TEAS Working Group
Internet-Draft
Intended Status: Standards Track
Expires: December 5, 2016

M. Taillon
T. Saad, Ed.
R. Gandhi, Ed.
Z. Ali
Cisco Systems
June 3, 2016

Extensions to Resource Reservation Protocol For Fast Reroute of Traffic Engineering GMPLS LSPs
draft-ietf-teas-gmpls-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 a 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.

Table of Contents

1. Introduction

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

The 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], but they do not 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 the 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 for the RSVP signaling refreshes to stop reaching some nodes along the primary LSP path after the PLRs finish rerouting signaling onto the bypass tunnels. This may occur when using node protection bypass tunnels after a link failure event and when RSVP signaling is sent in-fiber and in-band with data. 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 causes the protected bidirectional LSP to be destroyed, with subsequent traffic loss after FRR.

Protection State Coordination Protocol [RFC6378] is applicable to FRR [RFC4090] for local protection of bidirectional co-routed LSPs in order to minimize traffic disruptions in both directions. However, this does not address the above mentioned problem of RSVP soft-state timeout in control plane.

This document proposes solutions to the above mentioned problems by providing mechanisms in the control plane to complement the 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 after 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.

Procedures defined in this document apply to co-routed GMPLS signaled PSC bidirectional TE primary and FRR bypass LSPs. Unless otherwise specified in this document, the FRR procedures defined in [RFC4090] are not modified by this document.

2. Conventions Used in This Document

2.1. Key Word Definitions

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].

2.2. Terminology

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

LSR: An MPLS Label-Switch Router.

LSP: An MPLS Label-Switched Path.

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.

PSC: Packet Switched Capable.

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. A downstream PLR has a corresponding downstream MP.

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

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

The 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 defined 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 [RFC7551].

4. 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. Bidirectional GMPLS Bypass Tunnel Direction

This document defines procedures where GMPLS bypass tunnels are provisioned in the same direction as the GMPLS primary LSPs. In other words, the GMPLS bypass tunnels originate on the downstream PLR and terminate on the downstream MP. As the originating downstream PLR node has the policy information about the locally provisioned bypass tunnels, it always initiates the bypass tunnel assignment. The GMPLS bypass tunnels originating from the upstream PLR and terminating on the upstream MP are outside the scope of this document.

4.2. 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 MP labels so that data in the forward and reverse directions can be redirected through the bypass tunnel after 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, the procedures specified in [RFC4090] are used to record the upstream MP label in the RRO of the RSVP Path message. The upstream PLR obtains the upstream MP label from the recorded labels in the RRO of the received RSVP Path message.

4.3. Merge Point Addresses

To correctly assign a bidirectional bypass tunnel, the downstream and upstream PLRs have to know, in advance, the downstream and upstream MP addresses.

[RFC4561] defines procedures for the downstream PLR to obtain the protected LSP's downstream MP 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, the procedures specified in [RFC4561] are used to record upstream MP node-ID in the RRO of the RSVP Path message. The upstream PLR obtains the upstream MP address from the recorded node-IDs in the RRO of the received RSVP Path message.

4.4. RRO IPv4/IPv6 Subobject Flags

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

The procedures defined in [RFC4090] are used by the downstream PLR to signal the IPv4/IPv6 subobject flags upstream in the RRO of the RSVP Resv message. Similarly, these procedures are used by the downstream PLR to signal the IPv4/IPv6 subobject flags downstream in the RRO of the RSVP Path message.

4.5. Bidirectional Bypass Tunnel Assignment Co-ordination

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

4.5.1. Bidirectional Bypass Tunnel Assignment 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 after 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 SHOULD be 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. The upstream PLR (downstream MP) simply reflects the bypass tunnel assignment in the reverse direction.

When the BYPASS_ASSIGNMENT subobject is added in the RECORD_ROUTE Object:

- The BYPASS_ASSIGNMENT subobject MUST be added prior to the Node-ID subobject containing the node's address.
- The Node-ID subobject MUST also be added.
- The IPv4 or IPv6 subobject MUST also be added.
- The Label subobject MUST also be added.

In the absence of BYPASS_ASSIGNMENT subobject, the upstream PLR (downstream MP) SHOULD NOT assign a bypass tunnel in the reverse direction. This allows the downstream PLR to always initiate the bypass assignment and upstream PLR (downstream MP) to simply reflect the bypass assignment.

The upstream PLR (downstream MP) that detects a BYPASS_ASSIGNMENT subobject, selects a reverse bypass tunnel that terminates locally with the matching tunnel-ID and has a source address matching the node-ID sub-object received in the subobject. The RRO containing BYPASS_ASSIGNMENT subobject(s) is then simply forwarded downstream in the RSVP Path message.

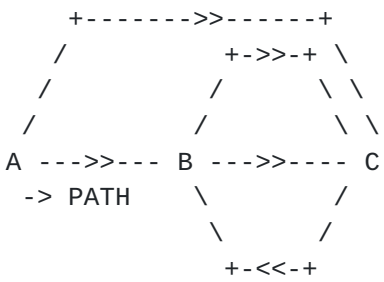
An upstream PLR (downstream MP) SHOULD examine the entire Path RRO and look at all BYPASS_ASSIGNMENT subobjects in order to assign a reverse bypass tunnel. The choice of a reverse bypass tunnel (if multiple bypass tunnels exist) is based on the local policy on the downstream MP and is discussed in Section 4.5.2 of this document.

The bypass assignment co-ordination procedure described in this Section 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].

4.5.2. Bidirectional Bypass Tunnel Assignment Policy

In the case of upstream PLR receiving multiple BYPASS_ASSIGNMENT subobjects from multiple downstream PLRs, the selection of a bypass tunnel in the reverse direction can be based on local policy. Examples of such a policy could be to prefer link protection over node protection, or to prefer the bypass tunnel to the furthest upstream node. When different policies are used for bypass tunnel assignment on the LSP path, it may result in some links in the reverse direction not assigned bypass protection during LSP setup as shown in examples below.

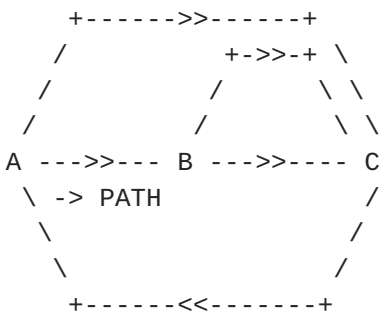
As shown in Example 1, node A assigns a node protection bypass tunnel in the forward direction but node C does not reflect the node protection bypass tunnel in the reverse direction for a protected bidirectional GMPLS LSP A-B-C. Both nodes B and C assign a link protection bypass tunnel. As a result, there is no fast reroute protection available in the reverse direction for link A-B for this LSP during the LSP setup. Note that this is corrected by node C during the re-coroute procedure after the FRR failure on link A-B as specified in Section 6 of this document since GMPLS bypass tunnels are bidirectional.



Example 1: An example of different bypass assignment policy

As shown in Example 2, nodes A and C assign a node protection bypass tunnel for a protected bidirectional GMPLS LSP A-B-C. Node B assigns a link protection bypass tunnel but node C does not reflect the reverse link protection bypass tunnel. As a result, there is no fast reroute protection available in the reverse direction for link A-B for this LSP during the LSP setup. Note that this is corrected by node C during the re-coroute procedure after the FRR failure on link

A-B as specified in Section 6 of this document since GMPLS bypass tunnels are bidirectional.



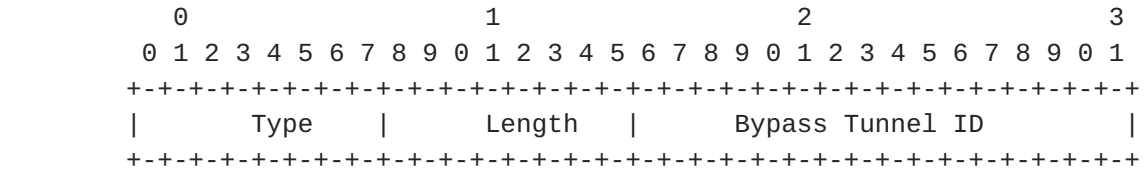
Example 2: An example of different bypass assignment policy

4.5.3. BYPASS_ASSIGNMENT Subobject

The BYPASS_ASSIGNMENT subobject is used to inform the downstream MP of the bypass tunnel being assigned by the PLR. This can be used to coordinate the bypass tunnel assignment for the protected LSP by the downstream and upstream PLRs in the forward and reverse directions respectively prior or after the failure occurrence.

This subobject SHOULD be inserted into the Path RRO by the downstream PLR. It SHOULD NOT be inserted into an RRO by a node which is not a downstream PLR. It MUST NOT be changed by downstream LSRs and MUST NOT be added to a Resv RRO.

The BYPASS_ASSIGNMENT subobject in RRO has the following format:



Type

Downstream Bypass Assignment. Value is TBA by IANA.

Length

The Length contains the total length of the subobject in bytes, including the Type and Length fields. The length is always 4 bytes.

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, the downstream PLR reroutes traffic and RSVP messages over the bypass tunnel using the procedures defined in [RFC4090]. Upstream PLR reroutes traffic upon detecting the link failure or upon receiving RSVP Path message over a bidirectional bypass tunnel. Upstream PLR reroutes RSVP Resv signaling 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 LSP.

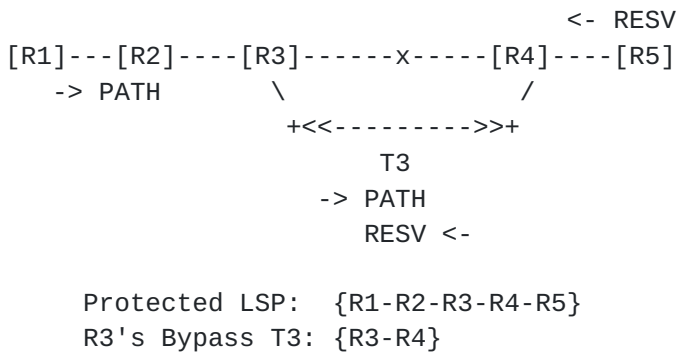


Figure 1: Flow of RSVP signaling after FRR and link failure

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 head-end is on node R1 and tail-end is on node R5, each traversed node (a potential PLR) assigns a link protection bidirectional co-routed bypass tunnel.

5.1. Behavior After Link Failure After FRR

Consider a link R3-R4 on the protected LSP path fails. 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.

5.2. Revertive Behavior After Link Failure After FRR

Revertive behavior as defined in [RFC4090], Section 6.5.2, is applicable to the link protection of GMPLS bidirectional LSPs. When using the local revertive mode, when downstream MP receives Path messages over the restored path, it starts sending Resv over the restored path and stops sending Resv over the reverse bypass tunnel. No additional procedure other than that specified in [RFC4090] is introduced for revertive behavior by this document.

6. Node Protection Bypass Tunnels for Bidirectional GMPLS LSPs

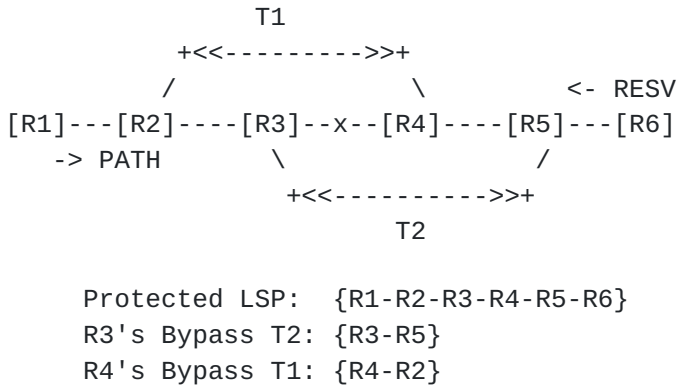


Figure 2: Flow of RSVP signaling after FRR and link failure

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 head-end is on node R1 and tail-end is on node R6, each traversed node (a potential PLR) assigns a node protection bidirectional co-routed bypass tunnel.

The proposed solution introduces two phases to invoking FRR procedures by the PLR after 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 After FRR and Link Failure

Consider a link R3-R4 on the protected LSP path fails. 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, node R4 stops receiving RSVP Path refreshes for the protected bidirectional LSP while primary protected traffic continues to flow over bypass tunnels.

6.2. Behavior After Link Failure To Re-coroute

The downstream 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 can be extracted from the "IPV4 tunnel sender address" in the SENDER_TEMPLATE Object of the primary LSP (see [RFC4090], Section 6.1.1).

- If reverse bypass tunnel is found and the primary LSP traffic is not already rerouted over the found bypass tunnel T2, the PRR R5 activates FRR reroute procedures to direct traffic over the found bypass tunnel T2 in the reverse direction. In addition, the PRR R5 also reroutes RSVP Resv over the bypass tunnel T2 in the reverse direction.

- If reverse bypass tunnel is not found, the PRR R5 immediately tears down the primary LSP.

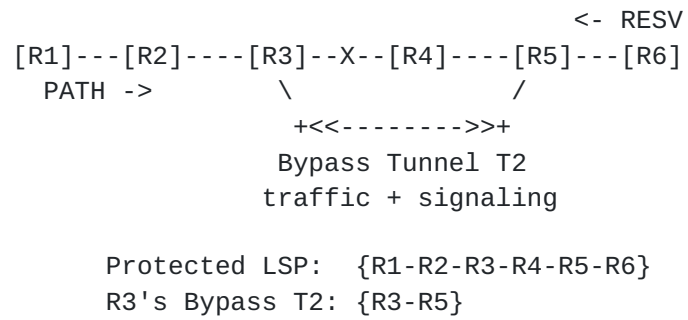


Figure 3: Flow of RSVP signaling after FRR and re-corouted

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. Node R4 will stop receiving the Path and Resv messages and it will timeout the RSVP soft-state,

however, this will not cause the LSP to be torn down. RSVP signaling at node R2 is not affected by the FRR and re-corouting.

If the link failure is unidirectional in the direction of R4 to R3, node R3 will stop receiving the RSVP Resv messages from node R4 and this will cause RSVP soft-state to timeout on node R3. However, unidirectional link failure in the opposite direction will not result in RSVP soft-state timeout as node R5 will trigger the re-coroute procedure after receiving RSVP Path message over the bypass tunnel from node R3.

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 LSP head-end) and activate the reroute procedures mentioned above.

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 LSP packets on this bypass tunnel, it MAY switch the upstream traffic on to this bypass tunnel. In order to identify the primary LSP 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 LSP traffic and signaling is already rerouted over the found bypass tunnel, if not, perform the above signaling procedure.

6.3. Revertive Behavior After Link Failure

Revertive behavior as defined in [RFC4090], Section 6.5.2, is applicable to node protection of GMPLS bidirectional LSPs. When using the local revertive mode, when downstream MP (R4) (before re-corouting) and PRR (R5) (after re-corouting) receive Path messages over the restored path, they start sending Resv over the restored path and stop sending Resv over the reverse bypass tunnel. No additional procedure other than that specified in [RFC4090] is introduced for revertive behavior by this document.

7. Compatibility

New RSVP subobject BYPASS_ASSIGNMENT is defined for RECORD_ROUTE Object 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 a new BYPASS_ASSIGNMENT subobject for the RECORD_ROUTE Object that is carried in an RSVP signaling message. Thus in the event of the interception of a signaling message, more information about LSP's fast reroute protection can 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

IANA manages the "RSVP PARAMETERS" registry located at <http://www.iana.org/assignments/rsvp-parameters>. IANA is requested to assign a value for the new BYPASS_ASSIGNMENT subobject in the "Class Type 21 ROUTE_RECORD - Type 1 Route Record" registry.

This document introduces a new subobject for RECORD_ROUTE Object:

Value	Description	Carried in Path	Carried in Resv	Reference
TBA By IANA	BYPASS_ASSIGNMENT subobject	Yes	No	This document

10. References

10.1. Normative References

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

[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.

[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.

[RFC7551] Zhang, F., Ed., Jing, R., and Gandhi, R., Ed., "RSVP-TE Extensions for Associated Bidirectional LSPs", RFC 7551, May 2015.

10.2. Informative References

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

[RFC6378] Weingarten, Y., Bryant, S., Osborne, E., Sprecher, N., and A. Fulignoli, "MPLS Transport Profile (MPLS-TP) Linear Protection", RFC 6378, October 2011.

Acknowledgements

Authors would like to thank George Swallow for his detailed and useful comments and suggestions. Authors would also like to thank Nobo Akiya, Loa Andersson, Matt Hartley and Gregory Mirsky for reviewing this document.

Contributors

Frederic Jounay
Orange CH

EMail: frederic.jounay@orange.ch

Manav Bhatia Ionos Networks Banglore India

EMail: manav@ionosnetworks.com

Lizhong Jin Shanghai, China

EMail: lizho.jin@gmail.com

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