AWG-BASED WDM SWITCHING ARCHITECTURES FOR ALL-OPTICAL IP NETWORKS

Stefano Bregni
Politecnico di Milano
bregni@elet.polimi.it

Achille Pattavina
Politecnico di Milano
pattavina@elet.polimi.it

Gianluca Vegetti
Politecnico di Milano

Abstract As the demand of transmission capacity grows, optical packet switching technology is emerging as a strong candidate, promising to 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 switching node for all-optical packet switching of IP traffic flows. Since limited hardware complexity and cost is a key requirement for all-optical switches, this new node architecture is compared in terms of packet loss probability with a previously proposed architecture of the same cost order.

1. Introduction

The growth of existing and new services is creating a large increase of badwidth request in telecommunication networks, which is likely to grow in the near future. In the latest years, great efforts were made to bring fast traffic channels up to the final user, thus making the Broadband Integrated Service Digital Network (B-ISDN) finally real.

In order to support this huge amount of traffic, optical backbone networks are being employed, exploiting the virtually unlimited bandwidth provided by Dense Wavelength Division Multiplexing (DWDM). However, nowadays packets undergo an Optical/Electrical (O/E) conversion before being switched and an other Electrical/Optical (E/O) to be transmitted. Therefore, these optical-electrical-optical (OEO) switches suffer from a bottleneck in the switching phase, since all these operations (i.e. conversions and electronic switching) cannot be performed at an adequate rate.

The development of devices able to switch optical signals resulted in a burst of research works, promising to make switching directly available in the optical domain. Moreover, recently DWDM has evolved to support some network functions as circuit routing and wavelength conversion and assignment, leading to the idea of all-optical networks, employing optical-optical-optical (OOO) switching nodes. Unfortunately, today optical devices used in market equipment are still too crude to allow packet-by-packet operation, so the upcoming of full-optical networks will probably occur gradually.

The simplest solutions are WDM-routed networks, where a wavelength is assigned to each connection in such a way that all traffic is handled in the optical domain, without any electrical processing on transmission. This kind of networks are basically circuit-switched.

An interesting solution which tries to represent a balance between Circuit-switching low hardware complexity and packet-switching efficient bandwidth utilization is the Optical Burst Switching (OBS, [1], [2]). In an optical burst switching system, 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 electronically the necessary resources on the intermediate nodes of the transport network. Such operation results in a lower average processing and synchronization 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.

The last solution is Optical Packet Switching (OPS), which promises to perform traffic switching with packet granularity, like with current electronic switches, in order to exploit optical bandwidth more efficiently. Many hurdles are still to be overcome in this field, such as the high reconfiguration speed needed and, mainly, the impossibility to store optical packets to resolve packet contention. The two most promising OPS architecture are KEOPS ([3]) and WASPNET ([4], [5]) projects.

We address here the long-term view of a full packet switching network performing IP packet transport, in which optical operations are performed as much as possible exploiting the currently available optical device technology. Apparently most of the operations related to the

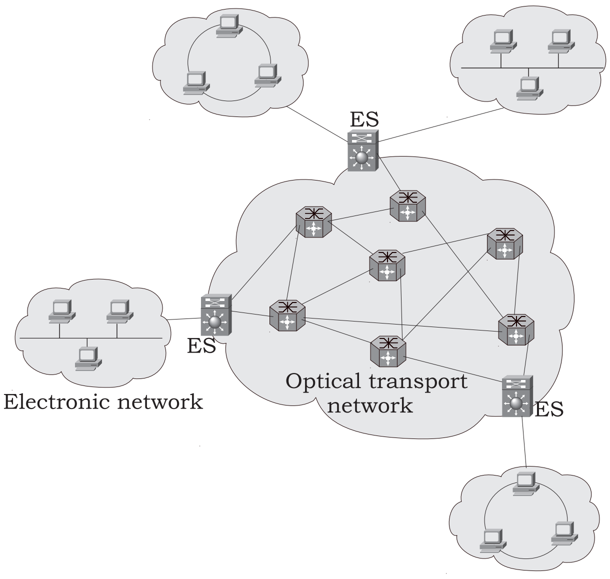

Figure 1. The optical transport network architecture

packet header processing needs to be done nowadays in the electronic domain. In this paper we present an alternative to the architecture of an optical packet switching node first proposed in [6][7], which is equipped with a fiber delay line stage used both as input and shared buffer for optical packets.

The paper is organized as follows. Sections 2 and 3 describe the optical network architecture we envision and the proposed architectures of an optical packet switched node. Traffic performance results attainable with different node configurations are described in section 4.

2. Optical Transport Network

The architecture of the optical transport network we propose consists of M *optical packet-switching nodes*, each denoted by an optical address made of $\lceil m = log_2 M \rceil$ 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. 1).

Edge systems multiplex IP datagrams from electronic networks and encapsulate them into optical packets with no fragmentation or aggragation. The optical packet is composed of a simple optical header, which comprises the m-bit 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 resources to switch packet payloads directly in the optical domain.

The transport network operation is *asynchronous*: packets are transmitted between nodes without enforcing any time alignment. The internal operation of the optical nodes, on the other hand, is *synchronous* or *slotted*, since the behavior of packets in an unslotted node is less regulated and more unpredictable, resulting in a larger contention probability. 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 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 to provide guard times.

3. Optical Node Architecture

The general architecture of a network node consists of N incoming fibers with W wavelengths per fiber. It includes three stages: a first stage of channel demultiplexing, a second stage of switching and a third stage of channel multiplexing. In the first stage the incoming fiber signals are demultiplexed and G wavelengths per channel are fed into each of the W/G second-stage switching planes, which constitute the switching fabric core. Once signals have been switched in one of the switching planes, packets can reach every output port through multiplexing carried out in the third stage. Fig. 2 shows this structure in the case of $W/G = 2$. We note that the number of inlets in each third-stage multiplexer varies, depending on the specific structure of the switching planes. Wavelength conversion must be used for contention resolution, since up to G packets can be concurrently transmitted by each second-stage plane on the same output link.

3.1 Switching Plane Structure

The detailed structure of one of the W/G switching planes is presented in Fig. 3. It includes three main blocks. An input *synchronization unit*, as the node is slotted and incoming packets need to be slot-aligned. This unit consists of a series of 2×2 optical switches interconnected by

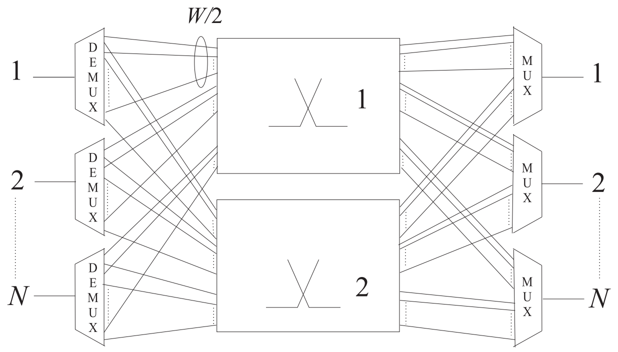

Figure 2. The optical packet-switching node architecture

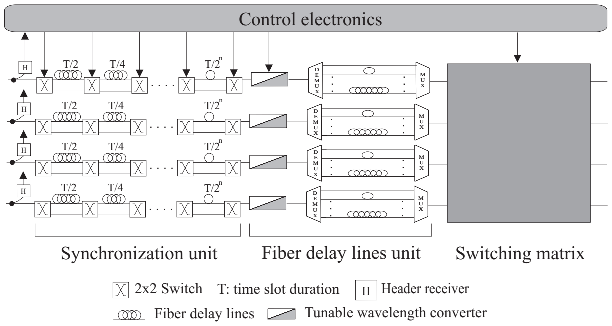

Figure 3. Architecture of the switching plane

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 for a variable amount of time, ranging between $\Delta t_{min} = 0$ and $\Delta t_{max} = 2(1 - (1/2)^{n+1}) \times T$, with a resolution of $T/2^n$, where T is the time slot duration and n the number of delay line stages. Due to the fast reconfiguration speed needed, fast 2×2 switching devices, such as 2×2 semiconductor optical amplifier (SOA) switches [8] that have a switching time in the nanosecond range, must be used.

A *fiber delay line* (FDL) *unit* is then used to store packets for contention resolution. As was pointed out in [6] a better performance is obtained by using this stage as an *optical scheduler*, by proper operation of a set of tunable wavelength converters (see, e.g., [9], [10]), 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*, described in section 3.2, is used to achieve switching of signals.

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, so that these enter the FDL stage slot-aligned;
- tunable wavelength converters managing, in order to properly delay and route incoming packets.

3.2 Switching matrix unit structure

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]. 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. Commercially available 40 channel devices have a channel spacing of 100 Ghz and show an insertion loss of less than 7.5 dB [12].

Two different structures are proposed here for the accomplishing the tasks of this stage, referred to as structures (a) and (b). We will consider multi-plane implementations of nodes with structure (a) switching matrix, while, for reasons that will be clear in section 3.2.2, only single-plane structures will be considered for structure (b). The reader interested in a multi-plane extension or in a deeper analysis of these two structures is referred to [13]. Actually, in both switching structures R AWG ports are reserved to manage recirculation delay lines, which act as a shared buffer. Each of these lines delays a packet by the same amount of time, $D_{rec} = kT$. These recirculation lines can make easier than pure input beffers the provision of different switching service classes by dinamically adjusting the priorities of packets being recirculated.

3.2.1 Structure (a). The simplest switching structure, first introduced in [6], is shown in Fig. 4. In this solution, each signal is fed to an AWG inlet and its wavelength is converted in order to route the packet to a desired output port. Only one packet at a time can be routed to each AWG outlet. After exiting the AWG, each packet is then converted to one of the G wavelengths used for transmission on the considered

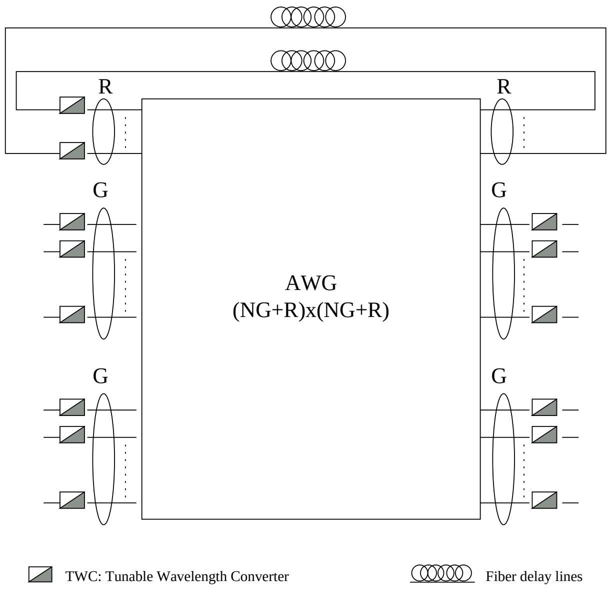

Figure 4. Structure (a) switching matrix

switching plane. Thus, in this structure each switching plane requires an $(NG + R) \times (NG + R)$ AWG.

3.2.2 Structure (b). We will now first describe this structure in the case of $R = 0$ (i.e. no recirculation lines are employed). Since for this structure we only consider single-plane structures, $G = W$ and so NW signals are fed into this switching matrix. The key feature of the structure in Fig. 5 is its ability to route more than one packet from a single AWG inlet to different outlets. In fact, the AWG working principle is such that signals with different wavelengths entering the same inlet will emerge on different outlets and different wavelengths are needed to route signals from different inlets to the same outlet. If N wavelengths from each DWDM channel are multiplexed on the same AWG inlet, this structure requires a $W \times W$ AWG. So, if the AWG works on the same wavelengths used in the outgoing fibers, no final wavelength conversion is needed. A deeper analysis of this structure implementation issues and a performance comparison with structure (a) has been carried out in [13],

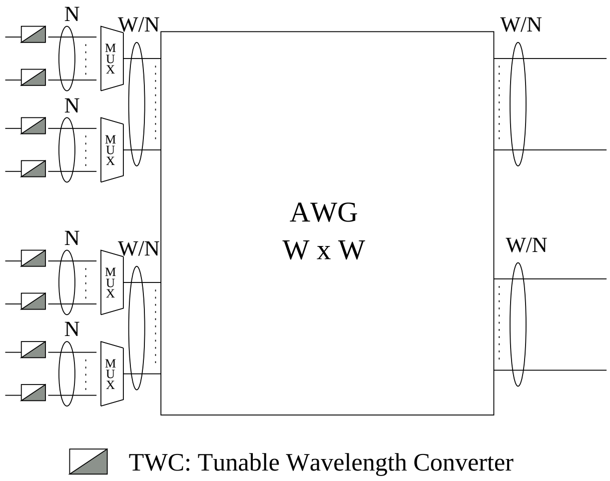

Figure 5. Structure (b) switching matrix with $R = 0$

where it is shown that this structure outperforms structure (a) when $R = 0$ and the two structures are compared with the same AWG size[1].

It is clear that, since the key property of this structure is to have a $W \times W$ AWG, if R AWG ports are reserved to recirculation lines, a different number k of wavelengths are to be multiplexed on each AWG inlet. The number of required AWG ports A can be expressed as $A = R + NW/k$. Therefore, since the relationship $A = W$ must be satisfied

$$W = R + \frac{NW}{k}.$$

Given the number of wavelengths per channel W and the number of recirculation ports R, the number of wavelengths fed to each AWG inlet k is expressed as

$$k = \frac{NW}{W - R}, \tag{1}$$

or alternatively, given W and k, R is obtained by

$$R = \frac{W\,(k - N)}{k}, \tag{2}$$

where obviously N, W, k and R must be integer values.

[1]Note that under this hypothesis, implementation of structure (a) leads to a multi-plane structure, since the relationship $NG = W$ must be satisfied to keep AWG size constant, and therefore $W/G > 1$.

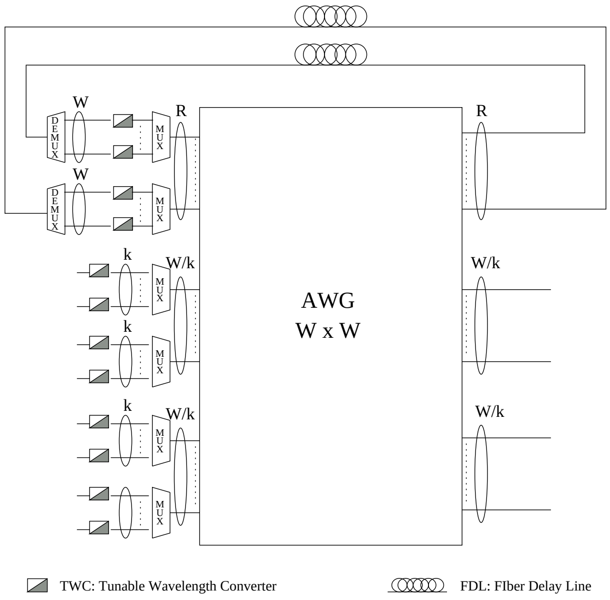

Figure 6. Structure (b) switching matrix

Given these considerations, structure (b) is shown in Fig. 6. Differently from structure (a), up to W packets on different wavelengths can be concurrently routed to each recirculation line. While allowing us to better exploit the buffering capacity of each delay line, this solution implies that before re-entering the AWG, the different wavelengths travelling through the FDL must be splitted and sent to different TWC. The TWC count is a good index of component cost, since TWCs are among the most expensive components of this architecture and a high TWC count is needed in each implementation.

The number of TWC needed to implement a single recirculation line is equal to the number needed to implement W recirculation lines in structure (a). In order to reduce the structure cost, P recirculation ports can be connected to each recirculation line, as shown in Fig. 7. Hence, the number of recirculation lines will be R/P and W/P TWC per recirculation port will be required.

4. Simulation Results

We now show some traffic performance results given by the different node architecture configurations obtained through computer simulation. Packet interarrival has been modelled as a Poisson process with negative exponential interarrival times. Based on measurement of real IP traffic

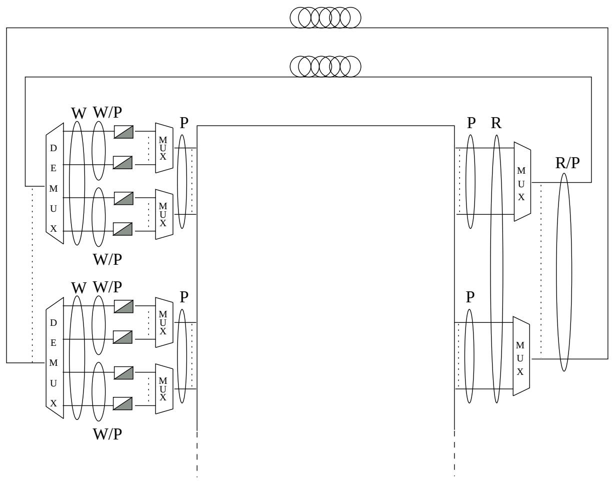

Figure 7. Recirculation lines architecture with $P > 1$

[14], the following packet length distribution has been assumed:

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

In this traffic model, the resulting average packet length is 393 bytes.

We evaluate how the number of recirculation ports R affects the overall packet loss performance of the node. To this aim we have chosen to compare multi-plane (a) structures with single-plane (b) structures, both employed using 12×12 AWG[2]. Given this AWG port count, the number of wavelengths per channel has been set to $W = 12$ and the number of WDM channels to $N = 2$. The FDL maximum delay is $D_{max} = 8$.

We first compare for both structures the two configurations with and without recirculation lines. Therefore, to obtain a structure (a) without recirculation lines and with the desired AWG size $R = 0$ and $G = 6$ have been chosen, and thus $W/G = 2$ parallel switching planes have been employed. As for the configuration with recirculation lines, our aim is to implement one recirculation line per lightpath (i.e. 24 recirculation lines), and so $R = 6$ recirculation ports have been set in every switching plane. Since $G = 3$ has been selected, $W/G = 4$ switching planes must be employed. Fig. 8 shows the packet loss probability for these configurations. It can be pointed out that the reduction of the *grouping*

[2]This value has been chosen in order to give us possibility to compare different architecture implementations.

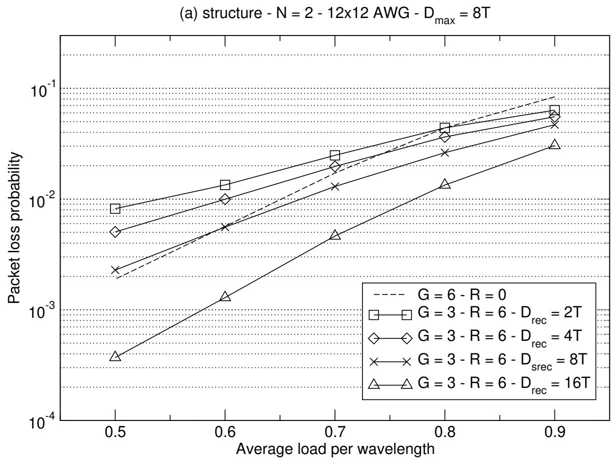

Figure 8. Packet loss probability of (a) structure with $R = 0$, $R = 6$

factor G from $G = 6$ ($R = 0$) to $G = 3$ ($R = 6$ and $D_{rec}\{2T, 4T, 8T\}$), yields a higher loss probability. This performance worsen for lower traffic values, since, as pointed out in [6] and [7], the effect of the grouping factor is more apparent for low levels of the offered traffic load. In the case of $R = 6$ and $D_{rec} = 16T$, the recirculation delay is large enough to compensate the loss probability increase due to the grouping factor reduction, but will lead to higher delays, as shown later.

The implementation of a structure (b) without recirculation lines leads to the choice of $k = N = 2$ wavelengths per AWG port. When shared buffer is employed, we consider the implementation of a structure (b) with the same number of recirculation ports as the (a) structure, $R = 6$. Therefore, from (1), the k parameter has been set to $k = 4$. Fig. 9 shows the packet loss probability for both these configurations. In this case, the implementation of recirculation lines always yields to lower loss probability. This is due to the fact that in structures (b), the implementation of recirculation lines brings higher buffering capacity, being able to buffer up to W packets per line, while reducing lightly the node capability to resolve contention in the wavelength domain.

Next, considering the higher buffering capacity of a single recirculation line, we consider a structure (b) with the minimum number of recirculation ports. To this aim, since for $R = 0$ we have $k = 2$, we consider $k = 3$ and see if (2) gives an integer value for R. The result of (2) for $k = 3$ is $R = 4$. In Fig. 10 a comparison between the two latter structures is traced for $D_{rec} = \{8, 16\}$. Reducing the number of recirculation lines gives a lower loss probability, especially for low levels of offered traffic.

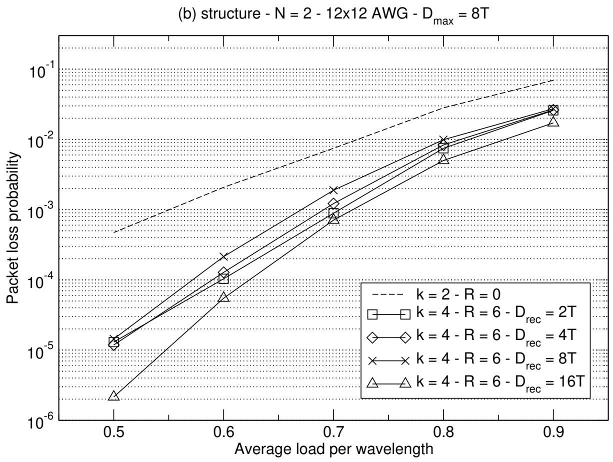

Figure 9. Packet loss probability of (b) structure with $R = 0$, $R = 6$

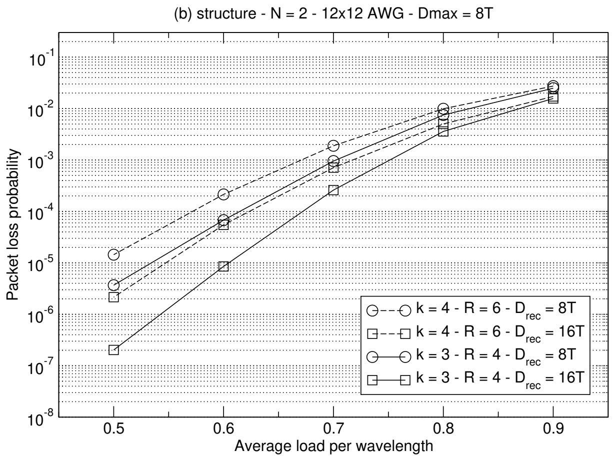

Figure 10. Packet loss probability of (b) structure with $R = \{4, 6\}$ and $D_{rec} = \{8, 16\}$

Once again, this is due to the higher number of contention resolution performed in wavelength domain.

We now evaluate how the P parameter affects node performance. The reduction of the number of recirculation lines, keeping the number of recirculation ports constant, will obviously lead to higher loss probability, since it reduces the buffering capacity of the shared buffer, while not increasing the number of contentions resolved in the wavelength domain. These considerations are confirmed by Figs. 11 and 12, which show the

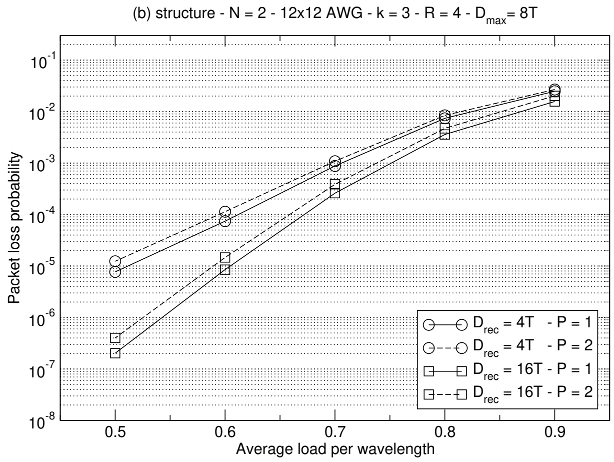

Figure 11. Packet loss probability of (b) structure with $R = 4$ and $P = \{1, 2\}$

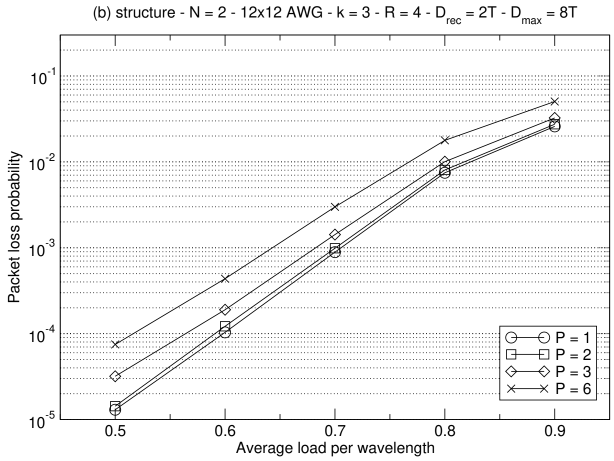

Figure 12. Packet loss probability of (b) structure with $R = 6$ and $P = \{1, 2, 3, 6\}$

packet loss probability of structures (b) with $R = 4$ and $R = 6$, for various values of P.

Finally, it is interesting to compare structures (a) and (b) with the same number of TWC needed in the shared buffer structure. To this aim, we consider a structure (a) with $G = 3$, $R = 6$ and a structure (b) with $k = 3$, $R = 4$, $P = 2$. Obviously both structures have the same parameters N and W, $N = 2$ and $W = 12$. The number of TWCs needed to implement recirculation lines in both structures is 24. Fig. 13 shows the packet loss performance of tese two structures, while in Fig. 14 the average packet delay is traced. Even with a reduced number

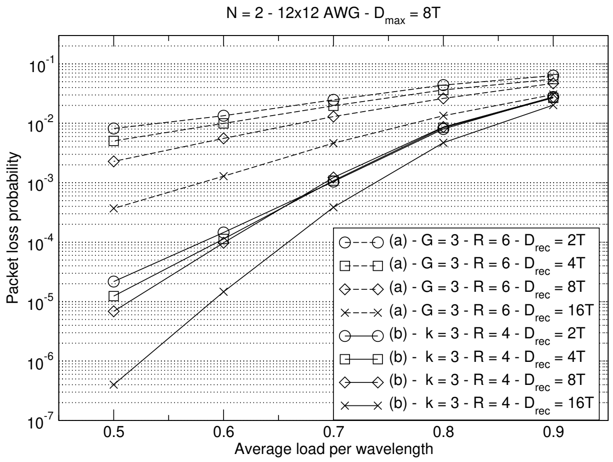

Figure 13. Packet loss probability of (a) and (b) structure with the same TWC count

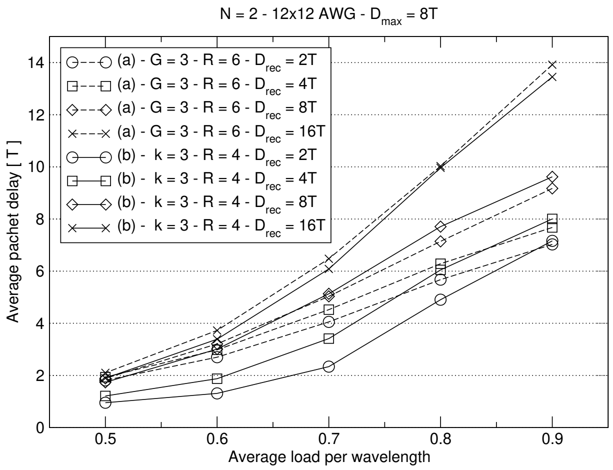

Figure 14. Average delivered packets delay of (a) and (b) structure with the same TWC count

of recirculation lines, structure (b) outperforms structure (a). As far as packet delay is concerned, the two structures give similar performance and structure (b) delays are always smaller for low offered traffic and larger for high offered traffic. Moreover, it is clear that, while improving node performance, the implementation of recirculation lines yields higher packet delays.

5. Conclusions

In this work, we have proposed an architecture for the switching core of an IP over WDM switching fabric. Starting from previous proposals of AWG-based optical switching node, it has been shown how to arrange the switch core in order to use fiber delay lines to implement a shared buffer. The results obtained through computer simulation are quite promising in that this new structure outperforms the original one. Other issues will have to be addressed in the future, such as the behavior of a switching node without input buffering and the number of recirculation lines needed to obtain reasonable performance anyway, or the behaviour of a shared-buffered node when dealing with packet flows belonging to different priority classes.

References

[1] Yoo, M. and Qiao, C.:Just-Enough-Time(JET): A High Speed Protocol for Bursty Traffic in Optical Networks., Proc. IEEE/LEOS Tech. for a Global Info Infrastructures, Aug. 1997, pp. 26-27.

[2] Qiao, C.: Labeled Optical Burst Switching for IP-over-WDM Integration, IEEE Commun. Mag., Sep. 2000, pp. 104-114.

[3] Guillemot, C. et al.: Transparent Optical Packet Switching: the European ACTS KEOPS Project Approach, IEEE Journal of Lightwave Technology, Dec. 1998, Vol. 16, No. 12, pp. 2117-2134.

[4] Hunter, D.K., Guild, D.K. and Bainbridge, J.D.: WASPNET: a Wavelength Switched Packet Network, IEEE Communication Magazine, Mar. 1999, Vol. 37, No. 3, pp. 120-129.

[5] Chia, M.C., Hunter, D.K. Andonovic, I., Ball, P., Ferguson, S., Guild K.M. and O'Mahony, M.J.: Packet Loss and Delay Performance of Feedback and Feed-Forward Arrayed-Waveguide Grating-Based Optical Packet Switches with WDM Inputs-Outputs, IEEE Journal of Lightwave Technology, Sep. 2001, Vol. 19, No. 9, pp. 1241-1254.

[6] Bregni, S., Guerra, G. and Pattavina, A.: Optical Packet Switching of IP Traffic, Proceedings of 6^{th} Working Conference on Optical Network Design and Modeling (ONDM), 2002.

[7] Bregni, S., Guerra, G. and Pattavina, A.: Node Architecture Design for All-Optical IP Packet Switching, *Accepted for publication in Proceedings of IEEE Globecom 2002*

[8] Dorgeuille, F., Mersali, B., Feuillade, M., Sainson, S., Slempkès, S. and Fouche, M.: Novel Approach for Simple Fabrication of High-Performance InP-Switch Matrix Based on Laser-Amplifier Gates, IEEE Photon. Techonl. Lett., 1996, Vol. 8, pp. 1178-1180

[9] Tzanakaki, A. and O'Mahony, M.J.: Analysis of Tunable Wavelength Converters Based on Cross-Gain Modulation in Semiconductor Optical Amplifiers Operating in the Counter Propagating Mode, IEEE Proc.-Optoelectron., Feb. 2000, Vol. 147, pp. 49-55.

[10] Mak, M.W.K. and Tsang, H.K.: Polarization-insensitive Widely Tunable Wavelength Converter Using a Single Semiconductor Optical Amplifier., IEEE Electron. Lett., 2000, Vol. 36, pp. 152-153.

[11] Parker, C. and Walker, S.D.: Design of Arrayed-Waveguide Gratings Using Hybrid Fourier-Fresnel Transform Techniques, IEEE J. Selec. Topics Quant. Electron., 1999, Vol. 5, pp. 1379-1384.

[12] Samsung Electronics: Fiberoptics products: Arrayed Waveguide Grating, http://www.samsungelectronics.com/fiberoptics/

[13] Bregni, S., Pattavina, A. and Vegetti, G.: Architecture and Performance of AWG-based Optical Switching Nodes for IP Networks, *Submitted for publication to JSAC special issue on High-Performance Optical/Electronic Switches/Routers for High-Speed Internet*

[14] Thompson, K., Miller, G.J. and Wilder, R.: Wide-Area Internet Traffic Patterns and Characteristics, IEEE Network Magazine, Nov. 1997, pp. 10-23.