Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 24, 2021

Z. Hu
Huawei Technologies
H. Chen
Futurewei
J. Yao
Huawei Technologies
C. Bowers
Juniper Networks
Y. Zhu
China Telecom
Y. Liu
China Mobile

February 20, 2021

SR-TE Path Midpoint Protection
draft-hu-spring-segment-routing-proxy-forwarding-13

Abstract

Segment Routing Traffic Engineering (SR-TE) supports explicit paths using segment lists containing adjacency-SIDs, node-SIDs and binding-SIDs. The current SR FRR such as TI-LFA provides fast re-route protection for the failure of a node along a SR-TE path by the direct neighbor or say point of local repair (PLR) to the failure. However, once the IGP converges, the SR FRR is no longer sufficient to forward traffic of the path around the failure, since the non-neighbors of the failure will no longer have a route to the failed node. This document describes a mechanism for fast re-route protection against the failure of a SR-TE path after the IGP converges. It provides fast re-route protection for an adjacency segment, a node segment and a binding segment of the 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 August 24, 2021.

Copyright Notice

Copyright (c) 2021 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 3
2. Proxy Forwarding 3
3. Extensions to IGP for Proxy Forwarding 4
  3.1. Extensions to OSPF 4
    3.1.1. Advertising Proxy Forwarding 4
    3.1.2. Advertising Binding Segment 7
  3.2. Extensions to IS-IS 10
    3.2.1. Advertising Proxy Forwarding 10
    3.2.2. Advertising Binding Segment 12
4. Building Proxy Forwarding Table 13
  4.1. Advertising Proxy Forwarding 15
  4.2. Building Proxy Forwarding Table 15
5. Node Protection for Segment List 16
  5.1. Next Segment is an Adjacency Segment 16
  5.2. Next Segment is a Node Segment 17
  5.3. Next Segment is a Binding Segment 17
6. Security Considerations 18
7. IANA Considerations 18
  7.1. OSPFv2 18
  7.2. OSPFv3 19
  7.3. IS-IS 20
8. Acknowledgements 20
9. References 21
  9.1. Normative References 21

9.2. Informative References 22
Authors' Addresses 22

1. Introduction

Segment Routing Traffic Engineering (SR-TE) is a technology that implements traffic engineering using a segment list. SR-TE supports the creation of explicit paths using adjacency-SIDs, node-SIDs, anycast-SIDs, and binding-SIDs. A node-SID in the segment list defining an SR-TE path indicates a loose hop that the SR-TE path should pass through. When the node fails, the network may no longer be able to properly forward traffic on that SR-TE path.

[I-D.ietf-rtgwg-segment-routing-ti-lfa] describes an SR FRR mechanism that provides fast re-route protection for the failure of a node on a SR-TE path by the direct neighbor or say point of local repair (PLR) to the failure. However, once the IGP converges, the SR FRR is no longer sufficient to forward traffic of the path around the failure, since the non-neighbors of the failure will no longer have a route to the failed node and drop the traffic.

To solve this problem, [I-D.ietf-spring-segment-protection-sr-te-paths] proposes that a hold timer should be configured on every router in a network. After the IGP converges on the event of a node failure, if the node-SID of the failed node becomes unreachable, the forwarding changes should not be communicated to the forwarding planes on all configured routers (including PLRs for the failed node) until the hold timer expires. This solution may not work for some cases such as some of nodes in the network not supporting this solution.

This document describes a proxy protection/forwarding mechanism, which provides more protection coverages. It considers the fast re-route protection capability of every node in the network and supports the fast re-route protection of the binding-SIDs on a failed node.

2. Proxy Forwarding

In the proxy forwarding mechanism, each neighbor of a possible failed node advertises its SR proxy forwarding capability in its network domain when it has the capability. This capability indicates that the neighbor (the Proxy Forwarder) will forward traffic on behalf of the failed node. A router receiving the SR Proxy Forwarding capability from the neighbors of a failed node will send traffic using the node-SID of the failed node to the nearest Proxy Forwarder after the IGP converges on the event of the failure.

Once the affected traffic reaches a Proxy Forwarder, it sends the traffic on the post-failure shortest path to the node immediately following the failed node in the segment list.

For a binding segment of a possible failed node, the node advertises the information about the binding segment, including the binding SID and the list of SIDs associated with the binding SID, to its direct neighbors only. Note that the information is not advertised in the network domain.

After the node fails and the IGP converges on the failure, the traffic with the binding SID of the failed node will reach its neighbor having SR Proxy Forwarding capability. Once receiving the traffic, the neighbor swaps the binding SID with the list of SIDs associated with the binding SID and sends the traffic along the post-failure shortest path to the first node in the segment list.

3. Extensions to IGP for Proxy Forwarding

This section defines extensions to IGP for advertising the SR proxy forwarding capability of a node in a network domain and the information about each binding segment (including its binding SID and the list of SIDs associated) of a node to its direct neighbors.

3.1. Extensions to OSPF

3.1.1. Advertising Proxy Forwarding

When a node P has the capability to do a SR proxy forwarding for all its neighboring nodes for protecting the failures of these nodes, node P advertises its SR proxy forwarding capability in its router information opaque LSA, which contains a Router Functional Capabilities TLV of the format as shown in Figure 1.

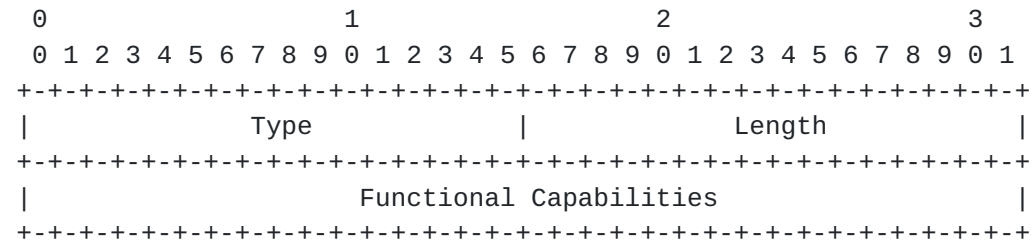


Figure 1: Router Functional Capabilities TLV

One bit (called PF bit) in the Functional Capabilities field of the TLV is used to indicate node P's SR proxy forwarding capability. When this bit is set to one by node P, it indicates that node P is capable of doing a SR proxy forwarding for its neighboring nodes.

For a node X in the network, it learns the prefix/node SID of node N, which is originated and advertised by node N. It creates a proxy prefix/node SID of node N for node P if node P is capable of doing SR proxy forwarding for node N. The proxy prefix/node SID of node N for node P is a copy of the prefix/node SID of node N originated by node N, but stored under (or say, associated with) node P.

In normal operations, node X prefers to use the prefix/node SID of node N. When node N fails, node X prefers to use the proxy prefix/ node SID of node N. Thus node X will forward the traffic targeting to the prefix/node SID of node N to node P when node N fails, and node P will do a SR proxy forwarding for node N and forward the traffic towards its final destination without going through node N. After node N fails, node X will keep the FIB entry to the proxy prefix/node SID of node N for a given period of time.

Note that the behaviors of normal IP forwarding and routing convergences in a network are not changed at all by the SR proxy forwarding. For example, the next hop used by BGP is an IP address (or prefix). The IGP and BGP converge in normal ways for changes in the network. The packet with its IP destination to this next hop is forwarded according to the IP forwarding table (FIB) derived from IGP and BGP routes.

If node P can not do a SR proxy forwarding for all its neighboring nodes, but for some of them, then it advertises the node SID of each of the nodes as a proxy node SID, indicating that it is able to do proxy forwarding for the node SID.

A new TLV, called Proxy Node SIDs TLV, is defined for node P to advertise the node SIDs of some of its neighboring nodes. It has the format as shown in Figure 2.

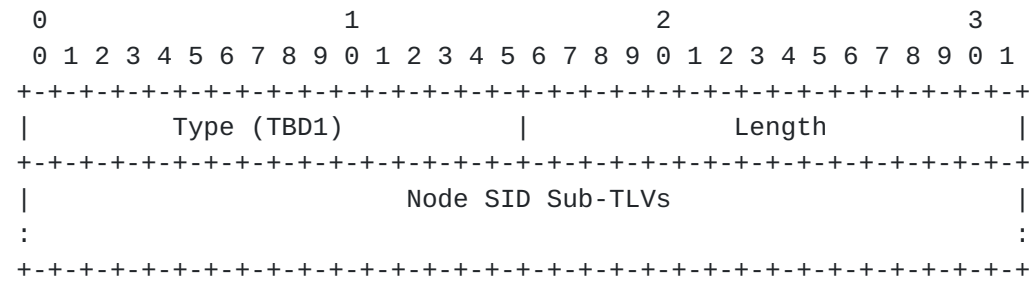


Figure 2: OSPF Proxy Node SIDs TLV

The Type (TBD1) is to be assigned by IANA. The TLV contains a number of Node SID Sub-TLVs. The Length is the total size of the Node SID Sub-TLVs included in the TLV. A Node SID Sub-TLV is the Prefix SID Sub-TLV defined in [RFC8665].

A proxy forwarding node P originates an Extended Prefix Opaque LSA containing this new TLV. The TLV includes the Node SID Sub-TLVs for the node SIDs of some of P's neighboring nodes. For each of some of P's neighboring nodes, the Node SID Sub-TLV for its prefix/node SID is included the TLV. This prefix/node SID is called a proxy prefix/ node SID.

A proxy forwarding node will originate an Extended Prefix Opaque LSA, which includes a Proxy Node SIDs TLV. The format of the LSA is shown in Figure 3.

For a proxy forwarding node P, having a number of neighboring nodes, P originates and maintains an Extended Prefix Opaque LSA, which includes a Proxy Node SIDs TLV. The TLV contains the Prefix/Node SID Sub-TLV for each of some of the neighboring nodes after node P creates the corresponding proxy forwarding entries for protecting the failure of some of the neighboring nodes.

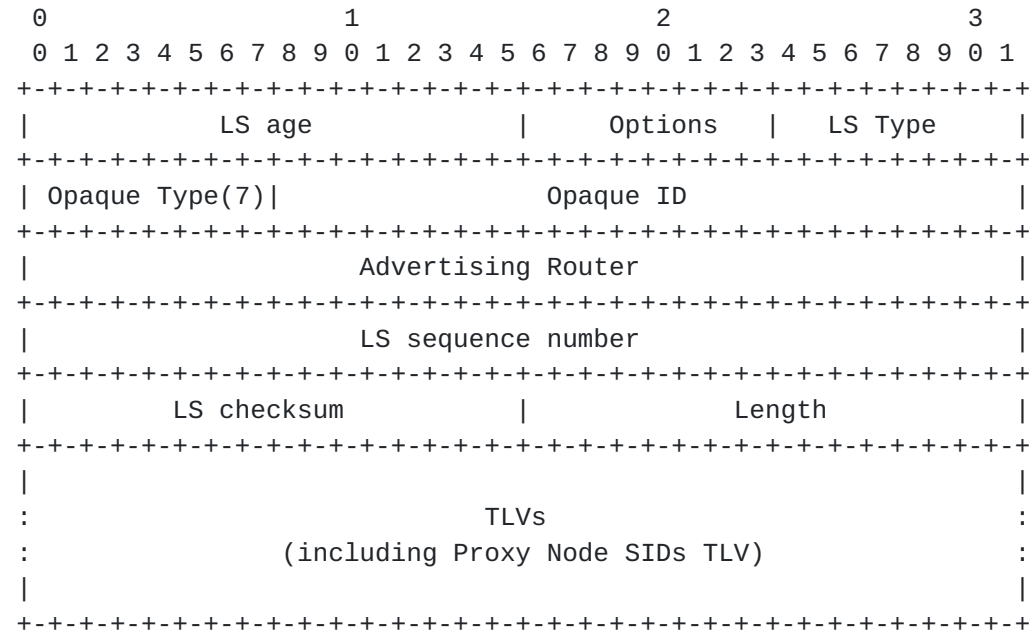


Figure 3: OSPFv2 Extended Prefix Opaque LSA

When an neighboring node fails, P maintains the LSA with the TLV containing the Prefix/Node SID Sub-TLV for the neighboring node for a given period of time. After the given period of time, the Prefix/ Node SID Sub-TLV for the neighboring node is removed from the TLV in the LSA and then after a given time the corresponding proxy forwarding entries for protecting the failure of the neighboring node is removed.

For a node X in the network, it learns the prefix/node SID of node N and the proxy prefix/node SID of node N. The former is originated and advertised by node N, and the latter is originated and advertised by the proxy forwarding node P of node N. Note that the proxy Prefix/Node SID Sub-TLV for node N does not contain a prefix of node N, and the prefix is the prefix associated with the prefix/node SID of node N originated by node N.

In normal operations, node X prefers to use the prefix/node SID of node N. When node N fails, node X prefers to use the proxy prefix/ node SID of node N. Thus node X will forward the traffic targeting to node N to node P when node N fails, and node P will do a proxy forwarding for node N and forward the traffic towards its destination without going through node N.

3.1.2. Advertising Binding Segment

For a binding segment (or binding for short) on a node A, which consists of a binding SID and a list of segments, node A advertises an LSA containing the binding (i.e., the binding SID and the list of the segments). The LSA is advertised only to each of the node A's neighboring nodes. For OSPFv2, the LSA is a opaque LSA of LS type 9 (i.e., a link local scope LSA).

A binding segment is represented by binding segment TLV of the format as shown in Figure 4.

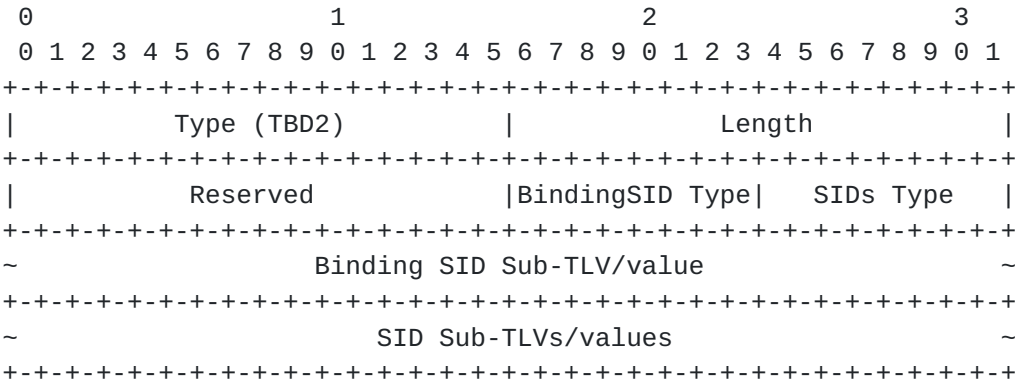


Figure 4: OSPF Binding Segment TLV

It comprises a binding SID and a list of segments (SIDs). The fields of this TLV are defined as follows:

Type: 2 octets, its value (TBD2) is to be assigned by IANA.

Length: 2 octets, its value is (4 + length of Sub-TLVs/values).

Binding SID Type (BT): 1 octet indicates whether the binding SID is represented by a Sub-TLV or a value included in the TLV. For the binding SID represented by a value, it indicates the type of binding SID. The following BT values are defined:

o BT = 0: The binding SID is represented by a Sub-TLV (i.e., Binding SID Sub-TLV) in the TLV. A binding SID Sub-TLV is a SID/Label Sub-TLV defined in [RFC8665]. BT != 0 indicates that the binding SID is represented by a value.

o BT = 1: The binding SID value is a label, which is represented by the 20 rightmost bits. The length of the value is 3 octets.

o BT = 2: The binding SID value is a 32-bit SID. The length of the value is 4 octets.

SIDs Type (ST): 1 octet indicates whether the list of segments (SIDs) are represented by Sub-TLVs or values included in the TLV. For the SIDs represented by values, it indicates the type of SIDs. The following ST values are defined:

o ST = 0: The SIDs are represented by Sub-TLVs (i.e., SID Sub-TLVs) in the TLV. A SID Sub-TLV is an Adj-SID Sub-TLV, a Prefix-SID Sub-TLV or a SID/Label Sub-TLV defined in [RFC8665]. ST != 0 indicates that the SIDs are represented by values.

o ST = 1: Each of the SID values is a label, which is represented by the 20 rightmost bits. The length of the value is 3 octets.

o ST = 2: Each of the SID values is a 32-bit SID. The length of the value is 4 octets.

The opaque LSA of LS Type 9 containing the binding segment (i.e., the binding SID and the list of the segments) has the format as shown in Figure 5. It may have Opaque Type of x (the exact type is to be assigned by IANA) for Binding Segment Opaque LSA.

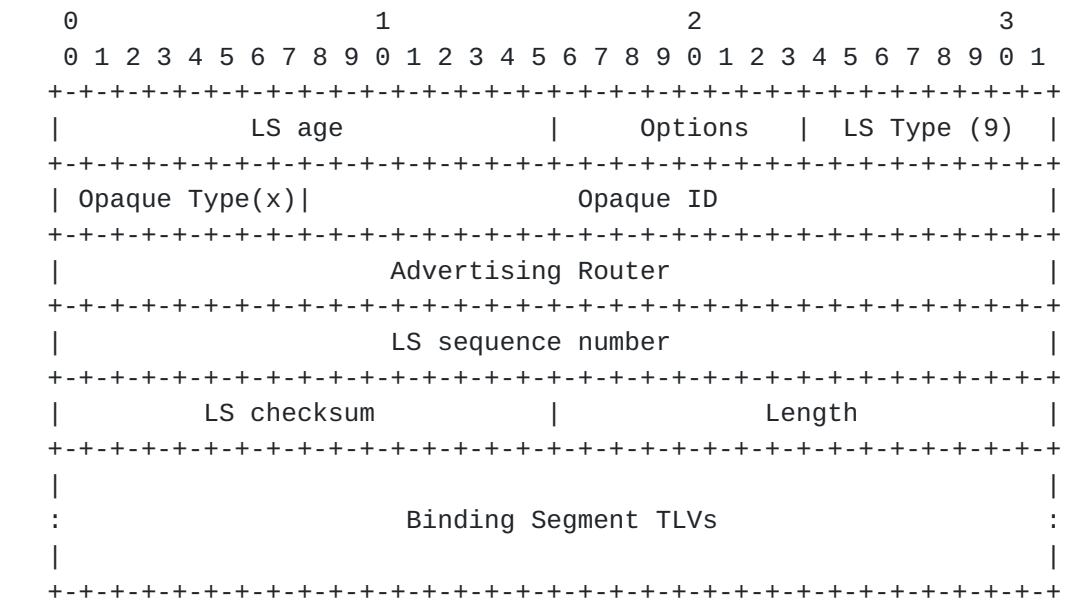


Figure 5: OSPFv2 Binding Segment Opaque LSA

For every binding on a node A, the LSA originated by A contains a binding segment TLV for it.

For node A running OSPFv3, it originates a link-local scoping LSA of a new LSA function code (TBD3) containing binding segment TLVs for the bindings on it. The format of the LSA is illustrated in Figure 6.

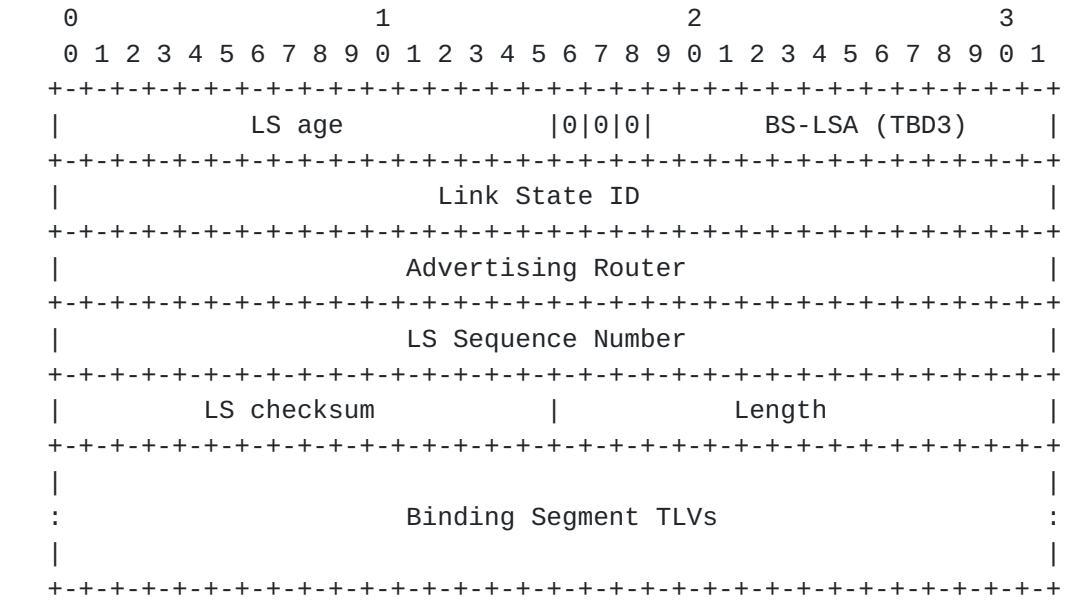


Figure 6: OSPFv3 Binding Segment Opaque LSA

The U-bit is set to 0, and the scope is set to 00 for link-local scoping.

3.2. Extensions to IS-IS

3.2.1. Advertising Proxy Forwarding

When a node P has the capability to do a SR proxy forwarding for its neighboring nodes for protecting the failures of them, node P advertises its SR proxy forwarding capability in its LSP, which contains a Router Capability TLV of Type 242 including a SR capabilities sub-TLV of sub-Type 2.

One bit (called PF bit as shown in Figure 7) in the Flags field of the SR capabilities sub-TLV is defined to indicate node P's SR proxy forwarding capability. When this bit is set to one by node P, it indicates that node P is capable of doing a SR proxy forwarding for its neighboring nodes.

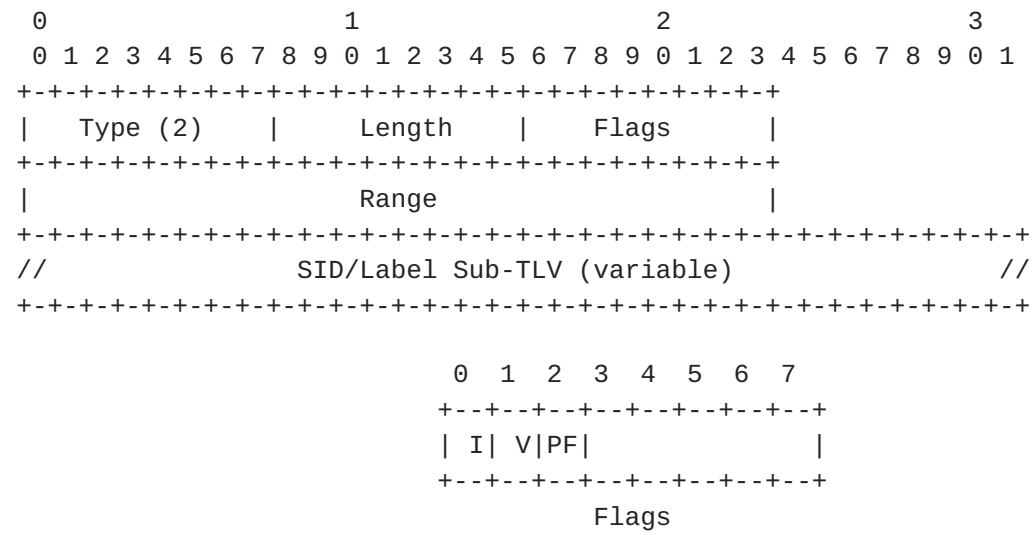


Figure 7: SR Capabilities sub-TLV

If node P can not do a SR proxy forwarding for all its neighboring nodes, but for some of them, then it advertises the node SID of each of the nodes as a proxy node SID, indicating that it is able to do proxy forwarding for the node SID.

The IS-IS SID/Label Binding TLV (suggested value 149) is defined in [RFC8667]. A Proxy Forwarder uses the SID/Label Binding TLV to advertise the node SID of its neighboring node. The Flags field of the SID/Label Binding TLV is extended to include a P flag as shown in Figure 8. The prefix/node SID in prefix/node SID Sub-TLV included in

SID/Label Binding TLV is identified as a proxy forwarding prefix/node SID.

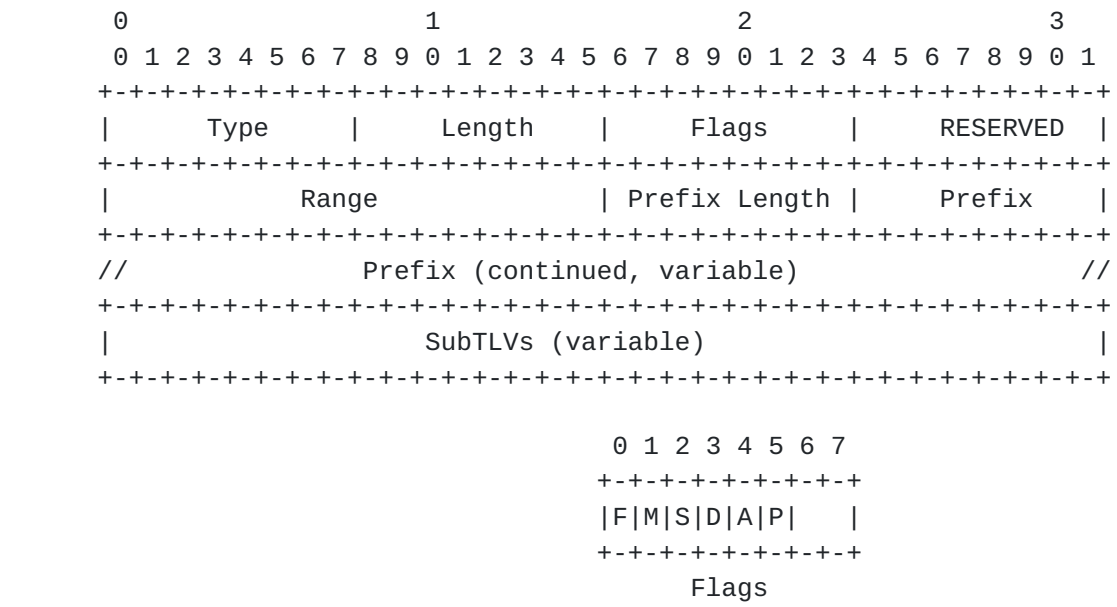


Figure 8: SID/Label Binding TLV

Where:

P-Flag: Proxy forwarding flag. If set, this prefix/node SID is advertised by the proxy node. This TLV is used to announce that the node has the ability to proxy forward the prefix/node SID.

When the P-flag is set in the SID/Label Binding TLV, the following usage rules apply.

The Range, Prefix Length and Prefix field are not used. They should be set to zero on transmission and ignored on receipt.

SID/Label Binding TLV contains a number of prefix/node SID Sub-TLVs. The TLV advertised by a proxy forwarding node P contains prefix/node SID Sub-TLVs for the node SIDs of P's neighbor nodes. Each of the Sub-TLVs is a prefix/node SID Sub-TLV defined in [RFC8667]. From the SID in a prefix/node SID Sub-TLV advertised by the Proxy Forwarding node, its prefix can be obtained through matching corresponding prefix/node SID advertised by the neighbor/protected node using TLV-135 (or 235, 236, or 237) together with the prefix/node SID Sub-TLV.

3.2.2. Advertising Binding Segment

[I-D.ietf-spring-segment-routing-policy] has defined the usage of binding-SID. For supporting binding SID proxy forwarding, a new IS-IS TLV, called Binding Segment TLV, is defined. It contains a binding SID and a list of segments (SIDs). This TLV may be advertised in IS-IS Hello (IIH) PDUs, LSPs, or in Circuit Scoped Link State PDUs (CS-LSP) [RFC7356]. Its format is shown in Figure 9.

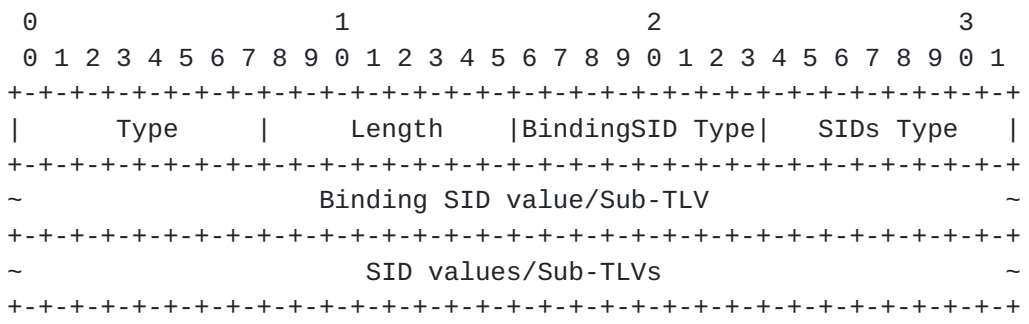


Figure 9: IS-IS Binding Segment TLV

The fields of this TLV are defined as follows:

Type: 1 octet Suggested value 152 (to be assigned by IANA)

Length: 1 octet (2 + length of Sub-TLVs/values).

Binding SID Type (BT): 1 octet indicates whether the binding SID is represented by a Sub-TLV or a value included in the TLV. For the binding SID represented by a value, it indicates the type of binding SID. The following BT values are defined:

o BT = 0: The binding SID is represented by a Sub-TLV (i.e., binding SID Sub-TLV) in the TLV. A binding SID Sub-TLV is a SID/Label Sub-TLV defined in [RFC8667]. BT != 0 indicates that the binding SID is represented by a value.

o BT = 1: The binding SID value is a label, which is represented by the 20 rightmost bits. The length of the value is 3 octets.

o BT = 2: The binding SID value is a 32-bit SID. The length of the value is 4 octets.

SIDs Type (ST): 1 octet indicates whether the SIDs are represented by Sub-TLVs or values included in the TLV. For the SIDs represented by values, it indicates the type of SIDs. The following ST values are defined:

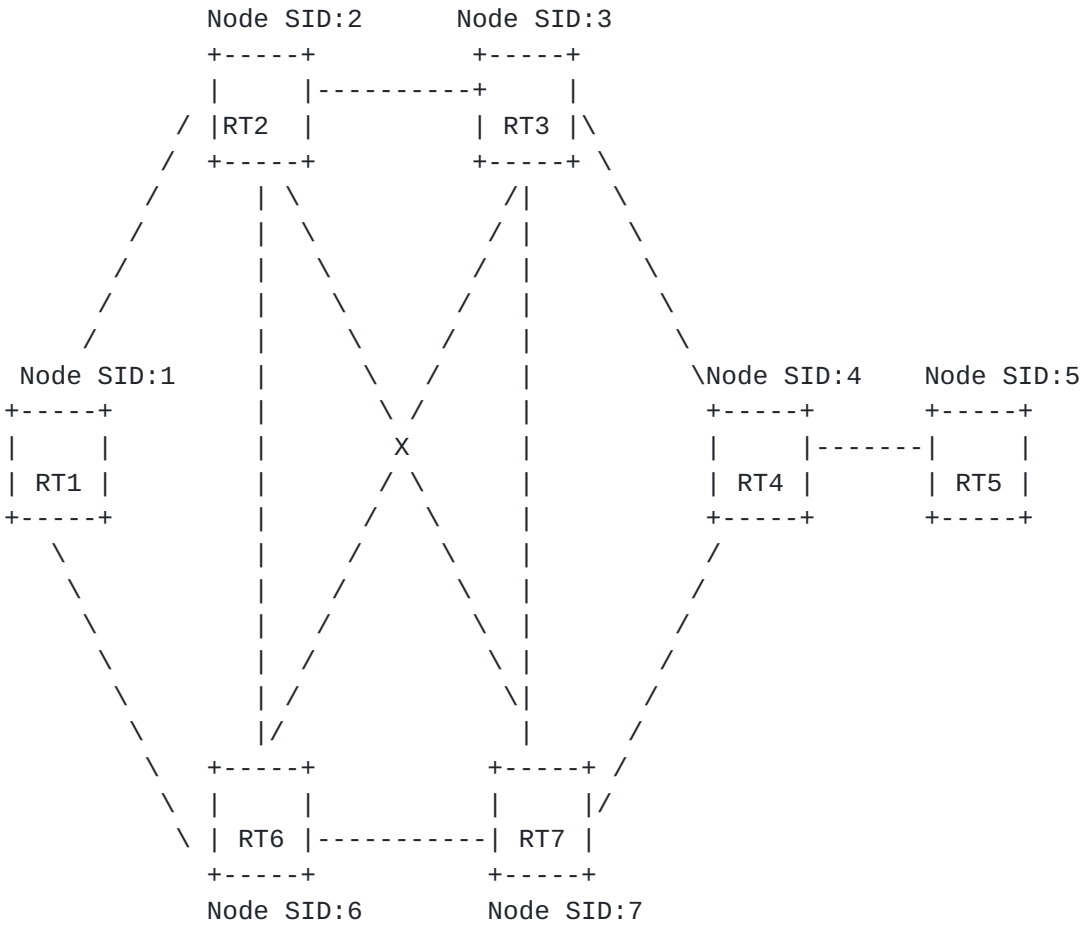
o ST = 0: The SIDs are represented by Sub-TLVs (i.e., SID Sub-TLVs) in the TLV. A SID Sub-TLV is an Adj-SID Sub-TLV, a Prefix-SID Sub-TLV or a SID/Label Sub-TLV defined in [RFC8667]. ST != 0 indicates that the SIDs are represented by values.

o ST = 1: Each of the SID values is a label, which is represented by the 20 rightmost bits. The length of the value is 3 octets.

o ST = 2: Each of the SID values is a 32-bit SID. The length of the value is 4 octets.

4. Building Proxy Forwarding Table

Figure 10 is used to illustrate the SR proxy forwarding approach. Each node N has SRGB = [N000-N999]. RT1 is an ingress node of SR domain. RT3 is a failure node. RT2 is a Point of Local Repair (PLR) node, i.e., a proxy forwarding node. Three label stacks are shown in the figure. Label Stack 1 uses only adjacency-SIDs and represents the path RT1->RT2->RT3->RT4->RT5. Label Stack 2 uses only node-SIDs and represents the ECMP-aware path RT1->RT3->RT4->RT5. Label Stack 3 uses a node-SID and a binding SID. The Binding-SID with label=100 at RT3 represents the ECMP-aware path RT3->RT4->RT5. So Label Stack 3, which consists of the node-SID for RT3 following by Binding-SID=100, represents the ECMP-aware path RT1->RT3->RT4->RT5.



Node SRGB	Adj-SID	Label Stack 1	Label Stack 2	Label Stack 3
RT1:[1000-1999]	RT1->RT2:10012	10012	1003	1003
RT2:[2000-2999]	RT2->RT3:20023	20023	3004	100
RT3:[3000-3999]	RT3->RT6:30036	30034	4005	100 is binding SID to {30034,40045}
RT4:[4000=4999]	RT3->RT7:30037	40045		
RT5:[5000-5999]	RT3->RT4:30034			
RT6:[6000-6999]	RT7->RT4:70074			
RT7:[7000-7999]	RT4->RT5:40045			

Figure 10: Topology of SR-TE Path

4.1. Advertising Proxy Forwarding

If the Point of Local Repair (PLR), for example, RT2, has the capability to do a SR proxy forwarding for all its neighboring nodes, it must advertise this capability. If the PLR can not do a SR proxy forwarding for all its neighboring nodes, but for some of them, for example, RT3, then it uses proxy Node SIDs TLV to advertise the prefix-SID learned from RT3. The TLV contains the Sub-TLV/value for the prefix/node SID of RT3 as a proxy SID. When RT3 fails, RT2 needs to maintain the Sub-TLV/value for a period of time. When the proxy forwarding table corresponding to the fault node is deleted (see section 3.2), the Sub-TLV/value is withdrawn. The nodes in the network (for example, RT1) learn the prefix/node SID TLV advertised by RT3 and the proxy Node SIDs TLV advertised by RT2. When RT3 is normal, the nodes prefer prefix/node SID TLV. When the RT3 fails, the proxy prefix/node SIDs TLV advertised by RT2 is preferred.

4.2. Building Proxy Forwarding Table

A SR proxy node P needs to build an independent proxy forwarding table for each neighbor N. The proxy forwarding table for node N contains the following information:

1: Node N's SRGB range and the difference between the SRGB start value of node P and that of node N;

2: All adjacency-SID of N and Node-SID of the node pointed to by node N's adjacency-SID.

3: The binding-SID of N and the label stack associated with the binding-SID.

Node P (PLR) uses a proxy forwarding table based on the next segment to find a node N as a backup forwarding entry to the adj-SID and Node-SID of node N. When node N fails, the proxy forwarding table needs to be maintained for a period of time, which is recommended for 30 minutes.

Node RT3 in the topology of Figure 1 is node N, and node RT2 is node P (PLR). RT2 builds the proxy forwarding table for RT3. RT2 calculates the proxy forwarding table for RT3, as shown in Figure 11.

In-label	SRGBDiffValue	Next Label	Action	Map Label
2003	-1000	30034	Fwd to RT4	2004
		30036	Fwd to RT6	2006
		30037	Fwd to RT7	2007
		100	Swap to { 30034, 40045 }	

Figure 11: RT2's Proxy Forwarding Table for RT3

5. Node Protection for Segment List

Segment Routing Traffic Engineering supports the creation of explicit paths using adjacency-SIDs, node-SIDs, and binding-SIDs. The label stack is a combination of one or more of adjacency-SIDs, node-SIDs, and binding-SIDs. This Section shows how a proxy node uses the SR proxy forwarding mechanism to protect traffic to the destination node when the next segment of label stack is adjacency-SIDs, node-SIDs, or binding-SIDs, respectively.

5.1. Next Segment is an Adjacency Segment

As shown in Figure 1, Label Stack 1 {10012, 20023, 30034, 40045} represents SR-TE strict explicit path RT1->RT2->RT3->RT4->RT5. When RT3 fails, node RT2 acts as a PLR, and uses next adj-SID (30034) of the label stack to lookup the proxy forwarding table built by RT2 locally for RT3. The path returned is the label forwarding path to RT3's next hop node RT4, which bypasses RT3. The specific steps are as follows:

a. RT1 pops top adj-SID 10012, and forwards the packet to RT2;

b. RT2 uses the label 20023 to identify the next hop node RT3, which has failed. RT2 pops label 20023 and queries the Proxy Forwarding Table corresponding to RT3 with label 30034. The query result is 2004. RT2 uses 2004 as the incoming label to query the label forwarding table. The next hop is RT7, and the incoming label is changed to 7004.

c. So the packet leaves RT2 out the interface to RT7 with label stack {7004, 40045}. RT4 forwards it to RT4, where the original path is rejoined.

d. RT2 forwards packets to RT7. RT7 queries the local routing table to forward the packet to RT4.

5.2. Next Segment is a Node Segment

As shown in Figure 1, Label Stack 2 {1003, 3004, 4005} represents SR-TE loose path RT1->RT3->RT4->RT5, where 1003 is the node SID of RT3.

When the node RT3 fails, the proxy forwarding TLV advertised by the RT2 is preferred to direct the traffic of the RT1 to the PLR node RT2. Node RT2 acts as a PLR node and queries the proxy forwarding table locally built for RT3. The path returned is the label forwarding path to RT3's next hop node RT4, which bypasses RT3. The specific steps are as follows:

a. RT1 swaps label 1003 to out-label 2003 to RT3.

b. RT2 receives the label forwarding packet whose top label of label stack is 2003, and searches for the local Routing Table, the behavior found is to lookup Proxy Forwarding table due to RT3 failure, RT2 pops label 2003.

c. RT2 uses 3004 as the in-label to lookup Proxy Forwarding table, The value of Map Label calculated based on SRGBDiffValue is 2004. and the query result is forwarding the packet to RT4.

d. Then RT2 queries the Routing Table to RT4, using the primary or backup path to RT4. The next hop is RT7.

e. RT2 forwards the packet to RT7. RT7 queries the local routing table to forward the packet to RT4.

f. After RT1 convergences, node SID 1003 is preferred to the proxy SID implied/advertised by RT2.

5.3. Next Segment is a Binding Segment

As shown in Figure 1, Label Stack 3 {1003, 100} represents SR-TE loose path RT1->RT3->RT4->RT5, where 100 is a Binding-SID, which represents segment list {30034, 40045}.

When the node RT3 fails, the proxy forwarding SID implied or advertised by the RT2 is preferred to forward the traffic of the RT1 to the PLR node RT2. Node RT2 acts as a PLR node and uses Binding-SID to query the proxy forwarding table locally built for RT3. The path returned is the label forwarding path to RT3's next hop node (RT4), which bypasses RT3. The specific steps are as follows:

a. RT1 swaps label 1003 to out-label 2003 to RT3.

b. RT2 receives the label forwarding packet whose top label of label stack is 2003, and searches for the local Routing Table, the behavior found is to lookup Proxy Forwarding table due to RT3 failure.

c. RT2 uses Binding-SID:100 (label 2003 has pop) as the in-label to lookup the Next Label record of the Proxy Forwarding Table, the behavior found is to swap to Segment list {30034, 40045}.

d. RT2 swaps Binding-SID:100 to Segment list {30034, 40045}, and uses the 3034 to lookup the Next Label record of the Proxy Forwarding table again. The behavior found is to forward the packet to RT4.

e. RT2 queries the Routing Table to RT4, using primary or backup path to RT4. The next hop is RT7.

f. RT2 forwards packets to RT7. RT7 queries the local routing table to forward the packet to RT4.

6. Security Considerations

The extensions to OSPF and IS-IS described in this document result in two types of behaviors in data plane when a node in a network fails. One is that for a node, which is a upstream (except for the direct upstream) node of the failed node along a SR-TE path, it continues to send the traffic to the failed node along the SR-TE path for an extended period of time. The other is that for a node, which is the direct upstream node of the failed node, it fast re-routes the traffic around the failed node to the direct downstream node of the failed node along the SR-TE path. These behaviors are internal to a network and should not cause extra security issues.

7. IANA Considerations

7.1. OSPFv2

Under Subregistry Name "OSPF Router Functional Capability Bits" within the "Open Shortest Path First v2 (OSPFv2) Parameters" [RFC7770], IANA is requested to assign one bit for Proxy Forwarding Capability as follows:

Bit number	Capability Name	Reference
31	Proxy Forwarding	This document

Under Registry Name "OSPFv2 Extended Prefix Opaque LSA TLVs" [RFC7684], IANA is requested to assign one new TLV value for OSPF Proxy Node SIDs as follows:

TLV Value	TLV Name	Reference
2	Proxy Node SIDs TLV	This document

Under Registry Name "Opaque Link-State Advertisements (LSA) Option Types" [RFC5250], IANA is requested to assign new Opaque Type registry values for Binding Segment LSA as follows:

Registry Value	Opaque Type	Reference
10	Binding Segment	This document

IANA is requested to create and maintain new registries:

- OSPFv2 Binding Segment Opaque LSA TLVs

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

Value	TLV Name	Definition
0	Reserved	
1	Binding Segment TLV	This Document
2-32767	Unassigned	
32768-65535	Reserved	

7.2. OSPFv3

Under Registry Name "OSPFv3 LSA Function Codes", IANA is requested to assign new registry values for Binding Segment LSA as follows:

Value	LSA Function Code Name	Reference
16	Binding Segment LSA	This document

IANA is requested to create and maintain new registries:

- OSPFv3 Binding Segment LSA TLVs

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

Value	TLV Name	Definition
0	Reserved	
1	Binding Segment TLV	This Document
2-32767	Unassigned	
32768-65535	Reserved	

7.3. IS-IS

Under Registration "Segment Routing Capability" in the "sub-TLVs for TLV 242" registry [RFC8667], IANA is requested to assign one bit flag for Proxy Forwarding Capability as follows:

Bit number	Capability Name	Reference
2	Proxy Forwarding (PF)	This document

Under Registration "Segment Identifier/Label Binding TLV 149" [RFC8667], IANA is requested to assign one bit P-Flag as follows:

Bit number	Flag Name	Reference
5	P-Flag	This document

Under Registry Name: IS-IS TLV Codepoints, IANA is requested to assign one new TLV value for IS-IS Binding Segment as follows:

Value	TLV Name	Reference
152	Binding Segment TLV	This Document

8. Acknowledgements

The authors would like to thank Peter Psenak, Acee Lindem, Les Ginsberg, Bruno Decraene and Jeff Tantsura for their comments to this work.

9. References

9.1. 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>.

[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>.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, DOI 10.17487/RFC5250, July 2008, <https://www.rfc-editor.org/info/rfc5250>.

[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>.

[RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", RFC 7684, DOI 10.17487/RFC7684, November 2015, <https://www.rfc-editor.org/info/rfc7684>.

[RFC7770] Lindem, A., Ed., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 7770, DOI 10.17487/RFC7770, February 2016, <https://www.rfc-editor.org/info/rfc7770>.

[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>.

9.2. Informative References

[I-D.ietf-rtgwg-segment-routing-ti-lfa]
Litkowski, S., Bashandy, A., Filsfils, C., Decraene, B., and D. Voyer, "Topology Independent Fast Reroute using Segment Routing", draft-ietf-rtgwg-segment-routing-ti-lfa-05 (work in progress), November 2020.

[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", draft-ietf-spring-segment-protection-sr-te-paths-00 (work in progress), September 2020.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-09 (work in progress), November 2020.

[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.

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

Email: huzhibo@huawei.com

Huaimo Chen
Futurewei
Boston, MA
USA

Email: Huaimo.chen@futurewei.com

Junda Yao
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: yaojunda@huawei.com

Chris Bowers
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
USA

Email: cbowers@juniper.net

Yongqing
China Telecom
109, West Zhongshan Road, Tianhe District
Guangzhou 510000
China

Email: zhuyq8@chinatelecom.cn

Yisong
China Mobile
510000
China

Email: liuyisong@chinamobile.com