

The Cisco Live! logo features the word "CISCO" in a bold, black, sans-serif font, followed by "Live!" in a black, cursive script font. The background is a vibrant, multi-colored abstract pattern of overlapping, wavy bands in shades of red, orange, yellow, green, and blue, radiating from a bright white center on the right side.

CISCO *Live!*

Let's go

#CiscoLive



The bridge to possible

# Preparing for a Successful Segment Routing Deployment

Jose Liste, Technical Marketing Engineer  
BRKMPLS-2135

CISCO *Live!*

#CiscoLive



# Cisco Webex App

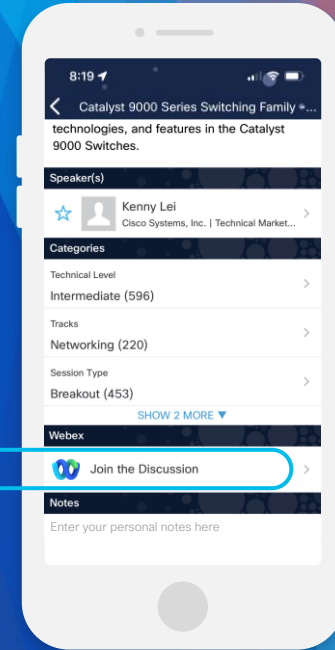
## Questions?

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 9, 2023.



<https://cislive.ciscoevents.com/cislivebot/#BRKMPL-2135>

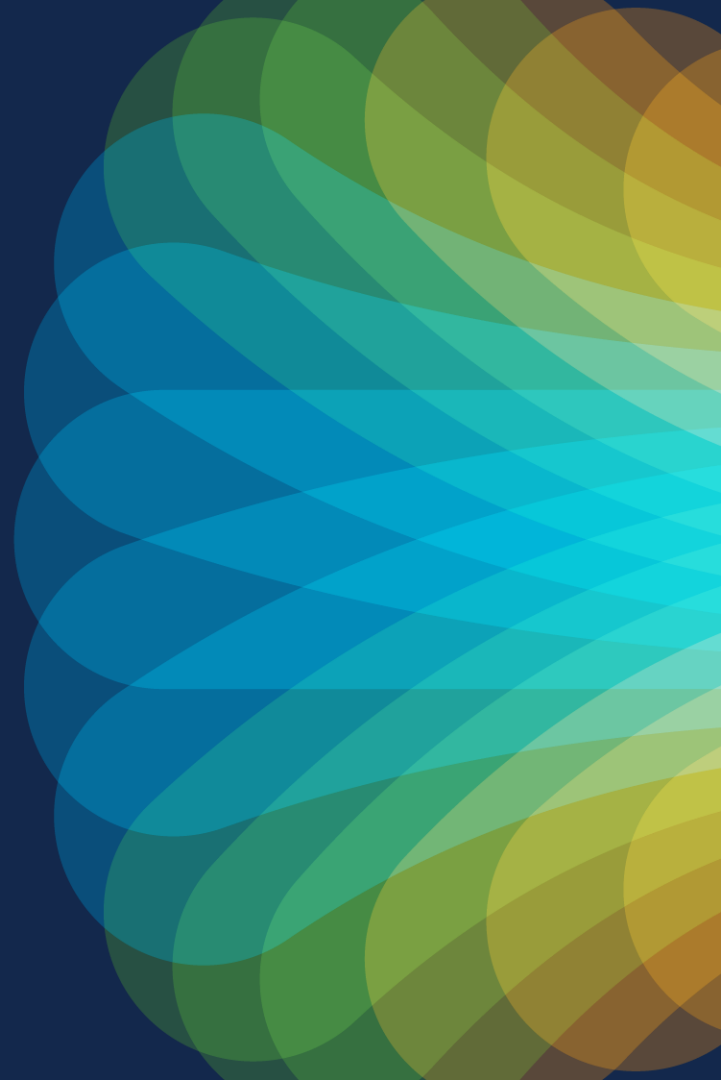
# Abstract

- Segment Routing is a de-facto industry standard architecture adopted by operators of all sizes.
- SR delivers a unified, end-to-end policy-aware network over a stateless IP fabric at uncompromised scale and simplicity.
- In this session, you will learn the fundamental deployment considerations, migration strategies and best practices for a successful rollout of SR in your network.
- Understanding of MPLS and IPv6 architectures recommended  
Understanding of SR fundamentals recommended

# Agenda

- Introduction
- Enabling SR
- SRGB planning
- Seamless Deployment
- SR BGP
- Prefix SID Redistribution
- Color-Aware Routing / SRTE
- SR-PCE
- Flexible Algorithm
- Conclusion

# Network Evolution with Segment Routing



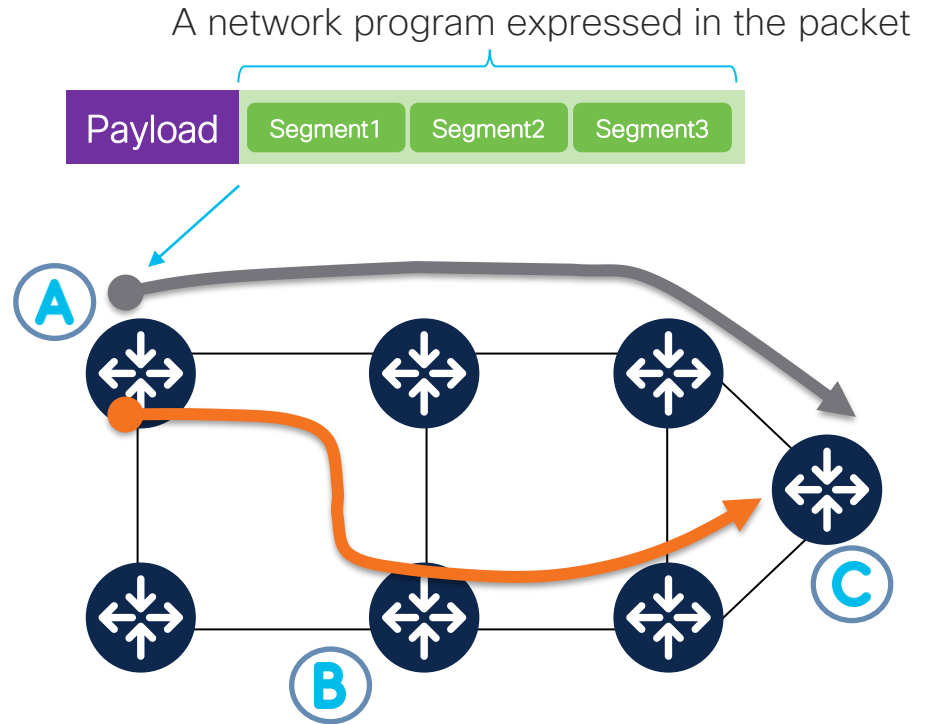
# What is Segment Routing?

- SR is an architecture that leverages the **Source Routing paradigm**
- Provides a **stateless IP fabric** with uncompromised **simplicity** and **scale**
- Delivers a unified, **end-to-end policy-aware** network infrastructure



# What is Segment Routing?

- An ingress node steers a packet through an **ordered list of instructions**, called **segments**
- A segment is locally defined and executed at a specific location in the network
- A segment can represent ANY function, topological or service-based or user-defined



# One Architecture / Two Data-Plane instantiations

## 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
- One or more segments are encoded with an IPv6 address

Segment Routing

# How does it work?

- **SR-MPLS**: the instantiation of SR on the MPLS data-plane
  - Segment ID (SID) → an **MPLS label** associated with the segment
- **SRv6**: the instantiation of SR on the IPv6 data-plane
  - SID → an **IPv6 address** associated with the segment



# Are there different types of Segments?

## Global Segments

- Segments learnt and **programmed by all nodes** in the SR domain
- SID is **operator-assigned** within the **SR Global Block (SRGB)**
- Example: prefix segment

## Local Segments

- Segments learnt by all nodes in the SR domain but **only programmed by the advertising node**
- SID is **dynamically allocated** by router and option for operator-assigned within the SR Local Block (SRLB)
- Example: adjacency segment, peering segment

# SR Toolkit - Are there different types of Segments?

## 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

# SR Toolkit - Are there different types of Segments?

## Binding SID

LOCAL segment representing an SR Policy  
Forward packet along the path(s) of the associated SR Policy

## Layer 2 Adjacency SID

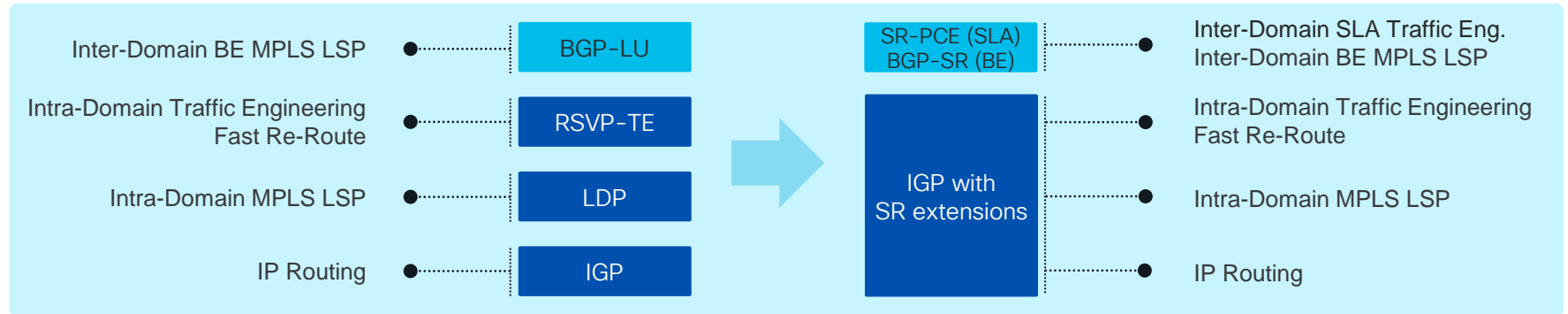
LOCAL segment representing a member interface of an Ethernet Bundle  
Forward packet over the Bundle member interface associated with the segment

# Network Evolution with SR-MPLS

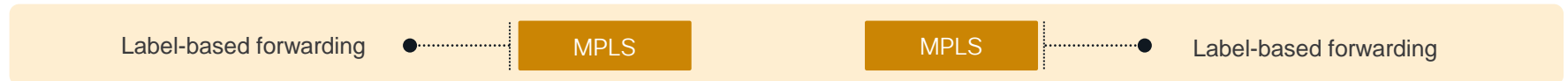
## Service Protocols



## Transport Protocols



## Data-Plane




LDP: Label Distribution Protocol, MP-BGP: Multi-protocol BGP, BGP-LU: BGP Labeled-Unicast, PCE: Path Computation Element, RSVP-TE: Reservation Protocol Traffic Engineering

# Enabling Segment Routing

# IS-IS Segment Routing Configuration

```
router isis 1
  address-family ipv4|ipv6 unicast
  metric-style wide
  segment-routing mpls
!
```



enable SR for the  
MPLS data plane

- MPLS forwarding is enabled on all non-passive IS-IS interfaces
- Adjacency-SIDs are allocated and distributed for all adjacencies
  - Non-protected adj-SIDs and protected adj-SIDs
  - Could be used by TI-LFA backup paths and SR-TE policies

# OSPF Segment Routing Configuration

```
router ospf 1  
  segment-routing mpls
```



- **segment-routing mpls** must be configured to install SIDs received by OSPF in the forwarding table
- MPLS forwarding is enabled on all OSPF interfaces
- Adjacency-SIDs are allocated and distributed for all SR-enabled adjacencies
- Configuration under ospf instance is recommended, but can be customized under each area

# Segment Routing Global Block

- Default Cisco's SRGB is [16,000-23,999]
- A non-default SRGB can be configured in global configuration

# Segment Routing Global Block (SRGB)

## Example

```
segment-routing
global-block 18000 19999
!
router ospf 1
segment-routing mpls
```

Configure a non-default  
global SRGB  
18,000 – 19,999

```
RP/0/0/CPU0:xrvr-1#show mpls label table detail
Table Label Owner State
-----
<...snip...>
0 18000 OSPF(A):ospf-1 InUse No
(Lbl-blk SRGB, vers:0, (start_label=18000, size=2000))
<...snip...>
```

OSPF SRGB


Start\_label = 18,000

Size = 2,000

Non-default SRGB  
label block allocation  
for OSPF  
[ 18,000 – 19,999 ]

# Prefix-SID / Node-SID Configuration

```
router isis 1
 interface Loopback0
  address-family ipv4|ipv6 unicast
  prefix-sid {absolute|index} {<SID value>|<SID index>}
```




```
router ospf 1
 area 0
 interface Loopback0
  prefix-sid {absolute|index} {<SID value>|<SID index>}
```



- Prefix-SID can be specified using:
  - an **absolute** value within the SRGB (“global mode”)
  - or an **index** (offset) from the lower bound of the SRGB.


# Prefix-SID / Node-SID Configuration

```
router isis 1
 interface Loopback0
  address-family ipv4 unicast
  prefix-sid absolute 16001
```

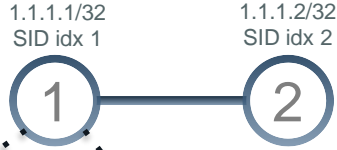


OR

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



# IS-IS example



```
interface Loopback0
  ipv4 address 1.1.1.1 255.255.255.255
!
router isis 1
  address-family ipv4 unicast
  metric-style wide
  segment-routing mpls
!
interface Loopback0
  address-family ipv4 unicast
  prefix-sid absolute 16001
  !! Or: prefix-sid index 1
```

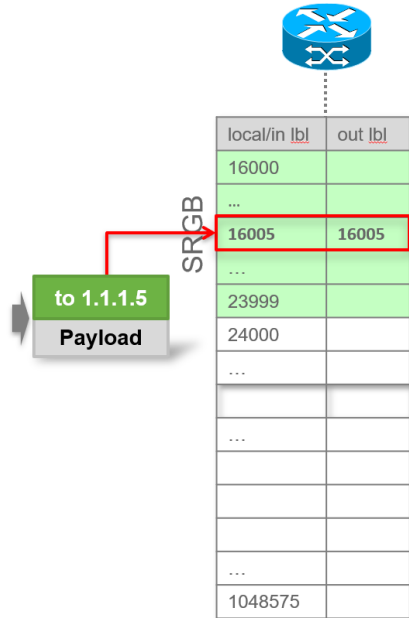
```
RP/0/0/CPU0:xrvr-1#show isis database verbose xrvr-1

IS-IS 1 (Level-2) Link State Database
LSPID                LSP Seq Num  LSP Checksum  LSP Holdtime  ATT/P/OL
xrvr-1.00-00         * 0x0000039b  0xfc27        1079          0/0/0
  Area Address: 49.0001
  NLPID:           0xcc
  Hostname:        xrvr-1
  IP Address:      1.1.1.1
  Router Cap:     1.1.1.1, D:0, S:0
    Segment Routing: I:1 V:1, SRGB Base: 16000 Range: 8000
<...>
Metric: 0          IP-Extended 1.1.1.1/32
  Prefix-SID Index: 1, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0
  Prefix Attribute Flags: X:0 R:0 N:1
  Source Router ID: 1.1.1.1
<...>
```

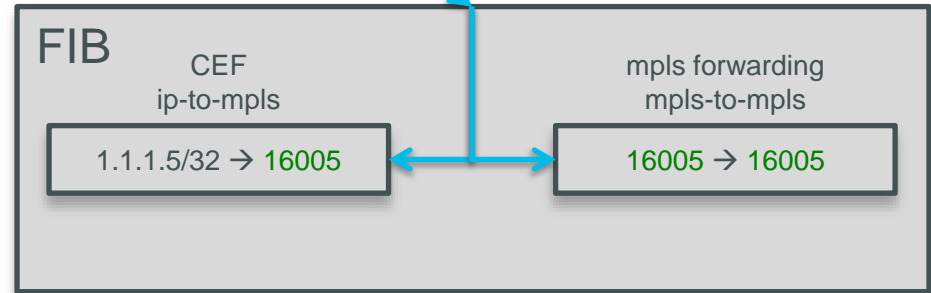
Prefix-SID of prefix 1.1.1.1/32

# IGP/SR programming FIB

- This diagram illustrates the FIB behaviors on a SR-enabled node



1.1.1.5/32  
Loc\_lbl 16005  
Out\_lbl 16005



Loc\_lbl: local label, allocated by local node  
Out\_lbl: outgoing label

16005: SR label

# SR Global Block (SRGB)

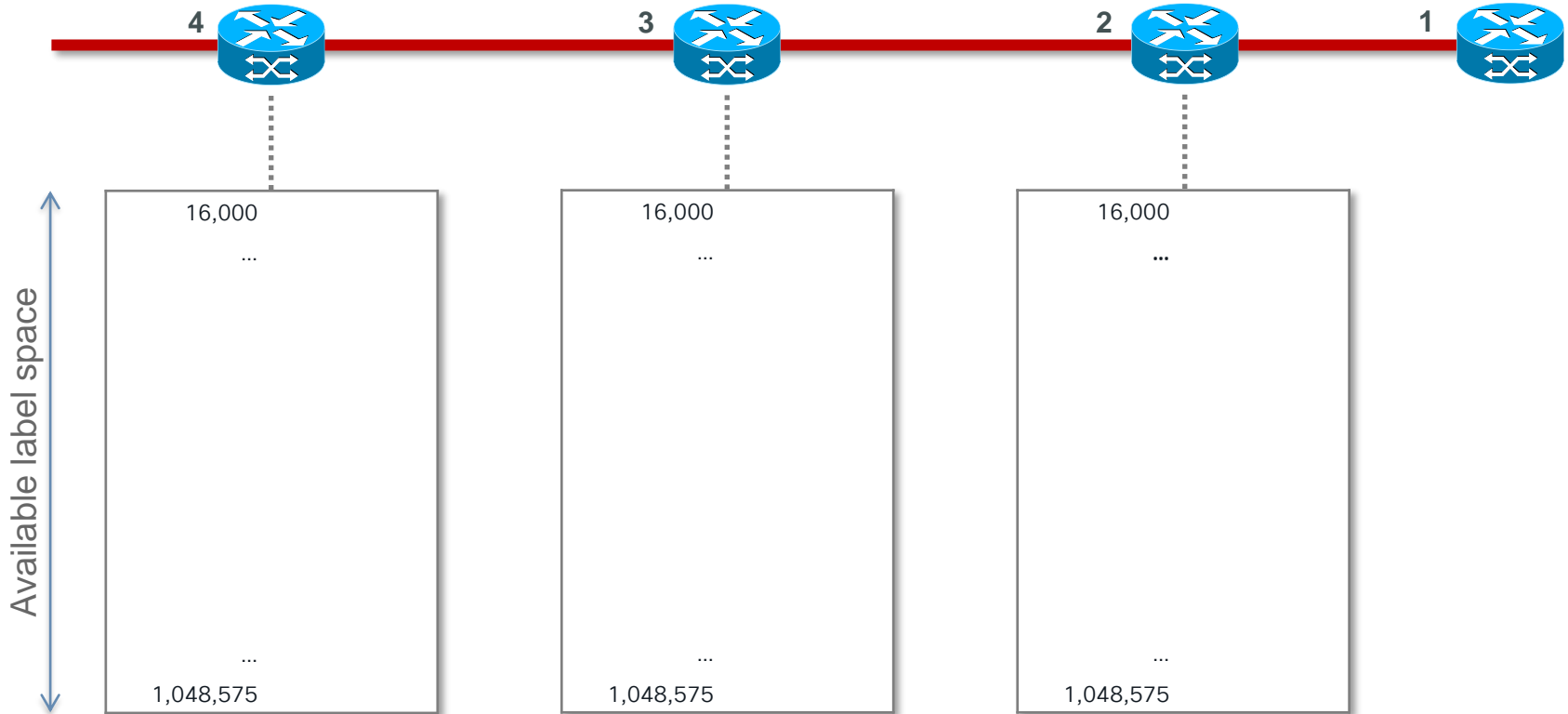
# Segment Routing Global Block (SRGB)

- Segment Routing Global Block
  - Range of labels reserved for Segment Routing Global Segments
  - Default Cisco's SRGB is 16,000 – 23,999
- A prefix-SID is advertised as a domain-wide **unique** index
- The Prefix-SID index points to a unique label within the SRGB
  - Index is zero based, i.e. first index = 0
  - Label = Prefix-SID index + SRGB base
  - E.g. Prefix 1.1.1.65/32 with prefix-SID index 65 gets label 16065

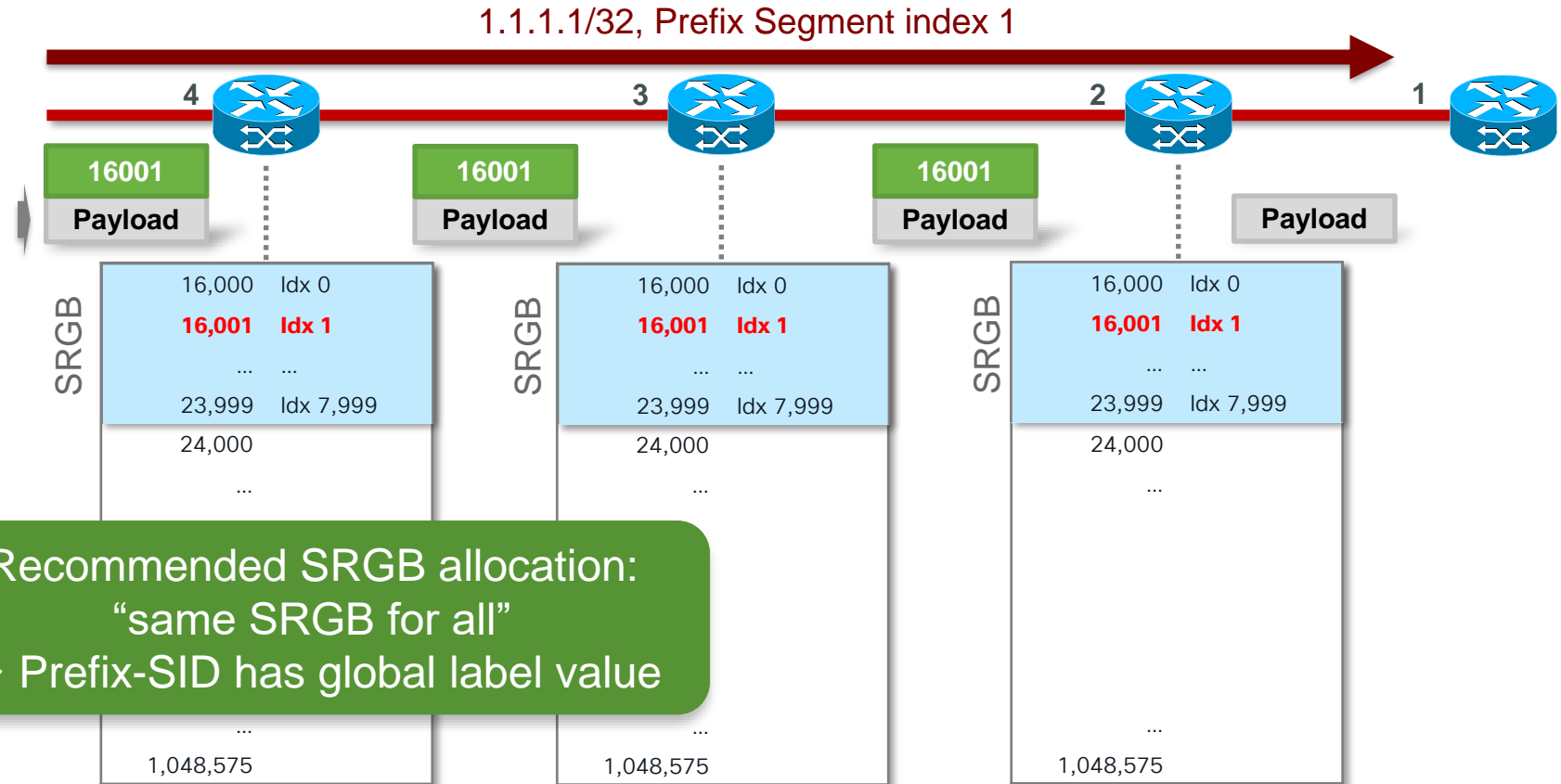
# Segment Routing Global Block (SRGB)

- Strongly recommended to use **same (homogeneous) SRGB on all nodes**
  - Simple, straightforward
  - Global Segment == Global Label value
  - Using different (heterogeneous) SRGBs is supported, but complicates operations for user
- A non-default SRGB can be allocated between 16,000 and 1,048,575
  - Or up to the platform limit, if any
- The size of the SRGB should be equal on all nodes
  - IOS-XR does not impose any constraints on the maximum size

# Segment Routing Global Block (SRGB)

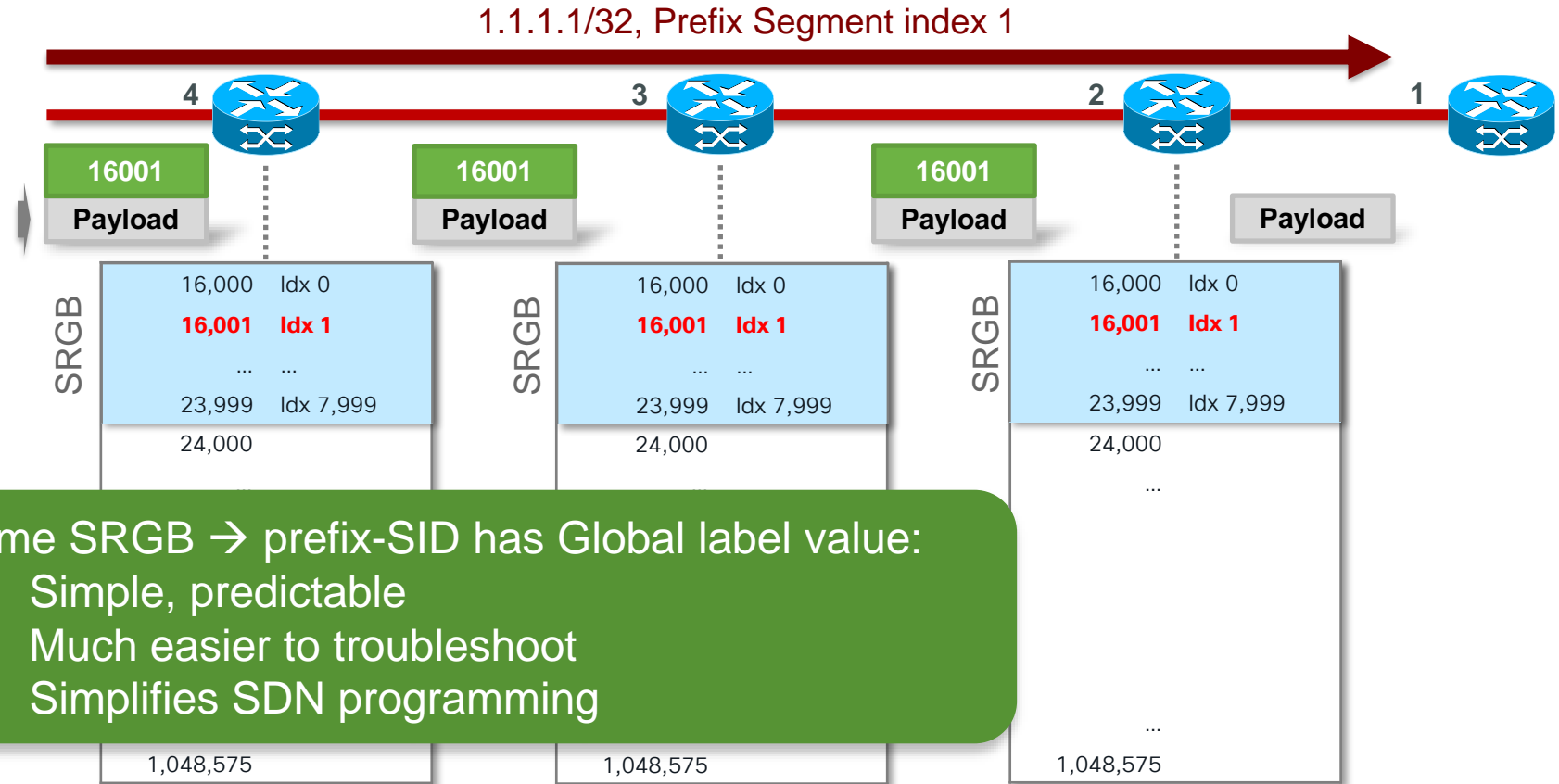


# Homogeneous SRGB allocation (Recommended)



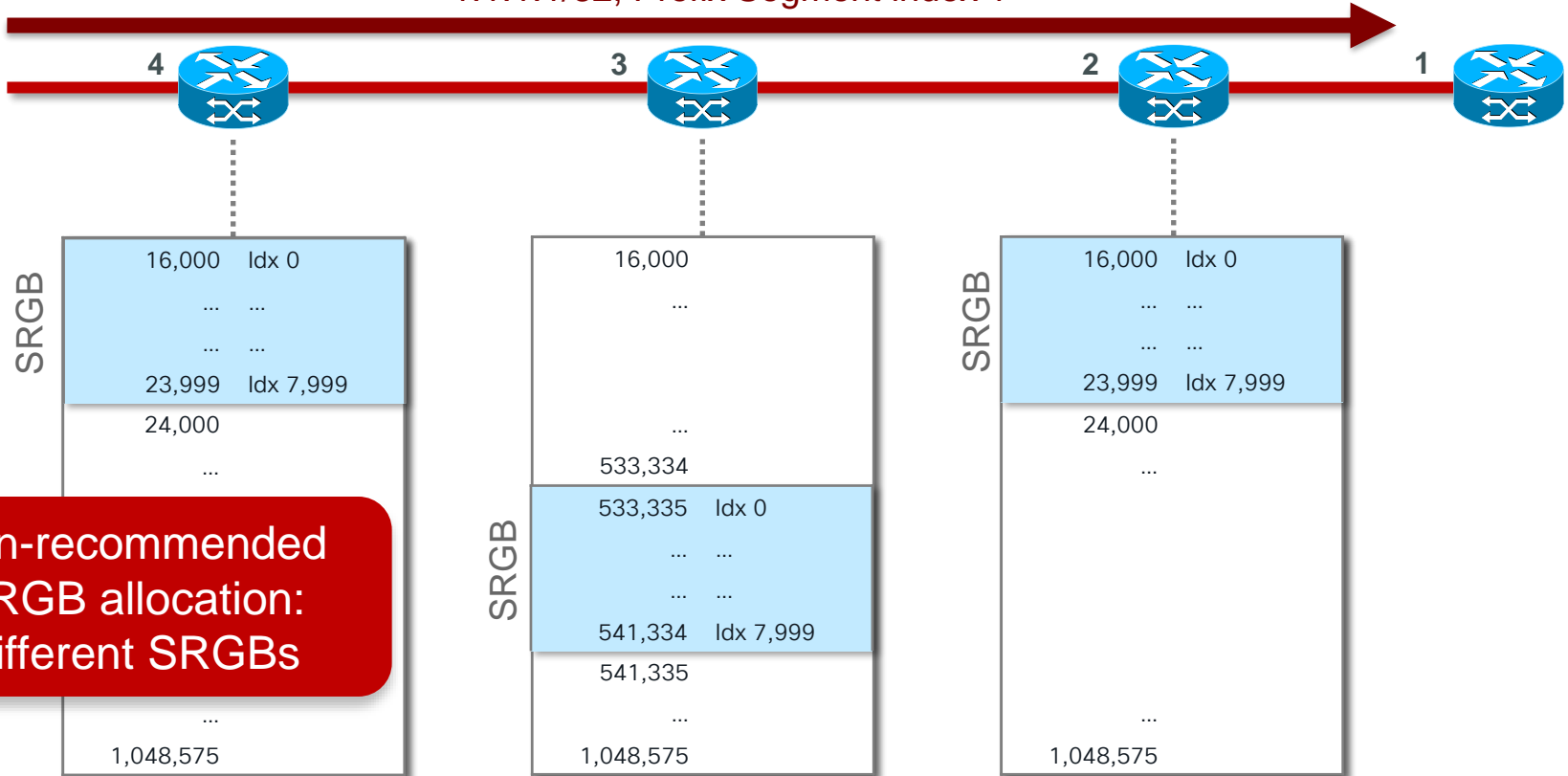
Recommended SRGB allocation:  
"same SRGB for all"  
→ Prefix-SID has global label value

# Homogeneous SRGB allocation (Recommended)



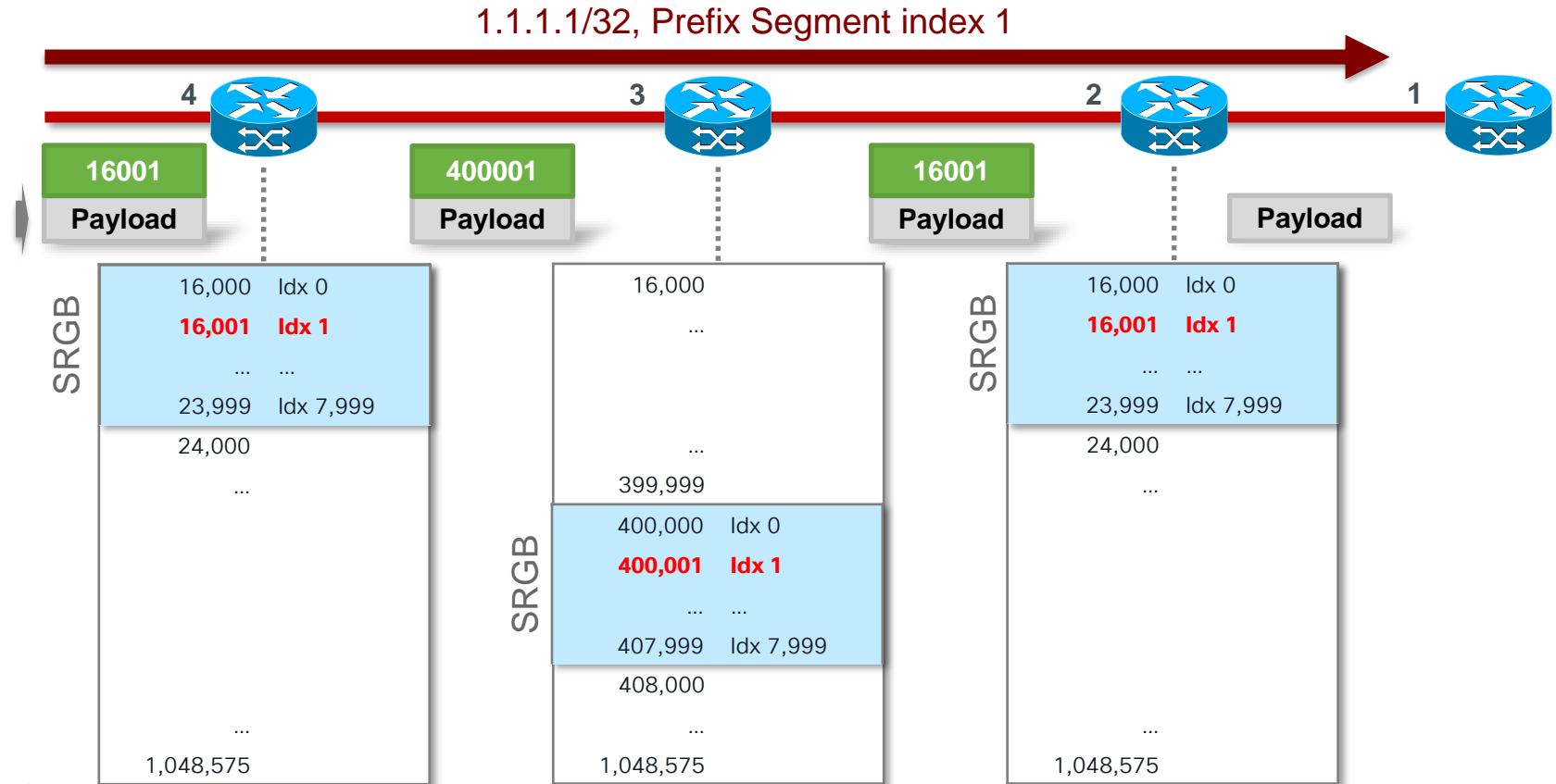
# Heterogeneous SRGB allocation (Not Recommended)

1.1.1.1/32, Prefix Segment index 1



Non-recommended  
SRGB allocation:  
Different SRGBs

# Heterogeneous SRGB allocation (Not Recommended)



# SRGB label range preservation

- LSD **preserves** the default SRGB label range [16,000–23,999]
  - In any Segment Routing capable software release
  - Even if Segment Routing is not enabled
  - Except if the configured mpls label range includes this default range
- LSD allocates **dynamic labels** starting from **24,000**

# SRGB label range preservation

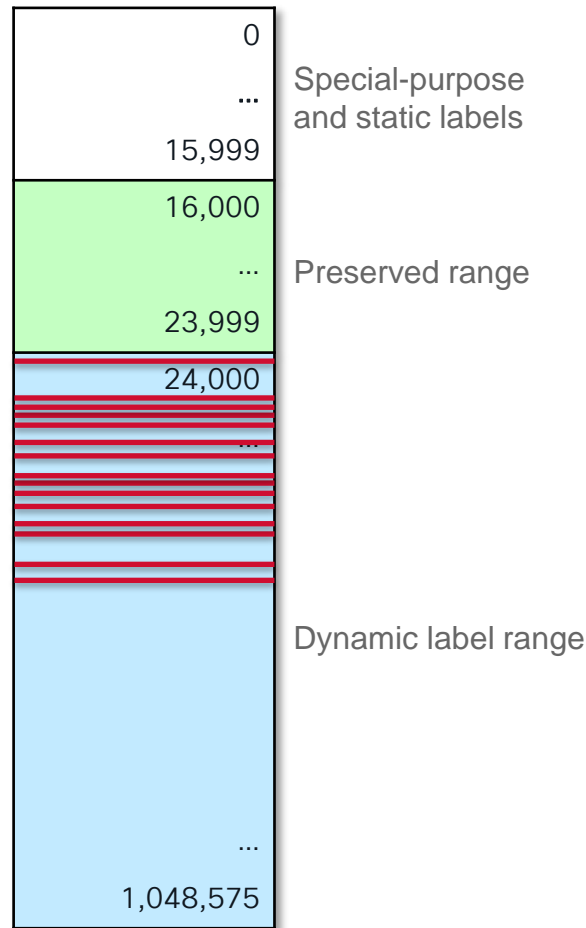
- Preservation of the default SRGB label range makes **future** Segment Routing **activation** possible **without reboot**
  - No labels are allocated from that preserved range. When enabling Segment Routing with default SRGB some time in the future, that label range is available and ready for use

# LSD SRGB preservation - Example

- An example sequence of Segment Routing activation:

1. No Segment Routing enabled, no SRGB allocated

- LSD preserves default SRGB label range
- Dynamic labels are allocated by various MPLS applications (— in diagram)



# LSD SRGB preservation - Example

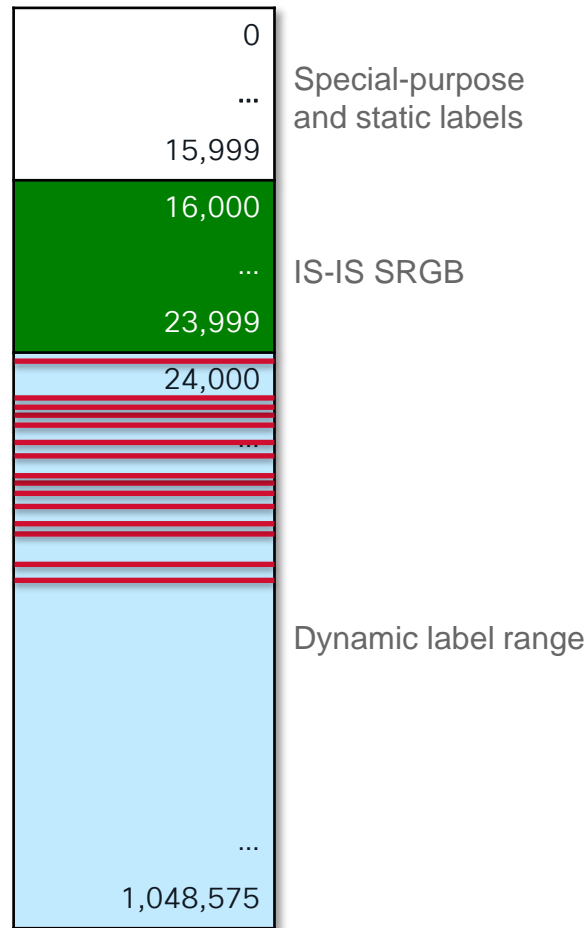
- An example sequence of Segment Routing activation:

## 1. No Segment Routing enabled, no SRGB allocated

- LSD preserves default SRGB label range
- Dynamic labels are allocated by various MPLS applications (— in diagram)

## 2. Sometime later, Segment Routing IS-IS is enabled with default SRGB

- SRGB label range is free (preserved), start using Segment Routing without reboot!



# Segment Routing Global Block (SRGB) Notes

- Allocating an SRGB label range does not mean that all the labels in this range are programmed in the forwarding table
  - But they are reserved (not available for other purposes)
  - A platform may limit the range of local labels that can be used (implicitly limits the number of local labels) – e.g., ASR 900 / ASR 920
  - A platform may limit the number of local labels that can be used (while not limiting the range) – e.g., NCS 5500 / NCS 500 families

# Segment Routing Global Block (SRGB) Notes

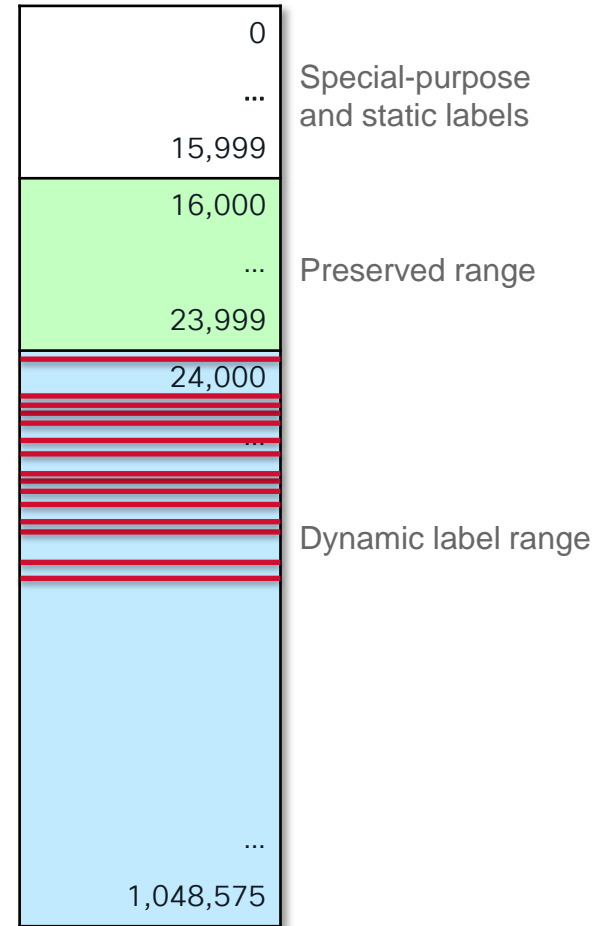
- On a node with multiple routing protocol instances, \*all\* should use the **same** SRGB:
  - SR-enabled IGP instances
  - SR-enabled BGP,

# Segment Routing Global Block (SRGB) Notes

- **Modifying** a SRGB configuration is **disruptive** for traffic
  - And may require a reboot if the new SRGB is not (entirely) available
  - **Allocating a non-default SRGB in the upper part of the MPLS label space increases the chance that the labels are free**

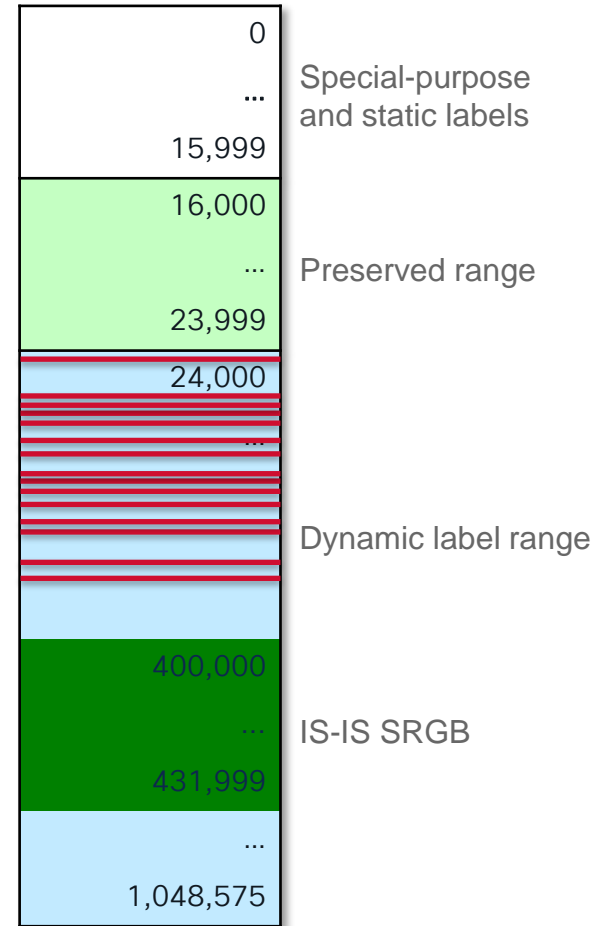
# LSD SRGB allocation - Example

- An example sequence of Segment Routing activation:
  1. No Segment Routing enabled, no SRGB allocated
    - LSD preserves default SRGB label range
    - Dynamic labels are allocated by various MPLS applications ( — in diagram)



# LSD SRGB allocation - Example

- An example sequence of Segment Routing activation:
  1. No Segment Routing enabled, no SRGB allocated
    - LSD preserves default SRGB label range
    - Dynamic labels are allocated by various MPLS applications ( — in diagram)
  2. Sometime later, SR IS-IS is enabled with **non-default SRGB in the upper label range (hence likely unused)**
    - SRGB label range is free, start using SR without reboot!



# SRGB Configuration

- The SRGB can be configured
  - Globally (Recommended)
    - By default, all IGP instances and BGP use this global SRGB
  - Per-IGP (Not Recommended)

```
segment-routing  
global-block 18000 19999
```



Recommended

```
router isis 1  
segment-routing global-block 18000 19999
```

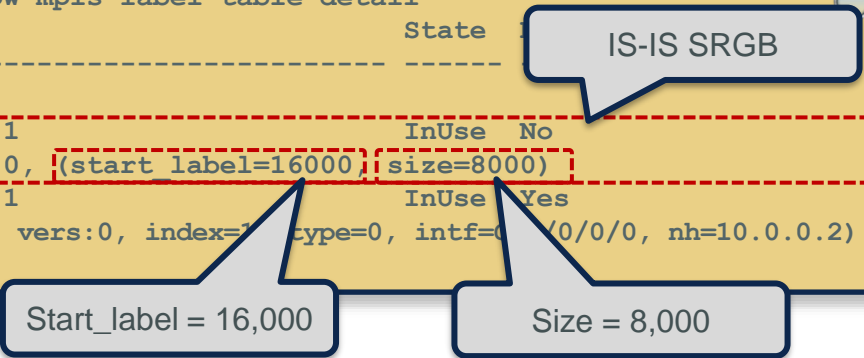


Not Recommended

# Segment Routing Global Block (SRGB)

## Default SRGB

```
RP/0/0/CPU0:xrvr-1#show mpls label table detail
Table Label Owner State
-----
<...snip...>
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=0/0/0/0, nh=10.0.0.2)
```



Default SRGB label block allocation for ISIS  
[ 16,000 – 23,999 ]

# Segment Routing Global Block (SRGB)

## Non-default SRGB Example

```
segment-routing
global-block 18000 19999
!
router isis 1
address-family ipv4 unicast
segment-routing mpls
```



Configure a non-default SRGB  
18,000 – 19,999

```
RP/0/0/CPU0:xrvr-1#show mpls label table detail
Table Label      Owner                               State
-----
<...snip...>
0      18000      ISIS(A) :1                          InUse No
(Lbl-blk SRGB, vers:0, (start_label=18000, size=2000))
0      24000      ISIS(A) :1                          InUse Yes
(SR Adj Segment IPv4, vers:0, index=1, type=0, intf=0/0/0/0, nh=10.0.0.2)
```

ISIS SRGB

Start\_label = 18,000

Size = 2,000

Non-default SRGB  
label block allocation  
for ISIS  
[ 18,000 – 19,999 ]

# SRGB Design Recommendations

# SRGB design

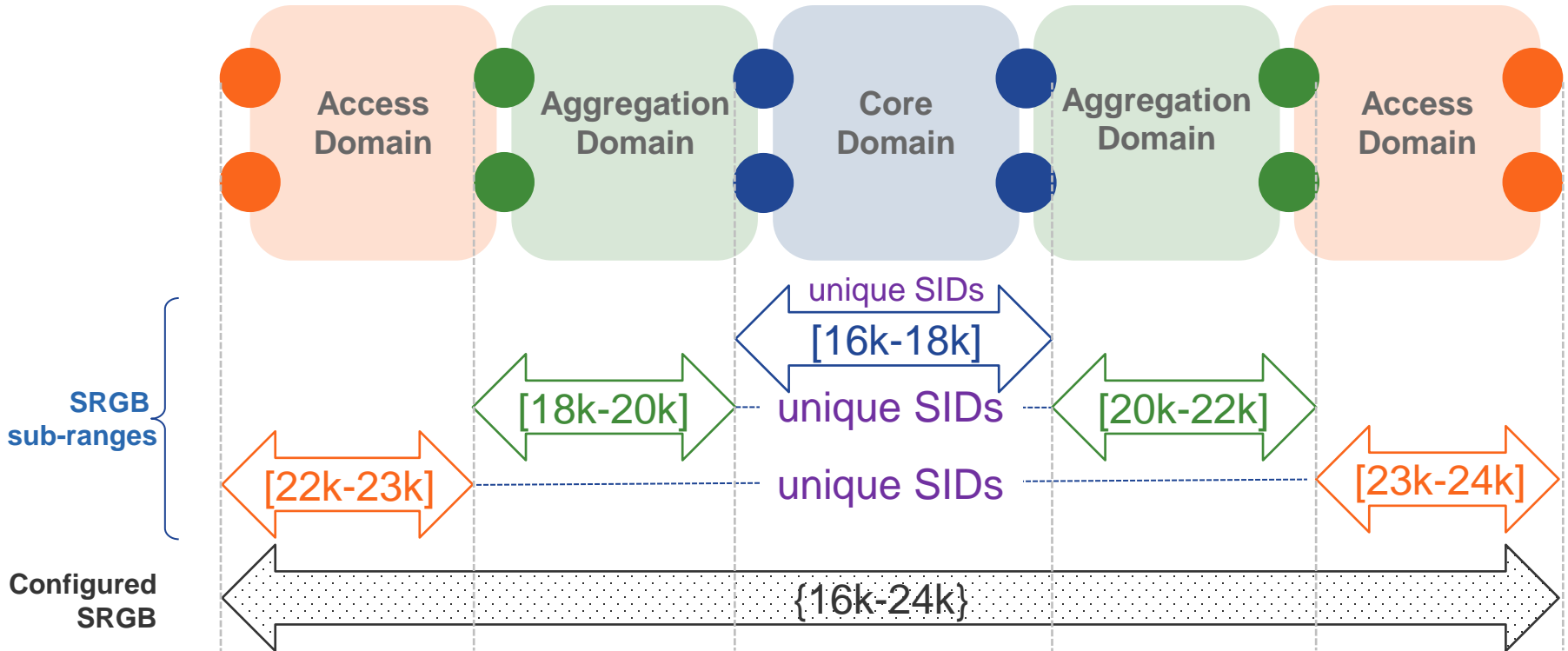
- SRGB planning should aim for the following goals:
  - Goal 1: Homogenous SRGB
  - Goal 2: Unique SID-to-prefix mappings
    - SRGB size > # required SIDs
    - Each SID can be allocated to a single prefix. No SID re-use among prefixes
- Large majority of deployments should be able to meet these goals
  - in some cases, they cannot be achieved due to scaling and platform limitations

# Configured SRGB and SRGB sub-ranges

- For ease of administration and operations, the **configured SRGB** is carved in **administrative sub-ranges**
  - Allocate a sub-range to each domain
  - The configured SRGB is still the entire SRGB, not the SRGB sub-range
- Alternatively, an operator could treat the configured SRGB as a global pool of SIDs
  - A global pool might lead to a more optimal use of the SRGB
  - More complicated administration

# SRGB and SRGB sub-ranges

Notation convention:  
SRGB {XXX-YYY}  
sub-range [QQQ-RRR]



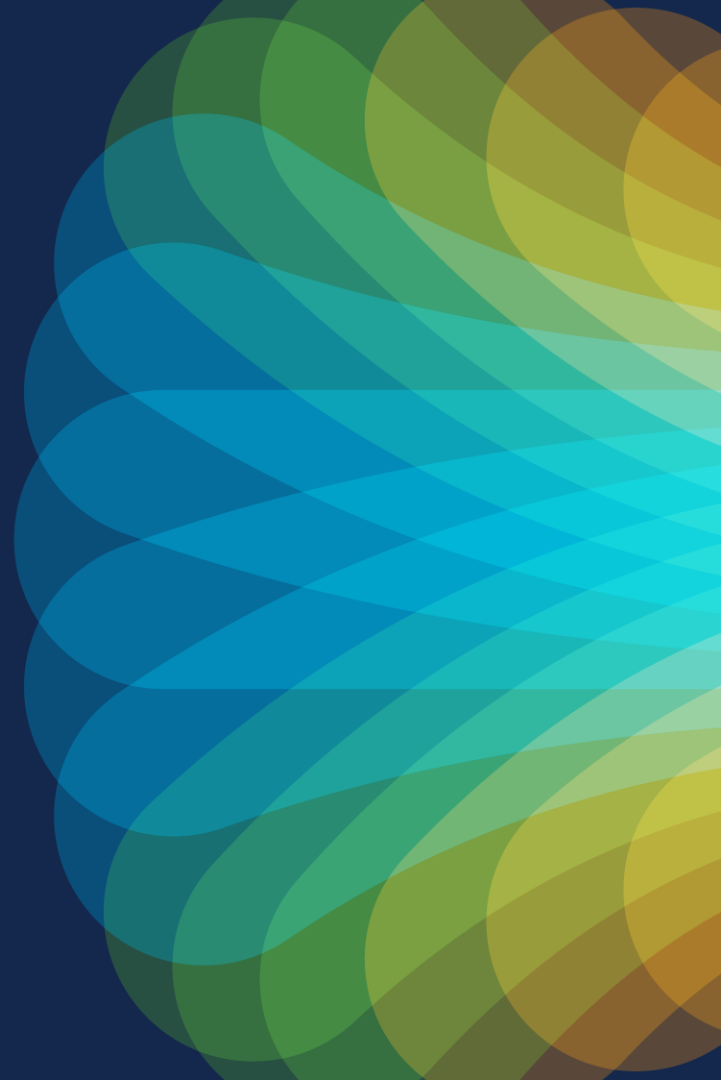
Note: [16k-17k] really means [16000-16999]

# How many SIDs are needed?

- Number of SIDs  $\neq$  number of nodes
- >1 SIDs needed per node
  - [Algo\(0\)](#) Prefix-SID
  - With [Flex-Algo](#), multiple SIDs are mapped to a prefix
  - How many Flex-Algo SIDs?
    - Delay metric: 1 Flex-Algo for low-delay service
    - TE metric: 1 Flex-Algo for e.g. premium service
    - IGP metric: 2 Flex-Algos for dual-plane
    - FA with affinity constraints? E.g., use encrypted links, avoid low BW links

# Seamless Deployment

SR insertion into Brownfield networks



# Seamless Deployment Models

- SR can be deployed incrementally and seamlessly operate in brownfield networks
- SR / LDP Co-existence Deployment Model
- SR / LDP Interworking Deployment Model

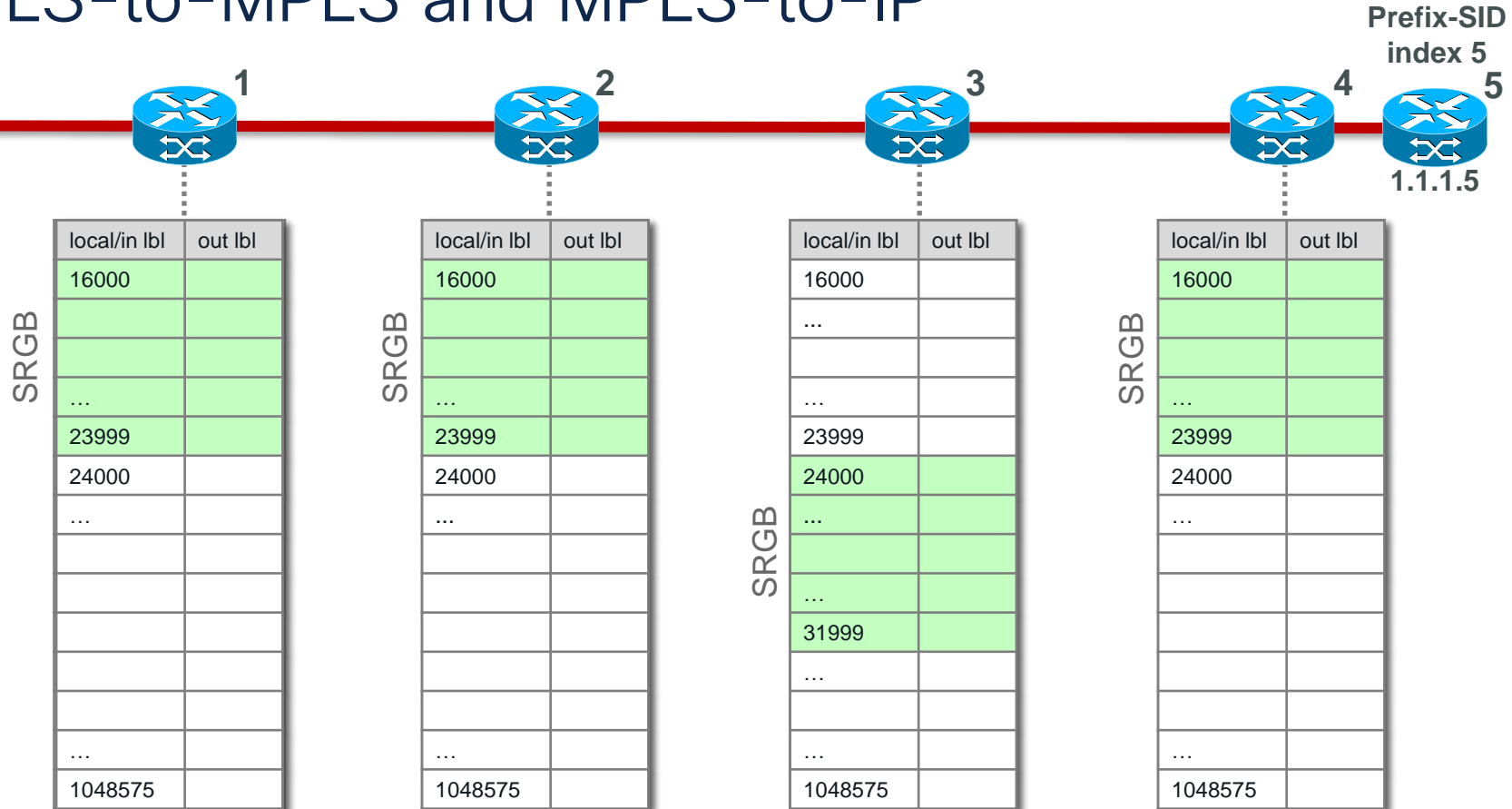
# Seamless Deployment

SR and LDP Co-existence Deployment Model

# Co-existence Deployment Model

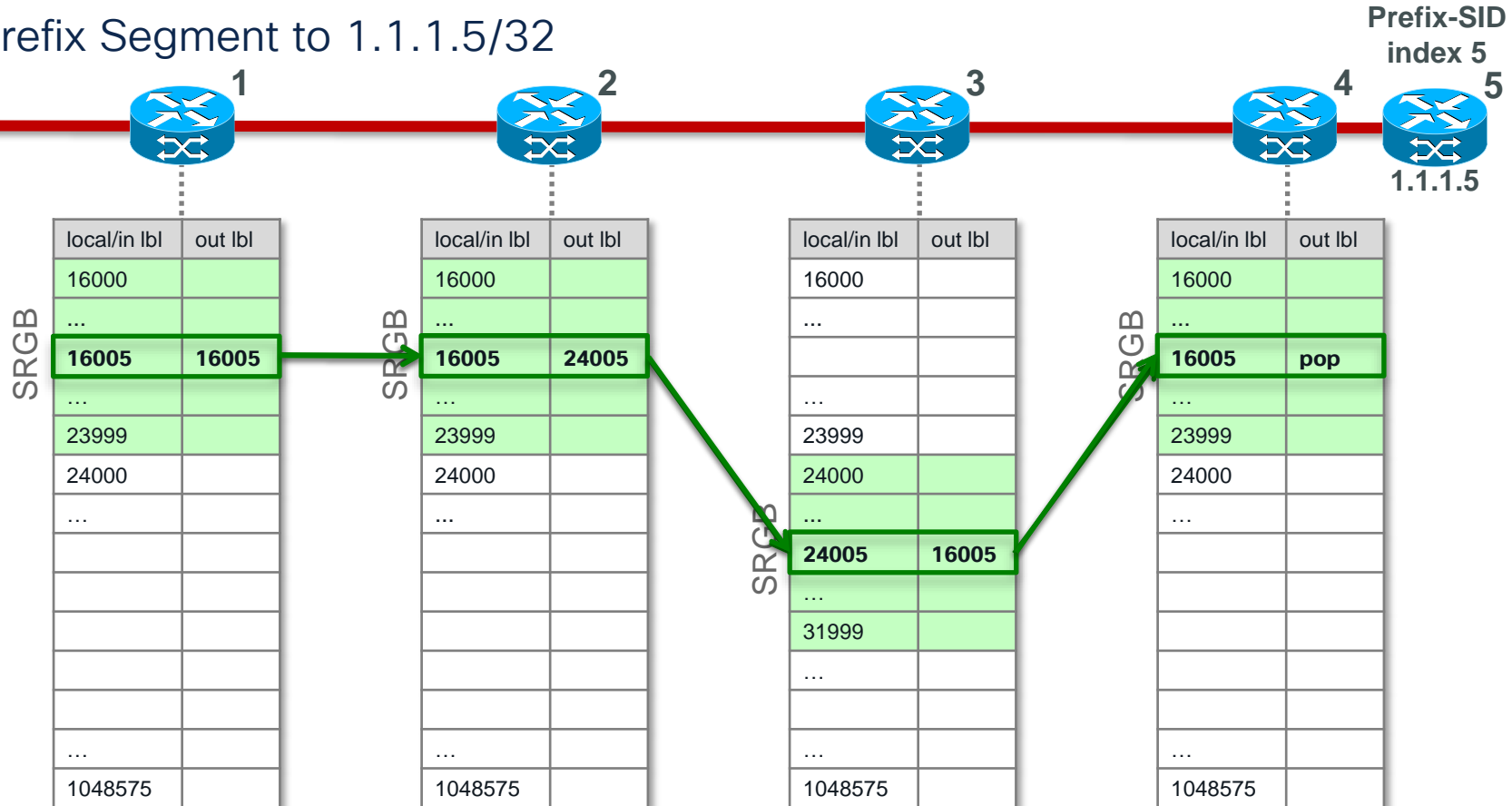
- SR and LDP control-plane and data-plane co-existence
- LDP and SR are kept independent
  - “Ships-in-the-night” model
- Deployment Model assumes that all nodes can be upgraded to SR
  - continuous SR connectivity between SR PEs
  - continuous LDP connectivity between LDP PEs

# MPLS-to-MPLS and MPLS-to-IP



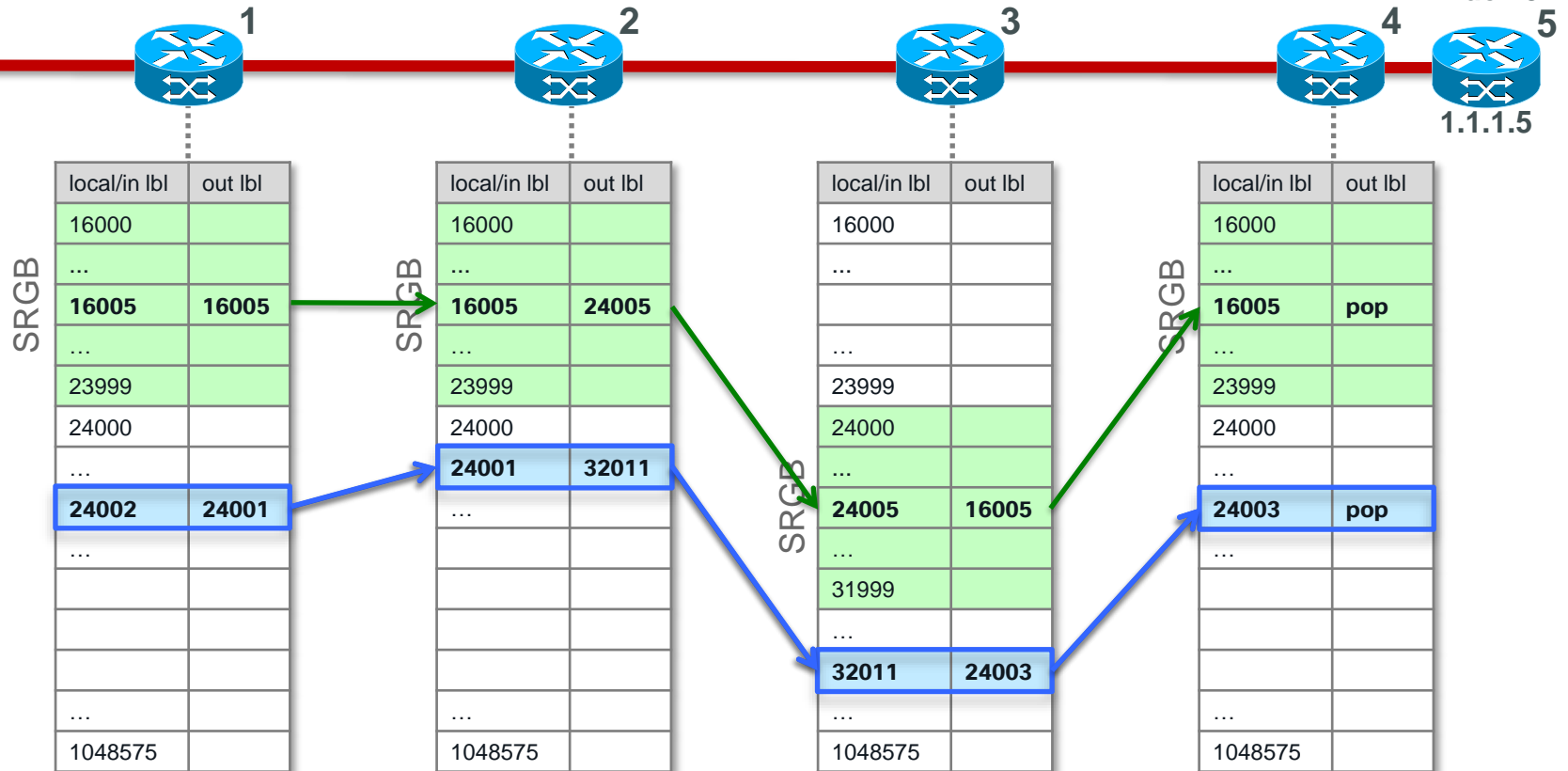
# MPLS-to-MPLS and MPLS-to-IP

## SR Prefix Segment to 1.1.1.5/32



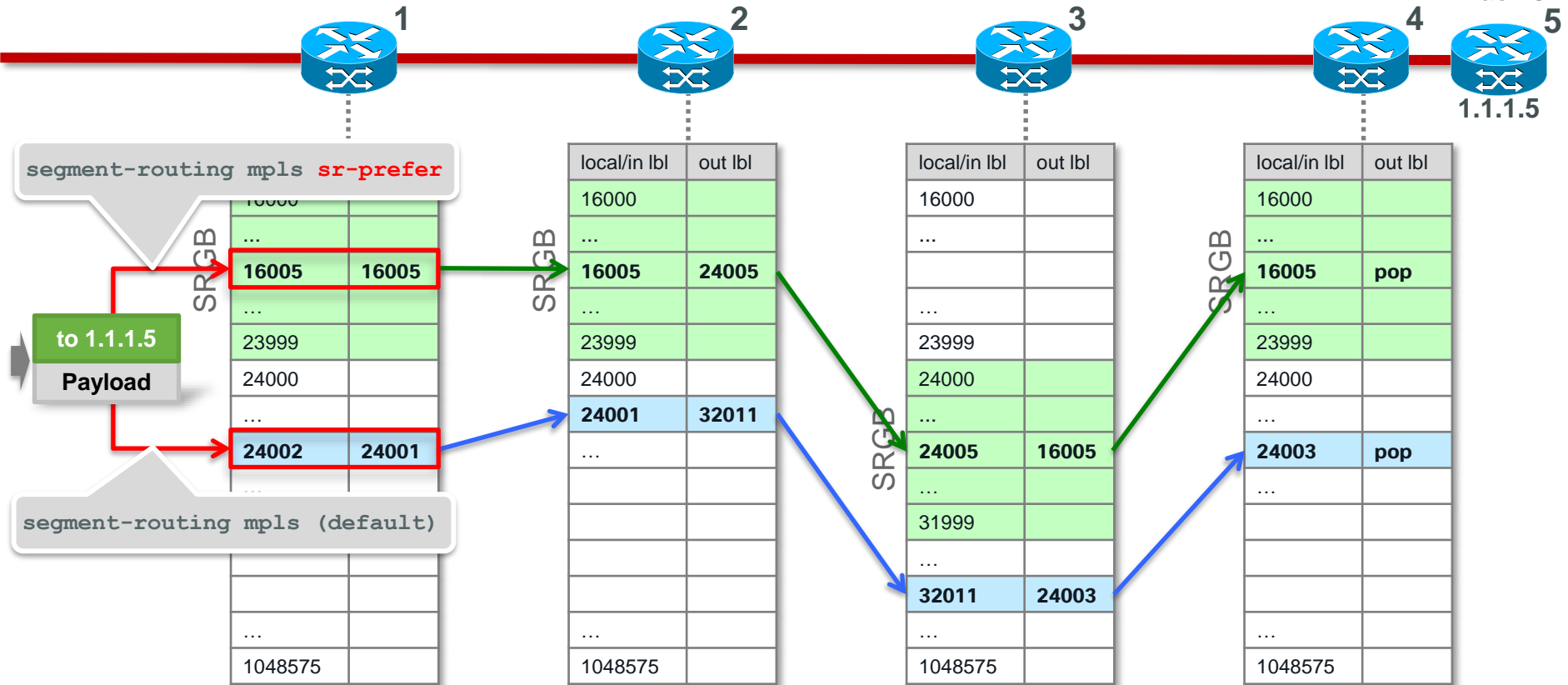
# MPLS-to-MPLS and MPLS-to-IP

LDP FEC to 1.1.1.5/32



# IP-to-MPLS: which label must be imposed?

Prefix-SID  
index 5  
1.1.1.5

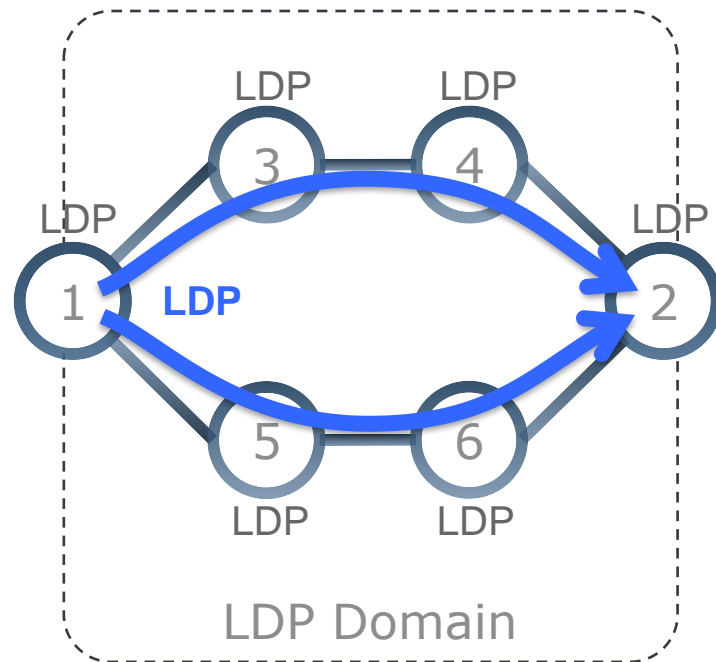


# Simplest migration LDP to SR

- **Initial state:** All nodes run LDP, not SR

Assumptions:

- all the nodes can be upgraded to SR
- all the services are supported over SR

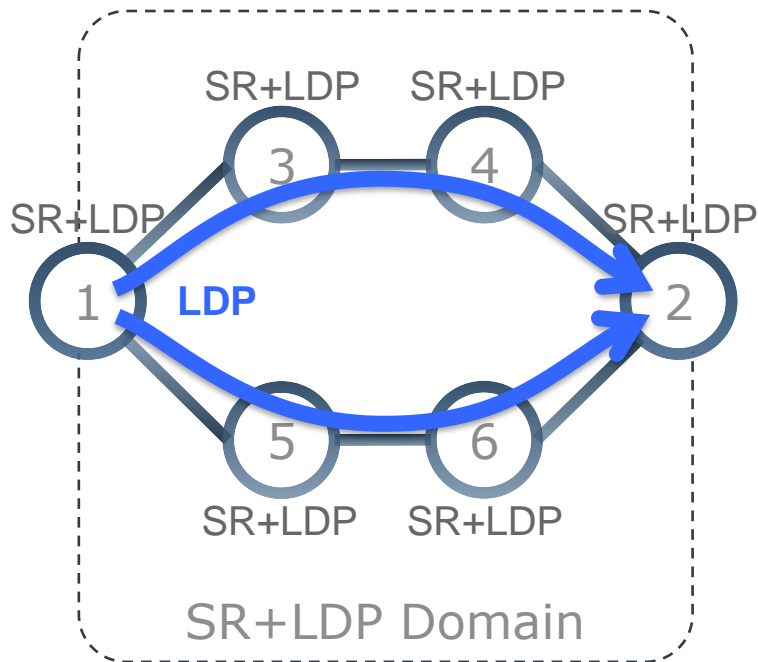


# Simplest migration LDP to SR

- **Initial state:** All nodes run LDP, not SR
- **Step 1:** All nodes are upgraded to SR
  - In no particular order
  - leave default LDP label imposition preference

Assumptions:

- all the nodes can be upgraded to SR
- all the services are supported over SR

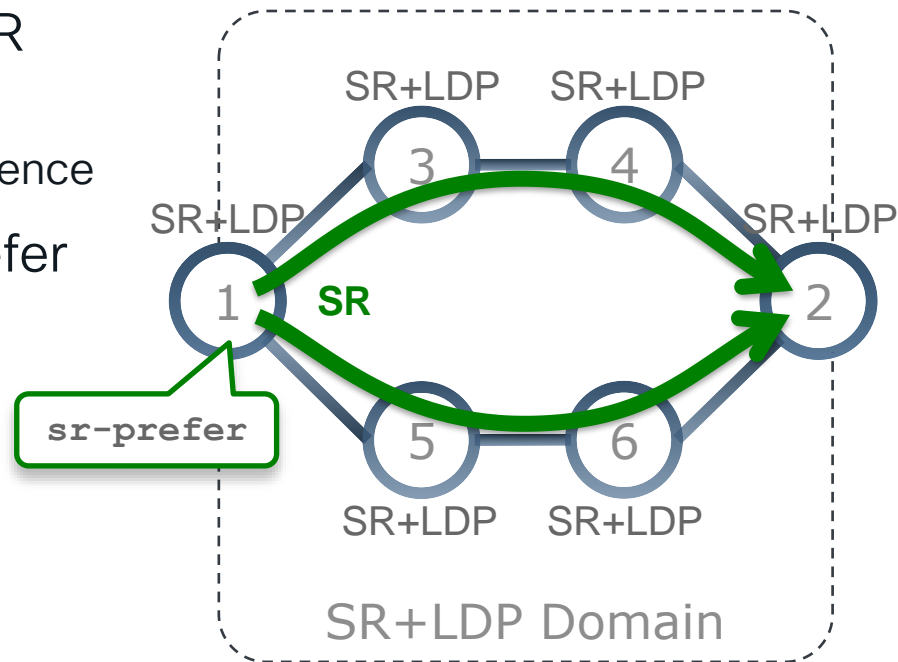


# Simplest migration LDP to SR

- **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

Assumptions:


- all the nodes can be upgraded to SR
- all the services are supported over SR




# IP-to-MPLS – label imposition

- For **IP2MPLS** forwarding, LDP XOR SR entry can be inserted into FIB
  - Only one IP2MPLS entry can exist for each prefix path
- Default: LDP label imposition is preferred

```
router isis 1
  address-family ipv4|ipv6 unicast
  segment-routing mpls sr-prefer
```

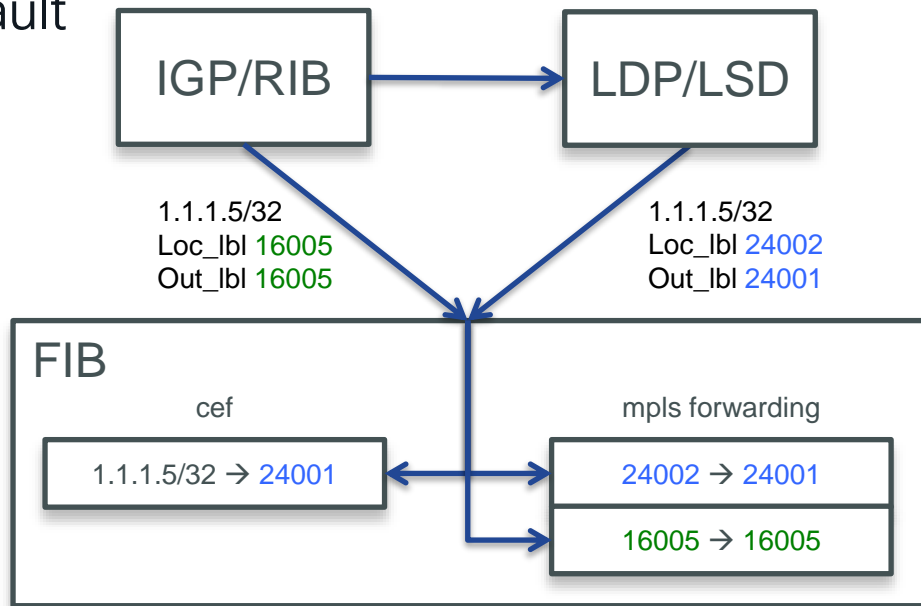
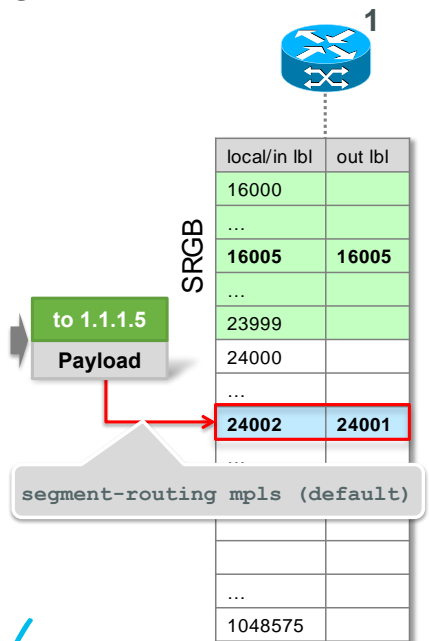


```
router ospf 1
  segment-routing mpls
  segment-routing sr-prefer
```



# IGP/SR and LDP programming FIB

- This diagram illustrates the default behavior: prefer LDP label imposition

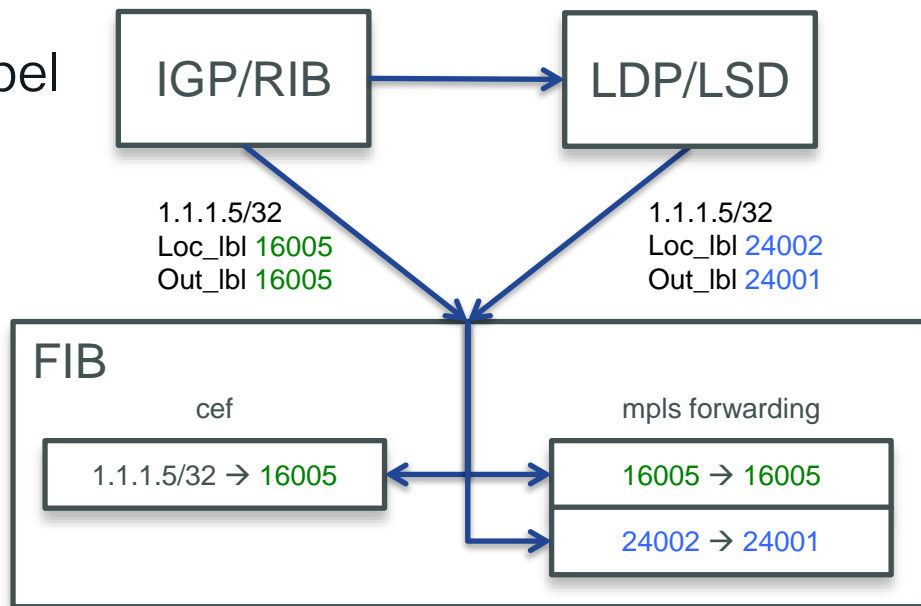
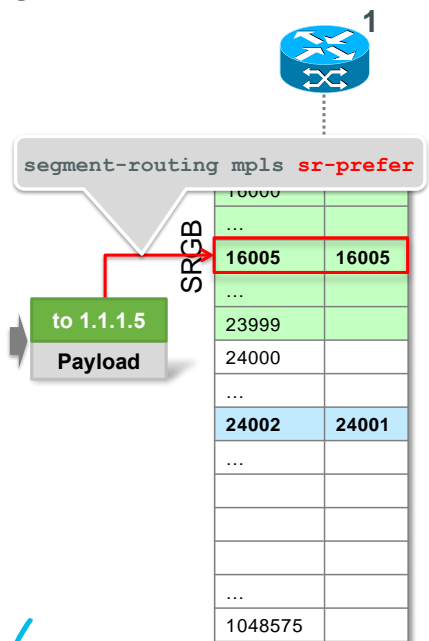


Loc\_lbl: local label, allocated by local node  
Out\_lbl: outgoing label

**16005**: SR label  
**24002**: LDP label

# IGP/SR and LDP programming FIB

- This diagram illustrates the behavior when preferring SR label imposition



Loc\_lbl: local label, allocated by local node  
Out\_lbl: outgoing label

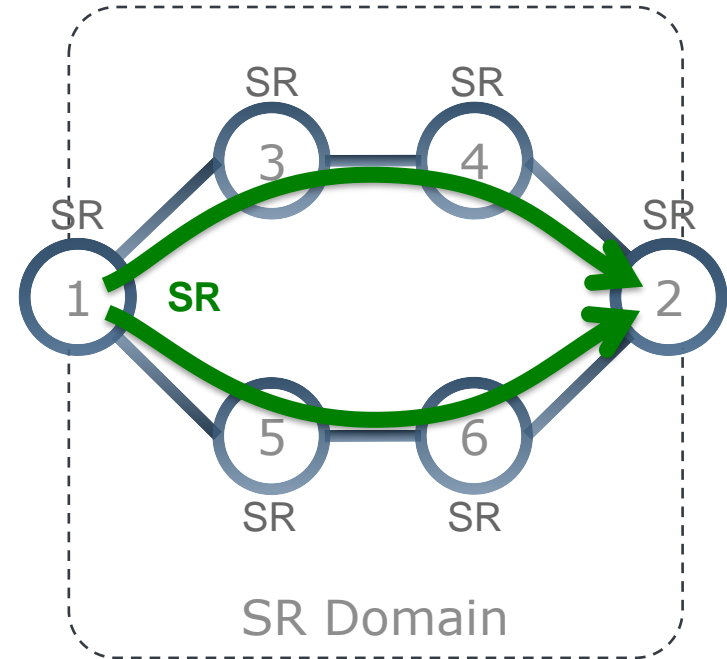
16005: SR label  
24002: LDP label

# Simplest migration LDP 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

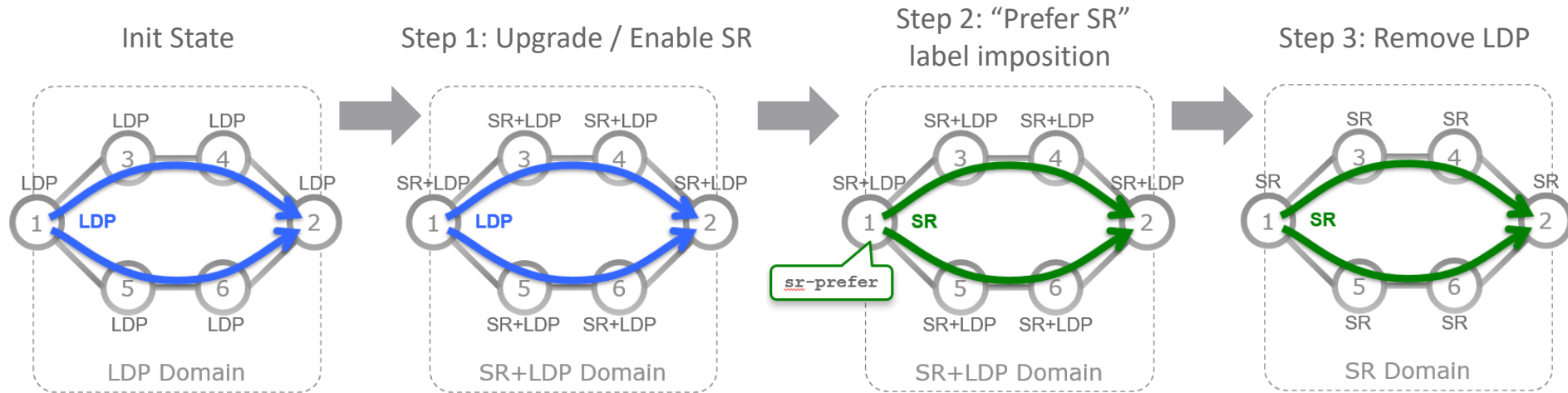
## Assumptions:

- all the nodes can be upgraded to SR
- all the services are supported over SR



# Co-existence Deployment Model

- Smooth migration steps where NE upgrades / configuration changes happen at the operator's pace

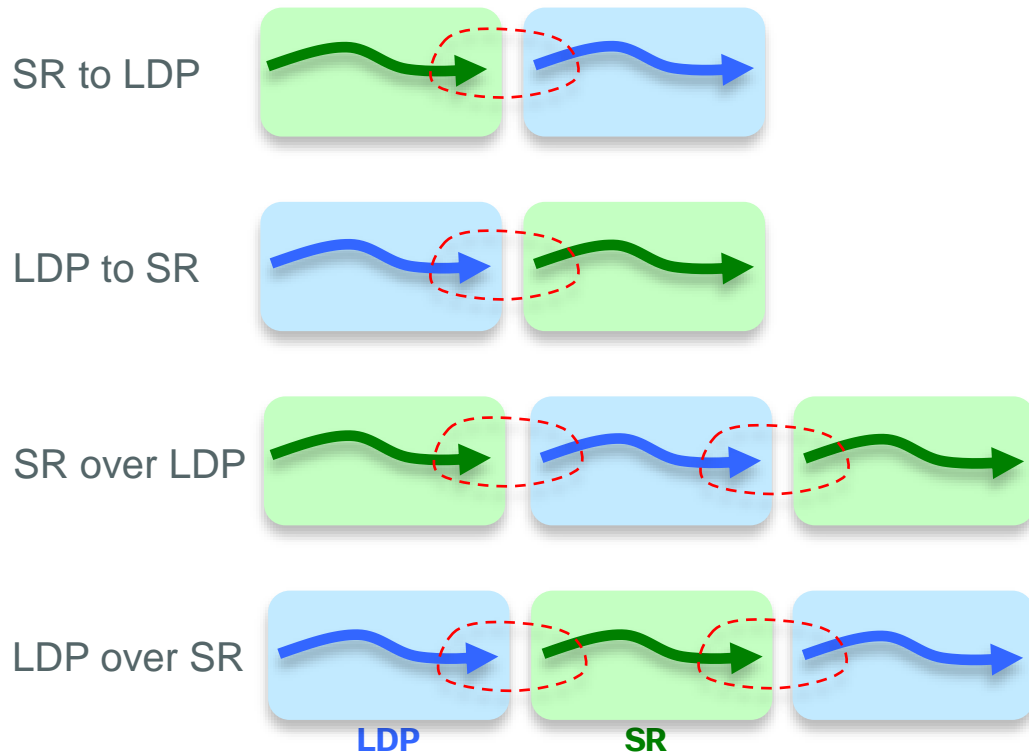


# Seamless Deployment

SR and LDP Interworking Deployment Model

# Interworking Deployment Model

- Simple, automatic and seamless interworking with LDP
- Based on:
  - SR Mapping Server (SRMS) control-plane function
  - SR / LDP data-plane interworking function



# SR Mapping Server

- Advertise **Prefix-to-SID mappings** in IGP on behalf of other **non-SR-capable** nodes
  - prefix-to-sid mappings are configured on the Mapping Server
- Enable SR-capable nodes to **interwork** with (non-SR-capable) LDP nodes, a Mapping Server is required for SR/LDP interworking

# Configuration Example – Mapping Server

```
segment-routing
 mapping-server
  prefix-sid-map
    address-family ipv4
      ! <prefix>/<len> <1st-SID> range <range>
      10.1.1.1/32 10 range 200
    !
  router isis 1
    address-family ipv4 unicast
    segment-routing prefix-sid-map advertise-local
```



10.1.1.1/32 – prefix-SID idx 10  
10.1.1.2/32 – prefix-SID idx 11  
...  
10.1.1.200/32 – prefix-SID idx 209

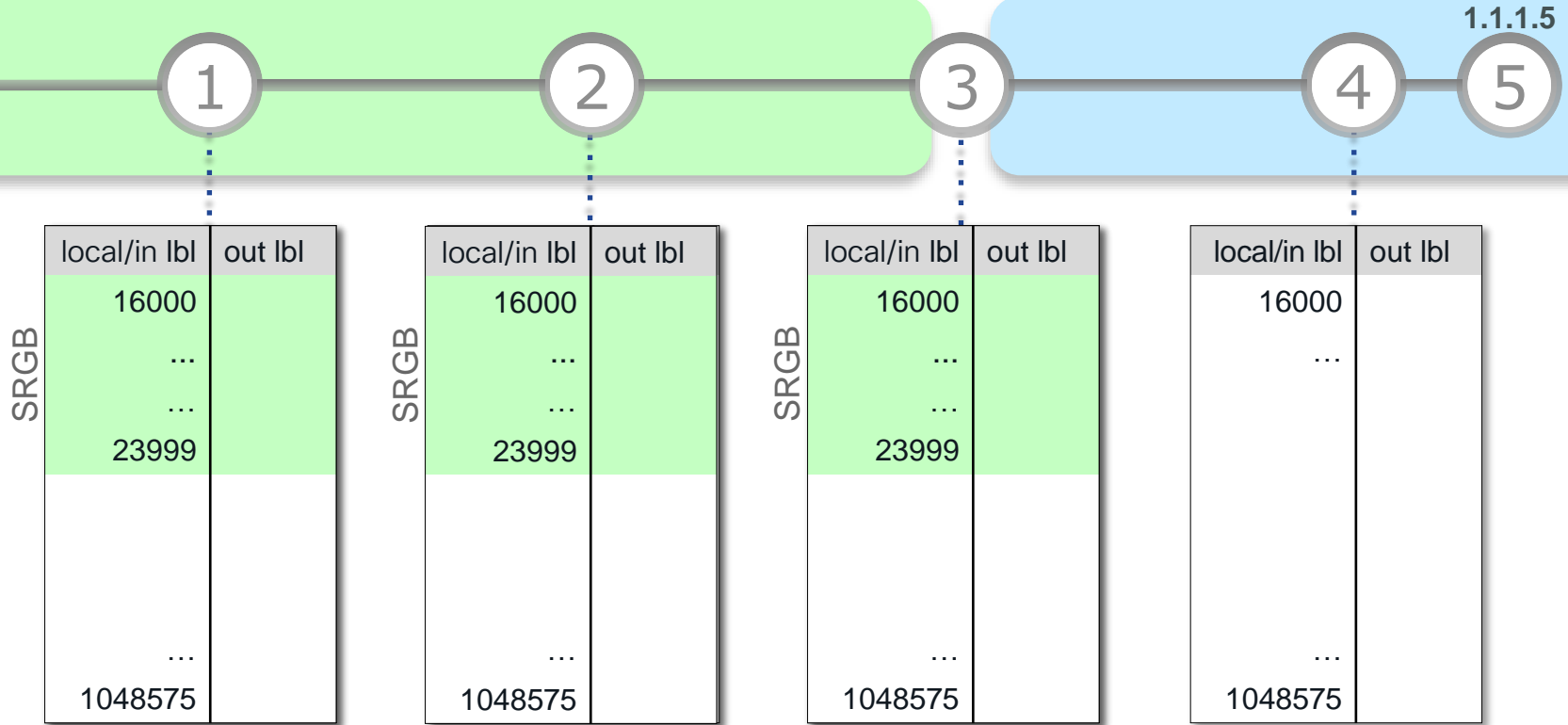


# SR to LDP traffic direction

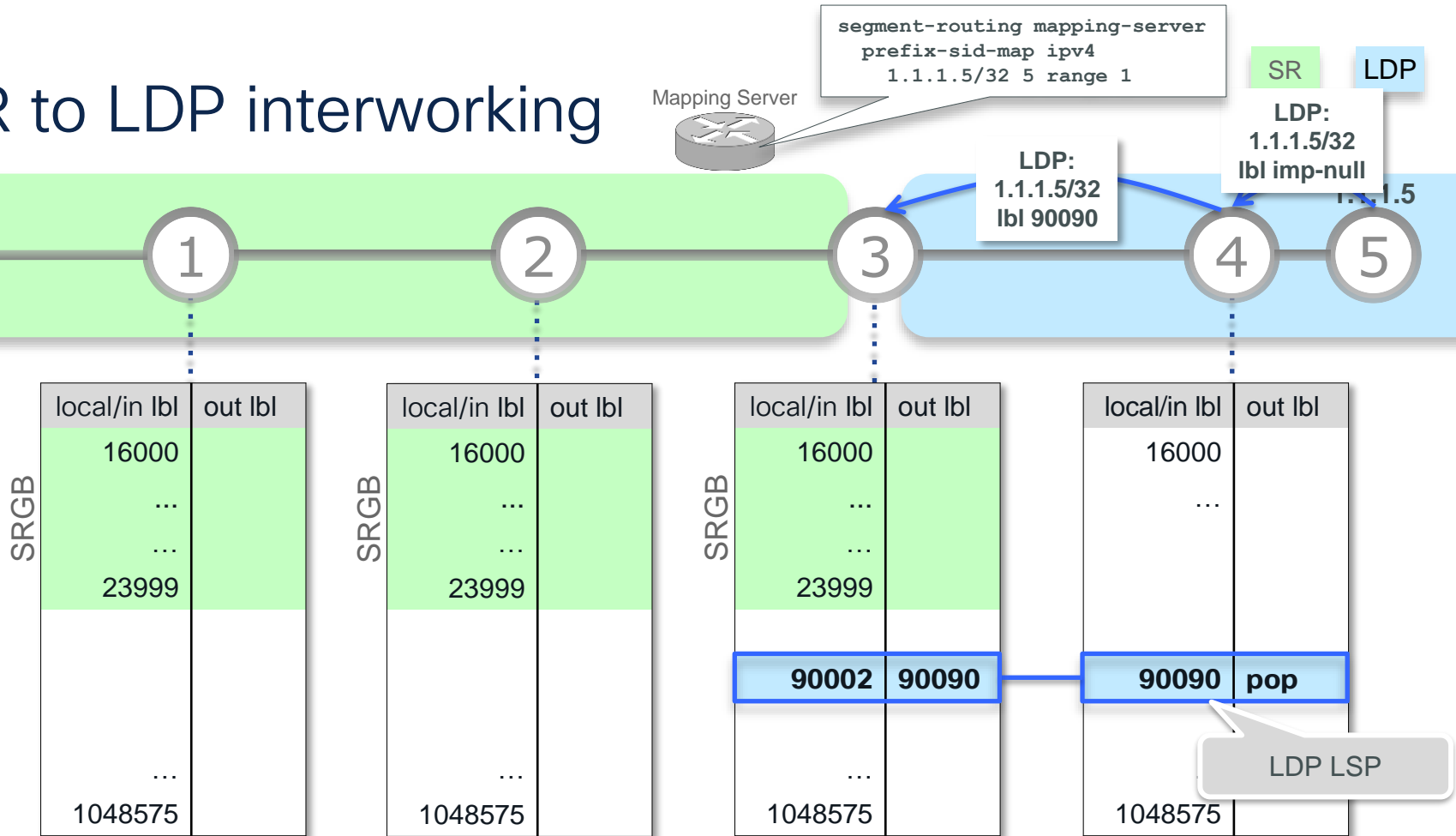
# SR to LDP interworking

SR

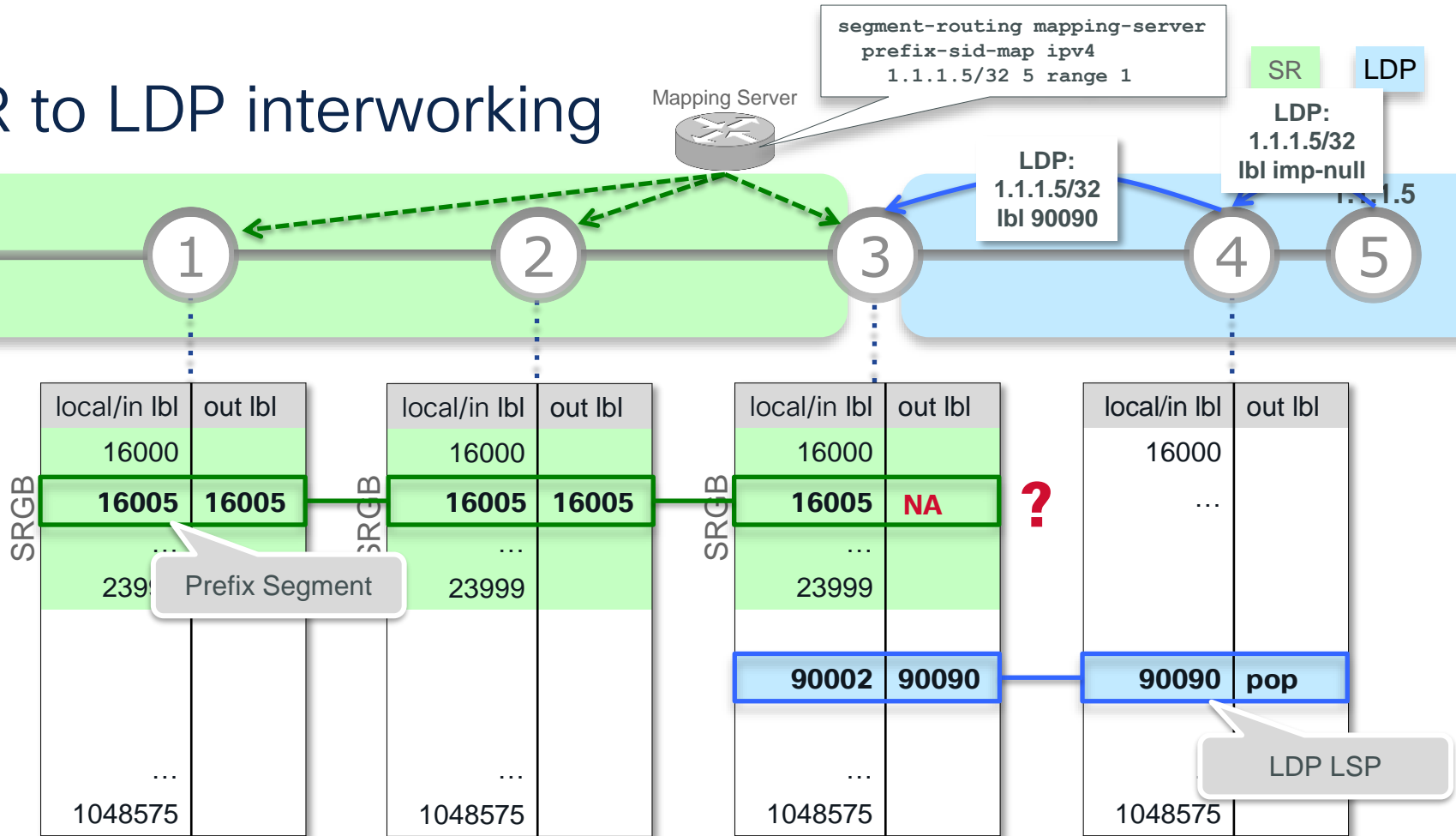
LDP



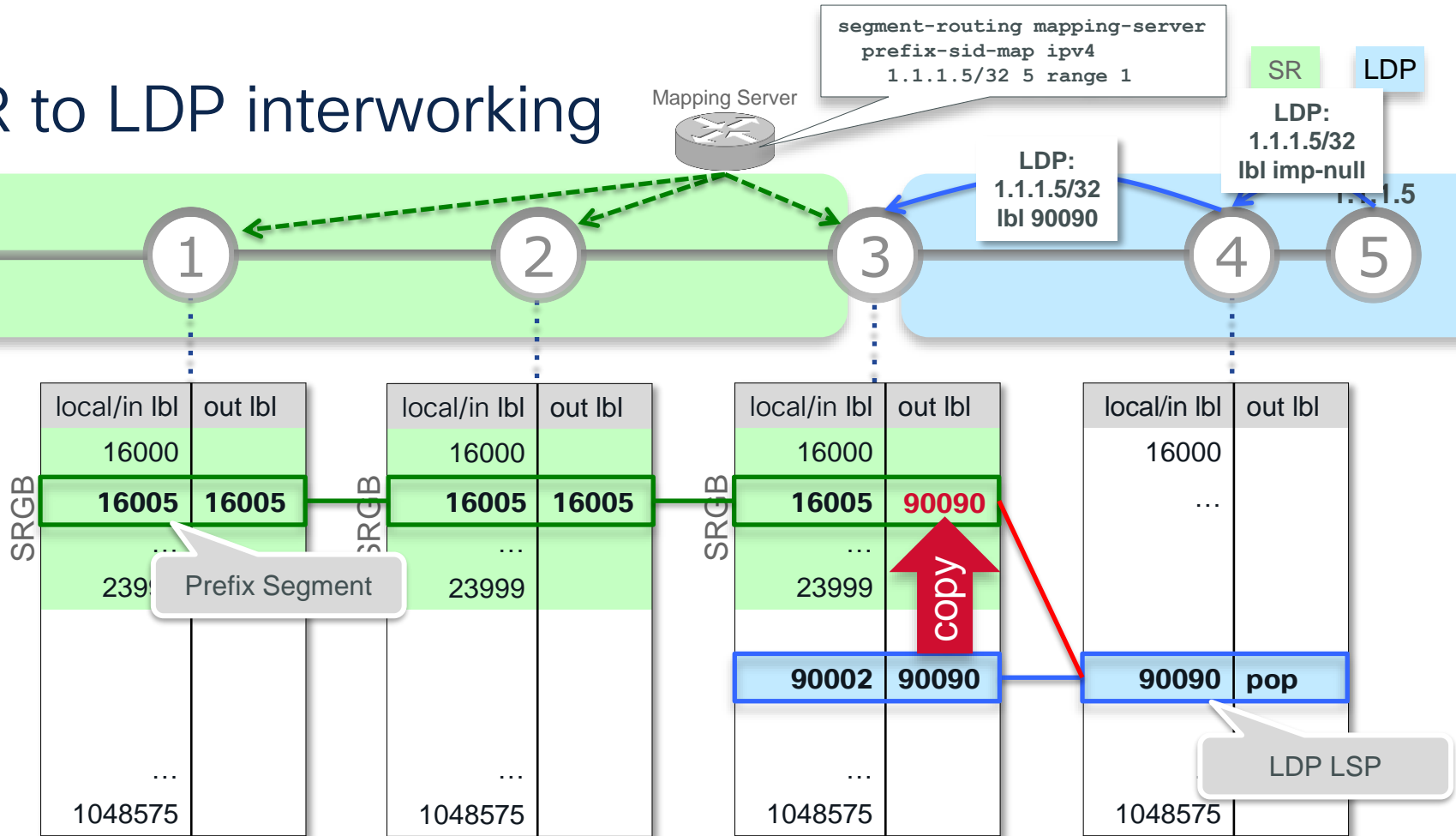
# SR to LDP interworking



# SR to LDP interworking



# SR to LDP interworking



# LDP to SR traffic direction

# LDP to SR interworking

SR

LDP



local/in lbl	out lbl
16000	
...	
...	
1048575	

local/in lbl	out lbl
16000	
...	
...	
1048575	

local/in lbl	out lbl
16000	
...	
...	
23999	
...	
1048575	

SRGB

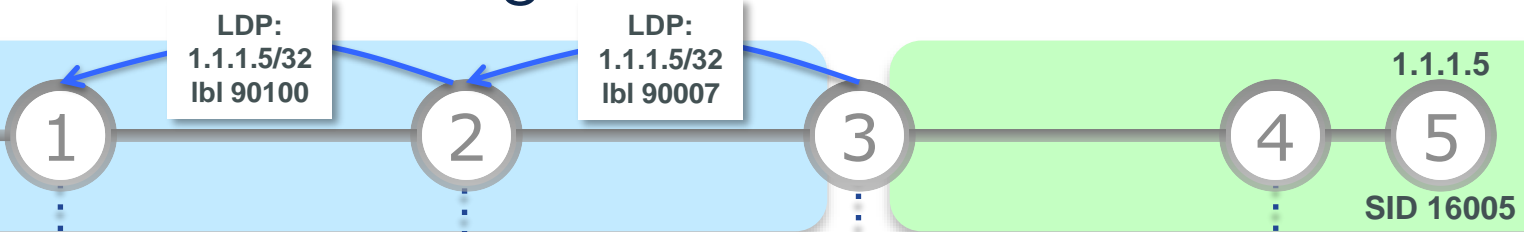
local/in lbl	out lbl
16000	
...	
...	
23999	
...	
1048575	

SRGB

# LDP to SR interworking

SR

LDP



local/in lbl	out lbl
16000	
...	
<b>90008</b>	<b>90100</b>
1048575	

LDP LSP

local/in lbl	out lbl
16000	
...	
<b>90100</b>	<b>90007</b>
1048575	

local/in lbl	out lbl
16000	
<b>16005</b>	<b>16005</b>
...	
23999	
...	
<b>90007</b>	<b>NA</b>
1048575	

SRGB

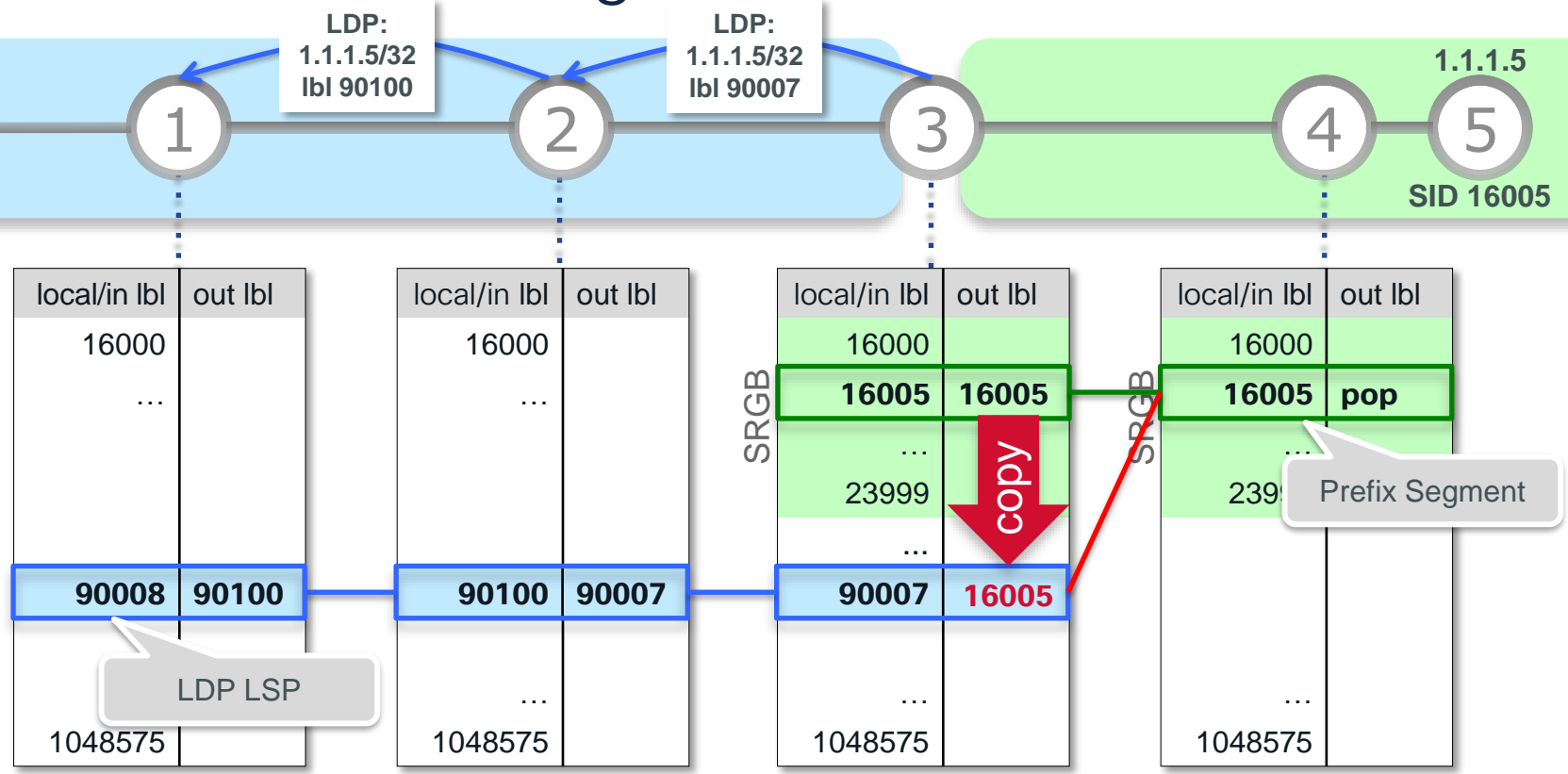
?

local/in lbl	out lbl
16000	
<b>16005</b>	<b>pop</b>
...	
23999	
...	
1048575	

SRGB

Prefix Segment

# LDP to SR interworking



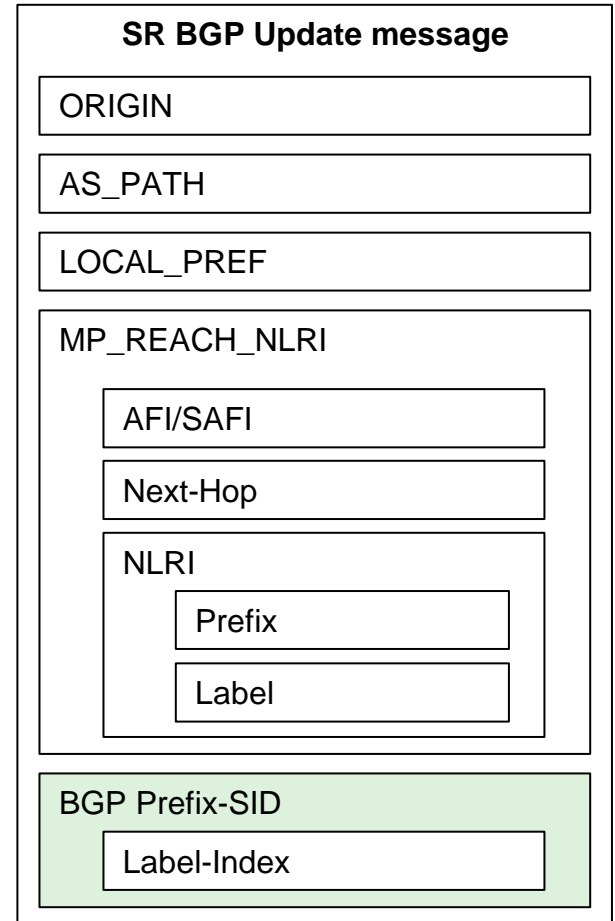
# BGP-SR / BGP Prefix-SID

# BGP Prefix-SID

- A BGP Prefix-SID is advertised with a prefix in BGP Labeled Unicast (BGP-LU)
  - BGP-LU = IPv4/IPv6 Labeled Unicast Address-families
- BGP Prefix-SIDs are global SIDs
- The instruction of the BGP Prefix-SID is to forward the packet over the ECMP-aware BGP best-path to the associated prefix

# BGP Prefix-SID advertisement

- Since the BGP Prefix-SID is a global SID, it is advertised as an index into the SRGB
  - BGP Prefix-SID label value =  $\text{SRGB}_{\text{base}} + \text{SID index}$
- SR BGP uses the BGP-LU address-family
- The Prefix-SID index is advertised in a **Label-Index TLV** of the **BGP Prefix-SID attribute** added to the BGP-LU Update message
  - BGP Prefix-SID attribute and Label-Index TLV are specified in draft-ietf-idr-bgp-prefix-sid



# SR BGP advertisement

- An SR BGP advertisement contains both the Prefix-SID label value and the Prefix-SID index
  - Allows interworking between classic BGP-LU and SR BGP
  - Allows using different SRGBs on the nodes (not recommended!)
- The **index is a hint to allocate the local label** for the route from the SRGB at the specified offset
- The **label value is the outgoing label** for the route

# SR BGP configuration

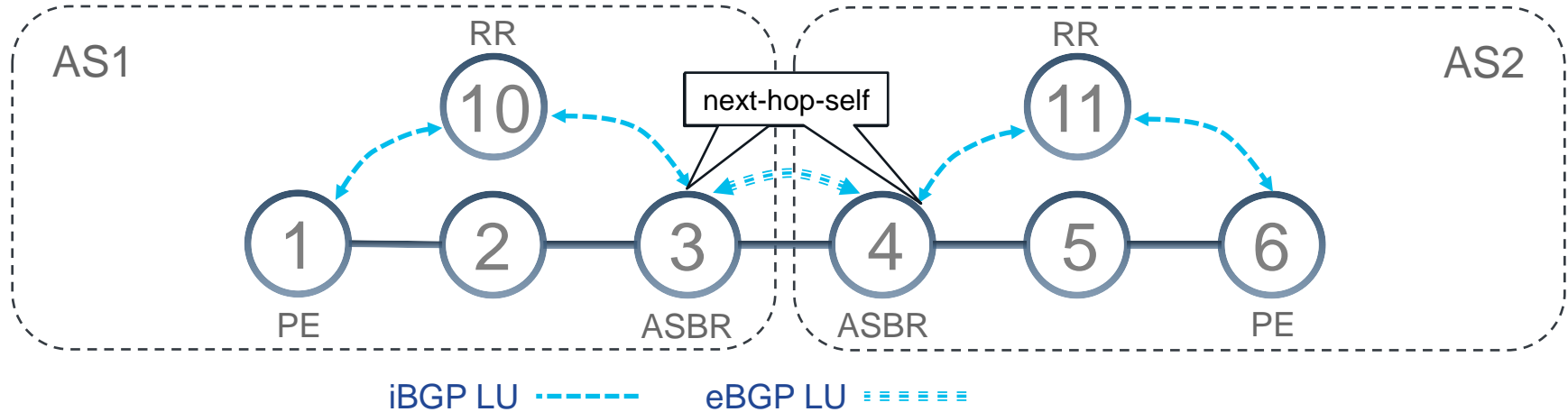
- SR BGP is automatically enabled when configuring a **global SRGB**

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



- BGP uses this globally configured SRGB
- Note 1: there is no default global SRGB
- Note 2: if a global SRGB is configured, the IGP's use it by default

# Reference Topology



- Distribution of PE routes:
  - BGP-LU between PE1 and ASBR3 and between PE6 and ASBR4 via their RRs Node10 and Node11


- Next-hop-self enabled on Node3's session to Node10 and Node4's session to Node11

# BGP Prefix-SID config – set label-index

- The Prefix-SID of a locally originated BGP route is set via a route-policy
- A route-policy with **set label-index <idx>** can be attached to:
  - a) **network** configuration
  - b) **redistribute** configuration


a)

```
route-policy SID($SID)
  set label-index $SID
end-policy
!
router bgp 1
  address-family ipv4 unicast
    network 1.1.1.1/32 route-policy SID(1)
    allocate-label all
```



b)

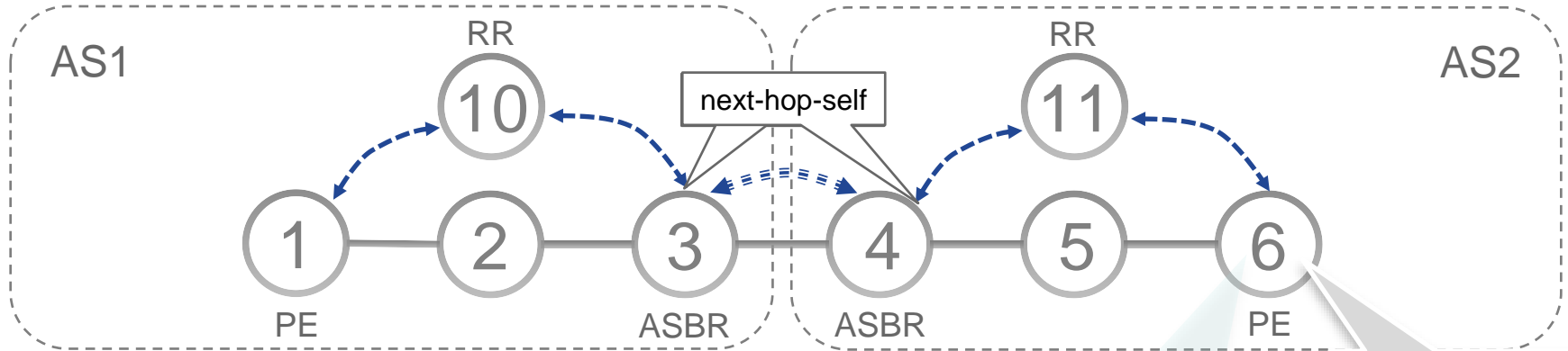
```
route-policy SIDs
  if destination in (1.1.1.1/32) then
    set label-index 1
  endif
end-policy
!
router bgp 1
  address-family ipv4 unicast
    redistribute connected route-policy SIDs
    allocate-label all
```



# Originate local route on Node6

iBGP LU -----

eBGP LU =====



This is an example, see next slide

```
route-policy SID($SID)
  set label-index $SID
end-policy
!
segment-routing
  global-block 16000 23999
!
router bgp 2
  bgp router-id 1.1.1.6
  address-family ipv4 unicast
  network 1.1.1.6/32 route-policy SID(6)
  allocate-label all
```

IP: 1.1.1.6/32  
Prefix-SID: 16006  
(SID index: 6)


Segment Routing  
Global Block (SRGB)

Associate prefix-SID  
index 6 with 1.1.1.6/32

# Use argument for Route-policy

- The route-policy SID(\$SID) on the previous slide uses an argument \$SID that is provided when configuring the route-policy
  - This allows to share a single route-policy for different attachment points
- \$SID in the route-policy is replaced by 6 for the network 1.1.1.6/32

```
route-policy SID($SID)
  set label-index $SID
end-policy
!
router bgp 2
  address-family ipv4 unicast
    network 1.1.1.6/32 route-policy SID(6)
```



# Prefix-SID Redistribution

# Prefix-SID redistribution

- Prefix-SID redistribution between SR domains
- Prefix-SID redistribution between SR and non-SR domain

# Redistribution reminder

- Redistribution uses the RIB as middle person:
- Routes installed by the **redistributed** (**source**) instance in RIB are redistributed into the **redistributing** (**destination**) instance
- The **redistributed** instance installs the RIB/FIB entry
- The **redistributing** instance does not install a forwarding entry

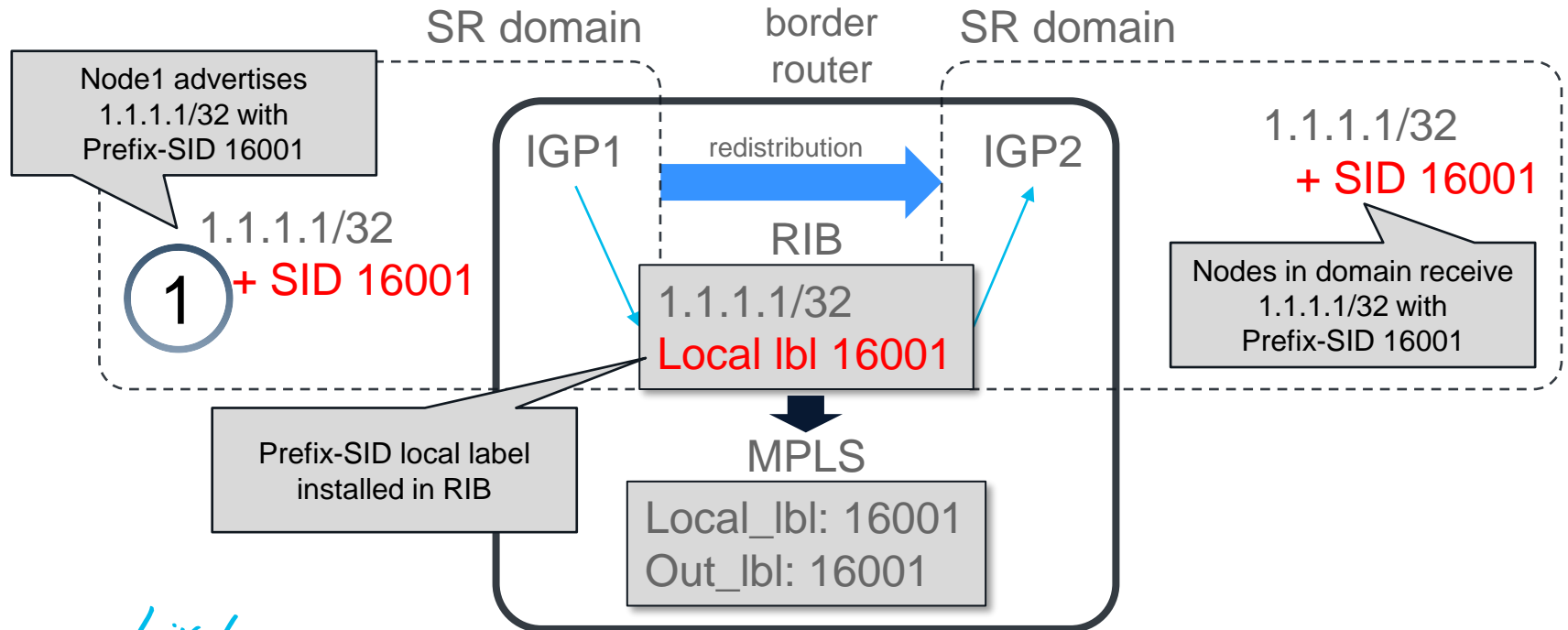
# Prefix-SID redistribution between SR domains

# Prefix-SID redistribution between SR domains

- Redistribution Scenarios:
  - IGP-SR  $\leftrightarrow$  IGP-SR
  - IGP-SR  $\leftrightarrow$  BGP-SR
- The Prefix-SID of a prefix is redistributed with the prefix
- Adj-SIDs are not redistributed as they belong an adjacency
- BGP Prefix-SID is used when redistributing to/from BGP(-LU)
- Both protocols/instances must have the **same SRGB** (SRGB is really a node property)

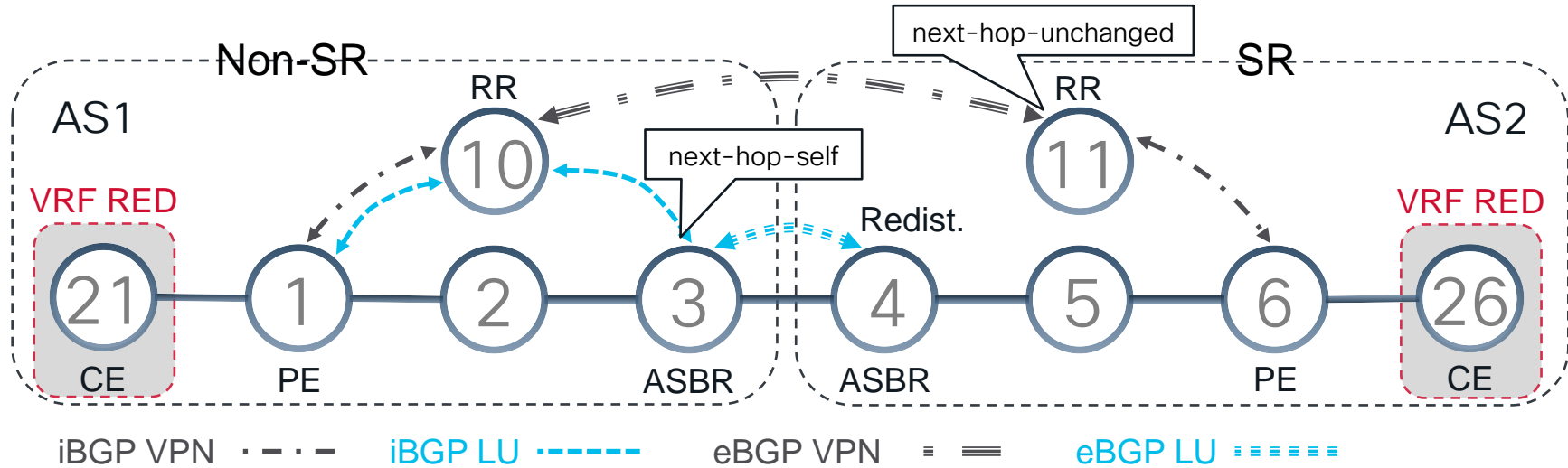
# Prefix-SID redistribution SR to SR

- A look inside the border router connecting two SR domains
  - Redistributing IGP1 into IGP2



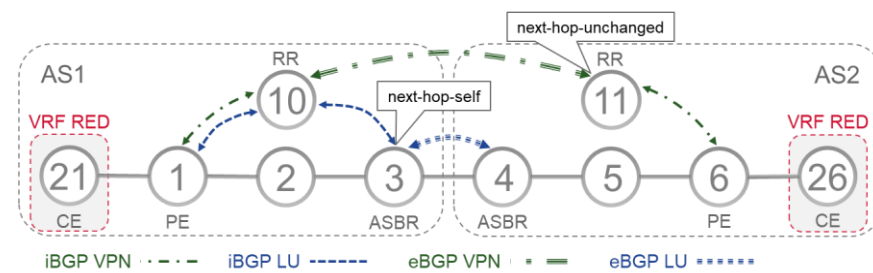
# Prefix-SID redistribution between SR and Non-SR domains

# SR VPN Inter-AS option C - Topology



- Distribution of PE routes:
  - In AS1: BGP-LU between ASBR3 and PE1
  - In AS2: mutually redistribute BGP in IGP
- Next-hop-self enabled on Node3's session to Node10
- Next-hop-unchanged enabled on Node11's session to Node10

# Topology description

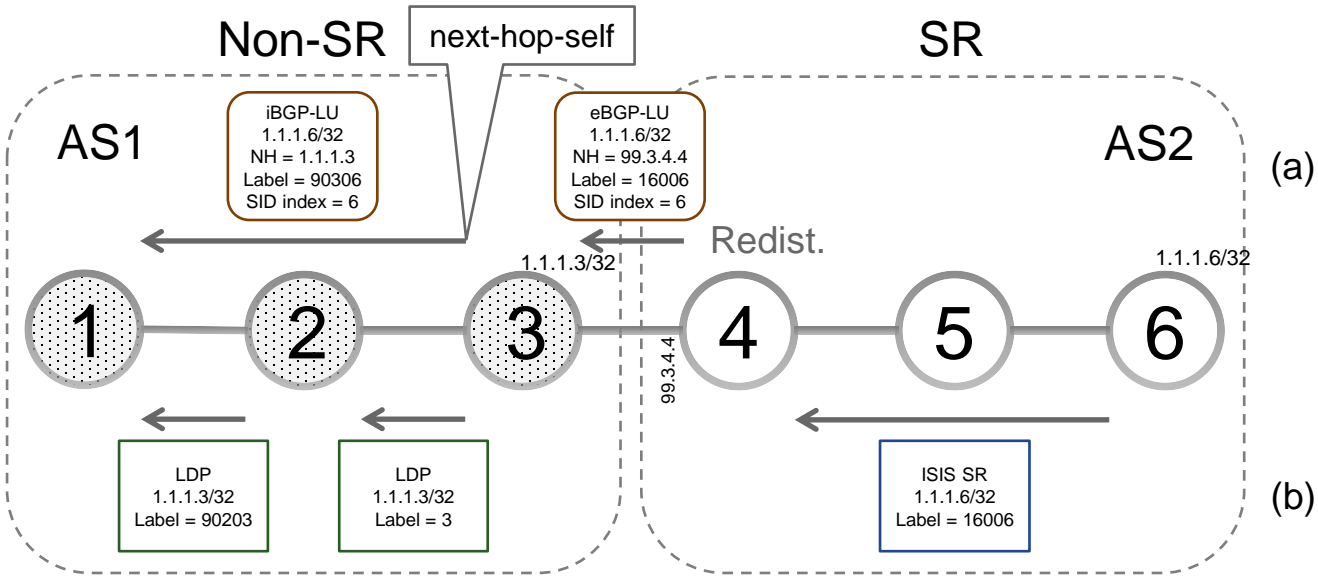


- Two IGP domains interconnected via BGP peering link between ASBR3 and ASBR4
- For illustration purposes, the remote PE reachability distribution uses a different method in both domains
- AS1 is non-SR-enabled (all nodes):
  - Uses LDP and BGP-LU
  - Remote PE reachability via BGP-LU between ASBR3 and PE1
- AS2 is SR enabled (all nodes):
  - Uses SR IGP and SR BGP
  - Remote PE reachability via mutually redistributing BGP ↔ IGP on ASBR4

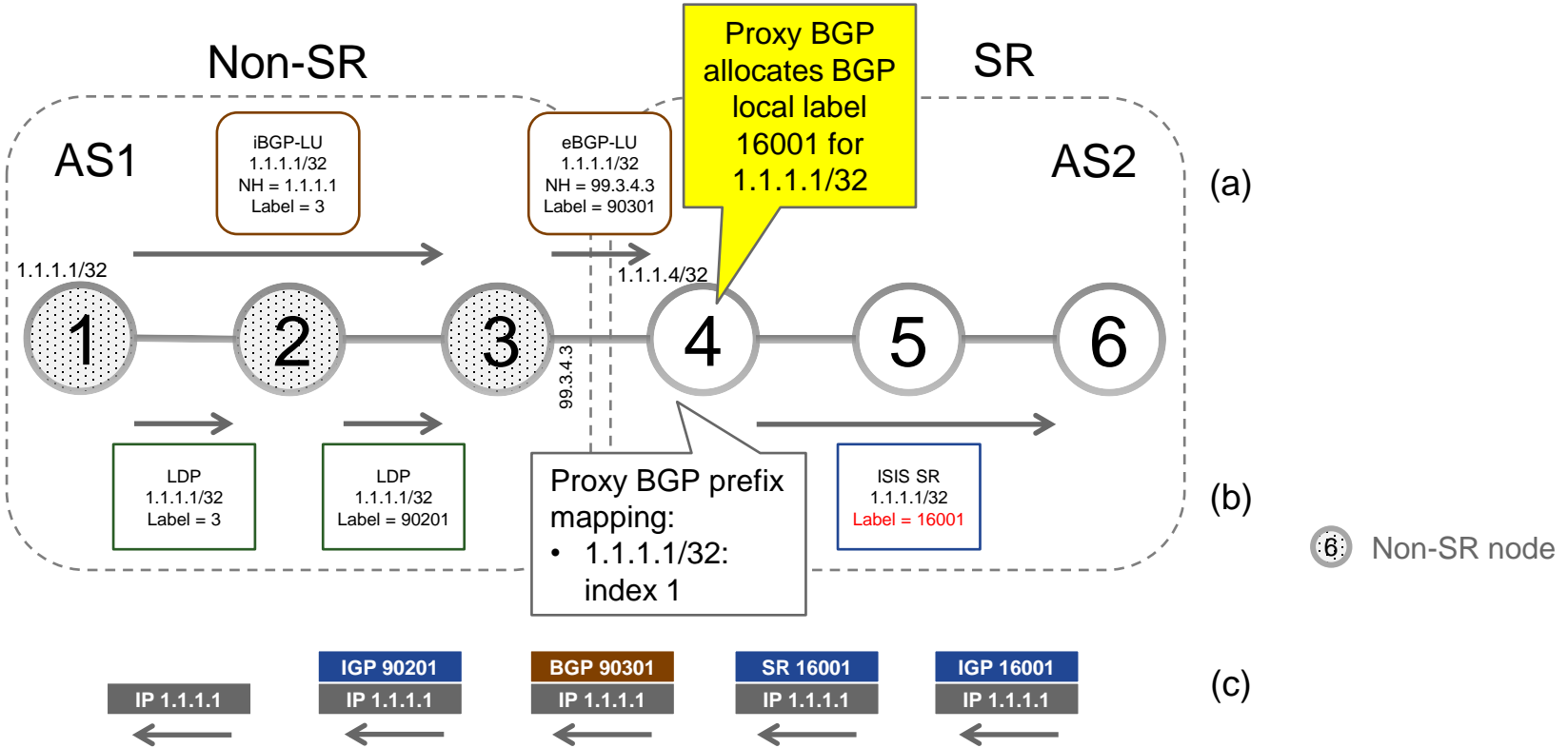
# Proxy BGP-SR

- Provide prefix-to-sid map to BGP for the prefixes that have no Prefix-SID
  - Use local prefix-sid-map and apply it under BGP
- BGP then allocates the Prefix-SID label as local label
  - BGP installs Prefix-SID forwarding entry
  - BGP provides Prefix-SID for redistribution

# Transport routes (1.1.1.6/32)



# Transport routes (1.1.1.1/32)



# Configuration Example – SR Mapping Server & BGP

```
segment-routing
  global-block 16000 23999
  mapping-server
    prefix-sid-map
      address-family ipv4
        ! <prefix>/<len> <1st-SID> range <range>
        1.1.1.1/32 1 range 200
        10.1.1.1/32 401 range 300
      !
```



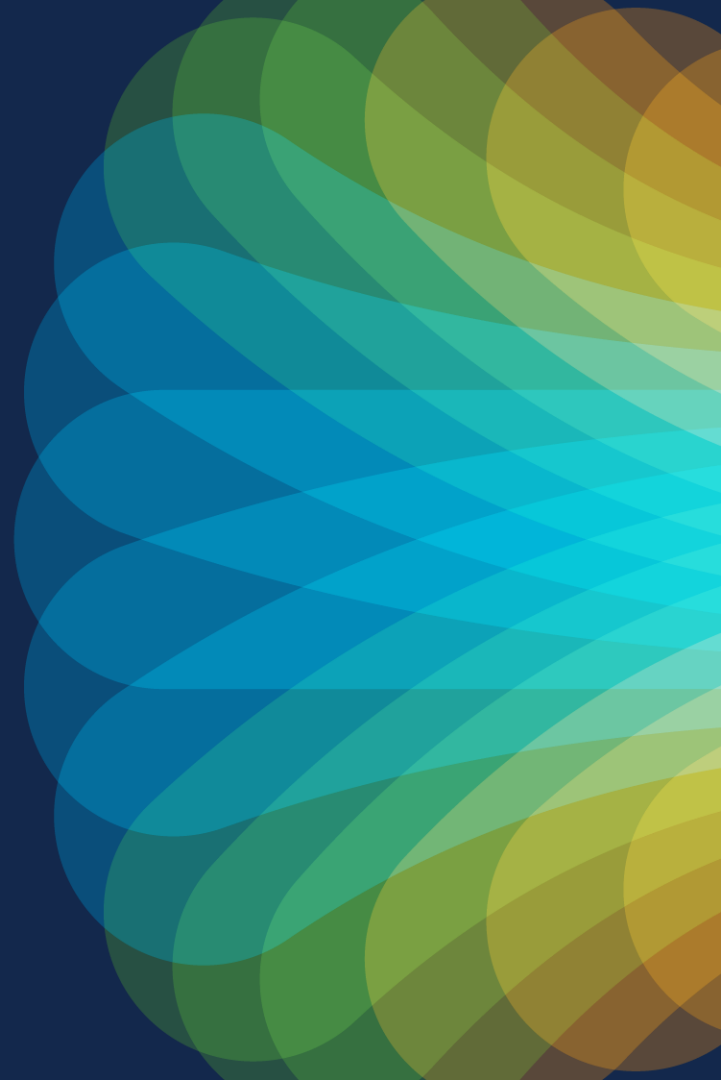
1.1.1.1/32 – prefix-SID idx 1  
1.1.1.2/32 – prefix-SID idx 2  
...  
1.1.1.200/32 – prefix-SID idx 200

20.1.1.1/32 – prefix-SID idx 401  
20.1.1.2/32 – prefix-SID idx 402  
...  
20.1.2.44/32 – prefix-SID idx 700

```
router bgp 1
  address-family ipv4 unicast
    segment-routing prefix-sid-map
```



# Color-Aware Routing Principles



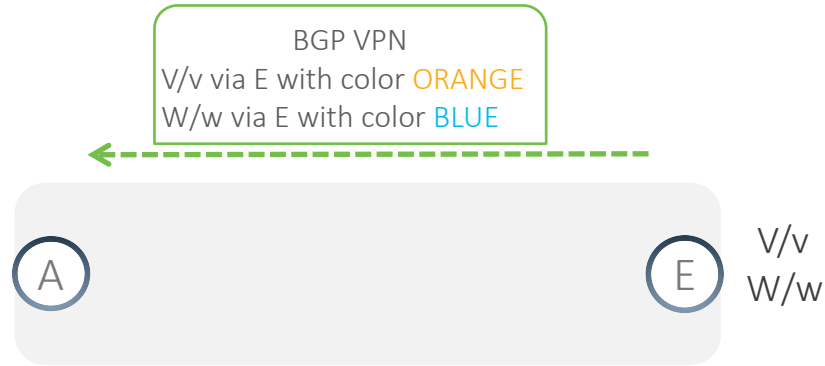
# Types of intent

- Topology path selection
  - Minimize expected delay
  - Minimize delay
  - Minimize cost per bit with a delay bound
  - Avoid resource
  - Disjoint paths
  - Disjoint planes
  - Data Sovereignty
- Others
  - Steer traffic along a service chain
- Any combination of the above

# Intent encoded as a color

- Color is a standard way to signal intent
  - A 32-bit number
- Mapping an intent to a color:
  - Low-latency: BLUE
  - Low-cost: ORANGE
- Colored Service Routes – requesting a particular intent
- Color-aware Transport Routes – satisfying a particular intent
- A colored service route is steered over a color-aware route of same intent

# Colored service route

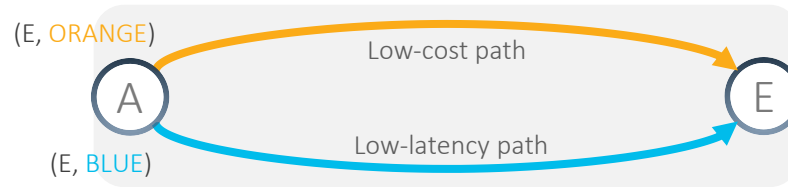


- E sends colored service routes, each **requesting** a particular intent
- Route coloring is done by using the BGP **Color** Extended-Community
  - Standard ([RFC5512](#) / [RFC9012](#)), supported by all major BGP implementations
- Any service route can be colored (L3VPN, EVPN, Internet routes)

# Color-aware transport routes

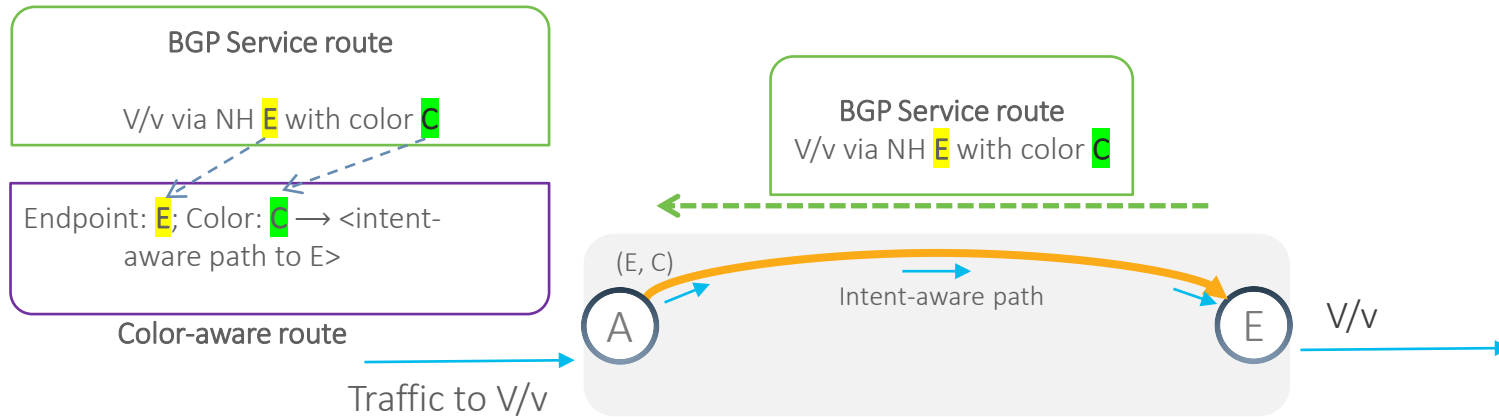
Color-aware routes @ A:

Endpoint	Color	Path
E	ORANGE	<low-cost path A to E>
E	BLUE	<low-latency path A to E>



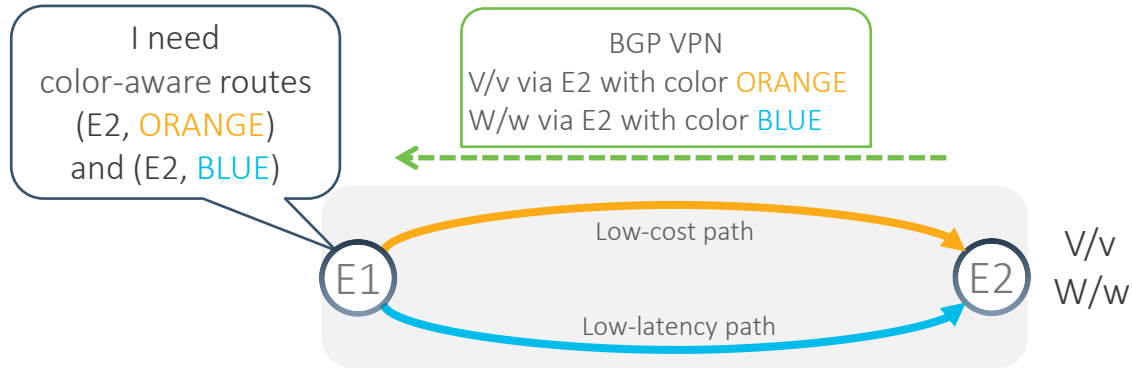
- A color-aware route satisfies a particular intent
- A color-aware route is identified by the tuple (Endpoint and Color); in short (E,C)
- A color-aware route can be signaled/instantiated by different mechanisms

# Service routes steered on color-aware route



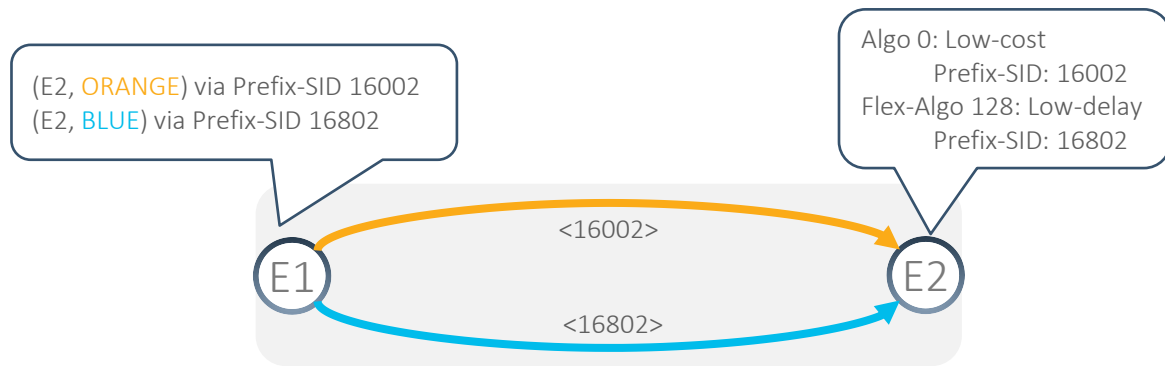
- A colored service route is steered over a color-aware route of same intent
- Traffic destined to prefix V/v via E with color C is steered over color-aware route (E, C)
- This is known as [Automated Steering](#)

# Service routes steered on color-aware route



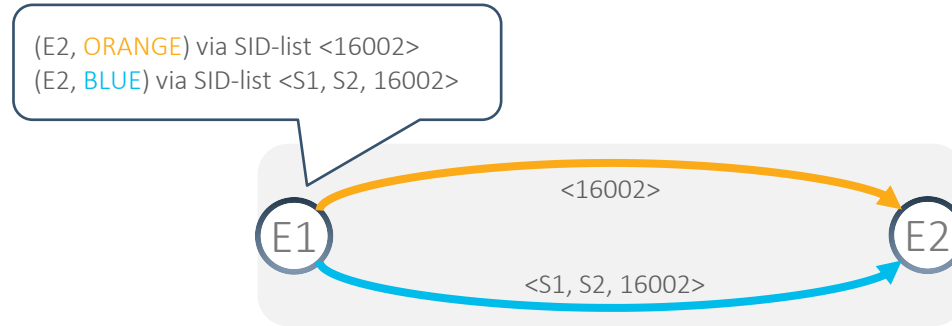
- A colored service route is steered over a color-aware route of same intent
  - V/v via E2 with color ORANGE is steered over (E2, ORANGE)
  - W/w via E2 with color BLUE is steered over (E2, BLUE)

# Color-aware route (E2, C) provided by IGP Flex Algo



- IGP Flex Algo ([RFC9350](#))
- E1 maps color to an IGP algorithm
  - Orange → Algo 0
  - Blue → Flex-Algorithm 128
- IOS-XR implementation available since 2019

# Color-aware route (E2, C) provided by SR Policy



- SR Traffic Engineering Policy – in short SR Policy ([RFC9256](#))
- E1 has two SR policies
  - (E2, ORANGE): Dynamic, low cost with SID-list <16002>
  - (E2, BLUE): Dynamic, low delay with SID-list <S1, S2, 16002>
- IOS-XR implementation available since 2017

# SR Policy

# SR Traffic Engineering

- Simple, Automated and Scalable
  - No core state: [state in the packet header](#)
  - No tunnel interface: “[SR Policy](#)”
  - uniquely identified by a tuple (head-end, color, end-point)
  - No head-end a-priori configuration: [on-demand](#) policy instantiation
  - No head-end a-priori steering: [automated](#) steering
- Multi-Domain
  - [SR PCE](#) for compute
  - [Binding-SID \(BSID\)](#) for scale
- Lots of Functionality
  - Designed with [lead operators](#) along their use-cases

# Key IETF document for SRTE

Internet Engineering Task Force (IETF)

Request for Comments: 9256

Updates: [8402](#)

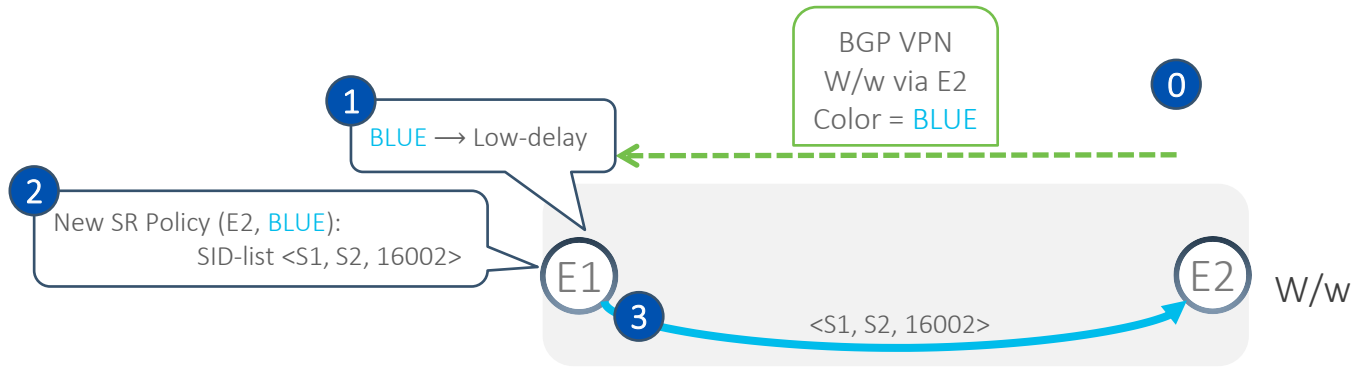
Category: Standards Track

ISSN: 2070-1721

C. Filsfils  
K. Talaulikar, Ed.  
Cisco Systems, Inc.  
D. Voyer  
Bell Canada  
A. Bogdanov  
British Telecom  
P. Mattes  
Microsoft  
July 2022

**Segment Routing Policy Architecture**

# SR Policy pull model: On-Demand Nexthop (ODN)



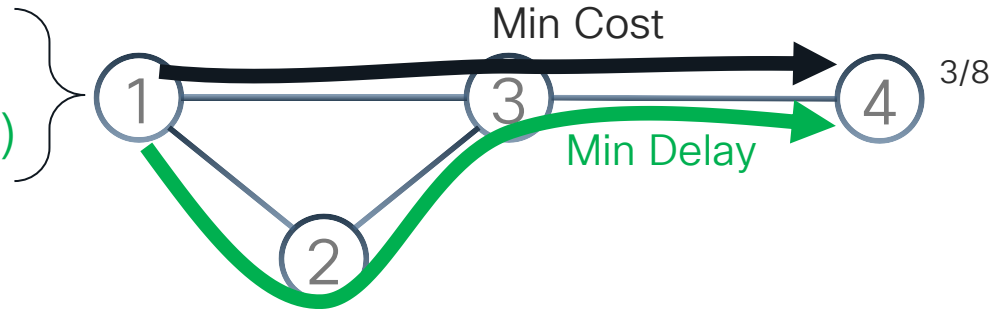
1. E1 maps color **BLUE** to the low-delay intent
2. Upon receiving a service route via E2 with color **BLUE**, E1 automatically instantiates the SR Policy (E2, **BLUE**)
  - This is called On-Demand Next-hop (ODN)
  - Each PE installs only the SR Policies that it needs
3. E1 steers the traffic for prefix W/w onto SR Policy (E2, **BLUE**)

# Need for Per-Flow Automated Steering

Same  
Destination

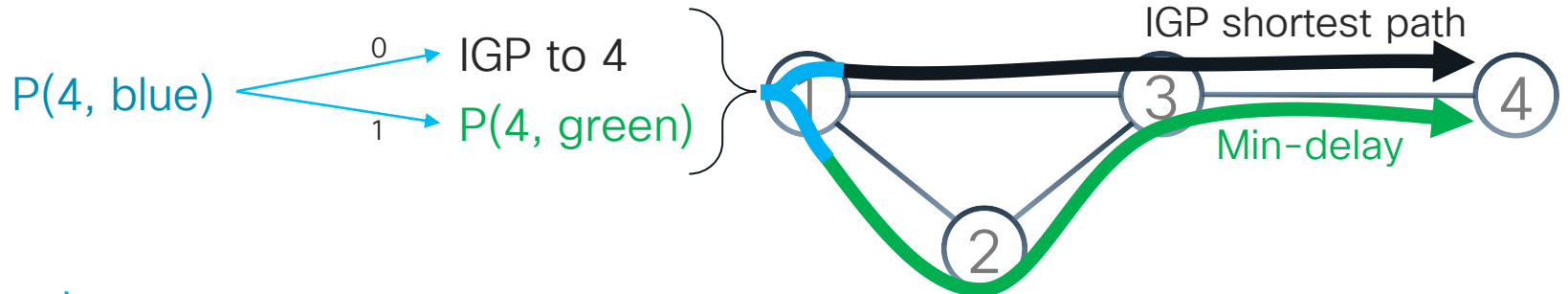


Different  
Flows (flow 1, flow 2)

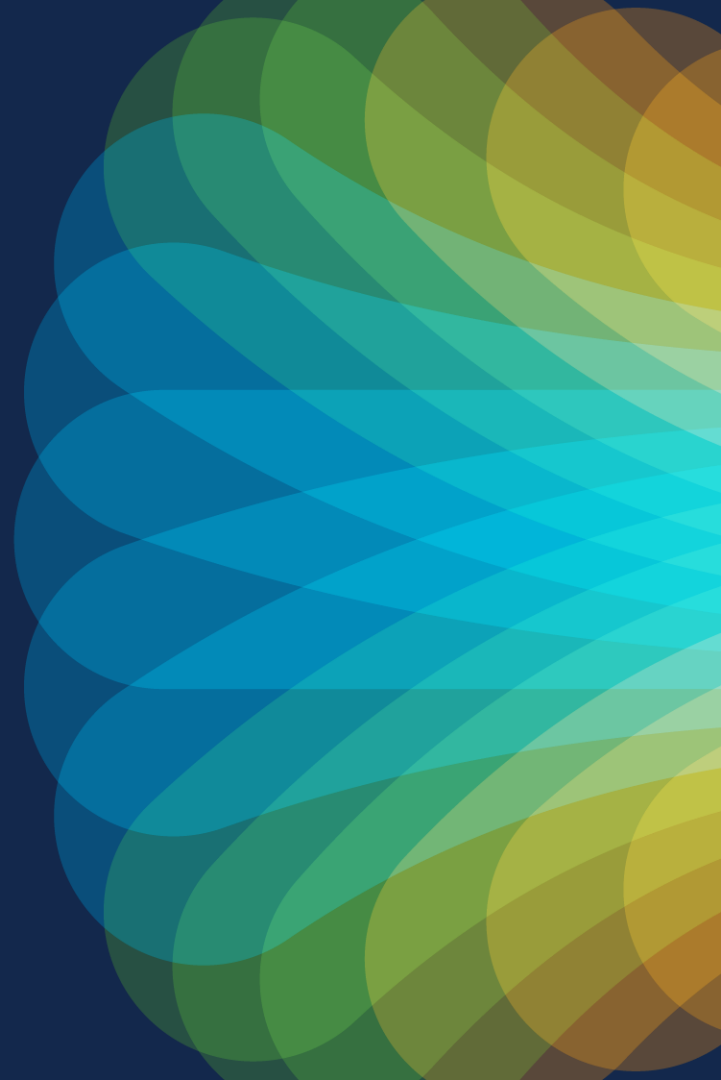


# Per-Flow SR Policy (PFP)

- Per-Flow Policy (Node4, blue) @ Node1
  - FC=0 → IGP shortest path == 16004
  - FC=1 → Per-Destination SR Policy (Node4, green)
- Per-Destination Policy (Node4, green)
  - Defined as Min Delay → <16002, 16004>



# SR Path Computation Element (SR-PCE)



# 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

Cisco Crosswork  
Optimization Engine



Single /  
Multi-  
Domain  
Topology

Computation  
algorithms

SR-PCE runs  
on virtual IOS-  
XR node

IGP  
BGP-LS

PCEP

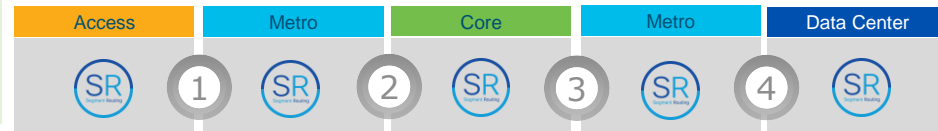
REST API

Topo  
DB

Compute

Collect

Deploy



# SRTE Use Cases

## Distributed or Centralized Path Computation?

Blue = computed by SR-PCE  
Green = computed by head-end

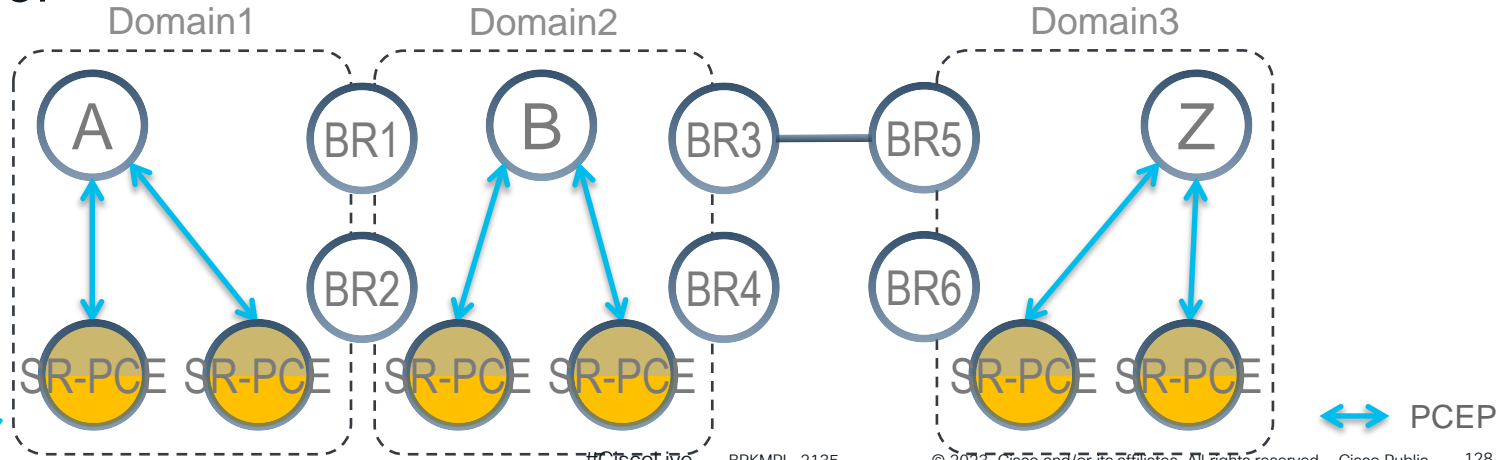
Use Case	Optimization objective / constraints	Single-Domain	Multi-Domain
Reachability	IGP metric + constraints	Distributed or Centralized	Centralized
Low Latency (TE metric)	TE metric + constraints	Distributed or Centralized	Centralized
Low Latency (actual)	Latency + constraints	Distributed or Centralized	Centralized
Path Disjointness	IGP / TE metric + PCEP association group	Centralized	Centralized
Tree-SID	P2MP	Centralized	Centralized

# SR-PCE – Fundamentally Distributed

- SR-PCE not to be considered as a single all-overseeing device
- SR-PCE deployment is closer to BGP RR deployment model
- Different service end-points can use different pairs of SR-PCEs
- Choice of SR-PCE can either be based on proximity or service

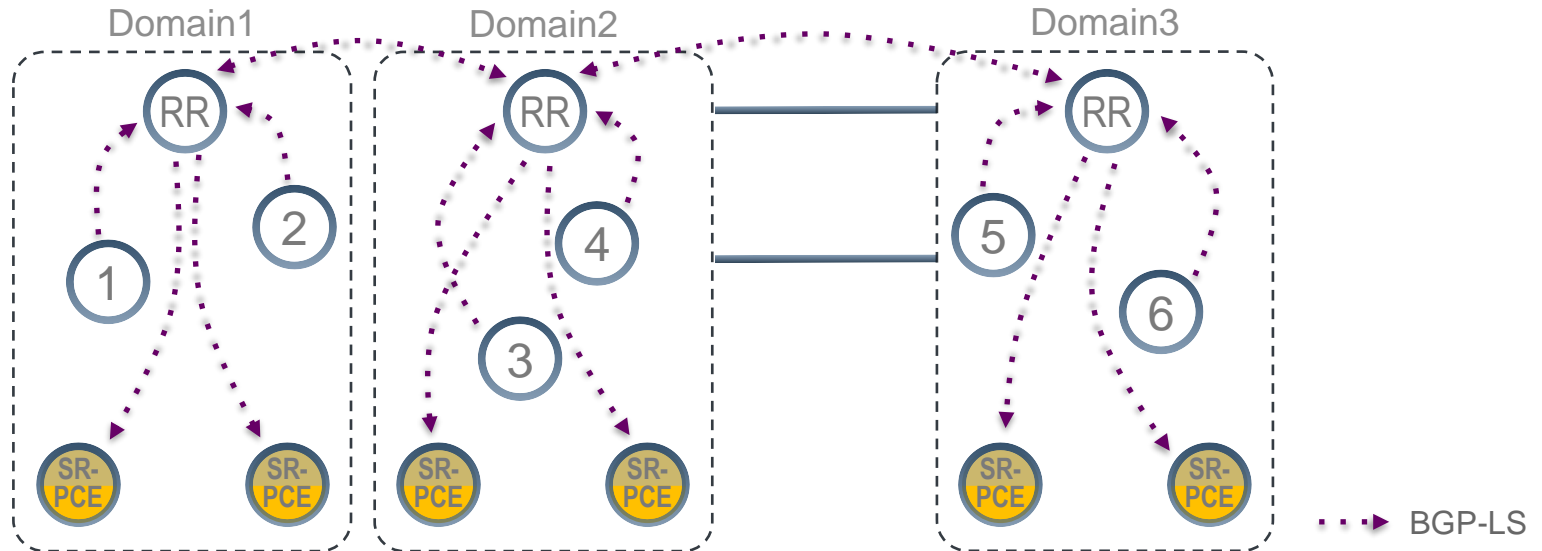
# SR-PCE – Fundamentally Distributed

- Add SR-PCE nodes where needed; per geographic region, per service, ...
  - SR-PCE needs to get the required topology information for its task
    - E.g. to compute inter-domain paths SR-PCE needs the topology of all domains
- Example:



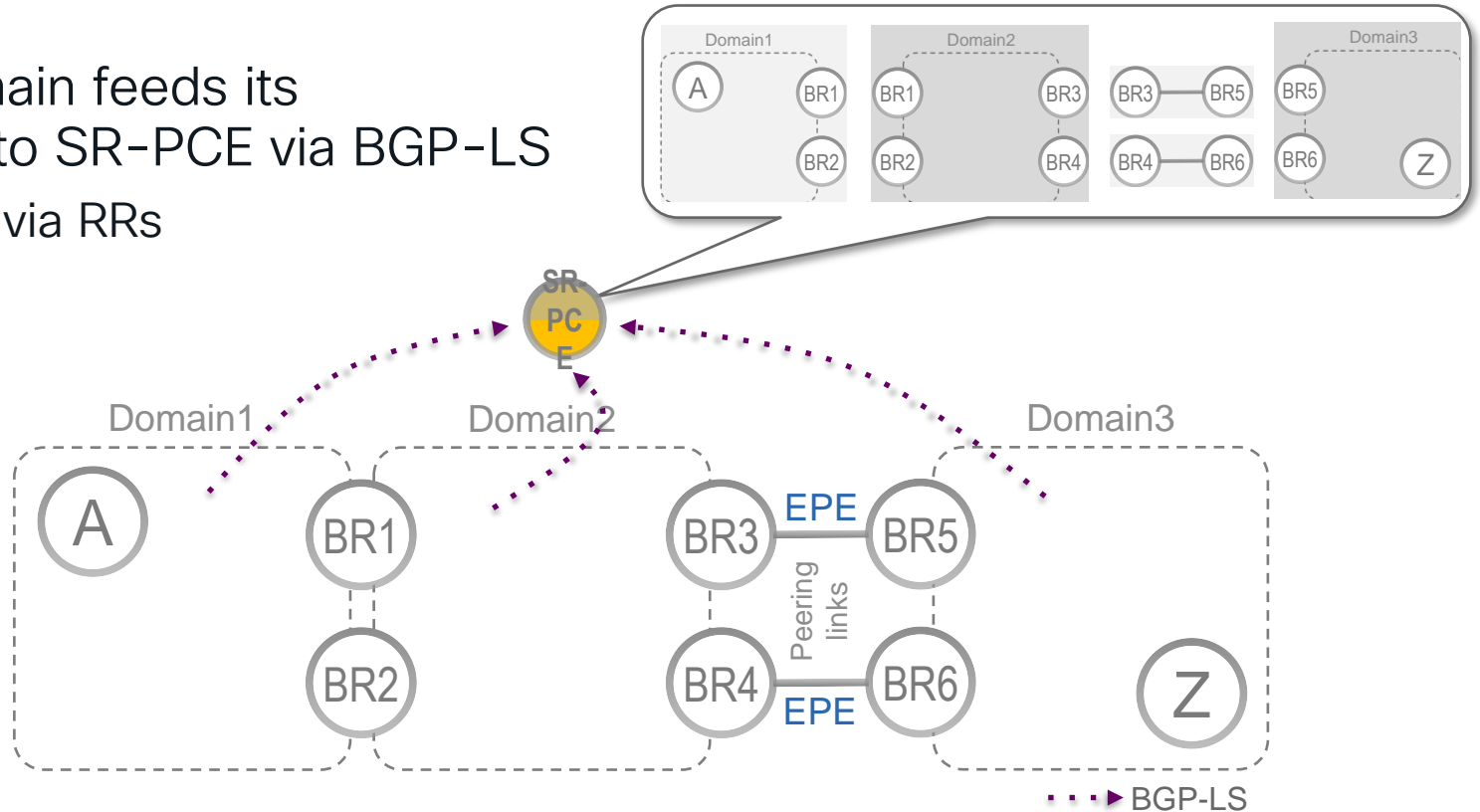
# SR-PCE – Fundamentally Distributed

- Using RRs to scale the BGP-LS topology distribution
- Any node can have a BGP-LS session to the RR



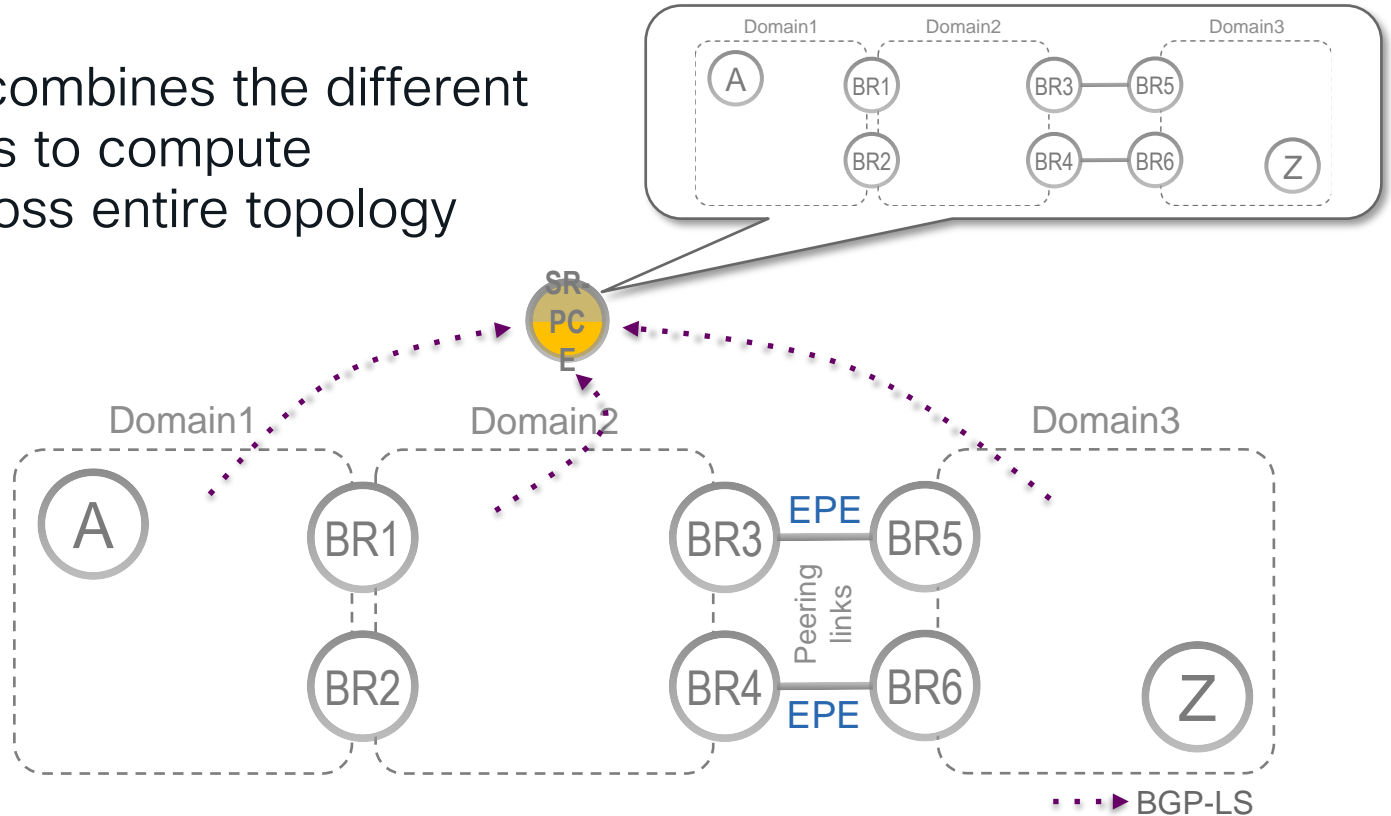
# SR-PCE receives topology of all domains

- Each domain feeds its topology to SR-PCE via BGP-LS
- Typically via RRs



# SR-PCE consolidates the topologies

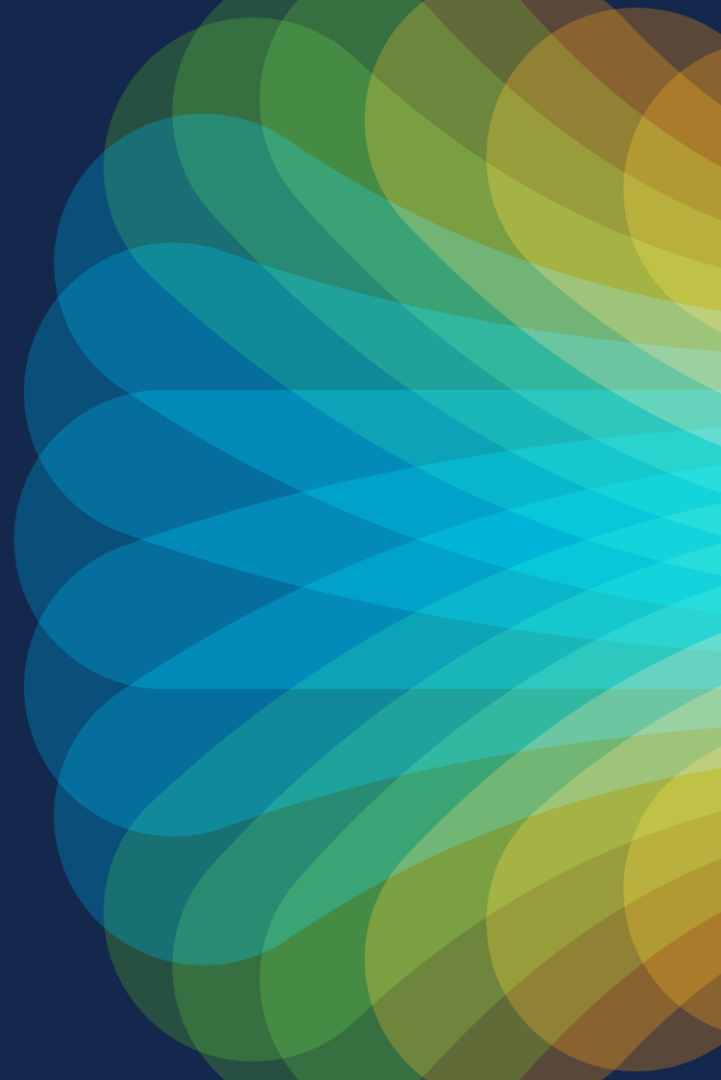
- SR-PCE combines the different topologies to compute paths across entire topology



# SR-PCE – High Availability (HA)

- SR-PCE leverages the well-known standardized PCE HA
- Head-end sends PCEP Report for its SR Policies to **all connected SR-PCE nodes**
- Head-end delegates control to its primary SR-PCE
  - Delegate flag (D) is set in PCRept to primary SR-PCE
- Upon failure of the primary SR-PCE, head-end re-delegates control to another SR-PCE

# SRTE with SR IGP Flexible Algorithm



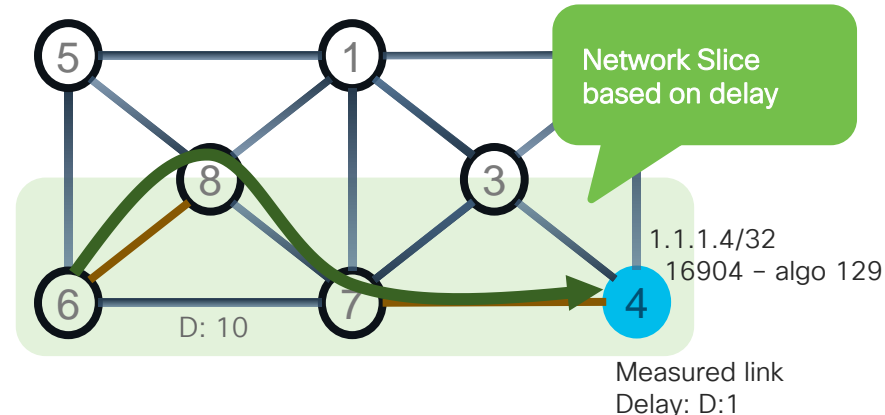
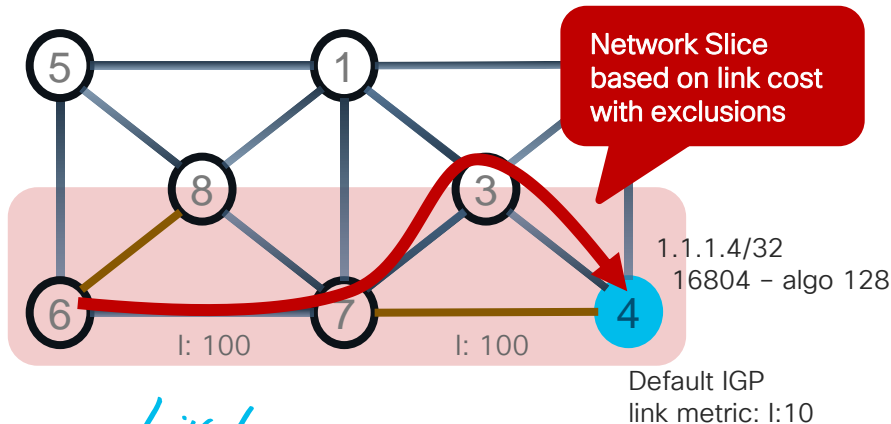
# SR IGP Flexible Algorithm (FA)

- Complements the SRTE solution with customizable Prefix-SIDs
- We call “Flex-Algo”
  - The algorithm is defined by the operator, on a per-deployment basis
- Flex-Algo K is defined as
  - The minimization of a specified metric: IGP, TE or delay
  - The exclusion of certain link properties: link-affinity, SRLG, ...

# SR IGP Flexible Algorithms

- Examples

- Operator defines Flex-Algo 128 as “minimize IGP metric while avoiding links with link-affinity brown”
- Operator defines Flex-Algo 129 as “minimize delay metric”



# SR IGP Flexible Algorithms

- Examples
  - Operator defines Flex-Algo 128 as “minimize IGP metric while avoiding links with link-affinity brown”
  - Operator defines Flex-Algo 129 as “minimize delay metric”

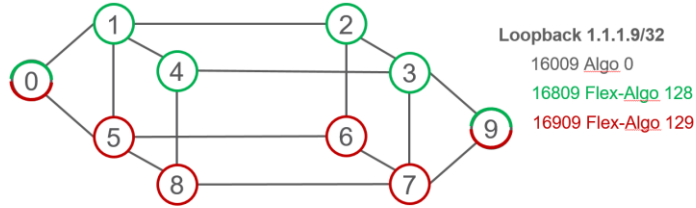
And the question is ... How many labels are needed to enforce traffic on the Flex-Algo path?

A single SID == SRTE with a single SID !!!

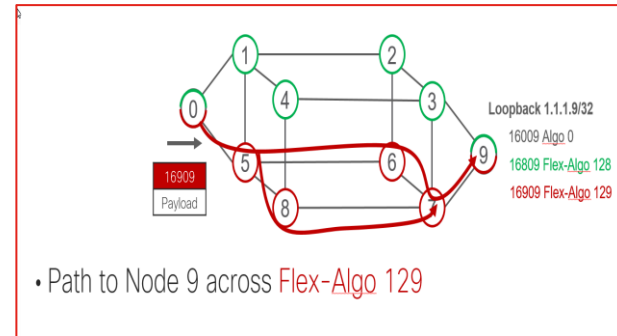
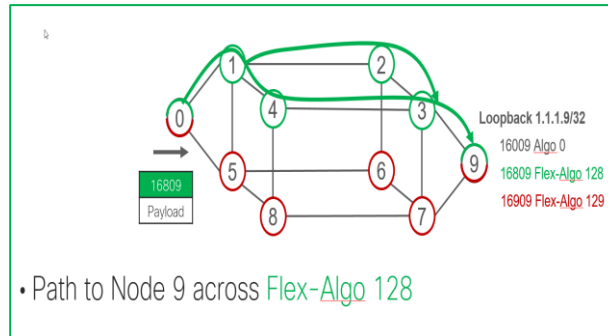
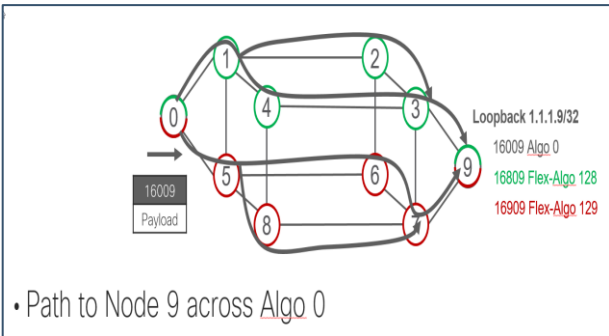
And realized with the simplicity of FA – single IGP process and single Loopback IP address

# Multi-Plane Networks

Powered by SR IGP Flex Algo

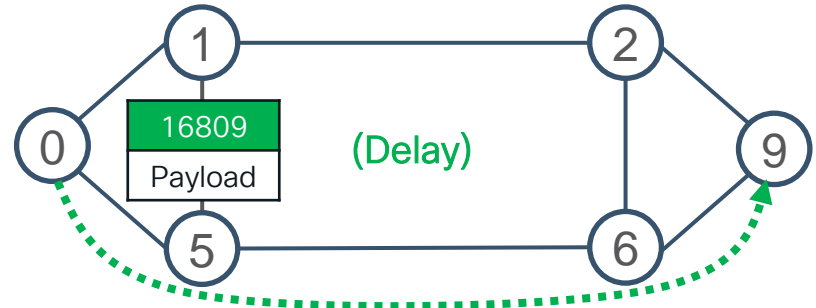
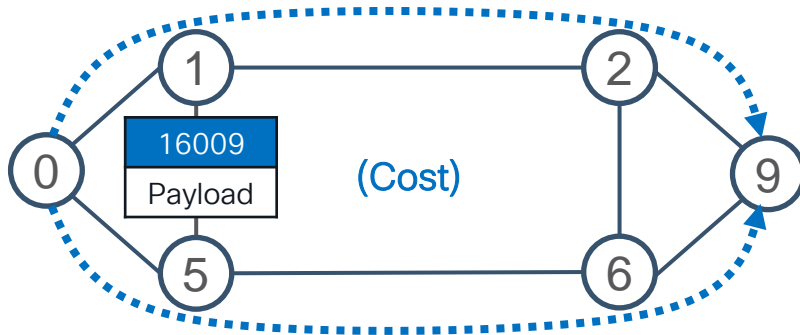
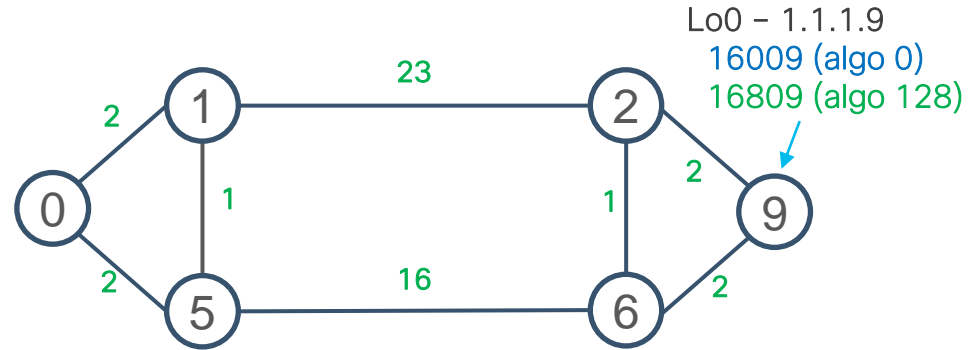


- All the nodes support Algo 0: minimize IGP metric
- Green nodes also support 128: minimize IGP metric
- Red nodes also support 129: minimize Delay



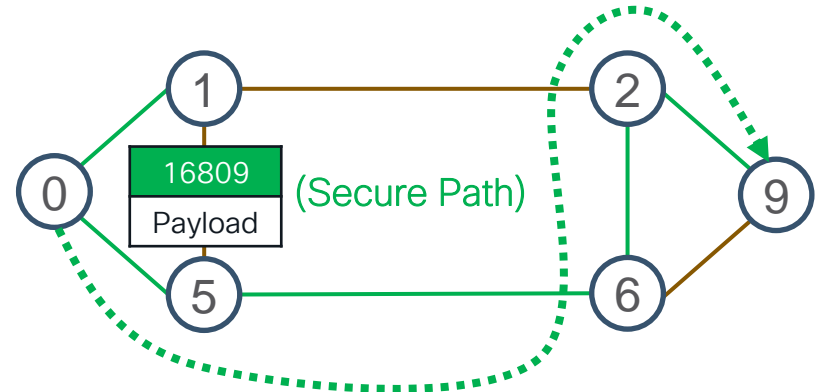
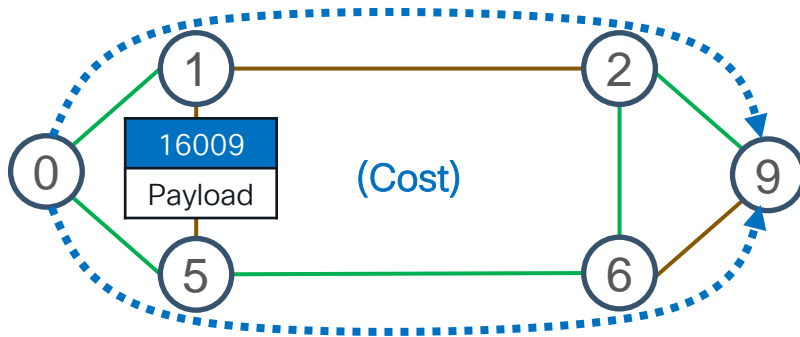
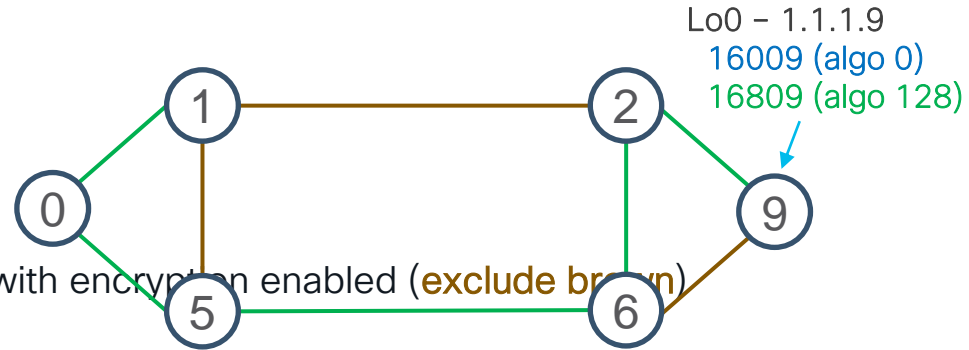
# Use-Case – Delay vs Cost of Transport

- All nodes support Algo 0 & 128
- ISIS link metric 10
- Algo 128: minimize delay metric
- Per-link measurement of delay and advertisement as delay metric via ISIS
- Delay metric at that time shown in green



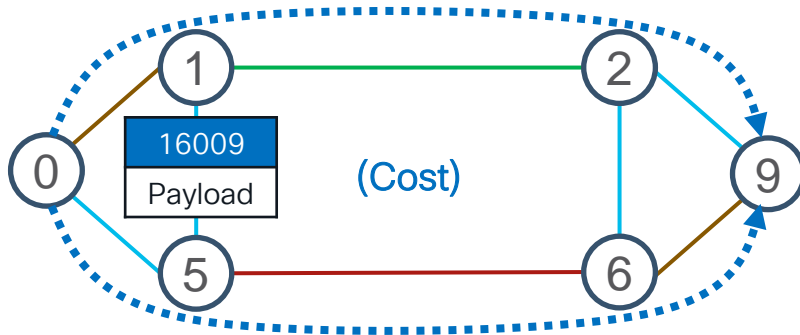
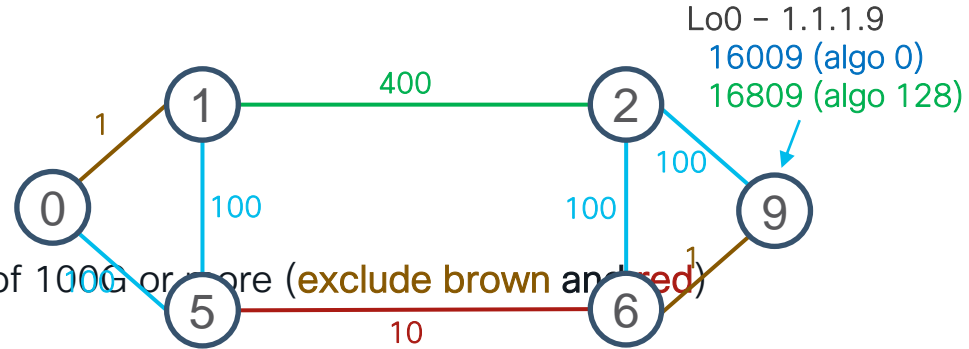
# Use-Case - SRTE for Intelligent Secure Paths

- ISIS link metric 10
- Link colors shown **Unencrypted** / **Encrypted**
- All nodes support Algo 0 & 128
- Algo 128: minimize IGP while traversing links with encryption enabled (**exclude brown**)
- Per-link colors flooded in IGP



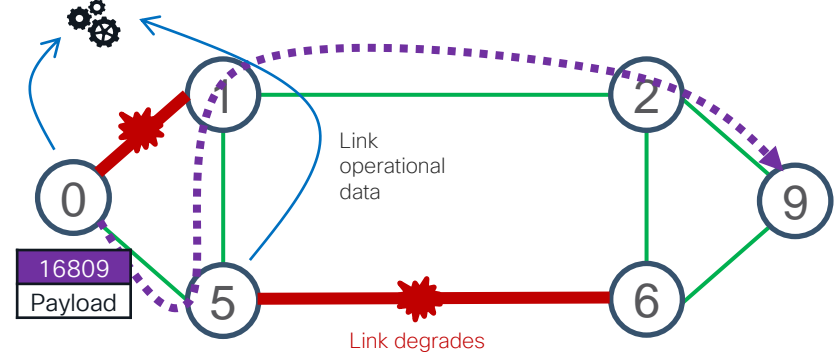
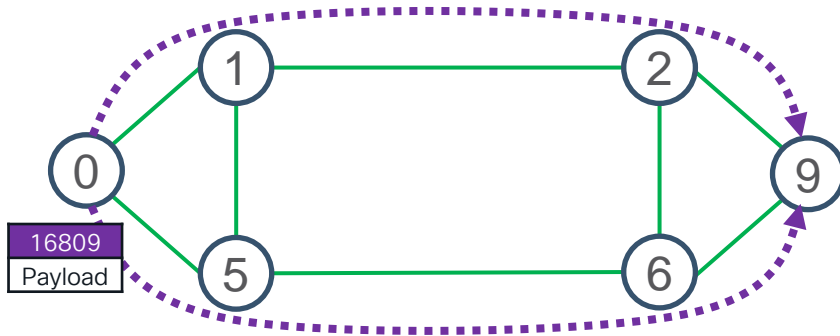
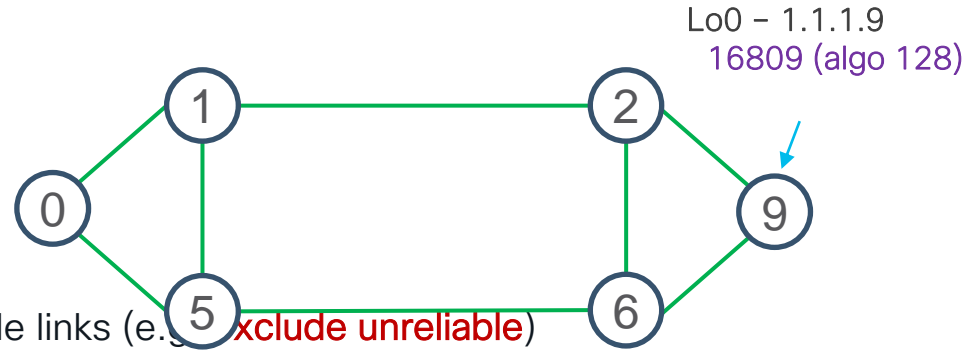
# Use-Case - SRTE for High-BW Links Preference

- ISIS link metric 10
- Link colors shown 1G / 10G / 100G / 400G
- All nodes support Algo 0 & 128
- Algo 128: minimize IGP while traversing links of 100G or more (exclude brown and red)
- Per-link colors flooded in IGP

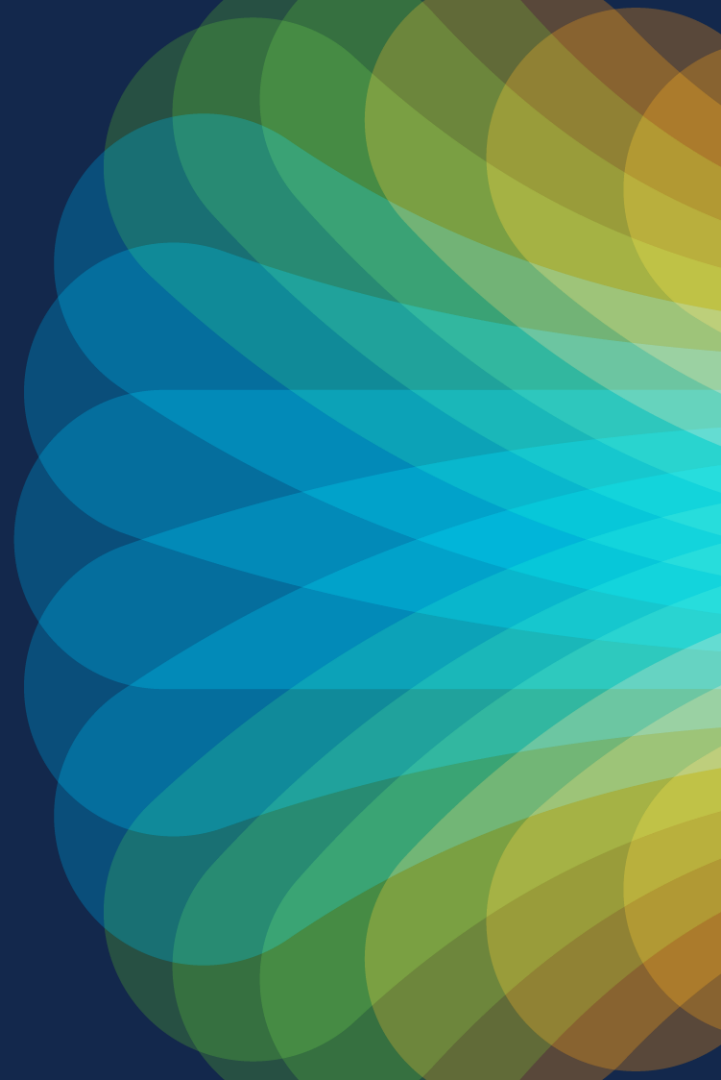


# Use-Case – “Clean” Slice

- ISIS link metric 10
- Link colors shown **reliable** / **unreliable**
  - Reliability of a link based on operator-defined factors
- All nodes support Algo 0 & 128
- Algo 128: minimize IGP while traversing reliable links (e.g. **exclude unreliable**)
- Per-link colors flooded in IGP



# Conclusion





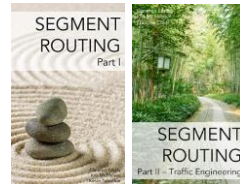
# Simplicity Always Prevails



Furthermore with more scale and functionality



# Stay up to date with...



Segment Routing, Part I / II Textbooks

[Available on Kindle and in paperback](#)

# Fill out your session surveys!



Attendees who fill out a minimum of four session surveys and the overall event survey will get **Cisco Live-branded socks** (while supplies last)!



Attendees will also earn 100 points in the **Cisco Live Challenge** for every survey completed.



These points help you get on the leaderboard and increase your chances of winning daily and grand prizes

# Continue your education

CISCO *Live!*

- 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](http://www.CiscoLive.com/on-demand)

# SR Learning Path

- **Deploying VPNs over SR Made Easy** - BRKMPL-2131, MON 06/05 8:00-9:00AM
- **High Value Wavelength / Private Line Services** - BRKOPT-1005, MON 06/05 8:30-10:00AM
- **SRv6 for Next-Gen Transport Networks** - BRKMPL-2205, MON 06/05 1:00-2:00PM
- **Preparing for a Successful SR Deployment** - BRKMPL-2135, MON 06/05 3:00-4:30PM
- **Enterprise E2E WAN Architectures** - BRKXAR-1003, MON 06/05 3:00-4:30PM
- **Simplify your journey to SR and SRv6 with Crosswork Automation** - BRKSPG-2043, MON 06/05 4:00-5:00PM

# SR Learning Path

- **SRv6 Fundamentals - BRKMPL-2203, TUE 06/06 3:00-4:30PM**
- Circuit-style SR and Service Emulation - BRKMPL-2133, TUE 06/06 4:00-5:00PM
- Deploying and troubleshooting multicast over SR - BRKMPL-2257, WED 06/07 10:30-12:00PM
- SR Implementation in Action: From CLI to Closed-Loop Automation - BRKMPLS-2147, THU 06/08 8:30-10:00AM

# SR Learning Path

- Troubleshooting SR - BRKMPL-3624, On-demand recording
- SRv6 Basics - LABMPL-1201, Walk-in Lab
- Implementing EVPN over SR using NCS 55xx/5xx platforms - LABSPG-3000, Walk-in Lab
- Implementing SRv6 on NCS 55xx/5xx platforms - LABSP-3393, Walk-in Lab
- How to use SR PCE for IOS-XR - LABMPL-2011, Walk-in Lab



Thank you

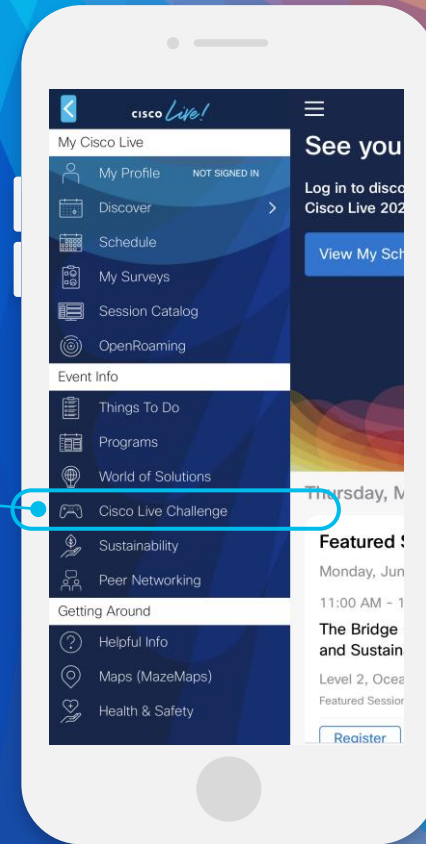
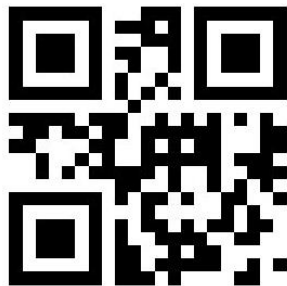


# Cisco Live Challenge

Gamify your Cisco Live experience!  
Get points for attending this session!

## How:

- 1 Open the Cisco Events App.
- 2 Click on 'Cisco Live Challenge' in the side menu.
- 3 Click on View Your Badges at the top.
- 4 Click the + at the bottom of the screen and scan the QR code:





You make **possible**