

IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: 27 October 2024

A. Wang
China Telecom
H. Chen
Futurewei
K. Talaulikar
Cisco Systems
S. Zhuang
Huawei Technologies
C. Lin
New H3C Technologies
25 April 2024

BGP-LS Extension for Inter-AS Topology Retrieval
draft-ietf-idr-bgppls-inter-as-topology-ext-15

Abstract

This document describes the process to build Border Gateway Protocol-Link State (BGP-LS) key parameters in inter-domain scenario, defines one new BGP-LS Network Layer Reachability Information (NLRI) type (Stub Link NLRI) and some new Type-Length-Values (TLVs) for BGP-LS to let Software Definition Network (SDN) controller retrieve the network topology automatically under various inter-AS environments.

Such extension and process can enable the network operator to collect the interconnect information between different domains and then calculate the overall network topology automatically based on the information provided by 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 27 October 2024.

Copyright Notice

Copyright (c) 2024 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
2. Conventions used in this document	3
3. Terminology	3
4. Inter-AS Domain Scenarios	3
5. Stub Link NLRI	4
6. IGP Information for Inter-AS Link	5
7. BGP-LS Extensions for Inter-AS Link	5
7.1. Remote AS Number TLV	6
7.2. IPv4 Remote ASBR ID	7
7.3. IPv6 Remote ASBR ID	7
8. Topology Reconstruction.	8
9. Security Considerations	8
10. IANA Considerations	8
10.1. New BGP-LS NLRI type	9
10.2. New Stub Link Descriptors	9
11. Acknowledgement	9
12. References	9
12.1. Normative References	9
12.2. Informative References	10
Authors' Addresses	10

1. Introduction

BGP-LS [RFC9552] describes the methodology that using BGP protocol to transfer the Link-State information. Such method can enable SDN controller to collect the underlay network topology automatically, but normally it can only get the information within one Interior Gateway Protocol (IGP) domain. If the operator has more than one IGP domain, and these domains interconnect with each other, there is no mechanic within current BGP-LS to transfer the interconnect topology information.

Draft [RFC9086] defines some extensions for exporting BGP peering node topology information (including its peers, interfaces and peering ASs) in a way that is exploitable in order to compute efficient BGP Peering Engineering policies and strategies. Such information can also be used to calculate the interconnection topology among different IGP domains, but it requires every border router to run BGP-LS protocol and report the information to SDN controller. Considering there will be plenty of border routers on the network boundary, such solution restricts its deployment flexibility.

This draft analysis the situations that the SDN controller needs to get the interconnected topology information between different AS domains, defines the new Stub Link NLRI and some new TLVs within the BGP-LS protocol to transfer the key information related to them. After that, the SDN controller can then deduce the multi-domain topology automatically based on the information from BGP-LS protocol.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119] .

3. Terminology

The following terms are defined in this document:

- * IDCs: Internet Data Centers
- * MAN: Metro-Area-Network
- * SDN: Software Definition Network

4. Inter-AS Domain Scenarios

Figure 1 illustrates the multi-domain scenarios that this draft discusses. Normally, SDN Controller can get the topology of IGP A and IGP B individually via the BGP-LS protocol, but it can't get the topology connection information between these two IGP domains because there is generally no IGP protocol run on the connected links.

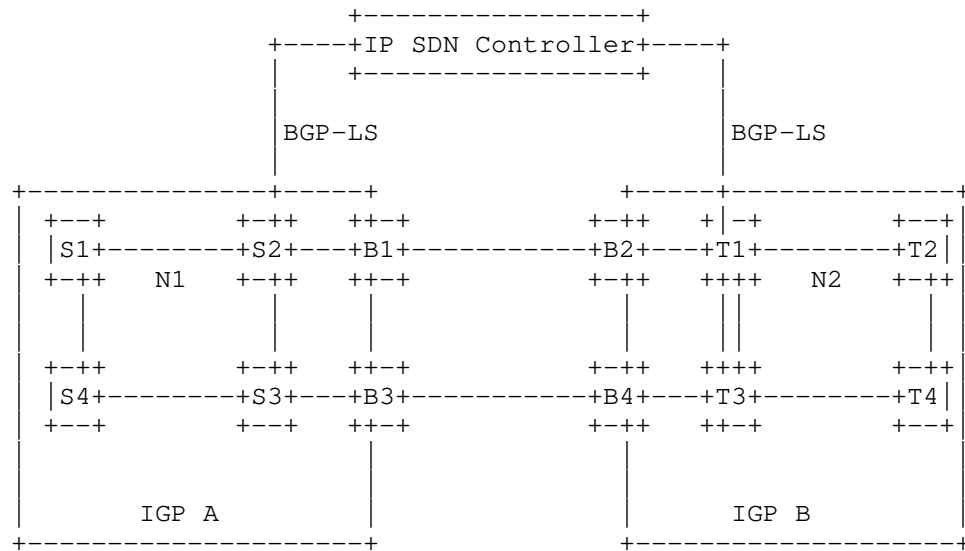


Figure 1: Inter-AS Domain Scenarios

5. Stub Link NLRI

[RFC9552] defines four NLRI types(Node NLRI, Link NLRI, IPv4 Topology Prefix NLRI, IPv6 Topology Prefix NLRI) to transfer the topology and prefix information. For inter-as link, the two ends of the link locates in different IGP domains, then it is not appropriate to transfer their information within the current defined NLRI types.

This draft defines one new NLRI type, called Stub Link NLRI, which is coded as the following format:

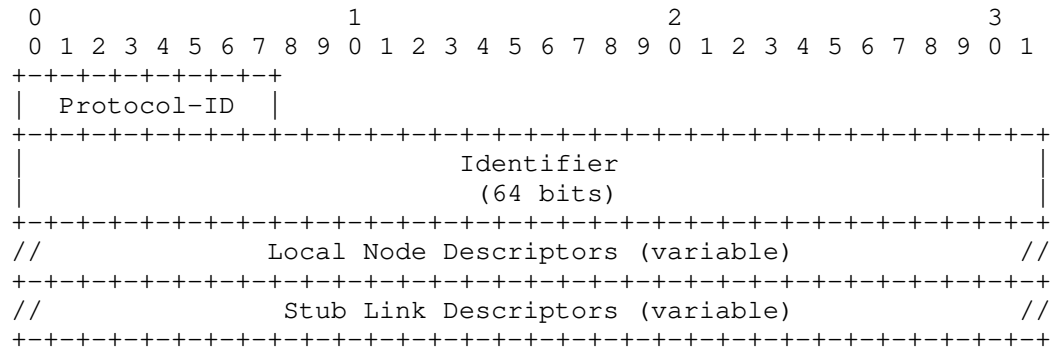


Figure 2: Stub Link NLRI Format

The "Protocol-ID" should be set to the value that indicates the source protocol of the stub link information, as indicated in [RFC9552]

The semantics of "Local Node Descriptors" and "Stub Link Descriptors" are same as that defined in [RFC9552] for "Node Descriptors" and "Link Descriptor".

This newly defined NLRI can be used to describe the link that has only one end located within the IGP domain, as described in the following sections. The Node/Link/Prefix Descriptor sub-TLVs and Node/Link/Prefix attributes that defined in [RFC9552] can be included in the NLRI if necessary. The interface and neighbor address sub-TLVs SHOULD be included in the Local Node Descriptors to differentiate the parallel links between two ASBRs.

6. IGP Information for Inter-AS Link

[RFC9346] and [RFC5392] define IS-IS and OSPF extensions respectively to deal with the situation for inter-AS link. Three sub-TLVs (Remote AS Number, IPv4 Remote ASBR ID, IPv6 Remote ASBR ID) which are associated with the inter-AS link are defined.

These TLVs are flooded within the IGP domain automatically. They should be carried within the newly defined Stub Link NLRI within the BGP-LS protocol, as the descriptors for the inter-AS stub link.

The "Local Node Descriptors" should describe the characteristics of ASBRs that are connected these inter-AS links.

If the SDN controller knows these information via one of the interior router that runs BGP-LS protocol, the SDN controller can rebuild the inter-AS topology correctly according to the procedure described in Section 8

7. BGP-LS Extensions for Inter-AS Link

This section proposes to add three new TLVs that is included within the Stub Link NLRI to transfer the information via BGP-LS, which are required to build the inter-AS related topology by the SDN controller.

The following Link Descriptor TLVs are added into the BGP-LS protocol :

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

Figure 3: Stub Link Descriptor TLVs

Detail encoding of these TLVs are synchronized with the corresponding parts in [RFC9346] and [RFC5392], which keeps the BGP-LS protocol agnostic to the underly protocol.

7.1. Remote AS Number TLV

A new TLV, the remote AS number TLV, is defined for inclusion in the link descriptor when advertising inter-AS links. 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 TBD (see Section 10) and is 4 octets in length. The format is as follows:

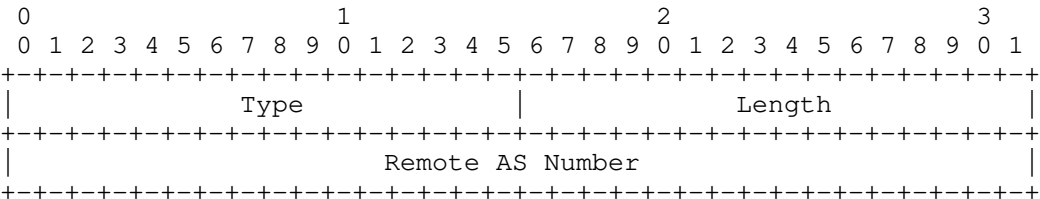


Figure 4: Remote AS Number TLV Format

The Remote AS number field has 4 octets. When only 2 octets are used for the AS number, as in current deployments, the left (high-order) 2 octets MUST be set to 0. The remote AS number TLV MUST be included when a router advertises an inter-AS link.

7.2. IPv4 Remote ASBR ID

A new TLV, which is referred to as the IPv4 remote ASBR ID TLV, is defined for inclusion in the link descriptor when advertising inter-AS links. The IPv4 remote ASBR ID TLV specifies the IPv4 identifier of the remote ASBR to which the advertised inter-AS link connects. This could be any stable and routable IPv4 address of the remote ASBR. 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 TBD (see Section 10) and is 4 octets in length. The format of the IPv4 remote ASBR ID TLV is as follows:

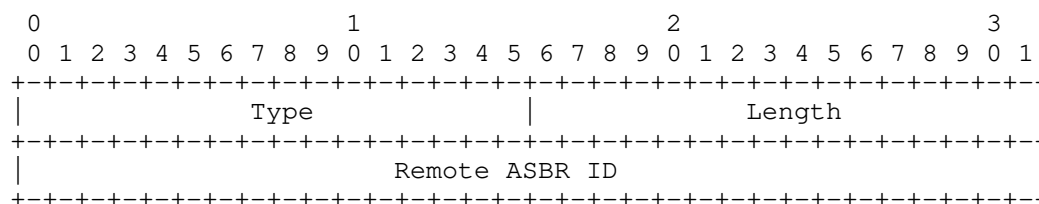


Figure 5: IPv4 Remote ASBR ID TLV Format

The IPv4 remote ASBR ID TLV MUST be included if the neighboring ASBR has an IPv4 address. If the neighboring ASBR does not have an IPv4 address (not even an IPv4 TE Router ID), the IPv6 remote ASBR ID TLV MUST be included instead. An IPv4 remote ASBR ID TLV and IPv6 remote ASBR ID TLV MAY both be present in an inter-AS stub link NLRI.

7.3. IPv6 Remote ASBR ID

A new TLV, which is referred to as the IPv6 remote ASBR ID TLV, is defined for inclusion in the link descriptor when advertising inter-AS links. The IPv6 remote ASBR ID TLV specifies the IPv6 identifier of the remote ASBR to which the advertised inter-AS link connects. This could be any stable and routable IPv6 address of the remote ASBR. 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 TBD (see Section 10) and is 16 octets in length. The format of the IPv6 remote ASBR ID TLV is as follows:

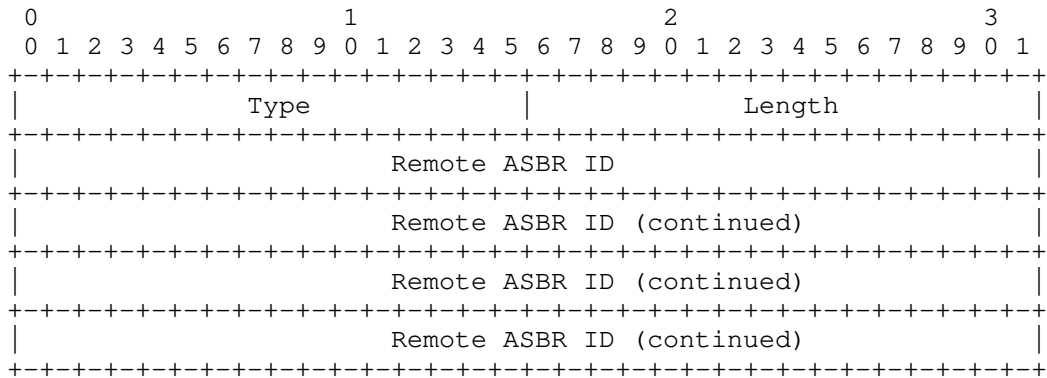


Figure 6: IPv6 Remote ASBR ID TLV Format

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. An IPv4 remote ASBR ID TLV and IPv6 remote ASBR ID TLV MAY both be present in an inter-AS stub link NLRI.

8. Topology Reconstruction.

When SDN controller gets such information from BGP-LS protocol, it should find the corresponding associated router information, build the link between these two border routers. After iterating the above procedures for all of the stub links, the SDN controller can then retrieve the connection topology between different domains automatically.

9. Security Considerations

It is common for one operator to occupy several IGP domains that are composited by its backbone network and several MAN(Metrio-Area-Network)s/Internet Data Centers (IDCs). When they do traffic engineering which spans MAN, Backbone and IDC, they need to know the inter-as topology via the process described in this draft. Using the passive interface features or configuring the Traffic Engineering (TE) parameters on the interconnect links will not spread the topology fluctuation across each other domain.

10. IANA Considerations

This document defines:

- * A new BGP NLRI Type: Stub Link NLRI. The codepoint is from the "BGP-LS NLRI Types"

- * Three new Link Descriptors TLV: Remote AS Number TLV, IPv4 Remote ASBR ID, IPv6 Remote ASBR ID. The codepoint are from "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" registry.

10.1. New BGP-LS NLRI type

This document defines a new value in the registry "BGP-LS NLRI Types":

Code Point	Description	Status
TBD(suggestion 7)	Stub Link NLRI	Allocation from IANA

Figure 7: Stub Link NLRI Codepoint

10.2. New Stub Link Descriptors

This document defines three new values in the registry "BGP-LS NLRI and Attribute TLVs":

Code Point	Description	Status
TBD(suggestion 270)	Remote AS Number	Allocation from IANA
TBD(suggestion 271)	IPv4 Remote ASBR ID	Allocation from IANA
TBD(suggestion 272)	IPv6 Remote ASBR ID	Allocation from IANA

Figure 8: BGP-LS Link Descriptors TLV

11. Acknowledgement

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

12. References

12.1. 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>>.
- [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>>.

12.2. 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
Futurewei
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