Segment Routing BGP Egress Peer Engineering over Layer 2 Bundle

draft-lin-idr-sr-epe-over-l2bundle-03

Changwang Lin (New H3C Technologies)

Zhenqiang Li (China Mobile)

Ran Pang (China Unicom)

Mengxiao Chen (New H3C Technologies)

IETF-118, November 2023

Background

There are deployments where the Layer 3 interface on which a BGP peer session is established is a Layer 2 interface bundle (L2 Bundle).

The operator of AS1 wishes to apply a BGP-EPE policy to steer the time-sensitive traffic from AS1 to AS2 via member link 1 of the Layer 2 bundle.

Time-Sensitive Traffic

Member Link 1: Low Delay

Member Link 2

Member Link 3

R1

R2

AS 1

AS 2

EBGP Session
over L2 Bundle
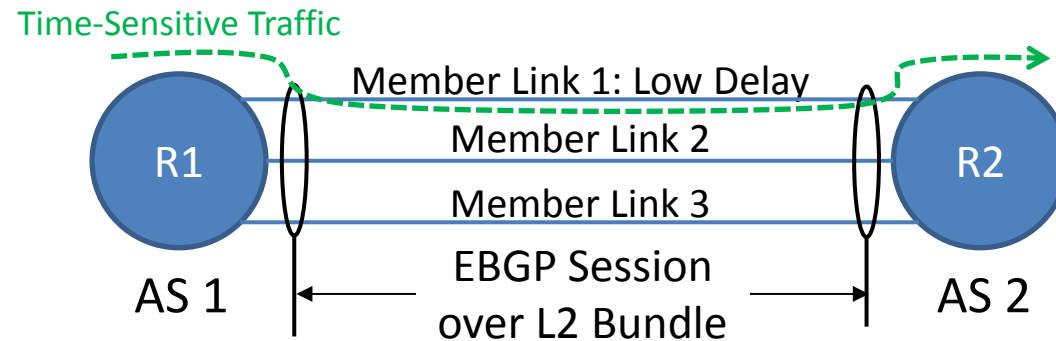
BGP Peering SIDs need to be allocated to individual bundle member links, and advertisement of such BGP Peering SIDs in BGP-LS is also required.

Why need L2 Bundle Member PeerAdj SID TLV ?

RFC9085

There is only the definition of IGP Adjacency SID TLV and LAN Adjacency SID TLV under the L2 Bundle Member Attributes TLV in RFC 9085, but the definition of **BGP L2 Bundle Adj SID is missing.**

RFC9086

RFC 9086 only defines the following three types:

- Peer Node Segment (PeerNode SID): an instruction to route to a specific peer node.
- Peer Adjacency Segment (PeerAdj SID): an instruction to route over a specific local interface to a specific peer node.
- Peer Set Segment (PeerSet SID): an instruction to load-balance to a set of specific peer nodes.

TLV Code Point	Description
1101	PeerNode SID
1102	PeerAdj SID
1103	PeerSet SID

Therefore, it is necessary to add the behavior definition for Peer Adjacency Segment for L2 Bundle Member Link, along with the corresponding TLV definition.

Peer Adjacency Segment for L2 Bundle Member Link

Semantics
- SR operation: NEXT.
- Next-Hop: forwarding across the bundle member link, which the segment is associated with, to the peer connected through the parent L3 interface.

Advertising in BGP-LS ← *BGP peering segments are generally advertised in BGP-LS from a BGP node along with its peering topology information, in order to enable computation of BGP-EPE policies and strategies.*
- BGP-LS Link NLRI (parent L3 link)
- Attributes:
  - ① MPLS-SR data plane:
    - PeerAdj SID TLV (Label for parent L3 link) ← RFC 9086
    - L2 Bundle Member Attribute TLV (member link 1) ← RFC 9085
      - **L2 Bundle Member PeerAdj SID TLV (Label for member link 1)** ← **this draft**
    - L2 Bundle Member Attribute TLV (member link 2)
      - L2 Bundle Member PeerAdj SID TLV (Label for member link 2)
    - …
  - ② SRv6 data plane:
    - SRv6 End.X SID TLV (SID for parent L3 link)
    - L2 Bundle Member Attribute TLV (member link 1) ← RFC 9085
      - SRv6 End.X SID TLV (SRv6 SID for member link 1) ← draft-ietf-idr-bgpls-srv6-ext
    - L2 Bundle Member Attribute TLV (member link 2)
      - SRv6 End.X SID TLV (SRv6 SID for member link 2)
    - …

L2 Bundle Member PeerAdj SID 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              |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Flags      |    Weight     |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    SID/Label/Index (variable)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- Type: TBD.
- Length: variable.
- Flags: one octet of flags. Same as the Flags field of BGP Peering SIDs TLV, as described in Section 5 of [RFC9086].
- Weight: 1 octet. The value represents the weight of the SID for the purpose of load balancing.
- SID/Index/Label. It contains either:
  - A 3-octet local label where the 20 rightmost bits are used for encoding the label value. In this case, the V- and L-Flags MUST be SET.
  - A 4-octet index defining the offset in the Segment Routing Global Block (SRGB) [RFC8402] advertised by this router. In this case, the SRGB MUST be advertised using the extensions defined in [RFC9085].

Next Steps

- Any questions or comments are Welcomed.

Thanks