Performance Analysis of a Traffic Engineering Solution for Multilayer Networks Based on the GMPLS Paradigm

Paola Iovanna, *Member, IEEE*, Marina Settembre, *Member, IEEE*, Roberto Sabella, *Senior Member, IEEE*,
Giulia Conte, and Lucia Valentini

***Abstract*—This paper presents an integrated traffic engineering (TE) system for new generation multilayer networks based on the generalized multiprotocol label switching (GMPLS) paradigm, and reports the performance analysis of such a system. The proposed TE system aims at dynamically reacting to traffic changes and, at the same time, fulfilling quality of service (QoS) requirements for different classes of service. The proposed solution consists of a hybrid routing approach and a bandwidth management strategy. The former makes use of both off-line and on-line methods to accommodate traffic requests. The latter, based on an "elastic" use of the bandwidth, allows the handling of different priorities among data flows, possible preemptions, and rerouting. The proposed TE permits the accommodation of the largest amount of traffic, while guaranteeing good performance to mission-critical services. The main building blocks and the operations of the system are reported and the major advantages are discussed. The performance of the proposed system are compared with the ones relating to a reasonable alternative system based on overprovisioning, to highlight its advantages of in terms of traffic volume that can be accommodated for a given network infrastructure.**

***Index Terms*—Generalized multiprotocol label switching (GMPLS), multilayer networks, routing, traffic engineering.**

I. INTRODUCTION

IT IS generally accepted that traffic will be increasingly dominated by Internet-based services, with respect to traditional voice traffic, owing to the increased adoption of high-speed access technology and the migration of more and more services toward the Internet Protocol (IP). However, IP is a connectionless, best-effort technology that was not designed for voice or any other real-time service. Moreover, the peculiar characteristics of Internet traffic [1], such as its unpredictability and instability, demand for new requirements for next-generation networks (NGNs): flexibility and ability to promptly react to traffic changes. Overprovisioning, which is the common solution to the problem of unpredictable bottlenecks in nowadays telecom networks, it is not a cost-effective solution for new generation networks. Moreover, the migration of all services over IP, including the real-time ones, requires guaranteeing quality of service (QoS) for a subset of services that should be comparable to that one provided by the telecom-based networks nowadays.

As a result, NGNs will have to be IP-centric, provide multiservice capabilities, which means being able to support several types of traffic with different requirements in terms of QoS [2], and be flexible and dynamic enough to use at the best their resources. Traffic engineering (TE) plays a key role to cope with these challenging requirements [3]. A promising solution to actualize TE in NGN is given by the generalized multiprotocol label switching (GMPLS) paradigm [4]. GMPLS extends the features of the well-known MPLS technique [5], [7] to both packet and circuit switching network, providing a common set of IP-based protocols to control heterogeneous network such as ATM, SONET/SDH, and WDM [8], [9]. However, in practice, the definition and analysis of a TE strategy exploiting the capabilities of GMPLS is a very challenging task.

Many papers deal with specific TE functions such as routing, wavelength assignment, and preemption algorithms [10]–[13] in an optical layer, possibly overlaid to an electrical layer. In those papers the two subproblems of 1) design of logical topology of the optical network (i.e. the set of wavelength paths) and 2) the routing of the data flows at the IP/MPLS layer onto the logical topology, are solved in a separate way (e.g., in two different steps). Differently, a multilayer approach would consist in simultaneously solving these two subproblems. A TE strategy involving and combining specific TE functions in a multilayer network has been reported in [14]. That paper also presented the main building blocks and the mode of operations, discussing the main characteristics of the system as a whole. Key building blocks of that solution were reported in [15]–[17]. In particular, the routing problem has been approached in a multilayer fashion, in the aforementioned sense. The present paper reports for the first time, to the best of our knowledge, an integrated solution of TE in the technical details, the performance analysis of this TE system as a whole, aiming at both assessing its feasibility and evidencing its advantages with respect to a relevant example among traditional overprovisioning approaches.

This paper describes in Section II the reference network scenario, and in Section III the realization of the network solution previously reported in [14], addressing the above-mentioned issues, by exploiting the GMPLS network model, in a multilayer scenario. Section IV reports the results of the performance analysis accomplished by means of a simulation tool. Conclusions and perspectives for future works are discussed in Section V.

Manuscript received May 27, 2003; revised March 3, 2004.

P. Iovanna, M. Settembre, and R. Sabella are with Ericsson Lab Italy, 00040 Rome, Italy (e-mail: paola.iovanna@ericsson.com).

G. Conte and L. Valentini are with CoRiTel-Research Consortium on Telecommunications, 00040 Rome, Italy.

Digital Object Identifier 10.1109/JSAC.2004.833839

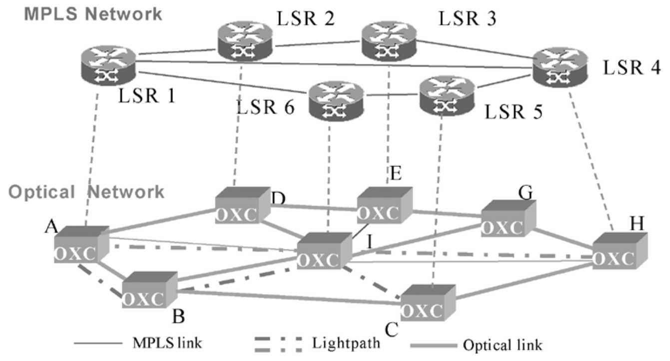

Fig. 1. Multilayer reference network scenario.

II. Reference Network Scenario

It is widely recognized that MPLS technology, together with proper constraint-based routing solutions, enables advanced TE capabilities and support of QoS in an IP-based network [18]. In fact, MPLS allows explicitly routing a traffic request through the network by forcing it on a specific path according to user and network constraints and reserving the resources for that path. Basically, MPLS re-proposes the concept of virtual connection, previously introduced with ATM, but adopting IP-based signaling and reservation protocols [19]–[21]. The virtual connection established by MPLS is called label switched path (LSP). The MPLS routers achieving label switching are called label switched routers (LSRs). An LSP can be set up, torn down, rerouted if needed, and modified by means of the variation of some of its attributes, including the bandwidth [21]. Furthermore, preemption mechanisms on LSPs can also be used in order to favor higher priority data flows at the expenses of lower priority ones, to avoid congestion in the network [22]. Another important feature of MPLS relates to the possibility of stacking labels that provides the means of nesting an LSP into another one of higher hierarchical level [5].

GMPLS extends the features of the MPLS technology [8], [9], [23] In particular, it can manage heterogeneous network elements, such as IP/MPLS routers, ATM switches, SDH/SONET elements, or even optical elements, using a suitably extended version of well-known IP protocol suite [24]. In fact, GMPLS allows a single control plane to handle heterogeneous LSPs [25], [26]. This means that a single instance of the control plane can span multiple technologies, and an LSP of low order can be tunneled into an already existing LSP of higher order that acts as a link. For sake of simplicity, but without losing generality, in this paper a two-layer network is considered as reference scenario. It consists of an IP/MPLS layer, whose network elements are LSRs, and a WDM transport layer, whose nodes are optical cross-connects (OXCs), as depicted in Fig. 1. Thus, just two types of LSPs are considered: MPLS LSPs and optical LSPs, which are usually called "lightpaths"; where a lightpath bundles several MPLS LSPs.

From a routing perspective, GMPLS is adopted to provide flexibility and efficiency in the use of network resource. In fact, GMPLS can exploit constraint-based routing (CBR) concept, already developed in MPLS based networks, and multilayer routing. CBR allows the calculation of the LSP routes taking into account of network status and user constraints (e.g., the actual link occupancy and the bandwidth requirement) by means of an extended routing protocol (e.g., OSPF-TE). Hence, CBR may find longer but less-congested paths instead of heavily loaded shortest paths, leading to a more uniform traffic distribution through the network and preventing congestions. Multilayer routing allows leading to a more effective use of network resources, considering the MPLS layer and the optical layer jointly. That means that in a single routing instance an LSP can be routed on a concatenation of optical paths.

Moreover, GMPLS can take advantage of suitable extension of signaling protocol, (e.g., RSVP-TE), to allow the reservation of network resources, and of priority mechanisms to assign resources to higher priority LSPs at expense of lower priority LSPs, in order to efficiently handling QoS support.

III. Traffic Engineering System for New Generation Multilayer Networks

The main goals of TE in new generation networks are the optimization of the use of network resources, the actualization of the "bandwidth-on-demand" concept, and the support of different classes of service by guaranteeing the required QoS. The proposed TE system aims at fulfilling those objectives, by means of a hybrid routing approach, based on off-line and on-line methods, and of a bandwidth engineering system that adopts an "elastic" use of the bandwidth resource and priority mechanisms.

The motivations for a hybrid routing originate from both traffic characteristics and practical implementation aspects relating to routing. Traffic entering a network can vary with time, both in predictable and unpredictable ways. For instance, legacy traffic carrying traditional telephone services is easily predictable, using well-known models, while Internet traffic is not. In general, the former type can be efficiently accommodated through an off-line routing approach, which is adequate for achieving a global optimization of route calculation based on a foreseen traffic matrix, particularly when a multilayer approach is adopted. Such an optimization of network resources requires long computational time, which increases with the network and traffic size. On the other hand, Internet traffic is quite unpredictable and unstable. In this case, a pure off-line approach can result unsatisfactory. In fact, the foreseen traffic matrix could strongly mismatch with the actual traffic entering the network. Overprovisioning, which is a common solution to the problem of unpredicted bottlenecks, does not seem a viable and cost-effective solution for the new generation IP networks, since it could lead to book a huge amount of network resources.

Notwithstanding, a pure on-line routing approach, which consists in evaluating the routes "on-demand," is more adequate to promptly react to traffic changes, but it does not lead to the same efficient use of the network resources as in the case of off-line approach, since it does not provide a global optimization.

As a result, the proposed hybrid routing solution combines the off-line and the on-line methods to efficiently manage both predictable and unpredictable components of traffic.

In the presence of more than one class of service, the flexibility provided by the hybrid routing can be enhanced by means of the module called bandwidth engineering (BE). The BE allows better exploiting network resources by taking advantage of an “elastic” use of the bandwidth and suitable priority and rerouting mechanisms, while at the same time fulfilling QoS requirements. In practice, the BE functions operate so that the temporarily unused reserved bandwidth of a higher priority LSP can be released and put at disposal of lower priority requesting LSPs, provided that the bandwidth is given back to higher priority LSP when needed. In other words, the bandwidth attribute of any existing LSP can be varied on-demand according to specific traffic requests, leading to an elastic bandwidth attribute. As soon as higher priority traffic needs the released bandwidth, a procedure that handles preemption of lower priority LSPs is activated. Moreover, the rerouting procedure can be used to move lower priority traffic on less-congested available routes, in order to serve as much traffic as possible. Essentially, the BE module accomplishes its functionalities by means of bandwidth modify mechanisms, preemption algorithms, and rerouting operations according to a defined priority policy.

For practical purposes, in the rest of the paper two main groups of LSPs are identified. The LSPs belonging to the first group relate to the traffic with very tight QoS requirements, and they can be referred as higher priority (HP) LSPs. HP LSPs are guaranteed at any time and in any traffic conditions, whatever is their bandwidth attribute, up to the maximum value previously agreed by the SLA. The HP traffic carries presumably mission-critical services, such as voice and video communications. The nature of this traffic and the fact that it is regulated by the SLA render it more easily predictable. Thus, the network operator has to infer the traffic matrix associated to this traffic taking care to dimension the connection requests on their peaks resulting by the SLA. Then HP LSP routes are calculated by means of the off-line procedure. On the other hand, the LSPs belonging to the second group, relating to all the other types of lower priority data flows can be referred as lower priority (LP) LSPs. The traffic carried by LP LSPs can be of various types of possible Internet services, with different QoS requirements. That traffic is much more unpredictable, and can be estimated by means of statistical evaluations and measurements. In addition, the LP LSPs are not guaranteed and can be preempted if they are using the bandwidth required by the HP traffic. In this case, the network operator has to infer the resulting portion of traffic matrix according to its specific policy. The issue stays in dimensioning each traffic request on a reasonable average estimation, without making an expensive overprovisioning of the network resources. The way this is done is above the scope of the present paper.

The considered TE solution employs the off-line procedure to configure the optical and the MPLS connections, basing on the traffic matrix that represents the requested connections for both HP and LP LSPs.

The TE system is designed to serve on-demand both HP and LP traffic, with a difference: HP traffic routes, once calculated, remain fixed during their life unless another off-line procedure is activated; while LP traffic routes can be dynamically changed from their originally assigned routes, according to the actual network status. In order to efficiently use the bandwidth capacity, the TE strategy allows that HP traffic consumes only the amount of bandwidth that it really needs and for the time it is necessary, and temporarily releases the unused bandwidth to LP traffic. Thus, bandwidth modify operations are dynamically performed for HP traffic, by means of the BE function that preempts those LP-LSPs to make available the required bandwidth for HP-LSP and tries to reroute the removed LP-LSPs on less-congested routes. Bandwidth modify operations referring to a bandwidth decreasing request are achieved by known modify MPLS mechanisms and the released bandwidth is put at disposal for accommodating other requests.

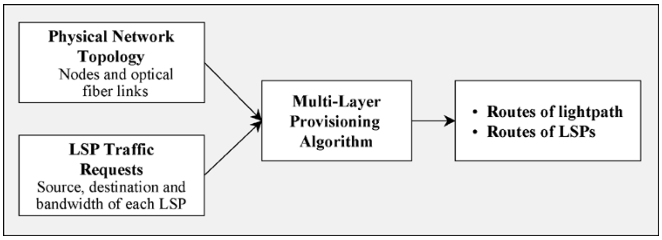

Fig. 2. Sketch of the provisioning module.

The description of the integrated TE system is reported in the following sections, where the hybrid routing solution and the bandwidth-engineering module are reported.

A. Hybrid Routing Solution

1) Off-Line Routing: The Global Path Provisioning: The off-line routing is actualized by the global path-PRovisioning (PR) module, whose input and output are schematically sketched in Fig. 2. Essentially, the PR module designs the optical logical topology and calculates the LSP routes, according to foreseen LSP traffic requests and to the physical topology of the network.

The foreseen traffic requests represent an input data. It is assumed that the network operators infer the traffic requests for both HP and LP LSPs through the knowledge of two types of information: agreements stipulated with clients and estimations made through statistical evaluations.

The physical topology of the optical network, assumed to be set during the network planning phase, is composed of a set of nodes connected by a set of links in a given mesh topology. Each link bundles a set of fibers between two adjacent nodes, and a single fiber carries a certain number of wavelengths. Each node can consist of either an LSR integrated with an OXC or a standalone OXC. The OXCs are assumed to have full wavelength conversion capability, without losing generality.

The output of the PR module consists of the set of lightpaths (i.e., optical LSPs according to GMPLS LSP hierarchy) that represent the logical topology of the optical layer, and the routes for all the LSPs groomed into the lightpaths of the logical topology.

Different objective functions can be defined for the path-provisioning problem according to the network operator policy. For instance, the objective function could be the maximization of the efficiency of network resources consumption (optical resources, electrical resources or both), the minimization of the traffic lost, or the average packet hop distance [11], [12]. The specific objective function considered here is the minimization of the congestion on the network resources. Formally, it is defined as the maximum ratio between used and available resources over all the optical resources, that is, wavelengths on each optical link, ports incoming to each LSR node, and ports outgoing from each LSR node. The rationale is that minimizing congestion facilitates the dynamic routing operations; that is the accommodation

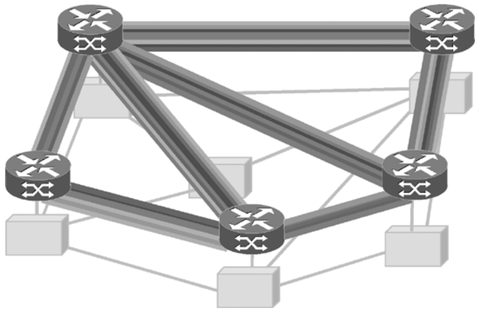

Fig. 3. Logical topology.

of new connection requests or the handling of fluctuations of the traffic demand.

The PR algorithm operates in a multilayer fashion. This means that the selection of the lightpaths and the calculation of their routes on the physical topology are performed *concurrently* with the calculation of the LSP routes on the logical topology [15]. Obviously, solving the provisioning problem with a multilayer approach increases the complexity of the algorithm, but it leads to a more efficient use of network resources.

A heuristic procedure has been used to implement the algorithm in the simulations. However, the "goodness" of the adopted method (i.e., the estimation of the distance between the obtained solution and the optimal one) has been tested by means of a comparison with an algebraic algorithm, achieved by solving the ILP formulation of the problem through the optimization solver CPLEX under different conditions. The results are reported in [15].

2) On-Line Routing: The Dynamic Path Selection: The DR module evaluates "on-line" the route for the entering LP LSP request, expressed in terms of source and destination nodes, and bandwidth requirements, taking into account the updated link state status of MPLS and WDM layers. In the proposed TE, it is also used to reroute LSPs that have been preempted by the bandwidth engineering operations.

For sake of simplicity, it has been assumed that the DR module cannot set up new lightpaths, but it can only operate on the logical topology derived during the off-line provisioning phase. As a result, the establishment of one or more lightpaths can only follow the decision of off-line providing a greater logical capacity to the network (e.g., when a new Internet Service Providers enters the network).

The logical topology provided by the PR module is enriched with the information of bandwidth availability on each logical link and on each lightpath constituting the logic link, learned by a suitable extension of signaling protocol (e.g., OSPF-TE). In fact, as shown in Fig. 3, each logic link between two LSRs is constituted by a set of lightpaths connecting the two OXCs integrated with those MPLS nodes.

The DR module aims at better utilizing network resources, by using less-congested paths instead of shortest, but heavily loaded paths. In order to accomplish this, the DR algorithm has to concurrently satisfy two criteria:

1) finding the shortest route that minimizes congestion, evenly distributing the traffic at MPLS layer;
2) selecting the lightpath in the logic link, privileging the choice of more filled wavelengths in order to facilitate the accommodation of subsequent requests with more severe bandwidth requirements.

The two criteria can be fulfilled by using a shortest path algorithm with a weight function that takes into account of number of hops between the source and destination nodes in the MPLS layer, the capacity availability in logical links, and the capacity availability on each lightpath in the logical link.

The weight function adopted in this paper has been derived by extending the least resistance routing weight method [10] to our GMPLS reference scenario, leading to the following formula:

$$w(i) = \frac{C^T}{C_i^A} + \begin{cases} 0, & \text{if } \underset{\max j}{C_{ij}^A} \le R \\ \infty, & \text{otherwise} \end{cases} \tag{1}$$

where C_i^A is the available bandwidth in the MPLS link i (as the sum of individual spare capacities inside the wavelengths), C^T is the maximum link capacity in the MPLS network, R is the bandwidth required by the LSP, and C_{ij}^A is the available bandwidth in the jth wavelength of the ith MPLS link. According to the formula, the weight of a link increases as the available aggregated capacity of that link decreases, while it is set to ∞ when there is no wavelength, whose unused capacity is greater than or equal to the required one. If it can be found a route with a finite cost, the lightpath selection is performed by privileging the choice of more filled wavelengths [16].

It is worth noting that the DR module can be regarded as a CBR algorithm in which the constraint is the bandwidth requirement associated to that request. Since the DR operates in a multilayer scenario, it has to consider that bandwidth constraint ranges in a continuous domain in the IP/MPLS layer, while the resource at the optical layer range in a discrete domain (number of wavelengths).

3) Hybrid Routing: It has to be highlighted that both the above described on-line and off-line routing modules aim at improving network performance (i.e., minimizing the utilization of network resources and the blocking probability), but while the former operates on the basis of a statistical estimation of traffic pattern, the latter operates on actual traffic requests. Clearly, the dynamic routing is able to handle the temporary congestion due to the increment of actual traffic volume and/or to the different traffic distribution among the nodes with respect to the estimated traffic considered in the provisioning phase.

To facilitate the integration of the PR and DR an opportune flexibility factor, α, is introduced during the provisioning phase. The basic idea is to suitably scale the physical topology during the off-line procedure by reducing the bandwidth of each wavelength, so that the PR module must select more lightpaths in order to accommodate the same amount of foreseen traffic. As a result, the task of the DR module is facilitated since it operates on an enforced topology, at the expense of an increment of physical network resources utilization. In other words, if the factor $\alpha \in ([0, 1]$ is introduced and if the wavelength capacity is b_w, the wavelength bandwidth used during the PR procedure, is limited to αb_w, while during the DR operation those lightpaths, constituting the logical topology, are considered with their actual bandwidth, i.e., b_w.

The value of the factor α, which leads to an improvement of the dynamic network performance with a minimum number of network resources, depends both on the network load and on the relationship between the expected and the actual traffic [27]. The impact of the factor α will be discussed in Section IV, where the performance of TE are reported in different conditions, to test the robustness of the solution itself to promptly react to traffic fluctuations and unpredictability.

B. Bandwidth Engineering

The TE system is based on an elastic use of the bandwidth. This means that bandwidth assigned to higher priority LSPs during the provisioning phase can be temporarily released for the amount of time in which it is not needed and put at disposal of all the other lower priority LSPs. This means that as soon as the HP LSPs require back their bandwidth, the TE system immediately has to satisfy that need in some way. In order to do that, a function that handles preemption of lower priority LSPs or, even better, that can move lower priority traffic on less-congested routes is needed. In [17] it was proposed a BE system for that scope, that in this paper has been integrated as a module invoked in the integrated TE system. Specifically, BE makes use of two key elements: 1) a bandwidth handling algorithm (BHA), which selects those LP LSPs that need to be moved to make available the bandwidth required by the HP LSPs and 2) the previously mentioned DR algorithm, which aims at rerouting those selected LP LSPs on alternatives paths. In this way, BE allows bandwidth resource to be managed in an effective way, with the aim of both accommodating more traffic with respect to classic (non-TE) networks, and guaranteeing the required QoS for different CoS.

In particular, the BHA is invoked when an HP LSP requires more bandwidth on its route and at least one link on that route is congested because of the presence of other LP LSPs. Its operation consists in selecting the LP LSPs that have to be moved and rerouted by means of the DR. It is applied on all the congested links of the HP path requiring more bandwidth. On each congested link, it works iteratively until there is enough free bandwidth to let the HP traffic pass. Parameters used to calculate weights are recalculated in each step of the iteration. Several solutions have been investigated in [17]. The simplest one, herein considered, is an implementation of the MinConn algorithm reported in [22] for an IP/MPLS network. The BHA works as follows.

1) Search for congested links: It proceeds sequentially along the HP LSP, starting from the first link of the HP LSP path.
2) Weight calculation of LP LSP, $w_{i,j}$, related to the ith LP LSP on the jth congested link:

$$w_{i,j} = \frac{B_{\mathrm{LP-LSP}(i),j} - \delta_j}{\delta_j} \tag{2}$$

where δ_j is the bandwidth to be released on the jth link to accommodate HP LSP request, $B_{\mathrm{LP-LSP}(i),j}$ is the bandwidth used by the ith LP LSP, crossing the link jth.
3) Weight sorting, by increasing order (taking into account the weight sign).

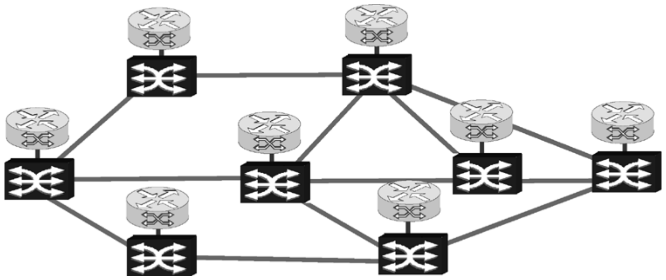

Fig. 4. Network topology.

4) Selection of LP LSP: If there exists at least one positive weight, the LP LSP with the lowest positive weight is selected to be torn down from the link under evaluation and the correspondent bandwidth is released on all the links crossed by that LP LSP. The selected LP LSP is submitted to a DR procedure so as trying to reroute it. If rerouting fails, the LP LSP is torn down from the network by the BE control. If there are only negative weights (when LP LSPs bandwidths are singularly smaller than the bandwidth to be free), iteratively, the algorithm selects more LP LSPs until the constraint of the bandwidth to be free is satisfied. Again, each selected LP LSP is passed to a DR procedure, which reroutes it. If rerouting fails, that LP LSP is torn down by the BE control.

IV. PERFORMANCE ANALYSIS

The objective of the performance analysis is to assess the feasibility of the proposed integrated TE system. In particular, it is interesting to evaluate:

1) the robustness of the solution with respect to different relevant situations of traffic demand;
2) the gain of adopting such a TE approach with respect to the case in which an over-provisioning approach is assumed;
3) the price to be paid, in terms of complexity, which can be measured in terms of number of operations (preemptions).

Section IV-A reports the details of the analysis environment. In particular, the network topology and the simulated traffic behavior for both HP and LP classes are presented. Section IV-B shows the achieved results for the multiservice network scenario, in which the proposed TE solution is applied.

A. Analysis Environment

The physical network topology is depicted in Fig. 4. It is composed by $\mathrm{N} = 8$ nodes and $\mathrm{L} = 12$ bidirectional optical links. Each optical link supports 16 wavelengths, with a wavelength capacity equal to $b_w = 2,5$ Gb/s.

Both for HP and LP traffic, the offered traffic can be described by a traffic matrix, whose generic element B_{ij} is the aggregated bandwidth considering the set of LSP requests between node i and node j

$$[B]^k = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & B_{ij}^k & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \quad k = \mathrm{HP}, \mathrm{LP}. \tag{3}$$

The sum of the bandwidth requested by all the LSPs for each pair of nodes is the traffic volume (TV)

$$\mathrm{TV} = \mathrm{TV}^{\mathrm{HP}} + \mathrm{TV}^{\mathrm{LP}} = \sum_{i,j\in V} B_{ij}^{\mathrm{HP}} + \sum_{i,j\in V} B_{ij}^{\mathrm{LP}}. \quad (4)$$

In order to characterize the traffic entering the network, a network load parameter ρ, defined as the ratio between the total offered bandwidth and the network available bandwidth, has been introduced:

$$\rho = \frac{\overline{h}\mathrm{TV}}{C_{\mathrm{net}}} \quad (5)$$

where $\overline{h}$ is the average minimum distance between each pair of source-destination nodes, and C_{net} is the total available bandwidth on the physical optical network.

In the simulations, the aggregated bandwidths are chosen randomly. In particular, for each pair of nodes, i and j, a random number uniformly distributed between 0 and 1, $C_{ij} \in [0,1]$, is picked so that the aggregated bandwidth of all the LSPs from node i to node j, B_{ij}^k, is $r^{(k)}C_{ij}$, where $r^{(k)}$ is a scaling parameter. The scaling parameter is chosen so that

$$\sum_{ij} r^{(k)} C_{ij} = \mathrm{TV}^k \quad k = \mathrm{HP}, \mathrm{LP}. \quad (6)$$

According to the proposed TE strategy, the off-line procedure operates using as input the estimated traffic matrix, while the on-line procedures (LP setup and HP bandwidth modify) operate using as input the actual traffic matrix.

In case of estimated traffic, B_{ij}^k represents the average expected aggregated bandwidth from node i to node j. It is determined by statistical evaluations, when relates to LP traffic. Differently, it represents the maximum allowed amount of traffic from node i to node j, agreed by SLAs, when relates to HP traffic. In the simulations, the estimated traffic matrix has been derived by generating a set of N_{ij} LSP requests with $b_{ij}^{(n)}$, representing the bandwidth associated to each LSP from node i to node j, so that

$$\sum_{n=1}^{N_{ij}} b_{ij}^{(n)} = B_{ij}^k, \quad b_{\min}^k \le b_{ij}^{(n)} \le b_{\max}^k \quad k = HP, LP \quad (7)$$

where $b_{\min}^k$ and $b_{\max}^k$ represent the minimum and the maximum estimated bandwidth requested by an LSP in case of LP traffic, while they represent the range values defined for the SLAs, for HP traffic.

In case of actual traffic matrix, a Poisson distribution has been assumed for the arrival process of LP connections between node i and node j follows. The rate λ_{ij} and the connection holding time follows a negative exponential distribution with mean $1/\mu$. The bandwidth of each LP LSP is uniformly distributed between $b_{\min}$ and $b_{\max}$, with mean $b = (b_{\max} - b_{\min})/2$. Thus, the average aggregated bandwidth, B_{ij}, between i and j, can be expressed as follows:

$$B_{ij} = \frac{\lambda_{ij}}{\mu} b. \quad (8)$$

In the simulations, by fixing μ, $b_{\min}$, and $b_{\max}$, from (8), it is possible to get λ_{ij} for each source-destination pair (i,j), and, hence, to generate the process.

As far as HP traffic is concerned, bandwidth modify events are generated for each HP LSP. The arrival time of bandwidth modify event is assumed to be uniformly distributed between $t_{\min}$ and $t_{\max}$, with mean $t = (t_{\max} - t_{\min})/2$. The amount of bandwidth modify is uniformly distributed between $b_{\min}$ and $b_{ij}^{(n)}$, with mean $b = (b_{ij}^{(n)} - b_{\min})/2$, where $b_{ij}^{(n)}$ is specified in the HP traffic matrix, and represents the SLA for each LSP.

Essentially, the traffic generation is based on the assumption that a traffic matrix, derived by SLAs, is at disposal of the provider for the HP traffic; hence, the maximum actual traffic is assumed consistent with the estimated traffic used in the provisioning phase. In the case of LP traffic, instead, the actual traffic can exceed and/or mismatch in spatial distribution the estimated one. Thus, in order to test the robustness of the proposed TE solution, three relevant case studies have been considered, which correspond to different relationships between estimated and actual traffic. In practice, the three cases differ from the level of accuracy of information available on both the traffic volume and the aggregated bandwidths.

Case 1: It corresponds to have an accurate *a priori* knowledge of the traffic behavior. That means that the information on both the traffic volume and the aggregated bandwidths between each pair of source-destination nodes are correct and, hence, that $(\rho)_a = (\rho)_e$, $(B_{ij})_a = (B_{ij})_e$ where the subscripts a and e refer to actual and estimated traffic and the relation on the aggregated bandwidths is valid for each (i,j).

Case 2: It corresponds to a case in which the information on the total traffic volume entering the network is correct, but it is not known how the traffic is distributed among the source-destination network nodes. That means that only parameter ρ is equal for the actual and estimated traffic: $(\rho)_a = (\rho)_e$.

Case 3: It corresponds to the worst case where the estimation of traffic volume and of the traffic distribution among network nodes is incorrect. Specifically, it has been assumed that the estimated ρ is 25% less than actual ρ.

In all the simulations the value of the holding time $1/\lambda$ is assumed constant for all the LP LSP connections and it is 200 s. The LP LSP bandwidths are assumed to be uniformly distributed from 1 to 500 Mb/s. For the HP traffic, the average *modify* holding time for each LSP is 2% of the simulation duration. The HP LSP SLAs range between 1 and 500 Mb/s and the modified bandwidth are assumed to be uniformly distributed from zero to the maximum bandwidth allowed by each LSP SLA.

In order to relate the HP and LP traffic load to the total network load a factor β has been defined, representing the percentage of HP traffic load with respect to the total network load such that

$$\rho = \rho_{\mathrm{HP}} + \rho_{\mathrm{LP}} = \beta \cdot \rho + (1 - \beta) \cdot \rho. \quad (9)$$

In the next section, results refer to the cases $\beta = 0.1$ and $\beta = 0.5$, which mean an HP traffic load corresponding to 10% and to 50% of the total network load, respectively.

B. Simulation Results

To test the robustness of the solution with respect to different relevant situations of traffic demand, the network performances

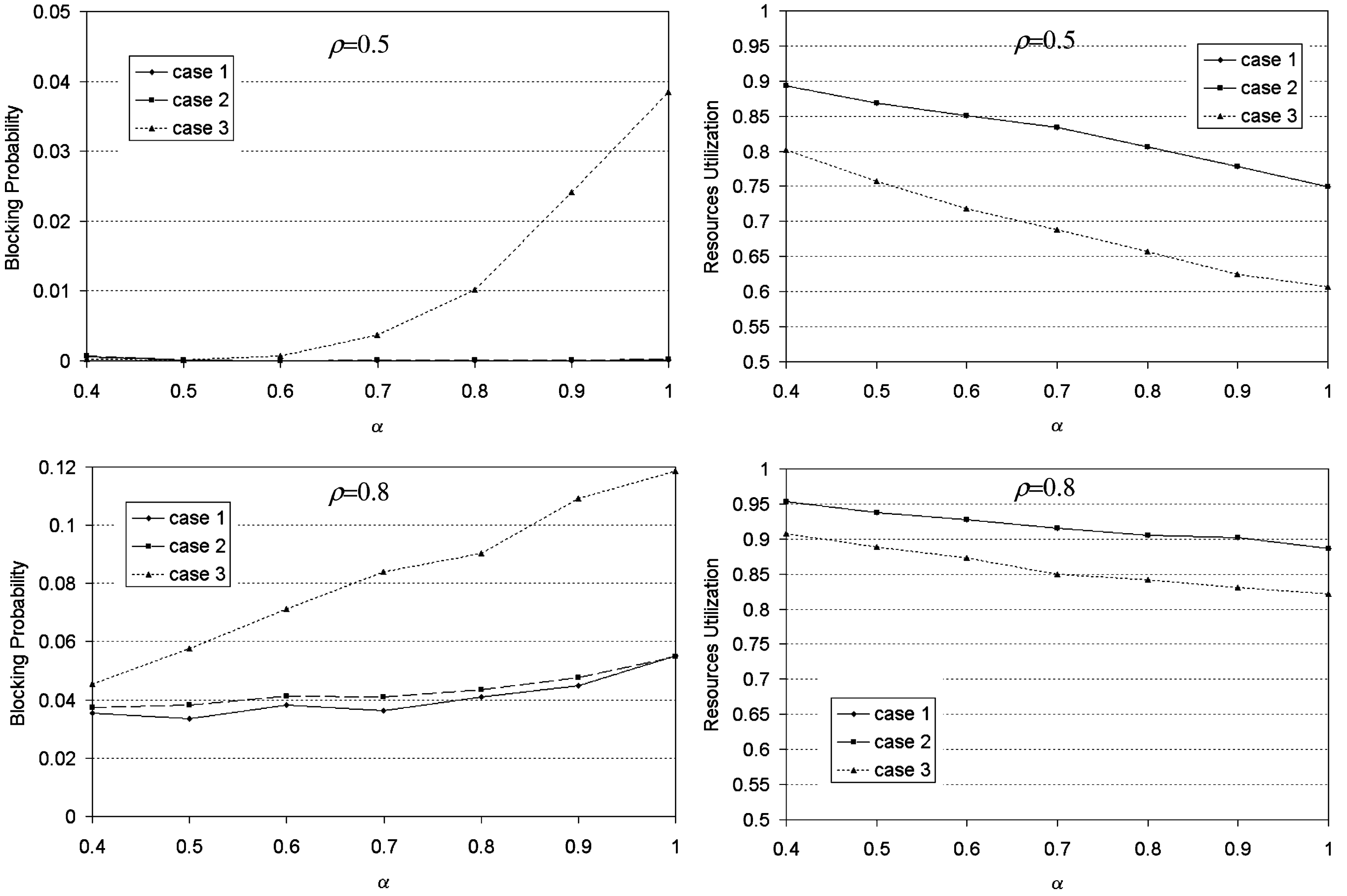

Fig. 5. Connection blocking probability and network resources utilization versus α, for the three different traffic relationship cases, with $\rho = 0.5$ (top) and $\rho = 0.8$ (bottom) and $\beta = 0.5$.

have been evaluated for the three aforementioned cases, and for two different values of β. The performances are calculated in terms of the connection blocking probability and the optical network resources utilization. The connection blocking probability is defined as the number of rejected connection requests with respect to the total number of connection requests. The optical network resources utilization is defined as the average ratio between the number of the used wavelengths and of the available wavelengths for each optical link.

By means of a simulative analysis, it has resulted that in order to make negligible the dependence of the simulation results on traffic matrix patterns, the simulative curves have been averaged on 20 different matrices.

Fig. 5 shows the connection blocking probability and the network resources utilization versus the flexibility factor α, for the three case studies, when the TE strategy is applied. Two different network load conditions ($\rho = 0.5$ and $\rho = 0.8$) are considered, assuming the HP traffic is a half of the total offered traffic ($\beta = 0.5$).

When $\alpha = 1$ (no enhanced flexibility is introduced during the provisioning phase), the blocking probability has the highest value, while the network resources utilization shows the lowest value, for both values of ρ.

Interestingly, Case 2 and Case 1 show the same behavior in terms of blocking probability. Specifically, for values of $\rho = 0.5$, the distance between the blocking probability curves is of the order of 10^{-5}, while for $\rho = 0.8$, it is in the range between 10^{-4} and 10^{-3}. This means that the dynamic routing is able to promptly react to traffic changes that could not be predicted during the provisioning phase. In fact, even though in Case 2 the traffic distribution varies significantly with respect to the one estimated by the traffic matrix (differently from Case 1), the TE system succeed in rearranging the routes' distribution in order to track the current traffic demand. In addition, the blocking probability, in both those cases, is almost independent of α; hence, the designer would favorably operate with high values of α in order to reduce network resource utilization. This is a further confirmation of the effective cooperation between off-line and on-line procedures. Clearly, the resource utilization curves in both Case 1 and 2 coincide, since the traffic matrices used in the network configuration phase are the same.

The role of α is much more evident in Case 3, where the actual traffic load becomes 25% bigger with respect to the estimated one. In fact, the blocking probability decreases rapidly as α decreases at the expense of resource utilization. This means that the choice of the α value depends on the trade off between blocking probability and resource utilization. This behavior is due to the way the routing algorithms operate. In fact, as α decreases, the provisioning algorithm provides a logical topology that is "more meshed" respect to the case of $\alpha = 1$, thus facilitating the task of the dynamic routing. In particular, an α value in the range [0.5–0.7] leads to an improvement of the network blocking probability of about 90% for $\rho = 0.5$ and 30% for $\rho = 0.8$. These results correspond to a reasonable increase of network resources utilization, i.e., 15% for $\rho = 0.5$ and 4% for $\rho = 0.8$.

The performance relating to Case 3 is worth being highlighted: even in presence of an appreciably increased traffic volume (25%) the TE is robust enough to cope with it, at a reasonable expense in terms of resource utilization. Of course, an overprovisioning approach would have led to similar performance, but at a much higher price in terms of network resources.

To better evidence this point, the performance of the TE system have been compared, in all the considered *case studies*, with an alternative approach based on overprovisioning, since, to the best of our knowledge, there are no integrated solutions of TE in multilayer networks reported in the literature. The alternative approach used as a reference, called dedicated bandwidth (DB), assumes that the bandwidth reserved for the HP traffic cannot be accessed by any LP traffic request, but it is completely dedicated to HP; even when the HP LSPs are not requiring the maximum bandwidth allowed by their SLAs. It can be regarded as a solution that uses overprovisioning just for the HP traffic that requires the mission-critical performance. Of course, any approach making a massive use of overprovisioning would lead to good performance at the expense of a huge utilization of network resources. Comparing the resource utilization in these cases with the one relating to TE would not add anything meaningful.

Fig. 6 shows in fact the blocking probability versus the traffic load ρ, in the three relevant cases, respectively. Each figure reports the comparison between the TE approach and DB approach, for two different values of percentage of HP traffic with respect to the total amount of traffic ($\beta = 0.1$ and $\beta = 0.5$). All the reported figures have been evaluated assuming the same value of the flexibility factor ($\alpha = 0.6$). This value corresponds to a reasonable tradeoff between performance and resource utilization, as derived from Fig. 5. In all the considered cases, it is evident the clear improvement of TE with respect to the DB approach. This is essentially due to the flexible use of the bandwidth resource that is achieved by the BE.

It can be observed from Fig. 6 that the advantage of TE with respect to the DB approach is more evident when the parameter β is higher. This is because the higher is the percentage of HP traffic with respect to the total offered traffic, the higher is the portion of bandwidth temporarily released by HP that can be utilized by BE to accommodate LP LSPs.

In particular, it is meaningful to evaluate the extra portion of traffic that the TE is able to accommodate with respect to the DB approach, for a given blocking probability (e.g., Blocking Probability $= 0.05$). Specifically, for $\beta = 0.5$ the gain obtained by TE is more than 50%, considering the cases 1 and 2 [Fig. 6(a) and (b)], and more than 300% in the Case 3 [Fig. 6(c)].

Basically, the advantages of TE with respect to an approach based on overprovisioning is obtained at the expense of the number of operations relating to preemption and rerouting that are needed to implement BE, and the amount of signaling required to actualize the solution.

In order to evaluate the amount of preemption that the strategy accomplishes, Fig. 7 reports the percentage of preempted LP LSPs with respect to the total number of LP LSPs accommodated in the network, for $\beta = 0.1$ and $\beta = 0.5$, versus the

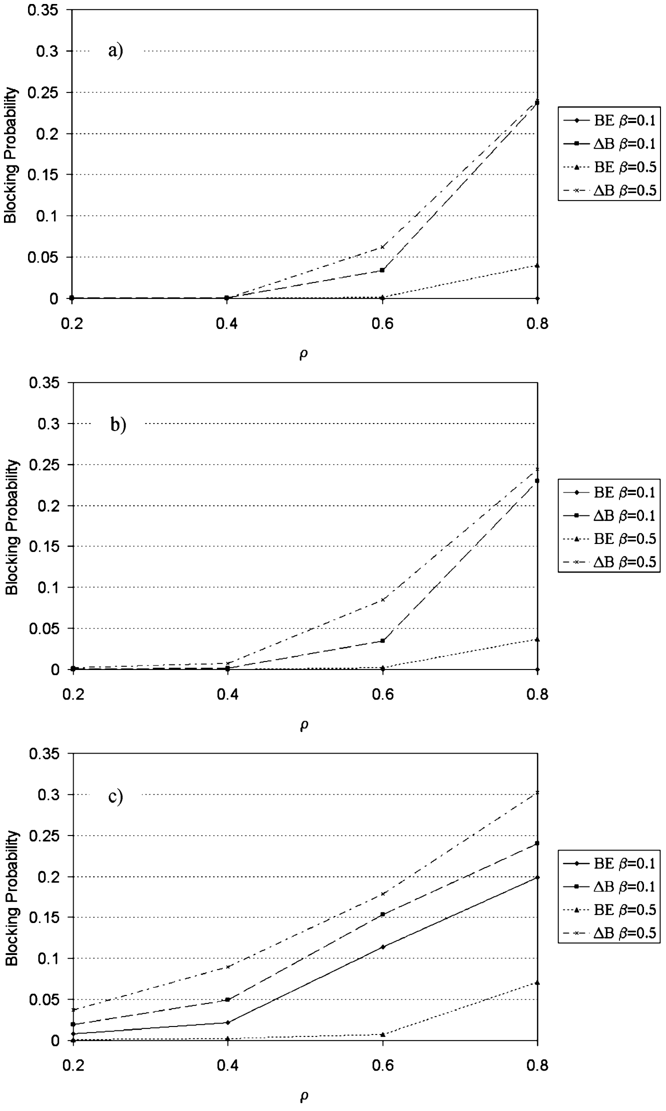

Fig. 6. Comparison between the BE and the DB strategies in terms of connection blocking probability versus ρ for different values of β with $\alpha = 0.6$: (a) case 1; (b) case 2; and (c) case 3.

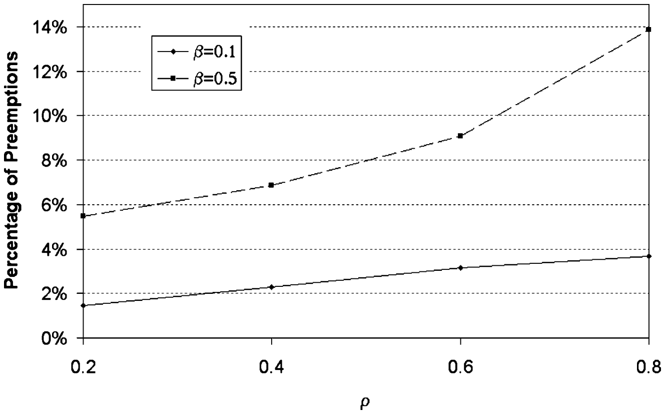

Fig. 7. Number of preemption operations required in the case $\beta = 0.1$ and in the case $\beta = 0.5$, normalized to the average number of LP LSP requests.

traffic load. For the sake of brevity, that figure has been only reported in the worst case, that is Case 3 where BE is more often invoked. The obtained low values indicate that the complexity of dynamic operations, required to achieve BE, is limited and does not prevent practical implementation.

The complexity of the TE system has been experienced during the realization of a test bed [28], which emulates the control plane of an IP/MPLS network, implemented in a distributed approach, supporting real routing and signaling protocols (OSPF-TE and RSVP-TE), in agreement with IETF standards. In practice, all the considered modules and related algorithms have been implemented in each node of the test bed network. The experiments made on the test bed show the feasibility of the key functions and of the TE system as a whole, in terms of scalability, stability, and QoS performance [28].

V. CONCLUSION

This paper reports an integrated TE system, which applies to a multilayer network, in a GMPLS-based multiservice scenario. A performance analysis assessing the feasibility of the proposed solution is also reported. In particular, the described TE system is based on two key novel components: 1) a hybrid routing scheme and 2) a system able to handle priority, preemption, and rerouting, called bandwidth engineering. The former allows an optimization of the use of the network resources and, at the same time, an improvement of the dynamic performance of the network and the robustness against traffic unpredictability. The latter further improves the performance of the network by achieving an elastic use of the bandwidth, so that the temporarily unused bandwidth by HP traffic is not wasted, but put at disposal of LP traffic. As a result, the proposed TE guarantees QoS requirements to be fulfilled, while at the same time it optimizes the use of the network resources, increases the flexibility of the network, and allows a large amount of traffic to be accommodated.

In order to assess the effectiveness of the considered TE system, the performances have been evaluated for three relevant case studies, according to the relationships between the traffic characteristics predicted off-line and the ones resulting from the current requests. Namely, a) Case 1: the current traffic distribution and volume do not appreciably differ from the ones predicted by the traffic matrix; b) Case 2: the traffic volumes are the same, but the current traffic distribution appreciably differs from the predicted one; and c) Case 3: the current traffic volume and distribution appreciably differ from the predicted ones.

The simulation results show the robustness of the proposed solution. In fact, the system is able to react to traffic changes by rearranging network resources, even when traffic volume and related distribution vary appreciably with respect to the predicted ones.

The advantage of TE with respect to an applicable overprovisioning approach (DB: dedicated bandwidth) has been shown in all the individuated case studies. Actually, for a given blocking probability, the gain obtained by TE with respect to DB in terms of accommodated traffic is quite evident.

In addition, it has been shown that since the number of preemption and rerouting operations is quite limited in practice, it is reasonable to assume that the complexity of TE is manageable.

ACKNOWLEDGMENT

This paper considers a TE strategy that has been partially developed in the framework of the Ericsson project Advance Solutions of Traffic Engineering: Research, Implementations and eXtensions (ASTERIX), and in part carried out within the National project TANGO, in the framework of the FIRB program. The authors wish to thank all the participants of the project. In particular, they are grateful to A. Bosco and E. Manconi for the development of the bandwidth-engineering concept and for fruitful discussions.

REFERENCES

[1] M. Listanti, V. Eramo, and R. Sabella, "Architectural and technological issues for future optical Internet networks," *IEEE Commun. Mag.*, vol. 38, pp. 82–92, Sept. 2000.

[2] X. Xiao and L. Ni, "Internet QoS: a big picture," *IEEE Network*, pp. 8–33, Mar./Apr. 1999.

[3] D. Awduche *et al.*, "Overview and principles of Internet traffic engineering," IETF, RFC 3272, May 2002.

[4] E. Mannie *et al.*, "Generalized multi-protocol label switching (GMPLS) architecture," work in progress IETF draft <draft-ietf-ccamp-gmpls-architecture-02.txt>.

[5] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol label switching architecture," IETF, RFC 3031, Jan. 2001.

[6] D. Awduche *et al.*, "Requirements for traffic engineering over MPLS," IETF RFC 2702, Sept. 1999.

[7] X. Xiao, A. Hannan, B. Bailey, and L. M. Ni, "Traffic engineering with MPLS in the Internet," *IEEE Network*, pp. 28–33, Mar./Apr. 2000.

[8] A. Banerjee *et al.*, "Generalized multiprotocol label switching: An overview of routing and management enhancements," *IEEE Commun. Mag.*, pp. 144–150, Jan. 2001.

[9] ——, "Generalized multiprotocol label switching: an overview of signaling enhancements and recovery techniques," *IEEE Commun. Mag.*, pp. 144–151, July 2001.

[10] B. Wen and K. M. Sivalingam, "Routing, wavelength and time-slot assignment in time division multiplexed wavelength-routed optical WDM networks," in *Proc. IEEE INFOCOM*, New York, June 2002.

[11] T. Fabry-Asztalos, N. M. Bhide, and K. M. Sivalingam, "Adaptive weight functions for shortest path routing algorithms for multi-wavelength optical WDM networks," in *Proc. IEEE Int. Conf. Communications*, New Orleans, LA, June 2000.

[12] D. Banerjee and B. Mukherjee, "Wavelength-routed optical networks: linear formulation, resource budgeting tradeoffs, and a reconfiguration study," in *Proc. Conf. Computer Communications (IEEE Infocom)*, Kobe, Japan, Apr. 1997, p. 269.

[13] T. Cinkler, D. Marx, C. P. Larsen, and D. Fogaras, "Heuristic algorithms for joint configuration of the optical and electrical layer in multi-hop wavelength routing networks," in *Proc. INFOCOM 2000*, pp. 1000–1009.

[14] P. Iovanna, M. Settembre, and R. Sabella, "A traffic engineering system for multi-layer networks based on the GMPLS paradigm," *IEEE Network*, Mar. 2003.

[15] R. Sabella, M. Settembre, G. Oriolo, F. Razza, F. Ferlito, and G. Conte, "A multilayer solution for path provisioning in new-generation optical/MPLS networks," *IEEE J. Lightwave Technol.*, vol. 21, pp. 1141–1155, May 2003.

[16] R. Sabella, P. Iovanna, G. Oriolo, and P. D'Aprile, "Strategy for dynamic routing and grooming of data flows into lightpaths in new generation network based on the GMPLS paradigm," *Kluwer Photon. Network Commun.*, vol. 7, pp. 131–144, Mar. 2004.

[17] A. Bosco, E. Manconi, R. Sabella, and L. Valentini, "An innovative solution for dynamic bandwidth engineering in IP/MPLS networks with QoS support," *Kluwer Photon. Network Commun.*, vol. 7, pp. 37–42, Jan. 2004.

[18] B. Davie and Y. Rekhter, *MPLS Technology and Applications.* New York: Academic, 2000.

[19] R. Coltun, "The OSPF opaque LSA option," RFC 2370, July 1998.

[20] D. Katz, "Traffic engineering (TE) extensions to OSPF version 2," IETF RFC 3630, Sept. 2003.

[21] D. Awduche *et al.*, "RSVP-TE: Extensions to RSVP for LSP tunnels," RFC 3209, Dec. 2001.

[22] M. Peyravian and A. D. Kshemkalyani, "Connection preemption: issues, algorithms, and a simulation study," in *Proc. IEEE INFOCOM*, vol. 1, 1997, pp. 143–151.

[23] E. Mannie *et al.*, "Generalized multi-protocol label switching (GMPLS) architecture," work in progress IETF draft <draft-ietf-ccamp-gmpls-architecture-07.txt>, Nov. 2003.

[24] L. Berger *et al.*, "Generalized multi-protocol label switching (GMPLS) signaling functional description," IETF RFC 3471, Jan. 2003.

[25] J. P. Lang, "Link management protocol," work in progress IETF draft <draft-ietf-ccamp-lmp-10.txt>, Oct. 2003.

[26] K. Kompella *et al.*, "LSP hierarchy with generalized MPLS TE," work in progress IETF draft <draft-ietf-mpls-lsp-hierarchy-08.txt>, Mar. 2002.

[27] G. Conte, P. Iovanna, M. Settembre, R. Sabella, and L. Valentini, "A traffic engineering solution for GMPLS network: a hybrid approach based on off-line and on-line routing methods," in *Proc. ONDM Conf.*, vol. 1, Budapest, Hungary, Feb. 4–6, 2003, pp. 313–331.

[28] A. Bosco, A. Botta, M. Intermite, P. Iovanna, and S. Salsano, "Distributed implementation of a pre-emption and re-routing mechanisms for a network control based on IP/MPLS paradigm," in *Proc. ONDM Conf.*, vol. 2, Budapest, Hungary, Feb. 4–6, 2003, pp. 821–839.

Paola Iovanna (M'03) was born in Rome, Italy, in 1971. She received the degree in electronics engineering from the University of Rome "Tor Vergata" in 1996.

From 1995 to 1997, she had collaboration as fellowship with a research center "U. Bordoni," Rome, where she dealt with advanced fiber-optic communications and optical networking issues. From 1997 to 2000, she was with "Telecom Italia," where she was involved in experimentation of new services based on different access technologies (as XDSL, Frame Relay, etc.). In 2000, she joined Ericsson Lab Italy, Rome, in the Research Department where she dealt with networking issues using MPLS and GMPLS technique. She is involved in the activity carried on in the framework of the Network of Excellence Euro-NGI, where she is responsible for the activity "Traffic Engineering for Next Generation Network." She holds two patents on dynamic routing solutions and traffic engineering system for multilayer networks based on the GMPLS paradigm for networks based on the GMPLS model, and she is author of several papers on traffic engineering in GMPLS-based networks.

Marina Settembre (M'03) was born in Rome, Italy, on July 6, 1960. She received the Laurea degree in physics from the University of Rome "La Sapienza" in 1985.

In the same year, she was granted a fellowship at the Fondazione Ugo Bordoni, working on innovative materials for optical devices. From 1986 to September 2000, she was a Research Scientist at the Fondazione Ugo Bordoni in the Optical Communication Division, focusing first on optical devices for signal routing/processing and then on physical layer modeling and on high-capacity optical transmission systems. In October 2000, she joined Ericsson Lab Italy, Rome, where she works in the Research Department on network architectures, traffic engineering strategies, and related algorithms and control plane definition for new generation networks based on the GMPLS paradigm. Her present research interests are also in the area of broadband wireless access technologies and architectures, focusing on traffic management, media access control protocols, and quality of service for packet-based communication in wireless systems. She has been actively involved in several European COST, ACTS, and IST projects (Cost 239, Cost 245, Cost 266, ACTS-Upgrade, ACTS- Esther, IST Atlas, IST Embrace) and is at the moment participating in the IST Integrated Project Broadwan in the Sixth Framework Programme. She has published more than 90 papers on international scientific journals and conferences and is author of the book *Non-Linear Optical Communication Networks* (New York: Wiley, 1998).

Roberto Sabella (M'92–SM'02) was born in Rome, Italy, on December 10, 1962. He received the Laurea degree in electronic engineering from the University of Rome "La Sapienza," Rome, Italy, in 1987.

He then joined Ericsson, Rome, where he was involved first in hardware design and subsequently in research activities on advanced fiber optic communication systems. His research interests have covered the fields of optical device technology, high-speed optical communication systems, WDM optical networks, and next generation Internet networks. In May 1997, he became the technical coordinator of the research consortium CoRiTeL. Since 1999, he has been the manager of the Research and Innovation Unit in Ericsson Lab Italy. He holds five patents relating to optical networks and traffic engineering strategies, is coauthor of a book on high-speed optical communications, and author or coauthor of about 100 papers published in either international scientific/technical journals or international conferences. He has been Lecturer (professore a contratto) at the University of Rome "Tor Vergata," and at the Polytechnic of Bari, both in Italy. Currently, he is Adjunct Professor of Design of Telecommunications Equipment at the University of Rome "La Sapienza." He is a member of the editorial board of the journal *Photonic Network Communications* (Kluwer). For the same journal, he has operated twice as Guest Editor for a special issue on WDM transport networks and on routing and failure restoration strategies in optical networks. He was also Guest Editor of a special issue on optical networks for the journal *Computer Network* (Elsevier). He will be the Chairman of the conference Next Generation Internet Networks, supported by the IEEE Communication Society, organized on behalf of the European Network of Excellence EURO-NGI.

Prof. Sabella has been Guest Editor of the IEEE COMMUNICATION MAGAZINE (Special Issue on "Optical Networking Solutions for Next Generation Internet Networks") and the IEEE NETWORK (Special Issue on "Traffic Engineering in Optical Networks).

Giulia Conte was born in Rome, Italy, in December 26, 1976. She received the degree in telecommunications engineering from "Tor Vergata" University, Rome, in 2001. Her thesis was about routing and failure restoration strategies in the presence of different traffic classes in WDM networks.

In February 2001 she joined CoRiTeL, Rome, where she takes part in the research activity of the "Optical Networks" group. Her present research interests include traffic engineering strategies and related algorithms, network architectures, and control plane definition for new generation networks based on the GMPLS paradigm.

Lucia Valentini was born in Perugia, Italy, on January 19, 1976. She received the degree in electronical engineering from the University of Perugia in November 2002. Her thesis was about traffic engineering strategies in presence of different traffic classes in GMPLS networks.

In December 2002 she joined CoRiTeL, Rome, Italy, where she involved in development of coding and transcoding functions for MPEG4 application transport.