Node Architecture Design

for All-Optical IP Packet Switching

Stefano Bregni, Giacomo Guerra, Achille Pattavina
Dept. of Electronics and Information, Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
bregni@elet.polimi.it, guerra@elet.polimi.it, pattavina@elet.polimi.it

***Abstract*— In order to support efficiently the fast-growing demand for transmission capacity, optical packet-switched systems seem to be strong candidates as they allow fast dynamic allocation of WDM channels combined with a high degree of statistical resource sharing. In this work, we propose the architecture of an input/shared-buffered optical packet switching node, for all-optical switching of IP traffic flows. Since a limited hardware complexity is a key requirement for all-optical switches, due to the high cost of optical components, different node configurations are compared in order to evaluate the effect of a hardware-complexity variation on the performance of the proposed architecture.**

I. Introduction

As telecommunication networks are experiencing a dramatic increase in demand for capacity, mainly related to the exponential growth of the IP traffic, they are evolving to provide a reconfigurable optical layer. This optical layer will 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, [1], [2], [3], [4], in order to study how to achieve the switching of packets directly in the optical domain, in a way that allows fast dynamic allocation of WDM channels, so as to improve the transport network performance.

In this context all-optical switching fabrics become more and more important since these systems allow the switching of signals directly in the optical domain, eliminating the need for many time-consuming optical-electrical-optical conversions. Unfortunately, today optical devices used in market equipment are still very expensive. Thus a key requirement in optical switches design is to limit hardware complexity, due to the high cost of optical components.

In this article we present the architecture of an optical packet switching node equipped with fiber delay lines used as both input and shared buffers to delay packets for contention resolution. Performance evaluation, in terms of packet loss probability and average delay, is carried out for this node when it is loaded by a real IP traffic pattern. Different node configurations are compared in order to evaluate the effect of a hardware-complexity variation on the performance of the proposed architecture.

The paper is organized as follows. Sections II and III describe the optical network architecture we envision and the proposed architecture of an optical packet switching node. Traffic performance results attainable with different node configurations are described in section IV. Some conclusions are finally given.

II. Optical Transport Network

The architecture of the optical transport network we propose consists of M optical packet switching nodes linked together in a mesh like topology, which allow switching of datagram, variable-length IP-like packets directly in the optical domain. These nodes are connected to IP legacy (electronic) networks by means of electro-optic edge systems (ES), as shown in Fig. 1.

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

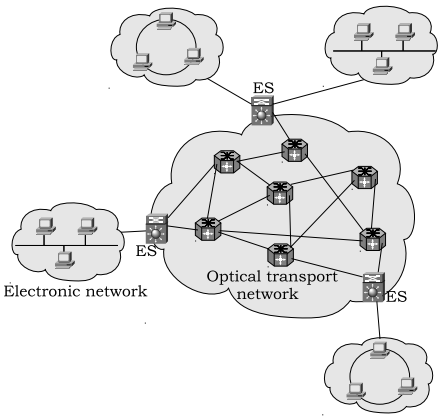

Fig. 1. The optical transport network architecture.

will properly configure the node devices 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. Supposing a bit rate of 10 Gbps per wavelength channel, a 40 ns slot duration seems to be appropriate, since the 40 bytes payload is transmitted in 32 ns, and the additional time can be used for the optical packet header transmission and for the provision of guard times.

III. Optical Node Architecture

The general architecture of a network node consists of N incoming fibers with W wavelengths per fiber (see Fig. 2).

The incoming fiber signals are demultiplexed and G wavelengths from each input fiber are fed into one of the W/G switching planes, which constitute the switching fabric core. Once signals have been switched in one of the second-stage parallel planes, packets can reach every output port on one of the G wavelengths that are directed to each output fiber. This allows the use of wavelength conversion for contention resolution, since G packets can be concurrently transmitted, by each second-stage plane, on the same output link. Thus the parameter G is here referred to as "channel group size".

A. Switching Plane Structure

The detailed structure of one of the W/G parallel switching planes is presented in Fig. 3. It consists of three main blocks. An n-stage input *synchronization unit*, as the node is slotted and incoming packets need to be aligned. This unit consists of a series of 2×2 SOA optical switches, [5], 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 by a variable amount of time, ranging between $\triangle t_{min} = 0$ and $\triangle t_{max} = (1 - (1/2)^n) \cdot T$, with a resolution of $T/2^n$, where T is the time slot duration and n the number of delay lines.

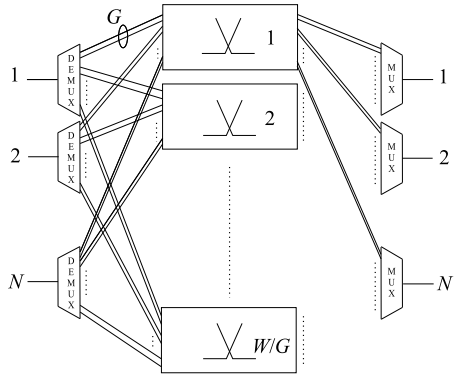

Fig. 2. The optical packet-switching node architecture.

A *fiber delay lines* (FDL) *unit* is then used to store packets for contention resolution. The delay lines unit is used as an *optical scheduler*, by proper operation of a set of tunable wavelength converters, [6], [7], in order to schedule the transmission of the maximum number of packets onto the correct output link. This is obtained delaying each packet by a variable amount of time, ranging between 0 and D_{max}.

Finally a *switching matrix unit* is used to achieve the switching of signals, using an AWG (Arrayed Waveguide Grating, [8]). The implementation of each unit of a switching plane is detailed in [9].

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

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

Actually the switching matrix is used to switch packets to the output ports or, if necessary, to a number of recirculation ports, in order to store them for an additional amount of time to avoid collisions. To this purpose R output ports of the AWG are connected, via fiber delay lines, to R input ports. Each switching plane is then characterized by $C = N \cdot G$ traffic channels and R recirculation channels, and the switching matrix permutations are managed using tunable wavelength converters.

As far as recirculation lines are concerned, two different structures are considered: the constant delay recirculation (CDR) and the variable delay recirculation (VDR). In the CDR structure all the recirculation ports delay each packet by the same amount of time, $D_{rec} = kT$, while in the VDR structure D_{rec} doubles every two ports, in such a way that the $i-th$ couple of recirculation lines delays each packet by $D_{rec,i} = 2^{(i-1)}T$. The first couple of ports will then have a recirculation delay of T, the second couple of $2T$, and so on.

B. Switching Plane Hardware Complexity

Since hardware complexity is a key constraint in optical switching node architecture design, in the following we discuss the complexity of a single switching plane of the model we propose. The hardware complexity is expressed in terms of number of SOA switches, n_{SOA}, mux/demux, n_{MUX}, tunable and fixed wavelength converters, n_{TWC} and n_{FWC}, as a function of the number of traffic channels $C = N \cdot G$:

$$\begin{cases} n_{SOA} = (n+1) \cdot C \\ n_{MUX} = 2C \\ n_{TWC} = 2C + R \\ n_{FWC} = C \end{cases} \quad (1)$$

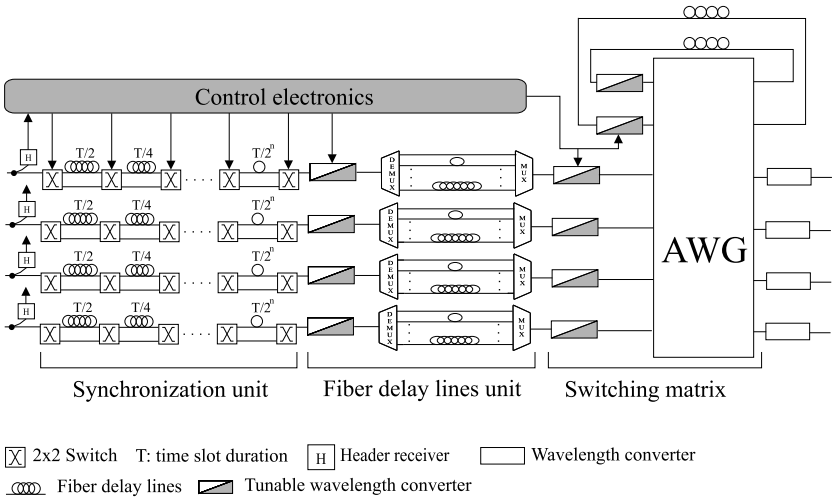

Fig. 3. Detailed structure of one of the W/G parallel switching planes.

It can be pointed out that each parameter varies linearly with the number of traffic channels C and, therefore, with the product $N \cdot G$.

C. Node Hardware Complexity

We would like to investigate the effect of a variation of the parameter G on the hardware complexity of the node, keeping the number N of input/output fibers and the number W of wavelength channels per fiber constant. A reduction of the parameter G, by a factor 2^k, yields a reduction of the hardware complexity of each switching plane, in such a way that:

$$\begin{cases} n'_{SOA} = (n+1) \cdot N\frac{G}{2^k} = (n+1) \cdot \frac{C}{2^k} \\ n'_{MUX} = 2\frac{C}{2^k} \\ n'_{TWC} = 2\frac{C}{2^k} + R \\ n'_{FWC} = \frac{C}{2^k} \end{cases} \tag{2}$$

However, if the number of wavelength channels per input/output fiber is constant, a reduction of G yields an increase in the number of switching planes, by the same factor 2^k. The node hardware complexity is then given by:

$$\begin{cases} n'_{SOA} \cdot \frac{W}{G}2^k = (n+1) \cdot N\frac{G}{2^k}\frac{W}{G}2^k = n_{SOA} \cdot \frac{W}{G} \\ n'_{MUX} \cdot \frac{W}{G}2^k = n_{MUX} \cdot \frac{W}{G} \\ n'_{TWC} \cdot \frac{W}{G}2^k = (n_{TWC} + (2^k - 1)R) \cdot \frac{W}{G} \\ n'_{FWC} \cdot \frac{W}{G}2^k = n_{FWC} \cdot \frac{W}{G} \end{cases} \tag{3}$$

Thus a reduction of the parameter G leaves the hardware complexity of the node unchanged, with the exception of the number of tunable wavelength converters used to manage the recirculation lines.

The same considerations are valid when the parameter G and the total number of external wavelength channels, $N \cdot W$, are kept unchanged while varying the number N of input/output fibers and the number W of wavelength channels per fiber. In fact, if N is reduced by a factor 2^k and W is increased by the same factor, to keep $N \cdot W$ constant, the number of traffic channels per switching plane, $C = N \cdot G$, and the number of switching planes, W/G, change in the same way as the previous situation.

IV. Simulation Results

In this section we show some simulation results of the operation of different configurations of the node architecture, which were carried out assuming a packet length distribution based on real measurements on IP traffic [10]:

$$\begin{cases} p_0 = P(L = 40 \text{ bytes}) = 0.6 \\ p_1 = P(L = 576 \text{ bytes}) = 0.25 \\ p_2 = P(L = 1500 \text{ bytes}) = 0.15 \end{cases} \tag{4}$$

In this model, packets have average length equal to 393 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 $10T$. Moreover, p_0, p_1 and p_2 represent the probability that the packet duration is T, $15T$ and $38T$ respectively.

We evaluate how the size of the parameter G, which represents the number of channels per input/output fibers handled by a single plane, affects the overall packet loss performance of the node. To this aim we have selected a node architecture with a single switching plane ($W = G$). We have first compared four switch configurations with the same external lightpath number ($N \cdot W$), and no recirculation lines ($R = 0$). By assuming the availability of a 32×32 AWG and an FDL

stage with maximum delay $D_{max} = 8T$, the switch size varies in the set $N = \{2, 4, 8, 16\}$ and the channel group size in the set $G = \{16, 8, 4, 2\}$, in such a way that $N \cdot G = 32$. Fig. 4 shows that for a given offered load the packet loss performance improves as G increases. In particular for low levels of the offered traffic the improvement can be of several orders of magnitude. This improvement is nothing else than that attained in any multiple-server system, in which all users fully share the set of servers. Traffic engineers well know this phenomenon under various names, among which perhaps the most common is "channel grouping" (or "trunk grouping").

We have then compared two configurations, with and without recirculation lines. A single switching plane node, equipped with an 8×8 AWG, has been selected. The number of input/output fibers is kept constant ($N = 2$), while the grouping factor varies in the set $G = \{2, 4\}$.

Figs. 5 and 6 show the packet loss probability and average delay for a node with and without fiber recirculation lines, for different values of the offered load; the FDL stage maximum delay is $D_{max} = 16T$ and the recirculation lines configuration is the CDR configuration. It can be pointed out that the reduction of the grouping factor G, from $G = 4$ ($R = 0$) to $G = 2$ ($R = 4$ and $D_{rec} = \{T, 4T\}$), yields higher loss probability and average delay. This performance worsens more as the average traffic load decreases, since the effect of the grouping factor variation is more evident for low levels of the offered load, as we pointed out before. A similar behavior can be observed comparing the curves obtained with the configurations $G = 4$, $R = 0$ and $G = 2$, $R = 4$, $D_{rec} = 16T$. Now the delay D_{rec} is large enough to compensate the loss probability increase due to the grouping factor reduction. As far as average delay is concerned, however, it can be pointed out that the node without recirculation lines always yields a better performance since more contentions can be resolved in the wavelength domain.

Similar considerations can be drawn observing Fig. 7, which plots the packet loss probability, at different traffic loads per wavelength, for a switching node with $R = 4$ recirculation lines (VDR configuration) and a node without recirculation lines ($R = 0$), for different values of D_{max}.

Finally, we have compared the performance of two architectures, with and without recirculation fibers, with the same number of input/output fibers $N = 2$, the same value of the grouping factor $G = 4$, varying the AWG dimension. An 8×8 AWG, without recirculation lines and a 16×16 AWG, with $R = 8$ recirculation lines have been selected to this aim. Figs. 8 and 9 show a comparison between the packet loss performance of these models. It can be pointed out that, as the grouping factor doesn't change, the nodes with recirculation lines always give a better performance, since they have a higher buffering capability than the nodes without recirculation lines, while the same number of contentions can be resolved in the wavelength domain.

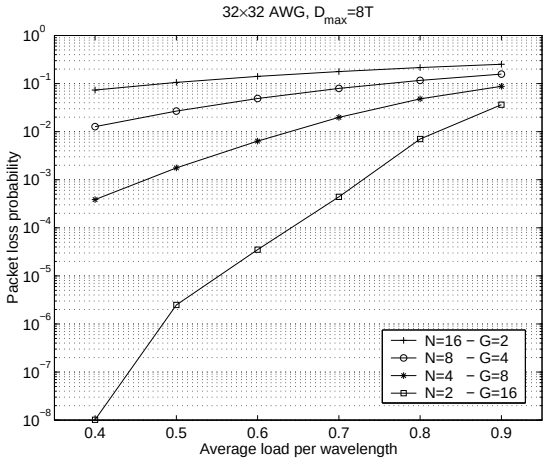

Fig. 4. Packet loss performance for different grouping factor and node dimension values.

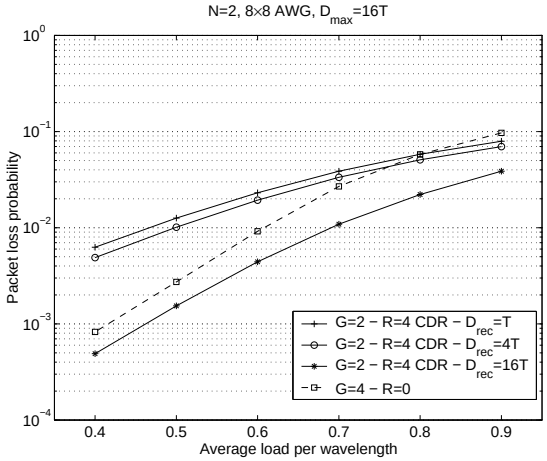

Fig. 5. Packet loss performance for different grouping factor values, with and without recirculation lines ($R = 0$ vs. $R = 4$ CDR).

V. Conclusions

In this work, we have proposed an architecture for optical packet-switched transport networks. The proposed structure of the optical switching nodes has been presented and the basic building blocks have been described. Simulation results have been also presented, showing a comparison between different node configurations.

It has been shown that the hardware complexity of the node, in terms of optical components number, varies linearly with the number of traffic channels C and, therefore, with the the number G of channels per input/output fibers handled by a single plane. It has been then found that the value of G has a great influence on the overall performance of the node as well. In fact a reduction of G, for the same value of the overall delay-lines buffer capacity, yields a reduction of the number of contentions that can be resolved in the wavelength domain.

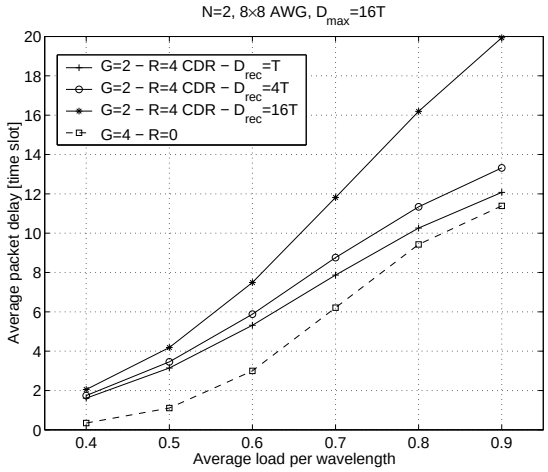

Fig. 6. Average packet delay for different grouping factor values, with and without recirculation lines ($R = 0$ vs. $R = 4$ CDR).

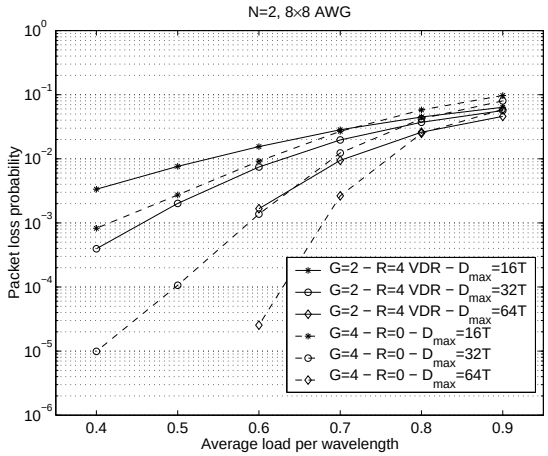

Fig. 7. Packet loss performance for different grouping factor values, with and without recirculation lines ($R = 0$ vs. $R = 4$ VDR).

Nevertheless, this reduction worsens the node performance especially in terms of packet loss probability.

REFERENCES

[1] M. Renaud, F. Masetti, C. Guillemot, and B. Bostica, "Network and System Concepts for Optical Packet Switching," *IEEE Communications Magazine*, vol. 35(4), pp. 96–102, April 1997.

[2] D.K. Hunter and I. Andonovic, "Approaches to Optical Internet Packet Switching," *IEEE Communications Magazine*, vol. 38(9), pp. 116–122, September 2000.

[3] S. Yao, B. Mukherjee, and S. Dixit, "Advances in Photonic Packet Switching: an Overview," *IEEE Communications Magazine*, vol. 38(2), pp. 84–94, February 2000.

[4] F. Callegati, G. Corazza, and C. Raffaelli, "Design of a WDM Optical Packet Switch for IP Traffic," in *Global Telecommunications Conference, 2000. GLOBECOM '00*, 2000, vol. 2, pp. 1283–1287.

[5] F. Dorgeuille, B. Mersali, M. Feuillade, S. Sainson, S. Slempkès, and M. Foucher, "Novel Approach for Simple Fabrication of High Performance InP-Switch Matrix Based on Laser-Amplifier Gates," *IEEE Photonics Technology Letters*, vol. 8(9), pp. 1178–1180, September 1996.

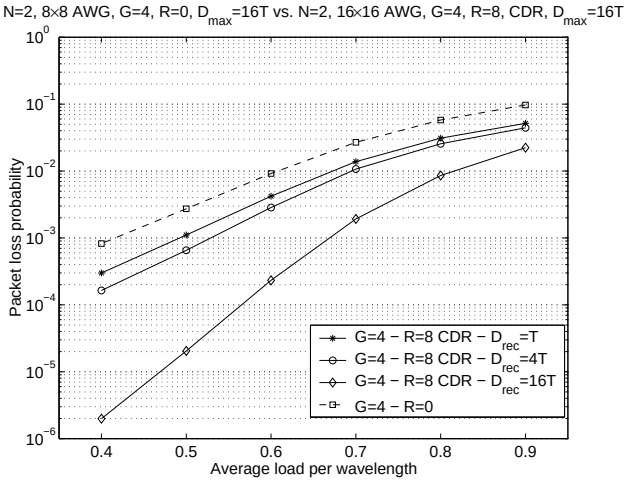

Fig. 8. Packet loss performance for the same grouping factor value, with and without recirculation lines ($R = 0$ vs. $R = 8$ CDR).

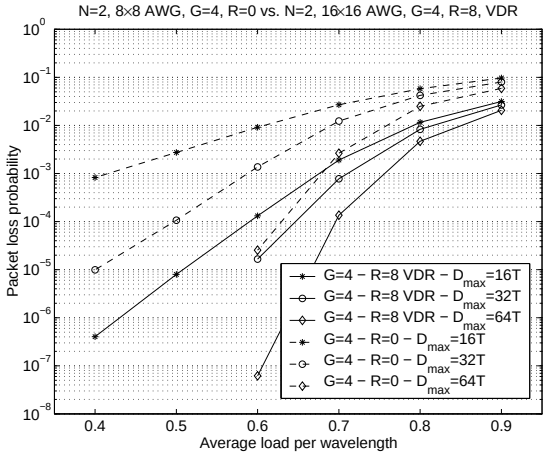

Fig. 9. Packet loss performance for the same grouping factor value, with and without recirculation lines ($R = 0$ vs. $R = 8$ VDR).

[6] A. Tzanakaki and M.J. O'Mahony, "Analysis of Tunable Wavelength Converters Based on Cross-Gain Modulation in Semiconductor Optical Amplifiers Operating in the Counter Propagating Mode," *IEE Proceedings-Optoelectronics*, vol. 147(1), pp. 49–55, February 2000.

[7] M.W.K. Mak and H.K. Tsang, "Polarization-Insnsitive Widely Tunable Wavelength Converter Using a Single Semiconductor Optical Amplifier," *IEEE Electronics Letters*, vol. 32(2), pp. 152–153, January 2000.

[8] M. Volanthen, M. Van der Vliet, V. Tandon, and J. Bonar, "Characterisation of Arrayed Waveguide Gratings," March 2001, Kymata Ltd. white paper at http://www.kymata.com.

[9] S. Bregni, G. Guerra, and A. Pattavina, "Optical Packet Switching of IP Traffic," in *Sixth Working Conference on Optical Network Design and Modelling (ONDM), Torino, Italy*, 2002.

[10] K. Thompson, G.J. Miller, and R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics," *IEEE Network*, vol. 11(6), pp. 10–23, 1997.