

**September, 2003**

Generalized MPLS Recovery Functional Specification

[draft-ietf-ccamp-gmpls-recovery-functional-01.txt](#)

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#) [[RFC2026](#)].

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference 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>.

Abstract

This document presents a functional description of the protocol extensions needed to support GMPLS-based recovery (i.e. protection and restoration). Protocol specific formats and mechanisms will be described in companion documents.



## Contributors

This document was the product of many individuals working together in the CCAMP WG Protection and Restoration design team. The following are the authors that contributed to this document:

Deborah Brungard (AT&T)  
Rm. D1-3C22 - 200 S. Laurel Ave.  
Middletown, NJ 07748, USA  
E-mail: dbrungard@att.com

Sudheer Dharanikota (Consult)  
E-mail: sudheer@ieee.org

Jonathan P. Lang  
Email: jplang@ieee.org

Guangzhi Li (AT&T)  
180 Park Avenue,  
Florham Park, NJ 07932, USA  
E-mail: gli@research.att.com

Eric Mannie (Consult)  
E-mail: eric\_mannie@hotmail.com

Dimitri Papadimitriou (Alcatel)  
Francis Wellesplein, 1  
B-2018 Antwerpen, Belgium  
E-mail: dimitri.papadimitriou@alcatel.be

Bala Rajagopalan  
Tellium  
2 Crescent Place - P.O. Box 901  
Oceanport, NJ 07757-0901, USA  
Email: braja@tellium.com

Yakov Rekhter (Juniper)  
1194 N. Mathilda Avenue  
Sunnyvale, CA 94089, USA  
E-mail: yakov@juniper.net

## **1. Introduction**

A requirement for the development of a common control plane for both optical and electronic switching equipment is that there must be signaling, routing, and link management mechanisms that support data plane fault recovery. In this document, the term "recovery" is generically used to denote both protection and restoration; the

specific terms "protection" and "restoration" are only used when differentiation is required. The subtle distinction between

protection and restoration is made based on the resource allocation done during the recovery period (see [TERM]).

A label-switched path (LSP) may be subject to local (span), segment, and/or end-to-end recovery. Local span protection refers to the protection of the link (and hence all the LSPs marked as required for span protection and routed over the link) between two neighboring switches. Segment protection refers to the recovery of an LSP segment (i.e., an SNC in the ITU-T terminology) between two nodes, i.e. the boundary nodes of the segment. End-to-end protection refers to the protection of an entire LSP from the ingress to the egress port. The end-to-end recovery models discussed in this draft apply to segment protection where the source and destination refer to the protected segment rather than the entire LSP. Multiple recovery levels may be used concurrently by a single LSP for added resiliency; however, the interaction between levels becomes affecting any one direction of the LSP results in both directions of the LSP being switched to a new span, segment, or end-to-end path.

Unless otherwise stated, all references to link in this draft indicate a bi-directional link (which may be realized as a pair of unidirectional links).

Consider the control plane message flow during the establishment of an LSP. This message flow proceeds from an initiating (or source) node to a terminating (or destination) node, via a sequence of intermediate nodes. A node along the LSP is said to be UPSTREAM from another node if the former occurs first in the sequence. The latter node is said to be DOWNSTREAM from the former node. That is, an UPSTREAM node is closer to the initiating node than a node further DOWNSTREAM. Unless otherwise stated, all references to UPSTREAM and DOWNSTREAM are in terms of the control plane message flow.

The flow of the data traffic is defined from ingress (source node) to egress (destination node). Note that for bi-directional LSPs there are two different data plane flows, one for each direction of the LSP. This document presents a protocol functional description to support GMPLS-based recovery (i.e., protection and restoration). Protocol specific formats and mechanisms will be described in companion documents.

## 2. Span Protection

Consider a (working) link *i* between two nodes A and B. There are two fundamental models for span protection. The first is referred to as 1+1 protection. Under this model, a dedicated link *j* is pre-assigned to protect link *i*. LSP traffic is permanently bridged onto both links *i* and *j* at the ingress node and the egress node selects the signal (i.e., normal traffic) from *i* or *j*, based on a selection

function (e.g., signal quality). Under unidirectional 1+1 span protection ([Section 2.1](#)), each node A and B acts autonomously to select the signal from the working link (i) or the protection link

(j). Under bi-directional 1+1 span protection ([Section 2.2](#)) the two nodes A and B coordinate the selection function such that they select the signal from the same link, i or j.

Under the second model, a set of N working links are protected by a set of M protection links, usually with  $M \leq N$ . A failure in any of the N working links results in traffic being switched to one of the M protection links that is available. This is typically a three-step process: first the data plane failure is detected at the egress node and reported (notification), then a protection link is selected, and finally, the LSPs on the failed link are moved to the protection link. If reversion is supported, a fourth step is included, i.e. return of the traffic to the working link (when the working link has recovered from the failure). In [Section 2.3](#), 1:1 span protection is described. In [Section 2.4](#), M:N span protection is described, where  $M \leq N$ .

## 2.1 Unidirectional 1+1 dedicated protection

Suppose a bi-directional LSP is routed over link i between two nodes A and B. Under unidirectional 1+1 protection, a dedicated link j is pre-assigned to protect the working link i. LSP traffic is permanently bridged on both links at the ingress node and the egress node selects the normal traffic from one of the links, i or j. If a node (A or B) detects a failure of a span, it autonomously invokes a process to receive the traffic from the protection span. Thus, it is possible that node A selects the signal from link i in the B to A direction of the LSP, and node B selects the signal from link j in the A to B direction.

The following functionality is required for 1+1 unidirectional span protection:

- o Routing: A single TE link encompassing both working and protection links should be announced with Link Protection Type "Dedicated 1+1" along with the bandwidth parameters for the working link. As the resources are consumed/released, the bandwidth parameters of the TE link are adjusted accordingly. Encoding of the Link Protection Type and bandwidth parameters in IS-IS is specified in [\[GMPLS-ISIS\]](#). Encoding of this information in OSPF is specified in [\[GMPLS-OSPF\]](#).
- o Signaling: The Link Protection object/TLV should be used to request "Dedicated 1+1" link protection for that LSP. This object/TLV is defined in [\[GMPLS-SIG\]](#). If the Link Protection object/TLV is not used, link selection is a matter of local policy. No additional signaling is required when a fail-over occurs.

- o Link management: Both nodes must have a consistent view of the link protection association for the spans. This can be done



using the Link Management Protocol (LMP), or if LMP is not used, this must be configured manually.

## **2.2 Bi-directional 1+1 dedicated protection**

Suppose an LSP is routed over link *i* between two nodes A and B. Under bi-directional 1+1 protection, a dedicated link *j* is pre-assigned to protect the working link *i*. LSP traffic is permanently duplicated on both links and under normal conditions, the traffic from link *i* is received by nodes A and B (in the appropriate directions). A failure affecting link *i* results in both A and B switching to the traffic on link *j* in the respective directions. Note that some form of signaling is required to ensure that both A and B start receiving from the protection link.

The basic steps in 1+1 bi-directional span protection are as follows:

1. If a node (A or B) detects the failure of the working link (or a degradation of signal quality over the working link), it should begin receiving on the protection link and send a switchover message reliably to the other node (B or A, respectively). This message should indicate the identity of the failed working link and other relevant information.
2. Upon receipt of the switchover message, a node MUST begin receiving from the protection link and send a switchover response message to the other node (A or B, respectively). Since both the working/protect spans are exposed to routing & signaling as a single link, the switchover should be transparent to routing and signaling.
  - o The routing procedures are the same as in 1+1 unidirectional.
  - o The signaling procedures are the same as in 1+1 unidirectional.
  - o In addition to the procedures described in 1+1 (unidirectional), a switchover request message must be used to signal the switchover request. This can be done using LMP. Note that GMPLS-based mechanisms may not be necessary when the underlying span (transport) technology provides such a mechanism.

## **2.3 Dedicated 1:1 protection with Extra Traffic**

Consider two adjacent nodes A and B. Under 1:1 protection, a dedicated link *j* between A and B is pre-assigned to protect working

link i. Link j may be carrying (preemptable) Extra Traffic. A failure affecting link i results in the corresponding LSP(s) being

restored to link j. Extra Traffic being routed over link j may need to be preempted to accommodate the LSPs that have to be restored.

Once a fault is isolated/localized, the affected LSP(s) must be moved to the protection link. The process of moving an LSP from a failed (working) link to a protection link must be initiated by one of the nodes, A or B. This node is referred to as the "master". The other node is called the "slave". The determination of the master and the slave may be based on configured information or protocol specific requirements.

The basic steps in dedicated 1:1 span protection (ignoring reversion) are as follows:

1. If the master detects/localizes a link failure event, it invokes a process to allocate the protection link to the affected LSP(s).
2. If the slave detects a link failure event, it informs the master of the failure using a failure indication message. The master then invokes the same procedure as (1) to move the LSPs to the protection link. If the protection link is carrying Extra Traffic, the slave stops using the span for the Extra Traffic.
3. Once the span protection procedure is invoked in the master, it requests the slave to switch the affected LSP(s) to the protection link. Prior to this, if the protection link is carrying Extra Traffic, the master stops using the span for this traffic (i.e., the traffic is dropped by the master and not forwarded into or out of the protection link).
4. The slave sends an acknowledgement to the master. Prior to this, the slave stops using the link for Extra Traffic (i.e., the traffic is dropped by the slave and not forwarded into or out of the protection link). It then starts sending the normal traffic on the selected protection link.
5. When the master receives the acknowledgement, it starts sending and receiving the normal traffic over the new link. The switchover of the LSPs is thus completed.

From the description above, it is clear that 1:1 span protection may require up to three signaling messages for each failed span: a failure indication message, an LSP switchover request message, and an LSP switchover response message. Furthermore, it may be possible to switch multiple LSPs from the working span to the protection span simultaneously.

- o Pre-emption MUST be supported to accommodate Extra Traffic.

o Routing: A single TE link encompassing both working and protection links is announced with Link Protection Type

"Dedicated 1:1". If Extra Traffic is supported over the protection link, then the bandwidth parameters for the protection link must also be announced. The differentiation between bandwidth for working and protect links is made using priority mechanisms. In other words, the network must be configured such that bandwidth at priority X or lower is considered Extra Traffic.

If there is a failure on the working link, then the normal traffic is switched to the protection link, preempting Extra Traffic if necessary. The bandwidth for the protection link must be adjusted accordingly.

- o Signaling: To establish an LSP on the working link, the Link Protection object/TLV indicating "Dedicated 1:1" should be included in the signaling request message for that LSP. To establish an LSP on the protection link, the appropriate priority (indicating Extra Traffic) should be used for that LSP. These objects/TLVs are defined in [\[GMPLS-SIG\]](#). If the Link Protection object/TLV is not used, link selection is a matter of local policy.
- o Link management: Both nodes must have a consistent view of the link protection association for the spans. This can be done using LMP or via manual configuration.
- o When a link failure is detected at the slave, a failure indication message must be sent to the master informing the node of the link failure.

#### **2.4 Shared M:N protection**

Shared M:N protection is described with respect to two neighboring nodes A and B. The scenario considered is as follows:

- o At any point in time, there are two sets of links between A and B, i.e., a working set of N (bi-directional) links carrying traffic subject to protection and a protection set of M (bi-directional) links. A protection link may be carrying Extra Traffic. There is no a priori relationship between the two sets of links, but the value of M and N may be pre-configured. The specific links in the protection set MAY be pre-configured to be physically diverse to avoid the possibility that failure events affect a large proportion of protection links (along with working links).
- o When a link in the working set is affected by a failure, the normal traffic is diverted to a link in the protection

set, if such a link is available. Note that such a link might be carrying more than one LSP, e.g., an OC-192 link carrying four STS-48 LSPs.

- o More than one link in the working set may be affected by the same failure event. In this case, there may not be an adequate number of protection links to accommodate all of the affected traffic carried by failed working links. The set of affected working links that are actually restored over available protection links is then subject to policies (e.g., based on relative priority of working traffic). These policies are not specified in this draft.
- o When normal traffic must be diverted from a failed link in the working set to a protection link, the decision as to which protection link is chosen is always made by one of the nodes, A or B. This node is considered the "master" and it is required to both apply any policies and select specific protection links to divert working traffic. The other node is considered the "slave". The determination of the master and the slave may be based on configured information, protocol specific requirements, or as a result of running a neighbor discovery procedure.
- o Failure events themselves are detected by transport layer mechanisms if available (e.g., SONET Alarm Indication Signal (AIS)/ Remote Defect Indication (RDI)). Since the bi-directional links are formed by a pair of unidirectional links, a failure in the link from A to B is typically detected by B and a failure in the opposite direction is detected by A. It is possible that a failure simultaneously affects both directions of the bi-directional link. In this case, A and B will concurrently detect failures, in the B-to-A direction and in the A-to-B direction, respectively.

The basic steps in M:N protection (ignoring reversion) are as follows:

1. If the master detects a failure of a working link, it autonomously invokes a process to allocate a protection link to the affected traffic.
2. If the slave detects a failure of a working link, it MUST inform the master of the failure using a failure indication message. The master then invokes the same procedure as above to allocate a protection link. (It is possible that the master has itself detected the same failure, for example, a failure simultaneously affecting both directions of a link).
3. Once the master has determined the identity of the protection link, it indicates this to the slave and requests

the switchover of the traffic (using a "switchover request" message). Prior to this, if the protection link is carrying Extra Traffic, the master stops using the link for this



traffic (i.e., the traffic is dropped by the master and not forwarded into or out of the protection link).

4. The slave sends a "switchover response" message back to the master. Prior to this, if the selected protection link is carrying traffic that could be preempted, the slave stops using the link for this traffic (i.e., the traffic is dropped by the slave and not forwarded into or out of the protection link). It then starts sending the normal traffic on the selected protection link.
5. When the master receives the switchover response, it starts sending and receiving the traffic that was previously carried on the now-failed link over the new link.

From the description above, it is clear that M:N span restoration (involving LSP local recovery) may require up to three messages for each working link being switched: a failure indication message, a switchover request message and a switchover response message.

- o Pre-emption MUST be supported to accommodate Extra Traffic.
- o Routing: A single TE link encompassing both sets of working and protect links should be announced with Link Protection Type "Shared M:N". If Extra Traffic is supported over set of the protection links, then the bandwidth parameters for the set of protection links must also be announced. The differentiation between bandwidth for working and protect links is made using priority mechanisms.

If there is a failure on a working link, then the affected LSP(s) must be switched to a protection link, preempting Extra Traffic if necessary. The bandwidth for the protection link must be adjusted accordingly.

- o Signaling: To establish an LSP on the working link, the Link Protection object/TLV indicating "Shared M:N" should be included in the signaling request message for that LSP. To establish an LSP on the protection link, the appropriate priority (indicating Extra Traffic) should be used for that. These objects/TLVs are defined in [GMPLS-SIG]. If the Link Protection object/TLV is not used, link selection is a matter of local policy.
- o For link management, both nodes must have a consistent view of the link protection association for the links. This can be done using LMP or via manual configuration.



## **2.6 Messages**

The following messages are used in local span protection procedures. All these messages must be transmitted reliably from the message source to the message destination.

### **2.6.1 Failure Indication Message**

This message is sent from the slave to the master to indicate the identities of one or more failed working links. (This message may not be necessary when the transport plane technology itself provides for such a notification).

The number of links included in the message would depend on the number of failures detected within a window of time by the sending node. A node may choose to send separate failure indication messages in the interest of completing the recovery for a given link within an implementation-dependent time constraint.

### **2.6.2 Switchover Request Message**

Under bi-directional 1+1 span protection, this message is used to coordinate the selecting function at both nodes. This message is originated at the node that detected the failure.

Under dedicated 1:1 and shared M:N span protection, this message is used as an LSP switchover request. This message is sent from the master node to the slave node (reliably) to indicate that the LSP(s) on the (failed) working link can be switched to an available protection link. If so, the ID of the protection link as well as the LSP labels (if necessary) must be indicated. These identifiers used must be consistent with those used in GMPLS signaling.

A working link may carry multiple LSPs. Since the normal traffic carried over the working link is switched to the protection link, it may be possible for the LSPs on the working link to be mapped to the protection link without re-signaling each individual LSP. For example, if link bundling [[BUNDLE](#)] is used where the working and protect links are mapped to component links, and the labels are the same on the working and protection links, it may be possible to change the component links without needing to re-signal each individual LSP. Optionally, the labels may need to be explicitly coordinated between the two nodes. In this case, the switchover request message should carry the new label mappings.

The master may not be able to find protection links to accommodate all failed working links. Thus, if this message is generated in response to a Failure Indication message from the slave then the set of failed links in the message may be a sub-set of the links received in the Failure Indication message. Depending on time

constraints, the master may switch the normal traffic from the set of failed links in smaller batches. Thus, a single failure indication message may result in the master sending more than one Switchover Request message to the same slave node.

### **2.6.3 Switchover Response Message**

This message is sent from the slave to the master (reliably) to indicate the completion (or failure) of switchover at the slave. In this message, the slave may indicate that it cannot switch over to the corresponding free link for some reason. The master and slave in this case notify the user (operator) of the failed switchover. A notification of the failure may also be used as a trigger in an end-to-end recovery.

### **2.7 Preventing Unintended Connections**

An unintended connection occurs when traffic from the wrong source is delivered to a receiver. This must be prevented during protection switching. This is primarily a concern when the protection link is being used to carry Extra Traffic. In this case, it must be ensured that the LSP traffic being switched from the (failed) working link to the protection link is not delivered to the receiver of the preempted traffic. Thus, in the message flow described above, the master node MUST disconnect (any) preempted traffic on the selected protection link before sending the Switchover Request. The slave node MUST also disconnect preempted traffic before sending the Switchover Response. In addition, the master node should start receiving traffic for the protected LSP from the protection link. Finally, the master node should start sending protected traffic on the protection link upon receipt of the Switchover Response.

### **3.0 End-to-End (Path) Protection and Restoration**

End-to-end path protection and restoration refer to the recovery of an entire LSP from the initiator to the terminator. Suppose the primary path of an LSP is routed from the initiator (Node A) to the terminator (Node B) through a set of intermediate nodes. In the following subsections, we describe three previously proposed end-to-end protection schemes and the functional steps needed to implement them.

#### **3.1 Unidirectional 1+1 Protection**

A dedicated, resource-disjoint alternate path is pre-established to protect the LSP. Traffic is simultaneously sent on both paths and received from one of the functional paths by the end nodes A and B.

There is no explicit signaling involved with this mode of protection.

#### **3.2 Bi-directional 1+1 Protection**

A dedicated, resource-disjoint alternate path is pre-established to protect the LSP. Traffic is simultaneously sent on both paths; under

normal conditions, the traffic from the working path is received by nodes A and B (in the appropriate directions). A failure affecting

the working path results in both A and B switching to the traffic on the protection path in the respective directions.

Note that this requires coordination between the end nodes to switch to the protection path.

The basic steps in bi-directional 1+1 path protection are as follows:

- o Failure detection: There are two possibilities for this.
  1. A node in the working path detects a failure event. Such a node must send a failure indication message towards the upstream or/and downstream end node of the LSP (node A or B). This message may be forwarded along the working path, or routed over a different path if the network has general routing intelligence. Mechanisms provided by the data transport plane may also be used for this, if available.
  2. The end nodes (A or B) detect the failure themselves (e.g., loss of light).
- o Switchover: The action when an end node detects a failure in the working path is as follows: Start receiving from the protection path. At the same time, send a switchover request message to the other end node to enable switching at the other end.

The action when an end node receives a switchover message is as follows:

- Start receiving from the protection path. At the same time, send a switchover response message to the other end node.

GMPLS signaling mechanisms may be used to (reliably) signal the switchover request. This message may be forwarded along the protection path if no other routing intelligence is available in the network.

### **3.2.1 Identifiers**

LSP Identifier: A unique identifier for each LSP. The LSP Identifier is within the scope of the Source ID and Destination ID.

Source ID: ID of the source (e.g., IP address).

Destination ID: ID of the destination (e.g., IP address).





### **3.2.2 Nodal Information**

Each node that is on the working or protection path of an LSP must have knowledge of the LSP identifier as well as the previous and next nodes in the LSP. This is so that restoration-related messages may be forwarded properly. The optical network may also have general routing intelligence. In this case, messages may be forwarded along paths different than that of the LSP.

The nodal information may be assembled when the working and protection paths of the LSP are provisioned using signaling, or may be configured when LSP provisioning does not involve signaling (e.g., provisioning through a management system). This information must remain until the LSP is explicitly de-provisioned.

### **3.2.3 End-to-End Failure Indication Message**

This message is sent (reliably) by an intermediate node towards the source of an LSP. For instance, such a node might have attempted local span protection and failed. This message may not be necessary if the data transport layer provides mechanisms for the notification of LSP failure by the endpoints (i.e. if LSP endpoints are co-located with a corresponding data (transport) maintenance/recovery domain).

Consider a node that detects a link failure. The node must determine the identities of all LSPs that are affected by the failure of the link, and send an end-to-end failure indication message to the source of each LSP. Each intermediate node receiving such a message must forward the message to the appropriate next node such that the message would ultimately reach the LSP source. Furthermore, if an intermediate node is itself generating a failure indication message, there SHOULD be a mechanism to suppress all but one source of failure indication messages. Finally, the failure indication message must be sent reliably from the node detecting the failure to the LSP source. Reliability may be achieved, for example, by re-transmitting the message until an acknowledgement is received.

### **3.2.4 End-to-End Failure Acknowledge Message**

This message is sent by the source node in response to an End-to-End failure indication message. This message is sent to the originator of the failure indication message. The acknowledge message should be sent for each failure indication message received. Each intermediate node receiving the acknowledge message must forward it towards the destination of the message.

### **3.2.5 End-to-End Switchover Request Message**

This message is generated by the source node receiving an indication

of failure in an LSP. It is sent to the LSP destination, and it carries the identifier of LSP being restored. The End-to-End

Switchover message must be sent reliably from the source to the destination of the LSP.

### **3.2.6 End-to-End Switchover Response Message**

This message is sent by the destination node receiving an End-to-End Switchover Request message towards the source of the LSP. This message should identify the LSP being switched over. This message must be transmitted in response to each End-to-End Switchover Request message received.

### **3.3 Shared Mesh Restoration**

Shared mesh restoration refers to schemes under which protection paths for multiple LSPs share common link and node resources. Under these schemes, the protection capacity is pre-reserved, i.e., link capacity is allocated to protect one or more LSPs but explicit action is required to instantiate a specific protection LSP. This requires restoration signaling along the protection path. Typically, the protection capacity is shared only amongst LSPs whose working paths are physically diverse. This criterion can be enforced when provisioning the protection path. Specifically, provisioning-related signaling messages may carry information about the working path to nodes along the protection path. This can be used as call admission control to accept/reject connections along the protection path based on the identification of the resources used for the primary path.

Thus, shared mesh restoration is designed to protect an LSP after a single failure event, i.e., a failure that affects the working path of at most one LSP sharing the protection capacity. It is possible that a protection path may not be successfully activated when multiple, concurrent failure events occur. In this case, shared mesh restoration capacity may be claimed for more than one failed LSP and the protection path can be activated only for one of them (at most). For implementing shared mesh restoration, the identifier and nodal information related to signaling along the control path are as defined for 1+1 protection in Sections [3.2.1](#) and [3.2.2](#). In addition, each node must also keep (local) information needed to establish the data plane of the protection path. This information must indicate the local resources to be allocated, the fabric cross-connect to be established to activate the path, etc. The precise nature of this information would depend on the type of node and LSP (the GMPLS signaling draft describes different type of switches [GMPLS\_SIG]). It would also depend on whether the information is fine or coarse-grained. For example, fine-grained information would indicate pre-selection of all details pertaining to protection path activation, such as outgoing link, labels, etc. Coarse-grained information, on the other hand, would allow some details to be determined during

protection path activation. For example, protection resources may be pre-selected at the level of a TE link, while the selection of the

specific component link and label occurs during protection path activation.

While the coarser specification allows some flexibility in selection of the precise resource to activate, it also brings in more complexity in decision making and signaling during the time-critical restoration phase. Furthermore, the procedures for the assignment of bandwidth to protection paths must take into account the total resources in a TE link so that single-failure survivability requirements are satisfied.

### **[3.3.1](#) End-to-End Failure Indication and Acknowledgement**

The End-to-End failure indication and acknowledgement procedures and messages are as defined in Sections [3.2.3](#) and [3.2.4](#).

### **[3.3.2](#) End-to-End Switchover Request**

This message is generated by the source node receiving an indication of failure in an LSP. It is sent to the LSP destination along the protection path, and it identifies the LSP being restored. If any intermediate node is unable to establish cross-connects for the protection path, then it is desirable that no other node in the path establishes cross-connects for the path. This would allow shared mesh restoration paths to be efficiently utilized.

The End-to-End Switchover message must be sent reliably from the source to the destination of the LSP along the protection path.

### **[3.3.3](#) End-to-End Switchover Response**

This message is sent by the destination node receiving an End-to-End Switchover Request message towards the source of the LSP, along the protection path. This message should identify the LSP that is being switched over. Prior to activating the secondary bandwidth at each hop along the path, Extra Traffic (if used) must be dropped and not forwarded

This message must be transmitted in response to each End-to-End Switchover Request message received.

## **[4](#). Reversion and other Administrative Procedures**

Reversion refers to the process of moving an LSP back to the original working path after a failure is cleared and the path is repaired. Reversion applies both to local span and end-to-end path protected LSPs. Reversion is desired for the following reasons. First, the protection path may not be optimal as compared to the working path from a routing and resource consumption point of view. Second, moving an LSP to its working path allows the protection

resources to be used to protect other LSPs. Reversion has the disadvantage of causing a second service disruption. Use of

reversion is at the option of the operator. Reversion implies that a working path remains allocated to the LSP that was originally routed over it even after a failure. It is important to have mechanisms that allow reversion to be performed with minimal service disruption to the customer. This can be achieved using a "bridge-and-switch" approach (often referred to as "make-before-break").

The basic steps involved in bridge-and-switch are:

1. The source node commences the process by "bridging" the signal onto both the working and the protection paths (or links in the case of span protection).
2. Once the bridging process is complete, the source node sends a Bridge and Switch Request message to the destination, identifying the LSP and other information necessary to perform reversion. Upon receipt of this message, the destination selects the signal from the working path. At the same time, it bridges the transmitted signal onto both the working and protection paths.
3. The destination then sends a Bridge and Switch Response message to the source confirming the completion of the operation.
4. When the source receives this message, it switches to receive from the working path, and stops transmitting traffic on the protection path. The source then sends a Bridge and Switch Completed message to the destination confirming that the LSP has been reverted.
5. Upon receipt of this message, the destination stops transmitting along the protection path and de-activates the LSP along this path. The de-activation procedure should remove the cross-LSPs along the protection path (and frees the resources to be used for restoring other failures).

Administrative procedures other than reversion include the ability to force a switchover (from working to protect or vice versa), and locking out switchover, i.e., preventing an LSP from moving from working to protect administratively. These administrative conditions have to be supported by signaling.

## **5. Discussion**

### **5.1 LSP Priorities During Protection**

Under span protection, a failure event could affect more than one working link and there could be fewer protection links than the number of failed working links. Furthermore, a working link may contain multiple LSPs of varying priority. Under this scenario, a decision must be made as to which working links (and therefore LSPs)

should be protected. This decision may be based on LSP priorities.  
In general, a node might detect failures sequentially, i.e., all  
failed working links may not be detected simultaneously, but only



sequentially. In this case, as per the proposed signaling procedures, LSPs on a working link may be switched over to a given protection link, but another failure (of a working link carrying higher priority LSPs) may be detected soon afterwards. In this case, the new LSPs may bump the ones previously switched over the protection link.

In the case of end-to-end shared mesh restoration, priorities may be implemented for allocating shared link resources under multiple failure scenarios. As described in [Section 3.3](#), more than one LSP can claim shared resources under multiple failure scenarios. If such resources are first allocated to a lower priority LSP, they may have to be reclaimed and allocated to a higher priority LSP.

## **6. Author's Addresses**

Jonathan P. Lang  
email: [jplang@ieee.org](mailto:jplang@ieee.org)

Bala Rajagopalan  
Tellium, Inc.  
2 Crescent Place  
P.O. Box 901  
Oceanport, NJ 07757-0901  
email: [braja@tellium.com](mailto:braja@tellium.com)

## **7. Intellectual Property Considerations**

This section is taken from [Section 10.4 of \[RFC2026\]](#).

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in [BCP-11](#). Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.



## **8. References**

### **8.1 Normative References**

[BUNDLE] Kompella, K., Rekhter, Y. and Berger, L., "Link Bundling in MPLS Traffic Engineering", [draft-ietf-mpls-bundle-04.txt](#) (work in progress).

[GMPLS-ISIS] Kompella, K., Rekhter, Y., Banerjee, A. et al, "IS-IS Extensions in Support of Generalized MPLS", [draft-ietf-isis-gmpls-extensions-16.txt](#) (work in progress).

[GMPLS-OSPF] Kompella, K., Rekhter, Y., Banerjee, A. et al, "OSPF Extensions in Support of Generalized MPLS", [draft-ietf-ccamp-ospf-gmpls-extensions-09.txt](#) (work in progress).

[GMPLS-SIG] Ashwood-Smith, P., Banerjee, A., et al, "Generalized MPLS - Signaling Functional Description," [RFC 3471](#).

[LMP] Lang, P, ed., "Link Management Protocol (LMP) v1.0" Internet Draft, Work in progress, [draft-ietf-ccamp-lmp-09](#).

### **8.2 Informative References**

[RFC2026] Bradner, S., "The Internet Standards Process -- Revision 3," [BCP 9](#), [RFC 2026](#), October 1996.

[TERM] Mannie, E., Papadimitriou, D., ed., "Recovery (Protection Internet Draft, [draft-mannie-gmpls-recovery-terminology-02.txt](#), (work in progress).

## **Full Copyright Statement**

©Copyright © The Internet Society (date). All Rights Reserved. This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.



This document and the information contained herein is provided on an  
ôAS ISö basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK  
FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING  
BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION  
HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF  
MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.ö

This draft expires in March, 2004.