Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: May 1, 2018

G. Dawra, Ed.
C. Filsfils
K. Talaulikar
A. Sreekantiah
L. Ginsberg
Cisco Systems

October 28, 2017

BGP Link State extensions for IPv6 Segment Routing(SRv6)
draft-dawra-idr-bgpls-srv6-ext-00.txt

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 sub-paths, called "segments". These segments are advertised by the various protocols such as (BGP, IGP, BGP-LS etc).

BGP Link-state(BGP-LS) address-family solution for SRv6 is consistent with BGP-LS for SR-MPLS [I-D.ietf-idr-bgp-ls-segment-routing-ext]. This draft defines extensions to the BGP-LS to carry SR/SRv6 Segments, Services and function information via BGP.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 1, 2018.

Table of Contents

1. Introduction

SRv6 refers to Segment Routing instantiated on the IPv6 dataplane [I-D.ietf-spring-segment-routing].

The network programming paradigm [I-D.filsfils-spring-srv6-network-programming] is central to SRv6. It describes how any function can be bound to a SID and how any network program can be expressed as a combination of SID's.

"SID" (abbreviation for Segment Identifier) is often used as a shorter reference for "SRv6 Segment". An SRv6-capable node N

maintains a "My Local SID Table". This table contains all the local segments explicitly instantiated at node N.

The IGP link-state routing protocols have been extended to advertise some of these SIDs and other SR-related information. IGP extensions are described in [I-D.bashandy-isis-srv6-extensions].

Applications that require topological visibility, syncing the "My Local SID Table" within a domain area, or even across Autonomous Systems (AS), is required.

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 attribute to advertise the SR/SRv6 "My Local SID Table" and some other related SR-information. Some of the information advertised in BGP-LS may not be advertised by IGP/BGP Protocol.

2. BGP-LS extensions for SRv6

BGP-LS[RFC7752] defines the BGP Node and Link 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.

This document defines BGP/IGP SR extensions to BGP-LS Node attribute TLVs in Section 2.1 and Link attribute TLVs in Section 2.2.

2.1. SRv6 Node Attributes

Node Attribute TLVs are used for the node level SRv6 capabilities and for SR-MPLS/SRv6 SIDs with their node level functions (e.g. END, END.T, Service function) that are signaled by the originating router for SR operations.

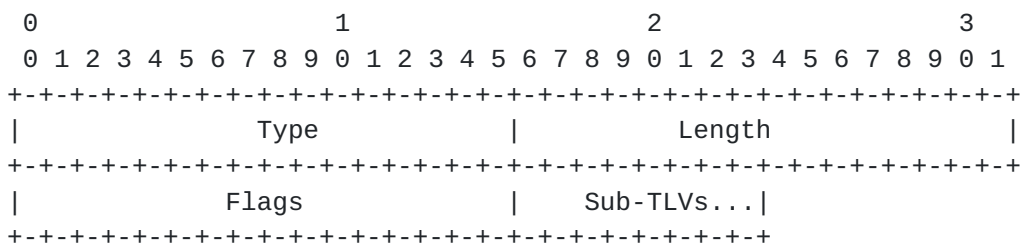
The following Node Attribute TLVs are defined:

TLV Code Point	Description	Length	Section
TBD	SRv6 Capabilities TLV	variable	Section 2.1.1
TBD	SID TLV	variable	Section 2.1.2

These TLVs can ONLY be added to the Node Attribute associated with the local node that is originating the corresponding IGP/BGP TLV.

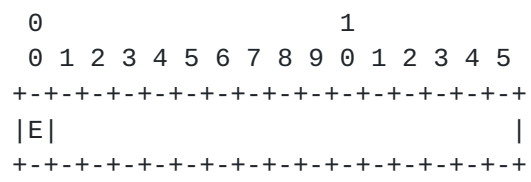
2.1.1. SRv6 Capability Attribute TLV

This TLV is used to announce the SRv6 capability of the router and to indicate the nature of its support for the SRH operations.



Where:

- Type: 16 bit field. TBD
- Length: 16 bit field. Length of Capability TLV + length of Sub-TLVs
- Flags: 16 bit field. The following flags are defined:

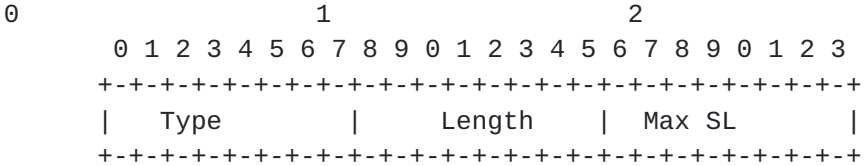


where: E-flag: If set, then router is able to apply "T.Encap" operation

The following sections define the supported sub-TLVs.

2.1.1.1. Maximum SL sub-TLV

The Maximum Segments Left sub-TLV specifies the maximum value of the "SL" field [SRH] in the SRH of a received packet before applying the function associated with a SID.

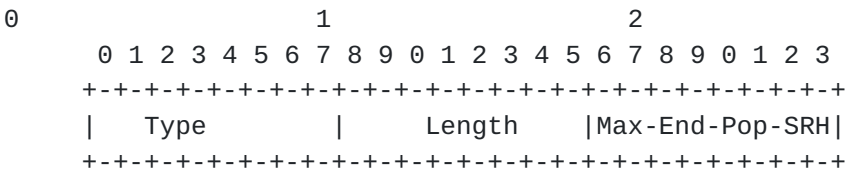


- Type: 1
- Length: 1
- SL Value: 1 octet
- An 8 bit unsigned integer.

If the sub-TLV is NOT advertised the value is assumed to be 0.

2.1.1.2. Maximum End Pop SRH sub-TLV

The Maximum End Pop SRH sub-TLV specifies the maximum number of SIDs in the top SRH in an SRH stack to which the router can apply "PSP" or USP" [NP] flavors.

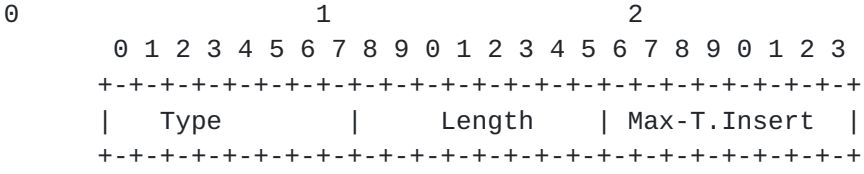


- Type: 2
- Length: 1
- Max-End-Pop-SRH Value: 1 octet
- An 8 bit unsigned integer.

If the value is zero or the sub-sub-TLV is NOT advertised, then it is assumed that the router cannot apply PSP or USP flavors.

2.1.1.3. Maximum T.Insert SRH sub-TLV

The Maximum T.Insert SRH sub-sub-TLV specifies the maximum number of SIDs that can be inserted as part of the "T.insert" behavior [NP].



- Type: 3
- Length: 1

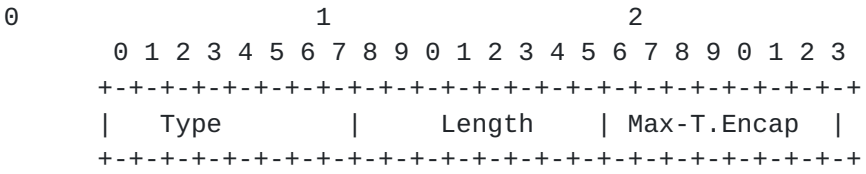
- Max-T.Insert Value: 1 octet

- An 8 bit unsigned integer.

If the value is zero or the sub-sub-TLV is omitted, then the router is assumed not to support any variation of the "T.insert" behavior.

2.1.1.4. Maximum T.Encap SRH sub-TLV

The Maximum T.Encap SRH sub-sub-TLV specifies the maximum number of SIDs that can be included as part of the "T.Encap" behavior [NP].



- Type: 4

- Length: 1

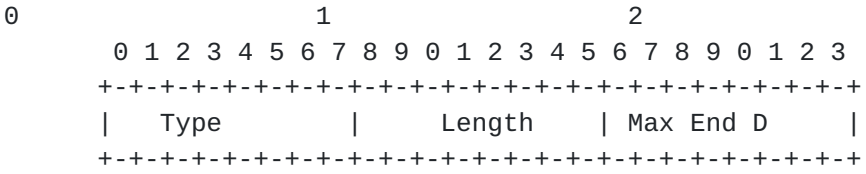
- Max-T.Encap Value: 1 octet

- An 8 bit unsigned integer.

If this value is zero or the sub-TLV is omitted and the "E" flag is set in the associated SRv6 Capabilities sub-TLV, then it is assumed that the router can apply T.Encap by encapsulating the incoming packet in another IPv6 header without SRH the same way IPinIP encapsulation is performed. If the "E" flag is clear, then this sub-sub-TLV SHOULD NOT be transmitted and MUST be ignored on receipt.

2.1.1.5. Maximum End D SRH sub-TLV

The Maximum End D SRH sub-sub-TLV specifies the maximum number of SIDs in an SRH when applying "End.DX6" and "End.DT6" functions.



- Type: 5

- Length: 1

o Max End D Value: 1 octet

o An 8 bit unsigned integer.

If this value is zero or the sub-TLV is omitted, then it is assumed that the router cannot apply "End.DX6" or "End.DT6" functions if the extension header right underneath the outer IPv6 header is an SRH.

Function Sub-TLV described in Section 2.3 is used to encode Function Identifier and Function Flags.

2.1.2. SID TLV

This TLV is used for the Segment Routing SID(SR-MPLS or SRv6) and associate with their node level functions (e.g. END, END.T, Service functions etc) that are signaled by the originating router. Function Sub-TLV described in Section 2.3 is used to encode Function Identifier and Function Flags.

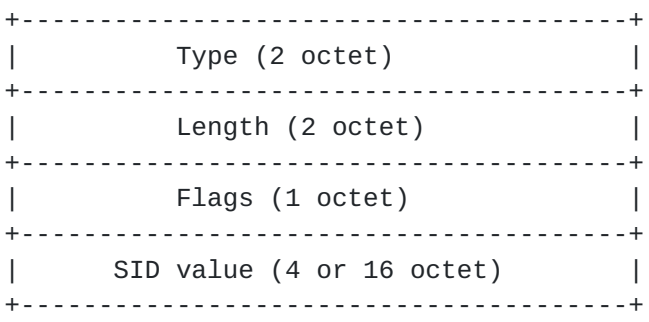


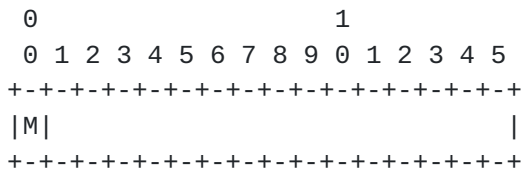
Figure 1: Segment Identifier(SID) TLV

Where:

Type: 16 bit field. TBD

Length: 16 bit field. The total length of the value portion of the TLV.

Flags: 8 bit field. The following flag is defined:



where: M-flag: If set, then SR-MPLS 20-bit local label is encoded. If not set, SRv6 SID is encoded.

Rest Flags SHOULD be set to zero on transmission and MUST be ignored on receipt.

SID Value: 4 octet or 16 octet. 4 octet of MPLS Local Label or 16 octet IPv6 address.

2.2. SRv6 Link Attributes

These TLVs are used for the SRv6 SIDs with their link or adjacency level functions (e.g. END.X function) that are signaled by the originating router for SRv6 operations. One or more of these link attribute TLVs may be attached to the BGP-LS update.

The following Link Attribute TLVs are defined:

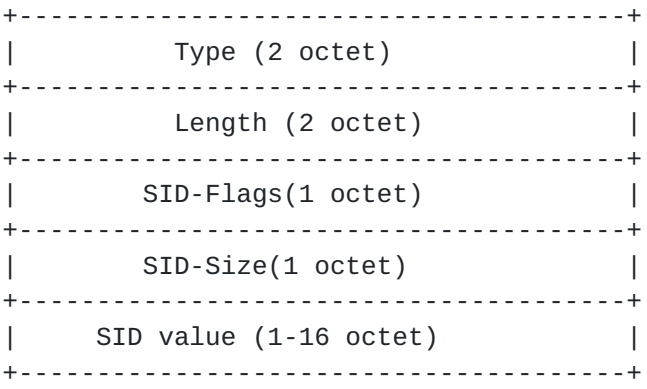
TLV Code Point	Description	Length	Section
TBD	SRv6 P2P X-SID TLV	variable	Section 2.2.1
TBD	SRv6 LAN X-SID TLV	variable	Section 2.2.2

These TLVs can ONLY be added to the Link Attribute associated with the link whose local node originates the corresponding TLV.

For a LAN, normally a node only announces its adjacency to the IS-IS pseudo-node (or the equivalent OSPF Designated and Backup Designated Routers). The SRv6 SID LAN Adjacency Attribute TLV allows a node to announce adjacencies to all other 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 SID TLVs for these neighbor adjacencies.

2.2.1. SRv6 P2P X-SID TLV

The SRv6 SID Adjacency TLV has the following format:



Where:

Type is TBD

Length: 16 bit field. The total length of the value portion of the TLV.

SID-Flags: For IGP as defined in [I-D.bashandy-isis-srv6-extensions]

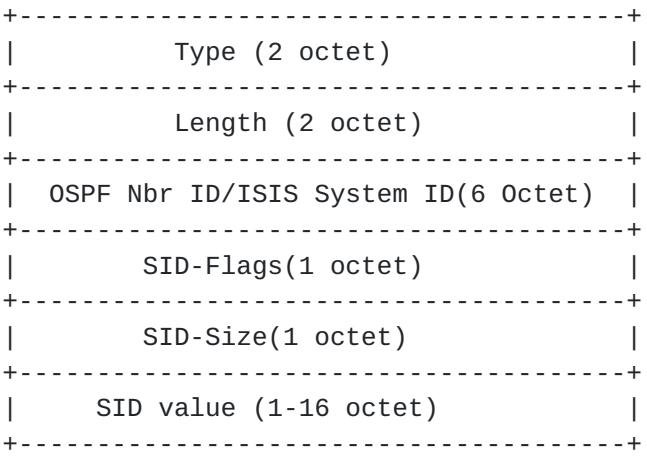
SID-size: Number of bits in the SID field.

SID: 1-16 octets. This field encodes the advertised SRv6 SID. The "SID-size" field can have the values 1-128 and indicates the number of bits in the SID. The SRv6 SID is encoded in the minimal number of octets for the given number of bits.

Function Sub-TLV described in Section 2.3 is used to encode Function Identifier and Function Flags.

2.2.2. SRv6 LAN X-SID TLV

The SRv6 SID LAN Adjacency TLV has the following format:



- Type: TBD

- Length: 16 bit value. Variable

- SID-Flags: For IGP as defined in [I-D.bashandy-isis-srv6-extensions]

- SID-size: Number of bits in the SID field.

- SID: 1-16 octets. This field encodes the advertised SRv6 SID. The "SID-size" field can have the values 1-128 and indicates the number of bits in the SID. The SRv6 SID is encoded in the minimal number of octets for the given number of bits.

- Neighbor : 4 octets for OSPFv2/v3 neighbor ID or 6 octets for IS-IS system-id of the neighbor

Function Sub-TLV described in Section 2.3 is used to encode Function Identifier and Function Flags.

2.3. Function Sub-TLV

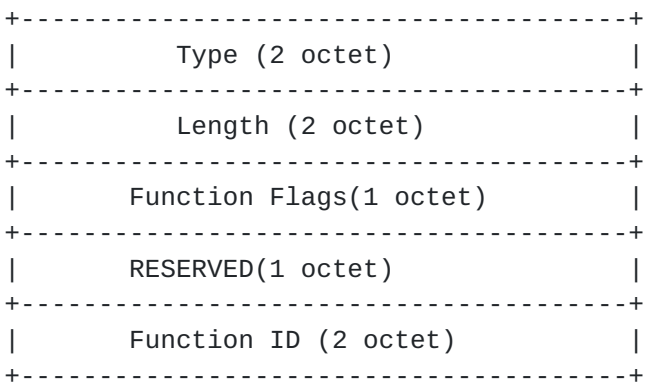


Figure 2: Function Sub-TLV

Where:

Type: 16 bit field: TBD

Length: 16 bit field. Length of the TLV

Function Flags: 8 bit field. For IGP as defined in [I-D.bashandy-isis-srv6-extensions]

RESERVED: 8 bit field. SHOULD be unset on transmission and MUST be ignored on receipt.

Function ID: 16 bit field. Function Identifier of the SID Encoded in the TLV. New function Top Level registry is defined in Section Section 4.2.

3. Function Mapping Summary

Below table illustrate the summary of functions and Node or Link Level TLV encodings. Table will encode Functions for both SR-MPLS/ SRv6 for both Layer-2 and Layer-3 functions with unique values. Table will also include encoding Attribute TLV for functions. Segment Routing functions are defined in[I-D.filsfils-spring-srv6-network-programming]

Function	Function Identifier	Encoding Attribute TLV
END	1	Node SID
END.X	2	Link LAN X-SID Link P2P X-SID
END.T	3	Node SID
END.DX6	4	Link LAN X-SID Link P2P X-SID
END.DX4	5	Link P2P X-SID
END.DT6	6	Node SID
END.DT4	7	Node SID

4. IANA Considerations

This document requests assigning code-points from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs".

4.1. TLV/Sub-TLV Code Points Summary

This section contains the global table of all TLVs defined in this document.

TLV Code Point	Description	Section
TBD	SRv6 Capabilities TLV	Section 2.1.1
TBD	SID TLV	Section 2.1.2
TBD	SRv6 P2P X-SID TLV	Section 2.2.1
TBD	SRv6 LAN X-SID TLV	Section 2.2.2
TBD	Function TLV	Section 2.3

4.2. Segment routing function Identifier(SFI)

IANA is request to create a new top-level registry called "Segment Routing Function Identifier(SFI)". Valid values are in the range 0 to 65535. Values 0 and 65535 are to be marked "Reserved, not to be allocated".

Function	Function Identifier
END	1
END.X	2
END.T	3
END.DX6	4
END.DX4	5
END.DT6	6
END.DT4	7

5. Manageability Considerations

This section is structured as recommended in[RFC5706]

6. Operational Considerations

6.1. Operations

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

7. 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.

8. Conclusions

This document proposes extensions to the BGP to allow advertising certain attributes and functionalities related to SRv6.

9. References

9.1. Normative References

[I-D.bashandy-isis-srv6-extensions]
Ginsberg, L., Bashandy, A., Filsfils, C., and B. Decraene, "IS-IS Extensions to Support Routing over IPv6 Dataplane", draft-bashandy-isis-srv6-extensions-01 (work in progress), September 2017.

[I-D.filsfils-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Raza, K., Liste, J., Clad, F., Lin, S., bogdanov@google.com, b., Horneffer, M., Steinberg, D., Decraene, B., and S. Litkowski, "Segment Routing Policy for Traffic Engineering", draft-filsfils-spring-segment-routing-policy-01 (work in progress), July 2017.

[I-D.filsfils-spring-srv6-network-programming]
Filsfils, C., Leddy, J., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Steinberg, D., Raszuk, R., Matsushima, S., Lebrun, D., Decraene, B., Peirens, B., Salsano, S., Naik, G., Elmalky, H., Jonnalagadda, P., Sharif, M., Ayyangar, A., Mynam, S., Henderickx, W., Bashandy, A., Raza, K., Dukes, D., Clad, F., and P. Camarillo, "SRv6 Network Programming", draft-filsfils-spring-srv6-network-programming-01 (work in progress), June 2017.

[I-D.ietf-6man-segment-routing-header]
Previdi, S., Filsfils, C., Raza, K., Leddy, J., Field, B., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Matsushima, S., Leung, I., Linkova, J., Aries, E., Kosugi, T., Vyncke, E., Lebrun, D., Steinberg, D., and R. Raszuk, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-07 (work in progress), July 2017.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <https://www.rfc-editor.org/info/rfc2460>.

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, DOI 10.17487/RFC3107, May 2001, <https://www.rfc-editor.org/info/rfc3107>.

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

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[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>.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012, <https://www.rfc-editor.org/info/rfc6514>.

[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>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[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>.

9.2. Informative References

[I-D.ietf-bess-evpn-prefix-advertisement]
Rabadan, J., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in EVPN", draft-ietf-bess-evpn-prefix-advertisement-08 (work in progress), October 2017.

[I-D.ietf-idr-bgp-ls-segment-routing-ext]
Previdi, S., Psenak, P., Filsfils, C., Gredler, H., and M. Chen, "BGP Link-State extensions for Segment Routing", draft-ietf-idr-bgp-ls-segment-routing-ext-03 (work in progress), July 2017.

[I-D.ietf-idr-bgp-prefix-sid]
Previdi, S., Filsfils, C., Lindem, A., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix SID extensions for BGP", draft-ietf-idr-bgp-prefix-sid-07 (work in progress), October 2017.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and j. jefftant@gmail.com, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-13 (work in progress), June 2017.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-13 (work in progress), October 2017.

[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>.

[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>.

[RFC4659] De Clercq, J., Ooms, D., Carugi, M., and F. Le Faucheur, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", RFC 4659, DOI 10.17487/RFC4659, September 2006, <https://www.rfc-editor.org/info/rfc4659>.

[RFC5549] Le Faucheur, F. and E. Rosen, "Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop", RFC 5549, DOI 10.17487/RFC5549, May 2009, <https://www.rfc-editor.org/info/rfc5549>.

Appendix A. Contributors

Authors' Addresses

Gaurav Dawra (editor)
Cisco Systems
USA

Email: gdawra@cisco.com

Clarence Filsfils
Cisco Systems
Belgium

Email: cfilsfil@cisco.com

Ketan Talaulikar
Cisco Systems
India

Email: ketant@cisco.com

Arjun Sreekantiah
Cisco Systems
USA

Email: asreekan@cisco.com

Les Ginsberg
Cisco Systems
USA

Email: ginsberg@cisco.com