rtgwg Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 26, 2022

W. Cheng
W. Jiang
China Mobile
C. Lin
Y. Qiu
New H3C Technologies
April 27, 2022

SRv6 Egress Protection in Multi-home scenario
draft-cheng-rtgwg-srv6-multihome-egress-protection-00

Status of this Memo

Copyright Notice

Abstract

This document describes a SRv6 egress node protection mechanism in multi-home scenarios.

Table of Contents

1. Introduction 2
2. Terminology 3
3. Multi-home SRv6 Egress Protection Mechanism 3
3.1. B-flag in Segment Routing Header 3
3.2. Procedure of Multi-home Egress Protection on SRv6 TE Path 3
3.2.1. Procedure on the Ingress Endpoint 4
3.2.2. Procedure on the Penultimate Endpoint 6
3.3. Procedure of Multi-home Egress Protection on SRv6 BE Path 7
4. Multi-home SRv6 Egress Protection Example 8
5. IANA Considerations 10
6. Security Considerations 10
7. References 11
7.1. Normative References 11
7.2. Informative References 12
8. Acknowledgments 12
Authors' Addresses 12

1. Introduction

The fast protection of a transit node of a Segment Routing (SR) path or tunnel is described in [I-D.ietf-rtgwg-segment-routing-ti-lfa] and [I-D.hu-spring-segment-routing-proxy-forwarding]. [RFC8400] specifies the fast protection of egress node(s) of an MPLS TE LSP tunnel including P2P TE LSP tunnel and P2MP TE LSP tunnel in details. However, these documents do not discuss the fast protection of the egress node of a Segment Routing for IPv6 (SRv6) path or tunnel.

[I-D.ietf-rtgwg-srv6-egress-protection] proposes mirror protection mechanism and presents protocol extensions for the fast protection of the egress node of a SRv6 path or tunnel. However, the mechanism provided in this document is relatively complex. It is necessary to configure the Mirror SID for the protected egress node on the backup egress node. The mirror relationship is distributed through IGP and BGP protocols to automatically create mapping entries.

This document introduces a simplified protection mechanism of the egress node of a SRv6 path. Only expanding the data plane can perform fast path switching in case of egress node failure.

2. Terminology

The following terminologies are used in this document.

SR: Segment Routing

SRv6: SR for IPv6

SRH: Segment Routing Header

SID: Segment Identifier

CE: Customer Edge

PE: Provider Edge

VPN: Virtual Private Network

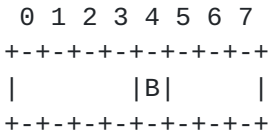
3. Multi-home SRv6 Egress Protection Mechanism

This section describes the mechanism of SRv6 path egress protection in multi-home scenarios and the extension of SRH extension header.

3.1. B-flag in Segment Routing Header

[RFC8754] describes the Segment Routing Header (SRH) and how SR capable nodes use it. The SRH contains an 8-bit "Flags" field.

This document defines the following bit in the SRH Flags field to carry the B-flag:



Where:

- B-flag: The marking bit of carrying backup SID in segment list. If the B-flag is set to 1, a backup SID is carried in the segment list.

3.2. Procedure of Multi-home Egress Protection on SRv6 TE Path

The Figure 1 is used to explain the multi-home egress node protection mechanism.

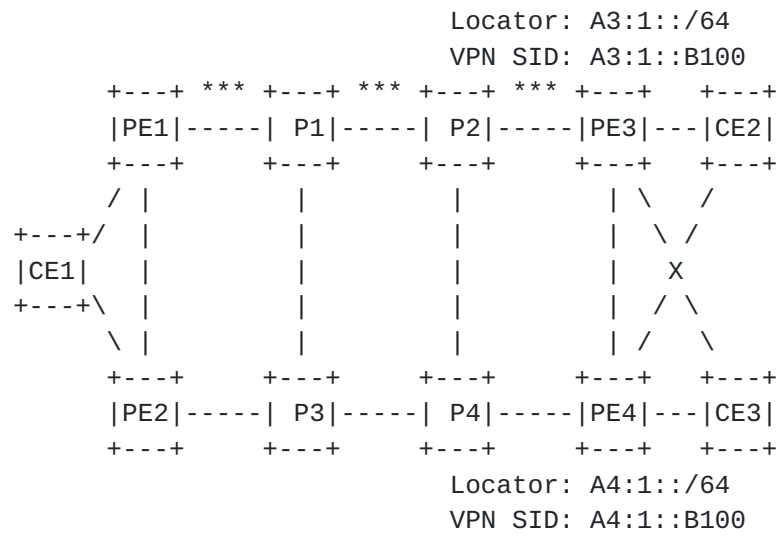


Figure 1

3.2.1. Procedure on the Ingress Endpoint

In the multi-home or dual-home scenario, after the ingress node learns the multi-home or dual-home route through routing protocol, it determines the optimal path and suboptimal path according to the route optimization strategy. The egress node on the optimal path is an primary egress, and the SID of the primary egress node is used as the primary SID The egress node on the suboptimal path is an backup egress,and the SID of the backup egress node is used as the backup SID.

On the path forwarded based on SRv6 TE policy, when the ingress node encapsulates the SRH extension header, judge whether the primary VPN SID of the egress node (PE1) has a backup SID. If yes, insert the backup SID into the position of SRH[Last Entry], and set B-flag to 1 to identify that the backup SID has been carried in the last

position of the segment list, then the value of SL is set to n-1. The format of SRH extension header filling is shown in the following figure 2.

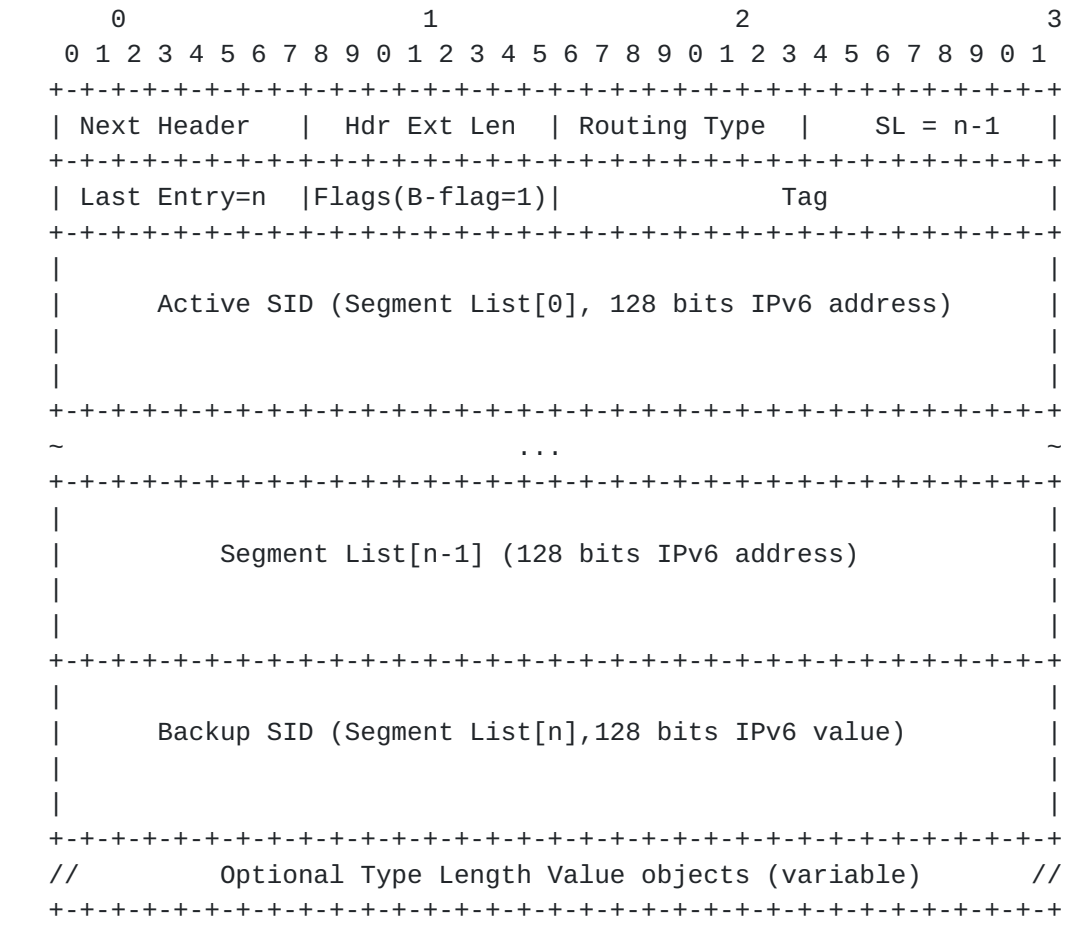


Figure 2

3.2.2. Procedure on the Penultimate Endpoint

Normally, the traffic is forwarded along the path P1->P2->PE3->CE2. When primary egress node (PE3) fails, P2 finds out that the PE3's SID is unreachable and the B-flag value is set. Then P2 modifies the destination address of the packet to SRH[Last Entry] which is the backup SID, and sends the modified packet to backup egress node (PE4). Through this method P2 can provide fast protection for the egress failure.

The detailed processing can be described in two cases according to the endpoint behavior of the destination address of the packet received by P2.

The behavior of the local endpoint is END.X

When receiving a packet destined to a local End.X SID whose outgoing interface is down, the penultimate endpoint acting as a Repair Node can provide fast protection for the failure of directly connected egress nodes after SL decreasing through executing the following procedures.

```
 IF B-flag = 1 THEN
    IF SL = 0 and the failed egress node is directly connected to
Repair Node THEN
          Update the IPv6 DA with SRH[Last Entry];
          FIB lookup on the updated DA;
          Forward the packet according to the matched entry;
    ELSE IF SL = 1 and SRH[1] and SRH[0] are the SIDs of the
failed egress node directly connected to Repair Node THEN
          Update the IPv6 DA with SRH[Last Entry];
          FIB lookup on the updated DA;
          Forward the packet according to the matched entry;
```

The behavior of the local endpoint is END

After looking up the FIB for the updated DA with Segment List[Segments Left] and SL decreasing, in the following two cases, the penultimate endpoint acting as a Repair Node can provide fast protections for the failure of directly connected egress nodes through executing the following procedure.

Case 1: For the packet whose Next Header is SRH and Segments Left is equal to 1, perform the following processing:

IF B-flag = 1 and SRH[1] and SRH[0] are the SIDs of the failed egress node directly connected to Repair Node THEN
Update the IPv6 DA with SRH[Last Entry];
FIB lookup on the updated DA;
Forward the packet according to the matched entry;

Case 2: For the packet whose Next Header is SRH and Segments Left is equal to 0, perform the following processing:

IF B-flag = 1 and the failed egress endpoint is directly connected to Repair Node THEN
Update the IPv6 DA with SRH[Last Entry];
FIB lookup on the updated DA;
Forward the packet according to the matched entry;

When the packet arrives at PE4, PE4 removes the outer IPv6 header, and forwards the exposed inner packet.

After the route convergence is completed, the ingress node (PE1) will reselect the forwarding path for the traffic to VPN, and switch the path P1->P3->P4->PE4->CE2 to the CE to the egress node (PE4). After that, P2 no longer needs to forward the packet with the destination address of PE3.

Considering that the egress node may check the consistency between the segment list and the destination address, for the packet with B-flag 1, as long as the destination address is the same as any one of SRH[0] or SRH[Last Entry], it is considered to be consistent.

In addition, when a penultimate endpoint using non-PSP-flavored SID receives a packet with B-flag of 1, it is recommended to directly remove the SRH extension header after replacing the destination address with SRH[Last Entry].

3.3. Procedure of Multi-home Egress Protection on SRv6 BE Path

The multi-home egress node protection processing on the SRv6 BE path is consistent with that on the SRv6 TE path, except that the ingress

node is required to add an SRH extension header with the active SID, backup SID and B-flag when encapsulating the outer IPv6 packet header.

In the multi-home scenario egress node scenario, the ingress node determines the active SID (PE3's SID) and the backup SID (PE4's SID) of the egress node through the optimization strategy of the routing protocol.

When the traffic from PE1 to CE2 is forwarded through the SRv6 BE path, in order to realize the fast protection of egress node failure, when the ingress node adds an outer IPv6 packet header to the forwarded packet, it must encapsulate the SRH extension header at the same time. The contents filled in the SRH extension header are the same as Figure 2 in Section 3.2.1, in which the segment list only fills in the active SID and backup SID, the SL is set to 0, the last entry is set to 1, and the B-flag is set to 1. The active SID is used as the destination address of the outer IP packet header.

Normally, because the destination address of the packet is the active SID (PE3's SID), P1 and P2 will forward the packet to PE3 according to the destination address.

Once PE3 fails, the processing of the penultimate endpoint is the same as that on the SRv6 TE path. When P2 finds out that the route to the directly connected egress node PE3 is unreachable, if the B-flag is 1, modify the destination address to the backup SID in SRH[1], and send the packet to the updated destination address.

4. Multi-home SRv6 Egress Protection Example

Figure 3 shows an example of protecting egress PE3 of a SRv6 TE path, which is from ingress PE1 to egress PE3.

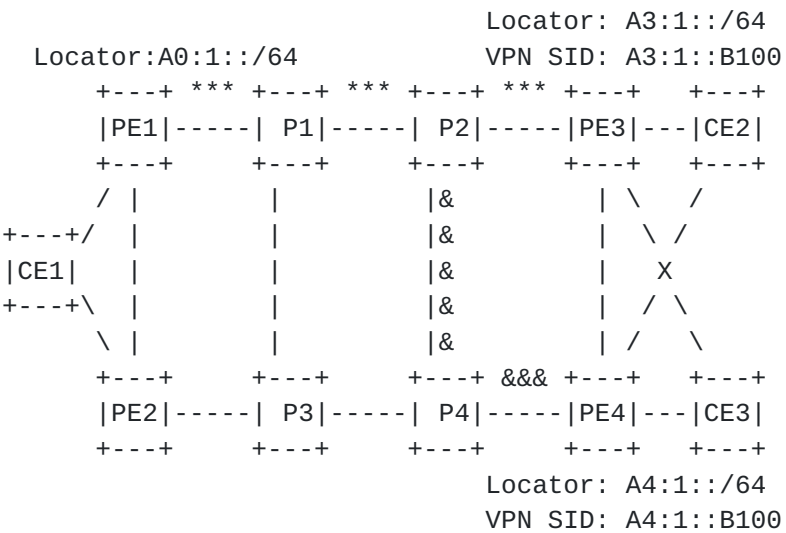


Figure 3

In this document, a SID list is represented as <S1, S2, S3> where S1 is the first SID to visit, S2 is the second SID to visit and S3 is the last SID to visit along the SRv6 path.

In Figure 3, Both CE2 and CE3 are dual home to PE3 and PE4. PE1 has a locator A0:1::/64. P1 has a locator A1:1::/64. P2 has a locator A2:1::/64 and END.X SID A2:1::A100. PE3 has a locator A3:1::/64 and a VPN SID A3:1::B100. PE4 has a locator A4:1::/64 and VPN SID A4:1::B100. The traffic from CE1 to CE2 is forwarded along the path PE1->P1->P2->PE3. After the configuration, PE1 determines that PE3's backup SID is PE4's VPN SID through the routing optimization strategy of BGP.

In normal operations, after receiving a packet with destination PE3, P2 forwards the packet to PE3 according to its FIB. When PE3 receives the packet, it sends the packet to CE2.

When PE1 receives the packet from CE1 to CE2, PE1 encapsulates the packet with IPv6 header. The segment list in SRH is designed as <A0:1::1, A1:1::1, A2:1::A100, A3:1::B100, A4:1::B100>. The SL is set to 3, the Last Entry is set to 4, and B-flag is set to 1.

When P2 receives a packet destined to END.X SID A2:1::A100, in normal operations, it forwards the packet with source A0:1::1 and

destination PE3's VPN SID A3:1::B100 from the link between P2 and PE3 according to END.X SID.

When PE3 fails, P2 receives the packet to be sent to PE3's VPN SID A3:1::B100. P2 finds that the outgoing interface is down. If the B-flag is 1, P2 changes the destination address of the packet with the backup SID of SRH[4], removes SRH extension header and sends the modified packet to A4:1::B100.

When PE4 receives the modified packet, it decapsulates the packet and forwards the decapsulated packet by executing End.DT6 behavior for an End.DT6 SID instance.

5. IANA Considerations

This document requests that IANA allocate the following registration in the "Segment Routing Header Flags" sub-registry for the "Internet Protocol Version 6 (IPv6) Parameters" registry maintained by IANA:

Bit	Description	Reference
4	B-flag	This document

6. Security Considerations

[RFC8754] defines the notion of an SR domain and use of SRH within the SR domain. The use of egress protection mechanism described in this document is restricted to an SR domain. For example, similar to the SID manipulation, B-flag manipulation is not considered as a threat within the SR domain. Procedures for securing an SR domain are defined the section 5.1 and section 7 of [RFC8754].

This document does not impose any additional security challenges to be considered beyond security threats described in [RFC8754], [RFC8679] and [RFC8986].

7. References

7.1. Normative References

[I-D.ietf-rtgwg-segment-routing-ti-lfa] Litkowski, S., Bashandy, A., Filsfils, C., Francois, P., Decraene, B., and D. Voyer, "Topology Independent Fast Reroute using Segment Routing", Work in Progress, Internet-Draft, draft-ietf-rtgwg-segment-routing-ti-lfa-08, 21 January 2022, <https://www.ietf.org/archive/id/draft-ietf-rtgwg-segment-routing-ti-lfa-07.txt>.

[I-D.hu-spring-segment-routing-proxy-forwarding] Hu, Z., Chen, H., Yao, J., Bowers, C., Yongqing, and Yisong, "SR-TE Path Midpoint Restoration", Work in Progress, Internet-Draft, draft-hu-spring-segment-routing-proxy-forwarding-18, 1 September 2022, <https://www.ietf.org/archive/id/draft-hu-spring-segment-routing-proxy-forwarding-18.txt>.

[I-D.ietf-rtgwg-srv6-egress-protection] Hu, Z., Chen, H., Chen, H., Wu, P., Toy, M., Cao, C., He T., Liu, L., Liu, X., "SRv6 Path Egress Protection", Work in Progress, Internet-Draft, draft-ietf-rtgwg-srv6-egress-protection-04, 17 October 2021, < https://www.ietf.org/archive/id/draft-ietf-rtgwg-srv6-egress-protection-04.txt >

[RFC8400] Chen, H., Liu, A., Saad, T., Xu, F., and L. Huang, "Extensions to RSVP-TE for Label Switched Path (LSP) Egress Protection", RFC 8400, DOI 10.17487/RFC8400, June 2018, <https://www.rfc-editor.org/info/rfc8400>.

[RFC8679] Shen, Y., Jeganathan, M., Decraene, B., Gredler, H., Michel, C., and H. Chen, "MPLS Egress Protection Framework", RFC 8679, DOI 10.17487/RFC8679, December 2019, <https://www.rfc-editor.org/info/rfc8679>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header(SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

[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>.

7.2. Informative References

TBD

8. Acknowledgments

The authors would like to thank the following for their valuable contributions of this document:

Yisong Liu
China Mobile

Authors' Addresses

Weiqiang Cheng
China Mobile

Email: chengweiqiang@chinamobile.com

Wenying Jiang
China Mobile

Email: jiangwenying@chinamobile.com

Changwang Lin
New H3C Technologies

Email: linchangwang.04414@h3c.com

Yuanxiang Qiu
New H3C Technologies

Email: qiuyuanxiang@h3c.com