Open Shortest Path First IGP
Internet-Draft
Intended status: Standards Track
Expires: August 17, 2014

P. Psenak, Ed.
S. Previdi, Ed.
C. Filsfils
Cisco Systems, Inc.
H. Gredler
Juniper Networks, Inc.
R. Shakir
British Telecom
W. Henderickx
Alcatel-Lucent

February 13, 2014

OSPFv3 Extensions for Segment Routing
draft-psenak-ospf-segment-routing-ospfv3-extension-01

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 OSPFv3 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].

Copyright Notice

Table of Contents

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). Prefix segments represent an ecmp-aware shortest-path to a prefix (or a node), 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 OSPFv3 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 defines various types of Segment Identifiers (SIDs): Prefix-SID, Adjacency-SID, LAN Adjacency SID and Binding SID.

2.1. SID/Label sub-TLV

SID/Label sub-TLV appears in multiple TLVs or Sub-TLVs defined later in this document. It is used to advertise SID or label associated with the prefix or adjacency. SID/Label TLV 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      SID/Label (variable)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: 1

Length: variable, 3 or 4 bytes

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

3. Segment Routing Capabilities

Segment Routing requires some additional capabilities of the router to be advertised to other routers in the area.

These SR capabilities are advertised in OSPFv3 Router Information Opaque LSA (defined in [RFC4970]).

3.1. SR-Algorithm TLV

SR-Algorithm TLV is a TLV of Router Information Opaque LSA (defined in [RFC4970]).

Router may use various algorithms when calculating reachability to other nodes in area or to prefixes attached to these nodes. Examples of these algorithms are metric based Shortest Path First (SPF), various sorts of Constrained SPF, etc. SR-Algorithm TLV allows a router to advertise algorithms that router is currently using to other routers in an area. SR-Algorithm TLV has following structure:

```
 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... |  Algorithm n |                 |
+-                                                             -+
|                                                               |
+                                                               +
```

where:

Type: 8

Length: variable

Algorithm: one octet identifying the algorithm. The following value has been defined:

0: IGP metric based SPT.

RI LSA can be advertised at any of the defined flooding scopes (link, area, or autonomous system (AS)). For the purpose of the SR-Algorithm TLV propagation area scope flooding is required.

3.2. SID/Label Range TLV

The SID/Label Range TLV is a TLV of Router Information Opaque LSA (defined in [RFC4970]).

SID/Label Sub-TLV MAY appear multiple times 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          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Range Size               |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Sub-TLVs (variable)                       |
+-                                                             -+
|                                                               |
+                                                               +
```

where:

Type: 9

Length: variable

Range Size: 3 octets of SID/label range

Currently the only supported Sub-TLV is the SID/Label TLV as defined in Section 2.1. SID/Label advertised in SID/Label TLV represents the first SID/Label from the advertised range.

RI LSA can be advertised at any of the defined flooding scopes (link, area, or autonomous system (AS)). For the purpose of the SR-Capability TLV propagation area scope flooding is required.

4. Prefix SID Identifier

A new extended OSPFv3 LSAs as defined in [I-D.acee-ospfv3-lsa-extend] are used to advertise prefix SID in OSPFv3.

4.1. Prefix SID Sub-TLV

The Prefix SID Sub-TLV is a Sub-TLV of the following OSPFv3 TLVs as defined in [I-D.acee-ospfv3-lsa-extend]:

Inter-Area Prefix TLV

External Prefix TLV

Intra-Area-Prefix TLV

It MAY appear more than once 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     |  Algorithm    |          Range Size           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              Index                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: 2.

Length: A 16-bit field that indicates the length of the value portion in octets. Set to 8.

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

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

where:

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.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.

M-Flag: Mapping Server Flag. If set, the SID is advertised from the Segment Routing Mapping Server functionality as described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

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

Algorithm: one octet identifying the algorithm the Prefix-SID is associated with as defined in Section 3.1.

Range Size: this field provides the ability to specify a range of addresses and their associated Prefix SIDs. It represents a compression scheme to distribute a continuous Prefix and their continuous, corresponding SID/Label Block. If a single SID is advertised then the Range Size field MUST be set to 1. For range advertisements > 1, Range Size represents the the number of addresses that need to be mapped into a Prefix-SID.

Index: 32 bits representing the offset to the advertised SID/Label range.

If multiple Prefix-SIDs are advertised for the same prefix, the receiving router MUST use the first encoded SID and MAY use the subsequent ones.

P-Flag (no-PHP) MUST be set on the Prefix-SIDs allocated to inter-area prefixes that are originated by the ABR based on intra-area or inter-area reachability between areas. In case the inter-area prefix is generated based on the prefix which is directly attached to the ABR, P-Flag SHOULD NOT be set

P-Flag (no-PHP) MUST NOT be set on the Prefix-SIDs allocated to redistributed prefixes, unless the redistributed prefix is directly attached to ASBR, in which case the P-Flag SHOULD NOT be set.

When M-Flag is set, PHP MUST NOT be done.

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

```
Router-A: 192::1/128, Prefix-SID: Index 1
Router-B: 192::2/128, Prefix-SID: Index 2
Router-C: 192::3/128, Prefix-SID: Index 3
Router-D: 192::4/128, Prefix-SID: Index 4
```

then the Address Prefix field in Intra-Area Prefix TLV, Inter-Area Prefix TLV or External Prefix TLV is set to 192::1, Prefix Length in these TLVs would be set to 128, Range Size in Prefix SID sub-TLV would be set to 4 and Index value would be set to 1.

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

```
10:1:1::0/120,   Prefix-SID: Index 51
10:1:1::100/120, Prefix-SID: Index 52
10:1:1::200/120, Prefix-SID: Index 53
10:1:1::300/120, Prefix-SID: Index 54
10:1:1::400/120, Prefix-SID: Index 55
10:1:1::500/120, Prefix-SID: Index 56
10:1:1::600/120, Prefix-SID: Index 57
```

then the Address Prefix field in Intra-Area Prefix TLV, Inter-Area Prefix TLV or External Prefix TLV is set to 10:1:1::0, Prefix Length in these TLVs would be set to 120, Range Size in Prefix SID sub-TLV would be set to 7 and Index value would be set to 51.

4.2. SID/Label Binding sub-TLV

SID/Label Binding sub-TLV is used to advertise SID/Label mapping for a prefix or a path to the prefix. SID/Label value advertised in this sub-TLV has local significance (to the router).

SID/Label Binding sub-TLV is a sub-TLV of the following OSPFv3 TLVs, as defined in [I-D.acee-ospfv3-lsa-extend]:

Inter-Area Prefix TLV

External Prefix TLV

Intra-Area-Prefix TLV

Multiple SID/Label Binding sub-TLVs can be present in above mentioned TLVs. SID/Label Binding sub-TLV 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     |   Weight      |          Range Size           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sub-TLVs (variable)                     |
+-                                                             -+
|                                                               |
```

where:

Type: 3

Length: variable

Flags: 1 octet field of following flags:

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

where:

M-bit - When the bit is set the binding represents the mirroring context as defined in [I-D.minto-rsvp-lsp-egress-fast-protection].

Weight: weight used for load-balancing purposes. The use of the weight is defined in [I-D.filsfils-rtgwg-segment-routing].

Range Size: usage is the same as described in Section 4.1

SID/Label Binding sub-TLV currently supports following Sub-TLVs:

SID/Label sub-TLV as described in Section 2.1. This sub-TLV MUST appear in the SID/Label Binding Sub-TLV and it MUST only appear once.

ERO Metric sub-TLV as defined in Section 4.2.1.

ERO sub-TLVs as defined in Section 4.2.2.

4.2.1. ERO Metric sub-TLV

ERO Metric sub-TLV is a Sub-TLV of the SID/Label Binding TLV.

The ERO Metric sub-TLV carries the cost of an ERO path. It is used to compare the cost of a given source/destination path. A router SHOULD advertise the ERO Metric sub-TLV. The cost of the ERO Metric sub-TLV SHOULD be set to the cumulative IGP or TE path cost of the advertised ERO. Since manipulation of the Metric field may attract or distract traffic from and to the advertised segment it MAY be manually overridden.

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Metric (4 octets)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

ERO Metric sub-TLV format

where:

Type: 12

Length: 4 bytes

Metric: 4 bytes

4.2.2. ERO sub-TLVs

All 'ERO' information represents an ordered set which describes the segments of a path. The last ERO sub-TLV describes the segment closest to the egress point, contrary the first ERO sub-TLV describes the first segment of a path. If a router extends or stitches a path it MUST prepend the new segments path information to the ERO list.

The above similarly applies to backup EROs.

All ERO Sub-TLVs must immediately follow the (SID)/Label Sub-TLV.

All Backup ERO sub-TLVs must immediately follow last ERO Sub-TLV.

4.2.2.1. IPv4 ERO sub-TLV

IPv4 ERO sub-TLV is a sub-TLV of the SID/Label Binding sub-TLV.

The IPv4 ERO sub-TLV describes a path segment using IPv4 Address style of encoding. Its semantics have been borrowed from [RFC3209].

```
 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     |                   Reserved                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     IPv4 Address (4 octets)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv4 ERO sub-TLV format

where:

Type: 4

Length: 8 bytes

Flags: 1 octet field of following flags:

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

where:

L-bit - If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

IPv4 Address - the address of the explicit route hop.

4.2.2.2. IPv6 ERO sub-TLV

IPv6 ERO sub-TLV is a sub-TLV of the SID/Label Binding sub-TLV.

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

```
 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     |                    Reserved                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+-                                                             -+
|                                                               |
+-                         IPv6 Address                        -+
|                                                               |
+-                                                             -+
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv6 ERO sub-TLV format

where:

Type: 5

Length: 8 bytes

Flags: 1 octet field of following flags:

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

where:

L-bit - If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

IPv6 Address - the address of the explicit route hop.

4.2.2.3. Unnumbered Interface ID ERO sub-TLV

Unnumbered Interface ID ERO sub-TLV is a sub-TLV of the SID/Label Binding sub-TLV.

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

The Unnumbered Interface-ID ERO sub-TLV 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.

```
 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     |                  Reserved                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Router ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Interface ID                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Unnumbered Interface ID ERO sub-TLV format

Type: 6

Length: 12 bytes

Flags: 1 octet field of following flags:

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

where:

L-bit - If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

Router-ID: Router-ID of the next-hop.

Interface ID: is the identifier assigned to the link by the router specified by the Router-ID.

4.2.2.4. IPv4 Backup ERO sub-TLV

IPv4 Prefix Backup ERO sub-TLV is a sub-TLV of the SID/Label Binding sub-TLV.

The IPv4 Backup ERO sub-TLV describes a path segment using IPv4

Address style of encoding. Its semantics have been borrowed from [RFC3209].

```
 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         |                     Reserved                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    IPv4 Address (4 octets)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv4 Backup ERO sub-TLV format

where:

Type: 7

Length: 8 bytes

Flags: 1 octet field of following flags:

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

where:

L-bit - If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

IPv4 Address - the address of the explicit route hop.

4.2.2.5. IPv6 Backup ERO sub-TLV

IPv6 ERO sub-TLV is a sub-TLV of the SID/Label Binding sub-TLV.

The IPv6 Backup ERO sub-TLV 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             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Flags     |                   Reserved                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+-                                                             -+
|                                                               |
+-                        IPv6 Address                         -+
|                                                               |
+-                                                             -+
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv6 Backup ERO sub-TLV format

where:

Type: 8

Length: 8 bytes

Flags: 1 octet field of following flags:

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

where:

L-bit - If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

IPv6 Address - the address of the explicit route hop.

4.2.2.6. Unnumbered Interface ID Backup ERO sub-TLV

Unnumbered Interface ID Backup sub-TLV is a sub-TLV of the SID/Label Binding sub-TLV.

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

The Unnumbered Interface-ID ERO sub-TLV 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.

```
 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     |                    Reserved                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Router ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Interface ID                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Unnumbered Interface ID Backup ERO sub-TLV format

where:

Type: 9

Length: 12 bytes

Flags: 1 octet field of following flags:

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

where:

L-bit - If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

Router-ID: Router-ID of the next-hop.

Interface ID: is the identifier assigned to the link by the router specified by the Router-ID.

5. Adjacency Segment Identifier (Adj-SID)

An Adjacency Segment Identifier (Adj-SID) represents a router adjacency in Segment Routing. At the current stage of Segment Routing architecture it is assumed that the Adj-SID value has local

significance (to the router).

5.1. Adj-SID sub-TLV

A new extended OSPFv3 LSAs, as defined in [I-D.acee-ospfv3-lsa-extend], are used to advertise prefix SID in OSPFv3

Adj-SID sub-TLV is an optional sub-TLV of the Router-Link TLV as defined in [I-D.acee-ospfv3-lsa-extend]. It MAY appear multiple times in Router-Link TLV. Examples where more than one Adj-SID may be used per neighbor are described in [I-D.filsfils-rtgwg-segment-routing-use-cases]. The structure of the 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    |             Reserved          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sub-TLVs (variable)                     |
+-                                                             -+
|                                                               |
```

where:

Type: 10.

Length: variable.

Flags. 1 octet field of following flags:

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

where:

B-Flag: Backup-flag: set if the Adj-SID refer 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: weight used for load-balancing purposes. The use of the weight is defined in [I-D.filsfils-rtgwg-segment-routing].

Adj-SID Sub-TLV supports following Sub-TLVs:

SID/Label sub-TLV as described in Section 2.1. This TLV MUST appear in the Adj-SID sub-TLV and it MUST only appear once.

A SR capable router MAY allocate an Adj-SID for each of its adjacencies and 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].

5.2. LAN Adj-SID Sub-TLV

LAN Adj-SID is an optional sub-TLV of the Router-Link TLV. It MAY appear multiple times in Router-Link TLV. It is used to advertise SID/Label for adjacency to non-DR node on broadcast or NBMA network.

```
 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    |             Reserved          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Neighbor ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sub-TLVs (variable)                     |
+-                                                             -+
|                                                               |
```

where:

Type: 11.

Length: variable.

Flags. 1 octet field of following flags:

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

where:

B-Flag: Backup-flag: set if the LAN-Adj-SID refer 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: weight used for load-balancing purposes. The use of the weight is defined in [I-D.filsfils-rtgwg-segment-routing].

LAN Adj-SID Sub-TLV supports following Sub-TLVs:

SID/Label sub-TLV as described in Section 2.1. This TLV MUST appear in the Adj-SID sub-TLV and it MUST only appear once.

6. Elements of Procedure

6.1. Intra-area Segment routing in OSPFv3

The OSPFv3 node that supports segment routing MAY advertise Prefix-SIDs for any prefix that it is advertising reachability for (e.g. loopback IP address) as described in Section 4.1.

If multiple routers advertise Prefix-SID for the same prefix, then the Prefix-SID MUST be the same. This is required in order to allow traffic load-balancing if multiple equal cost paths to the destination exist in the network.

Prefix-SID can also be advertised by the SR Mapping Servers (as described in [I-D.filsfils-rtgwg-segment-routing-use-cases]). The Mapping Server advertises Prefix-SID for remote prefixes that exist in the network. Multiple Mapping Servers can advertise Prefix-SID for the same prefix, in which case the same Prefix-SID MUST be advertised by all of them. SR Mapping Server could use either area scope or autonomous system flooding scope when advertising Prefix SID for prefixes, based on the configuration of the SR Mapping Server. Depending on the flooding scope used, SR Mapping Server chooses the LSA that will be used. If the area flooding scope is needed, E-Intra-Area-Prefix-LSA ([I-D.acee-ospfv3-lsa-extend]) is used. If autonomous system flooding scope is needed, E-AS-External-LSA ([I-D.acee-ospfv3-lsa-extend]) is used.

When Prefix-SID is advertised by the Mapping Server, which is indicated by the M-flag in the Prefix-SID sub-TLV (Section 4.1), route-type as indicated by the LSA type which is being used for

flooding is ignored. Prefix SID is bound to a prefix, in which case route-type becomes unimportant.

Advertisement of the Prefix-SID by the Mapping Server using Inter-Area Prefix TLV, External Prefix TLV or Intra-Area-Prefix TLV ([I-D.acee-ospfv3-lsa-extend]) does not itself contribute to the prefix reachability. NU-bit MUST be set in the PrefixOptions field of the LSA which is used by the Mapping Server to advertise SID or SID range, which prevents such advertisement to contribute to the prefix reachability.

6.2. Inter-area Segment routing in OSPFv3

In order to support SR in a multi-area environment, OSPFv3 must propagate Prefix-SID information between areas. The following procedure is used in order to propagate Prefix SIDs between areas.

When an OSPFv3 ABR advertises a Inter-Area-Prefix-LSA from an intra-area prefix to all its connected areas, it will also include Prefix-SID sub-TLV, as described in Section 4.1. The Prefix-SID value will be set as follows:

The ABR will look at its best path to the prefix in the source area and find out the advertising router associated with its best path to that prefix.

If no Prefix-SID was advertised for the prefix in the source area by the router that contributes to the best path to the prefix, then the ABR will use the Prefix-SID advertised by any other router (e.g.: a Prefix-SID coming from an SR Mapping Server as defined in [I-D.filsfils-rtgwg-segment-routing-use-cases]) when propagating Prefix-SID for the prefix to other areas.

When an OSPFv3 ABR advertises Inter-Area-Prefix-LSA LSAs from an inter-area route to all its connected areas it will also include Prefix-SID sub-TLV, as described in Section 4.1. The Prefix-SID value will be set as follows:

The ABR will look at its best path to the prefix in the source area and find out the advertising router associated with its best path to that prefix.

The ABR will then look if such router advertised a Prefix-SID for the prefix and use it when advertising the Prefix-SID to other connected areas.

If no Prefix-SID was advertised for the prefix in the source area by the ABR that contributes to the best path to the prefix, the originating ABR will use the Prefix-SID advertised by any other router (e.g.: a Prefix-SID coming from an SR Mapping Server as defined in [I-D.filsfils-rtgwg-segment-routing-use-cases]) when propagating Prefix-SID for the prefix to other areas.

6.3. SID for External Prefixes

AS-External-LSAs are flooded domain wide. When an ASBR, which supports SR, generates AS-External-LSA, it should also include Prefix-SID sub-TLV, as described in Section 4.1 Prefix-SID value will be set to the SID that has been reserved for that prefix.

When a NSSA ASBR translates NSSA-LSA into AS-External-LSA, it should also advertise the Prefix-SID for the prefix. The NSSA ABR determines its best path to the prefix advertised in the translated NSSA-LSA and finds the advertising router associated with such path. If such advertising router has advertised a Prefix-SID for the prefix, then the NSSA ASBR uses it when advertising the Prefix-SID in AS-External-LSA. Otherwise the Prefix-SID advertised by any other router will be used (e.g.: a Prefix-SID coming from an SR Mapping Server as defined in [I-D.filsfils-rtgwg-segment-routing-use-cases]).

6.4. Advertisement of Adj-SID

The Adjacency Segment Routing Identifier (Adj-SID) is advertised using the Adj-SID Sub-TLV as described in Section 5.

6.4.1. Advertisement of Adj-SID on Point-to-Point Links

Adj-SID MAY be advertised for any adjacency on p2p link that is in a state 2-Way or higher. If the adjacency on a p2p link transitions from the FULL state, then the Adj-SID for that adjacency MAY be removed from the area. If the adjacency transitions to a state lower then 2-Way, then the Adj-SID MUST be removed from the area.

6.4.2. Adjacency SID on Broadcast or NBMA Interfaces

Broadcast or NBMA networks in OSPFv3 are represented by a star topology where the Designated Router (DR) is the central point all other routers on the broadcast or NBMA network connect to. As a result, routers on the broadcast or NBMA network advertise only their adjacency to DR and BDR. Routers that are neither DR nor BDR do not form and do not advertise adjacencies between them. They, however, maintain a 2-Way adjacency state between them.

When Segment Routing is used, each router on the broadcast or NBMA

network MAY advertise the Adj-SID for its adjacency to DR using Adj-SID Sub-TLV as described in Section 5.1.

SR capable router MAY also advertise Adj-SID for other neighbors (e.g. BDR, DR-OTHER) on broadcast or NBMA network using the LAN ADJ-SID Sub-TLV as described in section 5.1.1.2. Section 5.2.

7. IANA Considerations

This specification updates OSPFv3 Extend-LSA sub-TLV registry and allocates following sub-TLVs:

o 1 - SID/Label sub-TLV

o 2 - Prefix SID sub-TLV

o 3 - SID/Label Binding sub-TLV

o 4 - IPv4 ERO sub-TLV

o 5 - IPv6 ERO sub-TLV

o 6 - Unnumbered Interface ID ERO sub-TLV

o 7 - IPv4 Backup ERO sub-TLV

o 8 - IPv6 Backup ERO sub-TLV

o 9 - Unnumbered Interface ID Backup ERO sub-TLV

o 10 - Adj-SID sub-TLV

o 11 - LAN Adj-SID sub-TLV

o 12 - ERO Metric sub-TLV

8. Security Considerations

Implementations must assure that malformed permutations of the newly defined sub-TLvs do not result in errors which cause hard OSPFv3 failures.

9. Contributors

The following people gave a substantial contribution to the content

of this document: Ahmed Bashandy, Martin Horneffer, Bruno Decraene, Stephane Litkowski, Igor Milojevic, Rob Shakir and Saku Ytti.

10. Acknowledgements

We would like to thank Anton Smirnov for his contribution.

Many thanks to Yakov Rekhter, John Drake and Shraddha Hedge for their contribution on earlier incarnations of the "Binding / MPLS Label TLV" in [I-D.gredler-ospf-label-advertisement].

11. References

11.1. Normative References

[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.

[RFC4970] Lindem, A., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 4970, July 2007.

11.2. Informative References

[I-D.acee-ospfv3-lsa-extend]
Lindem, A., Mirtorabi, S., Roy, A., and F. Baker, "OSPFv3 LSA Extendibility", draft-acee-ospfv3-lsa-extend-02 (work in progress), September 2013.

[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-01 (work in progress), October 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., Kini, S., and E. Crabbe, "Segment Routing Use Cases", draft-filsfils-rtgwg-segment-routing-use-cases-02 (work in progress), October 2013.

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

[I-D.minto-rsvp-lsp-egress-fast-protection]
Jeganathan, J., Gredler, H., and Y. Shen, "RSVP-TE LSP egress fast-protection", draft-minto-rsvp-lsp-egress-fast-protection-03 (work in progress), November 2013.

Authors' Addresses

Peter Psenak (editor)
Cisco Systems, Inc.
Apollo Business Center
Mlynske nivy 43
Bratislava 821 09
Slovakia

Email: ppsenak@cisco.com

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

Email: sprevidi@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels,
Belgium

Email: cfilsfil@cisco.com

Hannes Gredler
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: hannes@juniper.net

Rob Shakir
British Telecom
London
UK

Email: rob.shakir@bt.com

Wim Henderickx
Alcatel-Lucent
Copernicuslaan 50
Antwerp 2018
BE

Email: wim.henderickx@alcatel-lucent.com