

Capacity/Cost Tradeoffs in Optical Switching Fabrics for Terabit Packet Switches

D. Cuda, R. Gaudino, G.A. Gavilanes, F. Neri
Politecnico di Torino
C.so Duca degli Abruzzi 24
10129 Torino, Italy

G. Maier
Politecnico di Milano
Piazza Leonardo da Vinci 32
20133 Milan, Italy

C. Raffaelli, M. Savi
Universita di Bologna
Viale Risorgimento 2
40136 Bologna, Italy

Abstract—Future Internet packet switches and routers will have to overcome intrinsic limitations of current electronic switching fabrics, related to the extremely high amount and density of the information to be internally processed, and the consequent very large number of interconnections, and very high power consumption and dissipation. Motivated by this observation, we consider different architectures of optical switching fabrics, capable to interconnect the linecards of a packet switch with an overall switching capacity in the order of Tb/s, using both Wavelength Division Multiplexing (WDM) and space diversity. The physical-layer feasibility and the cost of the considered architectures are studied and compared using realistic models taken from commercially available optoelectronic devices.

I. INTRODUCTION

The continuous evolution of high-capacity packet switches and routers is today getting close to fundamental physical limits of electronic devices, mostly in terms of maximum clock rate, maximum number of gates inside a single core, power density and dissipation. A lively debate on how to overcome these limits is ongoing in the research community. Although high-capacity switches with aggregate capacities of few Tb/s are today commercially available, each new generation shows increasing component complexity and needs to dissipate more power than the previous one. The current architectural trend is to separate the switching fabric from the linecards and to employ “optical wires” to interconnect them. This solution results in a large footprint (current solutions are often multi-rack), poses serious reliability issues because of the large number of active devices, and it is extremely power-hungry.

Both academic and industrial research groups are today considering optical technologies to realize intra-switch interconnects, which promise better scalability to higher capacities, increased reliability, reduced footprint and lower power consumption [1], [2]. In this paper, we consider three optical architectures belonging to the well-known family usually referred to as “tunable transmitter, fixed receiver” (TTx-FRx), which have been deeply investigated in the past [3]. In particular, we consider optical interconnection architectures based on the use of broadcast-and-select or wavelength-routing techniques to implement packet switching through an optical fabric where both Wavelength Division Multiplexing (WDM) and space multiplexing are used. The novelty of this paper lies in the fact that we study these architectures not only from an architectural viewpoint, but we mainly consider the physical-layer feasibility, the scalability issues and the cost related to these solutions. An experimental prototype of a variant of these

architectures is being developed at the PhotonLab and LIPAR laboratories of Politecnico di Torino [4].

Many optical switching experiments published in the last 10-15 years have used advanced optical processing techniques such as wavelength conversion or 3R regeneration and have been often successfully demonstrated in a laboratory environment. However, they are far from being commercially feasible, since they require optical components which are still either in their infancy or, simply, too expensive. Instead, the optical architectures considered in this paper require optical components that are commercially available nowadays. Fast-tunable lasers with nanosecond switching times are probably the only significant exception, since they have not yet a real commercial availability, even though their feasibility has already been demonstrated in many experimental projects [1], and the first commercial products start being offered.

This paper is organized as follows, Sect. II provides the general framework and presents the three considered fabric architectures. Sect. III discusses the feasibility, the scalability and a cost model of the considered optical fabrics. Sect. IV presents a selection of performance results obtained with our approach, while some final comments are given in Sect. V.

II. OPTICAL SWITCHING FABRICS

The general architecture of the optical switching fabric considered in this paper is shown in Fig. 1: a set of N input linecards send packets to an optical interconnection structure that provides connectivity towards the N output linecards using both WDM and space multiplexing techniques. All linecards are operated towards the switching fabric at the same, typically large, bitrate, and can serve a number of input/output ports/lines possibly operating at different bitrates and according to different technologies. The optical switching fabric is organized in multiple switching planes, to which a subset of output linecards are connected. Fast optical switches (or in general plane distribution subsystems) allow input linecards to select the plane leading to the desired output linecard on a packet-per-packet basis. Within each plane, wavelength routing techniques select the proper output. Thus packet switching is controlled at each input linecard by means of a fast tunable laser (i.e., in the wavelength domain) and, possibly, a fast optical switch (i.e., in the space domain). Each linecard is equipped with one tunable transmitter (TTx) and one fixed-wavelength burst mode receiver (BMR) operating at the data

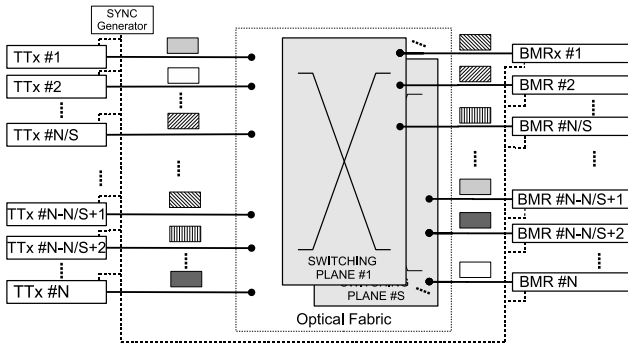


Fig. 1. Multiplane optical fabric architecture

rate of a single WDM channel. Burst-mode operation is required on a packet-by-packet basis. For simplicity, we assume that all the considered architectures have a synchronous and time-slotted behavior, as reported in [5] and [6], and that packets transmissions are scheduled so that at most one packet is sent to each receiver on a time slot (i.e., contentions are solved at transmitters).

The lasers tuning range, i.e., the number of wavelengths a transmitter is required to tune to, is a practical limiting factor. Even for prototypes, the maximum tuning range for tunable lasers is in the order of a few tens of wavelengths [8]. As a result, the wavelength “dimension” alone could not ensure input/output connectivity when the number N of linecards is large. Multiple switching planes, that is, space diversity dimension, was introduced to overcome this limitation. By doing so, since the same wavelengths can be reused on each switching plane, if S is the number of switching planes, a wavelength tunability equal to N/S (instead of N) is required.

In the three considered architectures, shown in Figs. 2–4, transmitters reach the different S planes by means of proper optical *distribution* stages that actually differentiate each of the three architectures. At the output of the switching plane, an $S : 1$ splitter collects the packets, which are then distributed to the N/S receivers. These fabric architectures are designed so that the number of couplers and other devices that packets have to cross is the same for all input/output paths. The EDFA WDM amplification stages can thus add equal gain to all wavelength channels since all packets arrive with the same power level. On each plane, signals are amplified by an EDFA, demultiplexed and finally received.

A. Wavelength-Selective “V” Architecture (WSV)

This architecture was originally proposed for optical packet switched WDM networks inside the e-Photon/ONE project [7] (the “V” name stems from an internal denomination used in the e-Photon/ONE community, that studied several variations of this architecture, called V_x). This optical switching fabric, presented in Fig. 2, connects N input/output linecards by means of broadcast-and-select stages. Sets of N/S TTx’s are multiplexed in a WDM signal by means of a $N/S : 1$ coupler; then this signal is split into S copies by means of an $1 : S$ splitter. Each copy is interfaced to a different switching plane by means of Wavelength Selectors (WSs), composed by a demux/mux pair separated by an array of $\frac{N}{S}$ SOA gates, which are responsible for both plane and wavelength selection.

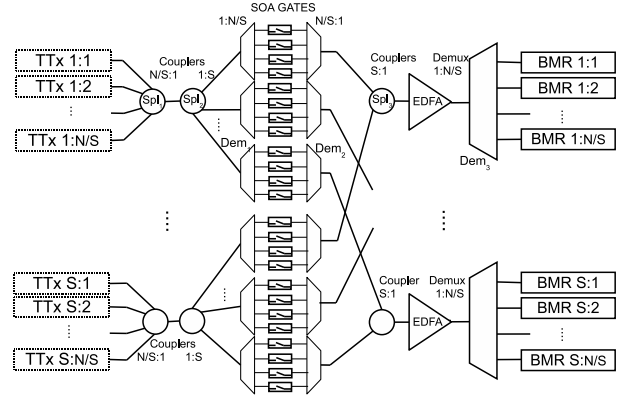


Fig. 2. Wavelength-Selective V (WSV) architecture

The WSV is a blocking architecture. Since within a group of input linecards all the transmitters must use a different wavelength, it is not possible for transmitters located in the same input group to transmit to output ports located on different switching planes that receive on the same wavelength. The basic WSV architecture can be modified to obtain a non-blocking behavior [7], but we do not further discuss this issue in this paper.

B. Wavelength-Routing-Space (WRS)

In this case the wavelength routing property of Arrayed Waveguide Gratings (AWGs), main component of the distribution stage, is exploited to perform both plane and destination selection: no space switch is necessary. Each collecting stage gathers then all the packets for a specific plane.

The WRS fabric can either exploit the cyclic behavior of AWGs or not. This property states that several homologous wavelengths belonging to different Free Spectral Ranges (FSR) are identically routed to the same AWG output. This situation is depicted in Fig. 3, which shows an instance of WRS architecture with $N = 9$ and $S = 3$. Note that each receiver is associated with a set of N/S different wavelengths (for receivers this is not a real limitation because the bandwidth of photodiodes is actually very large). By exploiting the cyclic property of the AWGs we can prevent that linecards reuse the same wavelength in different parts of the fabric. By doing so, the total TTx tunability range required has to be equal to N . The advantage is that no coherent crosstalk (see [5]) is introduced in the optical fabric, and only out-of-band crosstalk is present, which can be shown to introduce negligible penalties. Even though this solution can be almost optimal from the physical point of view, the AWGs must exhibit an almost identical transfer function over N/S FSRs, and the EDFA amplifying bandwidth has to be significantly larger. In this configuration, this architecture is called WRS-ZC, that stands for zero-crosstalk configuration. If instead the AWG cyclic property is not exploited, and the system operation is limited to a single FSR (all TTx’s are identical), some in-band crosstalk is introduced, which can severely limit scalability [5]. However, it is possible to show that these impairments can be limited or fully eliminated by proper packet scheduling algorithms, which avoid using the same wavelengths at different fabric inputs at the same time.

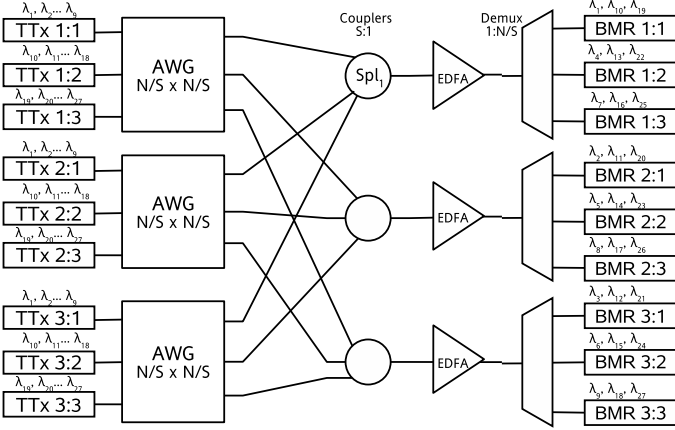


Fig. 3. Wavelength-Routing-Space (WRS) architecture

C. Multiplane Couple-Amplify-Demultiplex (MCAD)

In the MCAD fabric, depicted in Fig. 4, the distribution stage is implemented by splitting transmitter signal in S copies by a $1 : S$ splitter, and then sending signals to SOA gates. A given input wanting to transmit a packet to a given output must first select the destination plane by turning off all the SOA gates but the one associated with the destination plane and, then, use wavelength tunability to reach the desired output port. The following coupler stages are used to combine packets for each switching plane; to minimize excess losses in optical couplers, coupling stages are organized in two sections: the first is a distribution stage used by inputs to reach all planes, while the second is a collecting stage to gather packets for each plane. Each plane combines at most N/S packets coming from the N input linecards.

III. PHYSICAL IMPAIRMENTS AND SCALABILITY

A. Power Penalties

The considered optical fabrics do not include any signal regeneration but only pure linear optical amplification. Using a common terminology [9], we have 1R regeneration of the signal inside the optical fabric (no 2R and 3R regeneration). On a first attempt to assess the scalability of optical architectures, we observed that modeling fixed values for components losses yields unrealistic scalability results e.g., “infinite” port count scalability for fabrics whose total power penalties do not depend on the number of ports (N) and the number of planes (S). We adopted a more accurate second-order model capable of capturing other important effects such as insertion losses, excess losses, channel uniformity, polarization dependency and crosstalk, which were translated into power penalties and related to the fabric’s port and plane counts. This modeling is described in [5] and [6] (it is not further detailed here for space limitations), where the information quantifying these impairments has been collected from data sheets of commercially available components. According to the notation used in [5] and [6], hereafter $L_{Dem}(N)$, $L_{Spl}(N)$ and $L_{AWG}(N)$ will denote losses introduced by muxes/demuxes, couplers/splitters and AWGs of N ports, respectively. However, in this paper a more realistic model of TTxs was introduced, with respect to [5] and [6], to account for their noise contributions.

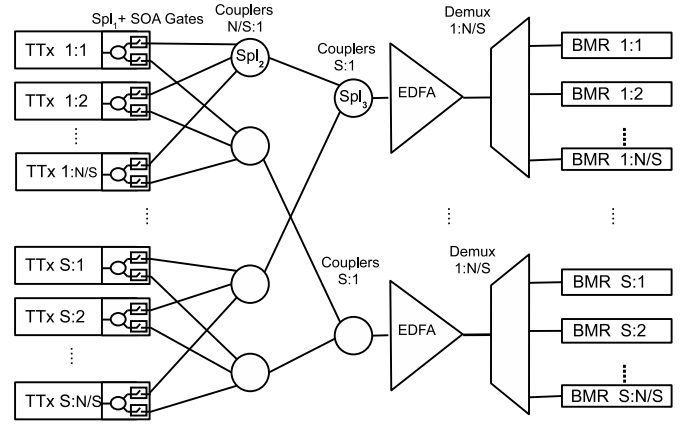


Fig. 4. Multiplane Couple-Amplify-Demultiplex (MCAD) architecture

B. Scalability analysis

By using the physical impairments mentioned above we conducted a scalability and feasibility study of the fabrics described in Sect. II. Firstly, we need to introduce some key parameters specific to the components used.

For the transmitter, an average transmitted power P_{TX} of 3 dBm is assumed. Generally, a typical tunable laser peak output power is of the order of +10 dBm, but we need to consider 6-7 dB of equivalent loss introduced by an external modulator (3 dB due to on-off keying and 3-4 dB due to insertion losses). WDM tunable transmitters were modeled as optical sources characterized by a given Optical Signal to Noise Ratio $OSNR_{TX}$, corresponding to the ratio between the useful laser power and the noise power due to spontaneous emission inside the laser. Using the laser $OSNR_{TX}$, we calculate the equivalent noise power spectral density (N_{0TX}) as $OSNR_{TX} = \frac{P_{TX}}{N_{0TX}B}$. We have assumed a flat noise behavior with an optimistic $OSNR_{TX}$ value ranging from 40 to 60 dB over a 0.1 nm reference bandwidth B in Hz.

For the receiver, we assume a target BER $T_{BER} = 10^{-12}$, for which best receivers at 10 Gb/s have today a typical receiver sensitivity P_S of around -26 dBm (without optical amplification). In order to address scalability at different bitrates, we followed the analysis presented in [11] which, inferring from different commercial data sheets, proposes a sensitivity slope vs. bitrate of 13.5 dB/decade. For instance, given the aforementioned P_S at 10 Gb/s, we infer $P_S = -17.8$ dBm at 40 Gb/s.

For what regards the SOA-based space switches (used in the MCAD architecture) and gates (used in WSV), we based our analysis on the characteristics of one of the few commercially available SOA-based switches [10], as described in [6]. In the *on* state, SOAs are characterized by both a gain A_{SOA} , set to recover preceding losses and by a noise figure F_{SOA} of 9 dB. In the *off* state, SOAs show a switching Extinction Ratio (ER) of 35 dB. This finite SOA’s ER, leads to significant crosstalk contributions among different planes, thus, it imposes a limit on the maximum amount of planes that the fabric admits. In the worst case, the number of coherent crosstalk contributions in a switching plane is one for each plane; henceforth, the total crosstalk penalty on an arbitrary plane depends both on S and on the nominal switch ER (in linear units). This result is

detailed in [6], where the crosstalk is modeled by the approach proposed in [12] and the penalty is defined for optimized decision-threshold receivers and for the required T_{BER} .

All architectures present a common amplification and demultiplexing output stage. We have assumed that EDFAs operate in saturated regime; that is, EDFAs show a constant output power, which is split over N/S channels that cross it simultaneously. The nominal EDFA output power is assumed to be $P_{tot,out}^{EDFA} = 17$ dBm, which can be set by means of gain locking techniques. Let A_{EDFA} be the EDFA gain. The A_{EDFA} used in noise calculations is obtained considering the ratio among $P_{tot,out}^{EDFA}$ and $P_{tot,in}^{EDFA}$. Furthermore, we characterized EDFAs by a noise figure F_{EDFA} of 5 dB.

In our architectures, scalability is limited by both the useful signal power at the receiver's photodiode, and by the ASE noise present at the EDFA output. Firstly, received signal power must be larger than the receiver's sensitivity; thus, at each output the following must hold:

$$P_{ch,out}^{EDFA}|_{dBm} - L_{Dem}(N/S)|_{dB} - \mu \geq P_S|_{dBm} \quad (1)$$

where $P_{ch,out}^{EDFA}|_{dBm} = P_{tot,out}^{EDFA}|_{dBm} - 10 \log(N/S)|_{dB}$ is the power per channel after amplification, and μ is a 3 dB margin to consider component aging, penalties of BMR with respect to standard continuous receivers, laser misalignments, and other effects that might degrade the received signal.

Even though the sensitivity constraint is the same for all the considered architectures, the noise-related constraint is not.

Being commonly accepted, we require a minimum optical signal to noise ratio T_{OSNR} of 17 dB (defined over a bandwidth equal to the bitrate) to guarantee the target T_{BER} of 10^{-12} . Moreover, this condition must hold independently of the linecard bitrate. In order to properly deal with the multiplane setup, as discussed in [5] and [6], we account for all the optical noise sources present in the fabric. All noise contributions (from lasers and SOAs), add up over a very large bandwidth (usually called a noise "floor" on power spectrum); as a result, each WDM channel will see a portion of this accumulated noise in band, leading to a non negligible scalability limitation.

The noise spectral density $G_N(f)$ (in linear units, being f the optical frequency in Hz) was evaluated for each architecture. Now let $G_{N,out}^{Amp}(f)$ be an amplifier output noise spectral density:

$$G_{N,out}^{Amp}(f) = G_{N,in}^{Amp}(f) \times A_{Amp} + hf(A_{Amp} - 1) \times F_{Amp} \quad (2)$$

where h is the Planck constant, A_{Amp} and F_{Amp} (either $Amp = EDFA$ or $Amp = SOA$) are the amplifier gain and noise figure, respectively, and $G_{N,in}^{Amp}(f)$ is the noise power spectral density entering the amplifier. For the final EDFA (common to all architectures), $G_N^{EDFA}(f)$ is different for each architecture, and we call it $G_N^{WSV}(f)$, $G_N^{MCAD}(f)$ or $G_N^{WRS}(f)$ for the WSV, the MCAD or the WRS fabric, respectively. Due to space limitations, we only show the final expressions derived for these quantities in Table I. In the first equation, $G_N^{Dem1}(f) = \frac{N_{0TX} \times N}{L_{Sp1}(\frac{N}{S})L_{Sp2}(S)L_{Dem1}(\frac{N}{S})S}$ is the power spectral density after first demux.

The EDFA noise power per output channel $P_{N-ch,out}^{EDFA}$ can be evaluated by integrating $G_{N,out}^{EDFA}(f)$ over a channel band-

TABLE I
POWER SPECTRAL DENSITIES FOR ALL THE ARCHITECTURES

$G_N^{WSV}(f)$	$= \frac{G_N^{Dem1}(f)G_{SOA} + hf(G_{SOA} - 1)F_{SOA}}{L_{Dem2}(\frac{N}{S})L_{Sp3}(S)}$
$G_N^{MCAD}(f)$	$= \frac{N_{0TX}G_{SOA} + hf(G_{SOA} - 1)F_{SOA}L_{Sp1}(S)}{L_{Sp1}(S)L_{Sp2}(\frac{N}{S})L_{Sp3}(S)} \times \frac{N}{S}$
$G_N^{WRS}(f)$	$= N_{0TX} \times \frac{N}{L_{AWG}(\frac{N}{S})L_{Sp1}(S)}$

width equal to the bitrate; thus $P_{N-ch,out}^{EDFA} = G_{N,out}^{EDFA}(f) \times R_b$; the OSNR constraint can be expressed as:

$$P_{ch,out}^{EDFA}|_{dBm} - P_{N-ch,out}^{EDFA}|_{dBm} - \mu \geq T_{OSNR}|_{dB} \quad (3)$$

with μ having the same meaning as in Eq. 1.

C. Cost Model

In this section, a cost model for each device is defined and, then, it is used to assess the cost of the entire architecture. Since the operational bandwidth of any device is finite, a limited number of wavelengths can be in general amplified or routed. Commercially available EDFAs and AWGs are usually designed to work over one of the C, L, XL, S bands (in order of preferred deployment). The considered optical fabrics heavily rely on wavelength agility and space switching; indeed, as the number of necessary SOAs increases, the required transmitter tunability decreases for both the MCAD and WSV architectures. We took this tradeoff into account.

Component costs in our model rely on the current market situation (discrete components, no on-chip integration, no economies of scale); thus, they should be used as a rough guide to assess the relative costs of the different architectures. Our goal is to discuss which are the main factors impacting the cost of an optical fabric and which are the cost-wise optimal choices to be taken when designing such architectures.

Splitters, Muxes/Demuxes: We assume that the main factor affecting both splitter and mux/demux costs is their port count n . From market data [13], by using a Minimum Square Error (MSE) fitting method, we inferred the following cost model:

$$C_{Spl}(n) = \frac{100}{41.7} \times n^{\frac{1}{4}} \quad (4)$$

In order to simplify, but also being compliant with market prices [13], we assumed that mux/demux costs behave as the splitter ones, scaled by a factor α_{Dem}^1 , i.e.:

$$C_{Dem}(n) = \alpha_{Dem} \times C_{Spl}(n) \quad (5)$$

where $\alpha_{Dem} = 10$ in our computations.

Arrayed waveguide gratings: AWG costs depend on both the number of ports and the number of wavelengths it must be able to route, hence on the number of FSRs over which the AWG Transfer Function (TF) should be approximately flat. The AWG dependency on the number of ports follows the mux/demux law scaled by a factor $\beta_{AWG} = 2.5$. Let N_{FSR} be the number of FSRs over which the AWG TF has to be flat. In the WRS architecture, if the AWG cyclic property is exploited to avoid crosstalk (WRS-ZC), $N_{FSR} = (N/S)^2$; otherwise

¹We assumed the same behavior for the AWG cost with respect to the demux cost.

(WRS), $N_{FSR} = N/S$. For commercially available AWGs, the TF sharply decreases outside the bandwidth over which the AWG was designed; thus, the AWG TF was assumed to be approximately flat over at most 3 FSRs. If an architecture requires a wider TF, we assume that more AWGs must be used in parallel. The whole AWG cost is:

$$C_{AWG}(n, N_{FSR}) = (1 + \lfloor \frac{N_{FSR}-1}{3} \rfloor \gamma_{AWG}) \beta_{AWG} C_{Dem}(n)$$

where $\gamma_{AWG} = 1.5$; indeed, the cost of adding more AWGs in parallel is larger than simply multiplying the number of AWGs by their cost (some other devices as splitters and filters are needed).

Laser and transmitters: Many manufacturers claim that the technology to produce cheap tunable lasers is going to be available soon and that their cost is going to be less than twice the cost of a fixed laser. Thus, if $C_{f-TX} \approx 1000$ \$ [13] is the cost of a fixed transmitter at a bitrate of $R_{fix} = 1$ Gb/s, we assume that a tunable laser is $\delta_{t-TX} = 1.5$ times more expensive than a fixed laser operating at the same bitrate. Even though tunability could be available at affordable costs, the tunability range as well as the transmitter bitrate need to be considered. Regarding the bitrate, we assume a transmitter cost that is linearly proportional to the bitrate. Regarding the tunability range, we assume that fast tunable lasers will be available soon over the C, L, XL, S bands. When the required tunability exceeds the one guaranteed by a single band, an additional laser is added to the transmitter. For instance, if a tunability of 200 wavelengths is required, then two lasers (one tunable over the C band and another one tunable over the L band) are employed. Finally, the total TTx cost can be evaluated as follows:

$$C_{t-TX}(W, R_b) = \left(1 + \lfloor \frac{W-1}{W_b} \rfloor \alpha_{TX}\right) \delta_{t-TX} C_{f-TX} \times \frac{R_b}{R_{fix}}$$

where W is the number of wavelengths the transmitter has to be able to tune to, W_b is the number of wavelengths in each band, and R_b is the transmitter bitrate. For the multi-band case, we set a multiplicative factor $\alpha_{TX} = 1.3$, since the cost of integrating more than one laser in parallel is larger than simply the product of the number of integrated lasers and the cost of a single laser.

Optical Amplifiers: Optical amplification is usually implemented using either an EDFA or a SOA. SOAs are used here as on/off gates to perform fast plane selection, and they have to deal with just one wavelength at a time; hence, we consider bandwidth as a cost factor only for EDFAs. The SOA cost is therefore fixed to $C_{SOA} = 1000$ \$ (complexity and technological skills needed to manufacture a SOA should not be exceeded much by those needed for a laser). Commercially available EDFAs are today designed to operate over one of the four different optical bands; thus, if a large bandwidth is needed, additional EDFAs are used in parallel and the device cost can be expressed as:

$$C_{EDFA}(W) = \left(1 + \lfloor \frac{W-1}{W_b} \rfloor \alpha_{EDFA}\right) C_{fix}$$

where $C_{fix} = 10000$ \$ (see [13]) is the cost of a single-band EDFA, W is the number of channels to be amplified, W_b is the number of wavelengths in the corresponding optical band and $\alpha_{EDFA} = 1.3$ accounts for the integration cost.

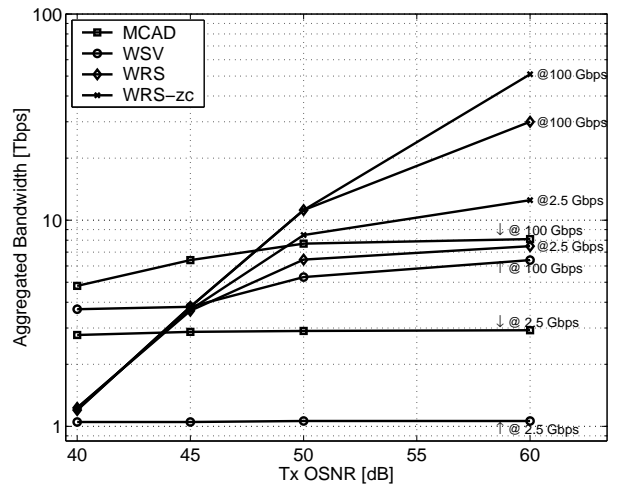


Fig. 5. Aggregate bandwidth as a function of TTx noise $OSNR_{TX}$

Finally, the overall cost of the considered switching fabrics was evaluated aggregating the relevant components cost (formulas are omitted for space limitations); taking into account both the mentioned technological limitations and the fabric dimensions (N and S) determined by the physical feasibility.

IV. RESULTS

A. Scalability Analysis

We first discuss the feasibility and scalability of the considered optical fabrics. Each architecture was evaluated in terms of its total aggregate bandwidth, which depends on R_b , S and N . Given a set of line bitrates, we evaluated the maximum achievable bandwidth (corresponding to the maximum number of linecards that can be supported). Although the number of switching planes S in principle can range from 1 to N , for the architecture configurations satisfying the feasibility limits given in Eq. 1 and Eq. 3, there is an optimum value of S for which the OSNR at the BMR is maximum. This value was obtained numerically in each case. For all architectures, the main limiting effect was observed to be the OSNR at the receiver: the maximum achievable bandwidth is mainly limited by the optical noise accumulation. Therefore, the nominal transmitter OSNR, being the earliest noise contribution in the fabric, deeply impacts the optical fabric scalability. Fig. 5 shows the maximum achievable aggregate bandwidth vs. $OSNR_{TX}$ for different bitrates. Firstly, for low bitrates (2.5 Gb/s per linecard), performance of both the WSV and the MCAD architectures are quite independent from the TTx noise, while for higher bitrates (100 Gb/s per linecard), the better is the $OSNR_{TX}$ the larger is the maximum achievable bandwidth. However, their performance dependency on noise is very weak; indeed, the number of noise floor sources is equal to N/S . Conversely, WRS and WRS-ZC performance are deeply affected by the $OSNR_{TX}$. If the transmitter noise is very low, WRS and WRS-ZC achieve the largest maximum aggregate bandwidth of several tens of Tb/s, but as the $OSNR_{TX}$ becomes closer to realistic values, performance severely falls. In these architectures the noise floor of all the N sources (instead of N/S as in the case of WSV and MCAD) accumulates as the signal propagates through the optical fabric. Note that, up to now, even technologically

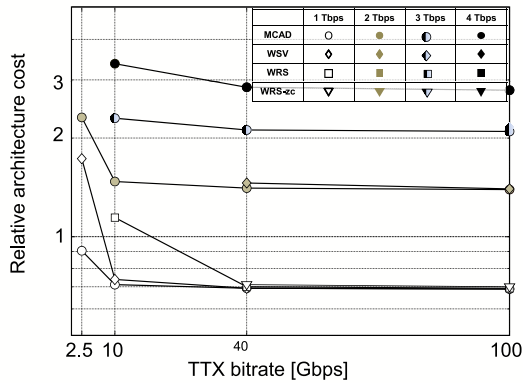


Fig. 6. Cost vs. transmitter bitrate for different aggregate bandwidth

advanced transmitters show an OSNR of at most 40 dB; thus, MCAD and WSV might be the best solutions to implement an optical fabric in the near future.

B. Cost analysis

Fig. 6 shows the relative architectures' cost vs. R_b for different aggregate bandwidths (from 1 Tb/s to 4 Tb/s) for TTX characterized by $OSNR_{TX} = 40$ dB. Different marker shapes identify the different architectures, while different filling patterns are associated with different aggregate capacities. Given the bitrate and aggregate bandwidth, cost was calculated by means of the models presented in Sec. III-C for all feasible configurations. Then those showing minimum cost for the same bitrate are plotted in Fig. 6. For the MCAD, WSV and WRS architectures, cost significantly decreases as the bitrate increases. Even though higher bitrate transmitters are more expensive (assuming cost linearly dependent on the bitrate), our calculations show that cost savings due to decreasing the number of transmitters, splitters, amplifiers and SOAs are larger than the additional cost of higher bitrate transmitters. WRS fabric costs decrease as lower tunability and lower number of AWGs is needed with higher bitrate. Consistently with the results shown in Fig. 5, for transmitters with a realistic OSNR of 40 dB, Fig. 6 shows that the MCAD (especially) and WSV can support higher aggregate bandwidths than the WRS-ZC and WRS architectures; indeed, the MCAD switching fabric achieves a switching capacity of 4 Tb/s, even though it is generally more expensive.

We focused the discussion above on relative trends of the curves in Fig. 6 because the absolute cost values are still quite high: the vertical scale was normalized to one million dollars. We must however remark that our cost analysis considered discrete components as they are commercially available today. It is reasonable to expect that economy of scale effects, technological improvements, and a larger integration of optical components will significantly cut down costs in a near future.

V. CONCLUSIONS

In this paper we considered three architectures suited for the realization of a fully optical switching fabric with a large aggregate switching capacity. These architectures are based on commercially available optical devices. To overcome the current technological limits of laser tunability and amplifiers bandwidth, space multiplicity was introduced by partitioning

the fabrics in multiple switching planes, thereby enabling wavelength reuse.

Feasibility and scalability studies were conducted to evaluate the maximum aggregate bandwidth of the considered optical fabrics, which were shown to be able to support tens of Tb/s. Moreover, being optical devices (in contrast to the electronic ones) far from their intrinsic physical limits, they can still be improved (for instance by reducing the components' noise or by increasing the devices' bandwidth), so that performance of optical fabrics can increase.

In addition to the scalability analysis, we performed simple estimations of the cost of the considered optical fabrics. Even though high costs are still a matter of fact, trends show a sublinear dependence on the aggregate switching capacity, and a limited dependence on the line bitrate (in particular it is more convenient to employ few high-bitrate linecards than several low-bitrate linecards). Moreover, several optical devices are still in the first phases of their commercial maturity, and there are still large margins to reduce costs. We think that larger on-chip integration, economy of scale and volume discount will knock down device costs.

In summary, considering the fact the electronic technologies in switching architectures are approaching intrinsic physical limits, and the fact that there is still room to reduce costs of optical fabrics, we think that optical technologies can play an important role in the realization of future generations of high-performance packet switches and routers.

ACKNOWLEDGMENT

This work was partially supported by the Italian project OSATE and by the Network of Excellence BONE ("Building the Future Optical Network in Europe"), funded by the European Commission through the 7th Framework Programme.

REFERENCES

- [1] J. Gripp, et al., "Optical switch fabrics for ultra-high-capacity IP routers", *IEEE/OSA JLT*, vol. 21, no. 11, Nov. 2003, pp. 2839-2850.
- [2] N. McKeown, "Optics inside Routers", *ECOC 2003*, Rimini, Italy, Sep. 2003.
- [3] B. Mukherjee, "WDM-Based Local Lightwave Networks – Part I: Single-Hop Systems," *IEEE Network*, vol. 6, no. 3, pp. 12-27, May 1992.
- [4] A. Antonino, R. Birke, V. De Feo, J. M. Finocchietto, R. Gaudino, A. La Porta, F. Neri, M. Petraccia, "The WONDER Testbed: Architecture and Experimental Demonstration", *ECOC 2007*, Berlin, Germany, Sept. 2007.
- [5] J. M. Finocchietto, R. Gaudino, G.A. Gavilanes Castillo, F. Neri, "Simple Optical Fabrics for Scalable Terabit Packet Switches", *ICC 2008*, Beijing, China, May 2008.
- [6] J. M. Finocchietto, R. Gaudino, G.A. Gavilanes Castillo, F. Neri, "Multi-plane Optical Fabrics for Terabit Packet Switches", *ONDM 2008*, Vilanova, Catalonia, Spain, March 2008.
- [7] A. Stadvas, et al. "New Designs for Optical Packet Switching Nodes", *D7.3: Final JP1 report*, e-Photon/ONe Project: <http://www.e-photon-one.org>.
- [8] A. Bhardwaj, et al. "Demonstration of Stable Wavelength Switching on a Fast Tunable Laser Transmitter", *IEEE Photonics Technology Letters*, Vol. 15, No. 7, July 2003.
- [9] R. Ramaswami, and K.N. Sivarajan, *Optical Networks - A Practical Perspective*, Morgan Kaufman, 1988.
- [10] Alphion, QLight I-Switch Model IS22, Advance Product Information, <http://www.alphion.com>
- [11] E. Sackinger, *Broadband Circuits for Optical Fiber Communication* New York, John Wiley & Sons, 2005
- [12] H. Takahashi, K. Oda, H. Toba, "Impact of crosstalk in an arrayed-waveguide multiplexer on NN optical interconnection", *J. of Lightwave Technology*, Vol.14, Iss.6, Jun 1996, Pages:1097-1105
- [13] Taken from www.go4fiber.com Product Datasheets