Inter-Domain Routing   G. Dawra, Ed.
Internet-Draft   LinkedIn
Intended status: Standards Track   C. Filsfils
Expires: July 10, 2020   K. Talaulikar, Ed.
                                                          Cisco Systems
                                                                M. Chen
                                                                 Huawei
                                                             D. Bernier
                                                            Bell Canada
                                                            B. Decraene
                                                                 Orange
                                                        January 7, 2020

BGP Link State Extensions for SRv6
draft-ietf-idr-bgpls-srv6-ext-02

Abstract

Segment Routing 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 the various protocols such as BGP, ISIS and OSPFv3.

BGP Link-state (BGP-LS) address-family solution for SRv6 is similar to BGP-LS for SR for MPLS dataplane. This draft defines extensions to the BGP-LS to advertise SRv6 Segments along with their behaviors and other attributes via BGP.

Table of Contents

1. Introduction

SRv6 refers to Segment Routing instantiated on the IPv6 dataplane [RFC8402]. Segment Identifier (SID) is often used as a shorter reference for "SRv6 Segment".

The network programming paradigm [I-D.ietf-spring-srv6-network-programming] 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 N maintains all the SRv6 segments explicitly instantiated at node N.

The IS-IS [I-D.ietf-lsr-isis-srv6-extensions] and OSPFv3 [I-D.li-ospf-ospfv3-srv6-extensions] link-state routing protocols have been extended to advertise some of these SRv6 SIDs and SRv6-related information. BGP ([I-D.ietf-bess-srv6-services]) has been extended to advertise some of these SRv6 SIDs for VPN services. Certain 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 dataplane instantiation is specified in [I-D.ietf-idr-bgp-ls-segment-routing-ext] and for the BGP Egress Peer Engineering (EPE) is specified in [I-D.ietf-idr-bgpls-segment-routing-epe]. On the 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 a 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 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 domain when sourced from link-state routing protocols and directly from individual SRv6 capable nodes when sourced from BGP.

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 BGP Node, Link and Prefix attributes. All non-VPN link, node, and prefix information SHALL be encoded using AFI 16388 / SAFI 71. VPN link, node, and prefix information SHALL be encoded using AFI 16388 / SAFI 72.

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

- o SRv6 Capabilities of the node is advertised via a new SRv6 Capabilities TLV

- o New MSD types introduced for SRv6 are advertised as new sub-TLVs of the Node MSD TLV specified in [I-D.ietf-idr-bgp-ls-segment-routing-msd].

- o Algorithm support for SRv6 is advertised via the existing SR Algorithm TLV specified in [I-D.ietf-idr-bgp-ls-segment-routing-ext].

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

- o SRv6 End.X SID of the link state routing adjacency or the BGP EPE Peer Adjacency is advertised via a new SRv6 End.X SID TLV

- o SRv6 LAN End.X SID of the link state routing adjacency to a non-DR/DIS router is advertised via a new SRv6 LAN End.X SID TLV

- o New MSD types introduced for SRv6 are advertised as new sub-TLVs of the Link MSD TLV specified in [I-D.ietf-idr-bgp-ls-segment-routing-msd].

The SRv6 Locator information of a node is advertised via the BGP-LS Prefix NLRI using the new SRv6 Locator TLV in the BGP-LS Attribute.

The SRv6 SIDs associated with the node are advertised as a newly introduce BGP-LS SRv6 SID NLRI. 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. New BGP-LS Attribute TLVs are introduced for the SRv6 SID NLRI as follows:

- The endpoint behavior of the SRv6 SID is advertised via a new SRv6 Endpoint Behavior TLV

- The BGP EPE Peer Node and Peer Set SID context is advertised via a new SRv6 BGP EPE Peer Node SID TLV

When the BGP-LS router is advertising topology information that it sources from the underlying link-state routing protocol, 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.li-ospf-ospfv3-srv6-extensions] protocols to their BGP-LS TLVs/ sub-TLVs for all SRv6 capable nodes in that routing protocol domain. When the BGP-LS router is advertising topology information from the BGP routing protocol [I-D.ietf-idr-bgpls-segment-routing-epe], then it advertises the SRv6 information from the local node alone (e.g. BGP EPE topology information or in the case of a data center network running BGP as the only routing protocol).

Subsequent sections of this document specify the encoding of the newly defined extensions.

3. SRv6 Node Attributes

SRv6 attributes of a node are advertised using the new 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. This TLV maps to the SRv6 Capabilities sub-TLV and the SRv6 Capabilities TLV of the IS-IS and OSPFv3 protocol SRv6 extensions respectively.

```
  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:

o Type: 2 octet field with value 1038

o Length : 2 octet field with value set to 4.

o Flags: 2 octet field. The following flags are defined:

```
  0                   1
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | |O|                           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: SRv6 Capability TLV Flags Format

* O-flag: If set, then router is capable of supporting SRH O-bit Flags, as specified in [I-D.ietf-6man-spring-srv6-oam].

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

3.2. SRv6 Node MSD Types

The Node MSD TLV [I-D.ietf-idr-bgp-ls-segment-routing-msd] of the BGP-LS Attribute of the Node NLRI is also used to advertise the limits and the supported Segment Routing Header (SRH) [I-D.ietf-6man-segment-routing-header] 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 codepoints 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 [I-D.ietf-lsr-isis-srv6-extensions].

Each MSD type is encoded as a one octet type followed by a one octet value.

4. SRv6 Link Attributes

SRv6 attributes and SIDs associated with a link or adjacency are advertised using the new 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 End.X SIDs that correspond to a point-to-point or point-to-multipoint link or adjacency of the local node for IS-IS and OSPFv3 protocols. This TLV

can also be used to advertise the End.X SRv6 SID corresponding to the underlying layer-2 member links for a layer-3 bundle interface using L2 Bundle Member Attribute TLV as specified in [I-D.ietf-idr-bgp-ls-segment-routing-ext] .

For the nodes running BGP routing protocol, this TLV is used to advertise the BGP EPE Peer Adjacency SID for SRv6 on the same lines as specified for SR/MPLS in [I-D.ietf-idr-bgpls-segment-routing-epe]. The SRv6 End.X SID for the BGP Peer Adjacency indicates the cross-connect to a specific layer-3 link to the specific BGP session peer (neighbor).

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 3: SRv6 End.X TLV Format

Where:

Type: 2 octet field with value 1106

Length: 2 octet field with the total length of the value portion of the TLV.

Endpoint Behavior: 2 octet field. The Endpoint Behavior code point for this SRv6 SID as defined in section 9.2 of [I-D.ietf-spring-srv6-network-programming].

Flags: 1 octet of flags with the following definition:

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

Figure 4: SRv6 End.X SID TLV 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 adjacencies (and therefore MAY be assigned to other adjacencies as well).

* 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 interface flap.

* Rsvd bits: Reserved for future use and MUST be zero when originated and ignored when received.

Algorithm: 1 octet field. Algorithm associated with the SID. Algorithm values are defined in the IGP Algorithm Type registry.

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 SHOULD be set to 0 and MUST be ignored on receipt.

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

Sub-TLVs : currently none defined. Used to advertise sub-TLVs that provide additional attributes for the given SRv6 End.X SID.

4.2. SRv6 LAN End.X SID TLV

For a LAN interface, normally a node only announces its adjacency to the IS-IS pseudo-node (or the equivalent OSPF Designated Router). The SRv6 LAN End.X SID TLV allows a node to announce SRv6 SID corresponding to behaviors like END.X for 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, the corresponding BGP-LS link NLRI would need to be originated for each

additional adjacency in order to advertise the SRv6 End.X SID TLVs for these neighbor adjacencies.

The IS-IS and 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    |   ISIS System-ID (6 octets)   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               |
| or OSPFv3 Router-ID (4 octets) of the neighbor                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    SID (16 octets) ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     SID (cont ...)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     SID (cont ...)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     SID (cont ...)                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Sub-TLVs (variable) . . .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: SRv6 LAN End.X SID TLV Format

Where:

- o Type: 2 octet field with value 1107 in case of IS-IS and 1108 in case of OSPFv3

- o Length: 2 octet field with the total length of the value portion of the TLV.

- o Endpoint Behavior: 2 octet field. The Endpoint Behavior code point for this SRv6 SID as defined in section 9.2 of [I-D.ietf-spring-srv6-network-programming].

- o Flags: 1 octet of flags with the following definition:

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

Figure 6: SRv6 LAN End.X SID TLV 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 adjacencies (and therefore MAY be assigned to other adjacencies as well).

* 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 interface flap.

* Rsvd bits: Reserved for future use and MUST be zero when originated and ignored when received.

o Algorithm: 1 octet field. Algorithm associated with the SID. Algorithm values are defined in the IGP Algorithm Type registry.

o 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].

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

o Neighbor ID : 6 octets of ISIS System ID of the neighbor for the ISIS SRv6 LAN End.X SID TLV and 4 octets of OSPFv3 Router-id of the neighbor for the OSPFv3 SRv6 LAN End.X SID TLV.

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

o Sub-TLVs : currently none defined. Used to advertise sub-TLVs that provide additional attributes for the given SRv6 LAN End.X SID.

4.3. SRv6 Link MSD Types

The Link MSD TLV [I-D.ietf-idr-bgp-ls-segment-routing-msd] of the BGP-LS Attribute of the Link NLRI is also used to advertise the limits and the supported Segment Routing Header (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 codepoints 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 [I-D.ietf-lsr-isis-srv6-extensions].

Each MSD type is encoded as a one octet type followed by a one octet value.

5. SRv6 Prefix Attributes

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

5.1. SRv6 Locator TLV

As described in [I-D.ietf-spring-srv6-network-programming], an SRv6 SID is 128 bits and comprises of 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 IPv6 Prefix matching the Locator MAY be also advertised as a prefix reachability by the underlying routing protocol. In this case, the Prefix NLRI would be also associated with the Prefix Metric TLV 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 consumer of the BGP-LS topology information MUST NOT interpret the Locator prefix as a prefix reachability routing advertisement.

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 7: SRv6 Locator TLV Format

Where:

Type: 2 octet field with value 1162

Length: 2 octet field with the total length of the value portion of the TLV.

Flags: 1 octet of flags with the following definition:

```
     0 1 2 3 4 5 6 7
    +-+-+-+-+-+-+-+-+
    |D|A| Reserved  |
    +-+-+-+-+-+-+-+-+
```

Figure 8: SRv6 Locator TLV Flags Format

* D-Flag: Indicates that the locator has been leaked into the IGP domain when set. IS-IS operations for this are discussed in [I-D.ietf-lsr-isis-srv6-extensions].

* A-Flag: When the Locator is associated with anycast destinations, the A flag SHOULD be set. Otherwise, this bit MUST be clear.

* Reserved bits: Reserved for future use and MUST be zero when originated and ignored when received.

Algorithm: 1 octet field. Algorithm associated with the SID. Algorithm values are defined in the IGP Algorithm Type registry.

Reserved: 2 octet field. The value MUST be zero when originated and ignored when received.

Metric: 4 octet field. The value of the metric for the Locator.

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

6. SRv6 SID NLRI

SRv6 SID information is advertised in BGP UPDATE messages using the MP_REACH_NLRI and MP_UNREACH_NLRI attributes [RFC4760]. 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 following:

o Link-State NLRI Type: SRv6 SID NLRI value 6.

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                          |
|                           (64 bits)                           |
++-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+|
|               Local Node Descriptors (variable)              //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               SRv6 SID Descriptors (variable)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9: SRv6 SID NLRI Format

Where:

o Protocol-ID: 1 octet field that specifies the protocol component through which BGP-LS learns the SRv6 SIDs of the node. The following Protocol-IDs apply to the SRv6 SID NLRI:

Protocol-ID	NLRI information source protocol
1	IS-IS Level 1
2	IS-IS Level 2
4	Direct
5	Static configuration
6	OSPFv3
7	BGP

Figure 10: Protocol IDs for SRv6 SID NLRI

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

- o Local Node Descriptors TLV: as defined in [RFC7752] for IGPs, local and static configuration and as defined in [I-D.ietf-idr-bgpls-segment-routing-epe] for BGP protocol.

- o SRv6 SID Descriptors: MUST include the SRv6 SID Information TLV defined in Section 6.1 and optionally MAY include the Multi-Topology Identifier TLV as defined in [RFC7752].

New TLVs carried in the BGP Link State Attribute defined in [RFC7752] are also defined in order to carry the attributes of a SRv6 SID in Section 7.

6.1. SRv6 SID Information TLV

A SRv6 SID is a 128 bit value [I-D.ietf-spring-srv6-network-programming] and is encoded using the SRv6 SID Information TLV.

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 11: SRv6 SID Information TLV Format

Where:

Type: 2 octet field with value 518

Length: 2 octet field with value set to 16.

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

7. SRv6 SID Attributes

This section specifies the new 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 a specific instruction bound to it. A set of well-known behaviors that can be associated with a SID are defined in [I-D.ietf-spring-srv6-network-programming].

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. 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 12: SRv6 Endpoint Behavior TLV

Where:

Type: 2 octet field with value 1250

Length: 2 octet field with the value 4.

Endpoint Behavior: 2 octet field. The Endpoint Behavior code point for this SRv6 SID as defined in section 9.2 of [I-D.ietf-spring-srv6-network-programming].

Flags: 1 octet of flags with the none defined currently. Reserved for future use and MUST be zero when originated and ignored when received.

Algorithm: 1 octet field. Algorithm associated with the SID. Algorithm values are defined in the IGP Algorithm Type registry.

7.2. SRv6 BGP Peer Node SID TLV

The BGP Peer Node SID and Peer Set SID for SR with MPLS dataplane are specified in [I-D.ietf-idr-bgpls-segment-routing-epe]. The similar Peer Node and Peer Set functionality can be realized with SRv6 using the END.X behavior. The SRv6 BGP Peer Node SID TLV is an optional TLV for use in the BGP-LS Attribute for an SRv6 SID NLRI corresponding to BGP protocol. This TLV MUST be included along with SRv6 End.X SID that is associated with the BGP Peer Node or Peer Set 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 13: SRv6 BGP Peer Node SID TLV Format

Where:

o Type: 2 octet field with value 1251

o Length: 2 octet field with the value 12.

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

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

Figure 14: SRv6 BGP Peer End.X SID TLV 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.

* Rsvd bits: Reserved for future use and MUST be zero when originated and ignored when received.

- o 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].

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

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

For a 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 (set/group) flag is SET.

7.3. SRv6 SID Structure TLV

SRv6 SID Structure TLV is used to advertise the length of each individual part of the SRv6 SID as defined in [I-D.ietf-spring-srv6-network-programming]. It is an optional TLV for use in the BGP-LS Attribute for an SRv6 SID NLRI. 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 15: SRv6 SID Structure TLV

Where:

Type: 2 octet field with value 1252

Length: 2 octet field with the value 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.

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

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

8. 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.

8.1. BGP-LS NLRI-Types

The following codepoints are assigned by IANA via the early allocation processs from within the sub-registry called "BGP-LS NLRI-Types":

Type	NLRI Type	Reference
6	SRv6 SID	this document

Figure 16: SRv6 SID NLRI Type Codepoint

8.2. BGP-LS TLVs

The following TLV codepoints are 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
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
518	SRv6 SID Information 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 17: SRv6 BGP-LS Attribute TLV Codepoints

9. Manageability Considerations

This section is structured as recommended in[RFC5706]

10. Operational Considerations

10.1. Operations

Existing BGP and BGP-LS operational procedures apply. No additional operation procedures are defined in this document.

11. Security Considerations

Procedures and protocol extensions defined in this document do not affect the BGP security model. See the 'Security Considerations' section of [RFC4271] for a discussion of BGP security. Also refer to[RFC4272] and [RFC6952] for analysis of security issues for BGP.

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 and Francois Clad for their review of this document and their comments.

14. References

14.1. Normative References

[I-D.ietf-6man-segment-routing-header]
Filsfils, C., Dukes, D., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-26 (work in progress), October 2019.

[I-D.ietf-6man-spring-srv6-oam]
Ali, Z., Filsfils, C., Matsushima, S., Voyer, D., and M. Chen, "Operations, Administration, and Maintenance (OAM) in Segment Routing Networks with IPv6 Data plane (SRv6)", draft-ietf-6man-spring-srv6-oam-03 (work in progress), December 2019.

[I-D.ietf-bess-srv6-services]
Dawra, G., Filsfils, C., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "SRv6 BGP based Overlay services", draft-ietf-bess-srv6-services-01 (work in progress), November 2019.

[I-D.ietf-idr-bgp-ls-segment-routing-ext]
Previdi, S., Talaulikar, K., Filsfils, C., Gredler, H., and M. Chen, "BGP Link-State extensions for Segment Routing", draft-ietf-idr-bgp-ls-segment-routing-ext-16 (work in progress), June 2019.

[I-D.ietf-idr-bgp-ls-segment-routing-msd]
Tantsura, J., Chunduri, U., Talaulikar, K., Mirsky, G., and N. Triantafillis, "Signaling MSD (Maximum SID Depth) using Border Gateway Protocol Link-State", draft-ietf-idr-bgp-ls-segment-routing-msd-09 (work in progress), October 2019.

[I-D.ietf-idr-bgpls-segment-routing-epe]
Previdi, S., Talaulikar, K., Filsfils, C., Patel, K., Ray, S., and J. Dong, "BGP-LS extensions for Segment Routing BGP Egress Peer Engineering", draft-ietf-idr-bgpls-segment-routing-epe-19 (work in progress), May 2019.

[I-D.ietf-lsr-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extension to Support Segment Routing over IPv6 Dataplane", draft-ietf-lsr-isis-srv6-extensions-03 (work in progress), October 2019.

[I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-ietf-spring-srv6-network-programming-07 (work in progress), December 2019.

[I-D.li-ospf-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", draft-li-ospf-ospfv3-srv6-extensions-07 (work in progress), November 2019.

[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>.

[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>.

14.2. Informative References

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[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>.

[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013, <https://www.rfc-editor.org/info/rfc6952>.

[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>.

Authors' Addresses

Gaurav Dawra (editor)
LinkedIn
USA

Email: gdawra.ietf@gmail.com

Clarence Filsfils
Cisco Systems
Belgium

Email: cfilsfil@cisco.com

Ketan Talaulikar (editor)
Cisco Systems
India

Email: ketant@cisco.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