POLITECNICO DI TORINO

SCUOLA DI DOTTORATO

Dottorato in Ingegneria Elettronica e delle Comunicazioni
XXV ciclo

Tesi di Dottorato

Optical architectures for high performance switching and routing

Miquel Garrich Alabarce

Tutore
prof. Paolo Giaccone

Coordinatore del corso di dottorato
prof. Ivo Montrosset

Febbraio 2013

Abstract

This thesis investigates optical interconnection networks for high performance switching and routing. Two main topics are studied.

The first topic regards the use of silicon microring resonators for short reach optical interconnects. Photonic technologies can help to overcome the intrinsic limitations of electronics when used in interconnects, short-distance transmissions and switching operations. This thesis considers the peculiar asymmetric losses of microring resonators since they pose unprecedented challenges for the design of the architecture and for the routing algorithms. It presents new interconnection architectures, proposes modifications on classical routing algorithms and achieves a better performance in terms of fabric complexity and scalability with respect to the state of the art. Subsequently, this thesis considers wavelength dimension capabilities of microring resonators in which wavelength reuse (i.e. crosstalk accumulation) presents impairments on the system performance. To this aim, it presents different crosstalk reduction techniques, a feasibility analysis for the design of microring resonators and a novel wavelength-agile routing matrix.

The second topic regards flexible resource allocation with adaptable infrastructure for elastic optical networks. In particular, it focus on Architecture on Demand (AoD), whereby optical node architectures can be reconfigured on the fly according to traffic requirements. This thesis includes results on the first flexible-grid optical spectrum networking field trial, carried out in a collaboration with University of Essex. Finally, it addresses several challenges that present the novel concept AoD by means of modeling and simulation. This thesis proposes an algorithm to perform automatic architecture synthesis, reports AoD scalability and power consumption results working under the proposed synthesis algorithm. Such results validate AoD as a flexible node concept that provides power efficiency and high switching capacity.

Acknowledgements

I am really grateful to Prof. Fabio Neri for his guidance, support and encouragement from the beginning of my research career, in September 2008. Under his supervision I carried out my master thesis and beginning of the PhD until he suddenly passed away in April 2011. I must thank him for sharing his principles, motivation and teachings that go beyond academic and research aspects. It has been an inestimable honor the opportunity to work under his supervision. Thank you, Fabio.

I am very grateful to Prof. Andrea Bianco and Prof. Paolo Giaccone for their time, kindness, guidance and very helpful advices during my master thesis and PhD. I thank Dr. Guido Gavilanes Castillo and Dr. Davide Cuda for their mentorship, encouragement and precious friendship. Indeed, their skills on presenting, writing and pursuing results have been a source of inspiration and a huge motivation for my research activities.

It has been a pleasure to be working in the Telecommunication Networks Group, at Politecnico di Torino, because their members' attitude eases the development of research, skills and background. I thank specially my closest colleagues during these years Ignacio Bermudez and Edoardo Bonetto for sharing research concerns and valuable friendship.

I am really thankful to Prof. Dimitra Simeonidou for allowing me to spend my abroad stage in the High Performance Networks group at University of Essex, U.K.. I thank specially Dr. George Zervas and Dr. Norberto Amaya for their supervision, advices, friendship and opportunity to do experimental research that allowed me to develop more skills and increase my background. I also thank my colleagues Mr. Bijan Rahimzadeh and Mr. Konstantinos Banias for their motivation and friendship during my abroad stage.

My deep gratitude to my family for their support and encouragement to overcome the difficulties to pursue my goals. My most sincere admiration is for my mother Maria Josefa for her precious incentive to develop myself in personal and professional aspects. Finally, I especially thank my beloved wife Patricia for being at my side when adversity happens but also for sharing gratification, prosperity and moments full of happiness.

Contents

Chapter 1

Thesis outline

Two different contexts present notable challenges for current optical networking technologies.

Short-reach optical interconnects

Chip multiprocessors (CMPs) are able to achieve performance improvements only if their increasing number of computational cores can collaborate with each other with efficient networks. In this context, optical networking technologies are becoming a viable option for solving major performance limitations in CMPs caused by bus bandwidth bottlenecks, latency and power consumption issues.

This thesis focuses on the use of silicon microring resonators for short reach optical interconnects. Microring resonators are small foot-print devices that can be directly integrated in complementary metal-oxide semiconductor (CMOS) technology and are suited to a wide range of applications including optical signal processing, filtering, delaying, modulating and, in particular, switching. Photonic technologies (and microring resonators in particular) can help to overcome the intrinsic limitations of electronics when used in interconnects, short-distance transmissions and switching operations. Indeed, microring resonators are key components to enable future optical networks-on-chip. The first part of this thesis that focuses on short-reach optical interconnects is organized as follows.

Chapter 2 studies the use of optical microring resonators as building blocks to create switching elements in on-chip optical interconnection networks. The peculiar asymmetric power losses of microring resonators are considered for the study of classical multistage interconnection networks, the introduction of new multistage interconnection networks and the modification of classical Paul algorithm used to set up connections in Clos networks. In this work it is assumed single wavelength operation. This chapter is partially derived from [1–3].

Chapter 3 explores the wavelength dimension capabilities of microring resonators to build optical switching fabrics based on wavelength routing matrices (WRM). Two approaches are considered. The first approach considers a *passive* WRM to carry out a crosstalk and feasibility analysis taking into account the geometrical limitations of microring resonators. The second approach introduces an *active* WRM to address the wavelength reuse problem. This chapter is partially derived from [4, 5].

Long-haul optical networks

Long-haul optical networks have to deal with an increasing highly variable and complex nature of internet traffic. To this aim, future carriers' backbone optical networks will be required to support a mixture of bit-rate channels allowing bandwidth-variable resource allocation. Indeed, there is a need of a transition from the current optical networks that exploit the well-defined slots of the (dense) wavelength division multiplexing (DWDM) standard towards a spectrum allocation performed in a flexible manner.

This thesis investigates the flexible resource allocation with adaptable infrastructure for elastic optical networks. In particular, it focus on Architecture on Demand (AoD), defined as an optical cross-connect that dynamically adapts its architecture in real time, in order to fulfill the switching and processing requirements of actual network traffic. AoD addresses many challenges, relevant for current operators of optical backbone networks, such as (i) enabling the coexistence of both legacy low bit-rate channels and future high-speed super-channels, while maintaining acceptable QoS levels, (ii) minimizing power consumption and (iii) achieving high spectral efficiency. Therefore, AoD is a novel optical interconnect concept suited for long-reach scenarios such as optical core networks. The second part of this thesis that focuses on long-haul optical networks is organized as follows.

Chapter 4 reports results from the first gridless networking field trial with flexible spectrum switching nodes where different bit-rate signals were successfully generated, transported, switched and defragmented using a flexible and custom spectrum allocation per channel. This chapter is partially derived from [6, 7].

Chapter 5 studies the Architecture-on-Demand (AoD) concept defined as an optical cross-connect (OXC) which dynamically synthesizes architectures suited to the switching and processing requirements of traffic. More in detail, AoD is based in a optical backplane (3D-MicroElectroMechanical Systems) that connects inputs, outputs and optical building modules providing an unprecedented flexibility. It is carried out a scalability analysis in terms of number of backplane cross-connections, and a power consumption analysis. This chapter is partially derived from [8, 9].

Part I

Short-reach optical interconnects

Chapter 2

Optical interconnection networks based on microrings

This chapter studies the use of optical microring resonators as building blocks to create switching elements in on-chip optical interconnection networks. Optical networks on-chip promise to overcome some of the limitations of current electronic networks on-chip. However, the peculiar asymmetric power losses of microring resonators impose new constraints on the design and control of on-chip optical networks.

The chapter begins with an introduction in Sec. 2.1 of optical networks on-chip, the usage of microring resonators and the motivation of this work. In Sec. 2.2, it describes simple microring-based SEs that can be used as building blocks to create larger interconnection networks. In Sec. 2.3, the chapter defines the *degradation index* and investigates the scalability properties of three classical multistage networks based on microring SEs in terms of area (measured as the number of deployed microrings) and degradation index. To optimize the tradeoff between area and degradation index, it also proposes two multistage networks in Sec. 2.4 and the mirroring technique in Sec. 2.5. In Sec. 2.6, it presents a variation of the classical Paul algorithm for Clos networks [10], to configure microring-based switching fabrics and drastically reduce the degradation index. The chapter ends with some design evaluations of the proposed solutions in Sec. 2.7, and with some conclusions in Sec. 2.8.

2.1 Introduction and motivation

Optical technologies appear as a viable alternative to electronics to build interconnection networks suited for on-chip integration. On the one hand,

predictions outlined by the International Technology Roadmap for Semiconductors [11] show that the main performance limitation of feasible on-die electronic interconnections is the length of metal-dielectric wirings, becoming critical for distance above the millimeters. Indeed, the higher the information densities to support, the stricter become the design constraints to be fulfilled, especially in terms of (i) electromagnetic compatibility issues, (ii) maximum distances which electronic signals can cover without regeneration, and (iii) power requirements. On the other hand, recent breakthroughs in CMOS-compatible silicon photonic integration [12] are boosting the penetration of optical technologies into interconnection systems [13, 14]. As a matter of fact, photonic technologies can transport huge information densities, their performance are roughly independent of the bitrate and offer the possibility to cover larger distances without regeneration. Several issues still remain to be solved, both at the component level, e.g., efficient light generation, and at the architectural level, where interconnection architectures exploiting the distinctive features available in the optical domain must be identified. In this chapter we focus on the architectural level, studying on-chip optical interconnection networks.

Silicon *microring resonators* represent one of the most promising optical devices to this end. Microring resonators are small foot-print devices, which have already proved to be suited for a wide range of applications, including signal processing, filtering, delaying or modulating optical signals. Furthermore, they have been considered to build sensors, modulators, microlasers and buffers [15, 16].

In this chapter, we investigate how to design and control interconnection networks based on microring resonators acting as optical Switching Elements (SEs), as proposed in [13, 17–19]. In a microring resonator SE (see for example Fig. 2.1(a)), an incoming optical signal can be either coupled to the ring (if the input signal wavelength is equal to the resonance wavelength of the microring) or it can continue along its path (if the input signal wavelength is different from the resonance wavelength). This allows to switch an incoming signal to a different output port, depending on the ring resonance wavelength.

In classical interconnection networks, input signals suffer power losses independent of the state of the SEs in the interconnection network. On the contrary, SEs based on microring resonators show an intrinsic asymmetric state-dependent behavior because optical signals coupled/not-coupled to the ring suffer different power penalties, depending on the specific microring design [19]. To enhance architecture scalability, it is crucial to minimize the power degradation experienced by *any* signal traveling through the interconnection network. To this aim, we propose different architecture designs and suitable algorithms to control the SEs configuration.

The major contributions in this chapter are: (i) two multistage interconnection networks that minimize the power losses, (ii) a mirroring technique to scale up the size of the network, and (iii) a control algorithm that considers the specific physical impairments. All these contributions allow to design and control the whole switching networks, leading to good performance in terms of complexity and port count scalability.

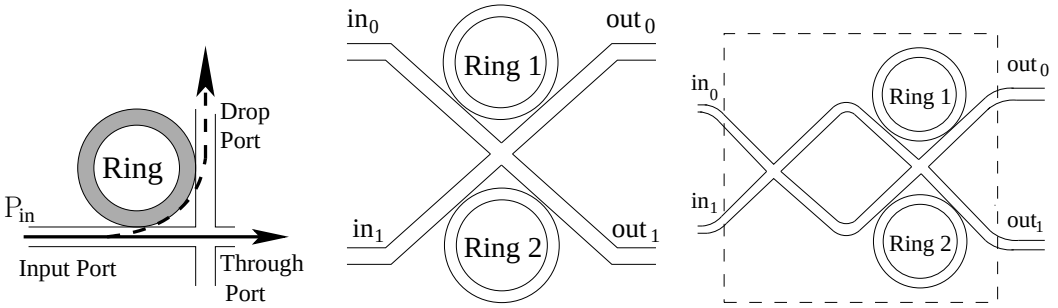
2.2 Microring-based switching elements

Microring resonators are basically composed by a waveguide bent to itself in a circular shape, coupled to one or two waveguides or to another microring. Fig. 2.1(a) shows an example of a microring coupled to two waveguides to build a 1×2 SE. This basic building block is called 1B-SE. Optical signals entering the *input* port can be either deflected to the *drop* port, when the ring is properly tuned (or in resonance) with the input signal wavelength, or continue along the input waveguide towards the *through* port in untuned state.

The resonance frequency can be modified exploiting either thermal-optical effects [20], carrier-injection [21, 22] or optical pumps [23]. Depending on the used technology, different tuning times, as well as different power penalties, can be observed. In the remainder of the chapter, we assume that microrings are controlled by carrier injection techniques because they ensure (short) switching times (few hundred ps according to [21]). We also assume a single wavelength operation; the incoming optical signal is coupled (or not) to the ring, depending on the wavelength to which the ring is tuned. All the proposed architectures can be extended to a Wavelength Division Multiplexing (WDM) scenario, but this is outside the scope of this work.

Dealing with a physical model of microrings (see for example [18]) is outside the scope of this work. Microring based 1B-SEs present an asymmetric power penalty depending on the switching state. According to [19], the design of the microring resonator is based on a tradeoff between different figures of merit. More precisely, a large coupling coefficient between microring and adjacent waveguides leads to low power degradation to signals deflected to the drop port but high degradation to signals sent towards the through port. On the contrary, a small coupling coefficient between microring and adjacent waveguides leads to high degradation to signals deflected to the drop port but low degradation to signals sent towards the through port. For instance, in the microrings controlled through laser pumps [23], input signals sent to the drop port experience larger power losses (around 1.5 dB) than signals routed to the through port (with negligible losses) because of the small coupling coefficient

between microring and drop/through waveguides. Therefore, we describe the loss behavior of the 1B-SE with a simplified model based on two loss states, assuming either a High-Loss State (HLS), or a Low-Loss State (LLS).



(a) 1×2 Basic SE(1B-SE) (b) 2×2 Basic SE (2B-SE) (c) 2×2 Mirrored SE (2M-SE)

Figure 2.1. Elementary microring-based switching elements (SE)

Extending the 1B-SE structure, it is possible to design more complex SEs that can be used as elementary switching elements in large interconnection networks. Fig. 2.1(b) depicts a possible implementation of a basic 2×2 SE (called 2B-SE). The 2B-SE exploits two 1B-SEs jointly controlled to provide two switching states. In the *bar* state ($in_0 \rightarrow out_0$, $in_1 \rightarrow out_1$), each ring deflects the corresponding optical input signal to the drop port of the respective 1B-SE. In the *cross* state ($in_0 \rightarrow out_1$, $in_1 \rightarrow out_0$), each ring lets the corresponding optical input signal pass to the through port of the respective 1B-SE. Also the 2B-SE can be modeled as a SE with two opposite loss states. Without loss of generality, we assume the bar state as a HLS, and the cross state as a LLS, coherently with either the experimental measurements of [17], or the results of [19] for a small coupling coefficient between microring and adjacent waveguides. Note that the penalty due to the waveguide crossing can be minimized applying an expansion technique to achieve a 0.2 dB of loss [24]. In [25], the loss due to the waveguide crossing is further reduced to 0.16 dB applying a double-etched technique.

We propose a modified version of the 2B-SE, useful to optimize the design of multistage interconnection networks. Fig. 2.1(c) shows a 2×2 Mirrored-SE (called 2M-SE). By cross-connecting the input ports, the bar state is now realized by setting up the internal 2×2 SE in the cross state (LLS), whereas the cross state is achieved with an internal bar state (HLS), thus exchanging the LLS with the HLS, with respect to a 2B-SE. The layout of the 2M-SE introduces additional bendings compared with the 2B-SE. However, the corresponding loss can be considered negligible for our analysis, because its value is approximately 0.005 dB for a $90°$ bend with a 6.5 μm radius [16].

Note that our characterization of the SE asymmetric loss behavior can be applied to other optical switching technologies. As an example, the Mach-Zehnder (MZ) based optical 2×2 switch presented in [26] shows a bar-state transmission loss 1 dB higher than the cross-state transmission loss. Similarly, [27] describes a MZ-based optical 2×2 switch with a bar-state transmission loss 0.7 dB higher than the cross-state transmission loss.

In this work we neglect the signal degradation due to crosstalk. Note that [19] shows that the effect of crosstalk depends on the specific switching configuration and on the coupling coefficients between the optical guides. Similarly to our assumptions, the signal impairment is asymmetric for each SE state, although it depends also on the traffic sent through the fabric. Following the approach described in [18], loss and crosstalk penalties could be integrated into a single "impairment index", thus permitting to extend the proposed methodology to devise a crosstalk-aware design. We leave this extension for future work. Instead, we focus on the design of interconnection network architectures able to scale to large size by reducing the number of SEs in HLS that optical signals should cross while traveling from input to output ports. Our methodology is independent of the specific switching configuration which creates either one of the loss states (HLS or LLS).

2.3 Basic multistage architectures

The SEs presented in Sec. 2.2 is used as building elements to assemble a larger interconnection network with N input and N output ports. Even if we borrow some terminology from the circuit-switching domain, the reference scenario is packet switching. The interconnection network must be configured to support a given set of input-output pairs defined by a permutation π: input i must be connected to output $\pi(i)$. The total number of possible complete permutations is $N!$. When we say that a connection from i to $\pi(i)$ must be established, we mean that a data packet from input i must be transferred to output $\pi(i)$ according to the decision of a packet scheduler. The sequence of SEs traversed by the connection defines a path in the switching network, computed by a proper routing algorithm.

We aim at maximizing the scalability of the architecture for large N by considering both the area and the power losses. The area, denoted by C, is evaluated as the *total number of microrings* present in the interconnection network.

The power losses are evaluated through the maximum number of SEs configured in HLS that the input signal must cross in the optical interconnection, considering all possible input-output permutations, i.e. taking a worst-case

approach. Such number is denoted by X and referred to as the *degradation index*. The architecture with the best scalability is achieved by minimizing C and X for a given N. More precisely, we consider a network design constrained by a given maximum degradation index $\hat{X}$, equal to the maximum number of HLS SEs that optical signals are allowed to cross. Whenever $X \leq \hat{X}$, then all the possible input-output permutations can be established and the corresponding architecture is said to be feasible. On the contrary, when $X > \hat{X}$, there exist some permutations for which some input-output connections cannot be established, because the corresponding path would violate the target $\hat{X}$. In such a case, the connection is said to be blocked, and the corresponding packet cannot be sent across the switching fabric.

In the following we will consider three basic architectures for interconnection networks: crossbars, Clos networks and Benes networks. Then, we will propose some hybrid architectures, to combine the different tradeoff between area and degradation index offered by each basic architecture.

2.3.1 Crossbar networks based on microrings

Firstly, we discuss the properties of the crossbar (XBAR) architecture, which will be used also as basic building block for the architectures that will be later proposed. Crossbars can be built exploiting 1B-SEs: Fig. 2.2 shows an example of crossbar, in which column waveguides are connected to the input ports and row waveguides to the outputs. The crossbar exhibits the best scalability in terms of degradation index, because each input can be connected to any output crossing a single 1B-SE in HLS. Hence, the maximum number of SEs in HLS that any optical signal crosses is $X_{\text{XBAR}}(N) = 1$. Routing is trivial. However, the crossbar exhibits the worst scalability in terms of area, because $C_{\text{XBAR}}(N) = N^2$.

2.3.2 Clos networks based on microrings

The definition of networks with lower number of SEs than the crossbar, naturally leads to multistage interconnection networks, like the well known Clos networks [10]. Fig. 2.3 shows a symmetric three-stage Clos network: each switching module (SM) of the first stage and of the third stage is connected with all the SMs of the middle stage. The first design parameter is the number of SMs in the first and the third stage, denoted by k. The second design parameter is the number of SMs in the second stage, which affects the blocking property. In the following, we consider $n = N/k$ SMs in the middle stage, because this is the minimum n that guarantees a non-blocking

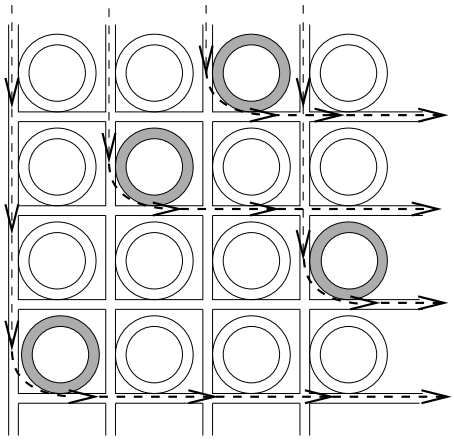


Figure 2.2. 4×4 microring-based crossbar connecting $1 \rightarrow 4$, $2 \rightarrow 2$, $3 \rightarrow 1$ and $4 \rightarrow 3$

rearrangeable network [1]. As a consequence, the first and third stage SMs are of size $n \times n$, and the second stage SMs are of size $k \times k$. When all SMs

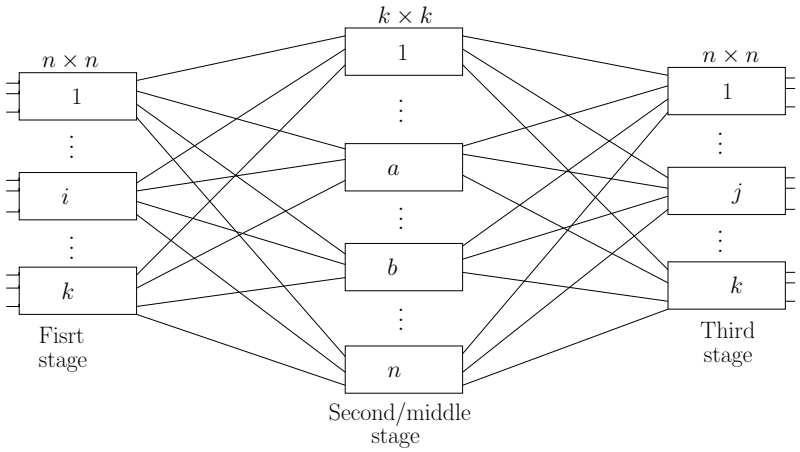


Figure 2.3. Rearrangeable non-blocking Clos network

are implemented by crossbars, $X_{\text{CLOS}} = 3$ and the minimum network area, obtained for $n = \sqrt{N/2}$ as shown in [10], is equal to $C_{\text{CLOS}}(N) = 2\sqrt{2}N^{\frac{3}{2}}$.

[1] As a reminder, a switching network is defined non-blocking if it is always possible to add a new connection from any free input to any free output port. A non-blocking network is defined rearrangeable, if it is possible to add such new connection by reconfiguring the pre-existing paths due to the previous connections [10]

2.3.3 Benes networks based on microrings

Benes networks are Clos networks that are recursively constructed with basic SMs of size 2×2; thus $N = 2^h$ for some integer h. From the area point of view, they offer the best scalability, being asymptotically optimal [10]. Indeed, an $N \times N$ Benes network has a number of stages (columns of 2B-SEs) equal to $2 \log_2 N - 1$, each stage including $N/2$ basic 2B-SEs. Hence, the area scales as

$$C_{\text{BENES}}(N) = 2N \log_2 N - N \tag{2.1}$$

because each 2×2 SE includes 2 rings. In terms of degradation index,

$$X_{\text{BENES}}(N) = 2 \log_2 N - 1 \tag{2.2}$$

since, in the worst case, there exists a path that passes through a HLS SE in each stage. Hence, Benes networks show a poor scalability in terms of degradation index, since X_{BENES} grows with the size N; given a maximum value $\hat{X}$ for the degradation index index, it is impossible to build networks with size $N > 2^{(\hat{X}+1)/2}$.

Table 2.1. Comparison between different network architectures, having defined $\hat{k} = 2^{(\hat{X}-1)/2}$.

Network	Area	Degradation index
Crossbar	N^2	$1 \leq \hat{X}$
Clos	$2\sqrt{2}N^{\frac{3}{2}}$	$3 \leq \hat{X}$
Benes	$2N \log_2 N - N$	$2 \log_2 N - 1 \leq \hat{X}$
M-Benes	$4N \log_2 N$	$\log_2 N \leq \hat{X}$
HCB	$2N^2/\hat{k} + N(\hat{X} - 2)$	$3 \leq \hat{X} \leq 2 \log_2 N - 1$
M-HCB	$4N^2/\hat{k}^2 + 2N(2\hat{X} - 3)$	$3 \leq \hat{X} \leq \log_2 N$
HBC	$N^2/\hat{k} + N(\hat{X} - 1)$	$1 \leq \hat{X} \leq 2 \log_2 N - 1$
M-HBC	$4N^2/\hat{k}^2 + 2N(\hat{X} - 2)$	$3 \leq \hat{X} \leq \log_2 N + 1$

2.4 Hybrid multistage architectures

The upper part of Table 2.1 provides a synoptic overview of the area and the degradation index for the basic architectures. To improve scalability for large N, we now combine the Clos network architecture (based on simple crossbar

with low degradation index) with the Benes network (with minimum area), considering two possible variants: i) the Hybrid Crossbar-Benes network and ii) the Hybrid Benes-Crossbar network.

2.4.1 Hybrid crossbar-Benes (HCB) architecture

The HCB network is depicted in Fig. 2.4 and consists of a Clos network in which the middle-stage modules are implemented through Benes networks, instead of crossbars. Like the original Clos network, the HCB network is constrained to $N = nk$, being $n \geq 2$ the size of the crossbars, and $k = 2^h$ the size of the Benes networks for some integer $h \geq 1$. Clearly, from the area perspective, the best solution is to make the Benes network as large as possible, up to $k = N/2$, when the HCB network degenerates into a Benes network.

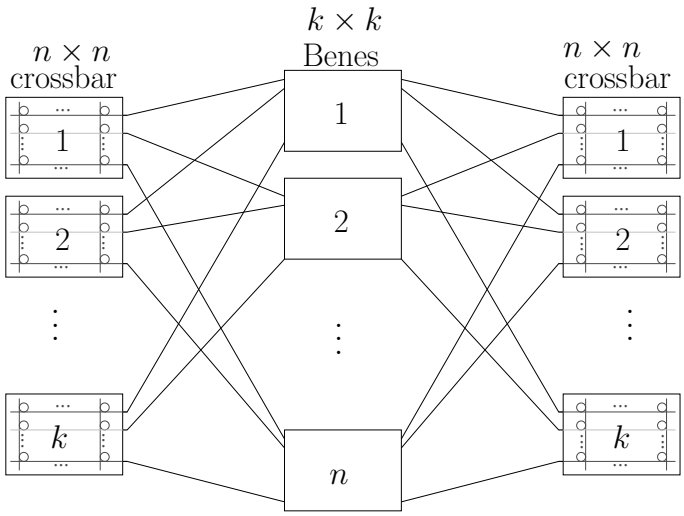


Figure 2.4. Hybrid Crossbar-Benes (HCB) network.

On the contrary, from the degradation index perspective, since

$$X_{\text{HCB}} = 2\log_2(N/n) + 1 \tag{2.3}$$

we must increase the crossbar size n. Thus, given a target $\hat{X} \geq 3$, it must be $X_{\text{HCB}} \leq \hat{X}$ and $n \geq N/\hat{k}$, having defined $\hat{k} = 2^{(\hat{X}-1)/2}$. The value $n = N/\hat{k}$

ensures both feasibility and minimum area. Finally,

$$C_{\text{HCB}}(N,\hat{X}) = 2\hat{k}C_{\text{XBAR}}(n) + nC_{\text{BENES}}(\hat{k}) = 2\frac{N^2}{\hat{k}} + N(2\log_2\hat{k} - 1) = 2\frac{N^2}{\hat{k}} + N(\hat{X} - 2) = \frac{N^2}{2^{(\hat{X}-3)/2}} + N(\hat{X} - 2) \quad (2.4)$$

for $3 \leq \hat{X} \leq 2\log_2 N - 1$. Note that (2.4) is lower than a crossbar for $N \geq (\hat{X} - 2)/(1 - 2^{-(\hat{X}-3)/2})$.

2.4.2 Hybrid Benes-crossbar (HBC) architecture

The HBC network is depicted in Fig. 2.5 (with $N = n2^h$) and consists of a Benes network factorized until a certain level $h \geq 0$, when the middle stage is substituted by $k = 2^h$ crossbars of size n.

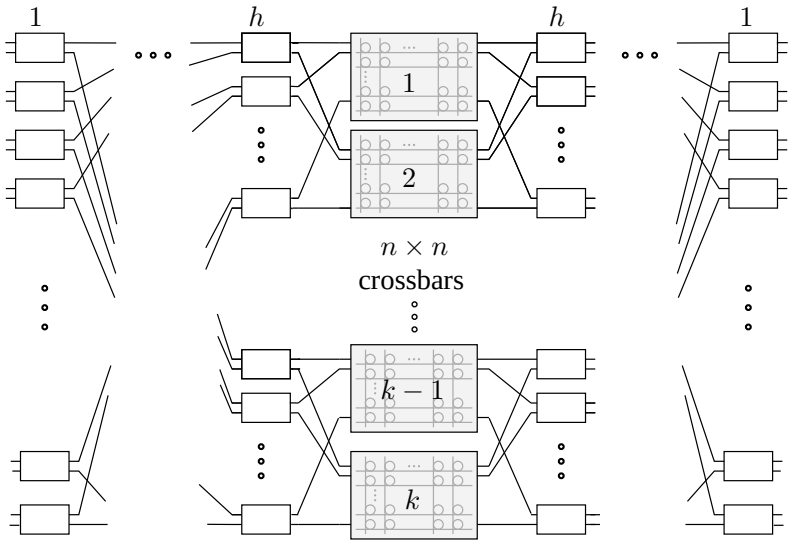


Figure 2.5. Hybrid Benes-Crossbar (HBC) network.

Again, from the area perspective, we must increase the "Benes part" up to $k = N/2$ when the HBC degenerates into a Benes network. However, from the degradation index point of view, since

$$X_{\text{HBC}} = 2h + 1 = 2\log_2(N/n) + 1 \quad (2.5)$$

the best solution is to decrease the level of factorization h (up to $h = 0$ when the HBC degenerates into a crossbar). Thus, to satisfy a given target $\hat{X} \geq 1$,

it must be $X_{\text{HBC}} \leq \hat{X}$ and $n \geq N/\hat{k}$, having defined $\hat{k} = 2^{\hat{h}}$ and $\hat{h} = (\hat{X} - 1)/2$. The value $n = N/\hat{k}$ ensures both feasibility and minimum area. Now the HBC area scales as:

$$C_{\text{HBC}}(N, \hat{X}) = \hat{k} C_{\text{XBAR}}(n) + 2(2\hat{h})(N/2) = \frac{N^2}{\hat{k}} + N(\hat{X} - 1) = \frac{N^2}{2^{(\hat{X}-1)/2}} + N(\hat{X} - 1) \quad (2.6)$$

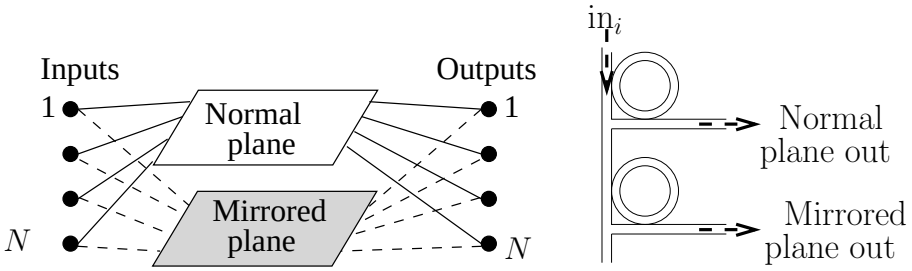
for $1 \leq \hat{X} \leq 2\log_2 N - 1$. Note that (2.6) is asymptotically half the area of the HCB and it is lower than a crossbar for $N \geq (\hat{X} - 1)/(1 - 2^{-(\hat{X}-1)/2})$.

2.5 Mirrored architectures

To further reduce the degradation index X or to achieve larger N for a given maximum degradation index $\hat{X}$, we propose the mirroring technique that exploits the spatial dimension, as shown in Fig. 2.6(a). We use two different switching planes, topologically identical: the *normal* plane is built with 2B-SEs only, whereas the *mirrored* plane with 2M-SEs only.

Depending on the architecture of the plane selector, the inputs are connected to both planes by means of either one of three solutions: i) the 1B-SE (Fig. 2.1(a)) which implicitly introduces a different degradation index to the signal depending on the plane, or ii) the plane selector depicted in Fig. 2.6(b), which introduces a constant degradation index to both planes, or iii), when possible, *integrating* the selector with an extension of the first stage of the architecture. For large N, in the i) and ii) plane selectors, the number of waveguide crossings may be significant, due to the connection from/to input/output ports to/from the fabric (see the solid and dashed lines in the left hand side of Fig. 2.6(a)). To address this issue we propose to adopt a vertically coupled microring resonator [28] for the 1B-SE in i) and for the microring connected to the normal plane in ii). By doing so, input ports are connected to the normal plane through waveguides on a "lower busline level" (solid lines in Fig.2.6(a)), and to the mirrored plane through waveguides in an "upper busline level" (dashed lines in Fig. 2.6(a)). According to [28], the required fabrication process for the two busline levels is feasible and would avoid power penalties due to waveguide crossing in our architecture. Since a similar design can be exploited with the couplers from the fabric to the outputs, we neglect losses due to waveguide crossing in our model.

We now consider the two fabrics, one for each plane. Since a HLS of a SE in one plane corresponds to a LLS in the other plane, a generic path that crosses X 2B-SEs in HLS in the normal plane, will cross $(S - X)$ 2M-SEs in HLS in



(a) Mirrored architecture (b) Plane selector

Figure 2.6. Mirrored architecture and plane selector

the mirrored plane, where S is the total number of traversed 2×2 SEs in each plane. This observation leads to a simple routing algorithm. After computing a path for the normal plane, whenever $X > S/2$, the routing algorithm will choose the path along the mirrored plane, i.e. the one with lowest degradation index; otherwise the path will be routed across the normal plane. As a consequence, we can start from a plane characterized by a degradation index X and build the whole network with degradation index:

$$X_M = \lfloor X/2 \rfloor + X_S \tag{2.7}$$

where X_S is the degradation index introduced by the selector; as a consequence, the degradation index is roughly halved by doubling the area of the "Benes part" of the network, constituted only by 2×2 SEs (i.e., doubling the number of 2×2 SEs). This allows to scale the size of the network given the same $\hat{X}$, as shown below for each architecture.

From the routing point of view, computing the path in the mirrored architecture has the same complexity as computing it for just a single plane, because of the identical topology of the normal plane and the mirrored one.

Note that mirroring the crossbars (based on 1B-SEs) does not provide any advantage in terms of degradation index. Two possible solutions to build the crossbars can be envisaged: i) the crossbars are replicated in both planes; ii) each crossbar is shared between the two planes, by adding one plane selector at the inputs of each crossbar, and one passive coupler at the outputs. Solution i) has the advantage of avoiding additional degradation index, but it introduces the area cost of replicating the crossbars. On the contrary, solution ii) shows a larger degradation index, but this is compensated by a smaller area. In the following, we considered solution i), whereas solution ii) is left for future works.

The mirroring technique will be denoted with the prefix *M-* and it is applied to Benes, HCB and HBC networks.

2.5.1 Mirrored Benes networks

In the M-Benes network, the selector is either a 1B-SE or the plane selector of Fig. 2.6(b). On the one hand, the 1B-SE introduces an asymmetric behavior in terms of degradation index, and makes more complex the layout to connect directly the drop port and the through port to each of the two planes. On the other hand, the plane selector adds a constant degradation index of one extra 1B-SE in HLS to all the paths, but it is simpler to implement due to its layout suited for an homogeneous selection between both planes. Therefore, we will assume always the plane selector for the M-Benes.

Since the Benes network presents an odd number of stages, when we build a plane with a maximum degradation index X_{BENES}, from (2.7) we obtain a M-Benes network with:

$$X_{\text{M-BENES}} = \frac{X_{\text{BENES}} - 1}{2} + 1 = \log_2 N \tag{2.8}$$

Let N be the maximum size of a Benes network satisfying $\hat{X}$; by (2.2), $\hat{X} = 2\log_2 N - 1$. A mirrored version satisfying $\hat{X}$ is built with two Benes networks in which the maximum degradation index is relaxed up to $2\hat{X} - 1$, according to (2.8). Hence, such Benes networks can have up to N' ports compatible with $2\hat{X} - 1 = 2\log_2 N' - 1$. By simple algebra, it can be shown that $N' = N^2/2$, i.e. the mirroring technique allows to scale the network by a quadratic factor.

The final area is simply obtained by considering the plane selectors and the two Benes networks:

$$C_{\text{M-BENES}}(N) = 2C_{\text{BENES}}(N) + 2N = 4N\log_2 N$$

2.5.2 Mirrored HBC networks

The M-HBC network requires a selector due to the Benes part at the edges of the network. Similarly to the M-Benes network, we assume the plane selector of Fig. 2.6(b).

Let X_{HBC} be the maximum degradation index in a single plane HBC network. Due to the even number of Benes stages that compose the HBC network and the single crossbar stage in the middle, the degradation index for the M-HBC network is:

$$X_{\text{M-HBC}} = \frac{X_{\text{HBC}} + 1}{2} + 1 = \log_2 \frac{N}{n} + 2 \tag{2.9}$$

Let $\hat{X}$ be the maximum degradation index satisfied in a HBC network. According to (2.9), we can build a M-HBC in which the maximum degradation

index is relaxed up to $2\hat{X} - 3$. Therefore, recalling that $\hat{k} = 2^{(\hat{X}-1)/2}$, the area scales as:

$$C_{\text{M-HBC}}(N, \hat{X}) = 2C_{\text{HBC}}(N, 2\hat{X} - 3) + 2N = 4\frac{N^2}{\hat{k}^2} + 2N(\hat{X} - 2) = \frac{N^2}{2^{\hat{X}-3}} + 2N(\hat{X} - 2) \quad (2.10)$$

for $3 \leq \hat{X} \leq \log_2 N + 1$. Note that for high values of N, (2.10) is lower than (2.6) for a degradation index $\hat{X} > 5$.

2.5.3 Mirrored HCB networks

Differently from the mirrored versions of Benes and HBC networks, it is possible to *integrate* the plane selector in the first stage crossbar. Indeed, in a mirrored HCB network, the first stage is composed by crossbars of size $n \times n$. The plane selector is connected to two crossbars (one for each plane) and allows to connect each input port with $2n$ output ports (n for each plane). This observation suggests a possible way to integrate n plane selectors and two $n \times n$ first-stage crossbars with just a single $n \times (2n)$ crossbar, reducing the area by $2N$ and the degradation index by one.

Let X_{HCB} be the maximum degradation index in a single plane HCB network. Due to the odd number of Benes stages that compose the HCB network and the crossbars stages at the edges, we obtain the following degradation index for the M-HCB:

$$X_{\text{M-HCB}} = \frac{X_{\text{HCB}} + 1}{2} = \log_2 \frac{N}{n} + 1 \quad (2.11)$$

Let $\hat{X}$ be the maximum degradation index satisfied in a HCB network. According to (2.11), we can build a M-HCB in which the maximum degradation index is relaxed up to $2\hat{X} - 1$. Therefore, recalling that $\hat{k} = 2^{(\hat{X}-1)/2}$, the area scales as:

$$C_{\text{M-HCB}}(N, \hat{X}) = 2C_{\text{HCB}}(N, 2\hat{X} - 1) = 4\frac{N^2}{\hat{k}^2} + 2N(2\hat{X} - 3) = \frac{N^2}{2^{\hat{X}-3}} + 2N(2\hat{X} - 3) \quad (2.12)$$

for $3 \leq \hat{X} \leq \log_2 N + 1$. Note that for high values of N, (2.12) is lower than (2.4) for a degradation index $\hat{X} > 3$.

2.6 Power-penalty-aware routing algorithm

In previous Sections 2.3-2.5 we have proposed different network architectures, whose area and degradation index have been summarized in Table 2.1. Independently of the configuration algorithm, both the crossbars and the Clos networks experience a fixed degradation index: $X_{\text{XBAR}} = 1$ and $X_{\text{CLOS}} = 3$. On the other hand, for all the other networks considered in this chapter, the degradation index depends on the path chosen by the routing algorithm to establish the required connections. In this section, we show the design of a routing algorithm, aware of HLS and LLS states, that reduces the degradation index.

Aim of the routing algorithm is to configure the state of each SE to satisfy any given input-output permutation π, as defined in Sec. 2.3. We consider the well-known Paull algorithm [10] which has been designed to configure a three-stage Clos network, but it is also suitable for multi-stage interconnection networks based on recursive Clos construction. At each recursion level, the network is abstracted as an equivalent three stage Clos network.

Referring to Fig. 2.3, consider a basic Clos network with I-stage and III-stage switching modules (SMs) of size $n \times n$; as a consequence, n modules are present in the II stage. The algorithm starts from the $N \times N$ network and from a given permutation π of size N, and then it computes the configuration of all the $2k + n$ SMs to establish the paths that route all the connections in π. Then, if the middle-stage SMs are, internally, Clos networks, the Paul algorithm is applied recursively to each individual $k \times k$ SM to establish all the required k connections.

Consider now just a basic Clos network that must be configured according to π. The Paull algorithm works in an incremental way, considering the inputs in an arbitrary order. When input x is considered, the algorithm computes the path to connect x to output $\pi(x)$ by configuring (see Fig. 2.3): (i) the I-stage SM where input x is located (we denote this as "module i"), (ii) the III-stage SM where output $\pi(x)$ is located (we denote this as "module j"), (iii) one or at most two SMs present in the II-stage. By construction, exactly one of the two cases can occur:

1. there exists a II-stage module a that can be connected to both modules i and j; in this case, II-stage module a is configured to support the connection from its internal input i to its internal output j; I-stage module i is configured to connect input x to its internal output a; III-stage module j is configured to connect its internal input a to $\pi(x)$;

2. otherwise, there exist two II-stage modules a and b, such that a can be

connected to module i, and b can be connected to module j; in this case, the algorithm moves a set of pre-existing connections from a to b, and another set from b to a, and accordingly recomputes the configurations of the I-stage modules and III-stage modules; then either a or b is used to support the new connection from x to $\pi(x)$, similarly to the previous case.

The Paull algorithm exploits many degrees of freedom to establish the connections given by π: (i) the sequence of inputs considered by adding the connections in π, (ii) the choice of a and b, (iii) the choice of the paths to be moved from/to a and b. Of course, each routing choice affects the degradation index experienced by the paths across the switching fabric.

We propose to modify the Paull algorithm to exploit such degrees of freedom in choosing a and b, among all the n possible II-stage modules, to minimize the degradation index. This modified version of the algorithm will be denoted by *PPA-Paull* (Power-Penalty-Aware Paull algorithm), in contrast with the classical version denoted simply by *Paull*. In the case $n > 2$, the I-stage and III-stage modules are crossbars and the degradation index introduced by them is always two, independently from the choice of the II-stage module: PPA-Paull chooses randomly a II-stage module that is currently available. In the case $n = 2$ (i.e. 2B-SEs at the I-stage and III-stage), the degradation index depends on the state of I-stage module i and III-stage module j. Let us assume that all the inputs and the outputs of the whole Clos network are numbered in increasing order starting from 1 and that the II-stage module a is the upper module, whereas b is the lower one. When an odd input is connected to an odd output, PPA-Paull will choose (if available) b to configure both 2B-SEs at the edges in LLS. Analogously, when an even input is connected to a even output, a will be chosen. Otherwise, either a or b will be chosen at random, since exactly one 2B-SE in the I-stage or III-stage will be in HLS (i.e., the degradation index will be always one).

2.7 Numerical evaluation

In the following Sec. 2.7.1 we investigate the scalability of different interconnection networks. Afterwards in Sec. 2.7.2 we investigate by means of simulations the degradation index reduction offered by the PPA-Paull algorithm with respect to the classical Paull algorithm.

2.7.1 Interconnection network design

We compare $N \times N$ interconnection networks by evaluating both the degradation index X and the area (in terms of number of microrings).

Fig. 2.7 shows the degradation index of different networks in function of N. Crossbars and Clos networks show a fixed degradation index, as discussed in Sections 2.3.1 and 2.3.2. On the contrary, the Benes network shows the worst scalability in terms of degradation index, because X scales logarithmically with respect to the number of inputs, coherently with (2.2). The HCB and the HBC networks (for the two different cases $n = 16$ and $n = 32$) show the same scalability law, as described by (2.3) and (2.5). As shown in Sec. 2.5, the mirrored technique roughly reduces X by a factor of two. Thus, if we impose a maximum degradation index $\hat{X}$, it is possible to build Benes networks with a number of ports equal to $N^2/2$ instead of N; similar gains apply to the Benes part of the other networks. Note that this advantage appears only for large N, as predicted by our models.

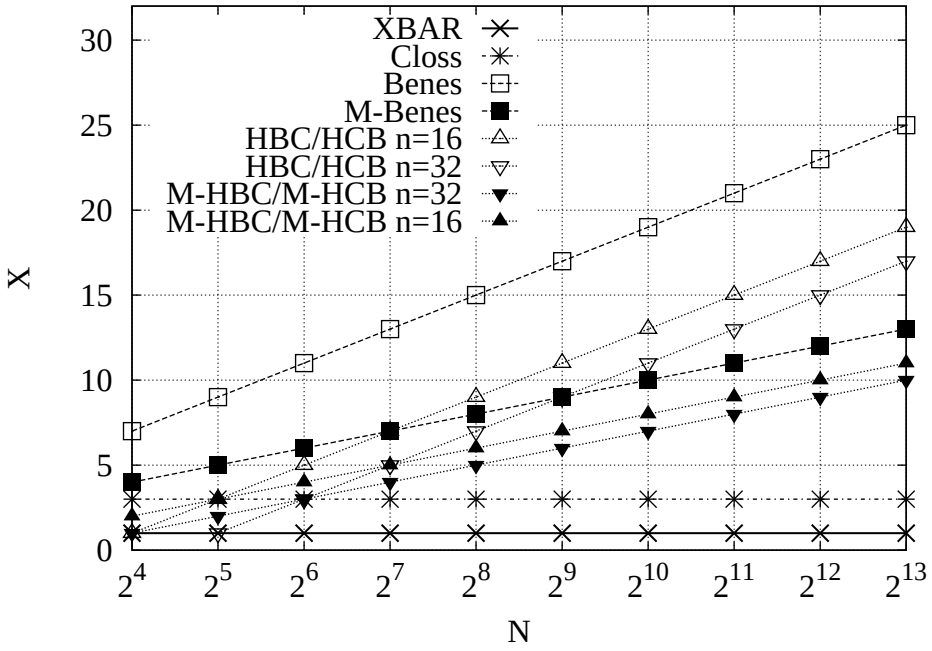


Figure 2.7. Degradation index for the studied architectures

Fig. 2.8 and Fig. 2.9 show the area for different interconnection networks for two different values of maximum power penalty; the points refer only to feasible configurations, i.e. compatible with the given target $\hat{X}$. The main figure refers to smaller networks (N ranging from 2^4 to 2^{10}), whereas the inset

figure refers to larger networks (N from 2^{10} to 2^{16}).

In general, the crossbar always shows the highest area, while Benes networks always the lowest one, whenever feasible.

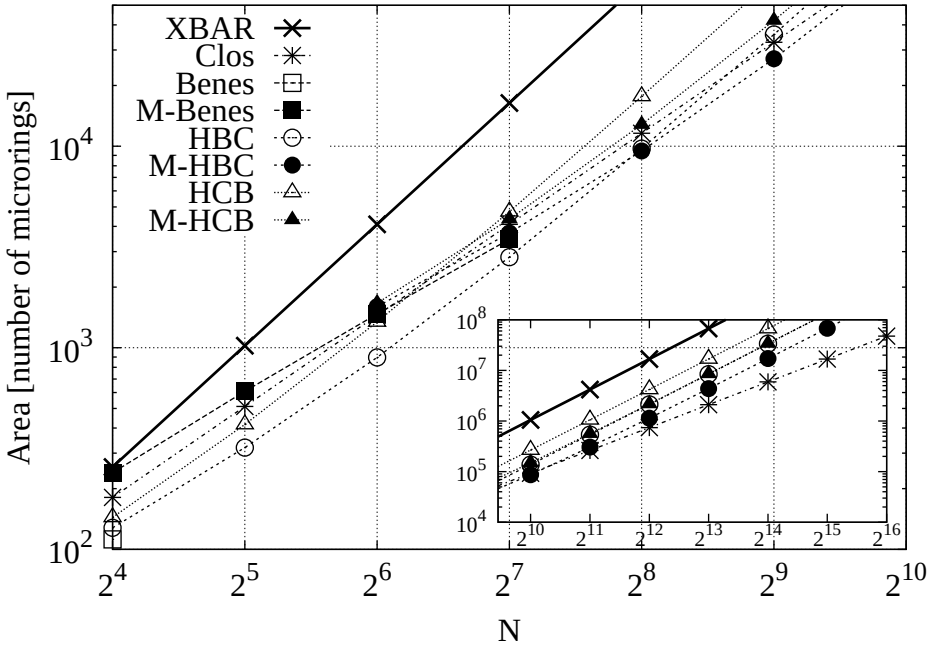


Figure 2.8. Area for maximum degradation index $\hat{X} = 7$

Fig. 2.8 shows that the only feasible Benes network, when $\hat{X} = 7$, is for $N = 16$. Instead, exploiting the mirroring technique, networks up to $N = 128$ can be built. Clos networks show the lowest area for very large N. Indeed, as the network size increases, the worst case degradation index grows. As a consequence, when $\hat{X}$ is smaller, edge or inner crossbars in HBC/HCB networks must be larger. On the contrary, Fig. 2.9 shows that, if the target $\hat{X}$ is larger, the size of crossbars inside the HBC and HCB networks decreases, reducing the overall area. Mirrored hybrid architectures reduce by a square factor the size of the internal crossbars. For low values of $\hat{X}$, the mirroring technique becomes area-effective for smaller N. HCB and HBC networks are feasible whenever the Benes network violates the HLS constraint. Among the two hybrid solutions, the HCB architecture exhibits a larger area than HBC network, and similar observations hold for their mirrored solutions.

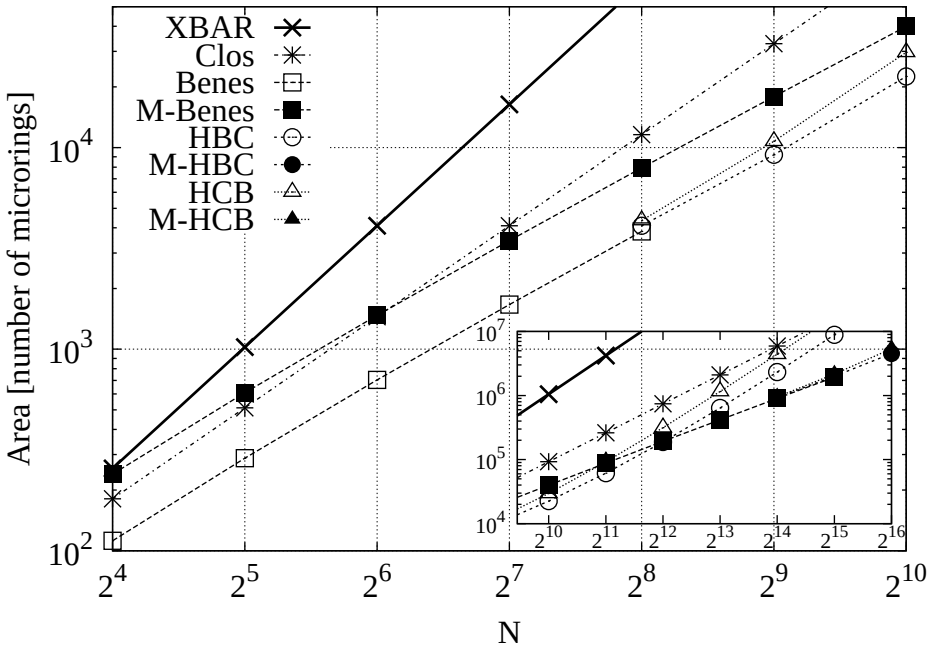


Figure 2.9. Area for maximum degradation index $\hat{X} = 15$

2.7.2 Performance of the routing algorithm

We compare the degradation index of the PPA-Paull algorithm with respect to the classical version of the Paull algorithm. We report results only for Benes networks, which provide an upper bound on the degradation index experienced by the other architectures.

We assume a microring based interconnection network, with synchronous operations, i.e., time is divided into intervals of fixed duration (timeslots) and the network transfers data units of fixed size (denoted as "cells"); the timeslot duration is equal to the transmission time of a cell. In the case of variable-size packets, incoming packets are chopped into cells, while outputs reassemble all the cells belonging to the same packet. We consider a uniform traffic scenario, with ρ being the average load at each input port. At each timeslot, we generate an input-output permutation π in which each input is active with probability ρ. Starting from a random input and considering all the other inputs in a sequential fashion, a sequence of consecutive connections is generated according to the rule: if input i is active, it is connected to $\pi(i)$. The routing algorithm adds each connection at the time in an incremental way, during the same timeslot.

Each path computed by the algorithm will result in a certain power penalty

index; if such index is larger than $\hat{X}$, the corresponding connection is blocked. To compare the routing algorithms, we measure the blocking probability $P_B(\hat{X})$ as the average fraction of blocked input-output pairs over the number of active inputs. As a complementary measurement, the throughput is evaluated as the average number of connections that are established without blocking in a generic timeslot; the maximum throughput is equal to the average load ρ and it is reached when a connection is never blocked, for any permutation. In the figures, the throughput and the blocking probability are averaged across many timeslots.

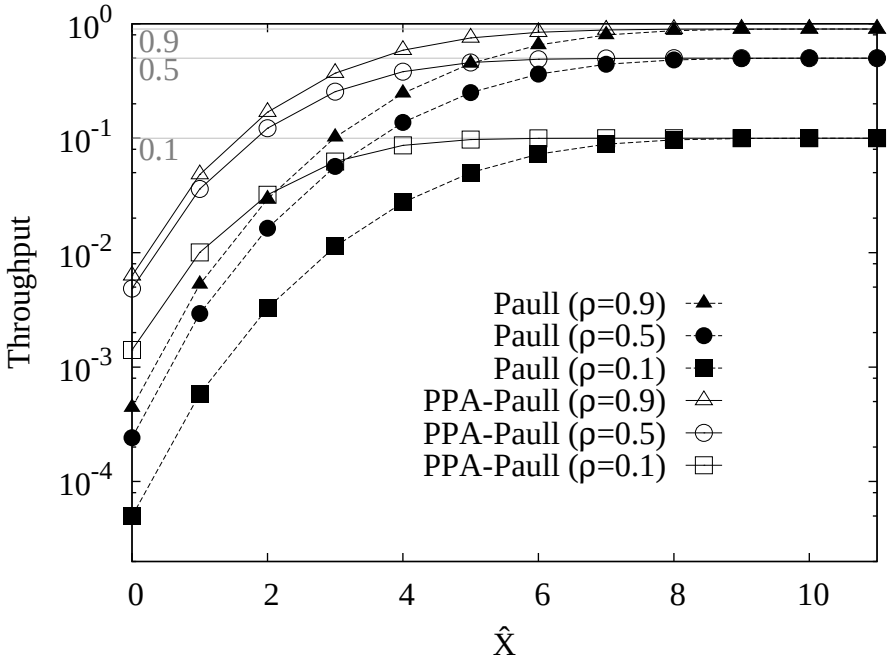


Figure 2.10. Throughput under uniform traffic for $N = 64$.

Fig. 2.10 shows the throughput in function of the maximum degradation index $\hat{X}$ for a 64×64 Benes network and different input loads: $\rho \in \{0.1,0.5,0.9\}$, corresponding to a lightly, medium and highly loaded network, respectively. For enough large $\hat{X}$, the throughput reaches its maximum value (0.1, 0.5 or 0.9 for each couple of curves), since the routing is not affected by the degradation index; in such case, all the algorithms behave the same. Note that the number of stages for the considered network is 11, hence larger values of $\hat{X}$ are not affecting the routing. On the other side, smaller values of $\hat{X}$ reduce the possibility of finding feasible paths; in the extreme case, the throughput approaches zero. In general, PPA-Paull achieves always a better throughput than Paull. When the input load ρ increases, the minimum $\hat{X}$ to achieve the

maximum throughput increases, since the routing is more constrained by a larger number of preliminary paths added during the current timeslot.

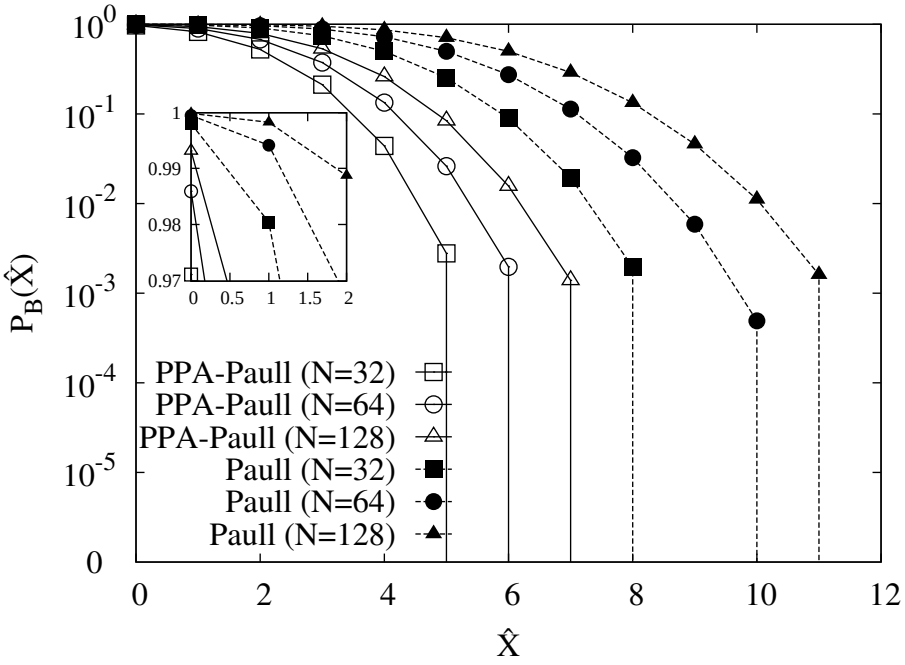


Figure 2.11. Blocking probability in Benes networks with $\rho = 0.1$

Fig. 2.11, Fig. 2.12 and Fig. 2.13 show the blocking probability in function of the maximum power penalty, each figure referring to a different value of input load. The smaller plot inside each figure details the blocking probability for low values of $\hat{X}$.

The total number of stages $S(N)$ in function of N are $S(32) = 9$, $S(64) = 11$ and $S(128) = 13$; whenever $\hat{X} > S(N)$, the blocking probability is zero by construction and the maximum throughput is achieved. On the contrary, when $\hat{X}$ approaches zero, the routing is severely constrained by the degradation index: the blocking probability increases and the throughput tends to zero. Furthermore, as N increases, in all the figures the blocking probability increases due to the larger network depth. In general, the reduction in the blocking probability due to PPA-Paull with respect to Paull is very large, reaching more than two orders of magnitude in some cases.

To better understand such results, consider the case in which just one path must be connected (this event may happen for low input load). If we set $\hat{X} = 0$, there will be only one specific destination (among N possible ones) reachable by each input; the corresponding path will be found by PPA-Paull. Thus, at low load and under uniform traffic, we can expect $P_B^{\text{PPA-Paull}}(0) = 1 - 1/N$

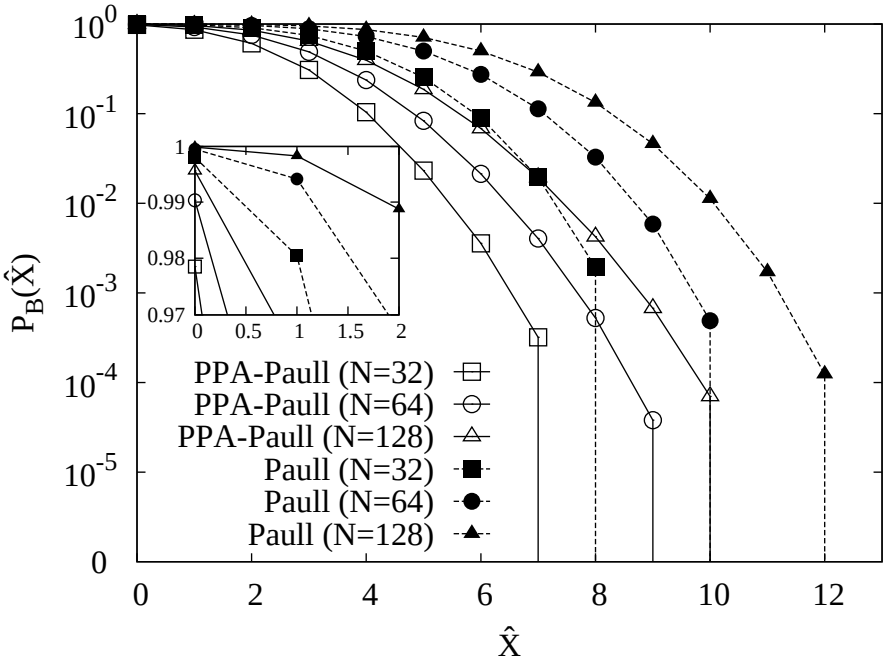


Figure 2.12. Blocking probability in Benes networks with $\rho = 0.5$

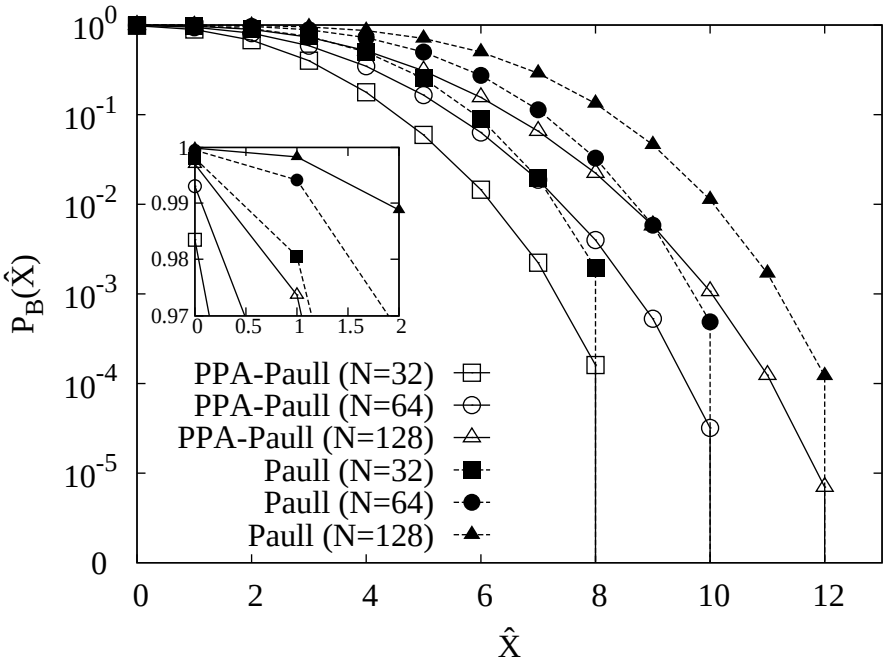


Figure 2.13. Blocking probability in Benes networks with $\rho = 0.9$

(consistently with the values in the figure).

Now observe that, in a Benes network, there exist always $2^{\log_2 N-1} = N/2$ different paths connecting any input to any output, since in the first $\log_2 N-1$ stages there are always two output ports in each module that can be used to reach any destination, whereas in the last $\log_2 N$ stages there exists just one output port in each module to reach the desired output. Given an input-output pair with a possible path compatible with $\hat{X} = 0$ (this pair is chosen with probability $1/N$ as shown above), the Paull algorithm will choose one random path among the $N/2$ available paths, but only one of them will be able to satisfy the constraint $\hat{X} = 0$. Hence, we can expect that $P_B^{\text{Paull}}(0) = 1 - 2/N^2$ (coherently with the values in the figure), which is larger than $P_B^{\text{PPA-Paull}}(0)$.

We now evaluate the maximum degradation index experienced by a single path computed by PPA-Paull. At each factorization level, PPA-Paull chooses the configuration of the first and the third stage to minimize the degradation index; in the case of a single path, the degradation index can increase by one at each factorization level. As a consequence, the maximum degradation index will be $\log_2 N - 1$, equal to the number of factorization levels. Hence, for low load, we expect that the blocking probability tends to zero when $\hat{X} \geq \log_2 N$. Indeed, Fig. 2.11 shows that the observed blocking probability goes to zero for $\hat{X} = 6,7,8$ when $N = 32,64,128$, which is very close to the bound found before.

2.8 Conclusions

In this chapter we analyzed the scalability in terms of area and power penalty of different interconnection networks based on microring resonators. We described the basic 1×2 and 2×2 switching elements, we highlighted the asymmetric power loss of the different switching states, and we defined the degradation index. Then, we obtained the degradation index of crossbar, Benes and Clos networks. To achieve a better compromise between area (in terms of number of microrings) and degradation index, we propose (i) two architectures based on different combinations of the Benes and crossbar networks and (ii) the mirroring technique. Finally, we presented a simple variation of the classical Paull algorithm to setup input-output connections, and we investigated the corresponding improvement in terms of the degradation index. Given the promising results obtained in our studies, we believe that the role of microring resonators will become more and more relevant in future high capacity photonic interconnection networks.

Chapter 3

Microring-based wavelength routing matrices

This chapter explores the use of optical microring resonators to build wavelength-routing-based optical switching fabrics. Similarly to the case of other devices used to build all-optical switching fabrics, such as arrayed waveguide gratings, microring-based switching fabrics might show a limited scalability in terms of port count due to the crosstalk accumulation caused by the spatial reuse of wavelengths.

The chapter begins with an introduction in Sec. 3.1 of optical switching fabrics, the usage of microring resonators and the motivation of this work. Then, two different approaches are used to face the crosstalk accumulation issue in microring-based wavelength routing matrices (WRMs).

The first approach, in Sec. 3.2, considers a *passive* WRM based in a static microring resonator routing elements. It is carried out a crosstalk analysis proposing two different solutions that exploit the periodicity of the transfer function of the microring. In addition, a feasibility analysis is conducted taking into account the geometrical limitations of microring resonators. Finally, scalability results are presented and some concluding remarks of this first approach are given.

The second approach, in Sec. 3.3, introduces an *active* WRM based in a dynamic microring resonator switching elements. It is described the considered scenario and the formulation of the wavelength reuse problem. Then it is presented the reconfigurable microring-based WRM to implement the optical switching fabric. Two different strategies to exploit the active wavelength routing matrix are presented and compared in terms of the offered level of wavelength reuse. The chapter ends with some conclusions for this second approach.

3.1 Introduction and motivation

The need to process and switch the continuously-increasing Internet bandwidth demands has driven the evolution of electronic switching technologies, pushing them almost to their physical limits. Although electronic routers with capacities above several Tb/s are commercially available, their packaging in a single rack of equipment is increasingly difficult given the number of backplane interconnections, the required power density and power dissipation, as well as the several electromagnetic compatibility issues they must face.

Today, high-end routers often comprise several racks: one or more racks host the electronic switching fabric and the control logic, while other racks host the line cards. In this scenario, optical technologies have found their way in the switching domain, since optical links are increasingly used to interconnect the fabric with the line cards. In addition, recent breakthroughs in CMOS-compatible silicon photonic integration are enabling the penetration of optical technologies for interconnections between chips on boards and up to the on-chip scale [29]. Several solutions in the literature propose to push optics beyond point-to-point transmission and interconnection, for instance performing switching operations directly in the optical domain [30], to ease the strict technological requirements and power consumption of today electronic routers [31].

These trend arise due to the interesting properties of photonic technologies: ability to transport huge data rates over larger distances, a switching complexity which is almost independent of the bit rate, good scalability to high-capacities, and lower power requirements. The usual scenarios to exploit the benefits of photonic technologies capitalize on Wavelength Division Multiplexing (WDM), and often are based upon a set of Tunable Transmitters (TTxs) sending data through a wavelength-agile optical switching fabric to reach a set of wide-band burst-mode receivers. In this work, we study the implementation of the optical switching fabric by using Wavelength Routing Matrices (WRMs). Despite the existence of several proposals using different devices, mainly based on Arrayed Waveguide Gratings (AWGs), our work is motivated by two main observations.

On the one hand, AWGs are mature and commercially available WRMs, but they show a limited scalability in terms of port count due to the crosstalk accumulation caused by the reuse of the same wavelength by several TTxs [32,33]. This drawback can be mitigated either $i)$ by employing a proper scheduling algorithm, which requires quite a large complexity [34, 35], or $ii)$ by using a wider tuning range at the TTxs, exploiting the periodicity of the AWGs' Transfer Function (TF). This resource exploitation is limited because both the number of wavelengths a TTx can tune to is limited and the AWG's

TF loses uniformity after a small number of periods.

On the other hand, silicon *microring resonators* (MRRs) are promising devices which can be used to realize all-optical reconfigurable switching fabrics [2]. MRRs are waveguide-based and wavelength-selective elements whose filtering function is defined by their geometrical properties. An additional and interesting property which makes MRRs appealing for building switching elements is that their filtering behavior can be modified by changing their waveguide refractive index by means of simple physical effects [20,21,23]. As such, these wavelength-selective properties make the MRRs very suitable for wavelength routing operations. However, in the work [36] it is observed that in MRR-based architectures the main factor limiting scalability is still the crosstalk accumulation due to the wavelength reuse. Hence, in this work, we explore a novel MRR-based crossbar architecture to mitigate this limitation.

The remainder of this chapter is divided in two sections. The major contributions in Sec. 3.2 are: (i) the introduction of a MRR-based passive wavelength routing matrix (WRM) logically equivalent to an array waveguide grating (AWG); (ii) an analysis of its accumulated coherent crosstalk and the proposal of two solutions exploiting the periodicity of the MRR TF to mitigate the ACC; and (iii) considering geometrical limitations we propose a new MRR design strategy that employs feasible MRRs to build WRMs supporting a large number of ports. Subsequently, the major contributions in Sec. 3.3 are: (i) the introduction of a MRR-based WRM that employs the switching capabilities of MRRs and (ii) the proposal of two different strategies that exploit the versatility of the newly proposed WRM to mitigate crosstalk limitation, the definition of the proper control algorithms and their performance analysis.

3.2 Passive microring-based WRM

3.2.1 Microring resonator basic routing element

Fig. 3.1 shows the microring-based 1x2 routing element (1x2-RE) used in this first approach. The 1x2-RE is the basic building block of our wavelength routing matrix and it consists of a waveguide bent into itself and side-coupled to two perpendicular waveguides. Input optical signals are coupled into the ring and, thus, to the drop port, if their frequencies match the resonance frequencies f_x of the microring resonator, given by Eq. (3.1) [37], where m is an integer number, R is the MRR radius, n_g is the waveguide group and c is the speed of light in vacuum. For instance, in Fig. 3.1, $f_x = \{f_0, f_4\}$. Conversely, if the signals' frequencies are different from f_x, they continue, almost unaffected, to the through port.

$$f_x = m\frac{c}{2\pi R n_g} \quad (3.1) \qquad FSR = \frac{c}{2\pi n_g R} \quad (3.2)$$

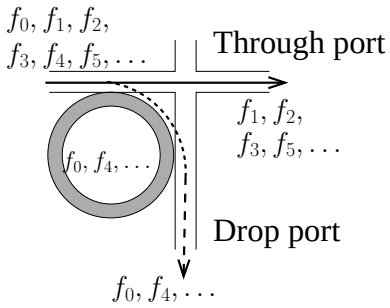


Figure 3.1. 1×2 static microring-based routing element (RE)

Microring resonators, like other resonant structures, present a periodic transfer function whose period is named Free Spectral Range (FSR) and depends on the geometry of the microring, as shown in Eq. (3.2) [37]. In addition, being a filtering device, microrings are characterized by its passband (BW), commonly called *full width half maximum* bandwidth and defined as the bandwidth where input signals suffer at most a 3dB power drop with respect to the maximum of the transfer function. The BW depends on the coupling coefficients between the ring and the other waveguides and it can be set as a design target [37].

3.2.2 Passive microring-based WRM

We propose a wavelength routing matrix logically equivalent to an array waveguide grating (AWG). Without loss of generality, we assume a $N \times N$ wavelength routing matrix with synchronous and time-slotted operation, controlled by a proper scheduling algorithm. Each input port is equipped with a tunable transmitter (TTx) and each output port with a waveband burst-mode receiver (WBMR), as depicted in Fig. 3.2. The 1x2-REs composing the wavelength routing matrix are devised such that the 1x2-RE in position (i, j) resonates on $f_k + mFSR$, being m a positive integer equal for all the microrings and f_k appearing at most once in each row and each column. As in typical AWGs and with no loss of generality, we assume that f_k is selected according to $k = (j - i) \mod N$. Finally, let w be the *fundamental* TTx tuning range, i.e., the *minimum* bandwidth through which TTxs must be able to sweep to provide full connectivity. Since each TTx must be able to tune over at least N distinct channels (at least one for each destination),

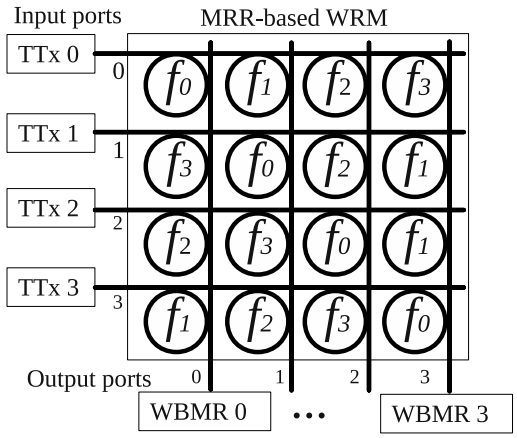


Figure 3.2. 4×4 microring-based passive wavelength routing matrix (WRM)

$w = N\Delta_{ch}$ Hz. Finally, given that f_k cannot be repeated for each column and row, $FSR \geq N\Delta_{ch}$.

3.2.3 Crosstalk analysis

Let us denote by X the maximum number of times an optical channel is used considering a complete input-output permutation (all inputs transmitting to a different output). X is a simplified measure of the accumulation of coherent crosstalk (ACC). In this work we do not present a detailed physical-layer analysis, but the methodology presented in [18] can be easily exploited. However, results presented in [18] showed that the scalability of MRR-based fabrics is mainly limited by the ACC. Thus, X provides a good insight on the scalability of the presented WRM.

X can be reduced exploiting the periodicity of the microring's transfer function through the *expansion* strategy which consists in increasing the TTx tuning range. Note that a TTx could tune over a bandwidth larger than w, defined as the expanded TTX tuning range $W \geq w$. Define the TTx expansion factor $K = \{1, 2, \ldots, N\}$, such that $W = Kw = KN\Delta_{ch}$ Hz. As K increases, the number of times an optical channel is reused decreases as $X = \left\lceil \frac{N}{K} \right\rceil$. Thus, if $K = 2$, $X = \lceil \frac{N}{2} \rceil$ because TTxs can use 2 channels to reach each output.

On the other hand, since the tuning range might become a limiting factor if it grows too high, we further propose the *juxtaposition* strategy, which consists in using TTxs which can tune over different and disjoint tuning ranges of size w. Let L be the number of different and disjoint fundamental tuning ranges over which TTxs can tune. If the TTxs are uniformly divided into L sets, the ACC decreases as $X = \left\lceil \frac{N}{L} \right\rceil$. Note that, the two strategies can

be combined together to reduce the ACC to $X = \lceil \frac{N}{LK} \rceil$ and that they do not influence the design of the microring's transfer function, which still requires a $FSR \geq N\Delta_{ch}$.

3.2.4 Feasibility Analysis

Depending on the filtering quality of the microrings's transfer function, the proposed wavelength routing matrix can tolerate different values of X. Indeed, in [18] we showed that a 50×50 crossbars is feasible using a single optical channel ($X = 50$). Furthermore, the above proposed strategies, largely reducing ACC issues, can enhance the physical scalability. In this context, the main limitations arise when considering the MRR geometry. Indeed, MRRs used in the proposed wavelength routing matrix are required to support a $FSR \geq N\Delta_{ch}$ but, as Eq. (3.2) shows, $FSR \propto 1/R$. As an example, if we consider a wavelength routing matrix with $N = 25$ ports and a channel spacing of $\Delta_{ch} = 100$ GHz on the DWDM ITU grid, then $FSR \geq 2.5$ THz, which approximatively corresponds to a radius $R \leq 12.7$ μm, considering $n_g = 1.5$ (typical of Silicon waveguides) and $f_x = 195.9$ THz (the first frequency of the conventional C band of the ITU grid). Note that a radius of few μm is close to technological limits [38]. Indeed, radii of hundreds of μm are more common, as a small radius implies high losses in traversing the ring waveguide because of the poor confinement of optical signals [37]. As such, the minimum feasible MRR radius R_{min} implies a maximum achievable FSR, FSR^{max}, and a maximum of $N^{\text{max}} < FSR^{\text{max}}/\Delta_{ch}$ ports.

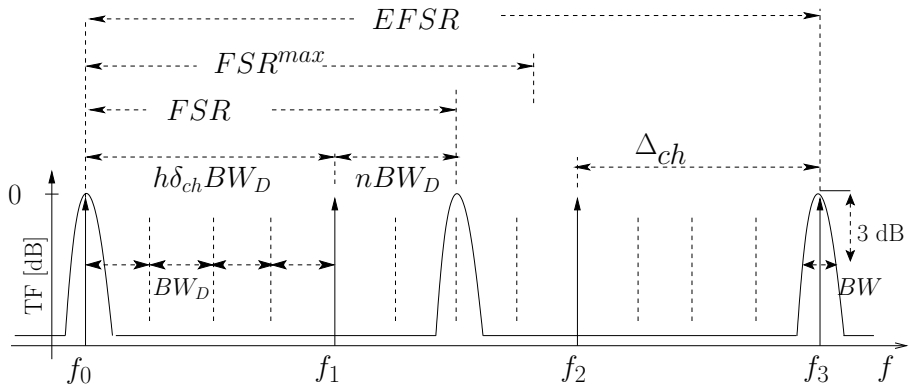


Figure 3.3. Transfer function (TF) at the drop port of the 1×2-RE

Hence, when the limitations on the MRR size are considered, the maximum number of ports that the proposed wavelength routing matrix supports is around ten. However, this limitation can be mitigated by carefully

designing the MRR TF. For instance, let us assume that $FSR = 2\Delta_{ch}$. Thus, one channel among the two available is dropped within the MRR's TF period, and the maximum achievable number of ports is $N^{\max} = 2$. Instead, if we design MRRs to present a $FSR = 3/2\Delta_{ch}$ (as in Fig. 3.3), the peaks of the MRR TF coincide with the channels f_0 and f_3. Therefore, $N^{\max}$ becomes 3. In this case we say that the MRR offers an equivalent FSR, $EFSR = 3\Delta_{ch}$.

Let BW_D be the maximum MRR bandwidth able to divide Δ_{ch} a finite number of times δ_{ch}, such that $\delta_{ch} = \lfloor \frac{\Delta_{ch}}{BW} \rfloor = \frac{\Delta_{ch}}{BW_D}$. Clearly $BW_D \geq BW$. Similarly, let $fsr = \lfloor \frac{FSR}{BW_D} \rfloor$ be the discretized MRR FSR, let $fsr^{(\max)} = \lfloor \frac{FSR^{(\max)}}{BW_D} \rfloor$ be the discretized $FSR^{(\max)}$ and $efsr = \lfloor \frac{EFSR}{BW_D} \rfloor$ be the discretized $EFSR$. Since fsr and $efsr$ can be expressed as multiples of δ_{ch}, i.e., $fsr = h\delta_{ch} + n$ and $efsr = M\, fsr$, the following equations must hold:

$$fsr = n \mod \delta_{ch} \tag{3.3}$$

$$efsr = Mfsr = Mn \mod \delta_{ch} \tag{3.4}$$

Eq. (3.3) and Eq. (3.4) derive from the definition of n and of $efsr$. Our objective is to find a triple $(h^{\max}, n^{\max}, M^{\max})$ maximizing $efsr$. Since $M^{\max}$ represents the maximum number of times the fsr can be repeated before a channel matches again the MRR's TF, then $nM^{\max} = 0 \mod \delta_{ch}$. It is easy to prove that the maximum $efsr$ is obtained if $M^{\max} = \delta_{ch}$, $h^{\max} = \lfloor \frac{FSR^{\max}}{BW_D} \rfloor$, $n^{\max} = \max_n \{fsr^{\max} \geq h^{\max}\delta_{ch} + n, \text{ s.t } \text{gdc}(n, \delta_{ch}) = 1\}$, where $\text{gdc}(x, y)$ is the greatest common divisor between x and y.

Table 3.1. Aggregate capacity in Gbps for different MRR $R_{\min}$ and Δ_{ch} using $EFSR$. Values in parenthesis report the gain with respect to a classic wavelength routing matrix.

$Rb = 10$ Gbps	Δ_{ch} [GHz]		
$R_{\max}$ $[\mu m]$	200	100	50
5	3180 (9.94)	3180 (4.97)	2530 (1.99)
10	1590 (9.94)	1590 (4.97)	1270 (1.98)
100	150 (7.5)	140 (4.67)	110 (1.83)
$Rb = 40$ Gbps			
5	2520 (1.97)	2520 (1)	0 (0)
10	1240 (1.94)	1240 (1)	0 (0)
100	120 (1.5)	120 (1)	0 (0)

Table 3.2. TTx tuning range w, in THz, needed to support aggregated capacity reported in Table 3.1 for a target ACC of $X = 50$. In parenthesis the expansion and juxtaposition factors L and K (being L and K numerically equal, only one value is reported).

$Rb = 10$ Gbps	Δ_{ch} [GHz]		
R_{max} $[\mu m]$	200	100	50
5	-	31.8 (7)	12.65 (6)
10	31.8 (4)	15.9 (4)	6.35 (3)
100	3 (1)	1.4 (1)	0.55 (1)
$Rb = 40$ Gbps			
5	12.6 (2)	6.3 (2)	0 (-)
10	6.2 (1)	3.1 (1)	0 (-)
100	0.6 (1)	0.3 (1)	0 (-)

3.2.5 Results

According to recent literature works, we considered MRRs with a minimum radius equal to $R_{\text{min}} = \{5, 10, 100\}$ μm, based on silicon waveguides ($n_g \approx 1.5$) and a BW equal to the minimum required for filtering the signal at the given bit rate. If TTxs perform basic NRZ modulation at a bitrate R_b, the effective optical filtering bandwidth is $BW = 2R_b$. In addition, we considered channel spacings in $\Delta_{ch} = \{50, 100, 200\}$ GHz, compliant to the ITU-T G.694.1 specifications. Table 3.1 shows the maximum aggregate capacity for several R_b and Δ_{ch}. Smaller MMR radii ensure larger capacities due to the larger FSR^{max}. The gain achieved by the EFSR extension technique is indicated in parenthesis and it is evaluated as the ratio between the capacity of wavelength routing matrix exploiting the $EFSR$ technique and a simple MMR-based wavelength routing matrix. The capacity gain is higher for higher channel spacings because it allows a larger δ_{ch}. A sensible gain in capacity is achieved with respect to the version without EFSR technique, i.e., it is possible to achieve several Tbps if the technology employed allows the construction of MMRs with 5μm radius. Furthermore, the use of smaller bitrates is related to a slightly larger scalability, due to the larger granularity in the use of the spectrum. Table 3.2 shows the final TTx tuning range w in THz needed to support the aggregated capacities reported in Table 3.1, when a maximum ACC $X = 50$ is allowed [18]. Dashes are placed where the resulting w was higher than 60 THz, a value today unreachable by TTx. In parenthesis, we reported the expansion and the juxtaposition factors K and L,

which are set to the same value. Both strategies mitigate the ACC, but they have different impact on the TTx side. For instance, for $R_b = 40$ Gbps and microrings with 5μm of radius, the ACC constraints ($X = 50$) can be satisfied either by the expansion strategy, employing TTxs with an extended tuning range of $W = 2 \times 6.3$ THz, or by the juxtaposition strategy, using two different sets of TTxs with $w = 6.3$ THz each.

3.2.6 Conclusions

We proposed a microring-based wavelength routing matrix for packet switching applications that is logically equivalent to an AWG. We analyzed its physical limitations related to ACC, and feasibility limitations due to the technological constraints on the minimum microring radius. We presented two solutions to limit the ACC that exploit the periodicity of the transfer function of the microring: the *expansion* strategy and the *juxtaposition* strategy. Finally, we introduced the concept of equivalent FSR which permits to overcome technological limitations related to the microring geometry. Hence, the proposed wavelength routing matrix is promising in terms of scalability, and is a good candidate for future optical switching fabrics.

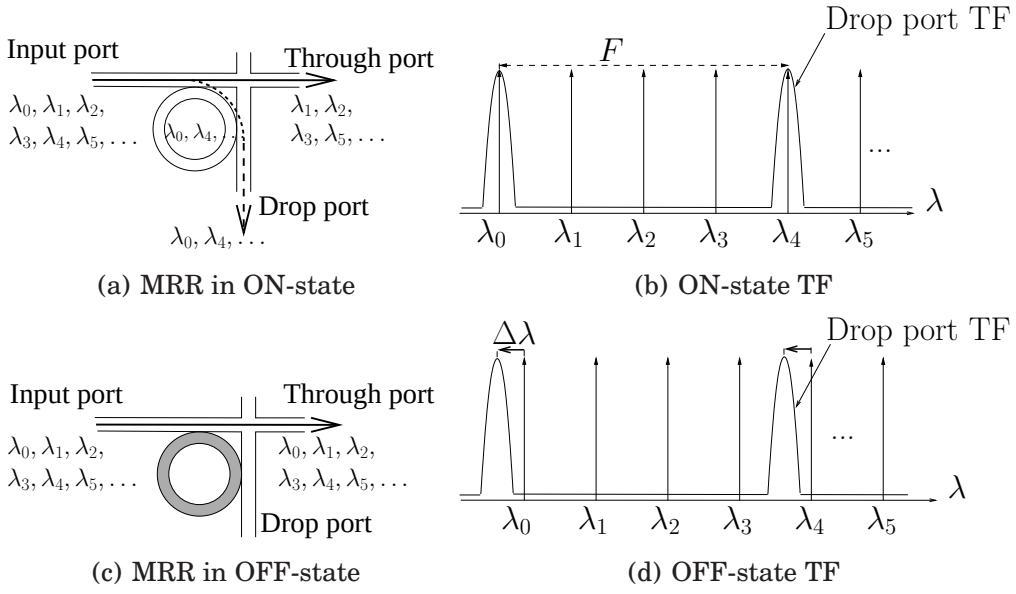
3.3 Active microring-based WRM

3.3.1 Microring resonator basic switching element

The MRR is the component used to perform filtering and switching operations. Fig. 3.4(a) shows that it mainly consists on a waveguide bent into itself and side-coupled to two perpendicular waveguides. The MRR exhibits a periodical Transfer Function (TF) of period F, usually referred to as Free Spectral Range (FSR): the optical signal coming from the input port is coupled to the ring and deflected to the drop port if its wavelength matches the MRR's resonance wavelengths $\lambda_k = \lambda_{0+kF}$, with $k \in Z$. Conversely, if the wavelength of the incoming signal is different from the MRR's λ_k, the signal continues straight to the through port. Note that the resonance wavelengths are determined by the geometrical and physical properties of the MRR [37]. Fig. 3.4(a) shows an example for a MRR routing element when $\{\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, ...\}$ are the incoming signals within a wavelength comb. The TF is depicted in Fig. 3.4(b): Since two peaks of the MRR drop TF match λ_0 and λ_4 ($F = 4$ wavelength channels), the optical signals at these wavelengths are deflected to the drop port. In this case, we say that the MRR is in the *ON-state*.

Another interesting property which actually differentiates MRRs from other devices as AWGs is that the MRRs' TF can be *shifted* of $\Delta\lambda$ by changing the refractive index of the waveguide composing the microring. This change of the refractive index can be achieved through thermo-optic [20], optical pump [23] or carrier injection [21] effects, each one ensuring different tuning times. In the remainder of the chapter, we consider MRRs controlled by carrier injection techniques because they ensure short switching times of few hundred ps [21]. More precisely, we exploit the possibility of shifting the MRR's TF to *detune* the MRR from its resonance wavelengths λ_k. When a shifting of $\Delta\lambda$ is applied we say that the MRR is configured in the *OFF-state*, in which signals are not deflected to the drop port, and they continue unaffected to the through port (see Fig. 3.4(c) and Fig. 3.4(d)).

Although a detailed physical analysis of MRR devices and MRR-based switching fabric is out of the scope of this work, we assume a non ideal MRR's TF causing a residual optical power leakage at the non-desired output. For instance, considering Fig. 3.4(a) and 3.4(b), we assume residual components of $\{\lambda_1, \lambda_2, \lambda_3, \lambda_5, \ldots\}$ at the drop port and of $\{\lambda_0, \lambda_4, ...\}$ at the through port. This non ideal MRR behaviour causes the presence of undesired signals interfering with useful signals, generating crosstalk.



(a) MRR in ON-state

(b) ON-state TF

(c) MRR in OFF-state

(d) OFF-state TF

Figure 3.4. 1×2 dynamic microring-based switching element (SE) and TFs

Normally MRR-based wavelength routing matrices (WRMs) operate on a single wavelength, leading to large coherent crosstalk values. In this work we exploit both wavelength agility of light sources and MRR tunability to reduce coherent crosstalk contributions.

3.3.2 Problem formulation

Fig. 3.5 depicts a classical $N \times N$ MRR-based architecture [36, 39]: It consists on a set of N TTxs that transfer data packets towards a set of N Wideband Burst Mode Receivers (WBMRs) through a MRR-based switching fabric which should be able to dynamically provide any input-output interconnection. We consider the availability of the set of N wavelengths $\Lambda = \{\lambda_0, \lambda_1, \ldots, \lambda_{N-1}\}$ at the TTxs.

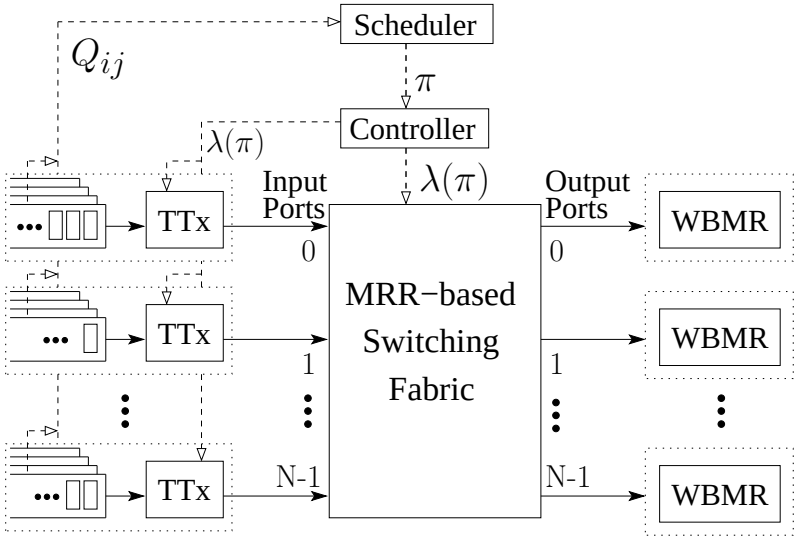


Figure 3.5. Scenario for the proposed architecture

We focus on a synchronous time-slotted architecture running a scheduling algorithm which ensures that, at each time slot, at most one packet is sent to each WBMR from any TTx, thus matching each input to each output. At each time slot, each TTx sends the information about its queue occupancy Q_{ij}, with $0 \leq i, j \leq N-1$, to the *scheduler*. The scheduler takes a scheduling decision choosing a permutation of port indexes to define packet transfer. We denote this permutation by a vector $\pi = [\pi[0], \pi[1], \ldots, \pi[N-1]]$, where $\pi[i]$ is the output port index to which input i transmits during the current time slot. Clearly, each output port index can appear at most once in π. The *controller* takes π as an input and performs the wavelength assignment phase, computing $\lambda(\pi)$, a vector of the wavelengths which must be used by TTxs to satisfy permutation π. Therefore, each input i must tune its TTx on $\lambda(\pi)[i]$ to reach output $\pi[i]$. On the other hand, given the possibility of tuning MRRs, the controller has also the option of configuring the state of the MRRs as a function of π and $\lambda(\pi)$. In other words, a MRR-based WRM operated according

to wavelength routing principles, can be controlled both with "external" (i.e. of TTxs) tunability, and with "internal" (i.e. of MRRs) tunability.

Our objective is to design a wavelength-routed MMR-based WRM and a controller which, given an input permutation π, computes the wavelength assignment $\lambda(\pi)$ characterized by the smallest number of wavelength reuse, defined as

$$C(\pi) = \max_i \sum_{k=0}^{N-1} \mathbb{1}_{\{\lambda(\pi)[k]=\lambda_i\}} \tag{3.5}$$

where $\mathbb{1}$ denotes the indicator function. $C(\pi)$ is the maximum number of inputs using the same wavelength in permutation π. Note that $1 \leq C(\pi) \leq N$, and $C(\pi) = N \ \forall \pi$ when the MRR-based WRM is operated on a single wavelength using internal tunability only. Similarly, when only external tunability is exploited and MRRs are fixed, $\exists \pi : C(\pi) = N$ (see Example 1 later). Since we aim at minimizing $C(\pi)$ given any permutation π to reduce crosstalk impact, we define the worst case wavelength reuse:

$$C = \max_{\pi} C(\pi) \tag{3.6}$$

3.3.3 Active microring-based WRM

The considered MRR-based switching fabric consists of a grid of $N \times N$ perpendicular waveguides forming a crossbar. The horizontal waveguides are connected to the inputs and the vertical waveguides are connected to the outputs. A MRR is placed at each crosspoint leading to a total number of N^2 MRRs as in Fig. 3.6.

MRR-based WRM can be designed to mimic the functionality of AWGs, passive interferometers that provide at each input a different wavelength to connect to each output. Given an input port i and an output port $\pi[i]$, the MRR of the corresponding crosspoint resonates at wavelength

$$\lambda_A(\pi)[i] = \lambda_a, \text{ where } a \triangleq (i - \pi[i]) \bmod N.^{1} \tag{3.7}$$

We call *wavelength assignment A* the assignment defined by (3.7). The MRR-based WRM using wavelength assignment A for $N = 5$ is shown in Fig. 3.6.

Remark 1. i) To effectively mimic the AWG functionality, the output is univocally determined by the input and the resonance wavelength, given that

[1]This means that λ_0 is used to connect input i to output i, λ_1 to connect input i to output $(i-1) \mod N$, λ_2 to connect input i to output $(i-2) \mod N$, etc.

no wavelengths are repeated in a row or in a column. In other words, let π and π' be two permutations such that $\pi[i] \neq \pi'[i]$. Then, $\lambda_A(\pi)[i] \neq \lambda_A(\pi')[i]$. Indeed, if we assume by contradiction that $\lambda_A(\pi)[i] = \lambda_A(\pi')[i]$, we get $(i - \pi[i]) = (i - \pi'[i]) \mod N$, and then $\pi[i] = \pi'[i] \mod N$, which is a contradiction.

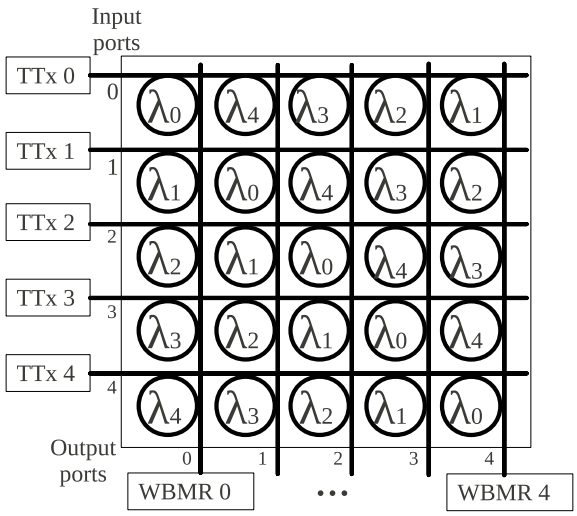


Figure 3.6. 5×5 microring-based fixed WRM

We denote by $C_A(\pi)$ the wavelength reuse when the assignment A is used to compute $\lambda(\pi)$. The following example shows that MRR-based routing matrices exploiting assignment A show a worst-case wavelength reuse equal to N as in AWGs.

Example 1. Consider the wavelength assignment A described in (3.7) and $N = 5$. Let π be a permutation such that, for each input i, $\pi[i] = (i + 2) \mod 5$. In this case, all the packets are sent using wavelength $\lambda_A(\pi)[i] = \lambda_a$ with $a = i - i - 2 = 3 \mod 5$ for all i. Hence, $C_A(\pi) = N$.

Motivated by the situation described in Example 1, our first contribution is to define a novel MRR-based WRM in which the TF of each MRR is designed assuming that each MRR can tune on an additional wavelength. This additional wavelength is selected according to *assignment B*:

$$\lambda_B(\pi)[i] = \lambda_b, \text{ where } b \triangleq (i + \pi[i]) \bmod N. \tag{3.8}$$

We denote by $C_B(\pi)$ the wavelength reuse when the assignment B is used to compute $\lambda(\pi)$. Fig. 3.7 depicts our proposed WRM for $N = 5$.

The following example shows that when both assignments A and B are available, the wavelength reuse can be reduced for some permutation π.

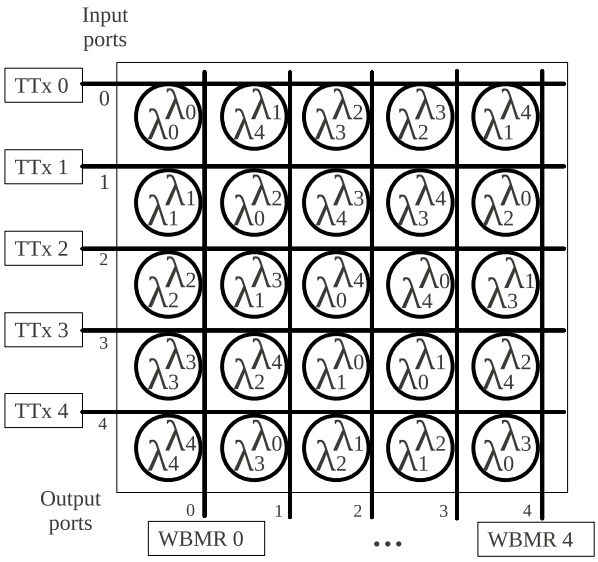


Figure 3.7. Proposed 5×5 microring-based active WRM

Example 2. Consider the wavelength assignment B described in (3.8) and $N = 5$. Let π be the same permutation of Example 1, i.e., $\pi[i] = (i+2) \mod 5 \forall i$. Packets can be sent using wavelength $\lambda_B(\pi)[i] = i + i + 2 = (2i + 2) \mod 5$. Note that given two different inputs i and i', $2i + 2 \neq (2i' + 2) \mod 5$. Thus $C_B(\pi) = 1$.

Given the above hints, we propose to define a WRM where each MRR can tune to the two above defined wavelengths and to dynamically select the best assignment between A or B, or a combination of them, depending on π.

In Sec. 3.3.4 we prove analytically that the property described in Ex. 2 holds in general, and that the additional wavelength assignment actually permits to reduce C from N to about $N/2$.

From the logical point of view, our proposal can be seen as a superposition of two WRMs characterized by a different wavelength assignment rule. However, instead of doubling the cost of the architecture by physically adding another WRM of N^2 MRRs, we exploit the tunability of MRRs in the ON/OFF states as follows:

- Consider MRRs resonating at the same wavelength in both assignments A and B. Due to Remark 1, easily extensible to assignment B, no other MRR in the same row or column resonates at that wavelength. Therefore, the microring can be fixed in its ON state moving that wavelength to the drop port.

- Focus now on MRRs with different resonance wavelengths in assignments A and B. Due to Remark 1 (i) it is required that the WRM does not repeat wavelengths dropped by the MRRs within the same row or column. Therefore, MRRs are set by default in the OFF state to let all wavelengths pass unaffected, and they are only switched to the ON state to satisfy the required permutation (dropping both λ_A and λ_B).

Note that the decomposition in two AWG switching stages proposed in [35] to limit the wavelength reuse to 4 is equivalent to duplicating complexity (or to providing a $2\times$ speedup in the space dimension). In this work, the complexity relies on the MRR design: MRR must resonate at the two wavelengths corresponding to assignments A and B. This leads to different MRRs depending on the position in the crossbar structure.

In the particular case shown in Fig. 3.7, the first column contains MRRs resonating at the same wavelength in both the assignments so they are fixed in the ON state. In the other columns, they are switched from OFF to ON to satisfy a given permutation. Notice that in Example 2, MRRs at $(i,j) =$ $(0,2)$;$(1,3)$; $(2,4)$; and $(4,1)$ need to switch from OFF to ON state to drop λ_3 (assignment A) and λ_2, λ_4, λ_1 and λ_0 respectively (assignment B), while the MRR at $(3,0)$ is fixed to λ_3.

3.3.4 Wavelength reuse reduction techniques

In this section, we analyze two different strategies to exploit the properties of the proposed WRM.

Matrix Selection (MS) strategy

The first approach considered to exploit both wavelength assignments A and B defined in Sec. 3.3.3 is to choose $\lambda(\pi)$ using the wavelength assignment that minimizes C. We call this approach Matrix Selection strategy.

Strategy 1. Matrix Selection. Consider a permutation π. The MS strategy makes the following choice:

$$\lambda(\pi) = \begin{cases} \lambda_A(\pi) & \text{if } C_A(\pi) \leq C_B(\pi) \\ \lambda_B(\pi) & \text{otherwise} \end{cases}$$

To prove the main Theorem, we start with two preliminary lemmas.

Lemma 1. Consider N odd, a permutation π and the wavelength assignments $\lambda_A(\pi), \lambda_B(\pi)$ defined respectively in (3.7) and (3.8). Suppose $\lambda_A(\pi)[i_1] = \lambda_A(\pi)[i_2]$ for $i_1, i_2 \in \{0, \ldots, N-1\}$ with $i_1 \neq i_2$. Then, $\lambda_B(\pi)[i_1] \neq \lambda_B(\pi)[i_2]$.

Proof. We prove the lemma by contradiction. Suppose that $\lambda_A(\pi)[i_1] = \lambda_A(\pi)[i_2]$ and $\lambda_B(\pi)[i_1] = \lambda_B(\pi)[i_2]$. Then, $i_1 - \pi[i_1] = i_2 - \pi[i_2] \bmod N$ and $i_1 + \pi[i_1] = i_2 + \pi[i_2] \bmod N$. If we sum each side of the two equations, we get $2i_1 = 2i_2 \bmod N$. By the properties of congruence mod N, this corresponds to $i_1 = i_2 \bmod N/(2, N)$, where $(2, N)$ is the greatest common divisor. Since N is odd, $(2, N) = 1$, and the equation is equivalent to $i_1 = i_2 \bmod N$, which is a contradiction. □

A result similar to Lemma 1 can be obtained for the case N even, as follows.

Lemma 2. Consider N even, a permutation π and the wavelength assignments $\lambda_A(\pi), \lambda_B(\pi)$ defined respectively in (3.7) and (3.8). Suppose $\lambda_A(\pi)[i_1] = \lambda_A(\pi)[i_2] = \lambda_A(\pi)[i_3]$ for $i_1, i_2, i_3 \in \{0, \dots, N-1\}$ with $i_1 \neq i_2 \neq i_3$. Then, $\lambda_B(\pi)[i_1] \neq \lambda_B(\pi)[i_2]$ or $\lambda_B(\pi)[i_2] \neq \lambda_B(\pi)[i_3]$.

Proof. Following the same reasonings of the proof of Lemma 1, we get the following equations:

$$\begin{cases} 2i_1 = 2i_2 \bmod N \\ 2i_2 = 2i_3 \bmod N \end{cases} \tag{3.9}$$

The first equation in (3.9) implies $i_1 = i_2 \bmod N/(N,2) = N/2$. Since $i_1, i_2 \in \{0, \dots, N-1\}$, if $i_1 = i_2 \mod N/2$ and $i_1 \leq N/2$, then $i_2 > N/2$, otherwise we would have $i_1 = i_2 \bmod N$, a contradiction. Suppose $i_1 \leq N/2$ and thus $i_2 > N/2$. For the same reasons, from the second equation in (3.9), if $i_2 > N/2$, then $i_3 \leq N/2$. From (3.9) we can also get $2i_1 = 2i_3$, that is $i_1 = i_3 \bmod N/2$: if $i_1 \leq N/2$ then $i_3 > N/2$, which is a contradiction. A similar contradiction is found if we assume $i_1 > N/2$. □

Theorem 1. The maximum wavelength reuse C using the MS strategy is

$$C = \lfloor N/2 \rfloor + 1.$$

Proof. Let $\tilde{\pi}$ be a permutation such that $C(\tilde{\pi}) = C$. We first prove the theorem for the case N odd. If $N = 1$ the thesis obviously holds. Thus, consider $N \geq 3$. Suppose without loss of generality that $\lambda = \lambda_A(\tilde{\pi})[0] = \lambda_A(\tilde{\pi})[1] = \dots = \lambda_A(\tilde{\pi})[C_A(\tilde{\pi}) - 1]$, while $\lambda_A(\tilde{\pi})[k] \neq \lambda$ for all $k \geq C_A(\tilde{\pi})$. From Lemma 1 we know that, necessarily, $\lambda_B(\tilde{\pi})[k] \neq \lambda_B(\tilde{\pi})[h]$ for all $k \neq h$, $0 \leq k, h < C_A(\tilde{\pi})$. The other $N - C_A(\tilde{\pi})$ positions $\lambda_B(\tilde{\pi})[k]$, $k \geq C_A(\tilde{\pi})$ are not constrained: in the worst case they are all equal to a wavelength λ', and there exists at most an index $k \leq C_A(\tilde{\pi})$ such that $\lambda_B(\tilde{\pi})[i_k] = \lambda'$. In this case $C_B(\tilde{\pi}) = N - C_A(\tilde{\pi}) + 1$. Therefore, we obtain:

$$\begin{cases} C(\tilde{\pi}) = \min\{C_A(\tilde{\pi}), N - C_A(\tilde{\pi}) + 1\} \\ C(\tilde{\pi}) = \max_{\pi} C(\pi) \end{cases} \tag{3.10}$$

with $1 \leq C_A(\tilde{\pi}) \leq N$. It is easy to verify that the solution $C_A(\tilde{\pi})$ of (3.10) is found when the two quantities argument of the minimum function are equal, that is when $C_A(\tilde{\pi}) = (N+1)/2$. The maximum wavelength reuse experienced using the MS strategy for N odd is $C = (N+1)/2 = \lfloor N/2 \rfloor + 1$.

We now prove the theorem for the case N even. If $N = 2$ the thesis obviously holds. Thus, consider $N \geq 4$. It is easy to see that $C_A(\tilde{\pi}) \geq 3$: if we assume that $C_A(\tilde{\pi}) \leq 2$, a permutation π such that $C(\pi) > C(\tilde{\pi})$ can be easily found, leading to a contradiction.

Suppose without loss of generality that $\lambda = \lambda_A(\tilde{\pi})[0] = \lambda_A(\tilde{\pi})[1] = \ldots = \lambda_A(\tilde{\pi})[C_A(\tilde{\pi}) - 1]$, while $\lambda_A(\tilde{\pi})[k] \neq \lambda$ for all $k \geq C_A(\tilde{\pi})$. From Lemma 2 we know that, given three different indices k, h, g smaller than $C_A(\tilde{\pi})$, necessarily, $\lambda_B(\tilde{\pi})[k] \neq \lambda_B(\tilde{\pi})[h]$ or $\lambda_B(\tilde{\pi})[h] \neq \lambda_B(\tilde{\pi})[g]$. The other $N - C_A(\tilde{\pi})$ positions $\lambda_B(\tilde{\pi})[k]$, $k \geq C_A(\tilde{\pi})$ are not constrained: in the worst case they are all equal to a wavelength λ', and there exist at most two indices $k, k' < C_A(\tilde{\pi})$ such that $\lambda_B(\tilde{\pi})[k] = \lambda_B(\tilde{\pi})[k'] = \lambda'$. In this case $C_B(\tilde{\pi}) = N - C_A(\tilde{\pi}) + 2$. Therefore, we obtain:

$$\begin{cases} C(\tilde{\pi}) = \min\{C_A(\tilde{\pi}), N - C_A(\tilde{\pi}) + 2\} \\ C(\tilde{\pi}) = \max_{\pi} C(\pi), \end{cases}$$

with $1 \leq C_A(\tilde{\pi}) \leq N$. From the previous equations $C_A(\tilde{\pi}) = (N+2)/2$ and the maximum wavelength reuse experienced using the Matrix Selection strategy for N even is $C = (N+2)/2 = \lfloor N/2 \rfloor + 1$. □

In addition to an upper bound on C, Theorem 1 also provides a WRM control algorithm, whose pseudo code is shown in Algorithm 1, in which this bound is not exceeded. Computing both $C_A(\pi)$ and $C_B(\pi)$ and selecting the wavelength assignment ensuring the minimum wavelength reuse is enough to achieve $C \leq \lfloor N/2 \rfloor + 1$.

Algorithm 1 Plane Selector (PS)

1: **Given:** a permutation π
2: **for** $i = 0$ to $N - 1$ **do**
3: compute $\lambda_A(\pi)[i]$
4: compute $\lambda_B(\pi)[i]$
5: **end for**
6: **if** $C_A(\pi) \leq C_B(\pi)$ **then** return $\lambda_A(\pi)$
7: **else** return $\lambda_B(\pi)$
8: **end if**

Matrix Combination (MC) strategy

The Matrix Selection strategy does not exploit the fact that wavelength assignments A and B can be combined together to further increase the scalability of the proposed WRM. We present here two approaches to combine together assignment A and B. We call this solution Matrix Combination (MC) strategy. The analytical derivation of the values of wavelength reuse obtained by the MC strategy is left for further study.

Strategy 2. Matrix Combination. Consider a permutation π. The MC strategy makes the following choice for each $i \in \{0, \ldots, N-1\}$.

$$\lambda(\pi)[i] = \begin{cases} \lambda_A(\pi)[i] \\ \lambda_B(\pi)[i] \end{cases} \quad \text{minimizing } C(\pi)$$

As a first approach, we implement the MC using Algorithm 2. Algorithm 2 is an Exhaustive Algorithm (EA) and explores of all the possible combinations offered by assignments A and B. We denote by $\lambda^k(\pi)$ the vector of the N indices of wavelengths required to satisfy the permutation π with the following characteristic: Each wavelength $\lambda^k(\pi)[i]$ with $i \in \{0, \ldots, N-1\}$ is either $\lambda_A(\pi)[i]$ or $\lambda_B(\pi)[i]$. Since each $\lambda^k(\pi)$ uses a different combination of the configurations A and B for each one of the N inputs, there exist 2^N combinations. Therefore, the EA complexity is $\Theta(2^N)$.

Algorithm 2 Exhaustive Algorithm (EA)

```
1: Given: a permutation π
2: for i = 0 to N − 1 do
3:     compute λ_A(π)[i] and λ_B(π)[i]
4: end for
5: for k = 0 to 2^N do
6:     compute the kth combination λ^k(π)
7:     compute C_k(π)
8: end for
9: return λ^k(π) with the lowest value of C_k(π)
```

Since the EA algorithm exhibits a complexity increasing exponentially with the number of inputs, we introduce a Greedy Algorithm (GA) able to ensure a low wavelength reuse at a reduced complexity. The pseudo code of GA is shown in Algorithm 3. At each time slot and for each input (starting randomly and proceeding sequentially), GA computes both $\lambda_A(\pi)[i]$ and $\lambda_B(\pi)[i]$ and selects wavelength assignment minimizing the current wavelength reuse. Hence, GA considers only local decisions and could not

minimize $C(\pi)$. However, it shows a complexity of $\Theta(N)$. Thus, it is more suitable than EA to be implemented in a real controller.

Algorithm 3 Greedy Algorithm (GA)

Given: a permutation π
v: vector of size N used as wavelength counter;
r: random input $r \in \{0, \ldots, N-1\}$
for $j = 0$ to $N-1$ **do**
    $i = (j + r) \mod N$
    compute $\lambda_A(\pi)[i]$ and $\lambda_B(\pi)[i]$
    if $v[\lambda_B(\pi)[i]] \leq v[\lambda_A(\pi)[i]]$ **then**
        $\lambda(\pi)[i] = \lambda_B(\pi)[i]$
    else
        $\lambda(\pi)[i] = \lambda_A(\pi)[i]$
    end if
    $v[\lambda(\pi)[i]]$++
end for
return $\lambda(\pi)$

3.3.5 Results

In this section, we present some scalability results on an optical switching fabric as the one depicted in Fig. 3.5. We consider complete input-output permutations π generated by the scheduler, which guarantees that N packets are transmitted at each timeslot. Since the scheduler could produce any π depending in the traffic pattern, all possible π instances should be considered. However, considering all the possible permutations as N increases becomes computationally intractable.

Fig. 3.8 shows C as a function of N. For $N \leq 11$ we could perform a "complete analysis": exploiting a backtracking technique, we generated all the possible permutations, applying on each permutation both EA and GA. Instead, for larger sizes of the WRM (N ranging from 12 to 30), we performed a "statistical analysis", generating 10^{13} random permutations for each N. We obtained the EA results for $N \leq 18$ only, due to its complexity.

AWG-WRM refers to the case when a single WRM is used, whereas the reuse C for the MS strategy is plotted according to Theorem 1. Thus, allowing a maximum wavelength reuse C, selecting a wavelength assignment between A and B (i.e. applying the MS strategy), it is possible to double the size of feasible optical fabric with respect the AWG-WRM case. The performance of the MC implementations are upper bounded by the MS strategy (which

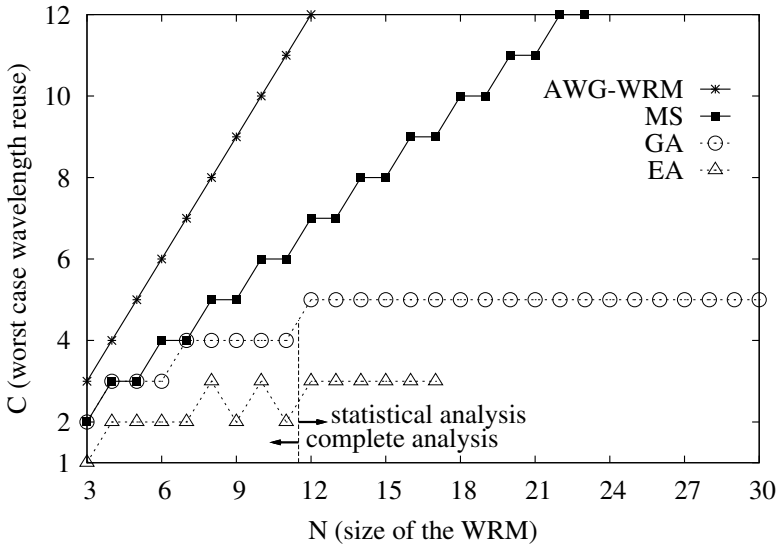


Figure 3.8. Maximum wavelength reuse C for different strategies and algorithms

reduces C by a factor of two). However, GA and EA results show that by combining assignments A and B it is possible to significantly reduce C, thus, making MRR-based WRM feasible also for large N.

3.3.6 Conclusions

In this section we proposed a MRR-based WRM suitable for a wavelength routed optical switching fabric exploiting MRRs as basic wavelength switching element. We considered the crosstalk limitation inherent to WRMs due to wavelength reuse and showed that a proper design and control of MRRs may enhance their scalability making them suitable for future high-capacity optical switching fabrics.

We first introduced a MRR-based switching fabric that uses MRR's periodical TF and tunability to implement multiple wavelength assignments. Then, we presented and analyzed the Matrix Selection (MS) and the Matrix Combination (MC) strategies to reduce the wavelength reuse exploiting the proposed active MRR-based WRM. The MS strategy roughly divides by two the wavelength reuse factor if compared with the single WRM configuration. The MC strategy leads to a further significant crosstalk reduction. We described two possible implementations for the MC strategy, namely the Exhaustive Algorithm (EA) and the Greedy Algorithm (GA). The GA exhibits a considerably lower complexity, and, at the same time, a significantly lower crosstalk with respect the MS strategy. It can thus be considered as a good candidate control algorithm for the proposed MRR-based WRM.

Part II

Long-haul optical networks

Chapter 4

Flexgrid optical networking field trial

This chapter presents results from the first gridless networking field trial with flexible spectrum switching nodes and 620 km of installed fibre links. Signals at 10G, 12.25G, 42.7G, DP-QPSK 40G, DP-QPSK 100G and 555G are generated, successfully transported and switched using flexible, custom spectrum allocation per channel. Spectrum defragmentation is demonstrated using integrated SOA-MZI wavelength converters. Results show error-free end-to-end performance (BER<1e-9) for the OOK channels and good pre-FEC BER performance with sufficient margin to FEC limit for the 40G and 100G coherent channels as well as for the 555G super-channel.

4.1 Introduction

Future transport networks will need to deal with a mix of providers' traffic representing services (i.e. 10Gb/s legacy channels), core traffic (i.e. at 100G, 400 Gb/s and beyond), as well as alien traffic, which could be variable bit rate and format channels. Hence, optical nodes will need to allocate resources in a flexible and efficient manner to support a mix of super-channels and lower speed channels. The nodes' complexity will largely depend on the network segment (i.e. inner core, metro), and should also facilitate transparent interoperability between segments. In addition, metro segments might carry legacy 10 Gb/s requiring dispersion compensated (DC) links whereas inner core segments might just use coherent compensation techniques (e.g. DSP) to support high speed and super-channels. To address increasing traffic growth, advances in modulation formats enable a 100G channel to fit in a standard 50-GHz WDM slot. However, this may not be the case for

higher bit-rate channels. For instance, super-channels at 400 Gb/s [40], 1 Tb/s [41] and beyond [42, 43] will occupy broader spectrum, which neither fits within the existing ITU grid nor is supported by conventional optical network infrastructures. Moreover, simultaneously supporting a combination of high-capacity super-channels and lower bit rate channels is critical [44].

Flexible and gridless optical networking is proposed to address such diverse requirements so as to switch and transport mix line rate technologies ranging from 10 Gb/s (25GHz spectrum) for better spectral efficiency to 555 Gb/s (650GHz spectrum). However, transporting mixed signal bit rates and modulation formats in such a flexible manner could lead to spectral fragmentation and increased blocking [45]. To overcome this, a super-channel or multiple lower bit rate channels need to be moved to a different spectral region. Wavelength conversion could provide a vital network function for such spectrum defragmentation optimizing spectral efficiency. To represent this evolving network scenario, we report results from the first, to the best of our knowledge, gridless optical networking field trial with adaptive and flexible spectrum inner-core node as well as flexible spectrum switching nodes placed in different geographical locations, connected by several field fiber links totaling 620 km. We have successfully demonstrated flexible switching and transport of mixed traffic including a high-speed super-channel at 555 Gb/s (650 GHz bandwidth), coherent 100G and 40G (50 GHz), 40G OOK NRZ (100 GHz), 40G OOK RZ (150 GHz), 12.25 and 10 Gb/s NRZ (25 GHz) signals. Also, the flexible-architecture inner-core node demonstrates adaptive architecture reconfiguration as in [46], mixed channels' switching and spectral defragmentation using wavelength converters based on cross-gain modulation (XGM) in a semiconductor-optical-amplifier Mach-Zehnder Interferometer (SOA-MZI).

4.2 Gridless network scenario

As shown in Fig. 4.1, the field trial gridless network is comprised of three optical nodes placed in different geographical locations and connected by 4 field fiber links with total 620-km of installed standard single mode fiber (SSMF). Links 2–4 are links of conventional design with in-line DCMs, whereas link 1 is a new DCM–less link of total 410km, which has five in-line amplifiers located in BT exchanges from Ipswich to London BT Tower, and looping back to Ipswich. The three optical nodes implement flexible spectrum switching whereby signals are switched all-optically with a custom bandwidth allocation per channel. Node-2 is a flexible-architecture node [46] that supports on-demand architecture reconfiguration to provide

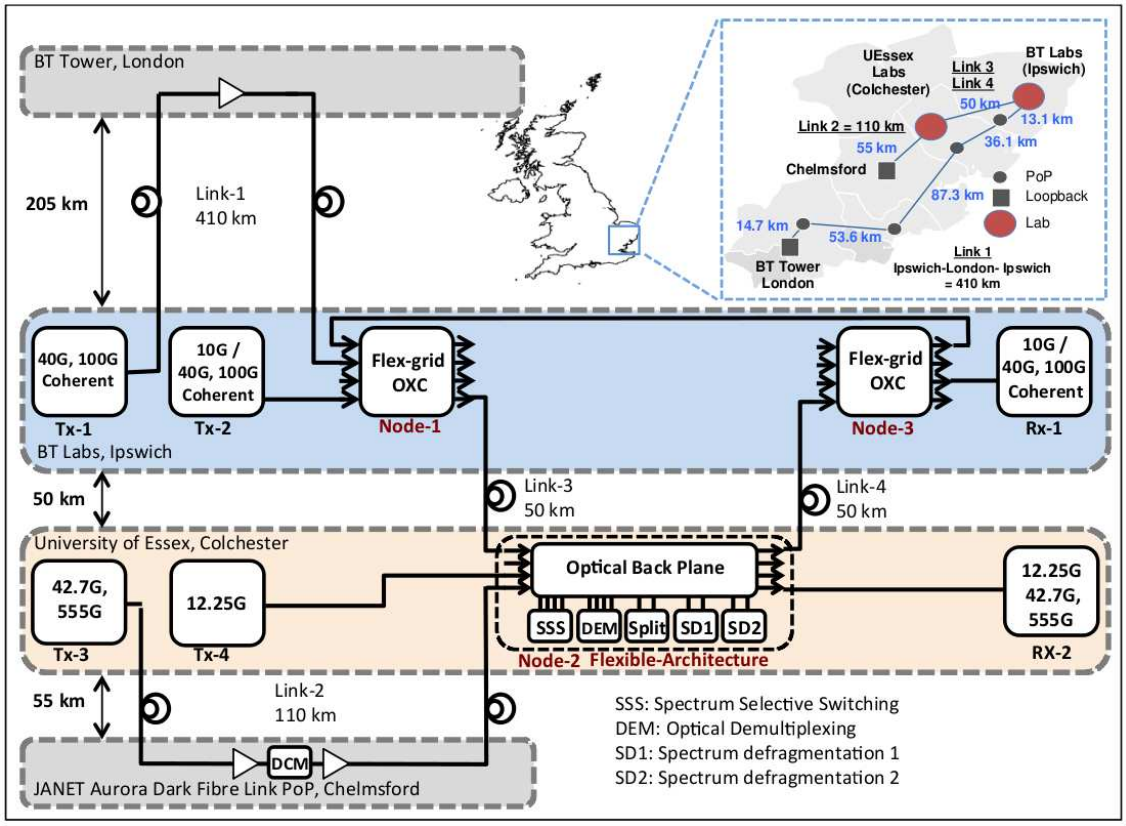


Figure 4.1. Gridless networking scenario and field trial map

multiple functions including spectrum switching and wavelength conversion for defragmentation. Hence, the field trial optical network represents a potential future flexible network, with different types of optical nodes with varying level of network functionality, e.g. in Core and Metro, with both conventional in-line DCM design and new DCM-less design. At Tx-1 and Tx-2, coherent 40G (DP-QPSK) and coherent 100G (dual carrier, DP-QPSK) are generated by commercial WDM transponders [47]. The coherent 100G and 40G signals from Tx-1 are transmitted over Link-1 (410-km DCM-less) to Node-1. Also, a coherent 100G, 40G and a standard 10G NRZ channels from Tx-2 are input to Node-1 but without prior transmission. At Node-1 channels are combined using flexible spectrum switching, with a custom bandwidth allocation per channel, and transmitted over Link-3 (50-km dispersion compensated) to Node-2. Meanwhile, in Tx-3 channels 1x555 Gb/s, 3x42.7 Gb/s OOK RZ, 1x42.7 Gb/s OOK NRZ are generated and transmitted over Link-2 (110-km dispersion compensated) to Node-2. Also, Tx-4 generates channels 3x10Gb/s OOK NRZ. Node-2 provides a flexible-architecture platform whereby modules (subsystems) are interconnected through a backplane

(3D-MEMS) to form optical nodes with on-demand functionality and able to reconfigure on the fly (20ms per 3D-MEMS cross-connection) according to traffic requirements, e.g. spectrum defragmentation when and where required. All input signals to Node-2 have to be transported over Link-4 to Node-3. However, there is contention between wavelengths, as shown in Fig. 4.2(b). Thus, Node-2 implements spectrum defragmentation and the signals are successfully switched, using flexible spectrum switching, to Node-3. In Node 3, the signals originally generated by Tx-1 and Tx-2 are dropped and input to the receiver Rx-1 for performance evaluation. Signals originally generated by Tx-3 and Tx-4 are spectrum-switched at Nodes 2, 3 and 1 and sent back to Node-2. At Node-2 they are dropped and input to the receiver (Rx-2) for performance evaluation.

4.3 Experimental setup and results

As shown in Fig. 4.2, the 555 Gb/s signal is generated from a 10.675-GHz 2-ps MLL pulse train followed by super-continuum generation in a highly non-linear fiber (HNLF), band-pass filtering and frequency-time transformation in dispersive medium (274.23-ps/nm to achieve a 23.4-ps delay between adjacent sub-carriers). Then, the signal is intensity modulated in a LiNbO3 MachâĂŞZehnder modulator (MZM) with a 42.7-Gb/s signal, which is composed of four electrically multiplexed pseudo-random bit sequences (PRBS) of length 2^7-1, 2^9-1, 2^{10}-1 and 2^{11}-1, at 10.675 Gb/s each. Thus, adjacent sub-carriers are modulated with a different PRBS. Additional 42.7 Gb/s and 12.25 Gb/s OOK signals are generated using LiNbO3 MZM at Tx-3 and Tx-4 respectively. Nodes 1 and 3 are flexible spectrum switching nodes using an LCoS-based spectrum selective switch (SSS) implemented with a WaveShaper [48]. The SSS switches programmable C-band spectrum slots from 10-GHz up to 5-THz with a 1-GHz resolution. The flexible-architecture Node-2 is implemented with a 96x96 3D-MEMS optical backplane that interconnects signal-processing modules such as the SSS and two SOA-MZI wavelength converter configurations at 12.25 Gb/s [49] and 42.7 Gb/s [50] , as shown in Fig. 4.2(c). In Nodes 1, 2 and 3, signals are allocated a customized spectral bandwidth per channel. For instance, a 650-GHz spectrum slot is allocated for the 555 Gb/s; 50 GHz for the 100 Gb/s DP-QPSK and 40 Gb/s DP-QPSK; 100 GHz for the 40 Gb/s NRZ; 150 GHz for the 42.7 Gb/s RZ; 50 GHz for the 12.25 Gb/s at $\lambda 8$/ $\lambda 8$' and 25 GHz for the remaining 12.25 Gb/s and 10 Gb/s signals.

The performance of coherent 100G and 40G is measured over the field trial network switching and transport; it is also compared with the point-to-point

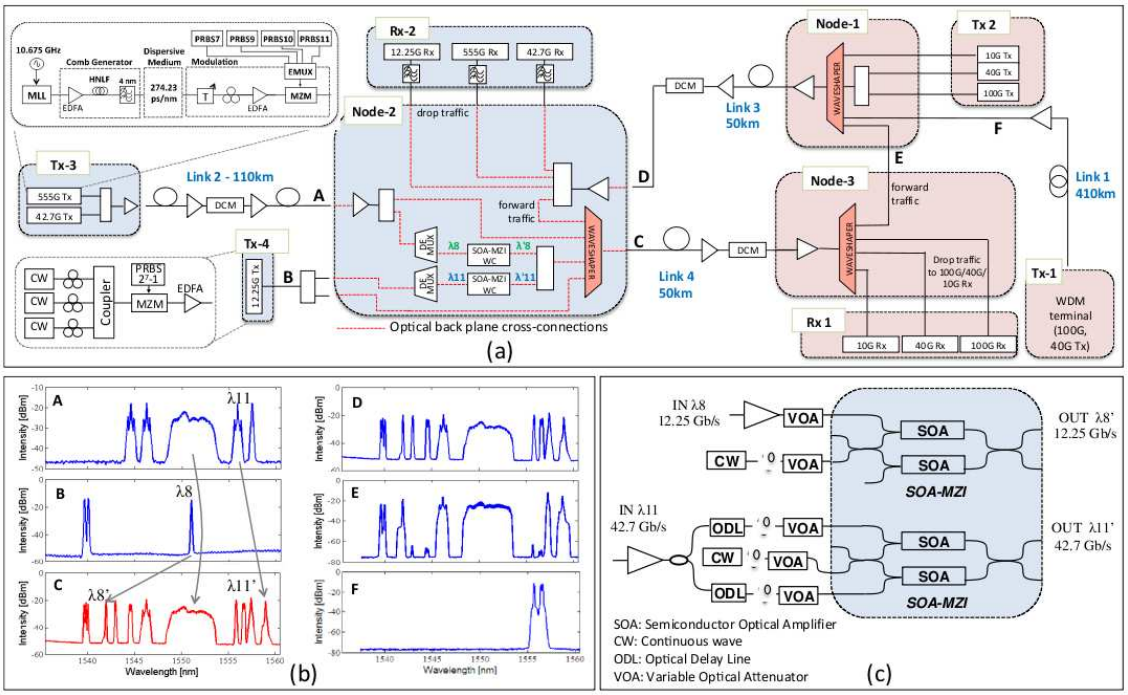


Figure 4.2. (a) Field trial gridless networking setup, (b) spectra at different points in the setup; A, B and C illustrate spectrum defragmentation, (c) setup used for SOA-MZI wavelength converters

transmission. Error free performance is achieved with long-term stability as shown in Fig. 4.3(a) with the measured Pre-FEC BER values over the time. Fig. 4.3(b) shows the spectrum of the 555 Gb/s signal at the receiving input to Node-2 and the end-to-end performance shows BER below 10^{-4}, which is a comfortable margin to the FEC limit of $2x10^{-3}$. Fig. 4.3(c) shows BER measurements of 42.7 Gb/s RZ and NRZ channels at various points in the experimental setup. Transmission over Link-2 introduces a 0.3-dB penalty to the RZ channel and a 1.6-dB penalty to the NRZ channel. Spectrum switching at Node-2 adds a penalty of 0.4 dB and 1 dB to the RZ and NRZ respectively. Although the combined length of Links 3 and 4 is less than the total length of Link-2, the penalty they introduce is much higher at 2 dB for the 42.7G RZ and an error floor (BER<1e-9) for 42.7G NRZ. This is due to the relatively high loss (∼28 dB) of each of the links (Link-3, Link-4), caused by additional loss from patch panel connections, and associated OSNR degradation when amplifying the weaker signals at the receiving end. As expected, the RZ format is more robust than the NRZ [51]. The performance of the wavelength converters is evaluated by means of bit error rate measurements and results are presented in Fig 4.3(d). The power penalty at BER = 10^{-9} is 1.5 dB for the 10-Gb/s SOA-MZI converter and 6 dB for the 42.7-Gb/s SOA-MZI converter.

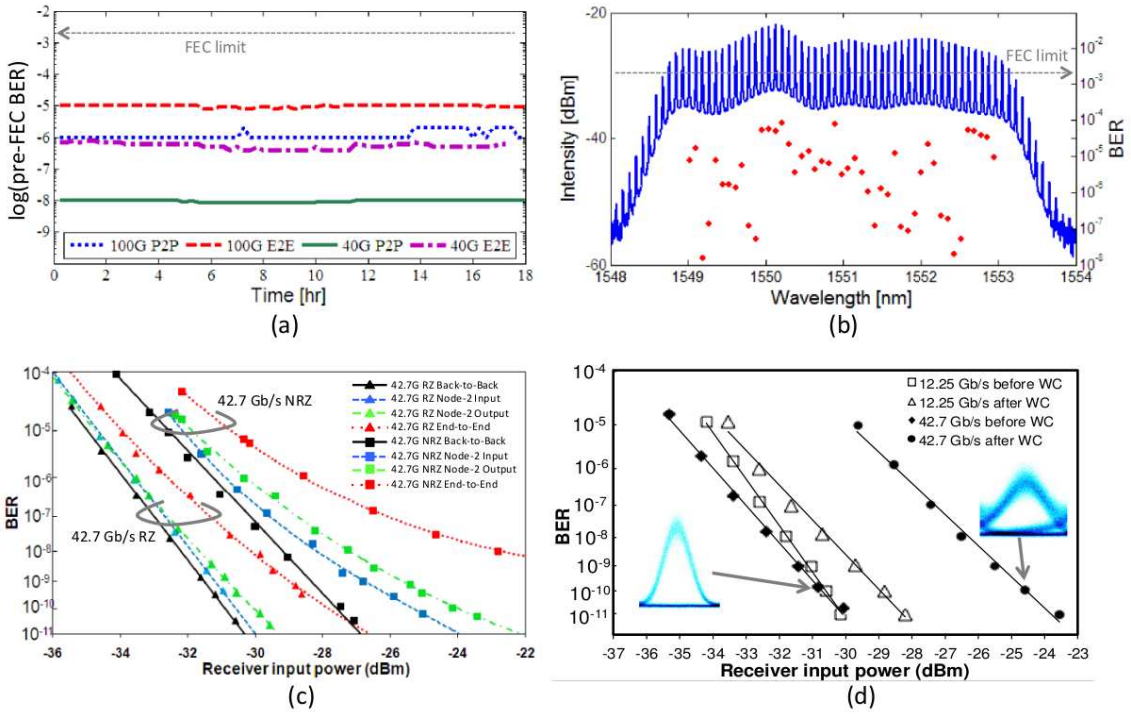


Figure 4.3. (a) BER of coherent 40G and 100G channels (P2P: 410km, E2E: Gridless network 510 km), (b) 555 Gb/s signal at Node 2 and its end-to-end BER, (c) Performance of 42.7 Gb/s RZ (Îż9) and 42.7 Gb/s NRZ (Îż12) at several points in the setup (d) BER performance of 12.25-Gb/s and 42.7-Gb/s SOA-MZI wavelength converters

4.4 Channel filtering and spacing

Flexible allocation of bandwidth per channel requires considering individual spectral requirements (i.e. bit-rate and modulation format) and the filter shape of the devices used for (de)muxing. In an all-optical network successive filter stages may result in a reduced end-to- end bandwidth, which may cause signal distortions with an associated power penalty [52]. On the other hand, if the allocated bandwidth is too wide for the transported channel spectral resources are wasted. In order to evaluate the effect of narrow filtering a 42.7 Gb/s RZ channel was passed through a co-centered filter. The filter bandwidth is decreased from 160 GHz down to 40 GHz while the spectrum, eye and sensitivity of the output signal are observed. Results are presented in Fig. 4.4(a). As the filter bandwidth is reduced the edges of the signal spectrum are attenuated. This gradually closes the eye and introduces an increasing power penalty. A 0.9-dB penalty is observed at 100

GHz filter bandwidth increasing rapidly for narrower bandwidths. Additional considerations are required if channels are to be tightly packed e.g. 10G at 12.5-GHz spacing [53]. Here, highly selective filters are required to reduce inter-channel crosstalk. Fig. 4.4(b) shows the performance of a 10G channel with an adjacent 10G channel at varying channel spacings demultiplexed using a WaveShaper as a filter with 10-GHz bandwidth. There is a flat region where the SNR (Q^2) shows little variation, from 50 GHz down to around 25 GHz. The penalty at 20-GHz spacing is 1 dB and increases rapidly for narrower channel spacings. Such degradation greatly depends on the selectivity of the filter used for channel (de)multiplexing; thus, it may be improved by using steeper filters. However, packing channels closer together also increases the interaction between them and may give rise to non-linear impairments such as XPM and FWM, which also constrain channel spacing.

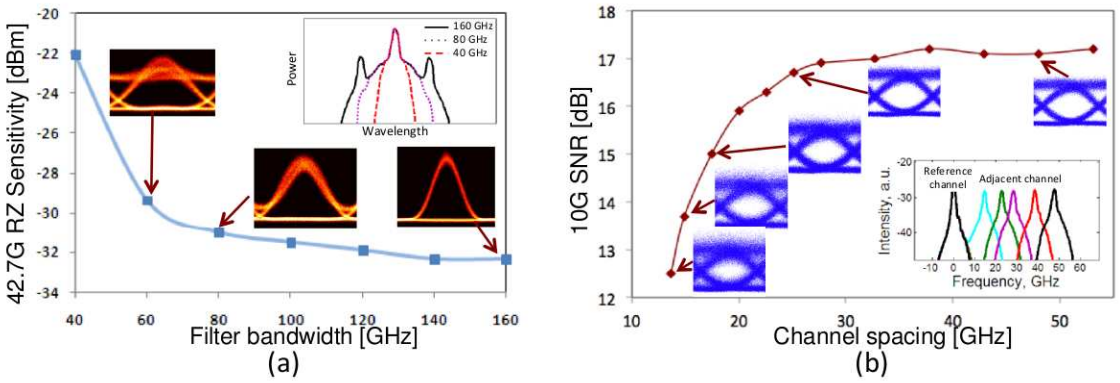


Figure 4.4. (a) Filtering effects on 42.7 Gb/s RZ signal and (b) 10G SNR performance for varying channel spacings

4.5 Conclusions

This work presents results from the first gridless optical networking field trial with geographically scattered flexible-spectrum-switching nodes linked by 620-km of field installed fibers, and spectrum defragmentation functionality. We have successfully demonstrated flexible spectrum switching and transport of mixed traffic with different bit rates and modulation formats including 555G, coherent 100G and 40G, as well as intensity modulated and wavelength converted 10G and 40G signals with good end-to-end BER performance. All channels are switched and transported using custom spectrum slots to support varying bandwidth requirement and optimize utilization (e.g. 555Gb/s on a 650 GHz slot, 3 adjacent 10Gb/s signals with a 25GHz spacing).

Chapter 5

Architecture on Demand

This chapter studies the Architecture-on-Demand (AoD) concept defined as an optical cross-connect (OXC) which dynamically synthesizes architectures suited to the switching and processing requirements of traffic. The design of OXCs is becoming very challenging since it has to fulfill requirements from legacy optical networks with low bit-rate channels and be future-proof to support high-speed super-channels. Therefore, optical node architectures require to handle the spectrum resources in a flexible manner. Indeed, AoD addresses the limitations of existing OXCs by providing flexible processing and switching of optical signals through customized/programmable architectures per degree according to traffic requirements.

The chapter begins with an introduction that includes current optical networking trends, the requirements for optical node architectures and the review of the AoD concept. Subsequently, the chapter is divided in two sections in order to perform an AoD scalability analysis in Sec. 5.2, and an AoD power consumption analysis in Sec. 5.3. The chapter ends with some conclusions for both analyses carried out.

5.1 Introduction and motivation

Current optical networks exploit the benefits of Wavelength Division Multiplexing (WDM). Conventional WDM scenarios capitalize on the huge bandwidth of wavelength multiplicity by means of well-defined slots in the optical spectrum. Therefore, traditional optical cross-connects (OXCs) accomplish the simple requirement of providing the same services per degree and for specific wavelengths. However, recent breakthroughs in optical networking are pushing towards flexible spectrum resource allocation. In particular, future optical networks may be able to handle channel bandwidth from

sub-wavelength to super-wavelength in a gridless manner so as to achieve higher spectral efficiency [54]. Three clear examples of the arising technologies that require high flexible spectrum allocation are: the SLICE concept which adjusts channel bandwidths by varying the number of sub-carriers in the transmitted OFDM signal [55]; optical packet switching systems that use several parallel wavelengths for packet transmission [56]; and elastic networks based on single carrier which adapt the transmission modulation format to transmit at higher data-rates by exploiting extra OSNR margins [57].

Such emerging technologies need the development of a new generation of OXCs and multi-degree reconfigurable optical add/drop multiplexers (MD-ROADMs) in particular. Hence, such optical systems must evolve from providing the same services per degree and for specific wavelengths towards arbitrary spectrum switching and custom functionalities per degree and wavelength. On the one hand, traditional OXC architectures are evolving from switching traffic all-optically at the wavelength level towards Multi-Granular OXC (MG-OXC) that perform fiber, waveband or wavelength switching [58, 59]. On the other hand, MD-ROADMs are changing from a simple aggregation of single degree ROADMs towards colorless, directionless and contentionless (C/D/C-less) ROADMs in which add/drop ports are not wavelength specific and transponders are not degree specific [60, 61].

Fig. 5.1 shows the usual ROADM architecture which is based on a broadcast-and-select configuration [60–62]. The main difference between these proposals relies on the composition of the transponder aggregator bank.

However, these approaches have three major drawbacks. First, they often consist of a hard-wired arrangement of devices restricting the upgradeability, limiting the support for new functionalities and reducing the capability of adapting the whole architecture to the network requirements. Even in the case that the architecture is mainly based on the usage of spectrum selective switches (SSS), functionalities such as spectrum defragmentation, time switching, regeneration or other methods of signal processing cannot be performed. Hence, these architectures lack on a desirable capability required for future optical networks: *flexibility*. Second,these recently proposed MG-OXCs and CDC-ROADMs are frequently designed with a scope of providing a service per granularity (i.e. degree, port, waveband or wavelength). The direct implication of such a design is the deployment of a device per granularity. In particular, the *scalability* of these architectures is hard limited by the number of devices required and the port-count of such devices. Third, the non-adaptable nature of these architectures implies that some devices contribute to the *power consumption* regardless of the traffic variations and the network requirements [63, 64].

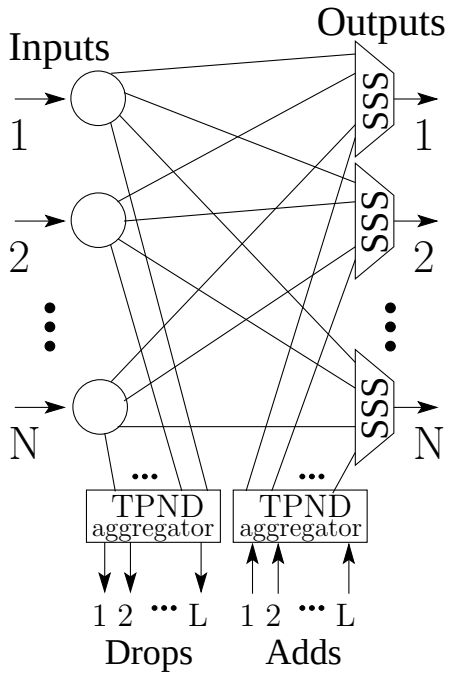


Figure 5.1. Usual ROADM architecture

In this work we assess these three limitations exploiting the benefits of Architecture on Demand (AoD) [65], defined as an optical cross-connect (OXC) that can implement and dynamically adapt its architecture in real time in order to provide the required functionality to fulfill the switching and processing requirements of network traffic. Recently, in [46] it has been experimentally demonstrated that AoD dynamically constructs architectures and supports elastic allocation of arbitrary spectral and time resources. In addition, in [7] are reported results from the first gridless networking field trial being AoD the main flexible spectrum switching node.

5.1.1 AoD concept

Unlike other architectures reported in the literature, Architecture on Demand (AoD) [65] dynamically synthesizes architectures suited to the switching and processing requirements of traffic. Fig. 5.2 depicts the implementation of AoD consisting of an optical backplane (e.g. 3D-MEMS) that interconnects inputs, outputs, adds, drops, single device building modules (e.g. MUX, coupler, spectrum selective switch (SSS)) and composed building modules (e.g. EDFA

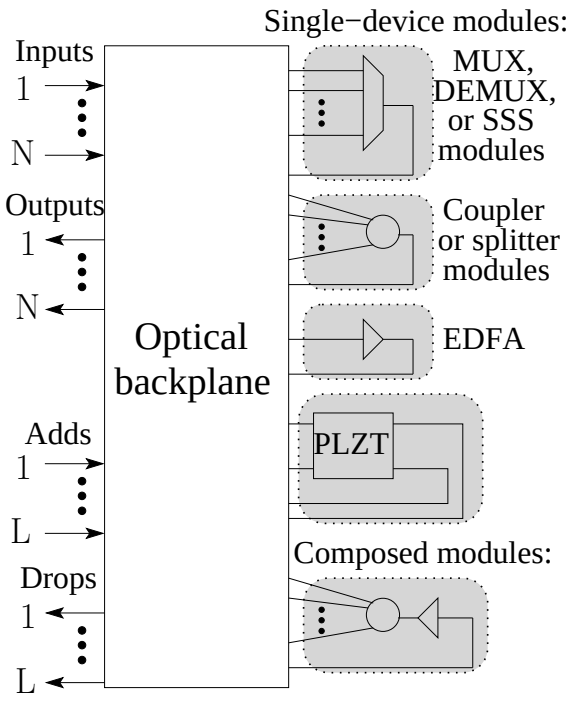


Figure 5.2. AoD implementation

+ splitter). AoD devises and constructs specific architectures interconnecting suitable building modules by means of backplane cross-connections according to the switching requests of the input signals. For instance, we recently reported [46] results from the first gridless networking field trial using an AoD node to provide on-demand functionalities such as spectrum switching and defragmentation.

The AoD node is an OXC that dynamically provides customisable architectures according to switching and processing requirements of network traffic. Fig. 5.2 shows a possible implementation of AoD consisting on an optical backplane that interconnects input ports, output ports and architecture-building modules. The optical backplane can be implemented with a large port-count optical switch (e.g. 3D-MEMS) and the building modules can be either single devices for optical processing such as MUX, DEMUX, spectrum selective switches (SSS), amplifiers, etc. or subsystems composed of several devices.

The AoD working principle is the following: given a set of switching requirements of arbitrary input signals, a specific AoD-OXC architecture is devised and constructed interconnecting suitable building modules by means of backplane cross-connections. Note that the resulting architecture is not hard-wired to accomplish particular requirements. Therefore, as

new processing or switching requirements arise, building modules can be dynamically added/removed from the implemented AoD-OXC architecture or re-located within the node to form new arrangements.

On the other hand, AoD also creates new challenges some of which are addressed in this work. First, AoD requires fast and efficient synthesis and selection algorithms that take into account the availability of building modules and traffic requests to build up the architecture. Second, given a number of building modules available for the AoD composition and a finite port-count optical backplane, it is crucial to know the maximum number and type of switching requests that can be handled, i.e. a scalability analysis. Other challenges, not addressed in this work, are the requirement for the AoD synthesis algorithms to work in conjunction with routing, spectrum and time assignment (RSTA) in order to maximise throughput or network utilisation and minimise blocking probability. Also for the implementation of AoD fast and ultra low-loss and high port count backplane technology is required.

5.2 Scalability analysis of AoD

This section reports an scalability analysis of AoD. It is first presented the AoD model as well as the model for the traffic requests in order to do the scalability analysis in Sec. 5.2.1. Subsequently, a preliminary technique suitable for AoD composition is presented in Sec. 5.2.2. Then, in Sec. 5.2.3 investigates the scalability properties of the architectures synthesized by the technique presented in terms of required cross-connections. Finally, the obtained results are compared with other proposed architectures in terms of required number of hardware modules.

5.2.1 AoD and traffic models

Fig. 5.3 depicts the model of the AoD node considered in this scalability analysis. A set of N input ports are connected to a set of N output ports through a maximum of two AoD modules. For instance, inputs may be connected either to a module or straight to an output, using an optical backplane cross-connection. In the case that a module is used, depending on the signal requirements a second module may be used before cross-connecting the signal to its required output. Here, we consider modules such as MUX, DEMUX, couplers, splitters and SSS.

We focus on a circuit switching scenario where a high layer control plane computes the switching requirements of the input channels (N ports with

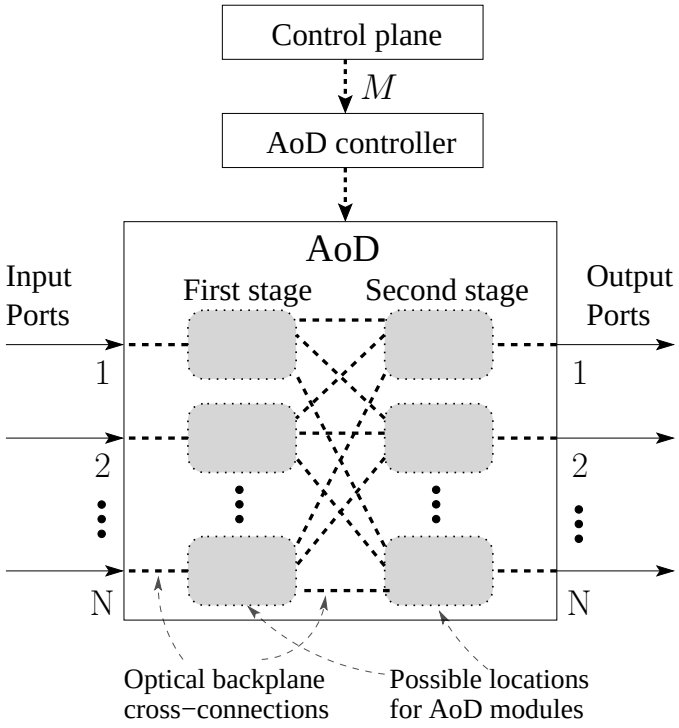


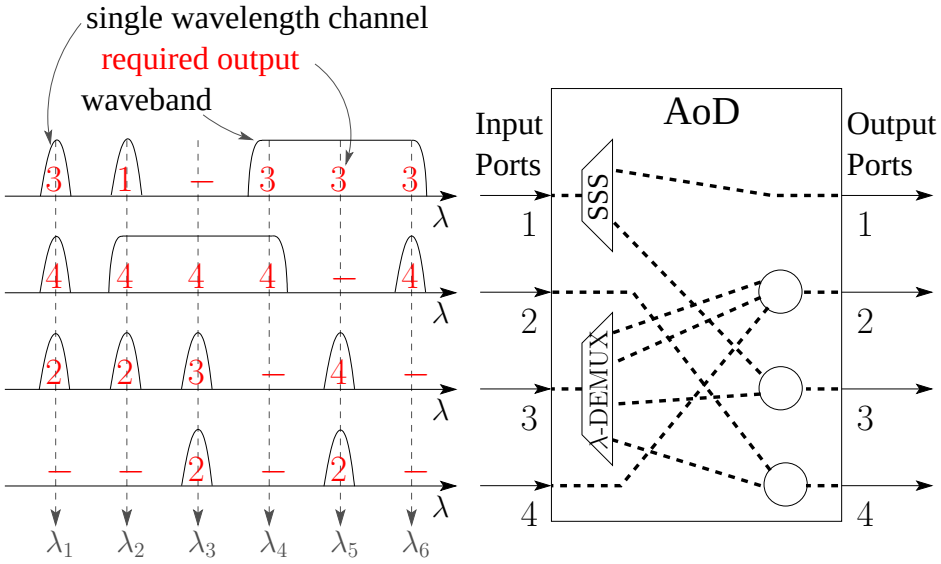
Figure 5.3. AoD two-stage model

W wavelengths[1]). We represent the switching requirements by a matrix M of size $N \times W$, where each $M(i, j)$ is the required output port of wavelength j from input i. In addition, wavebands are considered within M as aggregates of adjacent wavelengths. We assume that the routing and spectrum assignment algorithm (RSA) provides successful requests that can be accommodated i.e. there is no contention between signal destinations. Therefore, each output port index can appear at most once in each column of M. The controller of the AoD takes M as input and performs the synthesis algorithm phase computing the modules and cross-connections needed in order to provide the proper input-output interconnections (this is described in more detail in Sec. 5.2.2).

Fig. 5.4(a) shows an example of M for $N = 4$ and $W = 6$. It consists of a heterogeneous traffic scenario with 10 single wavelength channels and two waveband channels. The required output port for each channel is also shown in red in the same figure. Fig. 5.4(b) shows one possible AoD-OXC fulfilling the switching requirements of the considered channels. In particular, channels at input port 1 are fed to a SSS to perform flexible spectrum

[1]This means that we consider W available spectral slots.

switching functionality. Since all the channels fed to input port 2 require to be switched to output port 4, a cross-connection in the optical backplane is set. The same consideration is made at input port 4. A λ-DEMUX is placed at input port 3 since signals fed require to be wavelength-demultiplexed. At the second stage, output ports 2, 3 and 4 need to couple signals from multiple sources. Finally, note that in this example the number of hardware modules (components) required is 5 and the number of cross-connections is 13.



(a) Example M for N=4 and W=6 (b) AoD-OXC fulfilling M

Figure 5.4. Example of M and a possible AoD instance fulfilling it

The AoD model depicted in Fig. 5.3 employs a number of devices (D) such that

$$D \leq 2 \times N. \tag{5.1}$$

with equality (D_{MAX}) when the signals fed at each input port require to be switched to different output ports. Hence, the AoD controller needs to place one device per input port and one device per output port. As an example, D_{MAX} is achieved when M requires to wavelength-demultiplex all signals and also all outputs need to couple signals from multiple sources. In terms of backplane cross-connections, two different scenarios are taken into account. On the one hand, a maximum number of backplane cross-connections is achieved (C_{WB}) when all waveband channels fed at each input require to be switched to different outputs. For instance, it is required a first stage with SSS of $1 \times N$ ports and a second stage of $N \times 1$ couplers. Therefore, the number

of cross-connections is

$$C_{WB} = \overbrace{2 \times N}^{\substack{\text{cross-connections} \\ \text{between ports} \\ \text{and modules}}} + \overbrace{N^2}^{\substack{\text{cross-connections} \\ \text{between modules}}} . \tag{5.2}$$

On the other hand, for wavelength requests, a maximum number of backplane cross-connections is achieved (C_λ) when all available spectral slots W per input are occupied by single wavelength channels requiring to be demultiplexed to different outputs. In that case, the AoD controller needs to place a first stage of demultiplexers of $1 \times W$ ports and a second stage of $W \times 1$ couplers. Thus, the number of cross-connections is

$$C_\lambda = \overbrace{2 \times N}^{\substack{\text{cross-connections} \\ \text{between ports} \\ \text{and modules}}} + \overbrace{N \times W}^{\substack{\text{cross-connections} \\ \text{between modules}}} . \tag{5.3}$$

Equation (5.3) diverges from (5.2) due to the requirement of a cross-connection per wavelength between each module of the first stage and each module of the second stage.

5.2.2 AoD synthesis technique

The AoD controller, shown in Fig. 5.3, executes the algorithm to synthesise a suitable architecture according to the signals' switching requirements, stored in M.

The AoD Synthesis Algorithm (SA) is executed in three steps, from coarser to finer switching granularity. Fig. 5.5 shows the first step of the SA, which performs fiber switching. The destination of all signals from a given input is checked. In the case that they all go to the same output, a cross-connection is set (e.g. input ports two and four in Fig. 5.4). This is sequentially done until all inputs are considered.

Fig. 5.6 shows the second step of the SA where waveband switching is performed. It is divided in two parts. The first part, illustrated in the upper part of Fig. 5.6, places the building modules of the first stage and the cross-connections between the first and second stages. Thus, the presence of waveband channels is sequentially checked for all inputs. If waveband channels are found, a SSS is placed and a possible reuse of cross-connections is considered (due to the SSS arbitrary bandwidth switching capability),

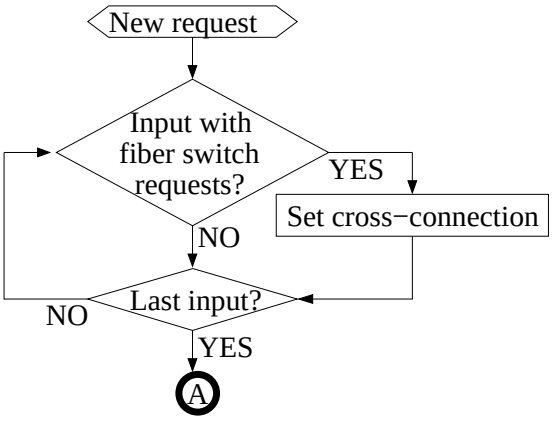


Figure 5.5. AoD Synthesis Algorithm (SA) step 1

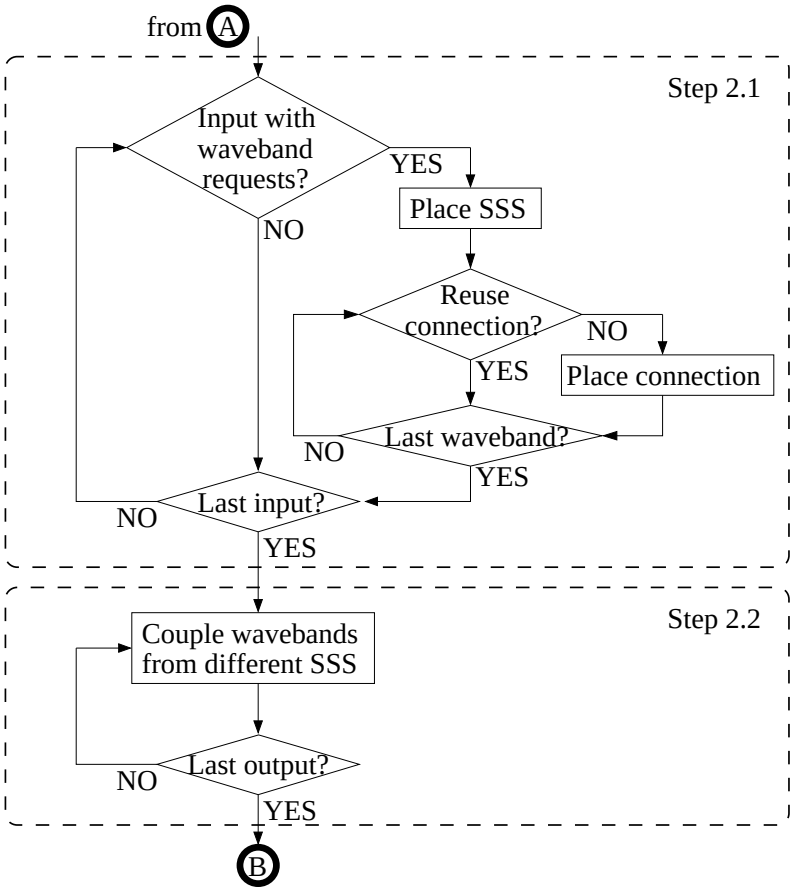


Figure 5.6. AoD Synthesis Algorithm (SA) step 2

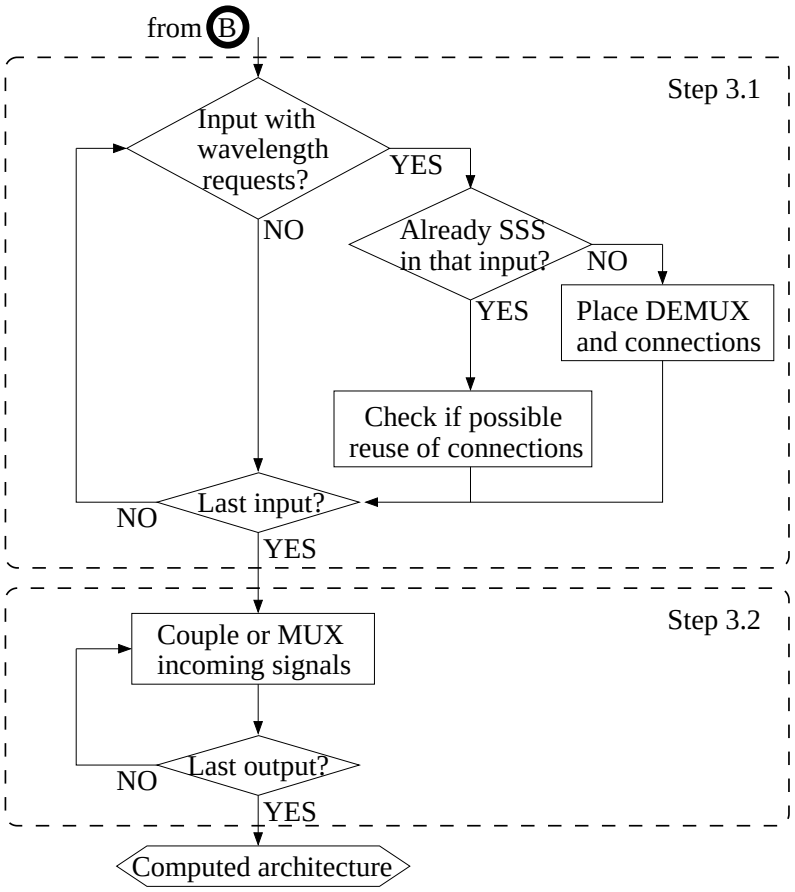


Figure 5.7. AoD Synthesis Algorithm (SA) step 3

otherwise the required connections are placed (e.g. input port one of Fig. 5.4). The second part, shown in the lower part of Fig. 5.6, places the building modules of the second stage. Here, all outputs are sequentially checked in order to couple wavebands from different sources.

The third SA step performs wavelength switching, as shown in Fig. 5.7. Like step two, it is divided into two parts. The first part checks if an input has wavelength channels to be switched and whether a SSS is already placed, which could possibly be reused together with already set cross-connections. Otherwise, a wavelength demultiplexer and the required connections are setup. The second part, shown in the lower part of Fig. 5.7, places the building modules of the second stage. Two types of building modules can be placed: a wavelength multiplexer when all the signals at the output considered are wavelengths, or a coupler in any case but mandatorily if any of the signals at the output considered is a waveband.

Note that the main motivation behind the SA sequential manner to proceed from coarser to finer granularity is the optimisation of the optical backplane cross-connections. For instance, fewer cross-connections are required if inputs are fiber-switch whenever possible or a reuse of cross-connections due to the SSS arbitrary bandwidth switching capability is exploited.

Finally, it is also worth to recall that each step of the SA explores all inputs in a sequential fashion and subsequently all outputs in a sequential fashion. Therefore, its complexity is $\Theta(N)$ making it a suitable algorithm to be implemented in a real AoD controller.

5.2.3 Simulations and results

In this section, we present some scalability results of the AoD model depicted in Fig. 5.3 working with the proposed SA and we compare the hardware requirements.

Our scenario focused on the off-time planning is divided in three phases for each switching requirement. First, we consider that no cross-connections are established in AoD and we generate requirement matrix, M, that must be fulfilled. Second, the AoD controller applies the proposed SA to compose an architecture that fulfills M. Third, as evaluation phase, we count the required number of optical backplane cross-connections used to setup the architecture computed. We repeat this process one thousand times for different set of requests M generated with some parameters. Results in terms of number of cross-connections (or number of components) are averaged and plotted in the figures. The number of cross- connections can be considered as an indicator of the optical backplane port-count required. We consider a uniform traffic scenario where a matrix M is generated for each switching requirement according to three parameters.

1. Port load P: We define P as the proportion of wavelength requirements at each input port. For instance, each row of M is generated with $P \times W$ switching wavelength requests (i.e. channels carrying signal). Hence, P sets the proportion of channels to switch in M.

2. Waveband requests ρ: In order to specify the level of channel aggregation into wavebands and their size appearing in each input port, we define a waveband load parameter ρ as

$$\rho = \frac{wB \times \Lambda}{W} \tag{5.4}$$

where wB denotes the number of wavebands and Λ their size. Therefore, at each input port, $\rho \times W$ wavelengths are grouped into wB wavebands of size Λ. Note that, according to the definition of waveband requests ρ and port load P, only scenarios where $\rho \leq P$ are considered. Once ρ and P have set the number of waveband and wavelength channels per input port, we determine their output port destination.

3. Fiber switch requests F: We define F as the proportion of fiber switch requests. Concretely, $F \times N$ input ports are assumed to request fiber switching in M and thus all wavelenght and waveband channels active in each of them are set to require the same destination. As an example, $F = 0.5$ is used at the scenario depicted in Fig. 5.4 since all channels fed at input two (four) require to be switched to output four (two). Conversely, the destination of channels that are not assumed as fiber switching is set according to a uniform distribution.

AoD scalability

Fig. 5.8 shows the number of cross-connections as a function of N for $\rho = 0$, $P = 0.9$, $W = 40$ and with F as a parameter. Also, the theoretical maximum number of cross-connections, C_λ, is plotted according to (5.3). A number of cross-connections close to C_λ is obtained when no fiber switching is requested ($F = 0$). Furthermore, as F increases, the number of cross-connections reduces until the minimum is achieved. Note that for $F = 1$, only N cross-connections are set, i.e. AoD performs only fiber switching for all ports.

Fig. 5.9(a) shows the number of cross-connections as a function of P with $F = 0$, $\rho = 0$ i.e. no waveband requests, $W = 20$ and N as a parameter. We observe that the number of optical backplane cross-connections linearly increases with the port load until it reaches the maximum according to (5.3). However, when the proportion of fiber switch requests is increased the number of cross-connections required by the architecture decreases as shown in Fig. 5.9(b). Therefore, the dimensioning of AoD optical backplane port-count must be tightly related with the port load and fiber switch requested and it shows that the gain of AoD depends on traffic aggregation. Note also that the number of cross-connections is roughly doubled from Fig. 5.9(b) to Fig. 5.8 since W is doubled.

Fig. 5.10 shows the number of cross-connections as a function of N and for $W = 40$, $F = 0$ i.e. no fiber switching and with a waveband requests ρ and port load P ranging jointly from 0 to 1. The number of wavebands and their size considered in ρ are listed in Table 5.1. In addition, the maximum number of cross- connections for this scenario, C_{WB}, is plotted according to (5.2).

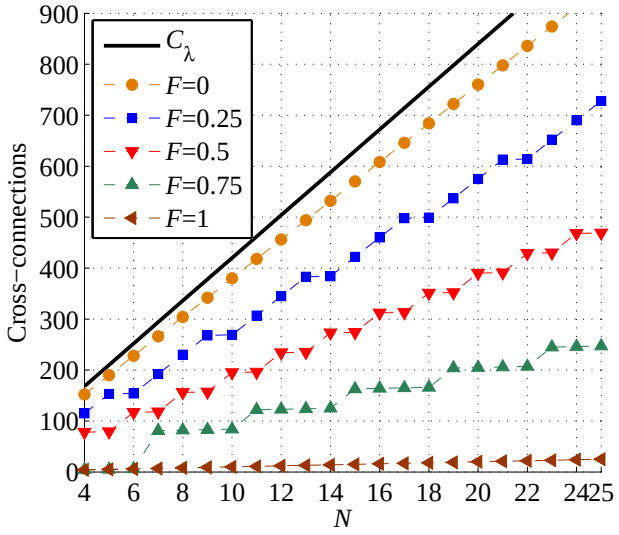


Figure 5.8. Cross-connections for $W = 40$, $\rho = 0$ and $P = 0.9$

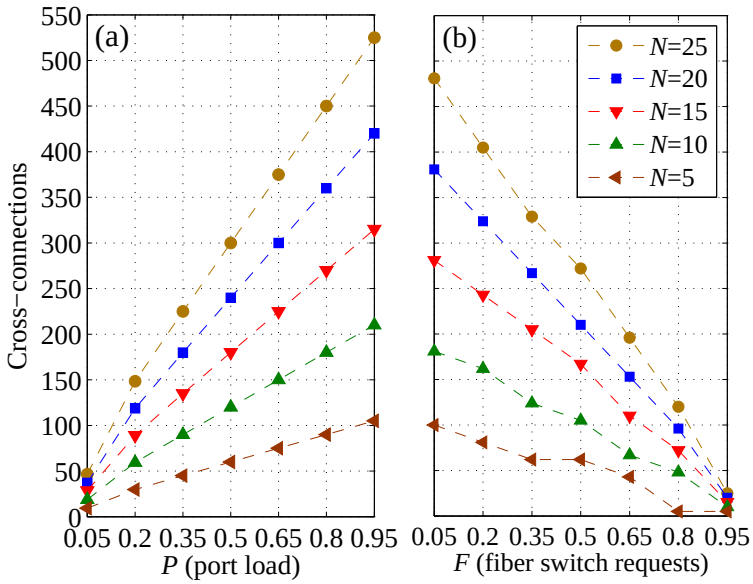


Figure 5.9. Cross-connections for W=20, ρ=0, with (a) F=0 and (b) P=0.9

For $\rho = 0$ all switching is carried out with (de)multiplexers. For $\rho > 0$ SSSs are introduced. We observe that fewer cross- connections are required when SSSs are employed for higher traffic loads ($\rho = P = 0.25$ compared with $\rho = 0$ $P = 0.1$). Higher values of ρ and P require a small increment in number of cross- connections. This is due to the capability of SSSs to aggregate several

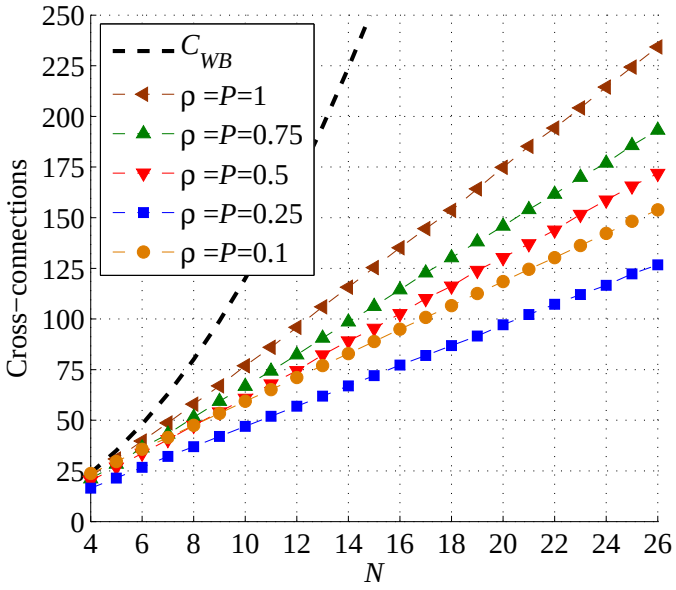


Figure 5.10. Cross-connections for $W = 40$ and $F = 0$

Table 5.1. Decomposition of ρ values for Fig. 5.10

ρ	wB^a	Λ^b
0.25	3	3
0.5	5	4
0.75	6	5
1	8	5

a. wB: number of wavebands b. Λ: size of the wavebands

wavelengths on a single port. Note that the trend expected from C_{WB} is not verified due to two factors: the number of wavebands per port considered is lower than N in most of the cases and SA proper reuses the optical backplane cross-connections.

Comparison of AoD with other architectures

We compare the CDC-ROADM architectures presented in [60, 61] with our AoD node model working under SA in terms of required number of hardware modules. In particular, CDC-ROADMs reported in [60] (cases from #2 to #5) and [61] are based on a broadcast-and-select configuration that use $2 \times N$ hardware modules at the input/output ports (e.g. N splitters and N SSSs respectively) and $2 \times N$ modules at the add/drop ports (e.g. N couplers and

N splitters or N multiplexers and N demultiplexers respectively). Therefore, these architectures use a total of $4 \times N$ modules. On the other hand, the flexibility of AoD to set any input (output) as add (drop) port enables its usage as CDC-ROADM with a maximum of $2 \times N$ modules according to (5.1). Concretely, our approach uses up to N hardware modules at the first stage that may be a combination of SSSs and demultiplexers and a second stage with up to N couplers or multiplexers. However, we also observe that the size of the AoD optical backplane is larger than the optical switches employed in these architectures [60, 61].

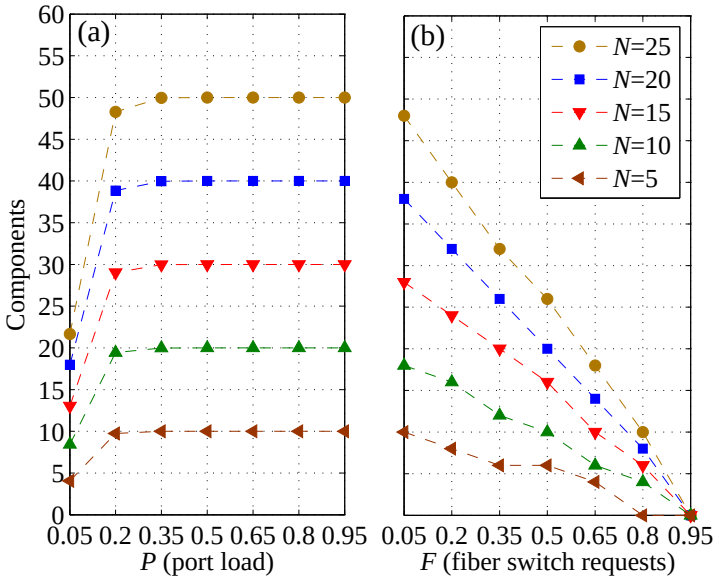


Figure 5.11. Cross-connections for W=20, ρ=0, with (a) F=0 and (b) P=0.9

Fig. 5.11(a) shows the number of components as a function of P with $F = 0$, $\rho = 0$ i.e. without wavebands, $W = 40$ and N as a parameter. According to our model, half the number of components consists of multiplexers placed at the first stage and half are couplers placed at the second stage. We observe that the number of components increases rapidly and reaches the maximum number ($2 \times N$) after a port load of 0.2. However, when the proportion of fiber switch request is increased there is an almost linear reduction on the number of components required from the maximum ($2 \times N$) to zero (when AoD performs completely fiber switching functionality) as shown in Fig. 5.11(b). Hence, AoD node uses fewer devices than the architectures reported in literature and also achieves a further reduction for different values of port load and fiber switch request. Note that the port load has a different impact on the number of cross-connections required and the number of components required (see

Fig. 5.11(a) and Fig. 5.9(a)). Therefore, the port load and the fiber switch request i.e. traffic aggregation are key parameters for dimensioning AoD.

5.3 Power consumption analysis of AoD

This section reports a power consumption analysis of AoD. First, several power consumption details are presented in Sec. 5.3.1. Then, in Sec. 5.3.2, it is detailed an AoD model and a model for the traffic requests with the addition of sub-wavelength time switching functionality. It is analyzed the power consumption of AoD according to this more complex model. Subsequently, an enhanced synthesis algorithm with respect to the previous section one is explained in Sec. 5.3.3. Then, in Sec. 5.3.4 a power consumption study of AoD is carried out and the obtained results are compared with other proposed architectures.

5.3.1 Power consumption details

Table 5.2 lists the number of devices required in several architectures reported in literature, being N the degree of the architecture and W the number of wavelengths per fiber. We have considered an add/drop ratio per degree of 25%. Note that we denote architectures #1 to #4 with the original numeric ID of [60]. The C/D/C-ROADM architecture #5 reported in [61] is based on the same structure of #2.

Table 5.2. Number of devices for different architectures

Arch.	**Active comp.**			**Passive comp.**	
	SSS	3D-MEMS[a]	Switch[b]	(de)Mux	Splitter
#1 [60]	0	1($3NW$)	0	$2N$	0
#2 [60]	N	0	$NW/2$	0	$2N$
#3 [60]	N	1(NW)	0	$2N$	N
#4 [60]	$2N$	2(NW)	0	0	$3N$
#5 [61]	N	0	$NW/2$	0	$2N$
#6 [62]	N	2(NW)	0	N	N

a 3D-MEMS slow switch (size) b Fast switch of size $1\times N$

Table 5.3 lists the power consumption values for the devices used in AoD and the considered architectures. The common equipment includes architecture controller, cooling fans and power supply. The high-speed piezoelectric lead lanthanum zirconate titanate (PLZT) switches are made

up of a switching device and a switch driver, with power consumptions in the order of mW and 8 W respectively. Hence, only the switch driver's contribution is considered. One or more 3D-MEMS switches are used to achieve the desired optical backplane port count for every particular architecture that AoD may compose. Therefore, the AoD backplane can be expanded by turning on successive 3D-MEMS switches.

Table 5.3. Consumption values

Device	**Power [W]**
Common equipment	100
3D-MEMS (360 ports) [66]	150
SSS [48]	40
Fast Switch (PLZT) [67]	8

5.3.2 AoD and traffic models including time switching

Fig. 5.12 depicts the model of the AoD node model for the power consumption analysis. It consist of a set of N input ports are connected to a set of N output ports either just through the optical backplane, or through 1, 2 or 3 stages. Indeed, N denotes the degree of the architecture. For instance, when input signals require both spectrum- and time-switching, the first stage performs spectrum routing by means of DEMUXes or spectrum selective switches (SSSs) [48], the second stage performs time-switching by means of piezoelectric lead lanthanum zirconate titanate (PLZT) [67] switches and the third stage couples or MUXes the output signals when needed.

We focus on a hybrid (i.e. fiber, waveband, wavelength and time) switching scenario where a high layer control plane gives the switching requests of the input channels to AoD represented by a matrix M of size $N \times W$. Element $M(i, j)$ is the destined output of wavelength j from input i which can be: a single number in case of wavelength channels, a "d" symbol in case of dropped wavelengths, and two numbers in case of sub-wavelength channels. Added wavelengths do not appear within M. Wavebands are represented within M as aggregates of adjacent wavelengths. We assume that a control plane provides successful requests that can be accommodated and all sub-wavelength time-sliced channels are fully synchronized. Therefore, each output index can appear in each column of M at most once in case of wavelength, or at most twice in case of sub-wavelength.

Fig. 5.13(a) shows an example of M for $N = 4$ and $W = 5$. It consists of a heterogeneous traffic scenario with 7 single wavelength channels,

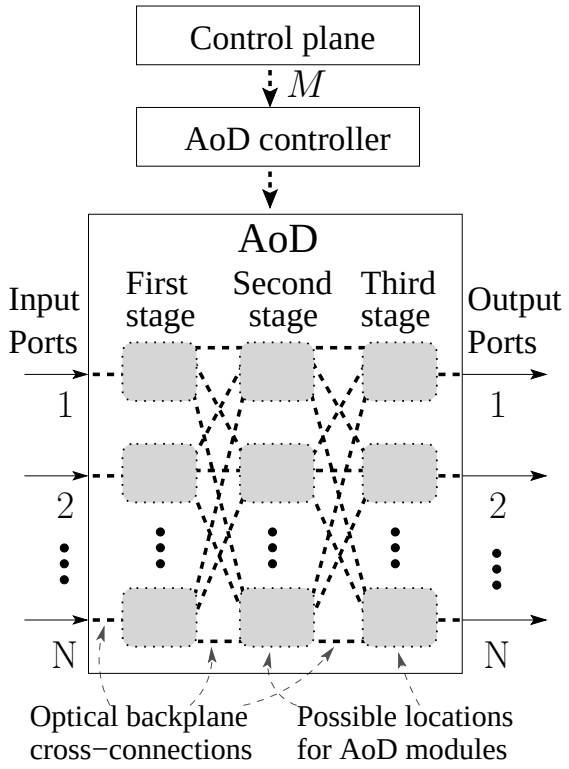
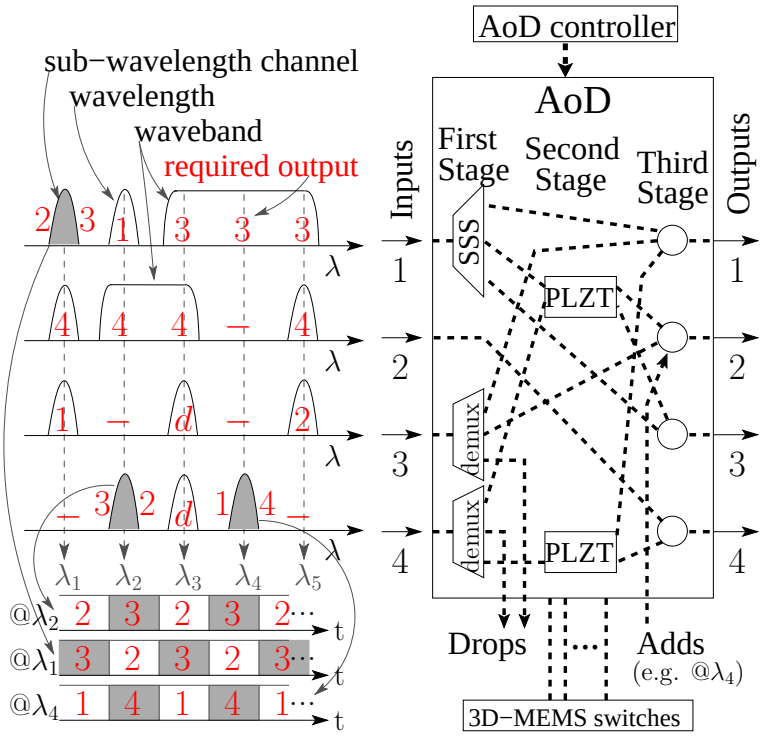


Figure 5.12. AoD three-stage model

two waveband channels and three sub-wavelength channels. The required output port for each channel is also shown in red in the same figure. Fig. 5.13(b) shows a possible AoD configuration fulfilling the switching requests. Channels at input 1 are fed to a $1\times N$ SSS to allow flexible spectrum switching and support the waveband channel. Channels fed to input 2 require to be switched to output 4, thus a cross-connection is set in the third stage. Two demuxes are placed at the first stage of inputs 3 and 4 since signals require to be demultiplexed. At the second stage, two PLZTs are placed to provide the fast time-switching functionality required by inputs 1 and 4, considering a possible reuse of the PLZT optical switch. Finally, note that in this example the number of hardware modules (devices) required is 9, the number of cross-connections is 22 and the power consumption is 306 Watts.

5.3.3 Enhanced AoD synthesis technique

The AoD controller executes the enhanced synthesis algorithm (E-SA) (see Fig. 5.14) divided in five steps to design the architecture in an automated fashion according to M. In our developed E-SA, four steps perform switching



(a) Ex. of M, N=4, W=5 (b) AoD fulfilling M

Figure 5.13. Example of M and its AoD configuration

functionalities from coarser to finer granularity (i.e. fiber, waveband, wavelength and sub-wavelength) and the fifth step couples signals from different sources.

More in detail, given M, the *first step* checks the destination of all signals from each input. In the case that they are all destined to the same output, a cross-connection is set (e.g. input port 2 in Fig. 5.13(b)).The *second step* checks the presence of waveband requests for each input. If waveband requests are found, a SSS is placed and a possible reuse of cross-connections is considered (due to the SSS arbitrary bandwidth switching capability), otherwise the required connections are placed (e.g. input port 1 of Fig. 5.13(b)). The *third step* checks the presence of wavelength channels to be switched for each input. It may reuse SSS and connections that have been already placed. Otherwise, a DEMUX and the required connections are setup. The *fourth step* checks the presence of sub-wavelength channels for each input. In such a case, a possible reuse of already placed SSS, DEMUXs and cross-connections is considered. Otherwise, a DEMUX is placed if needed. Subsequently, a possible reuse of already placed PLZTs is considered to provide time switching towards the

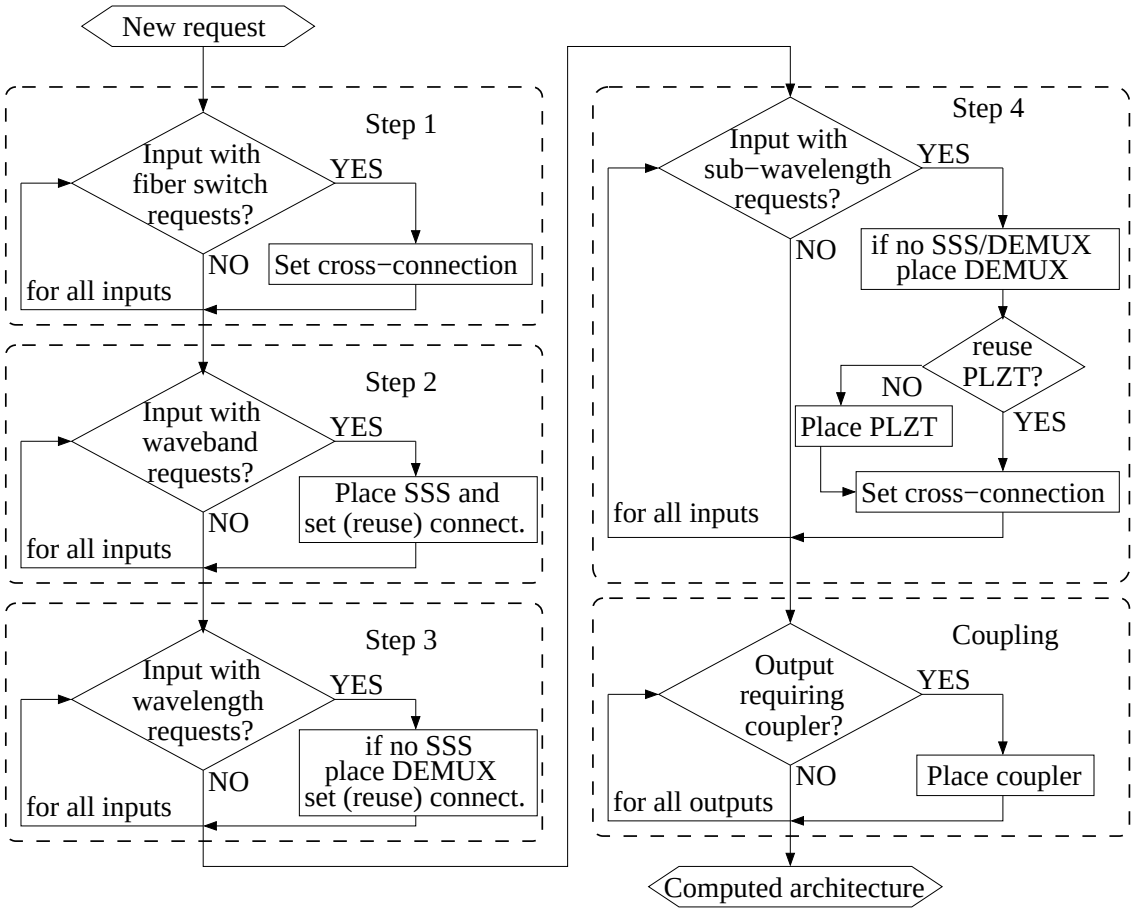


Figure 5.14. Enhanced synthesis algorithm (E-SA)

same two required outputs. Otherwise, a PLZT is placed to perform time switching between the two required outputs. Finally, in the *fifth step*, at each output the signals from different sources are coupled.

5.3.4 Simulations and results

We consider four off-line planning scenarios with different switching requirements. Given M, we start considering no cross-connections established in AoD, then, the AoD controller applies the developed SA to compose an architecture that fulfills M and finally we evaluate the metrics of the final architecture. We repeat this process one thousand times for different sets of requests, M, with a uniform traffic distribution.

The following four parameters are used in the simulations, being (1)–(3) recalled from Sec. 5.2.3 for the sake of readability:

1. Port load $P \in [0,1]$ is the fraction of requested wavelengths per input.

2. Waveband requests $\rho \in [0, P]$ is the fraction of requested wavelengths per input port that are set (randomly) into aggregates of two or three wavelengths.

3. Fiber switch requests F determines the proportion of inputs assumed to request fiber switching. Hence, all wavelength and waveband channels in each of these inputs require the same destination.

4. Sub-wavelength requests $\sigma \in [0, P]$ is the fraction of time-domain switching requests per input port

Tab. 5.4 describes the AoD stages and the parameters analyzed for the considered scenarios, assuming W=60 wavelengths per fiber.

Table 5.4. AoD stages configurations for each scenario

Scenario	1st stage	2nd stage	3rd stage	Parameters
1	demux	–	coupler	P, F
2	SSS	–	coupler	ρ
3	demux	PLZT	coupler	σ
4	demux/SSS	PLZT	coupler	–

Scenario 1: Fig. 5.15(a) shows the AoD power consumption as a function of P with F=0, ρ=0, σ=0 and N as a parameter, being Ⓐ at P=0.8. We observe that the power consumption increases with the port load since the number of required 3D-MEMS switches increases. However, the power consumption of AoD decreases, as shown in Fig. 5.15(b), as the channel aggregation into fiber switch requests (F) increases (being Ⓑ at F=0.5).

Scenario 2: Fig. 5.16(a) shows the AoD power consumption as a function of ρ with F=0, σ=0 and N as a parameter, being Ⓒ at ρ=0.7. Dashed lines and empty points show the power contribution due to SSSs. We observe that for low values of ρ the power consumption is mainly due to the common equipment and backplane. For values $\rho \geq 0.3$, SSSs are the main contributors to the total consumption, with a constant contribution.

Scenario 3: Fig. 5.16(b) shows the AoD power consumption as a function of σ with F=0, ρ=0 and N as a parameter, being Ⓓ at σ=0.7. Dashed lines and empty points show PLZT consumption. We observe that for low values of N, the power consumption is mainly due to the common equipment and backplane. However, as N and σ increase, the PLZT power consumption increases and becomes the major contributor to total power.

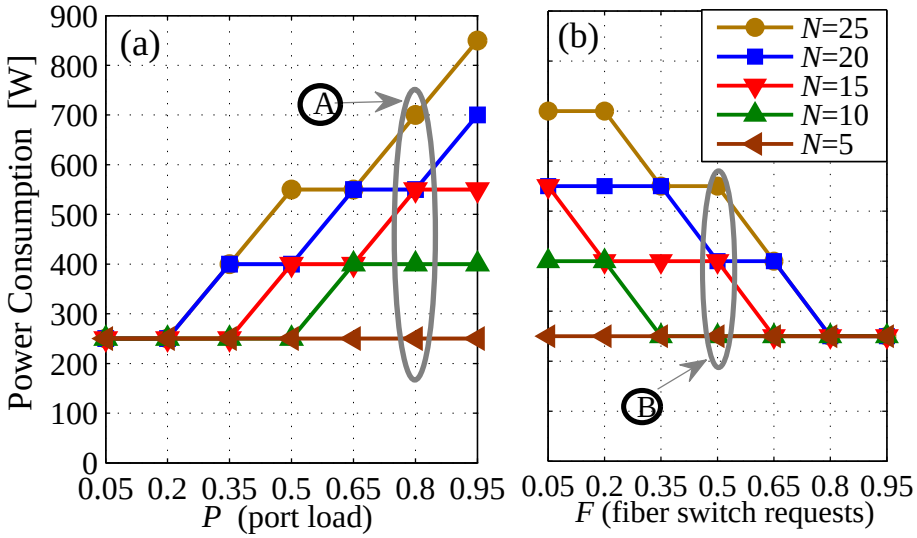


Figure 5.15. Power consumption for ρ=0, σ=0, (a) F=0 and (b) P=0.9

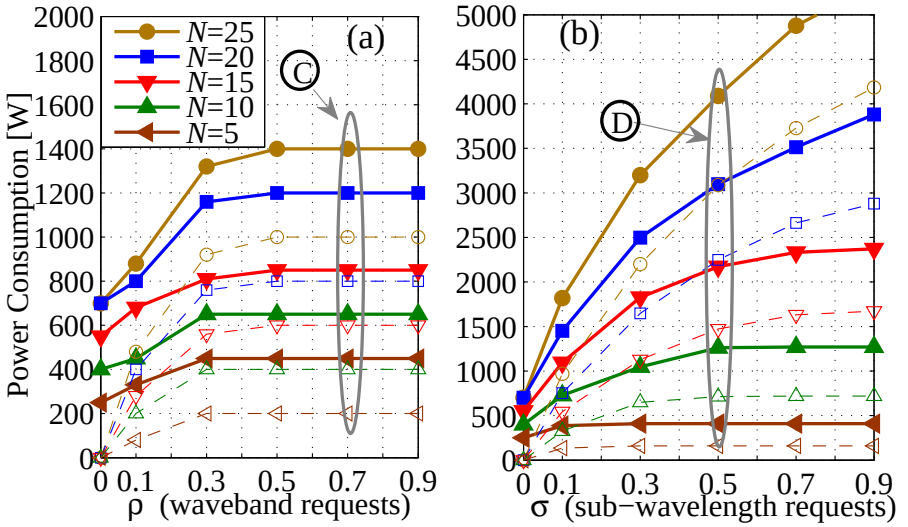


Figure 5.16. Power consumption for F=0, P=0.9, (a) σ=0 and (b) ρ=0

Scenario 4: Fig. 5.17 shows the power consumption for the architectures of Tab. 5.2. Notably, the inherent flexibility of AoD allows to save power compared to other architectures. For instance, more than a 70% reduction on the power consumption is achieved at Ⓑ due to traffic aggregation into fiber switching. In addition, time switching can be supported in AoD, unlike other architectures. However, when such capability is exploited, a higher power consumption can be experienced, as shown by curve Ⓓ in Fig. 5.17.

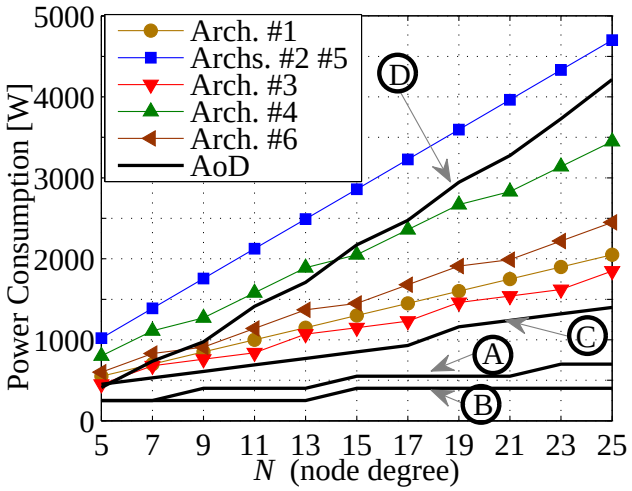


Figure 5.17. Power consumption comparison

5.4 Conclusions

In this work we reviewed the Architecture on Demand (AoD) concept and addressed its synthesis, scalability and its power consumption.

Regarding the scalability analysis, we first described the AoD node model and the switching requirements. Then, we presented the synthesis algorithm (SA) suited to compute the architecture. We discussed AoD scalability properties analysing the number of optical backplane cross-connections. Results show an efficient use of the cross-connections when SA performs switching at a coarser degree of granularity. In addition, the number of hardware modules used can be reduced at least by half compared to several optical node architectures reported in literature.

Regarding the power consumption analysis, we first characterized several ROADM architectures reported in the literature and the power consumption of active components. Then, we presented an AoD model which includes the sub-wavelength time switching functionality. In order to coop with the time-switching requests, we developed an enhanced synthesis algorithm (E-SA) suited to compute the architecture. Our results show that the adaptability of AoD offers significant power savings compared to other architectures unless power-demanding functionalities (e.g. time switching) are supported. Therefore, the AoD flexibility brings considerable power efficiency to the optical node.

Bibliography

[1] A. Bianco, D. Cuda, M. Garrich, R. Gaudino, G. Gavilanes, P. Giaccone, and F. Neri, "Optical interconnection architectures based on microring resonators," *IEEE International Conference on Photonics in Switching (PS'09)*, pp. 1–2, Sep. 2009.

[2] ——, "Optical interconnection networks based on microring resonators," *IEEE International Conference on Communications (ICC)*, pp. 1–5, May 2010.

[3] A. Bianco, D. Cuda, M. Garrich, R. Gaudino, G. Gavilanes, and P. Giaccone, "Optical interconnection networks based on microring resonators," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 4, no. 7, pp. 546–556, 2012.

[4] A. Bianco, D. Cuda, M. Garrich, G. Gavilanes, V. Martina, and F. Neri, "Crosstalk minimization in microring-based wavelength routing matrices," *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–5, 2011.

[5] A. Bianco, D. Cuda, M. Garrich, G. Gavilanes, and F. Neri, "Microring-based wavelength routing matrix," in *European Conference on Optical Communication (ECOC)*, 2011.

[6] N. Amaya, M. Irfan, G. Zervas, K. Banias, M. Garrich, I. Henning, D. Simeonidou, Y. Zhou, A. Lord, K. Smith, V. Rancano, S. Liu, P. Petropoulos, and R. D.J., "Gridless optical networking field trial: flexible spectrum switching, defragmentation and transport of 10G/40G/100G/555G over 620-km field fiber," *European Conference on Optical Communication (ECOC)*, 2011.

[7] ——, "Gridless optical networking field trial: flexible spectrum switching, defragmentation and transport of 10G/40G/100G/555G over 620-km field fiber," *Optics Express*, pp. B277–B282, 2011.

[8] M. Garrich, N. Amaya, G. S. Zervas, P. Giaccone, and D. Simeonidou, "Architecture on demand: Synthesis and scalability," *International Conference on Optical Network Design and Modeling (ONDM)*, pp. 1–6, 2012.

[9] ——, "Power consumption analysis of architecture on demand," *European Conference on Optical Communication (ECOC)*, pp. 1–3, 2012.

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

[11] S. I. A. ITRS, "International technology roadmap for semiconductors," 2009, http://www.itrs.net/Links/2009ITRS/Home2009.htm.

[12] A. V. Krishnamoorthy, X. Z. R. Ho, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, and J. E. Cunningham, "The integration of silicon photonics and vlsi electronics for computing systems intra-connect," in *Photonic in Switching*, 2009.

[13] M. Petracca, B. G. Lee, K. Bergman, and L. P. Carloni, "Design exploration of optical interconnection networks for chip multiprocessors," in *Hot Interconnects*, 2008.

[14] B. Lee, A. Biberman, J. Chan, and K. Bergman, "High-performance modulators and switches for silicon photonic networks-on-chip," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 1, pp. 6–22, 2010.

[15] L. Tobing and P. Dumon, *Photonic Microresonator Research and Applications*, I. Chremmos, O. Schwelb, and N. Uzunoglu, Eds. Springer, 2010.

[16] F. Xia, L. Sekaric, and Y. Vlasov, "Ultracompact optical buffers on a silicon chip," *Nature Photonics*, vol. 1, no. 1, pp. 65–71, 2006.

[17] B. Lee, A. Biberman, N. Sherwood-Droz, C. Poitras, M. Lipson, and K. Bergman, "High-speed 2×2 switch for multiwavelength silicon-photonic networks-on-chip," *Journal of Lightwave Technology*, vol. 27, no. 14, pp. 2900–2907, 2009.

[18] A. Bianco, D. Cuda, R. Gaudino, F. Neri, G. Gavilanes, and M. Petracca, "Scalability of optical interconnects based on microring resonators," *IEEE Photonic Technology Letters*, vol. 22, no. 15, pp. 1081–1083, July 2010.

[19] R. Ji, L. Yang, L. Zhang, Y. Tian, J. Ding, H. Chen, Y. Lu, P. Zhou, and W. Zhu, "Microring-resonator-based four-port optical router for photonic networks-on-chip," *Optics Express*, vol. 19, no. 20, pp. 18 945–18 955, Sep 2011.

[20] A. Biberman, N. Sherwood-Droz, B. Lee, M. Lipson, and K. Bergman, "Thermally active 4×4 non-blocking switch for networks-on-chip," in *IEEE 21st Annual Meeting of Lasers and Electro-Optics Society (LEOS)*, Nov 2008.

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

[22] A. Biberman, H. Lira, K. Padmaraju, N. Ophir, J. Chan, M. Lipson, and K. Bergman, "Broadband silicon photonic electro-optic switch for photonic interconnection networks," *IEEE Photonics Technology Letters*, 2011.

[23] B. G. Lee, A. Biberman, P. Dong, M. Lipson, and K. Bergman, "All-optical comb switch for multiwavelength message routing in silicon photonic networks," *IEEE Photonic Technology Letters*, vol. 20, no. 10, pp. 767–769, May 2008.

[24] P. Sanchis, P. Villalba, F. Cuesta, A. Håkansson, A. Griol, J. Galán, A. Brimont, and J. Martí, "Highly efficient crossing structure for silicon-on-insulator waveguides," *Optics letters*, vol. 34, no. 18, pp. 2760–2762, 2009.

[25] W. Bogaerts, P. Dumon, D. Thourhout, and R. Baets, "Low-loss, low-cross-talk crossings for silicon-on-insulator nanophotonic waveguides," *Optics letters*, vol. 32, no. 19, pp. 2801–2803, 2007.

[26] J. Van Campenhout, W. Green, S. Assefa, and Y. Vlasov, "Low-power, 2×2 silicon electro-optic switch with 110-nm bandwidth for broadband reconfigurable optical networks," *Optics Express*, vol. 17, no. 26, pp. 24 020–24 029, 2009.

[27] P. Dong, S. Liao, H. Liang, R. Shafiiha, D. Feng, G. Li, X. Zheng, A. Krishnamoorthy, and M. Asghari, "Submilliwatt, ultrafast and broadband electro-optic silicon switches," *Optics Express*, vol. 18, no. 24, pp. 25 225–25 231, 2010.

[28] Y. Kokubun, Y. Hatakeyama, M. Ogata, S. Suzuki, and N. Zaizen, "Fabrication technologies for vertically coupled microring resonator with multilevel crossing busline and ultracompact-ring radius," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 11, no. 1, pp. 4 – 10, Feb. 2005.

[29] A. V. Krishnamoorthy, X. Z. R. Ho, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, and J. E. Cunningham, "The integration of silicon photonics and vlsi electronics for computing systems intra-connect," in *Photonic in Switching*, 2009.

[30] N. Ophir, K. Padmaraju, A. Biberman, L. Chen, K. Preston, M. Lipson, and K. Bergman, "First demonstration of error-free operation of a full silicon on-chip photonic link," in *Optical Fiber Communication Conference*, 2011.

[31] E. Bonetto, L. Chiaraviglio, D. Cuda, G. Castillo, and F. Neri, "Optical technologies can improve the energy efficiency of networks," in *35th European Conference on Optical Communication (ECOC)*, 2009.

[32] R. Gaudino, G. Gavilanes, F. Neri, and J. Finochietto, "Simple optical fabrics for scalable terabit packet switches," in *IEEE International*

Conference on Communications, (ICC), 2008.

[33] ——, "Can simple optical switching fabrics scale to Terabit/s switches capacities?" in *IEEE/OSA Journal of Optical Communications and Networking (JOCN)*, vol. 1, no. 3, Aug 2009.

[34] A. Bianco, D. Cuda, G. Gavilanes Castillo, F. Neri, M. Rodelgo Lacruz, F. Gonzalez Castano, C. Lopez-Bravo, and M. Salvat, "Crosstalk limiting schedulers in awg based optical switches," *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–5, dec 2010.

[35] A. Bianco, D. Hay, and F. Neri, "Crosstalk-preventing scheduling in single- and two-stage awg-based cell switches," *Networking, IEEE/ACM Transactions on*, vol. 19, no. 1, pp. 142 –155, 2011.

[36] A. Bianco, D. Cuda, R. Gaudino, G. Gavilanes, F. Neri, and M. Petracca, "Scalability of optical interconnects based on microring resonators," *Photonics Technology Letters, IEEE*, vol. 22, no. 15, pp. 1081–1083, 2010.

[37] E. Klein, "Densely integrated microring-resonator based components for fiber-to-the-home applications," *PhD. Thesis,* University of Twente, PO. Box 217, 7500 AE Enschede, The Netherlands, April 2007.

[38] Q. Xu, D. Fattal, and R. Beausoleil, "Silicon microring resonators with 1.5-μm radius," *Optics Express*, vol. 16, no. 6, March 2008.

[39] A. Melloni and M. Martinelli, "The ring-based resonant router," in *International Conference on Transparent Optical Networks*, 2004.

[40] P. Winzer, A. Gnauck, S. Chandrasekhar, S. Draving, J. Evangelista, and B. Zhu, "Generation and 1,200-km transmission of 448-Gb/s ETDM 56-gbaud pdm 16-qam using a single I/Q modulator," *36th European Conference on Optical Communication (ECOC)*, Sept 2010.

[41] S. Chandrasekhar, X. Liu, B. Zhu, and D. W. Peckham, "Transmission of a 1.2-Tb/s 24-carrier no-guard-interval coherent OFDM superchannel over 7200-km of ultra-large-area fiber," *35th European Conference on Optical Communication (ECOC)*, Sept 2009.

[42] D. Hillerkuss, T. Schellinger, R. Schmogrow, M. Winter, T. Vallaitis, R. Bonk, A. Marculescu, J. Li, M. Dreschmann, J. Meyer, S. Ben Ezra, N. Narkiss, B. Nebendahl, F. Parmigiani, P. Petropoulos, B. Resan, K. Weingarten, T. Ellermeyer, and J. Lutz, "Single source optical OFDM transmitter and optical FFT receiver demonstrated at line rates of 5.4 and 10.8 Tb/s," *Optical Fiber Communication (OFC)*, March 2010.

[43] J. Yu, Z. Dong, X. Xiao, Y. Xia, S. Shi, C. Ge, W. Zhou, N. Chi, and Y. Shao, "Generation, transmission and coherent detection of 11.2 Tb/s (112$\times$100Gb/s) single source optical OFDM superchannel," *Optical Fiber Communication and Exposition (OFC)/NFOEC*, March 2011.

[44] T. Xia, G. Wellbrock, E. Yue-Kai Huang Ip, M.-F. Huang, Y. Shao, T. Wang, Y. Aono, T. Tajima, S. Murakami, and M. Cvijetic, "Field

experiment with mixed line-rate transmission (112-Gb/s, 450-Gb/s, and 1.15-Tb/s) over 3,560 km of installed fiber using filterless coherent receiver and EDFAs only," *Optical Fiber Communication and Exposition (OFC)/NFOEC*, March 2011.

[45] D. Geisler, Y. Yin, K. Wen, N. Fontaine, R. Scott, S. Chang, and S. Yoo, "Demonstration of spectral defragmentation in flexible bandwidth optical networking by FWM," *Photonic Technology Letters*, 2011.

[46] N. Amaya, G. Zervas, B. Rahimzadeh Rofoee, M. Irfan, Y. Qin, and D. Simeonidou, "Field trial of a 1.5 Tb/s adaptive and gridless oxc supporting elastic 1000-fold all-optical bandwidth granularity," *Optics Express*, vol. 19, pp. B235–B241, 2011.

[47] R. K., B. D., B. D., B. J., and L. C., "100G and beyond with digital coherent signal processing," *IEEE Communications Magazine*, no. 48, pp. 62–69, 2010.

[48] Spectrum selective switch, http://www.finisar.com/optical_instrumentation.

[49] D. Apostolopoulos, K. Vyrsokinos, P. Zakynthinos, N. Pleros, and H. Avramopoulos, "An SOA-MZI NRZ wavelength conversion scheme with enhanced 2r regeneration characteristics," *Photonic Technology Letters*, no. 21, pp. 1363–1365, 2009.

[50] M. Spyropoulou, N. Pleros, K. Vyrsokinos, D. Apostolopoulos, M. Bougioukos, D. Petrantonakis, and H. Miliou, A. Avramopoulos, "40 Gb/s NRZ wavelength conversion using a differentially-biased SOA-MZI: theory and experiment," *Journal of Lightwave Technology*, no. 29, pp. 1489–1499, 2011.

[51] M. Y. ., "4o-Gbit/s transport system: its WDM upgrade," *Optical Fiber Communication (OFC)*, pp. 323–325, March 2000.

[52] S. Tibuleac and M. Filer, "Transmission impairments in DWDM networks with reconfigurable optical add-drop multiplexers," *Journal of Lightwave Technology*, no. 28, pp. 557–598, 2010.

[53] H. Suzuki, M. Fujiwara, N. Takachio, K. Iwatsuki, T. Kitoh, and T. Shibata, "12.5 Ghz spaced 1.28 Tb/s (512-channel$\times$2.5 Gb/s) super-dense WDM transmission over 320 km SMF using multiwavelength generation technique," *Photonic Technology Letters*, no. 14, pp. 405–407, 2002.

[54] O. Gerstel, M. Jinno, A. Lord, and S. Yoo, "Elastic optical networking: A new dawn for the optical layer?" in *IEEE Communications Magazine*, Feb. 2012.

[55] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, and S. Matsuoka, "Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 66–73, 2009.

[56] N. Wada, S. Shinada, and H. Furukawa, "Modulation format free optical packet switching technology," in *International Conference on Transparent Optical Networks (ICTON)*, 2010.

[57] O. Rival and A. Morea, "Cost-efficiency of mixed 10-40-100Gb/s networks and elastic optical networks," in *Optical Fiber Communication Conference (OFC)*, 2011.

[58] M. Iyer, G. Rouskas, and R. Dutta, "A hierarchical model for multi-granular optical networks," in *International Conference on Broadband Communications, Networks and Systems, (BROADNETS)*, 2008.

[59] Y. Wang and X. Cao, "Multi-granular optical switching: A classified overview for the past and future," *IEEE Communications Surveys & Tutorials*, pp. 1–16, July 2011.

[60] P. Ji and Y. Aono, "Colorless and directionless multi-degree reconfigurable optical add/drop multiplexers," *Wireless and Optical Communications Conference (WOCC)*, pp. 1–5, 2010.

[61] S. Gringeri, B. Basch, V. Shukla, R. Egorov, and T. Xia, "Flexible architectures for optical transport nodes and networks," *IEEE Communications Magazine*, vol. 48, no. 7, pp. 40–50, 2010.

[62] R. Jensen, "Optical switch architectures for emerging colorless/directionless/contentionless roadm networks," in *Optical Fiber Communication Conference (OFC)*, 2011.

[63] A. Autenrieth, A. Tilwankar, C. Machuca, and J. Elbers, "Power consumption analysis of opaque and transparent optical core networks," in *International Conference on Transparent Optical Networks (ICTON)*, 2011.

[64] M. Murakami and K. Oda, "Power consumption analysis of optical cross-connect equipment for future large capacity optical networks," in *International Conference on Transparent Optical Networks (ICTON)*, 2009.

[65] N. Amaya, G. Zervas, and D. Simeonidou, "Architecture on demand for transparent optical networks," pp. 1–4, 2011.

[66] Calient's FiberConnect, http://www.calient.net.

[67] Piezoelectric lead zirconate titanate (PLZT) technology http://www.epiphotonics.com.