CCAMP Working Group
Internet-Draft
Intended Status: Standards Track
Expires: March 9, 2015

Mike Taillon
Tarek Saad, Ed.
Rakesh Gandhi, Ed.
Zafar Ali
(Cisco Systems, Inc.)
Manav Bhatia
Lizhong Jin
()
Frederic Jounay
(Orange CH)

September 5, 2014

# Extensions to Resource Reservation Protocol For Fast Reroute of Traffic Engineering GMPLS LSPs
## draft-tsaad-ccamp-rsvpte-bidir-lsp-fastreroute-05

Abstract

This document defines Resource Reservation Protocol - Traffic Engineering (RSVP-TE) signaling extensions to support Fast Reroute (FRR) of Packet Switched Capable (PSC) Generalized Multi-Protocol Label Switching (GMPLS) Label Switched Paths (LSPs). These signaling extensions allow the coordination of bidirectional bypass tunnel assignment protecting a common facility in both forward and reverse directions of a co-routed bidirectional LSP. In addition, these extensions enable the re-direction of bidirectional traffic and signaling onto bypass tunnels that ensure co-routedness of data and signaling paths in the forward and reverse directions after FRR to avoid RSVP soft-state timeout.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

Copyright and License Notice



Table of Contents

## 1. Introduction

Packet Switched Capable (PSC) bidirectional Traffic Engineering (TE) tunnels are signaled using Generalized Multi-Protocol Label Switching (GMPLS) signaling procedures specified in [RFC3473]. Fast Reroute (FRR) [RFC4090] has been widely deployed in the packet TE networks today and is preferred for bidirectional TE tunnels. Using FRR also allows to leverage existing mechanisms for failure detection and restoration in the deployed networks.

FRR procedures defined in [RFC4090] describe the behavior of the Point of Local Repair (PLR) to reroute traffic and signaling onto the bypass tunnel in the event of a failure for unidirectional LSPs. These procedures are applicable to unidirectional protected LSPs signaled using either RSVP-TE [RFC3209] or GMPLS procedures [RFC3473], however don't address issues that arise when employing FRR for bidirectional co-routed GMPLS Label Switched Paths (LSPs).

When bidirectional bypass tunnels are used to locally protect bidirectional co-routed GMPLS LSPs, the upstream and downstream PLRs may independently assign different bidirectional bypass tunnels in the forward and reverse directions. There is no mechanism in FRR procedures defined in [RFC4090] to coordinate the bidirectional bypass tunnel selection between the downstream and upstream PLRs.

When using FRR procedures with bidirectional co-routed GMPLS LSPs, it is possible in some cases (e.g. when using node protection bypass tunnels post a link failure event and when RSVP signaling is sent in-fiber and in-band with data), the RSVP signaling refreshes may stop reaching some nodes along the primary bidirectional LSP path after the PLRs complete rerouting traffic and signaling onto the bypass tunnels. This is caused by the asymmetry of paths that may be taken by the bidirectional LSP's signaling in the forward and reverse directions after FRR reroute. In such cases, the RSVP soft-state timeout eventually causes the protected bidirectional LSP to be destroyed, and consequently impacts protected traffic flow after FRR.

This document proposes solutions to the above mentioned problems by providing mechanisms in the control plane to complement FRR procedures of [RFC4090] in order to maintain the RSVP soft-state for bidirectional co-routed protected GMPLS LSPs and achieve symmetry in the paths followed by the traffic and signaling in the forward and reverse directions post FRR. The document further extends RSVP signaling so that the bidirectional bypass tunnel selected by the upstream PLR matches the one selected by the downstream PLR node for a bidirectional co-routed LSP.

Unless otherwise specified in this document, fast reroute procedures defined in [RFC4090] are not modified for bidirectional tunnels.

## 2. Terminology

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].

The reader is assumed to be familiar with the terminology in [RFC2205] and [RFC3209].

LSR: Label-Switch Router.

LSP: An MPLS Label-Switched Path. In this document, an LSP will always be explicitly routed.

Local Repair: Techniques used to repair LSP tunnels quickly when a node or link along the LSP's path fails.

PLR: Point of Local Repair. The head-end LSR of a bypass tunnel or a detour LSP.

Protected LSP: An LSP is said to be protected at a given hop if it has one or multiple associated bypass tunnels originating at that hop.

Bypass Tunnel: An LSP that is used to protect a set of LSPs passing over a common facility.

NHOP Bypass Tunnel: Next-Hop Bypass Tunnel. A bypass tunnel that bypasses a single link of the protected LSP.

NNHOP Bypass Tunnel: Next-Next-Hop Bypass Tunnel. A bypass tunnel that bypasses a single node of the protected LSP.

MP: Merge Point. The LSR where one or more bypass tunnels rejoin the path of the protected LSP downstream of the potential failure. The same LSR may be both an MP and a PLR simultaneously.

Downstream PLR: A PLR that locally detects a fault and reroutes traffic in the same direction of the protected bidirectional LSP RSVP Path signaling.

Upstream PLR: A PLR that locally detects a fault and reroutes traffic in the opposite direction of the protected bidirectional LSP RSVP Path signaling.

Point of Remote Repair (PRR): An upstream PLR that triggers reroute of traffic and signaling based on procedures described in this document.

## 3. Fast Reroute For Unidirectional GMPLS LSPs

FRR procedures defined in [RFC4090] are applicable to unidirectional protected LSPs signaled using either RSVP-TE or GMPLS procedures and are not modified by the extensions proposed in this document. These FRR procedures also apply to bidirectional associated GMPLS LSPs where two unidirectional GMPLS LSPs are bound together by using association signaling [BID-ASSOC].

## 4. Bidirectional Bypass Tunnel Assignment for Bidirectional GMPLS LSPs

This section describes signaling procedures for bidirectional bypass tunnel assignment for GMPLS signaled PSC bidirectional co-routed TE LSPs.

### 4.1. Merge Point Labels

To correctly reroute data traffic over a node protection bypass tunnel, the downstream and upstream PLRs have to know, in advance, the downstream and upstream Merge Point (MP) labels so that data in the forward and reverse directions can be tunneled through the bypass tunnel post FRR respectively.

[RFC4090] defines procedures for the downstream PLR to obtain the protected LSP's downstream MP label from recorded labels in the RRO of the RSVP Resv message received at the downstream PLR.

To obtain the upstream MP label, existing methods [RFC4090] to record upstream MP label are used in the RRO of the RSVP Path message. The upstream PLR can obtain the upstream MP label from the recorded label in the RRO of the received RSVP Path message.

### 4.2. Merge Point Addresses

To correctly assign a bidirectional bypass tunnel, the downstream and upstream PLRs have to know, in advance, the downstream and upstream Merge Point (MP) addresses. [RFC4561] defines procedures for the PLR to obtain the protected LSP's merge point address in multi-domain routing networks where a domain is defined as an Interior Gateway Protocol (IGP) area or an Autonomous System (AS).

[RFC4561] defines procedures for the downstream PLR to obtain the protected LSP's downstream merge point address from the recorded node-IDs in the RRO of the RSVP Resv message received at the

downstream PLR.

To obtain the upstream MP address, existing methods [RFC4561] to record upstream MP node-ID are used in the RRO of the RSVP Path message. The upstream PLR can obtain the upstream MP address from the recorded node-IDs in the RRO of the received RSVP Path message.

### 4.3. RRO IPv4/IPv6 Subobject Flags

RRO IPv4/IPv6 subobject flags are defined in [RFC4090], Section 4.4 and are applicable to the FRR procedure for the bidirectional tunnels.

[RFC4090] defined procedure is used by the downstream PLR independently to signal the Ipv4/IPv6 subobject flags in the RRO of the RSVP Path message. Similarly, this procedure is used by the upstream PLR independently to signal the IPv4/IPv6 subobject flags in the RRO of the RSVP Resv message.

### 4.4. Bypass Tunnel Assignment Co-ordination

This document defines a new BYPASS_ASSIGNMENT subobject in RSVP RECORD_ROUTE object used to co-ordinate the bidirectional bypass tunnel selection between the downstream and upstream PLRs.

#### 4.4.1. Bypass Tunnel Assignment Co-ordination Signaling Procedure

It is desirable to coordinate the bidirectional bypass tunnel selected at the downstream and upstream PLRs so that rerouted traffic and signaling flow on co-routed paths post FRR. To achieve this, a new RSVP subobject is defined for RECORD_ROUTE object (RRO) that identifies a bidirectional bypass tunnel that is assigned at a downstream PLR to protect a bidirectional LSP.

The BYPASS_ASSIGNMENT subobject is added by each downstream PLR in the RSVP Path RECORD_ROUTE message of the GMPLS signaled bidirectional primary LSP to record the downstream bidirectional bypass tunnel assignment. This subobject is sent in the RSVP Path RECORD_ROUTE message every time the downstream PLR assigns or updates the bypass tunnel assignment so the upstream PLR may reflect the assignment too. The BYPASS_ASSIGNMENT subobject is added in the RECORD_ROUTE object prior to adding the node's IP address in the node-ID subobject. A node MUST NOT add a BYPASS_ASSIGNMENT subobject without also adding a Node-ID subobject. A node MUST NOT add a BYPASS_ASSIGNMENT subobject without also adding an IPv4 or IPv6 subobject.

The upstream PLR (downstream MP) that detects a BYPASS_ASSIGNMENT

subobject whose bypass tunnel and the node-ID subobject when used as a bypass tunnel source terminates locally assigns the matching bidirectional bypass tunnel in the reverse direction, and forwards the RSVP Path message downstream. Otherwise, the bypass tunnel assignment subobject is simply forwarded downstream along in the RSVP Path message.

In the absence of BYPASS_ASSIGNMENT subobject, the upstream PLR does not assign a bypass tunnel in the reverse direction. This allows the downstream PLR to always initiate the bypass assignment and upstream PLR to simply reflect the bypass assignment.

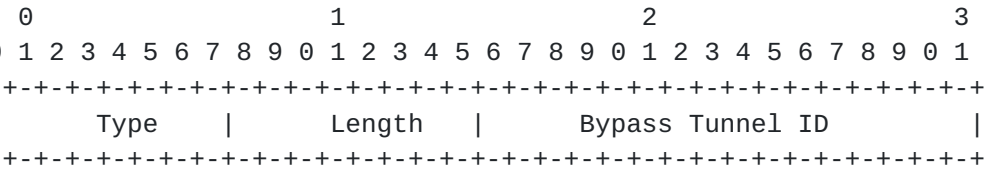
In the case of upstream PLR receiving multiple BYPASS_ASSIGNMENT subobjects from multiple downstream PLRs, the decision of selecting a bypass tunnel in the reverse direction can be based on local policy, for example, prefer link protection versus node protection bypass tunnel, or prefer the most upstream versus least upstream node protection bypass tunnel.

Bypass assignment co-ordination procedure described above can be used for both one-to-one backup described in Section 3.1 of [RFC4090] and facility backup described in Section 3.2 of [RFC4090].

### 4.4.2. BYPASS_ASSIGNMENT Subobject

The BYPASS_ASSIGNMENT subobject is used to inform the MP of the bypass tunnel being used by the PLR. This can be used to coordinate the bypass tunnel used for the protected LSP by the downstream and upstream PLRs in the forward and reverse directions respectively prior or post the failure occurrence. This subobject SHOULD only be inserted into the Path message by the downstream PLR and MUST NOT be changed by downstream LSRs.

The BYPASS_ASSIGNMENT subobject in RRO has the following format:




Type

Downstream Bypass Assignment.

Length

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

Bypass Tunnel ID

The bypass tunnel identifier (16 bits).

## 5. Link Protection Bypass Tunnels for Bidirectional GMPLS LSPs

When a bidirectional link protection bypass tunnel is used, after a link failure, downstream PLR reroutes RSVP Path and traffic over bypass tunnel using procedures defined in [RFC4090]. Upstream PLR may reroute traffic and RSVP Resv upon detecting the link failure or upon receiving RSVP Path message over a bidirectional bypass tunnel. This allows both traffic and RSVP signaling to flow on symmetric paths in the forward and reverse directions of a bidirectional tunnel.

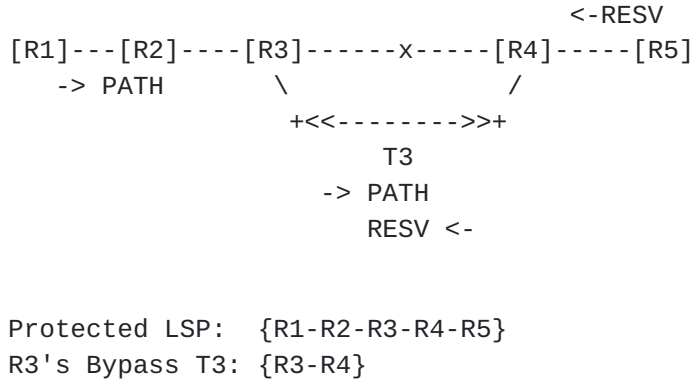



Figure 1: Flow of RSVP signaling post FRR after link failure

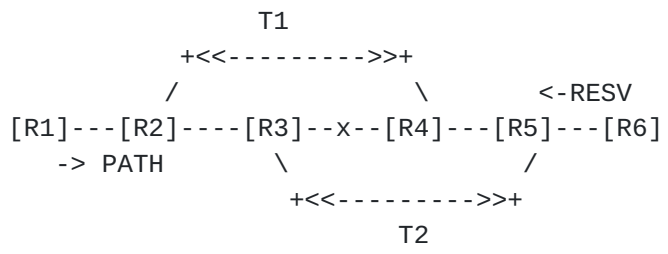
Consider the Traffic Engineered (TE) network shown in Figure 1. Assume every link in the network is protected with a link protection bypass tunnel (e.g. bypass tunnel T3). For the protected bidirectional co-routed LSP whose (active) head-end is on router R1 and (passive) tail-end is on router R5, each traversed router (a potential PLR) assigns a link protection bidirectional co-routed bypass tunnel. Consider a link R3-R4 on the protected LSP path fails.

### 5.1. Behavior Post Link Failure After FRR

The downstream PLR R3 and upstream PLR R4 independently trigger fast reroute procedures to redirect traffic onto bypass tunnels T3 in the forward and reverse directions. The downstream PLR R3 also reroutes

RSVP Path state onto the bypass tunnel T3 using procedures described in [RFC4090]. The upstream PLR R4 reroutes RSVP Resv onto the reverse bypass tunnel T3 upon receiving RSVP Path message over bypass tunnel T3.

## 6. Node Protection Bypass Tunnels for Bidirectional GMPLS LSPs




Protected LSP: {R1-R2-R3-R4-R5-R6}
R3's Bypass T2: {R3-R5}
R4's Bypass T1: {R4-R2}

Figure 2: Flow of RSVP signaling post FRR after link failure

Consider the Traffic Engineered (TE) network shown in Figure 2. Assume every link in the network is protected with a node protection bypass tunnel. For the protected bidirectional co-routed LSP whose (active) head-end is on router R1 and (passive) tail-end is on router R6, each traversed router (a potential PLR) assigns a node protection bidirectional co-routed bypass tunnel. Consider a link R3-R4 on the protected LSP path fails.

The proposed solution introduces two phases to invoking FRR procedures by the PLR post the link failure. The first phase comprises of FRR procedures to fast reroute data traffic onto bypass tunnels in the forward and reverse directions. The second phase re-coroutes the data and signaling in the forward and reverse directions after the first phase.

### 6.1. Behavior Post Link Failure After FRR

The downstream PLR R3 and upstream PLR R4 independently trigger fast reroute procedures to redirect traffic onto respective bypass tunnels T2 and T1 in the forward and reverse directions. The downstream PLR R3 also reroutes RSVP Path state onto the bypass tunnel T2 using procedures described in [RFC4090]. Note, at this point, router R4 stops receiving RSVP Path refreshes for the protected bidirectional LSP while primary protected traffic continues to flow over bypass tunnels.

### 6.2. Behavior Post Link Failure To Re-coroute

The downstream Merge Point (MP) R5 that receives rerouted protected LSP RSVP Path message through the bypass tunnel, in addition to the regular MP processing defined in [RFC4090], gets promoted to a Point of Remote Repair (PRR role) and performs the following actions to re-coroute signaling and data traffic over the same path in both directions:

- Finds the bypass tunnel in the reverse direction that terminates on the Downstream PLR R3. Note: the Downstream PLR R3's address is extracted from the "IPV4 tunnel sender address" in the SENDER_TEMPLATE object.

- If found, checks whether the primary LSP traffic and signaling are already rerouted over the found bypass tunnel. If not, PRR R5 activates FRR reroute procedures to direct traffic and RSVP Resv over the found bypass tunnel T2 in the reverse direction.

If downstream MP R5 receives multiple RSVP Path messages through multiple bypass tunnels (e.g. as a result of multiple failures), the PRR SHOULD identify a bypass tunnel that terminates on the farthest downstream PLR along the protected LSP path (closest to the primary bidirectional tunnel head-end) and activate the reroute procedures mentioned above.

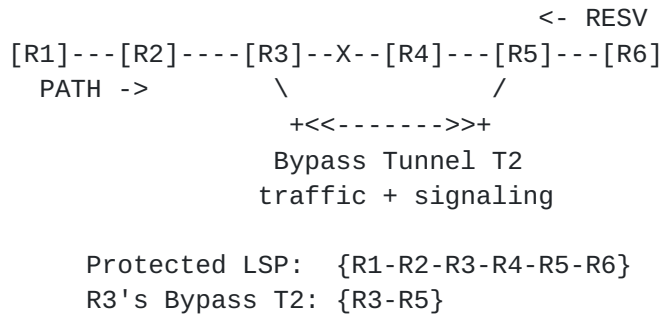



Figure 3: Flow of RSVP signaling post FRR after re-corouted

Figure 3 describes the path taken by the traffic and signaling after completing re-coroute of data and signaling in the forward and reverse paths described earlier.

The downstream MP MAY optionally support re-corouting in data plane as follows. If the downstream MP is pre-configured with bidirectional bypass tunnel, as soon as the MP node receives the

primary tunnel packets on this bypass tunnel, it MAY switch the upstream traffic on to this bypass tunnel. In order to identify the primary tunnel packets through this bypass tunnel, Penultimate Hop Popping (PHP) of the bypass tunnel MUST be disabled. The signaling procedure described above in this Section will still apply, and MP checks whether the primary tunnel traffic and signaling is already rerouted over the found bypass tunnel, if not, perform the above signaling procedure.

## 7. Compatibility

New RSVP subobject BYPASS_ASSIGNMENT is defined for RECORD_ROUTE in this document. Per [RFC2205], nodes not supporting this subobject will ignore the subobject but forward it without modification.

## 8. Security Considerations

This document introduces one new RSVP subobject that is carried in a signaling message. Thus in the event of the interception of a signaling message, slightly more information about the state of the network could be deduced than was previously the case. This is judged to be a very minor security risk as this information is already available by other means.

Otherwise, this document introduces no additional security considerations. For general discussion on MPLS and GMPLS related security issues, see the MPLS/GMPLS security framework [RFC5920].

## 9. IANA Considerations

A new type for the new BYPASS_ASSIGNMENT subobject for RSVP RECORD_ROUTE object is required.

## 10. Acknowledgements

Authors would like to thank George Swallow for his detailed and useful comments and suggestions.

## 11. References

### 11.1. Normative References

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

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

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

[BID-ASSOC] Zhang, F., Jing, R., and Gandhi, R., "RSVP-TE Extensions for Associated Bidirectional LSPs", July 2014.

### 11.2. Informative References

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

[RFC4561] Vasseur, J.-P., Ed., Ali, Z., and S. Sivabalan, "Definition of a Record Route Object (RRO) Node-Id Sub-Object", RFC 4561, June 2006.

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC5920, July 2010.

Authors' Addresses

Mike Taillon
Cisco Systems, Inc.

EMail: mtaillon@cisco.com

Tarek Saad (editor)
Cisco Systems, Inc.

EMail: tsaad@cisco.com

Rakesh Gandhi (editor)
Cisco Systems, Inc.

EMail: rgandhi@cisco.com

Zafar Ali
Cisco Systems, Inc.

EMail: zali@cisco.com

Manav Bhatia
India

Email: manav@ionosnetworks.com

Lizhong Jin
Shanghai, China

Email: lizho.jin@gmail.com

Frederic Jounay
Orange CH

Email: frederic.jounay@orange.ch