# An Architecture for Differentiated Protection Against Single and Double Faults in GMPLS

F. Ricciato*, M. Listanti

*INFOCOM Department, University of Rome ''La Sapienza'', V. Eudossiana, 18 00184 Roma*
*E-mail: ricciato@coritel.it, marco@infocom.uniromal.it*

S. Salsano

*DIE, University of Rome ''Tor Vergata'', Viale Politecnico, 1-00133 Roma, Italy*
*E-mail: stefano.salsano@uniromaz.it*

**Abstract.**  In the context of an optical network GMPLS can be used to provide network robustness to faults through end-to-end path protection techniques. In this paper, we present a dynamic distributed model supporting five different classes of protection, including protection against single and double fault, with and without sharing of backup bandwidth. Beyond link and node failures we also consider protection against shared risk link group (SLRG) failure. In this paper, we briefly describe the protection model and the underlying algorithms for route selection and backup bandwidth sharing. After that we face the following issue: Which subset out of the five possible protection classes is convenient for an operator to support on the same network infrastructure? To answer this question it is fundamental to have a clear view of the trade-offs between the costs and the performances associated to each class. To achieve that we carried out an extensive performance analysis by means of simulations. For each protection class, we evaluated two fundamental performance metrics: the recovery probability under multiple faults, and the average per-demand resource usage. On the basis of such results, we are able to identify some basic guidelines driving the choice of the more convenient subset of protection classes to be implemented within a single network.

**Keywords:**  protection and restoration, double fault recovery, GMPLS

## 1   Introduction

The capabilities introduced by the GMPLS control plane can be exploited to provide fault recovery in an optical network. An overview of the relevant concepts can be found in Lang and Drake [1]. The standardization process about GMPLS recovery is considerably active in the IETF community, but still in an early stage [2–4]. Among the various possible recovery schemes, those based on end-to-end path protection are attractive for their efficiency in mesh network [2]. Several recent works [7–10] aim at designing and evaluating dynamic algorithms and protocol mechanisms for path based protection in search of maximum network utilization efficiency. Another important issue is the ability to provide differentiated levels of fault-recovery performances to different classes of traffic [11,12], which adds flexibility in the design of the services offered to the customers.

In this paper, we present a model for fault recovery in an optical network. We assume full wavelength conversion capabilities at the optical switches. We consider a dynamic model (also called ''on-line'') as connection requests are allocated in the order they arrive, without rearrangement of previously established circuits. We consider path-based protection (also known as ''global'' protection) as opposed to link-based (or ''local'') protection. This means that the ingress node is able to detect a fault along the service circuit and readily switch the traffic on a pre-established backup circuit. In this work, a fault can refer to the failure of a single link, a node or a shared risk link group (SLRG) [5]. A SLRG identifies a set of links that can be contemporarily interrupted by the same event, for example a cut in a conduit breaking all the fibers through it. Furthermore, we also consider the case that some specific network links are not prone to failure. We call them no-risk link (NRL). An NRL can model, for example, a transmission link provided with some kind of protection switching mechanism acting locally, for example, a fiber protected $1 + 1$ locally.

*Corresponding author.

An important feature of our model is that it supports protection differentiation on a per-demand basis. Along with traditional dedicated/shared protection against a single fault, we also consider schemes for dedicated/shared protection against double fault. Together with the basic unprotected class, we end up with a total of five protection classes.

In this paper, we present a dynamic model supporting the full set of five protection classes. We describe the overall protection model in Section 2, along with a brief description of the underlying algorithms for route selection and bandwidth sharing (the interested reader is referred to Ricciato [13] and Ricciato et al. [14] for further algorithmic details). After that we face the following issue: which subset out of the five possible protection classes is convenient for an operator to support on the same network infrastructure? To answer this question it is fundamental to have a clear view of the trade-offs between the costs and performances associated to each class. To achieve that, we carried out an extensive performance analysis by means of simulations, whose main results are presented in Section 3. In particular, for each protection class we investigated two fundamental performance metrics: the recovery probability under multiple faults, and the average per-demand resources usage. On the basis of such results, we are able to identify some basic guidelines for the choice of the more convenient subset of protection classes to be supported by a single network, as discussed in the conclusive Section 4.

### 1.1 Related Work

The closest previous papers to this work are Grover and Clouqueur [12], Clouqueur and Grover [15,16]. In Clouqueur and Grover [15], the authors carried out an ''availability analysis'' under dual link failure. They found that a restorable mesh network designed against single link failure offers a considerable robustness to dual link failure also. Our results confirm those finding for the case of dual-fault. Furthermore, we extend the availability analysis to multiple contemporary faults—up to five—and to the case of mixed traffic where demands protected in different ways coexist in the same network. In Clouqueur and Grover [16], the same authors proposed methods for optimal capacity design achieving 100% robustness to dual-link failure. There the authors considered the static ''off-line'' problem—that is, joint allocation of all demands at the same time—and used LP optimization

formulations, while in the present work we face the ''on-line'' problem and adopt more scalable dynamic algorithms. In addition, we consider bandwidth sharing applied to both single and double fault protected demands, defining the related mechanisms for a distributed implementation. In Grover and Clouqueur [12] the authors start to investigate the convenience of per-demand protection differentiation, included preemptable service. In this sense, our work represents a complementary contribution carrying new results.

A model for a differentiated resilience scheme was also proposed in Autenrieth and Kirstädter [11] in the context of IP/MPLS networks. There the authors focused mainly on the differentiation in the recovery time requirements, limited to recovery from single fault. In this sense our work is orthogonal to that, as our focus is on the differentiation of recovery probability under multiple faults.

Recovery from double link failure was also considered in Choi et al. [17], where the authors adopted a local (link-based) protection scheme and did not consider SRLGs failures, while we follow the path-based approach.

Regarding the algorithmic details, the bandwidth sharing mechanism proposed here basically extends to double-fault protection what proposed in Doverspike et al. [9] for single-fault protection. The algorithms used to find SRLG-disjoint paths exploits elementary graph transformations jointly with the well-known Suurballe algorithm, as already presented in a previous work [14]. The same approach was independently found also in Ellinas et al. [18] limited to the case of single-fault protection.

## 2 The Model

### 2.1 General Description

In our model, each connection request can be associated to three different kinds of protection: Unprotected (UP), single-fault protected (SFP), double-fault protected (DFP). For UP demands only a service circuit $P_s$ is established, and no service continuity is guaranteed after the occurrence of a fault along $P_s$. For SFP demands a service circuit $P_s$ plus a single primary backup path $P_{r1}$ are allocated: upon failure of $P_s$ the ingress edge node readily switches the traffics on $P_{r1}$. In case of DFP demands, a service circuit $P_s$ plus two backup paths are allocated: a

*Table 1.* The five considered protection classes.

| H | Acronym | Protection Class | Number of Associated Circuits | Backup bw |
|---|---------|------------------|-------------------------------|-----------|
| 0 | UP | Unprotected | 1 (service circuit) | n.a. |
| 1 | Sh-SFP | Shared single-fault $P$ | 2 (1 service + 1 backup) | Shared |
| 2 | De-SFP | Dedicated single-fault $P$ | | Dedicated |
| 3 | Sh-DFP | Shared double-fault $P$ | 3 (1 service + 2 backup) | Shared |
| 4 | De-DFP | Dedicated double-fault $P$ | | Dedicated |

primary backup $P_{r1}$ and a secondary one $P_{r2}$. Upon failure of $P_s$ the traffic is switched on $P_{r1}$, and in case of contemporary interruption of $P_s$ and $P_{r1}$ it is switched to $P_{r2}$. Note that for DFP the order of preference between the two backup paths is fixed *a priori*.

Both SFP and DFP schemes can be implemented as dedicated or shared protection (following the terminology in Lang and Rajagopalan [2]), resulting in a total of five protection classes as sketched in Table 1. In case of dedicated protection the backup circuits are pre-established and ready to carry traffic. With shared protection during the setup of a backup circuit the resources are reserved and the signaling session is installed, but the circuit is not cross-connected at the optical level. In other words, backup circuits that are established on the control plane but not on the data plane. We will call such backup circuits as ''pre-qualified''. A pre-qualified circuit will be activated using a second signaling procedure only in response to a network fault. The activation phase adds a delay to the recovery time. Note that according to the terminology in Papadimitriou and Mannie [4] the shared protection scheme considered here would be classified as a restoration scheme.

With shared protection, it is possible to achieve a minor resources usage, as it is possible to reuse $n$ resources to protect a pool of $m \geq n$ service circuits whenever the pool is such that at most $n$ out of $m$ can be interrupted at the same time due to topological diversity [2]. On the other hand, this resources saving comes at the expense of a longer recovery delay due to the activation phase of pre-qualified circuits. The amount of additional delay is related to technological and/or equipment-specific factors. Furthermore, as will be discussed in Section 2.5, several additional capabilities are required at the intermediate nodes in support of shared protection, namely: (i) maintenance of local state to support computation of reserved shared bandwidth, (ii) contention resolution mechanisms based on pre-emption, and (iii) the capability to

fast cross-connect pre-qualified paths. The convenience of introducing such additional capabilities must be assessed with respect to the gain yielded by shared versus dedicated protection. The quantitative analysis presented later in Section 3 is a contribution towards such assessment.

### 2.2 Service Level Specifications

The amount of backup reserved resources on each network link and the reaction procedures to the faults will be designed in order to provide the following resilience guarantees with respect to the number $\varphi$ of contemporary faults in place in the network:

- In case of one single fault ($\varphi = 1$) all the affected SFP and DFP demands are guaranteed service continuity over the corresponding primary backup circuit.
- In case that a second fault occur afterwards ($\varphi = 2$) all the DFP demands experiencing interruption of the service circuit $P_s$ are guaranteed service continuity over the primary $P_{r1}$ or secondary $P_{r2}$ backup circuit, depending on the integrity of $P_{r1}$. For SFP demands, we distinguish between those affected—and recovered—by the first fault in chronological order (denote by $F_1$), and those affected by the second fault ($F_2$). All the SFP demands recovered from $F_1$ are guaranteed service continuity over the respective backup paths unless they are interrupted by $F_2$. There are no guarantees of successful recovery for all the SFP demands affected by $F_2$. Nevertheless they will be recovered in a sort of ''best-effort'' fashion.
- In case of successive fault ($\varphi > 2$), all the demands successfully recovered from the previous faults ($F_1, F_2, \ldots, F_{\varphi-1}$) are guaranteed service continuity over the respective backup paths unless they are interrupted by $F_\varphi$. There

are no recovery guarantees for those demands affected by $F_\varphi$, but again they are recovered in a sort of ''best-effort'' fashion.

We will show below that DFP demands are more likely to be successfully recovered than SFP when $\varphi > 2$, thus preserving a certain degree of service differentiation between DFP and SFP even under multiple contemporary faults.

## 2.3  Functional Description of the Allocation Procedure

We assume that connection requests arrive dynamically to the network. Each request is associated to the following attributes: ingress/egress node pair $\langle N_i, N_e \rangle$, requested bandwidth $B$ and protection class $H$. In an optically switched networks the requested bandwidth $B$ can be assumed equal to an integer number of wavelengths, usually $B = 1$. Five different values of $H$ are used to discriminate between the five classes as reported in Table 1. For each request, the ingress edge node $N_i$ must dynamically establish the service and backup paths towards $N_e$ using GMPLS signaling. The route selection process is run by a dedicated module, called the route selection engine (RSE). In the centralized approach a single RSE server is maintained for the whole network, and a communication procedure between the RSE server and the edge node is needed for requesting and transferring the computed routes for each demand. Alternatively, in the distributed approach the RSE module is duplicated on each edge node.

The route computation within the RSE takes into account the network topology and load, i.e., the amount of bandwidth (wavelengths) currently used on each link. Such information are collected in the so called network state database (NSD), which is local to the RSE. For each network link $m$, the NSD includes information about its capacity $u_m$ and the total reserved bandwidth $b_m^{tot}$. In an optical network each single traffic request will use an integer number of wavelengths on each link, therefore both $u_m$ and $b_m^{tot}$ can be expressed as integer numbers. The NSD needs to be updated following the reservation process. In the centralized approach, a unique NSD is associated to the centralized RSE and can be updated by the records of the RSE output itself. Instead in the distributed approach the NSDs associated to each edge node are fed by the flooding process of a link-state routing

protocol with traffic-engineering extensions as in Kompella and Rekhter [5].

After selected the routes, the edge nodes setup the service and backup circuits along the computed paths. The signaling procedures for all the implicated circuits can be run in parallel to speed up the total setup phase. During the circuit setup signaling, each intermediate node checks for local availability of the requested bandwidth. The eventual lack of resources along the computed paths can be due to race conditions or to inconsistency between the load information collected at the NSD and the real network state. This is particularly relevant in the distributed approach, due to the non-ideality of the flooding process. Furthermore, the adoption of flooding reduction techniques such as those proposed in Apostolopoulos et al. [19] and Shainkh et al. [20], which are needed to keep under control the flooding overhead, fatally increase the problem of potential inconsistency. The lack of available resources along the computed path will result in unsuccessful circuit setup. In this case, the edge node should tear-down the other service and/or backup circuits associated to the requesting demand. At this point the edge node must rejected the demand, or alternatively attempt a re-computation of a whole alternative set of disjoint paths. In this work we did not consider mechanisms for alternative re-computation, which are left for further study. Accordingly, in the successive simulations reported in Section 3 we always assumed that the unavailability of resources along a single path (service or backup) triggers the rejection of the demand. Regarding this point, an extensive comparison between the distributed and centralized implementation of our model, along with a quantitative assessment of the impact of flooding reduction techniques on the system performances can be found in a previous work [14].

In case of shared protection the intermediate nodes are also responsible for the evaluation of the bandwidth to be locally reserved for the shared backup circuits. In other words, the RSE computes the backup paths assuming that no bandwidth sharing can be applied for the new request, and it is the task of each intermediate node to decide about the actual amount of additional shared bandwidth to be reserved to support the new request. To accomplish that, certain information must be conveyed in the setup messages along the backup paths in order to identify the service path that was jointly computed. An addition to the

signaling protocol in support of this feature was already proposed in Lang and Rajagopalan [2]. A detailed description of the sharing mechanism for SFP and DFP are given in the Appendix.

This approach, inspired by the proposal made in Doverspike et al. [9] for single-fault protection, basically decouples the route selection (run by the RSE) from the sharing evaluation (run at the intermediate nodes). The main advantages of this approach are precision and robustness against information uncertainty. In fact, as the RSE does not consider the potential bandwidth sharing in selecting the backup path, it does not need to maintain the additional per-path information required to compute shared bandwidth on each link. Instead, such evaluation is done locally by intermediate nodes based on locally collected information, that can be easily maintained always updated. Furthermore, decoupling the route selection from the bandwidth sharing process fits well in a migration scenario where not all intermediate nodes implement backup bandwidth sharing. In fact, the handling of shared backup reservations requires additional capabilities to be installed at intermediate nodes, which are likely to be deployed node-by-node in an incremental fashion in an operational network. Finally, we remark that even in the case of shared protection applied to DFP only aggregate state information (i.e., per-link, not per-flow) has to be maintained at intermediate nodes, which is an important feature for preserving the scalability of the model.

## 2.4  The Route Selection Algorithm

For each incoming request, the RSE computes a number of paths between the end-nodes $\langle N_i, N_e \rangle$ based on the information collected at the local NSD database. The number of paths depends on the requested protection class: one, two or three, respectively, for UP, SFP and DFP demands. For protected demands the service and backup paths must be fault-disjoint, i.e., they cannot share a same link nor a same SRLG. Notably, two fault-disjoint paths may share a same NRL, as it is by definition not prone to any fault. In order to select the routes, the RSE needs a complete map of the network topology, included the complete knowledge of SRLGs and NRLs. This information can be embedded in the routing protocol, as already proposed in Kompella and Rekhter [5], and flooded to all network nodes. We believe this is the most convenient choice, as in case

of shared protection not only the edge nodes but also the intermediate nodes need such information. This is required for the exact computation of shareable backup bandwidth, as described in the Appendix.

The algorithm used by the RSE to find a set of fault-disjoint paths works as follows. First the topology graph $G$ is pre-processed by applying some basic graph transformations to each SRLG and NRL, if any. In force of such transformations, the modified graph, denoted by $G_{aux}$, is such that any set of link-disjoint paths on $G_{aux}$ corresponds to fault-disjoint paths on $G$. Afterwards, the RSE applies classical polynomial-time algorithms for finding link-disjoint paths are applied on $G_{aux}$, namely the Suurballe algorithm [21] and its extensions. Finally, link-disjoint paths found in $G_{aux}$ are reported back in the real graph $G$, resulting in a set of fault-disjoint paths. This procedure is sketched in Fig. 1, along with some exemplary transformations. The case A refers to the basic transformation applied to a single SRLG with a unique common node. The case B refer to the case of two SRLGs interlaced on a same link, and can be handled by applying twice the A transformation. Case C refer to the NRL transformation, and simply consists in the insertion of a supplementary link in association to the real NRL link: all the paths crossing the supplementary link in $G_{aux}$ are reported back on the real link in $G$. Finally note that this approach is not general, as some particular cases of SRLG in this way. For example, no transformation exist for the SRLG of case D, composed of two links with no common node. Such cases, if present in the network, should be handled by alternative approaches that have been left for further study (e.g., iterated Dijkstra on reduced topology, as in Doverspike et al. [9]). Further details about this technique can be found in Ricciato [13]. Notably the same approach was independently found in Ellinas et al. [18] limited to the case of protection from single fault, and without considering NRLs.

The path selection algorithm is designed in order to minimize and at the same time balance the overall resources usage. This is achieved by associating to each link a cost inversely proportional to the currently unused bandwidth, so that the less loaded links are preferred in the selection of the new paths. This helps in avoiding saturated links which could be critical for the allocation of future demands. Further details on the link metric can be found in Ricciato [13] and Ricciato et al. [14].
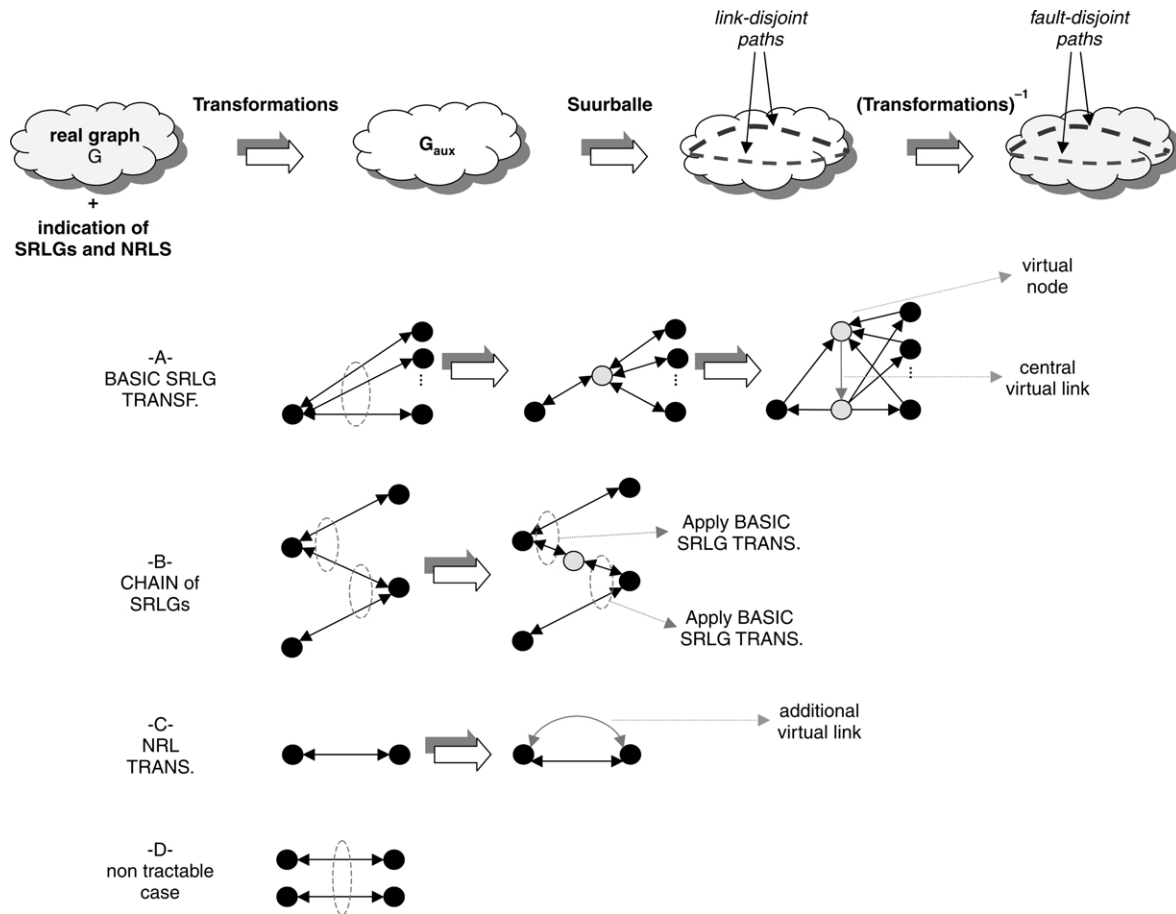
*F. Ricciato, M. Listanti, S. Salsano/Architecture for Differentiated Protection*



*Fig. 1.* Exemplary transformations for fault-disjoint route computation.

## 2.5 Functional Description of the Recovery Procedure

After established the service and backup circuits, the edge node is in charge of continuously monitoring their integrity, and react on the interruption of the current service ones. We assume that the generic ingress node $j$ is able to detect the interruption of any circuit originated by itself. This must be done for both established circuits (service circuits or dedicated backup circuits) and pre-qualified ones (shared backup circuits). A possible implementation could foresee an explicit notification message sent by the intermediate node local to the fault towards the edge nodes. Such notification could be either conveyed by the signaling protocol (as proposed in Lou Berger et al. [22]) or alternatively flooded by the routing protocol.

To achieve a fast recovery time, the reaction mechanism must not require any coordination between network nodes. In other words the requirement is that after detecting a network fault each node must know what to do without requiring communication with other network nodes.

Upon the interruption of the service circuit $P_s$ the ingress node must (i) do nothing for UP demands, (ii) switch the impacted SFP and DFP demands onto the respective primary backup path $P_{r1}$. In case that also the primary backup path is currently interrupted due to a previous fault, the ingress node must, (iii) do nothing for SFP demands, and (iv) switch the traffic onto the secondary backup path $P_{r2}$ for DFP demands only. Such simple reaction algorithm does not require any coordination between ingress nodes.

Let us now consider the case of shared protection. In case of single network fault it is guaranteed that enough resources are reserved on each network link to

accommodate all the required backup circuits, both for SFP and DFP demands. On the other hand, in case of two contemporary network faults, the amount of reserved backup bandwidth is enough to recover all the DFP demands plus some but not all the SFP ones. Furthermore, in case of more than two faults, there is no assurance to recovery all DFP demands neither. As a consequence, in case of multiple contemporary network faults there is a potential for conflicts on available resources during the activation of the shared backup circuits. A mechanism is needed to solve such conflicts in order to meet the recovery requirements listed above in Section 2.2. Our proposal is to apply preemption on the shared backup resources according to the following policy:

*Preemption policy:* In case of lack of available resources, a backup circuit (primary or secondary) for a Sh-DFP demand will preempt a backup circuit for a Sh-SFP demand.

In order to minimize service disruption we propose to always preempt the last activated Sh-SFP. It can be easily shown that such preemption policy ensures the service requirement listed above, and does not require any coordination between the network nodes. Therefore, the allocation policy of local backup resources as enforced by each intermediate node can be summarized as follows: requests are allocated on a first-come-first-served basis (independently from SFP versus DFP) until spare resources are available, then Sh-DFP requests preempt the last arrived Sh-SFP one. Note that resources used for dedicated protection are never preempted.

## 3   Simulation Results

In the previous section, we presented a dynamic protection model supporting a range of five different protection classes. However, there is no evidence of the convenience, from the business perspective, to support the full set of five classes within the same network infrastructure. The choice of the most convenient subset depends on the relative costs at the network level and performances at the service level of each class. In particular, one should consider for each class the following dimensions:

- *Complexity of the implementation.*

- *Recovery delay.*
- *Resources usage:* The average amount of bandwidth needed to accommodate a single demand.
- *Recovery level:* The probability of successful recovery under multiple contemporary faults.

Regarding the complexity of the implementation, we showed that shared protection requires additional capabilities to be installed at intermediate nodes, namely the maintenance of local state information (as specified in Appendix), handling of preemption, fast-activation of pre-qualified circuits. Whether or not these capabilities are available and their cost depends on the particular technology.

The recovery delay is in general larger for shared than for dedicated protection classes, as it includes the activation phase of pre-qualified backup paths. Again, the quantitative values of the recovery delay depends on technological aspects.

While the above two dimensions are related to technological and perhaps equipment-specific factors, the resource usage and the recovery level depend instead on the protection scheme and on the algorithms implemented in it. In order to quantify such metrics, we carried out extensive simulations on the network depicted in Fig. 2 (the same found in Irashko et al. [23] with the arbitrary addition of two more links to make the topology 3-connected). The link capacity was assumed fixed to 160 wavelengths. For sake of simplicity we did not include SRLGs nor NRLs, that is, all links are prone to simple failure. All demands require bidirectional wavelength-switched circuits, consuming exactly one wavelength on each link in both directions. Requests arrive randomly, with uniform intensity between each node pair, that is, the spatial traffic distribution is flat. The demand holding time was assumed infinite. For each experiment we loaded the network until the occurrence of the 10th rejection, which we call the ''stop-load point''. We considered it is as a reasonable working point for a network. In fact, it is a good compromise between a lightly loaded and a completely saturated network, which are two extreme situations of poor practical interest.

### 3.1   Analysis of Resource Usage
The resource usage can be quantified by a rather simple metric: the average amount of resources needed to accommodate a demand of class $H$, denoted
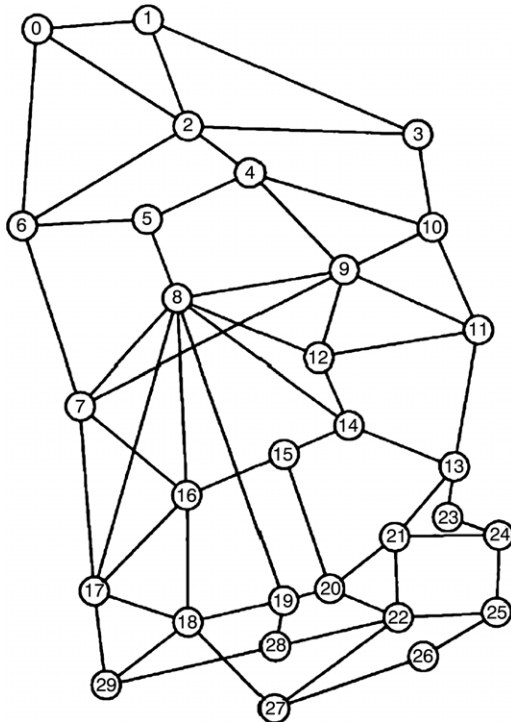
Fig. 2. Test-network used in simulations.

by $R_H$. Clearly, the value of $R_H$ depends on the network topology and on the spatial traffic pattern. In order to compare the relative resource usage for the different classes, we considered homogenous scenarios where all demands request the same protection class, and we run 100 different simulations for each scenario. Each time we measured (a) the total number of allocated demands and (b) the total amount of reserved wavelengths at the stop-load point. For each class $H$, the ratio between the number of wavelengths and the number of demands represents the average per-demand bandwidth usage. The results are reported in Table 2. It can be seen that unprotected demands

consume on average $R_0 = 6.46$ wavelengths, that means the mean hop-length of selected circuits in 3.23. By taking $R_0$ as a reference value, we can see (bottom line of Table 2) that Sh-SFP demands consume about 50% more bandwidth than UP demands, while De-DFP demands consume almost four times more. This last result is consistent with the findings in Clouqueur and Grover [16], that more than the triple than simple shortest-path resources are needed to support demands with complete double-fault protection. Furthermore, we note that the relative resource saving achievable with shared protection is greater for DFP than for SFP demands ($-45$ vs. $-34\%$). As a further remark, it can be seen that the cost of shared double-fault protection is about the same of dedicated single-fault protection in terms of resource usage. Such values should be carefully taken into account for example in the definition of the billing profile associated to each class.

## 3.2 Analysis of Recovery Level

Regarding the recovery level, let's denote by $\Pi_H(\varphi)$ the probability that a demand of class $H$ cannot be recovered when $\varphi$ contemporary network faults are present in the network. We will call $\Pi_H(\varphi)$ the unrecovery probability conditioned to $\varphi$. It can be seen that the larger number of alternative circuits for DFP versus SFP demands (3 vs. 2), coupled with the preemption policy introduced above, increases the level of service differentiation between the five proposed classes also under multiple ($> 2$) contemporary faults.

Consider a full-mix scenario where demands of different protection classes are contemporarily present in the network. Under a single network fault ($\varphi = 1$) service differentiation appears only between protected and unprotected demands: the former are guaranteed service continuity over the corresponding

Table 2. Resources usage for each different protection classes, as obtained by 100 different simulations with homogenous scenario on the test-network.

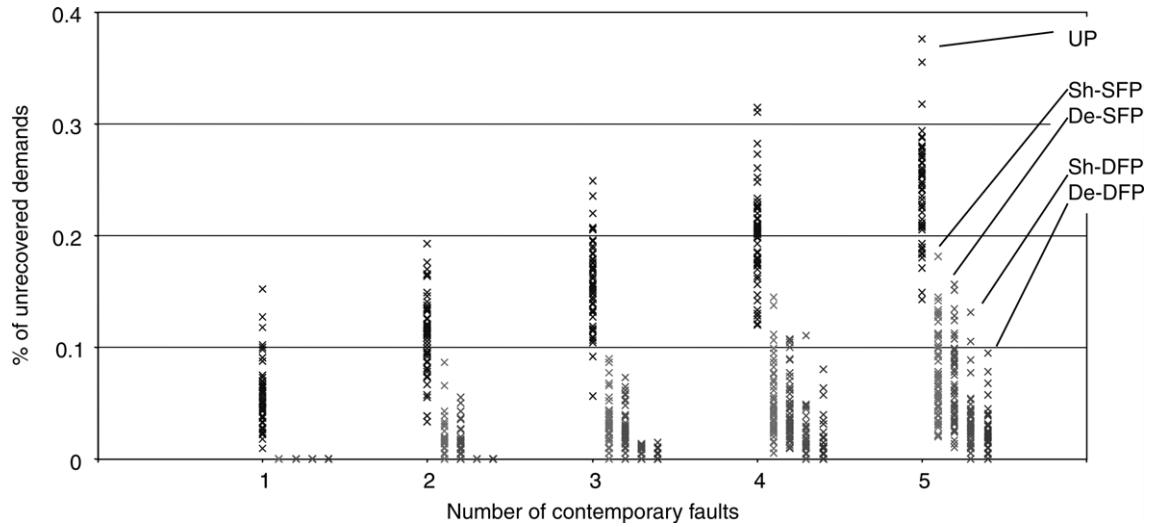|  |  | UP | SFP | | DFP | |
|---|---|---|---|---|---|---|
|  |  |  | Sh-SFP | De-SFP | Sh-DFP | De-DFP |
| Total accepted demands | Min-max | 1836–2125 | 1326–1552 | 918–1079 | 904–1085 | 538–612 |
|  | Mean | 1997 | 1433 | 996 | 1011 | 583 |
| Total used wavelengths | Min-max | 12,058–13,832 | 12,718–14,814 | 13,204–15,562 | 13,278–15,310 | 13,870–16,010 |
|  | Mean | 12,914 | 13,775 | 14,520 | 14,233 | 14,989 |
| $R_H$ = wavelengths/demands |  | 6.46 | 9.61 | 14.58 | 14.08 | 25.71 |
| $R_H/R_0$ |  | 1 | 1.49 | 2.25 | 2.18 | 3.97 |

*Fig. 3.* Percentage of non-recovered demands for each class in the full-mix scenario.

backup paths, the latter not. Under double network fault ($\varphi = 2$), differentiation emerges also between DFP and SFP demands, as only DFP ones are guaranteed service continuity. Nevertheless, it can be noted that differentiation is maintained between SFP and UP, as SFP demands are more likely to get service continuity than UP ones, as the former can potentially use two paths rather than just one. In principle, De-SFP demands are more likely to be recovered than Sh-SFP ones: in fact a De-SFP demand whose backup path is not affected by a fault is certainly recovered, while a Sh-SFP backup circuit has to potentially compete for resources with Sh-DFP and other Sh-SFP backup circuits. Under multiple ($\varphi > 2$) contemporary faults no protection class delivers guarantees of service continuity. However, similar considerations to those made for the case $\varphi = 2$ apply, resulting in an increasing level of recovery probability when moving from UP to SFP to DFP, and from shared to dedicated protection. In other words the service differentiation between the five considered protection classes can be interpreted in terms of service guarantees up to the case $\varphi = 2$, and in terms of recovery probability for $\varphi > 2$.

We are interested in evaluating the unrecovery probability under multiple faults for the different classes. To this scope, we run several simulations by loading the network described above with a balanced mixture of traffic. Each demand was associated to a randomly selected protection class, with uniform probability over the five possibilities. Therefore we

end up with a network loaded with 20% of traffic of each class on average. After reached the stop-load point, we simulated a series of five randomly selected link faults, and after each fault we counted the number of un-recovered demands for each class. These values are reported in Fig. 3 for 60 different simulations, as a fraction of the total number of demands in place for that class. The mean values of such percentages, reported in the top-most part of are taken as an estimate of the conditioned unrecovery probabilities $\Pi_H(\varphi)$.

As expected the unrecovery probability is null for SFP up to the first fault, and for DFP up to the second fault. An interesting result is that the differentiation between the recovery level of shared and dedicated protected demands is negligible. For ease of explication we need to introduce the following events associated to a generic allocated demand of class *H*:

- $A_H(\varphi)$ = all the service AND backup paths are interrupted by the current faults;
- $B_H(\varphi)$ = the service circuit is interrupted by the current fault AND there are no available resources along the backup path to establish the backup circuit.

With such notation, the probability $\Pi_H(\varphi)$ introduced above can be decomposed as follows:

- $\Pi_H(\varphi) = \Pr\{A_H(\varphi)\}$, for dedicated protection ($H \in \{$ De-SFP, De-DFP$\}$);

*Table 3*. Mean percentages of non-recovered demands for each class, for different mix-scenarios.

| | | | | Protection Class | | | |
| | | | | SFP | | DFP | |
| Mix Scenario | Total Demands | Fault ($\varphi$) | UP | Sh-SFP | De-SFP | Sh-DFP | De-DFP |
|---|---|---|---|---|---|---|---|
| *Scenario A* | | 1 | 5.76 | 0 | 0 | 0 | 0 |
| Full-mix, 20% traffic of each class | 1021 | 2 | 10.96 | 1.15 (0.0) | 0.92 | 0 | 0 |
| | | 3 | 15.51 | 2.53 (0.0) | 2.26 | 0.22 (0.0) | 0.26 |
| | | 4 | 20 | 4.94 (0.11) | 4.49 | 1.38 (0.1) | 1.18 |
| | | 5 | 24.3 | 7.47 (0.23) | 7.08 | 2.81 (0.27) | 2.25 |
| *Scenario B* | | 1 | 5.81 | — | 0 | — | 0 |
| 33% UP | | 2 | 10.57 | — | 0.77 | — | 0 |
| 33% De-SFP | 952 | 3 | 15.14 | — | 2.76 | — | 0.38 |
| 33% De-DFP | | 4 | 19.7 | — | 4.73 | — | 1.06 |
| | | 5 | 24.28 | — | 7.41 | — | 2.62 |
| *Scenario C* | | 1 | 5.39 | 0 | — | 0 | — |
| 33% UP | | 2 | 11.1 | 0.84 (0.0) | — | 0 | — |
| 33% Sh-SFP | 1371 | 3 | 15.71 | 2.45 (0.04) | — | 0.31 (0.05) | — |
| 33% Sh-DFP | | 4 | 20.54 | 4.93 (0.18) | — | 1.12 (0.12) | — |
| | | 5 | 25.08 | 7.83 (0.28) | — | 2.49 (0.25) | — |
| *Scenario D* | | 1 | — | — | 0 | — | 0 |
| 50% De-SFP | 742 | 2 | — | — | 0.78 | — | 0 |
| 50% De-DFP | | 3 | — | — | 2.31 | — | 0.27 |
| | | 4 | — | — | 4.41 | — | 0.9 |
| | | 5 | — | — | 6.69 | — | 2.09 |
| *Scenario E* | | 1 | — | 0 | — | 0 | — |
| 50% Sh-SFP | | 2 | — | 0.76 (0.0) | — | 0 | — |
| 50% Sh-DFP | 1199 | 3 | — | 2.31 (0.06) | — | 0.21 (0.0) | — |
| | | 4 | — | 4.73 (0.18) | — | 1.1 (0.12) | — |
| | | 5 | — | 7.36 (0.41) | — | 2.24 (0.28) | — |

- $\Pi_H(\varphi) = \text{Pr}\{A_H(\varphi) \cup B_H(\varphi)\}$, for shared protection ($H \in \{$ Sh-SFP, Sh-DFP$\}$).

For shared protection, Table 3 reports in brackets the fraction of demands that was not possible to recover due to lack of available wavelengths along the backup path. The small difference between the unrecovery probability of shared and dedicated suggests that the probability of event $B_H(\varphi)$ is in general very low. In fact, we verified in simulations that only a small portion of the resources available for shared backup circuits is effectively activated after the faults. For example, in the full-mix scenario after the 3rd fault only 3% of the links were using more than 50% of available wavelengths to shared backup paths, and less than 0.5% exceeded 80%. After the 5th fault such percentages reached 6.5% and 1.5%, respectively, with only about 0.5% of links using 100% of available resources.

In Fig. 4, we reported the average link bandwidth utilization with respect to the various components: (a)

used by service circuits, (b) reserved by dedicated backup circuits, (c) reserved for shared backup circuits, (d) free spare. In our model, we assumed that shared backup circuits can utilize bandwidth component (c) and (d), the latter representing a precious resource cushion for diminishing the blocking probability of shared versus dedicated backup circuits.

In a next series of simulations, we were interested in evaluating the impact of different traffic mix compositions on such results. In addition to the full-mix case (scenario A), we defined four alternatives scenarios where only a subset of the possible classes was present (Scenario B: UP/Sh-SFP/Sh-SFP; Scenario C: UP/De-SFP/De-SFP; Scenario D: Sh-SFP/Sh-SFP; Scenario E: De-SFP/De-SFP). In each scenario the traffic is evenly distributed between the present classes (two or three). In Table 3, we reported the estimated value of $\Pi_H(\varphi)$ up to the 5th fault, along with the mean number of totally allocated demands. Such values were averaged over 60 simulations. It can
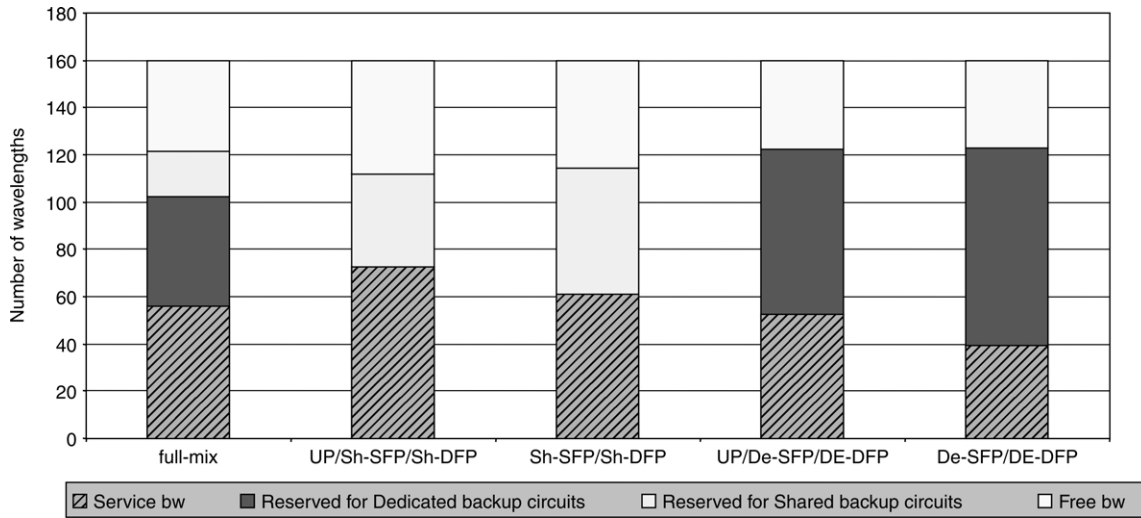
*Fig. 4.* Average profile of link-bandwidth allocation for different mix scenarios.

be seen that the values of unrecovery probability are poorly dependant on the particular traffic-mix. In particular it is interesting to compare the case B with C, and D with E. Where only shared protection is used (scenario C, E) the unrecovery probability for SFP and DFP are substantially similar to the case of dedicated protection (scenario B, D), but of course the number of allocated demands is sensibly higher: $+60\%$ from scenario D to E, $+30\%$ from scenario B to C. The conclusion is that shared protection achieve about the same recovery level of dedicated protection under multiple faults, but with a more efficient resource usage. Therefore, it is reasonable to expect that for both SFP and DFP demands the increase in network capacity yielded by shared protection pays-off its higher implementation complexity.

## 4 Conclusions

In this work, we proposed a dynamic model for differentiated protection in an optical network. Our model supports a range of up to five different protection schemes, including protection from single and double fault, with and without bandwidth sharing. The underlying algorithms are suited to be implemented in a distributed fashion.

We analyzed the performances of each protection class, in search of guidelines about the choice of the more convenient subset of classes to be supported by a

single network. We considered the metrics of resource usage and recovery level, in addition to the recovery delay and implementation complexity. In particular we investigated by simulation the first two metrics, as they depend directly from the protection scheme and associated algorithms, while the latter two are related to the equipment-specific technology.

Our findings globally show that the differentiation between shared and dedicated protection is negligible under the point of view of the recovery level. Instead, they could present an appreciable differentiation regarding the recovery delay, but this depends on the equipment-specific technology. On the other hand, with shared protection the resource usage is sensibly lower, leading to a higher network capacity in terms of allocated demands. Therefore, the contemporary coexistence of dedicated with shared protection schemes on the same network would be only justified in the case that sensible differences hold in the recovery delay for the two schemes with the particular network equipment.

In any case, a sensible differentiation hold at the service level between UP, SFP and DFP demands with respect to the recovery level performances. Such differentiation hold in terms of deterministic service continuity guarantees up to two contemporary faults. Even under multiple ($>2$) faults, sensible differences in terms of unrecovery probability still hold. From Table 3 it can be seen that the difference in the unrecovery probability between DFP and SFP is one order of magnitude at the 3rd fault (0.26% vs.

26%), but the distance diminishes with successive faults.

Whether such differences pay-off the major resource consumption associated with DFP versus SFP schemes depends on the criticality of the supported traffic. It is reasonable to expect that DFP schemes are a convenient choice for a minor portion of highly critical traffic in those networks where the probability of dual contemporary faults can not be neglected, typically large wide-area networks and/or networks with highly exposed physical deployment. In such cases, it becomes more compelling to adopt shared protection mechanisms that, still more complex to implement, allow almost to halve the resource consumption for DFP demands.

## Appendix. Algorithm for Shared Protection

Each network node $n$ maintains the following state information for each outbound link $m$ attached to it: the amount of bandwidth (i.e., the number of wavelengths) currently allocated to service circuits $b_m^s$, to dedicated backup circuits $b_m^d$, to shared backup circuits $b_m^r$, and the total link capacity $u_m$. The total reserved bandwidth is defined at any time as $b_m^{tot} = b_m^s + b_m^d + b_m^r$. The components $b_m^s$ and $b_m^d$ are updated during the signaling procedures in a very simple way: they are incremented (decremented) by $B$ at every circuit setup (release), where $B$ is the bandwidth requested by the demand. Instead the determination of the component $b_m^r$ relevant to the shared backup bandwidth requires the maintenance of additional data structures. In this subsection, we provide a detailed description of such structures and the related update algorithms run during the signaling procedures. This scheme basically extends to DFP what proposed in Quiao and Xu [7] for SFP. In the following $E$ will denote the number of possible fault events (link, node and SRLG failures) that are considered for the specific network, and $M$ the total number of network links. It is reasonable to expect that the values of $E$ and $M$ are close to each other, or at least in the same order of magnitude.

In order to implement Sh-SFP, a single data structure must be maintained by node $n$ for each outbound link $m$: the vector $FS_m$ of size $E$, whose generic component $FS_m[k]$ represents the bandwidth needed on link $m$ to carry all the Sh-SFP traffic rerouted on $m$ after the single fault event $k$. Obviously,

$FS_m[k] = 0$ if $k$ affects $m$ itself. The vector $FS_m[k]$ is maintained by letting the ingress node include in the signaling messages along the backup path the list $V_s$ of the links constituting the associated service path. From $V_s$ and from the knowledge of the SRLGs associated to each network link the intermediate node can easily derive the list $W_s$ of the faults affecting the service path. During the backup setup phase, the intermediate node will increment by $B$ the components $FS_m[k]$ for each $k \in W_s$. The same information $V_s$ should be advertised also during the circuit release phase, in order to decrement the relevant components of $FS_m[k]$.

The advertisement of $V_s$ requires only minor additions to the signaling protocol, that were already suggested in Lang and Rajagopalan [2], while the association between links and SRLGs, and eventually the NRLs, can be distributed by the routing protocol accordingly to Kompella and Rokhter [5].

In order to implement Sh-DFP two different data structure must be maintained by node $n$ for each outbound link $m$: the vector $FD1_m$ and the $E \times E$ matrix $FD2_m$. The vector $FD1_m$ for Sh-DFP has exactly the same meaning as $FS_m$ for Sh-SFP. Additionally, the generic component $FD2_m[k_1, k_2]$ represents the number of wavelengths needed on $m$ to support the demands whose service and primary backup paths are interrupted by the fault events $k_1$ and $k_2$, respectively. The vector $FD1_m$ is updated in the same way as $FS_m$, therefore the ingress node has to advertise the list $V_s$ of the links constituting the service path in the signaling messages along the primary backup path. On the other hand, along the secondary backup path it will advertise both the lists $V_s$ and $V_{r1}$, the latter referring to the links constituting the primary backup path. Similarly to above, the intermediate node $n$ will derive the fault lists $W_s$ and $W_{r1}$ from the link lists $V_s$ and $V_{r1}$, and during the setup (release) phase will increment (decrement) by $B$ the component $FD2_m[k_1, k_2]$ for each $k_1 \in W_s, k_2 \in W_{r1}$.

At every update of these data structures, the new value of the bandwidth reserved to shared backup circuits $b_m^r$ is computed according to the following allocation rule:

$$b_m^r = \max_{\substack{k_1, k_2 \\ k_1 \neq k_2}} \{FS_m[k_1] + FD1_m[k_1] + FD1_m[k_2] \\ + FD2_m[k_1, k_2] + FD2_m[k_2, k_1]\}. \qquad (1)$$

It can be shown that with this allocation rule the amount of shared backup resources is slightly over-estimated with respect to the minimum amount needed to meet the service requirements listed in Section 2.2. This is due to the fact that the information included in the data structures introduced above is partial and aggregated: $FS_m$, $FD1_m$ and $FD2_m$ enclose global information about the per-link reserved resources, not about the full set of per-demand paths. This information is enough for evaluating exactly the bandwidth needed to protect Sh-SFP demands against a single network fault. On the other hand the lack of complete information fatally leads to an imprecise evaluation of the minimum bandwidth needed to protect all DFP demands against a double network fault. Consider for example a DFP service circuit $P_s$ with associated bandwidth $B$ spanning two links $a_1$ and $a_2$, impacted by faults $k_1$ and $k_2$, respectively. This circuit is counted twice in Equation (1) in the $FD1_m[k_1] + FD1_m[k_2]$ term, therefore for such demand the generic node along the primary backup circuit will reserve $2B$ bandwidth while just $B$ would suffice. More formally, let us denote by $D_i (i = 1, 2)$ the set of demands whose service path includes link $a_i$, and by $D_{1,2} = D_1 \cap D_2$ the set of demands whose service path include both $a_i$ and $a_2$. Denote by Size $(D)$ the sum of bandwidths associated to the demands in the set $D$. Consider a third link $m$: the allocation rule in Equation (1) will compute $b_m^r$ so as the amount of shared reserved bandwidth equals $\text{Size}(D_1) + \text{Size}(D_2)$, while just $\text{Size}(D_1 \cup D_2)$ would suffice: therefore it over-reserves $\text{Size}(D_{1,2})$ bandwidth as demands in $D_{1,2}$ are counted twice. Note that this phenomenon does not apply to Sh-SFP, but exclusively to Sh-DFP. In order to quantify such inefficiency we compared the global amount of bandwidth reserved by our scheme with the minimum amount needed to protect all the demands for any possible pair of faults, which was evaluated by simulating all the possible fault pairs and considering the worst-case. We found that our scheme reserves about four 6.5% more backup wavelengths than the minimum required, corresponding to about 2 4% more total wavelengths. In order to eliminate such small inefficiency, the intermediate node $n$ should maintain complete per-demand information, i.e., the vectors $W_s$ and eventually $W_{r1}$ for all the demands having a shared backup path routed through it. This would increase the amount of state information required at each network node, and add complexity

(and time) to the computation of $b_m^r$. On the other hand the small values of the achievable gain mitigate the interest towards further refinements of the sharing mechanism exploiting complete state information.

## References

[1] J. P. Lang, J. Drake, Mesh network resilience using GMPLS, Proceedings of the IEEE, (Sept. 2002).

[2] J. P. Lang, B. Rajagopalan, Generalized MPLS recovery functional specification, draft-ietf-ccamp-gmpls-recovery-functional-00.txt, (Jan. 2003), work in progress.

[3] D. Papadimitriou, E. Mannie, Analysis of generalized MPLS-based recovery mechanisms (including protection and restoration), draft-ietf-ccamp-gmpls-recovery-analysis-00.txt, (Jan. 2003), work in progress.

[4] D. Papadimitriou, E. Mannie, Recovery (protection and restoration) terminology for GMPLS, draft-ietf-ccamp-gmpls-recovery-terminology-01.txt, (Nov. 2002), work in progress.

[5] K. Kompella, Y. Rekhter, Routing extensions in support of generalized MPLS, draft-ietf-ccamp-gmpls-routing-5.txt, (Aug. 2002), work in progress.

[6] D. Zhou, S. Subramaniam, Survivability in optical networks, IEEE Network, (Nov./Dec. 2000).

[7] C. Quiao, D. Xu, Distributed partial information management (DPIM) scheme for survivable networks—Part I, INFOCOM (2002).

[8] T. V. Lakshman, M. Kodialam, Dynamic routing of bandwidth guaranteed tunnels with restoration, INFOCOM (2000).

[9] R. Doverspike, C. Kalmanek, G. Li, D. Wang, Efficient distributed path selection for shared restoration connections, INFOCOM (2002).

[10] E. Boulliet, J. F. labordette, G. Ellinas, R. Ramamurthy, S. Chadhuri, Stochastic approaches to route shared mesh restored lightpaths in optical mesh networks, INFOCOM (2002).

[11] A. Autenrieth, A. Kirstädter, Engineering end-to-end IP resilience using resilience-differentiated QoS, IEEE Communications Magazine, (Jan. 2002).

[12] W. D. Grover, M. Clouqueur, Span-restorable mesh network design to support multiple quality of protection (QoP) service-classes, 1st International Conference on Optical Communications and Networks (ICOCN'02), (Singapore, Nov. 11–14, 2002), pp. 321–323.

[13] F. Ricciato, An architecture for dynamic differentiated end-to-end protection for connection-oriented networks, CoRiTeL internal report (available at ftp://ftp.coritel.it/pub/Publications/DynamicDifferentiatedProtection.pdf).

[14] F. Ricciato, M. Listanti, D. Perla, Performance evaluation of a distributed scheme for protection against single and double faults for MPLS, in Proceedings of 2nd International Workshop on QoS in Multiservice IP Networks (QoS-IP), Milan, (February 2003). Edited by Springer in Lecture Notes in Computer Science vol. 2601.

[15] M. Clouqueur, W. D. Grover, Dual-failure availability analysis of span-restorable mesh networks, IEEE JSAC Special Issue on Recent Advances in Fundamentals of Network Management, vol. 20, no. 4, (May 2002), pp. 810–821.

132                    *F. Ricciato, M. Listanti, S. Salsano/Architecture for Differentiated Protection*

[16] M. Clouqueur, W. D. Grover, Mesh-restorable networks with complete dual failure restorability and with selectively enhanced dual-failure restorability concepts, Opticom 2002, Boston, USA.

[17] H. Choi, S. Subramaniam, H. A. Choi, On double-link failure recovery in WDM optical networks, INFOCOM (2002).

[18] G. Ellinas, et al., Routing and restoration architectures in mesh optical networks, Optical Networks Magazine, vol. 4, no. 1, (Jan./Feb. 2003).

[19] G. Apostolopoulos, R. Guerin, S. Kamat, S. K.Tripathi, Quality of service based routing: A performance perspective, SIGCOMM (1998).

[20] A. Shainkh, J. Rexford, K. G. Shin, Evaluating the overheads of source-directed quality-of-service routing, Int'l Conference on Network Protocols (ICNP), (1998).

[21] J. W. Suurballe, R. E. Tarjan, A quick method for finding shortest pairs of disjoint paths. Networks, vol. 14, (1984), pp. 325–336.

[22] Lou Berger (ed.) et al., Generalized MPLS signaling-RSVP-TE extensions, RFC 3473.

[23] R. R. Irashko, W. D. Grover, M. H. MacGregor, Optimal capacity placement for path restoration in STM or ATM mesh-survivable networks, IEEE/ACM Transactions on Networking, (June 1998).

**Fabio Ricciato** received his ''Laurea'' degree in Electronics Engineering in 1999 and his ''Dottorato di Ricerca'' (Ph.D. degree) in Information and Communications Engineering in 2003, both from the University of Roma ''La Sapienza''. Since November 2002 he has been a research fellow in the Networking group of the INFOCOM Department. During 2000–2003, he collaborated with CoRiTeL (a research consortium sponsored by Ericsson Italy) and participated to the European research project called AQUILA. His research interests include transport networks technologies (IP, MPLS, Optics) and related topics (QoS, routing, traffic engineering, fault protection, network planning, GMPLS).

**Stefano Salsano** received his ''Laurea'' degree in 1994 (University of Rome ''Tor Vergata'') and his Ph.D. in 1998 (University of Rome ''La Sapienza''). From November 2000 he has been assistant professor at the University of Rome ''Tor Vergata''. From 1997 to 2000 he was with CoRiTeL, a research institute on telecommunications, where he has been coordinator of the research activities in the IP related area. He participated in several research projects founded by the EU (INSIGNIA, ELISA, AQUILA), by the European Space Agency and by the Italian Ministry of Research.

His current research interests include QoS and traffic engineering in IP networks, IP telephony, MPLS, IP over optics.

**Prof. Marco Listanti** received his ''Laurea'' degree in Electronics Engineering from the University ''La Sapienza'' of Rome in 1980. He joined the Fondazione Ugo Bordoni in 1981, where has been leader of the group ''TLC network architecture'' until 1991. In November 1991 joined the INFOCOM Department of the University of Roma ''La Sapienza'', where he is Full Professor of Switching Systems. Since 1994, he has collaborated with the Electronic Department of the University of Rome ''Tor Vergata'' where he holds courses in telecommunication networks. He is author of several papers published in the most important technical journals and conferences in the area of telecommunication networks and has been guest editor of the feature topic ''Optical Networking Solutions for Next Generation Internet Networks'', on *IEEE Communications Magazine*. His current research interests focus on traffic control in IP networks and on the evolution of techniques for optical networking.

Prof. Listanti has been representative of Italian PTT administration in international standardization organizations (ITU, ETSI) and has been coordinator of several national and international research projects (CNR, MURST, RACE, ACTS, ICT). He is also a member of IEEE Communications Society.