Segment Routing

Shivlu Jain
Technical Solution Architect
Blog: www.mplsvpn.info

Agenda

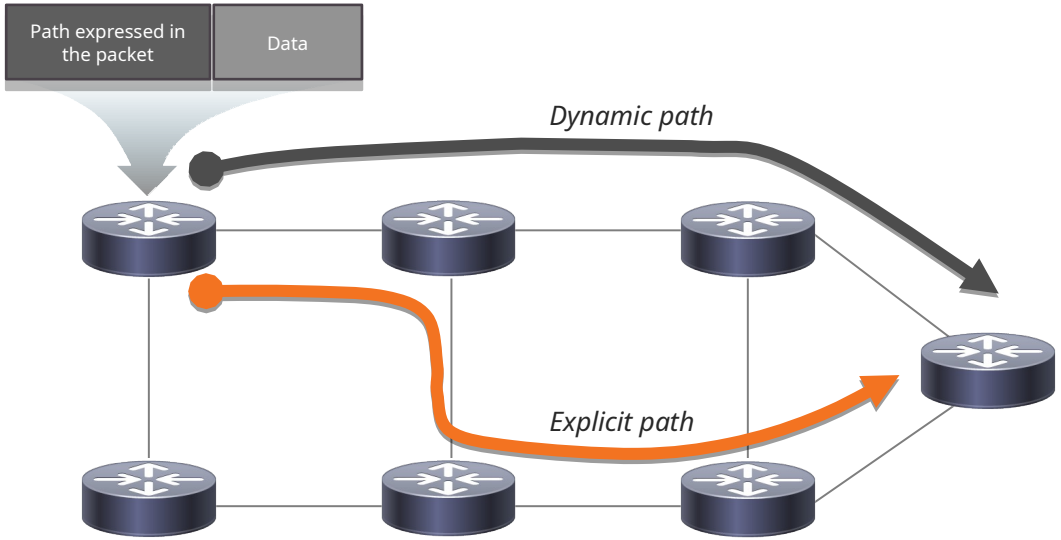
- **Current LDP and RSVP Challenges**
- **Brief About SR**
- **Data Plane Operation**
- **Network Slicing**
- **Data Plane Monitoring**
- **Migration**

Current Challenges of LDP and RSVP

- ✓ Multiple Layer Of Protocols
- ✓ Unable To Utilize ECMP
- ✓ Topology Dependent Protection
- ✓ States Are Everywhere
- ✓ Can’t Utilize Unused Paths To Serve Premium Or Delay Sensitive Customers

Segment Routing

An IP/MPLS source-routing architecture that seeks the **right balance** between **distributed intelligence** and **centralized optimization**



Control Plane

- Routing protocols with extensions (IS-IS,OSPF, BGP)
- SDN controller

Data Plane

- MPLS (segment labels)
- IPv6 (+SR header)

Paths options

- Dynamic (SPF computation)
- Explicit (expressed in the packet)

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

- MPLS: an ordered list of segments is represented as a stack of labels
  - Segment Routing re-uses MPLS data plane without any change
  - Segment represented as MPLS label
  - Applicable to IPv4 and IPv6 address families

- IPv6: an ordered list of segments is encoded in a routing extension header

Global and Local Segments

- **Global Segment**
  - ✓Any node in SR domain understands associated instruction
  - ✓Each node in SR domain installs the associated instruction in its forwarding table
  - ✓MPLS: global label value in Segment Routing Global Block (SRGB)

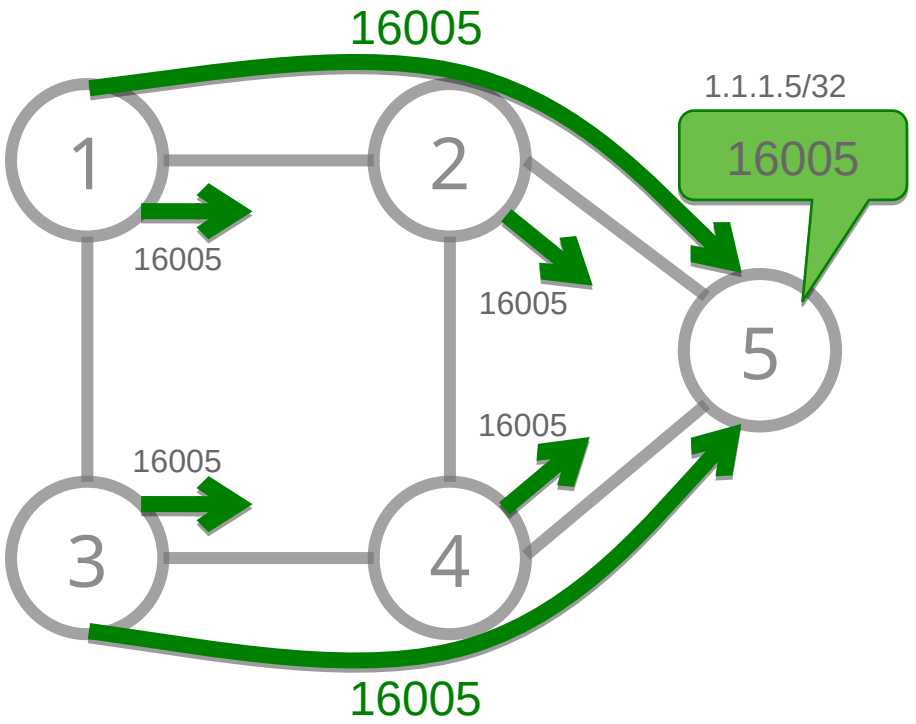
- **Local Segment**
  - ✓Only originating node understands associated instruction
  - ✓MPLS: locally allocated label

IGP segments

- Two basic building blocks distributed by IGP
  - Prefix Segments
  - Adjacency Segments

IGP Prefix Segment

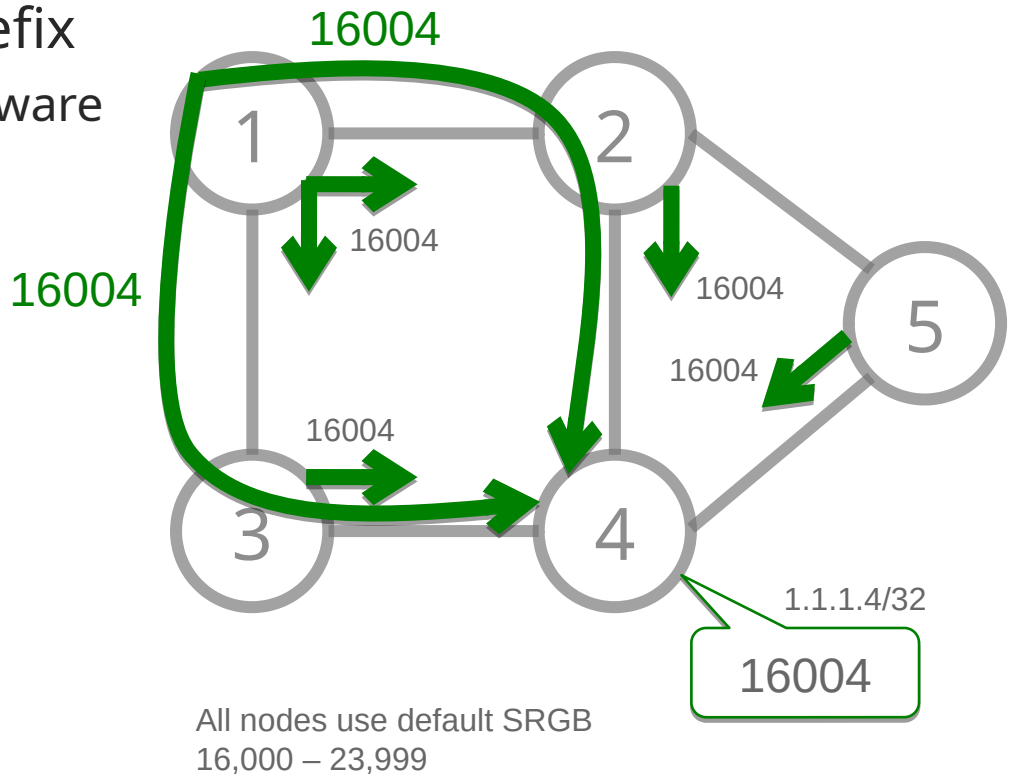
- Shortest-path to the IGP prefix
  - Equal Cost MultiPath (ECMP)-aware
- Global Segment
- Label = 16000 + Index
  - Advertised as index
- Distributed by ISIS/OSPF



All nodes use default SRGB
16,000 – 23,999

IGP Prefix Segment

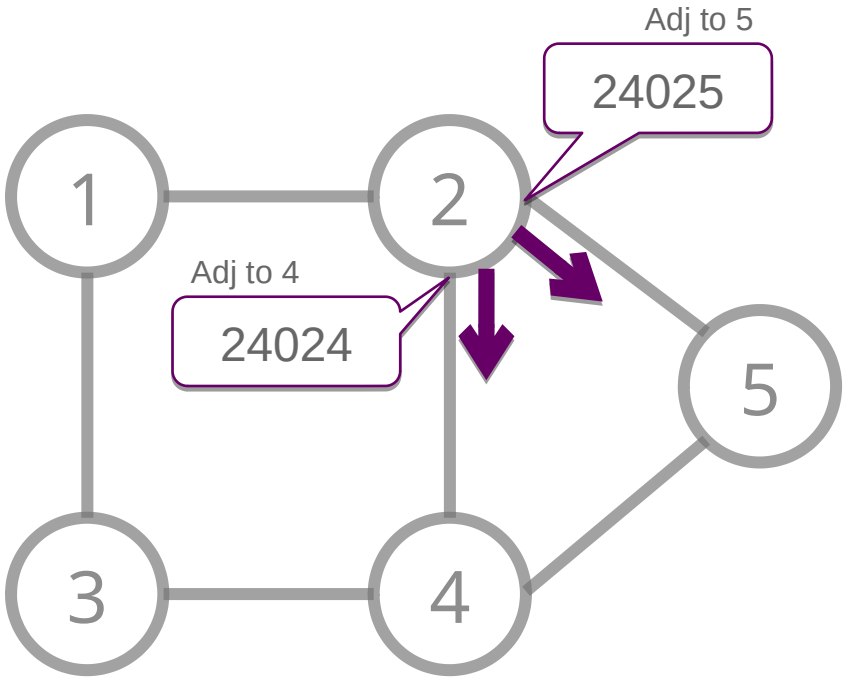
- Shortest-path to the IGP prefix
  - Equal Cost MultiPath (ECMP)-aware
- Global Segment
- Label = 16000 + Index
  - Advertised as index
- Distributed by ISIS/OSPF



All nodes use default SRGB
16,000 – 23,999

IGP Adjacency Segment

- Forward on the IGP adjacency
- Local Segment
- Advertised as label value
- Distributed by ISIS/OSPF

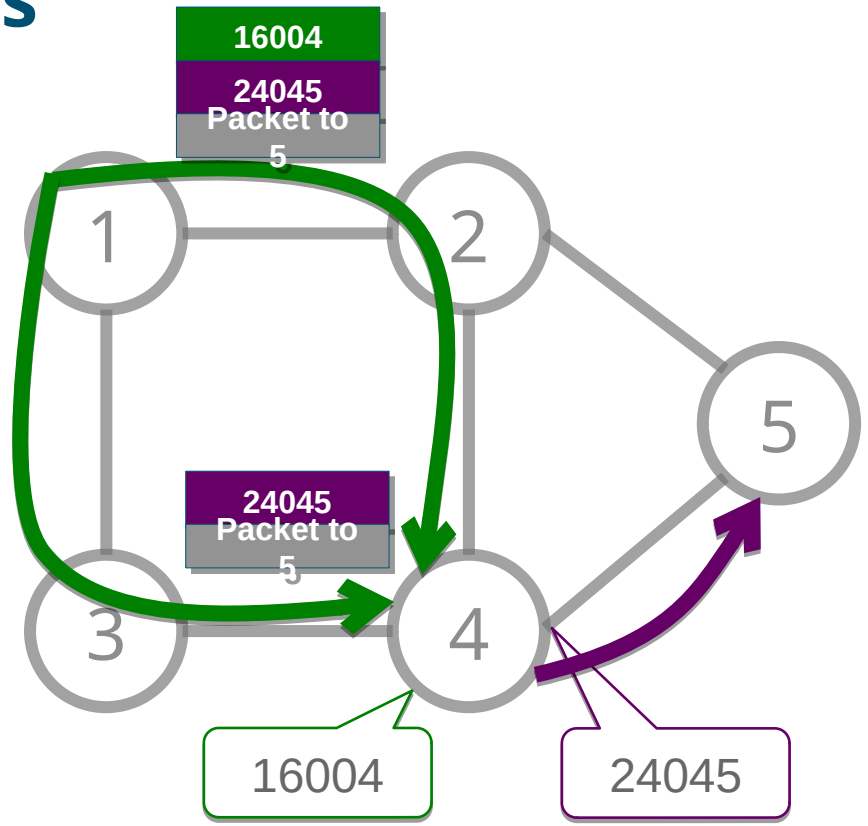


All nodes use default SRGB
16,000 – 23,999

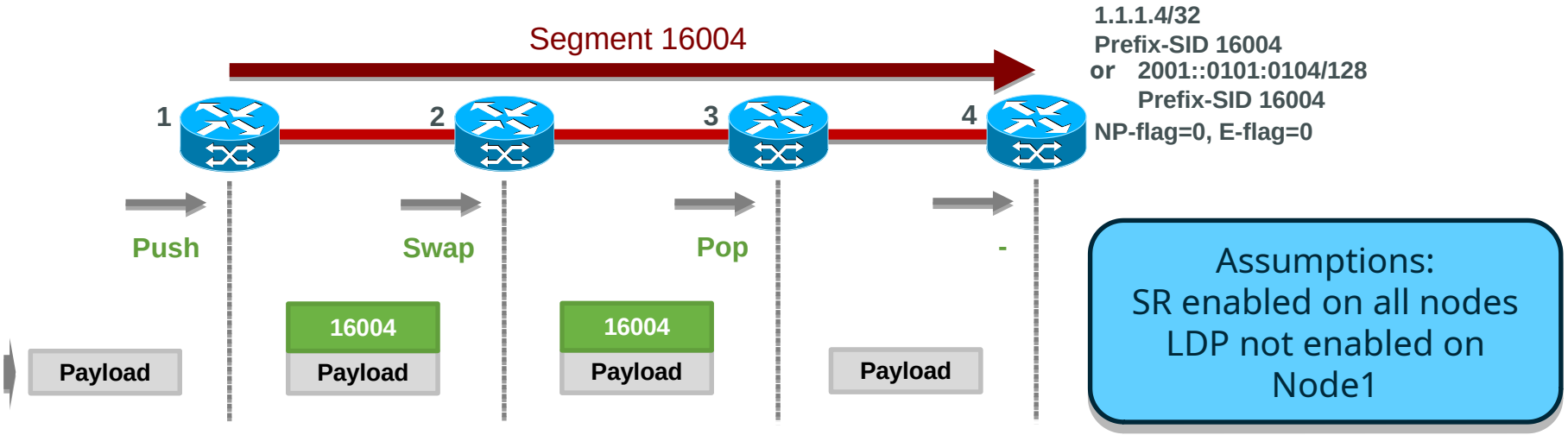
Combining IGP Segments

All nodes use default SRGB
16,000 – 23,999

- Steer traffic on any path through the network
- Path is specified by list of segments in packet header, a stack of labels
- No path is signaled
- No per-flow state is created
- Single protocol: IS-IS or OSPF



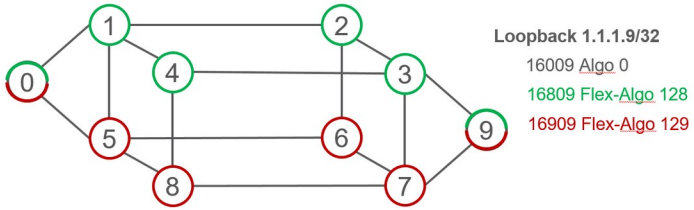
MPLS Data Plane Operations



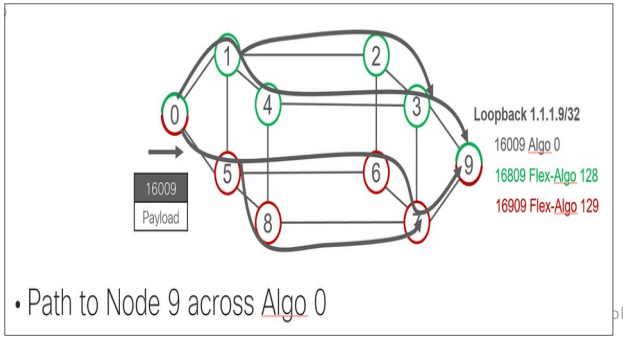
- Node4 advertises its loopback ipv4 prefix 1.1.1.4/32 with attached prefix-SID 16004
  - Or node4 advertises ipv6 prefix 2001::0101:0104/128 with attached prefix-SID 16004
- Node4 requests the default PHP functionality (noPHP-flag=0, ExpNull-flag=0)

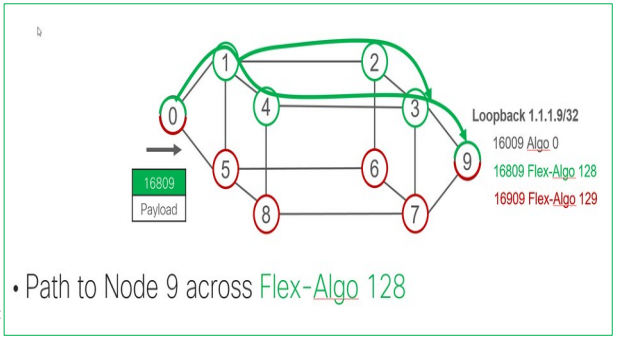
Network Slice

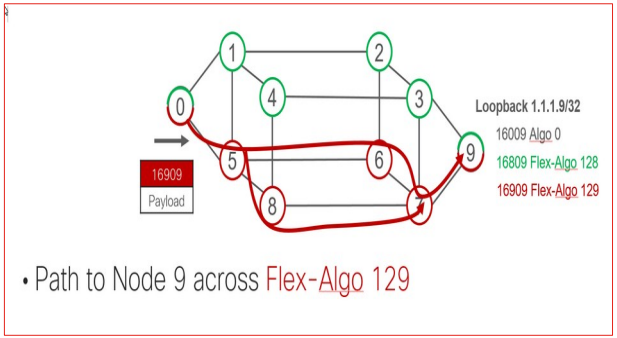
SR IGP Flex Algo



- All the nodes support Algo 0: minimize IGP metric
- Green nodes also support 128: minimize IGP metric
- Red nodes also support 129: minimize Delay







Advanced Monitoring

Problem

LSP black holes

Very difficult to detect, troubleshoot
Ex: CEF looks OK, IGP looks OK, etc. but half the data plane traffic is dropped

Solution

Data plane Monitoring;

Real packets are injected through each combination of neighbors and destinations.
All the ECMP **possibilities** are tested one by one

Technical Benefits

Proactive detection to enable fast reaction time to failures

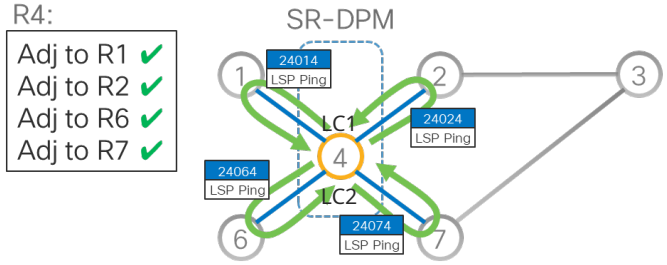
Scalability when comparing with end-to-end probing techniques

Business Benefits

Best Customer Experience

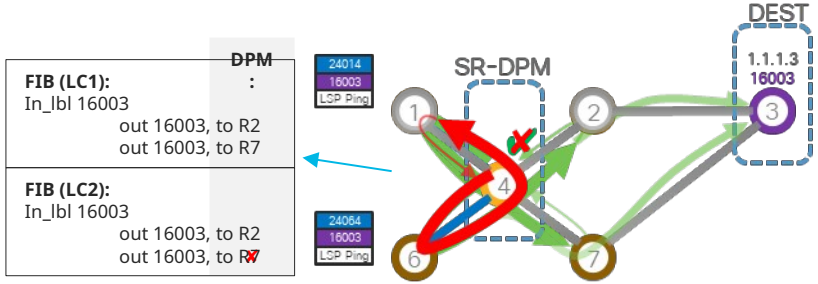
High UpTime

1- Adj-SIDs validation



2- Prefix Validation

Test “top” LC via 2, then via 7 to dest 3
Test “bottom” LC via 2, then via 7 to dest 3



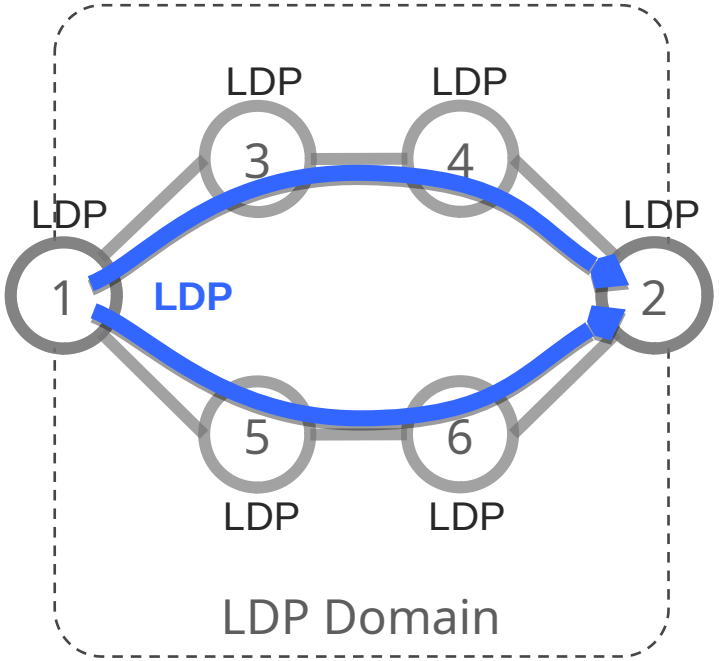
Dataplane FIB corruption: bottom LC **via 7** to dest 3: this entry wrongly points to intf to 1

2009: http://www.mpls-ipn.info/2009/01/missiing-labels-can-forward-routes-to.html

Simplest migration LDP to SR

Assumptions:
- all the nodes can be upgraded to SR
- all the services can be upgraded to SR

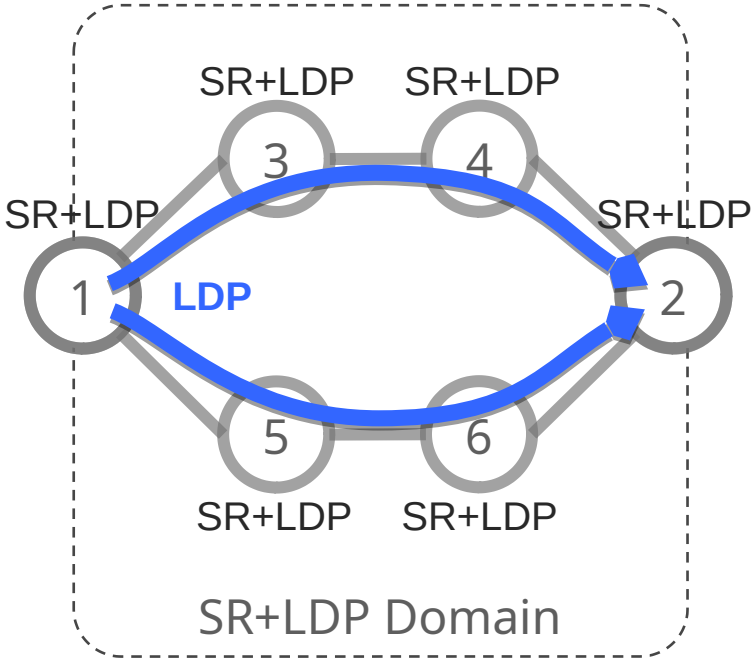
- Initial state: All nodes run LDP, not SR



Simplest migration LDP to SR

Assumptions:
- all the nodes can be upgraded to SR
- all the services can be upgraded to SR

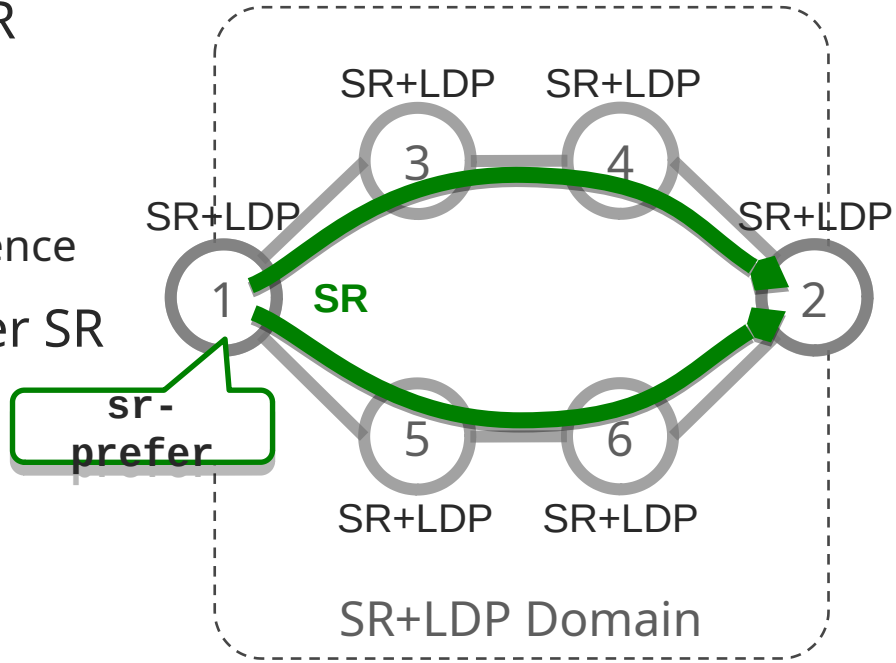
- Initial state: All nodes run LDP, not SR
- Step1: All nodes are upgraded to SR
  - In no particular order
  - leave default LDP label imposition preference



Simplest migration LDP to SR

Assumptions:
- all the nodes can be upgraded to SR
- all the services can be upgraded to SR

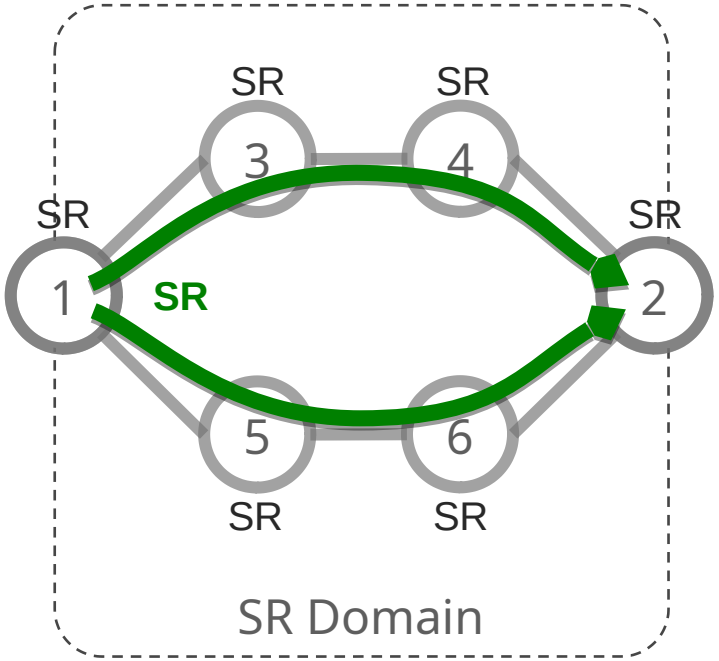
- Initial state: All nodes run LDP, not SR
- Step1: All nodes are upgraded to SR
  - In no particular order
  - leave default LDP label imposition preference
- Step2: All PEs are configured to prefer SR label imposition
  - In no particular order



Simplest migration LDP to SR

Assumptions:
- all the nodes can be upgraded to SR
- all the services can be upgraded to SR

- Initial state: All nodes run LDP, not SR
- Step1: All nodes are upgraded to SR
  - In no particular order
  - leave default LDP label imposition preference
- Step2: All PEs are configured to prefer SR label imposition
  - In no particular order
- Step3: LDP is removed from the nodes in the network
  - In no particular order
- Final state: All nodes run SR, not LDP



SR In Slide

Terminology

SR Segment
- Connects 2 points in SR Domain
- Represented By SID

Segment Identifier (SID)
- Identifies Path
- Global or Local

Labels

Prefix Sid
- Global – Least Cost IGP Path

Adjacency Sid
- IGP Adjacency

Binding
- Tunnel

BGP Prefix Sid
- BGP-LU

SR-TE

Simple, Automated and Scalable
- No State In The Core: “It’s Packet”
- No Tunnel Interface: “SR Policy”
- No Headend Config: “On-Demand”
- Automated Steering

Multi Domain
- SR-PCE
- Binding SID

Prog, DPM & Network Slicing

Data Plane Performance Monitoring

Per Domain Slice
- End-to-end packet latency
- Separation of use case traffic

Network Simplification Journey

	Legacy	Next Gen
Technology Arch.	Unified MPLS	Segment Routing
Provisioning	CLI Driven	NETCONF/YANG
Programmability	None	SR-PCE
Telemetry	SNMP	Model-Driven
Services (L2/L3 VPN)	LDP, BGP	BGP
Scaling Mechanism	BGP-LU	SR Segment Routing
TE, FRR	RSVP-TE	
MPLS Overlay Protocol	RSVP-TE, LDP	
Connectivity Protocol	IGP	


CISCO