

Performance Improvement of an Internally Blocking Optical Packet/Burst Switch

Jan Cheyns, Erik Van Breusegem, Chris Develder, Ann Ackaert, Mario Pickavet, Piet Demeester

Department Information Technology

Ghent University-IMEC

Ghent, Belgium

jan.cheyns@intec.rug.ac.be

Abstract— Optical packet/burst switching is considered a promising technique to improve the performance of optical networks. Key components in these technologies are the optical switching nodes. Some of these node architectures suffer from internal blocking. Synchronous operation allows overcoming most of the problems introduced by this internal blocking. However, in asynchronous networks internal blocking can have a more pronounced effect. In this paper we propose a windowing technique to improve the performance of internally blocking optical switching nodes in asynchronous operation. Simulations will show that significant improvements can be made.

Keywords *Optical Packet Switching; delay window; internal blocking.*

I. INTRODUCTION

In the past, optical networking experienced a first great boom with the introduction of optical fiber and a second time when optical amplifiers (e.g. Erbium Doped Fiber Amplifiers, EDFA) allowed the deployment of Dense Wavelength Division Multiplexing (DWDM). In these optical networks, bandwidth is usually assigned (semi-)statically on a per wavelength basis. These circuit switched networks have to be well dimensioned [1] and good survivability [2] can be realized. However the coarse granularity of assigning bandwidth per wavelength results in poor bandwidth efficiency, especially in current networks where bursty data traffic is becoming predominant. Several techniques are now under study to improve the bandwidth utilization: Optical Packet Switching (OPS) [3], Optical Burst Switching (OBS) [4], Wavelength-Routed OBS [5], etc. Common factor in all of these is the limited wavelength occupation time, allowing to benefit from statistical multiplexing directly at the optical layer.

The driving force in the success of optical communication up to now was concentrated around the continuous increase of available bandwidth, in several ways: (a) hardware evolutions, such as fiber improvements (lower loss, lower dispersion...) and optical amplifiers in a broad wavelength range allow higher bit rates and a broader wavelength range; (b) the use of time and wavelength division multiplexing. Further evolution of optical networks will however need to focus on other parts of the network, the switching nodes. As bandwidth keeps increasing, the use of switches that convert the optical input signal to the electronic domain, perform the switching functionality electronically, and then convert this signal back to the optical domain, will soon become very complicated and

expensive. Furthermore it can be expected that evolution in the speed of electronics will not be able to keep up with the fast growing bandwidth availability. These problems are driving the migration to implementing switching functionality in the optical domain. This evolution can also be seen in circuit switched optical networks [6], where its benefits can increase survivability [2]. For circuit switching as well as for packet/burst switching optically transparent switches are preferred, allowing smooth adaptation to any future changes in bit rate, data format, and so on. Clearly, circuit and packet switched networks also show important differences in the requirements for switching nodes. The most important issue is the switching time. In a circuit switched network a wavelength path is set up for a relatively long period, so a slow switch can suffice. However, if the holding time of the wavelength decreases, this switching overhead time leads to increased inefficiencies. Hence fast switches are required to make packet switching viable.

Key component in these fast optical switches is the switching matrix, where optical signals are routed from a certain input port to the destination output port. Doing this over the entire path keeps the payload in the optical domain from source to destination. Several architectures of these switching matrices are possible [7]. Some of these have the drawback of being internally blocking, limiting the throughput, and thus increasing the packet loss in such a node. The impact of this internal blocking can be limited in synchronous operation by using the knowledge of all simultaneously active input ports. We will shortly elaborate on this in Section III. However it is not possible to make a straightforward translation of this solution to the unsynchronized case. In Section IV we will describe a method that allows to partially circumvent the negative influences of the internal blocking, the proposed windowing technique.

II. NODE ARCHITECTURE

A. Generic description

An OPS/OBS node has a generic architecture as depicted in Figure 1. First of all the node structure considered here is hybrid. The payload is switched transparently in the optical domain, without any opto-electronic conversion along the path from source to destination, thus keeping the payload optical end-to-end. As optical processing is still in its infancy, the

control information contained in the headers is optoelectronically converted to allow electronic processing.

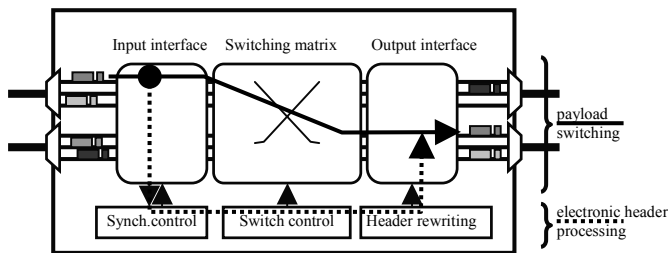


Figure 1. A generic representation of an optical packet/burst switch.

When a packet/burst arrives at the node, its header is extracted, while the payload is delayed in a fixed length fiber delay line. An alternative approach is to separate header and payload in time by using an offset [4], having some repercussions on node complexity [8]. This delay is needed in order to allow the processing of the header. The information contained in this header is then used to set up the switching matrix to switch the packet/burst to the correct output port. At the output interface of the switch a new (possibly modified) header is written for the outgoing packet.

Generally speaking, a network can be operated in two ways, either synchronized or asynchronous. In synchronized operation, the switch's input-output connections are reconfigured at discrete points in time. To this end, packets/bursts at different input ports should all arrive at the same time. Therefore a synchronization stage is necessary at the input interface when synchronized operation is chosen. In asynchronous operation however, packets can arrive at any time.

B. Internal blocking

The switching matrix is the core of the switching node, and several implementations are available [7]. In any switch contention occurs when two or more packets/burst want to use the same output port at the same time. Clearly only one of them can be sent, resulting in loss of the other(s) (in a bufferless switch). A node is said to be internally blocking if loss occurs even though the traffic does not show contention, usually caused by the physical properties of the switching device.

We will clarify this with an example. Consider the node structure of Figure 2. The core of this node is an Arrayed Waveguide Grating (AWG): a passive optical device where, when light enters at an input port, the output port where it will exit from depends on the wavelength used. Exploiting this feature with fast Tunable Wavelength Converters (TWC) at the input ports, it is possible to build a fast operating transparent optical packet/burst switch.

A detailed study of this architecture can be found in [9]. Depending on the input port where the signal enters, the incoming wavelength will have to be converted to a certain wavelength in order to be routed to the correct output fiber. For output fiber 0, the conversions are shown in TABLE I. From this (very simple) example, it is clear that the node is internally blocking, as a packet arriving at input port 4 will never be able

to reach output fiber 0, even if it is the only packet arriving (no contention). Now suppose packets come in at input ports 2 and 3. None of these packets gets lost if input port 2 is converted to λ_1 , and 3 to λ_0 . However, if we would have chosen to convert input port 2 to λ_0 this would lead to loss of the packet at input port 3. From this we can conclude that is important to make a good choice for the wavelength conversion. If it is possible to configure the switch in such a way that no loss occurs for any combination of active input ports (of course excluding contention), we call the switch rearrangeable nonblocking.

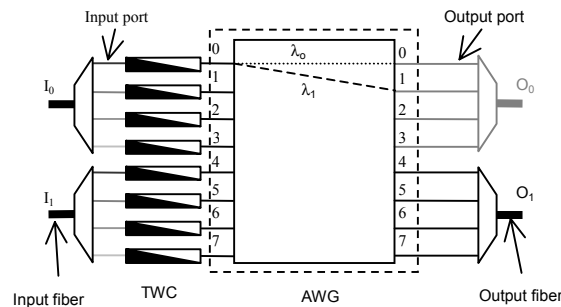


Figure 2. Example of an internally blocking node architecture, based on a single AWG and TWCs.

TABLE I. BEHAVIOR OF THE AWG-BASED SWITCH

		Input ports							
		0	1	2	3	4	5	6	7
Output ports	0	λ_0					λ_3	λ_2	λ_1
	1	λ_1	λ_0					λ_3	λ_2
	2	λ_2	λ_1	λ_0					λ_3
	3	λ_3	λ_2	λ_1	λ_0				

Note that although the architecture of Figure 2. is not completely rearrangeable nonblocking, it is still possible to make good configuration choices to improve throughput, as was shown by the example. This property will be exploited in synchronized operation. It will also form the basic condition for the windowed technique to function properly. It is also worth mentioning that making the correct choice on which output ports of the AWG to combine within an output fibre can optimize this node structure [9].

The TWC's used have an output wavelength range of W . Using a range of FW would allow to build a nonblocking switch, but this would come at the cost of extra wavelength converters at the output ports, in order to have a.o. correct operation at the next node, where these W wavelengths (and no others) are expected at the input. Also using a range of FW would scale poorly to nodes with high fiber and/or wavelength count.

III. SYNCHRONIZED OPERATION

Essential in synchronized operation is that all packets arrive at the same time. In this way the configuration of the switching matrix can be done based on information of all simultaneously

active input ports. We will call this condition “total input information availability” from now on. Like this, for synchronized operation, an internally completely rearrangeable node will have the same loss as an internally nonblocking switch: no loss will occur as long there is no contention.

Even when a node is not completely internally rearrangeable, synchronized operation allows for better performance. Again due to the fact that we know which input ports are active, the effect of internal blocking can be mostly alleviated. We repeat the stated example from above, having input ports 2 and 3 active. Since we know these input ports are active simultaneously we can decide to convert input port 2 to λ_1 , keeping λ_0 available for input port 3, so it will not get lost. We demonstrated in [9] that it is indeed possible to almost completely overcome the internal blocking, by using an algorithm based on Maximum Matching in bipartite graphs, for an architecture as the one shown in Figure 2.

IV. ASYNCHRONOUS OPERATION

A. Problem formulation

Not synchronizing in the node reduces the hardware complexity of the node. Of course this comes at a cost. The requirements for the switching matrix are now more demanding, since now rearrangeable nonblocking property does not suffice and strictly nonblocking is needed. This can be clearly seen on Figure 3. If there are already $N-1$ packets going out of output fiber O_0 , there still has to be room for an N^{th} packet to be switched to this output fiber, without changing the current configuration of the switching matrix. This condition should hold for any possible active input port configuration.

The main difference with synchronous operation is that we do not have total input information available. Since it is impossible to know which input ports will become active in the (near) future, we cannot predict which channels will overlap in time, and might contend. Consequently, no intelligent decision on the “best” configuration of the switching matrix can be made. This is the major reason for the poor behavior of internally blocking switching matrices when deployed in asynchronous regime. Again we go back to the example above. If input port 2 would become active at time t , a decision has to be made on which conversion to choose. Suppose input port 3 becomes active at $t+\tau$, when input port 2 is also still active. If the decision, made at time t , was to convert input port 2 to λ_0 , the packet subsequently arriving at input port 3 will be lost.

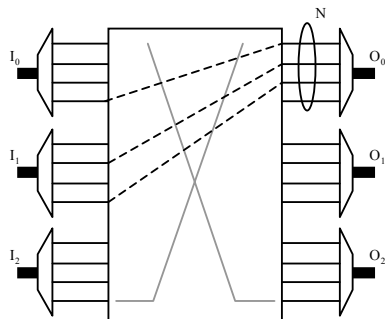


Figure 3. Generic node structure.

B. Proposed Solution

As discussed above, in order to improve the behavior of an internally blocking node, it is essential to have knowledge of which ports will be active at the same moment in time in order to be able to make a good decision for the configuration of the switching matrix. The availability of this total input information in synchronized operation allows largely decreased losses.

The key to improving loss characteristics in asynchronous operation is that we need to find a way to make more input information available when decisions have to be made. In fact what we need is a way to know which input ports will become active in the future. A possible way to realize this is enlarging the length of the fixed fiber delay line, that is present to compensate for the processing time of the opto-electronically converted header, by a length L (corresponding with an extra time delay $\delta=L/c$, with c the speed of light in the fiber). By doing this, the configuration decision can be postponed until more information is available on which other input ports become active in the future, since the headers of these future packets will already have arrived. This principle is further clarified in Figure 4.

Suppose the switch has to be configured correctly for packet/burst A. At that time however, due to the extra delay of δ , the headers of the packets B, C, D and E will have entered the switch controller too, thus we know that (and when) these ports will become active. In this case, we know that (the payload of) A, C and E overlap in time. This information allows for an intelligent decision for the switching matrix configuration for packet A, minimizing the impact on the loss probability of C and E. The same is done when packet C needs to be switched through the node, this process is repeated and goes on continuously for all packets, which is why we name δ the decision window duration. In short: for each packet we use future usage information on several input channels to make a decision for that packet.

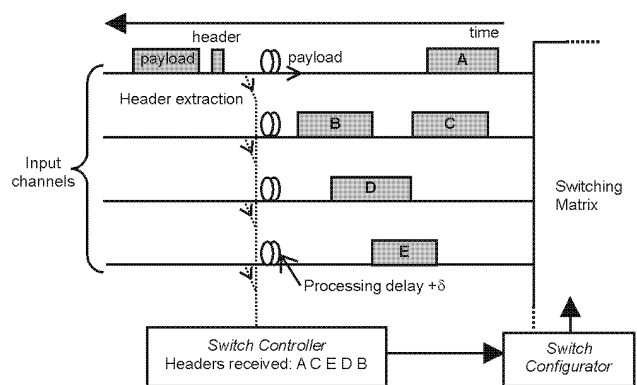


Figure 4. An example showing how the windowing technique works.

V. SIMULATION RESULTS

We will now show some results for asynchronous operation of the node architecture described in Figure 2. Although the following results are obtained on this specific architecture we

would like to stress the general applicability of the windowed technique for internally blocking switching nodes. The only condition needed is the existence of some algorithm to improve the blocking behavior when information is available on simultaneously active input ports (as the Maximum Matching based algorithm did for the described architecture).

We simulated a packet switch with three incoming and three outgoing fibers, each fiber carrying 9 wavelengths. The applied traffic pattern was Poisson, generated using a high quality pseudo-random generator, based on a Mersenne Twister [10], implemented in the COLT library for Java [11]. Simulation was stopped based on the size of the 95% confidence interval compared to the estimate of the mean value. When this ratio was smaller than 0.05, simulation was halted. Error flags are omitted on the figures in order to make them more readable. All of this allows us to have a high-quality reliable simulation [12]. Two distinct cases were studied, one with fixed packet sizes, described in subsection A, and another one in subsection B with burst sizes that had a negative exponential distribution.

A. Fixed Size Packets

The results of the fixed size packet simulations for an asynchronous switch are shown in Figure 5. The number in the legend denotes δ/L , this is the window delay time δ expressed in multiples of the fixed packet size L . Also indicated is the behavior when no windowing is used (none), and the performance of a perfect internally nonblocking switch as a way to assess the quality of the improvement obtained by the windowing technique.

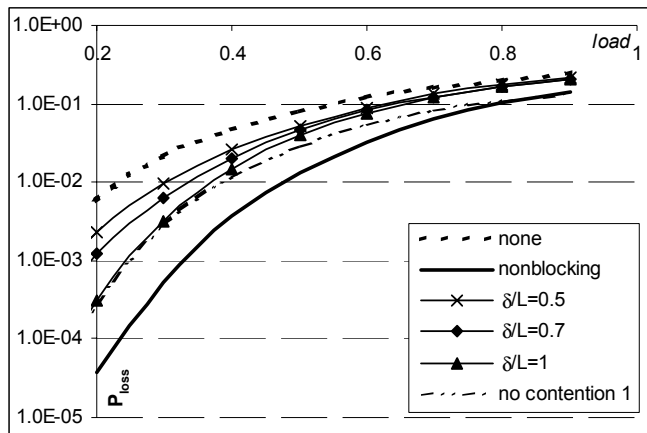


Figure 5. Simulation results for fixed packet size.

We see that introducing the windowed technique decreases loss significantly. It is clearly shown that the larger δ/L is, the higher the performance of the windowing technique becomes. However values higher than 1 have no or little use. Since the packet size is fixed, the packet for which a configuration decision has to be made can only possibly overlap with packets arriving L later. Another feature that is clear from the result is the better behavior of the windowed approach in regions of lower load, while the improvement is smaller when load is high. This is a quite logical evolution, as for higher loads, the

most important cause of loss is contention. It is only at low loads that the internal blocking is dominating the loss probability, since at these low loads the probability of contention is very small. This is again shown by the “no contention 1” curve, which is the result of a simulation with $\delta/L=1$ and where traffic applied to the input of the switch was contention free. We see that for low loads the curve coincides with the curve for traffic with contention: the case with contention has a small probability of contention, so it coincides with the “no contention” case. When the load increases, we see that the behavior of “no contention” is getting better than the case with contending traffic, showing that in this region contention is the most important cause of loss. It is impossible to avoid the contention losses by using a better configuration algorithm or increasing the window size; other solutions such as deflection routing, optical buffering... are needed to resolve these problems [13]. We do point out that for internally blocking nodes, the better the performance without any of these extra contention resolution mechanisms, the more effective these contention resolution features will be in reducing the loss due to contention. If the node suffers from severe internal blocking, the resource originally meant to resolve contention will in the first place be used up to compensate for this internal loss, leaving them to be less powerful for solving contention.

B. Variable size packets

We will now discuss the results for the simulation where packet length has a negative exponential distribution, with mean length L_m . We perform simulations for different values of δ/L_m .

The results for variable packet length are shown in Figure 6. They have some similarities to those of fixed size packet length, but also some important differences. As was the case for fixed size, there is an improvement by using the windowed technique, again the lower loss rate is more pronounced at lower loads, for the same reasons. We do see that with variable packet lengths the improvement is not that high as it was for fixed size packets. This is due to the higher overlap probability using variable lengths. Now it also makes sense to have δ/L_m values higher than 1. As bursts can be longer than L_m , overlap can occur on larger timescales, however there is a value of δ after which there is no longer a significant improvement visible, since it covers a very large portion of the overlap probability. Burst size distributions with an upper limit on the burst size are therefore interesting, since it allows to be certain of the largest possible overlap interval.

Using variable packet sizes thus gives slightly worse loss performance, and a higher δ value is needed to achieve significant performance improvement in terms of PLR. Clearly, this implies that more fiber delay should be implemented, delay is higher and of course memory and processing requirements in the electronic switch controller will be higher. Therefore we can conclude that the windowed technique is preferably to be used with fixed packet sizes, or at least packet size distributions with a maximum.

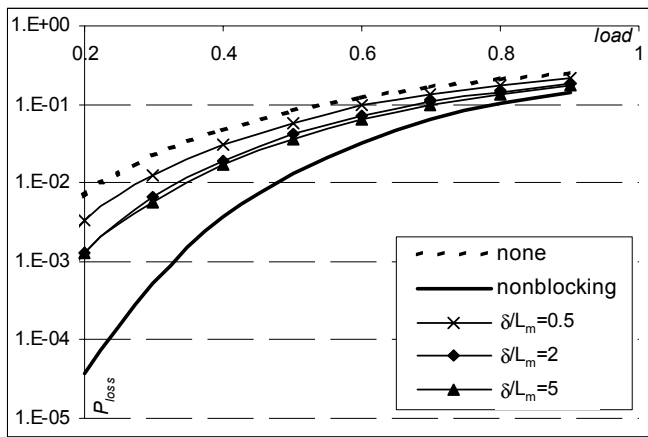


Figure 6. Simulation results for negative exponentially distributed packet size.

VI. CONCLUSION

In this paper we studied the logical performance of an OPS/OBS switch with internal blocking, both in synchronous and asynchronous operation. In the synchronized case, the complete knowledge of which input ports are simultaneously active can be exploited to considerably reduce the traffic loss. Based on these observations, we developed a technique to improve the behavior in asynchronous operation. The proposed windowing technique makes information available on future active input ports, allowing lower losses, due to better switching matrix configurations. However as opposed to the synchronous case (where the effect of internal blocking could be ruled out almost completely) the loss improvement is not as large. This shows internal blocking has a higher impact in asynchronous operation, even with the windowed technique. Using fixed size packets allows better performance of the windowing technique compared to using variable packet lengths.

ACKNOWLEDGMENT

This work was partly funded by the IST project STOLAS, under contract number IST-2000-28557 and the IWT-GBOU project 010058 "Optical Networking and Node Architectures."

Both Jan Cheyns and Chris Devellder are Research Assistants of the Fund for Scientific Research – Flanders

(Belgium) (F.W.O. – Vlaanderen). Erik Van Breusegem would like to thank the IWT for its financial support.

REFERENCES

- [1] N. Wauters, P. Demeester, "Design of the optical path layer in multiwavelength cross-connected networks," *IEEE J. on Sel. Areas in Comm.*, Vol. 14, no. 5, pp. 881–892, June 1996.
- [2] D. Colle, S. De Maesschalck, C. Devellder, P. Van Heuven, A. Groebbens, J. Cheyns, I. Lievens, M. Pickavet, P. Lagasse, P. Demeester, "Data-centric optical networks and their survivability," *IEEE J. on Sel. Areas in Comm.*, Vol. 20, no. 1, pp. 6–20, January 2002.
- [3] P. Gambini, M. Renaud, C. Guillemot, F. Callegati, I. Andonovic, B. Bostica, D. Chiaroni, G. Corazza, S. L. Danielsen, P. Gravey, P. B. Hansen, M. Henry, C. Janz, A. Kloch, R. Krahenbuhl, C. Raffaelli, M. Schilling, A. Talneau, L. Zucchelli, "Transparent optical packet switching: network architecture and demonstrators in the KEOPS project," *IEEE J. on Sel. Areas in Comm.*, Vol. 16, no. 7, pp. 1245–1259, September 1998.
- [4] C. Qiao, M. Yoo, "Optical burst switching (OBS) - A new paradigm for an optical internet," *J. High Speed Networks on WDM Networks*, Vol. 8, no. 1, pp. 69–84, January 1999.
- [5] M. Düser, P. Bayvel, "Performance of a dynamically wavelength-routed optical burst switched network," *IEEE Phot. Techn. Lett.*, Vol. 14, no. 2, pp. 239–241, February 2002.
- [6] C. R. Giles, V. Aksyuk, B. Barber, R. Ruel, L. Stulz, D. Bishop, "A silicon MEMS optical switch attenuator and its use in lightwave subsystems," *IEEE J. of Sel. Topics in Quantum Electronics* Vol. 5, no. 1, pp. 18–25, January–February 1999.
- [7] C. Devellder, J. Cheyns, E. Van Breusegem, E. Baert, A. Ackaert, M. Pickavet, P. Demeester, "Node architectures for optical packet and burst switching," (COIN + PS 2002, Cheju Island, Korea), (invited) paper PS.WeA1, July 2002.
- [8] E. Van Breusegem, J. Cheyns, B. Lannoo, M. Pickavet, P. Demeester, "Implications of using offsets in alloptical packet switched networks," *ONDM 2003, Budapest, Hungary*, February 2003.
- [9] J. Cheyns, C. Devellder, E. Van Breusegem, A. Ackaert, M. Pickavet, P. Demeester, "Routing in an AWG based optical packet switch," *Phot. Netw. Comm.*, Vol. 5, 1, January 2003, pp. 69–80.
- [10] M. Matsumoto, T. Nishimura, "Mersenne twister: a 623-dimensionally equidistributed Uniform pseudo-random number generator," *ACM Trans. Modeling and Comp. Sim.*, Vol. 8, no. 1, pp. 3–30, January 1998.
- [11] An implementation of the Mersenne Twister, colt distribution library, available online at <http://tildie-hoschek.home.cern.ch/~hoschek/colt/>.
- [12] K. Pawlikowski, H.-D. J. Jeong, J.-S. R. Lee, "On credibility of simulation studies of telecommunication networks," *IEEE Comm. Mag.*, Vol. 40, no. 1, pp. 132–139, January 2002.
- [13] S. Yao, B. Mukherjee, S. J. Ben Yoo, S. Dixit, "All-optical packet-switched networks: a study of contention-resolution schemes in an irregular mesh network with variable-sized packets," *Proc. SPIE* Vol. 4233, (OPTICOMM 2000, Plano, TX.), October 2000.