

# 瞭解Nexus NX-OS與Cisco IOS上iBGP通告中的下一跳設定

## 目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[背景資訊](#)

[瞭解iBGP通告](#)

[Nexus NX-OS案例](#)

[Cisco IOS案例](#)

[使用set ip next-hop redist-unchanged命令](#)

[初始裝置配置](#)

## 簡介

本文描述在Nexus NX-OS與基於Cisco IOS ( 包括Cisco IOS-XE ) 的平台上為內部邊界網關協定 (iBGP)通告設定時NEXT\_HOP路徑屬性的行為。這用於非本地路由的通告。

## 必要條件

### 需求

思科建議您瞭解以下主題：

- 邊界閘道通訊協定(BGP)
- 路由協定重分發

### 採用元件

本檔案所述內容不限於特定軟體和硬體版本：

- 執行NX-OS版本7.3(0)D1(1)的Nexus 7000
- 執行Cisco IOS版本15.6(2)T的Cisco路由器

本文中的輸出來自特定實驗室環境中的裝置。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

- 在基於Nexus NX-OS的平台上，對於非本地發起的路由，iBGP通告會修改NEXT\_HOP屬性並將其設定為自己的本地介面IP地址。
- 在基於Cisco IOS的平台上，對於非本地發起的路由，iBGP通告將保留原始路由的NEXT\_HOP屬性。

由於缺陷[CSCud20941](#)引入的代碼更改，Nexus NX-OS上的行為可以與Cisco IOS上的行為匹配（如果需要）。

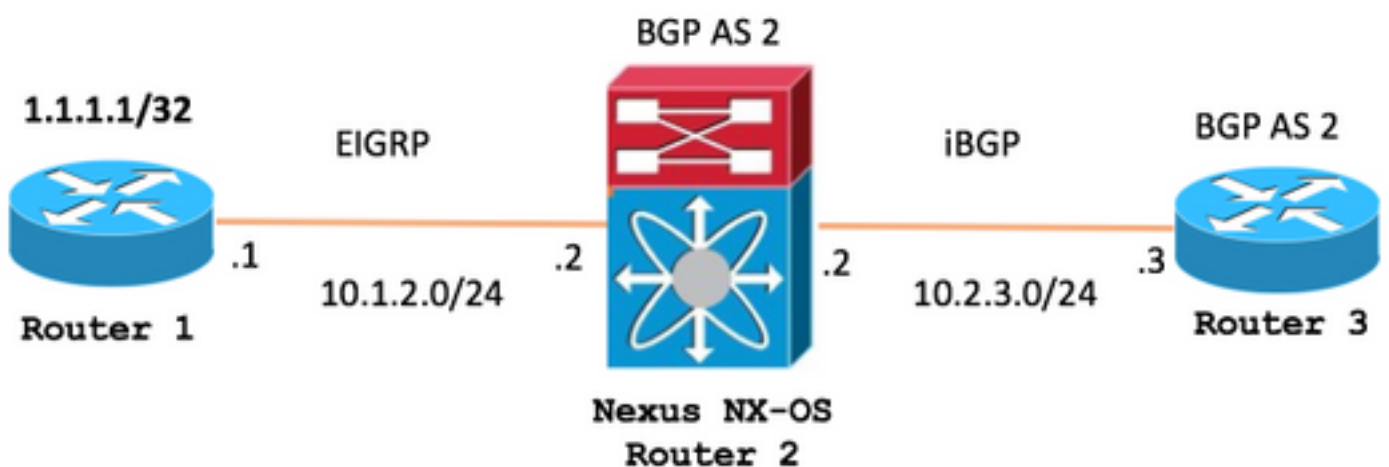
**注意：**這隻適用於iBGP通告，而不適用於eBGP。

**附註：**適用於配置為靜態路由或通過任何內部網關協定(IGP)(如增強型內部網關路由協定(EIGRP)、開放最短路徑優先(OSPF)或路由資訊協定(RIP))接收的非本地路由。

## 瞭解iBGP通告

為了瞭解iBGP通告中的NEXT\_HOP集，請以圖中所示的網路拓撲圖為例。

Nexus NX-OS機箱的拓撲

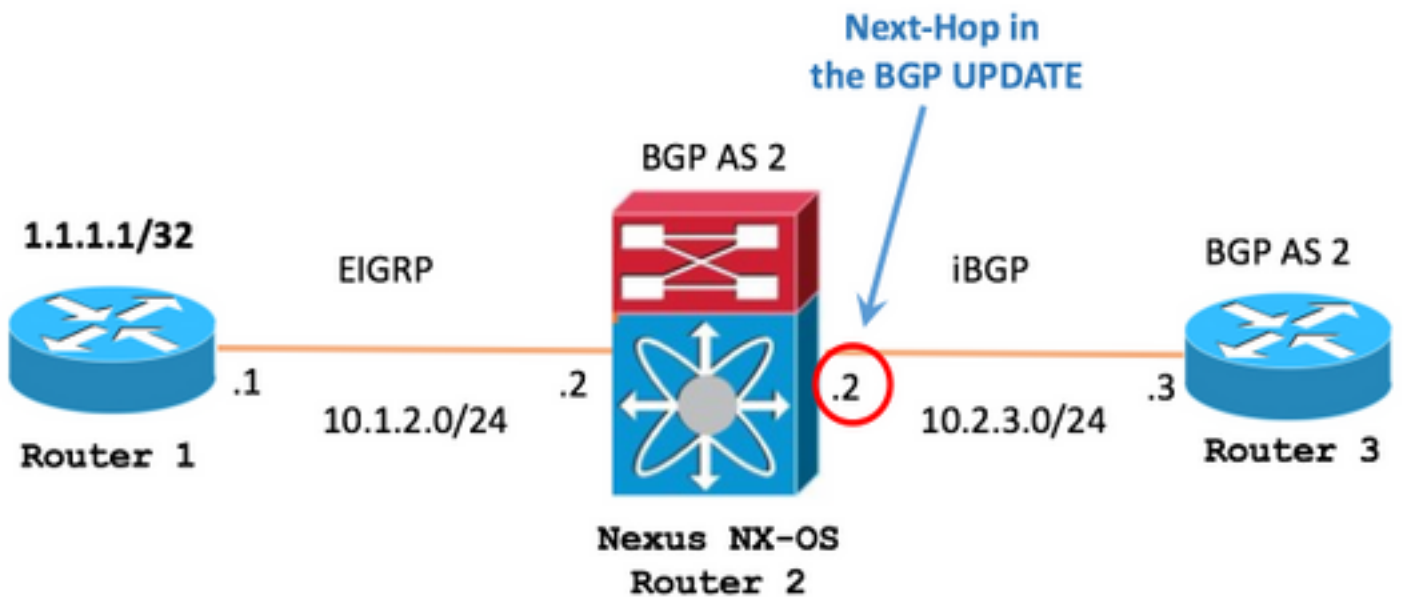


Cisco IOS案例的拓撲



## Nexus NX-OS案例

在Nexus NX-OS拓撲中，R2(Nexus NX-OS)通過EIGRP從路由器1接收1.1.1.1/32路由，並使用iBGP將其通告給路由器3，如下圖所示。



R2(Nexus NX-OS)路由表顯示通過EIGRP接收的路由1.1.1.1/32，其原始下一跳IP為10.1.2.1

### R2(Nexus NX-OS)

```
R2# show ip route 1.1.1.1/32
IP Route Table for VRF "default"
 '*' denotes best ucast next-hop
 '**' denotes best mcast next-hop
 '[x/y]' denotes [preference/metric]
 '%<string>' in via output denotes VRF <string>

1.1.1.1/32, ubest/mbest: 1/0
   *via 10.1.2.1, Eth2/1, [90/130816], 00:02:28, eigrp-1, internal
```

在BGP配置部分中，您可以看到通過iBGP將1.1.1.1/32通告給路由器3的命令。

## R2(Nexus NX-OS)

```
R2# show running-config bgp

!Command: show running-config bgp
!Time: -

version -
feature bgp

router bgp 2
  address-family ipv4 unicast
    network 1.1.1.1/32
  neighbor 10.2.3.3 remote-as 2
  address-family ipv4 unicast
```

在Router 3上，1.1.1.1/32路由是透過iBGP接收的，且下一個躍點現在已設定為R2(Nexus NX-OS)的IP位址(10.2.3.2)

— 路由器3 1.1.1.1/32的BGP表條目

## R3

```
R3# show bgp ipv4 unicast 1.1.1.1/32
BGP routing table entry for 1.1.1.1/32, version 8
Paths: (1 available, best #1, table default)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.2.3.2 from 10.2.3.2 (2.2.2.2)
      Origin IGP, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```

- 1.1.1.1/32的Router 3路由表條目

## R3

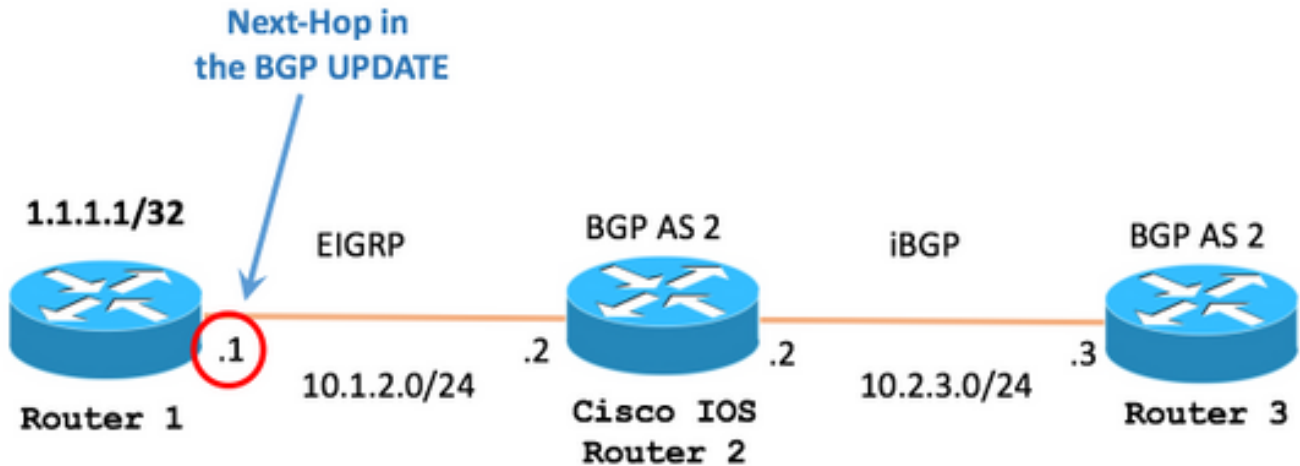
```
R3# show ip route bgp
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
       a - application route
       + - replicated route, % - next hop override, p - overrides from PfR

Gateway of last resort is not set

    1.0.0.0/32 is subnetted, 1 subnets
B       1.1.1.1 [200/0] via 10.2.3.2, 00:07:17
```

## Cisco IOS案例

在Cisco IOS案例的拓撲中，R2(Cisco IOS)通過EIGRP從路由器1接收1.1.1.1/32路由，並使用iBGP將其通告給路由器3，如下圖所示。



R2(Cisco IOS)路由表顯示了通過EIGRP接收的路由1.1.1.1/32，其原始下一跳IP為10.1.2.1

### R2(Cisco IOS)

```
R2# show ip route 1.1.1.1 255.255.255.255 longer-prefixes
Codes: L - local, C - connected, S - static, 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
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
       a - application route
       + - replicated route, % - next hop override, p - overrides from PfR
```

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

```
1.0.0.0/32 is subnetted, 1 subnets
D      1.1.1.1 [90/130816] via 10.1.2.1, 00:00:06, GigabitEthernet0/1
```

在BGP配置部分，您可以看到通過iBGP將1.1.1.1/32通告給路由器3的已設定命令

### R2(Cisco IOS)

```
R2# show running-config partition router bgp 2
Building configuration...
```

```
Current configuration : 210 bytes
!
! Last configuration change at -
!
!
```

```
!  
router bgp 2  
  bgp router-id 2.2.2.2  
  bgp log-neighbor-changes  
network 1.1.1.1 mask 255.255.255.255  
  neighbor 10.2.3.3 remote-as 2  
!  
!  
end
```

在Router 3上，您可以看到通過iBGP收到的1.1.1.1/32路由，其原始下一跳設定為路由器1上的IP(10.1.2.1)。

— 路由器3 1.1.1.1/32的BGP表條目

### R3

```
R3# show bgp ipv4 unicast 1.1.1.1/32  
BGP routing table entry for 1.1.1.1/32, version 0  
Paths: (1 available, no best path)  
  Not advertised to any peer  
  Refresh Epoch 1  
  Local  
    10.1.2.1 (inaccessible) from 10.2.3.2 (2.2.2.2)  
      Origin IGP, metric 130816, localpref 100, valid, internal  
      rx pathid: 0, tx pathid: 0
```

在此特定案例中，路由器3必須具有通往10.1.2.1（下一躍點）的路徑，因此BGP會認為路徑有效。否則，BGP會將路徑顯示為（無法存取）。

**附註：**這是在[BGP最佳路徑選取演演算法](#)中說明的基本檢查，目的是接受從BGP到路由表中的路由。

**debug ip bgp update**命令顯示Router 3不安裝路由的原因是，其路由表中沒有用於下一躍點的條目，在這種情況下，下一躍點是10.1.2.1

### R3

```
R3# debug ip bgp update  
*-*: BGP(0): 10.2.3.2 rcvd UPDATE w/ attr: nexthop 10.1.2.1, origin i, localpref 100, metric 130816  
*-*: BGP(0): 10.2.3.2 rcvd 1.1.1.1/32  
*-*: BGP(0): no valid path for 1.1.1.1/32
```

使下一跳可以訪問的方法之一是：

— 步驟1.在Router 3的路由表中配置一條靜態路由，以便為下一跳建立條目。

## R3

```
R3# configure terminal  
Enter configuration commands, one per line. End with CNTL/Z.  
R3(config)# ip route 10.1.2.1 255.255.255.255 10.2.3.2
```

— 步驟2.相同的debug命令顯示現在已接受路由。

## R3

```
R3# debug ip bgp update  
R3#  
*Mar 29 16:08:42.888: BGP(0): 10.2.3.2 rcvd UPDATE w/ attr: nexthop 10.1.2.1, origin i, localpref 100, 130816  
*Mar 29 16:08:42.890: BGP(0): 10.2.3.2 rcvd 1.1.1.1/32  
*Mar 29 16:08:42.892: BGP(0): Revise route installing 1 of 1 routes for 1.1.1.1/32 -> 10.1.2.1(global) main IP table  
R3#
```

— 步驟3. BGP表已刪除 ( 無法訪問 ) 狀態。

## R3

```
R3# show bgp ipv4 unicast 1.1.1.1/32  
BGP routing table entry for 1.1.1.1/32, version 6  
Paths: (1 available, best #1, table default)  
Not advertised to any peer  
Refresh Epoch 2  
Local  
10.1.2.1 from 10.2.3.2 (2.2.2.2)  
Origin IGP, metric 130816, localpref 100, valid, internal, best  
rx pathid: 0, tx pathid: 0x0
```

— 步驟4.路由表現在安裝到1.1.1.1/32的路由

## R3

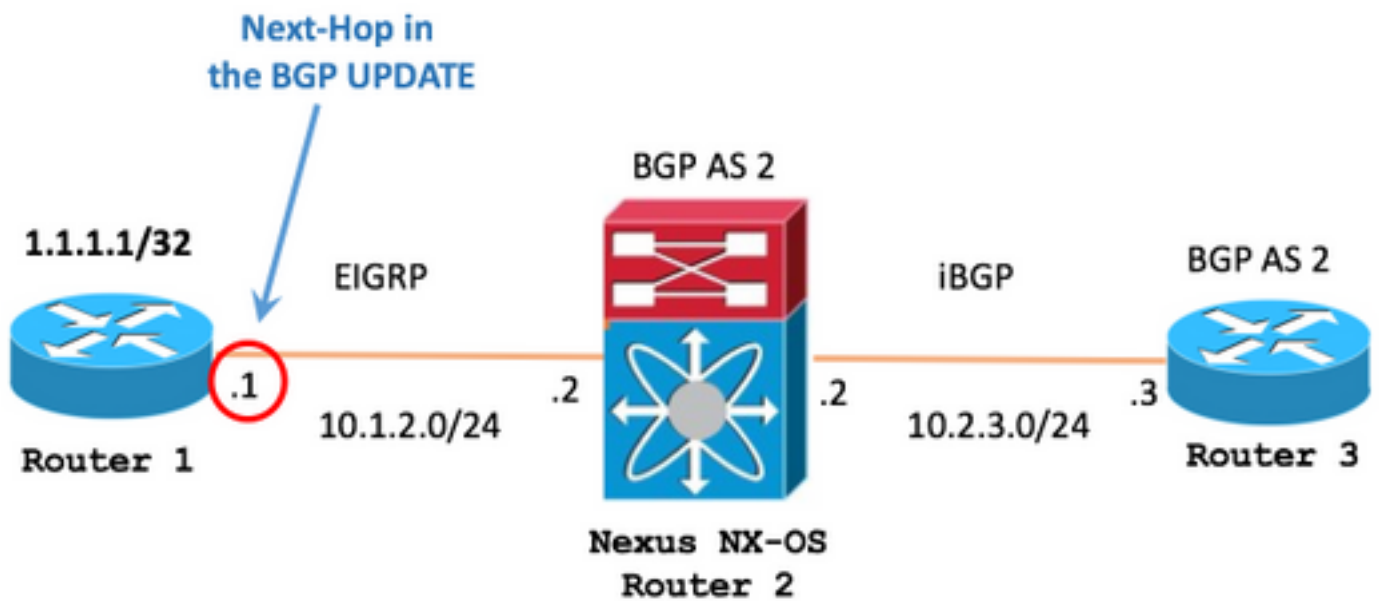
```
R3# show ip route bgp  
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP  
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area  
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2  
E1 - OSPF external type 1, E2 - OSPF external type 2  
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2  
ia - IS-IS inter area, * - candidate default, U - per-user static route  
o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP  
a - application route  
+ - replicated route, % - next hop override, p - overrides from PfR  
  
Gateway of last resort is not set  
  
1.0.0.0/32 is subnetted, 1 subnets  
B 1.1.1.1 [200/130816] via 10.1.2.1, 00:11:37
```

# 使用set ip next-hop redist-unchanged命令

自版本6.2(12)起，set ip next-hop redist-unchanged和set ipv6 next-hop redist-unchanged命令由缺陷CSCud20941引入，以使Nexus NX-OS映象Cisco IOS的行為。

**附註：**這些命令僅在用作路由對映中的引數時才起作用，並與redistribution命令結合使用。

在Nexus NX-OS的拓撲中，R2(Nexus NX-OS)通過EIGRP從路由器1接收1.1.1.1/32路由，並使用iBGP將其通告給路由器3，如下圖所示：



R2(Nexus NX-OS)路由表顯示通過EIGRP接收的路由1.1.1.1/32，其原始下一跳IP為10.1.2.1

## R2(Nexus NX-OS)

```
R2# show ip route 1.1.1.1/32
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'***' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

1.1.1.1/32, ubest/mbest: 1/0
  *via 10.1.2.1, Eth2/1, [90/130816], 04:38:21, eigrp-1, internal
```

命令set ip next-hop redist-unchanged在「route-map」配置模式下可用。



## R2(Nexus NX-OS)

```
R2(config)# route-map REDIST-UNCHANGED
R2(config-route-map)# set ip next-hop ?
A.B.C.D          IP address of next hop
load-share       Enables load sharing
peer-address     Use peer address (for BGP only)
redist-unchanged Use unchanged address during redistribution (for BGP
                  session only)
unchanged        Use unchanged address (for eBGP session only)
verify-availability Verify the reachability of the tracked object

R2(config-route-map)# set ip next-hop redist-unchanged
```

**route-map REDIST-UNCHANGED**應用為BGP中**redistribute**配置語句的引數。

## R2(Nexus NX-OS)

```
R2#
!
route-map REDIST-UNCHANGED permit 10 set ip next-hop redist-unchanged !

R2# show running-config bgp

!Command: show running-config bgp
!Time: -

version -
feature bgp

router bgp 2
  address-family ipv4 unicast
    redistribute eigrp 1 route-map REDIST-UNCHANGED
  neighbor 10.2.3.3 remote-as 2
  address-family ipv4 unicast
```

現在，Router3收到具有類似Cisco IOS的原始NEXT\_HOP設定的BGP更新。

## R3

```
R3# show ip bgp
BGP table version is 15, local router ID is 10.2.3.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop           Metric LocPrf Weight Path
* i 1.1.1.1/32      10.1.2.1          130816    100      0 ?
```

本文檔介紹Nexus NX-OS和Cisco IOS處理非本地生成的路由的iBGP通告的不同之處。

本文檔中描述的行為適用於大多數情形，並不影響通常的網路路由操作。

可在Nexus NX-OS上使用可選命令**set ip next-hop redist-unchanged**和**set ipv6 next-hop redist-unchanged**來維護BGP路由與RFC 4271的兼容性

## 初始裝置配置

### R1

```
hostname R1
!
interface Loopback0
 ip address 1.1.1.1 255.255.255.255
 ip ospf 1 area 0
!
interface GigabitEthernet0/1
 ip address 10.1.2.1 255.255.255.0
 ip ospf network point-to-point
 ip ospf 1 area 0
!
router ospf 1
!
```

### R2(Nexus NX-OS)

```
hostname R2
!
feature ospf
feature bgp
!
interface Ethernet2/1
 no switchport
 ip address 10.1.2.2/24
 ip ospf network point-to-point
 ip router ospf 1 area 0.0.0.0
 no shutdown
!
interface Ethernet2/2
 no switchport
 ip address 10.2.3.2/24
 no shutdown
!
router ospf 1
!
router bgp 2
 address-family ipv4 unicast
  network 1.1.1.1/32
 neighbor 10.2.3.3 remote-as 2
 address-family ipv4 unicast
!
```

### R2(Cisco IOS)

```
hostname R2
!
interface GigabitEthernet0/1
 ip address 10.1.2.2 255.255.255.0
 ip ospf network point-to-point
 ip ospf 1 area 0
!
interface GigabitEthernet0/2
 ip address 10.2.3.2 255.255.255.0
!
router ospf 1
!
router bgp 2
 bgp log-neighbor-changes
 network 1.1.1.1 mask 255.255.255.255
 neighbor 10.2.3.3 remote-as 2
!
```

### R3

```
hostname R3
!
interface GigabitEthernet0/1
 ip address 10.2.3.3 255.255.255.0
!
router bgp 2
 bgp log-neighbor-changes
 neighbor 10.2.3.2 remote-as 2
!
```