Performance Analysis and Dimensioning of Multi-Granular Optical Networks ⋆

Marc De Leenheer [∗,1], Chris Develder [2], Jens Buysse, Bart Dhoedt, Piet Demeester

Dept. of Information Technology
Ghent University - IBBT
Gaston Crommenlaan 8 bus 201
9050 Gent, Belgium

Abstract

Recent years have demonstrated the limited scalability of electronic switching to realize transport networks. In response, all-optical switching has been identified as a candidate solution to enable high-capacity networking in the future. One of the fundamental challenges is to efficiently support a wide range of traffic patterns, and thus emerges the need for equipment that is both practical and economical to construct and deploy. We have previously proposed the use of multi-granular optical cross-connects (MG-OXC), which support switching on both the wavelength and sub-wavelength level. To this end, the MG-OXCs are equipped with cheap, highly scalable slow switching fabrics, as well as a small number of expensive fast switching ports. The goal of this work is two-fold: first to demonstrate that a small number of fast switching ports suffice to support a wide range of traffic requirements, and second that multi-granular optical switching can offer cost-benefits on a network-wide scale. The first objective is studied through simulation analysis of a single switching node, and results indicate that a limited number of fast switching ports can significantly improve burst blocking performance over slow only switches. Furthermore, under certain circumstances, the MG-OXC can even approach the performance of a fast only switch design. Secondly, we introduce an Integer Linear Programming model for the total network installation cost, and our evaluation indicates that multi-granular optical switching can be a cost-effective solution on the network level, in comparison to slow only or fast only approaches. Furthermore, we can achieve reduced costs of individual OXC nodes, which allows us to minimize scalability problems corresponding to emerging fast switching fabrics.

Key words: multi-granular optical networks, simulation analysis, integer linear programming, dimensioning
PACS: 07.05.Tp, 42.79.Sz, 42.79.Ta

1 Introduction

Optical networks have a proven track-record in long-haul, point-to-point networking, where large amounts of data are transported in a cost-effective way. An enabling technology is Wavelength Division Multiplexing (WDM), as it allows multiple signals (wavelengths) concurrent access to a single fiber. However, interest is growing to use optical networks in edge and even access networks (e.g. Fiber To The Home or FTTH), mostly because of the predictable performance of photonic technology (i.e. high bandwidth, low latency). A major issue is O/E/O (optical/electronic/optical) conversions in the network, because the speed of electronic processing can not match the bandwidths currently offered in the form of wavelengths of 40 Gbps and higher. For this reason, most current research is focusing on all-optical networking solutions.

As of today, it is possible to create all-optical networks through the use of circuit-switched paths, which essentially reserve one or more full wavelengths between end points. For instance, Lambda Grids are a general term to refer to Grid applications making use of wavelengths (i.e. lambdas) to connect high-performance computing sites over an optical network [1]. However, novel applications are appearing which demand a much more fine-grained access to bandwidth capacity, as is demonstrated for instance in consumer Grids [2]. In such a scenario, data sizes become smaller, since aggregation of multiple data sources is much harder, and the bandwidth utilization would drop dramatically if full wavelengths were used by these applications. Consequently, the network must support reservation and allocation of bandwidth on a sub-wavelength scale. In this paper, we propose a generic multi-granular optical switch architecture, which supports both circuits (wavelength level) and bursts (sub-wavelength level).

As of today, it is possible to create all-optical networks through the use of circuit-switched paths, which essentially reserve one or more full wavelengths between end points. For instance, Lambda Grids are a general term to refer to Grid applications making use of wavelengths (i.e. lambdas) to connect high-performance computing sites over an optical network [1]. However, novel applications are appearing which demand a much more fine-grained access to bandwidth capacity, as is demonstrated for instance in consumer Grids [2].

⋆ This work was partially supported by the FP6 IP Phosphorus project.
∗ Corresponding author.
Email addresses: marc.deleenheer@intec.ugent.be (Marc De Leenheer), chris.develder@intec.ugent.be (Chris Develder), jens.buysse@intec.ugent.be (Jens Buysse), bart.dhoedt@intec.ugent.be (Bart Dhoedt), piet.demeester@intec.ugent.be (Piet Demeester).
[1] Funded by the IWT through a Ph.D. grant.
[2] Funded as a post-doctoral fellow by the F.W.O.

This paper will investigate the specific details of realizing optical networking solutions that allow efficient transfer of both large (circuit) and small scale (burst or packet) data sizes. This is useful for future optical (possible Grid) deployments, which must support a new and emerging generation of distributed network-based applications that combine scientific instruments, distributed data archives, sensors, computing resources and many others. Each application has its own traffic profile, resource usage pattern and different requirements originating in the computing, storage and network domains [3]. Dedicated networks do not offer sufficient flexibility to satisfy the requirements of each application type, nor are they economically acceptable. Hence it is vital to understand and redefine the role of networking, to support applications with different requirements and also offer service providers a flexible, scalable and cost effective solution. A dynamic optical network infrastructure with the ability to provide bandwidth granularity at different levels is a potential candidate. In this way, the network can adapt to application requirements and also support different levels of Quality of Service (QoS). However, care must be taken to devise a solution that remains scalable and cost effective.

More specifically, we define a *multi-granular optical switched network* as a network that is able to support dynamic wavelength and sub-wavelength bandwidth granularities with different QoS levels. As such, the network will support the three basic switching technologies in WDM networks; optical circuit switching (OCS), optical packet switching (OPS) and optical burst switching (OBS). In order to support these switching approaches, optical switching fabrics with speeds on the millisecond range down to nanosecond range must be considered. As we will demonstrate in the following section, OCS can utilize millisecond switching technologies efficiently, whereas this switching speed causes bandwidth inefficiency and unpredictability for the performance of OBS. This is mainly caused by the high overhead incurred by large offset times required to configure slow switches. Consequently, fast switching fabrics should be introduced in the network.

The ideal solution would thus consist of deploying fast switches of large dimension; however current technology can only realize fast fabrics of limited scalability at a very high cost (for more details, a review of existing switching technologies is provided in [4,5]). Therefore, one possible solution is an optical cross-connect (OXC) which combines both slow and fast switching elements, with careful consideration of scalability and cost properties. Furthermore, users and applications can decide on slow or fast network provisioning, and additionally the network service provider can optimize bandwidth utilization by allocating wavelengths or lightpaths according to the traffic's switching needs.

In summary, the multi-granular OXC (MG-OXC) has a number of distinct advantages over traditional single-fabric switches:

- Bandwidth provisioning and switching capability at fiber, wavelength and sub-wavelength granularities;
- Agility and scalability of switching granularities providing a dynamic solution;
- Fast reconfigurability and flexibility on the electronic control of switching technologies;
- Cost-performance efficiency by offering an optimal balance between slow and fast switch fabric technologies.

A number of authors have previously investigated multi-granular optical networks, although these generally focused on granularities higher than a single wavelength. For instance, multiple proposals have advocated the use of a multi-layer cross-connect to allow wavelength, waveband and even fiber switching [6,7]. In contrast, this work is the first to introduce multi-fabric cross-connects supporting wavelength and sub-wavelength switching in a single network layer. Furthermore, our solution contrasts with multi-layer approaches, such as IP over WDM, where multiple layers of control are introduced to support the wide range of traffic parameters. A related subject is the design of the optical path layer in networks with cross-connects; this issue is usually denoted as the routing and wavelength assigment (RWA) problem [8,9]. The complexity increases when considering multi-granular traffic, and two important subproblems can be identified: traffic grooming in optical networks [10,11], and the RWA-problem for multi-granular traffic (wavelength-scale and higher, i.e. waveband and fiber) [12]. Other research has focused on dimensioning individual nodes, such as [13]. In contrast, the algorithm presented in Section 4 focuses on optimizing total network and individual node costs, by appropriate routing decisions in a multi-granular (wavelength-scale and below) optical network.

The remainder of this paper is organized as follows. Section 2 further elaborates on the need for multi-granular switching, along with an indication of the most important challenges associated with the concept. Simulations are then used to evaluate a generic MG-OXC design for various traffic and design parameters in Section 3. A dimensioning study for the optimal design of a multi-granular optical network is presented in Section 4, while the concluding Section 5 summarizes our findings and discusses a number of remaining challenges.

2 Problem Statement

The basic function of an optical switch is straightforward: create a connection between an input and an output port for each incoming data packet. The decision which output port a data packet should be directed to is usually made

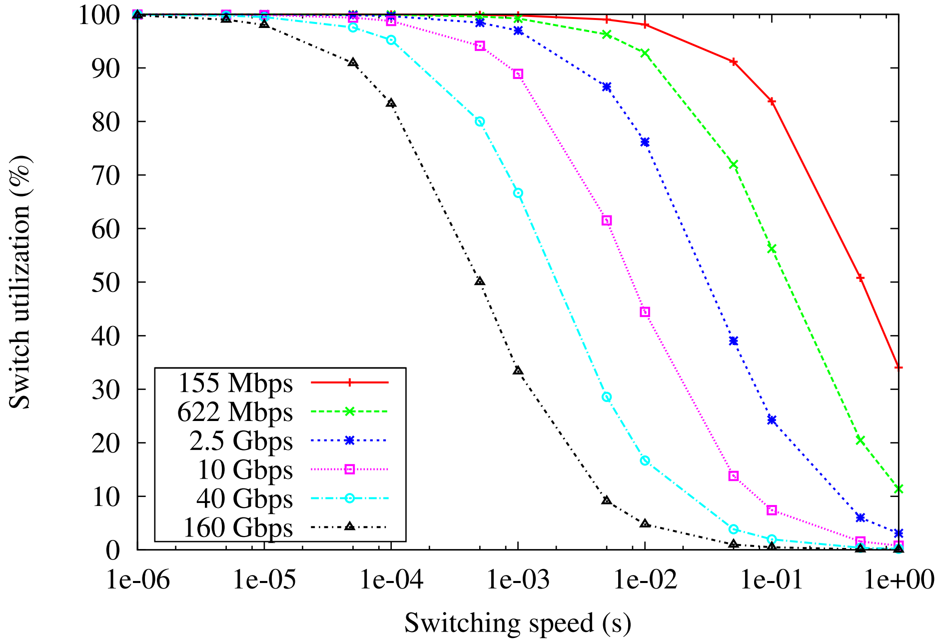

Fig. 1. Upper bound for utilization of an optical switch for different switch speeds and bandwidths (data size is 10 MB)

in a control unit available at each optical switch. This unit receives control information from each data transfer, which can be a reservation packet long in advance in the case of circuit switching, or a header prepended to the actual data in the case of packet switching. In this work, we assume data is sent in bursts (OBS), and control information is sent ahead of the actual data on a seperate control plane (i.e. out-of-band signaling). The time between the control packet and the actual data transfer is denoted by T_{offset}, and is the time available to the switch to reconfigure its internal cross-connections. Each switching fabric (see [4,5] for current technologies) is limited by its switching speed T_{switch}, and thus a data burst can only be switched successfully if $T_{switch} < T_{offset}$. T_{data} represents the length of the actual data in time, and thus is the time the switch's connection is in use. From this we can derive the *switch utilization*, i.e. the maximum fraction of time the switch is actually transferring data:

$$\frac{T_{data}}{T_{data} + T_{switch}}.$$

An illustration of this can be found in Figure 1, which shows the maximum utilization of an optical switch as a function of varying switching speeds. The data transferred has a size of 80 Mbit (10 MB), and the curves are shown for different link speeds. If we take, for instance, a switch speed of 10 ms (a representative value for micro-electro-mechanical systems or MEMS-based switches), we see that the switch utilization is 76% for a 2.5 Gbps link speed.

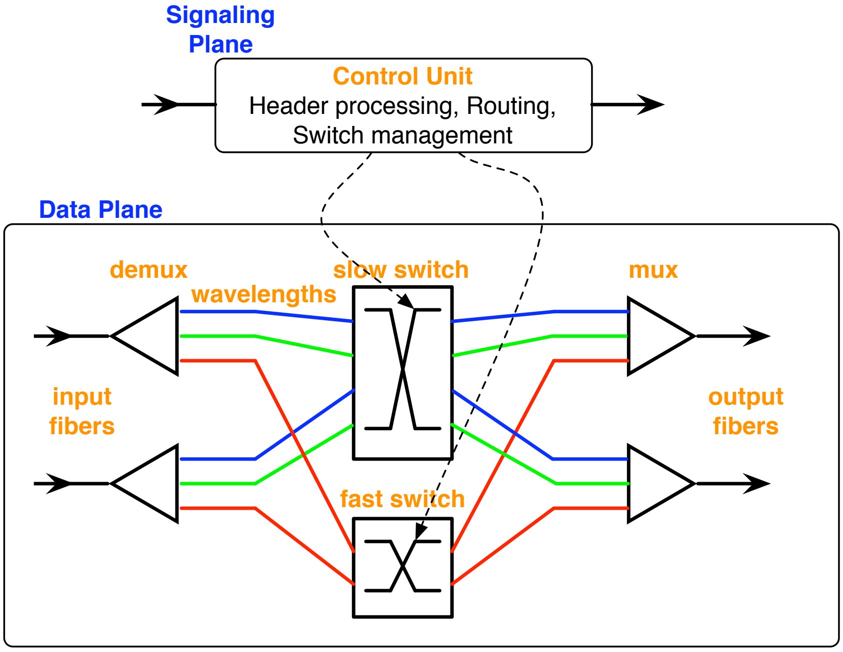

Fig. 2. Multi-granular optical switch supporting wavelength and sub-wavelength switching

This value drops to below 20% for 40 Gbps link speeds, and the situation clearly becomes worse for even higher bandwidths. Obviously, the same argument holds for a fixed bandwidth and decreasing data sizes. In contrast, a semiconductor optical amplifier or SOA-based switch can achieve nanosecond switching speeds, and is thus much better adapted to support the full range of data sizes and bandwidths required for OCS, OBS and OPS.

The example shows that, to support very long data transfers (i.e. circuits), slow switching speeds are usually sufficient to obtain a high switch utilization, even for very high speed link rates. However, for smaller data transfers (burst or even packet sizes), high speed switching fabrics are required to achieve acceptable throughput in optical switching nodes. As current and emerging applications generate data according to very diverse distributions (both the data sizes and the instants of time at which the data is created), the idea emerged to integrate multiple types of switching fabric into a single optical switch. This concept is generally referred to as *multi-granular optical switching*, and becomes essential if a single, unified data plane needs to support a wide range of users and applications. This is especially true if complex grooming is to be avoided, which can be implemented in either a single-layer or multi-layer approach. The single-layer variant corresponds to the use of burst assembly algorithms, which have a negative effect on latency, while multi-layer grooming is less dynamic as multiple layers of control need to be activated before actual data transmission.

Global network optimization not only depends on efficiency and utilization, but also on the feasibility to offer this technology in a cost-effective and practical way. Current optical switching technologies offer a broad range of switching speeds, but faster switching speeds generally have two distinct disadvantages: cost and scalability. For instance, MEMS switches have a typical switching time in the millisecond range, while it is technologically feasible to produce port counts of for instance 1000x1000. In contrast, SOA technology can only scale up to 32x32 port counts at very high cost, but at the same time can achieve switching speeds in the nanosecond range. Hence, cost-effectiveness is an important driver for hybrid optical switch designs requiring only a limited amount of expensive fast switching components. In response to this, Figure 2 presents the generic design of a multi-granular optical cross-connect (MG-OXC). The switch is composed of two separate switching fabrics, in order to support various application and QoS requirements on a common transport network infrastructure. As our results show (see for instance Section 3), even a minimal amount of fast switching fabrics can achieve considerable improvements in network performance. Hence, we can support OBS/OCS for latency-critical traffic with a cost-effective switch architecture, eliminating the need for switch designs composed completely out of expensive fast switch ports.

A final note is related to the practical realization of the MG-OXC, where several architectural choices remain an open research challenge. For instance, a sequential design (where the fast switching fabric is cascaded behind the slow fabric), allows reconfiguration of the fast wavelengths, at the expense of an increase in dimensionality of the slow switch. The design depicted in Figure 2 places the two switching fabrics in parallel, and results in a slightly smaller slow switching matrix, but loses the reconfigurability of the fast wavelengths. In Section 4.3, we will demonstrate that allowing reconfigurability of fast wavelengths has a negligable influence on the total network cost. Refer to [14] for further details regarding the architecture and performance of the multi-granular OXC.

In the following, we will show the potential improvements in blocking performance (Section 3), and that multi-granular switching also provides economic advantages on the network level (Section 4).

3 Simulation Analysis

In this section, simulation analysis is used to provide insight in the behaviour of the MG-OXC. The implementation allows us to evaluate an MG-OXC in a generic way, independent of architectural details. Note that for fixed F, W and Y values, the designs presented in the previous section are functionally equivalent. A comparison between MG-OXC and traditional, single-speed OXCs

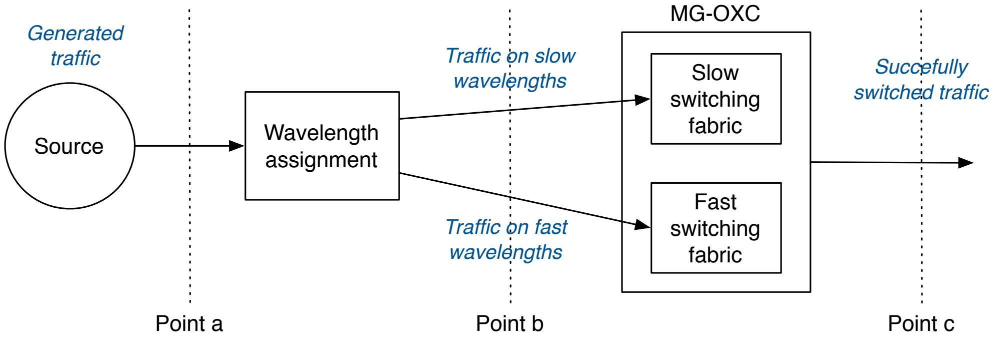

Fig. 3. Overview of node simulations

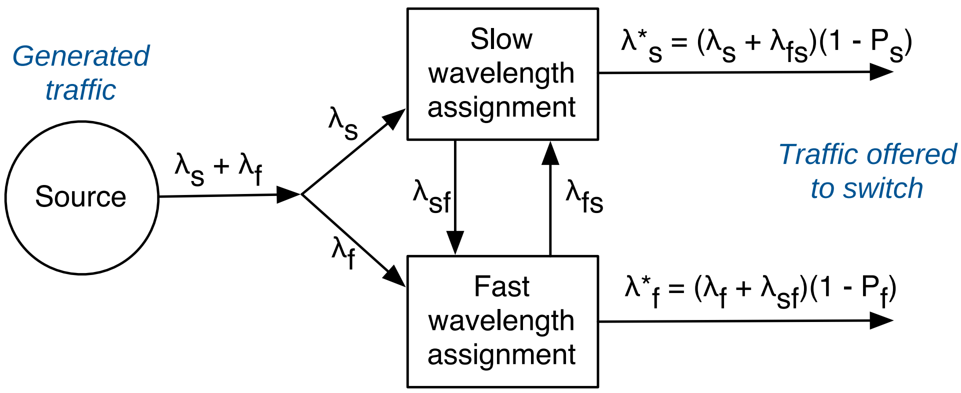

Fig. 4. Overview of wavelength assignment

(slow only, fast only) is presented, and results are given for varying traffic load, fractions of slow/fast traffic and number of slow/fast ports available. However, we start with introducing the different approaches to wavelength assignment, which are necessary for mapping incoming data bursts to a suitable wavelength. For a general overview of the node simulations, and to observe the different steps in which traffic is processed, refer to Figure 3.

3.1 Wavelength Assignment

The introduction of an MG-OXC in a network effectively creates a wavelength partitioning, by grouping wavelengths that are switched on the same type of switching fabrics. As such, an algorithm is required to assign generated traffic to a suitable wavelength partition, and the available wavelengths within a partition. This algorithm will be executed at the network's edge, thus before entering the all-optical data transport network (see Figure 3). In the following, the assumption was made that only two partitions (corresponding to slow and fast) are introduced.

As shown in Figure 4, generated traffic is first classified in slow (arrival rate λ_s) and fast (λ_f) traffic flows, by inspecting the offset time T_{offset} between the

	Simple	Slow-to-fast	Fast-to-slow	Greedy
λ_{sf}	0	$\lambda_s P_s$	0	$\lambda_s P_s$
λ_{fs}	0	0	$\lambda_f P_f$	$\lambda_f P_f$

Table 1
Transfer rates (number of bursts per time unit) between slow and fast wavelength assignment blocks

burst header and the actual data burst. Obviously, for slow traffic it holds that $T_{offset} > T_{slow}$ (T_{slow} the switching speed of the slow switch), while $T_{offset} < T_{slow}$ is true for fast traffic (T_{fast} the switching speed of the fast switch fabric). Based on this classification, a number of alternatives are now possible for assignment of traffic to the wavelength partitions.

The approaches differ in the way traffic is transferred between the slow and fast wavelengths partitions. *Simple* wavelength assignment is the most basic approach, whereby slow bursts are assigned to the slow wavelength partition, and the burst is dropped in case no free wavelength is available. Fast bursts are considered for assignment to the fast wavelength partition in a similar way. The *slow-to-fast* approach differs from the simple algorithm by allowing slow bursts on the fast wavelengths, only in case these can not be accommodated on the slow wavelengths. The corresponding *fast-to-slow* wavelength assignment allows transfer of fast bursts onto slow wavelengths (again only in case the fast burst can not be assigned to a fast wavelength). We motivate the use of this algorithm as follows: although the slow switch can not be configured in time for a fast burst, it is possible that the preceding (slow) burst requests the same output, and thus reconfiguration of the switch is not required. Finally, the *greedy* approach allows transfer of traffic between both wavelength partitions, again only when no available capacity can be found for the original wavelength assignment.

Let λ_{sf} be the transfer rate from the slow to the fast wavelength assignment block, and λ_{fs} from fast to slow. Then Table 1 shows the transfer rates for the different wavelength assignment approaches. Here, P_s and P_f represent the blocking probabilities of the slow and fast wavelength assignment blocks, and are given by:

$$P_s = Erl\left(\frac{\lambda_s + \lambda_{fs}}{B}, W_s\right) \text{ and } P_f = Erl\left(\frac{\lambda_f + \lambda_{sf}}{B}, W_f\right)$$

$$Erl(\rho, W) = \frac{\frac{\rho^W}{W!}}{\sum_{i=0}^{W} \frac{\rho^i}{i!}} \tag{1}$$

In these expressions, B represents the bandwidth of a wavelength, W_s and W_f are the number of slow and fast wavelengths in the slow and fast partitions

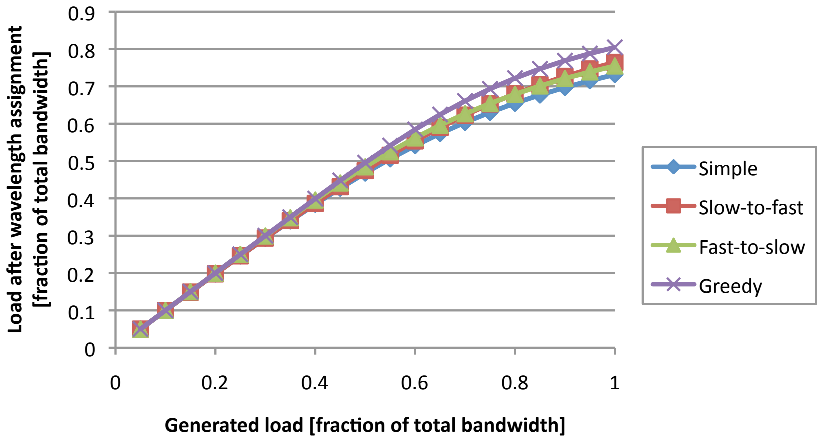

Fig. 5. Total load after wavelength assignment is similar for the various algorithms

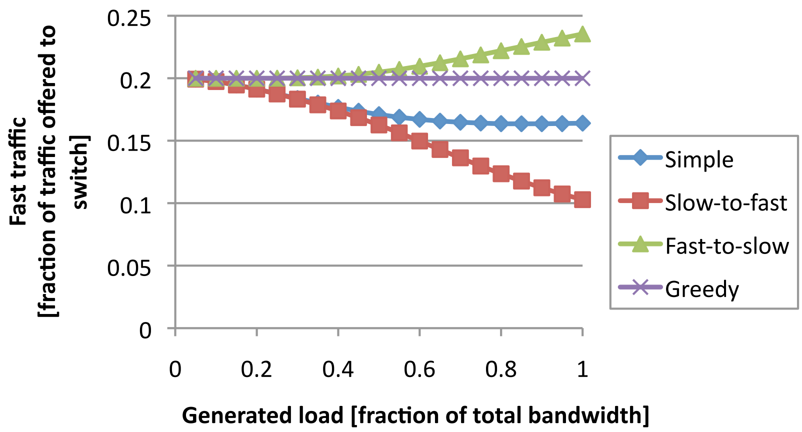

Fig. 6. Fraction of fast traffic after wavelength assignment: depending on wavelength assignment, the fraction of traffic sent on fast wavelengths deviates from the generated $\beta = 0.2$.

respectively, and $Erl(.)$ is the Erlang-B function as defined in Equation 1. Note that in case of *greedy* wavelength assignment, P_s and P_f depend on each other, and thus an iterative substitution is required to obtain the respective blocking probabilities.

To demonstrate the influence of these alternatives, Figure 5 shows the load remaining after wavelength assignment, the plot shows the total load at point b for varying generated loads at point a (Figure 3). This result was obtained by assuming $W_f = 2$ and $\beta = \frac{\lambda_f}{\lambda_s + \lambda_f} = .2$, which reflects the expected low number of (expensive) fast wavelengths. It also follows that β represents the fraction of generated fast traffic to the total generated (slow and fast) traffic. Even though the total load after wavelength assignment is similar for the various approaches, Figure 6 provides more insight into which type of traffic is favoured. The figure shows the fraction of fast traffic to the total traffic after wavelength assignment, and this for a varying total generated load. In other words, we plot the fraction of fast traffic to the total traffic at point b in Figure 3, and do this for various load averages at point a. Clearly, the *fast-to-slow* approach allows more fast traffic than the greedy approach, since the latter also allows slow bursts to use valuable fast wavelengths. It should be noted however that, although not shown, the behaviour of *fast-to-slow* converges to *greedy* for increasing values of β. Both the *simple* and *slow-to-fast* algorithms preserve only small fractions of fast traffic, and are thus not well-adapted to support a multi-granular optical network scenario, as a non-negligable amount of fast traffic will be lost because of inappropriate wavelength scheduling. Since our main interest is the effect of fast traffic and fast wavelengths, the simulation results presented in Section 3.2 have been made using the *fast-to-slow* wavelength assignment algorithm.

As mentioned before, another important decision is how bursts are assigned to individual wavelengths within a partition. Strategies such as first-fit or best-fit have previously been investigated in e.g. [15–17], however this subject falls outside the scope of this work. The simulation studies in the following Section 3.2 assume a first fit strategy.

3.2 Single Node Simulations

This section presents discrete event simulation results of several OXC alternatives (slow only, fast only and MG-OXC). All designs support 2 input and 2 output fibers, each fiber carrying 10 wavelengths. Neither wavelength conversion nor buffering capability is present in any of the switch designs. Each incoming data burst has a 50% probability of choosing the first output fiber. The bandwidth of each wavelength is 10 Gbps, and traffic is generated according to a Poisson process with an average inter-arrival time of 15 ms. Data sizes follow an exponential distribution, with a varying average to establish the generated load. Because of the limited scale of currently deployed OBS networks [3], there is no conclusive data available on a number of relevant traffic parameters. Thus, to control and evaluate the influence of different traffic types, the offset times between control packet and data are modeled as a 2-phase hyper-exponential distribution. The probability density function (pdf) f is given by: $f = \alpha \times f_{slow} + \beta \times f_{fast}$, with $\alpha + \beta = 1$ and α and β representing the fractions of generated slow and fast traffic [4]. The pdf of the slow

[3] OBS is still considered an immature technology, and as such OBS testbeds/prototypes are composed of at most a few nodes.

[4] Note that this arrival model does not generate these precise fractions of slow and fast traffic. For bursts generated according to f_{slow}, it is still possible that $T_{offset} < T_{slow}$. Using the cumulative distribution function of an exponentially distributed variable, this holds for the following fraction of traffic: $P[T_{offset} \leq$

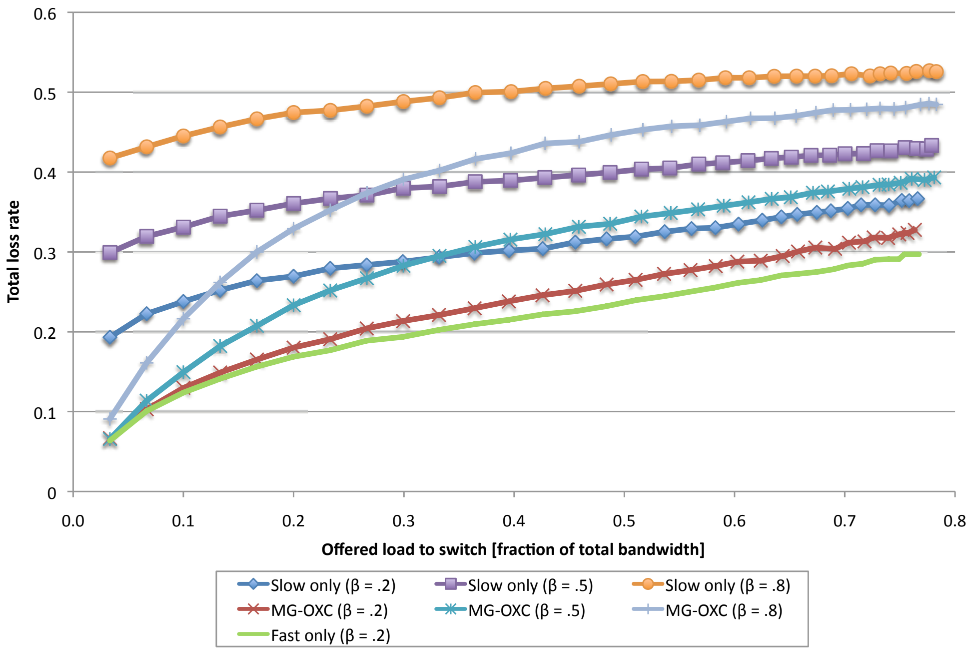

Fig. 7. Higher fractions of fast traffic increase the total loss rate for fixed number of fast wavelengths ($W_f = 2$)

f_{slow} (resp. fast f_{fast}) traffic is an exponential distribution with average 100 ms (resp. 10 ns). The slow switching fabric has a switching speed of $T_{slow} = 10$ ms, while the fast switch has $T_{fast} = 1$ ns. These values are representative for a MEMS-based (resp. SOA-based) switch [4,5]. This leads to $1 - e^{-.1} = 9.5\%$ of slow traffic that actually belongs to fast traffic, and an identical fraction of generated fast traffic will have $T_{offset} < T_{fast}$.

In the following sections, we show the performance of the MG-OXC switch, and compare the results to designs composed of a single switch fabric (slow only, fast only). To allow fair comparison of the results, the wavelength assignment algorithm is also applied in case single-fabric designs are used. This way, the offered traffic pattern at the switches' input ports is identical in all cases. As such, wavelengths are also partitioned in these single-fabric scenarios, even though the switching speeds are identical for all wavelengths. The *fast-to-slow* wavelength assignment algorithm (Section 3.1) was implemented, together with a first-fit approach for mapping data bursts on a specific wavelength within a partition. The results shown focus on the total loss rate of the switch; bursts can be lost either due to contention or because the switching speed is insufficient for a given burst.

3.2.1 Varying fraction of fast traffic

In the first experiment, 2 wavelengths are available in the fast partition, while the remaining 8 are allocated for the slow partition. Simulations were performed to evaluate the influence of the fraction of fast traffic for the three switch designs. The resulting Figure 7 shows the total loss rate (i.e. ratio of dropped traffic to the offered load) for a varying offered load. First observe that for low loads, the relatively high loss rates can be attributed to the fraction of fast traffic which has an offset time lower than the fast switching speed. However, some of these bursts can still be switched correctly as consecutive bursts taking the same output port does not require reconfiguration of the switch fabric (this explains the loss rate close to 6.5% of the fast only design in comparison to the predicted 9.5%). Then, an increasing fraction β of fast traffic causes higher loss rates, since the number of fast switching ports remains fixed (0 for the slow only, 2 for the MG-OXC). This does not apply to the fast only design (only shown for $\beta = .2$), whose performance is very similar for all fractions of fast traffic. Also, it is readily apparent that the MG-OXC outperforms the slow only design for all values of β. Another observation is that the MG-OXC offers loss rates similar to the fast only design, unless high fractions of fast traffic are generated ($\beta = .5$ and $.8$). This is not surprising considering the small number of fast switching ports available to the MG-OXC.

3.2.2 Varying number of fast wavelengths

In the following experiment, the generated traffic consisted of 80% fast traffic ($\beta = .8$). Now, simulations focus on varying the number of slow/fast wavelengths in each partition, and hence also the exact number of slow/fast wavelengths available to the MG-OXC. Figure 8 shows the total loss rate for a varying offered load, where one can immediately observe that an increased number of fast wavelengths results in a lower loss rate. That this result holds even for the slow only designs, is due to the simulation setup: the initial switch configuration connects the top input and output fibers (and likewise for the bottom fibers), and traffic is generated with a 50% probability of choosing either output fiber. Consequently, more or less half of the traffic on the W_f wavelengths can be switched correctly, and this explains why increasing values of W_f reduce the total loss rate. As before, the MG-OXC can provide an overall improved loss performance compared to the slow only design (behavior of slow only and MG-OXC are similar only for high loads and a severely under-dimensioned fast switching block). For high numbers of fast wavelengths ($W_f = 8$), the loss rate of the MG-OXC approaches the performance of the fast only design. Note again that results of the fast only design are shown

---

$T_{slow}] = 1 - e^{\frac{-T_{slow}}{T_{offset}}}$. The same argument holds for fast traffic, where $T_{offset} < T_{fast}$.

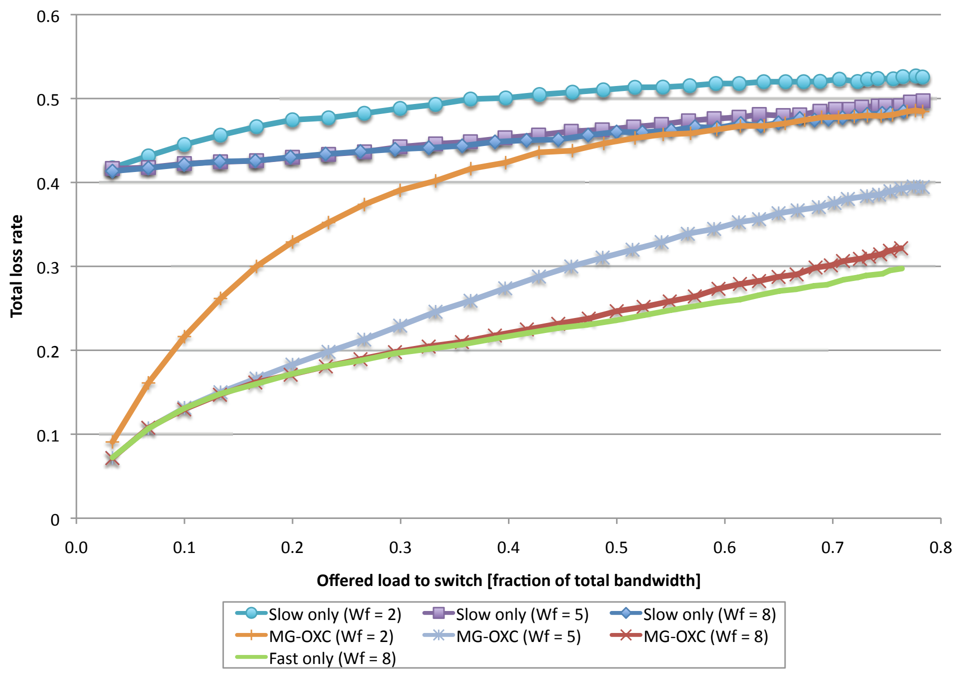

Fig. 8. Higher number of fast wavelengths decrease the total loss rate for a fixed fraction of fast traffic ($\beta = .8$)

only for $W_f = 8$, as other values for W_s lead to very similar loss rates. A final observation is that, although not shown, the loss rate of the slow only design is slightly higher in case *greedy* wavelength assignment is used, due to a assignment of fast bursts to slow wavelengths.

In conclusion, this section clearly demonstrated that an MG-OXC, equipped with only a limited amount of fast ports, can offer significant improvements in loss rates when compared to a slow only design. Furthermore, as long as the mismatch between fast traffic and fast wavelengths remains within acceptable bounds, the MG-OXC can approach the performance levels offered by a fast only design.

4 Network Dimensioning

In Section 2, we have motivated that multi-granular switching supports various application and QoS requirements on a common transport network infrastructure. We demonstrated that an MG-OXC can offer improved blocking performance for even a small number of fast ports, by using online simulation analysis. In the following, we will show that multi-granular switching also provides economic advantages on the network level. For this, an offline cost model for dimensioning a multi-granular optical network will be proposed, and results are obtained that illustrate the possible reductions in total network cost and improvements with regard to node scalability.

4.1 Problem Statement

Assume the network is composed of OXCs, capable of switching circuits or slow bursts on millisecond scale (slow MEMS switch), and fast bursts or packets on a nanosecond scale (fast SOA switch). Likewise, traffic is generated by clients requiring both fast and slow switching. The question arises how to dimension the network, given a static traffic demand with a given fraction of fast and slow traffic. The main objective is to minimize the network's cost, given a price ratio of slow over fast port costs. Another objective is to reduce the cost of the cross-connect with the highest cost, and as such obtain reduced node complexity.

The problem can be simplified by observing the very high cost and difficult scalability of fast switching fabrics [4,5]. To minimize the use of fast ports, slow traffic will be switched exclusively by slow switches, and thus this minimum cost network flow problem can be solved independently with known algorithms [18]. We do not consider this problem, and as such only need to plan the network for the remaining fast traffic. This fast traffic can be switched in one of the following ways: either on a fast switch, which can be shared between different demands [5], or on a slow switch that is then exclusively reserved for that particular demand.

Note that the proposed Integer Linear Programming (ILP) model [19] does not incorporate wavelength assignment, mainly because of complexity issues. Thus, in principle this model assumes full wavelength conversion is available in each OXC, which has consequences for the economics of the obtained solutions [20]. Otherwise, an additional wavelength assignment step is required; an overview of optimal and heuristic approaches to this problem can be found in [15,16,21]. We now proceed to the actual model, which has been formulated as a Linear Integer Programming model. The appeal of a linear model is that it allows the use of general purpose techniques (simplex and interior-point methods) to find optimal solutions.

[5] Sharing bandwidth among different demands is usually denoted as *grooming*.

wavelengths on link l, and are given by:

$$\forall (s,d), p : x_p^{sd} \geq \Lambda^{sd} \delta_p^{sd} \quad (2)$$

$$\forall l \, : x_l \, = \sum_{sd} \sum_{p} \pi_{pl}^{sd} x_p^{sd} \quad (3)$$

$$\forall l \, : y_l \, \geq \sum_{sd} \sum_{p} \pi_{pl}^{sd} \Lambda^{sd} \epsilon_p^{sd} \quad (4)$$

The auxiliary variables x_p^{sd} (integer-valued) represent the number of wavelengths required to carry the demand Λ^{sd}. Clearly, slow switching corresponds to reserving end-to-end circuits that are exclusively accessed by the source and destination, while fast switching allows grooming of traffic on a link-by-link basis. The following contraints enforce two requirements: (i) each demand can only use a single path, thereby excluding solutions based on multi-path routing, and (ii) a demand is either switched slow *or* fast, but not both.

$$\forall (s,d) : \sum_{p} (\delta_p^{sd} + \epsilon_p^{sd}) = 1. \quad (5)$$

The final step to obtain total network cost is to transform the variables for the wavelength count on each link $l = (u,v)$ (u and v represent nodes), into ports counts for each node n. For the parallel architectures (refer to Section 4.2.1 for the sequential designs), the slow and fast port counts are given by:

$$\forall n : x_n = \sum_{m} (x_{(m,n)} + x_{(n,m)}) \quad (6)$$

$$\forall n : y_n = \sum_{m} (y_{(m,n)} + y_{(n,m)}) \quad (7)$$

The first objective we propose is to minimize the total installation cost of the network, which in large part depends on the total number of installed switching ports:

$$\min \sum_{n} (x_n + C y_n). \quad (8)$$

A related objective function is to minimize the cost of the most expensive cross-connect. This objective is motivated by the limited scalability of OXC designs, especially when based on fast switching fabrics. This objective can be stated more formally as:

$$\min z \quad \text{where } \forall n : z \geq x_n + C y_n. \quad (9)$$

Design	Variables	Constraints
Slow only	$2 \cdot D \cdot P + L + N$	$D \cdot (1 + P) + L + N$
Multi-granular	$3 \cdot D \cdot P + 2 \cdot L + 2 \cdot N$	$D \cdot (1 \cdot P) + 2 \cdot L + 2 \cdot N$
Fast only	$D \cdot P + L + N$	$D + L + N$

Table 2
Complexity of ILP model for different OXC designs

4.2.1 OXC Architectures

In this section, we demonstrate how the proposed model can be adapted to support the different OXC architectures that were presented in [14]. More precisely, we show how slow only, fast only and the MG-OXC alternatives (parallel vs. sequential) can be incorporated in the model.

First of all, note that the model captures two related ILP problems, corresponding to scenarios in which either only slow or only fast switching is used. Indeed, in case $\forall (s,d), p : \epsilon_p^{sd} = 0$, all demands will be served by a slow only connection (i.e. $y_l = 0$). Likewise, in case $\forall (s,d), p : \delta_p^{sd} = 0$, only fast ports will be used ($x_l = 0$).

Furthermore, observe that in case slow only switching is used, the objective function (8) can be simplified to:

$$\sum_n (x_n + C y_n) = \sum_n x_n = \sum_n \sum_m \sum_{sd} \sum_p \lceil \Lambda^{sd} \rceil (\pi_{p(n,m)}^{sd} + \pi_{p(m,n)}^{sd})$$

which corresponds to the use of shortest path routing for all demands. This is however not the case when objective function (9) is used.

To differentiate between the parallel and sequential MG-OXC approaches, the number of slow ports in the latter case is given by:

$$\forall n : x_n^* = x_n + 2 \sum_m y_{(m,n)}. \tag{10}$$

This corresponds to the allocation of additional slow ports for each incoming fast wavelength that is introduced in a cross-connect. In the following section, we will demonstrate that network cost is only slightly increased, as a limited number of additional slow ports suffice to allow the configurability offered by the sequential switch designs (see [14] for more details).

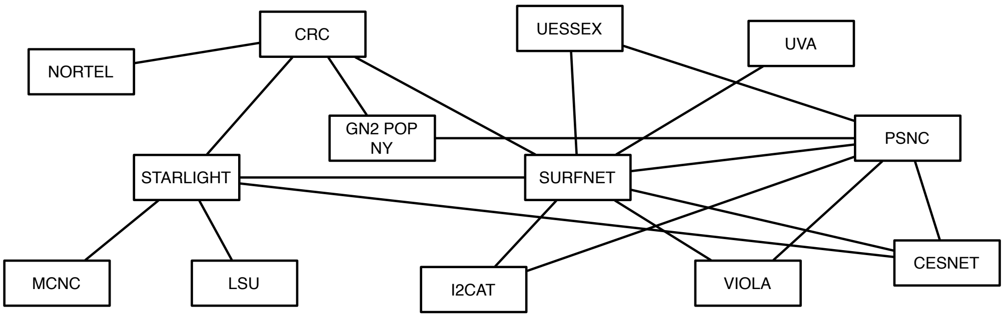

Fig. 9. Phosphorus simulated topology

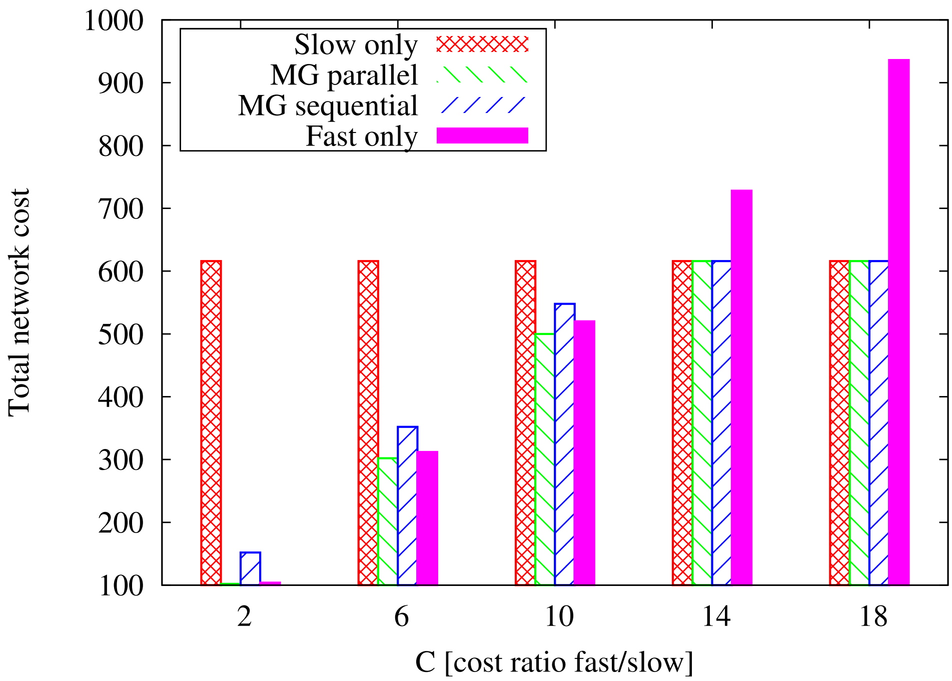

Fig. 10. Minimized total network cost

4.2.2 Complexity

Table 2 summarizes the complexity of the different ILP models. Here, N represents the number of OXC nodes, L the number of network links, D the number of demands, and P the number of paths that are considered for each demand (assumed identical for all demands). Observe that MG parallel and MG sequential have an identical complexity. The table lists the complexity when objective function (8) is used. When minimizing the highest node cost, the number of variables is increased by 1, and an additional N constraints are introduced.

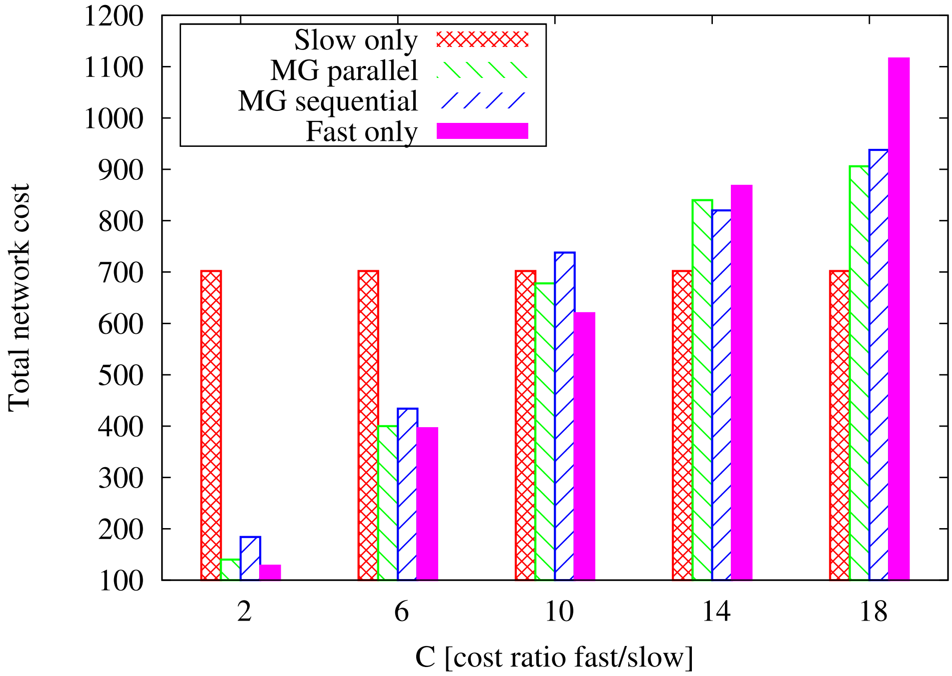

Fig. 11. Total network cost for minimized largest node cost

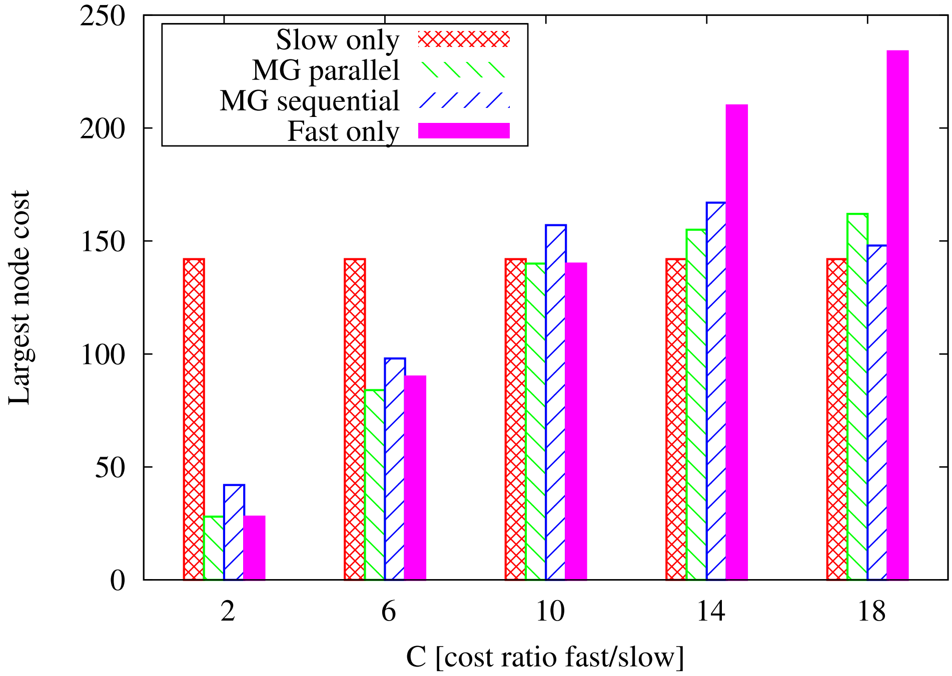

Fig. 12. Largest node cost for minimized total network cost

4.3 Evaluation

The ILP-formulated problems were implemented and solved through the use of the ILOG CPLEX library. All OXC design approaches are evaluated, including

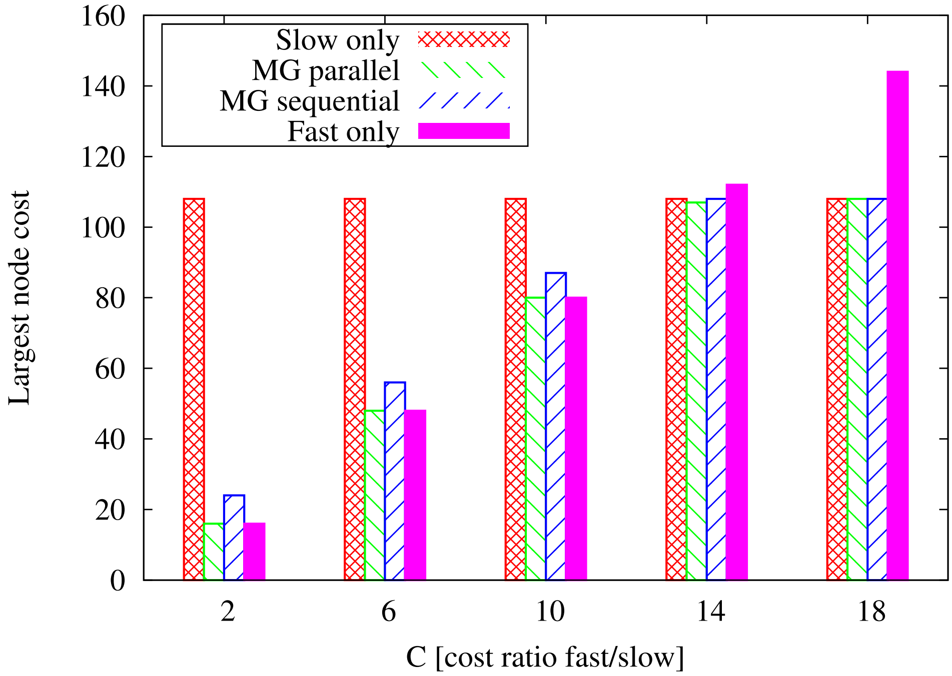

Fig. 13. Minimized largest node cost

slow only, fast only, and both multi-granular (parallel and sequential) architectures. Results are obtained for a specific scenario, defined by the Phosphorus topology depicted in Figure 9. The traffic demand matrix is fixed, and consists of uniformly generated traffic between all source-destination pairs with average $\Lambda = .05$. These low traffic demands were established in order to maximize the possibility of traffic grooming; observe that when shortest path routing is used for the given topology, the maximum number of demands making use of the same link is 15. To reduce computational complexity, we only considered the 5 shortest paths for each demand; this suffices for the topology considered, as the maximum distance between any node pair is 4 hops. Results show the total network cost and highest node cost, for both objective functions (8) and (9).

Comparing the total network cost when minimizing either total network cost (Figure 10) or highest node cost (Figure 11), a number of interesting observations can be made. First, note that slow only returns constant network costs, due to its independence of cost ratio C. As expected, minimizing the highest node cost slightly increases total network cost when compared to objective (8) (note the different Y-axis scale in Figures 10 and 11). Furthermore, MG sequential produces total network costs at least as large as MG parallel when minimizing network cost, although this is not the case when minimizing the highest node cost. Finally, for high values of C, the multi-granular approaches return identical results as the slow only design when using objective (8). In summary, significant cost savings are possible when using multi-granular optical switching, in comparison to slow only or fast only switching. Also, introducing reconfigurable fast wavelengths through the MG sequential design will only slightly increase total network cost.

We now consider the highest node cost when minimizing total network cost (Figure 12) or highest node cost (Figure 13). Again slow only produces constant results, but lower values are achieved by optimizing for objective (9). Observe that MG sequential returns highest node costs lower than MG parallel, only when minimizing the highest node cost. Multi-granular optical switching can thus clearly reduce the highest node cost, and consequently improve node complexity which is critical for scalability issues (see [14]).

5 Conclusions

In this paper, we described the trend towards all-optical switching where data remains in the optical domain from source to destination. We indicated a number of problems related to supporting a wide range of applications and services on a single, unified optical transport plane. A possible solution has been identified in the concept of multi-granular optical switching, where OXCs integrate different switching fabrics to support switching at different bandwidth granularities. An important driver of this technology is Grid computing, as a wide range of user requirements (going from consumer-oriented towards high-performance eScience applications) can be supported.

Simulation analysis of a single node compared the performance of a generic multi-granular node with OXCs composed of a single switching fabric. Analysis clearly demonstrated that even a minimal number of expensive fast switching ports can achieve significant performance improvements.

Finally, an ILP-based network dimensioning algorithm was introduced, and results indicated that significant cost savings can be obtained when implementing multi-granular optical switching. Furthermore, reduced node costs can be achieved as well, in order to minimize scalability problems corresponding to emerging fast switching fabrics.

In conclusion, this paper demonstrated the need for multi-granular optical switching, and the concept was validated through a simulation-based performance analysis. Further motivation was given by proving the possible advantages for total network cost and node complexity.

A number of research challenges remain before multi-granular switching can become a practical technology for deployment in the field. An obvious research challenge to pursue is a protocol for wavelength assignment and reconfiguration on the network level. Relevant objectives could be to minimize the number of expensive, fast wavelengths, to improve bandwidth utilization, or to reduce the influence of wavelength reconfigurations on existing traffic. Furtermore, bandwidth efficiency could be improved even further by combining (sub-)wavelength switching with waveband or even fiber-based switching. Another point of interest is an extensive dimensioning study on MG-OXCs, investigating issues such as traffic variability, multiple (> 2) switch fabrics, physical layer constraints (e.g. signal loss, BER), etc. Finally, the presented network dimensioning algorithms are based on ILP which does not scale properly for larger networks. Heuristic techniques are thus required in order to plan and dimension either new or existing, large-scale optical networks.

References

[1] T. DeFanti, C. de Laat, J. Mambretti, K. Neggers, B. St. Arnaud, *TransLight: A Global-Scale LambdaGrid for E-Science*, Communications of the ACM, 46(11):34–41, Nov 2003.

[2] M. De Leenheer, P. Thysebaert, B. Volckaert, F. De Turck, B. Dhoedt, P. Demeester, D. Simeonidou, R. Nejabati, G. Zervas, D. Klonidis, and M.J. O'Mahony,*A View on Enabling Consumer Oriented Grids through Optical Burst Switching*, IEEE Communications Magazine, 44(3):1240-131, Mar 2006.

[3] D. Simeonidou, R. Nejabati, G. Zervas, D. Klonidis, A. Tzanakaki, M.J. O'Mahony, *Dynamic Optical Network Architectures and Technologies for Existing and Emerging Grid Services*, Journal of Lightwave Technology, 23(10):3347–3357, Oct 2005.

[4] S.J. Ben Yoo, *Optical Packet and Burst Switching Technologies for the Future Photonic Internet*, Journal of Lightwave Technology, 24(12):4468–4492, Dec 2006.

[5] G.I. Papadimitriou, C. Papazoglou, A.S. Pomportsis, *Optical Switching: Switch Fabrics, Techniques, and Architectures*, Journal of Lightwave Technology, 21(2):384–405, Feb 2003.

[6] X. Cao, V. Anand, C. Qiao, *Framework for Waveband Switching in Multigranular Optical Networks: Part I - Multigranular Cross-Connect Architectures*, Journal of Optical Networking, 5(12):1043–1055, Dec 2006.

[7] R. Izmailov, S. Ganguly, T. Wang, Y. Suemura, Y. Maeno, S. Araki, *Hybrid Hierarchical Optical Networks*, IEEE Communications Magazine, 40(11):88–94, Nov 2002.

[8] H. Zhang, J.P. Jue, B. Mukherjee, *A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Networks*, Optical Networks Magazine, 1(1):47–60, Jan 2000.

[9] N. Wauters, P. Demeester, *Design of the Optical Path Layer in Multiwavelength Cross-Connected Networks*, IEEE Journal on Selected Areas in Communications, 14(5):881–892, Jun 1996.

[10] K. Zhu, B. Mukherjee, *Traffic Grooming in an Optical WDM Mesh Network*, IEEE Journal on Selected Areas in Communications, 20(1):122-133, Jan 2002.

[11] J.Q. Hu, B. Leida, *Traffic Grooming, Routing, and Wavelength Assignment in Optical WDM Mesh Networks*, Proc. 23rd Conference of the IEEE Communications Society (Infocom), Mar 2004.

[12] P.H. Ho, H.T. Mouftah, *Routing and Wavelength Assignment with Multigranularity Traffic in Optical Networks*, Journal of Lightwave Technology, 20(8):1292–1303, Aug 2002.

[13] J.Y. Choi, J.S. Choi, M. Kang, *Node Dimensioning in Optical Burst Switching Networks*, Photonic Network Communications, 13(2):207–216, Apr 2007.

[14] G. Zervas, M. De Leenheer, L. Sadeghioon, D. Klonidis, R. Nejabati, D. Simeonidou, C. Develder, B. Dhoedt, P. Demeester, *Multi-Granular Optical Cross-Connect: Design, Analysis and Demonstration*, Submitted to IEEE Journal on Selected Areas in Communications

[15] H.A. Choi, E.J. Harder, *On Wavelength Assignment in WDM Optical Networks*, Volume 468 of *The Springer International Series in Engineering and Computer Science*, pages 117–136, Springer US, 1998.

[16] M. Alanyali, *On Dynamic Wavelength Assignment in WDM Optical Networks*, In *Optical Networks: Recent Advances*, pages 1–17, Kluwer Academic Publishers, 2001.

[17] X. Sun, Y. Li, I. Lambadaris, Y.Q. Zhao, *Performance Analysis of First-Fit Wavelength Assignment Algorithm in Optical Networks*, Proc. of 7th International Conference on Telecommunications, Volume 2, pages 403–409, Jun 2003.

[18] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, *Introduction to Algorithms*, MIT Press, Sep 2001.

[19] G.L. Nemhauser, L.A. Wolsey, *Integer and Combinatorial Optimization*, Wiley-Interscience, Nov 1999.

[20] S. Subramaniam, M. Azizoglu, A.K. Somani, *All-Optical Networks with Sparse Wavelength Conversion*, IEEE/ACM Transactions on Networking, 4(4):544-557, Aug 1996.

[21] H. Ghafouri-Shiraz, G. Zhu, Y. Fei, *Effective Wavelength Assignment Algorithms for Optimizing Design Costs in SONET/WDM Rings*, Journal of Lightwave Technology, 19(10):1427–1439, Oct 2001.