Efficient Scheduling for High-Throughput OBS Networks with Partial Wavelength Conversion

Vivek Mehta, Neil Barakat, and Edward H. Sargent
Department of Electrical & Computer Engineering, University of Toronto
10 King's College Road, Toronto, ON, M5S 3G4, Canada
E-mail: {vivek.mehta, neil.barakat, ted.sargent}@utoronto.ca

Abstract

Cost-efficient contention resolution is an urgent topic in Optical Burst Switching. This paper examines the performance of a core OBS node equipped only with partial wavelength conversion, and investigates how the choice of burst scheduling algorithm can affect its blocking rate performance and converter usage. In particular, sensitivity of performance is examined with respect to burst length. A novel threshold-based scheduling scheme called Conversion Reduction Scheduling (CRS) is presented. We show that CRS can ensure that a node requires only 65-75% of the full number of wavelength converters under dynamic traffic loads. We show a further 35-40% converter savings through a proactive scheduling algorithm called Coordinated Wavelength Assignment (CWA) applied at the ingress nodes. By minimizing redundant use of wavelength conversion, the strategy reduces the required number of converters while preserving high burst throughput.

1. Introduction

Investigations of scheduling algorithms for OBS often assume that contention resolution is achieved in the wavelength domain through *full* wavelength conversion [1-3]. In large OBS networks with a high degree of connectivity between nodes, full conversion could represent a significant expense. Signal quality degradation may also result in some converters, constraining the number of allowable conversions per lightpath.

In reality only a subset of converters are in use at any node at any time. Given this fact, in combination with the cost of conversion, it is worthwhile to investigate whether full conversion is redundant and partial conversion sufficient.

Since the use of fibre delay lines to accomplish time-domain convention avoidance incurs added cost, complexity, jitter, and delay, we focus on scheduling algorithms designed for an OBS node that employs partial wavelength conversion and no FDL buffers. We consider a share-per-node architecture (Fig. 1) in which a bank of tunable wavelength converters is shared among output channels. This architecture employs a modest number of wavelength converters per node and features high converter utilization. A previous study has examined and quantified the potential converter savings achievable through partial wavelength conversion in OBS systems [4]. The effect of header offsets and the dependence of the results on burst length were not considered.

In the present work, we study *partial* wavelength conversion with realistic network traffic. We propose a dynamic scheduling algorithm which seeks to minimize redundant conversion operations. We also propose a proactive scheduling strategy to allow further converter savings without requiring extra fibres or FDL buffers. Recent investigations have shown significant performance gains from proactive scheduling algorithms [5].

Section 2 provides an overview of the common scheduling algorithms implemented with full wavelength conversion. Section 3 introduces a new scheduling algorithm, for use with partial wavelength conversion, which seeks to optimize the use of converters under dynamic traffic loads. In Section 4 we examine how a coordinated wavelength scheduling strategy between OBS nodes can further reduce the number of converters required. Section 5 shows simulation results carried out on a single node, and concluding remarks then follow in Section 6.

2. Scheduling algorithms

Upon receiving a control header packet, an OBS node's scheduler must quickly determine whether bandwidth can be reserved for the incoming burst. In most cases, several wavelengths will be available to accommodate the burst on the outgoing link. In OBS-JET, which we assume herein, there are two major challenges to the scheduling problem. First, a large number of idle

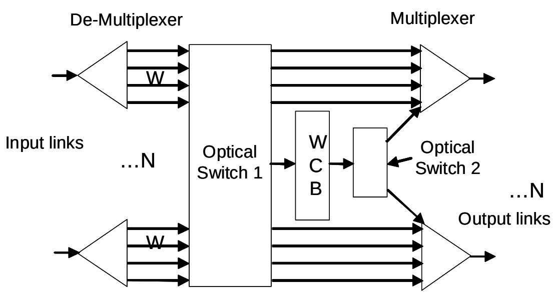

Figure 1. Model of an OBS node with partial wavelength conversion. The wavelength convertible bank (WCB) contains N_C converters where $N_C < W$.

periods (voids) are likely to be created on wavelength channels due to the dynamically random arrival of bursts. An inefficient scheduling algorithm would leave these channel fragments unused, wasting bandwidth. Secondly, bursts may be initialized with different offset lengths according to their path distance and, as the burst is processed at each intermediate node, these offsets will shrink. Thus, headers will often be processed, and scheduling decisions rendered, for bursts arriving out of order. Schedulers must therefore efficiently accommodate bursts in the absence of complete state information.

To overcome the problem of unused bandwidth due to channel voids, Xiong *et al.* [2] proposed a scheduling algorithm called LAUC-VF (Latest Available Unused Channel with Void Filling). LAUC-VF keeps track of all void intervals within the channel space and assigns the interval with the latest starting time that is still earlier than the arrival time of the incoming burst. This has the effect of filling channel space more effectively, while ensuring that any newly created voids would occur closer to the present time and hence more capable of being filled by newly arriving bursts. A study by Lannoo *et al.* [6] demonstrated the burst loss arising from LAUC-VF in nodes with full wavelength conversion to be just slightly higher than the theoretical Erlang-B blocking rate. Although void filling improves channel utilization, it comes at the expense of high converter utilization. This is detrimental to performance in partial conversion schemes.

Li *et al.* [5] later introduced a different class of scheduling algorithms to be used exclusively at edge nodes. They sought to reduce the number of overlapping bursts arriving simultaneously at a link, thereby proactively minimizing contentions. The BORA-FS (Burst Overlap Reduction Algorithm with Fixed-Order Search) method of scheduling relies on electronic buffers at the ingress in order to delay overlapping bursts and transmit them in serial fashion. Buffering the bursts at the ingress minimizes the instantaneous load arriving at the first hop, and hence proactively prevents burst contentions. A variant of this algorithm, named BORA-DS (BORA with Destination-Order Search), serializes bursts that have similar destinations, ensuring fewer overlapping bursts along much of the downstream route.

Because all previously proposed scheduling mechanisms have assumed full wavelength conversion, none has attempted to minimize the number of conversions required, nor the number of converters required in each node.

We now proceed to describe our proposed scheduling methods, based on the concepts of: (a) minimizing redundant conversions, and (b) coordinating primary wavelength selection between OBS nodes. We show a net reduction in the number of converters required without significant degradation to burst loss performance.

3. Conversion minimization

The LAUC-VF scheduling algorithm is unsuited to the use of partial wavelength conversion. Because voids are randomly distributed in the channel space and the minimum void is likely to arise on any of the W outgoing wavelength channels, incoming bursts have probability $(W - 1) / W$ of requiring a wavelength converter. In LAUC-VF, wavelength conversion is therefore primarily used to redistribute traffic for the purpose of better channel utilization.

If we distinguish two causes of burst contention:

§ Blocking due to Wavelength Unavailability (BWU); and

§ Blocking due to Converter Unavailability (BCU);

then it can be shown that as a node's conversion ratio, r_c, is decreased, i.e. we decrease the ratio of wavelength converters to output channels, the majority of blocking arising from the LAUC-VF algorithm will result from converter unavailability. The overall burst loss rate will rise dramatically for even an incremental reduction in conversion ratio from *full* (i.e. r_c =1) to *partial* (i.e. $r_c < 1$).

3.1. Conversion avoidance scheduling

We propose therefore a scheduling algorithm which we call **LAUC-VF with Conversion Avoidance Scheduling** (**LAUC-VF-CAS**). LAUC-VF scheduling is applied only if the channel occupied by the incoming burst is not available on the outgoing link. Otherwise, the burst is scheduled on its current channel. From Figs. 2 and 3, LAUC-VF-CAS greatly reduces the percentage of bursts requiring conversion, even for high traffic loads, and that the blocking due to converter unavailability is sharply reduced. Additionally, blocking due to wavelength unavailability is kept nearly constant with LAUC-VF-CAS, and for the system shown, we can achieve similar blocking performance to full conversion using only 65-75% of the converters. This range of conversion ratios has previously been cited in a study in which a form of conversion avoidance scheduling was used [4].

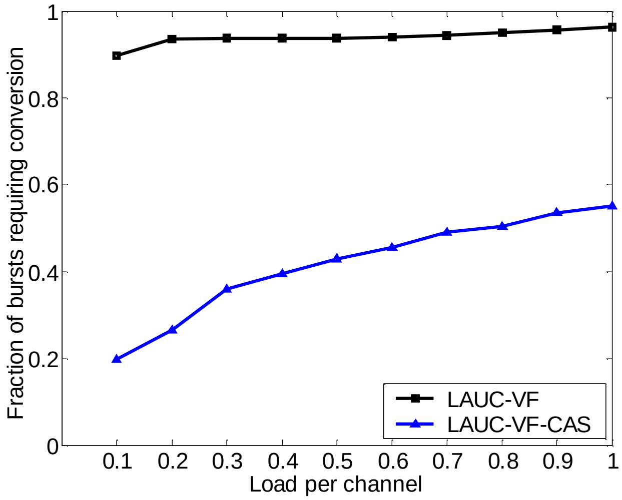

Figure 2. Fraction of bursts requiring conversion for various loads (W = 16 channels)

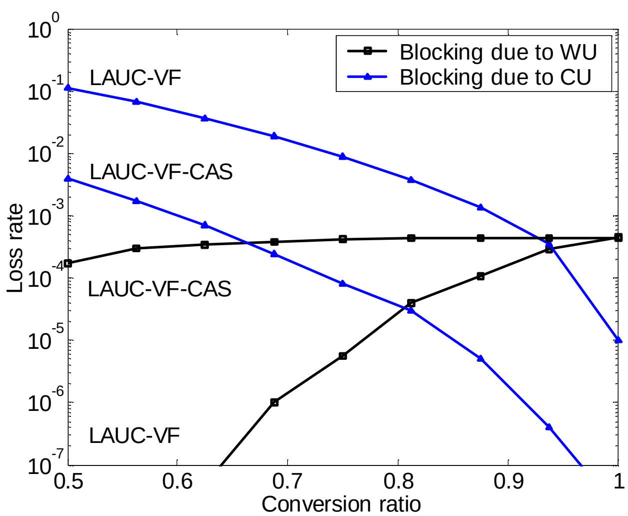

Figure 3. Blocking due to wavelength and converter unavailability (W = 16, load = 0.4 Erlang)

However, further investigation demonstrates the performance of LAUC-VF-CAS to be highly sensitive to the ratio of offset size to burst length. Larger offsets provide for more void filling possibility, and void filling can be extremely beneficial to the throughput of small bursts as these are more likely to fill voids and avoid bandwidth waste. Avoiding wavelength conversion entirely at the expense of void filling is therefore detrimental to throughput performance in this regime.

In Figs. 4a and 4b, burst loss probability is plotted as a function of conversion ratio for different offset-to-burst-length ratios. We observe that burst duration (L_b) need only be half the size of the mean offset time (T_{off}) to experience significant performance degradation. In this scenario, the use of LAUC-VF-CAS would generate nearly 3 times the blocking as LAUC-VF, even when full wavelength conversion is used. Conservatively estimating the typical network diameter as 5 and the per-node offset processing time as 10-20 µs, it is reasonable that the T_{off} /L_b ratio is as high as 4-5.

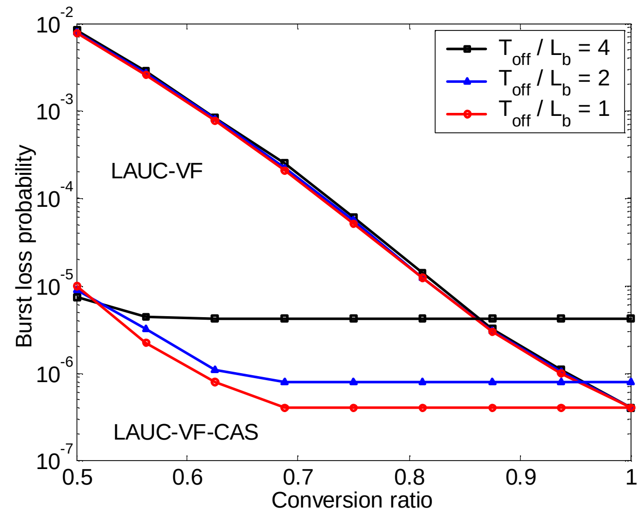

Figure 4a. Burst loss probability vs. Conversion ratio for varying burst lengths (W = 16, load = 0.2)

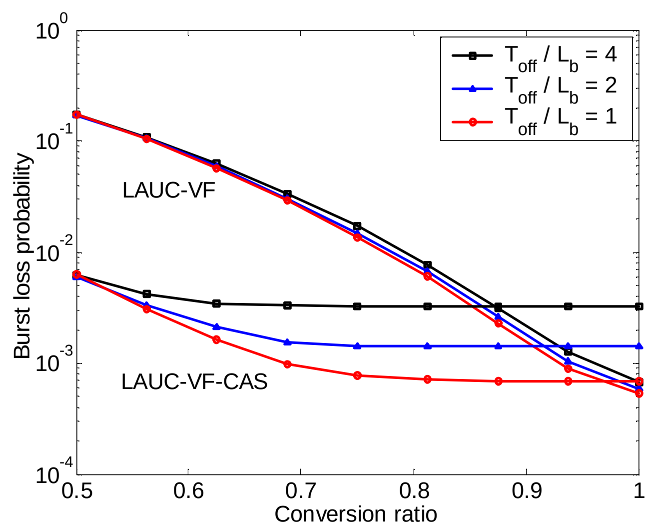

Figure 4b. Burst loss probability vs. Conversion ratio for varying burst lengths (W = 16, load = 0.4)

3.2. Conversion reduction scheduling

It was found in the preceding investigations that converter unavailability can account for the large majority of blocking in partial wavelength conversion regimes. A strict conversion avoidance policy is sub-optimal since the scheduler will avoid conversions even when there are free converters available; and because void-filling may help to improve channel utilization in these instances. Based on these conclusions, we propose a novel threshold-based scheduling algorithm called **LAUC-VF with Conversion Reduction Scheduling** (or **LAUC-VF-CRS**).

The algorithm works as follows. When a header packet arrives at the node to reserve resources for its associated burst, the scheduler determines the number of converters that are free during the burst's time interval. If this number exceeds a particular threshold value $N_{C\text{-}LIM}$, the scheduler uses normal LAUC-VF. However, below the threshold value, indicating that converter resources are scarce, the scheduler must conduct a further evaluation. If the burst's length is less than the k times the mean offset time, $T_{off\text{-}AVG}$, then the scheduler again resorts to LAUC-VF. Only otherwise does the scheduler use the LAUC-VF-CAS method in attempting to maintain the burst on its current channel and therefore avoid wavelength conversion. With appropriate choices of thresholds, $N_{C\text{-}LIM}$ and k, this ensures that conversion is avoided when there are scarce resources, and that void-filling is performed in instances where it would have high probability of improving channel utilization. We expect $N_{C\text{-}LIM}$ to be a function of the traffic load arriving at the node and for k to be related to the mean burst length.

Despite a number of soft decisions rendered by the scheduler, the order of complexity and processing speed of this algorithm does not differ greatly from LAUC-VF. LAUC-VF-CRS makes use the LAUC-VF algorithm the majority of the time. Additional complexity comes in form of finding the number of free wavelength converters, keeping a running average of the offset size, and comparing this to the time-occupancy (i.e. size) of the burst. The former is normally carried out by LAUC-VF in reserving a wavelength converter, and here amounts simply to storing an additional value in memory. The latter two are normally processed when the control packet is read, and thus only require additional memory storage and lookups. In cases where contention avoidance scheduling is used, the scheduler saves processing time that would otherwise be spent on searching the channel space for voids.

4. Wavelength assignment

After examining how to minimize redundant conversions within core nodes through a Conversion Reduction Scheduling algorithm, we will now examine a strategy for **Coordinated Wavelength Assignment** (**CWA**) for ingress traffic. By picking an initial wavelength strategically, in coordination with neighbouring ingress nodes, there is a higher probability of the burst avoiding conversions within the core of the network. The end result is that fewer converters are required within OBS nodes.

A formal graph colouring approach would yield optimal results for a large connected network. For the present work, however, we provide a simple heuristic to explain the wavelength assignment algorithm.

Assuming that identical wavelengths, numbered from W_1 to W_M, are used at each node, the algorithm works roughly as follows. For each core node v_i of the network, let $A_i = \{a_1, a_2, \ldots a_{Di}\}$ be the set of adjacent nodes, where D_i is the degree of node v_i. Assign bands of wavelengths of size $\lfloor M/D_i \rfloor$ to each adjacent node a_k. This band represents the preferred set of wavelengths to be used along the link between a_k and v_i. Thus, since traffic arriving at node v_i from its adjacent nodes will likely be filled on complementary bands of wavelengths, the large majority of bursts will be able to avoid conversion at the core node. Thus, converters remain free to perform contention resolution. The performance of this proactive scheduling algorithm is evaluated in the next section.

5. Simulation results

This section presents our results on the proposed Conversion Reduction Scheduling and Coordinated Wavelength Assignment algorithms that were earlier described. The major metrics used for evaluating performance are burst blocking probability and conversion ratio. A single node analysis was used in this simulation study in order to examine the conversion ratio required in a typical core node. The node connects traffic flowing from four ingress fibres onto one outgoing fibre, and there are 16 identical wavelengths on each.

The node's switching matrix is strictly non-blocking. Connections are blocked only due to unavailability of free wavelengths and/or wavelength converters. The node has a bank of tunable wavelength converters that can be shared among all wavelength channels, and which convert over the entire range of wavelengths.

Bandwidth reservation follows the standard OBS-JET strategy. Burst arrivals follow a Poisson process, and burst lengths are negative-exponentially distributed with a mean duration of 12.5 µs (i.e. 125 kb bursts on a 10 Gbps line). The maximum diameter of the network is taken to be 5, and offset times are thus variable in discrete sizes, with a per-hop processing time of 20 µs. Ingress bursts are randomly distributed across all channels, except for the study of Coordinated Wavelength Assignment.

5.1. Conversion reduction vs. conversion avoidance

To examine the performance of the newly proposed algorithm, LAUC-VF-CRS, we compare its burst loss probability at the single node with that of strict LAUC-VF-CAS. We conduct two sets of simulations, one examining the influence of the $N_{C\text{-}LIM}$ threshold and the other to examine the effect of the *k* value.

In both simulations, we use LAUC-VF to schedule bursts on the ingress links. The ratio of offset-time-to-burst length is $T_{off}/L_b = 4$, such that Conversion Avoidance Scheduling will have poor blocking probability performance. Fig. 5 shows the burst-blocking rate using Conversion Reduction Scheduling for different $N_{C\text{-}LIM}$ thresholds. The k-value was fixed at 0.5, so that even when few converters are free (below the threshold value), the scheduler will still use LAUC-VF if the burst time-occupancy is less than half the mean offset time.

From these results, we see that setting a low $N_{C\text{-}LIM}$ value (e.g. 1, 2) will result in low blocking near full

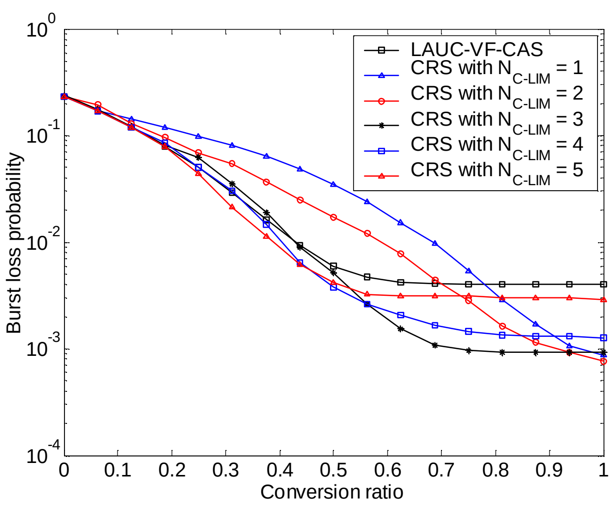

Figure 5. Loss rate of Conversion Reduction Scheduling for various converter threshold values (with parameters $k = 0.5$, $T_{off} / L_b = 4$, load = 0.4 Erlang), and a comparison with Conversion Avoidance Scheduling

wavelength conversion, but the loss will rise dramatically with decreasing conversion ratio. This is due to the fact that converters are not adequately being preserved. For higher $N_{C\text{-}LIM}$ values (e.g. 4, 5), not enough void filling is being performed, such that burst loss is higher for all values of conversion ratio. We determine an optimal value of $N_{C\text{-}LIM}$ at this load to be 3. We see that at this threshold value, the burst loss rate remains within 10% of the theoretical Erlang-B rate for a conversion ratio as low as 65%.

Different *k* values will affect the degree to which void-filling is performed, and thus to the extent the ***effective*** conversion ratio (lowest operating point for which burst block is no higher than 10% of the full wavelength conversion rate) can be reduced. For instance, in Fig. 6, we see that lowering the ***k***-value from 0.5 to 0.25 reduces the blocking probability floor, however the loss rate steeply climbs. Increasing the ***k***-value smoothes the burst loss for lowering conversion ratio, however the minimum loss rate achievable is much higher.

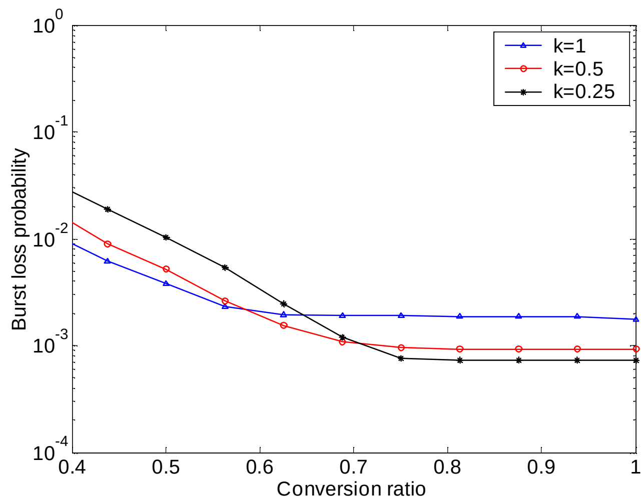

Figure 6. Loss rate of Conversion Reduction Scheduling for various *k* values

5.2. Coordinated Wavelength Assignment

We have also examined the degree to which a coordinated wavelength assignment for ingress traffic can reduce the number of converters required in an OBS node. The 16 channels were divided into bands of 4, and each one assigned to one of the incoming links as its preferred set of wavelengths. Fig. 7 demonstrates the burst loss performance on the outgoing link. We see that a 40% savings in converters is possible for a channel load of 0.2, and a 35% savings is possible for a load of 0.4. Future work consists of extending to study to a network-based simulation.

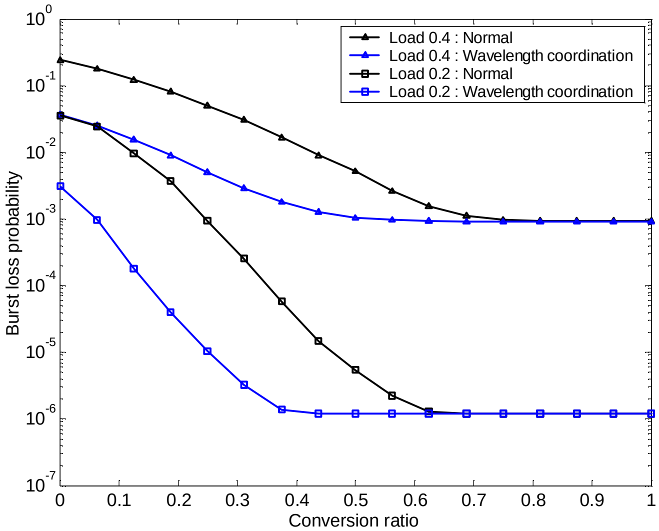

Figure 7. Comparison of Coordinated Wavelength Assignment scheduling to LAUC-VF-CRS scheduling

6. Conclusion

In this paper, we have examined the nature of burst contention arising in OBS nodes using partial wavelength conversion. We have shown that Conversion Avoidance scheduling can reduce the number of converters required, however that its performance is highly sensitive to burst lengths. We then presented a novel threshold-based Conversion Reduction scheduling algorithm, and found that it can provide more stable burst loss performance. Finally, we introduced a proactive scheduling method that relies on offline coordination of nodes in selecting a preferred set of wavelengths to route traffic. The loss rate downstream was shown to be significantly reduced, as bursts require a lesser degree of conversion.

7. References

[1] J. S. Turner, "Terabit burst switching progress (9/98-12/98)," in *Washington University at St. Louis Technical Report*, 1998.

[2] Y. Xiong, M. Vandenhoute, and H. C. Cankaya, "Control architecture in optical burst-switched WDM networks," *IEEE J. Sel. Areas Commun.* **18**, pp. 1838–1851, October 2000.

[3] Iizuka, *et al*, "A scheduling algorithm minimizing voids generated by arriving bursts in optical burst switched WDM network," in *Proceedings of IEEE GlobeCom 2002* (IEEE, Taipei, 2002).

[4] C. Gauger, "Performance of converter pools for contention resolution in optical burst switching," in *Proceedings of SPIE OptiComm 2002* (SPIE, Boston 2002).

5. J. Li and C. Qiao, "Schedule burst proactively for optical burst switching networks," in *Proceedings of IEEE Globecom 2003* (IEEE, San Francisco 2003).

6. B. Lannoo, *et al*, "A performance study of different OBS scheduler implementations," in *Proceedings of Symposium IEEE/LEOS Benelux* (IEEE/LEOS, Amsterdam 2002).