The bridge to possible

Multicast Segment Routing & Traffic Engineering

Multicast and segment routing

Mankamana Mishra
Technical Leader

BRKMPL-2123

Cisco Webex app

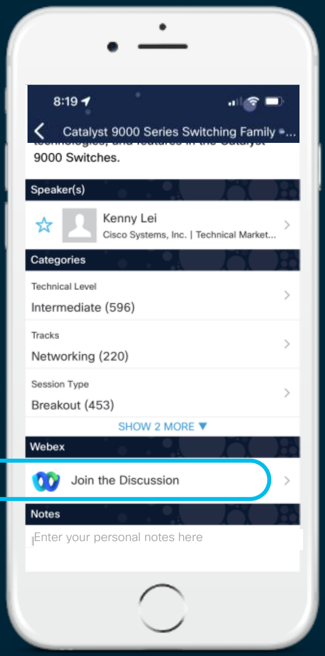
Questions?

Use Cisco Webex App 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 the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 17, 2022.



https://ciscolive.ciscoevents.com/ciscolivebot/#BRKMPL-2123

Agenda

- Introduction
- Basic Segment routing
- mLDP
- mLDP with flex algorithm
- Tree SID and traffic engineering
- Standardization

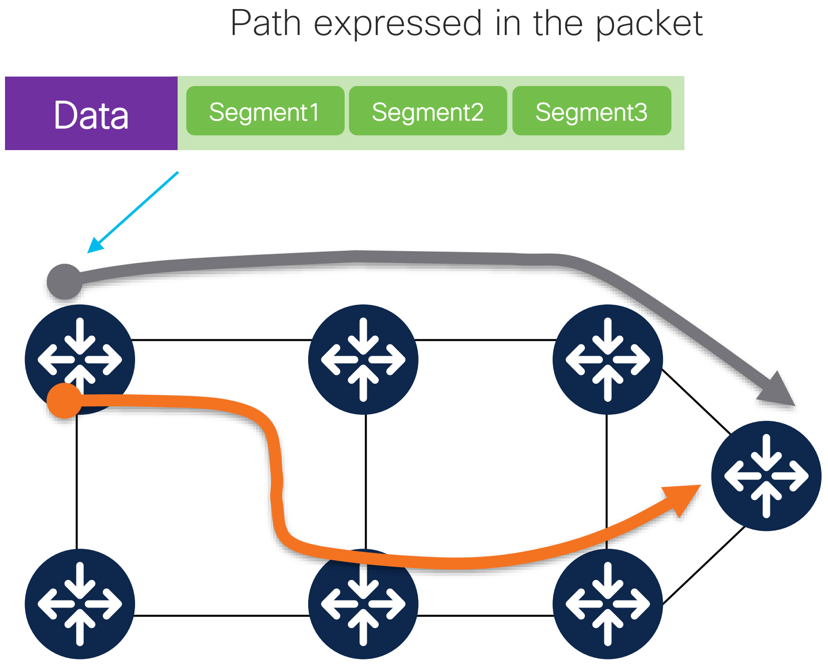
Introduction

- Segment routing is a technology that uses Source Routing to forward packets through the network.
- A packet is forwarded from Segment to Segment based on information carried in the packet.
- Due to adding more information in the packet, less state needs to be maintained in the network and can potentially be simplified.

SR Technology Overview

Segment routing

- Source Routing
  - the source chooses a path and encodes it in the packet header as an ordered list of segments
  - the rest of the network executes the encoded instructions
- Segment: an identifier for any type of instruction
  - forwarding or service



Segment routing – forwarding plane

Segment Routing

SR-MPLS

- Instantiation of SR on the MPLS data plane
- A segment is encoded with an MPLS label

SRv6

- Instantiation of SR on the IPv6 data plane
- A segment is encoded with an IPv6 address

A rich set of segment types

IGP Prefix SID	GLOBAL segment representing an IGP prefix Forward packet along shortest-path (ECMP-aware) to reach the prefix associated with the segment
IGP Node SID	An IGP-Prefix segment identifying a specific router (for example its loopback prefix)
IGP Anycast SID	An IGP-Prefix segment assign to an IGP prefix advertised by multiple routers (anycast prefix)
IGP Adjacency SID	LOCAL segment representing an IGP adjacency Forward packet over the interface where the adjacency is formed
BGP Prefix SID	GLOBAL segment representing a BGP prefix Forward packet along best-path to reach the prefix associated with the segment
BGP Peering SID	LOCAL segment representing a BGP neighbor Forward packet over the interface where the neighbor is formed

Traditional multicast options

- Deploying SR for unicast is orthogonal to solution used for Multicast.
- Nothing prevents existing protocols to continue to work, like:
  - Ingress Replication (IR)
  - PIM
  - mLDP
  - RSVP-TE
- In that sense, there is no requirement to change the Multicast deployment.
- However, if there is a technology that would benefit from being simplified and scale improved, its Multicast ☺

Simplifying multicast delivery

- SR architecture and its building blocks enhance existing multicast solutions and allow for new ones
- Depending on the requirements, we can choose the best fit from options such as:

Multicast strategy	Computation	SR-MPLS-Multicast	SRv6-Multicast
Non Traffic engineering (shortest path)	Distributed	mLDP	PIMv6
Traffic Engineering (Controlled path)	Distributed	mLDP+Flex-Algo	Work in Progress
	Centralized	TreeSID + PCE controller	TreeSID + PCE controller

mLDP

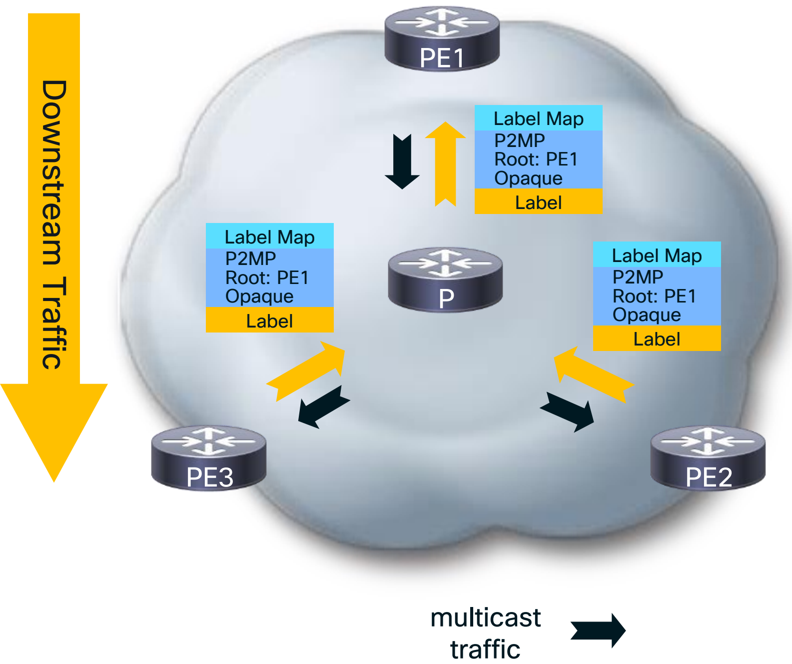
mLDP introduction

- mLDP is a protocol that builds
  - P2MP LSPs
  - MP2MP LSPs
- Very often and **preferred** deployed for Multicast VPNs
- It's a receiver driven tree building protocol like PIM.
- mLDP uses the LDP Transport to exchange Label Mappings.

mLDP signaling and packet forwarding

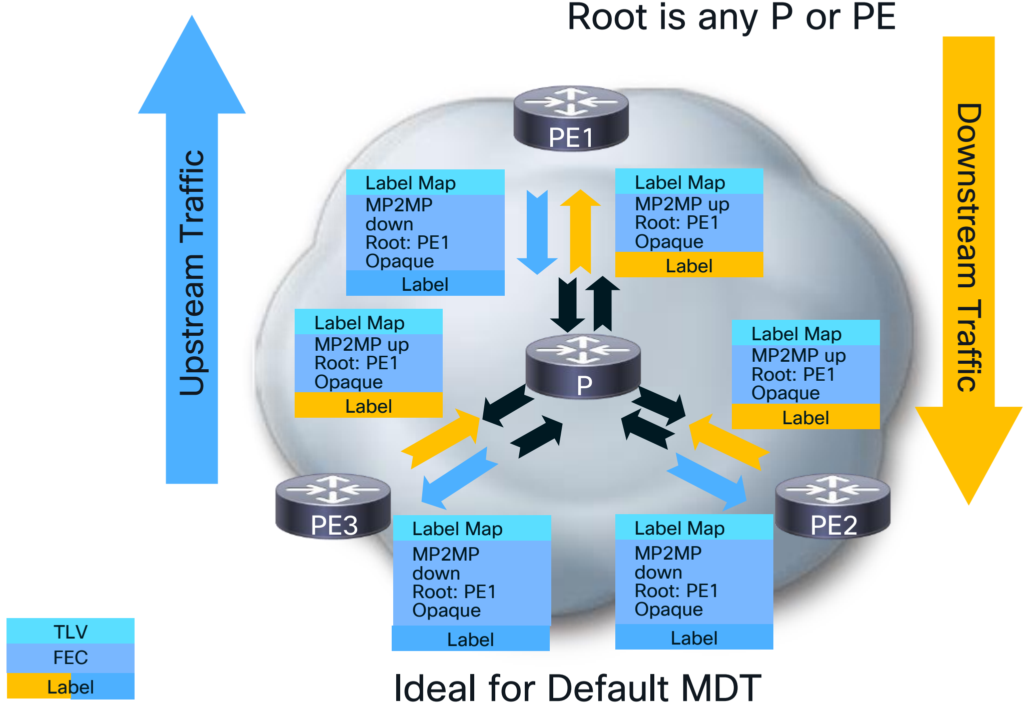
P2MP Tree

Root is ingress PE



MP2MP Tree

Root is any P or PE



Ideal for Default MDT
Data plane is still P2MP

Multicast LDP with transport differentiation using SR Flex-Algo

SR IGP Flex Algo

- Complements the SRTE solution by adding new Prefix-Segments with specific optimization objective and constraints
  - minimize igp-metric or delay or te-metric
  - avoid SRLG or affinity
- Leverages the SRTE benefits of simplicity and automation
  - Automated sub-50msec FRR (TILFA)
  - On-Demand Policy (ODN)
  - Automated Steering (AS)

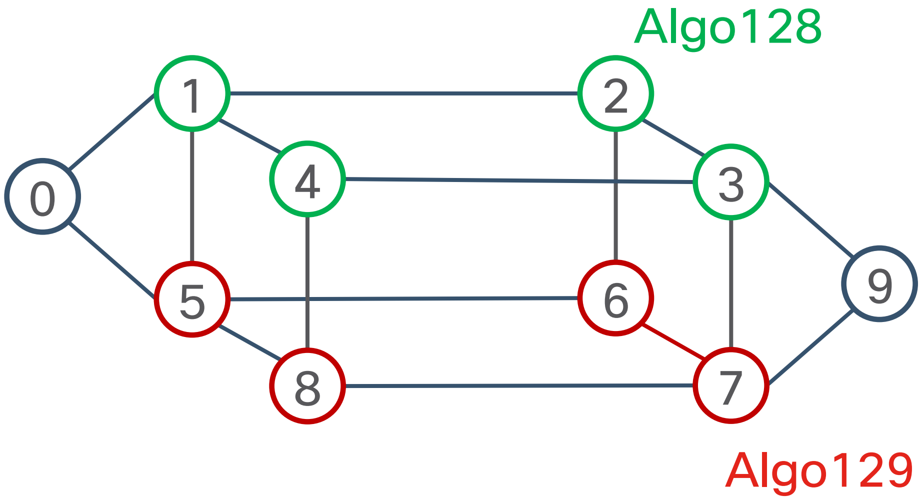
Flex-Algo participation advertisement

- Each node MUST advertise Flex-Algo(s) that it is participating in

Nodes 0 and 9 participate to Algo 0 and 128 and 129

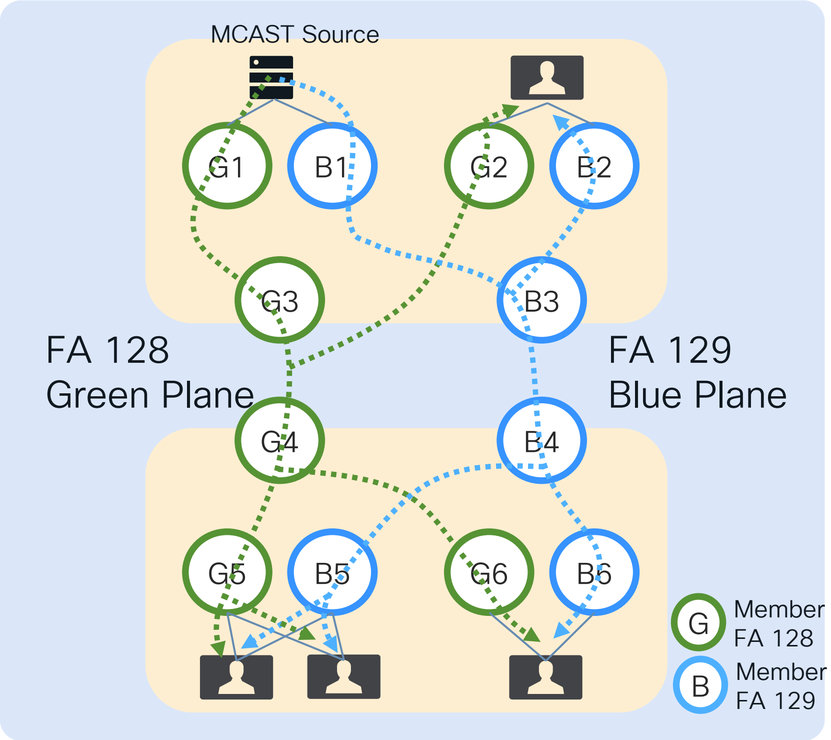
Nodes 1/2/3/4 participate to Algo 0 and 128

Nodes 5/6/7/8 participate to Algo 0 and 129



Flex-Algo aware MLDP – highlights

- MLDP-signaled Multicast Distribution Trees built within the Flex-Algo topology
- Value Proposition:
  - Multicast LDP with transport differentiation using SR Flex-Algo
  - Low latency routing
  - Disjoint paths (multicast Live-Live)
  - Paths avoiding specific links
  - Data sovereignty / region scope



Flex-Algo aware MLDP

- BGP MVPN discovery routes carry a P-MSI tunnel attribute (PTA) which identifies the transport used for mVPN
- For MLDP transport, MLDP FEC is carried in the PTA. The Flex-Algo instance value is stored in the PTA ID field.

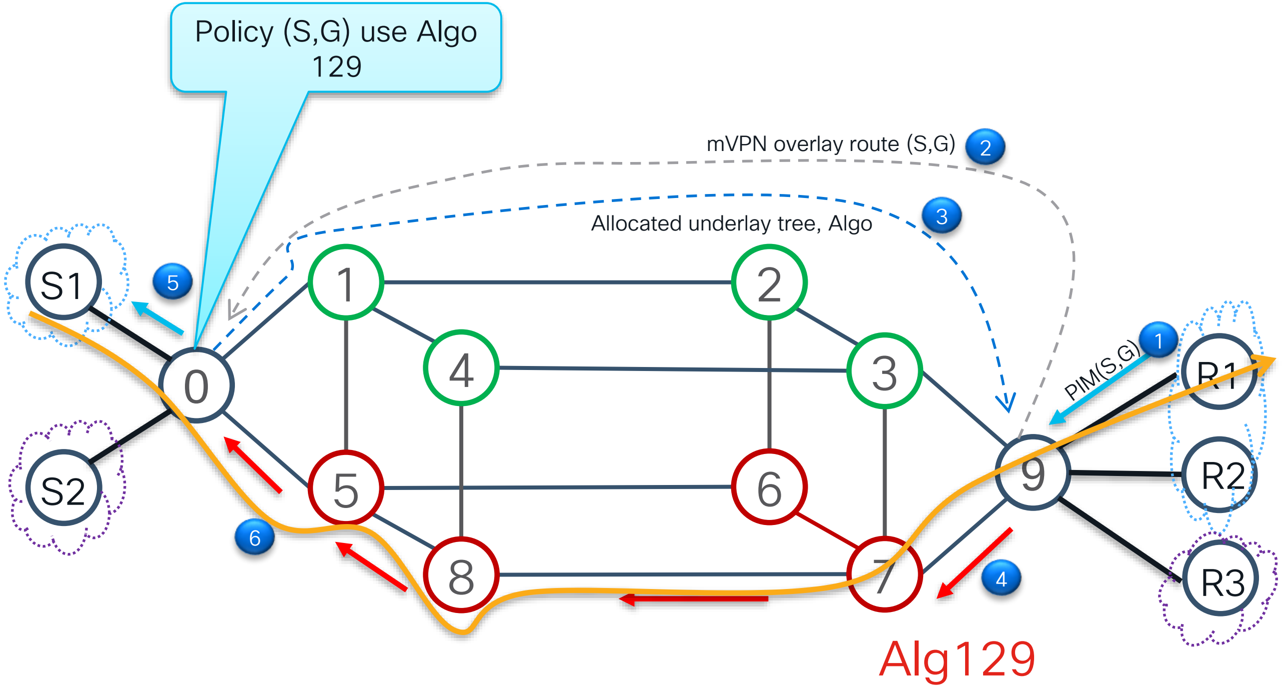
```
0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| MP FEC type   |  AF (MT IP/ MT IPv6)          |    AF Length  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Root Node Address                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Reserved   |       IPA     |           MT-ID               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Opaque Length              |          Opaque Value         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               +
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IGP Algorithm (IPA) field = carries SR Flex algo instance ID

Flex-Algo aware mLDP Cisco IOS-XR implementation

- mVPN profile 14 – Partitioned MDT mLDP P2MP with BGP-AD and BGP c-mcast signaling.
- mVPNv4 / mVPNv6 overlay
- Partitioned and Data MDTs
- Granular mapping of (C-S,C-G) to a Partitioned MDT / DATA MDT bound to a Flex-Algo instance
- PIM ASM, SSM, IGMPv2 and IGMPv3 as customer access protocols
- ECMP – A Flex-Algo topology may have ECMP and therefore multicast flows are load balanced if multiple paths are available.

mLDP signaling with Flex-Algo



Tree SID

Tree SID overview

- Tree-SID is a SDN controller-based approach to building P2MP trees in a SR domain
  - Cisco's SR Path Computation Element (SR-PCE) acts as controller
- With the central knowledge at the SR-PCE, the tree can be built using constraints.
- In this presentation we'll focus on Trees using MPLS as data-plane.

SR Path Computation Element (SR-PCE)

SRTE Head-End

Distributed Mode – SR-TE Head-End
Visibility is limited to its own IGP domain

Solution

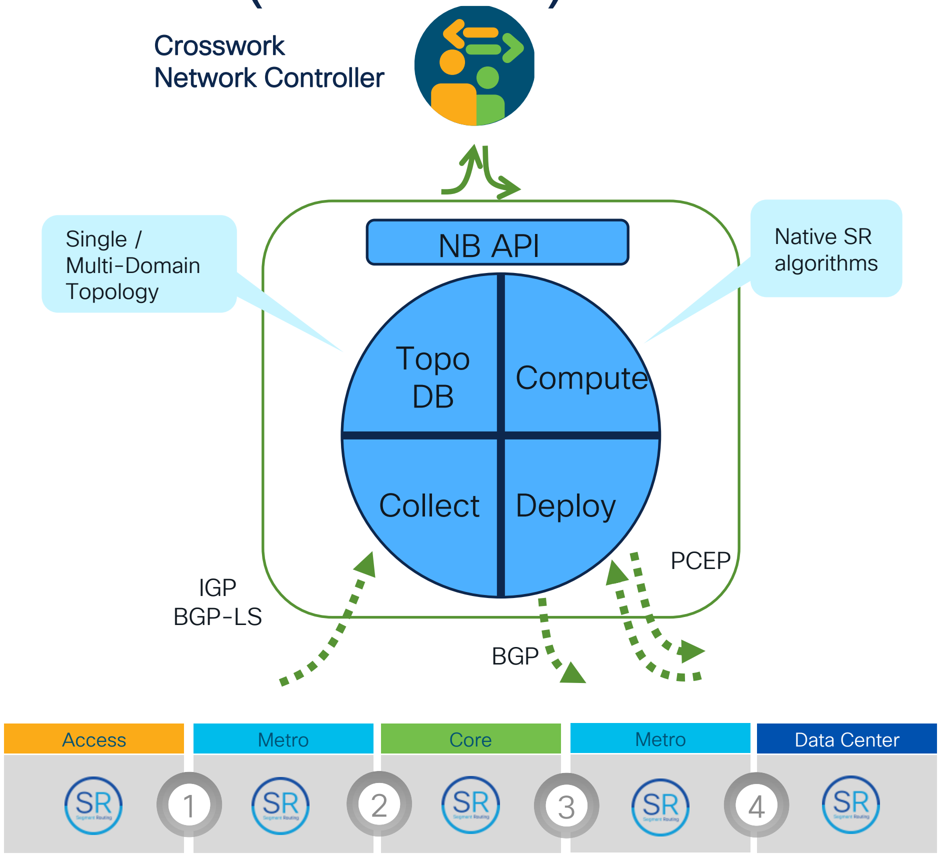
Multi-Domain SRTE Visibility
Centralized SR-PCE for Multi-Domain Topology view

Integration with Applications
North-bound APIs for topology/deployment

Delivers **across the unified SR Fabric** the SLA requested by the service

Benefits

Simplicity and Automation
End-to-End network topology awareness
SLA-aware path computation across network domains



SR Replication Segment - A segment for P2MP delivery

- Replication segment allows node (Replication Node) to replicate packets to a set of other nodes (Downstream Nodes) in a Segment Routing Domain
- A Replication segment is an MPLS label
- Replication segments provide building blocks for Point-to-Multipoint Service delivery via SR Point-to-Multipoint (SR P2MP) policy
- A Replication segment can replicate packet to directly connected nodes or to downstream nodes (without need for state on the transit routers)
- The use of one or more stitched Replication segments constructed for SR P2MP Policy tree

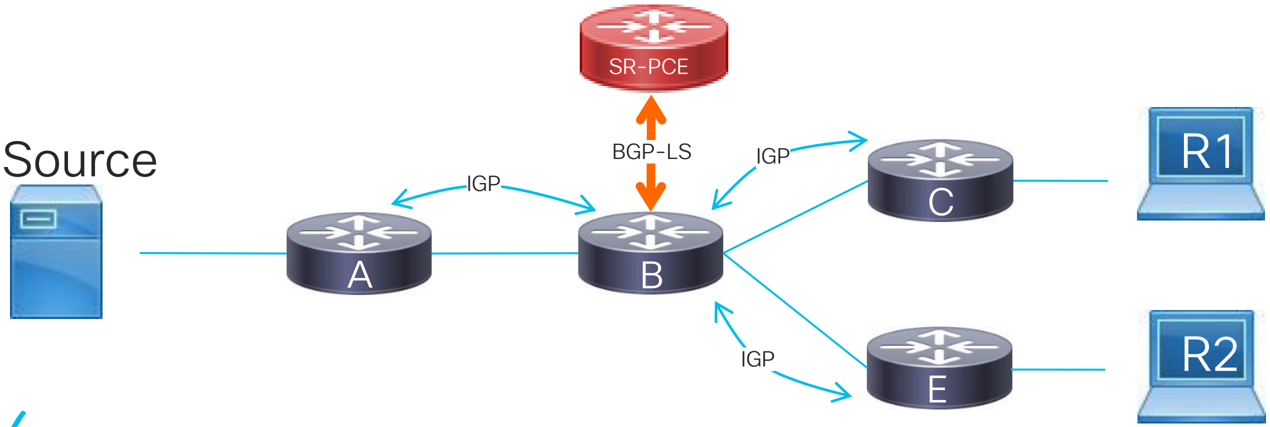
Tree-SID

SR-PCE is responsible for:

1. Learning the topology.
2. Learning the Root and Leaf's of the Tree.
3. Computing the Tree.
4. Knowing the MPLS Labels it can use.
5. Having a mechanism to program the Forwarding state.

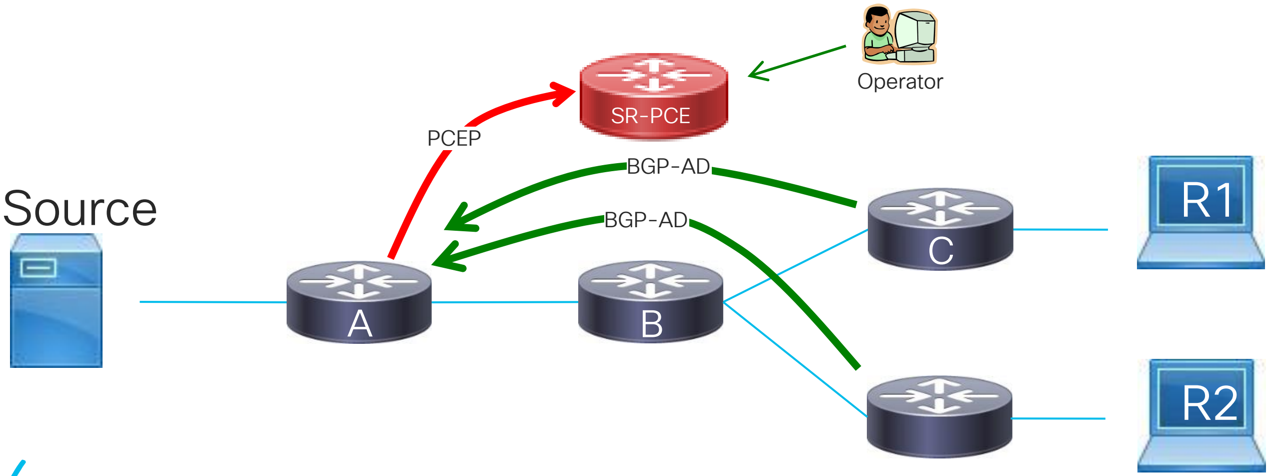
SR-PCE – learning the topology

- A common mechanism to learn the topology is using BGP Link State (LS).
- Through BGP-LS, the controller sucks up the Link State database.
- Through the LS database, the controller can use any sort of algorithm (like Dijkstra) to calculate paths.



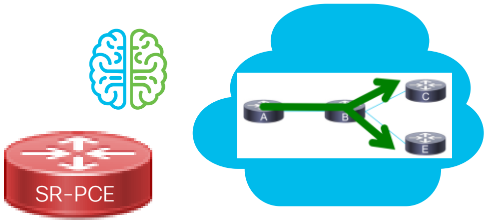
SR-PCE – learning the tree

- SR-PCE also needs to know the Tree Root and End-points.
  - This can be defined by an operator.
  - Dynamically through a protocol, like BGP Auto Discovery (AD).



SR-PCE – computing the tree

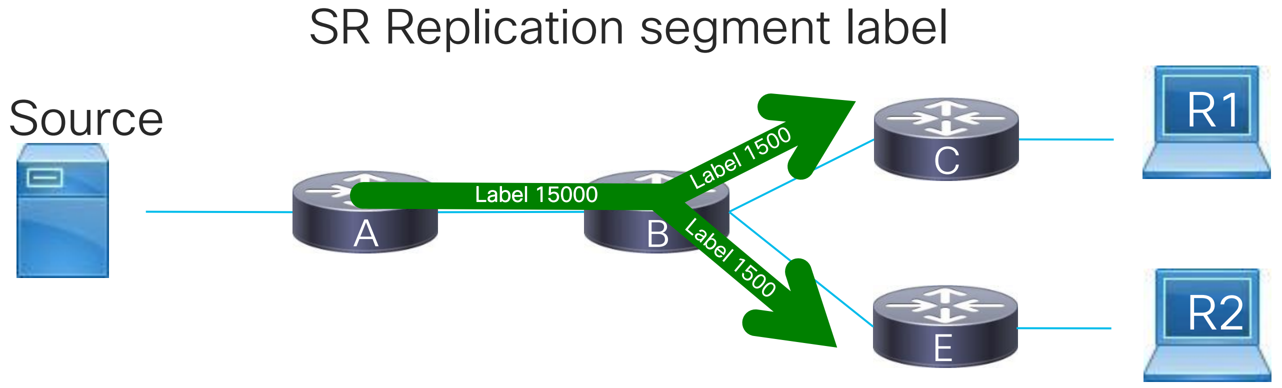
- With the central knowledge at the controller, the tree can be computed according to different metrics and constraints.
  - Optimization objective (metric)
    - IGP / TE / Delay
  - Affinity constraints



SR-PCE – MPLS label allocation

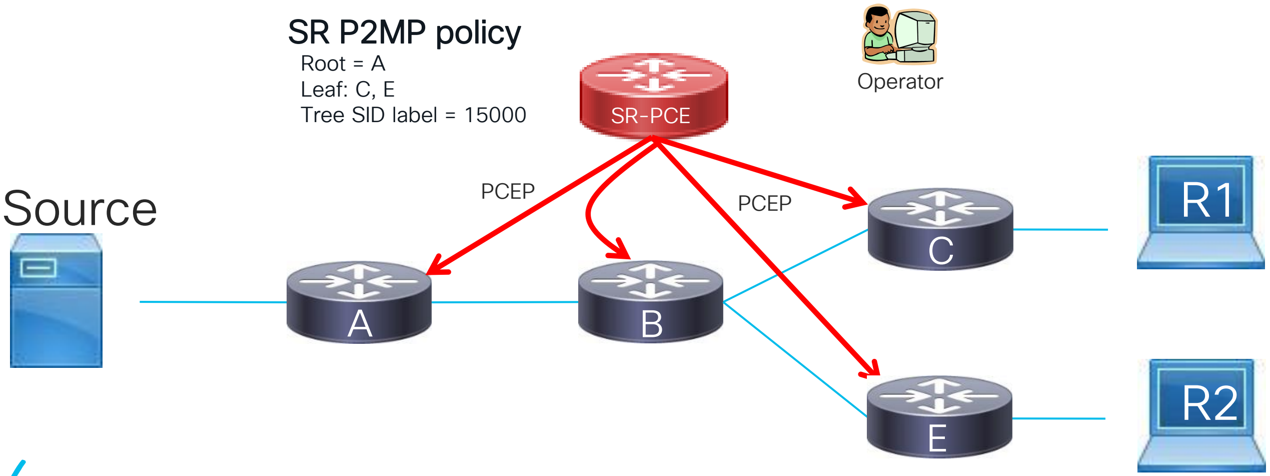
- The allocation and programming of the Label for each TreeSID is done by the SR-PCE.
- The entire Tree can be seen as a Segment.
  - All the routers in the network allocated the same Label range for TreeSID.
  - The controller assigns the same Label for a Tree on all the routers.
  - This means its well known and predictable.
  - It makes it easier to manage and troubleshoot the network.
  - Label range from the SR Local Block (SRLB).

SR-PCE – MPLS label allocation



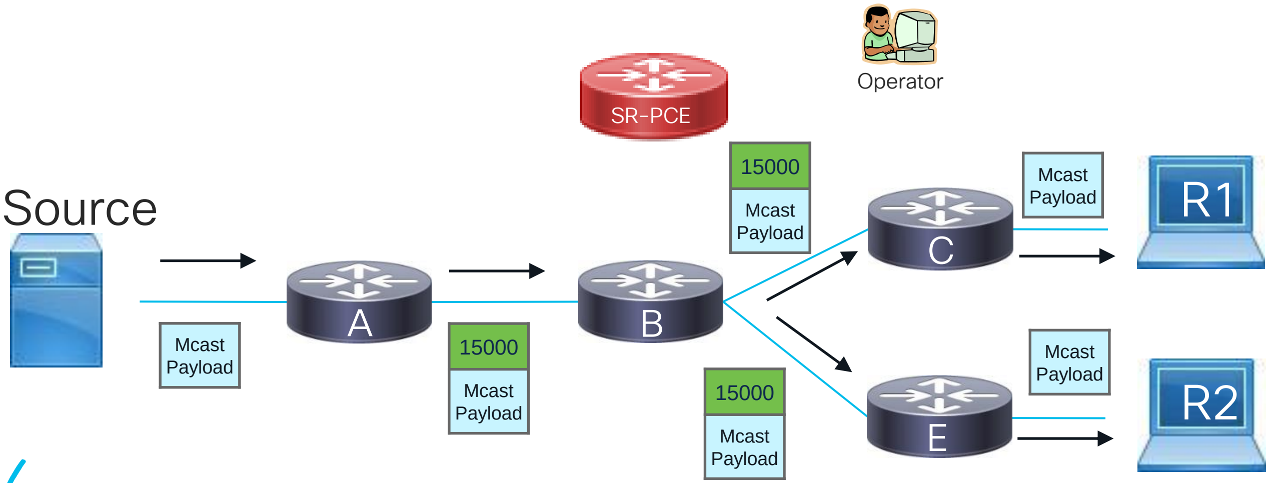
SR-PCE – programming the tree

- SR-PCE needs to program forwarding state on all the routers in the path of the Tree.
  - This is done via Path Computation Element Protocol (PCEP)



Data forwarding

- Once SR-PCE programs whole tree, data will be encapsulated with appropriate assigned label at ingress PE



Tree-SID types

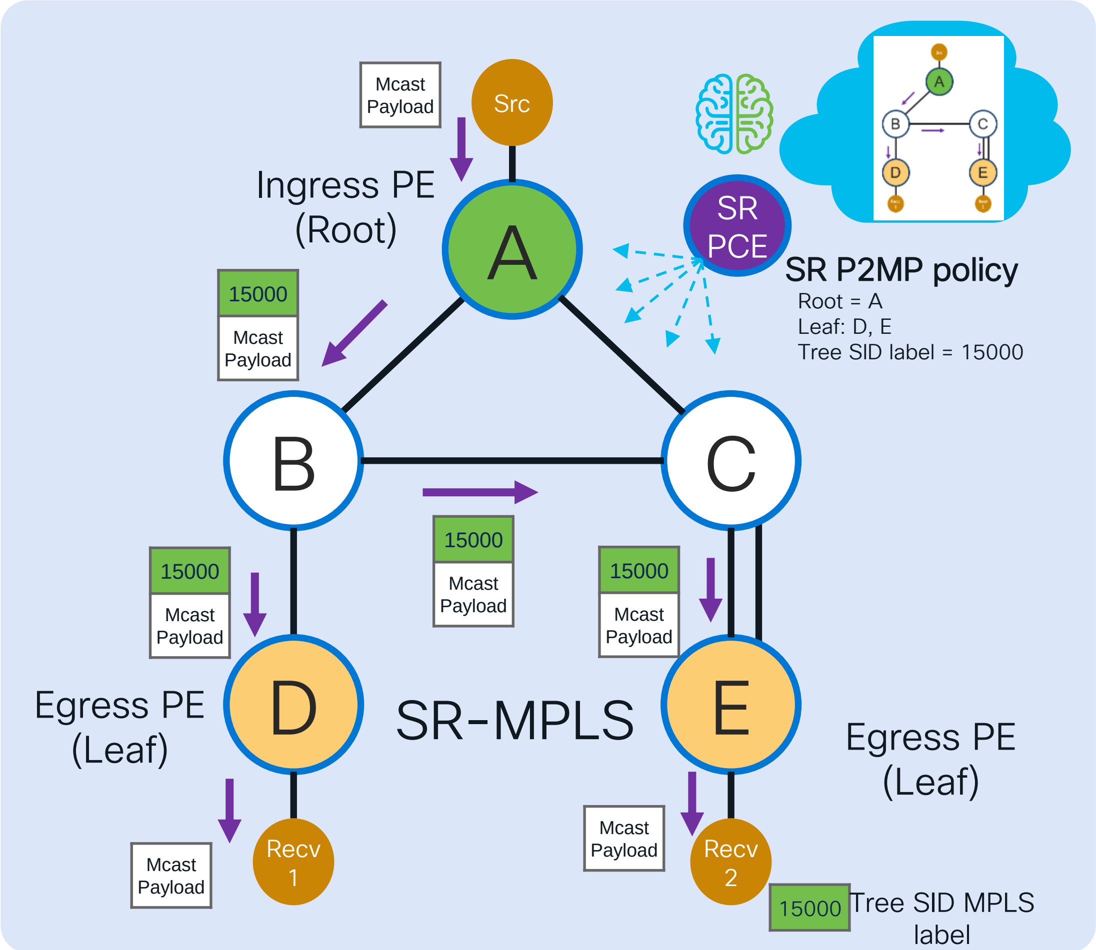
- Depending on how the Root and Leaves of a tree are learnt, the following Tree-SID types exist:
- Static SR P2MP Policies
  - User-defined root and leaves
- Dynamic SR P2MP Policies
  - Dynamically learnt root and leaves

Static SR P2MP policies

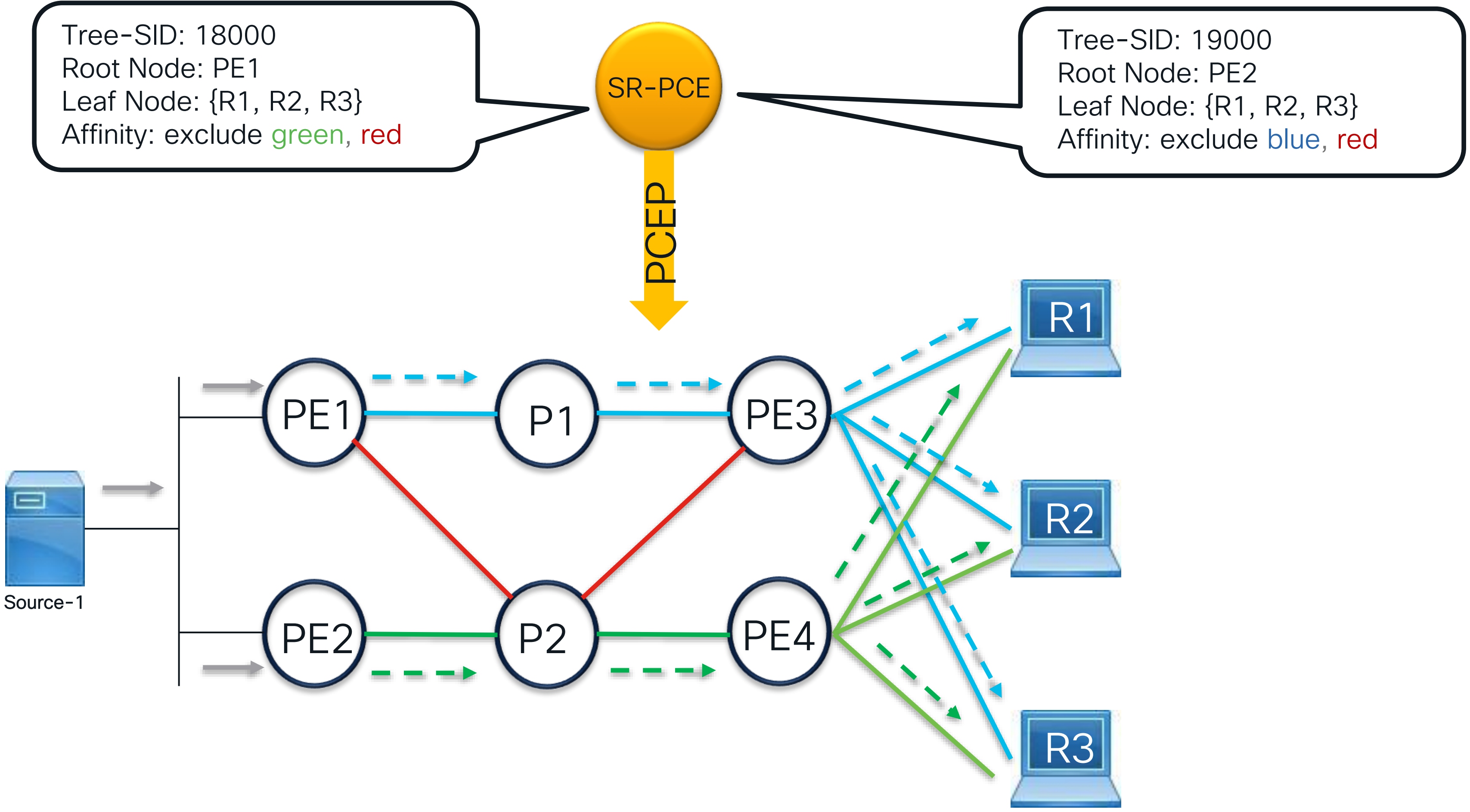
- Highlights:
  - Static Point-to-Multipoint (P2MP) trees to deliver Multi-point services in a SR domain
  - P2MP Provider tunnels (P-tunnels) instantiated via SR P2MP Policy computed by a PCE
  - Pre configured roots and leaf’s are required

Dynamic SR P2MP policies

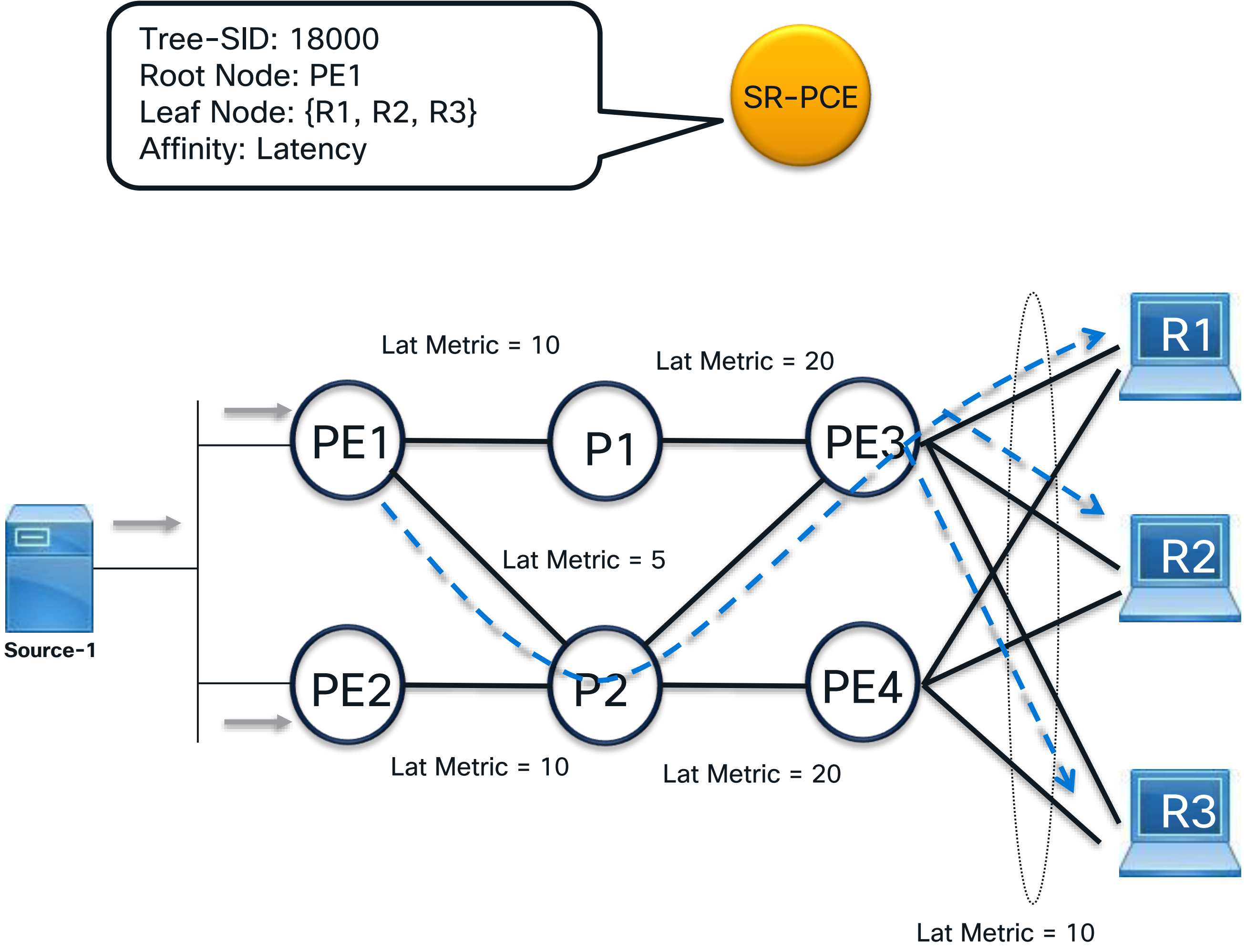
- Highlights:
  - Dynamic Point-to-Multipoint (P2MP) trees to deliver Multi-point services in a SR domain
  - P2MP Provider tunnels (P-tunnels) instantiated via SR P2MP Policy computed by a PCE
  - BGP Auto-Discovery for Distributed VPN end-point discovery and C-multicast flow mapping/signaling
- Use Case / Value Proposition:
  - BGP-based Multicast VPN (MVPN) without state in the core



Tree SID Use case
Disjoint Tree


Tree-SID: 18000
Root Node: PE1
Leaf Node: {R1, R2, R3}
Affinity: exclude green, red
SR-PCE
Tree-SID: 19000
Root Node: PE2
Leaf Node: {R1, R2, R3}
Affinity: exclude blue, red
PCEP
R1
PE1
P1
PE3
R2
Source-1
PE2
P2
PE4
R3

Tree SID Use case
Dynamic Min Latency Trees



- "Performance Measurement" enabled on all links to measure "Unidirectional min Link Delay" metric
- Latency metric advertised in IGPs
- SR-PCE learns link latency metric via BGP-LS

Standardization

Ongoing efforts to standardize in IETF

- https://datatracker.ietf.org/doc/html/draft-wijnands-mpls-mldp-multi-topology-04 - mLDP with flex algo
- https://datatracker.ietf.org/doc/html/draft-ietf-spring-sr-replication-segment-07 - SR Replication for multi-point Service Delivery
- https://datatracker.ietf.org/doc/html/draft-ietf-pim-sr-p2mp-policy – Segment routing point to multipoint policy
- https://datatracker.ietf.org/doc/draft-ietf-pce-sr-p2mp-policy/ - PCEP extensions for p2mp sr policy

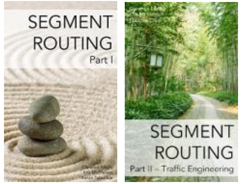
Stay up to date

Segment Routing
www.segment-routing.net/

LinkedIn
www.linkedin.com/groups/8266623

Twitter
www.twitter.com/SegmentRouting

Facebook
www.facebook.com/SegmentRouting/



Segment Routing, Part I / II Textbooks
Available on Kindle and in paperback

Technical session surveys

- Attendees who fill out a minimum of four session surveys and the overall event survey will get Cisco Live branded socks!
- Attendees will also earn 100 points in the Cisco Live Game for every survey completed.
- These points help you get on the leaderboard and increase your chances of winning daily and grand prizes.

Cisco learning and certifications

From technology training and team development to Cisco certifications and learning plans, let us help you empower your business and career. www.cisco.com/go/certs

Pay for Learning with Cisco Learning Credits

(CLCs) are prepaid training vouchers redeemed directly with Cisco.

Learn

Cisco U.
IT learning hub that guides teams and learners toward their goals

Cisco Digital Learning
Subscription-based product, technology, and certification training

Cisco Modeling Labs
Network simulation platform for design, testing, and troubleshooting

Cisco Learning Network
Resource community portal for certifications and learning

Train

Cisco Training Bootcamps
Intensive team & individual automation and technology training programs

Cisco Learning Partner Program
Authorized training partners supporting Cisco technology and career certifications

Cisco Instructor-led and Virtual Instructor-led training
Accelerated curriculum of product, technology, and certification courses

Certify

Cisco Certifications and Specialist Certifications
Award-winning certification program empowers students and IT Professionals to advance their technical careers

Cisco Guided Study Groups
180-day certification prep program with learning and support

Cisco Continuing Education Program
Recertification training options for Cisco certified individuals

Here at the event? Visit us at **The Learning and Certifications lounge at the World of Solutions**

Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Thank you

CISCO Live!

ALL IN

#CiscoLive