A PROTECTION-BASED APPROACH TO QOS IN PACKET OVER FIBER NETWORKS

Patrick Thiran†‡, Nina Taft †, Christophe Diot †, Hui Zang †and Robert Mac Donald †
†Sprint Advanced Technology Labs, 1, Adrian Court, Burlingame, CA 94010
‡ICA-DSC, EPFL, CH-1015 Lausanne, Switzerland*

Abstract

We propose a novel approach to Quality of Service, intended for IP over SONET (or IP over WDM) networks, that offers end-users the choice between two service classes defined according to their level of transmission protection. The first service class (called Fully Protected (FP)) offers end-users a guarantee of survivability: all FP traffic is protected in the case of a (single) failure. The second service class (called Best-Effort Protected (BEP)) does not offer any specific level of protection but is cheaper. When failures occur, the network does the best it can by only restoring as much BEP traffic as possible. We motivate the need for two classes of protection services based on observations about backbone network practices that include overprovisioning and an ongoing but unbalanced process of link upgrades. We use an ILP formulation of the problem for finding primary and backup paths for these two classes of service. As a proof of concept, we evaluate the gain of providing two protection services rather than one in a simple topology. These initial results demonstrate that it is possible to increase the network load (and hence revenue) without affecting users that want complete survivability guarantees.

1 Introduction

Today's internet backbone contains a large amount of unused capacity due primarily to the following three reasons: overprovisioning, duplication of equipment and unbalanced link upgrades. Overprovisioning is the current de facto solution to providing QoS.

A lot of effort is devoted to broadening the set of Internet services to a palette ranging from best-effort to real-time and streaming services. The proposed solutions for such services differ in the mechanisms they use - such as reservation or priority. However, their common goal is to provide users with a variety of service classes that differ based on their performance with respect to throughput, loss and/or delay measures. Such a differentiation is indeed useful when congestion occurs in portions of the network. But backbone networks are usually over-provisioned because it is often simpler and cheaper to buy additional hardware equipment than to run complex software

*This technical report will be published as a book chapter in the Lecture Notes of Computer Sciences, Springer-Verlag

for managing reservations and priorities in routers. Hence traffic rarely experiences congestion in the backbone [1], making service differentiation quite useless in practice. Overprovisioning allows carriers to provide everybody with the best class of service.

Not only is the backbone overprovisoned to offer low delay and losses to all traffic, but most equipment is duplicated for protection against failures. Carriers are not willing to forgo this additional redundancy because they do not want network services to be disrupted, even rarely. (Failures are actually less rare than one might expect; [2] has recently reported failure rates of 1 per year per 300km of fiber.) Avoiding service disruption is especially critical for backbone links, where a single failure may interrupt many channels. A large fraction of the capacity in the backbone links therefore remains unused, and this situation is likely to continue as long as the bottlenecks are in the access network rather than the backbone.

On the other hand, because traffic demands grow exponentially, network operators are continuously obliged to upgrade the capacity of their backbone links. Upgrading a backbone link can be a lengthy operation, and thus in practice links are upgraded one at a time. Many months can pass between the upgrading of two links. Providing protection means that an upgrade of the capacity for primary working links, should be matched by an equivalent upgrade of the redundant protection links. However, since the network is essentially in a continual state of flux, the typical network is quite heterogeneous containing some recently upgraded high-speed links (e.g. links with a DWDM system of 80 to 160 wavelengths operating at a 10 Gbps line speed), alongside older slower-speed links (e.g. WDM fibers with only 4 to 32 wavelengths at 2.5 Gbs line speed). This situation prevents operators from making use of the capacity in recently upgraded links. To see why, consider the following scenario. Suppose all links are initially 2.5 Gbps and then exactly one of them is upgraded to 10 Gbps. The full capacity of this link cannot be used for paths spanning multiple hops for two reasons. First, other links may not be able to support the growth in traffic, and second, it is unlikely that a backup path, on the other 2.5 Gpbs links, can be found for this additional traffic.

The combination of overprovisioning, redundant capacity for failures, and partial network upgrades creates a situation in which, on a day-to-day basis, there exists a large amount of unused bandwidth in the Internet backbone. In order to leverage this unused bandwidth we propose the use of two classes of service that differ according to the protection level provided. The two service classes are intended for either IP/SONET or IP/WDM networks with IP at the logical layer and either SONET or WDM systems at the physical layer. The first one, hereafter called the *Fully Protected* (`FP`) class, offers users the insurance that none of their traffic will be disrupted in case of a single point of failure. The second one, hereafter called the *Best-Effort Protected* (`BEP`) class, does not provide any specific guarantee on service disruption. Instead, in the case of failure, it offers to restore as much of the affected traffic as possible. What BEP offers to users, as a tradeoff for a lower amount of protection, is either a larger throughput, or a cheaper price. We will discuss how having two such services allows carriers to carry the BEP traffic on the excess capacity without impacting the FP traffic.

Many proposals for service classes differentiate the classes according to their delay, loss or throughput performance. Reliability is also an important QoS performance metric and the wide variety of applications that exist today demand different levels of availability guarantees.

Some applications, such as IP telephony, video-conferencing, and distance surveillance require 100% availability and hence full protection against network failures. Others, like on-line games, Web surfing, and Napster downloads are likely to be willing to tradeoff a partial and slower protection for increased throughput (or a lower price). Such tradeoffs are attractive as long as the probability of a service becoming unavailable is very small. Applications like e-mail can fall into either one of these service classes.

The reliability dimension of QoS can be quite independent of the traditional QoS parameters of delay, loss and throughput that are often correlated to one another. For example, two applications requiring similarly high levels of reliability need not have similar delay requirements. Most applications requiring full protection will be the priority traffic, but this may not always be the case. Table 1 demonstrates that categorizing applications by their protection needs can be different than categorizing them according to their traditional QoS needs. Reliability also differs from these traditional QoS measures in that delay, loss and throughput guarantees can be trivially satisfied by overprovisioning (if you are willing to pay for it), whereas reliability cannot because the amount of overprovisioning has to be carefully calculated. Overprovisioning to provide delay, loss and throughput guarantees can be done by simply inflating each link by say 20 or 30%, or by ensuring that the load on each link rarely exceeds specified thresholds (e.g., 60%). However such a per-link view of overprovisioning is insufficient for meeting reliability guarantees which requires a network-wide view of the capacity. This is because all links must be inflated proportionally if one wants to ensure that backup paths will exist for all source-destination pairs of flows. The slow and unbalanced process of link upgrades makes it very difficult to overprovision using a network-wide perspective.

The introduction of services offering different levels of protection guarantees at the WDM layer is gaining attention in the optical networking community. A classification in five classes is proposed in [3]. In our paper, we consider only two classes, but defined at the IP layer. This will result in making some SONET (or WDM) paths protected and others not, as in the work of Sridharan and Somani [4]. Despite some differences with this latter work in the problem formulation (for instance, we do not introduce different costs for each type of working or back-up paths, but we constrain the ratio between BEP and FP traffic to be less than a prescribed maximal value), we reach a similar conclusion, that the traffic load can be increased quite considerably when more then one class of protection is available in a homogeneous network, where all links have the same capacity. We show that this effect is even more accentuated in a heterogeneous network.

The rest of this paper is organized as follows. Section 2 briefly summarizes the kinds of mechanisms provided for protection at the optical and IP layers. We state our service definitions in Section 3 and describe which protection mechanisms are needed by each of the service classes. The ILP formulation of the resulting routing problem is given in Section 4. For a proof of concept demonstration, we provide an example in Section 5 that illustrates that a good deal of BEP traffic can be carried on the network without affecting the FP traffic, and thus we can substantially increase the load (and hence the revenue) the network carries. In Section 6, we extend our ILP formulation in order to secure a minimal amount of bandwdith to restore a fraction of the BEP traffic after a failure, so that this class of traffic does not suffer a complete service disruption in case of a failure, but a softer degradation. We conclude our proposal in Section 7.

Service	Fully Protected	Best Effort Protected
Low delay and losses	IP telephony, distance monitoring	cheap on-line games
Loose delay or loss requirement	professional e-mail	private e-mail, web surfing

Table 1: Service Categorization

2 Handling Failures at the IP and SONET Layers

Defining classes of service for protection requires specification of how protection is handled for each class. Before stating our proposal, we review the mechanisms that are available at each layer in the network. The optical layer provides *protection* that carries out very fast failure recovery but is often not bandwidth efficient [5, 6]. The IP layer can provide *restoration* that helps to determine more efficient routes but is typically not very fast. Most networks today rely on SONET to carry out protection.

Protection at the SONET layer. All protection techniques involve providing some redundant capacity within the network to reroute traffic in case of a failure. Protection is the mechanism by which traffic is switched to available resources when a failure occurs. It needs to be very fast; the commonly accepted standard for SONET is 50 ms. Protection routes must therefore be pre-computed, and wavelengths must be reserved in advance at the time of connection setup. Protection around the failed facility can be done at different points in the network: (i) around the two end-points of the the failed link, by *line* or *span protection* (in optical layer terminology this corresponds to protection at the line or multiplex sublayer), or (ii) by *path protection* which is between the source and destination of each connection traversing the failed link (in optical layer terminology, this corresponds to protection at the path sublayer) [7, 8, 9]. Line protection is simpler, but path protection requires less bandwidth and can better handle node failures. Here, we only consider path protection.

There are essentially two fundamental protection mechanisms. In *1+1 protection* traffic is transmitted simultaneously on two separate fibers on disjoint routes. The receiver accepts traffic from the primary fiber (also called working fiber) and only switches to accept the input from the other fiber (called protection or back-up fiber) in case the primary one fails. In *1:1 protection* traffic is transmitted only on the primary fiber. If this fiber is cut, the sender and receiver use simple signalling to jointly switch to the backup fiber. The generalization of 1:1 protection is 1:n protection where one back-up path protects n working paths. For our initial proof-of-concept analysis, we consider 1+1 and 1:1 protection schemes in this paper.

Restoration at IP layer. Since the IP layer is made up of a well meshed topology and its links are not fully loaded (due to overprovisioning), the IP layer is also capable of restoring traffic around a failed facility.

After SONET protection is done, today's routing protocols can discover routes in the new topology that are more efficient than the backup path used for failure recovery in the old topology. Within a carrier's backbone, Internal Gateway Protocols (IGP) are used for intradomain routing. IS-IS and OSPF are the most common protocols deployed today. In these protocols, routers periodically exchange hello messages to check the health of neighboring links and nodes. If a few successive messages are lost, a router deduces that a link or node is down, and begins the restoration process at the IP layer. After detection of a topology change, this process involves propagating the change information across the network and recomputing shortest paths. During the restoration process, a subset of destinations are reached through non-optimal routes (if the network supports SONET) or are briefly unreachable (otherwise). In IS-IS, the process of failure detection can take between 10-30 seconds depending upon the protocol configuration, and the rest of the recovery process can take another 10 seconds or so [10]. Although ISIS convergence today takes on the order of tens of seconds, it is believed [10] that these convergence times can be greatly reduced, potentially to the order of tens of milliseconds. The theoretical limit of link-state routing protocols to reroute is in link propagation time scales - in other words in the tens of milliseconds. Using today's technologies, restoration speed at the IP layer cannot compete with the protection and restoration speeds at SONET (or WDM) layers.

A difficulty that arises in today's networks, e.g., IP/SONET, is that each layer performs protection independently from the other layers. For example, IGP routing table updates are performed independently of SONET's line protection. This can lead to undesirable race conditions between different layers. Ideally IP and optical networks should be managed as an integrated network without overlap of functionality between layers and with sharing of information between layers. The issue of deciding exactly which aspects of protection and restoration should be carried out by which layer is still an open issue. The advantage of providing protection at the IP layer is the cost reduction that results from saving redundant equipment at the physical layer. The disadvantage is that it is slow. Providing protection at the SONET layer has the reverse tradeoff.

3 Definition and Provisioning of Service Classes

We now define our two service classes that differ in terms of their level of protection, their mechanism of protection and their cost.

Fully Protected (FP) service class.

- This service guarantees its customers that their traffic is protected against any single point of failure in the backbone within 50 msec.
- This service provides fast protection. Therefore FP traffic is protected via pre-computed, dedicated back-up paths at the SONET or WDM layer, using either by 1:1 or 1+1 protection. Failures are transparent to the IP layer for this class of traffic.
- This service is the more expensive of the two.

Best Effort Protected (BEP) service class.

- This service does not offer specific guarantees for protection against failures, but instead tries to restore as much of this traffic as possible after a the occurrence of a failure.

- For BEP traffic we offer restoration and not protection. When a failure occurs, BEP packets will 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. Actually this service can come in a variety of flavors. The simplest version of this service class is to leave BEP traffic entirely unprotected at the SONET (and/or WDM) layer. A more enhanced version of this service (and more difficult to implement) is to ensure users that in case of a single failure, they would not experience a complete service disruption but may experience a severe degradation.

- This service is cheaper than the FP service.

In order to implement two such service classes, packets would need to be marked according to their service class, and IP routers would need 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 are dropped until BEP traffic has been restored at the IP layer.

4 ILP formulation

We formulate the problem of routing traffic flows from two service classes over a physical and logical topologies as an Integer Linear Programming (ILP) problem whose objective is to maximize the total load carried by the network, which we denote by F.

We consider here that all *physical channels* are SONET paths. They could also be WDM lightpaths, if all optical cross-connects have full wavelength conversion possibilities, or if they perform electronic conversion before switching, so that wavelength continuity constraints [5] can be ignored. Each path is assigned a unit capacity, which represents the smallest granularity level of bandwdith of a SONET path. The total capacity C_l of physical link l, with $1 \leq l \leq L$ where L is the number of physical links, is thus an integer.

The *logical topology* is the set of logical links between IP routers. Let M be the numbers of routers that are connected by a logical link. Each logical link between router s and router t has capacity d_{st}, and is a set of consecutive physical links that form a route r. Logical links are considered here as bi-directional, i.e. d_{st} is the sum of the demand from s to t and from t to s. In the following, we need thus to consider only source-destination pairs (s, t) with $s < t$. This assumption can clearly be relaxed.

Each logical link presents a demand of d_{st} capacity units at the physical layer, for which one needs to find a route r among the set of all routes $\mathcal{R}_{st}$ between the source s and the destination t, such that the capacity constaints of all links l belonging to route r are satisfied. To keep routing at the physical layer simple, we do not allow multiple working paths between a given pair of nodes. Denoting by d_{st}^r the traffic flowing on route $r \in \mathcal{R}_{st}$, we have thus that for all

$1 \leq s < t \leq M$

$$d_{st} = \max_{r \in \mathcal{R}_{st}} \{d^r_{st}\}. \tag{1}$$

If multiple routes were allowed, one would have to change the maximum in this equation, by a sum.

A logical link between a given pair of nodes (s,t) carries d^{FP}_{st} traffic units of the FP class, and d^{BEP}_{st} traffic units of the BEP class:

$$d_{st} = d^{FP}_{st} + d^{BEP}_{st}. \tag{2}$$

Because of (1), both traffic classes are carried on the same route $r \in \mathcal{R}_{st}$.

In the simplest case, no precaution is taken to guarantee even a partial restoration of BEP traffic at the logical layer. This means that BEP traffic can be left unprotected, and will be restored only if resources are available after the failure has occurred. In the worst case, all BEP traffic may have to be dropped as a result of a failure.

On the other hand, FP traffic is protected on a 1+1 or 1:1 basis. This is the simplest and fastest recovery scheme, but also the most resources consuming. It requires that for each primary route $r \in \mathcal{R}_{st}$, we find a link disjoint route r' (if we have only link failures) or even a link and node disjoint route (for the general case where both link and node failures can occur) from r, that can carry d^{FP}_{st} traffic units. We consider here only the case of link failure. To state the resulting constraint, we first introduce the membership function

$$\delta^r_l = \begin{cases} 1 & \text{if} \quad l \in r \\ 0 & \text{if} \quad l \notin r \end{cases}$$

for any link $1 \leq l \leq L$ and any route $r \in \mathcal{R}_{st}$. We must therefore find a route $r' \in \mathcal{R}_{st}$ such that the traffic demand $d^{r',PR}_{st}$ on this protection route verifies for all $1 \leq s < t \leq M$, $r \in \mathcal{R}_{st}$

$$d^{r',PR}_{st} = d^{FP}_{st} \tag{3}$$

$$\sum_{1 \leq l \leq L} \delta^r_l \delta^{r'}_l d^r_{st} d^{r',PR}_{st} = 0. \tag{4}$$

Constraint (3) states that all traffic of the FP class, flowing on the primary route r, must be protected by a traffic allocation of the same amount on a back-up route r'. Constraint (4) ensures that it is link disjoint with r.

The finite link capacity imposes that for all $1 \leq l \leq L$

$$\sum_{s=1}^{M} \sum_{t=s+1}^{M} \sum_{r \in \mathcal{R}_{st}} \delta^r_l \left(d^r_{st} + d^{r,PR}_{st} \right) \leq C_l. \tag{5}$$

The two last constraints are provided by the actual traffic data.

The first one is the proportion of traffic belonging to both classes. We represent the amount of BEP traffic between node pairs as a given multiple ρ of the FP traffic. Clearly, FP traffic will require more resources than BEP traffic, so we need to set a maximum value to this ratio, since otherwise the optimal solution will always consist in having all traffic in the BEP class. Therefore we constrain ρ to be less than than a given maximal value $\rho_{\max}$:

$$d_{st}^{BEP} \leq \rho_{\max} \cdot d_{st}^{FP} \tag{6}$$

for all $1 \leq s < t \leq M$.

The second one is the fraction of the total load F that needs to be assigned between each pair of nodes, and which would be obtained by the IP traffic matrix data. By default, we assume here a balanced repartition of the load between each pair of IP nodes, so that the same fraction of the total load is assigned between each pair of nodes:

$$d_{st} = d_{s't'} \tag{7}$$

for all $1 \leq s < t \leq M, 1 \leq s' < t' \leq M$.

The problem amounts therefore to maximize

$$F = \sum_{s=1}^{M} \sum_{t=s+1}^{M} d_{st}$$

subject to constraints (1) to (7).

5 Example

In this section, we illustrate our ideas with a numerical example. The goal of this example is to serve as a proof of concept to demonstrate the gain that can be achieved by supporting more than one protection service class. In today's networks the only protection class is FP.

Figure 1 shows a network, consisting of $N = 6$ nodes and $L = 8$ links at the physical layer (SONET/WDM), and of $M = 3$ nodes and $M(M-1)/2 = 3$ links at the logical layer (IP). We consider here that all physical channels are SONET paths. Remember that the capacity unit is the smallest capacity of a SONET path, and that the capacity C_l of a link l is therefore an integer multiple of this capacity unit. In our example, the capacity of each phyiscal link is equal to 8, if the link has not been upgraded, and to 32, if the link has been upgraded. Figure 1 shows one possible mapping of the logical links (right) on the physical links (left), which is as follows:

- logical (IP) link (A, B) is mapped on working (SONET) physical path (or route) $\{(1,2)\}$ and back-up (SONET) physical path $\{(1,5),(5,6),(2,6)\}$;
- logical link (A, C) is mapped on working physical route $\{(1,4),(3,4)\}$ and back-up physical route $\{(1,5),(3,5)\}$;
- logical (IP) link (B, C) is mapped on working physical route $\{(2,3)\}$ and back-up physical route $\{(2,6),(5,6),(3,5)\}$.

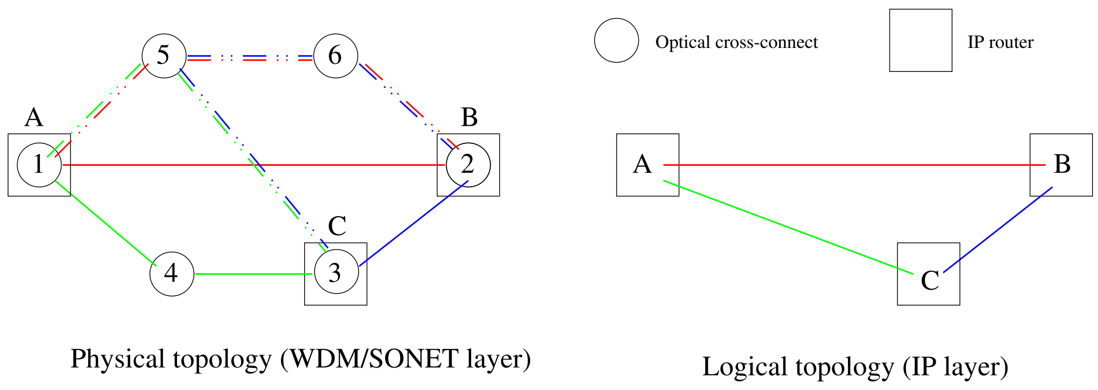

Figure 1: A SONET/WDM network (left) with working paths in plain and back-up paths in dashed lines. The logical topology at IP layer is represented on the right, and consists here of three logical links.

Other mappings are of course possible, the mapping which will be eventually adopted is the one that solves the ILP described in the previous section.

We use ILOG optimizer [11] to find the solution of the ILP. Figure 2 displays the results, when the following number of links have been upgraded: (i) none, (ii) two links ($(1,2)$ and $(2,3)$), (iii) four links ($(1,2)$, $(2,3)$, $(3,4)$ and $(1,4)$) and (iv) all eight links. The x-axis denotes $\rho_{\max}$, which is defined by (6).

First observe the case when all links have the same capacity, either before an upgrade or after an upgrade of all links. If we compare the scenario without any BEP traffic ($\rho_{\max} = 0$), and a scenario with BEP traffic ($\rho_{\max} = 1$), we see that we can nearly double the load on the network. As $\rho_{\max}$ denotes the maximal ratio between BEP and FP, it is natural that after some value of $\rho_{\max}$, the curves become flat because no more additional traffic can be added in the system.

Second, consider the case of a partial upgrade and say $\rho_{\max} = 4$ for example. If only two links are upgraded no additional FP traffic can be carried on the network. However a good deal of BEP traffic can be added after the upgrades. In case of an upgrade of four (appropriately chosen) links, one even reaches the same capacity as with a full upgrade of all eight links, for $\rho_{\max} \geq 7$.

6 Partial restoration of BEP traffic at IP layer

In the previous example, no precaution was taken to prevent BEP traffic from being dropped in case of a failure. It is however desirable that the connectivity of the IP layer be preserved after a single failure, so that BEP traffic is partly restorable (by partly restorable, we mean here that every IP node is reachable, but that queuing delays may become signifigant).

This imposes an additional constraint on the mapping of the logical topology on the physical topology, namely that a single failure leaves the logical topology connected. This problem has

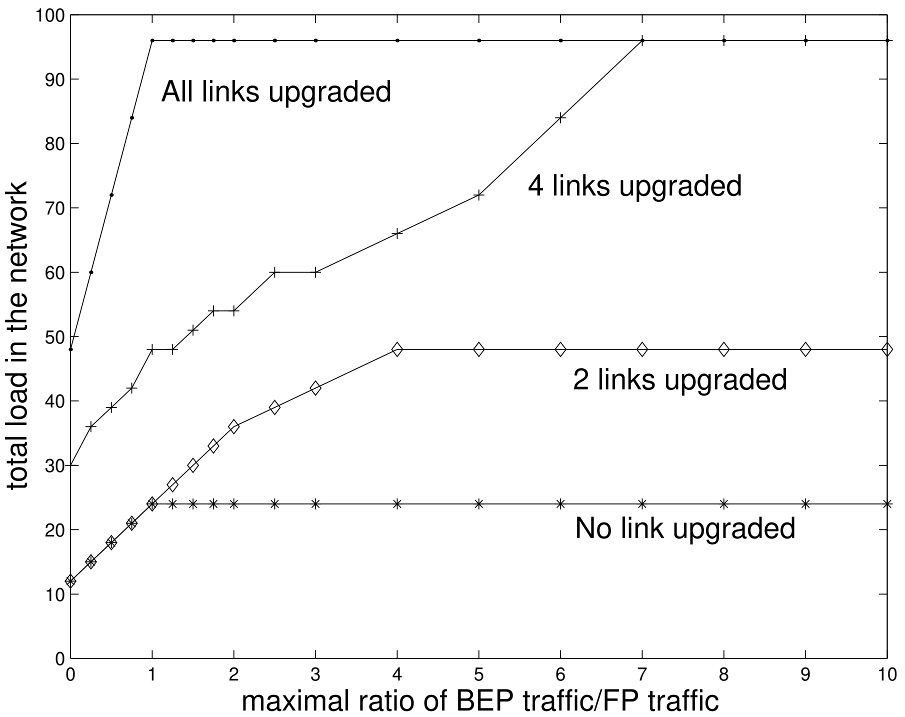

Figure 2: Total number of demands (total load) versus maximal ratio $\rho_{\max}$ of BEP traffic over FP traffic.

been shown to be NP-complete [12], and therefore requires heuristics for general logical and physical topologies. However, when the logical topology is a ring (as in our example), this constraint becomes particularly simple to state [13]: one must simply check that no physical link is shared by two logical links, since otherwise the failure of such a physical link would leave the logical topology un-connected. In other words, we now introduce the additional constraint that for all $1 \leq s < t \leq M$, $1 \leq s' < t' \leq M$, with $(s,t) \neq (s',t')$, and for any $r \in \mathcal{R}_{st}$ and $r' \in \mathcal{R}_{s't'}$

$$\sum_{1 \leq l \leq L} \delta_l^r \delta_l^{r'} d_{st}^r d_{s't'}^{r'} = 0. \tag{8}$$

In this case, the curve in Figure 2 for 2 upgraded links coincides with the curve for zero upgraded link. However, the curve for 4 upgraded links remains unchanged.

In this network, after a single failure on any link of the network, all FP traffic can therefore be rerouted on alternate routes offering the same capacity, whereas BEP traffic that used the broken link now needs to share routes with other BEP flows. As a result, congestion will occur for BEP traffic. In the example above, it is easy to check that the capacity offered to all BEP traffic after a failure is half the capacity it had before.

A better service would be provided for the BEP class, if we slightly over-provision the links taken by BEP traffic, so that it has some spare capacity from which it can benefit to absorb occassional bursts of traffic when no failure has occurred, and to offer a less severe degradation after the occurrence of failure, to rerouted BEP traffic.

Let us denote by ε the *amount of over-provisioning* we provide to the traffic class. This means that for every demand of d_{st}^{BEP} BEP traffic units between s and t, we will actually reserve $(1+\varepsilon)d_{st}^{BEP}$ capacity units. Because of the logical ring topology of our example, one can check that a single failure will then always leave a fraction $(1+\varepsilon)/2$ of the capacity needed for restoring BEP traffic at the IP layer. The value $\varepsilon = 0$ corresponds to the previous case, where BEP traffic receives half the traffic it has before a failure. A value $\varepsilon = 1$ corresponds to a fully restorable BEP traffic at the IP layer (in which case the only difference between the FP and BEP traffic is the layer, and thus the speed, at which traffic is restored).

This amounts to replacing constraint (5) by

$$\sum_{s=1}^{M} \sum_{t=s+1}^{M} \sum_{r \in \mathcal{R}_{st}} \delta_l^r \left(d_{st}^r + \varepsilon d_{st}^{r,BEP} + d_{st}^{r,PR} \right) \leq C_l$$

where $d_{st}^{r,BEP} = d_{st}^{BEP}$ because of (1) and (2). Of course, we need to keep (8) in the set of constraints.

Figure 3 shows the resulting total load when $\varepsilon = 0.5$. Because of the overprovisioning, the total load has decreased, compared to the scenario depicted in Figure 2. In this new scenario, upgrading 4 links no longer allows the network to reach the same total load level as in the case of upgrading all 8 links. This is to be expected as it illustrates the tradeoff between carrying extra load and providing (partial) restoration. However there is still a sizable gain in having two protection services. For example, in this case of partial restoration for BEP, with 4 links upgraded our approach can double the amount of new load carried as compared to a system with a single service ($\rho = 0$)

7 Conclusion

We proposed two service classes based on the level of reliability required by users. The FP class ensures fast protection at the SONET or WDM layer, and makes failures transparent to the IP layer. The BEP class does not offer any availability guarantees after a failure, and is left unprotected at the SONET and/or WDM layers. This proposal allows carriers to make good use of a few upgraded backbone links that otherwise would provide limited benefit until the majority of the backbone links have been similarly upgraded. Preliminary results show that in heterogeneous networks resulting from partial upgrades, our approach allows a substantial amount of additional traffic to be carried. In particular, we showed that in the case of our simple topology, when half of the network links are upgraded the amount of new load carried can be doubled or tripled (depending upon the amount of protection offered to BEP users) as compared to an environment that supported only a single full protection service. Our results demonstrate that by having a second protection class of service, carriers achieve a new method of generating revenue without harming their existing protection class of service.

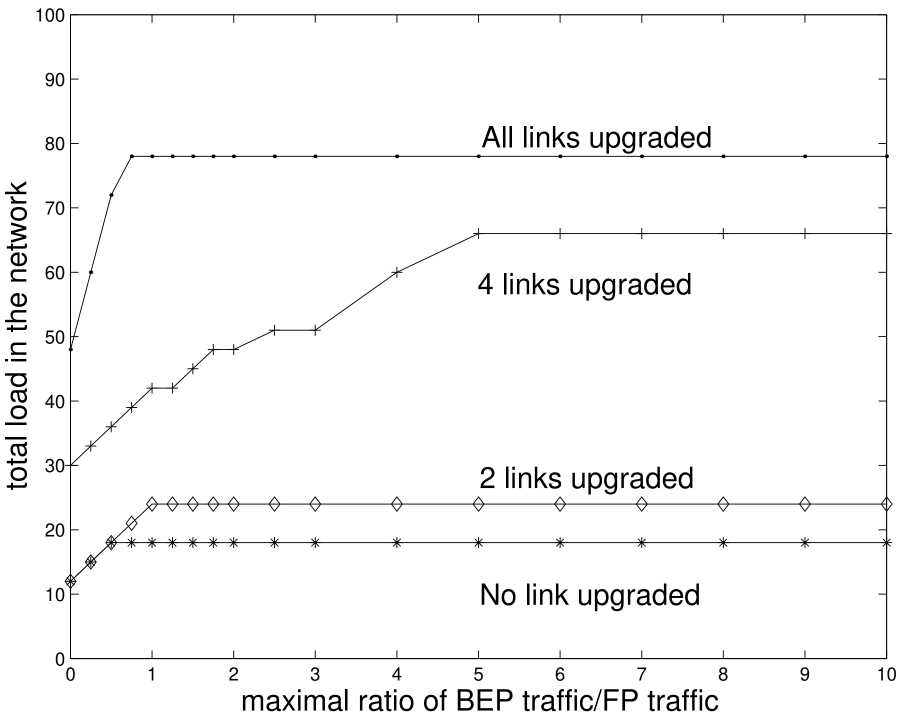

Figure 3: Total number of demands (total load) versus maximal ratio $\rho_{\max}$ of BEP traffic over FP traffic, when $\varepsilon = 0.5$, so that a fraction of 0.75 of the BEP traffic can be restored.

Further research should investigate these benefits for larger physical topologies, and meshed logical topologies. This approach should also be refined more generally (not just for ring topologies) so that the BEP traffic can secure some level of restoration at the IP layer.

Finally, if SONET is no longer the layer handling failures, and if optical cross-connects do not perform wavelength conversion, then MPLS may be needed to map IP traffic directly on the lightpaths [14, 15]. The MPLS protocol indeed offers a potential alternate mechanism for providing protection and restoration at layers 2/3. MPLS is a general purpose tunneling mechamisn that uses a simple label-swapping forwarding technique to transport IP packets across an IP network. It creates tunnels, called Label Switched Paths (LSPs) by distributing labels along a path of MPLS-capable routers. The LSP tunnel essentially sets up a path through a network of connectionless IP routers. MPLS is suited for survivability for a few reasons. LSPs can be used as backup paths and can be computed in advance. This requires storing extra labels in a forwarding table. Also, MPLS is not dependent upon IGP convergence since backup LSPs can be established a priori. Research in the performance of MPLS restoration mechanisms is still immature. However it is hypothesized that for link failures, link protection can occur within tens of milliseconds since no signalling is required. Yet path protection is expected to take on the order of seconds because this would require some signalling to inform the head of the tunnel about the topology change.

References

[1] Fraleigh, C., Moon, S. B., Diot, C., Lyles, B., and Tobagi, F.: Architecture of a Passive Monitoring System for backbone IP Networks. Sprint Technical Report TR00-ATL-101801 (2000).

[2] Arijs, P. , van Caenegem, P., Demmester, P., Lagasse, P., Van Parys, W., Achten, P.: Design of Ring and Mesh Based WDM Transport Networks, Optical Networks Magazine (2000) 25–40

[3] Gerstel, O., Ramswami, R.: "Optical Layer Survivability: A Services Perspective". IEEE Communication Magazine, **38**(3) (2000) 104–113.

[4] Sridharan, M., Somani, A. K.: "Revenue Maximization in Survivable WDM Networks." OPTICOMM 2000 (2000) Dallas.

[5] Ramamurthy R. and Mukherjee B., "Survivable WDM Mesh Networks". Proc. Infocom99 (1999) 744–751 New York.

[6] Doshi, B., Dravisa, S., Harshavadhana, Hauser O., Wang Y.: "Optical Network Design and Restoration." Bell Labs Technical Journal, **4**(1) (1999) 58–84.

[7] ITU-T G. 872, "Optical Transport Networks" (1999).

[8] Bonenfant, P., Rodriguez-Moral, A.: "Optical Data Networking". IEEE Communication Magazine, **38**(3) (2000) 63–70.

[9] Alferness, R.C., et al: "A Pratical Vision for Optical Transport Networking". Bell Labs Technical Journal, **4**(1) (1999) 3–18.

[10] Alaettinoglu, C., Jacobson, V., Yu, H.: Towards Milli-Second IGP Convergence, IETF Draft (ftp://ftp.isi.edu/internet-drafts/draft-alaettinoglu-ISIS-convergence-00.txt). (2000)

[11] http://www.ilog.com.

[12] Crochat, O., Le Boudec, J.-Y., Gerstel O: Protection Interoperability for WDM Optical Networks. IEEE/ACM Transactions on Networking **vol. 8** (2000) 384–395

[13] Modiano, E., Narula-Tam, A.: Survivable Routing of Logical Topologies in WDM Networks. Proc. Infocom2001 (2001) Anchorage.

[14] Awduche, D., et al: IETF Draft (ftp://ftp.isi.edu/internet-drafts/draft-awduche-mpls-te-optical-01.txt). (1999)

[15] Metz, C.: IP Protection and Restoration. IEEE Internet Computing (2000), 97–102.