Increasing the Link Utilization in IP over WDM Networks

Antonio Nucci[a,b], Nina Taft[a], Patrick Thiran[c], Hui Zang[a] and Christophe Diot[a]

[a] Sprint Advanced Technology Labs, Burlingame, CA 94010, USA
[b] Dipartimento di Elettronica - Politecnico di Torino - Torino - Italy
[c] ICA-DSC but LCA-ISC-I & C, EPFL, CH-1015 Lausanne, Switzerland

ABSTRACT

In this paper we study an approach to Quality of Service that offers end-users the choice between two classes of service defined according to their level of transmission protection. The first class of service, called *Fully Protected* (FP), offers end-users a guarantee of survivability in the case of a single failure; all FP traffic is protected using either a 1:1 or 1+1 protection scheme at the WDM layer. The second class of service, called *Best-Effort Protected* (BEP), is not protected; when a failure occurs, the network does the best it can by restoring at the IP layer only as much BEP traffic as possible. The FP service class mimics what Internet users receive today. The motivation of this approach is to increase the amount of bandwidth used on backbone networks by offering a lower quality of service that does not affect the current QoS provided by the network.

We design an ILP model, for finding primary and backup paths at the optical layer, that incorporates a number of carriers' common practices. Namely we allow the FP demand to be specified via a traffic matrix at the IP layer, we include an overprovisioning factor that specifies the portion of each link that must be left unused, and we incorporate a minimal fairness requirement on how the BEP traffic is allocated. Our goal is thus to quantify how much BEP traffic can be carried in addition to the FP traffic, without impacting the protection quality of the FP traffic even in the case of failure, and without impacting the FP load.

We show that by having two such classes of service, the load on a network can be increased by a factor of 4 to 7 (depending upon the network). Even if carriers want to overprovision their networks by 50%, we can still triple the total network load. We illustrate that the location of the bottleneck can affect whether or not we see a difference in performance between 1:1 or 1+1 protection schemes. Finally we evaluate the tradeoff between the two carrier requirements of overprovisioning and minimal fairness.

1. INTRODUCTION

Today's Internet backbone contains a large amount of unused capacity largely due to *redundancy of equipment* and *overprovisioning*. Redundancy of equipment is necessary to protect the backbone against failures. Single link failures are more common than one might expect; in the backbone of a large international Tier-1 carrier, about one link fails per week. Schroeder et al[1] show that multiple link failures are fortunately more rare, occurring roughly 2 or 3 times per year. With technologies such as WDM, a single fiber failure can bring down a large number of IP paths. Overprovisioning is the current solution adopted by carriers towards providing quality-of-service in the Internet. With overprovisioned networks, the delays and losses in the Internet are very small, as proved in Papagiannaki et al.[2] The simplicity of overprovisioning makes it cheaper in IP backbones than the alternative solutions of reservation-based and priority services, which can also offer users a low-delay, low-loss service. A side effect of overprovisioned backbones is that the service they offer is very good. Although the Internet service is called *best-effort*, ([2,3]) show that the service received by most far exceeds best-effort since there is essentially no loss and end-to-end delays are close to the speed of light.

Further author information: (Send correspondence to A.N.)
A.N. : E-mail: anucci@sprintlabs.com
N.T. : E-mail: nina@sprintlabs.com
P.T. : E-mail: Patrick.Thiran@epfl.ch
H.Z. : E-mail: hzang@sprintlabs.com
C.D. : E-mail: cdiot@sprintlabs.com

It is not mandatory for ISPs to provide this high quality of service to all its customers. In regular operation, over-provisioned backbone networks offer all users the same delay-free, loss-free service. We were thus motivated to propose two classes of service, one of which would mimic today's service and a second one that provide a lower quality of service. We propose to differentiate the two classes of service by their *level of protection* or resilience. The first class, called *Fully Protected* (FP), offers users the insurance that they will not suffer service interruption in the case of a single failure. Protection is provided at the WDM layer via either a 1+1 or 1:1 protection scheme that guarantees fast recovery after a single failure. This service class represents what an Packet-over-SONET backbone provides today. The second class of service, called *Best Effort Protected* (BEP)is new. It does not provide a specific guarantee on service disruption. Instead, in the case of failure, it offers to restore as much of the affected traffic as possible. This restoration is taken care of at the IP layer, via IGP protocols such as OSPF or IS-IS which have their own mechanisms for detecting failures and computing new paths. Protection at the WDM layer is very fast (e.g., 50 ms in SONET) because precomputed backup paths are used, while restoration at the IP layer is slower since new alternate paths are computed after the failure has been detected.

The idea behind two such services is for the lower-grade service to make use of the excess bandwidth in the backbone in such a way that has no impact on the Service Level Agreements (SLAs) promised to the higher-grade service. The majority of the time this excess bandwidth is unused; when it does become needed in a failure scenario, as much BEP traffic as necessary would be pre-empted, i.e., dropped to carry the FP traffic. By introducing a second service we enable an alternate vehicle for generating revenue.

We first proposed two such classes of service in Thiran.[4] The previous paper considered a small six node network, with 1+1 protection, and served as a proof of concept that with two such services, the load in a network can be significantly increased without impacting existing FP traffic. In this paper we develop a complete model that finds primary and backup paths for the FP traffic and that routes as much BEP traffic as possible on the network. Our new model incorporates a number of features not found in the previous paper: (1) it includes both 1+1 and 1:1 protection schemes, (2) it allows FP demands to be heterogeneous and specified via a traffic matrix at the IP layer; (3) it takes into account the constraint of mapping one IP layer logical link onto at most one lightpath, (4) it incorporates an overprovisioning factor, denoted β_{FREE}, that represents an operational requirement to leave a given fraction of each logical link unused; and (5) it incorporates a Z_{min} factor that represents the minimum amount of BEP traffic offered to each logical connection. We allow the BEP traffic to be routed on either the primary or the backup paths, but it cannot be split between them. In this paper we do not study the restoration problem at IP layer, instead we leave the BEP traffic unprotected at WDM layer.

We briefly comment on two of these key elements, the over-provisioning factor, and the minimum BEP offered to each logical connection. (These are motivated and discussed at length in Section 3.) As part of the overprovisioning practice, it is common for carrier's to require that a fixed percentage of each link remain unused. We incorporate this practice into our model to study the impact of such a policy. Clearly the more we over-provision, the less BEP traffic we can carry. We will quantify this relationship in our test cases. When we calculate how much BEP each logical connection can carry, i.e., how much we offer it, we must be careful as to how we distribute the unused capacity among the logical connections. If we simply try to add as much BEP as possible, network-wide, by using an optimization model that tries to maximize a global network load, it wouldn't be surprising if the distribution of the amounts of BEP offered to each logical connection were very inequitable. In order to introduce a minimal amount of fairness in how the BEP bandwidth is distributed we include the Z_{min} parameter in our model. This parameter requires that each logical connection be offered at least an amount Z_{min} of BEP traffic load.

The goal of this paper is to quantify how much BEP traffic can be carried on the network without impacting the FP service. We study two versions of a medium-sized network that contain different amounts of heterogeneity in the links. The gain of using our FP/BEP approach is that the total network load can be increased by a factor that varies between 4 and 7, depending upon the network. Even when the over-provisioning factor is at 50%, we still see a tripling of the total load that can be carried in a network if the FP/BEP approach is adopted. This shows that the potential of using a BEP traffic class to generate additional revenue is huge, and that this potential can be achieved without any impact to those desiring a high-grade protection service. We discuss the issue of whether bottlenecks lie in the WDM layer or the IP layer. We find that when the main bottlenecks lie in the WDM layer, we see marked differences in the performance of the 1:1 and 1+1 protection schemes. However,

when all the bottlenecks are at the IP layer, there is no difference in performance between the two protection schemes. We demonstrate that the BEP traffic that can be carried decreases linearly as the over-provisioning factor grows. Our performance tests also illustrate the tradeoff between the overprovisioning requirement and the minimal fairness requirement; namely, that if the overprovisioning requirement is too large, it may not be possible to satisfy the minimal fairness requirement.

Recently the problem of service management has gained a lot of attention in the optical community ([5–8]). Proposals for different service classes in optical networks are introduced in Gerstel and Ramaswami.[5] Ramamurthy and Mukherjee[6] study the traditional 1+1 and 1:1 protection strategies at the WDM layer for a single class of traffic. They formulate the corresponding ILP optimization problem applicable to small networks. Mohan and Somani[7] propose a class of service that offers a minimal level of protection to every connection. They claim that if the demands are highly dynamic, it is possible to select routes whose (shared) back-up paths have a specified maximal non-zero probability of being unavailable if a failure occurs. The difficulty here is to provide a tight upper bound on this probability, and to select lightpaths (i.e. logical connections) that have this specified degree of protection in a fluctuating environment. Sridharan and Somani[8] formulate the ILP problem when three different service classes co-exist. They try to minimize the capacity requested by all working and backup paths, weighted by the traffic class to which it belongs (since each class brings in a different amount of revenue). Ramamurthy and Mukherjee[6] proof that the general problem is NP-complete for a single class of traffic. Hence the recent proposal for three classes of traffic at the WDM layer may be too complex to apply to real networks.

The remainder of this paper is organized as follows. The FP and BEP classes of service are fully defined in Section 2 which also includes a brief summary of protection and restoration strategies. In Section 3 we explain which components of the overall problem belong to which layer (physical or logical), we describe our traffic matrices and give a formal problem statement. The ILP formulation for finding routes, in IP/WDM networks supporting FP/BEP traffic classes, is given in Section 4. Numerical results are presented and discussed in Section 5. Section 6 concludes the paper.

2. DEFINITION AND PROVISIONING OF CLASSES OF SERVICE

In this section, we fully specify the two classes of service introduced earlier. Following a slightly different taxonomy than those in Gerstel et al,[5] we can categorize them in the following sub-classes. For referral purposes, we put an abbreviation of each subclass.

The **Fully Protected (FP)** service guarantees its customers that their traffic is protected against any single point of failure in the backbone. FP traffic is protected via pre-computed, dedicated backup paths at the WDM layer, using either a 1:1 or 1+1 protection strategy. Failures are transparent to the IP layer for this class of traffic. It can be subdivided in two subclasses:

1. *Full 1+1 protection* (`FP/1+1`). In this scheme, FP traffic is transmitted simultaneously on two disjoint paths. The receiver selects the signal at the destination that has the better signal quality. If that path is cut, the receiver automatically switches to the other path to receive input. This is the fastest and simplest protection, because no signalling is needed. It is however very inefficient in terms of resources, as every unit of traffic is transmitted twice.

2. *Full 1:1 protection* (`FP/1:1`). In this scheme, FP traffic is transmitted only on one path (called the *working* or *primary* path). If this path fails, the sender and receiver both switch to the other path (called the *backup* path). This is not as fast nor as simple as 1+1 protection, because the node at the end of the failed link must detect the failure first and then signal it to the source, who would then switch over to the backup path. Our idea is to take advantage of 1:1 protection because the reserved but unused capacity on the backup path can be given to unprotected traffic that would be pre-empted in case of a failure.

The **Best Effort Protected (BEP)** service does not offer a guarantee of traffic survivability in case of failure. For BEP traffic we offer restoration and not protection; in other words, BEP traffic is entirely unprotected at the WDM layer, and instead we rely on the IP layer to carry out restoration. When a failure occurs, BEP packets may be dropped at the router before the point of congestion, until IP has been able to

restore this traffic by rerouting it on an alternate IP path. As mentioned in the previous section, failures are detected at the IP layer via IGP routing protocols such as OSPF or IS-IS. Thus the restoration offered to BEP is a slow one, in contrast to that offered to FP. The BEP traffic of each logical connection can be routed on either the primary or the backup path, but not on both (i.e., it cannot be split over two paths). The BEP class can be subdivided in three subclasses:

1. *No protection* (`BEP/UP`). This is the simplest of all schemes, since no backup path needs to be provided, nor special capacity be available to ensure a partial restoration at the IP layer. If a failure occurs and interrupts a lightpath with unprotected traffic, users may experience a total disruption of their service. On the other hand, if the lightpath is not interrupted by the failure, the connection is maintained.

2. *Pre-emption* (`BEP/PE`). Here traffic can also be pre-empted in case of a failure interrupting another lightpath. This traffic is the one flowing on the backup path of 1:1 or 1:n protected lightpaths, it can thus only be used in conjunction with `FP/1:1` traffic but not with `FP/1+1`. It allows to save bandwidth, but at the expense of more complexity in signalling the failure and pre-empting traffic.

3. *Partial Restoration* (`BEP/R`). In this scheme, spare capacity is left at the IP layer to restore some (unspecified) amount of the BEP traffic from the broken lightpath. No backup path is pre-computed at the WDM layer, and the (partial) restoration of this BEP traffic at the IP layer is much slower than any of the protection offered at the WDM layer for the FP class. However, a total disruption is avoided.

In this paper, we investigate the combination of `FP/1+1` traffic with either `BEP/R` or `BEP/UP` traffic, as well as `FP/1:1` traffic with all three subclasses of `BEP` traffic. We will see how the overprovisioning factor for $\beta_{free} > 0$ helps to ensure that some capacity is left at the IP layer for `BEP/R` traffic. Let us mention that other definitions of Best-Effort Protection exist such as Mohan and Somani[7] as discussed in Section 1. We point out that in an environment in which each logical connection is protected via either a 1:1 or 1+1 scheme at the WDM layer, and in which failures happen one at a time, the logical topology will always be connected. Thus the logical topology will be always able to apply a restoration strategy at the IP layer, and does not suffer from the *failure propagation* problem described in Crochat et al.[9–11]

In order to implement two such classes of service, packets need to be marked according to their class of service, and IP routers must implement class-based scheduling. In normal operation, differentiation is not needed between the two types of packets. However, upon notification of a failure, FP packets continue to be served as before, while BEP packets may be dropped until BEP traffic has been restored at the IP layer.

3. PROBLEM STATEMENT

The problem we address is to quantify how much BEP traffic can be carried in addition to FP traffic, while considering overprovisioning and fairness requirements, and without impacting the protection quality of the FP traffic even in the case of a single failure. To clarify which components of the problem are related to the logical (IP) layer and which are part of the physical (WDM) layer, we now discuss the elements of each layer and comment on the relationship between these elements. To be clear, we state some definitions of basic terms. We use the expression *logical link* to refer to a single link between two routers at the IP layer. We use the term *logical connection* to refer to a sequence of logical links. Each logical link corresponds to a sequence of one or more physical links interconnected via OXCs.

The FP traffic matrix and the IP routes are a part of the logical layer. In many formulations of optimization problems, one tries to maximize the total traffic. We decided to focus on maximizing the amount of BEP traffic carried while letting the FP traffic be specified by an input demand matrix. The reason for this is because capacity planning in the Internet is typically done using an IP layer traffic matrix that specifies the average amount of bandwidth that needs to flow between any two routers or POPs (Point-of-Presence) in a domain. After we choose an initial matrix, we scale the entire matrix up, in order to load the maximum amount of FP onto our network. By "scaling up" we mean that we multiply all elements in the matrix by a constant factor that is as large as possible. The limit on how much the matrix can be scaled up is defined by the maximum amount we can protect. In other words, we cannot scale the matrix any further if it means that some FP traffic could not be protected.

The IP routes are those given by either the OSPF or IS-IS protocol that operates at the IP layer. All carriers today use one of these two standardized and widely adopted protocols. Since these protocols are unlikely to undergo any fundamental changes, we assume they are a fixed component on our network environment. These protocols usually compute shortest-path routes between routers. A path specified by OSPF (or IS-IS) is thus a sequence of *logical links*. Both the FP traffic matrix and the IP routes are *inputs* to our problem. The problem we address is thus to find the routes at the physical layer for those IP logical links inside the IP routes.

The aggregated traffic on a logical link comes from the set of logical connections that share a given logical link. We determine this aggregate traffic demand for each logical link by routing the FP traffic matrix over the logical topology according to the OSPF routes. Once we have the aggregate demand for each logical link, the problem now passes to the physical layer.

At the physical layer, we have to find for each logical link, two link-disjoint physical paths, namely the *working* physical path and the *backup* physical path. By having two disjoint physical paths for each logical link, we ensure that the logical topology will always remain connected for each single physical failure at the WDM layer. In fact, our solution is robust to multiple failures as long as none of them is a *critical* failure. In this context, a *critical failure* is a multiple failure that brings down a set links such that both the working and backup path of the same logical link are interrupted.

As mentioned above, the FP traffic matrix is constructed so that the network is saturated by as much FP traffic as can be protected. No additional FP traffic can be accepted beyond the scaled up version of the FP traffix matrix. Having satisfied the demands for FP traffic, we then compute the amount of BEP traffic that can be added onto the network using the remaining capacity. Since the BEP traffic is allotted as a certain amount to each logical connection, the BEP traffic can also be described by a traffic matrix with the same rows and columns as the FP traffic matrix. The BEP traffic matrix that we search for is one that maximizes the network-wide load. We will see that it is indeed possible to add a significant amount of BEP traffic. However, the maximization of the total load will generally lead to a very unbalanced distribution of the additional BEP load. Some logical links are able to accomodate a huge additional BEP load in addition to the FP traffic they already carry, while other logical links may be full with FP traffic and hence cannot carry any BEP traffic. An operator may want to circumvent this unfairness by ensuring that the total BEP load be more equally distributed among the different logical links. We thus introduce a parameter Z_{min} that is the minimal amount of BEP traffic that needs to be offered on each logical link. A value of $Z_{min} = 0$ corresponds to the situation where no constraint of fairness is imposed on BEP traffic. We evaluate the maximum value of Z_{min} that can be reached, by starting to examine $Z_{min} = 0$ and then progressively increasing it until we cannot increase it anymore.

In the Internet today, it is common practice for carriers to require that a certain percentage of all links be left free. In other words, average link utilization levels are not supposed to exceed some threshold, say 60% or 70%, for any extended period of time. Once a link starts to repeatedly exceed the specified threshold, then plans for a link upgrade are usually put in place.This type of requirement also comes from router vendors who insist that link utilization levels shouldn't exceed about 80% or 90% otherwise routers can slow down to the point of introducing very large delays or even crash. We incorporated this practice into our model via a factor we call β_{FREE}, which represents the fraction of each logical link that a carrier desires to leave free. Hence a $\beta_{FREE} = 20$ means that 20% of a logical link is unavailable to either FP or BEP. The value of $\beta_{FREE} = 0$ represents the case where the full capacity of the logical link can be used.

This policy has an interesting side-effect. Indeed, it usually applies to average values of load, so that it is acceptable to exceed these thresholds for a limited amount of time. A positive value of $\beta_{FREE} > 0$ means that there is excess capacity available that could be used in the case of failure to reroute some BEP traffic *temporarily*, until the IP layer can find another route for the BEP traffic. For example, when a failure happens one could avoid dropping all affected BEP traffic by load balancing a fraction of it on another nearby link. This alternate link would be used only during the OSPF route convergence time. As a result, it may not be necessary to preempt *all* BEP traffic just after the occurrence of the failure and before OSPF has found new routes at the IP layer to restore this traffic.

We now give the formal problem statement, incorporating all of the elements above.

GIVEN:

i) a physical topology (which must be at least biconnected), whose nodes are optical cross connects (OXC) interconnected by optical fibers that support a limited number of wavelengths and have limited capacity.

ii) a logical topology whose nodes are IP routers interconnected by logical links. These links have a finite limit on the total amount of traffic they can carry (including both FP and BEP).

iii) an FP traffic matrix, denoted $D_{FP} = [d^{kh}(FP)] \geq 0$, that defines the FP traffic demand for each pair of routers (k, h) at the IP layer. We call these pair origin-destination (OD) pairs.

iv) The routing paths selected at the IP layer for each OD pair of routers. This set of routes is denoted by $\mathcal{R}$. These are the routes determined by OSPF and specify the path through the network of logical links. (We implemented a shortest path computation to mimic OSPF's routing decisions.)

v) an FP protection strategy at the WDM layer, either 1+1 or 1:1.

FIND

i) the primary and backup paths for each logical link in such a way that the network is able to carry all the demand specified in the FP traffic matrix D_{FP}.

ii) the amount of BEP that can be added to each logical connection so as to maximize the total network load without impacting the protection of the FP traffic. This output is specified in the form of a BEP traffic matrix $D_{BEP} = [d^{kh}(BEP)] \geq 0$.

4. PROBLEM FORMULATION

We formulate the problem as an Integer Linear Program (ILP) whose objective is to maximize the total load carried by the network and then maximize the total BEP carried, which we denote by $\mathcal{F}$. We first introduce the mathematical model for the 1+1 FP protection strategy, and then describe the modifications needed to extend the model to the 1:1 FP protection strategy.

4.1. Notation

We adopt the notation for multi-layered networks presented in Plante at al.[12] The supra-index indicates the layer, starting by the lowest layer, zero, that represents the physical network. Let $G^0 = (V, E^0)$ be a directed graph representing the physical topology. It is composed by OXC nodes V interconnected by optical fibers (i, j) belonging to the set E^0. We assume that OXC's are either OEO or have full wavelength conversion. Let $|V| = N$ be the cardinality of set V and $|E^0| = M$ that of set E^0. We assume that each fiber $(i, j) \in E^0$ is described by the set of parameters $\{n_{ij}, c^f_{ij}\}$, where n_{ij} represent the number of parallel WDM channels, each of which has a bandwidth equal to c^f_{ij}. Let $G^1 = (U, E^1)$ be an unidirected graph representing the logical topology. It is composed of IP routers U interconnected by lightpaths (s, t) belonging to the set E^1. Let $|U| = K$ be the cardinality of set U and $|E^1| = H$ that of set E^1. Let c^l_{st} be the capacity associated with the logical link $(s, t) \in E^1$, and be bounded by the electronic speed limit of each IP router interface. Let $\mathcal{C} = \{(k, h)\}$ be the set of all the logical connections (k, h). Let $D_{FP} = [d^{kh}(FP)] \geq 0$ be the FP traffic matrix, where each entry $d^{kh}(FP) \geq 0$ describes the FP traffic associated with the logical connection $(k, h) \in \mathcal{C}$. Let $\mathcal{R} = [r^{kh}_{st}]$ be the set of routes for each logical connection $(k, h) \in \mathcal{C}$ at the IP layer; $r^{kh}_{st} = 1$ if the logical connection (k, h) uses the logical link $(s, t) \in E^1$ in its own IP path, and 0 otherwise. For clarity of notation, in the following ILP formulation, the aggregated FP traffic on each logical link $(s, t) \in E^1$ will be described by $f^{st}(FP) = \sum_{(k,h)\in\mathcal{C}} r^{kh}_{st} d^{kh}(FP)$. We define Z_{min} to be the minimum amount of BEP traffic that each logical connection should carry.

4.2. Decision Variables

For uniformity of notation, let $D_{BEP} = [d^{kh}(BEP)] \geq 0$ be the BEP traffic matrix where each entry $d^{kh}(BEP)$ represents the BEP traffic exchanged between the IP layer OD pair (k,h). Then the aggregated BEP traffic on each logical link (s,t) will be described by the variables $f^{st}(BEP) = \sum_{(k,h)\in\mathcal{C}} r_{st}^{kh} d^{kh}(BEP)$. For each logical link (s,t) two disjoint physical paths are required. Let w_{ij}^{st} and b_{ij}^{st} be two binary variables used to describe the routing for the logical link (s,t) over the physical topology G^0, respectively for the working and the backup physical paths. The variables w_{ij}^{st} (b_{ij}^{st}) are equal to 1 if the physical link $(i,j) \in E^0$ is crossed by the working (backup) physical path associated with the logical link $(s,t) \in E^1$ and 0 otherwise. Let $\alpha_{ij}^{st}(BEP) \geq 0$ and $\beta_{ij}^{st}(BEP) \geq 0$ be the BEP traffic flow variables associated respectively with the working and backup physical paths for the logical link $(s,t) \in E^1$. To define which physical path is used to carry the BEP traffic (working or backup) we need to add to the model the new variables ϵ^{st} that are equal to 1 if the BEP traffic associated to the logical link $(s,t) \in E^1$ is physically sent on the working path and 0 if it is sent on the backup path.

4.3. Constraints

We now specify the relations among all the variables previously defined.

- The minimum amount of BEP traffic exchanged between any OD pairs is:

$$d^{kh}(BEP) \geq Z_{min} \qquad \forall (k,h) \in \mathcal{C} \tag{1}$$

Equation (1) ensures the minimal fairness requirement because this constraint forces each logical connection to get at least an amount Z_{min} of bandwidth for its BEP traffic.

- The aggregated BEP traffic on each logical link (s,t) is:

$$f^{st}(BEP) = \sum_{j\in V:(s,j)\in E^0} (\alpha_{sj}^{st}(BEP) + \beta_{sj}^{st}(BEP)) \qquad \forall (s,t) \in E^1 \tag{2}$$

Relation (2) ensures that the aggregated BEP traffic associated with each logical link (s,t) leaving node s can only traverse either the backup or working path at the WDM for that logical link. This general equation allows the traffic to be sent on *either* the working ($\alpha_{sj}^{st}(BEP)$) or the backup ($\beta_{sj}^{st}(BEP)$) path. Later we add a constraint that enables one of the two to be selected, but not both (because we do not support traffic splitting).

- The flow continuity constraint for the physical working path associated with logical link (s,t) is:

$$\sum_{j\in V:(i,j)\in E^0} w_{ij}^{st} - \sum_{j\in V:(j,i)\in E^0} w_{ji}^{st} = \begin{cases} 1 & \text{if} \quad i = s \\ -1 & \text{if} \quad i = t \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in V, \forall (s,t) \in E^1 \tag{3}$$

- The flow continuity constraint for the physical backup path of logical link (s,t) is:

$$\sum_{j\in V:(i,j)\in E^0} b_{ij}^{st} - \sum_{j\in V:(j,i)\in E^0} b_{ji}^{st} = \begin{cases} 1 & \text{if} \quad i = s \\ -1 & \text{if} \quad i = t \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in V, \forall (s,t) \in E^1 \tag{4}$$

Equations (3) and (4) define the two physical paths associated with each logical link.

- We force the working and backup physical paths to be disjoint via:

$$w_{ij}^{st} + w_{ji}^{st} + b_{ij}^{st} + b_{ji}^{st} \leq 1 \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{5}$$

- The BEP traffic on the backup path of logical link (s,t) is constrained by:

$$\beta_{ij}^{st}(BEP) \leq B b_{ij}^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{6}$$

where B is a large number chosen so that it is larger than $\alpha_{ij}^{st}(BEP)$ for any $(i,j) \in E^0$ and any $(s,t) \in E^1$. For example, we can take $B = \max_{(i,j)\in E^0, i<j} \left\{ c_{ij}^f \right\}$. Relation (6) forces $\alpha_{ij}^{st}(BEP)$ to be equal to 0 if $w_{ij}^{st} = 0$,

that is, if the working path selected for the logical link (s,t) does not cross over fiber (i,j). On the other hand, if $w_{ij}^{st}=1$, then relations (6) do not impose any restriction on $\alpha_{ij}^{st}(BEP)$. The same number B will be used in the realtions (7), (8) and (9).

• The following two equations determine which of the two physical paths will carry the BEP traffic. Note that only one path at a time is allowed to carry BEP.

$$\alpha_{ij}^{st}(BEP) \leq B\epsilon^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{7}$$

$$\beta_{ij}^{st}(BEP) \leq B(1-\epsilon^{st}) \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{8}$$

• The flow continuity constraints for the BEP class of service carried by logical link (s,t) on the selected physical path (working or backup) are:

$$\sum_{j \in V:(i,j)\in E^0} (\alpha_{ij}^{st}(BEP)+\beta_{ij}^{st}(BEP)) - \sum_{j \in V:(j,i)\in E^0} (\alpha_{ji}^{st}(BEP)+\beta_{ji}^{st}(BEP)) = \begin{cases} f^{st}(BEP) & \text{if} \quad i=s \\ -f^{st}(BEP) & \text{if} \quad i=t \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in V, \forall (s,t) \in E^1 \tag{9}$$

• The maximum number of wavelengths on each physical fiber is constrained by:

$$\sum_{(s,t)\in E^1} (w_{ij}^{st}+w_{ji}^{st}+b_{ij}^{st}+b_{ji}^{st}) \leq n_{ij} \qquad \forall (i,j) \in E^0 : i<j \tag{10}$$

Equation (10) ensures that the number of logical channels traversing each fiber can not be bigger than the number of available wavelengths.

• We have the following capacity constraints from the physical layer:

$$f^{st}(FP)(w_{ij}^{st}+w_{ji}^{st}) + (\alpha_{ij}^{st}(BEP)+\alpha_{ji}^{st}(BEP)) \leq c_{ij}^{f} \qquad \forall (i,j) \in E^0 : i<j, \forall (s,t) \in E^1 \tag{11}$$

$$f^{st}(FP)(b_{ij}^{st}+b_{ji}^{st}) + (\beta_{ij}^{st}(BEP)+\beta_{ji}^{st}(BEP)) \leq c_{ij}^{f} \qquad \forall (i,j) \in E^0 : i<j, \forall (s,t) \in E^1 \tag{12}$$

Equations (11) and (12) ensure that each logical link (s,t) does not carry more traffic than the bandwidth of the fibers in the working and backup paths, respectively.

• We also have capacity constraints from the logical layer:

$$f^{st}(FP)+f^{st}(BEP) \leq (1-\beta_{FREE})c_{st}^{l} \qquad \forall (i,j) \in E^0 : i<j, \forall (s,t) \in E^1 \tag{13}$$

Relation (13) ensures that the total amount of traffic carried by each logical channel is no bigger than its own electronic speed interface. This bound on the logical link's traffic rate is limited by the over-provisioning factor as well.

The objective function for our problem is to maximize:

$$\mathcal{F} = \sum_{(k,h)\in\mathcal{C}} d^{kh}(BEP) \tag{14}$$

4.4. Changes for 1:1 FP protection strategy

With a 1:1 protection strategy, the bandwidth on a working path carrying FP traffic must be protected with its own backup path. As long as no failure occurs in the network, the reserved bandwidth on backup paths can be used to carry BEP traffic. The physical capacity constraints (12) must be replaced by the following two constraints:

- The physical capacity constraints for 1:1 are:

$$f^{st}(FP)(b_{ij}^{st} + b_{ji}^{st}) \leq c_{ij}^{f} \qquad \forall (i,j) \in E^0 : i < j, \forall (s,t) \in E^1 \tag{15}$$

$$\beta_{ij}^{st}(BEP) + \beta_{ji}^{st}(BEP) \leq c_{ij}^{f} \qquad \forall (i,j) \in E^0 : i < j, \forall (s,t) \in E^1 \tag{16}$$

Relation (15) ensures that there is enough capacity on the backup path to protect the FP traffic flowing on the working path, and relation (16) ensures that the maximum amount of BEP traffic sent on the backup path cannot exceed the capacity of the WDM channel used in the corresponding fiber.

4.5. Basic Performance of the ILP model

We now present our first set of results using our ILP model. We use this initial test case as an example to illustrate all of the inputs and outputs of the model. Figure 1 shows a network consisting of 10 nodes (circle symbols) and 12 links at physical layer, and 6 nodes (rectangular boxes) and 9 links at the logical layer. This network represents the Italian backbone with heterogeneous links. Some of the physical links are OC-12 links while others are OC-48. When we speak of an OC-12 physical link we assume the link has 8 WDM channels each of which transmists at OC-12 rates (622 Mbps). Similarly, we assume an OC-48 link consists of 16 WDM channels at 2.448 Gbps each. We assume that the line cards in the routers are at OC-48 speeds.

Two assumptions are made: i) each logical link is bidirectional and ii) all the logical connections (k,h) and (h,k) use the same multihop path defined by OSPF (i.e. the sequence of logical links). The first one implies that each logical link carries the aggregated IP traffic in both the directions. The second one implies a triangular traffic matrix (see the bottom-right of Figure 1 for the FP traffic matrix), i.e. $d^{kh}(FP)$ is the sum of the aggregated FP traffic from k to h and from h to k.

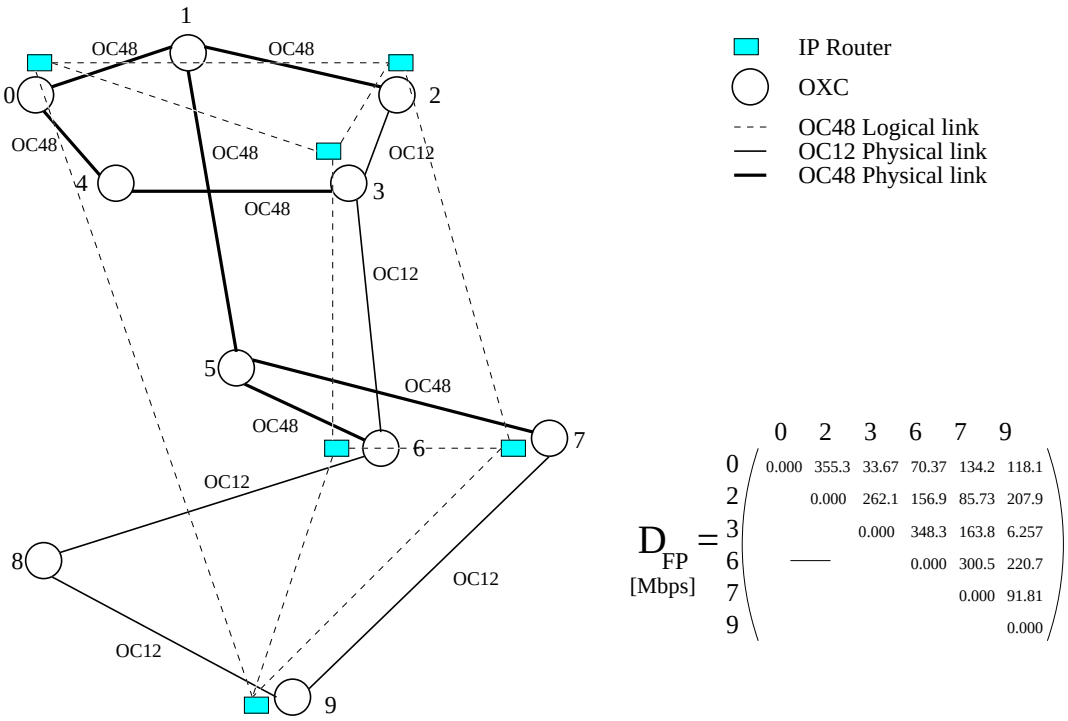
Figure 2 shows the optimal solution for 1:1 (on the top) and 1+1 (on the bottom) FP protection strategy obtained solving the mathematical model using a Branch and Bound technique running ILOG CPLEX optimizer[13] over a 800 MHz Pentium III PC running Linux 6.2. The solution shows the working and backup paths selected for each logical link, the path choosen for the BEP traffic ([p]=w: working, [p]=b: backup), and the BEP traffic matrix (D_{BEP}). Regardless of the FP protection strategy used at WDM layer, the gain realized, in terms of network load, by having two classes of service instead of only one is huge: 6.6 times bigger with 1:1 protection and 6 times with 1+1 protection.

This example illustrates how a bottleneck impacts the limitations on the scaling of the FP traffic matrix. In this example, we generated each element of the traffic matrix according to a uniform distribution between 1 and 50, and then scaled the matrix up as much as possible while still assuring all FP is protected. The amount of FP traffic that can be carried is limited because the router at node #9 has only OC12 interfaces. This bottleneck limits the amount of traffic that can be exchanged on logical links (9,0), (9,6) and (9,7). Since these cannot be scaled up by more than what's indicated in the traffic matrix, the rest of the FP matrix cannot be further scaled either.

Another important observation is the existence of a lot of zero entries in the BEP traffic matrix. This means there are a lot of logical connections that do not carry any BEP traffic. As indicated in Section 3, we expected this to happen and thus introduced the Z_{min} factor to circumvent this outcome.

5. NUMERICAL RESULTS

In this section we evaluate the performance of our FP/BEP approach on medium and large sized networks. We use our Italian backbone network from Fig. 1 to represent a medium sized network. We consider two versions of this network, the first is the same as before with five OC-12 and seven OC-48 links, and the second has four OC-12 links with eight OC-48 links. (We assume that physical link $(7,9)$ has been upgraded to an OC-48 system). For these medium-sized networks, all the results were obtained by solving our ILP model.



$$D_{FP}\ [\text{Mbps}] = \begin{array}{c} \\ 0 \\ 2 \\ 3 \\ 6 \\ 7 \\ 9 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 2 & 3 & 6 & 7 & 9 \end{array} \\ \begin{pmatrix} 0.000 & 355.3 & 33.67 & 70.37 & 134.2 & 118.1 \\ & 0.000 & 262.1 & 156.9 & 85.73 & 207.9 \\ & & 0.000 & 348.3 & 163.8 & 6.257 \\ & -- & & 0.000 & 300.5 & 220.7 \\ & & & & 0.000 & 91.81 \\ & & & & & 0.000 \end{pmatrix} \end{array}$$

Logical Connections	OSPF − IP Routing	Logical Connections	OSPF − IP Routing
[0,2]	(0,2)	[2,9]	(2,7)−(7,9)
[0,3]	(0,3)	[3,6]	(3,6)
[0,6]	(0,9)−(9,6)	[3,7]	(3,6)−(6,7)
[0,7]	(0,9)−(9,7)	[3,9]	(3,6)−(6,9)
[0,9]	(0,9)	[6,7]	(6,7)
[2,3]	(2,3)	[6,9]	(6,9)
[2,6]	(2,7)−(7,6)	[7,9]	(7,9)
[2,7]	(2,7)		

Figure 1. Italian backbone: physical and logical topologies - 10 nodes 12 links at WDM layer - 6 nodes 9 links at IP layer. 7 physical links upgraded to OC48. On the right is shown the FP traffic matrix considered D_{FP}.

5.1. Methodology

We use six performance metrics to assess our approach. The first three of our performance measures are computed in a network without failures. The last three metrics are calculated over a number of scenarios in which single failures occur. Since there are 12 links in the Italian backbone, there are 12 failure scenarios. In each scenario we assume a different single link fails. Our six metrics are as follows.

1. We measure the BEP load that can be carried when there are no failures. The FP load is defined by the input traffic matrix since we carry all of it. The total network load is thus the sum of the BEP and FP loads.
2. The average and maximum utilization of the logical links when no failures occur.
3. The average and maximum utilization of the physical links when no failures occur.
4. The average amount of BEP traffic lost when a failure happens. We compute the amount of BEP lost in each failure scenario and present the mean, averaged over all the failure scenarios.
5. The average and maximum utilization of the logical links when a failure happens.

1:1 FP Protection Strategy FP load=2556 Mbps BEP load=14313 Mbps

Logical Links	Working Paths	Backup paths	[p]
(0,2)	0->4->3->2	0->1->2	b
(0,3)	0->4->3	0->1->2->3	w
(0,9)	0->1->5->7->9	0->4->3->6->8->9	b
(2,3)	2->1->0->4->3	2->3	w
(2,7)	2->1->5->7	2->3->6->8->9->7	w
(3,6)	3->4->0->1->5->6	3->6	w
(6,7)	6->8->9->7	6->5->7	b
(6,9)	6->5->7->9	6->8->9	b
(7,9)	7->5->6->8->9	7->9	b

$$D_{\substack{BEP\\[Mbps]}} = \begin{array}{c} \\ 0 \\ 2 \\ 3 \\ 6 \\ 7 \\ 9 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 2 & 3 & 6 & 7 & 9 \end{array} \\ \begin{pmatrix} 0.000 & 2093 & 2414 & 0.000 & 0.000 & 622.0 \\ & 0.000 & 2186 & 0.000 & 1997 & 0.000 \\ & & 0.000 & 1930 & 0.000 & 0.000 \\ & & & 0.000 & 1827 & 622.0 \\ & & & & 0.000 & 622.0 \\ & & & & & 0.000 \end{pmatrix} \end{array}$$

1+1 FP Protection Strategy FP load=2556 Mbps BEP load=13259 Mbps

Logical Links	Working Paths	Backup paths	[p]
(0,2)	0->4->3->2	0->1->2	b
(0,3)	0->4->3	0->1->5->6->3	w
(0,9)	0->4->3->6->8->9	0->1->5->7->9	w
(2,3)	2->1->0->4->3	2->3	w
(2,7)	2->3->6->8->9->7	2->1->5->7	b
(3,6)	3->4->0->1->5->6	3->6	w
(6,7)	6->5->7	6->8->9->7	w
(6,9)	6->5->7->9	6->8->9	w
(7,9)	7->9	7->5->6->8->9	w

$$D_{\substack{BEP\\[Mbps]}} = \begin{array}{c} \\ 0 \\ 2 \\ 3 \\ 6 \\ 7 \\ 9 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 2 & 3 & 6 & 7 & 9 \end{array} \\ \begin{pmatrix} 0.000 & 2093 & 2414 & 0.000 & 0.000 & 299.3 \\ & 0.000 & 2186 & 0.000 & 1997 & 0.000 \\ & & 0.000 & 1930 & 0.000 & 0.000 \\ & & & 0.000 & 1827 & 324.7 \\ & & & & 0.000 & 188.1 \\ & & & & & 0.000 \end{pmatrix} \end{array}$$

Figure 2. Example of an optimal solution obtained by solving the ILP, with β_{FREE}=0 and $Z_{min} = 0$

6. The average and maximum utilization of the physical links when a failure happens.

In the last two metrics, the average and maximum are taken over all failure scenarios. In all the utilization graphs, the levels indicated include both FP and BEP traffic.

Our ILP model itself does not explicitly take into account the various failures scenarios. In other words, the model does not compute routes that are optimal over all failure scenarios, but rather that are optimal in the case of no failures (i.e., the full topology). To evaluate the performance of the ILP solution under a failure scenario we do the following. All of the FP and BEP traffic is routed according to the routes computed by the ILP model in the full topology. We then consider what happens if a single link fails. All logical connections not affected by the physical fiber failure retain their original physical routing paths. If the failure affects the working or backup path of a logical link, then the FP and BEP flows are swapped between the two physical paths according to the rules described in Section 2. Recall that a physical link failure can affect multiple logical connections. We compute how much BEP traffic can be retained in each of the failure scenarios. The numerical results we present are averaged over all possible failure scenarios.

We present our six performance metrics as a function of the overprovisioning factor β_{FREE}. Recall that the β_{FREE} factor is a requirement on the *logical* link. Both 1:1 and 1+1 protection strategies were analyzed. For each test case we considered 20 different FP traffic matrices, and thus each point in the graphs below represents a value averaged over the 20 traffic matrices. Each traffic matrix is generated randomly according to a uniform distribution, where each entry is selected uniformly between 1 and 50 Mbps. Each matrix is then scaled up as much as possible.

5.2. The Bottleneck Issue

Before presenting the results, we first describe what we mean when we say the bottleneck is either at the WDM layer or the IP layer. We will see further below how the location of the bottleneck impacts the results. In order to send packets over OC-48 links, both the router and the optical cross-connect need to have the appropriate interface card. The upper limit of a *logical link* will be 2.5 Gbps (622 Mbps) if both the source and destination *routers* have OC-48 (OC-12) interface cards, respectively. The upper speed limit of a *physical connection* will be 2.5 Gbps if all the *OXC's* in both the primary and backup paths have OC-48 interface cards. If an OXC in one of those paths uses OC-12 cards, then that path will be the bottleneck. Let $IP_c(l)$ denote the capacity limit of logical link l at the IP layer. Let $WDM_c(l)$ denote the capacity limit of the primary and backup paths at the optical layer for logical link l. More precisely, $WDM_c(l)$ is the maximum of the capacities of the primary and backup path if FP traffic is protected on a 1+1 basis, and it is the sum of the capacities of the primary and backup path if FP traffic is protected on a 1:1 basis. If $WDM_c(l) < IP_c(l)$ then the bottleneck for l is at the WDM layer, otherwise it is at the IP layer. In all of our sample networks, we assume that the routers have OC-48 interface cards. The interface cards of the OXC's are indicated in each of the figures.

To illustrate this bottleneck issue, we return to the example in Figure 2. If we look at the working and backup paths enumerated here for the nine logical links, we see that 6 of the logical links have one physical layer path that is OC-48 and the other that is OC-12. For these links the total capacity at the *logical* layer is OC-48, while the total capacity available to that logical link at the *physical* layer is the sum of OC-48 plus OC-12 (for the 1:1 case which allows us to use backup bandwidth). Hence we consider the bottleneck to be at the IP layer because the logical connection will not be able to fill all of the capacity available to it at the physical layer. (Recall that we allocate one wavelength to a logical connection.) However for 3 of these links, both their working and backup paths have OC-12 rates, and hence their bottleneck is at the WDM layer.

5.3. Medium-Sized Networks

Our first set of results is given in Figure 3 for the Italian network with 7 OC-48 links and $Z_{min} = 0$. When $\beta_{FREE} = 0$, the amount of BEP that can be carried is 4.5 times the FP load with a 1+1 protection strategy, and 5 times the FP load with a 1:1 protection strategy. Thus the total load increases by a factor of 5.5 under 1+1, and by 6 under 1:1. This demonstrates that the potential to increase the traffic carried on todays' networks is huge. Even with $\beta_{FREE} = 0.5$, the amount of BEP is 7.5 Gbps for 1:1, which still allows for a tripling (3 FP and 7.5 BEP) of the regular load (FP only). It is intuitive that the additional BEP load carriable decreases linearly as β_{FREE} increases, since β_{FREE} is defined at the logical layer. We see that more BEP can be carried under a 1:1 protection strategy than under 1+1 protection. This was expected as described in the example at the end of Section 2.

We computed (not shown in the graphs) that with FP alone, the average link utilization in the logical topology is 15%. (This is matches typical utilization levels observed in many commercial backbones.) We see in the top-middle figure that with BEP, the average logical link utilization increases to 80% (with 1:1) and to 75% (with 1+1) for $\beta_{FREE} = 0$. Even when links are overprovisioned to leave 50% free ($\beta_{FREE} = 0.5$) we can increase the average logical link utilization from 15% to to 50% (for 1:1) and to 44% (for 1+1). The fact that the utilization of the maximally loaded link decreases from 100% to 50% as β_{FREE} increases from 0 to 50% demonstrates the correctness of our implementation, because out target is for the maximally loaded link to be equal to $(1 - \beta_{FREE})c^l_{st}$ (as required by equation 13). The curve for the two maximums is in fact the line $(1 - \beta_{FREE})c^l_{st}$ converted to percentages.

In Figure 3 on the top-right, we see that the maximum physical link utilization is at 55%. The average physical link utilization lies between 30% for $\beta_{FREE} = 0$ and 20% for $\beta_{FREE} = 0.5$; this is intuitive since increasing the β_{FREE} factor leads to a decrease of the *logical* link utilization and yields less traffic in the physical network. The fact that the slope of the decrease in average utilization at the optical layer is smaller than at the logical layer, is because the physical layer is not highly used since each logical connection is mapped to a single channel and we have 16 channels per fiber. This comes from the fact that 6 of the 9 logical links have their bottleneck at the IP layer. We also see that the 1:1 protection strategy requires a little bit less physical bandwidth than 1+1.

On the bottom-left figure, we see that when failures happen, we lose on average about 20% of the BEP traffic for both the protection strategies. The maximum BEP lost is approximately the same for 1:1 (50%) and

1+1 (48%). As β_{FREE} increases, there is a slight increase in the amount lost by 1:1, relative to 1+1. This makes sense since 1:1 carries a larger amount of BEP traffic than 1+1, it also looses more.

Examining the utilization levels under failure scenarios, we see that the logical topology is on average still well loaded even under failures. This is because 80% (on average) of the BEP traffic is in fact retained. The average logical link utilization drops by a corresponding amount of 20% while the average physical link utilization drops by around 10% (regardless of the protection strategy).

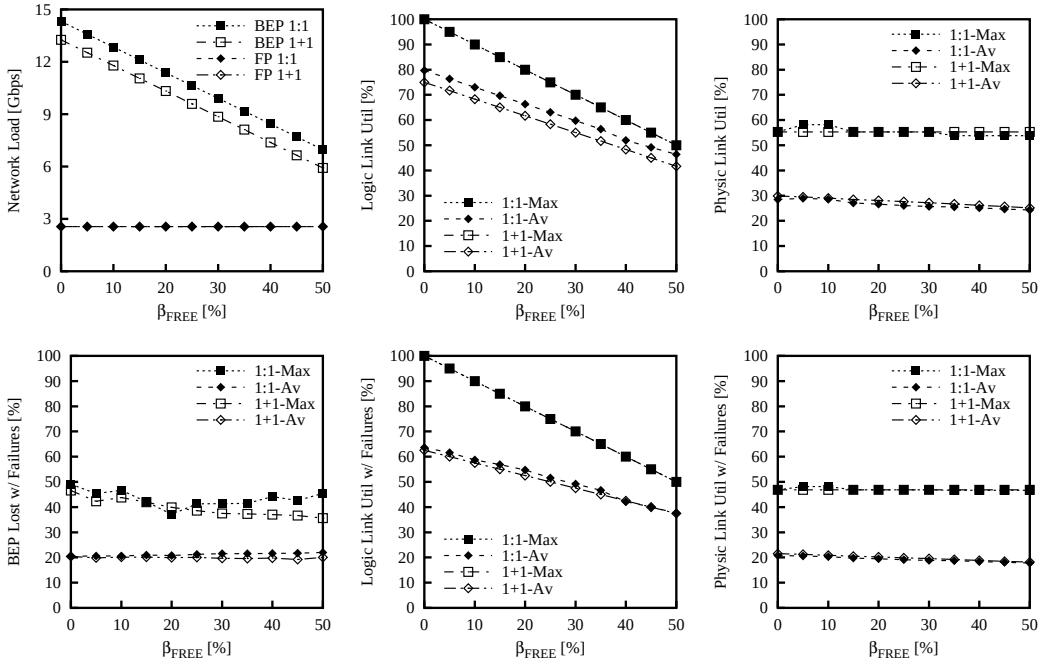


Figure 3. Italian backbone. ILP solution. 7 OC-48 links. Z_{min}=0.

We now examine the impact of the the Z_{min} factor in the second version of our medium-sized network in which physical link (7,9) has been upgraded to an OC-48 link. Figure 4 shows the amount of BEP load that can be carried for this network. First we see that the BEP load decreases by a seemingly fixed amount each time Z_{min} increases by 100 Mbps. This illustrates the tradeoff that in order to give *each* logical connection some BEP traffic, we must reduce the total amount of BEP traffic carried.

Second we observe that the curves for $Z_{min} = 300$ and $Z_{min} = 400$ stop at some point as we increase β_{FREE}. This illustrates an important impact of the overprovisioning factor, namely that if we require a certain amount of overprovisioning, it can limit the minimum amount of BEP traffic we can offer. For example, if a carrier requires β_{FREE} to be 40%, then it is not possible to guarantee each logical connection a minimum of 400 Mbps of BEP traffic. Hence there is a tradeoff between overprovisioning and fairness. Intuitively this is reasonable since both require additional bandwidth, a finite resource that needs to be partitioned between these two objectives.

We now present our six metrics for this second version of our medium-sized network in Figure 5. For this test case we set $Z_{min} = 200$. The impact of the upgrade of link (7,9) is the following. All of 9 logical connections now have one of their physical paths with an OC-48 rate and the other with an OC-12 rate. The bottleneck has thus been moved to the IP layer for all logical links. The most striking observation from all the graphs in Figure 5 is that the performance difference between 1:1 and 1+1 has disappeared. This reason for this is because the bottleneck is now at the IP layer. Recall that with two physical paths of OC-48 and OC-12, the total capacity available at the physical layer is 2.5 Gbps plus 622 Mbps. However since the routers have only OC-48 interface cards, the maximum amount of aggregated traffic a logical link can put into the network is 2.5 Gbps. When the bottleneck is at the IP layer, one cannot take advantage of idle backup bandwidth, such as in 1:1 protection.

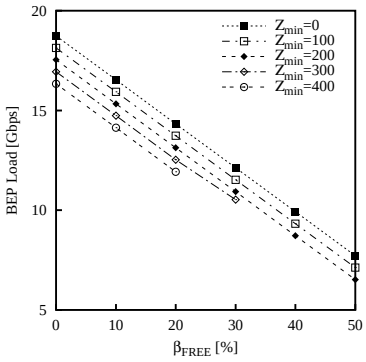


Figure 4. Italian backbone. ILP solution. 8 OC48 links.

We also note in this case that the total BEP load when $\beta_{FREE} = 0$ is at 17.5 Gbps; this is 6.25 times as much as the FP traffic, for an increase by a factor of roughly 7 for the total load. In the previous network with link (7,9) at OC-12 speeds, the maximum increase in the total load was a factor of 6. This factor of 7 would be even higher if we let $Z_{min} = 0$ as in the first version of the Italian backbone. The main point here is to illustrate that upgrading a single link can allow for a great deal of extra BEP to be carried in the network. This is advantageous because most carriers upgrade their backbone networks slowly; each link upgrade can take a few months. Thus carriers often find themselves in a position in which there backbone links are heterogneous. Our FP/BEP proposal allows carriers to extract benefit from this heterogeniety.

We see other ramifications of the fact that the bottleneck has been moved to the IP layer. First, the average and maximum logical link utilizations are the same. Second, note that the maximum physical link utilization in Figure 3 with seven OC-48 links and 14 Gbps of BEP load was at 55%, while in Figure 5 with eight OC-48 links and 18 Gbps of BEP load, the maximum physical link utilization is only at 50%. This is because we have upgraded link (7,9) and the bottleneck is now at the IP layer.

6. CONCLUSIONS

In this paper we have defined two classes of service and designed an ILP model that enables the use of two classes of service, differentiated by their type of protection, to be deployed in an IP-over-WDM network. Our model gives us a method, to find optimal paths, that incorporates a number of practical features including an over-provisioning requirememt, a minimal fairness factor for BEP traffic allocation, a specification of heterogeneous FP traffic demands via a traffic matrix at the IP layer, and heterogeneity in router and OXC interface cards.

We studied the gain of such an approach and found that it allows the total network load to be increased by a factor that varies between 4 and 7, depending upon the specific network scenario. This approach is appealing and timely because carriers today operate their networks at low utilizations levels and are interested in increasing their revenue without impacting the existing traffic. We found that even if carriers want to over-provision their networks by requiring that 50% of the bandwidth remain unused, we can still increase the total network load by a factor of 3.

If an operator provided FP traffic alone, the total amount of FP would be the same as seen in the three networks we evaluated. This was true was for all settings of the overprovisioning factor β_{FREE}, and illustrates that we were able to add BEP onto the network while still providing FP full protection and without impacting the FP load.

In our medium-sized network with heterogeneous links and interface cards, we observed that on average 20% of the BEP traffic gets preempted during a typical single link failure scenario. That means that even though we provide no protection guarantees at the WDM layer to BEP traffic, 80% of it was still unaffected by the failure. In our large network with homogeneous interface cards at both the OXC's and the routers, there was no loss of BEP traffic. The reason for this is that all of the bottlenecks were at the IP layer which was unable to take

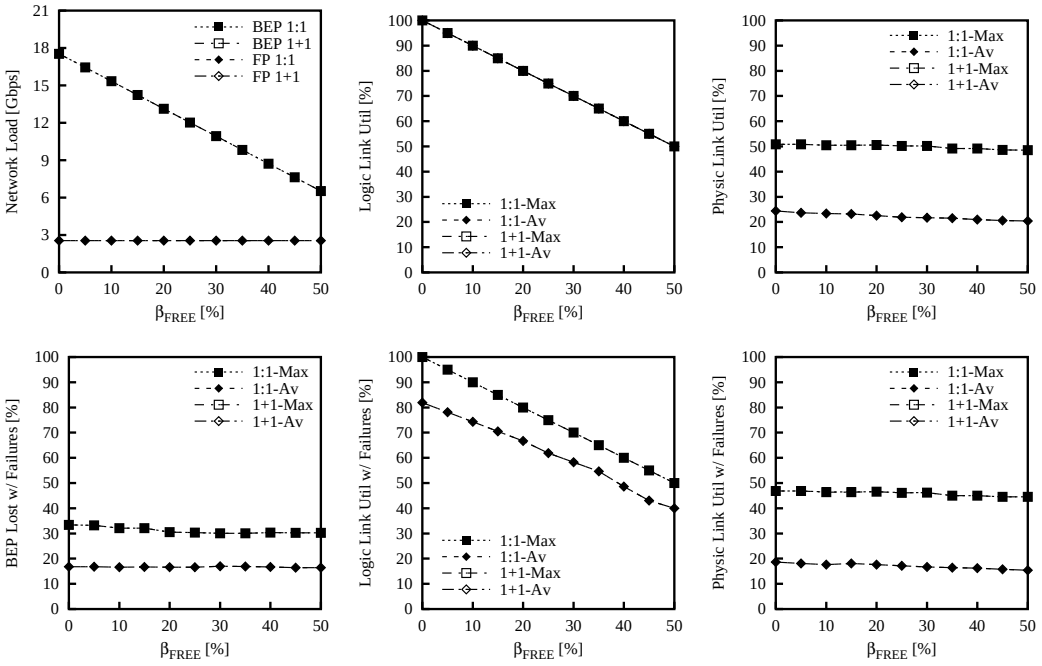


Figure 5. Italian backbone. ILP solution. 8 OC-48 links. $Z_{min} = 200$.

advantage of the excess capacity at the optical layer. We thus conclude that our proposal for FP/BEP service classes has the most benefit in networks that are heterogeneous in their links and interfaces.

We demonstrated that the location of the bottleneck plays an important role. If in a network, there are some logical links whose bottleneck resides at the WDM layer, then we see a difference between the 1:1 and 1+1 protection strategies. In particular, the 1:1 strategy can carry more BEP traffic than 1+1. However if all of the logical links have their bottleneck at the IP layer, then there is no difference between the two protection strategies.

We illustrated here that upgrading a single link can have a very large impact on the amount of BEP traffic carried. Since the process of upgrading all the links in a large backbone can take many months, our FP/BEP proposal allows a carrier to make use of dispersed pockets of additional bandwidth to increase their revenues.

Finally we demonstrated the tradeoff between two typical carrier requirements, an over-provisioning requirement and a minimal BEP fairness requirement. If both requirements are small, then they can both be met; however if they both become large (e.g., $\beta_{FREE} > 20\%$), then it is not always possible to satisfy both requirements simultaneously. The amount of overprovisioning essentially limits the minimal BEP allocation we can provide to logical connections.

In our ongoing efforts, we are incorporating the failure scenarios into our model so that the routes computed are optimal over all possible failure scenarios. We are also studying scenarios in which all the protection or restoration is done at the IP layer because some carriers are considering eliminating SONET from their backbones.

7. APPENDIX

Note that in the above formulation the problem related to the possible existence of *walks* or *isolated nodes* during the building of the working and backup physical paths was not taken in consideration. A *walk* (or *edge train* or *chain*) between two vertex v_s and v_f is defined as a finite alternating sequence of vertices and edges, starting from v_s and closing to v_f, such that each edge is incident with the vertices preceding and following it. No edge appears (is covered or traversed) more than once in a walk. A vertex, however, may appear more

than once. We define n_i as an *isolated node* if it can not be reached from any other vertex by traveling along the edges.[14] As a consequence, we could allocate resources on some fibers that could not be used to carry traffic. However, this is only a theoretical problem since the maximization of the total load in the network is equivalent to the minimization of the unused resources allocated for the logical connections. This phenomena can only occurs when the network has not much traffic to carry. To eliminate totally this problem we have to consider in the model the following variables to build correctly the *covering tree* for the working and backup physical paths and then use only the necessary resources in the network. Let θ_{ij}^{st} and γ_{ij}^{st} be integer variables used to build the *covering graph* respectively for the working and backup paths associated to the logical link $(s,t) \in E^1$ while the binary variables ξ_{ij}^{st} and η_{ij}^{st}) are used to transform the *covering graph* into a *covering tree*. Note that the *covering graph* is a graph whose each node is able to reach any other node and represents all the available links that the primary and backup paths can use to route the traffic. Building a *covery graph* eliminates totally the *isolated nodes* problem. A *covering tree* can be obtained from a *covering graph* simply imposing that the number of edges in the graph is exactly equal to the number of its nodes minus one. Building a *covery tree* eliminates totally the *walks* problem. We can write the following constraints:

- Covering graph for the working and backup paths associated to the logical link (s,t):

$$\sum_{j \in V:(i,j) \in E^0} \theta_{ij}^{st} - \sum_{j \in V:(j,i) \in E^0} \theta_{ji}^{st} = \begin{cases} N-1 & \text{if} \quad i = s \\ -1 & \text{if} \quad i \neq s \end{cases} \quad \forall i \in V, \forall (s,t) \in E^1 \tag{17}$$

$$\sum_{j \in V:(i,j) \in E^0} \gamma_{ij}^{st} - \sum_{j \in V:(j,i) \in E^0} \gamma_{ji}^{st} = \begin{cases} N-1 & \text{if} \quad i = s \\ -1 & \text{if} \quad i \neq s \end{cases} \quad \forall i \in V, \forall (s,t) \in E^1 \tag{18}$$

  where (17) and (18) ensure that starting from the source node $s \in U$ of the logical link (s,t) we can reach all the other network nodes. The source node s put in the network $N-1$ tokens to reach all the other $N-1$ network nodes. Each network node catch for himself one token and then all the network nodes are reached. The covering graph eliminates the islands problem but no the cycles problem. To eliminate also this last problem we must consider the other following constraints that force the graph to become a tree:

- Covering tree for the working and backup paths associated to the logical link (s,t):

$$\xi_{ij}^{st} \leq \theta_{ij}^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{19}$$

$$\xi_{ij}^{st}(N-1) \geq \theta_{ij}^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{20}$$

$$\sum_{(i,j) \in E^0} \xi_{ij}^{st} = (N-1) \qquad \forall (s,t) \in E^1 \tag{21}$$

  where (19) and (20) simply state that the working path associated to the logical link (s,t) cross over the physical link (i,j) if and only if there is a flow bigger than 0. The (21) ensure that the covering graph is a tree and then it lacks in any cycle. We can write the same set of equations for the backup physical paths:

$$\eta_{ij}^{st} \leq \gamma_{ij}^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{22}$$

$$\eta_{ij}^{st}(N-1) \geq \gamma_{ij}^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{23}$$

$$\sum_{(i,j) \in E^0} \eta_{ij}^{st} = (N-1) \qquad \forall (s,t) \in E^1 \tag{24}$$

The last two relations (25) and (26), link up the variables ξ_{ij}^{st} and η_{ij}^{st} respectively with the variables w_{ij}^{st} and b_{ij}^{st}:

$$w_{ij}^{st} \leq \xi_{ij}^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{25}$$

$$b_{ij}^{st} \leq \eta_{ij}^{st} \qquad \forall (i,j) \in E^0, \forall (s,t) \in E^1 \tag{26}$$

REFERENCES

1. B. Schroeder, S. Bhattacharyya, N. Taft and C. Diot, "IS-IS Link Weight Assignment for Transient Link Failures", *Submitted for Publication*, Sprint ATL Technical Report TR02-ATL020133, 2002.
2. K. Papagiannaki, S. Moom, C. Fraleigh, P. Thiran, F. Tobagi and C. Diot, "Analysis of Measured Single-Hop Delay from an Operational Backbone Network", *To appear in IEEE Infocom 2002*, New York, June, 2002.
3. C. Boutremans, G. Iannaccone and C. Diot, "Impact of link failures on VoIP performance", *To appear in NOSSDAV*, May, 2002.
4. P. Thiran, N. Taft, C. Diot, H. Zang and R. MacDonald, "A Protection-Based Approach to QoS in Packet-over-fiber Networks", *International Workshop on Digital Communications*, Taormina, ITALY. September, 2001
5. O. Gerstel and R. Ramaswami, "Optical Layer Survivability: A Services Perspective", *IEEE Communication Magazine*, vol. 38(3), pp. 104–113, March 2000.
6. R. Ramamurthy and B. Mukherjee, "Survivable WDM Mesh Networks", *Proc. Infocom 1999*, New York, March 1999.
7. G. Mohan and A. K. Somani, "Routing Dependable Connections with Specified Failure Restoration Guarantees in WDM Networks", *in Proc. INFOCOM '00*, Jerusalem, April 2000.
8. M. Sridharan and A. K. Somani, "Revenue Maximization in Survivable WDM Networks", *Opticomm 2000*, Dallas, October 2000.
9. O. Crochat and J.-Y. Le Boudec, "Design Protection for WDM Optical Networks", *IEEE Journal on Selected Areas in Communication*, vol.16, no.7, pp.1158-1165, September 1998.
10. O. Crochat, J.-Y. Le Boudec, and O. Gerstel, "Protection Interoperability for WDM Optical Networks", *IEEE Transaction on Networking*, vol.8, no.3, pp.384-395, June 2000.
11. E. Modiano and A. Narula-Tam, "Survivable routing of logical topologies in WDM networks", *in Proc. INFOCOM '01* vol.1, pp.348-357, Anchorage, Alaska, April 2001.
12. M. Plante and B. Sansó, "A Typology for Multi-technology Multi-service Network Synthesis Methods", *GERAD Publication*, G-98-51, July 1999.
13. `http://www.ilog.com/products/cplex/`
14. N. Deo, "Graph Theory with applications to engineering and computer science", *Prentice Hall*, 1974.