IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: 13 November 2026

A. Wang
China Telecom
H. Chen
Individual
K. Talaulikar
Cisco Systems
S. Zhuang
Huawei Technologies
C. Lin
New H3C Technologies

12 May 2026

BGP-LS Extension for Inter-AS Topology Retrieval
draft-ietf-idr-bgpls-inter-as-topology-ext-30

Abstract

This document specifies the procedure for distributing Border Gateway Protocol-Link State (BGP-LS) key parameters for inter-domain links between two Autonomous Systems (ASes). It defines a new type within the BGP-LS Network Layer Reachability Information (NLRI) for an inter-AS Link, as well as three new type-length-values (TLVs) for the BGP-LS inter-AS Link descriptor. These BGP-LS extensions enable Software-Defined Networking (SDN) controllers to retrieve network topology across inter-AS environments.

These extensions and procedures allow network operators to collect inter-domain interconnect information and automatically compute the end-to-end network topology using information provided by the BGP-LS protocol.

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 13 November 2026.

Table of Contents

1. Introduction

BGP-LS [RFC9552] describes the use of the BGP protocol for advertising Link-State topology information. It enables applications such as an SDN controllers to collect the underlay network topology. [RFC9552] covers the advertisement of topology information from within an Interior Gateway Protocol (IGP) domain. If the network has more than one IGP domain, and these domains interconnect with each other via inter-AS links, there is no mechanism within [RFC9552] to advertise the interconnect topology information.

[RFC9086] defines extensions for exporting BGP peering node topology information (including peers, interfaces, and peering ASes) in a way that is used to compute efficient BGP peering engineering policies and strategies. This information can also be used to compute interconnection topology among different IGP domains, but it requires every border router to run the BGP-LS protocol and report such information to SDN controllers. Considering there will be several border routers on the network boundary, such a solution restricts its deployment flexibility.

This document defines the inter-AS Link NLRI and some new TLVs for BGP-LS to cover scenarios where an SDN controller needs to get the interconnection topology information between different AS domains when sourced from IGPs.

2. 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.

3. Terminology

The following terms are defined in this document:

* DCs: Data Centers

* SDN: Software-Defined Network

4. Inter-AS Domain Scenarios

Figure 1 illustrates the multi-domain scenarios discussed in this document. Typically, the SDN controller can retrieve the topology of IGP A and IGP B individually via the BGP-LS protocol, but it cannot obtain topology connection information between these two IGP domains, as IGP protocols are generally not run on the inter-AS links.

In Figure 1, S2 (in IGP domain A) and T1 (in IGP domain B) are connected to the IP SDN controller via BGP-LS, but they can only report the topology information among the IGP A and IGP B themselves, and can’t report the inter-AS topology information among them because there is no IGP protocol running on the inter-AS links. The border routers, SB1/SB3 in IGP A and TB2/TB4 in IGP B know the inter-AS links among them, and can advertise such information via underlying OSPF [RFC5392] or IS-IS [RFC9346], but there is no place in [RFC9552] to transfer such information.

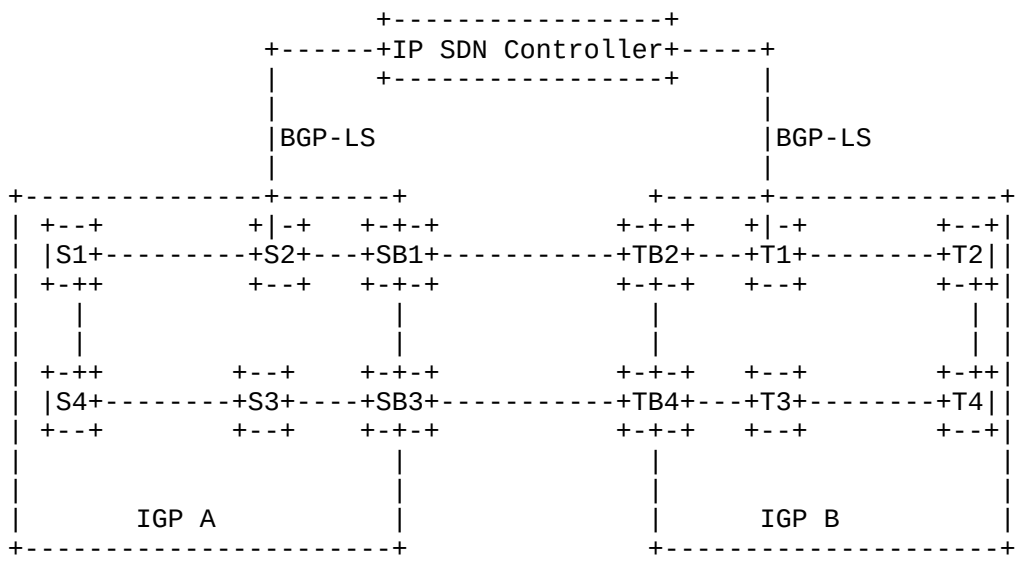

Figure 1: Inter-AS Domain Scenario

5. Inter-AS Link NLRI

[RFC9552] defines four NLRI types (Node, Link, IPv4 Topology Prefix, and IPv6 Topology Prefix) to transfer the topology and prefix information. For an inter-AS link, as the two ends of the link belong in different IGP domains and the link does not run an IGP protocol, it is not appropriate to advertise their information within the existing NLRI types listed above.

This document defines a new NLRI type 7, seeSection 11) within the BGP-LS NLRI, referred to as the inter-AS Link NLRI. The inter-AS Link NLRI is encoded in the format shown in Figure 2 as explained below:

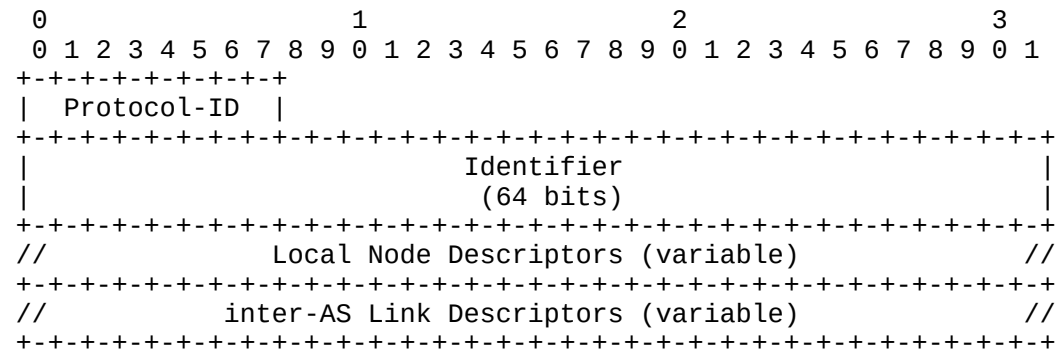

Figure 2: inter-AS Link NLRI Format

This document specifies the advertisement of inter-AS Links using the inter-AS Link NLRI when originating the information from the underlying OSPF [RFC5392] and IS-IS [RFC9346] advertisements.

This section describes the encoding of the inter-AS Link NLRI while the more detailed procedures for sourcing of this information from the underlying IGP are described in Section 7.

The "Protocol-ID" is set to the value indicating the source protocol of the inter-AS Link information, as specified in Section 5.2 of[RFC9552].

The semantics of the "Identifier" field are the same as defined in [RFC9552] and will be set to the BGP-LS Instance Identifier to identify the IGP domain into which the information associated with the inter-AS link is advertised. Therefore, the "Identifier" values for the two half-links (see Section 5.2 of [RFC9552]) of the inter-AS link could be different depending on the configuration of Identifiers for the two IGP domains.

The "Local Node Descriptors" field is encoded using the TLV 256 defined in section 5.2.1.2 of [RFC9552] to identify the ASBR associated with the specific half-link of the inter-AS link. The following Sub-TLVs MUST be included as the Local Node Descriptors:

- Autonomous System (TLV 512) [RFC9552].

- OSPF Area-ID (TLV 514) [RFC9552] to be included only in the case of OSPF, when the inter-AS TE LSA from which information is sourced is being flooded with an area-scope. It is not included when the LSA is flooded with AS-scope.

- IGP Router ID (TLV 515) encoded for either OSPF or IS-IS, depending on the source protocol as specified in section 5.2.1.4 of [RFC9552].

- One or both of IPv4 and IPv6 Router-ID of the ASBR using TLV 1028 and/or 1029 [RFC9552], depending on whether the ASBR is configured with one or both of the IPv4 and IPv6 TE Router-IDs. (Note: while [RFC9552] introduced these TLVs for use in the BGP-LS attribute, this document also leverages the same TLVs for use in the NLRI.)

Inter-AS Link Descriptors are encoded as TLVs that identify the specific half-link of the inter-AS link. Section 6 of this document introduces the TLVs that MUST be included as the inter-AS Link Descriptors:

- Remote AS Number (TLV 270), and

- One or both of IPv4 and IPv6 Remote ASBR ID using TLV 271 and/or TLV 272, depending on whether the Remote ASBR is configured with one or both of the IPv4 and IPv6 TE Router-IDs.

Additionally, the following TLVs MUST be included as inter-AS Link Descriptors if they are being advertised in the underlying IGP advertisement of the inter-AS link as they help identify individual links when there is more than one inter-AS link between two ASBRs.

- Link Local/Remote Identifiers (TLV 258) [RFC9552]

- IPv4 Interface Address (TLV 259) [RFC9552]

- IPv4 Neighbor Address (TLV 260) [RFC9552]

- IPv6 Interface Address (TLV 261) [RFC9552]

- IPv6 Neighbor Address (TLV 262) [RFC9552]

Use of any other TLVs as Local Node Descriptors or inter-AS Link Descriptors may cause challenges in the correlation of the two inter-AS Link NLRI half-links when the BGP-LS Producer implementations vary.

6. Inter-AS Link Descriptor TLVs

This document introduces three TLVs for inclusion as inter-AS Link Descriptors within the inter-AS Link NLRI for the advertisement of inter-AS link information via BGP-LS.

TLV Code Point	Description	IS-IS/OSPF TLV /Sub-TLV	Reference (RFC/Section)
270	Remote AS Number	24/21	[RFC9346]/3.4.1 [RFC5392]/3.3.1
271	IPv4 Remote ASBR ID	25/22	[RFC9346]/3.4.2 [RFC5392]/3.3.2
272	IPv6 Remote ASBR ID	26/24	[RFC9346]/3.4.3 [RFC5392]/3.3.3

Figure 3: inter-AS Link Descriptor TLVs

The encoding of these TLVs is aligned with the corresponding advertisements in [RFC9346] and [RFC5392], which keeps the BGP-LS protocol agnostic to the underlying protocol.

6.1. Remote AS Number TLV

The Remote AS Number TLV specifies the AS number of the neighboring AS to which the advertised link connects.

The Remote AS Number TLV is TLV Type 270 and is 4 octets in length. Its format is as follows:

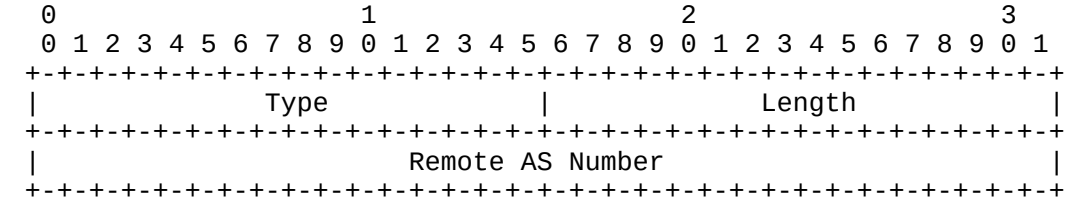

Figure 4: Remote AS Number TLV Format

The Remote AS Number field has 4 octets. When only 2 octets are used for the AS number (for example, when such information is advertised from OSPF), the left (high-order) 2 octets MUST be set to 0.

6.2. IPv4 Remote ASBR ID

The IPv4 Remote ASBR ID TLV specifies the IPv4 identifier of the remote ASBR to which the advertised inter-AS link connects. This can be any stable, routable IPv4 address of the remote ASBR. The use of the TE Router ID, as specified in the Traffic Engineering Router ID TLV [RFC9346] is RECOMMENDED.

The IPv4 Remote ASBR ID TLV is TLV Type 271 and is 4 octets in length. Its format is as follows:

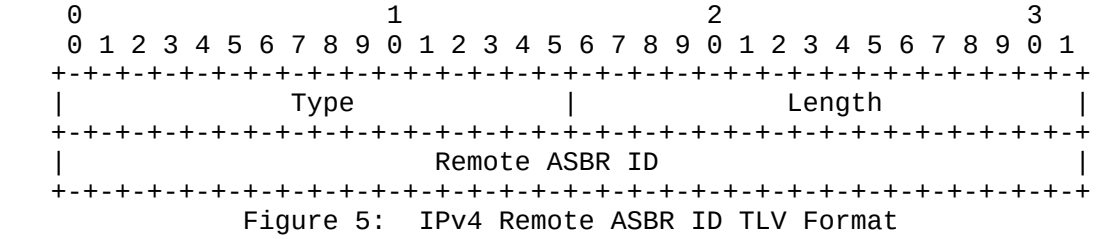

Figure 5: IPv4 Remote ASBR ID TLV Format

6.3. IPv6 Remote ASBR ID

The IPv6 Remote ASBR ID TLV specifies the IPv6 identifier of the remote ASBR to which the advertised inter-AS link connects. This can be any stable, routable IPv6 address of the remote ASBR. The use of the TE Router ID, as specified in the IPv6 Traffic Engineering Router ID TLV [RFC9346] is RECOMMENDED.

The IPv6 Remote ASBR ID TLV is TLV Type 272 and is 16 octets in length. Its format is as follows:

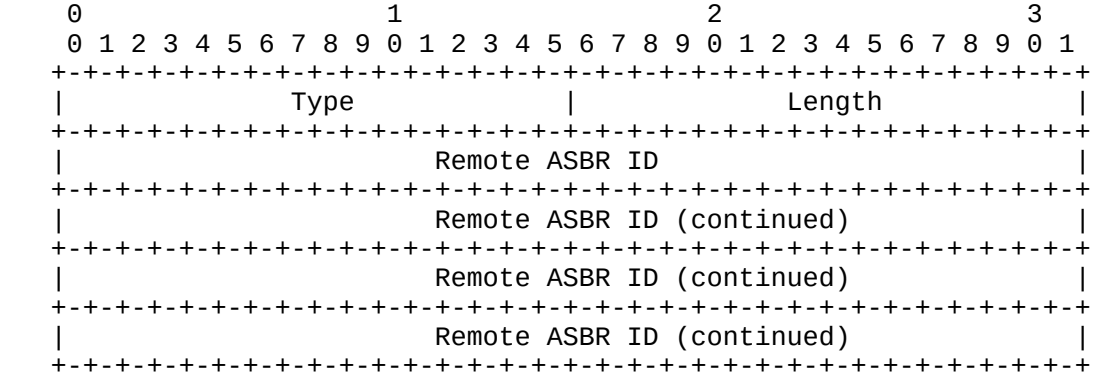

Figure 6: IPv6 Remote ASBR ID TLV Format

The IPv6 Remote ASBR ID TLV MUST be included if the neighboring ASBR has an IPv6 address. If the neighboring ASBR does not have an IPv6 address, the IPv4 Remote ASBR ID TLV MUST be included instead. Both an IPv4 Remote ASBR ID TLV and an IPv6 Remote ASBR ID TLV MAY be present in an inter-AS Link NLRI.

7. Advertisement of IGP Information for Inter-AS Links

Advertisement of inter-AS Links along with their TE information is done is done in IGPs as follows:

- In OSPFv2 via the inter-AS-TE-v2 LSA [RFC5392]

- In OSPFv3 via the inter-AS-TE-v3 LSA[RFC5392]

- In IS-IS via the inter-AS Reachability Information TLV (TLV 141) [RFC9346]

The routers that connect to the SDN controller via the BGP-LS protocol within each domain will advertise the information from the above IGPs for the inter-AS link. To retrieve the inter-AS topology, the Autonomous System TLV(TLV 512), Remote AS number, IPv6 Remote ASBR ID and/or IPv4 Remote ASBR ID MUST be presented within the inter-AS Link NLRI of each inter-AS link.

When advertising these inter-AS Links from the IGPs into BGP-LS as inter-AS Links, the sourcing of information for the inter-AS Link NLRI except for the inter-AS Link Descriptors follows the same procedures as specified in [RFC9552]. The information about the Remote AS Number and the IPv4/IPv6 Remote ASBR IDs specified in Section 6 are derived from the Remote AS Number and IPv4/IPv6 Remote ASBR ID TLVs specified for OSPF and IS-IS in [RFC5392] and [RFC9346] respectively. The rest of the inter-AS Link Descriptor TLVs of the inter-AS Link NLRI are sourced from the base OSPF/ISIS TE TLVs that were originally introduced for normal IGP links and which are also encoded for the inter-AS TE links as specified in [RFC5392] and [RFC9346]; their procedures are therefore the same as in [RFC9552].

The OSPF/ISIS inter-AS Link advertisements also include various link properties (e.g., TE metric, Admin Groups, SRLGs, etc.) which are encoded using the same TLVs as for normal IGP links. These link properties are advertised using their corresponding BGP-LS TLVs as specified in [RFC9552] and other BGP-LS extensions in the BGP-LS Attribute associated with the inter-AS Link NLRI of that specific link.

8. Solutions for other Alternative Scenarios

In some scenario, it is possible that the router running BGP-LS acts as also the role of the ASBR. Take the topology in following figure as the example:

In this alternative topology, it is SB1, which is also the ASBR of IGP A, runs BGP-LS protocol with the IP SDN controller. All other information are kept the same as that in Figure 1.

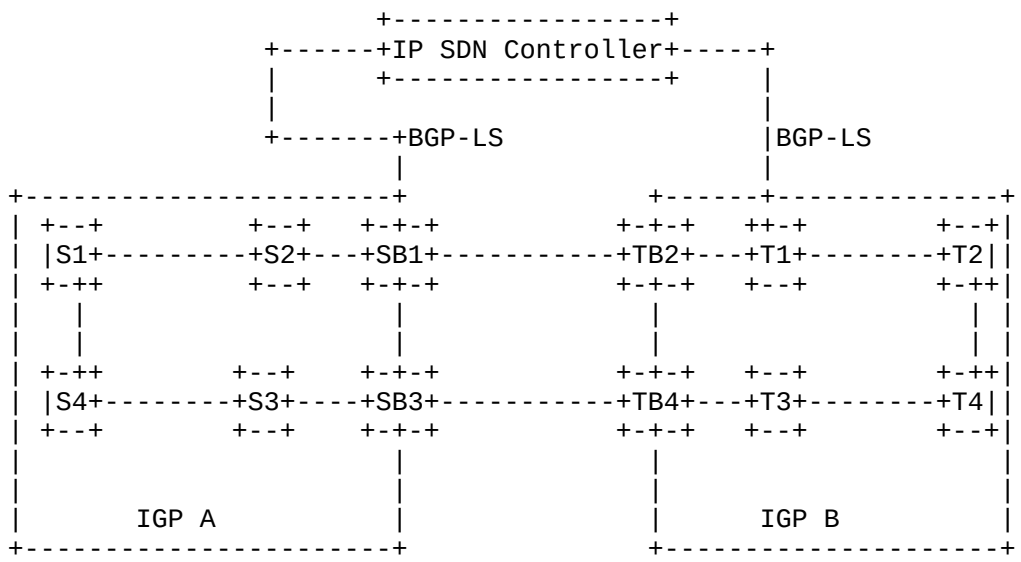

Figure 7: Alternative Inter-AS Domain Scenario

To retrieval the inter-AS topology among IGP A and IGP B in above alternative scenario, one possible solution is to utilize the BGP EPE [RFC9086] solution on SB1(reports the local information via the Link NLRI with 'protocol-id' set to 'BGP'), together with solution proposed in this document(reports the Inter-as link information on another ASBR via the inter-AS Link NLRI, with 'protocol-id' set to 'OSPF' or 'IS-IS').

Considering the inter-AS link information on SB1 is configured locally, and they are either from direct connect interfaces, or from static configuration routes, using also the inter-AS NLRI, with the 'protocol-id' set to either 'direct' or 'static' is aligned well with the definition of [RFC9552], and can also keep the implementation and deployment of the solution universal, regardless of whether the routers runs BGP-LS with the controller is within the AS or at the border of the AS.

All the inter-AS topology information will be coming from the inter-AS Link NLRI, and is easy to be distinguished with other information from Link NLRI of IGP.

9. End-to-End Inter-AS Topology Use Case

[RFC8735] section 3.3 "Traffic Engineering for Multi-domain" describes the scenario that the service provider needs to do traffic engineering from end to end that spans multiple domains. To program the end to end path, and let the traffic pass the non-congested links, especially the links that interconnect to the different domains, the SDN controller that covers all these domains which belong to the service provider needs to know the inter-AS topology information from the inter-AS Link NLRI via the BGP-LS. It binds the half link information from each domain together, calculates the end-to-end path for the assured service traffic and uses the final path information to control the transmission of the assured traffic.

The use case can also be explained via the scenario described in Figure 1.

If S1 wants to send traffic to T2 with assured performance, it should know the end-to-end path from the IP SDN controller, especially the connection topology among the four border routers(SB1, SB3, TB2 and TB4).

Such connection topology information is sent by S2 and T1 respectively via the BGP-LS protocol, to the IP SDN controller. The IP SDN controller can then stick the two IGP domain topologies together, calculate the end-to-end path that spans two IGP domains and their inter-connection links, guide the traffic from S1 (in IGP A) to T2 (IGP B) traverse the desired nodes and links.

If there are no inter-connection links from the inter-AS NLRI that is defined in this document, the IP SDN controller can’t stick the two IGP domains together and can’t then calculate the end-to-end path for the communication nodes located in different domains.

10. Security Considerations

BGP-LS security is specified in [RFC9552]. This extension to BGP-LS focuses on scenarios where a single entity-operated network includes multiple IGP domains composed of its backbone network, several Metropolitan-Area Networks (MANs), and Data Centers (DCs). The configuration of these networks, operated by a single administrative entity, creates a "walled garden". Within this single administrative domain, the network operator needs to monitor and engineer traffic flows traversing a network that spans multiple Autonomous Systems (ASes). The network operator can obtain this inter-AS topology information via the procedure described in this document.

A single administrative domain consisting of two ASes that passes information about inter-AS Link characteristics does not cause issues within a "walled garden". However, the inter-AS Link NLRI and its characteristics (Link/Local Identifier, IPv4 Interface Address, IPv4 Neighbor Address, IPv6 Interface Address, IPv6 Neighbor Address, Multi-Topology Identifier, Remote-AS Number, IPv4 Remote ASBR ID, and IPv6 Remote ASBR ID) constitute critical network information. As such, operators SHOULD handle this critical information in a manner that restricts it to the walled garden.

11. IANA Considerations

This document requests IANA to update the allocated codepoints from under the "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry group as follows:

11.1. New BGP-LS NLRI type

IANA has allocated a codepoint for the inter-AS Link NLRI type from the "BGP-LS NLRI Types" registry in the "Border Gateway Protocol Link State (BGP-LS) Parameter" Group:

Type	NLRI Type	Reference
7	inter-AS Link NLRI	This document

Figure 8: Inter-AS Link NLRI Codepoint

11.2. New Inter-AS Link Descriptors

IANA has allocated codepoints for the following TLVs from "BGP-LS NLRI and Attribute TLVs" registry in the "Border Gateway Protocol Link State (BGP-LS) Parameter" Group:

TLV Code Point	Description	Reference
270	Remote AS Number	This document
271	IPv4 Remote ASBR ID	This document
272	IPv6 Remote ASBR ID	This document

Figure 9: BGP-LS Link Descriptors TLV

12. Acknowledgement

The author would like to thank Susan Hares, Acee Lindem, Jie Dong, Shaowen Ma, Jeff Tantsura and Dhruv Dhody for their valuable comments and suggestions.

13. References

13.1. 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>.

[RFC5392] Chen, M., Zhang, R., and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5392, DOI 10.17487/RFC5392, January 2009, <https://www.rfc-editor.org/info/rfc5392>.

[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>.

[RFC8735] Wang, A., Huang, X., Kou, C., Li, Z., and P. Mi, "Scenarios and Simulation Results of PCE in a Native IP Network", RFC 8735, DOI 10.17487/RFC8735, February 2020, <https://www.rfc-editor.org/info/rfc8735>.

[RFC9346] Chen, M., Ginsberg, L., Previdi, S., and D. Xiaodong, "IS-IS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 9346, DOI 10.17487/RFC9346, February 2023, <https://www.rfc-editor.org/info/rfc9346>.

[RFC9552] Talaulikar, K., Ed., "Distribution of Link-State and Traffic Engineering Information Using BGP", RFC 9552, DOI 10.17487/RFC9552, December 2023, <https://www.rfc-editor.org/info/rfc9552>.

13.2. Informative References

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

Authors' Addresses

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing
Beijing, 102209
China
Email: wangaj3@chinatelecom.cn

Huaimo Chen
Individual
Boston, MA
United States of America
Email: hchen.ietf@gmail.com

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

Shunwan Zhuang
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing
100095
China
Email: zhuangshunwan@huawei.com

Changwang Lin
New H3C Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing
China
Email: linchangwang.04414@h3c.com