

# Design and Implementation of a GMPLS-Controlled Grooming-Capable Optical Transport Network

Fernando Agraz, Luis Velasco, Jordi Perelló, Marc Ruiz, Salvatore Spadaro, Gabriel Junyent, and Jaume Comellas

**Abstract**—Aiming at better resource utilization, an important requirement of future optical transport networks is the capability to accommodate subwavelength client flows efficiently. This can be put into action thanks to the enhanced traffic engineering (TE) protocols provided within the generalized multiprotocol label switching (GMPLS) standardization. The present paper concentrates on the design and implementation of a GMPLS-controlled grooming-capable transport infrastructure, namely, the automatically switched optical network (ASON)/GMPLS CARISMA test bed. Through the paper, the operation of a GMPLS-controlled multilayer network architecture is introduced, subsequently highlighting implementation issues that come to light. Special attention is devoted to a centralized flow reallocation module deployed in the CARISMA test bed to minimize the overall network cost. In this context, an integer linear programming (ILP) formulation to obtain its optimal cost is derived and low-weighted metaheuristics providing a nearly optimal solution are additionally proposed. All contributions in the paper are supported by illustrative experimental results.

**Index Terms**—Assignment and routing algorithms; Networks, circuit-switched; Network optimization.

## I. INTRODUCTION

Wavelength-routed optical networks have received increasing attention as a promising approach to deploy end-to-end transparent networks in a cost-effective way. Their main goal is to optically bypass highly overloaded electronic routers, which results in a significant reduction in optical-to-electrical (O/E) ports, thus decreasing the overall network cost. Such networks, however, have traditionally been rather static, due to the manual optical circuit provi-

sioning process. With the advent of the automatically switched optical network (ASON) architecture [1], the ITU-T has enhanced wavelength-routed optical networks with dynamic connection capability. This capability is accomplished by means of a control plane entity, responsible for the establishment, maintenance, and release of connections over the optical transport plane.

In parallel, the Internet Engineering Task Force (IETF) has standardized generalized multiprotocol label switching (GMPLS) [2] as a set of protocols to implement a common control plane, able to manage several switching regions in an integrated way. In fact, not only can packet-switched-capable interfaces be managed by the different GMPLS protocols, but they can also manage time-division multiplexing [e.g., synchronous optical network/synchronous digital hierarchy (SONET/SDH)], lambda, and even fiber-switched-capable interfaces. This makes GMPLS the most accepted solution for implementing the control plane functionalities in the ASON architecture. These ASON networks with a GMPLS-capable control plane will be hereafter referred to as ASON/GMPLS networks.

The role of IP as a convergent technology has triggered the development of a wide range of IP-based multimedia services, like HDTV, video conferencing, telemedicine applications, or Internet telephony, each having different bandwidth or quality of service (QoS) requirements. This huge, heterogeneous, and predominantly bursty generated amount of traffic poses new challenges to network operators to provide a cost-effective data transmission. Because the bandwidth granularity of wavelength-routed optical networks is very coarse, typically a whole wavelength supporting 10 or even 40 Gbps Ethernet or SONET/SDH tributaries, these networks lack the flexibility to support sub-wavelength traffic demands, which leads to poor bandwidth usage.

In this context, the term traffic grooming identifies the process of packing several low-speed traffic

Manuscript received November 7, 2008; revised January 23, 2008; accepted February 18, 2009; published July 1, 2009 (Doc. ID 103862).

The authors are with the Advanced Broadband Communications Center (CCABA), Universitat Politècnica de Catalunya (UPC), Jordi Girona 1-3, 08034 Barcelona, Spain (e-mail: perello@ac.upc.edu).

Digital Object Identifier 10.1364/JOCN.1.00A258

streams into higher-speed streams (e.g., see [3–5]). This problem, typically focused on reducing the number of add-and-drop multiplexers (ADM) in legacy SONET/SDH ring architectures [3], has also been extended to maximize optical channel bandwidth usage in general WDM meshed transport networks [4,5].

From the GMPLS point of view, the grooming problem is translated into merging several higher-order label-switched paths (LSPs) into a lower-order LSP (e.g., grooming packet LSPs carrying IP traffic into a  $\lambda$ -LSP). Such an LSP aggregation in GMPLS is accomplished by advertising newly created lower-order LSPs as forwarding adjacency LSPs (FA-LSPs, [6,7]), for instance, by means of the open shortest path first-traffic engineering (OSPF-TE) protocol [8]. In this way, conventional data links along with the previously advertised FA-LSPs can indistinctly enter the path computation process. Supposing that a valid route would be found, resource reservation would then be performed by resource reservation protocol traffic engineering (RSVP-TE) [9].

During this operation, a dynamic virtual topology is created and modified the whole time. This virtual topology is comprised of those existent single- or multiple-hop  $\lambda$ -LSPs, along with unallocated data links spanning one single hop in the physical topology. Because new  $\lambda$ -LSPs are dynamically allocated in the network, and connection holding times are typically random, a suboptimal allocation of resources may exist at any time. In this context, a centralized resource reallocation module could be deployed in the network. As its main objective, the module would be responsible for periodically checking the occupancy of the existent  $\lambda$ -LSPs. Thus, client connections supported on  $\lambda$ -LSPs having a low occupancy may be rearranged, if possible, onto alternative medium-loaded  $\lambda$ -LSPs, which would result in better bandwidth usage in the network, as well as in a release of O/E port pairs. In this way, network resource utilization could be improved in the network. Note that along the reallocation process, some working  $\lambda$ -LSPs need to be rerouted. In this regard, [10] specifies a rerouting process called “make-before-break.” This procedure consists of establishing a new LSP and transferring traffic from the old LSP to the new one before the old LSP tunnel is finally torn down. In this way, reallocation procedures do not cause any working traffic disruption, which becomes of critical importance for end users.

The goal of this paper is to introduce the design and further implementation of an experimental GMPLS-based grooming-capable network test bed. To this end, we first present a generic GMPLS-controlled multilayer network architecture. With this objective, the current standardization framework is reviewed, accompanied by some illustrative examples. The next step comprises the design of a centralized off-line re-

source reallocation module. To address this, we construct an integer linear programming (ILP) formulation to obtain the optimal results. Moreover, we provide light-weight metaheuristics for subsequent implementation, which obtains a nearly optimal solution. In the experimental evaluation, a discussion on the implementation of the FA-LSP functionality in the ASON/GMPLS CARISMA test bed is reported. Later, FA-LSP performance is assessed by experimental results. Furthermore, on the basis of an already operating multilayer network infrastructure, a centralized off-line resource reallocation module is additionally deployed, experimentally quantifying its benefits in front of simple FA-LSP operation.

The reconfiguration problem in traffic-grooming optical networks has been covered in the literature (e.g., see [11–13]). The authors in [11,12] redesign the network topology as soon as the offered traffic changes. Two costs are taken into consideration, the O/E port cost and the reconfiguration cost. In this scheme, the result of a reconfiguration is basically the set of  $\lambda$ -LSPs to be established, as well as the route for each client LSP in the new topology. Because during reconfiguration some  $\lambda$ -LSPs are set up while others are torn down, the reconfiguration cost minimizes the distance between the current topology and the targeted one. Alternatively, the authors of [13] propose an adaptation mechanism to follow traffic variations in the network. Therein, a mechanism monitors the load of the links and sets up or tears down  $\lambda$ -LSPs when the supported load is either higher or lower than two predefined thresholds. Specifically, this algorithm is periodically run several times per hour. In contrast to existent work in the literature, our approach aims to reduce the needed O/E port pairs. Hence, no new  $\lambda$ -LSPs are created. Conversely, whether a  $\lambda$ -LSP is decided to be torn down, the client LSPs are rerouted so that the number of allocated network resources is minimized. We observe in the experimental results presented later that running the proposed metaheuristic algorithm only a few times a day obtains highly valuable cost reduction in the network.

The remainder of this paper continues as follows. Section II presents a GMPLS-controlled multilayer network architecture. Section III introduces the design of a centralized resource reallocation module, which will be afterwards implemented in the CARISMA test bed. Section IV is devoted to GMPLS-based grooming implementation and validation. Subsection IV.A presents the ASON/GMPLS CARISMA test bed characteristics. Subsections IV.B and IV.C describe FA-LSP implementation and assess their benefits. Subsection IV.D quantifies performance improvements that stem from using a centralized resource reallocation module in the network. Finally, Section V concludes the paper.

II. GMPLS-CONTROLLED MULTILAYER NETWORK ARCHITECTURE

A. Principle of Operation

The ITU-T establishes in G.805 [14] a reference layered transport network architecture with technology-independent relationships among functional entities. Therein, each network layer has a twofold role, namely, a server role to the client layer above it as well as a client role to the network layer below it. In brief, a subnetwork describes the capacity to associate a set of connection points (CPs) to convey so-called characteristic information. With such an objective, two possible kinds of connection are defined. A link connection is a fixed and inflexible connection between two CPs. Conversely, a subnetwork connection (SNC) is a flexible connection that may be set up and released by either the control or the management plane. As a result, a network connection is a concatenation of subnetwork and link connections delimited by a termination connection point (TCP) pair. For the sake of generality, we present FA-LSP operation in a G.805-compliant context. Recall that correspondence between ITU-T and IETF terminology can be found in [15,16].

A single-layered four-node all-optical network is exemplified in Fig. 1(a). In such a scenario, link connections (representing the different wavelength channel data links) associate CPs at remote neighboring nodes. These link connection sets are bundled into network connections between remote TCPs, which respectively represent dense WDM (DWDM) traffic engineering (TE) links and optical network ports. Let us suppose now that a  $\lambda$ -LSP is set up between ingress node E and egress node D [Fig. 1(b)]. The incoming client signal at the optical node E is adapted and cross-connected by an SNC to an outgoing CP. This CP is, in turn, connected through a data link to an incoming CP in the neighbor. At the intermediate nodes F and G, an SNC binds incoming and outgoing CPs, which

should be mapped to the same wavelength in the case that no wavelength converter is used. As soon as the signal reaches the destination node D, this one is cross-connected, adapted, and sent to the optical access port.

Note that the whole process explained above would be transparent to a hypothetical client network. In such an overlaid network scenario [2], nodes D and E would appear to be directly connected. As a matter of fact, each layer runs its own control plane, so that operations in one layer become totally independent from those in other layers. While this approach has been typically deployed due to information exchange restrictions across different network domains, it often leads to suboptimal resource allocation in the network. Contrariwise, in a peer network model [2] there is a common control plane that, having complete network knowledge, governs all layers in a unified way. Alternatively, an augmented model [2] would lie between the overlay and peer models, where each layer runs its own control plane instance and only a limited amount of information is exchanged among them.

In this context, the enhanced TE protocols introduced in GMPLS pave the way to a peer multilayer network architecture, controlled by means of a GMPLS-enabled common control plane. The enabling entity to this goal is the FA. In GMPLS, those already-established lower-layer LSPs (e.g.,  $\lambda$ -LSPs) are advertised as FA-LSPs, which can be used to transport new client LSPs. In this way, lower-layer resources can be more effectively utilized.

Without loss of generality, a two-layered network peer architecture is assumed in this work, that is, an optical server layer and a client aggregation layer on top (e.g., SONET/SDH, MPLS, GbE, etc.). This layer allows the mapping of the client traffic to be transported over the DWDM physical layer. At the bottom, optical nodes provide network ports as well as client access ports, used to inject an aggregated client flow to the network. The incoming signal would be afterwards adapted, switched to a network port, multiplexed into a DWDM bundle, and finally transmitted to the next optical node. On top, the client aggregation layer includes generic nodes providing electrical switching, flow aggregation, and many other features. Client nodes are connected to optical nodes through the client access ports.

Figure 2 shows the architecture of the two-layered peer network under consideration. In Fig. 2(a) a client LSP is set up between nodes E and D. Imagine that the client LSP requested bandwidth becomes 1/4 of the total wavelength capacity. The incoming signal is adapted and further inserted into an outgoing  $\lambda$ -LSP aggregated signal, reaching in this way the destination node D. Note that no processing is needed at cli-

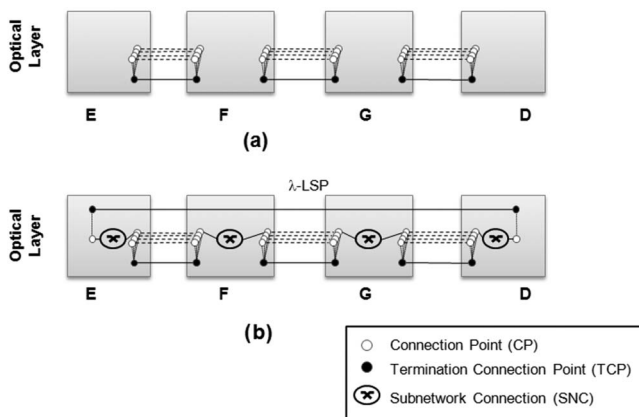


Fig. 1. Example of a single-layered network architecture.



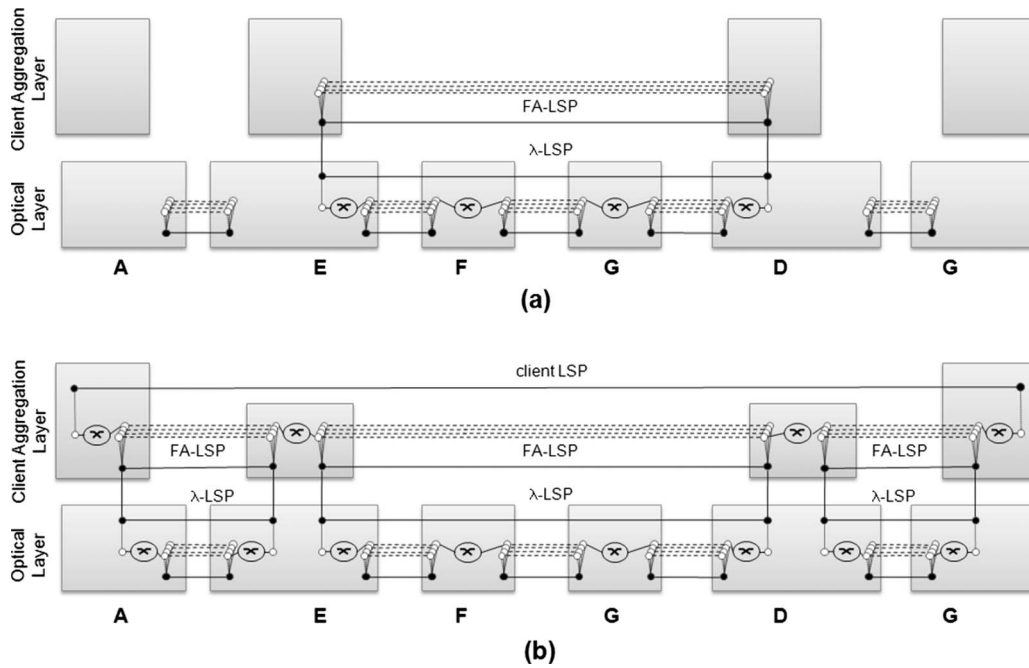


Fig. 2. Two-layer network architecture (a) before and (b) after setting up an end-to-end client LSP supported by three FA-LSPs.

ent intermediate nodes F and G, as the signal optically bypasses them through the  $\lambda$ -LSP E–D. At the destination, the signal is demultiplexed and the client signal is cross-connected, adapted, and delivered to the sink TCP (recall, termination connection point). In the resulting scenario, nodes E and D appear to be directly connected with an additional capacity of 3/4 of the total wavelength capacity. Further looking at Fig. 2(b), the previously established FA-LSP has now been used to create an additional client LSP between nodes A and G. To this end, two additional FA-LSPs between nodes A–E and D–G are set up, providing the required connectivity at the client layer. Specifically, an SNC associates incoming CPs with outgoing CPs at intermediate client nodes E and D. It is worth mentioning that FA-LSPs supported on different wavelength channels can be concatenated without wavelength converter requirements, as the signal is O/E converted at the client layer.

*B. Definition of the FA-LSP Routing Metric and Creation Cost Function*

In current GMPLS standardization [6] there is an intrinsic association between the signaling of new client LSPs and the creation of the required  $\lambda$ -LSP to support them. As will be later detailed, a route from source to destination is computed upon client LSP request, which may be constituted of both unallocated data links and already-existent FA-LSPs. In the case that no FA-LSP is comprised along the route, a new  $\lambda$ -LSP is typically set up from source to destination to support the incoming request. Otherwise,  $\lambda$ -LSPs are set up to provide connectivity on those route segments

where no FA-LSP is yet established. This operation, however, may lead to resource waste in the network. Notice that long FA-LSPs connecting far-off nodes are limited to be only reused by incoming LSP requests between remote end points. Hence, it appears more appropriate to separate the signaling functionality from the  $\lambda$ -LSP creation, so that  $\lambda$ -LSP placement can be decided based on network characteristics.

In the present paper, we set the routing metric of the already-established FA-LSPs to be  $\max(1, \text{FA-LSP hops}-1)$ , as described in [6]. Besides, the routing metric assigned to the unallocated data links spanning one single physical hop is set to 1. Aiming at better resource utilization, however, we dissociate  $\lambda$ -LSP establishment from network signaling functionality in the following way. Once the route from source to destination is calculated, the heuristic cost function,

$$C_{FA}(H) = H[(1 - p_H) + h/H], \tag{1}$$

is applied to the route segments where connectivity is not yet existent, with  $H$  standing for the number of hops of the yet-to-be-created  $\lambda$ -LSP and  $p_H$  standing for the probability that any incoming demand in the network has a certain number of hops  $H$ . The cost function provides us with the most appropriate  $\lambda$ -LSP configuration to optically connect the yet-uncovered route segment. As will be later depicted by example, the term  $(1 - p_H)$  encourages those  $\lambda$ -LSP lengths close to the average network distance, which are thus more likely to be reused. The term  $1/H$  identifies the use of O/E port pairs per hop, so that the larger the  $\lambda$ -LSP, the lower the use of O/E ports to connect its end points. The tunable  $h$  parameter fosters/penalizes the

use of O/E ports in the network. Finally, the total cost is multiplied by  $H$  as long as  $\lambda$ -LSPs need a higher number of unallocated data links. In this context, let us imagine that a new  $\lambda$ -LSP, which will afterwards act as FA-LSP, has to be established between a node pair distancing four hops. Supposing that a two-hop client LSP length is the most likely in the network, the combination  $C_{FA}(2)+C_{FA}(2)$ , that is, two  $\lambda$ -LSPs each spanning two hops, could have a lower cost than  $C_{FA}(4)$ , meaning one single end-to-end  $\lambda$ -LSP. Subsection IV.B particularizes  $C_{FA}(H)$  for the scenario under study. As will be highlighted, very short FA-LSP establishment (e.g., one hop) is also penalized by  $C_{FA}(H)$ , due to the large amount of required expensive O/E ports, as well as the large amount of bypass traffic to be electrically processed.

### III. CENTRALIZED RESOURCE REALLOCATION MODULE

#### A. Motivation

As mentioned before, dynamic connection establishment along with a randomness of connection holding times may lead to suboptimal resource allocation in the network at a certain time. For better understanding, consider the example depicted in Fig. 3 (left). It might happen that, due to the previous resource state in the network, a  $\lambda$ -LSP going through A–E–F–G–D–C would have to be created to support a client LSP from node A to node C. Supposing that the flow requested 1/4 of the total wavelength capacity, an FA-LSP from nodes A to C with 3/4 unreserved bandwidth would be created (step 1). Imagine now that a client LSP request arrives from node A to node B, also requesting 1/4 wavelength capacity. Provided that resources would be found on the direct link connecting both nodes, a direct  $\lambda$ -LSP would be set up, resulting as well in an FA-LSP from node A to C with 3/4 unreserved bandwidth (step 2). Finally, a client LSP re-

questing 1/2 of the total capacity reaches node B with destination node C. As the total routing metric of reusing FA-LSPs from B–A and A–C appears to be much higher than allocating a direct data link from B to C, a third  $\lambda$ -LSP is set up. This also triggers the establishment of a third FA-LSP with 1/2 unreserved bandwidth (step 3). Note that, if the client LSP request from A to C would have arrived now, rather than some time before, this one would have reused the A–B and B–C FA-LSPs.

This arouses concerns for deploying a centralized resource reallocation module in the network, which periodically checks the status of the already-deployed FA-LSPs and optimizes client LSP placement accordingly. If a resource reallocation would be triggered in the situation of Fig. 3 (left), this one could reallocate FA-LSP A–C into FA-LSPs A–B and B–C, which would result in resource savings of two client O/E ports and five optical data links as shown in Fig. 3 (right).

In this section, we address the reallocation process in a two-layered transport network. The main target is to minimize both the number of optical resources needed to carry the offered traffic to the network, as well as the number of hops of client LSPs. With such purposes in mind, we first propose an ILP formulation of the problem. Here, our objective lies in finding the optimal solution. In fact, solving the proposed ILP formulation is time consuming. Hence, we also introduce a metaheuristic to obtain a nearly-optimal resource reallocation in the network. The optimization process presented in the paper has been called optical resources optimization (ORO). Only the client LSPs (henceforth paths) are considered in the optimization procedures. As a matter of fact, the cost of the existent FA-LSPs (hereafter referred to as optical arcs) also reflects the underlying optical network topology, understanding a path as a feasible concatenation of optical arcs.

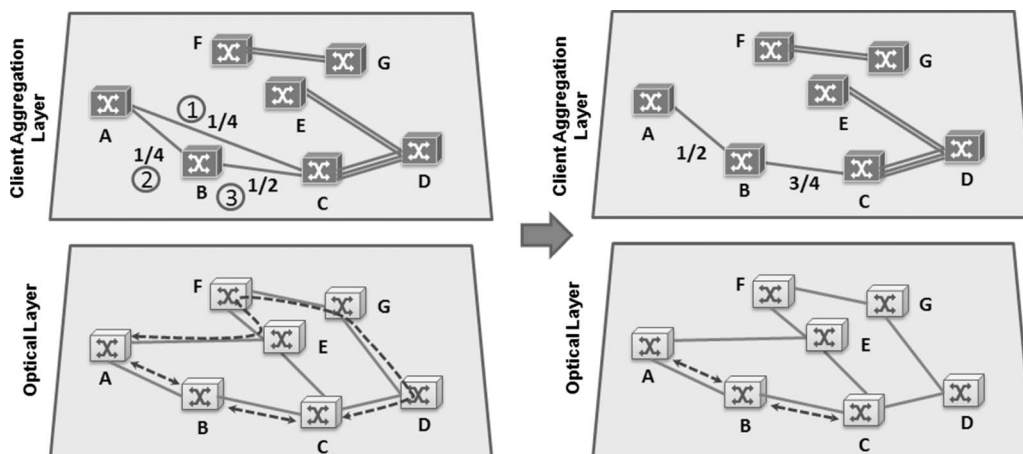


Fig. 3. Example of resource reallocation for optimization purposes in a two-layered network scenario.

### B. Optimal Solution: ILP Formulation

In the ILP formulation, the following notations are used for sets and parameters:

$E$	Set of paths (indexed by $i$ )
$R(i)$	Set of possible routes for path $i$ (indexed by $j$ )
$S$	Set of optical arcs (indexed by $k$ )
$C^k$	Cost of optical arc $k$
$M^k$	Capacity of optical arc $k$
$N_{ij}$	Equal to 1 if path $i$ was using route $j$ before optimization
$L_{ij}$	Cost of route $j$ for path $i$
$Q_{ij}^k$	Equal to 1 if route $j$ of path $i$ uses optical arc $k$
$W_i$	Bandwidth of path $i$

Additionally, the following notations are used for variables:

$\zeta^k$	Equal to 1 if optical arc $k$ is used after optimization (not removed)
$\delta^k$	Optical arc $k$ used bandwidth
$\eta_{ij}$	Equal to 1 if path $i$ uses route $j$ after optimization
$\rho_i$	Equal to 1 if path $i$ has been moved after optimization
$\psi_{ij}$	Equal to 1 if path $i$ has been moved to route $j$ after optimization

The ORO procedure releases as much optical resources as it can by releasing optical arcs so that the affected paths are rerouted using the minimum cost route, obtaining the most compact network. The formulation is based on an arc-path model, where the set of distinct routes for every path must be precomputed. Hence, the following objective function:  
Minimize

$$- \alpha_1 \times \sum_{k \in S} [C^k \times (1 - \zeta^k)] + \alpha_2 \times \sum_{i \in E} \sum_{j \in R(i)} (\psi_{ij} \times L_{ij}) + \alpha_3 \times \sum_{k \in S} \delta^k \quad (2)$$

subject to:

$$\sum_{j \in R(i)} \eta_{ij} = 1, \quad \forall i \in E, \quad (3)$$

$$\sum_{i \in E} \sum_{j \in R(i)} W_i \times Q_{ij}^k \times \eta_{ij} \leq \delta^k, \quad \forall k \in S, \quad (4)$$

$$\zeta^k \leq \delta^k \leq (M^k \times \zeta^k), \quad \forall k \in S, \quad (5)$$

$$\rho_i + \frac{\sum_{k \in S} [\zeta^k \times \sum_{j \in R(i)} (N_{ij} \times Q_{ij}^k)]}{\sum_{k \in S} \sum_{j \in R(i)} (N_{ij} \times Q_{ij}^k)} \geq 1, \quad \forall i \in E, \quad (6)$$

$$\sum_{k \in S} \left[ (1 - \zeta^k) \times \sum_{j \in R(i)} (N_{ij} \times Q_{ij}^k) \right] - \rho_i \geq 0, \quad \forall i \in E, \quad (7)$$

$$\psi_{ij} = \eta_{ij} - N_{ij} \times \eta_{ij}, \quad \forall i \in E, j \in R(i), \quad (8)$$

$$\sum_{j \in R(i)} \psi_{ij} = \rho_i, \quad \forall i \in E, \quad (9)$$

$$\zeta^k, \eta_{ij}, \rho_i, \psi_{ij} \in \{0, 1\}, \quad \delta^k \text{ integer}. \quad (10)$$

In the formulation, parameters  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  ( $\alpha_1 \gg \alpha_2 \gg \alpha_3$ ) in multiobjective function (2) give focus on the optical arc reduction objective. Constraint (3) ensures that every path has one and only one assigned route. Constraints (4) and (5) guarantee that if an optical arc is removed from the network, it does not contain any path. Moreover, the optical arc used bandwidth is stored and capacity restrictions are checked. Constraint (6) makes sure that each path using an optical arc to be removed is rerouted. Conversely, constraint (7) ensures that paths supported on optical arcs kept in the solution are not rerouted. Constraint (8) stores the new route to be used by the path being rerouted. Constraint (9) provides the paths to be rerouted. Note that constraints (8) and (9) basically capture information about the path routes, simplifying the structure of the objective function as well. Finally, constraint (10) defines variables as binary or integer. Additionally, we assume that every optical arc supports at least one path.

The formulation was run using the ILOG OPL [17] to implement the model and solved using the ILOG CPLEX v.11.0 [17] optimizer on a 3 GHz CPU machine with 1 GB RAM memory. Being that the model is quite complex, we observed in some example network scenarios that the solution lasted more than one hour. To enable ORO application in an off-line centralized module in the network, we also provide a metaheuristic based on greedy randomized adaptive search procedures (GRASP) [18] aiming at lower response optimization time.

---

#### Algorithm 1: ORO Constructive Phase

---

**Input:** Candidate list, network graph

**Output:** List of optical arcs to remove, list of paths to move with its new route and the cost of the solution

**while** candidate list size  $\neq 0$  **do**

RCL = {optical arc  $\in$  candidate list:  $Bw \leq Bw_{min} + \alpha(Bw_{max} - Bw_{min})$ };

Get a random optical arc from the RCL and remove it from the candidate list;

**if** reroute (arc, network graph) **then**

Add the optical arc to the solution;

Add every moved path to the solution;

If there is any path in its original route, remove it from the solution;

Compute solution cost;

---

**Algorithm 2:** ORO Reroute Algorithm**Input:** Arc to reroute, network graph**Output:** Successful or failed reroute

Create an empty path list;

Remove the arc from the graph;

**for every path using the optical arc do**

Get a copy of the current route of the path;

Tear down the path;

Compute shortest path from source to destination;

**if no route found then**

Restore the original route of the path and set it up;

Add the optical arc to the graph;

**for every path in the list do**

Tear down the path;

Establish the path in its original route;

Return (failed to reroute);

Establish the path;

Add the path to the list;

Mark every path in the list as moved;

Return (success);

### C. Proposed Heuristics

The ORO metaheuristic constructive phase (Algorithm 1) consists of trying to remove optical arcs from the network graph. At the beginning, a candidate list is built, containing every optical arc in the graph sorted in increasing order by used bandwidth ( $Bw$ ). For each iteration, an optical arc is selected from a restricted candidate list (RCL), which only contains a subset of the total candidate list composed of the least-congested optical arcs. In operation, the algorithm randomly chooses an optical arc and subsequently removes it from the candidate list. Then, it tries to reroute all paths supported on this selected optical arc.

The ORO reroute algorithm (Algorithm 2) uses the Dijkstra shortest path algorithm [19] to find an alternative route from source to destination, particularly avoiding the optical arc to be removed and with enough bandwidth to support the rerouted path. Note that the Dijkstra algorithm can only be used on connected graphs with no parallel arcs. This is not necessarily accomplished by the client aggregation layer as a whole. Because connections are dynamic in nature, the graph may not be connected, appearing as several connected subgraphs. This is illustrated in the example shown in Fig. 3. There, it can be realized that parallel arcs are easily found. In fact, the nodal degree  $d$  in the client layer is typically much higher than in the optical layer, as can be appreciated in node D where  $d$  equals 5. For instance, if a DWDM optical layer with 40 wavelengths per link and an average node degree  $\bar{d}=3$  would have been considered, a  $\bar{d}$  value equal to 120 could be reached in the client layer. Hence, before running the construction algorithm, a virtual node is added when a parallel arc between two

nodes is found. Besides, the Dijkstra algorithm is run over the connected subgraph containing the origin node. In the case that a path in the optical arc cannot be rerouted, already-processed paths are returned to the original route and the optical arc is placed again in the network graph.

The constructive phase analyzes all optical arcs ( $m$ ) and for each one it tries to reroute every supported path on the arc. Let us consider a worst-case scenario. As mentioned above, if no alternative route would be found for the last path under the reroute process, all already-rerouted paths contained in the arc would be placed again in their original route. This results in a computational complexity of  $O(m * p^2)$ , with  $p$  standing for the maximum number of paths contained in the optical arcs.

The ORO metaheuristic improvement phase consists of optical arc exchanges. An optical arc in the current solution is returned to the network and the improvement algorithm tries to get alternative optical arcs where paths may be able to be rerouted. Being that the exchange cost is lower than the cost of the returned arc, the current solution is updated. The computational complexity can be demonstrated to be  $O(m^2 * p^2)$ . Because the set of construction and improvement phases is repeated  $k$  times, updating the incumbent solution when the cost of the current solution is lower than the cost stored so far, the total computational complexity rises to  $O(k * m^2 * p^2)$ .

## IV. EXPERIMENTAL RESULTS

### A. ASON/GMPLS CARISMA Test Bed

The experimental evaluation has been carried out over the ASON/GMPLS CARISMA test bed [20], a configurable multitopology signalling communications network (SCN) running over wavelength-selective-switch- (WSS-) based optical cross-connect (OXC) emulators. In this configurable SCN, optical connection controllers (OCCs) are interconnected by 100 Mbps full-duplex point-to-point Ethernet links, describing the same physical topology of the emulated optical transport plane. This results in an out-of-fiber control plane architecture associated with the underlying transport plane. Running on top of the architecture, a network management system (NMS) has been implemented as a web-based application, enabling network configuration and both permanent and soft-permanent connection provisioning through the Internet.

In the test bed, OCCs are deployed by means of Pentium IV Linux-based routers at 2 GHz, so that each OCC implements the full GMPLS protocol set: RSVP-TE for signaling, OSPF-TE for routing and in-



formation advertisement, and, finally, link management protocol (LMP) for resource discovery and management [21]. In this work, we particularly build a nine-node meshed SCN configuration to evaluate the proposed contributions as depicted in Fig. 4.

### B. Forwarding Adjacencies Implementation

Because the CARISMA test bed was essentially a wavelength-routed optical network, resources were allocated for end-to-end connections with a whole wavelength granularity. This section reports the implementation and further evaluation of the entire FA-LSP functionality in the CARISMA test bed, enhancing it with multilayer support at the control plane level.

Looking at the standardization [22], a new RSVP-TE object named the `LSP_TUNNEL_INTERFACE_ID` was proposed to be used when signaling a new FA-LSP in the PATH and RESV RSVP-TE messages [9]. The object contains two fields, that is, the FA-LSP identifier and the router ID.

In the event of a new  $\lambda$ -LSP to be set up, the head-end OCC must allocate an identifier for the interface associated with the yet-to-be-created FA-LSP. Next, it originates an RSVP-TE PATH message containing a `LSP_TUNNEL_INTERFACE_ID` object filled with the selected local interface identifier, along with the local optical node identifier. When the PATH message arrives at the destination, the tail-end OCC must allocate an identifier for that FA-LSP end. This is called the remote FA-LSP interface identifier, which is reported back to the head end within the RSVP-TE RESV message. As soon as the  $\lambda$ -LSP has been cre-

ated, the head-end OCC advertises it as a forwarding adjacency by means of OSPF-TE. Being that the FA-LSP is bidirectional, it is also advertised by the tail-end OCC. All OCCs receiving the FA advertisement update their link state database adding a new link between the involved nodes.

Note that in the GMPLS standardization, the establishment of an FA-LSP is intrinsically associated with setting up a client layer LSP. As introduced in Subsection II.B, however,  $\lambda$ -LSP length may be limited to maximize resource reutilization in the network. To achieve such purposes,  $\lambda$ -LSP establishment must be dissociated from the client LSP setup procedure, allowing in this way the establishment of several underlying  $\lambda$ -LSPs while signaling only one client LSP request. To permit this separation between client LSP and  $\lambda$ -LSP setup, the head-end OCC evaluates the accumulated optical length, using loose hops in the RSVP-TE explicit route object (ERO) [22] when it decides to divide the whole route segment into two or more underlying  $\lambda$ -LSPs. The same mechanism is used also when an intermediate  $\lambda$ -LSP must be created. Every intermediate OCC receiving an RSVP-TE PATH message with a next hop set as loose must compute the next route segment possibly signaling a new  $\lambda$ -LSP.

It is worth pointing out that  $\lambda$ -LSPs created during client LSP setup signaling procedures have no specific RSVP-TE refresh messages. In fact, the  $\lambda$ -LSP must be torn down when releasing the last supported client LSP. In order to maintain the association between client LSPs and  $\lambda$ -LSP everywhere, the signaling of client LSPs using an already-existing FA-LSP follows, in our implementation, the same path in the control plane as the signal at the optical layer. This means that the same OCC can process a single RSVP-TE PATH or RESV message more than once.

In the example in Fig. 5, nodes E and D are directly connected through FA-LSP E–D. Let us suppose that the route of a new client LSP between nodes A–G includes FA-LSP E–D together with A–E and D–G links. In such a case, two new FA-LSPs A–E and D–G are created to support the new client LSP A–G. Node G in Fig. 5 processes RSVP-TE messages of the client LSP A–G two times, that is, the first time as an intermediate OCC and the second time as the tail-end OCC.

### C. Scenario Configuration and FA-LSP Functionality Evaluation

For the evaluation, we assume a nine-node meshed optical network with the same topology as the control plane depicted in Fig. 4 (i.e., the configured control plane would be deployed associated with a nine-node grooming-capable optical transport network). In such a scenario, we assume that each link carries eight bi-

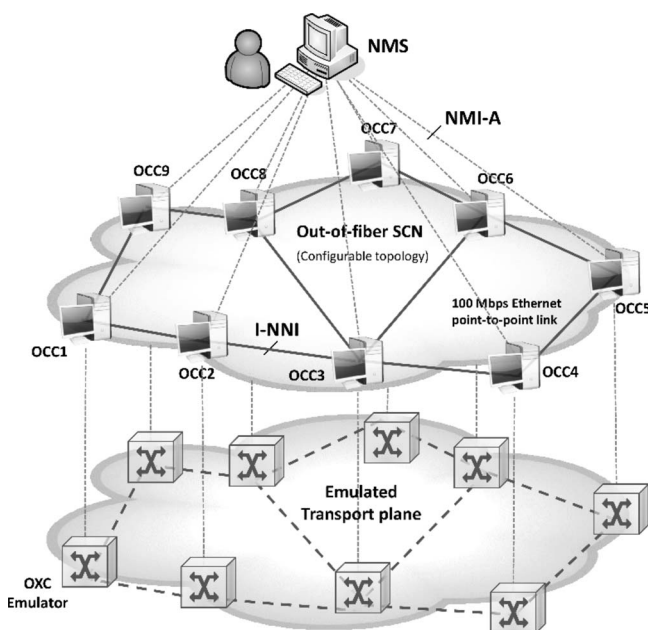


Fig. 4. The ASON/GMPLS CARISMA test bed describing a nine-node meshed scenario.



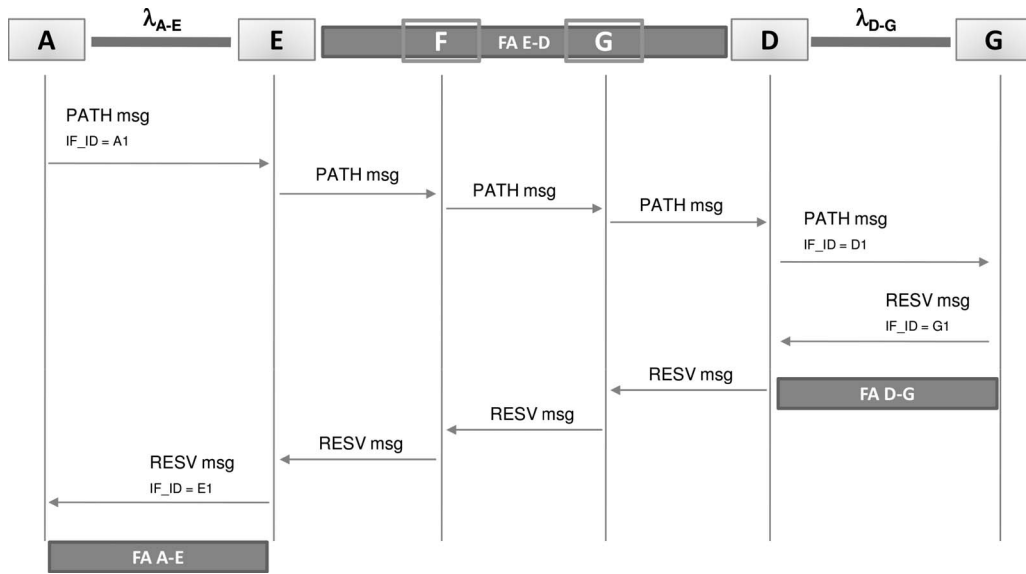


Fig. 5. Example of FA-LSP establishment.

directional wavelengths. For the traffic characteristics, we consider that uniformly distributed client LSP requests arrive at the network following a Poisson process with mean interarrival time (IAT) equal to  $1/\lambda$ . Besides, connection duration follows an exponential distribution with mean holding time (HT) set to  $1/\mu$ . In particular, the requested  $Bw$  of all incoming client LSP requests is considered to be  $1/4$  of the total wavelength capacity.

Figure 6 illustrates the obtained  $C_{FA}(H)$  function for the scenario under study. The bar graph plots the probability that an incoming client LSP request has a certain number of hops, assuming availability of resources through the shortest path. As seen, there is a 40% probability that an incoming request traverses

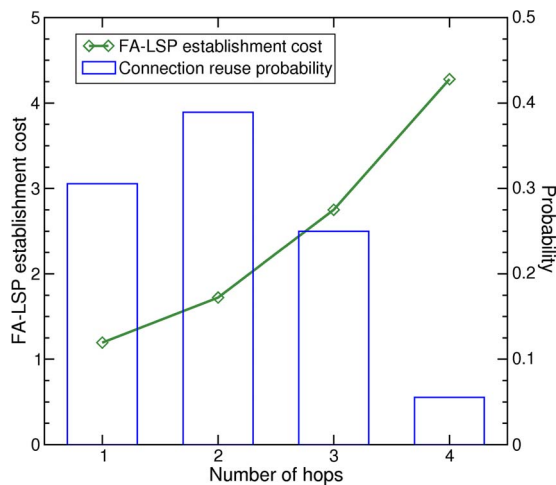


Fig. 6. (Color online) FA-LSP creation cost function for the nine-node topology under study. Bar graph shows the probability that a client LSP request has 1, 2, 3, or 4 hops.

two hops from source to destination. In contrast, only 5% of the incoming requests would traverse four hops. This validates our assumption in Subsection II.B, where we stated that, by splitting very long FA-LSPs into shorter ones, resources are much more likely to be reused. Values greater than the network diameter (i.e., four hops in our scenario) have  $p_H=0.0$  and are not depicted in the figure. To finally obtain  $C_{FA}(H)$  we fix  $h=0.5$ , as it provided the best network performance, while fulfilling our design criteria:  $C_{FA}(2) + C_{FA}(2) < C_{FA}(4)$  and  $C_{FA}(1) + C_{FA}(2) < C_{FA}(3)$ .

Figure 7 plots the obtained client LSP blocking probability as a function of the offered load to the network ( $\lambda/\mu=HT/IAT$ ) normalized to a value of 200.

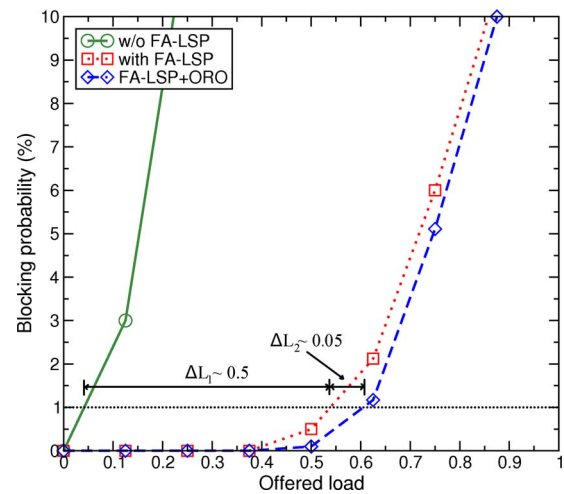


Fig. 7. (Color online) Client LSP blocking probability as a function of the offered load to the network. The situations without FA-LSP capability in the network, with FA-LSP capability, and with FA-LSP capability also deploying a centralized resource reallocation module are considered.

The curve without FA-LSP indicates the situation where no FA-LSP capabilities exist in the network, thus allocating a whole wavelength capacity for the incoming connection requests (even though they only request 1/4 of the total wavelength capacity). Conversely, the curve with FA-LSP evaluates the improvement obtained when enhanced subwavelength provisioning flexibility is provided to the network by implementing GMPLS-controlled grooming actions (i.e., with FA-LSP). The situation of additionally deploying a centralized resource reallocation module in the network is also depicted, but it will be detailed in the following subsection.

As expected, significantly better resource usage is achieved when implementing FA-LSP capabilities in the network. For instance, if 1% client LSP blocking probability would have to be ensured, a maximum load of  $L=0.05$  could be offered to a pure wavelength-routed optical network (i.e., without FA-LSP). Conversely, it could be further increased to  $L=0.55$  when FA-LSPs are implemented. This  $\Delta L_1=0.5$  experimentally assesses the implementation of FA-LSP capabilities to automatically manage grooming actions in future transport networks, given the lack of pure wavelength-routed optical networks to allocate incoming sub-lambda client LSP requests. In fact, as a whole wavelength is allocated in a per 1/4 wavelength capacity client LSP request in the without FA-LSP situation, 3/4 of the total network capacity is directly wasted. Qualitatively speaking, this approximately results in four times less carried traffic by the network.

#### D. Deployment and Validation of the Centralized Resource Reallocation Module

On the basis of an already-operative GMPLS-controlled multilayer network, a centralized resource reallocation module has been further deployed and evaluated in the CARISMA test bed. First of all, the performance of the ORO metaheuristic has been compared to the optimal solution obtained with the ORO ILP formulation. In the situations depicted in Fig. 7 low relative errors to the optimal solution around 5% are observed for low- and medium-loaded network scenarios (i.e., leading to client LSP request blocking probabilities lower than 1%). In fact, relative errors were significantly increased to 10%–20% in highly loaded network situations. Note, however, that these loads lead to unacceptable blocking probability values higher than 5%. Furthermore, it is worth highlighting that measures showed less than 0.3 s ORO metaheuristic running times. These short running times leverage its applicability for optimizing resource allocation in the network.

Because the client LSP requests incoming process

plays an important role in final network performance, resource reallocation actions are triggered every 50 incoming requests. In our scenario this is translated into applying the resource reallocation process a few times per day. Looking at the results previously obtained in Fig. 7, a reduction in the number of consumed resources and its consequent impact on the blocking probability can be appreciated. The random nature of the offered connections leads to suboptimal resource utilization, so the proposed periodic rearrangement module is useful to improve the whole network performance. When looking at the figure, it can be noticed that request blocking probability is clearly reduced for offered load values ranging from 0.4 to 0.65. For example, the request blocking probability is reduced from 2% to around 1% when the load is 0.625. This is due to the release of some O/E ports produced when reallocating client LSPs between FA-LSPs.

Port reduction is additionally plotted in Fig. 8 as a function of offered network load. There, it is shown that for low load values the gain is only marginal. Obviously, when the number of FA-LSPs is low, it is difficult to reallocate traffic. As a matter of fact, ORO benefits increase along with the offered load to the network, reaching its maximum for load values around 0.6. Under these loads, 14% port reduction can be achieved in the network, drastically decreasing network capital expenditures (CAPEX). Note, however, that if the load is further increased, reduction decreases again as the FA-LSP mean occupation reaches higher values, it thus being again difficult to reallocate client LSPs.

#### V. CONCLUDING REMARKS

This paper targeted at designing and implementing a GMPLS-controlled grooming-capable optical transport network. To this end, the current standardization

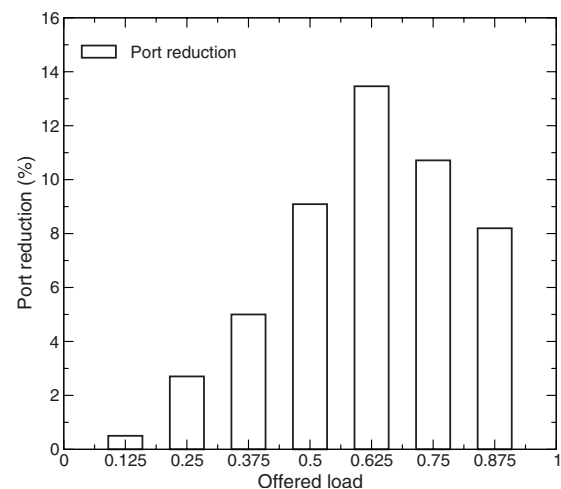


Fig. 8. (Color online) O/E port reduction in the network in percentages when the ORO metaheuristic is applied in the network.

framework concerning GMPLS-enabled multilayer networks was first reviewed and exemplified. Next, operation of a centralized resource reallocation module to be deployed in a multilayer optical network was discussed. In this context, an ILP formulation quantifying its optimal performance was derived. Being that the model is quite complex and its solving time is rather long, subsecond running time metaheuristics were also provided, which obtained less than 5% relative error to the optimal solution in the operating network scenario. The obtained experimental results validated GMPLS-controlled grooming actions, drastically improving client LSP blocking probability compared to a pure wavelength-routed optical network scenario. Besides, the deployment of the centralized resource reallocation module in the ASON/GMPLS CARISMA test bed resulted in noticeable improvements in terms of both O/E port and client LSP blocking probability reduction. In particular, focusing on a scenario with request blocking probabilities around 1%, O/E port and request blocking probability reductions of about 10% and up to 1% were respectively obtained.

The evaluation presented in this paper concerns a single nine-node ASON domain. Further work will extend the implementation of the FA-LSP functionality in larger network scenarios, including signaling and routing interworking issues in multidomain multilayer network environments. Particularly, it is our goal to assess not only the performance, but also the scalability of the GMPLS-controlled grooming and the proposed reallocation approach as the network gets larger, even spanning more than a single domain.

#### ACKNOWLEDGMENT

The work reported in this paper has been partially supported by the Spanish Science Ministry through the project Engineering Next Generation Optical Transport Networks (ENGINE), (TEC2008-02634). Moreover, the authors thank the support from the i2CAT Foundation through the project TRILOGY.

#### REFERENCES

- [1] ITU-T Rec. G.8080/Y.1304, "Architecture for the automatically switched optical networks," Nov. 2001.
- [2] E. Mannie, "Generalized multi-protocol label switching (GMPLS) architecture," IETF RFC 3945, Oct. 2004.
- [3] O. Gerstel, P. Lin, and G. Sasaki, "Combined WDM and SONET network design," in *Proc. IEEE INFOCOM*, 1998, pp. 734–743.
- [4] K. Zhu and B. Mukherjee, "Traffic grooming in an optical WDM meshed network," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 1, pp. 122–133, Jan. 2002.
- [5] R. Dutta and G. N. Rouskas, "Traffic grooming in WDM networks: past and future," *IEEE Network*, vol. 16, no. 6, pp. 46–56, Nov. 2002.
- [6] K. Kompella and Y. Rekhter, "Label switched paths (LSP) hi-

erarchy with generalized multi-protocol label switching (GMPLS) traffic engineering (TE)," IETF RFC 4206, Oct. 2005.

- [7] J. Comellas, R. Martínez, J. Prat, V. Sales, and G. Junyent, "Integrated IP/WDM routing in GMPLS-based optical networks," *IEEE Network*, vol. 17, no. 2, pp. 22–27, Mar. 2003.
- [8] D. Katz, K. Kompella, and D. Yeung, "Traffic engineering (TE) extensions to OSPF version 2," IETF RFC 3630, Sept. 2003.
- [9] L. Berger, "Generalized multi-protocol label switching (GMPLS) signaling resource reservation protocol-traffic engineering (RSVP-TE) extensions," IETF RFC 3473, Jan. 2003.
- [10] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: extensions to RSVP for LSP tunnels," IETF RFC 3209, Dec. 2001.
- [11] B. Ramamurthy and A. Ramakrishnan, "Virtual topology reconfiguration of wavelength routed optical WDM networks," in *Proc. IEEE GLOBECOM*, 2000, pp. 1269–1275.
- [12] R. Mahalati and R. Dutta, "Reconfiguration of traffic grooming optical networks," in *Proc. BROADNETS*, 2004, pp. 170–179.
- [13] A. Gencata and B. Mukherjee, "Virtual-topology adaptation for WDM mesh networks under dynamic traffic," *IEEE/ACM Trans. Netw.*, vol. 11, no. 2, pp. 236–247, April 2003.
- [14] ITU-T Rec. G.805, "Generic functional architecture of transport networks," March 2000.
- [15] D. Fedyk, O. Aboul-Magd, D. Brungard, J. Lang, and D. Papadimitriou, "A transport network view of the link management protocol (LMP)," IETF RFC 4394, Feb. 2006.
- [16] I. Bryskin and A. Farrell, "A lexicography for the interpretation of generalized multi-protocol label switching (GMPLS) terminology within the context of the ITU-T's automatically switched optical network (ASON) architecture," IETF RFC 4397.
- [17] ILOG Inc., [www.ilog.com](http://www.ilog.com)
- [18] T. Feo and M. Resende, "Greedy randomized adaptive search procedures," *J. Glob. Optim.*, vol. 6, pp. 109–133, June 1995.
- [19] R. Bhandari, *Survivable Networks: Algorithms for Diverse Routing*, Norwell, MA: Kluwer Academic, 1999.
- [20] J. Perelló, E. Escalona, S. Spadaro, J. Comellas, and G. Junyent, "Resource discovery in ASON/GMPLS transport networks," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 86–92, Aug. 2007.
- [21] J. Lang, "Link management protocol (LMP)," IETF RFC 4204, Oct. 2005.
- [22] K. Kompella and Y. Rekhter, "Signalling unnumbered links in resource reservation protocol—traffic engineering (RSVP-TE)," IETF RFC 3477, Jan. 2003.



networks.

**Fernando Agraz** received his M.Sc. degree in computer engineering in 2005 from the Polytechnic University of Catalonia (UPC). Since 2005 he has been working as a research engineer in the Optical Communications Group (GCO) at UPC, also preparing his Ph.D. He has also participated in various European research projects such as IST Nobel Phase 2 or E-Photon/ONe+. His current research focuses on network management and routing in GMPLS-based





**Luis Velasco** received the B.Sc. degree in telecommunications engineering from Universidad Politécnica de Madrid (UPM) in 1989, the M.Sc. degree in physics from Universidad Complutense de Madrid (UCM) in 1993, and the Ph.D. degree from Universitat Politècnica de Catalunya (UPC) in 2009. In 1989 he joined Telefónica of Spain and was involved in the specifications and first office application of the Telefónica SDH transport network. In 2003 he joined UPC,

where currently he is an Assistant Professor in the Computer Architecture Department (DAC) and a Researcher in the Optical Communications group (GCO) and the Advanced Broadband Communications Center (CCABA). His interests include signaling, routing, and resilience mechanisms in ASON/GMPLS-based networks.



**Jordi Perelló** received his M.Sc. degree in telecommunications engineering in 2005 from the Universitat Politècnica de Catalunya (UPC). Currently, he is an Assistant Professor in the Computer Architecture Department (DAC) at UPC, finishing his Ph.D. studies in the Advanced Broadband Communications Center (CCABA). He has participated in various IST FP-6 and FP-7 European research projects such as EU DICONET, BONE, IST NOBEL 2, e-Photon/

ONe+, and COST Action 291. His research interests concern resource management, quality of service issues, and survivability of next-generation optical transport networks.



**Marc Ruiz** received the B.Sc. degree in biology from Universitat de Barcelona (UB), Spain, in 2005 and the M.Sc. degree in statistics and operational research from Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2009. He is currently working towards the Ph.D. degree with the Optical Communication Group (UPC). His research interests include optimization and performance of next-generation optical networks.



**Salvatore Spadaro** received the M.Sc. (2000) and the Ph.D. (2005) degrees in telecommunications engineering from Universitat Politècnica de Catalunya (UPC). He also received the Dr.Ing. degree in electrical engineering from Politecnico di Torino (2000). He is currently an Associate Professor in the Optical Communications Group of the Signal Theory and Communications Department of UPC. Since 2000 he has been a staff member of the Advanced Broadband

Communications Center (CCABA) of UPC, and he is currently participating in the DICONET and BONE FP7 EU projects. He has co-authored about 80 papers in international journals and conferences. His research interests are in the field of all-optical networks with emphasis on traffic engineering and resilience.



**Gabriel Junyent** is a telecommunications engineer (Universidad Politécnica de Madrid, UPM, 1973) and holds a Ph.D. degree in communications (UPC, 1979). He has been a Teaching Assistant (UPC, 1973–1977), Adjunct Professor (UPC, 1977–1983), Associate Professor (UPC, 1983–1985), and Professor (UPC, 1985–1989) and has been a Full Professor since 1989. In the past 15 years he has participated in more than 30 national and international R&D projects

and has published more than 30 journal papers and book chapters and 100 conference papers.



**Jaume Comellas** received the M.S. (1993) and Ph.D. (1999) degrees in telecommunications engineering from UPC. His current research interests are optical transmission and IP over WDM networking topics. He has participated in many research projects funded by the Spanish government and the European Commission. He has co-authored more than 70 research articles in international journals and conferences. He is an associate professor in the Signal Theory and

Communications Department at UPC, also serving as International Affairs Vice-Dean at the Telecommunications Engineering School (Telecom BCN) of the same university.