UNIVERSITY
OF TRENTO

DEPARTMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY

38050 Povo – Trento (Italy), Via Sommarive 14
http://www.dit.unitn.it

TRAFFIC ENGINEERING IN G-MPLS NETWORKS WITH QOS GUARANTEES

Elio Salvatori and Roberto Battiti

September 2003

Technical Report # DIT-03-050

Traffic Engineering in G-MPLS networks with QoS guarantees

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

***Abstract*— In this paper a new Traffic Engineering (TE) scheme to efficiently route sub-wavelength requests with different QoS requirements is proposed for G-MPLS networks. In most previous studies on TE based on dynamic traffic grooming, the objectives were to minimize the rejection probability by respecting the constraints of the optical node architecture, but without considering service differentiation. In practice, some high-priority (HP) connections can instead be characterized by specific constraints on the maximum tolerable end-to-end delay and packet-loss ratio. The proposed solution consists of a distributed two-stage scheme: each time a new request arrives, an on-line dynamic grooming scheme finds a route which fulfills the QoS requirements. If a HP request is blocked at the ingress router, a preemption algorithm is executed locally in order to create room for this traffic. The proposed preemption mechanism minimizes the network disruption, both in term of number of rerouted low-priority connections and new set-up lightpaths, and the signaling complexity. Extensive simulation experiments are performed to demonstrate the efficiency of our scheme.**

I. Introduction

Generalized Multi-Protocol Label Switching (G-MPLS) is emerging as the control-plane solution for next-generation IP over WDM optical networks [1]. G-MPLS is an extension to MPLS which enables Generalized Label Switched Paths (G-LSPs) such as lightpaths to be automatically set-up or torn down by means of a signaling protocol. One of its most interesting application is Traffic Engineering (TE), whose main objective is to optimize the performance of a network through an efficient utilization of resources and to provide for Quality of Service (QoS) guarantees.

In an IP over WDM network, optical nodes can be optical cross-connects (OXC) only or IP/MPLS routers, namely Label-Switched Routers (LSRs), connected to OXCs or even with WDM interfaces on their ports. A sub-wavelength connection request (an *electronic* LSP) could be routed over a direct lightpath (a single-hop path at the IP level) connecting an ingress router to an egress router or over a sequence of lightpaths (a multi-hop path at the IP level), crossing many intermediate LSRs along its route. G-MPLS inherits most of the MPLS mechanisms to properly route sub-wavelength requests over optical connections [1]. As in MPLS, the ingress routers use the network resource information (periodically updated through specific link-state routing protocols, e.g. OSPF-TE [2]) to perform explicit routing of electronic LSPs, by using some constraint-based routing (CBR) scheme. When new lightpaths must be set-up in the optical network to route incoming requests, G-MPLS runs some routing and wavelength assignment (RWA) algorithm. Once the LSP path is decided, the ingress router uses a signaling mechanism such as RSVP-TE to effectively route the connection [3].

The problem of routing sub-wavelength requests in optical networks is called *traffic grooming* and it has been proved to be NP-hard [4], [5]. When dynamic traffic is considered this problem is called *dynamic traffic grooming* or *integrated routing*, to highlight the advantage of having a unified control plane to exploit resources both at the IP layer and the WDM optical layer. Many heuristic methods have been developed to deal with this problem [4], [6], [7], where most of the attention is focused on optimizing the physical resources usage by considering specific constraints on the optical node architecture (i.e. number of router ports, number of fiber per link, number of wavelength converters in the network, etc.).

While a significant amount of research has been done on guaranteeing Quality of Service in pure IP-based networks, the problem of providing QoS guarantees to different services carried over high-capacity optical channels remains largely unsolved for wavelength-routed networks [8]. In the literature, the concept of “Quality of Service” in such networks assumes two main meanings: *service differentiation* or *transmission quality*, seldom jointly considered. In the first case most papers propose Routing and Wavelength Assignment (RWA) algorithms which allow to dynamically assign a set of wavelengths to higher-priority traffic in order to maximize the revenue for the service provider [8]. In the second case the transmission impairments introduced by the physical layer are considered in the RWA algorithms, which assign optical paths to incoming requests only if the resulting lightpath is feasible (i.e. the output SNR is good enough to guarantee the required transmission quality to the bitstream). In this case no service differentiation is considered because the same QoS is guaranteed for all the carried traffic [9], [10].

All these proposals consider the routing of an entire lightpath over a single- or multi-hop path. While an optical network with traffic grooming is considered, very little attention has been paid so far on the effects of both the physical constraints characterizing the optical layer and the delay restrictions of a multi-hop path over the traffic carried on the wavelengths. If

This work was supported by the Italian Ministry for University and Scientific Research under the ADONIS (Algorithms for Dynamic Optical Networks based on Internet Solutions) project and FIRB-CNIT “Grid computing”.

for example some packet-loss sensitive traffic (e.g. real-time video) is carried over a lightpath experiencing strong transmission impairments, the resulting SNR degradation could be so high to compromise the signal quality requirements. If a connection carrying delay-sensitive traffic (e.g. Voice-over-IP) is routed over a multi-hop path, the output signal could suffer very high end-to-end delay due to the optical-to-electronic (o-e-o) conversions and queueing delays experienced along the path.

To guarantee the specific requirements of different applications, a "QoS-aware" dynamic grooming scheme should then consider these constraints when a new connection must be routed. Unfortunately such a mechanism could not be sufficient, because requests with tighter requirements (usually the more profitable from the Service Provider viewpoint, thus with High Priority - HP) are likely to be blocked by the requests with less or no QoS requirements (Low Priority - LP). In order to guarantee lower blocking probability to higher priority classes, different mechanisms can be applied. In [11], a dynamic grooming algorithm which admits HP requests in preference over the LP 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 disadvantage of such a technique is how to decide the right threshold for the resources dedicated to LP traffic, which can lead to block many requests even if the network has enough room for them. A more convenient option to deal with different classes of traffic is to implement a *preemption* scheme, a well-known mechanism both in general connection-oriented networks [12] and in pure MPLS networks [13], [14]. This mechanism is based on routing all the requests independently from their priority and on preempting lower-priority connections if a higher priority request can't find a path in the network. An application of this mechanism in the context of a "G-MPLS-enabled" network can be found in [15]. The drawback is that class priorities are not mapped onto the optical layer constraints, thus the HP traffic can be routed over bad paths in term of delay and packet-loss.

In this paper a novel Traffic Engineering scheme for G-MPLS networks is proposed, which is based on a dynamic grooming algorithm which finds a route to fulfill the QoS requirements of the incoming request and on a preemption mechanism which guarantees lowest blocking probability to HP traffic. The novelty of this integrated approach is on guaranteeing the Quality of Service in a wavelength-routed network from both the service differentiation and the transmission quality point of view: to the best of our knowledge, no heuristic algorithms for solving this problem for sub-wavelength requests arriving dynamically has been developed in previous work. Furthermore, the proposed preemption mechanism is based on an algorithm executed locally in the ingress router which minimizes the network disruption and the signaling complexity. A high number of rerouted traffic flows could potentially lead to instantaneous wavelength disruption in the network, which could have a dramatic impact over its stability. Moreover, while considering the rerouting of one or more electronic traffic flows, the complexity of the related signaling mechanism must be considered, because compared to MPLS, in G-MPLS the signaling overhead is much bigger due to the increased number of layers to consider.

The paper is organized as follows. Section II explains the motivations for our proposal and the problem definition, while Section III explains the proposed integrated scheme. Then Section IV analyzes the results.

II. Problem definition and system model

The considered network consists of n nodes interconnected by m optical links, where each link can carry up to K wavelengths. Each node in the network can be an LSR router with WDM interfaces or an OXC with or without wavelength conversion capability. Let R be the set of routers, T the set of OXCs without wavelength converters and S the set of OXCs with wavelength converters. The multiplexing or demultiplexing of electronic LSPs at any granularity is possible only in the nodes in R, where wavelength conversion is also possible, while in the nodes belonging to set T and S it is possible to perform traffic switching at wavelength granularity only. As in [6], we assume that all the LSRs in the network have enough ports to process all the traffic flowing through them [1], while edge LSRs can also act as intermediate LSRs to LSPs established between other edge LSRs. This means that in the multi-hop case, a connection can be dropped at an intermediate router and multiplexed with other low-speed connections on different lightpaths before it reaches its destination.

A connection request is defined by a quadruplet $D(s, d, b, q)$, where s and d specify the ingress and egress routers, b indicates the amount of bandwidth required and q specifies the Class Type (CT), by using the MPLS terminology to aggregate classes of traffic with similar requirements [16]. In the rest of the paper we will consider only the routing of bandwidth-guaranteed connections. We assume also an on-line context with connection requests arriving one at a time. When a new request arrives, the operator determines if it can be routed on the current set of lightpaths (the so-called Virtual Topology - VT) or if the VT needs to be modified by setting up new lightpaths in order to accommodate the request. Different decisions reflect different objectives in term of network resource utilization, and are referred to as *grooming policies* [4]. As a result, a request can be routed over a direct lightpath (a single-hop path at the IP level), if it crosses only nodes belonging to the set T between an ingress and an egress router, or over a sequence of lightpaths (a multi-hop path at the IP level), if it crosses nodes belonging to the set R and S as well. Note that a lightpath in the optical domain corresponds to a single wavelength crossing a certain number of nodes in T, because for nodes belonging to S we consider only electronic wavelength converters.

A connection request carried into a single or multi-hop path in the virtual topology can experience *delay* and *packet loss*.

[1]This assumption can be relaxed in this model by considering the case where an LSR has limited processing power.

Most of delay suffered by a request derives from the queueing delays in IP/MPLS routers and from o-e-o conversion delays in regenerators and electronic wavelength converters [11], [17]. In fact it has been proved that in a network implementing WFQ (weighted fair queueing) scheduling with leaky bucket traffic shaping at ingress nodes, the maximum packet delay a request can tolerate is proportional to the number of crossed LSRs [18]. In this work we assume that *delay-sensitive* applications can cross no more than C_{max} nodes belonging to the set R and S along their route, which means that their multi-hop path cannot consists of more than $C_{max} + 1$ lightpaths. The transmission impairments that digital transmission experiences along a lightpath can impact the packet loss ratio of the connection carried over the optical path [17]. 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) [9], [10]. In this work we assume that some packet-loss sensitive traffic can be routed only over lightpaths characterized by stringent BER requirements, while other applications can tolerate higher BER (i.e. they can be routed over any set-up lightpath), by managing packet losses through retransmission. Furthermore, by assuming the simplified hypothesis that all the fiber links introduce the same level of transmission impairments (i.e. all the links have equal length and the same type and number of optical amplifiers), we can reduce the problem of selecting a good lightpath for packet-loss sensitive traffic to the problem of limiting the maximum number of hops for the lightpath which carries this type of traffic. Note that this is valid for both the single- and the multi-hop case: in fact, in this last case if only one of the lightpaths introduce some strong impairments, producing a high BER, the carried applications would be penalized even though intermediate electronic regeneration. This assumption can be relaxed by considering a more realistic network such as in [9], [10], but at this stage we believe it is reasonable enough to study the specific problem of guaranteeing different QoS requirements to the sub-wavelength connection request in an optical network. Then we assume that *packet loss-sensitive* applications can be routed over lightpaths whose route is made of no more than H_{max} fiber links each.

In the rest of the paper we consider two CTs only: a HP class, characterized by minimum end-to-end delay and low packet loss probability (high-quality real-time services, such as interactive video), and a LP class with no QoS requirements (e.g. Best-Effort), 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.

Given a G-MPLS network with connection requests belonging to different CT arriving dynamically, the objective of our on-line TE scheme is twofold: first, it must route the request according to the specific QoS requirements and second, it must balance the allocation of the already established LP connections in the network to maximize the success probability for HP traffic demands.

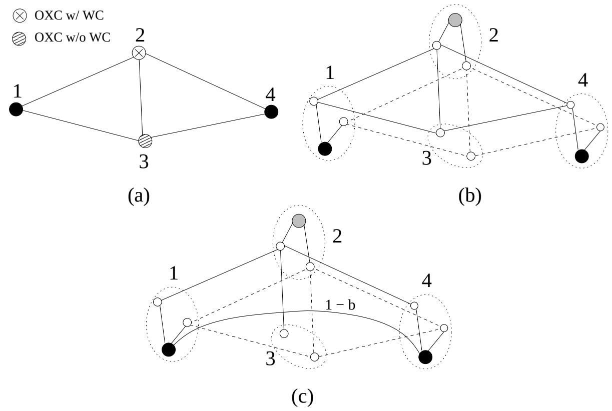

Fig. 1. The layered graph representation of an optical network

A. A layered graph representation

The optical network is modelled as a layered graph [6], where wavelengths on a link are separated into different graph edges. The routing algorithm used to dynamically route the connection request works on a graph which is modified after every successful connection.

In Figure 1 a network with two wavelengths per fiber is considered: the *extended graph* $\mathcal{G}$ in (b) is obtained by expanding each node in the network (a) into a number of sub-nodes, one per wavelength, and then by connecting each sub-node to a wavelength on each incoming and outcoming link. Let define as $\mathcal{E}$ the set of edges of this graph. LSRs and OXCs with wavelength converters are represented adding some *supernodes*, that topologically connect all lambda layers through “fictitious” links with infinite capacity.

If the initial (full) capacity of each edge is normalized to 1, each time the routing algorithm finds a path between an ingress-egress pair in $\mathcal{G}$ (e.g. from node 1 to 4 through 3 in (c)) we modify it by removing the graph edges traversed by a lightpath and by adding a direct edge, called *cut-through* with capacity equals to $1 - b$. The extended graph allows to model the wavelength availability per link and the residual bandwidth per logical link at the IP layer (see [6] for further details).

When an established lightpath is torn down because the last connection occupying it is ended, the cut-through arc is removed and the edges in the extended graph corresponding to the underlying physical links are set back with full capacity.

Different grooming policies can be realized by modifying the weights of the edges in $\mathcal{E}$ and then by running the Dijkstra shortest-path-first (SPF) algorithm on the extended graph. In fact, these weights reflect the cost of network elements such as o-e-o converters (routers) or free wavelength on some link. As it will be shown later, by modifying them according to the incoming CT request, it is possible to choose a path which minimizes the number of o-e-o conversions or which maximizes the usage of existing lightpaths.

Three possible kind of edges can be identified in $\mathcal{E}$, 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 (cut-through) 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 $h = 0$.

III. A Traffic Engineering scheme for QoS routing

The proposed Traffic Engineering scheme for QoS routing in IP over WDM network is composed of two main components: a dynamic grooming algorithm which routes connection requests with specific QoS requirements according to the constraints of the physical layer, and a preemption mechanism which is triggered every time an HP request is blocked and whose function is to preempt LP connections in the network, thus guaranteeing lowest blocking probability to high-priority connections.

A. QoS-aware Dynamic Grooming

Many dynamic grooming algorithms have been proposed recently, among them [6], [4], [7]. Unfortunately, no comparisons have been done among all these proposals in order to understand which algorithm performs the better, mainly because almost each proposal uses a specific representation model for the multi-layer IP over WDM network. For all schemes the basic idea is to perform a constraint-based routing algorithm in order to maximize the utilization of the network resources thus minimizing the blocking probability. The QoS requirements are not considered in all the proposed algorithms, thus no attention is paid on the impact of the route selection both in term of delay and signal degradation, which instead are fundamental when an operator wants to guarantee a certain quality to specific class types.

In order to take into consideration the specific physical constraints described in the previous Section, new "QoS-aware" dynamic grooming algorithms are needed. Then a HP request can be routed on an optical path characterized by a a maximum number of o-e-o conversions C_{max}, where every single lightpath is made of no more than H_{max} fiber links each. An LP request instead can be routed on any sequence of lightpaths, while crossing whatever number of router or OXCs.

As specified in [4], when a request $D(s, d, b, q)$ arrives in some ingress router, there are four possible operations that can be applied. For each of them it is possible to modify the weights assigned to all the edges in the extended graph in order to get different objectives when applying the SPF algorithm.

- *RouteDirect*: route the traffic onto an existing lightpath connecting directly s to d. In this case, no Dijkstra algorithm is applied on $\mathcal{G}$, but instead the request is routed over a LE having enough residual bandwidth to route the request.
- *RouteVT*: route the traffic over the existing VT. The SPF algorithm runs over a skimmed graph where all the WEs and LEs with capacity less than b are removed.
- *RouteNew*: set up a new lightpath l connecting s to d. The SPF algorithm runs over a skimmed graph where all the LEs are removed[2].
- *RouteMixed*: set up a certain 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. This is usually applied when some of the previous operations has not find any path, by trying to maximize the usage of the existing lightpaths just by adding as few new lightpaths as possible in the optical network.

Compared to the operations proposed in [4], here the constraints which characterize higher priority requests lead to some specific difference. For example, when a HP request must be routed, all the LEs l with $h_l > H_{max}$ are skimmed from $\mathcal{E}$ in both *RouteDirect* and *RouteVT*, while in *RouteNew* the resulting path is accepted only if $h_l \leq H_{max}$. Moreover, for an HP request, the resulting path is accepted if $h_l \leq H_{max}$ and the number of o-e-o conversions is no more than C_{max}. In particular, *RouteMixed* has different implementations according to q: for an LP request the objective is to maximize the existing lightpaths usage by assigning a small weight δ (with $\delta \ll 1$) to all LEs, while for all WEs $w_e = 1$. For a HP request instead, the objective is to assign a path which minimizes the number of conversions in the network. Here $w_e = \delta$ for all WEs apart from the ones going from a node belonging to set T to a node belonging to set R or S, for which $w_e = 1$: in fact, a path which crosses such an edge will experience an o-e-o conversion. For the same reason, for all the LE respecting $h_l \leq H_{max}$, $w_l = 1$ otherwise $w_l = \infty$.

Each operation can be applied only if some prerequisites are satisfied: for example when no direct lightpaths connect node s to d, *RouteDirect* cannot be applied.

Different grooming objectives can be achieved by modifying the sequence of operations. In this paper we are not interested in studying the impact of different grooming policies on the overall blocking probability *per-se* but instead we are interested in their impact over the entire TE scheme to guarantee specific QoS requirements for the incoming requests. In particular, every time a new HP request arrives at some ingress router s, we consider two "extreme" grooming policies:

[2]Here different Wavelength Assignment algorithms could be applied, Random-Fit, First-Fit, Least-Loaded or Most-Loaded among the others. In our simulations we apply the first one, because the obtained performances were quite similar.

1) `VT-first`: Maximize the Virtual Topology usage
RouteDirect ⇒ *RouteVT* ⇒ *RouteNew* ⇒ *RouteMixed*
2) `PT-first`: Maximize the optical resources usage
RouteDirect ⇒ *RouteNew* ⇒ *RouteVT* ⇒ *RouteMixed*

Basically, `VT-first` always tries to route the request over the existing lightpaths before modifying the VT; while the utilization of the optical resources is definitely improved, the required electronic processing at intermediate hops is also increased. `PT-first` instead increases the logical connectivity at IP level (the resulting VT will be more connected), but leads to a heavy usage of the available optical resources.

While instead an LP request must be routed, in order to minimize its impact over the network, a new direct lightpath from the ingress to the egress router should be installed only when no other possibilities are available. Then only one grooming policy is applied in this case:

RouteDirect ⇒ *RouteVT* ⇒ *RouteMixed*

Note that the last operation includes the possibility to set-up a new direct lightpath as a last option, only when the related path costs less than setting up a mixed path with existing and new lightpaths.

If the incoming request cannot find a path, the traffic must be blocked if it is low-priority, while a preemption algorithm is applied if it is high-priority.

B. Connection Preemption Algorithm

In MPLS the preemption is implemented in a distributed way by using the RSVP signaling protocol. Basically, when an LSP having priority p (where $0 \leq p < 7$ and 0 is the highest priority) needs to be set-up, its priority is sent in the `PATH` message along the route selected in the ingress router. Each time this message reaches an intermediate LSR, if the available capacity on the outgoing link is not sufficient to carry the request, a local selection of the lower-priority LSPs to be preempted is performed by the router, and then the proper notification messages is sent upstream to the routers which should try to reroute (or block) these LSPs. This mechanism can potentially involve all the edge LSRs in a network to finalize the set-up of one high-priority LSP.

Many distributed algorithms for routers implementing a preemption mechanism have been proposed in literature to select the best LSP to reroute, according to different objectives (typically: minimizing the number of reroutings in the network). In [12], [13] the proposed algorithms are optimal with respect to their objective functions: in the rest of the paper, we indicate them as *global* preemption algorithms (GPA), to highlight this optimal behavior and the involvement of many edge routers in the network. Thus from the point of view of the signaling, they could result in a large amount of RSVP messages around the network which increase the amount of overhead. When considering a network based on Generalized MPLS, the complexity of this mechanism is even more complex, because the signaling must take into account much more information regarding the optical layer.

The proposed TE scheme considers a preemption mechanism which is based on a simpler implementation both

```
1.  for each LP LSP_i crossing (s ∩ d) in (LE_i^sd = 1) hop
2.          assignWeightSH (LSP_i)
3.  <sortedLSPSet> ← sorting (LSP_i)
4.  if (sortedLSPSet ≠ ∅)
5.     <rerLSP_j> ← pickLowest (sortedLSPSet)
6.     dynGroomAlgo (rerLSP_j)
7.  else for each LP LSP_i crossing (s ∩ d) in (LE_i^sd) hops
8.             assignWeightMH (LSP_i)
9.     <sortedLSPSet> ← sorting (LSP_i)
10.    if (sortedLSPSet ≠ ∅)
11.       <rerLSP_j> ← pickLowest (sortedLSPSet)
12.       dynGroomAlgo (rerLSP_j)
```

Fig. 2. The local preemption algorithm (LPA)

from the algorithmic and the signaling point of view. In the following, the proposed algorithm is called *local* preemption algorithm (LPA) to distinguish it from the optimal one (GPA). The idea comes from the property of LSR routers in MPLS to have complete information about the crossing LSPs (each node mantains information about all the LSPs that originate, terminate or cross the node itself [15]). Then, by focusing only on the ingress routers of the network, the simplest suboptimal preemption mechanism to apply is to perform local selection of one or few low-priority LSPs to preempt when a high-priority LSP is blocked.

Figure 2 shows the pseudo-code of the *local* preemption algorithm in a G-MPLS network, which is triggered every time a HP request $D(s, d, b, q)$ is blocked in the ingress router s. The algorithm performs a simple local search through all the low-priority LSPs originated or crossing node s, with final or intermediate destination d, looking for the best LSP (or LSPs) to preempt in order to leave its route to the incoming request. When looking for (one or more) LSPs to preempt in the network, LPA first searches for LSPs carried over some direct lightpath from s to d (lines 1–6, where LE_i^{sd} is the number of crossed Lightpath Edges), and then for LSPs carried over a multi-hop lightpath (lines 7–12), simply because it is always better to route HP traffic over direct lightpaths, while leaving multi-hop paths to LP traffic.

As shown in the Figure, there is a different function which assigns the weights to the LSPs in the two cases: the single- or the multi-hop. By using the concept expressed in the $MinConn$ algorithm [12], $assignWeightSH$ assigns this weight ν_{ij} to each LP LSP i-th crossing the LE j-th (note that there could be multiple parallel LE connecting s to d, because of the multi-layer characteristic of a WDM network):

$$\nu_{ij} = \frac{b_{ij} - \rho_j}{\rho_j}$$

where b_{ij} is the bandwidth used by the LSP i-th and ρ_j is the difference between b and the residual bandwidth on LE j-th. The idea is to preempt LSP according to this order: if some LSP with positive weight exists, it should be rerouted first a single LSP with the lower bandwidth b_{ij} (but such that $b_{ij} + \rho_j = b$ if $\rho_j \neq 0$), then an LSP with bandwidth (equal or) bigger than b. If only LSPs with negative weights exist, two or more of them (in order to free at least bandwidth b, or $b - \rho_j$ if $\rho_j \neq 0$) should be rerouted.

The second function $assignWeightMH$ considers only LP LSPs crossing two or more LEs. The constraints for HP traffic impose to skim from $\mathcal{E}$ all the LEs j whose cost $h_j > H_{max}$. Furthermore, weights must be assigned in order to preempt first LSPs crossing the lower number of o-e-o conversions. In the multi-hop case is more difficult to find potential LSPs to preempt compared to the single-hop case: first of all, for each LSP i-th originated or crossing node s and with final or intermediate destination d, we must find the crossed LE j-th with the minimum residual bandwidth σ_j. Then, indicated with b_{ij} the bandwidth used by this LSP, the LSP itself is a preemptable connection only if $b_{ij} + \sigma_j \geq b$ (for $\sigma_j \neq 0$).

$$\begin{cases} \nu_i \overset{1}{=} c_i + b_i \cdot C_{max} & (b_i < b) \\ \nu_i \overset{2}{=} max(\nu^1) + c_i & (b_i = b) \\ \nu_i \overset{3}{=} max(\nu^2) + c_i + b_i \cdot C_{max} & (b_i > b) \end{cases} \quad (1)$$

Here ν_i is the weight assigned to each potential LSP i-th, b_i its bandwidth and c_i the number of experienced conversions (equal to $LE_i^{sd} - 1$), while $max(\nu^1)$ and $max(\nu^2)$ is the maximum value of ν_i when $(b_i < b)$ and $(b_i = b)$ respectively. With such weights, the LSP with the smaller bandwidth (but such that $b_i + \sigma = b$) and the minimum number of o-e-o conversions is given the smallest weight, then it is the first one selected to preempt.

In both cases if some LSP (or two or more LSPs with the same path) to reroute is found, then the incoming HP request must be routed over its path, while the LSP (or LSPs) must be rerouted by using some dynamic grooming algorithm (or blocked if necessary). We consider that proper MPLS mechanisms guarantee a hitless rerouting of an LSP inside the network.

Compared to GPA, LPA allows to minimize the number of LSPs to reroute, which are more than one only in the case we need to preempt some low-bandwidth LSP, with great benefit for the network disruption. Furthermore, by limiting the execution of the preemption algorithm to the ingress router s only, instead of having an algorithm executed in many LSRs along the request path as in GPA, very few RSVP messages would flow through the network to manage the preemption. In particular the signaling involves the ingress router and some edge LSRs only if the low-priority LSPs to reroute are not originated in s.

IV. Simulation results

In order to evaluate the performance of the proposed Traffic Engineering scheme, an extensive set of experiments have been executed. Simulation have been performed on different network topologies, both with low and high number of nodes: thanks to the consistency of the results, only the graphs relating to the well-known topology of [6] are shown.

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

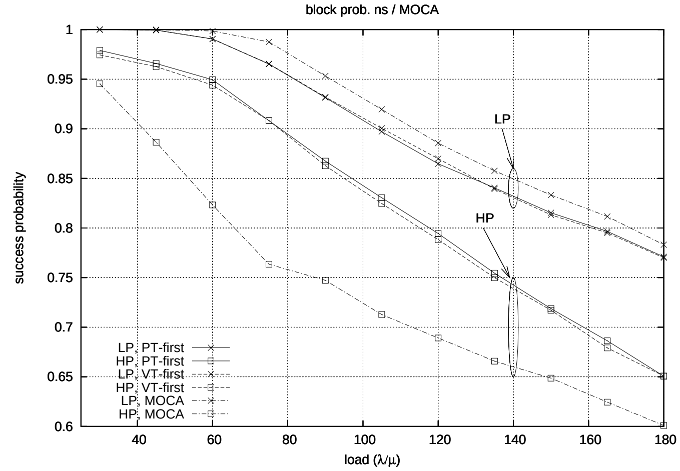

Fig. 3. Success probability for MOCA vs. `PT-first` and `VT-first`

priority. 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 $1/\mu$. Ingress and egress router pairs for each LSP set-up request are chosen randomly. The network is loaded with 50000 requests during one trial, and the performance are 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. These percentages have been chosen in order to have appropriate results from the point of view of the blocking probability when a dynamic grooming algorithm is applied. Regarding the physical constraints used in the experiments, $H_{max} = 4$ and $C_{max} = 1$ have been chosen for the High-Priority Class. 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. The number of wavelengths considered in all the tests is $K = 4$.

The first set of tests compare the "QoS-aware" grooming algorithms `VT-first`, `PT-first` and the Minimum Open Capacity Algorithm (MOCA) proposed in [6]. Traffic requests are limited only to some specific ingress and egress router pairs because MOCA can work only when this strong assumption is considered. In this first test, the same pairs as in [6] have been considered for comparison. However it is important to highlight that our scheme allows to relax this constraint.

Figure 3 shows the success probability of MOCA, `VT-first` and `PT-first` for each CT. MOCA performs the 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 due to the fact that MOCA is a dynamic grooming algorithm whose main objective is to perform load balancing distribution of the traffic: because the average number of physical hops crossed by each request is very high, it performs very badly when HP traffic must be routed over the network.

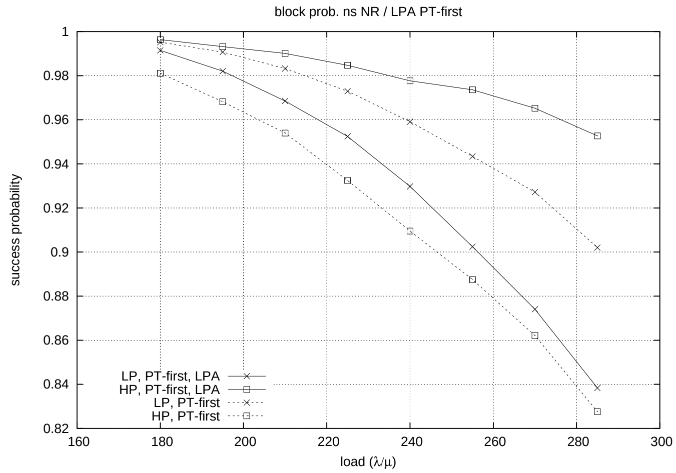

Fig. 4. Success probability with and without LPA: `PT-first`

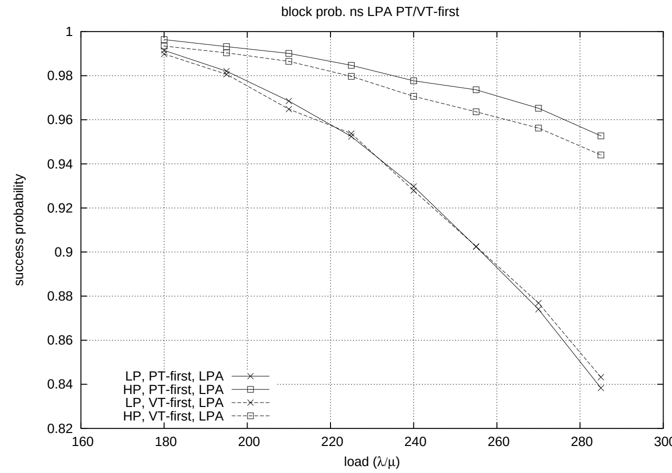

Fig. 5. Success probability for `PT-first` and `VT-first` with LPA

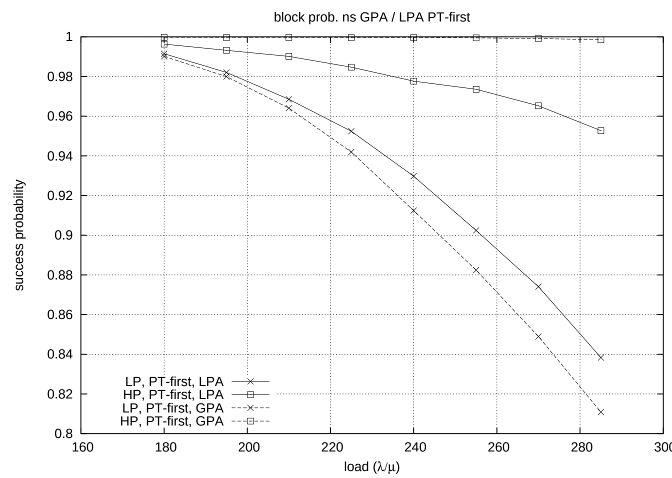

Fig. 6. Success probability with LPA and GPA: `PT-first`

Actually, the better performance for LP traffic can be justified also by the fact that more room is left free due to the higher probability of blocking for HP traffic. Furthermore this Figure shows that the resulting blocking probability is very high even for very low network load due to the restrictions on the ingress and egress router pairs.

Another important aspect to consider is the complexity of a dynamic grooming algorithms. In fact each time a new LSP needs to be routed over the network with MOCA, the maximum flow between each ingress-egress pair nodes has to be computed in $O(n^2\sqrt{m})$ [6] before applying the traditional $O(nm)$ SPF algorithm. Our grooming algorithms instead route each connection request by using the SPF algorithm only.

As expected, for all the proposed dynamic grooming algorithms a higher blocking probability is experienced by high-priority traffic. In the following, we analyze the impact of the preemption mechanism proposed in Section III-B. In the rest of this Section, we relax 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.

Figure 4 shows the success probability for the `PT-first` dynamic grooming algorithm when a “local” preemption mechanism (LPA) is applied. `VT-first` performs very similarly, thus results are not included: in both cases the obtained gain is quite high, and in particular it can be noticed that by using LPA, the success probability for HP traffic is increased by about 14%, while it decreases dramatically for LP traffic.

Figure 5 shows the performance of LPA when `PT-first` and `VT-first` is applied. Compared to Figure 3, the success probability is increased dramatically for HP traffic to the detriment of LP traffic. It can be noticed that when the `PT-first` grooming algorithm is applied, the success probability for the HP traffic class is always higher than the one obtained with `VT-first`. This can be explained by the implicit mechanism used in `PT-first`, which always tries to set-up a new lightpath (a direct one) when a new request is arrived, thus guaranteeing a a higher connectivity in the VT and then more available routes for HP traffic.

Figure 6 shows a comparison of the success probability when the proposed LPA and the optimal preemption algorithm (GPA) are applied: only the results when the `PT-first` grooming scheme is applied are indicated, because `VT-first` performs very similarly. The GPA mechanism considered in these simulations is implemented by using the mechanism proposed in the $MinConn$ algorithm [12]. LPA performs quite well compared to the optimal algorithm, which always find a route for HP requests. In fact the success probability is quite high (more than 90%) even at high network loads, when the LP traffic experience a higher blocking probability.

When considering the impact of the proposed TE scheme in term of network disruption there are two main parameters to consider: the percentage of rerouted or blocked LP LSPs and the number of lightpaths which are set-up when LP LSPs must be rerouted in the network to leave room for incoming HP connection requests.

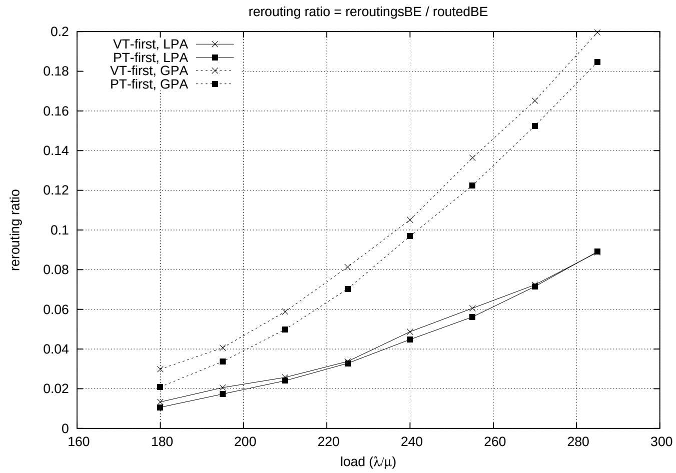

Fig. 7. Rerouting ratio for `PT-first` and `VT-first`

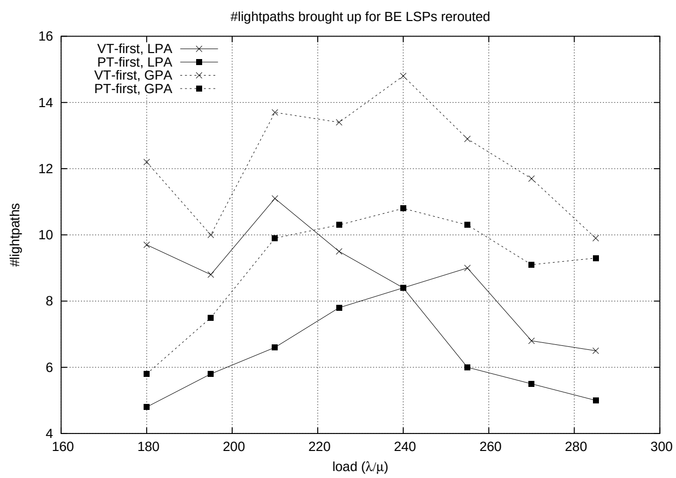

Fig. 8. Number of set-up lightpaths due to reroutings: `PT-first` and `VT-first`

Figure 7 shows the percentage of rerouted LSPs, calculated as number of rerouted LSPs over the total number of LP LSPs routed with success. As expected, when using the proposed LPA this ratio is much lower (roughly the half!) than in the optimal case. The same behavior has been verified when the ratio of blocked LP LSPs is considered. In both cases, the lowest ratio of preempted LSPs is obtained by using the `PT-first` grooming policy instead of `VT-first`. When instead the absolute number of set-up lightpaths when LP LSPs are rerouted is considered (Figure 8), GPA sets up between 40% and 60% more lightpaths on average compared to LPA, while the lower number of new lightpaths is reached when `PT-first` is used. In fact, by using `VT-first` the Virtual Topology would be highly loaded on average, thus forcing the set-up of new-lightpaths when a LP LSP needs to be rerouted.

V. Conclusions

A novel Traffic Engineering scheme for G-MPLS networks to efficiently route sub-wavelength requests with different QoS requirements has been proposed in this paper. Compared to previously proposed TE schemes, the objectives are to minimize the rejection probability for high-priority traffic by respecting specific constraints on the maximum tolerable end-to-end delay and packet-loss ratio at the same time. The proposed scheme consists of an on-line dynamic grooming scheme which routes an incoming request by respecting the specific QoS requirements, while a preemption algorithm guarantees that high-priority requests experience a reduced blocking probability compared to low-priority ones.

Simulations performed on different topologies shows that the proposed “local” preemption mechanism minimizes the network disruption both in term of number of preempted LP connections and new set-up lightpaths. The best results, even in term of success probability, are obtained when the `PT-first` grooming policy is applied, thus when the set-up of new lightpaths is preferred to the routing of incoming request over the existing Virtual Topology. An important aspect of this mechanism is the reduced signaling complexity: in fact, by reducing the number of LSPs to preempt, very few edge LSRs are involved in the signaling, thus very few RSVP messages flow through the network to manage the preemption.

Future improvements of the proposed TE scheme should consider a more accurate representation of the network which considers the number of ports per LSR in the network or even the specific physical properties of the optical devices (fibers, OXCs, amplifiers,...) in order to better evaluate the real impact of specific lightpaths to guarantee the transmission quality of high-priority connections.

Furthermore, by considering the two highlighted constraints, delay and packet-loss, it is possible to consider more intermediate (e.g. normal-priority, NP) classes of traffic: it could be interesting to study the impact of different preemption policy when more than two Class Types are considered in the network. For example, for NP connections with constraints on the maximum end-to-end delay only, it could be interesting to consider the case of preempt them when a HP request is blocked, but limiting the number of possible reroutings such a connection can experience during its lifetime.

Acknowledgments

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

References

[1] 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.

[2] K. Kompella and Y. Rekhter. OSPF Extensions in Support of Generalized MPLS. IETF Draft, *draft-ietf-ccamp-ospf-gmpls-extensions-09.txt*, work in progress, December 2002.

[3] L. Berger. GMPLS Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions. IETF RFC 3473, January 2003.

[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] M. Brunato and R. Battiti. A multistart randomized greedy algorithm for traffic grooming on mesh logical topologies. In *Proceedings of the 6th IFIP ONDM*, Torino, Italy, February 2002.

[6] 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.

[7] X. Niu, W.D. Zhong, G. Shen, and T. H. Cheng. Connection Establishment of Label Switched Paths in IP/MPLS over Optical Networks. *Photonic Network Communications*, 6:33–41, July 2003.

[8] 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.

[9] J.Y. Youe and S.W. Seo. An algorithm for virtual topology design in WDM optical networks under physical constraints. In *Proceedings of ICC*, volume 3, pages 1719–1723, New Orleans, LA, June 1999.

[10] 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.

[11] 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.

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

[13] J.C. de Oliveira, C. Scoglio, I.F. Akyildiz, and G. Ulh. A new preemption policy for diffserv-aware traffic engineering to minimize rerouting. In *Proceedings of INFOCOM*, pages 695 –704, 2002.

[14] B. Szviatovszki, A. Szentesi, and A. Jüttner. Minimizing re-routing in MPLS networks with preemption-aware constraint-based routing. *Journal of Computer Communications*, 25:1076–1083, 2002.

[15] A. Bosco, A. Botta, M. Intermite, P. Iovanna, and S. Salsano. Distributed Implementation of a Pre-Emption Mechanism for a Network Control Based on IP/MPLS Paradigm. In *Proceedings of the 7th IFIP ONDM*, Budapest, Hungary, February 2003.

[16] F. Le Faucheur. Protocol extensions for support of Diff-Serv-aware MPLS Traffic Engineering. IETF Draft, *ddraft-ietf-tewg-diff-te-proto-03.txt*, work in progress, February 2003.

[17] L. Gouveia, P. Patrício, A. de Sousa, and R. Valadas. MPLS over WDM Network Design with Packet Level QoS Constraints based on ILP Models. In *Proceedings of INFOCOM*, 2003.

[18] A.K. Parekh and R.G. 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.