

How to Optimize BGP Path using AIGP

Contents

[Introduction](#)

[Prerequisites](#)

[Requirements](#)

[Components Used](#)

[Background Information](#)

[Accumulated Interior Gateway Protocol Attribute](#)

[Sample use-cases](#)

[1. Multi-ASN Network](#)

[Configuration](#)

[Verification](#)

[2. Seamless MPLS \(Single ASN\)](#)

[Observations](#)

[Device Logs - Initial](#)

[AIGP Solution](#)

[Configuration](#)

[Sample Configuration](#)

[Device Logs - After AIGP Implementation](#)

[Things to Consider](#)

[Ignore AIGP](#)

[Conclusion](#)

[Related Information](#)

Introduction

This document describes what Accumulated Interior Gateway Protocol (AIGP) metric in Border Gateway Protocol (BGP) is and its use cases.

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- BGP
- Seamless Multiprotocol Label Switching (MPLS) network

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 our network is live, make sure that you understand the potential impact of any command.

Background Information

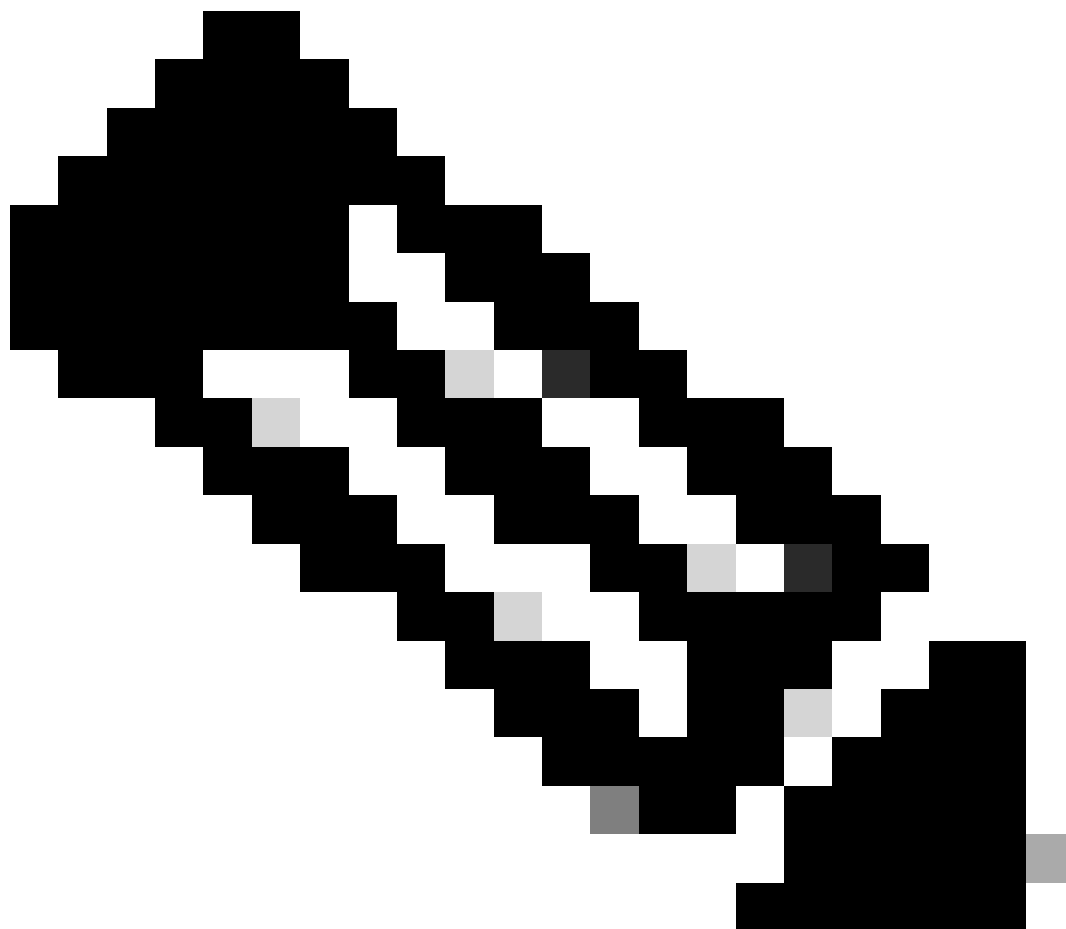
This section provides an overview of the AIGP metric and some important considerations regarding its use.

As you know, IGP stands for Interior Gateway Protocol and represents a group of routing protocols that run within a single administrative domain. IGP make a path-selection decision based on metric value.

BGP is designed to provide routing over a large number of independent Autonomous Systems (AS) with limited or no coordination among respective administrations. It does not make its path-selection decisions through the use of a metric. However, there are deployments in which a single administration runs several contiguous BGP networks. In such cases, it can be desirable within that single administrative domain, for BGP to select paths based on a metric, just as an IGP would do.

Accumulated Interior Gateway Protocol Attribute

The AIGP metric (defined via RFC7311) is an Optional Non-transitive BGP Path Attribute. The value field of the AIGP Attribute is defined as a set of Type/Length/Value elements (TLVs). The BGP AIGP TLV contains the Accumulated IGP Metric.



Note: BGP routers that do not support the optional non-transitive attributes (for example, AIGP)

must delete such attributes and must not pass them to other BGP peers. AIGP metric is not intended to be transitive between completely distinct autonomous systems (only across internal AS boundaries).

Sample use-cases

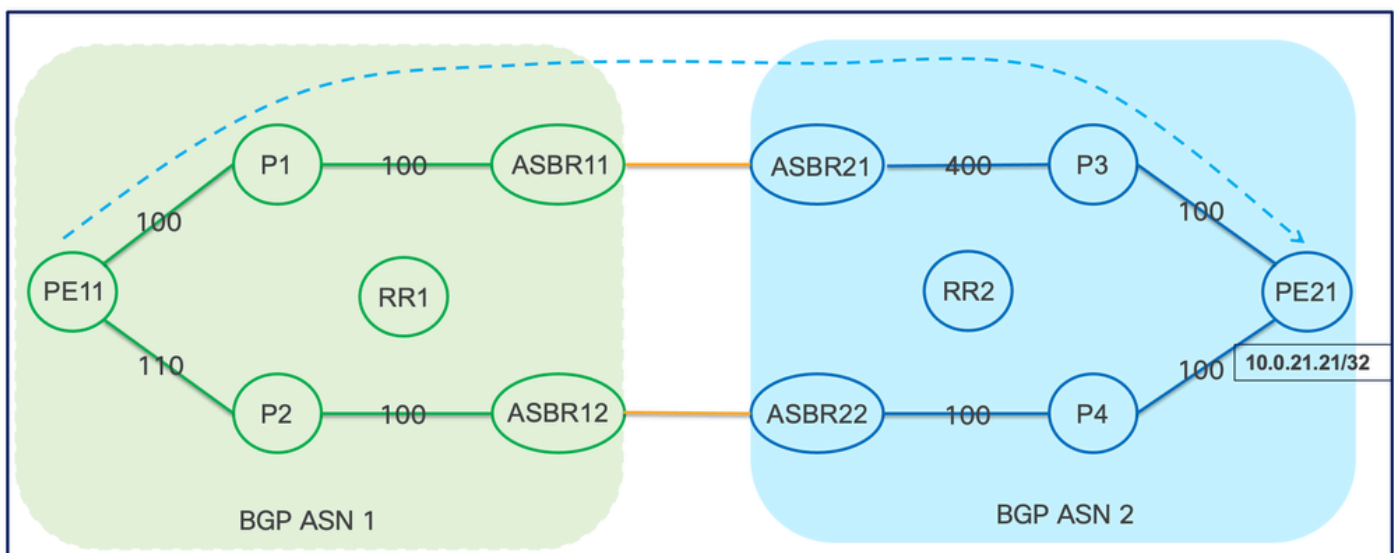
1. Multi-ASN Network

Today, there are many networks which are in a single administrative domain, which are subdivided into multiple ASNs for various reasons. There could be many possible reasons for that:

- IGP scale
- One provider network was bought by another network provider but still, they haven't merged their BGP ASNs internally
- Different business divisions have their separate network internally
- BGP confederations with Sub-AS
- Seamless MPLS etc.

In networks like these, it can be useful to allow BGP to make its decisions based on the IGP metric, so that BGP chooses the shortest end-to-end path between two nodes, even if the nodes are in two different ASNs.

For example: ABC network, which is subdivided into two BGP ASNs, ASN 1 and ASN 2. They are peering at ASBR and the link IGP costs are representing bandwidth. The goal here is to have an end-to-end optimal path between PE11 and PE21.



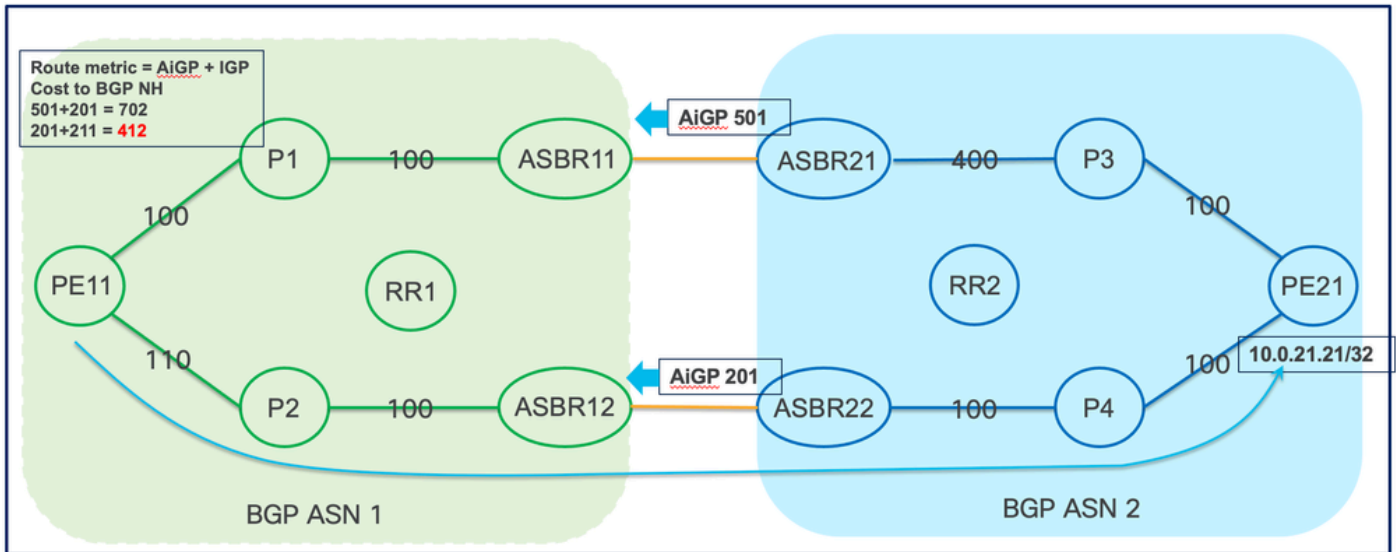
Multi-ASN Network without AIGP

Note:

1. Assuming that Add-path is enabled on RR1/2 to advertise both the NH towards PEx.
2. The devices used in the previously mentioned scenario are all running with Cisco IOS-XE.

```
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
```

With AiGP enabled in the topology (on PE11, PE32, ASBR1x, ASBR2x, RR1, RR2), PE11 now chooses the path with the lowest end-to-end IGP cost.



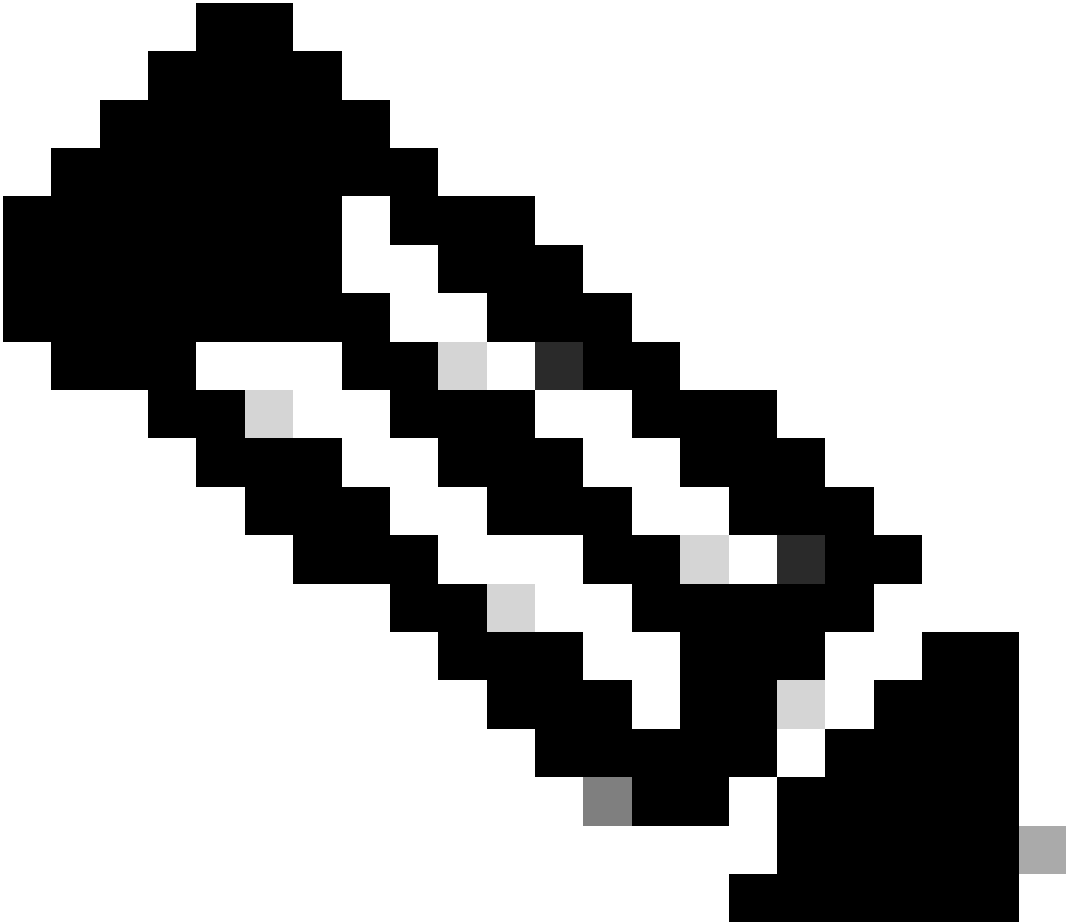
Multi-ASN Network with AIGP

Configuration

PE_x, ASBR_x, RR_n:

AIGP capability configuration:

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



Note: BGP peering drops and re-establishes to negotiate this new capability. So, it is advised to perform it in a maintenance window.

Advertise AIGP metric for a prefix.

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
!
```

Verification

```

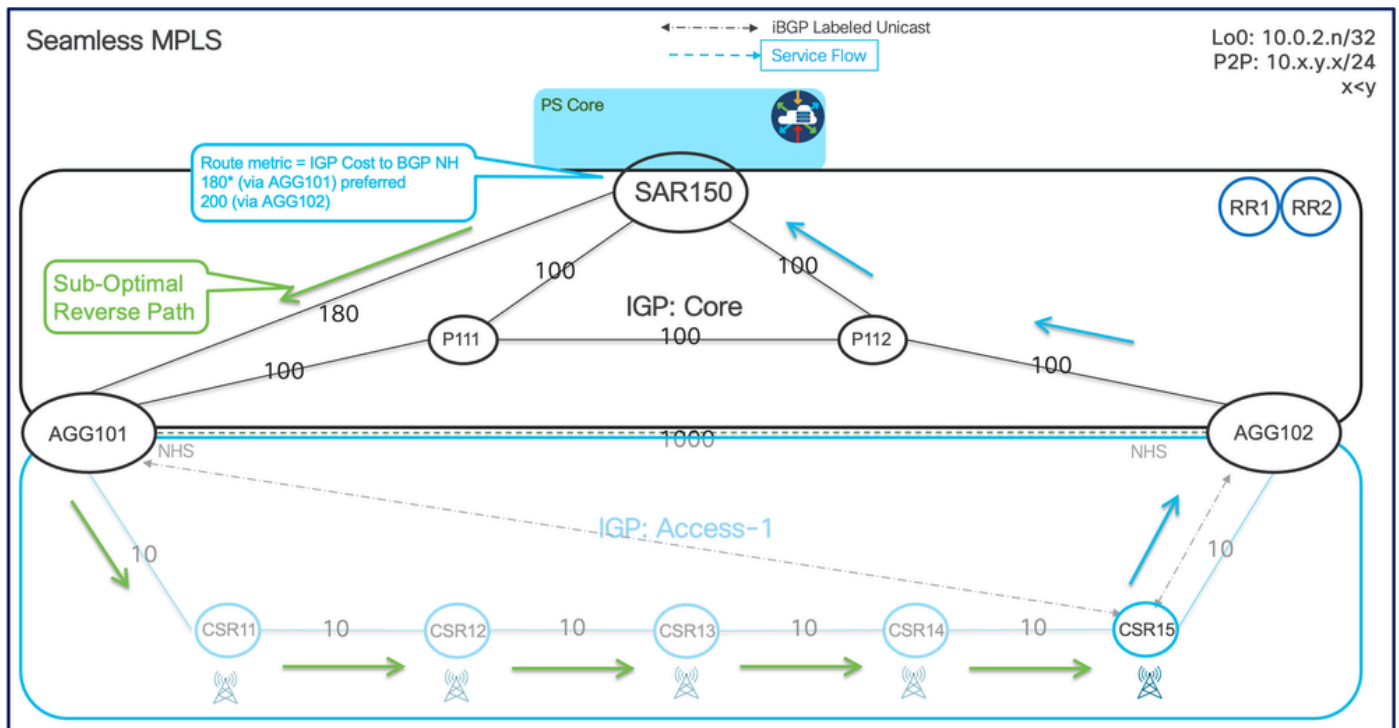
PE11#sh bgp ipv4 unicast 10.0.21.21/32
BGP routing table entry for 10.0.21.21/32, version 21
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 3
  2
    192.168.0.11 (metric 201) from 192.168.11.11 (192.168.11.11)
      Origin IGP, aigp-metric 501, metric 0, localpref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

In a large Service Provider Core network, the Transport Network is usually sub-divided into different IGP domains, stitched together using BGP Labeled Unicast to provide end-to-end Labeled Switched Path (LSP). Border routers perform Next Hop Self (NHS) in BGP LU AF.

IGP/LDP carries the prefix/label information only in the local area/domain. Then, BGP carries the prefix/label to all the remote areas/domains by redistributing the routes into BGP at area boundaries. The routes/labels are then advertised using LSPs. The next hop for the route is changed at each ABR to local router which removes the need to leak IGP routes across area/domain boundaries.

In this topology diagram, there is a single BGP domain divided into 2 IGP domains (CORE and Access-1). The number shown beside each link represents the IGP cost/metric of that link.



Seamless MPLS Network without AIGP

Challenge: Downwards traffic from PS-Core to eNB/gNB (connected to CSR15) is taking an Asymmetric

and Sub-Optimal path as compared to upward traffic from eNB/gNB (connected to CSR15) towards PS-Core, which is causing Latency issues in Mobility traffic.

Observations

1. Primarily observed in Geographical border scenarios where the same Aggregation router is acting as a common border router for multiple Access domains (for example, AGG102 in the previously mentioned image).
2. For upward traffic, the Cell-Site Router (CSR) router is choosing the nearest border router. For example, CSR15 selects AGG102 as NextHop.
3. For Downward traffic, the Service Aggregation Router (SAR) also chooses the nearest border router. For example, SAR150 selects AGG101 (cost 180 < 200).

Device Logs - Initial

Upstream traffic - CSR15 to 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
```

Downstream traffic - SAR150 to 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
```



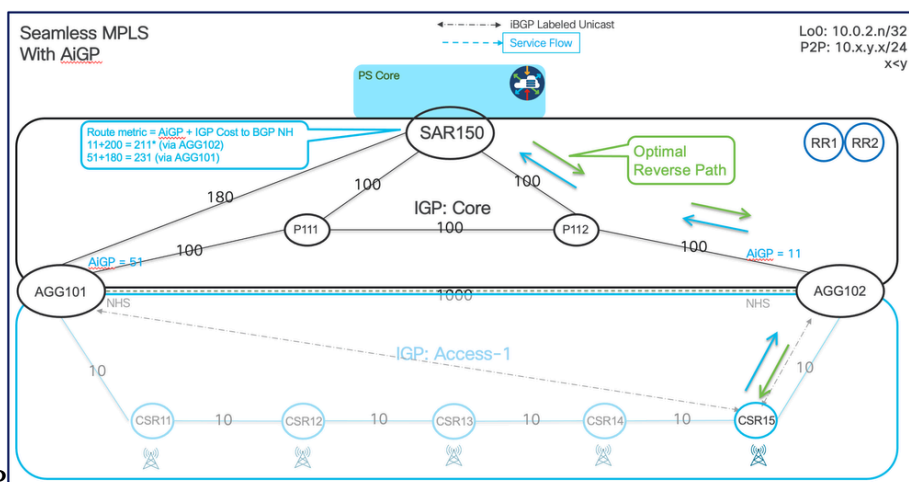
```

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 Solution

The goal here is to have an end-to-end optimal path between SAR routers and CSR routers. BGP Labeled Unicast (RFC 3107) is being used to compute the distance from SAR to CSR routers. The bandwidth available on each of the core links is mapped to IGP cost, hence BGP must carry this cost correctly between each of the PEs. This functionality is achieved by using the AiGP.



Seamless MPLS Network with AIGP

Seamless MPLS Network with AIGP

Note:

1. Assuming that Add-path is enabled on RR1/2, AGG, CSR, and SAR devices to advertise both the NH.
2. All devices used in the previously mentioned scenario are Cisco IOS-XR-based.

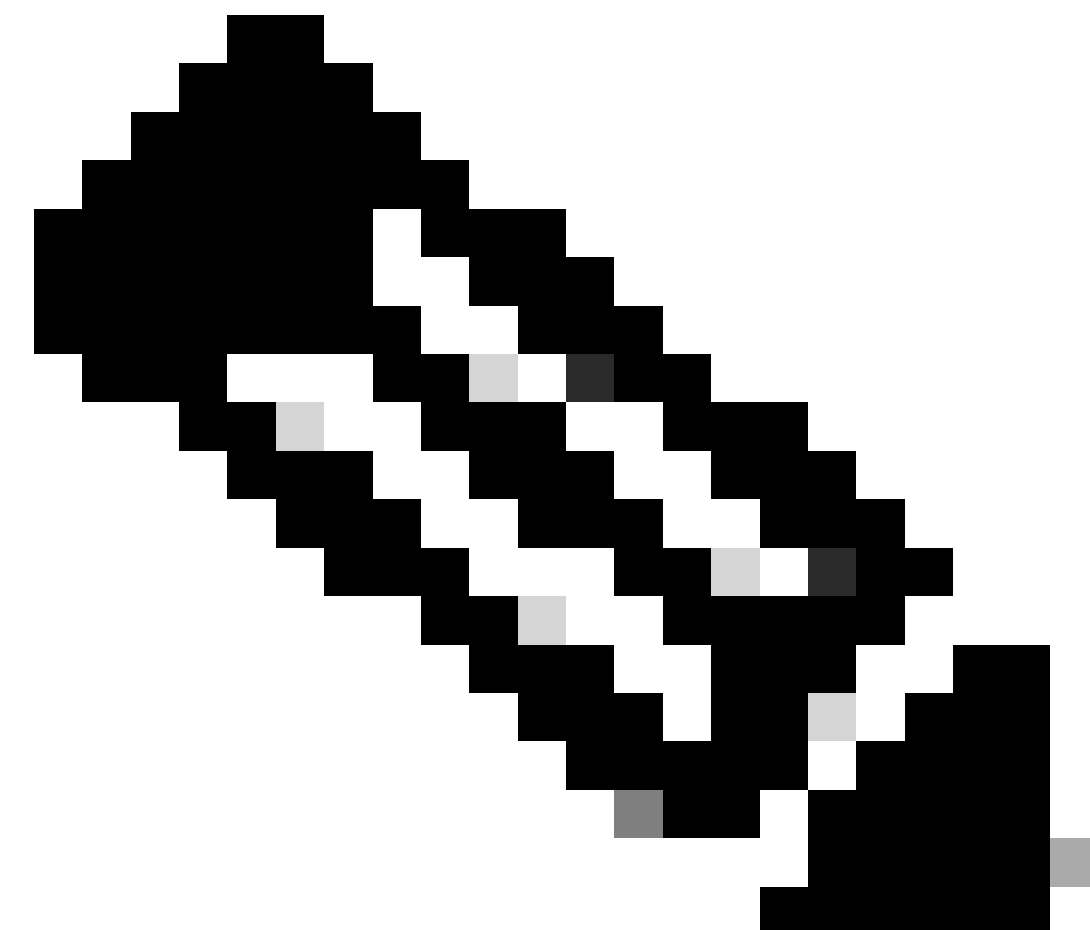
Configuration

AiGP Path Attribute capability must be agreed upon between the BGP peers. AiGP metrics are only included in prefix advertisements between AiGP-enabled peers. AIGP capability is configured for an individual BGP peer and a specific BGP address family.

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

AIGP metric is a 32-bit (0 to 4,294,967,295) value. It can be set during redistribution, route origination via network statement or during receipt of a prefix with a route map/route-policy.

```
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
!
```



Note:

1. For redistributed routes, the value assigned to the AiGP attribute is the value of iGP next hop to the route or as set by a route policy.
 2. For static routes redistributed into BGP, the value assigned is the value of the next hop to the
-

-
- route or as set by a route policy.
3. The route is imported into BGP through a network statement. The value assigned is the value of the next hop to the route or as set by a route policy.
-

Sample Configuration

```
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
    !
  !
!
```

Note: A similar configuration on all the respective BGP peering devices has been done.

Device Logs - After AIGP Implementation

Downstream traffic - SAR150 to 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

```

Upstream traffic - CSR15 to 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

```

Things to Consider

- With changes in IGP cost (for example, change in physical topology because of link cut and more), there are corresponding changes in BGP as well. This might cause BGP route churn. This convergence also depends upon the BGP Next Hop Trigger delay value configured for non-critical events.
- By Introducing AiGP, BGP's best path selection is modified. It puts BGP AIGP after BGP local preference and before the AS_PATH attribute. So, if there are any existing policies about AS_PATH prepending, they might need to be revisited.
- When there are two paths, one with the AIGP metric and the other without, BGP always prefers a path with the AIGP metric.

Ignore AIGP

A device that is running the Border Gateway Protocol (BGP) can also be configured to ignore the AIGP metric during the best path selection process between two paths when one path does not have the AIGP metric. Using the `bgp bestpath aigp ignore` command in router configuration mode. To return the device to default operation, use the no form of this command.

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

By default, BGP always prefers a path with the AIGP metric. When there are two paths, one with the AIGP metric and the other without, then executing the `bgp bestpath aigp ignore` command results in BGP performing best path computation as if neither path has the AIGP metric.

Conclusion

BGP AIGP attribute is certainly developed to solve certain niche use cases but it must be used cautiously.

Related Information

- [Configure the AIGP Metric Attribute for BGP](#)
- [Cisco Technical Support & Downloads](#)