MPLS Working Group   A. Atlas
Internet-Draft   R. Torvi
Intended status: Standards Track   M. Jork
Expires: January 10, 2013   Juniper Networks
                                                            July 9, 2012

Ingress Protection for RSVP-TE p2p and p2mp LSPs
draft-torvi-mpls-rsvp-ingress-protection-00

Abstract

Protection against node failure is important for RSVP-TE LSPs, whether point-to-point or point-to-multipoint. While [RFC4090] provides a mechanism for node protection, it does not specify how to protect against failure of the ingress node. This document specifies the RSVP extensions to support ingress node protection and describes the necessary processing behavior.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

It is desirable to protect RSVP-TE LSPs, whether p2p or p2mp, against ingress failure. To do this, a backup node must be pre-identified and prepared with the necessary state so that it can forward traffic when necessary.

Conceptually, a proxy ingress node is created that starts the RSVP signaling. The explicit path of the LSP goes from the proxy ingress node to the backup node and then to the real ingress node. The behavior and signaling for the proxy ingress node is done by the real ingress node.

The backup node must be only one logical hop away from the ingress, whether that be via a direct link or a tunnel.

```
                                 [ traffic source ]
                                  |             |
                                  |             |
                                  |             |
                          [ proxy ingress ]  [ backup ]
                          [ & ingress     ]     |
                                 |              |
                                 |----[ MP ]----|
```

Figure 1: Example Protected LSP with Proxy Ingress Node

There are three different scenarios that this document addresses for ingress protection. All three can be handled using the same set of signaling defined in this document.

A. Traffic Source detects failure The traffic source(s) can rapidly determine that the ingress has failed and switch over to sending traffic to the backup node. When this mode is specified, the backup node will forward any appropriately received traffic along its bypass tunnel to the merge point(s).

B. MP detects failure The traffic source(s) always send traffic to both the ingress and backup nodes. The backup node always forwards traffic along its bypass tunnel to the merge point(s). Each MP determines whether the ingress node has failed and, if so, switches over to accepting the traffic from the backup node.

C. Backup detects failure The traffic source(s) always send traffic to both the ingress and backup nodes. The backup node does not forward the received traffic from the traffic source under normal conditions. When the backup node determines that the ingress node has failed, the backup node starts forwarding the traffic alongs its bypass tunnel(s) to the merge point(s).

For all three scenarios, it is necessary for the backup node to know the merge point(s) and associated MPLS labels. This is accomplished by having the RSVP Path and RESV messages go through the backup node, although the forwarding path need not go through the backup node. There are two cases of interest - on-forwarding-path and off-forwarding-path. In the on-forwarding-path case, the backup node is already the immediate node after the ingress node for the LSP. In the off-forwarding-path, the backup node is not the immediate node after the ingress node for all asociated sub-LSPs.

For ingress protection to be functional, the backup node must have access and knowledge of the appropriate traffic to send into the protected LSP. The ingress node must be capable of describing the traffic to the backup node.

Once the backup node has the necessary state for the LSP, including the set of merge points, the backup node can use bypass tunnels as described in [RFC4090]. If the LSP is a point-to-multipoint, then the backup node has the option of choosing to use a bypass p2mp tunnel for protection.

Finally, an assumption of local protection is that a global repair mechanism will occur to replace the patched LSP with a new fully functional one. To do this, it is necessary that the backup node be able to enact a global repair that still allows sharing of bandwidth resources between the old and new LSPs.

2. Failure Detection Issues and Solution

For each of the different scenarios, the details of detecting an ingress node failure can vary. This document does not specify the details of how to do so, but it is possible using either approriately routed BFD sessions or direct link information.

For traffic-source detection and fail-over, the traffic source can merely monitor the state of the direct links over which traffic is sent to the ingress.

For MP detection, the MP can be configured with the appropriate BFD discriminators used on the BFD sessions. It is desirable for the MP

to know that the ingress can't send traffic to the MP or downstream (for when the ingress is protecting against the MP failure). The appropriate BFD discriminators will vary by MP; they are not signaled in the RSVP extensions described in this draft.

For backup node detection, the backup node can be configured with the appropriate BFD discriminators used on the BFD sessions. Again, they are not signaled in the RSVP extensions described in this draft.

3. Description of Behavior

3.1. Ingress Node

3.1.1. Required Configuration Information

The ingress node must be configured with four pieces of information for these extensions to work.

Backup Node Address The ingress node must know an IP address for the backup node that can be included in the ERO.

Protection Scenario The ingress node must know whether the traffic source, backup node, or merge point(s) will be responsible for handling fail-over.

Ingress-Protector-Context-Id The Ingress-Protector-Context-Id is used for the Extended Session ID in ingress-protected LSPs instead of using the ingress node's loopback address. The Ingress-Protector-Context-Id should not be the same as another address associated with a router that may signal TE LSPs. By having an Ingress-Protector-Context-Id, the backup node can perform global repair.

Application Traffic Identifier The ingress and backup node must both know what application traffic should be directed into the LSP. A commonly understood Application Traffic Identifier is sent between the ingress and backup nodes in RSVP signaling. The exact meaning of the identifier should be configured similarly at both the ingress and backup nodes. The Application Traffic Identifier is understood within the unique context of the Ingress-Protector-Context-Id.

With this additional information, the ingress node can create and signal the necessary RSVP extensions to support ingress protection.

3.1.2. Signaling Behavior

The ingress node is responsible for starting the RSVP signaling for the proxy-ingress node. To do this, the following is done for the RSVP Path message.

1. Compute the EROs for the LSP as normal for the ingress.

2. If the selected backup node is not the first node on the path (for all sub-LSPs), then insert at the beginning of the ERO first the backup node and then the ingress node.

3. Change the IPv4 tunnel sender address in the Sender Template Object to be that of the Ingress-Protector-Context-Id.

4. In the Path RRO, instead of recording the ingress node's address, replace it with the Ingress-Protector-Context-Id.

5. Leave the HOP object populated as usual with information for the ingress-node.

6. Add the INGRESS-PROTECTION object to the Path message. Allocate a second LSP-ID to be used in the INGRESS-PROTECTION object.

7. The RSVP Path message is sent to the backup node as normal. Since the backup node must be only one logical hop away from the ingress, normal RSVP signaling can be used.

When the backup node is off the forwarding path, there are additional behaviors for the ingress node to do when it is handling the associated PATH and RESV messages.

When the ingress node receives an RSVP Path message with an INGRESS-PROTECTION object and the object specifies that node as the ingress node and the PHOP as the backup node, the ingress node SHOULD check the Failure Scenario specified in the INGRESS-PROTECTION object and, if it is not the "MP detects failure" scenario, then the ingress node SHOULD remove the INGRESS-PROTECTION object from the PATH message before sending it out. Additionally, the ingress node must store that it will install ingress forwarding state for the LSP rather than midpoint forwarding.

When an RSVP RESV message is received by the ingress, it uses the NHOP to determine whether the message is received from the backup node or from a different node. The stored associated PATH message contains an INGRESS-PROTECTION object that identifies the backup node. If the RESV message is not from the backup node, then ingress forwarding state should be set up, and the INGRESS-PROTECTION object

MUST be added to the RESV before it is sent to the NHOP, which should be the backup node. If the RESV message is from the backup node, then the LSP should be considered available for use.

If the backup node is on the forwarding path, then a RESV is received with an INGRESS-PROTECTION object and an NHOP that matches the backup node. In this case, the ingress node's address will not appear after the backup node in the RRO. The ingress node should set up ingress forwarding state, just as is done if the LSP weren't ingress-node protected.

3.2. Backup Node

An LER determines that the LSP is ingress-protected based upon the presence of the INGRESS-PROTECTION object in the PATH message. An LER can further determine that it is the backup node if one of its addresses is listed as the backup node in the INGRESS-PROTECTION object.

3.2.1. Behavior for On-Forwarding-Path Backup Node

If the backup node is on the forwarding path, then the backup node MUST remove the INGRESS-PROTECTION object from the PATH message before forwarding it.

If the failure scenario is either "MP-detected" or "Backup-detected", then the backup node is responsible for determining if the ingress node has failed and forwarding the identified traffic from the traffic source(s) to the next-hop(s) on the LSP instead of forwarding the traffic from the ingress node.

When the backup node receives a RESV message, it should add back in the INGRESS-PROTECTION object before forwarding it.

3.2.2. Behavior for Off-Forwarding-Path Backup Node

When the backup node receives a PATH message with the INGRESS-PROTECTION object, the backup node examines the INGRESS-PROTECTION object to learn what traffic associated with the LSP and what failure scenario is being used. The backup node forwards the PATH message to the ingress node with the normal RSVP changes.

When the backup node receives a RESV message with the INGRESS-PROTECTION object, the backup node records an IMPLICIT-NULL label in the RRO. The backup node creates the appropriate forwarding state for the failure scenario specified. For the "MP-detected" and "traffic-source-detected", this means that backup node forwards any received identified traffic into the bypass tunnel(s) to the merge

point(s). For the "backup-detected", this means that the backup node creates state to quickly determine the ingress has failed and switch to sending any received identified traffic into the bypass tunnel(s) to the merge point(s). Then the backup node forwards the RESV message to the ingress node, which is acting for the proxy ingress.

If the backup node doesn't have a bypass tunnel to a merge point, then the backup node can wait to send the RESV until such has been created or it can send a Path Err with an Error Code of "Routing Problem (24)" and a new Error Value sub-code of "No Bypass Tunnel to Merge Point (TBD)".

3.3. Merge Node

An LSR that is serving as a Merge Node may need to support the INGRESS-PROTECTION object and functionality defined in this specification if the LSP is ingress-protected where the failure scenario is "MP-detected". An LSR can determine that it must be a merge point by examining the INGRESS-PROTECTION object and determining that it is neither the ingress node nor the backup node and the PHOP is the ingress node.

In that case, when the LSR receives a PATH message with an INGRESS-PROTECTION object, the LSR MUST remove the INGRESS-PROTECTION object before forwarding on the PATH message.

If the failure scenario specified is "MP-detected", the MP must connect up the fast-failure detection (as configured) to accepting backup traffic received from the backup node. There are a number of different ways that the MP can enforce not forwarding traffic normally received from the backup node. For instance, first, any LSPs set up from the backup node should not be signaled with an IMPLICIT NULL label and second, the associated label for the ingress-protected LSP could be set to normally discard inside that context.

When the MP receives a RESV message whose matching PATH state had an INGRESS-PROTECTION object, the MP SHOULD add the INGRESS-PROTECTION object to the RESV message before forwarding it.

3.4. Global Repair

When the backup node learns the ingress node has failed (e.g. via the IGP), then the backup node can compute new ERO(s) and signal the new LSP so that it no longer relies upon local repair. To do this, the backup node uses the same Ingress-Protector-Context-Id as the Ipv4 tunnel sender address in the Sender Template Object and uses the previously allocated second LSP-ID signaled in the INGRESS-PROTECTION object. This allows the new LSP to share resources with the old LSP.

3.5. Ingress Revival and Administrative Switching

In a future version, it is intended to describe the behavior when the ingress node comes back and how to handle management-triggered switches from ingress to backup node and vice versa.

4. RSVP Extensions

4.1. INGRESS-PROTECTION object

Class-Num = TBD
C-Type = TBD

```
       0             1              2             3
+-------------+-------------+-------------+-------------+
|       Length (bytes)      |  Class-Num  |   C-Type    |
+-------------+-------------+-------------+-------------+
|                Backup Node Address                    |
+-------------+-------------+-------------+-------------+
|                Ingress Node Address                   |
+-------------+-------------+-------------+-------------+
|           Application Traffic Identifier              |
+-------------+-------------+-------------+-------------+
|   Secondary LSP ID        | Protection  |  Flags      |
|                           | Scenario    |             |
+-------------+-------------+-------------+-------------+
```

Figure 2: INGRESS-PROTECTION object

Backup Node Address

Ingress Node Address

Application Traffic Identifier

Ingress-Protector-Context-Id

Secondary LSP ID

Protection Scenario Indicates if (1) traffic source(s), (2) backup node, (3) or merge point(s) will handle the fail-over.

Control Flags Backup sent flags: (0x01)Ingress-Protection in-use, (0x02)Ingress Detected Down, (0x04)Admin Override caused Ingress-Protection-in-use. Ingress sent flags: (0x08)Revert Control to Ingress, (0x10)Force control to Backup

5. Normative References

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

[RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

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

[RFC4875] Aggarwal, R., Papadimitriou, D., and S. Yasukawa, "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, May 2007.

Authors' Addresses

Alia Atlas
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Email: akatlas@juniper.net

Raveendra Torvi
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Email: rtorvi@juniper.net

Markus Jork
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Email: mjork@juniper.net