

Interdomain Routing Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 30, 2020

C. Li
Huawei Technologies
Y. Zhu
China Telecom
A. Sawaf
Saudi Telecom Company
Z. Li
Huawei Technologies
April 28, 2020

Segment Routing Path MTU in BGP
draft-ietf-idr-sr-policy-path-mtu-00

Abstract

Segment Routing is a source routing paradigm that explicitly indicates the forwarding path for packets at the ingress node. An SR policy is a set of candidate SR paths consisting of one or more segment lists with necessary path attributes. However, the path maximum transmission unit (MTU) information for SR path is not available in the SR policy since the SR does not require signaling. This document defines extensions to BGP to distribute path MTU information within SR policies.

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 October 30, 2020.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Terminology	3
2.1.	Requirements Language	3
3.	SR Policy for Path MTU	3
3.1.	SR Path MTU Sub-TLV	4
4.	Operations	5
5.	IANA Considerations	6
6.	Security Considerations	6
7.	Contributors	6
8.	Acknowledgements	6
9.	References	6
9.1.	Normative References	6
9.2.	Informative References	7
	Authors' Addresses	8

[1.](#) Introduction

Segment routing (SR) [[RFC8402](#)] is a source routing paradigm that explicitly indicates the forwarding path for packets at the ingress node. The ingress node steers packets into a specific path according to the Segment Routing Policy (SR Policy) as defined in [[I-D.ietf-spring-segment-routing-policy](#)]. In order to distribute SR policies to the headend, [[I-D.ietf-idr-segment-routing-te-policy](#)] specifies a mechanism by using BGP.

The maximum transmission unit (MTU) is the largest size packet or frame, in bytes, that can be sent in a network. An MTU that is too large might cause retransmissions. Too small an MTU might cause the router to send and handle relatively more header overhead and acknowledgments.

When an LSP is created across a set of links with different MTU sizes, the ingress router needs to know what the smallest MTU is on the LSP path. If this MTU is larger than the MTU of one of the intermediate links, traffic might be dropped, because MPLS packets cannot be fragmented. Also, the ingress router may not be aware of

this type of traffic loss, because the control plane for the LSP would still function normally. [RFC3209] specify the mechanism of MTU signaling in RSVP. Likewise, SRv6 packets will be dropped if the packet size is larger than path MTU, since IPv6 packet can not be fragmented on transmission [RFC8200] .

The host may discover the PMTU by Path MTU Discovery (PMTUD) [RFC8201] or other mechanisms. But the ingress still needs to examine the packet size for dropping too large packets to avoid malicious traffic or error traffic. Also, the packet size may exceeds the PMTU because of the new encapsulation of SR-MPLS or SRv6 packet at the ingress.

In order to check whether the Packet size exceeds the PMTU or not, the ingress node needs to know the Path MTU associated to the forwarding path. However, the path maximum transmission unit (MTU) information for SR path is not available since the SR does not require signaling.

This document defines extensions to BGP to distribute path MTU information within SR policies. The Link MTU information can be obtained via BGP-LS [I-D.zhu-idr-bgp-ls-path-mtu] or some other means. With the Link MTU, the controller can compute the PMTU and convey the information via the BGP SR policy.

2. Terminology

This memo makes use of the terms defined in [RFC8402] and [RFC3209].

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. SR Policy for Path MTU

As defined in [I-D.ietf-idr-segment-routing-te-policy] , the SR policy encoding structure is as follows:

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
      Segment
      Segment
      ...
    ...
```

As introduced in [Section 1](#), each SR path has it's path MTU. SR policy with SR path MTU information is expressed as below:

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
      ...
    ...
```

[3.1.](#) SR Path MTU Sub-TLV

An SR Path MTU sub-TLV is an Optional sub-TLV. When it appears, it must appear only once at most within a Segment List sub-TLV. If multiple Path MTU sub-TLVs appear within a Segment List sub-TLV, the first one will be processed, and the rest will be ignored. An SR Path MTU sub-TLV is associated with an SR path specified by a segment list sub-TLV or path segment as defined in [\[I-D.ietf-spring-mpls-path-segment\]](#) and [\[I-D.li-spring-srv6-path-segment\]](#). It has the following format:

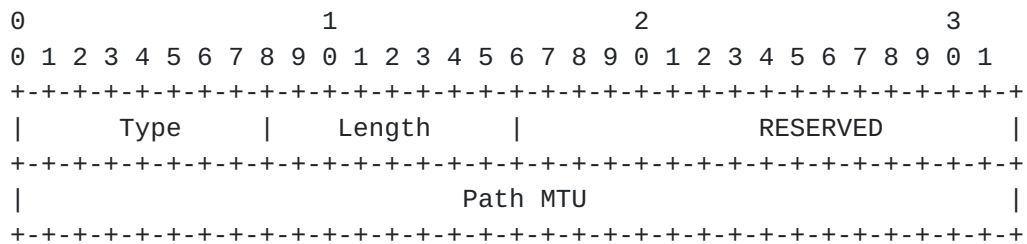


Figure 1. Path MTU sub-TLV

Where:

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 of 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.

4. Operations

The document does not bring new operation beyond the description of operations defined in [[I-D.ietf-idr-segment-routing-te-policy](#)]. The existing operations defined in [[I-D.ietf-idr-segment-routing-te-policy](#)] can apply to this document directly.

Typically but not limit to, the SR policies carrying path MTU information are configured by a controller.

After configuration, the SR policies carrying path MTU information will be advertised by BGP update messages. The operation of advertisement is the same as defined in [[I-D.ietf-idr-segment-routing-te-policy](#)], as well as the reception.

The consumer of the SR policies is not the BGP process. The operation of sending information to consumers is out of scope of this document.

5. IANA Considerations

This document defines a new Sub-TLV in registries "SR Policy List Sub- TLVs" [[I-D.ietf-idr-segment-routing-te-policy](#)]:

Value	Description	Reference

TBA	Path MTU sub-TLV	This document

6. Security Considerations

TBA

7. Contributors

Jun Qiu

Huawei Technologies

China

Email: qiujun8@huawei.com

8. Acknowledgements

Authors would like to thank Ketan Talaulikar, Aijun Wang, Weiqiang Cheng, Huanan Chen, Chongfeng Xie, Stefano Previdi, Taishan Tang, Keqiang Guo, Chen Zhang, Susan Hares, Weiguo Hao, Gong Xia, Bing Yang, Linda Dunbar, Shunwan Zhuang, Huaimo Chen, Mach Chen, Jingrui Xie, Zhibo Hu, Jimmy Dong and Jianwei Mao for their professional comments and help.

9. References

9.1. Normative References

[[I-D.ietf-idr-segment-routing-te-policy](#)]
Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., Rosen, E., Jain, D., and S. Lin, "Advertising Segment Routing Policies in BGP", [draft-ietf-idr-segment-routing-te-policy-08](#) (work in progress), November 2019.

- [I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", [draft-ietf-spring-segment-routing-policy-06](#) (work in progress), December 2019.
- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [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>>.

9.2. Informative References

- [I-D.ietf-spring-mpls-path-segment]
Cheng, W., Li, H., Chen, M., Gandhi, R., and R. Zigler, "Path Segment in MPLS Based Segment Routing Network", [draft-ietf-spring-mpls-path-segment-02](#) (work in progress), February 2020.
- [I-D.li-spring-srv6-path-segment]
Li, C., Cheng, W., Chen, M., Dhody, D., and R. Gandhi, "Path Segment for SRv6 (Segment Routing in IPv6)", [draft-li-spring-srv6-path-segment-05](#) (work in progress), March 2020.
- [I-D.zhu-idr-bgp-ls-path-mtu]
Zhu, Y., Hu, Z., Yan, G., and J. Yao, "BGP-LS Extensions for Advertising Path MTU", [draft-zhu-idr-bgp-ls-path-mtu-02](#) (work in progress), January 2020.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), DOI 10.17487/RFC3209, December 2001, <<https://www.rfc-editor.org/info/rfc3209>>.
- [RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, [RFC 8200](#), DOI 10.17487/RFC8200, July 2017, <<https://www.rfc-editor.org/info/rfc8200>>.

[RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed.,
"Path MTU Discovery for IP version 6", STD 87, [RFC 8201](#),
DOI 10.17487/RFC8201, July 2017,
<<https://www.rfc-editor.org/info/rfc8201>>.

Authors' Addresses

Cheng Li
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: chengli13@huawei.com

YongQing Zhu
China Telecom
109, West Zhongshan Road, Tianhe District.
Guangzhou
China

Email: zhuyq.gd@chinatelecom.cn

Ahmed El Sawaf
Saudi Telecom Company
Riyadh
Saudi Arabia

Email: aelsawaf.c@stc.com.sa

Zhenbin Li
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: lizhenbin@huawei.com

