Internet Engineering Task Force (IETF)
Request for Comments: 8666
Category: Standards Track
ISSN: 2070-1721
December 2019

P. Psenak, Ed.
S. Previdi, Ed.
Cisco Systems, Inc.

OSPFv3 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 OSPFv3 extensions required for Segment Routing with the MPLS data plane.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction
2. Terminology
3. Segment Routing Identifiers
   3.1. SID/Label Sub-TLV
4. Segment Routing Capabilities
5. OSPFv3 Extended Prefix Range TLV
6. Prefix-SID Sub-TLV
7. Adjacency Segment Identifier (Adj-SID)
   7.1. Adj-SID Sub-TLV
   7.2. LAN Adj-SID Sub-TLV
8. Elements of Procedure
   8.1. Intra-area Segment Routing in OSPFv3
   8.2. Inter-area Segment Routing in OSPFv3

8.3. Segment Routing for External Prefixes
8.4. Advertisement of Adj-SID
8.4.1. Advertisement of Adj-SID on Point-to-Point Links
8.4.2. Adjacency SID on Broadcast or NBMA Interfaces
9. IANA Considerations
9.1. "OSPFv3 Extended-LSA TLVs" Registry
9.2. "OSPFv3 Extended-LSA Sub-TLVs" Registry
10. TLV/Sub-TLV Error Handling
11. Security Considerations
12. References
12.1. Normative References
12.2. Informative References
Contributors
Authors' Addresses

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; the OSPFv3 extension for SR with the IPv6 data plane will be specified in a separate document. When used in MPLS networks, SR paths do not require any LDP or RSVP-TE signaling. However, SR can interoperate in the presence of Label Switched Paths (LSPs) established with RSVP or LDP.

This document describes the OSPFv3 extensions required for Segment Routing with the MPLS data plane.

Segment Routing architecture is described in [RFC8402].

Segment Routing use cases are described in [RFC7855].

2. Terminology

This section lists some of the terminology used in this document:

ABR: Area Border Router

Adj-SID: Adjacency Segment Identifier

AS: Autonomous System

ASBR: Autonomous System Boundary Router

DR: Designated Router

IS-IS: Intermediate System to Intermediate System

LDP: Label Distribution Protocol

LSP: Label Switched Path

MPLS: Multiprotocol Label Switching

OSPF: Open Shortest Path First

SPF: Shortest Path First

RSVP: Resource Reservation Protocol

SID: Segment Identifier

SR: Segment Routing

SRGB: Segment Routing Global Block

SRLB: Segment Routing Local Block

SRMS: Segment Routing Mapping Server

TLV: Type Length Value

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.

3. Segment Routing Identifiers

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

3.1. 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: 7

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.

4. 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 OSPFv3 Router Information Opaque LSA (defined in [RFC7770]) and specified in [RFC8665].

5. OSPFv3 Extended Prefix Range TLV

In some cases, it is useful to advertise attributes for a range of prefixes in a single advertisement. The SR Mapping Server, which is described in [RFC8661], is an example of where SIDs for multiple prefixes can be advertised. To optimize such advertisement in case of multiple prefixes from a contiguous address range, OSPFv3 Extended

Prefix Range TLV is defined.

The OSPFv3 Extended Prefix Range TLV is a top-level TLV of the following LSAs defined in [RFC8362]:

E-Intra-Area-Prefix-LSA

E-Inter-Area-Prefix-LSA

E-AS-External-LSA

E-Type-7-LSA

Multiple OSPFv3 Extended Prefix Range TLVs MAY be advertised in each LSA mentioned above. The OSPFv3 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: 9

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

Prefix Length: Length of prefix in bits.

AF: Address family for the prefix.

AF: 0 - IPv4 unicast

AF: 1 - IPv6 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:

Addresses from the IPv4 multicast address range (224.0.0.0/3), if the AF is IPv4 unicast.

Addresses other than the IPv6 unicast addresses, if the AF is IPv6 unicast.

Flags: Reserved. MUST be zero when sent and are ignored when received.

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.

For the address family IPv6 unicast, the prefix is encoded as an even multiple of 32-bit words and padded with zeroed bits as necessary. This encoding consumes ((PrefixLength + 31) / 32) 32-bit words.

Prefix encoding for other address families is beyond the scope of this specification. Prefix encoding for other address families can be defined in future Standards Track specifications from the IETF stream.

The range represents the contiguous set of prefixes with the same prefix length as specified by the Prefix Length field. The set starts with the prefix that is specified by the Address Prefix field. The number of prefixes in the range is equal to the Range Size.

If the OSPFv3 Extended Prefix Range TLVs advertising the exact same range appears in multiple LSAs of the same type, originated by the same OSPFv3 router, the LSA with the numerically smallest Instance ID MUST be used, and subsequent instances of the OSPFv3 Extended Prefix Range TLVs MUST be ignored.

6. Prefix-SID Sub-TLV

The Prefix-SID sub-TLV is a sub-TLV of the following OSPFv3 TLVs as defined in [RFC8362] and in Section 5:

Intra-Area Prefix TLV

Inter-Area Prefix TLV

External Prefix TLV

OSPFv3 Extended Prefix Range TLV

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     |  Algorithm    |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       SID/Index/Label (variable)              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: 4

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, 2 for IPv6) 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.

Algorithm: Single octet identifying the algorithm the Prefix-SID is associated with as defined in the IGP Algorithm Types registry [ALGOREG].

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 [RFC8665] MUST ignore the Prefix-SID sub-TLV.

SID/Index/Label: According to the V-Flag and L-Flag, 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 OSPFv3 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 prefixes that are propagated between areas by an ABR based on intra-area or inter-area reachability, unless the advertised prefix is directly attached to such 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 advertising 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 Traffic Class field [RFC5462].

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 setting the LA-bit in the Prefix Options as described in [RFC8362].

The Prefix is external type and the downstream neighbor is an ASBR, which is advertising prefix reachability and is setting the LA-bit in the Prefix Options as described in [RFC8362].

When a Prefix-SID is advertised in the OSPFv3 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: 2001:DB8::1/128, Prefix-SID: Index 1
Router-B: 2001:DB8::2/128, Prefix-SID: Index 2
Router-C: 2001:DB8::3/128, Prefix-SID: Index 3
Router-D: 2001:DB8::4/128, Prefix-SID: Index 4
```

then the Address Prefix field in the OSPFv3 Extended Prefix Range TLV would be set to 2001:DB8::1, the Prefix Length would be set to 128, the 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:

```
2001:DB8:1::0/120,   Prefix-SID: Index 51
```

```
2001:DB8:1::100/120, Prefix-SID: Index 52
2001:DB8:1::200/120, Prefix-SID: Index 53
2001:DB8:1::300/120, Prefix-SID: Index 54
2001:DB8:1::400/120, Prefix-SID: Index 55
2001:DB8:1::500/120, Prefix-SID: Index 56
2001:DB8:1::600/120, Prefix-SID: Index 57
```

then the Prefix field in the OSPFv3 Extended Prefix Range TLV would be set to 2001:DB8:1::0, the Prefix Length would be set to 120, the Range Size would be set to 7, and the Index value in the Prefix-SID sub-TLV would be set to 51.

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

The Adj-SID sub-TLV is an optional sub-TLV of the Router-Link TLV as defined in [RFC8362]. It MAY appear multiple times in the Router-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         |     Weight    |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   SID/Label/Index (variable)                  |
+---------------------------------------------------------------+
```

where:

Type: 5

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 (IPFRR) or MPLS-FRR (MPLS Fast Reroute)) as described in Section 3.4 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 the same 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.

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

SID/Index/Label: As described in Section 6.

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 [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.

7.2. LAN Adj-SID Sub-TLV

The LAN Adjacency SID is an optional sub-TLV of the Router-Link TLV. It MAY appear multiple times in the Router-Link TLV. It is used to advertise a SID/Label for an adjacency to a non-DR 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     |     Weight    |             Reserved          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Neighbor ID                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     SID/Label/Index (variable)                |
+---------------------------------------------------------------+
```

where:

Type: 6

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

Flags: Same as in Section 7.1.

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

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

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

SID/Index/Label: As described in Section 6.

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.

8. Elements of Procedure

8.1. Intra-area Segment Routing in OSPFv3

An OSPFv3 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 6).

A Prefix-SID can also be advertised by SR Mapping Servers (as described in [RFC8661]). A Mapping Server advertises Prefix-SIDs for remote prefixes that exist in the OSPFv3 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 SR Mapping Server could use either area flooding scope or autonomous system flooding scope when advertising Prefix-SIDs for prefixes, based on the configuration of the SR Mapping Server. Depending on the flooding scope used, the SR Mapping Server chooses the OSPFv3 LSA type that will be used. If the area flooding scope is needed, an E-Intra-Area-Prefix-LSA [RFC8362] is used. If autonomous system flooding scope is needed, an E-AS-External-LSA [RFC8362] is used.

When a Prefix-SID is advertised by the Mapping Server, which is indicated by the M-Flag in the Prefix-SID sub-TLV (Section 6), the route type as implied by the LSA type is ignored and the Prefix-SID is bound to the corresponding prefix independent of the route type.

Advertisement of the Prefix-SID by the Mapping Server using an Inter-Area Prefix TLV, External-Prefix TLV, or Intra-Area-Prefix TLV [RFC8362] does not itself contribute to the prefix reachability. The NU-bit [RFC5340] MUST be set in the PrefixOptions field of the LSA, which is used by the Mapping Server to advertise SID or SID Range, which prevents the advertisement from contributing to prefix reachability.

An SR Mapping Server MUST use the OSPFv3 Extended Prefix Range TLVs when advertising SIDs for prefixes. Prefixes of different route types can be combined in a single OSPFv3 Extended Prefix Range TLV advertised by an SR Mapping Server.

Area-scoped OSPFv3 Extended Prefix Range TLVs are propagated between areas, similar to propagation of prefixes between areas. Same rules that are used for propagating prefixes between areas [RFC5340] are used for the propagation of the prefix ranges.

8.2. Inter-area Segment Routing in OSPFv3

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

When an OSPFv3 ABR advertises an Inter-Area-Prefix-LSA from an intra-area prefix to all its connected areas, it will also include the Prefix-SID sub-TLV as described in Section 6. 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 OSPFv3 ABR advertises an Inter-Area-Prefix-LSA from an inter-area route to all its connected areas, it will also include the Prefix-SID sub-TLV as described in Section 6. The Prefix-SID value 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 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 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.

8.3. Segment Routing for External Prefixes

AS-External-LSAs are flooded domain wide. When an ASBR, which supports SR, originates an E-AS-External-LSA, it SHOULD also include a Prefix-SID sub-TLV as described in Section 6. 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 an E-NSSA-LSA into an E-AS-External-LSA, it SHOULD also advertise the Prefix-SID for the prefix. The NSSA ABR determines its best path to the prefix advertised in the translated E-NSSA-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 E-AS-External-LSA. Otherwise, the Prefix-SID advertised by any other router will be used.

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 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.

8.4.2. Adjacency SID on Broadcast or NBMA Interfaces

Broadcast, NBMA, or hybrid [RFC6845] networks in OSPFv3 are represented by a star topology where the 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 7.1.

SR-capable routers MAY also advertise a LAN Adjacency SID for other neighbors (e.g., Backup Designated Router (BDR), DR-OTHER, etc.) on the broadcast, NBMA, or hybrid network using the LAN Adj-SID sub-TLV as described in Section 7.2.

9. IANA Considerations

This specification updates two existing OSPFv3 registries.

9.1. "OSPFv3 Extended-LSA TLVs" Registry

The following values have been allocated:

Value	Description	Reference
9	OSPFv3 Extended Prefix Range TLV	This document

Table 1: OSPFv3 Extended-LSA TLVs

9.2. "OSPFv3 Extended-LSA Sub-TLVs" Registry

The following values have been allocated:

Value	Description	Reference
4	Prefix-SID sub-TLV	This document
5	Adj-SID sub-TLV	This document
6	LAN Adj-SID sub-TLV	This document
7	SID/Label sub-TLV	This document

Table 2: OSPFv3 Extended-LSA Sub-TLVs

10. 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. Errors SHOULD be logged subject to rate limiting.

11. Security Considerations

With the OSPFv3 Segment Routing extensions defined herein, OSPFv3 will now program the MPLS data plane [RFC3031]. 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 [RFC5340] and [RFC8362], apply to these Segment Routing extensions. While OSPFv3 is under a single administrative domain, there can be deployments where potential attackers have access to one or more networks in the OSPFv3 routing domain. In these deployments, stronger authentication mechanisms, such as those specified in [RFC4552] or [RFC7166], SHOULD be used.

Implementations MUST ensure that malformed TLVs and sub-TLVs defined in this document are detected and that they do not provide a vulnerability for attackers to crash the OSPFv3 router or routing process. Reception of a malformed TLV or sub-TLV 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 OSPFv3 control plane.

12. References

12.1. Normative References

[ALGOREG] IANA, "Interior Gateway Protocol (IGP) Parameters", <https://www.iana.org/assignments/igp-parameters>.

[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>.

[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>.

[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>.

[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>.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <https://www.rfc-editor.org/info/rfc5340>.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

[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>.

[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>.

[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>.

[RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and F. Baker, "OSPFv3 Link State Advertisement (LSA) Extensibility", RFC 8362, DOI 10.17487/RFC8362, April 2018, <https://www.rfc-editor.org/info/rfc8362>.

[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>.

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

[RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <https://www.rfc-editor.org/info/rfc8665>.

12.2. Informative References

[RFC4552] Gupta, M. and N. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, DOI 10.17487/RFC4552, June 2006, <https://www.rfc-editor.org/info/rfc4552>.

[RFC7166] Bhatia, M., Manral, V., and A. Lindem, "Supporting Authentication Trailer for OSPFv3", RFC 7166, DOI 10.17487/RFC7166, March 2014, <https://www.rfc-editor.org/info/rfc7166>.

[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>.

Contributors

The following people gave a substantial contribution to the content of this document and should be considered coauthors:

Clarence Filsfils
Cisco Systems, Inc.
Brussels
Belgium

Email: cfilsfil@cisco.com

Hannes Gredler
RtBrick Inc.
Austria

Email: hannes@rtbrick.com

Rob Shakir
Google, Inc.
United States of America

Email: robjs@google.com

Wim Henderickx
Nokia
Belgium

Email: wim.henderickx@nokia.com

Jeff Tantsura
Apstra, Inc.
United States of America

Email: jefftant.ietf@gmail.com

Thanks to Acee Lindem for his substantial contribution to the content of this document.

We would like to thank Anton Smirnov for his contribution as well.

Authors' Addresses

Peter Psenak (editor)
Cisco Systems, Inc.
Eurovea Centre, Central 3, Pribinova Street 10
81109 Bratislava
Slovakia

Email: ppsenak@cisco.com

Stefano Previdi (editor)
Cisco Systems, Inc.

Email: stefano@previdi.net