

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 19, 2020

Z. Hu
Huawei
H. Chen
Futurewei
H. Chen
China Telecom
P. Wu
Huawei
M. Toy
Verizon
C. Cao
T. He
China Unicom
L. Liu
Fujitsu
X. Liu
Volta Networks
March 18, 2020

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

Abstract

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

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 19, 2020.

Copyright Notice

Copyright (c) 2020 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 to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Terminologies	3
3. SR Path Egress Protection	4
3.1. Mechanism	4
3.2. Example	6
4. Extensions to IGP for Egress Protection	8
4.1. Extensions to IS-IS	8
4.2. Extensions to OSPF	10
5. Security Considerations	12
6. IANA Considerations	12
6.1. IS-IS	12
6.2. OSPFv3	12
7. Acknowledgements	13
8. References	13
8.1. Normative References	13
8.2. Informative References	14
Authors' Addresses	15

1. Introduction

The fast protection of a transit node of a Segment Routing (SR) path or tunnel is described in [I-D.ietf-rtgwg-segment-routing-ti-lfa] and [I-D.hu-spring-segment-routing-proxy-forwarding]. [RFC8400] specifies the fast protection of egress node(s) of an MPLS TE LSP tunnel including P2P TE LSP tunnel and P2MP TE LSP tunnel in details. However, these documents do not discuss the fast protection of the egress node of a Segment Routing for IPv6 (SRv6) path or tunnel.

This document fills that void and presents protocol extensions for the fast protection of the egress node of an SRv6 path or tunnel. Egress node and egress, fast protection and protection as well as SRv6 path and SRv6 tunnel will be used exchangeably below.

There are a number of topics related to the egress protection, which include the detection of egress node failure, the relation between egress protection and global repair, and so on. These are discussed in details in [RFC8679].

2. Terminologies

The following terminologies are used in this document.

SR: Segment Routing

SRv6: SR for IPv6

SRH: Segment Routing Header

SID: Segment Identifier

LSA: Link State Advertisement in OSPF

LSP: Label Switched Path in MPLS or Link State Protocol PDU in IS-IS

PDU: Protocol Data Unit

LS: Link State, which is LSA in OSPF or LSP in IS-IS

TE: Traffic Engineering

SA: Source Address

DA: Destination Address

P2MP: Point-to-MultiPoint

P2P: Point-to-Point

CE: Customer Edge

PE: Provider Edge

LFA: Loop-Free Alternate

TI-LFA: Topology Independent LFA

BFD: Bidirectional Forwarding Detection

VPN: Virtual Private Network

L3VPN: Layer 3 VPN

VRF: Virtual Routing and Forwarding

FIB: Forwarding Information Base

PLR: Point of Local Repair

BGP: Border Gateway Protocol

IGP: Interior Gateway Protocol

OSPF: Open Shortest Path First

IS-IS: Intermediate System to Intermediate System

3. SR Path Egress Protection

This section describes the mechanism of SR path egress protection and illustrates it through an example.

3.1. Mechanism

Figure 1 is used to explain the mechanism of SR path egress node protection.

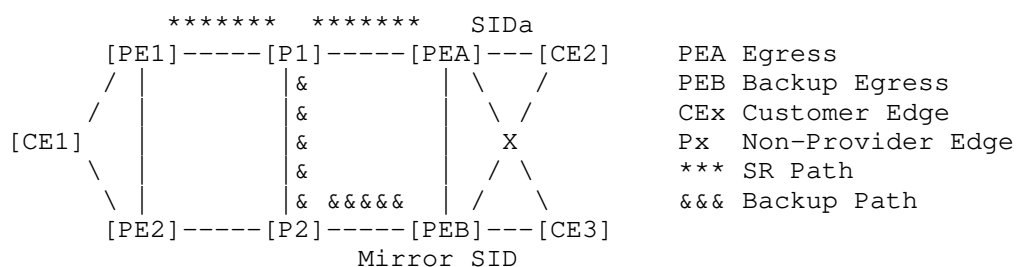


Figure 1: PEB Protects Egress PEA of SR Path

Where node PEA is the egress of the SR path from PE1 to PEA, and has SIDA which is the active segment in the packet from the SR path at PEA. Node PEB is the backup egress (or say protector) to provide the protection for egress (or say primary egress) PEA. Node P1 is the direct previous hop of egress PEA and acts as PLR to support the protection for PEA.

When PEB is selected as a backup egress to protect the egress PEA, a Mirror SID is configured on PEB to protect PEA. PEB advertises this information through IGP, which includes the Mirror SID and the egress PEA. The information is represented by <PEB, PEA, Mirror SID>, which indicates that PEB protects PEA with Mirror SID.

After PEA receives the information <PEB, PEA, Mirror SID>, it may send the forwarding behavior of the SIDA at PEA to PEB with the Mirror SID using some protocols such as BGP if PEB can not obtain this behavior from other approaches if PEB wants to protect SIDA of PEA. How to send the forwarding behavior of the SIDA to PEB is out scope of this document.

When PEB gets the forwarding behavior of the SIDA of PEA from PEA or other means, it adds a forwarding entry for the SIDA according to the behavior into the forwarding table for node PEA. This table is identified by the Mirror SID, which indicates node PEA's context. Using the forwarding entry for SIDA in this table, a packet with SIDA will be transmitted by PEB to the same destination as it is transmitted by PEA. For example, assume that the packet with SIDA is transmitted by PEA to CE2 through the forwarding behavior of the SIDA in PEA. The packet will be transmitted by PEB to the same CE2 through looking up the table identified by the Mirror SID.

After P1 as PLR receives the information <PEB, PEA, Mirror SID> and knows that PEB wants to protect SIDA of PEA, it computes a shortest path to PEB. A Repair List RL is obtained based on the path. It is one of the followings:

- o RL = <Mirror SID> if the path does not go through PEA; or
- o RL = <S1, ..., Sn, Mirror SID> if the path goes through PEA, where <S1, ..., Sn> is the TI-LFA Repair List to PEB computed by P1.

When PEA fails, P1 as PLR sends the packet with SIDA carried by the SR path to PEB, but encapsulates the packet before sending it by executing H.Encaps with the Repair List RL and a Source Address T.

Suppose that the packet received by P1 is represented by Pkt = (S, SIDA)Pkt0, where SA = S and DA = SIDA, and Pkt0 is the rest of the packet.

The execution of H.Encaps pushes an IPv6 header to Pkt and sets some fields in the outer and inner IPv6 header to produce an encapsulated packet Pkt'. Pkt' will be one of the followings:

- o Pkt' = (T, Mirror SID) (S, SIDA)Pkt0 if RL = <Mirror SID>; or

- o $\text{Pkt}' = (T, S1)(\text{Mirror SID}, S_n, \dots, S1; SL=n) (S, \text{SIDa})\text{Pkt0}$ if $RL = \langle S1, \dots, S_n, \text{Mirror SID} \rangle$.

When PEB receives the re-routed packet, which is (T, Mirror SID) (S, SIDA)Pkt0, it decapsulates the packet and forwards the decapsulated packet using the forwarding table identified by Mirror SID through executing End.DT6.

It obtains the Mirror SID in the outer IPv6 header of the packet, removes this outer IPv6 header with all its extension headers, and then processes the inner IPv6 packet (i.e., (S, SIDA)Pkt0, the packet without the outer IPv6 header). PEB finds the forwarding table for node PEA using the Mirror SID as the context ID, and submits the packet to this forwarding table lookup and transmission to the same destination as PEA does.

3.2. Example

Figure 2 shows an example of protecting egress PE3 of a SR path, which is from ingress PE1 to egress PE3.

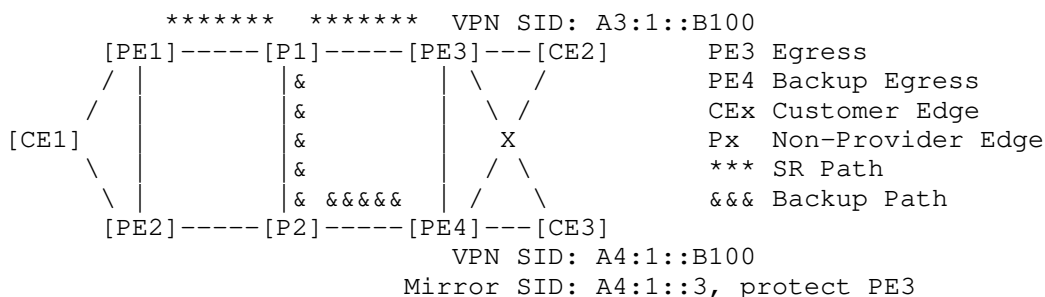


Figure 2: PE4 Protects Egress PE3 of SR Path

Where node P1's pre-computed 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 as PLR detects the failure through using a failure detection mechanism such as 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 2, Both CE2 and CE3 are dual home to PE3 and PE4. PE3 has a VPN SID A3:1::B100. PE4 has a VPN SID A4:1::B100. A Mirror SID A4:1::3 is configured on PE4 for protecting PE3.

After the configuration, PE4 advertises this information through an IGP LS (i.e., LSA in OSPF or LSP in IS-IS), which includes PE4's ID, PE3's ID and Mirror SID A4:1::3. Every node in the SR domain will receive this IGP LS, which indicates that PE4 wants to protect PE3 with Mirror SID A4:1::3.

When PE4 (e.g., BGP on PE4) receives a prefix whose VPN SID belongs to PE3 that is protected by PE4 through Mirror SID A4:1::3, it finds PE4's VPN SID corresponding to PE3's VPN SID. 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 knows that they are for the same VPN.

The forwarding behaviors for these two VPN SIDs are the same from function's point of view. If the behavior for PE3's VPN SID in PE3 forwards the packet with it to CE2, then the behavior for PE4's VPN SID in PE4 forwards the packet to the same CE2; and vice versa. PE4 creates a forwarding entry for PE3's VPN SID A3:1::B100 in the table (or FIB) identified by Mirror SID A4:1::3 according to the forwarding behavior for PE4's VPN SID A4:1::B100.

Node P1's pre-computed backup path for destination PE3 is from P1 to PE4 having mirror SID A4:1::3. 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, P1 as PLR sends the packet to PE4 via the backup path pre-computed. P1 encapsulates the packet using H.Encaps before sending it to PE4.

Suppose that the packet received by P1 is represented by $Pkt = (SA = A1:1::, DA = A3:1::B100)Pkt0$, where $DA = A3:1::B100$ is PE3's VPN SID, and $Pkt0$ is the rest of the packet. The encapsulated packet Pkt' will be one of the followings:

- o $Pkt' = (T, \text{Mirror SID } A4:1::3) (A1:1::, A3:1::B100)Pkt0$ if backup path not via PE3; or (otherwise)
- o $Pkt' = (T, S1) (\text{Mirror SID } A4:1::3, S_n, \dots, S1; SL=n) (A1:1::, A3:1::B100)Pkt0$.

where T is a Source Address, $\langle S1, \dots, S_n \rangle$ is the TI-LFA Repair List to PE4 computed by P1 when the backup path to PE4 goes through PE3.

When PE4 receives the re-routed packet, it decapsulates the packet and forwards the decapsulated packet by End.DT6. The packet received

by PE4 is (T, Mirror SID A4:1::3) (A1:1::, PE3's VPN SID A3:1::B100)Pkt0.

PE4 obtains Mirror SID A4:1::3 in the outer IPv6 header of the packet, removes this outer IPv6 header, and then processes the inner IPv6 packet (A1:1::, A3:1::B100)Pkt0. It finds the forwarding table for PE3 using Mirror SID A4:1::3 as the context ID, gets the forwarding entry for PE3's VPN SID A3:1::B100 from the table, and forwards the packet to CE2 using the entry.

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 Mirror SID sub-TLV, is defined. It is used in the SRv6 Locator TLV defined in [I-D.ietf-lsr-isis-srv6-extensions] to advertise SRv6 Mirror SID and the ID of the node to be protected. The SRv6 Mirror SID inherit the topology/algorithm from the parent locator. The format of the sub-TLV is illustrated below.

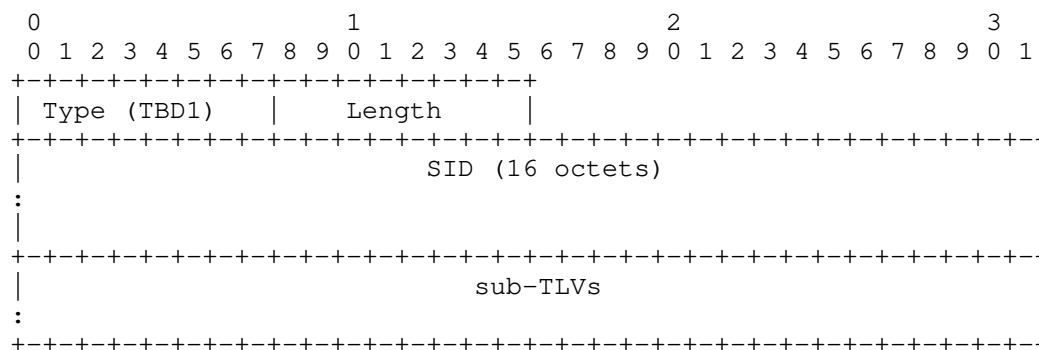


Figure 3: IS-IS SRv6 Mirror SID sub-TLV

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

Length: variable.

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

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

A protected node sub-TLV is used to carry the ID of the node to be protected by the SRv6 Mirror SID. It has the following format.

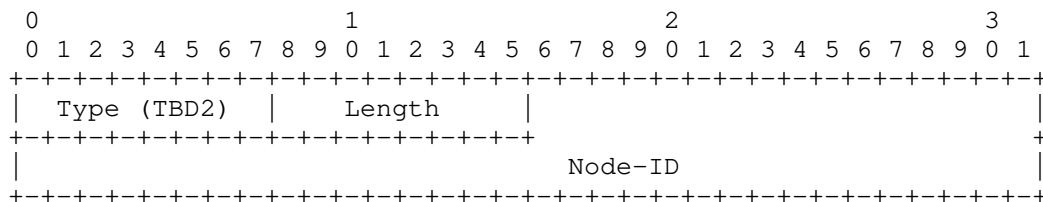


Figure 4: IS-IS Protected Node sub-TLV

Type: TBD2 (suggested value 1) is to be assigned by IANA.

Length: 1 octet. Its value is 6.

Node-ID: 6 octets. It contains a 6-octet ISO Node-ID (ISO system-ID).

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

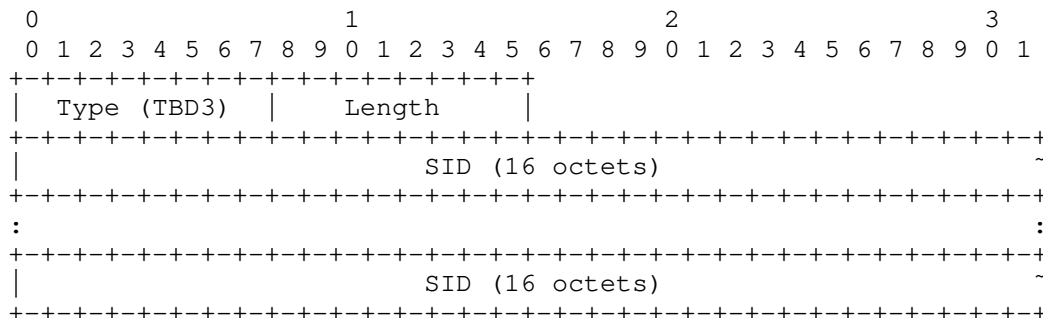


Figure 5: IS-IS Protected SIDs sub-TLV

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

Length: variable.

SID: 16 octets. This field encodes an SRv6 SID to be protected.

When node B advertises that B wants to protect node A with a Mirror SID through an LSP, the LSP contains an IS-IS SRv6 Mirror SID sub-TLV, which includes the Mirror SID and the node A's ID in an IS-IS Protected Node sub-TLV. If B wants to protect just a specific set of SIDs of node A, the Mirror SID sub-TLV includes these SIDs in an IS-

IS Protected SIDs sub-TLV; otherwise (i.e., B wants to protect all the SIDs of A) it does not contain any IS-IS Protected SIDs sub-TLV.

Note: the IS-IS extensions for SR MPLS is described in [RFC8667]. It says that the SID/Label Binding TLV may also be used to advertise a Mirror SID. For B to protect egress A of SR MPLS path, B may also use this TLV to advertise the node A's ID and a specific set of SIDs of A to be protected. An IS-IS SR MPLS Mirror SID sub-TLV may be obtained from an IS-IS SRv6 Mirror SID sub-TLV by replacing each SID field in the latter with an SID/Label sub-TLV. B may advertise a SID/Label Binding TLV including this IS-IS SR MPLS Mirror SID sub-TLV.

Alternatively, an IS-IS SR MPLS Mirror Supplement sub-TLV is defined from an IS-IS SRv6 Mirror SID sub-TLV by removing the SID field in the top level and replacing each other SID field with an SID/Label sub-TLV. That is that an IS-IS SR MPLS Mirror Supplement sub-TLV just contains a Protected Node sub-TLV and a Protected SIDs sub-TLV, which includes SID/Label sub-TLVs. When the SID/Label Binding TLV contains an SID/Label sub-TLV for the Mirror SID, it includes an IS-IS SR MPLS Mirror Supplement sub-TLV.

4.2. Extensions to OSPF

Similarly, a new sub-TLV, called OSPF Mirror SID sub-TLV, is defined. It is used to advertise SRv6 Mirror SID and the ID of the node to be protected. Its format is illustrated below.

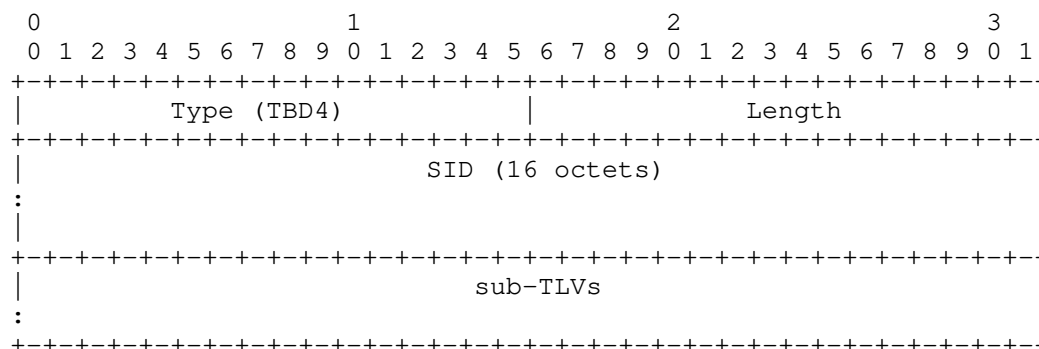


Figure 6: OSPF SRv6 Mirror SID sub-TLV

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

Length: variable.

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

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

A protected node sub-TLV is used to carry the ID of the node to be protected by the SRv6 Mirror SID. It has the following format.

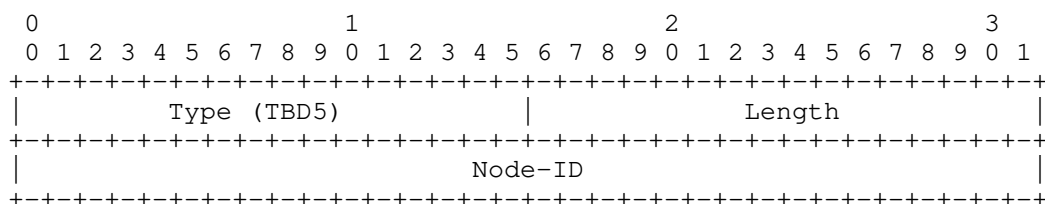


Figure 7: OSPF Protected Node sub-TLV

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

Length: 2 octets. Its value is 4.

Node-ID: 4 octets. It contains the ID of the OSPF node or router to be protected.

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

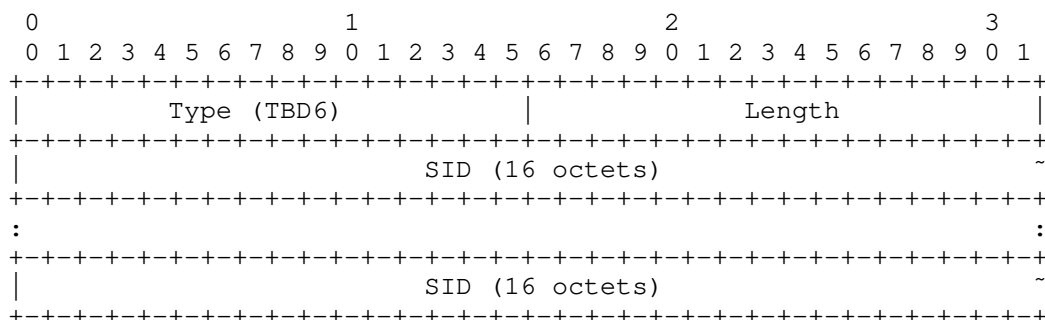


Figure 8: OSPF Protected SIDs sub-TLV

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

Length: variable.

SID: 16 octets. This field encodes an SRv6 SID to be protected.

5. Security Considerations

The security about the egress protection is described in in details in [RFC8679]. The extensions to OSPF and IS-IS described in this document for SRv6 path egress protection should not cause extra security issues.

6. IANA Considerations

6.1. IS-IS

Under "Sub-TLVs for TLVs 27, 135, 235, 236 and 237 registry" [I-D.ietf-lsr-isis-srv6-extensions], IANA is requested to add the following new Sub-TLV:

Sub-TLV Type	Sub-TLV Name	Reference
8	SRv6 Mirror SID Sub-TLV	This document

IANA is requested to create and maintain a new registry for sub-sub-TLVs of the SRv6 Mirror SID Sub-TLV. The suggested registry name is

- o Sub-Sub-TLVs for SRv6 Mirror SID Sub-TLV

Initial values for the registry are given below. The future assignments are to be made through IETF Review [RFC5226].

Value	Sub-Sub-TLV Name	Definition
0	Reserved	
1	Protected Node Sub-Sub-TLV	This Document
2	Protected SIDs Sub-Sub-TLV	
3-255	Unassigned	

6.2. OSPFv3

Under registry "OSPFv3 Locator LSA Sub-TLVs" [I-D.li-ospf-ospfv3-srv6-extensions], IANA is requested to assign the following new Sub-TLV:

Sub-TLV Type	Sub-TLV Name	Reference
8	SRv6 Mirror SID Sub-TLV	This document

IANA is requested to create and maintain a new registry for sub-sub-TLVs of the SRv6 Mirror SID Sub-TLV. The suggested registry name is

- o Sub-Sub-TLVs for SRv6 Mirror SID Sub-TLV

Initial values for the registry are given below. The future assignments are to be made through IETF Review [RFC5226].

Value	Sub-Sub-TLV Name	Definition
-----	-----	-----
0	Reserved	
1	Protected Node Sub-Sub-TLV	This Document
2	Protected SIDs Sub-Sub-TLV	
3-65535	Unassigned	

7. Acknowledgements

The authors would like to thank Peter Psenak, Yimin Shen, Zhenqiang Li, Alexander Vainshtein, Greg Mirsky, Bruno Decraene and Jeff Tantsura for their comments to this work.

8. References

8.1. Normative References

- [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-05 (work in progress), February 2020.
- [I-D.li-ospf-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", draft-li-ospf-ospfv3-srv6-extensions-07 (work in progress), November 2019.
- [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>>.

- [RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", RFC 7490, DOI 10.17487/RFC7490, April 2015, <<https://www.rfc-editor.org/info/rfc7490>>.
- [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>>.
- [RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <<https://www.rfc-editor.org/info/rfc8665>>.
- [RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", RFC 8667, DOI 10.17487/RFC8667, December 2019, <<https://www.rfc-editor.org/info/rfc8667>>.
- [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>>.

8.2. 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-05 (work in progress), July 2019.
- [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-07 (work in progress), January 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-02 (work in progress), January 2020.

- [I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Voyer, D., Bogdanov, A., and
P. Mattes, "Segment Routing Policy Architecture", draft-
ietf-spring-segment-routing-policy-06 (work in progress),
December 2019.
- [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.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an
IANA Considerations Section in RFCs", RFC 5226,
DOI 10.17487/RFC5226, May 2008,
<<https://www.rfc-editor.org/info/rfc5226>>.
- [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

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

Email: huzhibo@huawei.com

Huaimo Chen
Futurewei
Boston, MA
USA

Email: Huaimo.chen@futurewei.com

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

Email: chenhuan6@chinatelecom.cn

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

Email: baggio.wupeng@huawei.com

Mehmet Toy
Verizon
USA

Email: mehmet.toy@verizon.com

Chang Cao
China Unicom
Beijing China

Email: caoc15@chinaunicom.cn

Tao He
China Unicom
Beijing China

Email: het21@chinaunicom.cn

Lei Liu
Fujitsu
USA

Email: liulei.kddi@gmail.com

Xufeng Liu
Volta Networks
McLean, VA
USA

Email: xufeng.liu.ietf@gmail.com