Stream: Internet Engineering Task Force (IETF)
RFC: 9705
Updates: 4090
Category: Standards Track
Published: December 2024
ISSN: 2070-1721
Authors: C. Ramachandran, *Juniper Networks, Inc.*; T. Saad, *Cisco Systems, Inc.*; I. Minei, *Google, Inc.*; D. Pacella, *Verizon, Inc.*

RFC 9705
Refresh-Interval Independent Fast Reroute (FRR) Facility Protection

Abstract

The RSVP-TE Fast Reroute (FRR) extensions specified in RFC 4090 define two local repair techniques to reroute Label Switched Path (LSP) traffic over pre-established backup tunnels. Facility backup methods allow 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 techniques is attractive from a scalability point of view. This document enumerates facility backup procedures in RFC 4090 that rely on refresh timeout, hence, making facility backup methods 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, and hence, compatible with the Refresh-Interval Independent RSVP (RI-RSVP) capability specified in RFC 8370. Hence, this document updates RFC 4090 in order to support the RI-RSVP capability specified in RFC 8370.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

RSVP-TE relies on a periodic refresh of RSVP messages to synchronize and maintain the states related to the Label Switched Path (LSP) 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 increases in LSP scale.

[RFC2961] specifies mechanisms to eliminate the reliance on periodic refreshes and refresh timeouts of RSVP messages and enables a router to increase the message refresh interval to values much longer than the default 30 seconds defined in [RFC2205]. However, the protocol extensions defined in [RFC4090] for supporting Fast Reroute (FRR) using bypass tunnels implicitly rely on short refresh timeouts to clean up stale states.

In order to eliminate the reliance on refresh timeouts, the routers should unambiguously determine when a particular LSP state should be deleted. In scenarios involving FRR using bypass tunnels [RFC4090], additional explicit teardown messages are necessary. The Refresh-Interval Independent RSVP FRR (RI-RSVP-FRR) extensions specified in this document consist of procedures to enable LSP state cleanup that are essential in supporting the RI-RSVP capability for FRR using bypass tunnels from [RFC4090].

1.1. Motivation

Base RSVP [RFC2205] maintains state via the generation of RSVP Path and 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 and 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 [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 also 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 timeouts for stale state cleanup, refer to Section 3 of [RFC8370].

However, the facility backup protection procedures specified in [RFC4090] do not fully address stale state cleanup as the procedures depend on refresh timeouts for stale state cleanup. The updated facility backup protection procedures specified in this document, in combination with RSVP-TE Scaling Techniques [RFC8370], eliminate this dependency on refresh timeouts for stale state cleanup.

The procedures specified in this document assume reliable delivery of RSVP messages, as specified in [RFC2961]. Therefore, this document makes support for [RFC2961] a prerequisite.

2. Terminology

The reader is expected to be familiar with the terminology in [RFC2205], [RFC3209], [RFC4090], [RFC4558], [RFC8370], and [RFC8796].

Phop node: Previous-Hop router along the LSP

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

Nhop node: Next-Hop router along the LSP

NNhop node: Next-Next-Hop router along the LSP

PLR: Point of Local Repair router as defined in [RFC4090]

MP: Merge Point router as defined in [RFC4090]

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

PSB: Path State Block

RSB: Reservation State Block

RRO: Record Route Object as defined in [RFC3209]

TED: Traffic Engineering Database

LSP state: The combination of "path state" maintained as a PSB and "reservation state" maintained as an RSB forms an individual LSP state on an RSVP-TE speaker

RI-RSVP: The set of procedures defined in Section 3 of [RFC8370] to eliminate RSVP's reliance on periodic message refreshes

B-SFRR-Ready: Bypass Summary FRR Ready Extended Association object as defined in [RFC8796] and added by the PLR for each protected LSP

RI-RSVP-FRR: The set of procedures defined in this document to eliminate RSVP's reliance on periodic message refreshes when supporting facility backup protection [RFC4090]

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

Remote PathTear: A PathTear message sent from a PLR to the MP to delete the LSP state on the MP if the PLR had not previously sent the backup path state reliably

2.1. Requirements Language

The key words "**MUST**", "**MUST NOT**", "**REQUIRED**", "**SHALL**", "**SHALL NOT**", "**SHOULD**", "**SHOULD NOT**", "**RECOMMENDED**", "**NOT RECOMMENDED**", "**MAY**", and "**OPTIONAL**" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Problem Description

```
                E
              /   \
             /     \
            /       \
           /         \
          /           \
         /             \
        A ----- B ----- C ----- D
                \             /
                 \           /
                  \         /
                   \       /
                    \     /
                     \   /
                       F
```

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.

In addition, 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 PLR and C is the 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 LP-MP.

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

Step 1. B performs a local repair and redirects LSP traffic over the bypass LSP B -> F -> D.

Step 2. B also creates a backup state for the LSP and triggers the sending of a backup LSP state to D over the bypass LSP B -> F -> D.

Step 3. D receives the backup LSP states and merges the backups with the protected LSPs.

Step 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 the teardown messages for all LSPs. As each router should consider itself as an MP, C will time out the state only after waiting for an additional duration equal to the 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 a 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 the default refresh time.
- If C is to keep state until its timeout upon the link failure, then with a long refresh interval, this may result in a large amount of stale state on C. Alternatively, if C is to delete the state and send a PathTear to D upon the link failure, then this would result in deleting the state on D, thus deleting the LSP. D needs a reliable mechanism to determine whether or not it is an MP to overcome this problem.
- If head-end A attempts to tear down the LSP after Step 1 but before Step 2 of the above sequence, then B may receive the teardown message before Step 2 and delete the LSP state from its state database. If B deletes its state without informing D, with a long refresh interval, this could cause a (large) buildup of stale state on D.
- If B fails to perform a 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 a long refresh interval, this could cause a (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:

1. Utilize the MP determination mechanism specified in RSVP-TE Summary FRR [RFC8796] 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 sessions between the PLR and the MP to detect router failures and to determine capability. See Section 4.2 of this document for more details. This part of the solution reuses some of the extensions defined in [RFC8796] and [RFC8370], and the subsequent subsections will list the extensions in these documents that are utilized in this document.

2. Handle upstream link or node failures by cleaning up LSP states if the node has not found itself as an MP through the MP determination mechanism. See Section 4.3 of this document for more details.
3. Introduce extensions to enable a router to send a teardown message to the downstream router that enables the receiving router to conditionally delete its local LSP state. See Section 4.4 of this document for more details.
4. Enhance facility backup protection by allowing a PLR to directly send a teardown message to the MP without requiring the PLR to either have a working bypass LSP or have already signaled the backup LSP state. See Section 4.5 of this document for more details.
5. Introduce extensions to enable the above procedures to be backward compatible with routers along the LSP path running implementations that do not support these procedures. See Section 4.6 of this document for more details.

4.1. Requirement on RFC 4090 Capable Node to Advertise the RI-RSVP Capability

A node supporting facility backup protection [RFC4090] **MUST NOT** set the RI-RSVP flag (I-bit) that is defined in Section 3.1 of [RFC8370] unless it supports all the extensions specified in the rest of this document. Hence, this document updates [RFC4090] by defining extensions and additional procedures over facility backup protection [RFC4090] in order to advertise the RI-RSVP capability [RFC8370]. However, if a node supporting facility backup protection [RFC4090] does set the RI-RSVP capability (I-bit) but does not support all the extensions specified in the rest of this document, then it may result in lingering stale states due to the long refresh intervals recommended by [RFC8370]. This can also disrupt normal Fast Reroute (FRR) operations. Section 4.7 of this document delves into this in detail.

4.2. Signaling Handshake Between PLR and MP

4.2.1. PLR Behavior

As per the facility backup procedures [RFC4090], when an LSP becomes operational on a node and the "local protection desired" flag has been set in the SESSION_ATTRIBUTE object carried in the Path message corresponding to the LSP, then the node 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 an NP-bypass LSP to the NNhop node avoiding the Nhop node on a protected LSP path. In case node protection could not be made available, the node attempts to create an LP-bypass LSP to the Nhop node avoiding only the link that the protected LSP takes to reach the Nhop.
- If the "node protection desired" flag is not set, then the PLR attempts to create an LP-bypass LSP to the Nhop node avoiding the link that the protected LSP takes to reach the Nhop.

With regard to the PLR procedures described above and specified in [RFC4090], this document specifies the following additional procedures to support RI-RSVP [RFC8370].

- While selecting the destination address of the bypass LSP, the PLR **MUST** select the router ID of the NNhop or Nhop node from the Node-ID sub-object included in the RRO object that is carried in the most recent Resv message corresponding to the LSP. If the MP has not included a Node-ID sub-object in the 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 that is included by the MP in the RRO object. If the NP-MP in a different IGP area has not included a Node-ID sub-object in the RRO object, then the PLR **MUST** 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 **MUST** also include its router ID in a Node-ID sub-object in the RRO object that is carried in any subsequent Path message corresponding to the LSP. 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 an NP-bypass or an LP-bypass, the PLR **MUST** initiate a Node-ID-based Hello session to the NNhop or Nhop node respectively along the LSP to establish the RSVP-TE signaling adjacency. This Hello session is used to detect MP node failure as well as to determine the capability of the MP node. If the MP has set the I-bit in the CAPABILITY object [RFC8370] carried in the Hello message corresponding to the Node-ID-based Hello session, then the PLR **MUST** 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 the CAPABILITY object [RFC8370], then the PLR **MUST** execute backward compatibility procedures defined in Section 4.6.2.1 of this document.
- When the PLR associates a bypass to a protected LSP, it **MUST** include a B-SFRR-Ready Extended Association object [RFC8796] and trigger a Path message to be sent for the LSP. If a B-SFRR-Ready Extended Association object is included in the Path message corresponding to the LSP, the encoding and object ordering rules specified in RSVP-TE Summary FRR [RFC8796] **MUST** be followed. In addition to those rules, the PLR **MUST** set the Association Source in the object to its Node-ID address.

4.2.2. Remote Signaling Adjacency

A Node-ID-based RSVP-TE Hello session is one in which a 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 sessions 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 sessions between any two routers that are not immediate neighbors, the router that supports the extensions defined in the document **MUST** set the TTL to 255 in all outgoing Node-ID-based Hello messages exchanged between the PLR and the MP. The default hello interval for this Node-ID Hello session **MUST** be set to the default specified in RSVP-TE Scaling Techniques [RFC8370].

In the rest of the document, the terms "signaling adjacency" and "remote signaling adjacency" refer specifically to the RSVP-TE signaling adjacency.

4.2.3. MP Behavior

With regard to the MP procedures that are defined in [RFC4090], this document specifies the following additional procedures to support RI-RSVP as defined in [RFC8370].

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

A node receiving a Path message should determine:

- whether the message contains a B-SFRR-Ready Extended Association object with its own address as the bypass destination address and
- whether it has an operational Node-ID signaling adjacency with the Association source.

The node **MUST** execute the backward compatibility procedures defined in Section 4.6.2.2 of this document if:

- the PLR has not included the B-SFRR-Ready Extended Association object,
- there is no operational Node-ID signaling adjacency with the PLR identified by the Association source address, or
- the PLR has not advertised the RI-RSVP capability in its Node-ID-based Hello messages.

If a matching B-SFRR-Ready Extended Association object is found in the Path message and if there is an operational remote Node-ID signaling adjacency with the PLR (identified by the Association source) that has advertised the RI-RSVP capability (I-bit) [RFC8370], then the node **MUST** consider itself as the MP for the PLR. The matching and ordering rules for Bypass Summary FRR Extended Association specified in RSVP-TE Summary FRR [RFC8796] **MUST** be followed by the 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 **MUST** conclude it is the 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 **MUST** conclude it is the 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 as described in the preceding section, it **MUST** create a remote path state for the LSP. The only difference between the "remote" path state and the LSP state is the RSVP_HOP object. The RSVP_HOP object in a "remote" path state contains the address that the PLR uses to send Node-ID Hello messages to the MP.

The MP **MUST** consider the "remote" path state corresponding to the LSP automatically deleted if:

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

The purpose of "remote" path state is to enable the PLR to explicitly tear down the path and reservation states corresponding to the LSP by sending a tear message for the "remote" path state. Such a message tearing down the "remote" path state is called "Remote" PathTear.

The scenarios in which a "Remote" PathTear is applied are described in Section 4.5 of this document.

4.3. Impact of Failures on LSP State

This section describes the procedures that must be executed upon different kinds of failures by nodes along the path of the LSP. The procedures that must be executed upon detecting RSVP signaling adjacency failures do not impact the RSVP-TE graceful restart mechanisms [RFC3473] [RFC5063]. If a node executing these procedures acts as a helper for a neighboring router, then the signaling adjacency with the neighbor will be declared as having failed only after taking into account the grace period extended for the neighbor by this node acting as a helper.

Node failures are detected from the state of Node-ID Hello sessions established with immediate neighbors. RSVP-TE Scaling Techniques [RFC8370] recommends that each node establish Node-ID Hello sessions with all its immediate neighbors. A non-immediate PLR or MP 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 the Phop link or the Phop node failure for an LSP and the router is not an MP for the LSP, then it **MUST** send a Conditional PathTear (refer to Section 4.4 of this document) and delete the 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 an LP-MP for the LSP, the LP-MP **MUST** retain the PSB and RSB states corresponding to the LSP until the occurrence of any of the following events:

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

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

4.3.3. NP-MP Behavior

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

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

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

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

Receiving a Conditional PathTear from the Phop node will not impact the "remote" state from the PPhop PLR. Note that the Phop node must have sent the Conditional PathTear as it was not an MP for the LSP (see Section 4.3.1 of this document).

In the example topology in Figure 1, we assume C and D are the NP-MPs for the PLRs A and B, respectively. Now, when the A-B link fails, B will delete the LSP state, because B is not an MP and its Phop link has failed (this behavior is required for unprotected LSPs; refer to Section 4.3.1 of this document). In the data plane, that would require B to delete the label forwarding entry corresponding to the LSP. Thus, 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 the NP-MP for the 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 a Conditional PathTear from B, it decides to retain the LSP state as it is the NP-MP of the PLR A. It also checks whether Phop B had previously signaled availability of node protection. As B had previously signaled NP availability by including the B-SFRR-Ready Extended Association object, C removes the B-SFRR-Ready Extended Association object containing the Association Source set to B from the Path message and triggers a Path to D.
2. When D receives the Path message, 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 the 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 the LP-MP and NP-MP

A router may simultaneously be the LP-MP and the NP-MP for the Phop and PPhop nodes of an LSP, respectively. If the Phop link fails on such a node, the node **MUST** retain the PSB and RSB states corresponding to the LSP until the occurrence of any of the following events:

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

If a router that is both an LP-MP and an NP-MP detects Phop node failure, then the node **MUST** retain the PSB and RSB states corresponding to the LSP until the occurrence of any of the following events:

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

4.4. Conditional PathTear

In the example provided in Section 4.3.3 of this document, B deletes the PSB and RSB states corresponding to the LSP once B detects its Phop link that went down as B is not an MP. If B were to send a PathTear normally, then C would delete the 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 **MUST** add a new optional CONDITIONS object in the PathTear. The CONDITIONS object is defined in Section 4.4.3 of this document. If node C also understands the new object, then C **MUST NOT** delete the LSP state if it is an NP-MP.

4.4.1. Sending the Conditional PathTear

A router that is not an MP for an LSP **MUST** delete the PSB and RSB states corresponding to the LSP if the Phop link or the Phop Node-ID signaling adjacency goes down (see Section 4.3.1 of this document). The router **MUST** send a Conditional PathTear if the following are also true:

- the ingress has requested node protection for the LSP and
- no PathTear is received from the upstream node.

4.4.2. Processing the Conditional PathTear

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

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

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

4.4.3. CONDITIONS Object

Any implementation that does not support a Conditional PathTear needs to ignore the new object but process the message as a normal PathTear without generating any error. For this reason, the Class-Num of the new object follows the pattern 10bbbbbb, where "b" represents a bit. (The behavior for objects of this type is specified in Section 3.10 of [RFC2205].)

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

The object has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Length               |  Class        |     C-type    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Flags (Reserved)                           |M|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: CONDITIONS Object

Class: 135
C-type: 1
Flags: 32 bit field
M: Bit 31 is the Merge-point condition (M) bit. If the M bit is set to 1, then the PathTear message **MUST** be processed according to the receiver router role, i.e., if the receiving router is an MP or not for the LSP. If it is not set, then the PathTear message **MUST** be processed as a normal PathTear message for the LSP.

Bits 0-30 are reserved; they **MUST** be set to zero on transmission and **MUST** be ignored on receipt.

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 until the completion of backup LSP signaling to perform state cleanup. In this case, the PLR **MUST** send a "Remote" PathTear message instructing the MP to delete the PSB and RSB states corresponding to the LSP. The TTL in the "Remote" PathTear message **MUST** be set to 255. Doing this enables LSP state cleanup when the LSP is being locally repaired.

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

1. Ingress A receives a management event to tear down the LSP.
2. A sends a normal PathTear for the LSP to B.
3. Assume B has not initiated the backup signaling for the LSP during local repair. To enable LSP state cleanup, B sends a "Remote" PathTear with the destination IP address set to that of the node D used in the Node-ID signaling adjacency with D and the RSVP_HOP object containing the local address used in the Node-ID signaling adjacency.
4. B then deletes the PSB and RSB states corresponding to the LSP.
5. On D, there would be a remote signaling adjacency with B, and so D accepts the "Remote" PathTear and deletes the 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, the PLR **MUST** send a "Remote" PathTear to the MP. The purpose of this is to clean up LSP state from the PLR to the egress. Upon receiving the PathTear, the MP **MUST** delete the states corresponding to the LSP and also propagate the PathTear downstream, thereby achieving state cleanup from all downstream nodes up to the LSP egress. Note that in the case of link protection, the PathTear **MUST** be directed to the LP-MP's Node-ID IP address rather than the Nhop interface address.

4.5.2. PLR Behavior on Resv RRO Change

When a PLR router that has already made NP available for an LSP detects a change in the RRO carried in the Resv message that indicates that the router's former NP-MP is no longer present on the path of the LSP, then the router **MUST** send a "Remote" PathTear directly to its former NP-MP.

In the example topology in Figure 1, assume node A has made node protection available for an LSP and C has concluded it is the NP-MP for PLR A. When the B-C link fails, then C, implementing the procedure specified in Section 4.3.4 of this document, will retain the states corresponding to the LSP until one of the following occurs:

- the remote Node-ID signaling adjacency with A goes down or
- a PathTear or a ResvTear is received for its 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 a Resv to A with RRO changed. The new RRO of the LSP carried in the Resv will not contain C. When A processes the Resv message with a new RRO not containing C, its former NP-MP, A, sends a "Remote" PathTear to C. When C receives the "Remote" PathTear for its PSB state, C will send a normal PathTear downstream to D and delete both the PSB and RSB states corresponding to the LSP. As D has already received backup LSP signaling from B, D will retain the 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 an LP-MP after its Phop link has already failed but the backup LSP has not been signaled yet as part of the local repair procedure, then the node **MUST** send a normal PathTear and delete both the PSB and RSB states corresponding to the LSP. As the LP-MP has retained the LSP state expecting the PLR to initiate backup LSP signaling, preemption would bring down the LSP and the node would not be LP-MP anymore, requiring the node to clean up the LSP state.

4.5.3.2. Preemption on NP-MP After Phop Link Failure

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

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

1. The LSP is preempted on C.
2. C will delete the RSB state corresponding to the LSP. However, C cannot send a PathErr or a ResvTear to the PLR A because the 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 an NP-MP, C sends a normal PathTear to D and also deletes its PSB state. D would also delete the PSB and RSB states on receiving a PathTear from C.
4. B starts backup LSP signaling to D. However, as D does not have the LSP state, it will reject the backup LSP Path and send a PathErr to B.
5. B will delete its reservation and send a ResvTear to A.

4.6. Backward Compatibility Procedures

"Refresh-Interval Independent FRR" and "RI-RSVP-FRR" refer to the set of procedures defined in this document to eliminate the reliance on periodic refreshes. The extensions proposed in RSVP-TE Summary FRR [RFC8796] may apply to implementations that do not support RI-RSVP-FRR. On the other hand, RI-RSVP-FRR extensions relating to LSP state cleanup, namely Conditional and "Remote" PathTears, require support from one-hop and two-hop neighboring nodes along the LSP path. Thus, procedures that fall under the LSP state cleanup category **MUST NOT** be turned on if any of the nodes involved in the node protection FRR (i.e., the PLR, the MP, and the intermediate node in the case of NP) do not support RI-RSVP-FRR extensions. Note that for LSPs requesting link protection, only the PLR and the LP-MP **MUST** support the extensions.

4.6.1. Detecting Support for Refresh-Interval Independent FRR

An implementation supporting RI-RSVP-FRR extensions **MUST** set the flag "Refresh interval Independent RSVP" or RI-RSVP flag in the CAPABILITY object carried in Hello messages as specified in RSVP-TE Scaling Techniques [RFC8370]. If an implementation does not set the flag even if it supports RI-RSVP-FRR extensions, then its neighbors will view the node as any node that does not support the extensions.

- As nodes supporting the RI-RSVP-FRR extensions initiate Node-ID-based signaling adjacency with all immediate neighbors, such a node on the path of a protected LSP can determine whether its Phop and Nhop neighbors support RI-RSVP-FRR enhancements.
- As nodes supporting the RI-RSVP-FRR extensions also initiate Node-ID-based signaling adjacency with the NNhop along the path of the LSP requesting node protection (see Section 4.2.1 of this document), each node along the LSP path can determine whether its NNhop

node supports RI-RSVP-FRR enhancements. If the NNhop (a) does not reply to remote Node-ID Hello messages or (b) does not set the RI-RSVP flag in the CAPABILITY object carried in its Node-ID Hello messages, then the node acting as the PLR can conclude that NNhop does not support RI-RSVP-FRR extensions.

- If node protection is requested for an LSP and if (a) the PPhop node has not included a matching B-SFRR-Ready Extended Association object in its Path messages, (b) the PPhop node has not initiated remote Node-ID Hello messages, or (c) the PPhop node does not set the RI-RSVP flag in the CAPABILITY object carried in its Node-ID Hello messages, then the node **MUST** conclude that the PLR does not support RI-RSVP-FRR extensions.

4.6.2. Procedures for Backward Compatibility

Every node that supports RI-RSVP-FRR **MUST** support the procedures defined in this section in order to support backward compatibility for those subsets of LSPs that also traverse nodes that do not support RI-RSVP-FRR.

4.6.2.1. Lack of Support on Downstream Nodes

The procedures on the downstream direction are as follows:

- If a node finds that the Nhop node along the LSP does not support the RI-RSVP-FRR extensions, then the node **MUST** reduce the "refresh period" in the TIME_VALUES object carried in the Path messages to the default short refresh interval.
- If node protection is requested for the LSP and the NNhop node along the LSP path does not support the RI-RSVP-FRR extensions, then the node **MUST** reduce the "refresh period" in the TIME_VALUES object carried in the Path messages to the default short refresh interval.

If a node reduces the refresh time using the above procedures, it **MUST NOT** send any "Remote" PathTear or Conditional PathTear message to the downstream node.

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

- A and B reduce the refresh time to the default short refresh interval of 30 seconds and trigger a Path message.
- If B is not an MP and if the Phop link of B fails, B cannot send a Conditional PathTear to C but times out the PSB state from A normally. Note that B can only normally time out the PSB state A if A did not set the long refresh in the TIME_VALUES object carried in the Path messages sent earlier.

4.6.2.2. Lack of Support on Upstream Nodes

The procedures on the upstream direction are as follows:

- If a node finds that the Phop node along the LSP path does not support the RI-RSVP-FRR extensions, then the node **MUST** reduce the "refresh period" in the TIME_VALUES object carried in the Resv messages to the default short refresh interval.

- If node protection is requested for the LSP and the Phop node along the LSP path does not support the RI-RSVP-FRR extensions, then the node **MUST** reduce the "refresh period" in the TIME_VALUES object carried in the Path messages to the default short refresh interval (thus, the Nhop can use compatible values when sending a Resv).
- If node protection is requested for the LSP and the PPhop node does not support the RI-RSVP-FRR extensions, then the node **MUST** reduce the "refresh period" in the TIME_VALUES object carried in the Resv messages to the default short refresh interval.
- If the node reduces the refresh time using the above procedures, it **MUST 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 this 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 and 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 a 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.

4.7. Consequences of Advertising RI-RSVP Without RI-RSVP-FRR

If a node supporting facility backup protection [RFC4090] sets the RI-RSVP capability (I-bit) but does not support the RI-RSVP-FRR extensions, due to an implementation bug or configuration error, then it leaves room for the stale state to linger around for an inordinate period of time or for disruption of normal FRR operations (see Section 3 of this document). Consider the example topology (Figure 1) provided in this document.

- Assume node B does set the RI-RSVP capability in its Node-ID-based Hello messages even though it does not support RI-RSVP-FRR extensions. When B detects the failure of its Phop link along an LSP, it will not send a Conditional PathTear to C as required by the RI-RSVP-FRR procedures. If B simply leaves the LSP state without deleting, then B may end up holding on to the stale state until the (long) refresh timeout.
- Instead of node B, assume node C does set the RI-RSVP capability in its Node-ID-based Hello messages even though it does not support RI-RSVP-FRR extensions. When B details the failure of its Phop link along an LSP, it will send a Conditional PathTear to C as required by the RI-RSVP-FRR procedures. However, C would not recognize the condition encoded in the PathTear and end up tearing down the LSP.
- Assume node B does set the RI-RSVP capability in its Node-ID-based Hello messages even though it does not support RI-RSVP-FRR extensions. In addition, assume local repair is about to commence on node B for an LSP that has only requested link protection, that is, B has not initiated the backup LSP signaling for the LSP. If node B receives a normal PathTear at this time from ingress A because of a management event initiated on A, then B simply deletes the LSP state without sending a Remote PathTear to the LP-MP C, so C may end up holding on to the stale state until the (long) refresh timeout.

5. Security Considerations

The security considerations pertaining to the original RSVP protocols ([RFC2205], [RFC3209], and [RFC5920]) remain relevant. When using RSVP cryptographic authentication [RFC2747], more robust algorithms such as HMAC-SHA256, HMAC-SHA384, or HMAC-SHA512 [RFC2104] [FIPS-180-4] **SHOULD** be used when computing the keyed message digest where possible.

This document extends the applicability of Node-ID-based Hello sessions between immediate neighbors. The Node-ID-based Hello session between the PLR and the NP-MP may require the two routers to exchange Hello messages with a non-immediate neighbor. Therefore, the implementations **SHOULD** provide the option to configure a Node-ID neighbor specific or global authentication key to authentication messages received from Node-ID neighbors. The network administrator **SHOULD** utilize this option to enable RSVP-TE routers to authenticate Node-ID Hello messages received with a 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. CONDITIONS Object

IANA maintains the "Class Names, Class Numbers, and Class Types" registry in the "RSVP Parameters" registry group (see http://www.iana.org/assignments/rsvp-parameters/). IANA has extended these registries by adding a new Class Number (in the 10bbbbbb range) and assigning a new C-Type under this Class Number, as described below (see Section 4.4.3):

Class Number	Class Name	Reference
135	CONDITIONS	RFC 9705

Table 1: Class Names, Class Numbers, and Class Types

Value	Description	Reference
1	CONDITIONS	RFC 9705

Table 2: Class Type or C-Types - 135 CONDITIONS

IANA has added a subregistry called "CONDITIONS Object Flags" as shown below. Assignments of additional Class Type values for Class Name "CONDITIONS" are to be performed via "IETF Review" [RFC8126].

Bit Number	32-Bit Value	Name	Reference
0-30		Unassigned	
31	0x0001	Merge-point	RFC 9705

Table 3: CONDITIONS Object Flags

All assignments in this subregistry are to be performed via "IETF Review" [RFC8126].

7. References

7.1. Normative References

[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>.

[RFC2747] Baker, F., Lindell, B., and M. Talwar, "RSVP Cryptographic Authentication", RFC 2747, DOI 10.17487/RFC2747, January 2000, <https://www.rfc-editor.org/info/rfc2747>.

[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>.

[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>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[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>.

[RFC8796] Taillon, M., Saad, T., Ed., Gandhi, R., Deshmukh, A., Jork, M., and V. Beeram, "RSVP-TE Summary Fast Reroute Extensions for Label Switched Path (LSP) Tunnels", RFC 8796, DOI 10.17487/RFC8796, July 2020, <https://www.rfc-editor.org/info/rfc8796>.

7.2. Informative References

[FIPS-180-4] National Institute of Standards and Technology, "Secure Hash Standard", FIPS PUB 180-4, DOI 10.6028/NIST.FIPS.180-4, August 2015, <https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.180-4.pdf>.

[RFC2104] Krawczyk, H., Bellare, M., and R. Canetti, "HMAC: Keyed-Hashing for Message Authentication", RFC 2104, DOI 10.17487/RFC2104, February 1997, <https://www.rfc-editor.org/info/rfc2104>.

[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>.

Acknowledgements

We are very grateful to Yakov Rekhter for his contributions to the development of the idea and thorough review of the content of the document. We are thankful to Raveendra Torvi and Yimin Shen for their comments and inputs on early versions of the document. We also thank Alexander Okonnikov for his review and comments on the document.

Contributors

Markus Jork
Juniper Networks, Inc.
Email: mjork@juniper.net

Harish Sitaraman
Individual Contributor
Email: harish.ietf@gmail.com

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

Ebben Aries
Juniper Networks, Inc.
Email: exa@juniper.com

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

Authors' Addresses

Chandra Ramachandran
Juniper Networks, Inc.
Email: csekar@juniper.net

Tarek Saad
Cisco Systems, Inc.
Email: tsaad@cisco.com

Ina Minei
Google, Inc.
Email: inaminei@google.com

Dante Pacella
Verizon, Inc.
Email: dante.j.pacella@verizon.com