Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: June 1, 2014

H. Chen
Huawei Technologies
N. So
Tata Communications
A. Liu
Ericsson
F. Xu
Verizon
M. Toy
Comcast
L. Huang
China Mobile
L. Liu
UC Davis

November 28, 2013

Extensions to RSVP-TE for LSP Egress Local Protection
draft-chen-mpls-p2mp-egress-protection-10.txt

Abstract

This document describes extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for locally protecting egress nodes of a Traffic Engineered (TE) Label Switched Path (LSP) in a Multi-Protocol Label Switching (MPLS) and Generalized MPLS (GMPLS) network.

Status of this Memo

This Internet-Draft is submitted to IETF 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 http://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 June 1, 2014.

Copyright Notice

Copyright (c) 2013 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 (http://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

RFC 4090 describes two methods for protecting the transit nodes of a P2P LSP: one-to-one protection and facility bypass protection. RFC 4875 specifies how to use them to protect the transit nodes of a P2MP LSP. However, there is no mention of locally protecting any egress of a protected P2P or P2MP LSP in these RFCs.

To protect the egresses of an LSP (P2P or P2MP), an existing approach sets up a backup LSP from a backup ingress (or the ingress of the LSP) to the backup egresses, where each egress is paired with a backup egress and protected by the backup egress.

The main disadvantage of this approach is that more network resources such as double bandwidths may be used.

This document specifies extensions to RSVP-TE for locally protecting an egress of a P2MP or P2P LSP, which overcome this disadvantage.

1.1. An Example of Egress Local Protection

Figure 1 illustrates an example of using backup LSPs to locally protect egress nodes of a primary P2MP LSP, which is from ingress node R1 to two egress nodes: L1 and L2. The primary LSP is represented by star(*) lines and backup LSPs by hyphen(-) lines.

La and Lb are the designated backup egress nodes for egress nodes L1 and L2 of the P2MP LSP respectively. To distinguish between an egress (e.g., L1 in the figure) and a backup egress (e.g., La in the figure), an egress is called a primary egress if necessary.

The backup LSP for protecting primary egress L1 is from its upstream node R3 to backup egress La. The backup LSP for protecting primary egress L2 is from its upstream node R5 to backup egress Lb.

During normal operations, the traffic carried by the P2MP LSP is sent through R3 to L1, which delivers the traffic to its destination CE1. When R3 detects the failure of L1, R3 switches the traffic to the backup LSP to backup egress La, which delivers the traffic to CE1. The time for switching the traffic is within tens of milliseconds.

The failure of a primary egress (e.g., L1 in the figure) MAY be detected by its upstream node (e.g., R3 in the figure) through a BFD session between the upstream node and the egress in MPLS networks. Exactly how the failure is detected is out of scope for this document.

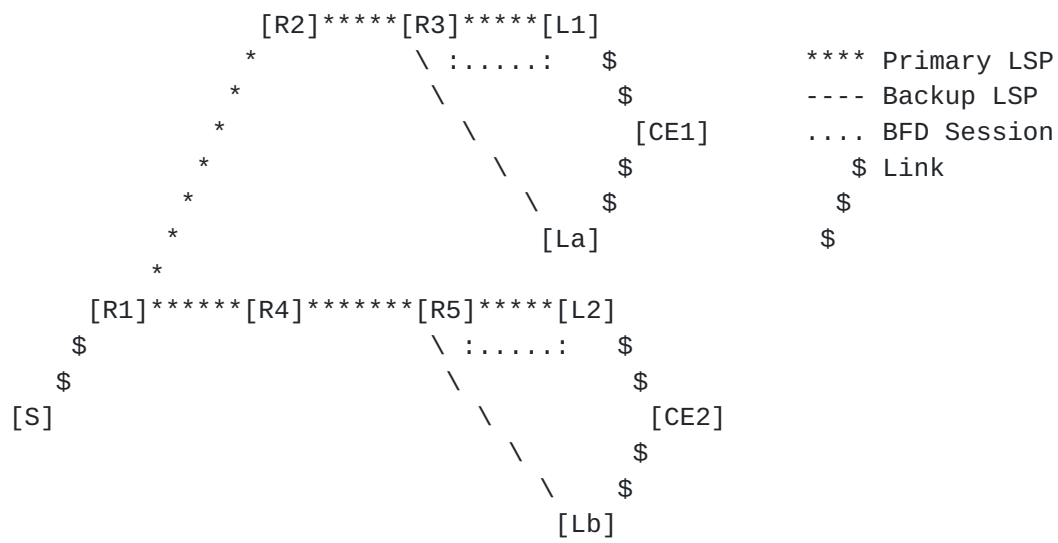


Figure 1: Backup LSP for Locally Protecting Egress

1.2. Egress Local Protection with FRR

Using the egress local protection and the FRR, we can locally protect the egresses, the links and the intermediate nodes of an LSP. The traffic switchover time is within tens of milliseconds whenever an egress, any of the links and the intermediate nodes of the LSP fails.

The egress nodes of the LSP can be locally protected via the egress local protection. All the links and the intermediate nodes of the LSP can be locally protected through using the FRR.

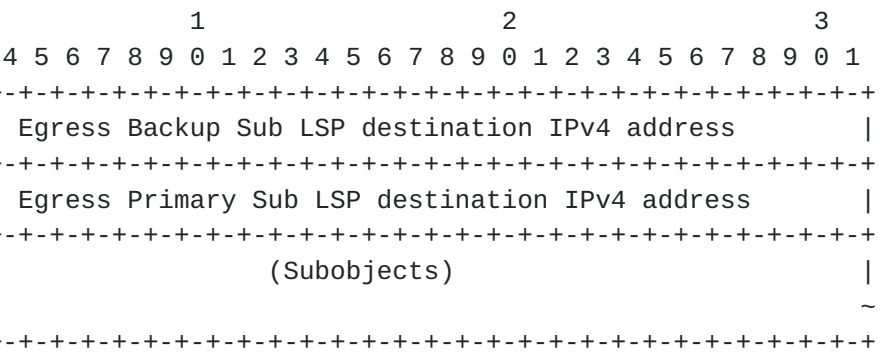
2. Protocol Extensions

A new object EGRESS_BACKUP_SUB_LSP is defined for signaling egress local protection. It contains a backup egress for a primary egress.

2.1. EGRESS_BACKUP_SUB_LSP IPv4/IPv6 Object

The class of the EGRESS_BACKUP_SUB_LSP IPv4/IPv6 object is 50, which is the same as that of the S2L_SUB_LSP IPv4/IPv6 object defined in RFC 4875. The C-Type of the EGRESS_BACKUP_SUB_LSP IPv4 object is a new number 3 or another number assigned by IANA.

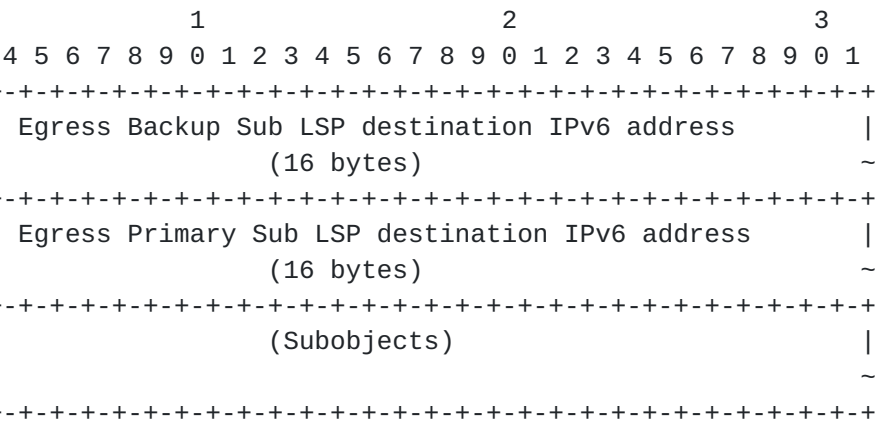
EGRESS_BACKUP_SUB_LSP_IPv4 C-Type = 3



Egress Backup Sub LSP destination IPv4 address
IPv4 address of the backup egress node
Egress Primary Sub LSP destination IPv4 address
IPv4 address of the primary egress node
Subobjects are optional

The C-Type of the EGRESS_BACKUP_SUB_LSP IPv6 object is a new number 4 or another number assigned by IANA.

EGRESS_BACKUP_SUB_LSP_IPv6 C-Type = 4



Egress Backup Sub LSP destination IPv6 address
IPv6 address of the backup egress node
Egress Primary Sub LSP destination IPv6 address
IPv6 address of the primary egress node
Subobjects are optional

2.2. EGRESS_BACKUP_SECONDARY_EXPLICIT_ROUTE Object

An EGRESS_BACKUP_SECONDARY_EXPLICIT_ROUTE (EB-SERO) object is defined for signaling protection for a primary egress of a P2MP LSP in a new S2L Sub LSP backup protection method. It contains a path from the

upstream node of the primary egress to a backup egress. Its format is identical to an ERO's.

The class of an EB-SERO is the same as that of a SERO defined in RFC 4873. The EB-SERO uses a new C-Type 3, or another number assigned by IANA. The formats of sub-objects in an EB-SERO are identical to those of sub-objects in an ERO defined in RFC 3209.

2.3. Path Message

A Path message is enhanced to carry the information about a backup egress for a primary egress of an LSP through including an egress backup sub LSP descriptor list. The format of the enhanced Path message is illustrated below.

```
<Path Message> ::=  <Common Header> [ <INTEGRITY> ]
                    [ [<MESSAGE_ID_ACK> | <MESSAGE_ID_NACK>] ...]
                    [ <MESSAGE_ID> ]
                    <SESSION> <RSVP_HOP> <TIME_VALUES>
                    [ <EXPLICIT_ROUTE> ]
                    <LABEL_REQUEST> [ <PROTECTION> ]
                    [ <LABEL_SET> ... ]
                    [ <SESSION_ATTRIBUTE> ] [ <NOTIFY_REQUEST> ]
                    [ <ADMIN_STATUS> ] [ <POLICY_DATA> ... ]
                    <sender descriptor>
                    [<S2L sub-LSP descriptor list>]
                    [<egress backup sub LSP descriptor list>]
```

The egress backup sub LSP descriptor list in the message is defined below. It is a sequence of EGRESS_BACKUP_SUB_LSP objects, each of which describes a pair of a primary egress and a backup egress.

```
<egress backup sub LSP descriptor list> ::=
                    <egress backup sub LSP descriptor>
                    [ <egress backup sub LSP descriptor list> ]

<egress backup sub LSP descriptor> ::=
                    <EGRESS_BACKUP_SUB_LSP>
                    [ <EGRESS_BACKUP_SECONDARY_EXPLICIT_ROUTE> ]
```

3. Egress Protection Behaviors

3.1. Ingress Behavior

To protect a primary egress of an LSP, a backup egress must be configured on the ingress of the LSP.

The ingress initiates a Path message for the LSP with an egress backup sub LSP descriptor list. For each primary egress of the LSP to be protected, the ingress MUST add an EGRESS_BACKUP_SUB_LSP object into the list. The object contains the primary egress and the backup egress for protecting the primary egress.

To protect a primary egress of an LSP via one-to-one backup or facility backup method, the ingress SHOULD include a FAST_REROUTE object and set the One-to-One Backup Desired or Facility Backup Desired flag.

To protect a primary egress of a P2MP LSP via S2L Sub LSP backup method, the ingress SHOULD add an EB-SERO object following the EGRESS_BACKUP_SUB_LSP object into the list. The EB-SERO object contains a path from the upstream node of the primary egress to the backup egress. The ingress computes the path if the P2MP LSP is in one area; otherwise, the path may be computed by the Path Computation Element (PCE).

3.2. Intermediate Node and PLR Behavior

If an intermediate node of an LSP receives the Path message with an egress backup sub LSP descriptor list and it is not an upstream node of any primary egress of the LSP, it forwards the list in the message unchanged.

If the intermediate node is the upstream node of a primary egress to be protected, it gets the backup egress for the primary egress from the EGRESS_BACKUP_SUB_LSP object in the list. It acts as a PLR to provide one-to-one or facility backup protection for the primary egress. It provides one-to-one backup protection if the One-to-One Backup Desired flag is set in the message; it provides facility backup protection if the Facility Backup Desired flag is set.

The PLR (upstream node of the primary egress) sets the protection flags in the RRO Sub-object for the primary egress in the Resv message according to the status of the primary egress and the backup LSP protecting the primary egress. For example, it will set the "local protection available" and the "node protection" flag to one indicating that the primary egress is protected when the backup LSP to the backup egress is set up for protecting the primary egress.

3.2.1. Signaling for One-to-One Protection

The behavior of the upstream node of a primary egress of an LSP as a PLR is the same as that of a PLR for one-to-one backup method described in RFC 4090 except for that the upstream node creates a backup LSP from itself to a backup egress.

In the case that the LSP is a P2MP LSP and a primary egress of the LSP is a transit node (i.e., bud node), the upstream node of the primary egress as a PLR also creates a backup LSP from itself to each of the next hops of the primary egress.

When the PLR detects a failure in the primary egress, it MUST rapidly switch the packets from the primary LSP to the backup LSP to the backup egress. For a failure in the bud node of an P2MP LSP, the PLR MUST also rapidly switch the packets to the backup LSPs to the bud node's next hops, where the packets are merged into the primary LSP.

3.2.2. Signaling for Facility Protection

Except for backup LSP and downstream label, the behavior of the upstream node of the primary egress as a PLR follows the PLR behavior for facility backup method described in RFC 4090.

For a number of primary P2P LSPs going through the same PLR to the same primary egress, the primary egress of these LSPs may be protected by one backup LSP from the PLR to the backup egress designated for protecting the primary egress.

The PLR selects or creates a backup LSP from itself to the backup egress. If there exists a backup LSP that satisfies the constraints given in the Path message, then this one is selected; otherwise, a new backup LSP to the backup egress will be created.

For a primary LSP carrying IP packets, the PLR does not need any downstream label as an inner label for the LSP when binding the primary LSP with the backup LSP. When the PLR detects a failure in the primary egress, it redirects the IP packets from the primary LSP into the backup LSP to the backup egress, where the IP packets are forwarded according to IP destinations in the packets.

For a primary LSP carrying packets with application or service labels, the PLR may not need any downstream label as an inner label for the LSP either when binding the primary LSP with the backup LSP. When the PLR detects a failure in the primary egress, it redirects the packets from the primary LSP into the backup LSP to backup egress through switching the top label with the backup LSP label. The backup egress delivers the packets to the same destinations as the

primary egress (see details in section "Considering Application Traffic" below).

3.2.3. Signaling for S2L Sub LSP Protection

The S2L Sub LSP Protection is used to protect a primary egress of a P2MP LSP. Its major advantage is that the application traffic carried by the P2MP LSP may be easily protected against the egress failure.

The PLR determines to protect a primary egress of a P2MP LSP via S2L sub LSP protection when it receives a Path message with an EB-SERO object following the EGRESS_BACKUP_SUB_LSP containing the primary egress and a backup egress.

The PLR sets up the backup S2L sub LSP to the backup egress, creates and maintains its state in the same way as of setting up a source to leaf (S2L) sub LSP defined in RFC 4875 from the signaling's point of view. It constructs and sends a PATH message along the path given in the EB-SERO for the backup LSP, receives and processes a RESV message that responses to the PATH message.

After receiving the RESV message for the backup LSP, the PLR creates a forwarding entry with an inactive state or flag called inactive forwarding entry. This inactive forwarding entry is not used to forward any data traffic during normal operations.

When the PLR detects a failure in the primary egress failure, it changes the forwarding entry for the backup LSP to active. Thus, the PLR forwards the traffic to the backup egress through the backup LSP, which sends the traffic to its destination.

3.2.4. PLR Procedures during Local Repair

When the upstream node of a primary egress of an LSP as a PLR detects a failure in the primary egress, it follows the procedures defined in section 6.5 of RFC 4090.

The PLR (i.e., the upstream node of the primary egress) SHOULD notify the ingress about the failure of the primary egress in the same way as a PLR notifies the ingress about the failure of an intermediate node.

In the local revertive mode, the PLR re-signals each of the primary LSPs that used to be routed over the restored resource once it detects that the resource is restored. Every primary LSP successfully re-signaled along the restored resource is switched back.

Moreover, the PLR lets the upstream part of the primary LSP stay after the primary egress fails. The downstream part of the primary LSP from the PLR to the primary egress SHOULD be removed.

4. Considering Application Traffic

This section focuses on the application traffic carried by P2P LSPs. When a primary egress of a P2MP LSP fails, the application traffic carried by the P2MP LSP may be delivered to the same destination by the backup egress since the inner label if any for the traffic is a upstream assigned label for every egress of the P2MP LSP.

4.1. A Typical Application

L3VPN is a typical application that an LSP carries. An existing solution (refer to Figure 2) for protecting L3VPN traffic against egress failure includes: 1) A multi-hop BFD session between ingress R1 and egress L1 of primary LSP; 2) A backup LSP from ingress R1 to backup egress La; 3) La sends R1 VPN backup label and related information via BGP; 4) R1 has a VRF with two sets of routes: one uses primary LSP and L1 as next hop; the other uses backup LSP and La as next hop.

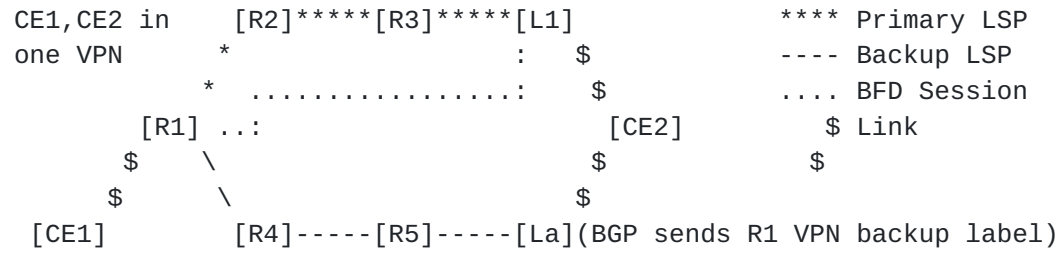


Figure 2: Protect Egress for L3VPN Traffic

In normal operations, R1 sends the traffic from CE1 through primary LSP with VPN label received from L1 as inner label to L1, which delivers the traffic to CE2 using VPN label.

When R1 detects a failure in L1, R1 sends the traffic from CE1 via backup LSP with VPN bakup label received from La as inner label to La, which delivers the traffic to CE2 using VPN backup label.

A new solution (refer to Figure 3) with egress local protection for protecting L3VPN traffic includes: 1) A BFD session between R3 and egress L1 of primary LSP; 2) A backup LSP from R3 to backup egress La; 3) L1 sends La VPN label as UA label and related information via BGP or another protocol; 4) L1 and La is virtualized as one from R1's

point of view.

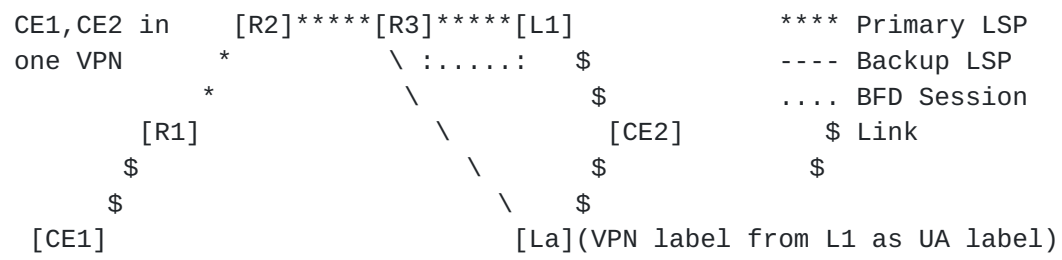


Figure 3: Locally Protect Egress for L3VPN Traffic

When R3 detects a failure in L1, R3 sends the traffic from primary LSP via backup LSP to La, which delivers the traffic to CE2 using VPN label under the backup LSP label as a context label.

4.2. PLR Procedure for Applications

When the PLR creates a backup LSP from itself to a backup egress for protecting a primary egress, it includes an EGRESS_BACKUP_SUB_LSP object in the Path message for the LSP. The object contains the primary egress and the backup egress and indicates that the backup egress SHOULD consider the backup LSP label as a context label and the inner label as application traffic label when needed.

4.3. Egress Procedures for Applications

When a primary egress of an LSP sends the ingress of the LSP a label for an application such as a VPN, it SHOULD sends the backup egress for protecting the primary egress the label as a upstream assigned label via BGP or another protocol. Exactly how the label is sent is out of scope for this document.

When the backup egress receives a upstream assigned label from the primary egress, it adds a forwarding entry with the label into the LFIB for the primary egress. Using this entry, the backup egress delivers the traffic with this label as inner label from the backup LSP to the same destination as the primary egress.

When the backup egress receives a packet from the backup LSP, it uses the top label as a context label to find the LFIB for the primary egress and the inner label to deliver the packet to the same destination as the primary egress according to the LFIB.

5. Security Considerations

In principle this document does not introduce new security issues. The security considerations pertaining to RFC 4090, RFC 4875 and other RSVP protocols remain relevant.

6. IANA Considerations

TBD

7. Contributors

Boris Zhang
Telus Communications
200 Consilium Pl Floor 15
Toronto, ON M1H 3J3
Canada

Email: Boris.Zhang@telus.com

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

Email: lizhenbin@huawei.com

Vic Liu
China Mobile
No.32 Xuanwumen West Street, Xicheng District
Beijing, 100053
China

8. Acknowledgement

The authors would like to thank Richard Li, Tarek Saad, Lizhong Jin, Ravi Torvi, Eric Gray, Olufemi Komolafe, Michael Yue, Rob Rennison, Neil Harrison, Kannan Sampath, Yimin Shen, Ronhazli Adam and Quintin Zhao for their valuable comments and suggestions on this draft.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3692] Narten, T., "Assigning Experimental and Testing Numbers Considered Useful", BCP 82, RFC 3692, January 2004.

[RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC4090] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

[RFC4875] Aggarwal, R., Papadimitriou, D., and S. Yasukawa, "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, May 2007.

[RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, August 2008.

[P2MP FRR] Le Roux, J., Aggarwal, R., Vasseur, J., and M. Vigoureux, "P2MP MPLS-TE Fast Reroute with P2MP Bypass Tunnels", draft-leroux-mpls-p2mp-te-bypass , March 1997.

9.2. Informative References

[RFC4461] Yasukawa, S., "Signaling Requirements for Point-to-Multipoint Traffic-Engineered MPLS Label Switched Paths (LSPs)", RFC 4461, April 2006.

Authors' Addresses

Huaimo Chen
Huawei Technologies
Boston, MA
USA

Email: huaimo.chen@huawei.com

Ning So
Tata Communications
2613 Fairbourne Cir.
Plano, TX 75082
USA

Email: ning.so@tatacommunications.com

Autumn Liu
Ericsson
CA
USA

Email: autumn.liu@ericsson.com

Fengman Xu
Verizon
2400 N. Glenville Dr
Richardson, TX 75082
USA

Email: fengman.xu@verizon.com

Mehmet Toy
Comcast
1800 Bishops Gate Blvd.
Mount Laurel, NJ 08054
USA

Email: mehmet_toy@cable.comcast.com

Lu Huang
China Mobile
No.32 Xuanwumen West Street, Xicheng District
Beijing, 100053
China

Email: huanglu@chinamobile.com

Lei Liu
UC Davis
USA

Email: liulei.kddi@gmail.com