Quality of Service in IP over WDM: considering both service differentiation and transmission quality

Elio Salvadori, Roberto Battiti
Università di Trento, Dipartimento di Informatica e Telecomunicazioni
via Sommarive 14, 38050 Povo (TN), Italy
{salvador,battiti}@dit.unitn.it

***Abstract*— IP over WDM networks are a promising candidate for the next generation optical Internet networks. A new Traffic Engineering (TE) scheme is proposed in this paper with the objective to route sub-wavelength connection requests with QoS constraints. In particular, we consider the routing of high-priority connections characterized by stringent requirements in term of delay and packet-loss ratio, by translating them into constraints at the physical layer. Furthermore, in order to provide efficient service differentiation, the impact of a sub-optimal preemption algorithm is analyzed through extensive simulation experiments considering both the blocking probability and the network disruption, while comparing it with an optimal mechanism proposed in literature.**

I. Introduction

IP over WDM (Wavelength Division Multiplexing) optical network is a model of transport network of growing interest. While the use of WDM circuit switching approach in wavelength routed networks allows the exploitation of the optical fiber huge bandwidth, it does not solve by itself the problem of providing Quality of Service (QoS) guarantees for advanced services, such as Voice-over-IP or real-time packet video [1]. Furthermore, the methods presented in the literature to solve QoS issues in IP networks exhibit strong differences when compared to the ones proposed to guarantee QoS in wavelength routed networks. In fact, QoS in optical networks traditionally means *service differentiation* or *transmission quality*. When considering service differentiation, most proposals are based on Lightpath Allocation (LA) algorithms which divide the available lightpaths into different subsets, assigning a service class to each of them. The proposed LA algorithms differ in the way they assign the available lightpaths to the connection requests, according to their class of service [1]. When instead transmission quality is considered, most papers propose Routing and Wavelength Assignment (RWA) algorithms which take into consideration the transmission impairments introduced by the physical layer [2].

The two aspects are seldom considered jointly in these works, thus leading to an incomplete analysis of QoS issues in optical networks. In addition, all proposals assume that each connection request uses the entire capacity of a single lightpath, which is a strong assumption when IP traffic is considered in the network. In fact, current optical technology allows to reach very high transmission speed per wavelength channel (up to 10 or even 40 Gbps), thus it is very important for network operators to be able to multiplex low-speed traffic connections into lightpaths in order to save network costs. The act of multiplexing, demultiplexing and switching low-speed traffic streams into high-capacity optical pipes is defined as *traffic grooming*. Traffic grooming has gained a lot of attention during recent years, and different studies were performed both on regular and arbitrary topologies, and on static or dynamic traffic.

Emerging IP-based applications motivate researchers to study in more depth the grooming problem when highly-dynamic data traffic is carried in an IP over WDM network. Many heuristic algorithms have been proposed lately to deal with dynamic grooming, where authors consider mainly the impact of the grooming node architecture or the effect of different routing algorithms at both IP and WDM level over the call blocking probability [3], [4]. Very few attention has been put so far on the QoS guarantees for the carried traffic, from the point of view of service differentiation *and* transmission quality. In [5] a dynamic grooming algorithm which admits high-priority (HP) requests in preference over the low-priority ones is proposed, by implementing an admission control mechanism which guarantees a limited portion of resources (in term of ports number) to low priority traffic. The main is how to decide the right threshold for the resources dedicated to HP traffic. An MPLS-based preemption scheme to deal with different traffic classes is used in [6], but the main drawback is that class priorities are not mapped onto the optical layer constraints, thus HP traffic can be routed over bad paths in term of delay and packet-loss.

In this paper a set of minimal QoS requirements for a high-priority class of service is defined first at the IP level, and then the corresponding QoS constraints are defined at the optical level. Then a novel Traffic Engineering (TE) scheme is proposed, based on two concurrent heuristics: a dynamic grooming algorithm, which routes incoming connection requests to guarantee their QoS constraints in term of transmission quality, and a preemption mechanism, which provide service differentiation by minimizing the blocking probability for HP requests.

The outline of the rest of the paper is as follows. Section II shows how QoS requirements for some IP-based applications can be mapped onto specific constraints in the optical layer. Section III describes the traffic engineering scheme, while Section IV presents and analyzes results from simulations.

This work was partially supported by the Italian Ministry of Education and Research (MIUR) through the GRID.IT and ADONIS projects

II. QoS requirements on groomed traffic

The network architecture considered in this work is IP over WDM where optical paths are opened on demand. The optical level is based on optical nodes interconnected by fiber links. Two node architectures are considered here: a node can be a pure Optical Crossconnect (OXC), which allows to switch entire lightpaths from an ingress port to an egress port, or it can be an IP/MPLS router (Label Switching Router - LSR) with WDM interfaces, which supports sub-wavelength traffic flows and groom them onto wavelength channels. These nodes can both terminate transit traffic or they can groom it into some optical pipe with incoming IP traffic. If OXCs have wavelength conversion capability, we assume they use electronic wavelength converters only.

In this architecture, a path connecting two routers in the IP layer is called a *virtual* or logical path, because it is created over some established lightpath in the optical layer. IP traffic dynamically follows the virtual topology built by the optical level underneath. A G-MPLS like control protocol is assumed [7], so that each node is always informed of the network status in term of wavelength usage and lightpath occupation.

In the rest of the paper we consider only the routing of bandwidth-guaranteed connections requests, which are carried over Label Switched Paths (LSP) set-up through an MPLS-based signalling plane, such as [8]. The decision to route incoming requests over the existing virtual topology or to establish new lightpaths to create more room for them can lead to different network performance. As a result, a request can be routed over a direct lightpath (a *single-hop* path at the IP level), if it crosses only pure OXCs between an ingress and an egress router, or over a sequence of lightpaths (a *multi-hop* path at the IP level), if it crosses intermediate LSR nodes. Furthermore, a network operator should take into account the specific QoS requirements of the incoming request when deciding its route along the network. In fact, a connection request routed over a single or multi-hop path in the virtual topology would experience different *delays* or *packet losses* according to the physical characteristics of the optical pipe carrying the request.

In the following, the impact of these two parameters is considered in more detail, and some specific constraints are highlighted when specific kind of traffic needs to be routed in the network:

Delay. Most of delay suffered by a traffic flow derives from the queueing delays in IP routers [9] and from optical-to-electronic-to-optical (*o-e-o*) conversion delays in regenerators and electronic wavelength converters [5]. Then we assume that, when some *delay sensitive* application needs to be routed in the optical network, the following constraint must be applied: the corresponding connection request cannot be carried over optical pipes consisting of more than $C_{max}+1$ lightpaths, i.e. it can't experience more than C_{max} o-e-o conversions.

Packet losses. The transmission impairments that digital transmission experiences along a lightpath without intermediate electronic regeneration can impact the packet loss ratio of the connection carried over it. In fact, ASE (amplified spontaneous emission) noise in optical amplifiers, insertion loss and crosstalk introduced by OXCs and attenuation and PMD (Polarization Mode Dispersion) effects introduced by the fibers can degrade the optical signal resulting in a very high BER (bit-error rate) [2]. In the rest of the paper we assume the simplified hypothesis that all the fiber links introduce the same level of transmission impairments[1], thus reducing the problem of selecting a good lightpath for packet-loss sensitive traffic to the problem of limiting the maximum number of hops (up to H_{max} fiber links) for the lightpath which carries such traffic.

In the rest of the paper we consider two classes of service only: a high-priority (HP) class, characterized by minimum end-to-end delay and low packet loss probability (e.g. high-quality real-time services), and a low-priority (LP) class with no QoS requirements, which can experience both high end-to-end delay and frequent retransmission when routed over lightpaths with higher BER or if disrupted due to rerouting (e.g. classical best-effort traffic).

III. A new Traffic Engineering scheme

Given an IP over WDM network with connection requests belonging to different classes of service arriving dynamically, the objective of the proposed traffic engineering scheme is twofold: it must route the request according to its specific QoS requirements while balancing the allocation of the already established connections in the network to minimize the blocking probability for high-priority traffic demands.

A. QoS-based grooming algorithms

Most of the grooming algorithm proposals work on a layered graph, representing the optical network, which is modified after every successfully routed connection. Many representations have been proposed in the last years, among them [3], [4], [5], which differ mostly in the way they represent the optical node architecture. Since we are not interested in the impact of the architecture on the network performance, we consider the simplified network model proposed by [3], which assumes that all the LSRs in the network have enough ports to process all the traffic flowing through them.

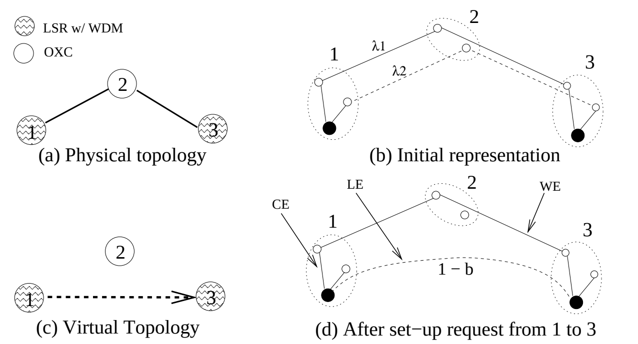

Fig. 1. The layered graph representation of an optical network

[1]This assumption can be relaxed by considering a more realistic network such as in [2], but at this stage we believe it is reasonable enough to study the specific problem of guaranteeing different QoS requirements to sub-wavelength connection requests.

In Figure 1 (a) a simple network with two wavelengths per fiber is considered, and its corresponding *extended graph* $\mathcal{G}$ is shown in (b), where LSRs are represented with *super-nodes* connecting all wavelength layers. According to [3], if the initial (full) capacity of each edge in $\mathcal{G}$ is normalized to 1, when a new sub-wavelength connection needs to be routed from node 1 to node 3, it is groomed into the lightpath crossing node 2. Thus two graph edges are removed from $\mathcal{G}$ and a new direct edge with capacity equals to $1 - b$ is added as shown in (d), where b is the amount of bandwidth required by the incoming request.

Let us define as $\mathcal{E}$ the set of edges of $\mathcal{G}$. As shown in Figure 1 (d), three possible kinds of edges can be identified in this set, each having a property tuple $P(c, w, h)$, where c is the edge capacity (if g is the wavelength capacity in bandwidth units, $c = g$ means full capacity), w the associated weight and h the cost metric which models the signal degradation introduced by the transmission link:

- *Wavelength Edges (WE)*. An edge $e \in \mathcal{E}$ from node i to j on wavelength layer k is a WE if there is a physical link from i to j and wavelength λ_k is free on this link. For such an edge: $c = g$ and $h = 1$.
- *Lightpath Edges (LE)*. An edge $l \in \mathcal{E}$ from node i to j on wavelength layer k is a LE if there is a direct lightpath from i to j on wavelength λ_k. For such an edge: $c = g - \sum_{i=1}^{M} b_i$, if M connections (LSPs) with bandwidth b_i are running over it ($b_i < g$), and $h = H_l$, if the lightpath crosses H_l fiber links.
- *Converter Edges (CE)*. These are all the so-called fictitious edges $f \in \mathcal{E}$ between the super-nodes in $\mathcal{G}$ and the nodes belonging to the wavelength layers. For such an edge: $c = \infty$ and $h = 0$.

When a request arrives in some ingress router s destined to some egress router d, there are four possible operations that can be applied [4]. For each of them it is possible to modify the weights assigned to all the edges in $\mathcal{G}$ in order to get different objectives when applying the shortest-path algorithm over it.

1) route the traffic onto an existing lightpath connecting directly (in one hop) s to d.
2) route the traffic over the current virtual topology (VT).
3) set up a new lightpath connecting s to d.
4) set up a set of new lightpaths, which do not connect s and d directly, and route the traffic through them and some existing lightpaths in order to maximize the usage of the VT.

Each operation can be applied only if some prerequisites are satisfied: for example, if s and d are disconnected in the current VT, both 1) and 2) cannot be applied. As in [4], these operations are applied sequentially in different priority order when a new connection request needs to be routed, thus different grooming objectives can be achieved by modifying the sequence of operations. But unlike [4], the constraints which characterize high-priority requests lead here to some specific difference. In particular, *RouteMixed* has different implementations according to the request's class of service: for a LP request the objective is to maximize the existing lightpaths usage, while for a HP request instead, the objective is to assign a path which minimizes the number of conversions in the network (see [10] for further details).

In our analysis of the TE scheme efficiency in guaranteeing specific QoS requirements to the incoming requests, we consider only few grooming policies for incoming requests. Each time a new HP request arrives at some ingress router s, the following two QoS-based grooming policies are considered:

- `VT-first`, which maximizes the VT usage, by applying these operations in sequence: $1) \Rightarrow 2) \Rightarrow 3) \Rightarrow 4)$
- `PT-first`, which maximizes the optical resources usage, by applying in the sequence: $1) \Rightarrow 3) \Rightarrow 2) \Rightarrow 4)$

If a LP request must be routed, only one "conservative" grooming policy is applied: $1) \Rightarrow 2) \Rightarrow 4)$

B. Preemption algorithms to provide service differentiation

Unfortunately, the main consequence of a QoS-based grooming algorithm is that requests with tighter requirements (usually the more profitable for Service Providers) are likely to be blocked by the requests with less or no QoS requirements. In order to guarantee lower blocking probability to higher priority classes, different mechanisms can be applied, such as call admission control or preemption. In the following, we focus our attention on this last mechanism, by considering two different algorithms, in order to study their impact when used jointly with the grooming algorithms presented in the previous section. In particular, we propose a sub-optimal preemption algorithm which allows to minimize the number of reroutings and the signalling overhead in an IP over WDM network, and compare its performance with a well-known optimal algorithm proposed in literature.

In connection-oriented networks where traffic belonging to different priority classes is considered, many distributed algorithms have been proposed, which are optimal with respect to their objective functions. A well-known algorithm is *Min_Conn* [11] which optimizes the criteria of (i) number of connections to be preempted, (ii) the bandwidth to be preempted and (iii) the priority of connections to be preempted, in that order. When this optimal mechanism is used in a G-MPLS based network, a large amount of RSVP messages around the network is generated each time a LP request needs to be preempted. In such cases in fact, an intermediate router who selected one or more connections to preempt needs to send upstream the proper notification messages to the edge routers, which should try to reroute (or block) these connections.

It is worth noticing that *Min_Conn* can potentially involve a lot of edge LSRs to finalize the set-up of one high-priority connection. More important, when considering the great amount of information needed for G-MPLS signalling, this mechanism could lead to very high amount of overhead in the network. In such networks there's the need then to consider preemption mechanisms based on a simpler implementation both from the algorithmic and the signaling point of view. Due to space constraints, we only present here an overview of the proposed *Local Preemption Algorithm* (`LPA`). For more details, see [10]. The idea is to run a "limited" version of

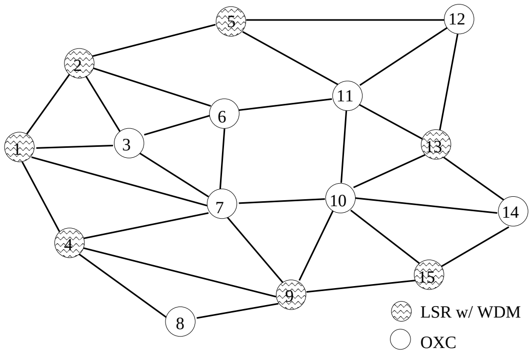

Fig. 2. Topology considered in the simulations.

Min_Conn only in the ingress router that experiments a high-priority request blocking. Because each LSR has complete information about the crossing connections, we restrict the search space of the preemption algorithm to all the LSPs which originate, terminate or cross the node itself. Unlike the original mechanism in [11], here we consider the specific constraints on the trasmission quality described in Section II. In particular, in order to find a candidate route for high-priority requests, `LPA` considers the reduced graph $\mathcal{G}' \subset \mathcal{G}$ made of LEs which respect the constraint $h_l \leq H_{max}$.

`LPA` is triggered in the ingress router s each time a high-priority request from s to d is blocked. The algorithm performs a simple local search through all the connections originated or crossing node s, with final or intermediate destination d, looking for the best low-priority LSP (or LSPs) to preempt in order to leave its route to the incoming request. When looking for connections to preempt in the network, it first searches for LSPs carried over some direct single-hop lightpath from s to d , and then for LSPs carried over a multi-hop lightpath. In fact, due to the constraint over the maximum number of o-e-o conversions, it is better to route HP traffic over direct lightpaths, while leaving multi-hop paths to LP traffic.

IV. Simulation results

The performance of the proposed TE scheme has been evaluated through extensive experimentation. Simulations have been performed on different network topologies, but thanks to the consistency of the results, only the graphs for the topology in Figure 2 are shown, where no wavelength conversion capability is considered in the OXCs.

Requests arrive between each ingress-egress pair according to a Poisson process with an average rate λ, and their holding times are exponentially distributed with mean $1/\mu$. Each wavelength has a full capacity $g = 10$ units, and connection requests have bandwidth demand b_i distributed uniformly between 1 and 3 units, independently from their priority. The number of wavelengths per fiber considered in all the tests is $K = 4$. The network is loaded with 50000 requests during one trial, and the performance is evaluated by considering average values calculated over 10 runs. The percentage of traffic routed in the network is 60% for LP traffic and 40% for HP traffic. The physical constraints $H_{max} = 4$ and $C_{max} = 1$ have been chosen for the high-priority class in the experiments[2].

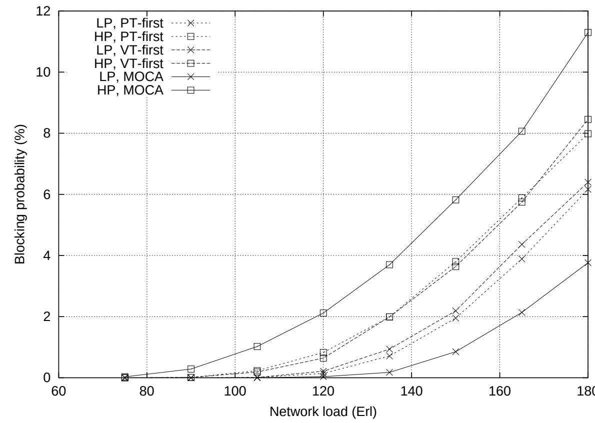

Fig. 3. Blocking probability for MOCA vs. `PT-first` and `VT-first`

The first set of tests compares the QoS-based grooming algorithms `VT-first`, `PT-first` with the Minimum Open Capacity Algorithm (MOCA) [3]. Figure 3 shows the corresponding blocking probability, when traffic requests are limited only to some specific ingress and egress router pairs (MOCA can work only when this strong assumption is considered). The standard deviations are hardly visible (in the order of 10^{-4})) and therefore not shown in this and other graphs related to blocking probability. MOCA performs better for LP traffic, when no requirements are needed to route successfully a connection, while its performance is much worse than our grooming algorithms for HP traffic. This behavior is justified by the fact that MOCA is a load balancing grooming algorithm, then spreading connections over longer paths on average, it performs very badly when HP traffic must be routed over the network.

As expected, using QoS-based grooming algorithms only penalizes high-priority traffic. In the following, we analyze the performance of our grooming algorithms when both *Min_Conn* and `LPA` preemption mechanisms are applied.

Figure 4 shows the performance of both *Min_Conn* and `LPA` when `PT-first` grooming algorithm is applied, and relaxing the assumption on the position of the ingress-egress router pairs, which are randomly selected every time a new request is loaded in the network. `VT-first` performs very similarly, thus results are not included for space reasons. As expected, by using a preemption mechanism, the blocking probability for HP traffic increases dramatically for HP traffic to the detriment of LP traffic. In particular, `LPA` performs quite well compared to the optimal algorithm, which always finds a route for HP requests. In fact, the blocking probability is quite low (less than 5%) even at high network loads, when the LP traffic experiences a higher blocking probability.

Figure 5 shows how the two types of traffic fare when one considers the average number of o-e-o conversions in the

[2] In fact these values are very dependent on the topology of the optical network; in particular H_{max} depends very much on the network diameter, while having $C_{max} = 1$ can be very restrictive for topologies where few core nodes have LSR capabilities.

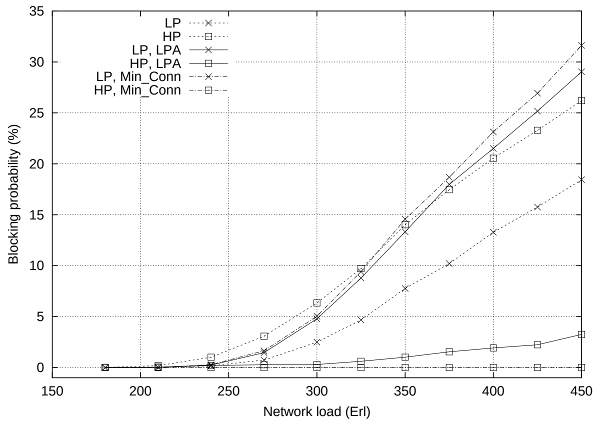

Fig. 4. Blocking probability for *Min_Conn* and `LPA` with `PT-first` grooming algorithm

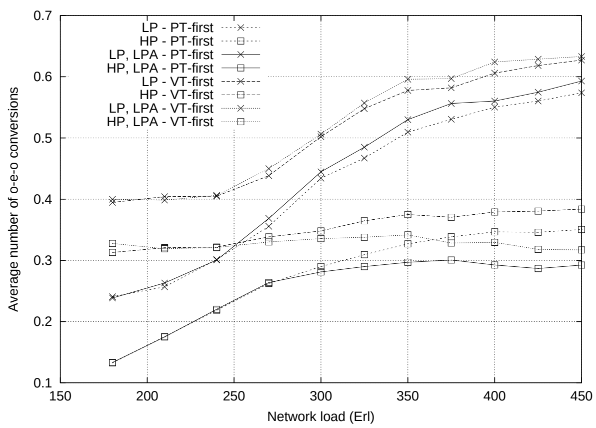

Fig. 5. Number of o-e-o conversions when using `VT-first` and `PT-first`

absence of service differentiation and when the proposed `LPA` preemption algorithm is adopted, jointly with `VT-first` and `PT-first`. This value gives in fact an estimate of the delay incurred by services carried over the two types of connections. When using `PT-first` the corresponding delay is lower for both classes of service. This can be explained by the implicit mechanism used by this grooming algorithm, which gives preference to direct lightpath when a new request arrives, thus reducing the average number of hops at IP level. Another interesting result is that the proposed preemption mechanism does not impact dramatically on the overall delay for high-priority requests, which is kept to very low values even at high network loads.

TABLE I

PERCENTAGE OF REROUTED AND BLOCKED LP CONNECTIONS

$\lambda/\mu = 450$	LPA		*Min_Conn*	
	VT-first	PT-first	VT-first	PT-first
Rerouted	19.16 %	18.3 %	32.49 %	27.94 %
Blocked	9.21 %	9.19 %	12.36 %	11.5 %

Another important aspect to consider when using preemption is its impact on network disruption. Table I shows the percentage of rerouted and blocked connections at high network load, calculated as number of rerouted (blocked) LSPs over the total number of low-priority LSPs routed with success. As expected, when using `LPA`, these ratios are much lower than in the optimal case. In both cases, the lowest ratios are obtained by using the `PT-first` grooming policy, again due to its stronger use of optical resources.

V. CONCLUSIONS

In this paper a new scheme for the routing of bandwidth guaranteed LSPs with different QoS requirements in IP over WDM networks has been proposed. The traffic engineering scheme consists of a distributed two-stage scheme: when a new request arrives, an on-line dynamic grooming scheme finds a route which fulfills some specific constraints in terms of maximum end-to-end delay and packet-loss ratio. If a high-priority request is blocked, a preemption algorithm is executed in order to create room for this traffic.

We demonstrate by using simulations that a sub-optimal preemption algorithm can guarantee comparable performance in comparison with an optimal mechanisms proposed in literature, while having a minimum impact over the transmission quality of the routed connections and a reduced network disruption.

ACKNOWLEDGMENTS

We would like to thank dr. Filippo Ardito for the implementation of the algorithms in the C++ software.

REFERENCES

[1] A. Kaheel, T. Khattab, A. Mohamed, and H. Alnuweiri. Quality-of-Service Mechanisms in IP-over-WDM Networks. *IEEE Communications Magazine*, pages 38–43, December 2002.

[2] B. Ramamurthy, D. Datta, H. Feng, J. Heritage, and B. Mukherjee. Impact of transmission impairments on the teletraffic performance of wavelength-routed optical network . *IEEE/OSA Journal of Lightwave Technology*, 10(17):1713–1723, October 1999.

[3] M. Kodialam and T.V. Lakshman. Integrated Dynamic IP and Wavelength Routing in IP over WDM Networks. In *Proceedings of INFOCOM*, volume 1, pages 358 –366, Anchorage - Alaska, 2001.

[4] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee. A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks. *IEEE/ACM Transactions on Networking*, 11(2):285 –299, April 2003.

[5] E.C. Tien and G. Mohan. Differentiated QoS routing in GMPLS-based IP/WDM networks. In *Proceedings of GLOBECOM*, volume 3, pages 2757–2761, 2002.

[6] A. Bosco and A. Botta *et al.* Distributed Implementation of a Pre-Emption Mechanism for a Network Control Based on IP/MPLS Paradigm. In *Proceedings of the 7th IFIP ONDM*, 2003.

[7] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella, and Y. Rekhter. Generalized Multiprotocol Label Switching: an Overview of Routing and Management Enhancements. *IEEE Communications Magazine*, 39(1):144–150, January 2001.

[8] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella, and Y Rekhter. Generalized Multiprotocol Label Switching: an Overview of Signaling Enhancements and Recovery Techniques. *IEEE Communications Magazine*, 39(7):144–151, July 2001.

[9] A. Parekh and R. Gallager. A generalized processor sharing approach to flow control in integrated services networks: the multiple node case. *IEEE/ACM Transactions on Networking*, 2(2):137 –150, April 1994.

[10] E. Salvadori and R. Battiti. Traffic Engineering in G-MPLS networks with QoS guarantees, DIT-03-050. Technical report, Università di Trento, Dipartimento di Informatica e Telecomunicazioni, September 2003.

[11] M. Peyravian and A.D. Kshemkalyani. Connection preemption: issues, algorithms, and a simulation study. In *Proceedings of INFOCOM*, volume 1, pages 143 –151, 1997.