Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: 2 April 2023

G. Dawra
LinkedIn
C. Filsfils
K. Talaulikar, Ed.
Cisco Systems
M. Chen
Huawei
D. Bernier
Bell Canada
B. Decraene
Orange

29 September 2022

BGP Link State Extensions for SRv6
draft-ietf-idr-bgpls-srv6-ext-10

Abstract

Segment Routing over IPv6 (SRv6) allows for a flexible definition of end-to-end paths within various topologies by encoding paths as sequences of topological or functional sub-paths, called "segments". These segments are advertised by various protocols such as BGP, IS-IS and OSPFv3.

This document defines extensions to BGP Link-state (BGP-LS) to advertise SRv6 segments along with their behaviors and other attributes via BGP. The BGP-LS address-family solution for SRv6 described in this document is similar to BGP-LS for SR for the MPLS data-plane defined in a separate document.

Status of This Memo

Table of Contents

1. Introduction

SRv6 refers to Segment Routing instantiated on the IPv6 data-plane [RFC8402]. An SRv6 Segment is often referred to by its SRv6 Segment Identifier (SID).

The network programming paradigm [RFC8986] is central to SRv6. It describes how different behaviors can be bound to SIDs and how a network program can be expressed as a combination of SIDs.

An SRv6-capable node maintains all the SRv6 segments explicitly instantiated locally.

The IS-IS and OSPFv3 link-state routing protocols have been extended to advertise some of these SRv6 SIDs and SRv6-related information [I-D.ietf-lsr-isis-srv6-extensions], [I-D.ietf-lsr-ospfv3-srv6-extensions]. Other SRv6 SIDs may be instantiated on a node via other mechanisms for topological or service functionalities.

The advertisement of SR related information along with the topology for the MPLS data-plane instantiation (SR-MPLS) is specified in [RFC9085] and for the BGP Egress Peer Engineering (EPE) is specified in [RFC9086]. On similar lines, introducing the SRv6 related information in BGP-LS allows consumer applications that require topological visibility to also receive the SRv6 SIDs from nodes across an IGP domain or even across Autonomous Systems (AS), as required. This allows applications to leverage the SRv6 capabilities for network programming.

The identifying key of each Link-State object, namely a node, link, or prefix, is encoded in the Network-Layer Reachability Information (NLRI) and the properties of the object are encoded in the BGP-LS Attribute [RFC7752].

This document describes extensions to BGP-LS to advertise the SRv6 SIDs and other SRv6 information from all the SRv6 capable nodes in the IGP domain when sourced from link-state routing protocols and directly from individual SRv6 capable nodes (e.g. when sourced from BGP for EPE).

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. BGP-LS Extensions for SRv6

BGP-LS [RFC7752] defines the Node, Link, and Prefix Link-State Network Layer Reachability Information (NLRI) types and the advertisement of their attributes via BGP.

When a BGP-LS router advertises topology information that it sources from the underlying link-state routing protocol (as described in [RFC7752]), then it maps the corresponding SRv6 information from the SRv6 extensions for IS-IS [I-D.ietf-lsr-isis-srv6-extensions] and OSPFv3 [I-D.ietf-lsr-ospfv3-srv6-extensions] protocols to their BGP-LS TLVs/sub-TLVs for all SRv6 capable nodes in that routing protocol domain. When a BGP-LS router advertises topology information from the BGP routing protocol (e.g., for EPE as described in [RFC9086]), then it advertises the SRv6 information from the local node alone.

The SRv6 information pertaining to a node is advertised via the BGP-LS Node NLRI and using the BGP-LS Attribute TLVs as follows:

* SRv6 Capabilities of the node are advertised via the SRv6 Capabilities TLV (Section 3.1).

* Maximum SID Depth (MSD) types introduced for SRv6 are advertised (Section 3.2) using the Node MSD TLV specified in [RFC8814]

* Algorithm support for SRv6 is advertised via the SR-Algorithm TLV specified in [RFC9085].

The SRv6 information pertaining to a link is advertised via the BGP-LS Link NLRI and using the BGP-LS Attribute TLVs as follows:

* SRv6 SID of the IGP Adjacency SID or the BGP EPE Peer Adjacency SID [RFC8402] is advertised via the SRv6 End.X SID TLV introduced in this document (Section 4.1).

* SRv6 SID of the IGP Adjacency SID to a non-Designated Router (DR) or non-Designated Intermediate-System (DIS) [RFC8402] is advertised via the SRv6 LAN End.X SID TLV introduced in this document (Section 4.2).

* MSD types introduced for SRv6 are advertised (Section 4.3) using the Link MSD TLV specified in [RFC8814].

The SRv6 information pertaining to a prefix is advertised via the BGP-LS Prefix NLRI and using the BGP-LS Attribute TLVs as follows:

* SRv6 Locator is advertised via the SRv6 Locator TLV introduced in this document (Section 5.1).

* The attributes of the SRv6 Locator are advertised via the Prefix Attribute Flags TLV specified in [RFC9085].

The SRv6 SIDs associated with the node are advertised using the BGP-LS SRv6 SID NLRI introduced in this document (Section 6). This enables the BGP-LS encoding to scale to cover a potentially large set of SRv6 SIDs instantiated on a node with the granularity of individual SIDs and without affecting the size and scalability of the BGP-LS updates. Had the SRv6 SIDs been advertised within the BGP-LS Link Attribute associated with the existing Node NLRI, the BGP-LS update would have grown rather large with the increase in SRv6 SIDs on the node and would have also required a large update message to be generated for any change to even a single SRv6 SID. BGP-LS Attribute TLVs for the SRv6 SID NLRI are introduced in this document as follows:

* The endpoint behavior of the SRv6 SID is advertised via the SRv6 Endpoint Behavior TLV (Section 7.1).

* The BGP EPE Peer Node context for a PeerNode SID, and the Peer Set context for a PeerSet SID [RFC8402] are advertised via the SRv6 BGP EPE Peer Node SID TLV (Section 7.2),

Subsequent sections of this document specify the encoding and usage of these extensions. All the TLVs introduced follow the formats and common field definitions provided in [RFC7752].

3. SRv6 Node Attributes

The SRv6 attributes of a node are advertised using the BGP-LS Attribute TLVs defined in this section and associated with the BGP-LS Node NLRI.

3.1. SRv6 Capabilities TLV

This BGP-LS Attribute TLV is used to announce the SRv6 capabilities of the node along with the BGP-LS Node NLRI and indicates the SRv6 support by the node. A single instance of this TLV MUST be included in the BGP-LS attribute for each SRv6 capable node. The IS-IS SRv6 Capabilities sub-TLV [I-D.ietf-lsr-isis-srv6-extensions] and the OSPFv3 SRv6 Capabilities TLV [I-D.ietf-lsr-ospfv3-srv6-extensions] that map to this BGP-LS TLV are specified with the ability to carry optional sub-sub-TLVs/sub-TLVs. However, no such extensions are currently defined. Moreover, the SRv6 Capabilities TLV defined below is not extensible. As a result, it is expected that any extensions will be introduced as top-level TLVs in the BGP-LS Attribute.

```
  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              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: SRv6 Capabilities TLV Format

Where:

* Type: 1038

* Length : 4.

* Flags: 2 octet field. The flags are derived from the SRv6 Capabilities sub-TLV/TLV of IS-IS (section 2 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 2 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

* Reserved: 2 octet that MUST be set to 0 and ignored on receipt.

3.2. SRv6 Node MSD Types

The Node MSD TLV [RFC8814] of the BGP-LS Attribute of the Node NLRI is also used to advertise the limits and the Segment Routing Header (SRH) [RFC8754] operations supported by the SRv6 capable node. The SRv6 MSD Types specified in section 4 of [I-D.ietf-lsr-isis-srv6-extensions] are also used with the BGP-LS Node MSD TLV as these code points are shared between IS-IS, OSPF and BGP-LS protocols. The description and semantics of these new MSD-types for BGP-LS are identical as specified in [I-D.ietf-lsr-isis-srv6-extensions].

Each MSD-type is encoded in the BGP-LS Node MSD TLV as a one-octet type followed by a one-octet value as derived from the IS-IS and OSPFv3 Node MSD advertisements as specified in [RFC8814].

4. SRv6 Link Attributes

SRv6 attributes and SIDs associated with a link or adjacency are advertised using the BGP-LS Attribute TLVs defined in this section and associated with the BGP-LS Link NLRI.

4.1. SRv6 End.X SID TLV

The SRv6 End.X SID TLV is used to advertise the SRv6 SIDs associated with an IGP Adjacency SID behavior that correspond to a point-to-point or point-to-multipoint link or adjacency of the node running the IS-IS and OSPFv3 protocols. The information advertised via this TLV is derived from the SRv6 End.X SID sub-TLV of IS-IS (section 8.1 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 9.1 of [I-D.ietf-lsr-ospfv3-srv6-extensions]). This TLV can also be used to advertise the SRv6 SID corresponding to the underlying layer-2 member links for a layer-3 bundle interface as a sub-TLV of the L2 Bundle Member Attribute TLV [RFC9085].

This TLV is also used by BGP-LS to advertise the BGP EPE Peer Adjacency SID for SRv6 on the same lines as specified for SR-MPLS in [RFC9086]. The SRv6 SID for the BGP Peer Adjacency using End.X behaviors (viz. End.X, End.X with PSP, End.X with USP, and End.X with PSP & USP) [RFC8986] indicates the cross-connect to a specific layer-3 link to the specific BGP session peer (neighbor).

More than one instance of this TLV can be included in the BGP-LS Attribute; one for each SRv6 End.X SID.

The 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               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Endpoint Behavior      |      Flags    |   Algorithm   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Weight    |   Reserved    |  SID (16 octets) ...          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)             | Sub-TLVs (variable) . . .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: SRv6 End.X TLV Format

Where:

Type: 1106

Length: variable

Endpoint Behavior: 2 octet field. The Endpoint Behavior code point for this SRv6 SID as defined in section 10.2 of [RFC8986].

Flags: 1 octet of flags. The flags are derived from the SRv6 End.X SID sub-TLV/TLV of IS-IS (section 8.1 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 9.1 of [I-D.ietf-lsr-ospfv3-srv6-extensions]). In the case of BGP EPE Peer Adjacency SID, the flags are as defined for the SRv6 BGP Peer Node SID TLV (Section 7.2).

Algorithm: 1 octet field. Algorithm associated with the SID.

Weight: 1 octet field. The value represents the weight of the SID for the purpose of load balancing. The use of the weight is defined in [RFC8402].

Reserved: 1 octet field that MUST be set to 0 and ignored on receipt.

SID: 16 octet field. This field encodes the advertised SRv6 SID as 128 bit value.

Sub-TLVs : Used to advertise sub-TLVs that provide additional attributes for the specific SRv6 SID. This document defines one in Section 8.

4.2. SRv6 LAN End.X SID TLV

For a LAN interface, an IGP node announces normally only its adjacency to the IS-IS pseudo-node (or the equivalent OSPF DR). The information advertised via this TLV is derived from the SRv6 LAN End.X SID sub-TLV of IS-IS (section 8.2 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 9.2 of [I-D.ietf-lsr-ospfv3-srv6-extensions]). The SRv6 LAN End.X SID TLV allows a node to announce the SRv6 SID corresponding to its adjacencies to all other (i.e., non-DIS or non-DR) nodes attached to the LAN in a single instance of the BGP-LS Link NLRI. Without this TLV, multiple BGP-LS Link NLRIs would need to be originated for each additional adjacency to advertise the SRv6 End.X SID TLVs for these neighbor adjacencies. The SRv6 SID for these IGP adjacencies using the End.X behaviors (viz. End.X, End.X with PSP, End.X with USP, and End.X with PSP & USP) [RFC8986] are advertised using the SRv6 LAN End.X SID TLV.

More than one instance of this TLV can be included in the BGP-LS Attribute; one for each SRv6 LAN End.X SID.

The BGP-LS IS-IS SRv6 LAN End.X SID and BGP-LS OSPFv3 SRv6 LAN End.X SID TLVs have 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               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Endpoint Behavior       |      Flags    |   Algorithm   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Weight    |   Reserved    |   Neighbor ID -               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               |
| IS-IS System-ID (6 octets) or OSPFv3 Router-ID (4 octets)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (16 octets) ...                                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Sub-TLVs (variable) . . .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: SRv6 LAN End.X SID TLV Format

Where:

* Type: 1107 in case of IS-IS and 1108 in case of OSPFv3

* Length: variable

* Endpoint Behavior: 2 octet field. The Endpoint Behavior code point for this SRv6 SID as defined in section 10.2 of [RFC8986].

* Flags: 1 octet of flags. The flags are derived from the SRv6 LAN End.X SID sub-TLV/TLV of IS-IS (section 8.2 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 9.2 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

* Algorithm: 1 octet field. Algorithm associated with the SID.

* Weight: 1 octet field. The value represents the weight of the SID for the purpose of load balancing.

* Reserved: 1 octet field that MUST be set to 0 and ignored on receipt.

* Neighbor ID : 6 octets of Neighbor System-ID in IS-IS SRv6 LAN End.X SID TLV and 4 octets of Neighbor Router-id in the OSPFv3 SRv6 LAN End.X SID TLV.

* SID: 16 octet field. This field encodes the advertised SRv6 SID as 128 bit value.

* Sub-TLVs : Used to advertise sub-TLVs that provide additional attributes for the specific SRv6 SID. This document defines one in Section 8.

4.3. SRv6 Link MSD Types

The Link MSD TLV [RFC8814] of the BGP-LS Attribute of the Link NLRI is also used to advertise the limits and the SRH operations supported on the specific link by the SRv6 capable node. The SRv6 MSD Types specified in section 4 of[I-D.ietf-lsr-isis-srv6-extensions] are also used with the BGP-LS Link MSD TLV as these code points are shared between IS-IS, OSPF, and BGP-LS protocols. The description and semantics of these new MSD types for BGP-LS are identical as specified in [I-D.ietf-lsr-isis-srv6-extensions].

Each MSD-type is encoded in the BGP-LS Link MSD TLV as a one-octet type followed by a one-octet value as derived from the IS-IS and OSPFv3 Link MSD advertisements as specified in [RFC8814].

5. SRv6 Prefix Attributes

SRv6 attributes with an IPv6 prefix are advertised using the BGP-LS Attribute TLVs defined in this section and associated with the BGP-LS Prefix NLRI.

5.1. SRv6 Locator TLV

As specified in [RFC8986], an SRv6 SID comprises Locator, Function and Argument parts.

A node is provisioned with one or more Locators supported by that node. Locators are covering prefixes for the set of SIDs provisioned on that node. These Locators are advertised as BGP-LS Prefix NLRI objects along with the SRv6 Locator TLV in its BGP-LS Attribute.

The information advertised via this TLV is derived from the SRv6 Locator TLV of IS-IS (section 7.1 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 7.1 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

The IPv6 Prefix matching the Locator may be also advertised as prefix reachability by the underlying routing protocol. In this case, the Prefix NLRI would be also associated with the Prefix Metric TLV [RFC7752] that carries the routing metric for this prefix. When the Locator prefix is not being advertised as a prefix reachability, then the Prefix NLRI would have the SRv6 Locator TLV associated with it but no Prefix Metric TLV. In the absence of Prefix Metric TLV, the Prefix NLRI advertisement corresponding to a Locator prefix MUST NOT be considered as prefix reachability advertisement in the underlying IGPs default SPF computation.

The SRv6 Locator 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    |   Algorithm   |           Reserved            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                            Metric                             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   Sub-TLVs (variable) . . .
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: SRv6 Locator TLV Format

Where:

Type: 1162

Length: variable

Flags: 1 octet of flags. The flags are derived from the SRv6 Locator TLV of IS-IS (section 7.1 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 7.1 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

Algorithm: 1 octet field. Algorithm associated with the SID.

Reserved: 2 octet field. The value MUST be set to 0 and ignored on receipt.

Metric: 4 octet field. The value of the metric for the Locator derived from the SRv6 Locator TLV of IS-IS (section 7.1 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 7.1 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

Sub-TLVs : Used to advertise sub-TLVs that provide additional attributes for the given SRv6 Locator. Currently none are defined.

6. SRv6 SID NLRI

The "Link-State NLRI" defined in [RFC7752] is extended to carry the SRv6 SID information.

A new "Link-State NLRI Type" is defined for SRv6 SID information as follows:

* Link-State NLRI Type: SRv6 SID NLRI (value 6).

The SRv6 SIDs associated with the node are advertised using the BGP-LS SRv6 SID NLRI.

The format of this new NLRI type is as shown in the following figure:

```
  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
 +-+-+-+-+-+-+-+-+-+
 |  Protocol-ID  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                         Identifier                               |
 |                         (8 octets)                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |               Local Node Descriptors (variable)                //
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |               SRv6 SID Descriptors (variable)                  //
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: SRv6 SID NLRI Format

Where:

* Protocol-ID: 1-octet field that specifies the information source protocol [RFC7752].

* Identifier: 8 octet value as defined in [RFC7752].

* Local Node Descriptors TLV: set of Node Descriptor TLVs for the local node, as defined in [RFC7752] for IGPs, local and static configuration and as defined in [RFC9086] for BGP protocol.

* SRv6 SID Descriptors: set of SRv6 SID Descriptor TLVs. This field MUST contain a single SRv6 SID Information TLV (Section 6.1) and MAY contain the Multi-Topology Identifier TLV [RFC7752].

New TLVs for advertisement within the BGP-LS Attribute [RFC7752] are defined in Section 7 to carry the attributes of an SRv6 SID.

6.1. SRv6 SID Information TLV

An SRv6 SID that is associated with the node and advertised using the SRv6 SID NLRI is encoded using the SRv6 SID Information TLV.

When advertising the SRv6 SIDs from the IGPs, the SID information is derived from the SRv6 End SID sub-TLV of IS-IS (section 7.2 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 8 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

The TLV carries the SRv6 SIDs corresponding to the BGP PeerNode and PeerSet SID [RFC8402] when SRv6 BGP EPE functionality is enabled in BGP.

The 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 (16 octets) ...                                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: SRv6 SID Information TLV Format

Where:

Type: 518

Length: 16.

SID: 16 octet field. This field encodes the advertised SRv6 SID as 128 bit value.

7. SRv6 SID Attributes

This section specifies the TLVs to be carried in the BGP Link State Attribute associated with the BGP-LS SRv6 SID NLRI.

7.1. SRv6 Endpoint Behavior TLV

Each SRv6 SID instantiated on an SRv6 capable node has specific instructions (called behavior) bound to it. [RFC8986] describes how behaviors are bound to a SID and also defines the initial set of well-known behaviors.

The SRv6 Endpoint Behavior TLV is a mandatory TLV that MUST be included in the BGP-LS Attribute associated with the BGP-LS SRv6 SID NLRI.

When advertising the SRv6 SIDs from the IGPs, the Endpoint behavior, Flags, and Algorithm are derived from the SRv6 End SID sub-TLV of IS-IS (section 7.2 of [I-D.ietf-lsr-isis-srv6-extensions]) and OSPFv3 (section 8 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

When advertising the SRv6 SIDs corresponding to the BGP EPE functionality, the Endpoint Behavior corresponds to End.X and similar behaviors. Flags are currently not defined and the algorithm value MUST be 0.

The 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               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |        Endpoint Behavior      |      Flags    |   Algorithm   |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: SRv6 Endpoint Behavior TLV

Where:

Type: 1250

Length: 4.

Endpoint Behavior: 2 octet field. The Endpoint Behavior code point for this SRv6 SID as defined in section 10.2 of [RFC8986].

Flags: 1 octet of flags. The flags map to the IS-IS or OSPFv3 encodings when advertising SRv6 SIDs corresponding to IGPs. For SRv6 SIDs corresponding to BGP EPE, none are defined currently and they MUST be set to 0 and ignored on receipt.

Algorithm: 1 octet field. Algorithm associated with the SID.

7.2. SRv6 BGP Peer Node SID TLV

The BGP PeerNode SID and PeerSet SID for SR-MPLS are specified in [RFC9086]. Similar Peer Node and Peer Set functionality can be realized with SRv6 using SIDs with END.X behavior. Refer to Appendix A for some differences between the signaling of these SIDs in SR-MPLS and SRv6. The SRv6 BGP Peer Node SID TLV is a mandatory TLV for use in the BGP-LS Attribute for an SRv6 SID NLRI advertised by BGP for the EPE functionality. This TLV MUST be included along with SRv6 SIDs that are associated with the BGP PeerNode or PeerSet functionality.

The 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             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                      Peer AS Number                           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                     Peer BGP Identifier                       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: SRv6 BGP Peer Node SID TLV Format

Where:

* Type: 1251

* Length: 12

* Flags: 1 octet of flags with the following definition:

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

Figure 9: SRv6 BGP EPE SID Flags Format

  - B-Flag: Backup Flag. If set, the SID is eligible for protection (e.g., using IPFRR) as described in [RFC8355].

  - S-Flag: Set Flag. When set, the S-Flag indicates that the SID refers to a set of BGP peering sessions (i.e., BGP Peer Set SID functionality) and therefore MAY be assigned to one or more End.X SIDs associated with BGP peer sessions.

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

  - Other bits are reserved for future use and MUST be set to 0 and ignored on receipt.

* Weight: 1 octet field. The value represents the weight of the SID for the purpose of load balancing. The use of the weight is defined in [RFC8402].

* Reserved: 2 octet field. The value MUST be set to 0 and ignored on receipt.

* Peer AS Number : 4 octets of BGP AS number of the peer router.

* Peer BGP Identifier : 4 octets of the BGP Identifier (BGP Router-ID) of the peer router.

For an SRv6 BGP EPE Peer Node SID, one instance of this TLV is associated with the SRv6 SID. For SRv6 BGP EPE Peer Set SID, multiple instances of this TLV (one for each peer in the "peer set") are associated with the SRv6 SID and the S-Flag is SET.

8. SRv6 SID Structure TLV

The SRv6 SID Structure TLV is used to advertise the length of each individual part of the SRv6 SID as defined in [RFC8986]. It is an optional TLV for use in the BGP-LS Attribute for an SRv6 SID NLRI and as a sub-TLV of the SRv6 End.X, IS-IS SRv6 LAN End.X, and OSPFv3 SRv6 LAN End.X TLVs.

When advertising SRv6 SIDs from the IGPs, the SRv6 SID Structure information is derived from the ISIS SRv6 SID Structure sub-sub-TLV (section 9 of [I-D.ietf-lsr-isis-srv6-extensions]) and the OSPFv3 SRv6 SID Structure sub-TLV (section 10 of [I-D.ietf-lsr-ospfv3-srv6-extensions]).

When advertising the SRv6 SIDs corresponding to the BGP EPE functionality, the SRv6 SID Structure information is derived from the locally provisioned SRv6 SID.

The 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               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    LB Length  |  LN Length    | Fun. Length   |  Arg. Length  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: SRv6 SID Structure TLV

Where:

Type: 1252

Length: 4

LB Length: 1 octet field. SRv6 SID Locator Block length in bits.

LN Length: 1 octet field. SRv6 SID Locator Node length in bits.

Fun. Length: 1 octet field. SRv6 SID Function length in bits.

Arg. Length: 1 octet field. SRv6 SID Argument length in bits.

The sum of the LB Length, LN Length, Func. Length, and Arg. Length MUST be less than or equal to 128.

9. IANA Considerations

This document requests assigning code points from the IANA "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry as described in the sub-sections below.

9.1. BGP-LS NLRI-Types

The following code points have been assigned by IANA via the early allocation process from within the sub-registry called "BGP-LS NLRI-Types":

Type	NLRI Type	Reference
6	SRv6 SID NLRI	this document

Figure 11: SRv6 SID NLRI Type Code Point

9.2. BGP-LS TLVs

The following TLV code points have been assigned by IANA via the early allocation process from within the sub-registry called "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs":

TLV Code Point	Description	Value defined in
518	SRv6 SID Information TLV	this document
1038	SRv6 Capabilities TLV	this document
1106	SRv6 End.X SID TLV	this document
1107	IS-IS SRv6 LAN End.X SID TLV	this document
1108	OSPFv3 SRv6 LAN End.X SID TLV	this document
1162	SRv6 Locator TLV	this document
1250	SRv6 Endpoint Behavior TLV	this document
1251	SRv6 BGP Peer Node SID TLV	this document
1252	SRv6 SID Structure TLV	this document

Figure 12: SRv6 BGP-LS Attribute TLV Code Points

9.3. SRv6 BGP EPE SID Flags

This document requests the creation of a new registry called "SRv6 BGP EPE SID Flags" under the "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry. The allocation policy of this registry is "Standards Action" according to [RFC8126].

The following flags are defined:

Bit	Description	Reference
0	Backup Flag (B-Flag)	This document
1	Set Flag (S-Flag)	This document
2	Persistent Flag (P-Flag)	This document
3-7	Unassigned	

Figure 13: SRv6 BGP EPE SID Flags

10. Manageability Considerations

This section is structured as recommended in [RFC5706].

The new protocol extensions introduced in this document augment the existing IGP topology information that is distributed via [RFC7752]. Procedures and protocol extensions defined in this document do not affect the BGP protocol operations and management other than as discussed in the Manageability Considerations section of [RFC7752]. Specifically, the malformed attribute tests for syntactic checks in the Fault Management section of [RFC7752] now encompass the new BGP-LS extensions defined in this document. The semantic or content checking for the TLVs specified in this document and their association with the BGP-LS NLRI types or their BGP-LS Attribute is left to the consumer of the BGP-LS information (e.g., an application or a controller) and not the BGP protocol.

The SR information introduced in BGP-LS by this specification may be used by BGP-LS consumer applications like an SR path computation engine (PCE) to learn the SRv6 capabilities of the nodes in the topology and the mapping of SRv6 segments to those nodes. This can enable the SR PCE to perform path computations based on SR for traffic engineering use-cases and to steer traffic on paths different from the underlying IGP based distributed best path computation. Errors in the encoding or decoding of the SRv6 information may result in the unavailability of such information to the SR PCE or incorrect information being made available to it. This may result in the SR PCE not being able to perform the desired SR-based optimization functionality or to perform it in an unexpected or inconsistent manner. The handling of such errors by applications like SR PCE may be implementation-specific and out of the scope of this document.

The manageability considerations related to BGP EPE functionality are discussed in [RFC9086] in the context of SR-MPLS and they also apply to this document in the context of SRv6.

The extensions, specified in this document, do not introduce any new configuration or monitoring aspects in BGP or BGP-LS other than as discussed in [RFC7752]. The manageability aspects of the underlying SRv6 features are covered by [I-D.ietf-spring-srv6-yang].

11. Security Considerations

The new protocol extensions introduced in this document augment the existing IGP topology information that is distributed via [RFC7752]. The advertisement of the SRv6 link-state information defined in this document presents a similar risk as associated with the existing set of link-state information as described in [RFC7752]. The Security Considerations section of [RFC7752] also applies to these extensions. The procedures and new TLVs defined in this document, by themselves, do not affect the BGP-LS security model discussed in [RFC7752].

The extensions introduced in this document are used to propagate IGP defined information ([I-D.ietf-lsr-isis-srv6-extensions] and [I-D.ietf-lsr-ospfv3-srv6-extensions]). These extensions represent the advertisement of SRv6 information associated with the IGP node, link, and prefix. The IGP instances originating these TLVs are assumed to support all the required security and authentication mechanisms (as described in [I-D.ietf-lsr-isis-srv6-extensions] and [I-D.ietf-lsr-ospfv3-srv6-extensions]).

The security considerations related to BGP EPE functionality are discussed in [RFC9086] in the context of SR-MPLS and they also apply to this document in the context of SRv6.

BGP-LS SRv6 extensions enable traffic engineering use-cases within the Segment Routing domain. SR operates within a trusted domain [RFC8402] and its security considerations also apply to BGP-LS sessions when carrying SR information. The SR traffic engineering policies using the SIDs advertised via BGP-LS are expected to be used entirely within this trusted SR domain (e.g., between multiple AS or IGP domains within a single provider network). Therefore, precaution is necessary to ensure that the link-state information (including SRv6 information) advertised via BGP-LS sessions is limited to consumers in a secure manner within this trusted SR domain. BGP peering sessions for address-families other than Link-State may be set up to routers outside the SR domain. The isolation of BGP-LS peering sessions is RECOMMENDED to ensure that BGP-LS topology information (including the newly added SR information) is not advertised to an external BGP peering session outside the SR domain.

12. Contributors

James Uttaro
AT&T
USA
Email: ju1738@att.com

Hani Elmalky
Ericsson
USA
Email: hani.elmalky@gmail.com

Arjun Sreekantiah
Individual
USA
Email: arjunhrs@gmail.com

Les Ginsberg
Cisco Systems
USA
Email: ginsberg@cisco.com

Shunwan Zhuang
Huawei
China
Email: zhuangshunwan@huawei.com

13. Acknowledgements

The authors would like to thank Peter Psenak, Arun Babu, Pablo Camarillo, Francois Clad, Peng Shaofu, Cheng Li, Dhruv Dhody, Tom Petch, and Dan Romascanu for their review of this document and their comments. The authors would also like to thanks Susan Hares for her shepherd review and Adrian Farrel for his detailed Routing Directorate review.

14. References

14.1. Normative References

[I-D.ietf-lsr-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Segment Routing over IPv6 Dataplane", Work in Progress, Internet-Draft, draft-ietf-lsr-isis-srv6-extensions-18, 20 October 2021, <https://www.ietf.org/archive/id/draft-ietf-lsr-isis-srv6-extensions-18.txt>.

[I-D.ietf-lsr-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", Work in Progress, Internet-Draft, draft-ietf-lsr-ospfv3-srv6-extensions-08, 14 September 2022, <https://www.ietf.org/archive/id/draft-ietf-lsr-ospfv3-srv6-extensions-08.txt>.

[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>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

[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>.

[RFC8814] Tantsura, J., Chunduri, U., Talaulikar, K., Mirsky, G., and N. Triantafillis, "Signaling Maximum SID Depth (MSD) Using the Border Gateway Protocol - Link State", RFC 8814, DOI 10.17487/RFC8814, August 2020, <https://www.rfc-editor.org/info/rfc8814>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

[RFC9085] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Gredler, H., and M. Chen, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing", RFC 9085, DOI 10.17487/RFC9085, August 2021, <https://www.rfc-editor.org/info/rfc9085>.

[RFC9086] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Patel, K., Ray, S., and J. Dong, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing BGP Egress Peer Engineering", RFC 9086, DOI 10.17487/RFC9086, August 2021, <https://www.rfc-editor.org/info/rfc9086>.

14.2. Informative References

[I-D.ietf-spring-srv6-yang]
Raza, S., Agarwal, S., Liu, X., Hu, Z., Hussain, I., Shah, H. C., Voyer, D., Matsushima, S., Horiba, K., Rajamanickam, J., and A. Abdelsalam, "YANG Data Model for SRv6 Base and Static", Work in Progress, Internet-Draft, draft-ietf-spring-srv6-yang-02, 23 September 2022, <https://www.ietf.org/archive/id/draft-ietf-spring-srv6-yang-02.txt>.

[RFC5706] Harrington, D., "Guidelines for Considering Operations and Management of New Protocols and Protocol Extensions", RFC 5706, DOI 10.17487/RFC5706, November 2009, <https://www.rfc-editor.org/info/rfc5706>.

[RFC8355] Filsfils, C., Ed., Previdi, S., Ed., Decraene, B., and R. Shakir, "Resiliency Use Cases in Source Packet Routing in Networking (SPRING) Networks", RFC 8355, DOI 10.17487/RFC8355, March 2018, <https://www.rfc-editor.org/info/rfc8355>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

Appendix A. Differences with BGP-EPE for SR-MPLS

The signaling of SRv6 SIDs corresponding to BGP-EPE functionality as defined in this document differ from the signaling of SR-MPLS BGP-EPE SIDs as specified in [RFC9086]. This section provides a high-level overview of the same.

There is no difference in the advertisement of the BGP Peer Adjacency SID in both SR-MPLS and SRv6 and it is advertised as an attribute of the Link NLRI which identifies a specific Layer 3 interface on the BGP Speaker. The difference is in the advertisement of the BGP Peer Node and Peer Set SIDs.

In case of SR-MPLS, an additional Link NLRI is required to be advertised corresponding to each BGP Peering session on the node. Note that, this is not the same Link NLRI associated with the actual layer 3 interface even when the peering is setup using the interface IP addresses. These BGP-LS Link NLRIs are not really links in the traditional link-state routing data model but instead identify BGP peering sessions. The BGP Peer Node and/or Peer Set SIDs associated with that peering session are advertised as attributes associated with this peering Link NLRI. In the case of SRv6, each BGP Peer Node or Peer Set SID is considered to be associated with the BGP Speaker

node and is advertised using the BGP-LS SRv6 SID NLRI while the peering session information is advertised as attributes associated with it.

The advertisement of the BGP Peer Set SID for SR-MPLS is done by including that SID as an attribute in all the Link NLRIs corresponding to the peering sessions that are part of the "set". The advertisement of the BGP Peer Set SID for SRv6 is advertised using a single SRv6 SID NLRI and all the peers associated with that "set" are indicated as attributes associated with the NLRI.

Authors' Addresses

Gaurav Dawra
LinkedIn
United States of America
Email: gdawra.ietf@gmail.com

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

Ketan Talaulikar (editor)
Cisco Systems
India
Email: ketant.ietf@gmail.com

Mach Chen
Huawei
China
Email: mach.chen@huawei.com

Daniel Bernier
Bell Canada
Canada
Email: daniel.bernier@bell.ca

Bruno Decraene
Orange
France
Email: bruno.decraene@orange.com