

# 技術說明：Next-Hop Mismatch and BGP Inactive Routes

## 目錄

[簡介](#)

[非活動路由和下一跳不匹配](#)

[拓撲示例](#)

[顯示輸出](#)

[抑制BGP配置中的非活動路由](#)

[新增靜態路由以匹配下一跳](#)

[ECMP對下一跳和非活動路由的意義](#)

## 簡介

本檔案將說明**bgp suppress-inactive**命令如何防止通告路由資訊庫(RIB)中未安裝的路由；還描述了非活動路由和下一跳不匹配之間的互動。

邊界閘道通訊協定(BGP)嘗試將bestpath首碼安裝到RIB中，但RIB拒絕BGP路由時，就會發生rib失敗，因為路由表中已存在具有更好管理距離的路由。非活動BGP路由是未安裝在RIB中，但作為rib-failure安裝在BGP表中的路由。

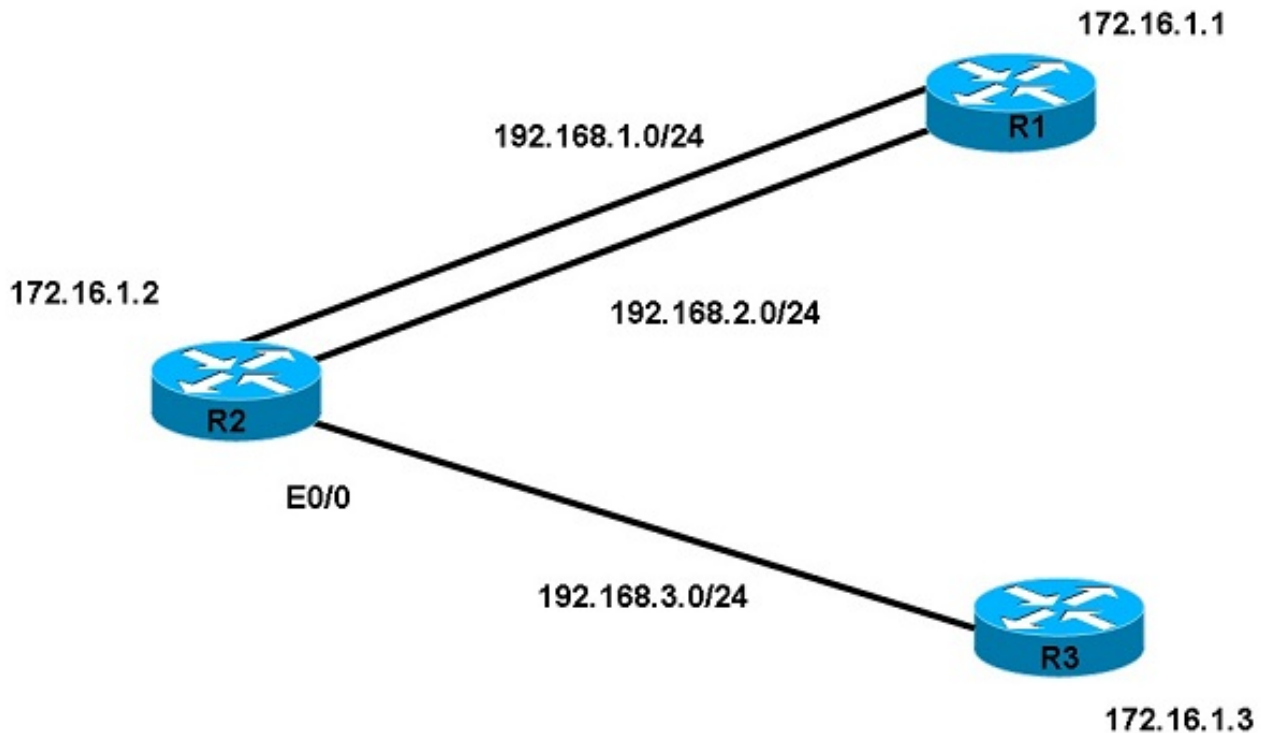
請參閱[禁止非活動路由的BGP通告](#)以瞭解其他詳細資訊。

## 非活動路由和下一跳不匹配

使用**bgp suppress-inactive**命令時，請務必瞭解下一個躍點不相符的影響。

### 拓撲示例

路由器1(R1)和路由器2(R2)有兩個並行鏈路；一條鏈路運行BGP AS 65535，另一條鏈路運行增強型內部網關路由協定(EIGRP)AS 1。BGP和EIGRP都在R1上通告網路10.1.1.1/32。



R2通過EIGRP和BGP獲知10.1.1.1/32路由，但由於管理距離較小，因此僅將EIGRP路由安裝在路由表中。由於BGP路由未安裝在R2路由表中，因此該路由在R2 BGP表中顯示為rib-failure。但是，R2會將BGP路由通告給路由器3(R3)，而不考慮rib-failure。

## 顯示輸出

對於R2，輸入**show ip route**命令以確定10.1.1.1上路由表的當前狀態，並輸入**show ip bgp**命令以顯示BGP路由表中的條目：

```

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
  
```

檢查下一躍點的遞迴路由，因為它是R1上的環回：

```
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也會將路由通告給R3，而R3會獲知該路由，因為不活動的路由不會被抑制：

```
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	I

## 抑制BGP配置中的非活動路由

輸入**bgp suppress-inactive**命令以抑制不活動的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
```

**附註：**只有BGP rib-failure路由的下一跳與路由表中當前安裝的同一路由的下一跳不同時，**bgp suppress-inactive**命令才會抑制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列中，請注意RIB下一跳不匹配。由於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。

換句話說，如果路由表中的下一跳與BGP下一跳匹配，**bgp suppress-inactive**命令將不再抑制。這意味著，即使R3的RIB發生故障，它也會再次開始接收10.1.1.1/32路由。

## 新增靜態路由以匹配下一跳

為此字首新增靜態路由，以便將其在RIB中的下一跳與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仍會通告路由，而R3仍會收到路由。

```
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
```

總而言之，只有當路由表中已安裝了具有更佳管理距離的路由，且其下一跳與BGP下一跳具有不同的相同路由時，**bgp suppress-inactive**命令才會允許BGP抑制通向鄰居的非活動路由通告。

## ECMP對下一跳和非活動路由的意義

在上一個示例中，如果安裝在RIB（來自EIGRP）中的路由是等價多路徑(ECMP)，並且抑制了非活動路由，則您只會看到被抑制的部分路由。

在R1和R2之間的兩條鏈路上運行EIGRP。R2從R1獲取一組字首作為192.168.1.1和192.168.2.1之間下一跳的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, viaTenGigabitEthernet0/0/0
  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, viaTenGigabitEthernet0/0/0
```

```
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中從R1獲取相同的字首集，並在兩條鏈路上獲取下一跳環回。

```
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, viaTenGigabitEthernet0/0/0
  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
```

由於下一跳路由是相同兩條鏈路上的ECMP，因此您會期望下一跳匹配BGP和R2中的所有字首，以便將所有字首通告給R3。當您檢視輸出的RIB-NH Matches列時，某些下一跳(NH)匹配為yes，而其它為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

RIB-NH匹配為yes的所有路由都會通告給R3;其他所有專案都遭到壓制。

```
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®軟體中，BGP只能選取一個下一躍點，並只會使用該下一躍點通告最佳路徑（沒有新增路徑、多重路徑、BGP最佳外部或其他功能）。

當RIB為目標安裝EIGRP路由時(注意輸出中的\*),RIB可能會選擇一個路徑作為最佳路徑。如果該路徑與BGP下一躍點的路徑相符，則下一個躍點配對將報告為yes。

在本示例中，RIB將192.168.1.1選為10.1.1.1/32網路的下一跳(注意sh ip route 172.16.1.1輸出中192.168.1.1上的\*)，該下一跳與BGP下一跳的路由172.16.1.1匹配；在下一跳匹配中將此報告為yes。RIB選擇192.168.2.1作為10.1.1.5/32的下一跳，它與BGP下一跳的路由不匹配；在下一個躍點不匹配中將其報告為no。

總而言之，僅當取消非活動路由時，下一跳匹配才重要；如果沒有匹配項，您會在RIB-NH匹配列中看到一個n/a標誌，並且R2會向R3通告所有路由。

```
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