Network Working Group
Internet Draft
Intended status: Standards Track
Expires: May 2019

A. Bashandy, Ed.
Arrcus
C. Filsfils, Ed.
S. Previdi,
Cisco Systems, Inc.
B. Decraene
S. Litkowski
Orange
R. Shakir
Google

November 19, 2018

Segment Routing with MPLS data plane
draft-ietf-spring-segment-routing-mpls-16

Abstract

Segment Routing (SR) leverages the source routing paradigm. A node steers a packet through a controlled set of instructions, called segments, by prepending the packet with an SR header. In the MPLS dataplane, the SR header is instantiated through a label stack. This document specifies the forwarding behavior to allow instantiating SR over the MPLS dataplane.

Copyright Notice

Table of Contents

1. Introduction

The Segment Routing architecture RFC8402 can be directly applied to the MPLS architecture with no change in the MPLS forwarding plane. This document specifies the forwarding plane behavior to allow Segment Routing to operate on top of the MPLS data plane. This document does not address the control plane behavior. Control plane behavior is specified in other documents such as [I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions], and [I-D.ietf-ospf-ospfv3-segment-routing-extensions].

The Segment Routing problem statement is described in [RFC7855].

Co-existence of SR over MPLS forwarding plane with LDP [RFC5036] is specified in [I-D.ietf-spring-segment-routing-ldp-interop].

Policy routing and traffic engineering using segment routing can be found in [I-D.ietf-spring-segment-routing-policy]

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. MPLS Instantiation of Segment Routing

MPLS instantiation of Segment Routing fits in the MPLS architecture as defined in [RFC3031] both from a control plane and forwarding plane perspective:

- o From a control plane perspective, [RFC3031] does not mandate a single signaling protocol. Segment Routing makes use of various control plane protocols such as link state IGPs [I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions] and [I-D.ietf-ospf-ospfv3-segment-routing-extensions]. The flooding mechanisms of link state IGPs fits very well with label stacking on ingress. Future control layer protocol and/or policy/configuration can be used to specify the label stack.

- o From a forwarding plane perspective, Segment Routing does not require any change to the forwarding plane because Segment IDs (SIDs) are instantiated as MPLS labels and the Segment routing header instantiated as a stack of MPLS labels.

We call "MPLS Control Plane Client (MCC)" any control plane entity installing forwarding entries in the MPLS data plane. IGPs with SR extensions [I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions], [I-D.ietf-ospf-ospfv3-segment-routing-extensions] and LDP [RFC5036] are examples of MCCs. Local configuration and policies applied on a router are also examples of MCCs.

In order to have a node segment to reach the node, a network operator SHOULD configure at least one node segment per routing instance, topology, algorithm. Otherwise, the node is not reachable within the routing instance, topology or along the routing algorithm, which restrict its ability to be used by a SR policy, including for TI-LFA. An implementation MAY check that an IGP node-SID is not associated with a prefix that is owned by more than one router within the same

routing domain. If so, it SHOULD NOT use this Node-SID, MAY use another one if available and SHOULD log an error.

2.1. Multiple Forwarding Behaviors for the Same Prefix

The SR architecture does not prohibit having more than one SID for the same prefix. In fact, by allowing multiple SIDs for the same prefix, it is possible to have different forwarding behaviors (such as different paths, different ECMP/UCMP behaviors,...,etc) for the same destination.

Instantiating Segment routing over the MPLS forwarding plane fits seamlessly with this principle. An operator may assign multiple MPLS labels or indices to the same prefix and assign different forwarding behaviors to each label/SID. The MCC in the network downloads different MPLS labels/SIDs to the FIB for different forwarding behaviors. The MCC at the entry of an SR domain or at any point in the domain can choose to apply a particular forwarding behavior to a particular packet by applying the PUSH action to that packet using the corresponding SID.

2.2. SID Representation in the MPLS Forwarding Plane

When instantiating SR over the MPLS forwarding plane, a SID is represented by an MPLS label or an index [RFC8402].

A global segment MUST be a label, or an index which may be mapped to an MPLS label within the Segment Routing Global Block (SRGB) of the node installing the global segment in its FIB/receiving the labeled packet. Section 2.4 specifies the procedure to map a global segment represented by an index to an MPLS label within the SRGB.

The MCC MUST ensure that any label value corresponding to any SID it installs in the forwarding plane follows the following rules:

- o The label value MUST be unique within the router on which the MCC is running. i.e. the label MUST only be used to represent the SID and MUST NOT be used to represent more than one SID or for any other forwarding purpose on the router.

- o The label value MUST NOT come from the range of special purpose labels [RFC7274].

Labels allocated in this document are considered per platform down-stream allocated labels [RFC3031].

2.3. Segment Routing Global Block and Local Block

The concepts of Segment Routing Global Block (SRGB) and global SID are explained in [RFC8402]. In general, the SRGB need not be a contiguous range of labels.

For the rest of this document, the SRGB is specified by the list of MPLS Label ranges [Ll(1),Lh(1)], [Ll(2),Lh(2)],..., [Ll(k),Lh(k)] where Ll(i) =< Lh(i).

The following rules apply to the list of MPLS ranges representing the SRGB

- o The list of ranges comprising the SRGB MUST NOT overlap.

- o Every range in the list of ranges specifying the SRGB MUST NOT cover or overlap with a reserved label value or range [RFC7274], respectively.

- o If the SRGB of a node does not conform to the structure specified in this section or to the previous two rules, then this SRGB MUST be completely ignored by all routers in the routing domain and the node MUST be treated as if it does not have an SRGB.

- o The list of label ranges MUST only be used to instantiate global SIDs into the MPLS forwarding plane

A Local segment MAY be allocated from the Segment Routing Local Block (SRLB) [RFC8402] or from any unused label as long as it does not use a special purpose label. The SRLB consists of the range of local labels reserved by the node for certain local segments. In a controller-driven network, some controllers or applications MAY use the control plane to discover the available set of local SIDs on a particular router [I-D.ietf-spring-segment-routing-policy]. The rules applicable to the SRGB are also applicable to the SRLB, except rule that says that the SRGB MUST only be used to instantiate global SIDs into the MPLS forwarding plane. The recommended, minimum, or maximum size of the SRGB and/or SRLB is a matter of future study

2.4. Mapping a SID Index to an MPLS label

This sub-section specifies how the MPLS label value is calculated given the index of a SID. The value of the index is determined by an MCC such as IS-IS [I-D.ietf-isis-segment-routing-extensions] or OSPF [I-D.ietf-ospf-segment-routing-extensions]. This section only specifies how to map the index to an MPLS label. The calculated MPLS

label is downloaded to the FIB, sent out with a forwarded packet, or both.

Consider a SID represented by the index "I". Consider an SRGB as specified in Section 2.3. The total size of the SRGB, represented by the variable "Size", is calculated according to the formula:

size = Lh(1)- Ll(1) + 1 + Lh(2)- Ll(2) + 1 + ... + Lh(k)- Ll(k) + 1

The following rules MUST be applied by the MCC when calculating the MPLS label value corresponding the SID index value "I".

- 0 =< I < size. If the index "I" does not satisfy the previous inequality, then the label cannot be calculated.

- The label value corresponding to the SID index "I" is calculated as follows

  - j = 1 , temp = 0

  - While temp + Lh(j)- Ll(j) < I

    - temp = temp + Lh(j)- Ll(j) + 1

    - j = j+1

  - label = I - temp + Ll(j)

An example for how a router calculates labels and forwards traffic based on the procedure described in this section can be found in Appendix A.1.

2.5. Incoming Label Collision

MPLS Architecture [RFC3031] defines Forwarding Equivalence Class (FEC) term as the set of packets with similar and / or identical characteristics which are forwarded the same way and are bound to the same MPLS incoming (local) label. In Segment-Routing MPLS, local label serves as the SID for given FEC.

We define Segment Routing (SR) FEC as one of the following [RFC8402]:

- (Prefix, Routing Instance, Topology, Algorithm [RFC8402]), where a topology identifies a set of links with metrics. For the purpose of incoming label collision resolution, the same Topology numerical value SHOULD be used on all routers to identify the same set of links with metrics. For MCCs where the "Topology" and/or "Algorithm" fields are not defined, the numerical value of zero MUST be used for these two fields. For the purpose of incoming label collision resolution, a routing instance is identified by a single incoming label downloader to FIB. Two MCCs running on the same router are considered different routing instances if the only way the two instances can know about the other's incoming labels is through redistribution. The numerical value used to identify a routing instance MAY be derived from other configuration or MAY be explicitly configured. If it is derived from other configuration, then the same numerical value SHOULD be derived from the same configuration as long as the configuration survives router reload. If the derived numerical value varies for the same configuration, then an implementation SHOULD make numerical value used to identify a routing instance configurable.

- (next-hop, outgoing interface), where the outgoing interface is physical or virtual.

- (number of adjacencies, list of next-hops, list of outgoing interfaces IDs in ascending numerical order). This FEC represents parallel adjacencies [RFC8402]

- (Endpoint, Color) representing an SR policy [RFC8042]

- (Mirrored SID) The Mirrored SID [RFC8042, Section 5.1] is the IP address advertised by the advertising node to identify the mirror-SID. The IP address is encoded as specified in Section 2.5.1.

This section covers the RECOMMENDED procedure to handle the scenario where, because of an error/misconfiguration, more than one SR FEC as defined in this section, map to the same incoming MPLS label. Examples illustrating the behavior specified in this section can be found in Appendix A.2.

An incoming label collision occurs if the SIDs of the set of FECs {FEC1, FEC2,..., FECk} maps to the same incoming SR MPLS label "L1".

An implementation MUST NOT allow the MCCs belonging to the same router to assign the same incoming label to more than one SR FEC. An implementation that allows such behavior is considered a faulty implementation and is not covered in this document.

The objective of the following steps is to deterministically install in the MPLS Incoming Label Map, also known as label FIB, a single FEC with the incoming label "L1". Remaining FECs may be installed in the IP FIB without incoming label.

The procedure in this section relies completely on the local FEC and label database within a given router.

The collision resolution procedure is as follows

1. Given the SIDs of the set of FECs, {FEC1, FEC2,..., FECk} map to the same MPLS label "L1".

2. Within an MCC, apply tie-breaking rules to select one FEC only and assign the label to it. The losing FECs are handled as if no labels are attached to them. The losing FECs with a non-zero algorithm are not installed in FIB.

   a. If the same set of FECs are attached to the same label "L1", then the tie-breaking rules MUST always select the same FEC irrespective of the order in which the FECs and the label "L1" are received. In other words, the tie-breaking rule MUST be deterministic. For example, a first-come-first-serve tie-breaking is not allowed.

3. If there is still collision between the FECs belonging to different MCCs, then re-apply the tie-breaking rules to the remaining FECs to select one FEC only and assign the label to that FEC

4. Install into the IP FIB the selected FEC and its incoming label in the label FIB.

5. The remaining FECs with the default algorithm (see the specification of prefix-SID algorithm [RFC8402]) are installed in the FIB natively, such as pure IP entries in case of Prefix FEC, without any incoming labels corresponding to their SIDs. The remaining FECs with a non-zero algorithm are not installed in the FIB.

2.5.1. Tie-breaking Rules

The default tie-breaking rules SHOULD be as follows:

1. if FECi has the lowest FEC administrative distance among the competing FECs as defined in this section below, filter away all the competing FECs with higher administrative distance.

2. if more than one competing FEC remains after step 1, select the smallest numerical FEC value

These rules deterministically select the FEC to install in the MPLS forwarding plane for the given incoming label.

This document defines the default tie breaking rules that SHOULD be implemented. An implementation MAY choose to implement additional tie-breaking rules. All routers in a routing domain SHOULD use the same tie-breaking rules to maximize forwarding consistency.

Each FEC is assigned an administrative distance. The FEC administrative distance is encoded as an 8-bit value. The lower the value, the better the administrative distance.

The default FEC administrative distance order starting from the lowest value SHOULD be

- o Explicit SID assignment to a FEC that maps to a label outside the SRGB irrespective of the owner MCC. An explicit SID assignment is a static assignment of a label to a FEC such that the assignment survives router reboot.

    - o An example of explicit SID allocation is static assignment of a specific label to an adj-SID.

    - o An implementation of explicit SID assignment MUST guarantee collision freeness on the same router

- o Dynamic SID assignment:

    - o For all FEC types except for SR policy, the FEC types are ordered using the default administrative distance ordering defined by the implementation.

    - o Binding SID [RFC8402] assigned to SR Policy always has a higher default administrative distance than the default administrative distance of any other FEC type

A user SHOULD ensure that the same administrative distance preference is used on all routers to maximize forwarding consistency.

The numerical sort across FECs SHOULD be performed as follows:

- o Each FEC is assigned a FEC type encoded in 8 bits. The following are the type code point for each SR FEC defined at the beginning of this Section:

- o 120: (Prefix, Routing Instance, Topology, Algorithm)

- o 130: (next-hop, outgoing interface)

- o 140: Parallel Adjacency [RFC8402]

- o 150: an SR policy [RFC8402].

- o 160: Mirror SID [RFC8402]

- o The numerical values above are mentioned to guide implementation. If other numerical values are used, then the numerical values must maintain the same greater-than ordering of the numbers mentioned here.

o The fields of each FEC are encoded as follows

- o Routing Instance ID represented by 16 bits. For routing instances that are identified by less than 16 bits, encode the Instance ID in the least significant bits while the most significant bits are set to zero

- o Address Family represented by 8 bits, where IPv4 encoded as 100 and IPv6 is encoded as 110. These numerical values are mentioned to guide implementations. If other numerical values are used, then the numerical value of IPv4 MUST be less than the numerical value for IPv6

- o All addresses are represented in 128 bits as follows

  - . IPv6 address is encoded natively

  - . IPv4 address is encoded in the most significant bits and the remaining bits are set to zero

- o All prefixes are represented by (128 + 8) bits.

  - . A prefix is encoded in the most significant bits and the remaining bits are set to zero.

  - . The prefix length is encoded before the prefix in a field of size 8 bits.

- o Topology ID is represented by 16 bits. For routing instances that identify topologies using less than 16 bits, encode the topology ID in the least significant bits while the most significant bits are set to zero

  o Algorithm is encoded in a 16 bits field.

  o The Color ID is encoded using 32 bits

- Choose the set of FECs of the smallest FEC type code point

- Out of these FECs, choose the FECs with the smallest address family code point

- Encode the remaining set of FECs as follows

  o Prefix, Routing Instance, Topology, Algorithm: (Prefix Length, Prefix, routing_instance_id, Topology, SR Algorithm,)

  o (next-hop, outgoing interface): (next-hop, outgoing_interface_id)

  o (number of adjacencies, list of next-hops in ascending numerical order, list of outgoing interface IDs in ascending numerical order). This encoding is used to encode a parallel adjacency [RFC8402]

  o (Endpoint, Color): (Endpoint_address, Color_id)

  o (IP address): This is the encoding for a mirror SID FEC. The IP address is encoded as described above in this section

- Select the FEC with the smallest numerical value

The numerical values mentioned in this section are for guidance only. If other numerical values are used then the other numerical values MUST maintain the same numerical ordering among different

2.5.2. Redistribution between Routing Protocol Instances

The following rule SHOULD be applied when redistributing SIDs with prefixes between routing protocol instances:

- If the receiving instance's SRGB is the same as the SRGB of origin instance, then

  o the index is redistributed with the route

- Else

o the index is not redistributed and if needed it is the duty of the receiving instance to allocate a fresh index relative to its own SRGB. Note that in that case, the receiving instance MUST compute its local label according to section 2.4 and install it in FIB.

It is outside the scope of this document to define local node behaviors that would allow to map the original index into a new index in the receiving instance via the addition of an offset or other policy means.

2.5.2.1. Illustration

```
A----IS-IS----B---OSPF----C-192.0.2.1/32 (20001)
```

Consider the simple topology above.

- o A and B are in the IS-IS domain with SRGB [16000-17000]

- o B and C are in OSPF domain with SRGB [20000-21000]

- o B redistributes 192.0.2.1/32 into IS-IS domain

- o In that case A learns 192.0.2.1/32 as an IP leaf connected to B as usual for IP prefix redistribution

- o However, according to the redistribution rule above rule, B decides not to advertise any index with 192.0.2.1/32 into IS-IS because the SRGB is not the same.

2.5.2.2. Illustration 2

Consider the example in the illustration described in Section 2.5.2.1.

When router B redistributes the prefix 192.0.2.1/32, router B decides to allocate and advertise the same index 1 with the prefix 192.0.2.1/32

Within the SRGB of the IS-IS domain, index 1 corresponds to the local label 16001

- o Hence according to the redistribution rule above, router B programs the incoming label 16001 in its FIB to match traffic arriving from the IS-IS domain destined to the prefix 192.0.2.1/32.

2.6. Effect of Incoming Label Collision on Outgoing Label Programming

For the determination of the outgoing label to use, the ingress node pushing new segments, and hence a stack of MPLS labels, MUST use, for a given FEC, the same label that has been selected by the node receiving the packet with that label exposed as top label. So in case of incoming label collision on this receiving node, the ingress node MUST resolve this collision using this same "Incoming Label Collision resolution procedure", using the data of the receiving node.

In the general case, the ingress node may not have exactly the same data of the receiving node, so the result may be different. This is under the responsibility of the network operator. But in typical case, e.g. where a centralized node or a distributed link state IGP is used, all nodes would have the same database. However to minimize the chance of misforwarding, a FEC that loses its incoming label to the tie-breaking rules specified in Section 2.5 MUST NOT be installed in FIB with an outgoing segment routing label based on the SID corresponding to the lost incoming label.

Examples for the behavior specified in this section can be found in Appendix A.3.

2.7. PUSH, CONTINUE, and NEXT

PUSH, NEXT, and CONTINUE are operations applied by the forwarding plane. The specifications of these operations can be found in [RFC8402]. This sub-section specifies how to implement each of these operations in the MPLS forwarding plane.

2.7.1. PUSH

PUSH corresponds to pushing one or more labels on top of an incoming packet then sending it out of a particular physical interface or virtual interface, such as UDP tunnel [RFC7510] or L2TPv3 tunnel [RFC4817], towards a particular next-hop. When pushing labels onto a packet's label stack, the Time-to-Live (TTL) field ([RFC3032], [RFC3443]) and the Traffic Class (TC) field ([RFC3032], [RFC5462]) of each label stack entry must, of course, be set. This document does not specify any set of rules for setting these fields; that is a matter of local policy. Sections 2.10 and 2.11 specify additional details about forwarding behavior.

2.7.2. CONTINUE

In the MPLS forwarding plane, the CONTINUE operation corresponds to swapping the incoming label with an outgoing label. The value of the outgoing label is calculated as specified in Sections 2.10 and 2.11.

2.7.3. NEXT

In the MPLS forwarding plane, NEXT corresponds to popping the topmost label. The action before and/or after the popping depends on the instruction associated with the active SID on the received packet prior to the popping. For example suppose the active SID in the received packet was an Adj-SID [RFC8402], then on receiving the packet, the node applies NEXT operation, which corresponds to popping the top most label, and then sends the packet out of the physical or virtual interface (e.g. UDP tunnel [RFC7510] or L2TPv3 tunnel [RFC4817]) towards the next-hop corresponding to the adj-SID.

2.7.3.1. Mirror SID

If the active SID in the received packet was a Mirror SID [RFC8402, Section 5.1] allocated by the receiving router, then the receiving router applies NEXT operation, which corresponds to popping the top most label, then performs a lookup using the contents of the packet after popping the outer most label in the mirrored forwarding table. The method by which the lookup is made, and/or the actions applied to the packet after the lookup in the mirror table depends on the contents of the packet and the mirror table. Note that the packet exposed after popping the top most label may or may not be an MPLS packet. A mirror SID can be viewed as a generalization of the context label in [RFC5331] because a mirror SID does not make any assumptions about the packet underneath the top label.

2.8. MPLS Label Downloaded to FIB for Global and Local SIDs

The label corresponding to the global SID "Si" represented by the global index "I" downloaded to FIB is used to match packets whose active segment (and hence topmost label) is "Si". The value of this label is calculated as specified in Section 2.4.

For Local SIDs, the MCC is responsible for downloading the correct label value to FIB. For example, an IGP with SR extensions [I-D.ietf-isis-segment-routing-extensions, I-D.ietf-ospf-segment-routing-extensions] allocates and downloads the MPLS label corresponding to an Adj-SID [RFC8402].

2.9. Active Segment

When instantiated in the MPLS domain, the active segment on a packet corresponds to the topmost label on the packet that is calculated according to the procedure specified in Sections 2.10 and 2.11. When arriving at a node, the topmost label corresponding to the active SID matches the MPLS label downloaded to FIB as specified in Section 2.4.

2.10. Forwarding behavior for Global SIDs

This section specifies forwarding behavior, including the calculation of outgoing labels, that corresponds to a global SID when applying PUSH, CONTINUE, and NEXT operations in the MPLS forwarding plane.

This document covers the calculation of the outgoing label for the top label only. The case where the outgoing label is not the top label and is part of a stack of labels that instantiates a routing policy or a traffic engineering tunnel is outside the scope of this document and may be covered in other documents such as [I-D.ietf-spring-segment-routing-policy].

2.10.1. Forwarding for PUSH and CONTINUE of Global SIDs

Suppose an MCC on a router "R0" determines that PUSH or CONTINUE operation is to be applied to an incoming packet related to the global SID "Si" represented by the global index "I" and owned by the router Ri before sending the packet towards a neighbor "N" directly connected to "R0" through a physical or virtual interface such as UDP tunnel [RFC7510] or L2TPv3 tunnel [RFC4817].

The method by which the MCC on router "R0" determines that PUSH or CONTINUE operation must be applied using the SID "Si" is beyond the scope of this document. An example of a method to determine the SID "Si" for PUSH operation is the case where IS-IS [I-D.ietf-isis-segment-routing-extensions] receives the prefix-SID "Si" sub-TLV advertised with prefix "P/m" in TLV 135 and the destination address of the incoming IPv4 packet is covered by the prefix "P/m".

For CONTINUE operation, an example of a method to determine the SID "Si" is the case where IS-IS [I-D.ietf-isis-segment-routing-extensions] receives the prefix-SID "Si" sub-TLV advertised with prefix "P" in TLV 135 and the top label of the incoming packet matches the MPLS label in FIB corresponding to the SID "Si" on the router "R0".

The forwarding behavior for PUSH and CONTINUE corresponding to the SID "Si"

o If the neighbor "N" does not support SR or advertises an invalid SRGB or a SRGB that is too small for the SID "Si"

  o If it is possible to send the packet towards the neighbor "N" using standard MPLS forwarding behavior as specified in [RFC3031] and [RFC3032], then forward the packet. The method by which a router decides whether it is possible to send the packet to "N" or not is beyond the scope of this document. For example, the router "R0" can use the downstream label determined by another MCC, such as LDP [RFC5036], to send the packet.

  o Else if there are other useable next-hops, then use other next-hops to forward the incoming packet. The method by which the router "R0" decides on the possibility of using other next-hops is beyond the scope of this document. For example, the MCC on "R0" may chose the send an IPv4 packet without pushing any label to another next-hop.

  o Otherwise drop the packet.

o Else

  o Calculate the outgoing label as specified in Section 2.4 using the SRGB of the neighbor "N"

  o If the operation is PUSH

    . Push the calculated label according the MPLS label pushing rules specified in [RFC3032]

  o Else

    . swap the incoming label with the calculated label according to the label swapping rules in [RFC3032]

  o Send the packet towards the neighbor "N"

2.10.2. Forwarding for NEXT Operation for Global SIDs

As specified in Section 2.7.3 NEXT operation corresponds to popping the top most label. The forwarding behavior is as follows

o Pop the topmost label

o Apply the instruction associated with the incoming label that has been popped

The action on the packet after popping the topmost label depends on the instruction associated with the incoming label as well as the contents of the packet right underneath the top label that got popped. Examples of NEXT operation are described in Appendix A.1.

2.11. Forwarding Behavior for Local SIDs

This section specifies the forwarding behavior for local SIDs when SR is instantiated over the MPLS forwarding plane.

2.11.1. Forwarding for PUSH Operation on Local SIDs

Suppose an MCC on a router "R0" determines that PUSH operation is to be applied to an incoming packet using the local SID "Si" before sending the packet towards a neighbor "N" directly connected to R0 through a physical or virtual interface such as UDP tunnel [RFC7510] or L2TPv3 tunnel [RFC4817].

An example of such local SID is an Adj-SID allocated and advertised by IS-IS [I-D.ietf-isis-segment-routing-extensions]. The method by which the MCC on "R0" determines that PUSH operation is to be applied to the incoming packet is beyond the scope of this document. An example of such method is backup path used to protect against a failure using TI-LFA [I-D.bashandy-rtgwg-segment-routing-ti-lfa].

As mentioned in [RFC8402], a local SID is specified by an MPLS label. Hence the PUSH operation for a local SID is identical to label push operation [RFC3032] using any MPLS label. The forwarding action after pushing the MPLS label corresponding to the local SID is also determined by the MCC. For example, if the PUSH operation was done to forward a packet over a backup path calculated using TI-LFA, then the forwarding action may be sending the packet to a certain neighbor that will in turn continue to forward the packet along the backup path

2.11.2. Forwarding for CONTINUE Operation for Local SIDs

A local SID on a router "R0" corresponds to a local label. In such scenario, the outgoing label towards a next-hop "N" is determined by the MCC running on the router "R0"and the forwarding behavior for CONTINUE operation is identical to swap operation [RFC3032] on an MPLS label.

2.11.3. Outgoing label for NEXT Operation for Local SIDs

NEXT operation for Local SIDs is identical to NEXT operation for global SIDs specified in Section 2.10.2.

3. IANA Considerations

This document does not make any request to IANA.

4. Manageability Considerations

This document describes the applicability of Segment Routing over the MPLS data plane. Segment Routing does not introduce any change in the MPLS data plane. Manageability considerations described in [RFC8402] applies to the MPLS data plane when used with Segment Routing. SR OAM use cases for the MPLS data plane are defined in [RFC8403]. SR OAM procedures for the MPLS data plane are defined in [RFC8287].

5. Security Considerations

This document does not introduce additional security requirements and mechanisms other than the ones described in [RFC8402].

6. Contributors

The following contributors have substantially helped the definition and editing of the content of this document:

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

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

Jeff Tantsura
Email: jefftant@gmail.com
Edward Crabbe
Email: edward.crabbe@gmail.com

Igor Milojevic
Email: milojevicigor@gmail.com

Saku Ytti
Email: saku@ytti.fi

7. Acknowledgements

The authors would like to thank Les Ginsberg, Chris Bowers, Himanshu Shah, Adrian Farrel, Alexander Vainshtein, Przemyslaw Krol, Darren Dukes, and Zafar Ali for their valuable comments on this document.

This document was prepared using 2-Word-v2.0.template.dot.

8. References

8.1. Normative References

[RFC8402] Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402 July 2018, <http://www.rfc-editor.org/info/rfc8402>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 0.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

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

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <http://www.rfc-editor.org/info/rfc3032>.

[RFC3443] P. Agarwal, P. and Akyol, B. "Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks", RFC 3443, DOI 10.17487/RFC3443, January 2003, <http://www.rfc-editor.org/info/rfc3443>.

[RFC5462] Andersson, L., and Asati, R., " Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <http://www.rfc-editor.org/info/rfc5462>.

[RFC7274] K. Kompella, L. Andersson, and A. Farrel, "Allocating and Retiring Special-Purpose MPLS Labels", RFC7274 DOI 10.17487/RFC7274, May 2014 <http://www.rfc-editor.org/info/rfc7274>

[RFC8174] B. Leiba, " Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", RFC7274 DOI 10.17487/RFC8174, May 2017 <http://www.rfc-editor.org/info/rfc8174>

8.2. Informative References

[I-D.ietf-isis-segment-routing-extensions] Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and j. jefftant@gmail.com, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-13 (work in progress), June 2017.

[I-D.ietf-ospf-ospfv3-segment-routing-extensions] Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPFv3 Extensions for Segment Routing", draft-ietf-ospf-ospfv3-segment-routing-extensions-09 (work in progress), March 2017.

[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-16 (work in progress), May 2017.

[I-D.ietf-spring-segment-routing-ldp-interop] Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., and S. Litkowski, "Segment Routing interworking with LDP", draft-ietf-spring-segment-routing-ldp-interop-08 (work in progress), June 2017.

[I-D.bashandy-rtgwg-segment-routing-ti-lfa], Bashandy, A., Filsfils, C., Decraene, B., Litkowski, S., Francois, P., Voyer, P. Clad, F., and Camarillo, P., "Topology Independent Fast Reroute using Segment Routing", draft-bashandy-rtgwg-segment-routing-ti-lfa-05 (work in progress), October 2018,

[RFC7855] Previdi, S., Ed., Filsfils, C., Ed., Decraene, B., Litkowski, S., Horneffer, M., and R. Shakir, "Source Packet Routing in Networking (SPRING) Problem Statement and Requirements", RFC 7855, DOI 10.17487/RFC7855, May 2016, <http://www.rfc-editor.org/info/rfc7855>.

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

[RFC5331] Aggarwal, R., Rekhter, Y., Rosen, E., " MPLS Upstream Label Assignment and Context-Specific Label Space", RFC5331 DOI 10.17487/RFC5331, August 2008, <http://www.rfc-editor.org/info/rfc5331>.

[RFC7510] Xu, X., Sheth, N., Yong, L., Callon, R., and D. Black, "Encapsulating MPLS in UDP", RFC 7510, DOI 10.17487/RFC7510, April 2015, <https://www.rfc-editor.org/info/rfc7510>.

[RFC4817] Townsley, M., Pignataro, C., Wainner, S., Seely, T., Young, T., "Encapsulation of MPLS over Layer 2 Tunneling Protocol Version 3", RFC4817, DOI 10.17487/RFC4817, March 2007, <https://www.rfc-editor.org/info/rfc4817>

[RFC8287] N. Kumar, C. Pignataro, G. Swallow, N. Akiya, S. Kini, 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" RFC8287, DOI 10.17487/RFC8287, December 2017, https://www.rfc-editor.org/info/rfc8287

[RFC8403] R. Geib, C. Filsfils, C. Pignataro, N. Kumar, "A Scalable and Topology-Aware MPLS Data-Plane Monitoring System", RFC8403, DOI 10.17487/RFC8403, July 2018, <https://www.rfc-editor.org/info/rfc8403>

[I-D.ietf-spring-segment-routing-policy] Filsfils, C., Sivabalan, S., Raza, K., Liste, J. , Clad, F., Voyer, D., Bogdanov, A., Mattes, P., " Segment Routing Policy for Traffic Engineering", draft-ietf-spring-segment-routing-policy-01 (work in progress), June 2018

9. Authors' Addresses

Ahmed Bashandy (editor)
Arrcus

Email: abashandy.ietf@gmail.com

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

Email: cfilsfil@cisco.com

Stefano Previdi
Cisco Systems, Inc.
Italy

Email: stefano@previdi.net

Bruno Decraene
Orange
FR

Email: bruno.decraene@orange.com

Stephane Litkowski
Orange
FR

Email: stephane.litkowski@orange.com

Rob Shakir
Google
US

Email: robjs@google.com

Appendix A. Examples

A.1. IGP Segments Example

Consider the network diagram of Figure 1 and the IP address and IGP Segment allocation of Figure 2. Assume that the network is running IS-IS with SR extensions [I-D.ietf-isis-segment-routing-extensions] and all links have the same metric. The following examples can be constructed.

```
                   +--------+
                  /          \
  R0-----R1-----R2----------R3-----R8
                 | \        / |
                 |  +--R4--+  |
                 |            |
                 +-----R5-----+
```

Figure 1: IGP Segments - Illustration

```
      +-------------------------------------------------------------+
      | IP address allocated by the operator:                       |
      |                     192.0.2.1/32 as a loopback of R1        |
      |                     192.0.2.2/32 as a loopback of R2        |
      |                     192.0.2.3/32 as a loopback of R3        |
      |                     192.0.2.4/32 as a loopback of R4        |
      |                     192.0.2.5/32 as a loopback of R5        |
      |                     192.0.2.8/32 as a loopback of R8        |
      |             198.51.100.9/32 as an anycast loopback of R4    |
      |             198.51.100.9/32 as an anycast loopback of R5    |
      |                                                             |
      | SRGB defined by the operator as 1000-5000                   |
      |                                                             |
      | Global IGP SID indices allocated by the operator:           |
      |                     1 allocated to 192.0.2.1/32             |
      |                     2 allocated to 192.0.2.2/32             |
      |                     3 allocated to 192.0.2.3/32             |
      |                     4 allocated to 192.0.2.4/32             |
      |                     8 allocated to 192.0.2.8/32             |
      |                  1009 allocated to 198.51.100.9/32          |
      |                                                             |
      | Local IGP SID allocated dynamically by R2                   |
      |                    for its "north" adjacency to R3: 9001    |
      |                    for its "north" adjacency to R3: 9003    |
      |                    for its "south" adjacency to R3: 9002    |
      |                    for its "south" adjacency to R3: 9003    |
      +-------------------------------------------------------------+
```

Figure 2: IGP Address and Segment Allocation - Illustration

Suppose R1 wants to send an IPv4 packet P1 to R8. In this case, R1 needs to apply PUSH operation to the IPv4 packet.

Remember that the SID index "8" is a global IGP segment attached to the IP prefix 192.0.2.8/32. Its semantic is global within the IGP domain: any router forwards a packet received with active segment 8 to the next-hop along the ECMP-aware shortest-path to the related prefix.

R2 is the next-hop along the shortest path towards R8. By applying the steps in Section 2.8 the outgoing label downloaded to R1's FIB corresponding to the global SID index 8 is 1008 because the SRGB of R2 is [1000,5000] as shown in Figure 2.

Because the packet is IPv4, R1 applies the PUSH operation using the label value 1008 as specified in Section 2.10.1. The resulting MPLS

header will have the "S" bit [RFC3032] set because it is followed directly by an IPv4 packet.

The packet arrives at router R2. Because the top label 1008 corresponds to the IGP SID "8", which is the prefix-SID attached to the prefix 192.0.2.8/32 owned by the node R8, then the instruction associated with the SID is "forward the packet using all ECMP/UCMP interfaces and all ECMP/UCMP next-hop(s) along the shortest/useable path(s) towards R8". Because R2 is not the penultimate hop, R2 applies the CONTINUE operation to the packet and sends it to R3 using one of the two links connected to R3 with top label 1008 as specified in Section 2.10.1.

R3 receives the packet with top label 1008. Because the top label 1008 corresponds to the IGP SID "8", which is the prefix-SID attached to the prefix 192.0.2.8/32 owned by the node R8, then the instruction associated with the SID is "send the packet using all ECMP interfaces and all next-hop(s) along the shortest path towards R8". Because R3 is the penultimate hop, we assume that R3 performs penumtimate hop popping, which corresponds to the NEXT operation, then sends the packet to R8. The NEXT operation results in popping the outer label and sending the packet as a pure IPv4 packet to R8.

In conclusion, the path followed by P1 is R1-R2--R3-R8. The ECMP-awareness ensures that the traffic be load-shared between any ECMP path, in this case the two links between R2 and R3.

A.2. Incoming Label Collision Examples

This section describes few examples to illustrate the handling of label collision described in Section 2.5.

For the examples in this section, we assume that Node A has the following:

o OSPF default admin distance for implementation=50

o ISIS default admin distance for implementation=60

A.2.1. Example 1

Illustration of incoming label collision resolution for the same FEC type using MCC administrative distance.

FEC1:

o OSPF prefix SID advertisement from node B for 198.51.100.5/32 with index=5

o OSPF SRGB on node A = [1000,1999]

o Incoming label=1005

FEC2:

o ISIS prefix SID advertisement from node C for 203.0.113.105/32 with index=5

o ISIS SRGB on node A = [1000,1999]

o Incoming label=1005

FEC1 and FEC2 both use dynamic SID assignment. Since neither ofthe FEC types is SR Policy, we use the default admin distances of 50 and 60 to break the tie. So FEC1 wins.

A.2.2. Example 2

Illustration of incoming label collision resolution for different FEC types using the MCC administrative distance.

FEC1:

o Node A receives an OSPF prefix sid advertisement from node B for 198.51.100.6/32 with index=6

o OSPF SRGB on node A = [1000,1999]

o Hence the incoming label on node A corresponding to 198.51.100.6/32 is 1006

FEC2:

ISIS on node A assigns the label 1006 to the globally significant adj-SID (I.e. when advertised the "L" flag is clear in the adj-SID sub-TLV as described in [I-D.ietf-isis-segment-routing-extensions]) towards one of its neighbors. Hence the incoming label corresponding to this adj-SID 1006. Assume Node A allocates this adj-SID dynamically, and it may differ across router reboots.

FEC1 and FEC2 both use dynamic SID assignment. Since neither of the FEC types is SR Policy, we use the default admin distances of 50 and 60 to break the tie. So FEC1 wins.

A.2.3. Example 3

Illustration of incoming label collision resolution based on preferring static over dynamic SID assignment

FEC1:
OSPF on node A receives a prefix SID advertisement from node B for 198.51.100.7/32 with index=7. Assuming that the OSPF SRGB on node A is [1000,1999], then incoming label corresponding to 198.51.100.7/32 is 1007

FEC2:
The operator on node A configures ISIS on node A to assign the label 1007 to the globally significant adj-SID (I.e. when advertised the "L" flag is clear in the adj-SID sub-TLV as described in [I-D.ietf-isis-segment-routing-extensions]) towards one of its neighbor advertisement from node A with label=1007

Node A assigns this adj-SID explicitly via configuration, so the adj-SID survives router reboots.

FEC1 uses dynamic SID assignment, while FEC2 uses explicit SID assignment. So FEC2 wins.

A.2.4. Example 4

Illustration of incoming label collision resolution using FEC type default administrative distance

FEC1:
OSPF on node A receives a prefix SID advertisement from node B for 198.51.100.8/32 with index=8. Assuming that OSPF SRGB on node A = [1000,1999], the incoming label corresponding to 198.51.100.8/32 is 1008.

FEC2:
Suppose the SR Policy advertisement from controller to node A for the policy identified by (Endpoint = 192.0.2.208, color = 100) and

consisting of SID-List = <S1, S2> assigns the globally significant Binding-SID label 1008

From the point of view of node A, FEC1 and FEC2 both use dynamic SID assignment. Based on the default administrative distance outlined in Section 2.5.1, the binding SID has a higher administrative distance than the prefix-SID and hence FEC1 wins.

A.2.5. Example 5

Illustration of incoming label collision resolution based on FEC type preference

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.110/32 with index=10. Assuming that the ISIS SRGB on node A is [1000,1999], then incoming label corresponding to 203.0.113.110/32 is 1010.

FEC2:
ISIS on node A assigns the label 1010 to the globally significant adj-SID (I.e. when advertised the "L" flag is clear in the adj-SID sub-TLV as described in [I-D.ietf-isis-segment-routing-extensions]) towards one of its neighbors).

Node A allocates this adj-SID dynamically, and it may differ across router reboots. Hence both FEC1 and FEC2 both use dynamic SID assignment.

Since both FECs are from the same MCC, they have the same default admin distance. So we compare FEC type code-point. FEC1 has FEC type code-point=120, while FEC2 has FEC type code-point=130. Therefore, FEC1 wins.

A.2.6. Example 6

Illustration of incoming label collision resolution based on address family preference.

FEC1:
ISIS on node A receives prefix SID advertisement from node B for 203.0.113.111/32 with index 11. Assuming that the ISIS SRGB on node A is [1000,1999], the incoming label on node A for 203.0.113.111/32 is 1011.

FEC2:
ISIS on node A prefix SID advertisement from node C for 2001:DB8:1000::11/128 with index=11. Assuming that the ISIS SRGB on node A is [1000,1999], the incoming label on node A for 2001:DB8:1000::11/128 is 1011

FEC1 and FEC2 both use dynamic SID assignment. Since both FECs are from the same MCC, they have the same default admin distance. So we compare FEC type code-point. Both FECs have FEC type code-point=120. So we compare address family. Since IPv4 is preferred over IPv6, FEC1 wins.

A.2.7. Example 7

Illustration incoming label collision resolution based on prefix length.

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.112/32 with index 12. Assuming that ISIS SRGB on node A is [1000,1999], the incoming label for 203.0.113.112/32 on node A is 1012.

FEC2:
ISIS on node A receives a prefix SID advertisement from node C for 203.0.113.128/30 with index 12. Assuming that the ISIS SRGB on node A is [1000,1999], then incoming label for 203.0.113.128/30 on node A is 1012

FEC1 and FEC2 both use dynamic SID assignment. Since both FECs are from the same MCC, they have the same default admin distance. So we compare FEC type code-point. Both FECs have FEC type code-point=120. So we compare address family. Both are IPv4 address family, so we compare prefix length. FEC1 has prefix length=32, and FEC2 has prefix length=30, so FEC2 wins.

A.2.8. Example 8

Illustration of incoming label collision resolution based on the numerical value of the FECs.

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.113/32 with index 13. Assuming that ISIS SRGB on node A is

[1000,1999], then the incoming label for 203.0.113.113/32 on node A is 1013

FEC2:
ISIS on node A receives a prefix SID advertisement from node C for 203.0.113.213/32 with index 13. Assuming that ISIS SRGB on node A is [1000,1999], then the incoming label for 203.0.113.213/32 on node A is 1013

FEC1 and FEC2 both use dynamic SID assignment. Since both FECs are from the same MCC, they have the same default admin distance. So we compare FEC type code-point. Both FECs have FEC type code-point=120. So we compare address family. Both are IPv4 address family, so we compare prefix length. Prefix lengths are the same, so we compare prefix. FEC1 has the lower prefix, so FEC1 wins.

A.2.9. Example 9

Illustration of incoming label collision resolution based on routing instance ID.

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.114/32 with index 14. Assume that this ISIS instance on node A has the Routing Instance ID 1000 and SRGB [1000,1999]. Hence the incoming label for 203.0.113.114/32 on node A is 1014

FEC2:
ISIS on node A receives a prefix SID advertisement from node C for 203.0.113.114/32 with index=14. Assume that this is another instance of ISIS on node A with a different routing Instance ID 2000 but the same SRGB [1000,1999]. Hence incoming label for 203.0.113.114/32 on node A 1014

These two FECs match all the way through the prefix length and prefix. So Routing Instance ID breaks the tie, with FEC1 winning.

A.2.10. Example 10

Illustration of incoming label collision resolution based on topology ID.

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.115/32 with index=15. Assume that this ISIS instance on

node A has Routing Instance ID 1000. Assume that the prefix advertisement of 203.0.113.115/32 was received in ISIS Multi-topology advertisement with ID = 50. If the ISIS SRGB for this routing instance on node A is [1000,1999], then incoming label of 203.0.113.115/32 for topology 50 on node A is 1015

FEC2:
ISIS on node A receives a prefix SID advertisement from node C for 203.0.113.115/32 with index 15. Assume that it is the same routing Instance ID = 1000 but 203.0.113.115/32 was advertised with a different ISIS Multi-topology ID = 40. If the ISIS SRGB on node A is [1000,1999], then incoming label of 203.0.113.115/32 for topology 40 on node A is also 1015

These two FECs match all the way through the prefix length, prefix, and Routing Instance ID. We compare ISIS Multi-topology ID, so FEC2 wins.

A.2.11. Example 11

Illustration of incoming label collision for resolution based on algorithm ID.

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.116/32 with index=16 Assume that ISIS on node A has Routing Instance ID = 1000. Assume that node B advertised 203.0.113.116/32 with ISIS Multi-topology ID = 50 and SR algorithm = 0. Assume that the ISIS SRGB on node A = [1000,1999]. Hence the incoming label corresponding to this advertisement of 203.0.113.116/32 is 1016.

FEC2:
ISIS on node A receives a prefix SID advertisement from node C for 203.0.113.116/32 with index=16. Assume that it is the same ISIS instance on node A with Routing Instance ID = 1000. Also assume that node C advertised 203.0.113.116/32 with ISIS Multi-topology ID = 50 but with SR algorithm = 22. Since it is the same routing instance, the SRGB on node A = [1000,1999]. Hence the incoming label corresponding to this advertisement of 203.0.113.116/32 by node C is also 1016.

These two FECs match all the way through the prefix length, prefix, and Routing Instance ID, and Multi-topology ID. We compare SR algorithm ID, so FEC1 wins.

A.2.12. Example 12

Illustration of incoming label collision resolution based on FEC numerical value and independent of how the SID assigned to the colliding FECs.

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.117/32 with index 17. Assume that the ISIS SRGB on node A is [1000,1999], then the incoming label is 1017

FEC2:
Suppose there is an ISIS mapping server advertisement (SID/Label Binding TLV) from node D has Range 100 and Prefix = 203.0.113.1/32. Suppose this mapping server advertisement generates 100 mappings, one of which maps 203.0.113.17/32 to index 17. Assuming that it is the same ISIS instance, then the SRGB is [1000,1999] and hence the incoming label for 1017.

The fact that FEC1 comes from a normal prefix SID advertisement and FEC2 is generated from a mapping server advertisement is not used as a tie-breaking parameter. Both FECs use dynamic SID assignment, are from the same MCC, have the same FEC type code-point=120. Their prefix lengths are the same as well. FEC2 wins based on lower numerical prefix value, since 203.0.113.17 is less than 203.0.113.117.

A.2.13. Example 13

Illustration of incoming label collision resolution based on address family preference

FEC1:
SR Policy advertisement from controller to node A. Endpoint address=2001:DB8:3000::100, color=100, SID-List=<S1, S2> and the Binding-SID label=1020

FEC2:
SR Policy advertisement from controller to node A. Endpoint address=192.0.2.60, color=100, SID-List=<S3, S4> and the Binding-SID label=1020

The FECs match through the tie-breaks up to and including having the same FEC type code-point=140. FEC2 wins based on IPv4 address family being preferred over IPv6.

A.2.14. Example 14

Illustration of incoming label resolution based on numerical value of the policy endpoint.

FEC1:
SR Policy advertisement from controller to node A. Endpoint address=192.0.2.70, color=100, SID-List=<S1, S2> and Binding-SID label=1021

FEC2:
SR Policy advertisement from controller to node A Endpoint address=192.0.2.71, color=100, SID-List=<S3, S4> and Binding-SID label=1021

The FECs match through the tie-breaks up to and including having the same address family. FEC1 wins by having the lower numerical endpoint address value.

A.3. Examples for the Effect of Incoming Label Collision on Outgoing Label

This section presents examples to illustrate the effect of incoming label collision on the selection of the outgoing label described in Section 2.6.

A.3.1. Example 1

Illustration of the effect of incoming label resolution on the outgoing label

FEC1:
ISIS on node A receives a prefix SID advertisement from node B for 203.0.113.122/32 with index 22. Assuming that the ISIS SRGB on node A is [1000,1999] the corresponding incoming label is 1022.

FEC2:
ISIS on node A receives a prefix SID advertisement from node C for 203.0.113.222/32 with index=22 Assuming that the ISIS SRGB on node A is [1000,1999] the corresponding incoming label is 1022.

FEC1 wins based on lowest numerical prefix value. This means that node A installs a transit MPLS forwarding entry to SWAP incoming label 1022, with outgoing label N and use outgoing interface I. N is determined by the index associated with FEC1 (index 22) and the SRGB advertised by the next-hop node on the shortest path to reach 203.0.113.122/32.

Node A will generally also install an imposition MPLS forwarding entry corresponding to FEC1 for incoming prefix=203.0.113.122/32 pushing outgoing label N, and using outgoing interface I.

The rule in Section 2.6 means node A MUST NOT install an ingress MPLS forwarding entry corresponding to FEC2 (the losing FEC, which would be for prefix 203.0.113.222/32).

A.3.2. Example 2

Illustration of the effect of incoming label collision resolution on outgoing label programming on node A

FEC1:

- o SR Policy advertisement from controller to node A
- o Endpoint address=192.0.2.80, color=100, SID-List=<S1, S2>
- o Binding-SID label=1023

FEC2:

- o SR Policy advertisement from controller to node A
- o Endpoint address=192.0.2.81, color=100, SID-List=<S3, S4>
- o Binding-SID label=1023

FEC1 wins by having the lower numerical endpoint address value. This means that node A installs a transit MPLS forwarding entry to SWAP incoming label=1023, with outgoing labels and outgoing interface determined by the SID-List for FEC1.

In this example, we assume that node A receives two BGP/VPN routes:

- o R1 with VPN label=V1, BGP next-hop = 192.0.2.80, and color=100,
- o R2 with VPN label=V2, BGP next-hop = 192.0.2.81, and color=100,

We also assume that A has a BGP policy which matches on color=100 that allows that its usage as SLA steering information. In this case, node A will install a VPN route with label stack = <S1,S2,V1> (corresponding to FEC1).

The rule described in section 2.6 means that node A MUST NOT install a VPN route with label stack = <S3,S4,V1> (corresponding to FEC2.)