Stream: Internet Engineering Task Force (IETF)
RFC: 9351
Category: Standards Track
Published: February 2023
ISSN: 2070-1721
Authors: K. Talaulikar, Ed. *Cisco Systems*
P. Psenak *Cisco Systems*
S. Zandi *LinkedIn*
G. Dawra *LinkedIn*

RFC 9351
Border Gateway Protocol - Link State (BGP-LS) Extensions for Flexible Algorithm Advertisement

Abstract

Flexible Algorithm is a solution that allows some routing protocols (e.g., 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 (FAD). This definition is provisioned on one or more routers and propagated through the network by OSPF and IS-IS flooding.

Border Gateway Protocol - 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 FAD as a part of the topology information from the network.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9351.

Copyright Notice

Table of Contents

1. Introduction

The classical IGP (e.g., OSPF and IS-IS) computation of best paths over the network is based on the IGP metric assigned to the links in the network. Many network deployments use solutions based on RSVP-TE [RFC3209] or Segment Routing (SR) Policy [RFC8402] to enforce traffic over a path that is computed using different metrics or constraints than the shortest IGP path. [RFC9350] defines the Flexible Algorithm solution that allows IGPs themselves to compute constraint-based paths over the network.

Flexible Algorithm is called so because 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), and
- 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 [RFC9350].

The BGP-LS extensions for SR are defined in [RFC9085] and [IDR-BGPLS-SRV6-EXT] for SR-MPLS and Segment Routing over IPv6 (SRv6), respectively. They include the extensions for advertisement of SR information including various types of Segment Identifiers (SIDs) as below:

- 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 a 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 used for setting up SR Policy paths end to end across domains by using the appropriate Flexible-Algorithm-specific SIDs in its segment list [RFC9256]. For example, picking the Flexible Algorithm Prefix-SID (in case of SR-MPLS) or End SID (in case of SRv6) of Area Border Routers (ABRs) or Autonomous System Border Routers (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

BGP-LS [RFC7752] specifies the Node Network Layer Reachability Information (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 FADs advertised by a node are considered as a node-level attribute and advertised as specified in Section 3.

Various link attributes, like affinities and Shared Risk Link Group (SRLG), 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 [RFC9350]. The BGP-LS extensions for ASLA advertisements are specified in [RFC9294].

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

3. Flexible Algorithm Definition TLV

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

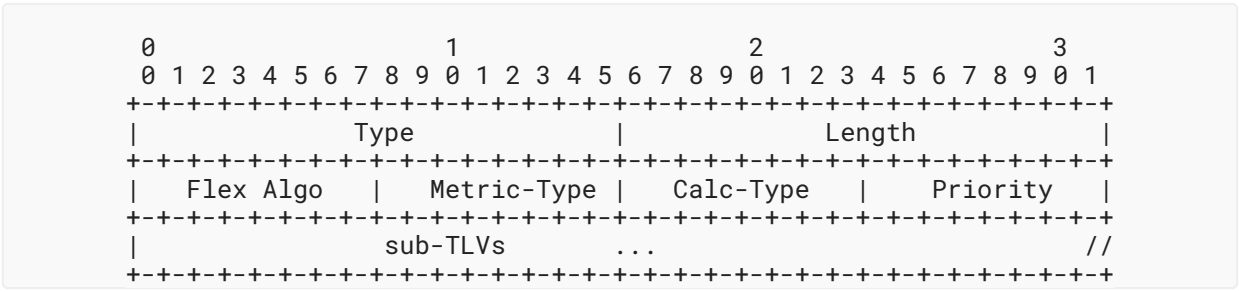

Figure 1: Flexible Algorithm Definition TLV

where:

Type: 1039

Length: The total length of the value field (including any sub-TLVs) 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 [RFC9350].

Metric-Type: Single octet value carrying the metric type, as defined in [RFC9350].

Calc-Type: Single octet value carrying the calculation type, as defined in [RFC9350].

Priority: Single octet value carrying the priority of the FAD advertisement, as defined in [RFC9350].

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 IS-IS Flexible Algorithm Definition sub-TLV in [RFC9350]
- in the case of OSPFv2/OSPFv3, from the OSPF Flexible Algorithm Definition TLV in [RFC9350]

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 subsections define sub-TLVs of the FAD TLV.

3.1. Flexible Algorithm Exclude-Any Affinity Sub-TLV

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 [RFC9350]. The affinity is expressed in terms of the Extended Admin Group (EAG), as defined in [RFC7308].

The sub-TLV has the following format:

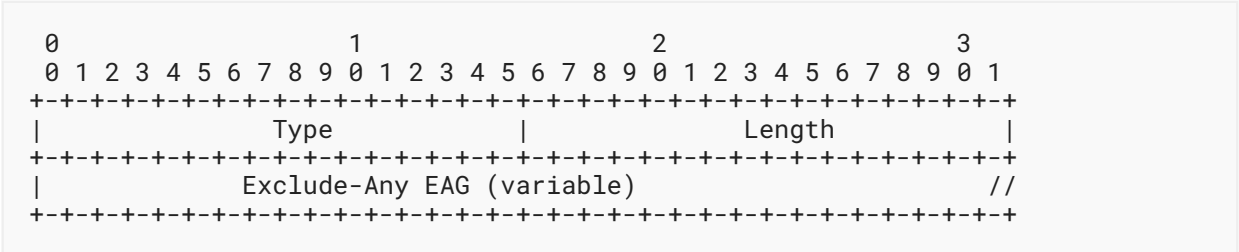

Figure 2: Flexible Algorithm Exclude-Any Affinity Sub-TLV

where:

Type: 1040

Length: The total length of the value field in octets dependent on the size of the EAG. It **MUST** be a non-zero value and a multiple of 4.

Exclude-Any EAG: The EAG value, as defined in [RFC9350].

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 [RFC9350].

3.2. Flexible Algorithm Include-Any Affinity Sub-TLV

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 [RFC9350]. The affinity is expressed in terms of the EAG, as defined in [RFC7308].

The sub-TLV has the following format:

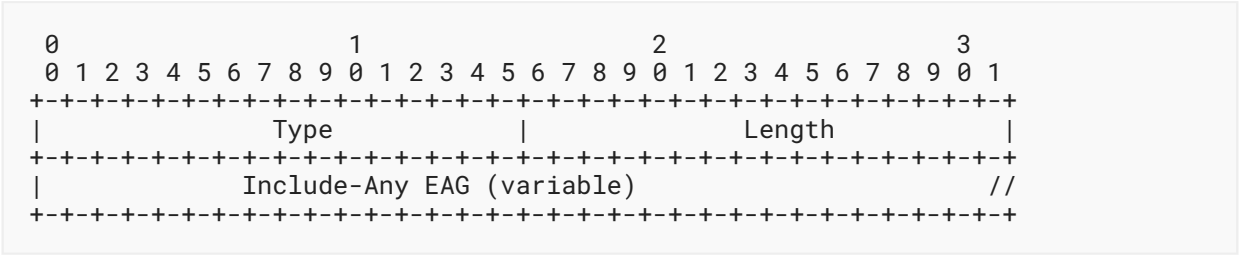

Figure 3: Flexible Algorithm Include-Any Affinity Sub-TLV

where:

Type: 1041

Length: The total length of the value field in octets dependent on the size of the EAG. It **MUST** be a non-zero value and a multiple of 4.

Include-Any EAG: The EAG value, as defined in [RFC9350].

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 [RFC9350].

3.3. Flexible Algorithm Include-All Affinity Sub-TLV

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 [RFC9350]. The affinity is expressed in terms of the EAG, as defined in [RFC7308].

The sub-TLV has the following format:

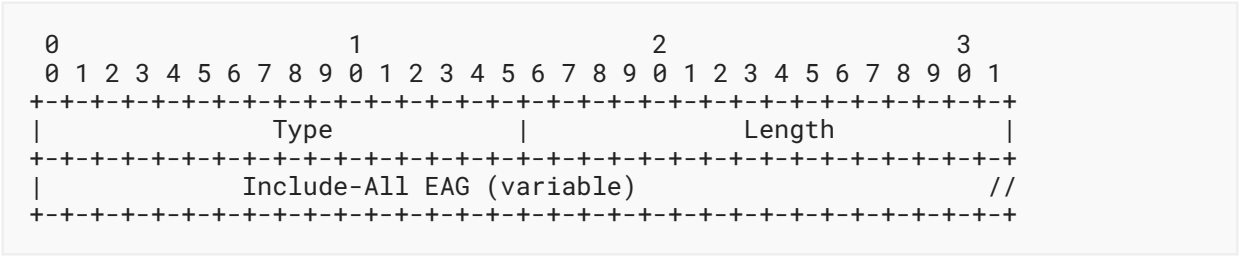

Figure 4: Flexible Algorithm Include-All Affinity Sub-TLV

where:

Type: 1042

Length: The total length of the value field in octets dependent on the size of the EAG. It **MUST** be a non-zero value and a multiple of 4.

Include-All EAG: The EAG value, as defined in [RFC9350].

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 [RFC9350].

3.4. Flexible Algorithm Definition Flags Sub-TLV

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 [RFC9350].

The sub-TLV has the following format:

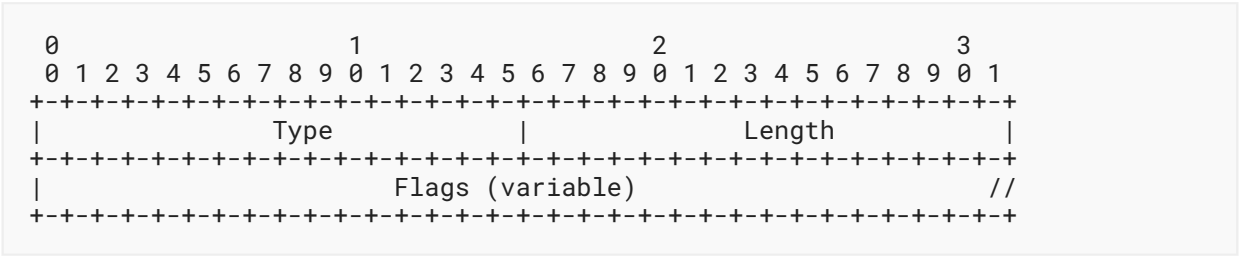

Figure 5: Flexible Algorithm Definition Flags Sub-TLV

where:

Type: 1043

Length: The total length of the value field in octets dependent on the size of the flags. 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 [RFC9350].

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 [RFC9350].

3.5. Flexible Algorithm Exclude SRLG Sub-TLV

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 [RFC9350]. The SRLGs associated with a link are carried in the BGP-LS Shared Risk Link Group (TLV 1096) [RFC7752].

The sub-TLV has the following format:

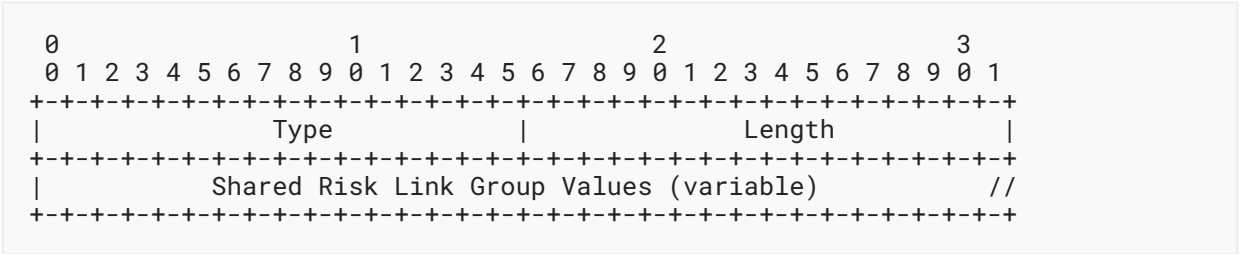

Figure 6: Flexible Algorithm Exclude SRLG Sub-TLV

where:

Type: 1045

Length: The total length of the value field in octets dependent on the number of SRLG values carried. It **MUST** be a non-zero value and a multiple of 4.

Shared Risk Link Group Values: One or more SRLG values, each with a size of 4 octets, as defined in [RFC9350].

The information in the Flexible Algorithm Exclude SRLG sub-TLV is derived from the IS-IS and OSPF protocol-specific Flexible Algorithm Exclude SRLG sub-TLV, as defined in [RFC9350].

3.6. Flexible Algorithm Unsupported Sub-TLV

The OSPF and IS-IS 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 [RFC9350], 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 unsupported sub-TLV information does not affect the base behavior.

The Flexible Algorithm Unsupported sub-TLV is an optional sub-TLV that is used to indicate the presence of 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:

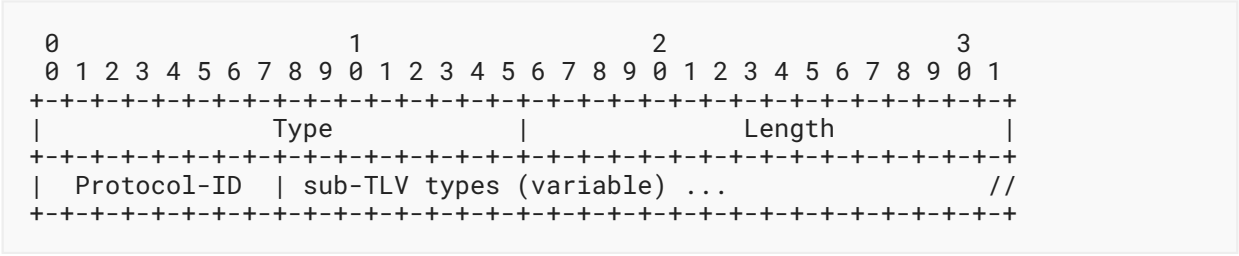

Figure 7: Flexible Algorithm Unsupported Sub-TLV

where:

Type: 1046

Length: The total length of the value field in octets (including any included sub-TLV types).

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 IANA "BGP-LS Protocol-IDs" subregistry under the "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry <https://www.iana.org/assignments/bgp-ls-parameters/>.

sub-TLV types: Zero or more sub-TLV types that are not supported 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. For example, for IS-IS, each sub-TLV type would be 1 octet in size, while for OSPF, each sub-TLV type would be 2 octets in size.

The node originating the advertisement **MUST** include the Flexible Algorithm Unsupported sub-TLV when it comes across an unsupported sub-TLV in the corresponding FAD in the IS-IS and OSPF advertisement. When advertising the Flexible Algorithm Unsupported sub-TLV, the protocol-specific sub-TLV types that are not supported **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 TLV

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

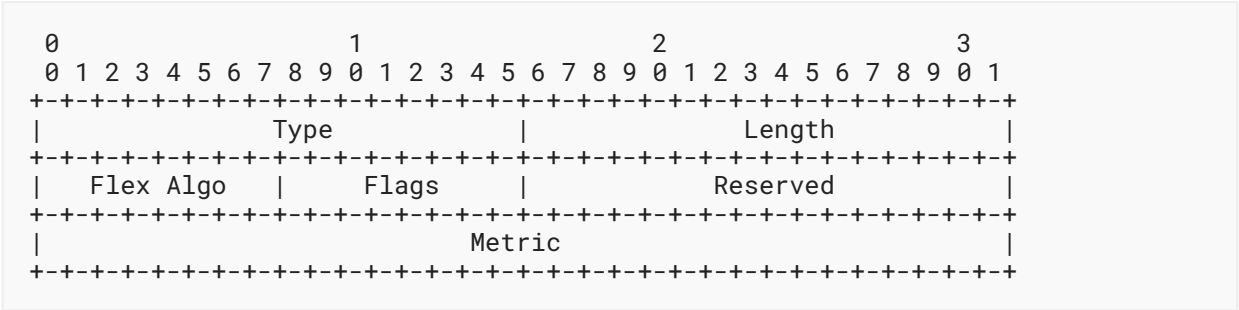

Figure 8: Flexible Algorithm Prefix Metric TLV

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 [RFC9350].

Flags: Single octet value and only applicable for OSPF, as defined in [RFC9350]. The value **MUST** be set to 0 for IS-IS.

Reserved: 2-octet value that **MUST** be set to 0 by the originator and **MUST** be ignored by the receiver.

Metric: 4-octet 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 IS-IS Flexible Algorithm Prefix Metric sub-TLV in [RFC9350]
- in the case of OSPFv2/OSPFv3, from the OSPF Flexible Algorithm Prefix Metric sub-TLV in [RFC9350]

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 in the "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" registry <https://www.iana.org/assignments/bgp-ls-parameters> based on the table below for the TLVs/sub-TLVs introduced by this document.

TLV 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	Flexible Algorithm Unsupported

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 what is 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 [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 [RFC9350]. It is assumed that the IGP instances originating these TLVs will support all the required security (as described in [RFC9350]) for Flexible Algorithm deployment.

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. References

8.1. Normative References

[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>.

[RFC9350] Psenak, P., Ed., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", RFC 9350, DOI 10.17487/RFC9350, February 2023, <https://www.rfc-editor.org/info/rfc9350>.

8.2. Informative References

[IDR-BGP-MODEL] Jethanandani, M., Patel, K., Hares, S., and J. Haas, "BGP YANG Model for Service Provider Networks", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-model-15, 13 October 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-model-15>.

[IDR-BGPLS-SRV6-EXT] Dawra, G., Filsfils, C., Talaulikar, K., Ed., Chen, M., Bernier, D., and B. Decraene, "BGP Link State Extensions for SRv6", Work in Progress, Internet-Draft, draft-ietf-idr-bgpls-srv6-ext-13, 14 January 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgpls-srv6-ext-13>.

[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>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

[RFC9294] Talaulikar, K., Ed., Psenak, P., and J. Tantsura, "Application-Specific Link Attributes Advertisement Using the Border Gateway Protocol - Link State (BGP-LS)", RFC 9294, DOI 10.17487/RFC9294, August 2022, <https://www.rfc-editor.org/info/rfc9294>.

Acknowledgements

The authors would like to thank Les Ginsberg, Amalesh Maity, Y. F. Siu, Vijay Gurbani, and Donald Eastlake 3rd 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.

Authors' Addresses

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

Peter Psenak
Cisco Systems
Slovakia
Email: ppsenak@cisco.com

Shawn Zandi
LinkedIn
United States of America
Email: szandi@linkedin.com

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