You make **possible**

Troubleshooting Segment Routing

Luc De Ghein Technical Leader

BRKRST-3009

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

Agenda

- Introduction
- Segment Routing (SR) Recap
- Troubleshooting Control Plane
- Troubleshooting Data Plane
- SR Traffic Engineering (SR-TE)
- ODN
- Ti-LFA
- Micro-loop avoidance
- SRv6
- SR PCE
- SR OAM
- Key Takeaways

Cisco Webex Teams

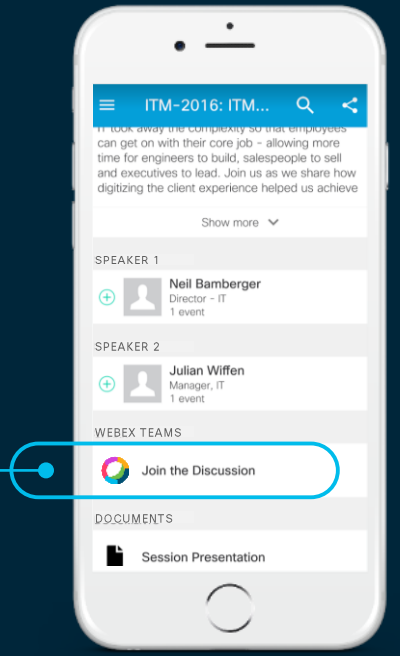
Questions?

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#BRKXXX-xxxx

Before We Get Started

- Fair basic knowledge on SR is required
- MPLS and IPv6 in data plane
  - This presentation mostly covers MPLS; there is some SRv6
- All is IOS-XR
  - Latest and greatest
  - Similar in IOS(-XE)

Stay Up-To-Date

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

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

https://twitter.com/SegmentRouting

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

SR Recap

You make networking **possible**

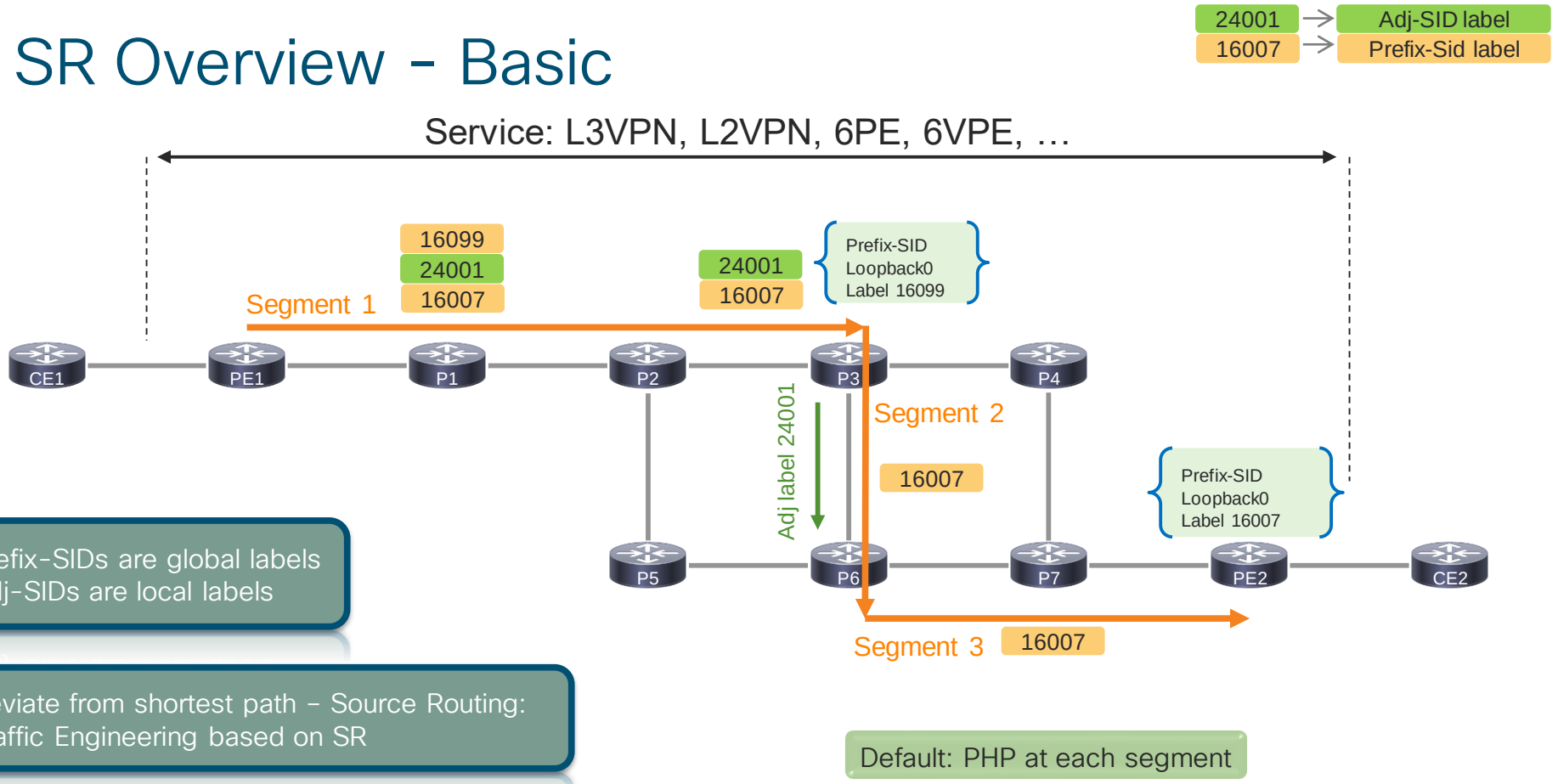
Cisco live!

Introduction

- A segment is an instruction
  - with MPLS forwarding: segment = label
- Forwarding is done by MPLS or IPv6
  - This session only covers MPLS
- Link-state protocol is needed to advertise
  - Segments (Prefix-SID, Adjacency-SID)
  - MPLS Label
- Removing the signaling and state (no LDP/ no RSVP-TE)
- Controller/SDN can be used if/when needed

*SID = Segment Identifier

SR Overview – Basic



Troubleshooting – Control Plane

You make networking **possible**

Cisco live!

Bringing up Segment Routing

```
router isis 1
 net 49.0001.0000.0000.0001.00
 address-family ipv4 unicast
  metric-style wide
  segment-routing mpls
```

metric-style wide must be enabled

Enable SR on all IS-IS IPv4 interfaces

Enable SR for all areas

Enable SR forwarding on interfaces

Disable SR for area 1

Disable SR forwarding on area 1 interfaces

```
router ospf 1
 segment-routing mpls
 segment-routing forwarding mpls ! On by default
 area 0
  interface Loopback0
  !
  interface GigabitEthernet0/0/0/0
 !
 area 1
  segment-routing forwarding disable
  segment-routing disable
```

```
RP/0/0/CPU0:PE1# show mpls interfaces

Interface                  LDP      Tunnel   Static   Enabled
-------------------------- -------- -------- -------- --------
GigabitEthernet0/0/0/0     No       No       No       Yes
GigabitEthernet0/0/0/1     No       No       No       Yes
GigabitEthernet0/0/0/2     No       No       No       Yes
GigabitEthernet0/0/0/3     No       No       No       Yes
```

Verify SRGB Block

```
RP/0/0/CPU0:PE1# show mpls label table detail

Table Label   Owner                           State  Rewrite
----- ------- ------------------------------- ------ -------
0     0       LSD(A)                          InUse  Yes
0     1       LSD(A)                          InUse  Yes
0     2       LSD(A)                          InUse  Yes
0     13      LSD(A)                          InUse  Yes
0     1000    Application-Controller(A):XTC   InUse  Yes
  (TE Binding, vers:0, identifier=3, type=2)
0     2000    Application-Controller(A):XTC   InUse  Yes
  (TE Binding, vers:0, identifier=4, type=2)
0     15000   LSD(A)                          InUse  No
  (Lbl-blk SRLB, vers:0, (start_label=15000, size=1000, app_notify=0)

0     16000   ISIS(A):1                       InUse  No
  (Lbl-blk SRGB, vers:0, (start_label=16000, size=8000)
0     24000   ISIS(A):1                       InUse  Yes
  (SR Adj Segment IPv4, vers:0, index=1, type=0, intf=Gi0/0/0/1, nh=10.1.6.6)
0     24001   ISIS(A):1                       InUse  Yes
  (SR Adj Segment IPv4, vers:0, index=3, type=0, intf=Gi0/0/0/1, nh=10.1.6.6)
0     24002   ISIS(A):1                       InUse  Yes
  (SR Adj Segment IPv4, vers:0, index=1, type=0, intf=Gi0/0/0/3, nh=10.1.5.6)
```

default SR Global Block: 16,000 – 23,999 (size 8,000)

dynamic label range: 24,000 – 1,048,575

labels for SR policies

first Adj-SID label (from dynamic range)

```
RP/0/0/CPU0:PE1# show mpls label range
Range for dynamic labels: Min/Max: 24000/1048575
```

dynamic range

IGP – LS Database

```
RP/0/0/CPU0:PE1# show isis database level 2 verbose P1.00

IS-IS 1 (Level-2) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime  ATT/P/OL
P1.00-00              0x00000017   0xa223        674           0/0/0
  Area Address: 49.0001
  NLPID:        0xcc
  Hostname:     P1
  IP Address:   10.100.1.5
  Router Cap:   10.100.1.5, D:0, S:0
    Segment Routing: I:1 V:0, SRGB Base: 16000 Range: 8000
  Metric: 10         IS-Extended P1.00
    Interface IP Address: 10.1.2.5
    Neighbor IP Address: 10.1.2.1
    ADJ-SID: F:0 B:1 V:1 L:1 S:0 weight:0 Adjacency-sid:24000
    ADJ-SID: F:0 B:0 V:1 L:1 S:0 weight:0 Adjacency-sid:24001

 Metric: 0          IP-Extended 10.100.1.5/32
   Prefix-SID Index: 5, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0
```

Regular LSP header

SR Capabilities

I:1	IPv4 capable
V:0	no IPv6 capable
SRGB block	

Adj-SID

F:0	address family IPv4
B:1	eligible for Backup
V:1	adj SID has a Value
L:1	Local significance
S:0	Set of adjacencies
Weight: 0	amount of load balancing (NA yet)

Prefix-SID

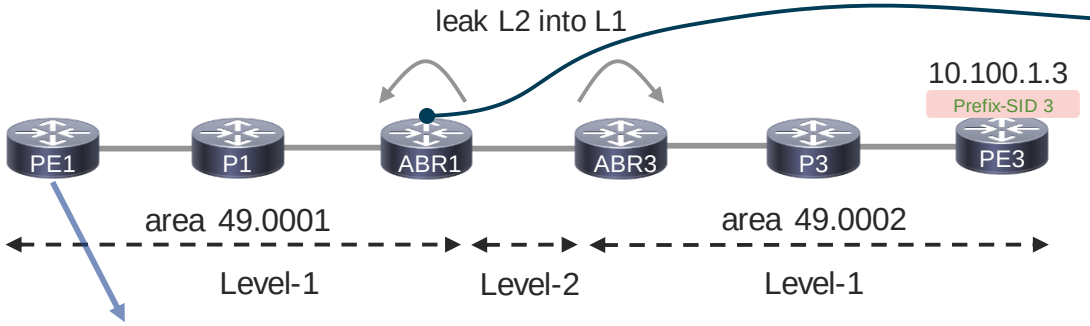
R:0	not Re-advertised
N:1	node SID
P:0	PHP on
E:0	explicit-NULL off
V:0	index
L:0	global significance

```
router isis 1
 interface Loopback0
  passive
  address-family ipv4 unicast
   prefix-sid index 5
```

```
router isis 1
 interface Loopback0
  passive
  address-family ipv4 unicast
   prefix-sid absolute 16005
```

*Prefix SID always advertised as relative index

Example of P and R Flags: Multi-Level ISIS




```
route-policy pass
  pass
end-policy
!
router isis 1
 net 49.0001.0000.0000.0009.00
 address-family ipv4 unicast
  metric-style wide
  mpls traffic-eng level-1-2
  router-id Loopback0
  propagate level 2 into level 1 route-policy pass
  segment-routing mpls
```

```
RP/0/0/CPU0:PE1# show isis database ABR1.00-00 verbose
..
 Metric: 40         IP-Extended-Interarea 10.100.1.3/32
    Prefix-SID Index: 3, Algorithm:0, R:1 N:1 P:1 E:0 V:0 L:0
    Prefix Attribute Flags: X:0 R:1 N:1
    Source Router ID: 10.100.1.3
…
```

Prefix SID

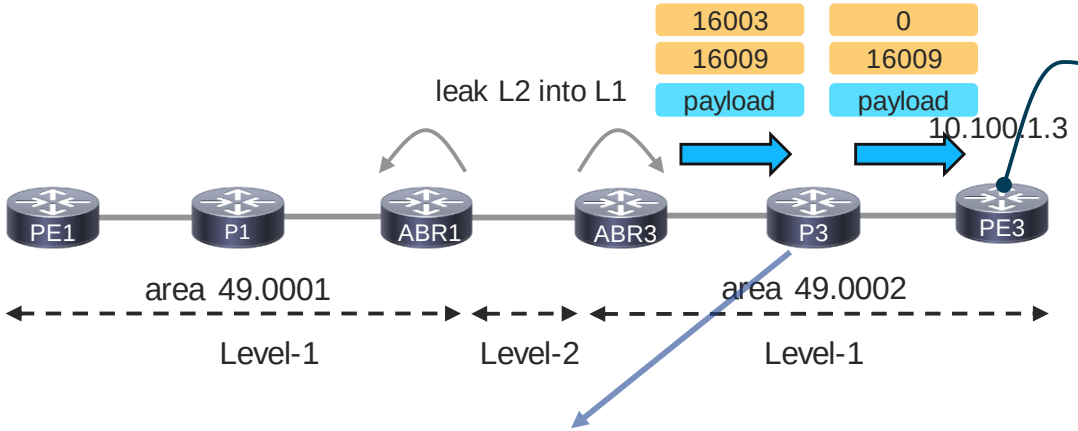
R:1 Re-advertised
P:1 PHP off

The R-Flag MUST be set for prefixes that are not local to the router and either: advertised because of propagation (Level-1 into Level-2); advertised because of leaking (Level-2 into Level-1); advertised because of redistribution (e.g.: from another protocol).

```
RP/0/0/CPU0:PE1# trace 10.100.1.3

 1  10.1.15.5 [MPLS: Label 16003 Exp 0] 29 msec  29 msec  39 msec
 2  10.1.59.9 [MPLS: Label 16003 Exp 0] 29 msec  29 msec  39 msec
 3  10.1.129.12 [MPLS: Label 16003 Exp 0] 39 msec  49 msec  29 msec
 4  10.1.126.6 [MPLS: Label 16003 Exp 0] 19 msec  19 msec  39 msec
 5  10.1.36.3 29 msec  *  19 msec
```

no PHP at ABR = LSP is uninterrupted

Example of E Flag




```
router isis 1
 is-type level-1
 net 49.0002.0000.0000.0003.00
 address-family ipv4 unicast
  metric-style wide
  router-id Loopback0
  segment-routing mpls
 !
 interface Loopback0
  address-family ipv4 unicast
   prefix-sid index 3 explicit-null
   !
```

```
RP/0/0/CPU0:P3# show isis database PE3.00-00 verbose
..
  Metric: 10         IP-Extended 10.100.1.3/32
    Prefix-SID Index: 3, Algorithm:0, R:0 N:1 P:1 E:1 V:0 L:0
    Prefix Attribute Flags: X:0 R:0 N:1
    Source Router ID: 10.100.1.3
```

Prefix SID

E:0 explicit-NULL on

```
RP/0/0/CPU0:P3# show mpls forwarding labels 16003

Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
16003  Exp-Null-v4 SR Pfx (idx 3)    Gi0/0/0/1    10.1.36.3       10721
```

LFIB on previous router shows explicit-null label

Troubleshooting – Data Plane

You make networking **possible**

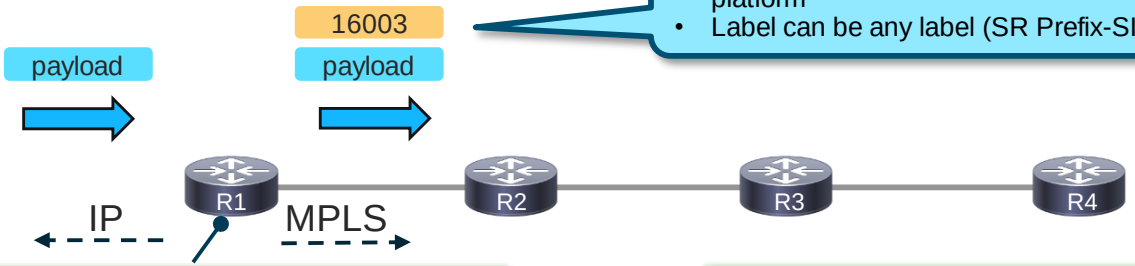
Cisco live!

SR has Regular MPLS Data Plane

- SR has segments: 1 segment represented by 1 label in MPLS label stack
- SR uses the existing MPLS data plane
  - No exceptions
  - MPLS label operations:
    - Push, Pop, and Swap
  - We have
    - Special labels {0 - 15}
    - PHP (default behavior, also for SR)
    - explicit-null for IPv4 and IPv6
    - Regular labels {16 - 1048575}
    - Static labels {16 - 4095}
    - SRGB {16000 - 23999} – Prefix-SIDs
    - Dynamic range {24000 - 1048575} – includes Adj-SIDs
    - QOS propagation (EXP bits)
    - Still uniform model, pipe, and short pipe model
    - TTL propagation as usual

MPLS Label Operation: Push Label(s)

- Push can occur at ingress of MPLS domain
  - MPLS label stack added in CEF (FIB) table
  - Top label is SR label; other labels can be service labels (e.g. MPLS VPN, BGP-LU, etc.)
- Push can occur at intermediate MPLS (P) router
  - MPLS label(s) added in LFIB (e.g. active FRR)



- Label stack added can be up to size supported by the platform
- Label can be any label (SR Prefix-SID, Adj-SID label, …)

```
RP/0/0/CPU0:R1# show route 10.100.1.3/32

Routing entry for 10.100.1.3/32
  Known via "isis 1", … , labeled SR, …
   Routing Descriptor Blocks
    10.1.15.5, from 10.100.1.3, via GigabitEthernet0/0/0/0
      Route metric is 60
```

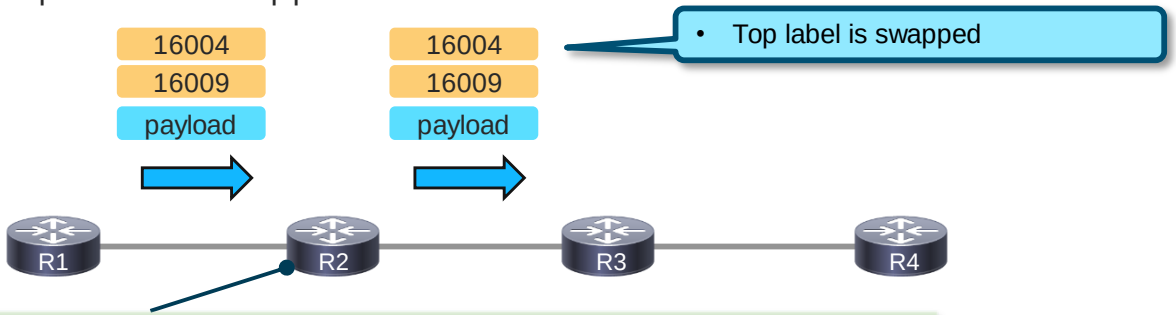
"labeled SR" must be present*

```
RP/0/0/CPU0:R1# show cef …

10.100.1.3/32, …  labeled SR, …

 via 10.1.12.2/32, GigabitEthernet0/0/0/1,…
 local label 16003      labels imposed {16003}
```

* not for Ti-LFA for link prefixes (see later)

MPLS Label Operation: Swap Label(s)

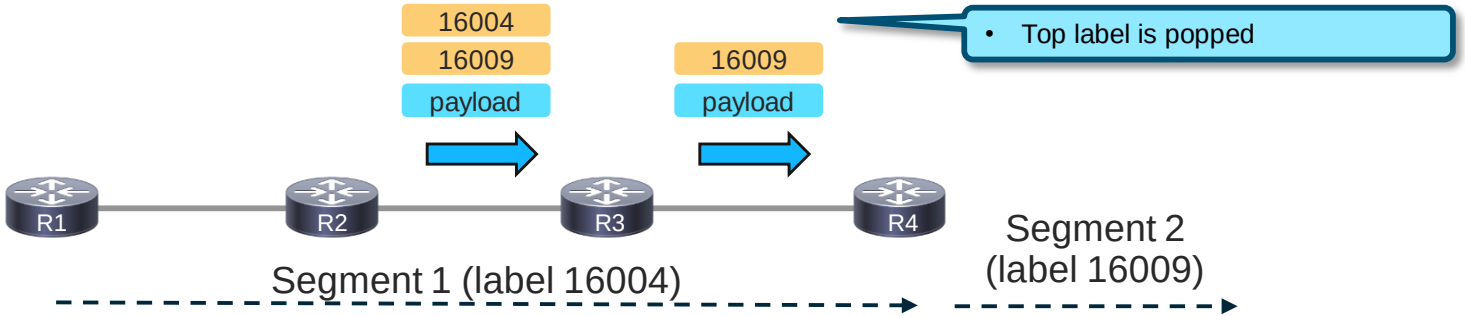
- Swap occurs at intermediate MPLS (P) router
  - Only top label is swapped
  - MPLS label is swapped in LFIB
  - Other labels are not touched (EXP bits, TTL)
  - Within one SR segment, top label is swapped with same label




```
RP/0/0/CPU0:R2# show mpls forwarding labels 16004

Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
16004  16004       SR Pfx (idx 4)    Gi0/0/0/1    10.1.59.9       1420
```

MPLS Label Operation: Pop Label(s)

- Pop occurs at intermediate MPLS (P) router: top label is removed
- By default on penultimate router of one SR segment
  - Label stack could become unlabeled
  - Label stack can still have other labels
    - e.g. when packet is moved from one SR segment to another SR segment




```
RP/0/0/CPU0:R3# show mpls forwarding labels 16004

Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
16004  Pop         SR Pfx (idx 4)    Gi0/0/0/3    10.1.46.4       1880280
```

LFIB

Nothing new!

```
RP/0/0/CPU0:P3# show mpls forwarding

Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
16001  16001       SR Pfx (idx 1)    Gi0/0/0/0    10.1.126.12     0
16002  16002       SR Pfx (idx 2)    Gi0/0/0/0    10.1.126.12     0
16003  Exp-Null-v4 SR Pfx (idx 3)    Gi0/0/0/1    10.1.36.3       43054
16004  16004       SR Pfx (idx 4)    Gi0/0/0/1    10.1.36.3       73402
       16004       SR Pfx (idx 4)    Gi0/0/0/2    10.1.68.8       0
16005  16005       SR Pfx (idx 5)    Gi0/0/0/0    10.1.126.12     0
16008  Pop         SR Pfx (idx 8)    Gi0/0/0/2    10.1.68.8       0
16009  16009       SR Pfx (idx 9)    Gi0/0/0/0    10.1.126.12     0
16010  16010       SR Pfx (idx 10)   Gi0/0/0/0    10.1.126.12     0
       16010       SR Pfx (idx 10)   Gi0/0/0/2    10.1.68.8       0
16012  Pop         SR Pfx (idx 12)   Gi0/0/0/0    10.1.126.12     0
16013  16013       SR Pfx (idx 13)   Gi0/0/0/0    10.1.126.12     0
       16013       SR Pfx (idx 13)   Gi0/0/0/2    10.1.68.8       0
24000  Pop         SR Adj (idx 1)    Gi0/0/0/0    10.1.126.12     0
24003  Pop         SR Adj (idx 2)    Gi0/0/0/0    10.1.126.12     0
24004  Pop         No ID             tt1          point2point     0
```

- Special labels {0-15} are still used
- ECMP, can only be Prefix-SID
- Data plane makes no distinction between Prefix-SID and Adj-SID
- Any Adj-SID will have pop operation
- Binding entry (used with SR-TE)

Load Balancing MPLS Traffic

Nothing new!

- Routers will try to load balance on the IP header, even when there is a label stack
  - IP traffic is best load-balanced by calculating hash over 3- or 7-tuple*
- PseudoWire traffic is load-balanced by calculating hash over the bottom label (PW service label)
  - Preserving per-flow load balancing
- Flow Aware Transport (FAT) Label can be used
  - Endpoints need to support this signaling
  - Endpoints classify traffic and pushes a unique flow label for each flow (each PW)
  - Load balancing on bottom (FAT) label
  - More granular than load balancing
- If many labels are present in label stack and the platform cannot look at/past last label, there is load balancing on higher label in the stack

Check platform!

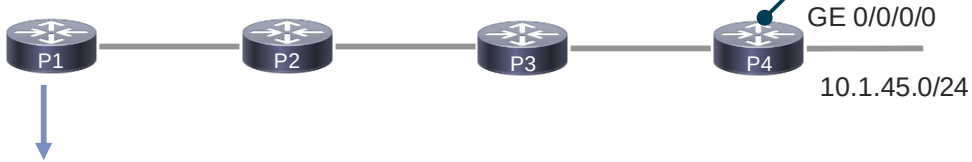
* 3- Source IP, Destination IP, Router ID
  7- Source IP, Destination IP, Router ID, Source port, Destination port, Protocol, Ingress interface handle

Pushing Label Stack at Ingress Router

- Considerations
  - MTU
  - Platform support

- What matters
  - The segment list is normally not large
  - Binding-SIDs are used

Labeling which Prefixes?



Prefix attached to P4	Outgoing label in CEF? Entry in LFIB?
Prefix-SID P4 (10.100.1.4/32 & 10.100.2.4/32)	Y
Prefix-SID P4 without Node flag (10.100.3.4/32)	Y
Loopback prefix without Prefix-SID (10.100.4.4/32)	N
Link prefix connected to P4 (10.1.45.0/24)	N

- So, this is the equivalent of LDP label prefix filtering: only assigning/advertising labels to /32 prefixes (loopback prefixes, used by service, (e.g. L3VPN), so BGP next hop IP addresses)
- Traffic to link prefixes is not labeled!

```
interface Loopback0
 ipv4 address 10.100.1.4 255.255.255.255
!
interface Loopback1
 ipv4 address 10.100.2.4 255.255.255.255
!
interface Loopback2
 ipv4 address 10.100.3.4 255.255.255.255
!
interface Loopback3
 ipv4 address 10.100.4.4 255.255.255.255
!
interface GigabitEthernet0/0/0/0
 ipv4 address 10.1.45.4 255.255.255.0

router isis 1
 net 49.0001.0000.0000.0004.00
 address-family ipv4 unicast
  metric-style wide
  router-id Loopback0
  segment-routing mpls
 !
 interface Loopback0
  address-family ipv4 unicast
   prefix-sid absolute 16004
!
 interface Loopback1
  address-family ipv4 unicast
   prefix-sid absolute 17004
 !
 interface Loopback2
  address-family ipv4 unicast
   prefix-sid absolute 18004 n-flag-clear
 !
 interface Loopback3
  adderss-family ipv4 unicast
  !
 interface GigabitEthernet0/0/0/0
  point-to-point
  address-family ipv4 unicast
```

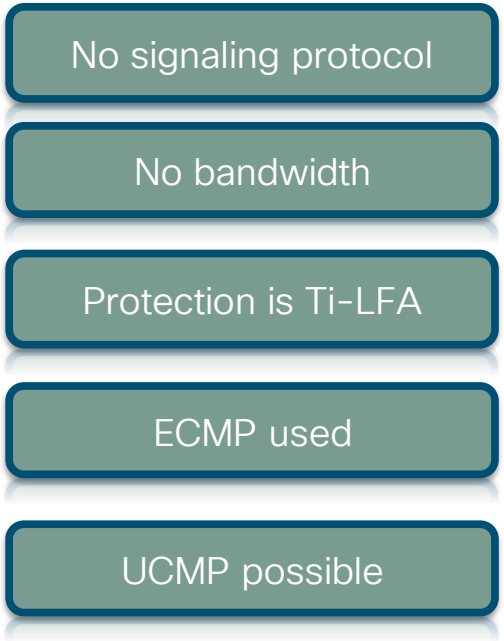
SR Policy

You make networking **possible**

Cisco live!

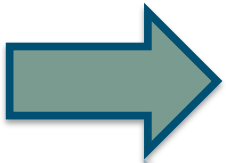
SR Policy

- Simple, automated, and scalable
  - No core state: state in the packet header
  - Traffic engineered “SR Policy”
  - No headend a-priori configuration: on-demand policy instantiation
  - No headend a-priori steering: on-demand-steering
- Dynamic or explicit path
  - Explicit-path is defined as list of segments:
    - All hops are IP addresses (link/node = loopback)
    - All hops are MPLS labels
    - Mix of IP addresses/MPLS labels
- PCE/PCC is possible (multi-domain)



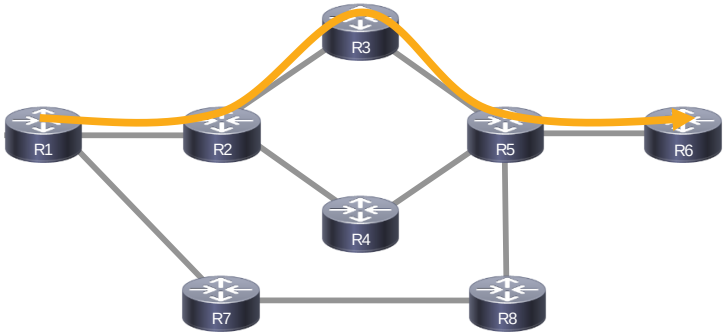
TE to SR Policy

- Forwarding (legacy)
  - Autoroute Announce (AA)
  - Autoroute Destination
  - Static route
  - Access-list Based Forwarding (ABF)

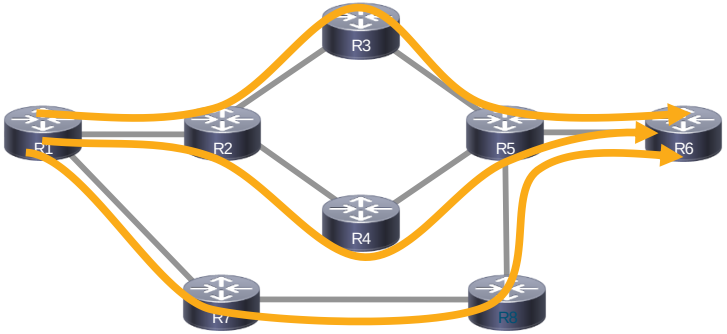


- Automated color-based steering
  - Path preference
    - IGP or TE metric/hopcount/measured latency
  - Constraints
    - Color/affinity
    - Cumulative metric bound
    - Disjoint paths
    - Prefix-SID algorithm

Circuit Optimization versus SR Optimization



- Classic TE is circuit-based
- CSPF: non-ECMP path
- SID-list: R2 - R3 - R5 - R6
- Poor ECMP, big SID-list, ATM-like



- SR native TE
- ECMP paths, no more circuits
- SID-list: R6
- ECMP, small SID-list, IP optimized

SR Policy

- An SR Policy is identified through the following tuple:
  - The head-end where the policy is instantiated/implemented
  - The endpoint (i.e.: the destination of the policy)
  - The color (an arbitrary numerical value)
- At a given head-end, an SR Policy is fully identified by the <color, endpoint> tuple
- An endpoint can be specified as an IPv4 or IPv6 address

Candidate-Path

- 2 options
  1. Dynamic: as before with RSVP-TE, but with SR-native algorithms
  2. Explicit:
     - SR-TE Policy path can be explicitly specified by configuring an ordered list of IP addresses and/or label values
     - The IP addresses in the explicit path will be mapped to label values by the head-end
- Any loopback prefix with associated prefix-SID can be used in path
  - If nodes have multiple loopbacks with prefix-SIDs, each of these can be used in the SR-TE Policy path
  - Including anycast prefixes/prefix-SIDs
- Result of computation is always a list of segments, so SR labels

First segment exception

If the first hop in the SR-TE Policy path is adjacency-SID or prefix-SID of adjacent node
Then the label for this first hop is not added in SR-TE Policy rewrite label stack this hop is only used to select outgoing interface(s)

SR-TE Configuration (new)

```
segment-routing
 global-block 16000 23999
 traffic-eng
  logging
   policy status
  !
  segment-list explicit-to-ABR-1
   index 5 address ipv4 10.1.3.3
   index 10 mpls label 16007
   index 20 mpls label 16009
   index 30 mpls label 16005
  !
  policy to-ABR1
   binding-sid mpls 1000
   color 1000 end-point ipv4 10.0.0.5
   candidate-paths
    preference 100
     dynamic
      metric
       type igp
      !
     !
    !
    preference 200
     explicit segment-list explicit-to-ABR-1
```

SR-TE

Explicit path definition

SR Policy

Dynamic path

Explicit path

candidate-paths with higher preference wins, if the path is valid

Verify SR Policy

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng policy name ?
  srte_c_1000_ep_10.0.0.5  Policy name (if contains space, enclose name with " ")
  WORD                     Policy name (if contains space, enclose name with " ")

RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng policy name srte_c_1000_ep_10.0.0.5

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

Color: 1000, End-point: 10.0.0.5
  Name: srte_c_1000_ep_10.0.0.5
  Status:
    Admin: up  Operational: up for 00:12:09 (since Apr 30 06:20:19.182)
  Candidate-paths:
    Preference: 200 (configuration) (active)
      Name: to-ABR1
      Requested BSID: 1000
      Explicit: segment-list explicit-to-ABR-1 (valid)
        Weight: 1, Metric Type: TE
          24001 [Adjacency-SID, 10.1.3.1 - 10.1.3.3]
          16007
          16009
          16005
    Preference: 100 (configuration)
      Name: to-ABR1
      Requested BSID: 1000
        Maximum SID Depth: 10
      Dynamic (invalid)
        Metric Type: IGP,   Path Accumulated Metric: 20
  Attributes:
    Binding SID: 1000
    Forward Class: 0
    Steering labeled-services disabled: no
    Steering BGP disabled: no
    IPv6 caps enable: yes
```

lookup SR policy by name = *srte_color_endpoint*

FIB @ headend
Incoming label: 1000
Action: pop and push <16007, 16009, 16005>

Explicit path configuration – Rules

- Ensure the path leads all the way to the destination
  - SR Policy end-point explicitly included in path
  - SR Policy end-point must be reachable (is verified by TE), if configured as IP address
  - One can mix Prefix-SIDs and Adj-SIDs
- When using Prefix-SIDs:
  - It's easy to skip hops
  - But Prefix-SID of tail end must be included
- When using Adj-SIDs:
  - Adj-SID has local significance, ensure the packet gets to the advertising node of this Adj-SID first
  - Precede Adj-SID hop by Prefix-SID or Adj-SID to node

Adj-SIDs are not recommended
- They can change (after reboot for example), if specified as label*
- No ECMP
- Worse Ti-LFA backup behavior (not using PC path to destination)

*unless when using manual Adj-SIDs

SR-TE Database

single domain

```
router isis 1
 net 49.0001.0000.0000.0001.00
 segment-routing global-block 16000 23999
 distribute link-state
```

needed for TE topology

or

multiple domain

```
router isis 1
 net 49.0001.0000.0000.0001.00
 segment-routing global-block 16000 23999
 distribute link-state instance-id <32-4294967295>
```

needed for TE topology

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng ipv4 topology summary

XTC Agent's topology database summary:
--------------------------------

Topology nodes:                 3
Prefixes:                       3
  Prefix SIDs:                  3
Links:                          4
  Adjacency SIDs:               4
```

SR-TE Database

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng ipv4 topology

SR-TE topology database
-----------------------

Node 1
  TE router ID: 10.0.0.1
  Host name: PE1
  ISIS system ID: 0000.0000.0001 level-1
  Prefix SID:
    Prefix 10.0.0.1, label 16001 (regular), flags: N:1, R:0, P:0, V:0,
E:0, L:0

  Link[0]: local address 10.1.3.1, remote address 10.1.3.3
    Local node:
      ISIS system ID: 0000.0000.0001 level-1
    Remote node:
      Host name: P1
      ISIS system ID: 0000.0000.0003 level-1
    Metric: IGP 10, TE 10, Latency 10 microseconds
    Bandwidth: Total link 125000000, Reservable 0
    Admin-groups: 0x00000000
    Admin-groups-detail:
    Adj SID: 24001 (unprotected)

  Link[1]: local address 10.1.7.1, remote address 10.1.7.7
    Local node:
      ISIS system ID: 0000.0000.0001 level-1
    …
```

link entry

Adj-SID

```
router isis 1
 is-type level-1
 net 49.0001.0000.0000.0001.00
 distribute link-state
 address-family ipv4 unicast
  metric-style wide
  router-id Loopback0
  segment-routing mpls
 !
 interface Loopback0
  passive
  address-family ipv4 unicast
   prefix-sid index 1
  !
 !
 interface GigabitEthernet0/0/0/0
  point-to-point
  address-family ipv4 unicast
  !
```

- Adj-SIDs with Backup flag is only advertised if LFA is configured (as of 6.2.1)

SR Policy Not Functional

- Inactive policy
  - No valid path found
- Invalid SID list
  - It is empty
  - The headend is unable to resolve the first SID into one or more outgoing interface(s) and next-hop(s)
  - The headend is unable to resolve any non-first SID expressed as an IP address
- Unreachable
  - The headend has no path to the SID in its SR-TE database
- Invalid path
  - A Path is invalid as soon as it has no valid SID list

The headend of an SR Policy updates the validity of a SID list upon network topological change.

Traffic To SR Policy

- Default: No traffic onto SR Policy
  - i.e. SR policy is not next hop in the RIB

- Steering
  - Automatic: matching color of service prefix
  - Static route to SR policy next hop

Traffic To SR Policy: Automatic

```
segment-routing
 global-block 16000 23999
 traffic-eng
  logging
   policy status
  !
  segment-list explicit-to-PE2
   index 10 mpls label 16005
   index 20 mpls label 16002
  !
 policy to-PE2
   binding-sid mpls 2000
   color 2000 end-point ipv4 10.0.0.2
   candidate-paths
    preference 100
     explicit segment-list explicit-to-PE2
     !
```

label 16005 to ABR

label 16002 to remote PE

- Next hop matches endpoint

prefix as received from remote PE

```
RP/0/RP0/CPU0:PE1# show bgp vpnv4 unicast  vrf one 10.0.0.14/32

BGP routing table entry for 10.0.0.14/32, Route Distinguisher: 65000:1
Paths: (1 available, best #1)
  65002
    10.0.0.2 C:2000 (bsid:2000) (metric 20) from 10.0.0.12 (10.0.0.2)
      Received Label 24004
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best,
import-candidate, imported
      Received Path ID 0, Local Path ID 1, version 12
      Extended community: Color:2000 RT:65000:100
      Originator: 10.0.0.2, Cluster list: 10.0.0.12
      SR policy color 2000, up, not-registered, bsid 2000
```

matching color

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng policy color 2000

SR-TE policy database
---------------------
Color: 2000, End-point: 10.0.0.2
  Name: srte_c_2000_ep_10.0.0.2
  Status:
    Admin: up  Operational: up for 01:27:13 (since Apr 30 07:05:04.832)
  Candidate-paths:
    Preference: 100 (configuration) (active)
      Name: to-PE2
      Requested BSID: 2000
      Explicit: segment-list explicit-to-PE2 (valid)
        Weight: 1, Metric Type: TE
          16005 [Prefix-SID, 10.0.0.5]
          16002
```

Traffic To SR Policy: Automatic

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng policy color 2000

SR-TE policy database
---------------------
Color: 2000, End-point: 10.0.0.2
  Name: srte_c_2000_ep_10.0.0.2
  Status:
    Admin: up  Operational: up for 01:27:13 (since Apr 30 07:05:04.832)
  Candidate-paths:
    Preference: 100 (configuration) (active)
      Name: to-PE2
      Requested BSID: 2000
      Explicit: segment-list explicit-to-PE2 (valid)
        Weight: 1, Metric Type: TE
          16005 [Prefix-SID, 10.0.0.5]
          16002
```

```
RP/0/RP0/CPU0:PE1# show route vrf one 10.0.0.14/32

Routing entry for 10.0.0.14/32
  Known via "bgp 65000", distance 200, metric 0
   Routing Descriptor Blocks
    10.0.0.2, from 10.0.0.12
      Nexthop in Vrf: "default", Table: "default", IPv4 Unicast,
```

```
RP/0/RP0/CPU0:PE1# show cef vrf one 10.0.0.14/32

10.0.0.14/32, version 13, internal 0x5000001 0x0
 Prefix Len 32, traffic index 0, precedence n/a, priority 3
   via local-label 2000, 3 dependencies, recursive [flags 0x6000]
    path-idx 0 NHID 0x0 [0xd30fa50 0x0]
    recursion-via-label
    next hop VRF - 'default', table - 0xe0000000
    next hop via 2000/0/21
     next hop srte_c_2000_ labels imposed {ImplNull 24004}
```

next hop

labels used

```
RP/0/RP0/CPU0:PE1# traceroute vrf one 10.0.0.14

 1  10.1.3.3 [MPLS: Labels 16005/16002/24004 Exp 0] 12 msec  6 msec  9 msec
 2  10.3.5.5 [MPLS: Labels 16002/24004 Exp 0] 7 msec  6 msec  6 msec
 3  10.5.6.6 [MPLS: Labels 16002/24004 Exp 0] 9 msec  4 msec  7 msec
 4  10.4.6.4 [MPLS: Labels 16002/24004 Exp 0] 14 msec  60 msec  10 msec
 5  10.2.4.2 [MPLS: Label 24004 Exp 0] 8 msec  7 msec  4 msec
 6  10.2.14.14 5 msec  *  5 msec
```

inter-area traffic steering !

Traffic To SR Policy: Static

```
segment-routing
 global-block 16000 23999
 traffic-eng
  logging
   policy status
  !
  segment-list explicit-to-PE2
   index 10 mpls label 16005
   index 20 mpls label 16002
  !
 policy to-PE2
   binding-sid mpls 2000
   color 2000 end-point ipv4 10.0.0.2
   candidate-paths
    preference 100
     explicit segment-list explicit-to-PE2
     !
```

label 16005 to ABR

label 16002 to remote PE

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng policy color 2000

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

Color: 2000, End-point: 10.0.0.2
  Name: srte_c_2000_ep_10.0.0.2
  Status:
    Admin: up  Operational: up for 02:04:14 (since Apr 30 07:05:04.832)
  Candidate-paths:
    Preference: 100 (configuration) (active)
      Name: to-PE2
      Requested BSID: 2000
      Explicit: segment-list explicit-to-PE2 (valid)
        Weight: 1, Metric Type: TE
          16005 [Prefix-SID, 10.0.0.5]
          16002
```

configure *srte_color_endpoint* as next hop for static route

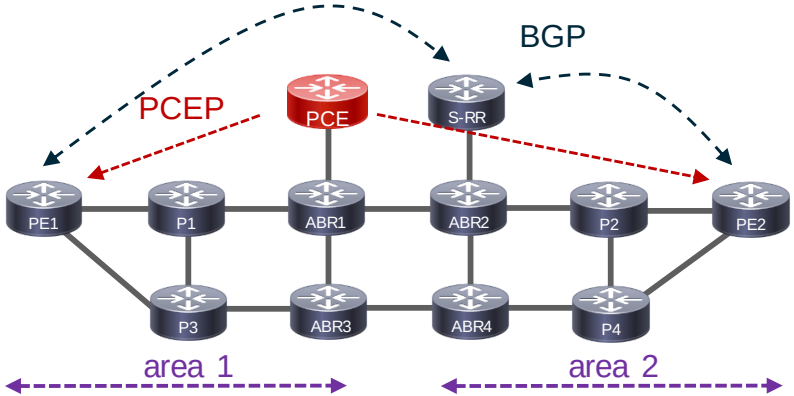
```
router static
 address-family ipv4 unicast
  10.0.0.0/24 10.0.0.5
 !
 vrf one
  address-family ipv4 unicast
   10.0.0.14/32 sr-policy srte_c_2000_ep_10.0.0.2
  !
 !
!
```

ODN

You make networking **possible**

Cisco live!

On-demand Next-hop (ODN)



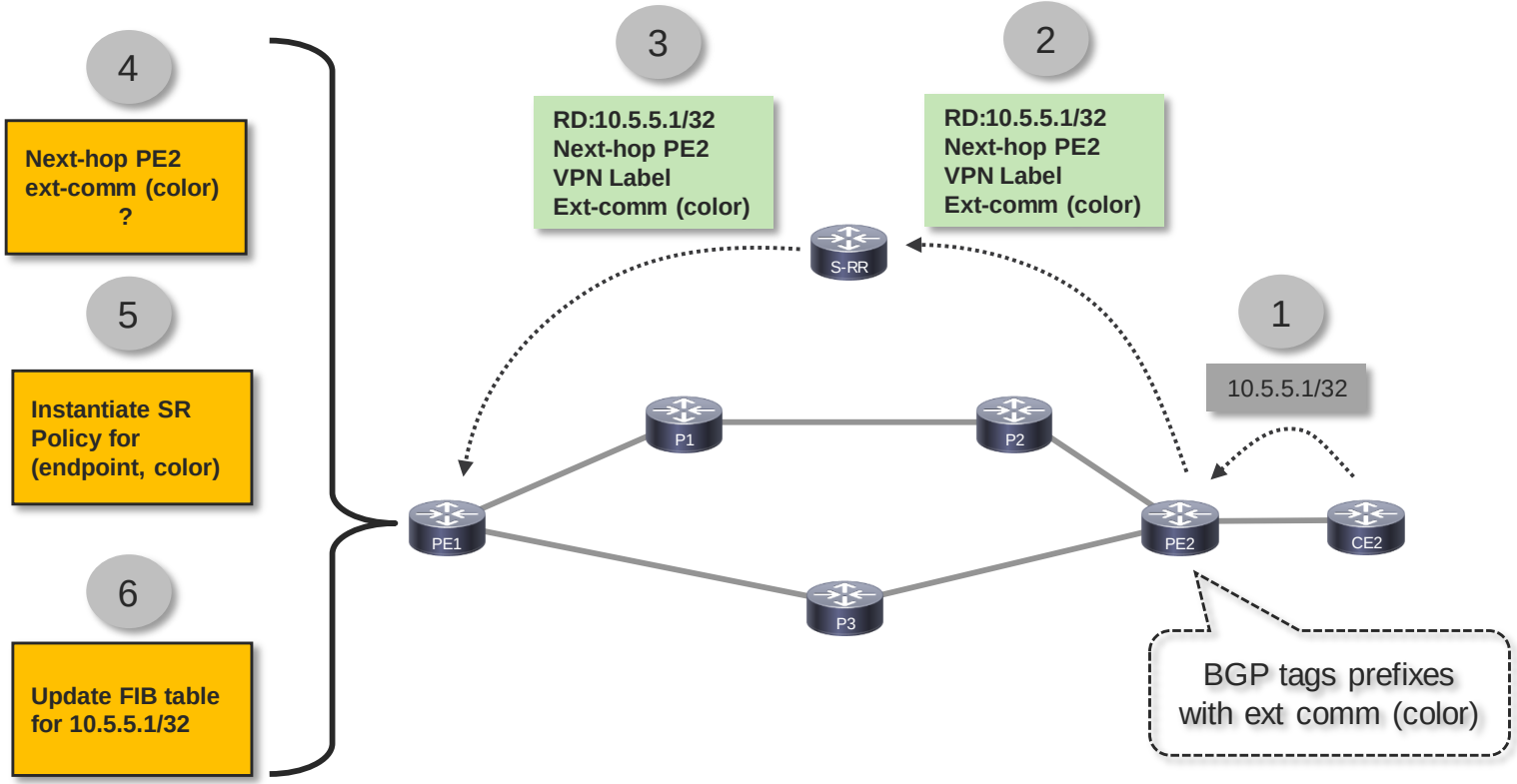
Automated SR Policy
Inter-AS & Seamless MPLS: no need for BGP-LU (RFC3107)
SLA-aware BGP service

- On-demand Next-hop automates and simplifies the service head end configuration
  - No SR Policy config on the head end router
  - No complex/explicit steering on the service head end for the service
    - For example: no autoroute-announce, no static routes
- The SR Policies deployed when needed
  - The learning of the service route, initiates the SR Policy, and traffic-to-SR Policy mapping
  - Example of a service route: vpnv4 route

ODN – On-demand SR Policy

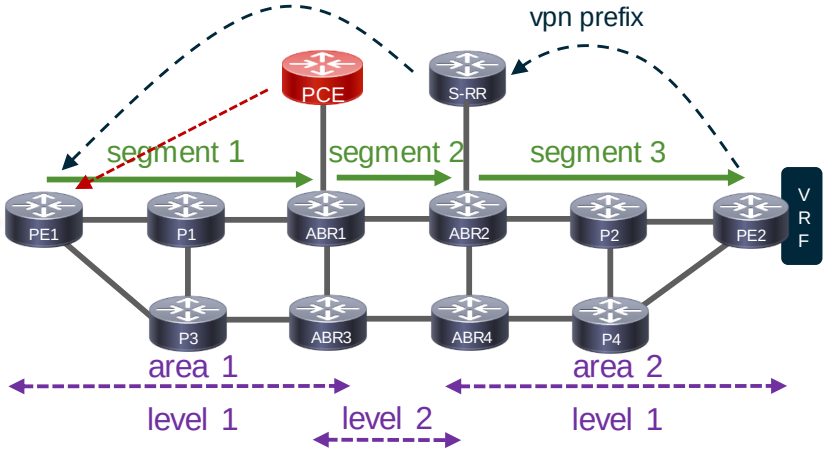
- A service head-end automatically instantiates an SR Policy to a BGP nhop when required (on-demand), automatically steering the BGP traffic into this SR Policy
- Color community is used as SLA indicator
- An SR policy is defined (endpoint, color)

ODN Workflow



ODN Troubleshooting

- receive vpn prefix
- next-hop is PE
- compute SR Policy
  - by head end or by PCE



head end configuration (PE1)

```
segment-routing
 global-block 16000 23999
 traffic-eng
  logging
   policy status
  !
  on-demand color 100
   dynamic
    pcep
    !
    metric
     type igp
    !
   !
  !
  pcc
   source-address ipv4 10.0.0.1
   pce address ipv4 10.0.0.11
   !
```

ODN color 100

SR policy path calculated by PCE

RPL inbound to set color to green (100)

tail end configuration (PE2)

```
extcommunity-set opaque green
  100
end-set
!
route-policy ODN-CE2
  set extcommunity color green
end-policy
!
router bgp 65000
!
 vrf one
  rd 65000:2
  address-family ipv4 unicast
  !
  neighbor 10.2.14.14
   remote-as 65002
   address-family ipv4 unicast
    route-policy ODN-CE2 in
    route-policy PASS out
```

Service Route

```
RP/0/RP0/CPU0:PE1# show bgp vpnv4 unicast vrf one 10.0.0.14/32

BGP routing table entry for 10.0.0.14/32, Route Distinguisher: 65000:1
 65002
    10.0.0.2 C:100 (bsid:24010) (metric 20) from 10.0.0.12 (10.0.0.2)
      Received Label 24004
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate, imported
      Received Path ID 0, Local Path ID 1, version 63
      Extended community: Color:100 RT:65000:100
      Originator: 10.0.0.2, Cluster list: 10.0.0.12
      SR policy color 100, up, registered, bsid 24010, if-handle 0x00000034

      Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65000:2
```

recursing on the BSID of the SR Policy

```
RP/0/RP0/CPU0:PE1# show cef vrf one 10.0.0.14/32

10.0.0.14/32, version 46, internal 0x5000001 0x0 (ptr 0xdf2505c) [1], 0x0 (0xe0e7c68), 0xa08 (0xe610228)
 Updated Apr 29 06:51:58.509
 Prefix Len 32, traffic index 0, precedence n/a, priority 3
   via local-label 24010, 3 dependencies, recursive [flags 0x6000]
    path-idx 0 NHID 0x0 [0xd4cc930 0x0]
    recursion-via-label
    next hop VRF - 'default', table - 0xe0000000
    next hop via 24010/0/21
     next hop srte_c_100_e labels imposed {ImplNull 24004}
```

display of resolved path

next hop points to the interface representing the SR Policy

service label (is VPN label here)

SR Policy Path

head end (PE1)

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng policy color 100

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

Color: 100, End-point: 10.0.0.2
  Name: srte_c_100_ep_10.0.0.2
  Status:
    Admin: up  Operational: up for 01:26:32 (since Apr 29 06:03:25.176)
  Candidate-paths:
    Preference: 200 (BGP ODN) (shutdown)
      Requested BSID: dynamic
        Maximum SID Depth: 10
      Dynamic (invalid)
    Preference: 100 (BGP ODN) (active)
      Requested BSID: dynamic
      PCC info:
        Symbolic name: bgp_c_100_ep_10.0.0.2_discr_100
        PLSP-ID: 4
        Maximum SID Depth: 10
      Dynamic (pce 10.0.0.11) (valid)
        Metric Type: IGP,   Path Accumulated Metric: 50
          16005 [Prefix-SID, 10.0.0.5]
          16002 [Prefix-SID, 10.0.0.2]
  Attributes:
    Binding SID: 24010
    Forward Class: 0
    Steering labeled-services disabled: no
    Steering BGP disabled: no
    IPv6 caps enable: yes
```

engineered path per SR Policy (PCE)

16005 is label to ABR1
16002 is label to PE2

PCE

```
RP/0/RP0/CPU0:PCE# show pce lsp detail

PCE's tunnel database:
----------------------
PCC 10.0.0.1:

Tunnel Name: bgp_c_100_ep_10.0.0.2_discr_100
 LSPs:
  LSP[0]:
   source 10.0.0.1, destination 10.0.0.2, tunnel ID 4, LSP ID 1
   State: Admin up, Operation up
   Setup type: Segment Routing
   Binding SID: 24010
   Maximum SID Depth: 10
   Bandwidth: signaled 0 kbps, applied 0 kbps
   PCEP information:
     PLSP-ID 0x4, flags: D:1 S:0 R:0 A:1 O:1 C:0
   LSP Role: Single LSP
   State-sync PCE: None
   PCC: 10.0.0.1
   LSP is subdelegated to: None
   Reported path:
     Metric type: IGP, Accumulated Metric 50
      SID[0]: Node, Label 16005, Address 10.0.0.5
      SID[1]: Node, Label 16002, Address 10.0.0.2
   Computed path: (Local PCE)
     Computed Time: Mon Apr 29 07:33:27 UTC 2019 (00:02:54 ago)
     Metric type: IGP, Accumulated Metric 50
      SID[0]: Node, Label 16005, Address 10.0.0.5
      SID[1]: Node, Label 16002, Address 10.0.0.2
   Recorded path:
     None
   Disjoint Group Information:
     None
```

```
CE1# trace 10.0.0.14 source loopback 0 numeric

Type escape sequence to abort.
Tracing the route to 10.0.0.14
VRF info: (vrf in name/id, vrf out name/id)
  1 10.1.13.1 2 msec 1 msec 1 msec
  2 10.1.3.3 [MPLS: Labels 16005/16002/24004 Exp 0] 14 msec 8 msec 9 msec
  3 10.3.5.5 [MPLS: Labels 16002/24004 Exp 0] 7 msec 8 msec 8 msec
  4 10.5.6.6 [MPLS: Labels 16002/24004 Exp 0] 7 msec 9 msec 8 msec
  5 10.4.6.4 [MPLS: Labels 16002/24004 Exp 0] 25 msec 8 msec 8 msec
  6 10.2.4.2 [MPLS: Label 24004 Exp 0] 41 msec 9 msec 12 msec
  7 10.2.14.14 6 msec *  8 msec
```

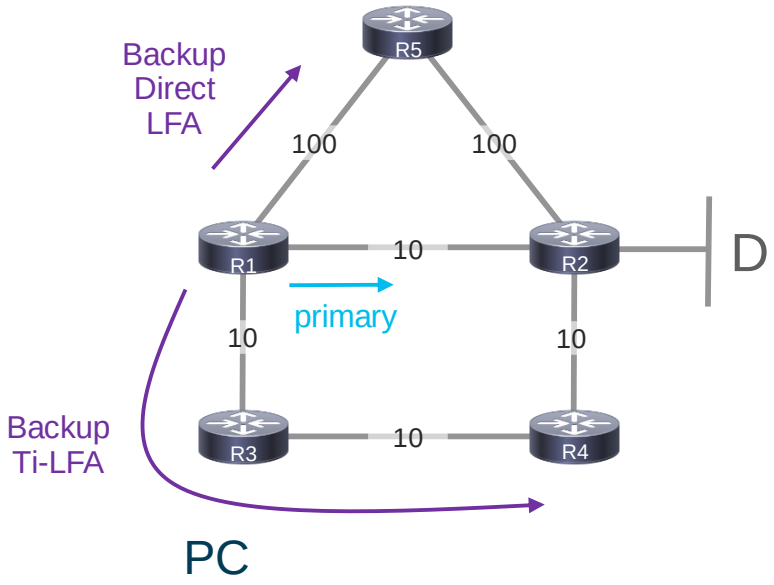
Ti-LFA

You make networking **possible**

LFA Recap

- Loop Free Alternate (LFA)
- Algorithm to calculate backup paths for IP (and MPLS)
- No signaling
- Link or node protection, and other tiebreakers
- Topology dependence: sometimes coverage issues
- Ti-LFA
  - Algorithm, with similar tiebreakers
  - Using segments to force traffic over backup path
    - 100% coverage
  - Protected traffic is on Post-Convergence (PC) path
    - Avoiding another path move at regular convergence after failure
    - Not available with LFA

Ti-LFA Uses PC Path

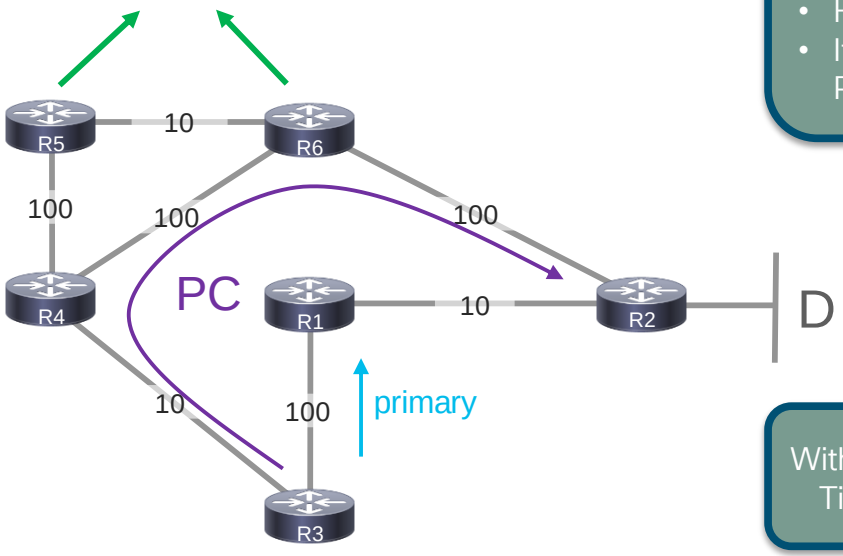


No Ti-LFA:
Even if Remote-LFA (targeted LDP) is used with LFA
Directly connected LFA is preferred over Remote LFA (can be tweaked with tie-breakers)

With Ti-LFA:
Ti-LFA uses PC

Remote LFA

2 PQ candidates, at equal cost from R3



No Ti-LFA:
- PQ node closest to calculating router is chosen
- Here: 2 PQ candidates
- If R5 is the PQ node: backup traffic is not on PC

With Ti-LFA:
Ti-LFA uses PC

Topology Independent LFA (Ti-LFA)

- Must have LFA enabled!
- Must have Ti-LFA enabled

```
router isis 1
 is-type level-2-only
 net 49.0001.0000.0000.0001.00
 address-family ipv4 unicast
  metric-style wide
  segment-routing mpls
 !
 interface Loopback0
  address-family ipv4 unicast
   prefix-sid absolute 16001
  !
 !
 interface GigabitEthernet0/0/0/0
  address-family ipv4 unicast
   fast-reroute per-prefix
   fast-reroute per-prefix ti-lfa
```

```
router ospf 1
 router-id 10.100.1.1
 fast-reroute per-prefix
 fast-reroute per-prefix ti-lfa enable
 address-family ipv4 unicast
 area 0
  segment-routing forwarding mpls !! On by default
  segment-routing mpls
  interface Loopback0
   prefix-sid absolute 16001
  !
  interface GigabitEthernet0/0/0/0
   network point-to-point
```

What is Protected?

- Prefix-SID prefix:
  - Ti-LFA backup

- For non-Prefix-SID/link prefix:
  - Ti-LFA backup is available (there is no Prefix label)
  - Same backup segments are used for Prefix-SID
  - While FRR is active, traffic for these prefixes are forwarded with labels, while when there is no FRR active, the traffic is not labeled!

- SR policy segments are protected by Ti-LFA

Protected/Non-Protected Adj-SID

```
RP/0/0/CPU0:R1# show isis database R1.00-00 verbose

IS-IS 1 (Level-2) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime  ATT/P/OL
R1.00-00            * 0x0000001b   0x448d        950             0/0/0
  Area Address:   49.0001
  NLPID:          0xcc
  Hostname:       R1
  IP Address:     10.100.1.1
  Router Cap:     10.100.1.1, D:0, S:0
    Segment Routing: I:1 V:0, SRGB Base: 16000 Range: 8000

  Metric: 10         IS-Extended R2.00
    Interface IP Address: 10.1.12.1
    Neighbor IP Address: 10.1.12.2
    ADJ-SID: F:0 B:1 V:1 L:1 S:0 weight:0 Adjacency-sid:28107
    ADJ-SID: F:0 B:0 V:1 L:1 S:0 weight:0 Adjacency-sid:28108
```

Different labels; one for protected, one for unprotected Adj-SID

```
RP/0/0/CPU0:R1# show isis adjacency detail

IS-IS 1 Level-2 adjacencies:
System Id      Interface        SNPA           State Hold Changed  NSF IPv4 IPv6
                                                                    BFD  BFD
R2             Gi0/0/0/0        *PtoP*         Up    24   00:01:24 Yes None None
  Area Address:            49.0001
  Neighbor IPv4 Address:   10.1.12.2*
  Adjacency SID:           28107
  Non-FRR Adjacency SID:   28108
  Topology:                IPv4 Unicast
```

Different labels; one for protected, one for unprotected Adj-SID

Ti-LFA Protecting Adj-SID

```
RP/0/RP0/CPU0:R1# show isis adjacency gigabitEthernet 0/0/0/0 detail

IS-IS 1 Level-2 adjacencies:
System Id      Interface                SNPA           State Hold Changed  NSF IPv4 IPv6
                                                                       BFD  BFD
R3             Gi0/0/0/0                *PtoP*         Up    24   2d21h    Yes None None
  Area Address:           49.0001
  Neighbor IPv4 Address:  10.1.3.3*
  Adjacency SID:          24004 (protected)
   Backup label stack:    [16003]
   Backup stack size:     1
   Backup interface:      Gi0/0/0/1
   Backup nexthop:        10.1.2.2
   Backup node address:   10.0.0.3
  Non-FRR Adjacency SID:  24005
  Topology:               IPv4 Unicast
  BFD Status:             BFD Not Required, Neighbor Useable

Total adjacency count: 1
```

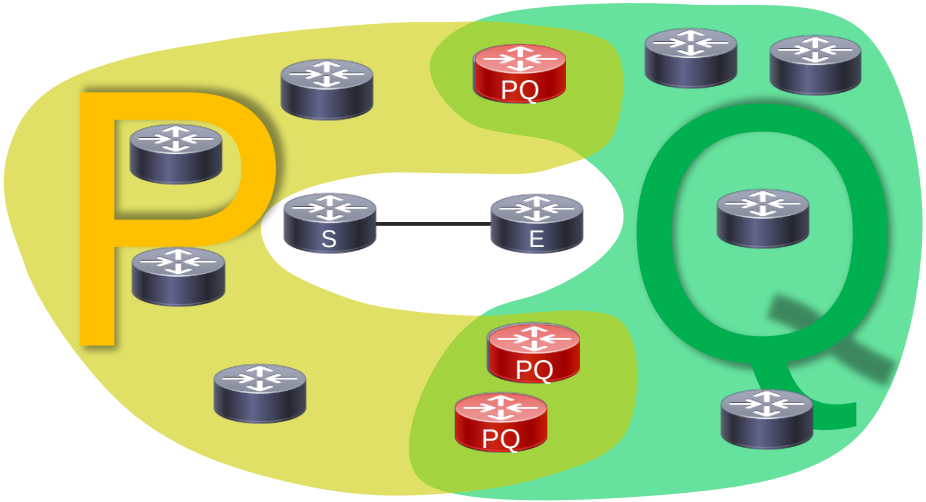
case of 3 or more backup labels:
SR Policy is used

```
RP/0/RP0/CPU0:R5# show isis adjacency gigabitEthernet 0/0/0/1 detail

IS-IS 1 Level-2 adjacencies:
System Id      Interface                SNPA           State Hold Changed  NSF IPv4 IPv6
                                                                       BFD  BFD
R7             Gi0/0/0/1                *PtoP*         Up    24   2d21h    Yes None None
  Area Address:           49.0001
  Neighbor IPv4 Address:  10.5.7.7*
  Adjacency SID:          24004 (protected)
   Backup label stack:    [16007]
   Backup stack size:     1
   Backup interface:      tt1000
   Backup nexthop:        0.0.0.0
   Backup node address:   10.0.0.7
   Backup for TE (no backup tunnel)
    Backup label stack:   [16008, 24003, 24003, 16007]
    Backup stack size:    4
    Backup interface:     Gi0/0/0/2
    Backup nexthop:       10.5.6.6
    Backup node address:  10.0.0.7
  Non-FRR Adjacency SID:  24005
  Topology:               IPv4 Unicast
  BFD Status:             BFD Not Required, Neighbor Useable

Total adjacency count: 1
```

LFA Theory: Calculating P-Space & Q-Space

P space of Router S

P-space of S and the link SE = set of routers that S can reach without passing through the link SE (including ECMP)



Q space of Router E

Q-space of the E and the link SE = set of routers that can reach the router E without passing through the link SE

PQ Common router P & Q

TI-LFA Selection Preference

- The post-convergence (PC) path is expressed as a list of segments

new

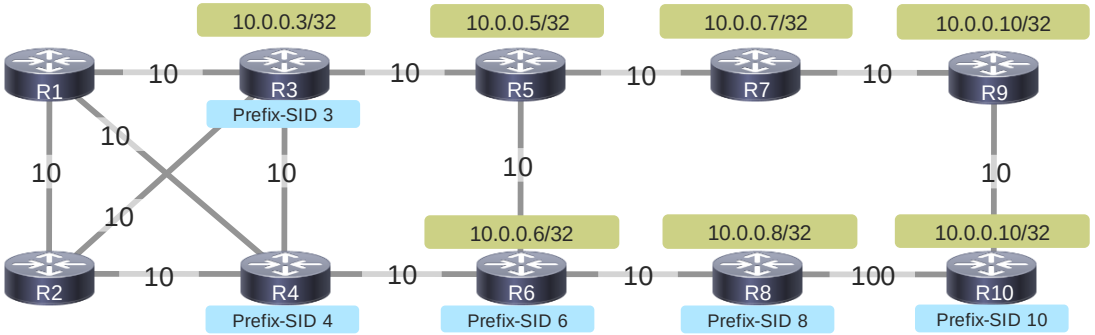
- Typically the shortest list
- Compute and express PC path in a SID-list

  * Use only P and Q segments (no more PQ)

TI-LFA Troubleshooting

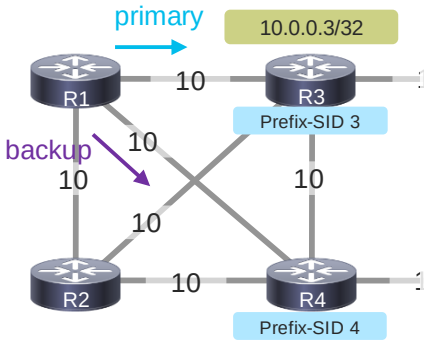
- show route <prefix>
  - Check for protected/backup path (!)
  - “Labeled SR” must be present
  - Check outgoing interfaces
- show cef <prefix>
  - Check for backup path and backup SR labels
- show isis fast-reroute <prefix> detail
  - Check for outgoing interface
  - Check for next hop
  - Check for P/Q node (single-segment) or P and Q nodes (double-segment)
  - “detail” for tie-breakers
- show ospf fast-reroute topology
- show mpls forwarding
  - Check for backup path in LFIB (!)

Ti-LFA Protection Examples



- Example network for the following slides
- A Prefix-SID label is used to get to PQ node (older IOS-XR code)
- A Prefix-SID label is used to get to P node
- An Adj-SID label is used to get to Q node

Ti-LFA – 0-Segment Example



RIB

```
RP/0/RP0/CPU0:R1# show route 10.0.0.3/32

Routing entry for 10.0.0.3/32
  Known via "isis 1", distance 115, metric 10, labeled SR, type level-2
  Installed Jun 11 17:53:51.454 for 00:00:21
  Routing Descriptor Blocks
    10.1.3.3, from 10.0.0.3, via GigabitEthernet0/0/0/0, Protected
      Route metric is 10
    10.1.4.4, from 10.0.0.3, via GigabitEthernet0/0/0/2, Backup (Local-LFA)
      Route metric is 20
  No advertising protos.
```

ISIS

```
RP/0/RP0/CPU0:R1# show isis fast-reroute 10.0.0.3/32

L2 10.0.0.3/32 [10/115]
     via 10.1.3.3, GigabitEthernet0/0/0/0, R3, SRGB Base: 16000, Weight: 0
       Backup path: LFA, via 10.1.4.4, GigabitEthernet0/0/0/2, R4, SRGB Base: 16000,
Weight: 0, Metric: 20
```

FIB

```
RP/0/RP0/CPU0:R1# show cef 10.0.0.3/32
10.0.0.3/32, version 131, labeled SR
remote adjacency to GigabitEthernet0/0/0/0
 Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 10.1.3.3/32, GigabitEthernet0/0/0/0, 10 dependencies, weight 0, class 0, protected [flags 0x400]
    path-idx 0 bkup-idx 1 NHID 0x0 [0xf377f90 0xf377eb0]
    next hop 10.1.3.3/32
     local label 16003      labels imposed {ImplNull}
   via 10.1.4.4/32, GigabitEthernet0/0/0/2, 10 dependencies, weight 0, class 0, backup (Local-LFA) [flags 0x300]
    path-idx 1 NHID 0x0 [0xf1242f0 0x0]
    next hop 10.1.4.4/32
    remote adjacency
     local label 16003      labels imposed {16003}
```

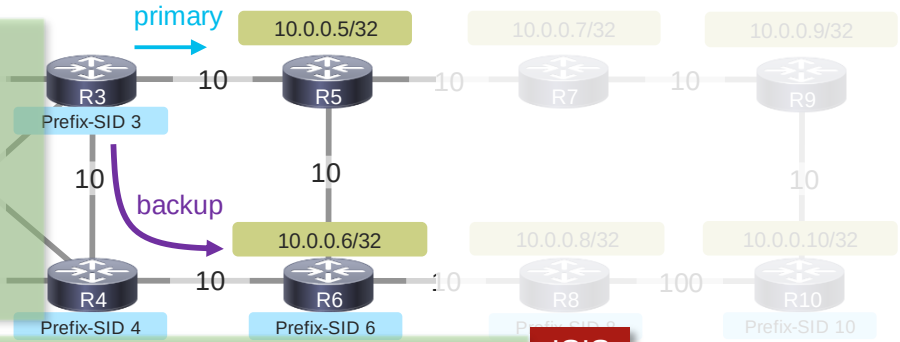
no additional label

no P/Q information

Ti-LFA – Single-Segment Example

RIB

```
RP/0/RP0/CPU0:R3# show route 10.0.0.5/32

Routing entry for 10.0.0.5/32
  Known via "isis 1", distance 115, metric 10, labeled SR, type level-2
  Installed Apr 26 13:59:29.323 for 3d00h
  Routing Descriptor Blocks
    10.3.4.4, from 10.0.0.5, via GigabitEthernet0/0/0/2, Backup (TI-LFA)
      Repair Node(s): 10.0.0.6
      Route metric is 30
    10.3.5.5, from 10.0.0.5, via GigabitEthernet0/0/0/1, Protected
      Route metric is 10
  No advertising protos.
```



ISIS

```
RP/0/RP0/CPU0:R3# show isis fast-reroute 10.0.0.5/32

L2 10.0.0.5/32 [10/115]
     via 10.3.5.5, GigabitEthernet0/0/0/1, R5, SRGB Base: 16000, Weight: 0
       Backup path: TI-LFA (link), via 10.3.4.4, GigabitEthernet0/0/0/2 R4, SRGB Base: 16000, Weight: 0, Metric: 30
         P node: R6.00 [10.0.0.6], Label: 16006
         Prefix label: 16005
         Backup-src: R5.00
```

P

one additional label

FIB

```
RP/0/RP0/CPU0:R3# show cef 10.0.0.5/32
10.0.0.5/32, version 212, labeled SR
 remote adjacency to GigabitEthernet0/0/0/1
 Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 10.3.4.4/32, GigabitEthernet0/0/0/2, 17 dependencies, weight 0, class 0, backup (TI-LFA) [flags 0xb00]
    path-idx 0 NHID 0x0 [0xf1244a0 0x0]
    next hop 10.3.4.4/32, Repair Node(s): 10.0.0.6
     local label 16005      labels imposed {16006 16005}
   via 10.3.5.5/32, GigabitEthernet0/0/0/1, 17 dependencies, weight 0, class 0, protected [flags 0x400]
    path-idx 1 bkup-idx 0 NHID 0x0 [0xf29e070 0xf29d0b0]
    next hop 10.3.5.5/32
     local label 16005      labels imposed {ImplNull}
```

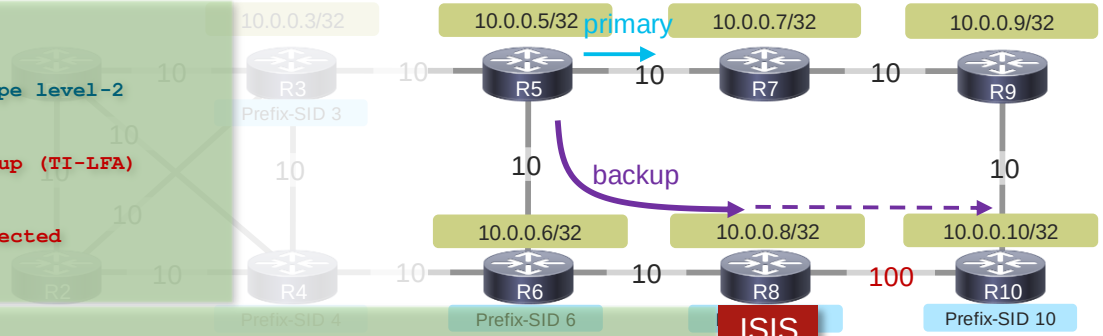
repair node

one additional label

Ti-LFA – Double-Segment Example

RIB

```
RP/0/RP0/CPU0:R5# show route 10.0.0.7/32

Routing entry for 10.0.0.7/32
  Known via "isis 1", distance 115, metric 10, labeled SR, type level-2
  Installed Jun 11 17:40:41.894 for 00:00:39
  Routing Descriptor Blocks
    10.5.6.6, from 10.0.0.7, via GigabitEthernet0/0/0/2, Backup (TI-LFA)
      Repair Node(s): 10.0.0.8, 10.0.0.10
      Route metric is 140
    10.5.7.7, from 10.0.0.7, via GigabitEthernet0/0/0/1, Protected
      Route metric is 10
  No advertising protos.
```




```
RP/0/RP0/CPU0:R5# show isis fast-reroute 10.0.0.7/32

L2 10.0.0.7/32 [10/115]
     via 10.5.7.7, GigabitEthernet0/0/0/1, R7, SRGB Base: 16000, Weight: 0
       Backup path: TI-LFA (link), via 10.5.6.6, GigabitEthernet0/0/0/2 R6, SRGB Base: 16000, Weight: 0, Metric: 140
         P node: R8.00 [10.0.0.8], Label: 16008
         Q node: R10.00 [10.0.0.10], Label: 24003
         Prefix label: 16007
         Backup-src: R7.00
```

P

Q

two additional labels

FIB

```
RP/0/RP0/CPU0:R5# show cef 10.0.0.7/32
10.0.0.7/32, version 138, labeled SR
  Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 10.5.6.6/32, GigabitEthernet0/0/0/2, 16 dependencies, weight 0, class 0, backup (TI-LFA) [flags 0xb00]
   next hop 10.5.6.6/32, Repair Node(s): 10.0.0.8, 10.0.0.10
    remote adjacency
     local label 16007      labels imposed {16008 24003 16007}
   via 10.5.7.7/32, GigabitEthernet0/0/0/1, 16 dependencies, weight 0, class 0, protected [flags 0x400]
    path-idx 1 bkup-idx 0 NHID 0x0 [0xf24da50 0xf24dc10]
    next hop 10.5.7.7/32
     local label 16007      labels imposed {ImplNull}
```

repair nodes

P

Q

two additional labels

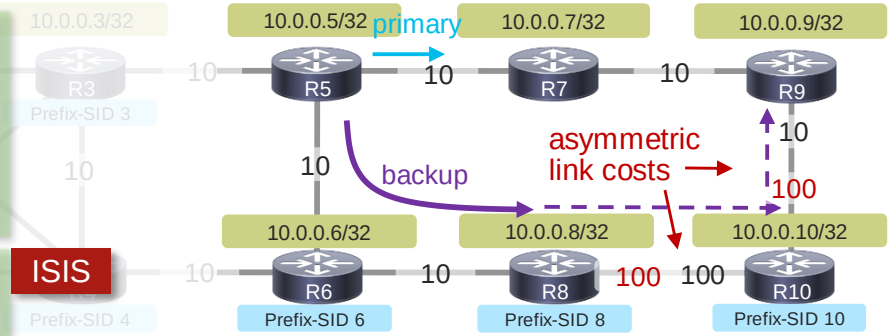
TI-LFA – Triple -Segment ???

RIB

```
RP/0/RP0/CPU0:R5# show route 10.0.0.7/32

Routing entry for 10.0.0.7/32
  Known via "isis 1", distance 115, metric 10, labeled SR, type level-2
  Installed Apr 29 15:36:21.174 for 00:06:59
  Routing Descriptor Blocks
    10.5.7.7, from 10.0.0.7, via GigabitEthernet0/0/0/1
      Route metric is 10
  No advertising protos.
```

```
RP/0/RP0/CPU0:R5# show isis fast-reroute  10.0.0.7/32

L2 10.0.0.7/32 [10/115]
     via 10.5.7.7, GigabitEthernet0/0/0/1, R7, SRGB Base: 16000, Weight: 0
       Backup path: TI-LFA (link), via 10.5.6.6, GigabitEthernet0/0/0/2 R6,
SRGB Base: 16000, Weight: 0, Metric: 230
       Backup tunnel: requested
         P node: R8.00 [10.0.0.8], Label: 16008
         Q node: R10.00 [10.0.0.10], Label: 24003
         Q node: R9.00 [10.0.0.9], Label: 24003
         Prefix label: 16007
         Backup-src: R7.00
```

ISIS



FAIL

FIB

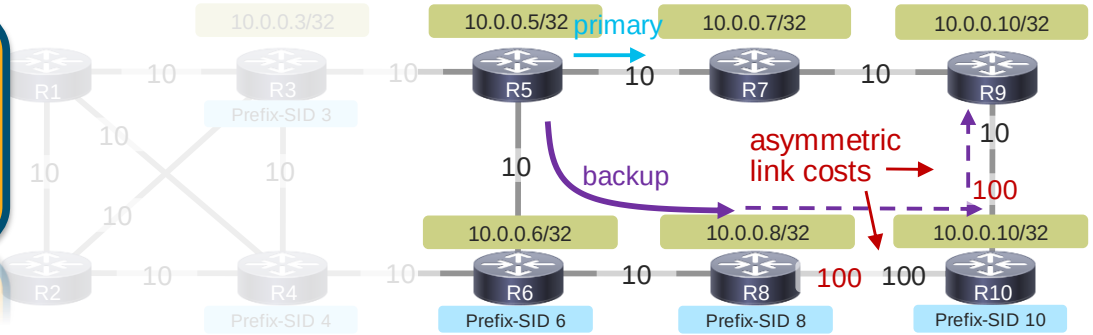
```
RP/0/RP0/CPU0:R5# show cef 10.0.0.7/32

10.0.0.7/32, version 10, labeled SR, internal 0x1000001 0x81 (ptr 0xdf25ab8) [1], 0x0 (0xe0e79e8), 0xa20 (0xe72a328)
 Updated Apr 29 15:36:21.181
 remote adjacency to GigabitEthernet0/0/0/1
 Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 10.5.7.7/32, GigabitEthernet0/0/0/1, 9 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 NHID 0x0 [0xf124140 0xf124380]
    next hop 10.5.7.7/32
    remote adjacency
     local label 16007      labels imposed {ImplNull}
```

no protected path

no backup path

TI-LFA – Triple (and more)-Segment – Solution

IOS-XR needs internal SR policy when there are more than 3 P/Q/repair nodes




```
ipv4 unnumbered mpls traffic-eng Loopback0
…
segment-routing
```

must have configuration for SR internal policy

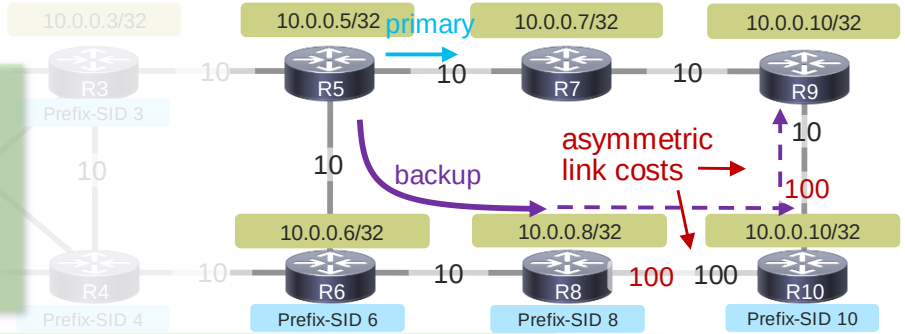
```
mpls traffic-eng
 auto-tunnel p2p
 tunnel-id min 1000 max 1999
```

optional configuration

TI-LFA – Triple-Segment

RIB

```
RP/0/RP0/CPU0:R5# show route 10.0.0.7/32

Routing entry for 10.0.0.7/32
  Known via "isis 1", distance 115, metric 10, labeled SR, type level-2
  Routing Descriptor Blocks
    directly connected, via tunnel-te1000, Backup (Local-LFA)
      Route metric is 230
    10.5.7.7, from 10.0.0.7, via GigabitEthernet0/0/0/1, Protected
      Route metric is 10
  No advertising protos.
```



ISIS

```
RP/0/RP0/CPU0:R5# show isis fast-reroute 10.0.0.7/32

L2 10.0.0.7/32 [10/115]
     via 10.5.7.7, GigabitEthernet0/0/0/1, R7, SRGB Base: 16000, Weight: 0
       Backup path: TI-LFA (link), via 10.5.6.6, GigabitEthernet0/0/0/2 R6, SRGB Base: 16000, Weight: 0, Metric: 230
       Backup tunnel: tunnel-te1000
         P node: R8.00 [10.0.0.8], Label: 16008
         Q node: R10.00 [10.0.0.10], Label: 24003
         Q node: R9.00 [10.0.0.9], Label: 24003
         Prefix label: 16007
         Backup-src: R7.00
```

P
Q
Q

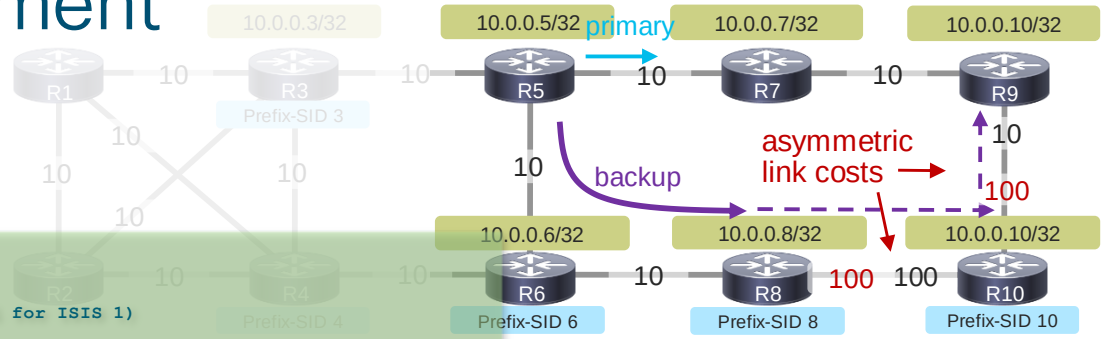
three additional labels

FIB

```
RP/0/RP0/CPU0:R5# show cef 10.0.0.7/32
10.0.0.7/32, version 120, labeled SR, internal 0x1000001 0x81 (ptr 0xe048be8) [1], 0x0 (0xe20bee8), 0xa28 (0xf19e600)
 remote adjacency to GigabitEthernet0/0/0/1
 Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 0.0.0.0/32, tunnel-te1000, 9 dependencies, weight 0, class 0, backup (Local-LFA) [flags 0x300]
    path-idx 0 NHID 0x0 [0xf124410 0x0]
    next hop 0.0.0.0/32
    local adjacency
     local label 16007      labels imposed {16007}
   via 10.5.7.7/32, GigabitEthernet0/0/0/1, 9 dependencies, weight 0, class 0, protected [flags 0x400]
    path-idx 1 bkup-idx 0 NHID 0x0 [0xf255dd0 0xf255cf0]
    next hop 10.5.7.7/32
     local label 16007      labels imposed {ImplNull}
```

internal SR policy

TI-LFA – Triple-Segment



MPLS TE

```
RP/0/RP0/CPU0:R5# show mpls traffic-eng tunnels 1000


Name: tunnel-te1000  Destination: 0.0.0.0  Ifhandle:0x1c (auto-tunnel for ISIS 1)
  Signalled-Name: auto_R5_t1000
  Status:
    Admin:    up Oper:   up   Path:  valid   Signalling: connected

    path option (_te1000), preference 10, (verbatim Segment-Routing) type explicit (_te1000) (Basis for Setup)
    G-PID: 0x0800 (derived from egress interface properties)
    Bandwidth Requested: 0 kbps  CT0
 Config Parameters:
    Bandwidth:        0 kbps (CT0) Priority:  7  7 Affinity: 0x0/0xffff
    Metric Type: TE (global)
     Forwarding-Adjacency: disabled
    Autoroute Destinations: 0
    …
    BFD Fast Detection: Disabled
    Reoptimization after affinity failure: Enabled
    SRLG discovery: Disabled

  Segment-Routing Path Info (IGP information is not used)
    Segment0[First Hop]: 10.5.6.6, Label: -          P
    Segment1[ - ]: Label: 16008                      Q
    Segment2[ - ]: Label: 24003                      Q
    Segment3[ - ]: Label: 24003

Displayed 1 (of 1) heads, 0 (of 0) midpoints, 0 (of 0) tails
Displayed 1 up, 0 down, 0 recovering, 0 recovered heads
```

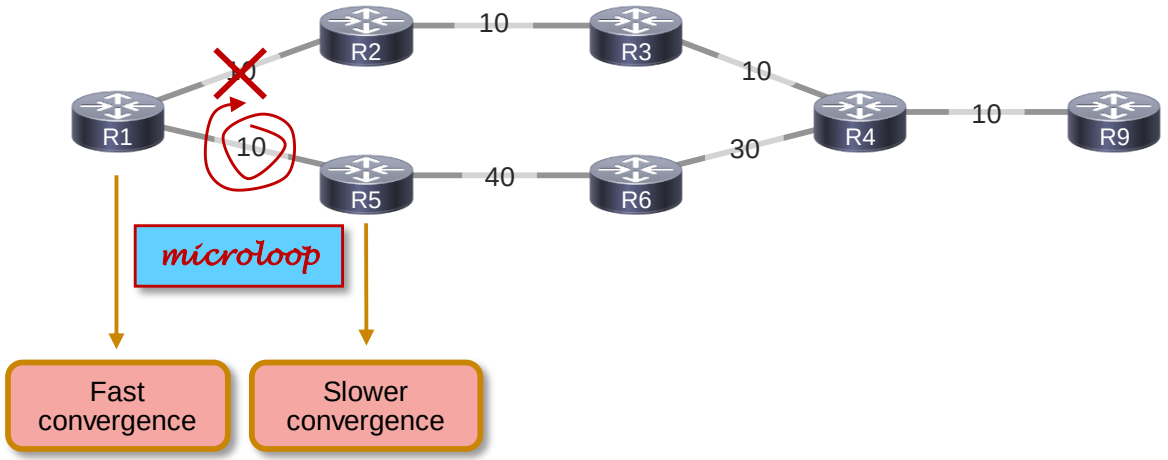
Microloop Avoidance

You make networking **possible**

Cisco live!

Microloop Avoidance

- Microloops =
  - Traffic moved to PC paths can suffer from microloops
  - Microloops are the result of difference of convergence on different routers
  - If faster converged router send traffic to not-yet converged router, the result is a microloop until slower router has converged



Microloop Avoidance: Old and New

- Before Ti-LFA
  - Microloop avoidance
    - Only for local link-down event
    - Microloop avoidance = “use backup path in case of local link-down” + “transition to PC with a delay”
    - Does not address all microloops
    - Enabled automatically when Remote LFA is enabled (not needed when directly connected LFA is available)
- New SR microloop avoidance
  - Local and remote – link-down and link-up events + metric changes
  - Ti-LFA still provides backup for local link-down events
  - Ti-LFA: no backup for local link-up and remote events
    - Remote events
    - Backup for remote events is not wanted/possible
    - At time of learning remote event: compute forced SR path (using SR segments) over new PC
    - Regular convergence (same path as stage 1)

no extra SPF needed

Stage 1

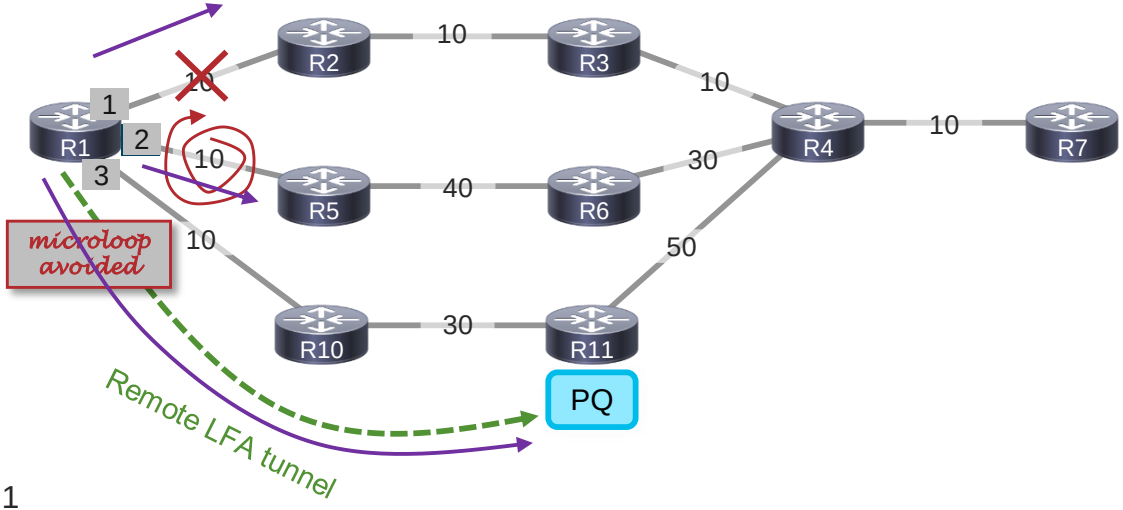
Stage 2

Microloop Avoidance - Old

- Only for local link-down event
- Microloop avoidance = “use backup path in case of local link-down” + “transition to PC with a delay”

No directly connected LFA

Only Remote LFA



Pre-failure: labels {16}, OIF 1

Post failure: labels {19, 200}, OIF 3, non-PC path, encapsulate to PQ node

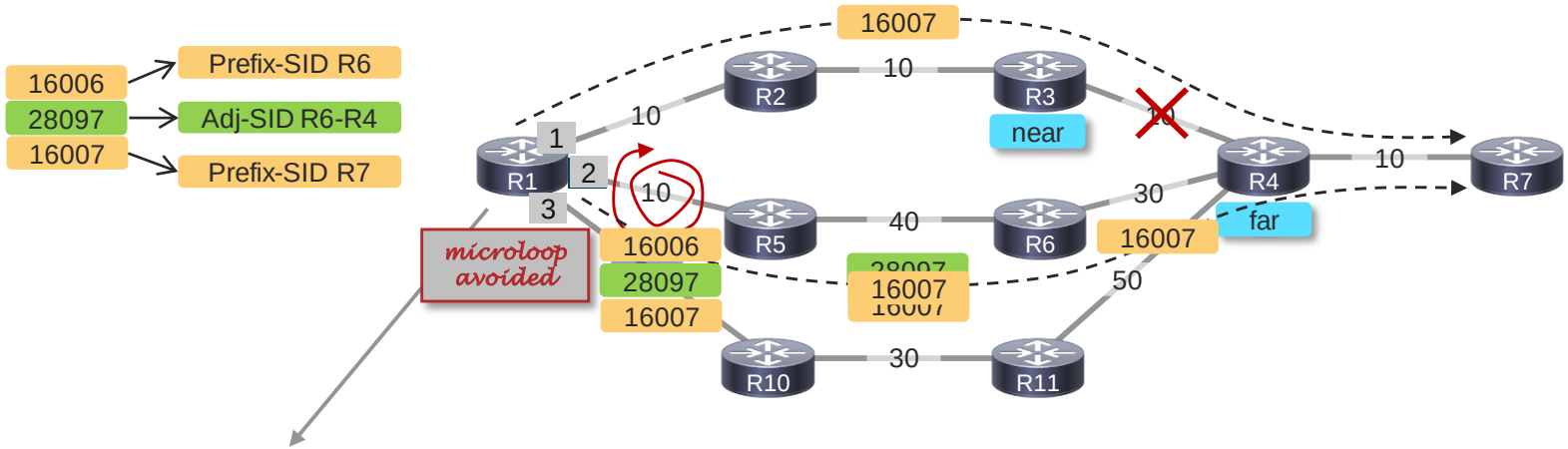
Post failure, after convergence: labels {17}, OIF 2, PC path

Microloop Avoidance – New

Remote link-down event

* Includes link cost increase

- Local and remote – link-down and link-up events
  - Microloop avoidance = {
    - Stage 1: At time of learning remote event: compute forced SR path over new PC
    - Stage 2: Regular convergence (same path as stage 1) }



Pre-failure: labels {16007}, OIF 1

Post failure – stage 1: labels {16006, 28097, 16007}, OIF 2, PC path (forced)

Post failure – stage 2: labels {16007}, OIF 2, PC path

Stage 1

Stage 2

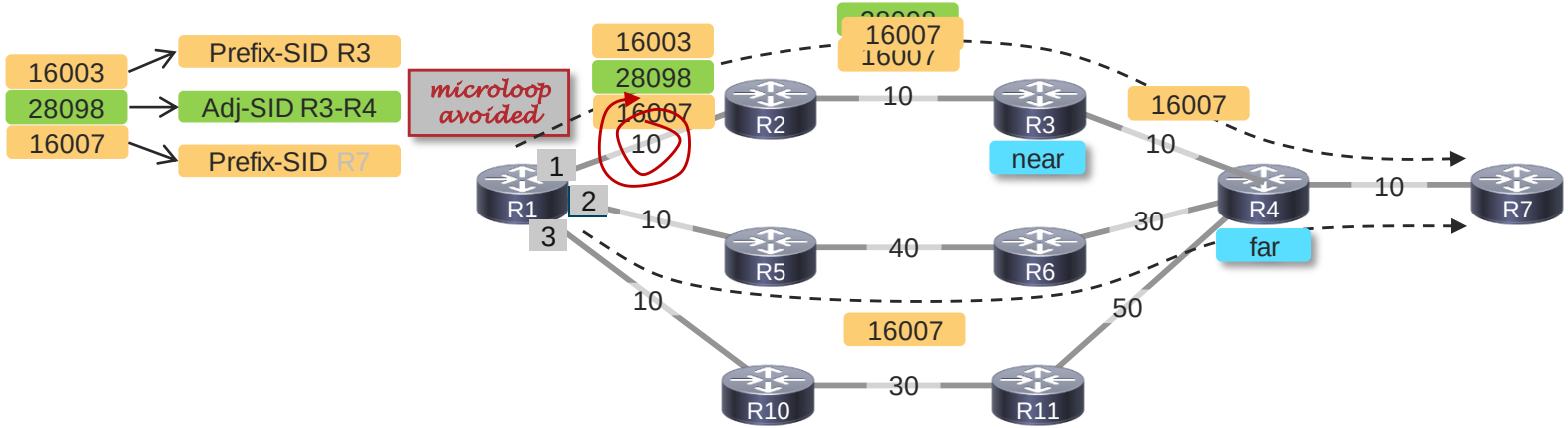
OIF/path did not change

Microloop Avoidance – New

Remote link-up event

* Includes link cost decrease

- Local and remote – link-down and link-up events
  - Microloop avoidance = {
    - Stage 1: At time of learning remote event: compute forced SR path over new PC
    - Stage 2: Regular convergence (same path as stage 1) }



Pre-failure: labels {16007}, OIF 2

Post failure – stage 1: labels {16003, 28098, 16007}, OIF 1, PC path (forced)

Post failure – stage 2: labels {16007}, OIF 1, PC path

Stage 1

Stage 2

→ OIF/path did not change

Microloop Avoidance CLI

- Old CLI

  microloop avoidance [protected]

  → 'protected' only for non-SR microloop avoidance

  → 'protected' = microloop avoidance only applies to prefixes with a valid backup path

- New CLI

  microloop avoidance segment-routing

  → either local or SR microloop avoidance is enabled

- Existing command for RIB update delay is used for both local and SR microloop avoidance feature

  microloop avoidance rib-update-delay <delay>

  → Default 5 seconds

  - local microloop : the time to keep the traffic on a backup-path
  - SR microloop : the time to enforce the traffic on a PC path

```
ipv4 unnumbered mpls traffic-eng Loopback0

mpls traffic-eng
 auto-tunnel p2p
  tunnel-id min 1000 max 1999
```

→ must have automatic instantiation of SR policies enabled for stage 1

→ optional

Microloop Avoidance

```
RP/0/0/CPU0:R1# show isis

IS-IS Router: 1
  System Id: 0000.0000.0001
Most recent startup mode: Cold Restart
  Topologies supported by IS-IS:
    IPv4 Unicast
      Wait for RIB redistribution complete
      Level-2
        Metric style (generate/accept): Wide/Wide
        Metric: 10
        ISPF status: Disabled
        Microloop avoidance: Enabled
          Configuration: Type: Segment routing, RIB update delay: 60000 msec
          State: Active, Duration: 4129 ms, Event Link down, Near: R3.00 Far: R4.00
      No protocols redistributed
```

shows near and far router

```
RP/0/0/CPU0:R1# show isis spf-log
…
16:22:32.323  FSPF     0     7     2             R3.00-00 LINKBAD PREFIXBAD
…
16:28:24.729  FSPF     0     7     2             R3.00-00 LINKGOOD PREFIX


RP/0/0/CPU0:R1# show isis spf-log detail | include SR uloop

  SR uloop:               Link Down
  SR uloop:               No
  SR uloop:               Link Up
```

Microloop Avoidance Active

```
RP/0/0/CPU0:R1# show cef 10.100.1.7

10.100.1.7/32, version 1773, labeled SR, …
Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 0.0.0.0/32, tunnel-te1004, 9 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 NHID 0x0 [0xa1641364 0x0]
    next hop 0.0.0.0/32
    local adjacency
     local label 16007      labels imposed {16007}
```

Original (destination) label

```
RP/0/0/CPU0:R1# show mpls traffic tunnel 1004

Name: tunnel-te1004  Destination: 10.100.1.4  Ifhandle:0x570 (auto-tunnel for OSPF 1)
  Signalled-Name: auto_R1_t1004
  Status:
    Admin:    up Oper:   up   Path:  valid   Signalling: connected
    path option 10, (verbatim Segment-Routing) type explicit (_te1004) (Basis for Setup)

  Segment-Routing Path Info (IGP information is not used)
    Segment0[First Hop]: 10.1.12.2, Label: -
    Segment1[ - ]: Label: 16003
    Segment2[ - ]: Label: 28098
```

SR added stack for near/far router

Microloop Avoidance Debug

```
debug isis segment-routing microloop-avoidance (detail)

isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: Uloop ON, start evaluation
isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: Evaluate added link R3.00-R4 (metric 10)
isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: Unusable R4.00: participant Y overloaded N metric 4294967295 link N
isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: Unusable link R3.00-R4.00. Event Up
isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: Evaluated added link R3.00-R4. NULL-NULL-NULL. Links 0 (metric 0)

…
isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: SPF starts with SR ULOOP. Link R3.00-R4.00, UP
…
isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: Node R4.00 is p2p far end. Add P/Q nodes R3.00/R4.00
…
te_control[1049]: %ROUTING-MPLS_TE-5-LSP_UPDOWN : tunnel-te1010 (signalled-name: auto_R1_t1010, LSP Id: 2) state changed to up
isis[1010]: IPv4 Unicast SR-ULOOP PFX 10.100.1.4/32: 1 path(s) updated with tt1010. Pfx downloadble
isis[1010]: IPv4 Unicast SR-ULOOP PFX 10.100.1.7/32: 1 path(s) updated with tt1010. Pfx downloadble
…

isis[1010]: IPv4 Unicast SR-ULOOP PFX 10.100.1.7/32: Download explicit paths
isis[1010]: IPv4 Unicast SR-ULOOP PFX 10.100.1.4/32: Download explicit paths
…
isis[1010]: Standard (IPv4 Unicast) L2 FSPF SR-ULOOP: SPF starts without SR ULOOP
…
```

SR PCE

You make networking **possible**

Cisco live!

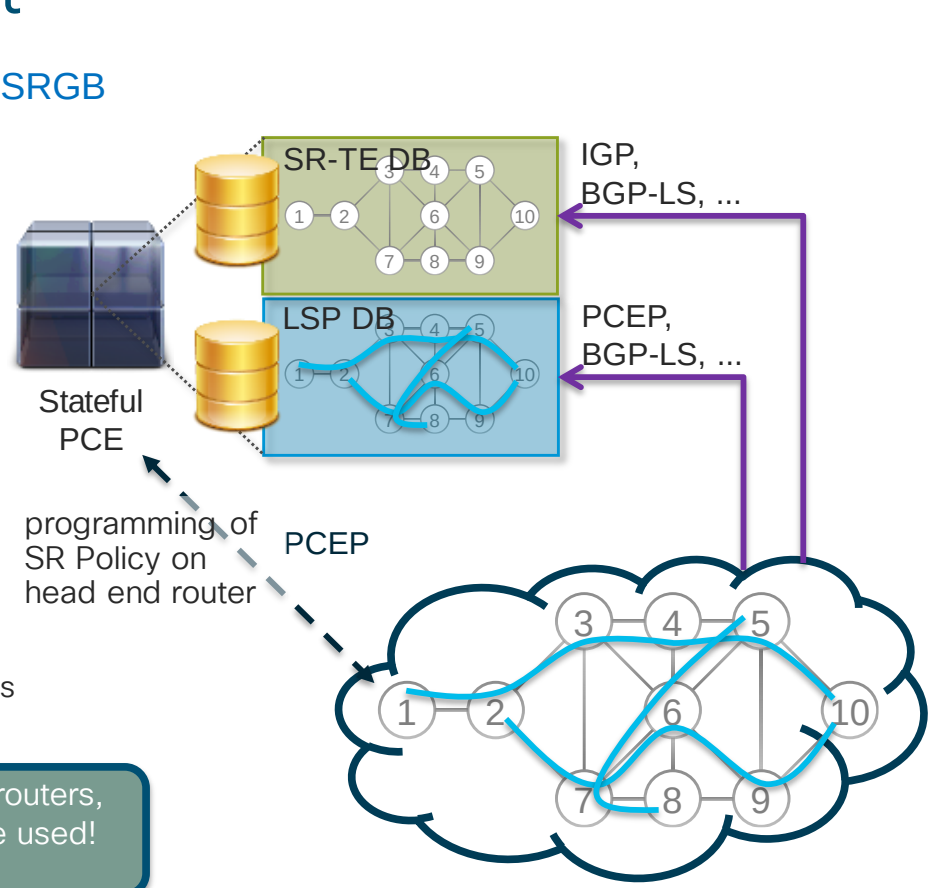
Path Computation Element

SR-TE DB = IGP Topology + TE link attributes + SIDs + SRGB

- PCE learns directly by participating in IGP or from BGP-LS

- PCE needed for inter-area, inter-AS
- Head end needs topology of all area's, domains

- PCCs need IP connectivity to PCE(s)
  - Important when doing inter-area/inter-AS
  - It's about getting all of the needed topology up to the PCE
- Direct vs BGP-LS
  - direct = distribute IGP-LS info into BGP on PCE
  - BGP-LS = PCE has BGP-LS real-time feeds from key routers (usually RRs) in each area/AS/domain



SR-TE needed on all intermediate routers, only if non-default TE attributes are used! No RSVP!

PCEP

1 PCRequest

Requesting

End points constraints:

BW

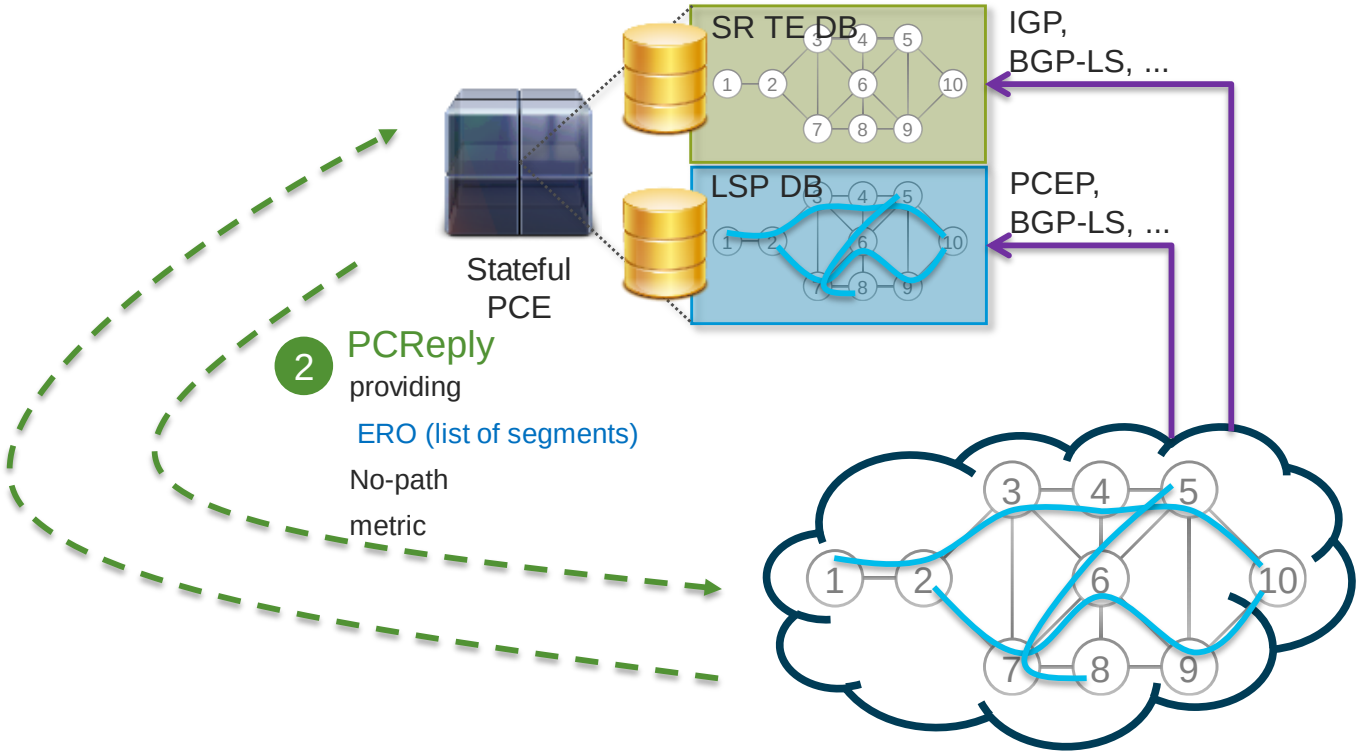
Metric (IGP, TE, hop count)

Affinities

Priorities

RRO

…



SR PCE

```
pce
 address ipv4 10.0.0.11
 segment-routing
  traffic-eng
   peer ipv4 10.0.0.1
   !
…
router ospf 1
 distribute link-state instance-id 33
```

SR PCE is IOS-XR router!

Local LS

There is no need for BGP-LS sessions

```
pce
 address ipv4 10.0.0.11
 segment-routing
  traffic-eng
   peer ipv4 10.0.0.1
   !
…
router ospf 1
 distribute link-state instance-id 33
```

Remote LS: through BGP-LS peering

Instance-id needed for multi-domain (one instance-id per domain)

```
router bgp 65000
 address-family ipv4 unicast
 !
 address-family link-state link-state
 !
 neighbor 10.100.1.9
  address-family link-state link-state
```

There are BGP-LS sessions

Path Computation Client (PCC)

```
segment-routing
 global-block 16000 23999
 traffic-eng

 pcc
   pce address ipv4 10.0.0.11
    precedence 0
   !
   pce address ipv4 10.0.0.12
    precedence 10
```

SR-TE is used

Lower precedence is more preferred PCE

SR Policy on PCC

```
segment-routing
 global-block 16000 23999
 traffic-eng
 !
 policy policy-1
   color 1000 end-point ipv4 10.0.0.2
   candidate-paths
    preference 100
     dynamic
      pcep
      !
      metric
       type igp
!
```

Path computed by PCE

No constraints in this example

- SR-TE only needed when non-default TE attributes are used
- No RSVP!

PCEP Session Verification - PCE

```
RP/0/0/CPU0:PCE# show pce ipv4 peer

PCE's peer database:
--------------------
Peer address: 10.100.1.1
  State: Up
  Capabilities: Stateful, Segment-Routing, Update

Peer address: 10.100.1.9
  State: Up
  Capabilities: Stateful, Segment-Routing, Update

Peer address: 10.100.1.12
  State: Up
  Capabilities: Stateful, Segment-Routing, Update
```

```
RP/0/0/CPU0:PCE# show pce ipv4 peer detail

PCE's peer database:
--------------------
Peer address: 10.100.1.1
  State: Up
  Capabilities: Stateful, Segment-Routing, Update
  PCEP has been up for: 02:06:56
  PCEP session ID: local 0, remote 0
  Sending KA every 30 seconds
  Minimum acceptable KA interval: 20 seconds
  Peer timeout after 120 seconds
  Statistics:
    Keepalive messages: rx      2 tx      2
    Request messages:   rx    810 tx      0
    Reply messages:     rx      0 tx    810
    Error messages:     rx      0 tx      0
    Open messages:      rx      1 tx      1
    Report messages:    rx      5 tx      0
    Update messages:    rx      0 tx      0
    Initiate messages:  rx      0 tx      0
  Last PCError:
    Received: None
    Sent: None

Peer address: 10.100.1.9
...
```

PCEP Session Verification – PCC

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng pcc ipv4 peer

PCC's peer database:
--------------------

Peer address: 10.0.0.11, Precedence: 0, (best PCE)
  State up
  Capabilities: Stateful, Update, Segment-Routing, Instantiation

Peer address: 10.0.0.12, Precedence: 10
  State TCP pending
  Capabilities: Stateless
```

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng pcc ipv4 peer detail

PCC's peer database:
--------------------

Peer address: 10.0.0.11 (best PCE)
  State up
  Capabilities: Stateful, Update, Segment-Routing, Instantiation
  PCEP has been up for: 00:30:01
  Local keepalive timer is 30 seconds
  Remote keepalive timer is 30 seconds
  Local dead timer is 120 seconds
  Remote dead timer is 120 seconds
  Statistics:
    Open messages:       rx 1        |  tx 1
    Close messages:      rx 0        |  tx 0
    Keepalive messages:  rx 61       |  tx 61
    Error messages:      rx 0        |  tx 0
    Report messages:     rx 0        |  tx 0
    Update messages:     rx 0        |  tx 0
```

Notice the difference in command for checking the peering if done on PCE or PCC!

Verify Topology on PCE

```
RP/0/RP0/CPU0:PCE# show pce ipv4 topology isis summary

PCE's topology database summary:
--------------------------------
Showing summary data for ISIS

Topology nodes:                  7
Prefixes:                       14
Prefix SIDs:
  Total:                        14
  Regular:                      14
  Strict:                        0
Links:
  Total:                        28
  EPE:                           0
Adjacency SIDs:
  Total:                        28
  Unprotected:                  28
  Protected:                     0
  EPE:                           0

Private Information:
Lookup Nodes                    14
Consistent                     yes

Update Stats (from IGP and/or BGP):
  Noded added:                  26
  Noded deleted:                 0
  Links added:                  54
  Links deleted:                 0
  Prefix added:                121
  Prefix deleted:                0
```

Verify presence of all:

- Nodes
- Prefix-SIDs
- Adj-SIDs

If BGP is used: verify sessions

- Check the session status and # of prefixes

```
RP/0/0/CPU0:PCE# show bgp link-state link-state summary

Process       RcvTblVer   bRIB/RIB   LabelVer  ImportVer  SendTblVer  StandbyVer
Speaker             332        332        332        332         332           0

Neighbor        Spk    AS MsgRcvd MsgSent   TblVer  InQ OutQ  Up/Down  St/PfxRcd
10.100.1.9        0 65000      92       75      332    0    0 00:17:16        164
10.100.1.10       0 65000      89       54      332    0    0 00:11:25        164
10.100.1.12       0 65000      89       54      332    0    0 00:11:23        164
10.100.1.13       0 65000      73        4      332    0    0 00:00:08        164
```

PCE Topology Checking

```
RP/0/0/CPU0:PCE# show pce ipv4 topology | utility egrep -A5 -B5 10.100.1.13

  Link[2]: local address 10.1.113.10, remote address 10.1.113.13
    Local node:
      OSPF router ID: 10.100.1.10 area ID: 0
    Remote node:
      OSPF router ID: 10.100.1.13 area ID: 0
    Metric: IGP 1, TE 1
    Bandwidth: Total link 125000000, Reservable 0
    Adj SID: 24002 (protected) 24003 (unprotected)
    Excluded from CSPF: no

..
  Link[0]: local address 10.1.123.12, remote address 10.1.123.13
    Local node:
      OSPF router ID: 10.100.1.12 area ID: 0
    Remote node:
      OSPF router ID: 10.100.1.13 area ID: 0
    Metric: IGP 1, TE 1
    Bandwidth: Total link 125000000, Reservable 0
    Adj SID: 24002 (protected) 24003 (unprotected)
    Excluded from CSPF: no

…
```

Nodes/links/IP addresses should be in IPv4 topology on PCE

PCE learns this from BGP-LS sessions or direct from IGP

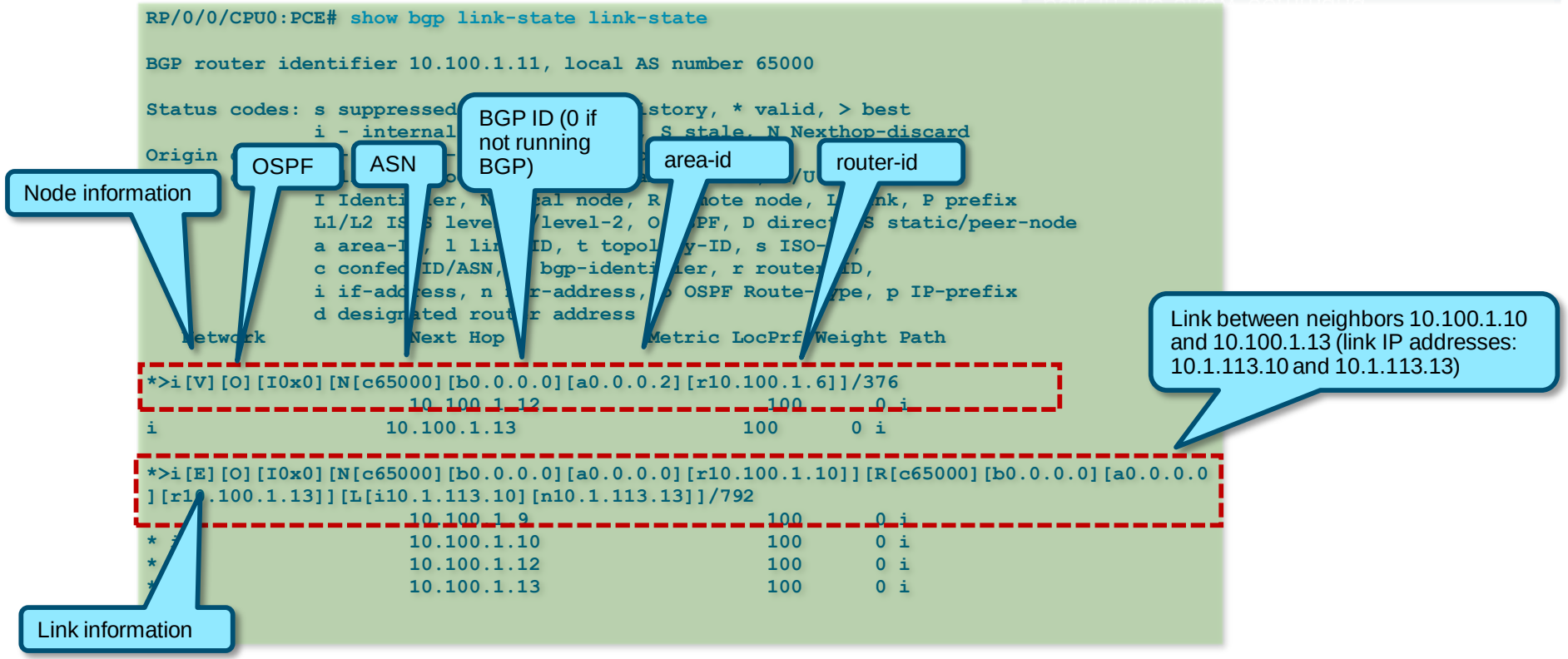
Debugging on the PCE

```
debug pce cspf

pce_server[1114]: DBG-CSPF:pce_cspf_request_path_sr:5770 CSPF Path Request. lsp-name: PE1_t2, type: SR-TE ,
src: 10.100.1.1 , dst: 10.100.1.3, SP:0, Protected
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.17.1---10.1.17.7
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.12.1---10.1.12.2
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.15.1---10.1.15.5
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.107.7---10.1.107.10
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.59.5---10.1.59.9
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.113.10---10.1.113.13
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.119.9---10.1.119.11
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.129.9---10.1.129.12
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.138.13---10.1.138.8
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.124.12---10.1.124.14
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.126.12---10.1.126.6
pce_server[1114]: DBG-CSPF:dijkstra_dest:2039 Link 10.1.48.8---10.1.48.4
pce_server[1114]: DBG-Error:shortest_single_path:4300 No path found.
pce_server[1114]: DBG-CSPF:pce_cspf_request_path_sr:5851 Computation Status: No path found
pce_server[1114]: DBG-CSPF:pce_cspf_request_path_sr:5876 No Path Found
```

debug pce error

BGP LS: Nodes and Links

You can copy/paste the complete NLRI part in the show command



BGP LS: Prefixes

You can copy/paste the complete NLRI part in the show command

```
RP/0/0/CPU0:PCE# show bgp link-state link-state

BGP router identifier 10.100.1.11, local AS number 65000

Status codes: s suppressed,               tory, * valid, > best
              i - internal,               S stale, N Nexthop-discard
Origin codes: i - IGP, e - E              ete
           es: E link, V node             le route, u/U unknown
                 fier, N                                L              x
                                  evel-2, O                 ect,          r-node
                 area-I              , t topolo       ID, s ISO-ID,
              c confed-    /ASN,     bgp-identif    r, r router-ID,
              i if-addr  ss, n n    -address,    OSPF Route-typ    p IP-prefix
              d design   ted rout   r address
   Network                ext Hop            Metric LocPrf Weight Path
*>i[T][O][I0x0][N[c65000][b0.0.0.0][a0.0.0.1][r10.100.1.9]][P[o0x02][p10.100.1.6/32]]/488
                      10.100.1.9                        100       0 i
* i                   10.100.1.10                       100       0 i
```

IP reachable route

OSPF

ASN

BGP ID (0 if not running BGP)

area-id 1

router-id

Prefix 10.100.1.6/32

OSPF route type 2

BGP LS: Node in Detail

```
RP/0/0/CPU0:PCE# show bgp link-state link-state [V][O][I0x0][N[c65000][b0.0.0.0$
BGP routing table entry for [V][O][I0x0][N[c65000][b0.0.0.0][a0.0.0.2][r10.100.1.6]]/376
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker               234         234
Paths: (2 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  Local
    10.100.1.12 (metric 3) from 10.100.1.12 (10.100.1.12)
      Origin IGP, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 0, version 234
      Link-state: Local TE Router-ID: 10.100.1.6, SRGB: 16000:8000
                  SR-ALG: 0 SR-ALG: 1
```

SR-ALG: 1

SRGB: 16000 (+8000)

```
RP/0/0/CPU0:ABR4# show ospf database opaque-
area adv 10.100.1.10
…
  LS age: 79
  Options: (No TOS-capability, DC)
  LS Type: Opaque Area Link
  Link State ID: 4.0.0.0
  Opaque Type: 4
  Opaque ID: 0
  Advertising Router: 10.100.1.6
  …
    Router Information TLV: Length: 4
    Capabilities:
      Graceful Restart Helper Capable
      Stub Router Capable
      Traffic Engineering enabled area
      All capability bits: 0x70000000

    Segment Routing Algorithm TLV: Length: 2
      Algorithm: 0
      Algorithm: 1

    Segment Routing Range TLV: Length: 12
      Range Size: 8000

        SID sub-TLV: Length 3
         Label: 16000
```

BGP LS: Link in Detail

```
RP/0/0/CPU0:PCE# show bgp link-state link-state
[E][O][I0x0][N[c65000][b0.0.0.0][a0.0.0.0][r10.100.1.10]][R[c65000][b0.0.0.0][
a0.0.0.0][r10.100.1.13]][L[i10.1.113.10][n10.1.113.13]]/792

BGP routing table entry for
[E][O][I0x0][N[c65000][b0.0.0.0][a0.0.0.0][r10.100.1.10]][R[c65000][b0.0.0.0][a0.0.0.0][
r10.100.1.13]][L[i10.1.113.10][n10.1.113.13]]/792
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker               256         256
Paths: (4 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  Local
    10.100.1.9 (metric 2) from 10.100.1.9 (10.100.1.9)
      Origin IGP, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 0, version 256
      Link-state: Local TE Router-ID: 10.100.1.10, admin-group: 0x00000000
                  max-link-bw (kbits/sec): 1000000, max-reserv-link-bw (kbits/sec): 0
                  max-unreserv-link-bw (kbits/sec): 0 0 0 0 0 0 0 0,
                 TE-default-metric: 1 metric: 1, ADJ-SID: 24002(e0) ,
                 ADJ-SID: 24003(60)
```

Adj-SID: 24003

Adj-SID: 24002

Adj-SID flags in hex

```
RP/0/0/CPU0:ABR4# show ospf database
opaque-area adv 10.100.1.10
…
 LS age: 1921
  Options: (No TOS-capability, DC)
  LS Type: Opaque Area Link
  Link State ID: 8.0.0.3
  Opaque Type: 8
  Opaque ID: 3
  Advertising Router: 10.100.1.10
  LS Seq Number: 80000003
  Checksum: 0x4d68
  Length: 68

    Extended Link TLV: Length: 44
      Link-type : 1
      Link ID   : 10.100.1.13
      Link Data : 10.1.113.10

     Adj sub-TLV: Length: 7
        Flags     : 0xe0
        MTID      : 0
        Weight    : 0
        Label     : 24002

     Adj sub-TLV: Length: 7
        Flags     : 0x60
        MTID      : 0
        Weight    : 0
        Label     : 24003

     Remote If Address sub-TLV: Length: 4
        Neighbor Address: 10.1.113.13
```

BGP LS: Prefix in Detail

```
RP/0/0/CPU0:PCE# show bgp link-state link-state
[T][O][I0x0][N[c65000][b0.0.0.0][a0.0.0.1][r10.100.1.10]][P[o0x02][p10.100.1.6/32]]/488

BGP routing table entry for
[T][O][I0x0][N[c65000][b0.0.0.0][a0.0.0.1][r10.100.1.10]][P[o0x02][p10.100.1.6/32]]/488
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker                 84          84
Last Modified: Nov 14 16:04:31.364 for 00:30:18
Paths: (2 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  Local
    10.100.1.9 (metric 2) from 10.100.1.9 (10.100.1.9)
      Origin IGP, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 0, version 84
      Link-state: Metric: 4, PFX-SID: 6(40/0)
```

Prefix-SID: index 6

```
RP/0/0/CPU0:ABR4# show ospf database opaque-
area adv 10.100.1.10
…
  LS age: 79
  Options: (No TOS-capability, DC)
  LS Type: Opaque Area Link
  Link State ID: 7.0.0.1
  Opaque Type: 7
  Opaque ID: 1
  Advertising Router: 10.100.1.6
  LS Seq Number: 80000003
  Checksum: 0x2c80
  Length: 44

    Extended Prefix TLV: Length: 20
      Route-type: 1
      AF        : 0
      Flags     : 0x40
      Prefix    : 10.100.1.6/32

      SID sub-TLV: Length: 8
        Flags     : 0x0
        MTID      : 0
        Algo      : 0
        SID Index : 6
```

Dynamic PCE SR Policy

```
segment-routing
 global-block 16000 23999
 traffic-eng
  logging
   policy status
  !
 policy to-PE2-PCE
   binding-sid mpls 1234
   color 3000 end-point ipv4 10.0.0.2
   candidate-paths
    preference 100
     dynamic
      pcep
      !
      metric
       type igp
```

Binding-SID

Local label for the SR Policy

```
RP/0/RP0/CPU0:PE1# show mpls forwarding

Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
1000   Pop         No ID              srte_c_1000_ point2point     0
1234   Pop         No ID              srte_c_3000_ point2point     0
2000   Pop         No ID              srte_c_2000_ point2point     1192
16003  Pop         SR Pfx (idx 3)     Gi0/0/0/0    10.1.3.3        0
       Unlabelled  SR Pfx (idx 3)     Gi0/0/0/2    10.1.7.7        0
(!)
16005  16005       SR Pfx (idx 5)     Gi0/0/0/0    10.1.3.3        108346
       Unlabelled  SR Pfx (idx 5)     Gi0/0/0/2    10.1.7.7        0
(!)
24000  Pop         SR Adj (idx 0)     Gi0/0/0/0    10.1.3.3        0
       Unlabelled  SR Adj (idx 0)     Gi0/0/0/2    10.1.7.7        0
(!)
24001  Pop         SR Adj (idx 2)     Gi0/0/0/0    10.1.3.3        0
24002  Unlabelled  10.0.0.13/32[V]    Gi0/0/0/1    10.1.13.13      0
24003  16005       SR TE: 2 [TE-INT]  Gi0/0/0/0    10.1.3.3        3864
24004  Aggregate   one: Per-VRF Aggr[V]   \
                                      one                          14868
24005  16005       SR TE: 4 [TE-INT]  Gi0/0/0/0    10.1.3.3        0
24006  16007       SR TE: 1 [TE-INT]  Gi0/0/0/0    10.1.3.3        0
```

```
RP/0/RP0/CPU0:PE1# show mpls forwarding labels 1234 detail
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
1234   Pop         No ID              srte_c_3000_ point2point     0
     Label Stack (Top -> Bottom): { Unlabelled Imp-Null }
     Outgoing Interface: srte_c_3000_ep_10.0.0.2 (ifhandle 0x00000034)
     Packets Switched: 0
```

```
RP/0/RP0/CPU0:PE1# show mpls forwarding labels 24005 detail
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24005  16005       SR TE: 4 [TE-INT]  Gi0/0/0/0    10.1.3.3        0
     Label Stack (Top -> Bottom): { 16005 16002 }
     Outgoing Interface: GigabitEthernet0/0/0/0 (ifhandle 0x01000018)
     Packets Switched: 0
```

PCC: Debug SR TE Policy LSP

```
RP/0/RP0/CPU0:PE1# debug segment-routing traffic-eng policy lsp

RP/0/RP0/CPU0:config[68253]: %MGBL-CONFIG-6-DB_COMMIT : Configuration committed by user 'cisco'. Use 'show configuration commit changes 1000000036' to view the
changes.
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_create:787  [LSP-ID: 0 POL-ID: 0] LSP created for POL-ID: 4, num LSPs: 3
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_set_ident:1088  [LSP-ID: 1 POL-ID: 4] Set ident
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_set_cpath:1156  [LSP-ID: 1 POL-ID: 4] Set CP-ID: 1
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_set_bsid:1478  [LSP-ID: 1 POL-ID: 4] Set BSID-ID: 6
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_egress_paths_populate:3391  [LSP-ID: 1 POL-ID: 4] Prefix 16005, set new egress path(s)
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_lsd_rewrite_fill_path_moi:442 (Policy: srte_c_3000_ep_10.0.0.2 #4)  [LSP-ID: 1 POL-ID: 4] Fill path moi path flag
0x0, load metric 64, num_of_labels : 2
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_lsd_rewrite_create_lsp_paths:1091  [LSP-ID: 1 POL-ID: 4] create_lsp_paths: SL num_valid_mois=1,
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_post_event:117  [LSP-ID: 1 POL-ID: 4] Current state Init, event posted: Label alloc
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_label_alloc_cb:941  [LSP-ID: 1 POL-ID: 4] Entered FSM callback
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_set_local_label:1227  [LSP-ID: 1 POL-ID: 4] Set local label: 24005
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_post_event:117  [LSP-ID: 1 POL-ID: 4] Current state Label allocation pending, event posted: Label RW
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_label_rewrite_cb:989  [LSP-ID: 1 POL-ID: 4] Entered FSM callback
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_lsd_rewrite_fill_label_fpi_data:1144  [LSP-ID: 1 POL-ID: 4] Filled label (24005) FPI data. Label flags: 0x00000000
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_lsd_rewrite_install_label:1290  [LSP-ID: 1 POL-ID: 4] [LSP-ID: 0 POL-ID: 0] MOI[1] Path Flags 0x1
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_post_event:117 (LSP: 1 T: 4)  [LSP-ID: 1 POL-ID: 4] Current state BSID rewrite pending, event posted:
Start install timer
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_post_event:117  [LSP-ID: 1 POL-ID: 4] Current state Install timer pending, event posted: Tunnel RW
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_tunnel_rewrite_cb:1067  [LSP-ID: 1 POL-ID: 4] Entered FSM callback
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_lsd_rewrite_fill_tunnel_moi:1449  [LSP-ID: 1 POL-ID: 4] Filled tunnel MOI. Label: 24005, Forward class: 0, Flags:
0x00000000, Ext-Flags: 0x00000080
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_sm_programmed_cb:901 (LSP: 1 T: 4)  [LSP-ID: 1 POL-ID: 4] Entered FSM callback
RP/0/RP0/CPU0:xtc_agent[1252]: %OS-XTC-5-SR_POLICY_UPDOWN : SR policy 'srte_c_3000_ep_10.0.0.2' (color 3000, end-point 10.0.0.2) state changed to UP
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_egress_paths_verify:3471 (Policy: srte_c_3000_ep_10.0.0.2 #4)  [LSP-ID: 1 POL-ID: 4] SL egress paths
verify: 16005, path_count=1, sl_changed=0,
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_lsp_egress_paths_verify:3489 (Policy: srte_c_3000_ep_10.0.0.2 #4)  [LSP-ID: 1 POL-ID: 4] LSP egress paths
verify: total=1, invalid=0, changed=0, drsn=0,
RP/0/RP0/CPU0:xtc_agent[1252]: DBG-Policy-LSP:xtc_policy_verify:2451  [LSP-ID: 1 POL-ID: 4] Verification indicated no path change
```

PCE: Debug PCE CSPF

```
RP/0/RP0/CPU0:PCE# debug pce cspf

RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_request_path_sr:8701 CSPF Path Request. lsp-name:
cfg_to-PE2-PCE_discr_100, type: SR , src: 10.0.0.1 , dst: 10.0.0.2, SP:0, Protected, Metric: IGP, msd: 10, Anycast:
Disabled
RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_should_use_te_latency_algo:8636 TE-LAT not required
RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_init_cache:5420 Cache is clear
RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_request_path_sr:8854 Computation Status: Short Path
Success
RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_request_path_sr:8859 Returned Path:
RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_print_sr_hop_list:7939 SR Path Hop = 16005 (10.0.0.5)
(unicast)
RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_print_sr_hop_list:7939 SR Path Hop = 16002 (10.0.0.2)
(unicast)
RP/0/RP0/CPU0:pce_server[1122]: DBG-CSPF:[140231257978624] pce_cspf_request_path_sr:8861 Returned Path Cost = 50
```

PCC: Verify SR-TE Policy

```
RP/0/RP0/CPU0:PE1# show segment-routing traffic-eng policy color 3000

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

Color: 3000, End-point: 10.0.0.2
  Name: srte_c_3000_ep_10.0.0.2
  Status:
    Admin: up  Operational: up for 00:08:04 (since Apr 30 13:34:51.041)
  Candidate-paths:
    Preference: 100 (configuration) (active)
      Name: to-PE2-PCE
      Requested BSID: 1234
      PCC info:
        Symbolic name: cfg_to-PE2-PCE_discr_100
        PLSP-ID: 2
        Maximum SID Depth: 10
      Dynamic (pce 10.0.0.11) (valid)
        Metric Type: IGP,   Path Accumulated Metric: 50
          16005 [Prefix-SID, 10.0.0.5]
          16002 [Prefix-SID, 10.0.0.2]
  Attributes:
    Binding SID: 1234
    Forward Class: 0
    Steering labeled-services disabled: no
    Steering BGP disabled: no
    IPv6 caps enable: yes
```

Binding SID

PCE must have LS database from all area's
e.g. multi-area link from ABR router to PCE

PCE: Verify LSP Paths

Verify SR-TE policies on PCE

```
RP/0/RP0/CPU0:PCE# show pce lsp pcc ipv4 10.0.0.1 detail

PCE's tunnel database:
----------------------
PCC 10.0.0.1:

Tunnel Name: cfg_to-PE2-PCE_discr_100
 LSPs:
  LSP[0]:
   source 10.0.0.1, destination 10.0.0.2, tunnel ID 4, LSP ID 1
   State: Admin up, Operation up
   Setup type: Segment Routing
   Binding SID: 1234
   Maximum SID Depth: 10
   Bandwidth: signaled 0 kbps, applied 0 kbps
   PCEP information:
     PLSP-ID 0x2, flags: D:1 S:0 R:0 A:1 O:1 C:0
   LSP Role: Single LSP
   State-sync PCE: None
   PCC: 10.0.0.1
   LSP is subdelegated to: None
   Reported path:
     Metric type: IGP, Accumulated Metric 50
       SID[0]: Node, Label 16005, Address 10.0.0.5
       SID[1]: Node, Label 16002, Address 10.0.0.2
   Computed path: (Local PCE)
     Computed Time: Tue Apr 30 13:34:51 UTC 2019 (00:0  2 ago)
     Metric type: IGP, Accumulated Metric 50
       SID[0]: Node, Label 16005, Address 10.0.0.5
       SID[1]: Node, Label 16002, Address 10.0.0.2
```

Filter on head end router

minimum set of labels from source to destination

This command can be used on PCE for any source and destination and provides the path
No LSP/policy needs to be actually requested/present!

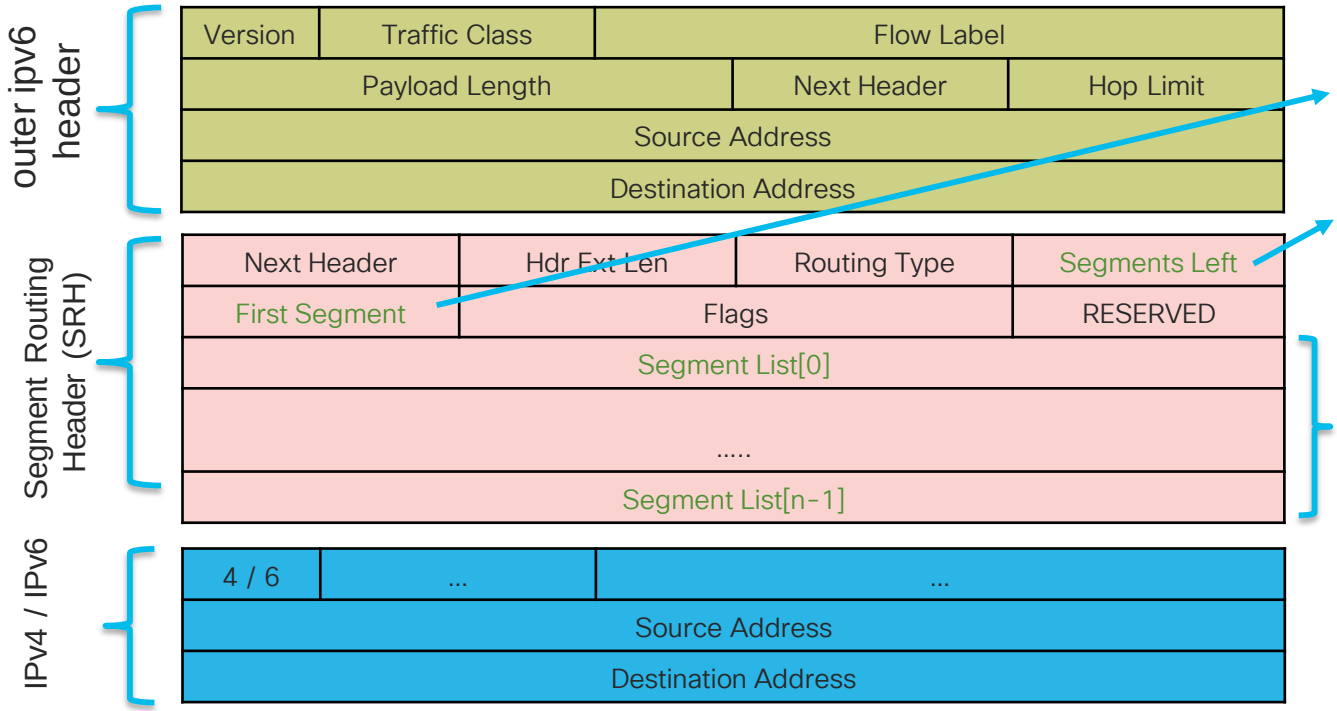
```
RP/0/RP0/CPU0:PCE# show pce ipv4 path source
10.0.0.1 destination 10.0.0.2

Path:
----:
Hop0: 10.1.3.1
Hop1: 10.3.5.3
Hop2: 10.5.6.5
Hop3: 10.4.6.6
Hop4: 10.2.4.4
```

SRv6

You make networking **possible**

Cisco live!

SRv6 Encap Packet



First Segment: offset in the SRH, not including the first 8 octets and expressed in 16-octet units, pointing to the last element of the Segment List (i.e.: that contains the first segment of the path).

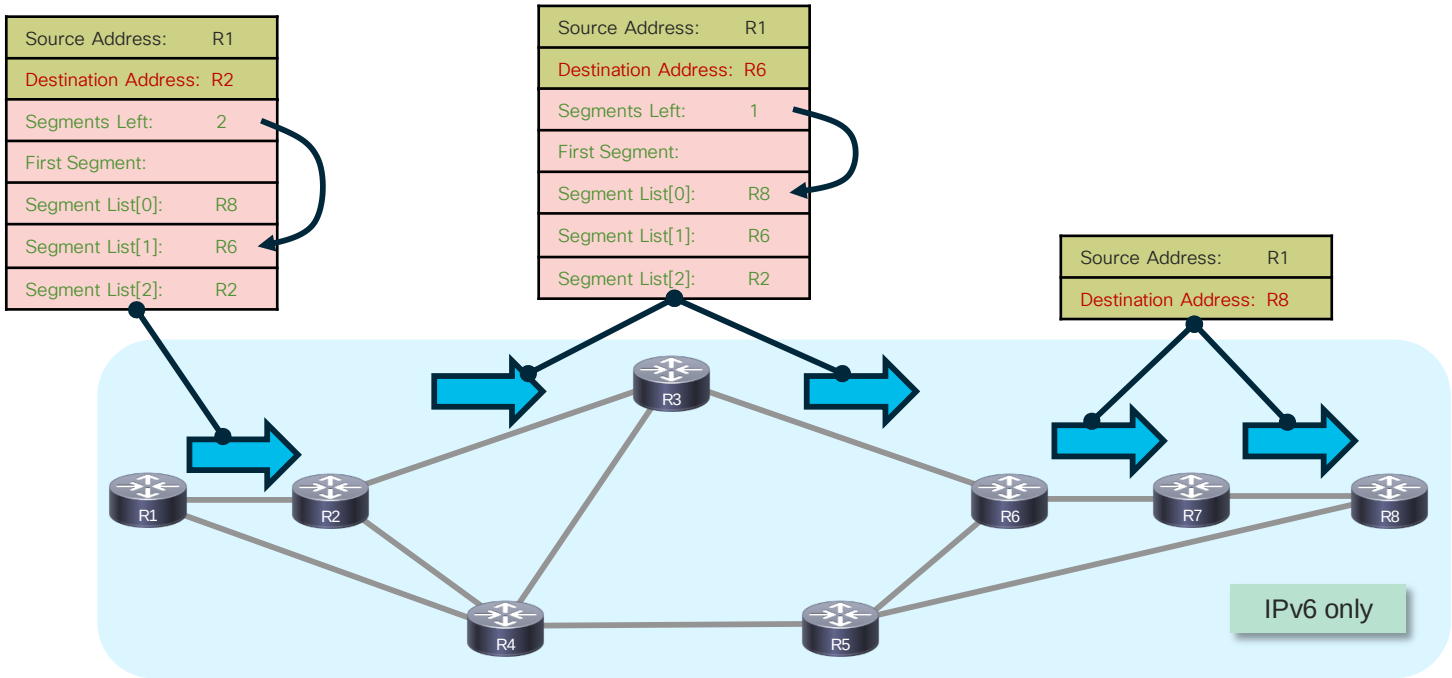
Segments Left: index, in the Segment List, of the current active segment in the SRH. Decremented at each segment endpoint.

SRH: Segment List[n]: 128 bit IPv6 addresses representing each segment of the path. The segment list is encoded in the reverse order of the path: the last segment is in the first position of the list and the first segment is in the last position.

SRv6 Forwarding

Forwarding rule:

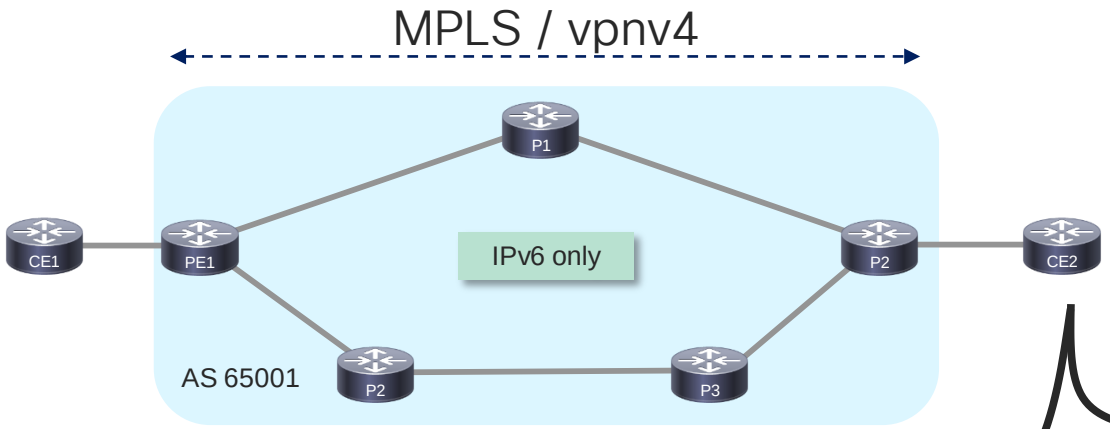
if DA is myself, and if Segments Left > 0, then update DA with segment list[Segments Left] and decrement Segments Left



Sid Functions

Function	Meaning	Comments
End	Endpoint function	SRv6 instantiation of a prefix SID
End.X	Endpoint with Layer-3 cross-connect	SRv6 instantiation of a Adj SID
End.DX4	Endpoint with decaps and IPv4 cross-connect	IPv4-L3VPN (equivalent to per-CE VPN label)
End.DT4	Endpoint with decaps and IPv4 table lookup	IPv4-L3VPN (equivalent to per-VRF VPN label)
End.B6.Insert	Endpoint bound to an SRv6 policy	SRv6 instantiation of a Binding SID
End.B6.Encaps	Endpoint bound to an SRv6 policy with encaps	SRv6 instantiation of a Binding SID
End.DT2U	Endpoint with decaps and unicast MAC L2table lookup	EVPN Bridging unicast use-cases
…		

SRv6 Based L3 vpnv4: Configuration

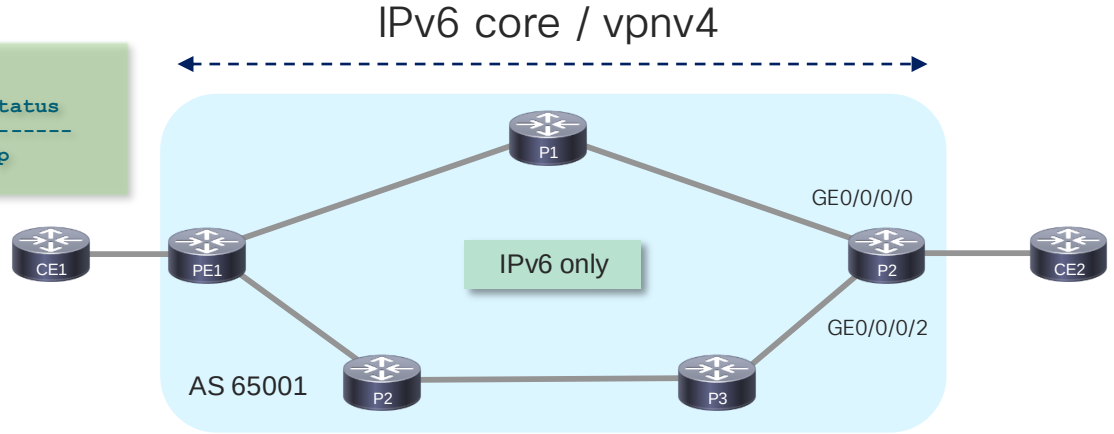



```
router bgp 65001
 bgp router-id 10.100.1.2
 address-family ipv4 unicast
 !
 address-family vpnv4 unicast
  segment-routing srv6
   locator default
  !
 !
!
 neighbor 2001:db8:100::1      # PE1
  remote-as 65001
  update-source Loopback0
  address-family vpnv4 unicast
  !
 !
 vrf one
  address-family ipv4 unicast
   segment-routing srv6
    alloc mode per-ce  # default mode
   !
   redistribute connected
  !
  address-family ipv6 unicast
  !
  neighbor 10.2.7.7       # CE2
   remote-as 65003
   address-family ipv4 unicast
```

```
segment-routing
 srv6
  encapsulation
   source-address 2001:100::2
  !
  locators
   locator default
    prefix 2001:db8:2::/64
```

SRv6 Based L3 vpnv4: Verification

```
RP/0/RP0/CPU0:PE2# show segment-routing srv6 locator

Name                  ID       Prefix                    Status
--------------------  -------  ------------------------  -------
default*              1        2001:db8:2::/64           Up
```




```
RP/0/RP0/CPU0:PE2# show segment-routing srv6 sid all

*** Locator: 'default' ***

SID                      Function     Context                          Owner               State  RW
-----------------------  -----------  -------------------------------  ------------------  -----  --
2001:db8:2:0:1::         End (PSP)    'default':1                      sidmgr              InUse  Y
2001:db8:2:0:11::        End.OP       'default'                        sidmgr              InUse  Y
2001:db8:2:0:40::        End.X (PSP)  [Gi0/0/0/2, Link-Local]          isis-1              InUse  Y
2001:db8:2:0:41::        End.X (PSP)  [Gi0/0/0/0, Link-Local]          isis-1              InUse  Y
2001:db8:2:0:42::        End.DT4      'one'                            bgp-65001           InUse  Y
2001:db8:2:0:43::        End.DX4      'one':2                          bgp-65001           InUse  Y
```

used for IP lookup in VRF context

used for forwarding to destinaton on CE2

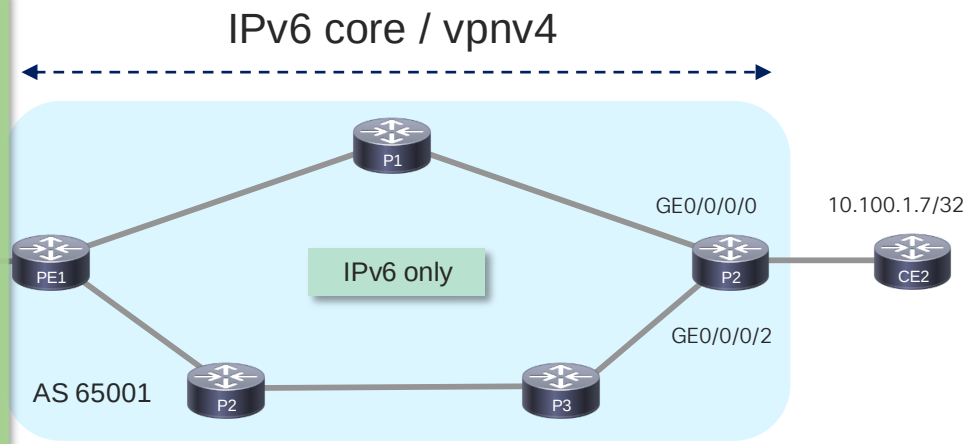
PSP: Penultimate Segment Pop of the SRH
USP: Ultimate Segment Pop of the SRH
USD: Ultimate Segment Decapsulation

SRv6 Based L3 vpnv4: Verification

```
RP/0/RP0/CPU0:PE1# show isis database PE2.00-00 verbose

IS-IS 1 (Level-2) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime/Rcvd  ATT/P/OL
PE2.00-00             0x00000b9b   0xa52b        1132 /1200         0/0/0
  Area Address:   49.0001
  NLPID:          0x8e
  Hostname:       PE2
  IPv6 Address:   2001:db8:100::2
  Metric: 1           MT (IPv6 Unicast) IPv6 2001:db8:2::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 10          MT (IPv6 Unicast) IPv6 2001:db8:2:3::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 10          MT (IPv6 Unicast) IPv6 2001:db8:2:5::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 0           MT (IPv6 Unicast) IPv6 2001:db8:100::2/128
    Prefix Attribute Flags: X:0 R:0 N:1
  Router Cap:     0.0.0.0 D:0 S:0
    IPv6 Router ID: 2001:db8:100::2
    SRv6: O:0
    Node Maximum SID Depth:
      SRH Max SL:       4
      SRH Max End Pop:  4
      SRH Max T.insert: 4
      SRH Max T.encaps: 5
      SRH Max End D:    5
  SRv6 Locator:   MT (IPv6 Unicast) 2001:db8:2::/64 D:0 Metric: 0 Algorithm: 0
    Prefix Attribute Flags: X:0 R:0 N:0
    END SID: 2001:db8:2:0:1:: End (PSP)
  MT:             IPv6 Unicast                                0/0/0
  Metric: 10          MT (IPv6 Unicast) IS-Extended P3.00
    Interface IPv6 Address: 2001:db8:2:5::2
    Neighbor IPv6 Address: 2001:db8:2:5::5
    END.X SID: 2001:db8:2:0:40:: B:0 S:0 P:0 End.X (PSP)
  Metric: 10          MT (IPv6 Unicast) IS-Extended P1.00
    Interface IPv6 Address: 2001:db8:2:3::2
    Neighbor IPv6 Address: 2001:db8:2:3::3
    END.X SID: 2001:db8:2:0:41:: B:0 S:0 P:0 End.X (PSP)
```

link in global context

link in global context



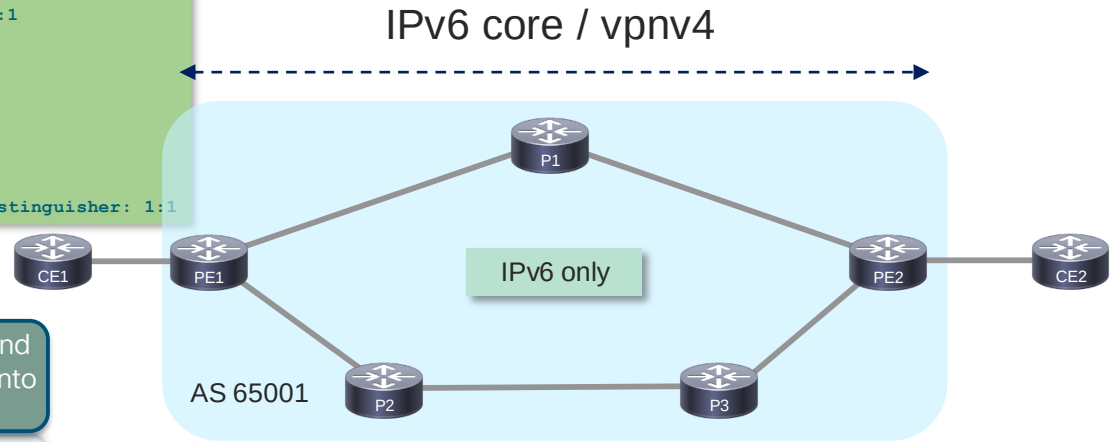
SRv6 Based L3 vpnv4: Verification

```
RP/0/RP0/CPU0:PE1# show bgp vpnv4 unicast vrf one 10.100.1.7/32
BGP routing table entry for 10.100.1.7/32, Route Distinguisher: 1:1
Paths: (1 available, best #1)
 65003
    2001:db8:100::2 (metric 20) from 2001:db8:100::2 (10.100.1.2)
      Received Label 3
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Received Path ID 0, Local Path ID 1, version 34
      Extended community: RT:1:1
      SRv6-VPN-SID: T1-2001:db8:2:0:43:: [total 1]
      Source AFI: VPNv4 Unicast, Source VRF: one, Source Route Distinguisher: 1:1
```

NH = loopback PE2

End.DX4 'one':2

meaning: endpoint with decaps and IPv4 cross-connect -> forward onto link to CE2




```
RP/0/RP0/CPU0:PE1# show cef vrf one 10.100.1.7/32
10.100.1.7/32, version 22, SRv6 Transit
  Prefix Len 32, traffic index 0, precedence n/a, priority 3
   via 2001:db8:2::/128, 3 dependencies, recursive [flags 0x6000]
    path-idx 0 NHID 0x0 [0xe014724 0x0]
    next hop VRF - 'default', table - 0xe0800000
    next hop 2001:db8:2::/128 via 2001:db8:2::/64
     SRv6 T.Encaps.Red SID-list {2001:db8:2:0:43::}
```

End.DX4 'one':2

T	Transit behavior
T.Insert	Transit behavior with insertion of an SRv6 policy
T.Insert.Red	Transit behavior with reduced insert of an SRv6 policy
T.Encaps	Transit behavior with encapsulation in an SRv6 policy
T.Encaps.Red	Transit behavior with reduced encaps in an SRv6 policy
T.Encaps.L2	T.Encaps applied to received L2 frames
T.Encaps.L2.Red	T.Encaps.Red applied to received L2 frames

SRv6 Based L3 vpnv4: Encapsulation

IPv6 header

Transported IPv4 packet

```
Internet Protocol Version 6, Src: 2001:100::1, Dst: 2001:db8:2:0:43::
    0110 .... = Version: 6
  > .... 0000 0000 .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
    .... .... .... 1111 1101 0001 1111 1001 = Flow Label: 0xfd1f9
    Payload Length: 100
    Next Header: IPIP (4)
    Hop Limit: 255
    Source: 2001:100::1
    Destination: 2001:db8:2:0:43::
Internet Protocol Version 4, Src: 10.100.1.6, Dst: 10.100.1.7
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 100
    Identification: 0x0019 (25)
  > Flags: 0x0000
    Time to live: 254
    Protocol: ICMP (1)
    Header checksum: 0xa5ab [validation disabled]
    [Header checksum status: Unverified]
    Source: 10.100.1.6
    Destination: 10.100.1.7
> Internet Control Message Protocol
```

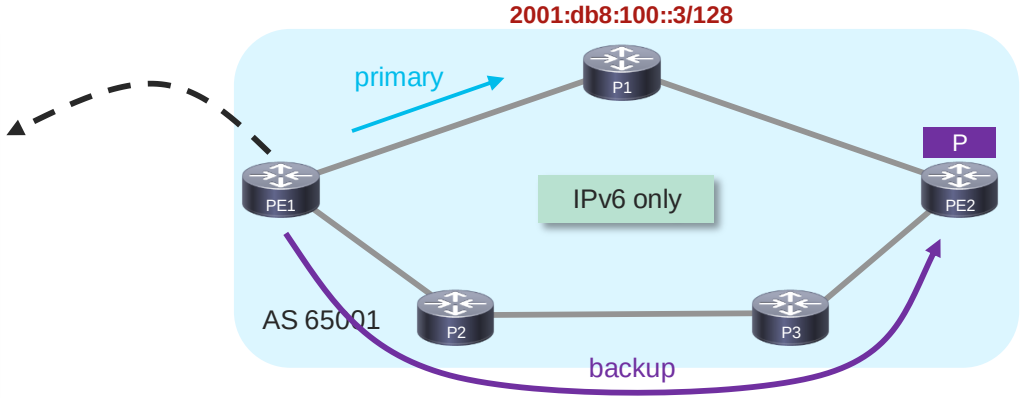
next header : IPinIP

source address: loopback ingress PE

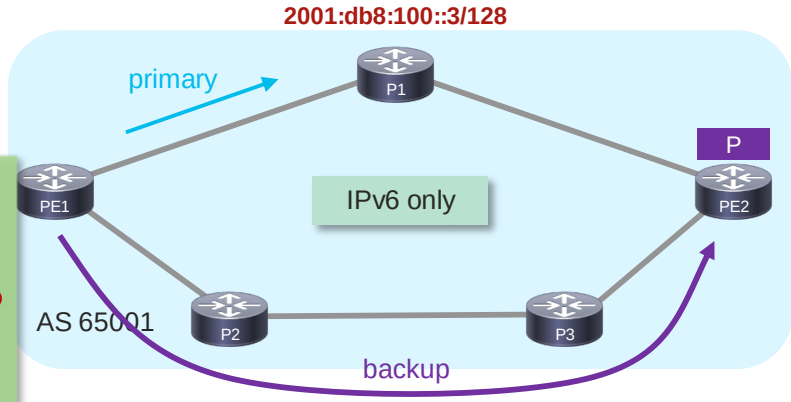
destination address: locator address : function

SRv6 – Ti-LFA

```
router isis 1
 is-type level-2-only
 net 49.0001.0000.0000.0001.00
 address-family ipv6 unicast
  metric-style wide
  segment-routing srv6
   locator default
   !
  !
 !
 interface Loopback0
  passive
  address-family ipv6 unicast
  !
 !
 interface GigabitEthernet0/0/0/1
  point-to-point
  address-family ipv6 unicast
   fast-reroute per-prefix
   fast-reroute per-prefix ti-lfa
  !
 !
!
```



SRv6 – Ti-LFA




```
RP/0/RP0/CPU0:PE1# show route ipv6 2001:db8:100::3/128

Routing entry for 2001:db8:100::3/128
  Known via "isis 1", distance 115, metric 10, type level-2
  Installed May 22 06:02:27.072 for 05:38:36
  Routing Descriptor Blocks
    fe80::f816:3eff:fe5b:4605, from 2001:db8:100::3, via GigabitEthernet0/0/0/2, Backup (TI-LFA)
      Repair Node(s): 2001:db8:100::2
      Route metric is 40
    fe80::f816:3eff:fee4:3e5c, from 2001:db8:100::3, via GigabitEthernet0/0/0/1, Protected
      Route metric is 10
```

```
RP/0/RP0/CPU0:PE1# show isis ipv6 fast-reroute 2001:db8:100::3/128

L2 2001:db8:100::3/128 [10/115]
     via fe80::f816:3eff:fee4:3e5c, GigabitEthernet0/0/0/1, P1, Weight: 0
       Backup path: TI-LFA (link), via fe80::f816:3eff:fe5b:4605, GigabitEthernet0/0/0/2 P2, Weight: 0, Metric: 40
         P node: PE2.00 [2001:db8:100::2], SRv6 SID: 2001:db8:2:0:1:: End (PSP)
         Backup-src: P1.00
```

End (PSP)

```
RP/0/RP0/CPU0:PE1# show cef ipv6 2001:db8:100::3/128
2001:db8:100::3/128, version 282, SRv6 Transit, internal 0x1000001 0x2 (ptr 0xe014c34) [1], 0x0 (0xe1ea6e8), 0x0 (0xf03c0a8)
 remote adjacency to GigabitEthernet0/0/0/1
 Prefix Len 128, traffic index 0, precedence n/a, priority 1
   via fe80::f816:3eff:fe5b:4605/128, GigabitEthernet0/0/0/2, 11 dependencies, weight 0, class 0, backup (TI-LFA) [flags 0xb00]
    path-idx 0 NHID 0x0 [0xd949738 0x0]
    next hop fe80::f816:3eff:fe5b:4605/128, Repair Node(s): 2001:db8:100::2
    remote adjacency
     SRv6 T.Insert.Red SID-list {2001:db8:2:0:1::}
   via fe80::f816:3eff:fee4:3e5c/128, GigabitEthernet0/0/0/1, 11 dependencies, weight 0, class 0, protected [flags 0x400]
    path-idx 1 bkup-idx 0 NHID 0x0 [0xf081190 0x0]
    next hop fe80::f816:3eff:fee4:3e5c/128
```

SRv6 – Ti-LFA

```
RP/0/RP0/CPU0:PE1# show isis database PE2.00-00 verbose

IS-IS 1 (Level-2) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime/Rcvd  ATT/P/OL
PE2.00-00             0x00000eb2   0x6e48        562  /1200         0/0/0
  Area Address:   49.0001
  NLPID:          0x8e
  Hostname:       PE2
  IPv6 Address:   2001:db8:100::2
  Metric: 1          MT (IPv6 Unicast) IPv6 2001:db8:2::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 10         MT (IPv6 Unicast) IPv6 2001:db8:2:3::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 10         MT (IPv6 Unicast) IPv6 2001:db8:2:5::/64
    Prefix Attribute Flags: X:0 R:0 N:0
  Metric: 0          MT (IPv6 Unicast) IPv6 2001:db8:100::2/128
    Prefix Attribute Flags: X:0 R:0 N:1
  Router Cap:     0.0.0.0 D:0 S:0
    IPv6 Router ID: 2001:db8:100::2
    SRv6: O:0
    Node Maximum SID Depth:
      SRH Max SL:       4
      SRH Max End Pop:  4
      SRH Max T.insert: 4
      SRH Max T.encaps: 5
      SRH Max End D:    5
  SRv6 Locator:   MT (IPv6 Unicast) 2001:db8:2::/64 D:0 Metric: 0 Algorithm: 0
    Prefix Attribute Flags: X:0 R:0 N:0
    END SID: 2001:db8:2:0:1:: End (PSP)
  MT:             IPv6 Unicast                                  0/0/0
  Metric: 10         MT (IPv6 Unicast) IS-Extended P3.00
    Interface IPv6 Address: 2001:db8:2:5::2
    Neighbor IPv6 Address: 2001:db8:2:5::5
    END.X SID: 2001:db8:2:0:40:: B:0 S:0 P:0 End.X (PSP)
  Metric: 10         MT (IPv6 Unicast) IS-Extended P1.00
    Interface IPv6 Address: 2001:db8:2:3::2
    Neighbor IPv6 Address: 2001:db8:2:3::3
    END.X SID: 2001:db8:2:0:41:: B:0 S:0 P:0 End.X (PSP)
```

P node

Endpoint function of PE2 (P node)

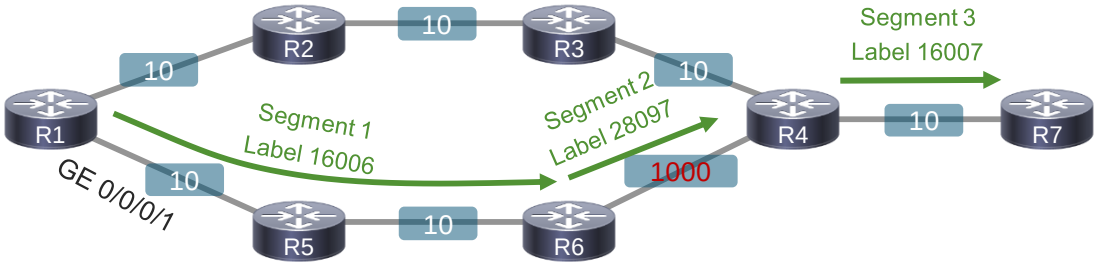
SR Operations, Administration, and Maintenance (OAM)

You make networking **possible**

NIL-FEC

- Nil FEC (defined in RFC4379), specifies that no explicit FEC (Control Plane) is associated with the label
- Typically used to carry labels in reserved range (explicit-null or router alert) for diagnostic purpose
- Ping and traceroute
- But very powerful tool to check any combination of segments on any path!
- Does not carry any information to identify the intended target
  - The packet may be forwarded wrongly somewhere, but still make it
  - No control plane validation is performed at originator or responder
- This was an interim solution
- Can force traffic over non-least cost path

NIL-FEC Example



User specifies:

- Outgoing label stack (one or more labels)
- Outgoing interface
- Next-hop interface address

Specify outgoing interface and next hop

```
RP/0/0/CPU0:R1# trace mpls nil-fec labels 16006,28097,16007 output interface gigabitEthernet 0/0/0/1 nexthop 10.1.15.5

Tracing MPLS Label Switched Path with Nil FEC with labels [16006,28097,16007], timeout is 2 seconds

Codes: '!' - success, 'Q' - request not sent, '.' - timeout,
  'L' - labeled output interface, 'B' - unlabeled output interface,
  'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch,
  'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label,
  'P' - no rx intf label prot, 'p' - premature termination of LSP,
  'R' - transit router, 'I' - unknown upstream index,
  'X' - unknown return code, 'x' - return code 0

Type escape sequence to abort.

  0 10.1.15.1 MRU 1500 [Labels: 16006/28097/16007/explicit-null Exp: 0/0/0/0]
L 1 10.1.15.5 MRU 1500 [Labels: implicit-null/28097/16007/explicit-null Exp: 0/0/0/0] 10 ms
L 2 10.1.56.6 MRU 1500 [Labels: implicit-null/16007/explicit-null Exp: 0/0/0] 0 ms
L 3 10.1.46.4 MRU 1500 [Labels: implicit-null/explicit-null Exp: 0/0] 10 ms
! 4 10.1.47.7 10 ms
```

Specify segments as list of labels in comma separated list (first label is top label)

28097 is adj-SID label from R6 to R4

SR OAM (Work-in-Progress)

- Ping, traceroute for Prefix SIDs and Adj-SIDs for IGP
- Regular MPLS OAM works for SR

```
ping mpls ipv4 10.1.1.1/32
traceroute mpls ipv4 10.1.1.1/32
```

- OAM gives you extra (above normal ping and traceroute):
  - Consistency check
  - Path discovery
  - MPLS traffic black hole
  - Path divergence detection
  - Premature IP header exposition
  - Can detect inconsistencies between control plane and forwarding

```
ping mpls ipv4 10.1.1.1/32 fec-type generic
traceroute mpls ipv4 10.1.1.1/32 fec-type generic
```

```
ping sr-mpls 10.1.1.1/32 fec-type igp <isis/ospf>
traceroute sr-mpls 10.1.1.1/32 fec-type igp <isis/ospf>
```

- OAM was expanded with SR OAM
  - Only prefix-SID for now
  - Only new Target FEC Stack TLV for SR is added

```
RP/0/RP0/CPU0:PE1# trace sr-mpls policy ?
  binding-sid  Specify the binding-sid of the SR policy
  color        Specify the color of the SR policy
  name         Specify the name of the SR policy
```

Debugging SR OAM

```
debug mpls oam packet
debug mpls oam tlv

lspv_server[1113]: [3808093952] DBG-TLV: : Echo Hdr encode: version:[1], msg type:[1], reply mode:[2],
lspv_server[1113]: [3808093952] DBG-TLV: :   return_code:[0], return_subcode:[0], sender handle:[5b9d40ec],
lspv_server[1113]: [3808093952] DBG-TLV: :   sequence number:[2],
lspv_server[1113]: [3808093952] DBG-TLV: :   timestamp sent:[E092609D.9D5A9BC6 (12:22:53.614 UTC Fri May 24 2019)],
lspv_server[1113]: [3808093952] DBG-TLV: :   timestamp rcvd:[00000000.00000000 (00:00:00.000 UTC Thu Jan  1 1970)]
lspv_server[1113]: [3808093952] DBG-TLV: : Cisco ext subTLV encode: type:[1], length:[4], tlv revision:[0x4]
lspv_server[1113]: [3808093952] DBG-TLV: : SR IGP IPv4 Prefix SID encode: destaddr 10.0.0.2/32, protocol OSPF
lspv_server[1113]: [3808093952] DBG-Pkt: : TFS TLV added for request (sender_handle:[0x5b9d40ec])
lspv_server[1113]: [3808093952] DBG-TLV: : DSMAP encode:
lspv_server[1113]: [3808093952] DBG-TLV: :   addr_type:[1], rtr_id:[10.3.5.5], mtu:[1500],
lspv_server[1113]: [3808093952] DBG-TLV: :   intf_addr:[10.3.5.5], flags:[0x0], hashkey:[0], depth limit:[0],
lspv_server[1113]: [3808093952] DBG-TLV: :   multipath length:[0], [16002]
lspv_server[1113]: [3808093952] DBG-Pkt: : DSMAP TLV added for request (sender_handle:[0x5b9d40ec])
lspv_server[1113]: [3808093952] DBG-Pkt: : UDP checksum:[0xcfa9] <10.1.3.1,3503> -> <127.0.0.1,3503>, len:[96]
lspv_server[1113]: [3808093952] DBG-Pkt: : Echo packet built successfully for request (sender_handle:[0x5b9d40ec]), pak size:[124]

lspv_server[1113]: [3808093952] DBG-Pkt: : Processing received ipv4 packet
lspv_server[1113]: [3808093952] DBG-Pkt: : Echo packet received: rx interface:[0x1000018], src:[10.3.5.5], dst:[10.1.3.1],
lspv_server[1113]: [3808093952] DBG-Pkt: : Getting rx info
lspv_server[1113]: [3808093952] DBG-TLV: : Echo Hdr decode: version:[1], msg type:[2], reply mode:[2],
lspv_server[1113]: [3808093952] DBG-TLV: :   return_code:[8], return_subcode:[1], sender handle:[5b9d40ec],
lspv_server[1113]: [3808093952] DBG-TLV: :   sequence number:[2],
lspv_server[1113]: [3808093952] DBG-TLV: :   timestamp sent:[E092609D.9D5A9BC6 (12:22:53.614 UTC Fri May 24 2019)],
lspv_server[1113]: [3808093952] DBG-TLV: :   timestamp rcvd:[E092609E.3DD46742 (12:22:54.241 UTC Fri May 24 2019)]
lspv_server[1113]: [3808093952] DBG-TLV: : Downstream Mapping, tlvtype:[0x2], tlvlength:[0x14]
lspv_server[1113]: [3808093952] DBG-TLV: :   epkt->dsmap_tlvs.echo_reply_dsmap_queue inited
lspv_server[1113]: [3808093952] DBG-TLV: : DSMAP decode:
lspv_server[1113]: [3808093952] DBG-TLV: :     addr_type:[1], rtr_id:[10.5.6.6], mtu:[1500],
lspv_server[1113]: [3808093952] DBG-TLV: :     intf_addr:[10.5.6.6], flags:[0x0], hashkey:[0],
lspv_server[1113]: [3808093952] DBG-TLV: :     depth limit:[0], multipath length:[0],
lspv_server[1113]: [3808093952] DBG-TLV: :     Labels:[16002]
lspv_server[1113]: [3808093952] DBG-TLV: :   DSMAP enqueued to epkt->dsmap_tlvs.echo_reply_dsmap_queue
lspv_server[1113]: [3808093952] DBG-TLV: : Echo packet decoded assuming tlv version:[4]
```

Key Takeaway's

- SR is simpler and easier to troubleshoot than LDP
- No changes in MPLS forwarding
- Ti-LFA
  - Built from same fundaments as LFA
  - But much better and much easier
- SR Policy (SR-TE) is simpler than RSVP-TE
- Controller
  - Understand basic PCEP
  - Understand BGP-LS if used

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

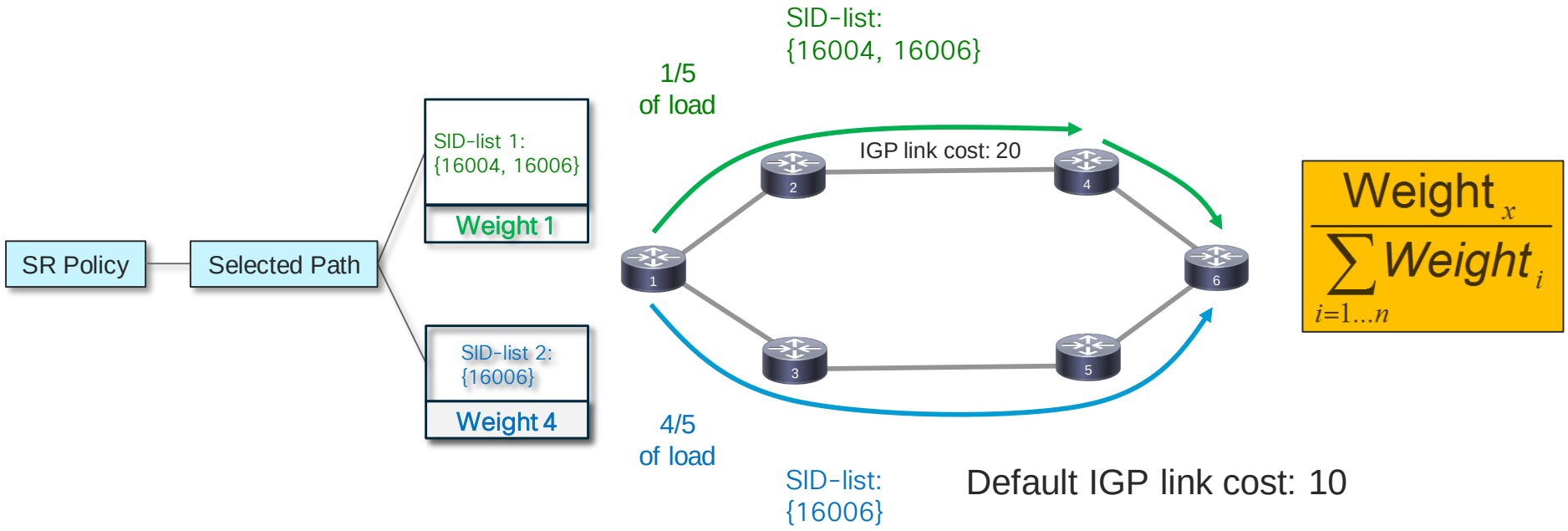
You make **possible**

Appendix

You make networking **possible**

Unequal Cost Multi-Path (UCMP)

- If a set of SID lists is associated with the selected path of the policy, then the steering is flow and UCMP-based according to the relative weight of each SID list



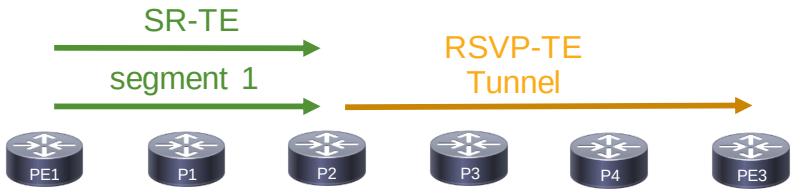
Min-metric with Margin

- Headend computes a SID list whose cumulated optimized metric is within range of the [shortest-path metric, shortest-path metric + margin]
- Margin is expressed in absolute or relative value
- If this is not possible because of the number of SIDs constraint, then the solution SID list minimizes the optimized metric while meeting the maximum number of SIDs constraints

```
segment-routing
 traffic-eng
  policy policy-1
   candidate-paths
    preference 100
     dynamic
      metric
       sid-limit 5
       margin absolute 100
```

SR-TE to RSVP-TE

- PE-to-PE




```
PE1

explicit-path name PE1-P2
 index 10 next-address strict ipv4 unicast 10.100.1.9
 index 20 next-label 16
!
interface tunnel-te1
 ipv4 unnumbered Loopback0
 autoroute destination 10.100.1.3
 autoroute destination 10.100.1.9
 destination 10.100.1.9
 path-option 1 explicit name PE1-P2 segment-routing
```

```
segment-routing
 global-block 16000 23999
 traffic-eng
  logging
   policy status
  !
  policy to-P2
   binding-sid mpls 16
   color 1000 end-point ipv4 10.0.0.2
   candidate-paths
    preference 100
     dynamic
     !
    !
   !
  !
```

```
RP/0/0/CPU0:PE1# traceroute vrf one 10.2.100.3

Type escape sequence to abort.
Tracing the route to 10.2.100.3

 1  10.1.15.5 [MPLS: Labels 16009/16/24004 Exp 0] 29 msec  39 msec  89 msec
 2  10.1.59.9 [MPLS: Labels 16/24004 Exp 0] 49 msec  29 msec  49 msec
 3  10.1.129.12 [MPLS: Labels 24008/24004 Exp 0] 39 msec  39 msec  29 msec
 4  10.1.126.6 [MPLS: Labels 24010/24004 Exp 0] 39 msec  39 msec  29 msec
 5  10.1.36.3 29 msec  *  29 msec
```

24004 is service (vpnv4) label

Protected/Non-Protected Adj-SID for SR-TE

- For each adjacency two Adj-SIDs can be advertised
- Adj-SID with Backup-flag = 1 (only advertised if Ti-LFA is configured)
  - This adj-SID may be protected
- Adj-SID with Backup-flag = 0
  - This adj-SID is not protected
  - In case of failure, the traffic is switched to other converged path (no FRR)
  - Example use-case: TDM based service: If a link fails then the end-to-end path must fail, higher layer provides redundancy
- With path-selection for SR Policy, you can select the SR Policy to use (if Adj-SIDs are present):
  - Only protected Adj-SIDs
  - Only unprotected Adj-SIDs

```
segment-routing
 traffic-eng
  policy POLICY_1
   end-point ipv4 1-.1.1.3 color 1
    candidate-paths
     preference 50
     dynamic mpls pce
      unprotected
      invalidation drop
```

drop traffic if path is invalid

- The default is to use any Adj-SID, preferring the protected ones
- Routes pointing to a SR Policy do no have backup paths in RIB because tunnel itself has backup

SR BGP EPE

You make networking **possible**

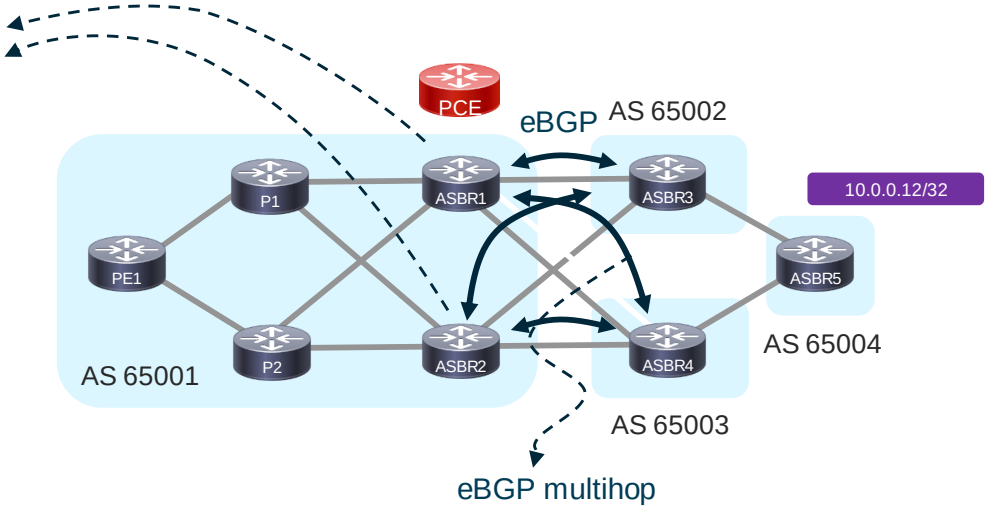
Cisco live!

Egress Peer Engineering (EPE)

- Controller instructs ingress PE to use a specific egress PE and egress peer
  - External interface
  - eBGP neighbor
- EPE functionality is needed only at egress PE and controller
- Controller:
  - gathers topology info + EPE info (through BGP-LS)
  - Compute and implement per-flow traffic steering
- Delivery:
  - SR-TE (PCEP/CLI/XML)
  - NetConf
  - BGP-LU (RFC3017)
  - …
- Ingress PE creates a lists of segments per flow, indicating intra-AS path + egress hop

Egress PE Configuration

```
router bgp 65001

 neighbor 10.4.6.6
  remote-as 65002
  egress-engineering
  address-family ipv4 unicast
   route-policy pass in
   route-policy pass out
  !
 !
 neighbor 10.0.0.7
  remote-as 65003
  ignore-connected-check
  egress-engineering
  update-source Loopback0
  address-family ipv4 unicast
   route-policy pass in
   route-policy pass out
  !
 !
```



- PeerNode SID: to eBGP Peer
  - MPLS Dataplane: Pop and Forward on any interface to the peer
- PeerAdj SID: to eBGP Peer via interface
  - MPLS Dataplane: Pop and Forward on the related interface
- PeerSet SID: to set of eBGP peers
  - MPLS Dataplane: Pop and Forward on any interface to the set of peers
  - All the peers in a set might not be in the same AS
  - Not available yet

EPE on ASBR

```
RP/0/0/CPU0:ASBR1# show bgp egress-engineering

 Egress Engineering Peer Set: 10.0.0.7/32 (15637008)
       Nexthop: 10.0.0.7
       Version: 2, rn_version: 2
         Flags: 0x00000006
     Local ASN: 65001
    Remote ASN: 65003
     Local RID: 10.0.0.4
    Remote RID: 10.0.0.7
  Local Address: 10.0.0.4
     First Hop: 10.4.7.7, 10.104.107.7
          NHID: 12, 13
           IFH: 0x60, 0x80
         Label: 24000, Refcount: 3
       rpc_set: 1626b864, ID: 1
...
 Egress Engineering Peer Set: 10.104.107.7/32 (15636f64)
       Nexthop: 10.104.107.7
       Version: 3, rn_version: 5
         Flags: 0x0000000a
     Local ASN: 65001
    Remote ASN: 65003
     Local RID: 10.0.0.4
    Remote RID: 10.0.0.7
  Local Address: 10.104.107.4
     First Hop: 10.104.107.7
          NHID: 13
           IFH: 0x80
         Label: 24001, Refcount: 3
       rpc_set: 1626b8dc, ID: 2
```

multiple next hops

PeerNode-SID

1 next hop

PeerAdj-SID

```
RP/0/0/CPU0:ASBR2# show bgp egress-engineering

 Egress Engineering Peer Set: 10.5.6.6/32 (15636f64)
       Nexthop: 10.5.6.6
       Version: 3, rn_version: 3
         Flags: 0x00000006
     Local ASN: 65001
    Remote ASN: 65002
     Local RID: 10.0.0.5
    Remote RID: 10.0.0.6
  Local Address: 10.5.6.5
     First Hop: 10.5.6.6
          NHID: 7
           IFH: 0x60
         Label: 24009, Refcount: 3
       rpc_set: 1642b674, ID: 2

 Egress Engineering Peer Set: 10.5.7.7/32 (15637008)
       Nexthop: 10.5.7.7
       Version: 2, rn_version: 2
         Flags: 0x00000006
     Local ASN: 65001
    Remote ASN: 65003
     Local RID: 10.0.0.5
    Remote RID: 10.0.0.7
  Local Address: 10.5.7.5
     First Hop: 10.5.7.7
          NHID: 8
           IFH: 0x80
         Label: 24003, Refcount: 3
       rpc_set: 1633af28, ID: 1
```

To ASBR3

PeerAdj-SID

PeerAdj-SID

Peer-Node-SID and Peer-Adj-SID labels are installed as local labels on ASBRs

Example EPE with PCEP

```
RP/0/0/CPU0:PCE# show bgp link-state link-state

Prefix codes: E link, V node, T IP reacheable route, u/U unknown
              I Identifier, N local node, R remote node, L link, P prefix
              L1/L2 ISIS level-1/level-2, O OSPF, D direct, S static/peer-node
              a area-ID, l link-ID, t topology-ID, s ISO-ID,
              c confed-ID/ASN, b bgp-identifier, r router-ID,
              i if-address, n nbr-address, o OSPF Route-type, p IP-prefix
              d designated router address
   Network            Next Hop            Metric LocPrf Weight Path
*>i[E][B][I0x0][N[c65001][b0.0.0.0][q10.0.0.4]][R[c65002][b0.0.0.0][q10.0.0.6]][L[i10.4.6.4][n10.4.6.6]]/664
                      10.0.0.4                 100      0 i
*>i[E][B][I0x0][N[c65001][b0.0.0.0][q10.0.0.4]][R[c65003][b0.0.0.0][q10.0.0.7]][L[i10.0.0.4][n10.0.0.7]]/664
                      10.0.0.4                 100      0 i
*>i[E][B][I0x0][N[c65001][b0.0.0.0][q10.0.0.4]][R[c65003][b0.0.0.0][q10.0.0.7]][L[i10.4.7.4][n10.4.7.7]]/664
                      10.0.0.4                 100      0 i
*>i[E][B][I0x0][N[c65001][b0.0.0.0][q10.0.0.4]][R[c65003][b0.0.0.0][q10.0.0.7]][L[i10.104.107.4][n10.104.107.7]]/664
                      10.0.0.4                 100      0 i

Processed 4 prefixes, 4 paths
```

PeerAdj-SID

PeerNode-SID

```
RP/0/0/CPU0:PCE# show bgp link-state link-state
[E][B][I0x0][N[c65001][b0.0.0.0][q10.0.0.4]][R[c65003][b0.0.0.0][q10.0.0.7]][L[i10.0.0.4][n10.0.0.7]]/664
 BGP routing table entry for
[E][B][I0x0][N[c65001][b0.0.0.0][q10.0.0.4]][R[c65003][b0.0.0.0][q10.0.0.7]][L[i10.0.0.4][n10.0.0.7]]/664
Local
    10.0.0.4 (metric 2) from 10.0.0.4 (10.0.0.4)
      Origin IGP, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 1, version 3
      Link-state: Peer-SID: 24000
```

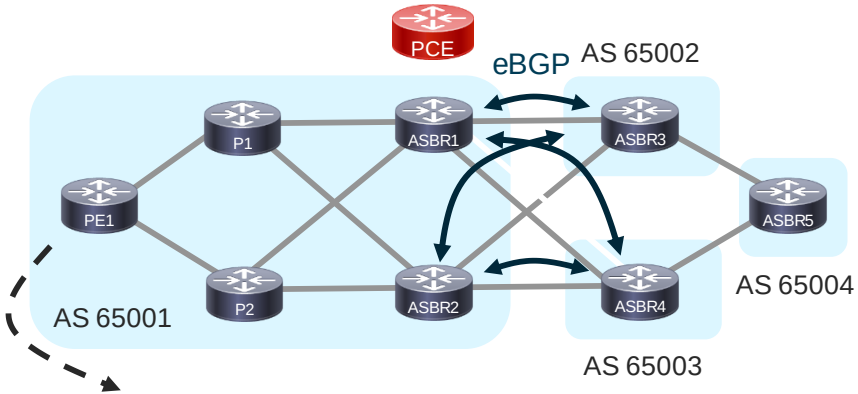
PCE learns EPE labels by BGP Link-State

Example EPE with PCEP

```
RP/0/0/CPU0:PCE# show pce ipv4 topology

PCE's topology database - detail:
--------------------------------
Node 1
  TE router ID: 10.0.0.4
  BGP router ID: 10.0.0.4 ASN: 65001

  Link[0]: local address 10.4.6.4, remote address 10.4.6.6
    Local node:
      BGP router ID: 10.0.0.4 ASN: 65001
    Remote node:
      BGP router ID: 10.0.0.6 ASN: 65002
    Metric: IGP 0, TE 0, Latency 0 microseconds
    Bandwidth: Total 0 Bps, Reservable 0 Bps
    Admin-groups: 0x00000000
    Adj SID: 24006 (epe)

  Link[1]: local address 10.0.0.4, remote address 10.0.0.7
    Local node:
      BGP router ID: 10.0.0.4 ASN: 65001
    Remote node:
      BGP router ID: 10.0.0.7 ASN: 65003
    Metric: IGP 0, TE 0, Latency 0 microseconds
    Bandwidth: Total 0 Bps, Reservable 0 Bps
    Admin-groups: 0x00000000
    Adj SID: 24000 (epe)

 …

Node 2
  BGP router ID: 10.0.0.6 ASN: 65002
```

EPE labels are dynamic labels!

Example EPE with PCEP




```
segment-routing
 traffic-eng
 !
  policy EPE-PCEP
   color 100 end-point ipv4 10.0.0.6
   candidate-paths
    preference 100
     dynamic
      pcep
```

end-point is ASBR3

end point is BGP next hop address !

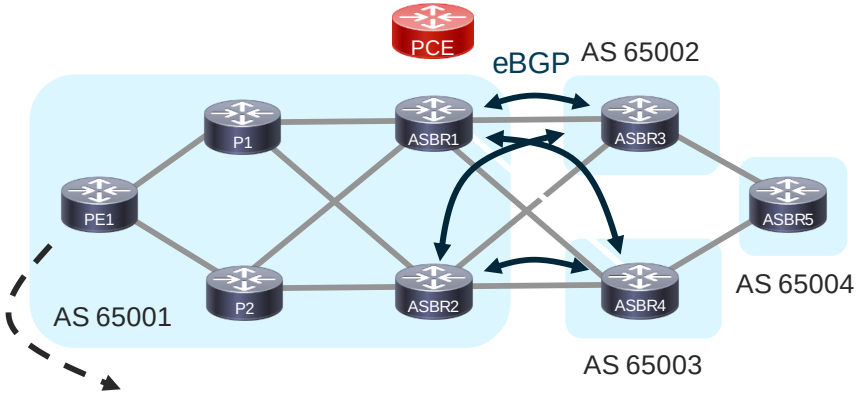
```
RP/0/0/CPU0:PE1# show seg traffic-eng policy color 100 detail

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

Color: 100, End-point: 10.0.0.6
  Name: srte_c_100_ep_10.0.0.6
  Status:
    Admin: up  Operational: up for 00:25:13 (since May  2 15:51:09.025)
  Candidate-paths:
    Preference: 100 (configuration) (active)
      Name: EPE-PCEP
      Requested BSID: dynamic
      PCC info:
        Symbolic name: cfg_EPE-PCEP_discr_100
        PLSP-ID: 3
      Dynamic (pce 10.0.0.10) (valid)
        Metric Type: TE,    Path Accumulated Metric: 2
          24009 [Adjacency-SID, 10.1.3.1 - 10.1.3.3]    -> P1
          24004 [Adjacency-SID, 10.3.4.3 - 10.3.4.4]    -> ASBR1
          24011 [Adjacency-SID, 10.4.6.4 - 10.4.6.6]
  LSPs:
    LSP[0]:
      LSP-ID: 1 policy ID: 7 (acti
      Local label: 24006
      State: Programmed
      Binding SID: 24007
```

PeerAdj-SID
ASBR1 to ASBR3

```
RP/0/0/CPU0:PE1# show mpls forwarding labels 24006
Local  Outgoing    Prefix          Outgoing     Next Hop        Bytes
Label  Label       or ID           Interface                    Switched
------ ----------- --------------- ------------ --------------- ------------
24006  24004       SR TE: 7 [TE-INT]  Gi0/0/0/1    10.1.3.3        0
```

Traffic goes ASBR1 to ASBR3, even though ASBR1 has ASBR4 as BGP next-hop for routes with color 100

Example EPE with Segment List




```
segment-routing
 traffic-eng
  segment-list to-ASBR3-via-ASBR1
   index 10 mpls label 16004
   index 20 mpls label 24011
  !
 policy EPE-1
   color 100 end-point ipv4 10.0.0.6
   candidate-paths
    preference 100
     explicit segment-list to-ASBR3-via-ASBR1

router static
 address-family ipv4 unicast
  10.0.0.12/32 sr-policy srte_c_100_ep_10.0.0.6
```

label ASBR1

EPE label: ASBR1 to ASBR 3

end-point is ASBR3

```
RP/0/0/CPU0:PE1# show seg traffic-eng policy color 100

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

Color: 100, End-point: 10.0.0.6
  Name: srte_c_100_ep_10.0.0.6
  Status:
    Admin: up  Operational: up for 00:04:55 (since May  2
16:36:06.560)
  Candidate-paths:
    Preference: 100 (configuration) (active)
      Name: EPE-1
      Requested BSID: dynamic
      Explicit: segment-list to-ASBR3-via-ASBR1 (valid)
        Weight: 1, Metric Type: TE
          16004
          24011
  Attributes:
    Binding SID: 24012
    Forward Class: 0
    Steering labeled-services disabled: no
    Steering BGP disabled: no
    IPv6 caps enable: yes
```