You make **possible**

How to migrate a MPLS network from RSVP-TE to SR-TE

Thomas Wang, Technical Marketing Engineer

BRKMPL-2130

Cisco live!
June 9-13, 2019 · San Diego, CA
#CLUS

Cisco Webex Teams

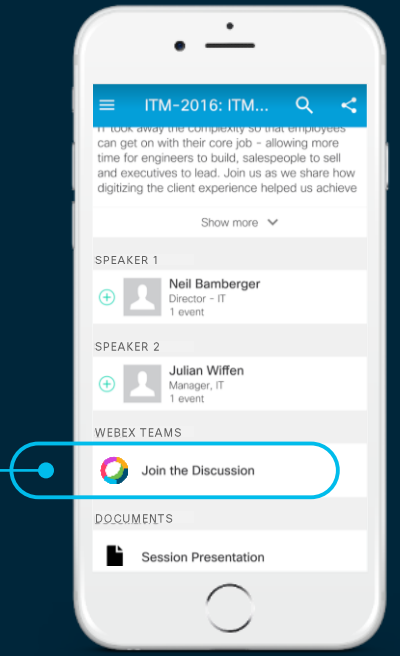
Questions?

Use Cisco Webex Teams to chat with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install Webex Teams or go directly to the team space
4. Enter messages/questions in the team space

Webex Teams will be moderated by the speaker until June 16, 2019.



cs.co/ciscolivebot# BRKMPL-2130

Agenda

- Welcome and Introduction
- RSVP-TE Network Challenges
- Solution Overview of Network Migration
- SR-TE Policies and Traffic Steering
- RSVP-TE Bandwidth Accounting
- Key Takeaways

Objective

- Introduce a solution that facilitates migration by allowing the coexistence of non-zero bandwidth RSVP-TE tunnel and Segment Routing in same network domain
  - SR-TE Policy Setup and Traffic Steering Techniques
  - SR Traffic Accounting and RSVP-TE resvBW Refresh

Reason of Network Migration

RSVP-TE Network Challenges and SR Benefits

You make networking **possible**

Cisco live!

Today Business Drivers

Scalability

- Distribute massive traffic volumes
- High-scale system

Optimization

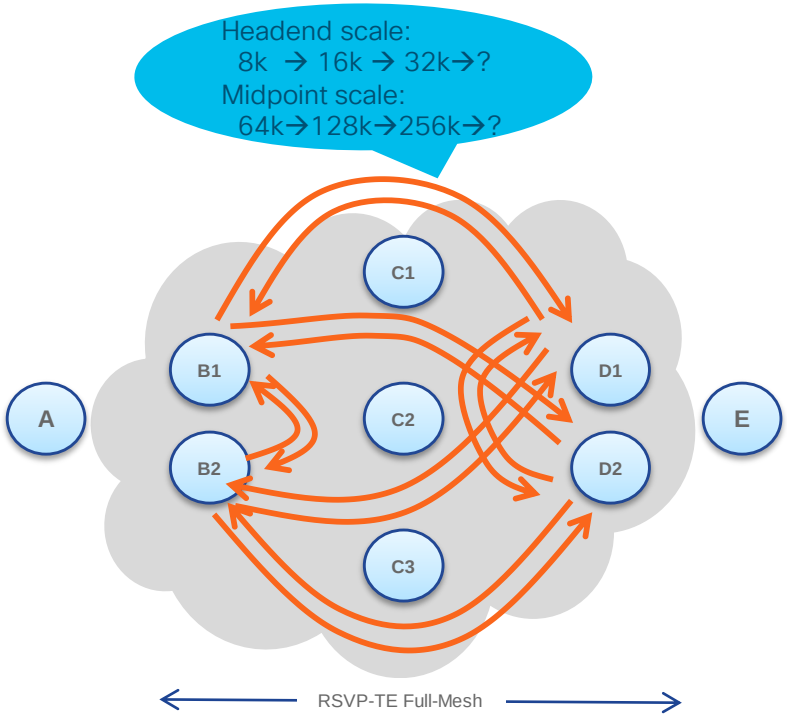
- Feature richness to optimize traffic steering
- API programmability

Automation

- Telemetry and visibility
- Model driven and machine learning

RSVP-TE Network Challenges

- Complicated MPLS WAN network
- Full-mesh of RSVP-TE Tunnels
  - ❖ RSVP state maintained at every hop
  - ❖ Large headend and midpoint scale
  - ❖ Core state n^2 problem
    - ❖ Compounded by emerging new service SLAs
- Doesn't scale well as network continues to growth

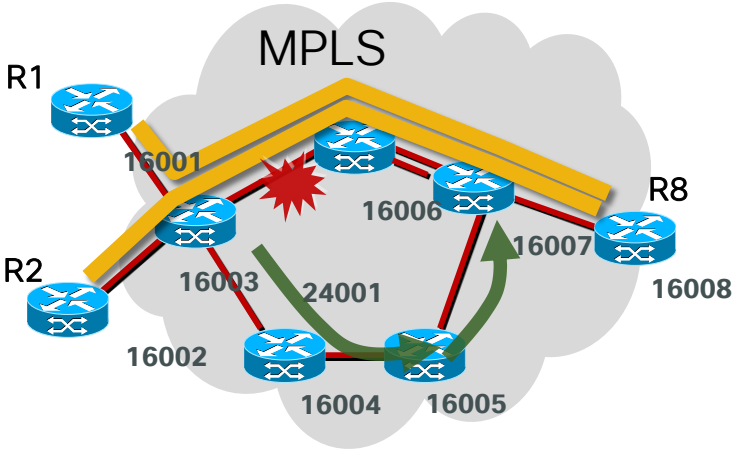


Why Segment Routing?

- Simple
  - Reduced control plane state
  - Utilize existing MPLS data plane
  - Path Control with SR source routing label stack
- Agile and Scalable
  - Eliminate LDP, RSVP protocols
- Programmable
  - Centralized path computation option
  - Ability for application integration

SR Path Control and Ti-LFA

Segment Routing Traffic Engineering (SR-TE)



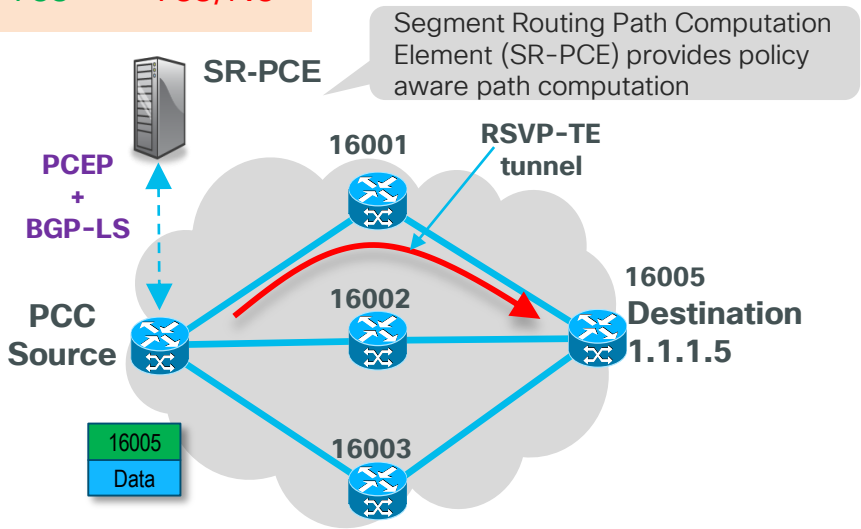
- Source Routing
  - Source chooses a path and encodes in the packet header as an ordered list of segments
  - The rest of the network nodes executes the encoded instructions
- SR-TE Policy Path Control
  - Policy label stack with Node-SID, or Adj-SID
  - Each Policy assigned unique Binding-SID
  - ECMP paths load-balance by IGP Nature
- Topology Independent LFA
  - Automated 1:N path protection
  - Local reroute comparable to MPLS TE Link / Node, but no RSVP or LDP
  - IGP algorithm, support microloop avoidance

SR Centralized Path Computation

	SR-TE	RSVP-TE
TE state only at head-end	Yes	No
ECMP-capability for TE	Yes	No
Engineered for SDN	Yes	Yes/No

```
segment-routing
 traffic-eng
 pcc
   pce address ipv4 192.99.1.1
 !
  policy foo123
   color 10 end-point ipv4 1.1.1.5
   autoroute
   candidate-paths
    preference 100
     dynamic pcep
```

SR Centralized Path Computation



RSVP-TE to SR-TE Seamless Migration

WHY

- Simple, easy and scalable
  - Avoid LDP or RSVP-TE sessions, less protocols to operate
  - Avoid midpoint of RSVP-TE LSP's n^2 scale problem
- Enables Application Driven Programmability:
  - Programmatic interfaces and Orchestration

HOW

- MPLS data plane leveraged without any modification
  - Push, swap and pop: all SR forwarding need
  - Inherit SR label ECMP load balance
- SR Traffic engineering via IGP Source Routing
  - Source routing path control push as a SR label or stack of labels

Solution Overview of Network Migration

SR-TE Policies and Traffic Steering

You make networking **possible**

Cisco live!

Network Migration Approaches

Interworking

- SR and RSVP-TE deployed in non overlapping areas of network
  - Isolated SR domains interconnected over RSVP-TE core
  - Within RSVP-TE domain, SR-TE label stack is carried inside RSVP-TE path

Ships-in-the-night

- SR-TE and RSVP-TE coexist over same network domain
  - Some services are migrated to SR, others retained on existing RSVP-TE
  - RSVP-TE link-management admission-control unaware of SR traffic utilization

RSVP-TE and SR-TE Coexistence

- In control plane: RSVP-TE and SR-TE independently program MPLS label forwarding entries
- In data plane: RSVP-TE and SR-TE use distinct MPLS labels

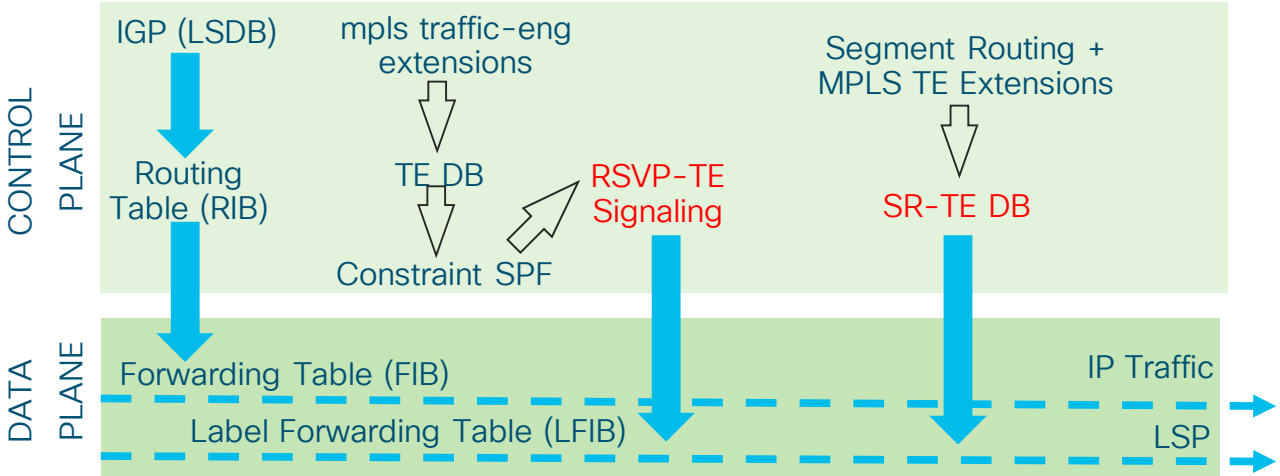
IP and Label Switch Path Install

RSVP-TE

- RSVP-TE Signaling (label distribution)
- Autoroute Announce via IGP
- Fast-reroute link/node protection

SR-TE

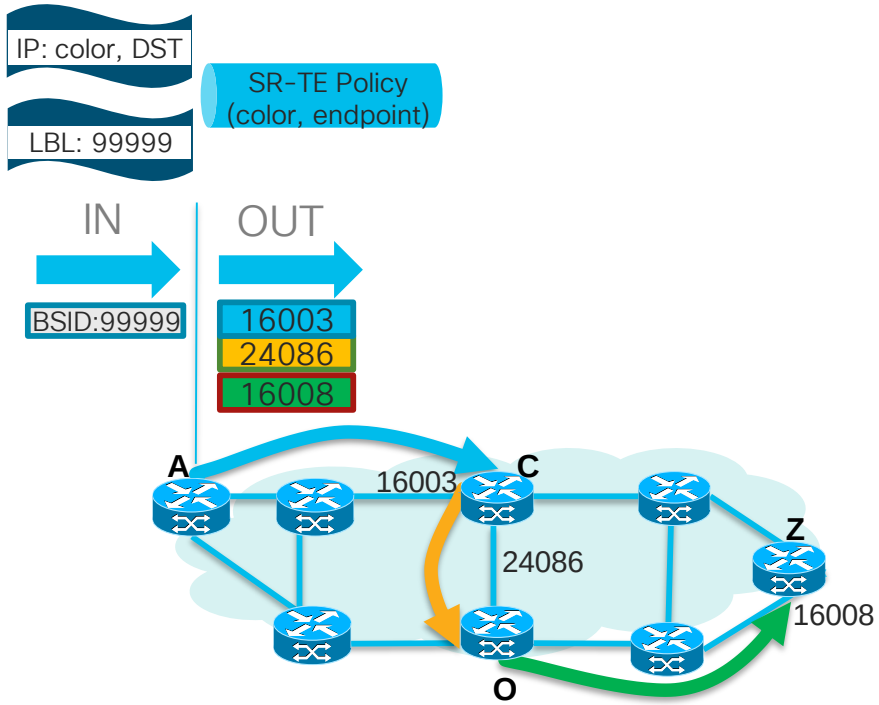
- IGP Label distribution, inherent ECMP
- Auto Steering via BGP Prefix Coloring
- Topology Independent LFA



SR-TE Auto-Steering with BGP Prefix Coloring

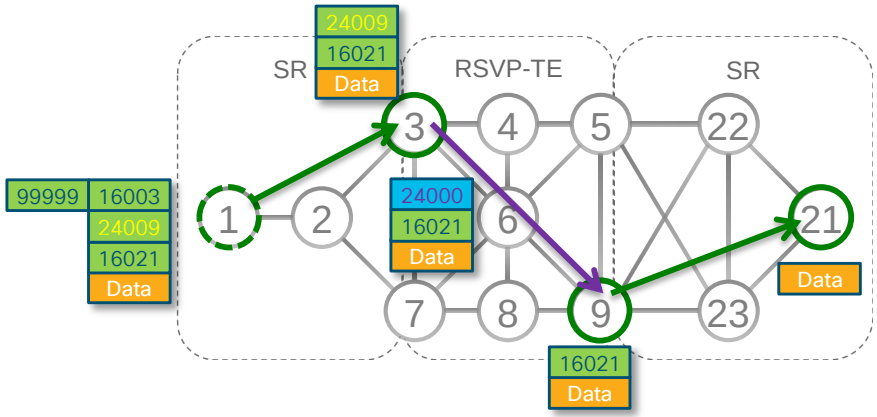
SR-TE provides more granular and automated steering techniques

- Egress BGP route-policy advertised prefixes colors using extended community attributes
- Incoming traffic is steered over SR-TE policy:
  - IP per flow traffic is steered if BGP prefix matches color and endpoint
  - labeled traffic is steered if packet top label matches policy Binding-SID
- Auto-steering applies in same way regardless of local or automatic SR-TE policy types
- If no valid SR-TE policy exists, prefixes will be installed "classically" in the forwarding plane by recursive lookup on the route to next-hop



SR-TE to RSVP-TE Interworking

- Extends an existing network where RSVP-TE is deployed
  - ❖ Enable SR in the network in islands, leave legacy RSVP-TE network in places
  - ❖ SR-TE interwork with RSVP-TE through Binding-SID as part of SRTE label stack
- SR-TE Policy "stitches" RSVP-TE tunnel by using Binding-SID of RSVP-TE tunnel

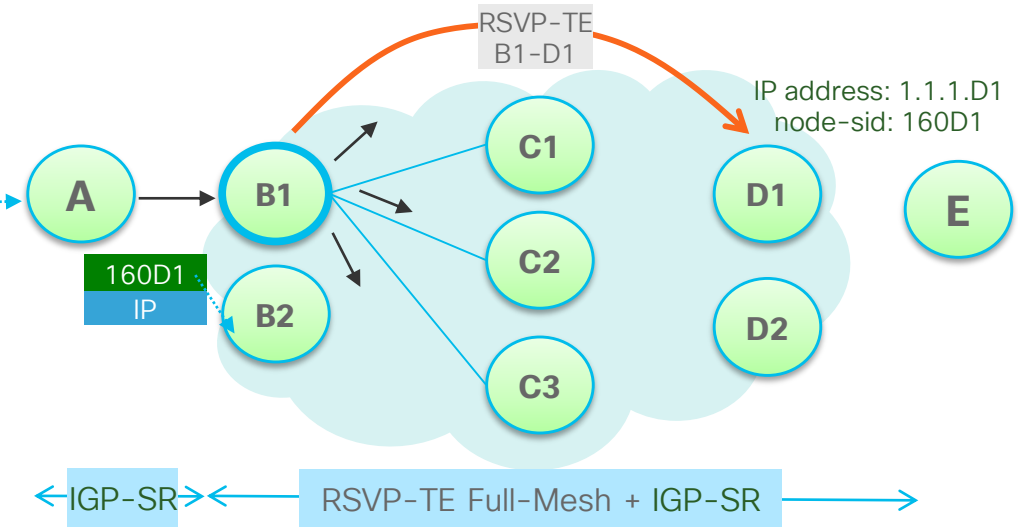


Ships-in-the-night: SR-TE and RSVP-TE Coexist

- Some services deployed over SR, others over existing transport (LDP or RSVP-TE)
- Data plane routes to native IGP SR path or SRTE policy by color and B-SID match
- Gradually move traffic from RSVP-TE Autoroute Announce steering to SR-TE color based Auto-steering

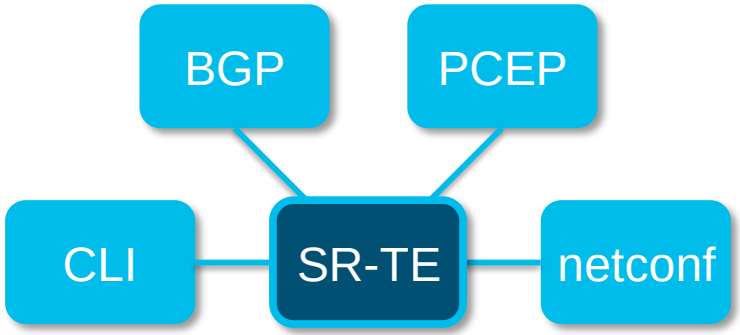
```
int tunnel-te 1000
 destination D1
 autoroute-announce
   exclude-traffic
    segment-routing
```

```
R1# show mpls traffic-eng autoroute
  Destination 1.1.1.5 has 1 tunnels in IS-IS
10 level 2
    tunnel-te1000 (traffic share 0, nexthop
1.1.1.5,  metric 0)
      (IS-IS 10 level-2, IPV4 Unicast)
      Signalled-Name: R1_t1000
        Exclude-traffic: segment-routing
```



Traffic Engineering with Segment Routing

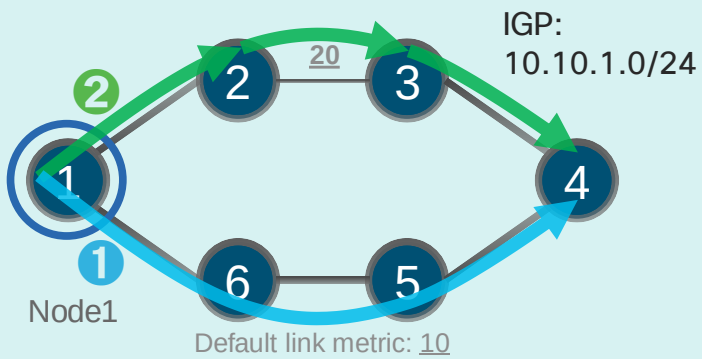
- Automated On-Demand Policy
  - No core state: state in the packet header
  - Automated steering: BGP prefix coloring
  - No tunnel interface: on-demand policy instantiation
- CLI or Controller Based Policy
  - Support constraint-based routing
  - CLI is just one way, more programmable ways
  - Support controller based BGP sr-policy instantiation
- Centralized Multi-Domain Policy
  - SR Path Compute Element (SR-PCE) for path compute
  - Binding-SID (BSID) for traffic steering and scale
  - Supports centralized policy via PCEP provision



Use Case1: CLI SRTE Policy with Autoroute

```
segment-routing
 traffic-eng
  policy POLICY1
   color 2 end-point ipv4 1.1.1.4
   autoroute
    include ipv4 10.10.1.0/24
   binding-sid mpls 999
   candidate-paths
1   preference 100
     dynamic
      metric
       type te
2   preference 200
     explicit segment-list SIDLIST1
   !
   segment-list name SIDLIST1
    index 10 mpls label 16002
    index 20 mpls label 24004
    index 30 mpls label 16004
```

- BGP installs SR-TE policy data path with policy B-SID
  - Path with higher preference is preferred
- SR-TE policy autoroute includes specific prefix
  - Remaining IGP autoroute traffic over RSVP-TE tunnels is not impacted



L2VPN Service Preferred-path Steering

- Use preferred-path configuration to specify SR-TE Policy used to transport Pseudowire traffic

```
l2vpn
 pw-class EoMPLS-PWCLASS
  encapsulation mpls
   preferred-path sr-te policy policy
srte_c_2_ep_1.1.1.4
 !
 xconnect group vpws
  p2p vpws1
   interface Bundle-Ether2.2
   neighbor ipv4 1.1.1.4 pw-id 1
    pw-class EoMPLS-PWCLASS
!
```

```
R1# show l2vpn xc pw-id 1 detail
Group vpws, XC vpws1, state is up; Interworking none
  AC: Bundle-Ether2.2, state is up
...
  PW: neighbor 1.1.1.4, PW ID 1, state is up ( established )
    PW class xc-vpls, XC ID 0xa000001f
    Encapsulation MPLS, protocol LDP
    Source address 1.1.1.1
    PW type Ethernet, control word disabled, interworking
none
    PW backup disable delay 0 sec
    Sequencing not set
    Preferred path Active : SR TE srte_c_2_ep_1.1.1.4,
Statically configured, fallback disabled
    Tunnel: Up
```

```
R1# show segment-routing traffic-eng policy name
srte_c_2_ep_1.1.1.4
Color: 2, End-point: 1.1.1.4
  Name: srte_c_2_ep_1.1.1.4
  Status:
    Admin: up  Operational: up for 02:48:30 (since Apr  1
12:22:57.663)
  Candidate-paths:
    Preference: 200 (configuration) (active)
      Name: POLICY1
      Requested BSID: 999
      PCC info:
        Symbolic name: cfg_ POLICY1_discr_200
        PLSP-ID: 4
      Explicit: segment-list SIDLIST1  (valid)
        Weight: 1, Metric Type: TE
          16002 [Prefix-SID, 1.1.1.2]
          24004 [Adjacency-SID, 10.23.0.1 - 10.23.0.2]
          16004 [Prefix-SID, 1.1.1.4]
    Preference: 100 (configuration)
      Requested BSID: 999
      PCC info:
        Symbolic name: cfg_POLICY1_discr_100
        PLSP-ID: 6
```

Use Case2: On-demand SRTE Policy

- Service source automatically instantiates an SRTE Policy to a BGP next-hop on-demand based on a BGP prefix color-template
  - One binding-sid per auto-policy, no need to pre-configured
- Service headend installs Policy data path based on binding-sid
  - Color match Auto-steering, no impact on RSVP-TE IGP autoroute traffic
- BGP prefix traffic automatically steered over SR-TE Policy based on Color community and BGP next-hop
  - Color community is used as SLA indicator
  - Each SR Policy is defined by (color, endpoint)

Internet Service On-demand Policy Steering

IF

BGP prefix next-hop and color match SR-TE Policy endpoint and color

=>

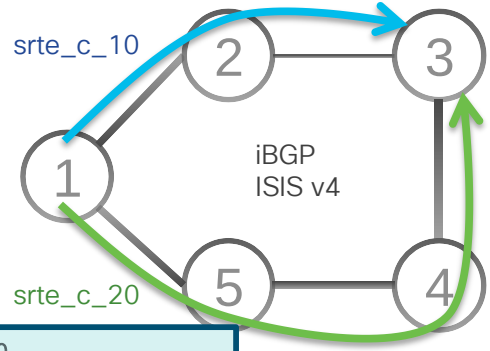
THEN

BGP route is installed and resolved with SR-TE Policy BSID

```
segment-routing
 traffic-eng
 on-demand color 10
  dynamic metric type igp
 on-demand color 20
  dynamic metric type te
!

110.1.1.3/32 (color 10, NH 1.1.1.3, srte_c_10)
120.1.1.3/32 (color 20, NH 1.1.1.3, srte_c_20)
```

IGP shortest path

better path with lowest TE metric




```
R1# show bgp ipv4 unicast 120.1.1.0/24 detail | beg C:20
 1.1.1.3 C:20 (bsid:24025) (metric 20) from 1.1.1.3 (1.1.1.3)
     Origin incomplete, metric 0, localpref 100, valid, internal, best, group-best
     Received Path ID 0, Local Path ID 1, version 25
     Extended community: Color:20
     SR policy color 20, up, registered, bsid 24025, if-handle 0x00000270
R1# show cef ipv4 120.1.1.1/32 detail | beg local-label
   via local-label 24025, 3 dependencies, recursive [flags 0x6000]
    Hash  OK  Interface                 Address
    0     Y   srte_c_20_ep_1.1.1.3      point2point
```

```
extcommunity-set opaque BLUE
  10
end-set
!
extcommunity-set opaque GREEN
  20
end-set
!
route-policy SET_COLOR
  if destination in (110.1.1.0/24) then
    set extcommunity color BLUE
  endif
  if destination in (120.1.1.0/24) then
    set extcommunity color GREEN
  endif
  pass
end-policy
!
router bgp 1
 neighbor 1.1.1.1
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
   route-policy SET_COLOR out
```

Use Case3: Dynamic VPN SRTE Policy

- Egress PE5 populates a vrf prefix 11.11.1.0/24 requires low latency service

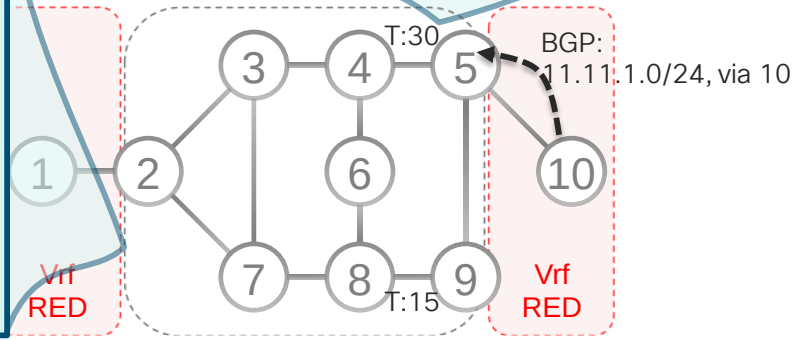
  => PE5 tags with extcommunity “color2”

- Ingress PE2 initiates on-demand policy by color template

  => PE2 finds a path to PE5 with optimized metric latency

```
extcommunity-set opaque color2
  2
end-set
!
route-policy bgp_col
 if destination in (11.11.1.0/24)  then
    set extcommunity color color2
  endif
end-policy
!
router bgp 1
 neighbor 1.1.1.2
  remote-as 1
  address-family vpnv4 unicast
  route-policy bgp_col out
```

```
router bgp 1
 neighbor 1.1.1.5
  address-family vpnv4 unicast
 vrf RED
  rd 2:2
  address-family ipv4 unicast
!
segment-routing
 traffic-eng
  on-demand color 2
   dynamic
     metric
     type latency
!
performance-measurement
 interface GigabitEthernet0/1/1/0
  delay-measurement
```



measure link delay to compute path latency
 # show isis database R1 verbose | inc Delay
Link Average Delay: 7 us
Link Min/Max Delay: 7/7 us

L3VPN Service Policy Work-flow and Steering

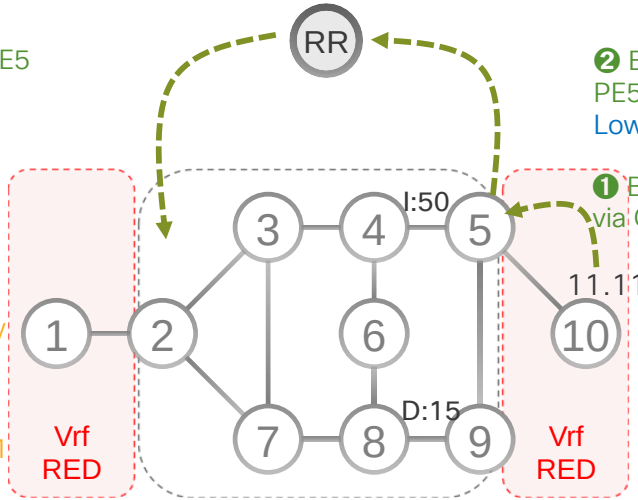
```
R1# show segment-routing traffic-eng policy color 2
Color: 2, End-point: 1.1.1.5
  Name: srte_c_2_ep_1.1.1.5
  Status:
    Admin: up  Operational: up for 00:39:14  (since Mar 31)
  Candidate-paths:
    Preference: 200 (BGP ODN) (active)
      Requested BSID: dynamic
      PCC info:
        Symbolic name: bgp_c_2_ep_1.1.1.5_discr_200
        PLSP-ID: 2
      Dynamic (valid)
        Metric Type: LATENCY,   Path Accumulated Metric: 10
          16005 [Prefix-SID, 1.1.1.5]
    Preference: 100 (BGP ODN)
      Requested BSID: dynamic
      PCC info:
        Symbolic name: bgp_c_2_ep_1.1.1.5_discr_100
        PLSP-ID: 1
      Dynamic
        Metric Type: NONE,   Path Accumulated Metric: 0
  Attributes:
    Binding SID: 24031
    Forward Class: 0
    Steering BGP disabled: no
    IPv6 caps enable: yes

# show cef vrf RED 11.11.1.0/24 | inc local-label
   via local-label 24031, 3 dependencies, recursive
```

❸ BGP: 110/24 via PE5
VPN-LABEL: 24016
Low-Delay (color 2)

❹ PE2 check prefix from PE5 with Low-latency (color 2)

❺ Initiate SRTE policy use template color 2

❻ Policy → SID-list <16005> BSID 24031



❷ BGP: 11.11.1.0/24 via PE5 VPN-LABEL: 24016
Low-Delay (color2)

❶ BGP: 11.11.1.0/24 via CE10

```
R1# show bgp vrf RED 11.11.1.0/24
Paths: (1 available, best #1)
Local
   2.2.2.2 C:2 (bsid:24031) (metric 10) from 2.2.2.2 (2.2.2.2)
     Received Label 24016
...
     Extended community: Color:2 RT:2:2
     SR policy color 2, up, registered, bsid 24031, if-handle 0x02000fe0
     Source AFI: VPNv4 Unicast, Source VRF: RED, Source Route Distinguisher: 2:2
```

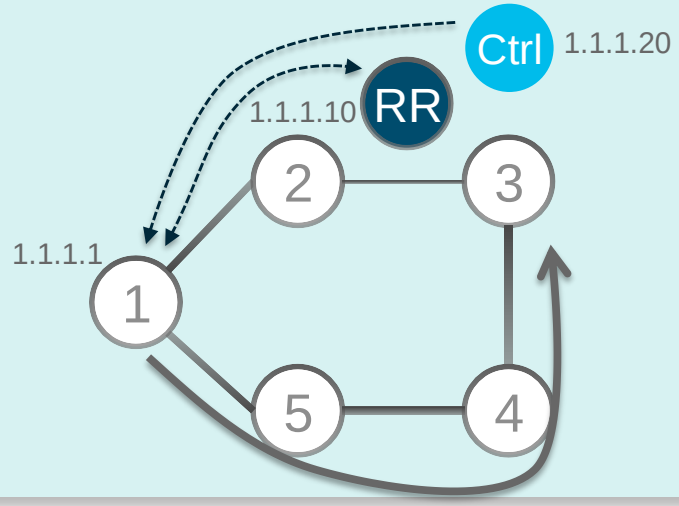
Use Case4: BGP SR-Policy Auto Initiated Policy

```
router bgp 1
 bgp router-id 1.1.1.1
 address-family ipv4 unicast
 address-family vpnv4 unicast
 address-family ipv4 sr-policy
 !
 neighbor 1.1.1.10
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
  address-family vpnv4 unicast
 !
 neighbor 1.1.1.20
  remote-as 1
  update-source Loopback0
  address-family ipv4 sr-policy
```

To Service RR

To SRTE Controller

New BGP SAFI “sr-policy” conveys SR-TE policy and NLRI information to signal a candidate paths from controller to headend



IPv4 – SR Policy NLRI
- Color 10
- End-point 1.1.1.3
- Distinguisher 2

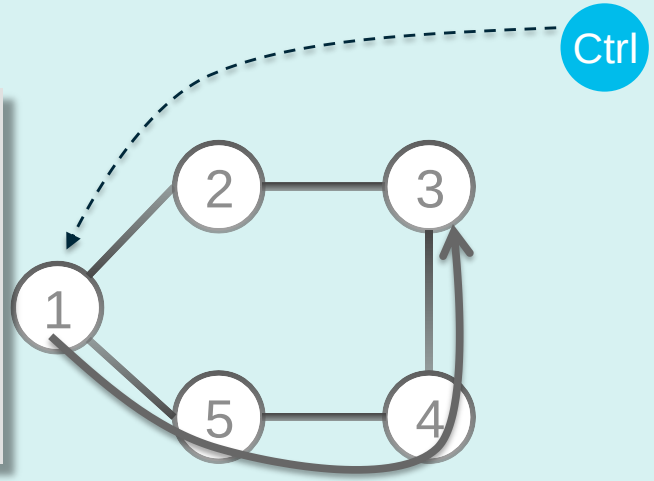
Tunnel encaps attr
- Preference 100
- Binding SID 999
- Segment List
  - Weight: 1
  - <16004, 16003>

```
curl http://1.1.1.1:8088/srpolicy-install -H "Content-Type: application/json" -X POST -d '{"source" : "1.1.1.1", "end-point" : "1.1.1.3", "binding-sid" : 999 "color" : 10, "preference" : 100, "route-distinguisher" : 0, "path-list" : [{"label-stack" : [16400 16300], "metric_te" : 30, "type" : sr-te}]}'
```

Controller Based NLRI Auto-Policy Steering

```
R1# show segment-routing traffic-eng policy
SR-TE policy database
---------------------
Name: bgp_AP_1 (Color: 10, End-point: 1.1.1.3)
   Status:
     Admin: up  Operational: up for 00:08:19 (since Jun 13 21:18:10.469)
   Candidate-paths:
     Preference 100:
       Explicit: segment-list Autolist_3_1* (active)
         Weight: 1, Metric Type: IGP
           16004 [Prefix-SID, 1.1.1.4]
           16003
   Attributes:
     Binding SID: 999
       Allocation mode: explicit
       State: Programmed
       Policy selected: yes
   Auto-policy info:
     Creator: BGP
```

IPv4 – SR Policy NLRI
Color 10
End-point 1.1.1.3
Distinguisher 2
Tunnel encaps attr
Preference 100
Binding SID 999
Segment List
Weight: 1
<16004, 16003>



Use Case5: Controller Multi-domain SRTE Policy

- SR-PCE build inter-domain topology based on the SR-TE DB
  - Learns an attached network domain topology via IGP or BGP-LS
  - Learns a remote domain topology via BGP-LS, directly or through a Route Reflector

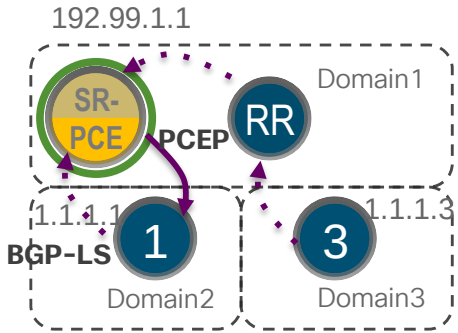
On SR-PCE:

```
pce
 address ipv4 192.99.1.1
!
router bgp 1
 address-family link-state link-state
 neighbor 1.1.1.1
  remote-as 101
  address-family link-state link-state
!
```

BGP-LS: Collection of ISIS / OSPF link state

On PCC:

```
segment-routing traffic-eng pcc
   source-address ipv4 1.1.1.1
   pce address ipv4 192.99.1.1
router isis 1    !! or ospf
 distribute link-state instance-id 101
router bgp 101
 address-family link-state link-state
 neighbor 192.99.1.1
  remote-as 1
  address-family link-state link-state
```



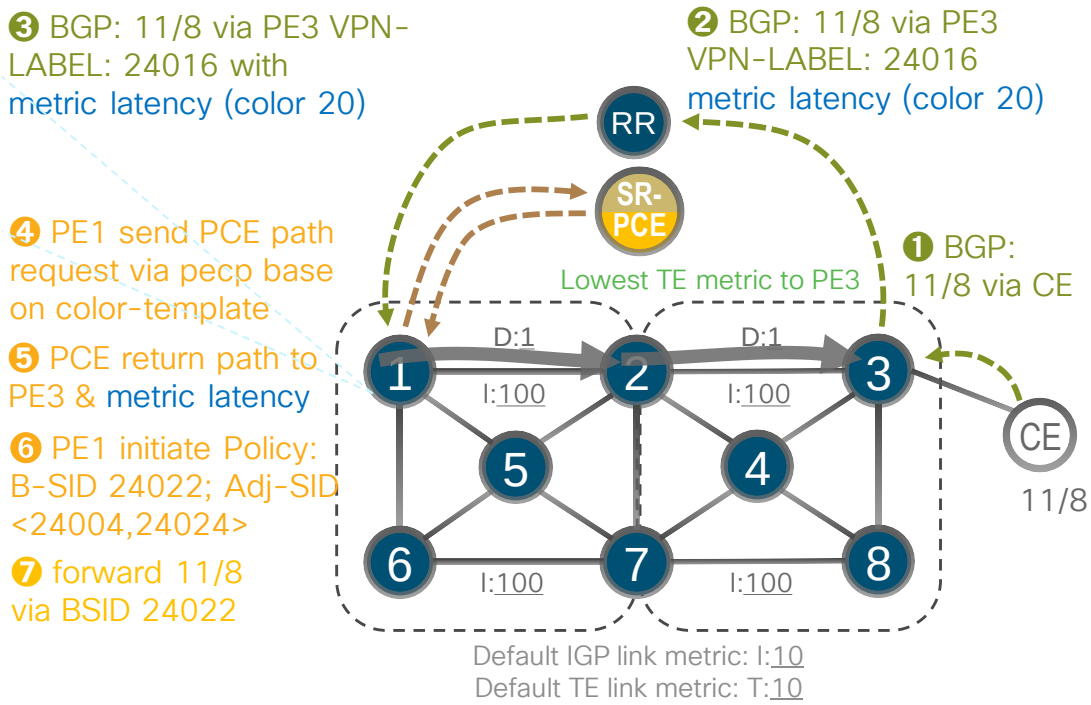

```
curl --raw -vN
"http://192.99.1.1:8080/lsp/create/simple?peer=1.1.1.1&name=rest_1_5&type=sr&
source=1.1.1.1&destination=1.1.1.3&color=10"
```

PCEP used to deploy policy

Centralized Inter-Domain SR-PCE Policy Steering

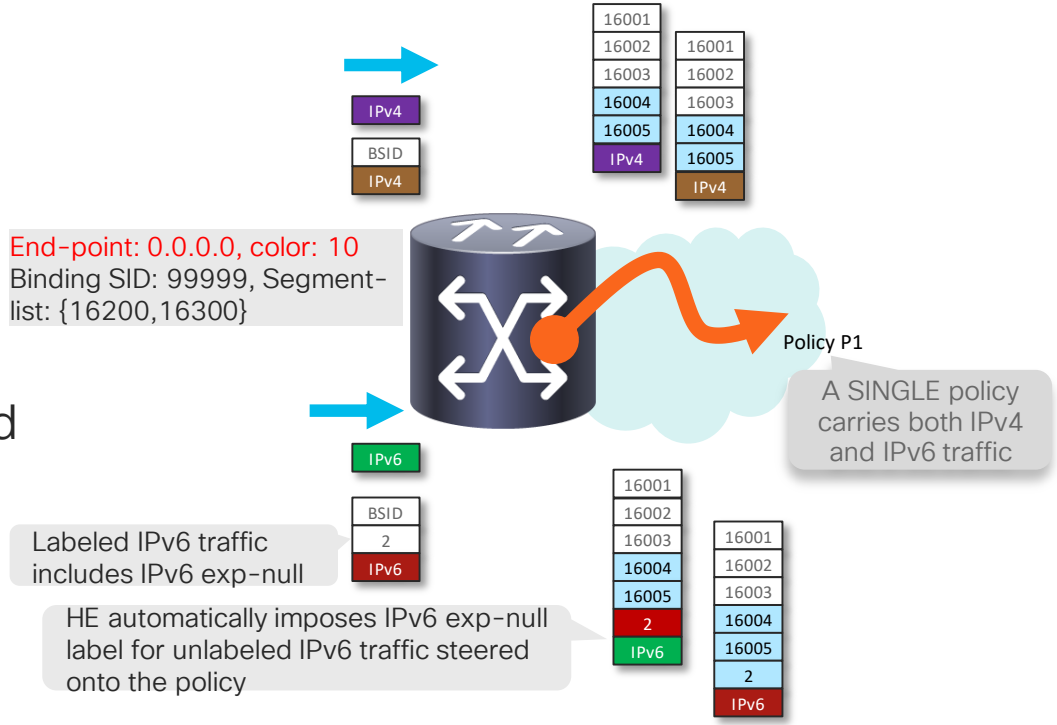
```
segment-routing
 traffic-eng
  policy pcep1
   color 20 end-point ipv4 1.1.1.3
    candidate-paths
     preference 200
      dynamic pcep
        metric type latency
```

```
R1#show segment-routing traffic-eng policy color 20
Color: 20, End-point: 1.1.1.3
  Name: srte_c_20_ep_1.1.1.3
  Status:
    Admin: up  Operational: up for 00:00:05 (Apr  9 09:58:11)
  Candidate-paths:
    Preference: 200 (configuration) (active) (reoptimizing)
      Name: pcep1
      Requested BSID: dynamic
      PCC info:
        Symbolic name: cfg_pcep1_discr_200
        PLSP-ID: 4
      Dynamic (pce 192.99.1.1) (valid)
        Metric Type: LATENCY,   Path Accumulated Metric: 20
          24004 [Adjacency-SID, 10.12.0.1 - 10.12.0.2]
          24024 [Adjacency-SID, 10.23.1.2 - 10.23.1.2]
  Attributes:
    Binding SID: 24022
```



Use Case6: Color-only Null-end Policy

- Single SR policy per color
  - Color-only and IPv6 enabled
  - Address Family Agnostic
- SR policy carries:
  - Labeled v4, Unlabeled v4
  - Labeled v6, Source expected to add v6-exp-null
  - Unlabeled v6, HE performs automatic v6-exp-null label push for unlabeled v6 traffic



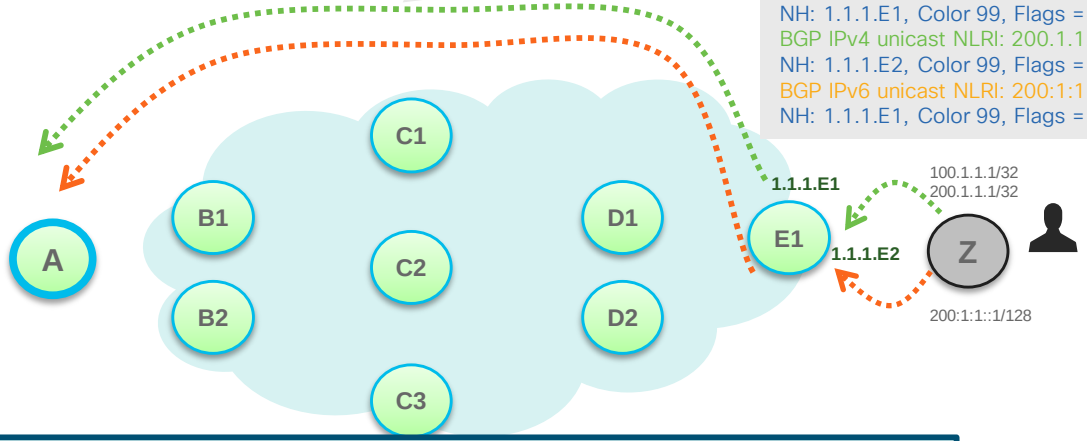
Color-only & AF-agnostic Steering

```
segment-routing
 traffic-eng
  policy c99-null-v4
   binding-sid mpls 999
   color 99 end-point ipv4 0.0.0.0
   ipv6 enable
   candidate-paths preference 111
     explicit segment-list SL-2-via-1
 segment-list name SL-2-via-1
   index 20 mpls label 16200
   index 30 mpls label 16300
```

```
R1# show segment traf policy binding-sid 99999
Color: 99, End-point: 0.0.0.0
  Name: srte_c_99_ep_0.0.0.0
  Status:
    Admin: up  Operational: up for 00:04:40 (since
Apr  9 18:59:37.047)
  Candidate-paths:
    Preference: 111 (configuration) (active)
      Name: c99-null-v4
      Requested BSID: 99999
      Explicit: segment-list SL-2-via-1 (valid)
        Weight: 1, Metric Type: TE
          16200 [Prefix-SID, 2.2.2.2]
          16300
  Attributes:
    Binding SID: 999
    Forward Class: 0
    Steering BGP disabled: no
    IPv6 caps enable: yes
```

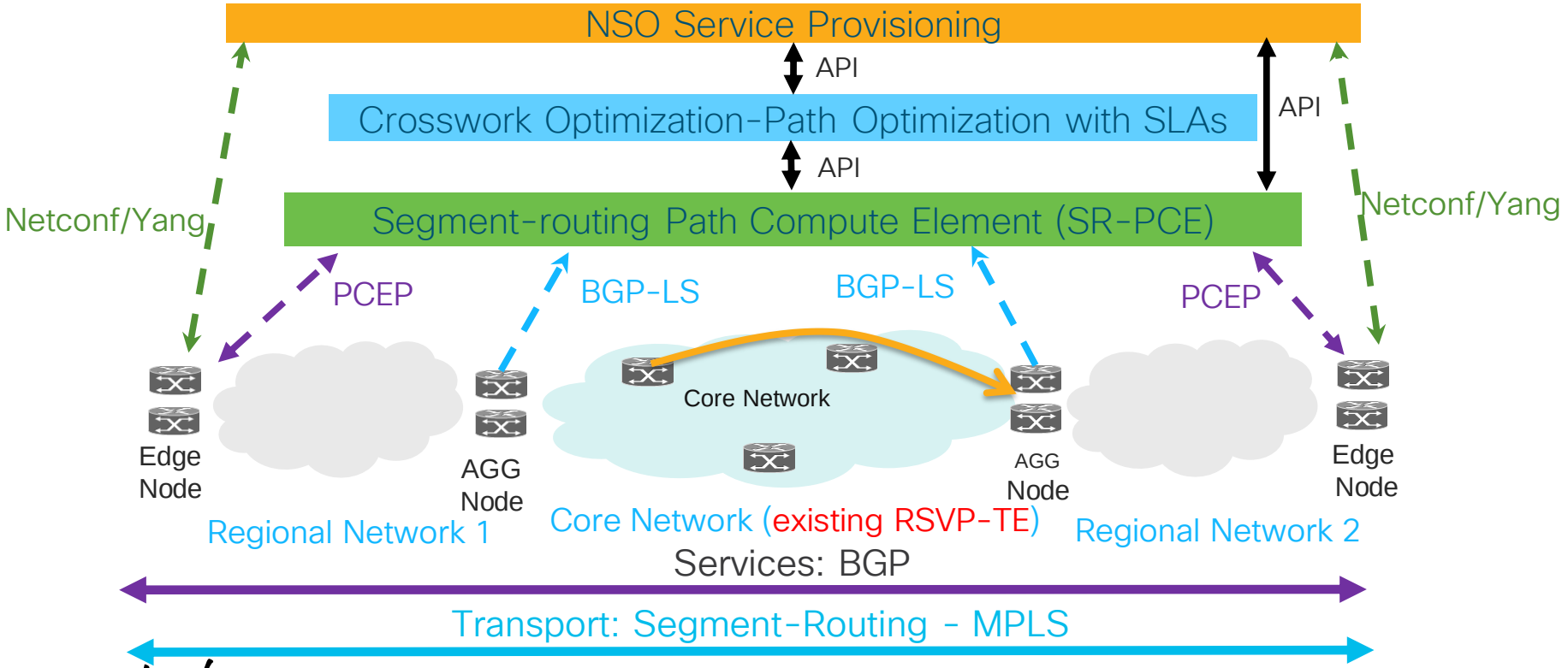
When "co-flag" set to 01, BGP prefix steering first match SR-TE policy color and endpoint. Otherwise, prefix color-only match going steer over a null-end Policy (C, null).

```
extcommunity-set opaque color99
  99 co-flag 01
end-set
!
BGP IPv4 unicast NLRI: 100.1.1.1/32
NH: 1.1.1.E1, Color 99, Flags = 01
BGP IPv4 unicast NLRI: 200.1.1.0/24
NH: 1.1.1.E2, Color 99, Flags = 01
BGP IPv6 unicast NLRI: 200:1:1::1/128
NH: 1.1.1.E1, Color 99, Flags = 01
```




```
R1# show cef 200.1.1.0/24
10.16.3.0/24, version 218468, internal 0x5000001 0x0 (ptr 0x784a1df8) [1], 0x0 (0x0), 0x0 (0x0)
 Updated Apr 10 08:18:57.113
 Prefix Len 24, traffic index 0, precedence n/a, priority 4
   via local-label 99999, 3 dependencies, recursive [flags 0x6000]
    path-idx 0 NHID 0x0 [0x78fbe2b8 0x0]
    recursion-via-label
    next hop via 999/1/21
```

Evolved Network to Unified SR Fabric



RSVP-TE Bandwidth Accounting

SR Accounting & RSVP resvBW Refresh

You make networking **possible**

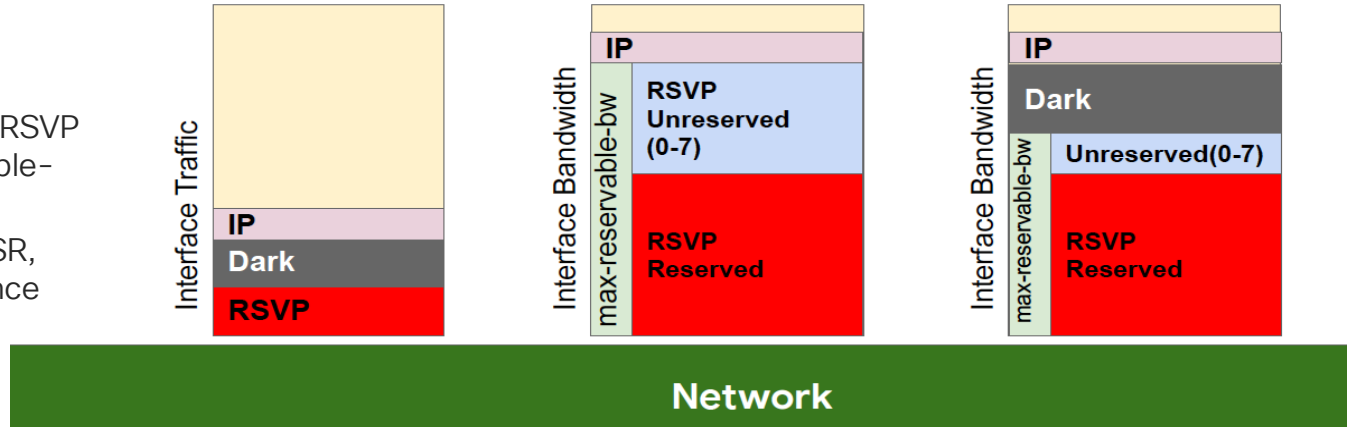
Cisco live!

RSVP-TE AdmissionCtrl Aware of SR Accounting

- RSVP-TE non-zero Bandwidth
  - Web/OTT customers: full-mesh rsvp-te + auto-BW
  - Voice/Video Service Providers: full-mesh rsvp-te + auto-BW (or signalled-BW)
- SR-TE and RSVP-TE Coexistence
  - One of Web/OTT customer introduced the network migration base on RSVP-TE Bandwidth Accounting at MPLS World Congress

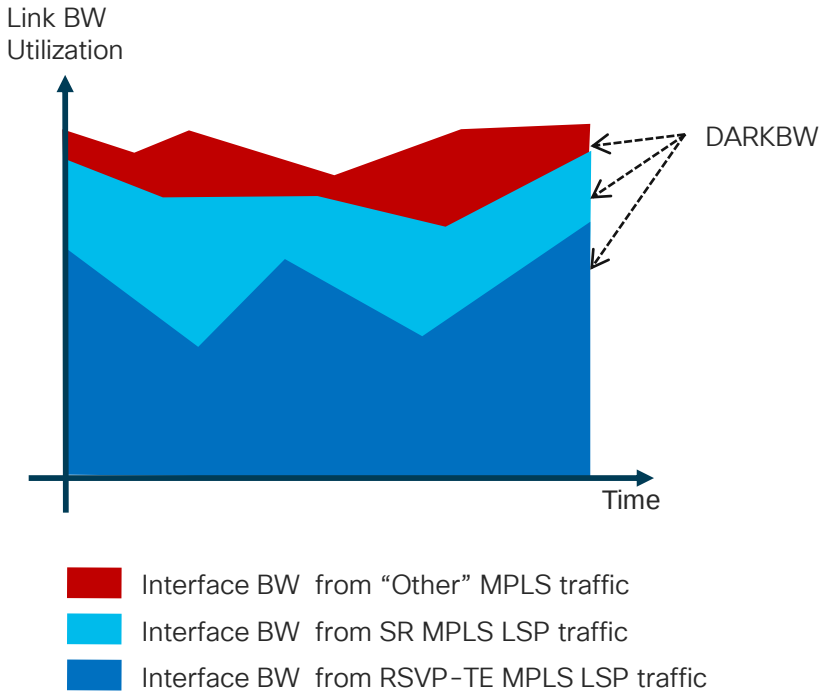
Dark Bandwidth

- SR traffic can alter RSVP Maximum-Reservable-Bandwidth
- Recommended in SR, RSVP-TE coexistence scenarios

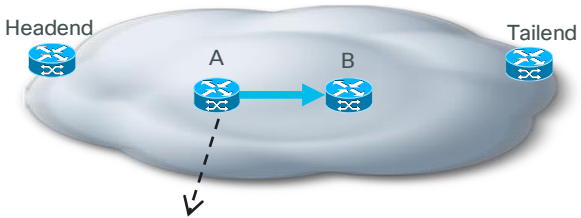


SR Accounting vs RSVP-TE “Dark Bandwidth”

- Subset of the traffic that is not explicitly admission controlled by RSVP-TE
- Those traffic bandwidth isn’t considered during RSVP-TE path computation as well as admission control
- It is assumed that SR traffic has higher priority than any other traffic transport with distributed RSVP-TE LSPs
- SR-accounting on RSVP links to trigger ResvBW refresh and gradually reduce the RSVP-TE bandwidth pool

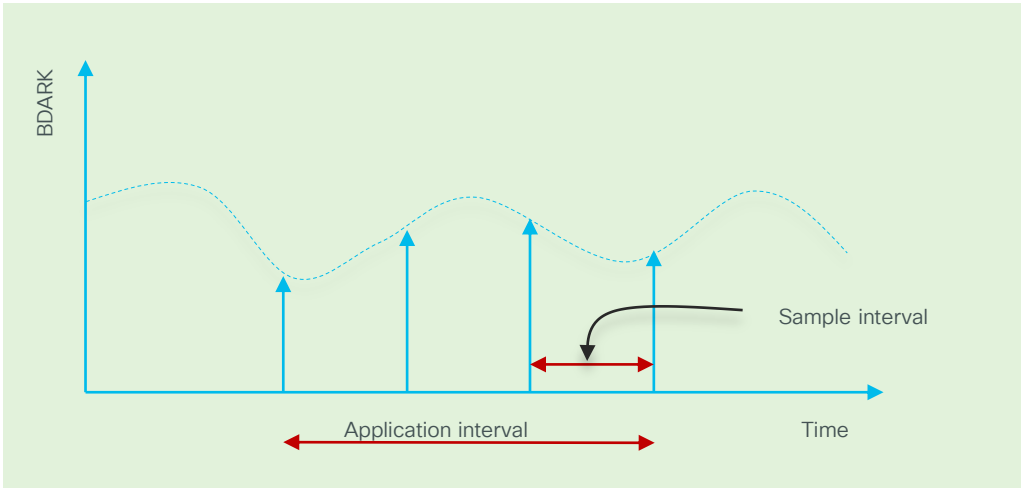


Dark Bandwidth Measurement



SR Traffic Link Utilization Collection

BDARK & “effective” resvBW Calculation

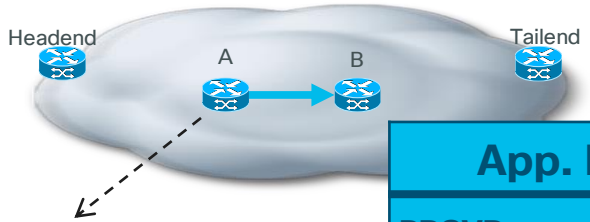


- BDARK @ application interval
  - An average of BDARK [t] samples is calculated
  - Configured Maximum Reservable-Bandwidth (BMRc)
  - “effective” resvBW flooded if change exceed threshold
- BDARK samples & “effective” resvBW (BMRe) calculated:
  - BDARK [t] = SR_IPv4 + SR_IPv6 + SR_MPLS
  - resvBW[t] = BMRc - BDARK [t]

RSVP-TE resvBW and Flooding Threshold

- RSVP-TE resvBW Counters:
  - BMRc = Configured Maximum Reservable Link Bandwidth
  - BMRe = Effective BMR
  - BMRef= last flooded BMRe
- "Effective" BMR (BMRe) collection @ application time [t]:
  - BMRe[t] = BMRc – BDARK[t] * K (K is a configurable adjustment-factor)
- Implementation provides BMRe flooding threshold as:
  - BMRe changes as a percentage of "last flooded BMRe" (BMRef)
  - Calculated as: $\frac{|\mathrm{BMRe} - \mathrm{BMRef}|}{\mathrm{BMRef}}$

Solution Overview – Flooding Action By Timeframe



Interface TenGigE0/0/1/0	
Physical BW	= 1,000
BMRc	= 900
Adjustment factor	= 100%
Flooding thr. Up	= 10%
Flooding thr. Down	= 10%

[1] $BDARK = Avg\ SR\ Samples$
[2] $\boldsymbol{BMRe = BMRc - BDARK}$
[3] $\Delta BMRe = \frac{|BMRe - BMRef|}{BMRef}$
[4] Assumes BMRef of 900 from T_0
[5] (|800 – 900|) / 900
[6] (|860 – 800|) / 800
[7] (|350 – 800|) / 800

App. Interval	T_1	T_2	T_3	T_4
BRSVP – RSVP utilized link BW	→ 200	→ 200	→ 200	→ 200
BDARK – SR utilized link BW	→ 0	↗ 100	↘ 40	↗ 550
BUTIL – Total utilized link BW	→ 200	↗ 300	↘ 240	↗ 750
BResvd – Total reserved link BW	300	300	300	300
BMRe – Effec. max reservable link BW	900	800	860	350
BMRef – Last flooded BMRe	900[4]	→ 900	→ 800	→ 800
ΔBMRe[3] – Change of BMRe	0	11.11%[5]	7.5% [6]	56.2% [7]
Action	No Flooding BMRef= 900	Flooding BMRef= 800	No Flooding BMRef= 800	Flooding BMRef= 350

Migration Solution Building Blocks

SR MPLS

- Source-routing
- Architecture brings right balance of distributed intelligence and centralized optimization
- Lightweight extensions to IP control-plane
- Leverages MPLS data and control plane

Auto-Steering

- Steer SR traffic away from RSVP-TE LSP by IGP sr-prefer, BGP color / endpoint match, and binding SID install

Dark Bandwidth

- Co-existence of non-zero bandwidth RSVP tunnels and SR in the same network domain
- Measure link utilization due to SR MPLS
- Adjust Maximum Link ResvBW (BMR) according to SR accounting

SR Traffic Engineering

- Simple, Automated and Scalable
- No core state – state in the packet header
- No tunnel interface – “SR Policy”
- BGP ODN candidate path
- Explicit local configured candidate paths
- Automate traffic steering
- Color-only match and AF- agnostic automated steering

RSVP-TE Bandwidth Accounting Provision

- BMRe[t] = BMRc – BDARK[t] * K, where:
  - K is a customer configurable adjustment-factor (0% to 200%)
  - K decide how aggressively react to RSVP-TE effective **resvBW** adjustment
  - resvBW refresh like "make before break", gradually reduce the RSVP-TE bandwidth pool

```
accounting interfaces
  segment-routing mpls
   ipv4
   ipv6
!
mpls traffic-eng
 bandwidth-accounting
  application
   enforced
   interval 90
  sampling-interval 10
  adjustment-factor 150
  flooding threshold up 10 down 10
```

K = 150%; by default K = 100%

SR Accounting Statistics

R1# show mpls traffic-eng link-management interfaces Gi0/1/1/0 detail

```
System Information::
   Links Count        : 16 (Maximum Links Supported 800)

Link ID:: GigabitEthernet0/1/1/0 (10.12.110.1)
  Local Intf ID: 22
  Link Status:

    Link Label Type          : PSC
    Physical BW              : 1000000 kbits/sec
    BCID                     : RDM
    Max Reservable BW        : 621981 kbits/sec (reserved: 94% in, 94% out)
    Flooded Max Reservable BW: 621981 kbits/sec
    BC0 (Res. Global BW)     : 529309 kbits/sec (reserved: 94% in, 94% out)
    BC1 (Res. Sub BW)        : 0 kbits/sec (reserved: 100% in, 100% out)
    MPLS TE Link State       : MPLS TE on, RSVP on, admin-up
    IGP Neighbor Count       : 1
    Max Res BW (RDM)         : 900000 kbits/sec
    BC0 (RDM)                : 900000 kbits/sec
    BC1 (RDM)                : 0 kbits/sec
    Max Res BW (MAM)         : 0 kbits/sec
    BC0 (MAM)                : 0 kbits/sec
    BC1 (MAM)                : 0 kbits/sec

    Bandwidth Accounting: Segment-Routing
    Bandwidth Accounting Enforced: Yes

(output continues on the right side)
* ResvBW = BMRc-BMRe = 621981
```

```
Bandwidth Utilization Details:
  Sampling Interval          : 10 sec
  Application Interval       : 90 sec
  Adjustment Factor          : 150%
  Max Reservable BW Up Threshold  : 10
  Max Reservable BW Down Threshold: 10
   Last Application at:  19:41:14 Sun 30 Apr 2017 (51 seconds ago)
    Segment-Routing BW Utilization     : 185346 kbits/sec
    Adjusted BW Utilization            : 278019 kbits/sec
    Enforced BW Utilization            : 278019 kbits/sec
  Next Application at:  19:42:43 Sun 30 Apr 2017 (in 38 seconds)
  Last Collection at :   19:40:42 Sun 30 Apr 2017 (13 seconds ago)
  Next Collection at :  19:40:51 Sun 30 Apr 2017 (in 6 seconds)

      Bandwidth Samples (Kbps):
  Timestamp                              Segment-Routing
       19:40:12 Sun 30 Apr 2017                  187961
       19:40:22 Sun 30 Apr 2017                  180130
       19:40:32 Sun 30 Apr 2017                  187949
       19:40:42 Sun 30 Apr 2017                  180124
Attributes        : 0x0
Ext Admin Group     :
   Length : 256 bits
   Value  : 0x::
Attribute Names     :
Flooding Status: (1 area)
  IGP Area[1]: IS-IS 0 level 2, flooded
    Nbr: ID 0000.0000.0002.00, IP 10.12.110.2 (Up)
    Admin weight: not set (TE), 10 (IGP)
Lockout Status: Never
```

SR Accounting & RSVP resvBW Refresh Monitoring

R1# show mpls traffic-eng link-management summary

```
System Information::
    Links Count         : 16 (Maximum Links Supported 800)
    Flooding System     : enabled
    IGP Areas Count     : 1

IGP Areas
----------

IGP Area[1]:: IS-IS 0 level 2
    Flooding Protocol   : IS-IS
    Flooding Status     : flooded
    Periodic Flooding   : enabled (every 180 seconds)
    Flooded Links       : 7
    IGP System ID       : 0000.0000.0001
    MPLS TE Router ID   : 1.1.1.1
    IGP Neighbors       : 7

Bandwidth accounting:
    Sampling interval: 10 seconds, Next in 9 seconds
    Application interval: 90 seconds, Next in 1 seconds
```

SR accounting intervals when bandwidth-accounting enabled

R1# show mpls traffic-eng link-management advertisements

```
Flooding Status             : Ready
Last Flooding               : 470 seconds ago
Last Flooding Trigger       : Link BW changed
Next Periodic Flooding In   : 136 seconds
Diff-Serv TE Mode           : Not enabled
Configured Areas            : 1

IGP Area[1]:: IS-IS 0 level 2
    Flooding Protocol   : IS-IS
    IGP System ID       : 0000.0000.0001
    MPLS TE Router ID   : 1.1.1.1
    Flooded Links       : 5

    Link ID:: 0 (GigabitEthernet0/1/1/0)
        Link IP Address     : 10.12.110.1
        O/G Intf ID         : 22
        Neighbor            : ID 0000.0000.0002.00, IP 10.12.110.2
        TE Metric           : 10
        IGP Metric          : 10
        Physical BW         : 1000000 kbits/sec
        BCID                : RDM
        Max Reservable BW   : 621981 kbits/sec
        Res Global BW       : 899999 kbits/sec
        Res Sub BW          : 0 kbits/sec
```

Effective BMRe triggers new resvBW flooding

Tunnel Pre-emption When Not Enough ResvBW

```
R1# show mpls traffic-eng link-management admission-control
    System Information::
    Tunnels Count     : 3
    Tunnels Selected  : 3
Bandwidth descriptor legend:
  B0 = bw from pool 0, B1 = bw from pool 1, R = bw locked, H = bw held
TUNNEL ID            UP IF     DOWN IF   PRI STATE        BW (kbits/sec)
----------------------- ---------- ---------- --- ------------ ---------------
1.1.1.1 10_6         -         Gi0/1/1/0   3/3 Resv Admitted 0              B0
1.1.1.1 10_7         -         Te0/3/0/1   3/3 Resv Admitted 800000         RB0

RP/0/RSP0/CPU0:Apr 30 19:43:51.209 : te_control[1178]: %ROUTING-MPLS_TE-5-LSP_REOPT : tunnel-te10 (signalled-name:
SRTE_RTR1_t10, old LSP Id: 6, new LSP Id: 7) has been reoptimized; reason: Soft Preemption.
```

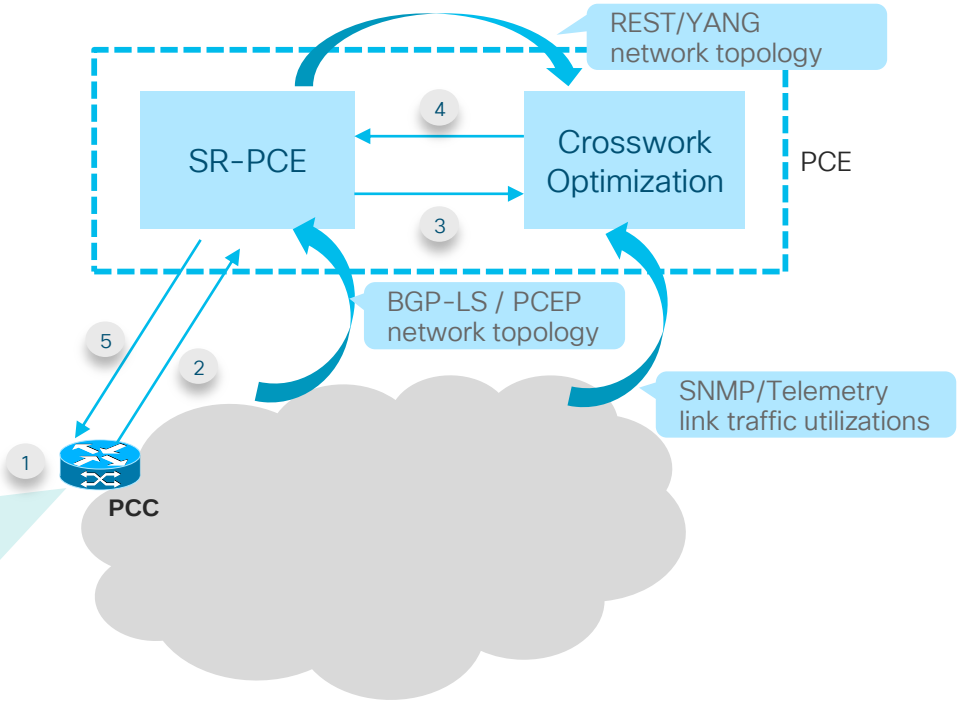
```
R1# show mpls traffic-eng link-management admission-control
            System Information::
    Tunnels Count     : 3
    Tunnels Selected  : 3
Bandwidth descriptor legend:
  B0 = bw from pool 0, B1 = bw from pool 1, R = bw locked, H = bw held
TUNNEL ID            UP IF     DOWN IF   PRI STATE        BW (kbits/sec)
----------------------- ---------- ---------- --- ------------ ---------------
1.1.1.1 10_3         -         Gi0/1/1/0   1/1 Resv Admitted 800000      RB0

RP/0/RSP0/CPU0:Apr 30 19:29:08.551 : te_control[1178]: %ROUTING-MPLS_TE-5-LSP_UPDOWN : tunnel-te10 (signalled-name:
SRTE_RTR1_t10, LSP Id: 3) state changed to down
 RP/0/RSP0/CPU0:Apr 30 19:29:25.131 : te_control[1178]: %ROUTING-MPLS_TE-5-LSP_UPDOWN : tunnel-te10 (signalled-name:
SRTE_RTR1_t10, LSP Id: 4) state changed to up
```

Use Case7: SRTE Policy with Bandwidth OnDemand

1. Operator configures SR-TE policy with bandwidth constraint
2. PCC sends PCReq to SR-PCE controller
3. SR-PCE requests BW-path from Crosswork Optimization
4. Crosswork Optimization returns BW-path (or no-path) to SR-PCE
5. SR-PCE sends BW-path (or no-path) to PCC

```
segment-routing
  traffic-eng
    policy BW123
      bandwidth 888
      color 10 end-point ipv4 3.3.3.3
      candidate-paths
      preference 100
       dynamic pcep
```

CLI inter-domain policy path calculated by PCE



BWoD Policy Configure and Show Output

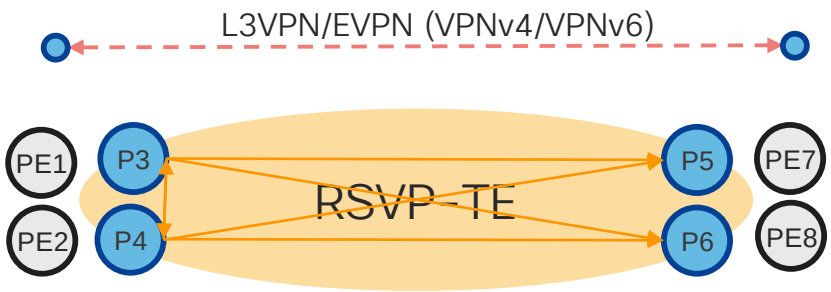
```
segment-routing
 traffic-eng
  pcc
   pce address ipv4 192.99.1.1
  !
  policy BW123
   bandwidth 888
   color 10 end-point ipv4 3.3.3.3
   autoroute
   !
   candidate-paths
    preference 100
     dynamic pcep
      metric
       type te
  !
```

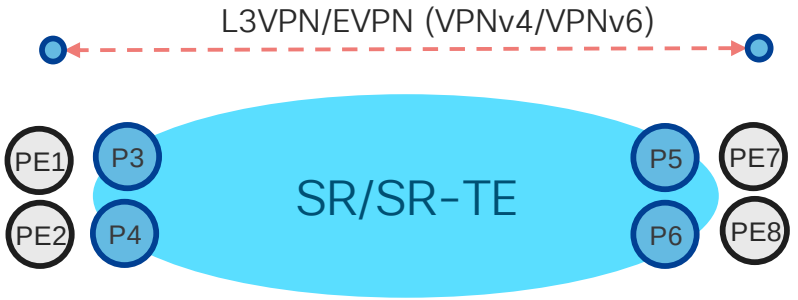
```
pcc# show segment-routing traffic-eng policy
SR-TE policy database
---------------------
Name: BW123 (Color: 10, End-point: 3.3.3.3)  ID: 1
  Status:
    Admin: up  Operational: up for 2d19h (since Jan 27
18:10:16.545)  Candidate-paths:
    Preference 100:
      Dynamic (pce 192.99.1.1) (active), valid yes
        Weight: 0, Metric Type: TE
        IGP area: 0
          16300 [Prefix-SID, 3.3.3.3]
  Attributes:
    Binding SID: 25056
      Allocation mode: dynamic
      State: Programmed
      Policy selected: yes
    Forward Class: 0
    Bandwidth Requested: 888 kbps
    Bandwidth Current: 888 kbps
```

Key Takeaways

You make networking **possible**

Migrate to Scalable and Programmable Network





RSVP-TE

- RSVP-TE Mesh for FRR
- Config+State for N(N-1) Tunnels
- Traffic steering by IGP autoroute

Segment Routing

- Automatic 100% FRR with TI-LFA
- Explicit/Dynamic Stateless Policy
- Traffic steering by policy B-SID

SR Unified Fabric Attributes

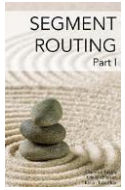
Stay Up-To-Date

http://www.segment-routing.net/

https://www.linkedin.com/groups/8266623

https://twitter.com/SegmentRouting

https://www.facebook.com/SegmentRouting/



Segment Routing, Part I - Textbook

Complete your online session evaluation

- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live water bottle.
- All surveys can be taken in the Cisco Live Mobile App or by logging in to the Session Catalog on ciscolive.cisco.com/us.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.cisco.com.

Continue your education

Demos in the Cisco campus

Walk-in labs

Meet the engineer 1:1 meetings

Related sessions

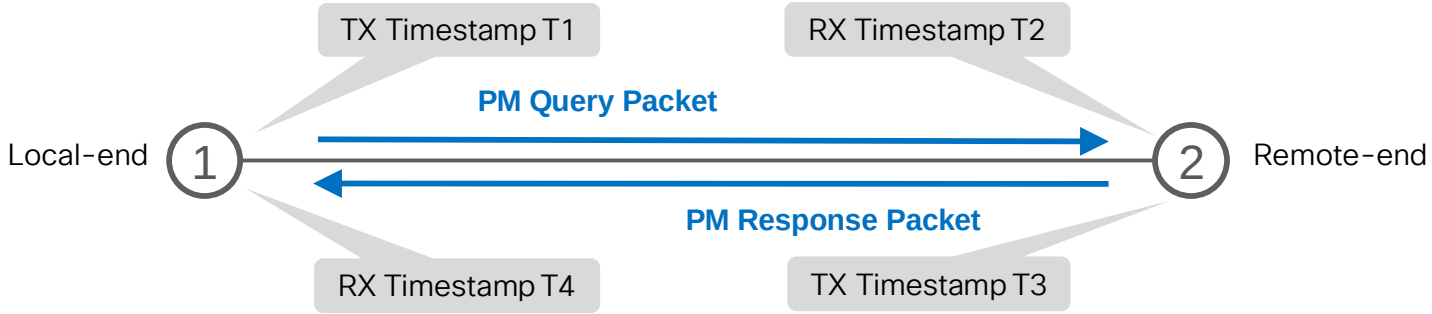
Thank you

You make **possible**

Acronym

- SR-TE — Segment Routing Traffic Engineering
- RSVP-TE — Resource Reservation Protocol Traffic Engineering
- TI-LFA — Path Protection with Topology Independent Loop Free Alternative
- RESTAPI — An application program interface (API) that uses HTTP requests to GET, PUT, POST and DELETE data.
- node-SID — A special type of prefix SID with node flag that identifies a specific node.
- adjacency SID — A dynamically allocated segment SID on an interface
- BSID — Binding SID used to identify an active candidate path of a policy
- PCE — Path Computation Element
- PCC — Path Computation Client
- NLRI — Network Layer Reachability Information
- BMR — Maximum Reservable Link Bandwidth
- BMRc — Configured Maximum Reservable Link Bandwidth
- BMRe — Effective BMR
- BMRef — last flooded BMRe

Link Delay Measurement

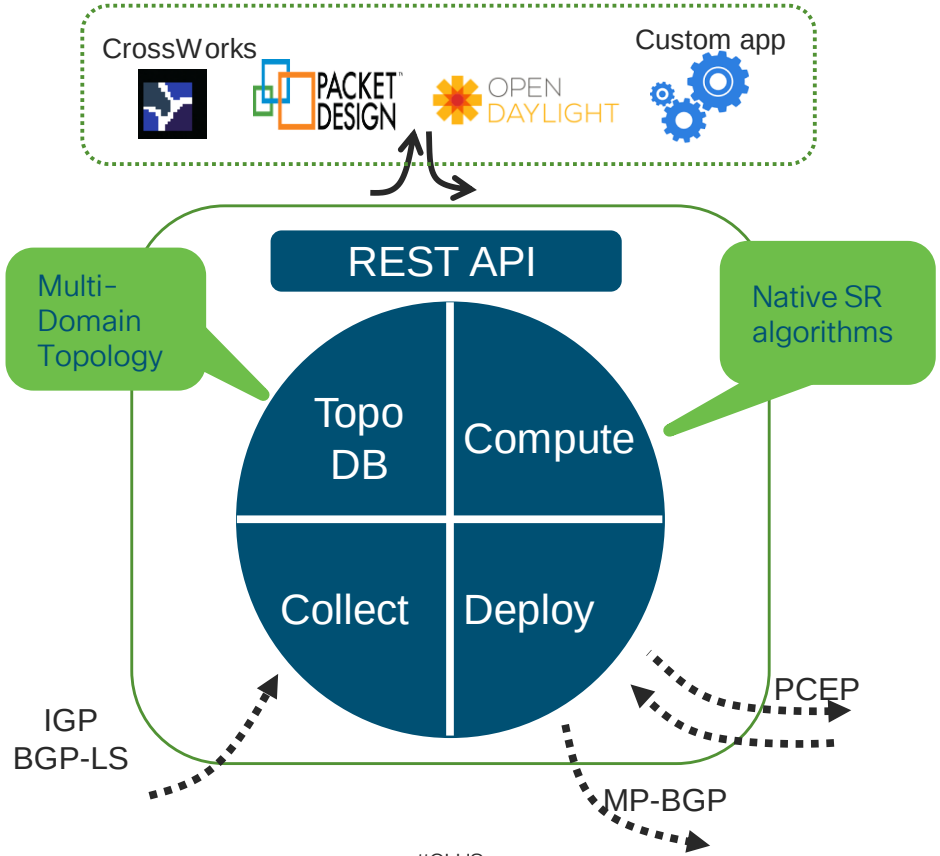


- Two-Way Delay = (T2 – T1) + (T4 – T3)
  = (T4 – T1) – (T3 – T2)
- One-way Delay = (Two-Way Delay)/2

- Measure link delay to compute path latency
- Advertised via IGP TE Metric Extensions
- SRTE optimize path on min-delay

```
# show isis database R1 verbose | i Delay
    Link Average Delay: 7 us
    Link Min/Max Delay: 7/7 us
```

SR-PCE Controller Architecture



Egress Peer Engineering (EPE)

Applicability Examples

Internet Peering

Engineered
Exit Points

Solution

SDN controller based approach to instruct an **ingress PE** to use a **specific egress PE** and a **specific external interface/neighbor** to reach a **particular destination.**

Benefits

Intent-based

SLA-aware BGP service

Optimal decisions based on cost, latency, loss

Optimal use of resources