Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 4 August 2024

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

1 February 2024

SRv6 Midpoint Protection
draft-chen-rtgwg-srv6-midpoint-protection-14

Abstract

The current local repair mechanism, e.g., TI-LFA, allows the upstream neighbor of the failed node or link to fast re-route traffic around the failure. This mechanism does not work properly for SRv6 TE path after the failure happens in an endpoint node and IGP converges on the failure. This document defines midpoint protection for SRv6 TE path, which enables the upstream endpoint node of the failed node to perform the endpoint behavior for the faulty node and fast re-route traffic around the failure after IGP converges on the failure.

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

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/.

Copyright Notice

Table of Contents

1. Introduction

The current local repair mechanism, e.g., Topology-Independent Loop-Free Alternate (TI-LFA) ([I-D.ietf-rtgwg-segment-routing-ti-lfa]), allows the upstream neighbor of the failed node or link to fast re-route traffic around the failure. This mechanism does not work properly after the failure happens in an endpoint node and IGP converges on the failure.

In SRv6, the IPv6 destination address (DA) in the outer IPv6 header could be a segment endpoint node (or endpoint for short) of an SRv6 TE path (or SRv6 path for short) rather than the destination of the

SRv6 path ([RFC8986]). After the endpoint fails and IGP converges on the failure, the packet with the failed endpoint as DA will be dropped since there is no FIB entry for DA (i.e., no route to this endpoint). The upstream non-endpoint neighbor of the failed endpoint will not receive the packet for the SRv6 path. [I-D.ietf-spring-segment-protection-sr-te-paths] and [I-D.hu-spring-segment-routing-proxy-forwarding] propose midpoint protection for SR-MPLS TE path after IGP converges on the failure of a node along the path.

This document defines midpoint protection for SRv6 path after IGP converges on the failure of an endpoint on the path, which enables the upstream endpoint of the failed endpoint to perform the endpoint behavior for the failed endpoint and fast re-route traffic around the failure after IGP converges on the failure.

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. Only endpoint node can process SRH, so, only endpoint nodes can perform midpoint protection. 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 upstream neighbor of the failed endpoint node. After detecting the failure of its interface to the failed endpoint node, the neighbor forwards the packet around the failed endpoint node using TI-LFA.

During the second time period, there is no FIB entry for the failed endpoint. When a upstream/previous endpoint of the failed endpoint has no FIB entry for the failed endpoint, it changes the DA of the packet to the IPv6 address of the next endpoint (of the failed endpoint) and forwards the packet using the FIB entry for the next endpoint. Note that the upstream/previous endpoint node may not be the upstream neighbor of the failed endpoint.

3. SRv6 Midpoint Protection Example

Figure 1 illustrates an example of network topology with SRv6 enabled on each node. The cost of each link is 1 by default, except for the costs of the links indicated by numbers on the links.

In this document, an end SID at node Ni with locator block B is represented as B:i. 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.

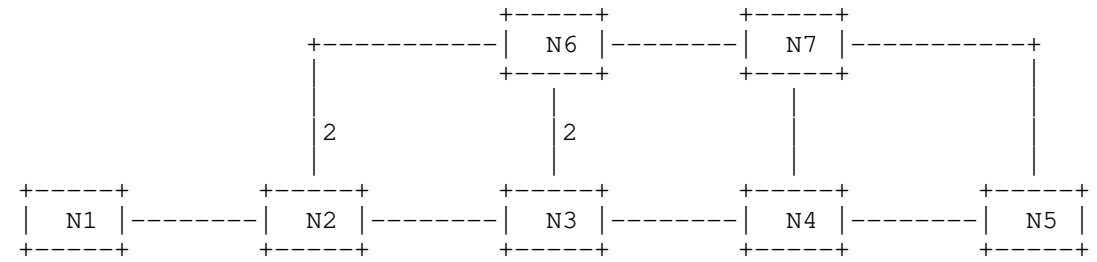

Figure 1: An example of network for midpoint protection

In the reference topology, suppose that there are two SRv6 paths having node N1 as ingress. The first path is from N1 through endpoint nodes N4 and N5, which is represented by SID list <B:4, B:5>. The second path is from N1 through endpoint nodes N2, N4 and N5, which is represented by SID list <B:2, B:4, B:5>. For a packet to be transported by the first path, N1 encapsulates the packet with <B:4, B:5>. For a packet to be transported by the second path, N1 encapsulates the packet with <B:2, B:4, B:5>.

When N4 fails, the packet on each of the two paths needs to bypass the failed endpoint N4 and be forwarded to the next endpoint N5 after the failed endpoint.

During the first time period (i.e., after N4 fails and before IGP converges on the failure), N3 (upstream neighbor of N4) as a Repair Node receives the packet for each of the two SRv6 paths. It forwards the packet around the failed endpoint N4 after detecting the failure of the outbound interface/link to the endpoint B:4. It uses the TI-LFA to forward the packet through encapsulating the packet with SID list <B:6, B:7> as a TI-LFA repair path.

During the second time period (i.e., after N4 fails and after IGP converges on the failure):

* For the first path, N3 (upstream endpoint neighbor of N4) as a Repair Node receives the packet, N3 has no FIB entry for the failed endpoint N4. N3 forwards the packet around the failed endpoint N4 to the next endpoint (e.g., N5) using the FIB entry for the next endpoint. N3 changes the DA of the packet to the next SID B:5 and forwards the packet using the FIB entry for DA = B:5 (i.e., using IGP SPF path to B:5).

* For the second path, N3 (upstream non-endpoint neighbor of N4) will not receive any packet for the path. The upstream endpoint N2 of the failed endpoint N4 will not send any packet for the path to N3. N2 has no FIB entry for the failed N4. N2, as a Repair Node, sends the packet around N4 to the next endpoint (e.g., N5) using the FIB entry for the next endpoint. N2 changes the DA of the packet to the next SID B:5 and sends the packet using the FIB entry for DA = B:5 (i.e., using IGP SPF path to B:5).

4. SRv6 Midpoint Protection Behavior

Figure 2 shows the procedure of a upstream (endpoint) node of an endpoint node on an SRv6 path for midpoint protection in pseudo code.

When the endpoint (e.g., N4) fails and before IGP converges on the failure (i.e., in the first period), if the upstream node (e.g., neighbor N3) of the failed endpoint detects the failure of the link used to send the packet for the path and the FIB entry for the DA of the packet exists, then it uses TI-LFA to fast re-route the packet around the failure (refer to line 1 and 2 of the procedure); otherwise, the link used to send the packet works (i.e., no failure) or no FIB entry for DA of the packet (i.e., after the failure and IGP converges on the failure, or say in the second period).

If the upstream node (e.g., endpoint N2 for the second path) has no FIB entry for the DA of the packet, then it changes the DA to the next SID and sends the packet using the FIB entry for DA = next SID when the packet has a SRH with SIDs as a next header (refer to lines 3 to 6). When the packet has no SRH with SIDs as a next header, the packet is dropped (refer to line 7 to 8).

When the upstream node has a FIB entry for the DA of the packet (i.e., no failure or there is a failure and before IGP converges on the failure which is in the first time period), it sends the packet using the FIB entry for the DA (refer to line 9 to A).

```
 1: IF link for sending packet failed and FIB entry for DA exists THEN
 2:   use TI-LFA to re-route packet around failure; // in 1st period
 3: ELSE IF no FIB entry for DA of packet (i.e., in 2nd period) THEN
 4:   IF NH = SRH && SL != 0 THEN // if next header is SRH with SIDs
 5:     SL--; DA = SRH[SL]; // change DA of packet to next SID/endpoint
 6:     forward packet using FIB entry for DA;//send packet to next SID
 7:   ELSE // next header is not SRH with SIDs
 8:     drop the packet;
 9: ELSE //has FIB entry for DA of packet(i.e.,normal or in 1st period)
 A:   forward packet accordingly to FIB entry for DA;
```

Figure 2: Procedure of Upstream Node for Midpoint Protection

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 enable SRv6 midpoint protection or not by defining a "no bypass" flag for the SIDs in IGP.

6. Security Considerations

To ensure that the Repair node does not modify the SRH header Encapsulated by nodes outside the SRv6 Domain, the segment within the SRH needs to be in the same domain as the repair node. So it is necessary to check the skipped segment has the same block as the 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. References

8.1. Normative References

[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>.

[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>.

8.2. Informative References

[I-D.hu-spring-segment-routing-proxy-forwarding]
Hu, Z., Chen, H., Yao, J., Bowers, C., Zhu, Y., and Y. Liu, "SR-TE Path Midpoint Restoration", Work in Progress, Internet-Draft, draft-hu-spring-segment-routing-proxy-forwarding-24, 21 August 2023, <https://datatracker.ietf.org/doc/html/draft-hu-spring-segment-routing-proxy-forwarding-24>.

[I-D.ietf-rtgwg-segment-routing-ti-lfa]
Bashandy, A., Litkowski, S., 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-13, 16 January 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-segment-routing-ti-lfa-13>.

[I-D.ietf-spring-segment-protection-sr-te-paths]
Hegde, S., Bowers, C., Litkowski, S., Xu, X., and F. Xu, "Segment Protection for SR-TE Paths", Work in Progress, Internet-Draft, draft-ietf-spring-segment-protection-sr-te-paths-05, 27 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-segment-protection-sr-te-paths-05>.

[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-09, 19 October 2023, <https://datatracker.ietf.org/doc/html/draft-li-rtgwg-enhanced-ti-lfa-09>.

Acknowledgments

The authors would like to thank Bruno Decraene, Jeff Tantsura, Ketan Talaulikar, Yingzhen Qu and Parag Kaneriya for their comments to this work.

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: hchen.ietf@gmail.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