Routing area S. Hegde
Internet-Draft K. Arora
Intended status: Standards Track M. Srivastava
Expires: May 3, 2020 S. Ninan
Juniper Networks Inc.
October 31, 2019

Label Switched Path (LSP) Ping/Traceroute for Segment Routing (SR) Egress Peer engineering Segment Identifiers (SIDs) with MPLS Data Planes

draft-hegde-mpls-spring-epe-oam-03

Abstract

Egress Peer Engineering is an application of Segment Routing to solve the problem of egress peer selection. The SR-based BGP-EPE solution allows a centralized (Software Defined Network, SDN)controller to program any egress peer. The EPE solution requires a node to program PeerNodeSID, PeerAdjSID, PeerSetSID as described in [I-D.ietf-spring-segment-routing-central-epe]. This document provides new sub-TLVs for EPE SIDs that would be used in Target stack TLV (Type 1) as defined in [RFC8029] for the EPE SIDs.

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

Egress Peer Engineering (EPE) as defined in [I-D.ietf-spring-segment-routing-central-epe] is an effective mechanism to select the egress peer link based on different criteria. The EPE SIDs provide means to represent egress peer links. Many network deployments have built their networks consisting of multiple Autonomous Systems either for ease of operations or as a result of network mergers and acquisitons. The inter-AS links connecting the two Autonomous Systems could be traffic engineered using EPE-SIDs in this case as well. It is important to be able to validate the control plane to forwarding plane synchronization for these SIDs so that any anomaly can be detected easily by the operator.

This document provides Target FEC stack TLV definitions for EPE SIDs. Other procedures for mpls ping and traceroute as defined in [RFC8287] are applicable for EPE-SIDs as well.

2. FEC Definitions

As described in [RFC8287] sec 5, 3 new type of sub-TLVs for the Target FEC Stack TLV are defined for the Target FEC stack TLV corresponding to each label in the label stack. If a malformed FEC sub-TLV is received, then a return code of 1, "Malformed echo request received" as defined in [RFC8029] SHOULD be sent.

2.1. PeerAdjSID Sub-TLV

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Type = TBD                       |          Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Local AS Number (4  octets)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote As Number (4 octets)                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Local BGP router ID (4 octets)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote BGP Router ID (4 octets)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Local Interface address (4/16 octets)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote Interface address (4/16 octets)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: PeerAdjSID Sub-TLV

Type : TBD

Length : variable based on ipv4/ipv6 interface address

Local AS Number :

4 octet unsigned integer representing the Member ASN inside the Confederation.[RFC5065]. The AS number corresponds to the AS to which PeerAdjSID advertising node belongs to.

Remote AS Number :

4 octet unsigned integer representing the Member ASN inside the Confederation.[RFC5065]. The AS number corresponds to the AS of the remote node for which the PeerAdjSID is advertised.

Local BGP Router ID :

4 octet unsigned integer of the advertising node representing the BGP Identifier as defined in [RFC4271] and [RFC6286].

Remote BGP Router ID :

4 octet unsigned integer of the receiving node representing the BGP Identifier as defined in [RFC4271] and [RFC6286].

Local Interface Address :

In case of PeerAdjSID Local interface address corresponding to the PeerAdjSID should be apecified in this field. For IPv4,this field is 4 octets; for IPv6, this field is 16 octets.

Remote Interface Address :

In case of PeerAdjSID Remote interface address corresponding to the PeerAdjSID should be apecified in this field. For IPv4,this field is 4 octets; for IPv6, this field is 16 octets.

2.2. PeerNodeSID Sub-TLV

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Type = TBD                     |          Length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Local AS Number (4  octets)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote As Number (4 octets)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Local BGP router ID (4 octets)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote BGP Router ID (4 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   No.of IPv4 interface pairs  |   No.of IPv6 interface pairs  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Local Interface address1 (4/16 octets)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote Interface address1 (4/16 octets)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Local Interface address2 (4/16 octets)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              ......                                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: PeerNodeSID Sub-TLV

Type : TBD

Length : variable based on ipv4/ipv6 interface address

Local AS Number :

4 octet unsigned integer representing the Member ASN inside the Confederation.[RFC5065]. The AS number corresponds to the AS to which PeerNodeSID advertising node belongs to.

Remote AS Number :

4 octet unsigned integer representing the Member ASN inside the Confederation.[RFC5065]. The AS number corresponds to the AS of the remote node for which the PeerNodeSID is advertised.

Local BGP Router ID :

4 octet unsigned integer of the advertising node representing the BGP Identifier as defined in [RFC4271] and [RFC6286].

Remote BGP Router ID :

4 octet unsigned integer of the receiving node representing the BGP Identifier as defined in [RFC4271] and [RFC6286].

Number of IPv4 interface pairs:

Total number of IPV4 local and remote interface address pairs.

Number of IPv6 interface pairs:

Total number of IPV6 local and remote interface address pairs.

There can be multiple Layer 3 interfaces on which a peerNodeSID loadbalances the traffic. All such interfaces local/remote address MUST be included in the FEC.

When a PeerNodeSID load-balances over few interfaces with IPv4 only address and few interfaces with IPv6 address then the FEC definition should list all IPv4 address pairs together followed by IPv6 address pairs.

Local Interface Address :

In case of PeerNodeSID, the interface local address ipv4/ipv6 which corresponds to the PeerNodeSID MUST be specified. For IPv4,this field is 4 octets; for IPv6, this field is 16 octets.

Remote Interface Address :

In case of PeerNodeSID, the interface remote address ipv4/ipv6 which corresponds to the PeerNodeSID MUST be specified. For IPv4,this field is 4 octets; for IPv6, this field is 16 octets.

2.3. PeerSetSID Sub-TLV

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Type = TBD                     |          Length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Local AS Number (4  octets)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Local BGP router ID (4 octets)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   No.of elements in set       |          Reserved             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote As Number (4 octets)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote BGP Router ID (4 octets)                  |
++-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-++
|   No.of IPv4 interface pairs  |   No.of IPv6 interface pairs  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Local Interface address1 (4/16 octets)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Remote Interface address1 (4/16 octets)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Local Interface address2 (4/16 octets)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              ......                                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: PeerSetSID Sub-TLV

Type : TBD

Length : variable based on ipv4/ipv6 interface address and number of elements in the set

Local AS Number :

4 octet unsigned integer representing the Member ASN inside the Confederation.[RFC5065]. The AS number corresponds to the AS to which PeerSetSID advertising node belongs to.

Remote AS Number :

4 octet unsigned integer representing the Member ASN inside the Confederation.[RFC5065]. The AS number corresponds to the AS of the remote node for which the PeerSetSID is advertised.

Advertising BGP Router ID :

4 octet unsigned integer of the advertising node representing the BGP Identifier as defined in [RFC4271] and [RFC6286].

Receiving BGP Router ID :

4 octet unsigned integer of the receiving node representing the BGP Identifier as defined in [RFC4271] and [RFC6286].

No.of elements in set:

Number of remote ASes, the set SID load-balances on.

PeerSetSID may be associated with a number of PeerNodeSIDs and PeerAdjSIDs. Link address details of all these SIDs should be included in the peerSetSID FEC so that the data-plane can be correctly verified on the remote node.

Number of IPv4 interface pairs:

Total number of IPV4 local and remote interface address pairs.

Number of IPv6 interface pairs:

Total number of IPV6 local and remote interface address pairs.

There can be multiple Layer 3 interfaces on which a peerNodeSID loadbalances the traffic. All such interfaces local/remote address MUST be included in the FEC.

When a PeerSetSID load-balances over few interfaces with IPv4 only address and few interfaces with IPv6 address then the Link address TLV should list all IPv4 address pairs together followed by IPv6 address pairs.

Local Interface Address :

In case of PeerNodeSID/PeerAdjSID, the interface local address ipv4/ipv6 which corresponds to the PeerNodeSID/PeerAdjSID MUST be specified. For IPv4,this field is 4 octets; for IPv6, this field is 16 octets.

Remote Interface Address :

In case of PeerNodeSID/PeerAdjSID, the interface remote address ipv4/ipv6 which corresponds to the PeerNodeSID/PeerAdjSID MUST be specified. For IPv4,this field is 4 octets; for IPv6, this field is 16 octets.

3. Security Considerations

The EPE SIDs are advertised for egress links for Egress Peer Engineering purposes or for inter-As links between co-operating ASes. When co-operating domains are involved, they can allow the packets arriving on trusted interfaces to reach the control plane and get processed. When EPE SIDs which are created for egress TE links where the neighbor AS is an independent entity, it may not allow packets arriving from external world to reach the control plane. In such deployments mpls OAM packets will be dropped by the neighboring AS.

4. IANA Considerations

New Target FEC stack sub-TLV from the "sub-TLVs for TLV types 1,16 and 21" subregistry of the "Multi-Protocol Label switching (MPLs) Label Switched Paths (LSPs) Ping parameters" registry

PeerAdjSID segment ID Sub-TLV : TBD

PeerNode segment ID Sub-TLV : TBD

PeerSetSID segment ID Sub-TLV : TBD

5. Acknowledgments

6. References

6.1. Normative References

[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.

[RFC8287] Kumar, N., Ed., Pignataro, C., Ed., Swallow, G., Akiya, N., Kini, S., and M. Chen, "Label Switched Path (LSP) Ping/Traceroute for Segment Routing (SR) IGP-Prefix and IGP-Adjacency Segment Identifiers (SIDs) with MPLS Data Planes", RFC 8287, DOI 10.17487/RFC8287, December 2017, <https://www.rfc-editor.org/info/rfc8287>.

6.2. Informative 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>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

Authors' Addresses

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

Email: shraddha@juniper.net

Kapil Arora
Juniper Networks Inc.

Email: kapilaro@juniper.net

Mukul Srivastava
Juniper Networks Inc.

Email: msri@juniper.net

Samson Ninan
Juniper Networks Inc.

Email: samsonn@juniper.net