



SEVENTH FRAMEWORK PROGRAMME

Project Number: FP7-ICT-2007-1 216863

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

CEC Deliverable Number: FP7-ICT-216863/RACTI/R/PU/D14.1

Contractual Date of Deliverable: 30/11/09

Actual Date of Delivery: 30/11/09

Title of Deliverable: D14.2: Report on Y2 activities and new integration strategy

Workpackage contributing to the Deliverable: WP14 : Virtual Center of Excellence on Optical Switching Systems (VCE S)

Nature of the Deliverable : R (Report)

Dissemination level of Deliverable : PU (Public)

Editors: RACTI / Kyriakos Vlachos
COM-DTU / Martin Nordal Petersen

Abstract:

This document is the report of the second year of activities in the VCE-WP14 Virtual Center of Excellence on Optical Switching Systems (VCE S). It includes the achievement and also reports the status of the joint activities proposed within VCE-WP14.

Keyword list:

Wavelength Converter Usage Reduction, Switching for Network Recovery, Quality of Service in switches, New switching paradigms, Optical Clock Recovery, Wavelength Conversion by nonlinear effects, Optical Multicast Architecture, Switching and network reliability, benchmarking and cost analysis, OCDM encoders/decoders, 2R Regeneration, Optical flip-flops, Optical packet switching, Hybrid Switch Architectures, GMPLS optical switch nodes, Contention Resolution Schemes, Optical Buffering, OTDM time-slot switching, Multi-wavelength regeneration, Optical Cross Connects, Power Issues in Switching Systems.



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

DISCLAIMER	2
TABLE OF CONTENTS	3
1. EXECUTIVE SUMMARY :	4
2. INTRODUCTION	5
2.1 VCE-S PARTICIPANTS	6
3. LIST OF JOINT ACTIVITIES IN VCE-S, WP14	7
4. TECHNICAL REPORT ON VCE-S JOINT ACTIVITIES	9
4.1 JA1 - POWER-COST-EFFECTIVE NODE ARCHITECTURE FOR MULTICAST LIGHT-TREE ROUTING IN WDM NETWORKS	9
4.2 JA2 - FEASIBLE PARALLEL SCHEDULERS FOR OBS/OPS NODES	11
4.3 JA4 – PERFORMANCE OF OPTICAL SWITCHING SYSTEM ARCHITECTURES WITH SHARED WAVELENGTH CONVERTERS	16
4.4 JA5 - CODE-BASED OPTICAL NODES	20
4.5 JA6 – REPACKING AND REARRANGING ALGORITHMS FOR MULTI-PLANE BANYAN TYPE SWITCHING FABRICS	27
4.6 JA7 – POWER CONSUMPTION AND SUPPLY OF HIGH-PERFORMANCE SWITCHING ELEMENTS	31
4.7 JA9 - ALL-OPTICAL LABEL PROCESSING TECHNIQUES FOR ULTRA-FAST OPTICAL PACKET SWITCHES	34
4.8 JA11 – MULTI-GRANULAR DESIGNS OPTICAL SWITCH ARCHITECTURES USING QD-SOAS FOR OPTICAL BUFFERING AND	36
4.9 JA12 - ENCOMPASSING SWITCH NODE IMPAIRMENTS AND CAPABILITIES IN DYNAMIC OPTICAL NETWORKS	39
4.10 JA14 - PHOTONIC CODE LABEL PROCESSORS FOR ULTRAFAST ROUTING	44
4.11 JA16 - LOW-CROSSTALK OPTICAL PACKET-SWITCHING ARCHITECTURES BASED ON WAVELENGTH-SWITCHING AND WAVELENGTH-SENSITIVE DEVICES	47
4.12 JA17 - NOVEL MULTI-GRANULARITY OPTICAL SWITCHING NODE WITH WAVELENGTH MANAGEMENT POOL RESOURCES	61
4.13 JA18 – COMPARISON OF THE SYNCHRONOUS/ASYNCHRONOUS OPERATION PARADIGM IN OPTICAL SWITCHES	63



1. Executive Summary :

This document is the second deliverable from the work package “Virtual Center of Excellence on Optical Switching Systems (VCE-S)”.

The main objectives of this deliverable are to collect the partner expertise and describe the status of their planned joint activities in the framework of BONE project. Since the last deliverable, some of the JAs are in the process of being finished while others continue. There has been one new JA proposal in 2009.

In the first deliverable sixteen (16) joint activities were reported to be under planning and development and of these 10 are active (including one new) and two will close at the end of this year. Another five JAs have been moved to other work packages as it was believed that the technical content fitted better into the scope of other WPs. JA3, JA8, JA13 and JA15 are now located in WP25 while JA7 in WP21.

Generally this WP is running very satisfactory with a high level of participation and activity. WP14 integrates 39 partners and 119 researchers. 29 joint papers have been reported to date. One less satisfactory point to mention is the number of mobility actions that only amounts to 6 in 2009 (two of these are in progress) . Therefore, an encouragement was given to increase in mobility actions in 2010 during the last WP14 technical meeting in Poznan, Poland.



2. Introduction

Recent technology development has unlocked most of the fiber capacity making capacity an abundant resource. However, the majority of WDM (Wavelength Division Multiplexing) deployment has occurred in the form of point-to-point links with amplifiers in between if needed.

Optical WDM light paths are static and are seen as a scarce resource. Once set up, they remain in place, essentially forever. Therefore, switching is only to transform the raw bit rates into useful bandwidth.

The prime objective of this deliverable is to provide the viewpoint on photonic switching of the “Virtual Center of Excellence on Optical Switching Systems (VCE-S)”. Within these directions, VCE-S must craft R&D directions and define the position of photonic switching in future optical networks. The questions that must be answered are how and where photonic switching is positioned in future Internet and which problems must be solved to reach this goal. The incentive is to use optical switching to make spectrum efficient in terms of switching and not in terms of only capacity. Towards this goal VCE-S has charted a list of key issues and prime research objectives. Actual implementation is carried out during the project via a number of collaborative projects (JAs).



2.1 VCE-S Participants

The following table displays the MM spent by partners in Y2.

Participant	WP14	Participant	WP14
	Effort spent		Effort spent
1- IBBT	0.3	31-SSSUP	0.6
2- TUW	0	33 -UNIMORE	0.5
3- FPMs	0	34 - UniRoma1	0.3
4 - Fraunhofer	0.1	35 - TELENOR	0
5 - TUB	0	36 - TUE	0.2
6 - UDE	0.5	37 - IT	0
7 - UST-IKR	0	38 - AGH	0.25
8 - COM	0	39 - PUT	0
9 - CTTC	0.1875	40 - HWDE	0.8
10 - TID	0	41 - KTH	0.5
11 - UAM	0	42 - BILKENT	0
12 - UC3M	0.225	43 - UniRoma3	0.5
13 - UPC	0.37	44 - ORC	0
14 - UPCT	0	45 - UCAM	1
15 - UPVLC	0.45	46 - UCL	2.8
16 - UVIGO	0.53	47 - UEssex	0.38
17 - FT	0.78	48 - USWAN	0.45
18 - GET	0.13	49 - Ericsson	0.5
19 - AIT	0		
20 - ICCS-NTUA	0		
29-PoLIMI	0.5		



3. List of Joint Activities in VCE-S, WP14

JA#	JA title	JA-Leader	Participants	Status
JA1	Power-Cost-Effective Node Architecture for Multicast Light-Tree Routing in WDM Networks.	Gonzalo F.D. Carpio gmfernand@hotmail.com	UC3M, UPVLC	Active
JA2	Feasible parallel schedulers for OBS/OPS nodes.	Pablo Pavón Mariño Pablo.Pavon@upct.es	UPCT, UVIGO, DEIS-UNIBO	Active. Will continue in Y3
JA3	Performance and complexity analysis of optical switching fabrics	Fabio Neri neri@polito.it.	PoliTo, UniBO, PoliMI, TUW	Moved to WP25
JA4	Performance of optical switching system architectures with shared wavelength converters.	Nail Akar akar@ee.bilkent.edu.tr	Bilkent, UNIBO, UniRoma1, UNIMORE, FT, KTH	Active. Will continue in Y3
JA5	Code-based optical nodes.	Gabriella Cincotti, cincotti@uniroma3.it	UNIROMA3, UNIMORE, UNIBO, UNIROMA1, GET, NICT	Active. Will continue in Y3
JA6	Repacking and rearranging algorithms for multi-plane banyan type switching fabrics.	Wojciech Kabacinski, wojciech.kabacinski@et.put.poznan.pl	PUT, POLIMI	Active. Will continue in Y3
JA7	Power Consumption and Supply of High-Performance Switching Elements.	Slavisa Aleksic slavisa.aleksic@tuwien.ac.at	TUW, UPCT, BME, UNIMORE, PoliTo	Moved to the new WP21
JA8	All-optical switches utilizing microring resonators.	Adonis Bogris abogris@di.uoa.gr .	UoA, UC3M	Moved to WP25
JA9	All-optical label processing techniques for ultra-fast optical packet switches.	Piero Castoldi castoldi@sssup.it	SSSUP, UNIROMA3, NICT	Active. Action will close by Y2
JA10	Currently inactive / merged into JA12			
JA11	The Optical Switch Architecture with Recirculation Buffer and Wavelength Conversion.	Wojciech Kabaciński wojciech.kabacinski@et.	PUT, RACTI	Active. Action will close by Y2



		put.poznan.pl		
JA12	Encompassing switch node impairments and capabilities in dynamic optical networks	Nicola Andrioli nick@sssup.it	SSSUP , DTU	Active. Will continue in Y3
JA13	Reliability of various optical switching technology	Rebecca Chandy rebeccapchandy@gmail.com	Ericsson , KTH	Moved to the new WP25
JA14	Photonic code label processors for ultrafast routing	Ian White ihw3@cam.ac.uk	UCAM , TUE, Uniroma3	Active. Will continue in Y3
JA15	Hardware efficient optoelectronic switch fabrics	Ian White ihw3@cam.ac.uk	UCAM , TUE, PUT	Moved to the new WP25
JA16	Low-crosstalk optical packet-switching architectures based on wavelength-switching and wavelength-sensitive devices	Achille Pattavina pattavina@elet.polimi.it	POLIMI , UPCT, POLITO, OVIGO	Active. Will continue in Y3
JA17	Novel Multi-granularity Optical Switching Node with Wavelength Management Pool Resources	Yabin Ye yabin.ye@huawei.com	HUAWEI , RACTI	New Joint Activity
JA18	Comparison of the synchronous/asynchronous operation paradigm in optical switches	Fabio Neri neri@polito.it.	POLITO , RACTI	New Joint Activity

4. Technical Report on VCE-S Joint Activities

4.1 JA1 - Power-Cost-Effective Node Architecture for Multicast Light-Tree Routing in WDM Networks.

Responsible partner: UC3M (Carmen Vázquez, cvazquez@ing.uc3m.es)

Participants: UC3M, UPVLC

Description of work carried out.

The work carried out has focused on the design of a novel cost-effective multicast-capable optical cross connect (MC-OXC) node architecture that features both tap-and-continue and tap-and-binary-split functionality. This architecture intends to provide an interesting balance between simplicity, power efficiency and overall wavelength consumption with respect to models based on TaC (Tap and Continue) or SaD (Split-and-Delivery). The main component of this node is a novel Tap-and-2-Split Switch (Ta2S). We propose and analyse an implementation of this switch based on integrated optics (namely, MMI taps and MZI switches), and we characterize and compare it with other alternatives implemented with the same technology. The study shows that, thanks to the presented Ta2S design, the 2STC node scales better in terms of number of components than the other alternatives. Furthermore, it is more power efficient than the SaD design and requires less wavelengths than TaC thanks to the binary split capability. On the other hand, simulation results reveal that the 2-split condition does not add a significant additional wavelength consumption in usual network topologies with respect to SaD.

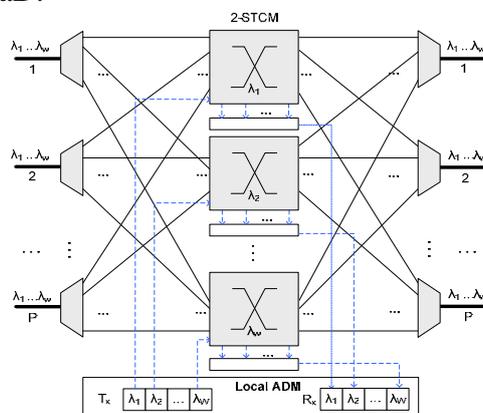


Fig. 1 General architecture of a $P \times P$ 2-STC MC-OXC node.

A comparison between different configurations is carried out as shown in Fig. 2(a). It can be seen that our proposal scales conveniently with respect to SaD and TaC modules. The number of components (optical switches, Ta2S switches, tunable 2x2 MMI splitters and tap devices) in 2-STCM is much lower than the other proposals, which contributes to improve the power efficiency.

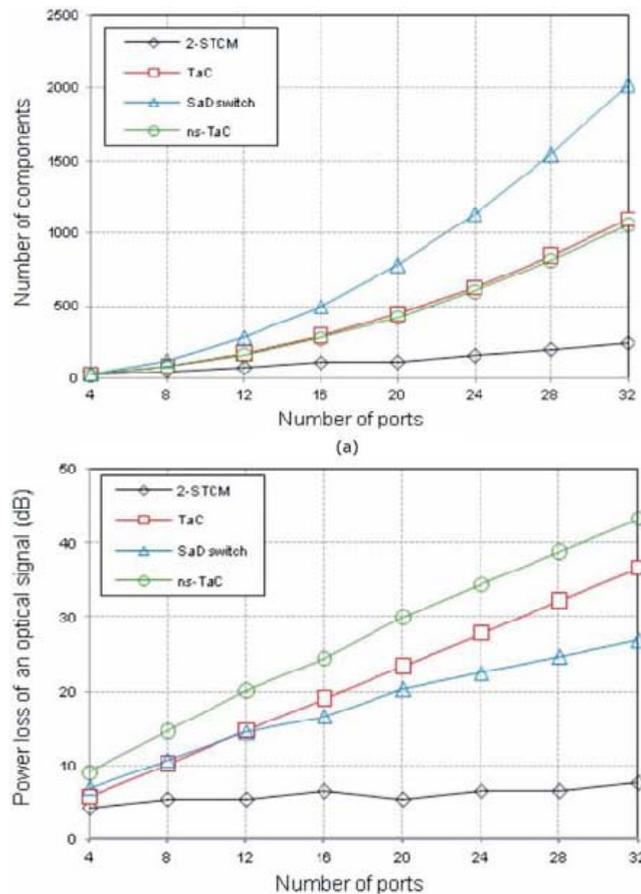


Fig. 2. Comparison of architectures. (a) Total number of components in 2-STCM, TaC and SaD modules. (b) Power loss (in dB) of a single optical signal for 2-STCM, TaC and SaD switch modules when a tap/drop-and-continue action is performed (average case).

Collaborative actions carried out

Meetings: Many teleconferes were held to coordinate joint work

Mobility actions:

Future Activities - Timescale

The JA has obtained several results. The possibilities of further joint collaboration UPVL are under study. This would take place during the second semester of 2010.

Outcome of the joint research activity:

[P1] Gonzalo M. Fernández, David Larrabeiti, Carmen Vázquez, Pedro Contreras Lallana “Power-Cost-Effective Node Architecture for Light-Tree Routing in WDM Networks”. New Orleans, IEEE Globecom 2008

[P2] Gonzalo M. Fernández, Carmen Vázquez, Pedro Contreras Lallana, and David Larrabeiti, “Tap-and-2-Split Switch Design Based on Integrated Optics for Light-Tree Routing in WDM Networks” JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 27, NO. 13, JULY 1, 2009



4.2 JA2 - Feasible parallel schedulers for OBS/OPS nodes

Responsible partner: UPCT

Participants: UPCT, UVIGO, UNIBO-DEIS, UNIROMA1

Description of work carried out.

The activities carried out in this JA can be classified into three main lines:

1. (UPCT, UNIBO): Design and performance evaluation of parallel scheduler for asynchronous arrivals, variable size packets in OPS/OBS switches. The PI-OBS scheduler.
2. (UVIGO, UPCT, UNIBO) Design and performance evaluation of parallel scheduler for IBWR switch, preserving packet sequence (O-IPDBM scheduler).
3. (UVIGO): Design and performance evaluation of different scheduling algorithms for asynchronous (not-aligned) Optical Cell Switching Networks.

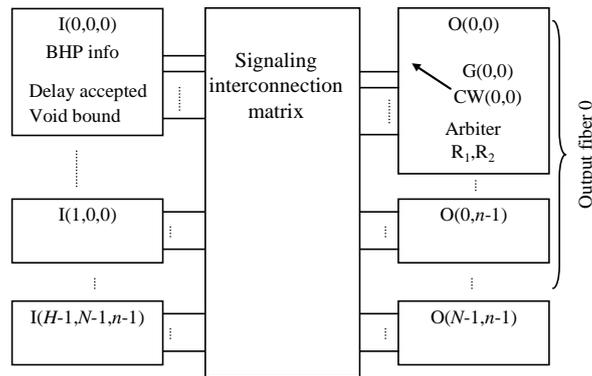
1. PI-OBS algorithm

UPCT and DEIS-UNIBO have investigated a proposal for parallel schedulers suitable for asynchronous arrival of packets of variable size, suitable for OBS or non-slotted OPS networks. As far as the partners know, this is the first attempt in this line. We have proposed the scheduler called PI-OBS (parallel-iterative scheduler for OBS), which has been presented in [Pav09].

The PI-OBS algorithm is designed as a parallel iterative algorithm, which is able to guarantee an upper bound to the response time. Let us denote this response time upper bound as T_A μ s. The algorithm is executed periodically, every T_I μ s. The constraint $T_I \geq T_A$ ensures that an algorithm execution starts strictly after the previous execution is finished. The algorithm execution starting at time $t=t_0$, is responsible for jointly processing the burst headers asynchronously received during the time interval $[t_0-T_I, t_0]$. We call this interval the *header arrival time window* of the algorithm execution.

We denote T_{WC} as the worst case time (μ s) spanning between the instant of header reception, and the moment in which a path is ready for the payload. T_{WC} is the sum of three time parameters: (i) T_I , as the worst case time from header arrival to algorithm execution (corresponding to a header received just at the start of a header arrival window), (ii) the algorithm response time T_A , and (iii) the reconfiguration time T_O of the optical components of the OSF. Two system parameters can be tuned to fulfill this constraint: (i) a minimum offset time δ_m between the burst header and the burst payload seen by S_E node, and (ii) an extra delay D_P added in the payload path, implemented by fixed duration FDLs in the data input ports.

Next figure depicts the main building blocks of the proposed scheduler architecture, for a switch fabric with N input and output fibers, and n wavelengths per fiber. It is based on the electronic interconnection of nNH input modules (left hand side), and nN output modules (right hand side), connected by means of a crossbar interconnection for inter-module signaling. H stands for the maximum number of horizons (horizons in this context are consecutive intervals of duration T_I , in which we organize the future payload arrivals, see [Pav09] for details)



PI-OBS scheduler architecture.

Performance results for PI-OBS show that its performance is very similar to that of the LAUC-VF scheduler, often considered an upper bound performance limit in OBS networks. This has been shown in [Pav09]. The current work to be carried out along 2010 is the implementation feasibility evaluation of the scheduler.

The PI-OBS is designed so that it can be implemented by the interconnection of fast combinational circuits, which may provide a total execution time independent on the switch size, and fast enough to be applicable for OBS networks.

2. Packet order in the IBWR switch. The OI-PDBM scheduler

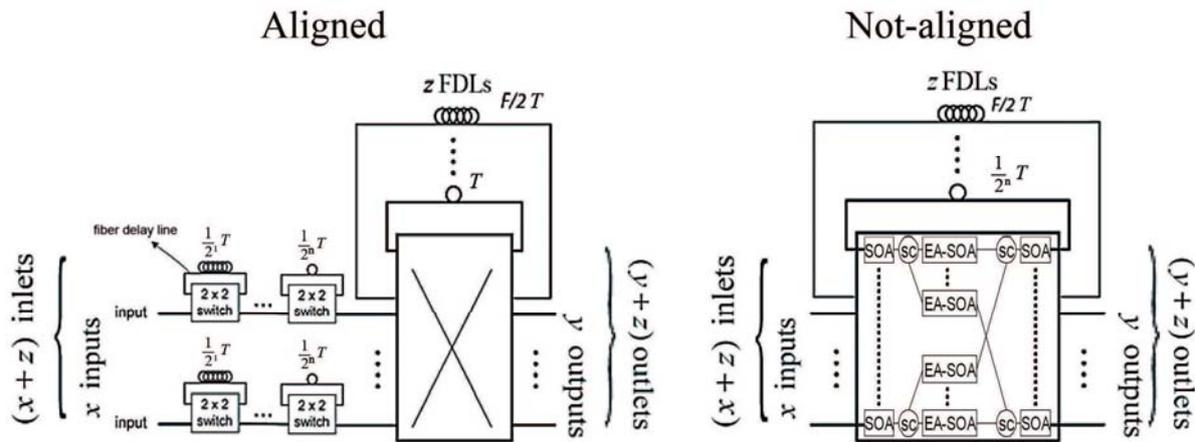
The IBWR switching architecture has been shown to have better hardware scalability properties than other optical switching fabrics for OPS networks. In the past, several algorithms have been proposed for scheduling this switch fabric under slotted traffic: fixed size packets, optically aligned at the switch inputs. PDBM [Pav07] and I-PDBM [Rod07] are two approaches in this line.

In this JA, UVIGO, UPCT and DEIS-UNIBO have investigated a variation of the PDBM-like algorithms to solve the traffic out-of-sequence problem that may arise in these architectures. The scheduler proposed is called OI-PDBM. It follows a modified request-grant scheduling, with respect to I-PDBM algorithm, as shown in [Rod09-A]. the advantages of OI-PDBM are (i) packet sequence is preserved, (ii) it maintains a simple and scalable implementation, (iii) it has shown good performance in terms of average packet delay, number of FDLs and algorithm convergence.

3. Design and performance evaluation of different scheduling algorithms for asynchronous (not-aligned) Optical Cell Switching Networks.

Most all-optical switching paradigms assume that different wavelengths are switched independently, which limits scalability. In optical cell switching (OCS), time is divided into time slots of fixed size by time-division multiplexing, and the wavelengths in a time slot are all bundled. Thus, each OCS switch has a single switching plane and performs mere time-space switching. Nevertheless, in an OCS network, each switching node (OCX) requires optical slot synchronizers (OSYNs) at all inputs for the arrival slots to be aligned, so that cells can be simultaneously forwarded.

To avoid this harmful process, UVIGO proposed a new switching paradigm, called Not-Aligned OCS [Rod09-C], where the OSYNs and the alignment process are no longer required. Cell shifting still takes place inside the OCXs for minimizing the gaps between cells, but it is not necessary to align them to a reference time. Not-aligned OCS has clear advantages over aligned OCS: the total number of fiber delay loops (FDLs) and the hardware cost are reduced, and the number of switching operations is also lower. Moreover, cell arrival time to the switch is not critical, and the network becomes simpler and more flexible. Next figure shows OCS and Not-Aligned OCS OCXs architecture.



As it can be in the previous figure, in Not-Aligned OCS FDLs are shared among the different input ports. Thus, to avoid possible contention problems, a scheduling algorithm is needed. First, we characterized the assignment of these shared FDLs as a min cost max flow problem to obtain a bound on OCS optimum performance (see [Rod09-B] for details). One of the conclusions, we took out of that paper, was that the performance of the SEFA [Lie05] scheduling algorithm was close to the optimum. So based on SEFA algorithm, we proposed several scheduling algorithms for Not-Aligned OCS networks that takes connection blocking probability to a reasonable values for practical loads.

References:

- [Pav07] Pavon-Marino, J. Garcia-Haro, A. Jajszczyk, "Parallel Desynchronized Block Matching: A Feasible Scheduling Algorithm for the Input-Buffered Wavelength-Routed Switch", *Computer Networks*, vol. 51, no. 15, pp. 4270-4283, October 2007.
- [Rod07] M. Rodelgo-Lacruz, P. Pavón-Mariño, F. J. González-Castaño, J. García Haro, C. López-Bravo and J. Veiga-Gontán, "Enhanced Parallel Iterative Schedulers for IBWR optical packet switches", *Lecture Notes in Computer Science (Springer-Verlag GmbH, ISSN 0302-9743)*, VOL. 4534, pp. 289-298. Proceedings of the 11th International Conference on Optical Networking Design and Modeling - ONDM 2007, Athens (Greece), May 2007.
- [Rod09-A] M. Rodelgo-Lacruz, P. Pavón-Mariño, F. J. González-Castaño, J. García-Haro, C. López-Bravo, J. Veiga-Gontán, and C. Raffaelli "Guaranteeing packet order in IBWR Optical



Packet Switches with Parallel Iterative Schedulers”, European Transactions on Telecommunications, DOI: 10-1002/ett1365..

[Rod09-B] M. Rodelgo Lacruz, C. López Bravo , F.J. González Castaño, F. J. Gil Castiñeira, H. J. Chao, ”Min-Cost Max-Flow Characterization of Shared-FDL Optical Switches,” IEEE Communications Letters, vol. 13, no. 7, pp. 540-542, August 2009.

[Rod09-C] M. Rodelgo Lacruz, C. López Bravo, F. J. González Castaño, F.J. Gil Castiñeira, H.J. Chao, “Not-Aligned Optical Cell Switching Paradigm”, IEEE/OSA Journal of Optical Communications and Networking, vol. 1, no. 3, pp. B70-B80, 2009.

[Rod09-D] M. Rodelgo Lacruz, C. López Bravo, F. J. González Castaño, F.J. Gil Castiñeira, H.J. Chao, “Distributed Resource Scheduling in Not-Aligned Optical Cell Switching ”, accepted for publication in IEEE/ACM Transactions on Communications.

[Lie05] S. Y. Liew, H. J. Chao and G. Hu, “Scheduling algorithms for shared fiber-delay-line optical packet switches. Part I: The single-stage case,” J. Lightw. Technol., vol. 23, no. 4, pp. 1586-1600, Apr. 2005

Collaborative actions carried out

Meetings (including tele-conferences)

Several conferences between UPCT and DEIS-UNIBO in the joint work around PI-OBS scheduler proposal.

Mobility actions: -

Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)

Follows a time plan for the 3rd year in this activity:

- UPCT and DEIS-UNIBO are working on a feasibility test for the electronic implementation of a proposed variation of the PI-OBS scheduler, which has shown to provide similar performance merits, allowing a simpler hardware implementation.
- UVIGO will continue working on Not-Aligned OCS networks, adapting Not-Aligned OCS network switching nodes to multicast traffic delivery.

Outcome of the joint research activity:

Joint publications:

- (UPCT, UNIBO) Pablo Pavon-Mariño, Juan Veiga-Gontan, Alejandro Ortuño-Manzanera, Walter Cerroni, Joan Garcia-Haro, " PI-OBS: a Parallel Iterative Optical Burst Scheduler for OBS networks", IEEE International Workshop on High Performance Switching and Routing 2009 (HPSR 2009), Paris, June 2009.
- (UVIGO, UPCT, UNIBO) M. Rodelgo-Lacruz, P. Pavón-Mariño, F. J. González-Castaño, J. García-Haro, C. López-Bravo, J. Veiga-Gontán, and C. Raffaelli “Guaranteeing packet order in IBWR Optical Packet Switches with Parallel Iterative Schedulers”, European Transactions on Telecommunications, DOI: 10-1002/ett1365.



Other:

- (UVIGO) M. Rodelgo Lacruz, C. López Bravo , F.J. González Castaño, F. J. Gil Castiñeira, H. J. Chao, "Min-Cost Max-Flow Characterization of Shared-FDL Optical Switches," IEEE Communications Letters, vol. 13, no. 7, pp. 540-542, August 2009.
- (UVIGO) M. Rodelgo Lacruz, C. López Bravo, F. J. González Castaño, F.J. Gil Castiñeira, H.J. Chao, "Not-Aligned Optical Cell Switching Paradigm", IEEE/OSA Journal of Optical Communications and Networking, vol. 1, no. 3, pp. B70-B80, 2009.
- (UVIGO) M. Rodelgo Lacruz, C. López Bravo, F. J. González Castaño, F.J. Gil Castiñeira, H.J. Chao, "Distributed Resource Scheduling in Not-Aligned Optical Cell Switching ", accepted for publication in IEEE/ACM Transactions on Communications.



4.3 JA4 – Performance of Optical Switching System Architectures with Shared Wavelength Converters

Responsible partner: Bilkent, UNIBO

Participants: Bilkent, UNIBO, UniRoma1, FT, FUB, KTH, UNIMORE

Description of work carried out.

The joint work between Bilkent and Unibo compares four different schemes to share wavelength converters in **asynchronous** optical packet switches with variable length packets. The first two architectures are the well-known Shared-Per-Node (SPN) and Shared-Per-Link (SPL) architectures, while the others are the Shared-Per-Input-Wavelength (SPIW) architecture (recently proposed as an optical switch architecture in synchronous context only, and a new scheme called Shared-Per-Output-Wavelength (SPOW) architecture. In this joint work, analytical models to evaluate packet loss in asynchronous switching systems are proposed for SPIW and SPOW architectures based on Markov chains and fixed-point iterations. The proposed models also account for unbalanced traffic whose impact is thoroughly studied. These models are validated by comparison with simulations which demonstrate that they are remarkably accurate. As an example, we consider an optical packet switch with 16 fibers, 32 wavelengths per fiber, asymmetric Poisson traffic for each output fiber and an overall traffic load of 60% on the switch. The packet loss probability obtained for the four conversion schemes using our proposed analytical model, namely SPN, SPL, SPIW, and SPOW, are depicted in **Figure 3**. We conclude that in terms of performance, the SPOW scheme provides blocking performance very close to the SPN scheme while maintaining almost the same complexity of the space switch, and employing less expensive wavelength converters. On the other hand, the SPIW scheme allows reduced implementation complexity while it substantially outperforms the widely accepted SPL scheme. The involved partners therefore believe that the SPIW and SPOW schemes are promising alternatives to the conventional SPN and SPL schemes for the implementation of next-generation optical packet switching systems.

The joint work between Unibo and Uniroma1 studies a multi-fiber all-optical **synchronous** switch which shares wavelength converters for contention resolution. The proposed switch architecture employs fixed-input/tunable-output wavelength converters (expected to be less complex than tunable-input/tunable-output ones), i.e., SPIW. The space switching matrix is modular and simple with respect to switching architectures with different wavelength converters sharing schemes (for example SPN). A parallel scheduling algorithm is proposed in this joint work to control optical packet forwarding in a synchronous scenario as well as an analytical model to evaluate packet loss performance. The analytical model is validated against simulation and the numerical results show good accuracy in most cases of interest for optical switch design. Another extension of this work is the management of quality of service in optical packet switches equipped with fixed input, tunable output wavelength converters. This kind of switch is studied in this joint work with the aim of keeping switch costs low by employing simple shared optical components and low-complexity space switching matrices. In particular, differentiation among classes of service in terms of loss performance for bufferless switches is addressed. The functionalities described refer to the high-speed

multiservice networking context foreseeable for the future Internet. A switch control algorithm to manage service classes in a slotted multiplexing context is proposed. An analytical model to evaluate the loss probability of each class is developed and validated against simulations and this model is used to prove the effectiveness of the quality-of-service approach. Numerical results show that the proposed scheme achieves basic class isolation as needed in core network optical packet switches.

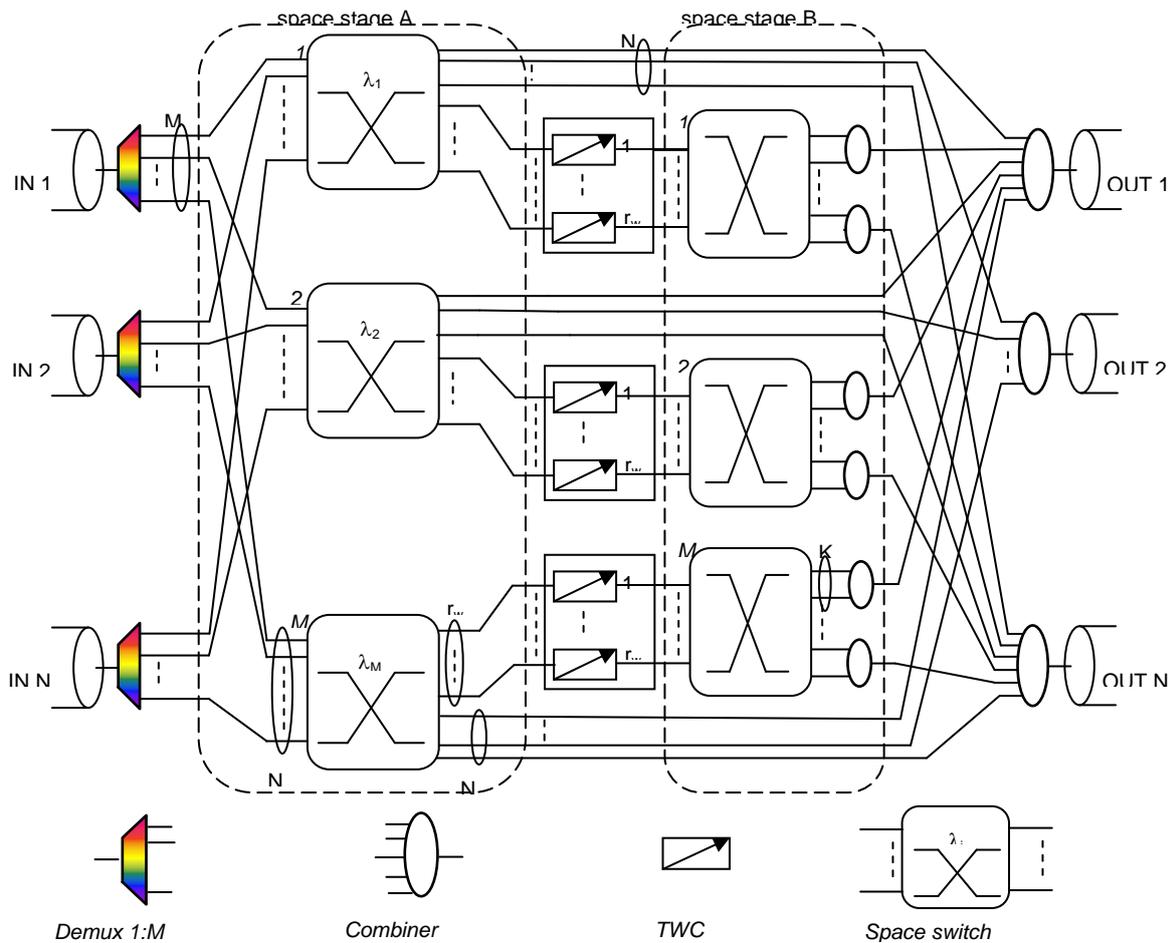


Figure 1 SPIW architecture

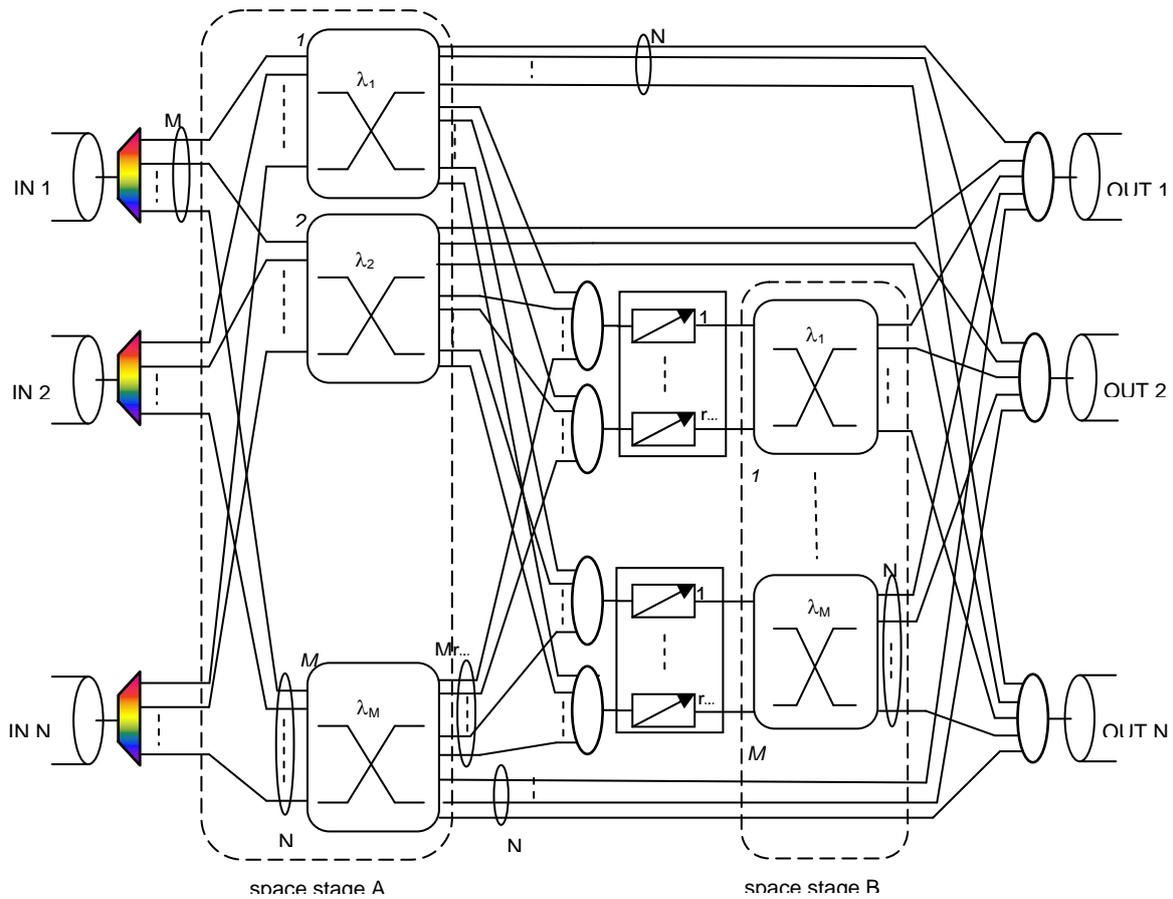


Figure 2 SPOW architecture

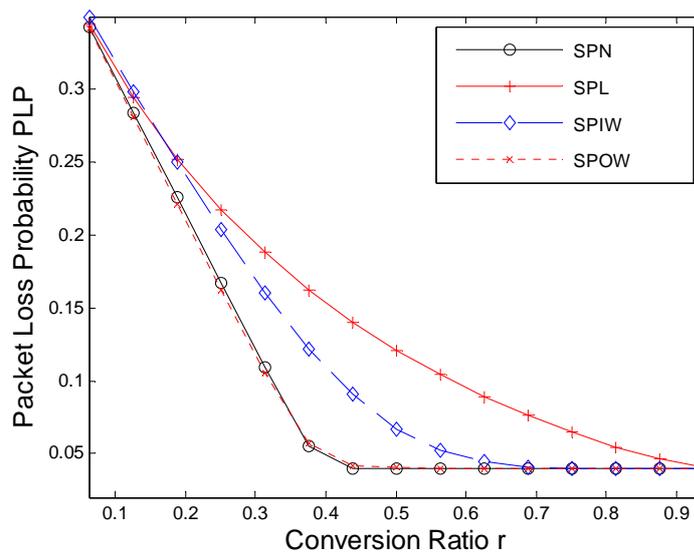


Figure 3 Packet loss probability for an optical switch with 16 fibers, 32 wavelengths per fiber, traffic load=60%, and asymmetric Poisson traffic using the four different wavelength conversion architectures.



Collaborative actions carried out

Meetings (including tele-conferences): Numerous e-mail exchanges and several face-to-face meetings

Mobility actions

Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)

Joint work between the parties will continue on optical switch architectures and their performance evaluation. Of particular interest to all involved partners are quality of service management in wavelength converting switches and their performance in addition to hybrid use of wavelength converters and fiber delay lines for contention resolution.

Outcome of the joint research activity:

Joint publications:

- V. Eramo, A. Germoni, C. Raffaelli, M. Savi, "Packet loss analysis of shared-per-wavelength multi-fiber all-optical switch with parallel scheduling", Elsevier Computer Networks, vol. 53, no. 2, pp. 202-216, February 2009.
- C. Raffaelli, M. Savi, A. Stavdas, "Multistage Shared-per-Wavelength Optical Packet Switch: Heuristic Scheduling Algorithm and Performance", IEEE Journal of Lightwave Technology, Vol. 27, No. 5, March 2009.
- N. Akar, E. Karasan, C. Raffaelli, "Fixed point analysis of limited range share per node wavelength conversion in asynchronous optical packet switching systems", Photonic Network Communications, , Volume 18, Number 2, pp. 255-263, October 2009.
- A. Cianfrani, V. Eramo, A. Germoni, C. Raffaelli, M. Savi, "Loss Analysis of Multiple Service Classes in Shared-Per-Wavelength Optical Packet Switches", IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 1, no. 2, pp. 69-80, July 2009.
- V. Eramo, A. Germoni, A. Cianfrani, C. Raffaelli, M. Savi, "Evaluation of QoS Differentiation Mechanism in Shared-Per-Wavelength Optical Packet Switches", ONDM 2009, Braunschweig, Germany.
- C. Raffaelli, M.Savi, 'Hybrid Contention Resolution in Optical Switching Fabric with QoS traffic', WOBS 2009, Madrid, Spain.
- N. Akar, C. Raffaelli, M. Savi, E. Karasan, "Shared-Per-Wavelength Asynchronous Optical Packet Switching: A Comparative Analysis", under revision for Computer Networks.

4.4 JA5 - Code-based optical nodes

Responsible partner: ROMA3

Participants: UNIROMA3, UNIMORE, UNIBO, UNIROMA1, GET

Description of work carried out.

The research activities related to the use of optical codes (OC) in Optical Packet Switching (OPS) and Optical Burst Switching (OBS) are proceeding following different but complementary directions. UNIROMA1 considers optical networks architectures where data are both Optical Code Division Multiplexed (OCDM) and Wavelength Division Multiplexed (WDM), whereas UNIBO and UNIMORE use optical codes only inside the node to solve packet contentions. UNIROMA3 and NICT have investigated network architectures where codes are used for packet labelling and WDM is used to increase the spectral efficiency of the payload. On the other hand, UNIBO uses codes to create paths in OBS systems and GET investigated a new label recognition and packet routing scheme, combining a SOA-MZI-based flip-flop and FBG-based OCDMA decoders. In the following, all the research activities performed are described in details.

The node switch investigate by UNIROMA1 is illustrated in Fig. 1 [1], and it is equipped with shared Wavelength Converters (WC) and uses both WDM and OCDM domains to solve the packet contentions. The packets are wavelength and code de-multiplexed, next are routed through the switching fabric where the contention resolutions are performed, finally they are encoded and wavelength multiplexed on the output fiber.

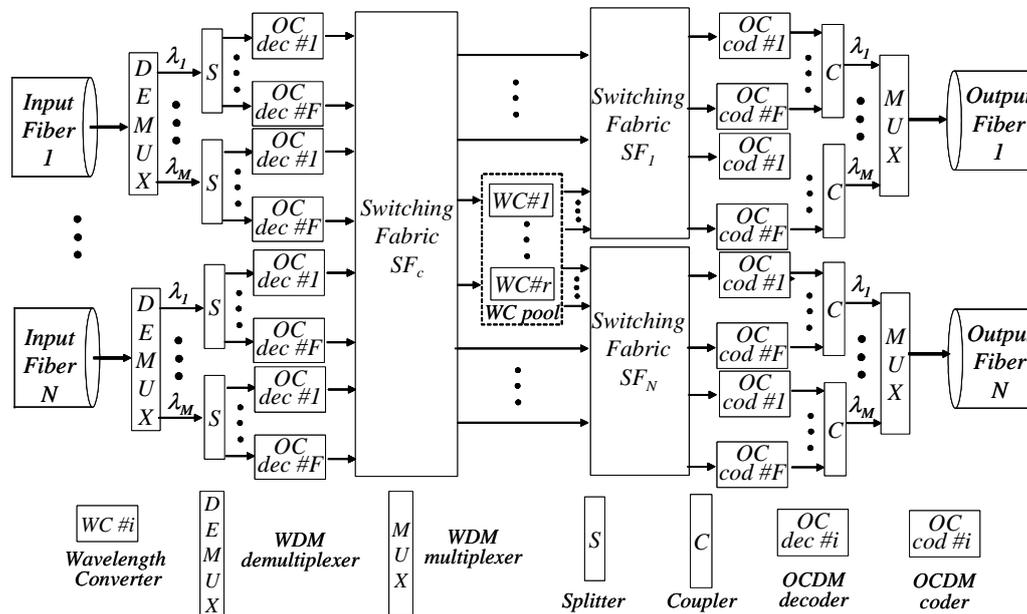


Fig. 1 WDM/OCDM Optical Packet Switch

Analytical models have been proposed to dimension the switch resources depending on the number F of optical codes carried on each wavelength, and the number r of the WCs. The packet loss probability due to output packet contentions is evaluated as a function of the main switch and traffic parameters, when coherent Gold codes are used.

Dimensioning aspects of the number r of WCs are coped with in Fig. 2, where the packet loss probability $P_{loss,c}$ due to output packet contentions is plotted as a function of r , for $N=8$ (number of input/output fibers), $M=2,8$ (number of wavelengths on a single fiber) and offered traffic $p=0,6, 0,8, 1$. The packet length is $H=500$ bytes, and to keep the bandwidth constant, the code length L is chosen so that the product $L \cdot M$ does not change. Therefore L is 31 or 127 for $M=8$ and $M=2$, respectively. The number F of OCs for each wavelength is selected so that the threshold Packet Loss Probability $P_{loss}^{MAI,th}$ due to MAI noise is below 10^{-10} . For instance, we obtain $p=0,6$ for $F=10, F=93$, and $L=31, L=127$, respectively. Observing Fig. 2, we can notice that all the curves have the same trend, decreasing with r with a threshold value r_{th} , for which the packet loss probability saturates; the saturation value corresponds to the packet loss probability of an OCDM/WDM switch using a full set of WCs, and therefore represents the packet loss probability due to the lack of output channels. The r_{th} value denotes the smallest number of WCs needed in an OCDM/WDM switch to reach the same Packet Loss Probability of an OCDM/WDM switch using a full set of WCs. The WC dimensioning is less severe as the number M of wavelengths decreases. For instance, when $p=0,6$, r_{th} is 20 and 28, for $M=2$ and $M=8$, respectively. This is to be expected, as the smaller M , the larger F and hence the probability that a contending packet can be forwarded on the same wavelength by changing the optical code and without the use of a WC.

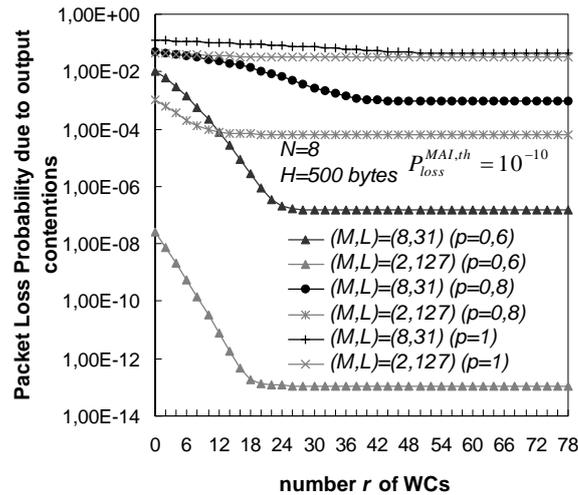


Fig. 2 Packet Loss Probability $P_{loss,c}$ due to packets output contentions as a function of the number r of used WCs for $p=0,6, 0,8, 1, N=8$, $H=500$ bytes and $(M,L)=\{(8,31), (2,127)\}$. The Packet Loss Probability due to MAI is 10^{-10} .

From an inspection of Fig. 2, we can also notice that when few WCs are employed, the lowest $P_{loss,c}$ is reached with fewer wavelengths, that is in the case $(M,F)=(2,127)$. For instance, in the case $p=0,8$, we have $P_{loss,c}=1,32 \cdot 10^{-4}$ when $(M,F)=(2,127)$. However, when a small number of WCs are used, the packet loss is smaller if fewer wavelength conversions are needed, but a larger number F of OCs is used. Even when many WCs are used, the lowest $P_{loss,c}$ is reached for small values of the number of wavelengths and of the code length. The effect is especially evident when low traffic is offered to the OCDM/WDM OPS. In this case, $P_{loss,c}$ has already reached the saturation value that depends on the total number $M \cdot F$ of available output channels. $P_{loss,c}$ reduces when $M \cdot F$ increases. For L increasing, instead, the statistical multiplexing gain leads to larger $M \cdot F$ and, consequently, smaller values of $P_{loss,c}$ are obtained, as shown in Fig. 2. As an example, for $p=0,8$, $P_{loss,c}$ equates $6,54 \cdot 10^{-5}$ for $(M,F)=(2,127)$.

As a complementary research field, UNIBO is investigating an optical switch, which uses optical codes to resolve contention [2, 3] only internally into the node, that has N input fibers (IFs) and N output fibers (OFs), each of them carrying a WDM signal with M wavelengths. Therefore, the node switch comprises a pool of WCs, and a set of encoding/decoding devices. Differently from the node architecture of Fig. 1, in this switching system the codes are removed before the packets exit the node; on the other hand, sharing schemes to reduce the number of WCs needed are considered, and simple WC solutions, such as fixed-input or fixed-output WCs. Therefore, the resulting switch architectures fall in the shared-per-wavelength context. In particular, UNIBO has focused its research on OPS architectures, which employ WCs shared-per-input (SPIW) or shared-per-output (SPOW) wavelength. Both architectures rely on optical devices that are expected to be mature in the near future or even already available (MUX/DEMUXes, couplers/splitters, optical encoder/decoder) on the market.

In a SPIW architecture, the WCs have fixed input wavelengths and full-range tunable outputs, as it is sketched in Fig. 3. Input wavelength channels are split and their packets are bit-by-bit encoded and forwarded to the star coupler. The code is chosen in relation to the output interface. Contention arises when two or more packets on the same wavelength are addressed to the same output interface. In this case one of them is encoded and forwarded, while the others are sent without encoding, to the proper WC to be shifted to another wavelength. The WC block is selected by controlling the optical gates. After wavelength conversion, the packets are encoded and broadcasted to the output interfaces, where they are properly selected, according the code, by decoders.

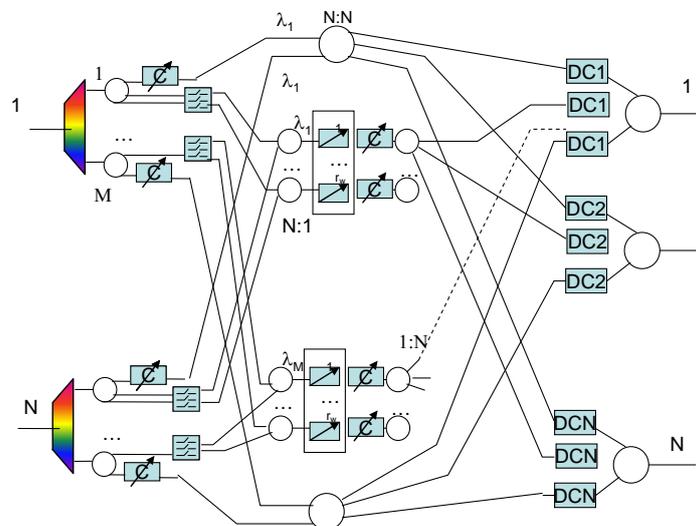


Fig. 3 Code-based SPIW optical switch with N input/output fiber interfaces each carrying M wavelengths each. The switch is equipped with N blocks of M WCs, each dedicated to the same input wavelength (SPIW).

The SPOW solution is defined to use fixed-output WCs with tunable-input. The blocks of optical gates are used to choose among the WC blocks to properly solve output contention. In this case, the encoder after conversion are required to operate on fixed wavelengths. The representation of the SPOW architecture is reported in Fig. 4, and the main difference from SPIW is represented by the gating pool and by the control scheduling.

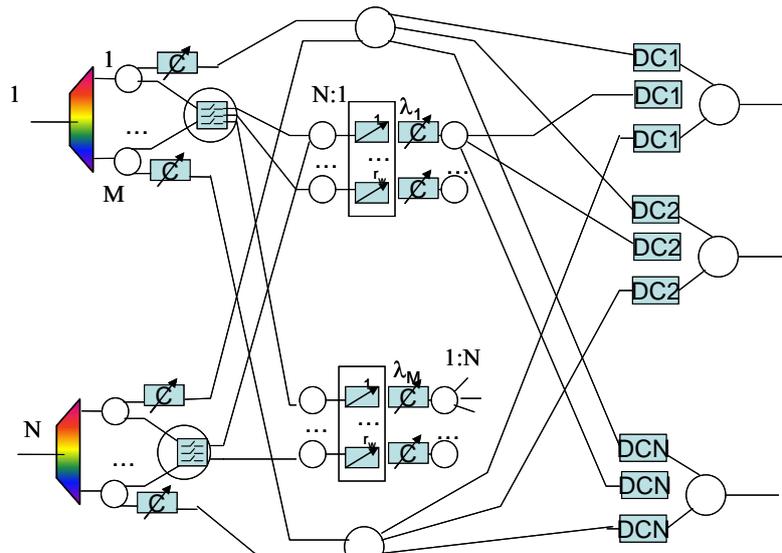


Fig. 4 Code-based SPOW optical switch with N input/output fiber interfaces each carrying M wavelengths each. The switch is equipped with rw blocks of M WCs each dedicated to the same output wavelength (SPOW).

In Fig. 5, a SPOW implementation, which employs a multi-port encoder is plotted. Here the multi-port encoders replace some star couplers, which provide packets on inputs to be encoded and appear on different output in relation to the destination interface.

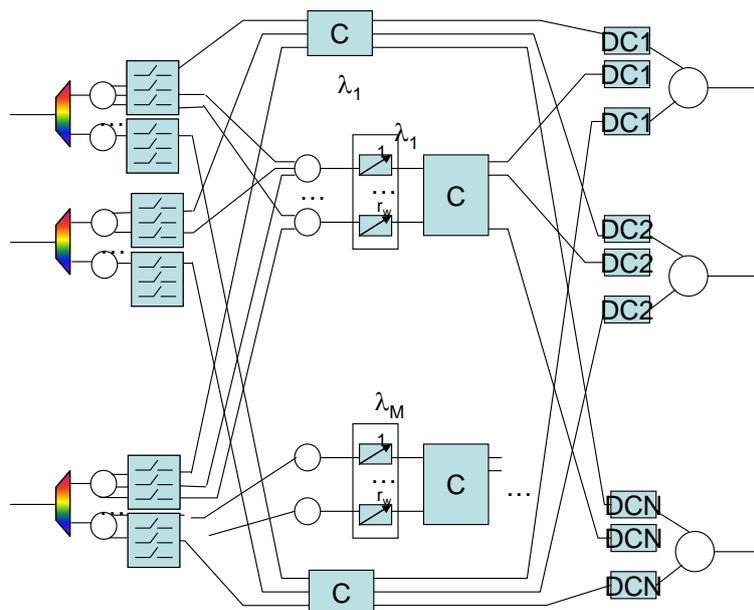


Fig. 5 Code-based SPOW optical switch with N input/output fiber interfaces each carrying M wavelengths each. The switch is equipped with rw blocks of M WCs each dedicated to an output wavelength (SPOW). A multi-port encoder is employed.

Performance of different schemes are analyzed in terms of packet loss in time slotted context with Bernoulli traffic.

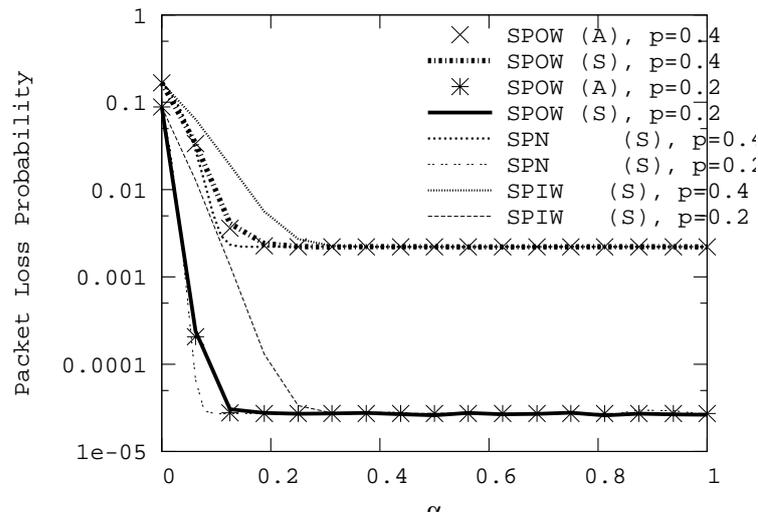


Fig. 6 Packet loss probability as a function of α =#wavelength converters/NM, with $N=16$, $M=8$

It can be seen, that the SPOW scheme performs much closer to the reference shared-per-node (SPN) architecture than SPIW, while employing simpler wavelength converters.

The multiport encoder/decoder is also one of the key devices of the node architectures investigated by UNIROMA3 and NICT, because its use allows us to accomplish three key requirements in OPS nodes: 1) the node architecture is modular and scalable; 2) the switching time is negligible with respect to the burst/packet duration; 3) the node provides large flexibility in bandwidth allocation, using interfaces between high-speed OPS/OBS and IP networks, that typically carries traffic at 10 Gbit/s or less. Using the multiport encoder/decoder, we can simultaneously generate and process N optical labels, and, using a multidimensional configuration, it is possible to increase the number of labels up to 2^N [4]. A large number of labels allow us not only to route all the packets, but we can use also the codes to accommodate different class of services, resolve contentions and/or implement multicast transmission. The switching time is extremely short, and it depends only on the time needed to process the optical label: using the multiport E/D, switching nodes at a speed of 13 Gpackets/second have been implemented. Finally, to solve the problem of the interface between an IP network and a high speed OPS network, a 160 Gb/s Ethernet (10 GbE) / 160 Gb/s OPS converter has been developed that solves also the contentions. Using coloured optical packets, where up to 16 wavelengths are transmitted simultaneously, it is possible to increase the spectral efficiency of OPS systems, and to bridge the gap between slow (10 Gb/s) IP networks and large bandwidth OPS systems.

UNIMORE is investigating the performance of a multi-stage Clos architecture having a buffer-less switch based on optical codes as basic switching element [5]. Optical codes are used for coding incoming bursts in each switching element in order to perform the switching function. To this end it employs encoders and decoders and possible output contentions are solved in the wavelength domain by means of variable-input fixed-output wavelength converters. A core node with this multi-stage architecture has been compared with a monolithic single-stage switch, with the final goal to evaluate the performance feasibility of a large Terabit optical switch. The main figure of merit investigated has been the burst blocking probability as a function of different load conditions and switch size. The main result obtained is that a 3-stage Clos architecture can be used for the realization of a terabit OBS core switch based on optical codes.

Fig. 7 shows the burst loss probability as a function of load for different multi stage switch size, with $\text{Prob}\{\text{Low Priority (LP) bursts}\}=0.8$ and $\text{Prob}\{\text{High Priority (HP) bursts}\}=0.2$. Performance for 32x32, 64x64 and 128x128 are also plotted. It is worth noting that performance improve as the size increase and 128x128 seems to represent the limit, which means that no meaningful improvements are expected for larger sizes. This is observed for LP bursts, since the performance of HP bursts are not here reported for 64x64 and 128x128 cases, because their very low burst loss values. Therefore, the 3-stage Clos architecture of a 128x128 switch has 128 input wavelengths operating at 10 Gbit/s, and it is a 1.2 Terabit/s non blocking switch for loads up to 0.6.

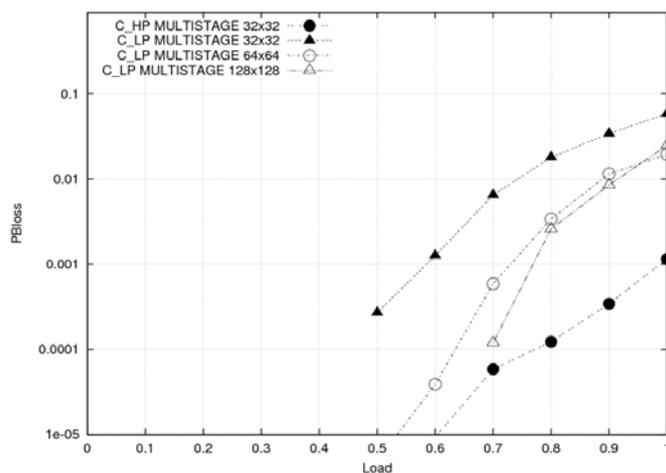


Fig. 7 Burst loss vs. load for different switch size

To achieve high-speed packet-switching, GET has proposed a new architecture taking advantage of the benefits of OCDM (fast label recognition, asynchronous, security, multiple access...) and all-optical processing offered by SOA (flip-flop). The design of router node combines OCDMA label switching with the concept of optical code gate. Depending on the matching address (correlator output pulse), a flip-flop will block or tolerate the propagation of the packet. The Ph. D. student Housseem Brahmi is currently in Athens (NTUA) for a preliminary experimental validation.

Collaborative actions carried out

Meetings (including tele-conferences)

Tele-conferences through Skype with UNIROMA3 and UNIBO

UNIROMA3 and NICT have met at ECOC and OFC conferences for technical discussions.

Mobility actions

Dimitrios Apostolopoulos visited GET.

Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)



UNIROMA1 will extend the analytical model of an optical OCDM/WDM network of a to take the beat noise into account.

UNIBO will make evaluations of the complexity and power consumption of the SPIW and SPOW architectures.

UNIMORE is planning to investigate other multistage architectures to improve the Clos architecture performance, evaluating of the overall feasibility.

UNIROMA3 will consider the possibility to use the coloured packets technique not to increase the spectral efficiency, but to resolve contention and accommodate different class of services.

Outcome of the joint research activity:

Joint publications:

- [1] V. Eramo, M. Listanti, A. Germoni, A. Cianfrani, "Impact of the MAI noise on the performance of OCDM/WDM optical packet switches using gold codes", submitted to Optics Express (2009).
- [2] G. Cincotti, N. Kataoka, N. Wada, X. Wang and K.-i. Kitayama, "Versatile usage of a multiport encoder/decoder in OPS and OBS nodes," invited paper Photonics in Switching Conference Pisa 2009.
- [3] 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, Issue 5, pages 538-551, March 2009.
- [4] V. Eramo, A. Germoni, C. Raffaelli, M. Savi, "Multi-Fiber Shared-Per-Wavelength All-Optical Switching: Architecture, Control and Performance", IEEE/OSA Journal of Lightwave Technology, Vol. 26, Issue 5, pages 537-551.

Other:

- [5] M. Casoni, A. Sacchi, "System Design and Evaluation of a Large Photonic Switch based on Optical Codes for Optical Burst Switched Networks", to be presented at next IEEE ANTS 2009, 14-16 December, New Delhi (India).

4.5 JA6 – Repacking and rearranging algorithms for multi-plane banyan type switching fabrics.

Responsible partner: Wojciech Kabacinski, PUT

Participants: POLIMI, PUT

Description of work carried out.

The multi-plane banyan type switching fabrics are composed of 2x2 switching elements grouped into planes. All these planes have the same structure. Examples of different structures of one plane are shown in Fig 1. Each plane has only one possible path to realize each possible connection. So, the control algorithm for path selection in such switching fabric must only chose a plane through which a new connection will be realized. Planes for particular connections have to be chosen very carefully, because in some situations wrong decision can cause blocking of planes for next connections. To avoid such situation or to resolve it when it occurs, special algorithms have to be used.

For different situations different algorithms will be used. We have proposed, investigated algorithms for setting up, rearranging and repacking connections in switching fabrics with even and odd number of stages. Some of them are being implemented in FPGA circuits.

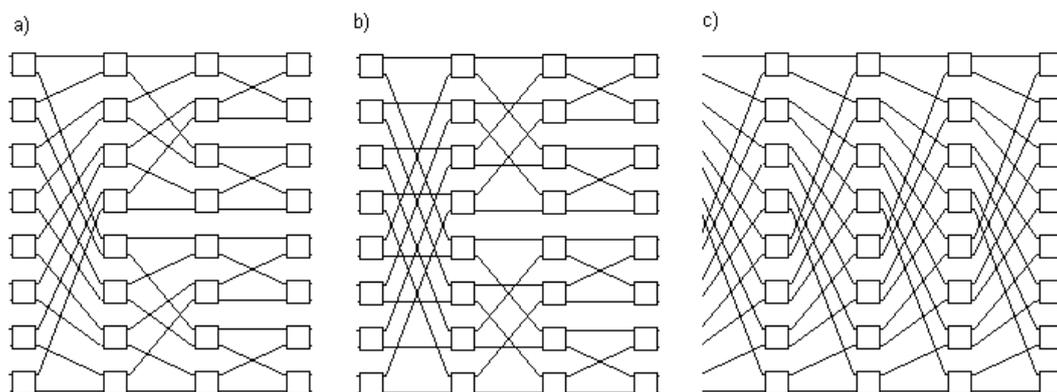


Figure 1. Different configurations of one plane: a) baseline, b) banyan, c) shuffle.

The Rearranging Algorithm for the Switching Fabric with Even Number of Stages which Requires Only One Rearrangement

All switching elements of the switching fabric are grouped into planes. Elements in each plane are organized as a two-dimensional matrix with certain number of columns and rows. The number of rows are 2 times lower than N (the total number of inputs or outputs) in the switching fabric, and the number of columns is equal to $\log_2 N$. Each such column is called a “stage”.

We proposed the algorithm, which can be used for rearranging connections in the rearrangeable switching fabric with the even number of stages. This algorithm will be used when there is no plane for a new connection (a blocking state appeared). An existing switching fabric’s state will be changed by re-switching existing connections to other planes



and at least one plane will be available for a new connection (i.e. this new connection will be unblocked). We have proved, that using this algorithm it is possible to unblock a blocked connection with only one rearrangement, and with the special version of this algorithm, such rearrangement can be done without interrupting other existing connections. During realization of this algorithm existing connections are grouped into three types according to their relation to the blocked connection. In Fig. 2 the general rearrangement algorithm is shown and its non-interruptive version is shown in Fig 3.

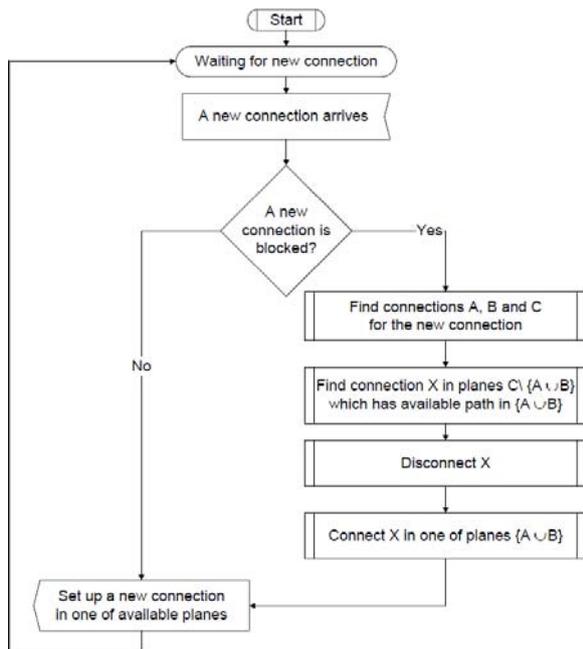


Figure 2. The algorithm for rearrangements in the switching fabric with even number of stages.

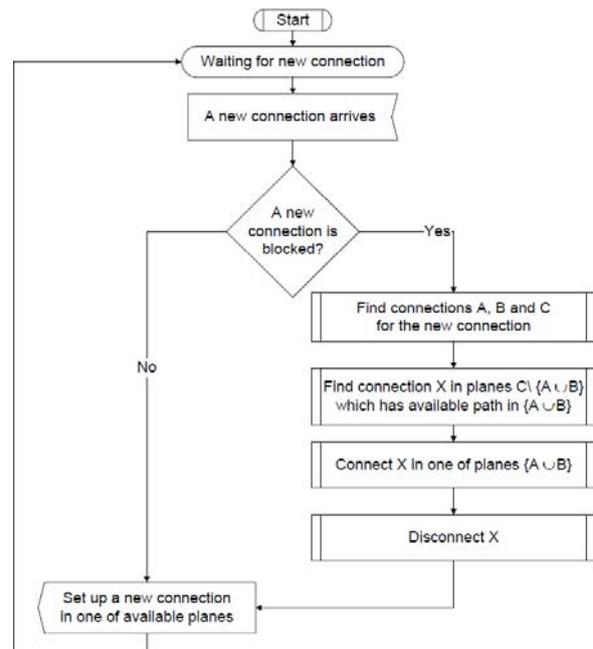


Figure 3. The algorithm for non-interruptive rearrangements in the fabric with the even number of stages.

The Rearranging Algorithm for the Switching Fabric with an Odd Number of Stages which Requires only Two Rearrangements.

Combinatorial properties of switching fabrics with the odd number of stages are different than properties of switching fabrics with the even number of stages. In switching fabrics with odd number of stages it is possible to obtain a state in which it is impossible to unblock some new connections only by one rearrangement. We proved, that in this state at least two rearrangements are required. We are still working on the formal proof, that necessary condition is also the sufficient one. All steps for this algorithm are shown in Fig 4. By changing order of some steps, we obtained another algorithm for rearrangements, but with no connection interruption during any rearrangement and a new connection will be also unblocked (as in the first version). Such non-interruptive implementation is very interesting, specially for switching fabrics realizing connections with a very high throughput (because even small break can cause a lost of a big amount of data from interrupted connection).

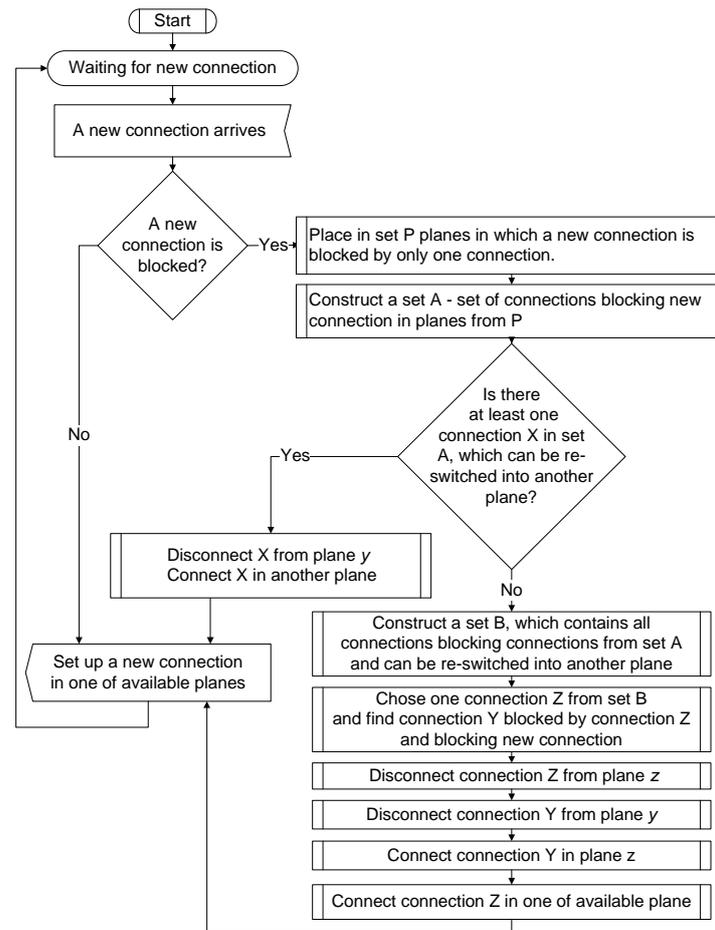


Figure 4. Algorithm for non-interruptive rearrangements in the switching fabric with odd number of stages.

The Implementation of Control Algorithms in FPGA Structures

All of the proposed algorithms are being implemented in hardware. We use FPGA (Field Programmable Gates Array) structures for this implementation. The proposed algorithms conversion of our into VHDL. To test and analyze the behavior of the programmed algorithms we use a special traffic generator (also prepared in the FPGA circuit). The prepared implementations allow to obtain results of algorithms in a very short time, even immediately. The path searching algorithm for instance can chose a plane for a new connection almost immediately – there is no delay between a request and an answer, and after one clock’s cycle the device is ready for setting up a next connection. This situation is shown in Fig 5. It can be seen, that at a rising edge of a clock signal, the plane number (signal with name “plane”) dedicated for a new connection (from input with number $input_i$ to output with number $output_j$) is known, and in the next clock cycle the next full and proper iteration can be done (four signals for each number represent its binary representation). We are working on an implementation of next algorithms (for rearrangements, repacking), optimization of them and making them more flexible.

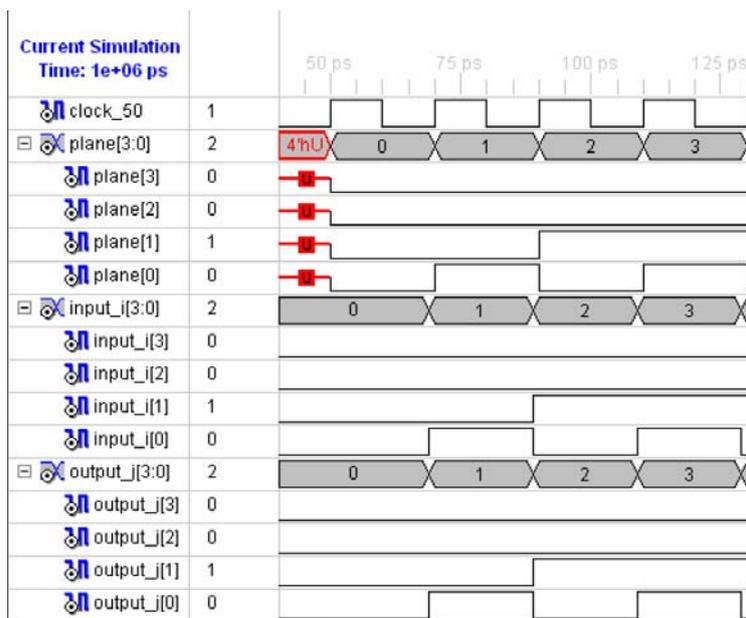


Figure 5. Simulation of a behavior of FPGA devices with implemented control algorithm (in 50 ps algorithm sends out information, that connection from input_i = 0 to output_j = 0 is to be set up in plane =0, in 70 ps – connection 1-1 is to be set up in plane 1, and so on...)

Collaborative actions carried out

Meetings (including tele-conferences)

Meetings with partners were held during the mobility action (see below), through internet calls (two in June and one in September), and also during occasional meetings at conferences (HPSR 2009, ICC 2009, ONDM 2009).

Mobility actions

Guido Meier – 4-9.05.2009, from POLIMI to PUT.

This mobility action was connected with WP25 and WP14. During this action intensive discussions were held on rearranging algorithms for multi-plane baseline type switching fabrics with odd number of stages. Several proposals for the proof have been considered. The next works have been carried out after the action separately.

Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)

Future activities will concentrate on proving the correctness of the algorithm for switching fabrics with odd number of stages and implementation of algorithms in FPGA.

Outcome of the joint research activity:

Joint publications:

W. Kabaciński, J. Kleban, M. Michalski, M. Żal, A. Pattavina, G. Maier: *Rearranging Algorithms for $\text{Log}_2(N, 0, p)$ Switching Networks with Even Number of Stages.* International Workshop on High Performance Switching and Routing, Paris, France, June 22 - 24, 2009,

4.6 JA7 – Power Consumption and Supply of High-Performance Switching Elements.

Responsible partner: Slavisa Aleksic, TUW

Participants: TUW, UPCT, BME, PoliTo, UNIMORE, UoP, USWAN, FUB.

Description of work carried out.

The work carried within the second year is mainly related to evaluation of the total electrical power consumption of large electronic and optical switching fabrics. In particular, four realization options of large switching fabrics including circuit and packet switches realized using either electronic (CMOS) or optical (SOA, MEMS) technologies have been considered.

Generally, optical transmission and processing technologies are able to provide high data rates per switching port, and thus, they may be used to increase capacity of switches and routers while keeping the switch port count at a reasonable level. However, complex data processing and buffering required in packet-switched network elements are very difficult to implement directly in the optical domain. Additionally, monolithic integration of complex photonic systems is currently at an early stage. Therefore, optical switching systems are mostly implemented using relatively simple structures with either a very limited or even not at all packet buffering. Thus, optical switches can not easily reach the features of current high-performance electronic switches, but they can potentially provide high switching speed and low power consumption.

Switching energy of a high-speed optical gate or a single switching cell is often reported to be about 100 fJ per one bit of operation, which is extremely low and at least an order of magnitude lower than that of current electronic packet switches. However, the total electrical power consumed by an optical gate is usually much higher. The process of generation of optical control pulses consumes electrical energy, and in some cases, a certain amount of energy has to be supplied to the gate itself because it may comprise active devices (see **Figure 4**). Even if the gate is completely passive, the optical data signal is attenuated and degraded in the gate due to insertion losses and various other impairments such as dispersion, crosstalk, and noise accumulation. Therefore, the signal has to be amplified and/or regenerated after a certain number of switching stages, which leads to increased total electrical power.

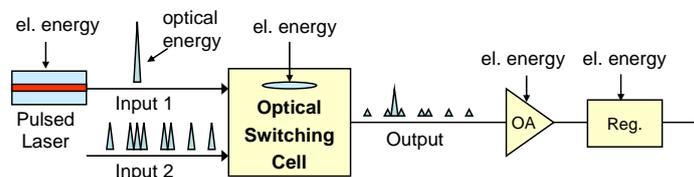
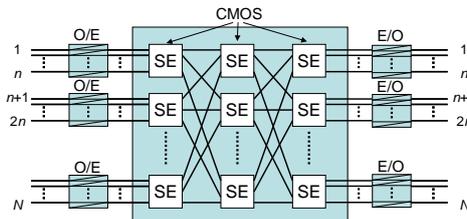


Figure 4: Contributors to the total electrical power consumption of an all-optical switching gate (OA: Optical Amplifier, Reg.: Regenerator).

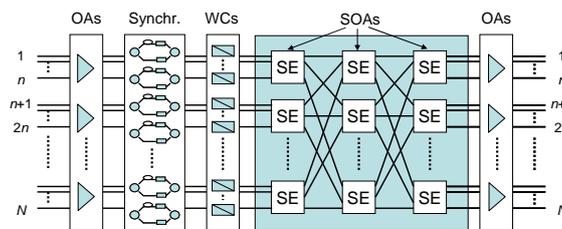
In this study, we considered four different options for the realization of large switching fabrics and estimate their total electrical energy consumption by taking into account all contributing elements as depicted in Figure 4. The first option shown in Figure 5a is a large electronic switching fabric that comprises a number of smaller switching elements (SEs) arranged in the non-blocking Clos network. Here, optical-electrical (O/E) and electrical-optical (E/O) conversion is applied at input and output ports, respectively. For the realization

of electronic packet switches (EPS) we assume SEs based on buffered packet switch devices, while electronic circuit switches (ECS) use cross-point switching elements. The switching elements of the considered optical packet switch (OPS) as shown in see Figure 5b are based on 1×2 SOA-based switching units arranged in the Spanke architecture. Only active semiconductor optical amplifiers (SOAs) are assumed to consume electrical power. All-optical 3R regenerators are placed after a cascade of 9 SOAs in order to compensate for signal degradation. Additionally, power consumption of all-optical wavelength converters, optical packet synchronizers, and optical amplifiers (EDFAs) are also taken into account. Finally, optical circuit switches (OCSs) are based on 3D-MEMS switching elements and make use of all-optical wavelength converters to reduce blocking probability (see Figure 5c).

a) EPS/ECS Switch



b) OPS Switch



c) OCS Switch

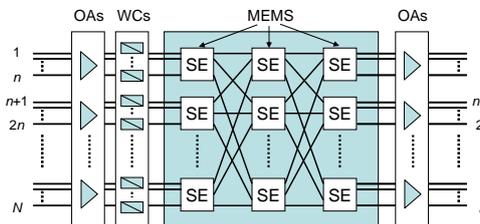


Figure 5: Generic architectures of the considered large switching fabrics

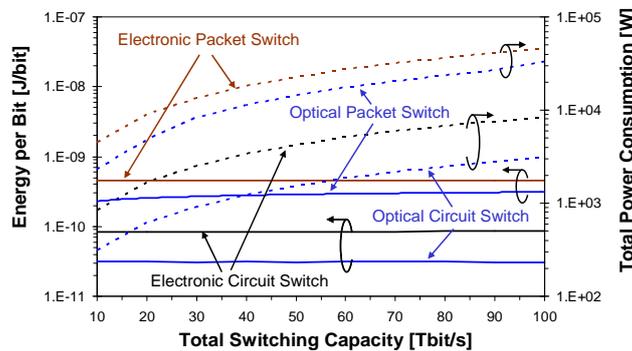


Figure 6: Comparison of optical and electronic realizations of packet and circuit switches with regard to total electrical power consumption and switching energy.

A detailed description of the model used for the estimation of power consumption can be found in [2]. As input to the model we used a large data base containing power consumption



values of various real components and subsystems obtained either from data sheets or from technical and research papers. **Figure 6** shows the energy consumption per bit and the total power consumption of the four considered implementation options. It is evident that the MEMS-based optical circuit switch is the most energy efficient one with switching energy in the order of several tens of pJ. The switching energy of a large electronic circuit switch lies below 100 pJ, while both packet switching options have considerably higher energy consumption of several hundreds of pJ.

References:

- [1] S. Aleksic, "Electrical Power Consumption of Large Electronic and Optical Switching Fabrics", accepted for publication in IEEE Photonics Society Winter Topicals 2010, Majorca, Spain, January 2010.
- [2] S. Aleksic, "Analysis of Power Consumption in Future High-Capacity Network Nodes" IEEE/OSA JOCN, vol.1, no. 3, pp. 245 – 258, 2009.

Collaborative actions carried out

Meetings were held during the ICTON2009 conference, Ponta Delgada, Portugal, July 2009 and at the BONE plenary meeting in Poznan, Poland, October 2009.

Mobility actions

There was no mobility actions performed during the 2nd year.

Future Activities – Timescale

Within the 3rd year, we intend to intensify collaborative work on evaluation of power consumption and other relevant performance parameters of large-scale switching fabrics. We will consider different architectures, implementations, and switching principles such as circuit-, packet, burst, hybrid- and flow-switching. We will work on the completion of the overview document comprising a description of all activities and results on power consumption of individual network elements obtained within both WP14 and WP21.

Outcome of the joint research activity:

Joint publications:

Other:

S. Aleksic, "Electrical Power Consumption of Large Electronic and Optical Switching Fabrics", accepted for publication in IEEE Photonics Society Winter Topicals 2010, Majorca, Spain, January 2010.

Remark:

This JA is under the common umbrella of the workpackages WP14 (Virtual Centre of Excellence on Optical Switching Systems) and WP21 (Topical Project on Green Optical Networks), providing the bridge between both domains. This positioning is crucial to allow for input from optical switching systems experts into the WP21 activities on power consumption.



4.7 JA9 - All-optical label processing techniques for ultra-fast optical packet switches.

Responsible partner: SSSUP

Participants: SSSUP, UNIROMA3, NICT.

Description of work carried out.

The joint activity had the goal of investigating all-optical processing techniques for the processing of the labels in ultra-fast optical packet switches. In the first part of the activity, the scenario of an optical packet switching (OPS) node with time-division multiplexed (TDM) label was considered. A combinatorial network for the resolution of the contentions was designed and implemented in a real 2×2 160 Gb/s/port OPS node test-bed at NICT in Koganei, Tokyo, Japan.

In the second part of the activity, a field trial experiment of DWDM-based optical packet switching has been successfully achieved, where packets with payload of 100 GHz (10 wavelengths \times 10 Gbit/s) have been transmitted over 64 km. The key devices of this new technology are newly-developed burst-mode EDFAs and an optical packet switch prototype with multiple all-optical label processors.

Figure 1(a) shows the concept of a proposed OPS network. The OPS network is connected to IP-technology-based metro or access networks. Namely, at an ingress edge node of the OPS network, the edge node system encapsulates an IP packet over Ethernet or SONET/SDH into an optical packet. Conversely, at an egress edge node, the IP packet is decapsulated from the optical packet.

In OPS networks, we introduce large-capacity colored optical packets using DWDM technology. As shown in Fig. 1(b), a DWDM-based optical packet consists of multiple 10 Gbit/s optical payloads of different N-wavelengths and an optical label. The 10 Gbit/s preamble is added to each payload. Thus, the data rate of multiple optical payloads in a DWDM-based optical packet is $N \times 10$ Gbit/s. Since 10 Gbit/s transmission systems are mature and are easy to interface with electronic systems, we can develop an OPS system more easily compared to a system involving transmission of an OTDM-based high-speed optical packet consisting of a single, large-bandwidth wavelength.

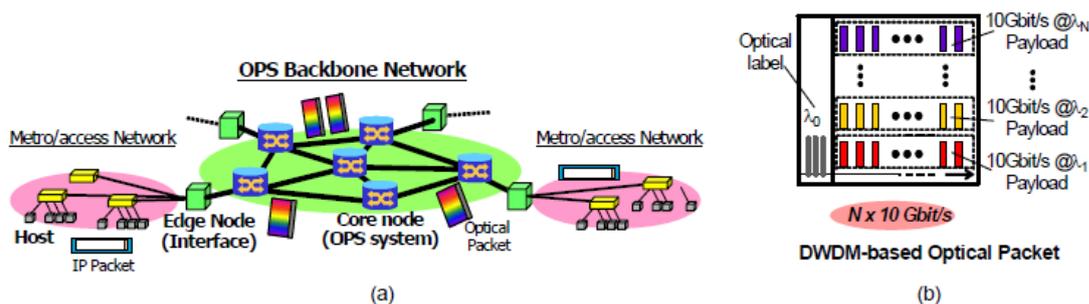


Fig. 1. (a) OPS backbone network. (b) Configuration of DWDM-based optical packet.

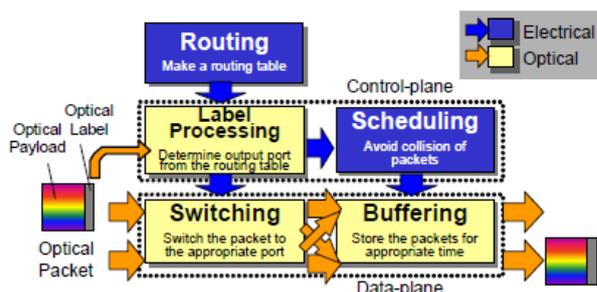


Fig. 2. Architecture of proposed OPS system.

Collaborative actions carried out

Meetings

UNIROMA3 and NICT have met at ECOC and OFC conferences for technical discussions.

Mobility actions: None

Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)

The JA achieved its goals by the end of the second year, so it can be closed and further activities will be addressed in a new JA.

Outcome of the joint research activity:

Joint publications:

Scaffardi M.; Lazzeri E.; Furukawa H.; Wada N.; Miyazaki T.; Poti L.; Bogoni, A.; 160 Gb/s/port 2 x 2 OPS Node Test-Bed Performing 50 Gchip/s All-Optical Active Label Processing with Contention Detection. to be published on IEEE J. of Lightwave Technol.

Other:



4.8 JA11 – Multi-granular designs optical Switch Architectures using QD-SOAs for optical buffering and.

Responsible partner: Wojciech Kabacinski, PUT; Kyriakos Vlachos, RACTI

Participants:, PUT, RACTI

Description of work carried out.

Optical Buffer design.

The project aims at the novel switch architecture with optical buffers. Nowadays, telecommunication networks consist of optical fiber links and electronic nodes. The common used technologies such as DWDM enable to transport data with tremendous bitrates, therefore the switching capacity is becoming the bottleneck of the telecommunication networks. Currently, our research group focuses on all-optical switching technologies, where optical/electrical conversions are limited to minimum. To accomplish that, an important issues that we have concerned are the proper switch architecture and buffering issues in optical domain. Currently, buffering is mainly implemented as electronic devices (RAM). To overcome expensive and slow E/O and O/E conversions, we have to store packets in fully optical domain. Such a buffering is typically done by fixed delay lines (FDLs). During this JA we are investigating some switch architectures with FDLs.

We proposed a new architecture for implementing optical buffers, with the use multi-wavelength selective elements like quantum dot semiconductor optical amplifiers (QD-SOAs) as multi-wavelength converters and fixed-length delay lines that are combined to form both an output queuing and a parallel buffer switch design. The proposed design consists of three stages, namely the input, the buffering and the output stage. The input stage only converts the incoming wavelength (λ_0) to a predefined set of wavelengths (i.e. $\lambda_1, \lambda_2, \lambda_3, \lambda_4$) for operating the multi-wavelength buffer stage. Similarly, the output stage converts back the data to their original wavelength. The buffering stage consists of different delay stages interconnected in line with fiber delay lines. This structure is a single time-slot-interchanger but in terms of functionality it corresponds to k parallel Time-Slot-Interchangers (TSIs), one per input port of the buffer. This is because the buffer fundamental switching blocks possess the capability to process simultaneously k wavelengths. The buffer architecture is a *shared* buffer architecture in the sense that buffer capacity is shared by all input ports.

The equivalent design is shown in Figure 7 (a) and consists of serially interconnected programmable delay stages in which the delay bank is accessible to all input ports, as detailed in Figure 7 (b). At each delay stage, k parallel wavelength converters assign the incoming packets/bursts to wavelengths that correspond to a pair of delay lines and output ports. The delay lines and output ports are accessed by the packets/bursts through all-passive space switches.

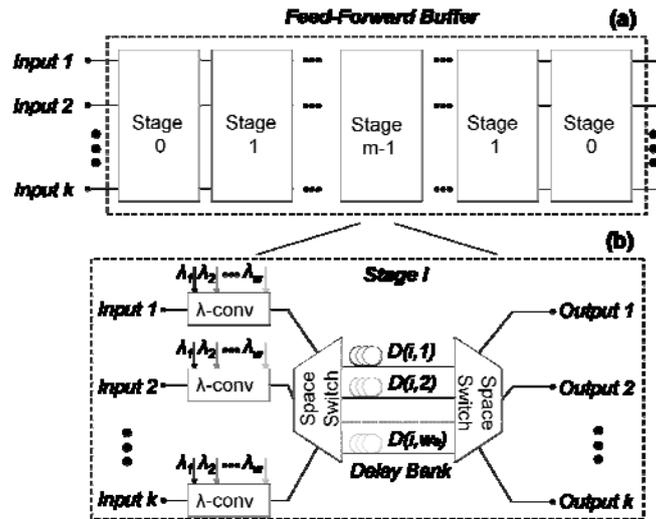


Figure 7. (a) The shared buffer architecture. (b) At the respective stages, each wavelength is assigned to a pair of delays and output ports. The wavelength converter tunability is w .

Operation of the QD SOA -buffer

In Figure 8a, a QD-SOA is used with two Arrayed Waveguide Grating (AWG) filters connected at its input and output. The operation of the switch is as follows. A set of equally-spaced data wavelengths λ_i carrying the incoming data on a predefined spectrum grid with wavelength spacing $\Delta\lambda_i$ ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) are wavelength-converted by the QD-SOA with the help of CW light at wavelengths λ_i^j . The data at λ_i modulate the carrier density with subsequent modulation of the refractive index, the phase and mainly the gain of the QD-SOA. This gain modulation effect is restricted to within a fairly narrow, homogeneously broadened spectral region around, say λ_1 . Figure 9a shows the relative position of, e. g., two channels within the homogeneously broadened part of the inhomogeneously broadened gain spectrum of the QD-SOA. Depending which CW laser oscillating at λ_i^j is switched on, the data will be transferred by cross-gain modulation (XGM) to the wavelength λ_i^j ; for a generalization see below.

outputs at $\lambda_1^1, \lambda_1^2, \lambda_1^3, \lambda_1^4$ for λ_1 input data signal, outputs at $\lambda_2^1, \lambda_2^2, \lambda_2^3, \lambda_2^4$ for λ_2 input data signal,

outputs at $\lambda_3^1, \lambda_3^2, \lambda_3^3, \lambda_3^4$ for λ_3 input data signal, outputs at $\lambda_4^1, \lambda_4^2, \lambda_4^3, \lambda_4^4$ for λ_4 input data signal.

The CW wavelengths λ_i^j are arranged on a grid with a frequency separation of about, say, $\Delta f = 200\text{GHz}$ such that $\lambda_1^1, \lambda_1^2, \lambda_1^3, \lambda_1^4$ are equidistantly spaced. The output AWG guides these five wavelengths to its five output ports, one of which (λ_1) is not used so that the input data wavelength (λ_1) is blocked. Therefore, depending on the choice of the CW source λ_i^j , the converted input data are directed to the four remaining output ports $j = 1..4$ of the output AWG. This AWG is designed to have a free spectral range (FSR) of, say, $\Delta\lambda_i = 20\text{nm}$. Therefore, all output wavelengths $\lambda_i^1, \lambda_i^2, \lambda_i^3$, and λ_i^4 belonging to different input data wavelengths λ_i are directed to the same physical output ports $j = 1, 2, 3$, and 4 of the output AWG. The AWG routing matrix for the example of Figure 8a is shown in Figure 9b.

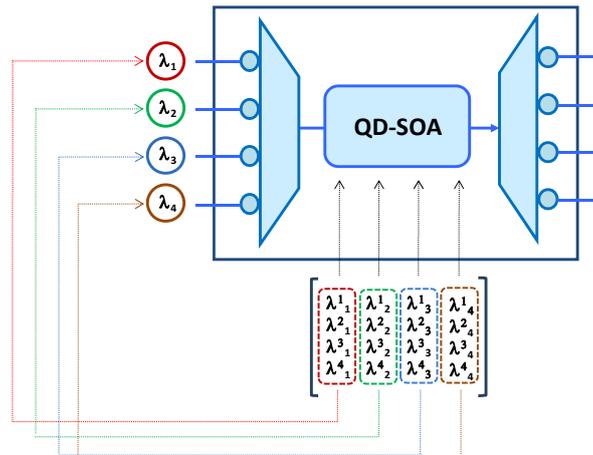


Figure 8. Buffer fundamental switching block using a QD-SOAs

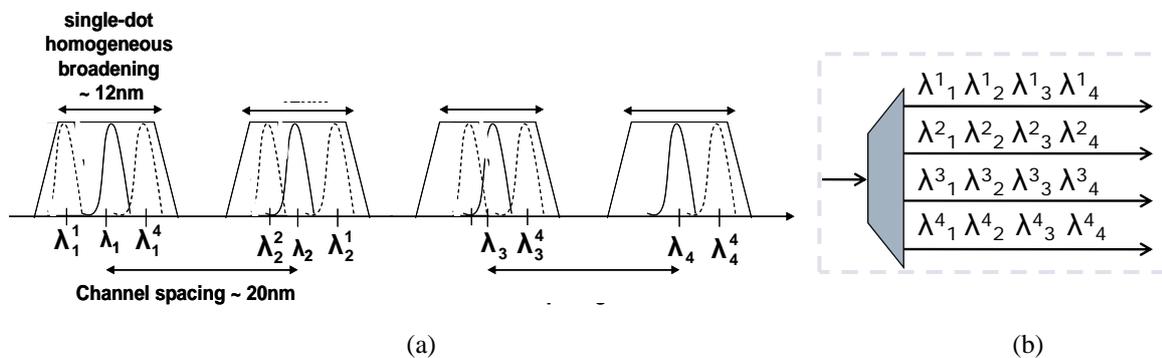


Figure 9. (a) Four channels lying inside four homogeneously broadened parts of an inhomogeneously broadened QD-SOA gain spectrum (b) corresponding AWG routing matrix

Collaborative actions carried out

Meetings (including tele-conferences): A face to face meeting was held in Poznan

Mobility actions:

Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)

The action will close.

Outcome of the joint research activity:

Joint publications:

Other:



4.9 JA12 - Encompassing switch node impairments and capabilities in dynamic optical networks.

Responsible partner: SSSUP-COM

Participants: SSSUP: Nicola Andriolli, Nicola Sambo, Mirco Scaffardi, Piero Castoldi; COM: Martin Nordal Petersen, Sarah Ruepp.

Description of work carried out.

In the second year of the Joint Activity, a mobility action performed by Dr. M. N. Petersen at SSSUP allowed to propose and realize a simple PMD monitoring method. This method does not require any alterations made to the transmitter and supports the commonly used non-return-to-zero (NRZ) modulation format. It is based on optical side-band filtering of the optical spectrum, which is broadened from self-phase-modulation (SPM) effects due to pulse shape changes caused by first order PMD. Experiments validating the technique have been carried out using 10 Gb/s NRZ data. Also, results from a long-term PMD monitoring experiment showing the method in use are presented. The experimental results are supported by numerical simulations carried out using Virtual Photonics Transmission maker VPI 8.0.

State of the art

Polarisation Mode Dispersion (PMD) has been identified as one of the important sources of signal distortion in modern high-speed optical communication systems [1]. With the increased interest on more transparent optical networks, PMD is getting more attention due to accumulation of the effect through optically switched paths in the network. As the size of the optical transparent network grows, the statistical possibility of having a situation with detrimental PMD effects grows. Thus, PMD monitoring becomes necessary in order to assure signal integrity.

Various methods of PMD monitoring based on different technologies have been developed [2,3]. Some of these require the signal format to be return-to-zero (RZ) [4] while others require alterations being made to the transmitter [5].

Principle of Operation and Experimental Setup

The proposed PMD monitoring technique exploits the effects of SPM induced spectral broadening in a non-linear medium due to high signal peak power [6]. Since PMD changes the shape of the optical pulse(s) any change due to PMD can be observed as a change in the optical spectrum. The degree of spectral broadening is dependant on SPM and the effect can be increased by utilizing a highly-non-linear-fibre (HNLf). Using a narrow optical filter the momentary changes can in turn be quantified and used as a 1st order PMD monitor signal.

The graph in Fig. 1 shows the effect of PMD, introduced with a PMD emulator, on the optical spectrum recorded after the HNLf. In the top of Fig. 1 is shown a diagram/setup explaining the concept of the PMD monitoring technique.

The influence of the differential-group-delay (DGD) on the spectrum can be clearly observed from the figure. The 0.1 nm enlarged window of the spectrum illustrates the optical filtering at -0.2 nm (-25GHz) respective to the carrier at 1550.350 nm.

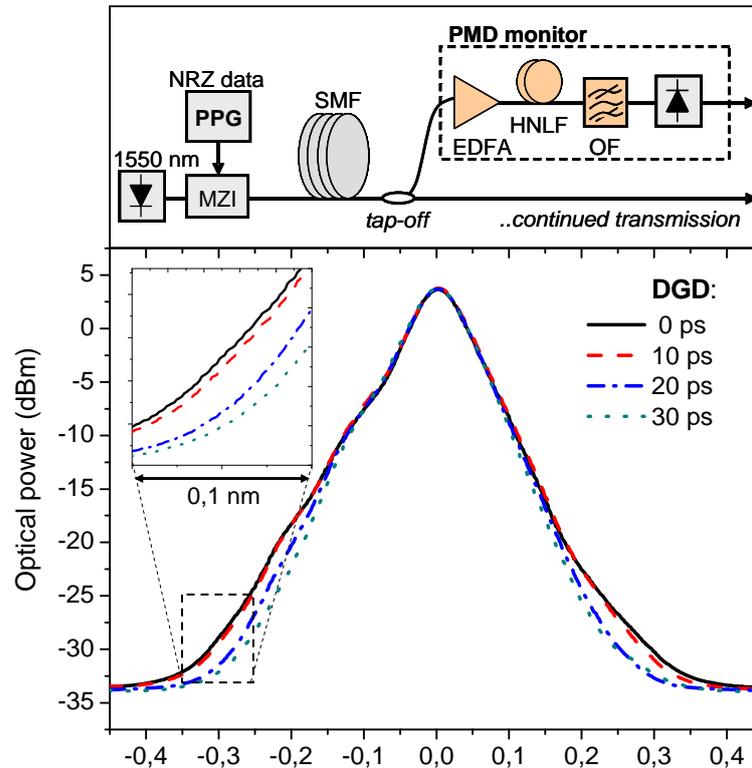


Figure 1 Setup and concept of PMD monitoring using sideband filtering of PMD-SPM induced spectral broadening. The spectrum is experimentally achieved using a 10 Gb/s NRZ signal.

The experimental setup consisted of a tuneable laser, Mach-Zehnder E/O modulator and a PPG creating the optical 10 Gb/s NRZ signal with a PRBS of 231 – 1. A spool of 10 year aged SMF fibre was available as PMD source. After transmission, a part of the PMD influenced signal was tapped off and sent to the PMD monitor, which is comprised of a EDFA with 23 dBm fixed output power, 1.0 km HNLF and a Gaussian optical filter with a 3 dB bandwidth of 0.1 nm. The slight detuning (-0.2 nm) of the optical filter is important in order to obtain a monitoring signal with high dynamic range. A simple power meter was used to read out the amount of filtered optical power. As will be seen in the next section, a 1st order PMD emulator was used to calibrate the monitor in order to be able to convert the monitor signal into a PMD value. It should be noted that only 1st order PMD (DGD) is monitored using this technique. 2nd order PMD is generally considered of less influence and can, for a majority of applications, be disregarded [7].

Experimental and numerical results

To verify the validity of the PMD monitoring method a 1st order PMD emulator from JDS Uniphase was used in the setup described above. While changing the DGD imposed on the signal, the monitor signal was recorded and the result can be seen in Fig. 2a. As expected, the spectral broadening and therefore also the monitor signal is reduced when the DGD increases. This is expected since the DGD causes a reduction in the peak power which in turn controls the level of SPM and spectral broadening.

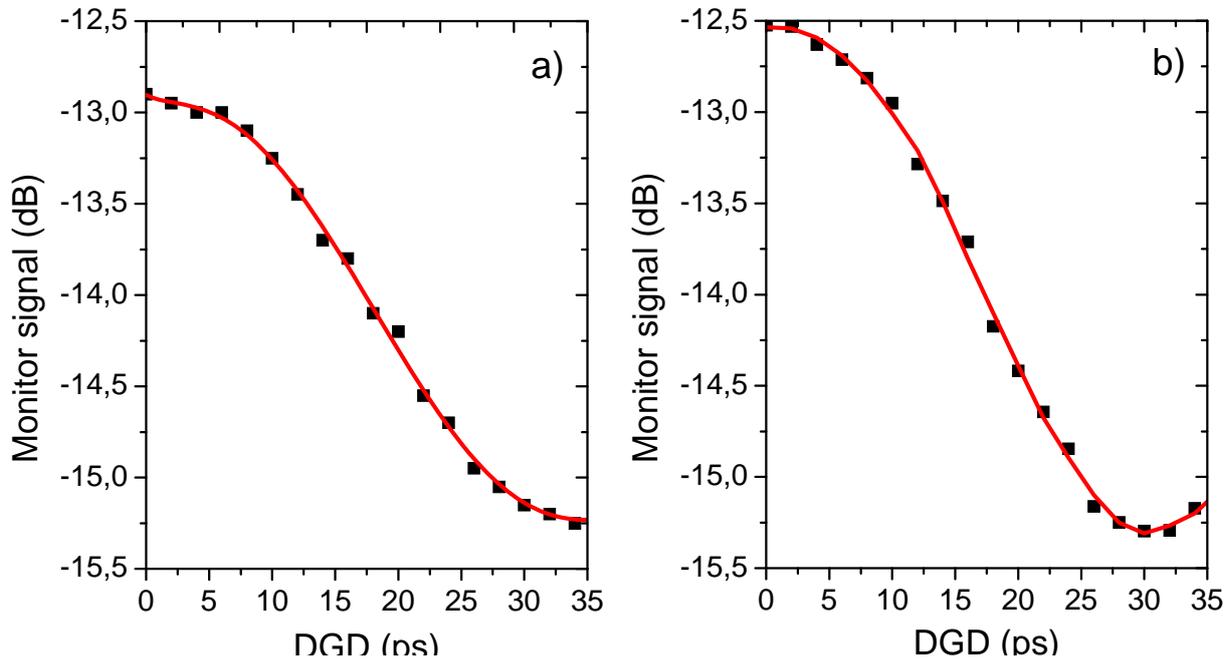


Figure 2 Monitoring signal versus DGD for a) experiments and b) numerical simulations. Both are for a 10 Gb/s NRZ signal and 0.1 nm optical filtering bandwidth.

The same was done numerically using VPI 8.0 and the result of the simulations can be seen in Fig. 2b. In both the experimental and numerical trials the bit rate was 10 Gb/s NRZ and the 0.1 nm Gaussian filter was detuned -0.2 nm respective to the carrier at 1550.35 nm. Comparing Fig. 2 a) and b) it is evident that the predictions of the behaviour from the numerical simulations are fulfilled also in the experimental case. Using the monitor signal, the above results thus give us a one-to-one measurement of the 1st order PMD in the region from 0 to ~ 33 ps. Beyond 33 ps DGD a single monitor signal cannot assure a full injective curve, which means that the monitored link or network should not give rise to DGD values above this value. This problem can be however solved with two distinct monitoring signals, obtained by simply filtering at two positions. For each position of the optical filter generating the monitor signal, a different curve is observed with a different minimum. This in turn gives the required information to deduce the DGD, also for DGD values beyond 33 ps.

Since 30 ps DGD corresponds to 1 dB power penalty for a 10 Gb/s NRZ signal, it was in these trials assumed to be a sufficient dynamic range.

In order to further test the usefulness of the PMD monitoring method reported here, a real-time trial was conducted using a 10 years old span of 80 km SMF fibre. The PMD of the fibre has previously been asserted to be around 6-8 ps. In order to further increase the PMD variations, the fibre spool was inserted into a temperature controlled chamber and cooled to 5°C and after three hours re-heated to 22°C. This treatment causes stress in the fibre and increases the PMD of the fiber [8] allowing us to observed stronger variations within a relative short time frame. Fig. 3 below shows the result from the real-time PMD monitoring trial.

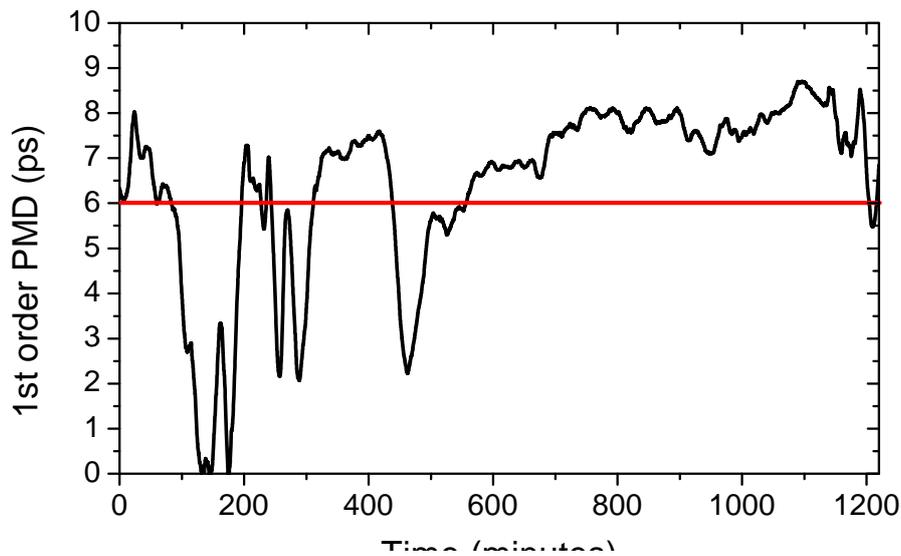


Figure 3 Real-time trial from a PMD measurement in a 80 km SMF + 10 km DCF span. The average DGD monitored is represented by the horizontal line at 6 ps.

At measurement start, time $T = 0$, the SMF fibre was placed in the 5°C chamber and allowed to cool for 180 minutes after which point it was removed, back to a room temperature of 22°C . The total measurement period went over a total of 20. hours. The rapid PMD variations in the beginning and at $T = 180$ minutes thus correspond to cooling and re-heating of the fibre respectively, while the relatively steady period after $T = 550$ minutes corresponds to the situation where the fibre temperature has settled to room temperature. At $T = 1150$ the spool was again cooled and again a more rapid PMD variation is observed. These measurements thus show a practical and real-time example of 1st order PMD monitoring.

Two day meeting in Pisa, 9-10 January 2009. Planning towards mobility action carried out in February 2009.

Mobility actions

Encompassing PMD impairments in a dynamic optical network

Martin Nordal, postdoc at COM, hosted by SSSUP from 16/02/2009 to 28/02/2009

Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)

As a future work, PMD monitoring will be exploited to provide a forecast of DGD dynamics. In the literature, typically a path is considered feasible if the average cumulated PMD is below a given threshold [9]. This is a pessimistic evaluation, since the PMD along rejected paths could be not so detrimental for the most of the time. To overcome this issue, the model in [10] based on DGD correlations can be exploited: given the measured DGD at time t_0 , the model provides an estimation of DGD at time t_1 (i.e., $\text{DGD}(t_1|\text{DGD}(t_0))$). By utilizing PMD monitoring and the model in [10], paths rejected by [9] could be accepted, resulting in a reduction of lightpath blocking probability and an increase of the network transparency domain.

**Outcome of the joint research activity:**

Joint publications: M. N. Petersen, N. Sambo, N. Andriolli, and M. Scaffardi; “PMD Monitoring using Optical Sideband Filtering”, in *Proceedings of 22nd Annual Meeting of the IEEE Photonics Society (LEOS2009)*, Belek-Antalya, Turkey, Oct. 4-8, 2009.

Other:

References

- [1] J. Strand, A. L. Chiu, and R. Tkach, “Issues for routing in the optical layer,” *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 81–87, Feb. 2001.
- [2] M. Brodsky, N. Frigo, M. Boroditsky, and M. Tur, “Polarization mode dispersion of installed fibers,” *J. Lightw. Technol.*, Dec. 2006.
- [3] J. Jiang, S. Sundhararajan, D. Richards, S. Oliva, M. Sullivan, and R. Hui, “PMD monitoring in traffic-carrying optical systems,” in *Proc. Of European Conference on Optical Communication, ECOC 2008*, Sep. 2008.
- [4] J. L. Blows et al., *ECOC 2005*, paper We.2.5.7.
- [5] S.M.R.M. Nezam, Y. Wang, M. Hauer, S. Lee, A.E. Willner, “Simultaneous PMD monitoring of several WDM channels using subcarrier tones”, *CLEO 2001*, pp. 561-563.
- [6] I.P. Kaminow et al., ‘*Optical Fiber Telecommunications V*,’ Elsevier, 2008.
- [7] M. Shtaif, A. Mecozzi, J.A. Nagel, “Mean-square magnitude of all orders of polarization mode dispersion and the relation with the bandwidth of the principal states”, *PTL*, 12(1), p. 53-5, 2000.
- [8] J. Cameron, L. Chen, X. Bao, J. Stears, “Time evolution of polarization mode dispersion in optical fibers”, *PTL*, 10(9), p. 1265-1267 1998.
- [9] D. Penninckx and C. Perret, “New physical analysis of 10-Gb/s transparent optical networks,” *IEEE Photon. Technol. Lett.*, vol. 15, no. 5, May 2003.
- [10] C. Antonelli, C. Colamarino, A. Mecozzi, and M. Brodsky, “A model for temporal evolution of PMD,” *IEEE Photon. Technol. Lett.*, vol. 20, no. 12, Jun. 2008.

4.10 JA14 - Photonic code label processors for ultrafast routing

Responsible partner: UoC

Participants: Prof Ian White, Prof Richard Penty, Dr Jose Rosas-Fernandez, Prof Gabriella Cincotti (Uniroma3), Prof Kevin Williams (TUE),

Description of work carried out

In this activity, we investigate the possibility of using chains or tree structures of ring resonators to implement multiport encoder/decoders (E/D). It is well known that ring resonators store a large amount of energy, since they can confine light into small cavities. Therefore, we expect to design new architectures of E/D with higher performances, with respect to arrayed waveguide (AWG)-based devices.

Different architectures for a multiport time/domain encoder/decoder have been considered: the simplest case is a chain or a tree of All Pass (AP) filters that introduce only a phase dispersion on the signal. Pass Band (PB) filters are obtained by considering a configuration where the ring is inserted between two parallel waveguides. An optimal choice of the coupling parameters allow us to design a device where the two output ports have the same energy. However the two outputs do not satisfy the half band property, that is required to design a multistage encoder/decoder.

UNIROMA3 have performed numerical simulations to investigate the behavior of ring resonators both in time and frequency domains (see Fig. 1)

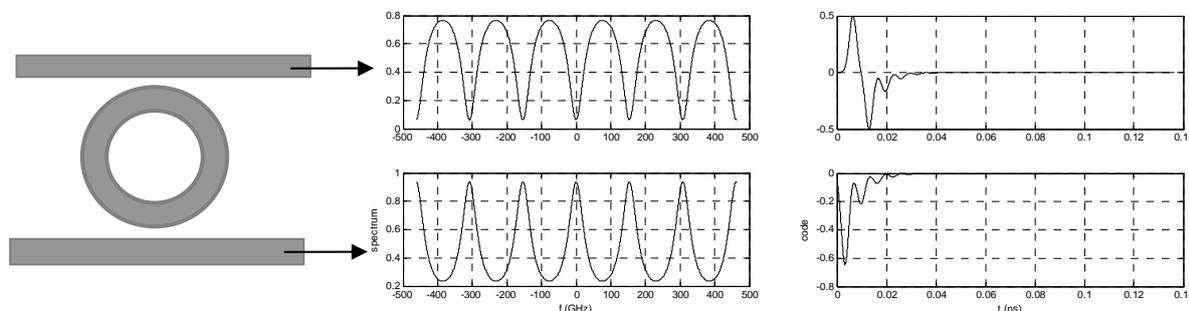


Figure 1: transfer function and impulse response of a ring resonator.

We have also been investigating the use of code labels for switching. In this case the labels use codes as headers and they are implemented electronically in order to compare it with optics implementation. This type of processing can find application in access systems. It has been argued that a fast recognition time would require optics. However, the recognition time is basically the transit time of the code label through the matched filter. UCAM has done some initial experiments which show that similar results can be obtained using recently available low cost electronic signal processing. Codes at 18 Gchip/s processing rates implemented by electronic finite impulse response (FIR) filters. They have already been demonstrated for PON applications. Thus, due to the matched filter technique used and the high-speed chip rate, fast label recognition techniques can be demonstrated. In this case, since

each tap has a separation of 55 ps and we use Hadamard codes with 8 chips length, a total recognition time for label switching is about 440 ps. This fast label recognition time, offering comparable recognition times to photonics, is demonstrated using low cost, low power consumption and flexible FIR filters.

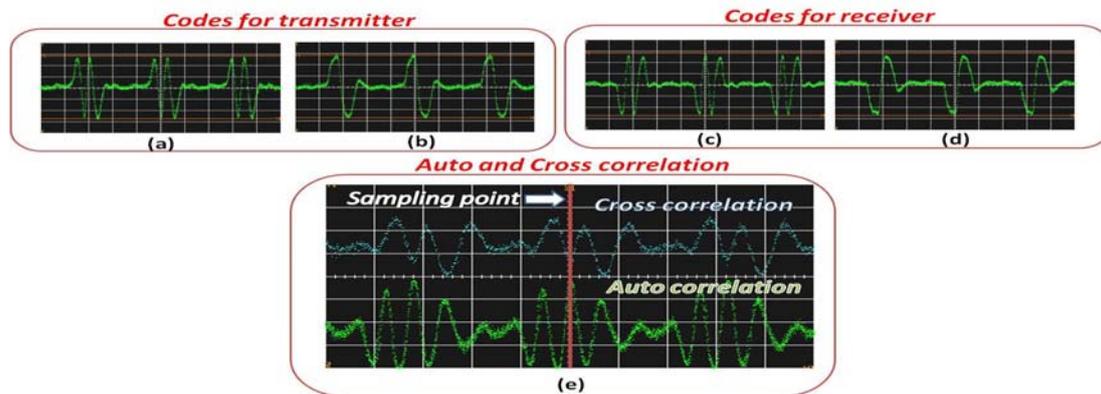


Figure 2 (a) Code 3 and (b) Code 5, of a Hadamard code family; (c) and (d) show the reverse codes of Code 3 and Code 5 for use in the receiver, respectively; (e) autocorrelation of Code 3 and cross-correlation of Code 3 with Code 5. All codes are programmed by the FIR filters for 18 Gchips/s label code generation; (a)-(d) are measured at point A and (e) at point C, of Fig. 2(a), respectively. Time axis 500 ns/div and voltage axis 20 mV/div.

The work at TU/e has focussed on the design and realisation of filter structures suited to fast tuning, and reliable fabrication in large scale circuits. A detailed 2D finite difference time domain study has been performed into fabrication tolerant microbends with controllable losses. Complex modal interaction is observed resulting in high sensitivity of losses to critical fab dimensions. Careful choice of radii does however allow optimum mode excitation and relaxed fabrication tolerance in low-loss microbends. A gated cyclic wavelength router suited to nanosecond timescale reconfigurability and high-connection architectures has been demonstrated for wavelength multiplexed end to end operation. A broad 100nm gain bandwidth with under 2dB gain variation is observed. Low power penalty routing of only 0.2dB is achieved for 3x10Gb/s multi-wavelength data.

Future study with investigate suitability for optical coding applications.

Collaborative actions carried out

Meetings (including tele-conferences)

UNIROMA3 and TUE had a technical discussion at OECC 2009, HongKong, China.

UNIROMA3 and UCAM had a technical discussion at ICTON 2009, Azores, Portugal.

UCAM and TUE has had three meetings in Cambridge on 23 Jul and 10 Sept 2009.

Mobility actions

Dr K. Williams will visit: Cambridge 20-22 January 2010



Future Activities - Timescale

(please include work plan for 3rd year with a timeplan for the work proposed)

We will proceed in this research considering the coupling between the ring and one or more waveguides using a multimode interference (MMI) coupler.

Outcome of the joint research activity:

Joint papers:

[1] "Loss mechanisms in dielectric optical micro-bends", R. Stabile, K.A. Williams, IEEE LEOS Benelux meeting, Brussels, 2009

[2] "Monolithic multiband wavelength router for fast reconfigurable data networking", A. Rohit, K.A. Williams, X.J.M. Leijtens, T. de Vries, Y.S. Oei, M.J.R. Heck, L.M. Augustin, R. Notzel, D.J. Robbins and M.K. Smit, Photonics in Switching, Pisa, 2009

[3] "Fast ECDMA labels for optical label switching", submitted to CLEO 2009.

Other:



4.11 JA16 - Low-crosstalk optical packet-switching architectures based on wavelength-switching and wavelength-sensitive devices.

Responsible partner: PoliMi

Participants: PoliMi, PoliTO, UPCT, UVIGO

Description of work carried out.

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 Tb/s [1]. The work has been focused on the issue of interconnecting line cards of routers/switches exploiting wavelength agility at transmitters to control routing of signals through an Arrayed Waveguide Grating (AWG) based structure.

This structure can be used as the central stage of the architecture represented in Fig. 10. The fabric can be divided into switching planes. Each plane hosts a subset of transmitters and receivers. Be S the number of switching planes and N/S the number of line-cards connected to each plane, in Fig. 10 an example with $S = 3$ and $N/S = 3$ is provided. The architecture is composed by:

- N line-cards subdivided in S planes. Each card consists of a tunable transmitter that operates at the bit rate of a single WDM channel, and of a receiver (ideally able to receive at any wavelength in use in the system);
- $N/S \times N/S$ AWGs (one per each plane). The well known routing properties of the AWG are exploited to steer each WDM channel from the source switching plane (the one containing the source card) to the appropriate destination switching plane (the one containing the target card);
- S (1: S) couplers. Each one collects channels coming from different switching planes;
- S EDFA amplifiers;
- S (1: N/S) WDM demultiplexers. Each one addresses WDM channel to the destination card within a plane. These demuxers may be implemented by using $N/S \times N/S$ AWGs again (with only one used inlet).

POLIMI approach to the feasibility and scalability analysis includes the following steps:

- 1) AWG design and crosstalk-impairment evaluation;
- 2) Scalability analysis;
- 3) Study of different strategies for crosstalk reduction;
- 4) Analysis of performance of the crosstalk-reduction strategies.

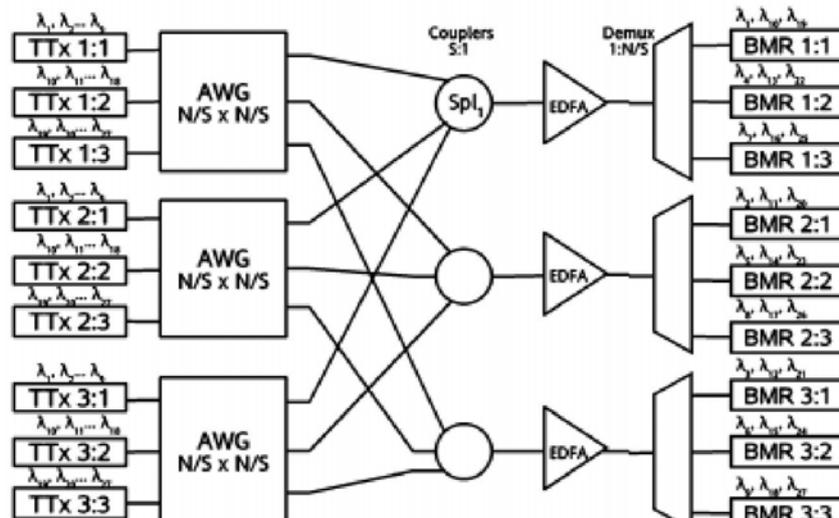


Fig. 10 AWG based architecture

Design of AWGs and crosstalk evaluation

POLIMI work aimed in first place at evaluating whether crosstalk is always a limiting factor in designing the proposed architectures. In order to analyse different feasible scenarios, this evaluation has been carried out with different values of inter-channel wavelength spacing $\Delta\lambda$ (25, 50, 100, 200 GHz) and for different transmission rates (2.5, 10, 40 Gb/s). With reference to the adopted technology, SiO_2/Si AWGs operating in third transmission window are assumed to be employed.

The design of the component goes through the following steps:

- (i) Calculate the Free Spectral Range (FSR): it is directly proportional to the number of channels (N/S) entering the AWG and to the spacing $\Delta\lambda$;
- (ii) Calculate the order of the transmission function (Q): it is inversely proportional to the FSR;
- (iii) Calculate the value of the geometrical distance between two adjacent waveguides in AWG central array (ΔL): it depends on Q ;
- (iv) Choose the number of guides in the central array (M): high values impact on the fabrication complexity of the device, but the number must be sufficient to guarantee enough bandwidth to switched WDM channels. Hence, channel bandwidth (B) has been set approximately equal to the channel bitrate, then M has been computed as a function of FSR and of B ;
- (v) Design the star couplers of the AWG (radius of curvature, angular distance between waveguides and width).

This design procedure leads to the computation of the transfer functions of the AWGs and consequently of the values for the adjacent and non-adjacent crosstalk coefficients X_A and X_N . These values are shown in Fig. 11. The column “ref.” contains reference values commonly adopted in literature [1] and similar to those declared by optical-component providers in the datasheets of commercial products.

	25 GHz	50 GHz	100 GHz	200 GHz	ref.	
2.5 Gb/s	-22 dB	-35 dB	-35 dB	-47 dB	-25 dB	X_A
	-28 dB	-40 dB	-42 dB	-52 dB	-30 dB	X_N
10 Gb/s	-18 dB	-22 dB	-29 dB	-36 dB	-25 dB	X_A
	-22 dB	-30 dB	-35 dB	-41 dB	-30 dB	X_N
40 Gb/s	-	-10 dB	-17 dB	-22 dB	-25 dB	X_A
	-	-18 dB	-22 dB	-29 dB	-30 dB	X_N

Fig. 11 Adjacent and Non-Adjacent crosstalk coefficients (X_A and X_N) for different values of $\Delta\lambda$ and bit rate

Model for scalability analysis

Scalability is evaluated in terms of total aggregate bandwidth by exploiting the model proposed in [2]. The aggregated bandwidth is calculated as the maximum number of allowed transmitters for the architecture multiplied by the bit-rate of a single channel. A power budget analysis has been conducted, taking physical impairments and penalties into account introduced by optical devices, in terms of:

- (i) average transmitted power for each transmitter;
- (ii) loss due to modulation and other effects;
- (iii) loss due to couplers (ideally, it logarithmically depend on the number of planes S);
- (iv) impairments related to AWGs:
 - insertion loss (function of the number of inputs N/S);
 - in-band and out-of-band crosstalk (functions of both adjacent and non-adjacent crosstalk coefficients);
- (v) loss due to demultiplexers (function of the number of inputs N/S).

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 performance¹.

The evaluation of the total aggregate bandwidth has been carried out considering:

- OSNR_{TX}² values ranging from 30 dB to 60 dB;
- wavelength spacing of 25, 50, 100 and 200 GHz between the WDM channels;
- channel transmission bit-rates of 2.5, 10 and 40 Gb/s.

The aggregate bandwidth is directly influenced by the channel transmission bit-rate (R_b) and by the value of N/S and S . The upper limit for the number of planes S is set to N/S (i.e. $S \leq N^{1/2}$). However, if both constraints on signal power at the receiver and on calculated OSNR at the photodiode must be satisfied, optimal S never reaches its maximum theoretical value.

¹ Defined over a bandwidth equal to the bit rate.

² It corresponds to the ratio between the useful laser power and the noise power due to spontaneous emission inside the laser.



Crosstalk reduction techniques

As one can argue, the first analysis made with $X_A = -25$ dB and $X_N = -30$ dB is not representative of all possible cases. In fact, some of these crosstalk values are not negligible and deeply impact on the scalability of AWG-based architectures. Hence, POLIMI considered two techniques that can be exploited to reduce crosstalk impairments.

Strategy 1: the interference given by adjacent crosstalk is higher for neighbouring AWG inlets and rapidly decreases as distance between inlets within an AWG increases. It is possible to design AWGs that use alternating input/output ports in order to suppress these contributions. Although this solution can potentially reduce total crosstalk power penalties, its effectiveness tends to vanish when the number of channels increases. As a matter of fact, for high values of N , the number of non-adjacent sources increases and non-adjacent crosstalk becomes the main limiting factor.

Strategy 2: coherent crosstalk represents an important source of power penalties and its impact can not be reduced by means of passive optical filtering. Coherent crosstalk can be reduced (or even eliminated) by exploiting the cyclic behaviour of the AWG routing function by using tunable transmitters spanning on different AWG Free Spectral Ranges (FSRs) [2]. Doing so, the same wavelength is never used more than once in different parts of the architecture and no coherent-crosstalk is introduced.

In order to evaluate the effectiveness of these strategies of crosstalk reduction on the scalability of the architecture, total aggregate bandwidth is computed by exploiting the previously proposed model. Five different solutions have been explored:

- A: Commonly-used commercial-product reference values of crosstalk coefficient ($X_A = -25$ dB and $X_N = -30$ dB);
- B: Values of crosstalk coefficient resulting from an accurate design of the AWG, according to actual channel spacing and channel bit-rate (X_A and X_N from Fig. 11);
- C: Reduced adjacent crosstalk exploiting strategy 1;
- D: Suppressed in-band (coherent) crosstalk exploiting strategy 2;
- E: Combination of C and D strategies.

Results are shown for three different scenarios in Figures 12, 13 and 14. In the first case (Fig 12, $\Delta\lambda = 50$ GHz, $R_b = 40$ Gb/s, $N/S_{MAX} = 100$), very high values of crosstalk are due to the design of the AWG. Hence, the application of the reduction techniques is fruitful. In this scenario and in the second one (Fig 13, $\Delta\lambda = 50$ GHz, $R_b = 10$ Gb/s, $N/S_{MAX} = 100$), the crosstalk given by the reference values underestimates the real crosstalk value which is provided by the design. In both cases, the reduction of adjacent crosstalk results in an improvement in the performance of the architecture. Nonetheless, since the number of managed channels is high, the penalties introduced by the adjacent crosstalk are not so important and the coherent crosstalk is the real bottleneck. As a matter of fact, employing both strategies does not give better results than the ones obtained with strategy D.

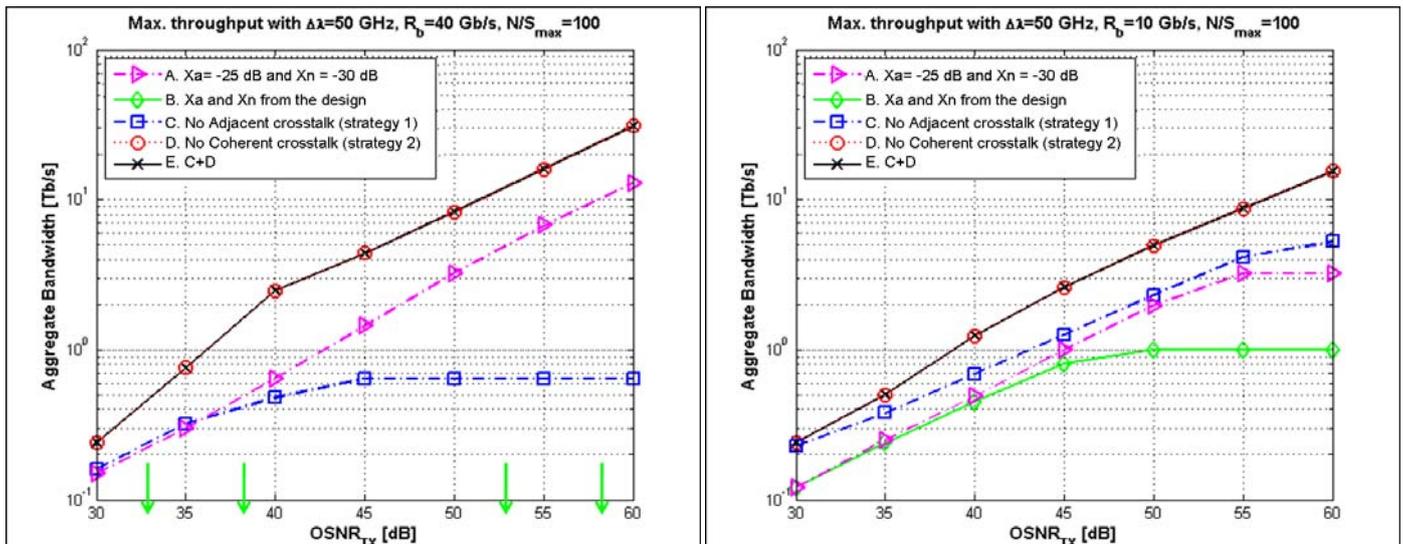


Fig. 12 and Fig. 13. Maximum throughput vs. OSNR with different crosstalk reduction techniques @ 40 and 10 Gbit/s

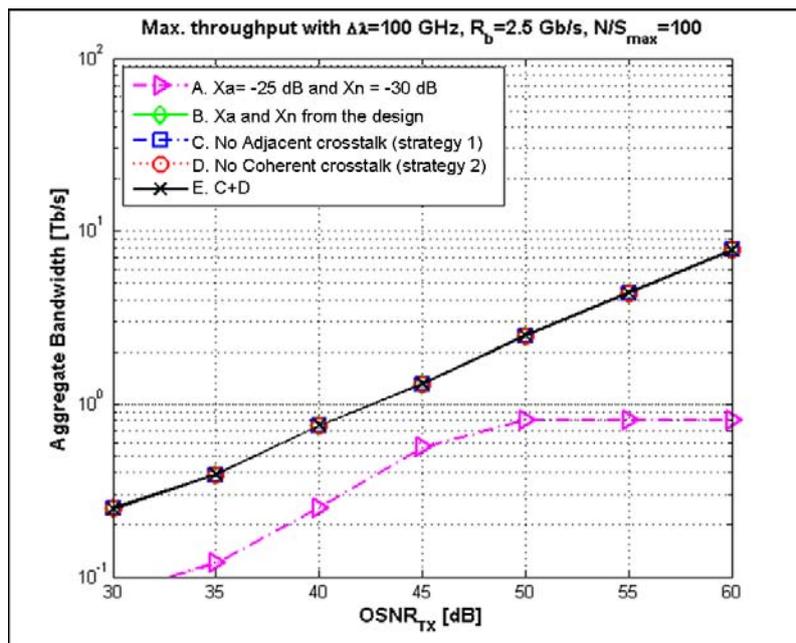


Fig. 14 Maximum throughput vs. OSNR with different crosstalk reduction techniques @ 2.5 Gbit/s

In Fig. 14, the third scenario is shown ($\Delta\lambda = 100$ GHz, $R_b = 2.5$ Gb/s, $N/S_{MAX} = 100$). In this case, the reference values of crosstalk coefficients ($X_A = -25$ dB and $X_N = -30$ dB) overestimate the real crosstalk value. It can be noticed that the performance of the architecture in this case is not bounded by the crosstalk, thus making crosstalk-reduction strategies useless.

Architecture scalability

The scalability of the architecture is presented below by referring to the case in which crosstalk reduction techniques are adopted when needed.

With reference to the number of transmitters/receivers in each plane, scalability with $N/S_{MAX} = 50$ and $N/S_{MAX} = 100$ has been investigated. Nominal transmitter OSNR allows calculating the first noise contribution to the architecture that deeply impacts on the scalability analysis. As the $OSNR_{TX}$ changes, the two constraints differently affect throughput: for values lower than 50 dB, OSNR at receivers is the limiting factor for the scalability; instead, for higher $OSNR_{TX}$ values, power constraint becomes more stringent.

Fig. 15 shows the maximum aggregate bandwidth as a function of $OSNR_{TX}$ for all the considered bit rates and for different values of N/S_{MAX} .

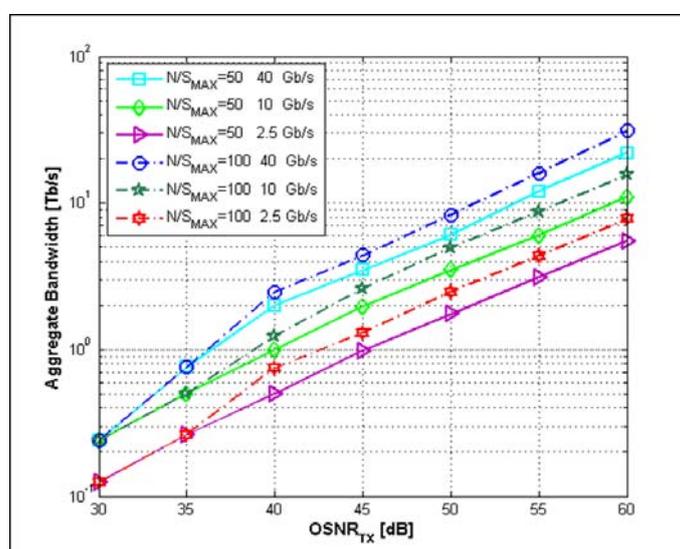


Fig. 15 Maximum throughput vs. OSNR for different bit rate and N/S_{MAX}

It can be noticed that very high capacities can be reached, in the order of tens of Tb/s. Moreover, being N/S_{MAX} equal, higher bit rates clearly allow higher bandwidth, and different bit rates show the same trend as $OSNR_{TX}$ increases. Finally, as expected, a higher bound on the maximum number of transmitters in each plane produces higher throughput for every bit rate considered.

References

- [1] R. Gaudino, G. Gavilanes Castillo, F. Neri, and J. Finochietto, "Simple optical fabrics for scalable terabit packet switches", in IEEE International Conference on Communications, 2008. ICC '08, May 2008, pp. 5331-5337.
- [2] D. Cuda, R. Gaudino, G. Gavilanes Castillo, F. Neri, G. Maier, C. Raffaelli and M. Savi, "Capacity/cost tradeoffs in optical switching fabrics for terabit packet switches", in International Conference on Optical Network Design and Modeling, 2009. ONDM 2009, Feb. 2009, pp. 1-6

Crosstalk-preventing Scheduling Algorithms for AWG-based Optical Switches (POLITO)

PoliTo is carrying analysis and experimentation by means of simulation on crosstalk-preventing scheduling algorithms for AWG-based optical switches. Figure 16 and Figure 17 explain the wavelength routing property in an $N \times N$ AWG with $N=4$, assuming that the n -th wavelength routes from input i to output j when $n = (j-i) \bmod N$; all transmitters at the inputs can tune an equal set of 4 wavelengths; graphically diagonals with the same color denote equal wavelengths. Despite the fact that they are used in different ports and have different routing properties, these equal wavelengths used at the same time lead to coherent crosstalk at the receiver [1]. We devised new scheduling schemes that limit the number of times each wavelength is used to perform the input-output matching in a given time slot (this number is denoted herein as k); the coherent crosstalk effect at the receiver is fully avoided as long as the target k is kept as low as 1.

Firstly PoliTo investigated the properties of k -legal matchings on a single-stage AWG-based switching fabric (See Figure 16) and showed that there are differences between an even or odd number of ports (N). For uniform traffic, the ideal case of $k=1$ (zero crosstalk switch) can be achieved for an odd number of ports, while a speedup of $1+1/N$ is required for an even number of ports, because no 1-legal permutation is possible in the latter case [2, 4]. For $k=2$ (one crosstalk channel is acceptable), no speedup is required for uniform traffic.

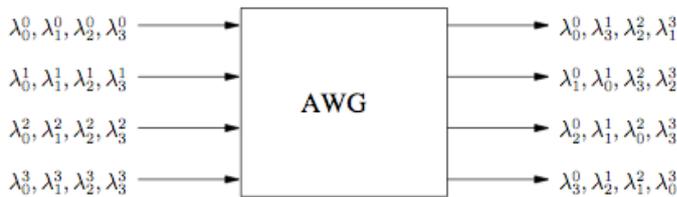


Figure 16. Arrayed Waveguide Grating (4x4) switching fabric

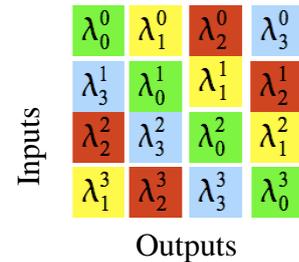


Figure 17. Wavelength routing property of an AWG, underlining identical wavelengths (with single color)

The worst case traffic scenario, that is, diagonal traffic, when the switch is forced to use most frequently the same wavelength (See the generalized diagonals with equally-subscripted wavelengths in Figure 17), requires a speedup of N/k . This makes the switch impractical, since the intention is to allow a value of k as low as possible at a reasonable cost.

PoliTO is testing a heuristically modified version of the i-SLIP scheduling algorithm to limit the amount of crosstalk in the switching fabric of input-queued cell-based switches. Other scheduling algorithms, such as a distributed version of iSLIP and 2DRR are being tested with the same objective in order to compare their performance.

An example of iSLIP scheduling algorithm modified with a lambda constraint is the following.

STEP 1: REQUEST. Each input i sends a request to an output j .

STEP 2: GRANT. Before granting to i, j checks the availability of lambda: A counter vector (lambda vector) has been defined for this purpose. Each field of the counter vector

corresponds to a lambda, so it counts the number of times a lambda is used. When i REQUESTs to transmit to j , j first checks in the lambda vector of wavelength $(j-i) \bmod N$ for the satisfaction of the k -legal constraint.

If value in lambda vector is lower than the target k , j grants i to transmit; otherwise j shifts to process the next request from another input port. A port pointer allows j to cycle among requests. On each timeslot (when a new matching calculation starts) the grant phase at j starts from a different port in a round robin basis. This step is performed for all j .

STEP 3: ACCEPT. i accepts one of the grants received (probably more than one output granted) and it increments and sets its pointer to $j \bmod N$ (in order to solve contentions). Similarly, j updates the lambda vector and sets its pointer to $i \bmod N$.

These three steps make an iteration. The algorithm can be iterated as many times as desired to achieve better performances. Figure 18 and Figure 19 show some preliminary results for uniform traffic on the achievable delay and maximum throughput as a function of the normalized input load for a 32 port fabric, comparing $k=1$ and 2, meaning zero crosstalk and 1 crosstalk channel respectively. The maximum sustainable throughput is reached when delays increase sharply, tending to a vertical asymptote (the increase is limited by the finiteness of input buffers). It is possible to observe the $1-1/N$ throughput limitation due to the even number of ports. It is also possible to observe that few iterations of the algorithm improve performance, while increasing the number of iterations from 16 to 32 is no longer beneficial.

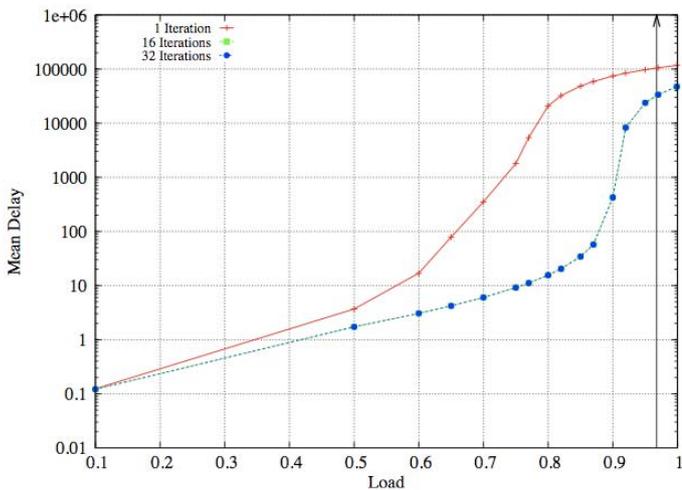


Figure 18. Results with 32 inputs and 1k-legal constraint.

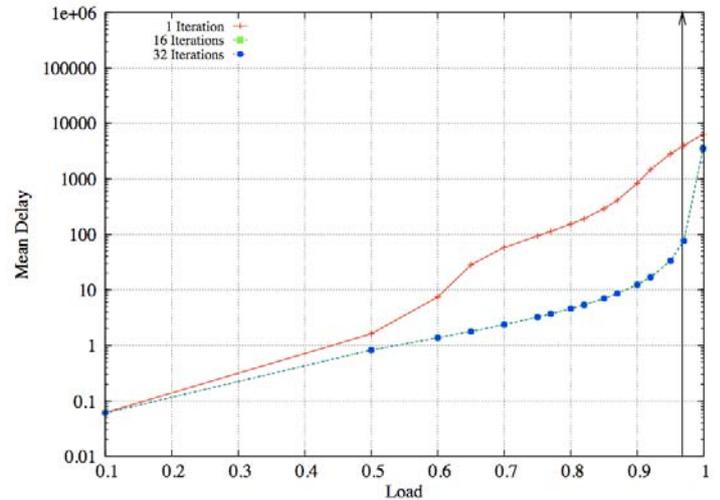


Figure 19. Results with 32 inputs and 2 k-legal constraint.

This work was in part done jointly with UVIGO, where the same problem was independently tackled, and with the support of the MS thesis work by an UPC student in mobility at PoliTO. Several fruitful discussions on the topic supported the joint research work. At the moment some other algorithms are being devised, and they are in a debugging phase and under discussion.

AWG crosstalk mitigation was faced also by exploiting horizontal extension of the switch in two cascaded stages (usually called Load Balanced Switch in the technical literature) as

depicted in Figure 20 and decomposing matchings accordingly. The goal is to maintain crosstalk within a certain limit by allowing an arbitrary k , and decomposing the matching that supports this k (usually called k -legal matching, or permutation) into two non-conflicting matchings.

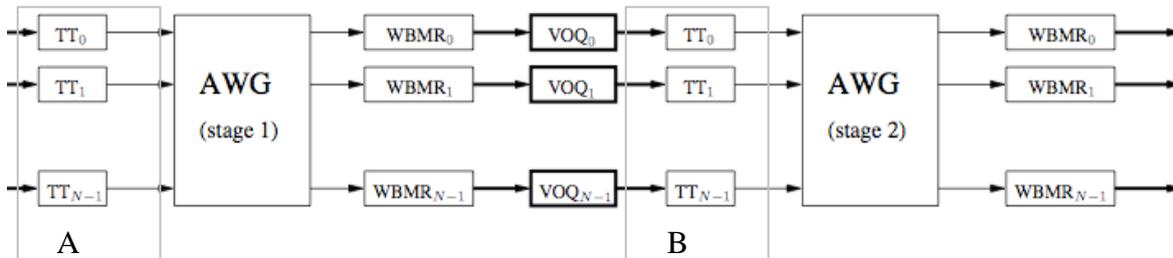


Figure 20. *Two-Stage AWG-based Switch.*

Since single-stage AWG-based switches require too large speedups to prevent crosstalk for general traffic patterns, the analysis for general traffic patterns was performed for two-stages load balanced switches. The first stage is used to redistribute traffic through all ports, and the second stage performs the scheduling for the resulting uniform traffic. This means that both switching stages are operated as in presence of uniform traffic, and that the previously obtained considerations for single-stage switches apply; Therefore a generic traffic matrix can be scheduled with a (space) speedup of 2 (for an odd number of ports) and $2+2/N$ in the case of even N). This represents a good advantage, since these spatial speedup avoids the use of faster (possibly unfeasible) components as expected in a time speedup.

Most of the load-balanced switch components are optical, except for the VOQs, which are constrained to be electronic. According to [3] there are non-stationary adversarial traffic patterns that make the load balanced switch's throughput to decrease arbitrarily. Additionally, the presence of the queues (and the corresponding resequencing logic) aggregate delays that reduce the QoS performance metrics. Under this motivation, a significant advantage can be obtained by a decomposition algorithm that is able to process the input-output matchings and give as a result a couple of k -legal matchings, for both switch stages to operate. Scheduling is performed in centralised way in one time slot, so the middle VOQs in the middle of the switch can be avoided, since the path through the whole two-stage switch (TT_i in A and B) is already set up once both k -legal matching are computed. The result is that a wavelength converter is enough to connect both stages (see [4]).

It was demonstrated that there is no 1-legal decomposition possible for any N ; for 2-legal it has been only possible to prove its existence by exhaustive search (with N up to 12). For k values higher than 2, PoliTo has provided a quadratic decomposition and correction algorithm (with $O(N^2)$ processing time) to transform k -legal to 4-legal matchings; these algorithms make use of transposition in the middle stage (which means wavelength conversion) to reduce wavelength overlapping on each AWG stage. Finally a proper estimation of the hardware resources required to run the algorithms is proposed, mapping the decomposition and correction algorithms to simple operations to implement in hardware.



With the above mentioned strategies, single stage heuristic algorithms and two-stage decomposition algorithms, coherent crosstalk can be avoided and making AWG-based optical architectures feasible.

References

- [1] H. Takahashi, K. Oda, H. Toba, *Impact of crosstalk in an arrayed-waveguide multiplexer on NN optical interconnection*, J. of Lightwave Technology, Vol.14, Iss.6, June 1996, pp. 1097-1105
- [2] M. Rodelgo-Lacruz, C. López-Bravo, F. J. González-Castaño, and H. J. Chao, *Practical Scalability of Wavelength Routing Switches*, in IEEE ICC 2009, 2009
- [3] I. Keslassy, *Load balanced router*, Ph.D. dissertation, Stanford University, June 2004.
- [4] D. Hay, A. Bianco, F. Neri, *Crosstalk-Preventing Scheduling in AWG-Based Cell Switches*, Globecom 2009, Honolulu, Hawaii, December 2009

Preventing Crosstalk in AWG-based Switches (UVIGO)

Although packet switches with aggregate bandwidth above several Terabits per second are currently commercially available, the number of interconnections and the power and information densities are reaching physical limits. These issues increase the difficulty of integrating a packet switch into a single equipment rack; thus, optical links are used to interconnect the line cards with the switching fabric. Packet switches with optical fabrics can potentially scale to higher capacities, improve reliability, and, at the same time, significantly reduce footprint and power consumption. A well-known alternative to implementing hardwired switches is the $N \times N$ Arrayed Waveguide Grating (AWG). Ideally, AWGs have insertion losses that do not depend on the number of input-output ports, thus leading to theoretical infinite scalability. However, an accurate second order assessment demonstrated that, in contrast to frequent assumptions in AWG-based designs, in-band crosstalk exponentially increases the power penalty, limiting the realistic useful size of AWG commercial devices to about 10-15 ports (13-18 dB) [ICC08].

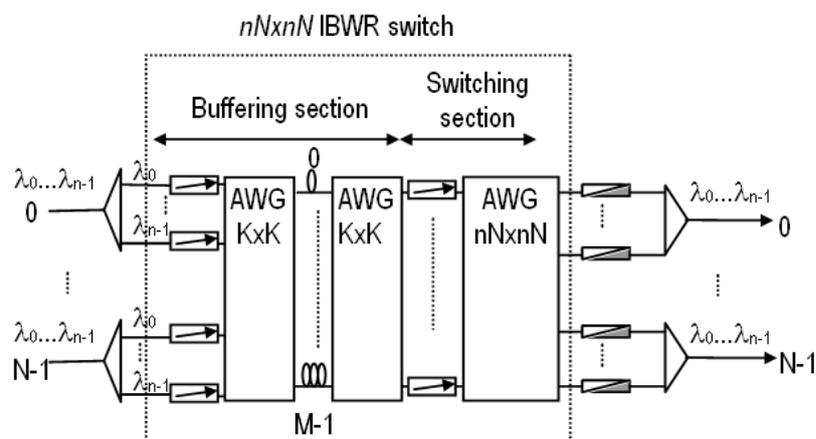
On the other hand, in-band crosstalk present at AWG outputs depends on the current connection pattern set by the scheduling algorithm. Port count limitations are calculated for the worst case. We have studied how this affected centralized and distributed schedulers, to determine a more realistic scalability limit. We have proved that the distributed schedulers that use predetermined connection patterns can avoid these harmful arrangements, and we have calculated the best permutation patterns to achieve the largest possible switch size. We have also shown that centralized general schedulers select those connection patterns with very low probability. For centralized general schedulers, we have calculated the probability of every connection pattern taking into account the in-band crosstalk penalty that allow us to dimension the switch discarding the connection patterns with probabilities far below the objective Bit Error Rate (BER). With these results, we have calculated more realistic port count limits for both scheduler types, concluding that was possible to implement much larger switches than predicted in the worst-case analysis (see [ICC09] and [NOC09] for details).

References

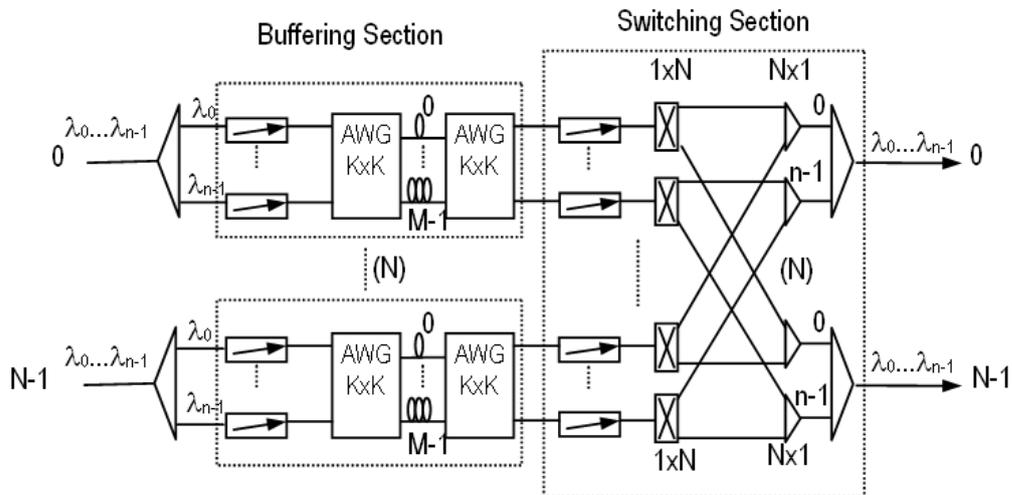
- [ICC08] J. M. Finochietto, R. Gaudino, G.A. Gavilanes Castillo, F. Neri, "Simple Optical Fabrics for Scalable Terabit Packet Switches," Proc. IEEE ICC 2008, Beijing, China, May 2008.
- [ICC09] M. Rodelgo Lacruz, C. López Bravo, F. J. González Castaño, "Preventing Crosstalk in AWG-based Switches with Distributed Schedulers", Proc. 14th Conference on Networks & Optical Communications, NOC 2009, Valladolid, Spain, June 2009.
- [NOC09] M. Rodelgo-Lacruz, C. López-Bravo, F. J. González-Castaño, H. J. Chao, "Practical Scalability of Wavelength Routing Switches," Proc. IEEE ICC 2009, Dresden, Germany, June 2009.

Evaluation of the Effects of In-Band Crosstalk in the IBWR Switch (UPCT)

UPCT investigated the effects of the in-band crosstalk in the Input-Buffered Wavelength-Routed switch, a switch fabric for Optical Packet Switching. The architecture of this switch fabric for N input and output fibers, and n wavelengths per fiber, adapted to WDM input and output ports, is shown below.



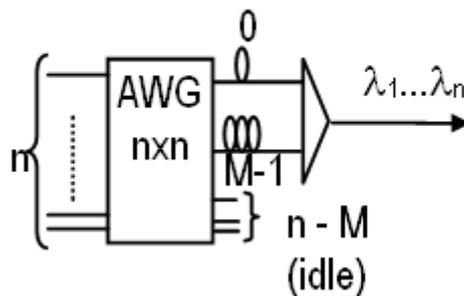
The scheduling of this architecture for slotted OPS networks has been studied in the past in several works like [Pav07] and [Cas09]. The coherent crosstalk in AWGs harms the switch performance, limiting the scalability of the architecture. However, the architecture can be modified to become scalable, as shown in the figure below:



From the scheduling point-of-view, this large-scale architecture is exactly equal to its lower-scale version. The size of the AWGs in the buffering section is now given by the number of wavelengths in the input fibers. The AWGs in the switching section can be eliminated by the combination of a tunable-wavelength converter followed by as many $1 \times N$ switches as input ports. This can be implemented by means $N^2 n$ optical gates, yielding a fairly scalable architecture (note that in the large scale DWDM scenario usually $n \gg N$). Note also, that in this version, nN fixed wavelength converters are eliminated thanks to the spatial switching capability given by the $1 \times N$ switches.

Observing the buffering stage, we note that the effects of crosstalk are limited in common large-scale situation, for which $M \ll n$. In this case, many of the output ports of the first AWG and their corresponding input ports of the second AWG are idle.

Buffering Section ($n > M$)



In this situation, the maximum number of times in which a wavelength is selected is limited by M (commonly from 2 to 4 according to the performance studies in [Pav07]). In practice, this mostly eliminates the scalability problems caused by coherent crosstalk.

The conclusion of the study is that effect of coherent crosstalk in the IB-WR architecture can be solved by using the large-scale version of the switching architecture shown above, without the need of particular changes to the scheduling algorithm.



References

[Pav07] P. Pavon-Marino, J. Garcia-Haro, A. Jajszczyk, "Parallel Desynchronized Block Matching: A Feasible Scheduling Algorithm for the Input-Buffered Wavelength-Routed Switch", *Computer Networks*, vol. 51, no. 15, pp. 4270-4283, October 2007.

[Cas09] F. J. González-Castaño, M. Rodelgo-Lacruz, P. Pavon-Marino, J. García-Haro, C. López-Bravo, J. Veiga-Gontán, C. Raffaelli, F. Gil-Castiñeira, "Guaranteeing packet order in IBWR Optical Packet Switches with Parallel Iterative Schedulers", to be published in *European Transactions on Telecommunications* journal.

Collaborative actions carried out

Meetings (including tele-conferences)

- 18 September teleconference. Participants: Miguel Rodelgo (UVIGO), Guido Gavilanes (PoliTo).
- 23 September teleconference. Participants: Miguel Rodelgo (UVIGO), Guido Gavilanes (PoliTo), Davide Cuda (PoliTo), M. Salvat (UPC).
- Meeting between POLIMI and PoliTo at Poznan, Poland 4 Oct 2009
- Meeting between UVIGO and PoliTo at Poznan, Poland 4 Oct 2009
- Partner's informal meetings during BONE Plenary Meeting 5-6 October, Poznan.
- 11 November: Meeting between D. Siracusa (POLIMI) and D. Cuda, G. Gavilanes and F. Neri (PoliTo), Torino, Italy

Mobility actions

1. Miguel Rodelgo Lacruz, PhD Student (UVIGO) visited PoliTo from 22/01/2009 to 23/01/2009
2. M. Salvat (UPC) worked at her MS thesis as PoliTo and is performing simulations on heuristically modified Schedulers to avoid crosstalk.

Future Activities - Timescale

PoliMi aims at continuing the study of the proposed architecture and the model adopted for the scalability analysis, comparing performance with other alternative optical-interconnection architectures.

PoliTO is continuing the work on heuristic input-queued scheduling algorithms with crosstalk prevention. A simulation-based comparison of some of these heuristics is being performed jointly with UVigo.

UVIGO have demonstrated that in-band crosstalk present at AWG outputs strongly depends on the connection pattern set by the scheduling algorithm. So, UVIGO designed variations of well-known scheduling algorithms that take into account in-band crosstalk and are able to



bound the number of ports that using the same wavelength. UVIGO is evaluating the algorithms and preparing a joint paper with the results.

UPCT will continue investigating the effects of in-band crosstalk in the scalability of AWG-based packet switching architectures.

Outcome of the joint research activity:

Joint publications

- D. Hay (PoliTO), A. Bianco (PoliTO), F. Neri (PoliTO), Crosstalk-Preventing Scheduling in AWG-Based Cell Switches, Globecom 2009, Honolulu, Hawaii, December 2009
- M. Rodelgo Lacruz, C. López Bravo, F. J. González Castaño, “Preventing Crosstalk in AWG-based Switches with Distributed Schedulers”, Proc. Of 14th Conference on Networks & Optical Communications, NOC 2009, Valladolid (Spain), June 2009.
- M. Rodelgo-Lacruz, C. López-Bravo, F. J. González-Castaño, H. J. Chao, ”Practical Scalability of Wavelength Routing Switches,” Proc. Of IEEE International Conference on Communications ICC 2009, Dresden (Germany), June 2009.

Planned joint publications

- A joint paper on switching architectures is planned between PoliMI and PoliTo.
- A joint paper on crosstalk-preventing heuristics is planned between PoliTo and UVigo.

4.12 JA17 - Novel Multi-granularity Optical Switching Node with Wavelength Management Pool Resources

Responsible partner: HUAWEI (Yabin Ye)

Participants: HUAWEI (Yabin Ye), RACTI (Kyriakos Vlachos)

Description of work carried out.

New, multi-granularity optical switching node architectures will be studied with wavelength management pool resource. Multi-granularity switching nodes play an important role in the optical switching mesh networks. The switching granularity can be classified into fiber switching (FXC), waveband switching (BXC), wavelength switching (WXC), and sub-wavelength switching (EXC). Previous multi-granularity nodes require a large switch core which is very complex, difficult to fabricate, and high cost [1, 2, 3]. Here, we will study novel multi-granularity optical switching node with new features. Compared to the conventional architectures published in the literatures [2, 3], the proposed architectures will be cost effective and more flexible.

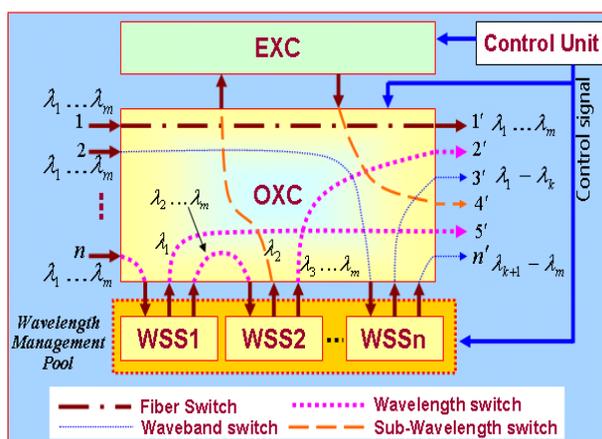


Fig.1 Proposed Node Architecture

The logical scheme of the proposed multi-granularity optical switching node is illustrated in Fig.1. The node is composed of three parts: the OXC (optical switch core) unit, the wavelength management pool unit, and the EXC unit. The center of the architecture is a bi-directional non-blocking optical switch core which can be realized by 3D MEMS or other technologies. While the wavelength management pool can be realized by bi-directional WSSs. Fig. 1 also shows the examples for multi-granularity switching functionalities:

- FXC: Any input fiber can be switched to any output port of the OXC directly, such as the $1 \rightarrow 1'$ lightpath illustrated in Fig.1.
- BXC: The input fiber is switched to the wavelength management pool, where wavelengths are divided into groups by WSS. For example, the input fiber 2 is switched to WSSn, where the wavelengths are divided into two groups: $\lambda_1 \dots \lambda_k$, $\lambda_{k+1} \dots \lambda_m$ ($1 < k < m$), and then routed to output port 3' and n' by optical switch respectively.



- WXC: The input fiber is also first switched to the wavelength management pool, by which the wavelengths are de-multiplexed. For instance, the input fiber n is switched to WSS1, and after de-multiplexing λ_1 is fed back to the optical switch and routed to any desired direction, e.g. the output port 5'.
- EXC: Sub-wavelength switching can be achieved by routing the de-multiplexed wavelengths into the EXC unit. Such as the λ_2 in fiber n , after it is de-multiplexed by WSS2, it is routed to EXC for sub-wavelength switching.

Apart from the multi-granularity switching functionalities, the proposed node architecture also has the following properties:

- Colorless Add/Drop: Since the output port of optical switch is wavelength independent, any wavelength can be routed to any output port.
- Directionless: Any wavelength can be routed to any direction.
- Convergence: Wavelengths from different fibers can be merged into one output fiber by using the WSS convergence characteristic
- Broadcast: The broadcast function can be provided by the WSS pool. Together with the optical switch, one any wavelength can be broadcasted to any output fibers.

Outcome of the joint research activity:

1 Joint paper to GLOBECOM 2010



4.13 JA18 – Comparison of the synchronous/asynchronous operation paradigm in optical switches

Responsible partner: PoliTo (Fabio Neri)

Participants: PoliTo, RACTI

Description

The constant evolution in optical switching has produced technologies and new approaches to optimize the forwarding operations in optical switches. Important questions such as the convenience of operating on fixed-size cells or variable-size packets, or of centralized versus distributed scheduling and resource allocation decisions have been addressed. Several alternatives in contention resolution, exploiting combinations of wavelength, space, and time domains have been studied. Almost all current solutions assume a synchronous operation of the switch, based on fixed-size time slots.

Switches are evolving towards multi-Terabit-per-second capacities, and thus must be highly scalable (in terms of port count) and decentralized devices; synchronous schedulers and optical interconnections are generally considered the best approach to enable this evolution.

Under the synchronous paradigm, much effort and a vast literature have been produced, both:

1. on the assumption of time-alignment of fixed-size data units (normally called cells) in the optical domain (from both transmission and scheduling points of view), and
2. on the attempt to maintain signal alignment (at the bit level) in practical implementations, using a single clock domain in the entire switch (while even small path length differences in fibers are critical in bit time-alignments).

As a consequence, for the former assumption, the arbitration schemes are often centralized, and additional effort is dedicated to support distributed operation (when the algorithms allow it). For the latter assumption, more complexity in transmission is required, since electronic/optical delay techniques should be employed to maintain time alignment. This constitutes a double effort that has driven research on the field and made part of the base for most of the state-of-the-art developments. It must be noted that time-domain operations are in general more difficult to realize in the optical domain than in the electronic domain.

The synchronous switch operation is in contrast with current variable-size Internet datagrams. Adaptation functions, including segmentation and reassembly, with the associated buffering, become necessary. Their complexity is obviously bit-rate dependent. Moreover, the synchronous operation transforms the statistical properties of network traffic (packet lengths, arrival times).

This joint activity faces the problem of considering the asynchronous operation of switches as an alternative paradigm that could unfold most of the usual assumptions and offer advantages in design, complexity and performance of future optical switch generations.

The motivation behind the analysis on synchronous/asynchronous switch operation is that there are many (possibly advantageous) alternatives to be explored behind this concept, and that asynchronous switches offer an option well worth to be explored. Preliminary results show that an asynchronous switch operation, based on variable-size packets and independent



scheduling decisions can offer good performance while reducing complexity (both at the technological level, and in switch control algorithms) significantly.

Several research areas are:

- Definition and performance of scheduling and resource allocation algorithms operating asynchronously
- Traffic scenarios where asynchronous/synchronous operation is relevant (beneficial and adversarial traffic patterns)
- Estimation of scheduling complexity in both paradigms
- Assessment of optical architecture complexity that exploit the asynchronicity
- Physical limitations arising with the asynchronous operation

Outcome of the joint research activity:

Joint paper submitted to an IEEE conference