A Performance Study of a Knock-Out Large-Scale Optical Packet Switching Architecture Under SCWP Operational Mode

Pablo Pavon-Marino, Joan Garcia-Haro, Josemaría Malgosa-Sanahuja, Fernando Cerdan
Department of Information Technologies & Communications, Polytechnic University of Cartagena, E-30202, Spain
Tel: +34 968 325952, Fax: +34 968 32 59 73
Email: {pablo.pavon, joang.haro, josem.malgosa, fernando.cerdan}@upct.es

Abstract— **Large-scale Optical Packet Switching is in a very immature research state. Specially, the effect of the application of the Scattered Wavelength Path operational mode to large scale architectures has not been well explored yet. In this paper, two variants of a knock-out large scale OPS architecture are presented: an output fiber distributed version, and an output wavelength distributed version. Evaluation of these architectures is conducted assuming a previously proposed output wavelength assignment algorithm. It is shown that the traffic spreading features of this algorithm, which provides an optimum behavior in terms of buffer overflow packet losses, also provide a performance improvement in terms of knock-out losses in the wavelength-distributed knock-out architecture. The hardware cost simplifications associated are illustrated by means of an example considering KEOPS switch fabric as the buffered output module of the knock-out architecture.**

I. INTRODUCTION

In the optical domain, Optical Packet Switching (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 unavoidable advantages in terms of bandwidth sharing and resource allocation since it operates on the granularity of a packet. Optical Packet Switching has been largely focused on fixed size packets, and this is also the case of this paper.

In the application of Optical Packet Switching on top of Wavelength Division Multiplexing (WDM) networks, the networking operation mode establishes the manner in which the permanent higher layer connections (Optical Packet Paths, OPPs) between backbone edge nodes are mapped onto the appropriate wavelengths in the WDM links. This topic has been addressed under the WASPNET project [1], where two possible methodologies were proposed: The Shared Wavelength Path (SHWP), and the Scattered Wavelength Path (SCWP). In SHWP, packets from the same OPP follow a fixed sequence of hops to the egress node, where transmission fiber and transmission wavelength are fixed for each hop. Under these considerations, each packet entering an OPS switch requests a fixed output fiber and a fixed output wavelength, whose values are stored in a look-up table on OPP provisioning. Unlike SHWP, in SCWP optical paths determine transmission fibers during OPP establishment, but do not fix transmission wavelength in each link. In each node, decision on output wavelength for each packet is the subject of an output wavelength assignment algorithm. The goal of this algorithm is to obtain the highest statistical multiplexing of the aggregated bandwidth of the wavelengths in the output fibers, optimizing buffer occupancy, and avoiding packet out-of-sequence. For this reason, SCWP switching offers higher throughput than SHWP operational mode, yielding to simplified architectures which require less optical buffers at the switch.

A. Large scale Optical Packet Switching

Fast packet-by-packet switching operation associated to Optical Packet Switching, imposes the highest constraints to the photonic switching function, involving very large hardware costs for the state-of-the-art technology. As a result, the deployment of an OPS backbone network is not envisaged up to the medium term. Very little can be said about the requirements of future OPS backbone network switch fabrics. The impressive growth in the number of transmission channels (even hundreds) that can be multiplexed onto each fiber, claims for OPS switch fabrics with a large count of input and output ports, one per wavelength and fiber. Scalability restrictions related to photonic devices, make that OPS proposed architectures are not feasible to implement even for *32x32* or *64x64* port switch sizes. Therefore, aggregation of optical switching elements into growable/scalable connection topologies is required [2]. Nevertheless, large-scale optical packet switching is still in its initial and immature research stage. Specially, the impact of the SCWP operation mode -the foreseen preferred alternative- on the large-scale designs has not been well investigated.

In this paper two variants of a knock-out [3] based large scale OPS architecture are presented: an output fiber distributed version, and an output wavelength distributed version. Evaluation of the knock-out losses for both architectures is performed by applying the output wavelength assignment algorithm described in [4], which offers the optimum performance in terms of buffer requirements. The benefits of the wavelength-distributed architecture in terms of reduction of knock-out packet loss probability, are compared in an exemplified dimensioning process.

The rest of the paper is organized as follows. Section 2 presents the two knock-out based architectures under study, together with the output wavelength assignment applied. Section 3, proposes a knock-out packet loss probability

0-7803-7978-0/03/$17.00 ©2003 IEEE

evaluation process, which is applied in a brief dimensioning example in Section 4. Finally, section 5 concludes the paper.

I. SWITCH FABRIC ARCHITECTURES

A. Hardware description

The two knock-out effect OPS architectures studied in this paper, are illustrated in figure 1-a and 1-b. The architectures are based on the connection of a memoryless distribution stage and a buffered output stage. In the output fiber distributed switch version (figure 1-a), one module exists per output fiber in the switch. On the other hand, in the output wavelength distributed version (figure 1-b), one module exists in the buffering section per output wavelength.

1) Distribution stage

The distribution stage in both switch fabrics is composed of a set of *nN* Tunable Wavelength Converters (TWC) and an Arrayed-Waveguide Grating device. Wavelength conversion of input packets in the distribution stage is employed to select packet output port of the AWG routing device. This output port is selected among the *L (L')* ports which connect to the module associated to the requested output fiber (wavelength). Therefore, knock-out losses arise in this stage if more than *L (L')* packets are destined to the same output module in the same time slot.

Progresses in the fabrication of large passive Arrayed-Waveguide-Grating (AWG) components using planar-lightwave circuit (PLC) technology, and the expansion of the tuning range of tunable wavelength converters (TWC) devices, allow for the implementation of large memoryless switch modules for optical packet switches. Feasibility of this architecture is envisaged in the short term for switch sizes close to 256 input and output ports.

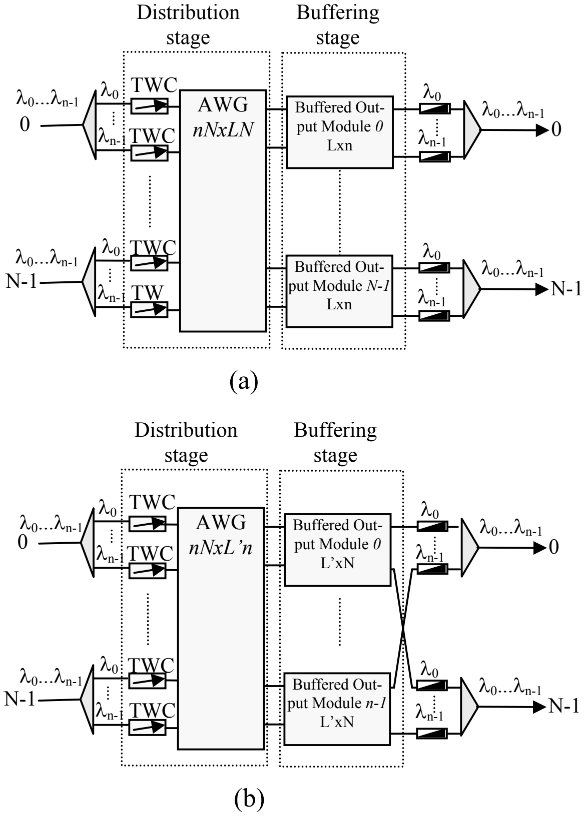
2) Buffering stage

Different OPS switch fabric architectures found in the open literature can be applied for the buffered switching modules in the buffering stage. Their mission is twofold: (1) appropriately store (delay) input packets, and (2) switch input packets to the requested output port of the buffered module. Buffer overflow is the source of packet losses in this stage.

B. Output wavelength selection algorithm

The algorithm of output wavelength selection applied to our performance study is specified in figure 2. This algorithm was previously presented in [4]. Given an SCWP switch with *N* input and *N* output fibers, *n* wavelengths per fiber, and *B* buffer positions per output wavelength and fiber, the algorithm is based on a round-robin assignment of output wavelengths for packets addressed to the same output fiber. For this, *N* round-robin pointers λ_i, $i=0..N-1$ are required (see figure 2 for further details).

As it is shown in [4], the algorithm provides the advantages of (1) simple implementation, (2) it maintains end-to-end packet sequence, and (3) provides an overall switch behavior equivalent to an optimum performance *n*-server queue model, with *Bn* buffer positions. Therefore, when applied to our large-scale architecture, the aforementioned algorithm achieves the optimum performance for buffer overflow packet loss, if output buffered switch fabrics are used in the modules [4].



(a)

(b)

Fig. 1. OPS knock-out architecture, (a) fiber-distributed, (b) wavelength-distributed.

```
/* N = number of I/O fibers */
/* n = number of wav. per fiber */
/* M = buffer depth */

for input i = 0 to nN-1 do
   if (packet p present on input i) then
     f = output fiber demanded by p
     if (delay [f] < B) then
         associate delay [f] to p
         associate wav. $\lambda_f$ to p
         /* $\lambda_f$ is a RR pointer */
         $\lambda_f$ ++
         if ($\lambda_f$ == n)
            $\lambda_f$ = 0
           delay [f] ++
         endif
      endif
   endif
endfor

/* decrement 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)
     $\lambda_f$ = 0 /* reset RR pointer */
  endif
endfor
```

Fig. 2. Round-robin SCWP scheduler (pseudocode).

Evaluation of the knock-out losses of this algorithm when applied to the architectures in figure 1-a and 1-b is a relevant point in this paper and it is presented in the next section.

II. KNOCK-OUT LOSS EVALUATION

This section analyzes the knock-out losses of OPS knock-out architectures described in the previous section. Bernouilli uniform input traffic of parameter ρ is assumed. Defining the random variable *A,* with values between *0* and A_{MAX}, as the number of packet arrivals destined to a tagged output module in a given time slot, we have that packet loss probability caused by the knock-out effect (P_{loss}) in both architectures is given by (1).

$$P_{loss} = \frac{E\left[packet\ loss\right]}{E\left[packet\ arrivals\right]} = \frac{\sum_{i=L+1}^{A_{MAX}} P(A=i)\cdot(i-L)}{\sum_{i=1}^{A_{MAX}} P(A=i)\cdot i} \tag{1}$$

Specification of the discrete density function of *A,* $P(A=i)$, $i=0..A_{MAX}$, is a required step to calculate the knock-out losses.

For the fiber-distributed architecture version, all the packets destined to the same output fiber are directed to the same output module. This is not affected by the output wavelength selection algorithm. The distribution of *A* is then given by the binomial formula.

$$P[A=k] = \binom{nN}{k}\left(\frac{\rho}{N}\right)^{k}\left(1-\frac{\rho}{N}\right)^{nN-k} \quad ; k = 0,1,...,nN = A_{MAX} \tag{2}$$

On the other hand, calculation of the density function of *A* for the wavelength-distributed architecture is affected by the output wavelength selection algorithm applied. For our tagged output module, we define random variable A_i $(i=0..N-1)$ as the number of packets arriving this module, which are destined to output fiber *i* (and thus destined to the *i*-th output port of the tagged module). In its turn, the random variable A_i is a function of two random variables:

- a_i $(i=0..N-1)$: Indicates the total number of packets destined to output fiber *i* in this time slot. Each random variable a_i can take values from *0* to *nN*.
- λ_i $(i=0..N-1)$: Indicates the initial position in this particular time slot of the round-robin pointer associated to output fiber *i.* Each random variable λ_i can take values from *0* to *n-1*. A value of *0* means that the pointer points to our tagged module, and thus the first packet destined to output fiber *i* will be switched to this module. A value of *n-1* indicates that the first *n-1* packets destined to output fiber *i* will be destined to other output modules.

Under these considerations, the random variable A_i is specified by (3), using the *ceil* function:

$$A_i = \left\lceil \frac{a_i - \lambda_i}{n} \right\rceil, i = 0..N-1 \tag{3}$$

Random variable *A* is then expressed as the sum of the random variables A_i, $i=0..N-1$.

$$A = A_0 + ... + A_{N-1} = A\left(\vec{a},\vec{\lambda}\right) \tag{4}$$

where $\vec{a}$ is the vector of arrivals $\vec{a} = \left(a_o,...,a_{N-1}\right)$, and $\vec{\lambda} = \left(\lambda_o,...,\lambda_{N-1}\right)$ is the vector of initial pointers positions. In each time slot, the random variables λ_i are jointly independent, and also independent from the arrival random variables a_i. However, the variables a_i, $i=0..N-1$ cannot be considered jointly independent. For example, while a_i can take the values *0..nN*, we have the constraint that the sum of arrivals to the knock-out switch $a_0+...+a_{N-1}$ is bounded by the number of input ports *nN*.

Exact calculation of the distribution of *A* is given by (5):

$$P\left[A=k\right] = \sum_D P\left[\vec{a} = \overrightarrow{a_d}, \vec{\lambda} = \overrightarrow{\lambda_d}\right] = \sum_D P\left[\vec{a} = \overrightarrow{a_d}\right] P\left[\lambda_1\right] \cdot ... \cdot P\left[\lambda_{N-1}\right] \tag{5}$$

Where *D* symbolizes the set of values of $\overrightarrow{a_d}$ and $\overrightarrow{\lambda_d}$ which provides that $A = \sum_{i=0}^{N-1} \left\lceil \frac{a_i - \lambda_i}{n} \right\rceil = k$. The arrivals distribution is given by the multinomial formula:

$$P\left[\vec{a} = (a_0,..,a_{N-1})\right] = \frac{(nN)!}{a_0!\cdot ...\cdot a_{N-1}!\cdot E!} \cdot (1-\rho)^{E} \left(\frac{\rho}{N}\right)^{nN-E} \tag{6}$$

Where *E* symbolizes the number of empty input ports in this time slot. While the independent probabilities $P[\lambda_i=k]$ are uniformly distributed.

$$P\left[\lambda_i = k\right] = \frac{1}{n}, k = 0..n-1 \tag{7}$$

Brute force solving of equation (5) by applying distributions in (6) and (7) is unfeasible even for moderate switch sizes. The calculation of the arrivals distribution process and knock-out losses have been performed by using a simplification method. The details of this method for the summation of expression (5) are not reproduced here due to the lack of space.

A. Upper bound of the number of arrivals

The maximum number of arrivals to the tagged wavelength-distributed output module in a time slot is obtained when:

1) Pointers λ_i, $i=0..N-1$, initially point to our tagged output module ($\lambda_i=0$, $i=0..N-1$).
2) N input packets are destined to N different output fibers, and thus, are routed to our tagged output module.
3) The remainder $(nN-N)$ input ports have packets destined to the same output fiber.

The value A_{MAX} is then bounded by the expression (8):

$$A_{MAX} = min\left(nN, N + \left\lceil \frac{nN - N - n + 1}{n} \right\rceil \right) \qquad (8)$$

If the parameter L' of the wavelength-distributed architecture is dimensioned with the value $L'=A_{MAX}$, a *0* knock-out loss probability is obtained for any input traffic pattern.

III. EVALUATION EXAMPLE

In table I, the values of the number of required knock-out links for a knock-out packet loss probability of 10^{-8} are computed. Input load of *0.8* is considered, for a switch size of *64* input and output ports $(nN=64)$, and different value arrangements of parameter N and n. A column for *0*-knock-out packet loss probability is appended for the wavelength-distributed approach.

In the buffering section, the saving in hardware cost obtained depends on the output module switch architecture selected. In this paper, a brief dimensioning process is introduced employing the KEOPS switch fabric [5] for these modules. The most impacting factor on hardware cost in KEOPS switch fabrics is provided by the amount of optical gates required. For a $P_{in}xP_{out}$ KEOPS switch with B buffer positions, this is given by $OG = P_{out}(B+P_{in})$. Then, in the fiber-distributed approach, the global amount (summing all output modules) of optical gates needed is given by $OG = N[n\cdot(B+L)] = nNB+nNL$. In its turn, in the wavelength-distributed version, the number of optical gates is given by $OG = n[N\cdot(B+L')] = nNB+nNL'$. Taking into consideration that the buffer requirements B of a $nNxnN$ switch are equal for both knock-out versions, the parameter of interest to compare is the second addend. Table II uses values in table I detailing both addends of optical gates count. The buffer requirements are calculated for a probability of packet loss associated to buffer overflow $<10^{-9}$ (see [4] for details). The benefits (in terms of hardware simplification) of the output wavelength assignment algorithm in the wavelength distributed architecture, when the number of wavelengths is greater than *4*, are clearly shown.

TABLE I
KNOCK-OUT LINKS DIMENSIONING EXAMPLE

Symbol	Fiber-distributed		Wavelength-distributed			
	KO loss < 10^{-8}		KO loss < 10^{-8}		KO loss = 0	
	L	*Mod.*	*L'*	*Mod.*	*L'*	*Mod.*
$N=32,n=2$	12	32	38	2	48	2
$N=16,n=4$	16	16	21	4	28	4
$N=8,n=8$	22	8	12	8	15	8
$N=4,n=16$	31	4	6	16	7	16
$N=2,n=32$	46	2	3	32	3	32

Mod.: Number of output modules required.
L (L'): Number of knock-out links per output module required.

TABLE II
OPTICAL GATES COUNT

Symbol	Fiber-distributed	Wavelength-distributed	
	KO loss < 10^{-8}	KO loss < 10^{-8}	KO loss = 0
$N=32,n=2, B=22$	1408+768	1408+2432	1408+3072
$N=16,n=4,B=11$	704+1024	704+1344	704+1792
$N=8,n=8,B=5$	320+1408	320+768	320+960
$N=4,n=16,B=3$	192+1984	192+384	192+448
$N=2,n=32,B=2$	128+2944	128+192	128+192

IV. CONCLUSIONS

The main contributions of this paper are the proposal and comparative evaluation of two OPS knock-out architectures and for Scattered Wavelength Path (SCWP) operational mode. A previously proposed output wavelength selection algorithm is applied and studied for both switching fabrics. Evaluation of the knock-out loss probability shows that the traffic spreading features of this algorithm can simplify hardware requirements for the wavelength-distributed large-scale switch version. A brief dimensioning process is exemplified, using the KEOPS architecture as the output buffered module choice. Results show this knock-out architecture as a very promising option in the field of large-scale Optical Packet Switching fabrics.

ACKNOWLEDGMENT

This work has been supported by the Spanish Research Council under projects FARIP (TIC2000-1737-C3-03) and MTCES (TIC2001-3339-C02-02).

REFERENCES

[1] D. Hunter *et al.*, "WASPNET: A Wavelength Switched Packet Network", *IEEE Communications Magazine*, vol. 37, no. 3, pp. 120-129, March 1999.

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

[3] Y. S. Yeh, M. G. Hluchyj, A. S. Acampora, "The knock-out switch: A simple, modular architecture for High Performance Packet Switching", *IEEE Journal on Selected Areas in Comm.*, vol. SAC-5, no. 8, pp. 223-231, October 1987.

[4] P. Pavon-Marino, J. Garcia-Haro, J. Malgosa-Sanahuja, F. Cerdan, "Optical Packet Switching Fabrics Comparison Under SCWP/SHWP Operational Modes", accepted for publication in *8th IEEE International Symposium on Computers and Communications (ISCC'2003)*, Antalya, Turkey, July 2003.

[5] C. Guillemot *et al.*, "Transparent optical packet switching: the European ACTS KEOPS project approach", *IEEE Journal of Lightwave Technology*, vol. 16, no. 12, pp. 2117-2134, Dec. 1998.