Workgroup: SPRING Working Group
Internet-Draft:
draft-dong-spring-srv6-inter-layer-
programming-04
Published: 24 October 2022
Intended Status: Standards Track
Expires: 27 April 2023
Authors: J. Dong, Huawei Technologies
L. Han, China Mobile
Z. Du, China Mobile
M. Wang, China Mobile

SRv6 for Inter-Layer Network Programming

Abstract

This document defines a new SRv6 function which can be used for SRv6 based inter-layer network programming. It is a variant of the SRv6 End.X behavior which is called "End.XU". Instead of pointing to an interface with layer-3 adjacency, the End.XU behavior points to an underlay interface which connects to a remote layer-3 node via underlying links or connections that may be invisible in the L3 topology.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

In many network scenarios, operator owns a multi-layered network. In layer-3, the technology has converged to IP, while there can be different technologies in the layer-2 and layer-1. In such networks, the cross-layer planning and optimization is considered more efficient than independent planning and operation of the layer-3 and the underlying networks in terms of resource utilization and SLA assurance, but are also considered more complicated. Thus a mechanism for flexible inter-layer network integration is desired.

Segment Routing over IPv6 (SRv6) [RFC8986] enables a network operator or an application to specify a packet processing program by encoding a sequence of instructions in the IPv6 packet header. Currently SRv6 does not consider about the network layers under the IP layer. However, with the capability of SRv6 network programming, it is possible to achieve seamless integration between IP (layer-3) and the underlying (layer-2 and layer-1) networks.

Following the SRv6 network programming concept, a new SRv6 function is defined for sending packet through an underlay interface, which connects to underlay links or connections between two layer-3 nodes. The underlay link or connection may be realized using either a ethernet link, a Metro Transport Network (MTN) path[ITU-T_G.8310], an ODUk or a DWDM connection. Such a SRv6 behavior can be considered as a variant of the SRv6 END.X behavior as defined in [RFC8986]. Instead of pointing to an interface with layer-3 adjacency, this new End.XU behavior points to an underlay interface which connects to a remote layer-3 node via an underlying link or connection that may be invisible in the L3 topology. The SRv6 End.XU SIDs can be used together with other types of SRv6 SIDs to build SRv6 SID lists for inter-layer network programming.

This document first describes the typical use cases of SRv6 inter-layer network programming, then new SRv6 End.XU behavior for inter-layer network programming is defined. The application of SRv6 End.XU in typical scenarios is also illustrated with examples.

1.1. 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.

2. Use Cases of SRv6 Inter-Layer Programming

2.1. IP and Optical Inter-layer Programming

In many network scenarios, the underlay of the IP network is an optical network. The IP network and optical network are usually managed separately, the optical network works as an underlay which is invisible to the IP network. In some cases, the optical path resource and the IP path resource may not be one-to-one mapping, the redundant optical paths may not be fully used by the IP layer. In some other cases, there may be optical paths between non-adjacent IP nodes thus they are not visible in the L3 IP topology, and thus they can not be used to carry IP traffic.

2.2. IP and MTN Inter-layer Programming

The architecture of Metro Transport Network (MTN) is defined in [ITU-T_G.8310]. In an MTN based network, network nodes can support two forwarding modes: per-hop IP packet forwarding and the MTN Path (MTNP) layer cross-connect. An MTN path is a multi-hop transport path which may be established between any two nodes in the MTN network, and the intermediate nodes of the MTN path will forward the traffic solely based on the pre-established MTN cross-connect without IP layer lookup. Thus an MTN path is an underlay connection between two remote MTN nodes. Although in some cases it is possible to set up a layer-3 adjacency between the two endpoints of the MTN path, it will make the provisioning of MTN path complicated.

Moreover, in some cases the two endpoints may reside in different IGP areas or ASes, which makes a layer-3 adjacency between them more challenging. Since the MTN paths are not visible to the IP layer topology, it is difficult to compute and establish a inter-layer path which consists of the layer-3 network segments with the MTN paths.

2.3. Steering Traffic to L2 bundle Member Link

In some network scenarios, L2 bundles which consists of a group of L2 member links are created to reduce the operational overhead of maintaining multiple parallel L3 links. On the L2 bundle, traffic are usually load balanced among all the member links. While for the purpose of traffic engineering, one specific L2 member link may be selected for specific service traffic.

In section 4.2 of [RFC8986], it is described that for an outgoing bundle interface, End.X SIDs might be allocated for both the bundle itself and for each of its member link. However, there is difference between the bundle interface and its layer-2 member links, as they are at different network layers, and there is no L3 adjacency on the layer-2 member link.

3. SRv6 END.XU behavior

This section defines a new SRv6 behavior for the underlay cross-connect.

The "Endpoint with Underlay cross-connect" behavior ("End.XU" for short) is a variant of the End.X behavior defined in [RFC8986]. Its main use is for inter-layer network programming and traffic engineering.

Any SID instance of this behavior is associated with an underlay interface, which connects to one or more underlay links or connections.

When N receives a packet destined to S and S is a local End.XU SID, N does the following:

```
S01. When an SRH is processed {
S02.   If (Segments Left == 0) {
S03.      Stop processing the SRH, and proceed to process the next
             header in the packet, whose type is identified by
             the Next Header field in the routing header.
S04.   }
S05.   If (IPv6 Hop Limit <= 1) {
S06.      Send an ICMP Time Exceeded message to the Source Address
             with Code 0 (Hop limit exceeded in transit),
             interrupt packet processing, and discard the packet.
S07.   }
S08.   max_LE = (Hdr Ext Len / 2) - 1
S09.   If ((Last Entry > max_LE) or (Segments Left > Last Entry+1)) {
S10.      Send an ICMP Parameter Problem to the Source Address
             with Code 0 (Erroneous header field encountered)
             and Pointer set to the Segments Left field,
             interrupt packet processing, and discard the packet.

S11.   }
S12.   Decrement IPv6 Hop Limit by 1
S13.   Decrement Segments Left by 1
S14.   Update IPv6 DA with Segment List[Segments Left]
S15.   Forward the packet through the underlay interface associated
          with SID S
S16. }
```

Note that the underlay interface and the associated connection in step 15 SHOULD be established before the associated End.XU SID is announced into the network.

End.XU SIDs MAY be announced using IGP or BGP-LS in a similar way to the announcement of End.X SIDs, while they need to be distinguished from the End.X SID by both the network nodes and the network controller. The detailed protocol extension will be described in a separate document. Then the network controller or a headend node could use the End.XU SIDs together with other types of SRv6 SIDs to build SID lists for inter-layer network paths.

4. Application of SRv6 End.XU

4.1. IP and Optical Integration

Assuming that an operator owns both the IP and optical network, and the operator needs to deploy E2E service across IP and optical network, with traditional approaches the planning and service provisioning would be complex and time consuming due of the manual synergy needed between the operator's IP team and optical team. With the introduction of SRv6 and the End.XU behavior, one simplified approach for IP and optical integration is to build a SID list that integrates the path in both the IP layer and the optical layer.

As the optical layer is not packet based, source routing mechanism can not be directly used in the optical network. However, the abstracted optical paths (e.g., with ODUk or DWDM) could be exposed to the control system of the IP network using the SRv6 End.XU SIDs, and some of the attributes of the optical paths may also be provided. Based on this information, IP-optical inter-layer paths can be programmed to meet some specific service requirements, such as low latency.

```
          -----          -----          -----
         |  P1 |--------|  P2 |--------|  P3 |
          -----          -----          -----
         /  |.             |.             |.  \
 -----  /   | .            | .            | .  \ -----
|  P7 |     |  .           |  .           |  .  |  P8 |
 ----- \    |   .          |   .          |   ./ -----
        \   |    .         |    .         |  / .
          -----   .      -----   .      -----   .
         |  P4 |-------|  P5 |--------|  P6 |   .
          -----    .     -----     .    -----     .
            .      .       .       .      .       .
            .    =====      .     =====    .     =====
             .  |  O1 |----------|  O2 |--------|  O3 |
              .  =====        .   =====      .   =====
               .    |          .    |         .    |
                .   |           .   |          .   |
                 .  |            .  |           .  |
                  . |             . |            . |
                   .|              .|             .|
                 =====            =====          =====
                |  O4 |----------|  O5 |--------|  O6 |
                 =====            =====          =====
```

Figure 1. IP and Optical Layered Network Topology

In Figure 1, P1 to P8 are IP nodes, and O1 to O6 are optical nodes. Assume the operator needs to deploy a low latency path between P7 and P8. With normal segment routing, an IP layer path with the segment list {P7, P1, P2, P3, P8} can be used. But if an optical path from O1 to O3 exists, and the End.XU SID defined in this document is used to announce this optical path as an underlay connection with specific attributes into the IP network, the headend node or the controller in IP layer can program an inter-layer path along {P7, P1, End.XU (O1, O2, O3), P3, P8} which may provide lower latency.

The optical path between O1 and O3 may be created in advance or as a result of the request from the IP layer. The creation should be done by the optical network controller (not shown in the Figure). The details of the process are out of scope of this document, and may refer to [I-D.ietf-teas-actn-poi-applicability].

There is also another case of IP and Optical integration. Assume there are two optical paths between P1 and P2. One is {P1, O1, O2, P2} , and the other is {P1, O1, O4, O5, O2, P2}. Two separate End.XU SIDs are allocated for these two underlay connections separately. One is End.XU P1::C2 for the underlay path {P1, O1, O2, P2}, and the other is End.XU P1::C45 for the path {P1, O1, O4, O5, O2, P2}. The headend P7 or the IP network controller will be informed about these two SRv6 End.XU SIDs and the associated path attributes, so that the headend or the controller can program different end-to-end inter-layer paths using SID lists with different End.XU SIDs for services with different SLA requirements.

4.2. MTN Networks

Assuming that an operator owns both an MTN network domain and an IP network domain. In the MTN network, each MTN node has both the layer-3 functionality and the MTN Path layer functionality. In layer-3, all the MTN nodes are in a layer-3 network topology, which connects to the IP network domain. In the MTN Path Layer, a set of MTN paths are provisioned between the selected pairs of MTN nodes. In the MTN network, different types of services may be carried using either a layer-3 path, or an MTN path, or an inter-layer path comprising of both the layer-3 links and the MTN path as different segments. In addition, For some type of services, end-to-end paths across the IP domain and the MTN domain are needed, which is comprised of both the layer-3 paths and the MTN path as different segments.

```
.......................................... ...........................
.                                        . .                         .
.          +----+     +----+     +----+  . . +----+     +----+       .
.          | M1 |-----| M2 |-----| M3 |------| P1 |-----| P2 |       .
.          +----+     +----+     +----+  . . +----+     +----+       .
.         /  |          |          |     . .   |          |  \       .
. +----+ /   |          |          |     . .   |          |   \+----+.
. | M7 |/    |          |          |     . .   |          |    | P5 |.
. +----+\    |          |          |     . .   |          |   /+----+.
.        \   |          |          |     . .   |          |  /       .
.         \+----+     +----+     +----+  . . +----+     +----+       .
.          | M4 |-----| M5 |-----| M6 |------| P3 |-----| P4 |       .
.          +----+     +----+     +----+  . . +----+     +----+       .
.                                        . .                         .
. Layer-3 Topology    MTN Network        . .        IP Network       .
.                                        . ...........................
----------------------------------------------------------------------
. MTN Path Layer Topology                .
.                                        .
.          +----+     +----+     +----+  .
.          | M1'|################| M3'|  .
.          +----+ ##  +----+  ## +----+  .
.                   ##      ##           .
. +----+              ##  ##             .
. | M7'|                ##               .
. +----+              ##  ##             .
.                   ##      ##           .
.          +----+ ##  +----+  ## +----+  .
.          | M4'|################| M6'|  .
.          +----+     +----+     +----+  .
.                                        .
.                                        .
..........................................
        .
```

Figure 2. A network with MTN Domain and IP Domain

Figure 2 gives an example of a network with a MTN domain and an IP domain. M1 to M7 are MTN nodes, and P1 to P4 are IP nodes. The same set of MTN nodes builds two separate network layers. The topology in the IP layer shows the layer-3 connectivity between the MTN nodes and the connectivity with the IP network domain, while the topology in the MTN Path layer shows the MTN paths between the selected pair of MTN nodes. An end-to-end path from M7 to P5 can be established in layer-3 using a SID list representing the layer-3 path {M7, M1, M2, M3, P1, P2, P5}. While for services which require low latency, an end-to-end path consisting of both the layer-3 segments and MTN paths could be established using an SRv6 SID list representing the path {M7, M1::C3, P1, P2, P5}, where the End.XU SID M1::C3 represents the MTN path M1'-M3'.

This shows that it is convenient to use an integrated SID list to program an inter-layer path both within the MTN domain, and across the IP and MTN domain using the combination of L3 SRv6 SIDs and the End.XU SIDs.

5. Security Considerations

TBD

6. IANA Considerations

This document defines a new SRv6 Endpoint behavior called END.XU.

IANA has allocated the following code points for different flavors of End.XU from the "SRv6 Endpoint Behaviors" sub-registry in the "Segment-routing with IPv6 data plane (SRv6) Parameters" registry:

Value	Hex	Endpoint Behavior	Reference
150	0x0096	End.XU	[This ID]
151	0x0097	End.XU with PSP	[This ID]
152	0x0098	End.XU with USP	[This ID]
153	0x0099	End.XU with USD	[This ID]
154	0x009A	End.XU with PSP, USP & USD	[This ID]
155	0x009B	End.XU with REPPLACE-CSID	[This ID]
156	0x009C	End.XU with REPPLACE-CSID & PSP	[This ID]
157	0x009D	End.XU with REPPLACE-CSID, PSP, USP & USD	[This ID]

7. Acknowledgements

The authors would like to thank Xiaodong Chang and Yongjian Hu for their review and comments.

8. References

8.1. Normative 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>.

[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>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

8.2. Informative References

[I-D.ietf-teas-actn-poi-applicability]
Peruzzini, F., Bouquier, J., Busi, I., King, D., and D. Ceccarelli, "Applicability of Abstraction and Control of Traffic Engineered Networks (ACTN) to Packet Optical Integration (POI)", Work in Progress, Internet-Draft, draft-ietf-teas-actn-poi-applicability-07, 10 July 2022, <https://www.ietf.org/archive/id/draft-ietf-teas-actn-poi-applicability-07.txt>.

[ITU-T_G.8310] ITU-T, "ITU-T G.8310: Architecture of the metro transport network", https://www.itu.int/rec/T-REC-G.8310-202012-I/en, December 2020.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Campus, No.156 Beiqing Road
Beijing, 100095
China

Email: jie.dong@huawei.com

Liuyan Han
China Mobile
No.32 XuanWuMen West Street
Beijing, 100053
China

Email: hanliuyan@chinamobile.com

Zongpeng Du
China Mobile
No.32 XuanWuMen West Street
Beijing, 100053

China

Email: duzongpeng@foxmail.com

Minxue Wang
China Mobile
No.32 XuanWuMen West Street
Beijing, 100053
China

Email: wangminxue@chinamobile.com