Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 6, 2019

A. Bashandy, Ed.
Individual
C. Filsfils, Ed.
S. Previdi
Cisco Systems, Inc.
B. Decraene
S. Litkowski
Orange

September 02, 2018

Segment Routing interworking with LDP
draft-ietf-spring-segment-routing-ldp-interop-15

Abstract

A Segment Routing (SR) node steers a packet through a controlled set of instructions, called segments, by prepending the packet with an SR header. A segment can represent any instruction, topological or service-based. SR allows to enforce a flow through any topological path while maintaining per-flow state only at the ingress node to the SR domain.

The Segment Routing architecture can be directly applied to the MPLS data plane with no change in the forwarding plane. This document describes how Segment Routing operates in a network where LDP is deployed and in the case where SR-capable and non-SR-capable nodes coexist.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any

time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on March 6, 2019.

Table of Contents

1. Introduction

Segment Routing, as described in [I-D.ietf-spring-segment-routing], can be used on top of the MPLS data plane without any modification as described in [I-D.ietf-spring-segment-routing-mpls].

Segment Routing control plane can co-exist with current label distribution protocols such as LDP ([RFC5036]).

This document outlines the mechanisms through which SR interworks with LDP in cases where a mix of SR-capable and non-SR-capable routers co-exist within the same network and more precisely in the same routing domain.

Section 2 describes the co-existence of SR with other MPLS Control Plane protocols. Section 3 documents the interworking between SR and LDP in the case of non-homogeneous deployment. Section 4 describes how a partial SR deployment can be used to provide SR benefits to LDP-based traffic including a possible application of SR in the context of inter-domain MPLS use-cases. Appendix A documents a method to migrate from LDP to SR-based MPLS tunneling.

Typically, an implementation will allow an operator to select (through configuration) which of the described modes of SR and LDP co-existence to use.

2. SR/LDP Ships-in-the-night coexistence

"MPLS Control Plane Client (MCC)" refers to any control plane protocol installing forwarding entries in the MPLS data plane. SR, LDP [RFC5036], RSVP-TE [RFC3209], BGP [RFC8277], etc are examples of MCCs.

An MCC, operating at node N, must ensure that the incoming label it installs in the MPLS data plane of Node N has been uniquely allocated to himself.

Segment Routing makes use of the Segment Routing Global Block (SRGB, as defined in [I-D.ietf-spring-segment-routing]) for the label allocation. The use of the SRGB allows SR to co-exist with any other MCC.

This is clearly the case for the adjacency segment: it is a local label allocated by the label manager, as for any MCC.

This is clearly the case for the prefix segment: the label manager allocates the SRGB set of labels to the SR MCC client and the operator ensures the unique allocation of each global prefix segment/ label within the allocated SRGB set.

Note that this static label allocation capability of the label manager has existed for many years across several vendors and hence is not new. Furthermore, note that the label-manager ability's to statically allocate a range of labels to a specific application is not new either. This is required for MPLS-TP operation. In this case, the range is reserved by the label manager and it is the MPLS-TP ([RFC5960]) NMS (acting as an MCC) that ensures the unique allocation of any label within the allocated range and the creation of the related MPLS forwarding entry.

Let us illustrate an example of ship-in-the-night (SIN) coexistence.

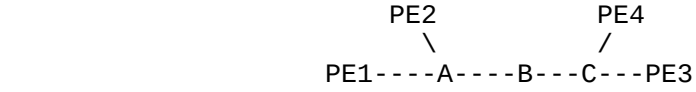


Figure 1: SIN coexistence

The EVEN VPN service is supported by PE2 and PE4 while the ODD VPN service is supported by PE1 and PE3. The operator wants to tunnel the ODD service via LDP and the EVEN service via SR.

This can be achieved in the following manner:

The operator configures PE1, PE2, PE3, PE4 with respective loopbacks 192.0.2.201/32, 192.0.2.202/32, 192.0.2.203/32, 192.0.2.204/32. These PE's advertised their VPN routes with next-hop set on their respective loopback address.

The operator configures A, B, C with respective loopbacks 192.0.2.1/32, 192.0.2.2/32, 192.0.2.3/32.

The operator configures PE2, A, B, C and PE4 with SRGB [100, 300].

The operator attaches the respective Node Segment Identifiers (Node-SID's, as defined in [I-D.ietf-spring-segment-routing]): 202, 101, 102, 103 and 204 to the loopbacks of nodes PE2, A, B, C and PE4. The Node-SID's are configured to request penultimate-hop-popping.

PE1, A, B, C and PE3 are LDP capable.

PE1 and PE3 are not SR capable.

PE3 sends an ODD VPN route to PE1 with next-hop 192.0.2.203 and VPN label 10001.

From an LDP viewpoint: PE1 received an LDP label binding (1037) for a forwarding equivalence class (FEC) 192.0.2.203/32 from its next-hop A. A received an LDP label binding (2048) for that FEC from its next-hop B. B received an LDP label binding (3059) for that FEC from its next-hop C. C received implicit-null LDP binding from its next-hop PE3.

As a result, PE1 sends its traffic to the ODD service route advertised by PE3 to next-hop A with two labels: the top label is 1037 and the bottom label is 10001. Node A swaps 1037 with 2048 and forwards to B. B swaps 2048 with 3059 and forwards to C. C pops 3059 and forwards to PE3.

PE4 sends an EVEN VPN route to PE2 with next-hop 192.0.2.204 and VPN label 10002.

From an SR viewpoint: PE2 maps the IGP route 192.0.2.204/32 onto Node-SID 204; node A swaps 204 with 204 and forwards to B; B swaps 204 with 204 and forwards to C; C pops 204 and forwards to PE4.

As a result, PE2 sends its traffic to the VPN service route advertised by PE4 to next-hop A with two labels: the top label is 204 and the bottom label is 10002. Node A swaps 204 with 204 and forwards to B. B swaps 204 with 204 and forwards to C. C pops 204 and forwards to PE4.

The two modes of MPLS tunneling co-exist.

The ODD service is tunneled from PE1 to PE3 through a continuous LDP LSP traversing A, B and C.

The EVEN service is tunneled from PE2 to PE4 through a continuous SR node segment traversing A, B and C.

2.1. MPLS2MPLS, MPLS2IP and IP2MPLS co-existence

MPLS2MPLS refers to the forwarding behavior where a router receives a labeled packet and switches it out as a labeled packet. Several MPLS2MPLS entries may be installed in the data plane for the same prefix.

Let us examine A’s MPLS forwarding table as an example:

Incoming label: 1037

- outgoing label: 2048
- outgoing next-hop: B

Note: this entry is programmed by LDP for 192.0.2.203/32

Incoming label: 203

- outgoing label: 203
- outgoing next-hop: B

Note: this entry is programmed by SR for 192.0.2.203/32

These two entries can co-exist because their incoming label is unique. The uniqueness is guaranteed by the label manager allocation rules.

The same applies for the MPLS2IP forwarding entries. MPLS2IP is the forwarding behavior where a router receives a label IPv4/IPv6 packet with one label only, pops the label, and switches the packet out as IPv4/IPv6. For IP2MPLS coexistence, refer to Section 6.1.

3. SR and LDP Interworking

This section analyzes the case where SR is available in one part of the network and LDP is available in another part. It describes how a continuous MPLS tunnel can be built throughout the network.

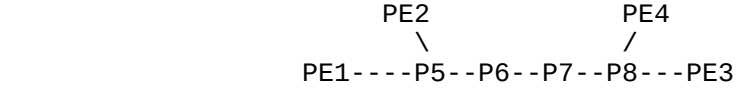


Figure 2: SR and LDP Interworking

Let us analyze the following example:

P6, P7, P8, PE4 and PE3 are LDP capable.

PE1, PE2, P5 and P6 are SR capable. PE1, PE2, P5 and P6 are configured with SRGB (100, 200) and respectively with node segments 101, 102, 105 and 106.

A service flow must be tunneled from PE1 to PE3 over a continuous MPLS tunnel encapsulation and hence SR and LDP need to interwork.

3.1. LDP to SR

In this section, a right-to-left traffic flow is analyzed.

PE3 has learned a service route whose next-hop is PE1. PE3 has an LDP label binding from the next-hop P8 for the FEC "PE1". Hence PE3 sends its service packet to P8 as per classic LDP behavior.

P8 has an LDP label binding from its next-hop P7 for the FEC "PE1" and hence P8 forwards to P7 as per classic LDP behavior.

P7 has an LDP label binding from its next-hop P6 for the FEC "PE1" and hence P7 forwards to P6 as per classic LDP behavior.

P6 does not have an LDP binding from its next-hop P5 for the FEC "PE1". However P6 has an SR node segment to the IGP route "PE1". Hence, P6 forwards the packet to P5 and swaps its local LDP-label for FEC "PE1" by the equivalent node segment (i.e. 101).

P5 pops 101 (assuming PE1 advertised its node segment 101 with the penultimate-pop flag set) and forwards to PE1.

PE1 receives the tunneled packet and processes the service label.

The end-to-end MPLS tunnel is built from an LDP LSP from PE3 to P6 and the related node segment from P6 to PE1.

3.1.1. LDP to SR Behavior

It has to be noted that no additional signaling or state is required in order to provide interworking in the direction LDP to SR.

A SR node having LDP neighbors MUST create LDP bindings for each Prefix-SID learned in the SR domain by treating SR learned labels as if they were learned through an LDP neighbot. In addition for each FEC, the SR node stitches the incoming LDP label to the outgoing SR label. This has to be done in both LDP independent and ordered label distribution control modes as defined in [RFC5036].

3.2. SR to LDP

In this section, the left-to-right traffic flow is analyzed.

This section defines the Segment Routing Mapping Server (SRMS). The SRMS is a IGP node advertising mapping between Segment Identifiers (SID) and prefixes advertised by other IGP nodes. The SRMS uses a dedicated IGP extension (IS-IS, OSPFv2 and OSPFv3) which is protocol specific and defined in [I-D.ietf-isis-segment-routing-extensions],

[I-D.ietf-ospf-segment-routing-extensions], and
[I-D.ietf-ospf-ospfv3-segment-routing-extensions].

The SRMS function of a SR capable router allows distribution of mappings for prefixes not locally attached to the advertising router and therefore allows advertisement of mappings on behalf of non-SR capable routers.

The SRMS is a control plane only function which may be located anywhere in the IGP flooding scope. At least one SRMS server MUST exist in a routing domain to advertise prefix-SIDs on behalf non-SR nodes, thereby allowing non-LDP routers to send and receive labeled traffic from LDP-only routers. Multiple SRMSs may be present in the same network (for redundancy). This implies that there are multiple ways a prefix-to-SID mapping can be advertised. Conflicts resulting from inconsistent advertisements are addressed by [I-D.ietf-spring-segment-routing-mpls].

The example diagram depicted in Figure 2 assumes that the operator configures P5 to act as a Segment Routing Mapping Server (SRMS) and advertises the following mappings: (P7, 107), (P8, 108), (PE3, 103) and (PE4, 104).

The mappings advertised by one or more SRMSs result from local policy information configured by the operator.

If PE3 had been SR capable, the operator would have configured PE3 with node segment 103. Instead, as PE3 is not SR capable, the operator configures that policy at the SRMS and it is the latter which advertises the mapping.

The mapping server advertisements are only understood by SR capable routers. The SR capable routers install the related node segments in the MPLS data plane exactly like the node segments had been advertised by the nodes themselves.

For example, PE1 installs the node segment 103 with next-hop P5 exactly as if PE3 had advertised node segment 103.

PE1 has a service route whose next-hop is PE3. PE1 has a node segment for that IGP route: 103 with next-hop P5. Hence PE1 sends its service packet to P5 with two labels: the bottom label is the service label and the top label is 103.

P5 swaps 103 for 103 and forwards to P6.

P6’s next-hop for the IGP route "PE3" is not SR capable (P7 does not advertise the SR capability). However, P6 has an LDP label binding

from that next-hop for the same FEC (e.g. LDP label 1037). Hence, P6 swaps 103 for 1037 and forwards to P7.

P7 swaps this label with the LDP-label received from P8 and forwards to P8.

P8 pops the LDP label and forwards to PE3.

PE3 receives the tunneled packet and processes the service label.

The end-to-end MPLS tunnel is built from an SR node segment from PE1 to P6 and an LDP LSP from P6 to PE3.

SR mapping advertisement for a given prefix provides no information about the Penultimate Hop Popping. Other mechanisms, such as IGP specific mechanisms ([I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions] and [I-D.ietf-ospf-ospfv3-segment-routing-extensions]), MAY be used to determine the Penultimate Hop Popping in such case.

Note: In the previous example, Penultimate Hop Popping is not performed at the SR/LDP border for segment 103 (PE3), because none of the routers in the SR domain is Penultimate Hop for segment 103. In this case P6 requires the presence of the segment 103 such as to map it to the LDP label 1037.

3.2.1. Segment Routing Mapping Server (SRMS)

This section specifies the concept and externally visible functionality of a segment routing mapping server (SRMS).

The purpose of a SRMS functionality is to support the advertisement of prefix-SIDs to a prefix without the need to explicitly advertise such assignment within a prefix reachability advertisment. Examples of explicit prefix-SID advertisment are the prefix-SID sub-TLVs defined in ([I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions], and [I-D.ietf-ospf-ospfv3-segment-routing-extensions]).

The SRMS functionality allows assigning of prefix-SIDs to prefixes owned by non-SR-capable routers as well as to prefixes owned by SR capable nodes. It is the former capability which is essential to the SR-LDP interworking described later in this section

The SRMS functionality consists of two functional blocks: the Mapping Server (MS) and Mapping Client (MC).

A MS is a node that advertises an SR mappings. Advertisements sent by an MS define the assignment of a prefix-SID to a prefix independent of the advertisment of reachability to the prefix itself. An MS MAY advertise SR mappings for any prefix whether or not it advertises reachability for the prefix and irrespective of whether that prefix is advertised by or even reachable through any router in the network.

An MC is a node that receives and uses the MS mapping advertisments. Note that a node may be both an MS and an MC. An MC interprets the SR mapping advertisment as an assignment of a prefix-SID to a prefix. For a given prefix, if an MC receives an SR mapping advertisement from a mapping server and also has received a prefix-SID advertisement for that same prefix in a prefix reachability advertisement, then the MC MUST prefer the SID advertised in the prefix reachability advertisement over the mapping server advertisement i.e., the mapping server advertisment MUST be ignored for that prefix. Hence assigning a prefix-SID to a prefix using the SRMS functionality does not preclude assigning the same or different prefix-SID(s) to the same prefix using explict prefix-SID advertisement such as the aforementioned prefix-SID sub-TLVs.

For example consider an IPv4 prefix advertisement received by an IS-IS router in the extended IP reachability TLV (TLV 135). Suppose TLV 135 contained the prefix-SID sub-TLV. If the router that receives TLV 135 with the prefix-SID sub-TLV also received an SR mapping advertisement for the same prefix through the SID/label binding TLV, then the receiving router must prefer the prefix-SID sub-TLV over the SID/label binding TLV for that prefix. Refer to ([I-D.ietf-isis-segment-routing-extensions], for details about the prefix-SID sub-TLV and SID/label binding TLV.

3.2.2. SR to LDP Behavior

SR to LDP interworking requires a SRMS as defined above.

Each SR capable router installs in the MPLS data plane Node-SIDs learned from the SRMS exactly like if these SIDs had been advertised by the nodes themselves.

A SR node having LDP neighbors MUST stitch the incoming SR label (whose SID is advertised by the SRMS) to the outgoing LDP label.

It has to be noted that the SR to LDP behavior does not propagate the status of the LDP FEC which was signaled if LDP was configured to use the ordered mode.

It has to be noted that in the case of SR to LDP, the label binding is equivalent to the independent LDP Label Distribution Control Mode ([RFC5036]) where a label in bound to a FEC independently from the received binding for the same FEC.

3.2.3. Interoperability of Multiple SRMSes and Prefix-SID advertisements

In the case of SR/LDP interoperability through the use of a SRMS, mappings are advertised by one or more SRMS.

SRMS function is implemented in the link-state protocol (such as IS-IS and OSPF). Link-state protocols allow propagation of updates across area boundaries and therefore SRMS advertisements are propagated through the usual inter-area advertisement procedures in link-state protocols.

Multiple SRMSs can be provisioned in a network for redundancy. Moreover, a preference mechanism may also be used among SRMSs so to deploy a primary/secondary SRMS scheme allowing controlled modification or migration of SIDs.

The content of SRMS advertisement (i.e.: mappings) are a matter of local policy determined by the operator. When multiple SRMSs are active, it is necessary that the information (mappings) advertised by the different SRMSs is aligned and consistent. The following mechanism is applied to determine the preference of SRMS advertisements:

If a node acts as an SRMS, it MAY advertise a preference to be associated with all SRMS SID advertisements sent by that node. The means of advertising the preference is defined in the protocol specific drafts e.g.,[I-D.ietf-isis-segment-routing-extensions] , [I-D.ietf-ospf-segment-routing-extensions], and [I-D.ietf-ospf-ospfv3-segment-routing-extensions]. The preference value is an unsigned 8 bit integer with the following properties:

0 - Reserved value indicating advertisements from that node MUST NOT be used.

1 - 255 Preference value (255 is most preferred)

Advertisement of a preference value is optional. Nodes which do not advertise a preference value are assigned a preference value of 128.

A MCC on a node receiving one or more SRMS mapping advertisements applies them as follows

- For any prefix for which it did not receive a prefix-SID advertisement, the MCC applies the SRMS mapping advertisments with the highest preference. The mechanism by which a prefix-SID is advertised for a given prefix is defined in the protocol specification , [I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions] and [I-D.ietf-ospf-ospfv3-segment-routing-extensions]

- If there is an incoming label collision as specified in [I-D.ietf-spring-segment-routing-mpls] , apply the steps specified in [I-D.ietf-spring-segment-routing-mpls] to resolve the collision.

When the SRMS advertise mappings, an implementation should provide a mechanism through which the operator determines which of the IP2MPLS mappings are preferred among the one advertised by the SRMS and the ones advertised by LDP.

4. SR/LDP Interworking Use Cases

SR can be deployed such as to enhance LDP transport. The SR deployment can be limited to the network region where the SR benefits are most desired.

4.1. SR Protection of LDP-based Traffic

In Figure 4, let us assume:

All link costs are 10 except FG which is 30.

All routers are LDP capable.

X, Y and Z are PE’s participating to an important service S.

The operator requires 50msec link-based Fast Reroute (FRR) for service S.

A, B, C, D, E, F and G are SR capable.

X, Y, Z are not SR capable, e.g. as part of a staged migration from LDP to SR, the operator deploys SR first in a sub-part of the network and then everywhere.

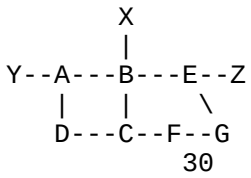


Figure 3: SR/LDP interworking example

The operator would like to resolve the following issues:

To protect the link BA along the shortest-path of the important flow XY, B requires a Remote Loop-Free alternate (RLFA, [RFC7490]) repair tunnel to D and hence a targeted LDP session from B to D. Typically, network operators prefer avoiding these dynamically established multi-hop LDP sessions in order to reduce the number of protocols running in the network and hence simplify network operations.

There is no LFA/RLFA solution to protect the link BE along the shortest path of the important flow XZ. The operator wants a guaranteed link-based FRR solution.

The operator can meet these objectives by deploying SR only on A, B, C, D, E, F and G:

The operator configures A, B, C, D, E, F and G with SRGB [100, 200] and respective node segments 101, 102, 103, 104, 105, 106 and 107.

The operator configures D as an SR Mapping Server with the following policy mapping: (X, 201), (Y, 202), (Z, 203).

Each SR node automatically advertises local adjacency segment for its IGP adjacencies. Specifically, F advertises adjacency segment 9001 for its adjacency FG.

A, B, C, D, E, F and G keep their LDP capability and hence the flows XY and XZ are transported over end-to-end LDP LSP’s.

For example, LDP at B installs the following MPLS data plane entries:

Incoming label: local LDP label bound by B for FEC Y
Outgoing label: LDP label bound by A for FEC Y
Outgoing next-hop: A

Incoming label: local LDP label bound by B for FEC Z
Outgoing label: LDP label bound by E for FEC Z

Outgoing next-hop: E

The novelty comes from how the backup chains are computed for these LDP-based entries. While LDP labels are used for the primary next-hop and outgoing labels, SR information is used for the FRR construction. In steady state, the traffic is transported over LDP LSP. In transient FRR state, the traffic is backup thanks to the SR enhanced capabilities.

The RLFA paths are dynamically pre-computed as defined in [RFC7490]. Typically, implementations allow to enable RLFA mechanism through a simple configuration command that triggers both the pre-computation and installation of the repair path. The details on how RLFA mechanisms are implemented and configured is outside the scope of this document and not relevant to the aspects of SR/LDP interwork explained in this document.

This helps meet the requirements of the operator:

Eliminate targeted LDP session.

Guaranteed FRR coverage.

Keep the traffic over LDP LSP in steady state.

Partial SR deployment only where needed.

4.2. Eliminating Targeted LDP Session

B’s MPLS entry to Y becomes:

- Incoming label: local LDP label bound by B for FEC Y
  Outgoing label: LDP label bound by A for FEC Y
  Backup outgoing label: SR node segment for Y {202}
  Outgoing next-hop: A
  Backup next-hop: repair tunnel: node segment to D {104}
  with outgoing next-hop: C

It has to be noted that D is selected as Remote Loop-Free Alternate (RLFA) as defined in [RFC7490].

In steady-state, X sends its Y-destined traffic to B with a top label which is the LDP label bound by B for FEC Y. B swaps that top label for the LDP label bound by A for FEC Y and forwards to A. A pops the LDP label and forwards to Y.

Upon failure of the link BA, B swaps the incoming top-label with the node segment for Y (202) and sends the packet onto a repair tunnel to

D (node segment 104). Thus, B sends the packet to C with the label stack {104, 202}. C pops the node segment 104 and forwards to D. D swaps 202 for 202 and forwards to A. A's next-hop to Y is not SR capable and hence node A swaps the incoming node segment 202 to the LDP label announced by its next-hop (in this case, implicit null).

After IGP convergence, B's MPLS entry to Y will become:

- Incoming label: local LDP label bound by B for FEC Y
  Outgoing label: LDP label bound by C for FEC Y
  Outgoing next-hop: C

And the traffic XY travels again over the LDP LSP.

Conclusion: the operator has eliminated the need for targeted LDP sessions (no longer required) and the steady-state traffic is still transported over LDP. The SR deployment is confined to the area where these benefits are required.

Despite that in general, an implementation would not require a manual configuration of LDP Targeted sessions however, it is always a gain if the operator is able to reduce the set of protocol sessions running on the network infrastructure.

4.3. Guaranteed FRR coverage

As mentioned in Section 4.1 above, in the example topology described in Figure 4, there is no RLFA-based solution for protecting the traffic flow YZ against the failure of link BE because there is no intersection between the extended P-space and Q-space (see [RFC7490] for details). However:

- G belongs to the Q space of Z.

- G can be reached from B via a "repair SR path" {106, 9001} that is not affected by failure of link BE (The method by which G and the repair tunnel to it from B are identified are out of scope of this document.)

B's MPLS entry to Z becomes:

- Incoming label: local LDP label bound by B for FEC Z
  Outgoing label: LDP label bound by E for FEC Z
  Backup outgoing label: SR node segment for Z {203}
  Outgoing next-hop: E
  Backup next-hop: repair tunnel to G: {106, 9001}

      G is reachable from B via the combination of a node segment to F {106} and an adjacency segment FG {9001}

      Note that {106, 107} would have equally work. Indeed, in many case, P's shortest path to Q is over the link PQ. The adjacency segment from P to Q is required only in very rare topologies where the shortest-path from P to Q is not via the link PQ.

In steady-state, X sends its Z-destined traffic to B with a top label which is the LDP label bound by B for FEC Z. B swaps that top label for the LDP label bound by E for FEC Z and forwards to E. E pops the LDP label and forwards to Z.

Upon failure of the link BE, B swaps the incoming top-label with the node segment for Z (203) and sends the packet onto a repair tunnel to G (node segment 106 followed by adjacency segment 9001). Thus, B sends the packet to C with the label stack {106, 9001, 203}. C pops the node segment 106 and forwards to F. F pops the adjacency segment 9001 and forwards to G. G swaps 203 for 203 and forwards to E. E's next-hop to Z is not SR capable and hence E swaps the incoming node segment 203 for the LDP label announced by its next-hop (in this case, implicit null).

After IGP convergence, B's MPLS entry to Z will become:

- Incoming label: local LDP label bound by B for FEC Z
  Outgoing label: LDP label bound by C for FEC Z
  Outgoing next-hop: C

And the traffic XZ travels again over the LDP LSP.

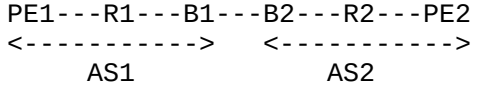
Conclusions:

- the operator has eliminated its second problem: guaranteed FRR coverage is provided. The steady-state traffic is still transported over LDP. The SR deployment is confined to the area where these benefits are required.

- FRR coverage has been achieved without any signaling for setting up the repair LSP and without setting up a targeted LDP session between B and G.

4.4. Inter-AS Option C, Carrier's Carrier

In inter-AS Option C [RFC4364], two interconnected ASes sets up inter-AS MPLS connectivity. SR may be independently deployed in each AS.



AS1 AS2

Figure 4: Inter-AS Option C

In Inter-AS Option C, B2 advertises to B1 a labeled BGP route [RFC8277] for PE2 and B1 reflects it to its internal peers, such as PE1. PE1 learns from a service route reflector a service route whose next-hop is PE2. PE1 resolves that service route on the labeled BGP route to PE2. That labeled BGP route to PE2 is itself resolved on the AS1 IGP route to B1.

If AS1 operates SR, then the tunnel from PE1 to B1 is provided by the node segment from PE1 to B1.

PE1 sends a service packet with three labels: the top one is the node segment to B1, the next-one is the label in the labeled BGP route provided by B1 for the route "PE2" and the bottom one is the service label allocated by PE2.

5. IANA Considerations

This document does not introduce any new codepoint.

6. Manageability Considerations

6.1. SR and LDP co-existence

When both SR and LDP co-exist, the following applies:

- If both SR and LDP propose an IP2MPLS entry for the same IP prefix, then by default the LDP route SHOULD be selected. This is because it is expected that SR is introduced into network that contain routers that do not support SR. Hence by having a behavior that prefers LDP over SR, traffic flow is unlikely to be disrupted

- A local policy on a router MUST allow to prefer the SR-provided IP2MPLS entry.

- Note that this policy MAY be locally defined. There is no requirement that all routers use the same policy.

6.2. Dataplane Verification

When Label switch paths (LSPs) are defined by stitching LDP LSPs with SR LSPs, it is necessary to have mechanisms allowing the verification of the LSP connectivity as well as validation of the path. These mechanisms are described in [RFC8287].

7. Security Considerations

This document does not introduce any change to the MPLS dataplane [RFC3031] and therefore no additional security of the MPLS dataplane is required.

This document introduces another form of label binding advertisements. The security associated with these advertisements is part of the security applied to routing protocols such as IS-IS [RFC5304] and OSPF [RFC5709] which both optionally make use of cryptographic authentication mechanisms. This form of advertisement is more centralized, on behalf of the node advertising the IP reachability, which presents a different risk profile. This document also specifies a mechanism by which the ill effects of advertising conflicting label bindings can be mitigated. In particular, advertisements from the node advertising the IP reachability is more preferred than the centralized one. Because this document recognizes that reachability, which presents a different risk profile. This document miscofiguration and/or programming may result in false or conflicting also specifies a mechanism by which the ill effects of advertising label binding advertisements, thereby compromising traffic conflicting label bindings can be mitigated. In particular, forwarding, the document recommends strict configuration/ advertisements from the node advertising the IP reachability is more programmability control as well as montoring the SID advertised and preferred than the centralized one. log/error messages by the operator to avoid or at least significantly minimize the possibility of such risk.

8. Acknowledgements

The authors would like to thank Pierre Francois, Ruediger Geib and Alexander Vainshtein for their contribution to the content of this document.

9. Contributors' Addresses

Edward Crabbe
Individual
Email: edward.crabbe@gmail.com

Igor Milojevic
Email: milojevicigor@gmail.com

Saku Ytti
TDC
Email: saku@ytti.fi

Rob Shakir
Google
Email: robjs@google.com

Martin Horneffer
Deutsche Telekom
Email: Martin.Horneffer@telekom.de

Wim Henderickx
Nokia
Email: wim.henderickx@nokia.com

Jeff Tantsura
Individual
Email: jefftant@gmail.com

Les Ginseberg
Cisco Systems
Email: ginsberg@cisco.com

10. References

10.1. Normative References

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", January 2018.

[I-D.ietf-spring-segment-routing-mpls]
Bashandy, A., Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-13 (work in progress), April 2018.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <https://www.rfc-editor.org/info/rfc5036>.

10.2. Informative References

[I-D.ietf-isis-segment-routing-extensions] Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-19 (work in progress), July 2018.

[I-D.ietf-ospf-ospfv3-segment-routing-extensions] Psenak, P., Filsfils, C., Previdi, S., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPFv3 Extensions for Segment Routing", draft-ietf-ospf-ospfv3-segment-routing-extensions-11 (work in progress), January 2018.

[I-D.ietf-ospf-segment-routing-extensions] Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-24 (work in progress), December 2017.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC3209] Awduche, D., Berger, L., Gan, G., Li, T., Srinivasan, V., and G. Srinivasan, "RSVP-TE: Extensions to RSVP for LSP Tunnels", December 2001.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, DOI 10.17487/RFC5304, October 2008, <https://www.rfc-editor.org/info/rfc5304>.

[RFC5709] Bhatia, M., Manral, V., Fanto, M., White, R., Barnes, M., Li, T., and R. Atkinson, "OSPFv2 HMAC-SHA Cryptographic Authentication", RFC 5709, DOI 10.17487/RFC5709, October 2009, <https://www.rfc-editor.org/info/rfc5709>.

[RFC5960] Frost, D., Ed., Bryant, S., Ed., and M. Bocci, Ed., "MPLS Transport Profile Data Plane Architecture", RFC 5960, DOI 10.17487/RFC5960, August 2010, <https://www.rfc-editor.org/info/rfc5960>.

[RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", RFC 7490, DOI 10.17487/RFC7490, April 2015, <https://www.rfc-editor.org/info/rfc7490>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", October 2017.

[RFC8287] Kumar, N., Pignataro, C., Swallow, G., Akiya, N., Kini, S., and M. Chen, "Label Switched Path (LSP) Ping/ Traceroute for Segment Routing (SR) IGP-Prefix and IGP-Adjacency Segment Identifiers (SIDs) with MPLS Data Planes", December 2017.

[RFC8355] Filsfils, C., Previdi, S., Decraene, B., and R. Shakir, "Resiliency Use Cases in Source Packet Routing in Networking (SPRING) Networks", March 2018.

Appendix A. Migration from LDP to SR

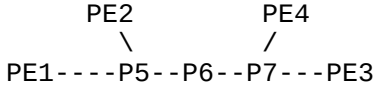


Figure 5: Migration

Several migration techniques are possible. The technique described here is inspired by the commonly used method to migrate from one IGP to another.

At time T0, all the routers run LDP. Any service is tunneled from an ingress PE to an egress PE over a continuous LDP LSP.

At time T1, all the routers are upgraded to SR. They are configured with the SRGB range [100, 300]. PE1, PE2, PE3, PE4, P5, P6 and P7 are respectively configured with the node segments 101, 102, 103, 104, 105, 106 and 107 (attached to their service-recursing loopback).

At this time, the service traffic is still tunneled over LDP LSP. For example, PE1 has an SR node segment to PE3 and an LDP LSP to PE3 but by default, as seen earlier, the LDP IP2MPLS encapsulation is preferred. However, it has to be noted that the SR infrastructure is usable, e.g. for Fast Reroute (FRR) or IGP Loop Free Convergence to protect existing IP and LDP traffic. FRR mechanisms are described in and [RFC8355].

At time T2, the operator enables the local policy at PE1 to prefer SR IP2MPLS encapsulation over LDP IP2MPLS.

The service from PE1 to any other PE is now riding over SR. All other service traffic is still transported over LDP LSP.

At time T3, gradually, the operator enables the preference for SR IP2MPLS encapsulation across all the edge routers.

All the service traffic is now transported over SR. LDP is still operational and services could be reverted to LDP.

At time T4, LDP is unconfigured from all routers.

Authors' Addresses

Ahmed Bashandy (editor)
Individual
USA

Email: abashandy.ietf@gmail.com

Clarence Filsfils (editor)
Cisco Systems, Inc.
Brussels
BE

Email: cfilsfil@cisco.com

Stefano Previdi
Cisco Systems, Inc.
IT

Email: stefano@previdi.net

Bruno Decraene
Orange
FR

Email: bruno.decraene@orange.com

Stephane Litkowski
Orange
FR

Email: stephane.litkowski@orange.com