BESS
Internet Draft
Intended status: Standards Track
Expires: July 26, 2026

Yisong Liu
China Mobile
C. Lin
New H3C Technologies
Y. Liu
ZTE

January 24, 2026

SRv6 L3VPN Fast Reroute
draft-liu-bess-srv6-l3vpn-fast-reoute-01

Abstract

In some multihoming SRv6 L3VPN scenarios, once fast reroute has taken place, a second fast reroute is undesirable and may cause looping. This document proposes a mechanism to prevent further fast reroutes by advertising No-Further-FRR behaviors for L3 SRv6 Service SIDs in BGP messages.

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 July 26, 2026.

Table of Contents

1. Introduction

[RFC9252] defines procedures and messages for SRv6-based BGP services, including Layer 3 Virtual Private Network (L3VPN), Ethernet VPN (EVPN), and Internet services. In some multihoming scenarios, two egress PEs may establish a backup path between them and use it as the protection of PE-CE link failure. Once fast reroute (FRR) has taken place, a second fast reroute is undesirable and may cause looping.

This document defines the No-Further-FRR behavior for L3 SRv6 Service SIDs carried in BGP messages and proposes a mechanism using the No-Further-FRR flavor to prevent further fast reroutes.

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 Case

2.1. SRv6 L3VPN Multihoming

In the multihoming SRv6 L3VPN scenarios, two egress PEs may establish a backup path between them and use it as the protection of PE-CE link failure.

Take the network in Figure 1 as an example. When traffic goes from CE1 to CE2, it may be load-balanced between PE2 and PE3 or only forwarded to the main egress PE. If the link PE2-CE2 fails, PE2 can still forward the traffic for CE2 by sending it over the backup path to PE3 (and similarly for PE3 if link2 fails).

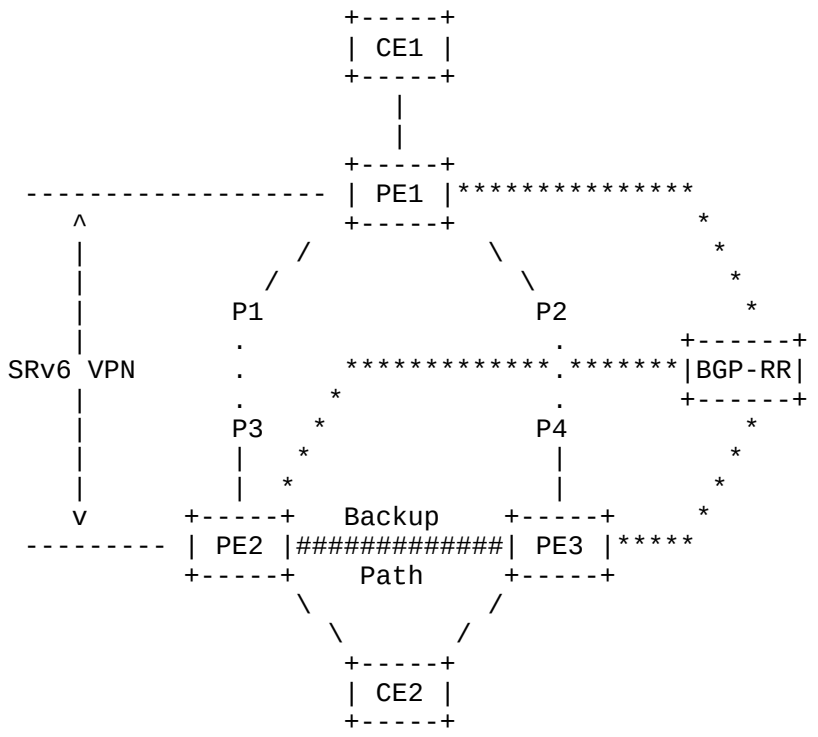

Figure 1

Examples of BGP routes advertised by PE2 and PE3 are as following:

```
BGP Route by PE2:
  VPN Prefix of CE2:
    BGP Prefix SID Attr:
      SRv6 L3 Service TLV:
        SRv6 SID Information sub-TLV:
          SID: SID-2
            Behavior: End.DT46

BGP Route by PE3:
  VPN Prefix of CE2:
    BGP Prefix SID Attr:
      SRv6 L3 Service TLV:
        SRv6 SID Information sub-TLV:
          SID: SID-3
            Behavior: End.DT46
```

Examples of FIB entries for L3VPN service SID on PE2 and PE3 are as following:

```
FIB on PE2:
  SID-2:
    Primary Next-hop: CE2
    Backup Next-hop: Service SRv6 SID-3

FIB on PE3:
  SID-3:
    Primary Next-hop: CE2
    Backup Next-hop: Service SRv6 SID-2
```

However, suppose CE2 is down. PE2 will think PE2-CE2 link is down and send traffic to PE3 over the backup path. PE3 will also think PE3-CE3 link is down and send the traffic back to PE2 over the backup path. So, traffic will loop between PE2 and PE3 until BGP convergence.

The traffic forwarding when CE2 fails is as following:

Node	Packet	Next	Comment
PE1	<SID-2> pkt	PE2	
PE2	pkt	CE2	PE2-CE2 down
PE2	<SID-3> pkt	PE3	FRR
PE3	pkt	CE2	PE3-CE2 down
PE3	<SID-3> pkt	PE2	FRR
PE2	--	CE2	PE2-CE2 down
PE2	<SID-3> pkt	PE3	FRR
...			Loop!

3. Solution

Each egress PE advertises an additional SRv6 Service SID in BGP routes which is called No-Further-FRR SID.

The owner of No-Further-FRR SID will not provide local FRR for it. When the next-hop of No-Further-FRR SID is down, like PE-CE link failure or CE node failure, the PE will drop packets rather than apply FRR.

The No-Further-FRR SID can used by other PE as the protection of local PE-CE link failure, without worrying about the looping problem.

To support backwards compatibility and BGP RR deployment, both the normal SRv6 Service SID and the No-Further-FRR SID MAY be advertised together. A No-Further-FRR behavior is used to indicate the No-Further-FRR SID.

Detailed BGP extensions will be described in Section 4.

Still taking the network in Figure 1 as an example, the BGP routes advertised by PE2 and PE3 are as following:

```
BGP Route by PE2:
  VPN Prefix of CE2:
    BGP Prefix SID Attr:
      SRv6 L3 Service TLV:
        SRv6 SID Information sub-TLV:
          SID: SID-21
            Behavior: End.DT46
        SRv6 SID Information sub-TLV:
          SID: SID-22
            Behavior: End.DT46.Reroute

BGP Route by PE3:
  VPN Prefix of CE2:
    BGP Prefix SID Attr:
      SRv6 L3 Service TLV:
        SRv6 SID Information sub-TLV:
          SID: SID-31
            Behavior: End.DT46
        SRv6 SID Information sub-TLV:
          SID: SID-32
            Behavior: End.DT46.Reroute
```

The FIB entries for L3VPN service SID on PE2 and PE3 are as following:

```
FIB on PE2:
  SID-21,Behavior: End.DT46:
    Primary Next-hop: CE2
    Backup Next-hop: Service SRv6 SID-32
  SID-22 (No-Further-FRR),Behavior: End.DT46.Reroute:
    Only Forwarding on Primary Next-hop: CE2

FIB on PE3:
  SID-31, Behavior: End.DT46:
    Primary Next-hop: CE2
    Backup Next-hop: Service SRv6 SID-22
  SID-32 (No-Further-FRR), Behavior: End.DT46.Reroute:
    Only Forwarding on Primary Next-hop: CE2
```

After adopting the proposed solution, if CE fails, PE2 will think PE2-CE2 link is down and send traffic to PE3 by using the No-Further-FRR SID-32. PE3 will also think PE3-CE3 link is down, but PE3 will drop the packets rather than apply FRR.

The traffic forwarding when CE2 fails is as following:

Node	Packet	Next	Comment
PE1	<SID-21> pkt	PE2	
PE2	pkt	CE2	PE2-CE2 down
PE2	<SID-32> pkt	PE3	FRR
PE3	pkt	CE2	PE3-CE2 down
PE3	-	-	Drop

4. Extensions for L3 SRv6 Endpoint Behaviors

4.1. End.DT4.Reroute : End.DT4 with Fast Reroute

The "End.DT4 with Fast Reroute" behavior ("End.DT4.Reroute" for short) is a variant of the End.DT4 behavior.

The End.DT4.Reroute behavior is defined for the fast-reroute application between two multi-homing peers, and extends the base End.DT4 behavior.

When processing the Upper-Layer header of a packet matching a FIB entry locally instantiated as an End.DT4 SID, N does the following:

```
S01. If (Upper-Layer header type == 4(IPv4) ) {
S02.    Remove the outer IPv6 header with all its extension headers
S03.    Set the packet’s associated FIB table to T
S04.    Submit the packet to the egress IPv4 FIB lookup for
           transmission to the new destination
S05.           if (The forwarding path for the new destination is
                   the backup path generated by Fast Reroute) {
S06.         Drop the packet
S07.           } Else {
S08.             Transmission to the new destination
S09.           }
S10. } Else {
S11.    Process as per Section 4.1.1
S12. }
```

4.2. End.DT6.Reroute : End.DT6 with Fast Reroute

The "End.DT6 with Fast Reroute" behavior ("End.DT6.Reroute" for short) is a variant of the End.DT6 behavior.

The End.DT6.Reroute behavior is defined for the fast-reroute application between two multi-homing peers, and extends the base End.DT6 behavior.

When processing the Upper-Layer header of a packet matching a FIB entry locally instantiated as an End.DT6 SID, N does the following:

```
S01. If (Upper-Layer header type == 41(IPv6) ) {
S02.    Remove the outer IPv6 header with all its extension headers
S03.    Set the packet’s associated FIB table to T
S04.    Submit the packet to the egress IPv6 FIB lookup for
           transmission to the new destination
S05.           if (The forwarding path for the new destination is
                   the backup path generated by Fast Reroute) {
S06.         Drop the packet
S07.           } Else {
S08.              Transmission to the new destination
S09.           }
S10. } Else {
S11.    Process as per Section 4.1.1
S12. }
```

4.3. End.DT46.Reroute : End.DT46 with Fast Reroute

The "End.DT46 with Fast Reroute" behavior ("End.DT46.Reroute" for short) is a variant of the End.DT46 behavior.

The End.DT46.Reroute behavior is defined for the fast-reroute application between two multi-homing peers, and extends the base End.DT46 behavior.

When processing the Upper-Layer header of a packet matching a FIB entry locally instantiated as an End.DT46 SID, N does the following:

```
S01. If (Upper-Layer header type == 4(IPv4) ) {
S02.    Remove the outer IPv header with all its extension headers
S03.    Set the packet's associated FIB table to T4
S04.    Submit the packet to the egress IPv4 FIB lookup for
           transmission to the new destination
S05.           if (The forwarding path for the new destination is
                  the backup path generated by Fast Reroute) {
S06.         Drop the packet
S07.           } Else {
S08.             Transmission to the new destination
S09.           }
S10. } Else If (Upper-Layer header type == 41(IPv6) ) {
S11.    Remove the outer IPv6 header with all its extension headers
S12.    Set the packet's associated FIB table to T6
S13.    Submit the packet to the egress IPv6 FIB lookup for
           transmission to the new destination
S14.           if (The forwarding path for the new destination is
                   the backup path generated by Fast Reroute) {
S15.         Drop the packet
S16.           } Else {
S17.              Transmission to the new destination
S18.           }
S19. } Else {
S20.    Process as per Section 4.1.1
S21. }
```

4.4. End.DX4.Reroute : End.DX4 with Fast Reroute

The "End.DX4 with Fast Reroute" behavior ("End.DX4.Reroute" for short) is a variant of the End.DX4 behavior.

The End.DX4.Reroute behavior is defined for the fast-reroute application between two multi-homing peers, and extends the base End.DX4 behavior.

When processing the Upper-Layer header of a packet matching a FIB entry locally instantiated as an End.DX4 SID, N does the following:

```
S01. If (Upper-Layer header type == 4(IPv4) ) {
S02.    Remove the outer IPv6 header with all its extension headers
S03.    If (L3 adjacency J is the backup path generated by Fast
            Reroute) {
S04.        Drop the packet
S05.    } Else {
S06.          Forward the exposed IPv4 packet to the L3 adjacency J
S07.    }
S08. } Else {
S09.    Process as per Section 4.1.1
S10. }
```

4.5. End.DX6.Reroute : End.DX6 with Fast Reroute

The "End.DT6 with Fast Reroute" behavior ("End.DX6.Reroute" for short) is a variant of the End.DX6 behavior.

The End.DX6.Reroute behavior is defined for the fast-reroute application between two multi-homing peers, and extends the base End.DX6 behavior.

When processing the Upper-Layer header of a packet matching a FIB entry locally instantiated as an End.DX6 SID, N does the following:

```
S01. If (Upper-Layer header type == 41(IPv6) ) {
S02.    Remove the outer IPv6 header with all its extension headers
S03.    If (L3 adjacency J is the backup path generated by Fast
            Reroute) {
S04.        Drop the packet
S05.    } Else {
S06.          Forward the exposed IPv6 packet to the L3 adjacency J
S07.    }
S08. } Else {
S09.    Process as per Section 4.1.1
S10. }
```

5. Backward Compatibility

To maintain backwards-compatibility, both End.DT4.Reroute and End.DT4 Behavior SIDs MAY be advertised together. Receiving PEs SHOULD use the SRv6 SID from the first instance of the Sub-TLV only (Section 3.1 of [RFC9252]), and ignore the SRv6 SID of unknown behavior End.DT4.Reroute (Section 3.2.1 of [RFC9252]).
The same compatibility handling applies to other behaviors such as End.DT6.Reroute and End.DT6, End.DT46.Reroute and End.DT46, End.DX4.Reroute and End.DX4, and End.DX6.Reroute and End.DX6.

6. SID Allocation Optimization Considerations

To reduce the allocation of SRv6 Service SIDs, only one SRv6 Service SID (with Behaviors End.DT4, End.DT6, End.DT46, End.DX4, End.DX6) can be allocated. Then, by setting the optional Fast Reroute argument "Arg.FR2," the SID can be distinguished as either for Fast Reroute or not. Specifically, "Arg.FR2" differentiates between Behaviors with Fast Reroute (End.DT4.Reroute, End.DT6.Reroute, End.DT46.Reroute, End.DX4.Reroute, End.DX6.Reroute) and those without Fast Reroute (End.DT4, End.DT6, End.DT46, End.DX4, End.DX6).

For example, using End.DT4, the implementation process is as follows:
The SRv6 L3 Service TLV in this case will carry two SRv6 SID Information sub-TLVs:

* the first one with the base End.DT4 behavior and

* the second one with the End.DT4.Reroute behavior variant. The second one will have a non-zero Arg length (AL) and convey Arg.FR2 embedded in the advertised SID.

Following is an example representation of the BGP Prefix-SID Attribute encoding in this case for a 16-bit argument Arg.FR2 (0x0001):

```
BGP Prefix SID Attr:
   SRv6 L3 Service TLV:
      SRv6 SID Information sub-TLV:
         SID: 2001:123:a:1:1234::
           Behavior: End.DT4
           SRv6 SID Structure sub-sub-TLV:
           LBL: 48, LNL: 16, FL: 16, AL: 0, TPOS-L: 0, TPOS-O: 0
      SRv6 SID Information sub-TLV:
         SID: 2001:123:a:1:1234:0001::
           Behavior: End.DT4.Reroute
           SRv6 SID Structure sub-sub-TLV:
           LBL: 48, LNL: 16, FL: 16, AL: 16, TPOS-L: 0, TPOS-O: 0
```

The processing of other types End.DT6, End.DT46, End.DX4, and End.DX6 is similar to that of DT4.

7. Security Considerations

TBD.

8. IANA Considerations

This document introduces two new Endpoint behaviors. This document requests IANA assign a two new values and update the "SRv6 Endpoint Behaviors" subregistry under the top-level "Segment Routing" registry as follows:

Value	Hex	Endpoint Behavior	Reference
TBD	TBD	End.DT4.Reroute	This document
TBD	TBD	End.DT6.Reroute	This document
TBD	TBD	End.DT46.Reroute	This document
TBD	TBD	End.DX4.Reroute	This document
TBD	TBD	End.DX6.Reroute	This document

Table 1: SRv6 Endpoint Behaviors Subregistry

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017

[RFC9252] Dawra, G., Ed., Talaulikar, K., Ed., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)", RFC 9252, DOI 10.17487/RFC9252, July 2022, <https://www.rfc-editor.org/info/rfc9252>.

Authors' Addresses

Yisong Liu
China Mobile
China
Email: liuyisong@chinamobile.com

Changwang Lin
New H3C Technologies
China
Email: linchangwang.04414@h3c.com

Yao Liu
ZTE
China
Email: liu.yao71@zte.com.cn