Optical switching nodes: architectures and performance

(Article begins on next page)

29 April 2024

Optical switching nodes: architectures and performance

Andrea Bianco *, Fabio Neri *, Chiara Piglione †

* Dipartimento di Elettronica, Politecnico di Torino, Italy, Email: {firstname.lastname}@polito.it
† DCBU Modelling and Simulation Team, CISCO Systems, San José, USA Email: {cpiglion}@cisco.com

Abstract—We propose and describe four optical buffer-less switching node architectures studied in the European Network of Excellence e-Photon/ONe. We consider two possible switching scenarios: data at inputs may require either an output fiber (fiber to fiber switching) or an output fiber and a specific wavelength (wavelength to wavelength switching). We first describe the four architectures; then, we propose both heuristic and optimal algorithms to solve contention among data at inputs, and study their loss performance. Although the modeling refers to synchronous fixed-size data units, most of the observations and results hold also when considering the proposed architectures as optical cross-connects for circuit-mode operations.

I. INTRODUCTION

Several optical cross-connect architectures were proposed and compared in the past, due to the significant interest in this technology for future high-capacity infrastructures [1]–[4]. New architectures, based on the broadcast and select paradigm, were recently proposed [5], [6], and some more are under study in the framework of the e-Photon/ONe European Network of Excellence [7]. In this paper, we present and describe these architectures; moreover, we analyze their performance, and we propose optimal, when possible, and simple, thus implementable, control algorithms to determine how to connect inputs to outputs according to data requirements.

Even if originally conceived as optical cross-connects (OXC) for circuit-mode operation, we envision that the proposed architectures may be used in a high-speed optical packet/burst switch as switching matrices, either in a completely buffer-less optical switch or in an electro-optical switch where data are buffered in electronics. Thus, we propose a migration scenario for four modular OXCs with increasing functionalities. In all cases the switching granularity can be either a single wavelength or a group of wavelengths provided that the corresponding technology is available.

We examine loss performance by both analysis and simulation. We assume synchronous and slotted operation, i.e., fixed-size data arrive at input fibers synchronously. However, the presented results have a more general value and are not limited to the specific fixed size packet switching scenario considered when performing the analysis. More precisely, we define the blocking properties of the proposed architectures using the same jargon of multi-stage switching architectures. Define as admissible traffic a traffic pattern that could be transferred through any strictly non blocking architecture (e.g. a crossbar); an OXC architecture will be classified as (i) blocking, if it is not able to transfer an admissible traffic pattern, (ii) strictly non blocking, if it can transfer any admissible traffic pattern regardless of the order in which the control algorithm chooses data to be transferred, and (iii) rearrangeable non blocking, if any admissible traffic pattern can be transferred but only if data are selected in a particular order. Blocking properties depend on the switching technique considered; fiber to fiber (F2F) switching implies that data received at inputs request only an output fiber, whereas in wavelength to wavelength (W2W) switching both an output fiber and an output wavelength (in general different from the wavelength where data are received at inputs) are requested.

As a direct consequence of the above definitions, loss performance depend on the control algorithm used to configure the OXC. For this reason, we discuss optimal control algorithms, i.e., algorithms that maximize the transferred traffic, and assess their performance by analysis; we also propose simple heuristics based on round-robin schemes and assess their performance by simulation, for uniform admissible and Bernoulli traffic.

II. OXC ARCHITECTURES

We focus on an OXC with N input and output fibers, and we assume that the same number M of wavelengths is available in each fiber. Thus, each OXC can be seen as an $MN \times MN$ buffer-less switch.

The studied architectures, named V1, V2, V3 and V4 in the paper, are depicted in Fig. 1 for $N = 2$ and $M = 4$. All architectures are of the broadcast and select type; switching is achieved by blocking the proper wavelength through the Wavelength Selector (WS) device. The WS consists of two gratings Mux/Demux in tandem, separated by any optical device that is able to operate as a "shutter" (on/off gating), named Switching Point (SP) in the remainder of the paper. The SP device is typically implemented as a Semiconductor Optical Amplifier (SOA), but free-space technology, e.g. Micro Electro Mechanical Systems (MEMS) can also be adopted.

The first architecture, named V1, was initially proposed in [5] and then further extended in [6]. The principle of operation is the following: at each node input, after optical amplification by means of an Erbium Doped Fiber Amplifier (EDFA), a power coupler is used to generate multiple copies of the multi-wavelength bundle of channels entering the node from this input. The N copies for each input fiber are directed to a group of N WSs. At the output, the WSs are interconnected by means of a ($N : 1$) power combiner, in such a way that
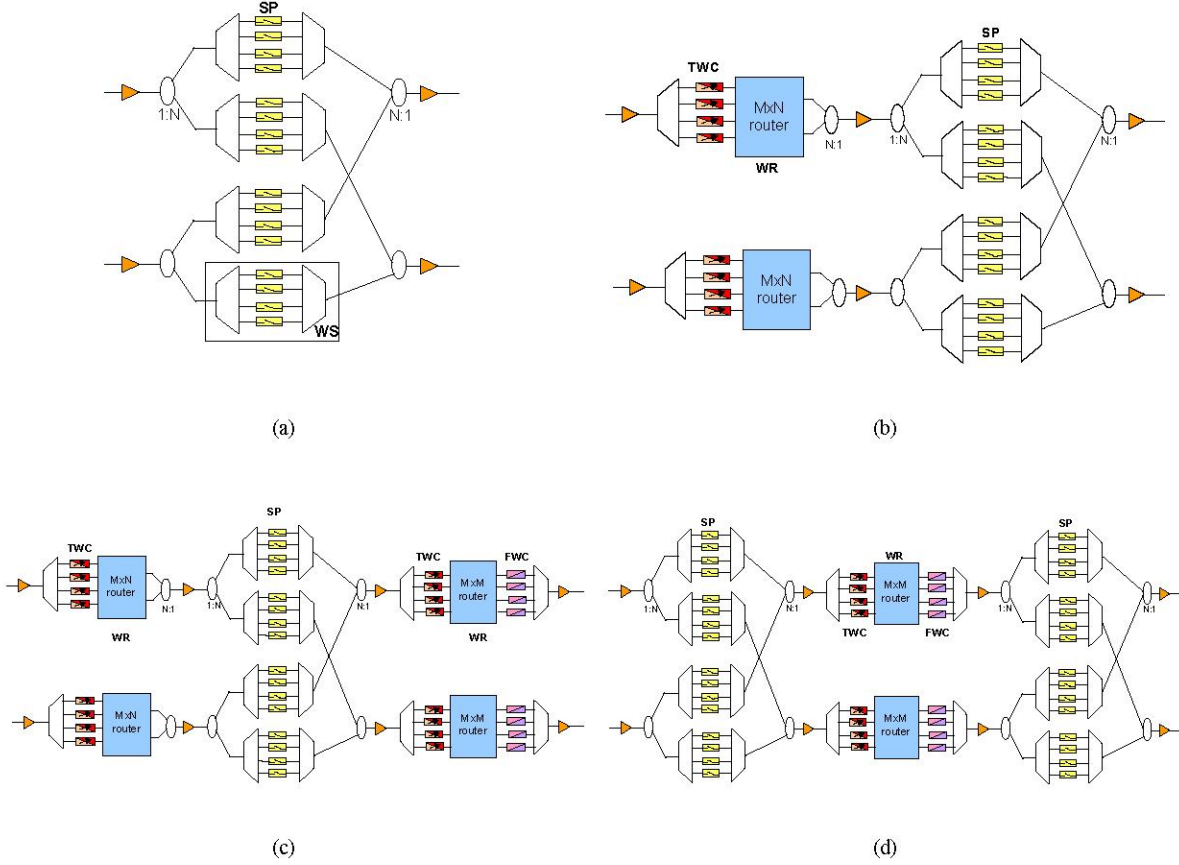
(a)

(b)

(c)

(d)

Fig. 1.   N=2 and M=4(a): V1 architecture; (b): V2 architecture; (c): V3 architecture; (d): V4 architecture;

only one WS from the same input fiber might be coupled to the same output power combiner. No wavelength conversion is available in V1 architecture; as such, there is no possibility of choosing wavelengths at outputs, and W2W switching cannot be supported. Moreover, it is not possible to transfer two data at different inputs on the same wavelength to the same output; thus, the architecture is blocking for F2F switching.

To avoid the blocking behavior of V1 architecture, V2 architecture (see Fig.1(b)) provides wavelength conversion capability at inputs; the number of components required is obviously increased with respect to V1. In the new structure, a new stage is added to V1, demonstrating a high degree of modularity of the entire structure. The new stage is composed by an array of tunable wavelength converters (TWCs) followed by an $M \times N$ Wavelength Router (WR), that can be realized by using a single Arrayed Waveguide Grating (AWG). V2 architecture is rearrangeable non blocking for F2F switching; however, V2 is blocking for W2W switching, since it is impossible to transfer two data at the same input willing to reach two different outputs on the same wavelength, due to the coupling stage at inputs after the router.

To overcome the main limitation of V2 architecture, the inability to fully support W2W switching, there are two possibilities. A third stage must be added, leading either to a wavelength-space-wavelength ($\lambda - S - \lambda$) structure, named V3

architecture (Fig.1(c)), or space-wavelength-space ($S - \lambda - S$) structure, named V4 architecture (Fig.1(d)). V3 architecture adds a wavelength conversion stage at outputs to permit fully rearrangeable non blocking behavior for both F2F and W2W switching. Similar properties hold for V4 architecture, which requires more switching points but less wavelength converters.

Note that the logical block encompassing the $M \times N$ WR and the ($N : 1$) coupler in the first stage of V2 and V3, and the logical block encompassing the $M \times M$ WR, the fixed wavelength converters (FWCs) and the Mux (third stage in V3 and second stage in V4), are logically equivalent to, and could be replaced by an ($M : 1$) coupler, while providing improved Optical Signal to Noise Ratio (OSNR).

III. BLOCKING PROPERTIES AND CONTROL ALGORITHMS

Define an admissible traffic pattern as a traffic that avoids input/output contention on a slot-by-slot basis. More precisely, for F2F switching, in each time slot there is no more than one data on each input wavelength, and there are no more than M data addressed to any output fiber. For W2W switching there is no more than one data on each input fiber wavelength and there is no more than one data addressed to each output wavelength.

Control algorithms can be classified as either optimal or heuristic. A control algorithms is optimal if it is able to transfer

any admissible traffic pattern with no losses. Due to lack of space, we describe only the general ideas behind the optimal algorithms, providing more details for heuristics in Sec. IV.

The proposed optimal algorithms (available for V2, V3 and V4 architectures in F2F switching and for V3 and V4 architectures in W2W switching) are based on the BvN (Birkhoff von Neumann) decomposition [9], which permits to decompose a doubly stochastic matrix into a sum of permutation matrices, i.e., matrices where all rows and columns contain at most one element equal to 1 and all other elements are 0. We create $\mathcal{R}$, a $N \times N$ matrix, representing the requests of data transfer in a given time slot: the element R_{ij} represents the number of data units willing to travel from input fiber i to output fiber j. First, the traffic is made admissible, according to either F2F or W2W switching constraints. Then, the traffic matrix $\mathcal{R}$ is completed by adding dummy requests if necessary, so that all rows and columns sum to M. The BvN algorithm is run on a normalized request matrix $\mathcal{R}^* = \mathcal{R}/M$, such that each row and column sums to 1. The outcome is equivalent to a schedule over an M slot frame or to run M sequential Maximum Size Matching (MSM) [8] on $\mathcal{R}$, one for each available wavelength. The algorithmic complexity is $O(N^{4.5})$.

This optimal solution can be used for V2 architecture in the case of F2F switching, and for V3 and V4 architectures for both F2F and W2W switching. Note that in the case of V4 architecture, for W2W switching, the request matrix is of size $M \times M$, with all rows and columns summing to N, and the algorithm is run N times.

IV. HEURISTIC CONTROL ALGORITHMS

A. V1 architecture: F2F switching

A simple and easily implementable *Round Robin* (RR) based heuristic control algorithm is enough to guarantee maximum performance for this blocking architecture. Note that this algorithm is not optimal in the sense that it allows to transfer admissible traffic, since this architecture is blocking; rather, it does not introduce any further limitation in the intrinsic architecture ability of data transfer.

In V1 architecture, the contention points are the N couplers at outputs. A RR counter is kept to indicate which is the input fiber that has to be served first in the considered time slot. At the beginning of each time slot, a set W_j of available wavelengths is associated with each output fiber j. Initially, $|W_j| = M$. Consider sequentially all wavelengths on the input fibers, starting from the ones on the fiber indicated by the RR counter. Suppose the data on the considered input fiber on wavelength λ^* is addressed to output fiber j^*; if λ^* is in W_{j^*} the data is served and λ^* is removed from W_{j^*}, otherwise, the data cannot be served and is lost. To ensure long term fairness among input fibers, at the end of the time slot, the RR counter is increased by one (modulo N). The algorithmic complexity is $O(MN)$.

B. V2 architecture: F2F switching

The heuristic associates with each input coupler/router a set of available wavelengths W_i^I (i indicates the router connected to input fiber i) and with each output coupler a set of available wavelengths W_j^O (j indicates output fiber j). A RR counter is used to determine which is the input fiber to be served first. Wavelengths are considered sequentially starting from the ones on the input fiber pointed by the RR counter. Suppose that on the considered wavelength λ on input fiber i^* there is a data addressed to output fiber j^*; convert λ into the first available wavelength $\lambda^* \in W_{i^*}^I \cap W_{j^*}^O$; if the intersection $W_{i^*}^I \cap W_{j^*}^O$ is empty, data is lost. Otherwise, remove λ^* from $W_{i^*}^I$ and $W_{j^*}^O$, close the SP connecting wavelength λ^* on input fiber i^* to output fiber j^* and transmit data. When all data have been considered (transmitted or discarded) increase the RR counter by one (modulo N). The algorithmic complexity is $O(M^2N)$.

C. V2 architecture: W2W switching

This heuristic is very similar to the case of F2F switching. With each output fiber and each router is associated a set of available wavelengths, W_j^O and W_i^I, respectively. A RR counter is used to determine which wavelength/fiber pair has to be served first. Consider sequentially all wavelength/fiber pairs: suppose that on the considered wavelength λ_i and input fiber i^*, there exists a data addressed to wavelength λ_j and output fiber j^*. If λ_j is in $W_{i^*}^I \cap W_{j^*}^O$, convert λ_i into λ_j and remove λ_j from sets $W_{i^*}^I$ and $W_{j^*}^O$; otherwise, discard the data. If successful, close the SP connecting wavelength λ_j on input fiber i^* to output fiber j^*. At the end of the time slot increase the RR counter by one (modulo NM). The algorithmic complexity is $O(MN)$.

D. V3 architecture: F2F switching

Under F2F switching, V3 architecture behaves like V2; thus, the same control algorithm can be adopted. However, the tunable converter in the third stage, which improves performance for W2W switching only, must be controlled. Suppose that the fixed converter at the i-th output fiber of each $M \times M$ router converts the incoming wavelength to wavelength i, and that each fixed converter operates on a single input wavelength at each time slot. Since no request on the output wavelength is associated to data at inputs, we may tune all the TWCs to the same wavelength, so that all data exit from each $M \times M$ router on different outlets; as a consequence, they will be converted to reach the desired output fiber on different wavelengths. The algorithmic complexity is $O(M^2N)$.

E. V3 architecture: W2W switching

In this case, we do not need to use the first stage tunable converters to directly convert the input wavelength to the desired output wavelength; indeed we can exploit input converters to pass without contention through the first two stages, and use third stage converters to tune to the desired wavelength. Again, we have to take control of the $M \times M$ router; the wavelengths at the inputs of each router should be selected (using the TWCs) so that data will go out on the output where they will be tuned to the proper wavelength by the FWCs. Thus, the same heuristic described for V2 architecture under F2F switching can be used. The algorithmic complexity is $O(M^2N)$.

F. V4 architecture: F2F and W2W switching

Since V4 architecture is an extension of V2 architecture, with an additional first stage of switching points, we exploit the same heuristics presented for V2. However, the additional switching stage allows to move each data to any router. As a consequence, whereas in V2 data can search for an available path only starting from the original fiber (and router) in which the data was received, in V4 all N paths (routers) can be searched for with a proper setting of the input switching stage. Complexities are $O(M^2N^2)$ and $O(MN^2)$ respectively.

V. DATA LOSS PROBABILITY ANALYSIS

We present data loss analysis for the proposed architectures

A. F2F switching

Data transmitted from an input fiber-wavelength reach the proper output fiber if there are no contentions on the same wavelength at the addressed output fiber (since there are no tunable elements). This means that V1 is blocking, i.e., it is not able to sustain admissible traffic. Indeed, V1 is equivalent to M $N \times N$ independent crossbars, one for each wavelength, which implies blocking behavior, since data are forced to use the crossbar corresponding to the wavelength over which they reached input ports.

Let us now evaluate the number of lost data under Bernoulli traffic for an input load per channel $\rho = 1$ (at each time slot, exactly M data arrive on each input fiber). Note that we use a different modeling approach with respect to the traditional Bernoulli-based analysis. This approach is less intuitive, but allows us to compute also the number of admissible traffic patterns among all the patterns that are generated according to the uniform Bernoulli arrival scenario.

Let L be the average number of lost data, and $P(loss \mid 1)$ the loss probability, when $\rho = 1$. Obviously:

$$P(loss \mid \rho = 1) = \frac{L}{MN} \qquad (1)$$

Since V1 does not include tunable elements, losses occur when there is more than one data addressed to the same output fiber on the same wavelength Thus, the loss probability is independent of M, and the computation of $P(loss \mid 1)$ can be performed by setting $M = 1$. In particular we evaluate L_1, i.e., the average number of data lost when setting $M = 1$; note that $L = ML_1$.

If only one wavelength on each input fiber is available, data are lost every time there is more than one data addressed to the same output fiber. Thus, if x data are addressed to a given fiber, the number of lost data is $\mathcal{R}_1(x)$, where $\mathcal{R}_1(x)$, the ramp function translated to 1, is defined as follow:

$$\mathcal{R}_1(x) = \begin{cases} x - 1 & \text{iff } x \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

Moreover, since $\rho = 1$, the sum of the number of data addressed to all output fibers is exactly N. The number of data addressed to each output fiber can be evaluated by computing the partitions[1] set $\mathcal{P}_N(N)$ of N over N parts. Each part of the partition indicates the number of data addressed to the corresponding output fiber; i.e., if $\phi_i \in \mathcal{P}_N(N), \phi_i = \{p_{i1}, p_{i2}, \ldots, p_{iN}\}$ then there are p_{i1} data addressed to output fiber 1, p_{i2} to output fiber 2, and so on. Once the number of data addressed to each output fiber is known, we need to determine at which input fiber they are generated. Given a partition ϕ_i, there are exactly:

$$\mathcal{N}_F(\phi_i) = \frac{N!}{\prod_{l=1}^{N} p_{il}!}$$

different ways to generate them at the N input fibers.

We now need to compute the number of partitions equivalent, i.e., with the same summands but in a different order, to each partition in set $\mathcal{P}_N(N)$, which comprises only non-equivalent partitions; the number of equivalent partitions to $\phi_i \in \mathcal{P}_N(N)$ is given by:

$$\mathcal{N}_E(\phi_i) = \frac{N!}{\prod_{l=0}^{N} \theta(l, \phi_i)!}$$

where $\theta(l, \phi_i)$ is the number of occurrence of l in ϕ_i; i.e., the number of parts of partition ϕ_i that are equal to l.

Finally, we can compute the average number of lost data, under Bernoulli input traffic when $M = 1$, as:

$$L_1 = \frac{1}{N^N} \sum_{\phi_i \in \mathcal{P}_N(N)} \left[\mathcal{N}_E(\phi_i)\mathcal{N}_F(\phi_i) \sum_{n=1}^{N} \mathcal{R}_1(p_{in}) \right]$$

Performance depends only on the number of input/output fibers and not on the number of wavelengths, as expected.

The introduction of the TWCs in V2, V3 and V4 architectures, makes these architectures rearrangeable non blocking when F2F switching is considered. Using the above notation, the average number of lost data and the loss probability when $\rho = 1$ for Bernoulli arrivals are evaluated. Losses occur every time there are more than M data addressed to the same output fiber. Thus, if x data will to reach a given fiber, the number of lost data is given by $\mathcal{R}_M(x)$, where $\mathcal{R}_M(x)$ is the ramp function translated to M.

Since the input load is 1, then the sum of the number of data addressed to all output fibers is exactly MN; we compute the number of data addressed to each output fiber evaluating the partition set in of MN over N parts $\mathcal{P}_N(MN)$. Then we compute at which input fiber wavelength they are generated. Given a partition ϕ_i, there are exactly:

$$\mathcal{N}_W(\phi_i) = \frac{(MN)!}{\prod_{l=1}^{N} p_{il}!}$$

different ways to generate data at the MN input wavelengths.

The number of equivalent partitions of $\phi_i \in \mathcal{P}_N(MN)$ is:

$$\mathcal{N}_E(\phi_i) = \frac{N!}{\prod_{l=0}^{MN} \theta(l, \phi_i)!}$$

[1] A partition of a positive integer N is a way of writing N as a sum of positive integers. Two sums which only differ in the order of their summands are considered to be the same partition. A summand in a partition is also called a part.

The average number of lost data, under Bernoulli traffic, is:

$$L = \frac{1}{N^{MN}} \sum_{\phi_i \in \mathcal{P}_N(MN)} \left[\mathcal{N}_W(\phi_i) \mathcal{N}_L(\phi_i) \sum_{n=1}^{N} \mathcal{R}_M(p_{in}) \right]$$

B. W2W switching

Since V2 architecture is blocking for W2W switching, we do not provide any analysis in this case. The analysis for V3 and V4 architectures, for W2W switching, is similar to the one presented in the Sect. V-A and the same formulation can be used with slight modifications. We must substitute $\mathcal{R}_M(x)$ with $\mathcal{R}_1(x)$, since losses occur whenever there is more than one data addressed to the same output wavelength/fiber. Also, $\mathcal{P}_N(MN)$ becomes $\mathcal{P}_{MN}(MN)$; in this case, each part of partition $\phi_i \in \mathcal{P}_{MN}(MN)$ indicates the number of data addressed to the corresponding output wavelength/fiber; i.e., if $\phi_i \in \mathcal{P}_{MN}(MN), \phi_i = \{p_{i1}, p_{i2}, \ldots, p_{iMN}\}$ there are p_{i1} data addressed to output wavelength/fiber 1, p_{i2} to output wavelength/fiber 2, and so on. As a consequence,

$$\mathcal{N}_W(\phi_i) = \frac{(MN)!}{\prod_{l=1}^{MN} p_{il}!} \quad \mathcal{N}_L(\phi_i) = \frac{(MN)!}{\prod_{l=0}^{MN} \theta(l, \phi_i)!}$$

Finally,

$$L = \frac{1}{(MN)^{MN}} \sum_{\phi_i \in \mathcal{P}_{MN}(MN)} \left[\mathcal{N}_W(\phi_i) \mathcal{N}_L(\phi_i) \sum_{n=1}^{MN} \mathcal{R}_1(p_{in}) \right]$$

VI. PERFORMANCE ANALYSIS BY SIMULATION

Performance results are obtained by simulation, for Bernoulli and admissible traffic when using heuristic algorithms to control the architectures, under uniform traffic pattern. At each time slot, for every wavelength at every input, data are generated with probability ρ. Destinations are randomly chosen among all outputs for Bernoulli traffic, and among free outputs for admissible traffic (the destination set can always be represented as input permutation). For Bernoulli traffic, which in general is not admissible, no major differences are observed when running heuristics (simulation) with respect to optimal algorithms (analysis); thus, only heuristic algorithms performance are reported in the plots. Optimal algorithms, instead, guarantee zero losses for admissible traffic.

In Fig. 2, the loss probability for V1 architecture is shown as a function of the input load for different numbers of input/output fibers. Performance improves when the number of input/output fibers decreases and does not depend on M, which was set to $M = 4$ in simulation. Loss probabilities are fairly high, as expected, since no buffering is available.

In Fig 3, we report loss probability for V2, V3 and V4 architectures for F2F switching when the heuristic algorithms are adopted. Under F2F switching, V2 and V3 are equivalent and are plotted together. Besides Bernoulli traffic, we report also results for admissible traffic. First, note that: (i) V2, V3 and V4 provide much lower loss probability than V1; (ii) performance are more sensitive to the number of wavelengths
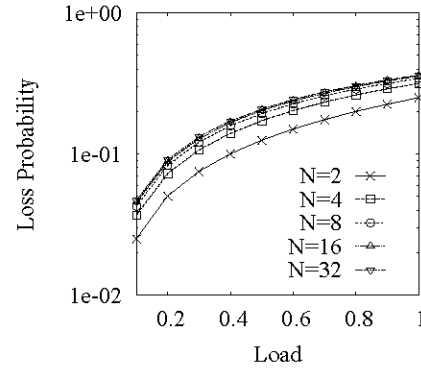

Fig. 2. V1 architecture: loss probability under Bernoulli traffic (M=4) for F2F switching.

rather than to the number of fibers, particularly for V2 and V3; this demonstrates that wavelength diversity is beneficial in contention resolution for F2F switching. Second, admissible traffic is obviously easier to deal with and architectures show smaller loss probability. Third, differences among architectures are marginal for Bernoulli traffic. Finally, under admissible traffic, V4 architecture performs much better than V2 and V3; differences are increasing for increasing values of the number of wavelengths M and of the number of fibers N. Note that V4 gains performance also for increasing N, as more input/output paths become available, so that space diversity, in addition to wavelength diversity, can be used to solve contentions.

Fig. 4 reports loss probability for V2, V3 and V4 architectures under W2W switching for admissible traffic. No differences are visible for Bernoulli traffic (not reported). For admissible traffic, V2 performs worse; V4 performs best when increasing the number of input fibers, whereas V3 benefits from an increase in the number of wavelengths per fiber. This is related to the peculiar characteristics of V3 and V4 architectures, which respectively exploit larger tunability and more space diversity to solve contentions. Moreover, as expected, performance degrade with respect to F2F switching, due to the additional constraint of W2W switching.

VII. CONCLUSION

We have described and analyzed four different OXC architectures. Heuristic control algorithm showed very good performance under Bernoulli traffic if compared to those observed when using optimal algorithms. When analyzing loss performance, under Bernoulli traffic differences among architectures tend to vanish. The admissible traffic pattern highlights the properties of the various architectures: V4 architecture provides best loss performance especially for large number of fibers. Only for W2W switching and a large number of wavelengths, V3 architectures behaves best.
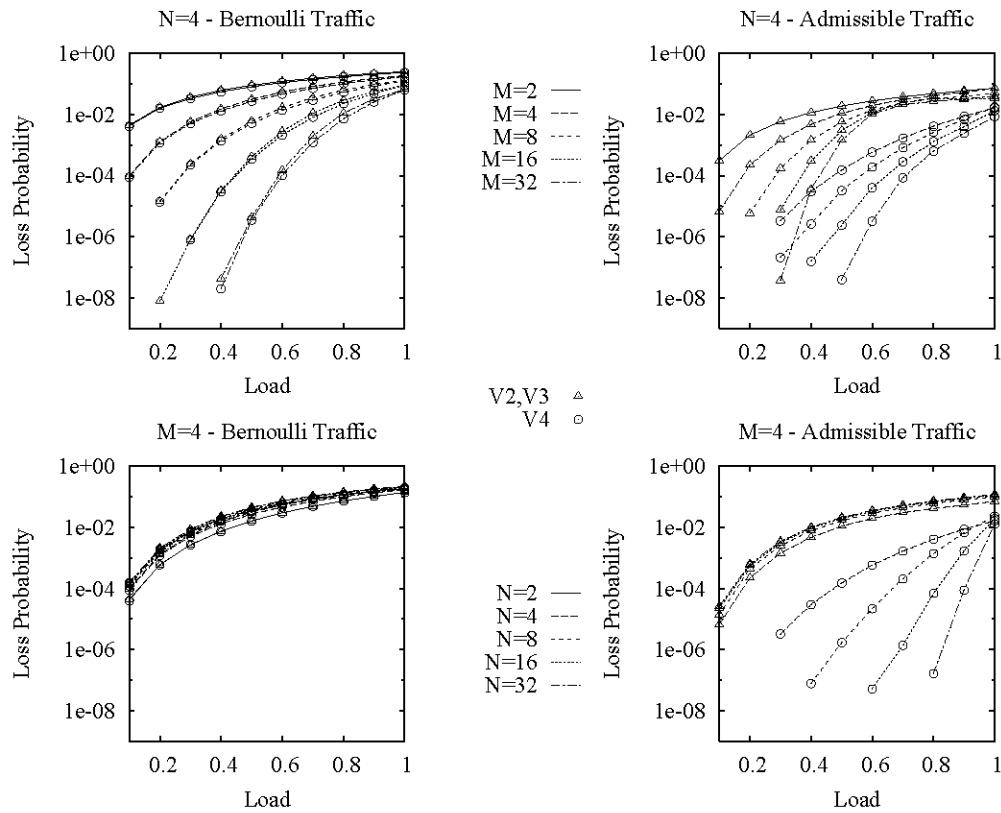
Fig. 3. V2, V3 and V4 architectures: loss probability under Bernoulli and admissible traffic for F2F switching.
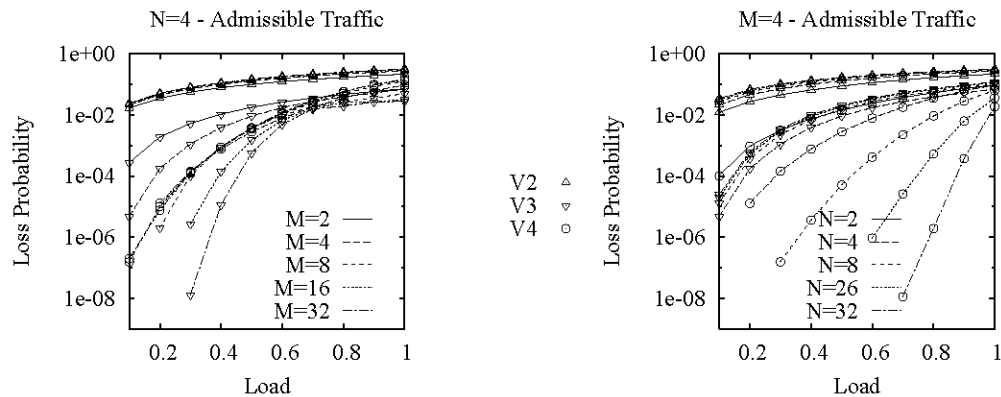

Fig. 4. V2, V3 and V4 architectures: loss probability under admissible traffic for W2W switching.

REFERENCES

[1] A.Jourdan, F.Masetti, M.Garnot, G.Soulage, M.Sotom, "Design and implementation of a fully reconfigurable all-optical cross-connect for high capacity multi-wavelength transport networks", IEEE/OSA Journal of Lightwave Technology, v.14, n.6, pp.1198-1206, June 1996

[2] M.Koga et al, "Design and performance of an optical path cross-connect system based on the wavelength path concept", IEEE/OSA Journal of Lightwave Technology, v.14, n.6, pp.1106-1118, June 1996

[3] S.Okamoto, A. Watanabe, K.Sato, "Optical path cross-connect node architecture for photonic transport networks", IEEE/OSA Journal of Lightwave Technology, v.14, n.6, pp.1410-1422, June 1996

[4] E.Iannone, R.Sabella, "A comparison among different cross-connect architectures", IEEE/OSA Journal of Lightwave Technology, v.14, n.10, pp.2184-2196, October 1996

[5] A.Stavdas, H.Avramopoulos, E.Protonotarios, J.E.Midwinter, "An OXC architecture suitable for high density WDM wavelength routed networks", Photonic Network Communications, v.1,n.1, pp.77-88, 1999

[6] A.Stavdas, "Architectures, Technology and Strategy for a Gracefully Evolving Optical Packet Switching Networks", SPIE Optical Networks Magazine, v.4, n.3, pp.92-107, 2003

[7] http://www.e-photon-one.org/

[8] Tarjan R.E. Data Structures and Network Algorithms. Society for Industrial and Applied Mathematics. Pennsylvania, November 1983.

[9] C.S.Chang, W.J.Chen, H.Huang, "Birkhoff von Neumann input buffered crossbar switches," IEEE INFOCOM'00, Tel Aviv, Israel, March 2000.