Workgroup: Network Working Group
Internet-Draft:
draft-chen-rtgwg-srv6-midpoint-protection-07
Published: 23 June 2022
Intended Status: Experimental
Expires: 25 December 2022

Authors: H. Chen, China Telecom; Z. Hu, Huawei Technologies; H. Chen, Futurewei; X. Geng, Huawei Technologies; Y. Liu, China Mobile; G. Mishra, Verizon Inc.

SRv6 Midpoint Protection

Abstract

The current local repair mechanism, e.g., TI-LFA, allows local repair actions on the direct neighbors of the failed node or link to temporarily route traffic to the destination. This mechanism could not work properly when the failure happens in the destination point. In SRv6 TE, the IPv6 destination address in the outer IPv6 header could be the segment endpoint of the TE path rather than the destination of the TE path. When the SRv6 endpoint fails, local repair couldn't work on the direct neighbor of the failed endpoint either. This document defines midpoint protection for SRv6 TE path, which enables other nodes on the network to perform endpoint behaviors instead of the faulty node, Update 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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction
2. SRv6 Midpoint Protection Mechanism
3. SRv6 Midpoint Protection Example
4. SRv6 Midpoint Protection Behavior
  4.1. Transit Node as Repair Node
  4.2. Endpoint Node as Repair Node
  4.3. Endpoint x Node as Repair Node
5. Determining whether the Endpoint could Be Bypassed
6. Security Considerations
7. IANA Considerations
8. Acknowledgements
9. References
  9.1. Normative References
  9.2. Informative References
Authors' Addresses

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 or link to temporarily route traffic to the destination. This mechanism could not work properly when the failure happens in the destination point. In SRv6 TE, the IPv6 destination address in the outer IPv6 header could be the segment endpoint of the TE path rather than the destination of the TE path ([RFC8986]). When the endpoint fails, local repair couldn't work on the direct neighbor of the failed endpoint either. This document defines midpoint protection for SRv6 TE path, which enables other nodes on the network to perform endpoint behaviors instead of the faulty node, Update 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. There are two stages or time periods after an endpoint node fails. The first is the time period from the failure until the IGP converges on the failure. The second is the time period after the IGP converges on the failure.

During the first time period, the packet will be sent to the direct neighbor of the failed endpoint node. After detecting the failure of its interface to the failed endpoint node, the neighbor forwards the packets around the failed endpoint node. It changes the IPv6 destination address with the IPv6 address of the next endpoint node (or the last or other reasonable endpoint node) which could avoid going through the failed endpoint.

During the second time period, the packet of a SRv6 TE path may not be sent to the direct neighbor of the failed endpoint node. There is no route to the failed endpoint node after the IGP converges. When a previous hop node of the failed endpoint node finds out that there is no route to the IPv6 destination address (of the failed endpoint node), it changes the IPv6 destination address with the IPv6 address of the next endpoint node. Note that the previous hop node may not be the direct neighbor of the failed endpoint node.

3. SRv6 Midpoint Protection Example

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

```
                     +------+          +------+
                     |  N5  |----------|  N6  |-------------+
                     +------+          +------+             |
                        |                 |                 |
                        |                 |                 |
                        |                 |                 |
   +------+          +------+          +------+          +------+
   |  N1  |----------|  N2  |----------|  N3  |----------|  N4  |
   +------+          +------+          +------+          +------+
```

Figure 1: An example of network for 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, suppose that Node N1 is an ingress node of SRv6 TE path going through N3 and N4. 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:

During the first time period (i.e., before the IGP converges), node N2 (direct neighbor of N3) as a Repair Node forwards the packets around the failed endpoint N3 after detecting the failure of the outbound interface to the endpoint B:3. It changes 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 repair 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.

During the second time period (i.e., after the IGP converges), node N1 does not have any route to the failed endpoint N3 in its FIB. Node N1, as a Repair Node, forwards the packets around the failed endpoint N3 to the next 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

A node N protecting the failure of an endpoint node on a SRv6 path may be one of the following types:

*a transit node: the destination address (DA) of the packet received by N is not N's local SID.

*an endpoint node: the destination address (DA) of the packet received by N is a N's local END SID.

*an endpoint x node (i.e., an endpoint with cross-connect node): the destination address (DA) of the packet received by N is a N's local End.X SID with an array of layer 3 adjacencies.

This section describes the behavior of each of these nodes as a repair node for the two time periods after the endpoint node fails.

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 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 IF there is no FIB entry for forwarding the packet THEN
  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;
```

*: SL could be decreased by any dedicated value from [1-N], where N is the current value of SL.

4.2. Endpoint Node as Repair 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 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 IF there is no FIB entry for forwarding the packet THEN
  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

When the Repair Node is an endpoint x node, it 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 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 IF there is no FIB entry for forwarding the packet THEN
    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.ietf-lsr-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Segment Routing over IPv6 Dataplane", Work in Progress, Internet-Draft, draft-ietf-lsr-isis-srv6-extensions-18, 20 October 2021, <https://www.ietf.org/archive/id/draft-ietf-lsr-isis-srv6-extensions-18.txt>.

[I-D.ietf-lsr-ospfv3-srv6-extensions] Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", Work in Progress, Internet-Draft, draft-ietf-lsr-ospfv3-srv6-extensions-03, 19 November 2021, <https://www.ietf.org/archive/id/draft-ietf-lsr-ospfv3-srv6-extensions-03.txt>.

[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>.

[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>.

9.2. Informative References

[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-19, 11 April 2022, <https://www.ietf.org/archive/id/draft-hu-spring-segment-routing-proxy-forwarding-19.txt>.

[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-08.txt>.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", Work in Progress, Internet-Draft, draft-ietf-spring-segment-routing-policy-22, 22 March 2022, <https://www.ietf.org/archive/id/draft-ietf-spring-segment-routing-policy-22.txt>.

[I-D.li-rtgwg-enhanced-ti-lfa] Li, C., Hu, Z., Zhu, Y., and S. Hegde, "Enhanced Topology Independent Loop-free Alternate Fast Re-route", Work in Progress, Internet-Draft, draft-li-rtgwg-enhanced-ti-lfa-06, 21 April 2022, <https://www.ietf.org/archive/id/draft-li-rtgwg-enhanced-ti-lfa-06.txt>.

[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.", Work in Progress, Internet-Draft, draft-sivabalan-pce-binding-label-sid-07, 8 July 2019, <https://www.ietf.org/archive/id/draft-sivabalan-pce-binding-label-sid-07.txt>.

[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,
United States of America

Email: Huaimo.chen@futurewei.com

Xuesong Geng
Huawei Technologies

Email: gengxuesong@huawei.com

Yisong Liu
China Mobile

Email: liuyisong@chinamobile.com

Gyan S. Mishra
Verizon Inc.
13101 Columbia Pike
Silver Spring, MD 20904
United States of America

Phone: 301 502-1347
Email: gyan.s.mishra@verizon.com