

了解Nexus NX-OS与Cisco IOS上iBGP通告中的下一跳集

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[背景信息](#)

[了解iBGP通告](#)

[Nexus NX-OS案例](#)

[Cisco IOS案例](#)

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

[初始设备配置](#)

简介

本文档介绍在Nexus NX-OS与基于Cisco IOS (包括Cisco IOS-XE) 的平台上为内部边界网关协议 (iBGP)通告设置时NEXT_HOP路径属性的行为。这用于通告非本地发起的路由。

先决条件

要求

Cisco 建议您了解以下主题：

- 边界网关协议 (BGP)
- 路由协议重分发

使用的组件

本文档不限于特定的软件和硬件版本：

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

本文档中的输出来自特定实验室环境中的设备。用于本文的所有设备从已清除的配置开始。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

背景信息

- 在基于Nexus NX-OS的平台上，对于非本地发起的路由，iBGP通告修改NEXT_HOP属性，并使用自己的本地接口IP地址设置该属性。
- 在基于Cisco IOS的平台上，对于非本地发起的路由，iBGP通告保留原始路由的NEXT_HOP属性。

Nexus NX-OS上的行为可以与Cisco IOS上的行为（如果需要）匹配，这要归功于缺陷CSCud20941引入的[代码更改](#)。

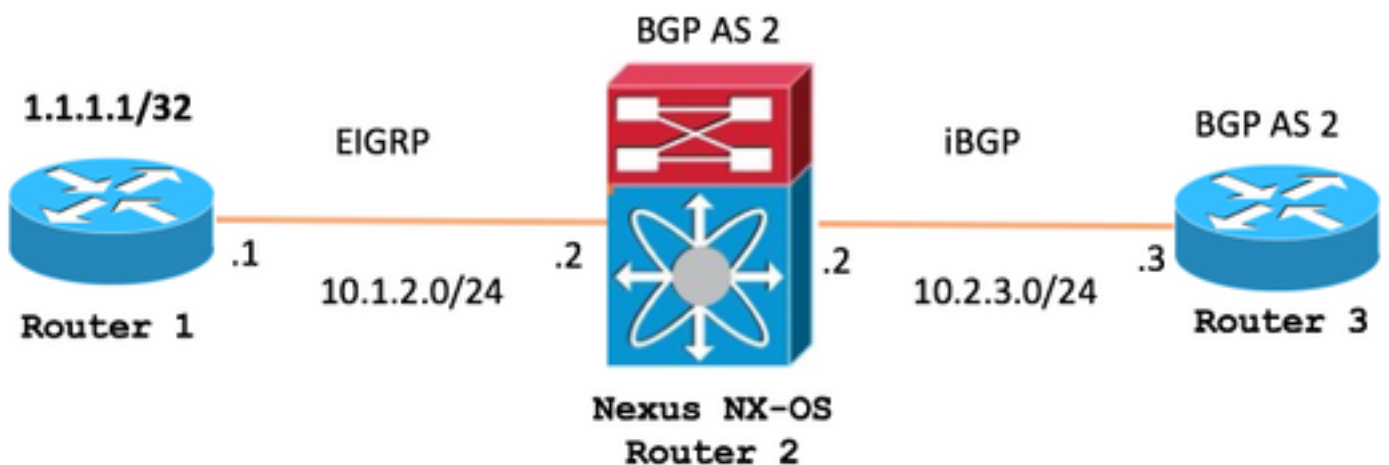
注意：这仅适用于iBGP通告，不适用于eBGP。

注意：适用于配置为静态路由或通过任何内部网关协议(IGP)(如增强型内部网关路由协议(EIGRP)、开放最短路径优先(OSPF)或路由信息协议(RIP))接收的非本地路由。

了解iBGP通告

要了解iBGP通告中的NEXT_HOP集，请以图像中显示的网络拓扑图为例。

Nexus NX-OS案例的拓扑

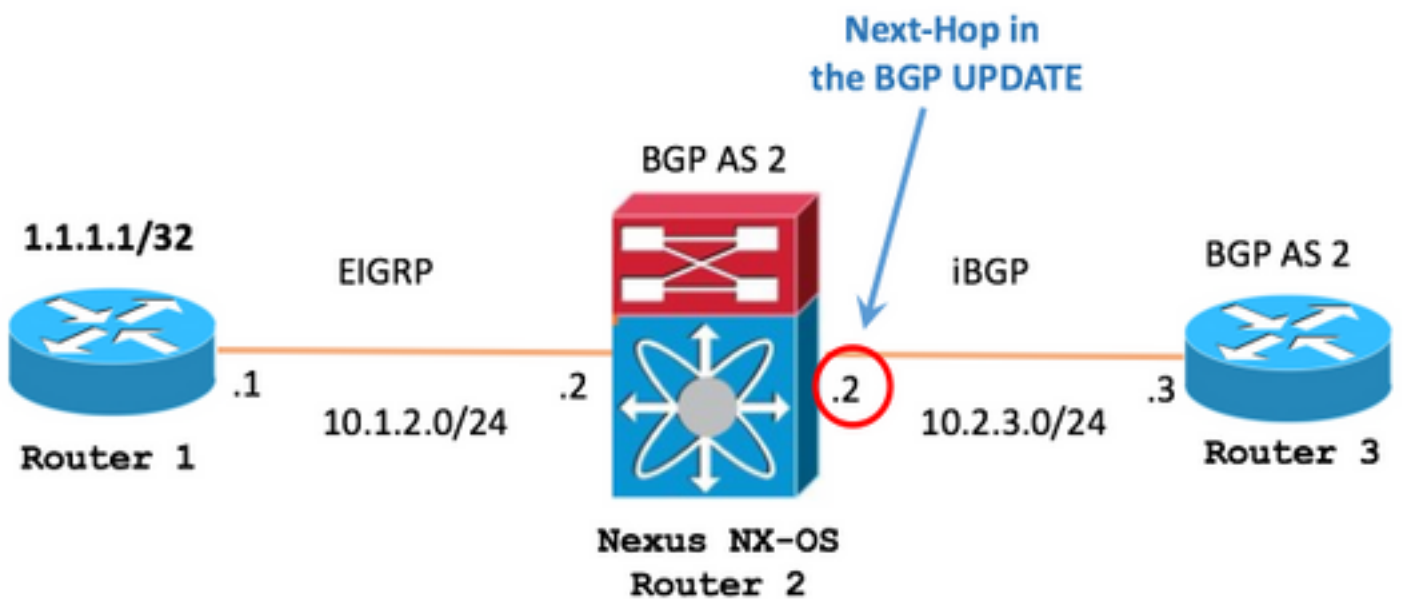


Cisco IOS案例的拓扑



Nexus NX-OS案例

在Nexus NX-OS的拓扑中，R2(Nexus NX-OS)通过EIGRP从路由器1接收1.1.1.1/32路由，并使用iBGP将其通告给路由器3，如图所示。



R2(Nexus NX-OS)路由表显示通过EIGRP接收的路由1.1.1.1/32，原始下一跳IP为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，您可以看到通过iBGP通告1.1.1.1/32到路由器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
```

在路由器3上，1.1.1.1/32路由通过iBGP接收，下一跳现在设置为R2(Nexus NX-OS)的IP地址10.2.3.2

- 1.1.1.1/32的路由器3 BGP表条目

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的路由器3路由表条目

R3

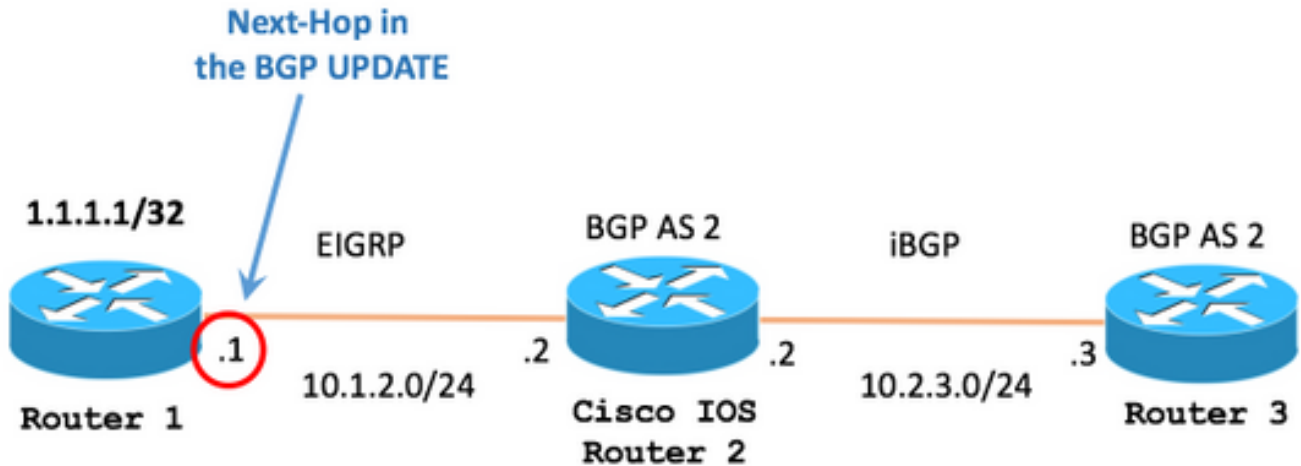
```
R3# show ip route bgp
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       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的拓扑中，R2(Cisco IOS)通过EIGRP从路由器1接收1.1.1.1/32路由，并使用iBGP将其通告给路由器3，如图所示。



R2(Cisco IOS)路由表显示通过EIGRP接收的路由1.1.1.1/32，原始下一跳IP为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，您可以看到通过iBGP通告1.1.1.1/32到路由器3的命令

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
```

在路由器3上，您可以看到通过iBGP收到的1.1.1.1/32路由，其原始下一跳设置为路由器1上的IP，即10.1.2.1。

- 1.1.1.1/32的路由器3 BGP表条目

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
```

在此特定场景中，路由器3必须具有到10.1.2.1（下一跳）的路径，以便BGP可以认为该路径有效。否则，BGP将路径显示为（不可访问）。

注意：这是BGP最佳路径选择算[法中描述的基本检查](#)，以便接受从BGP到路由表的路由。

命令**debug ip bgp update**显示路由器3不安装路由的原因是其路由表中没有下一跳的条目，在本例中，下一跳为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
```

使下一跳可访问的一种方法是：

— 步骤1.在路由器3的路由表中配置一条静态路由，以便为下一跳创建一个条目。

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
```

— 步骤2.相同的debug命令显示路由现在已被接受。

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#
```

— 步骤3. BGP表已删除 (不可访问) 状态。

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
```

— 步骤4.路由表现在安装到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

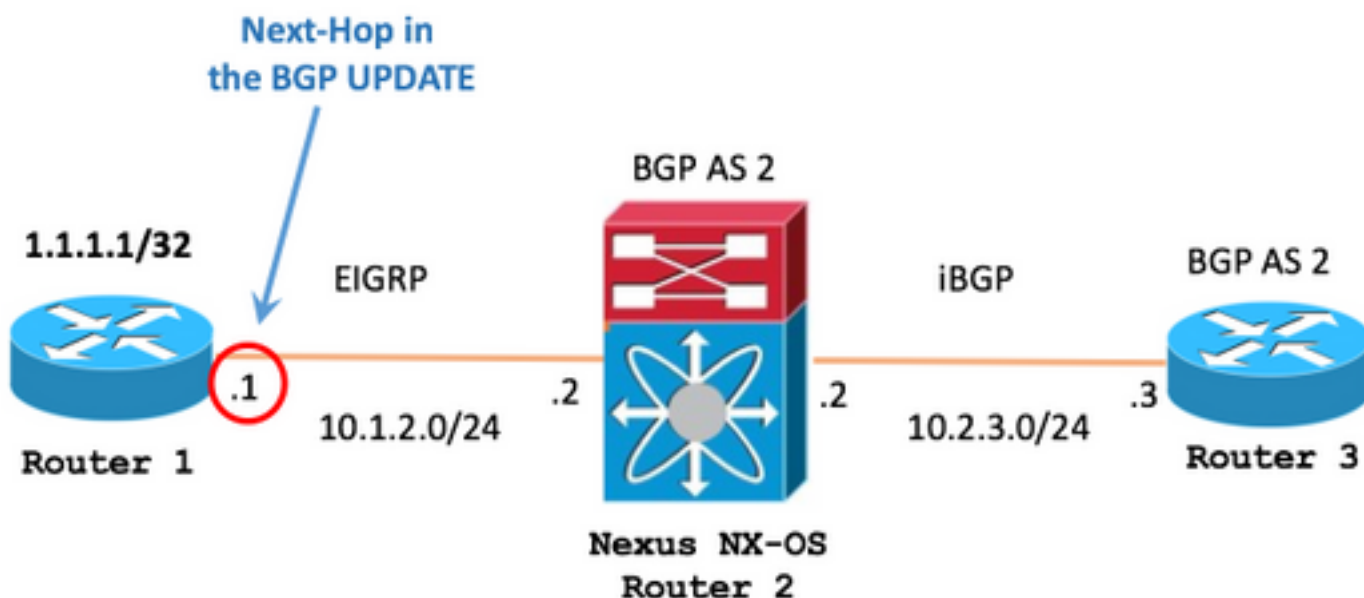
    1.0.0.0/32 is subnetted, 1 subnets
B       1.1.1.1 [200/130816] via 10.1.2.1, 00:11:37
```

使用set ip next-hop redist-unchanged命令

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

注意：这些命令仅在用作路由映射中的参数时起作用，并与redistribution命令结合使用。

在Nexus NX-OS的拓扑中，R2(Nexus NX-OS)通过EIGRP从路由器1接收1.1.1.1/32路由，并使用iBGP将其通告给路由器3，如图所示：



R2(Nexus NX-OS)路由表显示通过EIGRP接收的路由1.1.1.1/32，原始下一跳IP为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
```

路由映射REDIST-UNCHANGED将作为BGP中redistribute配置语句的参数应用。

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
```

现在，路由器3接收BGP更新，其原始NEXT_HOP设置与Cisco IOS类似。

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 ?
```

本文档介绍Nexus NX-OS和Cisco IOS处理非本地生成路由的iBGP通告的方式的区别。

本文档中描述的行为适用于大多数情况场景，并不影响常规网络路由操作。

可选命令`set ip next-hop redist-unchanged`和`set ipv6 next-hop redist-unchanged`可用于维护BGP路由与Nexus NX-OS上的RFC 4271兼容

初始设备配置

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  
!
```