Contents

Introduction

This document describes how the **bgp suppress-inactive** command prevents the advertisement of routes that are not installed in the routing information base (RIB); it also describes the interaction between inactive routes and next-hop mismatch.

A rib-failure occurs when Border Gateway Protocol (BGP) tries to install the bestpath prefix into the RIB, but the RIB rejects the BGP route because a route with better administrative distance already exists in the routing table. An inactive BGP route is a route that is not installed in the RIB, but is installed in the BGP table as rib-failure.

Refer to Suppress BGP Advertisement for Inactive Routes for additional details.

Inactive Routes and Next-Hop Mismatch

When you use the **bgp suppress-inactive** command, it is critical that you understand the impact of next-hop mismatch.

Example Topology

Router 1 (R1) and router 2 (R2) have two parallel links; one links runs BGP AS 65535 and the other link runs Enhanced Interior Gateway Routing Protocol (EIGRP) AS 1. Both BGP and EIGRP are advertising the network 10.1.1.1/32 on R1.
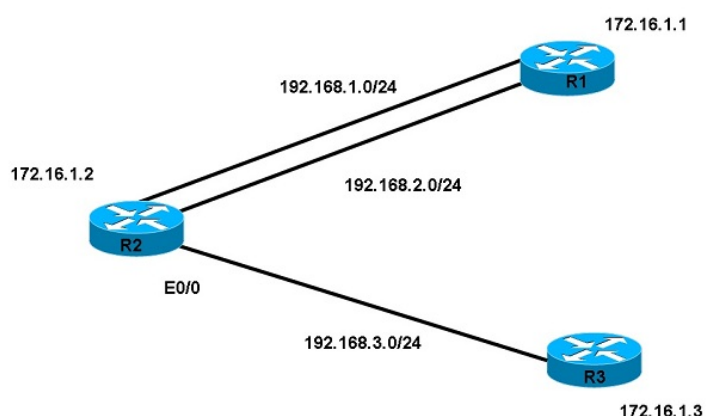
R2 learns about the 10.1.1.1/32 route through both EIGRP and BGP, but installs only the EIGRP route in the routing table because of the lower administrative distance. Since the BGP route is not installed in the R2 routing table, the route appears as a rib-failure in the R2 BGP table. However, R2 advertises the BGP route to router 3 (R3) regardless of the rib-failure.

Show Outputs

For R2, enter the **show ip route** command in order to determine the current status of the routing table on 10.1.1.1, and enter the **show ip bgp** command in order to display the entries in the BGP routing table:

```
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
```

Check the recursive route for the next-hop, since it is a loopback on R1:

```
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
```

Even though the next-hop is mismatched, R2 advertises the route to R3, and R3 learns about the route because inactive routes are not suppressed:

```
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
```

Suppress Inactive Routes in BGP Configuration

Enter the **bgp suppress-inactive** command in order to suppress the inactive BGP routes.

```
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
```

> **Note**: The **bgp suppress-inactive** command suppresses the rib-failed routes **only if** the next-hop of the BGP rib-failure route is different from the next-hop of same route currently installed in the routing table.

```
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
```

In the RIB-NH Matches column, notice that the RIB next-hop does not match. Because the next-hop for 10.1.1.1/32 route is different in EIGRP and BGP, you are able to suppress the rib-failed route with the **bgp suppress-inactive** command.

In other words, if the next-hop in the routing table matches the BGP next-hop, the **bgp suppress-inactive** command does not suppress anymore. That means R3 starts receiving 10.1.1.1/32 route again even if it is RIB failed.

Add Static Route to Match Next-Hop

Add a static route for the prefix in order to match its next-hop in RIB with the next-hop advertised by BGP:

```
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
```

Even with the **bgp suppress-inactive** command, R2 still advertises the route, and R3 still receives the route.

```
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
```

To summarize, the **bgp suppress-inactive** command lets BGP suppress the inactive route advertisement to the neighbors only if a route is already installed in the routing table with a better administrative distance and only if it has a different next-hop than the BGP next-hop for the same route.

Implication of ECMP on Next-Hop and Inactive Routes

In the previous example, if the routes installed in RIB (from EIGRP) are equal-cost multi-path (ECMP) and if inactive routes are suppressed, you see only a part of the routes that are suppressed.

Run EIGRP on both links between R1 and R2. R2 learns a set of prefixes from R1 as ECMP

between the two next hops 192.168.1.1 and 192.168.2.1. For example:

```
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 learns the same set of prefixes from R1 in BGP, and the next-hop loopback is learned on both links.

```
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
```

Since the next-hop route is an ECMP on the same two links, you would expect next-hop to match for all prefixes in BGP and R2 to advertise all of them to R3. When you look at the RIB-NH Matches column of the output, some next-hop (NH) matches are yes and others are 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
```

All routes with RIB-NH Matches of yes are advertised to R3; all others are suppressed.

```
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
```

In Cisco IOS® software, BGP can only pick one next-hop and advertises the best path only with that next-hop (without add-path, multi-path, BGP best-external, or other features).

While RIB installs the EIGRP routes for the destination (note the * in the output), RIB might pick one of the paths as the best path. If that path matches the one for the BGP next-hop, it is reported as yes for next-hop match.

In this example, RIB picked 192.168.1.1 as the next-hop for 10.1.1.1/32 network (note the * on 192.168.1.1 in the output from **sh ip route 172.16.1.1**), which matches with BGP next-hop's route 172.16.1.1; this is reported as a yes in next-hop match. RIB picked 192.168.2.1 as the next-hop for 10.1.1.5/32, which does not match with the BGP next-hop's route; this is reported as a no in next-hop mismatch.

In summary, the next-hop match is important only if you suppress inactive routes; if there is no match, you see an n/a flag in the RIB-NH Matches column, and R2 advertises all the routes to R3.

```
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
```