Network Working Group
Internet-Draft
Intended status: Experimental
Expires: April 18, 2021

H. Chen
China Telecom
Z. Hu
Huawei Technologies
H. Chen
Futurewei
X. Geng
Huawei Technologies

October 15, 2020

SRv6 Midpoint Protection
draft-chen-rtgwg-srv6-midpoint-protection-03

Abstract

The current local repair mechanism, e.g., TI-LFA, allows local repair actions on the direct neighbors of the failed node to temporarily route traffic to the destination. This mechanism could not work properly when the failure happens in the destination point or the link connected to the destination. In SRv6 TE, the IPv6 destination address in the outer IPv6 header could be the dedicated endpoint of the TE path rather than the destination of the TE path. When the endpoint fails, local repair couldn't work on the direct neighbor of the failed endpoint either. This document defines midpoint protection, which enables the direct neighbor of the failed endpoint to do the function of the endpoint, replace the IPv6 destination address to the other endpoint, and choose the next hop based on the new destination address.

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

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 April 18, 2021.

Table of Contents

1. Introduction

The current mechanism, e.g., TI-LFA ([I-D.ietf-rtgwg-segment-routing-ti-lfa]), allows local repair actions on the direct neighbors of the failed node to temporarily route traffic to the destination. This mechanism could not work properly when the failure happens in the destination point or the link connected to the destination. In SRv6 TE, the IPv6 destination address in the outer IPv6 header could be the dedicated endpoint of the TE path rather than the destination of the TE path

([I-D.ietf-spring-srv6-network-programming]). When the endpoint fails, local repair couldn't work on the direct neighbor of the failed endpoint either. This document defines midpoint protection, which enables the direct neighbor of the failed endpoint to do the function of the endpoint, replace the IPv6 destination address to the other endpoint, and choose the next hop based on the new destination address.

2. SRv6 Midpoint Protection Mechanism

When an endpoint node fails, the packet needs to bypass the failed endpoint node and be forwarded to the next endpoint node of the failed endpoint. On the Repair Node (i.e., the previous hop of the failed endpoint node), it performs the proxy forwarding as follows :

- Outbound interface failure happens in the Repair Node;

Case 1: Route to the failed endpoint could be found in the FIB of Repair Node:

- If the Repair Node is not directly connected to the failed endpoint, the normal Ti-LFA is executed;

- If the Repair Node is directly connected to the failed endpoint, the Repair Node forwards the packets through a bypass to the failed endpoint, changing the IPv6 destination address with the IPv6 address of the next, the last or other reasonable endpoint nodes, which could avoid going throw the failed endpoint.

Case 2: Route to the failed endpoint could not be found in the FIB of Repair Node:

- Repair Node forwards the packets through a bypass of the failed endpoint to the next, the last or other reasonable endpoint node directly . There is no need to check whether the failed endpoint node is directly connected to the Repair Node or not.

3. SRv6 Midpoint Protection Example

The topology shown in Figure 1 illustrates an example of network topology with SRv6 enabled on each node.

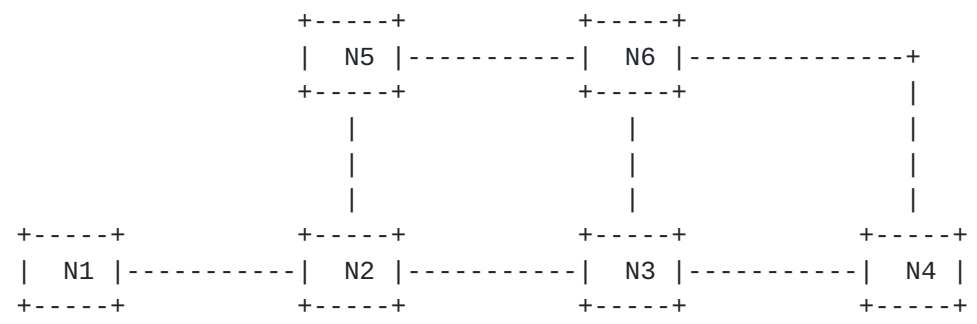

Figure 1: An example of midpoint protection

In this document, an end SID at node n with locator block B is represented as B:n. An end.x SID at node n towards node k with locator block B is represented as B:n:k. 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 TE path.

In the reference topology:

Node N1 is an ingress node of SRv6 domain. Node N1 steers a packet into a segment list < B:3, B:4>.

When Node N3 fails, the packet needs to bypass the failed endpoint node and be forwarded to the next endpoint node after the failed endpoint in the TE path. When outbound interface failure happens in the Repair Node (which is not limited to the previous hop node of the failed endpoint node), it performs the proxy forwarding as follows,:

For node N2, if the outbound interface to the endpoint B:3 is failed before IGP converges:

- Because node N2, as a Repair Node, is connected to the failed endpoint B:3 directly, node N2 forwards the packets through a bypass of the failed endpoint, changing the IPv6 destination address with the next sid B:4. N2 detects the failure of outbound interface to B:4 in the current route, it could use the normal Ti-LFA repait path to forward the packet, because it is not directly connected to the node N4. N2 encapsulates the packet with the segment list < B:5:6> as a repair path.

For node N1, route to the failed endpoint N3 could not be found in the FIB after IGP converges:

- Node N1, as a Repair Node, forwards the packets through a bypass of the failed endpoint to the next or endpoint node (e.g., N4) directly. There is no need to check whether the failed endpoint

node is directly connected to N1. N1 changes the IPv6 destination address with the next sid B:4. Since IGP has completed convergence, it forwards packets directly based on the IGP SPF path

4. SRv6 Midpoint Protection Behavior

4.1. Transit Node as Repair Node

When the Repair Node is a transit node, it provides fast protection against the endpoint node failure as follows after looking up the FIB.

```
  IF the primary outbound interface used to forward the packet failed
    IF NH = SRH && SL != 0, and
       the failed endpoint is directly connected to the Repair Node THEN
      SL decreases*; update the IPv6 DA with SRH[SL];
      FIB lookup on the updated DA;
      forward the packet according to the matched entry;
    ELSE
      forward the packet according to the backup nexthop;
  ELSE // there is no FIB entry for forwarding the packet
    IF NH = SRH && SL != 0 THEN
      SL decreases*; update the IPv6 DA with SRH[SL];
      FIB lookup on the updated DA;
      forward the packet according to the matched entry;
    ELSE
      drop the packet;
```

*: SL could decrease any dedicated value from [1-N], where N is the current value of SL.
The case is similar in the following examples.

4.2. Endpoint Node as Repair Node

When a node N receives a packet, if the destination address (DA) of the packet is a local END SID, then node N is an endpoint node. When the Repair Node is an endpoint node, it provides fast protections for the failure through executing the following procedure after looking up the FIB for the updated DA.

```
IF the primary outbound interface used to forward the packet failed
  IF NH = SRH && SL != 0, and
     the failed endpoint is directly connected to the Repair Node THEN
    SL decreases; update the IPv6 DA with SRH[SL];
    FIB lookup on the updated DA;
    forward the packet according to the matched entry;
  ELSE
    forward the packet according to the backup nexthop;
ELSE // there is no FIB entry for forwarding the packet
  IF NH = SRH && SL != 0 THEN
    SL decreases; update the IPv6 DA with SRH[SL];
    FIB lookup on the updated DA;
    forward the packet according to the matched entry;
  ELSE
    drop the packet;
ELSE
  forward accordingly to the matched entry;
```

4.3. Endpoint x Node as Repair Node

An endpoint node with cross-connect (End.X for short) is an endpoint node with an array of layer 3 adjacencies. When a node N receives a packet, if the destination address (DA) of the packet is a local END.X SID, then node N as Repair Node provides fast protections for the failure through executing the following procedure after updating DA.

```
  IF the layer-3 adjacency interface is down THEN
    FIB lookup on the updated DA;
    IF the primary interface used to forward the packet failed THEN
      IF NH = SRH && SL != 0, and
         the failed endpoint directly connected to the Repair Node THEN
        SL decreases; update the IPv6 DA with SRH[SL];
        FIB lookup on the updated DA;
        forward the packet according to the matched entry;
      ELSE
        forward the packet according to the backup nexthop;
    ELSE // there is no FIB entry for forwarding the packet
      IF NH = SRH && SL != 0 THEN
        SL decreases; update the IPv6 DA with SRH[SL];
        FIB lookup on the updated DA;
        forward the packet according to the matched entry;
      ELSE
        drop the packet;
  ELSE
    forward accordingly to the matched entry;
```

5. Determining whether the Endpoint could Be Bypassed

SRv6 Midpoint Protection provides a mechanism to bypass a failed endpoint. But in some scenarios, some important functions may be implemented in the bypassed failed endpoints that should not be bypassed, such as firewall functionality or In-situ Flow Information Telemetry of a specified path. Therefore, a mechanism is needed to indicate whether an endpoint can be bypassed or not. [I-D.li-rtgwg-enhanced-ti-lfa] provides method to determine whether enbale SRv6 midpoint protection or not by defining a "no bypass" flag for the SIDs in IGP.

6. Security Considerations

This section reviews security considerations related to SRv6 Midpoint protection processing discussed in this document.To ensure that the Repair node does not modify the SRH header Encapsulated by nodes outside the SRv6 Domain.Only the segment within the SRH is same domain as the repair node. So it is necessary to check the skipped segment have same block as repair node.

7. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

8. Acknowledgements

9. References

9.1. Normative References

[I-D.hu-spring-segment-routing-proxy-forwarding]
Hu, Z., Chen, H., Yao, J., Bowers, C., and Y. Zhu, "SR-TE Path Midpoint Protection", draft-hu-spring-segment-routing-proxy-forwarding-11 (work in progress), August 2020.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-25 (work in progress), May 2019.

[I-D.ietf-lsr-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extension to Support Segment Routing over IPv6 Dataplane", draft-ietf-lsr-isis-srv6-extensions-11 (work in progress), October 2020.

[I-D.ietf-lsr-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", draft-ietf-lsr-ospfv3-srv6-extensions-01 (work in progress), August 2020.

[I-D.ietf-ospf-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-27 (work in progress), December 2018.

[I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-ietf-spring-srv6-network-programming-24 (work in progress), October 2020.

[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>.

[RFC7356] Ginsberg, L., Previdi, S., and Y. Yang, "IS-IS Flooding Scope Link State PDUs (LSPs)", RFC 7356, DOI 10.17487/RFC7356, September 2014, <https://www.rfc-editor.org/info/rfc7356>.

9.2. Informative References

[I-D.hegde-spring-node-protection-for-sr-te-paths]
Hegde, S., Bowers, C., Litkowski, S., Xu, X., and F. Xu, "Node Protection for SR-TE Paths", draft-hegde-spring-node-protection-for-sr-te-paths-07 (work in progress), July 2020.

[I-D.ietf-rtgwg-segment-routing-ti-lfa]
Litkowski, S., Bashandy, A., Filsfils, C., Decraene, B., Francois, P., Voyer, D., Clad, F., and P. Camarillo, "Topology Independent Fast Reroute using Segment Routing", draft-ietf-rtgwg-segment-routing-ti-lfa-04 (work in progress), August 2020.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-08 (work in progress), July 2020.

[I-D.li-rtgwg-enhanced-ti-lfa]
Li, C. and Z. Hu, "Enhanced Topology Independent Loop-free Alternate Fast Re-route", draft-li-rtgwg-enhanced-ti-lfa-02 (work in progress), August 2020.

[I-D.sivabalan-pce-binding-label-sid]
Sivabalan, S., Filsfils, C., Tantsura, J., Hardwick, J., Previdi, S., and C. Li, "Carrying Binding Label/Segment-ID in PCE-based Networks.", draft-sivabalan-pce-binding-label-sid-07 (work in progress), July 2019.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

Authors' Addresses

Huanan Chen
China Telecom
109, West Zhongshan Road, Tianhe District
Guangzhou 510000
China

Email: chenhuan6@chinatelecom.cn

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

Email: huzhibo@huawei.com

Huaimo Chen
Futurewei
Boston, MA
USA

Email: Huaimo.chen@futurewei.com

Xuesong Geng
Huawei Technologies

Email: gengxuesong@huawei.com