Quality of Protection (QoP):

A Quantitative Unifying Paradigm to Protection Service Grades

Ori Gerstel and Galen Sasaki

Nortel Networks and the University of Hawaii. Email: ori@ieee.org, sasaki@spectra.eng.hawaii.edu

In this paper we provide a quantitative framework for best-effort protection of the optical layer. This framework allows to bridge the gap between two known protection grades of fully protected connections vis-a-vis unprotected protection. The framework allows to specify the probability with which the connection will be protected, providing the customer with a full range of protection guarantees at possibly different prices. Since connections may be partially protected, the required protection bandwidth can be reduced. The amount of protection bandwidth is shown to depend on an "equivalent survivable bandwidth." The framework also extends to preemptable (low priority) connections and to different ring architectures.

A. INTRODUCTION

Optical networks are emerging as the predominant transport layer technology for telecom service providers, replacing SONET/SDH in this role. As such, they are evolving from first-generation point-to-point systems that focused on efficient multiplexing using wavelength division multiplexing, to second-generation systems that provide more networking functions. In particular, fault-tolerance has been one of the main highlights of the SONET/SDH layer, and is now being added to optical layer equipment from optical add-drop multiplexers (OADMs) to optical cross-connects (OXCs).

The following protection classes[1] have been considered for ATM networks [VHS96] and later for the optical layer [GR00a]. In similarity to the concept of QoS they have been termed "Reliability of Service (RoS)" classes in [VHS96]:

(a) Guaranteed protection: the connection will be protected by the transport layer with very high likelihood (99.999% is typical),

(b) Best effort protection: using less protection bandwidth

(c) Unprotected traffic: the transport layer does not make an effort to protect the connection if a failure occurs, and

(d) Preemptable traffic: traffic that normally uses protection bandwidth for classes (a) and (b), and is preempted when the protection bandwidth is needed to protect against a failure. This class is termed "extra traffic" in SONET nomenclature.

While classes (a), (c) and (d) are well defined, and their implementation has been widely studied, e.g., in [RM99, GR00b], the grade of service for the best effort class (b) has never been quantified before. The only work that alludes to the need for a differentiated service model for the optical layer, and also addresses different grades of protection is [GNS00], but even it does not define what does, say, 40% protection really mean. Likewise, different service grades are also possible from class (d), but have not been addressed before systematically.

In this work we propose a unified paradigm that places all the above service classes on a continuous spectrum of protection grades (where different sub-ranges of grades map to a single protection class). This paradigm is based on assigning a guaranteed *quality of protection (or QoP)* to each connection. Upon a failure, the probability of a connection to survive the failure is determined by its QoP.

It should be noted that it is straightforward to assign different priorities to different connections and restore them based on their relative priority, but that this approach only provides a relative guarantee, which is insufficient if the service is priced and sold to a customer according to its grade. By contrast, our approach is based on an absolute guarantee and therefore provides a better option for a service level agreement (SLA).

Note that probabilistic guarantees addresses a weakness of relative priorities. Relative priorities do not address how best effort protection connections can be restored in a fair way. Higher priority connections always will be restored over lower priorities. If best effort protection connections were priced about the same then we would expect them to have the same chance of surviving. Then we should resort to something like probabilistic guarantees.

While the scheme is especially attractive for optical networks, given their emerging role as the protection layer of the transport network, it is equally viable for any connection-oriented network, such as MPLS or ATM.

We should also note that there is related work in [MS00]. The work focuses on D-connections, which are protected connections, but where the protection bandwidth can be used by working paths. Thus, D-connections are allowed

[1] We term these fixed protection alternatives "protection classes", while the continuous set of protection levels we propose below we term "protection grades" to make a distinction between the two approaches. (Thus a given range of protection grades may be mapped to a single protection class.

to become unprotected. Efficient routing algorithms and connection admission algorithms are given and simulated. However, the work does not provide a unified paradigm for protection covering the four classes of protection (guaranteed, best effort, unprotected, and preemtable) as we do in this paper. In addition, they do not have the same probabilistic guarantees that we introduce here.

In Section B, we provide motivation for our QoP paradigm. In Section C, we introduce the QoP framework using a simple two node network example. In addition, a measure for the amount of protection bandwidth required is given that is similar in concept to *equivalent bandwidth* for ATM networks, that is used to estimate how bandwidth is shared. Insights on how the protection may be implemented using probabilistic algorithms are discussed. An alternative to probabilistic protection is given in Section D referred to as the deterministic framework. This scheme is more appropriete for digital crossconnects whereas the probabilistic approach applies to all-optical cross-connects as well.

Some of the concepts in Section C are extended to ring networks in Section E. Ring topologies are important to optical networks since SONET/SDH are ring oriented. They are the prevalent topology deployed in existing transport networks. This section also provides insights into how to optimally route connections in the presence of mixed grades of protection. In particular, we have extensions for all survivable traffic excluding preemptable connections. However, under certain special cases, we can include preemptable connections and this is discussed in Section F, which is on extensions and related models.

In Section F, we also have a brief discussion of these results to mesh networks. Much future work must be done for mesh networks. However, the purpose of this paper is to introduce the QoP framework, which we believe may be the first unified paradigm to cover all four classes of protection in a seamless way.

Finally, a summary is given in Section G, including the many open problems that this paper instigates.

B. MOTIVATION

At present, SONET/SDH ring networks waste at least 50% of their bandwidth on protecting traffic. Since the protection is 100% guaranteed, the same amount of bandwidth that is allocated for working traffic along one side of the ring has to be reserved on the opposite side for protection. This waste is becoming unacceptable for many carriers, which is one of the main motivations for moving away from simple ring based protection, into more complex – yet more efficient – mesh protection schemes. Mesh protection reduces the amount of reserved protection bandwidth by sharing the same bandwidth among different working connections, as long as they are not likely to fail together. Thus, if connections A and B do not share any links and nodes along their working routes, they are highly unlikely to fail at the same time and can therefore share the same protection route.

Yet, even mesh protection schemes require considerable bandwidth on their protection path [IMG98, RM99] Furthermore, this waste grows as the network topology becomes sparser, as larger subnetworks resemble rings more and more. For example, if the mesh network is really a ring, then the same 50% protection bandwidth that SONET shared protection rings (e.g., SONET BLSR) need, is also required by mesh protection schemes.

Due to the cost structure of WDM, it is expected that the physical network topology will become sparser to take advantage of the relatively low cost of activating another channel on existing WDM links vis-à-vis installing a new link from scratch. This is due to the high upfront cost of optical (EDFA) amplifiers – especially for long-haul application, where many of them are needed along each link – and negligible added cost for turning up another channel once they are in place [S00]. This phenomenon has also been observed for non-WDM cost structures [DG00].

We shall see that our best-effort QoP framework allows to reduce the reserved bandwidth, even for rings, well below the mandatory 50% protection bandwidth and that the saving is proportional to the grade of protection required by the customers of the optical layer.

Another motivation for having a continuous range of protection grades (from the 100% protection traffic to the 0% protection, that unprotected traffic provides) is the fact that Internet traffic is often more sensitive to cost than to reliability. Furthermore, from real-life data obtained from a regional ISP during 1998, it is evident that most of the failures are at the IP layer and cannot be fixed by the optical layer [LAJ98]. This raises the following issue: why does the optical layer have to guarantee 99.999% service protection while the rest of the network components are quite far from this level of reliability? Our QoP scheme allows customers to gauge the grade of protection they are getting from the optical layer to the level of reliability that the rest of their network supports, while keeping the costs down.

Lastly, one can use low survivability lightpaths to construct a high-survivability network, as shown in the following example in Figure 1. In this figure, a 5 node full mesh IP network is depicted, whereby each link between routers uses an optical network lightpath. The probability of each lightpath to fail is high (0.1), but yet in order for the network to get disconnected at least 4 links have to fail, the likelihood of which is very low ($0.1^4 = 0.0001$ assuming the different lightpaths are independent).

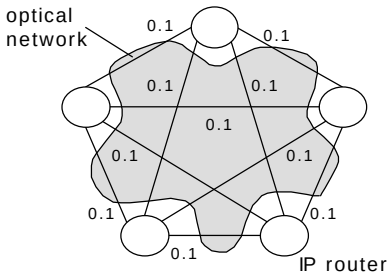


Figure 1: An IP network over an optical cloud

In the next section, we focus on a two node system, with one working fiber and one protection fiber. On this network we define our protection framework, starting with protected connections, and then preemptable connections.

C. The Quality of Protection Framework: The Two Node Case

We shall start with a few definitions. To illustrate these definitions, we refer to a two node network with two diverse parallel links *L1* and *L2* shown in *Figure 2*. We will first discuss a framework for protected, best-effort, and unprotected connections (lightpaths) only. Then we discuss preemptable connections.

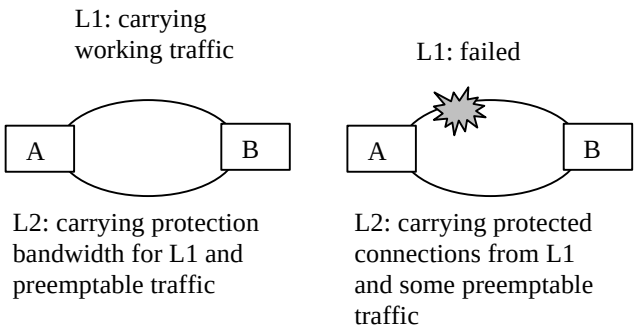


Figure 2: A two node network

For each connection *C* we associate a QoP grade $0 \le Q(C) \le 1$. If any link *L* along the working path of *C* fails, the *probability* that service will be immediately restored is *Q(C).* Note that these concepts are consistent with how the known protected and unprotected services behave*: Q(C) = 0* means that the connection *C* will not recover if a failure occurs (i.e., *unprotected*), and *Q(C)=1* means that *C* will recover (i.e., *guaranteed protection*). Any value between those two extremes is considered *best-effort protection*.

Assuming the connection *C* requires a bandwidth *B(C)*, define the *equivalent survivable bandwidth* (*ESB*) of the connection to be: $ESB(C) = Q(C) \cdot B(C)$. Note that this concept is similar to the concept of *equivalent bandwidth* for ATM networks in that it is used to estimate how bandwidth can be shared. In the rest of this paper we assume *B(C) = 1* for all connections *C*. Not only does this simplify the exposition, but also it is applicable to optical networks, were most connections typically run at a given fixed bandwidth (e.g., OC-48 or OC-192).

Also define the *working load* on a link *L*, *WL(L)* to be the number of working connections that use the link and its *equivalent survivable load*,

$$ESL(L) = \left\lceil \sum_{\text{Connection C uses link L}} ESB(C) \right\rceil,$$

to be the equivalent survivable bandwidth of all the connections that use the link *L*.

Theorem 1. *Consider the two node network of* Figure 2. *ESL(L1) is the necessary and sufficient amount of bandwidth on L2 required to ensure that each connection C with a working path on L1 is protected with probability Q(C) .*

Proof. We first show that ESL(L) is necessary. Recall that each connect *C* should have a probability *Q(C)* of recovering in case of a fault on *L1*. Then by definition $\sum_{\text{Connection C uses link L1}} ESB(C)$ is the average protection bandwidth used in Link *L2* in case of a fault on *L1* . Thus, the sum is a lower bound on the amount of protection bandwidth needed. Since bandwidth amounts are integer, *ESL(L1)* is also a lower bound.

To show sufficiency we will describe a scheme to randomly select a collection of connections that are protected in case of an *L1* fault. We denote by C_1, C_2,..., C_m the connections with working paths in *L1*, where m is the number of such connections. We let a binary *m*-

vector $b = (b_1, b_2,...,b_m)$ represent a collection of connections that are protected in case of a fault, where $b_k = 1$ if C_k is protected, and 0 otherwise. Note that for a binary m-vector b to be a proper representation, it must have at most $ESL(L1)$ ones. By Lemma A in the Appendix, we can find, for some n, a set of binary m-vectors $b(1), b(2),..., b(n)$ and a probabily vector $(f(1), f(2),...., f(n))$ with the following properties: (1) each $b(k)$ is a binary m-vector with at most $ESL(L1)$ ones, and (2) for each connection C_i,

$$Q(C_i) = \sum_k b_i(k)\ f(k)$$

The protection scheme is as follows. Whenever $L1$ incurs a fault, with probability $f(k)$ the connections corresponding to $b(k)$ are protected through $L2$. Notice that from property (1) only $ESL(L1)$ bandwidth is required for the protection. Also notice that from property (2) a connection C_i will be protected with probability $Q(C_i)$.

Q.E.D.

So far we have provided a framework to unify guaranteed protected connections, best effort connections and unprotected connections (protection grades (a)-(c) above). We refer to these connections as *survivable* connections since they should survive as long as their is no failure on their working paths.

We now extend the scheme to also support *preemptable* connections. Preemptable connections normally populate the protection bandwidth used to recover other survivable connections and are preempted when another connection requires the bandwidth. However, one should note that the order in which this traffic is preempted is unspecified. Therefore, it is possible to define a "prioritization order" which will define which of the preemptable connections is preempted first and which one is the most reliable of the preemptable connections, and will be the last to be disrupted.

For the sake of simplicity, we focus first on the two node network of *Figure 2* and assume that fiber $L1$ (resp. $L2$) carries $B1$ (resp. $B2$) channels. Suppose $L1$ carries protected traffic only, and $L2$ carries preemptable connections and free bandwidth to accommodate a failure on $L1$. Now suppose $L1$ has failed. If $B2 \geq WL(L2) + ESL(L1)$ then, by Theorem 1, the survivable connections on $L1$ will be protected on $L2$. However we are now interested in the case where $B2 < WL(L2) + ESL(L1)$. In this case, some of the connections in $L2$ must be preempted to make way for survivable connections on $L1$. Lets us first consider a straightforward extension of the best-effort model and see where the approach fails. Then we shall explain a better approach that does not suffer from these shortcomings.

1) The naive approach

There is no clear distinction between best-effort and preemptable traffic and the definitions from before are still valid. Each connection has its $Q(C)$, and in addition the maximum total number of channels $B2$ is given, such that $L2$ cannot protect all connections. As a result, both the disrupted survivable connections on link $L1$ (in need for protection) and the preemptable connections on $L2$ contend for the same $B2$ channels. Those connections on $L2$ with a lower QoP grade $Q(C)$ than that of the disrupted connections on $L1$ will be preempted and thus can be considered preemptable connections.

While this approach does provide a uniform framework for all protection classes, it has several shortcomings:

There is no fixed mapping between the QoP $Q(C)$ and the protection class: a connection on link $L2$ can be non-preempted (i.e., survivable) or preempted depending on the protection grades for other connections. Thus its class dynamically changes depending on its relative order amongst the connections on links $L1$ and $L2$.

An even more important issue concerns the definition of $Q(C)$: while it is well defined for the surviving connections, it provides no probabilistic guarantee for the preempted connections. In other words, $Q(C)$ becomes just a relative priority among the connections and not a (more tangible) protection guarantee.

For these reasons we turn to the following less intuitive approach which avoids these shortcomings.

2) The improved approach

We extend the range of QoP grades $Q(C)$ to be $-1 \leq Q(C) \leq 1$ to map as follows to the different protection classes:

Protection Service	Protection Grade
(a) guaranteed	$Q(C) = 1$
(b) best effort	$0 < Q(C) < 1$
(c) unprotected	$Q(C) = 0$
(d) preemptable	$-1 < Q(C) < 0$
(e) unused channel	$Q(C) = -1$

Thus, $Q(C) \geq 0$ means that the connection is survivable, while $Q(C)<0$ means that the connection is preemptable. The QoP grade $Q(C)$ is mapped into protection guarantees as follows:

The probability that a connection *C* will survive a failure on its working path is at least *SP(C)*, which is defined to be $SP(C) = Max\{Q(C),0\}$. *SP(C)* will be referred to as the *survivability probability*. Note that if *SP(C)* > 0 then the connection is a survivable connection.

The probability that a connection *C* will be preempted when there is a fault that is *not* on its working path is at most *PP(C)*, which is defined to be $PP(C) = Max\{-Q(C),0\}$. *PP(C)* will be referred to as the *preemptable probability*. Note that if *PP(C) > 0* then the connection is a preemptable connection.

We say that a a QoP grade *Q(C)* for a connection *C* is *valid* if it survives according to *SP(C)* and is preempted according to *PP(C)* when a fault occurs.

Next, we define the *equivalent preemptable bandwidth* (*EPB*) to be: $EPB(C) = PP(C) \cdot B(C)$. (Again, we make the assumption that *B(C)* = 1.) This is analagous to the equivalent survivable load for survivable connections defined earlier. We also define the *equivalent preemptable load* (*EPL*) to be

$$EPL(L) = \left\lfloor \sum_{\text{Connection C uses link L}} EPB(C) \right\rfloor.$$

The next theorem extends Theorem 1 to make use of preemptable connections for the protection of survivable connections.

Theorem 2. *Consider the two node network of Figure 2 and a failure on L1 A sufficient condition to insure that the QoP grades for each connection is valid for the failure is*

$$ESL(L1) \le EPL(L2)$$

Proof. From Theorem 1, we know that there is a scheme to such that *ESL(L1)* bandwidth on *L2* can protect survivable connections on *L1* according to their QoP grades. Let us refer to this as scheme 1. Using similar arguments in the proof of Theorem 2, we can define a scheme, referred to as scheme 2, such that if *EPL(L2)* bandwidth is preempted on *L2* then each preemptable connection *C* on *L2* is preempted with probability *PP(C)*. These two schemes can be combined to imply the theorem. In particular, upon a failure of link *L1*, the following algorithm is used: *EPL(L2)* preemptable connections on *L2* are preempted according to the scheme 2. This frees at least $ESL(L1) \le EPL(L2)$ bandwidth. on *L2*. *ESL(L1)* survivable connections on *L1* are chosen bo be protected on *L2* according to scheme 1.

Q.E.D.

D. A DETERMINISTIC QOP FRAMEWORK

In the above probabilistic scheme, survivable connections *share* protection bandwidth by using randomization. This is the only bandwidth sharing possible for a transparent truly optical network that carries different signal formats and is not even aware of their exact bit rates, since either a connection is completely recovered or not recovered at all.

Other optical networks (termed *opaque* networks) do have access to the carried formats, and can even multiplex/demultiplex them onto a single wavelength using time division multiplexing (TDM) techniques. Several electro-optical crossconnect vendors are boasting this capability (typically SONET OC-192 and OC-48 signals are broken into their consituent STS-1 level signals and are crossconnected at this level before being multiplexed back into OC-n signals).

This allows for a deterministic QoP model, whereby upon a failure, each survivable connection is guaranteed to have a deterministic *reduced protection bandwidth* $RPB(C) = SP(C) \cdot B(C)$. Note that here *SP(C)* is no longer the probability of protection the connection, but the reduced bandwidth available for the connection, assuming the protection connections are electrically multiplexed together. However, also note that *RPB(C)* is exactly equal to *ESL(C)*.

Likewise, upon a failure, each preemptable connection is guaranteed to have at most a *reducted working bandwidth* $RWB(C) = PP(C) \cdot B(C)$. Here, *PP(C)* is no longer the probability of preempting a connection, but the reduction of working bandwidth for the connection if a failure occurs. Note that *RWB(C)* is exactly equal to *EPL(C)*.

To illustrate these concepts we give an example of both the probabilistic and deterministic QoP frameworks.

Example. Consider the two node network in Figure 1. We have the following connections on *L1*: *C1* with *Q=0.5*, *C2* with *Q=0.5*, *C3* with *Q=0.25*, and *C4* with *Q=0.25*. On *L2* the following connections exist: *C5* with *Q= -0.5* and *C6* with *Q= -0.5*.

We also assume that there is a single unused channel on *L2*. Thus, the number channels on *L1* is *B1* = 4 (*C1* + *C2* + *C3* + *C4*), and the number of channels on *L2* is *B2* = 3 (*C5* + *C6* + *unused*). Notice that *C1*, *C2*, *C3*, and *C4* are survivable connections on *L1*, and *C5*, *C6*, and the unused channel are preemptable connections on *L2*. Also notice that *ESL(L1)* = 2 and *EPL(L2)* = 2.

The following is an example of how a probabilistic protection service may be implemented. Upon a failure on *L1*, with probability 0.5, {*C1, C3*} are protected and *C5* is preempted; and with probability 0.5, {*C2, C4*} are protected and *C6* is preempted. In both instances, two connections on *L1* are protected are protected through *L2* by using a preempted connection and the free channel. Also notice that *C3* and *C4* are given better protection (higher protection probabilities) than their prescribed QoP grades.

The next example is one of a deterministic protection service. Upon a failure on *L1*, both *C5* and *C6* reduce their bandwidth to 0.5. This increases the free bandwidth on *L2* to 2. Then *C1* and *C2* each send 0.5 bandwidth on the freed up bandwidth. *C3* and *C4* each send 0.25 bandwidth each on link *L2* on the unused channel. (Note that in this case, we could also have *C3* and *C4* restore more bandwidth than their prescribed QoP grade.)

The rest of the paper will focus on the probabilistic protection. It should be noted, however, that all the results for the probabilistic protection below can be translated to the deterministic protection. The reverse is not true, and some results are true only for the deterministic model. These cases are explicitly marked as such.

E. RING NETWORKS

SONET/SDH ring networks play a central role in the current telecommunications infrastructure, and have been widely considered for the optical layer as well. In this section we extend the definitions and results from the two node network to a general ring. We will simplify the discussion by focusing on survivable connections (i.e., $Q(C) \geq 0$) and ignoring preemptable ones. We have some results for preemptable connections and these are left in Section F. Even without preemptable connections, the results are nontrivial. For example, the type of protection on the ring (line or path protection) has bearings on the efficiency to the protection and the amount of bandwidth needed per link, which was not the case for the two node ring.

In the rest of this section we focus on how to realize the QoP grade assuming only survivable connections. We also assume that the ring has the same number of channels *Bmax* per link. We consider three typical protection schemes in a ring network:

Optical layer terminology (ITU)[2]	Sharing	Protected entity
OMS-SPRing	Shared	Line
OCh-DPRing	Dedicated	Path
OCh-SPRing	Shared	Path

These networks are graphically depicted in *Figure 3*, Figure 6, and Figure 7.

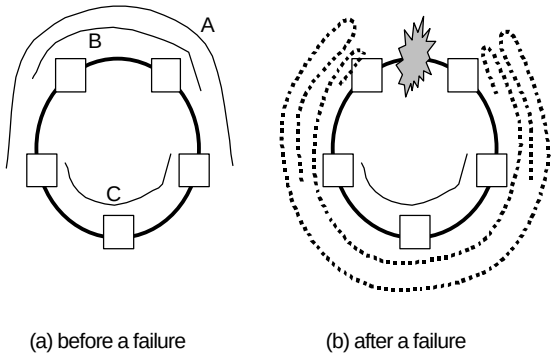


Figure 3: Line protection

E.1. Line Protection in Ring Networks

Line protection is based on switching all the connections that use a failed link together, around the ring. This approach has been proposed for the optical layer and has been termed "OMS-SPRing" by the standard bodies. It allows for substantial equipment saving if implemented optically [GR00b].

An example for the behavior of the scheme can be found in ks are graphically depicted in *Figure 3*, Figure 6, and Figure 7.

Figure 3, where a 5 node ring is depicted supporting 3 connections (*A*, *B* and *C*). Upon a failure of a link, the affected connections (*A* and *B*) are looped back at the nodes adjacent to the failure using dedicated protection bandwidth, while connection *C* remains intact. Notice that the protection path of *A* takes a long route that goes to the failed link rather than take a direct route. This type of inefficient routing is sometimes referred to as "backhauling".

With guaranteed connections, it is obvious that one needs to reserve half of the bandwidth around the ring for

[2] Similar terminology has been used by ITU for SDH networks as well. Optical multiplex section (or OMS) refers to the line layer, whereas Optical channel (OCh) refers to the path layer. SPRing stands for Shared Protection Ring and DPRing for Dedicated Rrotection Ring.

protection purposes, since if 2 links support more than 50% working connections, the failure of one of them will imply that the other link does not have enough bandwidth to protect the failed connections. With our QoP scheme, the amount of protection bandwidth needed to protect against a failure of link *L* is only *ESL(L)* – as proven in the following theorem.

Lemma 3. *For each link L in the ring, ESL(L) is the amount of bandwidth required on each of the other links to ensure that each connection C on link L is protected with probability SP(C) if L fails.*

Proof. *Consider the survivable connections using link* L. *We can use the scheme in the proof of Theorem 1 to randomly choose* ESL(L) *of them to be protected in case* L *fails so that each connection* C *is protected with probability at least* SP(C). *Thus,* ESL(L) *bandwidth on*

the other links is sufficient bandwidth to protect the randomly chosen connections.

Q.E.D.

Let us now calculate the capacity (number of channels) needed for a link *L*, *B(L)*. The number of working channels per link is *WL(L)*, and the number of protection channels needed on link *L* to protect against a failure of any other link *L'* is $\max_{L' \neq L} ESL(L')$. Thus,

$$B(L) = WL(L) + \max_{L' \neq L} ESL(L').$$

Note that some connections *C* contribute twice to some links: once as a working channel (1 unit of traffic) and again as a protection path, but then only contributing *ESB(C)*. For example, in Figure 3, connection A contributes 1 unit of traffic to link *L1* and upon the fault, it also contributes an additional *ESB(L)* as its protection path backhauls through *L1* again.

The amount of bandwidth required is

$$Bmax = \max_{L} B(L) = \max_{L} \left\{ WL(L) + \max_{L' \neq L} ESL(L') \right\}$$

We have a simpler upper bound for *Bmax*,

$$Bmax^* = \max_{L} WL(L) + \max_{L} ESL(L)$$

Note that *Bmax* = *Bmax** if there is more than one link *L* that maximizes $\max_{L' \neq L} ESL(L)$. Since in that case, for all links *L*, $\max_{L' \neq L} ESL(L) = \max_{L'} ESL(L)$.

Lemma 3 implies the following theorem.

Theorem 4. *Consider a ring network that is supporting survivable connections only. Suppose the survivable connections have been routed. A sufficient amount of bandwidth to provide protection according to the QoP grades is Bmax per link.*

Next we consider the problem of routing connections to minimize *Bmax*. This seems to be a difficult problem, so one may consider minimizing the simpler *Bmax** instead. The next theorem addresses this optimization problem assuming all connections have the same *EPB(C)*. The theorem uses the following definition:

$$WLmax = \max_{L} WL(L)$$

Theorem 5. *Consider a ring network with only survivable connections that are not routed yet. Suppose EPB(C)=Q for all connections C. The problem of routing the connections to minimize Bmax* is equivalent to the problem of minimizing Lmax.*

Outline of Proof. Note that *Bmax** is composed of two terms: *Lmax* and $\max_{L} ESL(L)$. Note that a routing that minimizes *Lmax* will minimize the maximum number of connections through any link. This is precisely the routing that will minimize $\max_{L} ESL(L)$ since all the QoP grades are the same. Since a routing that minimizes *Lmax* will also minimize the two terms of *Bmax**, it must also minimize *Bmax**.

Q.E.D.

This problem has been studied extensively in the context of circuit-switching and optimal solutions for it exist in ring networks [FNSST92, VSKW96].

When *EPB(C)* is different for different links, this equivalence does not hold, as shown by the following theorem. In addition, the theorem considers the maximum weighted load *Wmax*, where the weight for connection *C* is *1+Q(C)*, which may (wrongly) seem as an approximation of *Bmax**.

Theorem 6. Given a ring and a set of connection endpoints, with *Q=0* or *Q=1*. The following three design problems have conflicting optimization goals:

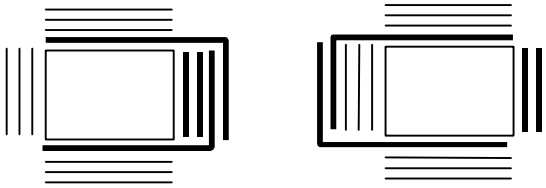
a) Find a routing that minimizes the maximum load *Lmax* on any link,

b) Find a routing that minimizes the maximum weighted load *Wmax* on any link, and

c) Find a routing that minimizes the maximum capacity *Bmax** on any link.

Proof. Consider two examples, one demonstrating that any configuration that minimizes *Lmax* on a given network does not minimize *Wmax* or *Bmax** in the network. The other example shows that minimizing *Wmax* does not minimize *Bmax**.

First, consider the example in Figure 4, in which a 4 node ring is depicted and 13 connections, out of which 4 have *Q=1* (depicted in bold) and 9 have *Q=0*.

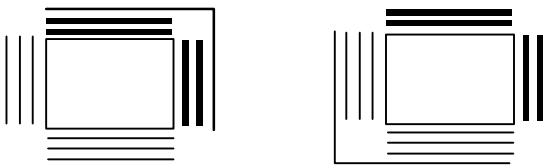


(a) optimal Lmax:
Lmax=4, Bmax*=8
Wmax=8

(b) optimal Bmax* & Wmax:
Lmax=5, Bmax*=7
Wmax=7

Figure 4: Lmax is not mutually achievable with Bmax and Wmax*

In part (a), the load is minimized (*Lmax = 4*). If any of the connections is rerouted around the ring, the load is increased to 5. However, to minimize *Bmax** and *Wmax*, such a rerouting is needed as shown in part (b) of the figure. One can exhaustively check that no other configuration optimizes both *Lmax* and *Bmax**, or both *Lmax* and *Wmax*.



(a) Optimal Bmax*:
Bmax*=5, Wmax=5

(b) Optimal Wmax:
Bmax*=6, Wmax=4

Figure 5: Bmax is not mutually achievable with Wmax*

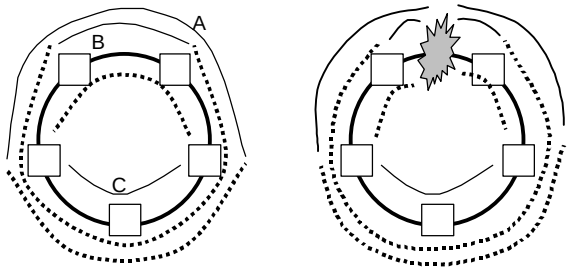
Now consider the example in Figure 5, in which 11 connections exist, 4 of which have *Q=1*and the rest have *Q=0*. In this example part(a) optimizes *Bmax** while part (b) optimizes *Wmax*. One can exhaustively check that no other configuration optimizes both metrices.

Q.E.D.

Observation. *Note that in the above examples Q=0 or Q=1. Thus they pertain to the known problem of mixing fully protected and unprotected traffic. It is interesting to note that the bandwidth optimization problem in this simple case(Bmax*) is still different from the Lmax solution with is used for fully protected SONET BLSR rings.*

E.2. Dedicated Path Protection in Ring Networks

Dedicated path protection has been widely deployed in SONET networks, where it is termed "unidirectional path switched ring" (or UPSR for short). The main advantage of this scheme is its simplicity, steming from the fact that no coordination is needed between the endpoints of a connection in order to switch to a protect path. Each node transmits two copies of the data in both clockwise and counterclockwise directions and it is totally up to the receiving node to decide which copy to choose.



(a) before a failure

(b) after a failure

Figure 6: Dedicated path protection

Under the QoP paradigm, deterministic protection makes more sense than probabilistic protection. With determistic protection, a survivable connection *C* will have a protection path that is running in parallel with the working path though at a lower data rate *ESL(C)*. With probabilistic protection, some survivable connections will have no protection paths.

With deterministic protection, the data flows of a survivable *C* must have its bit rate reduced to fit into the protection path. Packets (or tributary flows) within a connections data flow must be dropped, perhaps at random or by priorities, if such priorities exist. This protection scheme relies on flow control at higher layer protocols. Then a connection's data flow will be reduced to fit into the protection path.

Notice that each connection *C* contributes a bandwidth of 1 along its working path and a bandwidth of *EPB(C)* along its protection path. This is equivalent to saying that *C* contributes a bandwidth *EPB(C)* along every link in the ring, and an addition, *1-EPB(C)* along the working path. Therefore we can ignore the former contribution for all connections, and focus the optimization problem on the latter contribution, arriving at the following optimization problem.

Equivalent optimization problem. *Given a set of connections on a ring network, find a routing for them that minimizes the maximum "weighted load"(Wmax) over all links, where each connection C contributes 1-EPB(C) to the load of each link along its route.*

While this problem is known to be NP-Complete [CS94], good heuristics to it are known (e.g., [VSKW96]).

E.3. Shared Path Protection in Ring Networks

Shared path protection is based on switching each of the failed connections individually, along the shortest alternate route around the ring, as shown in Figure 7. Thus the "back hauling" effect for line protection is avoided. For example, in Figure 3 the working path of connection *A* for line protection is longer than the working path of the same connection using shared path protection in Figure 7. Consequently, the amount of protection bandwidth needed along the different links is different.

We proceed to compute the amount of protection bandwidth required for survivable connections. Let us determine the amount of bandwidth *B(L)* required for a link *L*. *B(L)* should include *WL(L)*. It should also have protection bandwidth when other links fail. The amount of protection bandwidth required if link *L'* fails is

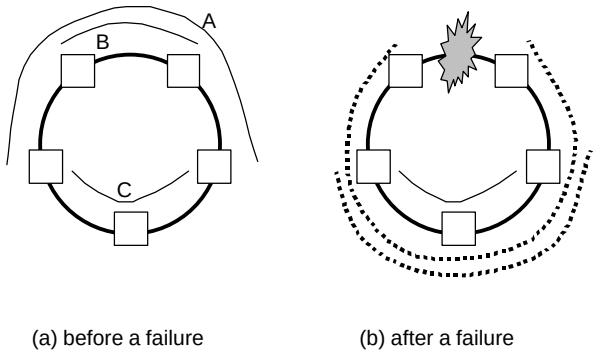
$$\sum_{C \text{ traverses } L' \text{ but not } L} ESB(C)$$

Thus, the amount of protection bandwidth required for link L is

$$\max_{L' \neq L} \sum_{C \text{ traverses } L' \text{ but not } L} ESB(C)$$

$$= \max_{L'} \sum_{C \text{ traverses } L' \text{ but not } L} ESB(C)$$

Let the last maximum be denoted by *PathPB(L)* (i.e., *path protection bandwidth* at link *L*). Notice that the summation in *PathPB(L)* is a "weighted load" of all the connections through a link *L'* excluding connections through link *L*, where the "weight" of a connection *C* is the equivalent survivable bandwidth *ESB(C)*. Thus, *PathPB(L)* is the maximum weighted load over all links excluding the connections that go through *L*.



(a) before a failure (b) after a failure

Figure 7: Shared path protection

Thus, the amount of bandwidth required on link L is *WL(L)* + *PathPB(L)*. This implies the following theorem.

Theorem 7. *Consider a ring network that is supporting survivable connections only. Suppose the connections have been routed. A necessary amount of bandwidth to provide shared path protection according to the QoP grades is Bmax =* $\max_{L}(WL(L) + PathPB(L))$

It is straightforward to prove that this condition is also a sufficient condition in the deterministic case.

Theorem 8. *Under the conditions of Theorem 7, Bmax is sufficient to realaize the QoP for a set of survivable connections in the deterministic model.*

The implementation of a probabilistic scheme to realize the QoP grade of service with only *Bmax* bandwidth per link is for future study.

An interesting question is whether shared path protection provides any gain over the simpler shared line protection. The following example shows that in a ring with 3 connections, line protection may require 6 units of bandwith in the worst case, whereas path protection only requires 3 units.

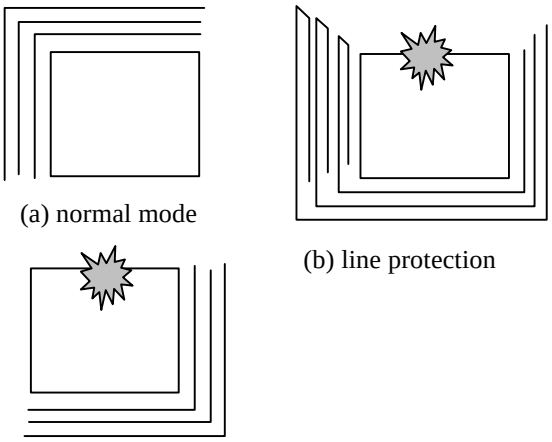


Figure 8: Path versus line protection

F. Extensions and Related Models

F.1. Preemptable Connection in Rings

So far, our results for ring networks do not include preemptable connections. We have the following theorem when preemptable connections are included if their QoP grades *Q(C)* are the same value.

Theorem 9. *Consider a line protection ring network. Suppose there is a value q > 0, such that each preemptable connection C has QoP grade Q(C) = - q. Then a sufficient condition to insure that the QoP grades are valid for each connection is for all pairs of links L and L': $ESL(L) \le EPL(L')$.*

Outline of Proof. Consider a link *L*. Now consider all preemptable connections. We can partition these connections into "strings", where a string is a sequence of preemptable connections starting from one end of *L* and ending at the other end of *L*. The sequence is such that the end of one connection is the beginning of another. These strings can be created in a "greedy" way by starting at one end of *L* and following connections around the ring until the other end of *L* is reached. During the traversal of the ring, whenever a connection of a string terminates, a new connection is added to the string. Such a greedy method will result in *ESL(L)/q* strings since on each link *L'* there are *EPL(L')/q* ≥ *ESL(L)/q* preemptable connections. Note that if we preempt each string with probability *q* whenever there is a fault on *L*, then the QoP grade will be valid.

Now we have survivable connections on *L* with equivalent survivable load *ESL(L)*. We also have *ESL(L)/q* strings on the rest of the ring, each will be interrupted with probability *q*. This is basically the same situation as the two node network, where our *L* is the *L1* in the two node network, and the rest of the ring is *L2* in the two node network. Theorem 8 then follows from Theorem 2.

Q.E.D.

F.2. Mesh Networks

We will discuss at least one difficulty in dealing with mesh networks. One of the advantages of our QoP paradigm is that it has a simple notion of equivalent survivable bandwidth. This bandwidth turns out to be an efficient amount of protection bandwidth for the two node network and ring with line protection. However, for mesh networks, the equivalent protection bandwidth may provide a necessary bandwidth requirement that is much smaller than what is sufficient. This is true for the probabilistic protection scheme as illustrated by the following example.

Example. Consider three survivable connections *C1*, *C2*, and *C3* in Figure 8. Their QoP grades are 0.5. Their working paths are shown as solid lines and their protection paths are dotted lines. Notice that their working paths meet pair-wise at links *L1*, *L2*, and *L3*, and all their protection paths meet at *L4*. Note that any link fault at *L1*, *L2*, and *L3* leads to an equivalent survivable bandwidth of 0.5 + 0.5 = 1 at link *L4*. This implies that only one channel is needed at *L4* to protect *C*1, *C2*, and *C3*. Yet, only one of the three can be recovered if there is only a single protection channel at *L4*, since the recovery of each one of the channels will prohibit the other two from being recovered. Thus, *L4* may require 2 channels to properly protect the connections. This requirement is twice the equivalent survivable bandwidth.

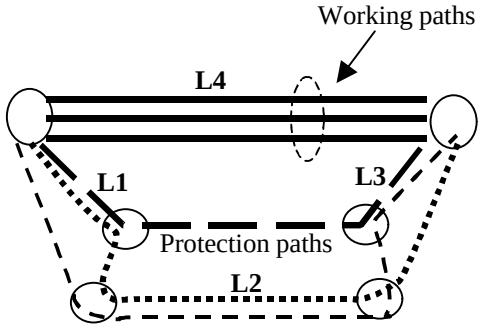


Figure 9: Equivalent protection bandwidth is insufficient for mesh networks

Finally, note that the deterministic protection scheme does not have these problems.

G. SUMMARY

In this paper we have suggested a new paradigm for guaranteed quality of protection. We have shown how the paradigm provides absolute guarantees for all classes of protection, including best-effort and preemptable (low-priority) connections. We have also shown how the paradigm extends to ring networks based on different protection schemes and pointed out some of the challenges of extending the paradigm to general mesh networks. This paper provides insights into this problem, most notably into the simpler and well-known case of a mixture of protected and unprotected connections – that is relevant even if additional grades of protection are not adopted.

Much work is still needed in defining efficient algorithms for choosing which survivable connections to protect and which preemptable connections to drop in rings networks. In addition, many network optimization problems for rings remain open for further research. Beyond rings, challenges exist even in redefining a sufficient equivalent protection bandwidth.

H. REFERENCES

[CS94] S. Cosares and I. Saniee, "An optimization problem to balancing loads on SONET rings," *Telecommunication Systems*, vol. 3, pp. 165-181.

[DDHHW99] B.T. Doshi, S. Dravida, P. Harshavardhana, O. Hauser, and Y. Wang, "Optical network design and restoration," *Bell Labs Technical Journal*, pp. 58-84, Jan-Mar 1999.

[DG00] J. Doucette and W.D. Grover, "Cost and topology-optimized mesh architectures for survivable WDM networks," *J. Selected Areas in Communications*, special issue on "Protocols and Architectures for Next Generation Optical WDM Networks," 4Q2000.

[FNSST92] A. Frank, T. Nishizeki, N. Saito, H. Suzuki, and E. Tardos, "Algorithms for routing around a rectangle," *Discrete Applied Math*, vol. 40, pp. 363-378, North Holland, 1992.

[GNS00] N. Bolmie, T.D. Ndousse, and D.H. Su, "A differentiated optical service for WDM networks," *IEEE Communications Magazine*, pp. 68-73, Feb. 2000.

[GR00a] O. Gerstel and R. Ramaswami, "Optical layer survivability – a service perspective," *IEEE Communications Magazine*, March 2000.

[GR00b] O. Gerstel and R. Ramaswami, "Optical layer survivability – an implementation perspective," *J. Selected Areas on Communications*, special issue on "Protocols and Architectures for Next Generation Optical WDM Networks," 4Q2000.

[IMG98] R.R. Iraschko, M.H. MacGregor and W.D. Grover, "Optimal capacity placement for path restoration in STM or ATM mesh-survivable networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 3, pp. 325336, June 1998.

[LAJ98] C. Labovitz, A. Ahuja, and F. Jahanian, "Experimental study of Internet stability and wide-area backbone failures," Technical report CSE-TR-382-98, U. Michigan, 1998. url:www.eecs.umich.edu/techreports/CSE/1998/CSE-TR-382-98.pdf

[MS00] G. Mohan and A. Somani, "Routing dependable connections with specified failure restoration guarantees in WDM networks," IEEE Infocom 2000, Tel-Aviv, March 2000.

[P97] V. Paxson, "End-to-end routing behavior in the Internet," *IEEE/ACM Transactions on Networking*, vol. 5, no. 5, pp. 601-615, Oct. 1997.

[RM99] S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks," *IEEE Infocom 99*, March 1999.

[S00] A.M. Saleh, "Transparent optical networking for the backbonenetwork," *Proc. OFC '00*, Paper ThD7, March 2000.

[VSKW96] R. Vachani, A. Shulman, P. Kubat, and J. Ward, "multicommodity flows in ring networks," *INFORMS Journal of Computing*, vol. 8, no. 3, pp. 235-242, 1996.

[VHS96] P. Veitch, I. Hawker, and G. Smith, "Administration of restorable virtual path mesh networks," *IEEE Communications Magazine*, December 1996.

I. APPENDIX

Lemma A. *Suppose we are given probabilities: Q(1), Q(2),..., Q(m) for some m. Let ESL(L1) denote* $\left\lceil \sum_k Q(k) \right\rceil$. *Then there is*

- *an integer n*
- *a collection of n binary m-vectors b(1), b(2),..., b(n) (Thus, b(k) = ($b_1(k)$, $b_2(k)$,..., $b_m(k)$)*
- *a probability vector (f(1), f(2), ..., f(n))*

with the following properties:

- *Each binary vector b(k) has at most ESL(L1) ones (and thus, it has m-ESL(L1) zeros).*
- *For each i, Q(i) =* $\sum_k f(k) b_i(k)$

Proof . Note that m denotes a collection of best-effort protection connections that share protection bandwidth, and in particular there are ESL(L1) channels of protection bandwidth. Each vector b(k) denotes a subset of ESL(L1) connections that can use the protection bandwidth simultaneously. Those connections that survive are denoted by "1" in the vector.

The probability of the configuration b(k) being chosen is f(k). Thus, $\sum_k f(k) b_i(k)$ is the probability that connection i is chosen to survive. Thus, we would want the sum to equal to Q(i).

To prove the lemma, we will show how to construct the binary vectors b() and the probability vector (f(1),...,f(n)). First, we construct an interval I* that starts from 0. This interval I* consists of m subintervals, where the kth subinterval corresponds to connection k and has length Q(k). We will refer to the kth subinterval by I(k). Note that I* will have length approximately equal to ESL(L1), but it may be smaller. If it is smaller, we extend it using a "dummy" subinterval so that it has length exactly equal to ESL(L1).

For each real value t that lands in interval I*, let H(t) denote the subinterval it lands in. Thus, if t lands in

subinterval I(k) then H(t) = k. Now we divide I* into subintervals of unit length. Note that there will be ESL(L1) of them. Refer to these intervals by J(0), J(1),..., J(ESL(L1)-1). Let U denote a uniform random variable distributed over the unit interval. Note that U is uniformly distributed over J(1). Therefore, H(U) is a random variable whose values are one of the best-effort protected connections. The probability of {H(U) = k} is equal to the length of subinterval I(k) in J(1). Similarly, for each integer r < ESL(L1), U+r is a random variable uniformly distributed over J(r). In addition, H(U+r) is a random variable whose values are from the best-effort protected connections. The probability of {H(U+r) = k} is equal to the length of subinterval I(k) in J(r).

Now, let h(0:U) = H(U+0), h(1:U) = H(U+1),..., h(ESL(L1)-1:U) = H(U+ESL(L1)-1). Then (h(0:U), h(1:U), ... h(ESL(L1)-1:U) is a random vector. Each best-effort protected connection k will appear in the vector with probability Q(k) because the likelihood of k appearing is equal to the length of subinterval I(k). We should note that a connection k will appear in the vector in at most one of the elements. To see why this is so, consider if it were not true. Then it must be that I(k) appears in two intervals J(x) and J(x+1) for some x, and there is a value y such that y is in $J(x) \cap I(k)$, and y+1 is in $J(x+1) \cap I(k)$. But this is impossible since I(k) has length less than one, so we cannot have it contain both y and y+1.

Now consider the vector (h(0:t), h(1:t), ... h(ESL(L1)-1:t)). Note that as t sweeps from 0 to 1, the vector changes at a finite number of points, and between consecutive points it is constant. We can divide the unit interval into a finite number of subintervals so that within a subinterval the vector is constant. We refer to these subintervals as f-subintervals. Let n denote the number of f-subintervals, and let f(k) denote the length of the kth f-subinterval. Let b(k) be a binary vector that represents those best-effort protected connections corresponding to the kth f-subinterval. These are exactly the f(k) and b(k) values we are looking for. Each b(k) corresponds to a value for the vector (h(0:t), h(1:t), ... h(ESL(L1)-1:t)). Thus, it is a binary m-vector with at most ESL(L1) ones. Note that $\sum_k f(k) b_i(k)$ equals the length of all the f-subintervals where $b_i(k)$ equals 1. Note that when equals $b_i(k)$ equals 1 it corresponds to some element in vector (h(0:t), h(1:t), ... h(ESL(L1)-1:t)) being equal to i. This means that t, t+1, t+2..., or t+ESL(L1)-1 must in I(i). Thus $\sum_k f(k) b_i(k)$ is equal to the length of I(i), which in turn is equal to Q(i).