Use Multi-Exit Discriminator to Select Best Path for BGP Router

Contents

Introduction

This document describes the bgp deterministic-med command and explains how it effects the path selection based on multi-exit discriminator (MED).

Prerequisites

Requirements

There are no specific requirements for this document.

Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

Conventions

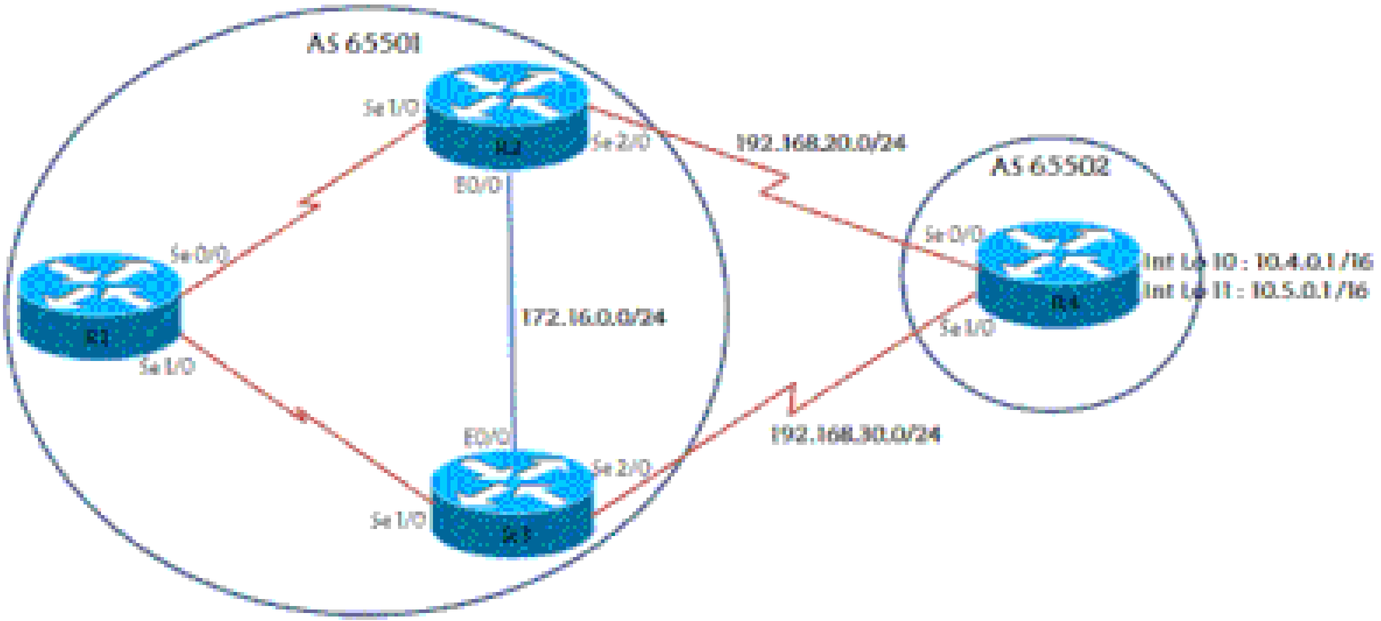
For more information on document conventions, refer to Cisco Technical Tips Conventions.

The MED Attribute

MED is an optional nontransitive attribute. MED is a hint to external neighbors about the preferred path into an autonomous system (AS) that has multiple entry points. The MED is also known as the external metric of a route. A lower MED value is preferred over a higher value.

This section describes an example of how to use MED to influence the routing decision taken by a neighboring AS.

Network Topology



Network Topology

Example

In this scenario, AS 65502 is a user of the ISP which has AS 65501. R4 is connected to two different routers on the ISP side for redundancy purposes, and advertises two networks to the ISP—10.4.0.0/16 and 10.5.0.0/16. Some of the relevant configuration is shown in this section.

R4

```
!
version 12.3
!
hostname r4
!
ip cef
!
!
interface Loopback10
 ip address 10.4.0.1 255.255.0.0
!
interface Loopback11
 ip address 10.5.0.1 255.255.0.0
!
interface Serial0/0
 ip address 192.168.20.4 255.255.255.0
!
interface Serial1/0
 ip address 192.168.30.4 255.255.255.0
```

```
!
router bgp 65502
 no synchronization
 bgp log-neighbor-changes
 network 10.4.0.0 mask 255.255.0.0
 network 10.5.0.0 mask 255.255.0.0
 neighbor 192.168.20.2 remote-as 65501
 neighbor 192.168.30.3 remote-as 65501
 no auto-summary
!
ip classless
!
!
line con 0
 exec-timeout 0 0
line aux 0
line vty 0 4
 exec-timeout 0 0
 login
!
!
end
```

R2

```
!
version 12.3
!
hostname r2
!
ip cef
!
!
interface Loopback0
 ip address 10.2.2.2 255.255.255.255
!
interface Ethernet0/0
 ip address 172.16.0.2 255.255.255.0
!
interface Serial1/0
 ip address 192.168.1.2 255.255.255.0
 serial restart-delay 0
!
interface Serial2/0
 ip address 192.168.20.2 255.255.255.0
 serial restart-delay 0
!
router ospf 1
 log-adjacency-changes
 redistribute connected
 passive-interface Serial2/0
 network 10.2.2.2 0.0.0.0 area 0
 network 172.16.0.2 0.0.0.0 area 0
 network 192.168.1.2 0.0.0.0 area 0
 network 192.168.20.2 0.0.0.0 area 0
!
router bgp 65501
 no synchronization
 bgp log-neighbor-changes
 neighbor 10.1.1.1 remote-as 65501
 neighbor 10.1.1.1 update-source Loopback0
 neighbor 10.3.3.3 remote-as 65501
 neighbor 10.3.3.3 update-source Loopback0
 neighbor 192.168.20.4 remote-as 65502
 no auto-summary
!
ip classless
!
!
line con 0
 exec-timeout 0 0
 transport preferred all
 transport output all
line aux 0
 transport preferred all
 transport output all
line vty 0 4
 exec-timeout 0 0
 login
 transport preferred all
 transport input all
 transport output all
!
end
```

The configurations of R1 and R3 are similar to R2. R3 has an eBGP which peers with R4 and an iBGP which peers with R1.

R1 has an iBGP which peers to R2 and one to R3. Look at what the R1, R2, and R3 BGP tables display for the two networks advertised by R4:

<#root>

```
r2#

show ip bgp 10.4.0.1

BGP routing table entry for 10.4.0.0/16, version 7
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  10.1.1.1 10.3.3.3
  65502
    192.168.20.4 from 192.168.20.4 (10.4.4.4)
      Origin IGP, metric 0, localpref 100, valid, external, best
  65502
    192.168.30.4 (metric 74) from 10.3.3.3 (10.3.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal

r2#

show ip bgp 10.5.0.1

BGP routing table entry for 10.5.0.0/16, version 6
Paths: (2 available, best #2, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  10.1.1.1 10.3.3.3
  65502
    192.168.30.4 (metric 74) from 10.3.3.3 (10.3.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal
  65502
    192.168.20.4 from 192.168.20.4 (10.4.4.4)
      Origin IGP, metric 0, localpref 100, valid, external, best

r3#

show ip bgp 10.4.0.1

BGP routing table entry for 10.4.0.0/16, version 8
Paths: (2 available, best #2, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  10.1.1.1 10.2.2.2
  65502
    192.168.20.4 (metric 74) from 10.2.2.2 (10.2.2.2)
      Origin IGP, metric 0, localpref 100, valid, internal
  65502
    192.168.30.4 from 192.168.30.4 (10.4.4.4)
      Origin IGP, metric 0, localpref 100, valid, external, best

r3#

show ip bgp 10.5.0.1

BGP routing table entry for 10.5.0.0/16, version 10
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  10.1.1.1 10.2.2.2
  65502
    192.168.30.4 from 192.168.30.4 (10.4.4.4)
      Origin IGP, metric 0, localpref 100, valid, external, best
  65502
    192.168.20.4 (metric 74) from 10.2.2.2 (10.2.2.2)
      Origin IGP, metric 0, localpref 100, valid, internal

r1#

show ip bgp 10.4.0.1

BGP routing table entry for 10.4.0.0/16, version 11
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Not advertised to any peer
  65502
    192.168.20.4 (metric 128) from 10.2.2.2 (10.2.2.2)
      Origin IGP, metric 0, localpref 100, valid, internal, best
  65502
    192.168.30.4 (metric 128) from 10.3.3.3 (10.3.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal

r1#

show ip bgp 10.5.0.1

BGP routing table entry for 10.5.0.0/16, version 10
Paths: (2 available, best #2, table Default-IP-Routing-Table)
  Not advertised to any peer
  65502
    192.168.30.4 (metric 128) from 10.3.3.3 (10.3.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal
  65502
    192.168.20.4 (metric 128) from 10.2.2.2 (10.2.2.2)
      Origin IGP, metric 0, localpref 100, valid, internal, best
```

Both R2 and R3 pick as best path the external route from R4 which is expected based on the BGP bestpath selection algorithm. Refer to BGP Best Path Selection Algorithm for more information.

Similarly, R1 chooses R2 to access the 2 networks, which is in accordance with the BGP best path rule—select the path with the lowest router ID. Because the R2 router ID is 10.2.2.2 and the R3 router ID is 10.3.3.3, R2 is chosen. In this basic configuration, all traffic to the two networks in AS 65502 passes from R1 through R2 and then to R4 by default. Now, suppose that R4 wants to load balance the traffic it receives from AS 65501. To do so without any R4 ISP modifications, you configure R4 to utilize MED to force traffic for one network down one path, and traffic for the other network down the other path.

This is the configuration of R4 after you apply the necessary configuration:

R4

```
<#root>

!
version 12.3
!
hostname r4
!
ip cef
!
!
!
interface Loopback10
 ip address 10.4.0.1 255.255.0.0
!
interface Loopback11
 ip address 10.5.0.1 255.255.0.0
!
interface Serial0/0
 ip address 192.168.20.4 255.255.255.0
!
interface Serial1/0
 ip address 192.168.30.4 255.255.255.0
!
router bgp 65502
 no synchronization
 bgp log-neighbor-changes
 network 10.4.0.0 mask 255.255.0.0
 network 10.5.0.0 mask 255.255.0.0
 neighbor 192.168.20.2 remote-as 65501


neighbor 192.168.20.2 route-map setMED-R2 out

 neighbor 192.168.30.3 remote-as 65501


neighbor 192.168.30.3 route-map setMED-R3 out

 no auto-summary
!
ip classless
no ip http server
!
!

access-list 1 permit 10.4.0.0 0.0.255.255
access-list 2 permit 10.5.0.0 0.0.255.255
```

```
!
route-map setMED-R3 permit 10
 match ip address 1
 set metric 200
!
route-map setMED-R3 permit 20
 match ip address 2
 set metric 100


!--- The route-map MED-R3 is applying a MED of 200 to the 10.4.0.0/16
!--- network and a MED of 100 to the 10.5.0.0/16 network.
!--- The route-map is being applied outbound towards R3.

! route-map setMED-R2 permit 10 match ip address 1 set metric 100 ! route-map setMED-R2 permit 20 match

 !--- The route-map MED-R2 is applying a MED of 100 to the 10.4.0.0/16
!--- network and a MED of 200 to the 10.5.0.0/16 network.
!--- The route-map is being applied outbound towards R2.  ! ! ! line con 0 exec-timeout 0 0 line aux 0 l
```

Note: You need to clear the BGP session with the clear ip bgp * soft out command, for example, to make these configurations take action.

R1 now sees the route over R2 as the best path for network 10.4.0.0/16 because the update received from R2 has a MED of 100 versus a MED of 200, which is what R3 advertises. Similarly, R1 uses R3 and the R3 - R4 link to access 10.5.0.0/16:

```
<#root>

r1#

show ip bgp 10.4.0.1

BGP routing table entry for 10.4.0.0/16, version 14
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Flag: 0x800
  Not advertised to any peer
  65502
    192.168.20.4 (metric 128) from 10.2.2.2 (10.2.2.2)
      Origin IGP,

metric 100

, localpref 100, valid, internal, best
r1#sh ip bgp 10.5.0.1
BGP routing table entry for 10.5.0.0/16, version 13
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Flag: 0x800
  Not advertised to any peer
  65502
    192.168.30.4 (metric 128) from 10.3.3.3 (10.3.3.3)
      Origin IGP,

metric 100

, localpref 100, valid, internal, best
```

Look at the R2 display:

<#root>

```
r2#

show ip bgp 10.4.0.1

BGP routing table entry for 10.4.0.0/16, version 10
Paths: (1 available, best #1, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  10.1.1.1 10.3.3.3
  65502
    192.168.20.4 from 192.168.20.4 (10.4.4.4)
      Origin IGP, metric 100, localpref 100, valid, external, best


r2#

show ip bgp 10.5.0.1

BGP routing table entry for 10.5.0.0/16, version 11
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  192.168.20.4
  65502
    192.168.30.4 (metric 74) from 10.3.3.3 (10.3.3.3)
      Origin IGP, metric 100, localpref 100, valid, internal, best
  65502
    192.168.20.4 from 192.168.20.4 (10.4.4.4)
      Origin IGP, metric 200, localpref 100, valid, external
```

The reason why R2 only shows one path for 10.4.0.0/16 is because R3 withdraws (sends an update with unreachable metric) the update for 10.4.0.0/16 once it notices that R3 uses R2 to access 10.4.0.0/16 (after you run BGP bestpath on all available paths):

<#root>

```
r3#

show ip bgp 10.4.0.0

BGP routing table entry for 10.4.0.0/16, version 20
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  192.168.30.4
  65502
    192.168.20.4 (metric 74) from 10.2.2.2 (10.2.2.2)
      Origin IGP, metric 100, localpref 100, valid, internal, best
  65502
    192.168.30.4 from 192.168.30.4 (10.4.4.4)
      Origin IGP, metric 200, localpref 100, valid, external
```

This allows R2 to save some memory since it does not have to store this useless information. In the event that the BGP session between R2 and R4 fails, R2 would send an unreachable update to R3 for 10.4.0.0/16. This update would trigger R3 to send an update with the R3 route for 10.4.0.0/16 via R4 to R2. R2 could start to route via R3.

The bgp deterministic-med Command

If you enable the **bgp deterministic-med** command, it removes any temporal dependency of MED-based best path decisions. It ensures that an accurate MED comparison is made across all routes received from the same autonomous system (AS).

If you disable **bgp deterministic-med** , the order in which routes are received can impact MED-based best path decisions. This can occur when the same route is received from multiple ASs or confederation sub-ASs, with exactly the same path length, but different MEDs.

Examples

For example, consider the next routes:

```
entry1: ASPATH 1, MED 100, internal, IGP metric to NEXT_HOP 10
entry2: ASPATH 2, MED 150, internal, IGP metric to NEXT_HOP 5
entry3: ASPATH 1, MED 200, external
```

The order in which the BGP routes were received is entry3, entry2, and entry1 (entry3 is the oldest entry in the BGP table and entry1 is the newest one).

A BGP Router with bgp deterministic-med Disabled

A BGP router with **bgp deterministic-med** disabled chooses entry2 over entry1, due to a lower IGP metric to reach the NEXT_HOP (MED was not used in this decision because entry1 and entry2 are from two different ASs). It then prefers entry3 over entry2 because it is external. However, entry3 has a higher MED than entry1. For more information about BGP path selection criteria, refer to BGP Best Path Selection Algorithm .

A BGP Router with bgp deterministic-med Enabled

In this case, routes from the same AS are grouped together, and the best entries of each group are compared. In the given example, there are two ASs, AS 1 and AS 2.

<#root>

```
Group 1:

entry1: ASPATH 1, MED 100, internal, IGP metric to NEXT_HOP 10

          entry3: ASPATH 1, MED 200, external
Group 2:

entry2: ASPATH 2, MED 150, internal, IGP metric to NEXT_HOP 5
```

In Group 1, the best path is entry1 because of the lower MED (MED is used in this decision since the paths are from the same AS). In Group 2, there is only one entry (entry2). The best path then is determined with a comparison of the winners of each group (MED is not used in this comparison by default because the winners of each group are from different ASs. When you enable **bgp always-compare-med** itchanges this default behavior). Now, when you compare entry1 (the winner from Group 1) and entry2 (the winner from Group 2), entry2 can be the winner since it has the better IGP metric to the next hop.

If **bgp always-compare-med** was also enabled when you compare entry1 (the winner from Group 1) and entry 2 (the winner from Group 2), entry 1 can be the winner because of lower MED.

Cisco recommends that you enable **bgp always-compare-med** in all new network deployments. In addition, if bgp always-compare-med is enabled, BGP MED decisions are always deterministic.

For more information on the **bgp deterministic-med** and the **bgp always-compare-med** commands, refer toHow the bgp deterministic-med Command Differs from the bgp always-compare-med Command.

Related Information

- Sample Configuration for BGP with Two Different Service Providers (Multihoming)
- IP/BGP Support
- BGP Case Studies
- Cisco Technical Support & Downloads