Elastic Traffic Effects on WDM Dynamic Grooming Algorithms

R. Lo Cigno[1], E. Salvadori[1], Z. Zsóka[2],
[1]Dipartimento di Informatica e Telecomunicazioni – Università di Trento
[2]Department of Telecommunications – Budapest University of Technology and Economics

Abstract— **Traffic grooming in IP over WDM networks introduces a coupling between the optical and the IP layer. Grooming algorithms are normally studied with a very simple traffic model that completely ignores this interaction. This paper compares the performance of two simple grooming algorithms with a traditional, Poisson based traffic model and a more complex one that takes into account the IP traffic elasticity and the inherent interaction between IP and the optical layer. Simulation results, supported by heuristic considerations highlighting the interaction effects, show that ignoring the two layer interaction may lead to wrong conclusions and waste of resources.**

I. INTRODUCTION

IP over WDM is one of the racehorses that pulls the train of large bandwidth networking. New services are continuously deployed over IP, and WDM evolution [1], [2] provides the transmission speed needed to pump the information through the network. One of the main issues in IP over WDM architectures, is the traffic aggregation or *grooming*. The traffic is generated as tiny trickles over IP, while the transmission pipe over a single λ within an optical fiber is enormous.

Many grooming algorithms were proposed in recent years (see [3], [4], [5] to cite just a few), and compared one another. Some works assume static grooming [5], [6], and generally tackle the problem with some optimization technique, while others assume that grooming is dynamic [4], [7], [8], [9]. All these works, however, disregard the elastic nature of TCP/IP traffic: IP over WDM is indeed modeled like a traditional circuit switched multiplexing problem!

As shown in [10], considering the adaptivity of traffic has a deep impact on the network performance and on routing algorithms in particular. The reason lies in the feedback nature of the interaction of elastic traffic with the network: the network status (e.g., congestion) induces a reaction (feedback) in the source behavior. If the feedback is positive it has, a noxious impact on performance, since congestion, is exasperated by the positive feedback.

As usual in closed loop systems with delay, the nature of feedback (positive or negative) can change with changing conditions, so that, for instance, a negative feedback at low loads can change to a positive feedback at high loads, leading to instability phenomena.

The aim of this paper is to investigate how traffic elasticity impacts on grooming. The problem in itself is rather complex, since it requires to take into account how competing groomed flows interact one another, e.g., sharing resources following a max-min criterion, as well as how the optical management plan behaves and assigns resources to traffic relations.

Elasticity in groomed traffic can arise due to a number of reasons and in very different scenarios. In emerging metro-area optical network, the foreseen trend is a very dynamic and aggressive use of optical paths, thus leading to traffic relations that are very close to a simple host-to-host IP flow. In more traditional wide-area optical networks, where traffic relations are peering contracts with highly aggregated flows, the elasticity still arises from the fact that all the flows within a traffic relation are elastic: if congestion arises, then all the flows will react reducing their offered load.

II. GROOMING ALGORITHMS

The network architecture considered in this work is IP over WDM with dynamic optical routing, i.e., optical paths are opened on demand. We assume an overlay model [11].

The optical level is based on Optical Crossconnects (OXC) interconnected by fiber links. Routing is shortest path with First-Fit wavelength assignment for the establishment of lightpaths. OXCs do not have wavelength conversion capabilities. The search for a lightpath is greedy, but it is terminated when a predefined maximum number of crossed links N_l is reached. This threshold is very important both to limit the complexity of routing and wavelength assignment and the waste of optical resources on very long lightpaths.

The IP level assumes traditional routers with shortest path routing based on the number of hops. An optical path is seen as a single hop regardless of the number of OXCs it crosses.

There are two node architectures: i) pure OXC, which allows to switch entire lightpaths from an ingress port to an egress port; ii) Grooming OXC (G-OXC), which supports sub-wavelength traffic flows and groom them onto wavelength channels. A G-OXC is also an IP router. Low-speed traffic can then be transmitted or received only in G-OXCs.

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 build by the optical level underneath. Using this information and the grooming strategy defined below G-OXCs decide whether a new traffic relation must be routed at the IP level or a new lightpath should be opened.

This work was supported by the Italian Ministry of Education and Research (MIUR) through the GRID.IT and ADONIS projects, and by the Hungarian Italian Intergovernmental S&T Cooperation Programme '04–'07 (I-17/03).

0-7803-8794-5/04/$20.00 © 2004 IEEE

The decision to route the incoming requests over the existing virtual topology or to establish new lightpaths to create more room for them can lead to different network performances. A general analysis of different "grooming policies" is carried out in [4] under the hypothesis of bandwidth-guaranteed (circuit-based) traffic. When elastic traffic is considered, there is no obvious upper limit to the possible number of flows which is routed onto the existing logical layer. In this case, the need for the establishment of new lightpaths must be introduced based on some suitable parameter. We introduce a parameter, called *optical opening threshold* th_o, as a threshold on the instantaneous throughput obtained by connections, defined as a fraction of the peak rate B_M required by each flow.

In this work we consider the following two grooming policies.

Virtual-topology First (***VirtFirst***) **—** Each time a new IP request arrives in some router, the current virtual topology is considered first to route the request. If, once routed, the amount of bandwidth for some flow (not necessarily the one being routed) is less than th_o, a new lightpath is set-up between source and destination (if possible). If the setup is successful, the IP request is routed over it (it is a one hop route at the IP level) and a new virtual topology is computed at the IP level. The new topology does not affect already routed requests (i.e., no re-routing is considered), but will be used for routing all new requests. If the new lightpath cannot be set up, then the request is routed based on the current virtual topology. Whenever a closing flow leaves a lightpath empty, the lightpath is closed too (after a suitable timeout) and the virtual topology is re-computed.

Optical-level First (***OptFirst***) **—** Each time a new IP request arrives in some router, the G-OXC always attempts first to set up a new lightpath in the optical layer, in order to route the request over it. If no free wavelengths are available, the incoming request is routed on the current virtual topology. As in *VirtFirst* if a closing flow leaves a lightpath empty, the lightpath is closed.

These two opposite policies have been often considered by different authors to perform comparisons with new proposals or to study the impact of some specific network constraints, such as OXC node's architecture. In this paper we considered them to study the impact of elastic traffic and analyze whether it affects them differently.

III. THE SIMULATION TOOL

The simulator we developed for this study, named GANCLES is described extensively in [12] and a web page [13] is maintained where the software is available. GANCLES is an extension of the connection level simulator ANCLES [14].

Several improvements were made to ANCLES over the years, some regarding the introduction of best-effort, elastic traffic as described in [10] and some related to the introduction of optical routing capabilities [15].

The key point required to jointly study the IP and the optical level is the capability of handling both a physical topology (a directed graph of links and OXCs) and a virtual topology (a directed graph of virtual links and the routers embedded in G-OXCs). The links of the virtual topology match lightpaths provided dynamically by the lower level. Dynamic grooming solutions presented in Sect. II (and others being added) correlate the optical and IP level during simulations.

Several routing and management schemes are available at both levels. We use only very simple routing algorithms both at the optical and at the IP level to highlight clearly the interaction of grooming algorithms and elastic traffic.

When elastic traffic is considered, no admission control is enforced, and no backpressure on traffic sources is available, the network can become instable, as the number of flows within the networks grows to infinity and their individual throughput goes to zero. To avoid this risk, and to build a more realistic scenario, we introduce a starvation threshold th_s expressed as a fraction of the peak bandwidth B_M required by the flow[1]. If at some time instant one or more flows receive a throughput smaller than th_s (due to the arrival of a new flow), the elastic flow with the highest backlog is immediately closed. An important performance meter is the rate of flows interrupted this way. We call this meter *starvation probability*. Notice that if admission control is enforced, this simply means refusing the arriving flow instead of closing a flow as just described. The two actions are however not equivalent, because: i) the arriving flow may not be a starved one (e.g., has a smaller required B_M); ii) blocking is not influenced by the flow dimension, while the starvation is higher for larger flows (in bytes); iii) starved flows waste network resources and may influence overall throughput, which is computed only on completed flows.

A. Traffic Models

We introduce two different models of elastic traffic. Both share the characteristic that a flow i arrives to the network with a backlog of data D_i to transmit and both include some form of elasticity, though very different one another.

The first model, that we name *time-based* (TB), assumes that the elasticity is taken into account only reducing the transfer rate when congestion arises. The flow duration τ_i is determined when the flow arrives to the network, based on its backlog D_i and its "requested bandwidth" B_{Mi} (e.g., the peak negotiated rate, or the access link speed) $\tau_i = (D_i)/(B_{Mi})$. During the connection lifetime bandwidth is then shared according to the max-min criterion. Congestion reduces the throughput, but the closing time is not affected. A consequence is that the data actually transferred by a flow i is generally less than the "requested" amount D_i. This model is very simple and does not grab all the complexity of the closed-loop interaction between the sources and the network. It simply models the fact that the more congested is the network, the smaller is the throughput the flows get.

The second model, that we name *data-based* (DB), assumes instead an ideal max-min sharing of the resources within the network at any given instant. Flows still arrive to the network with a backlog D_i, but the acceptance of a new flow will affect not only all the other flows on the same path, but indeed

---

[1] Notice that th_s is structurally identical to th_o introduced in Sect.II, but its meaning is very different and its numerical value can be different too.

all the flows in the network, since the max-min fair share is completely recomputed updating the estimated closing time of all the flows in the network. The same applies when flows close, freeing network resources. This model includes the most important feature of elastic traffic, which is the feedback on the flows duration. The more congested is the network, the longer flows remain in the network. Congestion spreads over time enhancing the possibility that still further flows arrive in the network worsening congestion.

The DB model is more accurate, mimicking the behavior of an ideal congestion control scheme; however, its complexity and computational burden are larger, specially for high loads. Investigating whether (or under which conditions) the simpler TB model is accurate enough in the context of IP over WDM with dynamic grooming, or if it leads to gross approximations can be very important.

IV. NUMERICAL EXAMPLES

As we discuss in Sect. IV-B, the phenomena involved in routing/grooming elastic traffic are complex, and often far from intuitive. As performance parameters we consider the following five: three at the IP level and two at the optical level.

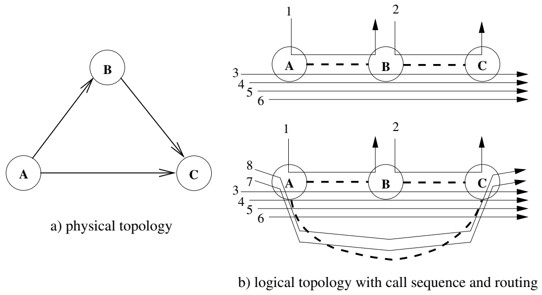
- T: The average throughput per flow $T = \frac{1}{N_c}\sum_{i=1}^{N_c} T_i$ where N_c is the number of observed flows. Notice that in a resource sharing environment this is not the average resource occupation divided by the number of flows, since flows have all the same weight, regardless of their dimension.
- p_s: The starvation probability as defined before.
- R_o: The ratio between the opening rate of optical paths and the arrival rate of flows at the IP level. It is a measure of the optical level routing effort. For an optical routed network without grooming $R_o = 1$, while for a purely IP routed network $R_o = 0$.
- N_{lo}: Average number of links per optical path.

The goal of a grooming algorithm is maximizing T while minimizing p_s, R_o and N_{lo}.

Before discussing results on a mesh topology, we highlight some peculiar behavior of the *VirtFirst* grooming in a very simple scenario, that will help in interpreting results in more complex scenarios.

A. A Trivial Example

Consider the simple 3-node topology of Fig. 1 a), where only a single wavelength per link is present and the active traffic relations are only A–B, B–C, and A–C. Assume that *VirtFirst* grooming is used and, starting with the network empty, the following sequence of flows arrives: AB, BC, AC, AC, AC, AC, ... (AB identifies a flow originating in A with destination B and so on). We set $th_o = 0.2$ and $th_s = 0$ and all flows are able to fully exploit the optical path capacity. The average throughput obtained by flows is represented by the solid line with cross marks in Fig. 2 (this curve refers to the top x-axis (numer of flows) and left y-axis (normalized throughput)), as can be easily seen following the logical topology evolution reported in Fig. 1 b).



a) physical topology

b) logical topology with call sequence and routing

Fig. 1. Simple 3-node topology used for the theoretic verification of results

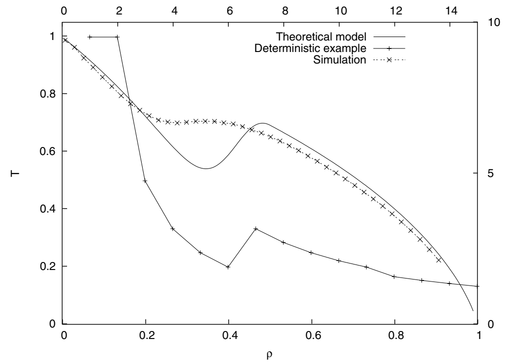


Fig. 2. Average throughput computed deterministically, via simulation and with a simple stochastic model for the scenario depicted in Fig. 1 a)

This example show that with *VirtFirst* grooming, it is possible that T increases while the load increases due to the interaction between the IP and optical layer. However, a deterministic example is not enough to draw conclusions. In order to investigate further in the behavior, we have set up a simple (and approximate) queuing model of the same scenario based on processor sharing queues that mimic the max-min resource division. Details about the model can be found in [16].

Fig. 2 reports, beside the simple deterministic example, results obtained with the simple stochastic model and with simulations (DB traffic model) for $th_o = 0.2$. These two curves are plotted versus the bottom (ρ) and right (absolute throughput) axes. The simulation curve does not show the same increase in throughput around the load $\rho = 0.5$ displayed by the model (however, we have observed it for smaller thresholds th_o). The reason is that the dynamic routing of flows makes the transition from routing the AC traffic mainly through A–B and B–C to routing it mainly over A–C smoother than in the approximate model. In this case we set $th_s = 0$, so that $p_s = 0$. Given the simple scenario $N_{lo} = 1$, while R_o and N_h are not of much interest.

This simple example give some insight on the complex behavior of grooming associated with elastic traffic, which, to the best of our knowledge was never observed in other works, that, using constant-bit-rate like traffic models, cannot observe throughput performance. In the following we study a more realistic scenario, to gain more insight on grooming and

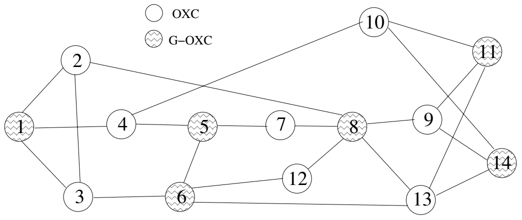


Fig. 3. NSFNET network

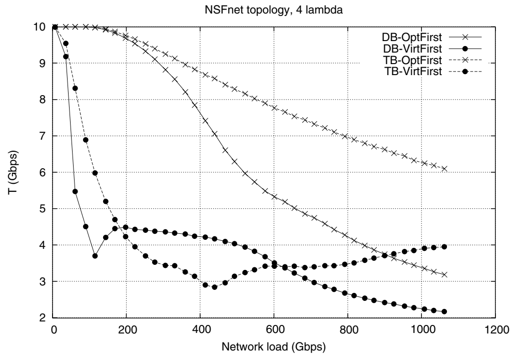


Fig. 4. Average throughput T for the DB and TB traffic models for the two grooming policies

elastic traffic interaction.

B. Results for a Mesh Topology

We present results obtained on the NSFNET network shown in Fig. 3, which has 14 nodes and 21 fiber links. Each fiber carry up to 4 wavelengths, and only 6 nodes out of 14 are G-OXCs. Each wavelength has a capacity of 20 Gbit/s. A traffic source is connected to each G-OXC, opening flows with $B_M = 10$ Gbit/s; each flow transfer data whose size is randomly chosen from an exponential distribution with average 12.5 Gbytes. A uniform traffic pattern is simulated, i.e., when a new traffic relation is generated, the source and destination are randomly chosen with the same probability; $th_s = 0.1$ in all simulations and $th_o = th_s$ for the sake of simplicity. All simulations are run until performance indices reach a 95% confidence level over a ±5% confidence interval around the point estimate. We have run simulations on other topologies obtaining similar results, not reported here for lack of space [16].

Fig. 4 presents a comparison of the average throughput T obtained modeling best-effort traffic relations using the TB approach (dotted lines) and the DB approach (solid lines) when the two grooming algorithms *VirtFirst* (round marks) and *OptFirst* (cross marks) are used. With the same graphic rules, Fig. 5 reports the starvation probability. The difference in performance results of the two approaches is striking.

Let's consider first the *OptFirst* grooming policy. Both approaches show T starting from 10 Gbit/s when the offered load is low; however, they immediately diverge as the offered load increases. Indeed, the DB traffic model shows much faster decrease in T as soon as the offered load increases and this is due to the spreading of congestion over time with a sort of snow-ball effect. On the contrary, the TB traffic model shows a smoother decrease of the average bandwidth.

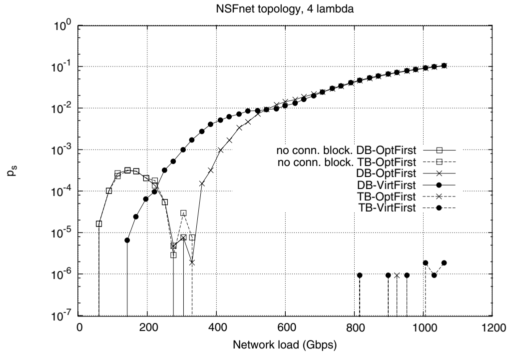


Fig. 5. Starvation probability p_s for the DB and TB traffic models for the two grooming algorithms

Analysing the starvation probability in Fig. 5 adds more insight. When the traffic is very low (below 350 Gbit/s) both traffic models show the same, very strange behavior: the starvation increases and then decreases sharply. This form of blocking is independent of the traffic model and it is due to a very aggressive and dynamic use of optical resources that sometimes leads to have no connectivity at the IP level, i.e., a flow request arrive and there is no possible path, neither optical, nor through multiple IP hops, between the source and the destination. When the load increases, however, lightpaths become more stable (because there is always traffic keeping lightpaths open) and the probability that the virtual topology is not completely connected becomes negligible. To highlight the difference of this phenomenon from the real starvation, in Fig. 5 the curves relative to it are plotted with square marks. When the load increases further, the two traffic models behavior diverges: the TB model show no starvation at all, apart from points at very high loads, which show a blocking probability around 10^{-6}, while the DB model show a starvation probability increasing steadily.

When considering the *VirtFirst* grooming policy instead, the behavior of both traffic model is different from the previous one. Both DB and TB T decrease sharply even when the offered load is low, due to the conservative policy of *VirtFirst*. In fact, *VirtFirst* sets up the minimum number of lightpaths in order to guarantee the minimum network connectivity, and keeps this configuration unchanged until some flow crosses the starvation threshold th_s. Only in this case *VirtFirst* increases the resources at IP level by setting up new lightpaths. In particular, the T for DB traffic relations decreases very rapidly, causing an earlier set-up of new lightpaths compared to TB traffic. This lead the DB traffic throughput T to "bounce" taking advantage of the higher number of lightpaths in the network, at a load much smaller than for the TB model, that starts increasing again at higher loads. Obviously, both models would show another (and definitive) decrease in T for higher loads, not shown here. Notice that the starvation rate (see Fig. 5) for the TB model is in this case always zero (apart from a single point around load 900 Gbit/s), while the starvation rate of the DB model increases steadily and shows a behavior

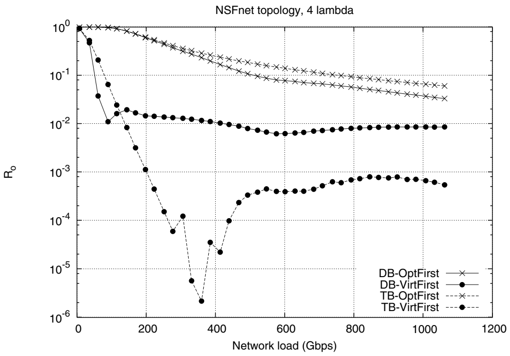


Fig. 6. Ratio R_o between the opening rate of optical paths and the arrival rate of flows at the IP level.

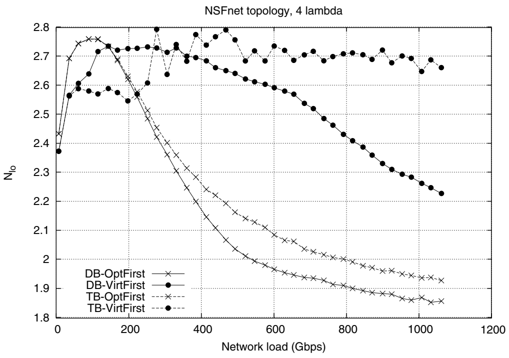


Fig. 7. Average number of links N_{lo} per optical path.

similar to the DB model in the *OptFirst* case.

Fig. 6 shows the ratio R_o. As expected, when *VirtFirst* grooming policy is used, R_o decreases quickly with the load, indicating a burden for the optical level that does not increase with the traffic (indeed, it might also decrease when the load is high). When the *OptFirst* grooming policy is adopted, R_o decreases slowly and smoothly, indicating a much higher burden for the optical level.

Fig. 7, finally, plots N_{lo}. Once again the behavior of the *OptFirst* policy is more predictable, with the number of links that decreases steadily with the load, and roughly converges to the weighted average distance in number of links between G-OXCs. The *VirtFirst* policy shows instead very long optical paths. This effect is due to the intrinsic behavior of this grooming policy: most of the lightpaths set up by *VirtFirst* are in fact never torn-down since they are carrying traffic almost all the time. Then, when new lightpaths must be established, it is more likely that they would be set up in the optical network through longer routes.

V. DISCUSSION & CONCLUSION

This paper introduced the analysis of dynamic grooming algorithms in IP over WDM with elastic traffic. The elasticity of traffic interacts with the grooming algorithms as well as with the routing both at the IP and optical level, leading to unexpected results.

Two basic grooming policies were considered, one privileging the opening of new optical paths (*OptFirst*), the other privileging the use of the already available IP logical topology (*VirtFirst*).

In both grooming algorithms the impact of elastic traffic, included with a sophisticated model in the simulations tool, is dramatic, showing clearly that approximating IP traffic with CBR-like traffic can lead to wrong conclusions when routing and grooming are considered.

The focus of the paper was on the impact of the traffic elasticity, thus little attention was placed on the "suitability" of the grooming algorithms analyzed. Both the *OptFirst* and *VirtFirst* algorithms, however, have clearly shown that they are not suited for the management of an IP over WDM network, since the lack of coordination between the IP and the optical level leads to waste resources. As shown on the NSFNET topology the *OptFirst* policy may even lead to block requests with very low network loads because a very aggressive use of optical resources may lead to IP-level virtual topologies that are not completely connected.

This observation open new and interesting questions on the heuristics that dynamic grooming algorithms in IP over WDM networks should pursue in order to optimize the use of resources and, at the same time, maximize the satisfaction of the end users.

REFERENCES

[1] N. Ghani, S. Dixit, T.S. Wang, "On IP-WDM Integration: A Retrospective," *IEEE Communications Magazine*, 41(9):42–45, Sept. 2003.

[2] P. Molinero-Fernandez, N. McKeown, H. Zhang, "Is IP going to take over the world (of communications)?" *ACM SIGCOMM Comp. Comm. Review*, 33(1):113–119, Jan. 2003.

[3] X. Zhang, C. Qiao, "An Effective and Comprehensive Approach to Traffic Grooming and Wavelength Assignment in SONET/WDM Rings," *IEEE/ACM Trans. on Networking*, 8(5):608–617, Oct. 2000.

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

[5] K. Zhu, B. Mukherjee, "Traffic grooming in an optical WDM mesh network," *IEEE JSAC*, 20(1):122–133, Jan. 2003.

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

[7] M. Kodialam, T.V. Lakshman, "Integrated Dynamic IP and Wavelength Routing in IP over WDM Networks," *In Proc. of INFOCOM 2001*, pp. 358–366, Anchorage, AK, USA, Apr. 22–26 2001.

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

[9] R. Srinivasan, A.K. Somani, "Dynamic Routing in WDM Grooming Networks," *Photonic Network Communications*, 5:123–135, Mar. 2003.

[10] C. Casetti, R. Lo Cigno, M. Mellia, M. Munafò, Z. Zsóka "A Realistic Model to Evaluate Routing Algorithms in the Internet," *In Proc. IEEE Globecom 2001*, S. Antonio, TX, USA, pp. 1882–85, Nov. 25–29, 2001.

[11] B. Rajagopalan, J. Luciani, D. Awduche, "IP over Optical Networks: A Framework," RFC 3717, IETF, Mar. 2004.

[12] E. Salvadori, Z. Zsóka, D. Severina, R. Lo Cigno, "GANCLES: A Network Level Simulator to Study Optical Routing, Wavelength Assignment and Grooming Algorithms," DIT-04-017, 2004. `www.dit.unitn.it/locigno/preprints/SZL4-17.pdf`

[13] GANCLES - Grooming cApable Network Call-Level Simulator. `ardent.unitn.it/gancles`

[14] ANCLES - A Network Call-Level Simulator. `www.tlc-networks.polito.it/ancles`

[15] ASONNCLES - ASOn Network Call-Level Simulator. `www.hit.bme.hu/∼zsoka/asoncles`

[16] E. Salvadori, R. Lo Cigno, Z. Zsóka, "Analysis of Elastic Traffic Effects on WDM Dynamic Grooming Algorithms (extended version)," DIT-04-091, 2004. `www.dit.unitn.it/locigno/preprints/SLZ4-91.pdf`