

Open Shortest Path First IGP
Internet-Draft
Intended status: Standards Track
Expires: June 5, 2015

P. Psenak, Ed.
S. Previdi, Ed.
C. Filsfils
Cisco Systems, Inc.
H. Gredler
Juniper Networks, Inc.
R. Shakir
British Telecom
W. Henderickx
Alcatel-Lucent
J. Tantsura
Ericsson
December 2, 2014

OSPF Extensions for Segment Routing
draft-ietf-ospf-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 OSPF extensions required 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 June 5, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) 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	3
2.	Segment Routing Identifiers	3
2.1.	SID/Label Sub-TLV	4
3.	Segment Routing Capabilities	4
3.1.	SR-Algorithm TLV	4
3.2.	SID/Label Range TLV	5
4.	OSPF Extended Prefix Range TLV	7
5.	Prefix SID Sub-TLV	9
6.	SID/Label Binding Sub-TLV	12
6.1.	ERO Metric Sub-TLV	14
6.2.	ERO Sub-TLVs	14
6.2.1.	IPv4 ERO Sub-TLV	15
6.2.2.	Unnumbered Interface ID ERO Sub-TLV	16
6.2.3.	IPv4 Backup ERO Sub-TLV	17
6.2.4.	Unnumbered Interface ID Backup ERO Sub-TLV	18
7.	Adjacency Segment Identifier (Adj-SID)	19
7.1.	Adj-SID Sub-TLV	19
7.2.	LAN Adj-SID Sub-TLV	20
8.	Elements of Procedure	22
8.1.	Intra-area Segment routing in OSPFv2	22
8.2.	Inter-area Segment routing in OSPFv2	22
8.3.	SID for External Prefixes	23
8.4.	Advertisement of Adj-SID	24
8.4.1.	Advertisement of Adj-SID on Point-to-Point Links	24
8.4.2.	Adjacency SID on Broadcast or NBMA Interfaces	24
9.	IANA Considerations	24
9.1.	OSPF OSPF Router Information (RI) TLVs Registry	25
9.2.	OSPF Extended Prefix LSA TLV Registry	25
9.3.	OSPF Extended Prefix LSA Sub-TLV Registry	25

9.4.	OSPF Extended Link LSA Sub-TLV Registry	25
10.	Security Considerations	25
11.	Contributors	26
12.	Acknowledgements	26
13.	References	26
13.1.	Normative References	26
13.2.	Informative References	27
	Authors' Addresses	27

[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 cases, is a one-hop path. SR's control-plane can be applied to both IPv6 and MPLS data-planes, and does not require any additional signalling (other than IGP extensions). For example, when used in MPLS networks, SR paths do not require any LDP or RSVP-TE signalling. However, SR can interoperate in the presence of LSPs established with RSVP or LDP.

This draft describes the OSPF extensions required 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 [\[RFC5250\]](#) are defined in [\[I-D.ietf-ospf-prefix-link-attr\]](#). These new LSAs are defined as generic containers that can be used to advertise any additional attributes associated with a prefix or link. These new Opaque LSAs are complementary to the existing LSAs and are not aimed to replace any of the existing LSAs.

algorithms are metric based Shortest Path First (SPF), various flavors of Constrained SPF, etc. The SR-Algorithm TLV allows a router to advertise the algorithms that the router is currently using to other routers in an OSPF area. The SR-Algorithm 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                 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Algorithm 1 | Algorithm... | Algorithm n |                               |
+-+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

where:

Type: TBD, suggested value 8

Length: variable

Algorithm: Single octet identifying the algorithm. The following value is defined by this document:

0: IGP metric based Shortest Path Tree (SPT)

The RI LSA can be advertised at any of the defined opaque 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 top-level TLV of the Router Information Opaque LSA (defined in [RFC4970](#)).

The SID/Label Range TLV MAY appear multiple times and has the following format:

[illegible]

where:

Type: TBD, suggested value 9

Length: variable

Range Size: 3 octets of the SID/label range

Initially, the only supported Sub-TLV is the SID/Label TLV as defined in [Section 2.1](#). The SID/Label advertised in the SID/Label TLV represents the first SID/Label in the advertised range.

Multiple occurrence of the SID/Label Range TLV MAY be advertised, in order to advertise multiple ranges. In such case:

- o The originating router MUST encode each range into a different SID/Label Range TLV.
- o The originating router decides the order in which the set of SID/Label Range TLVs are advertised inside the Router Information Opaque LSA. The originating router MUST ensure the order is same after a graceful restart (using checkpointing, non-volatile storage or any other mechanism) in order to assure the SID/label range and SID index correspondence is preserved across graceful restarts.
- o The receiving router must adhere to the order in which the ranges are advertised when calculating a SID/label from a SID index.

The following example illustrates the advertisement of multiple ranges:

The originating router advertises following ranges:

Range 1: [100, 199]
Range 2: [1000, 1099]
Range 3: [500, 599]

The receiving routers concatenate the ranges and build the Segment Routing Global Block (SRGB) is as follows:

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

The indexes span multiple ranges:

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

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

4. OSPF Extended Prefix Range TLV

In some cases it is useful to advertise attributes for the range of prefixes. Segment Routing Mapping Server, which is described in [[I-D.filsfils-rtgwg-segment-routing](#)] is an example, where we need a single advertisement to advertise SIDs for multiple prefixes from a contiguous address range.

OSPF Extended Prefix Range TLV, which is a new top level TLV of the Extended Prefix LSA described in [[I-D.ietf-ospf-prefix-link-attr](#)] is defined for this purpose.

Multiple OSPF Extended Prefix Range TLVs MAY be advertised in each OSPF Extended Prefix Opaque LSA, but all prefix ranges included in a single OSPF Extended Prefix Opaque LSA MUST have the same flooding scope. The OSPF Extended Prefix Range TLV has the following format:

[illegible]

where:

Type: TBD, suggested value 2.

Length: variable

Prefix length: length of the prefix

AF: 0 - IPv4 unicast

Range size: represents the number of prefixes that are covered by the advertisement. The Range Size MUST NOT exceed the number of prefixes that could be satisfied by the prefix length without including the IPv4 multicast address range (224.0.0.0/3).

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

0	1	2	3	4	5	6	7
+	-	+	-	+	-	+	-
	IA						
+	-	+	-	+	-	+	-

where:

IA-Flag: Inter-Area flag. If set, advertisement is of inter-area type. ABR that is advertising the OSPF Extended Prefix Range TLV between areas MUST set this bit.

This bit is used to prevent redundant flooding of Prefix Range TLVs between areas as follows:

An ABR always prefers intra-area Prefix Range advertisement over inter-area one.

An ABR does not consider inter-area Prefix Range advertisements coming from non backbone area.

An ABR propagates inter-area Prefix Range advertisement from backbone area to connected non backbone areas only if such advertisement is considered to be the best one.

Address Prefix: the prefix, encoded as an even multiple of 32-bit words, padded with zeroed bits as necessary. This encoding consumes $((\text{PrefixLength} + 31) / 32)$ 32-bit words. The Address Prefix represents the first prefix in the prefix range.

5. Prefix SID Sub-TLV

The Prefix SID Sub-TLV is a Sub-TLV of the OSPF Extended Prefix TLV described in [[I-D.ietf-ospf-prefix-link-attr](#)] and the OSPF Extended Prefix Range TLV described in [Section 4](#). It MAY appear more than once in the parent TLV 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                |          Reserved              |
|                               |          MT-ID                 |
|                               |          Algorithm              |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |          SID/Index/Label (variable) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

Type: TBD, suggested value 2.

Length: variable

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

```

0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+
|N|NP|M|E|V|L|||
+---+---+---+---+---+---+---+

```

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 to Prefix-SIDs corresponding 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](#)].

NP-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](#)].

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

V-Flag: Value/Index Flag. If set, then the Prefix-SID carries an absolute value. If not set, then the Prefix-SID carries an index.

L-Flag: Local/Global Flag. If set, then the value/index carried by the Prefix-SID has local significance. If not set, then the value/index carried by this Sub-TLV has global significance.

Other bits: Reserved. These MUST be zero when sent and are 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](#).

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

A 32 bit index defining the offset in the SID/Label space advertised by this router.

A 24 bit label where the 20 rightmost bits are used for encoding the label value.

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

When propagating Prefix-SIDs between areas, if multiple prefix-SIDs are advertised for a prefix, an implementation SHOULD preserve the original order when advertising prefix-SIDs to other areas. This allows implementations that only support a single Prefix-SID to have a consistent view across areas.

When calculating the outgoing label for the prefix, the router MUST take into account E and P flags advertised by the next-hop router, if next-hop router advertised the SID for the prefix. This MUST be done regardless of whether the next-hop router contributes to the best path to the prefix.

The NP-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. When the inter-area prefix is generated based on the prefix which is directly attached to the ABR, NP-Flag SHOULD NOT be set

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

If the NP-Flag is not set then any upstream neighbor of the Prefix-SID originator MUST pop the Prefix-SID. This is equivalent to the penultimate hop popping mechanism used in the MPLS dataplane. In such case, MPLS EXP bits of the Prefix-SID are not preserved for the final destination (the Prefix-SID being removed). If the NP-flag is clear then the received E-flag is ignored.

If the NP-flag is set then:

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

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

When M-Flag is set, NP-flag MUST be set and E-bit MUST NOT be set.

When a Prefix-SID is advertised in an Extended Prefix Range TLV, then the value advertised in Prefix SID Sub-TLV is interpreted as a starting SID value.

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
```

then the Prefix field in the Extended Prefix Range TLV would be set to 192.0.2.1, Prefix Length would be set to 32, Range Size would be set to 4 and the Index value in the Prefix-SID Sub-TLV would be set to 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
```

then the Prefix field in the Extended Prefix Range TLV would be set to 10.1.1.0, Prefix Length would be set to 24, Range Size would be 7 and the Index value in the Prefix-SID Sub-TLV would be set to 51.

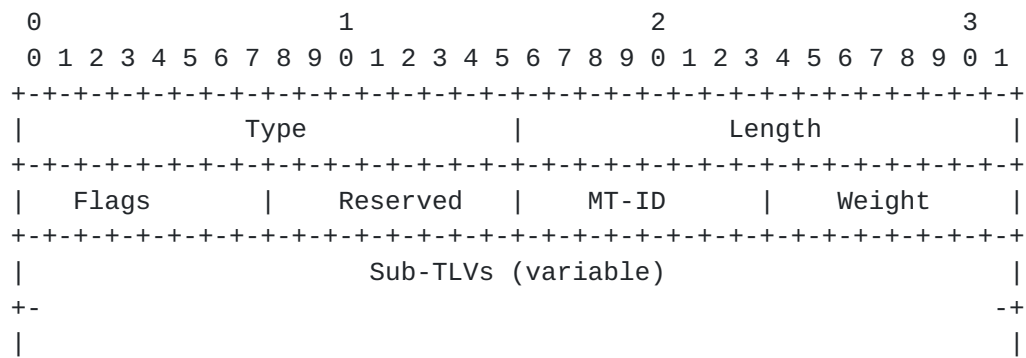
6. SID/Label Binding Sub-TLV

The SID/Label Binding Sub-TLV is used to advertise a SID/Label mapping for a path to the prefix.

The SID/Label Binding TLV MAY be originated by any router in an OSPF 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 in 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 anchoring SID/label bindings to both link and node IP addresses, any Label Switched Path (LSP) can be described. Additionally, SID/Label Bindings from external protocols can be easily re-advertised.

The SID/Label Binding TLV may be used for advertising SID/Label Bindings and their associated Primary and Backup paths. In a single TLV, a primary ERO Path, backup ERO Path, or both can be advertised. If a router wants to advertise multiple parallel paths, then it can generate several TLVs for the same Prefix/FEC. Each occurrence of a Binding TLV for a given FEC Prefix will add a new path.

The SID/Label Binding Sub-TLV is a Sub-TLV of the OSPF Extended Prefix TLV described in [[I-D.ietf-ospf-prefix-link-attr](#)] and the OSPF Extended Prefix Range TLV described in [Section 4](#). Multiple SID/Label Binding TLVs can be present in their parent TLV. The SID/Label Binding Sub-TLV has following format:

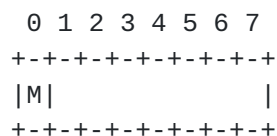


where:

Type: TBD, suggested value 3

Length: variable

Flags: 1 octet field of following flags:



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](#)].

All 'ERO' information represents an ordered set which describes the segments of a path. The first ERO Sub-TLV describes the first segment of a path. Similarly, the last ERO Sub-TLV describes the segment closest to the egress point. If a router extends or stitches

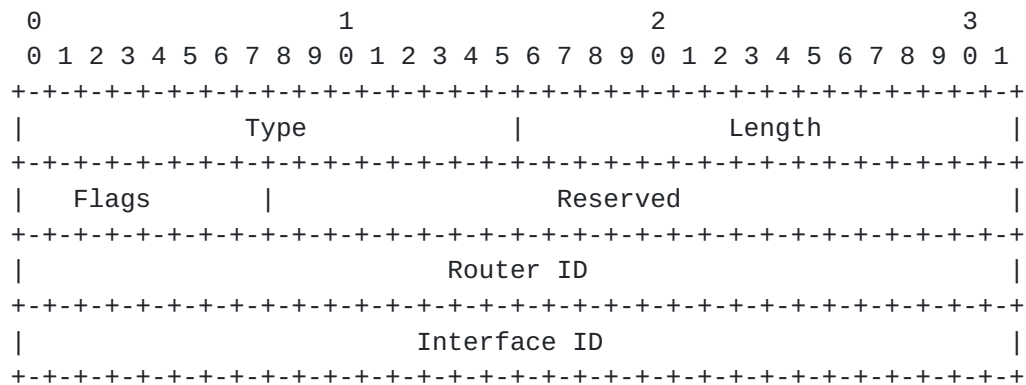
IPv4 Address - the address of the explicit route hop.

6.2.2. Unnumbered Interface ID ERO Sub-TLV

The 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 includes 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.



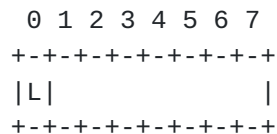
where:

Unnumbered Interface ID ERO Sub-TLV format

Type: TBD, suggested value 5

Length: 12 bytes

Flags: 1 octet field of following flags:



where:

L-bit - If the L-bit is set, then the segment path is designated as 'loose'. Otherwise, the segment path is designated as 'strict'.

IPv4 Address - the address of the explicit route hop.

L-bit - If the L-bit is set, then the segment path is designated as 'loose'. Otherwise, the segment path is designated as '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.

7. Adjacency Segment Identifier (Adj-SID)

An Adjacency Segment Identifier (Adj-SID) represents a router adjacency in Segment Routing.

7.1. Adj-SID Sub-TLV

Adj-SID is an optional Sub-TLV of the Extended Link TLV defined in [I-D.ietf-ospf-prefix-link-attr]. It MAY appear multiple times in the 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 Adj-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                   |      Reserved                |
|                               |      MT-ID                   |
|                               |      Weight                  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |                               |
|                               |                               |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

where:

Type: TBD, suggested value 2.

Length: variable.

Flags. 1 octet field of following flags:

```

0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+
|B|V|L|S|      |
+---+---+---+---+---+---+---+

```

where:

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].

The V-Flag: Value/Index Flag. If set, then the Prefix-SID carries an absolute value. If not set, then the Prefix-SID carries an index.

The L-Flag: Local/Global Flag. If set, then the value/index carried by the Prefix-SID has local significance. If not set, then the value/index carried by this Sub-TLV has global significance.

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

Other bits: Reserved. These MUST be zero when sent and are 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](#)].

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

A 32 bit index defining the offset in the SID/Label space advertised by this router.

A 24 bit label where the 20 rightmost bits are used for encoding the label value.

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

7.2. LAN Adj-SID Sub-TLV

LAN Adj-SID is an optional Sub-TLV of the Extended Link TLV defined in [[I-D.ietf-ospf-prefix-link-attr](#)]. It MAY appear multiple times in the Extended-Link TLV. It is used to advertise a SID/Label for an adjacency to a non-DR node on a broadcast or NBMA network.

[illegible]

where:

Type: TBD, suggested value 3.

Length: variable.

Flags. 1 octet field of following flags:

```

  0 1 2 3 4 5 6 7
+-+--+--+--+--+
|B|V|L|S|      |
+-+--+--+--+--+

```

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](#)].

The V-Flag: Value/Index Flag. If set, then the Prefix-SID carries an absolute value. If not set, then the Prefix-SID carries an index.

The L-Flag: Local/Global Flag. If set, then the value/index carried by the Prefix-SID has local significance. If not set, then the value/index carried by this Sub-TLV has global significance.

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

Other bits: Reserved. These MUST be zero when sent and are 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](#)].

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

A 32 bit index defining the offset in the SID/Label space advertised by this router.

A 24 bit label where the 20 rightmost bits are used for encoding the label value.

8. Elements of Procedure

8.1. Intra-area Segment routing in OSPFv2

An OSPFv2 router that supports segment routing MAY advertise Prefix-SIDs for any prefix to which it is advertising reachability (e.g., a loopback IP address as described in [Section 5](#)).

If multiple routers advertise a Prefix-SID for the same prefix, then the Prefix-SID MUST be the same. This is required in order to allow traffic load-balancing when 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-SIDs for remote prefixes that exist in the OSPFv2 routing domain. Multiple Mapping Servers can advertise Prefix-SIDs for the same prefix, in which case the same Prefix-SID MUST be advertised by all of them. The flooding scope of the OSPF Extended Prefix Opaque LSA that is generated by the SR Mapping Server could be either area scoped or AS scoped and is determined based on the configuration of the SR Mapping Server.

8.2. 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 [[I-D.ietf-ospf-prefix-link-attr](#)]. The flooding scope of the Extended Prefix Opaque LSA type will be set to area-scope. The route-type in the 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 the advertising router associated with the best path to that prefix.

The ABR will then determine 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 router 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 the 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 [[I-D.ietf-ospf-prefix-link-attr](#)]. 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 the advertising router associated with the best path to that prefix.

The ABR will then determine 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 the Prefix-SID for the prefix to other areas.

[8.3.](#) 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 an Extended Prefix Opaque LSAs, as described in [[I-D.ietf-ospf-prefix-link-attr](#)]. The flooding scope of the Extended Prefix Opaque LSA type is set to AS-scope. The route-type in the OSPF Extended Prefix TLV is set to external. The 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 an NSSA ABR 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 that path. If the advertising router has advertised a Prefix-SID for the prefix, then the NSSA ABR 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](#)]).

8.4. Advertisement of Adj-SID

The Adjacency Segment Routing Identifier (Adj-SID) is advertised using the Adj-SID Sub-TLV as described in [Section 7](#).

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

An Adj-SID MAY be advertised for any adjacency on a p2p link that is in neighbor 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 than 2-Way, then the Adj-SID advertisement MUST be removed from the area.

8.4.2. 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 to which all other routers on the broadcast or NBMA network connect. As a result, routers on the broadcast or NBMA network advertise only their adjacency to the DR. Routers that do not act as DR do not form or advertise adjacencies with each other. They do, however, maintain 2-Way adjacency state with each other and are directly reachable.

When Segment Routing is used, each router on the broadcast or NBMA network MAY advertise the Adj-SID for its adjacency to the DR using Adj-SID Sub-TLV as described in [Section 7.1](#).

SR capable routers MAY also advertise an Adj-SID for other neighbors (e.g. BDR, DR-OTHER) on the broadcast or NBMA network using the LAN ADJ-SID Sub-TLV as described in [Section 7.2](#).

9. IANA Considerations

This specification updates several existing OSPF registries.

9.1. OSPF OSPF Router Information (RI) TLVs Registry

- o 8 (IANA Preallocated) - SR-Algorithm TLV
- o 9 (IANA Preallocated) - SID/Label Range TLV

9.2. OSPF Extended Prefix LSA TLV Registry

Following values are allocated:

- o 2 - OSPF Extended Prefix Range TLV

9.3. OSPF Extended Prefix LSA Sub-TLV Registry

Following values are allocated:

- 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 - Unnumbered Interface ID ERO Sub-TLV
- o 6 - IPv4 Backup ERO Sub-TLV
- o 7 - Unnumbered Interface ID Backup ERO Sub-TLV
- o 8 - ERO Metric Sub-TLV

9.4. OSPF Extended Link LSA Sub-TLV Registry

Following initial values are allocated:

- o 1 - SID/Label Sub-TLV
- o 2 - Adj-SID Sub-TLV
- o 3 - LAN Adj-SID/Label Sub-TLV

10. Security Considerations

Implementations must assure that malformed TLV and Sub-TLV permutations do not result in errors which cause hard OSPF failures.

11. Contributors

The following people gave a substantial contribution to the content of this document: Acee Lindem, Ahmed Bashandy, Martin Horneffer, Bruno Decraene, Stephane Litkowski, Igor Milojevic, Rob Shakir and Saku Ytti.

12. 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](#)].

Thanks to Acee Lindem for the detail review of the draft, corrections, as well as discussion about details of the encoding.

13. References

13.1. 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.

13.2. 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-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.ietf-ospf-prefix-link-attr]

Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", [draft-ietf-ospf-prefix-link-attr-01](#) (work in progress), September 2014.

[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

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
US

Email: Jeff.Tantsura@ericsson.com