CCAMP Working Group   Mike Taillon
Internet-Draft   Tarek Saad
Intended Status: Standards Track   Rakesh Gandhi
Expires: April 24, 2014   Zafar Ali
                                                 (Cisco Systems, Inc)
                                                           Manav Bhatia
                                                      (Alcatel-Lucent)
                                                            Lizhong Jin
                                                                     ()
                                                        Frederic Jounay
                                                            (Orange CH)
                                                       October 21, 2013

Extensions to Resource Reservation Protocol For Fast Reroute of Bidirectional Co-routed Traffic Engineering LSPs
draft-tsaad-ccamp-rsvpte-bidir-lsp-fastreroute-02

Abstract

This document defines RSVP-TE signaling extensions to support Fast Reroute (FRR) of bidirectional co-routed Traffic Engineering (TE) LSPs. These extensions enable the re-direction of bi-directional traffic and signaling onto bypass tunnels that ensure co-routedness of data and signaling paths in the forward and reverse directions after FRR. In addition, the RSVP-TE signaling extensions allow the coordination of bypass tunnel assignment protecting a common facility in both forward and reverse directions prior to or post failure occurrence.

http://www.ietf.org/1id-abstracts.html

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

Copyright and License Notice

Table of Contents

1. Introduction

Co-routed bidirectional tunnels are signaled using GMPLS signaling procedures specified in [RFC3473] and [RFC3471]. Existing procedures defined in [RFC4090] describe the behavior of the Point of Local Repair (PLR) to reroute traffic and signaling onto the bypass tunnel in the event of a failure for unidirectional LSPs. These procedures are applicable to unidirectional protected LSPs, and don’t address issues that arise when employing FRR for bidirectional co-routed Label Switched Paths (LSPs).

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

When co-routed bidirectional bypass tunnels are used to locally protect bidirectional LSPs, the upstream and downstream PLRs may independently assign different bidirectional bypass tunnels in the forward and reverse direction. Currently, there is no means to coordinate the bypass tunnel selection between the downstream and upstream PLRs. In case of mismatch and after FRR, data traffic and signaling may flow over asymmetric paths in the forward and reverse directions which may be undesirable for certain applications.

This document proposes solutions to the above problems by providing corrective actions in the control plane to complement FRR procedures of [RFC4090] in order to maintain the RSVP soft-state for bidirectional protected LSPs and achieve symmetry in the paths followed by data and signaling in the forward and reverse directions post FRR. The document also extends RSVP signaling so it is possible that the bypass tunnel selected by the upstream PLR matches the one selected by the downstream PLR.

2. Terminology

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

The reader is assumed to be familiar with the terminology in [RSVP] and [RSVP-TE].

LSR: Label-Switch Router.

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

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

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

Facility Backup: A local repair method in which a bypass tunnel is used to protect one or more protected LSPs that traverse the PLR, the resource being protected, and the Merge Point in that order.

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

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

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

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

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

CSPF: Constraint-based Shortest Path First.

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

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

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

3. Link Failure With Node-protection Bypass Tunnels

```
                      T1
                 +------<<-----+
                /               \       <-RESV
          [R1]---[R2]----[R3]--x--[R4]---[R5]---[R6]
             ->          \               /
                PATH       +------>>-----+
                                    T2

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

Figure 1: Flow of RSVP signaling post FRR after failure

Consider the Traffic Engineered (TE) network shown in Figure 1. Assume every link in the network is protected with a node- protection bypass tunnel. For the protected bidirectional co-routed LSP whose active/head is on router R1 and passive/tail is on router R6, each traversed router (a potential PLR) independently assigns a node-protection bypass tunnel. Consider a link R3-R4 on the LSP path fails.

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

3.1. Behavior Before Local Repair

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

3.1.1. Downstream Merge Point Label Discovery

For unidirectional primary LSPs, [RFC4090] defines procedures for the downstream PLR to obtain the downstream MP label from recorded labels of the RSVP Resv message received at the downstream PLR.

3.1.2. Upstream Merge Point Label Discovery

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

3.2. Behavior Post Link Failure After FRR

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

3.3. Behavior Post Link Failure To Re-coroute

The downstream Merge Point (MP) R5 that receives rerouted protected LSP RSVP Path message through the bypass tunnel, in addition to the regular MP processing defined in RF4090, gets promoted to a Point of Remote Repair (PRR role) and performs the following actions to re-coroute signaling and data traffic over the same path in both directions:

For unidirectional bypass tunnels:

- Checks for presence of a bypass tunnel in the reverse direction that terminates on the Downstream PLR R3. Note: the Downstream PLR R3’s address is extracted from the "IPV4 tunnel sender address" in the SENDER_TEMPLATE object.

- If present, checks whether the primary LSP traffic and signaling is already rerouted over the found bypass tunnel. If not, PRR R5 activates FRR reroute procedures to direct traffic and signaling (RSVP Resv) over the found bypass tunnel T3 in reverse direction.

- If not present, PRR R5 attempts to auto-provision a bypass tunnel that terminates on the downstream PLR R3. For unidirectional bypass tunnels, if co-routedness in forward and reverse direction is desired, the reverse path bypass tunnel can be inferred from the forward bypass tunnel path (e.g. by reflecting the RRO recorded in the forward direction as ERO for the reverse direction).

- If PRR R5 is unable to successfully provision a bypass tunnel that terminates on the downstream PLR, it may send an immediate RSVP Notify message back to the head-end. The head-end may tear and re-setup the LSP immediately.

For bidirectional bypass tunnels:

- The PRR follows similar procedures described in the solution to second problem in order to identify the bypass tunnel, and reroute traffic and signaling in the reverse path.

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

```
                                         <- RSVP RESV
         [R1]---[R2]----[R3]--X--[R4]---[R5]---[R6]
             RSVP PATH -> \             /
                           +<<------->>+
                           Bypass Tunnel
                         traffic + signaling

              Protected LSP:  {R1-R2-R3-R4-R5-R6}
              R3’s Backup T2: {R3-R5}
              R5’s Backup T3: {R4-R2}
```

Figure 2: Flow of RSVP signaling post FRR after re-coroute

Figure 2 describes the path taken by traffic and signaling after completing re-coroute of data and signaling in the forward and reverse paths described earlier.

The MP MAY optionally support handling in data plane as follows. If the MP is preconfigured with bidirectional bypass tunnel (by DOWNSTREAM_BYPASS_ASSIGNMENT Subobject in Section 4), as soon as the MP node receives the primary tunnel packets on this bypass tunnel, it MAY switch the upstream traffic on to this bypass tunnel. In order to identify the primary tunnel packets through this bypass tunnel, PHP of the bypass tunnel MUST be disabled. The signaling procedure described above in this Section will still apply, and MP checks whether the primary tunnel traffic and signaling is already rerouted over the found bypass tunnel, if not, perform the signaling procedure.

4. Bypass Tunnel Assignment Coordination

This document defines a new subobject in RSVP RECORD_ROUTE object, DOWNSTREAM_BYPASS_ASSIGNMENT, to extend RSVP-TE for fast-reroute signaling. This object is backward compatible with LSRs that do not recognize it (see section 3.10 in [RSVP]).

4.1. DOWNSTREAM_BYPASS_ASSIGNMENT Subobject

The DOWNSTREAM_BYPASS_ASSIGNMENT subobject is used to inform the MP of the backup being used by the PLR. This can be used to coordinate the backup used for the protected LSP by the downstream and upstream PLRs in the forward and reverse direction respectively prior or post the failure occurrence. This subobject MUST only be inserted into the Path message by the downstream PLR and MUST NOT be changed by downstream LSRs. The DOWNSTREAM_BYPASS_ASSIGNMENT subobject has the following format:

The IPv4 DOWNSTREAM_BYPASS_ASSIGNMENT subobject 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Type    |      Length      |    Flags       |  C-Type    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Bypass Tunnel ID          |       Reserved                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                IPv4 Bypass Source Address                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                IPv4 Bypass Destination Address                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The IPv6 DOWNSTREAM_BYPASS_ASSIGNMENT subobject 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      Type     |     Length      |     Flags     |   C-Type    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |       Bypass Tunnel ID         |       Reserved                |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 |                IPv6 Bypass Source Address                     |
 |                                                               |
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 |                IPv6 Bypass Destination Address                |
 |                                                               |
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type

0x04 (TBD) Downstream Bypass Assignment

Length

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

Flags

TBD.

C-Type

The C-Type of the Downstream Bypass Assignment subobject

Bypass Source Address

The bypass tunnel source IPV4 or IPV6 address.

Bypass Destination Address

The bypass tunnel destination IPV4 or IPV6 address.

Bypass Tunnel ID

The bypass tunnel identifier.

4.2. Bypass Tunnel Assignment Signaling Procedure

In cases where bidirectional bypass tunnels are used for FRR Local Repair for a bidirectional co-routed LSP, it is desirable to coordinate the bypass tunnel selected at the downstream and upstream PLRs so that rerouted traffic and signaling flows on symmetrical paths post FRR. To achieve this, a new RSVP subobject is defined for RECORD_ROUTE object (RRO) that identifies a bidirectional bypass tunnel that is assigned at a downstream PLR to protect a bidirectional LSP.

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

The upstream PLR (downstream MP) that detects a DOWNSTREAM_BYPASS_ASSIGNMENT subobject whose bypass tunnel destination matching its own address assigns the matching bidirectional bypass tunnel in the reverse direction, and forwards the RSVP Path message downstream. Otherwise, the bypass tunnel assignment subobject is simply forwarded downstream along in the RSVP Path message.

In absence of DOWNSTREAM_BYPASS_ASSIGNMENT subobject, the downstream MP can independently assign a bypass tunnel in the reverse direction. In the case of downstream MP receiving multiple DOWNSTREAM_BYPASS_ASSIGNMENT subobjects from multiple downstream PLRs, the decision of selecting a bypass tunnel in the reverse direction can be based on local policy, for example, prefer link protection vs. node protection bypass, or prefer the most upstream vs. least upstream node protection bypass tunnel. Note, the bypass tunnel selection will be corrected after FRR based on the PRR behavior after failure.

5. Compatibility

The DOWNSTREAM_BYPASS_ASSIGNMENT subobject to be defined for RSVP RECORD_ROUTE object with class numbers in the form 11bbbbbb, which ensures compatibility with non- supporting nodes. Per [RSVP], nodes not supporting this extension will ignore the subobject but forward it, unexamined and unmodified, in all messages resulting from this message.

6. Security Considerations

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

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

7. IANA Considerations

A new type for the new DOWNSTREAM_BYPASS_ASSIGNMENT subobject for RECORD_ROUTE object is required.

8. Acknowledgements

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

9. References

9.1. Normative References

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

[RSVP-TE] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

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

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

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

9.1. Informative References

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

Authors' Addresses

Mike Taillon
Cisco Systems, Inc.

EMail: mtaillon@cisco.com

Tarek Saad
Cisco Systems, Inc.

EMail: tsaad@cisco.com

Rakesh Gandhi
Cisco Systems, Inc.

EMail: rgandhi@cisco.com

Zafar Ali
Cisco Systems, Inc.

EMail: zali@cisco.com

Manav Bhatia
Alcatel-Lucent
India

Email: manav.bhatia@alcatel-lucent.com

Lizhong Jin
Shanghai, China

Email: lizho.jin@gmail.com

Frederic Jounay
Orange CH

Email: frederic.jounay@orange.ch