Workgroup: Network Working Group

Internet-Draft: draft-hc-lsr-sr-proxy-fw-01

Published: 4 September 2022

Intended Status: Standards Track

Expires: 8 March 2023

Authors: Z. Hu H. Chen
Huawei Technologies Futurewei
J. Yao C. Bowers

Huawei Technologies Juniper Networks

Y. Zhu Y. Liu

China Telecom China Mobile

LSR for SR Proxy Forwarding

Abstract

This document describes extensions to OSPF and IS-IS to support SR proxy forwarding mechanism for fast protecting the failure of a node with segments on a SR-TE path. The segments of the node include adjacency, node or binding segments.

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 8 March 2023.

Copyright Notice

Copyright (c) 2022 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

- 1. Introduction
- 2. IGP Extensions/Re-uses for Proxy Forwarding
 - 2.1. OSPF Extensions/Re-uses
 - 2.1.1. Advertising Binding Segment
 - 2.1.2. Advertising Proxy Forwarding
 - 2.2. IS-IS Extensions/Re-uses
 - 2.2.1. Advertising Binding Segment
 - 2.2.2. Advertising Proxy Forwarding
- 3. Security Considerations
- 4. IANA Considerations
 - 4.1. <u>OSPFv2</u>
 - 4.2. OSPFv3
 - <u>4.3</u>. <u>IS-IS</u>
- Acknowledgements
- 6. References
 - 6.1. Normative References
 - 6.2. Informative References

Authors' Addresses

1. Introduction

[I-D.hu-spring-segment-routing-proxy-forwarding] describes a SR proxy forwarding for protection. Each neighbor of a possible failed node advertises its SR proxy forwarding capability 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 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 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/ segments associated with the binding SID and sends the traffic along the post-failure shortest path to the first node in the segment list.

2. IGP Extensions/Re-uses for Proxy Forwarding

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

2.1. OSPF Extensions/Re-uses

2.1.1. 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 1.

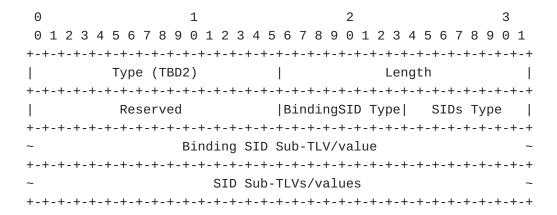


Figure 1: OSPF Binding Segment TLV

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 <u>Figure 2</u>. It may have Opaque Type of x (the exact type is to be assigned by IANA) for Binding Segment Opaque LSA.

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 2 3 4	5 6 7 8 9 0 1			
+-+-+-+-	+-+-+-+-+-+	-+-+-+-+-	+-+-+-+-+-+			
LS age		Options	LS Type (9)			
+-+-+-+-+-+-+-	+-+-+-+-+-+	-+-+-+-	+-+-+-+-+-+			
Opaque Type(x)	C	paque ID	1			
+-+-+-+-	+-+-+-+-+-+	-+-+-+-	+-+-+-+-+-+			
	Advertising F	outer	1			
+-+-+-+-	+-+-+-+-+-+	-+-+-+-	+-+-+-+-+-+			
1	LS sequence r	umber	1			
+-+-+-+-	+-+-+-+-+-+	-+-+-+-+-	+-+-+-+-+-+			
LS checksur	n	Leng	th			
+-						
1			1			
:	Binding Segm	ent TLVs	:			
			1			
+-						

Figure 2: 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 $\underline{\text{Figure}}$ $\underline{3}$.

0								1										2										3	
0	1 2	3	4 5	5 6	7	8	9	0	1	2	3	4	. !	5 6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
+-	+-+-	+-+	·-+·	-+-	+	+-+	-+	- +		- +	+-	+-	+	-+-	+-	+	+	+-	+-	+-	+-	+-	+-	+	+	+ - +	+ - +	+	- +
						age								•	•	•	•							•		D3)	•		
+-	+-+-	+-+	· - + ·	-+-	+	+-+	-+	+									+	+ -	+-	+-	+-	+-	+-	+	+	+	+ - +	+	·-+
														tat															ı
+-	+-+-	+-+	·-+·	-+-	+	+-+	-+	- +		+ -	+-	+-	+	-+-	+-	+	+	+ -	+-	+-	+-	+-	+-	+	+	+ - +	+ - +	+	+
									Α	vb	er	ti	S.	inç	j R	ou'	te	r											
+-	+-+-	+-+	·-+·	-+-	+	+-+	-+	- +		+ -	+-	+-	+	-+-	+-	+	+	+ -	+-	+-	+-	+-	+-	+	+	+-+	+ - +	+	- +
									LS	S	Se	qu	eı	nce	N	uml	be	r											
+-	+-+-	+-+	·-+·	-+-	+	+-+	-+	- +		+ -	+-	+-	+	-+-	+-	+	+	+-	+-	+-	+-	+-	+-	+	+	+	+ - +	+	- +
			LS	S C	he	cks	um	1						Ι						Le	ng	th							
+-	+-+-	+-+	·-+·	-+-	+	+-+	-+	- +		+ -	+-	+-	+	-+-	+-	+	+	+-	+-	+-	+-	+-	+-	+	+	+	+ - +	+	- +
I																													- 1
:									В:	in	di	na		Seg	ıme	nt	TI	LV	S										:
ı												J																	- 1
' +-	+-+-	+-+	+-	-+-	+	+-+	-+	· - +		+ -	+-	+-	+	-+-	+-	+	+	+ -	+-	+-	+-	+-	+-	+	+ - •	+	+ - +	+	ا + - ٠
			-	-											-														

Figure 3: OSPFv3 Binding Segment Opaque LSA

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

2.1.2. 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 these nodes, P advertises its capability for these nodes. The mirror SID [RFC8402] [RFC8667] for a node N (Neighbor of P) advertised by P indicates the capability of P for N.

Alternatively, P advertises its capability in its router information opaque LSA with Router Functional Capabilities TLV [RFC7770]. One bit (called PF bit) in the Functional Capabilities field of the TLV is used to indicate node P's 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. The route to the proxy prefix/node SID is through proxy forwarding capable nodes.

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.

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.

2.2. IS-IS Extensions/Re-uses

2.2.1. Advertising Binding Segment

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 is advertised in Circuit Scoped Link State PDUs (CS-LSP) [RFC7356]. Its format is shown in Figure 4.

Figure 4: 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).

The other fields are the same as those in <a>Figure 1.

2.2.2. Advertising Proxy Forwarding

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

One bit (called PF bit) in the Flags field of the SR capabilities sub-TLV is defined to indicate node P's 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.

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 5. 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.

```
1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
Flags
     Length
           Range
           | Prefix Length |
                    Prefix
Prefix (continued, variable)
SubTLVs (variable)
0 1 2 3 4 5 6 7
            +-+-+-+-+-+-+
            |F|M|S|D|A|P|
            +-+-+-+-+-+-+
              Flags
```

Figure 5: 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. 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 reroutes 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.

4. IANA Considerations

4.1. 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:

+:	====	======	=+=		+==	==========	=+
•			•	Capability Name	•		٠
+:	====	======	=+=		+==	===========	=+
		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:

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		'		, ,	•			
+:	-======	======+	-=:	======	=======	:+:	=====		=+
	10	I		Binding	Segment		This	document	
+.		+				+			-+

IANA is requested to create and maintain new registries:

o 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
Θ	Reserved	
1	Binding Segment TLV	This Document
2-32767	Unassigned	
32768-65535	Reserved	

4.2. OSPFv3

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

IANA is requested to create and maintain new registries:

o 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	

4.3. 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:

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

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

5. Acknowledgements

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

6. References

6.1. 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>.

- [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.

6.2. Informative References

[I-D.hu-spring-segment-routing-proxy-forwarding]

Hu, Z., Chen, H., Yao, J., Bowers, C., Yongqing, and Yisong, "SR-TE Path Midpoint Restoration", Work in Progress, Internet-Draft, draft-hu-spring-segment-routing-proxy-forwarding-20, 18 August 2022, https://www.ietf.org/archive/id/draft-hu-spring-segment-routing-proxy-forwarding-20.txt.

[I-D.ietf-spring-segment-routing-policy]

Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", Work in Progress, Internet-Draft, draft-ietf-spring-segment-routing-policy-22, 22 March 2022, https://www.ietf.org/archive/id/draft-ietf-spring-segment-routing-policy-22.txt.

Authors' Addresses

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: <u>Huaimo.chen@futurewei.com</u>

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 United States of America

Email: cbowers@juniper.net

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

Email: <u>zhuyq8@chinatelecom.cn</u>

Yisong China Mobile 510000 China

Email: liuyisong@chinamobile.com