

如何使用AIGP最佳化BGP路徑

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[背景資訊](#)

[累計的內部網關協定屬性](#)

[使用案例示例](#)

[1. 多ASN網路](#)

[組態](#)

[驗證](#)

[2. 無縫MPLS \(單一ASN\)](#)

[觀察](#)

[裝置日誌-初始](#)

[AIGP解決方案](#)

[組態](#)

[配置示例](#)

[裝置日誌-AIGP實施之後](#)

[需要考慮的事項](#)

[忽略AIGP](#)

[結論](#)

[相關資訊](#)

簡介

本檔案介紹邊界閘道通訊協定(BGP)中的累積內部閘道通訊協定(AIGP)測量結果及其使用案例。

必要條件

需求

思科建議您瞭解以下主題：

- BGP
- 無縫多協定標籤交換(MPLS)網路

採用元件

本文件所述內容不限於特定軟體和硬體版本。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除 (預設

) 的組態來啟動。如果您的網路運作中，請確保您已瞭解任何指令可能造成的影響。

背景資訊

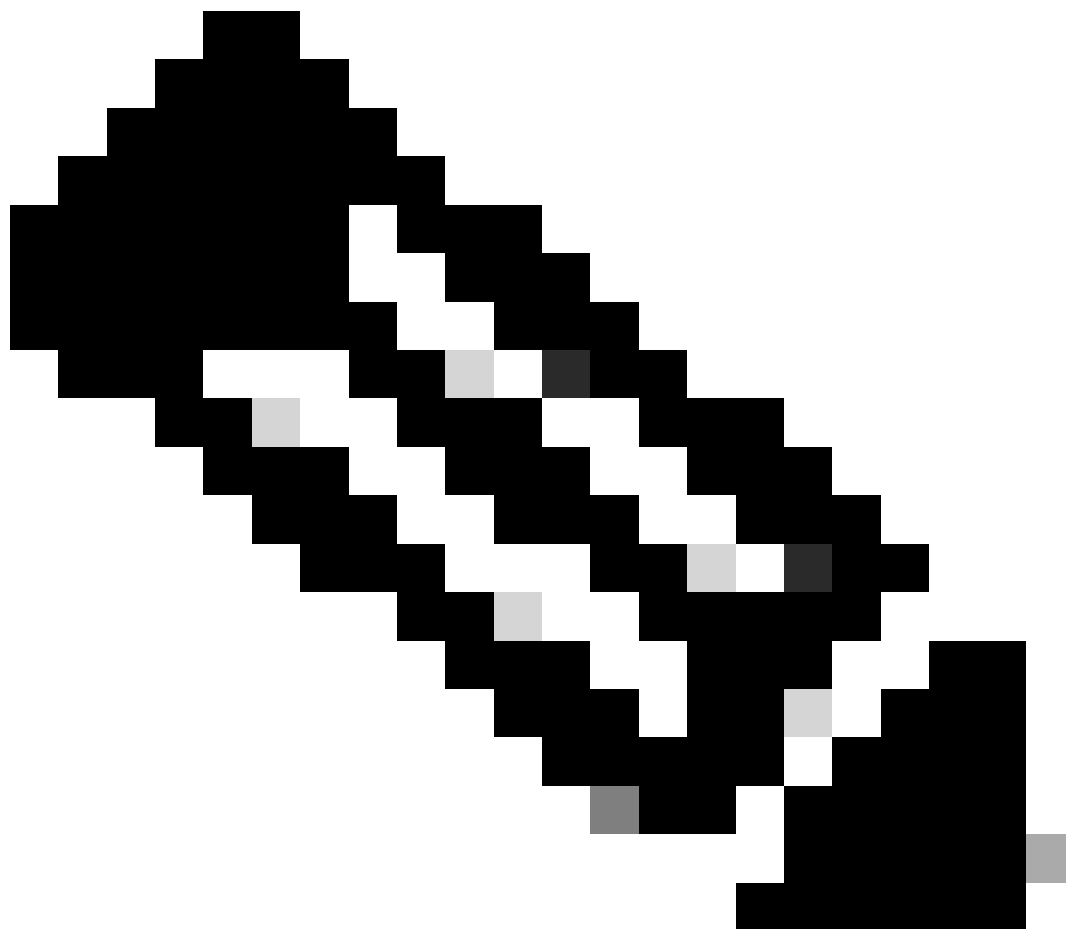
本節提供AIGP指標的概述以及與其使用相關的一些重要注意事項。

如您所知，IGP代表內部網關協定，代表在單個管理域內運行的一組路由協定。IGP根據度量值做出路徑選擇決策。

BGP旨在透過大量獨立的自治系統(AS)提供路由，但相關管理部門之間的協調有限或沒有協調。它不會透過使用度量來決定其路徑選擇。但是，有些部署中，單個管理運行多個連續的BGP網路。在這些情況下，單一管理域內可能是期望BGP根據度量選擇路徑，正如IGP將會做的那樣。

累計的內部網關協定屬性

AIGP指標 (透過RFC7311定義) 是可選的非傳遞性BGP路徑屬性。AIGP屬性的值欄位定義為一組型別/長度/值元素(TLV)。BGP AIGP TLV包含累計IGP指標。



注意：不支援可選非傳遞屬性（例如AIGP）的BGP路由器必須刪除此類屬性，且不得將其傳遞給其他BGP對等體。AIGP度量不打算在完全不同的自治系統之間傳遞（僅跨內部AS邊界）。

使用案例示例

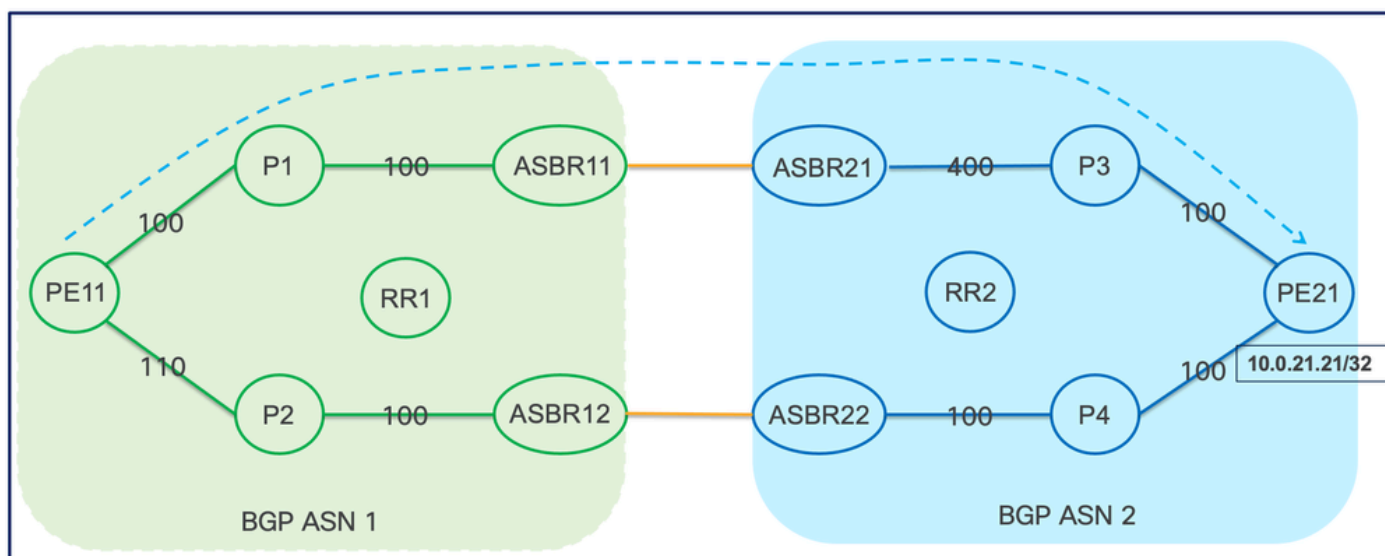
1. 多ASN網路

如今，許多網路位於單個管理域中，由於各種原因，這些網路被細分為多個ASN。這可能有許多原因：

- IGP規模
- 一家提供商網路被另一家網路提供商收購，但他們尚未在內部合併其BGP ASN
- 不同的業務部門在內部都有獨立的網路
- 具有子AS的BGP聯盟
- 無縫MPLS等

在這樣的網路中，允許BGP根據IGP指標做出決策很有用，這樣，即使兩個節點位於兩個不同的ASN中，BGP也會選擇兩個節點之間的最短端到端路徑。

例如：ABC網路，再細分為兩個BGP ASN，即ASN 1和ASN 2。它們在ASBR對等，鏈路IGP成本代表頻寬。此處的目標是在PE11和PE21之間建立端到端的最佳路徑。



沒有AIGP的多ASN網路

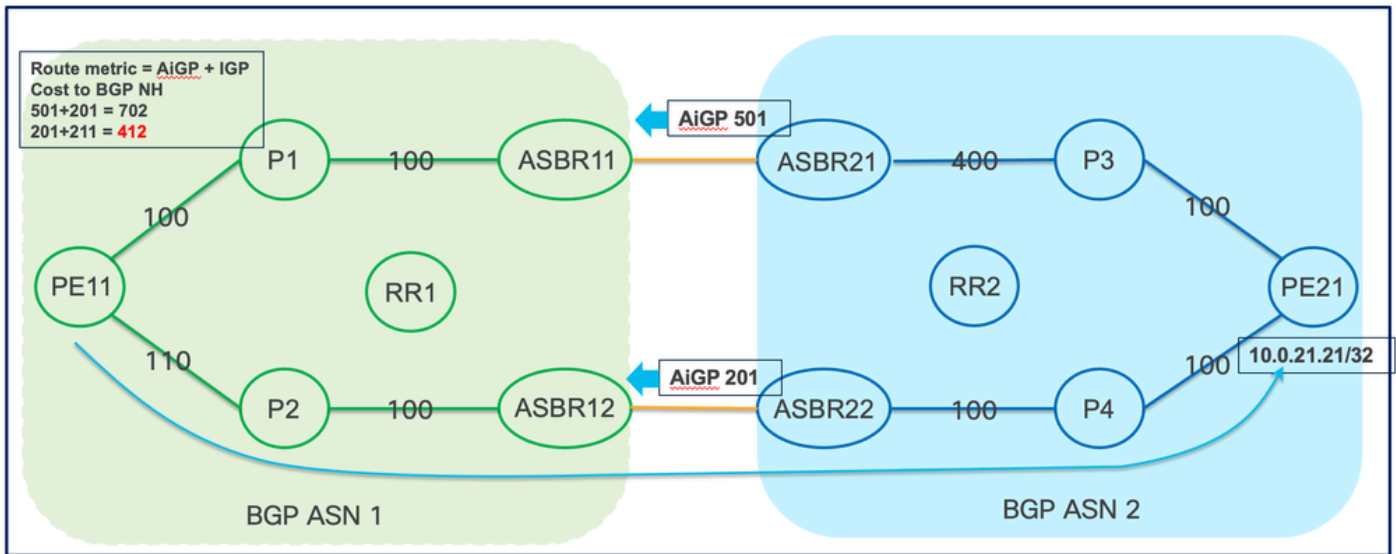
附註：

1. 假設在RR1/2上啟用了Add-path，以將兩個NH通告給PEX。
2. 上述方案中使用的裝置均使用Cisco IOS-XE運行。

```
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

在拓撲中啟用AiGP後 (在PE11、PE32、ASBR1x、ASBR2x、RR1、RR2上) , PE11現在將選擇端到端IGP成本最低的路徑。



具有AiGP的多ASN網路

組態

PE_x、ASBR_x、RR_N :

AiGP功能配置 :

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



注意：BGP對等丟棄並重新建立以協商此新功能。因此，建議您在維護時段內執行它。

通告字首的AIGP指標。

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
!
```

驗證

```

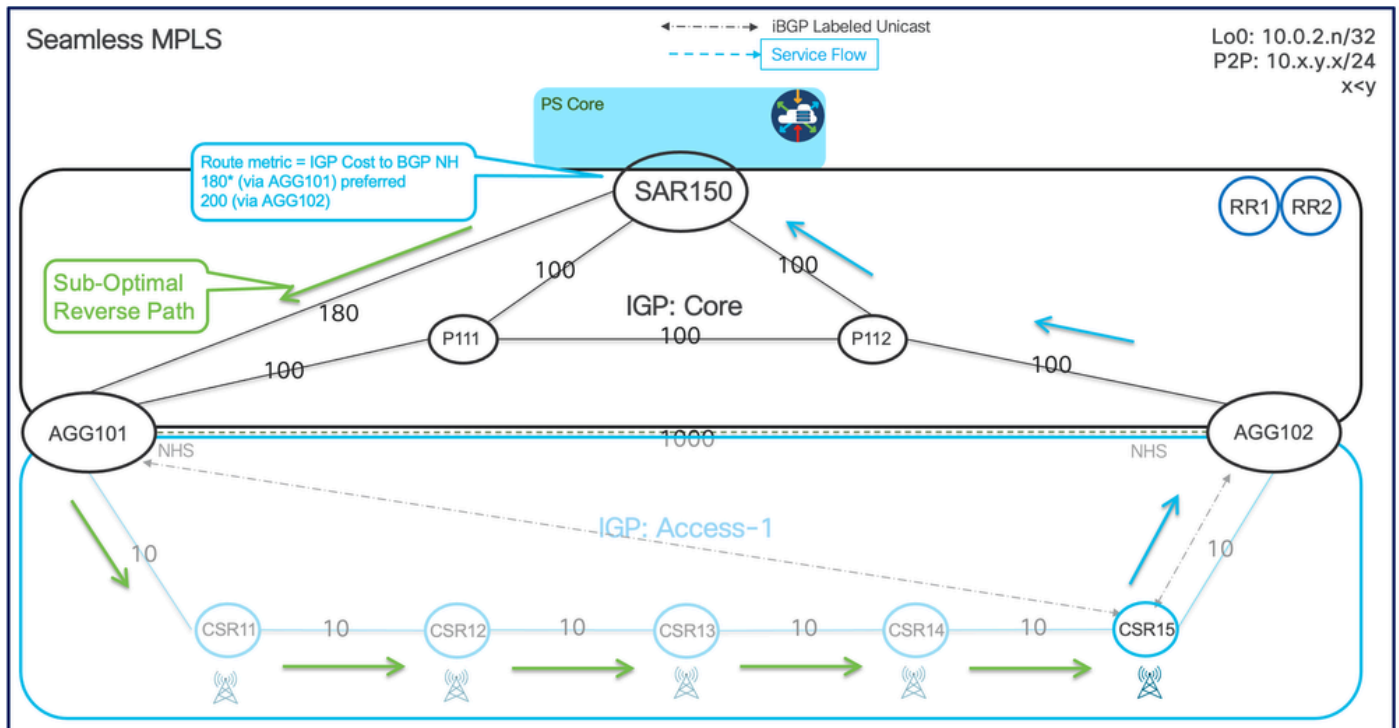
PE11#sh bgp ipv4 unicast 10.0.21.21/32
BGP routing table entry for 10.0.21.21/32, version 21
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 3
  2
    192.168.0.11 (metric 201) from 192.168.11.11 (192.168.11.11)
      Origin IGP, aigp-metric 501, metric 0, localpref 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. 無縫MPLS (單一ASN)

在大型服務提供商核心網路中，傳輸網路通常被細分為不同的IGP域，使用BGP標籤的單播進行縫合，以提供端到端標籤交換路徑(LSP)。邊界路由器在BGP LU AF中執行下一跳自身(NHS)。

IGP/LDP僅在本地/域中攜帶字首/標籤資訊。然後，BGP透過重新將路由分配到區域邊界的BGP，將字首/標籤傳送給所有遠端區域/域。然後使用LSP通告路由/標籤。路由的下一跳在每個ABR上更改為本地路由器，這樣便無需跨區域/域邊界洩漏IGP路由。

在此拓撲圖中，有一個BGP域被分為兩個IGP域 (CORE和Access-1)。每條鏈路旁邊顯示的數字代表該鏈路的IGP開銷/度量。



不使用AIGP的無縫MPLS網路

挑戰：從PS-Core到eNB/gNB (連線到CSR15) 的下行流量與從eNB/gNB (連線到CSR15) 到PS-

Core的上行流量相比，採用的是非對稱次優路徑，這導致了移動流量中的延遲問題。

觀察

1. 主要在地理邊界場景中觀察，在該場景中，同一個聚合路由器充當多個接入域（例如，前面提到的影象中的AGG102）的通用邊界路由器。
2. 對於向上流量，信元站點路由器(CSR)路由器選擇最近的邊界路由器。例如，CSR15選擇AGG102作為NextHop。
3. 對於下行流量，服務聚合路由器(SAR)也會選擇最近的邊界路由器。例如，SAR150會選取AGG101 (cost 180 < 200)。

裝置日誌-初始

上行流量- CSR15到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          !!!! AGG102
. 2 *                                                         !!!! P112 does not have a route t
! 3 10.112.150.150 20 ms                                       !!!! 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]
```



```

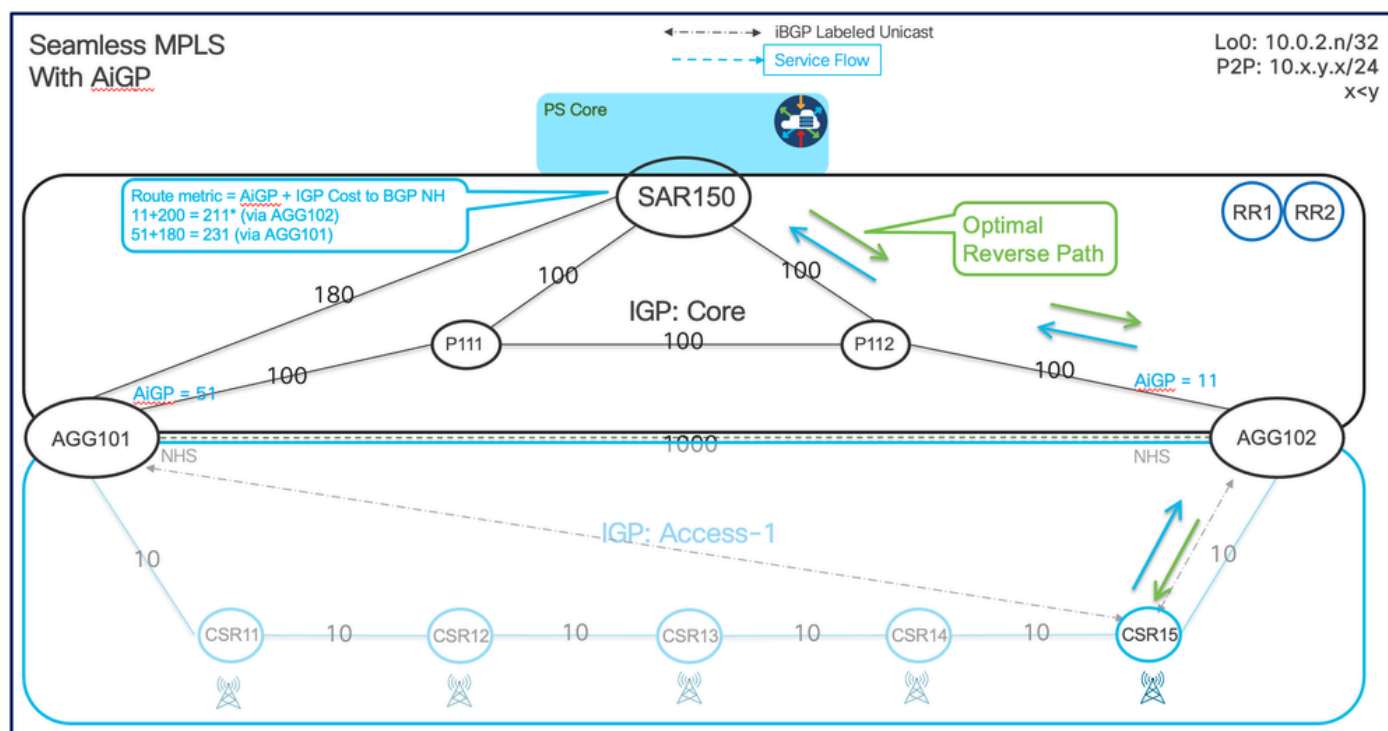
L 1 10.101.150.101 MRU 1500 [Labels: 16015 Exp: 0] 10 ms          !!! AGG101
L 2 10.11.101.11 MRU 1500 [Labels: 16015 Exp: 0] 10 ms         !!! CSR11
L 3 10.11.12.12 MRU 1500 [Labels: 16015 Exp: 0] 10 ms         !!! CSR12
L 4 10.12.13.13 MRU 1500 [Labels: 16015 Exp: 0] 20 ms         !!! CSR13
L 5 10.13.14.14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

```

AIGP解決方案

此處的目標是在SAR路由器和CSR路由器之間建立端到端最佳路徑。BGP標籤的單播(RFC 3107)用於計算SAR到CSR路由器的距離。每個核心鏈路上的可用頻寬都對映到IGP開銷，因此BGP必須在每個PE之間正確傳輸此開銷。使用AiGP可實現此功能。

具有AIGP的無縫MPLS網路



具有AIGP的無縫MPLS網路



附註：

1. 假設在RR1/2、AGG、CSR和SAR裝置上啟用了增加路徑，以通告兩個NH。
2. 上述場景中使用的所有裝置均基於Cisco IOS-XR。

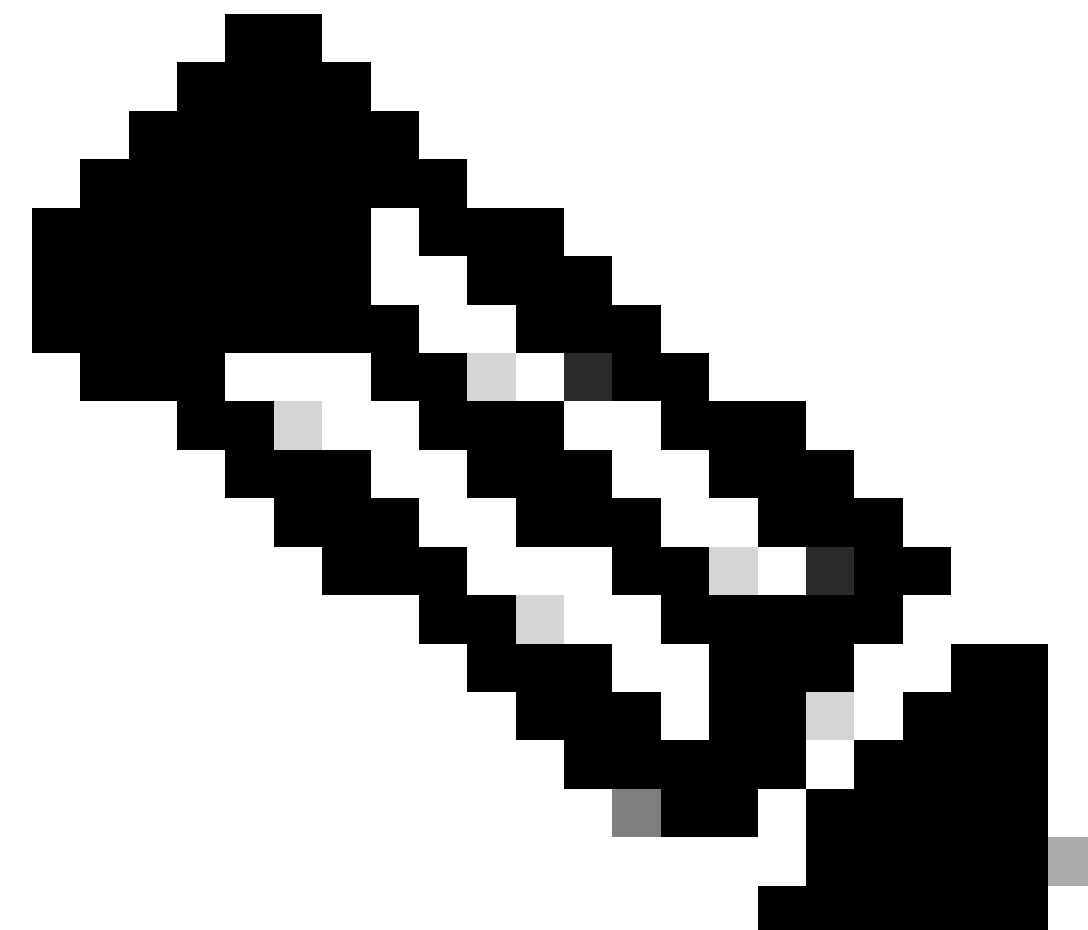
組態

AiGP路徑屬性功能必須在BGP對等體之間達成一致。AiGP度量僅包含在啟用AiGP的對等體之間的字首通告中。AIGP功能是為了個別BGP對等體和特定BGP地址系列配置的。

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

AIGP指標是32位元 (0到4,294,967,295) 的值。它可在重分配、透過network語句路由發起或接收帶有路由對映/路由策略的字首時設定。

```
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
!
```



附註：

1. 對於重分發的路由，分配給AIGP屬性的值是該路由的iGP下一跳的值或路由策略設定
-

-
- 的值。
2. 對於重分配到BGP中的靜態路由，分配的值是路由的下一跳值或路由策略設定的值。
 3. 該路由透過network語句導入到BGP中。分配的值是路由的下一跳的值或路由策略設定的值。
-

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：已在所有相應BGP對等裝置上進行了類似的配置。

裝置日誌- AIGP實施之後

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到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

```

需要考慮的事項

- 隨著IGP開銷的變化（例如，由於鏈路斷開等原因物理拓撲的變化），BGP中也有相應的變化。這可能會導致出現BGP路由抖動。此收斂還取決於為非關鍵事件配置的BGP下一跳觸發器延遲值。
- 透過引入AIGP，BGP的最佳路徑選擇將被修改。它將BGP AIGP置於BGP本地優先順序之後和AS_PATH屬性之前。因此，如果存在任何有關AS_PATH預置的策略，則可能需要重新檢視這些策略。
- 當有兩條路徑（一條具有AIGP指標，另一條沒有）時，BGP一律會偏好使用AIGP指標的路徑。

忽略AIGP

當一條路徑沒有AIGP指標時，運行邊界網關協定(BGP)的裝置也可以配置為在兩條路徑之間的最佳路徑選擇過程中忽略AIGP指標。在路由器配置模式下使用`bgp bestpath aigp ignore`命令。要將裝置恢復為預設操作，請使用此命令的`no`形式。

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

預設情況下，BGP一律優先使用具有AIGP指標的路徑。如果有兩個路徑，一個具有AIGP指標，另一個沒有，則執行`bgp bestpath aigp ignore`命令會使BGP執行最佳路徑計算，就像兩個路徑都不具有AIGP指標一樣。

結論

BGP AIGP屬性當然是為解決某些小眾使用情形而開發的，但必須謹慎使用。

相關資訊

- [設定BGP的AIGP指標屬性](#)
- [思科技術支援與下載](#)

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。