IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: June 22, 2018

S. Previdi, Ed.
L. Ginsberg, Ed.
C. Filsfils
A. Bashandy
Cisco Systems, Inc.
H. Gredler
RtBrick Inc.
S. Litkowski
B. Decraene
Orange
J. Tantsura
Individual

December 19, 2017

IS-IS Extensions for Segment Routing
draft-ietf-isis-segment-routing-extensions-15

Abstract

Segment Routing (SR) allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF).

This draft describes the necessary IS-IS extensions that need to be introduced for Segment Routing operating on an MPLS data-plane.

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 June 22, 2018.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Segment Routing Identifiers 3
2.1. Prefix Segment Identifier (Prefix-SID Sub-TLV) 3
2.1.1. Flags 6
2.1.2. Prefix-SID Propagation 8
2.2. Adjacency Segment Identifier 8
2.2.1. Adjacency Segment Identifier (Adj-SID) Sub-TLV . . . 9
2.2.2. Adjacency Segment Identifiers in LANs 11
2.3. SID/Label Sub-TLV 12
2.4. SID/Label Binding TLV 13
2.4.1. Flags 14
2.4.2. Range 15
2.4.3. Prefix Length, Prefix 17
2.4.4. Mapping Server Prefix-SID 17
2.4.5. SID/Label Sub-TLV 18
2.5. Multi-Topology SID/Label Binding TLV 18
3. Router Capabilities 19
3.1. SR-Capabilities Sub-TLV 19
3.2. SR-Algorithm Sub-TLV 22
3.3. SR Local Block Sub-TLV 23
3.4. SRMS Preference Sub-TLV 25
4. Non backward compatible changes with prior versions of this document 25
4.1. Encoding of Multiple SRGBs 25
5. IANA Considerations 26
5.1. Sub TLVs for Type 22,23,25,141,222, and 223 26
5.2. Sub TLVs for Type 135,235,236 and 237 27
5.3. Sub TLVs for Type 242 27
5.4. New TLV Codepoint and Sub-TLV registry 28

6. Security Considerations 29
7. Acknowledgements 29
8. Contributors 29
9. References 31
9.1. Normative References 31
9.2. Informative References 32
Authors' Addresses 33

1. Introduction

Segment Routing (SR) allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF). Two types of segments are defined, Prefix segments and Adjacency segments. Prefix segments represent an ecmp-aware shortest-path to a prefix, as per the state of the IGP topology. Adjacency segments represent a hop over a specific adjacency between two nodes in the IGP. A prefix segment is typically a multi-hop path while an adjacency segment, in most of the cases, is a one-hop path. SR's control-plane can be applied to both IPv6 and MPLS data-planes, and do not require any additional signaling (other than the regular IGP). For example, when used in MPLS networks, SR paths do not require any LDP or RSVP-TE signaling. Still, SR can interoperate in the presence of LSPs established with RSVP or LDP.

This draft describes the necessary IS-IS extensions that need to be introduced for Segment Routing operating on an MPLS data-plane.

Segment Routing architecture is described in [I-D.ietf-spring-segment-routing].

Segment Routing use cases are described in [RFC7855].

2. Segment Routing Identifiers

Segment Routing architecture ([I-D.ietf-spring-segment-routing]) defines different types of Segment Identifiers (SID). This document defines the IS-IS encodings for the IGP-Prefix-SID, the IGP-Adjacency-SID, the IGP-LAN-Adjacency-SID and the Binding-SID.

2.1. Prefix Segment Identifier (Prefix-SID Sub-TLV)

A new IS-IS sub-TLV is defined: the Prefix Segment Identifier sub-TLV (Prefix-SID sub-TLV).

The Prefix-SID sub-TLV carries the Segment Routing IGP-Prefix-SID as defined in [I-D.ietf-spring-segment-routing]. The 'Prefix SID' MUST

be unique within a given IGP domain (when the L-flag is not set). The 'Prefix SID' MUST carry an index (when the V-flag is not set) that determines the actual SID/label value inside the set of all advertised SID/label ranges of a given router. A receiving router uses the index to determine the actual SID/label value in order to construct forwarding state to a particular destination router.

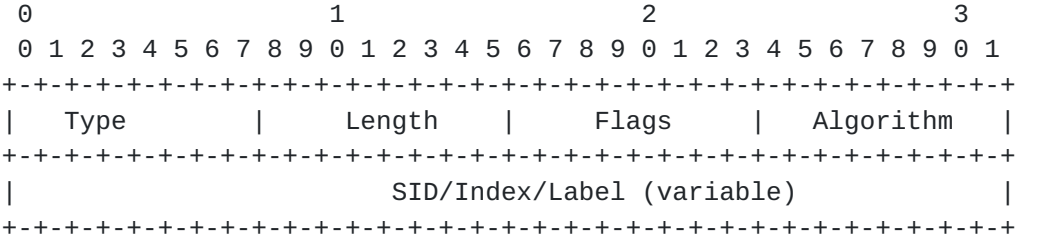
In many use-cases a 'stable transport' IP Address is overloaded as an identifier of a given node. Because the IP Prefixes may be re-advertised into other levels there may be some ambiguity (e.g. Originating router vs. L1L2 router) for which node a particular IP prefix serves as identifier. The Prefix-SID sub-TLV contains the necessary flags to disambiguate IP Prefix to node mappings. Furthermore if a given node has several 'stable transport' IP addresses there are flags to differentiate those among other IP Prefixes advertised from a given node.

A Prefix-SID sub-TLV is associated to a prefix advertised by a node and MAY be present in any of the following TLVs:

- TLV-135 (Extended IPv4 reachability) defined in [RFC5305].
- TLV-235 (Multitopology IPv4 Reachability) defined in [RFC5120].
- TLV-236 (IPv6 IP Reachability) defined in [RFC5308].
- TLV-237 (Multitopology IPv6 IP Reachability) defined in [RFC5120].
- Binding-TLV defined in Section 2.4.

When the IP Reachability TLV is propagated across level boundaries, the Prefix-SID sub-TLV SHOULD be kept.

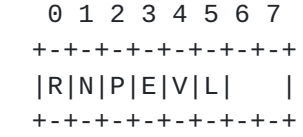
The Prefix-SID sub-TLV has the following format:



where:

- Type: TBD, suggested value 3
- Length: variable.

Flags: 1 octet field of following flags:



where:

R-Flag: Re-advertisement flag. If set, then the prefix to which this Prefix-SID is attached, has been propagated by the router either from another level (i.e., from level-1 to level-2 or the opposite) or from redistribution (e.g.: from another protocol).

N-Flag: Node-SID flag. If set, then the Prefix-SID refers to the router identified by the prefix. Typically, the N-Flag is set on Prefix-SIDs attached to a router loopback address. The N-Flag is set when the Prefix-SID is a Node-SID as described in [I-D.ietf-spring-segment-routing].

P-Flag: no-PHP flag. If set, then the penultimate hop MUST NOT pop the Prefix-SID before delivering the packet to the node that advertised the Prefix-SID.

E-Flag: Explicit-Null Flag. If set, any upstream neighbor of the Prefix-SID originator MUST replace the Prefix-SID with a Prefix-SID having an Explicit-NULL value (0 for IPv4 and 2 for IPv6) before forwarding the packet.

V-Flag: Value flag. If set, then the Prefix-SID carries a value (instead of an index). By default the flag is UNSET.

L-Flag: Local Flag. If set, then the value/index carried by the Prefix-SID has local significance. By default the flag is UNSET.

Other bits: MUST be zero when originated and ignored when received.

Algorithm: the router may use various algorithms when calculating reachability to other nodes or to prefixes attached to these nodes. Algorithms identifiers are defined in Section 3.2. Examples of these algorithms are metric based Shortest Path First (SPF), various sorts of Constrained SPF, etc. The algorithm field of the Prefix-SID contains the identifier of the algorithm the router has used in order to compute the reachability of the prefix the Prefix-SID is associated to.

At origination, the Prefix-SID algorithm field MUST be set to 0 or to any value advertised in the SR-Algorithm sub-TLV (Section 3.2).

A router receiving a Prefix-SID from a remote node and with an algorithm value that such remote node has not advertised in the SR-Algorithm sub-TLV (Section 3.2) MUST ignore the Prefix-SID sub-TLV.

SID/Index/Label: according to the V and L flags, it contains either:

* A 4 octet index defining the offset in the SID/Label space advertised by this router using the encodings defined in Section 3.1. In this case the V and L flags MUST be unset.

* A 3 octet local label where the 20 rightmost bits are used for encoding the label value. In this case the V and L flags MUST be set.

2.1.1. Flags

2.1.1.1. R and N Flags

The R-Flag MUST be set for prefixes that are not local to the router and either:

advertised because of propagation (Level-1 into Level-2);

advertised because of leaking (Level-2 into Level-1);

advertised because of redistribution (e.g.: from another protocol).

In the case where a Level-1-2 router has local interface addresses configured in one level, it may also propagate these addresses into the other level. In such case, the Level-1-2 router MUST NOT set the R bit. The R-bit MUST be set only for prefixes that are not local to the router and advertised by the router because of propagation and/or leaking.

The N-Flag is used in order to define a Node-SID. A router MAY set the N-Flag only if all of the following conditions are met:

The prefix to which the Prefix-SID is attached is local to the router (i.e., the prefix is configured on one of the local interfaces, e.g., a 'stable transport' loopback).

The prefix to which the Prefix-SID is attached MUST have a Prefix length of either /32 (IPv4) or /128 (IPv6).

The router MUST ignore the N-Flag on a received Prefix-SID if the prefix has a Prefix length different than /32 (IPv4) or /128 (IPv6).

[RFC7794] also defines the N and R flags and with the same semantics of the equivalent flags defined in this document. There will be a transition period where both sets of flags will be used and eventually only the flags of the Prefix Attributes will remain. During the transition period implementations supporting the N and R flags defined in this document and the N and R flags defined in [RFC7794] MUST advertise and parse all flags. In case the received flags have different values, the value of the flags defined in [RFC7794] prevails.

2.1.1.2. E and P Flags

When calculating the outgoing label for the prefix, the router MUST take into account E and P flags advertised by the next-hop router, if next-hop router advertised the SID for the prefix. This MUST be done regardless of next-hop router contributing to the best path to the prefix or not.

When propagating (either from Level-1 to Level-2 or vice versa) a reachability advertisement originated by another IS-IS speaker, the router MUST set the P-flag and MUST clear the E-flag of the related Prefix-SIDs.

The following behavior is associated with the settings of the E and P flags:

- If the P-flag is not set then any upstream neighbor of the Prefix-SID originator MUST pop the Prefix-SID. This is equivalent to the penultimate hop popping mechanism used in the MPLS dataplane which improves performance of the ultimate hop. MPLS EXP bits of the Prefix-SID are not preserved to the ultimate hop (the Prefix-SID being removed). If the P-flag is unset the received E-flag is ignored.

- If the P-flag is set then:

  * If the E-flag is not set then any upstream neighbor of the Prefix-SID originator MUST keep the Prefix-SID on top of the stack. This is useful when, e.g., the originator of the Prefix-SID must stitch the incoming packet into a continuing MPLS LSP to the final destination. This could occur at an inter-area border router (prefix propagation from one area to

another) or at an inter-domain border router (prefix propagation from one domain to another).

* If the E-flag is set then any upstream neighbor of the Prefix-SID originator MUST replace the PrefixSID with a Prefix-SID having an Explicit-NULL value. This is useful, e.g., when the originator of the Prefix-SID is the final destination for the related prefix and the originator wishes to receive the packet with the original EXP bits.

2.1.2. Prefix-SID Propagation

The Prefix-SID sub-TLV MUST be preserved when the IP Reachability TLV gets propagated across level boundaries.

The level-1-2 router that propagates the Prefix-SID sub-TLV between levels MUST set the R-flag.

If the Prefix-SID contains a global index (L and V flags unset) and it is propagated as such (with L and V flags unset), the value of the index MUST be preserved when propagated between levels.

The level-1-2 router that propagates the Prefix-SID sub-TLV between levels MAY change the setting of the L and V flags in case a local label value is encoded in the Prefix-SID instead of the received value.

2.2. Adjacency Segment Identifier

A new IS-IS sub-TLV is defined: the Adjacency Segment Identifier sub-TLV (Adj-SID sub-TLV).

The Adj-SID sub-TLV is an optional sub-TLV carrying the Segment Routing IGP-Adjacency-SID as defined in [I-D.ietf-spring-segment-routing] with flags and fields that may be used, in future extensions of Segment Routing, for carrying other types of SIDs.

IS-IS adjacencies are advertised using one of the IS-Neighbor TLVs below:

TLV-22 (Extended IS reachability)[RFC5305]

TLV-222 (Multitopology IS)[RFC5120]

TLV-23 (IS Neighbor Attribute)[RFC5311]

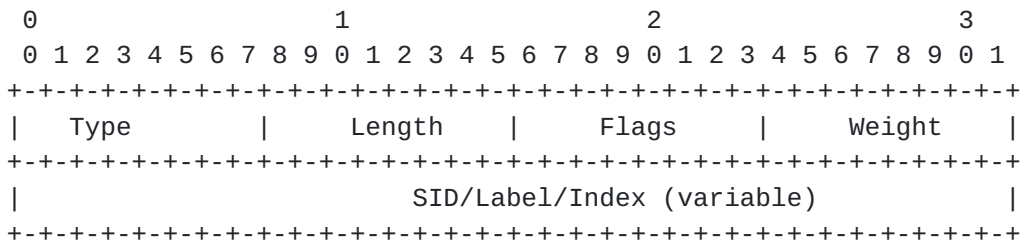
TLV-223 (Multitopology IS Neighbor Attribute)[RFC5311]

TLV-141 (inter-AS reachability information)[RFC5316]

Multiple Adj-SID sub-TLVs MAY be associated with a single IS-neighbor.

2.2.1. Adjacency Segment Identifier (Adj-SID) Sub-TLV

The following format is defined for the Adj-SID sub-TLV:

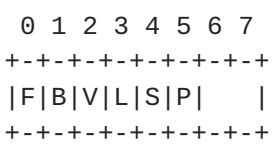


where:

Type: TBD, suggested value 31

Length: variable.

Flags: 1 octet field of following flags:



where:

F-Flag: Address-Family flag. If unset, then the Adj-SID refers to an adjacency with outgoing IPv4 encapsulation. If set then the Adj-SID refers to an adjacency with outgoing IPv6 encapsulation.

B-Flag: Backup flag. If set, the Adj-SID is eligible for protection (e.g.: using IPFRR or MPLS-FRR) as described in [I-D.ietf-spring-resiliency-use-cases].

V-Flag: Value flag. If set, then the Adj-SID carries a value. By default the flag is SET.

L-Flag: Local Flag. If set, then the value/index carried by the Adj-SID has local significance. By default the flag is SET.

S-Flag. Set flag. When set, the S-Flag indicates that the Adj-SID refers to a set of adjacencies (and therefore MAY be assigned to other adjacencies as well).

P-Flag. Persistent flag. When set, the P-Flag indicates that the Adj-SID is persistently allocated, i.e., the Adj-SID value remains consistent across router restart and/or interface flap.

Other bits: MUST be zero when originated and ignored when received.

Weight: 1 octet. The value represents the weight of the Adj-SID for the purpose of load balancing. The use of the weight is defined in [I-D.ietf-spring-segment-routing].

SID/Index/Label: according to the V and L flags, it contains either:

* A 3 octet local label where the 20 rightmost bits are used for encoding the label value. In this case the V and L flags MUST be set.

* A 4 octet index defining the offset in the SID/Label space advertised by this router using the encodings defined in Section 3.1. In this case V and L flags MUST be unset.

An SR capable router MAY allocate an Adj-SID for each of its adjacencies and SHOULD set the B-Flag when the adjacency is eligible for protection (IP or MPLS).

An SR capable router MAY allocate more than one Adj-SID to an adjacency.

An SR capable router MAY allocate the same Adj-SID to different adjacencies.

When the P-flag is not set, the Adj-SID MAY be persistent. When the P-flag is set, the Adj-SID MUST be persistent.

Examples of use of the Adj-SID sub-TLV are described in [I-D.ietf-spring-segment-routing].

The F-flag is used in order for the router to advertise the outgoing encapsulation of the adjacency the Adj-SID is attached to.

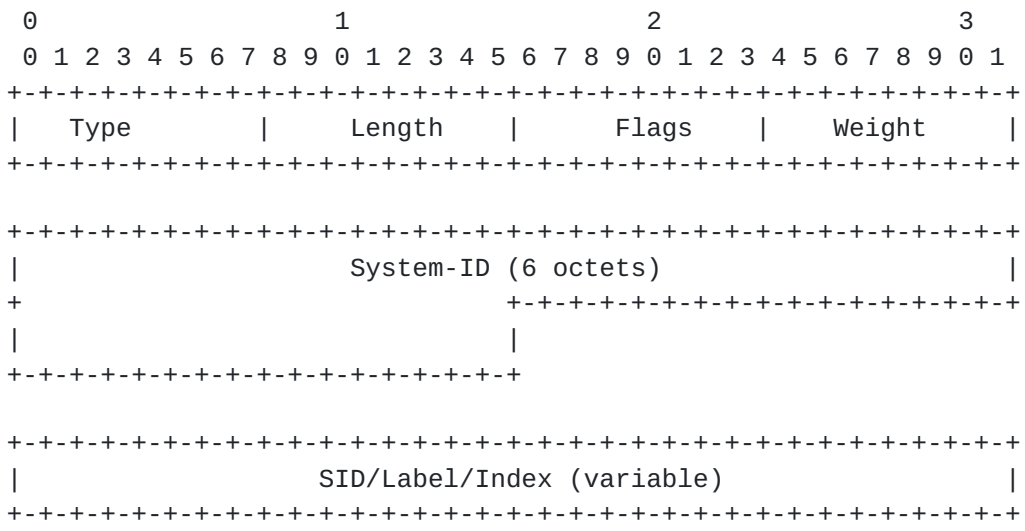
2.2.2. Adjacency Segment Identifiers in LANs

In LAN subnetworks, the Designated Intermediate System (DIS) is elected and originates the Pseudonode-LSP (PN-LSP) including all neighbors of the DIS.

When Segment Routing is used, each router in the LAN MAY advertise the Adj-SID of each of its neighbors. Since, on LANs, each router only advertises one adjacency to the DIS (and doesn't advertise any other adjacency), each router advertises the set of Adj-SIDs (for each of its neighbors) inside a newly defined sub-TLV part of the TLV advertising the adjacency to the DIS (e.g.: TLV-22).

The following new sub-TLV is defined: LAN-Adj-SID (Type: TBD, suggested value 32) containing the set of Adj-SIDs the router assigned to each of its LAN neighbors.

The format of the LAN-Adj-SID sub-TLV is as follows:

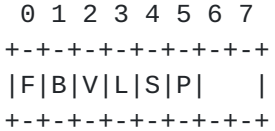


where:

Type: TBD, suggested value 32

Length: variable.

Flags: 1 octet field of following flags:



where F, B, V, L, S and P flags are defined in Section 2.2.1. Other bits: MUST be zero when originated and ignored when received.

Weight: 1 octet. The value represents the weight of the Adj-SID for the purpose of load balancing. The use of the weight is defined in [I-D.ietf-spring-segment-routing].

System-ID: 6 octets of IS-IS System-ID of length "ID Length" as defined in [ISO10589].

SID/Index/Label: according to the V and L flags, it contains either:

* A 3 octet local label where the 20 rightmost bits are used for encoding the label value. In this case the V and L flags MUST be set.

* A 4 octet index defining the offset in the SID/Label space advertised by this router using the encodings defined in Section 3.1. In this case V and L flags MUST be unset.

Multiple LAN-Adj-SID sub-TLVs MAY be encoded. Note that this sub-TLV MUST NOT appear in TLV 141.

When the P-flag is not set, the LAN-Adj-SID MAY be persistent. When the P-flag is set, the LAN-Adj-SID MUST be persistent.

In case one TLV-22/23/222/223 (reporting the adjacency to the DIS) can't contain the whole set of LAN-Adj-SID sub-TLVs, multiple advertisements of the adjacency to the DIS MUST be used and all advertisements MUST have the same metric.

Each router within the level, by receiving the DIS PN LSP as well as the non-PN LSP of each router in the LAN, is capable of reconstructing the LAN topology as well as the set of Adj-SID each router uses for each of its neighbors.

A label is encoded in 3 octets (in the 20 rightmost bits).

An index is encoded in 4 octets.

2.3. SID/Label Sub-TLV

The SID/Label sub-TLV may be present in the following TLVs/sub-TLVs defined in this document:

SR-Capabilities Sub-TLV (Section 3.1)

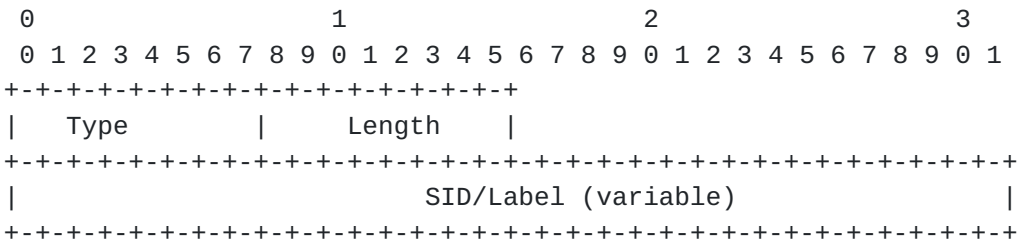
SR Local Block Sub-TLV (Section 3.3)

SID/Label Binding TLV (Section 2.4)

Multi-Topology SID/Label Binding TLV (Section 2.5)

Note that the code point used in all of the above cases is the SID/Label Sub-TLV code point assigned by IANA in the "sub-TLVs for TLV 149 and 150" registry.

The SID/Label sub-TLV contains a SID or a MPLS Label. The SID/Label sub-TLV has the following format:



where:

Type: TBD, suggested value 1

Length: variable

SID/Label: if length is set to 3 then the 20 rightmost bits represent a MPLS label.

2.4. SID/Label Binding TLV

The SID/Label Binding TLV MAY be originated by any router in an IS-IS domain.

The SID/Label Binding TLV is used to advertise prefixes to SID/Label mappings. This functionality is called the Segment Routing Mapping Server (SRMS). The behavior of the SRMS is defined in [I-D.ietf-spring-segment-routing-ldp-interop].

The SID/Label Binding TLV has Type TBD (suggested value 149), and has the following format:

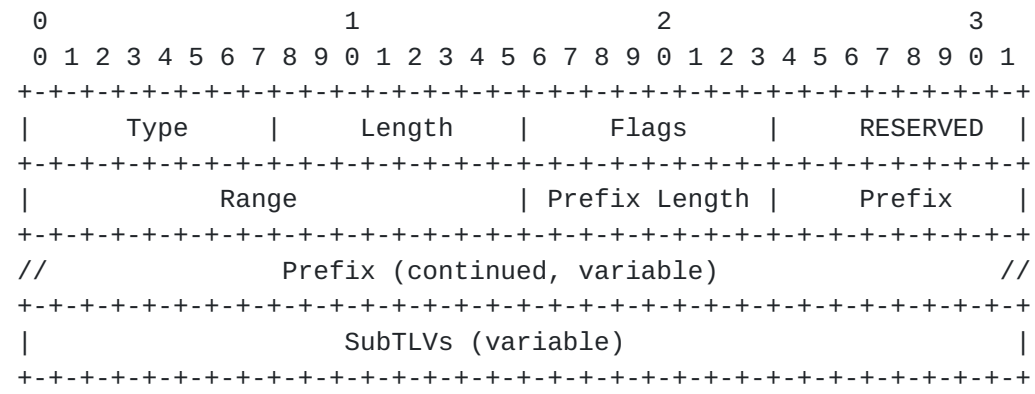
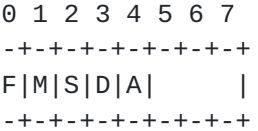


Figure 1: SID/Label Binding TLV format

- Type: TBD, suggested value 149
- Length: variable.
- 1 octet of flags
- 1 octet of RESERVED
- 2 octets of Range
- 1 octet of Prefix Length
- 0-16 octets of Prefix
- sub-TLVs, where each sub-TLV consists of a sequence of:
  - 1 octet of sub-TLV type
  - 1 octet of length of the value field of the sub-TLV
  - 0-243 octets of value

2.4.1. Flags

Flags: 1 octet field of following flags:



where:

F-Flag: Address Family flag. If unset, then the Prefix carries an IPv4 Prefix. If set then the Prefix carries an IPv6 Prefix.

M-Flag: Mirror Context flag. Set if the advertised SID corresponds to a mirrored context. The use of the M flag is described in [I-D.ietf-spring-segment-routing].

S-Flag: If set, the SID/Label Binding TLV SHOULD be flooded across the entire routing domain. If the S flag is not set, the SID/Label Binding TLV MUST NOT be leaked between levels. This bit MUST NOT be altered during the TLV leaking.

D-Flag: when the SID/Label Binding TLV is leaked from level-2 to level-1, the D bit MUST be set. Otherwise, this bit MUST be clear. SID/Label Binding TLVs with the D bit set MUST NOT be leaked from level-1 to level-2. This is to prevent TLV looping across levels.

A-Flag: Attached flag. The originator of the SID/Label Binding TLV MAY set the A bit in order to signal that the prefixes and SIDs advertised in the SID/Label Binding TLV are directly connected to their originators. The mechanisms through which the originator of the SID/Label Binding TLV can figure out if a prefix is attached or not are outside the scope of this document (e.g.: through explicit configuration). If the Binding TLV is leaked to other areas/levels the A-flag MUST be cleared.

An implementation MAY decide not to honor the S-flag in order not to leak Binding TLV's between levels (for policy reasons). In all cases, the D flag MUST always be set by any router leaking the Binding TLV from level-2 into level-1 and MUST be checked when propagating the Binding TLV from level-1 into level-2. If the D flag is set, the Binding TLV MUST NOT be propagated into level-2.

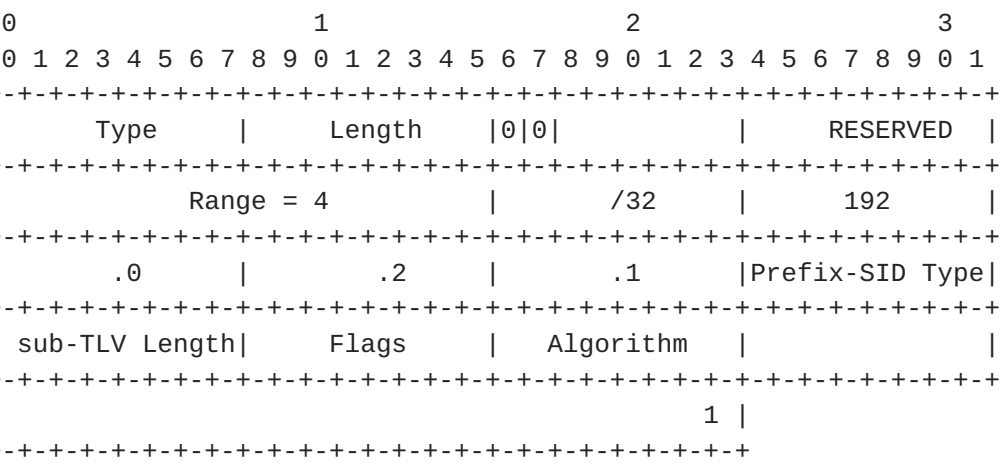
Other bits: MUST be zero when originated and ignored when received.

2.4.2. Range

The 'Range' field provides the ability to specify a range of addresses and their associated Prefix SIDs. This advertisement supports the SRMS functionality. It is essentially a compression scheme to distribute a continuous Prefix and their continuous, corresponding SID/Label Block. If a single SID is advertised then the range field MUST be set to one. For range advertisements > 1, the number of addresses that need to be mapped into a Prefix-SID and the starting value of the Prefix-SID range.

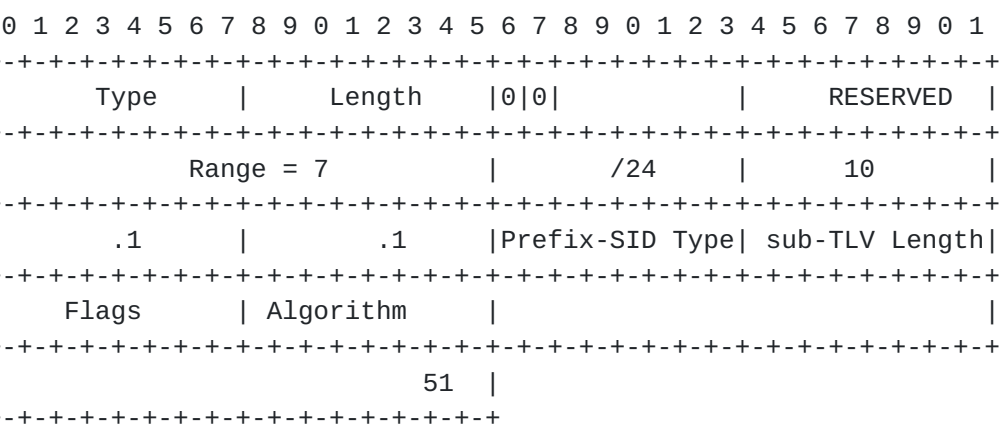
Example 1: if the following router addresses (loopback addresses) need to be mapped into the corresponding Prefix SID indexes.

Router-A: 192.0.2.1/32, Prefix-SID: Index 1
Router-B: 192.0.2.2/32, Prefix-SID: Index 2
Router-C: 192.0.2.3/32, Prefix-SID: Index 3
Router-D: 192.0.2.4/32, Prefix-SID: Index 4



Example-2: If the following prefixes need to be mapped into the corresponding Prefix-SID indexes:

10.1.1/24, Prefix-SID: Index 51
10.1.2/24, Prefix-SID: Index 52
10.1.3/24, Prefix-SID: Index 53
10.1.4/24, Prefix-SID: Index 54
10.1.5/24, Prefix-SID: Index 55
10.1.6/24, Prefix-SID: Index 56
10.1.7/24, Prefix-SID: Index 57



It is not expected that a network operator will be able to keep fully continuous Prefix / SID/Index mappings. In order to support noncontinuous mapping ranges an implementation MAY generate several instances of Binding TLVs.

For example if a router wants to advertise the following ranges:

Range 16: { 192.0.2.1-15, Index 1-15 }

Range 6: { 192.0.2.22-27, Index 22-27 }

Range 41: { 192.0.2.44-84, Index 80-120 }

A router would need to advertise three instances of the Binding TLV.

2.4.3. Prefix Length, Prefix

The 'Prefix' represents the Forwarding equivalence class at the tail-end of the advertised path. The 'Prefix' does not need to correspond to a routable prefix of the originating node.

The 'Prefix Length' field contains the length of the prefix in bits. Only the most significant octets of the Prefix are encoded. (i.e., 1 octet for prefix length 1 up to 8, 2 octets for prefix length 9 to 16, 3 octets for prefix length 17 up to 24 and 4 octets for prefix length 25 up to 32, ..., 16 octets for prefix length 113 up to 128).

2.4.4. Mapping Server Prefix-SID

The Prefix-SID sub-TLV (suggested value 3) is defined in Section 2.1 and contains the SID/index/label value associated with the prefix and range. The Prefix-SID SubTLV MUST be present in the SID/Label Binding TLV unless the M-flag is set in the Flags field of the parent TLV.

A node receiving a MS entry for a prefix MUST check the existence of such prefix in its link-state database prior to consider and use the associated SID.

2.4.4.1. Prefix-SID Flags

The Prefix-SID flags are defined in Section 2.1. The Mapping Server MAY advertise a mapping with the N flag set when the prefix being mapped is known in the link-state topology with a mask length of 32 (IPv4) or 128 (IPv6) and when the prefix represents a node. The mechanisms through which the operator defines that a prefix represents a node are outside the scope of this document (typically it will be through configuration).

The other flags defined in Section 2.1 are not used by the Mapping Server and MUST be ignored at reception.

2.4.4.2. PHP Behavior when using Mapping Server Advertisements

As the mapping server does not specify the originator of a prefix advertisement it is not possible to determine PHP behavior solely based on the Mapping Server Advertisement. However, if additional information is available PHP behavior may safely be done. The required information consists of:

- A prefix reachability advertisement for the prefix has been received which includes the Extended Reachability Attribute Flags sub-TLV ([RFC7794]).

- X and R flags are both set to 0 in the Extended Reachability Attribute Flags sub-TLV.

In the absence of an Extended Reachability Attribute Flags sub-TLV ([RFC7794]) the A flag in the binding TLV indicates that the originator of a prefix reachability advertisement is directly connected to the prefix and thus PHP MUST be done by the neighbors of the router originating the prefix reachability advertisement. Note that A-flag is only valid in the original area in which the Binding TLV is advertised.

2.4.4.3. Prefix-SID Algorithm

The algorithm field contains the identifier of the algorithm the router MUST use in order to compute reachability to the range of prefixes. Use of the algorithm field is described in Section 2.1.

2.4.5. SID/Label Sub-TLV

The SID/Label sub-TLV (Type: TBD, suggested value 1) contains the SID/Label value as defined in Section 2.3. It MUST be present in the SID/Label Binding TLV when the M-flag is set in the Flags field of the parent TLV.

2.5. Multi-Topology SID/Label Binding TLV

The Multi-Topology SID/Label Binding TLV allows the support of M-ISIS as defined in [RFC5120]. The Multi-Topology SID/Label Binding TLV has the same format as the SID/Label Binding TLV defined in Section 2.4 with the difference consisting of a Multitopology Identifier (MTID) as defined here below:

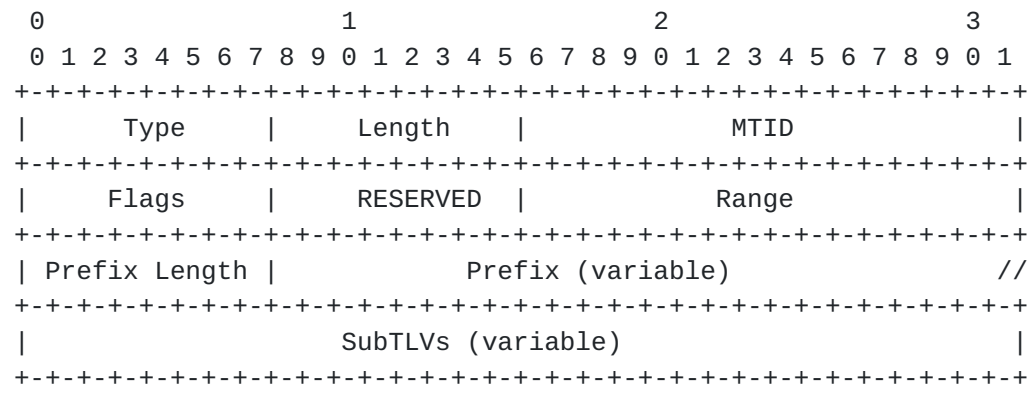


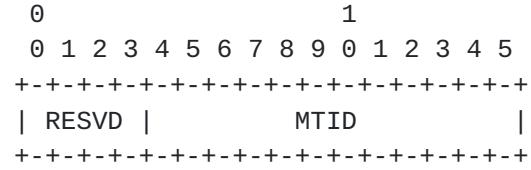
Figure 2: Multi-Topology SID/Label Binding TLV format

where:

Type: TBD, suggested value 150

Length: variable

MTID is the multitopology identifier defined as:



RESVD: reserved bits. MUST be reset on transmission and ignored on receive.

MTID: a 12-bit field containing the non-zero ID of the topology being announced. The TLV MUST be ignored if the ID is zero. This is to ensure the consistent view of the standard unicast topology.

The other fields and SubTLVs are defined in Section 2.4.

3. Router Capabilities

This section defines sub-TLVs which are inserted into the IS-IS Router Capability TLV-242 that is defined in [RFC7981].

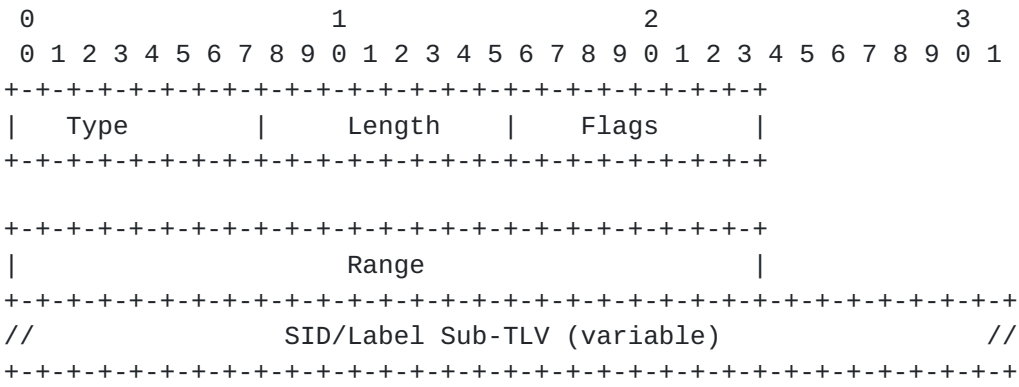
3.1. SR-Capabilities Sub-TLV

Segment Routing requires each router to advertise its SR data-plane capability and the range of MPLS label values it uses for Segment Routing in the case where global SIDs are allocated (i.e., global

indexes). Data-plane capabilities and label ranges are advertised using the newly defined SR-Capabilities sub-TLV.

The Router Capability TLV specifies flags that control its advertisement. The SR Capabilities sub-TLV MUST be propagated throughout the level and MUST NOT be advertised across level boundaries. Therefore Router Capability TLV distribution flags are set accordingly, i.e., the S flag in the Router Capability TLV ([RFC7981]) MUST be unset.

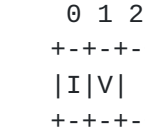
The SR Capabilities sub-TLV has following format:



Type: TBD, suggested value 2

Length: variable.

Flags: 1 octet of flags. The following are defined:



where:

I-Flag: MPLS IPv4 flag. If set, then the router is capable of processing SR MPLS encapsulated IPv4 packets on all interfaces.

V-Flag: MPLS IPv6 flag. If set, then the router is capable of processing SR MPLS encapsulated IPv6 packets on all interfaces.

One or more SRGB Descriptor entries, each of which have the following format:

Range: 3 octets.

SID/Label sub-TLV (as defined in Section 2.3).

SID/Label sub-TLV contains the first value of the SRGB while the range contains the number of SRGB elements. The range value MUST be higher than 0.

The SR-Capabilities sub-TLV MAY be advertised in an LSP of any number but a router MUST NOT advertise more than one SR-Capabilities sub-TLV. A router receiving multiple SR-Capabilities sub-TLVs, from the same originator, SHOULD select the first advertisement in the lowest numbered LSP.

When multiple SRGB Descriptors are advertised the entries define an ordered set of ranges on which a SID index is to be applied. For this reason changing the order in which the descriptors are advertised will have a disruptive effect on forwarding.

When a router adds a new SRGB Descriptor to an existing SR-Capabilities sub-TLV the new Descriptor SHOULD add the newly configured block at the end of the sub-TLV and SHOULD NOT change the order of previously advertised blocks. Changing the order of the advertised descriptors will create label churn in the FIB and blackhole / misdirect some traffic during the IGP convergence. In particular, if a range which is not the last is extended it's preferable to add a new range rather than extending the previously advertised range.

The originating router MUST ensure the order is same after a graceful restart (using checkpointing, non-volatile storage or any other mechanism) in order to guarantee the same order before and after GR.

The originating router MUST NOT advertise overlapping ranges.

When a router receives multiple overlapping ranges, it MUST conform to the procedures defined in [I-D.ietf-spring-conflict-resolution].

Here follows an example of advertisement of multiple ranges:

The originating router advertises following ranges:

```
SR-Cap: range: 100, SID value: 100
SR-Cap: range: 100, SID value: 1000
SR-Cap: range: 100, SID value: 500
```

The receiving routers concatenate the ranges in the received order and build the SRGB as follows:

```
SRGB = [100, 199]
       [1000, 1099]
       [500, 599]
```

The indexes span multiple ranges:

```
index=0   means label 100
...
index 99  means label 199
index 100 means label 1000
index 199 means label 1099
...
index 200 means label 500
...
```

3.2. SR-Algorithm Sub-TLV

The router may use various algorithms when calculating reachability to other nodes or to prefixes attached to these nodes. Examples of these algorithms are metric based Shortest Path First (SPF), various sorts of Constrained SPF, etc. The SR-Algorithm sub-TLV (Type: TBD, suggested value 19) allows the router to advertise the algorithms that the router is currently using. The following values have been defined:

0: Shortest Path First (SPF) algorithm based on link metric. This is the well-known shortest path algorithm as computed by the IS-IS Decision process. Consistent with the deployed practice for link-state protocols, algorithm 0 permits any node to overwrite the SPF path with a different path based on local policy.

1: Strict Shortest Path First (SPF) algorithm based on link metric. The algorithm is identical to algorithm 0 but algorithm 1 requires that all nodes along the path will honor the SPF routing decision. Local policy MUST NOT alter the forwarding decision computed by algorithm 1 at the node claiming to support algorithm 1.

The Router Capability TLV specifies flags that control its advertisement. The SR-Algorithm MUST be propagated throughout the

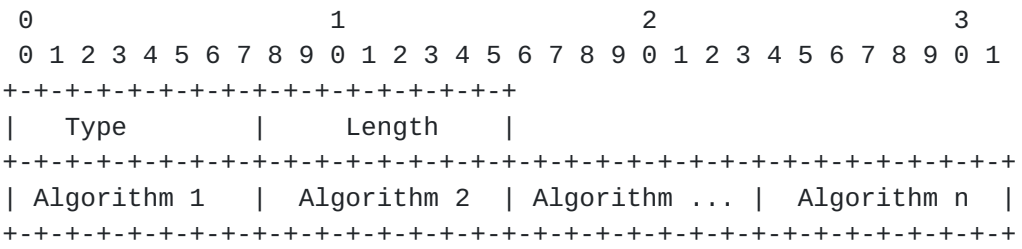
level and MUST NOT be advertised across level boundaries. Therefore Router Capability TLV distribution flags are set accordingly, i.e., the S flag MUST be unset.

The SR-Algorithm sub-TLV is optional, it MAY only appear a single time inside the Router Capability TLV.

When the originating router does not advertise the SR-Algorithm sub-TLV, then all the Prefix-SIDs advertised by the router MUST have algorithm field set to 0. Any receiving router MUST assume SPF algorithm (i.e., Shortest Path First).

When the originating router does advertise the SR-Algorithm sub-TLV, then algorithm 0 MUST be present while algorithm 1 MAY be present.

The SR-Algorithm sub-TLV has following format:



where:

Type: TBD, suggested value 19

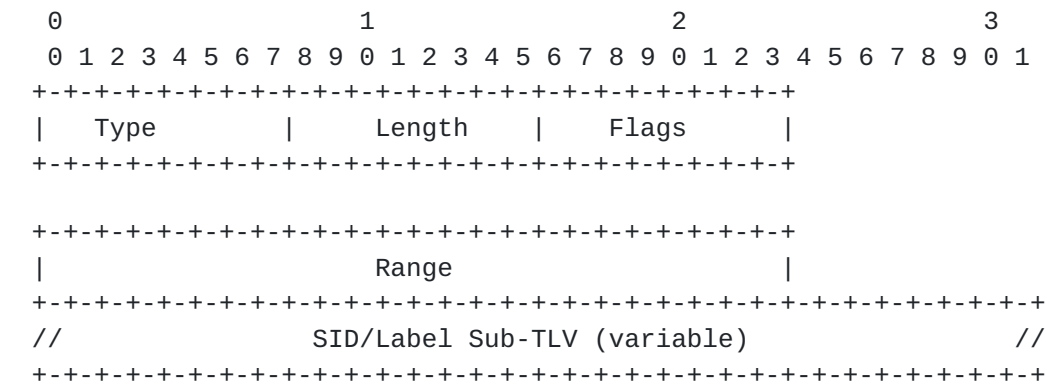
Length: variable.

Algorithm: 1 octet of algorithm Section 2.1

3.3. SR Local Block Sub-TLV

The SR Local Block (SRLB) Sub-TLV contains the range of labels the node has reserved for local SIDs. Local SIDs are used, e.g., for Adjacency-SIDs, and may also be allocated by other components than IS-IS protocol. As an example, an application or a controller may instruct the router to allocate a specific local SID. Therefore, in order for such applications or controllers to know what are the local SIDs available in the router, it is required that the router advertises its SRLB.

The SRLB Sub-TLV is used for that purpose and has following format:



Type: TBD, suggested value 22.

Length: variable.

Flags: 1 octet of flags. None are defined at this stage.

One or more SRLB Descriptor entries, each of which have the following format:

Range: 3 octets.

SID/Label sub-TLV (as defined in Section 2.3).

SID/Label sub-TLV contains the first value of the SRLB while the range contains the number of SRLB elements. The range value MUST be higher than 0.

The SRLB sub-TLV MAY be advertised in an LSP of any number but a router MUST NOT advertise more than one SRLB sub-TLV. A router receiving multiple SRLB sub-TLVs, from the same originator, SHOULD select the first advertisement in the lowest numbered LSP.

The originating router MUST NOT advertise overlapping ranges.

It is important to note that each time a SID from the SRLB is allocated, it SHOULD also be reported to all components (e.g.: controller or applications) in order for these components to have an up-to-date view of the current SRLB allocation and in order to avoid collision between allocation instructions.

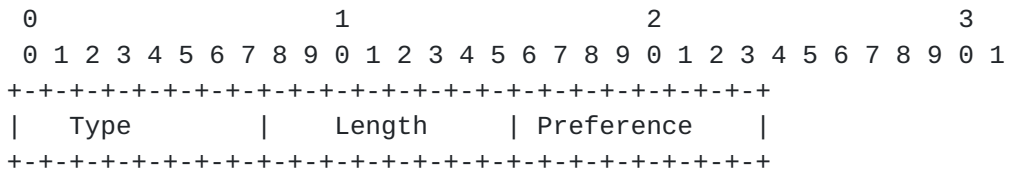
Within the context of IS-IS, the reporting of local SIDs is done through IS-IS Sub-TLVs such as the Adjacency-SID. However, the reporting of allocated local SIDs may also be done through other means and protocols which mechanisms are outside the scope of this document.

A router advertising the SRLB TLV may also have other label ranges, outside the SRLB, for its local allocation purposes which are NOT advertised in the SRLB. For example, it is possible that an Adjacency-SID is allocated using a local label not part of the SRLB.

3.4. SRMS Preference Sub-TLV

The Segment Routing Mapping Server (SRMS) Preference sub-TLV is used in order to associate a preference with SRMS advertisements from a particular source.

The SRMS Preference sub-TLV has following format:



Type: TBD, suggested value 24.

Length: 1.

Preference: 1 octet. Unsigned 8 bit SRMS preference.

The SRMS Preference sub-TLV MAY be advertised in an LSP of any number but a router MUST NOT advertise more than one SRMS Preference sub-TLV. A router receiving multiple SRMS Preference sub-TLVs, from the same originator, SHOULD select the first advertisement in the lowest numbered LSP.

The use of the SRMS Preference during the SID selection process is described in [I-D.ietf-spring-conflict-resolution].

4. Non backward compatible changes with prior versions of this document

This section describes the changes that have been applied to this document that are not backward compatible with previous versions.

4.1. Encoding of Multiple SRGBs

Version -04 of this document introduced a change in Section 3.1 regarding the encoding method for multiple SRGBs in the SR-Cap SubTLV and made the support of multiple SRGBs REQUIRED.

The modified method consists of having a single SR-Cap Sub-TLV where all SRGBs are encoded. In previous versions (prior to version -04)

of this document it was allowed to have multiple occurrences of the SR-Cap Sub-TLV.

At the time of writing this document, no existing implementations are affected by the change since no implementations actually (i.e., at the time of updating this document) encode multiple SRGBs anyway.

5. IANA Considerations

This documents request allocation for the following TLVs and subTLVs.

5.1. Sub TLVs for Type 22,23,25,141,222, and 223

This document makes the following registrations in the "sub-TLVs for TLV 22, 23, 25, 141, 222 and 223" registry.

Type: TBD (suggested value 31)

Description: Adjacency Segment Identifier

TLV 22: yes

TLV 23: yes

TLV 25: no

TLV 141: yes

TLV 222: yes

TLV 223: yes

Reference: This document (Section 2.2.1)

Type: TBD (suggested value 32)

Description: LAN Adjacency Segment Identifier

TLV 22: yes

TLV 23: yes

TLV 25: no

TLV 141: yes

TLV 222: yes

TLV 223: yes

Reference: This document (Section 2.2.2)

5.2. Sub TLVs for Type 135,235,236 and 237

This document makes the following registrations in the "sub-TLVs for TLV 135,235,236 and 237" registry.

Type: TBD (suggested value 3)

Description: Prefix Segment Identifier

TLV 135: yes

TLV 235: yes

TLV 236: yes

TLV 237: yes

Reference: This document (Section 2.1)

5.3. Sub TLVs for Type 242

This document makes the following registrations in the "sub-TLVs for TLV 242" registry.

Type: TBD (suggested value 2)

Description: Segment Routing Capability

Reference: This document (Section 3.1)

Type: TBD (suggested value 19)

Description: Segment Routing Algorithm

Reference: This document (Section 3.2)

Type: TBD (suggested value 22)

Description: Segment Routing Local Block (SRLB)

Reference: This document (Section 3.3)

Type: TBD (suggested value 24)

Description: Segment Routing Mapping Server Preference (SRMS Preference)

Reference: This document (Section 3.4)

5.4. New TLV Codepoint and Sub-TLV registry

This document registers the following TLV:

Type: TBD (suggested value 149)

name: Segment Identifier / Label Binding

IIH: no

LSP: yes

SNP: no

Purge: no

Reference: This document (Section 2.4)

Type: TBD (suggested value 150)

name: Multi-Topology Segment Identifier / Label Binding

IIH: no

LSP: yes

SNP: no

Purge: no

Reference: This document (Section 2.5)

This document creates the following sub-TLV Registry:

Registry: sub-TLVs for TLV 149 and 150

Registration Procedure: Expert review

Reference: This document (Section 2.4)

Type: TBD, suggested value 1

Description: SID/Label

Reference: This document (Section 2.4.5)

Type: TBD, suggested value 3

Description: Prefix-SID

Reference: This document (Section 2.1)

6. Security Considerations

With the use of the extensions defined in this document, IS-IS carries information which will be used to program the MPLS data plane [RFC3031]. In general, the same types of attacks that can be carried out on the IP/IPv6 control plane can be carried out on the MPLS control plane resulting in traffic being misrouted in the respective data planes. However, the latter may be more difficult to detect and isolate. Existing security extensions as described in [RFC5304] and [RFC5310] apply to these segment routing extensions.

7. Acknowledgements

We would like to thank Dave Ward, Dan Frost, Stewart Bryant, Pierre Francois and Jesper Skrivers for their contribution to the content of this document.

8. Contributors

The following people gave a substantial contribution to the content of this document and should be considered as co-authors:

Peter Psenak
Cisco Systems Inc.
US

Email: ppsenak@cisco.com

Martin Horneffer
Deutsche Telekom
DE

Email: Martin.Horneffer@telekom.de

Wim Henderickx
Nokia
BE

Email: wim.henderickx@nokia.com

Edward Crabbe
Individual
US

Email: edward.crabbe@gmail.com

Rob Shakir
Google
UK

Email: robjs@google.com

Igor Milojevic
Individual
RS

Email: milojevicigor@gmail.com

Saku Ytti
TDC
FI

Email: saku@ytti.fi

Steven Luong
Cisco Systems Inc.
US

Email: sluong@cisco.com

9. References

9.1. Normative References

[I-D.ietf-spring-conflict-resolution]
Ginsberg, L., Psenak, P., Previdi, S., and M. Pilka, "Segment Routing MPLS Conflict Resolution", draft-ietf-spring-conflict-resolution-05 (work in progress), July 2017.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-13 (work in progress), October 2017.

[ISO10589]
International Organization for Standardization, "Intermediate system to Intermediate system intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002, Second Edition, Nov 2002.

[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>.

[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>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[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>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <https://www.rfc-editor.org/info/rfc5308>.

[RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, DOI 10.17487/RFC5310, February 2009, <https://www.rfc-editor.org/info/rfc5310>.

[RFC7794] Ginsberg, L., Ed., Decraene, B., Previdi, S., Xu, X., and U. Chunduri, "IS-IS Prefix Attributes for Extended IPv4 and IPv6 Reachability", RFC 7794, DOI 10.17487/RFC7794, March 2016, <https://www.rfc-editor.org/info/rfc7794>.

[RFC7981] Ginsberg, L., Previdi, S., and M. Chen, "IS-IS Extensions for Advertising Router Information", RFC 7981, DOI 10.17487/RFC7981, October 2016, <https://www.rfc-editor.org/info/rfc7981>.

9.2. Informative References

[I-D.ietf-spring-resiliency-use-cases] Filsfils, C., Previdi, S., Decraene, B., and R. Shakir, "Resiliency use cases in SPRING networks", draft-ietf-spring-resiliency-use-cases-12 (work in progress), December 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-09 (work in progress), September 2017.

[RFC5311] McPherson, D., Ed., Ginsberg, L., Previdi, S., and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, DOI 10.17487/RFC5311, February 2009, <https://www.rfc-editor.org/info/rfc5311>.

[RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, DOI 10.17487/RFC5316, December 2008, <https://www.rfc-editor.org/info/rfc5316>.

[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, <https://www.rfc-editor.org/info/rfc7855>.

Authors' Addresses

Stefano Previdi (editor)
Cisco Systems, Inc.
IT

Email: stefano@previdi.net

Les Ginsberg (editor)
Cisco Systems, Inc.
IT

Email: ginsberg@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels
BE

Email: cfilsfil@cisco.com

Ahmed Bashandy
Cisco Systems, Inc.
170, West Tasman Drive
San Jose, CA 95134
US

Email: bashandy@cisco.com

Hannes Gredler
RtBrick Inc.

Email: hannes@rtbrick.com

Stephane Litkowski
Orange
FR

Email: stephane.litkowski@orange.com

Bruno Decraene
Orange
FR

Email: bruno.decraene@orange.com

Jeff Tantsura
Individual

Email: jefftant@gmail.com