SRv6 Path Egress Protection
draft-hu-rtgwg-srv6-egress-protection-00

Abstract

This document describes protocol extensions and procedures for protecting the egress node of a Segment Routing for IPv6 (SRv6) path.

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 September 6, 2019.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect
1. Introduction

Fast protection of a transit node of a Segment Routing (SR) path is described in [I-D.bashandy-rtgwg-segment-routing-ti-lfa] and [I-D.hu-spring-segment-routing-proxy-forwarding]. However, these documents do not discuss the procedures for fast protection of the egress node of a Segment Routing for IPv6 (SRv6) path.

This document fills that void and specifies protocol extensions and procedures for fast protection of the egress node of an SRv6 path. Egress node and egress as well as fast protection and protection will be used exchangeably.

2. Terminologies

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
Figure 1 shows an example of protecting egress PE3 of a SR path, which is from ingress PE1 to egress PE3.

Locator: A3:1::/64

***** ***** VPN SID: A3:1::B100

[PE1]-----[P1]-----[PE3]
/ | & | \ PE3 Egress
/ | & | \ CEx Customer Edge
[CE1] | & | [CE2] Px Non-Egress
\ | & | / *** SR Path
\ | & & & & & / & & & Backup Path
[PE2]-----[P2]-----[PE4]

Locator: A4:1::/64
VPN SID: A4:1::B100
Mirror SID: A4:1::3, protect A3:1::/64

Figure 1: Protecting SR Path Egress PE3
Node P1’s pre-computed TI-LFA backup path for PE3 is from P1 to PE4 via P2. In normal operations, after receiving a packet with destination PE3, P1 forwards the packet to PE3 according to its FIB. When PE3 receives the packet, it sends the packet to CE2.

When PE3 fails, P1 detects the failure through BFD and forwards the packet to PE4 via the backup path. When PE4 receives the packet, it sends the packet to the same CE2.

In Figure 1, CE2 is dual home to PE3 and PE4. PE3 has a locator A3:1::/64 and a VPN SID A3:1::B100. PE4 has a locator A4:1::/64 and a VPN SID A4:1::B100. A mirror SID A4:1::3 is configured on PE4 for protecting PE3 with locator A3:1::/64.

After the mirror SID is configured on a local PE (e.g., PE4), when the local PE (e.g., BGP on the local PE) receives a prefix whose VPN SID belongs to a remote PE (e.g., PE3) with the locator that is protected by the local PE through mirror SID, the local PE (e.g., PE4) creates a mapping from the remote PE’s (e.g., PE3’s) VPN SID and the mirror SID to the local PE’s (e.g., PE4’s) VPN SID. The remote PE is protected by the local PE.

For example, local PE4 has Prefix 1.1.1.1 with VPN SID:A4:1::B100, when PE4 receives prefix 1.1.1.1 with remote PE3’s VPN SID A3:1::B100, it creates a mapping from remote PE3’s VPN SID and the mirror SID (i.e., "A3:1::B100, A4:1::3") to local PE4’s VPN SID (i.e., "A4:1::B100").

Node P1’s pre-computed TI-LFA backup path for destination PE3 having locator A3:1::/64 is from P1 to PE4 having mirror SID A4:1::3. It is installed as a T.Insert transit behavior. When P1 receives a packet destined to PE3’s VPN SID A3:1::B100, in normal operations, it forwards the packet with source A1:1:: and destination PE3’s VPN SID A3:1::B100 according to the FIB using the destination PE3’s VPN SID A3:1::B100.

When PE3 fails, node P1 protects PE3 through sending the packet to PE4 via the backup path pre-computed. P1 modifies the packet before sending it to PE4. The modified packet has destination PE4 with mirror SID A4:1::3, and SRH with PE3’s VPN SID A3:1::B100 and the mirror SID A4:1::3 (i.e., "A3:1::B100, A4:1::3; SL=1").

When PE4 receives the packet, it forwards the packet to CE2 through executing END.M instruction according to the local VPN SID (i.e., A4:1::B100).
4. Extensions to IGP for Egress Protection

This section describes extensions to IS-IS and OSPF for advertising the information about SRv6 path egress protection.

4.1. Extensions to IS-IS

A new sub-TLV, called IS-IS SRv6 End.m SID sub-TLV, is defined. It is used in the SRv6 Locator TLV defined in [I-D.bashandy-isis-srv6-extensions] to advertise SRv6 Segment Identifiers (SIDs) with END.M function for SRv6 path egress protection. The SRv6 End.m SIDs inherit the topology/algorithim from the parent locator. The format of the sub-TLV is illustrated below.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type (TBD1)   |    Length     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Flags     |    SRv6 Endpoint Function     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         SID (16 octets)                       |
| :                      |                                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            sub-TLVs                           |
| :                      |                                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: IS-IS SRv6 End.m SID sub-TLV

Type:  TBD1 (suggested value 8) is to be assigned by IANA.

Length:  variable.

Flags:  1 octet. No flags are currently defined.

SRv6 Endpoint Function:  2 octets. Add a new endpoint function 40 for end.m SID.

SID:  16 octets. This field contains the SRv6 end.m SID to be advertised.

Two sub-TLVs are defined. One is the protected locators sub-TLV, and the other is the protected SIDs sub-TLV.

A protected locators sub-TLV is used to carry the Locators to be protected by the SRv6 mirror SID. It has the following format.
A protected SIDs sub-TLV is used to carry the SIDs to be protected by the SRv6 mirror SID. It has the following format.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type (TBD3) | Length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| SID (16 octets) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| SID (16 octets) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: IS-IS Protected SIDs sub-TLV

Type: TBD3 (suggested value 2) is to be assigned by IANA.

Length: variable.

SID: 16 octets. This field encodes an SRv6 SID to be advertised.
4.2. Extensions to OSPF

Similarly, a new sub-TLV, called OSPF SRv6 End.m SID sub-TLV, is defined. It is used to advertise SRv6 Segment Identifiers (SIDs) with END.M function for SRv6 path egress protection. Its format is illustrated below.

```
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Type (TBD4)           |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Flags     |    SRv6 Endpoint Function     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         SID (16 octets)                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       sub-TLVs                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: OSPF SRv6 End.m SID sub-TLV

- **Type**: TBD4 (suggested value 8) is to be assigned by IANA.

- **Length**: variable.

- **Flags**: 1 octet. No flags are currently defined.

- **SRv6 Endpoint Function**: 2 octets. Add a new endpoint function 40 for end.m SID.

- **SID**: 16 octets. This field contains the SRv6 end.m SID to be advertised.

Two sub-TLVs are defined. One is the protected locators sub-TLV, and the other is the protected SIDs sub-TLV.

A protected locators sub-TLV is used to carry the Locators to be protected by the SRv6 mirror SID. It has the following format.
Figure 6: OSPF Protected Locators sub-TLV

Type: TBD5 (suggested value 1) is to be assigned by IANA.

Length: variable.

Locator-Size: 1 octet. Number of bits (1 - 128) in the Locator field.

Locator: 1-16 octets. This field encodes an SRv6 Locator to be protected by the SRv6 mirror SID. The Locator is encoded in the minimal number of octets for the given number of bits.

A protected SIDs sub-TLV is used to carry the SIDs to be protected by the SRv6 mirror SID. It has the following format.

Figure 7: OSPF Protected SIDs sub-TLV

Type: TBD6 (suggested value 2) is to be assigned by IANA.

Length: variable.

SID: 16 octets. This field encodes an SRv6 SID to be advertised.
5. Behavior for SRv6 Mirror SID

The "Endpoint with mirror protection to a vpn SID" function (End.M for short) is a variant of the End function. The End.M is used for SRv6 VPN egress protection. It is described below.

End.M: Mirror protection
When N receives a packet destined to S and S is a local End.M SID, N does:
IF NH=SRH and SL = 1 ;; Ref1
    SL--
    Map to a local VPN SID based on Mirror SID and SRH[SL] ;; Ref1
    forward according to the local VPN SID ;; Ref2
ELSE
    drop the packet

Figure 8: SRv6 Mirror SID Procedure

Ref1: An End.M SID must always be the penultimate SID.

Ref2: The rest forwarding behavior is the same as the corresponding VPN sid.

6. Security Considerations

TBD

7. IANA Considerations

TBD

8. Acknowledgements

TBD

9. References

9.1. Normative References

[I-D.bashandy-isis-srv6-extensions]
[I-D.hu-spring-segment-routing-proxy-forwarding]

[I-D.ietf-isis-segment-routing-extensions]

[I-D.ietf-ospf-segment-routing-extensions]

[I-D.li-ospf-ospfv3-srv6-extensions]


9.2. Informative References

[I-D.bashandy-rtgwg-segment-routing-ti-lfa]

[I-D.hegde-spring-node-protection-for-sr-te-paths]
[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., 
bogdanov@google.com, b., and P. Mattes, "Segment Routing 
Policy Architecture", draft-ietf-spring-segment-routing-
policy-02 (work in progress), October 2018.

[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-06 (work in progress), February 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 

Authors’ Addresses

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

Email: huzhibo@huawei.com

Huaimo Chen
Huawei Technologies
Boston, MA
USA

Email: Huaimo.chen@huawei.com

Peng Wu
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing  100095
China

Email: baggio.wupeng@huawei.com