Deploy BGP Soft Next-Hop in Cisco IOS XR

Contents

Introduction

This document describes the behavior of Next-Hop treatment by Border Gateway Protocol (BGP) in Cisco IOS® XR. BGP requires the Next-Hop (NH) of a path to be reachable before it installs the path in Routing Information Base (RIB). This rule applies to all BGP speakers. This is the next-hop validation check. The BGP Soft Next-Hop feature ensures that there is no longer a need for the BGP next-hop to be reachable in the RIB.

Prerequisites

Requirements

There are no specific requirements for this document.

Components Used

This document is specific to Cisco IOS XR.

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.

Background Information

In a single Autonomous System (AS) network, multi-domain network, or Inter-AS scenario, there could be no reachability to the NH if it is not redistributed between the domains or Autonomous Systems.

The problem is not confined to headend Provider Edge (PE) routers, but also to the intermediate BGP speakers (for example Route-Reflector (RR) and Autonomous System Border Router (ASBR)) between egress and ingress PEs. An intermediate BGP speaker must have reachability to NH before it installs and propagates a route.

On-Demand Next-hop (ODN) is a Segment Routing (SR) application that installs SR policies on the router. The service routes attached to these SR policies can be BGP routes. These BGP routes can only be installed in the RIB and Cisco Express Forwarding (CEF) table if the next-hop is valid. There are designs like Seamless MPLS or Inter-AS MPLS Virtual Private Network (VPN) where the reachability to the BGP next-hop in another part of the network such as another area or another domain are not guaranteed by a route in the RIB. This is not an issue if the reachability is guaranteed by a controller or SR Path Computation Element (SR-PCE) that provides the reachability to the network elements throughout the network.

Currently, the BGP service route can only use the SR policy if the next-hop of the BGP route is in the RIB as a non-default route.

If the BGP speaker with the SR policy does not have a route (other than the default route) in the RIB for the BGP next-hop, then a workaround can be used. The workaround is to configure a specific (non-default) static route to null0 covering those unreachable NHs, inject the routes via BGP-LU, or redistribute them between IGP domains.

This is cumbersome and/or impacts the scalability.

Solution

The PE (headend) receives colored BGP L3VPN prefixes. It could learn the SR policy locally or request ODN SR policy for color and next-hop.

If NH validation is configured, BGP does soft validation of NH and applies the NH AD/metric when the command is enabled. For colored NH, the AD/metric comes from the SR controller. The soft validation of the next-hop means that there is no check for RIB reachability, but the check is performed on the SR policy information. This includes the SR policy route type and the admin distance and the metric value for that metric type.

A new command is introduced to do this soft next-hop validation on the headend router or the RR.

A new command is introduced for the RR, to skip the next-hop reachability validation for color-extcomm paths.

A new command is introduced for the RR so that SR policy is not used for BGP best-path calculation.

The feature was introduced in Cisco IOS XR releases 7.3.2 and 7.4.1.

Next-Hop Inaccessible for BGP Route

A BGP route with an inaccessible next-hop is not advertised.

This route is a VPNv4 route on a RR. Its next-hop (PE loopback) is inaccessible because there is no route for the next-hop address in the routing table.

```
RP/0/RP0/CPU0:RR#show bgp vpnv4 unicast rd 65001:2 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65001:2
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 0           0
Last Modified: Oct 26 10:40:12.136 for 00:03:07
Paths: (1 available, no best path)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002, (Received from a RR-client)
   10.0.0.5 (inaccessible) from 10.0.0.5 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, not-in-vrf
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
```

The BGP VPNv4 route is not advertised as a result.

```
RP/0/RP0/CPU0:RR#show route 10.0.0.5
Routing entry for 0.0.0.0/0
 Known via "isis 1", distance 115, metric 20, candidate default path, type level-1
 Installed Oct 25 09:35:07.256 for 1d01h
 Routing Descriptor Blocks
   10.2.7.2, from 10.0.0.3, via GigabitEthernet0/0/0/0
     Route metric is 20
```

```
 No advertising protos.
```

The current workaround is to configure a static route that covers the PE loopback addresses on the headend router. This is an example of such a static route to null0.

```
address-family ipv4 unicast
 10.0.0.0/24 Null0
 !
!
```

This static route to Null0 creates reachability in the RIB for all remote PE loopback addresses (the BGP next-hop addresses). This static route covers all addresses in the range 10.0.0.0 – 10.0.0.255.

The next-hop is resolved through the static route. You can see this with this command.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast nexthops 10.0.0.5 color 101
Nexthop: 10.0.0.5 C:101
 VRF: default
 Nexthop ID: 0x6000008, Version: 0x0
 Nexthop Flags: 0x00480002
 Nexthop Handle: 0x7fa734042e94

 RIB Related Information:
 Firsthop interface handle 0x0000000c
   Gateway TBL Id: 0xe0000000    Gateway Flags: 0x00000080
   Gateway Handle: 0x7fa7988c7ce8
   Gateway: reachable, non-Connected route, prefix length 24
   Resolving Route: 10.0.0.0/24 (static)
   Paths: 0
   RIB Nexhop ID: 0x0
   Status: [Reachable][Connected][Not Local]
   Metric: 0
   ORR afi bits: 0x0
   Registration: Synchronous, Completed: 01:22:27
   Events: Critical (0)/Non-critical (0)
   Last Received: 01:22:27 (Registration)
   Last gw update: (Crit-sync) 01:22:27(rib)
   Reference Count: 4

 Prefix Related Information
   Active Tables: [IPv4 Unicast][VPNv4 Unicast]
   Metrices: [0x0][0x0]
   Reference Counts: [0][4]
 Interface Handle: 0x0
 Attr ref-count: 7
 SR policy color 101, State: [Up]
   Not registered, bsid 24009
   Skip Reg on restart [No]
   First notif received [Yes]
   SR Policy Flags [0x2]
   BGP TE registered [No]
   ODN registered [No]

   IPv6 capability required/enabled: Yes/Yes
   Last SR policy update: 01:22:35
```

If an SR policy is used for the validation of the next-hop, then you see this output:

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast nexthops 10.0.0.5 color 101
```

```
Nexthop: 10.0.0.5 C:101
 VRF: default
 Nexthop ID: 0x6000008, Version: 0x0
 Nexthop Flags: 0x00480000
 Nexthop Handle: 0x7fa734042e94

 RIB Related Information:
 Firsthop interface handle 0x00000000
   Gateway TBL Id: 0xe0000000    Gateway Flags: 0x00000080
   Gateway Handle: 0x7fa7988c7ce8
   Gateway: unreachable, non-Connected route, prefix length 8192
   Resolving Route: 10.0.0.0/24 (static)
   Paths: 0
   RIB Nexhop ID: 0x0
   Status: [Unreachable]
   Metric: 4294967295
   ORR afi bits: 0x0
   Registration: Synchronous, Completed: 01:25:30
   Events: Critical (1)/Non-critical (0)
   Last Received: 00:00:43 (Critical)
   Last gw update: (Crit-notif) 00:00:43(rib)
   Reference Count: 2

 Prefix Related Information
   Active Tables: [IPv4 Unicast][VPNv4 Unicast]
   Metrices: [0xffffffff][0xffffffff]
   Reference Counts: [0][2]
 Interface Handle: 0x0
 Attr ref-count: 5
 SR policy color 101, State: [Up]
   Not registered, bsid 24009
   Skip Reg on restart [No]
   First notif received [Yes]
   SR Policy Flags [0x2]
   BGP TE registered [No]
   ODN registered [No]

   IPv6 capability required/enabled: Yes/Yes
   Last SR policy update: 01:25:38
```

Configuration

These configuration commands are new:

```
nexthop validation color-extcomm sr-policy
nexthop validation color-extcomm disable
bgp bestpath igp-metric sr-policy
bgp bestpath sr-policy prefer
bgp bestpath sr-policy force
nexthop validation color-extcomm disable
```

BGP ODN AS NextHop Soft Validation Knob

On PE (HE):

```
RP/0/RP0/CPU0:PE1(config)#router bgp 65001
RP/0/RP0/CPU0:PE1(config-bgp)#nexthop ?
 mpls        Configure next-hop related items for mpls
 resolution  Configure next-hop related items for resolution
 validation  Configure next-hop reachability validation
```

```
RP/0/RP0/CPU0:PE1(config-bgp)#nexthop validation ?
 color-extcomm  Configure next-hop reachability validation for color-extcomm paths

RP/0/RP0/CPU0:PE1(config-bgp)#nexthop validation color-extcomm ?
 disable    Disable next-hop reachability validation for color-extcomm paths
 sr-policy  Enable BGP next-hop reachability validation by SR Policy for color-extcomm paths

RP/0/RP0/CPU0:PE1(config-bgp)#nexthop validation color-extcomm sr-policy
RP/0/RP0/CPU0:PE1(config-bgp)#commit
```

This is the main command: it turns on the BGP soft next-hop behavior. The RIB validation is not performed if there is an SR policy that is up for the next-hop and color.

BGP ODN AS Next-Hop Hard Validation Knob

BGP Hard Next-Hop is the default behavior.

This command is the command to revert back to this behavior: **no nexthop validation color-extcomm.**

BGP Best Path Selection Considering SR Policy Path Metric

When we have Interior Gateway Protocol (IGP) reachability to the NHs, and if the algorithm reaches Step 8 in the BGP best-path selection process, the preferred BGP path is the one with the lowest (IGP) distance to the next-hop. This is the default behavior. See BGP Best Path Selection Algorithm.

This is true except if the command **bgp bestpath igp-metric ignore** is configured. In that case, the IGP cost is not considered at all.

Currently, only the IGP metric to the BGP NH is considered; not the metric provided by the SR policy path. This remains the default behavior, but there is a command that instructs BGP to use the SR policy path metric instead of the IGP metric for the BGP best-path selection algorithm.

```
RP/0/RP0/CPU0:PE1(config)#router bgp 65001
RP/0/RP0/CPU0:PE1(config-bgp)#bgp bestpath igp-metric ?
 ignore     Ignore IGP metric during path comparison
 sr-policy  Use next-hop admin/metric from SR policy at Next Hop metric compareson stage

RP/0/RP0/CPU0:PE1(config-bgp)#bgp bestpath igp-metric sr-policy
RP/0/RP0/CPU0:PE1(config-bgp)#commit
```

This command turns on the consideration of the PCE/path admin and metric values. These admin/metric values can only be passed to BGP if the SR policy is up. This command enables the BGP algorithm to pick the best path based on the admin and metric for the next-hop from the SR policy. Without this command, the default behavior is to only consider the IGP metric of the next-hop. This is referred to as 'RIB validation of the next-hop'.

BGP Best Path Selection Preferring SR Policy Paths

There are platforms that do not support the mix of paths that have either a native next-hop or an SR policy next-hop. The platform might not support that mix in forwarding over both path types. This is important considering the use of Equal Cost Multi-Path (ECMP) or Unequal Cost Multi-Path (UCMP) or backup paths. Any type of path can be the BGP's best path. The default behavior is to

only consider BGP paths that have the same next-hop type as the BGP best path.

This command instructs BGP to prefer routes for which there is an SR policy for the color/next-hop when the router performs the best path calculation. This means that paths where the SR policy is down, or where there is no SR policy, are not considered during the best-path calculation.

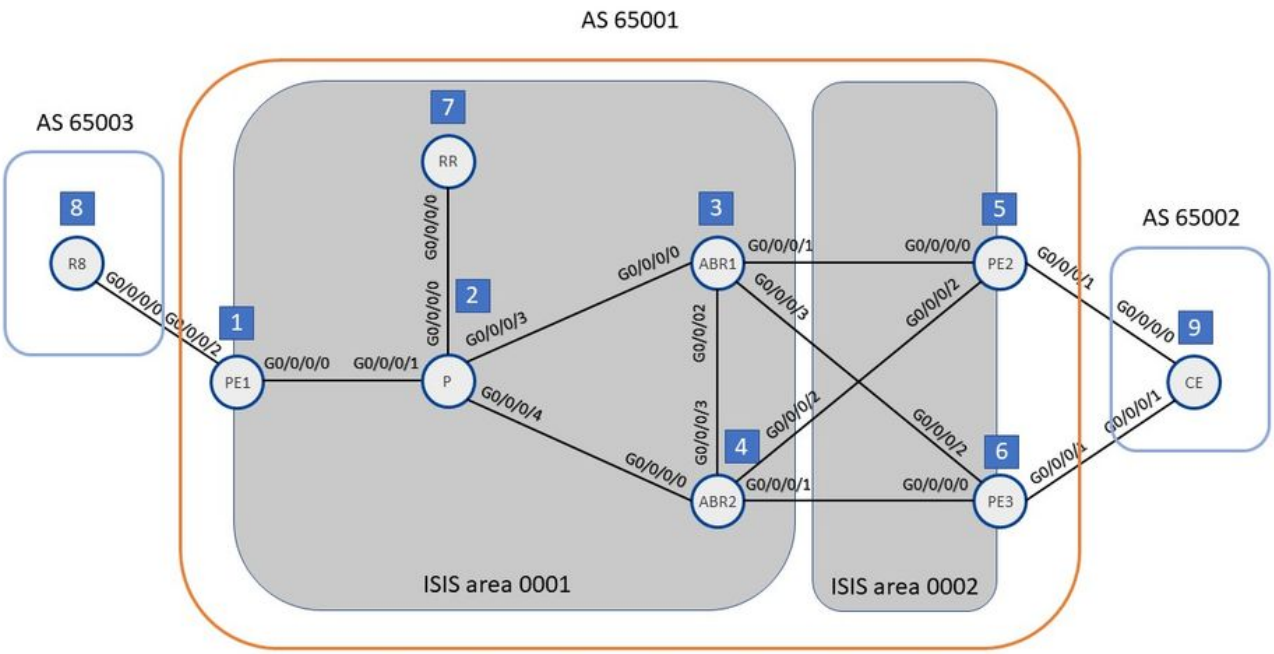
bgp bestpath sr-policy {force | prefer}

One of the two keywords must be configured.

```
RP/0/RP0/CPU0:PE1(config-bgp)#bgp bestpath sr-policy ?
 force   Consider only paths over SR Policy for bestpath selection, eBGP no-color ineligible
 prefer  Consider only paths over SR Policy for bestpath selection, eBGP no-color eligible
```

If you configure the preferred option, then the eBGP paths without color are marked eligible (so can be part of the best path). If this behavior is not wanted, you could add a dummy SR policy to the eBGP paths. Otherwise, you can configure the force option for this command so that the eBGP routes without color are ineligible.

Refer to the network as shown in the image.



There are three possible paths for the network 10.99.99.99/32 from the router PE1. The prefix 10.99.99.99/32 is advertised by R8, and the CE router.

BGP has 3 paths for the route 10.99.99.99/32: 2 iBGP (PE2 and PE3 are the BGP next-hop routers) and 1 eBGP paths (from R8).

The iBGP paths have next-hop 10.0.0.5 and 10.0.0.6. The eBGP path has next-hop 10.1.8.8.

The configuration does not have this command **bgp bestpath sr-policy**.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.9.9.9/32
BGP routing table entry for 10.9.9.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                474          474
   Local Label: 24005
Last Modified: Nov 29 09:04:07.948 for 00:00:49
Paths: (3 available, best #3)
 Advertised to PE peers (in unique update groups):
   10.0.0.4        10.0.0.3
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24007) (admin 20) (metric 23) from 10.0.0.3 (10.0.0.5)
     Received Label 24018
     Origin IGP, metric 0, localpref 100, valid, internal, group-best, imported
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24007
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
 Path #2: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.6 from 10.0.0.4 (10.0.0.6)
     Received Label 24004
     Origin IGP, metric 0, localpref 100, valid, internal, imported
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: RT:65001:101
     Originator: 10.0.0.6, Cluster list: 10.0.0.4
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:3
 Path #3: Received by speaker 0
 Advertised to PE peers (in unique update groups):
   10.0.0.4        10.0.0.3
 65003
   10.1.8.8 from 10.1.8.8 (10.0.0.8)
     Origin IGP, metric 0, localpref 100, valid, external, best, group-best, import-candidate
     Received Path ID 0, Local Path ID 1, version 474
     Extended community: RT:65001:101
     Origin-AS validity: (disabled)
```

The eBGP path does not have a color or an SR policy. It is the best path.

If the eBGP route does have a color, but no SR policy, it is still chosen as the best path.

If the eBGP route does have a color, and an SR policy, it is chosen as the best path.

Here is another example. The eBGP route does not have a color, and no SR policy and the command **bgp bestpath sr-policy prefer** is configured.

> **Note**: The eBGP neighbor is inside the VRF. This means that you must configure the command **bgp bestpath sr-policy prefer** under the VRF.

```
router bgp 65001
 nexthop validation color-extcomm sr-policy
 bgp unsafe-ebgp-policy
 bgp bestpath igp-metric sr-policy
 address-family vpnv4 unicast
 !
 neighbor 10.0.0.3
```

```
 remote-as 65001
 update-source Loopback0
 address-family vpnv4 unicast
 !
 !
 neighbor 10.0.0.4
 remote-as 65001
 update-source Loopback0
 address-family vpnv4 unicast
 !
 !
 neighbor 10.0.0.7
 remote-as 65001
 shutdown
 update-source Loopback0
 address-family vpnv4 unicast
 !
 !
 vrf one
 rd 65000:1
 bgp unsafe-ebgp-policy
 bgp bestpath sr-policy prefer
 address-family ipv4 unicast
  redistribute connected
 !
 neighbor 10.1.8.8
  remote-as 65003
  address-family ipv4 unicast
  !
 !
 !
!

RP/0/RP0/CPU0:PE1#show bgp vrf one 10.9.9.9/32 bestpath-compare
BGP routing table entry for 10.9.9.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                579          579
   Local Label: 24004 (no rewrite);
   Flags: 0x01343001+0x00020000;
Last Modified: Nov 30 07:36:55.948 for 00:03:05
Paths: (3 available, best #3)
 Advertised to PE peers (in unique update groups):
   10.0.0.4        10.0.0.3
 Path #1: Received by speaker 0
 Flags: 0x2000000001020005, import: 0x080
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24007) (admin 20) (metric 23) from 10.0.0.3 (10.0.0.5), if-handle
0x00000000
     Received Label 24018
     Origin IGP, metric 0, localpref 100, valid, internal, group-best, imported
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24007
     best of AS 65002
     An iBGP path, whereas best path (path #3) is an eBGP path
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
 Path #2: Received by speaker 0
 Flags: 0x2000000000020005, import: 0x0a0
 Not advertised to any peer
 65002
```

```
  10.0.0.6 from 10.0.0.4 (10.0.0.6), if-handle 0x00000000
    Received Label 24004
    Origin IGP, metric 0, localpref 100, valid, internal, imported
    Received Path ID 0, Local Path ID 0, version 0
    Extended community: RT:65001:101
    Originator: 10.0.0.6, Cluster list: 10.0.0.4
    Non SR-policy path is ignored due to config knob
    Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:3
 Path #3: Received by speaker 0
 Flags: 0x300000000d040003, import: 0x31f
 Advertised to PE peers (in unique update groups):
   10.0.0.4        10.0.0.3
 65003
   10.1.8.8 from 10.1.8.8 (10.0.0.8), if-handle 0x00000000
    Origin IGP, metric 0, localpref 100, valid, external, best, group-best, import-candidate
    Received Path ID 0, Local Path ID 1, version 579
    Extended community: RT:65001:101
    Origin-AS validity: (disabled)
    best of AS 65003, Overall best
```

The eBGP path is the best, even though it has no color. If you do not want the eBGP route without color as the best path, then configure the command **bgp bestpath sr-policy** with the **force** option.

Note: The local and redistributed paths are always eligible for the best path calculation.

Use this command to check if the platform supports the mix of forwarding over SR policy and native next-hop.

```
RP/0/RP0/CPU0:R1#show bgp process detail | include native
Platform support mix of sr-policy and native nexthop: No
```

Note: Routers NCS55xx and NCS560/NCS540 show no, and ASR9000 shows yes.

BGP Best Path Selection Forcing SR Policy Paths

The command instructs BGP to prefer routes with SR policy next-hop when performing best path calculation but excludes eBGP paths with no color.

```
RP/0/RP0/CPU0:PE1(config-bgp)#bgp bestpath sr-policy ?
 force   Consider only paths over SR Policy for bestpath selection, eBGP no-color ineligible
RP/0/RP0/CPU0:PE1(config-bgp)#bgp bestpath sr-policy force ?
  <cr>
```

next-hop reachability validation for color-extcomm paths is disabled

This is typically used on Route Reflectors (RRs).

On RR:

```
RP/0/RP0/CPU0:RR1(config-bgp)#nexthop validation color-extcomm disable
RP/0/RP0/CPU0:RR1(config-bgp)#commit
```

The next-hop reachability validation for color-extcomm paths is disabled. This is irrelevant to the state or presence of an SR policy.

Overview

The behavior on Headend and RR is driven by the configuration of the next-hop validation command and the **bgp best path igp-metric sr-policy** command. There are 4 scenarios. Each scenario has a combination of two configuration commands.

Default Behavior

Applicable on Headend router and RR.

Configuration:

```
no nexthop validation color-extcomm sr-policy
no bgp bestpath igp-metric sr-policy
```

Function:

```
Perform RIB validation (hard next-hop).
Do not use admin/metric from the sr-policy.
```

RIB Dependent Validation with SR Policy Metric

Applicable on Headend router and RR.

Configuration:

```
no nexthop validation color-extcomm sr-policy
bgp bestpath igp-metric sr-policy
```

Function:

```
Perform RIB validation (hard next-hop).
If NH is reachable in RIB:
 If policy is up:
  Use policy metric
 If policy is down:
  Use RIB metric
```

SR policy Dependent Validation with RIB Metric

This is the default behavior.

Applicable on Headend router.

Configuration:

```
nexthop validation color-extcomm sr-policy
no bgp bestpath igp-metric sr-policy
```

Function:

```
Do not perform RIB validation (soft next-hop).
```

```
Do not use admin/metric from the SR policy.
The RIB metric might not be available.
```

SR Policy Dependent Validation with SR Policy Metric

Applicable on Headend router.

Configuration:

```
nexthop validation color-extcomm sr-policy
bgp bestpath igp-metric sr-policy
```

Function:

```
Do not perform RIB validation (soft next-hop).  RIB reachability is not needed.
If policy is up:
 Use policy metric and validation, even if RIB reachability is present
If policy is down:
 Use RIB validation and metric if available. If not available, the route is not installed.
```

SR Policy Dependent Validation with RIB Metric and SR Policy not Used for Best Path Calculation

Applicable on RR router.

Configuration:

```
nexthop validation color-extcomm disable
no bgp bestpath igp-metric sr-policy
```

Function:

```
Use RIB metric if the next-hop is in the RIB. Else, use the gateway metric (the next-hop IGP
metric) 0.
Do not use SR policy for bestpath calculation. Do not use admin/metric from the SR policy.
```

SR Policy Dependent Validation with RIB Metric and SR Policy Used for Best Path Calculation

Applicable on RR router.

Configuration:

```
nexthop validation color-extcomm disable
bgp bestpath igp-metric sr-policy
```

Function:

```
Use RIB metric if the next-hop is in the RIB. Else, use the gateway metric 0.
Use sr-policy for bestpath calculation.

If policy is up:
```

```
 Use policy metric and validation, even if RIB reachability is present
If policy is down
 Use RIB validation and metric if available
 If RIB validation and metric is not available:
  use the gateway metric 0
```

Verification

This is how you verify which kind of next-hop validation is active and if the admin distance/metric of the SR policy is used during the best path calculation.

```
RP/0/RP0/CPU0:PE1#show bgp process detail | i Nexthop
Use SR-Policy admin/metric of color-extcomm Nexthop during path comparison: enabled
ExtComm Color Nexthop validation: SR-Policy then RIB
```

This is the default.

This is an example of SR policy Dependent Validation with RIB Metric and SR policy not Used for Best Path calculation.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast process detail | in Nexthop
Use SR-Policy admin/metric of color-extcomm Nexthop during path comparison: disabled
ExtComm Color Nexthop validation: RIBExtComm Color Nexthop validation: RIB
```

This is an example of an admin distance/metric attached to the BGP route.

```
RP/0/RP0/CPU0:PE1#show bgp vrf VRF1002 ipv4 unicast 10.77.2.0
BGP routing table entry for 10.77.2.0/24, Route Distinguisher: 18522:1002
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker            5232243     5232243
Paths: (1 available, best #1)
 Advertised to CE peers (in unique update groups):
   10.11.2.11      10.15.2.2
 Path #1: Received by speaker 0
 Advertised to CE peers (in unique update groups):
   10.11.2.11      10.15.2.2
 16611 770
   10.1.1.33 C:1129 (bsid:27163) (admin 20) (metric 25) from 10.1.1.100 (10.1.1.33)
     Received Label 24007
     Origin IGP, localpref 100, valid, internal, best, group-best, import-candidate, imported
     Received Path ID 1, Local Path ID 1, version 5232243
     Extended community: Color:1129 RT:17933:1002 RT:18522:1002
     Originator: 10.1.1.33, Cluster list: 10.1.1.100
     SR policy color 1129, up, registered, bsid 27163, if-handle 0x200053dc
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 18522:3002
```

Check if a Policy is Up or Down

This is how you verify if the SR policy is up or down.

```
RP/0/RP0/CPU0:PE1#show segment-routing traffic-eng pcc lsp
PCC's SR policy database:
-------------------------
Symbolic Name: cfg_ODN-policy-1_discr_100
 LSP[0]:
 Source 10.0.0.1, Destination 10.0.0.5, Tunnel ID 3, LSP ID 8
```

```
 State: Admin up, Operation up
 Setup type: SR
 Binding SID: 24005
```

Verification if Policy is Used

Use the BGP show command to look at the route.

If there is a Binding Segment IDentifier (BSID), then this route uses an SR policy.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                89           89
Last Modified: Oct 28 13:21:57.714 for 00:00:30
Paths: (1 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24004) from 10.0.0.3 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
imported
     Received Path ID 0, Local Path ID 1, version 87
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24004
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
```

The binding SID is an MPLS label here. This label is linked to one SR policy.

```
RP/0/RP0/CPU0:PE1#show mpls forwarding labels 24004
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24004  Pop         No ID              srte_c_101_e point2point     0
```

Show BGP Nexthops Command

You can see color, admin, and metric for the endpoint with the **show bgp nexthops** command.

```
RP/0/RP0/CPU0:RR#show bgp nexthops wide
Total Nexthop Processing
 Time Spent: 0.000 secs

Maximum Nexthop Processing
 Received: 00:21:57
 Bestpaths Deleted: 0
 Bestpaths Changed: 31
 Time Spent: 0.000 secs

Last Notification Processing
 Received: 00:01:22
 Time Spent: 0.000 secs

Gateway Address Family: IPv4 Unicast
Table ID: 0xe0000000
```

```
Gateway Reference Count: 8
Gateway AF Bits : 0x8011
Nexthop Count: 6
Critical Trigger Delay: 3000msec
Non-critical Trigger Delay: 10000msec

Nexthop Version: 1, RIB version: 1
EPE Table Version: 1, EPE Label version: 1
EPE Downloaded Version: 1, EPE Standby Version: 0

Status codes: R/UR Reachable/Unreachable
              C/NC Connected/Not-connected
              L/NL Local/Non-local
              PR   Pending Registration
              I    Invalid (Policy drop)

Next Hop                                Status          Metric           Tbl-ID
Notf      LastRIBEvent             RefCount
10.0.0.1                                [R][NC][NL]         30
e0000000                  6/0    00:01:22 (Cri)                  0/5
10.0.0.3                                [R][NC][NL]         20         e0000000
         6/0    00:01:22 (Cri)                 0/34
10.0.0.4                                [R][NC][NL]         30
e0000000                  6/0    00:01:22 (Cri)                 0/34
10.0.0.5                                [UR]        4294967295
        e0000000                   2/0    00:01:22 (Cri)                  0/4
10.0.0.5 T:101                          [UR]        4294967295
e0000000                  2/0    00:01:22 (Cri)                  0/3
10.0.0.6                                [UR]        4294967295
e0000000                  2/0    00:01:22 (Cri)                  0/3

RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast nexthops 10.0.0.5 color 101
Nexthop: 10.0.0.5 C:101
 VRF: default
 Nexthop ID: 0x6000006, Version: 0x0
 Nexthop Flags: 0x00480002
 Nexthop Handle: 0x7efc84043624
 RIB Related Information:
 Firsthop interface handle 0x0000000c
   Gateway TBL Id: 0xe0000000    Gateway Flags: 0x00000080
   Gateway Handle: 0x7efcadee6e98
   Gateway: reachable, non-Connected route, prefix length 8
   Resolving Route: 10.0.0.0/8 (static)
   Paths: 0
   RIB Nexhop ID: 0x0
   Status: Reachable via SR-TE
   Status: [Reachable][Connected][Not Local]
   Metric: 0 (SR-TE metric 333)
   ORR afi bits: 0x0
   Registration: Asynchronous, Completed: 2d05h
   Events: Critical (14)/Non-critical (0)
   Last Received: 02:15:15 (Critical)
   Last gw update: (Crit-notif) 02:15:15(rib)
   Reference Count: 2

 Prefix Related Information
   Active Tables: [IPv4 Unicast][VPNv4 Unicast]
   Metrices: [0x0][0x0]
   Reference Counts: [0][2]
 Interface Handle: 0x0
 Attr ref-count: 5
 SR policy color 101, State: [Up]
   Not registered, bsid 24004
```

```
  Skip Reg on restart [No]
  First notif received [Yes]
  SR Policy Flags [0x2]
  BGP TE registered [No]
  ODN registered [No]
  End-point admin/metric: 30/333
  IPv6 capability required/enabled: Yes/Yes
  Last SR policy update: 00:55:07
```

BGP Trace

Some entries in the output of **show bgp trace** refer to SR policy. Notice the presence of admin/metric.

```
default-bgp/spkr-tr2-sr 0/RP0/CPU0 t8885 [SR]:1323: SR-policy hdlr for reg nh with XTC af 0,
reg/unreg flag 1
default-bgp/spkr-tr2-sr 0/RP0/CPU0 t8885 [SR]:3394: SR-policy XTC nexthop 10.0.0.5/32 T:, color
101, register 1 with XTC done, v6-cap 1, rc 'Success', flags 0x480000
default-bgp/spkr-tr2-sr 0/RP0/CPU0 t8885 [SR]:3394: SR-policy XTC nexthop 10.0.0.6/32 T:, color
101, register 1 with XTC done, v6-cap 0, rc 'Success', flags 0x480000
default-bgp/spkr-tr2-sr 0/RP0/CPU0 t8885 [SR]:2424: SR-policy XTC notif NH end-point
color,gw_afi 0, [C:101][10.0.0.5] admin/metric 100/2147483647
default-bgp/spkr-tr2-sr 0/RP0/CPU0 t8885 [SR]:2424: SR-policy XTC notif NH end-point
color,gw_afi 0, [C:101][10.0.0.5] admin/metric 100/2147483647
default-bgp/spkr-tr2-sr 0/RP0/CPU0 t8885 [SR]:2424: SR-policy XTC notif NH end-point
color,gw_afi 0, [C:101][10.0.0.5] admin/metric 20/30
default-bgp/spkr-tr2-sr 0/RP0/CPU0 t8881 [SR]:1379: SR-policy trigger XTC for nh reg af 0,
reg/unreg flag 1
default-bgp/spkr-tr2-nh 0/RP0/CPU0 t8885 [NH]:7370: nexthop walk for AFI:'VPNv4 Unicast' start
default-bgp/spkr-tr2-nh 0/RP0/CPU0 t8885 [NH]:7425: nexthop walk for AFI:'VPNv4 Unicast', paths
deleted: 0, recalculated bestpaths: 2, color nh trigger for 2 nets, 0 msec
```

> **Note**: Cisco IOS XR Traffic Controller (XTC) refers to the SR controller.

Some entries in the BGP trace refer to the configuration change related to the next-hop processing.

```
default-bgp/spkr-tr2-prog 0/RP0/CPU0 t9036 [PROG]:724: 'Done VRF cfg notif init', name default
iid 0
default-bgp/spkr-tr2-prog 0/RP0/CPU0 t9036 [PROG]:792: 'Done cfg init', name default iid 0
default-bgp/spkr-tr2-gen 0/RP0/CPU0 t9048 [GEN]:17871: nh cfg change 2 sense 1
default-bgp/spkr-tr2-gen 0/RP0/CPU0 t9048 [GEN]:17920: nh cfg change 1 sense 1
```

Administrative Distance and Metric

The administrative distance (admin) is determined by the metric type in the SR policy. The metric type can be set on the headend router.

```
RP/0/RP0/CPU0:PE1#conf t
RP/0/RP0/CPU0:PE1(config)#segment-routing
RP/0/RP0/CPU0:PE1(config-sr)#traffic-eng
RP/0/RP0/CPU0:PE1(config-sr-te)#policy ODN-policy-1
RP/0/RP0/CPU0:PE1(config-sr-te-policy)#color 101 end-point ipv4 10.0.0.5
RP/0/RP0/CPU0:PE1(config-sr-te-policy)#candidate-paths
RP/0/RP0/CPU0:PE1(config-sr-te-policy-path)#preference 100
RP/0/RP0/CPU0:PE1(config-sr-te-policy-path-pref)#dynamic
RP/0/RP0/CPU0:PE1(config-sr-te-pp-info)#metric ?
```

```
 margin     Metric margin
 sid-limit  SID limit
 type       Metric type configuration
 <cr>

RP/0/RP0/CPU0:PE1(config-sr-te-pp-info)#metric type ?
 hopcount  Hopcount metric type
 igp       IGP metric type
 latency   Latency metric type
 te        TE metric type
```

Admin Values

These are the default SR policy admin values.

- latency 10
- TE 20
- IGP (default) 30
- hopcount 40
- NONE/UNKNOWN metric type (for explicit segment list policies) 100

If the metric type is **none**, then the metric value is 1.

The lower the admin value, the more preferred the path is to BGP.

The lower the metric, the more preferred the path is to BGP if the admin has the same value.

Verification of the Admin and Metric Type in ODN

```
RP/0/RP0/CPU0:PE1#show segment-routing traffic-eng policy color 101 endpoint ipv4 10.0.0.5

SR-TE policy database
---------------------

Color: 101, End-point: 10.0.0.5
 Name: srte_c_101_ep_10.0.0.5
 Status:
   Admin: up  Operational: up for 01:01:00 (since Oct 28 15:22:36.012)
 Candidate-paths:
   Preference: 100 (configuration) (active)
     Name: ODN-policy-1
     Requested BSID: dynamic
     PCC info:
       Symbolic name: cfg_ODN-policy-1_discr_100
       PLSP-ID: 4
       Protection Type: protected-preferred
       Maximum SID Depth: 10
     Dynamic (pce 10.0.0.7) (valid)
       Metric Type: IGP,   Path Accumulated Metric: 30
         16002 [Prefix-SID, 10.0.0.2]
         24009 [Adjacency-SID, 10.2.3.2 - 10.2.3.3]
         16005 [Prefix-SID, 10.0.0.5]
 Attributes:
   Binding SID: 24004
   Forward Class: Not Configured
   Steering labeled-services disabled: no
   Steering BGP disabled: no
   IPv6 caps enable: yes
   Invalidation drop enabled: no
```

Effective Metric

The effective metric forces the type and metric of the policy under which this command is configured.

```
RP/0/RP0/CPU0:PE1#conf t
RP/0/RP0/CPU0:PE1(config)#segment-routing
RP/0/RP0/CPU0:PE1(config-sr)#traffic-eng
RP/0/RP0/CPU0:PE1(config-sr-te)#policy ODN-policy-1
RP/0/RP0/CPU0:PE1(config-sr-te-policy)#candidate-paths
RP/0/RP0/CPU0:PE1(config-sr-te-policy-path)#preference 100
RP/0/RP0/CPU0:PE1(config-sr-te-policy-path-pref)#effective-metric ?
 value  Metric value, advertised to other protocols
 <cr>

RP/0/RP0/CPU0:PE1(config-sr-te-policy-path-pref)#effective-metric value 333 ?
 type  Metric type, advertised to other protocols
 <cr>

RP/0/RP0/CPU0:PE1(config-sr-te-policy-path-pref)#effective-metric value 333 type ?
 hopcount  HOPCOUNT metric type
 igp       IGP metric type
 latency   LATENCY metric type
 te        TE metric type

RP/0/RP0/CPU0:PE1(config-sr-te-policy-path-pref)#effective-metric value 333 type igp ?
 <cr>

RP/0/RP0/CPU0:PE1(config-sr-te-policy-path-pref)#effective-metric value 333 type igp
RP/0/RP0/CPU0:PE1(config-sr-te-policy-path-pref)#commit

RP/0/RP0/CPU0:PE1#show run segment-routing traffic-eng policy ODN-policy-1
segment-routing
 traffic-eng
 policy ODN-policy-1
  color 101 end-point ipv4 10.0.0.5
  candidate-paths
   preference 100
    dynamic
     pcep
     !
     metric
      type igp
     !
    !
    effective-metric
      value 333 type igp
```

You can verify the applied effective metric type (admin distance) and metric value in this way.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                131          131
Last Modified: Oct 28 15:22:35.714 for 00:03:42
Paths: (1 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
```

```
 65002
   10.0.0.5 C:101 (bsid:24004) (admin 30) (metric 333) from 10.0.0.7 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
imported
     Received Path ID 0, Local Path ID 1, version 130
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.7, 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24004
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
```

Comparisons of BGP Paths

The comparison of BGP paths is not changed by default.

If the command **bgp bestpath igp-metric sr-policy** is configured, then the admin distance and the metric of the SR policy is used in the BGP best-path selection algorithm.

The admin distance and the metric of the SR policy are tied to the SR policy. This is locally configured or received through PCEP (Path Computation Element Protocol) from the SR-PCE. This means that if an RR compares paths, it does not see the admin distance and metric, because it has no head-end functionality for ODN. Hence, it has no PCEP session to the SR PCE.

This example shows a prefix advertised by one remote PE router. This is the configuration.

```
segment-routing
 global-block 16000 23999
 traffic-eng
 logging
  policy status
 !
 policy ODN-policy-1
  color 101 end-point ipv4 10.0.0.5
  candidate-paths
   preference 100
    dynamic
     pcep
     !
     metric
      type te
     !
    !
   !
   preference 200
    dynamic
     pcep
     !
     metric
      type te
     !
```

The metric type is TE.

This head-end router sees a prefix with a color twice, with the same TE metric, because it is the same BGP next-hop for both paths.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast rd 65001:2 10.0.0.9/32 bestpath-compare
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65001:2
Versions:
```

```
 Process           bRIB/RIB  SendTblVer
 Speaker                  8           8
   Flags: 0x00040001+0x00010000;
Last Modified: Nov  2 09:21:55.948 for 00:00:32
Paths: (2 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Flags: 0xa000000025060005, import: 0x31f
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24018) (admin 20) (metric 23) from 10.0.0.3 (10.0.0.5), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
not-in-vrf
     Received Path ID 0, Local Path ID 1, version 8
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24018
     best of AS 65002, Overall best
 Path #2: Received by speaker 0
 Flags: 0x2000000024020005, import: 0x000
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24018) (admin 20) (metric 23) from 10.0.0.4 (10.0.0.5), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, import-candidate, not-in-vrf
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.4, 10.0.0.7, 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24018
     Longer cluster length than best path (path #1)
```

Because the admin distance and the metric are the same for both paths, the decision on which path is the best is taken further down in the BGP best-path selection algorithm.

This example shows a prefix advertised by two remote PE routers. One path has next-hop 10.0.0.5 and the other has next-hop 10.0.0.6. The prefix has color 101 from both remote PE routers. The headend router, PE1, has two ODN policies for this color.

```
segment-routing
 global-block 16000 23999
 traffic-eng
 logging
  policy status
 !
 policy ODN-policy-1
  color 101 end-point ipv4 10.0.0.5
  candidate-paths
   preference 100
    dynamic
     pcep
     !
      metric
       type igp
     !
    !
   !
   preference 200
    dynamic
     pcep
```

```
     !
     metric
       type te
     !
    !
   !
  !
 !
 policy ODN-policy-2
  color 101 end-point ipv4 10.0.0.6
  candidate-paths
   preference 100
    dynamic
     pcep
     !
      metric
       type igp
     !
```

The policy for endpoint 10.0.0.5 uses metric type TE and the policy for endpoint 10.0.0.6 uses metric type IGP.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast vrf one 10.0.0.9/32 bestpath-compare
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 25          25
   Flags: 0x00043001+0x00000000;
Last Modified: Nov  1 11:42:28.948 for 00:43:41
Paths: (2 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Flags: 0xa000000005060005, import: 0x080
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24007) (admin 20) (metric 30) from 10.0.0.4 (10.0.0.5), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, weight 65000, valid, internal, best, group-best,
import-candidate, imported
     Received Path ID 0, Local Path ID 1, version 25
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.4, 10.0.0.7, 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24007
     best of AS 65002, Overall best
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
 Path #2: Received by speaker 0
 Flags: 0x2000000000020005, import: 0x0a0
 Not advertised to any peer
 65002
   10.0.0.6 C:101 (bsid:24012) (admin 30) (metric 30) from 10.0.0.4 (10.0.0.6), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, weight 65000, valid, internal, imported
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.6, Cluster list: 10.0.0.4
     SR policy color 101, up, not-registered, bsid 24012
     Higher nexthop admin distance than best path (path #1)
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:3
```

The best path is the first one because it has a lower admin distance than the second path. The admin distance of metric type TE is lower than the one for metric type IGP.

The SR policy for ODN-policy-1 is up with precedence 200 and the SR policy for ODN-policy-2 is up with precedence 100.

```
RP/0/RP0/CPU0:PE1#show segment-routing traffic-eng pcc lsp detail

PCC's SR policy database:
-------------------------
Symbolic Name: cfg_ODN-policy-1_discr_100
 LSP[0]:
 Source 10.0.0.1, Destination 10.0.0.5, Tunnel ID 1, LSP ID 0
 State: Admin up, Operation down
 Setup type: SR
 Bandwidth: requested 0, used 0
 LSP object:
   PLSP-ID 0x1, flags: D:0 S:0 R:0 A:1 O:0 C:0
 Metric type: IGP, Accumulated Metric 30
 ERO:
   SID[0]: Node, Label 16004, NAI: 10.0.0.4
   SID[1]: Node, Label 16005, NAI: 10.0.0.5
Symbolic Name: cfg_ODN-policy-1_discr_200
 LSP[0]:
 Source 10.0.0.1, Destination 10.0.0.5, Tunnel ID 1, LSP ID 4
 State: Admin up, Operation up
 Binding SID: 24007
 Setup type: SR
 Bandwidth: requested 0, used 0
 LSP object:
   PLSP-ID 0x2, flags: D:0 S:0 R:0 A:1 O:1 C:0
 Metric type: TE, Accumulated Metric 30
 ERO:
   SID[0]: Adj, Label 24001, NAI: local 10.1.2.1 remote 10.1.2.2
   SID[1]: Adj, Label 24003, NAI: local 10.2.3.2 remote 10.2.3.3
   SID[2]: Node, Label 16005, NAI: 10.0.0.5
Symbolic Name: cfg_ODN-policy-2_discr_100
 LSP[0]:
 Source 10.0.0.1, Destination 10.0.0.6, Tunnel ID 2, LSP ID 2
 State: Admin up, Operation up
 Binding SID: 24012
 Setup type: SR
 Bandwidth: requested 0, used 0
 LSP object:
   PLSP-ID 0x3, flags: D:0 S:0 R:0 A:1 O:1 C:0
 Metric type: IGP, Accumulated Metric 30
 ERO:
   SID[0]: Node, Label 16004, NAI: 10.0.0.4
   SID[1]: Node, Label 16006, NAI: 10.0.0.6
```

Here is an example where the admin distance is the same, but the metric is different.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast vrf one 10.0.0.9/32 bestpath-compare
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 57           57
   Flags: 0x00043001+0x00010000;
Last Modified: Nov  2 07:54:20.948 for 00:00:04
Paths: (2 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Flags: 0xa000000005060005, import: 0x080
 Not advertised to any peer
```

```
 65002
   10.0.0.5 C:101 (bsid:24007) (admin 30) (metric 23) from 10.0.0.4 (10.0.0.5), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, weight 65000, valid, internal, best, group-best,
import-candidate, imported
     Received Path ID 0, Local Path ID 1, version 39
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.4, 10.0.0.7, 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24007
     best of AS 65002, Overall best
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
 Path #2: Received by speaker 0
 Flags: 0x2000000004020005, import: 0x080
 Not advertised to any peer
 65002
   10.0.0.6 C:101 (bsid:24012) (admin 30) (metric 30) from 10.0.0.4 (10.0.0.6), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, weight 65000, valid, internal, import-candidate,
imported
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.6, Cluster list: 10.0.0.4
     SR policy color 101, up, not-registered, bsid 24012
     Higher IGP metric than best path (path #1)
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:3
```

This is an example with metric-type hopcount.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast vrf one 10.0.0.9/32 bestpath-compare
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 99           99
   Flags: 0x00043001+0x00010000;
Last Modified: Nov  2 08:21:19.948 for 00:00:41
Paths: (2 available, best #2)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Flags: 0x2000000004020005, import: 0x080
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24007) (admin 40) (metric 4) from 10.0.0.4 (10.0.0.5), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, weight 65000, valid, internal, import-candidate,
imported
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.4, 10.0.0.7, 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24007
     Higher IGP metric than best path (path #2)
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
 Path #2: Received by speaker 0
 Flags: 0xa000000005060005, import: 0x080
 Not advertised to any peer
 65002
   10.0.0.6 C:101 (bsid:24010) (admin 40) (metric 3) from 10.0.0.4 (10.0.0.6), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, weight 65000, valid, internal, best, group-best,
import-candidate, imported
```

```
    Received Path ID 0, Local Path ID 1, version 95
    Extended community: Color:101 RT:65001:101
    Originator: 10.0.0.6, Cluster list: 10.0.0.4
    SR policy color 101, up, not-registered, bsid 24010
    best of AS 65002, Overall best
    Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:3
```

There are two competing BGP paths for two different endpoints. BGP decides which path wins and gets installed in the routing table. This in turn decides, based on the color and endpoint, which SR policy gets installed to forward the traffic towards the BGP VPNv4 prefix.

Comparing a BGP Path With and Without Color

In scenario four, the soft next-hop validation is enabled on the head end router and it receives two BGP paths for one prefix, one with and one without color. If there is no route for the next-hop, the path with no color does have next-hop inaccessible and is not considered for installment.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast | include 10.0.0.9/32
*>i10.0.0.9/32        10.0.0.5 C:101           0    100      0 65002 i
*>i10.0.0.9/32        10.0.0.5 C:101           0    100      0 65002 i
* i10.0.0.9/32        10.0.0.6                 0    100      0 65002 i
```

The last BGP path does not have the **>**, so the next-hop is inaccessible.

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast rd 65001:3 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65001:3
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                31           31
Last Modified: Nov  2 10:08:44.948 for 00:08:11
Paths: (2 available, no best path)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.6 (inaccessible) from 10.0.0.3 (10.0.0.6)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, not-in-vrf
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: RT:65001:101
     Originator: 10.0.0.6, Cluster list: 10.0.0.3, 10.0.0.7, 10.0.0.4
 Path #2: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.6 (inaccessible) from 10.0.0.4 (10.0.0.6)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, not-in-vrf
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: RT:65001:101
     Originator: 10.0.0.6, Cluster list: 10.0.0.4
```

The BGP path with the SR policy is used.

However, if the next-hop 10.0.0.6 is resolved because of a route in the RIB, then this path can be picked up as the best path. If it has no color though, it cannot be used for ODN and the SR policy would be down. However, the admin distance of this route is 100, so it is much higher than the path with color.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32 bestpath-compare
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 47           47
   Flags: 0x00043001+0x00000000;
Last Modified: Nov  2 10:30:55.948 for 00:00:21
Paths: (2 available, best #1)
 Advertised to CE peers (in unique update groups):
   10.1.8.8
 Path #1: Received by speaker 0
 Flags: 0xa000000005060005, import: 0x080
 Advertised to CE peers (in unique update groups):
   10.1.8.8
 65002
   10.0.0.5 C:101 (bsid:24021) (admin 20) (metric 23) from 10.0.0.3 (10.0.0.5), if-handle
0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
imported
     Received Path ID 0, Local Path ID 1, version 40
      Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24021
     best of AS 65002, Overall best
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
 Path #2: Received by speaker 0
 Flags: 0x2000000000020005, import: 0x0a0
 Not advertised to any peer
 65002
   10.0.0.6 from 10.0.0.4 (10.0.0.6), if-handle 0x00000000
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, imported
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: RT:65001:101
     Originator: 10.0.0.6, Cluster list: 10.0.0.4
     Higher nexthop admin distance than best path (path #1)
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:3
```

Scenarios in More Detail with Show Commands

Default Behavior

Applicable on Headend router and RR.

Configuration:

```
no nexthop validation color-extcomm sr-policy
no bgp bestpath igp-metric sr-policy
```

Function:

Perform RIB validation (hard next-hop).

BGP does not use admin/metric from the SR policy.

RIB validation is performed for the next-hop of the service route.

If there is no more specific route for the next-hop than the default route, the service route has an

inaccessible next-hop.

```
If the RIB metric is available:
RIB metric is used. Route is installed.
If policy is up:
Policy is used.
If policy is not up:
Policy is not used.
If the RIB metric is not available:
Route is not installed.
```

```
RP/0/RP0/CPU0:PE1#show bgp vpnv4 unicast rd 65001:2 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65001:2
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 31           31
Last Modified: Oct 26 14:21:56.714 for 00:01:32
Paths: (1 available, no best path)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
  65002
   10.0.0.5 C:101 (bsid:24005) (inaccessible) from 10.0.0.3 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, not-in-vrf
     Received Path ID 0, Local Path ID 0, version 0
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24005
```

This also leads to the fact that the service route is not imported into the VRF.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 37           37
Last Modified: Oct 26 14:24:36.714 for 00:00:03
Paths: (0 available, no best path)
 Not advertised to any peer
```

If you add a non-default static route on the headend router covering the next-hop of the service route it alleviates this issue. This is commonly used in ODN networks.

This static route covers the next-hop 10.0.0.5 and is not a default route.

```
router static
 address-family ipv4 unicast
 10.0.0.0/24 Null0
 !
!
```

It solves the next-hop inaccessible for ODN.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 27           27
```

```
Last Modified: Oct 26 14:19:06.714 for 00:00:26
Paths: (1 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24005) from 10.0.0.3 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
imported
     Received Path ID 0, Local Path ID 1, version 22
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24005
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
```

The same is true on the RR: if the next-hop of the service route is inaccessible, the route is not reflected in other iBGP speakers. The same workaround of a non-default static route can be used on a RR.

RIB Dependent Validation with SR Policy Metric

Applicable on Headend router and RR.

Configuration:

```
no nexthop validation color-extcomm sr-policy
bgp bestpath igp-metric sr-policy
```

Function:

The PCE/path admin and metric values are passed to BGP and are used for the best path calculation.

```
Perform RIB validation (hard next-hop).
If NH is reachable in RIB:
If policy is up:
 Use policy metric.
 If policy is down:
  Use RIB metric.
```

Headend Router

If the next-hop is not reachable in the RIB, then the service route has the next-hop inaccessible and it is not installed.

If the next-hop is reachable (possible via the use of a static route), then the service route is installed, now with the admin and metric values.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 43           43
Last Modified: Oct 26 14:42:54.714 for 00:00:03
Paths: (1 available, best #1)
 Not advertised to any peer
```

```
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24005) (admin 20) (metric 30) from 10.0.0.3 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
imported
     Received Path ID 0, Local Path ID 1, version 43
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24005
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
```

The policy is up.

If the policy is down, while the RIB has a route for the next-hop, then the service route is installed. However, the service route is not resolved in the CEF table. The SR policy no longer provides the connectivity (the MPLS label stack) to reach the endpoint.

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 57           57
Last Modified: Oct 26 15:13:46.714 for 00:01:39
Paths: (1 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.5 from 10.0.0.3 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
imported
     Received Path ID 0, Local Path ID 1, version 48
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
```

The route is installed, but CEF is not resolved for this service route without the SR policy.

```
RP/0/RP0/CPU0:PE1#show cef vrf one 10.0.0.9/32
10.0.0.9/32, version 36, drop adjacency, internal 0x5000001 0x30 (ptr 0xe3abf78) [1], 0x600
(0xe54a068), 0xa08 (0xec42558)
 Updated Oct 26 15:13:47.003
 Prefix Len 32, traffic index 0, precedence n/a, priority 3
 gateway array (0xe3b26b8) reference count 2, flags 0x3a, source rib (7), 0 backups
               [3 type 1 flags 0x88401 (0xec85888) ext 0x0 (0x0)]
 LW-LDI[type=1, refc=1, ptr=0xe54a068, sh-ldi=0xec85888]
 gateway array update type-time 3 Oct 26 15:16:24.524
 LDI Update time Oct 26 14:42:54.404
 LW-LDI-TS Oct 26 15:13:47.003
  via 10.0.0.5/32, 0 dependencies, recursive [flags 0x6000]
   path-idx 0 NHID 0x0 [0xd649400 0x0]
   recursion-via-/32
   next hop VRF - 'default', table - 0xe0000000
   unresolved
    labels imposed {24002}

   Load distribution: 0 (refcount 3)
```

```
   Hash  OK  Interface                Address
   0     Y   recursive                drop
```

RR Router:

If the SR policy is up or not, and if the RIB reachability is there, the RR advertises the service route.

SR Policy Dependent Validation with RIB Metric

Applicable on Headend router.

Configuration:

```
nexthop validation color-extcomm sr-policy
no bgp bestpath igp-metric sr-policy
```

Function:

The PCE/path admin and metric values are not passed to BGP.

```
If the RIB metric is available:
 RIB metric is used. Route is installed.
 If policy is up:
  Policy is used.
  If policy is not up:
   Policy is not used.

If the RIB metric is not available:
 Route is not installed.
```

SR Policy Dependent Validation with SR Policy Metric

Applicable on Headend router.

Configuration:

```
nexthop validation color-extcomm sr-policy
bgp bestpath igp-metric sr-policy
```

Function:

```
Do not perform RIB validation (soft next-hop).  RIB reachability is not needed.
If policy is up:
 Use policy metric and validation, even if RIB reachability is present.
If policy is down:
Use RIB validation and metric if available. If not available, the route is not installed.
```

If the SR policy is available:

```
RP/0/RP0/CPU0:PE1#show bgp vrf one 10.0.0.9/32
BGP routing table entry for 10.0.0.9/32, Route Distinguisher: 65000:1
Versions:
```

```
 Process           bRIB/RIB  SendTblVer
 Speaker                101         101
Last Modified: Oct 28 13:32:24.714 for 00:25:39
Paths: (1 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 65002
   10.0.0.5 C:101 (bsid:24008) (admin 30) (metric 30) from 10.0.0.3 (10.0.0.5)
     Received Label 24002
     Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
imported
     Received Path ID 0, Local Path ID 1, version 99
     Extended community: Color:101 RT:65001:101
     Originator: 10.0.0.5, Cluster list: 10.0.0.3
     SR policy color 101, up, not-registered, bsid 24008
     Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65001:2
```

SR Policy Dependent Validation with RIB Metric and SR Policy Not Used for Best Path Calculation

Applicable on RR router.

Configuration:

```
nexthop validation color-extcomm disable
no bgp bestpath igp-metric sr-policy
```

Function:

The first command means that the next-hop reachability validation for color-extcomm paths is disabled. There is a hard check for the next-hop reachability. The validation check for the soft next-hop reachability can be disabled as this router is a RR and only reflects the BGP service routes. The RR does not install an SR policy for them. Without this command, a soft check would be performed. If there is no other route for the next-hop in the routing table than the default route, then the next-hop is inaccessible. The route is then not reflected.

The second command means that the SR policy is not used for BGP best-path calculation. So, the admin/metric from the SR policy is not used. The RIB metric is used if the next-hop is in the RIB. Else, gateway metric 0 (the next-hop IGP metric) is used.

SR Policy Dependent Validation with RIB Metric and SR Policy Used for Best Path Calculation

Applicable on RR router.

Configuration:

```
nexthop validation color-extcomm disable
bgp bestpath igp-metric sr-policy
```

Function:

The first command means that the next-hop reachability validation for color-extcomm paths is disabled. There is a hard check for the next-hop reachability. The validation check for the soft next-hop reachability can be disabled as this is a RR and only reflects the BGP service routes.

The RR does not install an SR policy for them. Without this command, a soft check would be performed. If there is no other route for the next-hop in the routing table than the default route, then the next-hop is inaccessible. The route is then not reflected.

The second command means that SR policy is used for BGP best-path calculation.

```
If policy is up:
 Use policy metric and validation, even if RIB reachability is present
If policy is down
 Use RIB validation and metric if available
 If RIB validation and metric is not available:
  use the gateway metric 0
```