RSVP-TE Extensions For Shared-Mesh Restoration in Transport Networks

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026 [1].

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

Efficient techniques for rapid restoration must be addressed within GMPLS. This document describes extensions to RSVP-TE signaling in support of shared mesh restoration. Shared mesh restoration describes restoration plans in which restoration capacity is shared across multiple independent failures. In particular, this document proposes extensions enabling reservation of restoration capacity, LSP restoration, LSP reversion and LSP deletion.

1. Introduction

Rapid recovery (restoration) from network failures is a crucial aspect of current and future transport networks. Rapid restoration is required by transport network providers to support stringent Service Level Agreements (SLAs) that dictate high reliability and availability for customer connectivity.

The choice of a restoration policy is a tradeoff between network resource utilization (cost) and service interruption time. Clearly, minimized service interruption time is desirable, but schemes
achieving this usually do so at the expense of network resource utilization, resulting in increased cost to the provider. Different restoration schemes operate with different tradeoffs mainly among spare capacity requirements and service interruption time as well as complexity, robustness, etc.

In light of these tradeoffs, transport providers are expected to support a range of different service offerings, with a strong differentiating factor between these service offerings being service interruption time in the event of network failures. For example, a provider’s highest offered service level would generally ensure the most rapid recovery from network failures. However, such schemes (e.g., 1+1, 1:1 protection) generally use a large amount of spare restoration capacity, and are thus not cost effective for most customer applications. Significant reductions in spare capacity can be achieved by instead sharing this capacity across multiple independent failures.

GMPLS signaling proposals have primarily focused on the development of methods for label switched path (LSP) establishment and removal [1,2,3] with some fault recovery capabilities. Recent Internet drafts [4,9] examine how to realize some different restoration schemes for transport networks using GMPLS signaling. Other LSP restoration-related contributions [5,6,7,8,16,17,18] mainly focus on MPLS networks. This contribution motivates the need for path-based shared mesh restoration in transport networks, and defines extensions to support it. The proposal here primarily focuses on restoration within a single control domain.

Shared mesh restoration for transport networks was proposed in [13]. The basic functionality discussed within [13] is enabled using the GMPLS extensions proposed within this draft. Kini et al. have recently proposed shared mesh restoration for MPLS networks [5,6]. The fundamental difference between transport networks and packet networks (where MPLS applies) is that in packet networks we can establish an LSP without using any bandwidth. However, in transport networks, if an LSP is established, then by definition the full bandwidth requested by the LSP is consumed, independent of whether traffic is transmitted over this LSP or not. A LSP can be established before failure in MPLS, but not used until after failure, whereas this is not possible in transport networks. This contribution addresses the GMPLS-specific extensions required to support shared mesh restoration in transport networks.

The current GMPLS signaling specification is based on extensions to existing protocols â€” namely RSVP-TE [8] and CR-LDP [15]. The introduction of new signaling protocols for restoration [9] is likely to significantly complicate the standardization process and future implementations. Instead, we propose extending the existing signaling protocols to provide the necessary network failure restoration functionality. We demonstrated a reference implementation of the extensions to RSVP-TE described here for shared-mesh restoration in [10], and have successfully demonstrated...
that rapid end-to-end restoration signaling can be achieved using these extensions. Similar extensions are required for CR-LDP.

2. Restoration methods

We classify restoration techniques into path-based and link-based [16]. Path-based schemes are implemented via an alternate or backup path that may traverse multiple nodes. Failure recovery is typically provided on a per LSP basis between a pair of nodes. Different LSPs on a failed link, segment or path may use different restoration techniques and traverse different restoration routes. In contrast, link-based techniques are provided on a per link basis. Traffic on the failed link usually traverses on the same restoration route. Note that by "link" in this document we mean a "logical" link in the network layer of interest (e.g., one or more similar-routed channels between a pair of optical cross-connects).

In general, path-based schemes may protect an end-to-end path, a segment or a single link / node. The extensions proposed here are applicable to all of these cases, although we focus primarily on end-to-end path-based restoration. Depending on the degree to which a service provider wishes to protect LSPs, the service and restoration paths may be link-disjoint, node-disjoint or Shared Risk Link Group (SRLG)\[1,2,13\]-disjoint. SRLG-disjoint routes are important as they cover several common types of failure that must be protected against, including link failures, conduit cuts, etc.

There are a number of possible path-based restoration techniques for transport networks. The interested reader is referred to [16] for a complete taxonomy of MPLS-based restoration schemes. If the network pre-establishes a restoration path for a given service path, then restoration of the service path in the event of service path failure simply involves cross-connecting the add/drop ports at the source and destination from the failed path onto the restoration path. This is referred to as dedicated path protection. Dedicated path protection provides very rapid failure recovery, but is expensive in terms of the spare capacity requirements.

Alternatively, if the network searches for restoration capacity and establishes the restoration path only after service path failure, then the restoration scheme is referred to as dynamic restoration. Dynamic restoration may utilize techniques such as crankback [11] to successively try different paths until a path with sufficient resources is found. Dynamic restoration does not require pre-planning on a per LSP basis and as such may be more robust to (unanticipated) failures. The disadvantages of dynamic restoration schemes include long worst-case restoration times, lack of predictability and no guarantee of successful failure recovery. Dynamic restoration may be particularly useful as a backup restoration technique when other pre-established or pre-calculated restoration routes are not available (e.g., for multiple failure events in which insufficient restoration capacity has been established / reserved).
Another path-based restoration technique is instead based on pre-calculating restoration routes, with cross-connection performed after failure [10,12]. This approach allows efficient use of spare restoration capacity by sharing this capacity across multiple independent failures. In this scheme, when the service path for a LSP is established, resources may be reserved along the restoration path without allocating the resources to a specific LSP and configuring the cross-connects on the restoration path. The resources reserved for a particular restoration path can be shared with other restoration paths if their service paths do not have any (single) failure in common. In another words, if the service paths of two LSPs are failure disjoint, (e.g., they fail independently), the resources reserved for restoration can be shared on the common links of their restoration paths. We refer to this technique as shared mesh restoration. Note that for all-optical networks without wavelength conversion, restoration resources may have to be shared on a per-wavelength basis.

To implement shared mesh restoration, we require new extensions to the existing GMPLS signaling specifications [8,15] for bandwidth reservation, LSP restoration, LSP reversion and LSP deletion. These signaling procedures are discussed in the following section.

3. Shared mesh restoration

3.1 Resource reservation for restoration

A restorable LSP in a transport network supporting shared mesh restoration has both a service (primary) path and a restoration (secondary) path. During normal network operation (without failures), the LSP is established along the service path, with resources (optionally) reserved along the restoration path.

In implementing shared mesh restoration, capacity may be reserved along the restoration path during LSP provisioning [10,13]. The resources reserved on each link along a restoration path may be shared across different service LSPs that are not expected to fail simultaneously. The restoration capacity might be either idle or used for pre-emptable LSPs.

The amount of restoration capacity reserved on the restoration paths determines the robustness of the restoration scheme to failures. For example, a network operator may choose to reserve sufficient capacity to ensure that all shared mesh restorable LSPs can be recovered in the event of any single failure event (e.g., a conduit being cut). A network operator may instead reserve more or less capacity than that required to handle any single failure event, or may alternatively choose to reserve only a fixed pool independent of the number of LSPs requiring this capacity.

The sharing of restoration bandwidth across multiple independent failures can be simply illustrated using the topology depicted in
Figure 1. We consider an LSP established between A and C, and another between F and H. The service and restoration paths for the LSP between A and C are A-B-C and A-D-E-C, respectively, whilst the service and restoration paths for the LSP between F and H are F-G-H and F-D-E-H, respectively. Thus, the link between D and E has capacity reserved for the failure of both the service LSPs. If the service provider wishes to guarantee recovery from any single failure event, and if the links along the two service paths do not share any common failure (e.g., SRLG), then a single unit of capacity may be reserved on the D-E link for the restoration of either of the service LSPs. An example is provided in Section 6 that illustrates the reservation of restoration capacity when guaranteeing recovery from a single SRLG failure.

Figure 1. Example network topology.

When the amount of reserved capacity is a function of the number of LSPs that are to be restored on each link, signaling is required to reserve this capacity along the restoration path. Details of resource reservation are described in Section 4.1.

In general, depending on the network operator’s desired functionality, channel selection may be performed either during the reservation stage, or after failure. If channels are pre-selected, the channel selection is stored during the resource reservation phase as part of the reservation state along the LSP’s restoration path. Importantly, although the channels are pre-selected, the cross-connect is not established until after a failure. If channels are pre-selected during the reservation phase, then restoration message processing during restoration may be faster. However, if the pre-selected channels are dependent on the failure scenario, channel pre-selection may necessitate that fault isolation be performed before connectivity can be restored.

Alternatively, channel selection may be performed after failure on receipt of a signaling message for restoration. In this case, since restoration capacity along the restoration path is only reserved but not allocated, handling a fault translates into allocating the restoration LSP after failure. This requires efficient mechanisms for triggering and allocating the restoration LSP to meet the tight restoration timing constraints. The LSP restoration time will depend on the time to detect the failure, (possibly) localize the failure, notify the node(s) responsible for restoration, and finally activate the restoration LSP. Internet draft [16] shows a complete...
specification of the various cycle times involved in different recovery scenarios.

3.2 Interaction with failure detection and localization

Both failure detection and failure localization are technology and implementation dependent. In general, failures are detected by lower layer mechanisms (e.g., SONET/SDH, Loss-of-Light (LOL)). When a node detects a failure, an alarm may be passed up to a GMPLS entity, which will take appropriate action. This section discusses models for how failure detection interacts with and triggers end-to-end path-based restoration.

One model generates alarms upon failure detection and uses IP signaling to propagate a failure notification to the node(s) responsible for initiating restoration. Fault localization is important in this model to avoid having numerous alarms and IP messages generated for each failed LSP. Where hardware-based (e.g., SONET/SDH) fault localization techniques are not available, fault localization can be performed using IP-based protocols, such as the Link Management Protocol (LMP) [14]. Once the fault has been localized, the node(s) adjacent to the failure send a failure notification message to the node(s) responsible for restoring the failed LSP, which initiates restoration. In RSVP, the failure notification (NOTIFY) message is sent via normal IP forwarding with optional end-to-end reliable transmission.

Using this approach, restoration may be delayed due to the fact that failure localization needs to complete first. Additional delays may be incurred when sending failure notifications if normal IP routing has not converged. If the notification message is generated by a node downstream (upstream) of the failure and sent to a node upstream (downstream) of the failure, then normal IP forwarding may result in the message following a route that is broken as a result of the failure. The failure notification will thus not reach the node responsible for initiating restoration until IP routing has converged.

Another option is to trigger restoration based on failure detection at the nodes terminating the LSP. Failure localization is now targeted at the task of repairing the fault and becomes a background task that can be performed on a much slower time scale. However, it is important that valid signaling actions for planned events (e.g., LSP deletion) do not trigger failure notification and restoration actions along the path. For example, if LSPs are deleted in an all-optical network by sending a single deletion message, LOL resulting from disconnection at a node will propagate down the path faster than the LSP deletion message, potentially triggering restoration. Thus, for planned events that could result in LOL along the path, such as LSP deletion, all nodes must be informed of the upcoming event so that they may turn off alarms corresponding to the desired LSP so as not to initiate restoration.
For uni-directional LSPs, failures will be detected at the destination. For bi-directional LSPs, failures may be detected at either the source, the destination or both, depending on whether there is a uni-directional or bi-directional failure. Restoration should then be initiated by either the source, the destination or both. If restoration is initiated by the source (destination) and only the destination (source) detects the failure, then a failure notification must be propagated to the other end of the LSP. For all-optical networks, this failure notification may be done using IP messages, as above. However, most framing schemes in O-E-O networks will be capable of hardware level notification upstream of the failure, such as using SONET's Path AIS. Alternatively, restoration can be initiated by both the source and the destination, with restoration signaling meeting at an intermediate node along the pre-calculated restoration route.

All the above are potential implementations and therefore the extensions proposed herein are intended to work independent of the mechanism used for failure localization and notification.

4. Operations overview

The following discusses how shared-mesh restoration may be supported using extensions to RSVP-TE signaling.

4.1 Restoration path reservation

When a LSP requesting path-based restoration is established, the source node calculates the service and restoration paths for the LSP. To satisfy SLAs, the network may reserve resources along the chosen restoration path. To achieve this, the source node sends a PATH message along the restoration path with a new “shared reservation” flag (see Section 5.2) requesting a shared reservation along the path. The PATH message sent along the pre-calculated restoration path reserves the required restoration resources and establishes shared reservation state relating to the LSP without cross-connecting the channels (see the example in Section 6). A RESV message with the same flag is returned to acknowledge the resource reservation along the restoration path, but without establishing the restoration LSP.

In general, many carriers will want to protect their network against at least any single failure event, such as a fiber cut, or a conduit cut. If we generalize the SRLG concept, it may be used to represent different failure-prone network components, such as a fiber span, a node, a DWDM system or a conduit. Thus, for simplicity in the following description, we assume that we are protecting against SRLG failures.

The nodes along the restoration path need to know the path taken by the service LSP so that reservations can be shared among SRLG-disjoint failures along the service path. Thus, the PATH message sent along the restoration path includes information about the
service path. Two options for service path information are discussed in Section 5.3. The information can contain either a list of the links along the service LSP, or a list of the SRLGs traversed by the service LSP.

4.2 Restoration path setup operation

As described in Section 3.2, restoration path setup can be triggered in several ways. Path-based restoration may be triggered at either the source or destination node, or both [12].

If the restoration signaling is initiated by the source, the source node sends a PATH message along the restoration path with the "shared reservation" flag not set, indicating that the LSP should now be established. Since nodes along the path retained reservation state for the restoration LSP, this state can be used to ensure that restoration LSPs allocate resources out of the capacity reserved for restoration. Upon receipt of the PATH message, the nodes along the restoration path should check the cross-connect state for this LSP. (This is needed in case restoration triggered from the destination node has already performed the cross-connection.) If the cross-connection has not been performed for this LSP, the node should select channels for the LSP (if not already pre-selected), and perform the required cross-connections. In nodes with potentially slower cross-connect switching times (e.g., MEMS cross-connects) it is important to have the PATH message be forwarded without waiting for the cross-connection to be completed. The destination node sends a RESV message to the source to acknowledge the successful establishment of the restoration path.

If the signaling is initiated by the destination, then a RESV message is sent along the restoration path with the "shared reservation" flag not set. Upon receipt of the RESV message, the nodes along the restoration path should check the cross-connection states for this LSP. If the cross-connection has not been performed for this LSP, the node should select channels for the LSP (if not already pre-selected), and perform the required cross-connections. In nodes with potentially slower cross-connect switching times (e.g., MEMS cross-connects) it is important to have the RESV message forwarded without waiting for the cross-connection to be completed. The source node sends a RESV_CONF message to the destination to acknowledge the successful establishment of the restoration path.

If both ends initiate restoration, the PATH and RESV messages for the same LSP may meet at an intermediate node. This may result in label contention. For a uni-directional LSP, the contention is resolved using downstream label assignment. For a bi-directional LSP, the contention is resolved based on higher node-ID label assignment, as proposed for GMPLS [1,8]. When signaling messages from the two ends meet at an intermediate node, the node sends a RESV message to the source and RESV_CONF to the destination in response to the establishment of the restoration path.
When restoration is triggered from both source and destination, and PATH/RESV messages are forwarded without waiting for cross-connection as described above, the receipt of the RESV or RESV_CONF does not guarantee the success of restoration path establishment. In this case, a subsequent error message may override the acknowledgment. This behavior must be evaluated further.

One issue in establishing a restoration path using GMPLS LSP setup signaling is the contention resolution method. GMPLS allows upstream suggested label and resolves the contention via master/slave node relationship. During restoration process, two LSPs from different clients may be mis-connected when contention occurs. This may occur between two restoration LSPs or between a restoration LSP and a service LSP if they share the same label pool. One possible solution may be to only do label assignment from the master node, but this method may affect the restoration time. The detailed behavior must be evaluated further.

4.3 Error handling

In shared mesh restoration schemes, the reserved restoration resources may be limited. During a restoration path establishment, there may be scenarios in which the restoration path can’t be setup, for example, if there aren’t adequate reserved restoration resources due to any reason or if there is a failure along the restoration path. In this case, PATHERR and RESVERR messages may be used to report the failure of restoration path establishment. It is important that any resources allocated by the incomplete restoration path establishment be immediately released such that these resources can be used for other restoration paths.

In the RSVP-TE extensions proposed for GMPLS, the PATHERR message was extended to carry a ‘state_remove’ flag to release the resources consumed by incomplete LSP establishment. In shared mesh restoration schemes, we may borrow the same idea and define a new flag ‘allocation_remove’, which could be carried in both PATHERR and RESVERR messages. Upon receipt of PATHERR or RESVERR messages with this ‘allocation_remove’ flag, the node does not remove all local state but instead frees the cross-connect resources and releases the channels to the reserved capacity pool.

4.4 LSP reversion operation

After service path repair, most carriers prefer to cause the LSP to revert back to its original service path. Often, the routing of the restoration LSP may not be as efficient as the original service LSP. Additionally, once a restoration LSP is established, there is no guarantee that other service paths that were sharing its resources are protected, unless the other restoration routes are re-calculated.

Reverting back to the service path after a failure is repaired requires that the service LSP’s resources remain allocated during
the time that the LSP uses restoration resources. For RSVP, techniques must be developed that allow service path resources to remain allocated even though refreshes may be affected by failed signaling channels.

It is important to have mechanisms that allow LSP reversion to be performed without disrupting service to the customer. This can be achieved if LSP reversion is implemented using a "bridge and roll" approach. The source node commences the process by "bridging" the customer signal onto both the service and restoration paths. Once the bridge process has completed, the source node sends a Notification message to the destination, requesting that the destination "bridge and roll" the service and restoration paths. In this case, the "roll" function causes the destination to select the service path signal. Upon finishing the bridge and roll at the destination, the destination sends a Notification message to the source confirming the completion of the bridge and roll operation. When the source receives this Notification, it stops transmitting traffic along the restoration route, and sends another Notification message to the destination confirming that the LSP is reversed. Once the destination receives this Notification message, it issues a RESVTEAR message along the restoration path and stops transmitting along the restoration route. Additional mechanisms may be required in some cases (e.g., all-optical networks) to ensure that intermediate nodes do not alarm due to LOL during the teardown procedure (see Section 3.2). The RESVTEAR message informs the nodes along the restoration route to release the restoration resources if shared restoration is used for this LSP. This procedure achieves the "make-before-break" feature, that is, minimal service traffic interruption during the reversion process. Note that the RESVTEAR removes the cross-connection for the restoration path (and frees the resources to be used for restoring other failures), but does not delete the Path state along the restoration path. In this case, the RESVTEAR should not trigger a PATHTEAR message from the source since we want resources to continue to be reserved for this LSP. This allows the termination node to quickly re-establish the restoration path by sending either a RESV or PATH message if the service path fails again in the future. The protection object with "shared reservation" flag is carried in the RESVTEAR message to suppress the PATHTEAR. If the restoration paths are reoptimized periodically, the original restoration reservation state should be cleared and new restoration reservation state must be created.

4.5 LSP deletion operation

Once an LSP is no longer required, the LSP service path and its restoration resources should be released for future traffic. If the source node initiates the LSP deletion, it should send two PATHTEAR messages to the destination node: one along the service path and the other along the restoration path. The PATHTEAR along the restoration path should include information about the service path. The information can contain either a list of the links along the service LSP, or a list of the SRLGs traversed by the service LSP. If the
destination initiates the LSP deletion, it should send two RESVTEAR messages to the source. The RESVTEAR along the restoration path should include the information about the service path. Again, additional mechanisms may be required in some cases (e.g., all-optical networks) to ensure that nodes do not alarm due to LOL during the teardown procedure (see Section 3.2).

5. RSVP-TE restoration extensions

5.1 Current GMPLS fault restoration capabilities

The GMPLS signaling specifications [1] currently define protection information used in the LSP setup procedure. This protection information is carried in a new object/TLV that includes a bit flag that indicates whether the LSP is a primary (service) or a secondary (restoration) LSP.

GMPLS also specifies a Link Flags field in the protection information object. The Link Flags field indicates the link protection type desired by the LSP. If a particular type is requested, a new LSP request is processed only if the desired link protection type can be honored.

5.2 Shared reservation/allocation request

To implement restoration resource reservation for shared mesh restoration, a new mechanism must be introduced into PATH messages to distinguish between normal LSP establishment, reservation of shared resources, and allocation of shared resources to a particular LSP. The S (secondary) bit in the protection information object may be used to indicate that an LSP is a restoration/secondary path, not a service LSP.

The shared resource reservation and shared resource allocation can be explicitly indicated through a new Shared Reservation flag in the protection information object. The protection information object would be used in the PATH/RESV message forwarded along the restoration route during LSP resource reservation and resource allocation.

```
| S | R | Reserved | Link Flags |
+---+---+----------+------------|
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
```

Figure 3. Protection information object.

The Shared Reservation (R) flag described above may be encoded as follows:

```
0 allocation
1 reservation
```
If other flags are needed to support path-based restoration, the shared reservation flag can be included in a Path Flags field.

5.3 Service path information

To support shared reservations, intermediate nodes must compute the total resources that must be reserved to support service paths that are not subject to simultaneous failures. This requires identification of the specific failure events that are to be protected. If we wish to protect against link failures, then we must know the set of links used along the service path when reserving capacity on the restoration path. Alternatively, if we wish to protect (more generally) against SRLG failures, when a restoration LSP is reserved, the setup message must convey information about the SRLGs that are associated with the service LSP that it is protecting. Since a single restoration channel on a common link of multiple restoration paths can be shared by non-simultaneous fiber span failures.

This information is communicated by introducing a new object, the service path information object, in the PATH message. We propose two alternatives for information that might be conveyed:

(1) LINK_LIST SERVICE_PATH INFORMATION object

The LINK_LIST SERVICE_PATH INFORMATION object denotes the set of TE links [2,2] that are used along the service path. This information can be used directly when restoration bandwidth reservation accounts for link failures only. If we account for SRLG failures in our restoration reservations, then the use of the LINK_LIST requires the nodes along the restoration path to map from links to SRLGs.

(2) SRLG_LIST SERVICE_PATH INFORMATION object

If we account for SRLG failures in the restoration reservations, then transmitting the list of links along the restoration route would require that every node duplicate the calculation of the associated set of SRLGs for the primary links. This calculation could instead be performed only at the source node, with the set of SRLGs then carried in the PATH message. We thus propose a SRLG_LIST SERVICE_PATH INFORMATION object.

The SRLG_LIST carries the list of SRLGs that are used by the service path. Each SRLG is defined as a 32-bit unsigned number [2,3]. In this SRLG list, the order of specific SRLGs is not significant.

The information carried in the SRLG_LIST would be:
The use of the SRLG_LIST is more straightforward and requires less processing at each node than the LINK_LIST. However, the LINK_LIST is more generic and, in some realistic topologies, may be significantly shorter.

5.4 Path message format

The new proposed format for the PATH message is:

```
(PATH Message) ::= (Common Header)
                [ (INTEGRITY) ]
                [ <MESSAGE_ID_ACK> | <MESSAGE_ID_NACK> ]
                [ <MESSAGE_ID> ]
                <SESSION>
                <RSVP_HOP>
                <TIME_VALUES>
                [ <EXPLICITE_ROUTE> ]
                <LABEL_REQUEST>
                [ <SERVICE_PATH_INFORMATION> ]
                [ <PROTECTION> ]
                [ <LABEL_SET> ]
                [ <SESSION_ATTRIBUTE> ]
                [ <NOTIFY_REQUEST> ]
                [ <POLICY_DATA> ]
                <sender descriptor>
```

Shared restoration resource reservation is done if and only if the PATH message includes the <SERVICE_PATH_INFORMATION> and the <PROTECTION> objects with S and R (shared reservation) bits set. Otherwise, the <SERVICE_PATH_INFORMATION> is ignored and message processing is performed as usual. Shared restoration resource allocation is done if and only if the PATH/RESV message includes the <PROTECTION> object with S bit set and the R bit not set.

5.5 LSP establishment after failure

When a service path fails, the restoration LSP should be established along the restoration path using the reserved restoration bandwidth on each link. The LSP establishment along the restoration path may be signaled from the source and/or the destination. A PATH message is sent from the source including the <PROTECTION> object with S bit set and the R bit not set, and/or a RESV message is sent from the
destination including the <PROTECTION> object with S bit set and the R bit not set.

5.6 LSP reversion extension

It is proposed that LSP reversion be handled using the RSVP Notification message. The NOTIFICATION message should be extended to include a status field describing each of the different steps in the reversion process. The NOTIFY message includes the <ERROR_SPEC> object, which has four fields: node address, flags, code, and value. The node address represents the address of the node generating the notification. New codes/values in the <ERROR_SPEC> object could be reserved to support reversion. Three new codes/values are needed:

+ Bridging completed
+ Roll/bridge completed
+ Roll completed

5.7 Deletion extension

A PATHTEAR message or RESVTEAR message as defined in the GMPLS signaling specification [8] is used to remove (de-allocate) the service path. Additional mechanisms required to ensure that nodes do not alarm due to LOL during the teardown procedure are being developed for some network applications such as all-optical networks. Once a restoration LSP is no longer required, we must also release the reserved restoration resources and any allocated resources along the restoration path. To achieve this, the source sends a PATHTEAR message along the restoration path, including the <SERVICE_PATH_INFORMATION> object. Upon receipt of this message, each node along the restoration path should de-allocate any resources allocated to this LSP (e.g., if the LSP is currently using the restoration path) and decrement the reserved resources accordingly.

The new proposed format for the PATHTEAR message is:

```
<PATHTEAR Message> ::= <Common Header> 
[<INTEGRITY>] 
[ <MESSAGE_ID_ACK> | <MESSAGE_ID_NACK> ] 
[ <MESSAGE_ID> ] 
<SESSION> 
<RSPV_HOP> 
[ <EXPLICIT_ROUTE> ] 
[ <SERVICE_PATH_INFORMATION> ] 
[ <PROTECTION> ] 
[ <SESSION_ATTRIBUTE> ] 
<sender descriptor>
```

6. Example

We illustrate here how the above RSVP signaling messages can be used to implement resource reservation for shared mesh restoration in a network that aims to guarantee recovery from any single SRLG
failure. We also assume here that channels are selected after failure and that full wavelength conversion capabilities exist if we are considering an all-optical network.

With the GMPLS routing enhancements [2,3], each node will have a representation of the transport network topology, including the available bandwidth, and the list of SRLGs for each optical link.

When a new LSP request arrives in the network, the source node is responsible for computing two SRLG diverse paths. An RSVP PATH message is sent along the calculated service path to establish the service LSP. An RSVP PATH message containing a Protection information object with the S and R (shared reservation) bits set should also be forwarded along the restoration path with information that identifies the SRLGs of the service path. This information may be conveyed using either the LINK_LIST or the SRLG_LIST. Upon receipt of this message, each node should then update the restoration bandwidth reserved on the outgoing links of the restoration path. Assume that each link has a Reservation array \( R[i] \), \( i=1,2,\ldots,K \), where \( K \) is the maximum SRLG index. There are various techniques on how these arrays for each link can be maintained among the nodes. These methods are not specified here. \( R[i] \) indicates the bandwidth required on the link if the \( i \)-th SRLG in the network fails. The total reserved restoration capacity should be calculated as the maximum over all SRLGs (i.e., \( \max R[i] \), \( i=1,2,\ldots,K \)). When a node receives a new reservation message, it saves state relating to the LSP and updates the Reservation array on its link(s) in the following way: \( R[i] = R[i] + \text{reservation bandwidth} \) if the \( i \)-th SRLG is in the SRLGs associated with the \(<\text{SERVICE_PATH_INFORMATION}> \) object. Once \( R[i] \) has been re-calculated for all SRLGs associated with the service path, a new required reserved capacity is calculated (i.e., \( \max R[i]=1,2,\ldots,K \)). If inadequate capacity is available to support this new resource reservation, the LSP reservation process may be abandoned, with an error message (PATHERR) being returned to the source. The already reserved resources must then be removed. However, if the reservation is successful and the reserved capacity has changed as a result of this new LSP, then updated link resource information may be flooded to other nodes in the network for the purpose of path computation. For example, the reserved capacity may reduce the available bandwidth information that is flooded. If the GMPLS routing extensions were further extended to explicitly flood the bandwidth reserved on each link, some additional improvement in network utilization may be possible.

Similarly, when a node receives a message requesting the removal of reservations for an existing restoration LSP, the restoration capacity is updated for each of the SRLGs along the primary path: \( R[i] = R[i] - \text{reservation bandwidth} \) if the \( i \)-th SRLG is in the set of SRLGs along the service path. Again, this update may result in a change in the link information that is flooded throughout the network.
7. Discussion

7.1 Interaction with other restoration schemes

An operational transport network is expected to support multiple restoration schemes to satisfy different clients requirements. For example, a service provider may offer four different services based on dedicated protection (1:1, 1+1), shared mesh restoration, dynamic restoration, and no restoration. In fact, our shared mesh restoration can co-exist with dedicated protection, dynamic restoration, and other restoration schemes. Our RSVP-TE extensions can also be re-used for these schemes. For example, restoration path reversion messages and procedures can be used for 1:1 protection whilst dynamic restoration can re-use the restoration path creation message for purposes of bandwidth accounting, path reversion, and deletion.

7.2 Multi-domain restoration

This contribution focuses on shared mesh restoration within a single control domain, area, or sub-network. Realistically, each domain may implement different restoration schemes. If a LSP is routed over multiple domains, domain-by-domain restoration may be applied to recover from failures internal to each domain. External links between domains may be protected via link protection (e.g., 1:1 or 1+1 protection). In this way, the shared mesh restoration procedures proposed here are able to interoperate with other protection schemes crossing network-to-network interfaces.

Alternatively, the shared mesh restoration procedure proposed here may also be executed across multiple domains.

7.3 Restoration priority and pre-emption

The shared mesh restoration extensions proposed within this draft can support restoration priority and pre-emption using setup priority and holding priority. Our restoration messages are extended from RSVP-TE provisioning messages and inherit the pre-emption functionalities.

8. Security considerations

This draft introduces no new security considerations to [1,8].

9. References


10. Author’s Addresses

Guangzhi Li                           Charles Kalmanek
AT&T                                  AT&T
180 Park avenue                       180 Park avenue
Florham Park, NJ 07932                Florham park, NJ 07932
973-360-7376                          973-360-8720
gli@research.att.com                  crk@research.att.com