Restoration Mechanisms and Signaling in Optical Networks

<draft-many-optical-restoration-00.txt>

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.

Conventions used in this document:

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",

Abstract

With the advent of tunable lasers and optical switches, dynamic configuration of optical mesh networks has become possible. This Internet-Draft presents a signaling mechanism for provisioning schemes and dynamic lightpath restoration. The mechanisms proposed cover shared restoration (i.e. 1:1, 1:N and M:N) as well as provisioned and non-provisioned scheme. Provisioned restoration (equivalent to lightpath protection) implies the provisioning of backup lightpaths in advance before failures occur.

1. Introduction

In general, Internet backbone networks are overbuilt in comparison to average traffic volumes, in order to support fluctuations in traffic levels and to stay ahead of traffic growth rates. Therefore, any given time there are typically under utilized capacity in deployed optical network facilities. Overbuilding of the backbone to support traffic fluctuations is part of the optical network design to meet the expected performance and the QoS provided.

One of the most important concepts in network management is maintaining the survivability of networks and in particular the survivability of optical networks. When there are link failures or the like, any affected routes should be repaired as soon as possible. This mechanism is also referred to as fast restoration. In today’s SONET/SDH networks, one can achieve recovery times of 50 ms or even lower, but this mechanism depends on the 1+1 backup optical network link resources to achieve this performance. To avoid this resources waste in optical networks, an adaptive 1:N or M:N shared restoration mechanism can be applied that still provides enhanced network survivability while minimizing the waste in network resources. When using these shared mechanisms, the protection paths can be used to transport best-effort traffic throughout the optical network.

This Internet Draft proposes a mechanism for route computation and restoration processes that meet these goals in optical meshed networks, as well as the associated signaling extensions and notification message definitions.

Note: in order to avoid the use of RSVP-TE and CR-LDP message in the first four sections of this memo, a generic terminology described in Appendix 1.

2. Basic concepts for bandwidth provisioning and restoration
This section describes the basic concept of bandwidth provisioning and restoration mechanisms in optical networks.

2.1 OXC system architecture

The Optical Cross-Connect (OXC) system architecture for lambda switching has been introduced in the [IPO-FRAME] and [IPO-MPLS]. The OXC interfaces considered hereafter includes, as described in [GMPLS-SIG], Lambda Switch Capable (LSC) interfaces that switches the lightpath segment based on the incoming wavelength and Fiber-Switch Capable (FSC) interfaces that switches the lightpath based on the spatial position of the incoming data stream in the real world physical spaces. Our proposed bandwidth provisioning and restoration mechanisms are based on this architecture.

The internal processing in OXC system for lambda switching (i.e. wavelength switching) is briefly outlined hereafter. This description is presented only as an aid to understanding the content of the next sections.

An OXC has several incoming and outgoing LSC interfaces or ports, connected to adjacent OXCs, and several incoming and outgoing LSC interfaces or ports attached to an edge device that can be a router (or any other termination capable device supporting LSC interface connectivity see [IPO-MPLS]). An OXC includes mainly two functional parts: an OXC Switch Controller (OSCtrl) and OXC optical matrix.

The OSCtrl communicates through Optical Signalling Channels (OSC) i.e. out-of-band signaling transport mechanism or through dedicated and physically diverse control network i.e. out-of-network signaling transport mechanism as indicated in Figure 1. OSCtrls and OSCs together define the signalling plane of the optical network.

(Fig.1) OSC Channel, OSC Controller and OXC
The OSCtrl converts the received messages (mentioned in section 5) to the proper control command, and sends this command to the OXC optical matrix. The commands to control the OXC optical matrix are as follows: connect (and disconnect) between an incoming LSC interface and an outgoing LSC interface. The OSCtrl is also capable to process status requests, lightpath modification requests as well as notification messages. The same commands apply when the OXC is connected to an LSC capable edge device. However, within the optical transport plane, lightpaths are strictly defined between the ingress and the egress OXC LSC interfaces.

Based on these commands, a chain of connections through OXCs can form an end-to-end lightpath. The OXC initiating the lightpath creation process is called the ingress OXC, and the OXC ending a lightpath is called the egress OXC. The OXCs inside the lightpath are called the intermediate OXCs.

An OXC optical matrix receives commands from the OSCtrl, and replies whether the command was successful or not. The OSCtrl then converts the result into a message (described in section 5) that it sends via the OSC channel throughout the optical network.

(Fig.2) OXC system architecture

To cover any kind of optical network, we consider as specified in [IPO-FRAME] the following distinction between Optical Cross-Connect (OXC) in Transparent optical networks and Photonic Cross-Connect (PXC) in All-optical networks. Basically, OXC devices included
within Transparent optical networks performs optical/electrical/optical (O/E/O) conversion at each of their interfaces. Conversely, PXC devices included within All-optical networks do not perform O/E/O conversion at all so they are defined as pure O-O devices (including for most of them tunable lasers). As described in [GMPLS-SIG], we consider here that all the interfaces of a PXC are Lambda-Switch Capable (LSC) as well as all the interfaces of an OXC are Lambda Switch Capable.

2.2 Control channel between OXC switching controllers

The OSCtrl communicates through Optical Signalling (i.e. control) Channels (OSC) i.e. in-fiber/out-of-band signaling transport mechanism or dedicated physically diverse control network or through out-of-network signaling transport mechanism. OSC signalling transport mechanism is described in [IPO-ONNI].

We assume here that control channels between OSCtrl’s must be continuously available in order to enable the exchange of signaling information. The details of the signalling fault tolerance mechanisms required are detailed in sub-section 4.3.2 (but limited here to the scope of notification mechanisms).

2.3 Gathering of link state information for lightpath computation

Lightpath route computation (and selection) can be computed based on dynamic flooding of information related to optical network link resources, by using link state IGP protocol advertisements. Many mechanisms for exchanging the link state information have been introduced by extending existing IGP routing protocol such as OSPF [GMPLS-OSPF] or IS-IS [GMPLS-ISIS] for traffic-engineering purposes in optical network environment. This memo assumes that the link state information exchange mechanisms and eventually the bundling (or link bundling) mechanisms enabling the grouping of several links between adjacent OXCs is as described in [GMPLS-BUNDLE]. So, when a link state routing protocol is used, these links can be bundled together and be advertised in a single LSA.

The minimum link state information required to select lightpaths can be described as follows:
- the link identifier, status and the corresponding signal encoding
- the link bandwidth (minimum/maximum reservable bandwidth and available bandwidth)
- optionally the link protection type (unprotected 0:1, dedicated 1:1, shared protection M:N and dedicated 1+1)

Any ingress OXC has to know the SRLG (Shared Risk Link Group) list of every link to compute lightpath routes that do not share the same risk of potential damages. In the steady state, these SRLG values do not change, so this information is exchanged only once (at the initial optical network setup time). As described in [IPO-SRLG], the routing diversity constraint implies specific considerations with respect to the optical network topology.
In order to know whether it is possible to switch an incoming wavelength to a different outgoing wavelength, the ingress OXC could also need to know the availability of lambda converters (i.e. tunable lasers) in all OXCs.

Currently, we assume here that all OXCs have enough lambda converters for that purpose. If this assumption is not verified, then the switching process depends on the availability of the wavelength requested for a given lightpath. In this case, the wavelength blocking probability must be taken into account during the setup and the restoration process. This may lead to specific considerations like crank-back mechanism during lightpath setup process.

2.4 Lightpath Creation Process

It is assumed that ingress OXC (or boundary intermediate OXC see [IPO-ONNI] and [IPO_FRAME]) functions will perform lightpath route computation and selection. However, in this Draft, we only consider the signaling procedures after the route (or the partial) of the lightpath (of the lightpath segment, respectively) is computed by a distributed CSPF algorithm. Since the signaling mechanisms are independent on the route computation and selection, the restoration procedures described here are quite general and apply to distributed optical network environments.

In general, the establishment of an LSP by using GMPLS signaling [GMPLS-SIG] can be divided into two cases:

- signalling-driven approach: pre-establishment before data traffic arrives at the ingress router in response to an explicit signalling request
- data-driven approach: post-establishment after data traffic arrives at the ingress router

In the latter case, LSPs must be created quickly enough to handle the data traffic waiting at the ingress Label Switch Router (LSR) for transmission. However, the signalling-driven approach is the one considered here and we only refer to the data-driven approach for completeness.

Nevertheless, even with data-driven approach, one might expect that the lightpath creation process will not in general be so intensive, as lightpath will only be created in response to long-term changes in traffic volume (note: long-term comparing to the time scale of the optical restoration one). Since several LSPs are generally multiplexed within the same lightpath, the granularity of the lightpath is coarser than the one of the LSPs it carriers.

Moreover, since we assume that the permanent state [IPO-FRAME] is supported for the lightpath, the Bandwidth-on-Demand (BoD) model support will require high availability of the signaling plane (i.e. of the control channels). Consequently, some specific features such
as signaling fault-tolerance and signaling performance (i.e. signaling delay) are required when considering the lightpath create procedures.

3. Optical Restoration Mechanisms

The ingress OXC (or intermediate OXC) is in charge of the creation, modification, deletion, and subsequently the restoration (or partial restoration) of lightpaths. This means that the ingress OXC will also assign the restoration-related parameters during the lightpath creation phase.

First we compare the link protection to the link restoration mechanisms and analyze in detail the lightpath restoration procedures and the corresponding signalling enhancements in the remaining part of this document.

3.1 Link Protection and Restoration vs. Lightpath Restoration

If we compare the optical restoration to the link (1+1) protection mechanism between OXC, then the following protection mechanism can be considered. For each of the link included in the primary lightpath route that needs to be protected, a dedicated link-disjoint backup route is setup in advance during the lightpath creation process. Switching from the primary to the backup route is executed when a link failure is detected.

The link restoration mechanism is based on distributed recovery process. When a link failure is detected by the node(s) adjacent to the failure, the adjacent upstream node dynamically computes a route (for instance, from the adjacent upstream node to the adjacent downstream node) for each damaged lightpath traversing that failed link. Upon backup lightpaths establishment the upstream adjacent node and downstream adjacent nodes to the failure switch the failed primary lightpath to the corresponding backup lightpath.

The lightpath restoration mechanism is based on the following edge-to-edge principle. When a link failure is detected by the node(s) adjacent to the link failure (or the node failure), they notify the ingress and egress nodes of each of the lightpaths traversing that failed link (or failed node). Then depending on the restoration scheme, the ingress and egress nodes compute the route of an end-to-end backup lightpath or use the provisioned end-to-end backup lightpath. When the backup lightpath is established the ingress and egress nodes switch from the failed primary lightpath to the backup lightpath.

However, as described in the section 4 (see below), with non-revertive strategy, the provisioned restoration (which is equivalent to lightpath protection) operation is a particular case of the lightpath destructive modification procedure. On the other hand, the non-provisioned restoration is basically a triggered lightpath creation procedure.
3.2 Lightpath Failure Detection

If an optical link is damaged physically, all the lightpaths passing through this link will be affected. Generally, in case of an all-optical network (i.e. transparent optical network) where all devices do not use Optical/Electrical/ Optical (O/E/O) conversion, damage to a lightpath or drop of light signal level depends on the Loss of Light (LoL) detection capability of the PXC. For that purpose the AND function of the LoLs on all terminated wavelengths channels can be used to detect the loss of the physical section (more precisely, the wavelengths changing device which can be based on an adaptation of the laser/detector). In practice, a LoL detector will be essential to provide at the end of the physical optical section in order to provide fault detection and localization.

If the PXC supports this feature, a bi-directional LoL detection will be propagated from the source of the failure to the ingress and egress PXC by a Notification Message. Otherwise, if the PXC doesn’t support LoL detection, this lost of signal will be triggered by the ingress OXC or the egress OXC depending on the transmission direction since the devices support O/E/O conversion at their LSC interfaces. The unidirectional LoL detection mechanism is detailed in section 4.3.

Link failure detection used in all-optical networks can be compared with the current mechanisms used in SDH/Sonet networks:
- **LoF (Lost of Frame)** is detected by monitoring A1 and A2 RS/Section overhead bytes. LoF is declared when valid values for these bytes are not received for a 3-ms time period. LoF is cleared when two consecutive valid A1 and A2 bytes are received.
- **LoS (Lost of Signal)** occurs when an all zeroes pattern is received for 10 Å’s or longer (indicating an upstream transmit failure).
- **AIS (Alarm Indication Signal)** occurs as a result of a failure condition such as LoF or LoS and is used to notify downstream nodes that a failure has occurred. Line AIS is detected by 1’s in bits 6 - 8 of the K2 byte in five consecutive frames. AIS performs two functions: inform the intermediate nodes that a failure has been detected and notify that the connection end-point that the service is no longer available. Note that some proprietary implementations also provides to localization (i.e. the originating point) of the failure.
- **BER (Bit error rate)**: BER monitoring is handled by counting parity violations detected via the B2 MS/Line overhead byte or B1 RS/Section overhead byte and converting the number of violations over a time period into the corresponding bit error rate. The configured BER threshold is used to determine if the observed BER results in a signal failure (SF) or signal degradation (SD) condition.

Note here that LoF, LoS and AIS refer to Inherent path state monitoring and BER refers to Inherent and Non-intrusive path state monitoring.
In all-optical networks, LoS (or LoL) is directly applicable to optical link failure detection. BER is more complicated to achieve in all-optical networks since they include only PXC’s. Since the BER real-time measurement in all-optical networks is currently under definition, these considerations are left for further study.

Note also that the failure detection mechanisms considered in this memo are bit rate independent and transparent (to the client signal).

3.3 Strict vs. Loose explicit routing

In CR-LDP [MPLS-CRLDP], both strict explicit routing and loose explicit routing mechanisms may be used. However, in this scheme for lambda switching within a given IGP area, we only consider strict explicit routing to designate the route of the lightpath. By using strict explicit routing, optical network resources can be managed more precisely, and the rate of success for lightpath creation can be enhanced. Consequently, we only consider strict explicit routing for intra-area routing.

For inter-area routing, the knowledge of the topology of the neighboring areas depends on the information advertised through the IGP protocol (we consider here inter-area summary advertisements as described in [MPLS-TE]). So, we assume that the explicit route computation and selection is performed at the ingress OXC (default behavior) and explicitly at each border OXC. This means that at each border OXC (i.e. at the ingress of each sub-network or area), the explicit route for the lightpath segment within this area needs to be computed (or selected).

3.4 Decision of backup lightpath

The difference between a primary lightpath and backup lightpath is that at the edge of the optical network, the backup lightpath has no connections within the optical matrix between the input and the output LSC interface of the ingress OXC and egress OXC, whereas the primary lightpath has connection in both OXCs. The connections between these LSC interfaces are performed during the restoration process.

However, the backup lightpath resources used within the optical network could be used to transport best-effort traffic. In this case, the edge (ingress and egress) port connections for the backup lightpath are assigned to interfaces excluded from those allocated to the corresponding restoration group. The corresponding bandwidth resource is then considered as available but advertised with a lower priority. This is the more efficient way to use the optical network resources.

3.4.1 Restoration schemes

In our approach of the lightpath restoration, we consider two kinds of lightpath restoration schemes: provisioned or non-provisioned.
Provisioned restoration of lightpaths is also referred to as lightpath protection.

- Either provisioned: meaning that backup lightpaths are established in advance of a link failure. The method of pre-establishment of backup lightpaths in advance of link failure minimizes the restoration time, especially when the restoration process must be carried out simultaneously for a large number of damaged lightpaths sharing the same damaged link. In order to maximize the use of the optical network resources, the wavelengths allocated to the backup lightpaths could be used for carrying preemptable best-effort, low-priority traffic and this only during failure-free optical network operating conditions.

- Or non-provisioned with a maximum restoration time value: meaning that backup lightpaths are dynamically created when a link failure is detected or advertised. The maximum restoration time is related to the signaling delay needed to create the backup lightpath through the optical network. This implies the need to define an ultra fast restoration mechanism (we assume here a low latency signalling network).

This non-provisioned approach does not preclude the background pre-computing of the backup route for these paths (also referred to as non-static routing). However, in this case, some refreshment mechanism of the pre-calculated backup routes is needed in order to ensure that these routes are still "usable" regarding the topological and resource information advertised by the IGP protocol. The (non-)static vs dynamic route computation is described in section 3.4.

3.4.2 Dedicated vs. Shared Restoration

The following dedicated and shared restoration mechanisms for meshed optical networks are considered:
- Dedicated 1:1 where one primary lightpath shares one backup lightpath
- Group Shared 1:N where N primary lightpaths share one backup lightpath (still dedicated to the designated to a unique restoration group-N)
- Group Shared M:N (with N > M and M >= 2) where N primary lightpaths share M backup lightpaths (still dedicated to the designated to a unique restoration group-N)
- Multi-group Group Shared 1:N where N primary lightpaths share one backup lightpath (here the protection lightpath is shared between P restoration group-N)
- Multi-group Shared M:N (with N > M and M >= 2) where N primary lightpaths share M backup lightpaths (here the protection lightpaths are shared between P restoration group-N)

Since 1:1 and 1:N are particular case of the M:N restoration scheme we only consider this last case for the seek of generality of this draft. The size of N will depend on the topological characteristics of the network and the size of M depends on the restoration level
supported by the optical network. We say that the N primary lightpath and M backup lightpath share the same restoration group. A given restoration group is identified by a restoration identifier defined by the ingress OXC.

Note that the weakness of the 1:N restoration mechanism is that, after an initial failure and before re-provisioning, it cannot support restoration for additional failures to other lightpaths in the same restoration group. This defect is remedied through the general M:N restoration mechanism (M &geq; 2) which the simultaneous restoration of maximum M primary lightpaths.

To decide on the grouping of primary lightpaths and backup lightpaths, the SRLG values of links are considered. Because a lightpath travels through several links, it will have corresponding to it the set of SRLG values of its member links.

The lightpaths within a 1:N restoration group cannot share the same SRLG value (and subsequently set of SRLG value, since SRLG sets must be disjoint), as then a single failure could affect several lightpaths simultaneously. Therefore, when a new lightpath is created, if it cannot be placed within any existing restoration group, a new restoration group is created, along with a backup lightpath for the newly created primary lightpath.

Moreover, since SRLG sets must be completely disjoint more than M lightpaths within a M:N restoration group cannot share the same set of SRLG values since a single failure could affect more than M primary lightpath simultaneously.

3.4 Static vs. Dynamic Routing

The Optical Restoration Diameter (ORD) defines the Optical Restoration Scope (ORS) or domain. Optical Restoration Diameter is defined as the number of nodes (i.e. number OXC's and PXC's) in the routing path (i.e. node sequence) for which a given restoration mechanism and policy is applied providing the static or dynamic routing.

Regarding the lightpath restoration policy (provisioned or non-provisioned) and depending on the restoration diameter, two route computation schemes can be considered: static and dynamic routing.

- Static routing: for every end-to-end lightpath between border OXC, the routing path are predetermined through the application of Constraint-based SPF (CSPF) algorithm and does not change depending on the connection in progress in the network. Different routing constraints can be used to compute the routing path for every lightpath. Some of these constraints are:

  . to minimize the transport-plane delay
  . to minimize the signalling delay
  . to minimize the link and wavelength usage
  . to maximize the bandwidth usage
  . to optimize other network performance characteristics
  . to optimize the routing diversity between lightpath belonging to
the same restoration group

Dynamic routing: for any given end-to-end lightpath, the routing path to be used is determined when a connection between boundary OXCs is requested. Using a dynamic routing scheme, a lightpath is not predetermined and depends on the routing path for other connection in progress in the network. A dynamic scheme needs a setup time (i.e. a maximum restoration time) to determine the routing path such that the lightpath satisfies the routing constraints (delay, bandwidth and performance). From the link and wavelength usage at any instant, the lightpath is constructed dynamically. Consequently, the OXCs are not programmed initially and set to establish the lightpaths when such requests occurs.

When using static routing scheme, if faults occur to the nodes, links or channels, one or more predetermined lightpaths will become unusable. Consequently, since fault occurrence is inherently a dynamic network characteristic, achieving fault-tolerance in a network using static routing is more difficult than in a network using dynamic routing. The failure avoidance mechanism needed with static routing as well as specific dynamic routing mechanisms are detailed in the remaining document.

3.5 Partial Restoration

When the ORD does not include any node traversed by a lightpath (from the ingress to the egress OXC) but only some partial segment of the routing path, then the restoration mechanism applied for this lightpath is only partial.

Consequently, a restoration path has to be provided for each link or sequence of links covered by the lightpath and not only from the ingress to the egress OXC, source and destination of this lightpath. Depending on the routing plan (i.e. static or dynamic), partial restoration can be applied for instance per IGP area. In this case, the lightpath restoration procedure only applies between ABR belonging to the area experiencing the lightpath failure.

If the lightpath traverse multiple areas, partial restoration can independently handle simultaneous recovery of single failure per area. At the ABR, in case of simultaneous recovery, the synchronization between the ingress and the egress port switching connection is guaranteed by the bi-directionality of the lightpath restoration process.

Partial restoration offers the advantage of lightpath failure isolation per area (or any other logical structure) as well as a greater flexibility in the backup routing path selection. However, the main advantage of partial restoration is that this technique enables the partitioning of the restoration domains; this in turn limits the probability of simultaneous failures within the mean restoration time.

Partial restoration can also be applied for link failure restoration and more generally to node failure. In this case, one could expect
that each node is capable to compute the lightpath segment of detour [IPO-FT] between the upstream and the downstream node adjacent the link failure detection. This restoration mechanism is sometimes referred to as ‘local re-routing’.

3.6 Release of damaged lightpath and Restoration Strategies

Depending on the lightpath restoration strategy, a damaged lightpath could be released (non-revertive restoration strategy) or repaired (revertive restoration strategy):

- If we consider a non-revertive strategy, then the damaged lightpath must be released in order to make the resources of its undamaged segments available for future demands. Since a link failure generates a Notification Message in both ingress and egress OXC direction, the release of the corresponding lightpath resources is by default the ingress OXC’s responsibility. However, as explained in section 4.3, in order to decrease the restoration time from the primary to the backup path, the switching is performed ‘simultaneously’ at both ends. The released resources can be reused after announcement of the corresponding released status through IGP link-state advertisements.

- If we consider a revertive strategy, then the resources used by the damaged lightpaths must not be released in order to be re-used when the failed resource will be available again. By receiving the Notification Message, any OXC included within the route of the failed lightpath reserves the corresponding resources. Note that a specific time-out could be locally defined to release the resources if the failed resource status does not change within a certain period of time.

Note that by referring to MPLS specification [MPLS-ARCH] (and to the GMPLS signalling specification [GMPLS-SIG]), a non-revertive restoration strategy refers to a liberal (generalized) label retention mode and a revertive restoration to a conservative (generalized) label retention mode.

3.7 Failure of lightpath creation

The creation of lightpaths can fail for the following three reasons. First, failure can occur due to an inconsistency between the gathered link state information at the ingress OXC and the actual link status of optical network. Because link state information can be transmitted with some delay from every OXC, delayed updates can make for this inconsistency.

Secondly, failure can occur even if link state information is consistent with the actual link status. If two ingress OXCs decide simultaneously to use the same network resources, one of the lightpath setup attempts will fail because its resources have already been claimed by the other lightpath. However, contention resolution procedures are described in [GMPLS-SIG] and are not further considered within this document.
Finally, even if lightpath establishment completes, the lightpath may fail to carry data because the quality of transmission does not reach a minimal threshold level. The deterioration of transmission quality can occur due to several reasons as described in [IPO-CTRL] and [IPO-LMP].

The first two cases of failure during lightpath creation are reported to the ingress OXC by a Notification Message from the intermediate OXCs where the error is detected. Failure can be detected by measuring the Optical Signal to Noise Ratio (OSNR) of the received optical signal or the Bit Error Rate (BER) of Optical-to-Electric transformed data. In the all-optical domain, a new BER measurement has been proposed in [IPO-OPM], this technique performs sequential scanning of the input power of the optical signal after splitting (more accurately by using taps). If the quality of the optical signal does not reach the threshold level, then the lightpath is released, and a new lightpath is created.

4. Restoration Procedures

In this section, we detail the optical restoration procedures and mechanisms. More precisely, we analyze the setup procedures for primary and backup lightpaths, the reporting procedure for handling damaged lightpaths, the restoration procedure for replacing a damaged lightpath with one of the backup lightpaths included within the backup group, and the release procedure for unused lightpaths.

4.2 Decision of lightpath route based on link state information

If a primary lightpath cannot be created within an already existing restoration group, then a new restoration group needs to be defined. So, the primary lightpath and the corresponding backup lightpath are created within the new restoration group defined by a new restoration identifier.

If the primary lightpath can be created within an existing restoration group, then if the number of backup lightpath already created for this restoration group is less than M then a new backup lightpath is created. Otherwise, if the number of backup lightpath is greater than M, no new backup lightpath is created. However, if the number of established primary lightpath does not exceed the number of existing backup lightpath, no new backup lightpath is created.

4.2.1 Lightpath setup procedure for primary lightpath

Once the route of a lightpath is decided, the ingress OXC sends the Lightpath Create Request Message to the next intermediate OXC through the control channel. The sequence of intermediate OXCs to be followed by the lightpath is included within the explicit route field of the Lightpath Create Request Message; and at each OXC the wavelength to be allocated for the lightpath is deduced from the
An intermediate OXC receives the Lightpath Create Request Message from its adjacent upstream OXC. The OSCtrl of the intermediate OXC then attempts to configure the OXC switch fabric according to the Lightpath Create Request Message, and receives the result from the OXC switch fabric. If a successful result is returned, the OSCtrl of the intermediate OXC sends the Lightpath Create Request Message to its downstream intermediate OXC for that route. During this procedure, intermediate OXCs consume the optical network resources and therefore update the link state database by indicating the corresponding resources as reserved meaning that these resources won’t be allocated to another restoration group. This link state information will then be flooded to other OXC by using IGP link-state advertisement.

If a Lightpath Create Request Message successfully reaches the egress OXC, the egress OXC then returns the Lightpath Create Response Message noting the successful lightpath creation through the reverse direction of the Lightpath Create Request Message direction to the ingress OXC. Upon receiving this response message, each of the OXC updates the link state database by indicating the corresponding resources as used. This link state information will then be flooded to other OXC through IGP link-state advertisements.

When the Lightpath Create Response Message reaches the ingress OXC, it advertises this increment in available bandwidth capacity by flooding link-state IGP routing protocol advertisement at the client IP network control-plane level. Routers receiving this link status information apply the increased bandwidth to the generation of LSPs meeting traffic-engineering requirements [MPLS-LSPH].

If an intermediate OXC receiving the Lightpath Create Request Message cannot successfully configure the OXC switch fabric, the intermediate OXC replies with a Notification Message indicating the failure to the ingress OXC. At each of the OXC where the Lightpath Create Request Message has been successfully proceeded and the allocated resources marked as reserved, the corresponding resources previously marked as reserved are released and marked as unused in the link state database. An ingress OXC receiving this reply then sends a Lightpath Create Response Message (to the requesting client) to release any tentatively claimed resources. The detailed procedure is explained in section 4.3.

### 4.2.2 Lightpath setup procedure for backup lightpath

Backup lightpaths are created through the same procedure as primary lightpaths. Depending on the restoration policy defined regarding the backup lightpath provisioning, backup lightpath creation is triggered either during the primary lightpath setup (provisioned) or during the lightpath recovery (non-provisioned).

The backup lightpaths established by this procedure belong to the same restoration group than the primary group of lightpaths to which
they refers. Each of these lightpaths is also identified by the restoration Path_Status field that indicates the status of the lightpath within the range \([0, M]\) and the Restoration Path_Type bit \((= 1)\) defining it as a backup lightpath.

4.3 Procedure for reporting a damaged lightpath

Damaged lightpaths can be triggered by detecting a LoL or a LoS Alarm on the upstream and/or on the downstream OXC depending on failure experienced by the fiber link, the fiber sub-segment and the fiber segment connecting the optical device interfaces [IPO-SRLG].

4.3.1 Link Failure Detection

However, the resulting link failure detection will be one of the following:

- Uni-directional link detection: the failure affects only one fiber of the fiber pair; meaning that the LoS is detected by the downstream OXC or by the upstream OXC.

- Bi-directional link detection: the failure affects both fibers of a fiber pair; meaning that the LoS is detected by the downstream and by the upstream OXC.

In an end-to-end (or partial) restoration, by receiving the alarm signal, the OXC issues a Notification Message translating this physical network failure. The Notification Message comprises the Lightpath ID, the Restoration Group ID and the Restoration Path_Status of the damaged primary or backup lightpath. The Notification Message may optionally include the detailed reason why the impairment occurred.
To detect a unidirectional link failure the following mechanism is proposed: by detecting a LoS Alarm, the OXC generates a Notification in both the downstream and upstream direction of the damaged lightpath. If the upstream (or the downstream) OXC has also detected the LoS Alarm (i.e. bi-directional link failure detection), then it sends back a Notification Message to the downstream (or upstream) OXC. Otherwise (i.e. unidirectional link failure detection) the upstream (or the downstream) OXC forward the Notification toward the ingress (or the egress) OXC.

Depending on restoration mechanism supported within a given ORD:
- Either the OXCs detecting the failure restore the corresponding damaged lightpaths; in that case the restoration procedure is referred to as link restoration (as described in section 3.1). If the link restoration mechanism does not recover the lightpath after the expiration restoration time-out then the notification message is forwarded until reaching the ingress (or the egress) OXC.
- Or if these OXCs does not support (or are not allowed by their restoration policy restriction) to perform link restoration; in that case the notification message is directly sent to the ingress (or the egress) OXC who will apply the restoration procedure(s) described in section 4.4.

4.3.2 Notification Mechanism

The need for an upstream and downstream Notification depends on:
- the restoration policy (end-to-end or partial including link restoration).
Specific timers and thresholds are introduced in order to ensure the proper operation of the restoration mechanism. The following timers and thresholds are defined for that purpose:

- **Notification Timer (NTR):** when an OXC detects a failure, this timer defines the waiting period before sending a notification message. For instance, in the previous example (see Figure 3), if NRT = 0 then T1 - T0 = 0.

- **Forwarding Timer (FTR):** when an OXC receive a notification message, this timer defines the waiting period before forwarding the notification message to the upstream or the downstream OXC. A notification timer of 0 means that the notification messages must be directly sent to the upstream or downstream next OXC. For instance, in the previous example (see Figure 3), if FTR = 0 then T1 - T2 = 0.

- **Restoration Timer (RTR):** when an OXC detects a failure, this timer (also referred to as restoration time-out) defines the period during which the OXC can execute the local restoration procedure (as specified in the next section). A restoration timer value equal to 0 means that the OXC does not have the capability to perform the local path re-routing so that the corresponding forwarding timer equals to 0. For instance, in the previous example (see Figure 3), if RTR = 0 then T1 - T2 = 0 since it implies that FTR must be equal to 0.

### 4.3.3 Reliable Notification Mechanism

In order to provide a reliable notification mechanism, a Notification Counter (RCR) can be additionally defined. This counter provides the required retransmission capabilities in case of Notification Message loss and the continuous notification of the failure until this failure has been repaired (see section 4.4.1).

The continuous notification mechanism is regulated by an exponential back-off algorithm: \( T[n] = 2 \times k \times T[n-1] \leq RTR \) with \( T[0] = 1 \) and \( k =< 1 \).

However, reliability of the notification mechanism can be also achieved by extending the (CR-)LDP fault tolerant (FT) enhancements [MPLS-LDPFT]. This mechanism allows nodes to negotiate how long they will retain FT states such as learned (generalized) label bindings after the lightpath failure occurs as well as the corresponding damaged links but also after the potential failure of the (CR-)LDP session occurs (for signalling plane fault-tolerance purpose). Consequently, when sending the notification message, the OXC (i.e. the one who has detected the failure) is waiting for the acknowledgement of this notification message by the receiver. The receiver is defined as the OXC that takes the failure recovery procedure into consideration and applies the lightpath restoration procedure as described in section 4.4.

The definition of a reliable notification mechanism allows the OXC
to negotiate how long they will remember pre-failure switching state after experiencing a link failure.

4.4 Lightpath Restoration procedure

Restoration procedure can be applied to a damaged primary lightpath or backup lightpath.

4.4.1 Restoration procedure for single damaged primary lightpath

In case of failure of a primary lightpath, the (ingress) OXC is made aware of which lightpath is damaged by the Lightpath ID, the Restoration Group ID and the Restoration Path_Status field included within the Notification Message. The localization of the failure is inferred from the link ID (or the bundle ID) included in the notification message.

- If the provisioned restoration policy has been previously configured for this restoration group, the ingress OXC then replaces the damaged primary lightpath with the backup lightpath shared by the same restoration group. With M:N restoration, the backup lightpath Path_Status corresponds to the value of N (if this value of M exist).

- If the provisioned restoration policy has not been previously configured for this restoration group, then the ingress OXC initiates a lightpath create procedure for the backup lightpath, as described in section 4.2.2. Note that since the route computation for backup lightpath has been determined prior to the primary lightpath failure, the maximum restoration time depends here only on the signaling delay.

When provisioned restoration policy is defined, if the physical resource failure affect also the one computed for the backup lightpath, the restoration policy is automatically switched to the one executed for the non-provisioned case.

When receiving the failure Notification Message, the ingress OSCtrl reconfigures the connection within the OXC switch by switching the connection between the incoming and the outgoing lambda port of the damaged lightpath with that of the backup lightpath. This process is executed internally without exchanging messages but not activated. Then, the ingress OSCtrl reflects the result in the Notification Message it sends to the egress OSCtrl.

Since the failure Notification Message is sent simultaneously from the OXC detecting the failure to the ingress and the egress OXC, this Notification Message is also received by the egress OXCs. This in order to enable the egress OSCtrl to find the outgoing LSC interface for the damaged lightpath. The OSCtrl converts this message into an internal control command it sends to the OXC switch to enable the connection between the incoming and outgoing lambda port in the OXC switch fabric. The OXC switch returns a result status to the OSCtrl. Then, the egress OSCtrl reflects the result in the Notification Message it sends to the ingress OSCtrl.
When the ingress (egress) OSCtrl receives the Notification message sent by the egress (ingress) OSCtrl, corresponding to the lightpath failure, both OXC activate the backup connection.

Since one of the backup lightpath is consumed (or created depending on the restoration policy) by this restoration procedure, a new alternative backup lightpath must be created. This procedure is carried out according to the procedure of section 4.2.2. As we consider the Restoration Path_Status of each of the backup lightpath, this procedure must take into account the Path_Status value of the consumed lightpath.

4.4.2 Restoration procedure for multiple damaged lightpaths

The previous section describes the basic restoration procedure for multiple damaged lightpath. Depending on the restoration policy, the following alternative is considered:

- Provisioned restoration: in order to decrease the number of notification messages sent by the OXC detecting the failure, we consider a unique Notification Message including the Lightpath ID and Path_Status value of each of the damaged lightpath belonging to the same Restoration Group ID. At the ingress OXC, if the number of provisioned backup is not sufficient, then additional backup lightpath can be created as described in section 4.2.2. However, the egress OXC never initiates a backup Lightpath Create Request Message even if the number of backup lightpath is not sufficient to restore a sufficient number of damaged primary lightpaths.

- Non-provisioned restoration: the same notification mechanism as the one described for provisioned restoration is used. However, only the ingress OXC initiates a backup Lightpath Create Request Message in order to restore the damaged primary lightpaths.

4.4.3 Restoration procedure for damaged backup lightpath

In case of failure of a backup lightpath, the ingress OXC must be aware which lightpath is damaged by referring to the Lightpath ID, the Restoration Group ID and Restoration Path_Status of the backup lightpath.

The damage of a backup lightpath must not affect to the ongoing data transfer. However, since the primary lightpaths belonging to the same restoration group lose the restoration capability, an alternative backup lightpath must be created when a backup lightpath failure is detected. The procedure for creating alternative backup lightpath is executed according to the procedure of section 4.2.2.

4.4.4 Restoration procedure for damage primary and backup lightpaths

It is possible that a combination of failures will affect primary and backup lightpaths in a restoration group. In general, in an M:N restoration group, a set of failures can occur so that N' primary lightpaths and M' backup lightpaths are disabled. If the number of
remaining backup lightpaths $M - M'$ is less than the number of affected primary lightpaths $N'$, then it will not be possible for the network to restore service to all the affected primary lightpaths by using the standard restoration procedure described in the preceding sections.

In such a situation, the network should attempt to provide automatic restoration services to as many affected primary lightpaths as possible. If different priority (i.e. precedence) levels are assigned to the various primary lightpaths in the restoration group, those affected lightpaths with the highest priority will have their traffic routed onto the remaining backup lightpaths. Once all the highest-priority affected lightpaths have been served, the network will then attempt to provide backup lightpaths for the set of affected lightpaths that have the next highest precedence level. This process will continue until the supply of unaffected backup lightpaths has been exhausted. If all the affected lightpaths have the same priority level or if priority levels are not being used, backup paths will be activated for them on a first come, first served basis (FIFO-like service).

For the remaining $N' - M + M'$ lightpaths that do not have backup lightpaths available, the network will attempt to create backup lightpaths on the fly using the procedure in Section 4.2.2. It will do this by precedence level or on a first come, first served basis until all $N'$ affect primary lightpaths have been served or until $M$ backup lightpaths have been used. After the restoration procedures have been completed, the network will create $N$ new backup lightpaths, as described in Section 4.4.1, for the restoration group.

4.5 Lightpath release procedure

4.5.1 Release procedure with non-revertive restoration policy

The release of a lightpath can take place for several reasons:
(1) the primary or backup lightpath becoming impaired by the damage
(2) the backup lightpath becoming unneeded due to the release of a sufficient number of primary lightpaths belonging to the same restoration group
(3) the primary or backup lightpath not further processed due to the occurrence of an non-recoverable error during the lightpath creation procedure
(4) the preemption by a lightpath with a higher precedence when the preempting lightpath is unable to obtain free network resources to be created (this case include the blocking probability recovery)
(5) the primary lightpath becoming under-utilized due to a decreased volume of the data traffic (in a data-driven approach only and not considered in this memo since the procedure associated to such an event is a lightpath modification procedure)

A Lightpath Delete Request Message is sent from the ingress OXC to the egress OXC or to an intermediate OXC by passing through the intermediate OXCs comprising the target lightpath that experience a
failure. For this process, the Lightpath Delete Request Message includes the identifier of the damaged lightpath (which stands for the explicit route) in order to determine at each the sequence of intermediate OXCs that this message need to follow.

When an intermediate OXC receives the Lightpath Delete Request Message from its upstream OXC, its OSCtrl attempts to configure the OXC switch fabric according to the Lightpath Delete Request Message content, and wait to receive the result from the OXC switch. If a successful result is returned, the OSCtrl of the intermediate OXC transmits the Lightpath Delete Request Message to the next downstream intermediate OXC. During this procedure, intermediate OXCs release the optical network resources and then update their link state database by marking the corresponding resources as reserved.

If the Lightpath Delete Request Message successfully proceeds all the way to the egress OXC, the latter notifies the result of the delete request by generating a Lightpath Delete Response Message back to the ingress OXC. By receiving this message, intermediate OXCs update their link state database by marking the optical network resources as unused. This link state information will be then flooded through usual OSPF/IS-IS link-state advertisement.

If the intermediate OXC receiving the Lightpath Delete Request (or Response) Message fails to commit the release in the OXC switch fabric, an error message is returned to the OSCtrl and subsequently propagated back using a Notification Message.

### 4.5.2 Maintenance procedure with revertive restoration policy

If we consider a revertive approach, then the resource used by the damaged lightpath must not be released in order to be re-used when the failed resource will be available again. More precisely, the non-release procedure is a standby maintenance procedure since the resources can not be used either by the damaged primary lightpath or by any another lightpath. Note here that we assume that the notification message generated during the failure detection has already trigger usage of a backup lightpath between the ingress and the egress OXC.

By receiving the Notification Message sent by any OXC included within the route of the failed lightpath indicates the corresponding resources as reserved (i.e. locked) within the local link state database. The subsequent unlock process of the resources reserved for this lightpath depends on following alternative:

- if the failure is not recovered at expiration of the restoration time-out, then the OXC(s) still experiencing the damage sends a last Notification Message; and as described for the non-revertive approach, the ingress OXC initiates a primary lightpath deletion procedure;

- if the failure is recovered prior to the expiration of the
restoration time-out, then by reaching the ingress OXC, the last Notification Message initiates a Lightpath Create Request Message whose effect is to unlock the resources reserved for the "restored" lightpath. When both end of the restored lightpath are synchronized (i.e. the ingress OXC received the Lightpath Create Response Message back from the egress OXC) the switching from the backup to the primary lightpath can occur.

5. Signaling Extensions

In this section, we define the message types that are used for the bandwidth provisioning and restoration procedures. These message types and their corresponding can be further defined and included within optical CR-LDP (and within optical RSVP extensions) signaling extensions.

5.1 Lightpath Create Request Message

A Lightpath Create Request Message is issued to create a primary lightpath or backup lightpath by an ingress OXC. As defined in [MPLS-CRLDP] and generalized in [GMPLS-CRLDP], the explicit route field of the Label Request Message (i.e. the Lightpath Create Request Message) specifies the path to be taken by the lightpath being established.

The Label Request Message must include additionally to the parameters already defined, the following specific information defined within. The Restoration TLV format is defined as:

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0|        Type (TBD)         |      Message Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Restoration Group ID TLV                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Path Type & Status TLV                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Preemption TLV                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Restoration Policy TLV                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

The Restoration TLV (optional) which could include the following information:

- Restoration Group ID: 48-bit field identifying the restoration group to which the lightpath belongs. The Restoration Group ID is defined as a unique identifier of a Restoration Group defined within a GMPLS network. The Restoration Group ID is defined as been composed by the following fields:
  - the ingress LSR Router ID (i.e ingress OXC ID) or any of its own IPv4 addresses
- a locally unique Group ID (16-bit) to that LSR
- a locally unique P integer (8-bit) specifying the number restoration Group IDs sharing the backup LSPs, to that LSR

The Restoration Group ID TLV format is defined as:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0|        Type (TBD)          |      Length = 8               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    P Group    |  Res  |ActFlg |      Local Group ID           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Ingress LSR Router ID                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The ActFlg is a 4-bit field that indicates explicitly the action that should be taken if the Restoration Group already exists on the LSR receiving the message. The ActFlg value must be set to 0001 (modify LSP indication) except when a new Restoration Group has to be created.

- Restoration Path_Type: 1-bit field (T bit) enabling the distinction between a primary and backup lightpath is defined through the use of the Restoration Path_Type bit. When the

```
Internet Draft Å» Expires August 2001 24
Draft-many-optical-restoration-00.txt  February 2001
```

Restoration Path_Type bit equals to 0 the lightpath is defined as a primary path and when it equals to 1 the lightpath is defined as a backup path.

- Restoration Path_Status: for primary path this 15-bit integer assigned by the ingress OXC takes a value within the range [0, N] with N < 32767 and for backup path this integer takes a value within the range [0, M] with M < N.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0|        Type (TBD)          |      Length = 4               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T|         Path Status         |      Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- Restoration Priority: this optional parameter defines the setup priority (8-bit field) and holding priority (8-bit field) of the restoration group to which it refers. In [MPLS-CRLDP], the corresponding Preemption TLV is format is:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0|     Type = 0x0820          |      Length = 4               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  SetPriority  | HoldPriority  |      Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
Note that the default value of the setup and holding priorities should be in the middle of the range (e.g., 4).

- Restoration Policy: this optional parameter (8-bit field) includes
  . the restoration strategy bit: revertive (R bit = 0) or non-revertive (R bit = 1)
  . the restoration scheme bit: provisioned (P bit = 0) or non-provisioned (P bit = 1)
  . the restoration diameter bit: partial restoration (D bit = 0) or end-to-end restoration (D bit = 1)
  . the setup (S bit) and recovery (R bit) preemption behavior of the lightpath included within the corresponding restoration group: the setup and the recovery preemption bit values are defined by the ingress OXC;

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>++++++++++++++++++++++++++</td>
<td>Length = 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[R] [P] [D] [S] [R]</td>
<td>Reserved</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When the non-provisioned policy is selected, an additional parameter referred to as the maximum restoration time can be considered. This parameter enable to determine a "pre-selected" route for the backup lightpaths whose signaling delay is less than the maximum restoration time.

The restoration policy parameter values should be the same for all the lightpath belonging to the same restoration group.

5.2 Lightpath Create Response Message

This message is used as the response message for the Lightpath Create Request Message. As described in [MPLS-CRLDP] and generalized in [GMPLS-CRLDP], an ingress OXC identifies a Label Mapping Message (i.e. Lightpath Create Response Message) respectively by comparing the Message ID conveyed by a Label Request Message (i.e. Lightpath Create Request Message) with the Message ID that was issued by itself.

The Label Mapping Message must include additionally to the parameters already defined, the specific information defined within the Restoration TLV (optional); which could include the specific information within the Result Code TLV.

5.3 Notification Message

The lightpath restoration procedures imply the definition of an additional Notification Message type. This Notification Message must include the Lightpath ID (i.e. LSPID TLV - see section 5.3.2) as well as the Status of the corresponding lightpath (or the list of lightpaths).
The Status of the lightpath must include the current status of the lightpath (active, inactive, reserved, failed) as well as optionally the alarm detection such as a LoL, LoS, BER, OSNR, etc. resulting from the physical network damage or failure.

5.3.1 Notification Message format

As defined in [MPLS-LDP], an OXC sends a Notification Message to inform an OXC peer of significant event. A Notification Message signals a fatal error or provides advisory information such as the outcome of processing a CR-LDP message or the state of the LDP session.

Here, to distinguish between Notification Message sent in response to a Label Request Message and Unsolicited Notification Message, we define an Unsolicited Notification Message type. Consequently, to inform an OXC peer of significant event within the transport plane, an OXC sends an Unsolicited Notification Message.

The encoding for the Unsolicited Notification Message is:

```
0                   1                   2                   3
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|   Notification (0x0002)     |      Message Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Message ID                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Status (TLV)                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Optional Parameters                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type » Unsolicited Notification (0x0002)

Message ID
32-bit value used to identify this message.

Status TLV
Indicates the event being signaled. The encoding for the Status TLV is specified in Section "Status TLV" of [MPLS-LDP].

The Status Code included in the Status TLV must have the fatal error E bit = 0 (advisory notification) except if a fatal error is notified E bit = 1 (fatal error). If the Notification Message need to be forwarded the forward F bit = 1, otherwise F bit = 0. The latter case applies only for unidirectional link failure restoration procedure.

Optional Parameters
This variable length field contains 0 or more parameters, each encoded as a TLV. For the purpose of this memo, only the Extended Status TLV (type 0x0301 » length 32-bit) will appear in Notification Messages.
5.3.2 Notification Message Parameters

For the notification messages, the additional parameters defined for lightpath restoration purposes are:

- **Lightpath ID**: which corresponds to the LSPID TLV defined in [MPLS-CR-LDP]

  This identifier is mandatory and corresponds to the 48-bit value of the optical network identifier defining the lightpath. The Lightpath ID corresponds to the LSPID defined in [MPLS-CRLDP] which is a unique identifier of a Constraint-based LSP defined within a GMPLS network. The LSPID is defined as been composed by the ingress LSR Router ID (or any of its own IPv4 addresses) and a locally unique LSPID (16-bit) to that LSR. So, we define the Lightpath ID as been composed by the node ID of the source OXC and a locally unique identifier to that OXC.

  0                   1                   2                   3
  +---------------+---------------+---------------+---------------+
  | 0 | 0 |       Type = 0x0821       |      Length = 4               |
  +---------------+---------------+---------------+---------------+
  |       Reserved        |ActFlg |      Local LSP ID             |
  +---------------+---------------+---------------+---------------+
  |                       Ingress LSR Router ID                   |
  +---------------+---------------+---------------+---------------+

  The ActFlg is a 4-bit field that indicates explicitly the action that should be taken if the LSP already exists on the LSR receiving the message. The ActFlg value must be set to 0001 (modify LSP indication).

- **Link ID (or Bundle ID)**: 32-bit integer identifying the link (or the bundle) that experiences a failure. The Link (or the bundle) can be numbered or unnumbered. A list of Link IDs is needed in order to notify multiple simultaneous failures. Consequently, the format of the Link TLV includes at least one Link ID:

  0                   1                   2                   3
  +---------------+---------------+---------------+---------------+
  | 0 | 1 | 2 | 3 |       Type (TBD)       |      Length (4)       |
  +---------------+---------------+---------------+---------------+
  |                       Link ID                            |
  +---------------+---------------+---------------+---------------+

- **Status**: The additional status (values) included in the Status TLV are defined with respect to the restoration procedure they refer:
  - Active Connection Status defines an Active Notification Message
  - Inactive Connection Status defines an Inactive Notification Message
  - Reserved Connection Status defines a Reservation Notification Message
  - Suspend Connection Status defines a Suspend Notification Message
Failed Connection Status defines a Failure Notification Message

The optional Extended Status only refers to the Failed Connection Status Code. The Extended Status are defined as follows:
- LoL: Loss of Light detected
- LoS: Loss of Signal detected
- BER: BER Threshold exceeded
- OSNR: OSNR Threshold exceeded
- Jitter: Jitter Threshold exceeded

Consequently, the encoding of the Unsolicited Notification Message is defined as:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0| Notification (0x0002) | Message Length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

5.3.3 Fault Tolerant Notification Message

As described in section 4.3.3, a Fault Tolerant (FT) Notification mechanism can be extended to implement a reliable notification procedure after the lightpath failure has been detected.

As described in [MPLS-LDPFT], the corresponding FT Protection TLV is optional and is included only if Reliable Notification as defined in Section 4.3.3 is being negotiated during the FT CR-LDP session establishment. If reliable transport-plane notification is desired, FT status must be negotiated by the participants in the CR-LDP session during the exchange of LDP Initialization Messages.

These messages must contain the FT Session TLV as defined in [MPLS-LDPFT], which is also used to negotiate the amount of time that a node will maintain state for FT resources and labels in the event of LDP session failure. The new FT Path Protection TLV message format is defined as:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0| FT Path Protection (TBD) | Length (= 4) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| FT Sequence Number |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
FT Sequence Number:
32 bit unsigned integer that uniquely identifies each message in a CR-LDP session between two nodes. The sequence number is initialized to 0x00000001 when the CR-LDP session is created at FT status is negotiated. It increments by 1 when a new notification message is transmitted, and wraps from 0xFFFFFFF to 0x00000000.

If an LSR transmits a Fault Notification message using a FT CR-LDP session and does not receive an acknowledgment from the target node within the time allocated by the NCR, it will retransmit the message, without incrementing the FT Sequence Number, using the continuous notification mechanism defined in Section 4.3.2.

ACK Sequence Number:
32 bit unsigned integer that identifies the most recent FT Protection Sequence Number that was received by the sending node in the LDP session, as defined in [MPLS-LDPFT]. The ACK Sequence Number is cumulative, e.g. a value of 0x0000000F indicates that the first 15 messages in the session have been successfully received.

Consequently, the updated Un solicited Notification Message with Reliable Notification encoding is defined as:

While [MPS-LDPFT] does not require FT ACKs to be sent immediately, they should be sent as soon as possible in response to Failure Notification Messages by the LSR recovering the LSP failure. The corresponding FT ACK TLV message format is defined as:
6. Security Considerations

It is essential to maintain secure communication between the OXCs. However, this document does not address the detailed security consideration. These considerations are reserved for future study.

7. References


7. Acknowledgments

This work has been produced from the joint project between ETRI (Electronics and Telecommunications Research Institute in Korea) and NIST (National Institute Standards and Technology).

From this collaboration, the following draft was published: draft-hahm-optical-restoration-01.txt and used as guideline for the publication of the current document.

The authors would like also to thank Doug Montgomery and Mark Carson (NIST) and Bernard Sales, Emmanuel Desmet, Hans De Neve, Stefan Ansorge, Gert Grammel, Mike Sexton and Jim Jones (Alcatel) for their valuable comments, suggestions and inputs.

8. Author’s Addresses

Jin Ho Hahm
ETRI
820 West Diamond Avenue
Gaithersburg, MD 20899, USA
Phone: +1 301-975-8400
Email: hahm@antd.nist.gov & jhhahm@etri.re.kr

Kwang-il Lee
Advanced Network Technologies Division
National Institute of Standards and Technology (NIST)
820 West Diamond Avenue
Gaithersburg, MD 20899, USA
Phone: +1 301 975-8428
Email: kilee@antd.nist.gov

David W. Griffith
Advanced Network Technologies Division