Architectures and Performance of Optical Packet Switching Nodes for IP Networks

Achille Pattavina
Dept. of Electronics and Information, Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milan, Italy
pattavina@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. The traffic performance of such node when loaded by a typical IP traffic is evaluated through computer simulation; packet loss probability and average delay performance are shown for various load conditions.**

***Index Terms*— WDM network, optical switching, arrayed waveguide grating (AWG), IP packets, traffic performance.**

I. Introduction

Telecommunication 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, so as to improve transport network performance.

Two main alternative strategies have been proposed to reach this purpose: optical packet switching [1]–[5], and optical burst switching [6]–[8]. In this article, we describe an 'almost' all-optical switching architecture that uses an arrayed waveguide grating (AWG) as the packet router device. As usual, packet buffering is accomplished by fiber delay line units. Preliminary papers have already reported some traffic performance results for this switch, when the node supports packet buffering only at switch inputs (input queueing) [9] or by sharing delay lines among all switch inputs through packet recirculation (shared queueing) [10]. Here we present the overall switch structure in a deeper detail and evaluate the switch performance also with combined input/shared queueing.

[0]Work partially supported by MIUR, Italy, under FIRB project ADONIS.

In Section 2 we first introduce the basic concepts supporting the two types of switching, that is optical packet switching and optical burst switching. Then in Section 3 we introduce the optical network environment envisioned for the long-haul transport scenario, for which Section 4 describes the architecture of the AWG-based switching node. Traffic performance is evaluated in Section 5 for a single node supporting either pure input queueing or combined input/shared queuing.

II. Optical Packet and Burst Switching

Optical packet switching makes it possible 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. Two different basic approaches have been proposed to this aim, which differ in the switching matrix unit: *broadcast-and-select* architectures or *wavelength routing* architectures.

The first project based on broadcast-and-select principle is KEOPS [11]; Figure 1 shows its architecture that adopts input buffering. It is composed of two stages performing optical buffering and switching. In the first stage packets are delayed by a suitable amount of time in order to avoid collisions at the switch output ports; this function is accomplished by a set of tunable wavelength converters (TWC) whose task is to select the proper delay line to be accessed through the demultiplexers. Optical packets emerging from the multiplexers are given a new wavelength by the second set of TWCs so as to select the addressed switch outlet.

Since this solution does not allow packet recirculation, it cannot efficiently support different packet priorities, because, once a packet has been sent to a delay line, it cannot be stored longer than the fiber delay to eventually transmit a new packet with higher priority. This is a crucial shortcoming of this solution, since the need for some methods of providing differentiated classes of service for Internet traffic is growing, with the explosion of new possible applications. Actually the IPv4 TOS field or the IPv6 Traffic Class field are already used to give packets a particular forwarding treatment at each network node, and the availability in the network nodes of such feature is a fundamental requirement.

Recently another project has been proposed, further elaborating the broadcast-and-select solution: the DAVID [12]

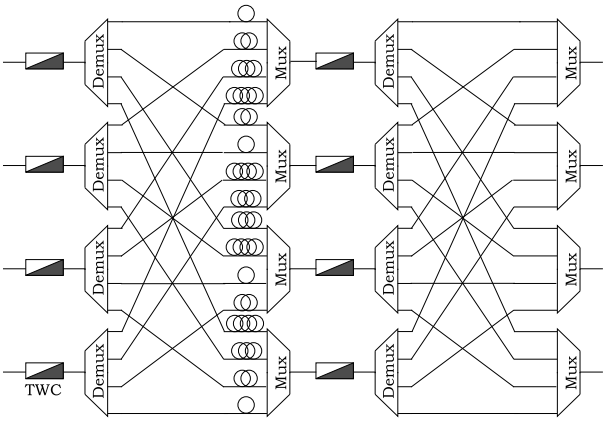


Fig. 1. KEOPS node architecture.

project proposes an optical networking solution viable for both metropolitan and wide-area networks.

Using a wavelength-routing device for optical packet switching has been proposed in the projects OPERA [13] and WASPNET switch [14]. In the former case only one arrayed waveguide grating (AWG) device is used for optical packet routing, whereas two AWGs are adopted in the latter project, which presents a more complete view from a networking standpoint. Figure 2 shows the WASPNET architecture that includes two switching stages. The first stage is used to route packets to the delay lines buffer, for contention resolution, or to the second stage. This latter stage is used to properly route packets to the desired output port. In both stages an AWG device is used to switch (route) packets to an outlet which is jointly identified by the AWG incoming port and the adopted transmission wavelength. TWCs at the AWG inputs perform this function, while the other TWCs feeding the demultiplexers are used to select the amount of recirculation delay. This architecture allowing packet recirculation accomplishes shared queueing, but the need for a second AWG to route packets to their addressed output link yields a considerable hardware overhead.

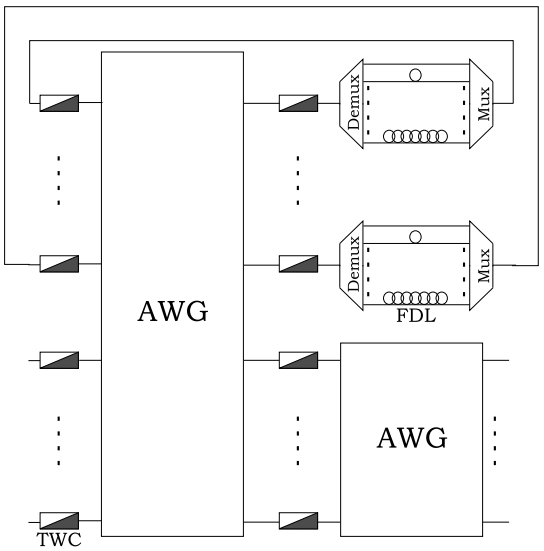


Fig. 2. WASPNET node architecture.

The systems presented insofar 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 availability.

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 Figure 3). This 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.

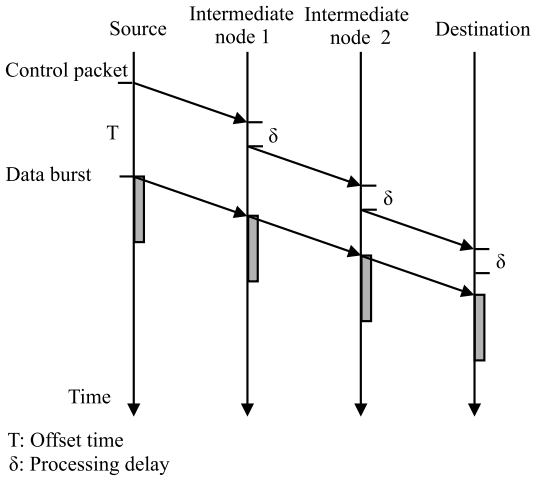


Fig. 3. The use of an offset time in optical burst switching.

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 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.

All the mentioned solutions accomplishing optical packet switching (KEOPS, WASPNET, DAVID, OPERA) define an environment suitable to switch fixed-length packets, whose transmission time is called *slot*, whose duration is in the order of 1 μs. This implies the development of complex segmentation and reassembly protocols at the optical network edges, if the offered traffic is composed of variable-length information units whose transmission time exceeds the slot time. On the other hand if the slot time is selected in such a way to fit the largest information unit, it is very likely that most of the slots will be partially used when small-size information

units are sent, thus wasting network resources.

The solution we propose here is to an optical packet switching node capable of switching variable-length optical packets; so the client layer (we assume IP in our scenario) can be interfaced more easily with the optical layer, thus avoiding a heavy packet processing overhead at the optical transport network edges. A *slot* concept is introduced also in our case, but here refers to the minimum size of optical packet that can be switched in a TCP/IP network environment.

III. Optical Transport Network Architecture

The architecture of the optical transport network we propose consists of M *optical packet switching nodes*, each denoted by an optical address made of $m = log_2\lceil M \rceil$ bits, which are linked together in a mesh-like topology. Edge systems (ES) interface the optical transport network with IP legacy (electronic) networks (see Figure 4).

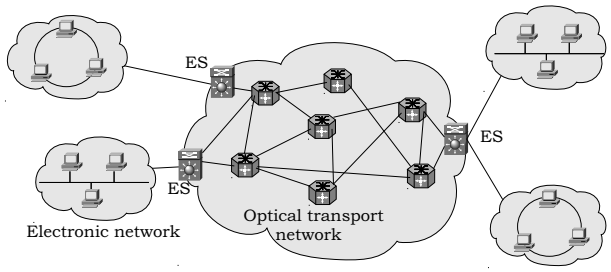


Fig. 4. The optical transport network architecture.

The optical packet is composed of a simple optical header, which comprises the m-bit destination address, and an optical payload containing an IP packet. In principle multiple IP packets could be packed in the same optical packet payload, if they are all addressed to the same ES. The optical packets are 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 the optical 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*; 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* (that is *slotted*), meaning that the optical packet switching must start at the beginning of a time slot. In the model we propose the time slot duration, T, to be equal to the amount of time needed to transmit an optical packet with a 40-byte payload from an input WDM channel to an output WDM channel. Such payload has been chosen as it is the minimum-size packet that can be transmitted in an IP-based network; actually it is an IP datagram transporting a TCP acknowledgement. Supposing a bit rate of 10 Gbps per wavelength channel, a 40 ns slot duration seems appropriate, since the 40-byte payload is transmitted in 32 ns (payload time, T_p) and the additional time can be used for the optical packet header transmission and to provide guard times.

Optically transporting variable-length packets in such a slotted environment is made possible by allowing an optical packet to engage several consecutive slots. We assume that the optical packet header needed for packet routing is present only in the first slot of the optical packet and that the payload per slot is always the same, that is $T_p = 32$ ns. For example an IP packet of 1500 bytes is transported by an optical packet engaging 38 slots. Therefore the bandwidth usage is kept under control by selecting carefully T and T_p.

Slotted operation has been assumed by all projects mentioned in the previous section, with slot duration equal to the time needed to transmit fixed-size packets. Only the project DAVID foresees the possibility of switching variable-length packets with a slot aggregation similar to that assumed here. Assuming slotted operation for variable-size packets compared to unslotted switching, makes simpler the switch control, for example handling the transition between different switch permutations. Other project proposals of optical packet switching with unslotted operations are not available in the technical literature, as far as the author's knowledge is concerned. In this paper we do not address the issue of comparing complexity or traffic performance between slotted and unslotted switching.

In our switch model a contention occurs every time 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 contending 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 result in paths longer than the minimum-distances.

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, eventually, 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 is lost or, if a suitable deflection algorithm is implemented, it can be deflected to a different output port than the correct one.

IV. Node Architecture

The general architecture of a network node is shown in Figure 5. It consists of N incoming fibers with W wavelengths per fiber. The incoming fiber signals are demultiplexed and G wavelengths from each input fiber are then 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 transmitted at the same time by each second-stage plane on the same output link. Apparently hardware simplicity requirements suggest to feed each plane with the same wavelengths from any input fiber. Nevertheless in principle there is no restriction in selecting the value of G, even if it will be shown that it has a significant impact on the switch traffic performance.

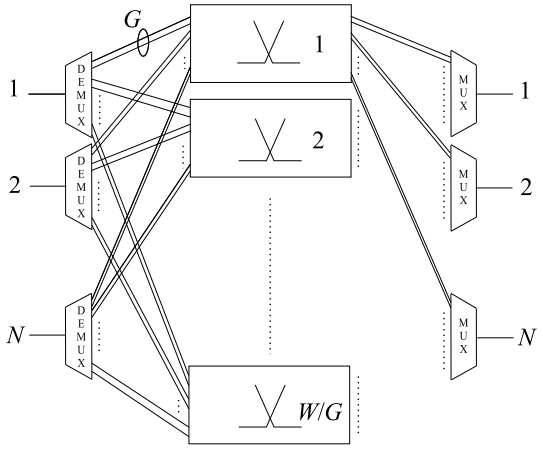


Fig. 5. Optical packet-switching node architecture.

The structure of one of the W/G parallel switching planes is presented in Figure 6. It interfaces single-wavelength input and output links and 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 (TWCs) in order to properly delay and route incoming packets in the second and third unit of the system, respectively.

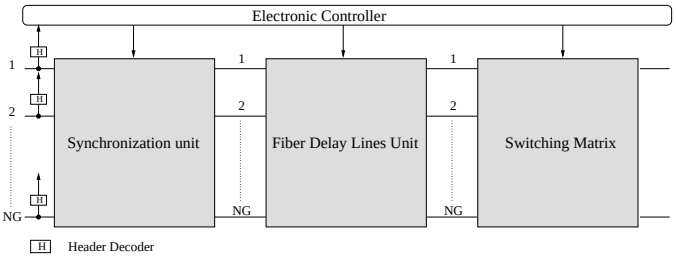


Fig. 6. Structure of one of the W/G parallel switching planes.

When packet recirculation is allowed, the AWG is used to switch packets to the output ports or, if necessary, to a recirculation port, in order to store them for an additional amount of time, to avoid collisions. Moreover recirculation ports allow the switch to support different priority classes, with service preemption. In fact, an optical packet, traveling through a recirculation port delay line, can always be preempted by a higher priority packet and be redirected to a recirculation port, instead of being transmitted.

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

A. Synchronization Unit

The synchronization unit is shown in Figure 7 and 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 by 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.

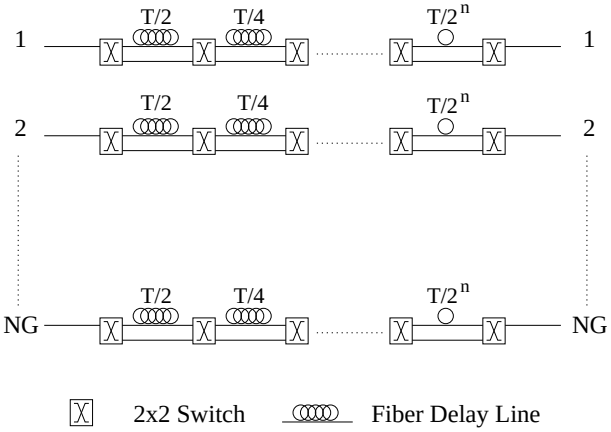


Fig. 7. Structure of the synchronization unit

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 [15], which have a switching time in

the nanosecond range, must be used. SOAs are all-optical amplification devices that are already used in a wide range of applications; they can be arranged in a particular structure (as shown in Figure 8), in order to achieve switching of optical signals. In this configuration SOAs are used as gates that let the signals go through or stop, depending on the permutation required. An interesting characteristic of SOA switches is that these devices allow the amplification of the traveling signals making it possible, besides routing functionalities, to restore a required given signal level.

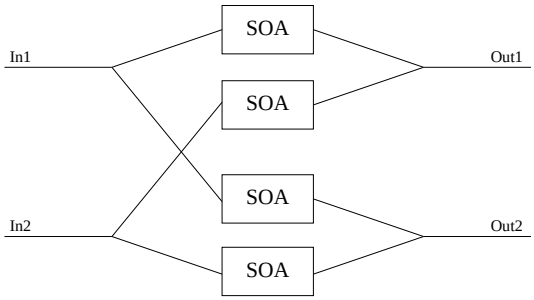


Fig. 8. 2×2 SOA switch.

B. 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 (TWCs), in order to deliver each packet to the correct delay line to resolve contentions (see Figure 9). On each of the NG inputs of the plane a delay can be applied that is multiple of the basic slot duration T and ranges up to D_{max} slots. An optical packet can be stored for a time slot, with a 40 ns duration, in about 8 meters of fiber at 10 Gbps. To achieve wavelength conversion several devices are available [16]–[19].

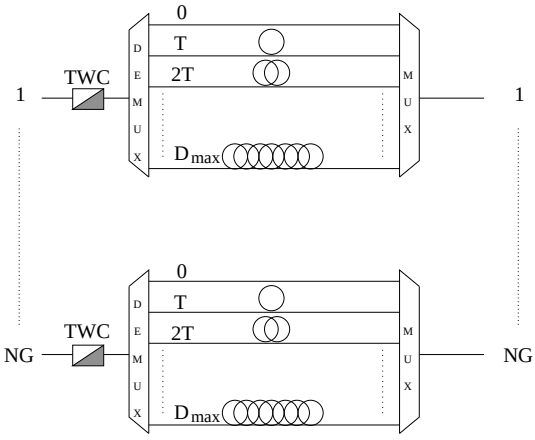


Fig. 9. Structure of the fiber delay unit

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 αT 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 βT, being $\beta > \alpha$. For example, suppose that packet P_1, of duration $l_1 T$, must be delayed d_1 time slots, in order to be transmitted onto the correct output port. This packet will then leave the optical scheduler at time $(\alpha + d_1)T$. 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 $\beta + d_2 + l_2 < \alpha + d_1$ 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 αT, 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 could be delayed for $d_2 < d_1$ slots to avoid conflict, its destination output port will be idle for $d_1 - d_2$ time slots, while there would be a packet to transmit.

C. 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 [20], as is shown in Figure 10a.

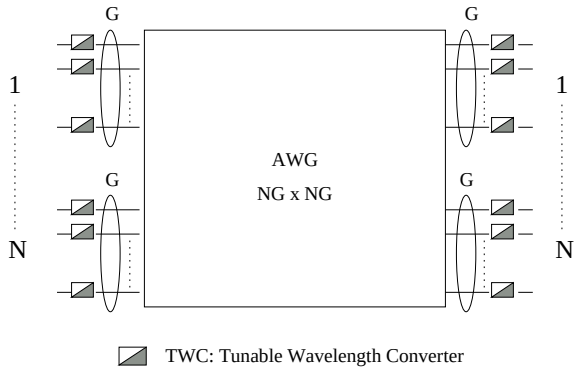
This device consists of two slab star couplers, interconnected by an array of waveguides. Each grating waveguide has a precise path difference with respect to its neighbors, ΔX, and is characterized by a refractive index of value n_w.

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 traveling through each couple of adjacent waveguides emerges with a phase delay difference given by:

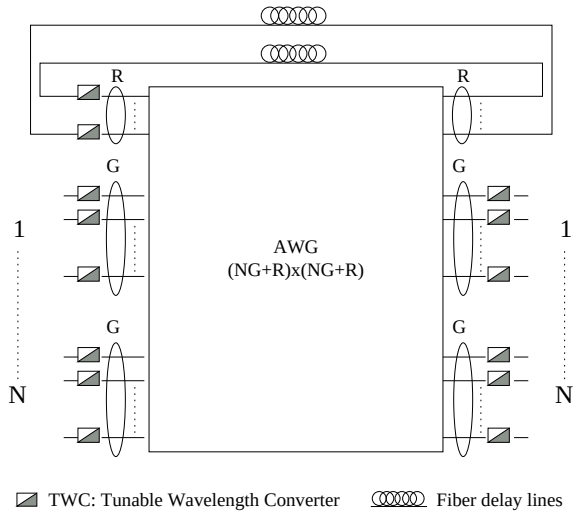
$$\Delta\phi = 2\pi n_w \times \frac{\Delta X}{\lambda}$$

where λ is 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, depending only on the input signal central wavelength.

Figure 11 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, the



(a)



(b)

Fig. 10. Structure of the switching unit: input queueing (a); combined input-shared queueing (b)

wavelength and the input port of the signals being the only parameters determining the switch permutation.

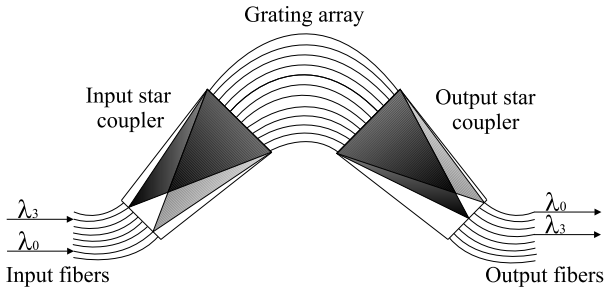


Fig. 11. Arrayed waveguide grating.

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. Commercially available 40 channel devices have a channel spacing of 100 GHz and show a typical insertion loss of 7.5 dB.

As we said before, to improve the system performance and to eventually support different priority classes with service preemption, some of the AWG ports can be reserved to allow packet recirculation (see Figure 10(b)). To this purpose R AWG output ports are connected, via fiber delay lines, to R input ports. Packet recirculation is then managed using tunable wavelength converters.

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.

V. Simulation Results

We present now some simulation results of the operation of an optical node with a single switching plane and hence $W = G$ is always assumed. These results have been obtained assuming that the node receives its input traffic directly from N edge systems. The edge system buffer capacity is supposed to be large enough to make packet loss negligible and each WDM channel is supposed to have a dedicated buffer in the edge system.

The packet arrival process has been modelled as a Poisson process, with packet interarrival times having a negative exponential distribution. As the node operation is slotted, the optical packets duration is 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-byte payload, from an input WDM channel to an output WDM channel.

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

1) *Empirical distribution*. Based on real measurements on IP traffic [21], [22], we have assumed the following probability distribution for the packet length, L:

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

In this model, packets have average length equal to 393 bytes. Since a 40-byte 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.

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 760 bytes (duration $19T$). Also in this model, packets have average duration of $10T$.

No deflection routing algorithm has been implemented. Under this assumption, a packet is supposed to be lost if it cannot be delayed by a suitable amount of time, in order to transmit it onto the correct output port, on any of the G available wavelengths. We will now present the performance results of both architectures, with and without packet recirculation ports in the AWG, remarking that all the plotted values have a 95% confidence interval not larger than 40% of the plotted values.

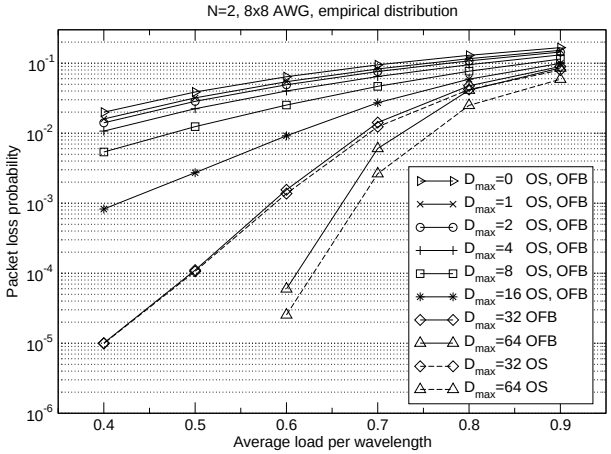


Fig. 12. Packet loss probability for the empirical distribution: OS vs. OFB.

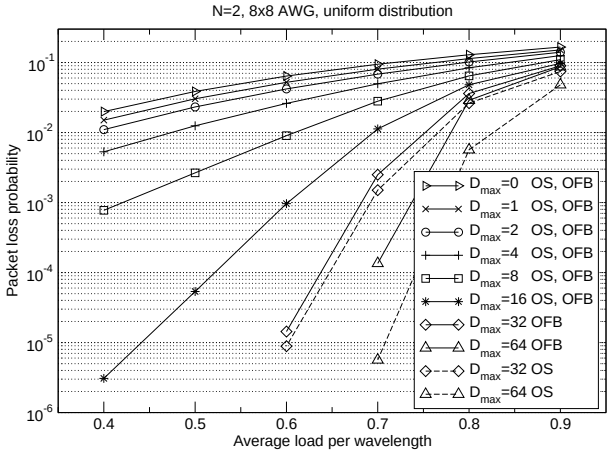


Fig. 13. Packet loss probability for the uniform distribution: OS vs. OFB.

A. Optical node without packet recirculation

A node without recirculation line is considered first in which we evaluate the effect on traffic performance of the two managing policies optical scheduling (OS) and the optical FIFO buffering (OFB). Figures 12 and 13 show the packet loss probability at different traffic loads per wavelength for various values of the input line delay D_{max} of an 8×8 AWG, with $N = 2$ and $W = G = 4$ wavelengths per fiber, for the empirical and uniform distributions. It can be seen that, regardless of packet length distribution, the OS policy yields a better performance than the OFB policy only when the maximum delay, D_{max}, becomes large enough to allow efficient packet scheduling. It must be remarked that an optical packet P_2, of duration l_2T entering the node at time βT from the i-th WDM input channel, can be transmitted before a packet P_1, entering the node on the same input channel at time $\alpha T < \beta T$, only if D_{max} is large enough to avoid collision at the fiber delay lines output, that is $D_{max} \geq (\beta + l_2 + d_2 - \alpha)T$. Further performance improvements are expected by equipping longer delay lines on the input side.

Figure 14 shows the packet loss probability for the empirical and uniform distributions, which have an L_{max} value of $38T$ and $19T$ respectively, for the OFB policy. It can be pointed out

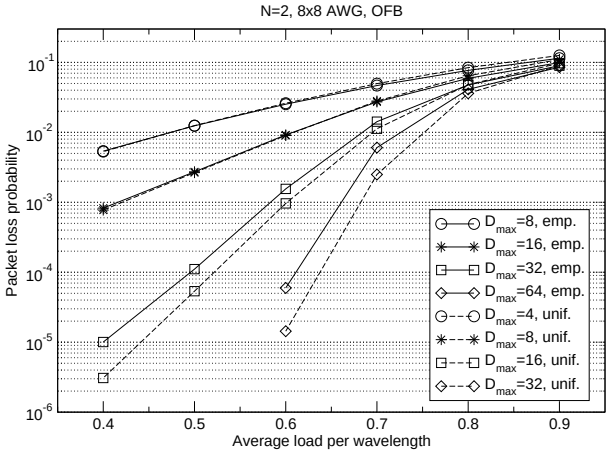


Fig. 14. Packet loss probability for the empirical and uniform distributions.

that, regardless of packet length distribution, the node almost shows the same loss probability for the same value of the D_{max}/L_{max} ratio. Furthermore, performance improves as this ratio increases. In a variable packet length environment, then, it is convenient to use the OFB policy for the fiber delay lines unit management, as it is simpler to implement than the OS policy and gives almost the same performance.

We evaluate how 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 under optical scheduling operation. To this aim we have selected a node architecture with a single switching plane ($W = G$). We have 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 $N \cdot G = 32$. Figure 15 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").

B. Optical node with packet recirculation

Here we present the simulation results for an optical switching plane with R recirculation ports with OS policy. Two different structures for the recirculation delay lines have been tested: 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, which, unless stated otherwise, is given by $D_{rec} = T$. In the VDR structure D_{rec} doubles every two ports, that is the first couple of ports will then have a recirculation delay of T, the second couple of $2T$, and so on. Such structure of delay lines for the VDR case has been selected to enable

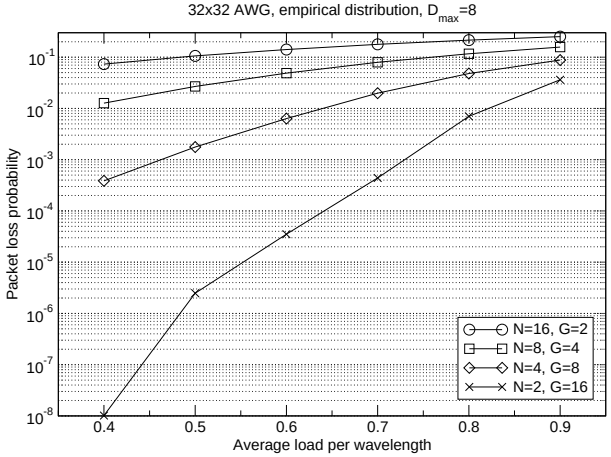


Fig. 15. Packet loss performance for different grouping factor and node size values.

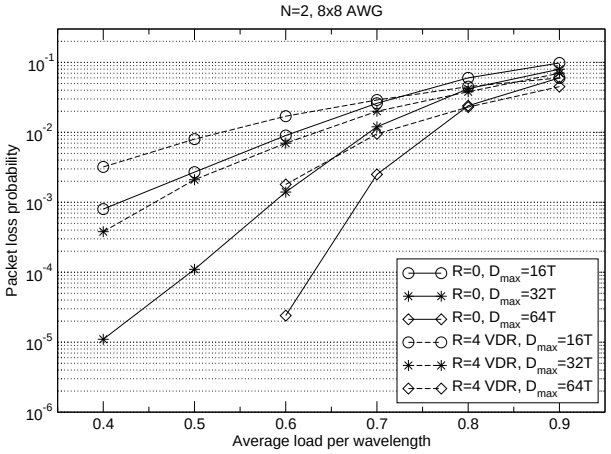


Fig. 17. Packet loss probability: 8×8 AWG, $R=0$ vs. 8×8 AWG, $R=4$ VDR.

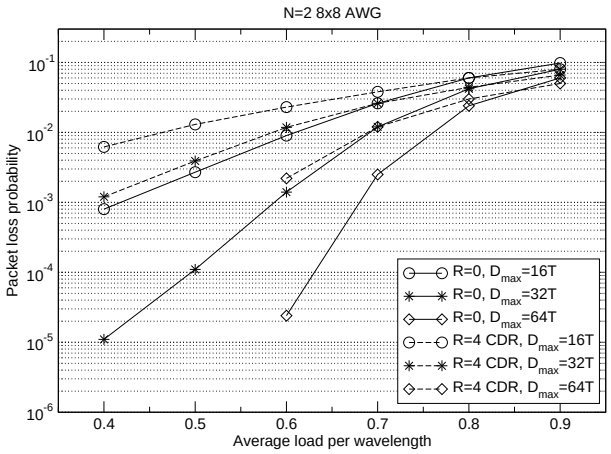


Fig. 16. Packet loss probability: 8×8 AWG, $R=0$ vs. 8×8 AWG, $R=4$, CDR.

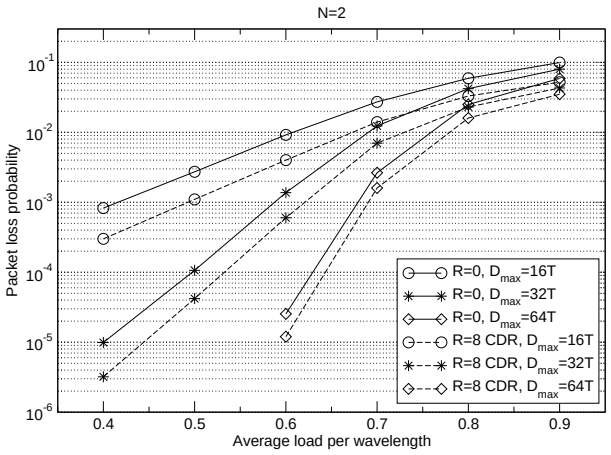


Fig. 18. Packet loss probability: 8×8 AWG, $R=0$ vs. 16×16 AWG, $R=8$, CDR.

the switch performance evaluation with recirculation lines of reasonable length for the AWG sizes considered here, e.g. up to $D_{rec}=16T$. Moreover multiple recirculations are allowed only if the packet duration L is lower than the recirculation delay to prevent long packets to occupy simultaneously more than one recirculation port.

Figures 16 through 19 show the packet loss probability of an optical switching node with recirculation ports ($R>0$) and compare them with a structure without recirculation ports ($R=0$), with different values of D_{max} for the empirical distribution.

Figures 16 and 17 plot the loss probability of an 8×8 AWG, with $R=0$ and $R=4$, for constant delay recirculation (CDR) and variable delay recirculation (VDR) structures, respectively. As the AWG dimension does not change, the system with the recirculation ports always gives higher loss probabilities than the other one since, in order to make packet recirculation possible, the grouping factor G has to be reduced, reducing the number of packets that can be transmitted at the same time on one output link.

Figures 18 and 19 show the comparison between an 8×8 AWG, with $R=0$, and a 16×16 AWG, with $R=8$. Now, as the grouping factor, G, does not change, the packet recirculation effect on the system performance is apparent: larger buffers in the recirculation lines (adopted in the VDR case) improve the packet loss performance. As the introduction of the recirculation ports allows a longer packet storage, the average delay grows (see Figure 20), yielding a lower loss probability for both the CDR and VDR structures.

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

Figure 21 shows 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 $D_{rec}=\{T,4T,16T\}$), 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.

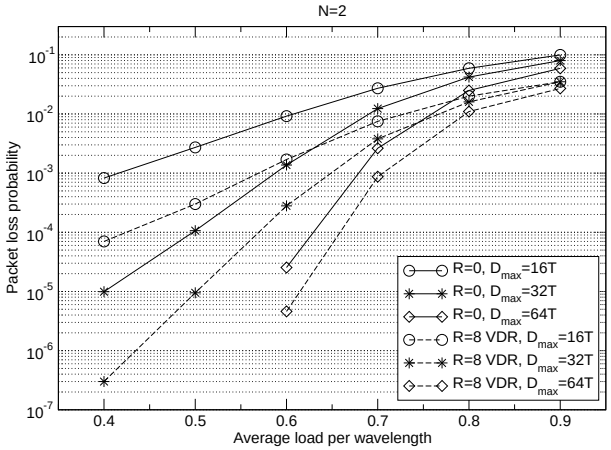


Fig. 19. Packet loss probability: 8×8 AWG, $R = 0$ vs. 16×16 AWG, $R = 8$, VDR.

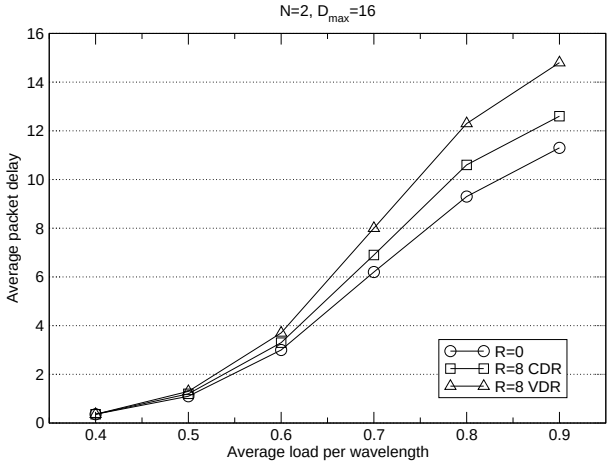


Fig. 20. Average packet delay at different loads per wavelength for an 8×8 AWG, $R = 0$ and for a 16×16 AWG, $R = 8$, with CDR and VDR structures.

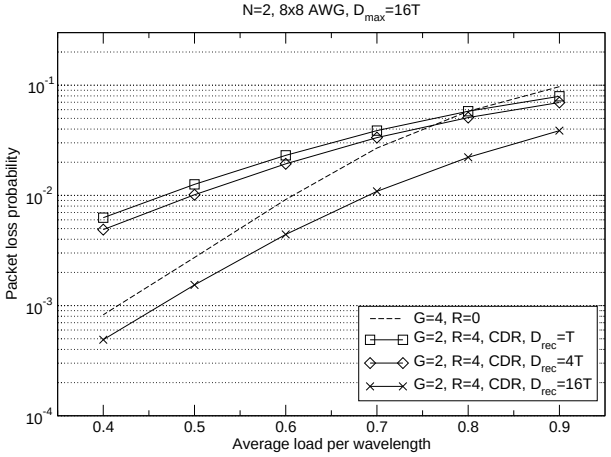


Fig. 21. Packet loss performance for different grouping factor values, with and without recirculation lines ($R = 0$ vs. $R = 4$ CDR).

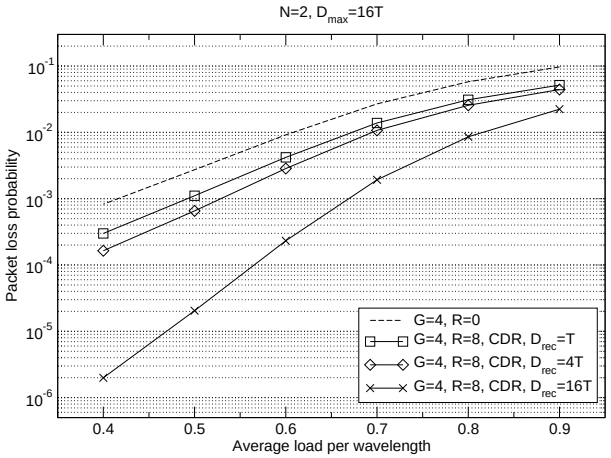


Fig. 22. Packet loss performance for the same grouping factor value, with and without recirculation lines ($R = 0$ vs. $R = 8$ CDR).

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. Figure 22 shows that, as the grouping factor does not 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.

VI. Conclusions and Topics for Further Research

A node architecture for optical packet-switched transport networks has been proposed that is based on an AWG device. Packet buffering is made possible by fiber delay lines accomplishing either input queueing only, or combined input/shared queueing. It has been shown that, unless the input buffer length exceeds the maximum packet size, optical scheduling and optical FIFO buffering give almost the same performance. On the other hand, when the input queue can hold at least one packet of maximum size, optical scheduling yields a better performance than optical FIFO buffering, because the output links can be more efficiently exploited. Adding shared buffering through recirculation lines can improve the system performance only if the grouping factor G is not reduced.

Many issues will have to be addressed in the future, such as the improvement attainable with the introduction of different priority classes adopting service preemption. Moreover, the behavior of a meshed optical transport network will have to be investigated, in which deflection routing policies can be adopted to enhance the overall system performance.

References

[1] S. Yao, B. Mukherjee, and S. Dixit. Advances in Photonic Packet Switching: An Overview. *IEEE Communications Magazine*, 38(2):84–94, Feb 2000.

[2] D. K. Hunter and I. Andonovic. Approaches to Optical Internet Packet Switching. *IEEE Communications Magazine*, 28(9):116–122, Sep. 2000.

[3] L. Xu, H. G. Perros, and G. Rouskas. Techniques for Optical Packet Switching and Optical Burst Switching. *IEEE Communications Magazine*, 39(1):136–142, Jan. 2001.

[4] M. J. O'Mahony, D. Simeonidou, D. K. Hunter, and A. Tzanakaki. The Application of Optical Packet Switching in Future Communication Network. *IEEE Communications Magazine*, 39(3):128–135, Mar. 2001.

[5] M. A. Bourouha, M. Bataineh, and M. Guizani. Advances in Optical Switching and Networking: Past, Present, and Future. In *Proceedings of IEEE SoutheastCon 2002*, pages 405–413, 2002.

[6] C. QIao and M.Yoo. Optical Burst Switching OBS - A New Paradigm for an Optical Internet. *Journal of High Speed Networks*, 8:69–84, 1999.

[7] C. Qiao. Labeled Optical Burst Switching for IP-over-WDM Integration. *IEE Commun. Mag.*, pages 104–114, Sep. 2000.

[8] M. Yoo and C. Qiao. Optical Burst Switching for Service Differentiation in the Next Generation Optical Internet. *IEEE Communication Magazine*, 32(2):98–104, Feb. 2001.

[9] S. Bregni, G. Guerra, and A. Pattavina. Optical Switching of IP Traffic Using Input Buffered Architectures. *Optical Network Magazine*, 3(6):20–29, Nov.-Dec. 2002.

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

[11] C. Guillemot et al. Transparent Optical Packet Switching: the European ACTS KEOPS Project Approach. *IEEE Journal of Lightwave Technology*, 16(12):2117–2134, Dec. 1998.

[12] L. Dittman et al. The European IST Project DAVID: A Viable Approach Toward Optical Packet Switching. Architecture and Performance of AWG-based Optical Switching Nodes for IP Networks. *J. on Selected Areas in Commun.*, 21(7): 1026–1040, Sept. 2003.

[13] A. Carena, M.D. Vaughn, R. Gaudino, M. Shell, D.J. Blumenthal. An Optical Packet Experimental Routing Architecture with Label Swapping Capability. *IEEE Journal of Lightwave Technology*, 16(12):2135–2145, Dec. 1998.

[14] D. K. Hunter, K. M. Guild, and J. D. Bainbridge. WASPNET: a Wavelength Switched Packet Network. *IEEE Communication Magazine*, 37(3):120–129, Mar. 1999.

[15] F. Dorgeuille, B. Mersali, M. Feuillade, S. Sainson, S. Slempkès, and M. Fouche. Novel Approach for Simple Fabrication of High-Performance InP-Switch Matrix Based on Laser-Amplifier Gates. *IEEE Photoncs Technology Letters*, 8:1178–1180, 1996.

[16] M.W.K. Mak and H.K. Tsang. Polarization-insensitive Widely Tunable Wavelength Converter Using a Single Semiconductor Optical Amplifier. *IEE Electronics Letters*, 36:152–153, 2000.

[17] 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. In *IEE Proceedings-Optoelectronics*, volume 147, pages 49–55, Feb. 2000.

[18] I. White, R. Penty, M. Webster, Y. J. Chai, A. Wonfor, and S. Shahkooh. Wavelength Switching Components for Future Photonic Networks. *IEEE Communications Magazine*, 40(9):74–81, Sep. 2002.

[19] O. A. Lavrova, L. Rau, and D. J. Blumenthal. 10-Gb/s Agile Wavelength Conversion With Nanosecond Tunig Times Using a Multisection Widely Tunable Laser. *IEEE Journal of Lightwave Technology*, 20(4):712–717, Apr. 2002.

[20] C. Parker and S.D. Walker. Design of Arrayed-Waveguide Gratings Using Hybrid Fourier-Fresnel Transform Techniques. *IEE Journal on Selected Topics in Quantum Electronics*, 5:1379–1384, 1999.

[21] K. Thompson, G. J. Miller, and R. Wilder. Wide-Area Internet Traffic Patterns and Characteristics. *IEEE Network Magazine*, pages 10–23, Nov. 1997.

[22] Generating the Internet Traffic Mix Using a Multi-Modal Length Generator. Spirent Communications white paper at http://www.netcomsystems.com.