Optical crossbars on chip, a comparative study based on worst-case losses

Sébastien Le Beux, Hui Li, Gabriela Nicolescu, Jelena Trajkovic, Ian O'connor

Optical Crossbars on Chip, A Comparative Study based on Worst-Case Losses

Sébastien Le Beux[1]*, Hui Li[1], Gabriela Nicolescu[2], Jelena Trajkovic[3] and Ian O'Connor[1]

[1] Lyon Institute of Nanotechnology, INL-UMR5270
Ecole Centrale de Lyon, Ecully, F-69134, France

[2] Computer and Software Engineering Dept.
Ecole Polytechnique de Montréal
Montréal (QC), Canada

[3] Electrical and Computer Engineering Department
Concordia University
Montreal, QC, Canada

* Contact author: sebastien.le-beux@ec-lyon.fr

Abstract — The many-core design research community have shown high interest in optical crossbar on chip for more than a decade. Key properties of optical crossbars, namely a) contention free data routing b) low latency communication and c) potential for high bandwidth through use of WDM, motivate several implementations of this type of interconnect. These implementations demonstrate very different scalability and power efficiency abilities depending on three key design factors: a) network topology, b) considered layout and c) insertion losses induced by the fabrication process. In this paper, the worst-case optical losses of crossbar implementations are compared according to the factors mentioned above. The comparison results have the potential to help many-core system designer to select the most appropriate crossbar implementation according to, for instance, the number of IP cores and the die size.

Keywords—Optical Network on Chip, crossbar, optical losses.

I. INTRODUCTION

Technology scaling down to the ultra-deep submicron domain provides for billions of transistors on chip, enabling the integration of hundreds of cores. Many-core designs, integrating interconnect that can support low latency and high data bandwidth, are being increasingly required in modern embedded systems to address the increasing power and performance constraints of embedded applications. Designing such systems using traditional electrical interconnect presents a significant challenge: due to capacitive and inductive coupling [11], interconnect noise and propagation delay of global interconnect increase. The increase in propagation delay requires global interconnect to be clocked at a very low rate, which limits the achievable bandwidth and overall system performance.

In this context, Optical Network-on-Chip (ONoC) is an emerging technology considered as one of the key solutions for the future generation of on-chip interconnects. It relies on optical waveguides to carry optical signals, so as to replace electrical interconnect, and provide the low latency and high bandwidth properties of the optical interconnect. Moreover, 3D integration technologies allow for both optical and electrical layers to be stacked. Proposals for ONoC can, thus, realistically envision the integration of sufficient photonic devices for fast chip-length communications [8][2][3].

Among the proposed ONoCs, the crossbar-based solutions gain considerable interest among the major players in the field. A dedicated place-and-route tool was proposed [18] to optimize layout with the main objective to reduce the power consumption. The efficient crossbar solutions rely on passive Microring Resonators (MRs), and do not require any arbitration [1][4] due to the dedicated point-to-point connections between IP cores. In these networks, the signal propagation relies on Wavelength Division Multiplexing (WDM). Comparing the proposed crossbars is a challenging task, since it requires considering both network characteristics and technological data, assuming layout constraints.

In this paper, we compare proposed crossbars according to their worst-case losses, which is a key metric to evaluate the ONoC scalability and power efficiency. The worst-case losses can be estimated by considering the losses in the network (e.g. induced by waveguide crossings and waveguide propagation) and the optical losses value (e.g. propagation loss).

The paper is structured as follows. Section II presents the considered architecture models, topologies and implementations. Section III presents the comparison methodology based on worst-case loss analysis. Section IV gives the comparison results and Section V concludes the paper.

II. ARCHITECTURE MODELS, TOPOLOGIES AND IMPLEMENTATIONS

In this section, we describe the considered ONoC architecture models, the associated topologies and implementations.

A. Architecture model overview

Figure 1 illustrates the considered 3D architecture model. It is composed of an electrical layer implementing *NxN* IP cores and an optical layer implementing ONoC. In our study, we assume *N* is an even number, but the work could be easily extended for odd values and for *NxM* IP cores architectures. The optical network in the optical layer is composed of on-chip laser sources [9], MRs, and photodetectors. The ONoC is connected to the IP cores by using Through Silicon Vias (TSV) [16]. Numerous ONoCs relying on WDM have been proposed. These networks can use wavelength routing scheme to propagate data from a source IP core to a destination IP core, thus leading to a contention-free network (without need for arbitration), with high throughput and low latency.

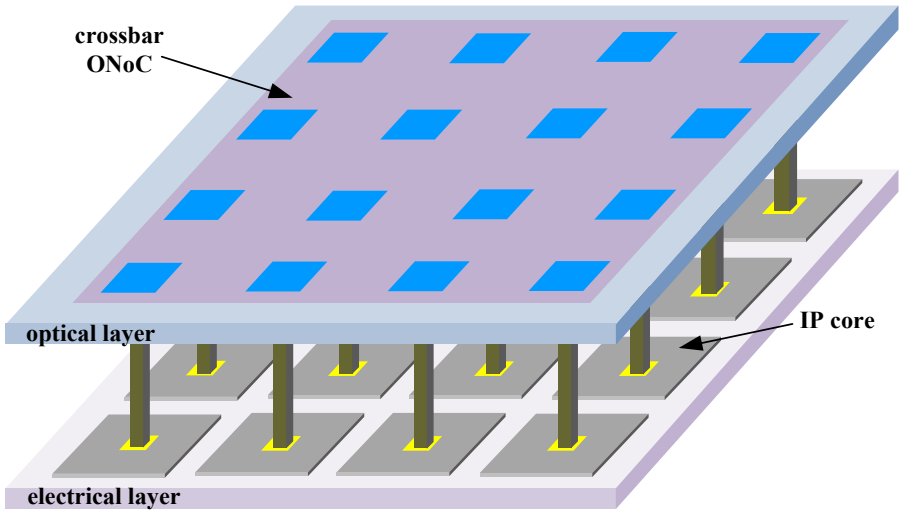


Figure 1: The crossbar ONoC is implemented in the optical layer and it interconnects IP cores located in the electrical layer

In this work, we compare ONoC architectures implementing crossbar functionality by considering the use of on-chip lasers, which allows direct modulation of the optical signals. Indeed, efficient on-chip lasers usually require the inclusion of III–V semiconductors: gallium arsenide (GaAs) or indium phosphide (InP) are currently considered to be the best options. Microlasers, based on microdisk structures coupling light evanescently from the cavity resonant mode to the guided mode in an adjacent silicon waveguide, are sufficiently compact as to be implemented in a large number and at any position. For a given wavelength, the size of an on-chip laser is of the same order of magnitude as the size of an MR used to modulate continuous waves emitted by off-chip lasers, which leads to a similar on-chip size for both approaches. While on-chip laser sources require the use of less mature technologies compared to their off-chip counterpart, they have the potential to provide the following three key advantages:

- Easier and more efficient integration by relaxing layout constraints: in case of on-chip lasers, it is not necessary to distribute the light from an external source to the modulators (e.g. through the, so called, *power waveguides* [7]). Relaxing such constraints contributes to reducing the number of waveguide crossings or even avoiding them altogether in the ring topology.
- Higher scalability by keeping the architecture fully distributed, which is not achievable by considering centralized off-chip lasers.
- Lower power by reducing the worst-case communication distance. This corresponds to the distance from the source IP to the destination IP for on-chip laser based architectures, while for off-chip laser based architecture, this distance also includes the distance from the off-chip laser to the source IP. Shorter distance, consequently, reduces the optical losses, and hence the minimum required laser output power. Moreover, the power consumption can be further improved by locally turning off the on-chip laser when no communication is required.

B. Passive ONoC architecture implementations

The crossbar network topologies considered in this study are 1) Matrix [17], 2) λ-router [1], 3) Snake [10], and 4) ORNoC [4], as shown in Figure 2. In the figure, each column is dedicated to a topology and the rows show their i) structural views, ii) implementation characteristics, and iii) layouts. This section briefly introduces these implementations and illustrates the way they can be used to interconnect 2x2 IP cores. We also illustrate the layout for 4x4 IP cores architecture, and evaluate the number of required optical devices, assuming N is an even number.

1) Matrix

The first crossbar topology relies on a traditional Matrix-like structure. Figure 2 1) illustrates a simple example where 4 IP cores are interconnected using the Matrix. Full connectivity is considered, which leads to a total of 16 MRs in this example. By considering only inter-IP core communications, one MR per line of the Matrix can be removed. Thus, for NxN IP cores, $(N^2\text{-}1)xN^2$ passive MRs are used to implement the crossbar itself. The transmitters are composed of on-chip laser sources, and the receivers are composed of photodetectors and passive MRs that drop the signal into the photodetector (not illustrated in the figure for sake of clarity). N_{MR}, $N_{MR,det}$, N_{laser} and N_{wl}, represent the number of MRs in the network itself after the reduction, the number of MRs in the receiver interface, the number of laser sources, and the number of wavelengths. Because we focus on crossbar networks, we assume dedicated communication between all the IP cores through spatial WDM. As a consequence, $(N^2\text{-}1)xN^2$ laser sources, with the same number of photodetectors and passive MRs, are required in the network interface. It is worth noticing that all topologies considered in this paper require the same number of laser sources and photodetectors. Matrix topology uses $N^2\text{-}1$ wavelengths to implement all the communications.

In this work, we consider that IP cores are uniformly placed on the chip. Figure 2 also presents two possible layouts: $layout_A$ and $layout_B$, which are designed to A) avoid any waveguide crossing between the network interfaces and the crossbar network itself and B) reduce the worst-case waveguide length between IP cores. For layout optimization purposes, the crossbar network is located in the middle of the optical layer, which also allows keeping a symmetrical structure (it is represented as a box for the sake of clarity in the figure). It interconnects 16 inputs (in red lines) with 16 outputs (in blue lines) through 240 MRs. For each IP core, we consider that optical signals are injected and received on opposite sides of the interface (i.e. with the transmitter and the receiver part being located on the different side of the interface). This allows a) to keep the layout regular, b) to avoid the use of extra waveguide crossing and c) to reduce the waveguide length. In order to match with the layout constraints from the regular NxN IP cores architecture, the transmitter and the receiver part of the IP cores are connected to the inputs and the outputs of the crossbar with waveguides, respectively. In this work, we consider only X and Y directions for the waveguides placing and routing, which simplifies the design rules, but can lead to extra waveguide length. The same layout rules will be assumed to interface with both λ-router and Snake networks.

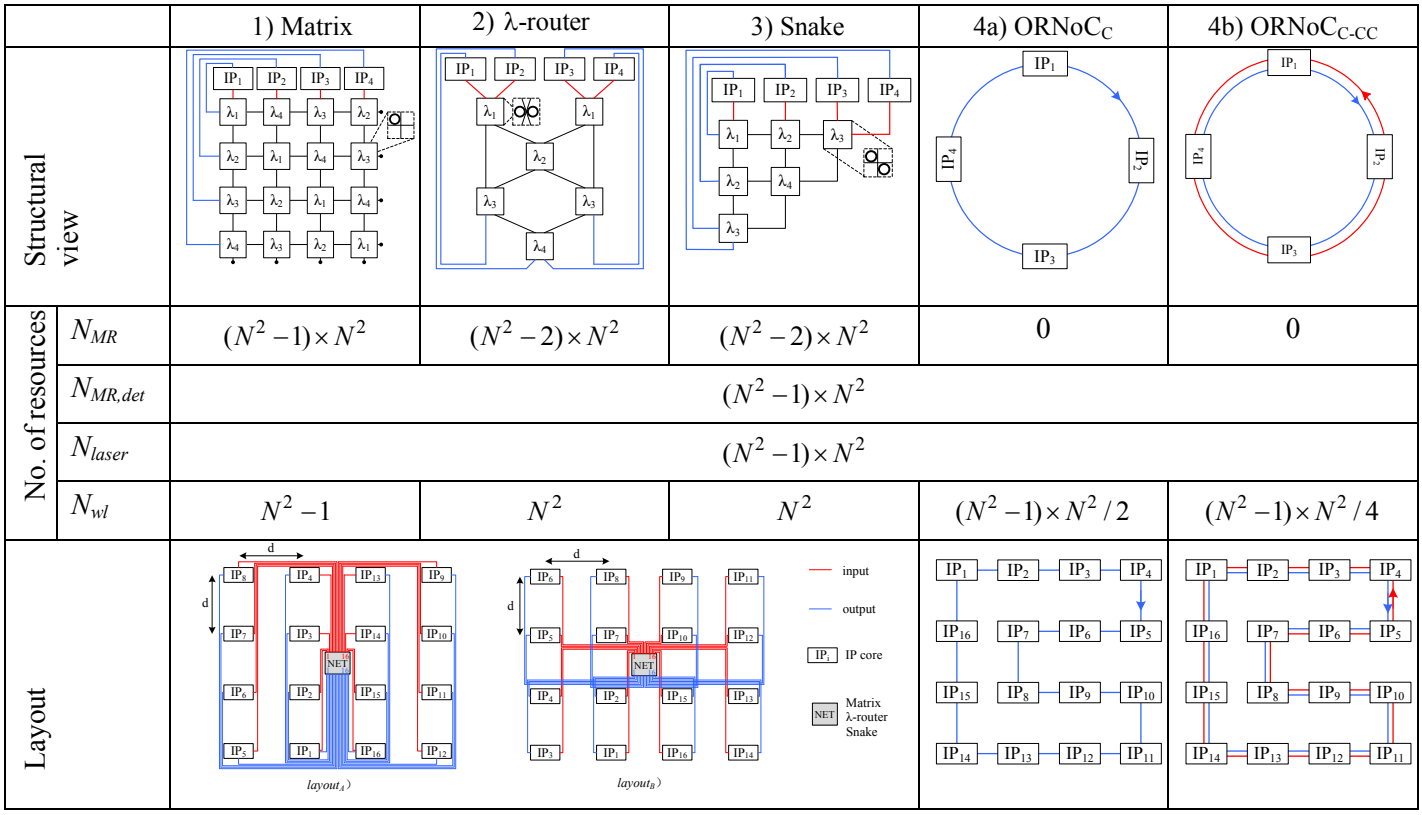


Figure 2: Summary of considered ONoCs: 1) Matrix, 2) λ-router, 3) Snake, 4a) ORNoC$_C$ and 4b) ORNoC$_{C\text{-}CC}$

2) *λ-router*

λ-router is a multistage network topology relying on WDM and wavelength routing to propagate optical signals from input to output ports. Compared to the Matrix, the multistage structure allows reducing the number of waveguide crossings in the worst-case scenario (6 and 3, for Matrix and λ-router, respectively, in the case of 4 IP cores). This is achieved by assuming symmetric 2x2-switch structure relying on 2 identical MRs. The initial structure of λ-router would assume 240 MRs, for the architecture with 4x4 IP cores, but a reduction method [1] reduces the network complexity by managing only the required optical connections and by removing the unused MRs. As a result, 224 MRs are required to implement the network.

3) *Snake*

Snake is, also, a multistage network topology. It has the same properties as the λ-router. The only difference is the distribution of the MRs in the network, which leads to the more compact layout compared to λ-router, with the side effect of different waveguide lengths between different input and output pairs. Similarly to λ-router, a reduction method adapted from [1] is applied to remove the unused MRs.

4) *ORNoC (Optical Ring Network-on-Chip)*

In ORNoC, the inter-IP core communication is realized through waveguides forming a ring. The following operations are performed:

- Injection: the IP core injects an optical signal into a waveguide through its output port. The wavelength of the signal specifies the destination of the IP core;
- Pass through: the incoming signal propagates along the waveguide (i.e. no MR with the same resonant wavelength is located along the waveguide);
- Ejection: the incoming optical signal is ejected from the waveguide and is redirected to the destination IP core. This is achieved by an MR located along the waveguide that has the same resonant wavelength as the signal.

Due to partitioning of the ring in ORNoC, the same wavelength can be reused to realize multiple communications in the same waveguide, at the same time, but on a different set of partitions of the ring. Furthermore, multiple waveguides can be used to interface IP cores. Both clockwise (C) and counter-clockwise (CC) directions can be considered for signal propagation, where each direction is realized in a separate waveguide. For this comparative study, we consider two versions of ORNoC: $ORNoC_C$ and $ORNoC_{C\text{-}CC}$ relying on only the C direction, and both C and CC directions, respectively. They are illustrated in Figure 2 4a) and 4b): blue and red lines represent C and CC directions separately. Compared to the other networks, no MR is used in the network itself (i.e. only passive MRs are used in the network interfaces), which leads to a reduction of the total number of optical devices. However, the ring structure implies crossing intermediate network interfaces, which leads to an increase of the minimum number of wavelengths to be used (6 and 3 wavelengths are required to interconnect 4 IP cores with $ORNoC_C$ and $ORNoC_{C\text{-}CC}$, respectively). If the maximum number of wavelengths in one waveguide is reached, then additional waveguides can be added in order to realize all the communications, without any impact on the layout complexity and the waveguide crossing, by considering a serpentine layout.

III. COMPARISON METHODOLOGY

The optical crossbars we consider are compared according to their worst-case losses. In this section, we first formulate the optical loss model for optical paths and then we detail the method to evaluate the worst-case loss.

A. Opitcal Loss Model

This section presents the proposed optical loss models for crossbar comparison. We assume on-chip laser sources for all topologies, which contribute to reduced number of waveguide crossings, compared to the off-chip laser counterpart.

In an optical crossbar, the total loss along an optical path L_{total} is given by the equation (1). L_{total} depends on: the total propagation loss in the waveguide $L_{waveguide}$, the total loss due to waveguide crossing $L_{crossing}$, the drop loss L_{drop} when a signal encounters a MR with the same wavelength, the total through loss $L_{through}$ when a signal passes by a non-resonant MR and the bending loss $L_{bending}$ of the waveguide, all given in dBs. In this work, we assume a negligible through loss and bending loss. Furthermore, we also assume negligible crosstalk between waveguides, which can be obtained by considering a 5μm distance between parallel waveguides. Indeed, for 5mm parallel waveguides assuming 500nm width and 220nm height,

the power coupling between the waveguides will be lower than -40dB, when the gap side by side is 3 μm or more [19]. The parameter description is given in Table 1 and Table 2.

$$L_{total}^{dB} = L_{waveguide}^{dB} + L_{crossing}^{dB} + L_{drop}^{dB} + L_{through}^{dB} + L_{bending}^{dB} \quad (1)$$

- $L_{waveguide}=P_{propagation}\times l_{s\text{-}d}$;
- $L_{crossing}=P_{crossing}\times N_{crossing}$;
- $L_{drop}=P_{drop}\times N_{drop}$;

Table 1 Insertion Loss Parameters

Parameter	Description
$P_{propagation}$ *(dB/cm)*	Intrinsic propagation loss in a waveguide
$P_{crossing}$ *(dB)*	Crossing loss
P_{drop} *(dB)*	Drop loss

Table 2 Network Implementation Characteristics

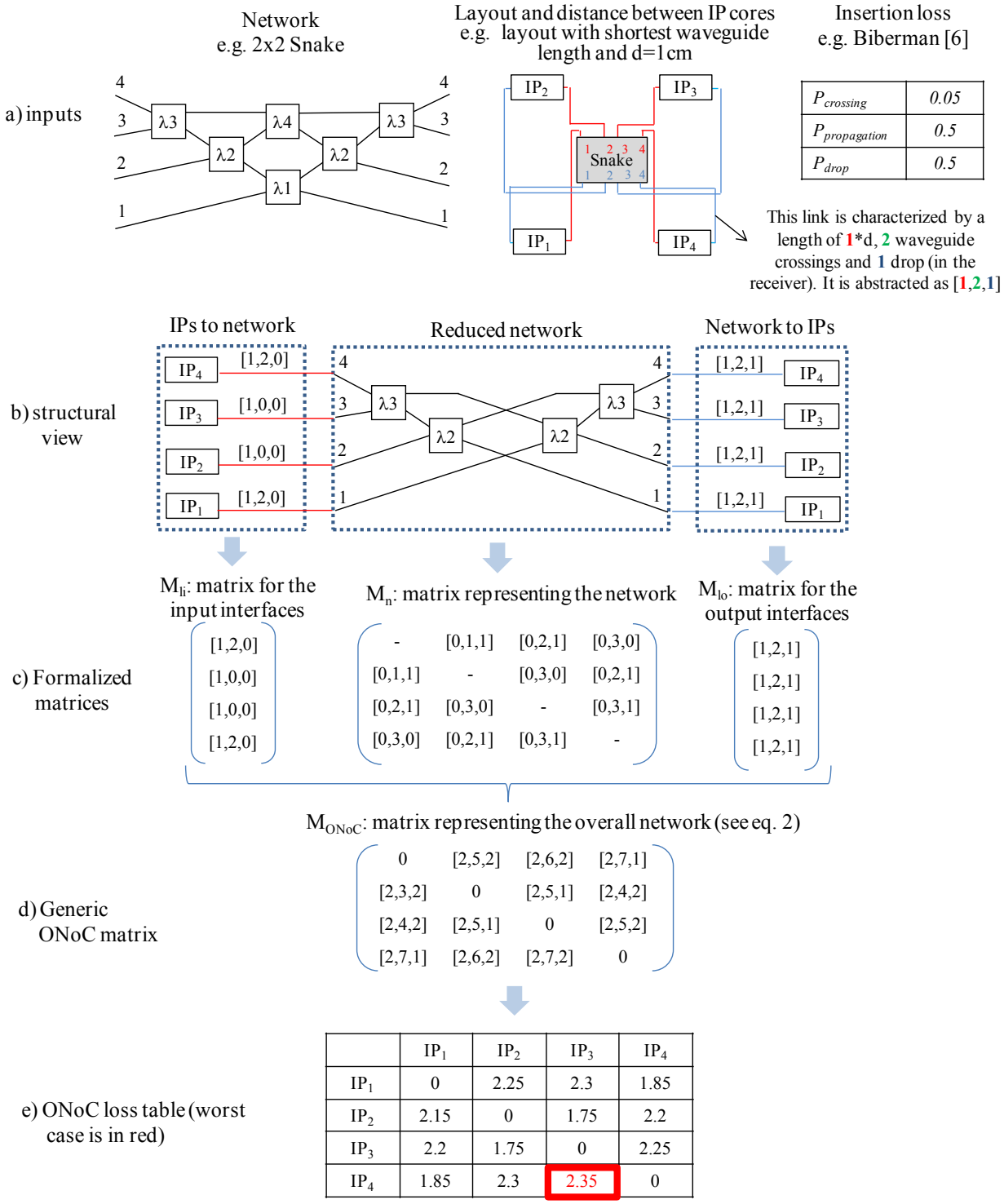
Parameter	Description
$l_{s\text{-}d}$	Waveguide length between a source and a destination
$N_{crossing}$	Number of waveguide crossings
N_{drop}	Number of drop operations

The waveguide length between the source and destination $l_{s\text{-}d}$ depends on the layout represented in Figure 2. The key metrics to evaluate $L_{waveguide}$, $L_{crossing}$, and L_{drop} are $l_{s\text{-}d}$, $N_{crossing}$, and N_{drop}. For Matrix, λ-router and Snake networks, both layouts (i.e. $layout_A$ and $layout_B$) are considered, which leads to: bigger longest waveguide length in layout A, and more additional waveguide crossings in layout B. We do not consider the distance between inputs and outputs of the network itself. For all aforementioned networks, two drop operations occur: one in the network itself, and the other one in the receiver interface to drop the signal into the photodetector.

Both $ORNoC_C$ and $ORNoC_{C\text{-}CC}$ do not suffer from any waveguide crossing and the signal is dropped only once in the receiver part ($N_{crossing}$=0 and N_{drop}=1). However, the considered serpentine layout implies that $l_{s\text{-}d}$ increases more rapidly when compared to the other networks. It is worth noticing that $l_{s\text{-}d}$ is significantly reduced for the C-CC case compared to the C case. This will result in a lower worst-case loss, which directly contributes to the energy-efficiency of $ORNoC_{C\text{-}CC}$, as detailed in the results section of the paper.

By applying the maximum function on the values of L_{total} for all the communications in the architecture, the worst-case loss in the optical path L_{wc} of each network is obtained, assuming *NxN* IP cores. L_{wc} is a key metric to measure the ONoC power efficiency, since lower loss leads to lower laser output power, i.e. reduced energy/bit for all communications. Considering both technological and structural values, which are related to the fabrication process and the network topology, enables a fair comparison of networks. The following section gives the worst-case loss analysis for the networks in details.

B. Worst-Case Loss Evaluation Methodology



	IP_1	IP_2	IP_3	IP_4
IP_1	0	2.25	2.3	1.85
IP_2	2.15	0	1.75	2.2
IP_3	2.2	1.75	0	2.25
IP_4	1.85	2.3	2.35	0

Figure 3: The worst-case evaluation methodology of ONoCs

The worst-case loss L_{wc} is evaluated following the methodology, illustrated in Figure 3, based on the loss model presented in the previous section. The inputs of the methodology (Figure 3 a)) are: i) the network topology, ii) the considered layout, and iii) the set of

insertion loss values. In the example illustrated in Figure 3, we consider a 2x2 Snake network topology, the layout assuming shortest waveguide length and insertion losses values from Biberman [6].

The structural view of the resulting implementation is represented in Figure 3 b). Each communication between IP cores is divided in three routing parts: i) routing from the transmitter part of IP cores to the input of the network (represented with red lines), ii) routing in the network itself, and iii) routing from the output of the network to the receiver part of IP cores (represented with blue lines). Each link between an IP core output/input and a network input/output is represented with red/blue lines, respectively. It is characterized according to a) the number of the length *d*, b) the number of waveguide crossings, and c) the number of the drop operations, represented as a triplet [a, b, c]. For instance, the link from IP_4 to input 4 of the network is characterized by [1, 2, 0], meaning that the length is d, there are 2 waveguide crossings, and no drop operation.

From this structural representation, three matrices are generated: M_{li}, M_n, and M_{lo} (Figure 3 c)). They correspond to the three above mentioned routing parts. Regarding the matrix characterizing the network, M_n, no values correspond to in the diagonal, since only inter-IP core communications are considered.

Based on M_{li}, M_n, and M_{lo}, a single matrix named M_{ONoC} is obtained according to equation (2), as illustrated by Figure 3 d). This matrix represents each IP core to IP core communication using a triplet consisting of the total distance propagated by signals, the number of waveguide crossings, and drop operations.

$$M_{ONoC} = \left(M_{li} \bullet \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} + M_n + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \bullet M_{lo}^{T} \right) \circ \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \quad (2)$$

From the given set of insertion loss values (Figure 3 a)) and M_{ONoC} representation, the loss for each pair of the IP core communication is obtained by applying the equation (1), as illustrated in Figure 3 e). Finally, the worst-case loss in the network is extracted by identifying the maximum value in the table (highlighted in red in the figure). The mean value of the communications gives the average loss (L_{avg}).

This methodology is used to evaluate the losses in each ONoC identified in Figure 2, and is generic enough to address any *NxN* network size.

IV. COMPARATIVE STUDY

We compare the topologies and corresponding layouts according to the worst-case and average losses. In a first comparison, all the related networks with two layouts ($Matrix_A$, $Matrix_B$, λ-$router_A$, λ-$router_B$, $Snake_A$, $Snake_B$) and ORNoC ($ORNoC_C$ and $ORNoC_{C\text{-}CC}$) are considered by assuming a given set of technological values extracted from Table 3. In a second comparison, we further compare the networks assuming various design parameters.

Table 3: Insertion Loss Parameters

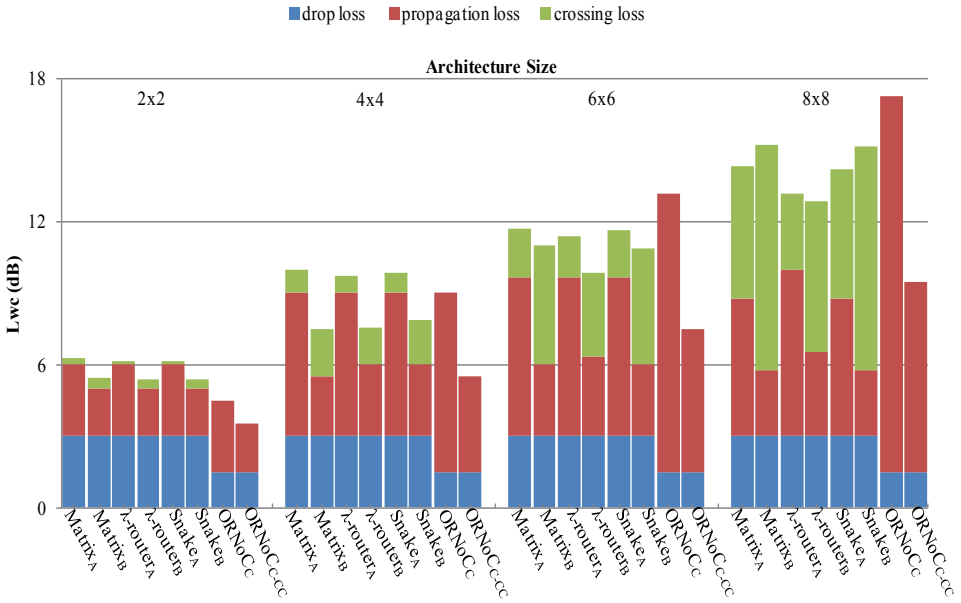
Optical loss	$P_{crossing}$ (dB)	$P_{propagation}$ (dB/cm)	P_{drop} (dB)
Pan[12]	0.05	1	1.5
Kirman[13]	0.12	1	1
Biberman[6]	0.05	0.5	0.5
Koka[14]	0.2	0.1	1.5

A. Worst-case and average losses evaluation

We first assume a fixed $4cm^2$ die size, and evaluate the losses for different architecture sizes: 2x2, 4x4, 6x6 and 8x8, where the distance between IP cores decreases as the number of IP cores increases, i.e., d=10mm, 5mm, 3.33mm and 2.5mm, respectively. Figure 4 a) and b) illustrate the evaluation results for worst-case losses with the parameter values given by Pan [12] and Biberman [6], respectively.

We compare different layouts for Matrix, λ-router and Snake topologies using the values from Pan, as shown in Figure 4 a), where, for example $Matrix_A$ refers to Matrix topology in $layout_A$. It can be seen that the $layout_B$ outperforms $layout_A$ for 2x2 to 6x6 network sizes, but $layout_A$ shows better scalability, since the loss, especially the propagation loss, is less influenced by the increase of the IP number. For an architecture of 8x8 IP cores, $layout_A$ exhibits lower losses for Matrix and Snake. By considering values from Biberman (Figure 4 b)), the same conclusion can be made for architectures containing up to 4x4 IP cores. However, for architectures with 6x6 and 8x8 IP cores, worst-case loss is lower for $layout_A$, due to the lower propagation loss in the waveguide, thus again highlighting the better scalability of this layout. From the comparison results in Figure 4 a) and b), $ORNoC_{C\text{-}CC}$ is the most scalable network despite the long waveguide length introduced by the serpentine layout. By considering values from Biberman, $ORNoC_{C\text{-}CC}$ is the most scalable network, with 4.5dB in the worst-case path for the architecture with 8x8 IP cores, followed by $\lambda\text{-router}_A$ with 7.65dB, thus achieving 41.2% improvement compared to $\lambda\text{-router}_A$. By assuming parameters from Koka (not shown in the results), the worst-case loss in $ORNoC_{C\text{-}CC}$ and $\lambda\text{-router}_A$ become 2.3dB and 16.3dB, respectively, thus leading to a 85.9% improvement for $ORNoC_{C\text{-}CC}$ over $\lambda\text{-router}_A$. Because of the rather large distance implied by the die size considered here, $ORNoC_C$ does not exhibit as good scalability as $ORNoC_{C\text{-}CC}$.

a)



b)

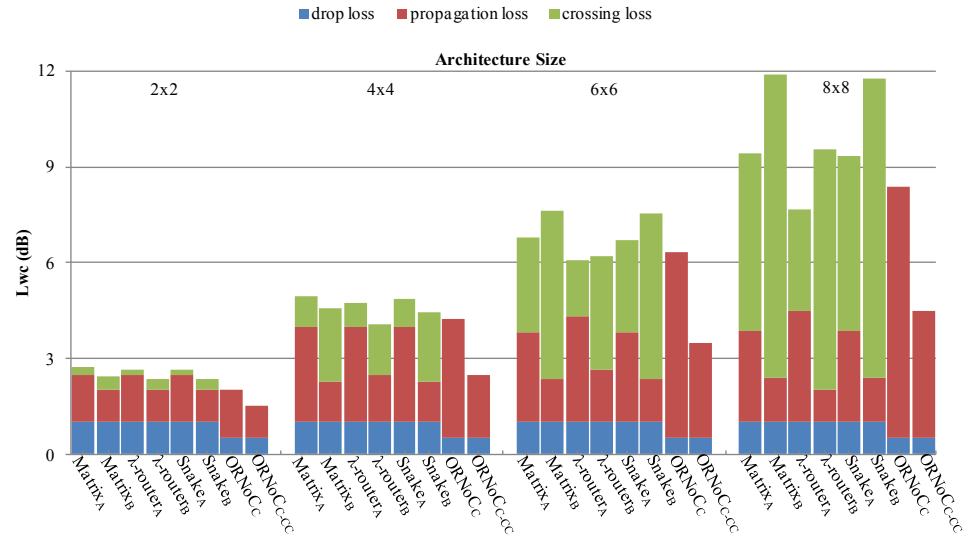
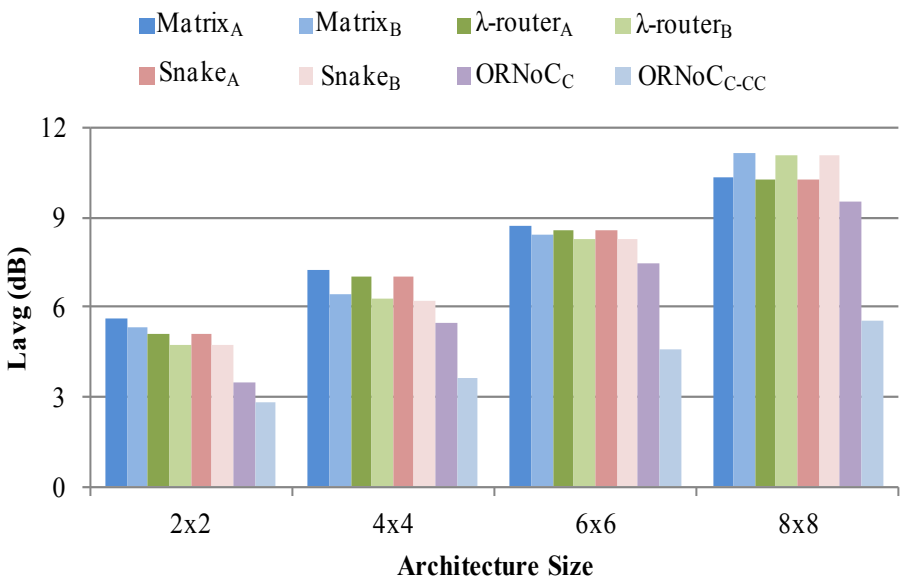


Figure 4: Worst-case losses evaluation for various number of IP cores assuming insertion loss parameters from a) Pan [12] and b) Biberman [6].

Figure 5 shows results of the average loss comparison. Compared with all the implementations of Matrix, λ-router and Snake (i.e. using both layouts), $ORNoC_{C\text{-}CC}$ demonstrates, on average, 46% and 56% reduction of the losses by considering Pan and Biberman values, respectively. This is achieved by avoiding the waveguide crossings in the serpentine layout, and by reducing the signal propagation distance thanks to the use of both C and CC directions. For 8x8 IP cores and by considering the value set from Biberman, the improvement over $Snake_B$ reaches 69% ($ORNoC_{C\text{-}CC}$ and $Snake_B$ demonstrate 2.5dB and 8dB, respectively). Moreover, thanks to this reduction in the average loss, additional significant power reduction can be achieved by considering the use of tunable lasers output power [3].

a)



b)

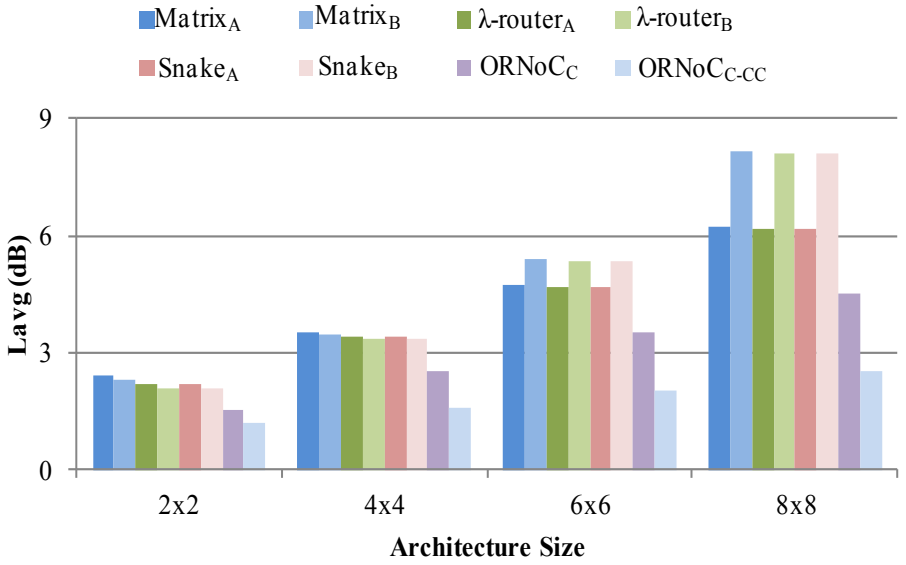
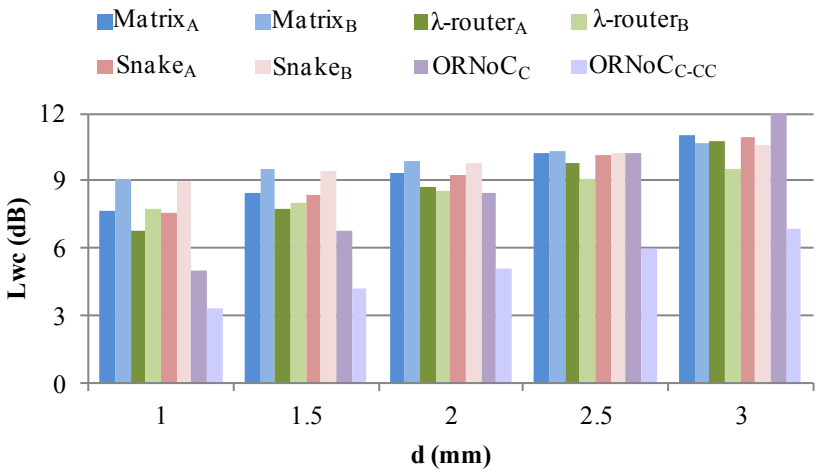


Figure 5 Average losses evaluation for various number of IP cores assuming insertion loss parameters from a) Pan [12] and b) Biberman [6]

Figure 6 represents the worst-case loss for a fixed size of 6x6 IP cores, and various distances between them (d=1, 1.5, 2, 2.5 and 3mm) by assuming insertion loss parameters given by Pan [12] and Biberman [6], separately. The impact of the distance increase is higher for the networks relying on $layout_A$ than $layout_B$. The relative increase of the loss with the distance is the greatest for ORNoC$_C$ and ORNoC$_{C\text{-}CC}$, due to the serpentine layout. Still, even for a 3mm distance, which implies a realistic 3.24cm² die size, ORNoC$_{C\text{-}CC}$ remains the most power-efficient network for both sets of the insertion loss parameters.

a)



b)

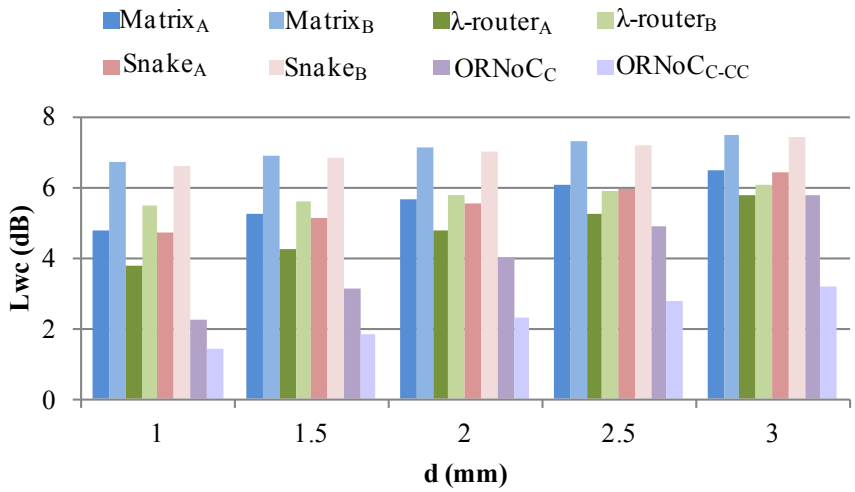


Figure 6: Evaluation of the impact of the distance between 6x6 IP cores on the worst-case losses assuming insertion loss parameters from a) Pan [12] and b) Biberman [6]

B. Implementation comparison

In order to further explore the design space, we observe the example of a set of architectures with 6x6 IP cores, and various distances d (in the range of 1mm-3mm with 0.5mm increments). We consider a range of 0-2dB for propagation loss and a range of 0-0.15dB for waveguide crossing loss. Figure 7 illustrates comparison results for the implementations of λ-router with $layout_A$ (i.e. without waveguide crossing) and $layout_B$ (with shorter waveguide length), by assuming P_{drop}=1dB. We also plot the values for ($P_{crossing}$, $P_{propagation}$) from Table 3. The area below each line represents the design space for which the worst-case loss is lower for $layout_A$; the area above the line gives the design space where the worst-case loss is lower for $layout_B$, and the line itself represents the designs with the same worst-case losses for both layouts. For example, $layout_A$ shows lower worst-case loss for the values from Pan [12], Biberman [6], Kirman [13] when the distance is smaller than or equal to 1.5 mm. In addition, the result indicates that $layout_A$ performs better as the distance gets

smaller, since the loss is less impacted by the waveguide propagation. Overall, this further helps to determine the most appropriate layout for a given set of insertion loss values and a given distance between IP cores.

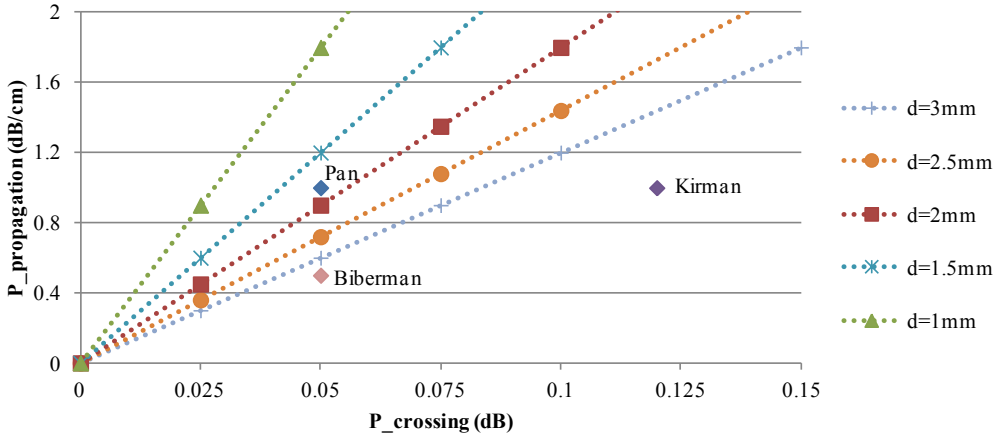


Figure 7: Comparison of λ-router$_A$ and λ-router$_B$ (6x6 IP cores, P_{drop}=1dB)

These comparisons highlight the importance of technological parameters, layout and network topology to evaluate the worst-case optical loss. We can see that for a given set of technological values (e.g. crossing loss and propagation loss), certain topology and layout may be more advantageous, which may significantly impact the overall power efficiency of the crossbar.

C. Designing complexity comparison

For a fair comparison of the considered optical crossbars, we also evaluate their implementation complexity while interconnecting N^2 IP cores. We compare implementation complexity using the numbers of the required MRs and waveguides.

The number of MRs is an important metric since it affects the network scalability. In general, with increase in the number of the IP cores, the number of the MRs also increases. It is worth noticing that all the networks require the same number of the lasers and photodetectors, and that one MR per photodetector is required in the interface. As a consequence, we only evaluate the number of MRs inside the network itself. The implementation of Matrix network requires $(N^2-1)xN^2$ MRs, which is slightly higher than the number required by λ-router and Snake ($(N^2-2)xN^2$). ORNoC does not require any MRs in the network, which leads to a lower design complexity, and, therefore, better scalability.

In ORNoC, the number of waveguides required for the implementation depends on the maximum number of wavelengths per waveguide. Indeed, ORNoC$_C$ (respectively ORNoC$_{C-CC}$) requires $(N^2-1)xN^2/2$ (respectively $(N^2-1)xN^2/4$) waveguides, assuming a single wavelength per waveguide. By considering N^2 possible wavelengths per waveguide (i.e. the number of wavelengths required by both λ-router and Snake), the number of the required waveguides for the implementation of ORNoC$_C$ (respectively ORNoC$_{C-CC}$) is reduced to $(N^2-1)/2$ (resp. $(N^2-1)/4$)). Hence, for the same number of wavelengths, both implementations of

ORNoC require fewer waveguides compared to Matrix ($2N^2$), λ-router (N^2) and Snake (N^2). For example, for a 8x8 architecture with a single waveguide per direction, $ORNoC_{C\text{-}CC}$ requires 1008 wavelengths in one waveguide compared to 63 wavelengths for Matrix and 64 for Snake and λ-router. 1008 is not a realistic value for the number of wavelengths. Following the methodology from [4], ORNoC would require 16 waveguides if we consider a maximum number of 64 wavelengths per waveguide, while Matrix, λ-router and Snake, would require 128, 64 and 64 waveguides, respectively. This also contributes to the lower design complexity and the better scalability. In addition, when the network size increases, there is a need for increasing number of the waveguides and wavelengths for implementation of optical crossbars. If a constraint such as the maximum number of wavelengths per waveguide (with optimistic value of 64) must be respected for all the considered networks, Matrix, λ-router and Snake could satisfy the constraint by considering the use of multiple networks, which implies additional waveguide crossings [15]. However, it is important to notice that additional waveguides can be used in $ORNoC_{C\text{-}CC}$ to satisfy the given constraint, and this can be achieved without any waveguide crossing because of the 3D architecture and the use of on-chip laser sources. Moreover, the layout of $ORNoC_{C\text{-}CC}$ is regular, and its $l_{s\text{-}d}$ is reduced compared to $ORNoC_C$, which makes the network implicitly scalable without need for a custom place-and-route tool such as in [10][5][18].

V. CONCLUSION

Optical crossbars on chip represent an efficient interconnect solution for many-core architectures. Various crossbar implementations have been proposed, and their worst-case losses depend on topological, physical and technological aspects. In this paper, we compare possible crossbar implementations relying on matrix, multistage and ring-based network topologies. For a given number of IP cores, and a given die size, ring-based networks facilitate the implementation characterized by the lower worst-case optical losses, yielding the most power-efficient solution. For the explored design space, ring-based topology implementations exhibit higher power efficiency compared to matrix-based and multistage-based network implementations. The approach of power loss analysis was applied to passive and fully interconnected networks, but it can be extended to active networks requiring resource allocation mechanism. We will focus on this aspect in our future work.

ACKNOWLEDGMENT

Sébastien Le Beux is supported by a Région Rhône-Alpes grant.

REFERENCES

[1] I. O'Connor, et al. "Reduction Methods for Adapting Optical Network on Chip Topologies to Specific Routing Applications". In Proceedings of DCIS, November 2008.

[2] A. Shacham, K. Bergman, L.P. Carloni, "Photonic Networks-on-Chip for Future Generations of Chip Multi-Processors," IEEE Transactions on Computers 57 (9), pp. 1246-1260, 2008.

[3] Y. Ye et al. "A Torus-based Hierarchical Optical-Electronic Network-on-Chip for Multiprocessor System-on-Chip", ACM Journal on Emerging Technologies in Computing Systems, 2012.

[4] S. Le Beux, et al. “Layout Guidelines for 3D Architectures including Optical Ring Network-on-Chip (ORNoC)”. In 19th IFIP/IEEE VLSI-SOC International Conference, 2011.

[5] L. Ramini, D. Bertozzi, and L. P. Carloni. “Engineering a Bandwidth-Scalable Optical Layer for a 3D Multi-Core Processor with Awareness of Layout Constraints”. Proceedings of the Third International Symposium on Networks-on-Chip (NOCS), 2012.

[6] A. Biberman, K. Preston, G. Hendry, N. Sherwood-Droz, J. Chan, J. S. Levy, M. Lipson, K. Bergman. “Photonic Network-on-Chip Architectures Using Multilayer Deposited Silicon Materials for High-Performance Chip Multiprocessors”, ACM Journal on Emerging Technologies in Computing Systems 7 (2) 7:1-7:25, 2011.

[7] D. Vantrease, et al. Corona: System Implications of Emerging Nanophotonic Technology. In Proceedings of the 35th Annual International Symposium on Computer Architecture (ISCA) pages 153–164, 2008.

[8] J. Psota, et al. ATAC: Improving Performance and Programmability With on-Chip Optical Networks. In Proceedings of IEEE International Symposium on Circuits and Systems, ISCAS, pages 3325–3328, 2010.

[9] J. Van Campenhout et al., “A compact SOI-integrated multiwavelength laser source based on cascaded InP microdisks,” IEEE Photon. Technol. Lett., vol. 20, no. 16, pp. 1345–1347, 2008.

[10] Luca Ramini, Paolo Grani, Sandro Bartolini, and Davide Bertozzi. Contrasting wavelength-routed optical NoC topologies for power-efficient 3D-stacked multicore processors using physical-layer analysis. In Proceedings of the Conference on Design, Automation and Test in Europe (DATE), 2013.

[11] R. Ho, K.W. Mai, and M.A. Horowitz. The Future of Wires. Proceedings of the IEEE, 89(4):490–504, April 2001.

[12] Y. Pan, J. Kim, and G. Memik. “FlexiShare: Channel Sharing for an Energy-Efficient Nanophotonic Crossbar”. In IEEE 16th International Symposium on HPCA, 2010.

[13] N. Kirman and José F. Martinez. “A power-efficient all-optical on-chip interconnect using wavelength-based oblivious routing”. in Proceedings of the ASPLOS, 2010.

[14] P. Koka, M.O. McCracken, H. Schwetman, C.-H.O. Chen, X. Zhang, R. Ho, K. Rai, and A.V. Krishnamoorthy. “A micro-architectural analysis of switched photonic multi-chip interconnects”. In 39th Annual International Symposium on Computer Architecture, 2012.

[15] S. Le Beux, J. Trajkovic, I. O’Connor, G. Nicolescu, G. Bois and P. Paulin. Multi-Optical Network on Chip for Large Scale MPSoC. In IEEE Embedded Systems Letters, Vol. 2, Issue 3, Pages 77-80, Sept. 2010.

[16] Igor Loi, Federico Angiolini, and Luca Benini. Supporting Vertical Links for 3D Networks-on-Chip: Toward an Automated Design and Analysis Flow. In Proceedings of the 2nd international conference on Nano-Networks, Nano-Net, pages 1–5, 2007.

[17] A. Bianco, D. Cuda, M. Garrich, G. G. Castillo, P. Giaccone. “Optical Interconnection Networks based on Microring Resonators”. In Proceedings of IEEE International Conference on Communications, 2010.

[18] A. Boos, L. Ramini, U. Schlichtmann, and D Bertozzi. “PROTON: An Automatic Place-and-Route Tool for Optical Networks-on-Chip”. ICCAD, pp. 138-145, 2013.

[19] V. Donzella, S. T. Fard and L. Chrostowski. “Study of waveguide crosstalk in silicon photonics integrated circuits”. in Proc. of SPIE 8915, Photonics North 2013, 89150Z.