

下一跳不匹配和BGP非活动路由技术说明

目录

[简介](#)

[非活动路由和下一跳不匹配](#)

[示例拓扑](#)

[显示输出](#)

[在BGP配置中抑制非活动路由](#)

[添加静态路由以匹配下一跳](#)

[ECMP对下一跳和非活动路由的影响](#)

简介

本文档介绍**bgp suppress-inactive**命令如何防止未安装在路由信息库(RIB)中的路由的通告；它还描述非活动路由与下一跳不匹配之间的交互。

边界网关协议(BGP)尝试将最佳路径前缀安装到RIB时，RIB会失败，但RIB会拒绝BGP路由，因为路由表中已存在管理距离更远的路由。非活动BGP路由是未安装在RIB中，但作为rib-failure安装在BGP表中的路由。

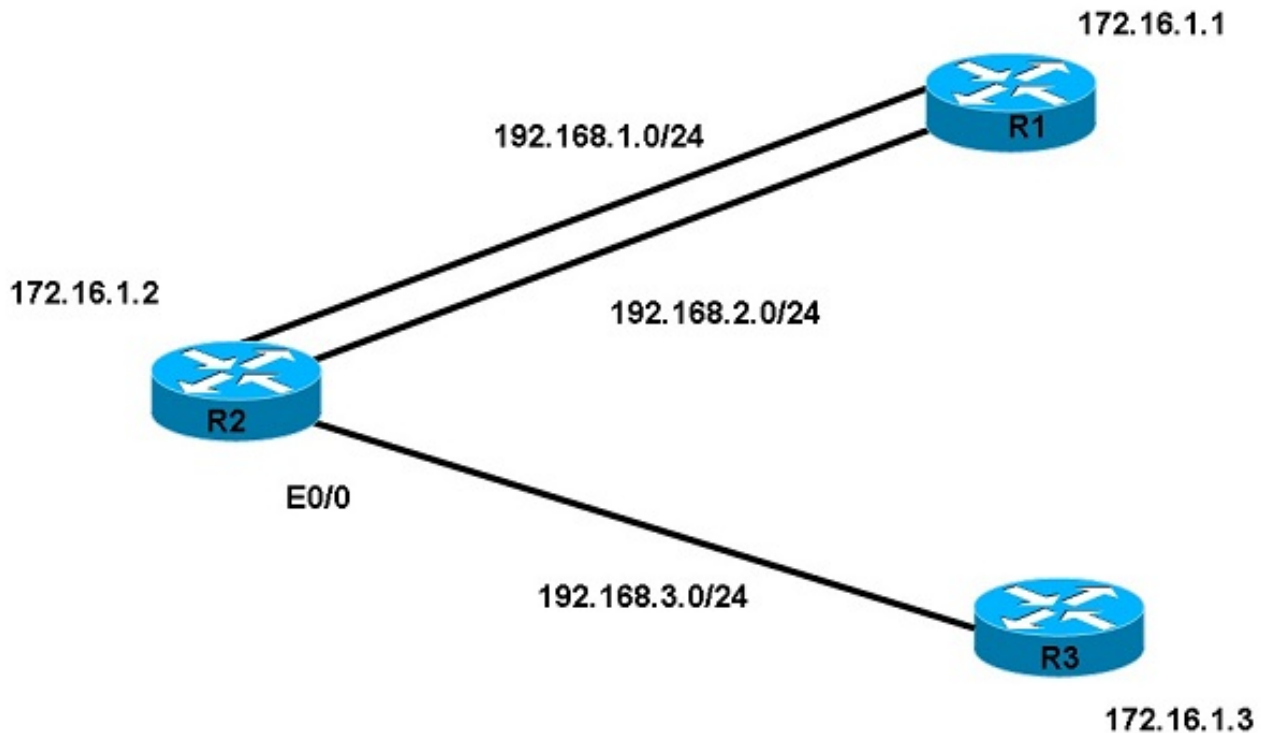
有关其他[详细信息](#)，请[参阅抑制非活动路由的BGP通告](#)。

非活动路由和下一跳不匹配

使用**bgp suppress-inactive**命令时，了解下一跳不匹配的影响至关重要。

示例拓扑

路由器1(R1)和路由器2(R2)有两条并行链路；一条链路运行BGP AS 65535，另一条链路运行增强型内部网关路由协议(EIGRP)AS 1。BGP和EIGRP都在R1上通告网络10.1.1.1/32。



R2通过EIGRP和BGP获知10.1.1.1/32路由，但由于管理距离较短，因此只将EIGRP路由安装到路由表中。由于R2路由表中未安装BGP路由，因此该路由在R2 BGP表中显示为rib-failure。但是，R2将BGP路由通告给路由器3(R3)，而不考虑rib-failure。

显示输出

对于R2，输入**show ip route** 命令以确定10.1.1.1上路由表的当前状态，并输入**show ip bgp** 命令以显示BGP路由表中的条目：

```

Router2#show ip route 10.1.1.1
Routing entry for 10.1.1.1/32
  Known via "eigrp 1", distance 90, metric 409600, type internal
  Last update from 192.168.1.1 on Ethernet0/2, 00:07:15 ago
  Routing Descriptor Blocks:
  * 192.168.1.1, from 192.168.1.1, 00:07:15 ago, via Ethernet0/2
  >>>>>>NEXT HOP IS LINK A
    Route metric is 409600, traffic share count is 1
    Total delay is 6000 microseconds, minimum bandwidth is 10000 Kbit
    Reliability 255/255, minimum MTU 1500 bytes
    Loading 1/255, Hops 1

Router2#show ip bgp
BGP table version is 4, local router ID is 172.16.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
r>i10.1.1.1/32      172.16.1.1              0     100     0    I
  
```

检查下一跳的递归路由，因为它是R1上的环回：

```
Router2#show ip route 172.16.1.1
Routing entry for 172.16.1.1/32
  Known via "eigrp 1", distance 90, metric 409600, type internal
  Last update from 192.168.2.1 on Ethernet0/1, 00:07:15 ago
  Routing Descriptor Blocks:
  * 192.168.2.1, from 192.168.2.1, 00:07:15 ago, via Ethernet0/1
  >>>>>>NEXT HOP IS LINK B
    Route metric is 409600, traffic share count is 1
    Total delay is 6000 microseconds, minimum bandwidth is 10000 Kbit
    Reliability 255/255, minimum MTU 1500 bytes
    Loading 1/255, Hops 1
```

即使下一跳不匹配，R2也会向R3通告该路由，而R3会获知该路由，因为非活动路由不会被抑制：

```
Router3#show ip bgp
BGP table version is 2, local router ID is 172.16.1.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric   LocPrf   Weight    Path
*> 10.1.1.1/32      172.16.1.2         0         0         0         I
```

在BGP配置中抑制非活动路由

输入**bgp suppress-inactive**命令以抑制非活动BGP路由。

```
Router2(config)#router bgp 65535
Router2(config-router)#bgp suppress-inactive
Router2(config-router)#end

Router2#show ip bgp neighbors 192.168.3.3 advertised-routes
Total number of prefixes 0
```

注意：仅当BGP rib-failure路由的下一跳与当前安装在路由表中的同一路由的下一跳不同时，**bgp suppress-inactive**命令才会抑制rib-failed路由。

```
Router2#show ip bgp rib-failure
Network          Next Hop          RIB-failure          RIB-NH Matches
10.1.1.1/32      172.16.1.1       Higher admin distance  No    <<<<< No match
```

在RIB-NH Matches (RIB-NH匹配) 列中，注意RIB下一跳不匹配。由于10.1.1.1/32路由的下一跳在EIGRP和BGP中不同，因此您可以使用**bgp suppress-inactive**命令抑制rib-failed路由。

换句话说，如果路由表中的下一跳与BGP下一跳匹配，则**bgp suppress-inactive**命令不再抑制。这意味着R3即使RIB发生故障，也会再次开始接收10.1.1.1/32路由。

添加静态路由以匹配下一跳

为前缀添加静态路由，以便将其RIB中的下一跳与BGP通告的下一跳匹配：

```
Router2(config)#ip route 10.1.1.1 255.255.255.255 192.168.2.1
```

```
Router2#show ip bgp rib-failure
```

```
Network      Next Hop      RIB-failure      RIB-NH Matches
10.1.1.0/24  192.168.2.1  Higher admin distance  Yes <<<< Next-Hop matches
```

即使使用**bgp suppress-inactive**命令，R2仍会通告该路由，而R3仍会接收该路由。

```
Router3#show ip bgp
```

```
BGP table version is 6, local router ID is 172.16.1.3
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
Network      Next Hop      Metric  LocPrf  Weight  Path
*> 10.1.1.0/24  172.16.1.2    0           1       i
```

总之，**bgp suppress-inactive**命令允许BGP仅在路由表中已安装具有更好管理距离的路由且该路由具有与同一路由的BGP下一跳不同的下一跳时，才禁止向邻居发送非活动路由通告。

ECMP对下一跳和非活动路由的影响

在上一个示例中，如果RIB（来自EIGRP）中安装的路由是等价多路径（ECMP），并且如果非活动路由被抑制，则您只会看到部分被抑制的路由。

在R1和R2之间的两条链路上运行EIGRP。R2从R1获知一组前缀，作为下一跳192.168.1.1和192.168.2.1之间的ECMP。例如：

```
R2#sh ip route 10.1.1.1
```

```
Routing entry for 10.1.1.1/32
```

```
Known via "eigrp 1", distance 170, metric 40030720, type internal
```

```
Last update from 192.168.1.1 on TenGigabitEthernet0/0/0, 2d02h ago
```

```
Routing Descriptor Blocks:
```

```
*192.168.1.1, from 192.168.1.1, 2d02h ago, via TenGigabitEthernet0/1/0
```

```
Route metric is 40030720, traffic share count is 1
```

```
Total delay is 1200 microseconds, minimum bandwidth is 64 Kbit
```

```
Reliability 255/255, minimum MTU 1500 bytes
```

```
Loading 32/255, Hops 2
```

```
192.168.2.1, from 192.168.2.1, 2d02h ago, via TenGigabitEthernet0/0/0
```

```
Route metric is 40030720, traffic share count is 1
```

```
Total delay is 1200 microseconds, minimum bandwidth is 64 Kbit
```

```
Reliability 255/255, minimum MTU 1500 bytes
```

```
Loading 32/255, Hops 2
```

```
R2#sh ip route 10.1.1.5
```

```
Routing entry for 10.1.1.5/32
```

```
Known via "eigrp 1", distance 170, metric 40030720, type internal
```

```
Last update from 192.168.1.1 on TenGigabitEthernet0/0/0, 2d02h ago
```

```
Routing Descriptor Blocks:
```

```
192.168.1.1, from 192.168.1.1, 2d02h ago, via TenGigabitEthernet0/1/0
```

```
Route metric is 40030720, traffic share count is 1
```

```
Total delay is 1200 microseconds, minimum bandwidth is 64 Kbit
```

```
Reliability 255/255, minimum MTU 1500 bytes
```

```
Loading 32/255, Hops 2
```

```
* 192.168.2.1, from 192.168.2.1, 2d02h ago, via TenGigabitEthernet0/0/0
```

```
Route metric is 40030720, traffic share count is 1
```

Total delay is 1200 microseconds, minimum bandwidth is 64 Kbit
Reliability 255/255, minimum MTU 1500 bytes
Loading 32/255, Hops 2

R2在BGP中从R1获取相同的前缀集，并在两条链路上获取下一跳环回。

Router2#show ip bgp

BGP table version is 4, local router ID is 172.16.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
r>i10.1.1.1/32	172.16.1.1	0	100	0	I
r>i10.2.2.2/32	172.16.1.1	0	100	0	I
r>i10.3.3.3/32	172.16.1.1	0	100	0	I
r>i10.4.4.4/32	172.16.1.1	0	100	0	I
r>i10.5.5.5/32	172.16.1.1	0	100	0	I
r>i10.6.6.6/32	172.16.1.1	0	100	0	I
r>i10.7.7.7/32	172.16.1.1	0	100	0	I
r>i10.8.8.8/32	172.16.1.1	0	100	0	I
r>i10.9.9.9/32	172.16.1.1	0	100	0	I
r>i10.10.10.10/32	172.16.1.1	0	100	0	I

R2#sh ip route 172.16.1.1

Routing entry for 172.16.1.1/32
Known via "eigrp 1", distance 170, metric 40030720 type internal
Redistributing via eigrp 109
Last update from 192.168.1.1 on TenGigabitEthernet0/0/0, 2d02h ago
Routing Descriptor Blocks:
* 192.168.1.1, from 192.168.1.1, 2d02h ago, via TenGigabitEthernet0/1/0
Route metric is 40030720, traffic share count is 1
Total delay is 1200 microseconds, minimum bandwidth is 64 Kbit
Reliability 255/255, minimum MTU 1500 bytes
Loading 32/255, Hops 2

192.168.2.1, from 192.168.2.1, 2d02h ago, via TenGigabitEthernet0/0/0
Route metric is 40030720, traffic share count is 1
Total delay is 1200 microseconds, minimum bandwidth is 64 Kbit
Reliability 255/255, minimum MTU 1500 bytes
Loading 32/255, Hops 2

由于下一跳路由是同一两条链路上的ECMP，因此您预计下一跳将匹配BGP和R2中的所有前缀，并将所有前缀通告给R3。当您查看输出的RIB-NH Matches列时，某些下一跳(NH)匹配是yes，而其他匹配是no。

Router2#sh ip bgp rib-failure

Network	Next Hop	RIB-failure	RIB-NH Matches
10.1.1.1/32	172.16.1.1	Higher admin distance	Yes
10.2.2.2/32	172.16.1.1	Higher admin distance	Yes
10.3.3.3/32	172.16.1.1	Higher admin distance	Yes
10.4.4.4/32	172.16.1.1	Higher admin distance	Yes
10.5.5.5/32	172.16.1.1	Higher admin distance	No
10.6.6.6/32	172.16.1.1	Higher admin distance	No
10.7.7.7/32	172.16.1.1	Higher admin distance	No
10.8.8.8/32	172.16.1.1	Higher admin distance	No
10.9.9.9/32	172.16.1.1	Higher admin distance	No
10.10.10.10/32	172.16.1.1	Higher admin distance	No

所有RIB-NH匹配为yes的路由都会通告给R3;其他的都被压制了。

```
R3#sh ip bgp
```

```
BGP table version is 17, local router ID is 172.16.1.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, x best-external,
f RT-Filter
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.1.1.1/32	172.16.1.2	0	2	1	i
*> 10.2.2.2/32	172.16.1.2	0	2	1	i
*> 10.3.3.3/32	172.16.1.2	0	2	1	i
*> 10.4.4.4/32	172.16.1.2	0	2	1	i

在Cisco IOS®软件中，BGP只能选择一个下一跳，并仅通过该下一跳通告最佳路径（无附加路径、多路径、BGP最佳外部或其他功能）。

当RIB为目的地安装EIGRP路由(在输出中注意*)时，RIB可能会选择其中一条路径作为最佳路径。如果该路径与BGP下一跳的路径匹配，则下一跳匹配的路径将报告为是。

在本示例中，RIB选择192.168.1.1作为10.1.1.1/32网络的下一跳(注意sh ip route 172.16.1.1输出中的*)，它与BGP下一跳的路由1匹配72.16.1.1;在下一跳匹配中，这被报告为是。RIB选择192.168.2.1作为10.1.1.5/32的下一跳，该跳与BGP下一跳的路由不匹配；这被报告为下一跳不匹配。

总之，下一跳匹配仅在抑制非活动路由时才重要；如果没有匹配项，您会在RIB-NH Matches列中看到n/a标志，R2将所有路由通告给R3。

```
Router2#sh ip bgp rib-failure
```

Network	Next Hop	RIB-failure	RIB-NH Matches
10.1.1.1/32	172.16.1.1	Higher admin distance	n/a
10.2.2.2/32	172.16.1.1	Higher admin distance	n/a
10.3.3.3/32	172.16.1.1	Higher admin distance	n/a
10.4.4.4/32	172.16.1.1	Higher admin distance	n/a
10.5.5.5/32	172.16.1.1	Higher admin distance	n/a
10.6.6.6/32	172.16.1.1	Higher admin distance	n/a
10.7.7.7/32	172.16.1.1	Higher admin distance	n/a
10.8.8.8/32	172.16.1.1	Higher admin distance	n/a
10.9.9.9/32	172.16.1.1	Higher admin distance	n/a
10.10.10.10/32	172.16.1.1	Higher admin distance	n/a