Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: November 29, 2013

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

May 28, 2013

# Extensions to RSVP-TE for P2MP LSP Egress Local Protection
## draft-chen-mpls-p2mp-egress-protection-09.txt

## Abstract

This document describes extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for locally protecting egress nodes of a Traffic Engineered (TE) point-to-multipoint (P2MP) Label Switched Path (LSP) in a Multi-Protocol Label Switching (MPLS) and Generalized MPLS (GMPLS) network.

## Status of this Memo

## Copyright Notice

Table of Contents

## 1. Introduction

RFC 4090 "Fast Reroute Extensions to RSVP-TE for LSP Tunnels" describes two methods for protecting P2P LSP tunnels or paths at local repair points. The first method is a one-to-one protection method, where a detour backup P2P LSP for each protected P2P LSP is created at each potential point of local repair. The second method is a facility bypass backup protection, where a bypass backup P2P LSP tunnel is created using MPLS label stacking to protect a potential failure point for a set of P2P LSP tunnels. The bypass backup tunnel can protect a set of P2P LSPs having similar backup constraints.

RFC 4875 "Extensions to RSVP-TE for P2MP TE LSPs" describes how to use the one-to-one protection method and facility bypass backup protection method to protect a link or intermediate node failure on the path of a P2MP LSP. However, there is no mention of locally protecting any egress node failure in a protected P2MP LSP.

An existing method for protecting the egress nodes of a P2MP LSP sets up a backup P2MP LSP from a backup ingress node to the backup egress nodes, where each egress node is paired with a backup egress node and protected by the backup egress node. The backup P2MP LSP carries the same traffic as the P2MP LSP at the same time. A traffic receiver from the P2MP LSP is normally connected to an egress node and its paired backup egress node. It receives the traffic from the egress node in normal situations.

The receiver selects the egress or backup egress node for receiving the traffic according to the route to the source through RPF. In a normal situation, it selects the egress node. When the egress node fails, it selects the backup egress for receiving the traffic since the route to the source through the egress node is gone and the route to the source through the backup egress node is active.

The main disadvantage of this method is that double network resources such as double bandwidths are used for protecting the egress nodes since the backup P2MP LSP consumes the same amount of network resource as the primary P2MP LSP. The impact on network efficiency can be significant in case of large P2MP deployments.

This document proposes a new method to locally protect the egress nodes of a P2MP LSP, which is called Egress Local Protection. It specifies the mechanism and extensions to RSVP-TE for locally protecting an egress node of a Traffic Engineered (TE) point-to-multipoint (P2MP) Label Switched Path through using a backup P2MP sub LSP. The new method overcomes the disadvantages described above. The same extensions and mechanism can also be used to protect the egress node of a TE P2P LSP.

## 2. Terminology

This document uses terminologies defined in RFC 2205, RFC 3031, RFC 3209, RFC 3473, RFC 4090, RFC 4461, and RFC 4875.

## 3. Conventions Used in This Document

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.

## 4. Mechanism

This section briefly describes a solution that locally protects an egress node of a P2MP LSP through using a backup sub LSP. We first show an example, and then present different parts of the solution, which includes the creation of the backup sub LSP, the forwarding state for the backup sub LSP, and the detection of a failure in the egress node.

### 4.1. An Example of Egress Local Protection

Figure 1 below illustrates an example of using backup sub LSPs to locally protect egress nodes of a P2MP LSP. The P2MP LSP is from ingress node R1 to three egress nodes: L1, L2 and L3. It is represented by double lines in the figure.

La, Lb and Lc are the designated backup egress nodes for the egress nodes L1, L2 and L3 of the P2MP LSP respectively. In order to distinguish an egress node (e.g., L1 in the figure) and a backup egress node (e.g., La in the figure), an egress node is called a primary egress node in the follwing description.

The backup sub LSP used to protect the primary egress node L1 is from its previous hop node R3 to the backup egress node La. The backup sub LSP used to protect the primary egress node L2 is from its previous hop node R5 to the backup egress node Lb. The backup sub LSP used to protect the primary egress node L3 is from its previous hop node R5 to the backup egress node Lc via the intermediate node Rc.

During normal operation, the traffic transported by the P2MP LSP is forwarded through R3 to L1, then delivered to its destination CE1. When the failure of L1 is detected, R3 forwards the traffic to the backup egress node La, which then delivers the traffic to its destination CE1. The time for switching the traffic after L1 fails

is within tens of milliseconds.

L1's failure CAN be detected by a BFD session between L1 and R3.

```
                         [R2]=====[R3]=====[L1]----[CE1]
                        //           \            /
                       //             \          /
                      //               \___[La]_/
                     //
                    //
                   //
                  //
              +---[R1]====[R4]====[R5]=====[L2]----[CE2]
              |                   |\\ \           /
              |                   | \\ \         /
         [S]--|                   |  \\ \__[Lb]_/
              |                   |   \\
              |                   |    \\
                                  |     \\
                                  |      \\
                                  |       \\
                                  |        [L3]----[CE3]
                                  |               /
                                  |              /
                                [Rc]_______[Lc]_/
```

Figure 1: Sub LSP for Locally Protecting Egress

## 4.2. Set up of Backup sub LSP

A backup egress node is designated for a primary egress node of a LSP. The previous hop node of the primary egress node sets up a backup sub LSP from itself to the backup egress node after receiving the information about the backup egress node.

The previous hop node sets up the backup sub LSP, creates and maintains its state in the same way as of setting up a source to leaf (S2L) sub LSP from the signalling's point of view. It constructs and sends a RSVP-TE PATH message along the path for the backup sub LSP, receives and processes a RSVP-TE RESV message that responses to the PATH message.

## 4.3. Forwarding State for Backup sub LSP(s)

The forwarding state for the backup sub LSP is different from that for a P2MP S2L sub LSP. After receiving the RSVP-TE RESV message for the backup sub LSP, the previous hop node 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. It SHALL only be used after the failure of the primary egress node.

Upon detection of the primary egress node failure, the state or flag of the forwarding entry for the backup sub LSP is set to be active. Thus, the previous hop node of the primary egress node will forward the traffic to the backup egress node through the backup sub LSP, which then sends the traffic to its destination.

### 4.4. Detection of Egress Node Failure

The previous hop node of the primary egress node SHALL detect the failures described below:

- The failure of the primary egress node (e.g. L1 in Figure 1)

- The failure of the link between the primary egress node and its previous hop node (e.g. the link between R3 and L1 in Figure 1)

- The failure of the link between the primary egress node and its destination node (e.g. the failure of the link between L1 and CE1 in Figure 1).

Failure of the primary egress node and the link between itself and its previous hop node CAN be detected through a BFD session between itself and its previous hop node in MPLS networks.

In the GMPLS networks where the control plane and data plane are physically separated, the detection and localization of failures in the physical layer can be achieved by introducing the link management protocol (LMP) or assisting by performance monitoring devices.

Failure of the destination node and the link between the primary egress node and the destination node CAN be detected by a BFD session between the previous hop node and the destination node.

Upon detecting any above mentioned failures, the previous hop node imports the traffic from the LSP into the backup sub LSP. The traffic is then delivered to its destination through the backup egress node.

When we use the egress local protection to protect a primary egress node, we SHOULD NOT use any fast re-route to protect the link between the primary egress node and its previous hop node. The failure of the link is protected by the egress protection.

## 5. Egress Local Protection with FRR

RFC4875 "Extensions to RSVP-TE for P2MP TE LSPs" describes how to use RFC 4090 "Fast Reroute Extensions to RSVP-TE for LSP Tunnels" (FFR for short) to locally protect failures in a link or intermediate node of a P2MP LSP. However, there is not any standard that locally protects the egresses of the P2MP LSP. The egress local protection mechanism proposed in this document fills this gap. Thus, through using the egress local protection and the FRR, we can locally protect the egress nodes, all the links and the intermediate nodes of a P2MP LSP. The traffic switchover time is within tens of milliseconds whenever any of the egresses, the links and the intermediate nodes of the P2MP LSP fails.

All the egress nodes of the P2MP LSP can be locally protected through using the egress local protection. All the links and the intermediate nodes of the LSP can be locally protected by using the FRR. Note that the methods for locally protecting all the links and the intermediate nodes of a P2MP LSP are out of scope of this document.

## 6. Representation of a Backup Sub LSP

A backup sub LSP exists within the context of a P2MP LSP in a way similar to a S2L sub LSP. It is identified by the P2MP LSP ID, Tunnel ID, and Extended Tunnel ID in the SESSION object, the tunnel sender address and LSP ID in the SENDER_TEMPLATE object, and the backup sub LSP destination address in the EGRESS_BACKUP_SUB_LSP object (to be defined in the section below).

An EGRESS_BACKUP_SECONDARY_EXPLICIT_ROUTE Object (EB-SERO) is used to specify the explicit route of a backup sub LSP that is from a previous hop node to a backup egress node. The EB-SERO is defined in the following section.

### 6.1. EGRESS_BACKUP_SUB_LSP Object

An EGRESS_BACKUP_SUB_LSP object identifies a particular backup sub LSP belonging to the LSP.

#### 6.1.1. EGRESS_BACKUP_SUB_LSP IPv4 Object

The class of the EGRESS_BACKUP_SUB_LSP IPv4 object is the same as that of the S2L_SUB_LSP IPv4 object defined in RFC 4875. The C-Type of the object is a new number 3, or may be another number assigned by Internet Assigned Numbers Authority (IANA).

```
    EGRESS_BACKUP_SUB_LSP Class = 50,
    EGRESS_BACKUP_SUB_LSP_IPv4 C-Type = 3

        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
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |         Egress Backup Sub LSP IPv4 destination address        |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                     Egress IPv4 address                       |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+


    Egress Backup Sub LSP IPv4 destination address
       IPv4 address of the backup sub LSP destination is the backup
       egress node.
    Egress IPv4 address
       IPv4 address of the egress node
```

### 6.1.2. EGRESS_BACKUP_SUB_LSP IPv6 Object

The class of the EGRESS_BACKUP_SUB_LSP IPv6 object is the same as that of the S2L_SUB_LSP IPv6 object defined in RFC 4875. The C-Type of the object is a new number 4, or may be another number assigned by Internet Assigned Numbers Authority (IANA).

```
    EGRESS_BACKUP_SUB_LSP Class = 50,
    EGRESS_BACKUP_SUB_LSP_IPv6 C-Type = 4

        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
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |         Egress Backup Sub LSP IPv6 destination address        |
       |                         (16 bytes)                            |
       |                         ....                                  |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                     Egress IPv6 address                       |
       |                         (16 bytes)                            |
       |                         ....                                  |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+


    Egress Backup Sub LSP IPv6 destination address
       IPv6 address of the backup sub LSP destination is the backup
       egress node.
    Egress IPv6 address
       IPv6 address of the egress node
```

### 6.2. EGRESS_BACKUP_SECONDARY_EXPLICIT_ROUTE Object

The format of an EGRESS_BACKUP_SECONDARY_EXPLICIT_ROUTE (EB-SERO) object is defined as identical to that of the ERO. The class of the EB-SERO is the same as that of the SERO defined in RFC 4873. The EB-SERO uses a new C-Type 3, or may use another number assigned by Internet Assigned Numbers Authority (IANA). The formats of sub-objects in an EB-SERO are identical to those of sub-objects in an ERO defined in RFC 3209.

## 7. Path Message

This section describes extensions to the Path message defined in RFC 4875. The Path message is enhanced to transport the information about a backup egress node to the previous hop node of a primary egress node of a P2MP LSP through including an egress backup sub LSP descriptor list.

### 7.1. Format of Path Message

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 format of the egress backup sub LSP descriptor list in the enhanced Path message is defined as follows.

```
      <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> ]
```

### 7.2. Processing of Path Message

The ingress node of a LSP initiates a Path message with an egress backup sub LSP descriptor list for protecting primary egress nodes of the LSP. In order to protect a primary egress node of the LSP, the ingress node MUST add an EGRESS_BACKUP_SUB_LSP object into the list. The object contains the information about the backup egress node to be used to protect the failure of the primary egress node. An EGRESS_BACKUP_SECONDARY_EXPLICIT_ROUTE object (EB-SERO), which describes an explicit path to the backup egress node, SHALL follow the EGRESS_BACKUP_SUB_LSP.

#### 7.2.1. Backup LSP for One-to-One Protection

If the previous hop node of the primary egress node receives the Path message with an egress backup sub LSP descriptor list and the request for protection via the one-to-one backup method, it generates a new Path message based on the information in the EGRESS_BACKUP_SUB_LSP and EB-SERO containing the backup egress node.

The format of this new Path message is the same as that of the Path message defined in RFC 4875. This new Path message is used to signal the segment of a special S2L sub-LSP from the previous hop node to the backup egress node. The new Path message is sent to the next-hop node along the path for the backup sub LSP.

If an intermediate node receives the Path message with an egress backup sub LSP descriptor list, it MUST put the EGRESS_BACKUP_SUB_LSP containing a backup egress into a Path message to be sent towards the backup egress. This SHALL be done for each EGRESS_BACKUP_SUB_LSP containing a backup egress node in the list.

When a primary egress node of the LSP receives the Path message with an egress backup sub LSP descriptor list, it SHOULD ignore the egress backup sub LSP descriptor list and generate a PathErr message.

### 7.2.2. Backup LSP for Facility Protection

The facility backup method will be used for locally protecting a primary egress node if the previous hop node of the primary egress node receives the Path message with an egress backup sub LSP descriptor list and the request for protection via the facility backup method.

The previous hop node selects or creates a backup LSP tunnel from itself to the backup egress designated for protecting the primary egress. If there exists a backup LSP tunnel from itself to the backup egress that satisifies the constraints given in the PATH message, then this tunnel is selected; otherwise, a new backup LSP tunnel to the backup egress will be created.

After having a backup LSP tunnel, the previous hop node assigns the label allocated by the backup egress for the backup LSP as a top label (or called backup label).

When the previous hop node detects a failure in the primary egress, it has to imports the traffic for the protected P2MP LSP into the backup bypass tunnel using the backup label as the top label.

## 8. Processing of Resv Message

The format of the Resv Message is not changed. The processing of the Resv Message at the previous hop of a primary egress node is enhanced for reporting the status of the primary egress protection.

The previous hop node of the primary egress node sets the protection flags in the RRO IPv4/IPv6 Sub-object for the primary egress node according to the status of the primary egress node and the backup LSP protecting the primary egress node. For example, it will set the node protection bit to one indicating that the primary egress node is protected when the backup LSP to the backup egress node is set up for protecting the primary egress node. It will set the bandwidth protection bit to one when the backup LSP guarantees to provide the desired bandwidth that is specified in the FAST_REROUTE object or the bandwidth of the protected LSP.

## 9. IANA Considerations

TBD

## 10. Contributors

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

Email: Boris.Zhang@telus.com

## 11. Acknowledgement

The authors would like to thank Richard Li, Olufemi Komolafe, Michael Yue, Zhenbin Li, Rob Rennison, Neil Harrison, Kannan Sampath, Yimin Shen, Ronhazli Adam and Quintin Zhao for their valuable comments and suggestions on this draft.

## 12. References

### 12.1. 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.

[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.

## 12.2. 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

Lei Liu
UC Davis
USA

Email: liulei.kddi@gmail.com