TEAS Working Group X. Zhang
Internet-Draft H. Zheng, Ed.
Intended Status: Informational Huawei Technologies
Expires: July 30, 2017 R. Gandhi, Ed.
Z. Ali
Cisco Systems, Inc.
P. Brzozowski
ADVA Optical
January 26, 2017

RSVP-TE Signaling Procedure for End-to-End GMPLS Restoration and Resource Sharing
draft-ietf-teas-gmpls-resource-sharing-proc-08

Abstract

In non-packet transport networks, there are requirements where Generalized Multi-Protocol Label Switching (GMPLS) end-to-end recovery scheme needs to employ restoration Label Switched Path (LSP) while keeping resources for the working and/or protecting LSPs reserved in the network after the failure occurs.

This document reviews how the LSP association is to be provided using Resource Reservation Protocol - Traffic Engineering (RSVP-TE) signaling in the context of GMPLS end-to-end recovery scheme when using restoration LSP where failed LSP is not torn down. In addition, this document discusses resource sharing-based setup and teardown of LSPs as well as LSP reversion procedures. No new signaling extensions are defined by this document, and it is strictly informative in nature.

Table of Contents

1. Introduction

Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] defines a set of protocols, including Open Shortest Path First - Traffic Engineering (OSPF-TE) [RFC4203] and Resource ReserVation Protocol - Traffic Engineering (RSVP-TE) [RFC3473]. These protocols can be used to set up Label Switched Paths (LSPs) in non-packet transport networks. The GMPLS protocol extends MPLS to support interfaces capable of Time Division Multiplexing (TDM), Lambda Switching and Fiber Switching. These switching technologies provide several protection schemes [RFC4426][RFC4427] (e.g., 1+1, 1:N and M:N).

Resource Reservation Protocol - Traffic Engineering (RSVP-TE) signaling has been extended to support various GMPLS recovery schemes, such as end-to-end recovery [RFC4872] and segment recovery [RFC4873]. As described in [RFC6689], an ASSOCIATION object with Association Type "Recovery" [RFC4872] can be signaled in the RSVP Path message to identify the LSPs for restoration. Also, an ASSOCIATION object with Association Type "Resource Sharing" [RFC4873] can be signaled in the RSVP Path message to identify the LSPs for resource sharing. [RFC6689] Section 2.2 reviews the procedure for providing LSP associations for GMPLS end-to-end recovery and Section 2.4 reviews the procedure for providing LSP associations for sharing resources.

Generally GMPLS end-to-end recovery schemes have the restoration LSP set up after the failure has been detected and notified on the working LSP. For recovery scheme with revertive behavior, a restoration LSP is set up while working LSP and/or protecting LSP are not torn down in control plane due to a failure. In non-packet transport networks, as working LSPs are typically set up over preferred paths, service providers would like to keep resources associated with the working LSPs reserved. This is to make sure that the service can be reverted to the preferred path (working LSP) when the failure is repaired to provide deterministic behavior and guaranteed Service Level Agreement (SLA).

In this document, we review procedures for GMPLS LSP associations, resource sharing based LSP setup, teardown, and LSP reversion for non-packet transport networks, including the following:

- o Review the procedure for providing LSP associations for the GMPLS end-to-end recovery using restoration LSP where working and protecting LSPs are not torn down and resources are kept reserved in the network after the failure.

- o In [RFC3209], the make-before-break (MBB) method assumes the old and new LSPs share the SESSION object and signal Shared Explicit

(SE) flag in SESSION_ATTRIBUTE object for sharing resources. According to [RFC6689], an ASSOCIATION object with Association Type "Resource Sharing" in the Path message enables the sharing of resources across LSPs with different SESSION objects. The procedure for resource sharing using the SE flag in conjunction with an ASSOCIATION object is discussed in this document.

- o When using end-to-end recovery scheme with revertive behavior, methods for LSP reversion and resource sharing are summarized in this document.

This document is strictly informative in nature and does not define any RSVP-TE signaling extensions.

2. Conventions Used in This Document

2.1. Terminology

The reader is assumed to be familiar with the terminology in [RFC3209], [RFC3473], [RFC4872] and [RFC4873]. The terminology for GMPLS recovery is defined in [RFC4427].

2.2. Acronyms and Abbreviations

GMPLS: Generalized Multi-Protocol Label Switching

LSP: An MPLS Label Switched Path

MBB: Make Before Break

MPLS: Multi-Protocol Label Switching

RSVP: Resource ReSerVation Protocol

SE: Shared Explicit flag

TDM: Time Division Multiplexing

TE: Traffic Engineering

3. Overview

The GMPLS end-to-end recovery scheme, as defined in [RFC4872] and being considered in this document, switches normal traffic to an alternate LSP that is not even partially established only after the

working LSP failure occurs. The new alternate route is selected at the LSP head-end node, it may reuse resources of the failed LSP at intermediate nodes and may include additional intermediate nodes and/or links.

3.1. Examples of Restoration Schemes

Two forms of end-to-end recovery schemes, 1+R restoration and 1+1+R restoration are described in the following sections. Other forms of end-to-end recovery schemes also exist and they can use these signaling techniques.

3.1.1. 1+R Restoration

One example of the recovery scheme considered in this document is 1+R recovery. The 1+R recovery scheme is exemplified in Figure 1. In this example, a working LSP on path A-B-C-Z is pre-established. Typically after a failure detection and notification on the working LSP, a second LSP on path A-H-I-J-Z is established as a restoration LSP. Unlike a protecting LSP which is set up before the failure, a restoration LSP is set up when needed, after the failure.

```
+-----+    +-----+     +-----+     +-----+
|  A  +----+  B  +-----+  C  +-----+  Z  |
+--+--+    +-----+     +-----+     +--+--+
    \                                /
     \                              /
   +--+--+      +-----+        +--+--+
   |  H  +-------+  I  +--------+  J  |
   +-----+      +-----+        +-----+
```

Figure 1: An Example of 1+R Recovery Scheme

During failure switchover with 1+R recovery scheme, in general, working LSP resources are not released so that working and restoration LSPs coexist in the network. Nonetheless, working and restoration LSPs can share network resources. Typically when the failure has recovered on the working LSP, the restoration LSP is no longer required and is torn down while the traffic is reverted to the original working LSP.

3.1.2. 1+1+R Restoration

Another example of the recovery scheme considered in this document is 1+1+R. In 1+1+R, a restoration LSP is set up for the working LSP and/or the protecting LSP after the failure has been detected, and

this recovery scheme is exemplified in Figure 2.

```
             +-----+       +-----+        +-----+
             |  D  +-------+  E  +--------+  F  |
             +--+--+       +-----+        +--+--+
               /                              \
              /                                \
          +--+--+    +-----+     +-----+     +--+--+
          |  A  +----+  B  +-----+  C  +-----+  Z  |
          +--+--+    +-----+     +-----+     +--+--+
              \                                /
               \                              /
             +--+--+       +-----+        +--+--+
             |  H  +-------+  I  +--------+  J  |
             +-----+       +-----+        +-----+
```

Figure 2: An Example of 1+1+R Recovery Scheme

In this example, a working LSP on path A-B-C-Z and a protecting LSP on path A-D-E-F-Z are pre-established. After a failure detection and notification on the working LSP or protecting LSP, a third LSP on path A-H-I-J-Z is established as a restoration LSP. The restoration LSP in this case provides protection against failure of both the working and protecting LSPs. During failure switchover with 1+1+R recovery scheme, in general, failed LSP resources are not released so that working, protecting and restoration LSPs coexist in the network. The restoration LSP can share network resources with the working LSP, and it can share network resources with the protecting LSP. Typically, the restoration LSP is torn down when the traffic is reverted to the original LSP and it is no longer needed.

There are two possible models when using a restoration LSP with 1+1+R recovery scheme:

o A restoration LSP is set up after either a working or protecting LSP fails. Only one restoration LSP is present at a time.

o A restoration LSP is set up after both working and protecting LSPs fail. Only one restoration LSP is present at a time.

3.1.2.1. 1+1+R Restoration - Variants

Two other possible variants exist when using a restoration LSP with 1+1+R recovery scheme:

- o A restoration LSP is set up after either a working or protecting LSP fails. Two different restoration LSPs may be present, one for the working LSP and one for the protecting LSP.

- o Two different restoration LSPs are set up after both working and protecting LSPs fail, one for the working LSP and one for the protecting LSP.

In all these models, if a restoration LSP also fails, it is torn down and a new restoration LSP is set up.

3.2. Resource Sharing by Restoration LSP

```
                           +-----+      +-----+
                           |  F  +------+  G  +---------+
                           +--+--+      +-----+         |
                              |                         |
                              |                         |
       +-----+    +-----+     +--+--+      +-----+     +--+--+
       |  A  +----+  B  +-----+  C  +--X---+  D  +-----+  E  |
       +-----+    +-----+     +-----+      +-----+     +-----+
```

Figure 3: Resource Sharing in 1+R Recovery Scheme

Using the network shown in Figure 3 as an example using 1+R recovery scheme, LSP1 (A-B-C-D-E) is the working LSP, and assume it allows for resource sharing when the LSP traffic is dynamically restored. Upon detecting the failure of a link along the LSP1, e.g. Link C-D, node A needs to decide which alternative path it will use to signal restoration LSP and reroute traffic. In this case, A-B-C-F-G-E is chosen as the restoration LSP path and the resources on the path segment A-B-C are re-used by this LSP. The working LSP is not torn down and co-exists with the restoration LSP. When the head-end node A signals the restoration LSP, nodes C, F, G and E reconfigure the resources (as listed in Table 1 of this document) to set up the LSP by sending cross-connection command to the data plane.

In the recovery scheme employing revertive behavior, after the failure is repaired, the resources on nodes C and E need to be reconfigured to set up the working LSP (using a procedure described in Section 4.3 of this document) by sending cross-connection command to the data plane. The traffic is then reverted back to the original working LSP.

4. RSVP-TE Signaling Procedure

4.1. Restoration LSP Association

Where GMPLS end-to-end recovery scheme needs to employ a restoration LSP while keeping resources for the working and/or protecting LSPs reserved in the network after the failure, the restoration LSP is set up with an ASSOCIATION object that has Association Type set to "Recovery" [RFC4872], the Association ID and the Association Source set to the corresponding Association ID and the Association Source signaled in the Path message of the LSP it is restoring. For example, when a restoration LSP is signaled for a failed working LSP, the ASSOCIATION object in the Path message of the restoration LSP contains the Association ID and Association Source set to the Association ID and Association Source signaled in the working LSP for the "Recovery" Association Type. Similarly, when a restoration LSP is set up for a failed protecting LSP, the ASSOCIATION object in the Path message of the restoration LSP contains the Association ID and Association Source set to the Association ID and Association Source signaled in the protecting LSP for the "Recovery" Association Type.

The procedure for signaling the PROTECTION object is specified in [RFC4872]. Specifically, the restoration LSP used for a working LSP is set up with P bit cleared in the PROTECTION object in the Path message of the restoration LSP and the restoration LSP used for a protecting LSP is set up with P bit set in the PROTECTION object in the Path message of the restoration LSP.

4.2. Resource Sharing-based Restoration LSP Setup

GMPLS LSPs can share resources during LSP setup if they have Shared Explicit (SE) flag set in the SESSION_ATTRIBUTE objects [RFC3209] in the Path messages that create them and:

o As defined in [RFC3209], LSPs have identical SESSION objects and/or

o As defined in [RFC6689], LSPs have matching ASSOCIATION object with Association Type set to "Resource Sharing" signaled in their Path messages. LSPs in this case can have different SESSION objects i.e. different Tunnel ID, Source and/or Destination signaled in their Path messages.

As described in [RFC3209], Section 2.5, the purpose of make-before-break is not to disrupt traffic, or adversely impact network operations while TE tunnel rerouting is in progress. In non-packet transport networks during the RSVP-TE signaling procedure, the nodes set up cross-connections along the LSP accordingly. Because the

cross-connection cannot simultaneously connect a shared resource to different resources in two alternative LSPs, nodes may not be able to fulfill this request when LSPs share resources.

For LSP restoration upon failure, as explained in Section 11 of [RFC4872], the reroute procedure may re-use existing resources. The action of the intermediate nodes during the rerouting process to reconfigure cross-connections does not further impact the traffic since it has been interrupted due to the already failed LSP.

The node actions for setting up the restoration LSP can be categorized into the following:

Category	Action
Reusing existing resource on both input and output interfaces (nodes A & B in Figure 3).	This type of node needs to reserve the existing resources and no cross-connection command is needed.
Reusing existing resource only on one of the interfaces, either input or output interfaces and using new resource on the other interfaces. (nodes C & E in Figure 3).	This type of node needs to reserve the resources and send the re-configuration cross-connection command to its corresponding data plane node on the interfaces where new resources are needed and it needs to re-use the existing resources on the other interfaces.
Using new resources on both interfaces. (nodes F & G in Figure 3).	This type of node needs to reserve the new resources and send the cross-connection command on both interfaces.

Table 1: Node Actions During Restoration LSP Setup

Depending on whether the resource is re-used or not, the node actions differ. This deviates from normal LSP setup since some nodes do not need to re-configure the cross-connection. Also, the judgment whether the control plane node needs to send a cross-connection setup or modification command to its corresponding data plane node(s) relies on the check whether the LSPs are sharing resources.

4.3. LSP Reversion

If the end-to-end LSP recovery scheme employs the revertive behavior, as described in Section 3 of this document, traffic can be reverted from the restoration LSP to the working or protecting LSP after its failure is recovered. The LSP reversion can be achieved using two methods:

1. Make-while-break Reversion, where resources associated with a working or protecting LSP are reconfigured while removing reservations for the restoration LSP.

2. Make-before-break Reversion, where resources associated with a working or protecting LSP are reconfigured before removing reservations for the restoration LSP.

In non-packet transport networks, both of the above reversion methods will result in some traffic disruption when the restoration LSP and the LSP being restored are sharing resources and the cross-connections need to be reconfigured on intermediate nodes.

4.3.1. Make-while-break Reversion

In this reversion method, restoration LSP is simply requested to be deleted by the head-end. Removing reservations for restoration LSP triggers reconfiguration of resources associated with a working or protecting LSP on every node where resources are shared. The working or protecting LSP state was not removed from the nodes when the failure occurred. Whenever reservation for restoration LSP is removed from a node, data plane configuration changes to reflect reservations of working or protecting LSP as signaling progresses. Eventually, after the whole restoration LSP is deleted, data plane configuration will fully match working or protecting LSP reservations on the whole path. Thus reversion is complete.

Make-while-break, while being relatively simple in its logic, has a few limitations as follows which may not be acceptable in some networks:

o No rollback

If for some reason reconfiguration of data plane on one of the nodes to match working or protecting LSP reservations fails, falling back to restoration LSP is no longer an option, as its state might have already been removed from other nodes.

o No completion guarantee

Deletion of an LSP provides no guarantees of completion. In particular, if RSVP packets are lost due to a node or link failure it is possible for an LSP to be only partially deleted. To mitigate this, RSVP could maintain soft state reservations and hence eventually remove remaining reservations due to refresh timeouts. This approach is not feasible in non-packet transport networks however, where control and data channels are often separated and hence soft state reservations are not useful.

Finally, one could argue that graceful LSP deletion [RFC3473] would provide guarantee of completion. While this is true for most cases, many implementations will time out graceful deletion if LSP is not removed within certain amount of time, e.g. due to a transit node fault. After that, deletion procedures which provide no completion guarantees will be attempted. Hence, in corner cases a completion guarantee cannot be provided.

o No explicit notification of completion to head-end node

In some cases, it may be useful for a head-end node to know when the data plane has been reconfigured to match working or protecting LSP reservations. This knowledge could be used for initiating operations like enabling alarm monitoring, power equalization and others. Unfortunately, for the reasons mentioned above, make-while-break reversion lacks such explicit notification.

4.3.2. Make-before-break Reversion

This reversion method can be used to overcome limitations of make-while-break reversion. It is similar in spirit to MBB concept used for re-optimization. Instead of relying on deletion of the restoration LSP, the head-end chooses to establish a new reversion LSP that duplicates the configuration of the resources on the working or protecting LSP, and uses identical ASSOCIATION and PROTECTION objects in the Path message of that LSP. Only if setup of this LSP is successful will other (restoration and working or protecting) LSPs be deleted by the head-end. MBB reversion consists of two parts:

A) Make part:

Creating a new reversion LSP following working or protecting LSP's path. The reversion LSP shares all of the resources of the working or protecting LSP and may share resources with the restoration LSP. As reversion LSP is created, resources are reconfigured to match its reservations. Hence, after reversion LSP is created, data plane configuration reflects working or protecting LSP reservations.

B) Break part:

After "make" part is finished, the original working or protecting and restoration LSPs are torn down, and the reversion LSP becomes the new working or protecting LSP. Removing reservations for working or restoration LSPs does not cause any resource reconfiguration on reversion LSP's path - nodes follow same procedures as for "break" part of any MBB operation. Hence, after working or protecting and restoration LSPs are removed, data plane configuration is exactly the same as before starting restoration. Thus, reversion is complete.

MBB reversion uses make-before-break characteristics to overcome challenges related to make-while-break reversion as follow:

o Rollback

If "make" part fails, (existing) restoration LSP will still be used to carry existing traffic as the restoration LSP state was not removed. Same logic applies here as for any MBB operation failure.

o Completion guarantee

LSP setup is resilient against RSVP message loss, as Path and Resv messages are refreshed periodically. Hence, given that network recovers from node and link failures eventually, reversion LSP setup is guaranteed to finish with either success or failure.

o Explicit notification of completion to head-end node

Head-end knows that data plane has been reconfigured to match working or protecting LSP reservations on intermediate nodes when it receives Resv for the reversion LSP.

5. Security Considerations

This document reviews procedures defined in [RFC3209] [RFC4872] [RFC4873] and [RFC6689] and does not define any new procedure. This document does not introduce any new security issues other than those already covered in [RFC3209] [RFC4872] [RFC4873] and [RFC6689].

6. IANA Considerations

This informational document does not make any request for IANA action.

7. References

7.1. Normative References

[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.

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

[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, May 2007.

[RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.

[RFC6689] L. Berger, "Usage of the RSVP ASSOCIATION Object", RFC 6689, July 2012.

7.2. Informative References

[RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.

[RFC4203] Kompella, K., and Rekhter, Y., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

[RFC4426] Lang, J., Rajagopalan, B., and Papadimitriou, D., "Generalized Multiprotocol Label Switching (GMPLS) Recovery Functional Specification", RFC 4426, March 2006.

[RFC4427] Mannie, E., and Papadimitriou, D., "Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching", RFC 4427, March 2006.

Acknowledgements

The authors would like to thank George Swallow for the discussions on the GMPLS restoration. The authors would like to thank Lou Berger for the guidance on this work. The authors would also like to thank Lou Berger, Vishnu Pavan Beeram and Christian Hopps for reviewing this document and providing valuable comments. A special thanks to Dale Worley for his thorough review of this document.

Contributors

Gabriele Maria Galimberti
Cisco Systems, Inc.

EMail: ggalimbe@cisco.com

Authors' Addresses

Xian Zhang
Huawei Technologies
F3-1-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China

EMail: zhang.xian@huawei.com

Haomian Zheng (editor)
Huawei Technologies
F3-1-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China

EMail: zhenghaomian@huawei.com

Rakesh Gandhi (editor)
Cisco Systems, Inc.

EMail: rgandhi@cisco.com

Zafar Ali
Cisco Systems, Inc.

EMail: zali@cisco.com

Pawel Brzozowski
ADVA Optical

EMail: PBrzozowski@advaoptical.com