     Framework for Multi-Protocol Label Switching (MPLS)-based Recovery

Status of this Memo

Copyright Notice

Abstract

   Multi-protocol label switching (MPLS) integrates the label swapping
   forwarding paradigm with network layer routing. To deliver reliable
   service, MPLS requires a set of procedures to provide protection of
   the traffic carried on different paths. This requires that the label
   switching routers (LSRs) support fault detection, fault notification,
   and fault recovery mechanisms, and that MPLS signaling support the
   configuration of recovery. With these objectives in mind, this
   document specifies a framework for MPLS based recovery. Restart
   issues are not included in this framework.

Table of Contents

1. Introduction

   This memo describes a framework for MPLS-based recovery. We provide
   a detailed taxonomy of recovery terminology, and discuss the
   motivation for, the objectives of, and the requirements for MPLS-
   based recovery. We outline principles for MPLS-based recovery, and
   also provide comparison criteria that may serve as a basis for
   comparing and evaluating different recovery schemes.

At points in the document, we provide some thoughts about the
operation or viability of certain recovery objectives. These should
be viewed as the opinions of the authors, and not the consolidated
views of the IETF. The document is informational and it is expected
that a standards track document will be developed in the future to
describe a subset of this document as to meet the needs currently
specified by the TE WG.

1.1. Background

Network routing deployed today is focused primarily on connectivity,
and typically supports only one class of service, the best effort
class. Multi-protocol label switching [RFC3031], on the other hand,
by integrating forwarding based on label-swapping of a link local
label with network layer routing allows flexibility in the delivery
of new routing services. MPLS allows for using such media-specific
forwarding mechanisms as label swapping. This enables some
sophisticated features such as quality-of-service (QoS) and traffic
engineering [RFC2702] to be implemented more effectively. An
important component of providing QoS, however, is the ability to
transport data reliably and efficiently. Although the current
routing algorithms are robust and survivable, the amount of time they
take to recover from a fault can be significant, in the order of
several seconds (for interior gateway protocols (IGPs)) or minutes
(for exterior gateway protocols, such as the Border Gateway Protocol
(BGP)), causing disruption of service for some applications in the
interim. This is unacceptable in situations where the aim is to
provide a highly reliable service, with recovery times that are in
the order of seconds down to 10's of milliseconds. IP routing may
also not be able to provide bandwidth recovery, where the objective
is to provide not only an alternative path, but also bandwidth
equivalent to that available on the original path. (For some recent
work on bandwidth recovery schemes, the reader is referred to [MPLS-
BACKUP].) Examples of such applications are Virtual Leased Line
services, Stock Exchange data services, voice traffic, video services
etc, i.e., every application that gets a disruption in service long
enough to not fulfill service agreements or the required level of
quality.

MPLS recovery may be motivated by the notion that there are
limitations to improving the recovery times of current routing
algorithms. Additional improvement can be obtained by augmenting
these algorithms with MPLS recovery mechanisms [MPLS-PATH]. Since
MPLS is a possible technology of choice in future IP-based transport
networks, it is useful that MPLS be able to provide protection and
restoration of traffic. MPLS may facilitate the convergence of
network functionality on a common control and management plane.
Further, a protection priority could be used as a differentiating

mechanism for premium services that require high reliability, such as
Virtual Leased Line services, and high priority voice and video
traffic. The remainder of this document provides a framework for
MPLS based recovery. It is focused at a conceptual level and is
meant to address motivation, objectives and requirements. Issues of
mechanism, policy, routing plans and characteristics of traffic
carried by recovery paths are beyond the scope of this document.

1.2. Motivation for MPLS-Based Recovery

MPLS based protection of traffic (called MPLS-based Recovery) is
useful for a number of reasons. The most important is its ability to
increase network reliability by enabling a faster response to faults
than is possible with traditional Layer 3 (or IP layer) approaches
alone while still providing the visibility of the network afforded by
Layer 3. Furthermore, a protection mechanism using MPLS could enable
IP traffic to be put directly over WDM optical channels and provide a
recovery option without an intervening SONET layer or optical
protection. This would facilitate the construction of IP-over-WDM
networks that request a fast recovery ability (Note that what is
meant here is the transport of IP traffic over WDM links, not the
Generalized MPLS, or GMPLS, control of a WDM link).

The need for MPLS-based recovery arises because of the following:

I.   Layer 3 or IP rerouting may be too slow for a core MPLS network
      that needs to support recovery times that are smaller than the
      convergence times of IP routing protocols.

II.   Layer 3 or IP rerouting does not provide the ability to provide
      bandwidth protection to specific flows (e.g., voice over IP,
      virtual leased line services).

III. Layer 0 (for example, optical layer) or Layer 1 (for example,
      SONET) mechanisms may be wasteful use of resources.

IV.   The granularity at which the lower layers may be able to protect
      traffic may be too coarse for traffic that is switched using
      MPLS-based mechanisms.

V.   Layer 0 or Layer 1 mechanisms may have no visibility into higher
      layer operations. Thus, while they may provide, for example,
      link protection, they cannot easily provide node protection or
      protection of traffic transported at layer 3. Further, this may
      prevent the lower layers from providing restoration based on the
      traffic's needs. For example, fast restoration for traffic that
      needs it, and slower restoration (with possibly more optimal use
      of resources) for traffic that does not require fast

restoration. In networks where the latter class of traffic is
dominant, providing fast restoration to all classes of traffic
may not be cost effective from a service provider's perspective.

VI.   MPLS has desirable attributes when applied to the purpose of
recovery for connectionless networks. Specifically that an LSP
is source routed and a forwarding path for recovery can be
"pinned" and is not affected by transient instability in SPF
routing brought on by failure scenarios.

VII. Establishing interoperability of protection mechanisms between
routers/LSRs from different vendors in IP or MPLS networks is
desired to enable recovery mechanisms to work in a multivendor
environment, and to enable the transition of certain protected
services to an MPLS core.

1.3. Objectives/Goals

The following are some important goals for MPLS-based recovery.

I.   MPLS-based recovery mechanisms may be subject to the traffic
engineering goal of optimal use of resources.

II.   MPLS based recovery mechanisms should aim to facilitate
restoration times that are sufficiently fast for the end user
application. That is, that better match the end-user's
application requirements. In some cases, this may be as short
as 10s of milliseconds.

We observe that I and II may be conflicting objectives, and a trade
off may exist between them. The optimal choice depends on the end-
user application's sensitivity to restoration time and the cost
impact of introducing restoration in the network, as well as the
end-user application's sensitivity to cost.

III. MPLS-based recovery should aim to maximize network reliability
and availability. MPLS-based recovery of traffic should aim to
minimize the number of single points of failure in the MPLS
protected domain.

IV.   MPLS-based recovery should aim to enhance the reliability of
the protected traffic while minimally or predictably degrading
the traffic carried by the diverted resources.

V.   MPLS-based recovery techniques should aim to be applicable for
protection of traffic at various granularities. For example,
it should be possible to specify MPLS-based recovery for a
portion of the traffic on an individual path, for all traffic

on an individual path, or for all traffic on a group of paths.
Note that a path is used as a general term and includes the
notion of a link, IP route or LSP.

VI.   MPLS-based recovery techniques may be applicable for an entire
end-to-end path or for segments of an end-to-end path.

VII.   MPLS-based recovery mechanisms should aim to take into
consideration the recovery actions of lower layers. MPLS-based
mechanisms should not trigger lower layer protection switching
nor should MPLS-based mechanisms be triggered when lower layer
switching has or may imminently occur.

VIII.   MPLS-based recovery mechanisms should aim to minimize the loss
of data and packet reordering during recovery operations. (The
current MPLS specification itself has no explicit requirement
on reordering.)

IX.   MPLS-based recovery mechanisms should aim to minimize the state
overhead incurred for each recovery path maintained.

X.   MPLS-based recovery mechanisms should aim to minimize the
signaling overhead to setup and maintain recovery paths and to
notify failures.

XI.   MPLS-based recovery mechanisms should aim to preserve the
constraints on traffic after switchover, if desired. That is,
if desired, the recovery path should meet the resource
requirements of, and achieve the same performance
characteristics as, the working path.

We observe that some of the above are conflicting goals, and real
deployment will often involve engineering compromises based on a
variety of factors such as cost, end-user application requirements,
network efficiency, complexity involved, and revenue considerations.
Thus, these goals are subject to tradeoffs based on the above
considerations.

2.   Overview

There are several options for providing protection of traffic. The
most generic requirement is the specification of whether recovery
should be via Layer 3 (or IP) rerouting or via MPLS protection
switching or rerouting actions.

Generally network operators aim to provide the fastest, most stable,
and the best protection mechanism that can be provided at a
reasonable cost. The higher the levels of protection, the more the

resources consumed. Therefore it is expected that network operators
will offer a spectrum of service levels. MPLS-based recovery should
give the flexibility to select the recovery mechanism, choose the
granularity at which traffic is protected, and to also choose the
specific types of traffic that are protected in order to give
operators more control over that tradeoff. With MPLS-based recovery,
it can be possible to provide different levels of protection for
different classes of service, based on their service requirements.
For example, using approaches outlined below, a Virtual Leased Line
(VLL) service or real-time applications like Voice over IP (VoIP) may
be supported using link/node protection together with pre-
established, pre-reserved path protection. Best effort traffic, on
the other hand, may use path protection that is established on demand
or may simply rely on IP re-route or higher layer recovery
mechanisms. As another example of their range of application, MPLS-
based recovery strategies may be used to protect traffic not
originally flowing on label switched paths, such as IP traffic that
is normally routed hop-by-hop, as well as traffic forwarded on label
switched paths.

2.1. Recovery Models

There are two basic models for path recovery: rerouting and
protection switching.

Protection switching and rerouting, as defined below, may be used
together. For example, protection switching to a recovery path may
be used for rapid restoration of connectivity while rerouting
determines a new optimal network configuration, rearranging paths, as
needed, at a later time.

2.1.1 Rerouting

Recovery by rerouting is defined as establishing new paths or path
segments on demand for restoring traffic after the occurrence of a
fault. The new paths may be based upon fault information, network
routing policies, pre-defined configurations and network topology
information. Thus, upon detecting a fault, paths or path segments to
bypass the fault are established using signaling.

Once the network routing algorithms have converged after a fault, it
may be preferable, in some cases, to reoptimize the network by
performing a reroute based on the current state of the network and
network policies. This is discussed further in Section 3.8.

In terms of the principles defined in section 3, reroute recovery
employs paths established-on-demand with resources reserved-on-
demand.

2.1.2 Protection Switching

   Protection switching recovery mechanisms pre-establish a recovery
   path or path segment, based upon network routing policies, the
   restoration requirements of the traffic on the working path, and
   administrative considerations. The recovery path may or may not be
   link and node disjoint with the working path. However if the
   recovery path shares sources of failure with the working path, the
   overall reliability of the construct is degraded. When a fault is
   detected, the protected traffic is switched over to the recovery
   path(s) and restored.

   In terms of the principles in section 3, protection switching employs
   pre-established recovery paths, and, if resource reservation is
   required on the recovery path, pre-reserved resources. The various
   sub-types of protection switching are detailed in Section 4.4 of this
   document.

2.2.   The Recovery Cycles

   There are three defined recovery cycles: the MPLS Recovery Cycle, the
   MPLS Reversion Cycle and the Dynamic Re-routing Cycle. The first
   cycle detects a fault and restores traffic onto MPLS-based recovery
   paths. If the recovery path is non-optimal the cycle may be followed
   by any of the two latter cycles to achieve an optimized network
   again. The reversion cycle applies for explicitly routed traffic
   that does not rely on any dynamic routing protocols to converge. The
   dynamic re-routing cycle applies for traffic that is forwarded based
   on hop-by-hop routing.

2.2.1 MPLS Recovery Cycle Model

   The MPLS recovery cycle model is illustrated in Figure 1. Definitions
   and a key to abbreviations follow.

```
   --Network Impairment
   |    --Fault Detected
   |    |     --Start of Notification
   |    |     |     -- Start of Recovery Operation
   |    |     |     |    --Recovery Operation Complete
   |    |     |     |    |    --Path Traffic Recovered
   |    |     |     |    |    |
   |    |     |     |    |    |
   |    |     |     |    |    |
   v    v     v     v    v    v
   -------------------------------------------------------------
   | T1 | T2 | T3 | T4 | T5 |
```
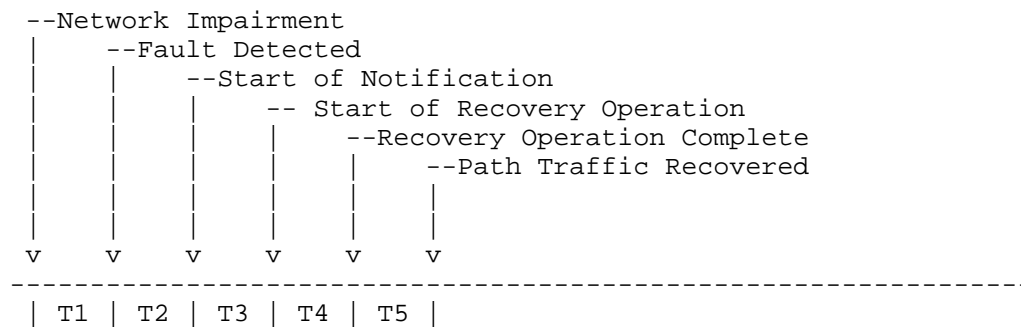
   Figure 1. MPLS Recovery Cycle Model

The various timing measures used in the model are described below.

    T1   Fault Detection Time
    T2   Fault Hold-off Time
    T3   Fault Notification Time
    T4   Recovery Operation Time
    T5   Traffic Recovery Time

Definitions of the recovery cycle times are as follows:

Fault Detection Time

    The time between the occurrence of a network impairment and the
    moment the fault is detected by MPLS-based recovery mechanisms.
    This time may be highly dependent on lower layer protocols.

Fault Hold-Off Time

    The configured waiting time between the detection of a fault and
    taking MPLS-based recovery action, to allow time for lower layer
    protection to take effect. The Fault Hold-off Time may be zero.

    Note: The Fault Hold-Off Time may occur after the Fault
    Notification Time interval if the node responsible for the
    switchover, the Path Switch LSR (PSL), rather than the detecting
    LSR, is configured to wait.

Fault Notification Time

    The time between initiation of a Fault Indication Signal (FIS) by
    the LSR detecting the fault and the time at which the Path Switch
    LSR (PSL) begins the recovery operation. This is zero if the PSL
    detects the fault itself or infers a fault from such events as an
    adjacency failure.

    Note: If the PSL detects the fault itself, there still may be a
    Fault Hold-Off Time period between detection and the start of the
    recovery operation.

Recovery Operation Time

    The time between the first and last recovery actions. This may
    include message exchanges between the PSL and PML (Path Merge LSR)
    to coordinate recovery actions.

   Traffic Recovery Time

      The time between the last recovery action and the time that the
      traffic (if present) is completely recovered. This interval is
      intended to account for the time required for traffic to once
      again arrive at the point in the network that experienced
      disrupted or degraded service due to the occurrence of the fault
      (e.g., the PML). This time may depend on the location of the
      fault, the recovery mechanism, and the propagation delay along the
      recovery path.

2.2.2 MPLS Reversion Cycle Model

   Protection switching, revertive mode, requires the traffic to be
   switched back to a preferred path when the fault on that path is
   cleared. The MPLS reversion cycle model is illustrated in Figure 2.
   Note that the cycle shown below comes after the recovery cycle shown
   in Fig. 1.

```
    --Network Impairment Repaired
    |    --Fault Cleared
    |    |    --Path Available
    |    |    |    --Start of Reversion Operation
    |    |    |    |    --Reversion Operation Complete
    |    |    |    |    |    --Traffic Restored on Preferred Path
    |    |    |    |    |    |
    |    |    |    |    |    |
    |    |    |    |    |    |
    v    v    v    v    v    v
   ----------------------------------------------------------------
    | T7 | T8 | T9 | T10| T11|
```

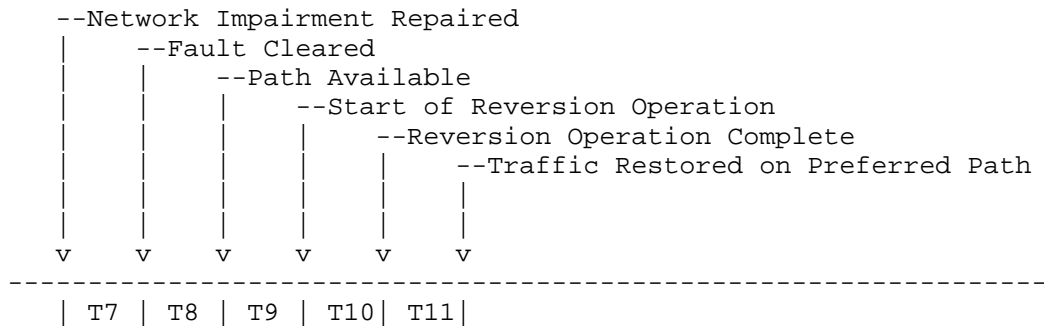   Figure 2. MPLS Reversion Cycle Model

   The various timing measures used in the model are described below.

   T7   Fault Clearing Time
   T8   Clear Hold-Off Time
   T9   Clear Notification Time
   T10 Reversion Operation Time
   T11 Traffic Reversion Time

   Note that time T6 (not shown above) is the time for which the network
   impairment is not repaired and traffic is flowing on the recovery
   path.

Definitions of the reversion cycle times are as follows:

Fault Clearing Time

   The time between the repair of a network impairment and the time
   that MPLS-based mechanisms learn that the fault has been cleared.
   This time may be highly dependent on lower layer protocols.

Clear Hold-Off Time

   The configured waiting time between the clearing of a fault and
   MPLS-based recovery action(s). Waiting time may be needed to
   ensure that the path is stable and to avoid flapping in cases
   where a fault is intermittent. The Clear Hold-Off Time may be
   zero.

   Note: The Clear Hold-Off Time may occur after the Clear
   Notification Time interval if the PSL is configured to wait.

Clear Notification Time

   The time between initiation of a Fault Recovery Signal (FRS) by
   the LSR clearing the fault and the time at which the path switch
   LSR begins the reversion operation. This is zero if the PSL
   clears the fault itself.

   Note: If the PSL clears the fault itself, there still may be a
   Clear Hold-off Time period between fault clearing and the start of
   the reversion operation.

Reversion Operation Time

   The time between the first and last reversion actions. This may
   include message exchanges between the PSL and PML to coordinate
   reversion actions.

Traffic Reversion Time

   The time between the last reversion action and the time that
   traffic (if present) is completely restored on the preferred path.
   This interval is expected to be quite small since both paths are
   working and care may be taken to limit the traffic disruption
   (e.g., using "make before break" techniques and synchronous
   switch-over).

   In practice, the most interesting times in the reversion cycle are
   the Clear Hold-off Time and the Reversion Operation Time together
   with Traffic Reversion Time (or some other measure of traffic

disruption). The first interval is to ensure stability of the
repaired path and the latter one is to minimize disruption time
while the reversion action is in progress.

Given that both paths are available, it is better to wait to have
a well-controlled switch-back with minimal disruption than have an
immediate operation that may cause new faults to be introduced
(except, perhaps, when the recovery path is unable to offer a
quality of service comparable to the preferred path).

2.2.3 Dynamic Re-routing Cycle Model

Dynamic rerouting aims to bring the IP network to a stable state
after a network impairment has occurred. A re-optimized network is
achieved after the routing protocols have converged, and the traffic
is moved from a recovery path to a (possibly) new working path. The
steps involved in this mode are illustrated in Figure 3.

Note that the cycle shown below may be overlaid on the recovery cycle
shown in Fig. 1 or the reversion cycle shown in Fig. 2, or both (in
the event that both the recovery cycle and the reversion cycle take
place before the routing protocols converge), and occurs if after the
convergence of the routing protocols it is determined (based on on-
line algorithms or off-line traffic engineering tools, network
configuration, or a variety of other possible criteria) that there is
a better route for the working path.

```
    --Network Enters a Semi-stable State after an Impairment
   |       --Dynamic Routing Protocols Converge
   |      |        --Initiate Setup of New Working Path between PSL
   |      |       |                                          and PML
   |      |       |      --Switchover Operation Complete
   |      |       |     |       --Traffic Moved to New Working Path
   |      |       |     |      |
   |      |       |     |      |
   |      |       |     |      |
   v      v       v     v      v

   ----------------------------------------------------------------
     | T12 | T13 | T14 | T15 |
```
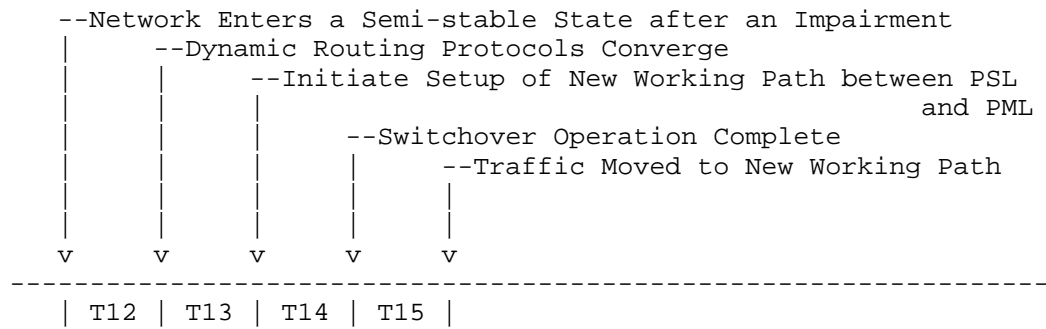
Figure 3. Dynamic Rerouting Cycle Model

The various timing measures used in the model are described below.

T12 Network Route Convergence Time
T13 Hold-down Time (optional)
T14 Switchover Operation Time
T15 Traffic Restoration Time

   Network Route Convergence Time

      We define the network route convergence time as the time taken for
      the network routing protocols to converge and for the network to
      reach a stable state.

   Holddown Time

      We define the holddown period as a bounded time for which a
      recovery path must be used. In some scenarios it may be difficult
      to determine if the working path is stable. In these cases a
      holddown time may be used to prevent excess flapping of traffic
      between a working and a recovery path.

   Switchover Operation Time

      The time between the first and last switchover actions. This may
      include message exchanges between the PSL and PML to coordinate
      the switchover actions.

   Traffic Restoration Time

      The time between the last restoration action and the time that
      traffic (if present) is completely restored on the new preferred
      path.

2.2.4 Example Recovery Cycle

   As an example of the recovery cycle, we present a sequence of events
   that occur after a network impairment occurs and when a protection
   switch is followed by dynamic rerouting.

     I. Link or path fault occurs
    II. Signaling initiated (FIS) for the detected fault
   III. FIS arrives at the PSL
    IV. The PSL initiates a protection switch to a pre-configured
        recovery path
     V. The PSL switches over the traffic from the working path to the
        recovery path
    VI. The network enters a semi-stable state
   VII. Dynamic routing protocols converge after the fault, and a new
        working path is calculated (based, for example, on some of the
        criteria mentioned in Section 2.1.1).
  VIII. A new working path is established between the PSL and the PML
        (assumption is that PSL and PML have not changed)
    IX. Traffic is switched over to the new working path.

2.3.   Definitions and Terminology

   This document assumes the terminology given in [RFC3031], and, in
   addition, introduces the following new terms.

2.3.1 General Recovery Terminology

   Re-routing

      A recovery mechanism in which the recovery path or path segments
      are created dynamically after the detection of a fault on the
      working path. In other words, a recovery mechanism in which the
      recovery path is not pre-established.

   Protection Switching

      A recovery mechanism in which the recovery path or path segments
      are created prior to the detection of a fault on the working path.
      In other words, a recovery mechanism in which the recovery path is
      pre-established.

   Working Path

      The protected path that carries traffic before the occurrence of a
      fault. The working path can be of different kinds; a hop-by-hop
      routed path, a trunk, a link, an LSP or part of a multipoint-to-
      point LSP.

      Synonyms for a working path are primary path and active path.

   Recovery Path

      The path by which traffic is restored after the occurrence of a
      fault. In other words, the path on which the traffic is directed
      by the recovery mechanism. The recovery path is established by
      MPLS means. The recovery path can either be an equivalent
      recovery path and ensure no reduction in quality of service, or be
      a limited recovery path and thereby not guarantee the same quality
      of service (or some other criteria of performance) as the working
      path. A limited recovery path is not expected to be used for an
      extended period of time.

      Synonyms for a recovery path are: back-up path, alternative path,
      and protection path.

   Protection Counterpart

      The "other" path when discussing pre-planned protection switching
      schemes. The protection counterpart for the working path is the
      recovery path and vice-versa.

   Path Switch LSR (PSL)

      An LSR that is responsible for switching or replicating the
      traffic between the working path and the recovery path.

   Path Merge LSR (PML)

      An LSR that is responsible for receiving the recovery path
      traffic, and either merging the traffic back onto the working
      path, or, if it is itself the destination, passing the traffic on
      to the higher layer protocols.

   Point of Repair (POR)

      An LSR that is setup for performing MPLS recovery. In other
      words, an LSR that is responsible for effecting the repair of an
      LSP. The POR, for example, can be a PSL or a PML, depending on
      the type of recovery scheme employed.

   Intermediate LSR

      An LSR on a working or recovery path that is neither a PSL nor a
      PML for that path.

   Path Group (PG)

      A logical bundling of multiple working paths, each of which is
      routed identically between a Path Switch LSR and a Path Merge LSR.

   Protected Path Group (PPG)

      A path group that requires protection.

   Protected Traffic Portion (PTP)

      The portion of the traffic on an individual path that requires
      protection. For example, code points in the EXP bits of the shim
      header may identify a protected portion.

Bypass Tunnel

   A path that serves to back up a set of working paths using the
   label stacking approach [RFC3031]. The working paths and the
   bypass tunnel must all share the same path switch LSR (PSL) and
   the path merge LSR (PML).

Switch-Over

   The process of switching the traffic from the path that the
   traffic is flowing on onto one or more alternate path(s). This
   may involve moving traffic from a working path onto one or more
   recovery paths, or may involve moving traffic from a recovery
   path(s) on to a more optimal working path(s).

Switch-Back

   The process of returning the traffic from one or more recovery
   paths back to the working path(s).

Revertive Mode

   A recovery mode in which traffic is automatically switched back
   from the recovery path to the original working path upon the
   restoration of the working path to a fault-free condition. This
   assumes a failed working path does not automatically surrender
   resources to the network.

Non-revertive Mode

   A recovery mode in which traffic is not automatically switched
   back to the original working path after this path is restored to a
   fault-free condition. (Depending on the configuration, the
   original working path may, upon moving to a fault-free condition,
   become the recovery path, or it may be used for new working
   traffic, and be no longer associated with its original recovery
   path, i.e., is surrendered to the network.)

MPLS Protection Domain

   The set of LSRs over which a working path and its corresponding
   recovery path are routed.

MPLS Protection Plan

   The set of all LSP protection paths and the mapping from working
   to protection paths deployed in an MPLS protection domain at a
   given time.

   Liveness Message

      A message exchanged periodically between two adjacent LSRs that
      serves as a link probing mechanism. It provides an integrity
      check of the forward and the backward directions of the link
      between the two LSRs as well as a check of neighbor aliveness.

   Path Continuity Test

      A test that verifies the integrity and continuity of a path or
      path segment. The details of such a test are beyond the scope of
      this document. (This could be accomplished, for example, by
      transmitting a control message along the same links and nodes as
      the data traffic or similarly could be measured by the absence of
      traffic and by providing feedback.)

2.3.2 Failure Terminology

   Path Failure (PF)

      Path failure is a fault detected by MPLS-based recovery
      mechanisms, which is defined as the failure of the liveness
      message test or a path continuity test, which indicates that path
      connectivity is lost.

   Path Degraded (PD)

      Path degraded is a fault detected by MPLS-based recovery
      mechanisms that indicates that the quality of the path is
      unacceptable.

   Link Failure (LF)

      A lower layer fault indicating that link continuity is lost. This
      may be communicated to the MPLS-based recovery mechanisms by the
      lower layer.

   Link Degraded (LD)

      A lower layer indication to MPLS-based recovery mechanisms that
      the link is performing below an acceptable level.

   Fault Indication Signal (FIS)

      A signal that indicates that a fault along a path has occurred.
      It is relayed by each intermediate LSR to its upstream or
      downstream neighbor, until it reaches an LSR that is setup to
      perform MPLS recovery (the POR). The FIS is transmitted

periodically by the node/nodes closest to the point of failure,
for some configurable length of time or until the transmitting
node receives an acknowledgement from its neighbor.

Fault Recovery Signal (FRS)

A signal that indicates a fault along a working path has been
repaired. Again, like the FIS, it is relayed by each intermediate
LSR to its upstream or downstream neighbor, until is reaches the
LSR that performs recovery of the original path. The FRS is
transmitted periodically by the node/nodes closest to the point of
failure, for some configurable length of time or until the
transmitting node receives an acknowledgement from its neighbor.

2.4.   Abbreviations

FIS: Fault Indication Signal.
FRS: Fault Recovery Signal.
LD: Link Degraded.
LF: Link Failure.
PD: Path Degraded.
PF: Path Failure.
PML: Path Merge LSR.
PG: Path Group.
POR: Point of Repair.
PPG: Protected Path Group.
PTP: Protected Traffic Portion.
PSL: Path Switch LSR.

3.   MPLS-based Recovery Principles

MPLS-based recovery refers to the ability to effect quick and
complete restoration of traffic affected by a fault in an MPLS-
enabled network. The fault may be detected on the IP layer or in
lower layers over which IP traffic is transported. Fastest MPLS
recovery is assumed to be achieved with protection switching and may
be viewed as the MPLS LSR switch completion time that is comparable
to, or equivalent to, the 50 ms switch-over completion time of the
SONET layer. Further, MPLS-based recovery may provide bandwidth
protection for paths that require it. This section provides a
discussion of the concepts and principles of MPLS-based recovery.
The concepts are presented in terms of atomic or primitive terms that
may be combined to specify recovery approaches. We do not make any
assumptions about the underlying layer 1 or layer 2 transport
mechanisms or their recovery mechanisms.

3.1.   Configuration of Recovery

   An LSR may support any or all of the following recovery options on a
   per-path basis:

   Default-recovery (No MPLS-based recovery enabled): Traffic on the
   working path is recovered only via Layer 3 or IP rerouting or by some
   lower layer mechanism such as SONET APS. This is equivalent to
   having no MPLS-based recovery. This option may be used for low
   priority traffic or for traffic that is recovered in another way (for
   example load shared traffic on parallel working paths may be
   automatically recovered upon a fault along one of the working paths
   by distributing it among the remaining working paths).

   Recoverable (MPLS-based recovery enabled): This working path is
   recovered using one or more recovery paths, either via rerouting or
   via protection switching.

3.2.   Initiation of Path Setup

   There are three options for the initiation of the recovery path
   setup. The active and recovery paths may be established by using
   either RSVP-TE [RFC2205][RFC3209] or CR-LDP [RFC3212], or by any
   other means including SNMP.

   Pre-established:

      This is the same as the protection switching option. Here a
      recovery path(s) is established prior to any failure on the
      working path. The path selection can either be determined by an
      administrative centralized tool, or chosen based on some algorithm
      implemented at the PSL and possibly intermediate nodes. To guard
      against the situation when the pre-established recovery path fails
      before or at the same time as the working path, the recovery path
      should have secondary configuration options as explained in
      Section 3.3 below.

   Pre-Qualified:

      A pre-established path need not be created, it may be pre-
      qualified. A pre-qualified recovery path is not created expressly
      for protecting the working path, but instead is a path created for
      other purposes that is designated as a recovery path after
      determining that it is an acceptable alternative for carrying the
      working path traffic. Variants include the case where an optical
      path or trail is configured, but no switches are set.

Established-on-Demand:

This is the same as the rerouting option. Here, a recovery path
is established after a failure on its working path has been
detected and notified to the PSL. The recovery path may be pre-
computed or computed on demand, which influences recovery times.

3.3. Initiation of Resource Allocation

A recovery path may support the same traffic contract as the working
path, or it may not. We will distinguish these two situations by
using different additive terms. If the recovery path is capable of
replacing the working path without degrading service, it will be
called an equivalent recovery path. If the recovery path lacks the
resources (or resource reservations) to replace the working path
without degrading service, it will be called a limited recovery path.
Based on this, there are two options for the initiation of resource
allocation:

Pre-reserved:

This option applies only to protection switching. Here a pre-
established recovery path reserves required resources on all hops
along its route during its establishment. Although the reserved
resources (e.g., bandwidth and/or buffers) at each node cannot be
used to admit more working paths, they are available to be used by
all traffic that is present at the node before a failure occurs.
The resources held by a set of recovery paths may be shared if
they protect resources that are not simultaneously subject to
failure.

Reserved-on-Demand:

This option may apply either to rerouting or to protection
switching. Here a recovery path reserves the required resources
after a failure on the working path has been detected and notified
to the PSL and before the traffic on the working path is switched
over to the recovery path.

Note that under both the options above, depending on the amount of
resources reserved on the recovery path, it could either be an
equivalent recovery path or a limited recovery path.

3.3.1   Subtypes of Protection Switching

   The resources (bandwidth, buffers, processing) on the recovery path
   may be used to carry either a copy of the working path traffic or
   extra traffic that is displaced when a protection switch occurs. This
   leads to two subtypes of protection switching.

   In 1+1 ("one plus one") protection, the resources (bandwidth,
   buffers, processing capacity) on the recovery path are fully
   reserved, and carry the same traffic as the working path. Selection
   between the traffic on the working and recovery paths is made at the
   path merge LSR (PML). In effect the PSL function is deprecated to
   establishment of the working and recovery paths and a simple
   replication function. The recovery intelligence is delegated to the
   PML.

   In 1:1 ("one for one") protection, the resources (if any) allocated
   on the recovery path are fully available to preemptible low priority
   traffic except when the recovery path is in use due to a fault on the
   working path. In other words, in 1:1 protection, the protected
   traffic normally travels only on the working path, and is switched to
   the recovery path only when the working path has a fault. Once the
   protection switch is initiated, the low priority traffic being
   carried on the recovery path may be displaced by the protected
   traffic. This method affords a way to make efficient use of the
   recovery path resources.

   This concept can be extended to 1:n (one for n) and m:n (m for n)
   protection.

3.4. Scope of Recovery

3.4.1 Topology

3.4.1.1 Local Repair

   The intent of local repair is to protect against a link or neighbor
   node fault and to minimize the amount of time required for failure
   propagation. In local repair (also known as local recovery), the
   node immediately upstream of the fault is the one to initiate
   recovery (either rerouting or protection switching). Local repair
   can be of two types:

Link Recovery/Restoration

In this case, the recovery path may be configured to route around
a certain link deemed to be unreliable. If protection switching
is used, several recovery paths may be configured for one working
path, depending on the specific faulty link that each protects
against.

Alternatively, if rerouting is used, upon the occurrence of a
fault on the specified link, each path is rebuilt such that it
detours around the faulty link.

In this case, the recovery path need only be disjoint from its
working path at a particular link on the working path, and may
have overlapping segments with the working path. Traffic on the
working path is switched over to an alternate path at the upstream
LSR that connects to the failed link. Link recovery is
potentially the fastest to perform the switchover, and can be
effective in situations where certain path components are much
more unreliable than others.

Node Recovery/Restoration

In this case, the recovery path may be configured to route around
a neighbor node deemed to be unreliable. Thus the recovery path
is disjoint from the working path only at a particular node and at
links associated with the working path at that node. Once again,
the traffic on the primary path is switched over to the recovery
path at the upstream LSR that directly connects to the failed
node, and the recovery path shares overlapping portions with the
working path.

3.4.1.2 Global Repair

The intent of global repair is to protect against any link or node
fault on a path or on a segment of a path, with the obvious exception
of the faults occurring at the ingress node of the protected path
segment. In global repair, the POR is usually distant from the
failure and needs to be notified by a FIS.

In global repair also, end-to-end path recovery/restoration applies.
In many cases, the recovery path can be made completely link and node
disjoint with its working path. This has the advantage of protecting
against all link and node fault(s) on the working path (end-to-end
path or path segment).

However, it may, in some cases, be slower than local repair since the
fault notification message must now travel to the POR to trigger the
recovery action.

3.4.1.3 Alternate Egress Repair

It is possible to restore service without specifically recovering the
faulted path.

For example, for best effort IP service it is possible to select a
recovery path that has a different egress point from the working path
(i.e., there is no PML). The recovery path egress must simply be a
router that is acceptable for forwarding the FEC carried by the
working path (without creating looping). In an engineering context,
specific alternative FEC/LSP mappings with alternate egresses can be
formed.

This may simplify enhancing the reliability of implicitly constructed
MPLS topologies. A PSL may qualify LSP/FEC bindings as candidate
recovery paths as simply link and node disjoint with the immediate
downstream LSR of the working path.

3.4.1.4 Multi-Layer Repair

Multi-layer repair broadens the network designer's tool set for those
cases where multiple network layers can be managed together to
achieve overall network goals. Specific criteria for determining
when multi-layer repair is appropriate are beyond the scope of this
document.

3.4.1.5 Concatenated Protection Domains

A given service may cross multiple networks and these may employ
different recovery mechanisms. It is possible to concatenate
protection domains so that service recovery can be provided end-to-
end. It is considered that the recovery mechanisms in different
domains may operate autonomously, and that multiple points of
attachment may be used between domains (to ensure there is no single
point of failure). Alternate egress repair requires management of
concatenated domains in that an explicit MPLS point of failure (the
PML) is by definition excluded. Details of concatenated protection
domains are beyond the scope of this document.

3.4.2   Path Mapping

   Path mapping refers to the methods of mapping traffic from a faulty
   working path on to the recovery path. There are several options for
   this, as described below. Note that the options below should be
   viewed as atomic terms that only describe how the working and
   protection paths are mapped to each other. The issues of resource
   reservation along these paths, and how switchover is actually
   performed lead to the more commonly used composite terms, such as 1+1
   and 1:1 protection, which were described in Section 4.3.1..

   1-to-1 Protection

      In 1-to-1 protection the working path has a designated recovery
      path that is only to be used to recover that specific working
      path.

   n-to-1 Protection

      In n-to-1 protection, up to n working paths are protected using
      only one recovery path. If the intent is to protect against any
      single fault on any of the working paths, the n working paths
      should be diversely routed between the same PSL and PML. In some
      cases, handshaking between PSL and PML may be required to complete
      the recovery, the details of which are beyond the scope of this
      document.

   n-to-m Protection

      In n-to-m protection, up to n working paths are protected using m
      recovery paths. Once again, if the intent is to protect against
      any single fault on any of the n working paths, the n working
      paths and the m recovery paths should be diversely routed between
      the same PSL and PML. In some cases, handshaking between PSL and
      PML may be required to complete the recovery, the details of which
      are beyond the scope of this document. n-to-m protection is for
      further study.

   Split Path Protection

      In split path protection, multiple recovery paths are allowed to
      carry the traffic of a working path based on a certain
      configurable load splitting ratio. This is especially useful when
      no single recovery path can be found that can carry the entire
      traffic of the working path in case of a fault. Split path
      protection may require handshaking between the PSL and the PML(s),
      and may require the PML(s) to correlate the traffic arriving on

multiple recovery paths with the working path. Although this is
an attractive option, the details of split path protection are
beyond the scope of this document.

3.4.3   Bypass Tunnels

It may be convenient, in some cases, to create a "bypass tunnel" for
a PPG between a PSL and PML, thereby allowing multiple recovery paths
to be transparent to intervening LSRs [RFC2702]. In this case, one
LSP (the tunnel) is established between the PSL and PML following an
acceptable route and a number of recovery paths can be supported
through the tunnel via label stacking. It is not necessary to apply
label stacking when using a bypass tunnel. A bypass tunnel can be
used with any of the path mapping options discussed in the previous
section.

As with recovery paths, the bypass tunnel may or may not have
resource reservations sufficient to provide recovery without service
degradation. It is possible that the bypass tunnel may have
sufficient resources to recover some number of working paths, but not
all at the same time. If the number of recovery paths carrying
traffic in the tunnel at any given time is restricted, this is
similar to the n-to-1 or n-to-m protection cases mentioned in Section
3.4.2.

3.4.4   Recovery Granularity

Another dimension of recovery considers the amount of traffic
requiring protection. This may range from a fraction of a path to a
bundle of paths.

3.4.4.1 Selective Traffic Recovery

This option allows for the protection of a fraction of traffic within
the same path. The portion of the traffic on an individual path that
requires protection is called a protected traffic portion (PTP). A
single path may carry different classes of traffic, with different
protection requirements. The protected portion of this traffic may
be identified by its class, as for example, via the EXP bits in the
MPLS shim header or via the priority bit in the ATM header.

3.4.4.2 Bundling

Bundling is a technique used to group multiple working paths together
in order to recover them simultaneously. The logical bundling of
multiple working paths requiring protection, each of which is routed
identically between a PSL and a PML, is called a protected path group

(PPG). When a fault occurs on the working path carrying the PPG, the
PPG as a whole can be protected either by being switched to a bypass
tunnel or by being switched to a recovery path.

3.4.5   Recovery Path Resource Use

In the case of pre-reserved recovery paths, there is the question of
what use these resources may be put to when the recovery path is not
in use. There are two options:

Dedicated-resource: If the recovery path resources are dedicated,
they may not be used for anything except carrying the working
traffic. For example, in the case of 1+1 protection, the working
traffic is always carried on the recovery path. Even if the recovery
path is not always carrying the working traffic, it may not be
possible or desirable to allow other traffic to use these resources.

Extra-traffic-allowed: If the recovery path only carries the working
traffic when the working path fails, then it is possible to allow
extra traffic to use the reserved resources at other times. Extra
traffic is, by definition, traffic that can be displaced (without
violating service agreements) whenever the recovery path resources
are needed for carrying the working path traffic.

Shared-resource: A shared recovery resource is dedicated for use by
multiple primary resources that (according to SRLGs) are not expected
to fail simultaneously.

3.5. Fault Detection

MPLS recovery is initiated after the detection of either a lower
layer fault or a fault at the IP layer or in the operation of MPLS-
based mechanisms. We consider four classes of impairments: Path
Failure, Path Degraded, Link Failure, and Link Degraded.

Path Failure (PF) is a fault that indicates to an MPLS-based recovery
scheme that the connectivity of the path is lost. This may be
detected by a path continuity test between the PSL and PML. Some,
and perhaps the most common, path failures may be detected using a
link probing mechanism between neighbor LSRs. An example of a
probing mechanism is a liveness message that is exchanged
periodically along the working path between peer LSRs [MPLS-PATH].
For either a link probing mechanism or path continuity test to be
effective, the test message must be guaranteed to follow the same
route as the working or recovery path, over the segment being tested.
In addition, the path continuity test must take the path merge points

into consideration. In the case of a bi-directional link implemented
as two unidirectional links, path failure could mean that either one
or both unidirectional links are damaged.

Path Degraded (PD) is a fault that indicates to MPLS-based recovery
schemes/mechanisms that the path has connectivity, but that the
quality of the connection is unacceptable. This may be detected by a
path performance monitoring mechanism, or some other mechanism for
determining the error rate on the path or some portion of the path.
This is local to the LSR and consists of excessive discarding of
packets at an interface, either due to label mismatch or due to TTL
errors, for example.

Link Failure (LF) is an indication from a lower layer that the link
over which the path is carried has failed. If the lower layer
supports detection and reporting of this fault (that is, any fault
that indicates link failure e.g., SONET LOS (Loss of Signal)), this
may be used by the MPLS recovery mechanism. In some cases, using LF
indications may provide faster fault detection than using only MPLS-
based fault detection mechanisms.

Link Degraded (LD) is an indication from a lower layer that the link
over which the path is carried is performing below an acceptable
level. If the lower layer supports detection and reporting of this
fault, it may be used by the MPLS recovery mechanism. In some cases,
using LD indications may provide faster fault detection than using
only MPLS-based fault detection mechanisms.

3.6.   Fault Notification

MPLS-based recovery relies on rapid and reliable notification of
faults. Once a fault is detected, the node that detected the fault
must determine if the fault is severe enough to require path
recovery. If the node is not capable of initiating direct action
(e.g., as a point of repair, POR) the node should send out a
notification of the fault by transmitting a FIS to the POR. This can
take several forms:

(i)   control plane messaging: relayed hop-by-hop along the path
      upstream of the failed LSP until a POR is reached.
(ii) user plane messaging: sent downstream to the PML, which may take
      corrective action (as a POR for 1+1) or communicate with a POR
      upstream (for 1:n) by any of several means:
   - control plane messaging
   - user plane return path (either through a bi-directional LSP or
      via other means)

Since the FIS is a control message, it should be transmitted with
high priority to ensure that it propagates rapidly towards the
affected POR(s). Depending on how fault notification is configured
in the LSRs of an MPLS domain, the FIS could be sent either as a
Layer 2 or Layer 3 packet [MPLS-PATH]. The use of a Layer 2-based
notification requires a Layer 2 path direct to the POR. An example
of a FIS could be the liveness message sent by a downstream LSR to
its upstream neighbor, with an optional fault notification field set
or it can be implicitly denoted by a teardown message.
Alternatively, it could be a separate fault notification packet. The
intermediate LSR should identify which of its incoming links to
propagate the FIS on.

3.7.   Switch-Over Operation

3.7.1 Recovery Trigger

   The activation of an MPLS protection switch following the detection
   or notification of a fault requires a trigger mechanism at the PSL.
   MPLS protection switching may be initiated due to automatic inputs or
   external commands. The automatic activation of an MPLS protection
   switch results from a response to a defect or fault conditions
   detected at the PSL or to fault notifications received at the PSL.
   It is possible that the fault detection and trigger mechanisms may be
   combined, as is the case when a PF, PD, LF, or LD is detected at a
   PSL and triggers a protection switch to the recovery path. In most
   cases, however, the detection and trigger mechanisms are distinct,
   involving the detection of fault at some intermediate LSR followed by
   the propagation of a fault notification to the POR via the FIS, which
   serves as the protection switch trigger at the POR. MPLS protection
   switching in response to external commands results when the operator
   initiates a protection switch by a command to a POR (or alternatively
   by a configuration command to an intermediate LSR, which transmits
   the FIS towards the POR).

   Note that the PF fault applies to hard failures (fiber cuts,
   transmitter failures, or LSR fabric failures), as does the LF fault,
   with the difference that the LF is a lower layer impairment that may
   be communicated to MPLS-based recovery mechanisms. The PD (or LD)
   fault, on the other hand, applies to soft defects (excessive errors
   due to noise on the link, for instance). The PD (or LD) results in a
   fault declaration only when the percentage of lost packets exceeds a
   given threshold, which is provisioned and may be set based on the
   service level agreement(s) in effect between a service provider and a
   customer.

3.7.2 Recovery Action

   After a fault is detected or FIS is received by the POR, the recovery
   action involves either a rerouting or protection switching operation.
   In both scenarios, the next hop label forwarding entry for a recovery
   path is bound to the working path.

3.8. Post Recovery Operation

   When traffic is flowing on the recovery path, decisions can be made
   as to whether to let the traffic remain on the recovery path and
   consider it as a new working path or to do a switch back to the old
   or to a new working path. This post recovery operation has two
   styles, one where the protection counterparts, i.e., the working and
   recovery path, are fixed or "pinned" to their routes, and one in
   which the PSL or other network entity with real-time knowledge of
   failure dynamically performs re-establishment or controlled
   rearrangement of the paths comprising the protected service.

3.8.1   Fixed Protection Counterparts

   For fixed protection counterparts the PSL will be pre-configured with
   the appropriate behavior to take when the original fixed path is
   restored to service. The choices are revertive and non-revertive
   mode. The choice will typically be dependent on relative costs of
   the working and protection paths, and the tolerance of the service to
   the effects of switching paths yet again. These protection modes
   indicate whether or not there is a preferred path for the protected
   traffic.

3.8.1.1 Revertive Mode

   If the working path always is the preferred path, this path will be
   used whenever it is available. Thus, in the event of a fault on this
   path, its unused resources will not be reclaimed by the network on
   failure. Resources here may include assigned labels, links,
   bandwidth etc. If the working path has a fault, traffic is switched
   to the recovery path. In the revertive mode of operation, when the
   preferred path is restored the traffic is automatically switched back
   to it.

   There are a number of implications to pinned working and recovery
   paths:

   -    upon failure and after traffic has been moved to the recovery
        path, the traffic is unprotected until such time as the path
        defect in the original working path is repaired and that path
        restored to service.

-   upon failure and after traffic has been moved to the recovery
    path, the resources associated with the original path remain
    reserved.

3.8.1.2 Non-revertive Mode

In the non-revertive mode of operation, there is no preferred path or
it may be desirable to minimize further disruption of the service
brought on by a revertive switching operation. A switch-back to the
original working path is not desired or not possible since the
original path may no longer exist after the occurrence of a fault on
that path. If there is a fault on the working path, traffic is
switched to the recovery path. When or if the faulty path (the
originally working path) is restored, it may become the recovery path
(either by configuration, or, if desired, by management actions).

In the non-revertive mode of operation, the working traffic may or
may not be restored to a new optimal working path or to the original
working path anyway. This is because it might be useful, in some
cases, to either: (a) administratively perform a protection switch
back to the original working path after gaining further assurances
about the integrity of the path, or (b) it may be acceptable to
continue operation on the recovery path, or (c) it may be desirable
to move the traffic to a new optimal working path that is calculated
based on network topology and network policies. Once a new working
path has been defined, an associated recovery path may be setup.

3.8.2   Dynamic Protection Counterparts

For dynamic protection counterparts when the traffic is switched over
to a recovery path, the association between the original working path
and the recovery path may no longer exist, since the original path
itself may no longer exist after the fault. Instead, when the
network reaches a stable state following routing convergence, the
recovery path may be switched over to a different preferred path
either optimization based on the new network topology and associated
information or based on pre-configured information.

Dynamic protection counterparts assume that upon failure, the PSL or
other network entity will establish new working paths if another
switch-over will be performed.

3.8.3   Restoration and Notification

   MPLS restoration deals with returning the working traffic from the
   recovery path to the original or a new working path. Restoration is
   performed by the PSL either upon receiving notification, via FRS,
   that the working path is repaired, or upon receiving notification
   that a new working path is established.

   For fixed counterparts in revertive mode, an LSR that detected the
   fault on the working path also detects the restoration of the working
   path. If the working path had experienced a LF defect, the LSR
   detects a return to normal operation via the receipt of a liveness
   message from its peer. If the working path had experienced a LD
   defect at an LSR interface, the LSR could detect a return to normal
   operation via the resumption of error-free packet reception on that
   interface. Alternatively, a lower layer that no longer detects a LF
   defect may inform the MPLS-based recovery mechanisms at the LSR that
   the link to its peer LSR is operational. The LSR then transmits FRS
   to its upstream LSR(s) that were transmitting traffic on the working
   path. At the point the PSL receives the FRS, it switches the working
   traffic back to the original working path.

   A similar scheme is used for dynamic counterparts where e.g., an
   update of topology and/or network convergence may trigger
   installation or setup of new working paths and may send notification
   to the PSL to perform a switch over.

   We note that if there is a way to transmit fault information back
   along a recovery path towards a PSL and if the recovery path is an
   equivalent working path, it is possible for the working path and its
   recovery path to exchange roles once the original working path is
   repaired following a fault. This is because, in that case, the
   recovery path effectively becomes the working path, and the restored
   working path functions as a recovery path for the original recovery
   path. This is important, since it affords the benefits of non-
   revertive switch operation outlined in Section 4.8.1, without leaving
   the recovery path unprotected.

3.8.4   Reverting to Preferred Path (or Controlled Rearrangement)

   In the revertive mode, "make before break" restoration switching can
   be used, which is less disruptive than performing protection
   switching upon the occurrence of network impairments. This will
   minimize both packet loss and packet reordering. The controlled
   rearrangement of paths can also be used to satisfy traffic
   engineering requirements for load balancing across an MPLS domain.

3.9. Performance

   Resource/performance requirements for recovery paths should be
   specified in terms of the following attributes:

   I.   Resource Class Attribute:
         Equivalent Recovery Class: The recovery path has the same
         performance guarantees as the working path. In other words, the
         recovery path meets the same SLAs as the working path.

         Limited Recovery Class: The recovery path does not have the same
         performance guarantees as the working path.

         A. Lower Class:
             The recovery path has lower resource requirements or less
             stringent performance requirements than the working path.

         B. Best Effort Class:
             The recovery path is best effort.

   II.   Priority Attribute:
         The recovery path has a priority attribute just like the working
         path (i.e., the priority attribute of the associated traffic
         trunks). It can have the same priority as the working path or
         lower priority.

   III. Preemption Attribute:
         The recovery path can have the same preemption attribute as the
         working path or a lower one.

4. MPLS Recovery Features

   The following features are desirable from an operational point of
   view:

   I.   It is desirable that MPLS recovery provides an option to
         identify protection groups (PPGs) and protection portions
         (PTPs).

   II.   Each PSL should be capable of performing MPLS recovery upon the
         detection of the impairments or upon receipt of notifications of
         impairments.

   III. A MPLS recovery method should not preclude manual protection
         switching commands. This implies that it would be possible
         under administrative commands to transfer traffic from a working
         path to a recovery path, or to transfer traffic from a recovery

        path to a working path, once the working path becomes
        operational following a fault.

    IV. A PSL may be capable of performing either a switch back to the
         original working path after the fault is corrected or a
         switchover to a new working path, upon the discovery or
         establishment of a more optimal working path.

    V.   The recovery model should take into consideration path merging
         at intermediate LSRs. If a fault affects the merged segment,
         all the paths sharing that merged segment should be able to
         recover. Similarly, if a fault affects a non-merged segment,
         only the path that is affected by the fault should be recovered.

5. Comparison Criteria

    Possible criteria to use for comparison of MPLS-based recovery
    schemes are as follows:

    Recovery Time

        We define recovery time as the time required for a recovery path
        to be activated (and traffic flowing) after a fault. Recovery
        Time is the sum of the Fault Detection Time, Hold-off Time,
        Notification Time, Recovery Operation Time, and the Traffic
        Restoration Time. In other words, it is the time between a
        failure of a node or link in the network and the time before a
        recovery path is installed and the traffic starts flowing on it.

    Full Restoration Time

        We define full restoration time as the time required for a
        permanent restoration. This is the time required for traffic to
        be routed onto links, which are capable of or have been engineered
        sufficiently to handle traffic in recovery scenarios. Note that
        this time may or may not be different from the "Recovery Time"
        depending on whether equivalent or limited recovery paths are
        used.

    Setup vulnerability

        The amount of time that a working path or a set of working paths
        is left unprotected during such tasks as recovery path computation
        and recovery path setup may be used to compare schemes. The
        nature of this vulnerability should be taken into account, e.g.,
        End to End schemes correlate the vulnerability with working paths,

   Local Repair schemes have a topological correlation that cuts
   across working paths and Network Plan approaches have a
   correlation that impacts the entire network.

Backup Capacity

   Recovery schemes may require differing amounts of "backup
   capacity" in the event of a fault. This capacity will be
   dependent on the traffic characteristics of the network. However,
   it may also be dependent on the particular protection plan
   selection algorithms as well as the signaling and re-routing
   methods.

Additive Latency

   Recovery schemes may introduce additive latency for traffic. For
   example, a recovery path may take many more hops than the working
   path. This may be dependent on the recovery path selection
   algorithms.

Quality of Protection

   Recovery schemes can be considered to encompass a spectrum of
   "packet survivability" which may range from "relative" to
   "absolute". Relative survivability may mean that the packet is on
   an equal footing with other traffic of, as an example, the same
   diff-serv code point (DSCP) in contending for the resources of the
   portion of the network that survives the failure. Absolute
   survivability may mean that the survivability of the protected
   traffic has explicit guarantees.

Re-ordering

   Recovery schemes may introduce re-ordering of packets. Also the
   action of putting traffic back on preferred paths might cause
   packet re-ordering.

State Overhead

   As the number of recovery paths in a protection plan grows, the
   state required to maintain them also grows. Schemes may require
   differing numbers of paths to maintain certain levels of coverage,
   etc. The state required may also depend on the particular scheme
   used for recovery. The state overhead may be a function of
   several parameters. For example, the number of recovery paths
   and the number of the protected facilities (links, nodes, or
   shared link risk groups (SRLGs)).

Loss

   Recovery schemes may introduce a certain amount of packet loss
   during switchover to a recovery path. Schemes that introduce loss
   during recovery can measure this loss by evaluating recovery times
   in proportion to the link speed.

   In case of link or node failure a certain packet loss is
   inevitable.

Coverage

   Recovery schemes may offer various types of failover coverage.
   The total coverage may be defined in terms of several metrics:

   I.   Fault Types: Recovery schemes may account for only link faults
        or both node and link faults or also degraded service. For
        example, a scheme may require more recovery paths to take node
        faults into account.

   II. Number of concurrent faults: dependent on the layout of recovery
        paths in the protection plan, multiple fault scenarios may be
        able to be restored.

   III. Number of recovery paths: for a given fault, there may be one or
        more recovery paths.

   IV. Percentage of coverage: dependent on a scheme and its
        implementation, a certain percentage of faults may be covered.
        This may be subdivided into percentage of link faults and
        percentage of node faults.

   V.   The number of protected paths may effect how fast the total set
        of paths affected by a fault could be recovered. The ratio of
        protection is n/N, where n is the number of protected paths and
        N is the total number of paths.

6. Security Considerations

   The MPLS recovery that is specified herein does not raise any
   security issues that are not already present in the MPLS
   architecture.

   Confidentiality or encryption of information on the recovery path is
   outside the scope of this document, but any method designed to do
   this in other contexts may be used with the methods described in this
   document.

7. Intellectual Property Considerations

   The IETF has been notified of intellectual property rights claimed in
   regard to some or all of the specification contained in this
   document. For more information consult the online list of claimed
   rights.

8. Acknowledgements

   We would like to thank members of the MPLS WG mailing list for their
   suggestions on the earlier versions of this document. In particular,
   Bora Akyol, Dave Allan, Dave Danenberg, Sharam Davari, and Neil
   Harrison whose suggestions and comments were very helpful in revising
   the document.

   The editors would like to give very special thanks to Curtis
   Villamizar for his careful and extremely thorough reading of the
   document and for taking the time to provide numerous suggestions,
   which were very helpful in the last couple of revisions of the
   document. Thanks are also due to Adrian Farrel for a through reading
   of the last version of the document, and to Jean-Phillipe Vasseur and
   Anna Charny for several useful editorial comments and suggestions,
   and for input on bandwidth recovery.

9. References

9.1 Normative

   [RFC3031]   Rosen, E., Viswanathan, A. and R. Callon,
                  "Multiprotocol Label Switching Architecture", RFC 3031,
                  January 2001.

   [RFC2702]   Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M. and
                  J. McManus, "Requirements for Traffic Engineering Over
                  MPLS", RFC 2702, September 1999.

   [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.

   [RFC3212]   Jamoussi, B. (Ed.), Andersson, L., Callon, R., Dantu,
                  R., Wu, L., Doolan, P., Worster, T., Feldman, N.,
                  Fredette, A., Girish, M., Gray, E., Heinanen, J.,
                  Kilty, T. and A. Malis, "Constraint-Based LSP Setup
                  using LDP", RFC 3212, January 2002.

9.2 Informative

   [MPLS-BACKUP] Vasseur, J. P., Charny, A., LeFaucheur, F., and
                 Achirica, "MPLS Traffic Engineering Fast reroute:
                 backup tunnel path computation for bandwidth
                 protection", Work in Progress.

   [MPLS-PATH]   Haung, C., Sharma, V., Owens, K., Makam, V. "Building
                 Reliable MPLS Networks Using a Path Protection
                 Mechanism", IEEE Commun. Mag., Vol. 40, Issue 3, March
                 2002, pp. 156-162.

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

10. Contributing Authors

   This document was the collective work of several individuals over a
   period of three years. The text and content of this document was
   contributed by the editors and the co-authors listed below. (The
   contact information for the editors appears in Section 11, and is not
   repeated below.)

   Ben Mack-Crane
   Tellabs Operations, Inc.
   1415 West Diehl Road
   Naperville, IL 60563

   Phone: (630) 798-6197
   EMail: Ben.Mack-Crane@tellabs.com

   Srinivas Makam
   Eshernet, Inc.
   1712 Ada Ct.
   Naperville, IL 60540

   Phone: (630) 308-3213
   EMail: Smakam60540@yahoo.com

Ken Owens
Edward Jones Investments
201 Progress Parkway
St. Louis, MO 63146

Phone: (314) 515-3431
EMail: ken.owens@edwardjones.com

Changcheng Huang
Carleton University
Minto Center, Rm. 3082
1125 Colonial By Drive
Ottawa, Ont. K1S 5B6 Canada

Phone: (613) 520-2600 x2477
EMail: Changcheng.Huang@sce.carleton.ca

Jon Weil

Brad Cain
Storigen Systems
650 Suffolk Street
Lowell, MA 01854

Phone: (978) 323-4454
EMail: bcain@storigen.com

Loa Andersson

EMail: loa@pi.se

Bilel Jamoussi
Nortel Networks
3 Federal Street, BL3-03
Billerica, MA 01821, USA

Phone:(978) 288-4506
EMail: jamoussi@nortelnetworks.com

   Angela Chiu
   AT&T Labs-Research
   200 Laurel Ave. Rm A5-1F13
   Middletown , NJ 07748

   Phone: (732) 420-9061
   EMail: chiu@research.att.com

   Seyhan Civanlar
   Lemur Networks, Inc.
   135 West 20th Street, 5th Floor
   New York, NY 10011

   Phone: (212) 367-7676
   EMail: scivanlar@lemurnetworks.com

11. Editors' Addresses

   Vishal Sharma (Editor)
   Metanoia, Inc.
   1600 Villa Street, Unit 352
   Mountain View, CA 94041-1174

   Phone: (650) 386-6723
   EMail: v.sharma@ieee.org

   Fiffi Hellstrand (Editor)
   Nortel Networks
   St Eriksgatan 115
   PO Box 6701
   113 85 Stockholm, Sweden

   Phone: +46 8 5088 3687
   EMail: fiffi@nortelnetworks.com

12. Full Copyright Statement

Acknowledgement