Stream: Internet Engineering Task Force (IETF)
RFC: 8665
Category: Standards Track
Published: December 2019
ISSN: 2070-1721
Authors:
P. Psenak, Ed. *Cisco Systems, Inc.*
S. Previdi, Ed. *Cisco Systems, Inc.*
C. Filsfils *Cisco Systems, Inc.*
H. Gredler *RtBrick Inc.*
R. Shakir *Google, Inc.*
W. Henderickx *Nokia*
J. Tantsura *Apstra, Inc.*

RFC 8665
OSPF Extensions for Segment Routing

Abstract

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

This document describes the OSPFv2 extensions required for Segment Routing.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc8665.

Copyright Notice

Table of Contents

1. Introduction

Segment Routing (SR) allows a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological subpaths 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 it does not require any additional signaling (other than IGP extensions). The IPv6 data plane is out of the scope of this specification; it is not applicable to OSPFv2, which only supports the IPv4 address family. When used in MPLS networks, SR paths do not require any LDP or RSVP-TE signaling. However, SR can interoperate in the presence of LSPs established with RSVP or LDP.

There are additional segment types, e.g., Binding Segment Identifier (SID) defined in [RFC8402].

This document describes the OSPF extensions required for Segment Routing.

Segment Routing architecture is described in [RFC8402].

Segment Routing use cases are described in [RFC7855].

1.1. Requirements Language

The key words "**MUST**", "**MUST NOT**", "**REQUIRED**", "**SHALL**", "**SHALL NOT**", "**SHOULD**", "**SHOULD NOT**", "**RECOMMENDED**", "**NOT RECOMMENDED**", "**MAY**", and "**OPTIONAL**" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Segment Routing Identifiers

Segment Routing defines various types of Segment Identifiers (SIDs): Prefix-SID, Adjacency SID, LAN Adjacency SID, and Binding SID.

Extended Prefix/Link Opaque Link State Advertisements (LSAs) defined in [RFC7684] are used for advertisements of the various SID types.

2.1. SID/Label Sub-TLV

The SID/Label Sub-TLV appears in multiple TLVs or sub-TLVs defined later in this document. It is used to advertise the SID or label associated with a prefix or adjacency. The SID/Label Sub-TLV has the following format:

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

where:

Type: 1

Length: 3 or 4 octets

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

3. Segment Routing Capabilities

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

These SR capabilities are advertised in the Router Information Opaque LSA (defined in [RFC7770]). The TLVs defined below are applicable to both OSPFv2 and OSPFv3; see also [RFC8666].

3.1. SR-Algorithm TLV

The SR-Algorithm TLV is a top-level TLV of the Router Information Opaque LSA (defined in [RFC7770]).

The SR-Algorithm TLV is optional. It **SHOULD** only be advertised once in the Router Information Opaque LSA. If the SR-Algorithm TLV is not advertised by the node, such a node is considered as not being Segment Routing capable.

An SR Router can use various algorithms when calculating reachability to OSPF routers or prefixes in an OSPF area. Examples of these 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 currently used by the router to other routers in an OSPF area. The SR-Algorithm 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Algorithm 1 | Algorithm...  |   Algorithm n |               |
+-                                                             -+
|                                                               |
+                                                               +
```

where:

Type: 8

Length: Variable, in octets, depending on the number of algorithms advertised

Algorithm: Single octet identifying the algorithm. The following values are defined by this document:

0: Shortest Path First (SPF) algorithm based on link metric. This is the standard shortest path algorithm as computed by the OSPF protocol. Consistent with the deployed practice for link-state protocols, Algorithm 0 permits any node to overwrite the SPF path with a different path based on its local policy. If the SR-Algorithm TLV is advertised, Algorithm 0 **MUST** be included.

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

When multiple SR-Algorithm TLVs are received from a given router, the receiver **MUST** use the first occurrence of the TLV in the Router Information Opaque LSA. If the SR-Algorithm TLV appears in multiple Router Information Opaque LSAs that have different flooding scopes, the SR-Algorithm TLV in the Router Information Opaque LSA with the area-scoped flooding scope **MUST** be used. If the SR-Algorithm TLV appears in multiple Router Information Opaque LSAs that have the same flooding scope, the SR-Algorithm TLV in the Router Information (RI) Opaque LSA with the numerically smallest Instance ID **MUST** be used and subsequent instances of the SR-Algorithm TLV **MUST** be ignored.

The RI LSA can be advertised at any of the defined opaque flooding scopes (link, area, or Autonomous System (AS)). For the purpose of SR-Algorithm TLV advertisement, area-scoped flooding is **REQUIRED**.

3.2. SID/Label Range TLV

Prefix-SIDs **MAY** be advertised in the form of an index as described in Section 5. Such an index defines the offset in the SID/Label space advertised by the router. The SID/Label Range TLV is used to advertise such SID/Label space.

The SID/Label Range TLV is a top-level TLV of the Router Information Opaque LSA (defined in [RFC7770]).

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

where:

Type: 9

Length: Variable, in octets, depending on the sub-TLVs

Range Size: 3-octet SID/label range size (i.e., the number of SIDs or labels in the range including the first SID/label). It **MUST** be greater than 0.

Reserved: **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

Initially, the only supported sub-TLV is the SID/Label Sub-TLV as defined in Section 2.1. The SID/Label Sub-TLV **MUST** be included in the SID/Label Range TLV. The SID/Label advertised in the SID/Label Sub-TLV represents the first SID/Label in the advertised range.

Only a single SID/Label Sub-TLV **MAY** be advertised in the SID/Label Range TLV. If more than one SID/Label Sub-TLV is present, the SID/Label Range TLV **MUST** be ignored.

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

- The originating router **MUST** encode each range into a different SID/Label Range TLV.
- 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 the same after a graceful restart (using checkpointing, nonvolatile storage, or any other mechanism) in order to ensure the SID/Label range and SID index correspondence is preserved across graceful restarts.

- The receiving router **MUST** adhere to the order in which the ranges are advertised when calculating a SID/Label from a SID index.
- The originating router **MUST NOT** advertise overlapping ranges.
- When a router receives multiple overlapping ranges, it **MUST** conform to the procedures defined in [RFC8660].

The following example illustrates the advertisement of multiple ranges.

The originating router advertises the following ranges:

```
       Range 1: Range Size: 100   SID/Label Sub-TLV: 100
       Range 1: Range Size: 100   SID/Label Sub-TLV: 1000
       Range 1: Range Size: 100   SID/Label Sub-TLV: 500
```

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

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

The indexes span multiple ranges:

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

The RI LSA can be advertised at any of the defined flooding scopes (link, area, or autonomous system (AS)). For the purpose of SID/Label Range TLV advertisement, area-scoped flooding is **REQUIRED**.

3.3. SR Local Block TLV

The SR Local Block TLV (SRLB TLV) contains the range of labels the node has reserved for Local SIDs. SIDs from the SRLB **MAY** be used for Adjacency SIDs but also by components other than the OSPF protocol. As an example, an application or a controller can instruct the router to allocate a specific Local SID. Some controllers or applications can use the control plane to discover the available set of Local SIDs on a particular router. In such cases, the SRLB is advertised in the control plane. The requirement to advertise the SRLB is further described in [RFC8660]. The SRLB TLV is used to advertise the SRLB.

The SRLB TLV is a top-level TLV of the Router Information Opaque LSA (defined in [RFC7770]).

The SRLB TLV **MAY** appear multiple times in the Router Information Opaque LSA 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Range Size                   |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Sub-TLVs (variable)                      |
+-                                                             -+
|                                                               |
+                                                               +
```

where:

Type: 14

Length: Variable, in octets, depending on the sub-TLVs

Range Size: 3-octet SID/Label range size (i.e., the number of SIDs or labels in the range including the first SID/Label). It **MUST** be greater than 0.

Reserved: **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

Initially, the only supported sub-TLV is the SID/Label Sub-TLV as defined in Section 2.1. The SID/Label Sub-TLV **MUST** be included in the SRLB TLV. The SID/Label advertised in the SID/Label Sub-TLV represents the first SID/Label in the advertised range.

Only a single SID/Label Sub-TLV **MAY** be advertised in the SRLB TLV. If more than one SID/Label Sub-TLV is present, the SRLB TLV **MUST** be ignored.

The originating router **MUST NOT** advertise overlapping ranges.

Each time a SID from the SRLB is allocated, it **SHOULD** also be reported to all components (e.g., controller or applications) in order for these components to have an up-to-date view of the current SRLB allocation. This is required to avoid collisions between allocation instructions.

Within the context of OSPF, the reporting of Local SIDs is done through OSPF sub-TLVs, such as the Adjacency SID (Section 6). However, the reporting of allocated Local SIDs can also be done through other means and protocols, which are outside the scope of this document.

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

The RI LSA can be advertised at any of the defined flooding scopes (link, area, or autonomous system (AS)). For the purpose of SRLB TLV advertisement, area-scoped flooding is **REQUIRED**.

3.4. SRMS Preference TLV

The Segment Routing Mapping Server Preference TLV (SRMS Preference TLV) is used to advertise a preference associated with the node that acts as an SR Mapping Server. The role of an SRMS is described in [RFC8661]. SRMS preference is defined in [RFC8661].

The SRMS Preference TLV is a top-level TLV of the Router Information Opaque LSA (defined in [RFC7770]).

The SRMS Preference TLV **MAY** only be advertised once in the Router Information Opaque LSA 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Preference  |                 Reserved                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: 15

Length: 4 octets

Preference: 1 octet, with an SRMS preference value from 0 to 255

Reserved: **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

When multiple SRMS Preference TLVs are received from a given router, the receiver **MUST** use the first occurrence of the TLV in the Router Information Opaque LSA. If the SRMS Preference TLV appears in multiple Router Information Opaque LSAs that have different flooding scopes, the SRMS Preference TLV in the Router Information Opaque LSA with the narrowest flooding scope **MUST** be used. If the SRMS Preference TLV appears in multiple Router Information Opaque LSAs that have the same flooding scope, the SRMS Preference TLV in the Router Information Opaque LSA with the numerically smallest Instance ID **MUST** be used and subsequent instances of the SRMS Preference TLV **MUST** be ignored.

The RI LSA can be advertised at any of the defined flooding scopes (link, area, or autonomous system (AS)). For the purpose of the SRMS Preference TLV advertisement, AS-scoped flooding **SHOULD** be used. This is because SRMS servers can be located in a different area than consumers of the SRMS advertisements. If the SRMS advertisements from the SRMS server are only used inside the SRMS server's area, area-scoped flooding **MAY** be used.

4. OSPF Extended Prefix Range TLV

In some cases, it is useful to advertise attributes for a range of prefixes. The SR Mapping Server, which is described in [RFC8661], is an example where we need a single advertisement to advertise SIDs for multiple prefixes from a contiguous address range.

The OSPF Extended Prefix Range TLV, which is a top-level TLV of the Extended Prefix LSA described in [RFC7684] 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:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Prefix Length |     AF        |           Range Size          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Flags       |                    Reserved                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Address Prefix (variable)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Sub-TLVs (variable)                      |
+-                                                             -+
|                                                               |
```

where:

Type: 2

Length: Variable, in octets, depending on the sub-TLVs

Prefix Length: Length of prefix in bits

AF: Address family for the prefix. Currently, the only supported value is 0 for IPv4 unicast. The inclusion of address family in this TLV allows for future extension.

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: Single-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. An Area Border Router (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 only propagates an inter-area Prefix Range advertisement from the backbone area to connected nonbackbone areas if the advertisement is considered to be the best one. The following rules are used to select the best range from the set of advertisements for the same Prefix Range:

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

An ABR does not consider inter-area Prefix Range advertisements coming from nonbackbone areas.

Reserved: **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

Address Prefix: For the address family IPv4 unicast, the prefix itself is encoded as a 32-bit value. The default route is represented by a prefix of length 0. Prefix encoding for other address families is beyond the scope of this specification.

5. Prefix-SID Sub-TLV

The Prefix-SID Sub-TLV is a sub-TLV of the OSPF Extended Prefix TLV described in [RFC7684] 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: 2

Length: 7 or 8 octets, depending on the V-Flag

Flags: Single-octet field. The following flags are defined:

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

where:

NP-Flag: No-PHP (Penultimate Hop Popping) Flag. If set, then the penultimate hop **MUST NOT** pop the Prefix-SID before delivering packets to the node that advertised the Prefix-SID.

M-Flag: Mapping Server Flag. If set, the SID was advertised by an SR Mapping Server as described in [RFC8661].

E-Flag: Explicit Null Flag. If set, any upstream neighbor of the Prefix-SID originator **MUST** replace the Prefix-SID with the Explicit NULL label (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.

Reserved: **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

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

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

A router receiving a Prefix-SID from a remote node and with an algorithm value that the remote node has not advertised in the SR-Algorithm TLV (Section 3.1) **MUST** ignore the Prefix-SID Sub-TLV.

SID/Index/Label: According to the V- and L-Flags, it contains:

V-Flag is set to 0 and L-Flag is set to 0: The SID/Index/Label field is a 4-octet index defining the offset in the SID/Label space advertised by this router.

V-Flag is set to 1 and L-Flag is set to 1: The SID/Index/Label field is a 3-octet local label where the 20 rightmost bits are used for encoding the label value.

All other combinations of V-Flag and L-Flag are invalid and any SID Advertisement received with an invalid setting for V- and L-Flags **MUST** be ignored.

If an OSPF router advertises multiple Prefix-SIDs for the same prefix, topology, and algorithm, all of them **MUST** be ignored.

When calculating the outgoing label for the prefix, the router **MUST** take into account, as described below, the E-, NP-, and M-Flags advertised by the next-hop router if that 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 and the E-Flag **MUST** be clear for Prefix-SIDs allocated to inter-area prefixes that are originated by the ABR based on intra-area or inter-area reachability between areas unless the advertised prefix is directly attached to the ABR.

The NP-Flag (No-PHP) **MUST** be set and the E-Flag **MUST** be clear for Prefix-SIDs allocated to redistributed prefixes, unless the redistributed prefix is directly attached to the Autonomous System Boundary Router (ASBR).

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 data plane.
>
> The received E-Flag is ignored.

If the NP-Flag is set and 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 needs to stitch the incoming packet into a continuing MPLS LSP to the final destination. This could occur at an ABR (prefix propagation from one area to another) or at an ASBR (prefix propagation from one domain to another).

If both the NP-Flag and E-Flag are set, then:

> Any upstream neighbor of the Prefix-SID originator **MUST** replace the Prefix-SID with an Explicit NULL label. 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 the M-Flag is set, the NP-Flag and the E-Flag **MUST** be ignored on reception.

As the Mapping Server does not specify the originator of a prefix advertisement, it is not possible to determine PHP behavior solely based on the Mapping Server Advertisement. However, PHP behavior **SHOULD** be done in the following cases:

> The Prefix is intra-area type and the downstream neighbor is the originator of the prefix.
>
> The Prefix is inter-area type and the downstream neighbor is an ABR, which is advertising prefix reachability and is also generating the Extended Prefix TLV with the A-Flag set for this prefix as described in Section 2.1 of [RFC7684].

The Prefix is external type and the downstream neighbor is an ASBR, which is advertising prefix reachability and is also generating the Extended Prefix TLV with the A-Flag set for this prefix as described in Section 2.1 of [RFC7684].

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

```
            192.0.2.0/30, Prefix-SID: Index 51
            192.0.2.4/30, Prefix-SID: Index 52
            192.0.2.8/30, Prefix-SID: Index 53
           192.0.2.12/30, Prefix-SID: Index 54
           192.0.2.16/30, Prefix-SID: Index 55
           192.0.2.20/30, Prefix-SID: Index 56
           192.0.2.24/30, Prefix-SID: Index 57
```

then the Prefix field in the Extended Prefix Range TLV would be set to 192.0.2.0, Prefix Length would be set to 30, Range Size would be 7, and the Index value in the Prefix-SID Sub-TLV would be set to 51.

6. Adjacency Segment Identifier (Adj-SID)

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

6.1. Adj-SID Sub-TLV

Adj-SID is an optional sub-TLV of the Extended Link TLV defined in [RFC7684]. It **MAY** appear multiple times in the Extended Link TLV. 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       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   SID/Label/Index (variable)                  |
+---------------------------------------------------------------+
```

where:

Type: 2

Length: 7 or 8 octets, depending on the V-Flag

Flags: Single-octet field containing the following flags:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|B|V|L|G|P|     |
+-+-+-+-+-+-+-+-+
```

where:

B-Flag: Backup Flag. If set, the Adj-SID refers to an adjacency that is eligible for protection (e.g., using IP Fast Reroute or MPLS-FRR (MPLS-Fast Reroute) as described in Section 2.1 of [RFC8402].

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

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

G-Flag: Group Flag. When set, the G-Flag indicates that the Adj-SID refers to a group of adjacencies (and therefore **MAY** be assigned to other adjacencies as well).

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

Other bits: Reserved. These **MUST** be zero when sent and are ignored when received.

Reserved: **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

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

Weight: Weight used for load-balancing purposes. The use of the weight is defined in [RFC8402].

SID/Index/Label: As described in Section 5

An SR-capable router **MAY** allocate an Adj-SID for each of its adjacencies and set the B-Flag when the adjacency is eligible for protection by an FRR mechanism (IP or MPLS) as described in Section 3.5 of [RFC8402].

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

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

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

6.2. LAN Adj-SID Sub-TLV

The LAN Adjacency SID is an optional sub-TLV of the Extended Link TLV defined in [RFC7684]. 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 (Designated Router) router on a broadcast, Non-Broadcast Multi-Access (NBMA), or hybrid [RFC6845] 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     |    Reserved   |     MT-ID     |    Weight     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Neighbor ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    SID/Label/Index (variable)                 |
+---------------------------------------------------------------+
```

where:

Type: 3

Length: 11 or 12 octets, depending on the V-Flag

Flags: Same as in Section 6.1

Reserved: **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

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

Weight: Weight used for load-balancing purposes. The use of the weight is defined in [RFC8402].

Neighbor ID: The Router ID of the neighbor for which the LAN Adjacency SID is advertised

SID/Index/Label: As described in Section 5

When the P-Flag is not set, the LAN Adjacency SID **MAY** be persistent. When the P-Flag is set, the LAN Adjacency SID **MUST** be persistent.

7. Elements of Procedure

7.1. 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).

A Prefix-SID can also be advertised by the SR Mapping Servers (as described in [RFC8661]). A 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.

An SR Mapping Server **MUST** use the OSPF Extended Prefix Range TLV when advertising SIDs for prefixes. Prefixes of different route types can be combined in a single OSPF Extended Prefix Range TLV advertised by an SR Mapping Server. Because the OSPF Extended Prefix Range TLV doesn't include a Route-Type field, as in the OSPF Extended Prefix TLV, it is possible to include adjacent prefixes from different route types in the OSPF Extended Prefix Range TLV.

Area-scoped OSPF Extended Prefix Range TLVs are propagated between areas. Similar to propagation of prefixes between areas, an ABR only propagates the OSPF Extended Prefix Range TLV that it considers to be the best from the set it received. The rules used to pick the best OSPF Extended Prefix Range TLV are described in Section 4.

When propagating an OSPF Extended Prefix Range TLV between areas, ABRs **MUST** set the IA-Flag. This is used to prevent redundant flooding of the OSPF Extended Prefix Range TLV between areas as described in Section 4.

7.2. Inter-area Segment Routing in OSPFv2

In order to support SR in a multiarea environment, OSPFv2 **MUST** propagate Prefix-SID information between areas. The following procedure is used 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 OSPF Extended Prefix Opaque LSA as described in [RFC7684]. The flooding scope of the OSPF Extended Prefix Opaque LSA type will be set to area-local 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 this 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 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 OSPF Extended Prefix Opaque LSA as described in [RFC7684]. The flooding scope of the OSPF Extended Prefix Opaque LSA type will be set to area-local 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 will be set as follows:

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

The ABR will then determine if such a 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 backbone 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 when propagating the Prefix-SID for the prefix to other areas.

7.3. Segment Routing for External Prefixes

Type-5 LSAs are flooded domain wide. When an ASBR, which supports SR, generates Type-5 LSAs, it **SHOULD** also originate OSPF Extended Prefix Opaque LSAs as described in [RFC7684]. The flooding scope of the OSPF Extended Prefix Opaque LSA type is set to AS-wide 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 a Not-So-Stubby Area (NSSA) [RFC3101] 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.

7.4. Advertisement of Adj-SID

The Adjacency Segment Routing Identifier (Adj-SID) is advertised using the Adj-SID Sub-TLV as described in Section 6.

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

An Adj-SID **MAY** be advertised for any adjacency on a point-to-point (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 withdrawn from the area.

7.4.2. Adjacency SID on Broadcast or NBMA Interfaces

Broadcast, NBMA, or hybrid [RFC6845] 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, NBMA, or hybrid network connect. As a result, routers on the broadcast, NBMA, or hybrid 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, NBMA, or hybrid network **MAY** advertise the Adj-SID for its adjacency to the DR using the Adj-SID Sub-TLV as described in Section 6.1.

SR-capable routers **MAY** also advertise a LAN Adjacency SID for other neighbors (e.g., Backup Designated Router, DR-OTHER, etc.) on the broadcast, NBMA, or hybrid network using the LAN Adj-SID Sub-TLV as described in Section 6.2.

8. IANA Considerations

This specification updates several existing OSPF registries and creates a new IGP registry.

8.1. OSPF Router Information (RI) TLVs Registry

The following values have been allocated:

Value	TLV Name	Reference
8	SR-Algorithm TLV	This document
9	SID/Label Range TLV	This document
14	SR Local Block TLV	This document
15	SRMS Preference TLV	This document

Table 1: OSPF Router Information (RI) TLVs

8.2. OSPFv2 Extended Prefix Opaque LSA TLVs Registry

The following values have been allocated:

Value	Description	Reference
2	OSPF Extended Prefix Range TLV	This document

Table 2: OSPFv2 Extended Prefix Opaque LSA TLVs

8.3. OSPFv2 Extended Prefix TLV Sub-TLVs Registry

The following values have been allocated:

Value	Description	Reference
1	SID/Label Sub-TLV	This document
2	Prefix-SID Sub-TLV	This document

Table 3: OSPFv2 Extended Prefix TLV Sub-TLVs

8.4. OSPFv2 Extended Link TLV Sub-TLVs Registry

The following initial values have been allocated:

Value	Description	Reference
1	SID/Label Sub-TLV	This document
2	Adj-SID Sub-TLV	This document
3	LAN Adj-SID/Label Sub-TLV	This document

Table 4: OSPFv2 Extended Link TLV Sub-TLVs

8.5. IGP Algorithm Types Registry

IANA has set up a subregistry called "IGP Algorithm Type" under the "Interior Gateway Protocol (IGP) Parameters" registry. The registration policy for this registry is "Standards Action" ([RFC8126] and [RFC7120]).

Values in this registry come from the range 0-255.

The initial values in the IGP Algorithm Type registry are as follows:

Value	Description	Reference
0	Shortest Path First (SPF) algorithm based on link metric. This is the standard shortest path algorithm as computed by the IGP protocol. Consistent with the deployed practice for link-state protocols, Algorithm 0 permits any node to overwrite the SPF path with a different path based on its local policy.	This document
1	Strict Shortest Path First (SPF) algorithm based on link metric. The algorithm is identical to Algorithm 0, but Algorithm 1 requires that all nodes along the path will honor the SPF routing decision. Local policy at the node claiming support for Algorithm 1 **MUST NOT** alter the SPF paths computed by Algorithm 1.	This document

Table 5: IGP Algorithm Types

9. TLV/Sub-TLV Error Handling

For any new TLVs/sub-TLVs defined in this document, if the length is invalid, the LSA in which it is advertised is considered malformed and **MUST** be ignored. An error **SHOULD** be logged subject to rate limiting.

10. Security Considerations

With the OSPFv2 Segment Routing extensions defined herein, OSPFv2 will now program the MPLS data plane [RFC3031] in addition to the IP data plane. Previously, LDP [RFC5036] or another label distribution mechanism was required to advertise MPLS labels and program the MPLS data plane.

In general, the same types of attacks that can be carried out on the IP control plane can be carried out on the MPLS control plane resulting in traffic being misrouted in the respective data planes. However, the latter can be more difficult to detect and isolate.

Existing security extensions as described in [RFC2328] and [RFC7684] apply to these Segment Routing extensions. While OSPF is under a single administrative domain, there can be deployments where potential attackers have access to one or more networks in the OSPF routing domain. In these deployments, stronger authentication mechanisms such as those specified in [RFC7474] **SHOULD** be used.

Implementations **MUST** assure that malformed TLVs and sub-TLVs defined in this document are detected and do not provide a vulnerability for attackers to crash the OSPFv2 router or routing process. Reception of malformed TLVs or sub-TLVs **SHOULD** be counted and/or logged for further analysis. Logging of malformed TLVs and sub-TLVs **SHOULD** be rate limited to prevent a Denial of Service (DoS) attack (distributed or otherwise) from overloading the OSPF control plane.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <https://www.rfc-editor.org/info/rfc2328>.

[RFC3101] Murphy, P., "The OSPF Not-So-Stubby Area (NSSA) Option", RFC 3101, DOI 10.17487/RFC3101, January 2003, <https://www.rfc-editor.org/info/rfc3101>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC6845] Sheth, N., Wang, L., and J. Zhang, "OSPF Hybrid Broadcast and Point-to-Multipoint Interface Type", RFC 6845, DOI 10.17487/RFC6845, January 2013, <https://www.rfc-editor.org/info/rfc6845>.

[RFC7120] Cotton, M., "Early IANA Allocation of Standards Track Code Points", BCP 100, RFC 7120, DOI 10.17487/RFC7120, January 2014, <https://www.rfc-editor.org/info/rfc7120>.

[RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", RFC 7684, DOI 10.17487/RFC7684, November 2015, <https://www.rfc-editor.org/info/rfc7684>.

[RFC7770] Lindem, A., Ed., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 7770, DOI 10.17487/RFC7770, February 2016, <https://www.rfc-editor.org/info/rfc7770>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

[RFC8661] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., and S. Litkowski, "Segment Routing Interworking with LDP", RFC 8661, DOI 10.17487/RFC8661, December 2019, <https://www.rfc-editor.org/info/rfc8661>.

11.2. Informative References

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <https://www.rfc-editor.org/info/rfc5036>.

[RFC7474] Bhatia, M., Hartman, S., Zhang, D., and A. Lindem, Ed., "Security Extension for OSPFv2 When Using Manual Key Management", RFC 7474, DOI 10.17487/RFC7474, April 2015, <https://www.rfc-editor.org/info/rfc7474>.

[RFC7855] Previdi, S., Ed., Filsfils, C., Ed., Decraene, B., Litkowski, S., Horneffer, M., and R. Shakir, "Source Packet Routing in Networking (SPRING) Problem Statement and Requirements", RFC 7855, DOI 10.17487/RFC7855, May 2016, <https://www.rfc-editor.org/info/rfc7855>.

[RFC8666] Psenak, P., Ed. and S. Previdi, Ed., "OSPFv3 Extensions for Segment Routing", RFC 8666, DOI 10.17487/RFC8666, December 2019, <https://www.rfc-editor.org/info/rfc8666>.

Acknowledgements

We would like to thank Anton Smirnov for his contribution.

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

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, and Saku Ytti.

Authors' Addresses

Peter Psenak (EDITOR)
Cisco Systems, Inc.
Apollo Business Center, Mlynske nivy 43
821 09 Bratislava
Slovakia
Email: ppsenak@cisco.com

Stefano Previdi (EDITOR)
Cisco Systems, Inc.
Via Del Serafico, 200
00142 Rome
Italy
Email: stefano@previdi.net

Clarence Filsfils
Cisco Systems, Inc.
Brussels
Belgium
Email: cfilsfil@cisco.com

Hannes Gredler
RtBrick Inc.
Email: hannes@rtbrick.com

Rob Shakir
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
United States of America
Email: robjs@google.com

Wim Henderickx
Nokia
Copernicuslaan 50
2018 Antwerp
Belgium
Email: wim.henderickx@nokia.com

Jeff Tantsura
Apstra, Inc.
Email: jefftant.ietf@gmail.com