Can DiffServ Guarantee IP QoS Under Failures?

Brunilde Sansò and Christian Awad, Ecole Polytechnique de Montréal
André Girard, INRS-EMT

Abstract

As ISPs and telcos converge toward all-IP network infrastructures, the problem of service reliability becomes more acute. In this article, we investigate to what extent DiffServ can provide service protection against optical failures in IP over optical (WDM) networks. For this, we propose two mechanisms: a DiffServ/WDM method and a WDM differentiated protection technique that we call DiffProtect. Results show that the priority traffic is equally well protected by both techniques. However, for medium-priority traffic, there is an improvement in delay and jitter under failure when DiffServ is used. Thus, the proposed DiffServ implementation can be used as an affordable and effective fault management technique to protect high-priority traffic with little delay.

We are currently witnessing a radical change of telecommunications where all services will be offered on an IP platform. Internet service providers (ISP) increasingly want to sell services with quality of service (QoS) guarantees, and in that context reliability, defined as the ability to sustain that QoS in the event of failures, becomes a very important issue.

Current IP networks have a number of traffic, or logical layer, mechanisms that can adapt to changing network conditions: adaptive routing like open shortest path first (OSPF), congestion control like random early detection (RED), and service differentiation like integrated services (IntServ), differentiated services (DiffServ), and multiprotocol label switching (MPLS). Although these mechanisms were not designed with reliability in mind, they could in principle be used to maintain some level of QoS in case of failures. In this article we concentrate on DiffServ, which is a popular, scalable, low-complexity alternative for supporting QoS in IP networks.We evaluate how and to what extent it can enhance reliability and whether this is enough to maintain a suitable QoS at the IP layer in case of failure of transmission equipment. If it does, it could provide significant savings for ISPs who would not have to pay extra for protected transmissions systems but could use the mechanisms already in place to provide a reliable service. For this reason, we first propose a DiffServ/wavelength-division multiplexing (WDM) mechanism and then compare the protection level that it can offer with that provided by another scheme based on differentiation in the optical layer that we called DiffProtect.

We have divided this article as follows. A brief review of the state of the art in multilayer protection is presented. We devote a section to the two models used for comparison: Diffserv and DiffProtect. The simulation modeling is presented as well as performance and simulation results. Finally, we provide conclusions and recommendations for further work.

Related Work

The area of protection and restoration for synchronous digital hierarchy (SDH) transmission systems has been maturing for many years [1–3]. The traffic carried in these networks was made up mostly of telephone calls and there was no need to even consider differentiated protection. The emphasis was placed on restoring whole transmission systems fast enough that the telephone users would be unaware that a failure had happened. This approach is still used today for optical networks carrying IP services, but the diversity of these services makes it worthwhile to consider more sophisticated protection techniques.

These new techniques rely on two complementary concepts. The first one is *multilayer* protection and is based on the fact that communication components can fail at different network layers and can sometimes also be restored at different layers. Consider for example the loss of an optical channel in a fiber. This failure can be restored by having the optical switching equipment reroute the traffic that was using this channel to another one on the same or some other fiber. This is restoration in the transmission layer. If it is not restored in the transmission layer, the IP routers will eventually detect that a link has stopped operating and will trigger some form of rerouting, such as is done by OSPF. Here, the restoration is being done at the IP layer.

The other new concept is that of differentiated reliability or resilience. This is based on the fact that not all applications need the same level of protection or speed of restoration. A mail service can be delayed for some minutes, while the transmission of a videoconference cannot be interrupted for more than a fraction of a second before it becomes noticeable. Differentiated reliability is simply a set of techniques that allow an operator to match the level of reliability to different service and application requirements.

A multilayer restoration scheme in IP/WDM networks is examined in [4]. The authors evaluate the performance of

0890-8044/06/$20.00 © 2006 IEEE

both IP network resilience capabilities and optical protection at the WDM level. They consider only one QoS parameter: the length of time during which traffic was lost using the two protection schemes, both singly and jointly. IP differentiation is provided by assigning two priorities to traffic, gold and best effort (BE). No traffic differentiation is considered at the transmission layer.

A study of single fault management in IP-over-WDM networks is carried out in [5]. The authors propose two failure recovery techniques. One provides protection at theWDM layer by setting up backup paths, whereas the other provides restoration at the IP layer by overprovisioning the network. Heuristics are developed to investigate the maximum guaranteed network capacity and recovery time for the two fault management techniques. This article is one among many that study the problem of failure propagation between the transmission and IP layer and provide solutions on efficient multilayered resilience mechanisms. However, it does not cover the problem where several traffic classes must be given differentiated protection in any chosen resilience mechanism.

A study of IP restoration and WDM protection is presented in [6], which compares the most common restoration and protection methods available at the IP and WDM layers. It proposes a heuristic that optimizes the simultaneous use of the described IP restoration andWDM protection schemes. The heuristic is designed to optimally solve what the authors call the IP-aware wavelength mileage (IWM) problem. This minimizes the overall network cost which is defined as the sum of the total wavelength mileage for both working and spare lighpaths in the optical layer, and the cost of implementing resilience in the IP layer. According to the authors, this resilience cost can be either a performance-related quantity such as the cost of traffic disruption and QoS degradation due to IP rerouting or an abstract parameter that represents the designer's willingness to provide more or less reliability at either network layer. The optimal solution specifies the percentage of traffic protected by the WDM layer and that of traffic relying on IP restoration. The authors suggest that further work is required in order to take into account application requirements beyond the cost of traffic protection.

The authors of both [6] and [5] propose the combination of WDM protection mechanisms with IP restoration capabilities to minimize bandwidth provisioning costs at the WDM layer. The results of [5] show that WDM shared-path protection outperforms IP restoration for maximum guaranteed network capacity in case of a single fiber failure. Results also show that WDM protection offers much faster recovery times than IP restoration techniques and that the lengthy restoration process offered by the IP layer can severely affect QoS until rerouting is complete.

Differentiated multilayer resilience in IP over optical networks is examined in [7]. After presenting the different IP over optical architectures available, several single-layered resilience schemes such as failure detection, failure notification, and signaling and recovery mechanisms in the optical layer are discussed.

The notion of differentiated protection is discussed in [8]. The authors analyze the gain obtained by deploying in an IP-over-WDM network two classes of service differentiated by their type of protection. The first class offers full 1:1 protection (FP) at the optical layer while the second only offers best effort protection (BEP) when failure occurs. The problem is to maximize the network's revenue defined as the amount of BEP traffic admitted in the network in both regular (no failure) and single failure modes. The authors do not, however, carry out an in-depth comparison between DiffServ and differentiated optical protection.

None of the surveyed articles present the approach we propose in this article, which is to compare the natural resilience mechanisms of DiffServ with complete differentiated protection (DiffProtect) at the optical level. However, our proposals can be framed in multilayer protection specifications such as the one outlined in [9] for metro Ethernet networks. According to [9], the entities that can be protected can be links, paths, and multipoint communication trees, and this can be done at any level in a multilayer structure. There is also a description of the need to transfer information from one layer to the other efficiently and to coordinate the action of the various layers so as to avoid conflicting restoration actions. Our proposals can thus be viewed as two different implementations of such a multilayer protection architecture where the two layers considered are the IP and WDM layers.

The Protection Models

We now describe briefly two techniques that can be used to protect the QoS of services against failures of transmission equipment. One is based on DiffServ, and the other relies exclusively on protection by optical means.

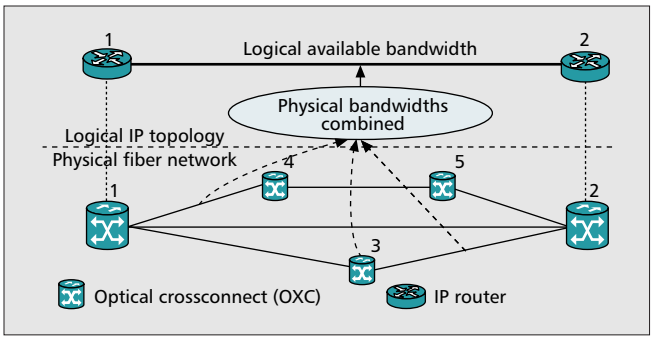
The DiffServ Model

To address the QoS issues that operators are currently facing, the IETF introduced in [10–12] the Differentiated Services (DiffServ) architecture and classes. Three main traffic classes are specified that are served according to a priority hierarchy. The highest class, known as expedited forwarding (EF), is made up of traffic that can tolerate very little loss and delay, and must have ultimate pass-through guarantees. For these reasons, it is served with absolute priority above all other classes. The assured forwarding (AF) class is made up of traffic that can tolerate some amount of loss and delay without suffering an important degradation of the QoS offered to the applications This class is served with medium priority. The final class is the best effort (BE) class which contains traffic that does not require any QoS guarantees. It is served at the lowest priority.

In the presence of congestion, DiffServ does not offer absolute service guarantees, such as a bound on the delay, but only *relative* guarantees for one class with respect to the others. Hence, the only guarantee provided by the priority structure is that EF traffic will always get better service than the others, and that AF traffic will get better service than BE traffic.

DiffServ as a Protection Mechanism

In a network with sufficient bandwidth in normal conditions, DiffServ is able to serve all IP flows at their maximum throughput. When the network's bandwidth becomes insufficient, congestion occurs, and DiffServ is set to provide a differentiated treatment to each class of traffic. Because of this differential handling of traffic classes, we propose that DiffServ can also be viewed as a mechanism for protecting traffic against failures of transmission equipment. In order for this to happen, a transmission failure must trigger the natural differentiated treatment of DiffServ. This happens when the rate at which packets are taken out of the queues is reduced because of the failure. In other words, the failure of one or more optical paths *propagates* as a bandwidth reduction at the IP layer. This causes congestion at the IP layer, and the DiffServ mechanism can then protect each class according to its priority. Since physical resources are no longer sufficient to carry the totality of the IP flows, the DiffServ scheduler indirectly divides the bandwidth among the traffic classes according to their priorities. The EF packets get the largest share of the bandwidth, AF packets are next. Having the lowest priority,



■ Figure 1. *IP layer with the DiffServ protection model.*

BE packets gain access to the remaining physical bandwidth only after EF and AF packets are served.

In order for this failure propagation to happen, the traffic coming out of the DiffServ-enabled router must not be sent on a single transmission system. If this were the case, the failure of this system would appear as a complete link failure, and there would be no possible protection of traffic. Instead, the traffic coming out of the router must be shared between a number of optical systems, preferably independent from each other. When one of these systems fails, we can notify the DiffServ router to reduce the service rate so as not to exceed the remaining transmission capacity; congestion will occur, and the DiffServ mechanisms will provide the necessary differentiation.

This is described in detail in Fig. 1, which shows both the physical and logical layers of what we call the DiffServ protection model. The IP layer is aware of the existence of only one direct IP route between routers 1 and 2, even though this logical route is made up of three optical paths, to be discussed further shortly. To simulate DiffServ protection, all three physical paths are grouped into one logical path connecting routers 1 and 2. The bandwidth available at the IP level is the sum of the bandwidths of all three physical paths.

All the IP traffic arriving at router 1 is already preclassified under the three DiffServ classes EF, AF, and BE. The packets are assigned to their respective queues, and a static priority scheduler serves them according to their priority. In practice, some form of congestion control such as a leaky bucket is put on the sources so that high-priority traffic cannot cause complete starvation of the lower classes in case of congestion. In order to be able to use DiffServ for protection, the packets are then assigned randomly to one of the three available optical paths by the random electronic to optical scheduler (REOS), as seen in Fig. 2. Since the physical bandwidth is randomly shared among these flows, packets of any class can be carried by either one of the three physical paths of the optical layer. After they enter the REOS queue, the packets are served on a first-in first-out (FIFO) discipline by the electronic to optical modulator (EOM), and the DiffServ priorities are no longer used. This is shown on Fig. 2 by the different shapes for the packets after they cross the (fictional) IP to EOM boundary.

Note also that the DiffServ technique does not rely on any IP restoration mechanism to offer reliability in case of fiber failure.

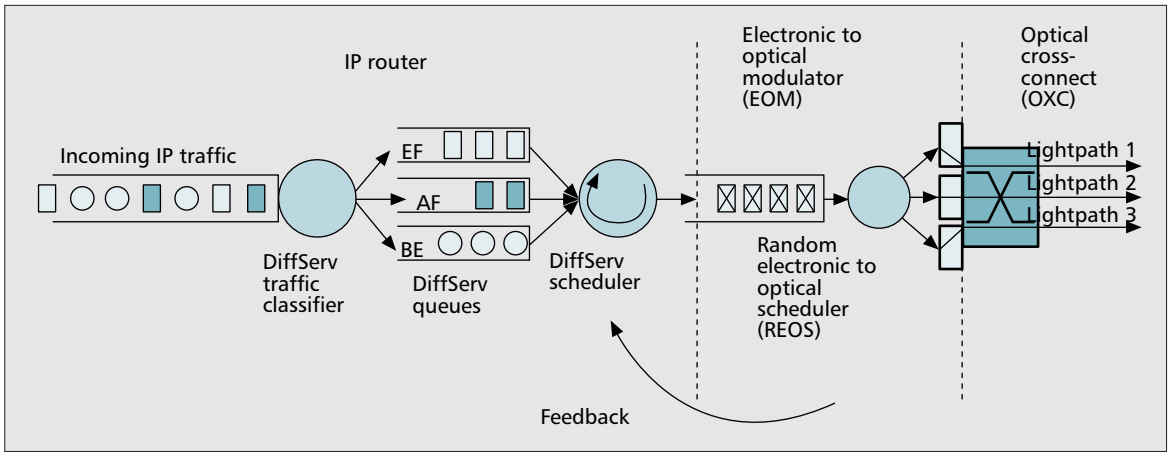
In the optical layer, crossconnects 1 and 2 are connected by three physically disjoint paths: {1,2}, {1,3,2}, and {1,4,5,2}. These paths are set up when the routing and wavelength assignment is done [3] for the optical network and are considered to be fixed. Nodes 3, 4, and 5 represent intermediate crossconnects, whereas nodes 1 and 2 are both crossconnects and routers. Traffic from router 1 is handed down to crossconnect 1 and transmitted over any one of the three optical paths to crossconnect 2 to be passed on to router 2. Lightpaths {1,2}, {1,3,2}, and {1,4,5,2} are assumed to be fiber-disjoint and are known only by crossconnect 1.

Single or multiple failures can occur in this system. A lightpath is said to fail when any segment of that lightpath fails. For example, the {1,3,2} lightpath fails when one of its segments {1,3} or {3,2} fails. We have single, double, or triple failures when one, two, or three lightpaths fail at the same time.

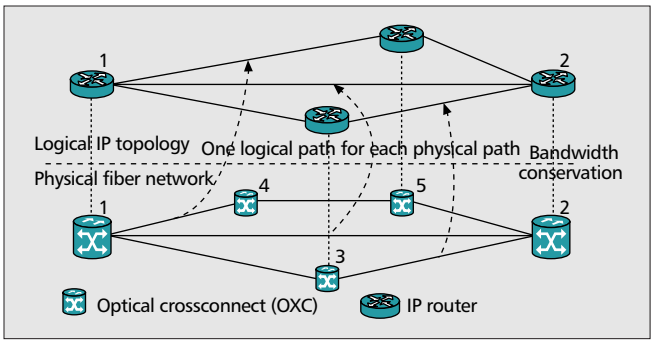
In case of failure(s), the REOS of Fig. 2 detects the event and feeds that information back to the DiffServ scheduler, causing it to reduce its output rate of traffic into the EOM. This causes congestion, and the priority structure of the DiffServ scheduler will then provide differentiated service to the traffic classes.

Given the ability to split IP traffic onto several fiber-disjoint optical systems, a single or multiple failure rarely causes a total loss of connectivity between the affected routers. Going back to Fig. 1, a cut of a fiber between optical crossconnects (OXCs) 4 and 5 does not completely disconnect routers 1 and 2 since some of the traffic can still use the {1,2} and {1,3,2} paths. The DiffServ model translates the partial loss of connectivity into an IP congestion and immediately protects high-priority traffic until the restoration process, if needed, is complete.

Where the network topology is such that three fiber-disjoint lightpaths are not available between a set of two consecutive routers, lightpaths might be required to share the same



■ Figure 2. *IP to optical traffic switching in the DiffServ protection model.*



■ Figure 3. *IP layer with the DiffProtect model.*

fiber. In this case a fiber cut might cause a multiple lightpath failure and thus a more severe connectivity reduction. In this study, we consider that all lightpaths are fiber-disjoint and that they fail independently from one another.

The Differentiated Optical Protection Model

One objective of this work is to estimate the efficiency of DiffServ as a service protection technique. To do this, we define another protection model called differentiated optical protection (DiffProtect), shown in Fig. 3. It represents a simple implementation of differentiated protection at the optical layer and is used in the following as the point of comparison with DiffServ protection.

Optical Protection — In the DiffProtect model, three different levels of protection are offered to the optical systems. With dedicated protection, a fiber is protected by another one that is only used if the first one fails. In Fig. 3 this type of protection is applied to lightpath {1,2}.

With shared protection, several optical paths are protected by a shared path. This could be done in the six-node network of Fig. 5 where we could have decided that the traffic from C_1 to C_3 will be routed on the path $\{C_1, C_2, C_5, C_6, C_3\}$ whenever link $\{C_2, C_3\}$ fails and the traffic from C_4 to C_6 on the path $\{C_4, C_1, C_2, C_5, C_6\}$ whenever link $\{C_4, C_5\}$ fails. The $\{C_2, C_5\}$ link is used to protect two traffic flows: C_1 to C_3 and C_4 to C_6. If we have provisioned one protection channel on this link, we can restore only one of the two failures at a time.

This type of protection is equivalent to dedicated protection in the case of single failures but provides degraded protection when multiple lightpath failures occur, depending on the probability of having a double failure. The degraded protection is expressed as the loss probability, that is, the probability that the path cannot be restored in case of failure. In the case of the network of Fig. 3, shared protection is applied to lightpath {1,3,2}. In a simple topology such as this, there is no other lightpath that can compete with {1,3,2} for its protection resources.

We model the loss probability by assuming that only a given fraction of lightpath {1,3,2}'s traffic is recovered in case of failure. Finally, we have unprotected paths where all the traffic is lost in case of failure. This option is used for lightpath {1,4,5,2}.

Service Protection in the Optical Layer — OXCs cannot access individual packets and the optical layer is incapable of differentiating between the three traffic classes. In case of failure, an OXC can only switch whole optical channels onto the available backup lighpaths. This is why each traffic class must be carried *completely* on a given optical system: no sharing is possible between classes. The technique used to offer service protection in the optical layer is described in Fig. 4. Because the model does not use any IP differentiated protection mechanism, each traffic class is assigned to the optical path with the appropriate protection level based on the DiffServ classification.

Traffic can receive the appropriate protection because a traffic classifier ensures that the highest-priority traffic is assigned to lightpath {1,2}, the medium-priority traffic to lightpath {1,3,2}, and the best effort traffic to lightpath {1,4,5,2}. The EF, AF, and BE queues depicted in Fig. 4 are served in FIFO fashion.

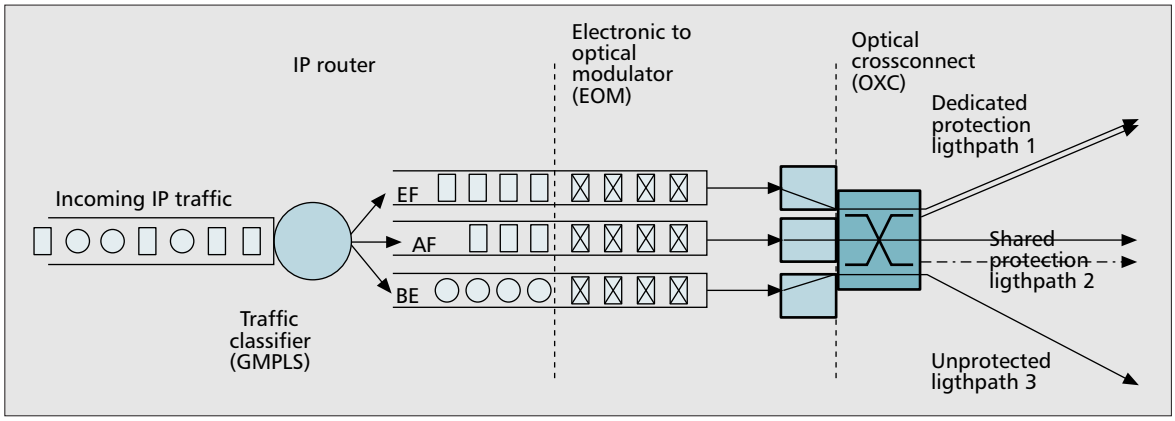
Again, the dashed barrier between the IP router domain and that of the EOM is only conceptual. The assignment of traffic to different queues can only be done at the IP layer. Once the packets enter the optical domain, they lose their DiffServ priority, and the protection they receive is determined exclusively by the lightpath they are in.

This assignment of traffic from the IP to the optical level is possible using some type of joint control between the two layers, such as the one provided by generalized multiprotocol label switching (GMPLS) [13].

Simulation Parameters and Modeling

Traffic Sources

In our simulation we used voice over IP as the highest priority traffic. Each voice source is modeled as an on-off process. The length of an on period is an exponential random variable with mean $1/\mu = 400$ ms; that of an off period is also exponential with a mean of $1/\lambda = 600$ ms. During an on period, the source is active and produces one fixed-length packet of 120 bytes every 15 ms. The off period corresponds to the time during which the source is silent. The peak rate of the source



■ Figure 4. *IP to optical traffic switching in the DiffProtect model.*

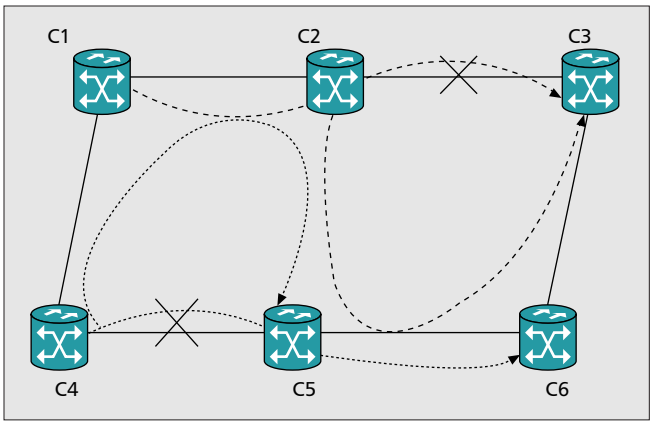


Figure 5. *A shared protection example.*

is then 64 kb/s, the average rate 25.6 kb/s. In the simulation each source starts transmitting at a random interval T_0 from the start of the simulation. This is computed by drawing a number p from a uniform distribution $U(0, 1)$. If $p < P_{On}$, $T_0 = 0$, where P_{On} is the probability of being in the on state and is given by

$$P_{On} = \frac{1}{1+\mu/\lambda}. \quad (1)$$

Otherwise, T_0 is drawn from an exponential distribution with parameter λ. This is to avoid synchronization of the sources during the first few on-off cycles (Fig. 6).

Medium-priority traffic is video streaming, such as video on demand (VoD). For this traffic, we have chosen a video traffic generator based on the transform expand sample (TES) model of MPEG-4 trace files as described in [14]. It generates traffic that has the same first- and second-order statistics as an original MPEG-4 trace. Video packets have variable sizes ranging from 100 to 1000 bytes. This type of traffic was assigned to the AF class as proposed in [15].

Finally, two UDP sources are used to simulate regular best effort traffic. This could represent a data file transfer with no particular reliability requirements. The sources are also on-off exponential sources whose burst and idle times follow exponential distributions with means varying from 1 and 0.5 to 1.5 and 1 s, and generate large 1000-byte packets.

We have not considered any of the physical layer overheads in this study. Thus, we assume that the link bandwidth needed for a transmission is equal to the source rate of that transmission (e.g., each voice source generates a bitstream that occupies 64 kb/s of bandwidth of a transmission link). The number of sources and their characteristics were chosen so that the amount of traffic generated for each type of service was approximately the same. There were a total of 75 VoIP, 1 video, and 2 data sources. Figure 7 shows the throughput distribution of the aggregate voice flow. The average throughput of the voice sources is approximatively 3.39 Mb/s, while the maximum throughput reached 4.58 Mb/s with an absolute maximum of 4.8 Mb/s when all sources are active at the same time.

Figure 8 shows the throughput distribution of the video source used in the simulations. It generates an average rate of 3.08 Mb/s, and its maximum throughput was 3.85 Mb/s.

Each UDP source is set to generate a total of 2.5 Mb/s during its on periods. Figure 9 shows the physical throughput distributions of the two sources combined. The sources generate a mean rate of 3.16 Mb/s with a maximum rate of 5 Mb/s.

The capacity of each lightpath in Figs. 1 and 3 is 5 Mb/s. By switching each flow on its reserved lightpath, DiffProtect can provide a maximum capacity of 5 Mb/s to each class of traffic. Since the maximum throughput of any source is less than 5 Mb/s, each lightpath has enough capacity to carry its offered flow.

In the DiffServ model, all sources share the three lightpaths randomly. The total capacity of the link between routers 1 and 2 is thus 15 Mb/s. As in the DiffProtect case, each traffic class should normally require no more than 5 Mb/s of the available physical capacity and the service discipline will have the effect that each traffic effectively receives an approximately equal share of the 15 Mb/s bandwidth. Figure 10 shows the throughput distribution of all three traffic flows combined. The combined sources generate an average of 9.63Mb/s up to a maximum of 12.6Mb/s which does not exceeds the 15 Mb/s transmission capacity.

Finally, note that only the UDP protocol has been used in our simulations. Unlike TCP, UDP does not provide retransmission in case of congestion, and protection becomes totally dependent on the mechanisms offered by the network. The results presented here are thus a worst-case scenario where no source protection is offered to the traffic flows. TCP could conceivably be viewed as an additional protection mechanism operating at a layer above DiffServ.

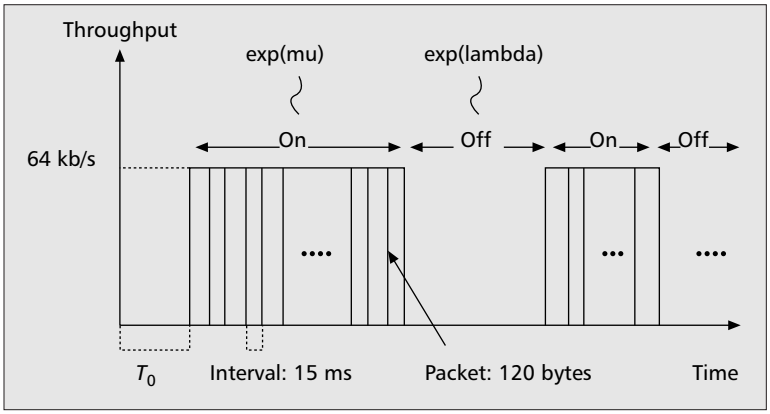


Figure 6. *Throughput of a VoIP source.*

Simulation Model

The simulator used in this study was ns-2 [16], an open source software designed to simulate the IP layer of networks. Currently, it cannot model the lower layers such as optical transmission systems and crossconnects.

We want to measure the network's performance under the failure of various elements. Any given set of failed components defines what we call a *failure configuration*. For small networks, we can enumerate all possible failure configurations and run a complete simulation for each case. This approach does not scale when networks get larger, and with this in mind we have chosen to *sample* the set of configurations instead of making a complete enumeration. The sampling is done by generating many configurations randomly. During a run, each lightpath experiences a sequence of intervals with different failure configurations. The intervals between failures and their duration are generated following exponential distributions. In what follows, a *failure scenario* is defined as a particular sequence of single, double, triple, or no failure intervals.

A run lasts 2000 s of simulation time during which we collect performance statistics. At the end, we compute the performance statistics for each configuration that occurred during the run. Runs

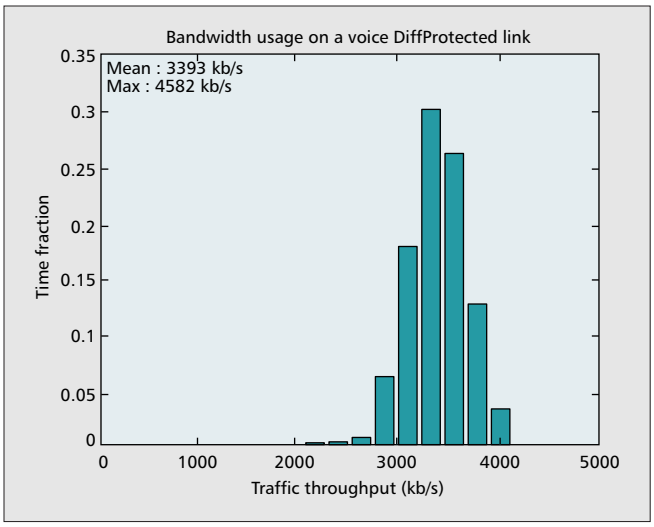


Figure 7. *Voice throughput distribution on a DiffProtect link.*

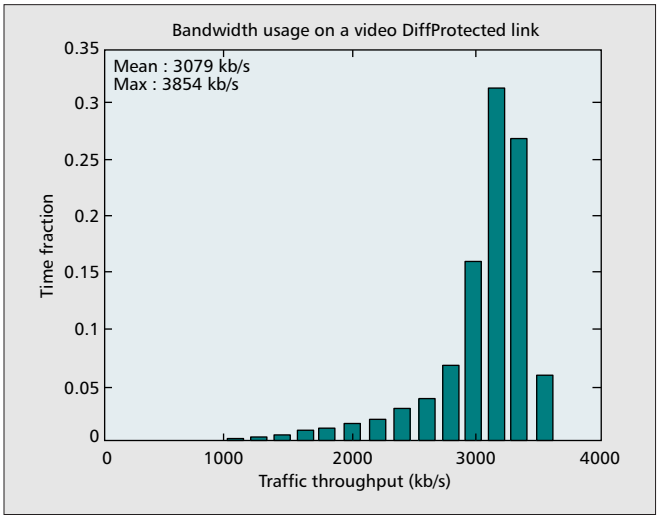


Figure 8. *Video throughput distribution on a DiffProtect link.*

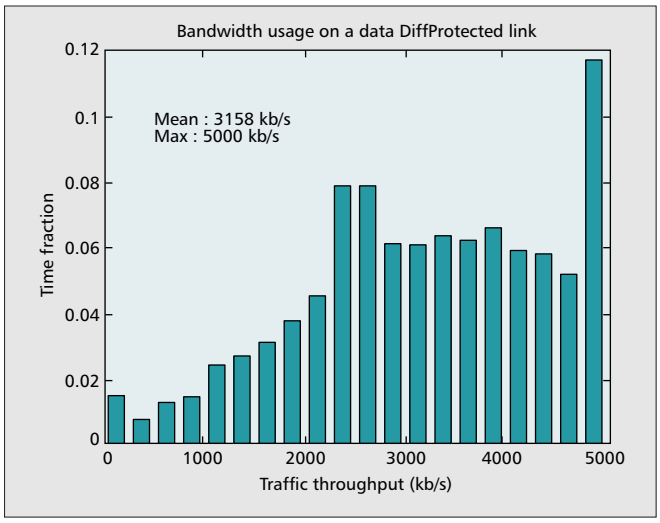


Figure 9. *Data throughput distribution on a DiffProtect link.*

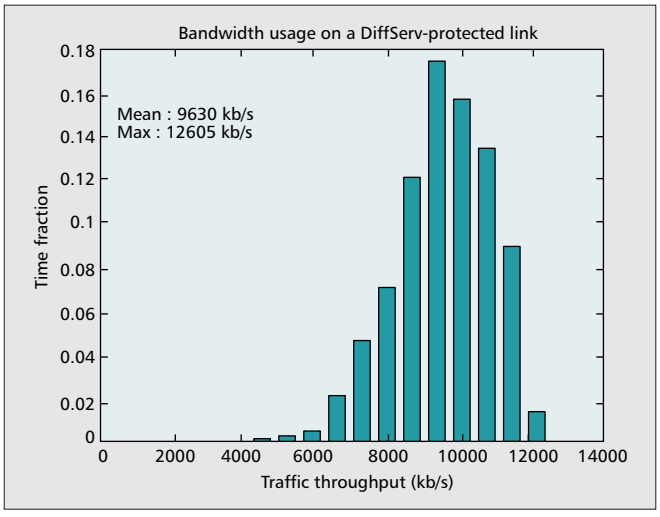


Figure 10. *Throughput distribution on a DiffServ link.*

for each simulated case were repeated 20 times, which allowed us to collect performance statistics with 95 percent confidence intervals.

As we said earlier, each traffic class normally requires no more than a third of the available network capacity in both the DiffServ and DiffProtect models. Physical resources are thus adequately dimensioned to carry all traffic flows at their maximum throughput when there is no failure.

In the DiffServ model a failure of one or more lightpaths translates into reduction of the available bandwidth at the IP layer. Since all three lightpaths have a capacity of 5 Mb/s, when one lightpath fails, the IP bandwidth is reduced by a third to 10 Mb/s. The area to the right of 10 Mb/s under the distribution shown in Fig. 10 indicates that this causes congestion about 50 percent of the time. In that case the DiffServ mechanism deals with the congestion by serving the EF, AF, and BE packets according to their priority.

The network simulator ns-2 is only aware of IP links and routers. To simulate the DiffProtect model, we had to add intermediate routers at the IP layer to mimic the fact that each class of traffic is taking a different optical route. A failure in the optical path for the EF class produces no effect at the IP layer other than a very fast recovery time of 50 ms because that path is supposed to be 100 percent protected. Medium class protection for the AF class is simulated as a 50 percent reduction of the IP bandwidth for that class. Finally, when a failure occurs in the unprotected optical path, all the packets of that class are lost.

Performance Under Failures

In what follows we study the performance of DiffServ and DiffProtect as standalone protection schemes in a two-node network. Similar results were also produced for a different architecture of a six-node network to check if the two-node network conclusions are still valid.

For each simulation of both two- and six-node networks, we evaluated the average packet loss, end-to-end delay, and jitter. The average packet loss is defined as the number of packets lost divided by the number of packets generated. The average packet delay is equal to the sum of the delays of all received packets divided by the number of received packets. And finally, let d_i be the delay of packet i and n the total number of received packets. The average jitter is calculated as follows:

$$\frac{\sum_{i=2}^{n} |d_i - d_{i-1}|}{n-1}.$$

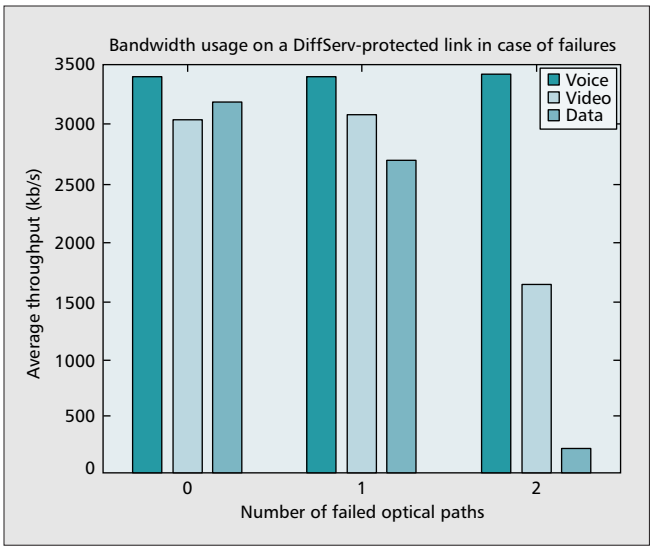


Figure 11. *Bandwidth usage per class under DiffServ protection.*

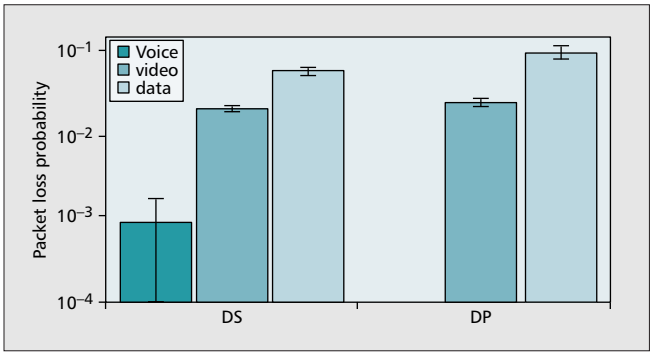


Figure 12. *Average losses under failures.*

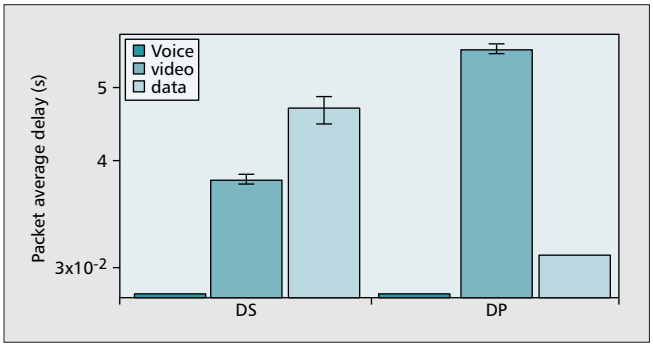


Figure 13. *Average end-to-end delay under failures.*

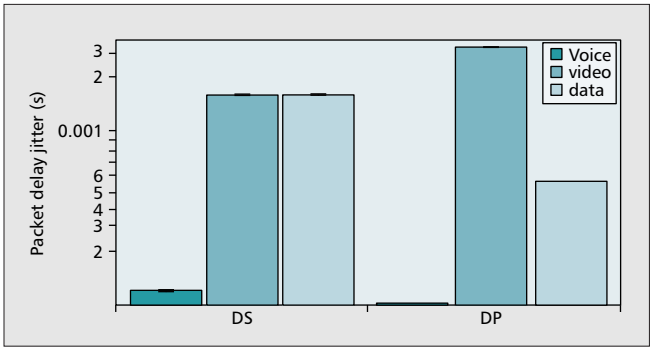


Figure 14. *Average packet jitter under failures.*

Two-Node Network

Figures 12, 13, and 14 show the average packet loss, end-to-end delay, and jitter for each protection model and their respective 95 percent confidence intervals. Figure 12 shows that the high-priority voice traffic is 100 percent protected against losses in both cases. In the case of DiffServ this is a direct consequence of the priority scheduling, which gives absolute priority to the EF class. This can be seen in Fig. 11, which shows that the amount of bandwidth used by each class under each failure configuration is directly related to the priority of that class. In the case of DiffProtect, this is due to the fact that the EF traffic is offered to a backup server at a rate high enough to carry all the traffic.

DiffServ, however, offers better loss guarantees than DiffProtect for both video and data traffic since all traffic classes can share the same physical bandwidth. When failure(s) occur, the bandwidth available to IP traffic is reduced. The lower-priority packets can wait in their queue until all the higher-priority traffic has been served, and this tends to reduce traffic losses by queuing packets rather than dropping them. With DiffProtect, each traffic has its own dedicated share of the available physical resources. A particular traffic suffer losses only when a failure affects the lightpath on which it is carried. In that case, there is no sharing of physical resources and packets are dropped rather than queued which tends to reduce the delay for the packets that are not dropped.

Figure 13 shows average end-to-end packet delays for both DiffServ and DiffProtect under both normal and failure modes. Voice traffic is equally well protected against delays in both models, and delays for video traffic are much lower under DiffServ protection than DiffProtect. Finally, data packets are guaranteed smaller delays in the DiffProtect model. The reason is that we compute the delay only for those packets that reach their destination, and the loss rate is quite large. There are fewer surviving packets but they get to their destination faster.

Figure 14 shows the average jitter of all packet types under both protection schemes. Voice packets are equally well protected against jitter in both DiffServ and DiffProtect. On the other hand, video jitter is higher and data jitter lower with DiffProtect than with DiffServ.

The jitter of a bursty (video) traffic source is directly related to the capacity used to serve it. Only a very high link capacity can guarantee minimal (close to zero) jitter. When a failure affects the video-carrying lightpath in the DiffProtect case, its bandwidth is reduced by half, which dramatically increases its traffic jitter.

As for data jitter, the DiffProtect model tends to drop BE packets rather than queue them in case of failures so that the BE packets that do make it to their destinations will do so with minimal delay and jitter. In the DiffServ case, BE packets suffer large delays and jitter since their service rate is severely limited by the traffic of higher-priority classes.

From this first set of results, we can conclude that either of the two schemes can be safely used for voice, but DiffServ produced better performance for video traffic, especially in terms of delay and jitter. Such improved performance from the AF class on DiffServ comes, as expected, at the expense of BE traffic, whose QoS measures degrade more under DiffServ.

Six-Node Network

We also carried out some tests for the six-node network shown in Fig. 15 to check whether the results obtained with the two-node network extend to larger networks as well.

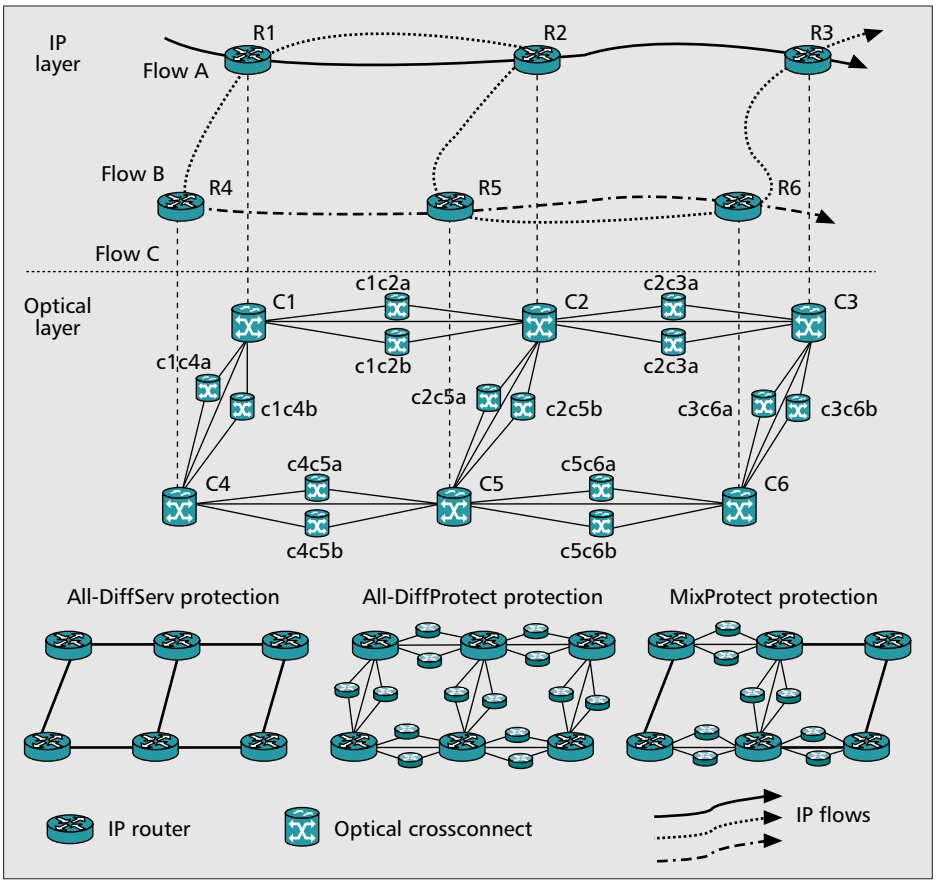
We have simulated three types of networks: one where all the links have DiffProtect, one with all the links protected with DiffServ, and one that has some links with DiffServ and

others with DiffProtect, called *MixProtect*. Each of the IP networks presented in the lower part of Fig. 15 was simulated a total of 10 times with random failure scenarios. The averaged results are presented in Figs. 16, 17, and 18. Figure 16 shows the average packet loss for each IP flow carried by the simulated network. We can clearly see for all three flows that high-priority voice packets are well protected against losses in any of the three protection schemes. On the other hand, losses for video and data packets tend to increase when using DiffProtect inside the network.

Figure 17 shows the average packet delays for all traffic classes. While voice packets are offered the best protection against increasing delays, video delays increase when the network is partially or totally protected by DiffProtect. Since data packets are more likely to be dropped than queued, the delays of those data packets that are not dropped tend to decrease with the increasing use of DiffProtect in the network.

The average jitter performance is shown in Fig. 18. High-priority packets are always well protected against increasing jitter. Video jitter is lowest when DiffServ is the protection mechanism of choice used throughout the network. Similar to data delay behavior, data jitter tends to also decrease when using DiffProtect.

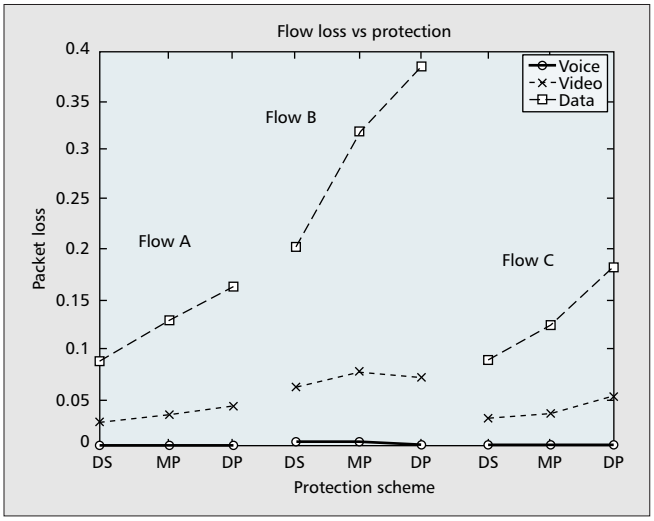
These results show that high-priority traffic is equally well protected in either protection scheme. The protection for video against losses, delays, and jitter increases with increasing use of DiffServ, and is best protected against losses, delays, and jitter when only DiffServ is used. The QoS offered to video traffic gradually degrades with increasing use of DiffProtect. While data packets are best protected against losses in an all-DiffServ scheme, their delay and jitter decrease with increasing use of DiffProtect. The conclusions from the observation on this type of network confirm those made with the two-node network: the EF class is well protected in all cases, the AF class presents better performance with DiffServ, and BE traffic sees its jitter and delay values degraded and its loss improved when DiffServ is used.



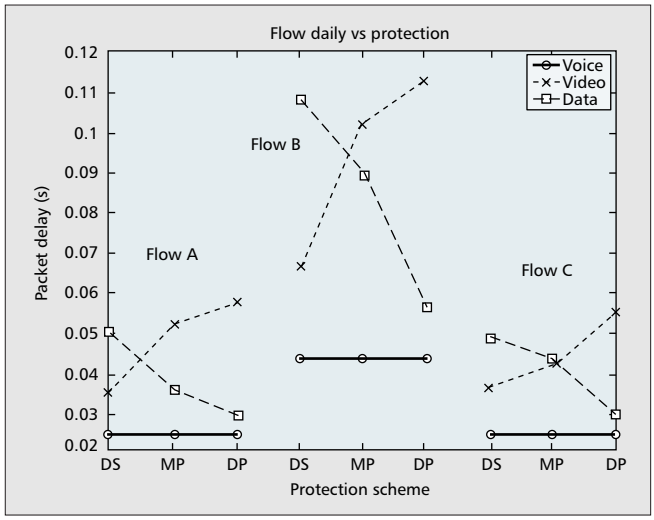
■ Figure 15. *Six-node network.*

Conclusion

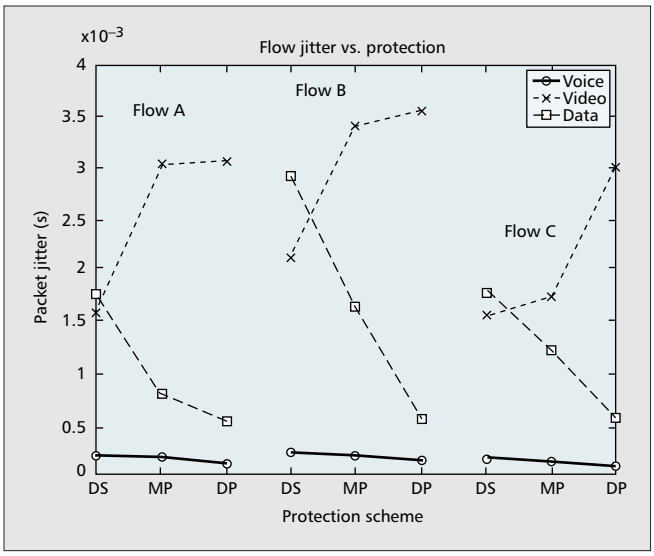
Even though DiffServ was designed without any particular consideration for reliability, we investigated to what extent differentiated services at the IP level of an IP/WDM system could be used to design a network with more resilient QoS in the



■ Figure 16. *Average flow loss.*



■ Figure 17. *Average flow delay.*



■ Figure 18. *Average flow jitter.*

presence of failures. For this, we first proposed a particular DiffServ/WDM model mechanism that behaves as a fault management technique to protect high-priority traffic with little delay.We have also proposed the DiffProtect model that provides reliability in the optical layer. These types of models can be implemented within the framework of new multilayer architectural proposals such as the metro Ethernet.

The conclusions that can be drawn from comparison of the two schemes can be summarized as follows. High-priority traffic is equally well protected by both DiffServ and DiffProtect for all measures of QoS: delay, jitter, and loss. This would suggest that no optical protection is needed for high-priority traffic. This is a significant advantage with respect to standard SDH protection techniques when there is a significant amount of EF traffic. As expected, low-priority traffic suffers large losses in all cases, but somewhat less with DiffServ than with DiffProtect. DiffProtect tends to produce better delay and jitter performance for best effort.

The most noteworthy differences between the two protection mechanisms show up for medium-priority traffic, for which the DiffServ protection presents the best averaged behavior for the two QoS mechanisms and their combinations. This difference for the AF class is a crucial point for ISPs if they want to offer more than just the two EF and BE services, as is currently done.

So, as a result of this study, the answer to our initial question, "Can DiffServ guarantee IP QoS under failures?" seems to be yes. The average behavior of the DiffServ model under failures indicates that it performs as well and even better than more costly and sophisticated differentiated optical protection mechanisms. Moreover, the method does not suffer the long delays introduced by IP layer restoration techniques. As for DiffProtect, according to our results, it can protect the system against some extreme behaviors. A relevant question that arises in this context is what is the ideal mix of DiffServ and DiffProtect for a given network. We have explored this issue by using total enumeration [17] on a small configuration with only three links. Clearly, total enumeration cannot be used in large-scale networks. Therefore, there is an obvious need to find analytical models to compute the optimal mix of the two techniques for networks of any size to combine their respective advantages. We can then carry out a systematic cost/benefit analysis of the technique to maximize traffic protection and QoS in case of failure while minimizing the cost of physical protection.

Acknowledgments

This work was partly supported by the National Sciences and Engineering Research Council of Canada and Quebec's NATEQ. The authors would also like to thank the reviewers and the Editor-in-Chief for their constructive suggestions, which significantly improved this article.

References

[1] B. Mukherjee, *Optical Communication Networks*, McGraw-Hill, 1997.
[2] W. D. Grover, *Mesh-Based Survivable Transport Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking*, Prentice Hall PTR, 2003.
[3] M. Pióro and D. Medhi, *Routing, Flow and Capacity Design in Communication and Computer Networks*, Morgan Kaufman, 2004.
[4] C. Guillemot, B. Jalliffier, and M. Le Foll, "Design and Experimental Assessment of a Multi-Level Restoration Scheme on the Broadband IP/WDM VTHD Network," *Proc. World Telecommun. Cong.*, Sept. 2002.
[5] L. Sahasrabuddhe, S. Ramamurthy, and B. Mukherjee, "Fault Management in IP-over-WDM Networks: WDM Protection Versus IP Restoration," *IEEE JSAC*, vol. 20, no. 1, Jan. 2002, pp. 21–33.
[6] A. Fumagalli and L. Valcarenghi, "IP Restoration vs. WDM Protection: Is There an Optimal Choice?," *IEEE Network*, vol. 14, no. 6, Nov. 2000, pp. 34–41.
[7] A. Autenrieth, "Differentiated Multilayer Resilience in IP Over Optical Networks," *Proc. SSGRR 2002, Int'l. Conf. Advances in Infrastructure for eBusiness, e-Education, e-Science, and e-Medicine*, July 2002.
[8] A. Nucci *et al.*, "Exploiting Failure Recovery for the Robust Support of Two Service Classes in IP over WDM networks," Tech. rep. TR02ATL-071001, Sprint Labs, July 2002.
[9] "Requirements and Framework for Ethernet Service Protection in Metro Ethernet Networks," http://www. metroethernetforum.org/TechSpec.htm, 2004.
[10] S. Blake *et al.*, "An Architecture for Differentiated Services," IETF RFC 2475, Dec. 1998.
[11] J. Heinanen *et al.*, "Assured Forwarding PHB Group," IETF RFC 2597, June 1999.
[12] V. Jacobson, K. Nichols, and K. Poduri, "An Expedite Forwarding PHB," IETF RFC 2598, June 1999.
[13] E. Mannie, "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," IETF RFC 3945, Oct. 2004.
[14] A. Matrawy, I. Lambadaris, and C. Huang, "MPEG4 Traffic Modeling Using the Transform Expand Sample Methodology," *Proc. 4th IEEE Int'l. Wksp. Network Appliances*, Gaithersburg, MDD, Jan. 2002.
[15] Y. Koucheryavy, D. Moltchanov, and J. Harju, "A Topdown Approach to VoD Traffic Transmission Over DiffServ Domain Using AF PHB Class," *Proc. ICC 2003*, May 2003, vol. 26, pp. 243–49.
[16] "The Network Simulator ns-2," http://www.isi.edu/ nsnam/ns/
[17] B. Sansò, C. Awad, and A. Girard, "Network Reliability Under Mixed IP and Optical Protection," *Proc. DRCN 2005*, Oct. 2005, pp. 187–94.

Biographies

BRUNILDE SANSÒ (brunilde.sanso@polymtl.ca) is a full professor of electrical engineering at Ecole Polytechnique de Montréal and director of the LORLAB. Her interests are in performance, reliability, design, and optimization of wireless and wireline networks. She is a recipient of several awards, Associate Editor of *Telecommunication Systems*, and editor of two books on planning and performance.

CHRISTIAN AWAD (christian.awad@polymtl.ca) received his Bachelor's degree in computer engineering in 2003 from Ecole Polytechnique of Montréal. He is currently pursuing his doctoral studies at Ecole Polytechnique. His current research interests are in the area of multimedia and data communication networks, and he specializes in quality of service, differentiated services and differentiated protection, restoration, and reliability in IP over WDM networks.

ANDRÉ GIRARD (andre@emt.inrs.ca) is an honorary professor at INRS-EMT and an adjunct professor at Ecole Polytechnique de Montréal. His interest is in the modeling, performance analysis, and optimization of telecommunication networks. He is Associate Editor for *Telecommunication Systems* and author of the book *Routing and Dimensioning for Circuit-Switched Networks* as well as of many journal and conference publications.