Open Shortest Path First IGP
Internet-Draft
Intended status: Standards Track
Expires: January 13, 2014

P. Psenak, Ed.
S. Previdi, Ed.
C. Filsfils
Cisco Systems, Inc.
H. Gredler
Juniper Networks, Inc.
R. Shakir
British Telecom
July 12, 2013

OSPF Extensions for Segment Routing
draft-psenak-ospf-segment-routing-extensions-02

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 OSPF 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

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 OSPF 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.

For the purpose of the advertisements of various SID values new Opaque LSAs (defined in [RFC5250]) are defined. These new LSAs are defined as generic containers that can be used in order to advertise any additional attributes associated with the prefix or link. These new Opaque LSAs are complementary to the existing LSAs and are not aimed to replace any of the existing LSAs.

2.1. SID/Label TLV

SID/Label TLV appears as Sub-TLV 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/Lable 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: TBA

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 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: TBA

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: TBA

Length: variable

Range Size: size of the 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. OSPFv2 Extended Prefix Opaque LSA type

A new Opaque LSA (defined in [RFC5250]) is defined in OSPFv2 in order to advertise additional prefix attributes: OSPFv2 Extended Prefix Opaque LSA.

Multiple OSPFv2 Extended Prefix Opaque LSAs can be advertised by a single router. Flooding scope of the OSPFv2 Extended Prefix Opaque LSA depends on the content inside the LSA and is in control of the originating router.

The format of the OSPFv2 Extended Prefix Opaque LSA 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            LS age             |     Options   |  9, 10, or 11 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Opaque type  |                   Instance                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+-                            TLVs                             -+
|                             ...                               |
```

Opaque type used by OSPFv2 Extended Prefix Opaque LSA is TBA.

The format of the TLVs within the body of the LSA is the same as the

format used by the Traffic Engineering Extensions to OSPF defined in [RFC3630]. The LSA payload consists of one or more nested Type/ Length/Value (TLV) triplets. The format of each TLV is:

```
 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            |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Value...                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Length field defines the length of the value portion in octets. The TLV is padded to 4-octet alignment; padding is not included in the length field. Nested TLVs are also 32-bit aligned. Unrecognized types are ignored.

4.1. OSPF Extended Prefix TLV

The OSPF Extended Prefix TLV is used in order to advertise additional attributes associated with the prefix. Multiple OSPF Extended Prefix TLVs MAY be carried in each OSPFv2 Extended Prefix Opaque LSA, however all prefixes included in the single OSPFv2 Extended Prefix Opaque LSA MUST have the same flooding scope. The structure of the OSPF Extended Prefix 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Route Type   | Prefix Length |     AF        |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Address Prefix (variable)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Sub-TLVs (variable)                      |
+-                                                             -+
|                                                               |
```

where:

Type is TBA.

Length is variable

Route type: type of the OSPF route. Supported types are:

                  0 - unspecified
                  1 - intra-area
                  3 - inter-area
                  5 - external
                  7 - NSSA external

If the route type is 0 (unspecified) the information inside the OSPF External Prefix TLV applies to the prefix regardless of what route-type it is. This is useful when some prefix specific attributes are advertised by some external entity, which is not aware of the route-type associated with the prefix.

Prefix length: length of the prefix

AF: 0 - IPv4 unicast

Address Prefix: the prefix itself encoded as an even multiple of 32-bit words, padded with zeroed bits as necessary. This encoding consumes ((PrefixLength + 31) / 32) 32-bit words. The default route is represented by a prefix of length 0.

4.2. Prefix SID Sub-TLV

The Prefix SID Sub-TLV is a Sub-TLV of the OSPF Extended 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     |     MT-ID     |   Algorithm   |    Reserved   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Index                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBA.

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.

MT-ID: Multi-Topology ID (as defined in [RFC4915]).

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

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.

PHP flag MUST NOT be set on the Prefix-SIDs allocated to inter- area prefixes that are originated by the router based on intra-area or inter-area reachability between areas.

4.3. SID/Label Binding TLV

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

SID/Label Binding TLV is a Sub-TLV of the OSPF Extended Prefix TLV. Multiple SID/Label Binding TLVs can be present in OSPF Extended

Prefix TLV. SID/Label Binding 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       |    MT-ID      |    Weight     |    Reserved   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Sub-TLVs (variable)                         |
+-                                                             -+
|                                                               |
```

where:

Type: TBA

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].

MT-ID: Multi-Topology ID (as defined in [RFC4915]).

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

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

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

ERO TLVs as defined in Section 4.3.1.

4.3.1. ERO TLVs

All 'ERO' information represents an ordered set which describes the segments of a path. The last ERO TLV describes the segment closest to the egress point, contrary the first ERO 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 TLVs must immediately follow last ERO Sub-TLV.

4.3.1.1. IPv4 ERO TLV

IPv4 ERO TLV is a Sub-TLV of the SID/Lable Binding TLV.

The IPv4 ERO 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 TLV format

where:

Type: TBA

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.

4.3.1.2. Unnumbered Interface ID ERO TLV

Unnumbered Interface ID ERO TLV is a Sub-TLV of the SID/Lable Binding TLV.

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

The Unnumbered Interface-ID ERO 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 TLV format

Type: TBA

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.

4.3.1.3. IPv4 Backup ERO TLV

IPv4 Prefix Backup ERO TLV is a Sub-TLV of the SID/Lable Binding TLV.

The IPv4 Backup ERO 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 TLV format

where:

Type: TBA

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.

4.3.1.4. Unnumbered Interface ID Backup ERO TLV

Unnumbered Interface ID Backup TLV is a Sub-TLV of the SID/Lable Binding TLV.

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

The Unnumbered Interface-ID ERO 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 TLV format

where:

Type: TBA

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. OSPFv2 Extended Link Opaque LSA

A new Opaque LSA (defined in [RFC5250] is defined in OSPFv2 in order to advertise additional link attributes: the OSPFv2 Extended Link Opaque LSA.

The OSPFv2 Extended Link Opaque LSA has an area flooding scope. Multiple OSPFv2 Extended Link Opaque LSAs can be advertised by a single router in an area.

The format of the OSPFv2 Extended Link Opaque LSA 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            LS age             |     Options   |       10      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Opaque type  |                   Instance                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Advertising Router                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      LS sequence number                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+-                            TLVs                             -+
|                             ...                               |
```

Opaque type used by OSPFv2 Extended Link Opaque LSA is TBA

The format of the TLVs within the body of LSA is the same as the format used by the Traffic Engineering Extensions to OSPF defined in [RFC3630]. The LSA payload consists of one or more nested Type/Length/Value (TLV) triplets. The format of each TLV is:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Value...                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Length field defines the length of the value portion in octets. The TLV is padded to 4-octet alignment; padding is not included in the length field. Nested TLVs are also 32-bit aligned. Unrecognized types are ignored.

5.2. OSPFv2 Extended Link TLV

OSPFv2 Extended Link TLV is used in order to advertise various attributes of the link. It describes a single link and is constructed of a set of Sub-TLVs. There are no ordering requirements for the Sub-TLVs. Only one Extended Link TLV SHALL be carried in each Extended Link Opaque LSA, allowing for fine granularity changes in the topology.

The Extended Link 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Link-Type |                Reserved                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Link ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Link Data                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sub-TLVs (variable)                     |
+-                                                             -+
|                                                               |
```

where:

Type is TBA.

Length is variable.

Link-Type: as defined in section A.4.2 of [RFC2328].

Link-ID: as defined in section A.4.2 of [RFC2328].

Link Data: as defined in section A.4.2 of [RFC2328].

5.3. Adj-SID sub-TLV

Adj-SID is an optional Sub-TLV of the Extended Link TLV. It MAY appear multiple times in Extended 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     |    MT-ID      |  Weight       |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Sub-TLVs (variable)                   |
+-                                                             -+
|                                                               |
```

where:

Type: TBA.

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.

MT-ID: Multi-Topology ID (as defined in [RFC4915].

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 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].

5.4. LAN Adj-SID/Label Sub-TLV

LAN Adj-SID is an optional Sub-TLV of the Extended Link TLV. It MAY appear multiple times in Extended 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     |     MT-ID     |    Weight     |    Reserved   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Neighbor ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
|                         Sub-TLVs (variable)                   |
+-                                                             -+
|                                                               |
```

where:

Type: TBA.

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.

MT-ID: Multi-Topology ID (as defined in [RFC4915].

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 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 OSPFv2

The OSPFv2 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 4.2.

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 advertise 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. Flooding scope of the OSPF Extended Prefix Opaque LSA that is generated by the SR Mapping Server could be either area scope or autonomous system scope and is decided based on the configuration of the SR Mapping Server.

6.2. Inter-area Segment routing in OSPFv2

In order to support SR in a multi-area environment, OSPFv2 must propagate Prefix-SID information between areas. The following procedure is used in order to propagate Prefix SIDs between areas.

When an OSPF ABR advertises a Type-3 Summary LSA from an intra-area prefix to all its connected areas, it will also originate an Extended Prefix Opaque LSA, as described in Section 4. The flooding scope of the Extended Prefix Opaque LSA type will be set to area-scope. The route-type in OSPF Extended Prefix TLV is set to inter-area. The Prefix-SID Sub-TLV will be included in this LSA and 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 OSPF ABR advertises Type-3 Summary LSAs from an inter-area route to all its connected areas it will also originate an Extended Prefix Opaque LSA, as described in Section 4. The flooding scope of the Extended Prefix Opaque LSA type will be set to area-scope. The route-type in OSPF Extended Prefix TLV is set to inter-area. The Prefix-SID Sub-TLV will be included in this LSA and the Prefix-SID 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

Type-5 LSAs are flooded domain wide. When an ASBR, which supports SR, generates Type-5 LSAs, it should also originate Extended Prefix Opaque LSAs, as described in Section 4. The flooding scope of the Extended Prefix Opaque LSA type is set to AS-scope. The route-type in OSPF Extended Prefix TLV is set to external. Prefix-SID Sub-TLV is included in this LSA and the Prefix-SID value will be set to the SID that has been reserved for that prefix.

When a NSSA ASBR translates Type-7 LSAs into Type-5 LSAs, it should also advertise the Prefix-SID for the prefix. The NSSA ABR determines its best path to the prefix advertised in the translated Type-7 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 for the Type-5 prefix. 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 OSPF 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 5.3.

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 5.4.

7. IANA Considerations

TBD

8. Manageability Considerations

TBD

9. Security Considerations

TBD

10. 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, Saku Ytti and Wim Henderickx.

11. 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].

12. References

12.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[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.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, June 2007.

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

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.

12.2. 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., 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 Use Cases", draft-filsfils-rtgwg-segment-routing-use-cases-00 (work in progress), June 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-02 (work in progress), April 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