     **RSVP-TE Signaling Procedure for End-to-End GMPLS Restoration and
                            Resource Sharing**
            draft-ietf-teas-gmpls-resource-sharing-proc-03

Abstract

   In 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 for transport
   networks. No new signaling extensions are defined by this document,
   and it is strictly informative in nature.

Status of this Memo

   material or to cite them other than as "work in progress."

   The list of current Internet-Drafts can be accessed at
   http://www.ietf.org/ietf/1id-abstracts.txt.

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

Copyright Notice

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 setup Label Switched Paths (LSPs) in 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], ASSOCIATION object can be used
to identify the LSPs for restoration using Association Type set to
"Recovery" [RFC4872] and also identify the LSPs for resource sharing
using Association Type set to "Resource Sharing" [RFC4873].
[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.

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

In this document, procedures for transport networks are reviewed for
LSP associations, resource sharing based LSP setup, teardown and LSP
reversion, including 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], ASSOCIATION object with Association Type "Resource
Sharing" enables the sharing of resources across LSPs with different
SESSION objects. Procedure for resource sharing using the SE flag in
conjunction with ASSOCIATION object is discussed in this document.

o When using end-to-end recovery with revertive mode, 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. Overview

The GMPLS end-to-end recovery scheme, as defined in [RFC4872] and
being considered in this document, "fully dynamic rerouting 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". Two examples, 1+R and
1+1+R are described in the following sections.

2.1. 1+R Restoration

One example of the recovery scheme considered in this document is 1+R
recovery. The 1+R recovery is exemplified in Figure 1. In this
example, 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 protection LSP, restoration LSP is signaled per need basis.

```
    +-----+     +-----+     +-----+     +-----+
    |  A  +----+  B  +-----+  C  +-----+  Z  |
    +--+--+     +-----+     +-----+     +--+--+
        \                                  /
         \                                /
     +--+--+        +-----+        +--+--+
     |  H  +-------+  I  +--------+  J  |
     +-----+        +-----+        +-----+
```
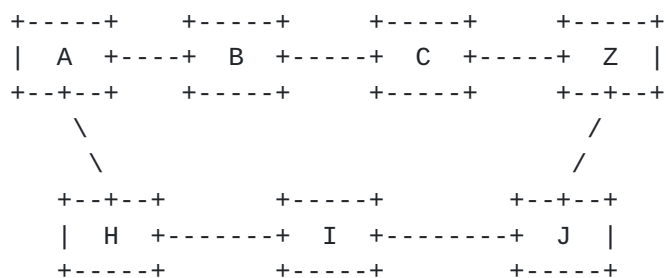
Figure 1: An Example of 1+R Recovery Scheme

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 failure
is recovered on the working LSP, restoration LSP is no longer
required and torn down, while the traffic is reverted to the original
working LSP.

2.2. 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 signaled for the working LSP
and/or the protecting LSP after the failure has been detected, and
this recovery is exemplified in Figure 2.

```
        +-----+         +-----+          +-----+
        |  D  +-------+  E  +--------+  F  |
        +--+--+         +-----+          +--+--+
          /                                  \
         /                                    \
     +--+--+     +-----+      +-----+      +--+--+
     |  A  +----+  B  +-----+  C  +-----+  Z  |
     +--+--+     +-----+      +-----+      +--+--+
        \                                  /
         \                                /
       +--+--+         +-----+          +--+--+
       |  H  +-------+  I  +--------+  J  |
       +-----+         +-----+          +-----+
```
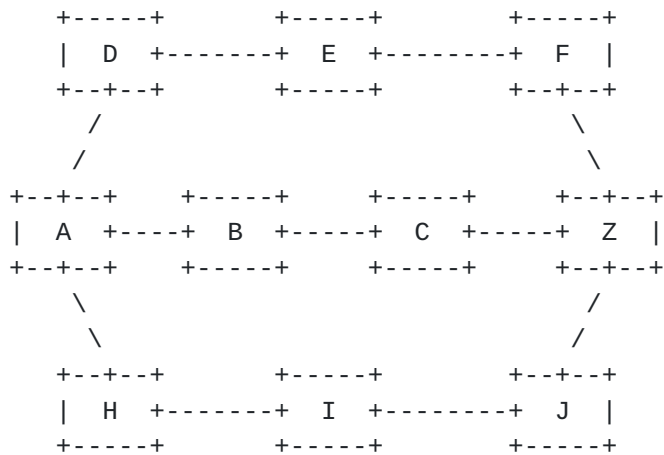
Figure 2: An Example of 1+1+R Recovery Scheme

In this example, working LSP on path A-B-C-Z and protecting LSP on
path A-D-E-F-Z are pre-established. After a failure detection and
notification on a 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 a second order failure.
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. Nonetheless, restoration
LSP with working LSP it is restoring as well as restoration LSP with
protecting LSP it is restoring can share network resources.
Typically, restoration LSP is torn down when the failure on the
original (working or protecting) LSP is repaired and the traffic is
reverted to the original LSP.

There are four possible models when using restoration LSP with 1+1+R

recovery scheme:

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

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

o A restoration LSP is signaled after both working and protecting
LSPs fail. Only one restoration LSP is present.

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

In all models discussed, if the restoration LSP also fails, it is
torn down and a new restoration LSP is signaled.

2.3. Resource Sharing By Restoration LSP

```
                        +-----+       +-----+
                        |  F  +------+  G  +--------+
                        +--+--+       +-----+       |
                           |                        |
                           |                        |
     +-----+     +-----+   +--+--+       +-----+     +--+--+
     |  A  +----+  B  +-----+  C  +--X---+  D  +-----+  E  |
     +-----+     +-----+     +-----+       +-----+     +-----+
```
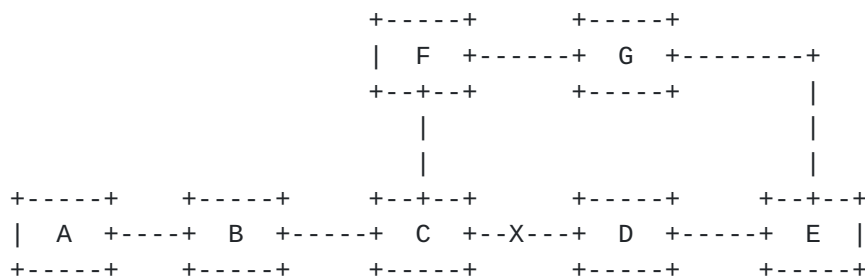
Figure 3: Resource Sharing in 1+R Recovery Scheme

Using the network shown in Figure 3 as an example, LSP1 (A-B-C-D-E)
is the working LSP and it allows for resource sharing when the LSP
traffic is dynamically restored after the link failure. 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 when the working LSP is not
torn down (e.g. in 1+R recovery scheme).

3. **RSVP-TE Signaling Procedure**

3.1. **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
   signaled 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 LSP it is restoring. For example, when a restoration
   LSP is signaled for a failed working LSP, the ASSOCIATION object in
   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 signaled for a failed protecting LSP, the
   ASSOCIATION object in 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 signaled with P bit cleared in the PROTECTION object and the
   restoration LSP used for a protecting LSP is signaled with P bit set
   in the PROTECTION object.

3.2. **Resource Sharing-based Restoration LSP Setup**

   GMPLS LSPs can share resources during LSP setup if they have Shared
   Explicit (SE) flag set in their SESSION_ATTRIBUTE objects 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". LSPs in this case
   can have different SESSION objects i.e. different Tunnel ID, Source
   and/or Destination.

   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 transport
   networks, the label has a mapping into the data plane resource used
   and the nodes along the LSP need to send triggering commands to data
   plane for setting up cross-connections accordingly during the RSVP-TE
   signaling procedure. Due to the nature of the transport networks,
   node may not be able to fulfill this purpose when sharing resources

in some scenarios.

For LSP restoration upon failure, as explained in Section 11 of
[RFC4872], reroute procedure may re-use existing resources. The
behavior of the intermediate nodes during 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 behavior for setting up the restoration LSP can be
categorized into the following three categories:

        Table 1: Node Behavior during Restoration LSP Setup

    ---------+---
    Category | Node Behavior during Restoration LSP Setup
    ---------+---
       C1   + Reusing existing resource on both input and output
             + interfaces (node A & B in Figure 3).
             +
             + This type of nodes only needs to book the existing
             + resources and no cross-connection setup
             + command is needed.
    ---------+---
       C2   + Reusing existing resource only on one of the interfaces,
             + either input or output interfaces and need to use new
             + resource on the other interface. (node C & E in Figure 3).
             +
             + This type of nodes needs to book 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 re-use the
             + existing resources on the other interfaces.
    ---------+---
       C3   + Using new resources on both interfaces.
             + (node F & G in Figure 3).
             +
             + This type of nodes needs to book the new resources
             + and send the cross-connection setup
             + command on both interfaces.
    ---------+---

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

3.3. LSP Reversion

If the end-to-end LSP recovery is revertive, as described in Section 2, 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
   working or protecting LSP are reconfigured while removing
   reservations for the restoration LSP.

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

In 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.

3.3.1. 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 working or protecting LSP on every node where resources are shared. Whenever reservation for restoration LSP is removed from a node, data plane configuration changes to reflect reservations of working or protection 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 few limitations as follows which may not be acceptable in some networks:

o No rollback

Deletion of restoration LSPs is not a revertive process. If for some reason reconfiguration of data plane on one of the nodes to match working or protection 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 nodal or DCN failures 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 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 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 protection 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.

3.3.2. 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
   restoration LSP, head-end chooses to establish a new LSP to
   reconfigure resources on the working or protection LSP path, and uses
   identical ASSOCIATION and PROTECTION objects from the LSP it is
   replacing. Only if setup of this LSP is successful will other
   (restoration and working/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 protection LSP's
   path. Reversion LSP is sharing resources both with working and
   restoration LSPs. As reversion LSP is created, resources are
   reconfigured to match its reservations. Hence, after reversion LSP
   is created, data plane configuration essentially reflects working or
   protecting LSP reservations.

   B) Break part:

After "make" part is finished, working and restoration LSPs are torn
down. Removing reservations for working and 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 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. 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 its DCN 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 protection LSP reservations on intermediate nodes when it receives
Resv for the reversion LSP.

4. 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].

5. IANA Considerations

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

6. References

6.1. 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.

6.2. Informative References

[PCEP-RSO]   X. Zhang, et al, "Extensions to Path Computation Element
             Protocol (PCEP) to Support Resource Sharing-based Path
             Computation", work in progress, February 2014.

[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

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

   Gabriele Maria Galimberti
   Cisco Systems, Inc.

   EMail: ggalimbe@cisco.com

   Pawel Brzozowski
   ADVA Optical

   EMail: PBrzozowski@advaoptical.com