Scattered Versus Shared Wavelength Path Operation, Application to Output Buffered Optical Packet Switches. A comparative study

P. Pavon-Marino, J. Garcia-Haro, J. Malgosa-Sanahuja, F. Cerdan

Department of Information Technologies and Communications

Polytechnic University of Cartagena. Campus Muralla del Mar s/n, Cartagena. E-30202. Spain.

Phone: +34 968325952 Fax: +34 968325338

E-mail: {pablo.pavon, joang.haro, josem.malgosa, fernando.cerdan}@upct.es

Abstract: Commercial application of Optical Packet Switching (OPS) to backbone WDM networks is far from reality mainly due to hardware costs. The two envisaged alternatives for network control and operation of this type of networks, proposed under the WASPNET project, are Scattered Wavelength Path (SCWP) and Shared Wavelength Path (SHWP). The advantages and drawbacks of both approaches involve many different issues, some of them still open. In this paper, the impact of SCWP and SHWP switching modes on performance and hardware costs is discussed and evaluated for three different prominent OPS switch fabric designs: KEOPS switch, space switch, and output-buffered wavelength-routed switch. To achieve it, we conduct a necessary set of modifications for these architectures, which were not originally conceived with SCWP/SHWP modes in mind. This normalization process is unavoidable to provide a fair comparison among the switch fabric designs, and to advance to a global comparison between SCWP and SHWP operational modes, which to the knowledge of the authors does not exist yet.

Keywords: Optical Packet Switching, Performance Evaluation, High Performance Networking and Protocols

1 Introduction

The impressive annual growth of Internet traffic is demanding and spurring more and more advances in telecommunications technology. Regarding core networks, one of the greatest progresses has been Wavelength Division Multiplexing (WDM) application to already installed fiber links, magnifying available transmission bandwidth. The problem of the sharing of this huge amount of bandwidth follows three different switching strategies [1]: Wavelength Routing, Optical Burst Switching, and Optical Packet Switching. The highest channel utilization due to statistical multiplexing, and the most direct mechanism to adapt to Internet traffic is achieved by the third alternative: Optical Packet Switching (OPS). Under the optical domain, OPS is similar to traditional electronic packet switching, except that packet payload transparently remains in the optics, while its header is processed electronically. This paradigm yields much more flexibility and bandwidth efficiency since it operates on the

granularity of a packet. On the other hand, fast packet-by-packet switching operation imposes the highest constraints to the photonic switching function, involving very large hardware costs under the state-of-the-art technology. For this reason, the deployment of an OPS backbone network is not envisaged up to the medium term.

The conceptual architecture for an OPS network is based on a set of Optical Cross-connect (OXC) nodes, interconnected by WDM links, as shown in figure 1. Edge routers of the backbone network are the source and termination of optical packets, which, in this paper, are assumed to be of fixed size. Electronic terabit IP switch routers with the necessary multi-protocol operation and a large electronic buffering are foreseen to be the winning products in this market segment [2]. Neither the mapping of variable size packets (e.g. IP datagrams) over these fixed length frames nor the frame-size have been optimized yet. The best performance would be obtained for small packet segmentation, which is carried out in the network edge nodes. However, packet segmentation is limited by the efficiency reduction due to packet header and guard times to payload ratio, and by the packet processing time. A realistic payload duration of *1*μ*s* transports *5,000* bytes in a *40* Gbps channel, which is an order of magnitude larger than the largest IP datagram generated by most applications. Therefore, packet aggregation is considered for OPS ingress edge nodes in a similar way as burst aggregation exists in Optical Burst Switching.

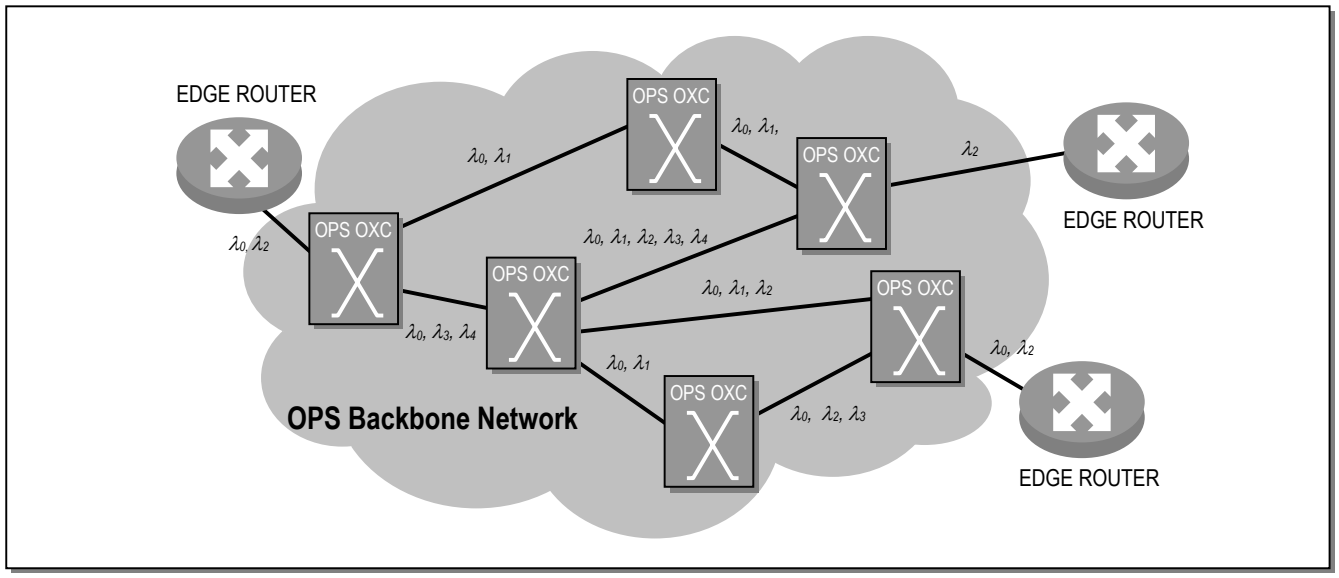
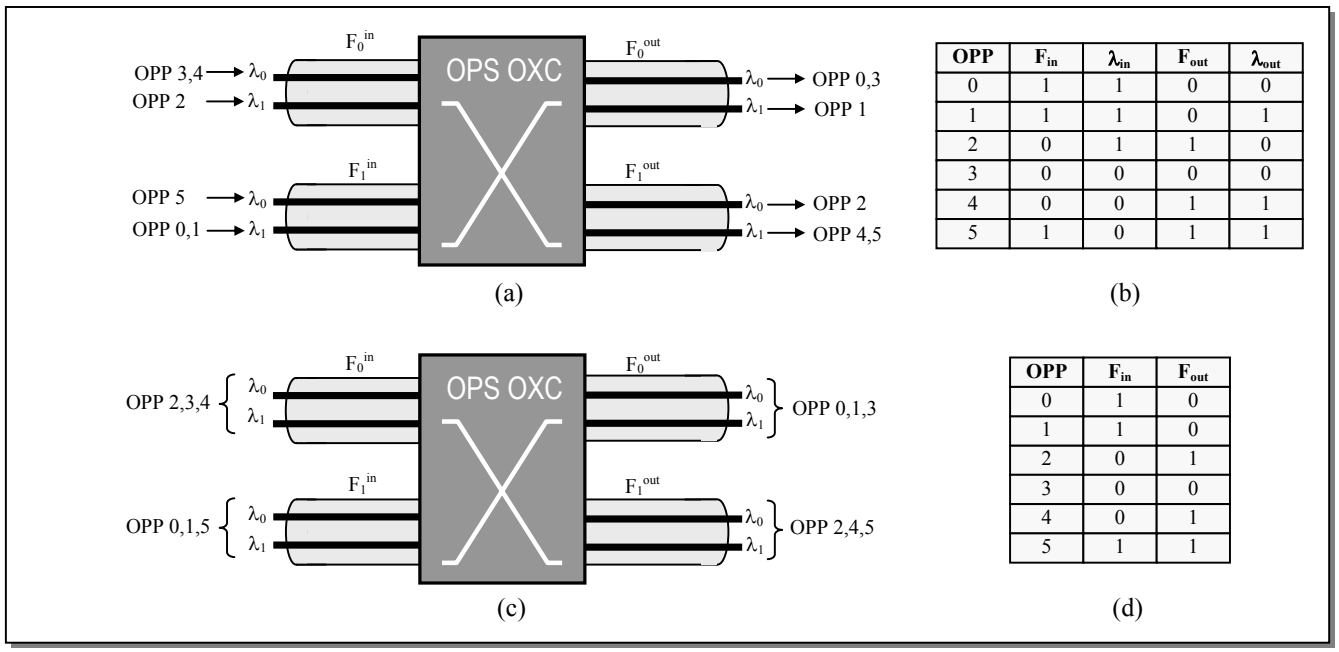


Figure 1. Example of Optical Packet Switching network architecture

It is generally accepted that higher layer traffic connections between edge routers of the OPS backbone network would be established as a type of optical permanent virtual circuits (in the following Optical Packet Paths -OPP). Attending to the virtual circuit concept, packets belonging to the same OPP follow a fixed sequence of hops from ingress to egress node. The establishment of this hop sequence is performed during OPP provisioning. The efficient use of bandwidth attained by OPS in

WDM networks results in a number of simultaneous OPPs that can be provisioned through a link much greater than the number of wavelengths. It is a responsibility of the network control and operation to decide on the most proficient method to distribute traversing OPPs traffic among the wavelengths of the links. This subject has been addressed under the WASPNET project [3-5], where two possible methodologies for OPS were proposed: The Scattered Wavelength Path (SCWP) and the Shared Wavelength Path (SHWP). In SHWP, OPP transmission fiber and transmission wavelength are statically fixed for each hop. Under these considerations, each packet entering an OPS node requests to switch control a fixed output fiber and a fixed output wavelength, whose values are stored in a look-up table on OPP provisioning (figure 2-a,b). Unlike SHWP, in SCWP optical packet paths determine transmission fibers during OPP establishment, but do not fix transmission wavelength in each link. Therefore, a switching node is free to dynamically transmit packets belonging to the same OPP over any wavelength of the required output fiber, according to internal concerns like buffer occupancy optimization (figure 2-c,d).



OPP	F_{in}	λ_{in}	F_{out}	λ_{out}
0	1	1	0	0
1	1	1	0	1
2	0	1	1	0
3	0	0	0	0
4	0	0	1	1
5	1	0	1	1

OPP	F_{in}	F_{out}
0	1	0
1	1	0
2	0	1
3	0	0
4	0	1
5	1	1

Figure 2. Example of SHWP/SCWP operation modes (a) SHWP switch OPPs provision, (b) SHWP switch lookup table, (c) SCWP switch OPPs provision, (d) SCWP switch lookup table

SCWP and SHWP are the foreseen competing candidates for a future deployment of a WDM OPS backbone network, like the one shown in figure 1. The adoption of one of them should be preceded by a deep analysis involving many factors. SCWP offers higher throughput than SHWP because of the flexibility to spread input packets across the wavelength dimension, yielding to lower packet loss probability and requiring smaller optical buffers at the switch. On the other hand, SCWP requires

a potentially complex scheduling decision on output wavelength instead of a fixed table lookup. Also, SCWP control operation may add extra restoration complexity, as explained in [5].

In this paper we study the influence of SHWP and SCWP operation modes on the design of three OPS switch fabric architectures, in terms of performance and hardware complexity evaluation. These three architectures are (1) broadcast-and-select KEOPS switch [6], (2) space switch [7], and (3) wavelength-routed output buffered switch [8]. They have been selected as representative approaches of current output buffered OPS switch fabrics. These architectures, as well as others referenced in the open literature (but the ones proposed under WASPNET project), were designed without any of the SCWP or SHWP mode in mind. Consequently, a normalization process to a unifying model involving, when required, several modifications of the original architectures is needed. We believe this adaptation of switch fabric designs per SCWP/SHWP operation mode should not be observed as a distortion, but a way to achieve a valuable and necessary comparison among competing architectures.

The remainder of this paper is organized as follows. Section 2 focuses on the switch fabric normalization process, along with its application to OPS output buffered architectures. The three selected switch fabric designs are described in section 3. Section 4 provides the performance evaluation results for both SHWP and SCWP output buffered versions. In section 5, a comparison of hardware costs of the different architectures is also presented. Finally, section 6 concludes the paper.

2 Output buffered WDM switch fabrics

In a context of disparity of OPS switch proposals and a 'hand-made' state-of-art, a need for normalization of designs arises under SHWP/SCWP operational schemes. This fact contributes to fairly compare equally controlled switches among them, and to provide a global comparison between SCWP and SHWP operation modes, allowing and easing, in its turn, the evaluation of switch fabric performance and the estimation of hardware costs. In order to incorporate a switch architecture to this evaluation process, three specification steps should be accomplished:

(1) Non-WDM architectures should be adapted to the multiwavelength input and output ports environment, to allow the application of SHWP/SCWP operation modes. The required hardware modifications to adapt to the WDM environment should be made, and computed in a hardware cost balance.

(2) SHWP scheduling process should be specified, constrained to the objective of maximizing switch performance for a given distribution of input packets, and a given output fiber and wavelength demand.

(3) A wavelength assignment algorithm should be determined for deciding on packet output wavelengths under SCWP operation mode.

In this paper, this normalization process is followed for the evaluation of three prominent output buffered OPS architectures: KEOPS switch, space switch and wavelength-routed output buffered switch. These architectures, together with the required hardware modifications associated to step (1), are depicted in section 3. In the next paragraphs, specification of the SHWP scheduling process and the output wavelength assignment algorithm for the three architectures are derived from the assumption of slotted operation and output buffered behavior, i.e. constant size packets are buffered in independent queues of fixed size, one per output fiber and wavelength.

Our goal for step 2, is to establish a scheduling process which optimizes switch throughput and average packet delay, preserving packet sequence, at an acceptable cost. In SHWP WDM networks end-to-end packet sequence is preserved by assuring that no internal packet misorder is introduced inside each switching node. Under these concerns, first-in-first-out (FIFO) queuing is the obvious optimum solution for SHWP output buffered architectures. In delay-loop based switches, like the ones considered in this paper, this is obtained by a per-output-port-counter which assigns packet delay in an increasing manner.

The specification of SCWP output wavelength assignment algorithm (step 3) has also the requirements of optimizing switch throughput and average delay, while maintaining packet sequence. The problem of end-to-end packet sequence preservation in SCWP OPS networks is not trivial, as the undefined wavelength path may result in out-of-order packets within the OPP arriving at destination nodes. The origin of this problem is that consecutive packets from the same OPP arriving a node in consecutive time slots, may be scheduled to leave the OPS switch in the *same* time slot, but using different output wavelengths, due to buffer occupancy. This is, for instance, the case of packets 0_2 and 0_3 of OPP *0* in the example shown in figure 3-a,b,c, which assumes OPP provisioning as displayed in figure 2-d.

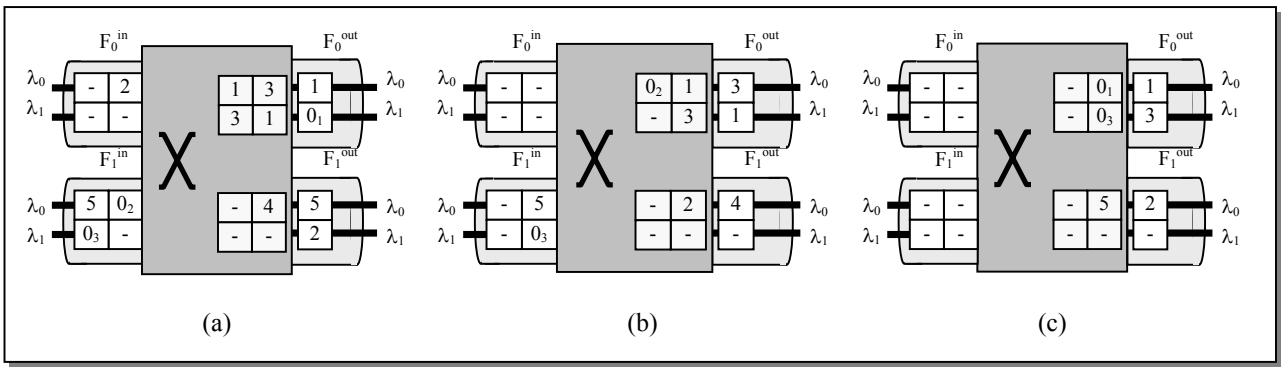


Figure 3. Example of packet sequence control in SCWP switching. (a) Initial switch state and packet arrivals, (b) packet 0_1 is scheduled to the lowest available wavelength λ_0, (c) simultaneous transmission of 0_1 and 0_2

In the limit, all of the output packets for a given fiber and a given time slot might belong to the same OPP. Then, a mechanism is required so that the next OPS node recognizes which is the order among simultaneous arrivals of input packets associated to the same OPP and therefore, the packet sequence could be maintained by the scheduler.

The packet misorder issue in SCWP networks has been described in [5], and an algorithm to sequence packets without the requirement of a global packet counter was provided for the feedback version of the WASPNET switch [3]. A per-OPP packet header counter solution is not desirable as it increases header size and adds management complexity to the network. The scheme in [5] lies on transmitting simultaneous packets of the same OPP ordered by their output wavelength, so that lower order packets are transmitted by using lower wavelengths. Given p_i, p_j, $i<j$, packets associated to the same OPP, the method requires that: (1) p_i is transmitted before p_j, or (2) p_i is transmitted during the same time slot as p_j, and $\lambda_i<\lambda_j$. End-to-end packet sequence is then maintained for the OPP, if the traversed OPS nodes satisfy the aforementioned conditions.

The pseudocode of the algorithm proposed in this paper for output wavelength assignment is written in figure 4, for a symmetric switch of N input and output fibers, n wavelengths per fiber, and M buffer positions. It is a simplification of the one presented in [5], accommodated for non-deflection routing output buffered switches. The algorithm is based on a round-robin assignment of output wavelengths from λ_0 to λ_{n-1}, for packets addressed to the same output fiber. Also, the round-robin pointer is reset when the system composed of the n queues is empty. As a result, a packet is assigned a delay d and a wavelength w if and only if queues $0..w-1$ have d $(0 \le d \le M-1)$ packets stored, and queues $w..n-1$ have $max(0,d-1)$ packets stored. Under these considerations, a packet is always assigned the minimum available delay. Also, throughput is maximized, since a packet is lost when the n queues corresponding to the n output wavelengths are full, and $M \cdot n$ packets are stored into the buffer. Furthermore, packet order is also preserved, in the terms described in this section: Input ports are examined sequentially, and lower wavelength input ports are allocated output wavelengths first. As an example, this algorithm is illustrated in figure 3, where packets 0_2 and 0_3 have to be transmitted in the same time slot, satisfying that $\lambda(0_2) < \lambda(0_3)$.

```
/* N = number of input/output fibers */
/* n = number of wavelengths per fiber */
/* M = buffer depth */

for input i = 0 to n·N-1 do
   if (packet is present on input i) then
     f = output fiber associated to input packet opp
     if (delay [f] < M) then
        associate delay [f] to packet
        associate output wavelength wav [f] to packet
        /* wav [f] is a Round-Robin pointer */
        wav [f] ++
        if (wav [f] == n)
           wav [f] = 0
          delay [f] ++
        endif
     endif
   endif
endfor

/* decrease every delay [f], f=0..N-1 after each time slot */
for output fiber i=0 to N-1 do
  delay [f] = max (0 , delay [f]-1)
  if (delay [f] == 0)
     wav [f] = 0 /* reset RR pointer if no packets are stored into this output fiber */
  endif
endfor
```

Figure 4. Round-Robin SCWP output wavelength assignment algorithm (pseudocode)

3 Switch fabric architectures description

3.1 KEOPS switch

In the mid 90's, the European ACTS KEOPS project (KEys to Optical Packet Switching) did an intensive research around optical packet switching to study its feasibility as the basis for the Optical Transparent Packet Network (OTP-N) [6]. A broadcast-and-select switch was proposed for optical packet switching and a *16x16* switch operating at *10* Gbps, with a *1.646μs* time slot (*1,680* bytes for payload at *10* Gbps) was demonstrated. KEOPS switch (figure 5-a) was a development of the OASIS switch, from the RACE R2039 ATMOS (Asynchronous Transfer Mode Optical Switching) project, also supported by the European Commission.

In the KEOPS switch, packets from different inputs are encoded on different wavelengths by fixed wavelength converters, and broadcast to *M* delay lines, which are shared by all packets. In each time slot, packets arriving the switch in any input port, for the last *M* time slots, are selectable for transmission through any output port. The first stage of Semiconductor Optical Amplifiers (SOA) optical gates selects a particular delay-line through which the packet arrives. The second stage of optical gates selects the wavelengths. KEOPS architecture does not require any tunable component, supports multicast traffic and

packet priorities, and emulates output queuing. Its main disadvantage is that it is not suitable for even medium-size switches, as it has an optical power loss proportional to NM^2 and requires $(NM+N^2)$ optical gates.

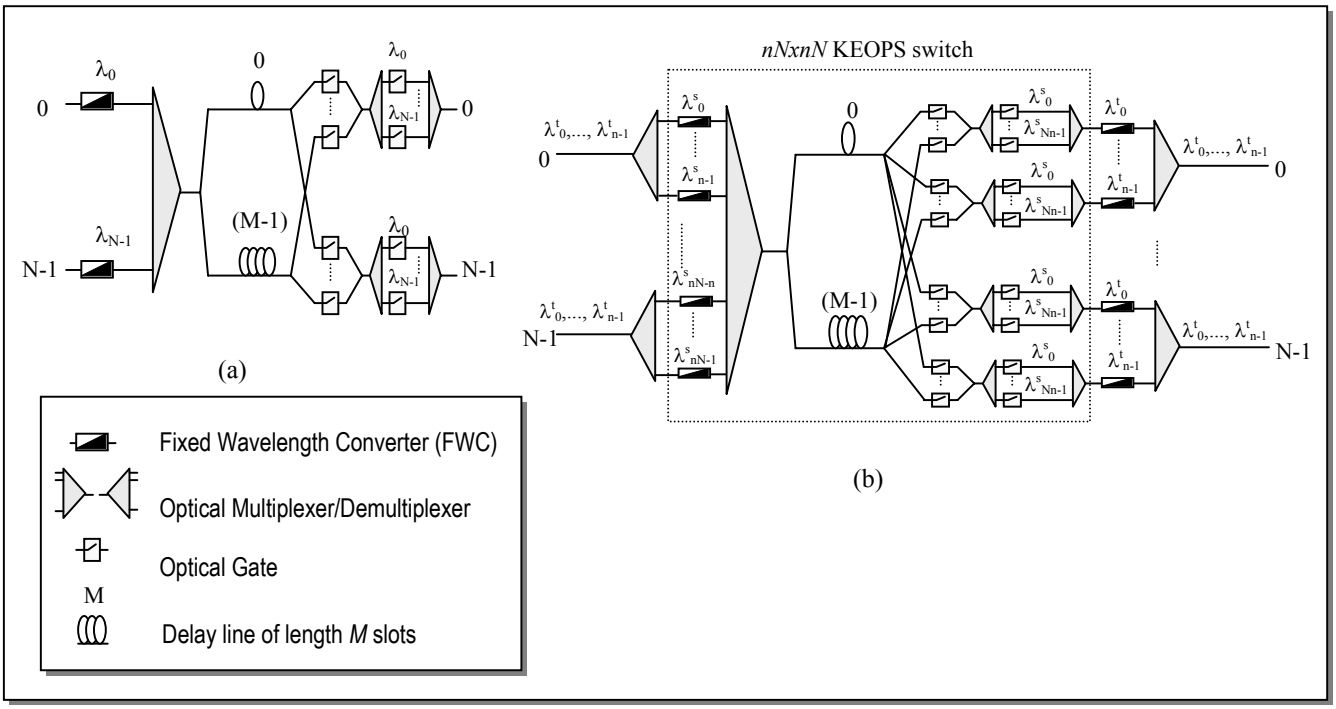


Figure 5. Broadcast-and-select KEOPS switch fabric (a) original architecture, (b) WDM architecture

The KEOPS switch, as displayed in figure 5-a, was not designed contemplating WDM input and output fibers. Thus, integration of this switch in an WDM OPS network requires the adaptation to a WDM environment (step 1 of the proposed normalization process). The adapted architecture applied in this paper is illustrated in figure 5-b. A wavelength demultiplexing stage is attached to the input side of an underlying *nNxnN* KEOPS switch. In the egress side of this switch, *n* output ports are wavelength converted and multiplexed in each WDM output fiber. Fixed wavelength conversion in the output stage means that packet output wavelength and fiber are fixed by selecting underlying KEOPS switch output port. Then, for the SHWP version switching function, packet output port of the internal KEOPS switch is defined by packet OPP. For the SCWP switching approach, a decision between *n* output ports of the internal KEOPS switch, associated to the same output fiber, should be taken, according to the algorithm of figure 4.

3.2 Space switch

A space-division based optical switch for an OPS network is described in [7]. The original switch shown in figure 6 interconnects *N* WDM inputs and *N* WDM outputs, each of them with *n* wavelengths (no WDM adaptation is required). Input packets arriving to the switch are converted to the appropriate output wavelength *($\lambda_0..\lambda_{n-1}$)* in the output fiber *(0..N-1)* selected. The switch consists of three main blocks: 1) a demultiplexer and *Nxn* Tunable Wavelength Converters (TWC) which encode the packet on the selected output wavelength, 2) an *nNxMN* space switch, and 3) a packet buffering stage which consists of fiber delay lines from *0* to *M-1*, for each one of the *N* output fibers. The original switching fabric was presented and evaluated only assuming the SCWP operation mode to exploit the wavelength dimension. Nevertheless, in this paper both the SCWP and SHWP versions are studied, to show how the large amount of buffering *(M)* required by the SHWP operational mode impacts on implementation feasibility for this particular architecture.

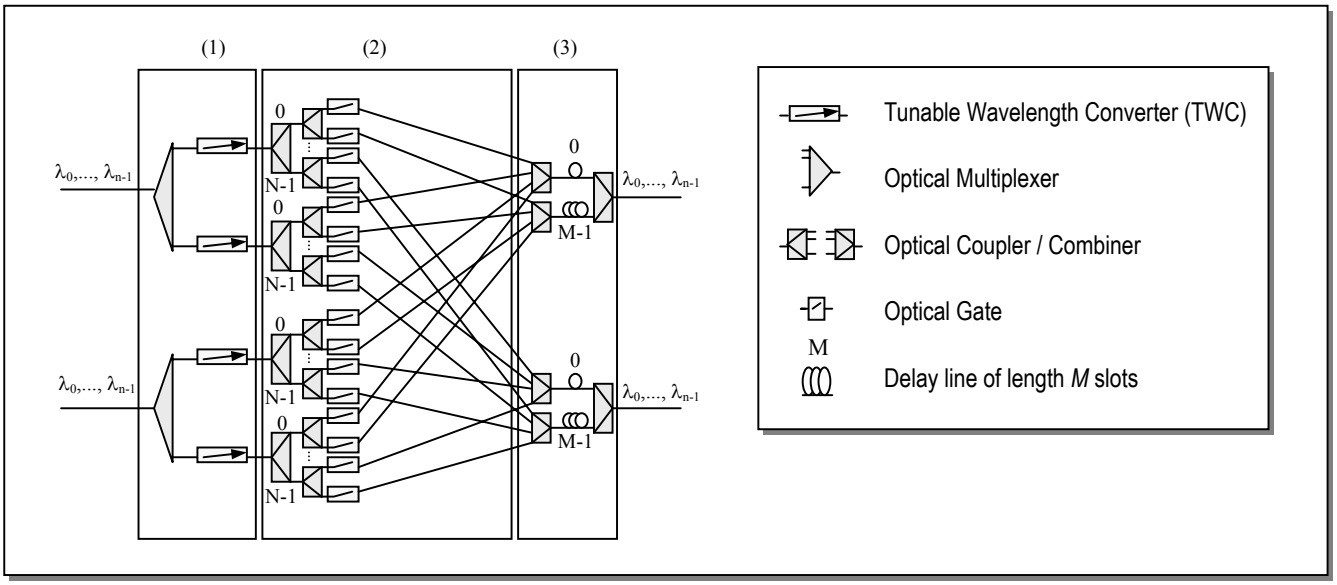


Figure 6. Space switch architecture, (1) packet encoding section, (2) space switching section, (3) buffering stage

3.3 Wavelength-routed output buffered switch

Progresses in fabrication of passive Arrayed-Waveguide-Grating (AWG) components using planar-lightwave circuit (PLC) technology empowered the study of optical packet switches taking advantage of their routing capability. This event allowed the application of AWGs to build compact buffering modules, as shown in figure 7-a. A *NxN* buffering section, with delays from *0* to *M-1*, can be designed by using *N* TWCs with a tuning range from λ_0 to λ_{K-1}, and two *KxK* AWG interconnected by the *M* delay lines, where *K=max (N,M)*. The cyclic routing operation of AWGs, depicted in figure 7-b, makes that a packet entering the *i*-th input port, leaves the buffering section by the *i*-th output port, whatever wavelength the packet is converted to. The

wavelength is then employed to determine the first AWG output port, and thus, the delay the packet will suffer, according to the rule λ=*(I/O port + delay) mod K*.

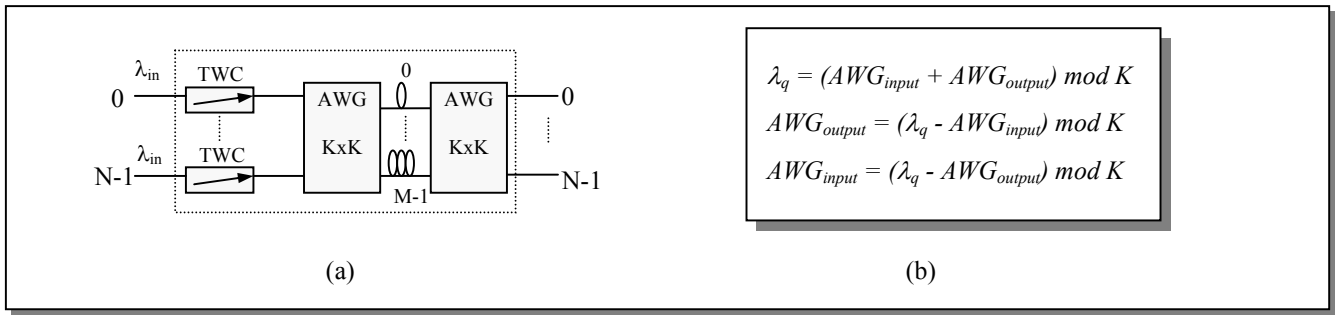


Figure 7. (a) AWG-based buffering section of *M* buffer positions, (b) routing features of AWG devices

A wavelength-routed output-buffered OPS switch fabric was presented in [8], which benefits from this buffering section. The original *NxN* output-buffered wavelength-routed (OB-WR) switch is shown in figure 8-a. It consists of a set of *N* TWC, a *NxN* non-blocking optical space switch matrix, and a *NxN* wavelength-routed packet buffer. The *NxN* space switch transfers input packets to their corresponding output port destination in a non-blocking manner so that, in an extreme case, up to *N* input packets can simultaneously be switched to the same output port of the space switch. The scheduling decision on packet delay consists of the selection of an appropriate wavelength in the wavelength conversion stage; thereby packets to the same output port are shifted to different wavelengths, and thus, routed through different delay lines in the buffering section.

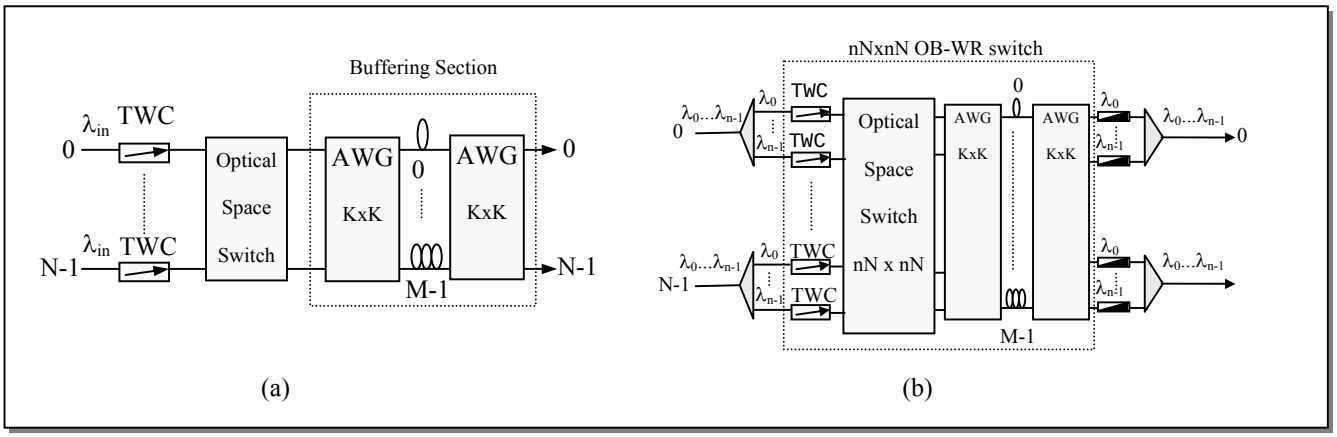


Figure 8. Output-buffered wavelength-routing (OB-WR) switch fabric (a) original architecture, (b) WDM adapted architecture

The output-buffered wavelength-routed switches must be adapted to the WDM input and output ports environment (step 1 of the normalization process) in order to apply SCWP or SHWP operation modes. Figure 8-b shows the required set of changes.

For this switch, wavelength demultiplexers and multiplexers are attached to the input and output sections of an original OB-WR switch. Consequently, $n \cdot N$ fixed wavelength converters should also be added. As in the KEOPS switch WDM version, the output wavelength selection is converted into output port selection of the underlying $nNxnN$ OB-WR switch.

4 Performance evaluation of SHWP and SCWP switches

In this section, performance evaluation of the previously introduced optical packet switches, for both SHWP and SCWP switching modes, is presented. The comparison is based on queuing analysis, validated by simulations. The optical packet switch is assumed to be symmetric, therefore having N input and output fibers, and n wavelengths $\lambda_0,..,\lambda_{n-1}$ for each fiber. Queue analysis is performed for independent uniform Bernouilli *(Be)* input traffic, of parameter ρ_{in}.

The scheduling processes for SHWP switches offer traditional output queuing behavior of buffer M. Evaluation of this model is based on the study of a tagged output queue of finite storage (M packet positions), fed by the aggregation of $N \cdot n$ Bernouilli sources of load ρ_{in}/N (figure 9-a). Under these conditions, evaluation of the switch is conducted by traditional output queuing analysis (see [9] for further details).

The scheduler for the SCWP switches, applying the algorithm described in section 2, is based on a round-robin dispatcher of input packets destined to the tagged output fiber. This equally spreads the traffic among n output queues of size M (figure 9-b). The fixed order in queue filling designated by this algorithm, allows the establishment of a one-to-one identification between buffer position M_b, $0 \le b \le M-1$ in queue of output wavelength λ_w, $0 \le w \le n-1$, and buffer position $M_{b \cdot n+w}$ of an equivalent n-server queue of $M \cdot n$ positions. The queue analysis employed for the evaluation of this multiserver finite buffered queue, fed by the aggregation of $n \cdot N$ independent input Bernouilli arrivals of offered load ρ_{in}/N (figure 9-c), is described in Appendix I.

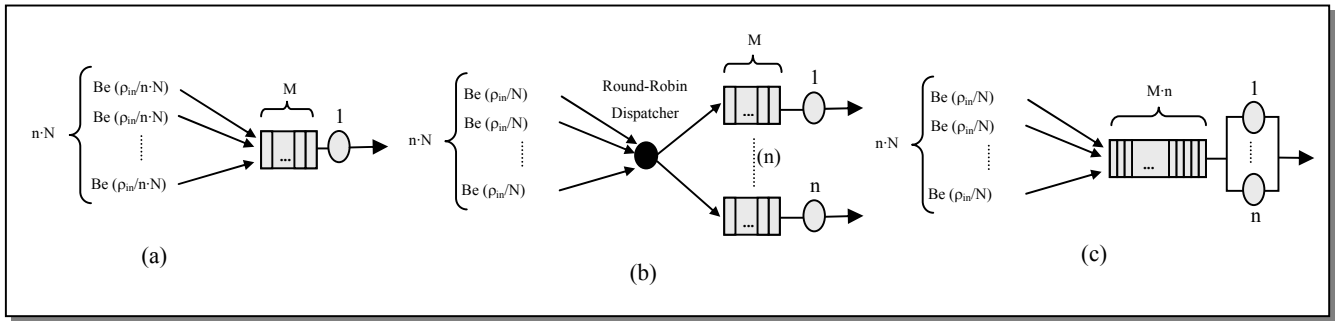
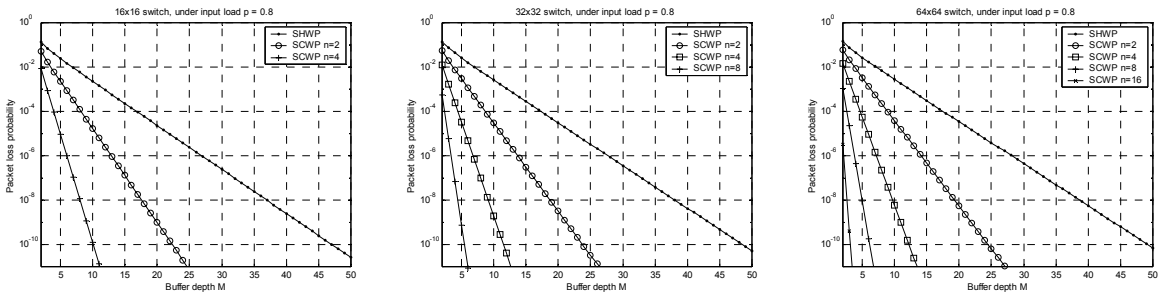


Figure 9. Queuing model for a $n \cdot N$ switch under *Be* (ρ_{in}) uniform input traffic, (a) SHWP switch queue model, (b) SCWP queue model under round-robin scheduler, (c) SCWP equivalent model

In a first step, both SHWP and SCWP evaluation methods have been applied to symmetric *16x16*, *32x32* and *64x64* switches, where switch size is determined as the number of input and output fibers (*N*), multiplied by the number of wavelengths per fiber (*n*). Figure 10-a,b,c, shows the packet loss probability performance of SHWP and SCWP switches as a function of the buffer depth *M* for different values of *n,* and offered loads $\rho_{in} = 0.8$. As no advantage is taken by this operation mode from the wavelength dimension, switch performance is independent of parameter *n* for SHWP switching and for a given switch size. On the other hand, a strong effect of *n* is observed in the graphs corresponding to SCWP switches, which clearly outperforms SHWP versions. As an example, in a *16x16* switch with *4* fibers and *4* wavelengths per fiber, *42* delays are required for achieving a packet loss probability $< 10^{-9}$ under SHWP operation mode, while only *11* are needed for SCWP switching. Figure 10 also shows a slight increase of packet loss probability as the switch size grows, in terms of number of input/output fibers. Regarding SCWP switches, the curve for $N=\infty$ can be considered an accurate pessimistic approximation when $N>32$. This is also the case for SHWP switches with the parameter $n \cdot N>32$.



(a) (b) (c)

Figure 10. Packet loss probability vs. buffer depth for *nNxnN* output queued SHWP and SCWP OPS switches, under input load ρ_{in} *0.8*, and *n = {2,4,8,16}*, (a) *16x16* switch, (b) *32x32* switch, (c) *64x64* switch

The comparison of the influence on packet delay of SHWP and SCWP modes is displayed in figure 11 normalized to packet duration, assuming a large switch size, and a sufficient buffer depth to provide a negligible packet loss probability. Again, a strong improvement in packet delay is observed in SCWP switches in relation to SHWP counterparts. This improvement is more evident for larger values of *n*, even under high input loads.

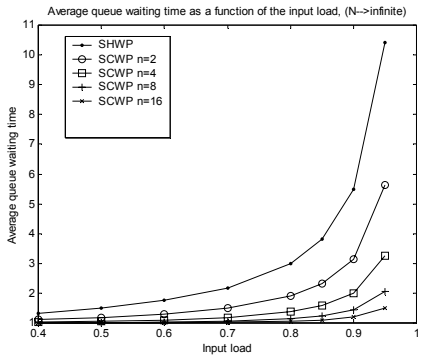


Figure 11. Normalized average packet delay vs. input load ρ_{in} for output queued SHWP and SCWP OPS switches as a function of the number of wavelengths *n* (assuming $N \rightarrow \infty$)

In table I, the required buffer depth for different switch sizes has been calculated considering a packet loss probability $< 10^{-9}$ under input load *0.8*. As an example, the obtained values show a *50%* saving in buffer positions for SCWP switches and for a wavelength dimension of *2*. Larger values of *n*, yield decreasing buffer requirements, confirming the benefits of the SCWP operation mode. For instance, a buffer depth of only *5* positions is required for a *32x32* switch with *8* wavelengths per fiber. The last column shows the impact of the size of the wavelength set on switch performance for $N \rightarrow \infty$, which can be used as an accurate pessimistic approximation for large-scale switches.

		16x16	**32x32**	**64x64**	**N→∞**
	n				
SHWP	**---**	42	44	44	45
SCWP	**2**	19	22	22	23
SCWP	**4**	10	11	11	12
SCWP	**8**	4	5	6	7
SCWP	**16**	***	3	3	4
SCWP	**32**	***	***	2	3
SCWP	**64**	***	***	***	2

Table I. Number of buffer positions required for output queued SHWP and SCWP OPS switches, for a packet loss probability $< 10^{-9}$, $n=\{2,4,8,16,32,64\}$, switch sizes *16x16, 32x32, 64x64*, and $N \rightarrow \infty$

5 Hardware cost comparison

In this section, the buffer requirements of SHWP and SCWP output queued switches analyzed in section 4, are used as the basis for hardware cost balance, which complements the evaluation process. The goal is to compare the different selected switch

fabric designs under SHWP and SCWP switching, as a contribution to a global comparison between SHWP and SCWP operation modes. As a first step, table II summarizes the hardware requirements for each of the selected architectures, as a function of the number of input/output fibers *(N),* wavelengths per fiber *(n)*, and buffer depth *(M).* These hardware requirements are calculated separately for each one of the components. An overall cost function for each architecture combining individual costs of the components has no clear meaning under current immature and boiling photonic research scenario, which may change the relative cost of the optical components in our medium term objective. Two more ideas should be taken into consideration:

(1) Hard limitations exist for the architectures described when applied to a large number of input/output ports. Distinct arrangements have been proposed to overcome the impractical complexity growth [10] of the selected architectures, however scalability and growability issues are not the central discussion of this paper. Then, the formulas in table II are applied in this paper to switch sizes up to *32* input/output channels.

(2) Also, it has to be remarked, that even under the switch sizes considered in this work, some of the alternatives exceeds current integration capacity in fabrication of optical components (still in a hand-made stage). Also, the fiber length required assuming fiber delay lines optical memories may be extremely large (although the final amount depends on slot time). Again, these values should be seen as a reference to gain insight, under the medium term objective in mind.

	Fixed Wav. Converters	**Optical Gates**	**Tunable Wav. Converters (tuning range)**	**Number of delay loops**	**AWG Size**
KEOPS switch	2nN	$MnN+n^2N^2$	0	M	0
OB-WR switch	nN	n^2N^2	nN (nN)	M	max (nN,M)
Space switch	0	nN^2M	nN (nN)	NM	0

Table II. Hardware cost computation

The values in Table I and II have been further employed to fill Table III, which calculates the number of components for a *32x32* channels switch for both SCWP and SHWP operation switching control modes. The calculus of the number of optical gates is provided assuming a crossbar implementation of space switches. Observing the SCWP versus SHWP hardware costs obtained for each architecture, the following points should be highlighted:

- For the KEOPS switch architecture, the influence of operation mode primarily impacts on the optical gates count. Assuming a constant switch size *(nN)*, this value shows a *linear* increase with the factor *M*. For this reason, the severe (and not linear) decay observed in buffering requirements *(M)* under SCWP operation mode, is linearly translated to hardware cost simplification. As an example, the reduction in the number of optical gates in SCWP switching versus SHWP version, reaches the *50%* for *n=8*.
- Minor improvements in AWG size and TWC tuning range are obtained under SCWP operation mode in OB-WR architecture. Specially, the number of optical gates remains constant, which is envisaged as a high impacting factor on final cost.
- Assuming a constant switch size *(nN)*, the number of optical gates and the number of delay lines required for the space switch grow linearly with the value *M/n*. In SHWP mode, the amount of buffer requirements *M* is large, and constant with *n*. This provides a linear decrease of the number of optical gates and delay loops as *n* grows. Nevertheless, the combined effect of buffer requirement reduction as *n* grows obtained under SCWP switching also has some repercussion on the *M/n* factor, pointing to a much stronger simplification (in the order of *7* to *1* for *n = 8*).

Analyzing hardware cost of each architecture for SHWP operation mode, KEOPS and OB-WR switches are observed as the two leading architecture candidates. OB-WR switch requires one half of the fixed wavelength converters, and a lesser amount of optical gates than KEOPS switch. On the other hand, KEOPS switch does not entail any tunable component or AWG devices. In addition, KEOPS switch admits an easier implementation of multicast (due to the broadcast-and-select function), and an easier implementation of quality-of-service issues (as all packets are available for transmission along *M* time slots, allowing a more prioritized packet to overtake a previous one).

The comparison among SCWP switching fabrics shows that space switch architecture may be the preferred alternative for higher values of *n*, taking benefit from the higher hardware simplification in terms of number of optical gates, fixed wavelength converters (which are not required), and tuning range of tunable converters. Nevertheless, the hypothetical requirements of multicast and/or QoS issues in a future backbone OPS network may favor KEOPS architecture.

A general vision on this section establishes clear benefits of SCWP control respect to SHWP switching mode, in terms of hardware complexity for the selected switch fabric designs. Also, application of the round-robin output wavelength assignment algorithm depicted in figure 4 under SCWP operation mode, preserves end-to-end packet sequence at a reasonable cost. Nevertheless, a final and definitive decision between SHWP and SCWP operation modes is out of the scope of this paper, due to the very different aspects that should be considered.

		Fixed Wav. Converters		Optical Gates		Tunable Wav. Converters (tuning range)		Number of delay loops		AWG (size)	
	n	**SHWP**	**SCWP**	**SHWP**	**SCWP**	**SHWP**	**SCWP**	**SHWP**	**SCWP**	**SHWP**	**SCWP**
KEOPS switch	**2**	64	64	2432	1728	0 (0)	0 (0)	44	22	0 (0)	0 (0)
	4	64	64	2432	1376	0 (0)	0 (0)	44	11	0 (0)	0 (0)
	8	64	64	2432	1216	0 (0)	0 (0)	44	6	0 (0)	0 (0)
OB-WR switch	**2**	32	32	1024	1024	32 (44)	32 (32)	44	22	2 (44)	2 (32)
	4	32	32	1024	1024	32 (44)	32 (32)	44	11	2 (44)	2 (32)
	8	32	32	1024	1024	32 (44)	32 (32)	44	6	2 (44)	2 (32)
Space switch	**2**	0	0	22528	11264	32 (2)	32 (2)	704	352	0 (0)	0 (0)
	4	0	0	11264	2816	32 (4)	32 (4)	352	88	0 (0)	0 (0)
	8	0	0	5632	768	32 (8)	32 (8)	176	24	0 (0)	0 (0)

Table III. Hardware cost evaluation for the selected architectures, for a switch size of *32x32*, *n={2,4,8}*, attending to buffer depth detailed in table I

6 Conclusions and further work

Optical packet switching for backbone networks is far from commercial reality. It is highly likely that its deployment will be related to the WDM concept. In the frontier of the two worlds, both SHWP and SCWP operation modes were proposed under the WASPNET project [3-5] as the competing switching modes in the network. A normalization effort is required to evaluate the OPS switch fabric architectures proposed in the open literature under SCWP and SHWP alternatives, to allow a fair comparison among the architectures, and globally, between both operational modes. This objective has been accomplished for three relevant OPS architectures based on the output queued concept. Attending hardware cost, for these switch architectures, it is revealed that SCWP operation mode is clearly advantageous. Performance evaluation of other OPS architectures under SCWP/SHWP operation modes is required to definitely establish the benefits. In this way, a wider comparison involving two more architectures: input-buffered wavelength routing switch [8], and FRONTIERNET switch [11] is a current line of work of our research group.

Acknowledgements

This work was supported by the Spanish Research Council under projects FAR-IP (TIC 2000-1734-C03-03) and MTCES (TIC 2001-3339-C02-02).

References

[1] Murthy C., Gurusamy M., "WDM optical networks. Concepts, design and algorithms". Prentice Hall PTR, 2002.

[2] Yao S., Xue F., Mukherjee B., Yoo B., Dixit S., “Electrical ingress buffering and traffic aggregation for optical packet switching and their effect on TCP-level performance in optical mesh networks”, *IEEE Communications Magazine*, vol. 40, no. 9, Sep. 2002, pp. 66-72.

[3] Hunter D., Nizam M., Chia M., Andonovic I., Guild K., Tzanakaki A., O’Mahony J., Bainbridge J., Stephens M., Penty R., White I., "WASPNET: A Wavelength Switched Packet Network", *IEEE Communications Magazine,* vol. 37, no. 3, March 1999, pp. 120-129.

[4] Chia M., Hunter D., Andonovic I., Ball P., Wright I., Ferguson S., Guild K., O’Mahony M., "Packet loss and delay performance of feedback and feed-forward arrayed-waveguide gratings-based optical packet switches with WDM inputs-outputs", *IEEE Journal of. Lightwave Technology*, vol. 19, no. 9, Sept. 2001, pp. 1241-1254.

[5] Nizam M.H.M., Hunter D.K., Andonovic I., "Designing an optimum WDM transport network: control architectures, node requirements and performance", *Proc. Soc. Photo-Optical Instrumentation Engineers (SPIE),* vol. 3531, Oct. 1998, pp. 244-255.

[6] Guillemot C., Renaud M., Gambini P., Janz C., Andonovic I., Bauknecht R., Bostica B., Burzio M., Callegati F., Casoni M., Chiaroni D., Clérot F., Danielsen S., Dorgeuille F., Dupas A., Franzen A., Hansen P., Hunter D., Kloch A., Krähenbühl R., Lavigne B., Le Corre A., Raffaelli C., Schilling M., Simon J., Zucchelli L., "Transparent optical packet switching: the European ACTS KEOPS project approach", *IEEE Journal of Lightwave Technology*, vol. 16, no. 12, Dec. 1998, pp. 2117-2134.

[7] Danielsen S. L., Joergensen C., Mikkelsen B., Stubkjaer K., "Analysis of a WDM packet switch with improved performance under bursty traffic conditions due to tunable wavelength converters", *IEEE J. Lightwave Technol.*, vol. 16, no. 5, May 1998, pp. 729-735.

[8] Zhong W., Tucker R., "Wavelength routing-based photonic packet buffers and their applications in photonic packet switching systems", *IEEE Journal of Lightwave Technology*, vol. 16, no. 10, Oct. 1998, pp. 1737-1745.

[9] Hluchyj M., Karol M., “Queueing in high-performance packet switching”, *IEEE Journal on Selected Areas in Communications.,* vol. 6, no. 9, Dec. 1988, pp. 1587-1597.

[10] Pavon-Marino P., Garcia-Haro J., Malgosa-Sanahuja J., “Scaling strategies survey for envisaged backbone optical packet switches”, Proc. of IASTED Comm. Systems and Networks (CSN 2002), Malaga, Spain, Sep. 2002, pp. 178-183.

[11] Sasayama K., Yamada Y., Habara K., Yukimatsu K., "FRONTIERNET: frequency-routing-type time-division interconnection network", *IEEE Journal of Lightwave Technology*, vol. 15, no. 3, March 1997, pp. 417-419.

[12] Kleinrock L., “Queueing systems. Volume I: Theory”, John Wiley & Sons, 1975.

Appendix I

This section analyzes the performance of an n-multiserver queue of $M \cdot n$ buffer positions, like the one shown in figure 9-c. Queue input traffic is assumed to be the aggregation of $n \cdot N$ Bernouilli independent sources of parameter ρ_{in}/N, which corresponds to the traffic arrivals to a tagged output queue of an SCWP switch. Defining the random variable A as the number of packet arrivals destined for the tagged output in a given time slot, we have

$$a_k = \Pr[A = k] = \binom{nN}{k} \left(\frac{\rho_{in}}{N}\right)^k \left(1 - \frac{\rho_{in}}{N}\right)^{nN-k} ; k = 0,1,....,nN \qquad (1)$$

which for $N \rightarrow \infty$, becomes

$$a_k = \Pr[A = k] = \frac{\left(\frac{\rho_{in}}{n}\right)^k e^{-\frac{\rho_{in}}{n}}}{k!} ; k = 0,1,....,\infty \qquad (2)$$

Denoting Q_m as the number of packets in the multiserver queue at the end of the m-th time slot, and A_m as the number of packet arrivals during the m-th time slot, we obtain

$$Q_{m+1} = min\{max\{0, Q_m - n\} + A_m, n \cdot M\} \qquad (3)$$

Q_m is modelled as a finite-state Markov chain with state transition probabilities $P_{i,j} = \Pr[Q_{m+1} = j \,|\, Q_m = i]$ given by

$$P_{i,j} = \begin{cases} a_j & if\ i \le n,\ j \le nN \\ a_{j-i+n} & if\ n+1 \le i \le nM,\ i-n \le j \le min(nM-1, nN+i-n) \quad when\ M \ge N \\ \sum_{s=nM+n-i}^{nN} a_s & if\ n(M-N+1) \le i \le nM,\ j = nM \end{cases} \qquad (4)$$

and

$$P_{i,j} = \begin{cases} a_j & if\ i \le n,\ j < nM \\ \sum_{s=nM}^{nN} a_s & if\ i \le n,\ j = nM \\ a_{j-i+n} & if\ n+1 \le i \le nM,\ i-n \le j < nM \\ \sum_{s=nM+n-i}^{nN} a_s & if\ n+1 \le i \le nM,\ j = nM \end{cases} \quad when\ N \ge M \qquad (5)$$

The steady-state queue size probabilities q_i, $i=0..nM$ can be directly obtained from the Markov chain balance equations [12]. System throughput ρ_{out} is then calculated by observing the number of packets transmitted on the tagged output during a time slot, which yields

$$\rho_{out} = \sum_{s=1}^{nM} q_s \cdot min(s,n) \quad < n \qquad (6)$$

Packet loss probability is calculated as a function of the packets offered to the queue and being served by the system (6).

$$\Pr[\,packet\ loss\,] = 1 - \frac{\rho_{out}}{n\rho_{in}} \qquad (7)$$

Finally, the mean waiting time $\overline{W}$ for a packet in the queue can be obtained by applying Little's result.

$$\overline{W} = \frac{\overline{Q}}{\rho_{out}} = \frac{\sum_{s=1}^{nM} s \cdot q_s}{\sum_{s=1}^{nM} min(s,n) \cdot q_s} \qquad (8)$$