Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 20, 2011

M. Bhatia
Alcatel-Lucent
L. Jin
ZTE
F. Jounay
France Telecom

May 19, 2011

Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Bi-directional Label Switched Paths (LSPs)
draft-bhatia-mpls-rsvp-te-bidirectional-lsp-01

Abstract

There are several applications that require symmetric Multiprotocol Label Switching (MPLS) path between two points. This cannot be achieved with regular MPLS as the LSPs are unidirectional. If symmetry is required, a separate LSP in each direction is required for bidirectional traffic flow. Generalized MPLS on the other hand, has provisions for setting up a bidirectional LSP. This document uses the extensions introduced for GMPLS and applies it to regular MPLS for establishing bidirectional LSPs. Additionally, it also describes how bi-directional symmetrical Fast Reroute using both one-to-one and facility backup can be achieved.

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].

When used in lower case, these words convey their typical use in common language, and are not to be interpreted as described in RFC2119 [RFC2119].

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

There are several applications that require symmetrical paths between a pair of speakers. One such application is 1588 [IEEE-1588] which requires that the Delay_Resp message takes the same path as the associated Delay_Req message. [I-D.ietf-tictoc-1588overmpls] describes a method for transporting PTP messages (PDUs) over an MPLS network to enable proper handling of these packets. Currently, the only way to ensure that the different PTP messages follow a symmetrical path is by statically configuring the RSVP-TE LSPs. This is unscalable and will not work in case of network failures as MPLS FRR may not guarantee symmetrical alternate paths.

This document describes how RSVP-TE can be used for setting up bi-directional LSPs for regular MPLS and the extensions required in FRR to ensure that the alternate paths are also symmetrical.

2. RSVP-TE to signal Bi-directional LSP

[RFC3473] describes a point-to-point bidirectional LSP mechanism for the GMPLS architecture, where a bidirectional LSP setup is indicated by the presence of an Upstream Label in the Path message. The Upstream_Label object has the same format as the generalized label, and uses Class-Number 35 (of form 0bbbbbbb) and the C-Type of the label being used.

For regular MPLS the Upstream_Label object will be used with C-Type value of 1.

Typically, a node initiates an RSVP session by adding the RRO to the Path message. The initial RRO contains only one subobject - the sender's IP addresses. If the node also desires label recording, it sets the Label_Recording flag in the SESSION_ATTRIBUTE object. This document extends this mechanism by also adding the Upstream label that has been advertised in the RRO subobject. Thus the initial RRO will now contain the sender's IP address and the Upstream label advertised by it. The upstream label subobject in RRO object will be with type 0x04 and same C-type with label object.

It is necessary to ensure the PLR and MP to bind to the same bidirectional protection tunnel (bypass tunnel or detour tunnel), this draft introduces a new subobject in RRO object to indicate the tunnel that PLR or MP binds.

```
       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
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |     Type      |     Length    |         Tunnel ID             |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                     Extended Tunnel ID                        |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |           LSP ID              |         Reserved              |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-++-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type 0x05 Protection bidirectional tunnel ID

Length The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length is always 8.

The tunnel ID and extended tunnel ID is derived from session object, LSP ID is derived from sender_template object of protection LSP.

3. Fast Reroute mechanisms

[RFC4090] extensions can be used to perform fast reroute for the mechanism described in this document when applied within packet networks. This section only applies to LSRs that support [RFC4090].

This section uses terminology defined in [RFC4090], and fast reroute procedures defined in [RFC4090] MUST be followed unless specified below. The head-end and transit LSRs MUST follow the SESSION_ATTRIBUTE and FAST_REROUTE object processing as specified in [RFC4090] for each Path message.

Since its a bi-directional LSP the detour LSPs and the bypass tunnels that are used for the protected LSP must also be bi-directional. This is required so that path symmetry is maintained even in an event of a network failure.

It should be noted that in case of bi-directional LSPs, the LSRs involved will play the role of both the Point-of-Local-Repair (PLR) and Merge Point (MP) at the same time during the failure. The router that is the PLR will become the MP for the traffic thats coming from the opposite direction.

In the Figure 1 assume that ABCD is the protected LSP. For protecting link BC, there is a bidirectional bypass tunnel BEFC (or a detour LSP in case of 1-on-1). B is the PLR and C is the MP for the traffic flowing from A towards D and B is the MP, and C the PLR for traffic flowing in the opposite direction (from D towards A).

```
         A - B ------- C - D
             |         |
             +- E - F -+

         Fig 1: Topology for
          link protection
```

In the Figure 2 ABCDE is the protected LSP and BFGD is the bypass tunnel for protecting the node C. In this case B is the PLR and D the MP for traffic from A towards E, and the roles reverse, i.e. B becomes the MP and D the PLR for traffic flowing in the opposite direction (from E towards A).

```
         A - B -- C -- D - E
             |         |
             +- F - G -+

          Fig 2: Topology for
            node protection
```

3.1. Discovering Upstream Labels

To support facility backup, the PLR must determine a label that will indicate to the MP that packets received with that label should be switched along the protected LSP. This can be done without explicitly signaling the backup path if the MP uses a label space global to that LSR.

As described in [RFC4090], the head-end LSR MUST set the "label recording requested" flag in the SESSION_ATTRIBUTE object for LSPs requesting local protection. This will cause (as specified in [RFC3209]) all LSRs to record their INBOUND labels and to note via a flag whether the label is global to the LSR. Thus, when a protected LSP is first signaled through a PLR, the PLR can examine the RRO in the Resv message and learn about the incoming labels that are used by all downstream nodes for this LSP. Similarly the MP, which will become the PLR for the reverse direction will learn about the upstream labels that are being used by the upstream nodes for this LSP by examining the upstream label subobject in RRO in the Path message.

The bypass tunnels and the detour tunnels that are set up for a bidirectional LSP must be bidirectional as well. They can internally use the Upstream_label technique that was described earlier to establish a bidirectional LSP.

3.2. Failure detection between PLR and MP

It is required that PLR and MP should detect the failure at the same time, then the two nodes could switch with traffic to the protection tunnel (bypass tunnel or detour tunnel) simultaneously. Such kind of detection mechanism could be BFD [RFC5880], RSVP-TE hello, or other proper mechanism.

For the link protection scenario, the detection mechanism should be enabled between PLR and MP. When a failure happens, both PLR and MP could detect the failure simultaneously, and switch the traffic to the protection tunnel.

For the node protection scenario, it is required to setup two co-related detection sessions. For the figure2 topology in section 3, the PLR node B and MP node D will do the node protection for the protected tunnel. There will be a detection session1 on the link between B and C, and session2 between C and D.. When a link failure happens between B and C, B could detect the failure by the session1, C should notify the link failure event to D by setting the diagnostic code to 6 (Concatenated Path Down) in BFD control packet [RFC5880]. Then D could detect failure through BFD control packets in session 2.

An alternative way is to do protected LSP segment detection between PLR and MP. When the link or node failed, the protected LSP segment detection session will be down, and both PLR and MP could detect the failure.

4. Behavior of various network elements in FRR

When a failure happens in the network the PLR router closest to the failure must perform the traffic protection. The MP router is the router that is the next hop to the failure point and merges the protected traffic back to the original path. In case of bidirectional LSPs, the same LSR is PLR in one direction and the MP for the other. Let us examine in detail what each network element does for the MPLS FRR.

4.1. The Head-End Router Behavior

The Head-End router originates the bi-directional LSP that needs to be protected. It's here that the desired protection type (one-to-one or facility backup) is also defined.

The Path message which has the FRR information in the SESSION_ATTRIBUTE object is propagated from the head-end LSR to the Tail router. Each hop sees the FRR flags and assumes the PLR role

and tries to establish a bi-directional tunnel. Every hop reports the availability of the FRR protection if its able to establish a bi-directional tunnel successfully. This is done via setting the RRO flags in the Resv message.

When a network failure occurs the PLR, or router upstream of the failure to be precise, uses FRR to reroute the traffic around the failure, and notifies the head-end LSR by (i) setting the FRR "Local protection in use" flag (0x2) in the RRO object of the Resv message and (ii) by sending a PathErr message with an ERROR object with code 0x19 - RSVP Notify Error and error value 0x3 - Tunnel locally repaired. The router that is downstream of the failure (traditionally the MP in case of unidirectional LSPs) also uses FRR to reroute the traffic around the failure. It does not send any message to the head-end LSR.

The head-end LSR upon recieving this indication tries to switch the traffic to a secondary LSP if its available. In case its not active, the head-end LSR signals this LSP via make-before-break mechanism.

4.2. The Point of Local Repair (PLR) Behavior

The PLR router of the protected LSP is also the origination point (head-end Router) of the protection tunnel (detour LSP or bypass tunnel). It is also the MP for the reverse protection tunnel at the same time. When an intermediate LSR receives a Path message carrying a SESSION_ATTRIBUTE with the FRR flags set, it assumes the role of a PLR and starts signaling a bi-directional FRR protection tunnel. In case facility backup is requested by the head-end LSR, the PLR signals a new bi-directional tunnel only if a bypass tunnel fulfilling the requirements does not already exist.

In the sections that follow the terms upstream and downstream are used in reference to the direction of traffic flow from the head-end towards the tail end. Thus router the tail router is downstream to the head-end.

When a network failure happens, the upstream router local to the failure assumes the role of the PLR and switches the traffic to the protection tunnel. This PLR is from now on referred to as the "upstream PLR". The downstream router, local to the failure also assumes the role of the PLR and switches the traffic to the bidirectional protection tunnel that is set up. This PLR is referred to as the "downstream PLR". These routers can use either the bi-directional detour LSP or a bi-directional bypass tunnel, depending upon what was requested by the head-end LSR.

The egress label that each PLR uses depends upon the kind of

protection provided. The subsequent sections only describe the behavior of the "upstream PLR" that is different with the protection mechanisms as described in [RFC4090].

Once the traffic gets switched to the protection path, the "downstream" PLR does not need to inform the HE router about the network failure.

4.2.1. PLR Behavior during one-to-one backup for a node failure

For the one-to-one backup, MP should bind the backup tunnel to protected LSP before replying the RESV message of detour LSP. When the PLR setup the detour LSP and bind to the protected LSP successfully, that also indicates that MP has bound successfully.

In case of one-to-one backup, the protection or the detour tunnel is a regular LSP. The downstream PLR uses the label that was distributed by the immediate upstream router on the detour LSP (detour label) to detour traffic arriving from the downstream router of the protected LSP. The label arriving from the immediate downstream router of the protected tunnel is swapped with the detour label, and the traffic is sent through the detour LSP.

```
         Head   Upstream         Downstream    Tail
         End      PLR              PLR        Router

             <-uL1   <-uL2     <-uL3     <- uL4
          A ------ B ------ C ------ D --------- E
                   |                 |
                 | |                 | ^
            udL1 | |                 | | udL2
                 V |                 | |
                   |                 |
                   +------- F -------+
```

Fig 3: one-to-one FRR protection

ABCDE is a bi-directional protected LSP
 BFD is a bi-directional detour LSP

The above figure describes this mechanism. udL1 is the Upstream_label advertised by B when setting up the bi-directional detour LSP from B to D. Similarly, udL2 is the Upstream_label advertised by F to D, when setting up this LSP. uL1, uL2, uL3 and uL4 are the Upstream_labels advertised when setting up the bi-directional protected tunnel ABCDE.

When a network failure happens, in this case the LSR router between B and D fails, the node D will assume the role of a downstream PLR and would need to switch the traffic from the protected LSP to the detour LSP. D does this by programming a Swap operation on the egress of the protected LSP path to the egress of the detour LSP. The uL4 label is thus swapped with udL2 during the failure, instead of label uL3.

4.2.2. PLR Behavior during facility backup for a node failure

For the facility backup, when the PLR successfully bind the protection tunnel to the protected LSP, it SHOULD insert the Protection Tunnel subobject in RRO object in the path message, and send downstream.

In the case of facility backup, the data from the protected LSP is tunneled through the bypass tunnel. Therefore, the outer label of the tunneled packet in the reverse direction is the label distributed by the immediate upstream router of the bypass tunnel. The "downstream PLR" also needs to know what label was expected by the router where this tunneled traffic merges (MP) at the upstream. The record label option makes this information available from the RRO in the Path messages for the protected LSP. This is the inner label that must be in the tunneled packet. Thus, the "downstream PLR" swaps the incoming label from the immediate downstream router in the protected path with these two labels and sends the path through the bypass tunnel.

```
         Head   Upstream         Downstream    Tail
         End      PLR               PLR       Router

             <-uL1   <-uL2     <-uL3     <- uL4
          A ------ B ------ C ------ D --------- E
                   |                 |
                 | |                 | ^
            udL1 | |                 | | udL2
                 V |                 | |
                   |                 |
                   +------- F -------+
```

Fig 4: Facility backup protection

ABCDE is a bi-directional protected LSP
BFD is a bi-directional bypass tunnel

The above figure describes the topology and the labels exchanged.

udL1 is the Upstream_label advertised by B when setting up the bi-directional facility bypass tunnel from B to D. Similarly, udL2 is the Upstream_label advertised by F to D, when setting up this tunnel. uL1, uL2, uL3 and uL4 are the Upstream_labels advertised when setting up the bi-directional protected tunnel ABCDE.

The "downstream" PLR router (D in this case) knows the label (uL2 in this case) that the upstream NNHop router expects because it has received a Path message which had this Upstream_label recorded in the RRO.

When a network failure happens, in this case the LSR router between B and D fails, the node D will assume the role of a "downstream" PLR and reroutes the traffic from the protected LSP through the bypass tunnel as follows:

- PLR D performs a swap operation to change the transport label. Since it knows that its doing node protection over the bypass tunnel, it will use the label that the NNHop router ("upstream" MP) expects instead of the label that the Nhop router (failed LSR) expects. D thus, swaps out uL4 and replaces it with uL2, instead of uL3 as it would normally have done.

- D also pushes the label udL2 on top of the label stack. This label would be used to switch the packet on the bypass tunnel and would finally reach the MP, which happens to be B in our case.

4.3. The Merge Point (MP) Router Behavior

The MP router is the LSR where the protection tunnel (detour LSP or bypass tunnel) and the protected LSP meet. It is the termination point (Tail router) of the protection tunnel. For a bi-directional protection tunnel the MP router in one direction becomes the PLR in the other.

```
         Head   Upstream         Downstream    Tail
         End       MP                MP        Router

          A ------ B ------ C ------ D --------- E
                   |                 |
                   |                 |
                   |                 |
                   |                 |
                   |                 |
                   +------- F -------+
```

Fig 5: Merge Points in bi-directional FRR

ABCDE is a bi-directional protected LSP
BFD is a bi-directional protection tunnel

Figure 5 shows two MPs associated with a bi-directional protection tunnel. This document refers to the MP defined in [RFC4090] as a downstream MP. This document does not change the behavior of the downstream MP. This means that it is still responsible for maintaining the protection tunnel's state by sending the Resv messages to the PLR and is also responsible for maintaining the state of the protected tunnel during the network failure. The upstream MP, defined in this document, is not required to do any of these. Its only responsible for merging the reverse traffic back to the protected path.

In one-to-one backup, the tail-end of backup LSP should consider it as MP. When the tail-end receives the Path message, and before sending RESV, it should try to bind the backup tunnel to protected tunnel. When binding successfully, MP sends the RESV message upstream for the backup tunnel.

During one-to-one backup the MP performs a swap operation on the ingress label of bi-directional detour LSP with the egress label of the bi-directional protected LSP.

In facility protection, when the LSR receives the Path message with RRO object, indicating the Previous_Hop or Previous_Previous_Hop with Protection Tunnel subobject, it should consider itself as MP. And it SHOULD try to bind the same protection tunnel indicated by Protection Tunnel subobject to the protected LSP. The protection tunnel would be expected to be from MP to PLR, with same tunnel-ID and LSP-ID indicated by the subobject.

During facility protection, the traffic arrives with a bypass tunnel label. The MP pops out this label to expose the original protected tunnel label that was distributed to the immediate downstream router

via the Upstream_label mechanism in the Path message on the protected tunnel. Since this label is already programmed, the traffic is switched out correctly.

The Resv message from MP to PLR should be sent in the protection LSP since there is a LSP path from MP to PLR.

5. Security Considerations

This document raises no new security concerns.

6. IANA Considerations

No requests for IANA at this point of time.

7. References

7.1. Normative References

[IEEE-1588]
"IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", 2008.

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

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

[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.

[RFC5467] Berger, L., Takacs, A., Caviglia, D., Fedyk, D., and J. Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label Switched Paths (LSPs)", RFC 5467, March 2009.

7.2. Informative References

[I-D.ietf-tictoc-1588overmpls]
Davari, S., Oren, A., Martini, L., Bhatia, M., and P. Roberts, "Transporting PTP messages (1588) over MPLS

Networks", draft-ietf-tictoc-1588overmpls-00 (work in progress), January 2011.

[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.

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

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

Authors' Addresses

Manav Bhatia
Alcatel-Lucent
India

Email: manav.bhatia@alcatel-lucent.com

Lizhong Jin
ZTE
889, Bibo Road
Shanghai, 201203, China

Email: lizhong.jin@zte.com.cn

Frederic Jounay
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex, FRANCE

Email: frederic.jounay@orange-ftgroup.com