使用MultiExit标识符影响BGP最佳路径选择

目录

简介

先决条件

- 要求
- 使用的组件
- 规则

MED 属性

- 网络拓扑
- 示例

MED命令

- Examples
  - 禁用MED的BGP路由器
  - 启用了MED的BGP路由器

相关信息

简介

本文档介绍bgp deterministic-med命令以及它如何影响基于多出口标识符(MED)的路径选择。

先决条件

要求

Cisco 建议您了解以下主题：

- 基本 IP 路由
- 边界网关协议 (BGP)

使用的组件

本文档不限于特定的软件和硬件版本。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。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。

规则

有关Cisco文档规则的详细信息，请参阅Cisco技术提示规则。

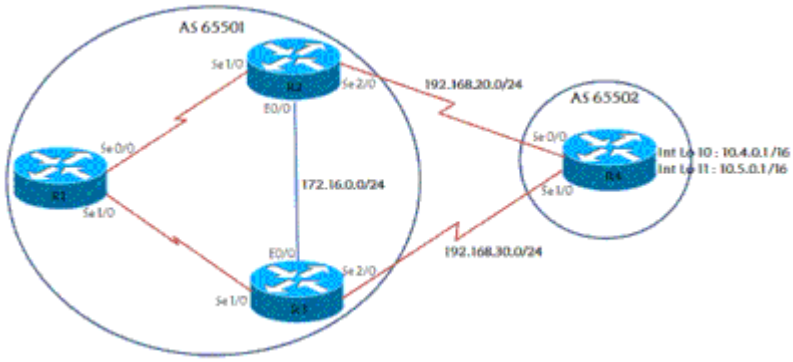
MED 属性

MED是一个可选的非传递性BGP属性，它向外部邻居提供有关进入具有多个入口点的自治系统(AS)的首选路径的提示。MED也称为路由的外部度量，并且更低的MED值优先于更高的值。默认情况下，BGP仅比较从同一邻居AS接收的路径中的MED值。



注意：默认情况下，除非配置了bgp always-compare-med，否则BGP只会比较从同一个邻居AS接收的路径的MED值。有关详细信息，请参阅bgp deterministic-med命令与bgp always-compare-med命令有何不同。

网络拓扑



网络拓扑

示例

在本场景中，AS 65502是具有AS 65501的ISP的用户。R4连接到ISP端的两台不同的路由器以实现

冗余，并将两个网络（10.4.0.0/16和10.5.0.0/16）通告给ISP。部分相关配置如本部分所示。

R4

```
!
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

```
!
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
```

相关的R1和R3配置使用与R2相同的设计。R1、R2和R3通过环回接口具有iBGP可达性，而IGP提供与BGP下一跳的可达性。R3与R4具有eBGP会话，与R1和R2具有iBGP会话。

R1有一个与R2对等的iBGP，以及一个与R3对等的iBGP。查看R4通告的两个网络的R1、R2和R3 BGP表显示内容：

<#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
```

R2和R3都选择最佳路径作为来自R4的外部路由，根据BGP最佳路径选择算法，这是预期的。有关详细信息，请参阅 BGP 最佳路径选择算法。

同样，R1选择R2来访问这两个网络，因为较早的BGP最佳路径属性相同，而最低的BGP路由器ID用作连接断开器。R2的路由器ID为10.2.2.2，而R3的路由器ID为10.3.3.3。由于R2的路由器ID为10.2.2.2，而R3的路由器ID为10.3.3.3，因此选择R2。在此基本配置中，默认情况下，发往AS 65502中两个网络的所有流量从R1通过R2，然后到R4。现在，假设R4要对其从AS 65501收到的流量进行负载均衡。要在不修改R4 ISP的情况下这样做，请将R4配置为使用MED强制一个网络的流量沿一条路径传输，另一个网络的流量沿另一条路径传输。

注意：请注意，所描述的MED配置示例提供基于前缀的流量工程，而不是真正的等价负载均衡。

以下是R4在应用必要配置后的配置：

R4

```
<#root>

!
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 setMED-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 ip address 2
set metric 200



!--- The route-map setMED-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
line vty 0 4
exec-timeout 0 0
login
!
!
end
```



注意:在本示例中，仅通告10.4.0.0/16和10.5.0.0/16。如果向同一邻居通告其他前缀，请将其添加到适当的访问列表或路由映射序列。如果必须通告不匹配的前缀而不通过出站路由映射过滤，则添加最终的允许序列。



注意：应用出站路由映射后，通过运行clear ip bgp * soft out命令刷新出站BGP更新。在生产中，尽可能使用特定于邻居的soft outbound clear，例如clear ip bgp 192.168.20.2 soft out和clear ip bgp 192.168.30.3 soft out。

应用MED后，R2选择R2到R4的路径/链路作为10.4.0.0/16的最佳路径/链路，R3选择R3到R4的路径/链路作为10.5.0.0/16的最佳路径/链路。在所示的融合状态下，R1接收每个前缀的符合条件的最佳iBGP通告，并使用R2作为10.4.0.0/16，使用R3作为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#

show 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
```

有关R2的显示信息，请参见下一个示例：

<#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
```

R2显示了一条通向10.4.0.0/16的路径，因为R3选择通过R2的iBGP获知的路径作为最佳路径后，不再通告其R4获知的路径。当R3注意到R3使用R2访问10.4.0.0/16时，R3会撤消10.4.0.0/16的更新（发送BGP路由撤消，此更新包含不可达度量）。标准iBGP通告规则阻止R3将获知iBGP的路径通告回本拓扑中的另一个iBGP对等体:

<#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
```

这允许R2节省一些内存，因为它不必存储这些无用的信息。如果R2和R4之间的BGP会话失败，R2会向R3发送有关10.4.0.0/16的不可达更新。此更新会触发R3通过R4向R2发送有关10.4.0.0/16的R3路由的更新。R2可以开始通过R3进行路由。

MED命令

要启用确定性MED比较，请在BGP路由器配置模式下配置bgp deterministic-med命令。启用时，bgp deterministic-med通过在比较之前将来自同一AS的路径分组来从基于MED的最佳路径决策中删除临时相关性。这样可以确保在从同一自治系统(AS)收到的所有路由之间进行准确的MED比较。

如果禁用bgp deterministic-med，则接收路由的顺序可能会影响基于MED的最佳路径决策。如果从多个 AS 或联盟的子 AS 处接收具有完全相同的路径长度、不同 MED 的相同路由，则可能会发生这种情况。

Examples

例如，考虑以下路由：

条目	AS路径	中文	本地首选	AS路径长度	原点
entry1	AS 65001	100	100	1	IGP
entry2	AS 65002	50	100	1	IGP
entry3	AS 65001	20	100	1	IGP

禁用MED的BGP路由器

到达订单1示例

收到的BGP路由的顺序为（entry1是BGP表中最旧的条目，entry3是最新条目）：

1. entry1
2. entry2
3. entry3

最初，只有一条路由，因此entry1成为最佳路径。将AS65002的entry2与AS65001的entry1进行比较时，由于路由来自不同的邻居AS，因此忽略MED;由于所有剩余属性都相同，路由器将保留当前的最佳路径entry1。

将MED为20的AS65001的entry3与MED为100的AS65001的entry1进行比较时，会对MED进行评估，因为两条路由都来自同一个相邻AS。由于20小于100,entry3成为最终最佳路径。

到达订单2示例

现在，假设相同的路由以不同的顺序到达：

1. entry2
2. entry3
3. entry1

最初，entry2被选为最佳路径。将AS65001的entry3与AS65002的entry2进行比较时，不会评估MED，因为路由来自不同的邻居AS;由于所有剩余属性都相同，路由器将保留当前的最佳路径entry2。

将AS1的entry65001与来自AS65002的entry2进行比较时，由于路由来自不同的邻居AS，因此再次忽略MED。因此，路由器会将entry2保留为最终最佳路径。

禁用MED后，路由到达顺序会影响比较序列。请注意，收到的路由器是完全相同的三条路由，但它选择了不同的最佳路径，因为路由到达的顺序不同。这是bgp deterministic-med旨在消除的时间相关性。

注意：有关BGP路径选择标准的详细信息，请参阅BGP最佳路径选择算法。

启用了MED的BGP路由器

启用MED后，路由器首先通过相邻AS对路由进行分组，然后做出任何基于MED的决策。在这种情况下，来自同一 AS 的路由器被分为一组，并且每组中的最佳项会进行比较。在给定的示例中，有两个 AS：AS 65001 和 AS 65002。

组	AS路径	条目	所选路由	原因
第 1 组	AS 65001	entry1(MED 100)、entry3(MED 20)	entry3	较低MED(20 < 100)
第 2 组	AS 65002	entry2(MED 50)	entry2	唯一应聘者

在组1中，最佳路径是entry3，因为最低的MED（由于路径来自同一AS，因此在此决策中使用MED）。 在 Group 2 中，只有一项 (entry2)。 然后，通过比较各组优胜者的优胜者确定最佳路径，由于各组的优胜者来自不同的AS，因此默认情况下不使用MED。

此时，不再考虑MED，因为剩余路由来自不同的相邻AS。BGP最佳路径算法继续使用后续属性（例如eBGP与iBGP、到下一跳的IGP度量、最旧路径、路由器ID等，具体取决于不同的属性）。

主要优点是此过程与接收的路由无关。路由器是否按以下方式获取路由：

- entry1 → entry2 → entry3
- entry2 → entry3 → entry1
- entry3 → entry1 → entry2

因此，到货订单产生相同的结果，因为决策基于路由属性而不是接收更新的订单。

注意：如果在比较entry3（组1的获胜者）和entry 2（组2的获胜者）时也启用了bgp always-compare-med;由于中位数的降低，第3条是获胜方。

注意：启用bgp deterministic-med命令可以确保在选择由同一AS中的不同对等体通告的路由时对MED变量进行比较。启用bgp always-compare-med命令可以确保为来自不同AS中邻居的路径比较MED。

此示例假定所有优先级较高的BGP属性(如“权重”(Weight)、“本地优先级”(Local Preference)、“AS路径长度”(AS Path length)、“原点”(Origin)等)相同。目的是隔离MED的影响，并解释为什么禁用bgp deterministic-med可确保最佳路径决策取决于路由到达顺序。

思科建议您在所有新网络部署中启用bgp always-compare-med。此外，如果启用了 bgp always-compare-med，则 BGP MED 决策始终是决定性的。

有关 bgp deterministic-med 命令和 bgp always-compare-med 命令的详细信息，请参阅 bgp deterministic-med 命令和 bgp always-compare-med 命令的区别。

相关信息

- 两个不同服务提供商（多宿）间的 BGP 的示例配置
- IP/BGP支持
- BGP 案例分析
- 思科技术支持和下载

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。