Deploying Segment Routing for PON Aggregation in Cox's Metro Network

A Technical Paper prepared for SCTE by

Deependra Malla
Lead Network Design Engineer
Cox Communication Inc.
6305 Peachtree Dunwoody Road, Atlanta, GA 30328
602-694-4429
deependra.malla@cox.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

1. Introduction

The continuous advancement of IP networks marked by increased capacity requirements, increased network assets, and elevated customer expectations on service continuity and quality of service, has placed substantial demands on the network infrastructure of Multi-Service Operators (MSOs), Internet Service Providers (ISPs), and large enterprises. To help meet these demands and improve network performance, network operators have started adopting Segment Routing (SR) in the various parts of their networks - access, metro, and backbone.

The evolution of access technology from cable Internet to Passive Optical Network (PON) involves a transition in the technology and architecture used to deliver high-speed broadband services to end-users. Both cable Internet and PON are broadband access technologies, but they differ significantly in their underlying infrastructure and methods of data delivery. The primary concept behind PON is to bring fiber optic cables directly to the end-users, eliminating the need for intermediate coaxial cables or copper lines. This is often referred to as Fiber-to-the-Home (FTTH) or Fiber-to-the-Premises (FTTP). Although cable access is Cox's legacy access technology, it introduced FTTH to its customers in 2014 under the brand name "Gigablast". Since then, Cox has continuously invested in network infrastructure and improving the quality of Gigablast services offered to its customers. Adoption of Segment Routing in the Cox's PON aggregation and metro network is a crucial step towards making Cox's access network more optimized, scalable, and resilient.

As the networking landscape evolves, the industry is leaning towards more flexible and software-driven networking paradigms. So, transitioning to Segment Routing positions Cox metro network to embrace these future changes and innovations. Adopting the SR also helps Cox to provide improved network performance and service continuity to its customers.

The Cox metro networks is an MPLS enabled network and runs LDP as the label switching protocol. LDP is used for following purposes in Cox metro networks:

- Label switch BGP traffic,
- Transport label for L2VPN, and
- Transport label for L3VPN

In the Cox metro network, Segment Routing is used to distribute MPLS labels and accomplish all the above services. Following are the motivations for enabling SR in Cox metro networks:

- Protocol simplification and unified control plane
- Achieving <50ms failover during link failure
- Tactical traffic engineering capability
- New protocol knowledge acquisition

Since Segment Routing is a more recent approach to label switch traffic, it is not available in legacy hardware. Hence, during the initial phase of SR deployment in Cox, LDP and SR will co-exist. Wherever the SR is enabled along with the LDP, the SR is made a preferred label switching protocol.

2. Cox Metro Network Overview

PON network in Cox metro is a multi-vendor hub-and-spoke topology with hub routers aggregating multiple pairs of PON aggregation routers. These PON aggregation routers aggregate multiple Optical Line Terminals (OLTs). The Cox metro network runs IS-IS as an IGP protocol, LDP as a label distribution protocol and hierarchical BGP. The high-level topology of Cox metro network is shown in figure 1.

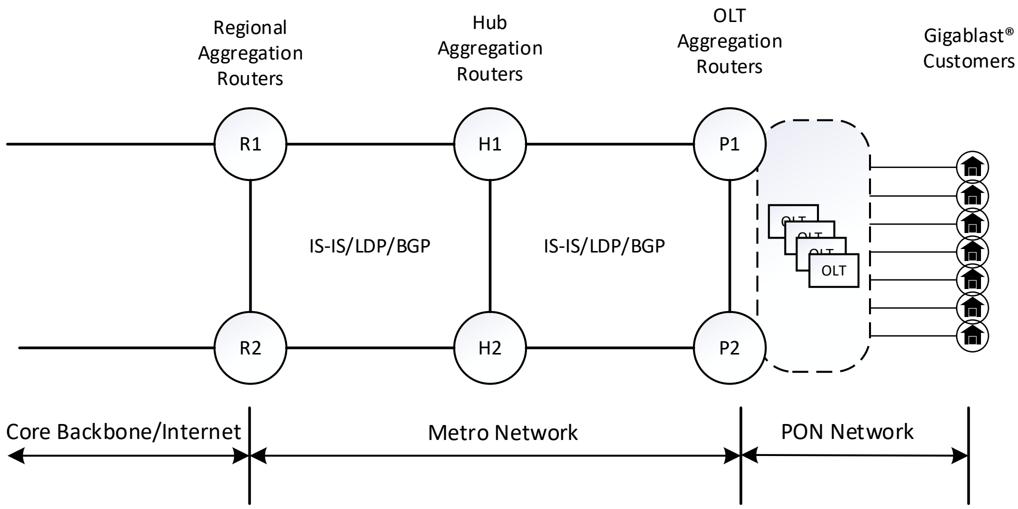

Figure 1 – Cox Communication High Level Metro Network Architecture

Regional aggregation routers are the gateway to the Internet and other Cox markets, and they aggregate several hub site's hub aggregation routers. Hub aggregation routers further aggregates several edge routers such as PON routers, CIN routers, local service/management routers, CMTS/CCAP etc. These hub aggregation routers function as Provider (P) routers in the Cox metro MPLS domain with no services provisioned on them. Regional aggregation routers and edge routers function as Provider Edge (PE) routers. The Internet traffic to and from OLT aggregation routers to regional aggregation routers are LDP label switched.

3. Label Distribution Protocol (LDP) Limitations

The Cox metro network is an MPLS enabled network with LDP used to distribute labels for loopback interfaces of metro routers. LDP is a traditional method for assigning and distributing labels in an MPLS network. It relies on label distribution from the headend router to downstream routers, establishing label-switched paths (LSPs) along which packets are forwarded based on labels. Today, LDP generated labels are used to label switch customer Internet traffic in Cox metro network. It is also used to provide a transport label for services like L2VPN and L3VPN. Although, LDP has its own advantages such as its widespread adoption, simplicity of basic MPLS and predictable behavior, there are certain limitations too.

Disadvantages of LDP:

Label Distribution Protocol (LDP) is a widely used technology for label assignment and distribution in Multiprotocol Label Switching (MPLS) networks. However, like any technology, LDP has its disadvantages and limitations. At Cox, we have noticed following limitations of LDP:

- **LDP-IGP Synchronization Requirement**: LDP must be fully established and synchronized with the underlying interior gateway protocol (IGP) before an IGP path is used for forwarding traffic. If the LDP and IGP are not synchronized, packet loss can occur. During link or node failures, we must wait until LDP-IGP synchronizes before putting traffic into LDP LSP.
- **Limited Traffic Engineering:** LDP provides limited capabilities for fine-grained traffic engineering and path optimization. It lacks the flexibility to specify explicit paths and perform dynamic traffic steering based on real-time network conditions. This can be a limitation in networks that require efficient resource utilization and optimized routing.

- **Scalability Concerns:** As networks grow larger and more complex, LDP's signaling overhead can become a scalability concern. The protocol generates a significant amount of control plane traffic to establish label-switched paths (LSPs), which can lead to increased processing and memory requirements on routers.
- **Limited Support for Programmability**: LDP's lack of programmability and flexibility can hinder its compatibility with modern network automation and software-defined networking (SDN) principles.
- **Incompatibility with IPv6 Transition**: LDP does not inherently support IPv6 traffic over an IPv4 infrastructure, making the transition to IPv6 more challenging.

Segment Routing (SR) and Label Distribution Protocol (LDP) are both networking technologies used in the context of MPLS (Multiprotocol Label Switching) networks. Each technology has its own advantages and use cases. However, transitioning from Label Distribution Protocol (LDP) to Segment Routing (SR) in Cox network can have following added benefits:

- Ability to introduce traffic engineering and path optimization,
- Simplicity and reduced network states,
- Improve scalability,
- Flexibility and Programmability,
- Unified Control Plane
- Faster traffic re-routes during link and node failure,
- Integration with SDN and automation

4. Segment Routing

Segment Routing (SR), as defined in RFC 8402,(Clarence Filsfils et al., 2018) is a modern network architecture and forwarding paradigm that simplifies the way packets are routed through networks. It leverages source-routing principles and enables efficient traffic engineering, optimal path selection, and seamless integration with software-defined networking (SDN) principles. SR is particularly suited for Multiprotocol Label Switching (MPLS) and IPv6 networks. There are two types of Segment Routing technologies - SR with MPLS data plane (SR-MPLS) and SR with IPv6 data plane (SRv6). The current deployment of Segment Routing in Cox metro network uses SR-MPLS and is primarily targeted for label switching IPv4 traffic. This document will focus on SR-MPLS deployment in Cox metro network.

In SR, the concept of a "segment" is central. A segment can be thought of as an ordered list of instructions that guides the path a packet should take through the network. These segments are represented as a stack of instructions embedded in the packet header. The segments define the specific nodes and links that the packet should traverse, allowing for explicit path control and optimization.

4.1. SR-MPLS Operation

SR-MPLS, also known as Segment Routing with MPLS data plane, is an implementation of Segment Routing (SR) within Multiprotocol Label Switching (MPLS) networks. It combines the benefits of SR's explicit path control and MPLS's label-switching capabilities. SR-MPLS offers efficient traffic engineering, simplified network management, and seamless integration with existing MPLS infrastructure.

In SR-MPLS, segments are represented as labels in the MPLS label stack. Each label corresponds to a specific segment or node in the path. The stack of labels in the packet header is known as the "segment list." The top label indicates the next hop, and subsequent labels indicate the subsequent nodes or segments in the path. At the source node, the path is encoded into the packet header as a stack of labels (the segment list). The labels in the segment list represent the desired path that the packet should follow through the network. As the packet traverses the network, routers perform MPLS label swapping according to the label stack in the packet header. Each router examines the top label, determines the next hop or segment, and forwards the packet accordingly.

The SR MPLS data plane operations use the existing MPLS forwarding operation. The segment list operations (PUSH, CONTINUE, POP) are mapped to the MPLS data plane operations in Table 1.

Table 1 – Segment list mapping to MPLS label stack operation

SR Segment List Operation	MPLS Label Stack Operation
PUSH	PUSH
CONTINUE	SWAP
POP	POP

As described earlier, the concept of segment in central to SR architecture. A segment is an instruction that an ingress node inserts on the packet header and is identified by a Segment Identifier (SID). The segment can be a global segment known as Segment Routing Global Block (SRGB) or a local segment known as Segment Routing Local Block (SRLB). The implementation of SID in Cox network for IPv4 FEC is an MPLS label (32-bits). Although there are many types of SIDs, for Cox's SR deployment, IGP Node Segment (Node-SID) and IGP Adjacency Segment (Adj-SID) are discussed in this paper. Node-SID is globally significant and should be unique in a SR-MPLS domain whereas Adj-SID is locally significant to the router.

Although, SR architecture does not assume a specific data plane implementation, for SR deployment in Cox metro network we use existing MPLS data plane to forward SR label switched traffic.

4.2. Segement Routing Global Block (SRGB)

The SRGB is a range of labels reserved for Segment Routing global segments when using MPLS data plane for segment routing (Clarence Filsfils et al., 2017). SRGB needs to be configured on every router that participates in the segment routing domain. Varying size of label spaces on various vendors poses a significant challenge on designing of SRGB. If the SRGB space is allocated from the already used label space in a router, then the router needs hard reboot for that label to take effect. To minimize the number of router reboots in the Cox network, Cox has allocated 300,000 through 559,999 label space for SRGB. This range of SRGB will be used homogenously and contiguously across the entire Cox network.

4.3. SRGB Design and Management

For Cox deployment, SRGB is used for global labels like Node-SIDs. Node-SIDs are globally unique identifiers that require static assignments. To make the allocation of Node-SID unique and easy, each market in Cox network is allocated a block of SID indices from the global range. This range is used for Node-SID index for IPv4, IPv6 address families and any other SR algorithm Node-SID requirements. Cox is not using "absolute" SID assignment for Node-SID.

For Cox's deployment, SRGB parameters are defined as follows:

- SRGB Base value = 300,000
- SRGB Block Size = 260,000
- SRBG Global Range = 300,000 – 559,999 (Across all Cox domains)

Each router in SR domain requires a unique Node SID which is generated by using unique node SID index assigned to each router in each market. Formally, a Node SID is a unique index in the SRGB. Any node in the SR domain derives the local label associated to the Node-SID as SRGB + index.

Different vendors implement SRGB in different ways – global allocation and/or per protocol allocation. The SRGB design and implementation in a network depends on network operators design requirement, topology, and size of the network.

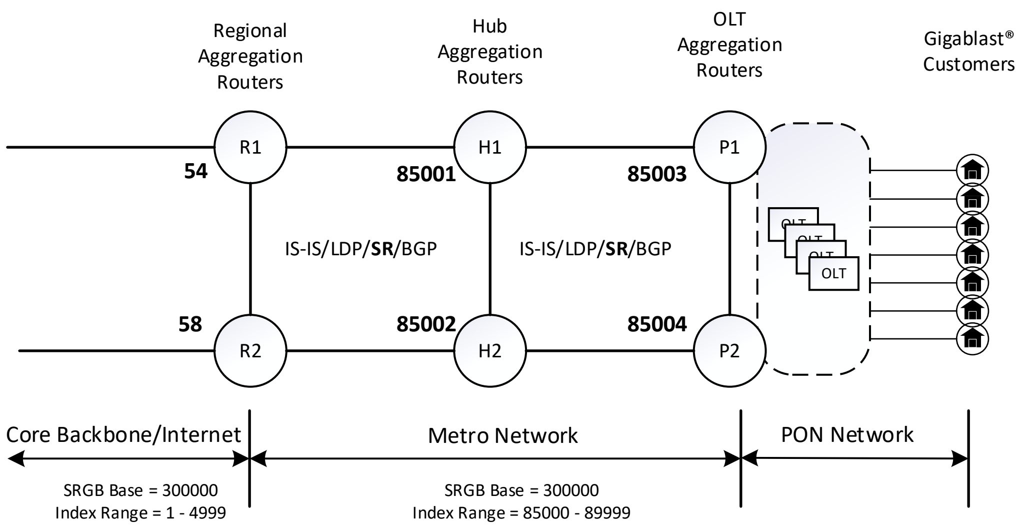

Figure 2 – Cox metro SRGB design and index assignment example

Example:

If the node-sid index for the regional aggregation router R1 in figure 1 is 54, then hub aggregation (H1, H2) and OLT aggregation routers (P1, P2) will calculate local label for R1 as follows:

Local label on P1 router for R1 = SRGB base on P1 + R1's node-sid index

= 300000 + 54

= 300054

The outgoing label is the local label that the downstream neighbor (the next-hop router on the shortest path to destination router) allocated for the Node-SID.

Outgoing label on P1 router for R1 = SRGB base advertised by H1 + R1's node-sid index

= 300000 + 54

= 300054

If other routers in the network are configured for same SRBG base value, all remaining routers in figure 1 topology will have label 300054 programmed in their LSD (Label Switching Database) for R1.

To avoid duplication of SID index, Cox will either create user defined field (UDF) in IPAM or create a database from where SID index can be assigned and stored for each router in the Cox network.

4.4. SR-MPLS Data Plane Operation

The traditional MPLS data plane is reused without any modification for Segment Routing in Cox metro network. Label operations (Push, Swap/Continue, Pop), PHP behavior, MPLS TTL processing, EXP bits processing, ECMP load-balancing techniques and MTU (Maximum Transmit Unit) handling are used similarly for Segment Routing like its being used for LDP today.

4.5. Segment Routing IS-IS Control Plane

Link state routing protocols like IS-IS and OSPF (Open Shortest Path First) have been extended to support distribution of SR information in IGP domain. RFC 8667 added several new sub-TLVs to IS-IS to signal various SR capabilities (Clarence Filsfils et al., 2019). The SR information includes Prefix-SID index, Adjacency-SID, SR data plane capabilities and label range used by SR. These new sub-TLVs are used to attach Prefix-SIDs and Adj-SIDs to various prefix and adjacency advertisement TLVs. Following are some useful TLV/sub-TLVs used by IS-IS:

- TLV = 22, Extended IS reachability
  - Sub-TLV = 31, Adjacency Segment Identifier
- TLV = 135, Extended IP reachability
  - Sub-TLV = 3, Prefix Segment Identifier
- TLV = 242, Router capability
  - Sub-TLV = 2, SR capability

The following example (output from Cisco IOS-XR) shows the IS-IS LSP flooded by R1 in figure 2 into the network. The example shows that both IPv4 and IPv6 address families are enabled in multi-topology mode. We can see several valuable information regarding segment routing information being advertised into the LSP such as:

- Router capability TLV (`Router Cap:      10.0.0.1, D:0, S:0`)
- IPv4 Node-SID index, default algo, and flags (`Prefix-SID Index: 54, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0`)
- IPv6 Node-SID index, default algo, and flags (`Prefix-SID Index: 654, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0`)
- Various Adj-SID and flags (`ADJ-SID: F:0 B:0 V:1 L:1 S:0 weight:0 Adjacency-sid:16`)
- SRGB Range (`SRGB Base: 300000 Range: 260000`)

This information is useful in validating the design and troubleshooting of SR if there are any issues.

```
RP/0/0/CPU0:H1#sh isis database R1.00-00 verbose
Wed Aug 16 19:18:43.915 UTC

IS-IS ISIS (Level-1) Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime  ATT/P/OL
R1.00-00              0x00000e37   0x1f53        753             0/0/0
  Area Address:   49.0090
  TLV 14:         Length: 2
  NLPID:          0xcc
  NLPID:          0x8e
  MT:             Standard (IPv4 Unicast)
  MT:             IPv6 Unicast                                 0/0/0
  Hostname:       R1
  Metric: 40         IS-Extended R2.00
    Interface IP Address: 192.168.10.1
    Neighbor IP Address: 192.168.10.2
    Local Interface ID: 330, Remote Interface ID: 330
    ADJ-SID: F:0 B:0 V:1 L:1 S:0 weight:0 Adjacency-sid:16
```



```
:
Metric: 40         MT (IPv6 Unicast) IS-Extended R2.00
  ADJ-SID: F:1 B:0 V:1 L:1 S:0 weight:0 Adjacency-sid:17
:
Metric: 0          IP-Extended 10.0.0.1/32
  Prefix-SID Index: 54, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0
:
Metric: 0          MT (IPv6 Unicast) IPv6 2001:BEEF:1:0:10::1/128
  Prefix-SID Index: 654, Algorithm:0, R:0 N:1 P:0 E:0 V:0 L:0
Router Cap:     10.0.0.1, D:0, S:0
  Segment Routing: I:1 V:1, SRGB Base: 300000 Range: 260000
  SubTLV 19 Length: 1
```

Maximum Segment Depth (MSD)

In SR, maximum segment depth refers to the number of segments (labels in SR-MPLS) a router is capable of imposing on a packet and varies across the vendors and the platforms (software & hardware). SR will increase the depth of label stacks if SR features such as TI-LFA, SR-TE, and traffic accounting are enabled on the routers. The topology and metrices of a network determine the segment depth of the labeled packets. Since Cox metro topology is a hub-and-spoke topology with static metric, the number of SIDs required for TI-LFA link-protection is 2.

5. Topology Independent Loop Free Alternate (TI-LFA)

Cox metro network provides critical services to customers such as Metro Ethernet over HFC (Hybrid Fiber Coaxial), voice and streaming services apart from Internet traffic. These critical services are the driving factors for enablement of fast reroute in metro network. In SR enabled network, TI-LFA provides the desired protection mechanism to such critical traffic. Topology Independent Loop-Free Alternate (TI-LFA) is a fast reroute (FRR) mechanism used in networks to provide protection against link and node failures. It is specifically designed to work in IP networks, including those running Segment Routing (SR), and aims to quickly restore connectivity in case of failures while avoiding the creation of forwarding loops. TI-LFA enhances network resilience and ensures that traffic continues to flow smoothly even during network disruptions. There are three types of FRR – Classic Loop Free Alternate (cLFA), Remote Loop Free Alternate (rLFA) and Topology Independent Loop Free Alternate (TI-LFA). cLFA and rLFA do not provide 100% coverage and TI-LFA does. So, at Cox we have deployed TI-LFA with link protection option. TI-LFA provides 100% link and node protection and micro loop avoidance. It always routes protected traffic on the post convergence path. Since Cox metro topology is dual egress hub-and-spoke topology, the primary benefit of TI-LFA implementation is micro-loop avoidance rather than post-convergence optimization.

Since TI-LFA uses Segment Routing for the repair path, SR must be deployed in network for TI-LFA to work. Following are the benefits of TI-LFA:

- TI-LFA provides less than 50ms link, node and SRLG (Shared Risk Link Group) protection with 100% coverage.
- The repair path is automatically computed by IGP.
- TI-LFA uses a post-convergence path as backup path.
- TI-LFA can be incrementally deployed; it is locally significant.
- TI-LFA also protects LDP and IP traffic in addition to SR traffic.

5.1. TI – LFA Protection Options

TI-LFA protection options ensure a repair path along the post convergence path in any topology. TI-LFA offers following protection options:

- **Link Protection**: Link protection is the most basic form of the options that is available for TI-LFA and all it needs is to be just enabled for all the links that need protection. Link protection is the only protection option that is enabled in Cox metro network. All interfaces that are running MPLS/LDP are enabled for link protection.
- **Node Protection**: In the node protection, the neighbor node is excluded during the post convergence backup path calculation.
- **SRLG protection**: SRLG refers to the scenario where network links share the same fiber or a common physical attribute. Such links are susceptible to collective risk. TI-LFA SRLG protection attempts to find the post-convergence backup path that excludes the SRLG of the protected link. All local links that share any SRLG with the protecting link are excluded.

The following output on H1 router shows the post convergence backup path calculated by H1 to reach R1 in figure 2 using TI-LFA.

```
RP/0/0/CPU0:H1#show isis fast-reroute 10.0.0.1/32
Thu Aug 17 06:30:28.144 UTC

L1 10.0.0.1/32 [16/115]
     via 100.120.100.36, GigabitEthernet0/0/0/1, R1, SRGB Base: 300000, Weight: 0
         Backup path: TI-LFA (link), via 100.120.100.39, GigabitEthernet0/0/0/0 H2,
SRGB Base: 300000, Weight: 0
           P node: R2.00 [10.0.0.2], Label: 300058
           Prefix label: 300054
```

6. LDP to SR Migration

Migrating label switching mechanism in a network from Label Distribution Protocol (LDP) to Segment Routing (SR) involves a well-planned process to transition the way traffic is routed and forwarded. This migration aims to leverage the benefits of SR, such as fast reroute, enhanced traffic engineering and path control, while ensuring minimal disruption to ongoing network operations. Before migrating LDP to SR, it is important to assess the network's current architecture, topology, and traffic patterns. Identify the specific areas or services that would benefit from SR's capabilities and then plan the migration strategy, considering factors such as network complexity, critical services, and potential impact on end-users.

Cox chose to deploy Segment Routing along with the existing LDP in its metro network. Adopting a parallel and incremental deployment approach allows for a gradual transition without affecting ongoing operations. This approach is also called “ships-in-the-night” and avoids the use of Segment Routing Mapping Server (SRMS) as Cox will not be enabling “interworking” between LDP and SR. The "ships-in-the-night" method is a gradual and non-disruptive approach for migrating from Label Distribution Protocol (LDP) to Segment Routing (SR) within a network. It is designed to minimize service disruption and ensure a smooth transition from one technology to another. The concept behind the ships-in-the-night migration is to allow both LDP and SR to coexist temporarily, with traffic being migrated gradually over time without affecting ongoing network operations.

Below is the high-level migration plan to enable Segment routing and removal of LDP protocol. The implementation will be split into three phases.

Phase 1: SR and LDP co-existence by configuring SR and letting LDP be the preferred label imposition method.

Phase 2: Preferring SR over LDP as a label imposition method.

Phase 3: Remove LDP.

Protocol Preferences

Like different routing protocols have different preferences, label switching protocols (LDP, SR, RSVP-TE etc.) also have different protocol preferences. When an operator changes the label preference on the node, the LSR will reprogram its FIB to "push" labels based on the preferred preference value (usually lower value is considered preferable). Thus, LSPs for which the LSR is the headend are impacted by this change. When a LSR terminates a service and has the preference changed, a service interruption can occur. This scenario is applicable to the PE routers in the network. Since LSPs are unidirectional, the remote LSR's LSP is not impacted by the change. The following table shows how vendors can use protocol preference values for LDP and SR to influence which label switching protocol to use to switch labeled traffic in the network.

Table 2 – Example protocol preference values

Vendor	Protocols and default preference values
Vendor A	LDP = 9 SR (L-ISIS) = 14
Vendor B	LDP = 55 IS-IS SR = 65
Vendor C	LDP = - (default preferred label switching protocol) SR = sr-prefer (makes SR preferable over LDP)
Vendor D	LDP = 9 SR-ISIS = 11

Network operators can lower the preference value of SR than that of LDP or raise the preference value of LDP to be greater than SR to make SR a preferred label switching protocol in the network.

7. Conclusion

As the networking landscape continues to evolve, transitioning from LDP to SR can offer significant advantages to network operators. Segment Routing is a flexible and efficient source routing mechanism that would simplify the control plane of a network. It can provide several benefits such as less than 50ms fast reroute for traffic in link and node failure scenario, deterministic and efficient traffic engineering capabilities, and network programmability that are otherwise not available using LDP. As Cox introduces SR in Cox metro network, the experience and knowledge acquired can be easily transformed into other part of the network such as core backbone, data center network, and business network. Cox considers the introduction of SR in Cox metro network as a solid foundation for the evolution of programmable network, and it will have a tremendous impact on Cox's next generation network evolution and modernization.

Abbreviations

SR	Segment Routing
LDP	Label Distribution Protocol
LSD	Label Switching Database
MPLS	Multi-Protocol Label Switching
IS-IS	Intermediate System-Intermediate System
TI-LFA	Topology Independent - Loop Free Alternate
SRLG	Shared Risk Link Group
EXP	MPLS Experimental Bits
ECMP	Equal Cost Multi Paths
MTU	Maximum Transmission Unit
PON	Passive Optical Network
BGP	Broader Gateway Protocol
LSR	Label Switch Router
SRMS	Segment Routing Mapping Server
MSD	Maximum Segment Depth

Bibliography & References

Clarence Filsfils, Ahmed Bashandy, Hannes Gredler, & Bruno Decraene. (2019). IS-IS Extensions for Segment Routing—RFC 8667. *Internet Engineering Task Force (IETF)*. https://doi.org/10.17487/RFC8667

Clarence Filsfils, Kris Michielsen, & Ketan Talaulikar. (2017). *Segment Routing Part I*. Cisco Systems.

Clarence Filsfils, Stefano Previdi, Les Ginsberg, Bruno Decraene, Stephane Litkowski, & Rob Shakir. (2018). Segment Routing Architecture—RFC 8402. *Internet Engineering Task Force (IETF)*. https://doi.org/10.17487/RFC8402

Ventre, P. L., Salsano, S., Polverini, M., Cianfrani, A., Abdelsalam, A., Filsfils, C., Camarillo, P., & Clad, F. (2021). Segment Routing: A Comprehensive Survey of Research Activities, Standardization Efforts, and Implementation Results. *IEEE Communications Surveys & Tutorials*, *23*(1), 182–221. https://doi.org/10.1109/COMST.2020.3036826

