IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2014

S. Previdi, Ed.
C. Filsfils
A. Bashandy
Cisco Systems, Inc.
H. Gredler
Juniper Networks, Inc.
S. Litkowski
Orange

October 21, 2013

IS-IS Extensions for Segment Routing
draft-previdi-isis-segment-routing-extensions-03

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.

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 http://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 April 24, 2014.

Copyright Notice

Copyright (c) 2013 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 (http://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 4
2. Segment Routing Identifiers 4
  2.1. SID/Label Sub-TLV 4
  2.2. Prefix Segment Identifier (Prefix-SID Sub-TLV) 5
  2.3. Adjacency Segment Identifier (Adj-SID) Sub-TLV 8
    2.3.1. Adjacency Segment Identifier (Adj-SID) Sub-TLV 9
    2.3.2. Adjacency Segment Identifiers in LANs 10
  2.4. SID/Label Binding TLV 12
    2.4.1. Flags 13
    2.4.2. Weight 13
    2.4.3. Range 13
    2.4.4. Prefix Length, Prefix 15
    2.4.5. SID/Label Sub-TLV 15
    2.4.6. IPv4 ERO subTLV 15
    2.4.7. IPv6 ERO subTLV 16
    2.4.8. Unnumbered Interface ID ERO subTLV 16
    2.4.9. IPv4 Backup ERO subTLV 17
    2.4.10. IPv6 Backup ERO subTLV 18
    2.4.11. Unnumbered Interface ID Backup ERO subTLV 18
    2.4.12. Prefix ERO and Prefix Backup ERO subTLV path semantics 19
3. Router Capabilities 20
  3.1. SR-Capabilities Sub-TLV 20
  3.2. SR-Algorithm Sub-TLV 21
4. IANA Considerations 22
5. Manageability Considerations 22
6. Security Considerations 22
7. Contributors 22
8. Acknowledgements 22
9. References 22
  9.1. Normative References 22
  9.2. Informative References 23

Authors' Addresses 24

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.

Segment Routing architecture is described in [I-D.filsfils-rtgwg-segment-routing].

Segment Routing use cases are described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

2. Segment Routing Identifiers

Segment Routing architecture ([I-D.filsfils-rtgwg-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. SID/Label Sub-TLV

The SID/Label Sub-TLV is present in multiple Sub-TLVs defined in this document and contains a SID or a MPLS Label. The SID/Label Sub-TLV has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          SID/Label (variable)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBA

Length: variable (3 or 4)

SID/Label: if length is set to 3 then the 20 rightmost bits represent a MPLS label. If length is 4 then the value represents a 32 bits SID.

2.2. 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.filsfils-rtgwg-segment-routing]. The 'Prefix SID' must be unique within a given IGP domain. The 'Prefix SID' is an index to determine 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 dissambiguate IP Prefix to node mappings. Furthermore if a given node has several 'stable transport' IP adresses 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 (IPv4) defined in [RFC5305].

TLV-235 (MT-IPv4) defined in [RFC5120].

TLV-236 (IPv6) defined in [RFC5308].

TLV-237 (MT-IPv6) defined in [RFC5120].

The Index inside the Prefix-SID Sub-TLV MUST be preserved when an IP Reachability TLV gets propagated across level boundaries.

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |     Flags     |   Algorithm   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        SID/Index                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBA

Length: variable.

Flags: 1 octet field of following flags:

```
0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|R|N|P|         |
+-+-+-+-+-+-+-+-+
```

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. Optional and, 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.filsfils-rtgwg-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.

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. Examples of these algorithms are metric based Shortest Path First (SPF), various sorts of Constrained SPF, etc. The Algorithm field allows a router to advertise algorithms that router is currently using. SR-Algorithm TLV has following structure: one octet identifying the algorithm to which the Prefix-SID is associated. Currently, the following value has been defined:

0: Shortest Path First (SPF) algorithm based on link metric.

Definitions and use of algorithms in Segment Routing are described in [I-D.filsfils-rtgwg-segment-routing]

SID/Index: 32 bit index defining the offset in the SID/Label space advertised by this router using the encodings defined in Section 3.1.

Multiple Prefix-SIDs Sub-TLVs MAY appear on the same prefix in which case each SID is encoded as a separate Sub-TLV. When multiple Prefix-SID Sub-TLVs are present, the receiving router MUST use the first encoded SID and MAY use the subsequent ones.

The No-PHP flag MUST be set on the Prefix-SIDs associated with reachability advertisements which were originated by other routers and leaked (either from Level-1 to Level-2 or vice versa).

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 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.: '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).

The router behavior determined by the P, R and N flags are described in [I-D.filsfils-rtgwg-segment-routing].

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

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.filsfils-rtgwg-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 [RFC5305]

TLV-222 [RFC5120]

TLV-23 [RFC5311]

TLV-223 [RFC5311]

TLV-141 [RFC5316]

Multiple Adj-SID Sub-TLVs MAY be associated with a single IS-neighbor. Examples where more than one Adj-SID may be used per IS-neighbor are described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

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

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |     Flags     |     Weight    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               SID/Label Sub-TLV (variable)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBA

Length: variable.

Flags: 1 octet field of following flags:

```
0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+
|F|B|         |
+-+-+-+-+-+-+-+
```

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 refers to an adjacency being protected (e.g.: using IPFRR or MPLS-FRR) as described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

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.filsfils-rtgwg-segment-routing].

SID/Label Sub-TLV: contains the SID/Label value as defined in Section 2.1.

An SR capable router MAY allocate an Adj-SID for each of its adjacencies and SHOULD set the B-Flag when the adjacency is protected by a FRR mechanism (IP or MPLS) as described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

The F-flag is used in order for the router to advertise the outgoing encapsulation of the adjacency the Adj-SID is attached to. Use cases of the use of the F-flag are described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

2.3.2. 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 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |      Flags    |    Weight     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     System-ID (6 octets)                      |
+                               +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   SID/Label Sub-TLV (variable)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBA.

Length: variable.

Flags: 1 octet field of following flags:

```
0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+
|F|B|         |
+-+-+-+-+-+-+-+
```

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 LAN-Adj-SID refers to an adjacency being protected (e.g.: using IPFRR or MPLS-FRR) as described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

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.filsfils-rtgwg-segment-routing].

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

SID/Label Sub-TLV: contains the SID/Label value as defined in Section 2.1.

Multiple LAN-Adj-SID Sub-TLVs MAY be encoded.

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, MUST have the same metric and SHOULD be inserted within the same LSP fragment.

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.

2.4. SID/Label Binding TLV

The SID/Label Binding TLV MAY be originated by any router in an IS-IS domain. The router may advertise a SID/Label binding to a FEC along with at least a single 'nexthop style' anchor. The protocol supports more than one 'nexthop style' anchor to be attached to a SID/Label binding, which results into a simple path description language. In analogy to RSVP the terminology for this is called an 'Explicit Route Object' (ERO). Since ERO style path notation allows to anchor SID/label bindings to to both link and node IP addresses any label switched path, can be described. Furthermore also SID/Label Bindings from external protocols can get easily re-advertised.

The SID/Label Binding TLV may be used for advertising SID/Label Bindings and their associated Primary and Backup paths. In one single TLV either a primary ERO Path, a backup ERO Path or both are advertised. If a router wants to advertise multiple parallel paths then it can generate several TLVs for the same Prefix/FEC. Each occurence of a Binding TLV with respect with a given FEC Prefix has accumulating and not canceling semantics. Due the space constraints in the 8-Bit IS-IS TLVs an originating router MAY encode a primary ERO path in one SID/Label Binding TLV and the backup ERO path in a second SID/Label Binding TLV. Note that the FEC Prefix and SID/Label Sub-TLV MUST be identical in both TLVs.

The SID/Label Binding TLV has type TBA and has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |     Length    |     Flags     |     Weight    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Range              | Prefix Length |  FEC Prefix   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//           FEC Prefix (continued, variable)                  //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    optional subTLVs (variable)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: SID/Label Binding TLV format

- o 1 octet of flags
- o 1 octet of Prefix length
- o 0-16 octets of FEC Prefix

- o 2 octets of Range
- o 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
  - * 0-255 octets of value

2.4.1. Flags

Flags: 1 octet field of following flags:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|F|M|X|S|       |
+-+-+-+-+-+-+-+-+
```

where:

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

M-Flag: Mirror Context flag. Set if the advertised SID/path corresponds to a mirrored context.

X-Flag: Index flag. Set if the value of the SID/Label Sub-TLV carries an index. Unset if the value of the SID/Label Sub-TLV carries a local SID/Label.

S-Flag: subTLV present 'S' flag: Set if there are subTLVs present.

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

2.4.2. Weight

Weight: 1 octet: The value represents the weight of the path for the purpose of load balancing. The use of the weight is defined in [I-D.filsfils-rtgwg-segment-routing].

2.4.3. Range

The 'Range' field provides the ability to specify a range of addresses and their associated Prefix SIDs. 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 advertisments > 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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |     Length    |0|0|1|1|       |     Weight    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Range = 4          |       /32     |      192      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       .0      |        .2     |       .1      |  Sub-TLV Type |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Sub-TLV Length|                                               1 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |     Length    |0|0|1|1|       |     Weight    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Range = 7          |       /24     |      10       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       .1      |        .1     |  Sub-TLV Type | Sub-TLV Length|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                              51 |
```

+-+

It is not expected that a network operator will be able to keep fully continuous FEC 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: { 192.168.1.1-15, Index 1-15 }

Range 6: { 192.168.1.22-27, Index 22-27 }

Range 41: { 192.168.1.44-84, Index 80-120 }

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

2.4.4. Prefix Length, Prefix

The 'FEC Prefix' represents the Forwarding equivalence class at the tail-end of the advertised path. The 'FEC 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 FEC are encoded. I.e. 1 octet for FEC prefix length 1 up to 8, 2 octets for FEC prefix length 9 to 16, 3 octets for FEC prefix length 17 up to 24 and 4 octets for FEC prefix length 25 up to 32, ..., 16 octets for FEC prefix length 113 up to 128.

2.4.5. SID/Label Sub-TLV

The SID/Label Sub-TLV contains the SID/Label value as defined in Section 2.1. It MUST be present in every SID/Label Binding TLV.

2.4.6. IPv4 ERO subTLV

The IPv4 ERO subTLV (Type TBA) describes a path segment using IPv4 address style of encoding. Its semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Length    |L|   Reserved  |  IPv4 address |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          IPv4 address  (continued)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: IPv4 ERO subTLV format

2.4.7. IPv6 ERO subTLV

The IPv6 ERO subTLV (Type TBA) describes a path segment using IPv6 Address style of encoding. Its semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Length    |L|   Reserved  |  IPv6 address |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: IPv6 ERO subTLV format

2.4.8. Unnumbered Interface ID ERO subTLV

The appearance and semantics of the 'Unnumbered Interface ID' have been borrowed from Section 4 [RFC3477].

The Unnumbered Interface-ID ERO subTLV (Type TBA) describes a path segment that spans over an unnumbered interface. Unnumbered interfaces are referenced using the interface index. Interface indices are assigned local to the router and therefore not unique within a domain. All elements in an ERO path need to be unique within a domain and hence need to be disambiguated using a domain unique Router-ID.

The 'Router-ID' field contains the router ID of the router which has assigned the 'Interface ID' field. Its purpose is to disambiguate the 'Interface ID' field from other routers in the domain.

IS-IS supports two Router-ID formats:

o (TLV 134, 32-Bit format) [RFC5305]

o (TLV 140, 128-Bit format) [RFC6119]

The actual Router-ID format gets derived from the 'Length' field.

o For 32-Bit Router-ID width the subTLV length is set to 8 octets.

o For 128-Bit Router-ID width the subTLV length is set to 20 octets.

The 'Interface ID' is the identifier assigned to the link by the router specified by the router ID.

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |     Length    |L|   Reserved  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                   Router ID (32 or 128 bits)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Interface ID (32 bits)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Unnumbered Interface ID ERO subTLV format

2.4.9. IPv4 Backup ERO subTLV

The IPv4 Backup ERO subTLV (Type TBA) describes a Backup path segment using IPv4 Address style of encoding. Its appearance and semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Length    |L|   Reserved  |  IPv4 address |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          IPv4 address  (continued)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: IPv4 Backup ERO subTLV format

2.4.10. IPv6 Backup ERO subTLV

The IPv6 Backup ERO subTLV (Type TBA) describes a Backup path segment using IPv6 Address style of encoding. Its appearance and semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Length    |L|   Reserved  |  IPv6 address |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: IPv6 Backup ERO subTLV format

2.4.11. Unnumbered Interface ID Backup ERO subTLV

The appearance and semantics of the 'Unnumbered Interface ID' have been borrowed from Section 4 [RFC3477].

The Unnumbered Interface-ID Backup ERO subTLV (Type TBA) describes a Backup LSP path segment that spans over an unnumbered interface. Unnumbered interfaces are referenced using the interface index. Interface indices are assigned local to the router and therefore not unique within a domain. All elements in an ERO path need to be unique within a domain and hence need to be disambiguated using a domain unique Router-ID.

The 'Router-ID' field contains the router ID of the router which has assigned the 'Interface ID' field. Its purpose is to disambiguate the 'Interface ID' field from other routers in the domain.

IS-IS supports two Router-ID formats:

o (TLV 134, 32-Bit format) [RFC5305]

o (TLV 140, 128-Bit format) [RFC6119]

The actual Router-ID format gets derived from the 'Length' field.

o For 32-Bit Router-ID width the subTLV length is set to 8 octets.

o For 128-Bit Router-ID width the subTLV length is set to 20 octets.

The 'Interface ID' is the identifier assigned to the link by the router specified by the router ID.

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |     Length    |L|   Reserved  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                   Router ID (32 or 128 bits)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Interface ID (32 bits)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: Unnumbered Interface ID Backup ERO subTLV format

2.4.12. Prefix ERO and Prefix Backup ERO subTLV path semantics

All 'ERO' and 'Backup ERO' information represents an ordered set which describes the segments of a path. The last ERO subTLV describes the segment closest to the egress point of the path. Contrary the first ERO subTLV describes the first segment of a path. If a router extends or stitches a label switched path it MUST prepend the new segments path information to the ERO list. The same ordering applies for the Backup ERO labels. An implementation SHOULD first encode all primary path EROs followed by the bypass EROs.

3. Router Capabilities

3.1. SR-Capabilities Sub-TLV

Segment Routing requires each router to advertise its SR data-plane capability and the range of SID/Label values it uses for Segment Routing. Data-plane capabilities and SID/Label ranges are advertised using the newly defined SR-Capabilities Sub-TLV inserted into the IS-IS Router Capability TLV-242 that is defined in [RFC4971].

The Router Capability TLV specifies flags that control its advertisement. The SR Capabilities Sub-TLV MUST be propagated throughout the level and need not to be advertised across level boundaries. Therefore Router Capability TLV distribution flags MUST be set accordingly, i.e.: the S flag MUST be unset.

The SR Capabilities Sub-TLV is optional, MAY appear multiple times inside the Router Capability TLV and has following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |    Flags      |    Range      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Range (cont.) |      SID/Label Sub-TLV (variable size)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBA.

Length: variable.

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

```
0
0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|I|V|           |
+-+-+-+-+-+-+-+-+
```

where:

I-Flag: IPv4 flag. If set, then the router is capable of outgoing IPv4 encapsulation on all interfaces.

V-Flag: IPv6 flag. If set, then the router is capable of outgoing IPv6 encapsulation on all interfaces.

Range: 2 octets value defining the number of values of the range from the starting value defined in the SID/Label Sub-TLV.

SID/Label Sub-TLV: SID/Label value as defined in Section 2.1.

If multiple occurrence of the SR-Capabilities Sub-TLV are advertised by the same router, only the Flags in the first occurrence of the Sub-TLV are to be taken into account.

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 allows the router to advertise the algorithms that the router is currently using. The following value has been defined:

0: Shortest Path First (SPF) algorithm based on link metric.

The SR-Algorithm Sub-TLV is inserted into the IS-IS Router Capability TLV-242 that is defined in [RFC4971].

The Router Capability TLV specifies flags that control its advertisement. The SR-Algorithm MUST be propagated throughout the level and need not to be advertised across level boundaries. Therefore Router Capability TLV distribution flags MUST be 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. If the SID-Label Capability Sub-TLV is advertised then the SR-Algorithm Sub-TLV MUST also be advertised.

It has following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Algorithm 1   |  Algorithm 2  | Algorithm ... |  Algorithm n  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBA.

Length: variable.

Algorithm: 1 octet of algorithm Section 2.2

4. IANA Considerations

TBD

5. Manageability Considerations

TBD

6. Security Considerations

TBD

7. Contributors

The following people gave a substantial contribution to the content of this document: Martin Horneffer, Bruno Decraene, Igor Milojevic, Rob Shakir, Saku Ytti and Wim Henderickx.

8. Acknowledgements

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

Many thanks to Yakov Rekhter and Ina Minei for their contribution on earlier incarnations of the "Binding / MPLS Label TLV" in [I-D.gredler-isis-label-advertisement].

9. References

9.1. Normative References

[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, March 1997.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.

[RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

[RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, October 2008.

[RFC5311] McPherson, D., Ginsberg, L., Previdi, S., and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, February 2009.

[RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, December 2008.

[RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", RFC 6119, February 2011.

9.2. Informative References

[I-D.filsfils-rtgwg-segment-routing]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing Architecture", draft-filsfils-rtgwg-segment-routing-00 (work in

progress), June 2013.

[I-D.filsfils-rtgwg-segment-routing-use-cases]
Filsfils, C., Francois, P., Previdi, S., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing Use Cases", draft-filsfils-rtgwg-segment-routing-use-cases-01 (work in progress), July 2013.

[I-D.gredler-isis-label-advertisement]
Gredler, H., Amante, S., Scholl, T., and L. Jalil, "Advertising MPLS labels in IS-IS", draft-gredler-isis-label-advertisement-03 (work in progress), May 2013.

Authors' Addresses

Stefano Previdi (editor)
Cisco Systems, Inc.
Via Del Serafico, 200
Rome 00142
Italy

Email: sprevidi@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
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: hannes@juniper.net

Stephane Litkowski
Orange
FR

Email: stephane.litkowski@orange.com