Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: January 24, 2023

K. Talaulikar, Ed.
Arrcus, Inc
P. Psenak
Cisco Systems
S. Zandi
G. Dawra
LinkedIn
July 23, 2022

Flexible Algorithm Definition Advertisement with BGP Link-State draft-ietf-idr-bgp-ls-flex-algo-09

Abstract

Flexible Algorithm is a solution that allows routing protocols (viz. OSPF and IS-IS) to compute paths over a network based on user-defined (and hence, flexible) constraints and metrics. The computation is performed by routers participating in the specific network in a distributed manner using a Flexible Algorithm definition. This definition is provisioned on one or more routers and propagated (viz. OSPF and IS-IS flooding) through the network.

BGP Link-State (BGP-LS) enables the collection of various topology information from the network. This document defines extensions to the BGP-LS address family to advertise the Flexible Algorithm Definition as a part of the topology information from the network.

Table of Contents

1. Introduction

IGP protocols (OSPF and IS-IS) traditionally compute best paths over the network based on the IGP metric assigned to the links. Many network deployments use RSVP-TE-based [RFC3209] or Segment Routing (SR) Policy-based [RFC8402] solutions to enforce traffic over a path that is computed using different metrics or constraints than the shortest IGP path. [I-D.ietf-lsr-flex-algo] defines the Flexible Algorithm solution that allows IGPs themselves to compute constraint based paths over the network.

Flexible Algorithm is called so as it allows a user the flexibility to define

- the type of calculation to be used (e.g. shortest path)
- the metric type to be used (e.g. IGP metric or TE metric)
- the set of constraints to be used (e.g. inclusion or exclusion of certain links using affinities)

The operations of the IGP flexible algorithm solution are described in detail in [I-D.ietf-lsr-flex-algo].

The BGP-LS extensions for SR are defined in [RFC9085] and [I-D.ietf-idr-bgpls-srv6-ext]. They include the

- SR Algorithm TLV to indicate the participation of a node in a flexible algorithm computation
- Prefix-SID TLV to indicate the association of the Prefix-SIDs to a specific flexible algorithm for SR-MPLS forwarding
- SRv6 Locator TLV to indicate the Locator for specific flexible algorithm for SRv6 forwarding

This document defines extensions to BGP-LS for the advertisement of the Flexible Algorithm Definition (FAD) information to enable learning of the mapping of the flexible algorithm number to its definition in each area/domain of the underlying IGP. This definition indicates the type of computation used and the constraints for a given flexible algorithm. This information can then be leveraged for setting up SR Policy paths end to end across domains by leveraging the appropriate flexible algorithm-specific SIDs in its Segment List [I-D.ietf-spring-segment-routing-policy]. e.g. picking the flexible algorithm Prefix SID (in case of SR-MPLS) or End SID (in case of SRv6) of ABRs/ASBRs corresponding to a definition that optimizes on the delay metric enables the building of an end to end low latency path across IGP domains with minimal SIDs in the SID list.

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. Overview of BGP-LS Extensions for Flexible Algorithm

The BGP-LS [RFC7752] specifies the Node NLRI for the advertisement of nodes along with their attributes using the BGP-LS Attribute, the Link NLRI for the advertisement of links along with their attributes using the BGP-LS Attribute and the Prefix NLRI for the advertisement of prefixes along with their attributes using the BGP-LS Attribute.

The Flexible Algorithm Definition(s) (FAD) advertised by a node is considered as a node-level attribute and advertised as specified in Section 3.

Various link attributes like affinities and SRLGs that are used during the Flexible Algorithm route calculations in IS-IS and OSPF are advertised in those protocols using the Application Specific Link Attribute (ASLA) advertisements as described in [RFC8919], [RFC8920], and [I-D.ietf-lsr-flex-algo]. The BGP-LS extensions for ASLA advertisements are specified in [I-D.ietf-idr-bgp-ls-app-specific-attr].

The Flexible Algorithm Prefix Metric (FAPM) is considered as a prefix attribute and advertised as specified in Section 4.

3. Flexible Algorithm Definition

This document defines a new optional BGP-LS Attribute TLV associated with the Node NLRI called the Flexible Algorithm Definition (FAD) TLV and its format is as follows:

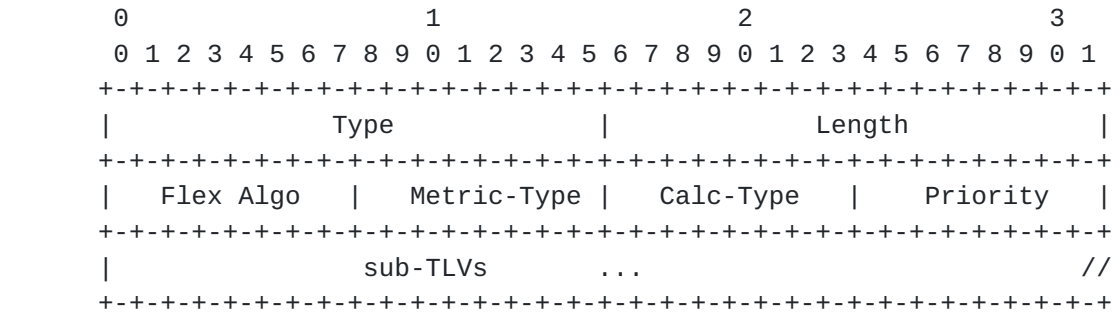

Figure 1: Flexible Algorithm Definition TLV

where:

- Type: 1039
- Length: variable length that represents the total length of the value field in octets. The length value MUST be 4 or larger.

- Flexible Algorithm (Flex Algo): Single octet value carrying the flexible algorithm number between 128 and 255 inclusive, as defined in [I-D.ietf-lsr-flex-algo].

- Metric-Type: Single octet value carrying the metric type, as defined in [I-D.ietf-lsr-flex-algo].

- Calc-Type: Single octet value carrying the calculation type, as defined in [I-D.ietf-lsr-flex-algo].

- Priority: Single octet value carrying the priority of the FAD advertisement, as defined in [I-D.ietf-lsr-flex-algo].

- sub-TLVs: zero or more sub-TLVs may be included as described further in this section.

The FAD TLV that is advertised in the BGP-LS Attribute along with the Node NLRI of a node is derived from the following IGP protocol-specific advertisements:

- In the case of IS-IS, from the ISIS Flexible Algorithm Definition sub-TLV in [I-D.ietf-lsr-flex-algo].

- In the case of OSPFv2/OSPFv3, from the OSPF Flexible Algorithm Definition TLV in [I-D.ietf-lsr-flex-algo].

The BGP-LS Attribute associated with a Node NLRI may include one or more FAD TLVs corresponding to the FAD for each algorithm that the particular node is advertising.

The following sub-sections define the sub-TLVs for the FAD TLV.

3.1. Flexible Algorithm Exclude Any Affinity

The Flexible Algorithm Exclude Any Affinity sub-TLV is an optional sub-TLV that is used to carry the affinity constraints associated with the FAD and enable the exclusion of links carrying any of the specified affinities from the computation of the specific algorithm as described in [I-D.ietf-lsr-flex-algo]. The affinity is expressed in terms of Extended Admin Group (EAG) as defined in [RFC7308].

The sub-TLV has the following format:

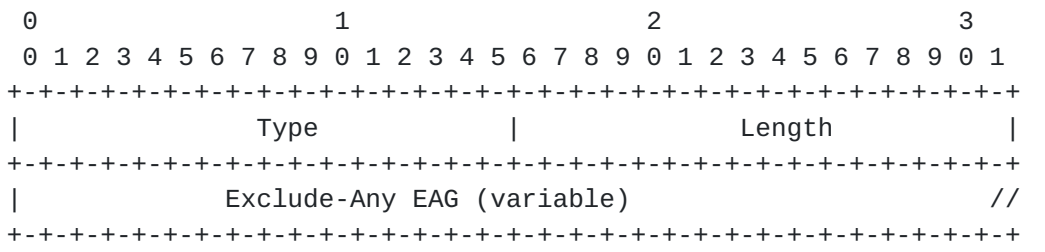

Figure 2: Flexible Algorithm Exclude Any Affinity sub-TLV

where:

- Type: 1040
- Length: variable, dependent on the size of the Extended Admin Group. It MUST be a non-zero value and a multiple of 4.
- Exclude-Any EAG: the EAG value as defined in [I-D.ietf-lsr-flex-algo].

The information in the Flexible Algorithm Exclude Any Affinity sub-TLV is derived from the IS-IS and OSPF protocol-specific Flexible Algorithm Exclude Admin Group sub-TLV as defined in [I-D.ietf-lsr-flex-algo].

3.2. Flexible Algorithm Include Any Affinity

The Flexible Algorithm Include Any Affinity sub-TLV is an optional sub-TLV that is used to carry the affinity constraints associated with the FAD and enable the inclusion of links carrying any of the specified affinities in the computation of the specific algorithm as described in [I-D.ietf-lsr-flex-algo]. The affinity is expressed in terms of Extended Admin Group (EAG) as defined in [RFC7308].

The sub-TLV has the following format:

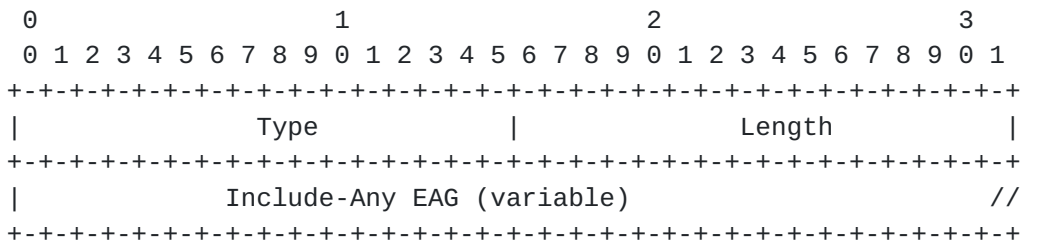

Figure 3: Flexible Algorithm Include Any Affinity sub-TLV

where:

- Type: 1041
- Length: variable, dependent on the size of the Extended Admin Group. It MUST be a non-zero value and a multiple of 4.
- Include-Any EAG: the EAG value as defined in [I-D.ietf-lsr-flex-algo].

The information in the Flexible Algorithm Include Any Affinity sub-TLV is derived from the IS-IS and OSPF protocol-specific Flexible Algorithm Include-Any Admin Group sub-TLV as defined in [I-D.ietf-lsr-flex-algo].

3.3. Flexible Algorithm Include All Affinity

The Flexible Algorithm Include All Affinity sub-TLV is an optional sub-TLV that is used to carry the affinity constraints associated with the FAD and enable the inclusion of links carrying all of the specified affinities in the computation of the specific algorithm as described in [I-D.ietf-lsr-flex-algo]. The affinity is expressed in terms of Extended Admin Group (EAG) as defined in [RFC7308].

The sub-TLV has the following format:

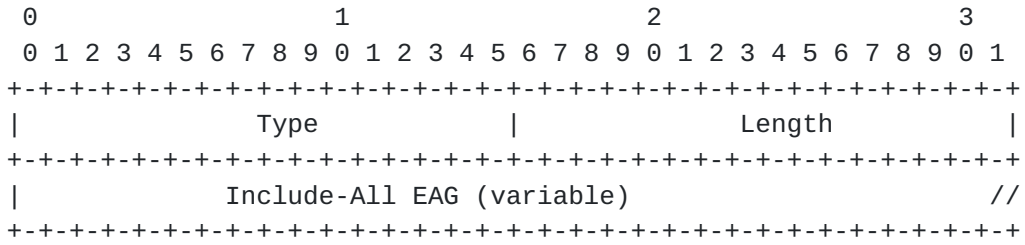

Figure 4: Flexible Algorithm Include All Affinity sub-TLV

where:

- Type: 1042
- Length: variable, dependent on the size of the Extended Admin Group. It MUST be a non-zero value and a multiple of 4.
- Include-All EAG: the EAG value as defined in [I-D.ietf-lsr-flex-algo].

The information in the Flexible Algorithm Include All Affinity sub-TLV is derived from the IS-IS and OSPF protocol-specific Flexible Algorithm Include-All Admin Group sub-TLV as defined in [I-D.ietf-lsr-flex-algo].

3.4. Flexible Algorithm Definition Flags

The Flexible Algorithm Definition Flags sub-TLV is an optional sub-TLV that is used to carry the flags associated with the FAD that are used in the computation of the specific algorithm as described in [I-D.ietf-lsr-flex-algo].

The sub-TLV has the following format:

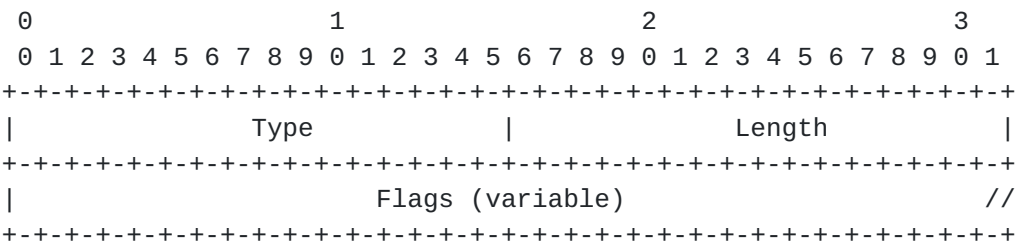

Figure 5: Flexible Algorithm Definition Flags sub-TLV

where:

- Type: 1043
- Length: variable. It MUST be a non-zero value and a multiple of 4.
- Flags: the bitmask used to represent the flags for the FAD, as defined in [I-D.ietf-lsr-flex-algo].

The information in the Flexible Algorithm Definition Flags sub-TLV is derived from the IS-IS and OSPF protocol-specific Flexible Algorithm Definition Flags sub-TLV as defined in [I-D.ietf-lsr-flex-algo].

3.5. Flexible Algorithm Exclude SRLG

The Flexible Algorithm Exclude SRLG sub-TLV is an optional sub-TLV that is used to carry the Shared Risk Link Group (SRLG) information associated with the FAD and enable the exclusion of links that are associated with any of the specified SRLG in the computation of the specific algorithm as described in [I-D.ietf-lsr-flex-algo]. The SRLGs associated with a link are carried in the BGP-LS Shared Link Risk Group (TLV 1096) [RFC7752].

The sub-TLV has the following format:

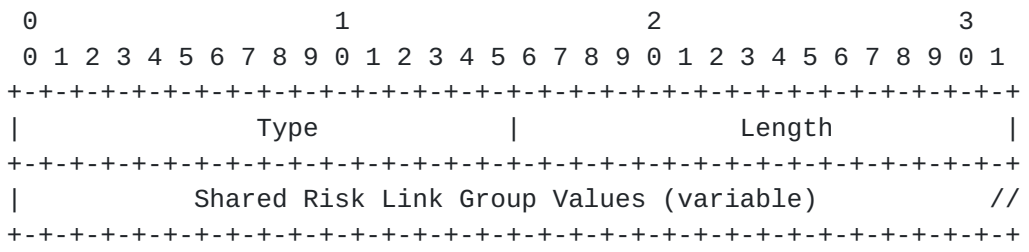

Figure 6: Flexible Algorithm Exclude SRLG sub-TLV

where:

- Type: 1045
- Length: variable, dependent on the number of SRLG values. It MUST be a non-zero value and a multiple of 4.
- Shared Risk Link Group Values: One or more SRLG values, each of 4 octet size, as defined in [I-D.ietf-lsr-flex-algo].

The information in the Flexible Algorithm SRLG Exclude sub-TLV is derived from the IS-IS and OSPF protocol-specific Flexible Algorithm Exclude SRLG sub-TLV as defined in [I-D.ietf-lsr-flex-algo].

3.6. Flexible Algorithm Unknown

The OSPF and ISIS signaling for FAD allows for extensions via new sub-TLVs under the respective IGP's Flexible Algorithm Definition TLV. As specified in section 5.3 of [I-D.ietf-lsr-flex-algo], it is important that the entire FAD be understood by anyone using it for computation purposes. Therefore the FAD is different from most other protocol extensions where the skipping or ignoring of unknown or unsupported sub-TLV information does not affect the base behavior.

The Flexible Algorithm Unknown sub-TLV is an optional sub-TLV that is used to indicate the presence of unknown or unsupported FAD sub-TLVs. The need for this sub-TLV arises when the BGP-LS implementation on the advertising node does not support one or more of the FAD sub-TLVs present in the IGP advertisement.

The sub-TLV has the following format:

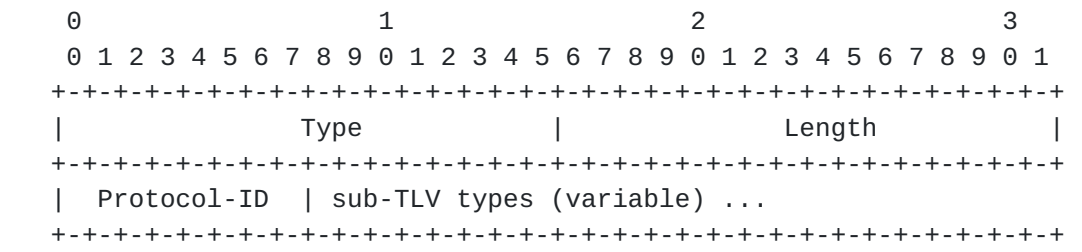

Figure 7: Flexible Algorithm Unknown sub-TLV

where:

- Type: TBD
- Length: variable length that represents the total length of the value field in octets.
- Protocol-ID: Indicates the BGP-LS Protocol-ID of the protocol from which the FAD is being advertised via BGP-LS. The values are from the "BGP-LS Protocol-IDs" registry under the IANA BGP-LS Parameters registry.
- sub-TLV types: Zero or more sub-TLV types that are unknown or unsupported by the node originating the BGP-LS advertisement. The size of each sub-TLV type depends on the protocol indicated by the Protocol-ID field e.g., for ISIS each sub-TLV type would be of size 1 byte while for OSPF each sub-TLV type would be of size 2 bytes.

The node originating the advertisement MUST include the Flexible Algorithm Unknown sub-TLV when it comes across an unsupported or unknown sub-TLV in the corresponding FAD in the IS-IS and OSPF advertisement. When advertising the Flexible Algorithm Unknown sub-TLV, the protocol-specific sub-TLV types that are unsupported or unknown SHOULD be included. This information serves as a diagnostic aid.

The discussion on the use of the FAD information by the consumers of the BGP-LS information is beyond the scope of this document. However, it is RECOMMENDED that the choice of the node used for originating the IGP topology information into BGP-LS be made such that the advertising node supports all the FAD extensions in use in its part of the network. This avoids the scenario where an incomplete FAD gets advertised via BGP-LS.

4. Flexible Algorithm Prefix Metric

This document defines a new optional BGP-LS Attribute TLV associated with the Prefix NLRI called the Flexible Algorithm Prefix Metric (FAPM) TLV and its format is as follows:

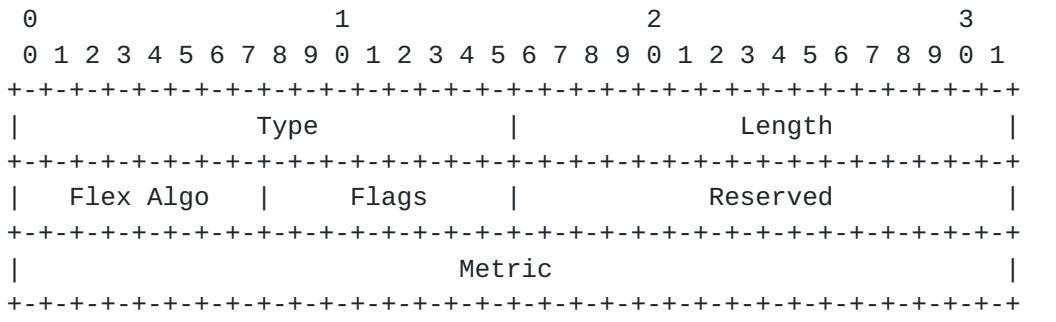

Figure 8: Flexible Algorithm Prefix Metric TLV

where:

- Type: 1044
- Length: 8 octets.
- Flexible Algorithm (Flex Algo): Single octet value carrying the flexible algorithm number between 128 and 255 inclusive, as defined in [I-D.ietf-lsr-flex-algo].
- Flags: single octet value and only applicable for OSPF as defined in [I-D.ietf-lsr-flex-algo]. The value MUST be set to 0 for ISIS.
- Reserved: 2 octet value that MUST be set to 0 by the originator and MUST be ignored by the receiver.
- Metric: 4 octets field to carry the metric information.

The FAPM TLV that is advertised in the BGP-LS Attribute along with the Prefix NLRI from a node is derived from the following IGP protocol-specific advertisements:

- In the case of IS-IS, from the ISIS Flexible Algorithm Prefix Metric sub-TLV in [I-D.ietf-lsr-flex-algo].
- In the case of OSPFv2/OSPFv3, from the OSPF Flexible Algorithm Prefix Metric sub-TLV in [I-D.ietf-lsr-flex-algo].

The BGP-LS Attribute associated with a Prefix NLRI may include one or more FAPM TLVs corresponding to the Flexible Algorithm Prefix Metric for each algorithm associated with that particular prefix.

5. IANA Considerations

IANA has allocated code points from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" <https://www.iana.org/assignments/bgp-ls-parameters/bgp-ls-parameters.xhtml#node-descriptor-link-descriptor-prefix-descriptor-attribute-tlv> based on the table below for most of the TLVs/sub-TLVs. This document requests IANA to allocate the pending code point for the Flexible Algorithm Unknown sub-TLV as suggested below. The column "IS-IS TLV/Sub-TLV" defined in the registry does not require any value and should be left empty.

Code Point	Description
1039	Flexible Algorithm Definition
1040	Flexible Algorithm Exclude Any Affinity
1041	Flexible Algorithm Include Any Affinity
1042	Flexible Algorithm Include All Affinity
1043	Flexible Algorithm Definition Flags
1044	Flexible Algorithm Prefix Metric
1045	Flexible Algorithm Exclude SRLG
1046 (Sugg)	Flexible Algorithm Unknown

Table 1: Flexible Algorithm Code Points

6. Manageability Considerations

The new protocol extensions introduced in this document augment the existing IGP topology information that can be 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 NLRIs attribute tests in the Fault Management section of [RFC7752] now encompass the new TLVs for the BGP-LS NLRI in this document.

The extensions specified in this document do not specify any new configuration or monitoring aspects in BGP or BGP-LS. The specification of BGP models is an ongoing work based on [I-D.ietf-idr-bgp-model].

7. Security Considerations

Security considerations for acquiring and distributing BGP-LS information are discussed in [RFC7752].

The TLVs introduced in this document are used to propagate the IGP Flexible Algorithm extensions defined in [I-D.ietf-lsr-flex-algo]. It is assumed that the IGP instances originating these TLVs will support all the required security (as described in [I-D.ietf-lsr-flex-algo]) to prevent any security issues when propagating the TLVs into BGP-LS.

This document specifies extensions for the advertisement of node and prefix related flexible algorithm information. Tampering with this flexible algorithm-related information may affect applications using it, including impacting route calculation and programming. As the advertisements defined in this document are related to a specific flexible algorithm topology, the impact of tampering is similarly limited in scope.

8. Acknowledgements

The authors would like to thank Les Ginsberg, Amalesh Maity, and Y F Siu for their reviews and contributions to this work. The authors would like to thank Jie Dong for his shepherd review. The authors would like to thank Alvaro Retana for his detailed AD review and suggestions for improving this document.

9. References

9.1. Normative References

[I-D.ietf-lsr-flex-algo]
Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", draft-ietf-lsr-flex-algo-20 (work in progress), May 2022.

[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>.

[RFC7308] Osborne, E., "Extended Administrative Groups in MPLS Traffic Engineering (MPLS-TE)", RFC 7308, DOI 10.17487/RFC7308, July 2014, <https://www.rfc-editor.org/info/rfc7308>.

[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>.

9.2. Informative References

[I-D.ietf-idr-bgp-ls-app-specific-attr]
Talaulikar, K., Psenak, P., and J. Tantsura, "Application-Specific Attributes Advertisement with BGP Link-State", draft-ietf-idr-bgp-ls-app-specific-attr-16 (work in progress), July 2022.

[I-D.ietf-idr-bgp-model]
Jethanandani, M., Patel, K., Hares, S., and J. Haas, "BGP YANG Model for Service Provider Networks", draft-ietf-idr-bgp-model-14 (work in progress), July 2022.

[I-D.ietf-idr-bgpls-srv6-ext]
Dawra, G., Filsfils, C., Talaulikar, K., Chen, M., Bernier, D., and B. Decraene, "BGP Link State Extensions for SRv6", draft-ietf-idr-bgpls-srv6-ext-09 (work in progress), November 2021.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-22 (work in progress), March 2022.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[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>.

[RFC8919] Ginsberg, L., Psenak, P., Previdi, S., Henderickx, W., and J. Drake, "IS-IS Application-Specific Link Attributes", RFC 8919, DOI 10.17487/RFC8919, October 2020, <https://www.rfc-editor.org/info/rfc8919>.

[RFC8920] Psenak, P., Ed., Ginsberg, L., Henderickx, W., Tantsura, J., and J. Drake, "OSPF Application-Specific Link Attributes", RFC 8920, DOI 10.17487/RFC8920, October 2020, <https://www.rfc-editor.org/info/rfc8920>.

[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>.

Authors' Addresses

Ketan Talaulikar (editor)
Arrcus, Inc
India

Email: ketant.ietf@gmail.com

Peter Psenak
Cisco Systems
Slovakia

Email: ppsenak@cisco.com

Shawn Zandi
LinkedIn
USA

Email: szandi@linkedin.com

Gaurav Dawra
LinkedIn
USA

Email: gdawra.ietf@gmail.com