SEVENTH FRAMEWORK PROGRAMME

D25.3: “Report on Y2”

FP7-ICT-216863/PoliMI/R/PU/D25.3

Project Number: FP7-ICT-2007-1 216863

Project Title: Building the Future Optical Network in Europe (BONE)

Contractual Date of Deliverable: 30/12/10

Actual Date of Delivery:

Workpackage contributing to the Deliverable: WP 25 TP: Optical Interconnects

Nature of the Deliverable R

Dissemination level of Deliverable PU

Editors:

Partner / names
PoliMI / Guido Maier, Achille Pattavina

Abstract:

After two-year activity, WP25 has reached the end of its planned duration. This is the second deliverable of WP25 reporting actual technical content. This is also the final deliverable of the workpackage.

The first four general sections compiled by the WP leader are followed by section 5 reporting about each Joint Activity (JA) of the workpackage. This latter section has been created by merging the contributions received by the JA leaders.

Keyword list:

Optical interconnect (OI), board-to-board OI, rack-to-rack OI, shelf-to-shelf OI, on-chip OI, energy efficiency, power saving, optical backbone

Disclaimer

The information, documentation and figures available in this deliverable, is written by the BONE ("Building the Future Optical Network in Europe*) – project consortium under EC co-financing contract FP7-ICT-216863 and does not necessarily reflect the views of the European Commission*

Table of Contents

1. Executive Summary

The general purpose of WP25 Topical Project “Optical Interconnects” is to study the optical implementation of the interconnection systems inside a high-performance switching/routing system. The target is to develop the capability to design and optimize an optical backplane, finding the technology and architecture best-fitting system requirements.

After a two-year activity, WP25 has reached the end of its planned duration. This is the second deliverable of WP25 reporting actual technical content. This is also the final deliverable of the workpackage.

Eleven partners have actively collaborated to the Joint Activities (JAs) of this workpackage during the second year. This document contains the detailed description of the JAs at their final status of development.

At the beginning of the last year of WP25 (BONE Y3), eight JAs were in operation in the workpackage.

The first three JAs, dedicated to survey and requirement definition, were planned to be in operation in the first year only of the TP. Thus D25.3 does not cover JA1, JA2 and JA3 in detail (their results have been reported in D25.2). However, activity on these JAs has continued also in the last year in terms of preparatory work for three journal papers to disseminate the results obtained by the JAs. The manuscripts of these papers will be submitted for publication in the next months. For this reason we have included in this deliverable the abstracts of the papers.

JA4 was planned to be developed just in the last year of the TP, as it has been conceived to take advantage of the findings on optical technologies, architectures and requirements carried out in JA1, JA2 and JA3. This deliverable reports on the results concerning a reference architecture for the optical backplane.

The JAs from JA5 to JA8 were planned to span over the two years of the TP. All these activities have concluded their work, and the results achieved are reported in this deliverable. The reports show a summary of technical content and joint/individual publications of the involved partners. In some cases, mobility actions carried out in Y3 under the umbrella of WP25 are reported.

The only exception to this rule is JA8, dedicated to On-chip optical switching networks. The work in this activity started by overviewing optical network-on-chip architectures (reported in D25.2). After this overview we realized that the work carried out in other JAs was already covering the main topic of JA8. These activities concern utilization of microring resonators and on-chip implementation of optoelectronic switching networks, carried out in JA6 and JA7, respectively. Therefore, we chose not to include in this deliverable a specific section for JA8, in consideration of the fact that JA6 and JA7 already provide ample research results which cover the topic of on-chip optical interconnections.

2. Introduction

After two-year activity, WP25 has reached the end of its planned duration. This is the second and last deliverable of WP25 reporting actual technical content.

The next two general sections compiled by the WP leader are followed by the section reporting about each Joint Activity (JA) of the workpackage, created by merging the contributions received by the JA leaders.

The WP leader would like to thank all the BONE partners contributing to this deliverable for the highly-valued technical content they have provided to the workpackage.

3. Participants

Eleven partners have actively collaborated to the JAs of this workpackage during the second year. Table 1 shows the list of the active partners.

A detailed description of the joint activities is provided in Section 5.

Partner #	Member	Country
29	**PoliMI**	IT
2	TUW	AU
12	UC3M	ES
14	UPCT	ES
22	UoA	GR
23	UoP	GR
30	PoliTO	IT
32	UniBO	IT
36	TUE	ND
39	PUT	PL
45	UCAM	UK

Table 1: Partners actively contributing to the joint activities of WP25 during BONE Y3

4. General status of the workpackage

As well described in the first WP25 deliverable D25.1, the main goal of WP25 Topical Project on “Optical Interconnects” can be summarized by the following statement:

- This project studies the **optical implementation of the interconnection systems inside a high-performance** switching/**routing system**. The target is to develop the capability to design and optimize an **optical backplane**, finding the technology and architecture best-fitting system requirements

The accomplishment of project goals implies sharing expertise and previous work among the involved partners, merging the complementary visions of the participants on the optical implementation of interconnections within the systems.

With deliverable D25.2 this Topical Project has been organized in a structure represented in Table 2. The five Tasks initially proposed by the WP leader group the JAs of the partners according to general topics and provide specific timing to the JAs.

JA #	JA title	JA leader	Timing	Task
1	Identification of requirements of the optical backplane	**TUW**	13-24	T1: Survey and requirements
2	Survey on photonic technologies for optical interconnections	**PoliMI**		
3	Survey on optical interconnection architectures and solutions	**PoliMI**		
4	The "BONE switch": a reference architecture for the optical backplane	**UPCT**	25-36	T2: “BONE switch”
5	Performance and complexity analysis of optical switching fabrics	**PoliTO**	13-36	T3-5: OI-solution proposal and validation at different interconnection layers (rack-to-rack / shelf-to-shelf, board-to-board, chip-to-chip / on-chip)
6	Optical backplanes utilizing microring resonators	**UoA**		
7	Hardware efficient optoelectronic switch fabric	**UCAM**		
8	On-chip optical switching networks	**(*)**		

Table 2. WP25 Topical Project structure during BONE Y3.
(*): JA8 topic has been covered by JA6 and JA7 (see the text for further details)

At the beginning of the last year of WP25 (BONE Y3), eight JAs were in operation in the workpackage.

The first three JAs, dedicated to survey and requirement definition, were planned to be in operation in the first year only of the TP. Thus D25.3 does not cover JA1, JA2 and JA3 in detail (their results have been reported in D25.2). However, activity on these JAs has continued also in the last year in terms of preparatory work for three journal papers to disseminate the results obtained by the JAs. The manuscripts of these papers will be submitted for publication in the next months. For this reason we have included in this deliverable the abstracts of the papers.

JA4 was planned to be developed just in the last year of the TP, as it has been conceived to take advantage of the findings on optical technologies, architectures and requirements carried out in JA1, JA2 and JA3. This deliverable reports on the results concerning a reference architecture for the optical backplane.

JAs from JA5 to JA8 were planned to span over the two years of the TP. All these activities have concluded their work, and the results achieved are reported in this deliverable. The reports show a summary of technical content and joint and individual publications of the involved partners. In some cases, mobility actions carried out in Y3 under the umbrella of WP25 are reported.

The only exception to this rule is JA8, dedicated to On-chip optical switching networks. The work in this activity started by overviewing optical network-on-chip architectures (reported in D25.2). After this overview we realized that the work carried out in other JAs was already covering the main topic of JA8. These two activities are:

- JA6 (Optical backplanes utilizing microring resonators): indeed, microrings are probably the best suited technology for on-chip integration, and the vast majority of currently studied on-chip optical networks are based on microrings. Physical-layer and architectural analyses of interconnection architectures based upon microring resonators are reported under JA6 in this document. The considered architectures can be successful in several application domains, and in particular, for on-chip applications.
- JA7 (Hardware efficient optoelectronic switch fabric): this JA presents in this deliverable a prototype implementation of an on-chip optoelectronic switching network (see the section dedicated to JA7 further below in this document), which is a clear example of on-chip optical interconnection based on another technology, i.e. semiconductor optical amplifiers (SOAs).

Therefore, in conclusion we do not include in this deliverable a specific section for JA8, in consideration of the fact that JA6 and JA7 already provide ample research results which cover the topic of on-chip optical interconnections.

5. Joint-activity detailed reports

In the following the content of the JAs of WP25 is described in details.

5.1 JA1: Identification of requirements of the optical backplane (closed)

Participants: TUW, PoliTO, UniBO

Responsible person: Slavisa Aleksic (TUW)

Deadline: Month 24

5.1.1 Description of the work carried out

This JA achieved in BONE Y2 the completion of the targeted survey, reported in D25.2. In BONE Y3 work on this activity has however continued beyond the official end of activity in order to prepare a journal paper that will be submitted for publication soon. We include hereafter the abstract of the paper.

5.1.1.1 Abstract

We show a detailed analysis of systems where optical interconnects can be applied. The application targets are the high-performance network nodes that can more benefit from an optical implementation of the backbone subsystem. Such nodes may comprise switches, routers, digital or optical cross-connects, and add-drop multiplexers with different number of ports and different data rates. In light of these requirements, we will identify potential bottlenecks of “traditional” electronic interconnection systems, in terms of scalability towards the upgraded performance required by the future networks. This document will provide an overview on the state-of-the-art structures and technologies in electronic switching systems to implement interconnections. It will then analyze possible limitations of these solutions in front of new requirements and show in which cases and how optics can offer a real competitive advantage to equipment providers.

5.1.2 Publications

The content of the survey presented by this joint activity will be submitted for publication.

5.2 JA2: Survey on photonic technologies for optical interconnections (closed)

Participants: PoliMI, PoliTO, TUE, TUW

Responsible person: Guido Maier (PoliMI)

Deadline: Month 24

5.2.1 Description of the work carried out

This JA achieved in BONE Y2 the completion of the targeted survey, reported in D25.2. In BONE Y3 work on this activity has however continued beyond the official end of activity in order to prepare a journal paper that will be submitted for publication soon. We include hereafter the abstract of the paper.

5.2.1.1 Abstract

For transmitting very high data rates inside systems, photonics allows to overcome several limitations of electrical wiring. We show how this can occur on the typical interconnection distances rack-to-rack, board-to-board and even on-chip. The comparison between optical and electronic transmission is carried out both for serial and parallel links. Then in this survey we report about the state-of-the-art photonic technologies and devices to implement all the functions needed in an interconnection subsystem or backplane, including: signal generation, signal detection, amplification and switching. Latest advances of research in this field are reviewed. Finally, a comparative analysis of power consumption in optical and electronic interconnection systems is reported.

5.2.2 Publications

The content of the survey presented by this joint activity will be submitted for publication.

5.3 JA3: Survey on optical interconnection architectures and solutions (closed)

Participants: PoliMI, PoliTO, PUT, TUW, UoP, UNIBO

Responsible person: Guido Maier (PoliMI)

Deadline: Month 24

5.3.1 Description of the work carried out

This JA achieved in BONE Y2 the completion of the targeted survey, reported in D25.2. In BONE Y3 work on this activity has however continued beyond the official end of activity in order to prepare a journal paper that will be submitted for publication soon. We include hereafter the abstract of the paper.

5.3.1.1 Abstract

Large-capacity interconnection systems only based on electronic technology suffer technical bottlenecks that bound their scalability. Optical interconnection has proved to be a valid alternative solution in many applications, ranging from circuit switching to packet switching, to super-computing, to systems embedded on-chip. In this survey we analyze the main applications of optical interconnection. First we describe the main type of system architectures classified according to the application, analyzing their most relevant requirements of interconnection. Then, for each application class, we present the main architectures that have been proposed so far in literature and prototyped in labs or, in some cases, developed as products available on the market. Our overview aims at testifying that optical interconnection is not only a subject for theoretical studies; it is instead an important resource to increase the capability of actual high-speed systems in telecommunications and computing.

5.3.2 Publications

The content of the survey presented by this joint activity will be submitted for publication.

5.4 JA4: The "BONE switch": a reference architecture for the optical backplane

Participants: UPCT, PoliMI, PUT

Responsible person: Pablo Pavon (UPCT)

Deadline: Month 36 (starting from Month 24)

5.4.1 Description of the work carried out

5.4.1.1 Introduction

The objective of this joint activity is the definition of a future high-performance network element which is able to take full advantage of OI in order to overcome limitations of current top-line switches and routers, still based on electronic interconnection. Two lines are proposed to identify a next-generation model able to achieve breakthrough targets in terms of scalability and power/energy-savings. The first idea is to introduce optical interconnection elements in the commercial routers already tested to reach advantages in consumption and in scalability. Moreover, better behavior is expected with higher bit rates than the switches and routers based on electronic elements. The second line proposed is a novel architecture for Strictly Non-Blocking (SNB) multistage photonic switches implemented using Tunable Wavelength Converters (TWCs) and Array Waveguide Gratings (AWGs). In this case the coherent crosstalk is completely suppressed so these AWG-based switches can be used in asynchronous networks. In the following sub-sections, the advances related to both lines are described.

5.4.1.2 Optical switch fabrics in electronic switches

Electronics limitations about scalability and consumption are widely known. Besides, it is necessary to consider that as line rates increase, capacity and consumption limitations get more restrictive. As a result, electronic backplanes have become a bottleneck. The approach we study to address this bottleneck, is the replacement of some parts of existing high-capacity commercial routers by optical architectures. First, we study the architectures of some commercial high-performance routers. In this step the objective is identifying elements that can be replaced by optical systems without affecting the rest of the architecture. After the identification step, we compute the benefits obtained if the replacement was made, in terms of power consumption, switch capacity and scalability.

In our study, we focus on the commercial high-capacity routers of Cisco and Juniper. We identify two models which have a separated bufferless switch fabric. We are interested in observing how these switch fabrics can be substituted, without affecting the router operation. We choose Juniper TX Matrix and Cisco 12816.

Juniper TX Matrix

Juniper Networks presents two proposals to terabits routers. Both suggestions are based on a set of routing nodes and a platform which could be upgraded to scale to terabits of bandwidth in a multi-chassis router configuration. In this case we study the proposal composed of T640 Routing Nodes and TX Matrix Platform. TX Matrix is the central switching and routing

element that interconnects the T640 Routing Nodes to form a single routing entity in what is called a routing matrix configuration. The TX Matrix platform chassis is a rigid sheet-metal structure which has a maximum weight of 225 kg and measures 113 cm high, 44.2 cm wide, and 76.2 cm deep. A TX-Routing Matrix reaches up to 2.56 terabits per second (Tbps) (4x640 Gbps) of switching capacity and consumes about 4560 W.

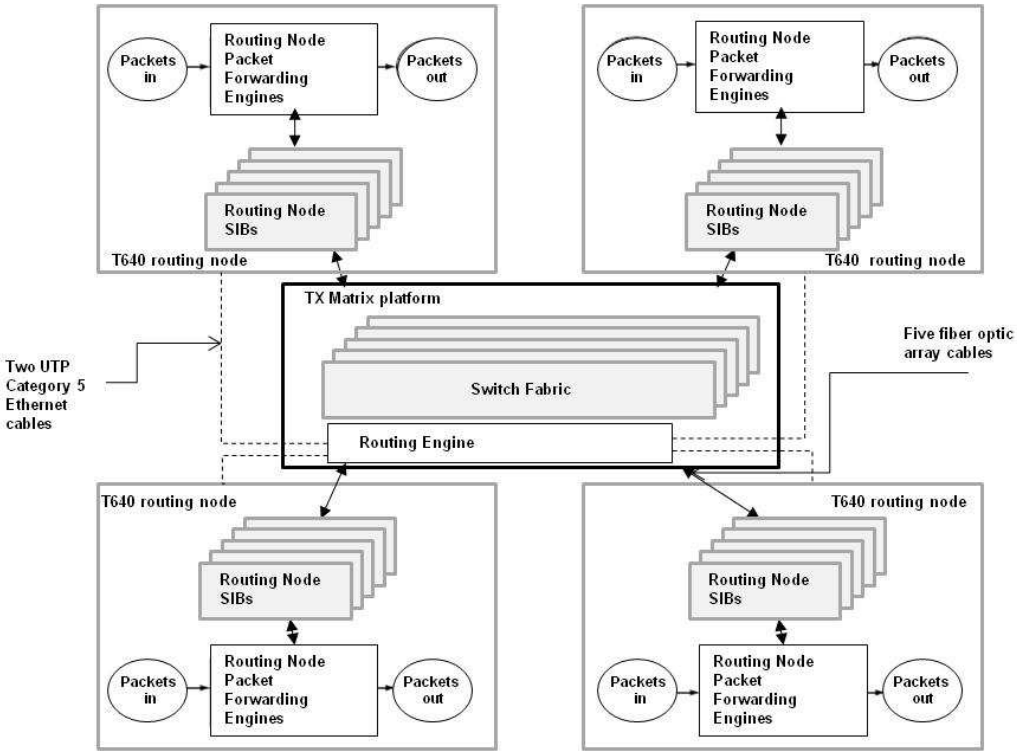

Figure 1. Juniper Networks Routing Matrix

A TX-Routing Matrix (shown in Figure 1) is composed of three major elements following the same structure in both router configurations proposed.

- From one to four T640 routing nodes which provide the network interfaces for the routing matrix and make distributed packet forwarding decisions. Each chassis contains a total of 16 Packet Forwarding Engines (PFEs) (up to 20 Gbps each for T640 and 50 Gbps for T1600) which communicate across the switch fabric.
- A TX Matrix (Plus) platform, which executes the routing protocols for the routing matrix, maintains system state, and provides the core of the switch fabric that interconnects the individual routing nodes.
- A set of cables, which interconnect the data and the control planes of each individual chassis into a unified routing matrix. Two UTP Category 5 Ethernet cables are used between the TX platform and each routing node to the control plane communications. Five fiber-optic array cables per node are utilized to connect the data plane.

In a TX-Routing Matrix, the switch fabric provides data plane connectivity among all the nodes in the matrix. The TX-Routing Matrix uses a multistage Clos network. The Clos stages are distributed between the routing nodes and the TX platform: the routing nodes have the first and the last stages, and the TX platform holds the intermediate stages. Five switch fabric planes are included in a TX-Routing Matrix and each stage is supported by Switch Interface Boards (SIBs) (see Figure 1). At a given time four of them are used in a round-robin fashion

to distribute packets from the ingress interface to the egress interface. The fifth one is utilized as a hot-backup in case of failures. The five switch fabric planes contained in TX Matrix platform are the electronic part to be replaced by the optics elements, as will be described later.

Note how the Clos switch fabric is distributed in TX Matrix. In TX Matrix case (see Figure 2), a T640 routing node in a routing matrix implements both the first and third stages Clos functionality. Each T640-SIB card in the T640 routing node performs a switch plane. The TX Matrix platform functions as the switching core of a routing matrix and contains five SIBs (Switch Interface Boards), which are connected to the T640-SIB cards in each T640 routing node. Each TX-SIB operating as the second stage of a Clos network functions as a switch plane that provides connectivity between the ingress and egress T640 routing nodes, and delivers 640 Gbps of switching capacity.

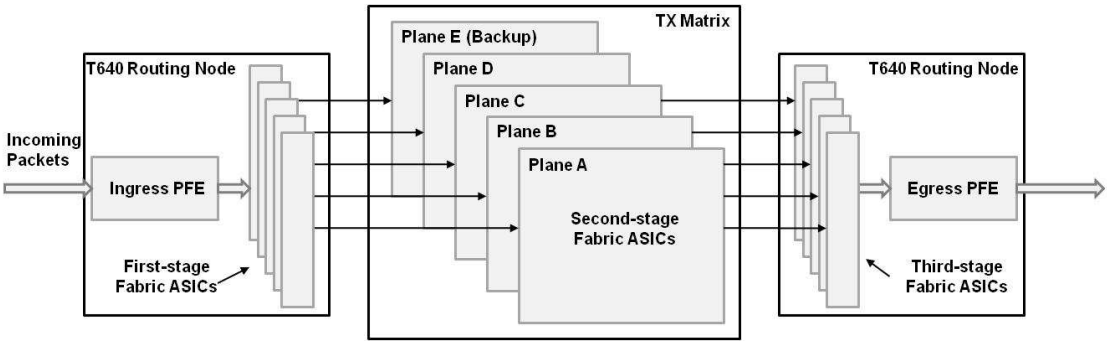

Figure 2. Routing Matrix Clos Switch Fabric Implementation in TX Matrix.

In this Juniper router the process of transmitting a data cell across the switch fabric involves the steps describe next. When a packet enters a routing node from the network, the ingress PFE creates the packet notification and segments the packet into 64-byte data cells. The data cells are then written into ingress memory, a forwarding table lookup is performed, ingress packet filtering is applied, and the data cells representing the packet are transmitted across the switch fabric to the egress PFE. The process of transmitting each data cell across the switch fabric involves the following request-grant mechanism. The request for each cell of a packet is transmitted in a round-robin order from the source PFE to the destination PFE crossing the TX platform. The destination transmits a grant to the source using the same switch plane from which the corresponding request was received. Source PFE then transmits the cell to the destination PFE on the same switch plane. When all the data cells arrive at the egress PFE, they are written into egress memory, a second forwarding table lookup is performed, egress packet filtering is applied, and the data cells representing the packet are reassembled to form the original packet.

Cisco 12816

The Cisco 12000 Series routers (see Figure 3) deliver capacity and services with its fully distributed forwarding architecture and crossbar switch fabric. We studied the product with the highest aggregate switching capacity (1280 Gbps) in 12000 Series, i.e. Cisco 12816. This router has 16 slots whose full-duplex throughputs are 40 Gbps/slot. The physical dimensions of the chassis are 181.6 cm high, 43.8 wide and 55.9 cm deep, and its maximum weight is 177 kg. The power consumption of the router considered is 4800 W. To describe Cisco 12816 we divide it into two main parts:

(i) Backplane: Gigabit Route Processors (GRPs) and Line Cards (LCs) are installed from the front of the chassis and plug into a passive backplane. This backplane contains serial lines that interconnect all of the line cards to the switch fabric cards, as well as other connections for power and maintenance functions.

(ii) 40 Gbps Switch Fabric: the core of the router is a crossbar switch fabric that provides synchronized connections between the line cards and the Router Processor (RP). The switch fabric consists of 2 clock scheduler cards (CSCs) and 3 switch fabric cards (SFCs). One CSC and the three SFCs are the active switch fabric; the second CSC provides redundancy for the other 4 cards. The SFCs and CSC provide the physical switch fabric for the system as well as the clocking for the Cisco cells that carry data and control packets among the line cards and route processors.

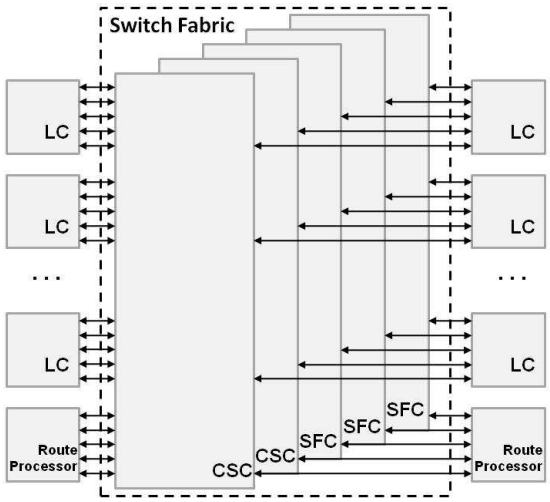

Figure 3. Cisco 12000 Series Internet Router.

The switch fabric can be imagined as an $N \times N$ non-blocking crossbar switch fabric where N stands the maximum number of LCs that can be supported in the chassis including the GRP (N = 16). Each LC has N+1 virtual output queuing (VOQ); one for each possible line card destination and one for multicast (Figure 4).

When a packet comes in an interface, a lookup is performed to determine the output LC, interface, and appropriate Media Access Control (MAC) layer rewrite information. Before the packet is sent to the output LC through the fabric, the packet is chopped into Cisco cells. A request is then made to the clock scheduler for permission to transmit a Cisco cell to the given output LC. The output LC then reassembles these Cisco cells into a packet, uses the MAC rewrite information sent with the packet to perform the MAC layer rewrite, and queues the packet for transmission on the appropriate interface.

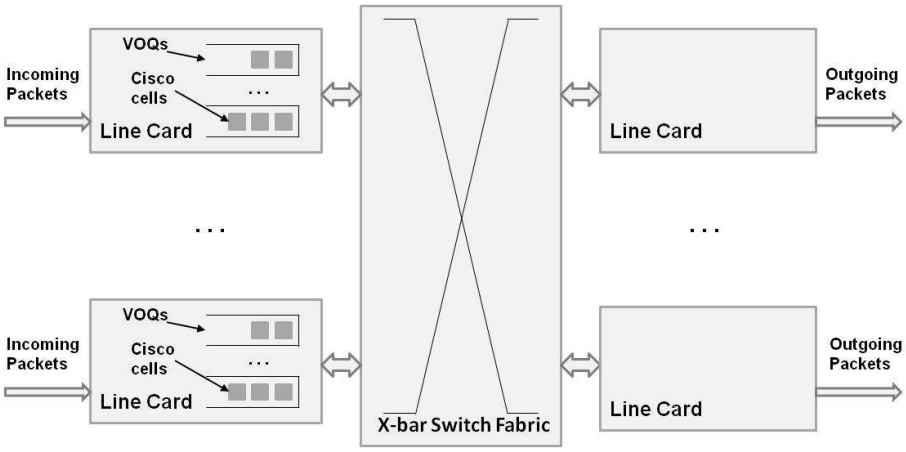

Figure 4. Crossbar switch fabric with VOQs.

Substitution stage

In our study, the electronic parts identified in each router, i.e., the five switch fabric cards in TX Matrix and the SFCs in Cisco 12816, are replaced by an AWG-based architecture such as the ones shown in Figure 5. The architecture in Figure 5(a) is composed of an array of tunable transmitters (one per switch port), followed by an arrayed waveguide grating (AWG), and an array of wide-band receivers. Considering the last researches [1], we must say AWGs present some constraints which must be taken into account. The main problem occurs when several input ports are using the same wavelength, so their corresponding output ports will present in-band crosstalk. Thus, we must control the levels of coherent crosstalk if we want to reach large-size switching fabric based on the AWG architecture. So, to this end we consider an AWG-based architecture with two stages instead of one (Figure 5(b)).

We study the substitution of the switch fabrics of the routers using the two solutions shown in Figure 5. The maximum number of ports of the one-stage solution (N_{max}) is obtained as the maximum port count possible taking into account the constraints due to the impairment effects, using the methodology in [1]. For computing the in-band effects, we assume the worst case situation in which all the transmitters in the array are using the same wavelength. For the two-stage case, it is no necessary to consider the worst-case situation. It has been shown in [2], that for this architecture it is possible to switch the packets so that at most 4 transmitters in any of the two stages are using the same wavelength. Thanks to this, the in-band crosstalk impairment is limited, and the AWG scalability is greatly improved.

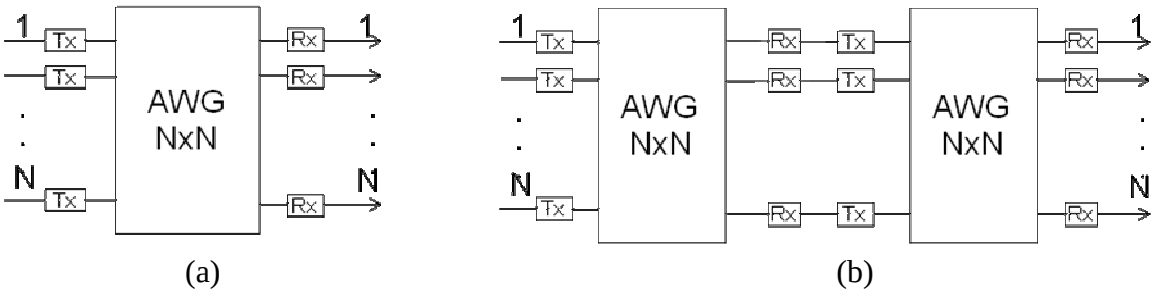

Figure 5. AWG-based one-stage and two-stage solutions.

Results

Table 1 and 2 show some results for Juniper TX Matrix. Table 3 collects consumption information for different line rates, where 50 Gbps is the line rate used by TX Matrix, and for the two proposed solutions, and maintaining the number of ports from the original router. We indicate if the solution is feasible or not according to the physical impairments. We observe that the consumption of the optical architectures is very low, in the order of mW. This suggests that the total consumption of the TX Matrix would be considerably reduced if this replacement was performed. Table 2 shows the throughput which could be reached if we exploited the scalability of the AWG-based architectures at its maximum that means using with the maximum port count allowed by the physical impairments $\boldsymbol{N_{max}}$. Results show that the throughput of the architectures can be increased in all the cases respect to its original number (2.56 Tbps), and/or the number of nodes which could be potentially connected to the TX Matrix platform is increased.

	Line Rate (Gbps)	10	40	50	100
One-stage solution	**No. AWGs**	5	5	5	5
	***N*×*N* AWG size**	32×32	8×8	7×7	4×4
	Impairment feasibility	Not feasible	Feasible	Feasible	Feasible
	Consumption(mW)	-	6560	7000	7600
Two-stage solution	**No. AWGs**	10	10	10	10
	***N*×*N* AWG size**	32×32	8×8	7×7	4×4
	Impairment feasibility	Feasible	Feasible	Feasible	Feasible
	Consumption (mW)	17920	13120	14000	15200

Table 1. Consumption Data TX Matrix (2.56 Tbps / 4 nodes)

	Line Rate (Gbps)	10	40	50	100
One-stage solution	**No. AWGs**	5	5	5	5
	***N*×*N* AWG size**	32×32	14×14	13×13	9×9
	Throughput (Gbps)	-	4480	5200	7200
	Total no. nodes	-	7	8	11
Two-stage solution	**No. AWGs**	10	10	10	10
	***N*×*N* AWG size**	127×127	67×67	57×57	27×27
	Throughput (Gbps)	10160	21440	22800	21600
	Total no. nodes	15	33	35	33

Table 2. Throughput Data TX Matrix (2.56 Tbps / 4 nodes)

5.4.1.3 Asynchronous AWG-based architecture

Building upon the outcomes of previous JAs, we propose a modular AWG-based optical interconnection architecture for B2B and R2R scenarios. Our network has a controllable amount of crosstalk (approximately to zero), high throughput, low optical path loss, lower power consumption. The proposed architecture complies to the fundamental guidelines coming from JA1, JA2 and JA3, which are: feasibility with current or near-future technology, low power consumption, high resilience, competitive advantage compared to electronic solutions.

Arrayed Waveguide Grating (AWG) is a very attractive passive optical device for constructing high-speed large-capacity WDM optical switches, because it is scalable to large size, consumes little power, offers high wavelength selectivity, low insertion loss, small size,

potentially low cost and fast switching time. As a matter of fact its switching speed is determined only by the speed of wavelength conversion, which is in the order of nanoseconds and sub-nanosecond, in the Tunable Wavelength Converters (TWCs) located at the ingress ports of the AWG. Moreover, AWGs and TWCs have the advantage to be commercially available components. These are key features that enable the design of an optical AWG-based switch that is cost-efficient.

AWGs are passive devices behaving as multiport interferometers. In the 1xN (Nx1) configuration, AWGs act as wavelength multiplexers (demultiplexers). In the $N \times N$ configuration, AWGs behave as wavelength routers. The information at an input port is forwarded to an output port depending on the selected wavelength. More specifically, at each input port, different wavelengths are used to reach different output ports.
Commercial AWGs provide uniform transfer functions, and extinction ratios among adjacent channels in the order of 30-40 dB. These physical layer characteristics are largely sufficient for multiplexers and demultiplexers, which are indeed commonly used in commercial WDM systems. However, significant coherent crosstalk figures were reported in $N \times N$ AWGs with large port counts. If the same wavelength is used at all AWG inputs, the maximum admissible value of N is severely limited.

We propose a method to design Zero-Coherent- Crosstalk (ZCC) and Strictly Non-Blocking (SNB) AWGs-based switches. An AWG-based switch is ZCC if each wavelength is not reused more than one time in any instant time. Moreover it is also SNB if the information can be transferred by an unused wavelength from any unused switch input to any unused switch output without changing the routing/ wavelength of any previously established connection. The coherent crosstalk suppression allows us using these AWG-based switches in asynchronous networks. With our approach, we are always able to establish an interconnection between an idle input port and an idle output port without reuse wavelengths utilized for connection demands already established.

To eliminate the problem of the coherent crosstalk, in the following we propose an optical switch fabric using combinations of an AWG and WDM couplers.

Without loss of generality, Figure 6 shows a 3×3 and a 4×4 implementation of the proposed ZCC architecture. Focusing on Figure 6b, the switch fabric supports up to $N = 4$ simultaneous connections. At most one cell can be sent from each input i ($i \mid 0 \leq i \leq N$-1), in the set of the switch input ports I, and at most one cell can be sent to each output j ($j \mid 0 \leq j \leq N$-1), in the set of the switch output ports J. The proposed architecture comprises an AWG with a set of $|S| = 8$ input ports $S = [s]$ ($s \mid 0 \leq s \leq |S|$-1) and a set of $|D| = 8$ output ports $D = [d]$ ($d \mid 0 \leq d \leq |D|$-1), where $|S|$ is always equal to $|D|$. Moreover 4 couplers are employed. Note that the switch is characterized by two different set of inputs (outputs), the switch input ports I (O) and the AWG input ports S (D). Each switch input port i is connected with a different even AWG input port s, leaving the odd AWG input ports unconnected. Each output port j is connected via coupler with a specific pair of AWG output ports following a specific algorithm (in this case $d = j$ and $d = j + 4$). Each input port i of the switch is equipped with a single TWC that assigns to each cell a wavelength out of a predetermined set of $|S| = |D| = 8$ wavelengths $\Lambda = \{\lambda_0, \lambda_1, \ldots, \lambda_7\}$.

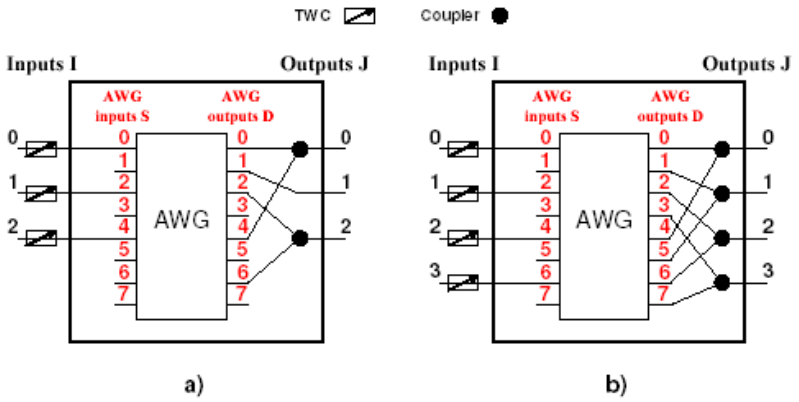

Figure 6: a) 3x3 and b) 4x4 ZCC/SNB AWG-based switches realized via an 8×8 AWG.

By way of illustration, Figures 7a and 7b show the 5x5 and 6x6 optical switches. Figure 8 and in Figure 9 report how the wavelength interconnections can be established into the respective optical switch [7].

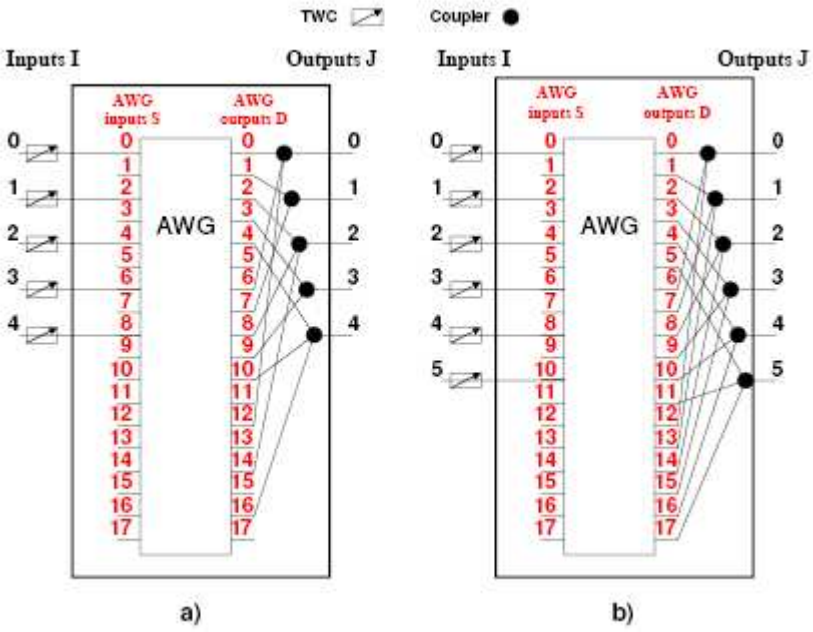

Figure 7. a) 5x5 and b) 6x6 ZCC/SNB AWG-based switches realized via a 16×16 AWG.

		Input Ports				
		0	1	2	3	4
Output Ports	0	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$
	1	$\lambda_1\lambda_7$	$\lambda_5\lambda_{17}$	$\lambda_3\lambda_{15}$	$\lambda_1\lambda_{13}$	$\lambda_{11}\lambda_{17}$
	2	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$
	3	$\lambda_3\lambda_9$	$\lambda_1\lambda_7$	$\lambda_5\lambda_{17}$	$\lambda_3\lambda_{15}$	$\lambda_1\lambda_{13}$
	4	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$

Figure 8. The wavelength interconnections that can be established for the optical switch in Figure 7a

Output Ports \ Input Ports	0	1	2	3	4	5
0	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$
1	$\lambda_1\lambda_7\lambda_{13}$	$\lambda_5\lambda_{11}\lambda_{17}$	$\lambda_3\lambda_9\lambda_{15}$	$\lambda_1\lambda_7\lambda_{13}$	$\lambda_5\lambda_{11}\lambda_{17}$	$\lambda_3\lambda_9\lambda_{15}$
2	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$
3	$\lambda_3\lambda_9\lambda_{15}$	$\lambda_1\lambda_7\lambda_{13}$	$\lambda_5\lambda_{11}\lambda_{17}$	$\lambda_3\lambda_9\lambda_{15}$	$\lambda_1\lambda_7\lambda_{13}$	$\lambda_5\lambda_{11}\lambda_{17}$
4	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$	$\lambda_4\lambda_{10}\lambda_{16}$	$\lambda_2\lambda_8\lambda_{14}$	$\lambda_0\lambda_6\lambda_{12}$
5	$\lambda_5\lambda_{11}\lambda_{17}$	$\lambda_3\lambda_9\lambda_{15}$	$\lambda_1\lambda_7\lambda_{13}$	$\lambda_5\lambda_{11}\lambda_{17}$	$\lambda_3\lambda_9\lambda_{15}$	$\lambda_1\lambda_7\lambda_{13}$

Figure 9. The wavelength interconnections that can be established for the optical switch in Figure 7b

As shown in Figures 8 and 9, the coherent crosstalk constrain is suppressed thanks to the multiple choice of wavelength to reach the output ports *J* from the input ports *I*.
In Figure 10 we report the number of AWG input/output ports required to implement the ZCC/SNB AWG-based switches. The implementation of a single-stage ZCC/SNB AWG-based switch with a port counts greater than 9-10 is an impracticable solution because it requires an AWG with more than 50 input/output ports. Therefore, our study is focused on the AWG with port counts lower than 40 channel devices with a channel spacing of 100 GHz which are commercially available.

AWG-based switch ports	AWG ports
3x3	8x8
4x4	8x8
5x5	18x18
6x6	18x18
7x7	32x32
8x8	32x32
9x9	50x50
10x10	50x50

practical

impracticable

Figure 10. Number of AWG ports required to realize a N×N ZCC/SNB AWG-based switch.

We have generalized the coherent crosstalk suppression technique for any *N*×*M* AWG-based switch (with $N > M$, $N < M$). Further details can be found in Ref. [7].
The multi-stage approach proposed by Clos can be adopted to solve scalability issues and construct switches with a larger port counts. In Figure 11 a 32×32 ZCC/SNB three-stages Clos switch is reported. Networks with more than three stages can be built by iterating the basic three-stages construction.

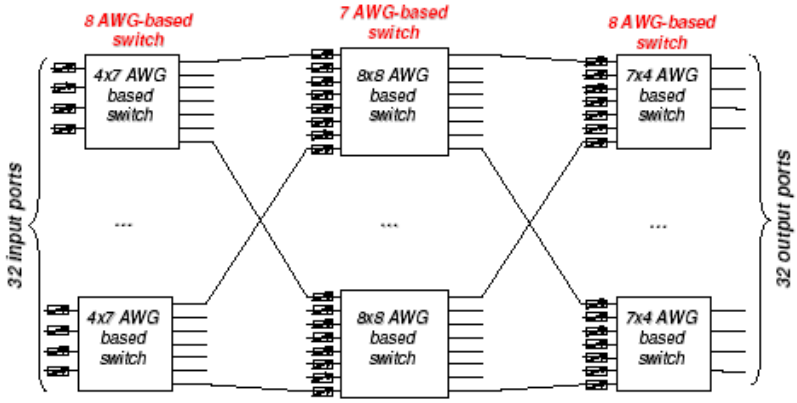

Figure 11. 32×32 ZCC/SNB three-stages Clos switch.

Let us now consider the feasibility and scalability of the optical fabric architectures described previously. The number of AWG and TWC devices required by iterative Clos architectures construction is reported in Figure 12 [7].

Architecture	Port count	#TWCs	#AWGs
3-stage	32×32	144	23
5-stage	128×128	1360	255
7-stage	512×512	10928	1831
9-stage	2048×2048	82128	13841

Figure 12. Number of devices for the proposed Clos architectures.

The number of TWCs primarily determines the cost, the power consumption and reliability of the switch-fabric. Even though TWCs are, at this point in time, rather complex and thus expensive devices, we do expect that their cost will drop severely. Indeed, research on these devices continues and integration of the converters with tunable lasers has already been proposed, allowing production at a substantially lower price.

Power loss in our switching architectures, which is experienced by AWG loss and coupler loss, is independent of the number of stages that characterize the Clos networks. As a matter of fact, at each stage ingress the signals traverse an array of Electrical-Optical-Electrical (EOE) TWCs where they are 3R regenerated. Figure 13 shows the optical signal path through 3 stages of a multistage Clos networks. Each stage is connected with the following one by very low-loss propagation media provided by the single-mode silica optical fibers with 0.2 dB/km (at 1.55 µm) loss. Given that typical rack-to-rack interconnection distances are lower than 100 m, the propagation attenuation is negligible in our model. For the TWC, a typical range transmitted power is from about 1 mW (0 dBm) up to a few 10 mW (10 dBm). In our analysis, we assume a transmitted power of 3 dBm. This is consistent with typical tunable laser peak output power on the order of 10 dBm, and with the 6-7 dB equivalent loss introduced by an external modulator (3 dB due to on-off keying and 3-4 dB due to additional insertion loss). The power penalty of passive devices like AWGs and couplers is dependent on the number of ports. The insertion loss that is inferred from data sheets for AWGs shows a dependency on the number N of ports that contributes logarithmically to the power penalty, i.e., $2 + 1.2\log(N/4)$ dB. Regarding crosstalk effects, the suppression of the coherent (in-band)

crosstalk leads to crosstalk-related contributions to penalty that are very weak for all architectures, which creates only out-of-band crosstalk. When *N* is equal to 32 and 14 the insertion loss of an AWG is approximately 3 dB and 2.6 dB, respectively. Successively, when the signal traverses *N*-port couplers (or splitters), the coupling process gives a minimum theoretical loss increasing with 10 log(*N*) dB. With 32×32 AWGs the outputs are coupled with a coupler degree of 4, which implies an insertion loss equals to 6 dB. Otherwise, in 14×14 AWGs the coupler degree is equal to 2, then the insertion loss is 3 dB. For the power received by TWCs, we assume a target Bit Error Rate (BER) equal to 10^{-12}, for which the receiver in the TWCs have a receiver sensitivity equal to -17.8 dBm and -12.5 dBm with 40 Gbps and 100 Gbps [3], respectively.

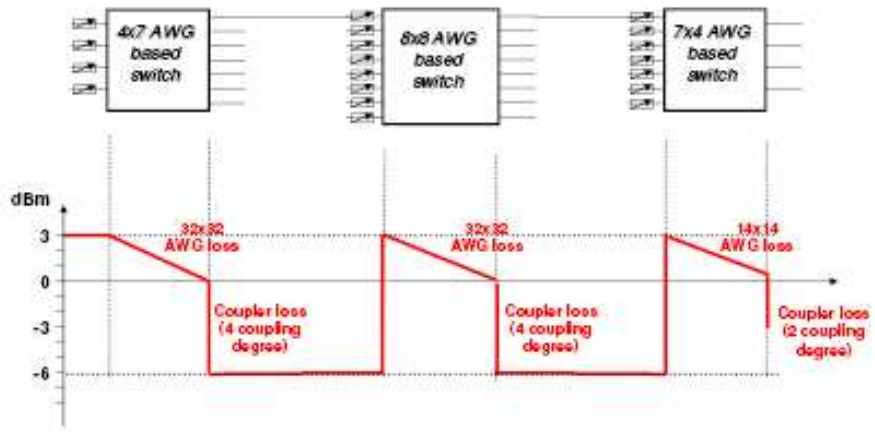

Figure 13. Power budget analysis in ZCC/SNB 3-stages Clos network.

Figure 13 shows that, the output power from each ZCC/SNB AWG-based switch is at least - 6 dBm, which is much higher than required given the TWC sensitivity in the ingress of the downsatream ZCC/SNB AWG-based switch, also with 100 Gbps channels. Therefore, also with a 32×32 ZCC/SNB three-stages Clos architecture, characterized by a restricted number of TWCs and AWGs, a throughput of 3.2 Tbps is offered. If the 2048×2048 configuration is adopted, the throughput is increased up to 204.8 Tbps. These throughput values are comparable with the ones offered by current electronic commercial routers Juniper TX Matrix (with 4xT640) and Cisco CRS-1, which are equal to 2.5 Tbps and 92 Tbps, respectively. Figure 14 shows the systems throughput under different channel rates.

Architecture	**Port count**	**Throughput (40 Gbps)**	**Throughput (100 Gbps)**
3-stage	32×32	1.3 Tbps	3.2 Tbps
5-stage	128×128	5.1 Tbps	12.8 Tbps
7-stage	512×512	20.5 Tbps	51.2 Tbps
9-stage	2048×2048	81.9 Tbps	204.8 Tbps

Figure 14. Throughput for the proposed Clos network.

Another issue with core routers is the amount of power required to operate them. In this section we estimate the power consumption of each multistage Clos architecture. For this

purpose, in the following we specify the energy per bit of common OEO TWCs required to implement the following interconnection models. In our calculation we assume that the TWC is composed by a receiver as its input and the modulator as its output, and the transmitter must be tunable. The total power consumed by the fixed receiver is 1.1 mW/Gbps [4].
Analogously, the power supply for a Vertical Cavity Surface Emitting Laser (VCSEL) is 2.5 mW/Gbps [5] plus 10mW to tuning laser [6]. Although the proposed multistage Clos architectures requires a high number of TWCs, only a subset of the installed TWCs may be simultaneously in an active state: this subset has a cardinality of *N*x*k*, where *N* is the switch port counts and *k* is the number of stages in the Clos architecture. Figure 15 reports the power consumption for different Clos architectures based on two different channel rates, i.e. 40 Gbps and 100 Gbps. For equal values of throughput, our architecture requires lower power consumption than the one required by Juniper TX Matrix (with 4xT640) and Cisco CRS-1, which, in their maximum configuration, is equal to 31 kW and 1 MW, respectively.

Architecture	Port count	#TWCs utilized	Power (40 Gbps)	Power (100 Gbps)
3-stage	32×32	96	14.7 W	35.5 W
5-stage	128×128	640	98.5 W	236.8 W
7-stage	512x512	3584	552 W	1.3 kW
9-stage	2048x2048	18432	2.8 kW	6.8 kW

Figure 15. Power consumption for the proposed Clos network.

5.4.2 References

[1] R. Gaudino, G. A. G. Castillo, F. Neri, and J.M. Finochietto, "Can simple optical switching fabrics scale to terabit per second switch capacities?," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 1, no. 3, pp. B56–B69, Aug. 2009.

[2] A. Bianco, D. Hay, and F. Neri, "Crosstalk-Preventing Scheduling in Single- and Two-Stage AWG-Based Cell Switches," *IEEE/ACM Transactions on Networking*, DOI 10.1109/TNET.2010.2054105 , 2010.

[3] E. Sackinger, "Broadband Circuits for Optical Fiber Communication," Wiley, Aug. 2005.

[4] Steven J. Koester and Clint L. Schow and Laurent Schares and Gabriel Dehlinger and Jeremy D. Schaub and Fuad E. Doany and Richard A. John, "Ge-on-SOI-Detector/Si-CMOS-Amplifier Receivers for High-Performance Optical-Communication Applications," *IEEE/OSA Journal of Lightwave Technology*., vol. 25, no. 1, pp. 46–57, 2007.

[5] C. Kromer, G. Sialm, C. Berger, T. Morf, M. L. Schmatz, F. Ellinger, D. Erni, G.-L. Bona, and H. Jackel, "A 100-mW 4 x 10 Gb/s transceiver in 80-nm CMOS for high-density optical interconnects," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2667–2679, 2005.

[6] R. S. Tucker, "The Role of Optics and Electronics in High-Capacity Routers," *IEEE/OSA Journal of Lightwave Technology*., vol. 24, no. 12, pp. 4655–4673, 2006.

5.4.3 Publications

[7] D. Lucerna, G. Maier, A. Pattavina, "AWG-based architecture for optical interconnection in asynchronous sysmes", submitted to *HPSR 2011*, Cartagena, Spain.

5.4.4 Mobility actions

No.	**Brief Description of Mobility**	**Participating partners (Indicate the host institute)**
1	Dates: 01.06 – 07.06.2010 Visitor: Belen Garcia-Manrubia (UPCT) Host: Guido Maier (PoliMI) Collect and analyze material about integration of optical technologies for high performance interconnection produced in the Joint Activity “BONE Switch”.	PoliMI

5.5 JA5: Performance and complexity analysis of optical switching fabrics

Participants: PoliTO, PoliMI, UniBO, TUW, UPCT, UoP

Responsible person: Guido Gavilanes, Fabio Neri (PoliTO)

Deadline: Month 36

5.5.1 Description of the work carried out

5.5.1.1 Optical packet switching architectures with shared wavelength converters

All research efforts in this JA have been directed to the analysis of architectures proposed by partners, such as the wavelength selective switch at UniBO, and classical architectures with microring resonator based filtering devices.

In the work by UniBO photonic switch architectures with shared wavelength converters [1], [10] are considered mainly in the perspective of signal path analysis, by accounting for the physical properties of the employed optical components. In particular, the different forwarding paths which support data transfer are investigated and represented by block diagrams to take into account physical properties like power loss, noise and cross-talk effects. To this end passive and active optical components are characterized using off-the-shelf data sheets or existing models. Two different, recently proposed, architectures are considered, the Shared-Per-Input-Wavelength (SPIW) [1] and the Shared-Per-Output Wavelength (SPOW) [10]. They are sketched in Figure 1(a) and Figure 1(b), respectively.

SPIW connects N input/output fibers carrying M wavelengths, and shares r_w Wavelength Converters (WCs) among packets coming on the same wavelength. There are M pools (dedicated to different wavelengths) of r_w WCs each. The total amount of WCs is Mr_w. The required WCs are Fixed-input/Tunable-output WCs (FTWCs). The organization is very modular to ensure flexibility and to reduce the size of the needed splitters/couplers. Further details can be found in [1].

SPOW is similar to the SPIW, but it employs Fixed output WCs (FWCs), which are simpler to implement. In this case, there are M pools of WCs, and WCs in the same pool are shared among the packets that are scheduled for forwarding on the same wavelength, in different output fibers. Further details can be found in [10].

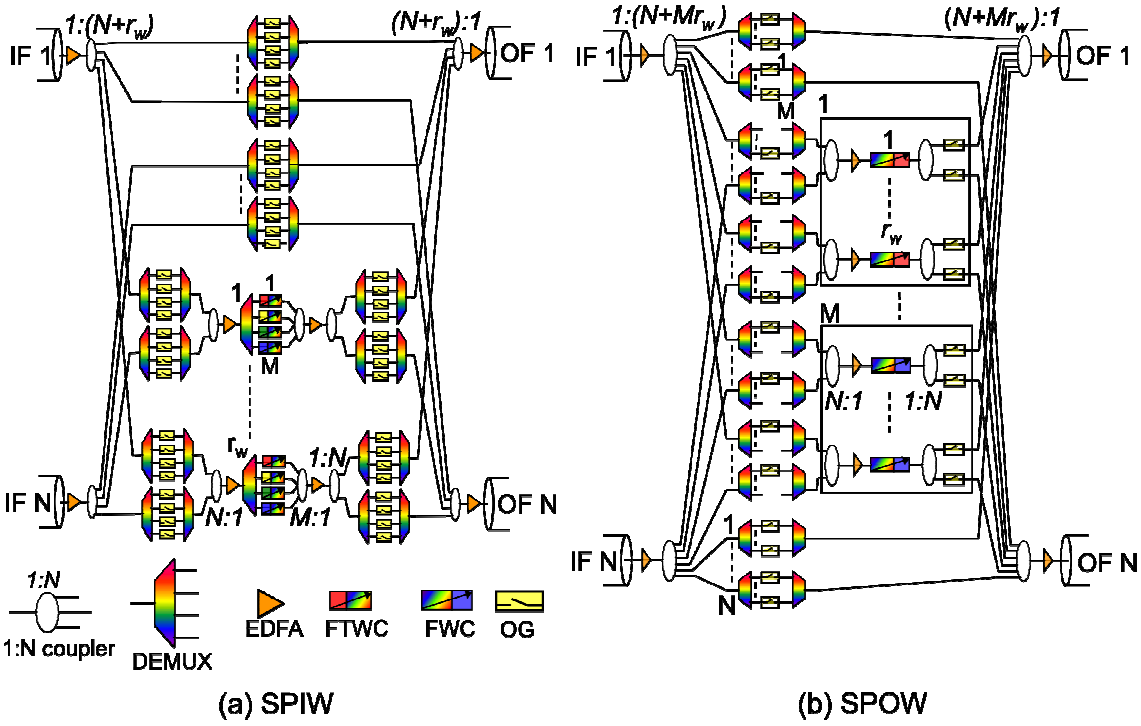

Figure 1. SPIW (a) and SPOW (b) architectures based on SOA technology. The architectures connect N input/output fibers carrying M wavelengths, and are provided with r_w wavelength converters per wavelength.

Loss and noise in optical signal paths

Forwarding paths in SPIW and SPOW architectures are here considered. In the proposed switching architectures, the optical signal carrying a data packet can pass from the input to the output fiber through two different paths, depending on the need for wavelength conversion. These paths are shown for the SPIW and SPOW architectures in Figure 2(a) and Figure 2(b), respectively.

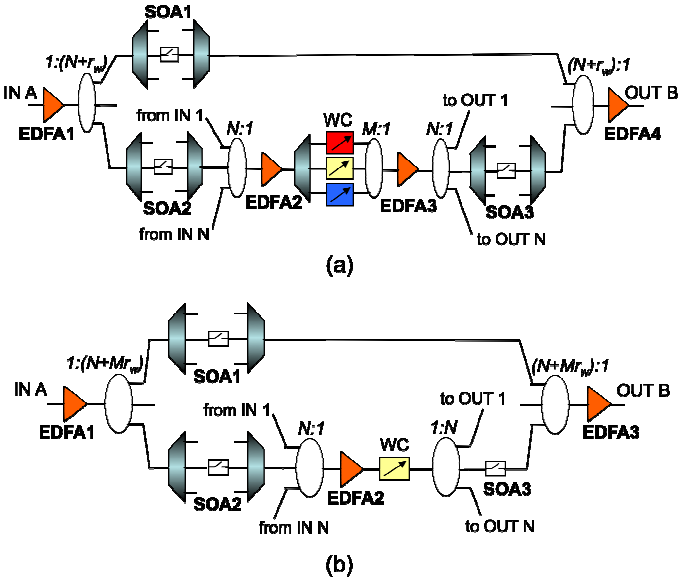

Figure 2. Optical forwarding paths in SPIW (a) and SPOW (b).

Different passive and active optical devices, with different number of ports, are employed. Passive devices are used to split/couple, multiplex/demultiplex the optical WDM signal, in order to connect paths to different subsystems. Active optical devices are used for switching and amplification purposes. Passive devices introduce optical power loss; active devices amplify the optical signal, while introducing additional noise, which impacts on the Optical Signal-to-Noise Ratio (OSNR). The physical characterization of these devices is now presented, by applying a methodology based on [3], [4], which was specialized for the SPIW and SPOW forwarding paths.

Passive devices

The passive devices are splitters/couplers (S/C) and demultiplexers/multiplexers (D/M). The ideal power loss introduced by a splitter (coupler) with N outputs (inputs), is given by $10\log_{10}N$ dB, while an ideal MUX (DEMUX), able to multiplex (demultiplex) M wavelengths, does not introduce any loss.

To provide a quite realistic loss analysis, the actual power loss introduced by these devices has been evaluated, taking second order effects into account, by interpolation of the values reported in data sheets [4], [5]. Three different effects are considered to contribute to the actual loss: insertion loss (*IL*), non-uniformity (NU) and polarization dependence (PD). The resulting losses, indicated with $L_s(N)$ and $L_c(N)$ for a splitter/coupler with N outputs (inputs), respectively, are given by:

$$L_S(N)\Big|_{dB} = L_C(N)\Big|_{dB} = \underbrace{1+11\log_{10}(N)}_{IL} + \underbrace{0.55N^{0.3}-0.05}_{NU} + \underbrace{0.3}_{PD} \quad dB$$

while the losses introduced by a MUX/DEMUX with M wavelengths, indicated with $L_M(M)$ and $L_D(M)$, respectively, are:

$$L_M(M)\Big|_{dB} = L_D(M)\Big|_{dB} = \underbrace{1.2\log_2(N)}_{IL} - 0.4 + \underbrace{1.0}_{NU} + \underbrace{0.05\log_2(N)+0.05}_{IL} \quad dB$$

Active devices

The active devices used in the proposed architectures are Semiconductor Optical Amplifiers (SOAs), which are used as ON/OFF Optical Gates (OG), and Erbium Doped Fiber Amplifiers (EDFAs), used for optical signal amplification purposes. As a matter of fact, SOAs along the optical paths are mainly used for switching purposes, by allowing/blocking the propagation of the optical signal through the architecture. They also amplify the optical signal. Instead, EDFAs are used to recover the losses introduced by passive optical devices.

Both SOAs and EDFAs, being active devices, introduce additional noise. The spectral noise density at the output of SOA and EDFA amplifiers $N_{OUT,S}$ and $N_{OUT,E}$, respectively, can be expressed (in dB) as [3]:

$$N_{OUT,SOA} = N_{IN,SOA}G_S + \underbrace{hf(G_S-1)F_S}_{ASE\ \ noise} \qquad N_{OUT,EDFA} = N_{IN,EDFA}G_E + \underbrace{hf(G_E-1)F_E}_{ASE\ \ noise}$$

where h is the Plank's constant, f is the frequency in Hertz, G_E, F_E and G_S, F_S are the gain and noise figure of the EDFA and SOA amplifiers, respectively. $N_{IN,SOA}$, $N_{OUT,SOA}$, $N_{IN,EDFA}$,

$N_{OUT,EDFA}$ represents the input and output noise spectral densities for SOA and EDFA, respectively. All these values are here considered in dB or dBm. Typical values are F_E = 5 dBm, F_S = 9 dBm [4]. Both components amplify the noise on input and add additional Amplified Spontaneous Emission (ASE) noise. It was assume that amplifiers provide an output power (among all the channels) given by $P^{tot}_{OUT,SOA} = 0$ dBm and $P^{tot}_{OUT,EDFA} = 25$ dBm for SOA and EDFA, respectively. In the proposed architectures the SOA operates on a single wavelength channel, so the output power for that channel is $P_{OUT,SOA} = P^{tot}_{OUT,SOA} = 0$ dBm, while an EDFA should amplify a WDM signal with up to M wavelengths. Assuming uniformity among the channels, the output power for a single channel is $P_{OUT,EDFA} = P^{tot}_{OUT,EDFA} - 10\log_{10}(M)$ dBm.

To ensure a proper functionality, the input power of an amplifier must be higher than the amplifier sensitivity, here assumed to be S_{SOA} = –10 dBm and S_{EDFA} = –29 dBm, respectively. The SOA and EDFA gains, G_S and G_E are determined by the difference between the (constant) output power and the (variable) input power.

Moreover, due to their finite Extinction Ratio (ER), which is the ratio between the power in the ON and OFF states (ideally ∞), SOAs do not behave exactly as ON/OFF gates. In practice, when an SOA is turned to the OFF state, part of the signal, instead of being blocked, still passes through the device. The higher this power is, the lower the ER. This way, in a coupler that multiplexes signals coming from different paths, the fraction of a signal exiting from an SOA in the OFF state becomes an interference for the one exiting from an SOA in the ON state. There are two different types of crosstalk: coherent (or in-band) crosstalk and incoherent crosstalk. The most relevant type for such architectures is the first one, which is given by [5]:

$$IX(N)\Big|_{dB} = 10\log_{10}(1 - \frac{I}{ER}Q^2) \quad dB$$

where I represents the number of interfering channels, ER is the SOA extinction ratio, $Q = \sqrt{2}erfcinv(2BER)$ represents the target quality factor (expressed in linear units) for OOK modulated systems, given a target Bit Error Rate (BER) [3]. Typical values of Q lie in the range of 6-7 for BERs between 10^{-9}–10^{-12}, while the extinction ratio here considered is ER = 35 dB.

As far as the WCs, it is here assumed that they are all-optical and based on SOA technology. They are assumed to generate a signal on a different wavelength with output power $P_{OUT,WC}$ = 3 dBm, without introducing further noise (in fact they regenerate the signal) and eliminating the crosstalk previously accumulated, given that the signal is translated on a new wavelength. This component is not mature yet, so further characterization should be considered as a future work [5].

Path analysis

To evaluate the loss and noise in such architectures, it is assumed that the optical signal carried by a wavelength is received by a burst mode receiver, with sensitivity S_R, given by [3]:

$$S_R(R_b)\Big|_{dBm} = -26 + 13.5\log_{10}(\frac{R_b}{10Gbit/s}) \quad dBm$$

where the bit rate R_b is expressed in Gbps (linear scale). A target $OSNR_T$ = 20 dB is assumed for this receiver.

The actual and target OSNR are calculated by considering the spectral noise density over a conventional bandwidth equal to the value of the channel bit rate. It is possible to represent the different optical paths illustrated in Figure 2 as block diagrams, as shown in Figure 3 for the SPIW architecture.

Each block describing an amplifier is characterized by gain and noise figure, indicated with *G*, while blocks accounting for passive devices are characterized by attenuation, indicated with *L*. All values are considered in dB. A final block takes crosstalk into account as additional interfering sources, as explained later on.

As stated above, an amplifier needs a given input power, which must be higher than its sensitivity. On the other hand an amplifier provides an output signal at acceptable power level. The optical path can be divided into different sub-paths, each going from the output of an amplifier to the input of the next one. The level of power provided to the amplifier must be evaluated for each sub-path. As an example, for the path without conversion of SPIW (Figure 3(a)), there are three sub-path:

i. from input to SOA1;
ii. from output of SOA1 to EDFA4;
iii. from output of EDFA4 to the receiver.

It should be ensured, respectively, that

i. $P_{IN,SOA1} = P_{OUT,EDFA1} - L_S(N+r_w) - L_D(M) \geq S_{SOA1}$;
ii. $P_{IN,EDFA4} = P_{OUT,SOA1} - L_M(M) - L_C(N+ r_w) \geq S_{EDFA4}$;
iii. $P_{OUT} = P_{OUT,EDFA4} \geq S_R$.

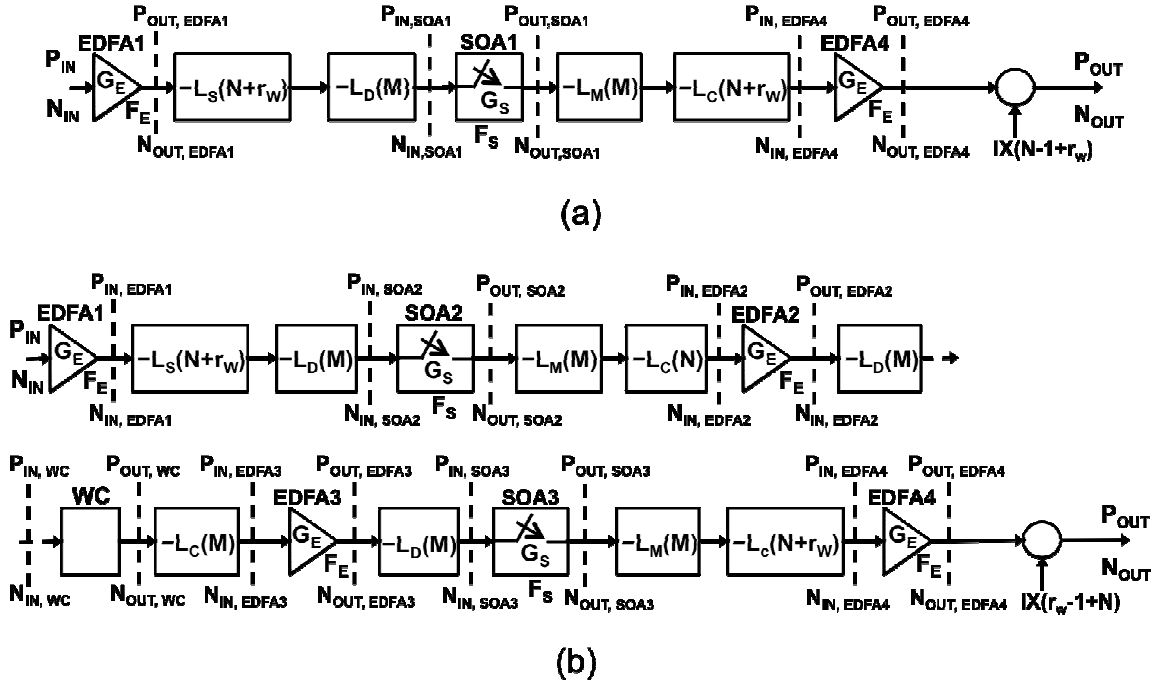

Figure 3. Block diagrams representing the optical devices through the forwarding paths in SPIW architecture: (a) path without wavelength conversion, (b) path with wavelength conversion.

At the receiver, it is also necessary to verify that the *OSNR* is larger than the target $OSNR_T$ of the receiver ($OSNR = P_{OUT} - (N_{OUT} + R_b) \geq OSNR_T$).

The coherent crosstalk, indicated by IX in Figure 3, is accounted at the receiver as N-1+r_w interfering sources. In fact, N–1 paths connect the input fibers other than IN A to the desired output fiber OUT B directly (without wavelength conversion) and r_w paths connect the r_w WC pools to the OUT B. So, there are up to N–1+r_w interfering sources on the same wavelength, and the worst case was always considered. These interfering sources can be viewed as additional noise, so the crosstalk must be accounted when considering the OSNR at the receiver.

In case the output of the node is connected to further nodes, the crosstalk contribution accumulates through cascading nodes and will be accounted at the final receiver.

Numerical results

The following parameters are considered to obtain results for SPIW and SPOW paths: bit rate R_b = 10 Gbps (leading to S_R = –26 dBm), P_{IN} = 3 dBm for an input wavelength channel, input $OSNR$ = 60 dB, $P_{OUT,EDFA}$ = 25 – 10log$_{10}$(M) dBm, F_E = 5 dBm, S_{EDFA} = –29 dBm, $P_{OUT,SOA}$ = 0 dBm, F_S = 9 dBm, S_{SOA} = –10 dBm, ER = 35 dB, Q = 6, $P_{OUT,WC}$ = 3 dBm, receiver $OSNR_T$ = 20 dB.

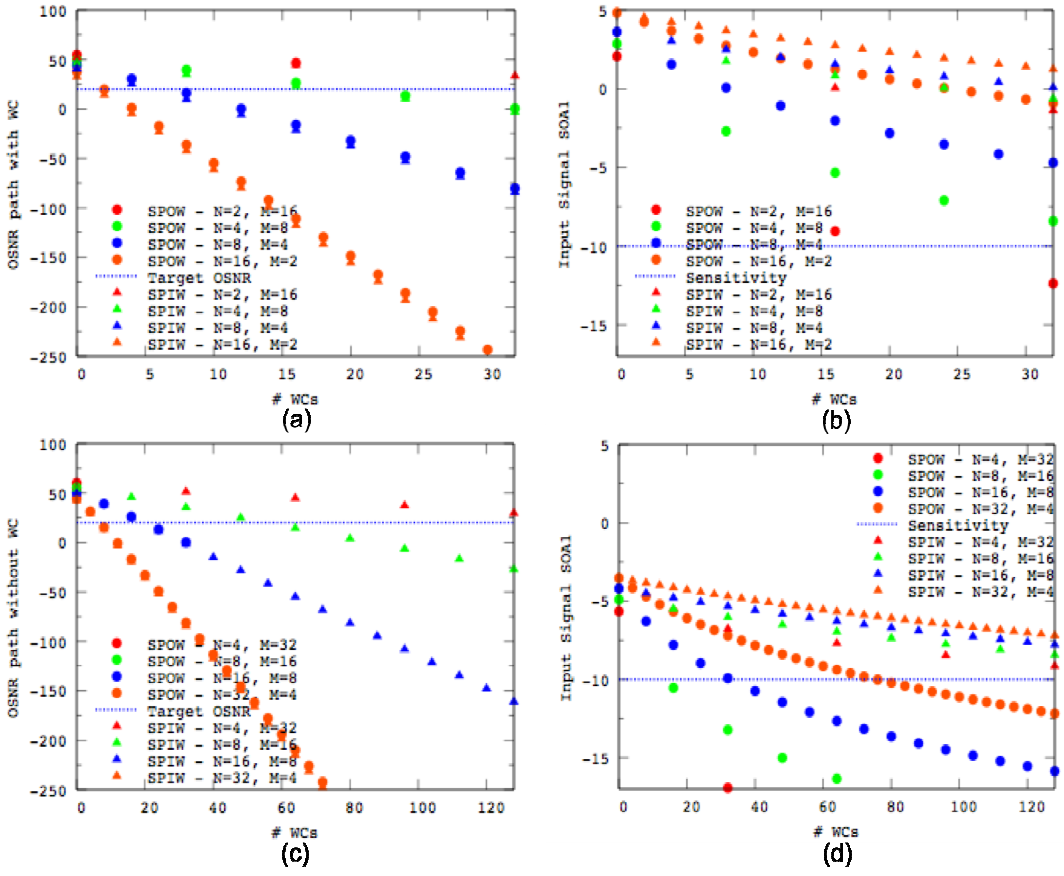

Figure 4. Results for SPIW and SPOW. 4(a) and 4(b) show the OSNR for the path with WC and the power before SOA1 as in the case NM=32 channels. 4(c) and 4(d) show the same results for the case NM=128 channels.

Figure 4(a) and Figure 4(b) plot results for SPIW and SPOW equipped with a total of NM = 32 channels. Figure 4(a) shows the OSNR at the receiver for the path with WC, respectively. The architectures provide almost the same results. The configurations N = 16,

$M = 2$ and $N = 8$, $M = 4$ appear to be not feasible due to noise problems; the others are feasible, at least with a small number of WCs. Figure 4(b) plots the power level before SOA1 and SOA3, which are the most stressed devices from power viewpoint. The figure shows how SPIW performs better than SPOW in all configurations. All configurations are feasible, only SPOW with $N = 2$, $M = 16$ is below the threshold S_{SOA} for very high number of WCs.

Figure 4(c) shows the same results for $NM = 128$ channels. As expected, performance degrades in terms of both OSNR and power. All the configurations are difficult to be implemented with this target $OSNR_T = 20$ dB, for both architectures. Finally Figure 4(d) shows how for the SPOW the power before SOA1 and SOA3 is often under their sensitivity, while the SPIW is shown to be feasible in terms of power requirements.

5.5.1.2 Optical interconnects using microring-based demultiplexers

The contribution of PoliMI to this Joint Activity is based on a feasibility and scalability analysis of switching architectures with overall capacity in the order of Tbps [3]. The main effort of PoliMI has been put on the issue of interconnecting line cards of routers/switches, exploiting wavelength agility at transmitters to control signal routing across a Ring-Resonator based structure. In Y3, starting from the architecture described in the previous deliverable, PoliMI deeply investigated the characteristics of the aforementioned structure that is adopted as the central stage of the switching backplane. To better understand the work that has been done during Y3, some concepts already described in BONE deliverable D25.2 are now recalled.

The proposed optical interconnection architecture is presented in Figure 5, in which N transmitters are divided in S switching planes, each one comprising N/S transmitters. The multi-plane architecture [6] allows us to further improve scalability by overcoming limitations due to differential loss between the drop and through ports of the rings. Starting from the left side, a stage of N/S:1 couplers collects signals coming from each single plane. Beside the ring structure, there is a second stage of couplers that are needed to achieve connectivity between input/output ports belonging to different switching planes. S EDFA amplifiers are then placed after the couplers, and finally S demultiplexers route wavelengths to different output ports within each plane. These demultiplexers can be realized with a broadcast and select structure made of a coupler and a filter.

The ring structure for the plane selection can be built with a variable number of stages indicated as TS (Total number of Stages). This parameter is important, as well as N/S and S, because it influences the number and the characteristics of the devices connected downstream. However, it is necessary to employ a TS big enough to completely discriminate all the S switching planes. This result is achieved by constraining TS to the following value:

$$TS = \lceil \log_2 S \rceil$$

Assuming to have identical transmitters tunable on the same range of N wavelengths, it can be noted that a single wavelength is assigned to each receiver. Moreover, with this configuration, the wavelength used to discriminate a receiver is always the same independently of the transmission plane and of the position of the transmitter in the plane. Thus, to communicate with the i-th receiver of the k-th plane, the j-th transmitter of the h-th plane should be tuned on the following wavelength:

$$TX_{h,j} \rightarrow RX_{k,i} : \lambda_f,\ f = k + i \cdot 2^{TS};(h,k) \in [0; S-1],(j,i) \in [0; N/S-1]$$

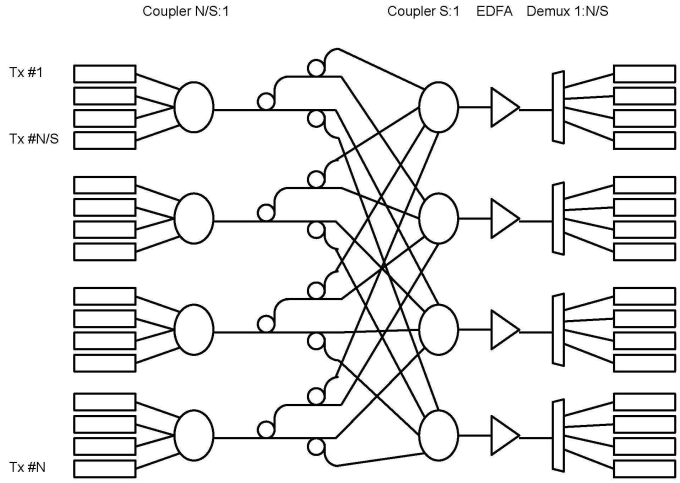

Figure 5. Ring-resonator-based architecture

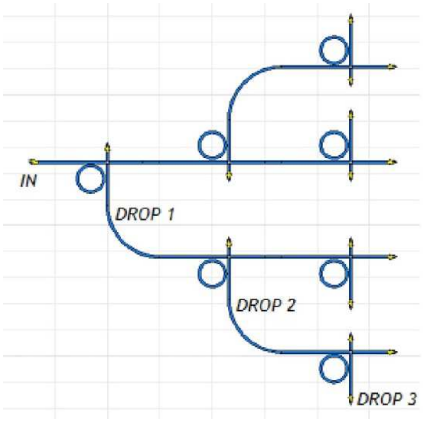

Figure 6. Ring-resonator Tree-based demultiplexing structure

As it can be appreciated from this wavelength assignment rule, the transmitter coordinates (*h* and *j*) are not present in the assignment.

The working principle of the microring-based demultiplexing structure is the same as the one on which the common cascaded Mach-Zender-filter demultiplexers are based: the *interleaving*. A 3-stage example of this architecture is shown in Figure 6. Rings are arranged in a tree-like topology, in which the *j*-th stage ($1 \le j \le \log_2 N$) contains 2^{j-1} rings. Rings belonging to same stage have equal radius (and thus equal FSR). *K* (eight) WDM signals, evenly spaced in wavelength by $\Delta\lambda$, enter the structure at the IN port, and they have to be demultiplexed and transmitted at the outlets (the output ports of all the rings of the last stage). In order to obtain one wavelength at each output port, the FSR parameter of the *j*-th stage should be set as it follows: $FSR_j = 2^j \cdot \Delta\lambda$ ($FSR_1 = 2\Delta\lambda$, $FSR_2 = 4\Delta\lambda$ and $FSR_3 = 8\Delta\lambda$). By adopting this approach, the FSR of the rings of the last stage (the maximum FSR in the network) is necessarily directly proportional to the number of managed wavelengths *K*. Since the FSR is inversely proportional to the ring radius, this implies that, as *K* increases, rings must be designed smaller and smaller. But given a certain technology, there is always a lower bound on the ring radius (below which waveguides are no longer able to confine optical beams and loss starts to rise very quickly) that imposes a scalability bound on *K*. To overcome this problem, in the tree-based structure, it is possible to adopt the *Vernier Effect* [7] which allows us to design the rings with a limited FSR while the periodicity of the transfer function remains suitable to correctly demultiplex the input channels.

PoliMI work has been also focused on solving another issue: in the design of a ring resonator it is necessary to impose the resonance condition of the device on the wavelength desired to be filtered from the ring. This centering action is made by manipulating the geometrical length of the ring, which is also inversely proportional to the FSR. Thus, the FSR is adjusted in order to be able to separate incoming channels but at the same time it is necessary to centre the resonator in order to let the rings perform their filtering action. Roughly, two equations are stressed, insisting on the same physical parameter.

The centering issue can be partially managed exploiting again the Vernier effect. To impose the ring resonance frequency means to translate the transfer function peak of the *drop* port from the "casual" position in which it is located after the initial design to the desired

resonance wavelength (Figure 7(a)). This lambda translation (Δλ) can be reduced in a variation of the physical length of the ring and therefore of its FSR (that in the worst case can be equal to $\Delta\lambda = FSR/2$). Here the Vernier Effect can be exploited to place some intermediate peaks inside the FSR of the transfer function. The presence of these peaks allows us to move, among them, the nearest peak to the wavelength to filter, thus reducing at most the Δλ (Figure 7 (b)).

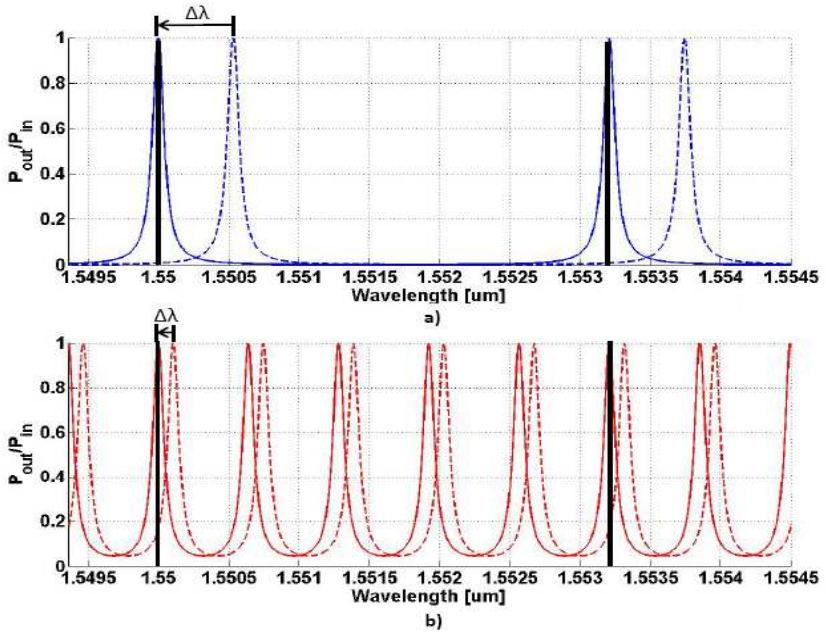

Figure 7. Effect of the centering procedure of the tree-based structure rings In two cases: ideal planning (a) or with Vernier effect exploitation (b).

After that, it has been found that it is possible to further optimize the tree-based ring structure parameters in order to improve the outgoing signal quality. First of all, it is necessary to clarify that reducing the FSR value of the rings with the Vernier Effect implies that the *crosstalk* values produced by all the other input channels of the demultiplexing structure are increased [7]. Thus, a procedure has been realized that optimizes the Vernier Effect adoption and the 3 dB bandwidth of the resonator with the aim of minimizing the crosstalk level.

Finally, PoliMI realized a modular simulator working with the transfer functions of the various devices of the architecture. This simulator has been used to validate results coming from the analytical worst-case estimation of the architecture scalability already described in D25.2. In particular, the simulator was employed to evaluate the effectiveness of both the centering and optimization procedures.

As mentioned in D25.2, scalability is evaluated in terms of total aggregated bandwidth by exploiting the model proposed in [6]. The aggregated bandwidth is calculated as the maximum number of allowed transmitters for the architecture multiplied by the bit rate of a single channel. The characterization of the different components taking into account physical impairments and power penalties have been defined in terms of: *(i)* average transmitted power for each transmitter, *(ii)* losses due to modulation and other effects, *(iii)* losses due to couplers, ring structures and demultiplexers. Two further constraints are considered: *(a)* signal power at the receiver must be greater than receiver sensitivity and *(b)* calculated Optical Signal-Noise

Ration (OSNR) at the photodiode must be greater than a target OSNR value that guarantees an acceptable BER performance.

The total aggregate bandwidth evaluation has been carried out considering $OSNR_{TX}$ values at the transmitter ranging from 40 dB to 70 dB. Ring resonators operating with spacing of 25, 50, 100 and 200 GHz between the WDM channels have been designed, and transmission rates of 2.5, 10 and 40 Gbps are considered. Aggregate bandwidth is directly influenced by the bit rate and by the value of *N/S* and *S*.

Figure 8 shows the throughput vs the $OSNR_{TX}$ for the different bit-rate/channel spacing configurations. Here the scalability is evaluated through the worst case analytical model. In particular, in results, PoliMI shows the comparison between the tree-based multiplexing structure in which Vernier effect is exploited with (case (b)) or without (case (a)) the optimization procedure. As it can be noted, the optimization procedure is effective in terms of aggregated bandwidth. Two different and distinct areas of the graphs identifying the bottlenecks of the architecture can be observed. The first one, characterized by curves with a linear slope, indicates that the limit to the scalability is given by the low OSNR values at receivers that cause an excessive noise accumulation through the various transmission planes until the signal reaches the EDFA and finally the receiver. On the contrary, the second part is characterized by a constant slope which indicates the scalability independence from the OSNR values. Here the bottleneck is the received power.

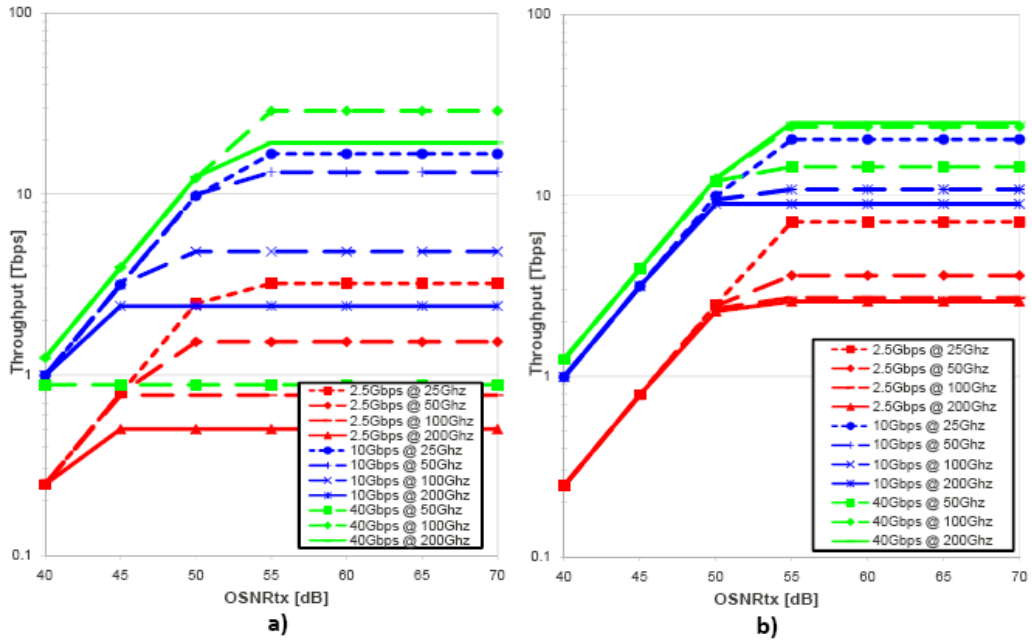

Figure 8. Analytical model: maximum throughput VS $OSNR_{TX}$ for all bit rate – channel spacing combinations.

Finally, Figure 9 shows the comparison between the analytical model (AM) and the simulation (SIM). The bit-rate /channel spacing configurations that give the best scalability were chosen. The first important outcome to highlight is the similarity between the results obtained with the worst-case analytical tool and with the simulator. The estimation values are always lower than the ones obtained with the simulation, this is due to the conservative attitude in the analytical analysis. Moreover, the optimization function effectiveness is also confirmed by these results.

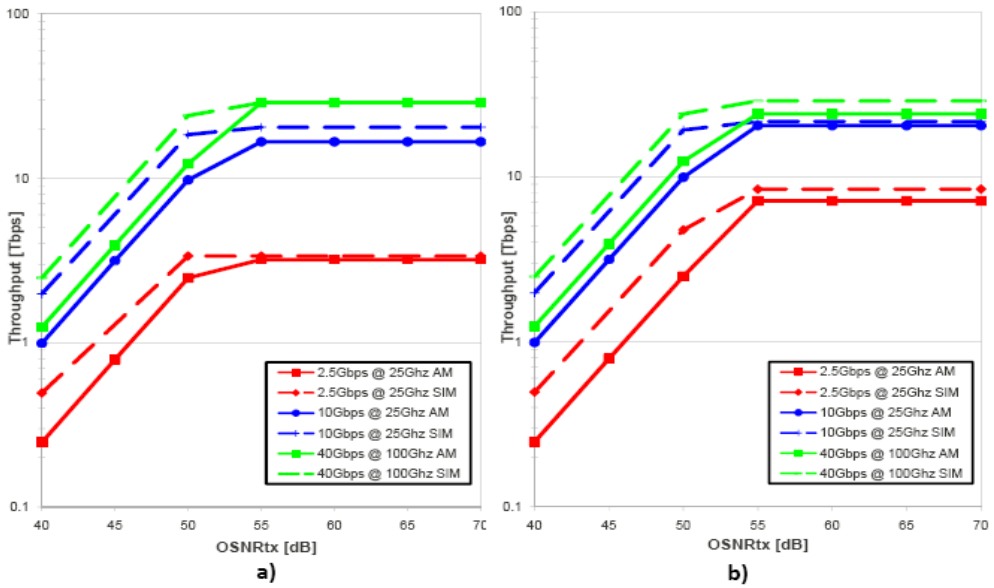

Figure 9. Analytical model VS simulations: maximum throughput calculation for some bit rate – channel spacing combinations.

The proposal shown in Figure 6 can be seen as an optical fabric that uses a demultiplexer based on microring structure. At the moment, PoliMI and PoliTO have been studying the use of other possible microring-based elements used as filters or even switching capabilities; a preliminary comparative analysis has been carried out with demultiplexers, in which the tree based architecture in Figure 6 is compared with the Bus-based structure in Figure 10. From the layout complexity point of view, both would have the same number of rings and the same reasoning can be applied to build the corresponding wavelength plan using the Vernier effect. The main differences appear in their transfer function. Since the Bus architecture is linear, the last drop ports are more penalized due to losses, and on the other hand, the first drop ports have a higher leakage power from the other ports, giving place to incoherent crosstalk. In both, Tree and Bus based demultiplexers there is a different scalability law; Figure 11 shows the advantage of Bus-based power penalties respect to Tree and the traditional AWG demultiplexer, for small demultiplexer sizes. The penalties due to incoherent crosstalk are comparable among all multiplexer types; but losses are the main limiting factor, since they are linearly dependent on the number of ports in the worst case, with respect to the logarithmic dependence of the AWG multiplexer [3], or the Tree (due to TS scalability law).

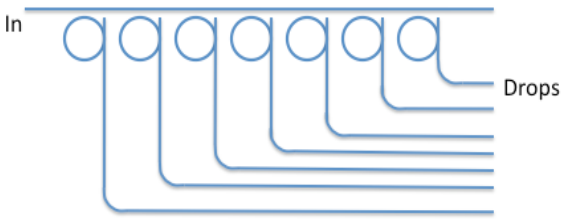

Figure 10. Bus-based demultiplexer

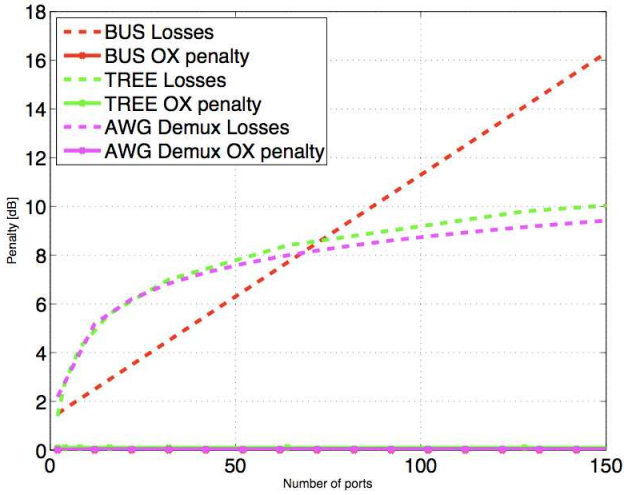

Figure 11. Comparison between power penalties of AWG-based, Microring Tree-based and Microring Bus-based demultiplexers

Currently this initial analysis is giving birth to other passive devices in the search for a higher scalability of the traditional architectures that PoliMI, UniBO and PoliTO have proposed. And it intends to be worth to produce results as a joint work.

5.5.1.3 Efficiency Assessment in OBS-CB networks

Within the scope of JA5, TUW continued the research work (on the basis of the first results presented in the previous reporting period) on scalability issues and performance evaluation of optical switching fabrics for optical burst-switched networks with channel bonding (OBS-CB). In OBS-CB networks, bursts are aggregated and transmitted on multiple wavelength channels in parallel, and therefore, the complexity and cost of core nodes can be reduced. Due to the fact that both the port count of core nodes and the temporal size of large data bursts are reduced, the scheduling algorithm applied is expected to be simpler than that used in conventional unbonded OBS systems with wavelength switching (OBS-WS). First results on the OBS-CB architecture were published in D25.2.

In particular, TUW extended the study by an estimation of the efficiency of OBS-CB networks when using the proposed switching fabrics [8]. The OBS-CB efficiency can be defined similarly to the waveband switching efficiency [9] as the relative reduction of the total number of optical ports in a network when OBS-CB is used instead of wavelength-switched OBS (WS-OBS). When assuming the case without wavelength conversion and separation of bonded channels in the intermediate core nodes, the number of switch ports needed to establish a lightpath over K wavelength channels and Q hops in a wavelength-switched OBS network, $n_{\text{OBS-WS}}$, and in an OBS-CB system, $n_{\text{OBS-CB}}$, can be calculated as follows:

$$n_{OBS-WS} = 2Ku(Q+1) \quad \text{and} \quad n_{OBS-CB} = 4Ku + 2\left(Q+1\right),$$

where

$$u = \frac{L_b}{\left(T_s + T_g\right)K}$$

denotes the average utilization of a burst slice, i.e., the efficiency of the burst assembly method used, while Q is the average hop number, T_s is the burst slice length, T_g is the guard time (see Figure 12). L_b is the mean burst length in the system. Thus, the efficiency is then given by

$$e = 1 - \frac{n_{OBS-CB}}{n_{OBS-WS}} = \frac{Q-1}{Q+1} - \frac{1}{Ku}$$

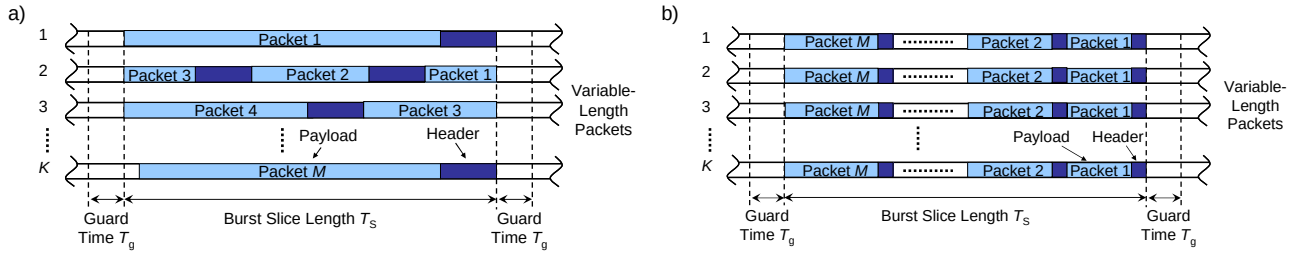

Figure 12. Channel-bonded burst assembly (a) with packet cut (b) with parallelized packets.

In general, there are several methods that can be applied for assembling and transmitting bursts in OBS-CB networks. The easiest one could be to collect a number of packets, let say *M* packets, and to transmit each packet as a whole on a wavelength channel, so that *M* packets are transmitted in parallel on *K* wavelengths. However, if the packet length varies –which is usually the case– the efficiency of this assembly scheme becomes very low. The burst transmission efficiency can be improved to some extent by collecting and transmitting several small packets on the same wavelength channel when the total length of the collected packets is smaller or equal to the burst slice length TS. Two more efficient burst assembly methods are depicted in Figure 12. The one shown in Figure 12(a) allows packets to be cut at boundaries of a burst slice. Here, only a small unused space can appear in the last channel *K*, which results in an improved transmission efficiency. The other method depicted in Figure 12(b) assembles packets belonging to a burst and transmit them bit-parallelized in a consecutive manner on all *K* wavelengths. The bit-parallel transmission allows maximum utilization of the burst slice length, but it requires bit-level synchronization of all *K* channels at the receiver.

Figure 13 shows the OBS-CB efficiency versus number of cascaded nodes for four different values of *K* (2, 4, 8, and 16) and two values of *u* (70% and 95%).

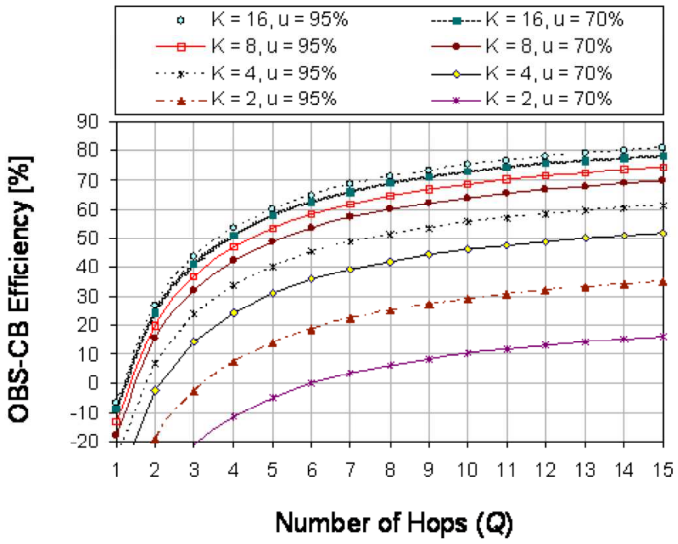

Figure 13. OBS-CB efficiency.

It is apparent from Figure 13 that a significant gain in efficiency can be achieved for more than three hops when using an efficient method for burst assembly on at least 4 bonded wavelength channels, i.e., when $K \geq 4$. However, increasing the parameter K, and consequently, decreasing the number of ports will most probably have an impact on blocking performance of OBS-CB nodes. Therefore, it would be of interest to find a trade-off between improving the burst-loss performance and decreasing the number of switching ports. In further work, TUW intends to evaluate blocking performance of OBS-CB network nodes and to find out the ranges of system parameters for which OBS-CB would be more efficient than WS-OBS by considering both the relative difference in the total number of optical ports and the difference in burst-loss probability.

5.5.2 References

[1] C. Raffaelli, M. Savi, A. Stavdas, "Multistage Shared-per-Wavelength Optical Packet Switch: Heuristic Scheduling Algorithm and Performance", *IEEE/OSA Journal of Lightwave Technology*, Vol. 27, No. 5, March 1, 2009.

[2] R. Gaudino, G.A. Gavilanes Castillo, F. Neri, J. M. Finochietto, "Can Simple Optical Switching Fabrics Scale to Terabit per Second Switch Capacities?", *Journal of Optical Communications and Networking*, Vol. 1, Issue 3, pp. B56-B69 (2009).

[3] R. Gaudino, G.A. Castillo, F. Neri, J. Finochietto, "Simple optical fabrics for scalable terabit packet switches," in *Proceedings of IEEE ICC*, May 2008, pp. 5331–5337.

[4] www.futurexusa.com/products.htm, www.accelink.com/weben/, www.avanex.com, www.avanex.com/Amplification.php, www.accelink.com/weben/index.jsp, www.inphenix.com/products.html.

[5] H. Takahashi, K. Oda, H. Toba, "Impact of crosstalk in an arrayed-waveguide multiplexer on NxN optical interconnection", *IEEE/OSA Journal of Lightwave Technology*, vol. 14, no. 6, pp. 1097–1105, June 1996.

[6] D. Cuda, R. Gaudino, G.A. Gavilanes Castillo, F. Neri, G. Maier, C. Raffaelli and M. Savi, "Capacity/cost tradeoffs in optical switching fabrics for terabit packet switches," in *Proceedings International Conference on Optical Network Design and Modeling 2009*, February 2009, pp. 1-6.

[7] E. J. Klein, "Densely integrated microring-resonator based components for Fibre-To-The-Home applications," *Ph.D. dissertation*, University of Twente, 2007.

5.5.3 Publications

[8] S. Aleksic, "Optically Transparent Core Nodes for Burst-Switched Networks with Channel Bonding", in *Proceedings of the World Telecommunications Congress 2010*, Vienna, Austria, September 2010.

[9] P. Torab, V. Hutcheon, D. Walters, A. Battou, "Waveband Switching Efficiency in WDM Networks: Analysis and Case Study", *Optical Fiber Communication Conference 2006*, Anaheim, California, USA, OTuG3, 2006.

[10] N. Akar, E. Kharazan, C. Raffaelli, M. Savi, "Shared-Per-Wavelength Asynchronous Optical Packet Switching: A Comparative Analysis", accepted for publication in *Elsevier Computer Networks*, 2010.

[11] C. Raffaelli, M. Savi, "Multi-granular traffic scheduling in hybrid optical interconnection", in *Proceedings Optical Network Design and Modelling 2010*, Kyoto, Japan, February 2010.

[12] C. Raffaelli, M. Savi, G. Tartarini, D. Visani, "Physical path analysis in photonic switches with shared wavelength converters", in *Proceedings International Conference on Transparent Optical Networks 2010*, Munich, Germany, July 2010.

5.5.4 Joint initiatives

Professor Fabio Neri (PoliTO) has supervised the research of Domenico Siracusa (PhD student at PoliMI) on Micro-Ring architecture scalability.

5.6 JA6: Optical backplanes utilizing microring resonators

Participants: UoA, UC3M, PoliTO, PoliMI

Responsible person: Antonis Bogris (UoA)

Deadline: Month 36

5.6.1 Description of the work carried out

5.6.1.1 Optical interconnection networks based on microring resonators

Each new generation of high-capacity routers and switches must process an always increasing amount of data traffic. Indeed, current trends are to increase the processed information densities while trying to maintain a reasonable power demand.

The International Technology Roadmap for Semiconductors [1] remarks that the most critical performance limitations of electronic interconnects depend on the latency and power requirements of metal wires, especially for lengths above the millimeter. On the other hand, the breakthroughs in CMOS-compatible silicon photonics are increasing the interest in optical technologies for interconnection networks.

In this research work, PoliTO proposes new solutions for large integrated optical switching fabrics, which can be used to interconnect routed/switch linecards. PoliTO is interested in the use of microring resonators as switching devices, such as 1×2 and 2×2 switching blocks [2].

Recalling the previous year's results, PoliTO uses three Switching Elements (SEs) as building blocks for proposing architectures:

- 1B-SE (Figure 1): The microring is coupled to two perpendicular waveguides. Optical signals entering the input port can be deflected to the drop port, when the ring is properly tuned (or in resonance) with the input signal wavelength, or just continue along the input waveguide towards the through port in the normal untuned state. The structure can be tuned or untuned thanks to the modulation of the refractive index of the material, which can be achieved by thermo-optic effects [3], carrier injection [4],[5] or optical pumps [6].
- 2B-SE (Figure 2): This SE exploits two 1B-SEs jointly controlled to provide two switching states. In the *bar* state (in_1 to out_1, in_2 to out_2), each ring deflects the corresponding optical input signal to the drop port of the respective 1B-SE. In the *cross* state (in_1 to out_2, in_2 to out_1), each ring lets the corresponding optical input signal pass to the through port of the respective 1B-SE.
- 2M-SE (Figure 3): This SE is a modified version of the 2B-SE. By cross-connecting input ports, the bar state is now achieved without tuning the microrings, and the cross state is achieved by tuning the microrings and deflecting the corresponding signals.

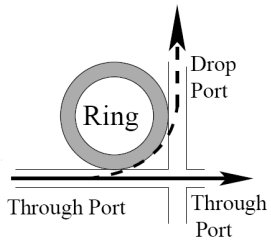

Figure 1. 1B-SE

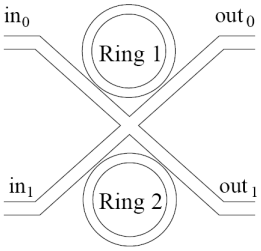

Figure 2. 2B-SE

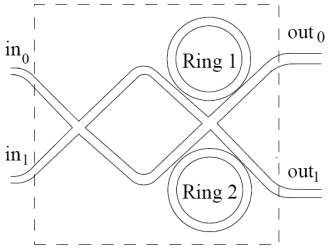

Figure 3. 2M-SE

The three SEs present an asymmetric behavior. Experimental measurements [2], [6] show that when a signal is coupled to the ring to reach the desired output of the SE it suffers a non-negligible power penalty (i.e., drop port in 1B-SE, bar state in 2B-SE and cross state in 2M-SE). Indeed, we call this state of the SE as High Loss State (HLS) and the complementary as Low Loss State (LLS, i.e., through port in 1B-SEs, cross state in 2B-SEs and bar state in 2M-SEs). Notice that to ensure a high scalability in terms of number of ports and in terms of performance (aggregate bandwidth), it is important to minimize the times in which this HLS occurs in the interconnection. In the following, we denote by X the number of SEs configured in HLS that a signal crosses in the worst case to satisfy any possible input-output permutation.

The target of PoliTO in this work is to reduce X by means of *(i)* the proposal of a proper architecture design technique, based on classical interconnection networks, and *(ii)* the proposal of a configuration algorithm able to further improve the performance of operation.

Microring Based Architecture Design

From the architectural point of view, PoliTO has first characterized in terms of cost (number of rings employed, *C*) and in terms of performance (architecture impairment, *X*) the classical crossbar (Figure 4) and Benes (Figure 5) networks. On the one hand, crossbar structure presents the best performance since all the signals cross only one 1B-SE in HLS (i.e., $X = 1$) but exhibits the worst scalability in terms of cost, since $C = N^2$. On the other hand, Benes networks exhibit a better scalability in terms of cost, since it increases as $C = NlogN$, but presents a worse performance, since in the worst case, there exists a path that passes through a SE in HLS in each stage.

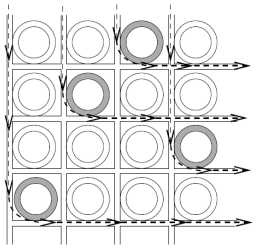

Figure 4. 4 × 4 microring-based crossbar connecting 1 to 4, 2 to 2, 3 to 1 and 4 to 3

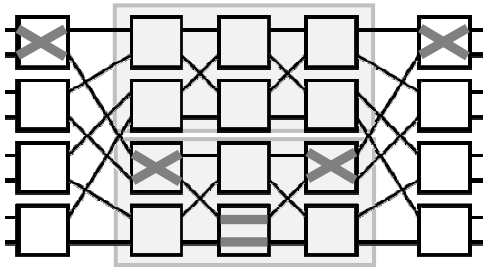

Figure 5. 8 × 8 Benes network, and sample connection setup from input 1 to ouput 1 using four cross states.

Based on these characterizations, PoliTO has proposed two hybrid architectures: the *Hybrid Crossbar-Benes (HCB)* which consists of a Clos network with the middle-stage modules based in Benes networks, while the first and third stage are crossbars; and the *Hybrid Benes-Crossbar (HBC)* which consists of a Benes network factorized until a certain recursion level in which the middle stage is substituted by crossbars [7].

On the other hand, PoliTO proposed the *mirroring* technique to improve the performance of microring-based Benes architectures. The mirroring technique uses the vertical dimension building two topologically identical Benes planes: one with 2B-SEs, and a second one implemented with 2M-SEs. In this way, the resulting architecture exploits the fact that the SEs configured in HLS in one plane correspond to SEs configured in LLS in the other plane. Hence, the mirroring technique roughly divides the physical impairment by a factor of two by doubling the cost of the architecture and adding microring-based plane selectors at each input. In the following, the mirroring technique is denoted by the prefix *M-*.

Table 1 summarizes the cost and the power penalty for all the considered architectures. We denote by $\hat{X}$ the maximum power penalty allowed in the architecture; it represents the target of maximum impairment.

Network	Cost	Power penalty
Crossbar	N^2	$1 \le \hat{X}$
Clos	$2\sqrt{2}N^{\frac{3}{2}}$	$3 \le \hat{X}$
Benes	$2N\log_2 N - N$	$2\log_2 N - 1 \le \hat{X}$
M-Benes	$4N\log_2 N$	$\log_2 N \le \hat{X}$
HCB	$\frac{2N^2}{2^{(\hat{X}-1)/2}} + N(\hat{X}-2)$	$3 \le \hat{X} \le 2\log_2 N - 1$
M-HCB	$\frac{4N^2}{2^{(\hat{X}-1)}} + 2N(2\hat{X}-3)$	$3 \le \hat{X} \le \log_2 N$
HBC	$\frac{N^2}{2^{(\hat{X}-1)/2}} + N(\hat{X}-1)$	$1 \le \hat{X} \le 2\log_2 N - 1$
M-HBC	$\frac{4N^2}{2^{(\hat{X}-1)}} + 2N(\hat{X}-2)$	$3 \le \hat{X} \le \log_2 N + 1$

Table 1. Performance comparison between different network architectures

Power Penalty Aware Routing Algorithm

Independently of the configuration algorithm, crossbar and crossbar-based Clos networks present a fixed power penalty of 1 and 3 respectively. On the other hand, for all the other networks considered, the power penalty depends on the path chosen by the routing algorithm in the configuration phase. Therefore, PoliTO did work on the design of a routing algorithm that is aware of the HLS, reducing the power penalty.

The well-known Paull algorithm [8] is often used to configure the connections in any multi-stage interconnection network based on recursive Clos construction. In this study, PoliTO has considered the use of the Paull algorithm for a pure Benes network, since it can be considered as the upper bound for the power penalty experienced by the other architectures reported in

Table 1. The Paull algorithm presents certain degrees of freedom when selecting the switching matrices for each connection at each recursion level. The modification of the Paull algorithm proposed by PoliTO exploits local decisions to choose for each connection the path that minimizes the power penalty. In particular, when multiple paths are available to establish a connection, Paull algorithm usually performs a random selection. PoliTO has modified this Paull's randomness to reduce the power penalty due to asymmetric behavior of the 2×2 SEs, giving priority to their LLS, whenever possible. This modified version of the algorithm is denoted by *PPA-Paull* (Power-Penalty-Aware Paull algorithm), in contrast with the classical version denoted simply by *Paull*.

As an example, the Benes network based on 2B-SEs shown in Figure 5 presents two choices when setting the path from 1 to 1 in the empty network: the first choice must be taken to configure the SEs at the first recursion level (first and fifth stages) and the second choice occurs at the second recursion level (second and fourth stages). Paull algorithm chooses randomly one of the 4×4 Benes blocks (shaded) while PPA-Paull deterministically chooses the lower 4×4 Benes block. Indeed, this choice implies configuring the SEs at the first and fifth stages in LLS instead of HLS. Analogously, Paull selects one of the third stage SEs inside the lower 4×4 Benes block randomly, while PPA-Paull deterministically chooses the lower SE. Thus, the SEs in the second and fourth stages are, again, in LLS instead of HLS.

Figure 6 shows the throughput (throughput losses are due to the unavailability of feasible paths) as a function of the maximum power penalty target $\hat{X}$ for a 64×64 Benes network (i.e., a Benes network with 11 stages) and different input loads: ρ = {0.1, 0.5, 0.9}. For enough large target $\hat{X}$, the throughput is maximum for both algorithms, since the routing is not affected by the relaxed power penalty. Smaller values of $\hat{X}$ reduce the possibility of finding feasible paths; in the extreme case, the throughput approaches zero. In general, PPA-Paull achieves always a better throughput than Paull.

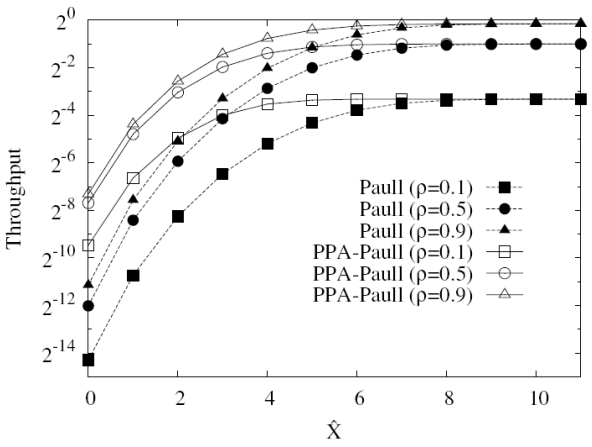

Figure 6. Throughput under uniform traffic and N = 64

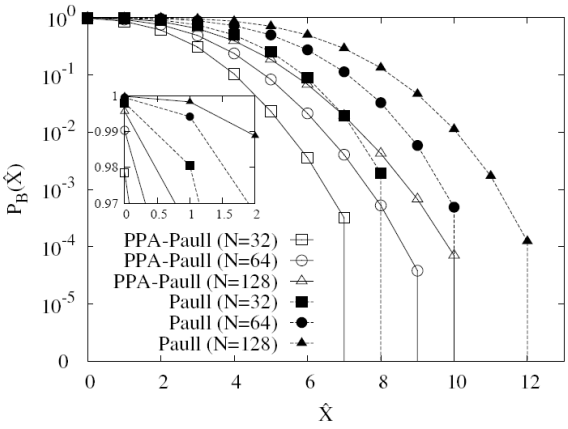

Figure 7. Blocking probability in Benes networks under uniform traffic with ρ=0.5

Figure 7 shows the blocking probability for a medium load and uniform traffic scenario, as a function of the maximum power penalty. The smaller plot, inside of each figure, details the blocking probability for low values of $\hat{X}$. Notice that the number of stages is 9, 11 and 13 for the N = 32, 64 and 128 Benes networks respectively. Whenever $\hat{X}$ is higher than the number of stages, the blocking probability is zero by construction and the maximum throughput is achieved. On the contrary, when $\hat{X}$ approaches zero, the routing is severely constrained by the

power penalty: the blocking probability increases and the throughput tends to zero. Furthermore, as *N* increases, the blocking probability increases due to the larger network depth. In general the reduction in the blocking probability due to PPA-Paull with respect to Paull is very large, reaching more than two orders of magnitude in some cases.

In conclusion, this part of the work *(i)* summarizes and compares several microring-based interconnection architectures already reported in the previous deliverable with their cost-performance laws and *(ii)* presents and analyses a modification of the Paull algorithm that further improves the performance of all the architectures defined by PoliTO.

Scalability Assessment of Microring Based Interconnection Networks

In addition to considering the physical impairment of interconnection networks as a discrete quantity, the physical impairments can be analyzed for the worst case of propagation, using (i) approximated microring transmission models and (ii) by means of travelling wave/transfer matrix simulation. In this direction, NKUA and PoliTO have been collaborating to test the above-mentioned microring-based structures with both approaches. The approximated microring transmission model for crossbar and Benes interconnection networks was presented by PoliTO in deliverable D25.2 and reported in a journal publication [9].

Regarding the simulative analysis, the static characteristics of 2B-SEs were studied with respect to the coupling coefficient (κ), which is a free parameter that quantifies the amount of field coupling between neighbouring waveguides; so κ is a measure of how much power is injected into the ring waveguides (optical cavity).

All the waveguides considered in simulations (micro rings and buses) are considered as built in Si with an effective index $n_{eff} \sim 3.5$. The microring radius is $R = 20$ μm which gives a free-spectral range $FSR = 681$ GHz. The total loss inside the cavity is 5 db/cm assuming the ON-state. For all simulations with modulated signals, both input ports are fed with signals of the same wavelength modulated with NRZ ON-OFF keying. For the detection, an ideal photo-detector followed by an electrical filter with 12.5 GHz bandwidth has been assumed.

The static transmission characteristics in the ON state are shown in Figure 8 and Figure 9 and they are extracted from the calculated transfer functions for the output ports. Considering operation at 10 Gbps, the required optical bandwidth of the device should be at least 20 GHz, which leads to an optimal coupling coefficient higher than 0.2. This value also leads to a reasonable value of rejection at 100 GHz away of the mode (23 dB), which ensures low crosstalk from adjacent channels in case of WDM operation. In particular for $k = 0.275$, the transmission model matches the measurements for a real device implemented in [6]; so this value is considered as the *working point* for the approximated model herein. Figure 9 shows the calculated transfer functions for the ON-state for a change in the refractive index $\Delta n \sim 10^{-3}$.

The next step of this work is to implement a Benes switching fabric incorporating the 2B-SE following the non-static approach, estimating the signal quality at the receiver. A 4×4 Benes fabric was configured accordingly, using the 6 blocks of 2B-SEs, and it is modeled as described in Figure 10 and Figure 11. All simulations consider signals with zero frequency detuning between them (that is, ideal laser conditions). Two cases are investigated: the worst loss, and the worst crosstalk following the work in publication [9].

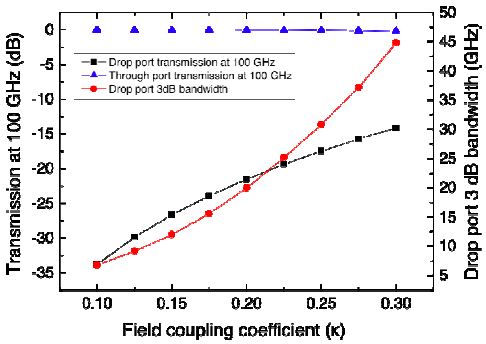

Figure 8. Transmission characteristics in ON state, with signal away from resonance

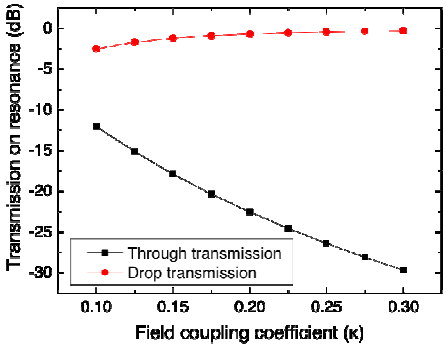

Figure 9. Transmission characteristics in ON state, with signal on resonance

The Loss Worst Case Path (LWCP) throughout the 4×4 Benes network is shown in Figure 10. This case considers the routing of an input signal to the output port assuming a path with all 2B-SEs at the ON-state: the input at port 2 is routed to the output port 2 with the 3 switches set at ON-state.

The Crosstalk Worst Case Path (XWCP, Figure 11) considers the routing of an input signal to the output port assuming a path with all switching blocks at OFF-state along the path: the input at port 2 is routed to the output port 4 through 3 2B-SEs set at OFF-state.

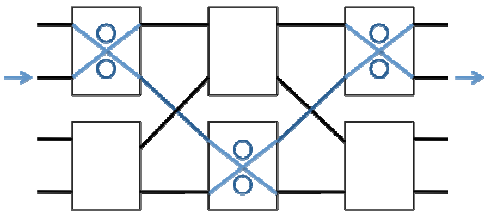

Figure 10. Loss Worst Case Path in a 4×4 microring-based Benes network

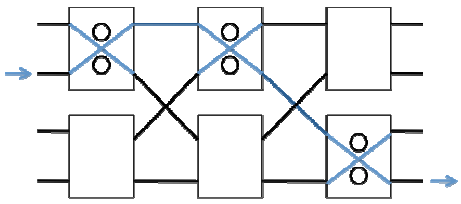

Figure 11. Crosstalk Worst Case Path in a 4×4 microring-based Benes network

The BER calculations for the 2B-SE is shown in Figure 12 (left), while the BER of the output signal on port 4 of the 4x4 Benes fabric is shown in Figure 12 (right) both versus the coupling strength of the MR switches. This first attempt of simulation showed acceptable performance for data rates of 2.5 and 5 Gbps. Regarding 10 Gbps rates, operation with moderate BER values is supported (BER quantified accurately) only for high coupling coefficients, where the notch transmission of the 2B-SE switch is higher.

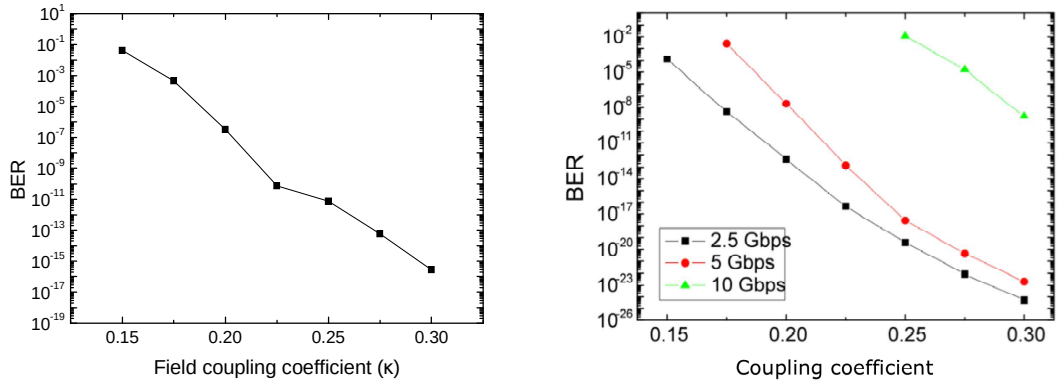

Figure 12. Left: BER calculations of the 2B-SE at 10 Gbps as a function of the coupling coefficient. Right: The corresponding BER calculations for the XWCP case for the 4×4 Benes network using 6 2B-SEs switching blocks as a function of the coupling coefficient.

Based on the performance of a single 2B-SE (2×2 interconnection network) in the working point, the output BER found at 10 Gbps, is just near 10^{-13} that is, practically an error-free transmission (considering a well accepted target of BER ~10^{-12}). Table 2 shows the different performances obtained by means of simulation for 2 different network sizes. Crosstalk was found to be the dominant limitation, since LWCP simulations showed quite ideal BER curves, whereas simulations of XWCP resulted more penalized. This behaviour was expected from the scalability model presented in [9], which empirically estimates a limited scalability due to the power penalty induced by crosstalk.

Switching fabric	2.5 Gbps	5 Gbps	10 Gbps
2×2 (2B-SE)	« 10^{-13}	« 10^{-13}	5×10^{-14}
4×4	« 10^{-13}	« 10^{-13}	1.6×10^{-5}

Table 2. Performance results for switching fabrics

In summary, several of the first simulation measurements obtained are presented here for optical interconnection networks based on microring resonators; UoA and PoliTO are giving continuity to this research, studying the use of new microring based switching blocks, such as the dilated 2×2 switching block (2D-SE) proposed by PoliTO in [9] as an alternative to boost scalability of the microring based interconnections.

5.6.2 References

[1] S. I. Association, *International Technology Roadmap for Semiconductors*, http://www.itrs.net/Links/2009ITRS/2009Chapters_2009Tables/2009_Interconnect.pdf

[2] B. G. Lee, A. Biberman, N. Sherwood-Dorz, M. Lipson and K. Bergman, "High speed 2x2 switch for multi-wavelength message routing in on-chip silicon photonic networks" *European Conference on Optical Communication (ECOC)*, May 2008.

[3] B.G. Lee, A. Biberman, N. Sherwood-Droz, M. Lipson, and K. Bergman, "Thermally active 4x4 non-blocking switch for networks-on-chip", *Annual Meeting of the Lasers and Electro-Optics Society (LEOS)*, p.TuBB3, Nov 2008.

[4] C. Li and A. Poon, "Silicon electro-optic switching based on coupled microring resonators," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2007.

[5] C. Li, L. Zhou, and A. W. Poon, "Silicon microring carrier-injection-based modulators/switches with tunable extinction ratios and OR-logic switching by using waveguide cross-coupling," *Opt. Express*, vol. 15, no. 8, pp. 5069–5076, Apr. 16, 2007.

[6] B.G. Lee, A. Biberman, P. Dong, M. Lipson, K. Bergman, "All-Optical Comb Switch for Multiwavelength Message Routing in Silicon Photonic Networks" *IEEE Photonics Technology Letters*, vol. 20, no. 10, 15 May 2008.

[7] A. Bianco, D. Cuda, M. Garrich, R. Gaudino, G. Gavilanes, P. Giaccone, F. Neri, "Optical Interconnection Networks based on Microring Resonators", *IEEE ICC 2010, Optical Networks and Systems Symposium*, Cape Town, South Africa, May 2010.

[8] J.Y. Hui, Switching and traffic theory for integrated broadband networks. Kluwer, 1990.

5.6.3 Publications

[9] A. Bianco, D. Cuda, R. Gaudino, G. Gavilanes, F. Neri, M. Petracca, "Scalability of Optical Interconnects Based on Microring Resonators", *IEEE Photonics Technology Letters*, Vol. 22, No. 15, pp. 1081-1083, August 2010.

[10] Bianco, D. Cuda, M. Garrich, R. Gaudino, G. Gavilanes, P. Giaccone, F. Neri, "Optical Interconnection Networks based on Microring Resonators", *IEEE ICC* 2010, *Optical Networks and Systems Symposium*, Cape Town, South Africa, May 2010.

5.7 JA7: Hardware efficient optoelectronic switch fabric

Participants: UCAM, PUT, TUE

Responsible person: Jose Bernardo Rosas Fernandez, M. Zol, Kevin Williams (UCAM)

Deadline: Month 36

5.7.1 Description of the work carried out

5.7.1.1 Hardware implementation of the control algorithm for monolithic optoelectronic multistage interconnection networks.

Switching network architecture and environment description

The purpose of this part of JA carried out by Poznan University of Technology (PUT) and Eindhoven University of Technology (TUE with assistance of the University of Cambridge was the hardware implementation of control algorithms for monolithic optoelectronic interconnection networks, which has been prepared by researchers from TUE. The considered switching network, as it is shown in Figure 1, contains 4 inputs and 4 outputs, and is composed of 16 semiconductor optical amplifiers (SOA). Formally, the switching network used in the experiment is the cross-bar switching network. The total area of this integrated circuit, i.e. the switching fabric and control interfaces use only 4 mm^2, presenting very high density of optical waveguides and electrodes. The control interface was used to connect SOA to the fast current driver, required to switch on and off selected SOA. The test bed used in this joined experiment is presented in Figure 2. The test bed can be divided into 2 parts: equipment of TUE laboratory (i.e. lasers, modulator, signal and pattern generators, sampling oscilloscope) and the control module based on the FPGA board with Virtex-5 (ML 505).

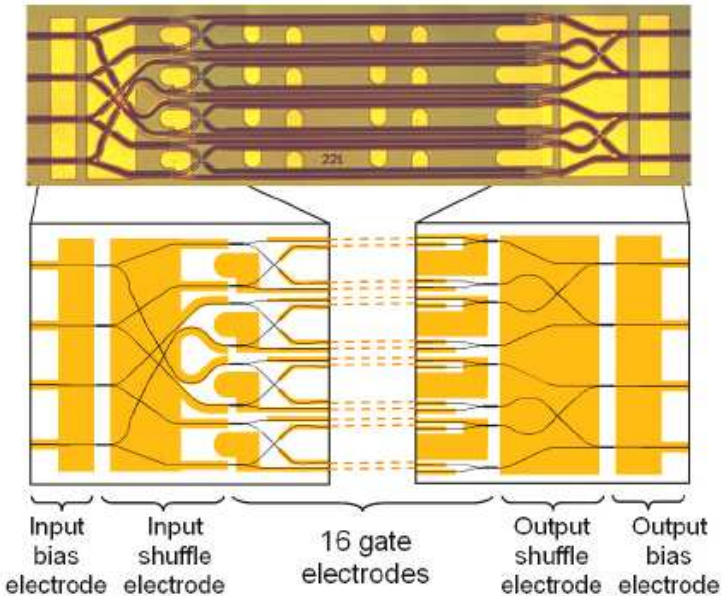

Figure 1. The 4x4 cross-bar switching network based on SOA

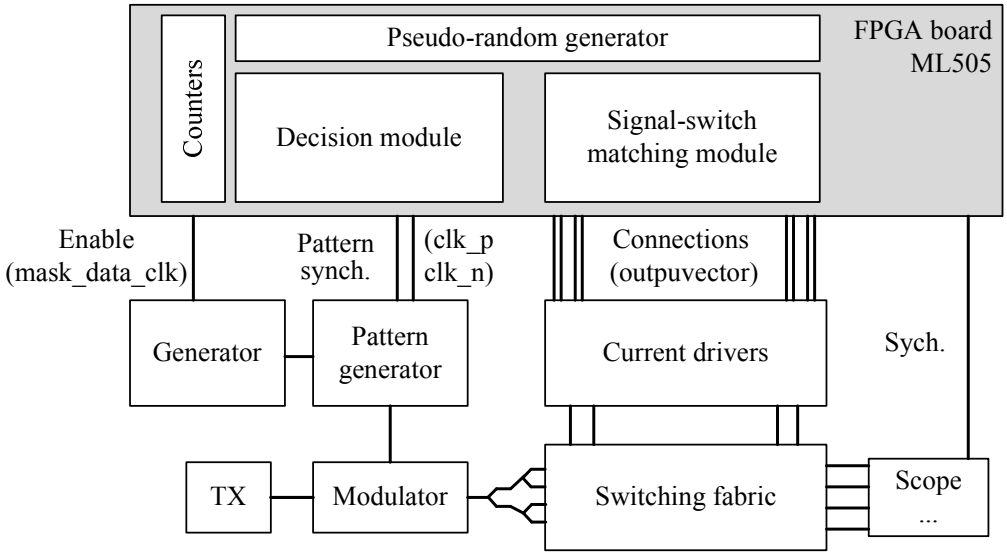

Figure 2. The test bed

Experiment description

Together with researchers from TUE several experiments realizing following scenarios were prepared:

- manual path selection,
- pseudo-random path selection,
- one-by-one path selection (based on Johnson counter),
- pseudo-random path selection with resolving problem of external blocking in the switching network.

All experiments were prepared as control algorithms implemented in VHDL language. Originally, these control algorithms should allow to set-up a connecting path for a new packet/connection in one cycle of an external clock, derived from a laboratory generator and connected into the FPGA board. Synchronization between the internal FPGA clock and the external clock was done by signal-switch matching module. To confirm correctness of VHDL code simulation tests were prepared. Next, VHDL codes were synthesized and implemented in the FPGA structure.

The manual path selection (pseudo)algorithm was required for optical path optimization, i.e. for the first end very important step in preparation of other experiments. The pseudo-random path selection algorithm, based on LSFR (Left Shift Feedback Register) modules, generates pseudo-random couple of number of range from 0 to 3, which represent a connection in the switching network. The aim of this experiment was observation of the case, when signals from two inputs are directed to one output. It should result in the increased power level at receivers. The timing diagram presenting the example of simulation result for pseudo-random path selection algorithm is shown in Figure 3.

In the one-by-one path selection, based on the Johnson counter, in every time slot the control module sequentially sets up four unicast connection paths from four inputs to four outputs. In this experiment the power level of received signals at every receiver were always the same. The timing diagram presenting the example of simulation result for the one-by-one path

selection algorithm is shown in Figure 4. All clock signals are input signals. In one time slot, it means in one cycle of “clk_p” (or “clk_n”), one pattern of connections was generated by the FPGA control module and was sent to the fast current driver through 16 bits different bus “outputvector_p” and “outputvector_n (not presented in Figure 3). Every pattern consists of four signals equal to “1” and 12 signals equal to “0” and in every time slot exactly one input was connected with every output. Because switching network is nonblocking one, in each time slot four optical signals appearat outputs . Time traces of the routed data are shown in Figure 5.

The last experiment, the pseudo-random path selection with resolving the problem of external blocking in the switching network was not implemented, because of insufficient time for this experiment. The aim of this experiment was to observe the case, when one of two connections directed to one output is set up in the switching network. As it is easy to imagine, sometimes the power level of received signals should be decreased, because one of two connections must be lost.

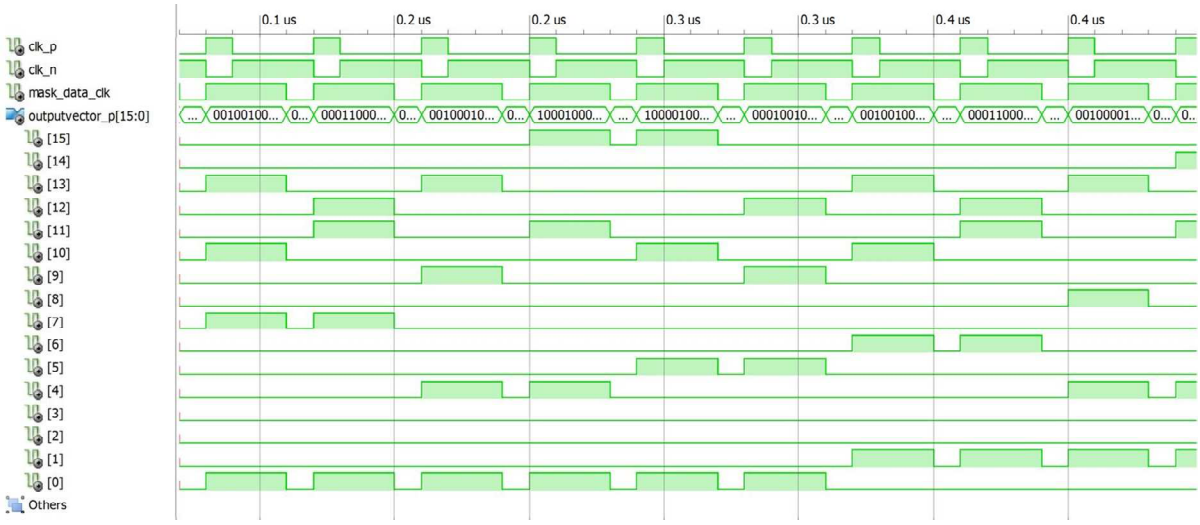

Figure 3. The timing diagram of the pseudo-random path selection

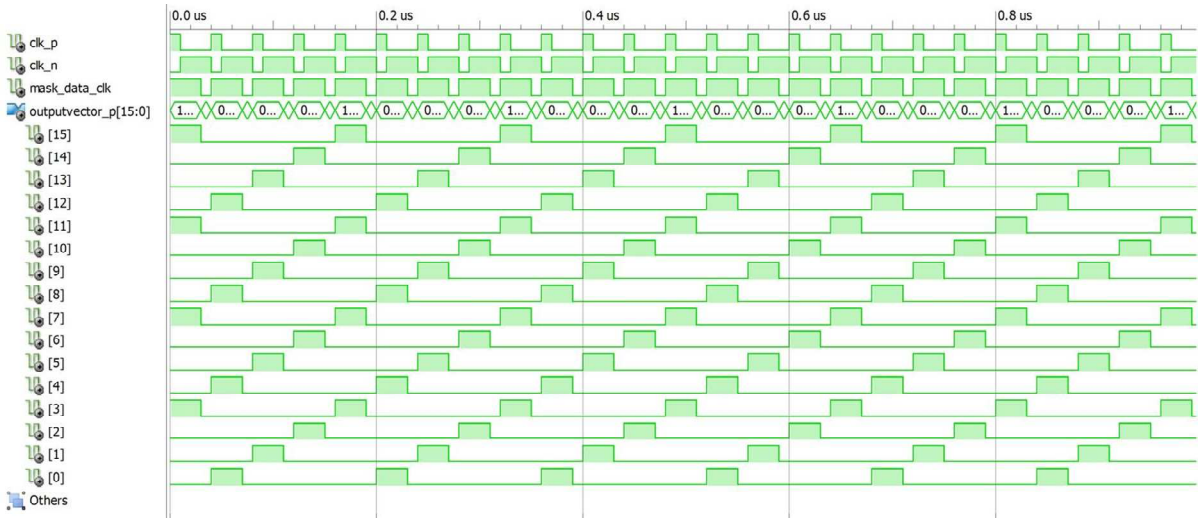

Figure 4. The timing diagram of the one-by-one path selection algorithm

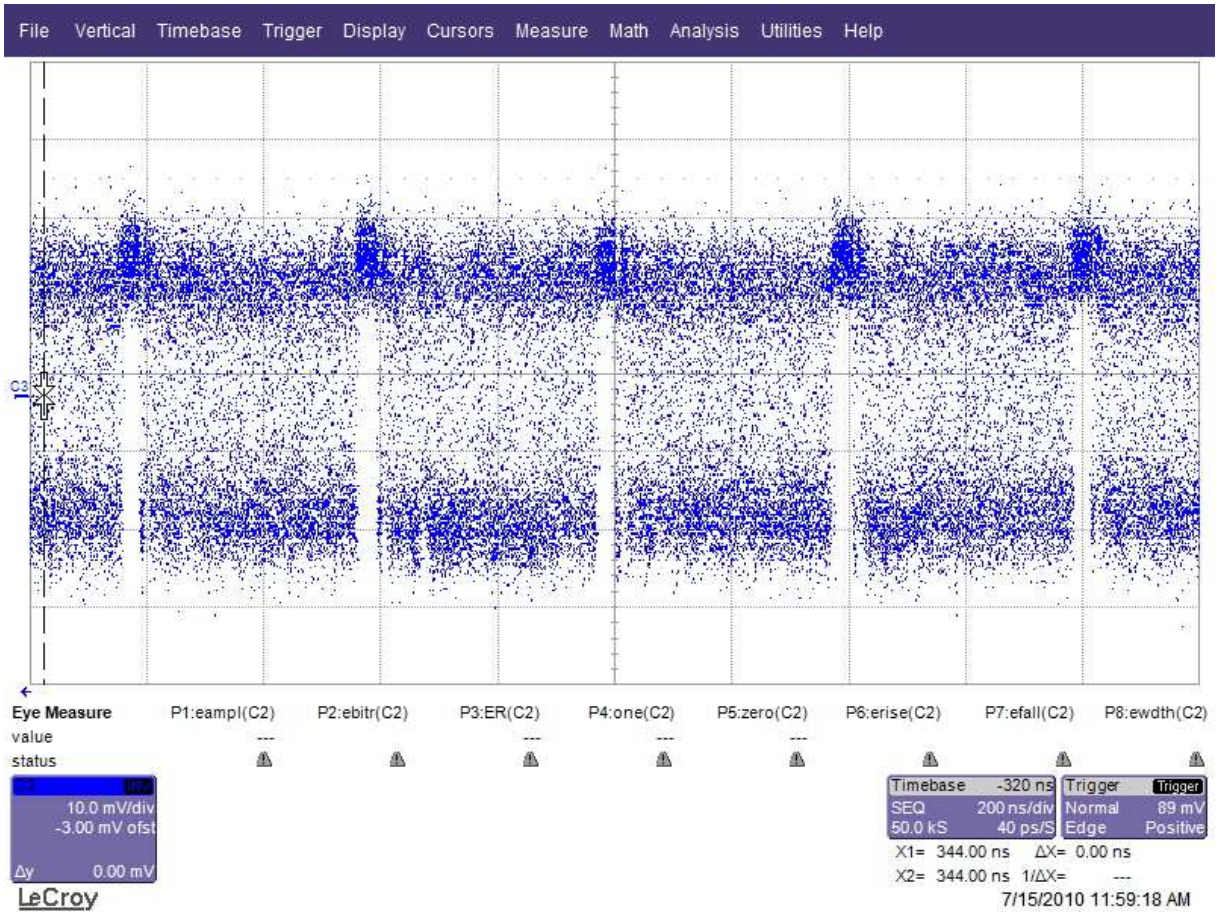

Figure 5. Time traces of the one-by-one path selection algorithm

5.7.1.2 Lossless Monolithic 16×16 QW semiconductor optical amplifier

In recent years, the demand for optical networks allows to rapidly re-route signals between moderate numbers of input and output fibres and it has encouraged research into the use of photonics in switching. As optical switches using MEMs technology and the like have limited speed, semiconductor optical amplifier (SOA) based photonic switches have attracted growing interest owing to their ability to achieve lossless switching of high capacity data with nanosecond switching timescales. Such switches have recently undergone much development, with integrated 2×2, 4×4 and 8×8 port switches being reported, the latter being achieved with multiple integrated 1x8 switching elements. Scaling further requires unprecedented levels of component integration. However, for practical applications, devices with port counts of at least 16×16 are needed for systems applications. Recently, feasibility studies have shown that 16×16 port counts can be achieved using cascaded 4x4 devices with fibre interconnections.
In this WP, it has been developed the first monolithically integrated 16×16 port SOA-based optical switch, incorporating ~1100 individual components. It consists in a 3-stage architecture as it is shown in Figure 6. The switch is rearrangeably non-blocking and has a chip area of 40 mm^2

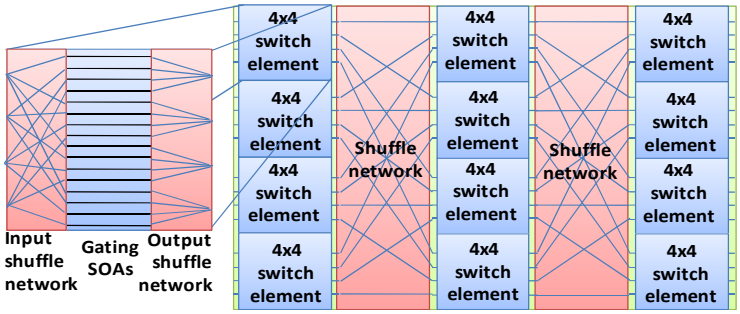

Figure 6. Schematic of 3-stage 16×16 switch architecture based on 4×4 switch elements

In Figure 7, it is shown the schematic of the 16×16 switch showing the 4×4 switch elements and shuffle network sections. The switch has a facet-to-facet gain of 2 dB at 1557.5 nm with fibre coupling losses measured to be 4 dB per facet. The TIR mirror loss has been measured to be 4 dB. The 3 dB on-chip output saturation power is -6.3 dBm with a 3 dB spectral bandwidth of 9 nm. The switch output has an in-band optical signal-to-noise ratio (OSNR) of 15.5 dB.

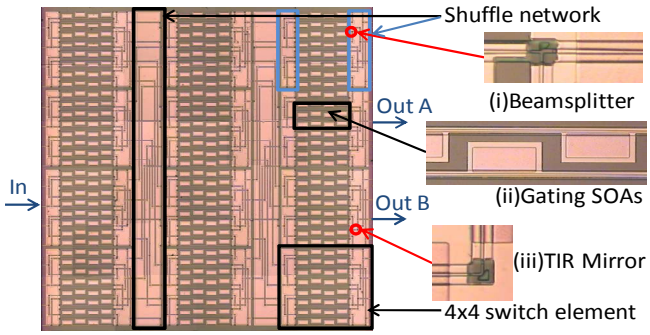

Figure 7 Schematic of the 16×16 switch showing the 4x4 switch elements and shuffle network sections (left). Insets show (i) waveguide beam splitter, (ii) gating SOA waveguides and (iii) TIR turning mirror.

As can be seen from Figure 8(a), the switch operates with a power penalty of 2.5 dB at a BER of 10^{-9} for an on-chip input power of -15 dBm. An input power dynamic range (IPDR) of more than 4 dB is measured for a power penalty of less than 4 dB, as can be seen in Figure 8(b). Due to its high capacity of routing it can be calculated that it consume 1mA/Gbps.

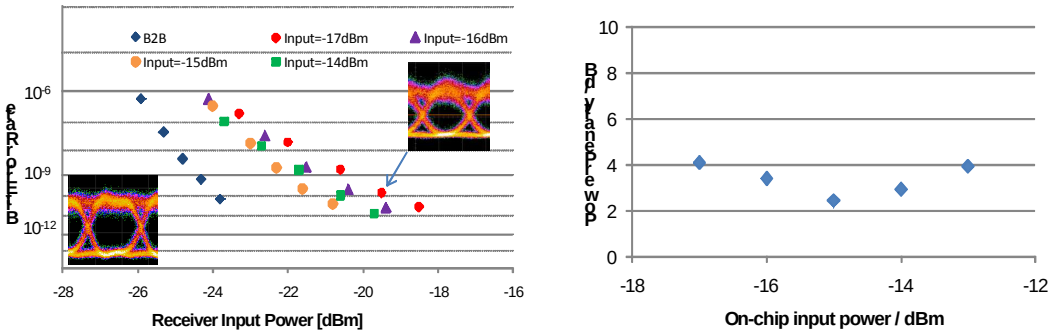

Figure 8: (a) Bit error rates and eye diagrams showing low power penalty for a number of different input powers (b) Power penalty as a function of on-chip input power.

Conclusions

The hardware implementation of control algorithms are always hard work and always show new problematic areas. First, the processing of any input signal (even very stable) adds jitter to the output signal (result). The jitter value can be decreased only by improving electrical connections. The hardware implementation of control algorithms with many input clock signals should be realized on FPGA board, which allow to connect clock signals to the main clock bank. In the other case, clock signals are degraded. The simulation of VHDL code confirms correctness of this code, but not correctness of hardware implementation. Also we have developed the 1st 16×16 monolithic switch based in SOAs. It has OSNR of 15.5 dB power penalties of 2.5 dB at a BER of 10^{-9} for an on-chip input power of -15 dBm.

5.7.2 Joint initiatives

Remigiusz Rajewski took part in JePPIX course/training organized by Eindhoven University of Technology. During this course he was studying design of photonic circuits and photonic integration technology. JePPIX is a technology platform for the development and promotion of Application Specific Photonic Integrated Circuits (ASPICs) in Indium Phosphide. JePPIX aims at a generic integration technology in which a variety of photonic integrated circuits can be designed and fabricated using the same process. ASICs are well known in Electronics but in Photonics they are new. They can lead to huge cost reductions by realising PICs for a broad variety of applications using a small set of basic building blocks that are integrated in a generic foundry process.

ASPICs will become a major driver for innovation in a broad variety of products, by integrating advanced optical functionality in a low-cost Photonic IC.

Remigiusz Rajewski works at the new optical switching fabric architecture and cost reduction is very important for him, since the switching networks used not electrical but optical elements. That is why the JePPIX is so helpfully. It gives possibility to design a new structure and do not spent extra money for physical implementation. Using JePPIX it is easier to find errors and eliminate them during design process. When this process is finished and everything works fine, then physical implementation starts. It allows save money and time since physical process takes few months.

Remigiusz Rajewski is planning to continue cooperation with Eindhoven University of Technology and he wants to implement his structure in indium phosphide or in the similar source.

Patty Stabile from Eindhoven University of Technology performed collaborative experiments at Cambridge University where she assessed the photonic integrated switch circuits she made at Eindhoven in the testbed configured at Cambridge. This lead to a joint publication at ECOC 2010 on fully scheduled switching fabrics.

Kevin Williams from Eindhoven University of Technology made regular visits to Cambridge University to develop avenues of research in photonic integration and optical switching systems.

5.7.3 Publications

[1] G. Danilewicz, W. Kabaciński, R. Rajewski, “The $\log_2$N-1 Optical Switching Fabrics”, *IEEE Transactions on Communications*, to be published.

[2] G. Danilewicz, R. Rajewski, "The New Modified Banyan-Based Switching Fabric Composed of Symmetrical and Asymmetrical Switching Elements", submitted to *IEEE International Conference on Communications 2011 (ICC 2011).*

[3] W. Kabacinski, M. Zal, "Banyan Type Switching Network with Wavelength Routing", submitted to *15th International Conference on Optical Network Design and Modeling 2011, Bologna, Italy.*

[4] R. Stabile, H. Wang, A. Wonfor, K. Wang, R.V. Penty, I.H. White and K.A. Williams, "Multipath routing in a fully scheduled integrated optical switch fabric", proceedings European Conference on Optical Communications, We.8.A.6, 2010

[5] H. Wang, A. Wonfor, K.A. Williams, R.V. Penty, I.H. White, "A compact lossless integrated 16x16 QW SOA switch", proceedings European Conference on Integrated Optics, ThG2, 2010.

5.7.4 Mobility actions

No.	**Brief Description of Mobility**	**Participating partners (Indicate the host institute)**
1	Dates: 05.07 – 16.07.2010 Visitor: Mariusz Zal (Assistant Professor at PUT) Host: TUE Hardware implementation of the control algorithm for monolithic optoelectronic multistage interconnection networks in the Joint Activity "Hardware efficient optoelectronic switch fabric".	TUE