IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 21, 2025

Y. Liu
China Mobile
C. Lin
New H3C Technologies
Ran.Chen
ZTE

February 21, 2025

BGP Extension for SRv6 Policy Segment List optimization

draft-liu-idr-sr-segment-list-optimize-01

Abstract

In some use cases, an SRv6 policy’s segment list ends with the policy endpoint’s node SID, and the traffic steered (over policy) already ensures that it is taken to the policy endpoint. In such cases, the SID list can be optimized by excluding the endpoint Node SID when installing the policy.

This document specifies a BGP extension to indicate whether the endpoint’s node SID needs to be included or excluded when installing the SRv6 Policy. This optimization can improve the forwarding efficiency of data packets when End SID and Service SID are present.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on August 21 2025.

Table of Contents

1. Introduction

Segment Routing (SR) [RFC8402] allows a node to steer a packet flow along any path. A Segment Routing Policy (SR Policy) [RFC8402] is an ordered list of segments that represent a source-routed policy. The headend node is said to steer a flow into an SR Policy. The packets steered into an SR Policy have an ordered list of segments associated with that SR Policy written into them. Segment Routing Policy Architecture [RFC9256] updates [RFC8402] as it details the concepts of SR Policy and steering into an SR Policy. [RFC8986] describes the representation and processing of this ordered list of segments for Segment Routing over IPv6 (SRv6). [I-D. draft-ietf-

idr-sr-policy-safi] document specifies how BGP may distribute SR Policy candidate paths. [I.D.draft-ietf-idr-bgp-ls-sr-policy] defines a mechanism to collect the Segment Routing Policy information that is locally available in a node and advertise it into BGP Link-State (BGP-LS) updates.

This document introduces an optimization method for segment list arrangement to solve the problem of the penultimate segment node being unable to perform PSP behavior when the egress node has both End SID and service SID, and improve the forwarding efficiency of data packets.

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

PSP: Penultimate Segment Pop

3. Motivation

In SRv6 networks, some functions can only be executed on the penultimate SR Segment Endpoint Node, such as Penultimate Segment Pop (PSP) behavior. However, if both the End SID and service SID of the egress node are encapsulated in SRH.SegmentList, the endpoint will not be able to identify itself as the penultimate SR Segment Endpoint Node based on the SRH.SL field after receiving the packet.

For example, in the following scenarios, the Segment List of SRv6 Policy must include the End SID of the egress node. The SRH extension header of VPN user’s data packets forwarded based on this SRv6 Policy tunnel will simultaneously encapsulate the End SID and VPN SID of the egress node.

Scenario 1

In tunnel splicing scenarios and cross domain path splicing scenarios, usually based on binding SID to steer traffic. The Segment List of SRv6 Policy on the head node must include the End SID of the egress node.

Scenario 2

When the head node enables end-to-end fast fault detection of SRv6 Policy, OAM messages are sent to the egress node. The End SID of the egress node must be specified in the Segment List of this SRv6 Policy.

In this way, the following two problems will arise:

Problem 1: PSP behavior may not be executable.

If the head node encapsulates both the End SID and VPN SID of the egress node in the SRH.SegmentList, the penultimate SR Segment Endpoint Node will find that local SID is not in the position with SL=1 after receiving the packet.

After executing SL--, SL is still greater than 0. Because the condition of (SL==0) is not met, the penultimate SR Segment Endpoint Node will not be able to perform the processing of removing the SRH from the IPv6 extension header.

Problem 2: The forwarding efficiency of egress node decreases.

If the egress node receives a packet with both a local End SID and a VPN SID, it needs to first look up the table based on the End SID. Then, based on the VPN SID, execute the VPN SID instruction, and finally remove the outer IPv6 packet header and forward it to VPN network.

The data packet needs to look up the SID table twice within the egress node. For some chips, the second SID table lookup requires a loopback interface to be implemented. Due to the bandwidth limitations and the possibility of other service packets coexisting on the loopback interface, the forwarding efficiency of packets to VPN will be greatly affected.

Problem 3: Increase the overhead of the packet header.

Carrying both the End SID and VPN SID of the egress node in the SRH.SegmentList will increase the overhead of the packet header. Especially in environments that require SRv6 header compression, arranging End SID for egress node will reduce compression efficiency.

Therefore, this document proposes a method to optimize the SRH.SegmentList encapsulated by the head node. When there are End SID and service SID of egress node on the path at the same time, only the service SID is encapsulated in the SRH.SegmentList.

This can solve the problem of the penultimate segment node being unable to perform PSP behavior when the egress node has both End SID and service SID, and improve the forwarding efficiency of data packets on the egress node.

4. Extension

N-flag (endpoint node SID inclusion flag) is proposed in the Candidate Path Administrative Flags Sub-TLV specified in [I.D-draft-lin-idr-sr-policy-admin-flags].The bit position for the flag is to be defined by IANA.

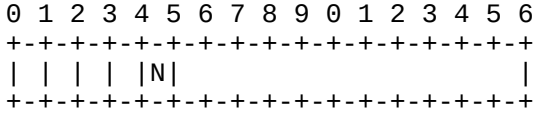

Figure 1: Administrative Flags

where:

- N-Flag(Bit TBD): indicate the endpoint node SID is included in installing SID list(s) of the Candidate Path (CP) when set.
  - * If set to 1, the endpoint node SID MUST be included when installing the SR Policy SID list(s) used to carry the data traffic.
  - * If set to 0, the endpoint node SID MUST NOT be included when installing the SR Policy SID list(s) used to carry the data traffic.

5. Operation

When the controller distributes the SRv6 Policy configuration to the head node through BGP, set the N-flag (endpoint node SID inclusion flag) is proposed in the Candidate Path Administrative Flags Sub-TLV specified in [I.D-draft-lin-idr-sr-policy-admin-flags].

After receiving the SRv6 Policy configuration with the N-Flag of the Candidate Path Administrative Flags, the ingress node will not simultaneously arrange the End SID and Service SID of the egress node into the SRH.SegmentList of packet.

For data packets forwarded to VPN through this SRv6 Policy, the SRH.SegmentList will not encapsulate the End SID corresponding to the egress node in the SID list of SRv6 Policy.

If the forwarding path does not include the service SID of the egress node, then the End SID of the egress node should be encapsulated in SRH.SegmentList.

For OAM detection packets of the SR policy, the SRH.SegmentList is encapsulated according to the SID list of the SR policy, only encapsulating node SIDs.

6. Use Case

Taking Figure 1 as an example, describe how SRv6 data packets and OAM packets are forwarded in the SRv6 network based on the optimized Segment List arrangement mechanism.

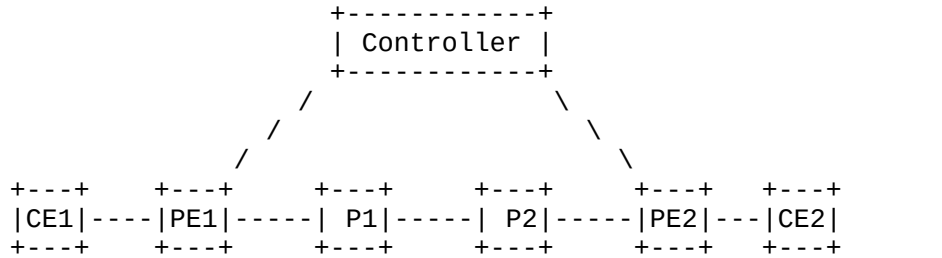

Figure 1

CE1 and CE2 are VPN access devices that connect to the IPv6 backbone network through PE. PE1 has a locator 2::/64. P1 has a locator 3::/64. P2 has an End SID 4::4 with PSP Flavor. PE2 has a locator 5::/64 and a VPN SID 5::100. The traffic from CE1 to CE2 is forwarded along the path PE1->P1->P2->PE2.

P2 needs to perform the PSP behavior to remove the SRH extension header.

The controller calculates the SRv6 forwarding path from PE1 to PE2 based on the collected topology and configuration information, and distributes the SRv6 Policy to PE1 through BGP. The Endpoint address is 5::5 of PE2. There is only one candidate path. The candidate path contains a Segment list <3::3, 4::4, 5::5> with N-Flag set in in the Candidate Path Administrative Flags Sub-TLV.

PE2 advertises a BGP VPN route to PE1, and the next hop of the BGP route is the endpoint address 5::5. After receiving the BGP route, PE1 iterates to the SRv6 Policy using the color and the next hop of the route.

There are two types of packets sent from PE1 to PE2: data packets and OAM packets.

6.1. Data packet Processing to VPN

After PE1 receives the data packet from CE1 to CE2, it looks up the VPN instance routing table and iterates to SRv6 Policy.

PE1 adds the SRH extension header to the packet and encapsulates the Segment List of the SRv6 Policy. The Segment List in the SRH extension header is encapsulated as <3::3, 4::4, 5::100>, and the SL is set to 2.

The Segment List in SRH is shown in Figure 2.

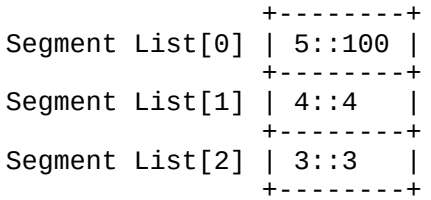

Figure 2

The segment list optimization method proposed in this document is suitable for both SRv6 SID compressed and non-compressed scenarios. If the END SID and VPN SID of the egress node share a common Locator-Block with a sequence of consecutive nodes, the SIDs of the egress node can also be arranged in a compressed Segment List.

In order to improve compression efficiency and reduce the overhead of SRv6 packet header, the compressed Segment List can only contain the compressed VPN SID.

As shown in Figure 3, PE1, P1, P2, and PE3 share the common Locator-block A:0:0:0/64 (represented by LB in Figure3).

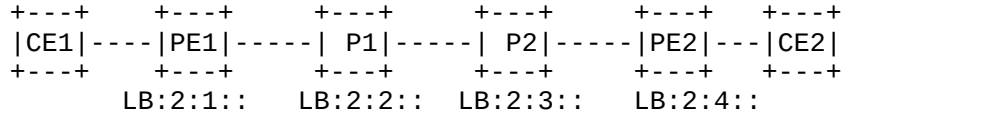

Figure 3

The compressed Segment List optimized in SRH is shown in Figure 4.

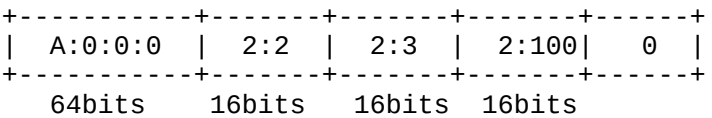

Figure 4

6.2. OAM Packet Processing to the Egress Node

If the head node enables OAM function and detects a fault in the SRv6 Policy forwarding path, PE1 will send OAM detection messages to PE2, such as BFD packets.

The OAM detection message sends by PE1 encapsulate the segment list corresponding to the SRv6 Policy. Since the message does not need to be sent to VPN, the Segment List of the SRH extension header is encapsulated as <3::3, 4::4, 5::5>.

The Segment List in SRH is shown in Figure 5.

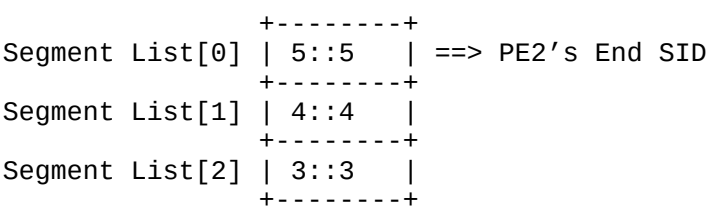

Figure 5

7. IANA Considerations

N-flag (endpoint node SID inclusion flag) is proposed in the Candidate Path Administrative Flags Sub-TLV specified in [I.D-draft-lin-idr-sr-policy-admin-flags].

- N-Flag(Bit TBD): indicate the endpoint node SID is included in installing SID list(s) of the Candidate Path (CP) when set.

* If set to 1, the endpoint node SID MUST be included when installing the SR Policy SID list(s) used to carry the data traffic.

* If set to 0, the endpoint node SID MUST NOT be included when installing the SR Policy SID list(s) used to carry the data traffic.

8. 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. 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].

9. References

9.1. Normative References

[I-D.draft-ietf-idr-sr-policy-safi]Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., and D. Jain, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-sr-policy-safi-06, 26 July 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-sr-policy-safi-06>.

[I-D.draft-ietf-idr-bgp-sr-segtypes-ext] Talaulikar, K., Filsfils, C., Previdi, S., Mattes, P., and D. Jain, "Segment Routing Segment Types Extensions for BGP SR Policy", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-sr-segtypes-ext-04, 30 July 2024,<https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-sr-segtypes-ext-04>.

[I-D.draft-lin-idr-sr-policy-admin-flags] C, Lin, "BGP SR Policy Extensions for Administrative Flags ", Work in Progress, Internet-Draft, draft-lin-idr-sr-policy-admin-flags-00, 30 July 2024,<https://datatracker.ietf.org/doc/html/ draft-lin-idr-sr-policy-admin-flags-00>.

[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>.

9.2. Informative References

TBD

10. Acknowledgments

TBD

Authors' Addresses

Yisong Liu
China Mobile

Email: liuyisong@chinamobile.com

Changwang Lin
New H3C Technologies

Email: linchangwang.04414@h3c.com

Ran Chen
ZTE Corporation

Email: chen.ran@zte.com.cn