Optical Packet Switching for IP-over-WDM Transport Networks

Stefano Bregni, Giacomo Guerra, Achille Pattavina

Department of Electronics and Information, Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milan, Italy
bregni@elet.polimi.it, guerra@cerbero.elet.polimi.it,
pattavin@elet.polimi.it

Abstract. As new bandwidth-hungry IP services are demanding more and more capacity, transport networks are evolving to provide a reconfigurable optical layer in order to allow fast dynamic allocation of WDM channels. To achieve this goal, optical packet-switched systems seem to be strong candidates as they allow a high degree of statistical resource sharing, which leads to an efficient bandwidth utilization. In this work, we propose an architecture for optical packet-switched transport networks, together with an innovative switching node structure based on the concept of per-packet wavelength routing. Some simulations results of node operation are also presented. In these simulations, the node performance was tested under three different traffic patterns.

1 Introduction

Telecomunication networks are currently experiencing a dramatic increase in demand for capacity, driven by new bandwidth-hungry IP services. This will lead to an explosion of the number of wavelengths per fiber, that can't be easily handled with conventional electronic switches. To face this challenge, networks are evolving to provide a reconfigurable optical layer, which can help to relieve potential capacity bottlenecks of electronic-switched networks, and to efficiently manage the huge bandwidth made available by the deployment of dense wavelength division multiplexing (DWDM) systems.

As current applications of WDM focus on a relatively static usage of single wavelength channels, many works have been carried out in order to study how to achieve switching of signals directly in the optical domain, in a way that allows fast dynamic allocation of WDM channels, to improve transport network performance. Two main alternative strategies have been proposed to reach this purpose: optical packet switching [1], [2] and optical burst switching [3], [4].

In this article, we first introduce optical packet and burst switching approaches. Then an architecture for packet-switched WDM transport networks and a novel optical switching node are proposed. Some simulation results of node operation, under different traffic patterns, are also presented.

2 Optical Packet and Burst Switching

Optical packet switching allows to exploit single wavelength channels as shared resources, with the use of statistical multiplexing of traffic flows, helping to efficiently manage the huge bandwidth of WDM systems. Several approaches have been proposed to this aim [5], [6].

Most proposed systems carry out header processing and routing functions electronically, while the switching of optical packet payloads takes place directly in the optical domain. This eliminates the need for many optical-electrical-optical conversions, which call for the deployment of expensive opto-electronic components, even though most of the optical components, needed to achieve optical packet switching, still remain too crude for commercial availment.

Optical burst switching aims at overcoming these technological limitations. The basic units of data transmitted are bursts, made up of multiple packets, which are sent after control packets, carrying routing information, whose task is to reserve the necessary resources on the intermediate nodes of the transport network (see Fig. 1). This results in a lower average processing and synchroniza-

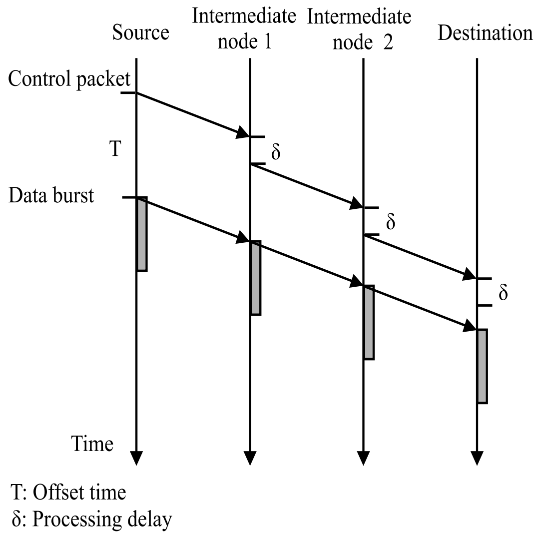

Fig. 1. The use of an offset time in optical burst switching

tion overhead than optical packet switching, since packet-by-packet operation is not required. However packet switching has a higher degree of statistical resource sharing, which leads to a more efficient bandwidth utilization in a bursty, IP-like, traffic environment.

Since optical packet-switching systems still face some technological hurdles, the existing transport networks will probably evolve through the intermediate step of burst-switching systems, which represent a balance between circuit and packet switching, making of the latter alternative a longer term strategy for network evolution.

In this work, we have focused our attention on optical packet switching, since it offers greater flexibility than the other relatively coarse-grained WDM techniques, aiming at efficient system bandwidth management.

3 Optical Transport Network Architecture

The architecture of the optical transport network we propose consists of $M = 2^m$ *optical packet-switching nodes*, each denoted by an optical address made of $m = \log_2 M$ bits, which are linked together in a mesh-like topology. A number of *edge systems* (ES) interfaces the optical transport network with IP legacy (electronic) networks (see Fig. 2).

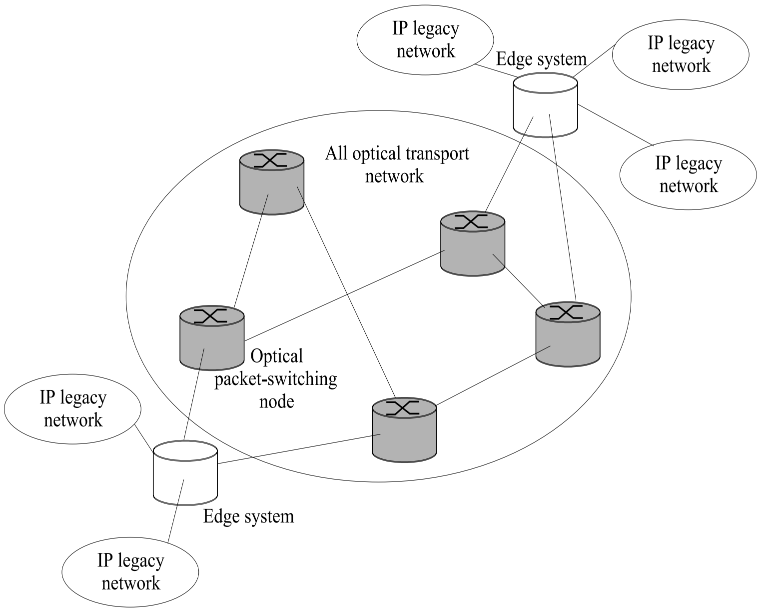

Fig. 2. The optical transport network architecture

An ES receives packets from different electronic networks and performs traffic aggregation in order to build *optical packets*. The optical packet is composed of a simple optical header, which comprises the m-bits long destination address, and of an optical payload made of a single IP packet, or, alternatively, of an aggregate of IP packets.

The optical packets are then buffered and routed through the optical transport network to reach their destination ES, which delivers the traffic it receives to its destination electronic networks.

At each intermediate node, in the transport network, packet headers are received and electronically processed, in order to provide routing information to the control electronics, which will properly configure the node's resources to switch packet payloads directly in the optical domain.

The transport network operation is asynchronous; that is, packets can be received by nodes at any instant, with no time alignment. The internal operation of the optical nodes, on the other hand, is synchronous (slotted). In the model we propose, the time slot duration, T, is equal to the amount of time needed to transmit an optical packet, with a 40-bytes long payload, from an input WDM channel to an output WDM channel.

The operation of the optical nodes is slotted since the behavior of packets, in an unslotted node, is less regulated and more unpredictable, resulting in a larger contention probability.

A contention occurs every time that two or more packets are trying to leave a switch from the same output port. How contentions are resolved has a great influence on network performance. Three main schemes are generally used to resolve contention: wavelength conversion, optical buffering and deflection routing.

In a switch node applying *wavelength conversion*, two packets trying to leave the switch from the same output port are both transmitted at the same time but on different wavelengths. Thus, if necessary, one of them is wavelength converted to avoid collision. In the *optical buffering* approach, one or more conteding packets are sent to fixed-length fiber delay lines, in order to reach the desired output port only after a fixed amount of time, when no contention will occur. Finally, in the *deflection routing* approach, contention is resolved by routing only one of the contending packets along the desired link, while the other ones are forwarded on paths which may lead to longer than minimum-distance routing paths.

Implementing optical buffering gives good network performance, but involves a great amount of hardware and electronic control. On the other hand, deflection routing is easier to implement than optical buffering, but network performance is reduced since a portion of network capacity is taken up by deflected packets.

In the all-optical network proposed, in order to reduce complexity while aiming at attaining good network performance, the problem of contention is resolved combining a small amount of optical buffering with wavelength conversion and deflection routing. Our policy can be summarized as follows:

1. When a contention occurs, the system first tries to transmit the conflicting packets on different wavelengths.
2. If all of the wavelengths of the correct output link are busy at the time the contention occurs, some packets are scheduled for transmission in a second time, and are forwarded to the fiber delay lines.

3. Finally, if no suitable delay line is available, at the time the contention occurs, for transmission on the correct output port, a conflicting packet can be deflected to a different output port than the correct one.

4 Node Architecture

The general architecture of a network node is shown in Fig. 3. It consists of N incoming fibers with W wavelengths per fiber. The incoming fiber signals

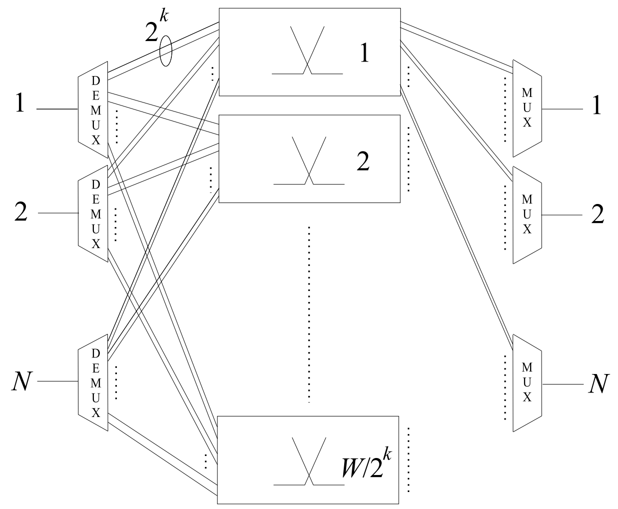

Fig. 3. Optical packet-switching node architecture

are demultiplexed and 2^k wavelengths, from each input fiber, are then fed into one of the $W/2^k$ switching planes, which constitute the switching fabric's core. Once signals have been switched in one of the second stage parallel planes, packets can reach every output port on one of the 2^k wavelengths that are directed to each output fiber. This allows the use of wavelength conversion for contention resolution, since 2^k packets can be contemporarely be transmitted, by each second-stage plane, on the same output link.

The detailed structure of one of the $W/2^k$ parallel switching planes is shown in Fig. 4. Each incoming link carries a single wavelength and the switching plain consists of three main blocks: an input *synchronization unit*, as the node is slotted and incoming packets need to be aligned, a *fiber delay lines unit*, used to store packets for contention resolution, and a *switching matrix unit*, to achieve the switching of signals.

These three blocks are all managed by an *electronic control unit* which carries out the following tasks:

– optical packet header recovery and processing;
– managing the synchronization unit in order to properly set the correct path through the synchronizer for each incoming packet;
– managing the tunable wavelength converters in order to properly delay and route incoming packets.

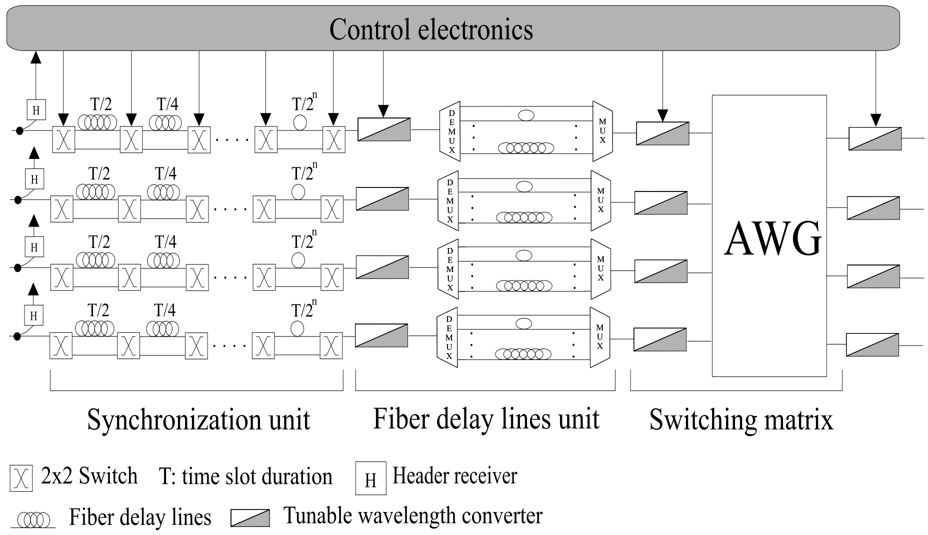

Fig. 4. Detailed structure of one of the $W/2^k$ parallel switching planes

We will now describe the second-stage switching planes mentioned above, detailing their implementation.

4.1 Synchronization Unit

This unit consists of a series of 2×2 optical switches interconnected by fiber delay lines of different lengths. These are arranged in a way that, depending on the particular path set through the switches, the packet can be delayed of a variable amount of time, ranging between $\triangle t_{min} = 0$ and $\triangle t_{max} = (1 - (1/2)^n) \times T$, with a resolution of $T/2^n$, where T is the time slot duration and n the number of delay lines.

The synchronization is achieved as follows: once the packet header has been recognized and packet delineation has been carried out, the packet start time is identified and the control electronics can calculate the necessary delay and configure the correct path of the packet through the synchronizer.

Due to the fast reconfiguration speed needed, fast 2×2 switching devices, such as 2×2 semiconductor optical amplifier (SOA) switches [7], which have a switching time in the nanosecond range, must be used.

4.2 Fiber Delay Lines Unit

After packet alignment has been carried out, the routing information carried by the packet header allows the control electronics to properly configure a set of tunable wavelength converters, in order to deliver each packet to the correct delay line to resolve contentions. To achieve wavelength conversion several devices are available [8], [9], [10].

Depending on the managing algorithm used by control electronics, the fiber delay lines stage can be used as an *optical scheduler* or as an *optical first-in-first-out* (FIFO) *buffer.*

- *Optical scheduling*: this policy uses the delay lines in order to schedule the transmission of the maximum number of packets onto the correct output link. This implies that an optical packet P_1, entering the node at time t_1 from the i-th WDM input channel, can be transmitted after an optical packet P_2, entering the node on the same input channel at time t_2, being $t_2 > t_1$. For example, suppose that packet P_1, of duration $l_1 T$, must be delayed of d_1 time slots, in order to be transmitted onto the correct output port. This packet will then leave the optical scheduler at time t_{1+d_1}. So, if packet P_2, of duration $l_2 T$, has to be delayed for d_2 slots, it can be transmitted before P_1 if $t_{2+d_2+l_2} < t_{1+d_1}$ since no collision will occur at the scheduler output.
- *Optical FIFO buffering*: in the optical FIFO buffer the order of the packets entering the fiber delay lines stage must be maintained. This leads to a simpler managing algorithm than the one used for the optical scheduling policy, yielding, however, a sub-optimal output channel utilization. In fact, suppose that optical packet P_1, entering the FIFO buffer at time t_1, must be delayed for d_1 time slots. This implies that packet P_2, behind packet P_1, must be delayed of, at least, d_1 time slots, in order to maintain the order of incoming packets. Due to this rule, if packet P_2 has to be delayed for $d_2 < d_1$ slots, in order to avoid conflict, its destination output port is idle for $d_1 - d_2$ time slots, while there would be a packet to transmit.

4.3 Switching Matrix Unit

Once packets have crossed the fiber delay lines unit, they enter the switching matrix stage in order to be routed to the desired output port. This is achieved using a set of tunable wavelength converters combined with an arrayed waveguide grating (AWG) wavelength router [11].

This device consits of two slab star couplers, interconnected by an array of waveguides. Each grating waveguide has a precise path difference with respect to its neighbours, $\triangle X$, and is characterized by a refractive index of value n_w.

Once a signal enters the AWG from an incoming fiber, the input star coupler divides the power among all waveguides in the grating array. As a consequence of the difference of the guides lengths, light travelling through each waveguide emerges with a different phase delay given by:

$$\triangle \Phi = 2\pi n_w \times \frac{\triangle X}{\lambda} \quad (1)$$

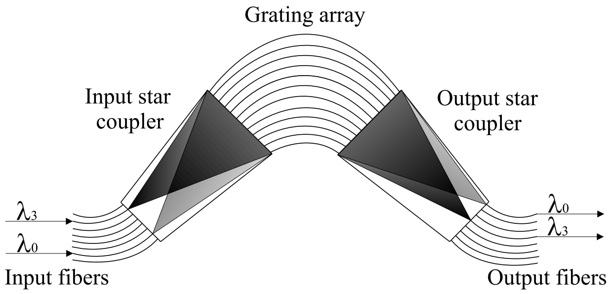

Fig. 5. Arrayed waveguide grating

being λ the incoming signal central wavelength. As all the beams emerge from the grating array they interfere constructively onto the focal point in the output star coupler, in a way that allows to couple an interference maximum with a particular output fiber , dependig only on the input signal central wavelength.

Figure 5 shows the mechanism described above. Two signals of wavelength λ_0 and λ_3 entering an 8×8 AWG, from input fibers number 6 and number 1 respectively, are correctly switched onto the output fibers number 0 and number 3, being the wavelength of signals the only routing information needed to achieve the required permutation.

The AWG is used as it gives better performance than a normal space switch interconnection network, as far as insertion losses are concerned. This is due to the high insertion losses of all the high-speed all-optical switching fabrics available at the moment, that could be used to build a space switch interconnection network. Moreover AWG routers are strictly non-blocking and offer high wavelength selectivity.

After crossing the three stages previously described, packets undergo a final wavelength conversion, to avoid collisions at the output multiplexers, where W WDM channels are multiplexed on each output link.

5 Simulation Results

In this section, we present some simulation results of the operation of one among the $W/2^k$ parallel switching planes, which structure has been shown in Fig. 4.

These results have been obtained assuming that the node receives its input traffic directly from N edge systems. The edge systems buffers capacity is supposed to be large enough to make packet loss negligible. Each WDM channel is supposed to have a dedicated buffer in the edge system.

The packet arrival process has been modeled as a Poisson process, with packet interarrival times having a negative exponential distribution. As the node operation is slotted, the packets duration was always assumed to be multiple of the time slot duration T, which is equal to the amount of time needed to transmit an optical packet, with a 40-bytes long payload, from an input WDM channel to an output WDM channel.

As far as packet length is concerned, the following probability distributions were considered:

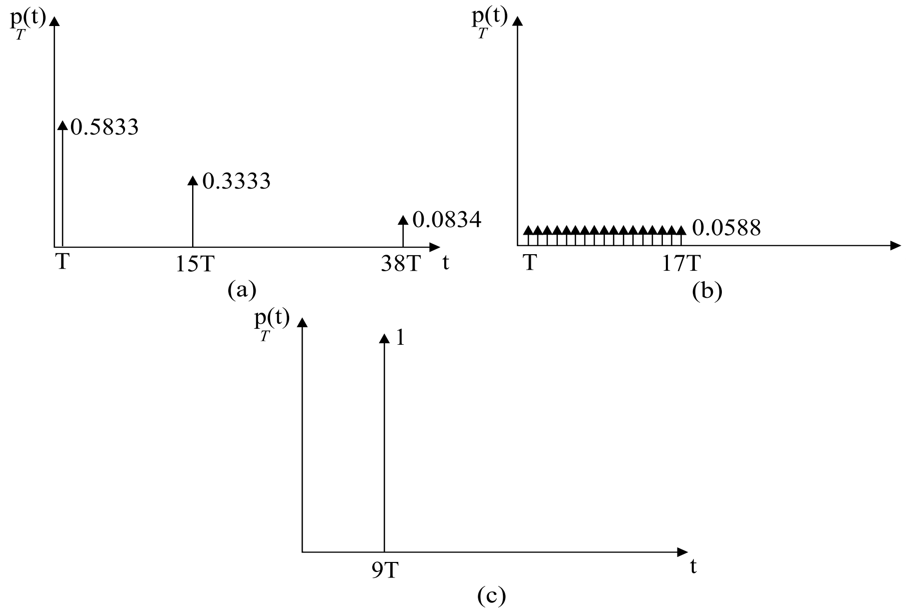

Fig. 6. Packet duration probability distributions: empirical distribution (a), uniform distribution (b), constant duration (c)

1. *Empirical distribution.* Based on real measurements on IP traffic [12], [13], we assumed the following probability distribution for the packet length L:

$$\begin{cases} p_0 = P(L = 40 \text{ bytes}) = 0.5833 \\ p_1 = P(L = 576 \text{ bytes}) = 0.3333 \\ p_2 = P(L = 1500 \text{ bytes}) = 0.0834 \end{cases} \quad (2)$$

   In this model, packets have average length equal to 341 bytes. Since a 40-bytes long packet is transmitted in one time slot of duration T, the average duration of an optical packet is approximatively $9T$. Moreover, p_0, p_1 and p_2 represent the probability that the packet duration is T, $15T$ and $38T$ respectively (see Fig. 6 (a)).
2. *Uniform distribution.* To show a comparison with the empirical model described above, we have modeled the optical packet length as a stochastic variable, uniformly distributed between 40 bytes (duration T) and 680 bytes

(duration $17T$). Also in this model, packets have average duration of $9T$ (see Fig. 6 (b)).

3. *Constant length.* We have also investigated the behaviour of the system when packets have a constant duration of value $9T$ (see Fig. 6 (c)).

These simulations were carried out assuming that no deflection routing algorithm is implemented. Under this assumption, a packet is supposed to be lost if it can't be delayed of a suitable amount of time, in order to transmit it onto the correct output port. Figures 7 through 12 show the packet loss probability at different traffic loads per wavelength, for different values of the maximum delay attainable by the fiber delay lines unit, $D = iT$ $(i = 0, 1, 2, \cdots)$.

Figures 7, 9 and 11 report the simulation results for the optical FIFO buffering (OFB) policy, in the fiber delay lines unit, while Figs. 8, 10 and 12 report the results for the optical scheduling (OS) policy.

It can be seen that, regardless of packet length distribution, the OS policy yields a better performance than the OFB policy, with an increasing improvement as D grows.

Figures 13, 14 and 15 show the values of the ratio

$$\eta = \frac{\Pi_{OS}}{\Pi_{OFB}} \tag{3}$$

for different values of the maximum delay achievable, D, at different traffic loads per wavelength, where Π_{OS} and Π_{OFB} are the packet loss probability for the optical scheduling and optical FIFO buffering policy, respectively. It can be pointed out that no significative improvement is experienced as D value is 0, 1, 2 or 4.

It can also be seen that this increase is more evident for the uniform distribution, and even more for the constant length packets. This happens because the system performance is not only influenced by the maximum delay achievable, D, but also by the maximum optical packet length.

In fact, what really influences the system performance is the L_M/D rate, being L_M the maximum packet duration and D the maximum delay attainable. So as the value of D is much smaller than L_M, the influence of the fiber delay lines managing discipline is negligible. When D becomes much larger than L_M, on the other hand, the OS policy efficiency improvement becomes more and more evident.

It is now interesting to show the efficiency improvement variation, yielded by the OS policy, depending on the optical packet length. Figure 16 plots this variation at different loads per wavelength, for three values of the packet length, for constant length packets, when the maximum delay attainable is $D = 16$. It can be seen that, for $L_M = 8$, that is $L_M = D/2$, a significative efficiency improvement is experienced, while for $L_M = 32$, that is $L_M = 2D$, the optical scheduling and optical FIFO buffering policies almost give the same performance.

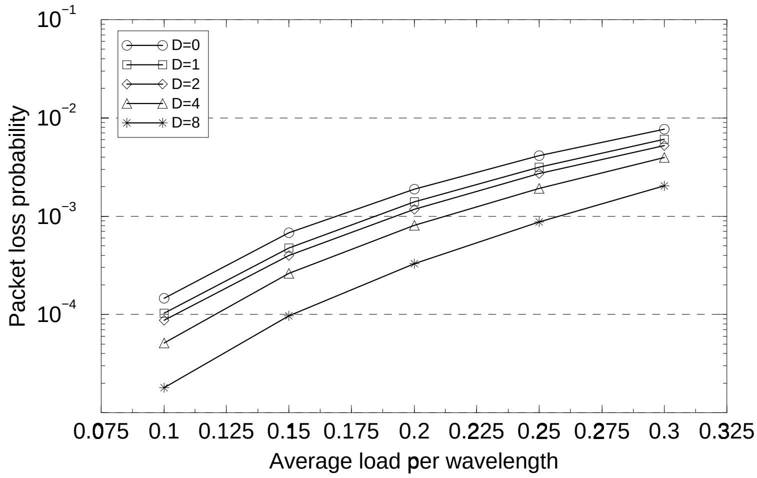

Fig. 7. Packet loss probability for the empirical distribution, with FIFO policy, at different loads per wavelength

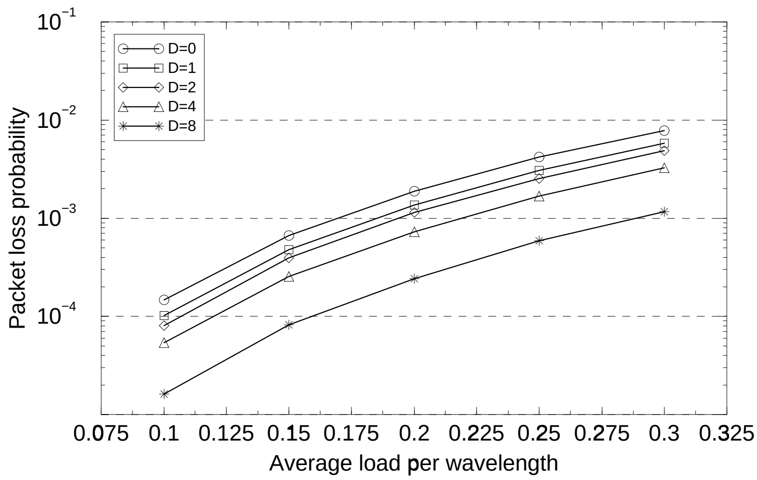

Fig. 8. Packet loss probability for the empirical distribution, with scheduling policy, at different loads per wavelength

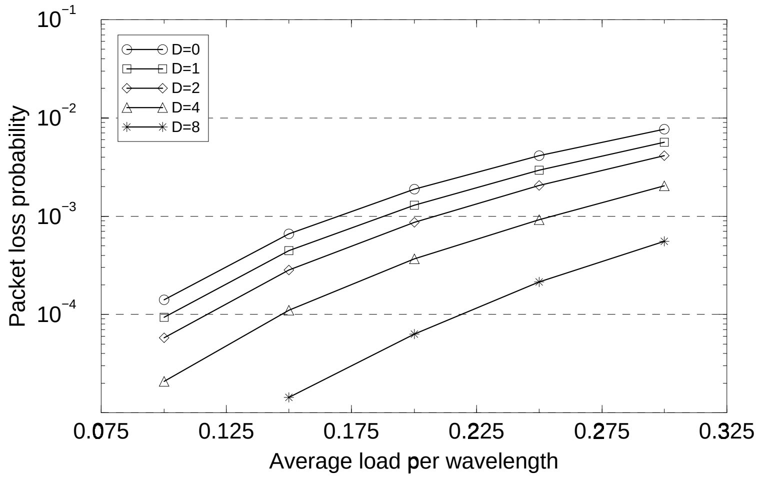

Fig. 9. Packet loss probability for the uniform distribution, with FIFO policy, at different loads per wavelength

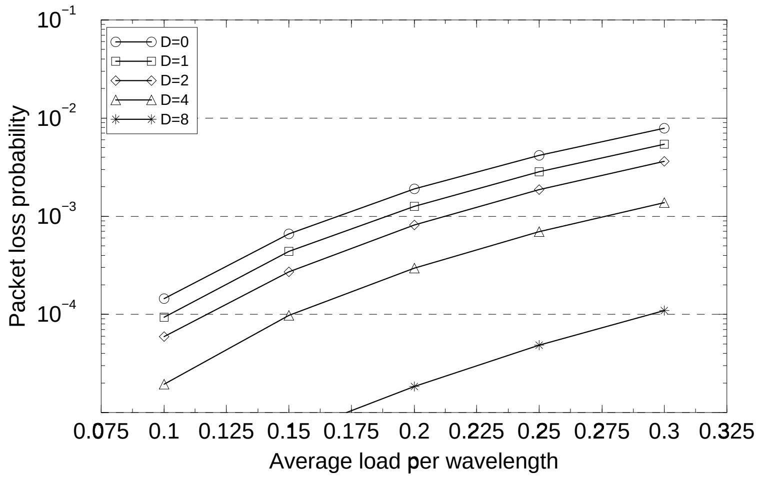

Fig. 10. Packet loss probability for the uniform distribution, with scheduling policy, at different loads per wavelength

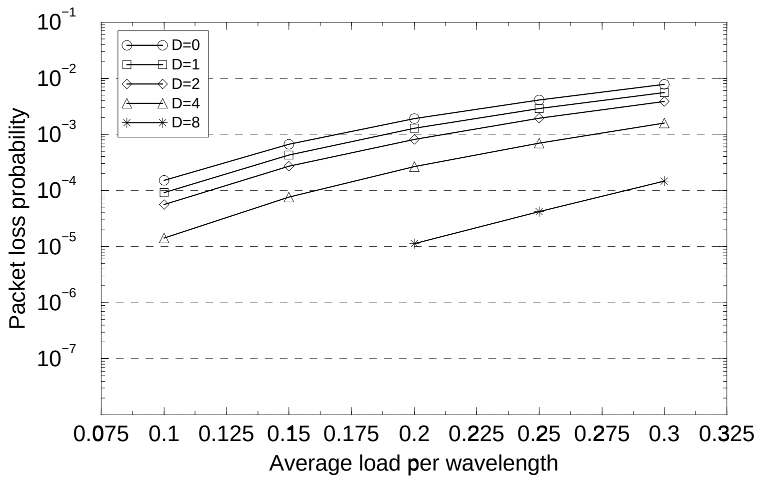

Fig. 11. Packet loss probability for constant packet length, with FIFO policy, at different loads per wavelength

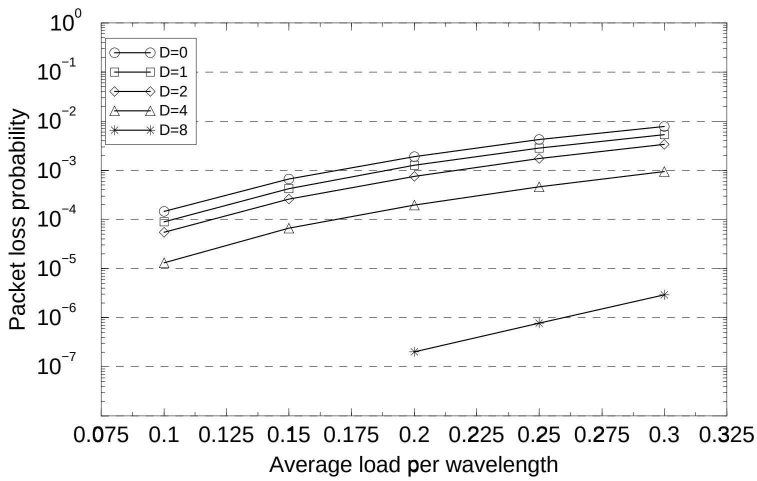

Fig. 12. Packet loss probability for constant packet length, with scheduling policy, at different loads per wavelength

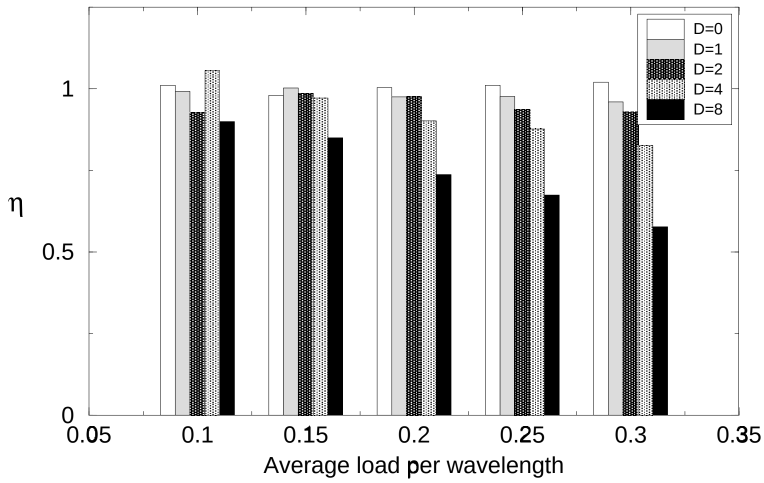

Fig. 13. Empirical distribution: values of η at different loads per wavelength and for different values of the maximum delay attainable D

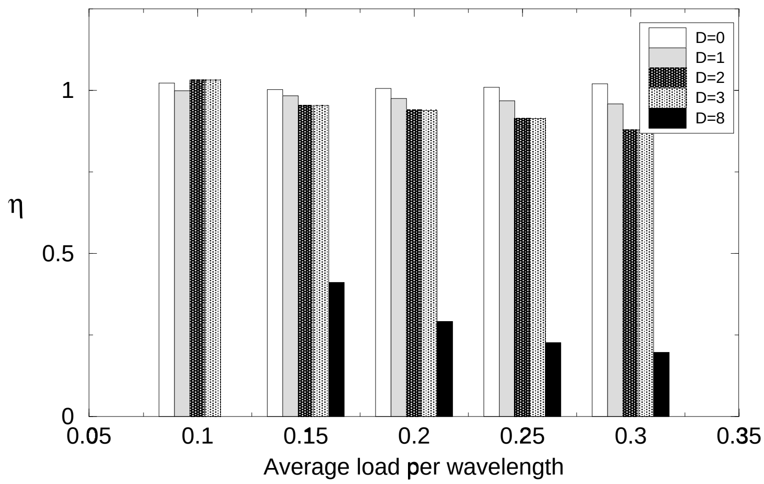

Fig. 14. Uniform distribution: values of η at different loads per wavelength and for different values of the maximum delay attainable D

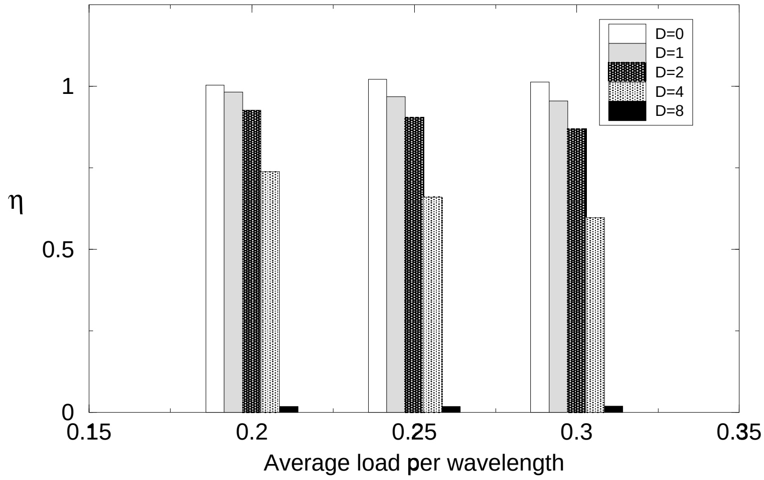

Fig. 15. Constant packet length: values of η at different loads per wavelength and for different values of the maximum delay attainable D

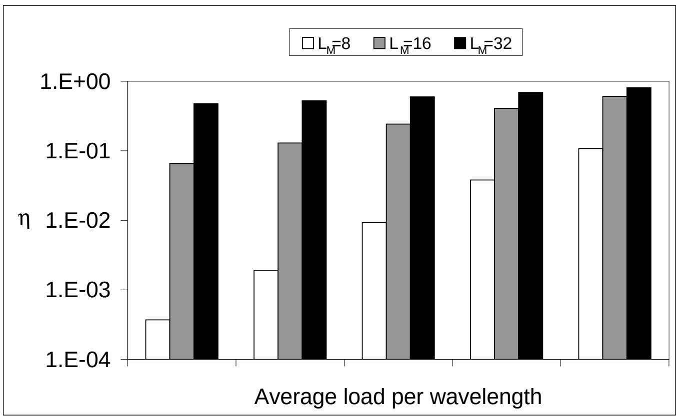

Fig. 16. Constant packet length: values of η at different loads per wavelength, for $D = 16$, $L_M = 8$, $L_M = 16$ and $L_M = 32$

6 Conclusions and Topics for Further Research

In this work, we proposed an architecture for optical packet-switched transport networks. The structure of the optical switching nodes was detailed and the basic building blocks were described. Some simulation results were also presented, showing a comparison between two different managing policies for the fiber delay lines stage: *optical scheduling* and *optical FIFO buffering*.

It was shown that, for $D << L_M$, OS and OFB almost give the same performance. For $D >> L_M$, on the other hand, the optical scheduling policy yields a better performance than the optical FIFO buffering policy, because the output links are more efficiently exploited.

Many issues will have to be addressed in the future, such as the detailed study of the improvement attainable with the optical scheduling policy depending on the optical packet length. Moreover, the behaviour of an optical transport network, as a whole, will have to be investigated, since a single node operation was simulated for this work. Another interesting issue is the implementation of a suitable deflection routing algorithm in order to improve network performance, varying the optical network topology.

References

1. Hunter, D.K., Andonovic, I.: Approaches to Optical Internet Packet Switching. IEEE Commun. Mag. (Sep. 2000) 116-122
2. Yao, S., Mukherjee, B., Dixit, S.: Advances in Photonic Packet Switching: An Overview. IEEE Commun. Mag. (Feb. 2000) 84-94
3. Yoo, M., Qiao, C.: Just-Enough-Time(JET): A High Speed Protocol for Bursty Traffic in Optical Networks. Proc. IEEE/LEOS Tech. for a Global Info. Infrastructure (Aug. 1997) 26-27
4. Qiao, C.: Labeled Optical Burst Switching for IP-over-WDM Integration. IEEE Commun. Mag. (Sep. 2000) 104-114
5. Hunter, D.K. et al.: WASPNET: A Wavelength Switched Packet Network. IEEE Commun. Mag. (Mar. 1999) 120-129
6. Renaud, M., Masetti, F., Guillemot, C., Bostica, B.: Network and System Concepts for Optical Packet Switching. IEEE Commun. Mag. (Apr. 1997) 96-102
7. Dorgeuille, F., Mersali, B., Feuillade, M., Sainson, S., Slempkès, S., Foucher, M.: Novel Approach for Simple Fabrication of High-Performance InP-Switch Matrix Based on Laser-Amplifier Gates. IEEE Photon. Technol. Lett., Vol. 8. (1996) 1178-1180
8. Stephens, M.F.C. et al.: Low Input Power Wavelength Conversion at 10 Gb/s Using an Integrated Amplifier/DFB Laser and Subsequent Transmission over 375 km of Fibre. IEEE Photon. Technol. Lett., Vol. 10. (1998) 878-880
9. Owen, M. et al.: All-Optical 1x4 Network Switching and Simultaneous Wavelength Conversion Using an Integrated Multi-Wavelength Laser. Proc. ECOC '98, Madrid, Spain, (1998)
10. Tzanakaki, A. et al.: Penalty-Free Wavelength Conversion Using Cross-Gain Modulation in Semiconductor Laser Amplifiers with no Output Filter. Elec. Lett., Vol. 33. (1997) 1554-1556

11. Parker, C., Walker, S.D.: Design of Arrayed-Waveguide Gratings Using Hybrid Fourier-Fresnel Transform Techniques. IEEE J. Selec. Topics Quant. Electron., Vol. 5. (1999) 1379-1384
12. Thompson, K., Miller, G.J., Wilder, R.: Wide-area Internet Traffic Patterns and Characteristics. IEEE Network, Vol 11. (1997) 10-23
13. Generating the Internet Traffic Mix Using a Multi-Modal Length Generator. Spirent Communications white paper. http://www.netcomsystems.com