Segment Routing and Enterprise: What It Is and Why It Matters

Replacing LDP in the Metropolitan Network

A Technical Paper prepared for SCTE•ISBE by

Jason Cole
Senior IP Engineer
Cox Communications
Atlanta, GA USA
404-971-9307
jason.cole@cox.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

1. Introduction

1.1. Purpose

This document details the deployment of Segment Routing in the residential domains. It lays the foundation for the sustaining engineering team to begin low-level planning and testing. The earliest sections of the document explain the value of Segment Routing; primarily in the residential domain. This is done by comparing SR's technical features against LDP and RSVP-TE. Most of this document describes the operational challenges of implementation. The challenges are arranged into three categories:

1. Hardware and Software
2. Global Block (SRGB) Standardization
3. Interworking

Matrices describe vendor software support, as well as hardware and software readiness of each market. The same section details why Segment Routing will not be deployed on all hardware, even though supported in code. Recommendations for code versions are provided for initial lab testing. The document also describes each platform's label depth, factors impacting label depth and why, for the foreseeable future, label depth will not be a concern in the residential network. Index assignments, which are derived from Segment Routing Global Blocks (SRGB), require planning. This document describes factors impacting planning and will recommend the assignment method. Last of the challenges is interworking. This will be illustrated in various drawings and described using CLI configurations. The concluding section will describe the implementation process and a FOA recommendation. It will detail the steps required to get to a "SR Core" in a residential market, using illustrations and configurations.

1.2. Scope

All residential domains will be impacted by these changes. The business and data center domains are not specifically covered by this document. In the initial deployment, only MX and NCS chassis will be impacted. At the time of this writing, there is a plan to integrate a Nokia SR-1 platform as the replacement for the ASR9K COI router. Depending on when the Segment Routing project is implemented, the scope will include that platform. The same could be said about a replacement chassis for the Cisco Nexus router. The Cisco Nexus is currently IP-only (no MPLS) but will likely be replaced by a MPLS enabled device.

1.3. Prerequisites

Readers of this document require a solid understanding of SR-MPLS. This document was not created to explain the technology. Rather, it describes the high-level application and deployment process. This document has not validated code and only serves to provide a code *recommendation*.

Disclaimer: Code testing will be performed by the sustaining engineering team, so no specific code mentioned in this document is guaranteed to be used for deployment.

2. Segment Routing

2.1. History

Because SR (RFC 8402) was designed to replace both RSVP-TE (RFC 3209) and LDP (RFC 3036) for most service provider deployments, RSVP-TE and LDP are briefly described.

RSVP-TE is a feature-rich MPLS protocol. Its most commonly deployed solely for Fast-Reroute (FRR). In fact, a survey[1] indicates that 90% of RSVP-TE deployments are solely for FRR[2]. It's less commonly deployed to leverage advanced features such as bandwidth reservation and auto-bandwidth. Cox deploys RSVP-TE for both reasons described previously. The business domain requires FRR for Metro-E services and the backbone domain leverages FRR, bandwidth reservation and auto-bandwidth for lowering cost in the long-haul fiber network. LDP is, for the most part, a simple protocol. It was the first plug-and-play, non-proprietary MPLS protocol. Its provides the immediate benefits of enabling MPLS in a service provider network. These have traditionally been "BGP-free core" and transport for MPLS-based services (L2VPN/EVPN/L3VPN). Features have continued to be developed even after LDP's inception, but most development has slowed or stopped. Modern LDP code supports IPv6 (RFC 7552), FRR (RFC 8102) and multicast (RFC 6826). These features will be discussed in a later sub-section.

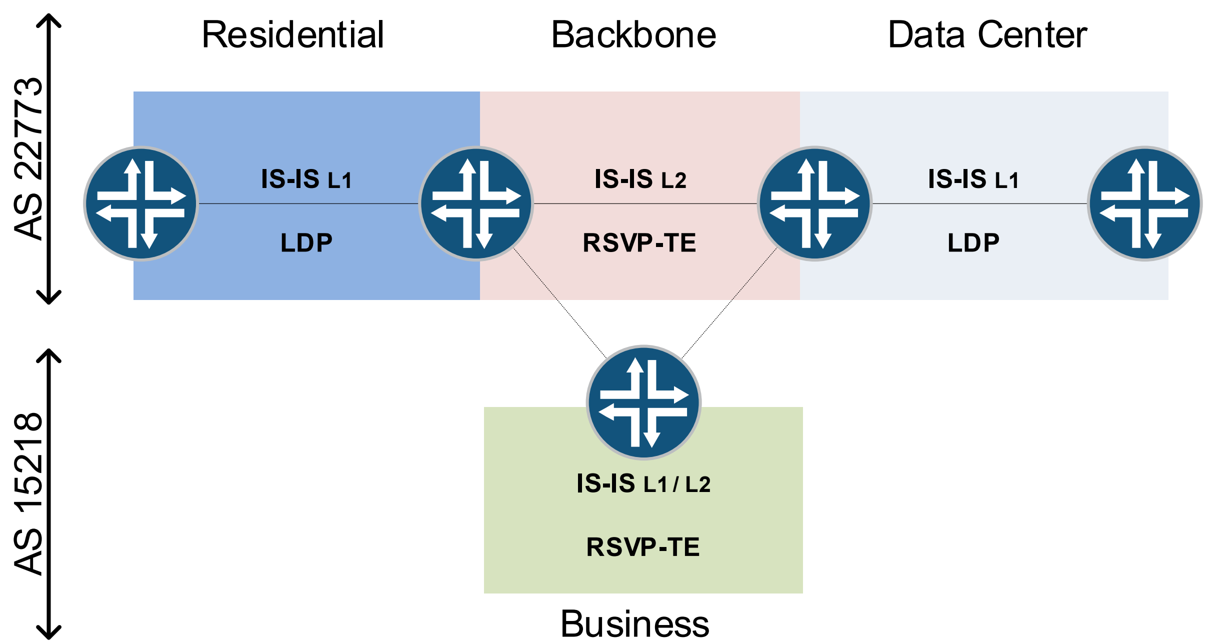

Figure 1 - The four domains and their protocols

Segment Routing is emerging as an elegant technology. It was designed to operate with the simplicity of LDP – being plug-and-play. It also provides *native* support for features like IPv6 and FRR. SR was also designed to support advanced features, such as On-Demand Next-hop (ODN), Flexible-Algorithm and multicast, by leveraging a SR-Path Computational Element (SR-PCE). By decoupling some control-plane functions, the protocol becomes even more flexible. It was designed for this at its inception. Some of the

---

[1] Insert here

advanced features for SR are still in development. Table 1 below describes each domain within the Cox network. The "next-generation protocol" is the strategic vision of Cox's architecture.

Table 1 - Cox's domain protocols

Domain	Existing Protocol	Next-Gen Protocol	SR-MPLS Comments
Backbone	IS-IS L2 / RSVP-TE (FRR / Bandwidth Reservation / Auto-BW)	RSVP-TE	Requires SR-PCE Requires multicast Requires bandwidth reservation Requires high visibility (statistics)
Residential	IS-IS L1 / LDP	SR-MPLS (IPv4/IPv6)	Superior to LDP Requires SRMS
Data Center	IS-IS L1 / LDP	SR-MPLS (IPv4/IPv6)	Superior to LDP
Business	IS-IS L1, L2 / RSVP-TE (FRR)	SR-MPLS (IPv4/IPv6)	Superior to LDP More scalable SR-PCE use-case (inter-area LSPs)

2.2. Evaluation

This sub-section does not compare advanced features of Segment Routing like ODN and Flexible-Algorithm. These features are not applicable to the residential domain, and it is undetermined whether they'll be leveraged in the backbone and business domains, as of the time of this writing. A brief description of these two technologies is provided to reduce obscurity. ODN leverages a SR-PCE to provide a label stack to an ingress LSR, when the LSR does not have reachability information about the egress LSR. This scenario occurs when both LSRs are in different domains (IGP areas or autonomous-systems) and an end-to-end LSP is required. The most likely application of this is in the business domain. Flexible-Algorithm uses interface monitoring probes to derive delay. These can then be translated and used to create a dual plane, IGP topology which may route based on delay. To achieve this, delay-based metrics are created and advertised by the IGP. Then, the nodes create an additional topology map which is independent from the standard, static-metric topology. The ingress LSR will associate a MPLS service to the label stack derived from either topology. The most likely application of Flex-Algorithm are the backbone and business domains. These are advanced features which distinguish Segment Routing from LDP and RSVP-TE, but they're not applicable to the residential domain.

2.2.1. RSVP

To understand SR and its holistic deployment in the Cox network, a few points will address the practical differences between RSVP-TE and SR. When RSVP-TE is used solely for FRR, as is the case in Cox's business domain, SR is the preferred protocol. SR is plug-and-play whereas RSVP-TE requires tunnel configurations. The elimination of MPLS tunnel configuration simplifies automation. RSVP-TE maintains lots of network state. When a network element fails, it creates churn and overhead. Depending on the number of LSPs, this can be quite significant. Segment Routing reduces state, configuration, complexity and convergence time, especially when replacing RSVP-TE. There are many other caveats to RSVP-TE when used solely for FRR. The points above are a shortlist which are most relevant. The backbone domain's deployment of RSVP-TE will not be directly compared against Segment Routing, because a one-for-one swap is not technically feasible at this time, but it is something that will be investigated in the near-term future.

2.2.2. LDP

The residential and data center domains in the Cox network run LDP. LDP adds an additional layer of complexity. It is often referred to as a "piggyback protocol", because LDP relies on IGP entries in the RIB to perform label allocations. Thus, it can be assumed that the IGP has already converged before the LDP process begins. The process begins with each node allocating labels for the prefixes in its RIB. The label information is stored and maintained in the LDP binding table. With LDP, the number of labels is proportional to the number of nodes and links. However, this can be adjusted so that only host routes are allocated labels. Label information is exchanged via LDP adjacencies on a hop-by-hop basis like distance vector routing protocols which "route-by-rumor". Each node in the domain performs the same task until the network has converged. Whenever a network element fails, LDP relies on the piggyback process to converge. This increases the overall convergence time.

Segment Routing eliminates a layer of protocol machinery required for label distribution – LDP is no longer required for label allocation, storage and exchange. SR is unique from LDP in that it leverages the IGP to propagate and store label information. This makes the network more efficient, thus reducing convergence time. In Segment Routing, the number of labels is proportional to the number of global SIDs. Leveraging the IGP for label distribution and storage reduces the overall cost on the protocol machinery. SR inherently eliminates LDP-IGP synchronization issues.

There are three other key features which are supported by both LDP and Segment Routing that may be immediately deployable or are of some interest to Cox in the future. These three features are: MPLS for IPv6, Fast Re-Route (FRR) and MPLS multicast. LDP has been extended to support LDPv6, R-LFA and mLDP. These enhancements are for the most part successful and effective. However, some LDP extensions are not as efficient as Segment Routing's built-in features. Segment Routing's equivalent features are: SR-MPLSv6 (not "SRv6"), Topology Independent-Loop Free Alternate (TI-LFA) and Tree-SID. The three features require a brief comparison.

2.2.2.1. IPv6

IPv6 support in LDP (LDPv6) creates additional state in the network; both LDPv4 and LDPv6 processes run concurrently. This means that additional adjacencies and tables are required for label exchange and storage. SR uses a small TLV for the IPv6 node-SID, which is for the most part identical to the IPv4 node-SID. It is transported in the same IGP advertisement. It requires less memory for state and less information is exchange than LDPv6. Thus, SR-MPLSv6 is more efficient than LDPv6.

2.2.2.2. Multicast

Cox's current multicast deployment in the residential and data centers domains is referred to as "IP multicast" and not "MPLS multicast"; multicast forwarding is done via PIM and multicast packets are not label switched. Multicast LDP (mLDP) is an effective and simple protocol. However, this add-on, like LDPv6, adds additional state to the LDP binding database that is not required using SR. SR's Tree-SIDs avoid additional overhead on nodes by using a SR-PCE. Thus, mLDP adds more cost to the routers than Tree-SIDs. However, Tree-SIDs move cost from the routers to a SR-PCE. An alternative to Tree-SIDs is Bit Index Explicit Replication (BIER). Prototypes have shown that it may be the next-generation MPLS multicast protocol of choice, due to its efficient addressing schema. BIER is recommended to be run in tandem with SR for maximum efficiency. As of the time of this writing, MPLS multicast is not planned for implementation in the residential and data center domains. This comparison is provided for a comprehensive analysis.

2.2.2.3. Fast-Reroute

TI-LFA provides 100% link and node protection and micro-loop avoidance in all topologies without the need for T-LDP, unlike R-LFA. Another plus for TI-LFA is that it will always route traffic on the post-convergence path, where R-LFA may have to move the protected traffic off the repair path and then onto the post-convergence path. It is worth noting that SR supports SRLG and TI-LFA can account for SRLGs, unlike R-LFA. As a side not, both TI-LFA and R-LFA support link and node protection. Contrary to the previous points describing TI-LFA's benefits, Cox's residential and data center domains use hub-and-spoke topologies. Such a topology should be fully covered by R-LFA, because there should only ever be two paths out of a hub-site. For instance, regarding TI-LFA's post-convergence optimization, the repair path should also be the post-convergence in link or node failure scenarios using R-LFA. Thus, the primary benefit of TI-LFA in the residential and business domains is micro-loop avoidance and not post-convergence optimization.

Table 2 - Summary of technology comparison

Feature	SR-MPLS	LDP
Configuration	IS-IS	LDP & IS-IS
Synchronization	Not required	IGP/LDP
State Information	No state	Minimal state
Protection	Link/Node (TI-LFA), SRLG	Link/Node (R-LFA) using T-LDP
IPv6 Support	IPv6 over SR	LDPv6
Multicast/P2MP	Not supported as of 2018	Supported
Anycast	Supported	Not supported
Flex Algorithm	Supported	Not supported
Label Learning	Leverages IGP	Piggybacking protocol
No. of Labels	SIDs proportional to nodes	Proportional to nodes or nodes & links
Label Scope	Local or Globally Unique	Locally unique
Egress Peering Engineering	Supported	Not Supported

2.3. Value

Traditional MPLS networks lack flexibility and programmability. Segment Routing provides this, at scale, with advanced capabilities such as On-Demand Next-Hop, Service Disjointness, Flexible-Algorithm and Service-Chaining. These advanced features, along with others, can be leveraged using a SR-PCE deployed in a centralized or hybrid deployment model. New capabilities supported in Segment Routing, which are being driven by new customer services, are causing network architects to prepare their infrastructure for the future. The new infrastructure should be able to provide end-to-end paths, between multiple domains, that meet higher-bandwidth and lower-latency requirements than before.

Although Segment Routing provides a new architecture for service providers, it will be deployed in its most basic form in the residential domain. This will introduce support for IPv6 and FRR. These features should be standard is today's networks. It will eliminate LDP-IGP synchronization outages. It is plug-and-play and will require minimal configuration[3]. This will remove one layer of protocols and simplify the network. However, the transitional state, which will run both LDP and Segment Routing, will introduce complexity. This complexity can be overcome, if new software/tools and proper training is provided. Given the Segment Routing architecture, it makes sense for Cox to begin migrating its network domains to Segment Routing to prepare for the future.

3. Hardware and Software

3.1. Platforms

There are five platforms in the residential domain which are Segment Routing capable. The ASR9K is SR capable, however, it has been decided to not enable SR on it; this is explained in a later sub-section. Also, the QFX[4] and SR-1 have not yet been deployed in the Cox network, so these are not an immediate concern. The five platforms are listed below:

1. Vendor A / platform A
2. Vendor A / platform B
3. Vendor B / platform A
4. Vendor B / platform B
5. Vendor C / platform A

Each platform's support for SR is detailed in the table below. "SR Support" does not include sub-features required to support Cox's deployment of SR; it only indicates minimum support. Also detailed in the table below is Cox's current and future codes. There is a subset of SR features that each platform's code should support. Cox should plan to deploy Segment Routing based on the dates of these feature's availability. There are some features which are mandatory and others which are not. Mandatory features should determine target dates. Optional features are flexible when determining target dates.

The Juniper QFX will be deployed for two different use-cases: Remote-Phy (R-PHY) and service aggregation[5]. When deployed for R-PHY, the QFX does not require MPLS. Thus, Segment Routing is not applicable in this scenario. Using the QFX for service aggregation is still in the high-level design phase. However, it will be deployed with MPLS, unlike the R-PHY deployment. Thus, Segment Routing should

[3] A Segment Routing Mapping Server will be required to interwork LDP/SR. This requires using a static one-to-one or dynamic block mapping between a label(s) and prefix(s)
[4] Deploying the QFX as a service layer router (SLR) as a Nexus replacement is a current discussion.
[5] Currently, this has not been finalized.

be enabled on the QFX in this instance. The two deployments are not discussed any further in this document, and the specific design documents should be referenced for further details

The Nokia SR-1 is the replacement for the Cisco ASR used for COI services. Currently, the ASR used for COI may also terminate Metro-E services. This will not be the case with the SR-1. The current design, which is near finalization, will deploy it identically to the Cisco ASR. Thus, it will be in the residential domain and MPLS enabled. Because this device is required to only support Internet traffic, it does not technically require MPLS now.

3.2. Label Depth

A platform's supported label depth is a concern with SR-MPLS. SR will increase the size of the label stack in the following cases: TI-LFA, SR-TE and traffic accounting. In Cox's residential deployment, TI-LFA is the concerning factor. Label depth is limited by both software and hardware. Typically, large differences in label depth limits are seen in platforms leveraging merchant and custom silicon. "Silicon", in the context of this sub-section, refers to chipsets, or components, within a platform that impose forwarding limitations like label depth. The time it takes for a vendor to adjust forwarding capabilities, like label depth, is often determined by whether the platform is merchant or custom silicon based.

Merchant silicon is designed and manufactured by a 3rd party – not by a vendor. The 3rd parties produces platform components, like ASICs, that adhere to industry standards. A vendor will then implement it as a component within their platform. Leveraging a 3rd party reduces cost. However, 3rd party manufacturers tend to be slower to adapt to industry changes. Examples of platforms using merchant silicon are the Juniper PTX 10K, Juniper QFX 5K and Cisco NCS 5K.

Custom silicon is completely designed and produced in-house by a vendor. This often makes a product costlier than if it were to leverage merchant silicon. However, it allows the vendor full control and flexibility to modify their platform to meet new requirements as the industry changes. These platforms tend to be richer in features. They also provide more accountability, because the vendor owns the entire product. The Juniper MX, Juniper QFX 10K and Cisco ASR are examples of custom silicon.

3.2.1. TI-LFA

Metro Ethernet over HFC is a commercial VPN service that exists in the residential domain. Metro-E over HFC, along with voice services, are drivers for implementing TI-LFA. TI-LFA inherently protects label switched traffic. Because video services are IP multicast in the residential domain, they will not be protected. Services may be disrupted by fiber cuts or node failures, when a fast re-route mechanism is not in place. TI-LFA can operate in either link protection or node protection modes.

Both link and node protection leverage additional SIDs, ultimately more MPLS labels, to route around link or node failures in under 50ms. The question becomes, "how many labels are required for link protection and node protection?" The case study below provides insight to this answer. Clearly, link and node protection have different label requirements. Also, the number of labels required by either protection mechanism varies. The additional labels for FRR capabilities adds cost and complexity *to the data-plane*. This is different from traditional FRR techniques provided in RSVP-TE for example, which add additional network state *to the control-plane*.

The topology and metrics of a network determine the number of SIDs required in either mode. The residential and business domains are "hub-and-spoke" topologies with static metrics. The number of labels required for each mode was found by modeling the deployment of both using the Phoenix market's IS-IS database, because it contains the most network elements. The results are below and the model has been attached.

Node Protection SIDs: 4

Link Protection SIDs[6]: 2

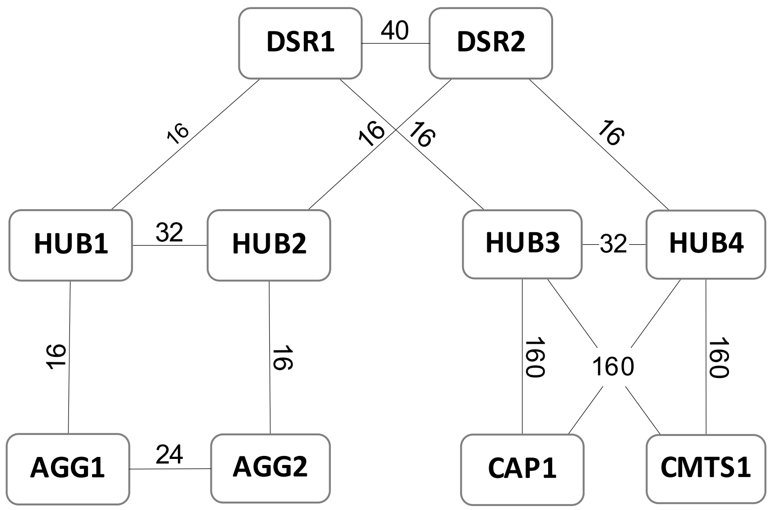

Figure 2 - Current residential topology and metrics used by Phoenix

Once the number of labels required for both protection modes was found, the cost of additional labels, per protection mode, had to be weighed against the likelihood of network element failures within the path of a Metro-E service. As this time, the majority of Metro-E over HFC services are type E-LINE and E-TREE. The remote ends of the service most commonly terminate in the business domain. Cox has two different methods for inter-domain services: BGP-LU and L2VPN Inter-AS Option A. These methodologies and designs are beyond the scope of this document. The key relating to TI-LFA is that the LSPs for the service terminate between the CAP/CMTS and DSR/COI router. The network elements in this path carry the most weight in the decision.

The illustration below describes two LSPs for both inter-domain methods. A service can use either LSP, depending on how it is built. CAP is a headend for both LSPs. The blue LSP's tailend is the DSR. The red LSP's tailend is COI. Both LSPs traverse the link which is most susceptible to fiber cuts (HUB-DSR). None of the nodes associated to these two LSPs are deployed with a non-redundant control-plane (RE/RSP/supervisor)[7]. Thus, node protection provides little value. It is worth noting that only "P" nodes without redundant control-planes are relevant to node protection. For instance, if the ingress or egress nodes have a non-redundant control-plane, that failure would cause the service to go down entirely; TI-LFA provides no protection in that case.

6 Technically, the model revealed that 3 labels were required, but this was due to an incorrect metric on a link. So, 2 is the true number based on a standard topology and metrics.

7 The business team plans to deploy a Nokia SR-1 as the new COI platform. This platform does not provide control-plane redundancy.

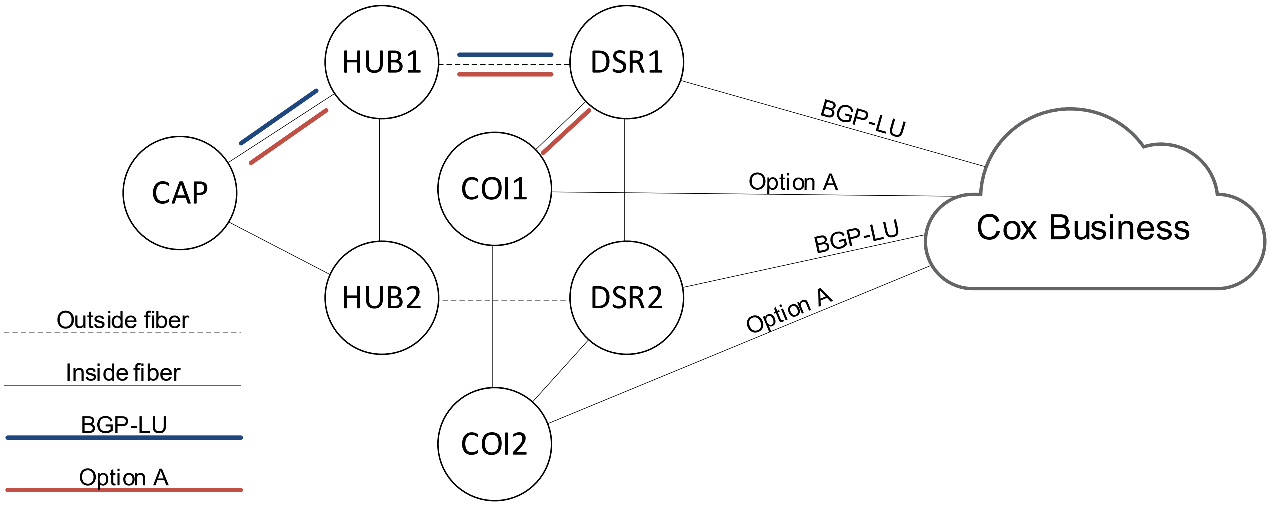

Figure 3 - Inter-domain services using BGP-LU and Inter-AS Option A

4. Global Block

4.1. Overview

MPLS nodes use the Label Switching Database (LSD) to store labels. The total number of labels supported by the LSD is defined by the operating system, it is normally fixed and only adjustable in software upgrades. The LSD is partitioned such that "label spaces", which are unique ranges of labels, are defined for different MPLS applications. Examples of MPLS applications include: RSVP-TE, SR, LDP, BGP-LU, L3VPN, L2VPN, etc. Normally, applications share label spaces, and they're most commonly group by their application type: "service" or "transport". When an application requires a label, it will make a request to the label manager, the process responsible for allocating labels to the applications, it will receive one from its designated label space. Segment Routing is unique from traditional MPLS applications, because a unique label space must be defined by the network operator that only SR-MPLS will use. This is known the Segment Routing Global Block (SRGB).

In SR-MPLS, there are two general types of label spaces. The SRGB is used for *global* labels like node-SIDs. Node-SIDs are globally unique identifiers which require static assignments. Thus, planning is required. The Segment Routing *Local* Block (SRLB) is used for *local* labels like Adjacency-SIDs. Adjacency-SIDs are dynamically assigned by the router without any type of operator configuration. The SRLB is the term used to describe the existing, "dynamic" label space. This label space is shared with other MPLS applications; only the SRGB label space is dedicated to SR-MPLS. SID type details can be found in RFC8402.

Each node in an IGP area, which is enabled with SR-MPLS, requires static label assignments from the SRGB. These label assignments are referred to as "indexes" and are relative to their position within the SRGB range. The example below provides an example of a label assignment using a SRGB and index.

If SRGB is [16,000 (min) - 23,999 (max)]:
Index Range = 1 – 7,999

If index is 1:
Node Label = 16000 + 1

Figure 4 - SRGB and index example

Each vendor defines their platform's label spaces differently. There are several attributes that may differ between vendors. A few significant examples: the number of unique label spaces, the size of the labels spaces, association of MPLS applications to label spaces and the default label spaces.

4.2. Planning

This sub-section explains the planning of the SRGB and index assignments for all domains at a high-level and provides recommended SRGB and index assignment methods. To do this properly, the network should be approached holistically like IP addressing. The approach assumes Segment Routing in *all* network domains – backbone, data center, residential, business. The proceeding sub-sections will describe:

1. Features requiring assignments
2. Reuse
3. Using the same SRGB
4. Selecting the size of the SRGB
5. Deriving index values
6. Partitioning the indexes/SRGB

4.2.1.1. Mapping Server (SRMS)

SRMS is currently in draft[8] but is supported by almost all vendors already. The Segment Routing Mapping Server feature allows for LDP/SR interoperability. Using IS-IS TLVs for Segment Routing, each SR-MPLS node advertises its SRGB. Each node also advertises the loopback prefix and node-SID binding within the IS-IS TLV. Thus, any node not advertising this is not SR-MPLS enabled. Nodes which do not support Segment Routing are allocated a node-SID by the mapping server and this is advertised using another IS-IS TLV. These allocations can be done in two ways: using 1-to-1 (node-SID to node) or using a pool of node-SIDs. Either way, a block of node-SIDs must be reserved for the SRMS(s). It is required that Cox plan for SRMS index allocation in the residential domain.

4.2.1.2. Flexible Algorithm (Flex-Algo)

Flexible Algorithm is currently in draft[9]. Flex-Algo can be used to create a dual-plane topology. Traditionally, there has a been a "single-plane", or "single topology". A topology consists of nodes, links and metrics. Metrics often reflect the bandwidth of a link. In this case, the preferred path is the one with the most bandwidth. Cox does not use bandwidth-based metrics in the residential domain. Rather, static metrics are used to try to load-balancing or avoid certain links. In either case, a "single-plane" is constructed and used to route labeled and unlabeled traffic.

[8] https://datatracker.ietf.org/doc/draft-ietf-spring-segment-routing-ldp-interop/
[9] https://datatracker.ietf.org/doc/draft-ietf-lsr-flex-algo/

"Dual-plane" introduces an additional topology. A second topology can be constructed using the same nodes and links. The metric, or "constraint", can be unique to that topology. For instance, the metric for the second topology could be based on delay. Thus, two topologies exist. One topology based on bandwidth. Another based on delay. The shortest path between two nodes could be different whether you route based on bandwidth or delay. A service can then be associated to either topology – service to LSP mapping.

4.2.2. Reuse

The planning of SRGBs and SIDs is like IP addressing. Like IP addressing, SIDs *can* be reused. However, there are restrictions, similarly to when duplicate IP addresses are assigned. For instance, duplicate IP addresses can be used when the nodes do not require direct communication. Issues may arise down the line when requirements change. This may cause a massive undertaking to re-IP the nodes[10]. In an ideal SR-MPLS deployment, every node in the autonomous-system is assigned a unique SID and a consistent SRGB. This allows end-to-end, inter-domain LSPs via IGP route leaking or BGP-SR. However, when reuse is leveraged, inter-domain LSPs via these two methods is not available and only a SR-PCE can accomplish this. Currently, Segment Routing is in its infancy and side-effects from reuse are not fully realized. Thus, it is a best practice to avoid reuse when possible.

Reuse allows SID assignments to scale "infinitely". It is only the largest carriers, having a single domain requiring more SIDs than the largest available SRGB, that may *intentionally and initially* design for reuse.

A completely different carrier may *unintentionally* reuse SIDs when the number of required SIDs exceeds the number of available SIDs than the *existing* SRGB can accommodate. Growth in SID requirements may be caused by drastic, unanticipated changes in the network. For instance, applications[11], planning[12], mergers, acquisitions, collapsing[13], etc. It is challenging to plan for changes that are unlikely to occur. By doing so, the designer may over plan and create an entirely different set of challenges. Rather than completely avoiding planning for unanticipated, drastic, network changes, it is recommended that Cox plans for reuse but only "tactically"[14]. In this case, reuse is only to be leveraged when an emergency arises.

For example, suppose the number of LSRs (MPLS enabled routers) in ASN22773 is 2,000. A SRGB supporting 16,000 SID assignments accommodates this requirement, as well as provides room-to-grow and flexibility in planning. Suppose the business domain collapses into the residential domain and 1,000 new LSRs are installed in each market in ASN22773. If Cox has 20 markets, 20,000 additional, unique SIDs are required. Thus, a total of 22,000 assignments are needed. The current SRGB of 16,000 configured on LSRs cannot accommodate this. It is a large undertaking to change the SRGB on the 2,000 existing LSRs[15]. If tactical reuse is initially planned for, it accommodates this requirement. Tactical reuse is described in further detail in this section.

---

[10] At one point, Cox used duplicate IP addresses in markets, when it was expected that the nodes would not require direct communication. Since that time, extensive work has been done to undo the planning error.
[11] SR-MPLSv6, Flex-Algo, etc.
[12] Partitioning node SIDs via geographical location, owner, function, etc.
[13] Collapsing the business and residential networks
[14] This is not a technology specific term
[15] Each LSR would require a reboot

If just 1,000 SIDs are set aside for reuse at the initial stage of planning, all assignments can be accommodated[16]. No configuration changes are required on the existing LSRs; they would continue to have globally unique SIDs; and there would still be ~15,000 unused SIDs for growth. The new, business LSRs in market #1 can be assigned SIDs from the reserved pool. The new LSRs in market #2 will be assigned the *same* SIDs from the reserved pool as market #1. If end-to-end inter-domain LSPs are required between the new devices, a SR-PCE is required if SR is to be leveraged. This is a "tactical" approach for reuse. It provides an additional degree of flexibility without over accommodating for large, unlikely changes.

Because of this, domain owners in the same autonomous-system can use the same SRGB and scale for unexpected changes. Thus, it is recommended to use the same SRGB in the backbone, data center and residential domains. Like all addressing deployments, the selected SRGB must be partitioned in a way that scales for the domains, markets, applications, projected nodal growth and unexpected tactical reuse. This allows for maximum scale, simplifies planning and eases troubleshooting concerns.

NON-TACTICAL SCENARIO

Rather than explaining tactical reuse first, this section describes the Cox network if it were designed with reuse for all nodes at initial deployment. **This is not recommended and is only used to illustrate the concept of reuse**. When planning index assignments for a single autonomous system,

:

1. Leaking from L1 to L2 should not be performed
2. Indexes on L1 routers must be unique within the area but can be re-used in other L1 areas
3. Indexes on L1/L2 and L2-only routers cannot be re-used

The figure below illustrates index assignment rules and tactical index reuse using ASN22773. This case requires a SR-PCE to facilitate the creation of an end-to-end LSP. The LSP is explicitly routed at each hop in the path. Index assignments are unique on all nodes in the topology except for the edge nodes (100,100) in each L1 area. There are two "core" nodes (101,102 & 103,104) in each L1 area. There are four L1/L2 nodes[17] (201,202,203,204). There are two L2-only nodes[18] (301,302). L1/L2 and L2-only indexes must be unique in this design. Also, this design cannot leak IS-IS prefixes between L1 and L2 areas. The diagram describes the stack, which can be thought of as indexes or MPLS labels, that is provided to the ingress LSR by the SR-PCE. The label of interest is label 100 which is being reused tactically. Although the ingress LSR has been assigned index 100 to itself, it is also able to use a label stack which contains its own label. This will be properly routed to the remote edge node (100).

[16] This assumes that route leaking between the markets does not occur and when/if an end-to-end LSP between the markets is required, a SR-PCE is used to build the SID list.
[17] DSRs/HDRs are L1/L2 routers.
[18] BBRs are L2 routers

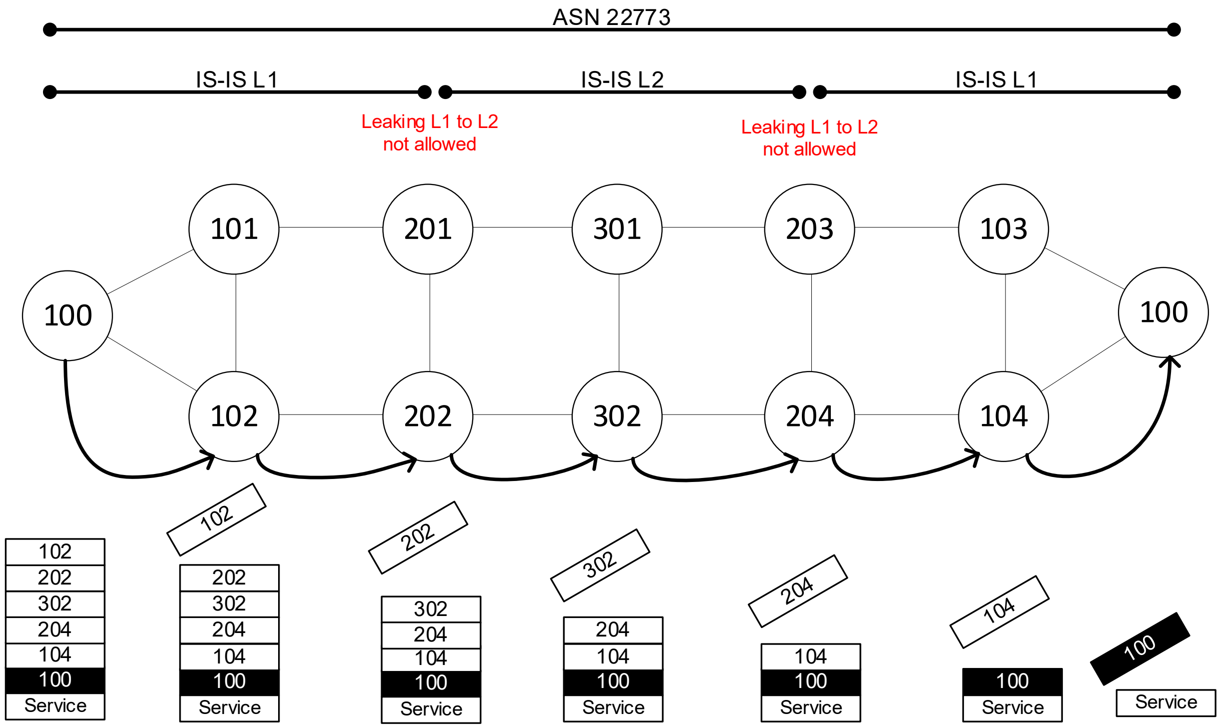

Figure 5 - Reuse within an autonomous-system using a SR-PCE and explicit hops

4.2.3. Consistency

As described in the previous sub-section, index assignments can overlap in different domains; exceptions to this were also described. Examples used a consistent SRGB; the SRGB was the same on all nodes in all domains. Vendors recommend this method. If domain owners select different SRGBs upon deployment of SR-MPLS, inconsistency occurs. Although this is technically feasible, it complicates troubleshooting and verification, particularly with SR-TE and a controller. Another reason for inconsistent SRGBs is when a domain grows beyond the total number supported indexes within the SRGB. If this occurs the SRGB must be adjusted on nodes in the domain[19]. It is recommended that Cox select a consistent SRGB for all domains that scales.

4.2.4. Size

Below is a table summarizing SRGB parameters of platform A, platform B, platform C, platform D. These values are used to determine available SRGBs. Default values are not defined in JunOS and SR-OS, because they do not perform SRGB label preservation like Cisco software. Label preservation was described in an earlier sub-section. From the table below, various SRGBs can be defined. Although the CBR-8 does not support Segment Routing currently, the latest IOS-XE version supports a SRGB size up to 64K indexes. This may change when the CBR-8 platform supports SR-MPLS, but it will be used as the lowest common denominator of all operating systems.

---

19 Although it is technically possible to use different SRGBs within a domain/IGP area, this is not recommended. Thus, all node's SRGB would have to be modified.

Table 3 - SRGB parameters summarization

Parameter	Platform A	Platform B	Platform C	Platform D	Platform E
Min (default)	none	16000	16000	16000	none
Max (default)	none	23999	23999	23999	none
Min	16	16000	16000	16000	18432
Max	1048575	1048575	1048575	1048575	1048575
Total Limit	1000000	65536	262144	65536	131072

The table below summarizes node counts in all domains. Domain owners provided estimates of future node counts. The values will be used to determine the requirements for the SRGB size based on node counts only. The table does not consider features described in a previous sub-section that require additional indexes from the SRGB.

Table 4 - Node requirements. Largest IS-IS L1 areas.

Domain	LSR Count (2018)[20]	LSR Count (2028)	Major Impacts
Backbone	< 110	< 120	None
Residential[21]	< 650	< 1,000	vCCAP buildout Data Center buildout
Data Center[22]	< 40	< 1,000	More virtual routers
Business Core	< 520	< 1,100	Additional S-PEs in each market
Business Access[23]	< 250	< 500	Possible standard change

Based on the estimated node counts, it is apparent that a small SRGB suffices. However, there are always challenges with forecasting node counts due to new products and services. For instance, the access engineering team has not decided on the ratio of virtual CAPs to RPDs. Because the CAP will be virtualized as a container or virtual machine, and automation and orchestration will be used to simplify deployments, it is possible that the ratio could be 1:1 to simplify the automation architecture. If this is the case, the number of SR-MPLS enabled nodes would increase dramatically in the largest market – estimate at around 50,000 in 10 years. Although an option, it is unlikely to be adopted since it would introduce FIB scaling and address allocation problems. Nonetheless, it emphasizes the difficulties of forecasting unknowns. The table below illustrates the forecasting of RPDs in the Phoenix market.

The table below describes a SRGB with 64K indexes. 64K is the recommended SRGB size, and it will be used in the examples to describe partitioning. 20000-83999 is the recommended SRGB range, given the available minimum and maximum limits on the different platforms. Label preservation is lost, but the solution scales. When the SRGB is configured, the operating system adjusts all label blocks. All four network operating systems support 1 million labels. Thus, less than 7% of all total label will be reserved for the SRGB. Vendors have noted that reserving 64K indexes should not cause problems, especially when SRGB size limitations are imposed[24]. However, label usage should always be tested.

[20] All data is collected
[21] Nodes in the markets owned by the data center team are classified as residential.
[22] Duke and Deer Valley
[23] Current standard as of 2018 is less than 250 nodes in an IS-IS L1 area or OSPF non-area zero
[24] JunOS does not limit the SRGB size.

Table 5 - 64K indexes

SRGB Size	+ / -	Description
64K	+	Size currently supported on all platforms
	+	Administered by each autonomous-system owner
	+	Features, immediate and future, supported
	+	Node growth supported
	+	Flexible index partitioning – features, reserved, etc.
	+/-	Utilizes <7% of total labels in the LSD
	-	Label preservation

The table below describes a SRGB with 32K indexes. This solution also sacrifices label preservation scale. Less than 4% of all total label will be reserved for the SRGB. This is less of a concern than 64K labels but should still be tested.

Table 6 - 32K indexes

SRGB Size	+ / -	Description
32K	+	Size currently supported on all platforms
	+	Administered by each autonomous-system owner
	+	Features, immediate and future, supported
	+	Node growth supported
	+	Flexible index partitioning – features, reserved, etc.
	+	Utilizes <4% of total labels in the LSD
	-	Label preservation

As described earlier, 8K index trades simplicity for scaling. It is ideal for the smallest deployments of SR-MPLS on Cisco software. The initial deployment of SR-MPLS in the residential domain is on less than 20 nodes in the largest market, so the initial deployment scope is relatively small. Thus, this SRGB can be eliminated.

Table 7 - 8K indexes

SRGB Size	+ / -	Description
8K (eliminated)	+	Size currently supported on all platforms
	+	Administered by each autonomous-system owner
	+	Features, immediate and future, supported
	-	Node growth supported
	-	Flexible index partitioning – features, reserved, etc.
	+	Utilizes <1% of total labels in the LSD
	+	Label preservation

4.2.5. Derivation

Index values must be assigned to each SR-MPLS node for them to generate node-SIDs. The method for deriving index values per node must be defined. There were two high-level options for this:

1. Based on static information residing on the node e.g. IP address

2. Partitioning indexes and tracking assignments using a database

The first option *appears* to be the most ideal. There is no point in maintaining an additional database if the index value can be derived from some value that is unique to the node. Using the IP address, an index value can be constructed so long as the index value can be large enough to support represented all or at least some of the octets. As stated in an earlier sub-section, lowest common denominator of currently available SRGB sizes is 64K. A future version of IOS-XE, may allow for 256K. With 256,000 index values, the last two octets of an IP address could be represented. This is illustrated in the figures below.

Loopback Example: 172.[Market].[0-255].[0-255]

Loopback Range – Low: 172.29.0.1

Loopback Range – High: 172.29.255.255

Index Values Required: 1 - 255,255

Example #1	Example #2
SRGB Base: 16,000 – 271,255	SRGB Base: 16,000 – 271,255
Loopback: 172.29.193.119	Loopback: 172.29.210.24
Index: 193,119	Index: 210,024
Label: 209,119	Label: 226,024

Figure 6 - Deriving indexes via IP addresses with an SRGB size of 256K

There are many challenges associated with this method. For instance, markets may not always use the same first two octets on all loopbacks. Phoenix uses 10.119.x.x, 10.120.x.x, 10.122.x.x, 172.29.x.x, 172.16.x.x, etc. addresses. Thus, it is possible that two nodes have the same last two octets, and this method begins to breakdown[25]. Another challenge is wasted indexes. No market should ever need 256K indexes, however, that is the minimum number required to derive two octets. Another downside is that there is no room for partitioning. Global indexes are used for various features – tactical reuse, IPv4, IPv6, SRMS, Anycast, Flex-Algo, etc. This method does not allow for partitioning by ownership either.

On the other hand, leveraging a database so that indexes can be partitioned providing structure based on function and group is most ideal. As mentioned above, it allows for clear delineation of ownership. For example, the data center and residential teams can each be allocated their own index blocks. Each team can further partition it, if they choose. For instance, indexes 1-99 can be allocated for nodes at the top tier of the hierarchy, and indexes 200-299 can be allocated for the next tier, etc. [26] As a recommendation, domain owners reserve a sub-block for any future technologies they may be considering that were not

[25] Using this method, no market node could share the same last two octets as a backbone node either, because backbone node indexes must be unique within the AS.
[26] A partitioning structure recommendation is provided in the next section.

addressed in this document. This method of index assignment also uses a smaller SRGB, because the indexes are used efficiently. This reduces the likelihood of label depletion. Database tracking, which is admittedly cumbersome in some ways, is required. This creates additional management overhead. However, it is still the recommended method to assign indexes.

4.2.6. Partitioning

This document has probably blurred the lines of domain ownership. It has, at a high-level, assessed all domains in the Cox network at one point or another, but it has tried to avoid specific statements about the business and data center domains. It has assessed features – whether needs are immediate, future or uncertain – for all domains e.g. Anycast, IPv6, Multicast, Flex-Algo, end-to-end LSPs, etc. It has addressed platform support, label depth, obsolesce, forecasting, etc. These factors were used to determine the SRGB size. Once that has been determined, along with the method to assign indexes, domain owners can plan their blocks. The tables in this section are the recommendations for label partitioning.

The first step is to get holistic view of the MPLS enabled routers in ASN22773 and to begin the initial phase of partitioning the SRGB indexes. The first phase partitions the block by domain ownership. The two tables below describes the total number of MPLS enabled nodes in each domain, as well as the percentage of nodes per domain. A domain is then allocated X indexes which is relative to the percentage of nodes in the domain. An additional column adds growth to the relative percentages. The actual index block assignments are also described. These numbers are based on a SRGB of 64,000. It is recommended that the data center and residential domain owners carves their index blocks by geolocation, mapping server and reuse.

Table 8 - MPLS Nodes and Index Requirements by Domain

Devices	MPLS Node Count (Q1 '19)	Percentage of Total MPLS Nodes	# of Indexes Relative to %	# of Indexes Relative to % + Growth
Backbone	106	5%	3,200	5,000
Residential	2000	93%	59,520	54,000
Data Center	42	2%	1,280	5,000

5. Interworking

5.1. Overview

Segment Routing provides two deployment models for integrating SR into a LDP-based network. One of the models, “ships in the night”, describes dual-stacking LDP and SR. This is the simplest implementation but is only leveraged when all nodes support SR. The other model, “interworking”, allows LSPs between LDP-only and SR-only nodes (and vice versa); Cox requires interworking. In this model, functions are required at both control-plane and data-plane layers. The mapping server (SRMS) facilitates control-plane learning. Certain nodes are required for perform SR-to-LDP and LDP-to-SR label swaps in the data-plane. The following sub-sections describe interworking as well as the challenges within Cox.

“Ships in the night”, or dual-stacking, is the simpler deployment model and will be briefly described here for completeness. There are four stages when transitioning.

1. LDP only
2. Dual stack
3. Prefer SR

4. SR only

The starting state has only LDP configured on all nodes. The ending state has only SR configured on all nodes. The second state dual stacks SR on LDP. Thus, the control-plane for SR is synchronized via the IGP. However, the FIB, or data-plane, is not programmed to push, pop or swap SR labels[27] until the third stage. In the third stage, SR is configured as the preferred protocol. On Juniper platforms, this is done by manipulating the route-preference. On Cisco platforms, "sr-prefer" is a macro which makes SR preferred over LDP. When a change is made on a node and because LSPs are unidirectional, only the LSP(s) for which that node is the ingress LSR will be impacted. This must be done on a node-by-node basis, until all nodes have transitioned to prefer SR. In the final stage, LDP is removed and the domain becomes SR only. The stages are illustrated below.

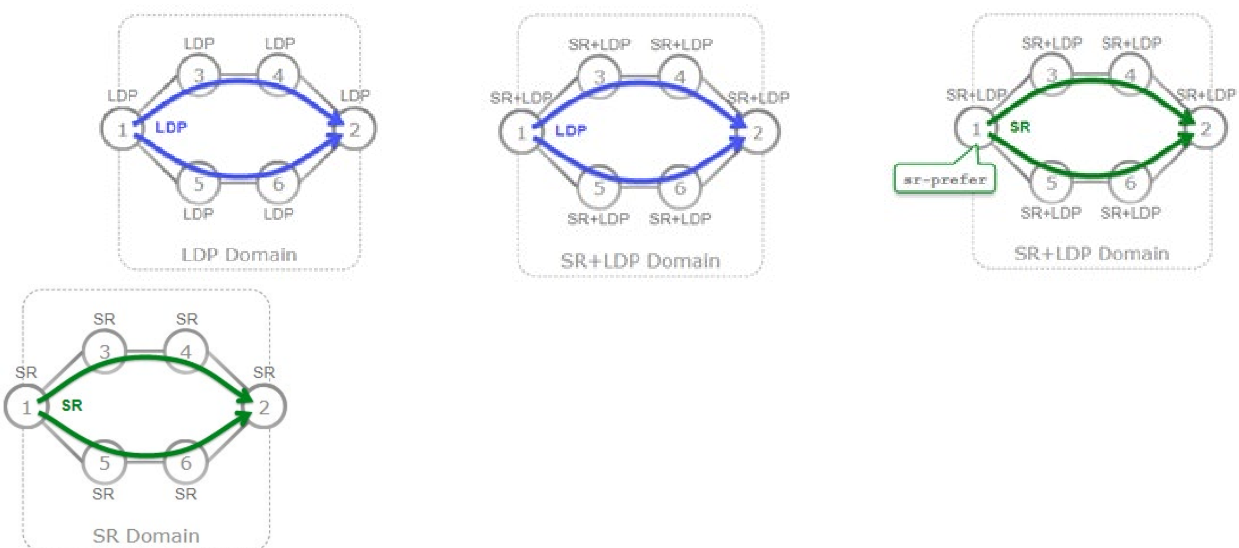

Figure 7 - Ships in the night

Interworking expands on ships in the night. The difference between ships in the night and interworking is that interworking accounts for LDP remaining in the network. Terminology has been created to define new nodal functionalities introduced in this deployment model. In the end state, the SR deployment can be classified into two sections: "the core" and "edges"[28] . Unique nodal functions occur in both sections. "SR LSRs" perform label swaps using SR labels. "Interworking LSRs" swap SR-to-LDP and LDP-to-SR labels. SR LSRs runs SR only, and an interworking LSRs run both SR and LDP. However, the interworking LSR does not have dual stacked interfaces. The "edge" is where "LDP LSRs" resides. These nodes only run LDP. This is illustrated in the figure below.

[27] These are traditional MPLS terms, and they have been redefined in the SR framework.
[28] This is not an industry or technology term. It has not been defined and is made up here.

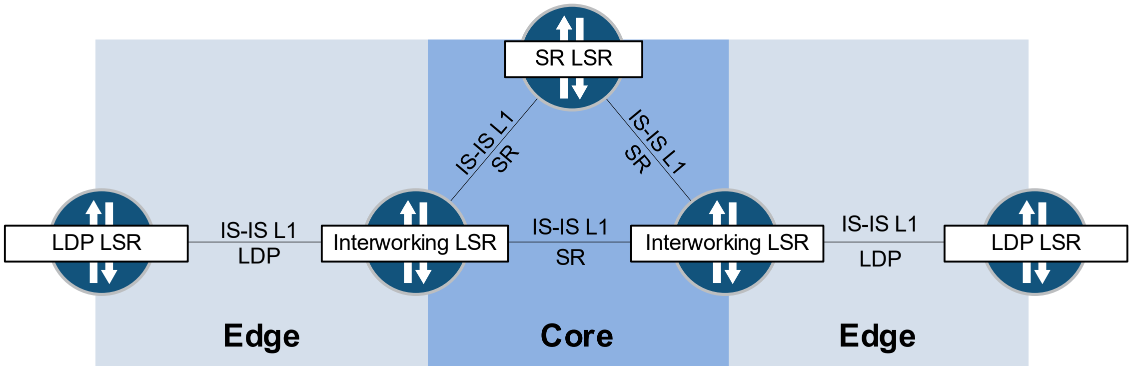

Figure 8 - Interworking terminology in the SR IGP area

It's important to keep in mind traffic flows and LSP termination points i.e. the ingress and egress LSRs of an LSP. Traffic flows and LSPs determine which node types the LSP must traverse. An LSP could terminate on any of the nodal types defined in the interworking model. For instance, a service is defined on two LDP LSRs. Thus, an LSP must exist. If the LDP LSRs are in different edges, the LSP must be programming in at the LDP LSRs and Interworking LSRs FIBs, using the diagram above. This requires swapping LDP-to-SR and SR-to-LDP at the interworking LSRs. However, if the service is defined on two LDP LSRs in the *same* edge, LDP only swaps occur. These scenarios are illustrated in the figure below.

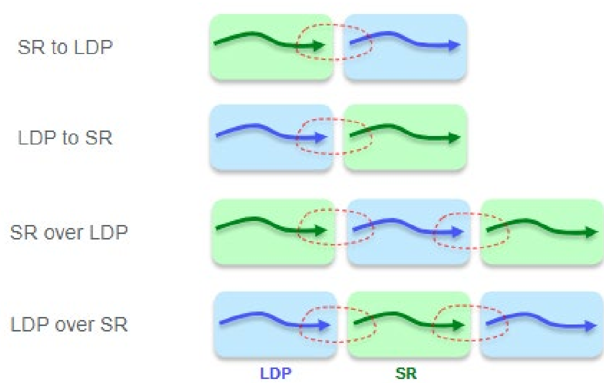

Figure 9 - Label swapping

To understand the type of label operation being performed, one must identify the LSP's termination points. Knowing the termination points also allows determining whether a mapping server is required for the LSP. A matrix has been created to reflect the residential domain and identify when the mapping server is required.

5.2. Mapping Server

A mapping server is configured with individual or blocks of IP addresses and indexes to be assigned to LDP LSRs. The mapping server will identify LDP LSRs via its IGP database. It then allocates and advertises, via its own LSP, an index on the nodes' behalf using a new TLV. From the perspective of a SR LSR, an index assignment exists for every node, including LDP LSRs. Thus, all nodes appear SR enabled.

This allows SR LSRs to push SR labels for services residing LDP LSRs. Once the mapping server allocates indexes on behalf of LDP nodes, the interworking LSR can program the FIB for label swapping between protocols. . This process is out of the scope of this document and should be understand via RFCs, technical documentation, training and testing.

At least two mapping servers should be defined within an IS-IS area. It functions like a BGP route-reflector; it is not required to be in the data-plane and is purely a control-plane function. The mapping server can reside on any node within the IGP area. Keep in mind that it will serve all nodes in the IGP area; this includes nodes owned by the data center team. Thus, the location of the mapping server should account for index block planning, because it may serve data center nodes[29]. The selection of mapping server placement is highly dependent upon operational perspectives no recommendation will be made on mapping server placement.

The mapping server is responsible for advertising prefix and index associations of LDP LSRs. It is recommended to dedicate a range for mapping server allocations. A recommended SRGB plan was defined in the Global Block Planning section. Currently, there are two methods to perform label allocation:

1. **1-to-1** using individual prefixes and individual indexes
2. **Many-to-many** using prefix and indexes ranges

The mapping server allocation method impacts operations and provisioning teams. Typically, new LSR deployments in the residential domain are irregular, compared to the business domain. In the business domain, or a domain were new LSR deployments are more frequent, this brings more value than in the residential domain. In today's state, simplifying provisioning does not outweigh the increase in operational complexity. Also, most mapping server entry challenges will be on "day one" deployments – not "day two" provisioning challenges. It is recommended that 1-to-1 be deployed to reduce the complexity of the interworking deployment. Each time a LDP LSR is added to the network, the mapping server must be updated. This assumes that the 1-to-1 entry type is being used. Operations and provisioning teams should be aware of this new, additional step in the deployment process.

With the mapping server, an error can occur when two mapping servers advertise different index bindings. This is likely an operational issue, rather than by design. In this case, a preference selection can occur, so that all nodes within the IGP area have selected the same mapping server for bindings. However, this was not the case in the initial phase of testing in the lab. In fact, this function is being defined within the IETF and is still in the "draft" state[30]. However, vendors often implement features before the IETF standardizes it. The current mapping entry preference is summarized below. However, this is subject to change.

6. Implementation

6.1. Prerequisites

After the base configuration is applied to all Vendor A and Vendor B platforms, it is recommended a tool be used to aide in verification as well as to generate and validate mapping server entries. This tool should be provided to the field along with the LLD. It should leverage NETCONF transport and YANG data models to collect IS-IS and LDP database information. Collection should occur via two nodes for

[29] If the data center receives their own index block from the residential SRGB, that team may not want the allocations performed on the NCS.
[30] https://datatracker.ietf.org/doc/draft-ietf-spring-conflict-resolution/

redundancy. By cross referencing both IS-IS and LDP data collection, a node can quickly be classified as one of the following:

1. **SR LSR** – index in IS-IS database
2. **LDP LSR** – entry in LDP database
3. **Interworking LSR** – index and entry in both databases
4. **IP-only** – no index or entry in either databases

All nodes generate and advertise an IS-IS LSP. These contain hostnames, prefixes and metrics. Nodes enabled with SR also advertise their index and SRGB assignments. Mapping servers also advertise the entries used for interworking. All information should be made available to an operator via a GUI. This allows them to quickly parse the SR information of any node which is not an easy task otherwise. This also allows an operator to easily identify LDP LSRs and generate mapping server configurations. Using this same information, two automation tasks can be performed. These tasks require cross examining LDP LSRs against the mapping server's advertisement. By cross examining these, it can be determined which nodes require an entry. Thus, mapping server configurations can automatically be generated. This is ideal for "day one" deployments where markets have 400+ LDP LSRs. The tool can also routinely validate the mapping server entries and generate reports when there is a LDP LSR without a mapping server entry. This is ideal for "day two". If a tool is constructed with these features, it will greatly reduce the workload of operators.

6.2. Current

This section addresses the high-level implementation phases. It also defines specific challenges and concerns. The current network state is illustrated below. All nodes are LDP LSRs. Only MPLS-enabled nodes are shown; IP nodes are not impact by SR-MPLS. Topologies and hostnames vary market to market.

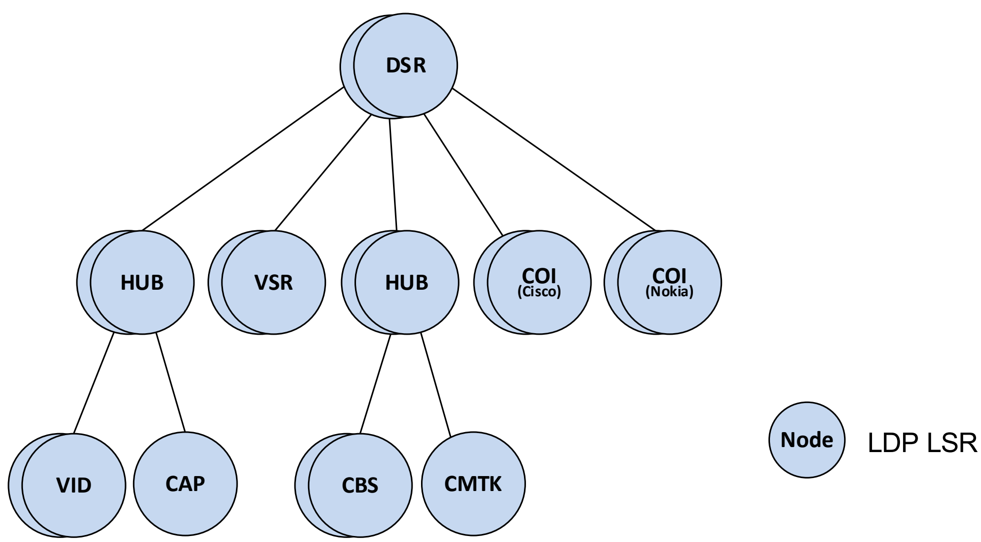

Figure 10 - Sample topology - current state

6.3. Dual Stack

After upgrading both MX and NCS platforms to the preferred software versions for SR-MPLS, the base SR configuration can be applied. This should be non-service impacting[31]. Like administrative distance (Cisco) and route preference (Juniper), MPLS transport protocol preferences exist. This allows an operator to toggle between different MPLS transport protocols, if label switched paths exist for both. By default, LDP is preferred over SR-MPLS in both platforms[32].

SR-MPLS and LDP are dual stacked with LDP remaining the preferred MPLS transport protocol. SR LSRs will program an additional LFIB entry for other SR LSRs. Thus, two entries for an SR LSR will exists; an entry to reach the SR LSR via LDP and another entry via SR. In addition to this, link protection also requires additional hardware resources to program backup routes for FRR capabilities. It is recommended to test the change in FIB/LFIB utilization caused by dual stacking and TI-LFA.

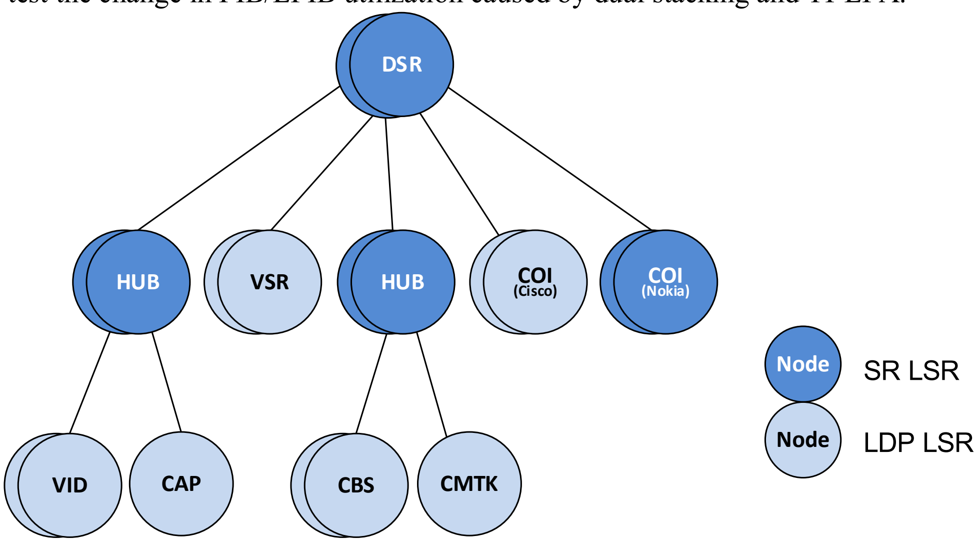

Figure 11 - Sample topology - Dual Stack

6.4. Mapping Server

This task, and the following, will leverage the tool described in the earlier sub-section. This should be non-service impacting. Since the mapping servers' location is currently undecided, the illustration uses the DSRs for explanatory purposes. As noted earlier, the mapping server is control-plane only. Thus, the functionality is the same no matter where implemented in the IGP area. After implementing the mapping server, every SR LSR should have an index for all MPLS nodes. SR LSRs will program an additional LFIB entry for LDP LSRs. Thus, the LFIB utilization on the NCS and MX will increase. The workload in this phase is greatly reduce if the SR tool mentioned in the previous section is available. It can be used for an immediate validation of mapping server entries and determining if entries are missing.

[31] Only nodes dual stacking will be impacted – MX/NCS. This most likely reason for service impacts, if any would be scale. Dual stacking on the MX and NCS should be scale tested prior to deployment
[32] This should be revalidated with the selected SR-MPLS codes.

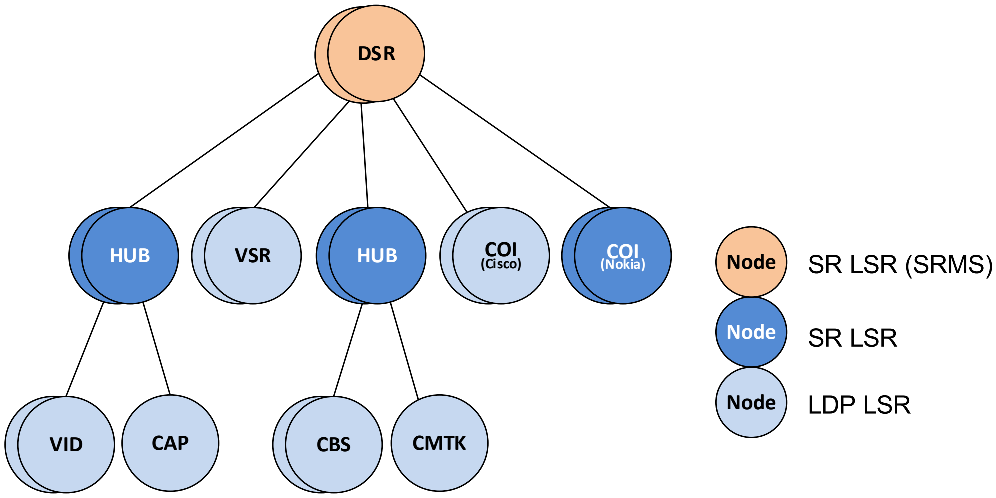

Figure 12 - Topology - mapping server

6.5. Protocol Preference

When an operator changes the label preference on the node, the LSR will reprogram its FIB to "push" SR labels instead of pushing LDP labels. Thus, LSPs for which the LSR is the headend are impacted. When a LSR terminates a service and has the preference changed, a service interruption can occur. This is scenario is applicable to the MX. This may apply to the Nokia COI router, depending on its deployment. Since LSPs are unidirectional, the remote LSR's LSP is not impacted by the change. The operation at the headend LSR is summarized in the following points:

- Push LDP, by default
- Push SR, if configured[33]

When the preference is modified on a the NCS, a "P" LSR, transit LSPs should not be impacted. Changing the preference on a P LSR does not cause it to perform the interworking functionality. In other words, it will not change its swap operation from a LDP-to-LDP swap to a LDP-to-SR swap. It is recommended to validate that the preference does not delegate interworking for transit LSPs using the software versions selected for deployment. This being the case, modifying the preference is less likely to impact services than the next stage. The operation at the transit LSR is summarized in the following points:

- Preference does not influence the swap operation
- Swap to LDP, when a SR LSP exists but leveraging a mapping server[34]
- Swap to SR, when a SR LSP exists without leveraging a mapping server

An illustration of how the protocol preference at the transit LSR and SRMS does not impact the LSP. The MX and NCS are dual stacked with LDP and SR. The DSR is performing the mapping server function. Thus, both the MX and NCS having SR labels for both CAP and COI. A service terminates on the CAP and COI. The preference is modified on the MX and NCS. However, if the transport layer does not change, the service is not impacted. The NCS continues to swap LDP labels because of the rules stated above.

---

[33] Whether the SR LSP is derived from mapping server or is end-to-end is irrelevant.
[34] The MPLS protocol preference does not impact this decision, per code testing.

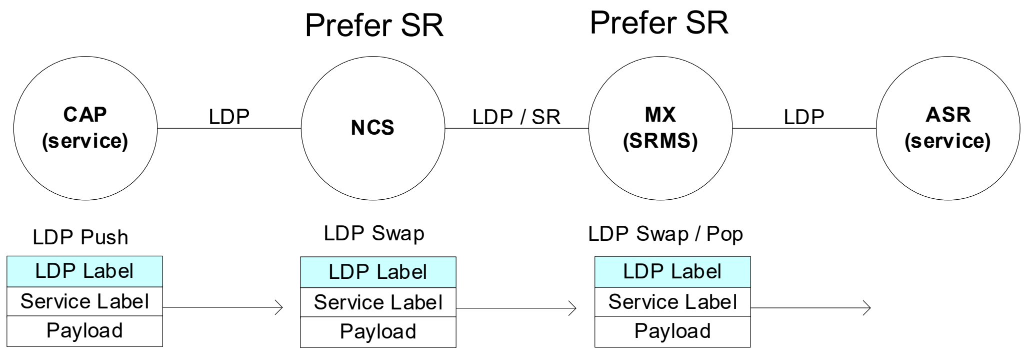

Figure 13 – SR preference should not impact transit LSPs

The example topology that has been referenced in the previous sub-sections is supplied below. The MX and NCS are configured to prefer SR over LDP.

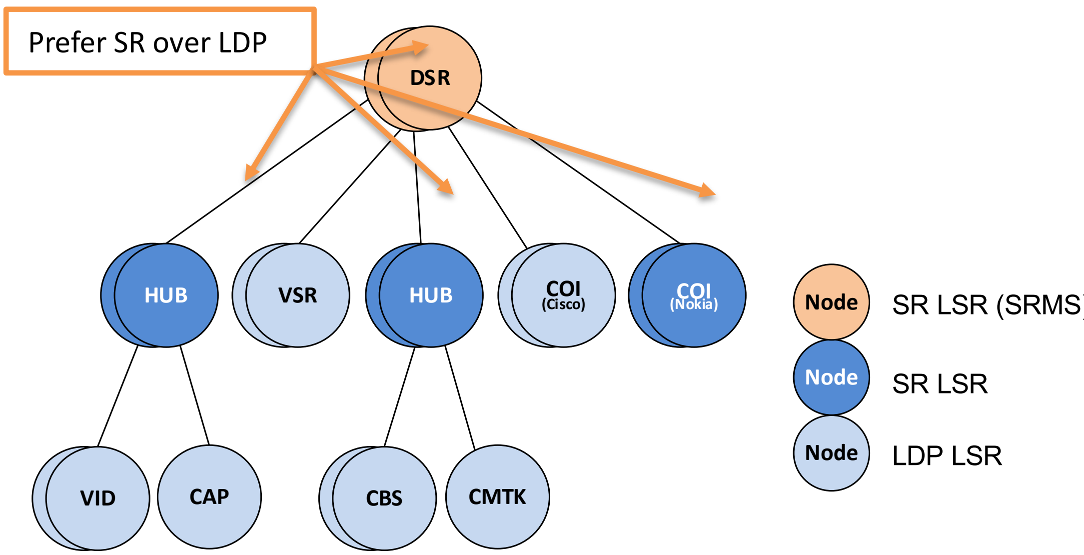

Figure 14 – Same topology – Protocol Preference

6.6. Interworking

At this point, all MX and NCS nodes should prefer SR as the MPLS transport protocol. Every MPLS node requires an index. SR LSRs have the index defined in their configuration. LDP LSRs are assigned indexes via the mapping servers. The final stage of implementation is disabling LDP and relying on interworking. In other words, LDP-to-SR and SR-to-LDP swaps.

When a site is selected to transition to Segment Routing, LDP should be removed on the NCS uplinks and interconnects. LDP-IGP synchronization should also be removed from those links or the IGP will advertise them with "max-metrics" as designed. LDP should remain on links to CBR-8, uBR, ASR, etc. platforms. Once the configurations have been committed, the NCS will perform the interworking function and forward the label switched traffic using the new label operation.

It is less likely that Internet and video services being impacted, because they are, or can be, transport via native IP. Thus, L2VPN and L3VPN customers, whether internal and external, should be aware of the potential for service impact. The most likely reason for impact to services is incorrect or incomplete mapping server configurations. After decommissioning LDP on the defined links at all hub-sites, the market will have a "SR core" and this project is complete.

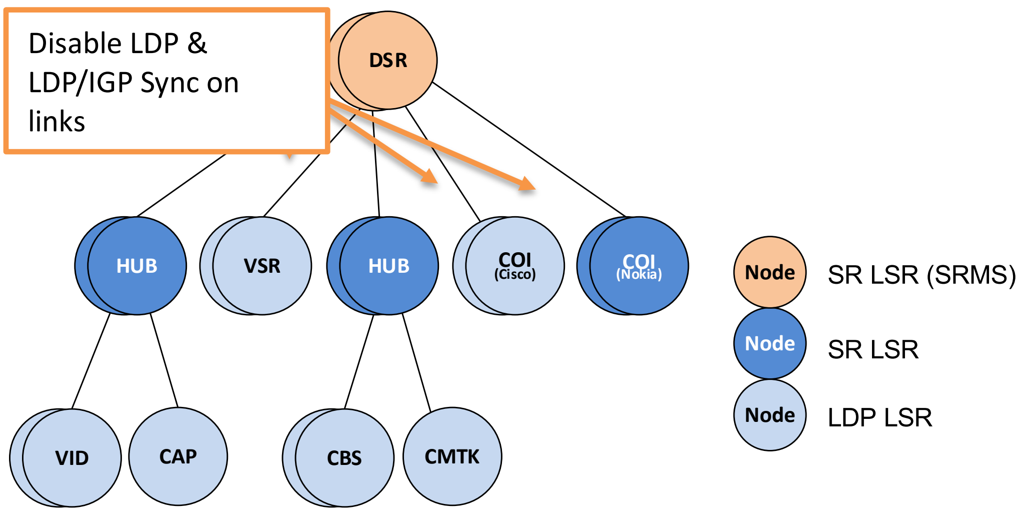

Figure 15 - Topology - Interworking

Abbreviations

Architectures	
SR-MPLS	A deployment of SR where MPLS is the data-plane and IPv4 is the control-plane
SR-MPLSv6	A deployment of SR where MPLS is the data-plane and IPv6 is the control-plane
SRv6	A deployment of SR where IPv6 is the data-plane and IPv6 is the control-plane
Unified SR	An architecture using BGP-LU to unify multiple SR domains
Unified MPLS	An architecture using BGP-LU to unify multiple MPLS domains
Distributed TE / SR-TE	A TE deployment not utilizing a SR-PCE, thus having policies defined on the node
Centralized TE / SR-TE	A TE deployment utilizing a PCE, thus having policies defined on the PCE
Terminology	
Backbone Domain	ASN22773: L1/L2 or L2-only nodes
Data Center Domain	ASN22773: L1 nodes – Duke and Deer Valley
Residential Domain	ASN22773: L1 nodes – Phoenix, Las Vegas, Hampton Roads, Rhode Island, etc.
Business Domain	ASN15218: All nodes
Business Core	ASN15218: L1/L2 or L2-only nodes

Business Access	ASN15218: L1 nodes
Segment Routing (SR)	
Source Packet Routing in Networking (SPRING)	Juniper's name for Segment Routing
Segment ID (SID)	A MPLS label identifying a node, IGP adjacency, prefix, LSP, etc.
Segment List	The MPLS label stack derived via Segment Routing
Global Block (SRGB)	A label range for global SID assignments. E.G. Node-SID
Lobal Block (SRLB)	A label range for local SID assignments. E.G. Adjacency-SID
Mapping Server (SRMS)	Devices which assign SR labels for non-SR devices via IGP advertisements
Mapping Client (SRMC)	Devices which receive SR labels from the SRMS
Path Computation Element (PCE)	A controller for RSVP-TE, SR or SR's ODN
Path Computation Client (PCC)	A client of the controller
MPLS-TE	MPLS protocols which provide traffic engineering capabilities such as SR and RSVP-TE
TE / SR-TE	Non-default routing – may be defined on either a PCE or directly on a node via CLI
On-Demand Next-Hop (ODN)	A model used to compute end-to-end LSPs across multiple domains via a PCE