

Path MTU in BGP/BGP-LS

draft-zhu-idr-bgp-ls-path-mtu-01
draft-li-idr-sr-policy-path-mtu-02

Yongqing Zhu/Cheng Li/Zhibo Hu/Gang Yan/Junda Yao/Zhenbin Li

IETF#105

Motivation

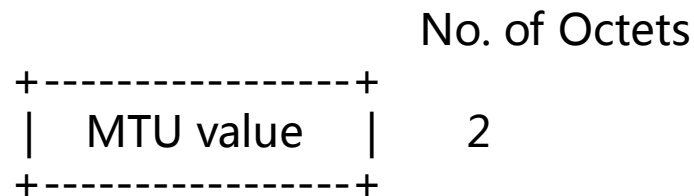
- In traditional MPLS, the Path MTU can be signaled via signaling protocols like RSVP-TE[3209] and LDP[RFC3988]
- However, there is no additional signaling to establish Segment Routing(SR) paths, so the SR tunnel cannot currently support the negotiation mechanism of the Path MTU.
- When SIDs (Label or IPv6 address) are inserted into a packet, the packet will be dropped (In IPv6) or fragmented in forwarding since the packet size may exceed the Path MTU.
 - Operator A: When using leased line over multi-domains, MTU should be learned to avoid dropping packets.
- Thus, we proposed two drafts:
 - **draft-zhu-idr-bgp-ls-path-mtu-01**: specify the extension to BGP Link State (BGP-LS) to carry maximum transmission unit (MTU) messages of link.
 - **draft-li-idr-sr-policy-path-mtu-02**: defines extensions to BGP to distribute path MTU information within SR policies.

draft-zhu-idr-bgp-ls-path-mtu-01

[RFC7752] defines the TLVs that map link-state information to BGP-LS NLRI and the BGP-LS attribute. This document, a new sub TLV is added to the Link Attribute TLV, called PMTU TLV.

Since [RFC 6326] (Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS) has defined a Sub-TLV to advertise the MTU of a Link, this document reuses the TLV. The format of the sub-TLV is as shown below.

- TYPE - TBD
- LENGTH - Total length of the value field, it should be 3
- Reserved Byte
- VALUE - 2-byte MTU value of the link



Whenever there is a change in MTU value represented by Link Attribute TLV, BGP-LS should re-originate the respective TLV with the new MTU value. Then, the controller can calculate the Path MTU.

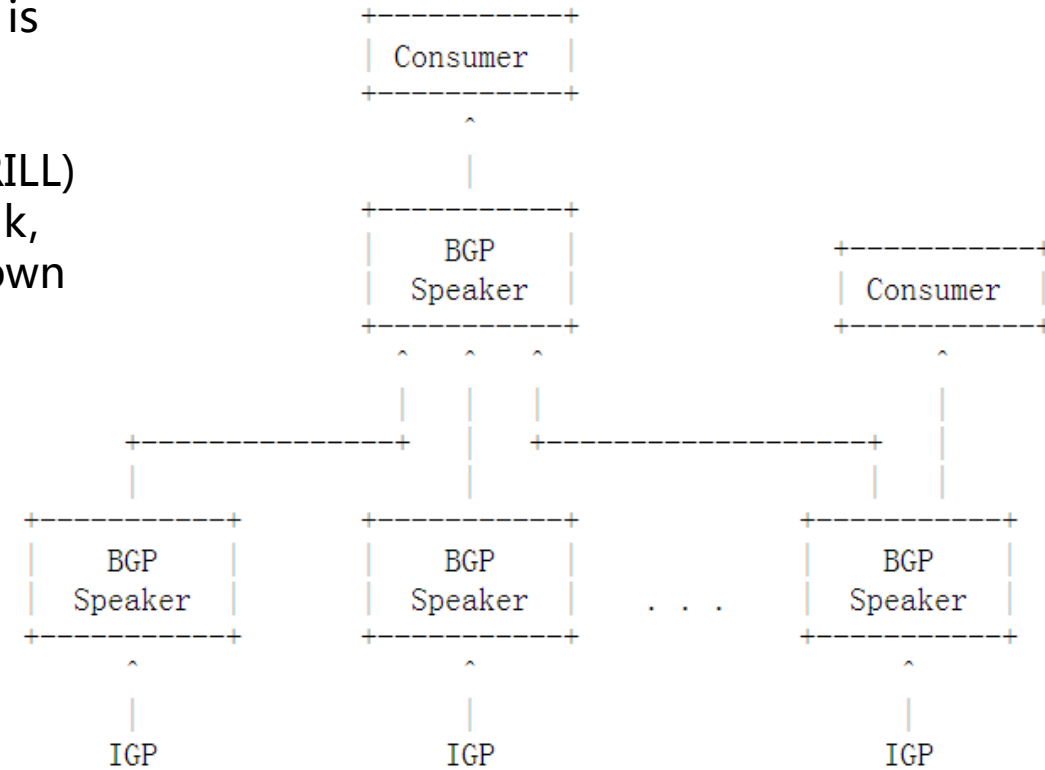


Figure 1: Collection of Link-State and TE Information

draft-li-idr-sr-policy-path-mtu-02

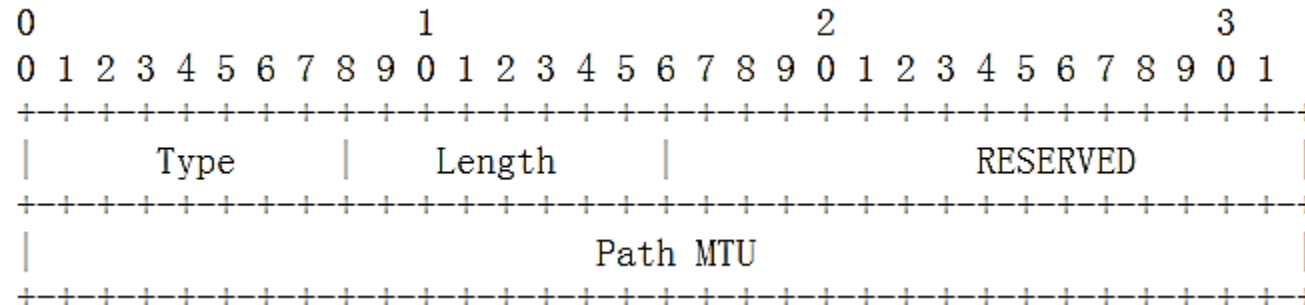
- [I-D.ietf-idr-segment-routing-te-policy](#) defines the SR Policy structure as per Figure 1.
- Each SR path has its path MTU, so a PMTU TLV should be inserted into the Segment List sub-TLV

```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
  Tunnel Encaps Attribute (23)
    Tunnel Type: SR Policy
    Binding SID
    Preference
    Priority
    Policy Name
    Explicit NULL Label Policy (ENLP)
    Segment List
      Weight
      Path MTU
      Segment
      Segment
      ...
    ...
```

Figure 1. Path MTU TLV in SR policy

draft-li-idr-sr-policy-path-mtu-02

- Type: to be assigned by IANA.
- Length: the total length of the value field not including Type and Length fields.
- Reserved: 16 bits reserved and MUST be set to 0 on transmission and MUST be ignored on receipt.
- Path MTU: 4 bytes value of path MTU in octets. The value can be calculated by a central controller or other devices based on the information that learned via IGP or BGP-LS or other means.



Whenever the path MTU of a physical or logical interface is changed, a new SR policy with new path MTU information should be updated accordingly by BGP.

Next Step

- Comments are welcome!
- Sync up the Path MTU TLV format in SR policy with the one in BGP-LS?

Any Questions?