Network Working Group C. Hu
Internet-Draft Huawei Technologies
Intended status: Standards Track July 02, 2018
Expires: January 3, 2019

Enhance IPv6-Segment-Routing-based EVPN VPWS All Active Usage
draft-hu-bess-srv6-vpn-bypass-sid-00

Abstract

This document describes the extensions to enhance SRv6 EVPN VPWS all-active Reliability.

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

In SRv6 EVPN VPWS all-active scenario, a router or switch (CE1) is dual-homed to enterprise site (PE1 and PE2). SRv6 EVPN VPWS service is run between enterprise sites (PE1, PE2, and CPE). When one PE fails, services can be rapidly switched to the other PE, minimizing the impact on services.

As shown in Figure 1, deploy fast reroute(FRR) service on PE1 and PE2. When the AC(attachment circuit) link on PE1 fails, PE1 receives downlink traffic and can bypass it to the PE2 device for forwarding. PE2 is also the same. If the AC side links on PE1 and PE2 fail together, a brief traffic loop between PE1 and PE2 occurs. The traffic loop will waste the forwarding resources of the equipment and cause performance pressure. The length of the traffic loop depends on the convergence of the control plane. That is, PE1 withdraws the per-EVI Ethernet A-D route advertised to PE2. The FRR backup path on PE2 is destroyed. PE2 does not send traffic to PE1. In order to solve the above problem, this document defines a sub type of the SRv6 VPN SID attribute [draft-dawra-idr-srv6-vpn], to be included with per-EVI Ethernet A-D routes.

```
                                      +-----+
                                      | CE2 |
                                      +-----+
                                         |
                                      +-----+
                                      |EVPL1| Local/Remote Ethernet Tag ID-
>100/200
                   -------------------| PE3 |
                      |               +-----+
                      |               /      \
                      |              /        \
                 SRv6 EVPN VPWS     /          \
                      |            /            \
                      |           /              \
                      |      +-----+SRv6 Bypass  +-----+
                   --------- | PE1 | Tunnel      | PE2 |
L/R Ethernet Tag ID->200/100 |EVPL1|-------------|EVPL1| L/R Ethernet Tag ID-
>200/100
                             +-----+             +-----+
                                   \             /
                                    \           /
                                    ESI1      ESI1
                                      \ Trunk /
                                      +\-----/+
                                      | \   / |
                                      +---+---+
                                          |
                                       +-----+
                                       | CE1 |
                                       +-----+
```

Figure 1: Basic networking of the SRv6 EVPN VPWS all-active scenario

2. SRv6 VPN Bypass SID Attribute

The SRv6 VPN Bypass SID is a sub type of the SRv6 VPN SID. The SRv6 VPN SID has been defined in draft-dawra-idr-srv6-vpn as follows:

```
 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           |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  SRv6 SID information(Variable)                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

SRv6 SID information is encoded as follows:

```
+---------------------------------------+
|  SID Type (1 Octet)                   |
+---------------------------------------+
|  SRv6 SID (16 octet)                  |
+---------------------------------------+
```

Figure 2: SRv6 VPN Bypass SID Attribute

Because the SID Type values 1 and 2 have already been defined, the SID Type of the SRv6 VPN Bypass SID is a value to be defined that is different from 1, 2. Current Type of SID defined as:

o Type-3(Type value is TBD) - corresponds to the equivalent functionality provided by a MPLS Label1 for EVPN Route-Types as defined in [RFC7432]. Some functions which MAY be encoded are End.DX2L, End.DT2UL etc.

We define hereafter a set of new functions that can be associated with a SID. As in draft-filsfils-spring-srv6-network-programming, a function is locally defined on the node where it is executed and may range from simply moving forward in the segment list to any complex user-defined behavior.

End.DX2L Endpoint with decapsulation and Layer-2 cross-connect to local access

End.DT2UL Endpoint with decapsulation and unicast Local MAC L2 table lookup

2.1. End.DX2L: Endpoint with decapsulation and Layer-2 cross-connect to local access

The "Endpoint with decapsulation and Layer-2 cross-connect to local access OIF" function (End.DX2L for short) is a variant of the endpoint function.

When N receives a packet destined to S and S is a local End.DX2L SID, N does:

1. IF NH=SRH and SL > 0

2. drop the packet ;; Ref1

3. ELSE IF ENH = 59 ;; Ref2

4. pop the (outer) IPv6 header and its extension headers

5. forward the resulting frame via local access OIF associated to the SID

6. ELSE

7. drop the packet

Ref1: An End.DX2L SID must always be the last SID, or it can be the Destination Address of an IPv6 packet with no SRH header.

Ref2: We conveniently reuse the next-header value 59 allocated to IPv6 No Next Header [RFC8200]. When the SID corresponds to function End.DX2L and the Next-Header value is 59, we know that an Ethernet frame is in the payload without any further header.

An End.DX2L function could be customized to expect a specific VLAN format and rewrite the egress VLAN header before forwarding on the outgoing interface.

One of the applications of the End.DX2L function is the L2VPN use-case.

2.2. End.DT2UL: Endpoint with decapsulation and unicast Local MAC L2 table lookup

The " Endpoint with decapsulation and unicast Local MAC L2 table lookup " function (End.DT2UL for short) is a variant of the endpoint function.

When N receives a packet destined to S and S is a local End.DT2UL SID, N does:

1. IF NH=SRH and SL > 0

2. drop the packet ;; Ref1

3. ELSE IF ENH = 59 ;; Ref2

4. pop the (outer) IPv6 header and its extension headers

5. learn the exposed inner MAC SA in L2 table T ;; Ref3

6. lookup the exposed inner MAC DA in L2 table T(Local)

7. forward via the matched T entry else to all L2OIF in T(Local)

8. ELSE

9. drop the packet

Ref1: An End.DT2UL SID must always be the last SID, or it can be the Destination Address of an IPv6 packet with no SRH header.

End.DT2UL and the Next-Header value is 59, we know that an Ethernet frame is in the payload without any further header.

Ref3: In EVPN, the learning of the exposed inner MAC SA is done via control plane.

The End.DT2UL is used for EVPN Bridging unicast Local use cases.

3. Control Plane Processing

As shown in Figure 1:

1. PE1 advertises per-EVI Ethernet A-D routes to PE2 and PE3. The route carries the SRv6 VPN SID (SID Type=2, End.DX2) sid1 and SRv6 VPN Bypass SID sid11 allocated by the EVPL1 service on PE1.

2. The PE2 device receives the per-EVI Ethernet A-D route advertised by PE1 and finds that it is the same as the Local/Remote Ethernet Tag ID and ESI1 of its own EVPL1. PE2 considers it to be a dual-homing relationship with PE1. PE2 uses the SRv6 VPN Bypass SID to establish an SRv6 bypass path to PE1. The tunnel is marked as sid11. The SRv6 VPN Bypass SID takes effect when its EVPL Local/Remote Ethernet Tag ID and ESI are the same as the per-EVI Ethernet A-D route received.

3. The EVPL1 Local/Remote Ethernet Tag ID of the PE3 device matches PE1. PE3 uses the SRv6 VPN SID to establish an EVPN VPWS service to PE1. The service is marked as sid1. PE3's EVPL1 Local/Remote Ethernet Tag ID and ESI are different from the per-EVI Ethernet A-D routes received. PE3 should ignore this attribute.

4. In the same way, PE2 advertises per-EVI Ethernet A-D routes to PE1 and PE3. The routes carry the SRv6 VPN SID sid2 and SRv6 VPN Bypass SID sid22 allocated by EVPL1 services on PE2.

5. Finally, the primary path from PE1 to CE1 is the local AC port and the bypass path is the SRv6 tunnel labeled by sid22. The primary path from PE2 to CE1 is the local AC port and the bypass path is the SRv6 tunnel labeled by sid11. Paths from PE3 to PE1 and PE2 are marked as sid1 and sid2.

4. Data Packets Processing

This section will describe the processes of the downlink Layer 2 packet forwarding cases.

As shown in Figure 1:

1. After receiving a Layer 2 packet sent by the CE2, PE3 encapsulates the packet with the EVPL1 sid1 as the destination IPv6 of the SRH header, and forwards the packet to PE1.

2. After receiving a Layer 2 packet sent by the PE3, PE1 parses the EVPL1 sid1 of the SRH header and forwards it according to the function End.DX2 of sid1. When the primary path from PE1 to CE1 fails, PE1 encapsulates the packet with the EVPL1 bypass sid22 as the destination IPv6 of the SRH header, and forwards the packet to PE2.

3. After receiving a Layer 2 packet sent by the PE1, PE2 parses the EVPL1 bypass sid22 of the SRH header and forwards it according to the function End.DX2L of sid22. When the primary path from PE2 to CE1 fails, PE2 discards the packet and successfully breaks the loop.

4. As above, if PE2 receives a Layer 2 packet from PE3, EVPL1 bypass sid11 can also break the loop.

5. EVPN Multipoint to Multipoint (MP2MP) services

In SRv6 EVPN Multipoint to Multipoint (MP2MP) all-active scenario, function End.DT2UL of SRv6 VPN Bypass SID Attribute also has a similar effect. When the AC side links on PE1 and PE2 fail together, downlink Layer 2 unicast packet will not traffic loop.

6. IANA Considerations

TBD.

7. Security Considerations

TBD.

8. Acknowledgements

The authors of this document would like to thank xxx for their comments and review of this document.

9. Contributors

The following individuals gave significant contributions to this document:

Bingshe Liu

Huawei Technologies

liubingshe@huawei.com

Haibo Wang

Huawei Technologies

rainsword.wang@huawei.com

10. References

[I-D.dawra-idr-srv6-vpn]
Dawra, G., Filsfils, C., Dukes, D., Brissette, P., Camarillo, P., Leddy, J., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Steinberg, D., Raszuk, R., Decraene, B., Matsushima, S., and S. Zhuang, "BGP Signaling of IPv6-Segment-Routing-based VPN Networks", draft-dawra-idr-srv6-vpn-04 (work in progress), June 2018.

[I-D.filsfils-spring-srv6-network-programming]
Filsfils, C., Li, Z., Leddy, J., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Steinberg, D., Raszuk, R., Matsushima, S., Lebrun, D., Decraene, B., Peirens, B., Salsano, S., Naik, G., Elmalky, H., Jonnalagadda, P., and M. Sharif, "SRv6 Network Programming", draft-filsfils-spring-srv6-network-programming-04 (work in progress), March 2018.

[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>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[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>.

[RFC8214] Boutros, S., Sajassi, A., Salam, S., Drake, J., and J. Rabadan, "Virtual Private Wire Service Support in Ethernet VPN", RFC 8214, DOI 10.17487/RFC8214, August 2017, <https://www.rfc-editor.org/info/rfc8214>.

Author's Address

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

Email: huchongyang@huawei.com