Network Working Group
Request for Comments: 3717
Category: Informational

B. Rajagopalan
Consultant
J. Luciani
Marconi Communications
D. Awduche
MCI

March 2004

IP over Optical Networks: A Framework

Status of this Memo

Copyright Notice

Abstract

The Internet transport infrastructure is moving towards a model of high-speed routers interconnected by optical core networks. The architectural choices for the interaction between IP and optical network layers, specifically, the routing and signaling aspects, are maturing. At the same time, a consensus has emerged in the industry on utilizing IP-based protocols for the optical control plane. This document defines a framework for IP over Optical networks, considering both the IP-based control plane for optical networks as well as IP-optical network interactions (together referred to as "IP over optical networks").

Table of Contents

1. Introduction

Optical network technologies are evolving rapidly in terms of functions and capabilities. The increasing importance of optical networks is evidenced by the copious amount of attention focused on IP over optical networks and related photonic and electronic interworking issues by all major network service providers, telecommunications equipment vendors, and standards organizations. In this regard, the term "optical network" is used generically in practice to refer to both SONET/SDH-based transport networks, as well as switched optical networks (including all-optical networks).

It has been realized that optical networks must be survivable, flexible, and controllable. There is, therefore, an ongoing trend to introduce intelligence in the control plane of optical networks to make them more versatile [1]. An essential attribute of intelligent optical networks is the capability to instantiate and route optical layer connections in real-time or near real-time, and to provide capabilities that enhance network survivability. Furthermore, there is a need for multi-vendor optical network interoperability, when an optical network may consist of interconnected vendor-specific optical sub-networks.

The optical network must also be versatile because some service providers may offer generic optical layer services that may not be client-specific. It would therefore be necessary to have an optical network control plane that can handle such generic optical services.

There is general consensus in the industry that the optical network control plane should utilize IP-based protocols for dynamic provisioning and restoration of optical channels within and across optical sub-networks. This is based on the practical view that signaling and routing mechanisms developed for IP traffic engineering applications could be re-used in optical networks. Nevertheless, the issues and requirements that are specific to optical networking must be understood to suitably adopt and adapt the IP-based protocols. This is especially the case for restoration, and for routing and signaling in all-optical networks. Also, there are different views on the model for interaction between the optical network and client networks, such as IP networks. Reasonable architectural alternatives in this regard must be supported, with an understanding of their relative merits.

Thus, there are two fundamental issues related to IP over optical networks. The first is the adaptation and reuse of IP control plane protocols within the optical network control plane, irrespective of the types of digital clients that utilize the optical network. The

second is the transport of IP traffic through an optical network together with the control and coordination issues that arise therefrom.

This document defines a framework for IP over optical networks covering the requirements and mechanisms for establishing an IP-centric optical control plane, and the architectural aspects of IP transport over optical networks. In this regard, it is recognized that the specific capabilities required for IP over optical networks would depend on the services expected at the IP-optical interface as well as the optical sub-network interfaces. Depending on the specific operational requirements, a progression of capabilities is possible, reflecting increasingly sophisticated interactions at these interfaces. This document therefore advocates the definition of "capability sets" that define the evolution of functionality at the interfaces as more sophisticated operational requirements arise.

This document is organized as follows. In the next section, terminology covering some basic concepts related to this framework are described. The definitions are specific to this framework and may have other connotations elsewhere. In Section 3, the network model pertinent to this framework is described. The service model and requirements for IP-optical, and multi-vendor optical internetworking are described in Section 4. This section also considers some general requirements. Section 5 considers the architectural models for IP-optical interworking, describing the relative merits of each model. It should be noted that it is not the intent of this document to promote any particular model over the others. However, particular aspects of the models that may make one approach more appropriate than another in certain circumstances are described. Section 6 describes IP-centric control plane mechanisms for optical networks, covering signaling and routing issues in support of provisioning and restoration. The approaches described in Section 5 and 6 range from the relatively simple to the sophisticated. Section 7 describes a number of specialized issues in relation to IP over optical networks. Section 8 describes a possible evolution path for IP over optical networking capabilities in terms of increasingly sophisticated functionality that may be supported as the need arises. Section 9 considers security issues pertinent to this framework. Finally, the summary and conclusion are presented in Section 10.

2. Terminology and Concepts

This section introduces terminology pertinent to this framework and some related concepts. The definitions are specific to this framework and may have other interpretations elsewhere.

WDM

Wavelength Division Multiplexing (WDM) is a technology that allows multiple optical signals operating at different wavelengths to be multiplexed onto a single optical fiber and transported in parallel through the fiber. In general, each optical wavelength may carry digital client payloads at a different data rate (e.g., OC-3c, OC-12c, OC- 48c, OC-192c, etc.) and in a different format (SONET, Ethernet, ATM, etc.). For example, there are many commercial WDM networks in existence today that support a mix of SONET signals operating at OC-48c (approximately 2.5 Gbps) and OC-192 (approximately 10 Gbps) over a single optical fiber. An optical system with WDM capability can achieve parallel transmission of multiple wavelengths gracefully while maintaining high system performance and reliability. In the near future, commercial dense WDM systems are expected to concurrently carry more than 160 wavelengths at data rates of OC-192c and above, for a total of 1.6 Tbps or more. The term WDM will be used in this document to refer to both WDM and DWDM (Dense WDM).

In general, it is worth noting that WDM links are affected by the following factors, which may introduce impairments into the optical signal path:

1. The number of wavelengths on a single fiber.
2. The serial bit rate per wavelength.
3. The type of fiber.
4. The amplification mechanism.
5. The number and type of nodes through which the signals pass before reaching the egress node or before regeneration.

All these factors (and others not mentioned here) constitute domain specific features of optical transport networks. As noted in [1], these features should be taken into account in developing standards based solutions for IP over optical networks.

Optical cross-connect (OXC)

An OXC is a space-division switch that can switch an optical data stream from an input port to a output port. Such a switch may utilize optical-electrical conversion at the input port and electrical-optical conversion at the output port, or it may be all-optical. An OXC is assumed to have a control-plane processor that implements the signaling and routing protocols necessary for computing and instantiating optical channel connectivity in the optical domain.

Optical channel trail or Lightpath

An optical channel trail is a point-to-point optical layer connection between two access points in an optical network. In this document, the term "lightpath" is used interchangeably with optical channel trail.

Optical mesh sub-network

An optical sub-network, as used in this framework, is a network of OXCs that supports end-to-end networking of optical channel trails providing functionality like routing, monitoring, grooming, and protection and restoration of optical channels. The interconnection of OXCs in this network can be based on a general mesh topology. The following sub-layers may be associated with this network:

(a) An optical multiplex section (OMS) layer network: The optical multiplex section layer provides transport for the optical channels. The information contained in this layer is a data stream comprising a set of optical channels, which may have a defined aggregate bandwidth.

(b) An optical transmission section (OTS) layer network: This layer provides functionality for transmission of optical signals through different types of optical media.

This framework does not address the interaction between the optical sub-network and the OMS, or between the OMS and OTS layer networks.

Mesh optical network (or simply, "optical network")

A mesh optical network, as used in document, is a topologically connected collection of optical sub-networks whose node degree may exceed 2. Such an optical network is assumed to be under the purview of a single administrative entity. It is also possible to conceive of a large scale global mesh optical network consisting of the voluntary interconnection of autonomous optical networks, each of which is owned and administered by an independent entity. In such an environment, abstraction can be used to hide the internal details of each autonomous optical cloud from external clouds.

Optical internetwork

An optical internetwork is a mesh-connected collection of optical networks. Each of these networks may be under a different administration.

Wavelength continuity property

A lightpath is said to satisfy the wavelength continuity property if it is transported over the same wavelength end-to-end. Wavelength continuity is required in optical networks with no wavelength conversion feature.

Wavelength path

A lightpath that satisfies the wavelength continuity property is called a wavelength path.

Opaque vs. transparent optical networks

A transparent optical network is an optical network in which optical signals are transported from transmitter to receiver entirely in the optical domain without OEO conversion. Generally, intermediate switching nodes in a transparent optical network do not have access to the payload carried by the optical signals.

Note that amplification of signals at transit nodes is permitted in transparent optical networks (e.g., using Erbium Doped Fiber Amplifiers << EDFAs).

On the other hand, in opaque optical networks, transit nodes may manipulate optical signals traversing through them. An example of such manipulation would be OEO conversion which may involve 3R operations (reshaping, retiming, regeneration, and perhaps amplification).

Trust domain

A trust domain is a network under a single technical administration in which adequate security measures are established to prevent unauthorized intrusion from outside the domain. Hence, it may be assumed that most nodes in the domain are deemed to be secure or trusted in some fashion. Generally, the rule for "single" administrative control over a trust domain may be relaxed in practice if a set of administrative entities agree to trust one another to form an enlarged heterogeneous trust domain. However, to simplify the discussions in this document, it will be assumed, without loss of generality, that the term trust domain applies to a single administrative entity with appropriate security policies. It should be noted that within a trust domain, any subverted node can send control messages which can compromise the entire network.

Flow

In this document, the term flow will be used to signify the smallest non-separable stream of data, from the point of view of an endpoint or termination point (source or destination node). The reader should note that the term flow is heavily overloaded in contemporary networking literature. In this document, we will consider a wavelength to be a flow, under certain circumstances. However, if there is a method to partition the bandwidth of the wavelength, then each partition may be considered a flow, for example using time division multiplexing (TDM), it may be feasible to consider each quanta of time within a given wavelength as a flow.

Traffic Trunk

A traffic trunk is an abstraction of traffic flow traversing the same path between two access points which allows some characteristics and attributes of the traffic to be parameterized.

3. The Network Model

3.1. Network Interconnection

The network model considered in this memo consists of IP routers attached to an optical core internetwork, and connected to their peers over dynamically established switched optical channels. The optical core itself is assumed to be incapable of processing individual IP packets in the data plane.

The optical internetwork is assumed to consist of multiple optical networks, each of which may be administered by a different entity. Each optical network consists of sub-networks interconnected by optical fiber links in a general topology (referred to as an optical mesh network). This network may contain re-configurable optical equipment from a single vendor or from multiple vendors. In the near term, it may be expected that each sub-network will consist of switches from a single vendor. In the future, as standardization efforts mature, each optical sub-network may in fact contain optical switches from different vendors. In any case, each sub-network itself is assumed to be mesh-connected internally. In general, it can be expected that topologically adjacent OXCs in an optical mesh network will be connected via multiple, parallel (bi-directional) optical links. This network model is shown in Figure 1.

In this environment, an optical sub-network may consist entirely of all-optical OXCs or OXCs with optical-electrical-optical (OEO) conversion. Interconnection between sub-networks is assumed to be implemented through compatible physical interfaces, with suitable

optical-electrical conversions where necessary. The routers that have direct physical connectivity with the optical network are referred to as "edge routers" with respect to the optical network. As shown in Figure 1, other client networks (e.g., ATM) may also connect to the optical network.

The switching function in an OXC is controlled by appropriately configuring the cross-connect fabric. Conceptually, this may be viewed as setting up a cross-connect table whose entries are of the form <input port i, output port j>, indicating that the data stream entering input port i will be switched to output port j. In the context of a wavelength selective cross-connect (generally referred to as a WXC), the cross-connect tables may also indicate the input and output wavelengths along with the input and output ports. A lightpath from an ingress port in an OXC to an egress port in a remote OXC is established by setting up suitable cross-connects in the ingress, the egress and a set of intermediate OXCs such that a continuous physical path exists from the ingress to the egress port. Optical paths tend to be bi-directional, i.e., the return path from the egress port to the ingress port is typically routed along the same set of intermediate interface cards as the forward path, but this may not be the case under all circumstances.

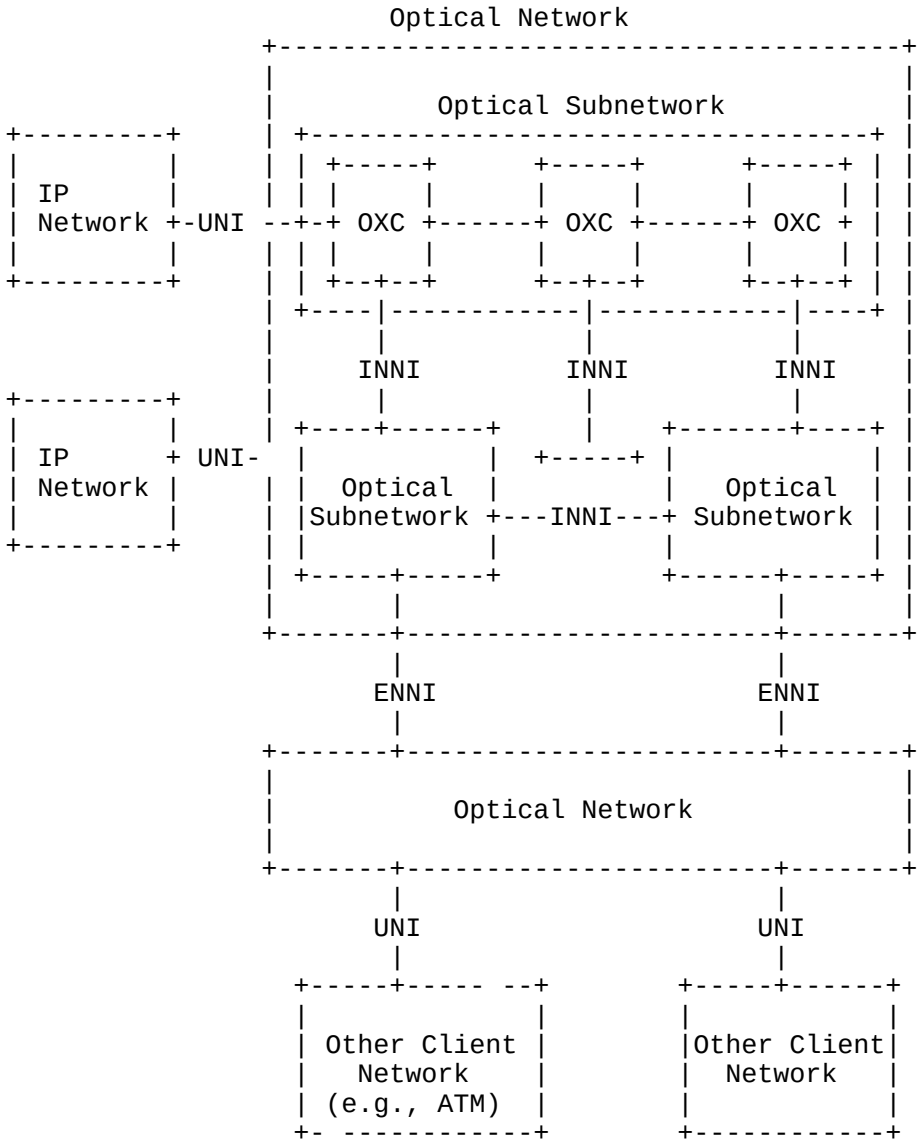

Figure 1: Optical Internetwork Model

Multiple traffic streams exiting from an OXC may be multiplexed onto a fiber optic link using WDM technology. The WDM functionality may exist outside of the OXC, and be transparent to the OXC. Or, this function may be built into the OXC. In the later case, the cross-connect table (conceptually) consists of pairs of the form, <{input port i, Lambda(j)}, {output port k, Lambda(l)}>. This indicates that

the data stream received on wavelength Lambda(j) over input port i is switched to output port k on Lambda(l). Automated establishment of lightpaths involves setting up the cross-connect table entries in the appropriate OXCs in a coordinated manner such that the desired physical path is realized.

Under this network model, a switched lightpath must be established between a pair of IP routers before the routers can transfer user traffic among themselves. A lightpath between IP routers may traverse multiple optical networks and be subject to different provisioning and restoration procedures in each network.

The IP-based control plane issue for optical networks pertains to the design of standard signaling and routing protocols for provisioning and restoration of lightpaths across multiple optical networks. Similarly, IP transport over optical networks involves establishing IP reachability and seamlessly constructing forwarding paths from one IP endpoint to another over an optical network.

3.2. Control Structure

There are three logical control interfaces identified in Figure 1. These are the client-optical internetwork interface, the internal node-to-node interface within an optical network (between OXCs in different sub-networks), and the external node-to-node interface between nodes in different optical networks. These interfaces are also referred to as the User-Network Interface (UNI), the internal NNI (INNI), and the external NNI (ENNI), respectively.

The distinction between these interfaces arises out of the type and amount of control information flow across them. The client-optical internetwork interface (UNI) represents a service boundary between the client (e.g., IP router) and the optical network. The client and server (optical network) are essentially two different roles: the client role requests a service connection from a server; the server role establishes the connection to fulfill the service request -- provided all relevant admission control conditions are satisfied.

Thus, the control flow across the client-optical internetwork interface is dependent on the set of services defined across it and the manner in which the services may be accessed. The service models are described in Section 4. The NNIs represent vendor-independent standardized interfaces for control flow between nodes. The distinction between the INNI and the ENNI is that the former is an interface within a given network under a single technical administration, while the later indicates an interface at the administrative boundary between networks. The INNI and ENNI may thus differ in the policies that restrict control flow between nodes.

Security, scalability, stability, and information hiding are important considerations in the specification of the ENNI. It is possible in principle to harmonize the control flow across the UNI and the NNI and eliminate the distinction between them. On the other hand, it may be required to minimize flow of control information, especially routing-related information, over the UNI; and even over the ENNI. In this case, UNI and NNIs may look different in some respects. In this document, these interfaces are treated as distinct.

The client-optical internetwork interface can be categorized as public or private depending upon context and service models. Routing information (i.e., topology state information) can be exchanged across a private client-optical internetwork interface. On the other hand, such information is not exchanged across a public client-optical internetwork interface, or such information may be exchanged with very explicit restrictions (including, for example abstraction, filtration, etc). Thus, different relationships (e.g., peer or over-lay, Section 5) may occur across private and public logical interfaces.

The physical control structure used to realize these logical interfaces may vary. For instance, for the client-optical internetwork interface, some of the possibilities are:

1. Direct interface: An in-band or out-of-band IP control channel (IPCC) may be implemented between an edge router and each OXC to which it is connected. This control channel is used for exchanging signaling and routing messages between the router and the OXC. With a direct interface, the edge router and the OXC it connects to are peers with respect to the control plane. This situation is shown in Figure 2. The type of routing and signaling information exchanged across the direct interface may vary depending on the service definition. This issue is addressed in the next section. Some choices for the routing protocol are OSPF or ISIS (with traffic engineering extensions and additional enhancements to deal with the peculiar characteristics of optical networks) or BGP, or some other protocol. Other directory-based routing information exchanges are also possible. Some of the signaling protocol choices are adaptations of RSVP-TE or CR-LDP. The details of how the IP control channel is realized is outside the scope of this document.

2. Indirect interface: An out-of-band IP control channel may be implemented between the client and a device in the optical network to signal service requests and responses. For instance, a management system or a server in the optical network may receive service requests from clients. Similarly, out-of-band signaling

may be used between management systems in client and optical networks to signal service requests. In these cases, there is no direct control interaction between clients and respective OXCs. One reason to have an indirect interface would be that the OXCs and/or clients do not support a direct signaling interface.

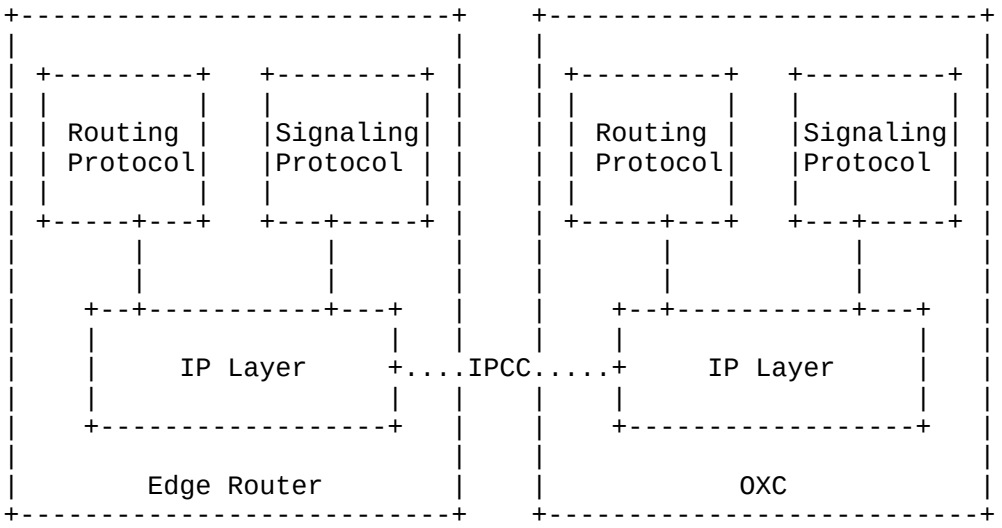

Figure 2: Direct Interface

3. Provisioned interface: In this case, the optical network services are manually provisioned and there is no control interactions between the client and the optical network.

Although different control structures are possible, further descriptions in this framework assume direct interfaces for IP-optical and optical sub-network control interactions.

4. IP over Optical Service Models and Requirements

In this section, the service models and requirements at the UNI and the NNIs are considered. Two general models have emerged for the services at the UNI (which can also be applied at the NNIs). These models are as follows.

4.1. Domain Services Model

Under the domain services model, the optical network primarily offers high bandwidth connectivity in the form of lightpaths. Standardized signaling across the UNI (Figure 1) is used to invoke the following services:

1. Lightpath creation: This service allows a lightpath with the specified attributes to be created between a pair of termination points in the optical network. Lightpath creation may be subject to network-defined policies (e.g., connectivity restrictions) and security procedures.

2. Lightpath deletion: This service allows an existing lightpath to be deleted.

3. Lightpath modification: This service allows certain parameters of the lightpath to be modified.

4. Lightpath status enquiry: This service allows the status of certain parameters of the lightpath (referenced by its ID) to be queried by the router that created the lightpath.

An end-system discovery procedure may be used over the UNI to verify local port connectivity between the optical and client devices, and allows each device to bootstrap the UNI control channel. Finally, a "service discovery" procedure may be employed as a precursor to obtaining UNI services. Service discovery allows a client to determine the static parameters of the interconnection with the optical network, including the UNI signaling protocols supported. The protocols for neighbor and service discovery are different from the UNI signaling protocol itself (for example, see LMP [2]).

Because a small set of well-defined services is offered across the UNI, the signaling protocol requirements are minimal. Specifically, the signaling protocol is required to convey a few messages with certain attributes in a point-to-point manner between the router and the optical network. Such a protocol may be based on RSVP-TE or LDP, for example.

The optical domain services model does not deal with the type and nature of routing protocols within and across optical networks.

The optical domain services model would result in the establishment of a lightpath topology between routers at the edge of the optical network. The resulting overlay model for IP over optical networks is discussed in Section 5.

4.2. Unified Service Model

Under this model, the IP and optical networks are treated together as a single integrated network from a control plane point of view. In this regard, the OXCs are treated just like any other router as far as the control plane is considered. Thus, in principle, there is no distinction between the UNI, NNIs and any other router-to-router

interface from a routing and signaling point of view. It is assumed that this control plane is IP-based, for example leveraging the traffic engineering extensions for MPLS or GMPLS, as described in [1]. The unified service model has so far been discussed only in the context of a single administrative domain. A unified control plane is possible even when there are administrative boundaries within an optical internetwork, but some of the integrated routing capabilities may not be practically attractive or even feasible in this case (see Section 5).

Under the unified service model and within the context of a GMPLS network, optical network services are obtained implicitly during end-to-end GMPLS signaling. Specifically, an edge router can create a lightpath with specified attributes, or delete and modify lightpaths as it creates GMPLS label-switched paths (LSPs). In this regard, the services obtained from the optical network are similar to the domain services model. These services, however, may be invoked in a more seamless manner as compared to the domain services model. For instance, when routers are attached to a single optical network (i.e., there are no ENNIs), a remote router could compute an end-to-end path across the optical internetwork. It can then establish an LSP across the optical internetwork. But the edge routers must still recognize that an LSP across the optical internetwork is a lightpath, or a conduit for multiple packet-based LSPs.

The concept of "forwarding adjacency" can be used to specify virtual links across optical internetworks in routing protocols such as OSPF [3]. In essence, once a lightpath is established across an optical internetwork between two edge routers, the lightpath can be advertised as a forwarding adjacency (a virtual link) between these routers. Thus, from a data plane point of view, the lightpaths result in a virtual overlay between edge routers. The decisions as to when to create such lightpaths, and the bandwidth management for these lightpaths is identical in both the domain services model and the unified service model. The routing and signaling models for unified services is described in Sections 5 and 6.

4.3. Which Service Model?

The relative merits of the above service models can be debated at length, but the approach recommended in this framework is to define routing and signaling mechanisms in support of both models. As noted above, signaling for service requests can be unified to cover both models. The developments in GMPLS signaling [4] for the unified service model and its adoption for UNI signaling [5, 6] under the domain services model essentially supports this view. The significant difference between the service models, however, is in routing protocols, as described in Sections 5 and 6.

4.4. What are the Possible Services?

Specialized services may be built atop the point-to-point connectivity service offered by the optical network. For example, optical virtual private networks and bandwidth on demand are some of the services that can be envisioned.

4.4.1. Optical Virtual Private Networks (OVPNs)

Given that the data plane links between IP routers over an optical network amounts to a virtual topology which is an overlay over the fiber optic network, it is easy to envision a virtual private network of lightpaths that interconnect routers (or any other set of clients) belonging to a single entity or a group of related entities across a public optical network. Indeed, in the case where the optical network provides connectivity for multiple sets of external client networks, there has to be a way to enforce routing policies that ensure routing separation between different sets of client networks (i.e., VPN service).

5. IP transport over Optical Networks

To examine the architectural alternatives for IP over optical networks, it is important to distinguish between the data and control planes. The optical network provides a service to external entities in the form of fixed bandwidth transport pipes (optical paths). IP routers at the edge of the optical networks must necessarily have such paths established between them before communication at the IP layer can commence. Thus, the IP data plane over optical networks is realized over a virtual topology of optical paths. On the other hand, IP routers and OXCs can have a peer relation with respect to the control plane, especially for routing protocols that permit the dynamic discovery of IP endpoints attached to the optical network.

The IP over optical network architecture is defined essentially by the organization of the control plane. The assumption in this framework is that an IP-based control plane [1] is used, such as GMPLS. Depending on the service model(Section 4), however, the control planes in the IP and optical networks can be loosely or tightly coupled. This coupling determines the following characteristics:

- The details of the topology and routing information advertised by the optical network across the client interface;
- The level of control that IP routers can exercise in selecting explicit paths for connections across the optical network;

- o Policies regarding the dynamic provisioning of optical paths between routers. These include access control, accounting, and security issues.

The following interconnection models are then possible:

5.1. Interconnection Models

5.1.1. The Peer Model

Under the peer model, the IP control plane acts as a peer of the optical transport network control plane. This implies that a single instance of the control plane is deployed over the IP and optical domains. When there is a single optical network involved and the IP and optical domains belong to the same entity, then a common IGP such as OSPF or IS-IS, with appropriate extensions, can be used to distribute topology information [7] over the integrated IP-optical network. In the case of OSPF, opaque LSAs can be used to advertise topology state information. In the case of IS-IS, extended TLVs will have to be defined to propagate topology state information. Many of these extensions are occurring within the context of GMPLS.

When an optical internetwork with multiple optical networks is involved (e.g., spanning different administrative domains), a single instance of an intra-domain routing protocol is not attractive or even realistic. In this case, inter-domain routing and signaling protocols are needed. In either case, a tacit assumption is that a common addressing scheme will be used for the optical and IP networks. A common address space can be trivially realized by using IP addresses in both IP and optical domains. Thus, the optical network elements become IP addressable entities as noted in [1].

5.1.2. The Overlay Model

Under the overlay model, the IP layer routing, topology distribution, and signaling protocols are independent of the routing, topology distribution, and signaling protocols within the optical domain. This model is conceptually similar to the classical IP over ATM or MPOA models, but applied to an optical internetwork instead. In the overlay model, a separate instance of the control plane (especially the routing and signaling protocols) would have to be deployed in the optical domain, independent of what exists in the IP domain. In certain circumstances, it may also be feasible to statically configure the optical channels that provide connectivity for the IP domain in the overlay model. Static configuration can be effected through network management functions. Static configuration, however,

is unlikely to scale in very large networks, and may not support the rapid connection provisioning requirements of future highly competitive networking environments.

5.1.3. The Augmented Model

Under the augmented model, there are separate routing instances in the IP and optical domains, but certain types of information from one routing instance can be passed through to the other routing instance. For example, external IP addresses could be carried within the optical routing protocols to allow reachability information to be passed to IP clients.

The routing approaches corresponding to these interconnection models are described below.

5.2. Routing Approaches

5.2.1. Integrated Routing

This routing approach supports the peer model within a single administrative domain. Under this approach, the IP and optical networks are assumed to run the same instance of an IP routing protocol, e.g., OSPF with suitable "optical" extensions. These extensions must capture optical link parameters, and any constraints that are specific to optical networks. The topology and link state information maintained by all nodes (OXCs and routers) may be identical, but not necessarily. This approach permits a router to compute an end-to-end path to another router across the optical network. Suppose the path computation is triggered by the need to route a label switched path (LSP) in a GMPLS environment. Such an LSP can be established using GMPLS signaling, e.g., RSVP-TE or CR-LDP with appropriate extensions. In this case, the signaling protocol will establish a lightpath between two edge routers. This lightpath is in essence a tunnel across the optical network, and may have capacity much larger than the bandwidth required to support the first LSP. Thus, it is essential that other routers in the network realize the availability of excess capacity within the lightpath so that subsequent LSPs between the routers can use it rather than instantiating a new lightpath. The lightpath may therefore be advertised as a virtual link in the topology as a means to address this issue.

The notion of "forwarding adjacency" (FA) described in [3] is essential in propagating existing lightpath information to other routers. An FA is essentially a virtual link advertised into a link state routing protocol. Thus, an FA could be described by the same parameters that define resources in any regular link. While it is

necessary to specify the mechanism for creating an FA, it is not necessary to specify how an FA is used by the routing scheme. Once an FA is advertised in a link state protocol, its usage for routing LSPs is defined by the route computation and traffic engineering algorithms implemented.

It should be noted that at the IP-optical interface, the physical ports over which routers are connected to OXCs constrain the connectivity and resource availability. Suppose a router R1 is connected to OXC O1 over two ports, P1 and P2. Under integrated routing, the connectivity between R1 and O1 over the two ports would have been captured in the link state representation of the network. Now, suppose an FA at full port bandwidth is created from R1 to another router R2 over port P1. While this FA is advertised as a virtual link between R1 and R2, it is also necessary to remove the link R1-O1 (over P1) from the link state representation since that port is no longer available for creating a lightpath. Thus, as FAs are created, an overlaid set of virtual links is introduced into the link state representation, replacing the links previously advertised at the IP-Optical interface. Finally, the details of the optical network captured in the link state representation is replaced by a network of FAs. The above scheme is one way to tackle the problem. Another approach is to associate appropriate dynamic attributes with link state information, so that a link that cannot be used to establish a particular type of connection will be appropriately tagged. Generally, however, there is a great deal of similarity between integrated routing and domain-specific routing (described next). Both ultimately deal with the creation of a virtual lightpath topology (which is overlaid over the optical network) to meet certain traffic engineering objectives.

5.2.2. Domain-Specific Routing

The domain-specific routing approach supports the augmented interconnection model. Under this approach, routing within the optical and IP domains are separated, with a standard routing protocol running between domains. This is similar to the IP inter-domain routing model. A specific approach for this is considered next. It is to be noted that other approaches are equally possible.

5.2.2.1. Domain-Specific Routing using BGP

The inter-domain IP routing protocol, BGP [8], may be adapted for exchanging routing information between IP and optical domains. This would allow routers to advertise IP address prefixes within their network to the optical internetwork and to receive external IP address prefixes from the optical internetwork. The optical internetwork transports the reachability information from one IP

network to others. For instance, edge routers and OXCs can run exterior BGP (EBGP). Within the optical internetwork, interior BGP (IBGP) is may be used between border optical switches, and EBGP may be used between different networks (over ENNI, Figure 1).

Under this scheme, it may be necessary to identify the egress points in the optical internetwork corresponding to externally reachable IP addresses. To see this, suppose an edge router intends to establish an LSP to a destination node across the optical internetwork. It may request a direct lightpath to that destination, without explicitly specifying the egress optical port for the lightpath because the optical internetwork has knowledge of externally reachable IP addresses. However, if the same edge router were to establish another LSP to a different external destination, then for efficiency reasons, it may first need to determine whether there is an existing lightpath (with sufficient residual capacity) to the target destination. For this purpose, it may be necessary for edge routers to keep track of which egress ports in the optical internetwork lead to which external destinations. Thus, a border OXC receiving external IP prefixes from an edge router through EBGP must include its own IP address as the egress point before propagating these prefixes to other border OXCs or edge routers. An edge router receiving this information need not propagate the egress address further, but it must keep the association between external IP addresses and egress OXC addresses. When optical VPNs are implemented, the address prefixes advertised by the border OXCs may be accompanied by some VPN specific identification.

There are however, some potential negative effects that could result from domain-specific routing using BGP in an IPO environment:

- The amount of information that optical nodes will have to maintain will not be bound by the size of the optical network anymore, but will have to include external routes as well.

- The stability of the optical network control plane will no longer be dictated solely by the dynamics emanating within the optical network, but may be affected by the dynamics originating from external routing domains from which external reachability information is received.

5.2.3. Overlay Routing

The overlay routing approach supports the overlay interconnection model. Under this approach, an overlay mechanism that allows edge routers to register and query for external addresses is implemented. This is conceptually similar to the address resolution mechanism used for IP over ATM. Under this approach, the optical network could

implement a registry that allows edge routers to register IP addresses and VPN identifiers. An edge router may be allowed to query for external addresses belonging to the same set of VPNs it belongs to. A successful query would return the address of the egress optical port through which the external destination can be reached.

Because IP-optical interface connectivity is limited, the determination of how many lightpaths must be established and to what endpoints are traffic engineering decisions. Furthermore, after an initial set of such lightpaths are established, these may be used as adjacencies within VPNs for a VPN-wide routing scheme, for example, OSPF. With this approach, an edge router could first determine other edge routers of interest by querying the registry. After it obtains the appropriate addresses, an initial overlay lightpath topology may be formed. Routing adjacencies may then be established across the lightpaths and further routing information may be exchanged to establish VPN-wide routing.

5.3. Signaling-Related

5.3.1. The Role of MPLS

It is possible to model wavelengths, and potentially TDM channels within a wavelength as "labels". This concept was proposed in [1], and "generalized" MPLS (GMPLS) mechanisms for realizing this are described in [4]. MPLS signaling protocols with traffic engineering extensions, such as RSVP-TE, can be appropriately extended and used for signaling lightpath requests. These protocols can be adapted for client/server signaling in the case of the domain services model, and for end-to-end integrated signaling in the case of the unified services model.

5.3.2. Signaling Models

With the domain-services model, the signaling control plane in the IP and optical network are completely separate as shown in Figure 3 below. This separation also implies the separation of IP and optical address spaces (even though the optical network would be using internal IP addressing). While RSVP-TE and LDP can be adapted for UNI signaling, the full functionality of these protocols will not be used. For example, UNI signaling does not require the specification of explicit routes. On the other hand, based on the service attributes, new objects need to be signaled using these protocols as described in [5, 6].

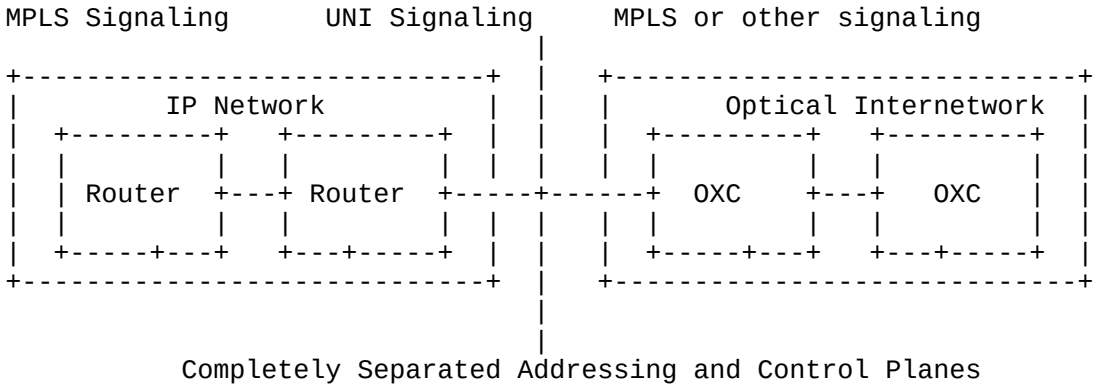

Figure 3: Domain Services Signaling Model

With the unified services model, the addressing is common in the IP network and optical internetwork and the respective signaling control are related, as shown in Figure 4. It is understood that GMPLS signaling is implemented in the IP and optical domains, using suitably enhanced RSVP-TE or CR-LDP protocols. But the semantics of services within the optical internetwork may be different from that in the IP network. As an example, the protection services offered in the optical internetwork may be different from the end-to-end protection services offered by the IP network. Another example is with regard to bandwidth. While the IP network may offer a continuum of bandwidths, the optical internetwork will offer only discrete bandwidths. Thus, the signaling attributes and services are defined independently for IP and optical domains. The routers at the edge of the optical internetwork must therefore identify service boundaries and perform suitable translations in the signaling messages crossing the IP-optical boundary. This may still occur even though the signaling control plane in both networks are GMPLS-based and there is tighter coupling of the control plane as compared to the domain services model.

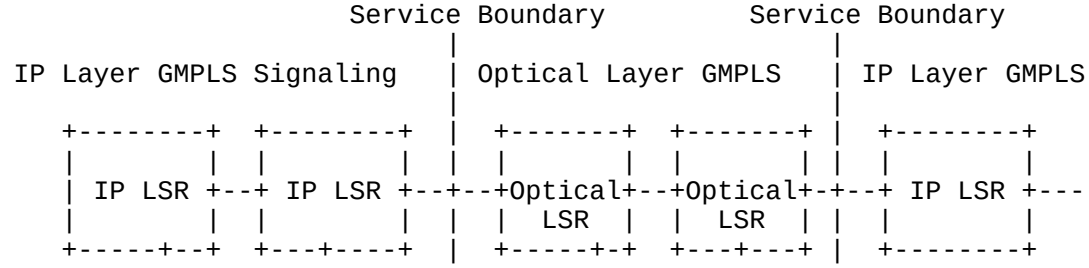

Figure 4: Unified Services Signaling Model

Thus, as illustrated in Figure 4, the signaling in the case of unified services is actually multi-layered. The layering is based on the technology and functionality. As an example, the specific adaptations of GMPLS signaling for SONET layer (whose functionality is transport) are described in [10].

5.4. End-to-End Protection Models

Suppose an LSP is established from an ingress IP router to an egress router across an ingress IP network, a transit optical internetwork and an egress IP network. If this LSP is to be afforded protection in the IP layer, how is the service coordinated between the IP and optical layers?

Under this scenario, there are two approaches to end-to-end protection:

5.4.1. Segment-Wise Protection

The protection services in the IP layer could utilize optical layer protection services for the LSP segment that traverses the optical internetwork. Thus, the end-to-end LSP would be treated as a concatenation of three LSP segments from the protection point of view: a segment in the ingress IP network, a segment in the optical internetwork and a segment in the egress IP network. The protection services at the IP layer for an end-to-end LSP must be mapped onto suitable protection services offered by the optical internetwork. Suppose that 1+1 protection is offered to LSPs at the IP layer, i.e., each protected LSP has a pre-established hot stand-by in a 1+1 or 1:1 configuration. In case of a failure of the primary LSP, traffic can be immediately switched to the stand-by. This type of protection can be realized end-to-end as follows. With reference to Figure 5, let an LSP originate at (ingress) router interface A and terminate at (egress) router interface F. Under the first protection option, a

primary path for the LSP must be established first. Let this path be as shown in Figure 5, traversing router interface B in the ingress network, optical ports C (ingress) and D (egress), and router interface E in the egress network. Next, 1+1 protection is realized separately in each network by establishing a protection path between points A and B, C and D and E and F. Furthermore, the segments B-C and D-E must themselves be 1+1 protected, using drop- side protection. For the segment between C and D, the optical internetwork must offer a 1+1 service similar to that offered in the IP networks.

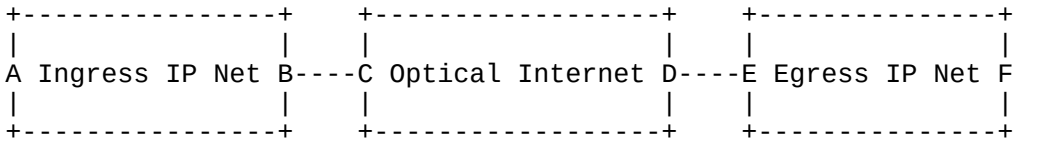

Figure 5: End-to-End Protection Example

5.4.2. Single-Layer Protection

Under this model, the protection services in the IP layer do not rely on any protection services offered in the optical internetwork. Thus, with reference to Figure 5, two SRLG-disjoint LSPs are established between A and F. The corresponding segments in the optical internetwork are treated as independent lightpaths in the optical internetwork. These lightpaths may be unprotected in the optical internetwork.

5.4.3. Differences

A distinction between these two choices is as follows. Under the first choice, the optical internetwork is actively involved in end-to-end protection, whereas under the second choice, any protection service offered in the optical internetwork is not utilized directly by client IP network. Also, under the first choice, the protection in the optical internetwork may apply collectively to a number of IP LSPs. That is, with reference to Figure 5, many LSPs may be aggregated into a single lightpath between C and D. The optical internetwork protection may then be applied to all of them at once leading to some gain in scalability. Under the second choice, each IP LSP must be separately protected. Finally, the first choice allows different restoration signaling to be implemented in the IP and optical internetwork. These restoration protocols are "patched up" at the service boundaries to realize end-to-end protection. A further advantage of this is that restoration is entirely contained within the network where the failure occurs, thereby improving the restoration latency, and perhaps network stability as a fault within

an optical domain is contained and corrected within the domain. For instance, if there is a failure in the optical internetwork, optical network protocols restore the affected internal segments. Under the second choice, restoration signaling is always end-to-end between IP routers, essentially by-passing the optical internetwork. A result of this is that restoration latency could be higher. In addition, restoration protocols in the IP layer must run transparently over the optical internetwork in the overlay mode. IP based recovery techniques may however be more resource efficient, as it may be possible to convey traffic through the redundant capacity under fault-free scenarios. In particular, it may be possible to utilize classification, scheduling, and concepts of forwarding equivalence class to route lower class traffic over protect facilities and then possibly preempt them to make way for high priority traffic when faults occur.

6. IP-based Optical Control Plane Issues

Provisioning and restoring lightpaths end-to-end between IP networks requires protocol and signaling support within optical sub-networks, and across the INNI and ENNI. In this regard, a distinction is made between control procedures within an optical sub-network (Figure 1), between sub-networks, and between networks. The general guideline followed in this framework is to separate these cases, and allow the possibility that different control procedures are followed inside different sub-networks, while a common set of procedures are followed across sub-networks and networks.

The control plane procedures within a single vendor sub-network need not be defined since these can be proprietary. Clearly, it is possible to follow the same control procedures inside a sub-network and across sub-networks. But this is simply a recommendation within this framework document, rather than an imperative requirement. Thus, in the following, signaling and routing across sub-networks is considered first, followed by a discussion of similar issues across networks.

6.1. Addressing

For interoperability across optical sub-networks using an IP-centric control plane, one of the fundamental issues is that of addressing. What entities should be identifiable from a signaling and routing point of view? How should they be addressed? This section presents some high level guidelines on this issue.

Identifiable entities in optical networks include OXCs, optical links, optical channels and sub-channels, Shared Risk Link Groups (SRLGs), etc. An issue here is how granular the identification should be as far as the establishment of optical trails are concerned. The scheme for identification must accommodate the specification of the termination points in the optical network with adequate granularity when establishing optical trails. For instance, an OXC could have many ports, each of which may in turn terminate many optical channels, each of which contain many sub-channels etc. It is perhaps not reasonable to assume that every sub-channel or channel termination, or even OXC ports could be assigned a unique IP address. Also, the routing of an optical trail within the network does not depend on the precise termination point information, but rather only on the terminating OXC. Thus, finer granularity identification of termination points is of relevance only to the terminating OXC and not to intermediate OXCs (of course, resource allocation at each intermediate point would depend on the granularity of resources requested). This suggests an identification scheme whereby OXCs are identified by a unique IP address and a "selector" identifies further fine-grain information of relevance at an OXC. This, of course, does not preclude the identification of these termination points directly with IP addresses(with a null selector). The selector can be formatted to have adequate number of bits and a structure that expresses port, channel, sub-channel, etc, identification.

Within the optical network, the establishment of trail segments between adjacent OXCs require the identification of specific port, channel, sub-channel, etc. With a GMPLS control plane, a label serves this function. The structure of the label must be such that it can encode the required information [10].

Another entity that must be identified is the SRLG [11]. An SRLG is an identifier assigned to a group of optical links that share a physical resource. For instance, all optical channels routed over the same fiber could belong to the same SRLG. Similarly, all fibers routed over a conduit could belong to the same SRLG. The notable characteristic of SRLGs is that a given link could belong to more than one SRLG, and two links belonging to a given SRLG may individually belong to two other SRLGs. This is illustrated in Figure 6. Here, the links 1,2,3 and 4 may belong to SRLG 1, links 1,2 and 3 could belong to SRLG 2 and link 4 could belong to SRLG 3. Similarly, links 5 and 6 could belong to SRLG 1, and links 7 and 8 could belong to SRLG 4. (In this example, the same SRLG, i.e., 1, contains links from two different adjacencies).

While the classification of physical resources into SRLGs is a manual operation, the assignment of unique identifiers to these SRLGs within an optical network is essential to ensure correct SRLG-disjoint path computation for protection. SRLGs could be identified with a flat identifier (e.g., 32 bit integer).

Finally, optical links between adjacent OXCs may be bundled for advertisement into a link state protocol [12]. A bundled interface may be numbered or unnumbered. In either case, the component links within the bundle must be identifiable. In concert with SRLG identification, this information is necessary for correct path computation.

6.2. Neighbor Discovery

Routing within the optical network relies on knowledge of network topology and resource availability. This information may be gathered and used by a centralized system, or by a distributed link state routing protocol. In either case, the first step towards network-wide link state determination is the discovery of the status of local links to all neighbors by each OXC. Specifically, each OXC must determine the up/down status of each optical link, the bandwidth and other parameters of the link, and the identity of the remote end of the link (e.g., remote port number). The last piece of information is used to specify an appropriate label when signaling for lightpath provisioning. The determination of these parameters could be based on a combination of manual configuration and an automated protocol running between adjacent OXCs. The characteristics of such a protocol would depend on the type of OXCs that are adjacent (e.g., transparent or opaque).

Neighbor discovery would typically require in-band communication on the bearer channels to determine local connectivity and link status. In the case of opaque OXCs with SONET termination, one instance of a neighbor discovery protocol (e.g., LMP [2]) would run on each OXC port, communicating with the corresponding protocol instance at the neighboring OXC. The protocol would utilize the SONET overhead bytes to transmit the (configured) local attributes periodically to the neighbor. Thus, two neighboring switches can automatically determine the identities of each other and the local connectivity, and also keep track of the up/down status of local links. Neighbor discovery with transparent OXCs is described in [2].

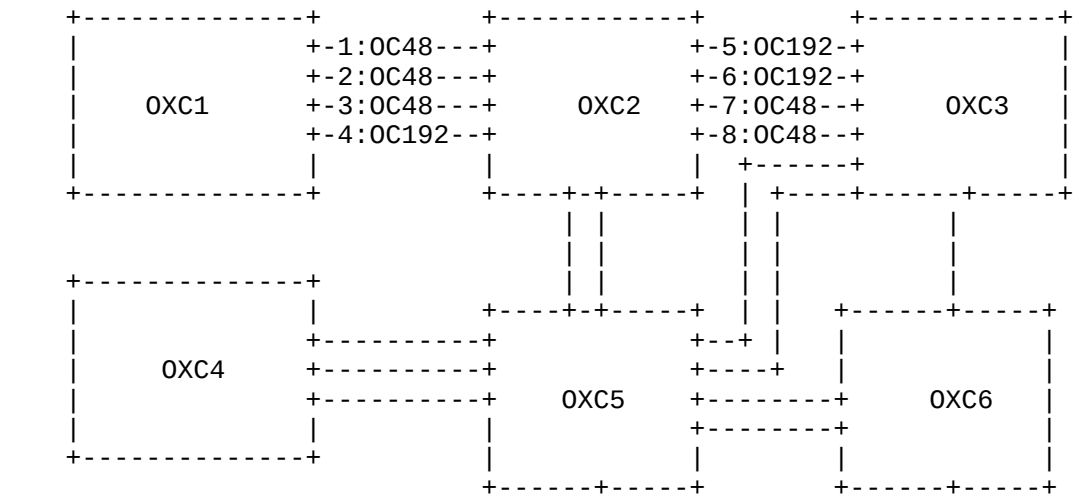

Figure 6: Mesh Optical Network with SRLGs

6.3. Topology Discovery

Topology discovery is the procedure by which the topology and resource state of all the links in a network are determined. This procedure may be done as part of a link state routing protocol (e.g., OSPF, ISIS), or it can be done via the management plane (in the case of centralized path computation). The implementation of a link state protocol within a network (i.e., across sub-network boundaries) means that the same protocol runs in OXCs in every sub-network. If this assumption does not hold then interworking of routing between sub-networks is required. This is similar to inter-network routing discussed in Section 6.7. The focus in the following is therefore on standardized link state routing.

In general, most of the link state routing functionality is maintained when applied to optical networks. However, the representation of optical links, as well as some link parameters, are changed in this setting. Specifically,

- The link state information may consist of link bundles [12]. Each link bundle is represented as an abstract link in the network topology. Different bundling representations are possible. For instance, the parameters of the abstract link may include the number, bandwidth and the type of optical links contained in the underlying link bundle [12]. Also, the SRLGs corresponding to each optical link in the bundle may be included as a parameter.

- The link state information should capture restoration-related parameters for optical links. Specifically, with shared protection (Section 6.5), the link state updates must have information that allows the computation of shared protection paths.

- A single routing adjacency could be maintained between neighbors which may have multiple optical links (or even multiple link bundles) between them. This reduces the protocol messaging overhead.

- Since link availability information changes dynamically, a flexible policy for triggering link state updates based on availability thresholds may be implemented. For instance, changes in availability of links of a given bandwidth (e.g., OC-48) may trigger updates only after the availability figure changes by a certain percentage.

These concepts are relatively well-understood. On the other hand, the resource representation models and the topology discovery process for hierarchical routing (e.g., OSPF with multiple areas) are areas that need further work.

6.4. Protection and Restoration Models

Automatic restoration of lightpaths is a service offered by optical networks. There could be local and end-to-end mechanisms for restoration of lightpaths within a network (across the INNI). Local mechanisms are used to select an alternate link (or network segment) between two OXCs across the INNI when a failure affects the primary link (or primary network segment) over which the (protected) lightpath is routed. Local restoration does not affect the end-to-end route of the lightpath. When local restoration is not possible (e.g., no alternate link is available between the adjacent OXCs in question), end-to-end restoration may be performed. Under this scenario this, the affected lightpath may be rerouted over an alternate diverse path to circumvent failed resources. For end-to-end restoration, alternate paths may be pre-computed to expedite the recovery time. End to end restoration may also be mixed with local recovery in various ways depending on acceptable tradeoffs between utilization of network resources and recovery times.

End-to-end protection may be based on two types of protection schemes; "1 + 1" protection or shared protection. Under 1 + 1 protection, a back-up path is established for the protected primary path along a physically diverse route. Both paths are active and the failure along the primary path results in an immediate switch-over to the back-up path. Under shared protection, back-up paths

corresponding to physically diverse primary paths may share the same network resources. When a failure affects a primary path, it is assumed that the same failure will not affect the other primary paths whose back-ups share resources.

It is possible that different restoration schemes may be implemented within optical sub-networks. It is therefore necessary to consider a two-level restoration mechanism. Path failures within an optical sub-network could be handled using procedures specific to the sub-network. If this fails, end-to-end restoration across sub-networks could be invoked. The border OXC that is the ingress to a sub-network can act as the source for restoration procedures within a sub-network. The signaling for invoking end-to-end restoration across the INNI is described in Section 6.6.3. The computation of the back-up path for end-to-end restoration may be based on various criteria. It is assumed that the back-up path is computed by the source OXC, and signaled using standard methods.

6.5. Route Computation

The computation of a primary route for a lightpath within an optical network is essentially a constraint-based routing problem. The constraint is typically the bandwidth required for the lightpath, perhaps along with administrative and policy constraints. The objective of path computation could be to minimize the total capacity required for routing lightpaths [13].

Route computation with constraints may be accomplished using a number of algorithms [14]. When 1+1 protection is used, a back-up path that does not traverse on any link which is part of the same SRLG as links in the primary path must be computed. Thus, it is essential that the SRLGs in the primary path be known during alternate path computation, along with the availability of resources in links that belong to other SRLGs. This requirement has certain implications on optical link bundling. Specifically, a bundled LSA must include adequate information such that a remote OXC can determine the resource availability under each SRLG that the bundled link refers to, and the relationship between links belonging to different SRLGs in the bundle. For example, considering Figure 3, if links 1,2,3 and 4 are bundled together in an LSA, the bundled LSA must indicate that there are three SRLGs which are part of the bundle (i.e., 1, 2 and 3), and that links in SRLGs 2 and 3 are also part of SRLG 1.

To encode the SRLG relationships in a link bundle LSA, only links which belong to exactly the same set of SRLGs must be bundled together. With reference to Figure 3, for example, two bundles can be advertised for links between OXC1 and OXC2, with the following information:

Bundle No.	SRLGs	Link Type	Number	Other Info
1	1,2	OC-48	3	---
2	1,3	OC-192	1	---

Assuming that the above information is available for each bundle at every node, there are several approaches possible for path computation. For instance,

1. The primary path can be computed first, and the (exclusive or shared) back-up is computed next based on the SRLGs chosen for the primary path. In this regard,

   - The primary path computation procedure can output a series of bundles the path is routed over. Since a bundle is uniquely identified with a set of SRLGs, the alternate path can be computed right away based on this knowledge. In this case, if the primary path set up does not succeed for lack of resources in a chosen bundle, the primary and backup paths must be recomputed.

   - It might be desirable to compute primary paths without choosing a specific bundle apriori. That is, resource availability over all bundles between a node pair is taken into account rather than specific bundle information. In this case, the primary path computation procedure would output a series of nodes the path traverses. Each OXC in the path would have the freedom to choose the particular bundle to route that segment of the primary path. This procedure would increase the chances of successfully setting up the primary path when link state information is not up to date everywhere. But the specific bundle chosen, and hence the SRLGs in the primary path, must be captured during primary path set-up, for example, using the RSVP-TE Route Record Object [15]. This SRLG information is then used for computing the back-up path. The back-up path may also be established specifying only which SRLGs to avoid in a given segment, rather than which bundles to use. This would maximize the chances of establishing the back-up path.

2. The primary path and the back-up path are computed together in one step, for example, using Suurbaale's algorithm [16]. In this case, the paths must be computed using specific bundle information.

To summarize, it is essential to capture sufficient information in link bundle LSAs to accommodate different path computation procedures and to maximize the chances of successful path establishment. Depending on the path computation procedure used, the type of support

needed during path establishment (e.g., the recording of link group or SRLG information during path establishment) may differ.

When shared protection is used, the route computation algorithm must take into account the possibility of sharing links among multiple back-up paths. Under shared protection, the back-up paths corresponding to SRLG-disjoint primary paths can be assigned the same links. The assumption here is that since the primary paths are not routed over links that have the same SRLG, a given failure will affect only one of them. Furthermore, it is assumed that multiple failure events affecting links belonging to more than one SRLG will not occur concurrently. Unlike the case of 1+1 protection, the back-up paths are not established apriori. Rather, a failure event triggers the establishment of a single back-up path corresponding to the affected primary path.

The distributed implementation of route computation for shared back-up paths require knowledge about the routing of all primary and back-up paths at every node. This raises scalability concerns. For this reason, it may be practical to consider the centralization of the route computation algorithm in a route server that has complete knowledge of the link state and path routes. Heuristics for fully distributed route computation without complete knowledge of path routes are to be determined. Path computation for restoration is further described in [11].

6.6. Signaling Issues

Signaling within an optical network for lightpath provisioning is a relatively simple operation if a standard procedure is implemented within all sub-networks. Otherwise, proprietary signaling may be implemented within sub-networks, but converted back to standard signaling across the INNI. This is similar to signaling across the ENNI, as described in Section 6.7. In the former case, signaling messages may carry strict explicit route information, while in the latter case the route information should be loose, at the level of abstraction of sub-networks. Once a route is determined for a lightpath, each OXC along the path must appropriately configure their cross-connects in a coordinated fashion. This coordination is conceptually analogous to selecting incoming and outgoing labels in a label-switched environment. Thus, protocols like RSVP-TE [9] may be adapted and used across the INNI for this purpose. The adaptation of IP-based signaling protocols must take into account a number of peculiar attributes of optical networks.

6.6.1. Bi-Directional Lightpath Establishment

Lightpaths are typically bi-directional. That is, the output port selected at an OXC for the forward direction is also the input port for the reverse direction of the path. Since signaling for optical paths may be autonomously initiated by different nodes, it is possible that two path set-up attempts are in progress at the same time. Specifically, while setting up an optical path, an OXC A may select output port i which is connected to input port j of the "next" OXC B. Concurrently, OXC B may select output port j for setting up a different optical path, where the "next" OXC is A. This results in a "collision". Similarly, when WDM functionality is built into OXCs, a collision occurs when adjacent OXCs choose directly connected output ports and the same wavelength for two different optical paths. There are two ways to deal with such collisions. First, collisions may be detected and the involved paths may be torn down and re-established. Or, collisions may be avoided altogether.

6.6.2. Failure Recovery

The impact of transient partial failures must be minimized in an optical network. Specifically, optical paths that are not directly affected by a failure must not be torn down due to the failure. For example, the control processor in an OXC may fail, affecting signaling and other internodal control communication. Similarly, the control channel between OXCs may be affected temporarily by a failure. These failure may not affect already established optical paths passing through the OXC fabric. The detection of such failures by adjacent nodes, for example, through a keepalive mechanism between signaling peers, must not result in these optical paths being torn down.

It is likely that when the above failures occur, a backup processor or a backup control channel will be activated. The signaling protocol must be designed such that it is resilient to transient failures. During failure recovery, it is desirable to recover local state at the concerned OXC with least disruption to existing optical paths.

6.6.3. Restoration

Signaling for restoration has two distinct phases. There is a reservation phase in which capacity for the protection path is established. Then, there is an activation phase in which the back-up path is actually put in service. The former phase typically is not subject to strict time constraints, while the latter is.

Signaling to establish a "1+1" back-up path is relatively straightforward. This signaling is very similar to signaling used for establishing the primary path. Signaling to establish a shared back-up path is a little bit different. Here, each OXC must understand which back-up paths can share resources among themselves. The signaling message must itself indicate shared reservation. The sharing rule is as described in Section 6.4: back-up paths corresponding to physically diverse primary paths may share the same network resources. It may therefore be necessary for the signaling message to carry adequate information that allows an OXC to verify that appropriateness of having a set of back-up paths sharing certain.

Under both 1+1 and shared protection, the activation phase has two parts: propagation of failure information to the source OXC from the point of failure, and activation of the back-up path. The signaling for these two phases must be very fast in order to realize response times in the order of tens of milliseconds. When optical links are SONET-based, in-band signals may be used, resulting in expedited response. With out-of-band control, it may be necessary to consider fast signaling over the control channel using very short IP packets and prioritized processing. While it is possible to use RSVP or CR-LDP for activating protection paths, these protocols do not provide any means to give priority to restoration signaling as opposed to signaling for provisioning. For instance, it is possible for a restoration-related RSVP message to be queued behind a number of provisioning messages thereby delaying restoration. It may therefore be necessary to develop a notion of prioritization for restoration signaling and incorporate appropriate mechanisms into existing signaling protocols to achieve this. Alternatively, a new signaling mechanism may be developed exclusively for activating protection paths during restoration.

6.7. Optical Internetworking

Within an optical internetwork, it must be possible to dynamically provision and restore lightpaths across optical networks. Therefore:

- A standard scheme for uniquely identifying lightpath end-points in different networks is required.
- A protocol is required for determining reachability of end-points across networks.
- A standard signaling protocol is required for provisioning lightpaths across networks.

- o A standard procedure is required for the restoration of lightpaths across networks.

- o Support for policies that affect the flow of control information across networks will be required.

The IP-centric control architecture for optical networks can be extended to satisfy the functional requirements of optical internetworking. Routing and signaling interaction between optical networks can be standardized across the ENNI (Figure 1). The functionality provided across ENNI is as follows.

6.7.1. Neighbor Discovery

Neighbor discovery procedure, as described in Section 6.2, can be used for this. Indeed, a single protocol should be standardized for neighbor discovery within and across networks.

6.7.2. Addressing and Routing Model

The addressing mechanisms described in Section 6.1 can be used to identify OXCs, ports, channels and sub-channels in each network. It is essential that the OXC IP addresses are unique within the internetwork.

Provisioning an end-to-end lightpath across multiple networks involves the establishment of path segments in each network sequentially. Thus, a path segment is established from the source OXC to a border OXC in the source network. From this border OXC, signaling across NNI is used to establish a path segment to a border OXC in the next network. Provisioning then continues in the next network and so on until the destination OXC is reached. The usage of protocols like BGP for this purpose need to be explored.

6.7.3. Restoration

Local restoration across the ENNI is similar to that across INNI described in Section 6.6.3. End-to-end restoration across networks is likely to be either of the 1+1 type, or segmented within each network, as described in Section 6.4.

7. Other Issues

7.1. WDM and TDM in the Same Network

A practical assumption would be that if SONET (or some other TDM mechanism that is capable partitioning the bandwidth of a wavelength) is used, then TDM is leveraged as an additional method to differentiate between "flows". In such cases, wavelengths and time intervals (sub-channels) within a wavelength become analogous to labels (as noted in [1]) which can be used to make switching decisions. This would be somewhat akin to using VPI (e.g., wavelength) and VCI (e.g., TDM sub-channel) in ATM networks. More generally, this will be akin to label stacking and to LSP nesting within the context of Multi-Protocol Lambda Switching [1]. GMPLS signaling [4] supports this type of multiplexing.

7.2. Wavelength Conversion

Some form of wavelength conversion may exist at some switching elements. This however may not be the case in some pure optical switching elements. A switching element is essentially anything more sophisticated than a simple repeater, that is capable of switching and converting a wavelength Lambda(k) from an input port to a wavelength Lambda(l) on an output port. In this display, it is not necessarily the case that Lambda(k) = Lambda(l), nor is it necessarily the case that the data carried on Lambda(k) is switched through the device without being examined or modified.

It is not necessary to have a wavelength converter at every switching element. A number of studies have attempted to address the issue of the value of wavelength conversion in an optical network. Such studies typically use the blocking probability (the probability that a lightpath cannot be established because the requisite wavelengths are not available) as a metric to adjudicate the effectiveness of wavelength conversion. The IP over optical architecture must take into account hybrid networks with some OXCs capable of wavelength conversion and others incapable of this. The GMPLS "label set" mechanism [4] supports the selection of the same label (i.e., wavelength) across an NNI.

7.3. Service Provider Peering Points

There are proposed inter-network interconnect models which allow certain types of peering relationships to occur at the optical layer. This is consistent with the need to support optical layer services independent of higher layers payloads. In the context of IP over optical networks, peering relationships between different trust domains will eventually have to occur at the IP layer, on IP routing

elements, even though non-IP paths may exist between the peering routers.

7.4. Rate of Lightpath Set-Up

Dynamic establishment of optical channel trails and lightpaths is quite desirable in IP over optical networks, especially when such instantiations are driven by a stable traffic engineering control system, or in response to authenticated and authorized requests from clients.

However, there are many proposals suggesting the use of dynamic, data-driven shortcut-lightpath setups in IP over optical networks. The arguments put forth in such proposals are quite reminiscent of similar discussions regarding ATM deployment in the core of IP networks. Deployment of highly dynamic data driven shortcuts within core networks has not been widely adopted by carriers and ISPs for a number of reasons: possible CPU overhead in core network elements, complexity of proposed solutions, stability concerns, and lack of true economic drivers for this type of service. This document assumes that this paradigm will not change and that highly dynamic, data-driven shortcut lightpath setups are for future investigation. Instead, the optical channel trails and lightpaths that are expected to be widely used at the initial phases in the evolution of IP over optical networks will include the following:

- Dynamic connections for control plane traffic and default path routed data traffic,
- Establishment and re-arrangement of arbitrary virtual topologies over rings and other physical layer topologies.
- Use of stable traffic engineering control systems to engineer lightpath connections to enhance network performance, either for explicit demand based QoS reasons or for load balancing).

Other issues surrounding dynamic connection setup within the core center around resource usage at the edge of the optical domain. One potential issue pertains to the number of flows that can be processed by an ingress or egress network element either because of aggregate bandwidth limitations or because of a limitation on the number of flows (e.g., lightpaths) that can be processed concurrently.

Another possible short term reason for dynamic shortcut lightpath setup would be to quickly pre-provision paths based on some criteria (e.g., a corporate executive wants a high bandwidth reliable connection, etc.). In this scenario, a set of paths can be pre-provisioned, but not actually instantiated until the customer

initiates an authenticated and authorized setup requests, which is consistent with existing agreements between the provider and the customer. In a sense, the provider may have already agreed to supply this service, but will only instantiate it by setting up a lightpath when the customer submits an explicit request.

7.5. Distributed vs. Centralized Provisioning

This document has mainly dealt with a distributed model for lightpath provisioning, in which all nodes maintain a synchronized topology database, and advertise topology state information to maintain and refresh the database. A constraint-based routing entity in each node then uses the information in the topology database and other relevant details to compute appropriate paths through the optical domain. Once a path is computed, a signaling protocol (e.g., [9]) is used to instantiate the lightpath.

Another provisioning model is to have a centralized server which has complete knowledge of the physical topology, the available wavelengths, and where applicable, relevant time domain information.

A corresponding client will reside on each network element that can source or sink a lightpath. The source client would query the server in order to set up a lightpath from the source to the destination. The server would then check to see if such a lightpath can be established based on prevailing conditions. Furthermore, depending on the specifics of the model, the server may either setup the lightpath on behalf of the client or provide the necessary information to the client or to some other entity to allow the lightpath to be instantiated.

Centralization aids in implementing complex capacity optimization schemes, and may be the near-term provisioning solution in optical networks with interconnected multi-vendor optical sub-networks. In the long term, however, the distributed solution with centralization of some control procedures (e.g., traffic engineering) is likely to be the approach followed.

7.6. Optical Networks with Additional Configurable Components

Thus far, this memo has focused mainly on IP over optical networks where the cross-connect is the basic dynamically re-configurable device in the optical network. Recently, as a consequence of technology evolution, various types of re-configurable optical components are now available, including tunable lasers, tunable filters, etc. Under certain circumstances, it may be necessary to

parameterize the characteristics of these components and advertise them within the control plane. This aspect is left for further study.

7.7. Optical Networks with Limited Wavelength Conversion Capability

At the time of the writing of this document, the majority of optical networks being deployed are "opaque". In this context the term opaque means that each link is optically isolated by transponders doing optical-electrical-optical conversions. Such conversions have the added benefit of permitting 3R regeneration. The 3Rs refer to re-power, signal retiming and reshaping. Unfortunately, this regeneration requires that the underlying optical equipment be aware of both the bit rate and frame format of the carried signal. These transponders are quite expensive and their lack of transparency constrains the rapid introduction of new services [17]. Thus there are strong motivators to introduce "domains of transparency" wherein all-optical networking equipment would transport data unfettered by these drawbacks.

Thus, the issue of IP over optical networking in all optical sub-networks, and sub-networks with limited wavelength conversion capability merits special attention. In such networks, transmission impairments resulting from the peculiar characteristics of optical communications complicate the process of path selection. These transmission impairments include loss, noise (due primarily to amplifier spontaneous emission -- ASE), dispersion (chromatic dispersion and polarization mode dispersion), cross-talk, and non-linear effects. In such networks, the feasibility of a path between two nodes is no longer simply a function of topology and resource availability but will also depend on the accumulation of impairments along the path. If the impairment accumulation is excessive, the optical signal to noise ratio (OSNR) and hence the electrical bit error rate (BER) at the destination node may exceed prescribed thresholds, making the resultant optical channel unusable for data communication. The challenge in the development of IP-based control plane for optical networks is to abstract these peculiar characteristics of the optical layer [17] in a generic fashion, so that they can be used for path computation.

8. Evolution Path for IP over Optical Architecture

The architectural models described in Section 5 imply a certain degree of implementation complexity. Specifically, the overlay model was described as the least complex for near term deployment and the peer model the most complex. Nevertheless, each model has certain advantages and this raises the question as to the evolution path for IP over optical network architectures.

The evolution approach recommended in this framework is the definition of capability sets that start with simpler functionality in the beginning and include more complex functionality later. In this regard, it is realistic to expect that initial IP over optical deployments will be based on the domain services model (with overlay interconnection), with no routing exchange between the IP and optical domains. Under this model, direct signaling between IP routers and optical networks is likely to be triggered by offline traffic engineering decisions. The next step in the evolution of IP-optical interaction is the introduction of reachability information exchange between the two domains. This would potentially allow lightpaths to be established as part of end-to-end LSP set-up. The final phase is the support for the full peer model with more sophisticated routing interaction between IP and optical domains.

Using a common signaling framework (based on GMPLS) from the beginning facilitates this type of evolution. In this evolution, the signaling capability and semantics at the IP-optical boundary would become more sophisticated, but the basic structure of signaling would remain. This would allow incremental developments as the interconnection model becomes more sophisticated, rather than complete re-development of signaling capabilities.

From a routing point of view, the use of Network Management Systems (NMS) for static connection management is prevalent in legacy optical networks. Going forward, it can be expected that connection routing using the control plane will be gradually introduced and integrated into operational infrastructures. The introduction of routing capabilities can be expected to occur in a phased approach.

It is likely that in the first phase, service providers will either upgrade existing local element management (EMS) software with additional control plane capabilities (and perhaps the hardware as well), or upgrade the NMS software in order to introduce some degree of automation within each optical subnetwork. For this reason, it may be desirable to partition the network into subnetworks and introduce IGP interoperability within each subnetwork (i.e., at the I-NNI level), and employ either static or signaled interoperability between subnetworks. Consequently, it can be envisioned that the first phase in the evolution towards network level control plane interoperability in IP over Optical networks will be organized around a system of optical subnetworks which are interconnected statically (or dynamically in a signaled configuration). During this phase, an overlay interconnection model will be used between the optical network itself and external IP and MPLS routers (as described in Section 5.2.3).

Progressing with this phased approach to IPO routing interoperability evolution, the next level of integration will be achieved when a single carrier provides dynamic optical routing interoperability between subnetworks and between domains. In order to become completely independent of the network switching capability within subnetworks and across domains, routing information exchange may need to be enabled at the UNI level. This would constitute a significant evolution: even if the routing instances are kept separate and independent, it would still be possible to dynamically exchange reachability and other types of routing information. Another more sophisticated step during this phase is to introduce dynamic routing at the E-NNI level. This means that any neighboring networks (independent of internal switching capability) would be capable of exchanging routing information with peers across the E-NNI.

Another alternative would be for private networks to bypass these intermediate steps and directly consider an integrated routing model from the onset. This direct evolution strategy is realistic, but is more likely to occur in operational contexts where both the IP (or MPLS) and optical networks are built simultaneously, using equipment from a single source or from multiple sources that are closely affiliated. In any case, due to the current lack of operational experience in managing this degree of control plane interaction in a heterogeneous network (these issues may exist even if the hardware and software originate from the same vendor), an augmented model is likely to be the most viable initial option. Alternatively, a very modular or hierarchical peer model may be contemplated. There may be other challenges (not just of a technical, but also administrative and even political issues) that may need to be resolved in order to achieve full a peer model at the routing level in a multi-technology and multi-vendor environment. Ultimately, the main technical improvement would likely arise from efficiencies derived from the integration of traffic-engineering capabilities in the dynamic inter-domain routing environments.

9. Security Considerations

The architectural framework described in this document requires a number of different protocol mechanisms for its realization. Specifically, the role of neighbor discovery, routing, and signaling protocols were highlighted in previous sections. The general security issues that arise with these protocols include:

- The authentication of entities exchanging information (e.g., signaling, routing, or link management) across a control interface;

- Ensuring the integrity of the information exchanged across the interface;
- Protection of the control mechanisms from intrusions and other modes of outside interference.

Because optical connections may carry high volumes of traffic and are generally quite expensive, mechanisms are required to safeguard optical networks against intrusions and unauthorized utilization of network resources.

In addition to the security aspects relating to the control plane, the data plane must also be protected from external interference.

An important consideration in optical networks is the separation of control channels from data channels. This decoupling implies that the state of the bearer channels carrying user traffic cannot be inferred from the state of the control channels. Similarly, the state of the control channels cannot be inferred from the state of the data channels. The potential security implications of this decoupling should be taken into account in the design of pertinent control protocols and in the operation of IPO networks.

Another issue in IPO networks concerns the fact that the underlying optical network elements may be invisible to IP client nodes, especially in the overlay model. This means that traditional IP tools such as traceroute cannot be used by client IP nodes to detect attacks within the optical domain.

For the aforementioned reasons, the output of the routing protocol security (RPSEC) efforts within the IETF should be considered in the design of control protocols for optical networks.

In Section 2, the concept of a trust domain was defined as a network under a single technical administration in which adequate security measures are established to prevent unauthorized intrusion from outside the domain. It should be strongly noted that within a trust domain, any subverted node can send control messages which can compromise the entire network.

9.1. General security aspects

Communication protocols usually require two main security mechanisms: authentication and confidentiality. Authentication mechanisms ensure data origin verification and message integrity so that intrusions and unauthorized operations can be detected and mitigated. For example, with reference to Figure 1, message authentication can prevent a malicious IP client from mounting a denial of service attack against

the optical network by invoking an excessive number of connection creation requests across the UNI interface. Another important security consideration is the need to reject replayed control packets. This capability can assist in countering some forms of denial of service attacks. Replay protection provides a form of partial sequence integrity, and can be implemented in conjunction with an authentication mechanism.

Confidentiality of signaling messages is also desirable, especially in scenarios where message attributes between communicating entities include sensitive or private information. Examples of such attributes include account numbers, contract identification information, and similar types of private data.

The case of equipment that are not co-located presents increased security threats. In such scenarios, the communicating entities engaged in protocol message transactions may be connected over an external network. Generally, the external network may be outside the span of control of the optical network (or client IP network) administrators. As a result, the protocol messages may be subject to increased security threats, such as address spoofing, eavesdropping, and intrusion. To mitigate such threats, appropriate security mechanisms must be employed to protect the control channels and associated signaling and routing messages.

Requests for optical connections from client networks must also be filtered using appropriate policies to protect against security infringements and excess resource consumption. Additionally, there may be a need for confidentiality of SRLGs in some circumstances.

Optical networks may also be subject to subtle forms of denial of service attacks. An example of this would be requests for optical connections with explicit routes that induce a high degree of blocking for subsequent requests. This aspect might require some global coordination of resource allocation.

Another related form of subtle denial of service attack could occur when improbable optical paths are requested (i.e., paths within the network for which resources are insufficiently provisioned). Such requests for improbable paths may consume ports on optical switching elements within the network resulting in denial of service for subsequent connection requests.

9.2. Security Considerations for Protocol Mechanisms

The security requirements for IP-centric control protocols employed in the control plane of optical networks would depend on the specific characteristics of the protocols and the security risks that exist in a particular operational context. Such details relating to particular operational contexts are beyond the scope of this document and hence are not considered further. Nevertheless, it must be stated that such control protocols must take into account the issues associated with the separation of control channels from data channels in switched optical networks, and the magnitude and extent of service interruptions within the IP domain that could result from outages emanating from the optical domain.

10. Summary and Conclusions

The objective of this document was to define a framework for IP over optical networks, considering the service models, and routing and signaling issues. There are a diversity of choices for IP-optical control interconnection, service models, and protocol mechanisms. The approach advocated in this document was to support different service models which allow for future enhancements, and define complementary signaling and routing mechanisms to enable these capabilities. An evolutionary scenario, based on a common signaling framework (e.g., based on GMPLS) was suggested, with the capability to increase the complexity of interworking functionality as the requirements become more sophisticated. A key aspect of this evolutionary principle is that the IP-optical control and service interaction is first based on the domain services model with overlay interconnection that will eventually evolve to support full peer interaction.

11. Informative References

[1] Awduche, D. and Y. Rekhter, "Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control With Optical Crossconnects", IEEE Communications Magazine, March 2001.

[2] Lang, J., et al., "Link Management Protocol", Work in progress.

[3] Kompella, K. and Y. Rekhter, "LSP Hierarchy with MPLS TE", Internet Draft, Work in progress.

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

[5] Rajagopalan, B., "Documentation of IANA Assignments for Label Distribution Protocol (LDP), Resource ReSeVation Protocol (RSVP), and Resource ReSeVation Protocol-Traffic Engineering (RSVP-TE) Extensions for Optical UNI Signaling", RFC 3476, March 2003.

[6] The Optical Interworking Forum, "UNI 1.0 Signaling Specification", December 2001.

[7] Kompella, K., et al., "OSPF Extensions in Support of Generalized MPLS," Work in Progress.

[8] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP4)", RFC 1771, March 1995.

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

[10] Mannie, E., "GMPLS Extensions for SONET/SDH Control", Work in Progress.

[11] Doshi, B., Dravida, S., Harshavardhana, P., et. al, "Optical Network Design and Restoration," Bell Labs Technical Journal, Jan-March, 1999.

[12] Kompella, K., et al., "Link Bundling in MPLS Traffic Engineering", Work in Progress.

[13] Ramamurthy, S., Bogdanowicz, Z., Samieian, S., et al., "Capacity Performance of Dynamic Provisioning in Optical Networks", Journal of Lightwave Technology, January 2001.

[14] Crawley, E., Nair, R., Rajagopalan, B. and H. Sandick, "A Framework for QoS-based Routing in the Internet", RFC 2386, August 1998.

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

[16] Suurballe, J., "Disjoint Paths in a Network", Networks, vol. 4, 1974.

[17] Chiu, A., et al., "Impairments and Other Constraints On Optical Layer Routing", Work in Progress.

12. Acknowledgments

We would like to thank Zouheir Mansourati (Movaz Networks), Ian Duncan (Nortel Networks), Dimitri Papadimitriou (Alcatel), and Dimitrios Pendarakis (Tellium) for their contributions to this document. The Security Considerations section was revised to reflect input from Scott Bradner and Steve Bellovin.

13. Contributors

Contributors are listed alphabetically.

Brad Cain
Cereva Networks
3 Network Dr.
Marlborough, MA 01752

EMail: bcain@cereva.com

Bilel Jamoussi
Nortel Networks
600 Tech Park
Billerica, MA 01821

Phone: 978-288-4734
EMail: jamoussi@nortelnetworks.com

Debanjan Saha

EMail: debanjan@acm.org

14. Authors' Addresses

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

EMail: braja@tellium.com

James V. Luciani
Marconi Communications
2000 Marconi Dr.
Warrendale, PA 15086

EMail: james_luciani@mindspring.com

Daniel O. Awduche
MCI
22001 Loudoun County Parkway
Ashburn, VA 20147

Phone: 703-886-1753
EMail: awduche@awduche.com

15. Full Copyright Statement

Intellectual Property

Acknowledgement

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