SPRING S. Hegde
Internet-Draft S. Sangli
Intended status: Standards Track Juniper Networks Inc.
Expires: September 12, 2019 March 11, 2019

BGP-LS Extensions for Inter-As TE using EPE based mechanisms
draft-hegde-idr-bgp-ls-epe-inter-as-00

Abstract

In certain network deployments, a single operator has multiple Autonomous Systems(AS) to facilitate ease of management. A multiple AS network design could also be a result of network mergers and acquisitions. In such scenarios, a centralized Inter-domain TE approach could provide most optimal allocation of resources and the most controlled path placement. BGP-LS-EPE [I-D.ietf-idr-bgpls-segment-routing-epe]describes an extension to BGP Link State (BGP-LS) for the advertisement of BGP Peering Segments along with their BGP peering node and inter-AS link information so that efficient BGP Egress Peer Engineering (EPE) policies and strategies can be computed based on Segment Routing. This document describes extensions to the BGP-LS EPE to enable it to be used for inter-AS Traffic-Engineering (TE) purposes.

Requirements Language

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].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Segment Routing (SR) leverages source routing. A node steers a packet through a controlled set of instructions, called segments, by prepending the packet with an SR header with segment identifiers (SID). A SID can represent any instruction, topological or service-based. SR segments allows to enforce a flow through any topological path or service function while maintaining per-flow state only at the ingress node of the SR domain.

As there is no per-path state in the network, the bandwidth management for the paths is expected to be handled by a controller which has a complete view of:

1. Up-to-date topology of the network

2. Resources, States and Attributes of links and nodes of the network

3. Run-time utilization/availability of resources

In the case of multi-AS networks, the controller learns the topology from all the involved ASes by participating in their IGPs or by BGP-LS [RFC7752] and the inter-AS link information via BGP-LS EPE[I-D.ietf-idr-bgpls-segment-routing-epe] along with extensions defined in this document.Then the controller merges above information sets to consolidated Traffic Engineering Database and computes end-to-end TE paths.

2. Reference Topology

The controller learns TE attributes of all the links, including the inter-AS links and uses the attributes to compute constrained paths. The controller should be able to correlate the inter-AS links for bidirectional connectivity from both ASes.

```
           +----------+
           |Controller|
           +----------+
                |
            |----------|
      +---------+      +------+
      |         |      |      |
      |    H    B------D      G
      |         | +---/| AS 2 |\  +------+
      |         |/     +------+ \ |      |---L/8
      A   AS1   C---+            \|      |
      |         |\\  \  +------+ /| AS 4 |---M/8
      |         | \\  +-E      |/ +------+
      |    X    |  \\   |      K
      |         |   +===F AS 3 |
      +---------+       +------+
```

Figure 1: Reference Diagram

The reference diagram from [I-D.ietf-spring-segment-routing-central-epe] represents multiple Autonomous Systems connected to each other. When the Multiple ASes belong to same operator and are organised into separate domains for operational purposes, it is advantageous to support Traffic-Engineering across the ASes including the inter-AS links. The controller has visibility of all of the ASes by means of IGP topology exported via BGP-LS [RFC7752], or other means. In addition, the inter-AS links and the labels associated with the inter-AS links are exported via [I-D.ietf-idr-bgpls-segment-routing-epe]. The controller needs to correlate the information acquired from all of the ASes, including the inter-AS links in order to get a view of the unified topology so that it can build end-to-end Traffic-Engineered paths.

3. Fast Reroute Label

[I-D.ietf-spring-segment-routing-central-epe] section 3.6 describes mechanisms to provide Fast Reroute (FRR) protection for the EPE Labels. The controller needs to know which links are used for protection so that admission control and failure simulation can be done effectively and appropriate inter-AS links used for path construction.

This document defines a new flag ’F" in the Peering SIDs TLV to indicate a SID as an FRR SID. The protection for any peering SID can be specified using another PeerAdjSID, PeerNodeSID or PeerSetSID to the controller. If the protection is achieved by fallback to local IP lookup, FRR SID SHOULD not be advertised . The link(s) represented by the FRR SID will carry the traffic when there is a failure. These SIDs are included as an FRR SIDs in the peerAdjSID, PeerNodeSID and PeerSetSID advertisements.

```
   0 1 2 3 4 5 6 7
  +-+-+-+-+-+-+-+-+
  |V|L|B|P|F|Rsvd |
  +-+-+-+-+-+-+-+-+
```

Figure 2: Peering SID TLV Flags Format

* F-Flag: FRR Label Flag: If set, the peer SID where the FRR Label appears is using backup links represented by FRR Label.

4. TE Link attributes of PeerNode SID

A Peer Node Segment is a segment describing a peer, including the SID (PeerNode SID) allocated to it. The link descriptors for the Peer Node SID include the addresses used by the BGP session encoded using TLVs as defined in [RFC7752]. In general case, eBGP session could be of multi-hop type, and use multiple underlaying IP interfaces. The IP addesess used by BGP session as local/neigbour are not sufficient to identify this underlaying interfaces. At the same time, the controller needs to know the links associated with the Peer Node SID, to be able derive TE link attributes. This can be achieved by including the interface local and remote addresses in the Link attributes in Peer Node SID NLRI.

PeerAdjSID MUST be advertised for each inter-AS link for the purposes of inter-AS TE. The PeerAdjSID should contain link TE attributes such as bandwidth, admin-group etc. The PeerAdjSID should also contain the local and remote interface ipv4/ipv6 addresses which is used for correlating the links. PeerNodeSID SHOULD contain the additional attribute of link local address which is used by the controller to find corresponding PeerAdjSID and hence the corresponding link TE attributes. PeerAdjSID advertisements MUST contain local and remote interface addresses for the purpose of inter-As TE to be help controller correlate the links. Unnumbered interface is not in the scope of this document.

TLV Code Point	Description	IS-IS TLV /Sub-TLV	Reference (RFC/Section)
TBD	IPv4 Local interface Address	22/6	[RFC5305]/3.2
TBD	IPv6 Local interface Address	22/12	[RFC6119]/4.2

Figure 3: Link Addresses carried as attributes

For Example

5. Backward Compatibility

The extension proposed in this document is backward compatible with procedures described in [I-D.ietf-idr-bgpls-segment-routing-epe] and [I-D.ietf-spring-segment-routing-central-epe]

6. Security Considerations

TBD

7. IANA Considerations

TLV Code Point	Description	IS-IS TLV /Sub-TLV	Reference (RFC/Section)
TBD	IPv4 Local interface Address	22/6	[RFC5305]/3.2
TBD	IPv6 Local interface Address	22/12	[RFC6119]/4.2

Figure 4: Link Attribute TLVs

8. Acknowledgements

Thanks to Julian Lucek and Rafal Jan Szarecki for careful review and suggestions.

9. References

9.1. Normative References

[I-D.ietf-idr-bgpls-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-bgpls-segment-routing-epe-17 (work in progress), October 2018.

[I-D.ietf-spring-segment-routing-central-epe]
Filsfils, C., Previdi, S., Dawra, G., Aries, E., and D. Afanasiev, "Segment Routing Centralized BGP Egress Peer Engineering", draft-ietf-spring-segment-routing-central-epe-10 (work in progress), December 2017.

[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>.

[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>.

9.2. Informative References

[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>.

Authors' Addresses

Shraddha Hegde
Juniper Networks Inc.
Exora Business Park
Bangalore, KA 560103
India

Email: shraddha@juniper.net

Srihari Sangli
Juniper Networks Inc.
Exora Business Park
Bangalore, KA 560103
India

Email: ssangli@juniper.net