

MPLS Working Group
Internet-Draft
Updates: [4090](#) (if approved)
Intended status: Standards Track
Expires: February 10, 2019

C. Ramachandran
Juniper Networks
I. Minei
Google, Inc
D. Pacella
Verizon
T. Saad
Cisco Systems Inc.
August 9, 2018

Refresh Interval Independent FRR Facility Protection
draft-ietf-mpls-ri-rsvp-frr-04

Abstract

RSVP-TE relies on periodic refresh of RSVP messages to synchronize and maintain the LSP related states along the reserved path. In the absence of refresh messages, the LSP related states are automatically deleted. Reliance on periodic refreshes and refresh timeouts are problematic from the scalability point of view. The number of RSVP-TE LSPs that a router needs to maintain has been growing in service provider networks and the implementations should be capable of handling increase in LSP scale.

[RFC 2961](#) specifies mechanisms to eliminate the reliance on periodic refresh and refresh timeout of RSVP messages, and enables a router to increase the message refresh interval to values much longer than the default 30 seconds defined in [RFC 2205](#). However, the protocol extensions defined in [RFC 4090](#) for supporting fast reroute (FRR) using bypass tunnels implicitly rely on short refresh timeouts to cleanup stale states.

In order to eliminate the reliance on refresh timeouts, the routers should unambiguously determine when a particular LSP state should be deleted. Coupling LSP state with the corresponding RSVP-TE signaling adjacencies as recommended in [RFC 8370](#) will apply in scenarios other than [RFC 4090](#) FRR using bypass tunnels. In scenarios involving [RFC 4090](#) FRR using bypass tunnels, additional explicit tear down messages are necessary. Refresh-interval Independent RSVP FRR (RI-RSVP-FRR) extensions specified in this document consists of procedures to enable LSP state cleanup that are essential in scenarios not covered by procedures defined in RSVP-TE Scaling Recommendations. Hence, this document updates the procedures defined in [RFC 4090](#) to support Refresh-Interval Independent RSVP (RI-RSVP) capability specified in [RFC 8370](#).

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](#)].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 10, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
1.1.	Motivation	4
2.	Terminology	5
3.	Problem Description	5
4.	Solution Aspects	7
4.1.	Requirement on RFC 4090 Capable Node to advertise RI-RSVP Capability	8
4.2.	Signaling Handshake between PLR and MP	8

4.2.1.	PLR Behavior	8
4.2.2.	Remote Signaling Adjacency	9
4.2.3.	MP Behavior	10
4.2.4.	"Remote" state on MP	11
4.3.	Impact of Failures on LSP State	12
4.3.1.	Non-MP Behavior	12
4.3.2.	LP-MP Behavior	12
4.3.3.	NP-MP Behavior	12
4.3.4.	Behavior of a Router that is both LP-MP and NP-MP . .	14
4.4.	Conditional Path Tear	14
4.4.1.	Sending Conditional Path Tear	14
4.4.2.	Processing Conditional Path Tear	15
4.4.3.	CONDITIONS object	15
4.5.	Remote State Teardown	16
4.5.1.	PLR Behavior on Local Repair Failure	17
4.5.2.	PLR Behavior on Resv RRO Change	17
4.5.3.	LSP Preemption during Local Repair	17
4.5.3.1.	Preemption on LP-MP after Phop Link failure . . .	17
4.5.3.2.	Preemption on NP-MP after Phop Link failure . . .	18
4.6.	Backward Compatibility Procedures	18
4.6.1.	Detecting Support for Refresh interval Independent FRR	19
4.6.2.	Procedures for backward compatibility	19
4.6.2.1.	Lack of support on Downstream Node	19
4.6.2.2.	Lack of support on Upstream Node	20
4.6.2.3.	Incremental Deployment	20
5.	Security Considerations	21
6.	IANA Considerations	22
6.1.	New Object - CONDITIONS	22
7.	Acknowledgements	22
8.	Contributors	22
9.	References	23
9.1.	Normative References	23
9.2.	Informative References	24
	Authors' Addresses	24

1. Introduction

RSVP-TE Fast Reroute [[RFC4090](#)] defines two local repair techniques to reroute label switched path (LSP) traffic over pre-established backup tunnel. Facility backup method allows one or more LSPs traversing a connected link or node to be protected using a bypass tunnel. The many-to-one nature of local repair technique is attractive from scalability point of view. This document enumerates facility backup procedures in [RFC 4090](#) that rely on refresh timeout and hence make facility backup method refresh-interval dependent. The RSVP-TE extensions defined in this document will enhance the facility backup

protection mechanism by making the corresponding procedures refresh-interval independent.

1.1. Motivation

Standard RSVP [[RFC2205](#)] maintains state via the generation of RSVP Path/Resv refresh messages. Refresh messages are used to both synchronize state between RSVP neighbors and to recover from lost RSVP messages. The use of Refresh messages to cover many possible failures has resulted in a number of operational problems.

- One problem relates to RSVP control plane scaling due to periodic refreshes of Path and Resv messages, another relates to the reliability and latency of RSVP signaling.
- An additional problem is the time to clean up the stale state after a tear message is lost. For more on these problems see [Section 1](#) of RSVP Refresh Overhead Reduction Extensions [[RFC2961](#)].

The problems listed above adversely affect RSVP control plane scalability and RSVP-TE [[RFC3209](#)] inherited these problems from standard RSVP. Procedures specified in [[RFC2961](#)] address the above mentioned problems by eliminating dependency on refreshes for state synchronization and for recovering from lost RSVP messages, and by eliminating dependency on refresh timeout for stale state cleanup. Implementing these procedures allows implementations to improve RSVP-TE control plane scalability. For more details on eliminating dependency on refresh timeout for stale state cleanup, refer to "Refresh Interval Independent RSVP" [section 3](#) of RSVP-TE Scaling Techniques [[RFC8370](#)].

However, the procedures specified in RSVP-TE Scaling Techniques [[RFC8370](#)] do not fully address stale state cleanup for facility backup protection [[RFC4090](#)], as facility backup protection still depends on refresh timeouts for stale state cleanup.

The procedures specified in this document, in combination with RSVP-TE Scaling Techniques [[RFC8370](#)], eliminate facility backup protection dependency on refresh timeouts for stale state cleanup including the cleanup for facility backup protection. The document hence updates the semantics of Refresh-Interval Independent RSVP (RI-RSVP) capability specified in [Section 3](#) of RSVP-TE Scaling Techniques [[RFC8370](#)].

The procedures specified in this document assume reliable delivery of RSVP messages, as specified in [[RFC2961](#)]. Therefore this document makes support for [[RFC2961](#)] a pre-requisite.

2. Terminology

The reader is expected to be familiar with the terminology in [\[RFC2205\]](#), [\[RFC3209\]](#), [\[RFC4090\]](#) and [\[RFC4558\]](#).

Phop node: Previous-hop router along the label switched path

PPhop node: Previous-Previous-hop router along the LSP

LP-MP node: Merge Point router at the tail of Link-protecting bypass tunnel

NP-MP node: Merge Point router at the tail of Node-protecting bypass tunnel

TED: Traffic Engineering Database

LSP state: The combination of "path state" maintained as Path State Block (PSB) and "reservation state" maintained as Reservation State Block (RSB) forms an individual LSP state on an RSVP-TE speaker

Conditional PathTear: PathTear message containing a suggestion to a receiving downstream router to retain Path state if the receiving router is NP-MP

Remote PathTear: PathTear message sent from Point of Local Repair (PLR) to MP to delete LSP state on MP if PLR had not reliably sent backup Path state before

3. Problem Description

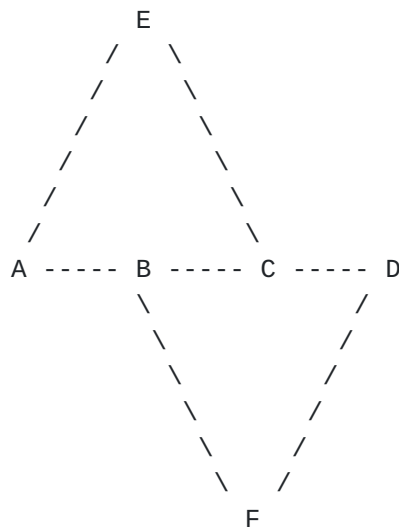


Figure 1: Example Topology

In the topology in Figure 1, consider a large number of LSPs from A to D transiting B and C. Assume that refresh interval has been configured to be long of the order of minutes and refresh reduction extensions are enabled on all routers.

Also assume that node protection has been configured for the LSPs and the LSPs are protected by each router in the following way

- A has made node protection available using bypass LSP A -> E -> C; A is the Point of Local Repair (PLR) and C is Node Protecting Merge Point (NP-MP)
- B has made node protection available using bypass LSP B -> F -> D; B is the PLR and D is the NP-MP
- C has made link protection available using bypass LSP C -> B -> F -> D; C is the PLR and D is the Link Protecting Merge Point (LP-MP)

In the above condition, assume that B-C link fails. The following is the sequence of events that is expected to occur for all protected LSPs under normal conditions.

1. B performs local repair and re-directs LSP traffic over the bypass LSP B -> F -> D.
2. B also creates backup state for the LSP and triggers sending of backup LSP state to D over the bypass LSP B -> F -> D.
3. D receives backup LSP states and merges the backups with the protected LSPs.
4. As the link on C, over which the LSP states are refreshed has failed, C will no longer receive state refreshes. Consequently the protected LSP states on C will time out and C will send tear down message for all LSPs. As each router should consider itself as a Merge Point, C will time out the state only after waiting for an additional duration equal to refresh timeout.

While the above sequence of events has been described in [\[RFC4090\]](#), there are a few problems for which no mechanism has been specified explicitly.

- If the protected LSP on C times out before D receives signaling for the backup LSP, then D would receive PathTear from C prior to receiving signaling for the backup LSP, thus resulting in deleting the LSP state. This would be possible at scale even with default refresh time.

- If upon the link failure C is to keep state until its timeout, then with long refresh interval this may result in a large amount of stale state on C. Alternatively, if upon the link failure C is to delete the state and send PathTear to D, this would result in deleting the state on D, thus deleting the LSP. D needs a reliable mechanism to determine whether it is MP or not to overcome this problem.
- If head-end A attempts to tear down LSP after step 1 but before step 2 of the above sequence, then B may receive the tear down message before step 2 and delete the LSP state from its state database. If B deletes its state without informing D, with long refresh interval this could cause (large) buildup of stale state on D.
- If B fails to perform local repair in step 1, then B will delete the LSP state from its state database without informing D. As B deletes its state without informing D, with long refresh interval this could cause (large) buildup of stale state on D.

The purpose of this document is to provide solutions to the above problems which will then make it practical to scale up to a large number of protected LSPs in the network.

4. Solution Aspects

The solution consists of five parts.

- Utilize MP determination mechanism specified in RSVP-TE Summary FRR [[I-D.ietf-mpls-summary-frr-rsvpte](#)] that enables the PLR to signal the availability of local protection to the MP. In addition, introduce PLR and MP procedures to establish Node-ID based hello session between the PLR and the MP to detect router failures and to determine capability. See [section 4.2](#) for more details. This part of the solution re-uses some of the extensions defined in RSVP-TE Summary FRR [[I-D.ietf-mpls-summary-frr-rsvpte](#)] and RSVP-TE Scaling Techniques [[RFC8370](#)], and the subsequent sub-sections will list the extensions in these drafts that are utilized in this document.
- Handle upstream link or node failures by cleaning up LSP states if the node has not found itself as MP through the MP determination mechanism. See [section 4.3](#) for more details.
- Introduce extensions to enable a router to send tear down message to the downstream router that enables the receiving router to conditionally delete its local LSP state. See [section 4.4](#) for more details.

- Enhance facility protection by allowing a PLR to directly send tear down message to MP without requiring the PLR to either have a working bypass LSP or have already signaled backup LSP state. See [section 4.5](#) for more details.
- Introduce extensions to enable the above procedures to be backward compatible with routers along the LSP path running implementation that do not support these procedures. See [section 4.6](#) for more details.

[4.1.](#) Requirement on [RFC 4090](#) Capable Node to advertise RI-RSVP Capability

A node supporting [RFC 4090](#) facility protection FRR MAY set the RI-RSVP capability (I bit) defined in [Section 3](#) of RSVP-TE Scaling Techniques [[RFC8370](#)] only if it supports all the extensions specified in the rest of this document. A node supporting [RFC 4090](#) facility bypass FRR but not supporting the extensions specified in this document MUST reset RI-RSVP capability (I bit) in the outgoing Node-ID based Hello messages. Hence, this document updates [RFC 4090](#) by defining extensions and additional procedures over facility protection FRR defined in [RFC 4090](#) in order to advertise RI-RSVP capability [[RFC8370](#)].

[4.2.](#) Signaling Handshake between PLR and MP

[4.2.1.](#) PLR Behavior

As per the procedures specified in [RFC 4090](#), when a protected LSP comes up and if the "local protection desired" flag is set in the SESSION_ATTRIBUTE object, each node along the LSP path attempts to make local protection available for the LSP.

- If the "node protection desired" flag is set, then the node tries to become a PLR by attempting to create a NP-bypass LSP to the Nnhop node avoiding the Nhop node on protected LSP path. In case node protection could not be made available, the node attempts to create a LP-bypass LSP to Nhop node avoiding only the link that protected LSP takes to reach Nhop
- If the "node protection desired" flag is not set, then the PLR attempts to create a LP-bypass LSP to Nhop node avoiding the link that the protected LSP takes to reach Nhop

With regard to the PLR procedures described above and that are specified in [RFC 4090](#), this document specifies the following additional procedures to support RI-RSVP defined in [RFC 8370](#).

- While selecting the destination address of the bypass LSP, the PLR SHOULD select the router ID of the NNhop or Nhop node from the Node-ID sub-object included RRO object carried in RESV message. If the MP has not included Node-ID sub-object in RESV RRO and if the PLR and the MP are in the same area, then the PLR may utilize the TED to determine the router ID corresponding to the interface address included by the MP in the RRO object. If the NP-MP in a different IGP area has not included Node-ID sub-object in RRO object, then the PLR SHOULD execute backward compatibility procedures as if the downstream nodes along the LSP do not support the extensions defined in the document (see [Section 4.6.2.1](#)).
- The PLR SHOULD also include its router ID in a Node-ID sub-object in RRO object carried in PATH message. While including its router ID in the Node-ID sub-object carried in the outgoing PATH message, the PLR MUST include the Node-ID sub-object after including its IPv4/IPv6 address or unnumbered interface ID sub-object.
- In parallel to the attempt made to create NP-bypass or LP-bypass, the PLR SHOULD initiate a Node-ID based Hello session to the NNhop or Nhop node respectively to establish the RSVP-TE signaling adjacency. This Hello session is used to detect MP node failure as well as determine the capability of the MP node. If the MP has set the I-bit in CAPABILITY object [[RFC8370](#)] carried in Hello message corresponding to Node-ID based Hello session, then the PLR SHOULD conclude that the MP supports refresh-interval independent FRR procedures defined in this document. If the MP has not sent Node-ID based Hello messages or has not set the I-bit in CAPABILITY object [[RFC8370](#)], then the PLR SHOULD execute backward compatibility procedures defined in [Section 4.6.2.1](#) of this document.
- If the bypass LSP comes up and the PLR has made local protection available for one or more LSPs, then the PLR SHOULD include B-SFRR-Ready Extended Association object and triggers PATH message to be sent for those LSPs. If a B-SFRR-Ready Extended Association object is included in the PATH message, then the encoding and ordering rules object specified in RSVP-TE Summary FRR [[I-D.ietf-mpls-summary-frr-rsvpte](#)] MUST be followed.

[4.2.2. Remote Signaling Adjacency](#)

A Node-ID based RSVP-TE Hello session is one in which Node-ID is used in the source and the destination address fields of RSVP Hello messages [[RFC4558](#)]. This document extends Node-ID based RSVP Hello session to track the state of any RSVP-TE neighbor that is not directly connected by at least one interface. In order to apply Node-ID based RSVP-TE Hello session between any two routers that are

not immediate neighbors, the router that supports the extensions defined in the document SHOULD set TTL to 255 in all outgoing Node-ID based Hello messages exchanged between PLR and MP. The default hello interval for this Node-ID hello session SHOULD be set to the default specified in RSVP-TE Scaling Techniques [[RFC8370](#)].

In the rest of the document the term "signaling adjacency", or "remote signaling adjacency" refers specifically to the RSVP-TE signaling adjacency.

4.2.3. MP Behavior

With regard to the MP procedures that are defined in [RFC 4090](#), this document specifies the following additional procedures to support RI-RSVP defined in [RFC 8370](#).

Each node along an LSP path supporting the extensions defined in this document SHOULD also include its router ID in the Node-ID sub-object in the RRO object carried in the RESV message of the LSPs. If the PLR has not included Node-ID sub-object in the RRO object carried in PATH message and if the PLR is in a different IGP area, then the router SHOULD NOT execute the MP procedures specified in this document for those LSPs. Instead, the node SHOULD execute backward compatibility procedures defined in [Section 4.6.2.2](#) as if the upstream nodes along the LSP do not support the extensions defined in this document.

The node should determine whether the incoming PATH messages contains B-SFRR-Ready Extended Association object with the Node-ID address of the PLR as the source and its own Node-ID as the destination. In addition the node should determine whether it has an operational remote Node-ID signaling adjacency with the PLR. If either the PLR has not included B-SFRR-Ready Extended Association object or if there is no operational Node-ID signaling adjacency with the PLR or if the PLR has not advertised RI-RSVP capability in its Node-ID based Hello messages, then the node SHOULD execute backward compatibility procedures defined in [Section 4.6.2.2](#).

If a matching B-SFRR-Ready Extended Association object is found in the PATH message and if there is an operational remote signaling adjacency with the PLR that has advertised RI-RSVP capability (I-bit) [[RFC8370](#)] in its Node-ID based Hello messages, then the node SHOULD consider itself as the MP for the corresponding PLR. The matching and ordering rules for Bypass Summary FRR Extended Association specified in RSVP-TE Summary FRR [[I-D.ietf-mpls-summary-frr-rsvpte](#)] MUST be followed by implementations supporting this document.

- If a matching Bypass Summary FRR Extended Association object is included by the PPhop node of an LSP and if a corresponding Node-ID signaling adjacency exists with the PPhop node, then the router SHOULD conclude it is NP-MP.
- If a matching Bypass Summary FRR Extended Association object is included by the Phop node of an LSP and if a corresponding Node-ID signaling adjacency exists with the Phop node, then the router SHOULD conclude it is LP-MP.

[4.2.4.](#) "Remote" state on MP

Once a router concludes it is the MP for a PLR running refresh-interval independent FRR procedures, it SHOULD create a remote path state for the LSP. The "remote" state is identical to the protected LSP path state except for the difference in RSVP_HOP object. The RSVP_HOP object in "remote" Path state contains the address that the PLR uses to send Node-ID hello messages to MP.

The MP SHOULD consider the "remote" path state automatically deleted if:

- MP later receives a PATH with no matching B-SFRR-Ready Extended Association object corresponding to the PLR's IP address contained in PATH RRO, or
- Node signaling adjacency with PLR goes down, or
- MP receives backup LSP signaling from PLR or
- MP receives PathTear, or
- MP deletes the LSP state on local policy or exception event

Unlike the normal path state that is either locally generated on the Ingress or created from a PATH message from the Phop node, the "remote" path state is not signaled explicitly from PLR. The purpose of "remote" path state is to enable the PLR to explicitly tear down path and reservation states corresponding to the LSP by sending tear message for the "remote" path state. Such message tearing down "remote" path state is called "Remote PathTear".

The scenarios in which "Remote" PathTear is applied are described in [Section 4.5](#).

4.3. Impact of Failures on LSP State

This section describes the procedures for routers on the LSP path for different kinds of failures. The procedures described on detecting RSVP control plane adjacency failures do not impact the RSVP-TE graceful restart mechanisms ([RFC3473], [RFC5063]). If the router executing these procedures act as helper for neighboring router, then the control plane adjacency will be declared as having failed after taking into account the grace period extended for neighbor by the helper.

Immediate node failures are detected from the state of Node-ID hello sessions established with immediate neighbors. RSVP-TE Scaling Techniques [RFC8370] recommends each router to establish Node-ID hello sessions with all its immediate neighbors. PLR or MP node failure is detected from the state of remote signaling adjacency established according to [Section 4.2.2](#) of this document.

4.3.1. Non-MP Behavior

When a router detects Phop link or Phop node failure and the router is not an MP for the LSP, then it SHOULD send Conditional PathTear (refer to [Section 4.4](#) "Conditional PathTear" below) and delete PSB and RSB states corresponding to the LSP.

4.3.2. LP-MP Behavior

When the Phop link for an LSP fails on a router that is LP-MP for the LSP, the LP-MP SHOULD retain PSB and RSB states corresponding to the LSP till the occurrence of any of the following events.

- Node-ID signaling adjacency with Phop PLR goes down, or
- MP receives normal or "Remote" PathTear for PSB, or
- MP receives ResvTear RSB.

When a router that is LP-MP for an LSP detects Phop node failure from Node-ID signaling adjacency state, the LP-MP SHOULD send normal PathTear and delete PSB and RSB states corresponding to the LSP.

4.3.3. NP-MP Behavior

When a router that is NP-MP for an LSP detects Phop link failure, or Phop node failure from Node-ID signaling adjacency, the router SHOULD retain PSB and RSB states corresponding to the LSP till the occurrence of any of the following events.

- Remote Node-ID signaling adjacency with PPhop PLR goes down, or
- MP receives normal or "Remote" PathTear for PSB, or
- MP receives ResvTear for RSB.

When a router that is NP-MP does not detect Phop link or node failure, but receives Conditional PathTear from the Phop node, then the router SHOULD retain PSB and RSB states corresponding to the LSP till the occurrence of any of the following events.

- Remote Node-ID signaling adjacency with PPhop PLR goes down, or
- MP receives normal or "Remote" PathTear for PSB, or
- MP receives ResvTear for RSB.

Receiving Conditional PathTear from the Phop node will not impact the "remote" state from the PPhop PLR. Note that Phop node would send Conditional PathTear if it was not an MP.

In the example topology in Figure 1, assume C & D are NP-MP for PLRs A & B respectively. Now when A-B link fails, as B is not MP and its Phop link has failed, B will delete LSP state (this behavior is required for unprotected LSPs - [Section 4.3.1](#)). In the data plane, that would require B to delete the label forwarding entry corresponding to the LSP. So if B's downstream nodes C and D continue to retain state, it would not be correct for D to continue to assume itself as NP-MP for PLR B.

The mechanism that enables D to stop considering itself as the NP-MP for B and delete the corresponding "remote" path state is given below.

1. When C receives Conditional PathTear from B, it decides to retain LSP state as it is NP-MP of PLR A. C also SHOULD check whether Phop B had previously signaled availability of node protection. As B had previously signaled NP availability by including B-SFRR-Ready Extended Association object, C SHOULD remove the B-SFRR-Ready Extended Association object containing Association Source set to B from the PATH message and trigger PATH to D.
2. When D receives triggered PATH, it realizes that it is no longer the NP-MP for B and so it deletes the corresponding "remote" path state. D does not propagate PATH further down because the only change is that the B-SFRR-Ready Extended Association object corresponding to Association Source B is no longer present in the PATH message.

4.3.4. Behavior of a Router that is both LP-MP and NP-MP

A router may be both LP-MP as well as NP-MP at the same time for Phop and PPhop nodes respectively of an LSP. If Phop link fails on such node, the node SHOULD retain PSB and RSB states corresponding to the LSP till the occurrence of any of the following events.

- Both Node-ID signaling adjacencies with Phop and PPhop nodes go down, or
- MP receives normal or "Remote" PathTear for PSB, or
- MP receives ResvTear for RSB.

If a router that is both LP-MP and NP-MP detects Phop node failure, then the node SHOULD retain PSB and RSB states corresponding to the LSP till the occurrence of any of the following events.

- Remote Node-ID signaling adjacency with PPhop PLR goes down, or
- MP receives normal or "Remote" PathTear for PSB, or
- MP receives ResvTear for RSB.

4.4. Conditional Path Tear

In the example provided in the [Section 4.3.3](#), B deletes PSB and RSB states corresponding to the LSP once B detects its link to Phop went down as B is not MP. If B were to send PathTear normally, then C would delete LSP state immediately. In order to avoid this, there should be some mechanism by which B can indicate to C that B does not require the receiving node to unconditionally delete the LSP state immediately. For this, B SHOULD add a new optional object called CONDITIONS object in PathTear. The new optional object is defined in [Section 4.4.3](#). If node C also understands the new object, then C SHOULD delete LSP state only if it is not an NP-MP - in other words C SHOULD delete LSP state if there is no "remote" PLR path state on C.

4.4.1. Sending Conditional Path Tear

A router that is not an MP for an LSP SHOULD delete PSB and RSB states corresponding to the LSP if Phop link or Phop Node-ID signaling adjacency goes down ([Section 4.3.1](#)). The router SHOULD send Conditional PathTear if the following are also true.

- Ingress has requested node protection for the LSP, and
- PathTear is not received from the upstream node

4.4.2. Processing Conditional Path Tear

When a router that is not an NP-MP receives Conditional PathTear, the node SHOULD delete PSB and RSB states corresponding to the LSP, and process Conditional PathTear by considering it as normal PathTear. Specifically, the node SHOULD NOT propagate Conditional PathTear downstream but remove the optional object and send normal PathTear downstream.

When a node that is an NP-MP receives Conditional PathTear, it SHOULD NOT delete LSP state. The node SHOULD check whether the Phop node had previously included B-SFRR-Ready Extended Association object in PATH. If the object had been included previously by the Phop, then the node processing Conditional PathTear from the Phop SHOULD remove the corresponding object and trigger PATH downstream.

If Conditional PathTear is received from a neighbor that has not advertised support (refer to [Section 4.6](#)) for the new procedures defined in this document, then the node SHOULD consider the message as normal PathTear. The node SHOULD propagate normal PathTear downstream and delete the LSP state.

4.4.3. CONDITIONS object

As any implementation that does not support Conditional PathTear SHOULD ignore the new object but process the message as normal PathTear without generating any error, the Class-Num of the new object SHOULD be 10bbbbbb where 'b' represents a bit (from [Section 3.10 of \[RFC2205\]](#)).

The new object is called as "CONDITIONS" object that will specify the conditions under which default processing rules of the RSVP-TE message SHOULD be invoked.

The object has the following format:

```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Length           | Class       | C-type   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Reserved                               |M|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Figure 2: CONDITIONS Object

Length

This contains the size of the object in bytes and should be set to eight.

Class

To be assigned

C-type

1

M bit

If M-bit is set to 1, then the PathTear message SHOULD be processed based on the condition if the receiver router is a Merge Point or not.

If M-bit is set to 0, then the PathTear message SHOULD be processed as normal PathTear message.

4.5. Remote State Teardown

If the Ingress wants to tear down the LSP because of a management event while the LSP is being locally repaired at a transit PLR, it would not be desirable to wait till the completion of backup LSP signaling to perform state cleanup. To enable LSP state cleanup when the LSP is being locally repaired, the PLR SHOULD send "remote" PathTear message instructing the MP to delete PSB and RSB states corresponding to the LSP. The TTL in "remote" PathTear message SHOULD be set to 255.

Consider node C in example topology (Figure 1) has gone down and B locally repairs the LSP.

1. Ingress A receives a management event to tear down the LSP.
2. A sends normal PathTear to B.
3. Assume B has not initiated backup signaling for the LSR. To enable LSP state cleanup, B SHOULD send "remote" PathTear with destination IP address set to that of D used in Node-ID signaling adjacency with D, and RSVP_HOP object containing local address used in Node-ID signaling adjacency.
4. B then deletes PSB and RSB states corresponding to the LSP.
5. On D there would be a remote signaling adjacency with B and so D SHOULD accept the remote PathTear and delete PSB and RSB states corresponding to the LSP.

4.5.1. PLR Behavior on Local Repair Failure

If local repair fails on the PLR after a failure, then this should be considered as a case for cleaning up LSP state from PLR to the Egress. PLR would achieve this using "remote" PathTear to clean up state from MP. If MP has retained state, then it would propagate PathTear downstream thereby achieving state cleanup. Note that in the case of link protection, the PathTear would be directed to LP-MP node IP address rather than the Nhop interface address.

4.5.2. PLR Behavior on Resv RRO Change

When a router that has already made NP available detects a change in the RRO carried in RESV message, and if the RRO change indicates that the router's former NP-MP is no longer present in the LSP path, then the router SHOULD send "Remote" PathTear directly to its former NP-MP.

In the example topology in Figure 1, assume A has made node protection available and C has concluded it is the NP-MP for A. When the B-C link fails then C, implementing the procedure specified in [Section 4.3.4](#) of this document, will retain state till: remote Node-ID signaling adjacency with A goes down, or PathTear or ResvTear is received for PSB or RSB respectively. If B also has made node protection available, B will eventually complete backup LSP signaling with its NP-MP D and trigger RESV to A with RRO changed. The new RRO of the LSP carried in RESV will not contain C. When A processes the RESV with a new RRO not containing C - its former NP-MP, A SHOULD send "Remote" PathTear to C. When C receives a "Remote" PathTear for its PSB state, C will send normal PathTear downstream to D and delete both PSB and RSB states corresponding to the LSP. As D has already received backup LSP signaling from B, D will retain control plane and forwarding states corresponding to the LSP.

4.5.3. LSP Preemption during Local Repair

4.5.3.1. Preemption on LP-MP after Phop Link failure

If an LSP is preempted on LP-MP after its Phop or incoming link has already failed but the backup LSP has not been signaled yet, then the node SHOULD send normal PathTear and delete both PSB and RSB states corresponding to the LSP. As the LP-MP has retained LSP state expecting the PLR to perform backup LSP signaling, preemption would bring down the LSP and the node would not be LP-MP any more requiring the node to clean up LSP state.

4.5.3.2. Preemption on NP-MP after Phop Link failure

If an LSP is preempted on NP-MP after its Phop link has already failed but the backup LSP has not been signaled yet, then the node SHOULD send normal PathTear and delete PSB and RSB states corresponding to the LSP. As the NP-MP has retained LSP state expecting the PLR to perform backup LSP signaling, preemption would bring down the LSP and the node would not be NP-MP any more requiring the node to clean up LSP state.

Consider B-C link goes down on the same example topology (Figure 1). As C is NP-MP for PLR A, C will retain LSP state.

1. The LSP is preempted on C.
2. C will delete RSB state corresponding to the LSP. But C cannot send PathErr or ResvTear to PLR A because backup LSP has not been signaled yet.
3. As the only reason for C having retained state after Phop node failure was that it was NP-MP, C SHOULD send normal PathTear to D and delete PSB state also. D would also delete PSB and RSB states on receiving PathTear from C.
4. B starts backup LSP signaling to D. But as D does not have the LSP state, it will reject backup LSP PATH and send PathErr to B.
5. B will delete its reservation and send ResvTear to A.

4.6. Backward Compatibility Procedures

The "Refresh interval Independent FRR" or RI-RSVP-FRR referred below in this section refers to the changes that have been proposed in previous sections. Any implementation that does not support them has been termed as "non-RI-RSVP-FRR implementation". The extensions proposed in RSVP-TE Summary FRR [[I-D.ietf-mpls-summary-frr-rsvpte](#)] are applicable to implementations that do not support RI-RSVP-FRR. On the other hand, changes proposed relating to LSP state cleanup namely Conditional and remote PathTear require support from one-hop and two-hop neighboring nodes along the LSP path. So procedures that fall under LSP state cleanup category SHOULD be turned on only if all nodes involved in the node protection FRR i.e. PLR, MP and intermediate node in the case of NP, support the extensions. Note that for LSPs requesting only link protection, the PLR and the LP-MP should support the extensions.

4.6.1. Detecting Support for Refresh interval Independent FRR

An implementation supporting the extensions specified in previous sections (called RI-RSVP-FRR here after) SHOULD set the flag "Refresh interval Independent RSVP" or RI-RSVP in CAPABILITY object carried in Hello messages. The RI-RSVP flag is specified in RSVP-TE Scaling Techniques [[RFC8370](#)].

- As nodes supporting the extensions SHOULD initiate Node Hellos with adjacent nodes, a node on the path of protected LSP can determine whether its PPhop or NPhop neighbor supports RI-RSVP-FRR enhancements from the Hello messages sent by the neighbor.
- If a node attempts to make node protection available, then the PLR SHOULD initiate remote Node-ID signaling adjacency with NNhop. If the NNhop (a) does not reply to remote node Hello message or (b) does not set RI-RSVP flag in CAPABILITY object carried in its Node-ID Hello messages, then the PLR can conclude that NNhop does not support RI-RSVP-FRR extensions.
- If node protection is requested for an LSP and if (a) PPhop node has not included a matching B-SFRR-Ready Extended Association object in PATH or (b) PPhop node has not initiated remote node Hello messages or (c) PPhop node does not set RI-RSVP flag in CAPABILITY object carried in its Node-ID Hello messages, then the node SHOULD conclude that the PLR does not support RI-RSVP-FRR extensions. The details are described in the "Procedures for backward compatibility" section below.

4.6.2. Procedures for backward compatibility

The procedures defined hereafter are performed on a subset of LSPs that traverse a node, rather than on all LSPs that traverse a node. This behavior is required to support backward compatibility for a subset of LSPs traversing nodes running non-RI-RSVP-FRR implementations.

4.6.2.1. Lack of support on Downstream Node

The procedures on the downstream direction are as follows.

- If the NPhop does not support the RI-RSVP-FRR extensions, then the node SHOULD reduce the "refresh period" in TIME_VALUES object carried in PATH to default short refresh default value.
- If node protection is requested and the NNhop node does not support the enhancements, then the node SHOULD reduce the "refresh

period" in TIME_VALUES object carried in PATH to a short refresh default value.

If the node reduces the refresh time from the above procedures, it SHOULD also not send remote PathTear or Conditional PathTear messages.

Consider the example topology in Figure 1. If C does not support the RI-RSVP-FRR extensions, then:

- A and B SHOULD reduce the refresh time to default value of 30 seconds and trigger PATH
- If B is not an MP and if Phop link of B fails, B cannot send Conditional PathTear to C but SHOULD time out PSB state from A normally. This would be accomplished if A would also reduce the refresh time to default value. So if C does not support the RI-RSVP-FRR extensions, then Phop B and PPhop A SHOULD reduce refresh time to a small default value.

4.6.2.2. Lack of support on Upstream Node

The procedures on the upstream direction are as follows.

- If Phop node does not support the RI-RSVP-FRR extensions, then the node SHOULD reduce the "refresh period" in TIME_VALUES object carried in RESV to default short refresh time value.
- If node protection is requested and the Phop node does not support the RI-RSVP-FRR extensions, then the node SHOULD reduce the "refresh period" in TIME_VALUES object carried in PATH to default value.
- If node protection is requested and PPhop node does not support the RI-RSVP-FRR extensions, then the node SHOULD reduce the "refresh period" in TIME_VALUES object carried in RESV to default value.
- If the node reduces the refresh time from the above procedures, it SHOULD also not execute MP procedures specified in [Section 4.3](#) of this document.

4.6.2.3. Incremental Deployment

The backward compatibility procedures described in the previous subsections imply that a router supporting the RI-RSVP-FRR extensions specified in this document can apply the procedures specified in the document either in the downstream or upstream direction of an LSP,

depending on the capability of the routers downstream or upstream in the LSP path.

- RI-RSVP-FRR extensions and procedures are enabled for downstream Path, PathTear and ResvErr messages corresponding to an LSP if link protection is requested for the LSP and the Nhop node supports the extensions
- RI-RSVP-FRR extensions and procedures are enabled for downstream Path, PathTear and ResvErr messages corresponding to an LSP if node protection is requested for the LSP and both Nhop & NNhop nodes support the extensions
- RI-RSVP-FRR extensions and procedures are enabled for upstream PathErr, Resv and ResvTear messages corresponding to an LSP if link protection is requested for the LSP and the Phop node supports the extensions
- RI-RSVP-FRR extensions and procedures are enabled for upstream PathErr, Resv and ResvTear messages corresponding to an LSP if node protection is requested for the LSP and both Phop and PPhop nodes support the extensions

For example, if an implementation supporting the RI-RSVP-FRR extensions specified in this document is deployed on all routers in particular region of the network and if all the LSPs in the network request node protection, then the FRR extensions will only be applied for the LSP segments that traverse the particular region. This will aid incremental deployment of these extensions and also allow reaping the benefits of the extensions in portions of the network where it is supported.

5. Security Considerations

The security considerations pertaining to the original RSVP protocol [[RFC2205](#)], [[RFC3209](#)] and [[RFC5920](#)] remain relevant.

This document extends the applicability of Node-ID based Hello session between immediate neighbors. The Node-ID based Hello session between PLR and NP-MP may require the two routers to exchange Hello messages with non-immediate neighbor. So, the implementations SHOULD provide the option to configure Node-ID neighbor specific or global authentication key to authentication messages received from Node-ID neighbors. The network administrator MAY utilize this option to enable RSVP-TE routers to authenticate Node-ID Hello messages received with TTL greater than 1. Implementations SHOULD also provide the option to specify a limit on the number of Node-ID based

Hello sessions that can be established on a router supporting the extensions defined in this document.

6. IANA Considerations

6.1. New Object - CONDITIONS

RSVP Change Guidelines [[RFC3936](#)] defines the Class-Number name space for RSVP objects. The name space is managed by IANA.

IANA registry: RSVP Parameters
Subsection: Class Names, Class Numbers, and Class Types

A new RSVP object using a Class-Number from 128-183 range called the "CONDITIONS" object is defined in [Section 4.4](#) of this document. The Class-Number from 128-183 range will be allocated by IANA.

7. Acknowledgements

We are very grateful to Yakov Rekhter for his contributions to the development of the idea and thorough review of content of the draft. Thanks to Raveendra Torvi and Yimin Shen for their comments and inputs.

8. Contributors

Markus Jork
Juniper Networks
Email: mjork@juniper.net

Harish Sitaraman
Juniper Networks
Email: hsitaraman@juniper.net

Vishnu Pavan Beeram
Juniper Networks
Email: vbeeram@juniper.net

Ebben Aries
Juniper Networks
Email: exa@juniper.net

Mike Taillon
Cisco Systems Inc.
Email: mtaillon@cisco.com

9. References

9.1. Normative References

- [I-D.ietf-mpls-summary-frr-rsvpte]
Taillon, M., Saad, T., Gandhi, R., Deshmukh, A., Jork, M.,
and V. Beeram, "RSVP-TE Summary Fast Reroute Extensions
for LSP Tunnels", [draft-ietf-mpls-summary-frr-rsvpte-01](#)
(work in progress), April 2018.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", [BCP 14](#), [RFC 2119](#),
DOI 10.17487/RFC2119, March 1997,
<<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S.
Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1
Functional Specification", [RFC 2205](#), DOI 10.17487/RFC2205,
September 1997, <<https://www.rfc-editor.org/info/rfc2205>>.
- [RFC2961] Berger, L., Gan, D., Swallow, G., Pan, P., Tommasi, F.,
and S. Molendini, "RSVP Refresh Overhead Reduction
Extensions", [RFC 2961](#), DOI 10.17487/RFC2961, April 2001,
<<https://www.rfc-editor.org/info/rfc2961>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
Tunnels", [RFC 3209](#), DOI 10.17487/RFC3209, December 2001,
<<https://www.rfc-editor.org/info/rfc3209>>.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label
Switching (GMPLS) Signaling Resource ReserVation Protocol-
Traffic Engineering (RSVP-TE) Extensions", [RFC 3473](#),
DOI 10.17487/RFC3473, January 2003,
<<https://www.rfc-editor.org/info/rfc3473>>.
- [RFC3936] Kompella, K. and J. Lang, "Procedures for Modifying the
Resource reSerVation Protocol (RSVP)", [BCP 96](#), [RFC 3936](#),
DOI 10.17487/RFC3936, October 2004,
<<https://www.rfc-editor.org/info/rfc3936>>.
- [RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast
Reroute Extensions to RSVP-TE for LSP Tunnels", [RFC 4090](#),
DOI 10.17487/RFC4090, May 2005,
<<https://www.rfc-editor.org/info/rfc4090>>.

- [RFC4558] Ali, Z., Rahman, R., Prairie, D., and D. Papadimitriou, "Node-ID Based Resource Reservation Protocol (RSVP) Hello: A Clarification Statement", [RFC 4558](#), DOI 10.17487/RFC4558, June 2006, <<https://www.rfc-editor.org/info/rfc4558>>.
- [RFC5063] Satyanarayana, A., Ed. and R. Rahman, Ed., "Extensions to GMPLS Resource Reservation Protocol (RSVP) Graceful Restart", [RFC 5063](#), DOI 10.17487/RFC5063, October 2007, <<https://www.rfc-editor.org/info/rfc5063>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8370] Beeram, V., Ed., Minei, I., Shakir, R., Pacella, D., and T. Saad, "Techniques to Improve the Scalability of RSVP-TE Deployments", [RFC 8370](#), DOI 10.17487/RFC8370, May 2018, <<https://www.rfc-editor.org/info/rfc8370>>.

9.2. Informative References

- [RFC5439] Yasukawa, S., Farrel, A., and O. Komolafe, "An Analysis of Scaling Issues in MPLS-TE Core Networks", [RFC 5439](#), DOI 10.17487/RFC5439, February 2009, <<https://www.rfc-editor.org/info/rfc5439>>.
- [RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", [RFC 5920](#), DOI 10.17487/RFC5920, July 2010, <<https://www.rfc-editor.org/info/rfc5920>>.

Authors' Addresses

Chandra Ramachandran
Juniper Networks

Email: csekar@juniper.net

Ina Minei
Google, Inc

Email: inaminei@google.com

Dante Pacella
Verizon

Email: dante.j.pacella@verizon.com

Tarek Saad
Cisco Systems Inc.

Email: tsaad@cisco.com