Multi-Layer Traffic Engineering through Adaptive λ-path Fragmentation and De-Fragmentation

Tibor Cinkler, Péter Hegyi, Márk Asztalos,
Géza Geleji, János Szigeti, András Kern

{cinkler,hegyi,asztalos,geleji,szigeti,kern}@tmit.bme.hu
HSN*Lab*, Department of Telecommunications and Media Informatics
Budapest University of Technology and Economics, Hungary
Magyar tudósok körútja 2, H-1117 Budapest

Abstract. In Multi-Layer networks, where more than one layer is dynamic, i.e., connections are set up using not only the upper, e.g., IP layer but the underlying wavelength layer as well leads often to suboptimal performance due to long wavelength paths, that do not allow routing the traffic along the shortest path. The role of MLTE (Multi-Layer Traffic Engineering) is to cut these long wavelength paths into parts (fragments) that allow better routing at the upper layer (fragmentation), or to concatenate two or more fragments into longer paths (defragmentation) when the network load is low and therefore less hops are preferred.
In this paper we present a new model (GG: Grooming Graph) and an algorithm for this model that supports Fragmentation and De-Fragmentation of wavelength paths making the network always instantly adapt to changing traffic conditions. We introduce the notion of *shadow capacities* to model "lightpath tailoring". We implicitly assume that the wavelength paths carry such, e.g., IP traffic that can be interrupted for a few microseconds and that even allows minor packet reordering.
To show the superior performance of our approach in various network and traffic conditions we have carried out an intensive simulation study.

Keywords. Adaptive Multi-Layer Traffic Engineering, Grooming Graph, Wavelength Path Fragmentation and De-Fragmentation

1 Introduction

The evolution of transport networks shows two main directions. First, there are multiple networking technologies layered one over the other. Second, it is required that not only the upper-most layer is dynamic, i.e., switched, but the upper two, or maybe all the layers.

If the layers of this vertical structure are run by different operators or providers then they must communicate to each other to exchange information necessary

This work has been done as a part of the European FP6 IP NOBEL (www.ist-nobel.org) and NoE e-Photon/ONe (www.e-photon-one.org) research projects.

for routing and other purposes. This vertical communication is referred to as Interconnection, and there are three defined Interconnection Models: (1) Overlay, (2) Augmented and (3) Peer model [1].

If all these layers are run by a single operator or provider then there is no need for communication interfaces between the layers. Therefore, a single unified integrated CP can be used for all the layers and then we have instead of the interconnection the so called *Integrated Model*. The forwarding units of all the layers of the data plane are connected to a single control plane unit.

Similarly, if such a Multi-Layer network has layers or some parts of certain layers built of interconnected elements of a unique networking technology then the set of these elements is referred to as a *Region*. Having multiple different regions within a network is referred to as a *Multi-Region* network [1] [2].

In switched multilayer transport networks (e.g. ASTN/GMPLS) the traffic demands have typically bandwidth by orders of magnitude lower than the capacity of λ-links. Therefore, it is not worth assigning exclusive end-to-end λ-paths to these demands, i.e., *sub-λ granularity* is required. Furthermore, the number of λs per fibre is limited and costy. To increase the throughput of a network with limited number of λs per fibre *traffic grooming* capability is required in certain nodes. There are many papers dealing with routing, traffic engineering and resilience in such multilayer networks, where grooming is one of the key issues [2] [3] [4] [5].

Here we consider the case of Wavelength Routing Dense Wavelength Division Multiplexing (WR-DWDM) Networks and one layer built over it. In the WR-DWDM layer a wavelength path (λ-path) connects two physically adjacent or distant nodes. These two physical nodes will seem adjacent for the upper layer built over it.

This upper layer is an "electronic" one, i.e., it can perform multiplexing different traffic streams into a single λ-path via simultaneous time and space switching. Similarly it can demultiplex different traffic streams of a single λ-path. Furthermore, it can perform re-multiplexing as well: Some of the demands de-multiplexed can be again multiplexed into some λ-paths and handled together along it. This is referred to as *traffic grooming* [6] [3]. Further on we will refer to it as *grooming*. This electronic layer is required for multiplexing packets coming from different ports (asynchronous time division multiplexing). It can be a classical or "next generation" SDH/SONET, MPLS, ATM, GbE, 10 GbE or it can be based on any other technology. However, in all cases the network carries mostly IP traffic. The only requirement is that it must be unique for all traffic streams that have to be de-multiplexed, and then multiplexed again, since we cannot multiplex e.g. ATM cells with Ethernet frames directly.

More generally, we can consider this two-layer approach as two layers of a 4-5 layer GMPLS/ASTN architecture [7] [8] [9]. However, not only the framing and layering structure is of interest, but also the control plane proposed in the GMPLS/ASTN framework.

Many excellent papers deal with design, configuration and optimisation of WDM Networks. Some of these methods can be generalised for on-line routing in two-layer networks as well using the model we propose in this paper.

There are also numerous papers dealing with on-line routing in WR-DWDM networks (see, e.g., Chapter 3 of [10]). There are multiple papers on grooming, mostly for the static case, i.e, when a two layer network is configured (see, e.g., Chapter 4 of [10]). Some papers consider the grooming capability in dynamic (on-line) routing [11]. There are also papers dealing with multilayer survivability, e.g., [12] and Chapter 5 of [10]).

However, there are only few papers, e.g., [4] [13] [14], that take all these into account *simultaneously*, using the peer or the MRN model. To our knowledge there is no paper that proposes any method for adaptive, automated, on-line and distributed Multi-Layer Traffic Engineering. The aim of our paper is to fill this gap.

Our objective is to perform distributed on-line routing of the on-line arriving demands with estimated effective bandwidths as constant bandwidth pipes over the two network-layers optimally in distributed way without separating these layers. The upper layer is assumed to support multiplexing (e.g., asynchronous TDM), while the lower layer is the λ-path system. Separating the two layers decreases the complexity, however, it also deteriorates the routing. According to the role of TE (Traffic Engineering) to increase throughput while it maintains QoS, grooming is performed jointly with adaptive, on-line, distributed and automated path fragmentation and defragmentation.

The rest of the paper is organised as follows. In Section 2 we present the "Grooming-Graph" and the "Shadow-Capacities" and explain how our model works with available routing protocols. In Section 3 the problem of λ-path fragmentation and defragmentation is explained and in Section 4 the simulation results are shown and discussed.

2 The Grooming Graph (GG) Model

The objective was to provide a general network model for routing in two layer networks with grooming, with different types of nodes and arbitrary topologies assuming peer/MRN-model, that allows optimal routing, using the resources of both layers jointly. The aim was to allow adaptive, automated, distributed MLTE (Multi-Layer Traffic Engineering) by the model used.

Although the most widespread topology is ring or interconnected rings, the model must be able to handle any regular or mesh topology. Furthermore, it must be able to handle any type of nodes of practical interest, e.g., OADM, OXC, EOXC, etc., all with or without grooming capability and with or without λ-conversion capability. Even limited grooming, and λ-conversion limited in number or range has to be supported. For this purpose we use our grooming graph (GG) model, where the node is substituted by a sub-graph.

The simpler version of this model that does not allow fragmentation and defragmentation was first proposed in [15]. ILP formulation of the static RWA problem with grooming and protection was given in [16], using the wavelength graph, while in [17] heuristics for solving the problem were proposed. [18] explains the used simpler model the "WG" and investigates the fairness issues of dynamic grooming.

In this paper we add the adaptivity to the model by defining the shadow links and their shadow capacities to be used for adaptive MLTE.

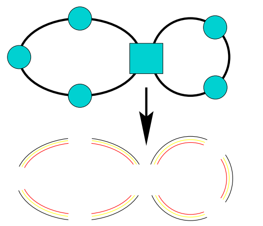

Fig. 1. Modelling edges in the GG.

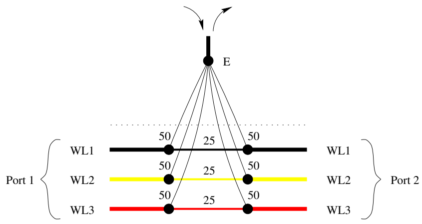

Fig. 2. Model of OADM nodes.

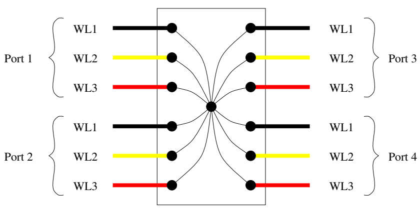

Fig. 3. Model of EOXC nodes.

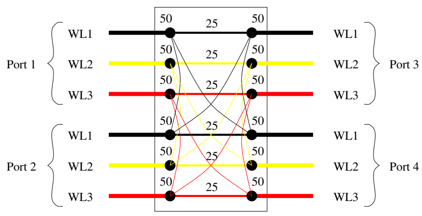

Fig. 4. Simple OXC (no λ-conv.).

2.1 Model of Links

A network consists of nodes, and links connecting the nodes. This can be modelled by a graph: a node is a vertex and a link is an edge. Having multiple λs (WL1-WL3) we will represent a λ of a link as an edge in the graph of wavelengths, according to Figure 1 for the network proposed in [19]. To prioritise filling up λs one-by-one we can assign slightly different weights to different λ-channels of one link. For example, edges representing WL1, WL2 and WL3 in Figure 1 will have weights 1,01, 1,02 and 1,03 respectively.

2.2 Model of Nodes

A node is modelled by a subgraph. The subgraph-nodes are the switch-ports, while the weighted edges represent the costs of transitions, terminations, conversions, etc. There are different types of nodes. Models of nodes differ for these. Some examples will be shown here. In similar manner a model can be derived for any additional node-type.

Optical Add-and-Drop Multiplexer (OADM): The OADM Nodes have in general two bi-directional ports (4 fibres). Their function is either to transmit a λ-path or to terminate it and usually they do not allow λ-conversion.

The weights assigned to edges representing termination (e.g., 50) are higher than weights of transition (e.g., 25), because transition is preferred to termination. According to the proposed model (Figure 2) the traffic streams can either enter or exit the OADM crossing vertex E or can be even re-multiplexed.

Cross-Connect with Electronic Core (EOXC): In the model shown in Figure 3 each pair of nodes should be connected by an edge, representing potential Cross-Connection. All edges should have equal weights. Instead of connecting all pairs using nxn edges we use n edges and one node. This simplifies the model. Each incoming channel is converted to electrical domain switched by a space-switch and again converted to the optical domain to arbitrary λ. Each termination, transition or λ-change has the same cost (e.g., 25). Therefore all edges have the same weight (e.g., 25/2).

Optical Cross-Connect (OXC): An optical Cross-Connect has more than two ports, e.g., four bi-directional ports according to Figure 1. In an OXC a light-path can make transition to any output port which supports that λ , and that λ is not yet used. This OXC type (without λ-conversion capability) will be referred to as *simple* OXC (see Figure 4). In this case one incoming channel can exit at any of the remaining output ports where that λ is supported and not yet used.

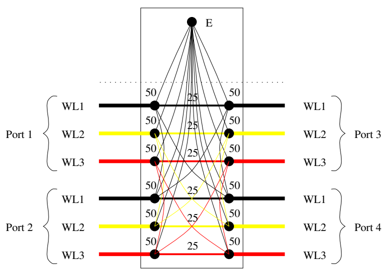

Fig. 5. OXC with λ-conversion.

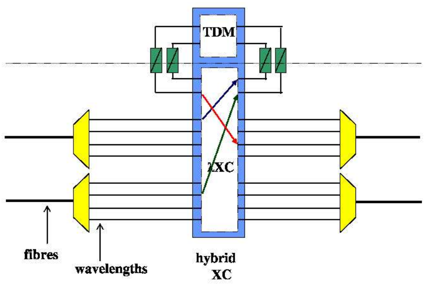

Fig. 6. The GG node model.

In some cases the traffic stream termination is also among the functions of an OXC. In that case the model does not need any change. The only difference will be that there will be some traffic offered to that OXC node. This can be modelled by offering traffic to node E and considering it as an end-node. In this case traffic-stream re-multiplexing capability is also required.

Modelling Grooming: Grooming can be modelled analogously to λ-conversion. The difference is that while in case of λ-conversion an incoming traffic stream can exit as a single outgoing stream at another λ, in case of grooming traffic streams can be multiplexed, i.e., instead of space switching space AND time switching/cross-connecting is performed.

These two functionalities can be combined as well within a single model.

Note, that λ-conversion is a special case of grooming. Therefore a node supporting only grooming can perform λ-conversion as well, while a λ-conversion node can not perform grooming.

The model we presented in Section 2 will be referred to as 'simple' grooming model. To make this model better adapt to the traffic and network conditions we extend it in Section 3.

3 Shadow Capacities for λ-path Fragmentation and De-Fragmentation

We assume either the peer interconnection model or the vertically integrated multi-region network (MRN) node model for multi-layer networks [2]. Then the network layers set the resources jointly, i.e., the control plane has knowledge of both the layers to best accommodate the arriving traffic demands.

This often leads to suboptimal performance, since the lightpaths will be routed depending on the arrival order of demands as well as on the load of the network. For instance in an empty network each arriving demand will be routed over an exclusive lightpath. This will result in a set of long lightpaths that will hinder routing the new demands, i.e., the network will become de-fragmented. After the transients the lightpaths will be configured more or less adequately. However, if the level of traffic grows short lightpaths with plenty of grooming are needed to accommodate it, i.e., lightpaths have to be fragmented into shorter parts.

To have always optimal performance the lightpath system has to adapt to the changing traffic conditions. Unfortunately, in the simple model the virtual topology offered by the wavelength system may not be changed until there is any traffic within the considered λ-paths.

3.1 Algorithm for Routing over Shadow Capacities

Figures 6 - 10 explain the use of shadow links and shadow capacities. Let us consider an example. Figure 6 shows a peer/MRN node that has two incoming and two outgoing fibres each carrying three λs. The bottom part is a wavelength cross-connect, that has two E/O and two O/E converters that connect to the electronic part of the node. In the upper part (marked as 'TDM') the signals can be groomed (or added, or dropped). The figure shows, that the content of two λ-paths is groomed into a single one.

Now, let us see the model of this node (Figure 7). On the left hand side part of the figure we show an example for setting up the internal link weights to be used for routing. Wavelength transition is cheaper (25 cost units) than using the electronic layer, that will cost at least $50 + 50 = 100$ cost units.

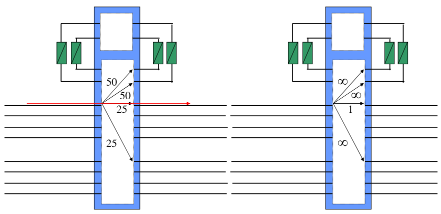

Fig. 7. Setting weights in the GG.

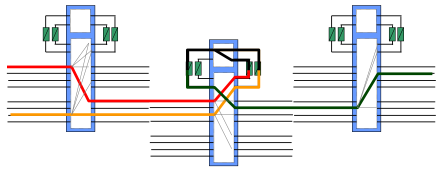

Fig. 8. Routing with grooming.

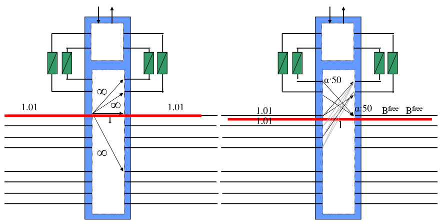

Fig. 9. Creating a "shadow link" of "shadow capacity" of B^{free}.

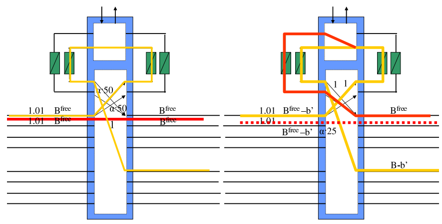

Fig. 10. If routing over the shadow capacities, the λ-paths will be cut.

Based on these weights set for all the internal and external links in the network model we search for a shortest path between certain nodes. In the right-hand side part of Figure 7 we have chosen a transition, while Figure 8 shows a grooming. Routing is always followed by re-setting the link weights. The right-hand side part of Figure 7 shows the approach used for the simple grooming model, while Figures 9 and 10 introduce the shadow links.

Figure 9 shows that after routing a demand using any of the shortest path algorithms (e.g., Dijkstra's), the internal links connected to internal nodes used by that demand will neither be deleted, nor will be their costs increased to infinity, but increased enough to avoid using those links until other wavelengths or other paths exist. In figure we have multiplied the weights of these links by parameter $\alpha >> 1$. It means that the model allows not only the used internal link, but introduces more expensive shadow links having as much shadow capacity as the free capacity of the internal link used by the considered demand is. This does not mean branching the optical signal, but it gives opportunity to choose instead of the current internal optical link cutting (fragmenting) the λ-path and going to the upper, electronic grooming layer.

Figure 10 shows such a case when there was no cheap alternative wavelength or path, and the more expensive shadow link of the GG had to be chosen while searching for the shortest path that resulted in cutting the λ-path. The right-hand side of the figure shows that the two traffic streams are now demultiplexed (de-groomed), a new shadow link with new shadow capacity has been defined (dotted thick line).

Until there is any free capacity in the λ-paths, they will have shadow links of shadow capacities equal to the free capacity.

In the upper example, we have shown how a λ-path can be cut for grooming purposes. Similarly, if a λ-path does not carry any traffic, it will be cut into λ-links, and the capacity and weight values of these links will be set to their initial values. We refer to this action as *λ-path fragmentation*.

Similarly, two λ-paths can be concatenated if they use the same wavelength AND they are connected to the same grooming node, but there is no third traffic that has to be added or dropped. Although it happens rarely, it is very useful in case when the number of grooming ports is the scarce resource. We refer to this action as *λ-path defragmentation*.

In the remaining part of this paper based on simulation results we show what parameters influence and how do they influence the performance and dynamic behaviour of the network. The blocking was in all cases the lowest for this adaptive grooming approach with λ-path fragmentation and de-fragmentation.

4 Numerical Results

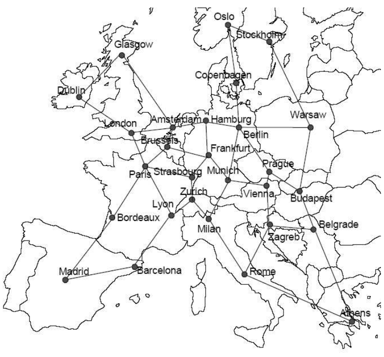

Fig. 11. Basic COST266 topology.

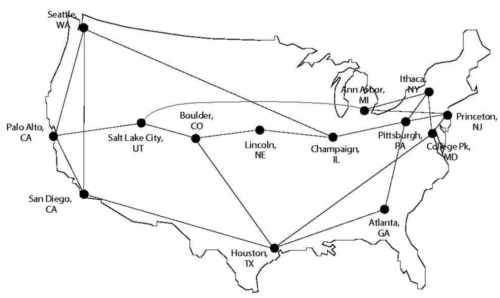

Fig. 12. The NSFnet topology.

The code was written in C++ under Linux and Windows operating systems, while the simulations were carried out on a Linux MSI K7Dual AMD Athlon 2000+ MP workstation with 2 GBytes of RAM, 2.4.18 kernel. We have applied DES (Discrete Event Simulation) where we route the demands in the given order, however, to speed up the simulation we do not wait between two demands as the time stamps determine, but route the next demand as soon as the last demand is routed.

The test networks were the COST 266 European reference Network [20] consisting of 25 nodes and 32 physical links shown in Figure 11 and the NSFnet consisting of 14 nodes and 21 links shown in Figure 12. We have used OADMs in all nodes of degree 2 and OXCs with grooming capability in all other nodes. We have compared the behaviour of three network node models:

– OXC: Optical cross-connect with no wavelength-conversion capability and no grooming capability.
– OGS: OXC with gromming capability. This is the *simple* grooming node model that we proposed earlier and was used by other authors as well.
– OGT: OXC with grooming capability with support for *"tailoring"* λ-paths, i.e., adaptive, distributed fragmentation and de-fragmentation of λ-paths. This is our new method proposed in this paper.

We investigate how the blocking ratio depends on three parameters, namely the bandwidth of demands, the holding time of sessions and the number of λs per link.

We have assumed 6 wavelengths per link, 1000 units of capacity for all wavelengths, 100 units of bandwidth on average and 8 units of holding time for the demands as the default values, for both, COST266 and NSF networks. Session arrival rate was 0.025 for the COST266 while it was 0.08 for the NSF network.

As a reference the OXC case was used, i.e., all nodes were OXCs without λ-conversion capability. In this case all the traffic demands have used exclusive λ-paths.

Bandwidth of Demands: First we tune the ratio of the average bandwidth of the demands to the capacity of λ-links (Figure 13). While the bandwidth ratio is significant, there is a huge difference in blocking. Adaptive grooming is superior to simple grooming. However, as the bandwidth ratio approaches 0,1 the blocking grows for both grooming approaches and they become comparable.

It is interesting to note, that blocking of both grooming approaches is larger than that of the approach with no grooming (OXC) as the bandwidth of the demands approaches the capacity of the λ-links. It is probably resulted by the long λ-paths that hinder routing demands over shorter paths. Note, that in our adaptive grooming framework we do not allow rerouting existing connections to other paths, but just cutting or concatenating the λ-path fragments they use for three reasons. Namely, to simplify the operation, keep the adaptive and automatic traffic engineering local, and to keep the interruption time very short.

However, in practice the typical operational region of networks falls out of this critical region, i.e., the typical bandwidth of demands is lower at least by one to two orders of magnitude than the capacity of λ-links.

Holding Time of Demands: Figure 14 shows that when increasing the holding time of connections the blocking grows. Our adaptive grooming approach (OGT) has significantly lower blocking than the other two methods, particularly for the NSF network (Figure 14). It is very interesting that simple grooming (OGS) has higher blocking for short holding times than in the case with no grooming at all (OXC)!

Number of Wavelengths: Figure 15 shows, that increasing the number of λs per link the blocking smoothly drops for the case with no grooming (OXC). The adaptive grooming model (OGT) has always better performance than the other two methods. Both grooming models have roughly the same blocking when the number of λs grow, while the performance of the model with no grooming improves. For large number of λs the simple grooming approach (OGS) has higher blocking than that with no grooming at all! The proposed grooming method has always the best performance. The curves for OXC are very smooth, while for grooming they fluctuate. This supports that grooming inherently introduces numerous anomalies.

5 Conclusion

In this paper we have proposed a new model, the Grooming Graph, that supports distributed, automatic, adaptive and on-line multi-layer traffic engineering performed through adaptive grooming using the *shadow links* and their *shadow capacities*. This approach allows the network to adapt well to changing traffic conditions. The λ-paths are *fragmented* and *de-fragmented* as the network and traffic conditions require in a fully automated, adaptive and distributed way without any centralised action or initialisation while simply using the available routing protocols!

The results show, that our approach yields the lowest blocking ratio in all cases for all scenarios studied. In some cases the blocking is by orders of magnitude lower than that achieved by known methods. Applying the proposed method in networks the throughput can be significantly increased and therefore the revenue as well, while minor investments are needed to upgrade to using this method.

The only limitation of the proposed approach is that separate wavelengths should be allocated for traffic that is sensitive even to these very short interrupts and delay variations needed for λ-path fragmentation and de-fragmentation.

References

1. E. Dotaro, M. Vigoureux, D. Papadimitriou: *"Multi-Region Networks: Generalized Multi-Protocol Label Switching (GMPLS) as Enabler for Vertical Integration"*, Alcatel Technology White Paper, February 2005, researchlibrary.theserverside.net/detail/RES/1109006898_409.html
2. M. Vigoureux, B. Berde, L. Andersson, T. Cinler, L. Levrau, D. Colle, J.F. Palacios, M. Jager: *"Multi-Layer Traffic Engineering for GMPLS-enabled Networks"*, *IEEE Communications Magazine*, July 2004, Vol.x, No.x, pp. x-x

3. E. Modiano, P.J. Lin: *"Traffic Grooming in WDM Networks"*, *IEEE Communications Magazine*, vol.39 no.7 pp.124-129, July 2001
4. T. Cinkler, Cs. Gáspár: "Fairness Issues of Routing with Grooming and Shared Protection", ONDM 2004, 8th Conference on Optical Network Design and Modelling, Ghent, Belgium, February 2-4, 2004
5. M. Perényi, J. Breuer, T. Cinkler, Cs. Gáspár: "Grooming Node Placement in Multilayer Networks", ONDM 2005, 9th Conference on Optical Network Design and Modelling, pp. 413-420, Milano, Italy, February 7-9 2005
6. T. Cinkler: *Traffic and λ Grooming"*, *IEEE Network*, March/April 2003, Vol.17, No.2, pp.16-21
7. A. Banerjee et al.: *"Generalized Multiprotocol Label Switching: An Overview of Signalling Enhancements and Recovery Techniques"*, *IEEE Communications Magazine*, pp. 144-151, July 2001, Vol. 39, No. 7
8. B. Rajagopalan et al.: *"IP over Optical Networks: Architectural Aspects"*, *IEEE Communications Magazine*, pp. 94-102, September 2000, Vol. 38, No. 9
9. R. Sabella, H. Zhang, eds.: *Traffic Engineering in Optical Networks"*, *IEEE Network*, March/April 2003, Vol.17, No.2
10. T. Cinkler et al. eds.: *Proceedings of ONDM 2003*, vol.1, ISBN-963206406 (http://www.hsnlab.hu/˜ ONDM2003)
11. K. Zhu, H. Zhu, B. Mukherjee: *"Traffic Engineering in Multigranularity Heterogeneous Optical WDM Mesh Networks through Dynamic Traffic Grooming"*, *IEEE Network*, March/April 2003, Vol.17, No.2, pp.8-15
12. S. De Maesschalck, et al.: *"Intelligent Optical Networking for Multilayer Survivability"*, *IEEE Communications Magazine*, January 2002, Vol.40, No.1, pp. 42-49
13. H. Zhu, H. Zang, K. Zhu, B. Mukherjee: *A Novel Generic Graph Model for Traffic Grooming in Heterogeneous WDM Mesh Networks*, *IEEE/ACM ToN: Transactions on Networking*, Vol.11, No.2, pp. 285-299, April 2003
14. C. Ou, K. Zhu, H. Zang, L.H. Sahasrabuddhe, B. Mukherjee: *Traffic Grooming for Survivable WDM Networks - Shared Protection*, IEEE JSAC: Journal on Selected Areas in Communications, Vol.21, No.9, pp. 1367-1382, November 2003
15. T. Cinkler, R. Castro, S. Johansson: *"Configuration and Re-Configuration of WDM Networks"*, NOC'98, European Conference on Networks and Optical Communications, Manchester, UK, June 1998
16. T. Cinkler: *"ILP Formulation of Grooming over Wavelength Routing with Protection"*, IFIP ONDM 2001, 5th Conference on Optical Network Design and Modeling, Wiena, February 2001
17. T. Cinkler, D. Marx, C.P. Larsen, D. Fogaras: *"Heuristic Algorithms for Joint Configuration of the Optical and Electrical Layer in Multi-Hop Wavelength Routing Networks"*, *IEEE INFOCOM 2000*, pp.1000-1009, Tel Aviv, March 2000
18. Cs. Gáspár, G. Makács, T. Cinkler, J. Tapolcai: *"Wavelength Routing with Grooming and Protection"*, IFIP ONDM 2003, Optical Network Design and Modelling, Budapest, Hungary, February 2003
19. S. Johansson: *"Transport Network Involving a Reconfigurable WDM Network Layer - A European Demonstration"*, *IEEE Journal on Lightwave Technology*, vol.14, no.6, pp. 1341-1348, June 1996
20. COST 266: http://www.ure.cas.cz/dpt240/cost266/index.html; *"COST 266 Reference Scenario"*, http://ibcn.atlantis.rug.ac.be/projects/COST266_IST_lion/NRS/index.html, January 2002

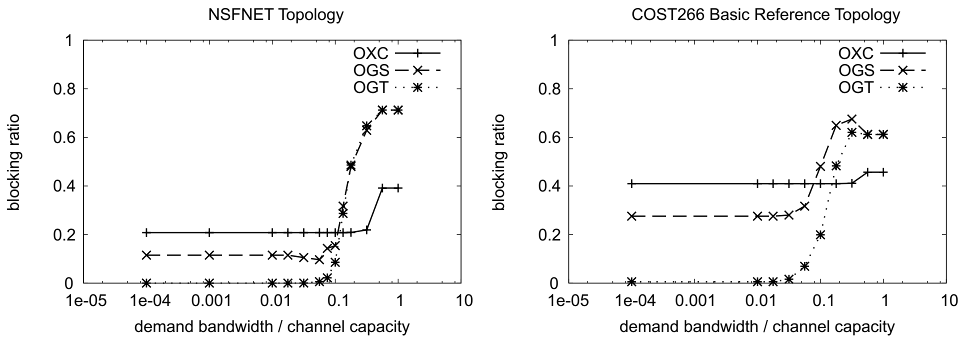

Fig. 13. Blocking ratio vs. the ratio of the demand bandwidth to the channel capacity

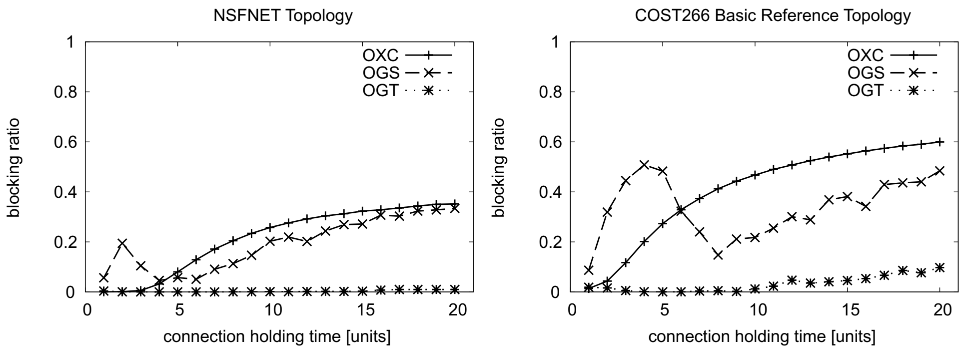

Fig. 14. Blocking ratio vs. the average connection holding time.

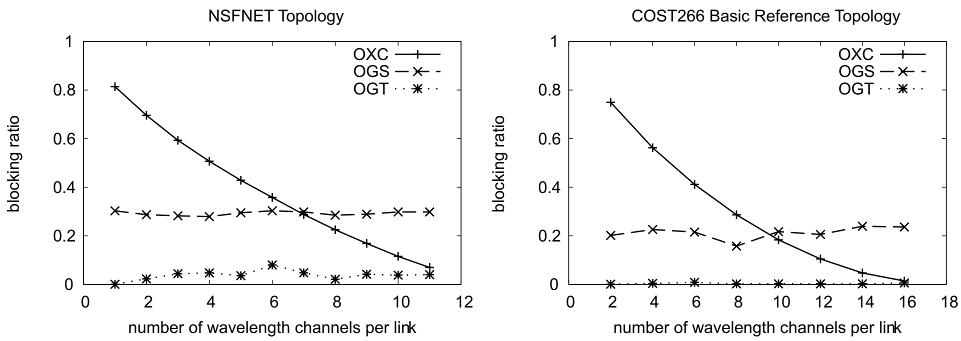

Fig. 15. Blocking ratio vs. the number of wavelengths.