   O_ASE 150 1      172.17.100.1    S2/1
192.168.99.0/24   O_ASE 150 1      172.17.100.1    S2/1
```

(4) 在Router C和Router D之间配置BGP，并引入OSPF路由

在Router C上配置Router D为其EBGP对等体，并引入OSPF路由。

```
[RouterC] bgp 65106
```

```
[RouterC-bgp] router-id 3.3.3.3
```

```
[RouterC-bgp] peer 10.220.2.217 as-number 64631
```

```
[RouterC-bgp] import-route ospf
```

在Router D上配置Router C为其EBGP对等体。

```
[RouterD] bgp 64631
```

```
[RouterD-bgp] router-id 4.4.4.4
```

```
[RouterD-bgp] peer 10.220.2.16 as-number 65106
```

```
[RouterD-bgp] quit
```

在Router D上查看路由表信息，可以看到Router D通过BGP学习到了到达192.168.74.0/24、192.168.99.0/24和192.168.64.0/18三个网段的路由。

```
[RouterD] display ip routing-table
```

Routing Tables: Public

Destinations : 8 Routes : 8

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

4.4.4.4/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.217	S2/0
10.220.2.217/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
192.168.64.0/18	BGP	255	1	10.220.2.16	S2/0
192.168.74.0/24	BGP	255	1	10.220.2.16	S2/0
192.168.99.0/24	BGP	255	1	10.220.2.16	S2/0

完成上述配置后，在Router D上可以ping通192.168.74.0/24、192.168.99.0/24和192.168.64.0/18网段内的主机。

(5) 在Router C上配置路由聚合

在Router C上将路由192.168.74.0/24、192.168.99.0/24和192.168.64.0/18聚合为192.168.64.0/18，并抑制发布具体路由。

```
[RouterC-bgp] aggregate 192.168.64.0 18 detail-suppressed
```

```
[RouterC-bgp] quit
```

在Router C上查看路由表信息，可以看到Router C上产生了一条聚合路由192.168.64.0/18，该聚合路由的出接口为Null0。

```
[RouterC] display ip routing-table
```

Routing Tables: Public

Destinations : 10 Routes : 10

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

3.3.3.3/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.16	S2/0
10.220.2.16/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
172.17.100.0/24	Direct	0	0	172.17.100.2	S2/1
172.17.100.2/32	Direct	0	0	127.0.0.1	InLoop0
192.168.64.0/18	BGP	130	0	127.0.0.1	NULL0
192.168.74.0/24	O_ASE	150	1	172.17.100.1	S2/1
192.168.99.0/24	O_ASE	150	1	172.17.100.1	S2/1

在Router D上查看路由表信息，可以看到Router D上到达公司内部三个网络的路由聚合为一条路由192.168.64.0/18。

```
[RouterD] display ip routing-table
```

Routing Tables: Public

Destinations : 6 Routes : 6

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

4.4.4.4/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.217	S2/0
10.220.2.217/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0

```
127.0.0.1/32    Direct 0   0       127.0.0.1    InLoop0
192.168.64.0/18 BGP 255 0       10.220.2.16  S2/0
```

完成上述配置后，成功实现了路由聚合。然而，在Router D上可以ping通192.168.74.0/24和192.168.99.0/24网段内的主机，但是无法ping通192.168.64.0/18网段内的主机。

三、过程分析：

在Router C上查看到达192.168.64.0/18网段路由的详细信息。可以看出存在两条到达该网段的路由，一条为BGP聚合路由，一条为通过OSPF学习到的路由。前者的优先级（130）高于后者（150），BGP聚合路由被优选。由于BGP聚合路由的出接口为Null0，因此，到达192.168.64.0/18网段的报文被丢弃，导致外部网络无法访问192.168.64.0/18网段。

```
[RouterC] display ip routing-table 192.168.64.0 verbose
```

```
Routing Table : Public
```

```
Summary Count : 2
```

```
Destination: 192.168.64.0/18
```

```
Protocol: BGP          Process ID: 0
```

```
Preference: 130        Cost: 0
```

```
IpPrecedence:          QoSLabel:
```

```
NextHop: 127.0.0.1     Interface: NULL0
```

```
BkNextHop: 0.0.0.0     BkInterface:
```

```
RelyNextHop: 0.0.0.0   Neighbor : 0.0.0.0
```

```
Tunnel ID: 0x0         Label: NULL
```

```
BKTunnel ID: 0x0       BKLabel: NULL
```

```
State: Active NoAdv    Age: 00h10m07s
```

```
Tag: 0
```

```
Destination: 192.168.64.0/18
```

```
Protocol: O_ASE        Process ID: 1
```

```
Preference: 150        Cost: 1
```

```
IpPrecedence:          QoSLabel:
```

```
NextHop: 172.17.100.1  Interface: Serial2/1
```

```
BkNextHop: 0.0.0.0     BkInterface:
```

```
RelyNextHop: 0.0.0.0   Neighbor : 0.0.0.0
```

```
Tunnel ID: 0x0         Label: NULL
```

```
BKTunnel ID: 0x0       BKLabel: NULL
```

```
State: Inactive Adv    Age: 00h16m07s
```

```
Tag: 1
```

四、解决方法：

可以采用如下方法解决上述问题：

？ 取消聚合路由的配置，通过路由策略抑制发布具体路由192.168.74.0/24和192.168.99.0/24，从而达到与路由聚合相同的效果。

？ 降低BGP聚合路由优先级。

？ 提高OSPF路由优先级。

？ 修改聚合路由使其与具体路由的目的网络地址和掩码不完全相同。

1、取消聚合路由，通过路由策略抑制发布具体路由

(1)故障处理步骤

取消聚合路由的配置。

```
[RouterC] bgp 65106
```

```
[RouterC-bgp] undo aggregate 192.168.64.0 18
```

配置通过路由策略specific-routes对发布给Router D的路由进行过滤。

```
[RouterC-bgp] peer 10.220.2.217 route-policy specific-routes export
```

```
[RouterC-bgp] quit
```

配置两个IP地址前缀列表，分别用来匹配网段192.168.74.0/24和192.168.99.0/24。

```
[RouterC] ip ip-prefix specific1 index 1 permit 192.168.74.0 24
```

```
[RouterC] ip ip-prefix specific2 index 1 permit 192.168.99.0 24
```

配置路由策略specific-routes的内容为：拒绝192.168.74.0/24和192.168.99.0/24网段通过，允许其他网段通过。

```
[RouterC] route-policy specific-routes deny node 0
```

```
[RouterC-route-policy] if-match ip-prefix specific1
```

```
[RouterC-route-policy] quit
```

```
[RouterC] route-policy specific-routes deny node 5
```

```
[RouterC-route-policy] if-match ip-prefix specific2
```

```
[RouterC-route-policy] quit
```

```
[RouterC] route-policy specific-routes permit node 10
```

(2)故障处理结果验证

在Router C上查看路由表信息，可以看到通过OSPF学习的路由为最优路由。

```
[RouterC] display ip routing-table
```

```
Routing Tables: Public
```

```
Destinations : 10    Routes : 10
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

```

3.3.3.3/32      Direct 0 0      127.0.0.1    InLoop0
10.220.2.0/24   Direct 0 0      10.220.2.16   S2/0
10.220.2.16/32  Direct 0 0      127.0.0.1    InLoop0
127.0.0.0/8     Direct 0 0      127.0.0.1    InLoop0
127.0.0.1/32    Direct 0 0      127.0.0.1    InLoop0
172.17.100.0/24 Direct 0 0      172.17.100.2  S2/1
172.17.100.2/32 Direct 0 0      127.0.0.1    InLoop0
192.168.64.0/18 O_ASE 150 1      172.17.100.1  S2/1
192.168.74.0/24 O_ASE 150 1      172.17.100.1  S2/1
192.168.99.0/24 O_ASE 150 1      172.17.100.1  S2/1

```

在Router D上查看路由表信息，可以看到只存在一条到达公司内部网络的路由192.168.64.0/18。

```

[RouterD] display ip routing-table
Routing Tables: Public
  Destinations : 6      Routes : 6

```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

4.4.4.4/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.217	S2/0
10.220.2.217/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
192.168.64.0/18	BGP	255	1	10.220.2.16	S2/0

在Router D上可以ping通192.168.74.0/24、192.168.99.0/24和192.168.64.0/18网段内的主机。

2、降低BGP聚合路由优先级

(1)故障处理步骤

在Router C上创建ACL 2010，用来匹配网段192.168.64.0/18。

```

[RouterC] acl number 2010
[RouterC-acl-basic-2010] rule permit source 192.168.64.0 0.0.63.255
[RouterC-acl-basic-2010] quit

```

在Router C上创建路由策略pre_mod，将192.168.64.0/18路由的优先级修改为200。

```

[RouterC] route-policy pre_mod permit node 0
[RouterC-route-policy] if-match acl 2010
[RouterC-route-policy] apply preference 200
[RouterC-route-policy] quit

```

根据路由策略pre_mod修改BGP路由的优先级，即192.168.64.0/18路由的优先级为200。

```

[RouterC] bgp 65106
[RouterC-bgp] preference route-policy pre_mod
[RouterC-bgp] quit

```

(2)故障处理结果验证

在Router C上查看路由表信息，可以看到通过OSPF学习的路由成为最优路由。

```

[RouterC] display ip routing-table
Routing Tables: Public
  Destinations : 10      Routes : 10

```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

3.3.3.3/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.16	S2/0
10.220.2.16/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
172.17.100.0/24	Direct	0	0	172.17.100.2	S2/1
172.17.100.2/32	Direct	0	0	127.0.0.1	InLoop0
192.168.64.0/18	O_ASE	150	1	172.17.100.1	S2/1
192.168.74.0/24	O_ASE	150	1	172.17.100.1	S2/1
192.168.99.0/24	O_ASE	150	1	172.17.100.1	S2/1

在Router D上可以ping通192.168.74.0/24、192.168.99.0/24和192.168.64.0/18网段内的主机。

3、提高OSPF路由优先级

(1)故障处理步骤

在Router C上配置OSPF外部路由的优先级为110，高于BGP聚合路由的优先级。

```

[RouterC] ospf
[RouterC-ospf-1] preference ase 110
[RouterC-ospf-1] quit

```

(2)故障处理结果验证

在Router C上查看路由表信息，可以看到通过OSPF学习的路由成为最优路由。

```

[RouterC] display ip routing-table

```

Routing Tables: Public

Destinations : 10 Routes : 10

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
3.3.3.3/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.16	S2/0
10.220.2.16/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
172.17.100.0/24	Direct	0	0	172.17.100.2	S2/1
172.17.100.2/32	Direct	0	0	127.0.0.1	InLoop0
192.168.64.0/18	O_ASE	110	1	172.17.100.1	S2/1
192.168.74.0/24	O_ASE	110	1	172.17.100.1	S2/1
192.168.99.0/24	O_ASE	110	1	172.17.100.1	S2/1

在Router D上可以ping通192.168.74.0/24、192.168.99.0/24和192.168.64.0/18网段内的主机。

4、修改聚合路由

(1)故障处理步骤

在Router C上将路由192.168.74.0/24、192.168.99.0/24和192.168.64.0/18聚合为192.168.0.0/16，并抑制发布具体路由。

```
[RouterC] bgp 65106
```

```
[RouterC-bgp] undo aggregate 192.168.64.0 18
```

```
[RouterC-bgp] aggregate 192.168.0.0 16 detail-suppressed
```

```
[RouterC-bgp] quit
```

(2)故障处理结果验证

在Router C上查看路由表信息，可以看到到达网段192.168.64.0/18的最优路由为通过OSPF学习到的路由。

```
[RouterC] display ip routing-table
```

Routing Tables: Public

Destinations : 11 Routes : 11

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
3.3.3.3/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.16	S2/0
10.220.2.16/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
172.17.100.0/24	Direct	0	0	172.17.100.2	S2/1
172.17.100.2/32	Direct	0	0	127.0.0.1	InLoop0
192.168.0.0/16	BGP	130	0	127.0.0.1	NULL0
192.168.64.0/18	O_ASE	150	1	172.17.100.1	S2/1
192.168.74.0/24	O_ASE	150	1	172.17.100.1	S2/1
192.168.99.0/24	O_ASE	150	1	172.17.100.1	S2/1

在Router D上查看路由表信息，可以看到到达公司内部三个网络的路由聚合为一条路由192.168.0.0/16。

```
[RouterD] display ip routing-table
```

Routing Tables: Public

Destinations : 6 Routes : 6

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
4.4.4.4/32	Direct	0	0	127.0.0.1	InLoop0
10.220.2.0/24	Direct	0	0	10.220.2.217	S2/0
10.220.2.217/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
192.168.0.0/16	BGP	255	0	10.220.2.16	S2/0

在Router D上可以ping通192.168.74.0/24、192.168.99.0/24和192.168.64.0/18网段内的主机。

5、建议与总结

BGP路由表中创建的聚合路由的出接口为Null0接口，聚合后可以减少向BGP对等体发布的路由数目。在使用中应注意不要使这条聚合路由成为本设备的优选路由，否则会导致报文转发失败。如果聚合路由的子网掩码长度和被聚合的某一条具体路由完全相同，且聚合路由由优先级高于具体路由，则聚合路由会成为优选路由，这种情况下需要通过修改路由优先级等方式，来确保优选的路由为具体路由。