Network Working Group
Request for Comments: 3945
Category: Standards Track

E. Mannie, Ed.

October 2004

Generalized Multi-Protocol Label Switching (GMPLS) Architecture

Status of this Memo

Copyright Notice

Abstract

Future data and transmission networks will consist of elements such as routers, switches, Dense Wavelength Division Multiplexing (DWDM) systems, Add-Drop Multiplexors (ADMs), photonic cross-connects (PXCs), optical cross-connects (OXCs), etc. that will use Generalized Multi-Protocol Label Switching (GMPLS) to dynamically provision resources and to provide network survivability using protection and restoration techniques.

This document describes the architecture of GMPLS. GMPLS extends MPLS to encompass time-division (e.g., SONET/SDH, PDH, G.709), wavelength (lambdas), and spatial switching (e.g., incoming port or fiber to outgoing port or fiber). The focus of GMPLS is on the control plane of these various layers since each of them can use physically diverse data or forwarding planes. The intention is to cover both the signaling and the routing part of that control plane.

Table of Contents

1. Introduction

The architecture described in this document covers the main building blocks needed to build a consistent control plane for multiple switching layers. It does not restrict the way that these layers work together. Different models can be applied, e.g., overlay, augmented or integrated. Moreover, each pair of contiguous layers may collaborate in different ways, resulting in a number of possible combinations, at the discretion of manufacturers and operators.

This architecture clearly separates the control plane and the forwarding plane. In addition, it also clearly separates the control plane in two parts, the signaling plane containing the signaling protocols and the routing plane containing the routing protocols.

This document is a generalization of the Multi-Protocol Label Switching (MPLS) architecture [RFC3031], and in some cases may differ slightly from that architecture since non packet-based forwarding planes are now considered. It is not the intention of this document to describe concepts already described in the current MPLS architecture. The goal is to describe specific concepts of Generalized MPLS (GMPLS).

However, some of the concepts explained hereafter are not part of the current MPLS architecture and are applicable to both MPLS and GMPLS (i.e., link bundling, unnumbered links, and LSP hierarchy). Since these concepts were introduced together with GMPLS and since they are of paramount importance for an operational GMPLS network, they will be discussed here.

The organization of the remainder of this document is as follows. We begin with an introduction of GMPLS. We then present the specific GMPLS building blocks and explain how they can be combined together to build an operational GMPLS network. Specific details of the separate building blocks can be found in the corresponding documents.

1.1. Acronyms & Abbreviations

AS	Autonomous System
BGP	Border Gateway Protocol
CR-LDP	Constraint-based Routing LDP
CSPF	Constraint-based Shortest Path First
DWDM	Dense Wavelength Division Multiplexing
FA	Forwarding Adjacency
GMPLS	Generalized Multi-Protocol Label Switching
IGP	Interior Gateway Protocol
LDP	Label Distribution Protocol
LMP	Link Management Protocol

LSA	Link State Advertisement
LSR	Label Switching Router
LSP	Label Switched Path
MIB	Management Information Base
MPLS	Multi-Protocol Label Switching
NMS	Network Management System
OXC	Optical Cross-Connect
PXC	Photonic Cross-Connect
RSVP	ReSource reserVation Protocol
SDH	Synchronous Digital Hierarchy
SONET	Synchronous Optical Networks
STM(-N)	Synchronous Transport Module (-N)
STS(-N)	Synchronous Transport Signal-Level N (SONET)
TDM	Time Division Multiplexing
TE	Traffic Engineering

1.2. Multiple Types of Switching and Forwarding Hierarchies

Generalized MPLS (GMPLS) differs from traditional MPLS in that it supports multiple types of switching, i.e., the addition of support for TDM, lambda, and fiber (port) switching. The support for the additional types of switching has driven GMPLS to extend certain base functions of traditional MPLS and, in some cases, to add functionality. These changes and additions impact basic LSP properties: how labels are requested and communicated, the unidirectional nature of LSPs, how errors are propagated, and information provided for synchronizing the ingress and egress LSRs.

The MPLS architecture [RFC3031] was defined to support the forwarding of data based on a label. In this architecture, Label Switching Routers (LSRs) were assumed to have a forwarding plane that is capable of (a) recognizing either packet or cell boundaries, and (b) being able to process either packet headers (for LSRs capable of recognizing packet boundaries) or cell headers (for LSRs capable of recognizing cell boundaries).

The original MPLS architecture is here being extended to include LSRs whose forwarding plane recognizes neither packet, nor cell boundaries, and therefore, cannot forward data based on the information carried in either packet or cell headers. Specifically, such LSRs include devices where the switching decision is based on time slots, wavelengths, or physical ports. So, the new set of LSRs, or more precisely interfaces on these LSRs, can be subdivided into the following classes:

1. Packet Switch Capable (PSC) interfaces:

   Interfaces that recognize packet boundaries and can forward data based on the content of the packet header. Examples include interfaces on routers that forward data based on the content of the IP header and interfaces on routers that switch data based on the content of the MPLS "shim" header.

2. Layer-2 Switch Capable (L2SC) interfaces:

   Interfaces that recognize frame/cell boundaries and can switch data based on the content of the frame/cell header. Examples include interfaces on Ethernet bridges that switch data based on the content of the MAC header and interfaces on ATM-LSRs that forward data based on the ATM VPI/VCI.

3. Time-Division Multiplex Capable (TDM) interfaces:

   Interfaces that switch data based on the data's time slot in a repeating cycle. An example of such an interface is that of a SONET/SDH Cross-Connect (XC), Terminal Multiplexer (TM), or Add-Drop Multiplexer (ADM). Other examples include interfaces providing G.709 TDM capabilities (the "digital wrapper") and PDH interfaces.

4. Lambda Switch Capable (LSC) interfaces:

   Interfaces that switch data based on the wavelength on which the data is received. An example of such an interface is that of a Photonic Cross-Connect (PXC) or Optical Cross-Connect (OXC) that can operate at the level of an individual wavelength. Additional examples include PXC interfaces that can operate at the level of a group of wavelengths, i.e., a waveband and G.709 interfaces providing optical capabilities.

5. Fiber-Switch Capable (FSC) interfaces:

   Interfaces that switch data based on a position of the data in the (real world) physical spaces. An example of such an interface is that of a PXC or OXC that can operate at the level of a single or multiple fibers.

A circuit can be established only between, or through, interfaces of the same type. Depending on the particular technology being used for each interface, different circuit names can be used, e.g., SDH circuit, optical trail, light-path, etc. In the context of GMPLS, all these circuits are referenced by a common name: Label Switched Path (LSP).

The concept of nested LSP (LSP within LSP), already available in the traditional MPLS, facilitates building a forwarding hierarchy, i.e., a hierarchy of LSPs. This hierarchy of LSPs can occur on the same interface, or between different interfaces.

For example, a hierarchy can be built if an interface is capable of multiplexing several LSPs from the same technology (layer), e.g., a lower order SONET/SDH LSP (e.g., VT2/VC-12) nested in a higher order SONET/SDH LSP (e.g., STS-3c/VC-4). Several levels of signal (LSP) nesting are defined in the SONET/SDH multiplexing hierarchy.

The nesting can also occur between interface types. At the top of the hierarchy are FSC interfaces, followed by LSC interfaces, followed by TDM interfaces, followed by L2SC, and followed by PSC interfaces. This way, an LSP that starts and ends on a PSC interface can be nested (together with other LSPs) into an LSP that starts and ends on a L2SC interface. This LSP, in turn, can be nested (together with other LSPs) into an LSP that starts and ends on a TDM interface. In turn, this LSP can be nested (together with other LSPs) into an LSP that starts and ends on a LSC interface, which in turn can be nested (together with other LSPs) into an LSP that starts and ends on a FSC interface.

1.3. Extension of the MPLS Control Plane

The establishment of LSPs that span only Packet Switch Capable (PSC) or Layer-2 Switch Capable (L2SC) interfaces is defined for the original MPLS and/or MPLS-TE control planes. GMPLS extends these control planes to support each of the five classes of interfaces (i.e., layers) defined in the previous section.

Note that the GMPLS control plane supports an overlay model, an augmented model, and a peer (integrated) model. In the near term, GMPLS appears to be very suitable for controlling each layer independently. This elegant approach will facilitate the future deployment of other models.

The GMPLS control plane is made of several building blocks as described in more details in the following sections. These building blocks are based on well-known signaling and routing protocols that have been extended and/or modified to support GMPLS. They use IPv4 and/or IPv6 addresses. Only one new specialized protocol is required to support the operations of GMPLS, a signaling protocol for link management [LMP].

GMPLS is indeed based on the Traffic Engineering (TE) extensions to MPLS, a.k.a. MPLS-TE [RFC2702]. This, because most of the technologies that can be used below the PSC level requires some

traffic engineering. The placement of LSPs at these levels needs in general to consider several constraints (such as framing, bandwidth, protection capability, etc) and to bypass the legacy Shortest-Path First (SPF) algorithm. Note, however, that this is not mandatory and that in some cases SPF routing can be applied.

In order to facilitate constrained-based SPF routing of LSPs, nodes that perform LSP establishment need more information about the links in the network than standard intra-domain routing protocols provide. These TE attributes are distributed using the transport mechanisms already available in IGPs (e.g., flooding) and taken into consideration by the LSP routing algorithm. Optimization of the LSP routes may also require some external simulations using heuristics that serve as input for the actual path calculation and LSP establishment process.

By definition, a TE link is a representation in the IS-IS/OSPF Link State advertisements and in the link state database of certain physical resources, and their properties, between two GMPLS nodes. TE Links are used by the GMPLS control plane (routing and signaling) for establishing LSPs.

Extensions to traditional routing protocols and algorithms are needed to uniformly encode and carry TE link information, and explicit routes (e.g., source routes) are required in the signaling. In addition, the signaling must now be capable of transporting the required circuit (LSP) parameters such as the bandwidth, the type of signal, the desired protection and/or restoration, the position in a particular multiplex, etc. Most of these extensions have already been defined for PSC and L2SC traffic engineering with MPLS. GMPLS primarily defines additional extensions for TDM, LSC, and FSC traffic engineering. A very few elements are technology specific.

Thus, GMPLS extends the two signaling protocols defined for MPLS-TE signaling, i.e., RSVP-TE [RFC3209] and CR-LDP [RFC3212]. However, GMPLS does not specify which one of these two signaling protocols must be used. It is the role of manufacturers and operators to evaluate the two possible solutions for their own interest.

Since GMPLS signaling is based on RSVP-TE and CR-LDP, it mandates a downstream-on-demand label allocation and distribution, with ingress initiated ordered control. Liberal label retention is normally used, but conservative label retention mode could also be used.

Furthermore, there is no restriction on the label allocation strategy, it can be request/signaling driven (obvious for circuit switching technologies), traffic/data driven, or even topology driven. There is also no restriction on the route selection; explicit routing is normally used (strict or loose) but hop-by-hop routing could be used as well.

GMPLS also extends two traditional intra-domain link-state routing protocols already extended for TE purposes, i.e., OSPF-TE [OSPF-TE] and IS-IS-TE [ISIS-TE]. However, if explicit (source) routing is used, the routing algorithms used by these protocols no longer need to be standardized. Extensions for inter-domain routing (e.g., BGP) are for further study.

The use of technologies like DWDM (Dense Wavelength Division Multiplexing) implies that we can now have a very large number of parallel links between two directly adjacent nodes (hundreds of wavelengths, or even thousands of wavelengths if multiple fibers are used). Such a large number of links was not originally considered for an IP or MPLS control plane, although it could be done. Some slight adaptations of that control plane are thus required if we want to better reuse it in the GMPLS context.

For instance, the traditional IP routing model assumes the establishment of a routing adjacency over each link connecting two adjacent nodes. Having such a large number of adjacencies does not scale well. Each node needs to maintain each of its adjacencies one by one, and link state routing information must be flooded throughout the network.

To solve this issue the concept of link bundling was introduced. Moreover, the manual configuration and control of these links, even if they are unnumbered, becomes impractical. The Link Management Protocol (LMP) was specified to solve these issues.

LMP runs between data plane adjacent nodes and is used to manage TE links. Specifically, LMP provides mechanisms to maintain control channel connectivity (IP Control Channel Maintenance), verify the physical connectivity of the data-bearing links (Link Verification), correlate the link property information (Link Property Correlation), and manage link failures (Fault Localization and Fault Notification). A unique feature of LMP is that it is able to localize faults in both opaque and transparent networks (i.e., independent of the encoding scheme and bit rate used for the data).

LMP is defined in the context of GMPLS, but is specified independently of the GMPLS signaling specification since it is a local protocol running between data-plane adjacent nodes.

Consequently, LMP can be used in other contexts with non-GMPLS signaling protocols.

MPLS signaling and routing protocols require at least one bi-directional control channel to communicate even if two adjacent nodes are connected by unidirectional links. Several control channels can be used. LMP can be used to establish, maintain and manage these control channels.

GMPLS does not specify how these control channels must be implemented, but GMPLS requires IP to transport the signaling and routing protocols over them. Control channels can be either in-band or out-of-band, and several solutions can be used to carry IP. Note also that one type of LMP message (the Test message) is used in-band in the data plane and may not be transported over IP, but this is a particular case, needed to verify connectivity in the data plane.

1.4. GMPLS Key Extensions to MPLS-TE

Some key extensions brought by GMPLS to MPLS-TE are highlighted in the following. Some of them are key advantages of GMPLS to control TDM, LSC and FSC layers.

- In MPLS-TE, links traversed by an LSP can include an intermix of links with heterogeneous label encoding (e.g., links between routers, links between routers and ATM-LSRs, and links between ATM-LSRs. GMPLS extends this by including links where the label is encoded as a time slot, or a wavelength, or a position in the (real world) physical space.

- In MPLS-TE, an LSP that carries IP has to start and end on a router. GMPLS extends this by requiring an LSP to start and end on similar type of interfaces.

- The type of a payload that can be carried in GMPLS by an LSP is extended to allow such payloads as SONET/SDH, G.709, 1Gb or 10Gb Ethernet, etc.

- The use of Forwarding Adjacencies (FA) provides a mechanism that can improve bandwidth utilization, when bandwidth allocation can be performed only in discrete units. It offers also a mechanism to aggregate forwarding state, thus allowing the number of required labels to be reduced.

- GMPLS allows suggesting a label by an upstream node to reduce the setup latency. This suggestion may be overridden by a downstream node but in some cases, at the cost of higher LSP setup time.

- GMPLS extends on the notion of restricting the range of labels that may be selected by a downstream node. In GMPLS, an upstream node may restrict the labels for an LSP along either a single hop or the entire LSP path. This feature is useful in photonic networks where wavelength conversion may not be available.

- While traditional TE-based (and even LDP-based) LSPs are unidirectional, GMPLS supports the establishment of bi-directional LSPs.

- GMPLS supports the termination of an LSP on a specific egress port, i.e., the port selection at the destination side.

- GMPLS with RSVP-TE supports an RSVP specific mechanism for rapid failure notification.

Note also some other key differences between MPLS-TE and GMPLS:

- For TDM, LSC and FSC interfaces, bandwidth allocation for an LSP can be performed only in discrete units.

- It is expected to have (much) fewer labels on TDM, LSC or FSC links than on PSC or L2SC links, because the former are physical labels instead of logical labels.

2. Routing and Addressing Model

GMPLS is based on the IP routing and addressing models. This assumes that IPv4 and/or IPv6 addresses are used to identify interfaces but also that traditional (distributed) IP routing protocols are reused. Indeed, the discovery of the topology and the resource state of all links in a routing domain is achieved via these routing protocols.

Since control and data planes are de-coupled in GMPLS, control-plane neighbors (i.e., IGP-learnt neighbors) may not be data-plane neighbors. Hence, mechanisms like LMP are needed to associate TE links with neighboring nodes.

IP addresses are not used only to identify interfaces of IP hosts and routers, but more generally to identify any PSC and non-PSC interfaces. Similarly, IP routing protocols are used to find routes for IP datagrams with a SPF algorithm; they are also used to find routes for non-PSC circuits by using a CSPF algorithm.

However, some additional mechanisms are needed to increase the scalability of these models and to deal with specific traffic engineering requirements of non-PSC layers. These mechanisms will be introduced in the following.

Re-using existing IP routing protocols allows for non-PSC layers taking advantage of all the valuable developments that took place since years for IP routing, in particular, in the context of intra-domain routing (link-state routing) and inter-domain routing (policy routing).

In an overlay model, each particular non-PSC layer can be seen as a set of Autonomous Systems (ASs) interconnected in an arbitrary way. Similarly to the traditional IP routing, each AS is managed by a single administrative authority. For instance, an AS can be an SONET/SDH network operated by a given carrier. The set of interconnected ASs can be viewed as SONET/SDH internetworks.

Exchange of routing information between ASs can be done via an inter-domain routing protocol like BGP-4. There is obviously a huge value of re-using well-known policy routing facilities provided by BGP in a non-PSC context. Extensions for BGP traffic engineering (BGP-TE) in the context of non-PSC layers are left for further study.

Each AS can be sub-divided in different routing domains, and each can run a different intra-domain routing protocol. In turn, each routing-domain can be divided in areas.

A routing domain is made of GMPLS enabled nodes (i.e., a network device including a GMPLS entity). These nodes can be either edge nodes (i.e., hosts, ingress LSRs or egress LSRs), or internal LSRs. An example of non-PSC host is an SONET/SDH Terminal Multiplexer (TM). Another example is an SONET/SDH interface card within an IP router or ATM switch.

Note that traffic engineering in the intra-domain requires the use of link-state routing protocols like OSPF or IS-IS.

GMPLS defines extensions to these protocols. These extensions are needed to disseminate specific TDM, LSC and FSC static and dynamic characteristics related to nodes and links. The current focus is on

intra-area traffic engineering. However, inter-area traffic engineering is also under investigation.

2.1. Addressing of PSC and non-PSC Layers

The fact that IPv4 and/or IPv6 addresses are used does not imply at all that they should be allocated in the same addressing space than public IPv4 and/or IPv6 addresses used for the Internet. Private IP addresses can be used if they do not require to be exchanged with any other operator; public IP addresses are otherwise required. Of course, if an integrated model is used, two layers could share the same addressing space. Finally, TE links may be "unnumbered" i.e., not have any IP addresses, in case IP addresses are not available, or the overhead of managing them is considered too high.

Note that there is a benefit of using public IPv4 and/or IPv6 Internet addresses for non-PSC layers if an integrated model with the IP layer is foreseen.

If we consider the scalability enhancements proposed in the next section, the IPv4 (32 bits) and the IPv6 (128 bits) addressing spaces are both more than sufficient to accommodate any non-PSC layer. We can reasonably expect to have much less non-PSC devices (e.g., SONET/SDH nodes) than we have today IP hosts and routers.

2.2. GMPLS Scalability Enhancements

TDM, LSC and FSC layers introduce new constraints on the IP addressing and routing models since several hundreds of parallel physical links (e.g., wavelengths) can now connect two nodes. Most of the carriers already have today several tens of wavelengths per fiber between two nodes. New generation of DWDM systems will allow several hundreds of wavelengths per fiber.

It becomes rather impractical to associate an IP address with each end of each physical link, to represent each link as a separate routing adjacency, and to advertise and to maintain link states for each of these links. For that purpose, GMPLS enhances the MPLS routing and addressing models to increase their scalability.

Two optional mechanisms can be used to increase the scalability of the addressing and the routing: unnumbered links and link bundling. These two mechanisms can also be combined. They require extensions to signaling (RSVP-TE and CR-LDP) and routing (OSPF-TE and IS-IS-TE) protocols.

2.3. TE Extensions to IP Routing Protocols

Traditionally, a TE link is advertised as an adjunct to a "regular" OSPF or IS-IS link, i.e., an adjacency is brought up on the link. When the link is up, both the regular IGP properties of the link (basically, the SPF metric) and the TE properties of the link are then advertised.

However, GMPLS challenges this notion in three ways:

- First, links that are non-PSC may yet have TE properties; however, an OSPF adjacency could not be brought up directly on such links.

- Second, an LSP can be advertised as a point-to-point TE link in the routing protocol, i.e., as a Forwarding Adjacency (FA); thus, an advertised TE link need no longer be between two OSPF direct neighbors. Forwarding Adjacencies (FA) are further described in Section 8.

- Third, a number of links may be advertised as a single TE link (e.g., for improved scalability), so again, there is no longer a one-to-one association of a regular adjacency and a TE link.

Thus, we have a more general notion of a TE link. A TE link is a logical link that has TE properties. Some of these properties may be configured on the advertising LSR, others may be obtained from other LSRs by means of some protocol, and yet others may be deduced from the component(s) of the TE link.

An important TE property of a TE link is related to the bandwidth accounting for that link. GMPLS will define different accounting rules for different non-PSC layers. Generic bandwidth attributes are however defined by the TE routing extensions and by GMPLS, such as the unreserved bandwidth, the maximum reservable bandwidth and the maximum LSP bandwidth.

It is expected in a dynamic environment to have frequent changes of bandwidth accounting information. A flexible policy for triggering link state updates based on bandwidth thresholds and link-dampening mechanism can be implemented.

TE properties associated with a link should also capture protection and restoration related characteristics. For instance, shared protection can be elegantly combined with bundling. Protection and restoration are mainly generic mechanisms also applicable to MPLS. It is expected that they will first be developed for MPLS and later on generalized to GMPLS.

A TE link between a pair of LSRs does not imply the existence of an IGP adjacency between these LSRs. A TE link must also have some means by which the advertising LSR can know of its liveness (e.g., by using LMP hellos). When an LSR knows that a TE link is up, and can determine the TE link's TE properties, the LSR may then advertise that link to its GMPLS enhanced OSPF or IS-IS neighbors using the TE objects/TLVs. We call the interfaces over which GMPLS enhanced OSPF or IS-IS adjacencies are established "control channels".

3. Unnumbered Links

Unnumbered links (or interfaces) are links (or interfaces) that do not have IP addresses. Using such links involves two capabilities: the ability to specify unnumbered links in MPLS TE signaling, and the ability to carry (TE) information about unnumbered links in IGP TE extensions of IS-IS-TE and OSPF-TE.

A. The ability to specify unnumbered links in MPLS TE signaling requires extensions to RSVP-TE [RFC3477] and CR-LDP [RFC3480]. The MPLS-TE signaling does not provide support for unnumbered links, because it does not provide a way to indicate an unnumbered link in its Explicit Route Object/TLV and in its Record Route Object (there is no such TLV for CR-LDP). GMPLS defines simple extensions to indicate an unnumbered link in these two Objects/TLVs, using a new Unnumbered Interface ID sub-object/sub-TLV.

   Since unnumbered links are not identified by an IP address, then for the purpose of MPLS TE each end need some other identifier, local to the LSR to which the link belongs. LSRs at the two end-points of an unnumbered link exchange with each other the identifiers they assign to the link. Exchanging the identifiers may be accomplished by configuration, by means of a protocol such as LMP ([LMP]), by means of RSVP-TE/CR-LDP (especially in the case where a link is a Forwarding Adjacency, see below), or by means of IS-IS or OSPF extensions ([ISIS-TE-GMPLS], [OSPF-TE-GMPLS]).

   Consider an (unnumbered) link between LSRs A and B. LSR A chooses an identifier for that link. So does LSR B. From A's perspective we refer to the identifier that A assigned to the link as the "link local identifier" (or just "local identifier"), and to the identifier that B assigned to the link as the "link remote identifier" (or just "remote identifier"). Likewise, from B's perspective the identifier that B assigned to the link is the local identifier, and the identifier that A assigned to the link is the remote identifier.

The new Unnumbered Interface ID sub-object/sub-TLV for the ER Object/TLV contains the Router ID of the LSR at the upstream end of the unnumbered link and the link local identifier with respect to that upstream LSR.

The new Unnumbered Interface ID sub-object for the RR Object contains the link local identifier with respect to the LSR that adds it in the RR Object.

B. The ability to carry (TE) information about unnumbered links in IGP TE extensions requires new sub-TLVs for the extended IS reachability TLV defined in IS-IS-TE and for the TE LSA (which is an opaque LSA) defined in OSPF-TE. A Link Local Identifier sub-TLV and a Link Remote Identifier sub-TLV are defined.

3.1. Unnumbered Forwarding Adjacencies

If an LSR that originates an LSP advertises this LSP as an unnumbered FA in IS-IS or OSPF, or the LSR uses this FA as an unnumbered component link of a bundled link, the LSR must allocate an Interface ID to that FA. If the LSP is bi-directional, the tail end does the same and allocates an Interface ID to the reverse FA.

Signaling has been enhanced to carry the Interface ID of a FA in the new LSP Tunnel Interface ID object/TLV. This object/TLV contains the Router ID (of the LSR that generates it) and the Interface ID. It is called the Forward Interface ID when it appears in a Path/REQUEST message, and it is called the Reverse Interface ID when it appears in the Resv/MAPPING message.

4. Link Bundling

The concept of link bundling is essential in certain networks employing the GMPLS control plane as is defined in [BUNDLE]. A typical example is an optical meshed network where adjacent optical cross-connects (LSRs) are connected by several hundreds of parallel wavelengths. In this network, consider the application of link state routing protocols, like OSPF or IS-IS, with suitable extensions for resource discovery and dynamic route computation. Each wavelength must be advertised separately to be used, except if link bundling is used.

When a pair of LSRs is connected by multiple links, it is possible to advertise several (or all) of these links as a single link into OSPF and/or IS-IS. This process is called link bundling, or just bundling. The resulting logical link is called a bundled link as its physical links are called component links (and are identified by interface indexes).

The result is that a combination of three identifiers ((bundled) link identifier, component link identifier, label) is sufficient to unambiguously identify the appropriate resources used by an LSP.

The purpose of link bundling is to improve routing scalability by reducing the amount of information that has to be handled by OSPF and/or IS-IS. This reduction is accomplished by performing information aggregation/abstraction. As with any other information aggregation/abstraction, this results in losing some of the information. To limit the amount of losses one need to restrict the type of the information that can be aggregated/abstracted.

4.1. Restrictions on Bundling

The following restrictions are required for bundling links. All component links in a bundle must begin and end on the same pair of LSRs; and share some common characteristics or properties defined in [OSPF-TE] and [ISIS-TE], i.e., they must have the same:

- Link Type (i.e., point-to-point or multi-access),
- TE Metric (i.e., an administrative cost),
- Set of Resource Classes at each end of the links (i.e., colors).

Note that a FA may also be a component link. In fact, a bundle can consist of a mix of point-to-point links and FAs, but all sharing some common properties.

4.2. Routing Considerations for Bundling

A bundled link is just another kind of TE link such as those defined by [GMPLS-ROUTING]. The liveness of the bundled link is determined by the liveness of each its component links. A bundled link is alive when at least one of its component links is alive. The liveness of a component link can be determined by any of several means: IS-IS or OSPF hellos over the component link, or RSVP Hello (hop local), or LMP hellos (link local), or from layer 1 or layer 2 indications.

Note that (according to the RSVP-TE specification [RFC3209]) the RSVP Hello mechanism is intended to be used when notification of link layer failures is not available and unnumbered links are not used, or when the failure detection mechanisms provided by the link layer are not sufficient for timely node failure detection.

Once a bundled link is determined to be alive, it can be advertised as a TE link and the TE information can be flooded. If IS-IS/OSPF hellos are run over the component links, IS-IS/OSPF flooding can be restricted to just one of the component links.

Note that advertising a (bundled) TE link between a pair of LSRs does not imply that there is an IGP adjacency between these LSRs that is associated with just that link. In fact, in certain cases a TE link between a pair of LSRs could be advertised even if there is no IGP adjacency at all between the LSR (e.g., when the TE link is an FA).

Forming a bundled link consist in aggregating the identical TE parameters of each individual component link to produce aggregated TE parameters. A TE link as defined by [GMPLS-ROUTING] has many parameters; adequate aggregation rules must be defined for each one.

Some parameters can be sums of component characteristics such as the unreserved bandwidth and the maximum reservable bandwidth. Bandwidth information is an important part of a bundle advertisement and it must be clearly defined since an abstraction is done.

A GMPLS node with bundled links must apply admission control on a per-component link basis.

4.3. Signaling Considerations

Typically, an LSP's explicit route (e.g., contained in an explicit route Object/TLV) will choose the bundled link to be used for the LSP, but not the component link(s). This because information about the bundled link is flooded but information about the component links is not.

The choice of the component link to use is always made by an upstream node. If the LSP is bi-directional, the upstream node chooses a component link in each direction.

Three mechanisms for indicating this choice to the downstream node are possible.

4.3.1. Mechanism 1: Implicit Indication

This mechanism requires that each component link has a dedicated signaling channel (e.g., the link is a Sonet/SDH link using the DCC for in-band signaling). The upstream node tells the receiver which component link to use by sending the message over the chosen component link's dedicated signaling channel. Note that this signaling channel can be in-band or out-of-band. In this last case, the association between the signaling channel and that component link need to be explicitly configured.

4.3.2. Mechanism 2: Explicit Indication by Numbered Interface ID

This mechanism requires that the component link has a unique remote IP address. The upstream node indicates the choice of the component link by including a new IF_ID RSVP_HOP object/IF_ID TLV carrying either an IPv4 or an IPv6 address in the Path/Label Request message (see [RFC3473]/[RFC3472], respectively). For a bi-directional LSP, a component link is provided for each direction by the upstream node.

This mechanism does not require each component link to have its own control channel. In fact, it does not even require the whole (bundled) link to have its own control channel.

4.3.3. Mechanism 3: Explicit Indication by Unnumbered Interface ID

With this mechanism, each component link that is unnumbered is assigned a unique Interface Identifier (32 bits value). The upstream node indicates the choice of the component link by including a new IF_ID RSVP_HOP object/IF_ID TLV in the Path/Label Request message (see [RFC3473]/[RFC3472], respectively).

This object/TLV carries the component interface ID in the downstream direction for a unidirectional LSP, and in addition, the component interface ID in the upstream direction for a bi-directional LSP.

The two LSRs at each end of the bundled link exchange these identifiers. Exchanging the identifiers may be accomplished by configuration, by means of a protocol such as LMP (preferred solution), by means of RSVP-TE/CR-LDP (especially in the case where a component link is a Forwarding Adjacency), or by means of IS-IS or OSPF extensions.

This mechanism does not require each component link to have its own control channel. In fact, it does not even require the whole (bundled) link to have its own control channel.

4.4. Unnumbered Bundled Link

A bundled link may itself be numbered or unnumbered independent of whether the component links are numbered or not. This affects how the bundled link is advertised in IS-IS/OSPF and the format of LSP EROs that traverse the bundled link. Furthermore, unnumbered Interface Identifiers for all unnumbered outgoing links of a given LSR (whether component links, Forwarding Adjacencies or bundled links) must be unique in the context of that LSR.

4.5. Forming Bundled Links

The generic rule for bundling component links is to place those links that are correlated in some manner in the same bundle. If links may be correlated based on multiple properties then the bundling may be applied sequentially based on these properties. For instance, links may be first grouped based on the first property. Each of these groups may be then divided into smaller groups based on the second property and so on. The main principle followed in this process is that the properties of the resulting bundles should be concisely summarizable. Link bundling may be done automatically or by configuration. Automatic link bundling can apply bundling rules sequentially to produce bundles.

For instance, the first property on which component links may be correlated could be the Interface Switching Capability [GMPLS-ROUTING], the second property could be the Encoding [GMPLS-ROUTING], the third property could be the Administrative Weight (cost), the fourth property could be the Resource Classes and finally links may be correlated based on other metrics such as SRLG (Shared Risk Link Groups).

When routing an alternate path for protection purposes, the general principle followed is that the alternate path is not routed over any link belonging to an SRLG that belongs to some link of the primary path. Thus, the rule to be followed is to group links belonging to exactly the same set of SRLGs.

This type of sequential sub-division may result in a number of bundles between two adjacent nodes. In practice, however, the link properties may not be very heterogeneous among component links between two adjacent nodes. Thus, the number of bundles in practice may not be large.

5. Relationship with the UNI

The interface between an edge GMPLS node and a GMPLS LSR on the network side may be referred to as a User to Network Interface (UNI), while the interface between two-network side LSRs may be referred to as a Network to Network Interface (NNI).

GMPLS does not specify separately a UNI and an NNI. Edge nodes are connected to LSRs on the network side, and these LSRs are in turn connected between them. Of course, the behavior of an edge node is not exactly the same as the behavior of an LSR on the network side. Note also, that an edge node may run a routing protocol, however it is expected that in most of the cases it will not (see also section 5.2 and the section about signaling with an explicit route).

Conceptually, a difference between UNI and NNI make sense either if both interface uses completely different protocols, or if they use the same protocols but with some outstanding differences. In the first case, separate protocols are often defined successively, with more or less success.

The GMPLS approach consisted in building a consistent model from day one, considering both the UNI and NNI interfaces at the same time [GMPLS-OVERLAY]. For that purpose, a very few specific UNI particularities have been ignored in a first time. GMPLS has been enhanced to support such particularities at the UNI by some other standardization bodies (see hereafter).

5.1. Relationship with the OIF UNI

This section is only given for reference to the OIF work related to GMPLS. The current OIF UNI specification [OIF-UNI] defines an interface between a client SONET/SDH equipment and an SONET/SDH network, each belonging to a distinct administrative authority. It is designed for an overlay model. The OIF UNI defines additional mechanisms on the top of GMPLS for the UNI.

For instance, the OIF service discovery procedure is a precursor to obtaining UNI services. Service discovery allows a client to determine the static parameters of the interconnection with the network, including the UNI signaling protocol, the type of concatenation, the transparency level as well as the type of diversity (node, link, SRLG) supported by the network.

Since the current OIF UNI interface does not cover photonic networks, G.709 Digital Wrapper, etc, it is from that perspective a subset of the GMPLS Architecture at the UNI.

5.2. Reachability across the UNI

This section discusses the selection of an explicit route by an edge node. The selection of the first LSR by an edge node connected to multiple LSRs is part of that problem.

An edge node (host or LSR) can participate more or less deeply in the GMPLS routing. Four different routing models can be supported at the UNI: configuration based, partial peering, silent listening and full peering.

- Configuration based: this routing model requires the manual or automatic configuration of an edge node with a list of neighbor LSRs sorted by preference order. Automatic configuration can be achieved using DHCP for instance. No routing information is

exchanged at the UNI, except maybe the ordered list of LSRs. The only routing information used by the edge node is that list. The edge node sends by default an LSP request to the preferred LSR. ICMP redirects could be send by this LSR to redirect some LSP requests to another LSR connected to the edge node. GMPLS does not preclude that model.

- Partial peering: limited routing information (mainly reachability) can be exchanged across the UNI using some extensions in the signaling plane. The reachability information exchanged at the UNI may be used to initiate edge node specific routing decision over the network. GMPLS does not have any capability to support this model today.

- Silent listening: the edge node can silently listen to routing protocols and take routing decisions based on the information obtained. An edge node receives the full routing information, including traffic engineering extensions. One LSR should forward transparently all routing PDUs to the edge node. An edge node can now compute a complete explicit route taking into consideration all the end-to-end routing information. GMPLS does not preclude this model.

- Full peering: in addition to silent listening, the edge node participates within the routing, establish adjacencies with its neighbors and advertises LSAs. This is useful only if there are benefits for edge nodes to advertise themselves traffic engineering information. GMPLS does not preclude this model.

6. Link Management

In the context of GMPLS, a pair of nodes (e.g., a photonic switch) may be connected by tens of fibers, and each fiber may be used to transmit hundreds of wavelengths if DWDM is used. Multiple fibers and/or multiple wavelengths may also be combined into one or more bundled links for routing purposes. Furthermore, to enable communication between nodes for routing, signaling, and link management, control channels must be established between a node pair.

Link management is a collection of useful procedures between adjacent nodes that provide local services such as control channel management, link connectivity verification, link property correlation, and fault management. The Link Management Protocol (LMP) [LMP] has been defined to fulfill these operations. LMP has been initiated in the context of GMPLS but is a generic toolbox that can be also used in other contexts.

In GMPLS, the control channels between two adjacent nodes are no longer required to use the same physical medium as the data links between those nodes. Moreover, the control channels that are used to exchange the GMPLS control-plane information exist independently of the links they manage. Hence, LMP was designed to manage the data links, independently of the termination capabilities of those data links.

Control channel management and link property correlation procedures are mandatory per LMP. Link connectivity verification and fault management procedures are optional.

6.1. Control Channel and Control Channel Management

LMP control channel management is used to establish and maintain control channels between nodes. Control channels exist independently of TE links, and can be used to exchange MPLS control-plane information such as signaling, routing, and link management information.

An "LMP adjacency" is formed between two nodes that support the same LMP capabilities. Multiple control channels may be active simultaneously for each adjacency. A control channel can be either explicitly configured or automatically selected, however, LMP currently assume that control channels are explicitly configured while the configuration of the control channel capabilities can be dynamically negotiated.

For the purposes of LMP, the exact implementation of the control channel is left unspecified. The control channel(s) between two adjacent nodes is no longer required to use the same physical medium as the data-bearing links between those nodes. For example, a control channel could use a separate wavelength or fiber, an Ethernet link, or an IP tunnel through a separate management network.

A consequence of allowing the control channel(s) between two nodes to be physically diverse from the associated data-bearing links is that the health of a control channel does not necessarily correlate to the health of the data-bearing links, and vice-versa. Therefore, new mechanisms have been developed in LMP to manage links, both in terms of link provisioning and fault isolation.

LMP does not specify the signaling transport mechanism used in the control channel, however it states that messages transported over a control channel must be IP encoded. Furthermore, since the messages are IP encoded, the link level encoding is not part of LMP. A 32-bit non-zero integer Control Channel Identifier (CCId) is assigned to each direction of a control channel.

Each control channel individually negotiates its control channel parameters and maintains connectivity using a fast Hello protocol. The latter is required if lower-level mechanisms are not available to detect link failures.

The Hello protocol of LMP is intended to be a lightweight keep-alive mechanism that will react to control channel failures rapidly so that IGP Hellos are not lost and the associated link-state adjacencies are not removed uselessly.

The Hello protocol consists of two phases: a negotiation phase and a keep-alive phase. The negotiation phase allows negotiation of some basic Hello protocol parameters, like the Hello frequency. The keep-alive phase consists of a fast lightweight bi-directional Hello message exchange.

If a group of control channels share a common node pair and support the same LMP capabilities, then LMP control channel messages (except Configuration messages, and Hello's) may be transmitted over any of the active control channels without coordination between the local and remote nodes.

For LMP, it is essential that at least one control channel is always available. In case of control channel failure, it may be possible to use an alternate active control channel without coordination.

6.2. Link Property Correlation

As part of LMP, a link property correlation exchange is defined. The exchange is used to aggregate multiple data-bearing links (i.e., component links) into a bundled link and exchange, correlate, or change TE link parameters. The link property correlation exchange may be done at any time a link is up and not in the Verification process (see next section).

It allows, for instance, the addition of component links to a link bundle, change of a link's minimum/maximum reservable bandwidth, change of port identifiers, or change of component identifiers in a bundle. This mechanism is supported by an exchange of link summary messages.

6.3. Link Connectivity Verification

Link connectivity verification is an optional procedure that may be used to verify the physical connectivity of data-bearing links as well as to exchange the link identifiers that are used in the GMPLS signaling.

This procedure should be performed initially when a data-bearing link is first established, and subsequently, on a periodic basis for all unallocated (free) data-bearing links.

The verification procedure consists of sending Test messages in-band over the data-bearing links. This requires that the unallocated links must be opaque; however, multiple degrees of opaqueness (e.g., examining overhead bytes, terminating the payload, etc.), and hence different mechanisms to transport the Test messages, are specified. Note that the Test message is the only LMP message that is transmitted over the data-bearing link, and that Hello messages continue to be exchanged over the control channel during the link verification process. Data-bearing links are tested in the transmit direction as they are unidirectional. As such, it is possible for LMP neighboring nodes to exchange the Test messages simultaneously in both directions.

To initiate the link verification procedure, a node must first notify the adjacent node that it will begin sending Test messages over a particular data-bearing link, or over the component links of a particular bundled link. The node must also indicate the number of data-bearing links that are to be verified; the interval at which the test messages will be sent; the encoding scheme, the transport mechanisms that are supported, the data rate for Test messages; and, in the case where the data-bearing links correspond to fibers, the wavelength over which the Test messages will be transmitted. Furthermore, the local and remote bundled link identifiers are transmitted at this time to perform the component link association with the bundled link identifiers.

6.4. Fault Management

Fault management is an important requirement from the operational point of view. Fault management includes usually: fault detection, fault localization and fault notification. When a failure occurs and is detected (fault detection), an operator needs to know exactly where it happened (fault localization) and a source node may need to be notified in order to take some actions (fault notification).

Note that fault localization can also be used to support some specific (local) protection/restoration mechanisms.

In new technologies such as transparent photonic switching currently no method is defined to locate a fault, and the mechanism by which the fault information is propagated must be sent "out of band" (via the control plane).

LMP provides a fault localization procedure that can be used to rapidly localize link failures, by notifying a fault up to the node upstream of that fault (i.e., through a fault notification procedure).

A downstream LMP neighbor that detects data link failures will send an LMP message to its upstream neighbor notifying it of the failure. When an upstream node receives a failure notification, it can correlate the failure with the corresponding input ports to determine if the failure is between the two nodes. Once the failure has been localized, the signaling protocols can be used to initiate link or path protection/restoration procedures.

6.5. LMP for DWDM Optical Line Systems (OLSs)

In an all-optical environment, LMP focuses on peer communications (e.g., OXC-to-OXC). A great deal of information about a link between two OXCs is known by the OLS (Optical Line System or WDM Terminal multiplexer). Exposing this information to the control plane can improve network usability by further reducing required manual configuration, and by greatly enhancing fault detection and recovery.

LMP-WDM [LMP-WDM] defines extensions to LMP for use between an OXC and an OLS. These extensions are intended to satisfy the Optical Link Interface Requirements described in [OLI-REQ].

Fault detection is particularly an issue when the network is using all-optical photonic switches (PXC). Once a connection is established, PXCs have only limited visibility into the health of the connection. Although the PXC is all-optical, long-haul OLSs typically terminate channels electrically and regenerate them optically. This provides an opportunity to monitor the health of a channel between PXCs. LMP-WDM can then be used by the OLS to provide this information to the PXC.

In addition to the link information known to the OLS that is exchanged through LMP-WDM, some information known to the OXC may also be exchanged with the OLS through LMP-WDM. This information is useful for alarm management and link monitoring (e.g., trace monitoring). Alarm management is important because the administrative state of a connection, known to the OXC (e.g., this information may be learned through the Admin Status object of GMPLS signaling [RFC3471]), can be used to suppress spurious alarms. For example, the OXC may know that a connection is "up", "down", in a "testing" mode, or being deleted ("deletion-in-progress"). The OXC can use this information to inhibit alarm reporting from the OLS when a connection is "down", "testing", or being deleted.

It is important to note that an OXC may peer with one or more OLSs and an OLS may peer with one or more OXCs. Although there are many similarities between an OXC-OXC LMP session and an OXC-OLS LMP session, particularly for control management and link verification, there are some differences as well. These differences can primarily be attributed to the nature of an OXC-OLS link, and the purpose of OXC-OLS LMP sessions. The OXC-OXC links can be used to provide the basis for GMPLS signaling and routing at the optical layer. The information exchanged over LMP-WDM sessions is used to augment knowledge about the links between OXCs.

In order for the information exchanged over the OXC-OLS LMP sessions to be used by the OXC-OXC session, the information must be coordinated by the OXC. However, the OXC-OXC and OXC-OLS LMP sessions are run independently and must be maintained separately. One critical requirement when running an OXC-OLS LMP session is the ability of the OLS to make a data link transparent when not doing the verification procedure. This is because the same data link may be verified between OXC-OLS and between OXC-OXC. The verification procedure of LMP is used to coordinate the Test procedure (and hence the transparency/opaqueness of the data links). To maintain independence between the sessions, it must be possible for the LMP sessions to come up in any order. In particular, it must be possible for an OXC-OXC LMP session to come up without an OXC-OLS LMP session being brought up, and vice-versa.

7. Generalized Signaling

The GMPLS signaling extends certain base functions of the RSVP-TE and CR-LDP signaling and, in some cases, adds functionality. These changes and additions impact basic LSP properties: how labels are requested and communicated, the unidirectional nature of LSPs, how errors are propagated, and information provided for synchronizing the ingress and egress.

The core GMPLS signaling specification is available in three parts:

1. A signaling functional description [RFC3471].
2. RSVP-TE extensions [RFC3473].
3. CR-LDP extensions [RFC3472].

In addition, independent parts are available per technology:

1. GMPLS extensions for SONET and SDH control [RFC3946].
2. GMPLS extensions for G.709 control [GMPLS-G709].

The following MPLS profile expressed in terms of MPLS features [RFC3031] applies to GMPLS:

- Downstream-on-demand label allocation and distribution.

- Ingress initiated ordered control.

- Liberal (typical), or conservative (could) label retention mode.

- Request, traffic/data, or topology driven label allocation strategy.

- Explicit routing (typical), or hop-by-hop routing.

The GMPLS signaling defines the following new building blocks on the top of MPLS-TE:

1. A new generic label request format.
2. Labels for TDM, LSC and FSC interfaces, generically known as Generalized Label.
3. Waveband switching support.
4. Label suggestion by the upstream for optimization purposes (e.g., latency).
5. Label restriction by the upstream to support some optical constraints.
6. Bi-directional LSP establishment with contention resolution.
7. Rapid failure notification extensions.
8. Protection information currently focusing on link protection, plus primary and secondary LSP indication.
9. Explicit routing with explicit label control for a fine degree of control.
10. Specific traffic parameters per technology.
11. LSP administrative status handling.
12. Control channel separation.

These building blocks will be described in more details in the following. A complete specification can be found in the corresponding documents.

Note that GMPLS is highly generic and has many options. Only building blocks 1, 2 and 10 are mandatory, and only within the specific format that is needed. Typically, building blocks 6 and 9 should be implemented. Building blocks 3, 4, 5, 7, 8, 11 and 12 are optional.

A typical SONET/SDH switching network would implement building blocks: 1, 2 (the SONET/SDH label), 6, 9, 10 and 11. Building blocks 7 and 8 are optional since the protection can be achieved using SONET/SDH overhead bytes.

A typical wavelength switching network would implement building blocks: 1, 2 (the generic format), 4, 5, 6, 7, 8, 9 and 11. Building block 3 is only needed in the particular case of waveband switching.

A typical fiber switching network would implement building blocks: 1, 2 (the generic format), 6, 7, 8, 9 and 11.

A typical MPLS-IP network would not implement any of these building blocks, since the absence of building block 1 would indicate regular MPLS-IP. Note however that building block 1 and 8 can be used to signal MPLS-IP as well. In that case, the MPLS-IP network can benefit from the link protection type (not available in CR-LDP, some very basic form being available in RSVP-TE). Building block 2 is here a regular MPLS label and no new label format is required.

GMPLS does not specify any profile for RSVP-TE and CR-LDP implementations that have to support GMPLS - except for what is directly related to GMPLS procedures. It is to the manufacturer to decide which are the optional elements and procedures of RSVP-TE and CR-LDP that need to be implemented. Some optional MPLS-TE elements can be useful for TDM, LSC and FSC layers, for instance the setup and holding priorities that are inherited from MPLS-TE.

7.1. Overview: How to Request an LSP

A TDM, LSC or FSC LSP is established by sending a PATH/Label Request message downstream to the destination. This message contains a Generalized Label Request with the type of LSP (i.e., the layer concerned), and its payload type. An Explicit Route Object (ERO) is also normally added to the message, but this can be added and/or completed by the first/default LSR.

The requested bandwidth is encoded in the RSVP-TE SENDER_TSPEC object, or in the CR-LDP Traffic Parameters TLV. Specific parameters for a given technology are given in these traffic parameters, such as the type of signal, concatenation and/or transparency for a SONET/SDH LSP. For some other technology there be could just one bandwidth parameter indicating the bandwidth as a floating-point value.

The requested local protection per link may be requested using the Protection Information Object/TLV. The end-to-end LSP protection is for further study and is introduced LSP protection/restoration section (see after).

If the LSP is a bi-directional LSP, an Upstream Label is also specified in the Path/Label Request message. This label will be the one to use in the upstream direction.

Additionally, a Suggested Label, a Label Set and a Waveband Label can also be included in the message. Other operations are defined in MPLS-TE.

The downstream node will send back a Resv/Label Mapping message including one Generalized Label object/TLV that can contain several Generalized Labels. For instance, if a concatenated SONET/SDH signal is requested, several labels can be returned.

In case of SONET/SDH virtual concatenation, a list of labels is returned. Each label identifying one element of the virtual concatenated signal. This limits virtual concatenation to remain within a single (component) link.

In case of any type of SONET/SDH contiguous concatenation, only one label is returned. That label is the lowest signal of the contiguous concatenated signal (given an order specified in [RFC3946]).

In case of SONET/SDH "multiplication", i.e., co-routing of circuits of the same type but without concatenation but all belonging to the same LSP, the explicit ordered list of all signals that take part in the LSP is returned.

7.2. Generalized Label Request

The Generalized Label Request is a new object/TLV to be added in an RSVP-TE Path message instead of the regular Label Request, or in a CR-LDP Request message in addition to the already existing TLVs. Only one label request can be used per message, so a single LSP can be requested at a time per signaling message.

The Generalized Label Request gives three major characteristics (parameters) required to support the LSP being requested: the LSP Encoding Type, the Switching Type that must be used and the LSP payload type called Generalized PID (G-PID).

The LSP Encoding Type indicates the encoding type that will be used with the data associated with the LSP, i.e., the type of technology being considered. For instance, it can be SDH, SONET, Ethernet, ANSI PDH, etc. It represents the nature of the LSP, and not the nature of the links that the LSP traverses. This is used hop-by-hop by each node.

A link may support a set of encoding formats, where support means that a link is able to carry and switch a signal of one or more of these encoding formats. The Switching Type indicates then the type of switching that should be performed on a particular link for that LSP. This information is needed for links that advertise more than one type of switching capability.

Nodes must verify that the type indicated in the Switching Type is supported on the corresponding incoming interface; otherwise, the node must generate a notification message with a "Routing problem/Switching Type" indication.

The LSP payload type (G-PID) identifies the payload carried by the LSP, i.e., an identifier of the client layer of that LSP. For some technologies, it also indicates the mapping used by the client layer, e.g., byte synchronous mapping of E1. This must be interpreted according to the LSP encoding type and is used by the nodes at the endpoints of the LSP to know to which client layer a request is destined, and in some cases by the penultimate hop.

Other technology specific parameters are not transported in the Generalized Label Request but in technology specific traffic parameters as explained hereafter. Currently, two set of traffic parameters are defined, one for SONET/SDH and one for G.709.

Note that it is expected than specific traffic parameters will be defined in the future for photonic (all optical) switching.

7.3. SONET/SDH Traffic Parameters

The GMPLS SONET/SDH traffic parameters [RFC3946] specify a powerful set of capabilities for SONET [ANSI-T1.105] and SDH [ITUT-G.707].

The first traffic parameter specifies the type of the elementary SONET/SDH signal that comprises the requested LSP, e.g., VC-11, VT6, VC-4, STS-3c, etc. Several transforms can then be applied successively on the elementary Signal to build the final signal being actually requested for the LSP.

These transforms are the contiguous concatenation, the virtual concatenation, the transparency and the multiplication. Each one is optional. They must be applied strictly in the following order:

- First, contiguous concatenation can be optionally applied on the Elementary Signal, resulting in a contiguously concatenated signal.

- Second, virtual concatenation can be optionally applied either directly on the elementary Signal, or on the contiguously concatenated signal obtained from the previous phase.

- Third, some transparency can be optionally specified when requesting a frame as signal rather than a container. Several transparency packages are defined.

- Fourth, a multiplication can be optionally applied either directly on the elementary Signal, or on the contiguously concatenated signal obtained from the first phase, or on the virtually concatenated signal obtained from the second phase, or on these signals combined with some transparency.

For RSVP-TE, the SONET/SDH traffic parameters are carried in a new SENDER_TSPEC and FLOWSPEC. The same format is used for both. There is no Adspec associated with the SENDER_TSPEC, it is omitted or a default value is used. The content of the FLOWSPEC object received in a Resv message should be identical to the content of the SENDER_TSPEC of the corresponding Path message. In other words, the receiver is normally not allowed to change the values of the traffic parameters. However, some level of negotiation may be achieved as explained in [RFC3946].

For CR-LDP, the SONET/SDH traffic parameters are simply carried in a new TLV.

Note that a general discussion on SONET/SDH and GMPLS can be found in [SONET-SDH-GMPLS-FRM].

7.4. G.709 Traffic Parameters

Simply said, an [ITUT-G.709] based network is decomposed in two major layers: an optical layer (i.e., made of wavelengths) and a digital layer. These two layers are divided into sub-layers and switching occurs at two specific sub-layers: at the OCh (Optical Channel) optical layer and at the ODU (Optical channel Data Unit) electrical layer. The ODUk notation is used to denote ODUs at different bandwidths.

The GMPLS G.709 traffic parameters [GMPLS-G709] specify a powerful set of capabilities for ITU-T G.709 networks.

The first traffic parameter specifies the type of the elementary G.709 signal that comprises the requested LSP, e.g., ODU1, OCh at 40 Gbps, etc. Several transforms can then be applied successively on the elementary Signal to build the final signal being actually requested for the LSP.

These transforms are the virtual concatenation and the multiplication. Each one of these transforms is optional. They must be applied strictly in the following order:

- First, virtual concatenation can be optionally applied directly on the elementary Signal,

- Second, a multiplication can be optionally applied, either directly on the elementary Signal, or on the virtually concatenated signal obtained from the first phase.

Additional ODUk Multiplexing traffic parameters allow indicating an ODUk mapping (ODUj into ODUk) for an ODUk multiplexing LSP request. G.709 supports the following multiplexing capabilities: ODUj into ODUk (k > j) and ODU1 with ODU2 multiplexing into ODU3.

For RSVP-TE, the G.709 traffic parameters are carried in a new SENDER-TSPEC and FLOWSPEC. The same format is used for both. There is no Adspec associated with the SENDER_TSPEC, it is omitted or a default value is used. The content of the FLOWSPEC object received in a Resv message should be identical to the content of the SENDER_TSPEC of the corresponding Path message.

For CR-LDP, the G.709 traffic parameters are simply carried in a new TLV.

7.5. Bandwidth Encoding

Some technologies that do not have (yet) specific traffic parameters just require a bandwidth encoding transported in a generic form. Bandwidth is carried in 32-bit number in IEEE floating-point format (the unit is bytes per second). Values are carried in a per protocol specific manner. For non-packet LSPs, it is useful to define discrete values to identify the bandwidth of the LSP.

It should be noted that this bandwidth encoding do not apply to SONET/SDH and G.709, for which the traffic parameters fully define the requested SONET/SDH or G.709 signal.

The bandwidth is coded in the Peak Data Rate field of Int-Serv objects for RSVP-TE in the SENDER_TSPEC and FLOWSPEC objects and in the Peak and Committed Data Rate fields of the CR-LDP Traffic Parameters TLV.

7.6. Generalized Label

The Generalized Label extends the traditional MPLS label by allowing the representation of not only labels that travel in-band with associated data packets, but also (virtual) labels that identify time-slots, wavelengths, or space division multiplexed positions.

For example, the Generalized Label may identify (a) a single fiber in a bundle, (b) a single waveband within fiber, (c) a single wavelength within a waveband (or fiber), or (d) a set of time-slots within a wavelength (or fiber). It may also be a generic MPLS label, a Frame Relay label, or an ATM label (VCI/VPI). The format of a label can be as simple as an integer value such as a wavelength label or can be more elaborated such as an SONET/SDH or a G.709 label.

SDH and SONET define each a multiplexing structure. These multiplexing structures will be used as naming trees to create unique labels. Such a label will identify the exact position (times-lot(s)) of a signal in a multiplexing structure. Since the SONET multiplexing structure may be seen as a subset of the SDH multiplexing structure, the same format of label is used for SDH and SONET. A similar concept is applied to build a label at the G.709 ODU layer.

Since the nodes sending and receiving the Generalized Label know what kinds of link they are using, the Generalized Label does not identify its type. Instead, the nodes are expected to know from the context what type of label to expect.

A Generalized Label only carries a single level of label i.e., it is non-hierarchical. When multiple levels of labels (LSPs within LSPs) are required, each LSP must be established separately.

7.7. Waveband Switching

A special case of wavelength switching is waveband switching. A waveband represents a set of contiguous wavelengths, which can be switched together to a new waveband. For optimization reasons, it may be desirable for a photonic cross-connect to optically switch multiple wavelengths as a unit. This may reduce the distortion on the individual wavelengths and may allow tighter separation of the individual wavelengths. A Waveband label is defined to support this special case.

Waveband switching naturally introduces another level of label hierarchy and as such the waveband is treated the same way, all other upper layer labels are treated. As far as the MPLS protocols are concerned, there is little difference between a waveband label and a

wavelength label. Exception is that semantically the waveband can be subdivided into wavelengths whereas the wavelength can only be subdivided into time or statistically multiplexed labels.

In the context of waveband switching, the generalized label used to indicate a waveband contains three fields, a waveband ID, a Start Label and an End Label. The Start and End Labels are channel identifiers from the sender perspective that identify respectively, the lowest value wavelength and the highest value wavelength making up the waveband.

7.8. Label Suggestion by the Upstream

GMPLS allows for a label to be optionally suggested by an upstream node. This suggestion may be overridden by a downstream node but in some cases, at the cost of higher LSP setup time. The suggested label is valuable when establishing LSPs through certain kinds of optical equipment where there may be a lengthy (in electrical terms) delay in configuring the switching fabric. For example, micro mirrors may have to be elevated or moved, and this physical motion and subsequent damping takes time. If the labels and hence switching fabric are configured in the reverse direction (the norm), the Resv/MAPPING message may need to be delayed by 10's of milliseconds per hop in order to establish a usable forwarding path. It can be important for restoration purposes where alternate LSPs may need to be rapidly established as a result of network failures.

7.9. Label Restriction by the Upstream

An upstream node can optionally restrict (limit) the choice of label of a downstream node to a set of acceptable labels. Giving lists and/or ranges of inclusive (acceptable) or exclusive (unacceptable) labels in a Label Set provides this restriction. If not applied, all labels from the valid label range may be used. There are at least four cases where a label restriction is useful in the "optical" domain.

Case 1: the end equipment is only capable of transmitting and receiving on a small specific set of wavelengths/wavebands.

Case 2: there is a sequence of interfaces, which cannot support wavelength conversion and require the same wavelength be used end-to-end over a sequence of hops, or even an entire path.

Case 3: it is desirable to limit the amount of wavelength conversion being performed to reduce the distortion on the optical signals.

Case 4: two ends of a link support different sets of wavelengths.

The receiver of a Label Set must restrict its choice of labels to one that is in the Label Set. A Label Set may be present across multiple hops. In this case, each node generates its own outgoing Label Set, possibly based on the incoming Label Set and the node's hardware capabilities. This case is expected to be the norm for nodes with conversion incapable interfaces.

7.10. Bi-directional LSP

GMPLS allows establishment of bi-directional symmetric LSPs (not of asymmetric LSPs). A symmetric bi-directional LSP has the same traffic engineering requirements including fate sharing, protection and restoration, LSRs, and resource requirements (e.g., latency and jitter) in each direction.

In the remainder of this section, the term "initiator" is used to refer to a node that starts the establishment of an LSP; the term "terminator" is used to refer to the node that is the target of the LSP. For a bi-directional LSPs, there is only one initiator and one terminator.

Normally to establish a bi-directional LSP when using RSVP-TE [RFC3209] or CR-LDP [RFC3212] two unidirectional paths must be independently established. This approach has the following disadvantages:

1. The latency to establish the bi-directional LSP is equal to one round trip signaling time plus one initiator-terminator signaling transit delay. This not only extends the setup latency for successful LSP establishment, but it extends the worst-case latency for discovering an unsuccessful LSP to as much as two times the initiator-terminator transit delay. These delays are particularly significant for LSPs that are established for restoration purposes.

2. The control overhead is twice that of a unidirectional LSP. This is because separate control messages (e.g., Path and Resv) must be generated for both segments of the bi-directional LSP.

3. Because the resources are established in separate segments, route selection is complicated. There is also additional potential race for conditions in assignment of resources, which decreases the overall probability of successfully establishing the bi-directional connection.

4. It is more difficult to provide a clean interface for SONET/SDH equipment that may rely on bi-directional hop-by-hop paths for protection switching. Note that existing SONET/SDH equipment transmits the control information in-band with the data.

5. Bi-directional optical LSPs (or lightpaths) are seen as a requirement for many optical networking service providers.

With bi-directional LSPs both the downstream and upstream data paths, i.e., from initiator to terminator and terminator to initiator, are established using a single set of signaling messages. This reduces the setup latency to essentially one initiator-terminator round trip time plus processing time, and limits the control overhead to the same number of messages as a unidirectional LSP.

For bi-directional LSPs, two labels must be allocated. Bi-directional LSP setup is indicated by the presence of an Upstream Label in the appropriate signaling message.

7.11. Bi-directional LSP Contention Resolution

Contention for labels may occur between two bi-directional LSP setup requests traveling in opposite directions. This contention occurs when both sides allocate the same resources (ports) at effectively the same time. GMPLS signaling defines a procedure to resolve that contention: the node with the higher node ID will win the contention. To reduce the probability of contention, some mechanisms are also suggested.

7.12. Rapid Notification of Failure

GMPLS defines several signaling extensions that enable expedited notification of failures and other events to nodes responsible for restoring failed LSPs, and error handling.

1. Acceptable Label Set for notification on Label Error:

   There are cases in traditional MPLS and in GMPLS that result in an error message containing an "Unacceptable label value" indication. When these cases occur, it can useful for the node generating the error message to indicate which labels would be acceptable. To cover this case, GMPLS introduces the ability to convey such information via the "Acceptable Label Set". An Acceptable Label Set is carried in appropriate protocol specific error messages. The format of an Acceptable Label Set is identical to a Label Set.

2. Expedited notification:

   Extensions to RSVP-TE enable expedited notification of failures and other events to determined nodes. For CR-LDP, there is not currently a similar mechanism. The first extension identifies where event notifications are to be sent. The second provides for general expedited event notification with a Notify message. Such extensions can be used by fast restoration mechanisms. Notifications may be requested in both the upstream and downstream directions.

   The Notify message is a generalized notification mechanism that differs from the currently defined error messages in that it can be "targeted" to a node other than the immediate upstream or downstream neighbor. The Notify message does not replace existing error messages. The Notify message may be sent either (a) normally, where non-target nodes just forward the Notify message to the target node, similar to ResvConf processing in [RFC2205]; or (b) encapsulated in a new IP header whose destination is equal to the target IP address.

3. Faster removal of intermediate states:

   A specific RSVP optimization allowing in some cases the faster removal of intermediate states. This extension is used to deal with specific RSVP mechanisms.

7.13. Link Protection

Protection information is carried in the new optional Protection Information Object/TLV. It currently indicates the desired link protection for each link of an LSP. If a particular protection type, i.e., 1+1, or 1:N, is requested, then a connection request is processed only if the desired protection type can be honored. Note that GMPLS advertises the protection capabilities of a link in the routing protocols. Path computation algorithms may consider this information when computing paths for setting up LSPs.

Protection information also indicates if the LSP is a primary or secondary LSP. A secondary LSP is a backup to a primary LSP. The resources of a secondary LSP are normally not used until the primary LSP fails, but they may be used by other LSPs until the primary LSP fails over the secondary LSP. At that point, any LSP that is using the resources for the secondary LSP must be preempted.

Six link protection types are currently defined as individual flags and can be combined: enhanced, dedicated 1+1, dedicated 1:1, shared, unprotected, extra traffic. See [RFC3471] section 7.1 for a precise definition of each.

7.14. Explicit Routing and Explicit Label Control

By using an explicit route, the path taken by an LSP can be controlled more or less precisely. Typically, the node at the head-end of an LSP finds an explicit route and builds an Explicit Route Object (ERO)/ Explicit Route (ER) TLV that contains that route. Possibly, the edge node does not build any explicit route, and just transmit a signaling request to a default neighbor LSR (as IP/MPLS hosts would). For instance, an explicit route could be added to a signaling message by the first switching node, on behalf of the edge node. Note also that an explicit route is altered by intermediate LSRs during its progression towards the destination.

The explicit route is originally defined by MPLS-TE as a list of abstract nodes (i.e., groups of nodes) along the explicit route. Each abstract node can be an IPv4 address prefix, an IPv6 address prefix, or an AS number. This capability allows the generator of the explicit route to have incomplete information about the details of the path. In the simplest case, an abstract node can be a full IP address (32 bits) that identifies a specific node (called a simple abstract node).

MPLS-TE allows strict and loose abstract nodes. The path between a strict node and its preceding node must include only network nodes from the strict node and its preceding abstract node. The path between a loose node and its preceding abstract node may include other network nodes that are not part of the loose node or its preceding abstract node.

This explicit route was extended to include interface numbers as abstract nodes to support unnumbered interfaces; and further extended by GMPLS to include labels as abstract nodes. Having labels in an explicit route is an important feature that allows controlling the placement of an LSP with a very fine granularity. This is more likely to be used for TDM, LSC and FSC links.

In particular, the explicit label control in the explicit route allows terminating an LSP on a particular outgoing port of an egress node. Indeed, a label sub-object/TLV must follow a sub-object/TLV containing the IP address, or the interface identifier (in case of unnumbered interface), associated with the link on which it is to be used.

This can also be used when it is desirable to "splice" two LSPs together, i.e., where the tail of the first LSP would be "spliced" into the head of the second LSP.

When used together with an optimization algorithm, it can provide very detailed explicit routes, including the label (timeslot) to use on a link, in order to minimize the fragmentation of the SONET/SDH multiplex on the corresponding interface.

7.15. Route Recording

In order to improve the reliability and the manageability of the LSP being established, the concept of the route recording was introduced in RSVP-TE to function as:

- First, a loop detection mechanism to discover L3 routing loops, or loops inherent in the explicit route (this mechanism is strictly exclusive with the use of explicit routing objects).

- Second, a route recording mechanism collects up-to-date detailed path information on a hop-by-hop basis during the LSP setup process. This mechanism provides valuable information to the source and destination nodes. Any intermediate routing change at setup time, in case of loose explicit routing, will be reported.

- Third, a recorded route can be used as input for an explicit route. This is useful if a source node receives the recorded route from a destination node and applies it as an explicit route in order to "pin down the path".

Within the GMPLS architecture, only the second and third functions are mainly applicable for TDM, LSC and FSC layers.

7.16. LSP Modification and LSP Re-routing

LSP modification and re-routing are two features already available in MPLS-TE. GMPLS does not add anything new. Elegant re-routing is possible with the concept of "make-before-break" whereby an old path is still used while a new path is set up by avoiding double reservation of resources. Then, the node performing the re-routing can swap on the new path and close the old path. This feature is supported with RSVP-TE (using shared explicit filters) and CR-LDP (using the action indicator flag).

LSP modification consists in changing some LSP parameters, but normally without changing the route. It is supported using the same mechanism as re-routing. However, the semantic of LSP modification will differ from one technology to the other. For instance, further

studies are required to understand the impact of dynamically changing some SONET/SDH circuit characteristics such as the bandwidth, the protection type, the transparency, the concatenation, etc.

7.17. LSP Administrative Status Handling

GMPLS provides the optional capability to indicate the administrative status of an LSP by using a new Admin Status object/TLV. Administrative Status information is currently used in two ways.

In the first usage, the Admin Status object/TLV is carried in a Path/Label Request or Resv/Label Mapping message to indicate the administrative state of an LSP. In this usage, Administrative Status information indicates the state of the LSP, which include "up" or "down", if it in a "testing" mode, and if deletion is in progress.

Based on that administrative status, a node can take local decisions, like inhibit alarm reporting when an LSP is in "down" or "testing" states, or report alarms associated with the connection at a priority equal to or less than "Non service affecting".

It is possible that some nodes along an LSP will not support the Admin Status Object/TLV. In the case of a non-supporting transit node, the object will pass through the node unmodified and normal processing can continue.

In some circumstances, particularly optical networks, it is useful to set the administrative status of an LSP to "being deleted" before tearing it down in order to avoid non-useful generation of alarms. The ingress LSR precedes an LSP deletion by inserting an appropriate Admin Status Object/TLV in a Path/Label Request (with the modification action indicator flag set to modify) message. Transit LSRs process the Admin Status Object/TLV and forward it. The egress LSR answers in a Resv/Label Mapping (with the modification action indicator flag set to modify) message with the Admin Status object. Upon receiving this message and object, the ingress node sends a PathTear/Release message downstream to remove the LSP and normal RSVP-TE/CR-LDP processing takes place.

In the second usage, the Admin Status object/TLV is carried in a Notification/Label Mapping (with the modification action indicator flag set to modify) message to request that the ingress node change the administrative state of an LSP. This allows intermediate and egress nodes triggering the setting of administrative status. In particular, this allows intermediate or egress LSRs requesting a release of an LSP initiated by the ingress node.

7.18. Control Channel Separation

In GMPLS, a control channel be separated from the data channel. Indeed, the control channel can be implemented completely out-of-band for various reason, e.g., when the data channel cannot carry in-band control information. This issue was even originally introduced to MPLS in the context of link bundling.

In traditional MPLS, there is an implicit one-to-one association of a control channel to a data channel. When such an association is present, no additional or special information is required to associate a particular LSP setup transaction with a particular data channel.

Otherwise, it is necessary to convey additional information in signaling to identify the particular data channel being controlled. GMPLS supports explicit data channel identification by providing interface identification information. GMPLS allows the use of a number of interface identification schemes including IPv4 or IPv6 addresses, interface indexes (for unnumbered interfaces) and component interfaces (for bundled interfaces), unnumbered bundled interfaces are also supported.

The choice of the data interface to use is always made by the sender of the Path/Label Request message, and indicated by including the data channel's interface identifier in the message using a new RSVP_HOP object sub-type/Interface TLV.

For bi-directional LSPs, the sender chooses the data interface in each direction. In all cases but bundling, the upstream interface is implied by the downstream interface. For bundling, the Path/Label Request sender explicitly identifies the component interface used in each direction. The new object/TLV is used in Resv/Label Mapping message to indicate the downstream node's usage of the indicated interface(s).

The new object/TLV can contain a list of embedded TLVs, each embedded TLV can be an IPv4 address, and IPv6 address, an interface index, a downstream component interface ID or an upstream component interface ID. In the last three cases, the embedded TLV contains itself an IP address plus an Interface ID, the IP address being used to identify the interface ID (it can be the router ID for instance).

There are cases where it is useful to indicate a specific interface associated with an error. To support these cases the IF_ID ERROR_SPEC RSVP Objects are defined.

8. Forwarding Adjacencies (FA)

To improve scalability of MPLS TE (and thus GMPLS) it may be useful to aggregate multiple TE LSPs inside a bigger TE LSP. Intermediate nodes see the external LSP only. They do not have to maintain forwarding states for each internal LSP, less signaling messages need to be exchanged and the external LSP can be somehow protected instead (or in addition) to the internal LSPs. This can considerably increase the scalability of the signaling.

The aggregation is accomplished by (a) an LSR creating a TE LSP, (b) the LSR forming a forwarding adjacency out of that LSP (advertising this LSP as a Traffic Engineering (TE) link into IS-IS/OSPF), (c) allowing other LSRs to use forwarding adjacencies for their path computation, and (d) nesting of LSPs originated by other LSRs into that LSP (e.g., by using the label stack construct in the case of IP).

ISIS/OSPF floods the information about "Forwarding Adjacencies" FAs just as it floods the information about any other links. Consequently to this flooding, an LSR has in its TE link state database the information about not just conventional links, but FAs as well.

An LSR, when performing path computation, uses not just conventional links, but FAs as well. Once a path is computed, the LSR uses RSVP-TE/CR-LDP for establishing label binding along the path. FAs need simple extensions to signaling and routing protocols.

8.1. Routing and Forwarding Adjacencies

Forwarding adjacencies may be represented as either unnumbered or numbered links. A FA can also be a bundle of LSPs between two nodes.

FAs are advertised as GMPLS TE links such as defined in [HIERARCHY]. GMPLS TE links are advertised in OSPF and IS-IS such as defined in [OSPF-TE-GMPLS] and [ISIS-TE-GMPLS]. These last two specifications enhance [OSPF-TE] and [ISIS-TE] that defines a base TE link.

When a FA is created dynamically, its TE attributes are inherited from the FA-LSP that induced its creation. [HIERARCHY] specifies how each TE parameter of the FA is inherited from the FA-LSP. Note that the bandwidth of the FA must be at least as big as the FA-LSP that induced it, but may be bigger if only discrete bandwidths are available for the FA-LSP. In general, for dynamically provisioned forwarding adjacencies, a policy-based mechanism may be needed to associate attributes to forwarding adjacencies.

A FA advertisement could contain the information about the path taken by the FA-LSP associated with that FA. Other LSRs may use this information for path computation. This information is carried in a new OSPF and IS-IS TLV called the Path TLV.

It is possible that the underlying path information might change over time, via configuration updates, or dynamic route modifications, resulting in the change of that TLV.

If forwarding adjacencies are bundled (via link bundling), and if the resulting bundled link carries a Path TLV, the underlying path followed by each of the FA-LSPs that form the component links must be the same.

It is expected that forwarding adjacencies will not be used for establishing IS-IS/OSPF peering relation between the routers at the ends of the adjacency.

LSP hierarchy could exist both with the peer and with the overlay models. With the peer model, the LSP hierarchy is realized via FAs and an LSP is both created and used as a TE link by exactly the same instance of the control plane. Creating LSP hierarchies with overlays does not involve the concept of FA. With the overlay model an LSP created (and maintained) by one instance of the GMPLS control plane is used as a TE link by another instance of the GMPLS control plane. Moreover, the nodes using a TE link are expected to have a routing and signaling adjacency.

8.2. Signaling Aspects

For the purpose of processing the explicit route in a Path/Request message of an LSP that is to be tunneled over a forwarding adjacency, an LSR at the head-end of the FA-LSP views the LSR at the tail of that FA-LSP as adjacent (one IP hop away).

8.3. Cascading of Forwarding Adjacencies

With an integrated model, several layers are controlled using the same routing and signaling protocols. A network may then have links with different multiplexing/demultiplexing capabilities. For example, a node may be able to multiplex/demultiplex individual packets on a given link, and may be able to multiplex/demultiplex channels within a SONET payload on other links.

A new OSPF and IS-IS sub-TLV has been defined to advertise the multiplexing capability of each interface: PSC, L2SC, TDM, LSC or FSC. This sub-TLV is called the Interface Switching Capability Descriptor sub-TLV, which complements the sub-TLVs defined in

[OSPF-TE-GMPLS] and [ISIS-TE-GMPLS]. The information carried in this sub-TLV is used to construct LSP regions, and determine region's boundaries.

Path computation may take into account region boundaries when computing a path for an LSP. For example, path computation may restrict the path taken by an LSP to only the links whose multiplexing/demultiplexing capability is PSC. When an LSP need to cross a region boundary, it can trigger the establishment of an FA at the underlying layer (i.e., the L2SC layer). This can trigger a cascading of FAs between layers with the following obvious order: L2SC, then TDM, then LSC, and then finally FSC.

9. Routing and Signaling Adjacencies

By definition, two nodes have a routing (IS-IS/OSPF) adjacency if they are neighbors in the IS-IS/OSPF sense.

By definition, two nodes have a signaling (RSVP-TE/CR-LDP) adjacency if they are neighbors in the RSVP-TE/CR-LDP sense. Nodes A and B are RSVP-TE neighbors if they directly exchange RSVP-TE messages (Path/Resv) (e.g., as described in sections 7.1.1 and 7.1.2 of [HIERARCHY]). The neighbor relationship includes exchanging RSVP-TE Hellos.

By definition, a Forwarding Adjacency (FA) is a TE Link between two GMPLS nodes whose path transits one or more other (G)MPLS nodes in the same instance of the (G)MPLS control plane. If two nodes have one or more non-FA TE Links between them, these two nodes are expected (although not required) to have a routing adjacency. If two nodes do not have any non-FA TE Links between them, it is expected (although not required) that these two nodes would not have a routing adjacency. To state the obvious, if the TE links between two nodes are to be used for establishing LSPs, the two nodes must have a signaling adjacency.

If one wants to establish routing and/or signaling adjacency between two nodes, there must be an IP path between them. This IP path can be, for example, a TE Link with an interface switching capability of PSC, anything that looks likes an IP link (e.g., GRE tunnel, or a (bi-directional) LSP that with an interface switching capability of PSC).

A TE link may not be capable of being used directly for maintaining routing and/or signaling adjacencies. This is because GMPLS routing and signaling adjacencies requires exchanging data on a per frame/ packet basis, and a TE link (e.g., a link between OXCs) may not be capable of exchanging data on a per packet basis. In this case, the

routing and signaling adjacencies are maintained via a set of one or more control channels (see [LMP]).

Two nodes may have a TE link between them even if they do not have a routing adjacency. Naturally, each node must run OSPF/IS-IS with GMPLS extensions in order for that TE link to be advertised. More precisely, the node needs to run GMPLS extensions for TE Links with an interface switching capability (see [GMPLS-ROUTING]) other than PSC. Moreover, this node needs to run either GMPLS or MPLS extensions for TE links with an interface switching capability of PSC.

The mechanisms for Control Channel Separation [RFC3471] should be used (even if the IP path between two nodes is a TE link). I.e., RSVP-TE/CR-LDP signaling should use the Interface_ID (IF_ID) object to specify a particular TE link when establishing an LSP.

The IP path could consist of multiple IP hops. In this case, the mechanisms of sections 7.1.1 and 7.1.2 of [HIERARCHY] should be used (in addition to Control Channel Separation).

10. Control Plane Fault Handling

Two major types of faults can impact a control plane. The first, referred to as control channel fault, relates to the case where control communication is lost between two neighboring nodes. If the control channel is embedded with the data channel, data channel recovery procedure should solve the problem. If the control channel is independent of the data channel, additional procedures are required to recover from that problem.

The second, referred to as nodal faults, relates to the case where node loses its control state (e.g., after a restart) but does not loose its data forwarding state.

In transport networks, such types of control plane faults should not have service impact on the existing connections. Under such circumstances, a mechanism must exist to detect a control communication failure and a recovery procedure must guarantee connection integrity at both ends of the control channel.

For a control channel fault, once communication is restored routing protocols are naturally able to recover but the underlying signaling protocols must indicate that the nodes have maintained their state through the failure. The signaling protocol must also ensure that any state changes that were instantiated during the failure are synchronized between the nodes.

For a nodal fault, a node's control plane restarts and loses most of its state information. In this case, both upstream and downstream nodes must synchronize their state information with the restarted node. In order for any resynchronization to occur the node undergoing the restart will need to preserve some information, such as it's mappings of incoming to outgoing labels.

These issues are addressed in protocol specific fashions, see [RFC3473], [RFC3472], [OSPF-TE-GMPLS] and [ISIS-TE-GMPLS]. Note that these cases only apply when there are mechanisms to detect data channel failures independent of control channel failures.

The LDP Fault tolerance (see [RFC3479]) specifies the procedures to recover from a control channel failure. [RFC3473] specifies how to recover from both a control channel failure and a node failure.

11. LSP Protection and Restoration

This section discusses Protection and Restoration (P&R) issues for GMPLS LSPs. It is driven by the requirements outlined in [RFC3386] and some of the principles outlined in [RFC3469]. It will be enhanced, as more GMPLS P&R mechanisms are defined. The scope of this section is clarified hereafter:

- This section is only applicable when a fault impacting LSP(s) happens in the data/transport plane. Section 10 deals with control plane fault handling for nodal and control channel faults.

- This section focuses on P&R at the TDM, LSC and FSC layers. There are specific P&R requirements at these layers not present at the PSC layer.

- This section focuses on intra-area P&R as opposed to inter-area P&R and even inter-domain P&R. Note that P&R can even be more restricted, e.g., to a collection of like customer equipment, or a collection of equipment of like capabilities, in one single routing area.

- This section focuses on intra-layer P&R (horizontal hierarchy as defined in [RFC3386]) as opposed to the inter-layer P&R (vertical hierarchy).

- P&R mechanisms are in general designed to handle single failures, which makes SRLG diversity a necessity. Recovery from multiple failures requires further study.

- Both mesh and ring-like topologies are supported.

In the following, we assume that:

- TDM, LSC and FSC devices are more generally committing recovery resources in a non-best effort way. Recovery resources are either allocated (thus used) or at least logically reserved (whether used or not by preemptable extra traffic but unavailable anyway for regular working traffic).

- Shared P&R mechanisms are valuable to operators in order to maximize their network utilization.

- Sending preemptable excess traffic on recovery resources is a valuable feature for operators.

11.1. Protection Escalation across Domains and Layers

To describe the P&R architecture, one must consider two dimensions of hierarchy [RFC3386]:

- A horizontal hierarchy consisting of multiple P&R domains, which is important in an LSP based protection scheme. The scope of P&R may extend over a link (or span), an administrative domain or sub-network, an entire LSP.

  An administrative domain may consist of a single P&R domain or as a concatenation of several smaller P&R domains. The operator can configure P&R domains, based on customers' requirements, and on network topology and traffic engineering constraints.

- A vertical hierarchy consisting of multiple layers of P&R with varying granularities (packet flows, STS trails, lightpaths, fibers, etc).

  In the absence of adequate P&R coordination, a fault may propagate from one level to the next within a P&R hierarchy. It can lead to "collisions" and simultaneous recovery actions may lead to race conditions, reduced resource utilization, or instabilities [MANCHESTER]. Thus, a consistent escalation strategy is needed to coordinate recovery across domains and layers. The fact that GMPLS can be used at different layers could simplify this coordination.

  There are two types of escalation strategies: bottom-up and top-down. The bottom-up approach assumes that "lower-level" recovery schemes are more expedient. Therefore we can inhibit or hold off

higher-level P&R. The Top-down approach attempts service P&R at the higher levels before invoking "lower level" P&R. Higher-layer P&R is service selective, and permits "per-CoS" or "per-LSP" re-routing.

Service Level Agreements (SLAs) between network operators and their clients are needed to determine the necessary time scales for P&R at each layer and at each domain.

11.2. Mapping of Services to P&R Resources

The choice of a P&R scheme is a tradeoff between network utilization (cost) and service interruption time. In light of this tradeoff, network service providers are expected to support a range of different service offerings or service levels.

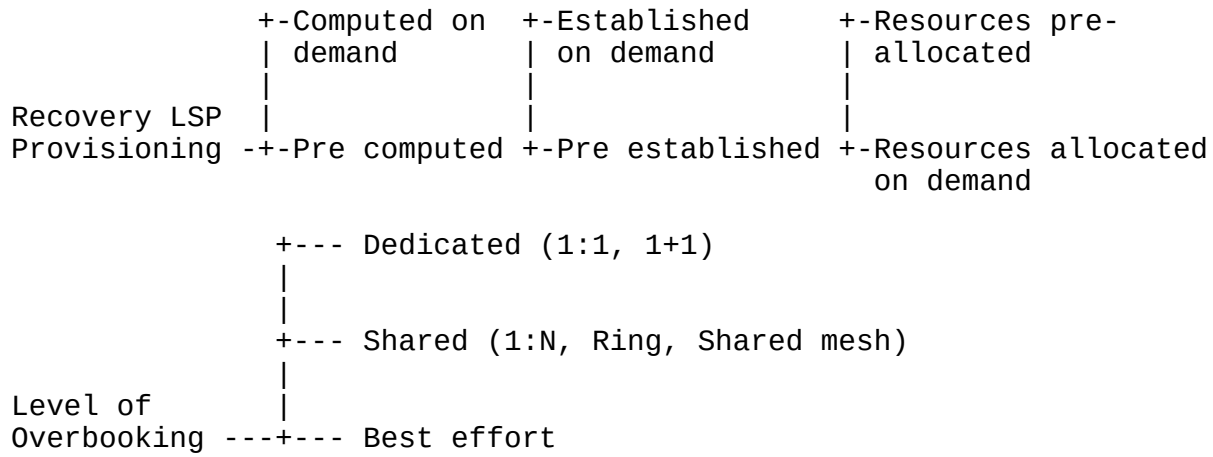
One can classify LSPs into one of a small set of service levels. Among other things, these service levels define the reliability characteristics of the LSP. The service level associated with a given LSP is mapped to one or more P&R schemes during LSP establishment. An advantage that mapping is that an LSP may use different P&E schemes in different segments of a network (e.g., some links may be span protected, whilst other segments of the LSP may utilize ring protection). These details are likely to be service provider specific.

An alternative to using service levels is for an application to specify the set of specific P&R mechanisms to be used when establishing the LSP. This allows greater flexibility in using different mechanisms to meet the application requirements.

A differentiator between these service levels is service interruption time in case of network failures, which is defined as the length of time between when a failure occurs and when connectivity is re-established. The choice of service level (or P&R scheme) should be dictated by the service requirements of different applications.

11.3. Classification of P&R Mechanism Characteristics

The following figure provides a classification of the possible provisioning types of recovery LSPs, and of the levels of overbooking that is possible for them.



11.4. Different Stages in P&R

Recovery from a network fault or impairment takes place in several stages as discussed in [RFC3469], including fault detection, fault localization, notification, recovery (i.e., the P&R itself) and reversion of traffic (i.e., returning the traffic to the original working LSP or to a new one).

- Fault detection is 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 actions, or the alarm may be propagated at the lower layer (e.g., SONET/SDH AIS).

- Fault localization can be done with the help of GMPLS, e.g., using LMP for fault localization (see section 6.4).

- Fault notification can also be achieved through GMPLS, e.g., using GMPLS RSVP-TE/CR-LDP notification (see section 7.12).

- This section focuses on the different mechanisms available for recovery and reversion of traffic once fault detection, localization and notification have taken place.

11.5. Recovery Strategies

Network P&R techniques can be divided into Protection and Restoration. In protection, resources between the protection endpoints are established before failure, and connectivity after failure is achieved simply by switching performed at the protection end-points. In contrast, restoration uses signaling after failure to allocate resources along the recovery path.

- Protection aims at extremely fast reaction times and may rely on the use of overhead control fields for achieving end-point coordination. Protection for SONET/SDH networks is described in [ITUT-G.841] and [ANSI-T1.105]. Protection mechanisms can be further classified by the level of redundancy and sharing.

- Restoration mechanisms rely on signaling protocols to coordinate switching actions during recovery, and may involve simple re-provisioning, i.e., signaling only at the time of recovery; or pre-signaling, i.e., signaling prior to recovery.

In addition, P&R can be applied on a local or end-to-end basis. In the local approach, P&R is focused on the local proximity of the fault in order to reduce delay in restoring service. In the end-to-end approach, the LSP originating and terminating nodes control recovery.

Using these strategies, the following recovery mechanisms can be defined.

11.6. Recovery mechanisms: Protection schemes

Note that protection schemes are usually defined in technology specific ways, but this does not preclude other solutions.

- 1+1 Link Protection: Two pre-provisioned resources are used in parallel. For example, data is transmitted simultaneously on two parallel links and a selector is used at the receiving node to choose the best source (see also [GMPLS-FUNCT]).

- 1:N Link Protection: Working and protecting resources (N working, 1 backup) are pre-provisioned. If a working resource fails, the data is switched to the protecting resource, using a coordination mechanism (e.g., in overhead bytes). More generally, N working and M protecting resources can be assigned for M:N link protection (see also [GMPLS-FUNCT]).

- Enhanced Protection: Various mechanisms such as protection rings can be used to enhance the level of protection beyond single link failures to include the ability to switch around a node failure or multiple link failures within a span, based on a pre-established topology of protection resources (note: no reference available at publication time).

- 1+1 LSP Protection: Simultaneous data transmission on working and protecting LSPs and tail-end selection can be applied (see also [GMPLS-FUNCT]).

11.7. Recovery mechanisms: Restoration schemes

Thanks to the use of a distributed control plane like GMPLS, restoration is possible in multiple of tenths of milliseconds. It is much harder to achieve when only an NMS is used and can only be done in that case in a multiple of seconds.

- End-to-end LSP restoration with re-provisioning: an end-to-end restoration path is established after failure. The restoration path may be dynamically calculated after failure, or pre-calculated before failure (often during LSP establishment). Importantly, no signaling is used along the restoration path before failure, and no restoration bandwidth is reserved. Consequently, there is no guarantee that a given restoration path is available when a failure occurs. Thus, one may have to crankback to search for an available path.

- End-to-end LSP restoration with pre-signaled recovery bandwidth reservation and no label pre-selection: an end-to-end restoration path is pre-calculated before failure and a signaling message is sent along this pre-selected path to reserve bandwidth, but labels are not selected (see also [GMPLS-FUNCT]).

  The resources reserved on each link of a restoration path may be shared across different working LSPs that are not expected to fail simultaneously. Local node policies can be applied to define the degree to which capacity is shared across independent failures. Upon failure detection, LSP signaling is initiated along the restoration path to select labels, and to initiate the appropriate cross-connections.

- End-to-end LSP restoration with pre-signaled recovery bandwidth reservation and label pre-selection: An end-to-end restoration path is pre-calculated before failure and a signaling procedure is initiated along this pre-selected path on which bandwidth is reserved and labels are selected (see also [GMPLS-FUNCT]).

  The resources reserved on each link may be shared across different working LSPs that are not expected to fail simultaneously. In networks based on TDM, LSC and FSC technology, LSP signaling is used after failure detection to establish cross-connections at the intermediate switches on the restoration path using the pre-selected labels.

- Local LSP restoration: the above approaches can be applied on a local basis rather than end-to-end, in order to reduce recovery time (note: no reference available at publication time).

11.8. Schema Selection Criteria

This section discusses criteria that could be used by the operator in order to make a choice among the various P&R mechanisms.

- Robustness: In general, the less pre-planning of the restoration path, the more robust the restoration scheme is to a variety of failures, provided that adequate resources are available. Restoration schemes with pre-planned paths will not be able to recover from network failures that simultaneously affect both the working and restoration paths. Thus, these paths should ideally be chosen to be as disjoint as possible (i.e., SRLG and node disjoint), so that any single failure event will not affect both paths. The risk of simultaneous failure of the two paths can be reduced by recalculating the restoration path whenever a failure occurs along it.

  The pre-selection of a label gives less flexibility for multiple failure scenarios than no label pre-selection. If failures occur that affect two LSPs that are sharing a label at a common node along their restoration routes, then only one of these LSPs can be recovered, unless the label assignment is changed.

  The robustness of a restoration scheme is also determined by the amount of reserved restoration bandwidth - as the amount of restoration bandwidth sharing increases (reserved bandwidth decreases), the restoration scheme becomes less robust to failures. Restoration schemes with pre-signaled bandwidth reservation (with or without label pre-selection) can reserve adequate bandwidth to ensure recovery from any specific set of failure events, such as any single SRLG failure, any two SRLG failures etc. Clearly, more restoration capacity is allocated if a greater degree of failure recovery is required. Thus, the degree to which the network is protected is determined by the policy that defines the amount of reserved restoration bandwidth.

- Recovery time: In general, the more pre-planning of the restoration route, the more rapid the P&R scheme. Protection schemes generally recover faster than restoration schemes. Restoration with pre-signaled bandwidth reservation are likely to be (significantly) faster than path restoration with re-provisioning, especially because of the elimination of any crankback. Local restoration will generally be faster than end-to-end schemes.

Recovery time objectives for SONET/SDH protection switching (not including time to detect failure) are specified in [ITUT-G.841] at 50 ms, taking into account constraints on distance, number of connections involved, and in the case of ring enhanced protection, number of nodes in the ring.

Recovery time objectives for restoration mechanisms are being defined through a separate effort [RFC3386].

- Resource Sharing: 1+1 and 1:N link and LSP protection require dedicated recovery paths with limited ability to share resources: 1+1 allows no sharing, 1:N allows some sharing of protection resources and support of extra (pre-emptable) traffic. Flexibility is limited because of topology restrictions, e.g., fixed ring topology for traditional enhanced protection schemes.

  The degree to which restoration schemes allow sharing amongst multiple independent failures is directly dictated by the size of the restoration pool. In restoration schemes with re-provisioning, a pool of restoration capacity can be defined from which all restoration routes are selected after failure. Thus, the degree of sharing is defined by the amount of available restoration capacity. In restoration with pre-signaled bandwidth reservation, the amount of reserved restoration capacity is determined by the local bandwidth reservation policies. In all restoration schemes, pre-emptable resources can use spare restoration capacity when that capacity is not being used for failure recovery.

12. Network Management

Service Providers (SPs) use network management extensively to configure, monitor or provision various devices in their network. It is important to note that a SP’s equipment may be distributed across geographically separate sites thus making distributed management even more important. The service provider should utilize an NMS system and standard management protocols such as SNMP (see [RFC3410], [RFC3411] and [RFC3416]) and the relevant MIB modules as standard interfaces to configure, monitor and provision devices at various locations. The service provider may also wish to use the command line interface (CLI) provided by vendors with their devices. However, this is not a standard or recommended solution because there is no standard CLI language or interface, which results in N different CLIs in a network with devices from N different vendors. In the context of GMPLS, it is extremely important for standard interfaces to the SP’s devices (e.g., SNMP) to exist due to the nature of the technology itself. Since GMPLS comprises many different layers of control-plane

and data-plane technology, it is important for management interfaces in this area to be flexible enough to allow the manager to manage GMPLS easily, and in a standard way.

12.1. Network Management Systems (NMS)

The NMS system should maintain the collective information about each device within the system. Note that the NMS system may actually be comprised of several distributed applications (i.e., alarm aggregators, configuration consoles, polling applications, etc.) that collectively comprises the SP's NMS. In this way, it can make provisioning and maintenance decisions with the full knowledge of the entire SP's network. Configuration or provisioning information (i.e., requests for new services) could be entered into the NMS and subsequently distributed via SNMP to the remote devices. Thus, making the SP's task of managing the network much more compact and effortless rather than having to manage each device individually (i.e., via CLI).

Security and access control can be achieved using the SNMPv3 User-based Security Model (USM) [RFC3414] and the View-based Access Control Model (VACM) [RFC3415]. This approach can be very effectively used within a SP's network, since the SP has access to and control over all devices within its domain. Standardized MIBs will need to be developed before this approach can be used ubiquitously to provision, configure and monitor devices in non-heterogeneous networks or across SP's network boundaries.

12.2. Management Information Base (MIB)

In the context of GMPLS, it is extremely important for standard interfaces to devices to exist due to the nature of the technology itself. Since GMPLS comprises many different layers of control-plane technology, it is important for SNMP MIB modules in this area to be flexible enough to allow the manager to manage the entire control plane. This should be done using MIB modules that may cooperate (i.e., coordinated row-creation on the agent) or through more generalized MIB modules that aggregate some of the desired actions to be taken and push those details down to the devices. It is important to note that in certain circumstances, it may be necessary to duplicate some small subset of manageable objects in new MIB modules for management convenience. Control of some parts of GMPLS may also be achieved using existing MIB interfaces (i.e., existing SONET MIB) or using separate ones, which are yet to be defined. MIB modules may have been previously defined in the IETF or ITU. Current MIB modules may need to be extended to facilitate some of the new functionality

desired by GMPLS. In these cases, the working group should work on new versions of these MIB modules so that these extensions can be added.

12.3. Tools

As in traditional networks, standard tools such as traceroute [RFC1393] and ping [RFC2151] are needed for debugging and performance monitoring of GMPLS networks, and mainly for the control plane topology, that will mimic the data plane topology. Furthermore, such tools provide network reachability information. The GMPLS control protocols will need to expose certain pieces of information in order for these tools to function properly and to provide information germane to GMPLS. These tools should be made available via the CLI. These tools should also be made available for remote invocation via the SNMP interface [RFC2925].

12.4. Fault Correlation between Multiple Layers

Due to the nature of GMPLS, and that potential layers may be involved in the control and transmission of GMPLS data and control information, it is required that a fault in one layer be passed to the adjacent higher and lower layers to notify them of the fault. However, due to nature of these many layers, it is possible and even probable, that hundreds or even thousands of notifications may need to transpire between layers. This is undesirable for several reasons. First, these notifications will overwhelm the device. Second, if the device(s) are programmed to emit SNMP Notifications [RFC3417] then the large number of notifications the device may attempt to emit may overwhelm the network with a storm of notifications. Furthermore, even if the device emits the notifications, the NMS that must process these notifications either will be overwhelmed or will be processing redundant information. That is, if 1000 interfaces at layer B are stacked above a single interface below it at layer A, and the interface at A goes down, the interfaces at layer B should not emit notifications. Instead, the interface at layer A should emit a single notification. The NMS receiving this notification should be able to correlate the fact that this interface has many others stacked above it and take appropriate action, if necessary.

Devices that support GMPLS should provide mechanisms for aggregating, summarizing, enabling and disabling of inter-layer notifications for the reasons described above. In the context of SNMP MIB modules, all MIB modules that are used by GMPLS must provide enable/disable objects for all notification objects. Furthermore, these MIBs must also provide notification summarization objects or functionality (as described above) as well. NMS systems and standard tools which

process notifications or keep track of the many layers on any given devices must be capable of processing the vast amount of information which may potentially be emitted by network devices running GMPLS at any point in time.

13. Security Considerations

GMPLS defines a control plane architecture for multiple technologies and types of network elements. In general, since LSPs established using GMPLS may carry high volumes of data and consume significant network resources, security mechanisms are required to safeguard the underlying network against attacks on the control plane and/or unauthorized usage of data transport resources. The GMPLS control plane should therefore include mechanisms that prevent or minimize the risk of attackers being able to inject and/or snoop on control traffic. These risks depend on the level of trust between nodes that exchange GMPLS control messages, as well as the realization and physical characteristics of the control channel. For example, an in-band, in-fiber control channel over SONET/SDH overhead bytes is, in general, considered less vulnerable than a control channel realized over an out-of-band IP network.

Security mechanisms can provide authentication and confidentiality. Authentication can provide origin verification, message integrity and replay protection, while confidentiality ensures that a third party cannot decipher the contents of a message. In situations where GMPLS deployment requires primarily authentication, the respective authentication mechanisms of the GMPLS component protocols may be used (see [RFC2747], [RFC3036], [RFC2385] and [LMP]). Additionally, the IPsec suite of protocols (see [RFC2402], [RFC2406] and [RFC2409]) may be used to provide authentication, confidentiality or both, for a GMPLS control channel. IPsec thus offers the benefits of combined protection for all GMPLS component protocols as well as key management.

A related issue is that of the authorization of requests for resources by GMPLS-capable nodes. Authorization determines whether a given party, presumable already authenticated, has a right to access the requested resources. This determination is typically a matter of local policy control [RFC2753], for example by setting limits on the total bandwidth available to some party in the presence of resource contention. Such policies may become quite complex as the number of users, types of resources and sophistication of authorization rules increases.

After authenticating requests, control elements should match them against the local authorization policy. These control elements must be capable of making decisions based on the identity of the

requester, as verified cryptographically and/or topologically. For example, decisions may depend on whether the interface through which the request is made is an inter- or intra-domain one. The use of appropriate local authorization policies may help in limiting the impact of security breaches in remote parts of a network.

Finally, it should be noted that GMPLS itself introduces no new security considerations to the current MPLS-TE signaling (RSVP-TE, CR-LDP), routing protocols (OSPF-TE, IS-IS-TE) or network management protocols (SNMP).

14. Acknowledgements

This document is the work of numerous authors and consists of a composition of a number of previous documents in this area.

Many thanks to Ben Mack-Crane (Tellabs) for all the useful SONET/SDH discussions we had together. Thanks also to Pedro Falcao, Alexandre Geyssens, Michael Moelants, Xavier Neerdaels, and Philippe Noel from Ebone for their SONET/SDH and optical technical advice and support. Finally, many thanks also to Krishna Mitra (Consultant), Curtis Villamizar (Avici), Ron Bonica (WorldCom), and Bert Wijnen (Lucent) for their revision effort on Section 12.

15. References

15.1. Normative References

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

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

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

[RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[RFC3472] Ashwood-Smith, P. and L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Constraint-based Routed Label Distribution Protocol (CR-LDP) Extensions", RFC 3472, January 2003.

[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

15.2. Informative References

[ANSI-T1.105] "Synchronous Optical Network (SONET): Basic Description Including Multiplex Structure, Rates, And Formats," ANSI T1.105, 2000.

[BUNDLE] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering", Work in Progress.

[GMPLS-FUNCT] Lang, J.P., Ed. and B. Rajagopalan, Ed., "Generalized MPLS Recovery Functional Specification", Work in Progress.

[GMPLS-G709] Papadimitriou, D., Ed., "GMPLS Signaling Extensions for G.709 Optical Transport Networks Control", Work in Progress.

[GMPLS-OVERLAY] Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter, "GMPLS UNI: RSVP Support for the Overlay Model", Work in Progress.

[GMPLS-ROUTING] Kompella, K., Ed. and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching", Work in Progress.

[RFC3946] Mannie, E., Ed. and Papadimitriou D., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) Control", RFC 3946, October 2004.

[HIERARCHY] Kompella, K. and Y. Rekhter, "LSP Hierarchy with Generalized MPLS TE", Work in Progress.

[ISIS-TE] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.

[ISIS-TE-GMPLS] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching", Work in Progress.

[ITUT-G.707] ITU-T, "Network Node Interface for the Synchronous Digital Hierarchy", Recommendation G.707, October 2000.

[ITUT-G.709] ITU-T, "Interface for the Optical Transport Network (OTN)," Recommendation G.709 version 1.0 (and Amendment 1), February 2001 (and October 2001).

[ITUT-G.841] ITU-T, "Types and Characteristics of SDH Network Protection Architectures," Recommendation G.841, October 1998.

[LMP] Lang, J., Ed., "Link Management Protocol (LMP)", Work in Progress.

[LMP-WDM] Fredette, A., Ed. and J. Lang Ed., "Link Management Protocol (LMP) for Dense Wavelength Division Multiplexing (DWDM) Optical Line Systems", Work in Progress.

[MANCHESTER] J. Manchester, P. Bonenfant and C. Newton, "The Evolution of Transport Network Survivability," IEEE Communications Magazine, August 1999.

[OIF-UNI] The Optical Internetworking Forum, "User Network Interface (UNI) 1.0 Signaling Specification - Implementation Agreement OIF-UNI-01.0," October 2001.

[OLI-REQ] Fredette, A., Ed., "Optical Link Interface Requirements," Work in Progress.

[OSPF-TE-GMPLS] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching", Work in Progress.

[OSPF-TE] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC1393] Malkin, G., "Traceroute Using an IP Option", RFC 1393, January 1993.

[RFC2151] Kessler, G. and S. Shepard, "A Primer On Internet and TCP/IP Tools and Utilities", RFC 2151, June 1997.

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

[RFC2385] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385, August 1998.

[RFC2402] Kent, S. and R. Atkinson, "IP Authentication Header", RFC 2402, November 1998.

[RFC2406] Kent, S. and R. Atkinson, "IP Encapsulating Security Payload (ESP)", RFC 2406, November 1998.

[RFC2409] Harkins, D. and D. Carrel, "The Internet Key Exchange (IKE)", RFC 2409, November 1998.

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

[RFC2747] Baker, F., Lindell, B., and M. Talwar, "RSVP Cryptographic Authentication", RFC 2747, January 2000.

[RFC2753] Yavatkar, R., Pendarakis, D., and R. Guerin, "A Framework for Policy-based Admission Control", RFC 2753, January 2000.

[RFC2925] White, K., "Definitions of Managed Objects for Remote Ping, Traceroute, and Lookup Operations", RFC 2925, September 2000.

[RFC3036] Andersson, L., Doolan, P., Feldman, N., Fredette, A., and B. Thomas, "LDP Specification", RFC 3036, January 2001.

[RFC3386] Lai, W. and D. McDysan, "Network Hierarchy and Multilayer Survivability", RFC 3386, November 2002.

[RFC3410] Case, J., Mundy, R., Partain, D., and B. Stewart, "Introduction and Applicability Statements for Internet-Standard Management Framework", RFC 3410, December 2002.

[RFC3411] Harrington, D., Presuhn, R., and B. Wijnen, "An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks", STD 62, RFC 3411, December 2002.

[RFC3414] Blumenthal, U. and B. Wijnen, "User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)", STD 62, RFC 3414, December 2002.

[RFC3415] Wijnen, B., Presuhn, R., and K. McCloghrie, "View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)", STD 62, RFC 3415, December 2002.

[RFC3416] Presuhn, R., "Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)", STD 62, RFC 3416, December 2002.

[RFC3417] Presuhn, R., "Transport Mappings for the Simple Network Management Protocol (SNMP)", STD 62, RFC 3417, December 2002.

[RFC3469] Sharma, V. and F. Hellstrand, "Framework for Multi-Protocol Label Switching (MPLS)-based Recovery", RFC 3469, February 2003.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.

[RFC3479] Farrel, A., "Fault Tolerance for the Label Distribution Protocol (LDP)", RFC 3479, February 2003.

[RFC3480] Kompella, K., Rekhter, Y., and A. Kullberg, "Signalling Unnumbered Links in CR-LDP (Constraint-Routing Label Distribution Protocol)", RFC 3480, February 2003.

[SONET-SDH-GMPLS-FRM] Bernstein, G., Mannie, E., and V. Sharma, "Framework for GMPLS-based Control of SDH/SONET Networks", Work in Progress.

16. Contributors

Peter Ashwood-Smith
Nortel
P.O. Box 3511 Station C,
Ottawa, ON K1Y 4H7, Canada

EMail: petera@nortelnetworks.com

Eric Mannie
Consult
Phone: +32 2 648-5023
Mobile: +32 (0)495-221775

EMail: eric_mannie@hotmail.com

Daniel O. Awduche
Consult

EMail: awduche@awduche.com

Thomas D. Nadeau
Cisco
250 Apollo Drive
Chelmsford, MA 01824, USA

EMail: tnadeau@cisco.com

Ayan Banerjee
Calient
5853 Rue Ferrari
San Jose, CA 95138, USA

EMail: abanerjee@calient.net

Lyndon Ong
Ciena
10480 Ridgeview Ct
Cupertino, CA 95014, USA

EMail: lyong@ciena.com

Debashis Basak
Accelight
70 Abele Road, Bldg.1200
Bridgeville, PA 15017, USA

EMail: dbasak@accelight.com

Dimitri Papadimitriou
Alcatel
Francis Wellesplein, 1
B-2018 Antwerpen, Belgium

EMail: dimitri.papadimitriou@alcatel.be

Lou Berger
Movaz
7926 Jones Branch Drive
MCLean VA, 22102, USA

EMail: lberger@movaz.com

Dimitrios Pendarakis
Tellium
2 Crescent Place, P.O. Box 901
Oceanport, NJ 07757-0901, USA

EMail: dpendarakis@tellium.com

Greg Bernstein
Grotto

EMail: gregb@grotto-networking.com

Bala Rajagopalan
Tellium
2 Crescent Place, P.O. Box 901
Oceanport, NJ 07757-0901, USA

EMail: braja@tellium.com

Sudheer Dharanikota
Consult

EMail: sudheer@ieee.org

Yakov Rekhter
Juniper
1194 N. Mathilda Ave.
Sunnyvale, CA 94089, USA

EMail: yakov@juniper.net

John Drake
Calient
5853 Rue Ferrari
San Jose, CA 95138, USA

EMail: jdrake@calient.net

Debanjan Saha
Tellium
2 Crescent Place
Oceanport, NJ 07757-0901, USA

EMail: dsaha@tellium.com

Yanhe Fan
Axiowave
200 Nickerson Road
Marlborough, MA 01752, USA

EMail: yfan@axiowave.com

Hal Sandick
Shepard M.S.
2401 Dakota Street
Durham, NC 27705, USA

EMail: sandick@nc.rr.com

Don Fedyk
Nortel
600 Technology Park Drive
Billerica, MA 01821, USA

EMail: dwfedyk@nortelnetworks.com

Vishal Sharma
Metanoia
1600 Villa Street, Unit 352
Mountain View, CA 94041, USA

EMail: v.sharma@ieee.org

Gert Grammel
Alcatel
Lorenzstrasse, 10
70435 Stuttgart, Germany

EMail: gert.grammel@alcatel.de

George Swallow
Cisco
250 Apollo Drive
Chelmsford, MA 01824, USA

EMail: swallow@cisco.com

Dan Guo
Turin
1415 N. McDowell Blvd,
Petaluma, CA 95454, USA

EMail: dguo@turinnetworks.com

Z. Bo Tang
Tellium
2 Crescent Place, P.O. Box 901
Oceanport, NJ 07757-0901, USA

EMail: btang@tellium.com

Kireeti Kompella
Juniper
1194 N. Mathilda Ave.
Sunnyvale, CA 94089, USA

EMail: kireeti@juniper.net

Jennifer Yates
AT&T
180 Park Avenue
Florham Park, NJ 07932, USA

EMail: jyates@research.att.com

Alan Kullberg
NetPlane
888 Washington
St.Dedham, MA 02026, USA

EMail: akullber@netplane.com

George R. Young
Edgeflow
329 March Road
Ottawa, Ontario, K2K 2E1, Canada

EMail: george.young@edgeflow.com

Jonathan P. Lang
Rincon Networks

EMail: jplang@ieee.org

John Yu
Hammerhead Systems
640 Clyde Court
Mountain View, CA 94043, USA

EMail: john@hammerheadsystems.com

Fong Liaw
Solas Research
Solas Research, LLC

EMail: fongliaw@yahoo.com

Alex Zinin
Alcatel
1420 North McDowell Ave
Petaluma, CA 94954, USA

EMail: alex.zinin@alcatel.com

17. Author's Address

Eric Mannie (Consultant)
Avenue de la Folle Chanson, 2
B-1050 Brussels, Belgium
Phone: +32 2 648-5023
Mobile: +32 (0)495-221775

EMail: eric_mannie@hotmail.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.