ALMA MATER STUDIORUM
UNIVERSITY OF BOLOGNA - ITALY

Ph.D. in Electronics, Computer Science and Telecommunications

HIGH-PERFORMANCE SWITCHING ARCHITECTURES FOR OPTICAL NETWORKS

Ph.D. thesis by

MICHELE SAVI

Coordinator

Prof. PAOLO BASSI

Supervisor

Prof. GIORGIO CORAZZA

Tutor

Prof. CARLA RAFFAELLI

ACADEMIC YEAR 2006-2007

ALMA MATER STUDIORUM
UNIVERSITÀ DEGLI STUDI DI BOLOGNA

Dottorato di ricerca in Ingegneria Elettronica, Informatica e delle Telecomunicazioni

ING-INF/03

CICLO XX

STUDIO DI ARCHITETTURE DI COMMUTAZIONE AD ALTE PRESTAZIONI PER LE RETI OTTICHE

Tesi di Dottorato di

MICHELE SAVI

Coordinatore

Chiar.mo Prof. Ing.
PAOLO BASSI

Tutor

Chiar.mo Prof. Ing.
GIORGIO CORAZZA

Correlatore

Prof. Ing.
CARLA RAFFAELLI

ANNO ACCADEMICO 2006-2007

to Fabiana, Antonio, Elide and Serena.

Contents

Acknowledgments

In this Ph.D. thesis I tried to summarize my research work performed with the Networks Research Group at the Department of Electronics, Computer Science and Systems (DEIS) of the University of Bologna. I first met the members of the group in 2003 when I started to work at my Graduate Thesis focused on routing algorithms for providing QoS in optical networks. Then they gave me the great chance to begin my Ph.D. course and the opportunity to go deeply in the subject of photonic packet switching.

I really would like to thank Professor Giorgio Corazza, the group leader and my supervisor, as well as Professors Carla Raffaelli and Franco Callegati for the invaluable help they gave me by means of a number of precious advices, coming from a long experience in the research field of optical networking. I also would like to thank Dr. Walter Cerroni and Dr. Giovanni Muretto and all the other people working with me during these three years for the constant support and useful cooperation.

The research work I performed was mainly supported to the European *e-Photon/ONe* and *e-Photon/ONe+* projects and the Italian OSATE project, so I also would like to thank all partners involved in these projects for the fruitful collaborations and discussions carried out during meetings and conferences.

Finally a special thanks to the people working at the Telecommunications section of the Department, my family and Fabiana for the constant support they gave me during these three years (thank you for the infinite patience).

Bologna - March 17, 2008

Summary

The need for high bandwidth, due to the explosion of new multimedia-oriented IP-based services, as well as increasing broadband access requirements is leading to the need of flexible and highly reconfigurable optical networks.

While transmission bandwidth does not represent a limit due to the huge bandwidth provided by optical fibers and Dense Wavelength Division Multiplexing (DWDM) technology, the electronic switching nodes in the core of the network represent the bottleneck in terms of speed and capacity for the overall network.

For this reason DWDM technology must be exploited not only for data transport but also for switching operations.

In this Ph.D. thesis solutions for photonic packet switches, a flexible alternative with respect to circuit-switched optical networks are proposed. In particular solutions based on devices and components that are expected to mature in the near future are proposed, with the aim to limit the employment of complex components.

The work presented here is the result of part of the research activities performed by the Networks Research Group at the Department of Electronics, Computer Science and Systems (DEIS) of the University of Bologna, Italy. In particular, the work on optical packet switching has been carried on within three relevant research projects: the *e-Photon/ONe* and *e-Photon/ONe+* projects, funded by the European Union in the Sixth Framework Programme, and the national project OSATE funded by the Italian Ministry of Education, University and Scientific Research.

The rest of the work is organized as follows. Chapter 1 gives a brief introduction to network context and contention resolution in photonic packet switches. Chapter 2 presents different strategies for contention

resolution in wavelength domain. Chapter 3 illustrates a possible implementation of one of the schemes proposed in chapter 2. Then, chapter 4 presents multi-fiber switches, which employ jointly wavelength and space domains to solve contention. Chapter 5 shows buffered switches, to solve contention in time domain besides wavelength domain. Finally chapter 6 presents a cost model to compare different switch architectures in terms of cost.

Chapter 1

Introduction to photonic packet switches and contention resolution

This chapter provides a brief description of the main concepts of photonic packet switches and contention resolution.

1.1 Network context

Dense Wavelength Division Multiplexing (DWDM) has emerged as the core transmission technology for next generation backbone networks with its ability to provide aggregate transmission rates in the Tbit/s range [1]. The application of this technology is now called to migrate from bare transmission to switching and meet both the bandwidth and flexibility requirements of future network contexts.

Different approaches have been proposed to achieve dynamic network connectivity at different levels of granularity, namely optical circuit switching, optical burst switching and optical packet switching [2], [3] [4], [5]. Correspondingly, dynamic switching node capability must be provided at different time scales. Optical switches are the key systems to enable the proliferation of transparent all optical networks. The introduction of photonic packet switches [6], [7], [8] could support this task as an important step for the proliferation of future Internet based services [9], [10], [11]. It is a common opinion that this scenario can turn into

a reality in an evolutionary perspective, where the introduction of limited cost modular switches is first adopted with design principles open to adaptability to new high performance technologies, when available and mature [12].

This work is focused on all-optical packet switches and in particular on contention resolution, one of the main and still unresolved aspects related to optical packet switches. Contention resolution concepts and the organization of this work are presented in detail in the next section.

1.2 Contention resolution in photonic packet switches

One of the main issues in photonic packet switches is contention resolution, which arises when two or more packets contend for the same node resource at the same time. While in electronic packet switches contentions are resolved mainly in the time domain by the employment of Random Access Memories (RAMs) as electronic buffers, in photonic switches this solution is more difficult to be applied, due to the lack of the optical equivalent of RAMs. For this reason in photonic switches often contentions are solved also in wavelength and space domains [13].

In this section a brief description of the techniques to manage contention resolution in these three domains is presented.

Due to the statistical multiplexing of the wavelengths obtained with DWDM technology, contention in a node can be resolved in the wavelength domain [14], [15], [16]. The principle is very simple: when packets coming to the switch from different input fibers on the same wavelength are directed to the same output fiber, one of them is directly forwarded, while the others are converted to different free wavelengths on the destination fiber, if available. In this way, loss occurs only when no more wavelength channels are available on the destination fiber. The advantage of this technique is the high flexibility and transparency assured when all-optical wavelength conversion is applied, jointly to the non-delay of the packets that are not buffered in the nodes. More, this technique takes advantage of benefits of statistical multiplexing, which assures low packet loss when the number of wavelengths becomes high enough. The main drawback is that the technology to realize the wavelength converters (es-

pecially the all-optical ones) is still immature, so they are very difficult to be implemented, and in any case their cost could be very high [17], [18]. Contention resolution in wavelength domain in treated in detail in chapter 2, where different schemes allowing to reduce both the number and the requirements of the wavelength converters are presented.

Contention resolution in space domain consists in avoiding collisions among packets on the same wavelength by exploiting spatial diversity (packets on the same wavelength are sent through disjoint physical paths inside the switch). The spatial diversity can also be exploited by providing the link of the network with a bundle of fibers instead of a single fiber (a solution called multi-fiber switch) [19], [20], [21], [22], [23]. In fact, in this case, packets on the same wavelength directed to the same link can be forwarded in different fibers. The advantage of this technique is its simplicity, while the drawback is the increasing of the number of switching elements needed to implement the switching fabrics. Multi-fiber switches to solve contention in both wavelength and space domains are presented in detail in chapter 4.

Contention resolution in time domain can be applied in two different ways. The first one is by maintaining the packets in the optical domain and exploit Fiber Delay Lines (FDLs) to delay packets [13] [24], [25], [26]. The advantage of this technique is that the packets are maintained in the wavelength domain thus assuring transparency, while the drawbacks are that FDLs only offer fixed delays (multiple of a value called granularity) and that they contributes to the degradation of the optical signal. The second alternative is to convert the packets that lose contention in electronic domain and store them in electronic buffers, obtaining an hybrid electro/optical switch [27]. This solution presents the typical advantage of RAMs (high memory capacity, possibility to extract and forward a packet in whatever moment, thus obtaining variable delay and relatively low cost) losing the optical transparency as main drawback. Two architectures, with contention resolution in wavelength and time domains, are presented in chapter 5. The first one exploits FDLs, the second one electronic buffers.

These three methods can be applied jointly to solve contention, and in the rest of the work these methods are faced in detail. In particular schemes and architectures for all-optical packet switches will be presented. Most part of the work is aimed at the reduction of both the

number and requirements of wavelength converters, that are complex and costly components. The wavelength converters are assumed to be all-optical and full-range (i.e., an input signal may be converted to any wavelengths on output), but the proposed solutions and results are also valid if electro/optical wavelength converters are considered, having in mind that, in this case, the optical transparency is lost.

First ideal schemes are illustrated, then the attention is focused on possible implementations relying on optical components that are expected to mature in the near future.

These architecture can be used in different network contexts, as optical circuit switching and optical burst/packet switching, obtaining different performance.

For all the presented architectures, scheduling algorithms are needed to manage packet forwarding through the node, in order to properly configure the hardware resources and perform contention resolution. This work is focused on optical packet switching in synchronous (time slotted) environment, so in the rest of the work proper scheduling algorithms to control the proposed architectures in this context are presented. Also various analytical models to evaluate performance in terms of packet loss probability are presented, as well as examples of results and performance graphs in order to present the effectiveness of the proposed techniques in solving contention. Particular attention is also given to the complexity of the proposed architectures in terms of expensive optical components, being the cost one of the main parameters that has to be taken into account in designing such next generation switches.

Chapter 2

Contention resolution in the wavelength domain: wavelength converters sharing strategies

This chapter first presents briefly the basic concept of contention resolution in the wavelength domain by exploiting Wavelength Converters (WCs), then illustrates different strategies to share them in all-optical buffer-less switching nodes. A simple example of contention resolution in the wavelength domain is presented in figure 2.1. In figure 2.1(a) an ideal switch without conversion capability is presented. In this situation if two packets coming on different input fibers (IFs) h and s, but in the same wavelength, are directed to the same output fiber (OF) at the same time, one of them is lost due to wavelength contention. Instead, in figure 2.1(b) a switch equipped with WCs is presented. In this case, if two packets on the same wavelength are directed to the same OF, one is sent without conversion as in the previous case, and the other is wavelength shift to another free wavelength by exploiting a WC.

To exploit properly this simple technique to solve contention, the first idea was to equip the switch with a WC for each input (or output) wavelength channel. As already mentioned, the main drawback is that WCs are the most complex components to be implemented in optical technology. For this reason, different strategies to limit the number of WCs

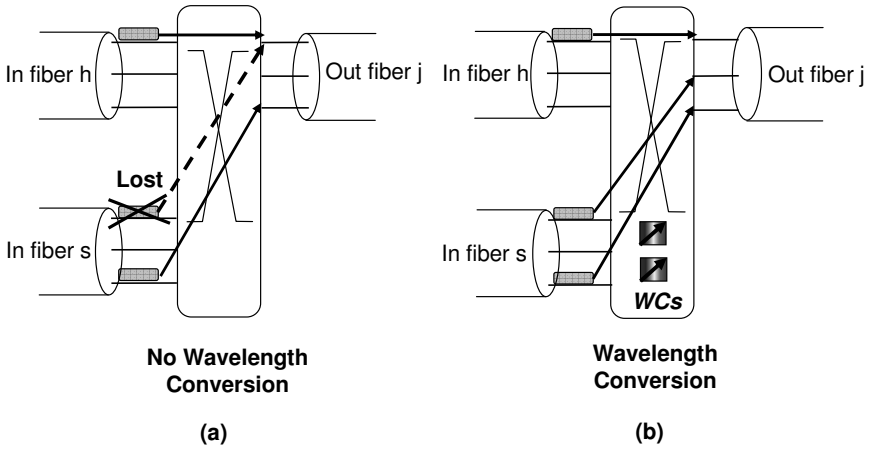


Figure 2.1: Example of contention resolution in the wavelength domain. In (a) the switch is not equipped with wavelength converters and a packet is lost due to wavelength contention, in (b) the packets can be forwarded by exploiting wavelength conversion.

employed in a node, by sharing them, have been proposed in the past [15], [16]. Two main strategies was proposed, the shared-per-link and the shared-per-node [15], [16]. In this work a third strategy is proposed, the so called shared-per-wavelength [28]. These strategies are described in detail in sections 2.1, 2.2, 2.3 respectively. Proper scheduling algorithms to control packet forwarding in synchronous context are presented, as well as analytical models to evaluate packet loss performance. More, in each section some results about the performance of the proposed switch are presented. The final section 2.4 is devoted to compare the three alternatives in terms of performance and complexity.

2.1 Shared-per-link strategy

In shared-per-link (SPL) the TWCs are shared on the output interfaces, meaning that each output interface has a dedicated pool of TWCs shared by packets directed to that interface [29]. The ideal scheme of the shared-per-link concept is presented in section 2.1.1, a proper scheduling algorithm to control packet forwarding is reported in section 2.1.2 and an

analytical model to evaluate Packet Loss Probability (PLP) is presented in section 2.1.3. In section 2.1.4 PLP, throughput and TWCs saving obtained for the shared-per-link concept are presented.

2.1.1 SPL scheme

The SPL concept, and the related ideal scheme are presented in figure 2.2. It represents a buffer-less packet switch with wavelength converters shared per output interface (optical link). It consists of N input and N output fibers each carrying a WDM signal with M wavelengths. A non-blocking space switching matrix is provided to transfer packets arriving on any input wavelength/any fiber to any output wavelength/any fiber. Each output interface is equipped with $R \leq M$ Full Range Tunable-input/Tunable-output Wavelength Converters (simply indicated with TWCs), and $M - R$ simple optical channels (OCs) without TWCs.

This switch configuration has the advantage to reduce the number of TWCs used and to avoid wavelength conversion if contention is absent. The special case $M = R$ has been studied in [30] for the synchronous environment. The WDM signal arriving to the switch from an input fiber is de-multiplexed and synchronized. Packets on input channels are then transferred by the space switching matrix to the proper output interface according to the routing information. The principle to solve contention is as follows: the control unit assigns the packet to a free OC if the wavelength of the packet is not already assigned on that output interface. If there is not free OC, a TWC is used to send the packet even if it is not wavelength shifted. If the wavelength on the output interface is busy, the packet is sent with wavelength conversion by exploiting a free TWC.

In fully equipped switch (equipped with one TWC per channel, for a total amount of NM TWCs) packet loss occurs only if all wavelengths on the destination fiber are busy [30]. Instead with the proposed switch configuration two different situations cause packet loss: i) the number of packets for the target output interface is greater than the number of wavelengths per fiber M (called output blocking), ii) the number of packets that effectively need wavelength conversion is greater than the number of TWCs R (loss due to lack of TWCs).

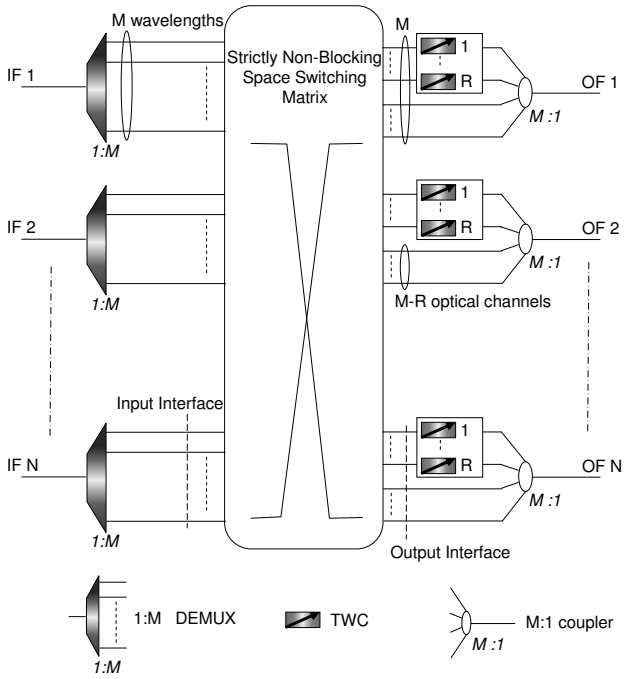


Figure 2.2: SPL switching node with N input/output fibers carrying M wavelengths each. The node is equipped with R TWCs on each output interface.

The basic all optical elements for the implementation of this architecture have been described in [31], [32], [33]. They are typically based on optical components like MEMS or SOAs, used as wavelength selectors, as well as optical filters and tunable wavelength converters [31], [32].

2.1.2 Scheduling algorithm for SPL

To perform the switching function a proper scheduling algorithm (SA) is needed to control the optical packet forwarding from the input to the output channels, by applying contention resolution through wavelength conversion. The resulting packet loss probability depends on the capability of the algorithm to find a matching between the transfer requests and available output (and internal) resources. In [15] a scheduling algorithm

that minimizes the packet loss probability is presented. This scheduling algorithm is applied to manage packet forwarding, and is here briefly described. It consists of three phases sequentially executed in each time slot (see figure 2.3). In the first step, headers of the incoming packets are read and packets coming on the same wavelength W and directed to the same output fiber j, are grouped in the set L_j^W. Packets on the same set contend for the same output channel, for this reason in the second phase one packet is randomly select from each set L_j^W and sent to the output fiber j without conversion. Such a packet can be sent on the output fiber by exploiting an OC (figure 2.3, packets from L_1^2, L_1^3, L_2^1) or a TWC, if no OCs are available (figure 2.3, the packet from L_1^4). In the third phase other packets randomly chosen are sent to the proper output fiber by exploiting wavelength conversion (figure 2.3, packets from L_1^2 and L_2^1), as long as there are available TWCs and there are free channels on output fiber. The remaining packets are lost.

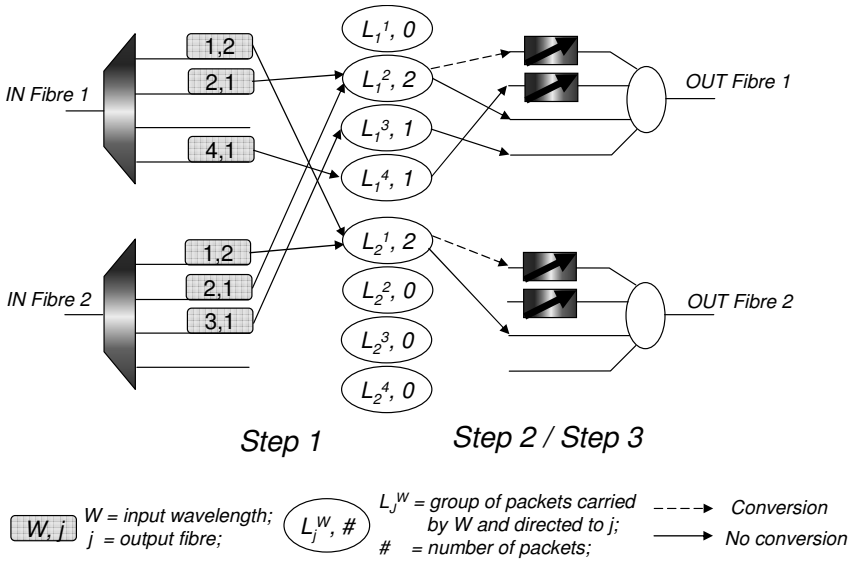


Figure 2.3: Example of packet forwarding in a time slot with the proposed scheduling algorithm: $N = 2$ input/output fibres with $M = 4$ wavelengths each, 4 packet arrivals directed to output 1 and 2 packet arrivals for output 2.

2.1.3 Analytical model of the packet loss for SPL

The proposed model assumes a synchronous packet switch and fixed length optical packets with single slot duration. Uniform traffic distribution is considered, meaning Bernoulli arrival process on input wavelengths and uniform balanced addressing to output fibers. The packet loss probability can be evaluated with reference to a generic output fiber j, due to the uniform traffic distribution and the organization of the TWCs (each output fiber has its own pool of TWCs). Moreover packet loss probability is calculated by means of the number G of different wavelengths sending packets to that fiber in a time slot. The problem is split into two distinct cases, given the total number of packets k to be transferred in a time slot:

- case 1: $M-R \leq G \leq M$. In this case G packets carried by different wavelengths can be sent without wavelength conversion. $M-R$ of them are transmitted to the simple OCs while $G-(M-R)$ packets are transmitted to the TWCs; loss occurs when $k > M$ and, in this case, $k-M$ packets are lost. The number of transferred packets is the same as in the full conversion.

- case 2: $G < M-R$. G packets carried by different wavelengths are transmitted to the output fiber; up to R packets can be transmitted to the output channels with TWCs. A total of $G+R < M$ packets can be transmitted, the remaining $k-(G+R)$ are lost.

In figure 2.4a an example of case 1 is given with $k = M = 6$ and $G = 4$; thanks to wavelength conversion no packet is lost. In figure 2.4b, instead, $k = M = 6$ and $G = 2$ so G is lower than $M-R$ (case 2); as a consequence, one packet is lost due to the lack of TWCs, while a simple OC is useless due to contention for the same wavelength.

Proper nomenclature to discuss the model is first introduced:
p probability of an arrival on a wavelength in a time slot (load per wavelength);
A_j probability of at least an arrival for output fiber j, given by $1-\left(1-\frac{p}{N}\right)^N$;
G number of different active wavelengths (with at least one packet for output j);

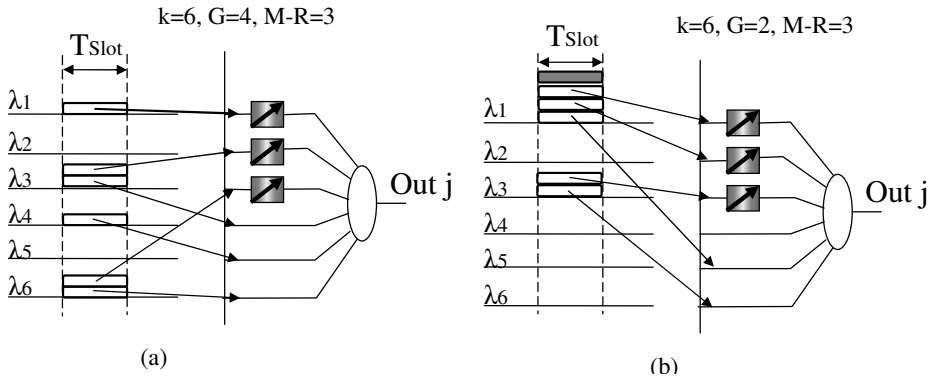


Figure 2.4: Examples of case 1 (a) and case 2 (b).

D_G probability of G active wavelengths;
$\Gamma_{h|G}$ probability of h arrivals to output j given G active wavelengths;
N_l^j number of lost packets on output j in a time slot;
N_o^j number of packets offered to output j in a time slot;
P_l packet loss probability.

P_l can be expressed in relation to the generic output j by definition as the ratio between the averages of N_l^j and N_o^j:

$$P_l = \frac{E\left[N_l^j\right]}{E\left[N_o^j\right]} \tag{2.1}$$

Under the assumption of uniform traffic the average number of packets offered to output j in a time slot, N_o^j, is given by

$$E\left[N_o^j\right] = M \cdot N \cdot p \cdot \frac{1}{N} = M \cdot p \tag{2.2}$$

and the average number of lost packets on output j in a time slot, N_l^j, is given by

$$E\left[N_l^j\right] = \sum_{G=1}^{M} E\left[N_l^j | G\right] \cdot D_G \tag{2.3}$$

where

$$D_G = \binom{M}{G} \cdot A_j^G \cdot (1 - A_j)^{M-G} \quad , \quad 1 \leq G \leq M \tag{2.4}$$

The expression given in (2.3) is calculated by taking into account the contributions of case 1 ($M-R \leq G \leq M$) and case 2 ($1 \leq G \leq (M-R)-1$) and by considering that, with G active wavelengths, the number of packet arrivals for output j is between G and GN. The following expression for $E\left[N_l^j\right]$ is obtained

$$E\left[N_l^j\right] = \sum_{G=1}^{M-R-1} \sum_{h=G+R+1}^{NG} (h-(G+R)) \cdot \Gamma_{h|G} \cdot D_G + \\ + \sum_{G=M-R}^{M} \sum_{h=M+1}^{NG} (h-M) \cdot \Gamma_{h|G} \cdot D_G \tag{2.5}$$

The expressions of the conditional probability $\Gamma_{h|G}$ can be directly calculated for $G=0$ and $G=1$:

$$\Gamma_{h|0} = \begin{cases} 1 & h=0 \\ 0 & h>0 \end{cases} \tag{2.6}$$

and

$$\Gamma_{h|1} = \frac{\binom{N}{h} \cdot \left(\frac{p}{N}\right)^h \cdot \left(1-\frac{p}{N}\right)^{N-h}}{1-\left(1-\frac{p}{N}\right)^N} \qquad 1 \leq h \leq N \tag{2.7}$$

where $1-\left(1-\frac{p}{N}\right)^N$ is the normalizing factor representing the probability of $h>0$ and $\Gamma_{h|1}=0$ for $h=0$ and $h>N$.
In order to write the recursive formula for the general case $G>1$, the example for $G=2$ is given. By considering all the combinations of arrivals from N input fibers with two active wavelengths, the following expression is obtained, where the factors $\Gamma_{x|1}$ and $\Gamma_{(h-x)|1}$ can be easily outlined:

$$\Gamma_{h|2} = \sum_{x=1}^{min\{N,h-1\}} \Gamma_{x|1} \cdot \Gamma_{(h-x)|1} = \\ = \sum_{x=1}^{min\{N,h-1\}} \frac{\binom{N}{x} \cdot \left(\frac{p}{N}\right)^x \cdot \left(1-\frac{p}{N}\right)^{N-x}}{1-\left(1-\frac{p}{N}\right)^N} \cdot \\ \cdot \frac{\binom{N}{h-x} \cdot \left(\frac{p}{N}\right)^{h-x} \cdot \left(1-\frac{p}{N}\right)^{N-(h-x)}}{1-\left(1-\frac{p}{N}\right)^N} \tag{2.8}$$

that is valid for $2 \leq h \leq 2N$, while $\Gamma_{h|2} = 0$ for $h < 2$ and for $h > 2N$.

In the general case $G \geq 2$ the accounting of all possible combinations can be again given as the product of $\Gamma_{x|1}$ and $\Gamma_{(h-x)|(G-1)}$ that, in its turn can be further factorized into a lower ordered calculation giving rise to the following recursive formula:

$$\Gamma_{h|G} = \sum_{x=1}^{min\{N,h-G+1\}} \Gamma_{x|1} \cdot \Gamma_{(h-x)|(G-1)} \quad G \leq h \leq GN \tag{2.9}$$

while $\Gamma_{h|G} = 0$ for $h < G$ and $h > GN$. This formula holds for $2 \leq G \leq M$.

Formulas (2.6), (2.7) and (2.9) allows the calculation of (2.5). Then the calculation of the packet loss probability given by (2.1) can be completed by using (2.2).

In particular the expression of the packet loss probability in the case of full wavelength conversion can be obtained by imposing $R = M$ in (2.5) given by:

$$P_l = \frac{1}{Mp} \sum_{G=0}^{M} \sum_{h=M+1}^{NG} (h - M) \cdot \Gamma_{h|G} \cdot D_G \tag{2.10}$$

In this case the packet loss takes place only when the number of packet arrivals for output j is greater than M. The resulting loss probability is the same obtaining with the formula presented in [30] for uniform Bernoulli traffic case.
Similarly, by imposing $R = 0$ the expression of packet loss without wavelength converters is obtained, that is

$$P_l = \frac{1}{Mp} \sum_{G=1}^{M} \sum_{h=G+1}^{NG} (h - G) \cdot \Gamma_{h|G} \cdot D_G \tag{2.11}$$

that is the ratio between the average number of packets per slot that need conversion and the average offered load, being G the number of packets that are transmitted on different wavelengths.

This model can be used to calculate the TWC and OC throughput, namely T_{TWC} and T_L, that represents the load per wavelength on output fibers. The TWC throughput is calculated, with reference to a generic

output j, through the average of the number of TWCs sending packets in a time slot in cases 1 ($M-R \leq G \leq M$) and 2 ($G < M-R$), weighted on the joint probability of h arrivals for output j and G active wavelengths, and by dividing this average by the total number of wavelengths, M, that can be accommodated on the output fiber. The resulting formula for TWC throughput is:

$$T_{TWC} = \frac{1}{M}\left[\sum_{G=1}^{M-R-1}\sum_{h=G}^{NG} min\left\{h-G, R\right\}\cdot\Gamma_{h|G}\cdot D_G + \right.$$

$$\left. + \sum_{G=M-R}^{M}\sum_{h=G}^{NG} min\left\{h-(M-R), R\right\}\cdot\Gamma_{h|G}\cdot D_G\right] \tag{2.12}$$

The OC throughput is obtained with similar considerations. Anyway in this case the number of OCs sending packets is only related to the number of (different) active wavelengths G. It results in:

$$T_L = \frac{1}{M}\left[\sum_{G=1}^{M} min\left\{G, M-R\right\}\right]\cdot D_G \tag{2.13}$$

2.1.4 Numerical results for SPL

Different switch dimensioning have been considered to validate the model and to give some results useful for switch design. The influence of the number R of TWCs per output fiber on PLP is first studied varying the load per wavelength as a parameter for $N = 8$ input/output fibers with $M = 16$ (figure 2.5) and $M = 32$ (figure 2.6) wavelengths. The results obtained by the application of the described analysis and by simulation with the scheduling algorithm described in section 2.1.2 show perfect agreement. Also, both figures outline that the number of TWCs per output fiber can be lower of the number of wavelengths per fiber. The number of TWCs to obtain the same performance of the fully equipped switch, that is $R = M$, ($M = 16$ in figure 2.5 and $M = 32$ in figure 2.6) is not sensibly dependent on load. Figure 2.7 shows PLP as a function of load, p, varying R as a parameter for $N = 8$ and $M = 16$. It can be seen that with $R = 8$ the same performance of the fully equipped switch

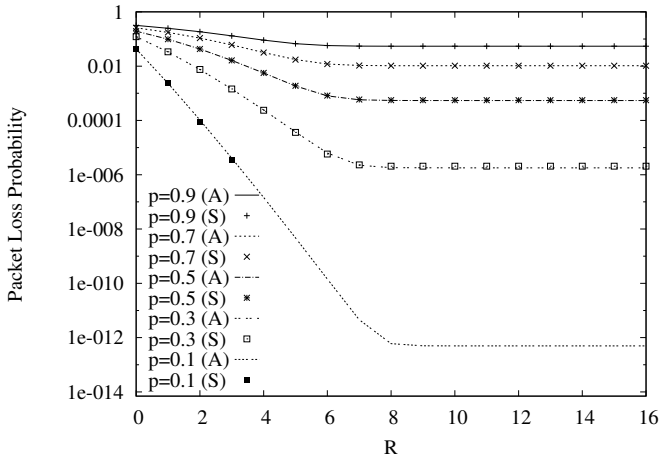


Figure 2.5: Packet loss probability as a function of R, varying load p in case N=8 and M=16. The results are obtained in both analysis (A) and simulation (S) cases.

is obtained, for all values of load per wavelength. In figure 2.8 the number of input and output fibers is varied as a parameter for $M = 16$ and $p = 0.7$. Also in this case a maximum number of converters can be found to obtain the optimal performance. Moreover it is possible to see that the PLP slightly depend on switch size. In figure 2.9 the packet loss probability as a function of R, varying the number of wavelengths M is plotted for $N = 8$ and $p = 0.7$. The number of TWCs to obtain the same performance as fully equipped switch is greater for high M respect to low M. This is mainly due to the greater number of multiplexed wavelengths. It results that the percentage of TWCs needed is almost equal to 50% of M, in the case $N = 8$, $p = 0.7$ and $M = 16$. This percentage being slightly lower for high M. Finally the TWC and simple OC throughput is analyzed by application of formulas 12 and 13, and plotted in figure 2.10 for $N = 8$, $M = 16$, $p = 0.3$ and 0.7. Throughput in the TWCs increases as R increases and, when $R = M$, the value of throughput is almost equal to the load per wavelength p, the only difference depending on PLP. The throughput on the OCs decreases as R increases and when $R = 0$, the throughput is lower than the load per wavelength due to the high PLP (being conversion not possible). In figure 2.11 the throughput

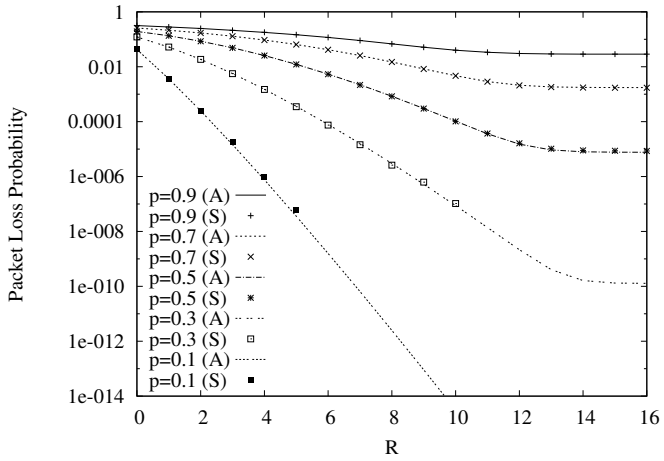


Figure 2.6: Packet loss probability as a function of R, varying load p in case N=8 and M=32.The results are obtained in both analysis (A) and simulation (S) cases.

per wavelength on switch output fiber is also presented, obtained by adding the values of TWC and simple OC throughput. It can be see that the throughput, except for very low values of R, is almost equal to the load per wavelength.

2.2 Shared-per-node strategy

In shared-per-node (SPN) concept the switching node is equipped with a single pool of TWCs that serves all input channels. In this way TWCs are shared among all incoming packets. A general scheme of the SPN switch, a proper scheduling algorithm and an analytical model to evaluate PLP are presented in sections 2.2.1, 2.2.2, 2.2.3 respectively, while in 2.2.4 performance in terms of PLP are presented.

2.2.1 SPN scheme

The reference SPN switch scheme is shown in figure 2.12. It is equipped with N input/output interfaces, each carrying M wavelengths. Full range

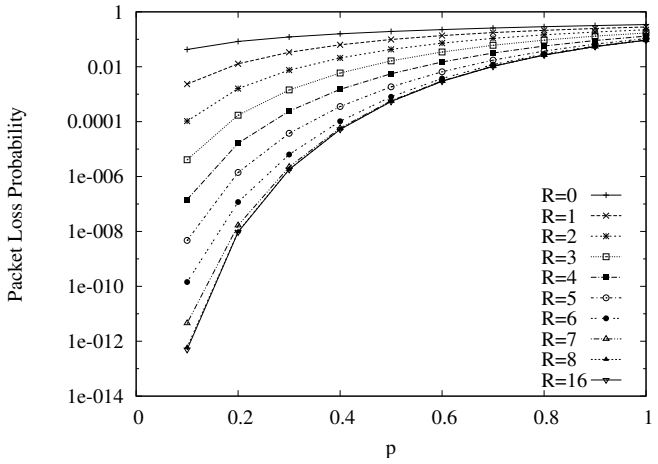


Figure 2.7: Packet loss probability as a function of the load per wavelength, p, varying the number of TWCs R for a switching node with N=8, M=16.

TWCs are grouped together in a single pool and shared among all input channels, so that an incoming packet can exploit whatever TWC. For this reason, also in this case Tunable-input/tunable-output TWCs are needed. A Fully equipped switch would require NM TWCs, one per channel, while in SPN $r \leq NM$ TWCs are considered so packet loss can occur due to the lack of TWCs.

In each time slot packets coming on different wavelengths in an input fiber are split and synchronized. A first attempt is made to forward incoming packets without wavelength conversion by exploiting the strictly non-blocking space switch. Otherwise the packet is sent to the TWC pool, if a free TWC is found, and forwarded after wavelength conversion. Channels on output interfaces are multiplexed by means of couplers. At the ingress of each coupler a maximum of M packets, each carried by a different wavelength, is allowed. Excess packets are lost due to output contention.

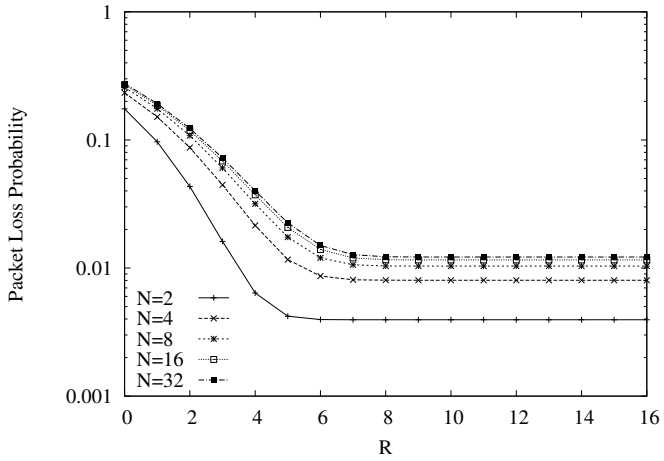


Figure 2.8: Packet loss probability as a function of the number of TWCs, R, varying the number of input/output fibers N, with M=16 wavelengths per fiber and load p=0.7.

2.2.2 Scheduling algorithm for SPN

A scheduling algorithm composed by three phases sequentially executed to control packet forwarding in SPN switches has been proposed in [15]. This scheduling algorithm provides a lower bound of packet loss given that it allows to forward the maximum number of packets in a time slot. An example of how the scheduling algorithm works is proposed in figure 2.13. In the first step, packets carried by wavelength j ($j = 1, \ldots, M$) and directed to output fiber k ($k = 1, \ldots, N$) are grouped (the corresponding group is called L_k^j). Packets in the same group contend for the same output channels, while packets on different groups are output contention free. In the second step one packet from each group (randomly chosen) is sent without conversion, so the maximum number of packets is forwarded without conversion. This two steps are the same as in the scheduling algorithm for SPL switch. In the third step the other packets are sent by exploiting wavelength conversion (the packet of L_1^3), until there are both free output channels and available TWCs in the pool.

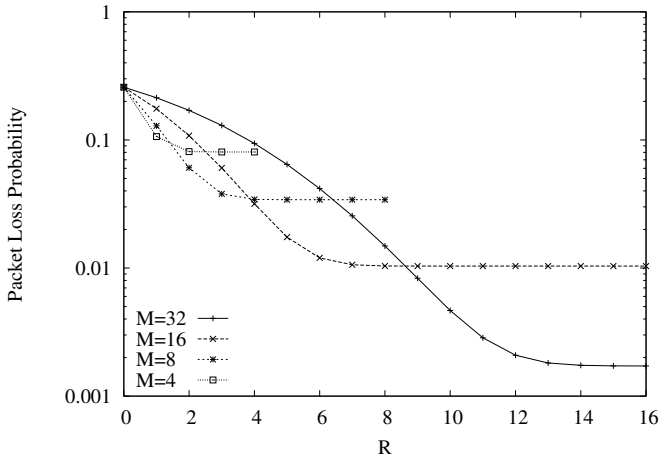


Figure 2.9: Packet loss probability as a function of R varying the number of wavelengths per fiber, with $N = 8$ input/output fiberss and load per wavelength $p = 0.7$.

2.2.3 Analytical model of the packet loss for SPN

The proposed model is based on the following hypothesis:

- independent Bernoulli arrivals on the input wavelengths with probability p in each time slot;
- arrivals on input wavelengths addressed to output fibers with uniform probability $1/N$;

The Bernoulli assumption is general but reasonably accurate. In fact it has been shown that the assembly process can absorb much of the correlation existing in the incoming peripheral traffic, e.g. IP traffic [34].

The model is developed taking into account the scheduling algorithm presented in section 2.2.2. In the proposed model packet loss probability is evaluated following a tagged incoming packet carried by wavelength j and directed to output fiber k. Two events lead to packet loss:

- more than M packets are directed to output fiber k, M of them are sent and the tagged packet is not one of them. Packet loss probability associated with this event is indicated as P_u.

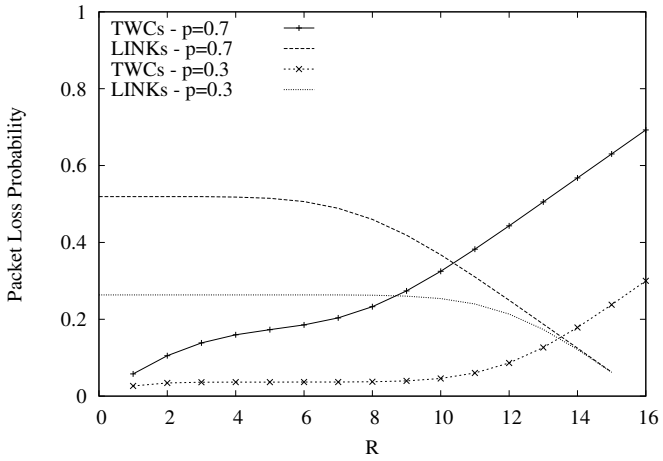


Figure 2.10: Throughput on the TWCs and Links as a function of the number of TWCs, R, in case N=8, M=16 and p=0.3, 0.7.

- more than r packets need conversion, r of them are sent to the proper output fibers by using TWCs and the tagged packet is not one of them. Packet loss probability related to this event is indicated as P_{bwc}.

The expression of the overall packet loss probability P_{loss} that takes the two above contributions into account is:

$$P_{loss} = P_u + P_b \left(1 - \frac{P_u}{P_b}\right) P_{bwc} \tag{2.14}$$

where the second term is the joint probability of P_{bwc} and $P_b \left(1 - \frac{P_u}{P_b}\right)$. The latter represents the probability that the tagged packet effectively requires conversion (joint probability that the tagged packet is blocked on its wavelength, P_b, and at least one free wavelength on the output fiber k is available, $1 - \frac{P_u}{P_b}$).

The probability P_u that the tagged packet is blocked on the destination output fiber results in:

$$P_u = \sum_{h=M+1}^{NM} \left(1 - \frac{M}{h}\right) \binom{NM-1}{h-1} \left(\frac{p}{N}\right)^{h-1} \left(1 - \frac{p}{N}\right)^{NM-h} \tag{2.15}$$

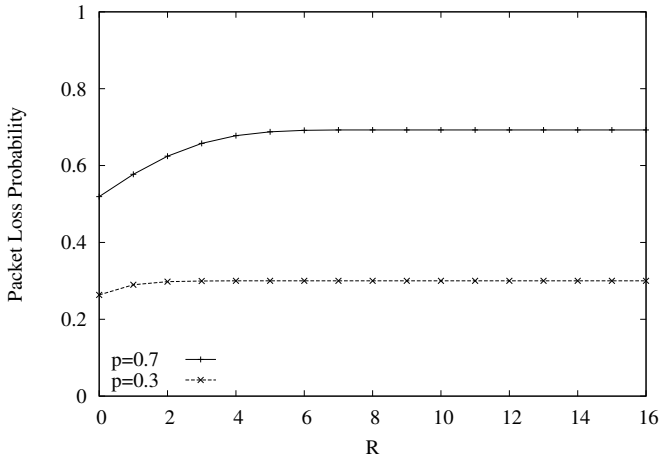


Figure 2.11: Throughput on output as a function of the number of TWCs, R, in case N=8, M=16 and p=0.3, 0.7. This is obtained by the sum of throughput on the TWCs and Links.

where the probability of h arrivals addressed to destination output fiber is expressed as the probability of $h-1$ arrivals at the other $MN-1$ input channels. Loss occurs when there are more than M arrivals and tagged packet is not among those chosen for transmission.

The probability P_b that the tagged packet is not forwarded into its wavelengths is given by:

$$P_b = \sum_{h=2}^{N} \left(1 - \frac{1}{h}\right) \binom{N-1}{h-1} \left(\frac{p}{N}\right)^{h-1} \left(1 - \frac{p}{N}\right)^{N-h} \tag{2.16}$$

by considering that there are N input fibers and the wavelengths are replicated in each of them, it is possible to have up to N packet arrivals directed to the same output fiber and carried by the same wavelength.

As a consequence the load offered to the TWC block by a single wavelength is:

$$A_{wc} = pP_b \left(1 - \frac{P_u}{P_b}\right) \tag{2.17}$$

Packet loss probability in the TWC block occurs when there are more than r conversion requests in the same time slot. The assumption of

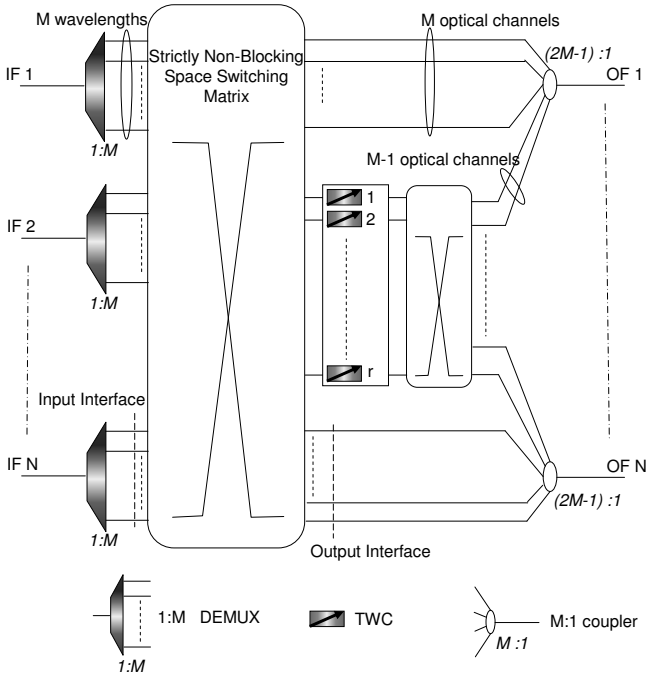


Figure 2.12: Shared-per-node (SPN) switch architecture with N input and output fibers, M wavelengths per fiber and a limited number r of TWCs.

NM independent Bernoulli arrivals at the TWC block in a time slot is made. As a matter of fact this arrivals are not independent and are negatively correlated since, for a switch with N input/output fibers, the total number of new packets arriving in each time slot at the same wavelength is no greater than N. As a consequence, each packet addressed to the output fiber g reduces the likelihood of packets destined for output fiber k, for $g \neq k$. In the extreme case, if N packets arrive during a time slot for a single output fiber g, no packet can arrive for any of the other output fibers [35], [36]. In [35], the effects of this correlation are shown to apply only when the load per wavelength is high, otherwise they can be neglected. In this context the correlation can be omitted, because, when the load is high, the packet loss due to the lack of TWCs is shadowed by the contention on output fiber. Further, the effect of this negative

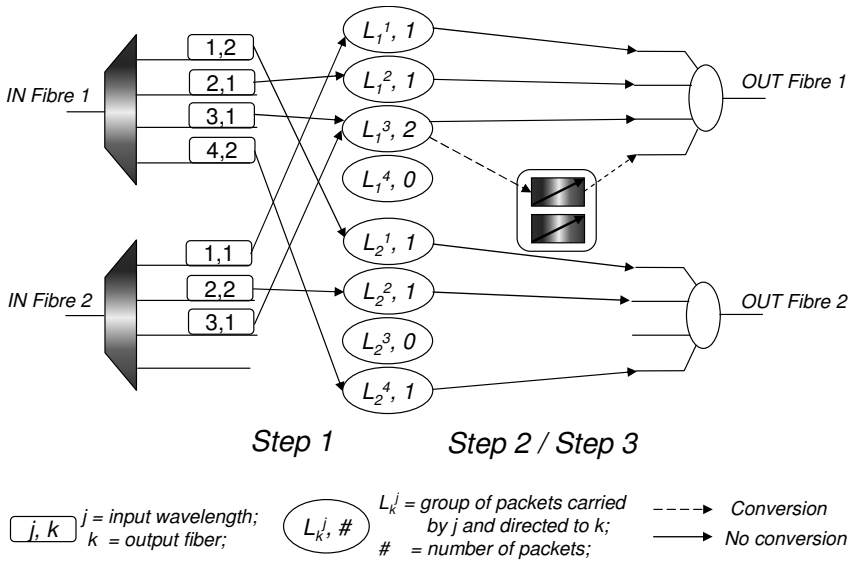


Figure 2.13: Example of the scheduling algorithm in SPN architecture with $N = 2$ input/output fibers, $M = 4$ wavelengths per fiber, $r = 2$ TWCs.

correlation decreases when the switching size N increases. Under this hypothesis, the packet loss probability due to the lack of TWCs, P_{bwc}, is calculated as:

$$P_{bwc} = \sum_{h=r+1}^{NM} \left(1 - \frac{r}{h}\right) \binom{NM-1}{h-1} (A_{wc})^{h-1} (1 - A_{wc})^{NM-h} \qquad (2.18)$$

When $r = MN$ (fully equipped architecture), $P_{bwc} = 0$ and $P_{loss} = P_u$, the same as the full wavelength conversion case. When $r = 0$, instead, $P_{bwc} = 1$ and $P_{loss} = P_b$, in fact if one packet is blocked on its wavelength, it is lost because conversion is not possible. This model makes it possible to find the minimum number of TWCs leading to the same packet loss as in the full wavelength conversion case.

2.2.4 Numerical results for SPN

In this section analytical (A) and simulation (S) results are compared. Simulation results, obtained by applying the scheduling algorithm pre-

sented in section 2.2.2 (which maximizes switch throughput), considers a confidence interval at 95% less than or equal to 5% of average. The results are presented as a function of the number of TWCs employed normalized to the number of TWCs in the fully equipped architecture, namely $\alpha = \frac{r}{NM}$. In figure 2.14 PLP is plotted as a function of α, varying the load per wavelength p, for different switching sizes ($M = 8, N = 4, 16, 64$). The figure shows very good agreement between analytical and simulation results, especially with high switch size. In addition, the same performance as fully equipped architecture ($\alpha = 1$) can be obtained with a limited number of shared TWCs. Figure 2.15 shows the PLP as a function of

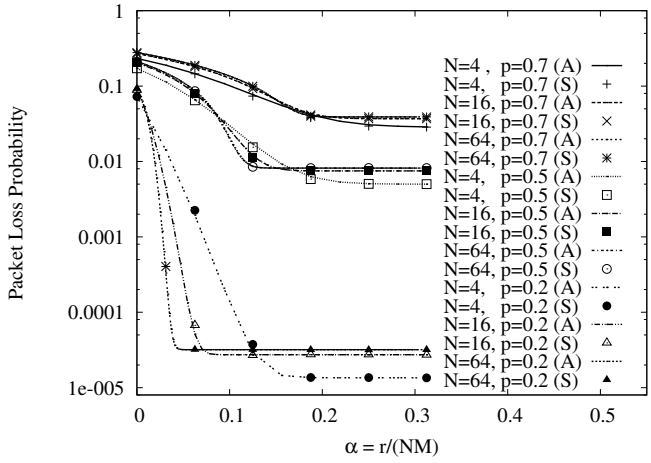


Figure 2.14: Packet loss probability of the SPN switch architecture as a function of the normalized number of TWCs, α, varying the load per wavelength, p, in case $M = 8, N = 4, 16, 64$. (A) is for analysis and (S) is for simulation.

α varying the load per wavelength p and the number of wavelengths per fiber M, for $N = 16$. Also in this case, analytical and simulation results exhibit good matching. When the number of wavelength increases, the packet loss decreases, as expected. Finally, in figure 2.16, P_{bwc} is plotted for $N = 16, M = 8$. When the load is high, the packet loss calculated with the analysis overestimates the packet loss obtained by simulation, due to the negative correlation discussed in section 2.2.3. However, in the evaluation of the packet loss, P_{loss}, when the load is high and α increases,

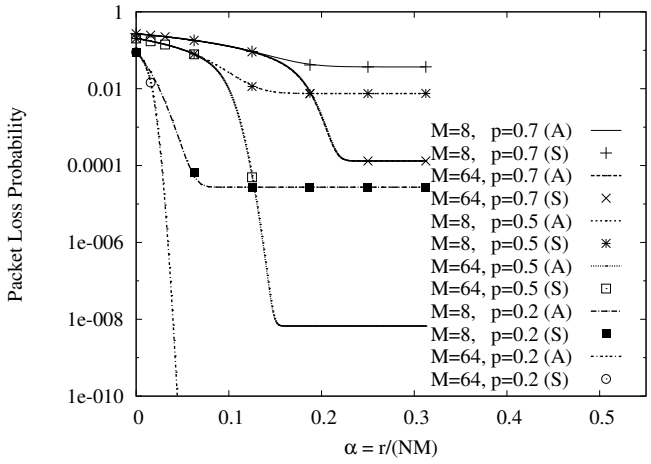


Figure 2.15: Packet loss probability of the SPN switch architecture as a function of the normalized number of TWCs, α, varying the load per wavelength, p, in case $N = 16, M = 8, 64$. (A) is for analysis and (S) is for simulation.

this difference is hidden by the high packet loss on output fiber, P_u. The model is slightly less precise when the switching size N is low ($N = 4$ for example) and both the number of wavelengths per fiber M and the load are high. In these cases the packet loss on output fiber is lower and the effect of negative correlation cannot be neglected.

2.3 Shared-per-wavelength strategy

In the shared-per-wavelength (SPW) concept the TWCs are partitioned among the wavelengths, so each wavelength of the system has its own pool of TWCs. A pool dedicated to a particular wavelength is shared among the packets coming on the same wavelength. Again, the SPW scheme, a scheduling algorithm, an analytical model to evaluate PLP and results are presented in sections 2.3.1, 2.3.2, 2.3.3 and 2.3.4 respectively.

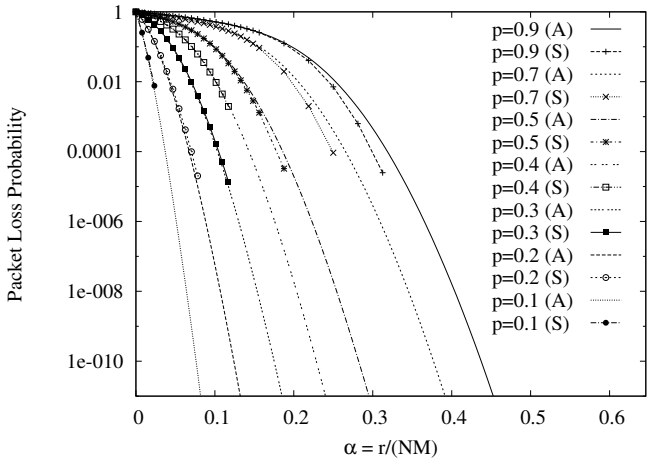


Figure 2.16: Packet loss probability due to the lack of TWCs, P_{bwc}, as a function of the normalized number of TWCs, α, varying the load per wavelength, p, in case $N = 16, M = 8$. (A) is for analysis and (S) is for simulation.

2.3.1 SPW scheme

The reference SPW scheme is presented in figure 2.17. It is equipped with N Input fibers/Output fibers each carrying a WDM signal with M wavelengths. A multi-fiber version of this concept is presented in [28], meaning that multiple fibers are available on each input and output interface. Packets arriving at the switch on different fibers on the same wavelength share the same pool of wavelength converters. M different pools of r_w wavelength converters are considered, for a total amount of Mr_w. The wavelength converters in the same pool have the same input wavelength, so Fixed-input/Tunable-output Wavelength Converters (FTWCs) are employed. In the rest of the section FTWC are again simply indicated by TWC, if not differently stated.

This kind of wavelength converters are expected to be simpler with respect to TTWCs, that makes the SPW architecture worth of interest. In addition, the SPW concept allows to organize the switch in a less costly way, given that M small switching fabrics dedicated to each wavelength can be employed instead of a single large switching fabric needed in the

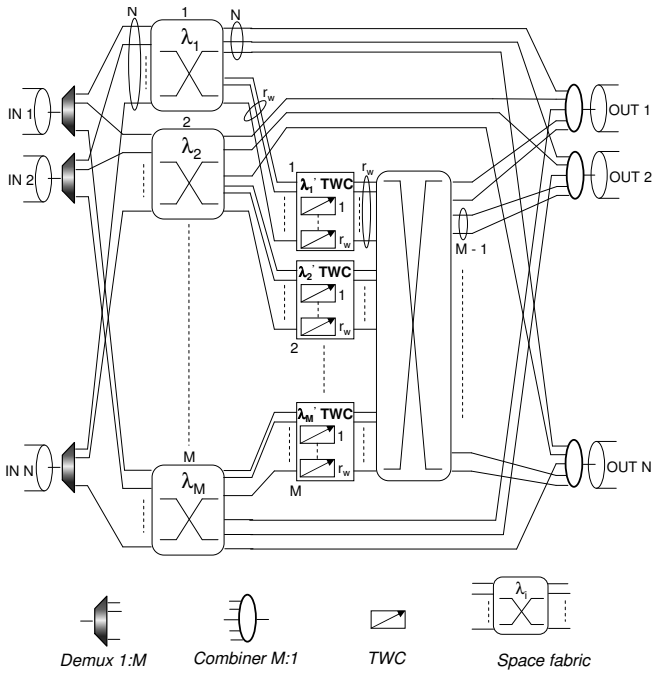


Figure 2.17: SPW reference architecture with N input and N output fiber interfaces, M wavelengths per fiber and r_w fixed-input/tunable-output wavelength converters shared-per-wavelength.

SPN and SPL schemes [28].

A packet can be lost either for output contention (not enough wavelengths in the destination output interface are available to accommodate all the packets directed to it) or inability to convert the packet (related to the lack of TWCs).

2.3.2 Scheduling algorithm for SPW

A proper scheduling algorithm is needed to manage packet forwarding in SPW switch. The scheduling algorithm proposed in [28] for the multi-fiber SPW switch can be used also for mono-fiber switch and is here briefly recalled. This algorithm aims at maximizing the number of packets forwarded without conversion thus minimizing the number of wavelength conversion requests in a time slot. In this way, it provides a lower bound of packet loss probability and is, in this sense, optimal [28]. The

scheduling algorithm is composed by three steps. Phase 1 is the initialization phase, where the variables used by the algorithm are initialized and the packets carried by the same wavelength and directed to the same output fiber are grouped together in the same set. Packets in the same set contend for the same output channel while packets belonging to different sets are channel contention free. For this reason, in phase 2 one and only one packet, randomly chosen, from each non-empty set is sent without conversion. In addition, in this phase the packets lost due to output contention are evaluated and discarded. Remaining packets, that are those not lost due to output contention and needing wavelength conversion, will contend in phase 3 for the available TWCs. In phase 3, up to r_w packets per wavelength, randomly selected, are sent to the TWC pool dedicated to that wavelength and forwarded by exploiting wavelength conversion. Remaining packets are those lost due to the lack of TWCs. It has been demonstrated in [28] that the computational complexity of phases 2 and 3 are $O(NM)$ and $O(NM + Mr_w)$ respectively, so the total complexity of the SA is $O(NM)$, given that $r_w < N$.

2.3.3 Analytical model of the packet loss for SPW

Here the attention is focused on synchronous optical packet-switched networks with fixed-size optical packets transferred through the network using a slotted statistical multiplexing scheme. Bernoulli traffic with probability p is considered, meaning that at input channels in a time slot there is a packet arrival with probability p. In a given input channel, independent arrivals in different time slots are considered due to the buffer-less nature of the proposed switching architecture. In such switches, performance are only related to the average load p, i.e. the correlation between different slots does not impact the performance [23]. Fiber-to-fiber switching is considered meaning that a packet arriving on an input fiber k and wavelength j could in principle be forwarded to any output l and wavelength m. Arrivals on different input wavelengths are independent and are addressed to the output fibers with the same probability $1/N$. Having in mind switch cost and processing optimization a first attempt is made to maintain the packet on the same wavelength to minimize wavelength conversion. An incoming optical packet is forwarded without conversion if its wavelength is not in use on the

requested output fiber, otherwise it is forwarded to the output fiber after wavelength conversion. Packet loss probability is the probability that an incoming packet on input wavelength j is discarded while it is directed to output fiber k during a given time slot. Packet loss occurs if one of the following events occurs:

- the packet loses contention on output fiber because excess packets require channels on that fiber; the probability of this event is indicated with P_u;
- the packet needs wavelength conversion, with probability P_b, but loses contention on wavelength converters because excess packets require conversion in the same time slot; the probability of this event is indicated with P_{bwc}.

By taking into account these two contributions, optical packet loss can be expressed as

$$P_{loss} = P_u + P_b \left(1 - \frac{P_u}{P_b}\right) P_{bwc} \tag{2.19}$$

where the first term P_u is the probability that packet is lost because of output blocking and the second term is the joint probability of P_{bwc} and the probability that the packet effectively requires conversion, $P_b(1 - \frac{P_u}{P_b})$. P_b is the probability that the packet is blocked on its wavelength, and $(1 - \frac{P_u}{P_b})$ is the probability that it is not discarded on destination fiber given that it is blocked on its wavelength. Each of these terms is calculated in the following.

Packet loss probability P_u is evaluated with infinite wavelength conversion capability. Let us consider a tagged arrival. For output fiber k there are up to NM packet arrivals in each time slot but only M packets can be served, so the tagged packet is lost when it is not one of them. It results in:

$$P_u = \sum_{h=M+1}^{NM} \left(1 - \frac{M}{h}\right) \binom{NM-1}{h-1} \left(\frac{p}{N}\right)^{h-1} \left(1 - \frac{p}{N}\right)^{NM-h} \tag{2.20}$$

where the probability of h arrivals is expressed as the probability of $h-1$ arrivals on input channels (wavelengths) other than the tagged.

The probability P_b that the tagged packet is blocked on its wavelength and needs wavelength conversion is given by

$$P_b = \sum_{h=2}^{N} \left(1 - \frac{1}{h}\right) \binom{N-1}{h-1} \left(\frac{p}{N}\right)^{h-1} \left(1 - \frac{p}{N}\right)^{N-h} \tag{2.21}$$

since there are N input fibers and up to N packet arrivals from wavelength j to output fiber k in each time slot.

The TWCs that perform wavelength conversion are organized in r_w blocks. The traffic entering the switch on a given wavelength is forwarded on simple optical channels if wavelength conversion is not needed, while it competes for a number of wavelength converters equal to $r_w \leq N$ to obtain wavelength conversion. Packet loss occurs if the tagged packet loses contention among packets carried by same wavelength for these r_w wavelength converters. To calculate the probability of this event (packet blocked by TWCs), the traffic offered to the wavelength converters by a single wavelength is first calculated as

$$A_{wc} = pP_b \left(1 - \frac{P_u}{P_b}\right) \tag{2.22}$$

where $1 - P_u/P_b$ takes into account that a packet is sent to the wavelength converters only if it is not blocked on output.

The final step is the evaluation of packet loss due to the lack of TWCs. Recalling that there are only r_w TWCs dedicated to one wavelength, packet loss occurs when $h > r_w$ arrivals from the same wavelength occurs at the corresponding TWC pool. Being the same wavelength replicated in N output fibers, there are up to N arrivals on a TWC pool, each with probability A_{wc}. If Bernoulli independent arrivals at TWCs is assumed P_{bwc} is given by

$$P_{bwc} = \sum_{h=r_w+1}^{N} \left(1 - \frac{r_w}{h}\right) \binom{N-1}{h-1} (A_{wc})^{h-1} (1 - A_{wc})^{N-h} \tag{2.23}$$

As a matter of fact the arrivals on TWCs in a given time slot are not independent and are negatively correlated since, for a switch with N input/output fibers, the total number of new packets arriving each time slot in the same wavelength is no larger than N. As a consequence

each packet addressed to output fiber j reduces the likelihood of packets destined for output fiber g, for $g \neq j$. In the extreme case, if N packets arrive during a time slot for a single output fiber j, no packets arrive for any of the other output fibers [35], [36]. In [35] it is shown that this dependence is evident only if the load per wavelength is high, while is slight when load is low. As a consequence, in this context the correlation can be omitted, because, when the load is high, the packet loss due to the lack of TWCs is shadowed by the contention on output fiber. Further, the effect of this negative correlation decreases when the switching size N increases.

When $r_w = N$, $P_{loss} = P_u$ the same loss as full wavelength conversion case is obtained. When $r_w = 0$, $P_{loss} = P_b$ the same loss as no conversion case is obtained, since, in this case, the packet blocked on its wavelength is lost because no conversion is possible.

2.3.4 Numerical results for SPW

Numerical results are carried out using both analysis and simulation. Simulation results are obtained by applying scheduling algorithm proposed in section 2.3.2, with confidence interval at 95% less than or equal to the 5% of the mean. The proposed analytical model is validated against and simulation results. The availability of such an accurate model can be very useful to obtain results when switch configuration leads to very low PLP, in ranges where simulation is hardly used to provide results.

Figures 2.18 and 2.19 show comparisons between simulation and analytical results for the SPW switch, with Bernoulli input traffic. PLP as a function of the number of TWCs per wavelength r_w is plotted, in the case $N = 64$, $M = 8$ and $N = 20$, $M = 80$ respectively. In the first case, with a few ($M = 8$) wavelengths per fiber, the PLP is dominated by output blocking in all regions of the graphs, and analytical results show very good agreement with simulation. In the second case, with higher number of wavelengths ($M = 80$), the output blocking is lower and, with high load, analysis overestimates simulation in the region where loss due to the lack of TWCs is high with respect to that due to output blocking. This is due to the independence assumption on the number of conversion requests in different OFs made in the evaluation of loss due to the lack

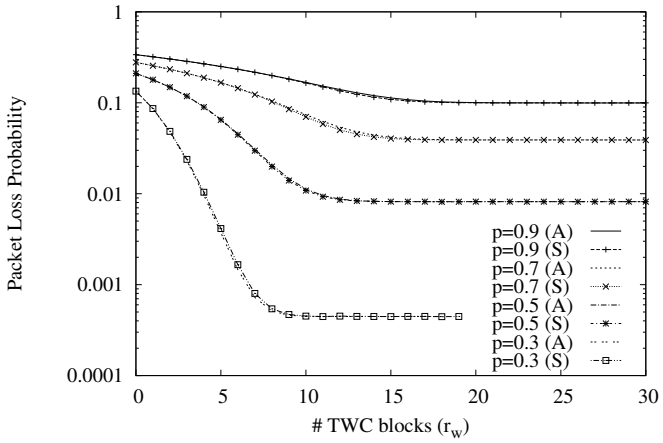


Figure 2.18: Comparison between analytical and simulation results for SPW switch with Bernoulli traffic on input. PLP is shown as a function of the number of TWC blocks, varying load, in the case $N = 64$, $M = 8$.

of TWCs, P_{bwc}. Anyway, the difference is not too high and, given that the model leads to an overestimation, it can be used to dimension the switch as a worst case. In fact the asymptotic values of PLP are exactly evaluated by the model.

2.4 Comparison among SPL, SPN, SPW

In this section the three proposed schemes are compared in terms of performance (PLP) and complexity (number of main optical components).

The SPN scheme represents the perfect sharing, given that any TWCs are shared among any packets. In SPL the TWCs are partitioned among N fibers, and each TWC pool is shared by the packets directed to the same output fiber, that are on average M. Instead, in SPW the TWCs are partitioned among the M wavelengths. The TWC pool dedicated to a wavelength is shared by N input channels (those related to the that wavelength on the input fibers). For this reason, the SPL scheme is more effective in solve contention when N is low and M is high. In fact in this case the TWCs are partitioned in few groups each shared by a relevant number of packets. On the contrary the SPW scheme performs better

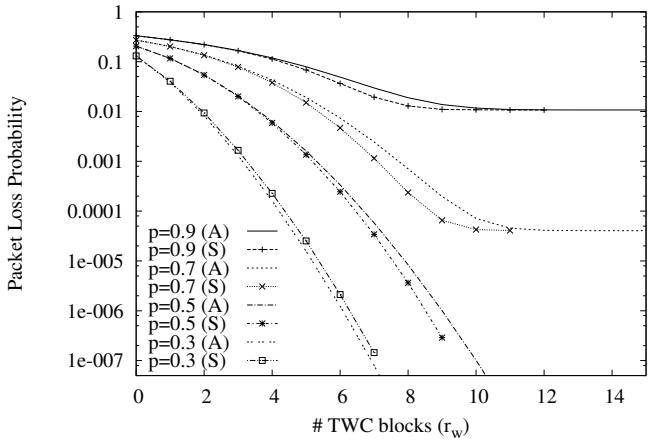


Figure 2.19: Comparison between analytical and simulation results for MS-SPW switch with Bernoulli traffic on input. PLP is shown as a function of the number of TWC blocks, varying load, in the case $N = 20$, $M = 80$.

when N is high and M is low, for the same reason. This is confirmed by observing figures 2.20 and 2.21 where the PLP obtained with SPL, SPN, SPW schemes is plotted, for different values of load (p=0.3, 0.5 and 0.8), in case $N = 32$, $M = 8$ and $N = 8$, $M = 32$ respectively. The PLP is plotted as a function of the total number of TWCs employed. In the first figure, where $N > M$, the PLP obtained with SPL switch is far to the one of the SPN switch, while the PLP obtained with SPW is nearer. In the second one, the PLP of SPL is lower than the one of SPW switch, that is far from the one of SPN.

All the schemes lead to the same asymptotic value of PLP, related to the output blocking, but the minimum number of TWCs needed to reach this asymptote differ for SPN, SPL and SPW. The number of TWCs needed is related to the switch dimensioning. For all switch dimensioning, the SPN requires the lower number of TWCs. The higher N with respect to M, the higher the additional number of TWCs needed for SPL scheme with respect to SPN. Instead, in this situation the additional number for SPW scheme is lower. The opposite is obtained when N is low with respect to M.

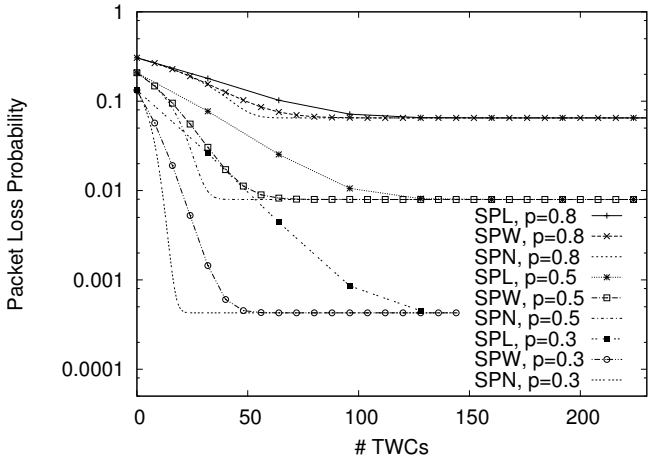


Figure 2.20: Comparison of SPL, SPN and SPW schemes, PLP as a function of the total number of TWCs varying load in case $N = 32$, $M = 8$.

Some considerations about the complexity of each scheme can be done. All schemes rely on strictly non-blocking space switching matrices with large size. The SPL scheme does not require any additional complexity with respect to the fully equipped architecture (equipped with one TWC for each output channel), while SPN and SPW schemes requires additional complexity to reach the shared TWCs and to connect them to the output fibers. This additional complexity is proportional to the total number of TWCs. In particular, as can be seen from figures 2.20 and 2.21, the SPW scheme requires a higher number of TWCs with respect to the SPN, so its complexity in terms of optical components is higher. On the other hand, the SPW scheme takes advantage by employing fixed-input/tunable-output TWCs, that are simpler to be implemented than tunable-input ones. In particular, in this work the space switching fabrics are considered as implemented by means of Semiconductor Optical Amplifiers (SOAs) used as optical gates. In fact SOAs used as gates present good properties, as very low switching time (in the range of few nanoseconds) and high extinction ratio. The number of SOAs needed for a switching fabric is given by the number of crossing point of the fabric. By observing figure 2.2 it can be deduced that the size of the switching fabric is $NM \times NM$ so the complexity in terms of SOAs of the SPL

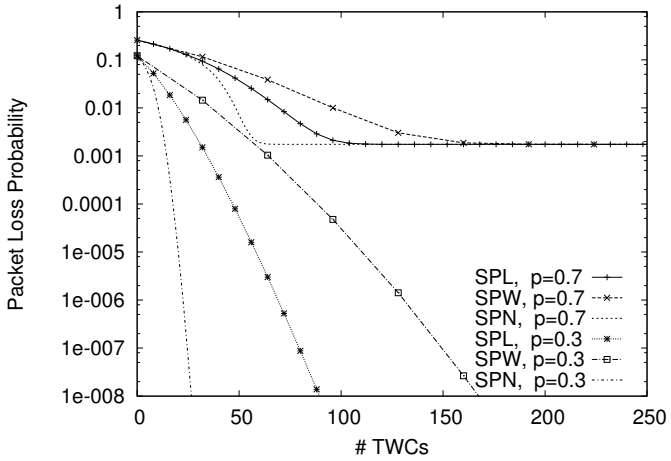


Figure 2.21: Comparison of SPL, SPN and SPW schemes, PLP as a function of the total number of TWCs varying load in case $N = 8$, $M = 32$.

scheme results in:

$$N_{SOA}^{SPL} = (NM)^2 \tag{2.24}$$

From figure 2.12 the size of the two switching fabrics needed in the SPN scheme results in $NM \times (NM + r)$ and $r \times N(M - 1)$ respectively, so the total number of SOAs is:

$$N_{SOA}^{SPN} = NM(NM + r) + rN(M - 1) = (NM)^2 + rN(2M - 1) \tag{2.25}$$

Finally by observing figure 2.17, the SPW scheme needs M switching fabrics with size $N \times (N + r_w)$ and an additional fabric with size $Mr_w \times N(M - 1)$ so the number of SOAs is:

$$N_{SOA}^{SPW} = MN(N + r_W) + Mr_wN(M - 1) = M(N^2 + NMr_w) \tag{2.26}$$

To reduce the complexity, practical implementations where strictly non-blocking matrices are replaced by simpler switching fabrics have to be defined. In this sense, the SPL scheme is not so easy to be defined, due to the fact that each TWC in an output interface should be reached from whatever incoming packet. In chapter 3 a practical implementation of the SPW concept is presented, while in chapter 4 practical implementations of both SPW and SPN schemes are presented for multi-fiber switches.

Chapter 3

Multi-stage SPW switch

In this chapter, a buffer-less multi-stage switch architecture, that is a practical implementation of the SPW concept, is presented. The multi-stage architecture allows to overcome limitations of single-stage optical cross-connects, which require either a large number of space switching elements or tunable wavelength converters that are tuned over a large number of wavelength channels [37]. Both these approaches do not scale easily in terms of capacity. The multi-stage architecture is implemented by optical components that are expected to be feasible in the near future [4]. It is based on wavelength selectors to implement the switching fabrics and on wavelength converters to solve contention. The detail of the architecture are presented in section 3.1.

A suitable scheduling algorithm to control packet forwarding by assigning switch resources on a time slot basis to incoming optical packets and resolve contentions is presented in section 3.2. In this section the need of heuristic scheduling algorithm to control the multi-stage architecture is discussed and a limited complexity solution is proposed. Discussions about traffic assumptions and analytical models are presented in section 3.3, while performance and complexity of multi-stage architecture are discussed in section 3.4.

3.1 Multi-stage shared-per-wavelength architecture

In this section the multi-stage implementation of the SPW scheme, called MS-SPW (Multi-Stage Shared-Per-Wavelength), is presented. It employs fast optical technology such as tunable transmitters and switching gates with tuning speed within the fraction of a microsecond, splitters, couplers, MUX/DEMUXes (see figure 3.1) that are expected to mature in the near future, as demonstrated in [4]. The proposed MS-SPW switch

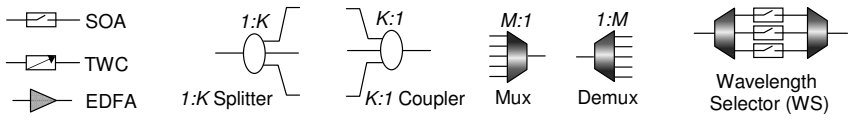


Figure 3.1: Optical components used in the implementation of the multi-stage SPW architecture.

is presented in figure 3.2. This switching matrix is equipped with N single fiber input/ output interfaces (II/OIs), each carrying M wavelengths, and is organized as a space-lambda-space (S-λ-S) architecture. In particular the S-stages rely on Wavelength Selectors (WSs), which consist of two grating Mux/Demuxes (or any device with equivalent functionality) in tandem separated by an array of M optical devices (each dedicated to one wavelength) which are able to operate as a ON/OFF gates (figure 3.1). The gate array of a WS can be implemented with Semiconductor Optical Amplifiers (SOAs) used as optical gates or with Micro-Electro-Mechanical Systems (MEMS) technology. In this work SOAs are considered given that they provide high extinction ratio and switching time in the order of few nanoseconds and their technology is quite mature.

The first and third S-stages are identical and exploit the broadcast-and-select principle, as reported previously in [39], [40] and extensively considered in literature. The principle of operation is the following: at each node input, after optical amplification by means of an EDFA (Erbium-Doped Fiber Amplifier), a power coupler is used to generate multiple copies of the multi-wavelength bundle of channels entering the node. The power coupler might have $N + 1$ outlets where one outlet per

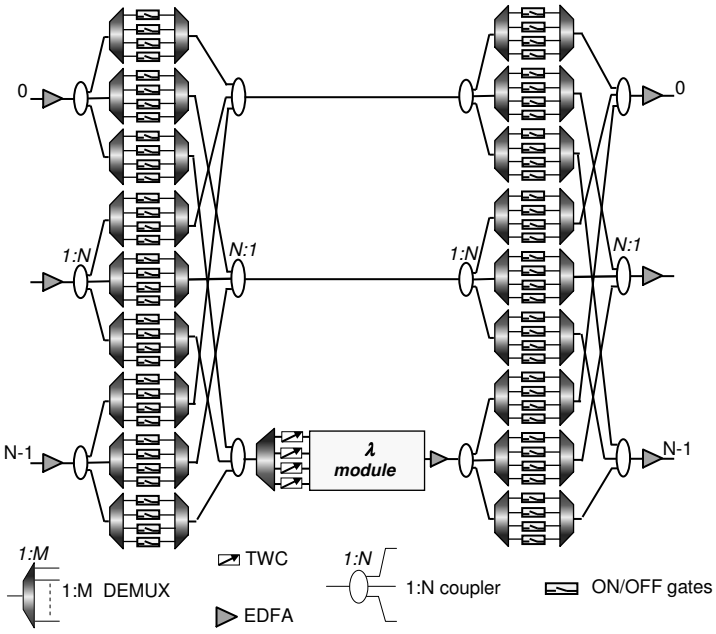


Figure 3.2: Multi-Stage Shared-Per-Wavelength (MS-SPW) switching node architecture with $N = 3$ input/output fiber interfaces, $M = 4$ wavelengths per fiber and r_w blocks of M TWCs shared per wavelength.

incoming fiber is reserved for a local drop, while N copies per input fiber are directed to a group of N WSs (only the N outlets for the transit traffic are shown in the figure). At the output side, the WSs are interconnected by means of a $N : 1$ power combiner so that only one WS from an input fiber can be coupled to the same output power combiner. Again, there might be $N + 1$ branches in the coupler to serve for local add.

Switching is achieved when the optical device is turned to the on-state, letting the wavelength pass through, whilst, when a particular wavelength must not appear on a particular output fiber, the 'gate' is switched to the off-state, blocking further propagation of the wavelength. The on-off ratio of the gate determines the level of in-band/out-of-band crosstalk. In [40] detailed information about this architecture is given. Since the architecture belongs to the broadcast-and-select family

of switching fabrics it easily allows broadcasting and multicasting.

The intermediate stage (λ) represents the conversion stage and consists of r_w TWC blocks with M TWCs each and $N - r_w$ simple optical fibers with M fixed optical channels between the two S-stages. In a TWC block, each TWC serves a different wavelength, so that the total number of TWCs dedicated to each wavelength is r_w, as in the SPW general scheme. Incoming packets on the same wavelength j (up to N) can be forwarded without wavelength conversion in the $N - r_w$ channels related to j on the optical fibers or exploit the r_w TWCs placed in different blocks. When some TWC blocks are removed and replaced with optical fibers, the cost of the switch drops but blocking rate increases. The optimum number of TWC blocks results from evaluating this cost-performance trade-off.

The outputs of the wavelength converters belonging to a block are interfaced to the λ-module. This module can be physically constructed from a variety of devices/sub-systems e.g. a) an $M : 1$ power coupler b) an $M \times K$ passive AWG router and $K : 1$ coupler c) an $M \times M$ passive router and an array of M fixed wavelength converters followed by a grating multiplexer. The role of the λ-module is to group all the M wavelengths to a single fiber. Although the options (a-c) are identical in terms of logical performance, their physical layer performance is radically different. Further, the cost difference between these three options is significant.

By comparing the MS-SPW (figure 3.2) and the SPW (figure 2.17) architectures, it is possible to observe that the 'conceptual' differences between them are represented by:

- number of optical channels. In SPW, there are M optical channels *dedicated* to each output fiber. Any channel can be used to forward a packet, independently of the wavelength of the packet. A total amount of NM optical channels is available. In MS-SPW, there are $N - r_w$ *shared* optical fibers, for a total amount of $N - r_w$ shared channels per wavelength. $(N - r_w)M$ channels are available for the whole architecture.

- organization of the TWCs. In SPW the r_w TWCs dedicated to a wavelength are grouped in a single pool. In MS-SPW the TWCs are partitioned in r_w blocks of M TWCs; each TWC in a block is

dedicated to one specific wavelength. In both cases, packets on the same wavelength can exploit up to r_w TWCs, for a total amount of Mr_w TWCs in the whole architecture.

- packet grooming. In SPW packets at the output of the TWC pools are directly sent to the proper output fibers using a strictly non-blocking space switching fabrics. In MS-SPW the wavelengths at the output of each TWC block are grouped in a single fiber. This can limit the possibility to find a matching wavelength to forward the packet (i. e. two packets in the same TWC block cannot be converted to the same wavelength even if they are directed to different output fibers).

In section 3.4 impact of packet grooming on loss performance will be shown. In any case it strongly influence the design of the scheduling algorithm, as will be thoughroghly discussed in section 3.2. Similar consideration will be done regarding the different number of optical channels.

A space equivalent of the proposed architecture is presented in figure 3.3, for $N = 3$ input and output fibers, $M = 4$ wavelengths per fiber and $r_w = 1$ TWC blocks [41]. This logical scheme helps in understanding contention occurrences and in defining a proper scheduling algorithm that optimizes resource utilization. It is composed by 3 stages, being the first and the third stages symmetrical and consisting of M $N \times N$ cross-bars, each one representing contention on the same wavelength. The intermediate stage is equivalent to r_w $M \times M$ cross-bars and $N - r_w$ groups of M simple channels. As can be seen from the figure, when fiber-to-fiber switching is considered, meaning that a packet arriving on an input fiber k and wavelength j could be forwarded to any output l and wavelength m, the fully equipped architecture ($r_w = N$) is rearrangeable non-blocking [42]. More, by observing figure 3.3 it is possible to see that a packet carried by a generic wavelength λ_i can exploit $N - r_w = 2$ optical channels and $r_w = 1$ TWCs. In this situation, if three packets carried by λ_i are addressed to the same output fiber k, one packet is sent on optical channels without wavelength conversion, one packet is sent to a TWC and one packet is lost because no further TWC block is available for wavelength conversion. When the number of TWC blocks r_w increases, the number of optical fibers $N - r_w$ decreases, so that if r_w is high, it

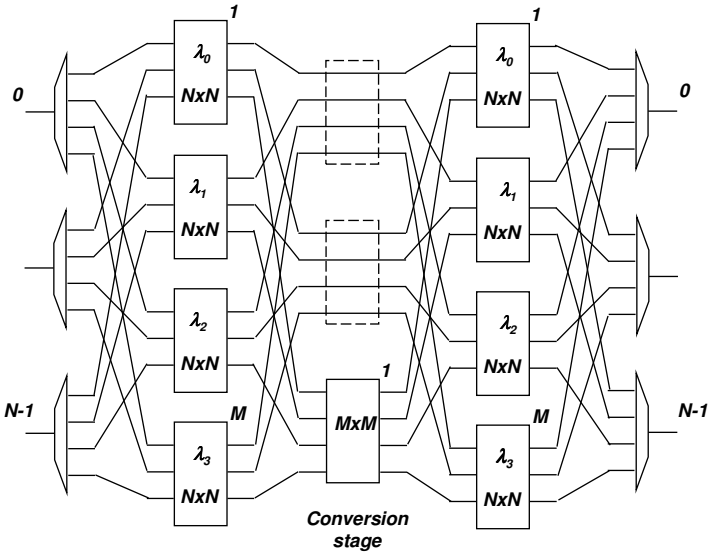


Figure 3.3: Space equivalent of multi-stage architecture with $N = 3$ input/output fiber interfaces, $M = 4$ wavelengths per fiber and $r_w = 1$ block of TWCs.

is possible that some packets that do not need conversion cannot exploit optical channels and must use TWCs even if they do not need conversion.

3.2 Scheduling algorithm for MS-SPW

The proposed scheduling algorithm aim at obtaining packet loss probability near to the one obtained with SPW switch managed with the 'optimal' scheduling algorithm (in the sense that it provides minimum PLP) presented in section 2.3.2. With respect to this scheduling algorithm, a computational complexity increases is in any case needed by the multi-stage architecture and the particular organization of the TWCs. Scheduling algorithms to manage MS-SPW architecture are presented in [38]. The reasons of computational complexity increase are discussed before describing possible scheduling algorithm implementation for MS-SPW.

3.2.1 Multi-stage organization: impact on the scheduling algorithm

The MS-SPW introduces two main constraints with respect to the SPW architecture, that are the limited number $(N - r_w)$ of optical fibers connecting input to output interfaces and the packet grooming at the output of each TWC block.

Due to the limited number of fibers $N - r_w$, only $N - r_w$ packets per wavelength can be forwarded by the second stage without wavelength conversion. When the number of packets in a particular wavelength is $h > N - r_w$, $N - r_w$ packets can be sent without conversion, while $h - (N - r_w)$ of them must be sent by the TWCs dedicated to that wavelength, each in a different block.

The packet grooming at the TWC outputs makes the definition of an optimal SA not straightforward.

In fact a packet coming on wavelength k and addressed to OI i could not be forwarded in a given TWC block j even if the TWC dedicated to k is free in that block. At least one of the free wavelengths on the destination OI i should be free also on output of block j to forward the packet. However, these wavelengths may be already in use to forward packets directed to different OIs, as shown in figure 3.4. The figure shows a possible contention situation during the execution of a generic SA. In the figure only a sub-set of the MS-SPW space equivalent is represented, that is an optical fiber a, a TWC block j and three generic OIs involved in the actual packet forwarding. Contention situation on TWC block j is shown. Suppose that the SA have already scheduled 4 packets: a packet on wavelength 1 directed to OI i is forwarded without conversion in the optical fiber a, a packet on wavelength 1 directed to OI f is forwarded in the TWC block j and converted to wavelength 2, two packets directed to OI h are forwarded in the TWC block j, the first one on wavelength 2 is converted to wavelength 4, while the second one on wavelength 4 is converted to wavelength 3. Now, the SA must decide how to schedule a new packet carried by wavelength 3 and directed to OI i. This packet cannot be forwarded by exploiting the TWC block j, even if the TWC dedicated to wavelength 3 is free (see the dotted line in the figure). In fact, the packet could only be converted to the wavelength 1 (given that the other wavelengths on output of block j are already assigned) but this

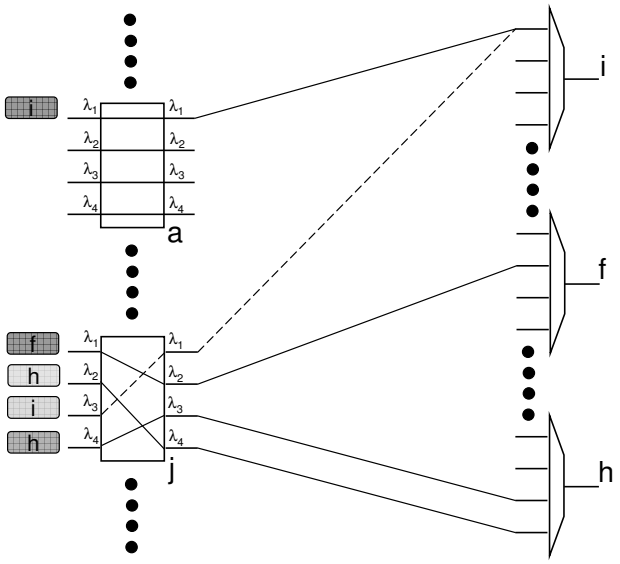


Figure 3.4: Example of the effect of packet grooming in the packet forwarding. The packet carried by wavelength $k = 3$ and directed to OI i cannot be forwarded in the TWC block j given that there are not any wavelength free in both the OI i and on output of block j.

wavelength is already in use on the OI i, to forward the packet that does not need conversion.

This example helps in understanding that a packet requiring conversion can be forwarded in a particular TWC block if and only if the intersection between the set of free wavelengths on the destination OI and the set of free wavelengths on the output of that TWC block is not empty. When this situation is not satisfied, the packet has to be forwarded in a different TWC block.

In the multi-stage MS-SPW architecture a free TWC does not assure that a packet can be converted and forwarded, differently from the SPW. The packet grooming effect limits the possibility to forward the packet in some TWC blocks. This condition leads to a double matching problem in bipartite graphs [43], where the first matching problem maximizes the number of packets allowed to exploit the TWC, and the second matching problem maximizes the number of packets forwarded by exploiting the TWC blocks. The computational complexity of this solution is very high

and a SA based on the double matching problem may be unfeasible within the time slot size used in optical packet switching. For this reason an heuristic SA with lower computational complexity is proposed.

3.2.2 Overview of the scheduling algorithm for MS-SPW

The philosophy behind the definition of the SA for MS-SPW is similar to the one proposed for the general SPW architecture and described in section 2.3.2, but it works accordingly to both the limited number of optical fibers in the second stage and the packet grooming at the output of each TWC block. The proposed SA is composed by three different phases executed sequentially in each time slot. In phase 1 variables and sets used in the SA are initialized. In phase 2 packets not requiring conversion are scheduled and packets lost due to output blocking are discarded. In phase 3 packets needing conversion are scheduled and packets lost due to conversion inability are discarded. In this phase, when a packet must be converted the TWC blocks are sequentially scanned and, as soon as a TWC block available to forward the packet is found, the packet is scheduled. A packet is lost when either a) no more TWCs dedicated to its wavelength are available or b) the TWC blocks where the TWC dedicated to that wavelength is free have been scheduled without finding one of them able to forward the packet.

Before giving the details of the SA for the MS-SPW architecture a proper nomenclature is introduced:

- $S_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is the set containing the couples (b, k), where b is a packet carried by wavelength k and directed to OI i. The information related to the wavelength is needed in phase 2 of the SA;

- $o_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is the channel corresponding to wavelength k on the OI i;

- $z_{j,k}$ $(j = 1, \cdots, r_w)$, $(k = 1, \cdots, M)$ is the channel corresponding to wavelength k on output of the TWC block j;

- Λ_i $(i = 1, \cdots, N)$ is the set containing the free channels on the OI i. The maximum cardinality of this set is M;

- NAC_i (Not Assigned Channel) $(i = 1, \cdots, N)$ represents the number of channels not already assigned on the OI i. Note that this value differs from the cardinality of Λ_i. In fact, a channel is removed from Λ_i only when that channel is scheduled for a given packet, while NAC_i is decremented by one when a given packet is considered for the transmission on the OI i but it is not scheduled for a given channel yet;

- W_j $(j = 1, \cdots, r_w)$ is the set containing the free channels on output of the TWC block j. The maximum cardinality of this set is M;

- S_i $(i = 1, \cdots, N)$ is the set containing all packets directed to OI i and not forwarded without conversion;

- $\Psi_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is the set containing the packets b directed to the OI i that must be converted;

- C_k^j $(k = 1, \cdots, M)$, $(j = 1, \cdots, r_w)$ represents the TWC dedicated to wavelength k in the j-th TWC block.

- Γ_k $(k = 1, \cdots, M)$ is the set of free TWCs dedicated to wavelength k. At each time slot, it is initialized as containing all TWCs C_k^j $(j = 1, \cdots, r_w)$, then it is updated by the SA. The maximum cardinality of this set is r_w;

- P_k $(k = 1, \cdots, M)$ is a counter where is stored the number of packets carried by wavelength k that have been already scheduled;

In phase 1 packets coming from the same wavelength k and directed to the same destination OI i are collected in the set $S_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$. Each set Λ_i $(i = 1, \cdots, N)$ is initialized as containing all the wavelength channels $o_{i,k}$ $(k = 1, \cdots, M)$ of the OI i and NAC_i) is initialized to M. Each set $\Psi_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is initialized as empty. Γ_k $(k = 1, \cdots, M)$ is initialized as containing all the TWCs C_k^j $(j = 1, \cdots, r_w)$ in the different r_w TWC blocks. The first element of Γ_k, C_k^1, correspond to the TWC dedicated to k in the first block, and so on.

The phases 2 and 3 of the SA, are described in sections 3.2.3 and 3.2.4, respectively.

3.2.3 Scheduling algorithm description: phase 2

The flow-chart of phase 2 is illustrated in figure 3.5. In this phase the

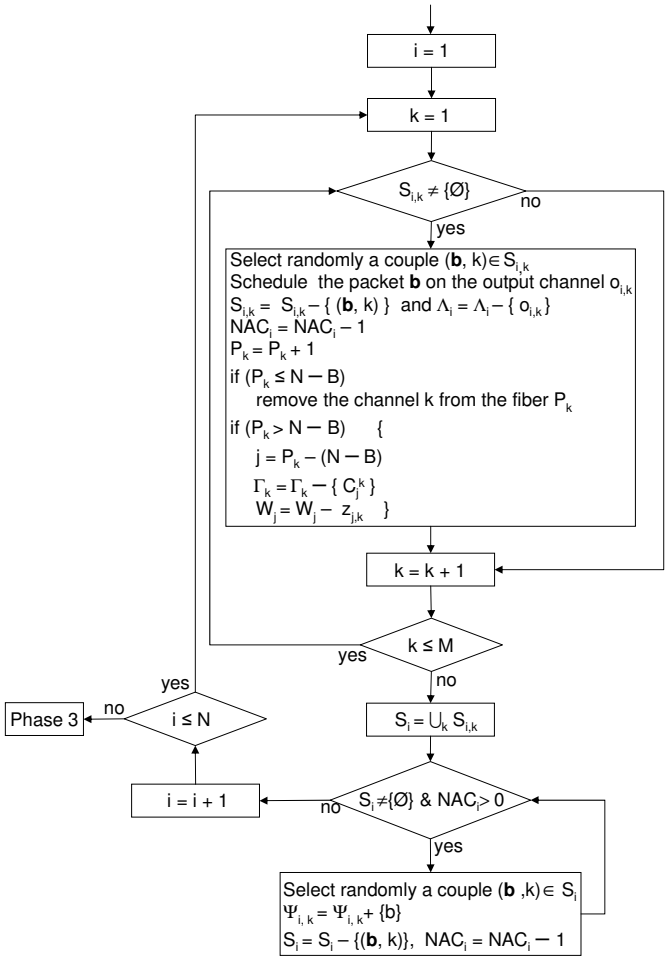


Figure 3.5: Scheduling algorithm for MS-SPW architecture: phase 2

packets forwarded without conversion are scheduled, and those lost due to *output blocking* (lack of free wavelength on the destination OI) are discarded. The OIs are considered starting from the OI 1 to the OI N, and the wavelengths of each OI are sequentially scanned, starting from wavelength 1 to M. When wavelength k on the OI i is considered a packet b randomly selected from $S_{i,k}$ (if $S_{i,k}$ is not empty) is scheduled

for the channel $o_{i,k}$ belonging to Λ_i. The packet b is removed from $S_{i,k}$, NAC_i is decremented by one given that a packet has been assigned to a channel on OI i. More, $o_{i,k}$ is removed from Λ_i given that the packet b has been scheduled for this channel. Finally, the counter P_k is incremented by one given that a packet carried by wavelength k has been scheduled. If the number of packets carried by wavelength k that have been already scheduled is less than or equal to the number of fibers in the second stage ($P_k \leq N - r_w$) the packet is sent in the wavelength channel k on the fiber P_k. In this way, the fibers are sequentially exploited.

When $P_k > N - r_w$, no more wavelength channel dedicated to k is available on the fibers. In this case the packet is sent to the TWC block even if it does not need conversion. The TWC used to forward the packet is C_k^j, with $j = P_k - (N - r_w)$. C_k^j is no longer available on the j-th TWC block, so C_k^j is removed from Γ_k. More, the channel associated to wavelength k on output of the TWC block j ($z_{j,k}$) is no longer available (the packet is forwarded without conversion), so $z_{j,k}$ is removed from W_j. After the packet is scheduled and switching resources assigned, next wavelength is considered.

When all the M wavelengths on the OI i have been considered, in the second part of this phase, the SA chooses which packets have to be forwarded with conversion and which are lost due to output blocking. Note that in this phase the SA evaluates, for each OI i, which packets have to be forwarded exploiting wavelength conversion, but it cannot assign an output wavelength channel (belonging to Λ_i) to these packets. This is due to the limits imposed by the packet grooming. The output channel assignment is left to the phase 3 of the SA.
The remaining packets destined to the OI i, after that those forwarded without conversion have been scheduled, are collected in the set S_i obtained by the union of the sets $S_{i,k}$ for $(k = 1, \cdots, M)$. After that, while both S_i is not empty and at least one output channel is available ($NAC_i > 0$), a couple (b, k) belonging to S_i is randomly chosen and the packet b is stored in the set $\Psi_{i,k}$. This set contains the packets directed to OI i and carried by wavelength k that must be wavelength shifted. The couple (b, k) is removed from S_i and the number of channels available on the destination OI i is decremented by one ($NAC_i - 1$). If NAC_i becomes equal to 0 and there are remaining packets in S_i, these packets are discarded due to output blocking on the OI i.

The set $\Psi_{i,k}$ contains the packets that are not forwarded without conversion and that are not lost due to output contention, those that must be effectively forwarded by exploiting wavelength conversion.

The sets $\Psi_{i,k}$ will be used in phase 3 to evaluate which packets may be forwarded and which are lost due to the incapability to convert the packet. The computational complexity of phase 2 is $O(NM)$, as in the SA for the SPW, given that in the worst case the number of operations needed for each OI is proportional to the number of the output channels, M.

3.2.4 Scheduling algorithm description: phase 3

In phase 3, the packets forwarded with conversion are scheduled and the remaining ones are discarded due to conversion inability. Conversion inability is related to either lack of TWCs or packet grooming. To assure fairness among the OIs, the SA starts the evaluation from the OI i indicated by a round robin counter RRF (modulo N). Then, the other fibers are cyclically considered, from the $i+1$ to the $i-1$. For each OI the SA scans cyclically the wavelengths starting from the wavelength k indicated by a round robin counter RRW (modulo M). RRF is incremented by one at each time slot while RRW is incremented by one when $RRF = 0$, every N time slots.
The flow chart of the phase 3 is presented in figure 3.6. When wavelength k on OI i is considered, until $\Psi_{i,k}$ is not empty and Γ_k is not empty, the SA exploits sequentially the TWC blocks to forward the packets belonging to $\Psi_{i,k}$. A counter n is used to scan the set Γ_k in order to find the first TWC block where it is possible to forward a packet. n is initialized to 1 every time a new set $\Psi_{i,k}$ is considered. When a TWC block j to transmit the packet is found, the packet is scheduled and C_k^j is removed from Γ_k.

The algorithm tries to employ the blocks sequentially, but some blocks are used out of order. For this reason, the n-th element of Γ_k is in general C_k^j with $j \neq n$. When a block j where C_k^j is free is found, the SA looks for a free wavelength in $\Lambda_i \cap W_j$. If at least one wavelength is found, a packet randomly chosen from $\Psi_{i,k}$ is scheduled for the first free wavelength h belonging to $\Lambda_i \cap W_j$. The packet is removed from $\Psi_{i,k}$, the channels $o_{i,h}$ and $z_{j,h}$ are removed from Λ_i and W_j, respectively, and the $n-th$ element

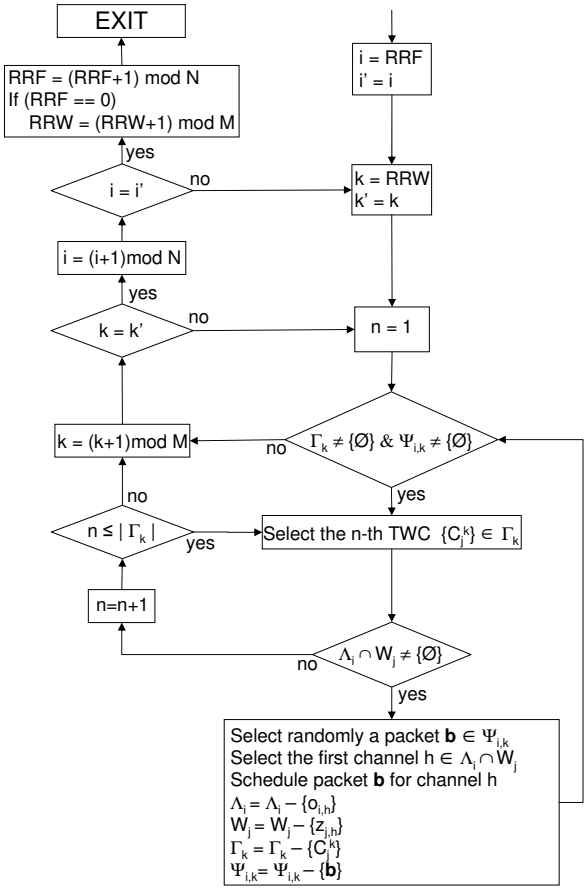


Figure 3.6: Scheduling algorithm for MS-SPW architecture: phase 3

(C_k^j) is removed from Γ_k. Then, another packet from $\Psi_{i,k}$ is considered. n is not updated given that for the next packet belonging to $\Psi_{i,k}$, the first available TWC is placed again in the n-th position of Γ_k.

Instead, when the packet cannot be forwarded in a particular block j ($W_j \cap \Lambda_i = \oslash$), n is incremented by one so that the next element of Γ_k is considered. When $\Psi_{i,k}$ becomes empty or all the element of Γ_k have been considered without find a block to forward a packet, the next wavelength is considered. When all the wavelengths of an OI have been scanned, next OI is considered. When all OIs have been considered, the SA ends.

If all packets could be accommodated in the first scanned block, the

same complexity as for the SA of SPW architecture, $O(NM + Mr_w)$, will be obtained. But, in MS-SPW the TWC blocks have to be scanned in order to forward a packet. This makes the computational complexity evaluation of the phase 3 a not trivial problem.
The most unlucky packets scans all blocks and, anyway, could not be forwarded. Anyway, it has been evaluated by simulation that the largest part of packets are sent in the first considered TWC block. More, also for the others, a few cycles are enough to find a block available to forward them. The complexity increase due to the scansion is evaluated by simulation, in the worst case, that is load $p = 1$ and number of TWC blocks $r_w = N$ (fully equipped architecture). The number of blocks scanned to forward is counted for each packet and added to a counter BC. This number is divided by the number of packets needing conversion ($PC = \sum_{i=1}^{N} \sum_{k=1}^{M} \Psi_{i,k}$). The obtained value represents the average number of cycles needed for each packet, and the values obtained for different switch configurations are presented in the table 3.1. In the SA for the SPW architecture (a sort of reference scheduling) the average number of cycles is equal to one, given that a packet is forwarded if a TCW is available, without any additional cycles. It can be seen that the num-

		M				
		4	**8**	**16**	**32**	**64**
N	**4**	1.1416	1.0903	1.0526	1.0283	1.0142
	8	1.3223	1.2383	1.1557	1.0935	1.0525
	16	1.5267	1.4054	1.2834	1.1833	1.1117
	32	1.7863	1.5955	1.4275	1.2942	1.1940

Table 3.1: Average number of cycles per packet for different switch configurations, in the case $p = 1$, $r_w = N$ (worst case).

ber of cycles needed decreases as the number of wavelengths increases. It means that the complexity of the SA does not increase as the number of wavelength increases, on the contrary the higher is the number of wavelength, the lower is the additional complexity with respect to the SPW architecture algorithm. When the number of input/output fibers increases, the number of cycles needed increases, the difference is smooth, in fact if $N = 4$, $M = 8$ the number of cycles per packet is 1.0903 while

when $N = 32$, $M = 8$ the number of cycles per packet is 1.5955.

In conclusion, the SA for the MF-SPW architecture has a computational complexity that is slightly higher than the SPW, but it is in the same magnitude of order, given by $O(NM + Mr_w) = O(NM)$.

3.3 Traffic assumption and blocking analysis for MS-SPW

Two main different traffic assumptions are considered regarding the arrivals on switch input channels:

- Bernoulli arrivals, meaning that arrivals in different slots are independent and characterized by the average arrival rate p
- Admissible traffic, meaning that arrivals are still characterized by mean p but no more than M packets arrive in a slot for the same output fiber

Bernoulli traffic can be considered as representative of the traffic in connection-less optical packet-switched networks as the result of statistical multiplexing of an high number of optical packets.

In a given input channel, independent arrivals in different time slots are considered due to the buffer-less nature of the proposed switching architecture. In such switches, performance are only related to the average load p, i.e. the correlation between different slots does not impact the performance [23].
Admissible traffic could, on the other hand, be considered as the result of the admission operation performed on optical packets that makes the traffic at each node to avoid switch output overbooking in each time slot, that is no more than M packets are admitted on the same output interface. Anyway also admissible traffic needs wavelength conversion to resolve contention in the wavelength domain and could run into switch internal blocking due to switch resource unavailability. Admissible traffic is useful to evaluate packet loss due to the unavailability of internal resources when the architecture is controlled with the proposed scheduling algorithm. For both kinds of traffic fiber-to-fiber switching is considered.

The model proposed for the SPW architecture in section 2.3.3 can be used to evaluate performance of the MS-SPW architecture with Bernoulli

input traffic, as will be shown in section 3.4. In fact, both the packet grooming and the limited number of optical fibers does not affect significantly the performance of the MS-SPW (see 3.4), when the proposed SA is applied, with respect to the SPW. More, note that the packet grooming is extremely hard to be analytically treated. This model for the SPW switch with Bernoulli input traffic is here extended to consider admissible traffic.

Analytical model with admissible traffic.
When admissible traffic is considered, no more than M packets addressed to the same output port arrive in a time slot. In this situation, the architecture is output contention free, so the term P_u evaluated in formula (2.20) becomes 0. Also, no more than M packets can arrive on the same wavelength addressed to the same output fiber, given that for the same output fiber there are maximum M packet arrivals in total. The traffic offered to the TWCs by a single output wavelength is evaluated by taking into account the constraint of maximum M packets addressed to the tagged output fiber. The expression of A_{wc} results in:

$$A_{wc} = pP_b \sum_{h=0}^{M} \binom{NM}{h} \left(\frac{p}{N}\right)^h \left(1 - \frac{p}{N}\right)^{NM-h} \tag{3.1}$$

By considering independent arrivals at the TWC blocks, P_{bwc} can be calculated by (2.23) and the final expression of the packet loss with admissible traffic is obtained by imposing $P_u = 0$ in (2.19):

$$P_{loss} = P_b P_{bwc} \tag{3.2}$$

where the expression of P_b is evaluated taking in account the maximum number of packets carried by the same wavelength and directed to the same output fiber, that is $min\{N, M\}$. In addition, a normalizing factor is introduced, so that P_b is given by:

$$P_b = \frac{\sum_{h=2}^{min\{N,M\}} \left(1 - \frac{1}{h}\right) \binom{N-1}{h-1} \left(\frac{p}{N}\right)^{h-1} \left(1 - \frac{p}{N}\right)^{N-h}}{\sum_{h=2}^{min\{N,M\}} \binom{N}{h} \left(\frac{p}{N}\right)^h \left(1 - \frac{p}{N}\right)^{N-h}} \tag{3.3}$$

In this case the independence assumption is less accurate than for Bernoulli traffic, due to the finite set of arrivals according to the admission procedure that enhance correlations as will be shown in model validation.

3.4 Numerical results for MS-SPW

Numerical results are carried out using both analysis and simulation. Simulation results are obtained with confidence interval at 95% less than or equal to the 5% of the mean for both MS-SPW and SPW architecture. These results, obtained by applying the SAs presented in sections 3.2 and 2.3.2 for MS-SPW and SPW respectively, are used to compare MS-SPW and SPW in both performance (3.4.1) and complexity (3.4.2) perspectives. Then, in section 3.4.3 the analytical models proposed in sections 2.3.3 and 3.3 for Bernoulli and admissible traffic respectively are validated against simulation.

3.4.1 Performance comparison for MS-SPW and SPW

The two architectures are compared in terms of PLP. The optimal SA (providing minimum packet loss) is applied to manage packet forwarding in the SPW switch. For this reason, results concerning SPW provide a lower bound of PLP obtainable with MS-SPW switch, and represent, in this sense, a sort of reference. As a consequence performance of MS-SPW architecture cannot be better than SPW switch performance, independently of the SA employed. The SA proposed in section 3.2 aims at obtaining performance close to that of the SPW switch.

In figure 3.7 SPW and MS-SPW switches are compared with Bernoulli input traffic. The PLP as a function of the number of TWC blocks r_w, varying the load is plotted, for $N = 64$, $M = 8$. As can be seen in figure 3.7, MS-SPW switch provides performance very close to the one of the SPW architecture. Also, the PLP presents asymptotic values when the number of TWC blocks increases, which is determined by loss due to output contention. For this reason, the same packet loss as fully equipped switch can be obtained with a limited number of TWCs. Figure 3.8 shows the same comparison in with higher number of wavelength $N = 20$, $M = 80$ which is representative of the fact that the number of wavelengths is expected to increase. Also in this case, the two switches provide quite the same performance and the asymptotic values of PLP are achieved by increasing r_w (note that when the number of wavelengths is higher, the asymptotic values of PLP rapidly decreases, due to the

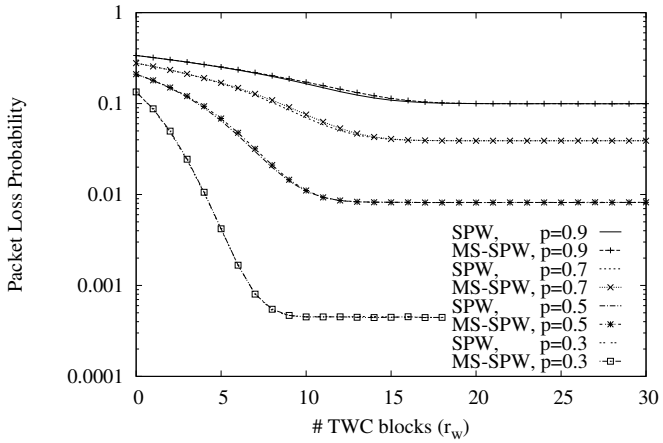


Figure 3.7: Comparison between packet loss probability of SPW and MS-SPW switches with Bernoulli traffic on input. Packet loss as a function of the number of TWC blocks r_w, varying load is plotted, in the case $N = 64$, $M = 8$.

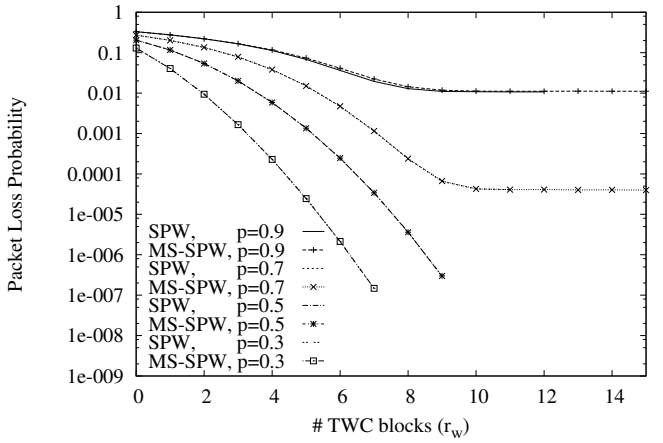


Figure 3.8: Comparison between packet loss probability of SPW and MS-SPW switches for Bernoulli traffic on input. Packet loss is shown as a function of the number of TWC blocks r_w, varying load, in the case $N = 20$, $M = 80$.

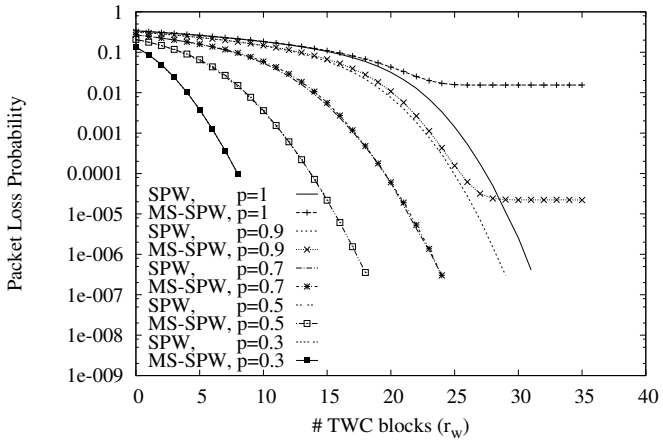


Figure 3.9: Comparison between packet loss probability of SPW and MS-SPW switches for admissible traffic on input. Packet loss is shown as a function of the number of TWC blocks, varying load, in the case $N = 64$, $M = 8$.

advantage of statistical multiplexing to solve output blocking). Only a slight difference in the region where the asymptotic values of PLP is not achieved can be observed, when load is high. With Bernoulli input traffic, the packet loss is dominated by output contention, so the effect of packet grooming is negligible and the MS-SPW architecture controlled by the proposed SA obtains results which are, in practice, equal to the lower bound of PLP. This confirm that such heuristic SA can be used instead of much more complex 'optimal' SA, while assuring very good performance anyway. Figures 3.9 and 3.10 show performance comparison between SPW and MS-SPW with admissible traffic on input, in the case $N = 64$, $M = 8$ and $N = 20$, $M = 80$, respectively. Again, PLP as a function of the number of TWC blocks r_w, varying load is plotted. Under admissible traffic, that is without any output contention, the PLP tends to zero in the SPW architecture when the number of TWC blocks r_w increases. In the MS-SPW architecture, the trend is the same but, under high load, the PLP presents an asymptote, which is related to the packet grooming effect. This effect is not negligible as in the Bernoulli case, where it is shadowed by the loss due to output contention.

In the figures this effect is relevant only for load values $p = 0.9$ and 1,

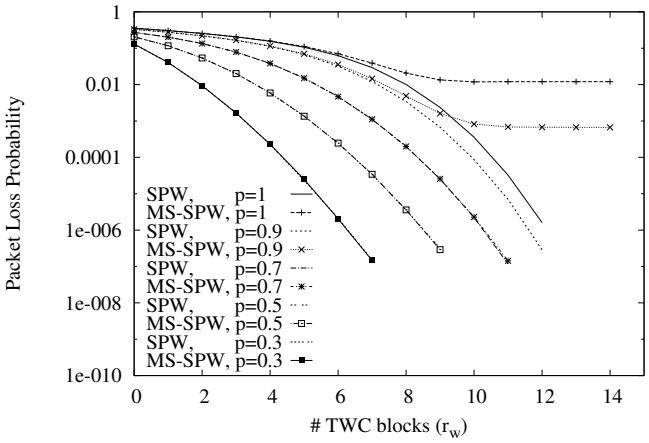


Figure 3.10: Comparison between packet loss probability of SPW and MS-SPW switches for admissible traffic on input. Packet loss is shown as a function of the number of TWC blocks, varying load, in the case $N = 20$, $M = 80$.

which are limiting loads and not typically applied in designing buffer-less switching architectures. In lower load range ($p = 0.3 : 0.7$) the asymptotic value is not present in the range of PLP evaluated by simulation, which is here greater than 10^{-6}. The asymptotic values are evidenced for higher PLP in figure 3.10 than in 3.9. This is due to the fact that the in the former case the number of IF/OF is $N = 20$, that represents also the maximum number of TWC blocks, while in the latter case, the number of IF/OFs is $N = 64$. The packet grooming effect becomes more and more relevant when the switch size N is low. In this case, at high load, the PLP related to packet grooming effect increases, given that a packet can exploit a fewer number of blocks to be forwarded.

By considering figures 3.7-3.10, and other evaluations not reported here due to space reasons, some more remarks can be carried out. The MS-SPW switch, when used in a typical packet switching scenario, can provide performance very close to SPW switch, for all switch dimensioning. Moreover, this scenario is here represented by Bernoulli traffic but, due to the buffer-less nature of the switch, the same results holds for traffic scenarios like circuit switching where circuit are constructed by concatenation of slots or ON/OFF traffic. The MS-SPW switch allows

to obtain the same performance as fully equipped architecture ($r_w = N$ TWC blocks) with a limited number of TWCs. Due to the SPW sharing strategy, these kind of architectures (both SPW and MS-SPW) obtains good performance when the number of IF/OFs is high. With high N, a high number packets on the same wavelength shares the same r_w TWCs, thus allowing to save a remarkable number of TWC blocks.

As the admissible traffic, the packet grooming has a relevant effect when the load per wavelength is very high, and in particular when N is small. Again this is related to the number of TWC blocks available in the MS-SPW architecture. In the scenario where the a pre-allocation of packets allowed on the OFs is made, that is under admissible traffic, the MS-SPW is a good solution only if the load is not too high, that seems to be in any case a realistic case.

To remove the asymptotic behavior of PLP under admissible traffic two different options are available. The first one consists in implementing more complex SAs, aiming at reducing the packet grooming effect on the TWC blocks, with consequently higher computational complexity. The second one consists in reducing the packet grooming effect occurrence by adding some optical fibers in the intermediate stage of the MS-SPW architecture, thus increasing the connectivity between the first and the third stage of the architecture. In this way, a higher number of packets can be forwarded through the simple optical fibers (without conversion) thus reducing the number of packets needing to exploit the TWC blocks. The TWC links are not more used to send packets which do not require conversion, as in the previous case, and the TWC blocks can be fully exploited by packets requiring conversion, thus reducing the packet grooming effect. In this case the computational complexity of the SA does not change, but the cost of the architecture increases, given that additional Wavelength Selectors (WSs) are needed to reach the additional K optical fibers in the intermediate stage, as will be shown in section 3.4.2. In this paper the second solution is presented, because it is preferable not further increase the complexity of the SA. In figure 3.11 the PLP when K additional fibers are added to the MS-SPW architecture is plotted, with admissible traffic on input. When the number of additional optical fibers K increases, the asymptotic value of PLP for MS-SPW architecture decreases, for both values of load $p = 0.9$ and 1. With few additional optical fibers ($K = 1, 2$), the PLP can be lowered

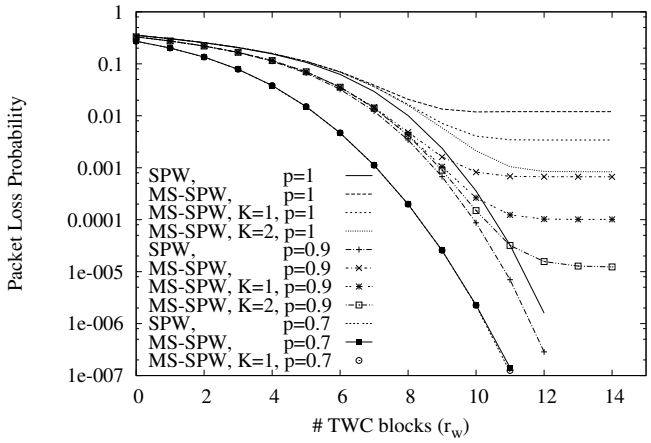


Figure 3.11: Comparison between packet loss probability of SPW and MS-SPW switches for admissible traffic on input. Packet loss is shown as a function of the number of TWC blocks, varying load, in the case $N = 20$, $M = 80$.

to negligible values even for $p = 0.9$, and the performance of MS-SPW tends to be close to the one of the SPW switch.

3.4.2 Complexity comparison for MS-SPW and SPW

The SPW and MS-SPW are compared in terms of optical components, when the SPW switch relies on the same SOA technology as the MS-SPW. For a given PLP the two architectures require the same number of TWCs, so the comparison is made in terms of the number of SOA gates needed.

The number of SOAs needed for the SPW has been evaluated in formula (2.26) of section 2.4, while the number of SOAs for MS-SPW switch is here evaluated. In the MS-SPW architecture, there are M SOA gates in each WS. In the first stage, there are N WSs for each IF, so the total number of WSs in the first stage is N^2 and the number of optical gates in the first stage is MN^2. The third stage is identical to the first,

so the total number of SOA gates in the MS-SPW results in:

$$N_{SOA}^{MS-SPW} = 2MN^2 \tag{3.4}$$

The index CSI (Cost Saving Index) is introduced to evaluate the cost saving of the MS-SPW architecture with respect to the SPW one, and defined as the ratio between the number of SOA gates needed for SPW (formula 2.26) and MS-SPW (formula 3.4), respectively:

$$CSI = \frac{N_{SOA}^{SPW}}{N_{SOA}^{MS-SPW}} = \frac{M(N^2 + Nr_wM)}{2MN^2} = \frac{1}{2} + \frac{r_w}{2}\frac{M}{N} \tag{3.5}$$

The higher the CSI is, the higher the saving allowed by MS-SPW with respect to SPW. As it can be deduced from formula (3.5), the cost saving allowed by MS-SPW architecture: a) increases as the number of TWCs employed increases (in fact, in MS-SPW the number of TWCs employed does not depend on r_w while in SPW it increases linearly with r_w) b) increases as the number of wavelengths per IF/OFs increases c) decreases as the number of IF/OFs increases. These switch architectures, under Bernoulli traffic, aim at obtaining the asymptotic values of PLP with the lowest number of TWCs. The maximum value of CSI is obtained when $r_w = N$, and is $CSI = \frac{M+1}{2}$.

If K additional optical fibers need to be added in the second stage, N WSs are needed to connect the IFs to an additional fiber, and N to connect each additional fiber to the OFs. $2NK$ additional WSs are needed, so $2MNK$ additional SOA gates are required. In this case, the total number of SOA gates in the MS-SPW architecture is:

$$N_{SOA}^{MS-SPW,K} = 2M(N^2 + NK) \tag{3.6}$$

and in this case CSI becomes:

$$CSI^K = \frac{N_{SOA}^{SPW}}{N_{SOA}^{MS-SPW,K}} = \frac{M(N^2 + Nr_wM)}{2M(N^2 + NK} = \frac{N + Mr_w}{2(N+K)} = \frac{N}{N+K}CSI \tag{3.7}$$

In table 3.2 values of CSI are presented as a function of N and M, with Bernoulli input traffic and load $p = 0.5$. The minimum number of TWC blocks r_w needed to achieve the asymptotic value of PLP is considered. As an example, by observing figure 3.7 the minimum number of TWC

		M				
		4	**8**	**16**	**32**	**64**
N	**4**	1.5	3.5	6.5	16.5	32.5
	8	1.25	2.5	5.5	12.5	32.5
	16	1	2	4.5	9.5	22.5
	32	0.875	1.5	3	6.5	16.5
	64	0.8125	1.3125	2.5	5.25	11.5

Table 3.2: Cost Saving Index (CSI) for different switch configurations, with Bernoulli input traffic and load $p = 0.5$.

blocks when $N = 64$, $M = 8$, that is $r_w = 13$ can be obtained, leading to a value of $CSI = 1.3125$. These values in the table confirm that the cost saving is highly influenced by the number of wavelengths in the system. In particular the MS-SPW switch becomes more and more convenient (in terms of optical components) with respect to the SPW one when the number of wavelengths increases. For example, for $N = 32$, $M = 32$, the MS-SPW is 6.5 times less expensive than SPW in terms of SOA gates. This confirm the better scalability of the proposed multi-stage solution, in terms of cost. Also, it can be seen that when N increases, CSI decreases, but with a lower rate than in the previous case. Anyway, if enough wavelengths are considered, the MS-SPW switch is always less costly than SPW.

3.4.3 Model validation

As shown in section 3.4.1, MS-SPW can provide performance very close to the one of SPW under Bernoulli and admissible traffic, if reasonable values of load are considered or if the switch is equipped with suitable additional fibers. For this reason, the analytical models proposed in sections 2.3.3 and 3.3 can be used in evaluating performance of MS-SPW switch under Bernoulli and admissible traffic even if they do not take into account the grooming constraint of such architecture. The availability of such an accurate model can be very useful to obtain results when switch configuration leads to very low PLP, in ranges where simulation is hardly used to provide results.

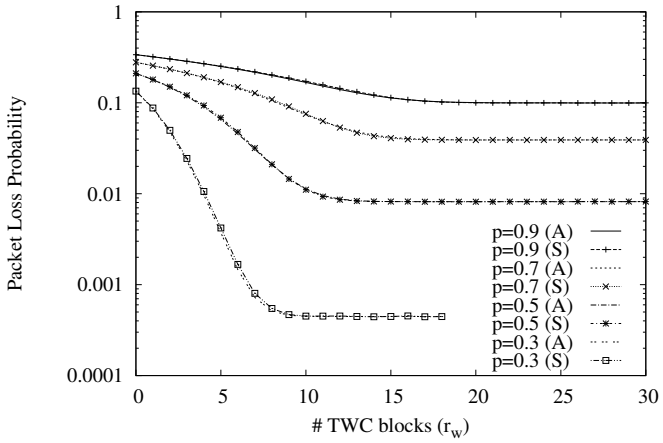


Figure 3.12: Comparison between analytical and simulation results for MS-SPW switch with Bernoulli traffic on input. PLP is shown as a function of the number of TWC blocks, varying load, in the case $N = 64$, $M = 8$.

Figures 3.12 and 3.13 show comparisons between simulation and analytical results for the MS-SPW switch, with Bernoulli input traffic. PLP as a function of the number of TWC blocks r_w is plotted, in the case $N = 64$, $M = 8$ and $N = 20$, respectively. PLP is evaluated as a function of the number of TWC blocks, varying load, in the case $N = 20$ and $M = 80$. In the first case, with a few ($M = 8$) wavelengths per fiber, the PLP is dominated by output blocking in all regions of the graphs, and analytical results show very good agreement with simulation. In the second case, with higher number of wavelengths ($M = 80$), the output blocking is lower and, with high load, analysis overestimates simulation in the region where loss due to the lack of TWCs is high with respect to that due to output blocking. This is due to the independence assumption on the number of conversion requests in different OFs made in the evaluation of loss due to the lack of TWCs P_{bwc}. Anyway, the difference is not too high and, given that the model leads to an overestimation, it can be used to dimension the switch as a worst case. In fact the asymptotic values of PLP are exactly evaluated by the model.

In figure 3.14 comparison between simulation and analysis for admissible traffic is presented, in the case $N = 64$, $M = 8$. In this case, also

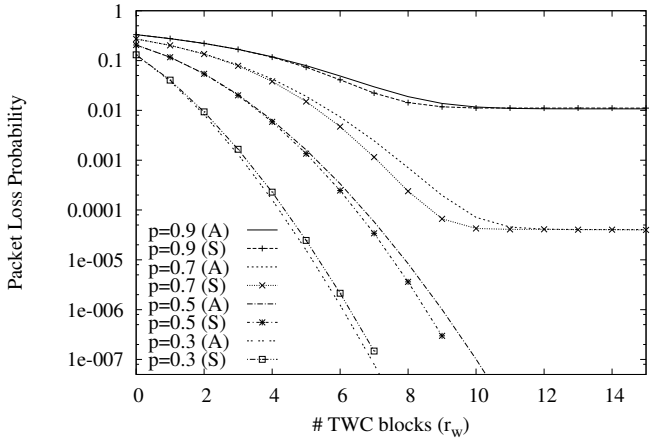


Figure 3.13: Comparison between analytical and simulation results for MS-SPW switch with Bernoulli traffic on input. PLP is shown as a function of the number of TWC blocks, varying load, in the case $N = 20$, $M = 80$.

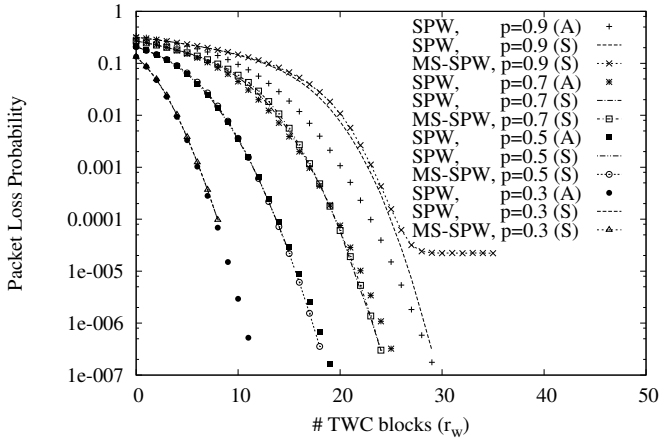


Figure 3.14: Comparison between analytical and simulation results for MS-SPW switch with admissible traffic on input. PLP is shown as a function of the number of TWC blocks, varying load, in the case $N = 64$, $M = 8$.

simulation results for SPW switch are plotted. Analytical results show good agreement only when traffic is far enough from 1. This is a consequence of correlation present in admissible traffic that is not accounted for in the model.

Chapter 4

Contention resolution in the wavelength and space domains: multi-fiber switches

In this chapter the SPN and SPW concepts are applied jointly to the multi-fiber concept in order to reduce the requirements of the TWCs employed inside the switch to solve contention. First, a brief presentation of the multi-fiber concept and its benefits in solving contention are presented. A schematic comparison between mono-fiber and multi-fiber switches is presented in figure 4.1. In the mono-fiber switch, depicted in figure 4.1a, each II/OI is equipped with one single optical fiber, as in the solutions previously presented in this work. Instead in the multi-fiber switch (figure 4.1b), the IIs/OIs are equipped with a bundle of F optical fibers. In this way, the same number W of input/output channels per interface can be obtained with a lower number M of wavelengths per fiber, and the same wavelengths can be repeated in the different fibers of the same interface. The F fibers provides spatial diversity, so that packets coming to the switch on the same wavelength can be forwarded in different fibers without wavelength conversion, assuring contention resolution in the space domain. For this reason, the number of conversion requests is lowered and the number of TWCs can be reduced with respect to the mono-fiber case. More, the tuning range of the TWCs is reduced, given that they are tuned over $M = \frac{W}{F}$ wavelengths instead of W as in the mono-fiber case.

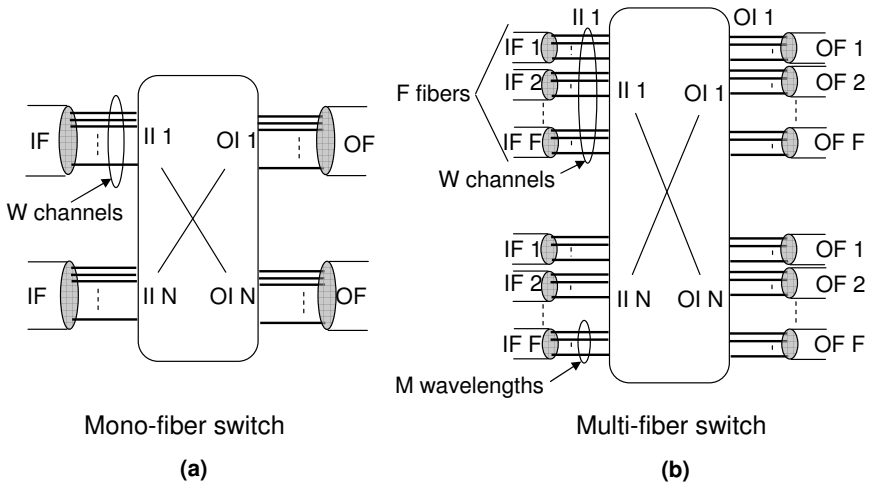


Figure 4.1: Comparison between mono-fiber (a) and multi-fiber (b) ideal switches.

Two multi-fiber switches are presented in this chapter, and the effectiveness of this solution is demonstrated by simulation and analytical results. The first section 4.1 presents the architecture (4.1.1) and the scheduling algorithm (4.1.2) for the multi-fiber SPN switch. Then section 4.2 shows the architecture (4.2.1), the scheduling algorithm (4.2.2) and the analytical model (4.2.3) for multi-fiber SPW switch. Finally section 4.3 shows a comparison between the SPN and SPW multi-fiber switches, in terms of performance (PLP) and complexity. In particular this work has been developed in collaboration with Prof. Vincenzo Eramo and Mr. Angelo Germoni, Infocom Department, University of Rome - Sapienza.

4.1 Multi-fiber SPN switch

Multi-fiber SPN switch (MF-SPN) comes from the SPN concept presented in section 2.2 jointly with the nulti-fiber concept. By exploiting the spatial diversity provided by the multi-fiber switch, the common pool of TWCs can be equipped with a lower number of TWCs. The attention is here focused on practical implementation of the MF-SPN, in particular

a solution based on WSs is here proposed as well as a proper scheduling algorithm.

4.1.1 Architecture for MF-SPN

The multi-fiber MF-SPN architecture is presented in figure 4.2, in the case $N = 2$ II/OIs with $F = 2$ fibers and $M = 2$ wavelengths each. In this architecture, r tunable-input/tunable-output TWCs are shared

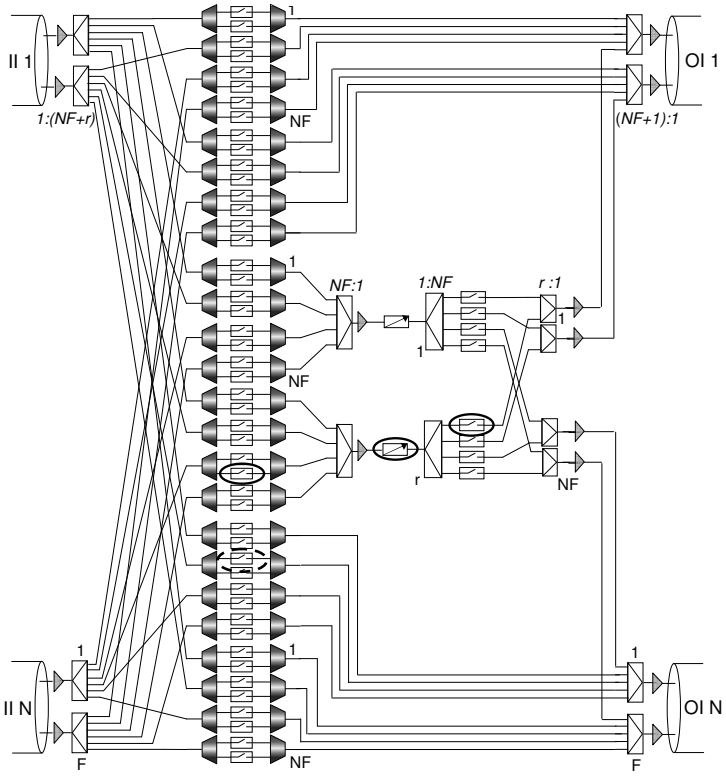


Figure 4.2: Example of practical architectural implementation of the multi-fiber shared-per-node (MF-SPN) concept. The architecture consists of $N = 2$ input/output interfaces (II/OIs) with $F = 2$ optical fibers, each one carrying $M = 2$ wavelengths. The architecture is equipped with $r = 2$ shared TWCs.

among all input channels (in figure 4.2, the switching node is equipped with $r = 2$ TWCs). Packets that do not need conversion can be directly

sent from the input fibers (IFs) to the destination output fibers (OFs). Each IF is connected to an OF by means of a WS, so that a packet can be directly forwarded from the IF to the destination OF by turning on the corresponding SOA.

For example, if the packet carried by the wavelength $k = 1$ in the IF $j = 2$ of the II $i = 1$ should be sent without wavelength conversion in the OF $j = 1$ of the OI $i = 2$, the first SOA in the seventh WS starting from the bottom must be turned on (this SOA is shown in the figure with a dotted circle), so letting the corresponding wavelength to pass through.

The number of WSs needed for an IF is NF, so that the total number of WSs needed to guarantee the direct connection from input to output fibers is $(NF)^2$ (these WSs can be seen at the top and bottom of the figure). This is the minimum number of WSs needed for the direct connection, independently of the sharing strategy adopted for the TWCs.

By observing figure 4.2 it is possible to see that the input WDM signal coming from an IF is split in $NF + r$ copies by means of a $1 : (NF + r)$ splitter. NF of these copies are connected to the NF WSs above mentioned, while other r copies allow the IF to reach the TWCs.

Now, the part of the architecture performing wavelength conversion (visible in the center of figure 4.2) is described. There are r WSs for each IF (for a total amount of NFr WSs) and r $NF : 1$ combiners, each one related to a TWC. Each WS which serves a given IF is connected to only one combiner to reach a particular TWC. By turning ON the *k-th* SOA in a particular WS, a signal carried by wavelength k in the IF connected to this WS is allowed to reach the related TWC. After wavelength conversion is accomplished, the signal from the output of a TWC is split by a $1 : NF$ splitter. The egresses of a splitter are connected to NF different $r : 1$ combiners by means of an array of SOAs. The needed number of SOAs is also here NFr. A $r : 1$ combiner is connected to a particular OF, and the SOAs allows to select which converted signals must reach this particular OF. Note that at each ingress of a combiner there are mono-wavelength signals, while at the output of a combiner there is a WDM signal that is sent to the OF. Finally, in each OF a $(NF + 1) : 1$ combiner is used to combine directly forwarded and wavelength converted signals. The $r : 1$ combiner allowing the converted signals to reach the proper OF and the final $(NF + 1) : 1$ combiner may be replaced with a single $(NF + r) : 1$ combiner, depending on

the maximum number of ingresses allowed by technological constraints. From the logical point of view, these two solutions are equivalent although physical performance can be different.

By considering the switching node presented in figure 4.2, suppose that a packet carried by wavelength $k = 2$ in the IF $j = 1$ of the II $i = 2$ must be converted to the wavelength $s = 1$ and sent to the OF $h = 1$ of the OI $g = 1$. In addition, the control unit chooses to use the second TWC. In this situation, the second SOA in the tenth WS starting from the bottom before the TWCs must be turned on to exploit the second TWC. Then, to forward the converted signal on the output of that TWC, the first SOA after the second splitter (related to the second TWC) must be turned on. All the SOAs and the TWC used are highlighted in figure 4.2 by solid circles.

The complexity of the MF-SPN architecture in terms of optical components is now given. The attention is focused to the evaluation of more expensive optical components, namely TWCs and SOAs. As mentioned above, $(NF)^2$ WSs are needed to directly connect IFs and OFs. Given that a WS is composed by M SOAs, the number of SOAs for the direct connection is $(NF)^2M$. To reach the TWCs, NFr WSs are needed, for a total amount of $NFMr$ SOAs. To allow the signals at the output of the TWCs to reach the proper OFs and avoid contention, additional NFr SOAs are employed. The total number of SOAs for the proposed MF-SPN implementation results in:

$$N_{SOA}^{MF-SPN} = (NF)^2M + NF(M+1)r \tag{4.1}$$

This number depends on the number of TWCs used.
The total number of TWCs is:

$$N_{TWC}^{MF-SPN} = r \tag{4.2}$$

and the number of TWCs r is related to the performance expected for the MF-SPN architecture (typically, r is equal to the minimum number of TWC allowing the same performance of an architecture equipped with the maximum number of TWCs, NMF, one TWC for each channel).

In addition, the MF-SPN architecture would require NF splitters $1 : (NF + r)$ (combiners $(NF + 1) : 1$) on input (output), r combiners $NF : 1$ (splitters $1 : NF$) before (after) the TWCs, NF combiners $r : 1$ to

group converted packets directed to the same OF. Also, $(NF)^2 + NFr$ MUX/DEMUX are needed to implement the WSs. $3NF + r$ EDFA are needed where the optical signal may be degradated (on input, given that the signal is coming after tens of kilometers, on output, given that the signal is forwarded in a long-haul fiber, and at the output of the combiners).

4.1.2 Scheduling algorithm for MF-SPN

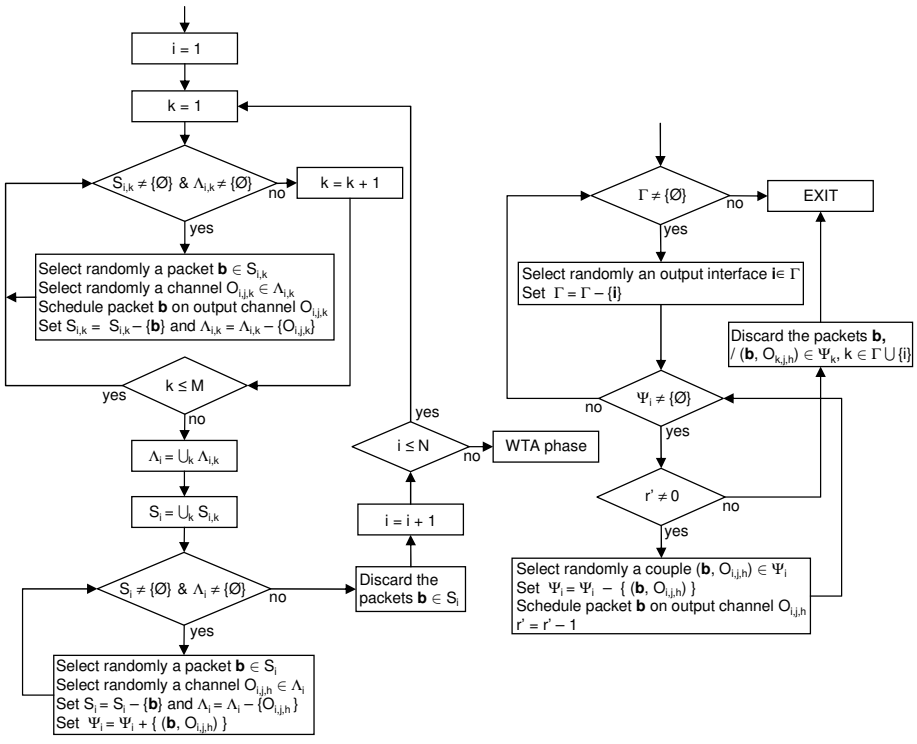
In the MF-SPN architecture, a pool of r shared TWCs is available to forward incoming packets by exploiting wavelength conversion.
A proper nomenclature for the description of the SA is first introduced:

- $S_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is the set containing the packets carried by wavelength k and directed to OI i;
- $O_{i,j,k}$ $(i = 1, \cdots, N)$, $(j = 1, \cdots, F)$, $(k = 1, \cdots, M)$ is the channel corresponding to the wavelength k in fiber j of OI i;
- $\Lambda_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is the set containing the free channels corresponding to the wavelength k in all fibers of OI i. The maximum cardinality of this set is F and it is initialized as containing the output channels $O_{i,j,k}$ for $j = 1, \cdots, F$;
- S_i $(i = 1, \cdots, N)$ is the set containing all packets directed to OI i and not forwarded without conversion;
- Λ_i is the set of free channels of the OI i after scheduling of packets not requiring conversion;
- Ψ_i $(i = 1, \cdots, N)$ is the set containing couples $(b, O_{i,j,h})$ where b is a packet directed to the OI i that must be converted and $O_{i,j,h}$ is the output channel assigned to that packet;
- Γ is the set of OIs not already served in the current time slot;
- r' is the number of available TWCs during the execution of the SA.

Each one of these sets and parameters is updated during the execution of the SA. In the INI phase each set $S_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is

firstly initialized as empty, then the input wavelength channels on each fiber of each input interface are scanned and the packets carried by wavelength k and directed to OI i are collected in the corresponding set $S_{i,k}$. In addition, in the INI phase, sets and variables used in the other phases are initialized.
Each set $\Lambda_{i,k}$ $(i = 1, \cdots, N)$, $(k = 1, \cdots, M)$ is re-initialized so that it contains the output channels $O_{i,j,k}$ with $j = 1, \cdots, F$, those corresponding to wavelength k in the F fibers of the OI i. The set Ψ_i $(i = 1, \cdots, N)$ is initialized as empty, the set Γ is initialized as containing all the OIs and r' is initialized as $r' = r$.

The OWCA phase is shown in figure 4.3(a). In this phase the packets forwarded without conversion are evaluated. The OIs are sequentially



(a) *OWCA phase* (b) *WTA phase*

Figure 4.3: Scheduling algorithm for MF-SPN architecture: a) OWCA phase b)WTA phase

considered and when the OI i is taken into account the wavelengths of this interface are sequentially scanned. When wavelength k of the OI i is considered, a packet b randomly selected from $S_{i,k}$ is scheduled for a channel $O_{i,j,k}$ ($j = 1, \cdots, F$) belonging to the $\Lambda_{i,k}$ randomly selected, as long as $S_{i,k}$ and $\Lambda_{i,k}$ are not empty. $S_{i,k}$ and $\Lambda_{i,k}$ are updated by removing the packet b and the channel $O_{i,j,k}$ respectively. Up to F packets randomly chosen from $S_{i,k}$ are scheduled and forwarded without wavelength conversion. In fact, packets on different sets $S_{i,k}$ are output wavelength blocking free, while packets belonging to the same set contend for the output channels of the same OI i and on the same wavelength k, those in the set $\Lambda_{i,k}$ (up to F). If $S_{i,k}$ or $\Lambda_{i,k}$ becomes empty, next wavelength is considered.

When all the M wavelengths of the OI i have been considered, remaining packets directed to the OI i are collected in the set S_i, obtained by the union of the sets $S_{i,k}$ for ($k = 1, \cdots, M$) in the second part of the OWCA phase. In addition, the free output channels on different wavelengths for the OI i are collected in the set Λ_i, obtained by the union of all sets $\Lambda_{i,k}$ for ($k = 1, \cdots, M$). After that, while both S_i is not empty and at least one output channel is free (Λ_i is not empty), a packet b belonging to S_i is randomly chosen as well as an output channel $O_{i,j,h}$ belonging to Λ_i. Then, the couple $(b, O_{i,j,h})$ is stored in the set Ψ_i. This set contains the packets directed to OI i that must be wavelength converted and the output channels on which this conversion will be accomplished if the number of available TWCs is high enough. If the set Λ_i is empty while S_i is not empty, these remaining packets are lost due to output wavelength blocking in the OI i.

Note that in this phase the TWCs are not needed, given that the packets are forwarded without conversion or collected in the sets Ψ_i ($i = 1, \cdots, N$). The sets Ψ_i will be used in the WTA phase to evaluate which packets may be forwarded and which are lost due to the lack of TWCs. The computational complexity of the OWCA phase is $O(NMF)$, given that in the worst case the number of operations needed for each OI is proportional to the number of the output channels in an OI, MF. The computational complexity is evaluated by considering the random operations being with the same weight (equal to one) of the other operations. This is not far from the reality, given that the random operations are only 'conceptual', and they can be replaced by round-robin operations,

allowing to obtain the same fairness, that can be implemented with the same complexity of a sum or decremental operations.

The WTA phase is illustrated in figure 4.3(b). Being the TWCs shared, they are randomly considered to assure the fairness among the OIs. When the OI i is taken into account, until the set Ψ_i is not empty, the SA looks for an available TWC. If at least one TWC is available, a couple $(b, O_{i,j,h})$ belonging to Ψ_i is randomly selected and the packet b is scheduled for the output channel $O_{i,j,h}$. Then the set Ψ_i is updated and the number of available TWC is decremented by one. If the set Ψ_i becomes empty, another OI is randomly selected, while if no TWC is available, the SA ends and the remaining packets are lost due to the lack of TWCs. The computational complexity of the WTA phase is $O(N + r)$ given that in the worst case all OIs are scanned and a number of operations proportional to the number of TWCs r is needed.

The computational complexity of the proposed SA is obtained by the sum of the complexity in the OWCA and WTA phases and results in $O(NMF + N + r) = O(NMF)$, given that $r \leq NMF$.

4.2 Multi-fiber SPW switch

In this section the multi-fiber SPW switch (MF-SPW) is presented. In particular the proposed architecture is presented in section 4.2.1, a proper scheduling algorithm in section 4.2.2 and an analytical model to evaluate PLP is proposed in section 4.2.3.

4.2.1 Architecture for MF-SPW

The practical implementation of the multi-fiber shared-per-wavelength concept is presented in figure 4.4, in the case $N = 2$ II/OIs with $F = 2$ fibers and $M = 2$ wavelengths each. In this example, the switching node is equipped with $r_w = 2$ TWCs for each wavelength. Also in this architecture, packets that do not need conversion are directly sent to the destination OFs, so the same number, $(NF)^2$, of WSs as in the MF-SPN architecture is needed to directly connect IFs and OFs.

Differently from the MF-SPN architecture the TWCs are grouped in r_w blocks (as can be seen in figure 4.4), each one including M TWCs dedicated to a different wavelength (to this end the TWCs in a block are

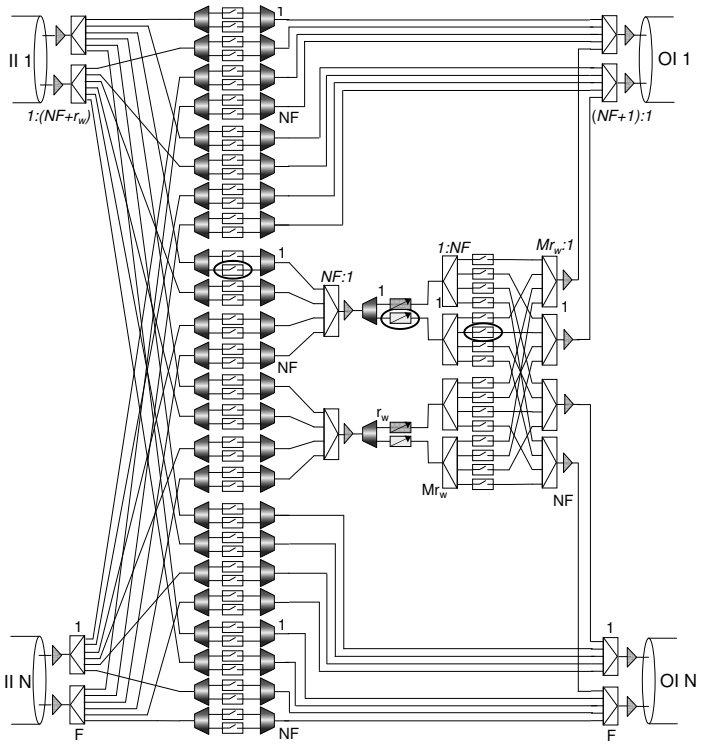


Figure 4.4: Example of practical architectural implementation of the multi-fiber shared-per-wavelength (MF-SPW) concept. The architecture consists of $N = 2$ input/output interfaces (II/OIs) with $F = 2$ optical fibers, each one carrying $M = 2$ wavelengths. The architecture is equipped with $r_w = 2$ shared TWCs for each wavelength. The total number of fixed-input/tunable-output FR-TWCs is Mr_w

preceded by a DEMUX). The IFs are connected to the r_w TWC blocks using WSs. Each WDM signal on input is split in $NF+r_w$ copies: the first NF are needed for the direct connection with the OFs (as in MF-SPN), the others r_w are interfaced to r_w $NF : 1$ combiners (each one related to a different TWC block) by means of WSs. So an IF is connected to all TWC blocks, and each TWC block may be reached from all IFs. Given that there are NF IFs, NFr_w WSs are needed to connect the IFs and the TWC blocks. The WSs are used to select the input channel which use a particular TWCs. Note that a wavelength k can exploit the *k-th* TWC in each TWC block, for a total amount of r_w TWCs. For example, the

wavelength 1 may exploit the first TWC in the first TWC block and the first TWC in the second block. The stage after the TWC blocks forwards wavelength-shifted packets to the destination OFs and is the same as in the MF-SPN.

By considering the switching node presented in figure 4.4, suppose that the packet carried by wavelength 2 in the IF $j = 1$ of the II $i = 1$ may be converted on the wavelength 1 in the OF $h = 2$ of the OI $g = 1$. In addition, the control unit chooses to use the second TWC dedicated to wavelength 2 in the first TWC block. The second SOA in the ninth WS starting from the top before the TWC blocks must be turned on as well as the second SOA of the second splitter after the TWC blocks. Both SOAs and the TWC used are highlighted in the figure by solid circles.

The complexity of this architecture is:

$$N_{SOA}^{MF-SPW} = (NF)^2 M + 2NFMr_w \tag{4.3}$$

Also in this case the number of SOAs depends on the number of TCWs used.
The total number of TWCs in the MF-SPW architecture is:

$$N_{TWC}^{MF-SPW} = Mr_w \tag{4.4}$$

In addition, the MF-SPW architecture requires NF splitters $1 : (NF + r_w)$ (combiners $(NF + 1) : 1$), Mr_w splitters $1 : NF$ after the TCWs, r_w combiners $NF : 1$ before the TWC blocks and NF combiners $Mr_w : 1$ to group converted packets directed to the same OF. The number of MUXs is $(NF)^2 + NFr_w$, equal to the number of WSs, while the number of DEMUXs is $(NF)^2 + (NF + 1)r_w$, given that an additional DEMUX is used on each TWC block. Also, $3NF + r_w$ EDFAs are employed.

4.2.2 Scheduling algorithm for MF-SPW

In the MF-SPW architecture, the TWCs are partitioned among the wavelengths, in the sense that each wavelength have a dedicated pool of r_w shared TWCs. The packets carried by a particular wavelength k (up to NF) can exploit r_w TWCs. For this reason a modified version of the SA proposed for MF-SPN is needed. The nomenclature used is almost the same as in the MF-SPN SA. The differences are:

- the set Ψ_i is replaced by M sets $\Psi_{i,k}$ $(k = 1, \cdots, M)$. A set $\Psi_{i,k}$ contains the packets directed to OI i and carried by the wavelength k that must be converted. Packets carried by different wavelengths exploit different TWCs, this is the reason why M different sets instead of one are considered;

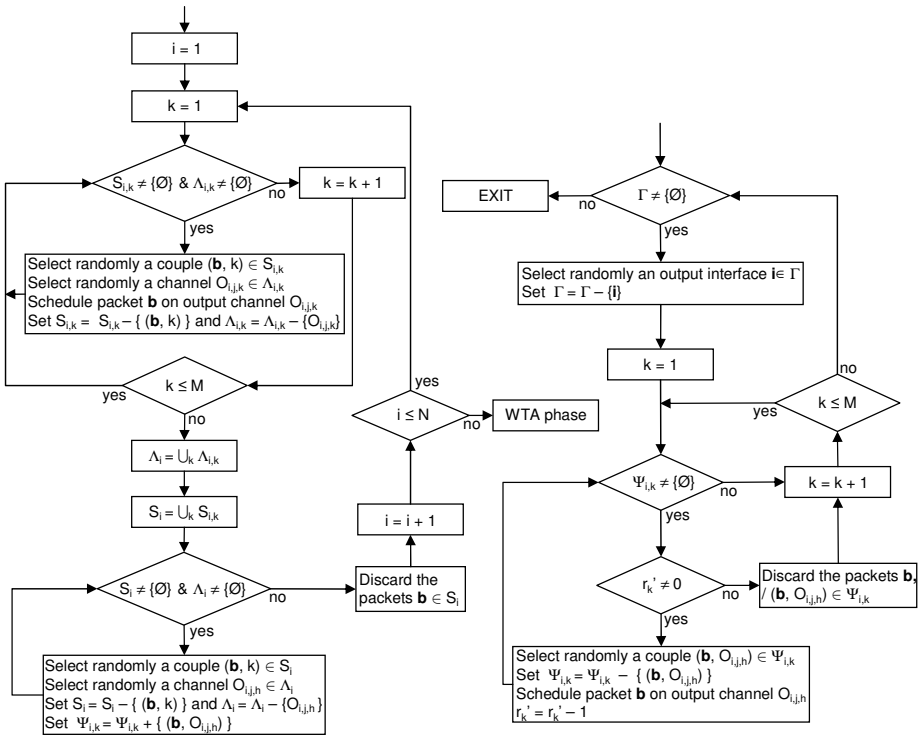
- the set $S_{i,k}$ does not contains only the packets carried by wavelength k and directed to OI i, but the couples (b, k) each one containing a packet and its encoding wavelength. This is due to the fact that the information related to the wavelength is necessary in the OWCA phase when the set $\Psi_{i,k}$ is updated;

- the number of available TWCs r' is replaced by M numbers $r_k^{'}$ $(k = 1, \cdots, M)$. $r_k^{'}$ indicates the number of TWCs dedicated to wavelength k available during the execution of the SA.

Having in mind these differences, in the INI phase the set $\Psi_{i,k}$ $(i = 1, \cdots, N)$ $(k = 1, \cdots, M)$ is initialized as empty and $r_k^{'}$ $(k = 1, \cdots, M)$ is initialized to $r_k^{'} = r_w$.

The OWCA phase is presented in figure 4.5(a). This phase is very similar to the OCWA phase in the MF-SPN SA. The only change is in the second part of the OWCA phase. In fact in this case when a packet b must be converted, it is not stored in the set Ψ_i as in the previous case, but it is stored in the set $\Psi_{i,k}$ together with the assigned output channel $O_{i,j,h}$. The wavelength the packet b is carried (k) is stored in the set $S_{i,k}$, given that this information is needed to insert the packet in the correct set $\Psi_{i,k}$. The computational complexity of the OWCA phase is also in this case $O(NMF)$.

The sets $\Psi_{i,k}$ are used in the WTA phase to evaluate which packets are forwarded by exploiting the TWCs. The WTA phase is shown in figure 4.5(b). As in the previous case the OIs are randomly considered for the fairness. When the OI i is taken into account, the wavelengths are sequentially evaluated. When the wavelength k is considered, until the set $\Psi_{i,k}$ is not empty, the algorithm looks for an available TWC dedicated to wavelength k ($r_k^{'} Neq 0$). If at least one available TWC is found, a packet randomly selected from $\Psi_{i,k}$ is scheduled for the corresponding output channel $O_{i,j,h}$, $\Psi_{i,k}$ is updated and $r_k^{'}$ is decremented by one. When $\Psi_{i,k}$ is empty or $r_k^{'} = 0$ the next wavelength is taken into account. The WTA



(a) *OWCA phase* (b) *WTA phase*

Figure 4.5: Scheduling algorithm for MF-SPW architecture: a) OWCA phase b)WTA phase

phase ends when all wavelengths of each OI have been evaluated. The packets that are not scheduled in the WTA phase are those lost due to lack of TWCs. The computational complexity of the WTA phase for the MF-SPW architecture is $O(N + NM + Mr_w) = O(NM + Mr_w)$, given that in the worst case all the N OIs are scanned, the NM sets $\Psi_{i,k}$ are scanned and a maximum number r_w of packets are sent for each of the M wavelengths.

The complexity of the SA, by taking into account the OWCA and WTA phases is, $O(NMF + NM + Mr_w) = O(NMF)$ given that $r_w < NF$.

4.2.3 Analytical model of packet loss for MF-SPW

The number of TWCs needed in the MF-SPN and MF-SPW switches is evaluated according to an appropriate dimensioning procedure that guarantees the same PLP as a switch equipped with a full set of TWCs. The analytical model to evaluate performance in MF-SPN switch is illustrated in [23]. In this paper an analytical model to evaluate performance in MF-SPW switch is presented. The model can be used to determine the minimum number of TWCs needed in MF-SPW. In particular the PLP of an MF-SPW switch is evaluated as a function of the number r_w of TWCs per wavelength, under the following assumptions:

- the operation mode of the switching node is synchronous [23] on a time-slot basis;
- packet arrivals at the NFM input wavelength channels are independent;
- the input traffic is symmetric, that is packet arrivals occur with probability p for all IFs. No assumption is made for the mutual dependence of packet arrivals at different time-slots, since, due to the buffer-less nature of the switch, performance and TWC dimensioning depend on p only;
- the output traffic is symmetric, that is a packet has the same probability $1/N$ to be directed to any given OI. The symmetric traffic scenario is assumed because it is the one requiring the most severe dimensioning of TWCs, as shown in [36]. In fact as packet loss is due to either lack of output wavelength channels or lack of TWCs, and TWC dimensioning aims at making the latter negligible with respect to the former, the lower the loss probability due to lack of output channels the higher the number of TWCs needed: hence the assumption of symmetric traffic, which allows for the lowest loss probability due to lack of output channels.

The analytical model here presented aims at evaluating an upper bound of the packet loss probability ($P_{l,MF-SPW}$) of an MF-SPW switch versus the number r_w of TWCs shared per each wavelength. The tightness of the obtained upper bound is demonstrated by means of simulations.

Finally, the values of $P_{l,MF-SPW}$ are compared with the packet loss probability ($P_{l,MF-SPN}$) of a MF-SPN architecture evaluated by means of the analytical model discussed in [23].

The probability $P_{l,MF-SPW}$ can be expressed according to the following expression:

$$P_{l,MF-SPW} = \frac{E[N_l]}{E[N_o]} \tag{4.5}$$

wherein:

- $E[N_o] = NMFp$ is the average number of packets entering the switch in a given time-slot;

- $E[N_l]$ is the average number of packets that are lost due to contentions inside the switch in a given time-slot.

A packet loss can be caused by two types of events:

- *output wavelength blocking*, i.e. the unavailability of a free wavelength on the OI towards the packet is directed;

- *converter blocking*, i.e. the unavailability of a TWC able to shift an incoming packet on a free wavelength.

So, $E[N_l]$ can be expressed as follows:

$$E[N_l] = E[N_{wl}] + E[N_{cl}] \tag{4.6}$$

wherein $E[N_{wl}]$ and $E[N_{cl}]$ are the average number of packets lost in a time slot because of the *output wavelength blocking* and the *converter blocking*, respectively. In evaluating $E[N_l]$ we take into account that the control algorithm proposed in section 4.2.2 firstly selects the packets to be dropped due to *output wavelength blocking* and, then, the packets to be lost due to *converter blocking*.

Due to the symmetric traffic assumption, the loss term $E[N_{wl}]$ is N times $E[N_{wl,OI}]$, that is the average number of packets directed to any OI and lost because of the *output wavelength blocking*. By considering that, if j packets are offered to an OI, $max(0, j - MF)$ out of them are

lost and taking into account that the number of packets directed to any OI has a *binomial*(NMF, p) distribution, we can write:

$$E[N_{wl}] = NE[N_{wl,OI}] =$$

$$= N \sum_{j=MF+1}^{NMF} (j - MF) \binom{NMF}{j} \left(\frac{p}{N}\right)^j \left(1 - \frac{p}{N}\right)^{NMF-j} \tag{4.7}$$

The loss term $E[N_{cl}]$ in (4.6) takes into account the event that, even if there are free wavelengths on an OI, some packets can be lost due to the lack of available TWCs to perform wavelength conversion. In the following we evaluate an upper bound $E[N_{cl,ub}]$ of the term $E[N_{cl}]$, by assuming that no *output wavelength blocking* takes place in the switch, i.e. $M = \infty$ on the OIs. As it is easy to understand, this condition requires a number of wavelength conversions greater than in the real scenario. By remembering that the input traffic is symmetrical and the choice of the packets to be shifted is purely random, we can evaluate the term $E[N_{cl,ub}]$ as M times the corresponding term for the packets arriving on a generic wavelength λ. Let W_λ be the random variable representing the number of wavelength conversions needed in a time slot for the packets arriving on a generic wavelength λ; if each wavelength is provided with r_w TWCs and $W_\lambda = h$, $max(0, h - r_w)$ packets are lost due to converter blocking for the wavelength λ considered, so

$$E\left[N_{cl,ub}\right] = M \sum_{h=r_w+1}^{NF} (h - r_w)\, p_{W_\lambda}(h) \tag{4.8}$$

where $p_{W_\lambda}(h = 0, \cdots, NF)$ denotes the W_λ's probabilities.

Exact evaluation the W_λ's probabilities is straightforward only when $F = 1$, that is in the case mono-fiber. In the case $F > 1$, the evaluation is performed by assuming statistical independence among the number of conversions required by the packets arriving on wavelength λ and directed to the different OIs. This assumption does not hold in practice but, being the dependence slight, as shown in [36], it leads to estimate a greater number of conversions with respect to the real case. This is due to the negative correlation of the number of conversion needed on the different OIs. The W_λ's probabilities will be now evaluated in the cases

$F = 1$ and $F > 1$.

Evaluation of W_λ's probabilities in the case $F = 1$

When the packet loss due to the lack of output wavelengths is neglected, W_λ can be expressed as follows:

$$W_\lambda = R_\lambda - (N - N_\lambda), \tag{4.9}$$

wherein:

- R_λ is the *Binomial*(N, p) random variable denoting the number of packets arriving on the wavelength λ.
- N_λ is the random variable denoting the number of OIs to which no packet, arriving at the wavelength λ, is directed.

By applying the total probability law and by conditioning to $R_\lambda = i$, we have:

$$p_{W_\lambda}(h) = \sum_{i=0}^{N} Pr\left(N_\lambda = h + N - i / R_\lambda = i\right) \binom{N}{i} p^i (1-p)^{N-i} \tag{4.10}$$

The N_λ's probabilities conditioned to $R_\lambda = i$ can be evaluated by interpreting our problem in terms of the classical occupancy problem [45]; that is we throw $R_\lambda = i$ balls at random into N urns and wish to know the distribution of the number of not occupied urns N_λ. The classical occupancy problem has been solved in literature and in particular in [45] the following formula is reported for the searched probabilities:

$$Pr\left(N_\lambda = h / R_\lambda = i\right) = \binom{N}{h} \sum_{Nu=0}^{N-h} (-1)^N u \binom{N-h}{Nu} \left(1 - \frac{h + Nu}{N}\right)^i \tag{4.11}$$

Finally by inserting (4.11) in (4.10) and (4.10) in (4.8) we can evaluate the upper bound $E[N_{cl,ub}]$ of the average number $E[N_{cl}]$ of packet lost because of the lack of converters in the case of mono-fiber MF-SPW switch ($F = 1$).

This model with $F = 1$ evaluates the PLP of the mono-fiber SPW, so it represents an alternative analysis with respect to the one proposed

in section 2.3.3. Note that the model here proposed in more precise than the one in 2.3.3, given that in this case the correlation among conversion requests for different OFs is not neglected, but exactly evaluated. For this reason this model is useful also to evaluate the performance of the multi-stage architecture (practical implementation of the SPW) proposed in chapter 3.

Evaluation of W_λ's probabilities in the case $F > 1$

By assuming statistical independence of the number of the conversions required by packets arriving on a generic wavelength and directed to the various OIs, we can write:

$$p_{W_\lambda}(h) = \overbrace{p_{W_{\lambda,OI}} \otimes \cdots\cdots \otimes p_{W_{\lambda,OI}}}^{N \ times} \tag{4.12}$$

where the symbol $\otimes$ denotes the convolution operator and $p_{W_{\lambda,OI}}(h)$ denotes the probability that h conversions are needed for the packets arriving on the generic wavelength λ and directed to a generic OI. Let $R_{\lambda,OI}$ be the $binomial(NF, \frac{p}{N})$ random variable denoting the number of packets arriving on wavelength λ and directed to a generic OI. When the loss due to the lack of output wavelengths is neglected, the number of conversions required by the packets arriving on wavelength λ equals the number of packets in excess with respect to the F output wavelength channels which allow the packets to be forwarded without wavelength conversion. Hence we can write:

$$p_{W_{\lambda,OI}}(h) = \begin{cases} \sum_{i=0}^{F} Pr\left(R_{\lambda,OI} = i\right) & h = 0 \\ Pr\left(R_{\lambda,OI} = F + h\right) & h = 1, \cdots, F\left(N-1\right) \\ 0 & otherwise \end{cases} \tag{4.13}$$

Finally by inserting (4.13) in (4.12) and (4.12) in (4.8) we can evaluate the upper bound $E[N_{cl,ub}]$ of the average number $E[N_{cl}]$ of packet lost because of the lack of converters in the case of multi-fiber MF-SPW switch ($F > 1$).

4.3 Numerical results for MF-SPN and MF-SPW

In this section the switching architectures illustrated in sections 4.1.1 and 4.2.1 for MF-SPN and MF-SPW respectively are numerically compared. In particular the accuracy of the analytical model introduced in section 4.2.3 to evaluate the PLP of MF-SPW is investigated in 4.3.1 by comparing analytical and simulation results. These results are obtained by applying the scheduling algorithm presented in section 4.2.2. All simulation results have been obtained with confidence interval at 95% less than or equal to 5% of the mean. Then in section 4.3.2 MF-SPN and MF-SPW switches are compared in terms of performance (PLP) and number of TWCs and SOA gates needed in order to guarantee a given value of PLP.

4.3.1 Model validation

Figure 4.6 shows the PLP as a function of the total number (Mr_w) of TWCs employed for the MF-SPW. Different configurations in terms of fibers/wavelengths per fiber are plotted by maintaining the same number ($FM = 16$) of channels per II/OI, in case $N = 16$, $p = 0.3, 0.7$. The figure shows very good agreement in the mono-fiber case ($F = 1$, $M = 16$) for both values of load, while there is a slight difference among simulation and analytical results in the multi-fiber case especially when load is high. As stated before analytical results tends to overestimate the simulated ones, anyway with this switch dimensioning there is good agreement between simulation and analysis.

Figure 4.7 plots the PLP as a function of the total number of TWCs in case $N = 10$, $p = 0.4$ and $FM = 24$ channels per II/OI. Again, the model for the mono-fiber case shows perfect agreement with simulation, while the model for the multi-fiber case overestimates simulation when the number of fibers F is low ($F = 2$).

In figure 4.8 the PLP is plotted in case $N = 20$, $p = 0.7$ and $FM = 100$ channels per II/OI. For the mono-fiber case, analysis and simulation perfectly agree, while for the multi-fiber case, analysis highly overestimates simulation. This is related to the correlation among the conversion requests on different OIs (relevant in this case, high load with

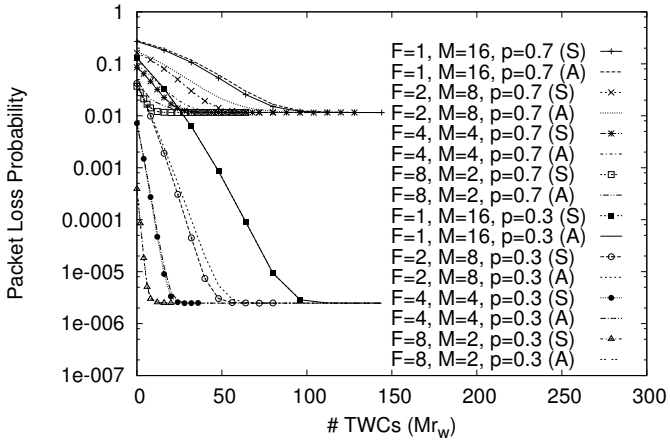


Figure 4.6: PLP of MF-SPW as a function of the number of TWCs, Mr_w, in case $N = 16$, $p = 0.3, 0.7$ and $F = 1, M = 16$, $F = 2, M = 8$, $F = 4, M = 4$ and $F = 8, M = 2$. (A) is for Analysis and (S) for simulation.

low loss due to output wavelength blocking), that is neglected in the multi-fiber model. Figures 4.6-4.8 shows that MF-SPW architecture obtains the same performance as fully equipped architectures (NFM TWCs) with a limited number of TWCs for all values of F, M. More, as F increases and M decreases, the saving in number of TWCs becomes more and more relevant.
By observing the presented results, some remarks about mono-fiber and multi-fiber SPW architectures are here given. SPW is a sharing strategy allowing to save in the number of TWCs employed in a node. The TWC saving is related to the *sharing degree* of the TWCs. The higher the sharing degree, the higher the TWC saving. In the SPW architectures the sharing degree is high when the number of packets carried by the same wavelength (those that share the TWCs) is high.

In a mono-fiber switch this happens when the number of input/output fibers N is high, so higher the number of fibers, better the SPW performs (same PLP as fully equipped architecture with lower number of TWCs). For multi-fiber architecture, the sharing degree depends on the number of fibers NF, so the architecture allows high TWC saving when this number is high (high number of II/OIs N or high number of fibers per

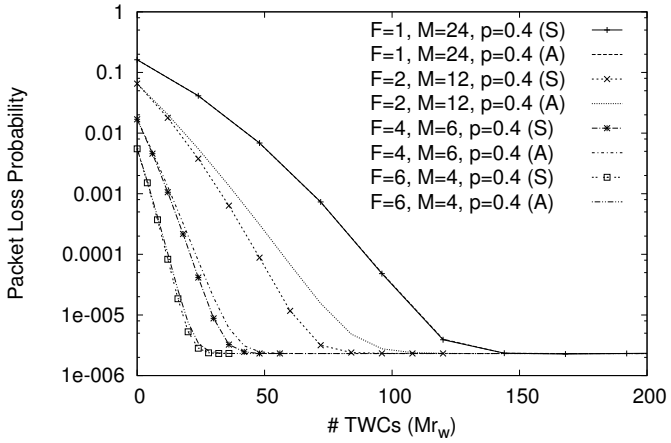


Figure 4.7: PLP of MF-SPW as a function of the number of TWCs, Mr_w, in case $N = 10$, $p = 0.4$ and $F = 1, M = 24$, $F = 2, M = 12$, $F = 4, M = 6$ and $F = 6, M = 4$. (A) is for Analysis and (S) for simulation.

interface F).

As the proposed models, the one for the mono-fiber case takes into account the correlation among the conversion requests. It is very precise under all scenarios, especially when the upper bound introduced is negligible (low wavelength blocking probability), so when the number of wavelengths per fiber M is high. Instead, the model for MF-SPW architecture does not take into account the correlation among the conversion requests, so it provides a good approximation of the PLP when the load is not too high. With high load and low output wavelength blocking probability (high MF), as in figure 4.8, the effect of the correlation is very relevant, so the overestimation provided by the model leads to results far from simulation.

4.3.2 Comparison of MF-SPN and MF-SPW

The MF-SPN and MF-SPW switch performance is compared in figures 4.9-4.10 in case $N = 16$, by fixing $MF = 32$ and $MF = 256$ respectively. PLP is reported as a function of the total number of TWCs for different switch configurations in terms of fibers/wavelengths (F, M). Results have

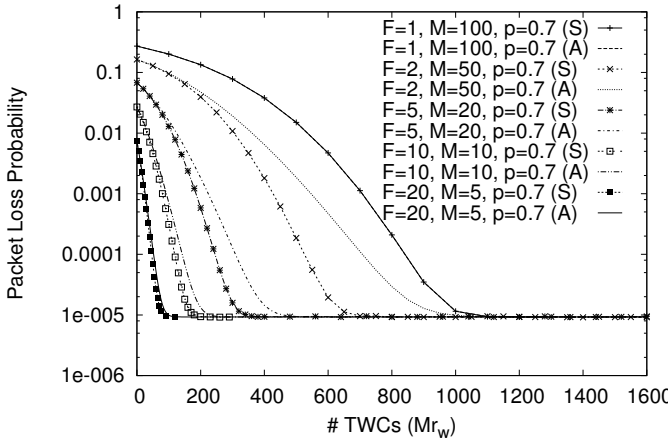


Figure 4.8: PLP of MF-SPW as a function of the number of TWCs, Mr_w, in case $N = 20$, $p = 0.7$ and $F = 1, M = 100$, $F = 2, M = 50$, $F = 5, M = 20$, $F = 10, M = 10$ and $F = 20, M = 5$. (A) is for Analysis and (S) for simulation.

been obtained by applying the analytical models proposed in [23] and section 4.2.3 for MF-SPN and MF-SPW switch respectively. In figure 4.9 the performance is evaluated in a low load scenario $p = 0.4$ and F ranging from 2 to 32. In figure 4.10 the performance is evaluated in a high load scenario $p = 0.8$ and F ranging from 2 to 128. All of the sketched curves show the same trend, the PLP decreases as the number of converters increases up to a threshold value beyond it the effect of the converter blocking is negligible and the PLP saturates. The saturation value for the PLP denotes the loss due to wavelength blocking and it represents the same PLP of a fully equipped multi-fiber switch, next referred to as Multi-Fiber Single-Per-Channel (MF-SPC) [15] switch. Let $r_{th,MF-SPN}$ and $r_{th,MF-SPW}$ be the threshold values for the MF-SPN and MF-SPW switch respectively. They represent the dimensioning values for the TWCs of the MF-SPN and MF-SPW switches, that is the minimum number of TWCs needed to the switches to reach the saturation PLP.

The TWCs dimensioning is more severe for MF-SPW switch with respect to MF-SPN switch (see figures 4.9-4.10), for example in figure 4.10 for the case $(F, M) = (2, 128)$ MF-SPW needs $r_{th,MF-SPW} = 2944$

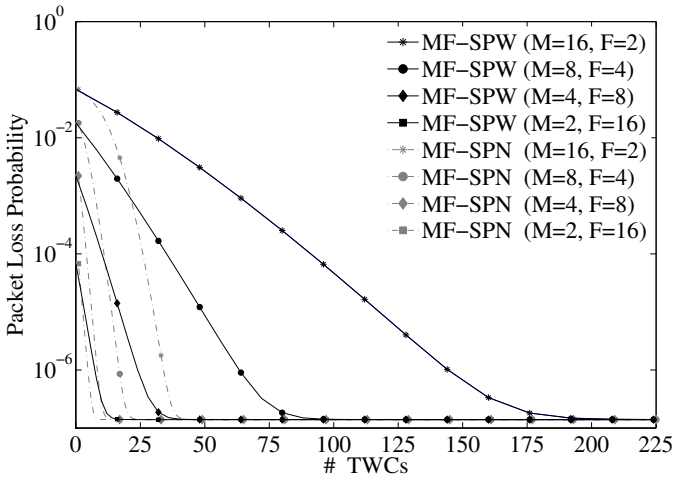


Figure 4.9: Packet loss probability of MF-SPN and MF-SPW switches as a function of the total number of TWCs in case $N = 16$, $p = 0.4$ and F ranging from 2 to 16 and M ranging from 16 to 2 accordingly.

TWCs, while $r_{th,MF-SPN} = 768$ only. This is due to the intrinsic characteristics of the MF-SPW switch that does not allow the perfect sharing of the available TWCs. In fact, the TWCs in the MF-SPW are fixed-input/tunable-output, so they are partitioned among the wavelengths. Instead, the MF-SPN switch is equipped with tunable-input/tunable-output TWCs, so each incoming packet can exploit whatever TWC (perfect sharing).

In the third and fifth row of tables 4.1-4.2 the number of TWCs needed for different switch configurations is presented, in the same cases as in figures 4.10-4.10, respectively. For both MF-SPN and MF-SPW switches, the TWC dimensioning is less severe as the number of fibers F used in each II/OI increases. This is as expected, as the higher F is, the higher the probability that a contending packet can be forwarded on the same wavelength it is arriving on. In the fourth and sixth row of tables 4.1-4.2 the TWCs percentage savings $g_{MF-SPN} = 100\ (1 - \frac{r_{th,MF-SPN}}{NMF})$ and $g_{MF-SPW} = 100\ (1 - \frac{r_{th,MF-SPW}}{NMF})$ of the MF-SPN and MF-SPW switches with respect to the MF-SPC switch (using one TWC for each input wavelength channel), is reported. Note that as F increases, the difference between the TWCs percentage savings g_{MF-SPN} and g_{MF-SPW} becomes

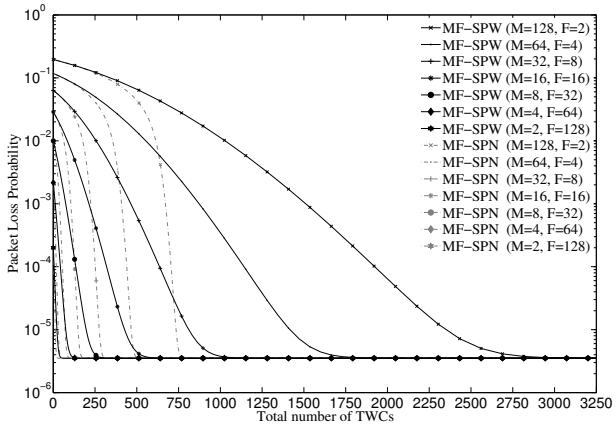


Figure 4.10: Packet loss probability of MF-SPN and MF-SPW switches as a function of the total number of TWCs, in case $N = 16$, $p = 0.8$ and F ranging from 2 to 128 and M ranging from 128 to 2 accordingly.

little. In fact, the SPW technique becomes more and more effective when, by increasing F, a higher number of input wavelength channels shares the r_w TWCs which convert packets arriving on the same input wavelength.

	(F, M)			
	(2, 16)	(4, 8)	(8, 4)	(16, 2)
$r_{th,MF-SPN}$	44	22	12	8
g_{MF-SPN}	91.41%	95.70%	97.66%	98.44%
$r_{th,MF-SPW}$	208	96	40	16
g_{MF-SPW}	59.38%	81.25%	92.19%	96.88%

Table 4.1: Values of $r_{th,MF-SPN}$, g_{MF-SPN}, $r_{th,MF-SPW}$ e g_{MF-SPW}, in MF-SPN and MF-SPW switches for $p = 0.4$, $N = 16$, F ranging from 2 to 16 and M ranging from 16 to 2 accordingly.

The MF-SPN and MF-SPW are then compared by evaluating the number of SOA gates needed to implement the switching fabric, by means of (4.1) and (4.3). In the third and fifth row of tables 4.3-4.4 the number of SOA gates needed when the MF-SPN and MF-SPW switches

	(F, M)					
	(2, 128)	(4, 64)	(8, 32)	(32, 8)	(64, 4)	(128, 2)
$r_{th,MF-SPN}$	768	501	302	90	43	22
g_{MF-SPN}	81.26%	87.77%	92.63%	97.80%	98.95%	99.46%
$r_{th,MF-SPW}$	2944	1792	1056	280	120	48
g_{MF-SPW}	28.16%	56.27%	74.23%	93.17%	97.07%	98.83%

Table 4.2: Values of $r_{th,MF-SPN}$, g_{MF-SPN}, $r_{th,MF-SPW}$ e g_{MF-SPW}, in MF-SPN and MF-SPW switches for $p = 0.8$, $N = 16$, F ranging from 2 to 128 and M ranging from 128 to 2 accordingly.

are equipped with the number of TWCs reported in tables 4.1-4.2, respectively, are reported. To quantify the switching fabric complexity increase, in the fourth and sixth rows of tables 4.3-4.4 the percentage increases α_{MF-SPN} and α_{MF-SPW} of the MF-SPN and MF-SPW switches with respect to MF-SPC switch are reported. Because the num-

	(F, M)			
	(2, 16)	(4, 8)	(8, 4)	(16, 2)
N_{SOA}^{MF-SPN}	$40.3 \cdot 10^3$	$45.4 \cdot 10^3$	$73.2 \cdot 10^3$	$137.2 \cdot 10^3$
α_{MF-SPN}	146.09%	38.67%	11.72%	4.69%
N_{SOA}^{MF-SPW}	$29.7 \cdot 10^3$	$45.1 \cdot 10^3$	$75.8 \cdot 10^3$	$139.3 \cdot 10^3$
α_{MF-SPW}	81.25%	37.50%	15.63%	6.25%

Table 4.3: Number of SOAs gates and percentage complexity increases α_{MF-SPN}, α_{MF-SPW} in MF-SPN and MF-SPW switches for $p = 0.4$, $N = 16$, F ranging from 2 to 16 and M ranging from 16 to 2 accordingly. The MF-SPN and MF-SPW switches are equipped with the number of TWCs reported in table 4.1.

ber of SOA gates needed in MF-SPC switch can be easily shown to be $(NF)^2M$, by using (4.1) and (4.3) the expressions $\alpha_{MF-SPN} = 100(M + 1)r_{th,MF-SPN}/(NMF)$ and $\alpha_{MF-SPW} = 200r_{th,MF-SPW}/(NMF)$ hold. From the results reported in tables 4.3-4.4, MF-SPW switch results to need, in most cases, a smaller number of SOAs gates than MF-SPN switch. As mentioned in section 4.2.1, this is due to the higher com-

plexity of the switching module in MF-SPN to allow incoming packets to reach the shared TWC pool. In MF-SPN NF WSs allow only one

	(F, M)					
	(2, 128)	(4, 64)	(8, 32)	(32, 8)	(64, 4)	(128, 2)
$\#SOA_{MF-SPN}$	$3.30 \cdot 10^6$	$2.35 \cdot 10^6$	$1.80 \cdot 10^6$	$2.51 \cdot 10^6$	$4.41 \cdot 10^6$	$8.39 \cdot 10^6$
α_{MF-SPN}	2418.75%	795.04%	243.31%	19.78%	5.25%	1.61%
$\#SOA_{MF-SPW}$	$0.32 \cdot 10^6$	$0.49 \cdot 10^6$	$0.79 \cdot 10^6$	$2.39 \cdot 10^6$	$4.44 \cdot 10^6$	$8.58 \cdot 10^6$
α_{MF-SPW}	143.75%	87.50%	51.56%	13.67%	5.86%	2.34%

Table 4.4: Number of SOAs gates and percentage complexity increases α_{MF-SPN}, α_{MF-SPW} in MF-SPN and MF-SPW switches for $p = 0.8$, $N = 16$, F ranging from 2 to 128 and M ranging from 128 to 2 accordingly. The MF-SPN and MF-SPW switches are equipped with the number of TWCs reported in table 4.2.

TWC to be reached. On the contrary, in MF-SPW, NF WSs allow one block of M TWCs to be reached. As can be seen from tables 4.3-4.4, when F increases and, consequently, M decreases, the MF-SPW switching fabric complexity becomes slightly higher than the one in MF-SPN switch. This is due to the following reasons: i) each TWC block contains a smaller number of TWCs in MF-SPW and the input switching modules allowing the packets to reach the TWCs tend to have the same complexity in MF-SPN and MF-SPW switches; ii) being the number of TWCs needed in MF-SPW greater than the one in MF-SPN, the output switching module, which allows converted packets to reach the OIs, has a higher complexity in MF-SPW switch with respect to MF-SPN switch. From the complexity analysis we have carried out, we can conclude that in an MF-SPW switch the number F of fibers per II/OI and, consequently, the number M wavelengths carried in each fiber must be chosen high enough to allow a good TWC saving but not too high to avoid more complex switching fabric with respect to MF-SPN.

By observing the third and fifth rows of tables 4.3, 4.4 it is possible to see that the number of SOA gates employed increases as the number of fibers per II/OI F increases. In fact, as shown in section 4.1.1, the number of SOA gates needed for the direct connection is $(NF)^2M$ (the same as for MF-SPC), so even if the total number of channels per II/OI MF is fixed, the number of SOA increases linearly with F.

From Tables 4.1-4.4 we can note as MF-SPW represents a compet-

itive architecture with respect to MF-SPN, especially when few fibers and many wavelengths are used. The currents trends in optical networks, where DWDM technology enables to use more wavelengths in one fiber, foresees that this will be the scenario expected in the future. In this case MF-SPW has the advantage to require a smaller number of SOAs with respect to MF-SPN, for example when p=0.8, N=16, F=4 and M=64, the number of SOAs is increased of 795% and 87% in MF-SPN and MF-SPW respectively. On the contrary MF-SPW allows a smaller number of TWCs to be saved. In the case study above mentioned, the saving of TWCs is 88% and 56% for MF-SPN and MF-SPW respectively. The total number of TWCs needed is higher in MF-SPW, but fixed-input/tunable-output TWCs can be employed instead of tunable-input/tunable-output. It is not easy to evaluate how much the employment of fixed-input/tunable-output can impact on the cost, but it is reasonable to think that these converters are easier to be realized and much less costly than the tunable-input/tunable-output one, so the MF-SPW architecture is a strong candidate to implement a switching node with shared TWCs in a cost-effective way.

Chapter 5

Contention resolution in the wavelength and time domains

In this chapter switching architectures relying on both wavelength and time domains to solve contention are presented. Contention resolution in wavelength domain is again performed by applying wavelength conversion, while two different alternatives are considered in the time domain. The first one is to maintain the packets in optical domain and exploit Fiber Delay Lines (FDLs) as optical buffers, while the second one is to convert the packets losing contention in electronic domain and store them in electronic buffers (RAMs) until an available time slot for forwarding is found.

In the former case, only discrete delays are admitted for a packet, and it can be only buffered for a given amount of time. An example of discrete delays provided by FDLs in synchronous time slotted environment is presented in figure 5.1. Different FDLs are available to delay a packet arriving at the instant time t_0. Each FDL provides a constant delay, associated to its length. The delays provided by FDLs are proportional to a constant value, called granularity and indicated with D, here equal to the time slot duration. In this way, when Q FDLs are available, a packet arriving at the instant t_0 can be delayed up to $Q-1$ time slots. In particular if the packet is broadcast to all FDLs, it is available for transmission not only in the current time slot but also for the next $Q-1$ time slots. If the packet is not forwarded during this period, it is discarded.

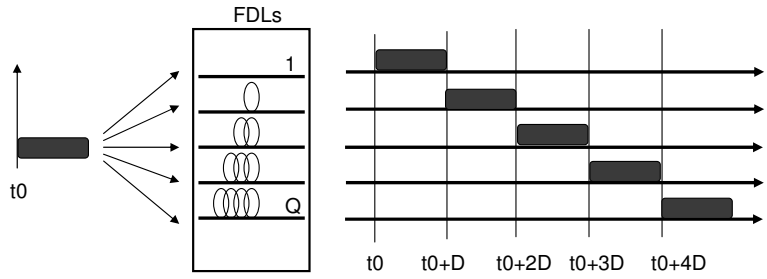


Figure 5.1: Example of discrete delays provided by Fiber Delay Lines. The packet is available for transmission in the $L-1$ time slots next the one it arrived.

In the latter case conversion to electronic domain is needed to avoid loss of data due to contention, so the resulting architecture is called hybrid electro/optical switch. In this kind of switches packets can be stored in the buffer for an arbitrary amount of time, and they can be extracted from the buffer in whatever instant. Loss occurs when no more rooms are available on the buffer to store packets losing contention.

An example of switching node relying on FDLs to solve contention is presented in section 5.1, while an example of hybrid switch with electronic buffers is presented in section 5.2.

5.1 All-optical input buffered switch

This section presents an input buffered switch architecture, shown in figure 5.2 in case $N = 3$ IF/OFs each carrying $M = 4$ wavelengths. The node consists of the same MS-SPW switch presented in chapter 3, with an input buffering stage (which confirms the modularity of this architecture) [44]. The resulitng switch is called IB-SPW (Input-Buffered multi-stage Shared-Per-Wavelength). Here the case with FDLs as input buffers is presented, but such input buffered switch can also be equipped with electronic buffers to solve contention, obtaining an hybrid switch. The input buffering stage is implemented by the same optical devices of the MS-SPW architecture, as described in section 3.1.

The principle of operation is the following: at each node input, af-

ter optical amplification by means of an EDFA (Erbium-Doped Fiber Amplifier), a power coupler is used to generate multiple copies of the multi-wavelength bundle of channels entering the node from this input. The power coupler should have $Q+1$ outlets where one outlet per incoming fiber is reserved for a local drop, while Q outlets forward wavelength channels to the Q FDLs that implement the input buffer. Optical packets arriving at the switch are broadcast to all the available FDLs, so they are available at the first S-stage with all possible delays between 0 and $L = Q - 1$ time slots, where L is called buffer size. Different queuing disciplines can be applied to manage queues and select the packets for forwarding. Each FDL is followed by a WS, which allows to transmit/block a packet in a particular wavelength. At the output side, the WSs are interconnected by means of a $Q : 1$ power combiner to feed the MS-SPW switch with the selected packets. Then the interconnection network behavior is the same as described in chapter 3.

The number of SOAs needed for the MS-SPW switch has been evaluated in formula (3.4). For the IB-SPW switch, additional $Q = L+1$ WSs per IF are needed, for a total amount of $(L + 1)N$ WSs with M SOAs each. By using (3.4) and taking into account the additional components for the buffering stage, the total complexity in terms of SOAs for the IB-SPW switch results in:

$$N_{SOA}^{IB-SPW} = 2N^2M + (L + 1)NM = NM(2N + L + 1) \qquad (5.1)$$

When contention arises, the scheduling algorithm chooses which packet is forwarded among those arriving on the IFs in the current time slot and those present in the buffers from previous time slots. Depending on the strategy adopted to extract the packets from FDLs, different queuing policies are obtained. Two different policies to manage the input buffering stage are here illustrated.

5.1.1 Scheduling algorithms for input buffered switch

Proper scheduling algorithms to manage the queues on the IFs are needed. A packet arriving in a time slot can be sent in the same time slot or in the next L time slot after. For this reason, it is necessary to decide as the packets are extracted of the FDLs. Due to the limited number of

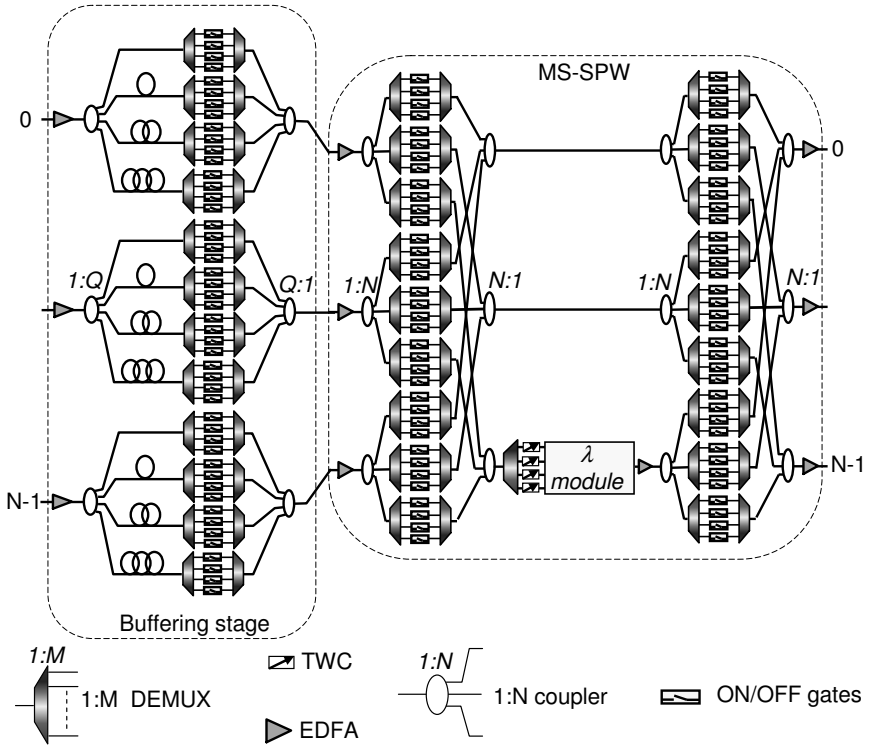


Figure 5.2: Input buffered switch with $N = 3$ input/output fibers carrying $M = 4$ wavelengths each. The node is equipped with $B = 1$ TWC block and $Q = 4$ FDLs per input fiber.

TWCs, to reduce the packet loss, the maximum number of packets must be forwarded without wavelength conversion. In addition in a time slot is preferable to forward the oldest packets, that are near to reach the maximum delay, after which they are discarded. Two scheduling disciplines are here illustrated for the buffers. First, a useful nomenclature is introduced:

- L_k^j the lists of packets carried by wavelength j on *k-th* input fiber
- T contains the total number of time slots
- C_k^j (Arrival Time Slot) contains the time slot in which the packet

on the head of L_k^j arrived

- OF_k the sets of free wavelengths on the *k-th* output fiber in the current time slot ($k = 0, \cdots, N-1$)
- BC_i the sets of free TWCs on the *i-th* block in the current time slot ($i = 0, \cdots, r_w - 1$)
- W_i^{BC} the sets of free wavelengths on output of the *i-th* TWC block in the current time slot ($i = 0, \cdots, r_w - 1$)
- SL_a the sets of free wavelengths on *a-th* optical fibers in the middle stage in the current time slot ($a = 0, \cdots, N - r_w - 1$)

To manage the packet forwarding, lists mechanism is introduced. Packets arriving in the same input channel are stored in the list L_k^j. A new packet arrival in a time slot is inserted on the tail of corresponding list. In this way, the packet on the head is the oldest of the list and the packet on the tail is the youngest. First policy that can be applied to manage the lists is the simple and well known FIFO (First-In First-Out) policy. In this case the algorithm is named FIFO-RR algorithm. The FIFO-RR algorithm serves each list in a first-in first-out manner. To assure the fairness among the fibers and wavelengths, two round robin counters (RRF and RRW)are used. RRF indicates which fiber has to be served first in a time slot (the other fibers are then scanned sequentially). RRF is incremented by one at each time slot. RRW indicates the wavelength that must be served first in each IF in a time slot. Differently from RRF, RRW is incremented by one when $RRF = 0$, every N time slots. The scheduling algorithm is composed by 2 steps executed in each time slot. In the first step, the packets arriving in the current time slot are inserted on the tail of the corresponding lists. In the second step, the lists are sequentially scanned and a function *sel_lambda* evaluates if the packets on the head of the lists can be forwarded. The pseudo-code of the FIFO-RR scheduling algorithm is here presented:

step 1:

```
for (i = 0 ; i < N ; i++)
    for (j = 0 ; j < M ; j++)
        if (there is a packet carried by λ_j on input fiber i)
            insert the packet in the tail of list L_i^j;
```

step 2:

for ($i = 0$; $i < N$; $i++$)
  for ($j = 0$; $j < M$; $j++$) {
    $in = (i + RRF) \bmod N$;
    $\lambda = (j + RRW) \bmod M$;
    if (L_{in}^{λ} != NULL) {
      out=output fiber of the packet in the head of L_{in}^{λ};
      tx=sel_lambda(λ, out);
      if ($tx > 0$)
        the packet is forwarded and removed from
        the head of L_{in}^{λ};
      if (($tx == -1$) && ($T - C_{in}^{\lambda} == L$))
        the packet is removed from the head of L_{in}^{λ};
    }
  }
$RRF = (RRF + 1) \bmod N$;
if ($RRF = 0$) $RRW = (RRW + 1) \bmod M$;

Note that the scheduling algorithm also removes the packets on the head of the lists if they are not forwarded in the current time slot and they have reached the maximum delay L (they are not available in the next time slot). *sel_lambda* procedure first tries to forward a packet without conversion, if this is not possible the packet is forwarded by exploiting wavelength conversion. *sel_lambda* returns the wavelength to forward the packet on the destination OF if possible, otherwise, when the packet cannot be forwarded *sel_lamnda* returns -1. The pseudo-code of the procedure *sel_lambda*, presented in detail in [38] presented below. By observing the architecture implementation, a path between input and output channels must be established to forward a packet. Suppose that the scheduling algorithm decide to forward a packet from $j-th$ (1) wavelength of the $k-th$ (2) IF to the $s-th$ (3) wavelength on the $h-th$ (4) OF. In addition suppose that the packet is stored in the switch from $T - C_k^j = u$ time slots and scheduling algorithm decide to forward it exploiting the $i-th$ (1) TWC block. In the first stage (buffering stage) the $j-th$ (1) gate on the $u-th$ WS connected to input k (2) has to be turned on. In the second stage the $j-th$ (1) gate in the $i-th$ (1)

WS connected to input k (2) has to be turned on. In the third stage, the $j-th$ (1) TWC on $i-th$ (1) block is used to convert a packet on the proper output wavelength s (3). Finally, in the $h-th$ (4) WS connected to $i-th$ (1) TWC block, the $s-th$ (4) gate must be turned on.

int sel_lambda(int λ_{in}***, int*** out***)*** {
  if ($\lambda_{in} \in OF_{out}$)
    for ($a = 0$; $a < N - r_w$; $a++$)
      if ($\lambda_{in} \in SL_a$) {
        λ_{in} is removed from SL_a and OF_{out};
        return λ_{in};
      }
  for ($i = 0$; $i < r_w$; $i++$)
    if (the TWC dedicated to $\lambda_{in} \in BC_i$)
      for ($j = 0$; $j < M$; $j++$)
        if ($\lambda_j \in W_i^{BC} \cap OF_{out}$) {
          remove λ_j from W_i^{BC} and OF_{out};
          remove the i-th TWC from BC_i;
          return λ_j;
        }
  return -1;
}

The FIFO-RR scheduling algorithm is affected by the Head Of Line (HOL) phenomenon, that is if a packet on the head of a list cannot be forwarded, it blocks the packets behind it even if they are directed to contention-free OFs. For this reason, a second policy to manage the lists is proposed in order to reduce the HOL effect. To do this the FIFO hypothesis is relaxed and a window (with size W) mechanism is introduced. When the list L_k^j is scanned, the algorithm tries to send the packet on the head (the oldest) but if this packet is blocked, one of the packets behind it belonging to a window W can be sent, if directed to a different (free) OF. In this paper the window size is equal to the buffer size, $W = L$.

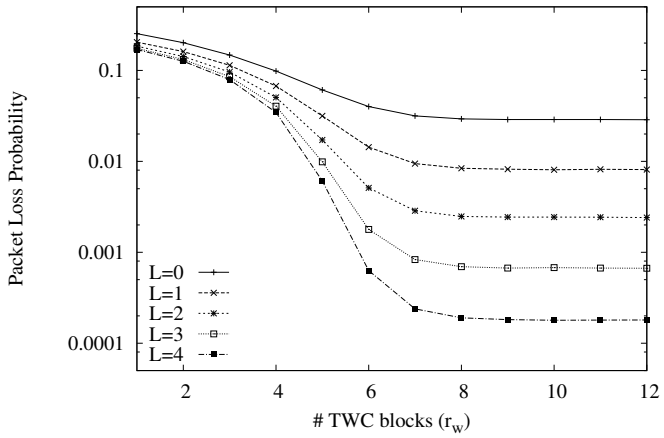


Figure 5.3: Packet loss probability as a function of the number of TWC blocks varying the buffer size, obtained by applying FIFO-RR algorithm in case $N = 16$, $M = 16$ and $p = 0.8$.

5.1.2 Numerical results

Results are obtained by simulation with confidence interval at 95% less than or equal to the 5% of the mean. Simulation are obtained considering Bernoulli independent arrivals on input wavelengths, where not differently specified.

First, in figure 5.3 PLP as a function of the number of TWC blocks r_w is plotted for the FIFO-RR algorithm, varying the buffer size L in case $N = 16$, $M = 16$ and $p = 0.8$. It can be seen that the same performance as fully equipped architecture can be obtained with limited number of TWC blocks. In addition, the figure shows that when the number of TWC blocks is high enough, by increasing by 1 the buffer size, PLP decreases about a half order of magnitude.

In figure 5.4 PLP as a function of the number of TWC blocks r_w varying the buffer size is plotted for both FIFO-RR and W-RR scheduling algorithms in case $N = 16$, $M = 16$ and $p = 0.9$. By increasing the number of the FDLs, packet loss decreases especially with W-RR. In addition the decrease as a function of the buffer size is more evident in the asymptotic value of the packet loss. More, for this high load, W-RR

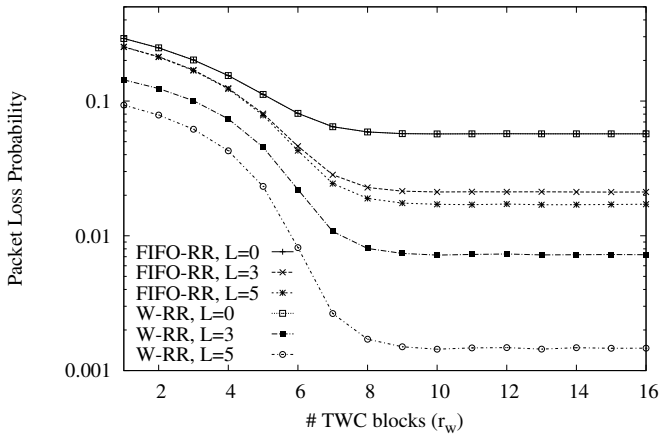


Figure 5.4: Comparison between the packet loss probability obtained by applying FIFO-RR and W-RR algorithms. PLP as a function of the number of TWC blocks varying the buffer size is plotted, in case $N = 16$, $M = 16$ and $p = 0.9$.

performs better than FIFO-RR, but the obtained PLP is high with this small buffer sizes.

Figure 5.5 shows the PLP obtained with FIFO-RR and W-RR algorithms as a function of the buffer size for the fully equipped ($r_w = N$) architecture in case $N = 16$, $M = 16$ and $p = 0.8, 0.9$. With high load ($p = 0.9$), even if L increases, the PLP obtained with FIFO-RR does not decrease significantly, due to the HOL phenomenon. W-RR algorithm provides better performance than FIFO-RR especially when the buffer size increases (4 order of magnitude when $L = 10$). When the load is $p = 0.8$, both algorithms improve the performance when L increases, and the gap between them are less evident, at least in the plotted region. When the load is below $p = 0.8$, FIFO-RR and W-RR provide almost the same packet loss (the HOL phenomenon becomes lower and lower as load decreases).

Finally in figure 5.6 the PLP with FIFO-RR is evaluated by taking into account admissible traffic (output contention-free) on input wavelengths in order to evaluate the packet loss due to the unavailability of the TWCs. PLP as a function of the number of TWC blocks is plot-

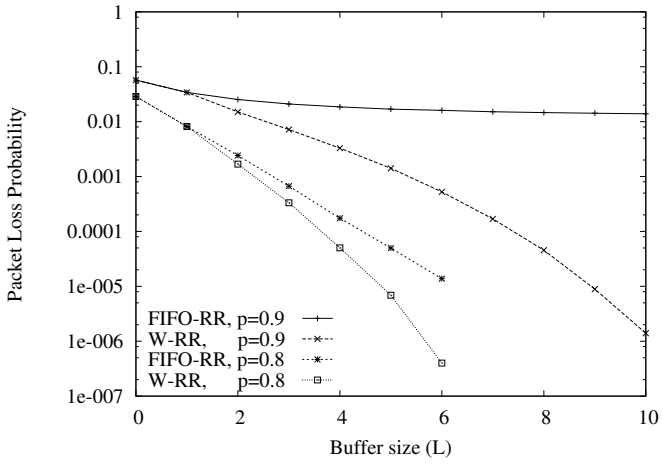


Figure 5.5: Comparison between the packet loss obtained by applying FIFO-RR and W-RR algorithms as a function of the buffer size for the fully equipped ($r_w = N$) architecture in case $N = 16$, $M = 16$ and $p = 0.8, 0.9$.

ted, varying the buffer size L, in case $N = 16$, $M = 16$ and $p = 0.8$. The PLP presents an asymptotic value, related to the limited capability of the scheduling algorithm for the MS-SPW switch in finding a match between input and output channels, due to the packet grooming effect, as explained in section 3.2. This effect is not related to the buffering stage, but to the conversion stage of the MS-SPW switch. Anyway, as the buffer size L increases, the asymptotic value of PLP decreases, by reaching negligible values even for not too high buffer size.

5.2 Electro/optical hybrid switch

In this section an hybrid electro/optical switch is presented. It is based on WSs to implement the switching fabric and on TWCs and electronic buffers to solve contention. The proposed architecture, called H-EOS (Hybrid Electro/Optical Switch) is presented in figure 5.7 in the case $N = 2$ IF/OF interfaces carrying $M = 4$ wavelengths each. An S-stage identical to the one presented in section 3.1 is used to directly connect IFs

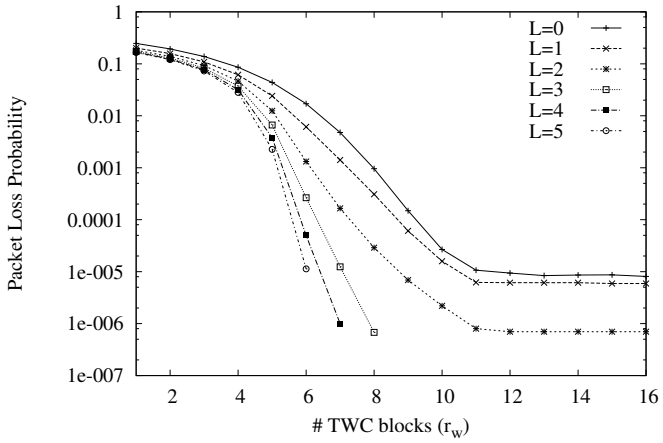


Figure 5.6: Packet loss probability as a function of the number of TWC blocks varying the buffer size L, for admissible traffic on input in case $N = 16$, $M = 16$ and $p = 0.8$

to the OFs. In this stage N WSs connects an IF to all the N OFs. A total amount of N^2 WSs are needed to obtain full connectivity among IFs and OFs. r_w TWC blocks are available to solve contention in the wavelength domain. Each TWC block is equipped with M TWCs dedicated per wavelength, exactly as for the MS-SPW (see section 3.1 for a detailed description of a TWC block). More, B blocks of electronic buffers are available to solve contention in time domain. Both TWC and buffer blocks are shared among the IFs, so each IF is connected to each TWC and buffer block by means of WSs. For this reason, $r_w + B$ additional WSs per IF are needed, for a total amount of $N(r_w + B)$ additional WSs to reach the blocks. Moreover, each of these blocks must be connected to all the OFs, so N WSs for each block are needed. Other $N(r_w + B)$ WSs are used. The total amount of WSs for the whole architecture is $N^2 + 2N(r_w + B) = N(N + 2(r_w + B))$, each equipped with M SOAs, so the complexity of the H-EOS results in:

$$N_{SOA}^{H-EOS} = NM(N + 2(r_w + B)) \tag{5.2}$$

Now, the main functionalities of the electronic buffer block, presented in figure 5.8, are described. Each buffer block is equipped with M elec-

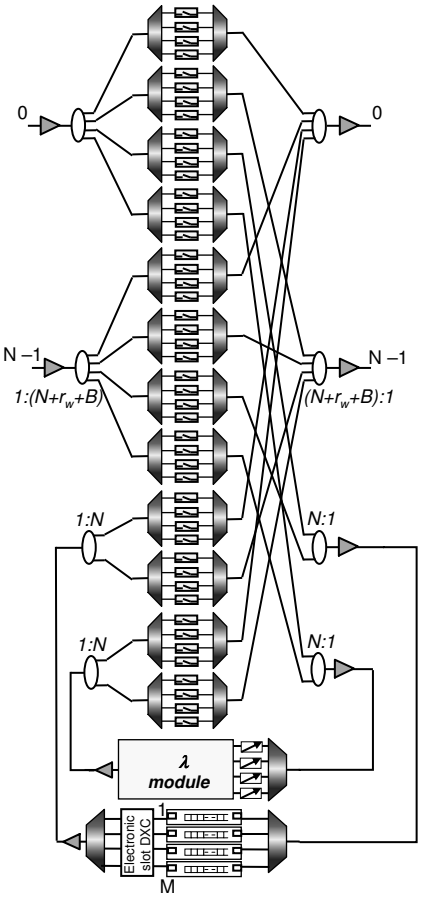


Figure 5.7: Hybrid electro/optical switch with $N = 2$ input/output fiber interfaces carrying $M = 4$ wavelengths each. The switch is equipped with $r_w = 1$ blocks of M TWCs dedicated per wavelength and $B = 1$ blocks of electronic buffers, each with M FIFO queues, one per wavelength.

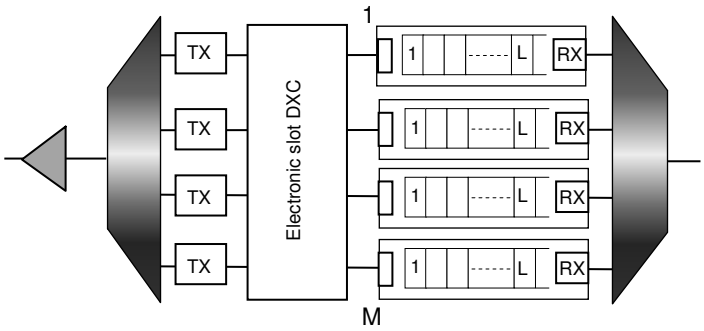


Figure 5.8: Example of electronic buffer block equipped with M line cards dedicated per wavelength.

tronic line cards, one per wavelength. First, the WDM signal on input is split by means of a DEMUX, and the M signals on different wavelength are interfaced with the M line cards. In each line card a packet coming on the corresponding wavelength is received and converted to the electronic domain, then it is stored in simple FIFO electronic queues, with L rooms each. At the transmission side, an electronic module synchronizes the packets that must be forwarded in a time slot, then they are converted to the optical domain by fixed transmitters in the same wavelength it arrived on the buffer. Finally the signals on different wavelengths are multiplexed in a single fiber by a MUX. As described before, this fiber is connected by WSs to the OFs.

A proper scheduling algorithm to control packet forwarding and resolve contention is needed. A scheduling algorithm is described in the next section.

5.2.1 Scheduling algorithm for the hybrid switch

In this kind of switches, the scheduling algorithm has to manage packets coming from IFs and packets stored in the electronic buffers. The hypothesis of this work is that the scheduler in each time slot gives priority to the packets in the electronic buffers, so these packets are served first. Another hypothesis is that packets stored in the buffers are sent without wavelength conversion, they cannot exploit the TWC blocks. The packets coming on the IFs are forwarded without wavelength conversion

if possible, otherwise they are sent by exploiting wavelength conversion. When there are not available wavelengths on the destination OF (output blocking) or there are not available TWCs to perform wavelength conversion, the packet is sent to the buffers. In particular the buffer block are scanned sequentially, and the packet is stored in the first block where the corresponding electronic queue (the one dedicated to the wavelength of the packet) has a free room. If all B buffer blocks are scanned without finding an available room for the packet, it is discarded.

The first part of the scheduling algorithm is devoted to schedule the packets on the electronic buffers. Given that these packets are forwarded without wavelength conversion, packets in queues dedicated to different wavelengths don't contend each other, so this part of the scheduling algorithm can be executed in parallel on the wavelengths. The flow chart of the scheduling algorithm for a generic wavelength k is presented in figure 5.9. The B buffer blocks are sequentially considered. When block

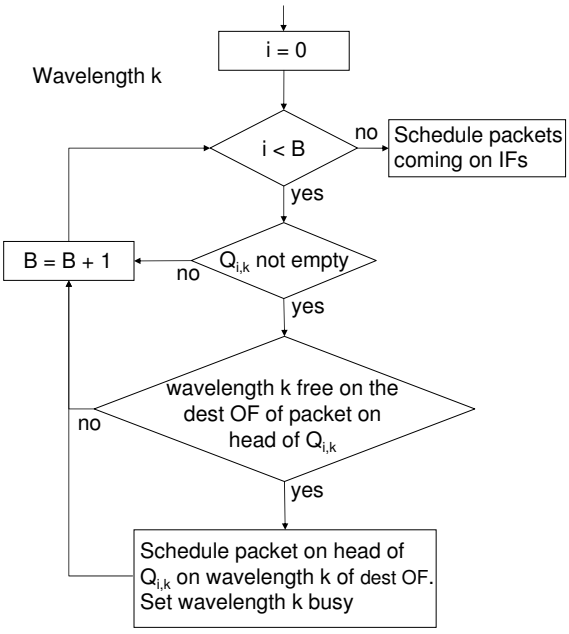


Figure 5.9: Flow chart of the first part of the scheduling algorithm. Packets in the electronic buffers are scheduled.

i is considered, if the queue dedicated to wavelength k in this block (in-

dicated with $Q_{i,k}$ is not empty, the algorithm controls if wavelength k is free on the destination OF of the packet at the head of the queue. If so, the packet is scheduled to that wavelength and it is occupied, otherwise the packet remains in the queue. Then, in both cases next block is considered. When all blocks has been considered, this phase ends and packets coming from IFs can be considered for possible forwarding. Note that only the packets at the head of the queues can be transmitted in a time slot, leading to the HOL phenomenon HOL influences the performance, accordingly to the available number of buffer blocks. The higher the number of blocks, the lower the effect of the HOL phenomenon.

When packets coming from IFs are considered, a scheduling algorithm very similar to the one presented in section 3.2 can be applied. It takes into account the packet grooming at the output of the TWC blocks. The only differences of the scheduling for this switch with respect to the algorithm for MS-SPW are
i) some wavelengths on the OFs can be already occupied by packets coming from electronic buffers
ii) the full connectivity among input and output fibers is assured by N^2 WSs (instead, in MS-SPW $N - r_w$ shared optical fibers connects IFs and OFs). A logical scheme of this phase of the scheduling algorithm is presented in figure 5.10. When a packet arrives, if no wavelengths

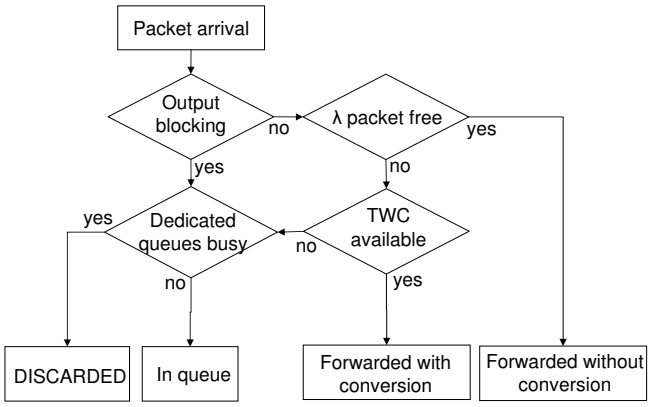


Figure 5.10: Logical example of packet scheduling for a packet coming on an input channel.

are available on the destination OF, the packet is blocked due to output

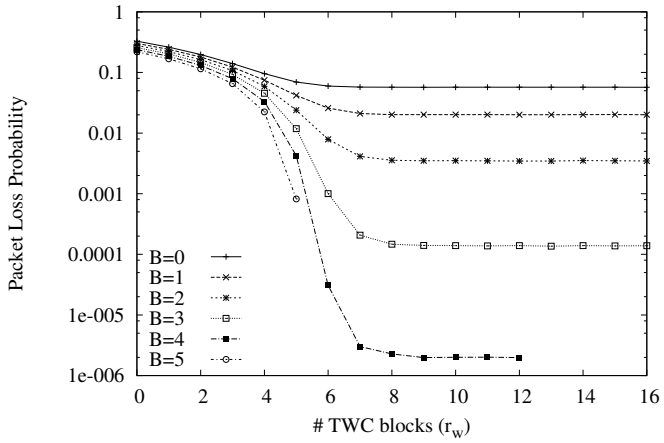


Figure 5.11: Packet loss probability as a function of the number of TWC blocks varying the number of buffer blocks in case $N = 16$ input/output fibers, $M = 16$ wavelengths per fiber, load $p = 0.9$ and buffer size $L = 5$.

blocking and it needs to be buffered. Otherwise, if at least one wavelength is available, it is sent without conversion if its wavelength is free on the corresponding OF, otherwise it is sent with wavelength conversion, if at least one available TWC is found. Otherwise, the packet is sent to the buffers. When this happens, the buffer blocks are sequentially considered to store the packet, and it is stored in the first available queue. Finally the packet is lost when all blocks are considered without finding a room to store the packet.

5.2.2 Numerical results for the hybrid switch

In this section some results of H-EOS switch, obtained by simulation, are presented. First in figure 5.11 PLP obtained with H-EOS switch as a function of the number of TWC blocks is plotted varying the number of buffer blocks B in case $N = 16$, $M = 16$, $p = 0.9$ and buffer size $L = 5$. The figure shows that the asymptotic values of PLP can be obtained with a reduced number of TWC blocks, almost independently from the number of buffer blocks. This asymptotic value of PLP rapidly decreases with the number of buffer blocks, and also for high load ($p = 0.9$) a few

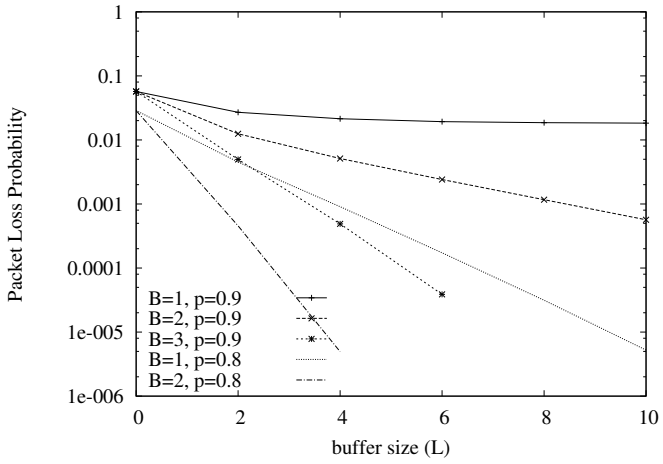


Figure 5.12: Packet loss probability as a function of the buffer size L varying the number of buffer blocks B for the fully equipped switch ($r_w = 16$), in case $N = 16$ input/output fibers, $M = 16$ wavelengths per fiber and load $p = 0.8, 0.9$.

buffer blocks (B=4) is enough to obtain relative low PLP. The figure also shows that the decrease of PLP becomes higher and higher as B increases.

Figure 5.12 shows the PLP as a function of the buffer size L, varying the number of buffer blocks B, for the fully equipped switch (in terms of TWCs) in case $N = 16$, $M = 16$, $p = 0.8, 0.9$. The figure shows how the PLP becomes low when the buffer size increases if enough buffer blocks are used ($B = 3$). In particular when the load is not very high ($p = 0.8$), the PLP becomes negligible as L increases even when a single buffer block is employed.

Chapter 6

Cost model for optical switches

In this chapter a simple cost model to compare in terms of cost different architectures relying on the same main optical components is presented [46]. To obtain a simple model, a parametric evaluation is here provided, and only the components that are very expensive and/or used in a relevant number in the architectures are taken into account. The model gives information about when an architecture is convenient with respect to another one, but does not provide information about the cost difference between the architectures. First, the parametric cost model is presented in section 6.1, then it is applied to compare SPN (section 2.2) and MS-SPW (chapter 3) switches in section 6.2. Finally in section 6.3 it is shown how the proposed model can be applied to compare different configurations of a multi-fiber switch.

6.1 Cost model

In this section the model is presented. The components taken into account in the definition of the cost model are TWCs (which are considered more complex than other optical devices) and SOA gates, which are used in a relevant number with respect to other components. The aim of the cost model is to compare two generic architectures (named A and B) relying on the same kind of TWCs and SOAs. In particular the TWCs should have the same tuning range. It is assumed that a TWC is α times

more expensive than an SOA, leading to the expression

$$C_{TWC} = \alpha C_{SOA} \quad , \quad \alpha \geq 1 \tag{6.1}$$

where C_{TWC} and C_{SOA} are the costs of a TWC and an SOA respectively. By indicating with N^A_{SOA}, N^A_{TWC} and N^B_{SOA}, N^B_{TWC} the number of TWCs and SOAs needed for architectures A and B respectively, the cost of these architectures can be expressed as:

$$\begin{aligned} C_A &= N^A_{SOA} C_{SOA} + N^A_{TWC} C_{TWC} = (N^A_{SOA} + \alpha N^A_{TWC}) C_{SOA} \\ C_B &= N^B_{SOA} C_{SOA} + N^B_{TWC} C_{TWC} = (N^B_{SOA} + \alpha N^B_{TWC}) C_{SOA} \end{aligned} \tag{6.2}$$

By imposing $C_A < C_B$ the conditions leading to architecture A less expensive than architecture B can be found. This inequality must be satisfied:

$$N^A_{SOA} + \alpha N^A_{TWC} < N^B_{SOA} + \alpha N^B_{TWC} \tag{6.3}$$

After some algebra, the condition $C_A < C_B$ results in:

$$\begin{aligned} &\alpha < \alpha_{th} \quad when \quad N^A_{TWC} - N^B_{TWC} > 0 \\ &\alpha > \alpha_{th} \quad when \quad N^A_{TWC} - N^B_{TWC} < 0 \end{aligned} \tag{6.4}$$

where

$$\alpha_{th} = \frac{N^B_{SOA} - N^A_{SOA}}{N^A_{TWC} - N^B_{TWC}} \tag{6.5}$$

is the threshold value that identifies the regions of convenience of the two architectures. Five cases are trivial to be evaluated, but are here reported for the sake of completeness.
i) **A and B use the same number of TWCs ($N^A_{TWC} = N^B_{TWC}$)**: by observing (6.4) the condition to have $C_A < C_B$ becomes $N^A_{SOA} < N^B_{SOA}$ independently from α (the most expensive architecture is the one that uses the higher number of SOAs). In this case $\alpha_{th} = +\infty$ when $N^B_{SOA} > N^A_{SOA}$ and $\alpha_{th} = -\infty$ when $N^B_{SOA} < N^A_{SOA}$.
ii) **A and B use the same number of SOAs ($N^A_{SOA} = N^B_{SOA}$)**, the condition $C_A < C_B$ results in $N^A_{TWC} < N^B_{TWC}$ (the most expensive architecture is the one that uses the higher number of SOAs). In this case $\alpha_{th} = 0$.

iii) **A and B use the same number of SOAs and TWCs ($N^A_{SOA} = N^B_{SOA}$ and $N^A_{TWC} = N^B_{TWC}$)**: A and B have the same cost. Note that in this case the inequality (6.4) is never satisfied, and the value of α_{th} is mathematically not defined ($\alpha_{th} = \frac{0}{0}$).
iv) **A uses more TWCs and SOAs than B ($N^A_{TWC} - N^B_{TWC} > 0$ and $N^B_{SOA} - N^A_{SOA} < 0$)**: $\alpha_{th} < 0$, so the condition $\alpha < \alpha_{th}$ is never satisfied, A is more expensive than B for all values of α.
v) **A uses less TWCs and SOAs than B ($N^A_{TWC} - N^B_{TWC} < 0$ and $N^B_{SOA} - N^A_{SOA} > 0$)**: $\alpha_{th} < 0$ and the condition $\alpha > \alpha_{th}$ is always satisfied, so architecture A is less expensive than B for all values of α.

The most interesting cases are $N^A_{TWC} - N^B_{TWC} > 0$, $N^B_{SOA} - N^A_{SOA} > 0$ and $N^A_{TWC} - N^B_{TWC} < 0$, $N^B_{SOA} - N^A_{SOA} < 0$ leading to $\alpha_{th} > 0$. In fact in this situation an architecture is convenient with respect to the other depending on the relative cost of an SOA and a TWC. By evaluating the number of SOAs and TWCs needed for the two architectures and by applying the proposed formulas (6.4) and (6.5) it is possible to find the less expensive architecture. In table 6.1 are shown the regions of convenience of A with respect to B accordingly to the different admitted values of α_{th}.

	$N^A_{TWC} = N^B_{TWC}$	$N^A_{TWC} > N^B_{TWC}$	$N^A_{TWC} < N^B_{TWC}$
$\alpha_{th} = -\infty$	never	-	-
$\alpha_{th} < 0$	-	never	always
$\alpha_{th} = 0$	-	never	always
$\alpha_{th} > 0$	-	$\alpha < \alpha_{th}$	$\alpha > \alpha_{th}$
$\alpha_{th} = +\infty$	always	-	-

Table 6.1: Regions of convenience of architecture A with respect to B accordingly to the values of α_{th}.

It is worthwhile outlining that the proposed cost model is valid in any traffic scenario, by providing the number of TWCs needed for a given switch configuration and loss target, and the corresponding number of SOAs.
Finally, some observations with respect to this cost model.
i) The proposed cost model only allows to determine less expensive configurations, but it does not provide a quantitative evaluation of cost saving.

To this end the exact value of the parameter α should be known;
ii) the values of α for different number of wavelengths per fiber (M) cannot be directly compared, given that the TWCs tuning range varies with M, so $\alpha = \frac{C_{TWC}}{C_{SOA}}$ also varies.

6.2 Cost comparison for SPN and MS-SPW

As an example, the proposed model is here applied to compare the SPN and MS-SPW switches proposed in section 2.2 and chapter 3 respectively. The same comparison can be found in [46]. The aim is here to evaluate and compare the cost of the architectures for a given loss performance. In figure 6.1 a comparison between the performance of the SPN and MS-SPW for Bernoulli input traffic, obtained with the analytical models presented in sections 2.2.3, 2.3.3 respectively, is shown. PLP is plotted

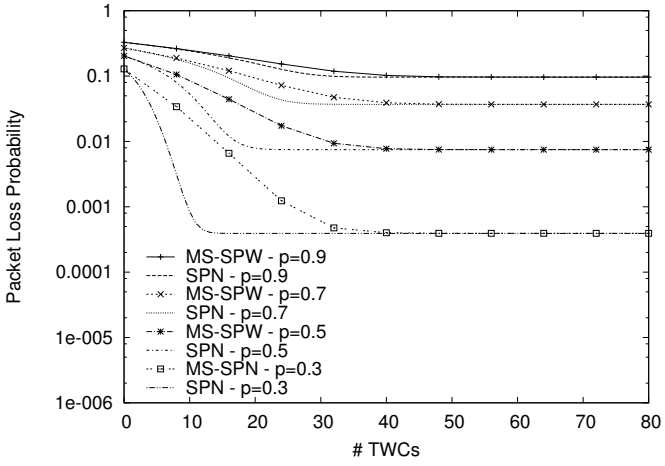


Figure 6.1: Comparison between SPN and MS-SPW architectures. PLP as a function of the total number of TWCs varying load, in the case $N = 16$, $M = 8$.

as a function of the total number of the TWCs varying load in case $N = 16, M = 8$. From this figure the minimum number of TWCs needed to obtain the asymptotic value of packet loss can be determined. The corresponding number of SOAs can be obtained by applying formulas (2.25) and (3.4) for SPN and MS-SPW respectively. SPN requires less

TWCs than MS-SPW, so the condition $C_{MS-SPW} < C_{SPN}$ is satisfied when $\alpha < \alpha_{th}$.
As an example in case $p = 0.5$ the same performance as fully equipped architecture is obtained with $N_{TWC}^{SPN} = 22$ and $N_{TWC}^{MS-SPW} = 48$ for SPN and MS-SPW respectively. By applying formula (6.5) the value $\alpha_{th} = 675.69$ is obtained. This value suggests that the MS-SPW architecture is always less expensive than the reference SPN architecture, given that it seems not realistic that a TWC is 675 times more expensive than an SOA. By exploiting graphs similar to the one presented in figure 6.1 the value of α_{th} for each value of a couple (N, M) can be found.
In figure 6.2 the value of α_{th} as a function of the number of input/output fibers N is plotted, in case $M = 16$ and $p = 0.3, 0.5, 0.7$. The value of

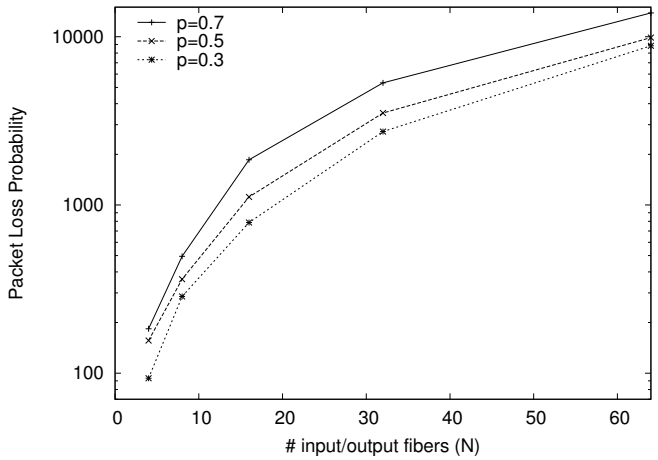


Figure 6.2: Values of α_{th} for MS-SPW and SPN architectures, as a function of N in the case $M = 16$, $p = 0.3, 0.5, 0.7$. MS-SPW is less expensive than SPN when $\alpha < \alpha_{th}$.

α_{th} increases as N increases. Given that the SPN architecture is less costly than the MS-SPW when $\alpha > \alpha_{th}$, when N is high the cost of a TWC must be very higher than the cost of an SOA in order that the SPN being convenient with respect the MS-SPW. In figure 6.3 the value of α_{th} is plotted as a function of the number of wavelengths per fiber M in case $N = 16$, $p = 0.3, 0.5, 0.7$. Also in this case, α_{th} increases as M increases, leading to similar results as in the previous case. The value of α_{th} is lower for light load, but is in any case high.

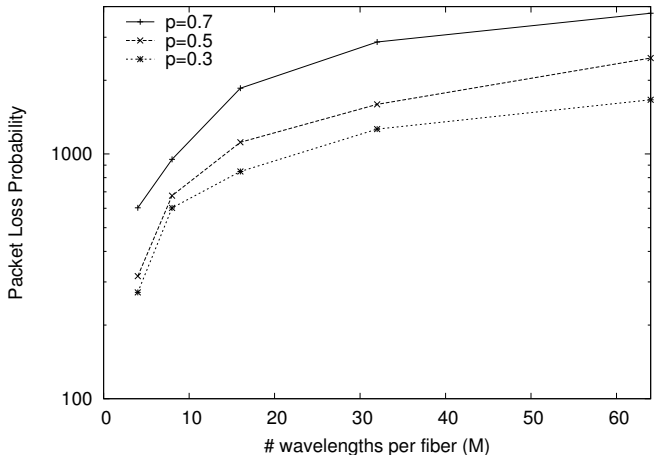


Figure 6.3: Values of α_{th} for MS-SPW and SPN architectures, as a function of M in the case $N = 16$, $p = 0.3, 0.5, 0.7$. MS-SPW is less expensive than SPN when $\alpha < \alpha_{th}$.

6.3 Cost comparison for multi-fiber switches

This model can also be applied to compare in cost different configurations of multi-fiber switches (MF-SPN or MF-SPW presented in chapter 4). Here the MF-SPW switch is considered as an example. The minimum number of TWCs N_{TWC} to obtain the asymptotic value of PLP is considered for each switch configuration

$$N_{TWC} = Mr_w^{th} \tag{6.6}$$

where r_w^{th} (th is for threshold) is the minimum number of TWCs per wavelength needed to obtain asymptotic PLP. The values of N_{TWC} for different configurations and switch dimensioning can be obtained from loss graphs similar to that presented in section 4.3.

Taking into account formula (6.6), the complexity in terms of SOAs, already presented in formula (4.3) in section 4.2.1, can be expressed as:

$$N_{SOA} = (NF)^2 M + 2NFN_{TWC} \tag{6.7}$$

The first part of this formula refers to the number of SOAs needed to connect IFs and OFs directly, while the second part refers to the number

of SOAs needed to connect the TWC pools to the IFs and OFs. When the number of channels per interface is maintained ($MF = W$), by varying the number of fibers per interface and the number of wavelengths per fiber accordingly, formula (6.7) can be rewritten as:

$$N_{SOA} = N^2FW + 2NFN_{TWC} \tag{6.8}$$

the first term increases linearly with F. In fact, the higher the number of fibers per interface, the higher the number of SOAs needed to connect IFs and OFs directly. The second term is proportional to both F and the total number of TWCs employed N_{TWC}. The second one decreases as F increases (see figures 4.9, 4.10). In fact, when F increases, the spatial diversity assured by the employment of F fibers per interface can be exploited to solve contention, so a lower number of TWCs is needed.

Tables 6.2, 6.3, 6.4 present the number of TWCs (first row) and SOAs (second row) needed for the switch configurations $N = 16$, $W = 16$ - $N = 10$, $W = 24$ and $N = 20$, $W = 100$ presented in figures 4.6, 4.7 and 4.8 respectively. The numbers in the tables have been carried out from simulation results, due to the limits of the analysis in the switch configurations plotted in these figures. Tables show that, as F increases, the number of TWCs decreases while the number of SOAs increases. More, a relevant number of SOAs is needed to implement the switching fabrics, especially when N and/or W is high (table 6.4). To apply the model to

	(F, M)			
	(1, 16)	(2, 8)	(4, 4)	(8, 2)
N_{TWC}	96	56	24	10
N_{SOA}	7168	11776	19456	35328
α_{th}	–	115.2	170.66	327.44

Table 6.2: Number of TWCs and SOAs for different switch configurations, in case $p = 0.3$, $N = 16$, $MF = 16$

the MF-SPW, it is necessary to assume that TWCs with different tuning range have the same cost. To compare different configurations in terms of number of fibers per interface/number of wavelengths per fiber, by fixing N and $W = MF$, a configuration is chosen as a reference, and the others

	(F, M)			
	(1, 24)	(2, 12)	(4, 6)	(6, 4)
N_{TWC}	144	84	48	28
N_{SOA}	5280	8160	13440	17760
α_{th}	–	48	85	107.58

Table 6.3: Number of TWCs and SOAs for different switch configurations, in case $p = 0.4$, $N = 10$, $MF = 24$

	(F, M)				
	(1, 100)	(2, 50)	(5, 20)	(10, 10)	(20, 5)
N_{TWC}	1100	700	450	200	90
N_{SOA}	84000	136000	290000	480000	872000
α_{th}	–	130	316.92	440	780.19

Table 6.4: Number of TWCs and SOAs for different switch configurations, in case $p = 0.7$, $N = 20$, $MF = 100$

are compared with this one. The number of TWCs and SOAs for the reference configuration are indicated with N_{TWC}^{ref} and N_{SOA}^{ref} respectively, and its the cost can be expressed as:

$$C_{ref} = N_{SOA}^{ref} C_{SOA} + N_{TWC}^{ref} C_{TWC} = (N_{SOA}^{ref} + \alpha N_{TWC}^{ref}) C_{SOA} \qquad (6.9)$$

Here the mono-fiber case is assumed as the reference and it is compared to the different multi-fiber configurations. The values of α_{th} obtained by applying (6.5) are presented in the third row of tables 4.1, 4.2, 4.3. In this case $C_{ref} < C_A$ when $\alpha < \alpha_{th}$, as deduced from (6.4) by observing the values of N_{TWC} in the tables. The obtained values for α_{th} are quite high, especially when N and/or W are high. In any case, the value of α_{th} rapidly increases with F. The condition $C_{ref} > C_A$, obtained when $\alpha > \alpha_{th}$ is not easily satisfied when F is high due to the high value of α_{th}.

In this cost evaluation, the cost of a TWC is assumed to be constant with respect to the tuning range. In the real case, the lower the tuning range, the lower the cost of a TWC, and this leads to lower values of α_{th}.

To take into account this in the cost model, two assumptions are made i) α is the relation between the cost of an SOA and a TWC with the tuning range needed for the mono-fiber case ii) the cost of a TWC is a linear function of the tuning range. The tuning range needed is related to the number of fibers per interface, and in particular the tuning range decreases as F increases. The tuning range of a multi-fiber configuration A is $\frac{F^A}{F^{ref}}$ times lower than the tuning range of the reference (mono-fiber) configuration, where F^{ref} and F^A are the number of fibers per interface of reference and A configurations respectively. For this reason, the cost of a TWC needed for the multi-fiber configuration A is $\frac{F^A}{F^{ref}}$ lower than the cost of a TWC in the reference configuration. The cost of the multi-fiber configuration A can be rewritten as:

$$C_A = N^A_{SOA}C_{SOA} + N^A_{TWC}C_{TWC}\frac{F^{ref}}{F^A} = (N^A_{SOA} + \alpha\frac{F^{ref}}{F^A}N^A_{TWC})C_{SOA} \tag{6.10}$$

and the expression of α_{th} becomes:

$$\alpha_{th} = \frac{N^A_{SOA} - N^{ref}_{SOA}}{N^{ref}_{TWC} - \frac{F^{ref}}{F^A}N^A_{TWC}} \tag{6.11}$$

With these assumptions the values of α_{th} in the case presented in table 6.3 becomes 28.23, 61.81, 89.57 instead of 48, 85, 107.58 respectively. The values of α_{th} are reduced due to the lower cost of TWCs when multi-fiber configurations are employed. Anyway, the hypothesis that the cost linearly decreases with the range is optimistic. In fact it is more probable that a cost of a TWC increases more rapidly when the tuning range is low and slowly when the tuning range is high. To have a better evaluation it is necessary to know the exact TWC cost function varying the tuning range. Anyway, the real value of α_{th} will be between the two values here proposed.
Some observations with respect to the cost evaluation in multi-fiber switches.
i) the values of α for different M and W cannot be directly compared, given that they influence the needed tuning range.
ii) Due to the still immature technology of the TWCs with respect to SOAs, also quite high values of α are considered to be reasonable.
In fact recently arrays of 32 SOAs have been shown to be feasible [47],

and it is reasonable to think that in a few time this devices will be available with a relative low cost, while when optical TWCs will be available, their cost should be very high due to technological constraints.
iii) α_{th} increases as F increases, but the real value of α decreases as F increases, given that the tuning range decreases with F, so also the TWC cost decreases. Also for this reason, the condition $C_{ref} > C_A$ ($\alpha > \alpha_{th}$) is very difficult to be satisfied with high F.
iv) This evaluation takes into account only the cost, so even if mono-fiber solution should be less costly under some switch dimensioning, the multi-fiber scheme could be anyway required due to the difficulty to realize TWCs with large tuning range. In addition, in some context the switching fabrics could be implemented with cheaper technology (like MEMS), anyway assuring a good trade-off between cost and performance even when the number of fibers per interface is high.

Chapter 7

Conclusions

The exploitation of DWDM photonic technology, not only for transport of information but also for switching operations, seems to be one of the most appealing and cost-effective solutions for the next generation networks. In this context photonic packet switches play a fundamental role to build high-speed service-oriented optical networks.

In this work different switch architectures able to exploit the available domains (wavelength, time and space) to solve contention have been proposed. Ideal schemes and possible implementations relying on optical devices expected to be mature in the near future have been proposed, in particular:

- a new shared-per-wavelength scheme for the wavelength converters has been presented, and a possible implementation with low complexity proposed;

- implementations for multi-fiber switches to solve contention in space and wavelength domains have been presented

- implementations of buffered switches, both all-optical and electro/optical have been presented

For all the proposed solutions, heuristic scheduling algorithms to control packet forwarding have also been studied. Performance has been carried out by means of analysis and simulation, showing the effectiveness of the proposed solutions in solving contention.

Finally, complexity in terms of expensive optical components has also been evaluated and a simple cost model to compare different architectures has been presented.

Bibliography

[1] R. Ramaswami, K. N. Sivarajan, "Optical networks: a practical perspective", *Morgan Kaufmann publishers, Inc.*, San Francisco, California, U.S.A., 1998.

[2] P. Gambini et al., "Transparent Optical Packet Switching: network architecture and demonstrators in the KEOPS project", *IEEE Journal on Selected Areas in Communications*, Vol. 16, pp. 1245-1253, Sept. 1998.

[3] C. Guillemot et al., "Tranparent Optical Packet Switching: The European ACTS KEOPS Approach", *Journal of Lightwave Technology*, Vol. 16, no. 12, pp. 2117-2134, Dec. 1998.

[4] L. Dittman et al., "The European IST project DAVID: A Viable approach towards Optical Packet Switching", *IEEE Journal on Selected Areas in Communications*, Vol. 21, n. 7, pp. 1026-1040, Sept. 2003.

[5] B. Mukherjee, F. Neri, "Report of US/EU Workshop on Key Issues and Grand Challenges in Optical Networking", June 2006, available at *http://networks.cs.ucdavis.edu/mukherje/US-EU-wksp-June05.html*.

[6] M. J. O'Mahony, D. Simeonidou, D. K. Hunter, A. Tzanakaki, "The application of Optical Packet Switching in Future Communications Networks", *IEEE Communication Magazine*, vol. 39, no. 3, pp. 128-135, Mar. 2001.

[7] M. Listanti, V. Eramo, R. Sabella, "Architectural and Technological Issues for Future Optical Internet Networks", *IEEE Communication Magazine*, Vol. 38, No. 9, pp. 82-92, Sept. 2000.

[8] B. Li, Y. Qin, X. Cao, K. Sivalingam, "Photonic Packet Switching: Architectures and Performance", *Opt. Network Magazine*, pp. 27-39, Jan/Feb. 2001.

[9] S Yao, B. Mukherjee, S. Dixit, "Advances in Photonic Packet Switching: An Overview", *IEEE Comm. Magazine*, Vol. 38, pp. 84-94, Feb. 2000.

[10] A. Jourdan, D. Chiaroni, E. Dotaro, G. Eilenberger, F. Masetti, M. Renaud, "The Perspective of Optical Packet Switching in IP-Dominant Backbone and Metropolitan Networks", *IEEE Comm. Magazine*, Vol.39 pp. 136-141, Mar. 2001.

[11] Xueli Hou, H.T. Mouftah, "Towards optimal design of wavelength-convertible optical switches for the all-optical next-generation internet", *IEEE CCECE 2001*, pp. 0441.

[12] A. Stavdas, H. Avramopoulos, E.N. Protonotarios, M.J. Midwinter, "An OXC Architecture suitable for High Density WDM Wavelength Routed Networks", *Photonic Network Communications*, Vol. 1 (1), pp. 77-88, 1999.

[13] F. Callegati, W. Cerroni, C. Raffaelli, P. Zaffoni: Wavelength and time domain exploitation for QoS management in optical packet switches, *Computer Networks*, vol. 44, no. 4, pp. 569-582, 15 Mar. 2004.

[14] N. Akar, E. Karasan, C. Raffaelli, M. Savi, "Packet Loss Analysis of Synchronous Buffer-less Optical Switch with Shared Limited Range Wavelength Converters", *In proceedings of High Performance Switching and Routing (HPSR 2007) IEEE Workshop*, New York, U.S.A., 31 May - 1 June 2007.

[15] V. Eramo, M. Listanti, G. Pacifici, "A Comparison Study on the Number of Wavelength Converters Needed in Synchronous and Asynchronous All- Optical Switching Architectures", *IEEE Journal of Lightwave Technology*, Vol. 21, No. 2, pp. 340-355, Feb. 2003.

[16] Eramo, M. Listanti, M. Spaziani, "Resource Sharing in Optical Packet Switches With Limited Range Wavelength Converters",

IEEE Journal of Lightwave Technology, Vol. 23, No. 2, pp. 671-686, Feb. 2005.

[17] C. Joergensen et al., "4 Gb/s optical wavelength conversion using semiconductor optical amplifier", *IEEE Photonic Technology Letters*, vol. 5, no. 6, pp 657-670, June 1993.

[18] S. L. Danielsen et al., "WDM packet switch architecture and analysis of the influence of tuneable wavelength converters on the performance", *Journal of Lightwave Technology*, vol. 16, no 12, pp. 2095-2108, Dec. 1998.

[19] V.Eramo, A.Germoni, C. Raffaelli, M. Savi, "Multi-Fiber Shared-Per-Wavelength All-Optical Switching: Architecture, Control and Performance" *to appear on Journal of Lightwave Technology.*

[20] H. Obara, H. Masuda, K. Suzuki, K Aida, "Multifiber Wavelength-division Multiplexed Ring Network Architecture for Tera-bit/s Throughput", *In proceedings of IEEE International Conference on Communications (ICC 1998)*, Vol. 2, pp. 921-925, June 1998.

[21] N. Wauters, P. Demeester, "Wavelength Conversion in Optical Multi-Wavelength Multifiber Transport Networks", *Int. J. Optoelectron.*, vol. 11, no. 1, pp. 53-70, 1997.

[22] F. A. Al-Zaharani, A.A. Habiballa, A. P. Jayasumana, "Performance Merits of Multi-Fiber DWDM Networks Employing Different Shared Wavelength Conversion Resources Architectures", *Region 5 Conference: Annual Technical and Leadership Workshop*, April 2004, pp. 59 - 67, published by IEEE.

[23] V. Eramo, "An Analytical Model for TOWC Dimensioning in a Multifiber Optical-Packet Switch", *Journal of Lightwave Technology*, Volume 24, Issue 12, pp. 4799 - 4810, Dec. 2006.

[24] F. Callegati. "Which packet length for a transparent optical network?" *In proceedings of SPIE 1997 Symposium on Broadband Networking Technology*, Dallas, USA, Nov. 1997.

[25] F. Callegati, W. Cerroni, "Wavelength Allocation Algorithms in Optical Buffers", *In proceedings of IEEE International Conference on Communications (ICC 2001)*, Helsinki, Finland, Vol. 2, pp. 499-503, June 2001.

[26] F. Callegati, W. Cerroni, "Time-Wavelength Exploitation in Optical Feedback Buffer with Trains of Packets", *3rd International Conference on Optical Networking and Communications (OptiComm 2002)*, Proc. of SPIE, Vol. 4874, pp. 274-285, 2002.

[27] T. G. Orphanoudakis, A. Drakos, C. Matrakidis, C. Politi, A. Stavdas, "A Hybrid Optical Switch Architecture with Shared Electronic Buffers", *In proceedings of International Conference on Transparent Optical Networks (ICTON) 2007*, Rome, Italy, July 2007.

[28] V.Eramo, A.Germoni, C. Raffaelli, M. Savi, "Performance Analysis of Multi-Fiber All-Optical Switches Employing Fixed-Input/Tunable-Output Wavelength Converters", *In proceedings of IT-NEWS (QoS-IP) 2008*, 4th International Telecommunication Network WorkShop on QoS Multiservice IP Networks, Venice, Italy, 13-15 Feb. 2008.

[29] C. Raffaelli, M. Savi, "Performance Modelling of Synchronous Buffer-less Optical Packet Switch with Partial Wavelength Conversion", *In proceedings of IEEE International Conference on Communications (ICC 2006)*, Istanbul, Turkey, 11 - 15 June 2006.

[30] H. Overby, "Performance Modelling of Synchronous Bufferless OPS Networks", *In proceedings of International Conference on Transparent Optical Networks (ICTON) 2004*, vol.1, pp. 22-28, 2004.

[31] G. Maier, A. Pattavina, "Generalized Space Equivalent Analysis of Optical Cross-Connect Architectures", *In proceedings of IEEE Infocom 2001*, pp. 159-168, 2001.

[32] A. Stavdas, "Architectures, Technology and Strategies for a Gracefully Evolving Optical Packet Switching Networks". *SPIE Optical Networks Magazine*, Vol.4 No.3, pp. 92-107, 2003.

[33] S. Okamoto, A. Watanabe, K.-I. Sato, "Optical path cross-connect node architectures for photonic transport network", *Journal of Lightwave Technology*, vol. 14 no. 6, pp. 1410-1422, June 1996.

[34] M. Izal, J. Aracil, "On the Influence of Self Similarity on Optical Burst Switching Traffic", *In proceedings of Globecom 2002*, 2002.

[35] A. E. Eckberg, T.-C. Hou "Effects of output buffer sharing on buffer requirements in an ATDM packet switching", *INFOCOM '88. Networks: Evolution or Revolution? Proceedings. Seventh Annual Joint Conference of the IEEE Computer and Communcations Societies.*, pp. 459 - 466 Mar 1988.

[36] V. Eramo, M. Listanti, "Packet loss in a bufferless optical WDM switch employing shared tunable wavelength converters", *Journal of Lightwave Technology*, Volume 18, Issue 12, pp. 1818 - 1833, Dec 2000.

[37] D. Hunter et al., "WASPNET: A Wavelength Switched Packet Network", *IEEE Communication Magazine*, pp. 120-129, Mar. 1999.

[38] C. Raffaelli, M. Savi, A. Stavdas "Performance of Scheduling Algorithms in Multi-Stage Optical Packet Switches with Sparse Wavelength Converters", *In proceedings of Globecom 2006*, San Francisco, U.S.A., 27 November-1 December 2006.

[39] A. Stavdas: Design of multiplexer/demultiplexer for dense WDM wavelength routed optical networks, *Ph.D. Thesis, University of London*, 1995.

[40] A. Stavdas, H. Avramopoulos, E. Protonotarios, J. E. Midwinter "A novel architecture for a wavelength crossconnect node", *Photonic Network Communications*, Vol. 1 (1), pp. 77-88, 1999.

[41] C. Raffaelli, M. Savi, A. Stavdas, "Sharing Wavelength Converters in Multistage Optical Packet Switches", *In proceedings of High Performance Switching and Routing (HPSR 2006) IEEE Workshop*, Poznan, Poland, pp. 203-208, June 2006.

[42] A. Bianco, F. Neri, C. Piglione, "Optical switching nodes: architectures and performance", *In proceedings of High Performance Switching and Routing (HPSR 2007) IEEE Workshop*, New York, U.S.A., June 2007.

[43] Zvi Galil, "Efficient algorithms for finding maximum matching in graphs", *ACM Computing Surveys (CSUR)*, Volume 18, Issue 1, pp. 23 - 38, Mar. 1986.

[44] C. Raffaelli, M. Savi, "Traffic performance of Buffered Multi-Stage Optical Packet Switch", *In proceedings of Photonics in Switching 2006*, Crete, Greece, Oct. 2006.

[45] W. Feller, "An Introduction to Probability Theory and its Applications", *Wiley, 1968* Volume 1.

[46] C.Raffaelli, M. Savi, "Cost Comparison of Optical Packet Switches with Shared Wavelength Converters", *In proceedings of International Conference on Transparent Optical Networks (ICTON) 2007*, Rome, Italy, July 2007.

[47] N. Sahri et al., "A highly integrated 32-SOA gates optoelectronic module suitable for IP multi-terabit optical packet routers", *Optical Fiber Communication Conference and Exhibit, OFC 2001* Volume 4, pp. PD32-1 - PD32-3, 2001.