

如何使用AIGP优化BGP路径

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[背景信息](#)

[累计的内部网关协议属性](#)

[使用案例示例](#)

[1. 多ASN网络](#)

[配置](#)

[确认](#)

[2. 无缝MPLS \(单ASN\)](#)

[观察结果](#)

[设备日志-初始](#)

[AIGP解决方案](#)

[配置](#)

[配置示例](#)

[设备日志-在AIGP实施之后](#)

[注意事项](#)

[忽略AIGP](#)

[结论](#)

[相关信息](#)

简介

本文档介绍边界网关协议(BGP)中的累计内部网关协议(AIGP)度量及其使用案例。

先决条件

要求

Cisco 建议您了解以下主题：

- 调试输出中显示“BGP”
- 无缝多协议标签交换(MPLS)网络

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。

背景信息

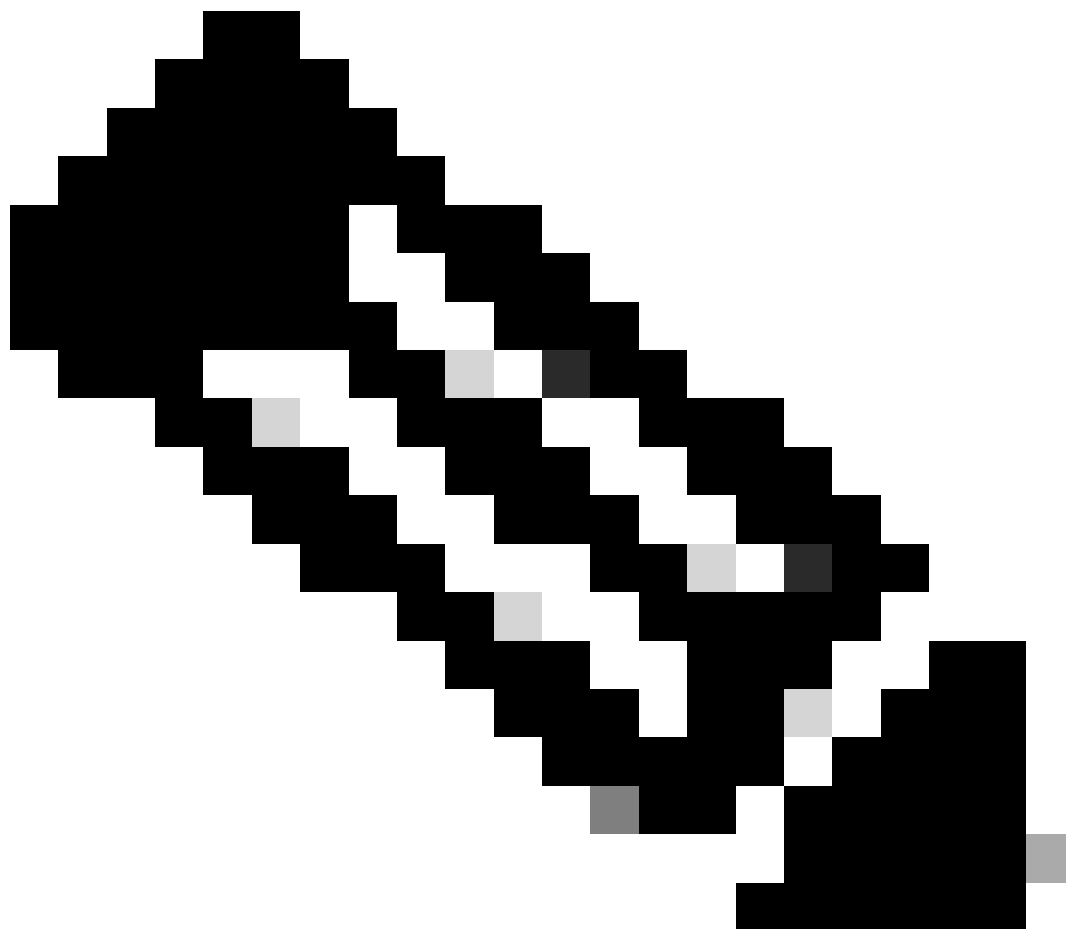
本节概述了AIGP度量以及有关其使用的一些重要注意事项。

如您所知，IGP代表内部网关协议，代表一组在单个管理域中运行的路由协议。IGP根据度量值做出路径选择决策。

BGP旨在通过大量独立的自治系统(AS)提供路由，但相关管理部门之间协调有限甚至不协调。它不会通过使用度量来作出路径选择决定。但是，在一些部署中，单一管理运行多个连续的BGP网络。在这种情况下，可能希望在单个管理域中，BGP可以基于度量选择路径，就像IGP需要做的那样。

累计的内部网关协议属性

AIGP度量（通过RFC7311定义）是可选的非传递BGP路径属性。AIGP属性的值字段定义为一组类型/长度/值元素(TLV)。BGP AIGP TLV包含累积的IGP度量。



注意：不支持可选非传递属性（例如，AIGP）的BGP路由器必须删除此类属性，并且不得将其传递给其他BGP对等体。AIGP度量不能在完全不同的自治系统之间传递（仅跨内部AS边界）。

使用案例示例

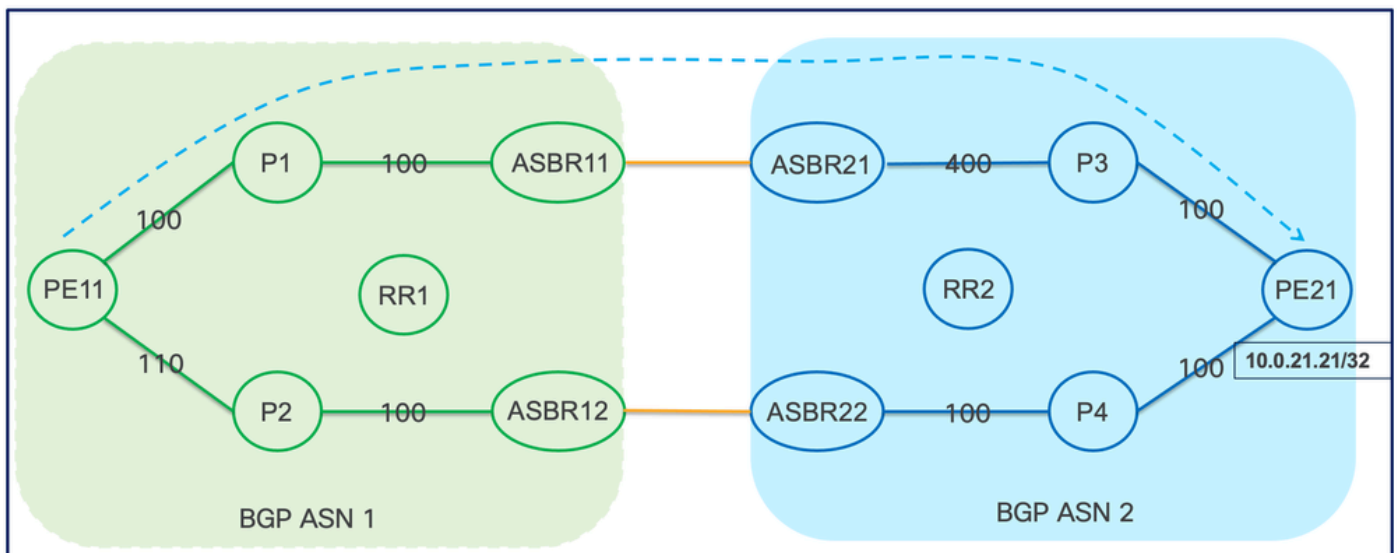
1. 多ASN网络

如今，许多网络位于单个管理域中，由于各种原因，这些网络被细分为多个ASN。这可能有许多原因：

- IGP扩展
- 一个提供商网络被另一个网络提供商收购，但他们尚未在内部合并其BGP ASN
- 不同的业务部门在内部都有独立的网络
- 具有子AS的BGP联盟
- 无缝MPLS等

在这样的网络中，允许BGP根据IGP度量做出决策很有用，这样，即使两个节点位于两个不同的ASN中，BGP也会选择两个节点之间的最短端到端路径。

例如：ABC网络，它可细分为两个BGP ASN，即ASN 1和ASN 2。它们在ASBR对等，链路IGP开销代表带宽。目标是在PE11和PE21之间具有端到端最佳路径。



没有AIGP的多ASN网络

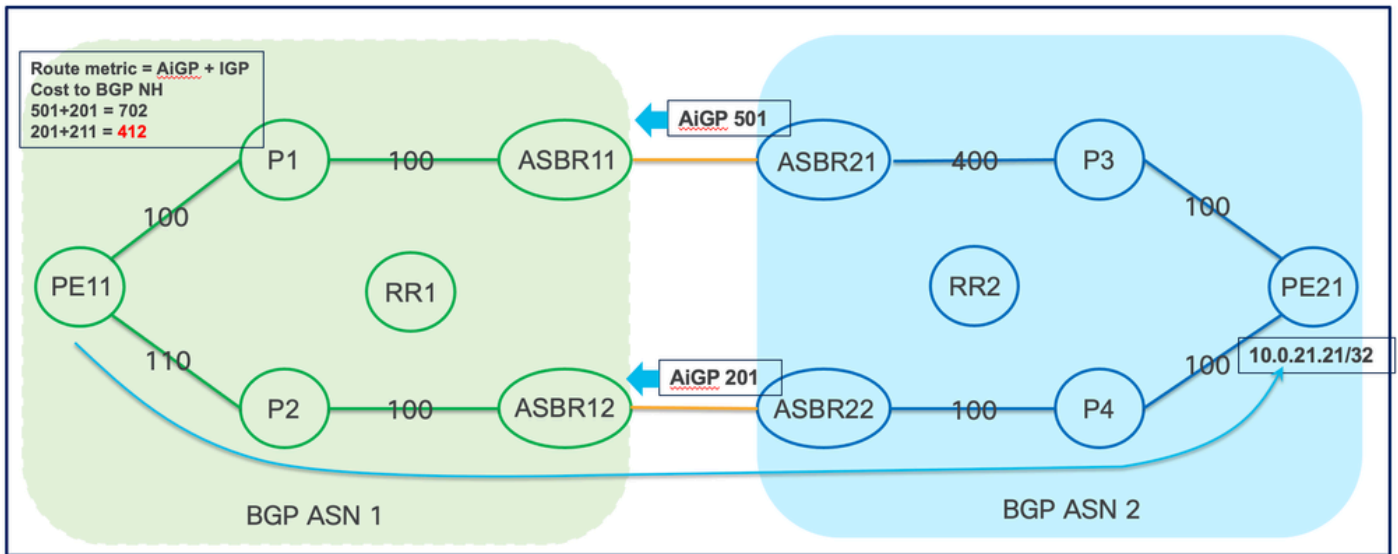
注意：

1. 假设在RR1/2上启用了Add-path以向PE通告两个NH。
2. 上述场景中使用的设备全部使用Cisco IOS-XE运行。

```
PE11#sh bgp ipv4 unicast 10.0.21.21/32
BGP routing table entry for 10.0.21.21/32, version 20
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 3
  2
    192.168.0.12 (metric 211) from 192.168.11.11 (192.168.11.11)
      Origin IGP, metric 0, localpref 100, valid, internal
      Originator: 192.168.0.12, Cluster list: 192.168.11.11
      rx pathid: 0x1, tx pathid: 0
  Refresh Epoch 3
  2
    192.168.0.11 (metric 201) from 192.168.11.11 (192.168.11.11)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Originator: 192.168.0.11, Cluster list: 192.168.11.11
```

rx pathid: 0x0, tx pathid: 0x0

在拓扑中启用AiGP后 (在PE11、PE32、ASBR1x、ASBR2x、RR1、RR2上) , PE11现在会选择端到端IGP开销最低的路径。



使用AiGP的多ASN网络

配置

PE_x、ASBR_x、RR_n :

AiGP功能配置 :

```
router bgp ASN
 neighbor <NBR_IP> aigp
!
```

注意：BGP对等会丢弃并重新建立以协商此新功能。因此，建议您在维护窗口中执行此任务。

通告前缀的AIGP度量。

PE21：

```
route-map SET_AIGP permit 10
  set aigp-metric igp-metric
!
router bgp 2
  address-family {ipv4|ipv6} unicast
    network 10.0.21.21 mask 255.255.255.255 route-map SET_AIGP
!
```

确认

```

PE11#sh bgp ipv4 unicast 10.0.21.21/32
BGP routing table entry for 10.0.21.21/32, version 21
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 3
  2
    192.168.0.11 (metric 201) from 192.168.11.11 (192.168.11.11)
      Origin IGP, aigp-metric 501, metric 0, localpref 100, valid, internal
      Originator: 192.168.0.11, Cluster list: 192.168.11.11
      rx pathid: 0x1, tx pathid: 0
  Refresh Epoch 3
  2
    192.168.0.12 (metric 211) from 192.168.11.11 (192.168.11.11)
  Origin IGP, aigp-metric 201, metric 0, localpref 100, valid, internal, best
  Originator: 192.168.0.12, Cluster list: 192.168.11.11
  rx pathid: 0x0, tx pathid: 0x0

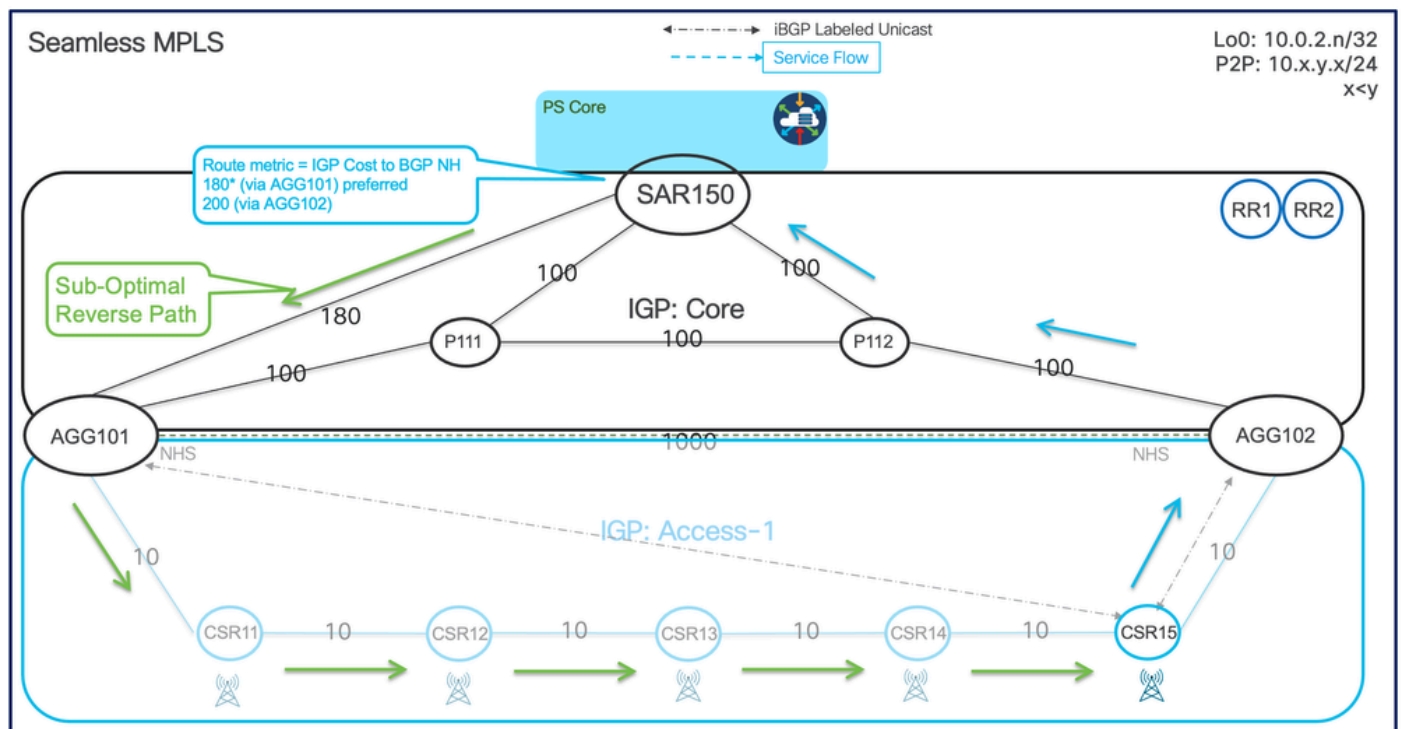
```

2. 无缝MPLS (单ASN)

在大型服务提供商核心网络中，传输网络通常被划分成不同的IGP域，使用BGP标记的单播拼接在一起以提供端到端标记交换路径(LSP)。边界路由器在BGP LU AF中执行下一跳自身(NHS)。

IGP/LDP仅在本地区域/域中传输前缀/标签信息。然后，BGP通过在区域边界将路由重分发到BGP来将前缀/标签传送到所有远程区域/域。然后使用LSP通告路由/标签。路由的下一跳在每个ABR上更改为本地路由器，这样就无需跨区域/域边界泄漏IGP路由。

在此拓扑图中，有一个BGP域被划分为2个IGP域 (CORE和Access-1)。每条链路旁边显示的数字代表该链路的IGP开销/度量。



无需AIGP的无缝MPLS网络

挑战：从PS-Core到eNB/gNB（连接到CSR15）的下行流量与从eNB/gNB（连接到CSR15）到PS-Core的上行流量相比，采用的是非对称次优路径，这导致了移动流量中的延迟问题。

观察结果

1. 主要在地理边界场景中观察到，其中同一个汇聚路由器充当多个接入域的公用边界路由器（例如，前面提到的映像中的AGG102）。
2. 对于向上流量，信元站点路由器(CSR)路由器选择最近的边界路由器。例如，CSR15选择AGG102作为NextHop。
3. 对于下行流量，服务聚合路由器(SAR)也选择最近的边界路由器。例如，SAR150选择AGG101(cost 180 < 200)。

设备日志-初始

上行流量- CSR15至SAR150

```
RP/0/0/CPU0:CSR15#traceroute mpls ipv4 10.0.2.150/32 so 10.0.2.15
Tracing MPLS Label Switched Path to 10.0.2.150/32, timeout is 2 seconds
```

```
Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
'L' - labeled output interface, 'B' - unlabeled output interface,
'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
'P' - no rx intf label prot, 'p' - premature termination of LSP,
'R' - transit router, 'I' - unknown upstream index,
'X' - unknown return code, 'x' - return code 0
```

Type escape sequence to abort.

```
0 10.15.102.15 MRU 1500 [Labels: explicit-null/16150 Exp: 0/0]
L 1 10.15.102.102 MRU 1500 [Labels: 16150 Exp: 0] 0 ms
. 2 *
! 3 10.112.150.150 20 ms
```

```
!!!! AGG102
!!!! P112 does not have a route to
!!!! SAR150
```

下行流量- SAR150至CSR15

```
RP/0/0/CPU0:SAR150#traceroute mpls ipv4 10.0.2.15/32 source 10.0.2.150
Tracing MPLS Label Switched Path to 10.0.2.15/32, timeout is 2 seconds
```

```
Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
'L' - labeled output interface, 'B' - unlabeled output interface,
'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
'P' - no rx intf label prot, 'p' - premature termination of LSP,
'R' - transit router, 'I' - unknown upstream index,
'X' - unknown return code, 'x' - return code 0
```

Type escape sequence to abort.


```

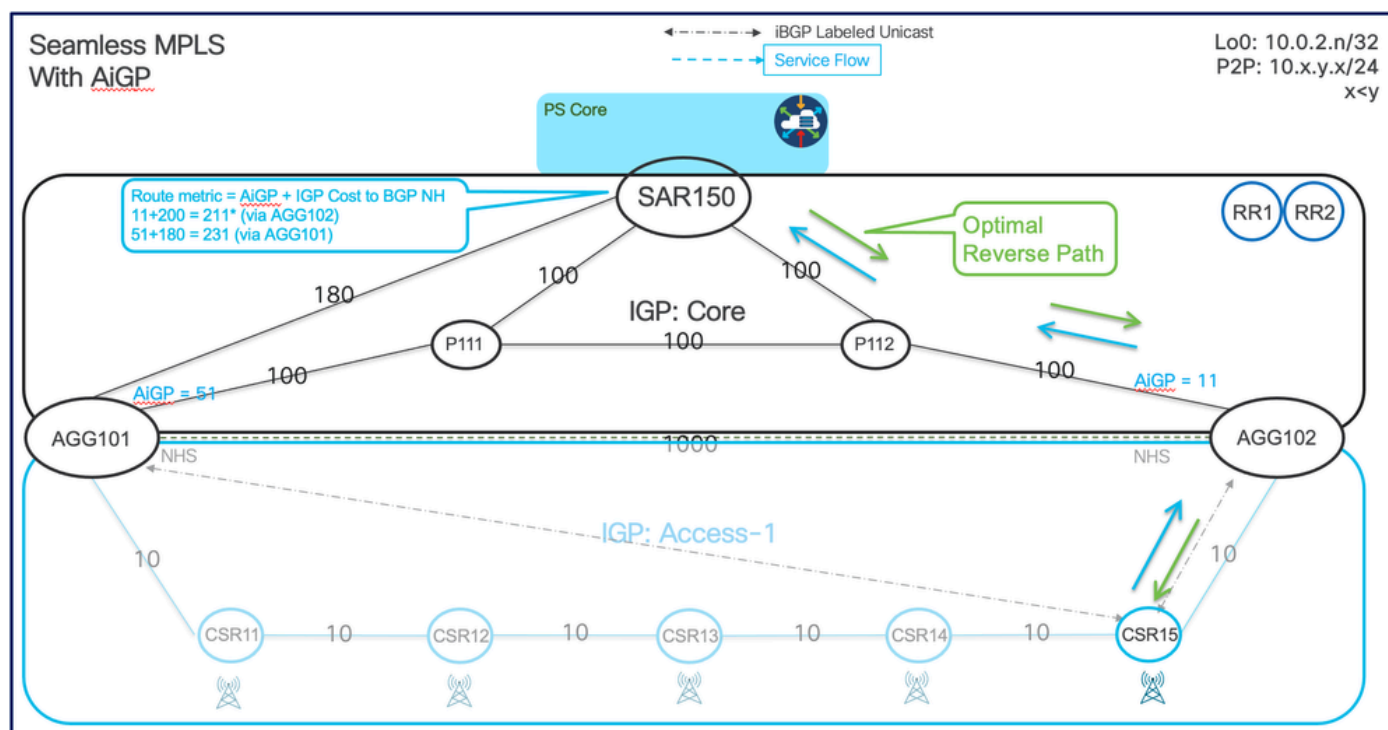
0 10.101.150.150 MRU 1500 [Labels: explicit-null/16015 Exp: 0/0]
L 1 10.101.150.101 MRU 1500 [Labels: 16015 Exp: 0] 10 ms          !!! AGG101
L 2 10.11.101.11 MRU 1500 [Labels: 16015 Exp: 0] 10 ms         !!! CSR11
L 3 10.11.12.12 MRU 1500 [Labels: 16015 Exp: 0] 10 ms         !!! CSR12
L 4 10.12.13.13 MRU 1500 [Labels: 16015 Exp: 0] 20 ms         !!! CSR13
L 5 10.13.14.14 MRU 1500 [Labels: explicit-null Exp: 0] 30 ms  !!! CSR14
! 6 10.14.15.15 30 ms                                          !!! CSR15

```

AIGP解决方案

此处的目标是在SAR路由器和CSR路由器之间建立端到端最佳路径。BGP标记单播(RFC 3107)用于计算从SAR到CSR路由器的距离。每个核心链路上的可用带宽映射到IGP开销，因此BGP必须在每个PE之间正确传输此开销。此功能通过使用AiGP来实现。

带AIP的无缝MPLS网络



带AIP的无线MPLS网络



注意：

1. 假设RR1/2、AGG、CSR和SAR设备上启用了添加路径以通告两个NH。
2. 上述场景中使用的所有设备均基于Cisco IOS-XR。

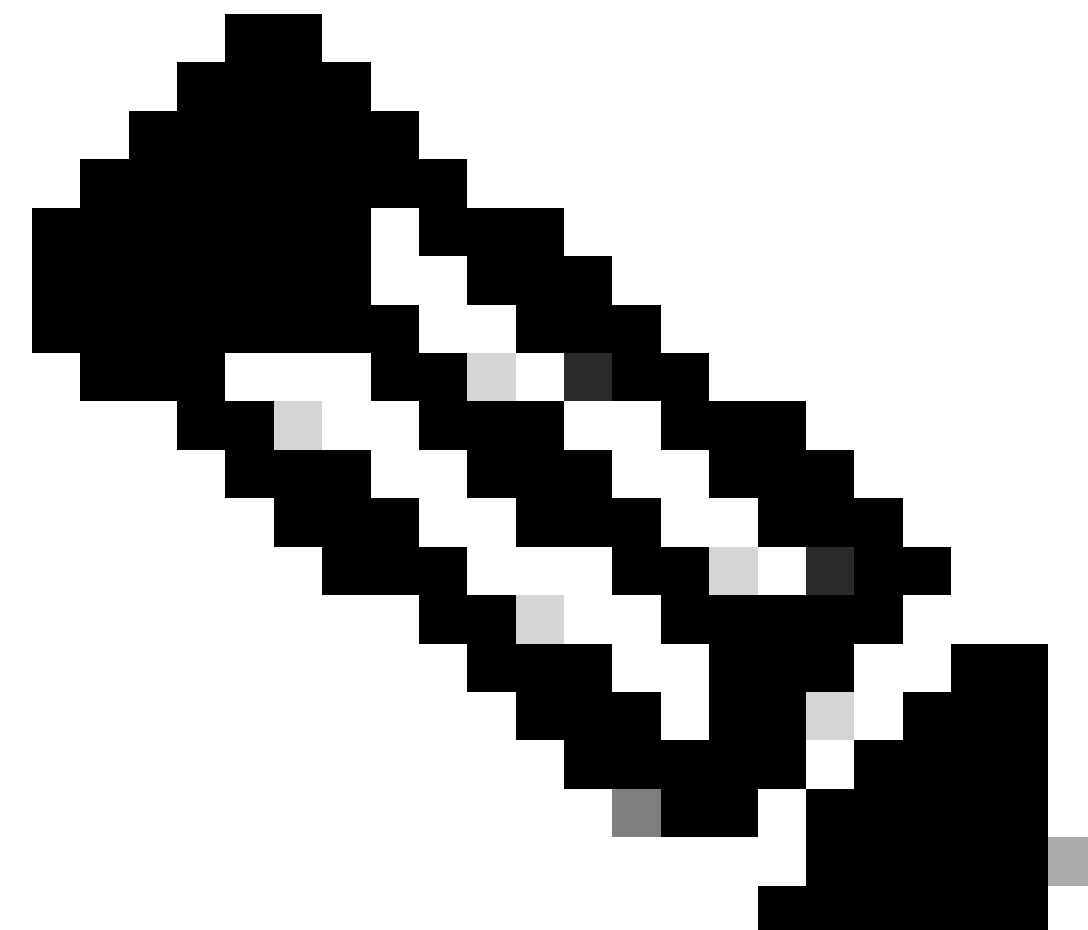
配置

必须在BGP对等体之间商定AiGP路径属性功能。AiGP度量仅包含在启用AiGP的对等体之间的前缀通告中。为单个BGP对等体和特定BGP地址系列配置AIGP功能。

```
router bgp ASN
neighbor <NBR_IP>
address-family ipv4 unicast
aigp [disable]
```

AIGP度量是一个32位 (0到4,294,967,295) 的值。它可以在重分发、通过network语句发起路由时设置，也可以在使用路由映射/路由策略接收前缀时设置。

```
route-policy AIGP_POLICY
  set aigp-metric igp-cost
end-policy
!
router bgp ASN
  address-family {ipv4|ipv6} unicast
    network <NETWORK/MASK> route-policy AIGP_POLICY
    or
    redistribute {ospf|isis} {process-id} route-policy AIGP_POLICY metric VALUE
  !
```



注意：

1. 对于重分发的路由，分配给AIGP属性的值是到路由的iGP下一跳的值或根据路由策略
-

-
- 设置。
2. 对于重分配到BGP中的静态路由，分配的值是路由的下一跳的值或路由策略设置的值。
 3. 该路由通过network语句导入到BGP中。分配的值是路由的下一跳的值或路由策略设置的值。
-

配置示例

```
CSR15:
! Additional config lines related to AIGP are marked in RED color
route-policy SID($SID)
  set label-index $SID
  set aigp-metric igp-cost
end-policy
!
router bgp 1
  address-family ipv4 unicast
    network 10.0.2.15/32 route-policy SID(15)
  neighbor-group RR
    address-family ipv4 labeled-unicast
      aigp
  !
  !
  !
```

注意：所有相应BGP对等设备上均已进行类似配置。

设备日志-在AIGP实施之后

下行流量- SAR150至CSR15

```
RP/0/0/CPU0:SAR150#sh bgp ipv4 labeled-unicast 10.0.2.15/32
BGP routing table entry for 10.0.2.15/32
Versions:
Process bRIB/RIB SendTblVer
Speaker 411 411
Local Label: 16015
Last Modified: Oct 24 11:05:26.796 for 00:00:04
Paths: (2 available, best #1)
Not advertised to any peer
Path #1: Received by speaker 0
Not advertised to any peer
Local
10.0.2.102 (metric 200) from 10.0.2.100 (10.0.2.15)
```

```

Received Label 16015
Origin IGP, metric 0, localpref 100, aigp metric 20, valid, internal, best, group-best, labeled-unicast
Received Path ID 1, Local Path ID 1, version 410
Originator: 10.0.2.15, Cluster list: 10.0.2.100, 10.0.2.102
Total AIGP metric 220
Label-Index: 15
Path #2: Received by speaker 0
Not advertised to any peer
Local
10.0.2.101 (metric 180) from 10.0.2.100 (10.0.2.15)
Received Label 16015
Origin IGP, metric 0, localpref 100, aigp metric 60, valid, internal, backup, add-path, labeled-unicast
Received Path ID 8, Local Path ID 7, version 411
Originator: 10.0.2.15, Cluster list: 10.0.2.100, 10.0.2.101
Total AIGP metric 240
Label-Index: 15

```

```

RP/0/0/CPU0:SAR150#traceroute mpls ipv4 10.0.2.15/32 so 10.0.2.150
Tracing MPLS Label Switched Path to 10.0.2.15/32, timeout is 2 seconds

```

```

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
'L' - labeled output interface, 'B' - unlabeled output interface,
'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
'P' - no rx intf label prot, 'p' - premature termination of LSP,
'R' - transit router, 'I' - unknown upstream index,
'X' - unknown return code, 'x' - return code 0

```

Type escape sequence to abort.

```

0 10.112.150.150 MRU 1500 [Labels: 16102/16015 Exp: 0/0]
L 1 10.112.150.112 MRU 1500 [Labels: explicit-null/16015 Exp: 0/0] 10 ms      !!! P112
L 2 10.102.112.102 MRU 1500 [Labels: explicit-null Exp: 0] 10 ms          !!! AGG102
! 3 10.15.102.15 20 ms                                                    !!! CSR15

```

上行流量- CSR15至SAR150

```

RP/0/0/CPU0:CSR15#traceroute mpls ipv4 10.0.2.150/32 source 10.0.2.15
Tracing MPLS Label Switched Path to 10.0.2.150/32, timeout is 2 seconds

```

```

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
'L' - labeled output interface, 'B' - unlabeled output interface,
'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
'P' - no rx intf label prot, 'p' - premature termination of LSP,
'R' - transit router, 'I' - unknown upstream index,
'X' - unknown return code, 'x' - return code 0

```

Type escape sequence to abort.

```

0 10.15.102.15 MRU 1500 [Labels: explicit-null/16150 Exp: 0/0]
L 1 10.15.102.102 MRU 1500 [Labels: 16150 Exp: 0] 10 ms      !!! AGG102
. 2 *                                                       !!! P112 does not have a route
! 3 10.112.150.150 30 ms                                       !!! SAR150

```

注意事项

- 随着IGP开销的变化（例如，由于链路断开等原因而导致的物理拓扑变化），BGP中也会有相应的变化。这可能导致BGP路由发生故障。此融合还取决于为非关键事件配置的BGP下一跳触发延迟值。
- 通过引入AiGP，可以修改BGP的最佳路径选择。它将BGP AIGP置于BGP本地优先级之后和AS_PATH属性之前。因此，如果存在任何有关AS_PATH预置的现有策略，可能需要重新访问这些策略。
- 当存在两条路径（一条具有AIGP度量，另一条没有）时，BGP始终首选具有AIGP度量的路径。

忽略AIGP

还可以将运行边界网关协议(BGP)的设备配置为在两条路径之间的最佳路径选择过程中，当一条路径没有AIGP度量时，忽略AIGP度量。在路由器配置模式下使用`bgp bestpath aigp ignore`命令要将设备恢复为默认操作，请使用此命令的`no`形式。

```
[no] bgp bestpath aigp ignore
```

默认情况下，BGP始终首选具有AIGP度量的路径。如果有两个路径，一个具有AIGP度量，另一个没有，则执行`bgp bestpath aigp ignore`命令会使BGP执行最佳路径计算，就像两个路径都不具有AIGP度量一样。

结论

BGP AIGP属性确实用于解决某些小生境使用案例，但必须谨慎使用。

相关信息

- [配置BGP的AIGP度量属性](#)
- [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）。