IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: February 15, 2019

A. Wang
China Telecom
H. Chen
Huawei Technologies

August 14, 2018

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

Abstract

This document describes the process to build BGP-LS key parameters in Native IP multi-domain scenario and defines 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 expand the usage of BGP-LS protocol to multi- domain, enable the network operator to collect the connection relationship between different AS 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 February 15, 2019.

Copyright Notice

Table of Contents

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 general TLV within current BGP- LS to transfer the interconnect 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 to collect this information and report them to the PCE/SDN controller, which restricts the deployment flexibility of BGP-LS protocol.

This draft analysizes the situations that the PCE/SDN controller needs to get about the inter-connected information between different AS domains, defines new TLVs to extend the BGP-LS protocol to

transfer the key information related to the interconnect TE topology. 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. Inter-AS Domain Scenarios.

Fig.1 illustrates the multi-domain scenarios that this draft discussed. 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.

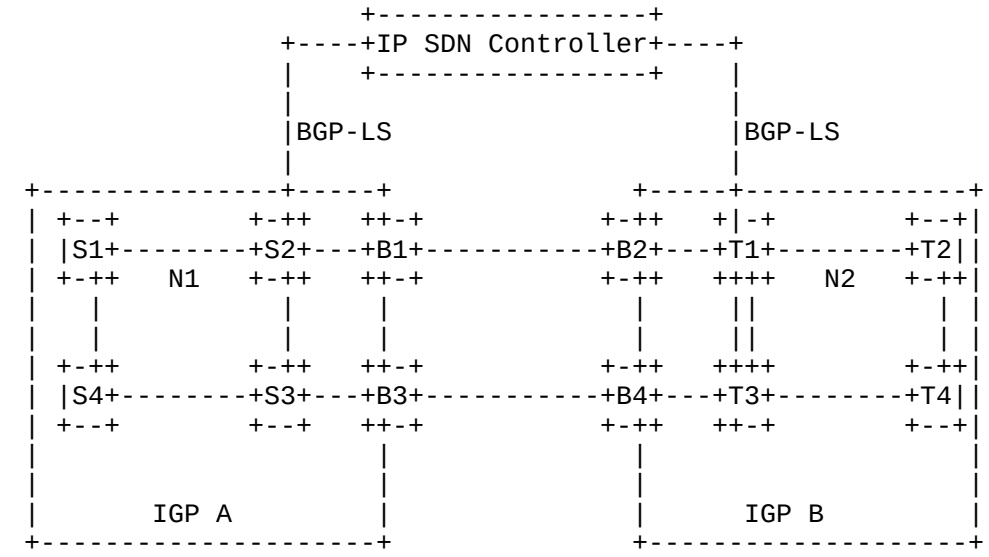


Figure 1: Inter-AS Domain Scenarios

3.1. IS-IS/OSPF Inter-AS Native IP Scenario

When the IGP A or IGP B runs native IS-IS/OSPF protocol, the operator often redistributes the IPv4/IPv6 prefixes of interconnect links into IS-IS/OSPF protocol to ensure the inter-domain connectivity.

If the IGP runs IS-IS protocol, the redistributed link information will be carried in IP External Reachability Information TLV within

the Level 2 PDU type that defined in [RFC1195], every router within the IGP domain can deduce the redistributed router from the IS-IS LSDB.

If the IGP runs OSPF protocol,[RFC2328]defines the type 5 external LSA to transfer the external IPv4 routes; [I-D.ietf-ospf-ospfv3-lsa-extend] defines the "External-Prefix TLV" to transfer the external IPv6 routes; these LSAs have also the advertising router information that initiates the redistribute activity. Every router within IGP domain can also deduce the redistributed router from the OSPF LSDB.

For prefix information that associated with each router, BGP-LS [RFC7752] defines the Prefix NLRI which is illustrated below:

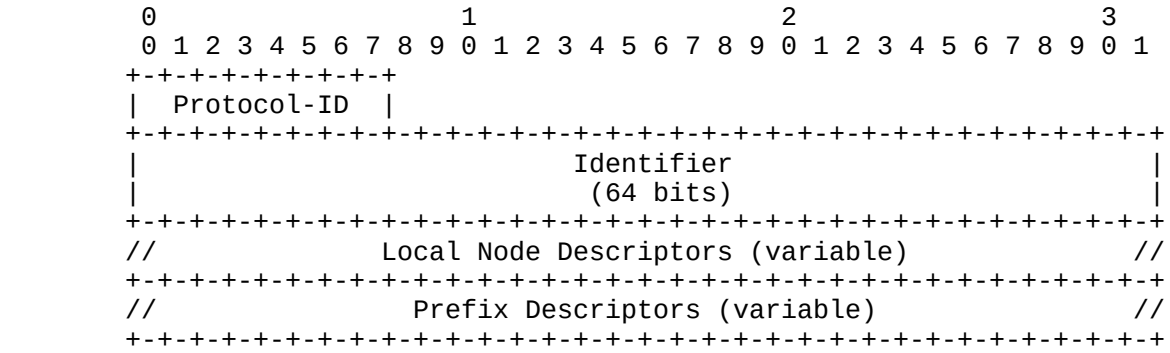


Figure 2: The IPv4/IPv6 Topology Prefix NLRI Format

For these redistributed inter-domain links, their prefix information should be included in the "Prefix Descriptor", and the associated redistributed router information should be included in the "Local Node Descriptors".

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 5

3.2. IS-IS/OSPF Inter-AS TE Scenario

[RFC5316] and [RFC5392] define the IS-IS and OSPF extensions respectively to deal with the requirements for inter-AS traffic engineering. They define some new sub-TLVs(Remote AS Number、IPv4 Remote ASBR ID、IPv6 Remote ASBR ID) which are associated with the inter-AS TE link TLVs to report the TE topology between different domains.

These TLVs are flooded within the IGP domain automatically. If the PCE/SDN controller can 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.

4. Inter-AS TE related TLVs

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

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

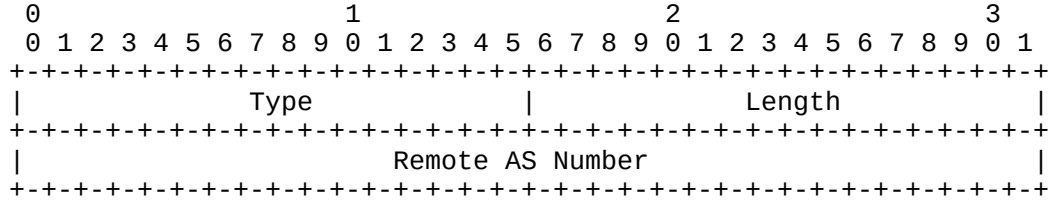
TLV Code Point	Description	IS-IS/OSPF TLV /Sub-TLV	Reference (RFC/Section)
TBD	Remote AS Number	24/21	[RFC5316]/3.3.1
			[RFC5392]/3.3.1
TBD	IPv4 Remote ASBR ID	25/22	[RFC5316]/3.3.2
			[RFC5392]/3.3.2
TBD	IPv6 Remote ASBR ID	26/24	[RFC5316]/3.3.3
			[RFC5392]/3.3.3

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.

4.1. 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 7) and is 4 octets in length. The format is as follows:

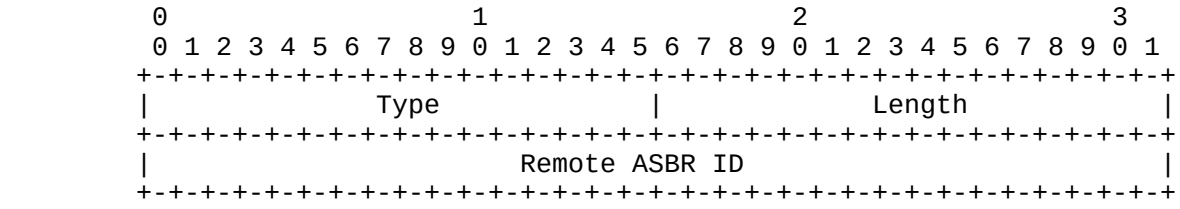


The Remote AS number field has 4 octets. When only 2 octets are used for the AS number, as in current deployments, the left (high-order) 2 octets MUST be set to 0. The remote AS number TLV MUST be included when a router advertises an inter-AS TE link.

4.2. 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 [RFC5305] is RECOMMENDED.

The IPv4 remote ASBR ID TLV is TLV type TBD (see Section 7) and is 4 octets in length. The format of the IPv4 remote ASBR ID sub-TLV is as follows:

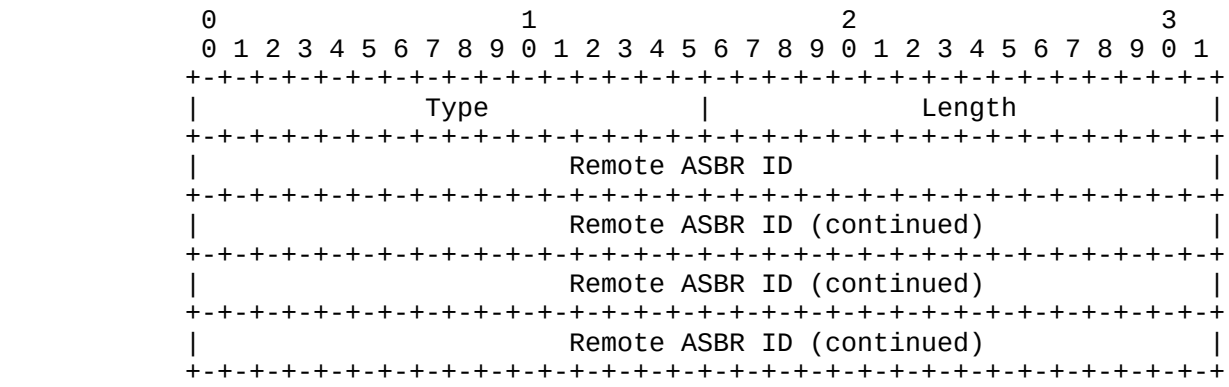


The IPv4 remote ASBR ID TLV MUST be included if the neighboring ASBR has an IPv4 address. If the neighboring ASBR does not have an IPv4 address (not even an IPv4 TE Router ID), the IPv6 remote ASBR ID TLV MUST be included instead. An IPv4 remote ASBR ID TLV and IPv6 remote ASBR ID TLV MAY both be present in an extended IS reachability TLV.

4.3. 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 7) and is 16 octets in length. The format of the IPv6 remote ASBR ID TLV is as follows:



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 extended IS reachability TLV.

5. Topology Reconstruction.

When SDN Controller gets such information from BGP-LS protocol, it should compares the proximity of the redistributed prefixes. 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 redistributed prefixes, the SDN controller can then retrieve the connection topology between different domains automatically.

6. Security Considerations

It is common for one operator to occupy several IGP domains that composited by its backbone network and several MAN(Metrio-Area-Network)s/IDCs. When they do traffic engineering from end to end that spans MAN-backbone-IDC, they need to know the inter-as topology via the process described in this draft. Then it is naturally to redistribute the interconnection prefixes in Native IP scenario.

If these IGP domains belong to different operators, it is uncommon do inter-as traffic engineering under one PCE/SDN controller, then it is unnecessary to get the inter-as topology. But redistributing the interconnection prefixes will do no harm to their networks, because the redistributed interconnection link prefixes belongs to both of them, they are also the interfaces addresses on the border routers. .

7. IANA Considerations

TBD.

8. Acknowledgement

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

9. 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-08 (work in progress), May 2018.

[I-D.ietf-idr-bgpls-segment-routing-epe]
Previdi, S., Filsfils, C., Patel, K., Ray, S., and J. Dong, "BGP-LS extensions for Segment Routing BGP Egress Peer Engineering", draft-ietf-idr-bgpls-segment-routing-epe-15 (work in progress), March 2018.

[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., Huang, L., and P. Mi, "CCDR Scenario, Simulation and Suggestion", draft-ietf-teas-native-ip-scenarios-01 (work in progress), June 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