

IDR Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: October 10, 2019

A. Wang  
China Telecom  
H. Chen  
Huawei Technologies  
S. Ma  
Mellanox Technologies  
April 8, 2019

**BGP-LS Extension for Inter-AS Topology Retrieval**  
**draft-ietf-idr-bgpls-inter-as-topology-ext-02**

Abstract

This document describes the process to build BGP-LS key parameters in multi-domain scenario, defines one new BGP-LS NLRI type(Stub Link NLRI) and some new inter-AS TE related TLVs for BGP-LS to let SDN controller retrieve the network topology automatically under various environments.

Such process and extension 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 October 10, 2019.

Copyright Notice

Copyright (c) 2019 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	Conventions used in this document . . . . .	<a href="#">3</a>
<a href="#">3.</a>	Inter-AS Domain Scenarios. . . . .	<a href="#">3</a>
<a href="#">4.</a>	Stub Link NLRI . . . . .	<a href="#">3</a>
<a href="#">4.1.</a>	Inter-AS Native IP Scenario . . . . .	<a href="#">4</a>
<a href="#">4.2.</a>	Inter-AS TE Scenario . . . . .	<a href="#">4</a>
<a href="#">5.</a>	Inter-AS TE NLRI related TLVs . . . . .	<a href="#">5</a>
<a href="#">5.1.</a>	Remote AS Number TLV . . . . .	<a href="#">5</a>
<a href="#">5.2.</a>	IPv4 Remote ASBR ID . . . . .	<a href="#">6</a>
<a href="#">5.3.</a>	IPv6 Remote ASBR ID . . . . .	<a href="#">7</a>
<a href="#">6.</a>	Topology Reconstruction. . . . .	<a href="#">7</a>
<a href="#">7.</a>	Security Considerations . . . . .	<a href="#">8</a>
<a href="#">8.</a>	IANA Considerations . . . . .	<a href="#">8</a>
<a href="#">9.</a>	Acknowledgement . . . . .	<a href="#">8</a>
<a href="#">10.</a>	Normative References . . . . .	<a href="#">8</a>
	Authors' Addresses . . . . .	<a href="#">9</a>

## [1.](#) Introduction

BGP-LS [[RFC7752](#)] 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 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 [[I-D.ietf-idr-bgpls-segment-routing-epe](#)] 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 the border routers to run BGP-LS protocol and report the information to the PCE/SDN controller, which restricts the solution deployment flexibility.



## 2. Conventions used in this document

### 3. Inter-AS Domain Scenarios.

```

+-----+
+-----+IP SDN Controller+-----+
|           +-----+           |
|           +-----+           |
|           BGP-LS           |           BGP-LS
|           |           |           |
+-----+-----+-----+-----+
| +--+ +--+ +--+ +--+ +--+ +--+ +--+ |
| |S1+-----+S2+-----+B1+-----+B2+-----+T1+-----+T2+ |
| +--+ N1 +--+ +--+ +--+ +--+ +--+ N2 +--+ |
| | | | | | | | | | | | | | |
| | | | | | | | | | | | | | |
| +--+ +--+ +--+ +--+ +--+ +--+ +--+ |
| |S4+-----+S3+-----+B3+-----+B4+-----+T3+-----+T4+ |
| +--+ +--+ +--+ +--+ +--+ +--+ +--+ |
| | | | | | | | | | | | | | |
| | | | | | | | | | | | | | |
| IGP A | IGP B |
+-----+-----+-----+-----+

```

Figure 1: Inter-AS Domain Scenarios

#### 4. Stub Link NLRI

[RFC7752] 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:

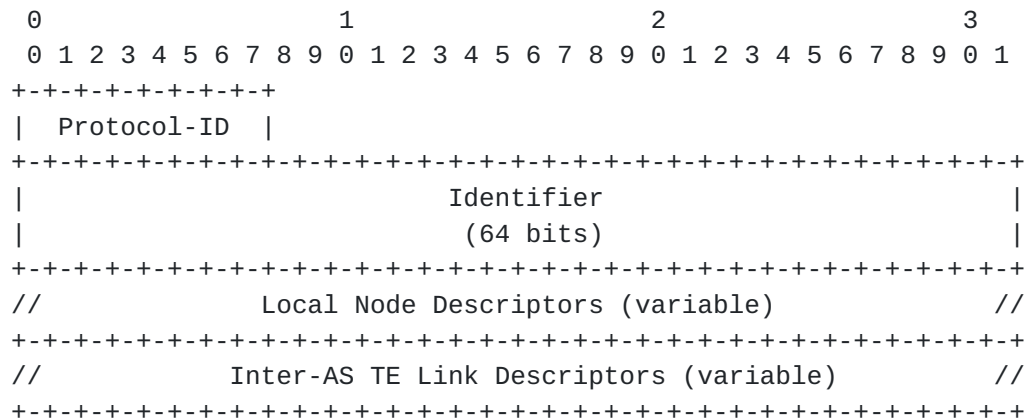


Figure 2: The Inter-AS TE Link NLRI Format

The semantics of "Stub Link Descriptors" is same as that defined in [\[RFC7752\]](#) for "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.

#### [4.1.](#) Inter-AS Native IP Scenario

When IGP A or IGP B runs native IS-IS/OSPF protocol, the operator can use passive feature for the inter-domain links to let the routers within the IGP domain know these links. Such stub links information can then be carried within the Stub Link NLRI to be reported via the BGP-LS protocol to the SDN controller.

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

When such information is reported via the BGP-LS protocol, the PCE/SDN controller can construct the underlay inter-domain topology according to procedure described in [section 6](#).

#### [4.2.](#) Inter-AS TE Scenario

When IGP A or IGP B run IS-IS TE/OSPF-TE protocol, [\[RFC5316\]](#) and [\[RFC5392\]](#) define the IS-IS and OSPF extensions respectively to deal with the situation for inter-AS traffic engineering. Three new sub-



TLVs(Remote AS Number&#12289;IPv4 Remote ASBR ID&#12289;IPv6 Remote ASBR ID) which are associated with the inter-AS TE link are defined.

These TLVs are flooded within the IGP domain automatically. They can also 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 the characteristics of ASBRs that are connected these inter-AS TE links.

If the PCE/SDN controller know these information via one of the interior router that runs BGP-LS protocol, the PCE/SDN controller can rebuild the inter-AS TE topology correctly according to the procedure described in [section 6](#)

## 5. Inter-AS TE NLRI related TLVs

This draft 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 TE related topology by the PCE/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	<a href="#">[RFC5316]</a> /3.3.1
			<a href="#">[RFC5392]</a> /3.3.1
TBD	IPv4 Remote ASBR ID	25/22	<a href="#">[RFC5316]</a> /3.3.2
			<a href="#">[RFC5392]</a> /3.3.2
TBD	IPv6 Remote ASBR ID	26/24	<a href="#">[RFC5316]</a> /3.3.3
			<a href="#">[RFC5392]</a> /3.3.3

Figure 3: Link Descriptor TLVs

Detail encoding of these TLVs are synchronized with the corresponding parts in [\[RFC5316\]](#) and [\[RFC5392\]](#), which keeps the BGP-LS protocol is agnostic to the underly protocol.

### 5.1. Remote AS Number TLV

A new TLV, the remote AS number TLV, is defined for inclusion in the link descriptor when advertising inter-AS TE links. The remote AS number TLV specifies the AS number of the neighboring AS to which the advertised link connects.









### 5.3. IPv6 Remote ASBR ID

A new TLV, which is referred to as the IPv6 remote ASBR ID TLV, is defined for inclusion in the inter-AS reachability TLV 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 [[RFC6119](#)] is RECOMMENDED.

The IPv6 remote ASBR ID TLV is TLV type TBD (see [Section 8](#)) and is 16 octets in length. The format of the IPv6 remote ASBR ID TLV is as follows:

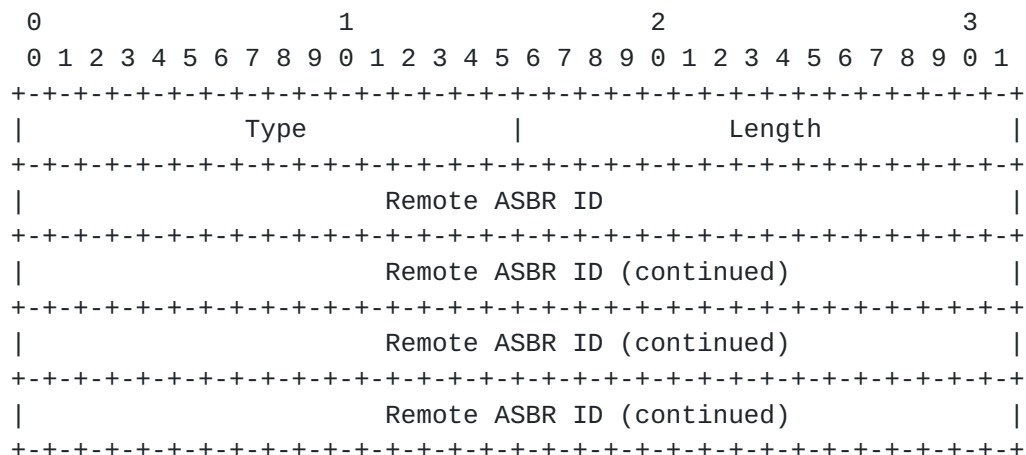


Figure 6: IPv6 Remote ASBR ID TLV Format

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 TE link NLRI.

## 6. Topology Reconstruction.

When SDN Controller gets such information from BGP-LS protocol, it should compares the proximity of these stub links. If they are under the same network scope, then 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.



## 7. 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/IDCs. When they do traffic engineering which spans MAN-backbone-IDC, they need to know the inter-as topology via the process described in this draft. Using the passive interface features or configuring the TE parameters on the interconnect links will not spread the topology fluctuation across each other domain.

## 8. IANA Considerations

TBD.

## 9. Acknowledgement

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

## 10. Normative References

- [I-D.ietf-idr-bgp-ls-segment-routing-ext]  
Previdi, S., Talaulikar, K., Filsfils, C., Gredler, H., and M. Chen, "BGP Link-State extensions for Segment Routing", [draft-ietf-idr-bgp-ls-segment-routing-ext-12](#) (work in progress), March 2019.
- [I-D.ietf-idr-bgp-ls-segment-routing-epe]  
Previdi, S., Talaulikar, K., Filsfils, C., Patel, K., Ray, S., and J. Dong, "BGP-LS extensions for Segment Routing BGP Egress Peer Engineering", [draft-ietf-idr-bgp-ls-segment-routing-epe-18](#) (work in progress), March 2019.
- [I-D.ietf-ospf-ospfv3-lsa-extend]  
Lindem, A., Roy, A., Goethals, D., Vallem, V., and F. Baker, "OSPFv3 LSA Extendibility", [draft-ietf-ospf-ospfv3-lsa-extend-23](#) (work in progress), January 2018.
- [I-D.ietf-teas-native-ip-scenarios]  
Wang, A., Huang, X., Qou, C., Li, Z., and P. Mi, "Scenario, Simulation and Suggestion of PCE in Native IP Network", [draft-ietf-teas-native-ip-scenarios-02](#) (work in progress), October 2018.
- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", [RFC 1195](#), DOI 10.17487/RFC1195, December 1990, <<https://www.rfc-editor.org/info/rfc1195>>.



- [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>>.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, [RFC 2328](#), DOI 10.17487/RFC2328, April 1998, <<https://www.rfc-editor.org/info/rfc2328>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", [RFC 5305](#), DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", [RFC 5316](#), DOI 10.17487/RFC5316, December 2008, <<https://www.rfc-editor.org/info/rfc5316>>.
- [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>>.
- [RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", [RFC 6119](#), DOI 10.17487/RFC6119, February 2011, <<https://www.rfc-editor.org/info/rfc6119>>.
- [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>>.
- [RFC7794] Ginsberg, L., Ed., Decraene, B., Previdi, S., Xu, X., and U. Chunduri, "IS-IS Prefix Attributes for Extended IPv4 and IPv6 Reachability", [RFC 7794](#), DOI 10.17487/RFC7794, March 2016, <<https://www.rfc-editor.org/info/rfc7794>>.
- [RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and F. Baker, "OSPFv3 Link State Advertisement (LSA) Extensibility", [RFC 8362](#), DOI 10.17487/RFC8362, April 2018, <<https://www.rfc-editor.org/info/rfc8362>>.

Authors' Addresses





Aijun Wang  
China Telecom  
Beiqijia Town, Changping District  
Beijing, Beijing 102209  
China

Email: wangaj.bri@chinatelecom.cn

Huaimo Chen  
Huawei Technologies  
Boston, MA  
USA

Email: Huaimo.chen@huawei.com

Shaowen Ma  
Mellanox Technologies

Email: mashaowen@gmail.com

