

# 抖動BGP路由故障排除 ( 遞迴路由故障 )

## 目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[背景資訊](#)

[慣例](#)

[問題](#)

[症狀](#)

[遞迴路由失敗](#)

[遞迴路由失敗的原因](#)

[解決方案](#)

[路由抑制](#)

[相關資訊](#)

## 簡介

本文說明如何解決由遞迴路由失敗引起的擺動邊界網關協定(BGP)路由問題。

## 必要條件

### 需求

本文件沒有特定需求。

### 採用元件

本文件所述內容不限於特定軟體和硬體版本。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除 ( 預設 ) 的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

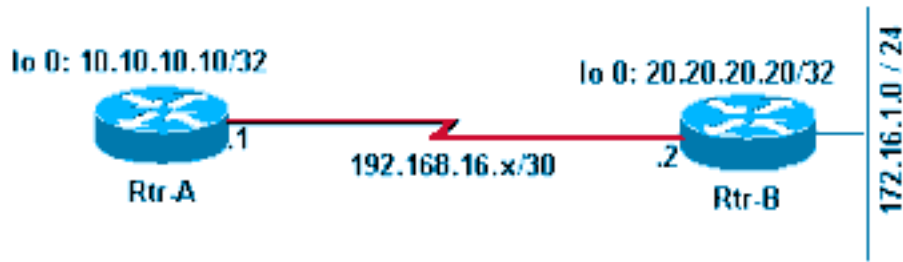
## 背景資訊

本文說明如何解決由遞迴路由失敗引起的擺動邊界網關協定(BGP)路由問題。

BGP中遞迴路由故障的常見症狀包括：

- 不斷將BGP路由刪除並重新插入路由表。
- 無法連線到通過BGP獲知的目標。

使用本檔案時請參閱以下網路圖表：



網路圖表

使用本檔案時請參閱以下設定：

## 路由器A

```
hostname RTR-A
!
interface Loopback0
 ip address 10.10.10.10 255.255.255.255
!
interface Serial8/0
 ip address 192.168.16.1 255.255.255.252
!
router bgp 1
 bgp log-neighbor-changes
 neighbor 10.20.20.20 remote-as 2
 neighbor 10.20.20.20 ebgp-multihop 2
 neighbor 10.20.20.20 update-source Loopback0
!
ip route 10.20.20.0 255.255.255.0 192.168.16.2
```

## Rtr-B

```
hostname RTR-B
!
interface Loopback0
 ip address 10.20.20.20 255.255.255.255
!
interface Ethernet0/0
 ip address 172.16.1.1 255.255.255.0
!
interface Serial8/0
 ip address 192.168.16.2 255.255.255.252
!
router bgp 2
 no synchronization
 bgp log-neighbor-changes
 network 10.20.20.20 mask 255.255.255.255
 network 172.16.1.0 mask 255.255.255.0
 neighbor 10.10.10.10 remote-as 1
 neighbor 10.10.10.10 ebgp-multihop 2
 neighbor 10.10.10.10 update-source Loopback0
 no auto-summary
!
ip route 10.10.10.0 255.255.255.0 192.168.16.1
!
```

## 慣例

如需文件慣例的詳細資訊，請參閱思科技術提示慣例。

## 問題

### 症狀

遞迴路由失敗時會出現以下兩個症狀：

- IP路由表中BGP獲知的路由的持續抖動。連續觀察路由表幾分鐘，以檢視其擺動。

```
RTR-A#show ip route
```

```
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - ISIS level-1, L2 - ISIS level-2, ia - ISIS inter are
       * - candidate default, U - per-user static route, o - ODR
       P - periodic downloaded static route
```

```
Gateway of last resort is not set
```

```
10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B    10.20.20.20/32 [20/0] via 10.20.20.20, 00:00:35
S    10.20.20.0/24 [1/0] via 192.168.16.2
172.16.0.0/24 is subnetted, 1 subnets
B    172.16.1.0 [20/0] via 10.20.20.20, 00:00:35
10.0.0.0/32 is subnetted, 1 subnets
C    10.10.10.10 is directly connected, Loopback0
192.168.16.0/30 is subnetted, 1 subnets
C    192.168.16.0 is directly connected, Serial8/0
```

**註：**使用**show ip route | include , 00:00** 命令以在處理大型路由表時觀察擺動路由。

等待大約一分鐘後，**show ip route** 命令結果會變為以下值：

```
RTR-A#show ip route
```

```
[...]
```

```
Gateway of last resort is not set
```

```
10.0.0.0/24 is subnetted, 1 subnets
S    10.20.20.0 [1/0] via 192.168.16.2
10.0.0.0/32 is subnetted, 1 subnets
C    10.10.10.10 is directly connected, Loopback0
192.168.16.0/30 is subnetted, 1 subnets
C    192.168.16.0 is directly connected, Serial8/0
```

**注意：**先前的路由表中缺少BGP路由。

- 當BGP路由存在於路由表中時，與這些網路的連線將失敗。為了觀察此情況，當Rtr-A的路由表在其路由表中具有BGP獲知的路由172.16.1.0/24時，對有效主機172.16.1.1執行ping操作失敗

。

```
RTR-A#show ip route 172.16.1.0
```

```
Routing entry for 172.16.1.0/24
```

```
Known via "bgp 1", distance 20, metric 0
```

```
Tag 2, type external
Last update from 10.20.20.20 00:00:16 ago
Routing Descriptor Blocks:
* 10.20.20.20, from 10.20.20.20, 00:00:16 ago
  Route metric is 0, traffic share count is 1
  AS Hops 1
```

```
RTR-A#ping 172.16.1.1
```

```
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
RTR-A#
```

## 遞迴路由失敗

在Rtr-A上，觀察指向BGP對等體10.20.20.20的路由。每分鐘左右兩個下一跳之間的路由會一致地跳動。

```
RTR-A#show ip route 10.20.20.20
Routing entry for 10.20.20.20/32
  Known via "bgp 1", distance 20, metric 0
  Tag 2, type external
  Last update from 10.20.20.20 00:00:35 ago
  Routing Descriptor Blocks:
  * 10.20.20.20, from 10.20.20.20, 00:00:35 ago
    Route metric is 0, traffic share count is 1
    AS Hops 1
```

通向BGP對等IP地址的路由是通過BGP本身獲取的；因此它會建立遞迴路由故障。

大約一分鐘後，路由將更改為：

```
RTR-A#show ip route 10.20.20.20
Routing entry for 10.20.20.0/24
  Known via "static", distance 1, metric 0
  Routing Descriptor Blocks:
  * 192.168.16.2
    Route metric is 0, traffic share count is 1
```

## 遞迴路由失敗的原因

以下步驟說明遞迴路由失敗的原因：

1. 請參閱[Rtr-A](#)的配置。在此配置中，靜態路由10.20.20.0/24配置為指向直連的下一跳192.168.16.2。透過此靜態路由，與對等Rtr-B 10.20.20.20建立BGP作業階段。
2. Rtr-B宣佈BGP路由172.16.1.0/24和10.20.20.20/32到Rtr-A，其環回IP地址10.20.20作為下一跳。
3. Rtr-A接收Rtr-B通告的BGP路由，並嘗試安裝10.20.20.20/32。這比已在Rtr-A中配置為靜態路由的10.20.20.0/24更為具體。因為首選最長的匹配路由，所以優先使用10.20.20.20/32而不是10.20.20.0/24。如需詳細資訊，請參閱[Cisco路由器中的路由選擇](#)。在路由表中，安裝的路由10.20.20.20/32的下一跳為10.20.20.20（Rtr-B對等IP地址）。這會導致遞迴路由失敗，因為通向10.20.20.20/32的路由本身具有下一躍點。為了瞭解在這種特定情況下遞迴路由失敗的原因，您需要瞭解路由演算法的工作原理。對於路由表中下一跳IP地址不是路由器直連介面的任何

非直連路由，該演算法會遞迴查詢路由表，直到找到可以轉發資料包的直連介面。在此特定情況下，Rtr-A獲知通往非直連網路10.20.20.20/32的路由，其非直連下一跳為10.20.20.20（其自身）。路由演算法會遇到遞迴路由環路故障，因為它找不到任何直接連線的介面，無法向其傳送目的地為10.20.20.20/32的資料包。

4. 路由器檢測到該非直連路由10.20.20.20/32存在遞迴路由故障，並從路由表中撤消10.20.20.20/32。因此，下一跳IP地址為10.20.20.20的所有BGP獲知的路由也會從路由表中撤消。
5. 整個過程從[步驟1重複](#)。如果您發出debug ip routing 指令，就可以確認這點。注意：在運行任何 debug 命令之前，請針對特定網路的訪問控制清單(ACL)運行debug 命令，以限制debug的輸出。在此範例中，設定ACL以限制偵錯輸出。

```
RTR-A(config)#access-list 1 permit 10.20.20.20
RTR-A(config)#access-list 1 permit 172.16.1.0
RTR-A(config)#end
```

```
RTR-A#debug ip routing 1
IP routing debugging is on for access list 1
```

```
00:29:50: RT: add 10.20.20.20/32 via 10.20.20.20, bgp metric [20/0]
00:29:50: RT: add 172.16.1.0/24 via 10.20.20.20, bgp metric [20/0]
00:30:45: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:45: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:45: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:46: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:46: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:48: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:48: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:50: RT: del 10.20.20.20/32 via 10.20.20.20, bgp metric [20/0]
00:30:50: RT: delete subnet route to 10.20.20.20/32
00:30:50: RT: del 172.16.1.0/24 via 10.20.20.20, bgp metric [20/0]
00:30:50: RT: delete subnet route to 172.16.1.0/24
```

6. 如果路由遞迴連續失敗，則出現以下錯誤消息：

```
%COMMON_FIB-SP-6-FIB_RECURSION: 10.71.124.25/32 has too many (8) levels of
recursion during setting up switching info
%COMMON_FIB-SP-STDBY-6-FIB_RECURSION: 10.71.124.25/32 has too many (8)
levels of recursion during setting up switching info
```

這是因為啟用MPLS的網路上發生了TCP重新傳輸。如果由於傳輸連結關閉，BGP keepalive訊息一次無法傳送到BGP對等點，則相鄰BGP對等點不會接受任何進一步的keepalive封包，即使TCP透過備份路徑重新傳輸失敗訊息，最終會導致具有保持時間到期的BGP對等點關閉。只有在Catalyst6500或Cisco7600上配置了MPLS時，才會出現此問題。此資訊包含在Cisco錯誤ID [CSCsj89544](#)中。附註：只有註冊思科使用者才能存取內部錯誤資訊和其他工具。

## 解決方案

下面詳細介紹此問題的解決方案。

在Rtr-A中為BGP對等體IP地址（本例中為10.20.20.20）新增特定靜態路由。

```
RTR-A#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
RTR-A(config)#ip route 10.20.20.20 255.255.255.255 192.168.16.2
```

為字首10.20.20.20/32配置靜態路由可確保不會將動態學習的BGP路由10.20.20.20/32安裝到路由表

中，從而避免遞迴路由環路情況。如需詳細資訊，請參閱[Cisco路由器中的路由選擇](#)。

**注意：**當EBGP對等體配置為使用預設路由相互訪問時，不會顯示BGP鄰居關係。這樣做是為了避免路由擺動和路由環路。

對172.16.1.1執行ping操作可確認解決方案。

```
RTR-A#ping 172.16.1.1
```

```
Type escape sequence to abort.
```

```
Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds:
```

```
!!!!
```

```
Success rate is 100 percent (5/5), round-trip min/avg/max = 20/24/40 ms
```

## 路由抑制

路由抑制是一項BGP功能，旨在最小化抖動路由在網際網路中的傳播。ISP建議的值是Cisco IOS®的預設值，您只需配置此命令即可啟用它。

```
router bgp <AS number>  
  bgp dampening
```

The `bgp dampening` command sets default values for the dampening parameters，例如Halftime= 15分鐘，reuse = 750, Suppress = 2000和Max Suppress Time= 60。這些值可由使用者配置，但思科建議保持其不變。

## 相關資訊

- [「#%BGP-3-INSUFCHUNKS: Insufficient chunk pools for aspath」錯誤消息表示什麼意思？](#)
- [為什麼BGP鄰居在空閒、連線和活動狀態之間切換？](#)
- [IP路由/BGP支援](#)
- [思科技術支援與下載](#)

## 關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。