MPLS Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 3, 2019

M. Taillon
Cisco Systems, Inc.
T. Saad, Ed.
Juniper Networks
R. Gandhi
Cisco Systems, Inc.
A. Deshmukh
Juniper Networks
M. Jork
128 Technology
V. Beeram
Juniper Networks
May 02, 2019

RSVP-TE Summary Fast Reroute Extensions for LSP Tunnels
draft-ietf-mpls-summary-frr-rsvpte-03

Abstract

This document defines Resource Reservation Protocol (RSVP) Traffic-Engineering (TE) signaling extensions that reduce the amount of RSVP signaling required for Fast Reroute (FRR) procedures and subsequently improve the scalability of the RSVP-TE signaling when undergoing FRR convergence after a link or node failure. Such extensions allow the RSVP message exchange between the Point of Local Repair (PLR) and the Merge Point (MP) to be independent of the number of protected Label Switched Paths (LSPs) traversing between them when facility bypass FRR protection is used. The signaling extensions are fully backwards compatible with nodes that do not support them.

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 November 3, 2019.

Table of Contents

1. Introduction

The Fast Reroute (FRR) procedures defined in [RFC4090] describe the mechanisms for the Point of Local Repair (PLR) to reroute traffic and signaling of a protected RSVP-TE LSP onto the bypass tunnel in the event of a TE link or node failure. Such signaling procedures are performed individually for each affected protected LSP. This may eventually lead to control plane scalability and latency issues on the PLR and/or the MP due to limited memory and CPU processing resources. This condition is exacerbated when the failure affects large number of protected LSPs that traverse the same PLR and Merge Point (MP) nodes.

For example, in a large scale RSVP-TE LSPs deployment, a single LSR acting as a PLR node may host tens of thousands of protected RSVP-TE LSPs egressing the same link, and also act as a MP node for similar number of LSPs ingressing the same link. In the event of the failure of the link or neighbor node, the RSVP-TE control plane of the node when acting as PLR becomes busy rerouting protected LSPs signaling over the bypass tunnel(s) in one direction, and when acting as an MP node becomes busy merging RSVP states from signaling received over bypass tunnels for LSP(s) in the reverse direction. Subsequently, the head-end LER(s) that are notified of the local repair at downstream LSR will attempt to (re)converge affected RSVP- TE LSPs onto newly computed paths - possibly traversing the same previously affected LSR(s). As a result, the RSVP-TE control plane at the PLR and MP becomes overwhelmed by the amount of FRR RSVP-TE processing overhead following the link or node failure, and the competing other control plane protocol(s) (e.g. the IGP) that undergo their convergence at the same time.

The extensions defined in this document enable a MP node to become aware of the PLR node's bypass tunnel assignment group and allow FRR procedures between PLR node and MP node to be signaled and processed on groups of LSPs.

As defined in [RFC2961], Summary Refresh procedures use MESSAGE_ID to refresh the RSVP Path and Resv states to help with the scale. The MESSAGE_ID information for the rerouted PATH and RESV states are exchanged between PLR and MP nodes between PLR and MP nodes a priori to the fault such that Summary Refresh procedures defined in [RFC2961] can continue to be used to refresh the rerouted state(s) after FRR has occurred.

2. Conventions Used in This Document

2.1. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [RFC2119].

2.2. Acronyms and Abbreviations

The reader is assumed to be familiar with terms and abbreviations used in [RFC3209] and [RFC4090].

The following abbreviations are also used in this document:

LSR: Label Switching Router

LER: Label Edge Router

MPLS: Multiprotocol Label Switching

LSP: Label Switched Path

MP: Merge Point node as defined in [RFC4090]

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

FRR: Fast Reroute as defined in [RFC4090]

B-SFRR-Ready: Bypass Summary FRR Ready Extended ASSOCIATION object. Added by the PLR for each LSP protected by the bypass tunnel.

B-SFRR-Active: Bypass Summary FRR Active Extended ASSOCIATION object. Used to notify the MP node of one ore more groups of protected LSP(s) that are being protected by the specified bypass tunnel are being rerouted.

3. Extensions for Summary FRR Signaling

The RSVP ASSOCIATION object is defined in [RFC4872] as a means to associate LSPs with each other. For example, in the context of GMPLS-controlled LSP(s), the object is used to associate recovery LSPs with the LSP they are protecting. The Extended ASSOCIATION object is introduced in [RFC6780] to expand on the possible usage of the ASSOCIATION object and generalize the definition of the Extended Association ID field.

This document proposes the use of the Extended ASSOCIATION object to carry the Summary FRR information and associate the protected LSP(s) with the bypass tunnel that protects them. Two new Association Types for the Extended ASSOCIATION object, and new Extended Association IDs are proposed in this draft to describe the Bypass Summary FRR Ready (B-SFRR-Ready) and the Bypass Summary FRR Active (B-SFRR-Active) associations.

The PLR creates and manages the Summary FRR LSP groups (Bypass_Group_Identifiers) and shares them with the MP via signaling. Protected LSPs sharing the same egress link and bypass assignment are grouped together and are assigned the same group. The MP maintains the PLR group assignments learned via signaling, and acknowledges the group assignments via signaling. Once the PLR receives the acknowledgment, FRR signaling can proceed as group based.

The PLR node that supports Summary FRR procedures adds the Extended ASSOCIATION object with Type B-SFRR-Ready and respective Extended Association ID in the RSVP Path message of the protected LSP to inform the MP of the PLR's assigned bypass tunnel, Summary FRR Bypass_Group_Identifier, and the MESSAGE_ID that the PLR will use to refresh the protected LSP PATH state after FRR occurs.

The MP node that supports Summary FRR procedures adds the B-SFRR-Ready Extended ASSOCIATION object and respective Extended Association ID in the RSVP Resv message of the protected LSP to acknowledge the PLR's bypass tunnel assignment, and provide the MESSAGE_ID object that the MP node will use to refresh the protected LSP RESV state after FRR occurs.

This document also defines a new Association Type for the Extended ASSOCIATION object and new Extended Association ID to describe the B-SFRR-Active association. The B-SFRR-Active Extended ASSOCIATION object and Extended Association ID are sent by PLR after activating FRR procedures on the PLR. The B-SFRR-Active Extended ASSOCIATION object and Extended Association ID are sent within the RSVP Path message of the bypass LSP to inform the MP node that one or more groups of protected LSPs protected by the bypass tunnel are now being rerouted over the bypass tunnel.

3.1. B-SFRR-Ready Extended ASSOCIATION Object

The Extended ASSOCIATION object is populated using the rules defined below to associate a protected LSP with the bypass LSP that is protecting it when Summary FRR procedures are enabled.

The Association Type, Association ID, and Association Source MUST be set as defined in [RFC4872] for the ASSOCIATION Object. More specifically:

Association Source:

The Association Source is set to an address of the PLR node.

Association Type:

A new Association Type is defined for B-SFRR-Ready as follows:

Value	Type
(TBD-1)	Bypass Summary FRR Ready Association (B-SFRR-Ready)

Extended ASSOCIATION ID for B-SFRR-Ready:

The B-SFRR-Ready Extended ASSOCIATION ID is populated by the PLR for the Bypass Summary FRR Ready association. The rules to populate the Extended ASSOCIATION ID in this case are described below.

3.1.1. IPv4 B-SFRR-Ready IPv4 Extended ASSOCIATION ID

The IPv4 Extended ASSOCIATION ID for the B-SFRR-Ready association type has the following format:

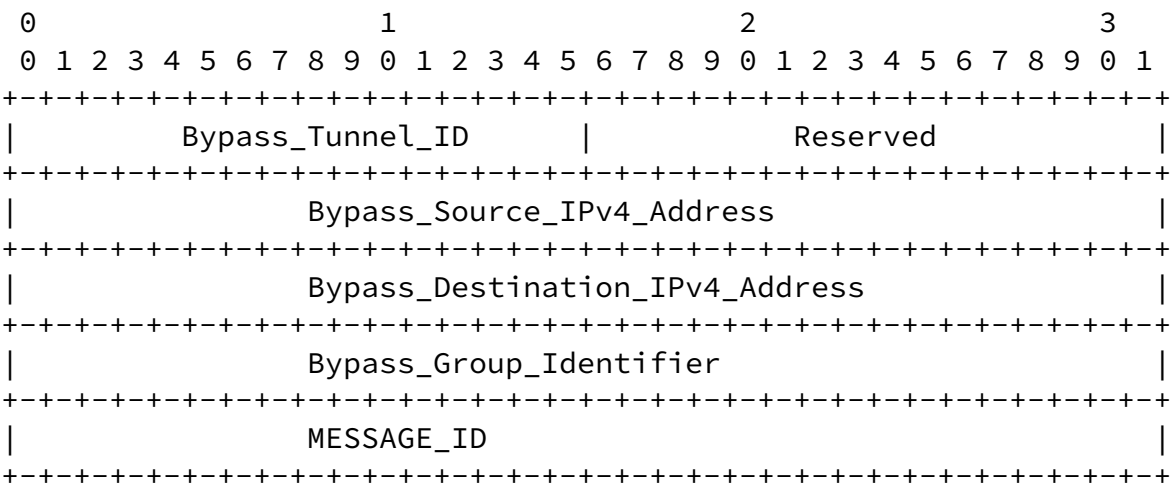

Figure 1: The IPv4 Extended ASSOCIATION ID for B-SFRR-Ready

Bypass_Tunnel_ID: 16 bits

The bypass tunnel identifier.

Reserved: 16 bits

Reserved for future use.

Bypass_Source_IPv4_Address: 32 bits

The bypass tunnel source IPV4 address.

Bypass_Destination_IPv4_Address: 32 bits

The bypass tunnel destination IPV4 address.

Bypass_Group_Identifier: 32 bits

The bypass tunnel group identifier.

MESSAGE_ID

A MESSAGE_ID object as defined by [RFC2961].

3.1.2. IPv6 B-SFRR-Ready IPv6 Extended ASSOCIATION ID

The IPv6 Extended ASSOCIATION ID field for the B-SFRR-Ready association type has the following format:

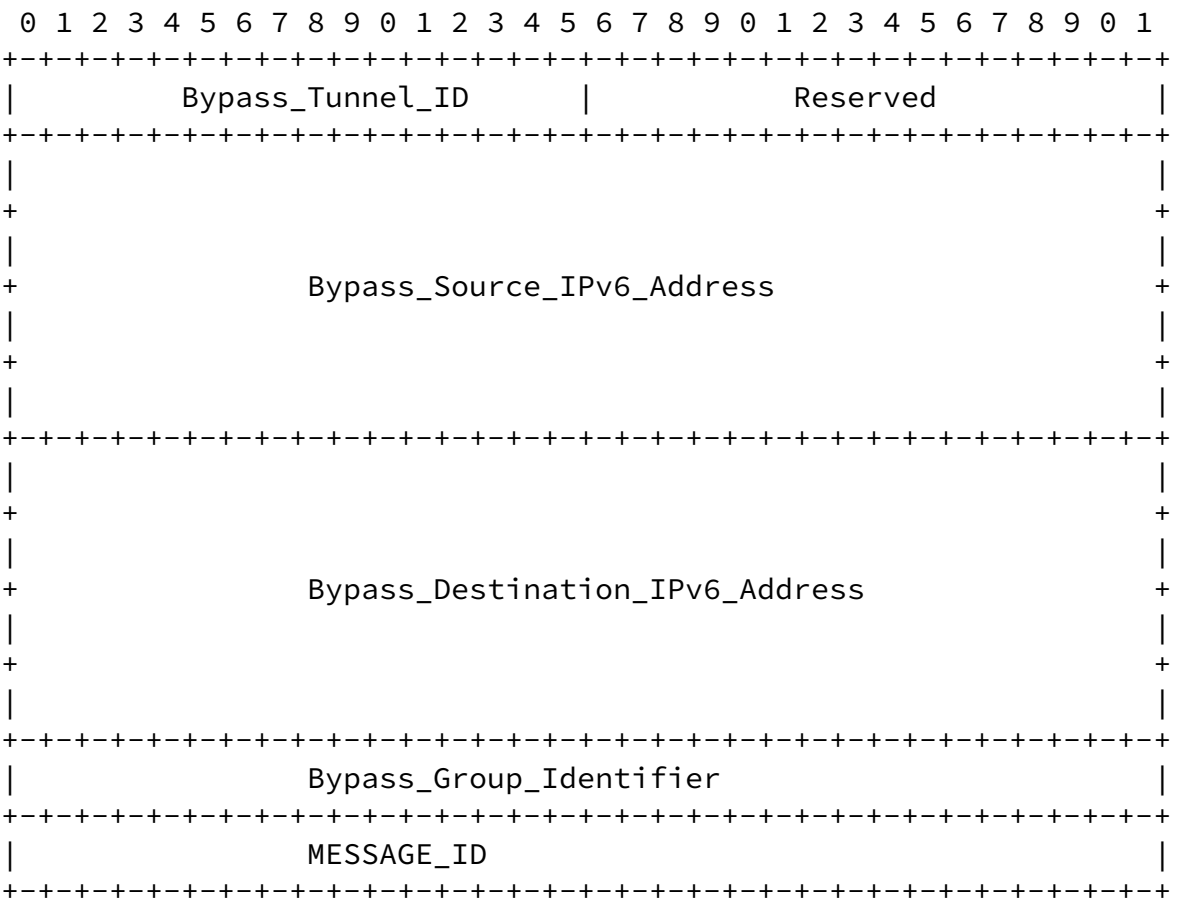

Figure 2: The IPv6 Extended ASSOCIATION ID for B-SFRR-Ready

Bypass_Tunnel_ID: 16 bits

The bypass tunnel identifier.

Reserved: 16 bits

Reserved for future use.

Bypass_Source_IPv6_Address: 128 bits

The bypass tunnel source IPV6 address.

Bypass_Destination_IPv6_Address: 128 bits

The bypass tunnel destination IPV6 address.

Bypass_Group_Identifier: 32 bits

The bypass tunnel group identifier.

MESSAGE_ID

A MESSAGE_ID object as defined by [RFC2961].

The PLR assigns a bypass tunnel and Bypass_Group_Identifier for each protected LSP. The same Bypass_Group_Identifier is used for the set of protected LSPs that share the same bypass tunnel and traverse the same egress link and are not already rerouted. The PLR also generates a MESSAGE_ID object (flags SHOULD be clear, Epoch and Message_Identifier MUST be set according to [RFC2961]).

The PLR MUST generate a new Message_Identifier each time the contents of the B-SFRR-Ready Extended ASSOCIATION ID changes; for example, when PLR node changes the bypass tunnel assignment.

The PLR node notifies the MP node of the bypass tunnel assignment via adding a B-SFRR-Ready Extended ASSOCIATION object and Association ID in the RSVP Path message for the protected LSP using procedures described in Section 3.4.

The MP node acknowledges the PLR node assignment by signaling the B-SFRR-Ready Extended ASSOCIATION object and Association ID within the RSVP Resv message of the protected LSP. With exception of the MESSAGE_ID objects, all other fields of the received in the B-SFRR-Ready Extended ASSOCIATION ID in the RSVP Path message are copied into the B-SFRR-Ready Extended ASSOCIATION ID to be added in the Resv

message. The MESSAGE_ID object is set according to [RFC2961] with

the Flags being clear. A new Message_Identifier MUST be used to acknowledge an updated PLR assignment.

The PLR considers the protected LSP as Summary FRR capable only if all the fields in the B-SFRR-Ready Extended ASSOCIATION ID that are sent in the RSVP Path message and the ones received in the RSVP Resv message (with exception of the MESSAGE_ID) match. If it does not match, or if B-SFRR-Ready Extended ASSOCIATION object is absent in a subsequent refresh, the PLR node MUST consider the protected LSP as not Summary FRR capable.

3.2. B-SFRR-Active Extended ASSOCIATION Object

The Extended ASSOCIATION object for B-SFRR-Active association type is populated by a PLR node to indicate to the MP node (bypass tunnel destination) that one or more groups of protected LSPs that are being protected by the specified bypass tunnel are being rerouted over the bypass tunnel.

The B-SFRR-Active Extended ASSOCIATION object is carried in the RSVP Path message of a bypass LSP and signaled downstream towards the MP (bypass LSP destination).

The Association Type, Association ID, and Association Source MUST be set as defined in [RFC4872] for the ASSOCIATION Object. More specifically:

Association Source:

The Association Source is set to an address of the PLR node.

Association Type:

A new Association Type is defined for B-SFRR-Active as follows:

Value	Type
(TBD-2)	Bypass Summary FRR Active Association (B-SFRR-Active)

Extended ASSOCIATION ID for B-SFRR-Active:

The B-SFRR-Active Extended ASSOCIATION ID is populated by the PLR for the Bypass Summary FRR Active association. The rules to populate the Extended ASSOCIATION ID in this case are described below.

3.2.1. B-SFRR-Active Extended ASSOCIATION ID

The Extended ASSOCIATION ID for the B-SFRR-Active association type has the following format:

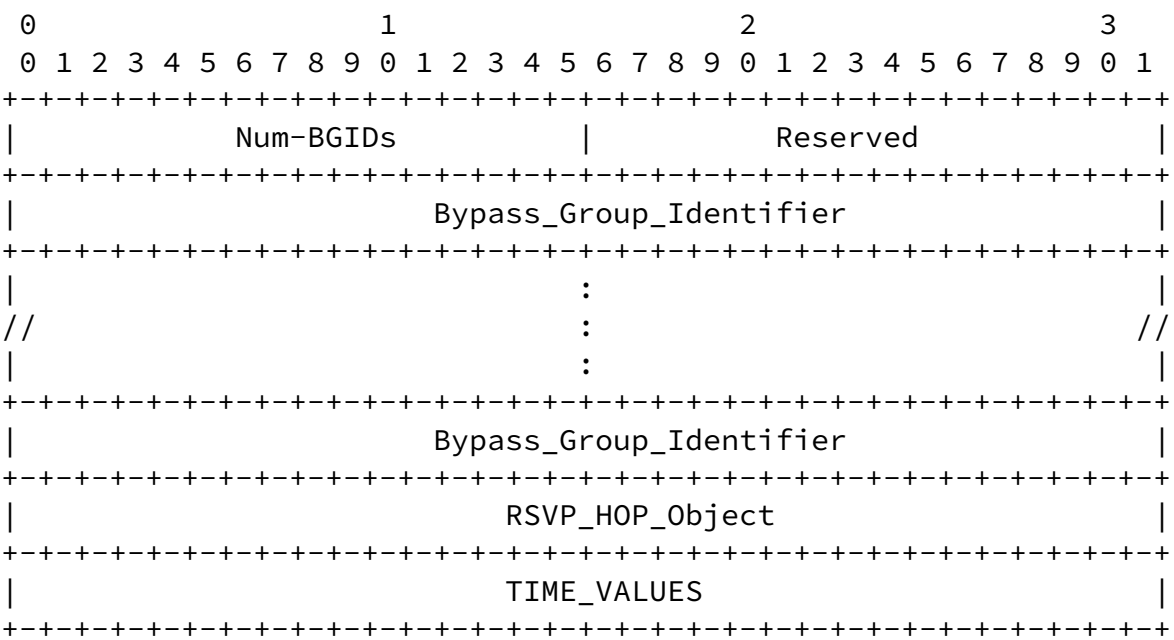

Figure 3: The Extended ASSOCIATION ID for B-SFRR-Active

Num-BGIDs: 16 bits

Number of Bypass_Group_Identifier fields.

Reserved: 16 bits

Reserved for future use.

Bypass_Group_Identifier: 32 bits

The Bypass_Group_Identifier that is previously signaled by the PLR using the Extended Association object. One or more Bypass_Group_Identifiers may be included.

RSVP_HOP_Object: Class 3, as defined by [RFC2205]

Replacement RSVP HOP object to be applied to all LSPs associated with each of the following Bypass_Group_Identifiers. This corresponds to C-Type = 1 for IPv4 RSVP HOP, or C-Type = 2 for IPv6 RSVP HOP depending on the IP address family carried within the object.

TIME_VALUES object: Class 5, as defined by [RFC2205]

Replacement TIME_VALUES object to be applied to all LSPs associated with each of the following Bypass_Group_Identifiers after receiving the B-SFRR-Active Extended ASSOCIATION Object.

3.3. Signaling Procedures Prior to Failure

Before Summary FRR procedures can be used, a handshake MUST be completed between the PLR and MP. This handshake is performed using Extended ASSOCIATION object that carries the B-SFRR-Ready Extended Association ID in both the RSVP Path and Resv messages of the protected LSP.

3.3.1. PLR Signaling Procedure

The B-SFRR-Ready Extended ASSOCIATION object is added by each PLR in the RSVP Path message of the protected LSP to record the bypass tunnel assignment. This object is updated every time the PLR updates the bypass tunnel assignment and that triggers an RSVP Path change message.

Upon receiving an RSVP Resv message with B-SFRR-Ready Extended ASSOCIATION object, the PLR node checks if the expected subobjects from the B-SFRR-Ready ASSOCIATION ID are present. If present, the PLR determines if the MP has acknowledged the current PLR assignment.

To be a valid acknowledgement, the received B-SFRR-Ready ASSOCIATION ID contents within the RSVP Resv message of the protected LSP MUST match the latest B-SFRR-Ready Extended ASSOCIATION object and Association ID contents that the PLR node had sent within the RSVP Path message (with exception of the MESSAGE_ID).

Note, when forwarding an RSVP Resv message upstream, the PLR node SHOULD remove any/all B-SFRR-Ready Extended ASSOCIATION objects whose Association Source matches the PLR node address.

3.3.2. MP Signaling Procedure

Upon receiving an RSVP Path message with a B-SFRR-Ready Extended ASSOCIATION object, the MP node processes all (there may be multiple PLRs for a single MP) B-SFRR-Ready Extended ASSOCIATION objects that have the MP node address as Bypass Destination address in the Association ID.

The MP node first ensures the existence of the bypass tunnel and that the Bypass_Group_Identifier is not already FRR active. That is, an LSP cannot join a group that is already FRR rerouted.

The MP node builds a mirrored Summary FRR Group database per PLR, which is determined using the Bypass_Source_Address field. The MESSAGE_ID is extracted and recorded for the protected LSP PATH state. The MP node signals a B-SFRR-Ready Extended Association object and Association ID in the RSVP Resv message of the protected LSP. With exception of the MESSAGE_ID objects, all other fields of the received B-SFRR-Ready Extended ASSOCIATION object in the RSVP Path message are copied into the B-SFRR-Ready Extended ASSOCIATION object to be added in the Resv message. The MESSAGE_ID object is set according to [RFC2961] with the Flags being clear.

Note, an MP may receive more than one RSVP Path message with the B-SFRR-Ready Extended ASSOCIATION object from different upstream PLR node(s). In this case, the MP node is expected to save all the received MESSAGE_IDs from the different upstream PLR node(s). After a failure, the MP node determines and activates the associated Summary Refresh ID to use once it receives and processes the RSVP Path message containing B-SFRR-Active Extended ASSOCIATION object

that is signaled over the bypass LSP from the PLR, as described Section 3.4

When forwarding an RSVP Path message downstream, the MP SHOULD remove any/all B-SFRR-Ready Extended ASSOCIATION object(s) whose Association ID contains Bypass_Destination_Address matching the MP node address.

3.4. Signaling Procedures Post Failure

Upon detection of the fault (egress link or node failure) the PLR first performs the object modification procedures described by Section 6.4.3 of [RFC4090] for all affected protected LSPs. For Summary FRR LSPs assigned to the same bypass tunnel a common RSVP_HOP and SENDER_TEMPLATE MUST be used.

The PLR MUST signal non-Summary FRR enabled LSPs over the bypass tunnel before signaling the Summary FRR enabled LSPs. This is needed to allow for the case when the PLR node has recently changed a bypass assignment and the MP has not processed the change yet.

The B-SFRR-Active Extended ASSOCIATION object is sent within the RSVP Path message of the bypass LSP to reroute RSVP state of Summary FRR enabled LSPs.

3.4.1. PLR Signaling Procedure

After a failure event, when using the Summary FRR path signaling procedures, an individual RSVP Path message for each Summary FRR LSP is not signaled. Instead, to reroute Summary FRR LSPs via the bypass

tunnel, the PLR adds the B-SFRR-Active Extended Association object in the RSVP Path message of the RSVP session of the bypass tunnel.

The RSVP_HOP_Object field in the B-SFRR-Active Extended ASSOCIATION ID is set to the common RSVP_HOP that was used by the PLR in Section 3.4 of this document.

The previously received MESSAGE_ID from the MP is activated. As a result, the MP may refresh the protected rerouted RESV state using Summary Refresh procedures.

For each affected Summary FRR group, its Bypass_Group_Identifier is added to B-SFRR-Active Extended ASSOCIATION ID.

3.4.2. MP Signaling Procedure

Upon receiving an RSVP Path message with a B-SFRR-Active Extended Association object, the MP performs normal merge point processing for each protected LSP associated with each Bypass_Group_Identifier, as if it received individual RSVP Path messages for the LSP.

For each Summary FRR LSP being merged, the MP first modifies the Path state as follows:

1. The RSVP_HOP object is copied from the B-SFRR-Active Extended ASSOCIATION ID.

2. The TIME_VALUES object is copied from the TIMES_VALUE field in the B-SFRR-Active Extended ASSOCIATION ID. The TIME_VALUES object contains the refresh time of the PLR to generate refreshes and that would have exchanged in a Path message sent to the MP after the failure when no SFRR procedures are in effect.

3. The SENDER_TEMPLATE object SrcAddress field is copied from the bypass tunnel SENDER_TEMPLATE object. For the case where PLR is also the head-end, and SENDER_TEMPLATE SrcAddress of the protected LSP and bypass tunnel are the same, the MP MUST use the modified HOP Address field instead.

4. The ERO object is modified as per Section 6.4.4. of [RFC4090]. Once the above modifications are completed, the MP then performs the merge processing as per [RFC4090].

5. The previously received MESSAGE_ID from the PLR is activated, meaning that the PLR may now refresh the protected rerouted PATH state using Summary Refresh procedures.

A failure during merge processing of any individual rerouted LSP MUST result in an RSVP Path Error message.

An individual RSVP Resv message for each successfully merged Summary

FRR LSP is not signaled. The MP node SHOULD immediately use Summary Refresh procedures to refresh the protected LSP RESV state.

3.5. Refreshing Summary FRR Active LSPs

Refreshing of Summary FRR active LSPs is performed using Summary Refresh as defined by [RFC2961].

4. Compatibility

The (Extended) ASSOCIATION object is defined in [RFC4872] with a class number in the form 11bbbbbb, which ensures compatibility with non-supporting node(s). Such nodes will ignore the object and forward it without modification.

5. Security Considerations

This document updates an existing RSVP object. Thus, in the event of the interception of a signaling message, a slightly more information could be deduced about the state of the network than was previously the case. Existing mechanisms for maintaining the integrity and authenticity of RSVP protocol messages [RFC2747] can be applied. Other considerations mentioned in [RFC4090] and [RFC5920] also apply.

6. IANA Considerations

IANA maintains the "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Parameters" registry (see http://www.iana.org/assignments/gmpls-sig-parameters [1]). The "Association Type" subregistry is included in this registry.

This registry has been updated by new Association Type for Extended ASSOCIATION Object defined in this document as follows:

Value	Name	Reference
TBD-1	B-SFRR-Ready Association	Section 3.1
TBD-2	B-SFRR-Active Association	Section 3.2

IANA also maintains and assigns the values for the RSVP-TE protocol parameters "Resource Reservation Protocol (RSVP) Parameters" (see http://www.iana.org/assignments/rsvp-parameters).

7. Acknowledgments

The authors would like to thank Loa Andersson, Lou Berger, Eric Osborne, Gregory Mirsky, and Mach Chen for reviewing and providing valuable comments to this document.

8. Contributors

Nicholas Tan
Arista Networks

Email: ntan@arista.com

9. References

9.1. 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>.

[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>.

[RFC4872] Lang, J., Ed., Rekhter, Y., Ed., and D. Papadimitriou, Ed., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, DOI 10.17487/RFC4872, May 2007, <https://www.rfc-editor.org/info/rfc4872>.

[RFC6780] Berger, L., Le Faucheur, F., and A. Narayanan, "RSVP ASSOCIATION Object Extensions", RFC 6780, DOI 10.17487/RFC6780, October 2012, <https://www.rfc-editor.org/info/rfc6780>.

9.2. Informative References

[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>.

9.3. URIs

[1] http://www.iana.org/assignments/gmpls-sig-parameters

Authors' Addresses

Mike Taillon
Cisco Systems, Inc.

Email: mtaillon@cisco.com

Tarek Saad (editor)
Juniper Networks

Email: tsaad@juniper.net

Rakesh Gandhi
Cisco Systems, Inc.

Email: rgandhi@cisco.com

Abhishek Deshmukh
Juniper Networks

Email: adeshmukh@juniper.net

Markus Jork
128 Technology

Email: mjork@128technology.com

Vishnu Pavan Beeram
Juniper Networks

Email: vbeeram@juniper.net