SR-TE On-Demand Next Hop

When redistributing routing information across domains, provisioning of multi-domain services (Layer 2 VPN and Layer 3 VPN) has its own complexity and scalability issues. Effective Cisco IOS XE Everest 16.6.1, On-Demand Next Hop (ODN) triggers delegation of computation of an end-to-end LSP to a PCE controller including constraints and policies without doing any redistribution. It then installs the reapplied multi-domain LSP for the duration of the service into the local forwarding information base (FIB).

Restrictions for SR-TE On-Demand Next Hop

- ODN anycast SID is not supported.
- ODN for IPv6 is not supported.

Information About SR-TE On-Demand Next Hop

On-Demand Next hop leverages upon BGP Dynamic SR-TE capabilities and adds the path computation (PCE) ability to find and download the end to end path based on the requirements. ODN triggers an SR-TE auto-tunnel based on the defined BGP policy. As shown in the following figure, an end-to-end path between ToR1 and
AC1 can be established from both ends based on low latency or other criteria for VRF (L3VPN) or IPv4 services. The work-flow for ODN is summarized as follows:

**Figure 1: ODN Operation**

1. PCE controller collects topology and SIDs information via BGP Link State (BGP-LS). For more information on BGP-LS, refer BGP Link-State.
2. If NSO controller is enabled, it configures L3VPN VRF or IPv4 prefixes and requests are sent to ToR1 and AC1.
3. ToR1 and AC1 checks if an LSP toward each other exists. If not, a request is sent to the PCE controller to compute that SR-TE path that matches SR-TE policy that is carried through BGP.
4. PCE controller computes the path and replies with a label stack (18001, 18002, 16001, example in ToR1).
5. ToR1 and AC1 create an SR-TE auto-tunnel and reply to the NSO controller indicating that the LSP for VRF or IPv4 is up and operational.

### Fast Convergence Default Optimize

The fast convergence "default optimize" feature modifies the default settings of all the protocols to recommended defaults for fast convergence. To revert the defaults to pre-fast-convergence settings for both IS-IS and OSPF, `routing-default-optimize` command is used. This command sends signals to IS-IS and OSPF and modifies the default configuration for these protocols.

By default, the fast convergence setting is enabled which means when you upgrade the software, you can automatically see the new behavior. This makes easier integration of the devices in a multivendor deployment and reduces support cases for poor convergence.

When default optimize is disabled, existing protocol default configuration is used. When default optimize is enabled, new protocol defaults are used. The show running configurations does not display configuration lines for default settings even when default settings are being used.

A configuration of a protocol overrides the default, but a change to default optimize does not override any configuration.

The following is the sample output of `spf-interval` command in IS-IS:
Device(config-if)# router isis
Device(config-router)# spf-interval 10 5500 5500

If a non-default value is configured, it is displayed in show running configuration output:

Device(config-router)# spf-interval 5 50 200
Device(config-router)# do show run | inc spf-interval
spf-interval 5 50 200

You can revert to the default values by configuring the default values or by removing the non-default configuration.

### Default Optimize Values for IS-IS

The following table summarizes the configuration impacted by default optimize:

<table>
<thead>
<tr>
<th>IS-IS command</th>
<th>Parameters</th>
<th>Default optimize disabled</th>
<th>Default optimize enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>fast-flood</td>
<td>number of LSPs flooded back-to-back</td>
<td>Disabled</td>
<td>10</td>
</tr>
<tr>
<td>spf-interval</td>
<td>Initial (milliseconds)</td>
<td>5500</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Secondary (milliseconds)</td>
<td>5500</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>Max (seconds)</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>pres-interval</td>
<td>Initial (milliseconds)</td>
<td>2000</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Secondary (milliseconds)</td>
<td>5000</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>Max (seconds)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>lsp-gen-interval</td>
<td>Initial (milliseconds)</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Secondary (milliseconds)</td>
<td>5000</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>Max (seconds)</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

### Default Optimize Values for OSPF

The following table summarizes the configuration impacted by default optimize for OSPFv2/v3:
<table>
<thead>
<tr>
<th>OSPF command</th>
<th>Parameters</th>
<th>Default optimize disabled</th>
<th>Default optimize enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>timers throttle spf</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial (milliseconds)</td>
<td>5500</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Secondary (milliseconds)</td>
<td>10000</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Max (seconds)</td>
<td>10</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>timers throttle lsa all</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial (milliseconds)</td>
<td>0</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Secondary (milliseconds)</td>
<td>5000</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Max (seconds)</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>timers lsa arrival</td>
<td>milliseconds</td>
<td>1000</td>
<td>100</td>
</tr>
</tbody>
</table>

The following is the sample output of show ip ospf command for OSPFv2 with the default-optimize values.

Device# show ip ospf
Routing Process "ospf 10" with ID 1.1.1.1
Start time: 00:00:01.471, Time elapsed: 03:00:34.706
Supports only single TOS(TOS0) routes
Supports opaque LSA
Supports Link-local Signaling (LLS)
Supports area transit capability
Supports NSSA (compatible with RFC 3101)
Supports Database Exchange Summary List Optimization (RFC 5243)
Event-log enabled, Maximum number of events: 1000, Mode: cyclic
Router is not originating router-LSAs with maximum metric
Initial SPF schedule delay 50 msecs
Minimum hold time between two consecutive SPF s 200 msecs
Maximum wait time between two consecutive SPF s 5000 msecs
Incremental-SPF disabled
Initial LSA throttle delay 50 msecs
Minimum hold time for LSA throttle 200 msecs
Maximum wait time for LSA throttle 5000 msecs
Minimum LSA arrival 100 msecs
LSA group pacing timer 240 secs
Interface flood pacing timer 33 msecs
Retransmission pacing timer 66 msecs
EXCHANGE/LOADING adjacency limit: initial 300, process maximum 300
Number of external LSA 18. Checksum Sum 0x075EB2
Number of opaque AS LSA 0. Checksum Sum 0x000000
Number of DCbitless external and opaque AS LSA 0
Number of DoNotAge external and opaque AS LSA 0
Number of areas in this router is 1. 1 normal 0 stub 0 nssa
Number of areas transit capable is 0
External flood list length 0
IETF NSF helper support enabled
Cisco NSF helper support enabled
Reference bandwidth unit is 100 mbps
Area BACKBONE(0)
  Number of interfaces in this area is 4 (2 loopback)
  Area has RRR enabled
The following is the sample output of `show ospf` command for OSPFv3 with the default-optimize values.

```
Device# show ospfv3
OSPFv3 10 address-family ipv6
Router ID 11.11.11.11
Supports NSSA (compatible with RFC 3101)
Supports Database Exchange Summary List Optimization (RFC 5243)
Event-log enabled, Maximum number of events: 1000, Mode: cyclic
Router is not originating router-LSAs with maximum metric
Initial SPF schedule delay 50 msecs
Minimum hold time between two consecutive SPF s 200 msecs
Maximum wait time between two consecutive SPF s 5000 msecs
Initial LSA throttle delay 50 msecs
Minimum hold time for LSA throttle 200 msecs
Maximum wait time for LSA throttle 5000 msecs
Minimum LSA arrival 100 msecs
LSA group pacing timer 240 secs
Interface flood pacing timer 33 msecs
Retransmission pacing timer 66 msecs
Retransmission limit dc 24 non-dc 24
EXCHANGE/LOADING adjacency limit: initial 300, process maximum 300
Number of external LSA 0. Checksum Sum 0x000000
Number of areas in this router is 1. 1 normal 0 stub 0 nssa
Graceful restart helper support enabled
Reference bandwidth unit is 100 mbps
RFC1583 compatibility enabled
Area BACKBONE(0)
  Number of interfaces in this area is 2
  SPF algorithm executed 7 times
  Number of LSA 3. Checksum Sum 0x012426
  Number of DCbitless LSA 0
  Number of indication LSA 0
  Number of DoNotAge LSA 0
  Flood list length 0
```

**SR-TE Policy, Color Extended Community, Affinity Constraint, and Disjointness Constraint**

Effective Cisco IOS XE Gibraltar 16.11.1, the router supports:

- MPLS TE new SR-TE policy command—`segment-routing traffic-eng`
- Color-extended community
- Affinity constraints
- Disjointness constraints
SR-TE Policy Command

Effective Cisco IOS XE Gibraltar 16.11.1, a new command (`segment-routing traffic-eng`) is added to configure the SR policy under segment routing. For information on using this command, see the "Configuring Color-Extended Community" section.

Color Extended Community

In earlier releases, the router created segment routed Traffic Engineering (SR-TE) tunnels based on a tunnel-profile or attribute set. As part of this functionality, an inbound route-map with a "match community" and "set attribute-set" was added on the ingress node and the route-map matched against communities received in the BGP updates. A BGP update with a matching community would initiate an SR-TE tunnel for the nexthop TE-profile.

However, effective Cisco IOS XE Gibraltar 16.11.1, ‘color extended’ community is now added as follows:

• An SR-TE policy is created on the ingress router for the Color-Endpoint pair.
• The egress router adds the ‘color extended’ community to the BGP updates that require a Traffic-Engineered path.

Affinity Constraint

Affinity is a 32-bit constraint used by the PCE and PCALC for calculating paths that take the "affinity constraint" into account.

Affinity constraints let you assign, or map, color names for path affinities. After mappings are defined, the attributes can be referred to by the corresponding color name in the command.

Affinity maps are used to map operator-defined color names to a bit position in the affinity bitmap.

Supported Affinity constraints are:

• include-all—indicates that constrained shortest path first (CSPF) includes a link when calculating a path, only if each link administrative group bit has the same name as each affinity bit.
• include-any—indicates that CSPF includes a link when calculating a path, if at least one link administrative group bit has the same name as an affinity bit.
• exclude-any—indicates that CSPF excludes a link when calculating a path, if any link administrative group bit has the same name as an affinity bit.

Disjointness Constraint

Disjointness is used to describe two or more services that must be completely disjoint of each other. Disjointness is useful for providing traffic flow redundancy in the network.

Disjointness is controlled by the PCE. The PCE learns of the network topology through an IGP (OSPF or IS-IS) through the BGP-LS protocol and is capable of computing paths based on the IGP or TE metric.

The PCE uses the disjoint policy to compute two lists of segments that steer traffic from the source node to the towards destination node along disjoint paths. Disjoint paths can originate from either the same or different head-ends.
A "disjoint level" refers to the type of resources that should not be shared by the two computed paths. The PCE supports the following disjoint path computations:

- Link
- Node
- Shared risk link group (SRLG)

When the first request is received from Path Computation client (PCC) or an ingress node, with a given disjoint-group ID, a list of segments is computed based on the metric requested, encoding the shortest path from source to destination.

When the second request is received with the same disjoint-group ID, based on the information received in both requests, the PCE computes two disjoint paths from the source to the destination.

Both paths are computed at the same time. The shortest list of segments is calculated to steer traffic on the computed paths.

**Limitations**

- On the Cisco ASR 920 routers, the total stack limit is five labels.

```
Router# show mpls outlabel-stack reservation
Application    Outlabels  Status
             Reserved
   bgp          0            COMPLETE
   transport    0            COMPLETE
    vpn         1            COMPLETE

Platform Max OutLabels: 5
Total Available OutLabels: 4
Reg apps: MPLS-TE(4) OSPF-1(4)
```

- The maximum number of SR-TE ODN policies supported is 500.
- Static SR-TE policy is not supported.
- SR-TE policy counters are not supported on the Cisco ASR 920 routers.
- SR-TE policy goes down only after 60 seconds when the connection to PCEP is disconnected due to any reason.
- 6PE and 6VPE services are supported only if the number of transport labels used for SR-TE policy does not exceed two labels.
How to Configure Segment Routing On-Demand Next Hops—Layer 3 and Layer 3 VPN

Configuring Segment Routing On-Demand Next Hops—Layer 3 and Layer 3 VPN

Perform the following steps to configure on-demand next hop for SR-TE. The following figure is used as a reference to explain the configuration steps.

Figure 2: ODN Auto-Tunnel Setup

1. Configure the router (R6 tailend) with VRF interface.

   vrf definition sr
   rd 100:100
   route-target export 100:100
   route-target import 100:100
   
   address-family ipv4
   exit-address-family
   
   address-family ipv6
   exit-address-family
   
   interface gigabitEthernet0/0/11
   vrf forwarding sr
   ip address 30.30.30.1 255.255.255.0

2. Tags VRF prefix with BGP community on R6 (tailend).

   route-map BGP_TE_MAP permit 9
   match ip address L3VPN_ODN_ROUTES
   set community 3276850
ip access-list extended L3VPN_ODN ROUTES
permit ip 30.30.30.1 255.255.0.0 any

3. Enable BGP on R6 (tailend) and R1 (headend) to advertise and receive VRF SR prefix and match on community set on R6 (tailend).

```plaintext
router bgp 100 <-- This BGP configuration applies to the tailend
  bgp router-id 172.16.0.1
  bgp log-neighbor-changes
  bgp graceful-restart
  no bgp default ipv4-unicast
  neighbor 10.0.0.2 remote-as 100
  neighbor 10.0.0.2 update-source Loopback0
  address-family ipv4
    neighbor 10.0.0.2 activate
    neighbor 10.0.0.2 send-community both
    neighbor 10.0.0.2 next-hop-self
  exit-address-family
  address-family vpnv4
    neighbor 10.0.0.2 activate
  exit-address-family
  address-family link-state link-state
  exit-address-family
  address-family ipv4 vrf sr
  redistribute connected
  exit-address-family
  route-map BGP_TE_MAP out
  exit-address-family
  address-family vpnv4
  neighbor 10.0.0.2 activate
  exit-address-family
  address-family link-state link-state
  exit-address-family
  address-family ipv4 vrf sr
  redistribute connected
  exit-address-family
  route-map BGP_TE_MAP in
  exit-address-family

route-map BGP_TE_MAP permit 9
match ip address traffic
set community 3276850
ip access-list extended traffic
permit ip 10.0.0.1 255.255.0.0 any

router bgp 100 <-- This BGP configuration applies to the headend
  bgp router-id 192.168.0.2
  bgp log-neighbor-changes
  bgp graceful-restart
  no bgp default ipv4-unicast
  neighbor 10.0.0.2 remote-as 100
  neighbor 10.0.0.2 update-source Loopback0
  address-family ipv4
    neighbor 10.0.0.2 activate
    neighbor 10.0.0.2 send-community both
    neighbor 10.0.0.2 next-hop-self
  exit-address-family
  address-family vpnv4
    neighbor 10.0.0.2 activate
  exit-address-family
  address-family link-state link-state
  exit-address-family
  address-family ipv4 vrf sr
  redistribute connected
  exit-address-family
  route-map BGP_TE_MAP in
  exit-address-family
  address-family vpnv4
  neighbor 10.0.0.2 activate
  exit-address-family
  address-family link-state link-state
  exit-address-family
  address-family ipv4 vrf sr
  redistribute connected
  exit-address-family
  route-map BGP_TE_MAP in
  exit-address-family

match community 1
set attribute-set BGP_TE5555
ip community-list 1 permit 3276850
mpls traffic-eng lsp attributes BGP_TE5555
```
path-selection metric igp
pce

4. Enable route-map or attribute set on headend (R1).

route-map BGP_TE_MAP permit 9
match community 1
set attribute-set BGP_TE5555
ip community-list 1 permit 3276850
mpls traffic-eng lsp attributes BGP_TE5555
path-selection metric igp
pce
end

5. Enable PCE and auto-tunnel configurations on R1.

mpls traffic-eng tunnels
mpls traffic-eng pcc peer 10.0.0.3 source 10.0.0.4 precedence 255
mpls traffic-eng auto-tunnel p2p tunnel-num min 2000 max 5000

6. Enable all core links with SR-TE configurations and ensure that they are enabled as point-to-point interfaces.

mpls traffic-eng tunnels
interface GigabitEthernet0/2/0
ip address 101.102.6.1 255.255.255.0
ip router isis 1
mpls traffic-eng tunnels
isis network point-to-point
interface GigabitEthernet0/3/1
vrf forwarding sr
ip address 101.107.3.1 255.255.255.0
negotiation auto
end

7. Enable R3 (RR) to advertise TED to the PCE server via BGP-LS.

router isis 1
net 49.0002.0000.0000.0003.00
ispf level-1-2
metric-style wide
nsf cisco
nsf interval 0
distribute link-state
segment-routing mpls
segment-routing prefix-sid-map advertise-local
redistribute static ip level-1-2
mpls traffic-eng router-id Loopback0
mpls traffic-eng level-1
mpls traffic-eng level-2
router bgp 100
bgp router-id 10.0.0.2
bgp log-neighbor-changes
bgp graceful-restart
no bgp default ipv4-unicast
neighbor 10.0.0.3 remote-as 100
neighbor 10.0.0.3 update-source Loopback0
address-family ipv4
neighbor 10.0.0.3 activate
exit-address-family

8. Enable PCE server configuration and verify that BGP-LS session is properly established with RR.
Device# `show bgp li li summary`
BGP router identifier 10.0.0.3, local AS number 100
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0x0 RD version: 1436
BGP main routing table version 1436
BGP NSR Initial initsync version 1 (Reached)
BGP NSR/ISSU Sync-Group versions 0/0
BGP scan interval 60 secs
BGP is operating in STANDALONE mode.
Process RcvTblVer bRIB/RIB LabelVer ImportVer SendTblVer StandbyVer
Speaker 1436 1436 1436 1436 1436
0
Neighbor Spk AS MagRcvd MsgSent TblVer InQ OutQ Up/Down St/PfxRcd
10.0.0.2 0 100 19923 17437 1436 0 0 1w2d 103
Device# `show pce ipv4 topo | b Node 3`
Node 3
TE router ID: 10.0.0.2
Host name: R3
ISIS system ID: 0000.0000.0003 level-1
ISIS system ID: 0000.0000.0003 level-2
Prefix SID:
  Prefix 10.0.0.2, label 20011 (regular)

---

**Note**
The above `show` command is run only on the PCE server, for example the Cisco ASR 9000 Series Aggregation Services Routers.

---

**Verifying Segment Routing On-Demand Next Hop—Layer 3 and Layer 3 VPN**

The ODN verifications are based on L3VPN VRF prefixes.

1. Verify that PCEP session between R1 (headend and PCE server) is established.

   Device# `show pce client peer`
   PCC’s peer database:
   ---------------
   Peer address: 10.0.0.3 (best PCE)
   State up
   Capabilities: Stateful, Update, Segment-Routing

2. Verify that PCEP session is established between all the peers (PCCs).

   Device# `show pce ipv4 peer`
   PCE’s peer database:
   ---------------
   Peer address: 10.0.0.4
   State: Up
   Capabilities: Stateful, Segment-Routing, Update
   Peer address: 172.16.0.5
   State: Up
   Capabilities: Stateful, Segment-Routing, Update
The above `show` command is run only on the PCE server, for example the Cisco ASR 9000 Series Aggregation Services Routers.

3. Verify that R1 (headend) has no visibility to R6 loopback address.

```
Device# show ip route 192.168.0.1
% Network not in table
```

4. Verify that VRF prefix is injected via MP-BGP in R1 VRF SR routing table.

```
Device# show ip route vrf sr
Routing Table: sr
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default, U - per-user static route
o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
a - application route
+ - replicated route, % - next hop override, p - overrides from PfR
Gateway of last resort is not set
C L
B
10.0.0.6/8 is variably subnetted, 2 subnets, 2 masks
  10.0.0.7/24 is directly connected, GigabitEthernet0/3/1
  10.0.0.7/32 is directly connected, GigabitEthernet0/3/1
10.0.0.8/24 is subnetted, 1 subnets
  10.0.0.9 [200/0] via binding label: 865, 4d21h
```

5. Verify that BGP is associating properly the policy and binding SID with the VRF prefix.

```
Device# show ip bgp vpnv4 vrf sr 106.107.4.0
BGP routing table entry for 100:100:106.107.4.0/24, version 3011
Paths: (1 available, best #1, table sr)
  Not advertised to any peer
  Refresh Epoch 4
  Local
    192.168.0.0/1 (metric 10) (via default) from 10.0.0.2 (10.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      Community: 3276850
      Extended Community: RT:100:100
      Originator: 192.168.0.1, Cluster list: 10.0.0.2
      mpls labels in/out nolabel/1085
      binding SID: 865 (BGP_TE5555)
      rx pathid: 0, tx pathid: 0x0
```

6. Verify binding label association with VRF prefix.

```
Device# show ip route vrf sr 106.107.4.0
Routing Table: sr
Routing entry for 106.107.4.0/24
  Known via "bgp 100", distance 200, metric 0, type internal
  Routing Descriptor Blocks:
    * Binding Label: 865, from 10.0.0.2, 4d22h ago
      Route metric is 0, traffic share count is 1
      AS Hops 0
```

---

**Verifying Segment Routing On-Demand Next Hop—Layer 3 and Layer 3 VPN**
7. Verify that VRF prefix is forwarded via ODN auto-tunnel.

Device# show ip cef label-table
Label Next Hop Interface
0 no route
865 attached Tunnel2000
Device# show ip cef vrf sr 106.107.4.0 detail
10.0.8.24, epoch 15, flags [rib defined all labels]
  recursive via 865 label 1085
  attached to Tunnel2000

8. Verify ODN auto-tunnel status.

Device# show mpls traffic-eng tunnels
P2P TUNNELS/LSPs:
Name: R1_t2000 (Tunnel2000) Destination: 192.168.0.1 Ifhandle: 0x6F5
(auto-tunnel for BGP TE)
Status:
  Admin: up Oper: up Path: valid Signalling: connected---.. auto-tunnel
Status:
  path option 1, (SEGMENT-ROUTING) (PCE) type dynamic (Basis for Setup, path weight
  10)
Config Parameters:
  Bandwidth: 0 kbps (Global) Priority: 7 7 Affinity: 0x0/0xFFFF
  Metric Type: IGP (interface)
  Path Selection:
  Protection: any (default)
  Path-selection Tiebreaker:
  Global: not set Tunnel Specific: not set Effective: min-fill (default)
  Hop Limit: disabled
  Cost Limit: disabled
  Path-invalidation timeout: 10000 msec (default), Action: Tear
  AutoRoute: disabled LockDown: disabled Loadshare: 0 [0] bw-based auto-bw: disabled
  Attribute-set: BGP_TE5555--- attribute-set
  Fault-OAM: disabled, Wrap-Protection: disabled, Wrap-Capable: No
  Active Path Option Parameters:
    State: dynamic path option 1 is active
    BandwidthOverride: disabled LockDown: disabled Verbatim: disabled
PCEP Info:
  Delegation state: Working: yes Protect: no
  Working Path Info:
  Request status: processed
  Created via PCRep message
  PCE metric: 30, type: IGP
  Reported paths:
    Tunnel Name: Tunnel2000_w
    LSPs:
      LSP[0]:
from PCE server: 10.0.0.3-- via PCE server
  source 10.0.0.4, destination 192.168.0.1, tunnel ID 2000, LSP ID 1
  State: Admin up, Operation active
  Binding SID: 865
  Setup type: SR
  Bandwidth: requested 0, used 0
  LSP object:
    PLSF-ID 0x807D0, flags: D:0 S:0 R:0 A:1 0:2
    Metric type: IGP, Accumulated Metric 0
    ERO:
      SID[0]: Adj, Label 2377, NAI: local 101.102.6.1 remote 10.0.0.10
      SID[1]: Unspecified, Label 17, NAI: n/a
SID(2): Unspecified, Label 20, NAI: n/a

History:
Tunnel:
  Time since created: 4 days, 22 hours, 21 minutes
  Time since path change: 4 days, 22 hours, 21 minutes
  Number of LSP IDs (Tun_Instances) used: 1
  Current LSP: [ID: 1]
    Uptime: 4 days, 22 hours, 21 minutes
  Tun_Instance: 1
Segment-Routing Path Info (isis level-1)
  Segment0[Link]: 101.102.6.1 - 10.0.0.10, Label: 2377
  Segment1[ - ]: Label: 17
  Segment2[ - ]: Label: 20


Device# show pce client lsp brief
PCC's tunnel database:
----------------------
Tunnel Name: Tunnel2000_w
  LSP ID 1
Tunnel Name: Tunnel2000_p
R1# sh pce client lsp detail
PCC's tunnel database:
----------------------
Tunnel Name: Tunnel2000_w
LSPs:
  LSP[0]:
    source 10.0.0.4, destination 192.168.0.1, tunnel ID 2000, LSP ID 1
    State: Admin up, Operation active
    Binding SID: 865
    Setup type: SR
    Bandwidth: requested 0, used 0
    LSP object:
      PLSP-ID 0x807D0, flags: D:0 S:0 R:0 A:1 O:2
      Metric type: IGP, Accumulated Metric 0
      ERO:
        SID[0]: Adj, Label 2377, NAI: local 101.102.6.1 remote 10.0.0.10
        SID[1]: Unspecified, Label 17, NAI: n/a
        SID[2]: Unspecified, Label 20, NAI: n/a

10. Verify ODN LSP status on the PCE server.

Device# show pce lsp summary
PCE's LSP database summary:
----------------------------
All peers:
  Number of LSPs: 1
  Operational: Up: 1 Down: 0
  Admin state: Up: 1 Down: 0
  Setup type: RSVP: 0 Segment routing: 1
  Peer 10.0.0.4:
    Number of LSPs: 1
    Operational: Up: 1 Down: 0
    Admin state: Up: 1 Down: 0
    Setup type: RSVP: 0 Segment routing: 1

11. Verify detailed LSP information on the PCE server.

Device# show pce lsp detail
PCE's tunnel database:
----------------------
12. Shut down the interface that is connected to VRF SR so that the prefix is no longer advertised by MP-BGP.

   Device# interface GigabitEthernet0/2
   Device(config-if)# shut

13. Verify that VRF prefix is no longer advertised to R1 (headend) via R6 (tailend).

   Device# show ip route vrf sr
   Routing Table: sr
   Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
         D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
         N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
         E1 - OSPF external type 1, E2 - OSPF external type 2
         i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
         *i - IS-IS inter area, * - candidate default, U - per-user static route
         G - OSPF external gateway link, a - application route
         + - replicated route, % - next hop override, p - overrides from PfR

   Gateway of last resort is not set
   10.0.0.6/8 is variably subnetted, 2 subnets, 2 masks
   C 10.0.0.7/24 is directly connected, GigabitEthernet0/3/1
   L 10.0.0.8/32 is directly connected, GigabitEthernet0/3/1

14. Verify that no ODN auto-tunnel exists.

   Device# show mpls traffic-eng tunnels
   P2P TUNNELS/LSPs:
   P2MP TUNNELS:
   P2MP SUB-LSPs:
Configuring Color Extended Community, Affinity Constraint, and Disjointness Constraint

Consider the following topology:

Configuring Color Extended Community

SR-TE Policy Headend Configuration on Node R3

```config
segment-routing traffic-eng
on-demand color 100
authorize restrict
ipv4 prefix-list R9350_BGP_INTER_DOMAIN
candidate-paths
preference 1
constraints
segments
dataplane mpls
!
!
dynamic
pcep
!
!
!
!
posr
pce address <pce loopback ip>source-address <pcc loopback ip>
!
```

SR-TE Policy Tailend Configuration on Node R9

```config
route-map R9_R3_R5_R2_BGP_INTER_DOMAIN permit 10
match ip address prefix-list R9350_BGP_INTER_DOMAIN
```
set extcommunity color 100  

--------------- Extended Color community configuration  
route-map R9_R3_R5_R2_BGP_INTER_DOMAIN permit 20  
ip prefix-list R9350_BGP_INTER_DOMAIN seq 35 permit 50.0.0.0/11 le 32  
router bgp 1  
    address-family vpnv4  
    neighbor 201.201.201.201 activate  
    neighbor 201.201.201.201 send-community both  
    neighbor 201.201.201.201 route-map R9_R3_R5_R2_BGP_INTER_DOMAIN out  
    neighbor 206.206.206.206 activate  
    neighbor 206.206.206.206 send-community both  
    neighbor 206.206.206.206 route-map R9_R3_R5_R2_BGP_INTER_DOMAIN out  
    exit-address-family  

In the SR-TE ODN color template, to select the metric type, choose either  
gp or te:

Router(config-srte-odn-path-pref-dyn-metric)# type ?  
  igp Specify IGP metric  
  te Specify TE metric  

Configuring Affinity Constraint

segment-routing traffic-eng  
interface GigabitEthernet0/2/3  
    affinity  
        name 1  
    on-demand color 100  
    authorize restrict  
        ipv4 prefix-list R9350_BGP_INTERDOMAIN  
    candidate-paths  
        preference 1  
    constraints  
        segments  
            dataplane mpls  

!  
affinity ------------------- Affinity configuration  
   include-any --------------- Affinity Type configuration  
       name 1 ------------------ Affinity Name configuration  
!
  !  
  !  
  !  
  !  
    dynamic  
    pcep  
    !  
    !  
    pcc  
    pce address <pce loopback ip> source-address <pcc loopback ip>  
!
affinity-map ------------------- Affinity Map configuration  
    name 1 bit-position 1  

Configuring Disjointness Constraint

segment-routing traffic-eng  
on-demand color 100  
authorize restrict  
ipv4 prefix-list R9350_BGP_INTER_DOMAIN  
candidate-paths  
preferred 1  
candidate-paths  
segments
Verifying SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint

SR-TE Policy Name: 209.209.209.209\|100

Router# show segment-routing traffic-eng policy name 209.209.209.209\|100
Name: 209.209.209.209\|100 (Color: 100 End-point: 209.209.209.209)
Status: Admin: up, Operational: up for 51:34:38 (since 01-07 06:19:08.040) Policy state is UP
Candidate-paths:
Preference 1:
Constraints:
Affinity:
include-any: -------------- Affinity Type
1 -------------- Affinity Name
Disjointness information:
Group ID: 1, Source: 1.0.0.0
Type: Node Disjointness ------ Disjointness Type
Dynamic (pce 12.12.12.12) (active) ------- PCE Computed Candidate-path
Weight: 0, Metric Type: TE -------------- Metric Type
Metric Type: TE, Path Accumulated Metric: 53 ---- Total IGP Metric from Source to
Destination
18007 [Prefix-SID, 211.211.211.211] --------------
18002 [Prefix-SID, 207.207.207.207] -------------- This Segment List should follow Affinity path
21 [Adjacency-SID, 10.10.20.2 - 10.10.20.1] --------------
Attributes:
Binding SID: 87 ------------------- Binding SID Allocated
Allocation mode: dynamic
State: Programmed
Auto-policy info:
To view detailed information about SR-TE Policy 209.209.209.209|100

Router# show segment-routing traffic-eng policy name 209.209.209.209|100 detail
Name: 209.209.209.209|100 (Color: 100 End-point: 209.209.209.209)
Status:
  Admin: up, Operational: up for 00:04:19 (since 01-10 06:20:57.810)
Candidate-paths:
  Preference 1:
  Constraints:
    Affinity: include-any:
    1
Disjointness information:
  Group ID: 1, Source: 1.0.0.0
  Type: Node Disjointness
  Dynamic (pce 12.12.12.12) (active)
  Weight: 0, Metric Type: TE
  Metric Type: TE, Path Accumulated Metric: 53
  18007 [Prefix-SID, 211.211.211.211]
  18002 [Prefix-SID, 207.207.207.207]
  21 [Adjacency-SID, 10.10.20.2 - 10.10.20.1]
Attributes:
  Binding SID: 87
  Allocation mode: dynamic
  State: Programmed
Auto-policy info:
  Creator: BGP SR Policy Client
  IPv6 caps enable: yes
Forwarding-ID: 65711 (0x44) ------- This FWD-ID is used for forwarding traffic
Stats:
  Packets: 8893 Bytes: 852848 ------- This counter indicates traffic flowing through this SRTE policy
Event history: --- This indicates event happened with this SRTE Policy
  Timestamp  Value  Client  Event type  Context:
  --------  -------  ------  ---------  --------:
  01-06 05:59:26.096  BGP SR Policy Cl  Policy created  Name:
  209.209.209.209|100
  01-06 05:59:26.096  BGP SR Policy Cl  Set colour  Colour: 100
  01-06 05:59:26.096  BGP SR Policy Cl  Set end point  End-point:
  209.209.209.209
  01-06 05:59:26.096  BGP SR Policy Cl  Set dynamic pce  Path option:
  dynamic pce
  01-06 05:59:26.480  FH Resolution  Policy state UP  Status:
  PATH RESOLVED
  01-06 05:59:40.424  FH Resolution  REOPT triggered  Status:
  REOPTIMIZED
  01-06 05:59:49.249  FH Resolution  REOPT triggered  Status:
  REOPTIMIZED
  01-06 05:59:56.469  FH Resolution  REOPT triggered  Status:
  REOPTIMIZED
  01-07 05:15:19.918  FH Resolution  Policy state DOWN  Status:
  PATH NOT RESOLVED
  01-07 06:15:55.739  FH Resolution  Policy state UP  Status:
  PATH RESOLVED
  01-07 06:16:08.552  FH Resolution  REOPT triggered  Status:
To check if the Affinity constraint is working, shut down any of the interfaces falling under the Affinity-defined path. If the constraint works, the SR-TE policy goes down instead of taking the another path (if available) to reach to the destination.

To check if the disjointness constraint is working, check the SR-TE policy information given by the PCE, which consists of Segment IDs used for the computed path from source to destination.

Disjointness constraint works, if the Segment IDs of both the SR-TE policies are different. For example:

**SRTE Policy 1:**
- SID[0]: Node, Label 16002, NAI: 207.207.207.207
- SID[1]: Node, Label 16004, NAI: 201.201.201.201
- SID[2]: Node, Label 16011, NAI: 205.205.205.205

**SRTE Policy 2:**
- SID[0]: Node, Label 16003, NAI: 208.208.208.208
- SID[1]: Node, Label 16004, NAI: 201.201.201.201
- SID[2]: Node, Label 16011, NAI: 205.205.205.205

SID[2] of policies 1 and 2 is the same since destination of both the SR-TE policies is the same.

To view the SR-TE policy and Affinity constraint in the PCE:
```
RP/0/RSP0/CPU0:ASR9K# show pce lsp pcc ipv4 213.213.213.213 private
```

Thu 00:11:52.983 UTC

PCE's tunnel database:
----------------------
PCC 213.213.213.213:
Tunnel Name: 209.209.209.209|100
LSPs:
- LSP[0]:
  - source 203.203.203.203, destination 209.209.209.209, tunnel ID 177, LSP ID 0
  - State: Admin up, Operation SRTE Policy is up
  - Setup type: Segment Routing
  - Binding SID: 87
  - Maximum SID Depth: 4
  - Absolute Metric Margin: 0
  - Relative Metric Margin: 0%
  - Affinity: exclude-any Ox0 include-any Ox2 include-all Ox0
    This indicates Affinity taken into account by PCE

PCEP information:
- PLSP-ID 0x800b1, flags: D:1 S:0 R:0 A:1 O:2 C:0
- LSP Role: Disjoint LSP
- State-sync PCE: None
- PCC: 213.213.213.213
- LSP is subdelegated to: None
Reported path:
Metric type: TE, Accumulated Metric 53
SID[1]: Node, Label 18007, Address 211.211.211.211
SID[2]: Node, Label 18002, Address 207.207.207.207
SID[3]: Adj, Label 21, Address: local 10.10.20.2 remote 10.10.20.1

Computed path: (Local PCE)
Computed Time: Thu Jan 10 00:09:36 UTC 2019 (00:02:17 ago)
Metric type: TE, Accumulated Metric 53
SID[1]: Node, Label 18007, Address 211.211.211.211
SID[2]: Node, Label 18002, Address 207.207.207.207
SID[3]: Adj, Label 21, Address: local 10.10.20.2 remote 10.10.20.1

Recorded path:
None

Disjoint Group Information:
Type Node-Disjoint, Group 1, Sub-Group 1.0.0.0

Event history (latest first):

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
</table>
| Thu Jan 10 00:09:37 UTC 2019 | Report from 213.213.213.213 (LSP owner)  
Symbolic-name: 209.209.209.209|100, LSP-ID: 0,  
Source: 203.203.203.203 Destination: 209.209.209.209,  
D:1, R:0, A:1 O:2, Sig.BW: 0, Act.BW: 0  
Reported Path: (Metric 53)  
Label 18007, Address 211.211.211.211  
Label 18002, Address 207.207.207.207 |
| Thu Jan 10 00:09:36 UTC 2019 | Update to 213.213.213.213 (PCC)  
Symbolic-name: 209.209.209.209|100, LSP-ID: 0, D:1  
Path: (Metric 53)  
Label 18007, Address 211.211.211.211  
Label 18002, Address 207.207.207.207  
Label 21, Address: local 10.10.20.2 remote 10.10.20.1  
Chng:0, AssoChng:0 |
| Thu Jan 10 00:09:36 UTC 2019 | Path Computation (Disjoint LSP)  
Symbolic-name: 209.209.209.209|100, LSP-ID: 0, D:1  
Status: Disjoint Path Success |
| Wed Jan 09 23:54:42 UTC 2019 | Update to 213.213.213.213 (PCC)  
Symbolic-name: 209.209.209.209|100, LSP-ID: 0, D:1  
Path: (Metric 53)  
Label 18007, Address 211.211.211.211  
Label 18002, Address 207.207.207.207  
Label 21, Address: local 10.10.20.2 remote 10.10.20.1 |
| Wed Jan 09 23:54:42 UTC 2019 | Path Computation (Disjoint LSP)  
Symbolic-name: 209.209.209.209|100, LSP-ID: 0, D:1  
Status: Fallback Node to Shortest Path  
Computed Path: (Metric 53)  
Label 18007, Address 211.211.211.211  
Label 18002, Address 207.207.207.207  
Label 21, Address: local 10.10.20.2 remote 10.10.20.1 |
| Wed Jan 09 23:54:21 UTC 2019 | Path Computation (Disjoint LSP)  
Symbolic-name: 209.209.209.209|100, LSP-ID: 0, D:1  
Troubleshooting the SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint

If SR-TE policy is down, check the status of the SR-TE Policy under the SR-TE policy information

Router# show segment-routing traffic-eng policy name 209.209.209.209|100
Name: 209.209.209.209|100 (Color: 106 End-point: 209.209.209.209)
Status:
  Admin: up, Operational: down for 00:00:18 (since 01-10 13:06:42.142)
Candidate-paths:
  Preference 1:
  Constraints:

This indicates that Disjointness between SRTE Policies is working
The possible reasons for the policy being down are:

- Connection to PCE is down.
- Max SID depth is exceeded.
- An interface falling under Affinity-defined path from source to destination has been shut down.

To check the SR-TE policy status on the PCE:

```
RP/0/RSP0/CPU0:ASR9K#show pce lsp pcc ipv4 213.213.213.213 private
```

```
Thu Jan 10 00:11:52.983 UTC
PCE's tunnel database:
----------------------
PCC 213.213.213.213:
Tunnel Name: 209.209.209.209|100
LSPs:
LSP[0]:
  source 203.203.203.203, destination 209.209.209.209, tunnel ID 177, LSP ID 0
  State: Admin up, Operation active
  ------------------ SRTE Policy is up
  Setup type: Segment Routing
  Binding SID: 87

Maximum SID Depth: 4
Absolute Metric Margin: 0
Relative Metric Margin: 0%

Affinity: exclude-any 0x0 include-any 0x2 include-all 0x0  ---□ This indicates Affinity is taken into account by the PCE

PCE is aware of the network topology. This information is used for path computation using the following command. This information is also used to determine if nodes and links are present and have the expected attributes (IGP/TE admin weights, SIDs and so on).

```
RP/0/RSP0/CPU0:ASR9K# show pce ipv4 topology
```

```
Tue Jan 15 01:36:20.298 UTC
PCE's topology database - detail:
-----------------------------
Node 1
  TE router ID: 207.207.207.207
  Host name: 920-R7
  ISIS system ID: 0000.0000.0207 level-1 ASN: 1
  ISIS system ID: 0000.0000.0207 level-2 ASN: 1
  Prefix SID:
```
ISIS system ID: 0000.0000.0207 level-1 ASN: 1 domain ID: 0
Prefix 207.207.207.207, label 16002 (regular), flags: N
ISIS system ID: 0000.0000.0207 level-1 ASN: 1 domain ID: 0
Prefix 207.207.207.207, label 18002 (strict), flags: N
ISIS system ID: 0000.0000.0207 level-2 ASN: 1 domain ID: 0
Prefix 207.207.207.207, label 16002 (regular), flags: N
ISIS system ID: 0000.0000.0207 level-2 ASN: 1 domain ID: 0
Prefix 207.207.207.207, label 18002 (strict), flags: N
SRGB INFO:
ISIS system ID: 0000.0000.0207 level-1 ASN: 1
SRGB Start: 16000 Size: 8000
ISIS system ID: 0000.0000.0207 level-2 ASN: 1
SRGB Start: 16000 Size: 8000
Link[0]: local address 10.10.21.1, remote address 10.10.21.2
Local node:
ISIS system ID: 0000.0000.0207 level-1 ASN: 1
Remote node:
  TE router ID: 208.208.208.208
  Host name: 920-R8
  ISIS system ID: 0000.0000.0208 level-1 ASN: 1
  Metric: IGP 10, TE 10, Latency 10
  Bandwidth: Total 1250000000 Bps, Reservable 0 Bps
  Admin-groups: 0x00000000
  Adj SID: 16 (unprotected) 17 (protected)
Link[1]: local address 10.10.21.1, remote address 10.10.21.2
Local node:
ISIS system ID: 0000.0000.0207 level-2 ASN: 1
Remote node:
  TE router ID: 208.208.208.208
  Host name: 920-R8
  ISIS system ID: 0000.0000.0208 level-2 ASN: 1
  Metric: IGP 10, TE 10, Latency 10
  Bandwidth: Total 1250000000 Bps, Reservable 0 Bps
  Admin-groups: 0x00000000
  Adj SID: 18 (unprotected) 19 (protected)
Link[2]: local address 10.10.21.1, remote address 10.10.21.2
Local node:
ISIS system ID: 0000.0000.0209 level-1 ASN: 1
ISIS system ID: 0000.0000.0209 level-2 ASN: 1
Prefix SID:
ISIS system ID: 0000.0000.0207 level-1 ASN: 1 domain ID: 0
Prefix 209.209.209.209, label 16001 (regular), flags: N
ISIS system ID: 0000.0000.0207 level-1 ASN: 1 domain ID: 0
Prefix 209.209.209.209, label 18001 (strict), flags: N
ISIS system ID: 0000.0000.0207 level-2 ASN: 1 domain ID: 0
Prefix 209.209.209.209, label 16001 (regular), flags: N
ISIS system ID: 0000.0000.0207 level-2 ASN: 1 domain ID: 0
Prefix 209.209.209.209, label 18001 (strict), flags: N
SRGB INFO:
ISIS system ID: 0000.0000.0209 level-1 ASN: 1
  SRGB Start: 16000 Size: 8000
Further troubleshooting tips:

- Enable the following debug commands on the PCCs:
  - debug segment-routing traffic-eng path
  - debug segment-routing traffic-eng pcalc
  - debug segment-routing traffic-eng policy
  - debug segment-routing traffic-eng topology
  - debug segment-routing traffic-eng ha

- Enable the following debug commands on the PCE:
  - debug pce pcep
  - debug pce cspf
  - debug pce cspf-internal
  - debug pce error
  - debug pce path
Troubleshooting the SR-TE ODN Color Extended Community, Affinity Constraint, and Disjointness Constraint