Network Working Group Z. Hu
Internet-Draft Huawei
Intended status: Informational Y. Zhu
Expires: January 1, 2019 China Telecom
Z. Li
L. Dai
Huawei
June 30, 2018

IS-IS Extensions for Path MTU
draft-hu-lsr-isis-path-mtu-00

Abstract

Segment routing (SR) leverages the source routing mechanism. It allows for a flexible definition of end-to-end paths with IGP topologies by encoding paths as sequences of topological sub-paths which is called segments. These segments are advertised by the link-state routing protocols (IS-IS and OSPF). Unlike the MPLS, SR does not have the specific path construction signaling so that it cannot support the Path MTU. This draft provides the necessary IS-IS extensions about the Path MTU that need to be used on SR.

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 the source routing mechanism. SR allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of toplogical sub-paths which is called segments. These segments are advertised by the link-state routing protocols (IS-SI and OSPF). The SR architecture as well as the routing policy is proposed in [I-D.ietf-spring-segment-routing] and [I-D.filsfils-spring-segment-routing-policy]. Two types of segments are defined, Prefix segments and Adjacency segments. Prefix segments represent an ecmp-aware shortest-path to a prefix, as per the state of the IGP topology. Adjacency segments represent a hop over a specific adjacency between two nodes in the IGP. A prefix segment is typically a multi-hop path while an adjacency segment, in most of the cases, is a one-hop path. SR can compute the paths from end to end and without requiring any LDP or RSVP-TE signaling. SR supports per-flow explict routing while just maintaining per-flow state only at the source node.

SR architecture supports the distributed scenario and the centralized scenario. In the distributed scenario, the segments are allocated and signaled by IGP or BGP and a node needs to compute the source-routed policy. Some necessary IS-IS extensions for SR are proposed in [I-D.ietf-isis-segment-routing-extensions]. In a centralized scenario, the SR controller decides which nodes need to steer which packets on which source-routed policies. However, in both conditions, the MTU is not included in the SR policy. As the SR may push more MPLS labels or SRv6 SIDs in the packet header, the packets are larger than the minimum MTU in the path compared to the traditional MPLS forwarding process. Unfortunately the paths do not provide the path MTU informaiton so that the path can not assure the packet size is less than the path MTU, which is the minimum link MTU of all the links in a path between a source node and a destination node. The definition of the path MTU is discussed in RFC1981 [RFC1981].

This draft describes the necessary IS-IS extensions about the path MTU that need to be used on SR. A new TLV is introduced into the IS-IS protocol. With the IGP flooding process in the distributed scenario or transmission to the controller by BGP, the ingress nodes or the controllers compute the Path MTU for the SR policy.

2. Terminology

router: a node that forwards IP packets not explicitly addressed to itself.

interface: a node’s attachment to a link.

Segment: an instruction a node executes on the incoming packet. For example, froward packet according to shortest path to destination or a specific interface, etc..

SR Policy: an ordered list of segments.

MTU: Maximum Transmission Unit, the size in bytes of the largest IP packet, including the IP header and payload, that can be transmitted on a link or path.

link MTU: the maximum transimission unit, i.e., maximum packet size in octets, that can be conveyed in one piece over a link.

path: the set of links traversed by a packet between a source node and a destination node

Path MTU: the minimum link MTU of all the links in a path between a source node and a destination node.

3. Extendsion of IS-IS

This document describes an IS-IS extension to flood the router interface MTU to each node with the IGP domain. Then the controller or the original node collects all the link MTUs from the routers. After the SR path is calculated, packet may be lost if the packet size is larger than the minimum MTU along the path. So the original node can compute the minimum link MTU of all the links in the path. The source node can limit the packet size less than the path MTU.

3.1. Protocol Extension

A new TLV called link MTU TLV is defined to be included in the Router Information LSP. The LSP transmitted by an interface in a router MUST include the TLV. Each such TLV is encoded as shown in Figure 1.

Type: MTU, 1 byte

Length: # of octets in the value field (1bytes)

Value: the value is the MTU size of a link.

```
 0                0
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Type = MTU   |
+-+-+-+-+-+-+-+-+
|  Length       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     MTU size                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1

The use and meaning of these fields are as follows:

Type - A single octet encoding the TLV type. Here the type is 1 octet.

Length - One octet encoding the length in octet of the TLV. This field identifies the length of the value part.

MTU size - This field identifies the size of the router interfaces. Two octets encoding the MTU size of the TLV.

This document defines a single MTU TLV, the codepoints need to be determined by the IANA.

4. Acknowledgements

TBD.

5. IANA Considerations

This document requests that IANA allocate from the IS-IS TLV Codepoints Registry a new TLV.

6. Security Considerations

This extension to IS-IS does not change the underlying security issues iherent in the existing IGP.

7. References

[I-D.filsfils-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Hegde, S., daniel.voyer@bell.ca, d., Lin, S., bogdanov@google.com, b., Krol, P., Horneffer, M., Steinberg, D., Decraene, B., Litkowski, S., Mattes, P., Ali, Z., Talaulikar, K., Liste, J., Clad, F., and K. Raza, "Segment Routing Policy Architecture", draft-filsfils-spring-segment-routing-policy-06 (work in progress), May 2018.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-18 (work in progress), June 2018.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.

[RFC1981] McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, DOI 10.17487/RFC1981, August 1996, <https://www.rfc-editor.org/info/rfc1981>.

[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>.

Authors' Addresses

Zhibo Hu
Huawei
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: huzhibo@huawei.com

Yongqing Zhu
China Telecom
109, West Zhongshan Rd.
Guangzhou 510000
China

Email: zhuyq@gsta.com

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

Email: lizhenbin@huawei.com

Longfei Dai
Huawei
Huawei Bld., No. 156 Beiqing Rd.
Beijing 100095
China

Email: larry.dai@huawei.com