Contrasting Laser Power Requirements of Wavelength-Routed Optical NoC Topologies Subject to the Floorplanning, Placement, and Routing Constraints of a 3-D-Stacked System

Marta Ortín-Obón, Mahdi Tala, Luca Ramini, Víctor Viñals-Yufera, and Davide Bertozzi

Abstract—A realistic assessment of optical networks-on-chip (ONoCs) can be performed only in the context of a comprehensive floorplanning strategy for the system as a whole, especially when the 3-D stacking of electronic and optical layers is implemented. This paper fosters layout-aware ONoC design by developing a physical mapping methodology for wavelength-routed ONoC topologies subject to the floorplanning, placement, and routing constraints that arise in a 3-D-stacked environment. As a result, this paper is able to compare the power efficiency and signal-to-noise ratio of ring-based versus filter-based wavelength-routed topologies as determined by their physical design flexibility.

Index Terms—Multiprocessor interconnection networks, optical fiber communication, wavelength routing.

I. INTRODUCTION

SILICON photonics is gaining momentum as the most promising emerging technology to deliver chip-level connectivity in the future large-scale many-core systems [1]. A major source of overhead of optical networks-on-chip (ONoCs) comes from static power, especially due to laser power and thermal tuning of optical devices. This cost is highly sensitive to ONoC static design choices, such as the connectivity pattern and the communication protocol on top of it. However, an additional yet significant contribution to static power arises during the physical mapping of the logical topology onto the chip floorplan. At that time, the actual distances that waveguides have to span over the chip surface are defined. However, other effects come into play. First, during the routing phase, unexpected waveguide crossings arise, which were not there in the logical scheme of the topology [25]. Second, floorplanning obstructions may not enable the monolithic placement of the topology as a whole, typically in the middle of the chip [29]. Therefore, the logical scheme might have to be deconstructed into physical partitions. In general, partitioned topology layouts are subject to higher optical power losses along their optical paths due to the increased wiring intricacy and larger spatial extension.

The above considerations are especially important for wavelength-routed topologies [5], which exhibit highly regular logical schemes built around geometrical patterns [13]–[15], and typically make strong assumptions on the location of the end nodes. Therefore, the physical mapping might cause a significant deviation between the logical scheme and the physical implementation (and between their quality metrics) as an effect of floorplanning, placement, and routing constraints.

Ring topologies are an apparent exception to this, because of their overly simple layout structure. However, they also suffer from the design predictability concern. In fact, optical rings can easily be configured in terms of the number of parallel ring waveguides they instantiate [17]. The laser power distribution network (PDN) needs to reach all of the waveguides, thus giving rise to larger and more lossy ring interfaces that only a layout-aware analysis can disclose.

Ultimately, the predictability gap between the insertion loss as estimated from the topology logical scheme versus from its physical implementation results in an increase of laser power requirements as the topologies go through their physical mapping process.

This paper takes on the challenge of assessing the static power efficiency of wavelength-routed ONoC (WRONoC) topologies using a comprehensive floorplanning strategy in a 3-D-stacked environment. Visibility of physical design steps, such as floorplanning, placement, and routing, enables this paper to shed light on the following novel aspects of ONoC topology evaluation.

A. Laser Power Requirements of Topologies Are Associated Not Only With the Properties of Their Logical Schemes But Also With Their Physical Design Flexibility

In principle, increasing the chip area relaxes the floorplanning constraints while increasing the propagation distance. This paper thus captures the nonintuitive relation between floorplan area and laser power requirements.

B. Physical Mapping of ONoCs Is Strictly Topology Specific

This paper thus defines a cross-layer synthesis methodology which yields visibility of the floorplanning implications of

ring topology configurations (e.g., degree of spatial multiplexing), or which searches for the most layout-friendly geometrical pattern inscribed into a filter-based topology for the sake of its physical partitioning.

While pursuing the above strategic goals, this paper aims at achieving the following more specific technical contributions.

1) We improve upon existing synthesis technologies for wavelength-routed ring topologies [17], [32], by spanning a better tradeoff between wavelength- and space-division multiplexing, which results in a lower static power and a more efficient exploitation of the available die space.

2) We bridge an existing gap in the literature between monolithic placement of topologies in the middle of the die [29], and its opposite solution, namely, their distributed (and automatic) place&route (P&R) at the granularity of each photonic switching element (PSE) [26]. We demonstrate the benefits of P&R at an intermediate granularity (i.e., partitions) for a more predictable physical mapping.

3) We analyze different partitioning patterns of the most relevant filter-based wavelength-routed topologies, in an attempt to preserve their geometrical properties while unfolding them to fit tight space constraints.

4) We account for the PDN in the topology comparison framework, namely for its structure, for its intersections with the layout of the main optical NoC, and for the equalization of laser power requirements across groups of optical paths.

5) We derive layout guidelines for the design of wavelength-routed topologies in a 3-D-stacked environment, accounting for the interdependences between stacked layers.

6) We point out the counterintuitive competitiveness of filter-based topologies with respect to ring-based ones, thus questioning the convergence of floorplanning-agnostic/oversimplified literature on ringlike structures for wavelength routing.

The definition of a systematic floorplanning strategy for topology comparison was achieved by setting up a complete physical mapping flow. While the merit of such a flow consists of identifying the methodological synthesis gaps at each abstraction layer, their horizontal and vertical interdependences, and their coherent vertical integration, the goal was not to bridge all of the design automation gaps. Rather, the goal was to identify such gaps, thus laying the groundwork for future evolution of design automation beyond its electronic roots.

II. Related Work

The significant amount of work on ONoC topologies, protocols, and architectures has fostered cross-layer methodologies for designing new optical networks [18], [28], [34], although the discipline is still admittedly in the early stage.

This has raised the interest in floorplanning and P&R approaches for ONoCs, since they greatly impact the performance and energy efficiency of the overall many-core system. The onset of additional waveguide crossings during P&R

of topologies has been pointed out in [25], and considered for topology comparison in [14]. Architecture and layout adaptation to the requirements of a 3-D-stacked architecture is another related research field [19], [24], [31]. The optical layer routing problem is formulated and solved in [23] and [24], so to minimize the optical loss in the ONoC given a fixed netlist. Other frameworks extend the scope from routing to the complete placement and routing process, either viewing these steps sequentially [18], or implementing some form of cross-layer optimization [22], [26]. Some frameworks augment P&R algorithms with thermal profile awareness [22]. A more comprehensive approach in [23] considers even scheduling policy, thermal tuning, and heterogeneity in chip power profiles. The above works suffer from one or more of the following issues.

1) They focus on optical routing geometries, thus failing to capture the topologic level [20], [21].

2) They tackle the P&R challenge mainly for physically distributed [27] or ring-equivalent structures [23], while centralized topologies are trivially placed in the middle of the die [29]. Therefore, the problem of unfolding a topology into physical partitions has never been thoroughly addressed before. Moreover, the side effects of the PDN on laser power requirements are never brought to the forefront.

3) P&R tools for generic optical topologies are typically not instructed to recognize geometrical properties to be exploited for better physical mapping [22], [26].

4) The focus for optical ring design is either on the routing pattern in a 3-D stacked setting [31], or on the efficient reuse of wavelength channels across ring waveguides [17]. Instead, hub design issues are typically overlooked.

In the context of WRONoCs, this paper aims at comprehensively bringing topologies from their logical scheme down to their layout planning. While complementing previous work at lower abstraction layers, this paper bridges the above existing gaps in the placement and routing of optical topologies subject to the (potentially tight) floorplanning constraints of a 3-D-stacked environment.

III. Wavelength-Routed ONoCs

In WRONoCs, all initiators can potentially communicate with all targets at the same time without any conflict, i.e., there is no need for arbitration to solve contention on shared resources. The underlying principle is that each initiator uses a different wavelength to reach each target, and each target receives packets from the different initiators on different wavelengths. Clearly, WRONoCs are a static power-sensitive technology, since scaling the system size comes at the cost of a proliferation of laser sources. Nonetheless, it has been demonstrated that for small-to-medium network sizes (up to 16 nodes), WRONoCs are even more power efficient than arbitrated nanophotonic crossbars [7], since they come with no arbitration overhead. Therefore, a 16×16 WRONoC is the target of this paper. This paper focuses on the physical mapping process of the two main categories of WRONoC topologies: filter-based (FbONoC) versus ring-based (ORing) Topologies.
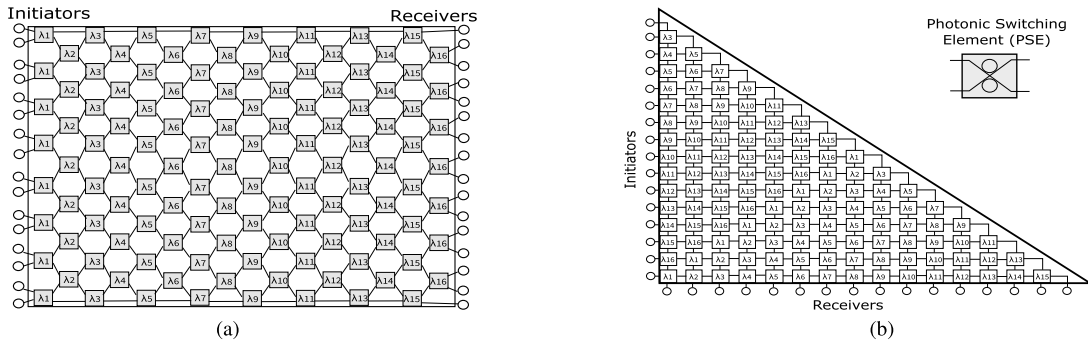
Fig. 1. Logical schemes of filter-based WRONoC topologies. (a) Lambda router. (b) Snake.

A. Filter-Based Topologies

Two representative FbONoC topologies have been selected from the literature, namely, the lambda router [13] and the snake [14] (logical schemes in Fig. 1), because they stand out as the most power efficient solutions from existing design space explorations [14], [29]. For a given connectivity requirement, both solutions are composed of the same number of 2×2 add-drop optical filters (also named PSEs), although tuned to different resonant wavelengths. The key difference lies in the geometrical properties of the topologies. The lambda router connects N nodes by means of N stages of alternately $N/2$ and $(N/2) - 1$ add-drop filters, and is fundamentally built around a diamond shape. Instead, the snake is shaped as a right-angled triangle, with $N-1$ add-drop filters on each side, and exhibits highly heterogeneous source-to-destination paths.

FbONoCs are typically laid out by assuming their monolithic placement in the middle of the optical layer [32]. If there is not enough spacing, the only alternative option explored in the literature consists of using ONoC-specific P&R tools (such as PROTON [29]) that operate at the granularity of the individual PSE. These tools are useful for the general case, when no geometrical properties can be easily identified in the topology. However, they leave the regularity of the logical scheme unexploited, thus potentially ending up in suboptimal physical designs. This paper pursues an intermediate approach between monolithic and fine-grained physical mapping: physical partitioning of the logical topology to fit the available interhub spacing.

B. Optical Ring

In contrast, an ORing has a straightforward routing pattern on the physical layout, at the cost of a more intricate hub design. In fact, two important ORing parameters are the number of ring waveguides [i.e., the spatial-division multiplexing (SDM) degree] and the number of wavelengths that are multiplexed on each waveguide [i.e., the wavelength-division multiplexing (WDM) degree]. On the one hand, reusing the same wavelength channels across different waveguides enables to reduce the required number of laser sources to meet preassigned connectivity requirements. On the other hand, increasing the number of ring waveguides for a more extensive reuse of wavelength channels gives rise to both an increased number of waveguide crossings inside the hubs (in order to bring the optical power to the innermost waveguides) and to an increased extension of the hubs themselves. While the former

effect results in larger static power overhead to compensate for the optical power waste, the latter may cause the designer to fail meeting the assigned area constraints.

So far, ring synthesis algorithms have been proposed to infer the right combination of SDM and WDM degree. The pioneer algorithm in [17] aims at minimizing the number of instantiated ring waveguides, but does not have visibility of the physical implementation tradeoffs. Moreover, it unnecessarily ends up using long paths on specific waveguides while underutilizing the remaining ones. In this paper, we aim at the generation of more static-power efficient ring configurations based on physical- and layout-layer analysis.

IV. TARGET 3-D ARCHITECTURE

Our target 3-D architecture (Fig. 2) is composed of an electronic layer and an optical layer vertically stacked on top of it. The latter is powered by an array of off-chip continuous-wave laser sources providing multiple optical carriers.[1] We augment the accuracy of the laser power subsystem model by accounting for the (PDN) to bring the optical carriers to all of the hubs for the modulation phase. The PDN is implemented as a binary tree: signals from the off-chip laser sources are multiplexed onto an input waveguide acting as the PDN root, while every hub is a leaf. Branches are implemented by means of splitting devices of the laser power, which can be implemented through Y-junctions, directional couplers of MMI devices.

We assume that the electronic layer is structured into 16 clusters, each one having its own gateway to the optical layer (and an associated hub on top of it), which delivers intercluster communications. Each hub is both initiator and target on the resulting 16×16 ONoC. Cluster sizes of at least 16 computing cores are not uncommon [3], which means that our target architecture can easily connect at least 256 cores. Gateways are positioned in the middle of each cluster on the electronic plane, with the corresponding hubs vertically aligned on top of them, forming a gridlike structure (Fig. 2).

Our baseline assumption is that most of the electronic circuits reside on the electronic plane (serializers and drivers

---

[1]The authors are aware of the debate about the use of off-chip versus on-chip laser sources [33], but make a conservative choice for off-chip ones. In our view, off-chip laser sources are likely to be a more practical solution for early prototypes in the near future, since they can be considered as external optical power supplies, featuring easy replacement, and temperature stability. Also, they do not contribute to the chip power budget, but to the system one, which gives more flexibility.
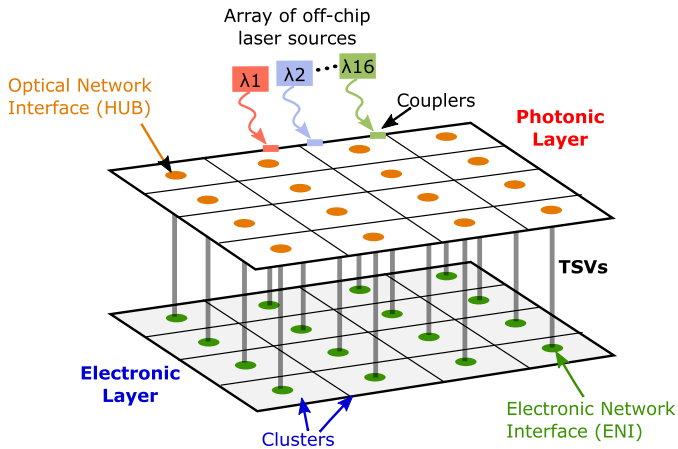
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4
IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS

Fig. 2.   Experimental setting: 3-D-stacked multicore processor.

Fig. 3.   Hub design for filter-based topologies in a 16 × 16 WRONoC.

Fig. 4.   Hub design for a two-waveguide ring topology in a 16 × 16 WRONoC.

in transmission interfaces, transimpedance amplifiers, digital comparators, and deserializers in the receiver interface), while the optical layer hosts modulators, photodetectors, and low-speed analog and digital circuitry for automatic wavelength tuning and thermal stabilization of microring resonators [35]. In practice, gateways implement the electronic network interface (ENI), while hubs implement the optical network interface (ONI) of an ONoC.

For the complete ENI architecture, the interested reader is referred to [36]. We obtained the area of the electronic components in that ENI by synthesizing all the blocks presented in [36] with a 40-nm industrial technology library, and obtained about 0.15 mm^2.

A. FbONoC Hub

Hubs for FbONoC topologies can be inferred in a regular and efficient way, as shown in Fig. 3 for the target 16×16 system. Interestingly, lambda router and snake differ only in the connectivity pattern, not in the hub design.

The transmission side accommodates an array of 15 ring modulators (self-communication is not allowed), each one serving a wavelength channel associated with a different destination. Modulators are placed directly along the PDN leaves, which are shaped as a serpentine in order to keep the hub more compact. The receiver side consists of an array of optical filters feeding photodiodes, which convert the optical signal back into the electronic domain. The photodetector outputs are directly delivered to the transimpedance amplifiers in the electronic layer by means of through silicon vias (TSVs). For modulators and PSEs, the actual ring radii depend on the wavelength channels they are tuned to. In order to model the area of these structures, we assume a reference ring radius of 15 μm for each of them (an average for the typical range of ring resonator radii from 5 to 25 μm [8]). A conservative spacing of 50 μm is left between components in order to avoid undesired coupling effects, as observed in other fabricated devices [9], [10]. This setting also allows to meet the minimum TSV pitch requirements for high yield in the fabrication of TSV arrays (tens of micrometers [11]), like those that connect the electrical layer with the optical one, and vice versa (see Fig. 3). Based on the above assumptions,
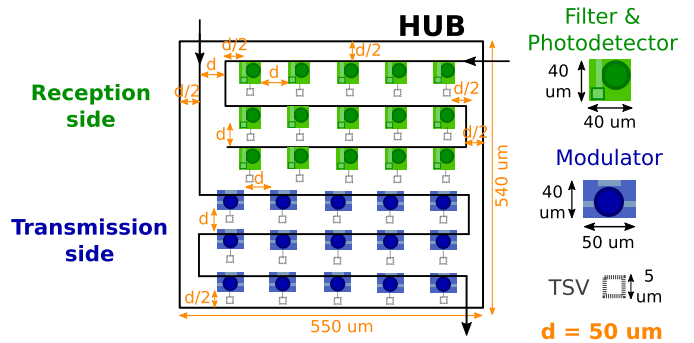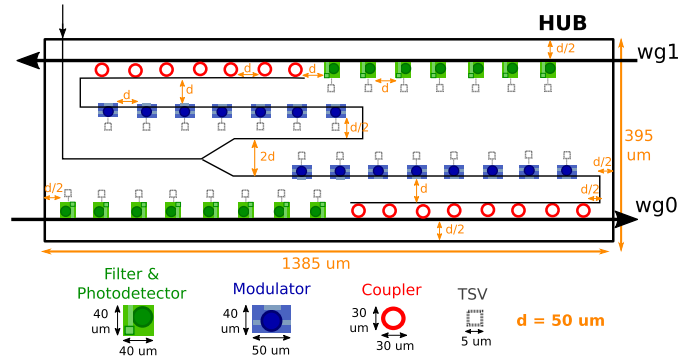
the hub for a 16 × 16 FbONoC has an area of about 0.3 mm^2.

B. ORing Hub

ORings feature a more complex hub, as can be observed in Fig. 4. In fact, modulated signals need to be coupled into the ring waveguides, and the number of waveguides itself is parametric. Fig. 4 shows the simplest case with two waveguides, and assumes that the 15 ring modulators/couplers for transmission and the 15 filters/receivers for reception are equally distributed between the two waveguides. However, the real mapping depends on the decisions that the ring synthesis algorithm takes.

For the sake of analysis, we propose a parametric hub layout and an analytical model for its area that can be adapted to any ring configuration. Each ring waveguide comes with an associated number of modulators, couplers, and ejection filters, depending on the number of wavelength channels that are injected into/ejected from it. Two horizontal lines are reserved on the parametric layout for each waveguide: one for couplers and filters, and one for modulators. The longest line across al waveguides determines the width of the hub, and the number of waveguides determines its height (results in Section IX-B2).

C. Vertical Alignment Dependencies

When placing the optical topology, we must make sure that the ONI area fits within the ENI area located beneath it in the electronic layer.

For FbONoC topologies, vertical alignment dependences are easily fulfilled: ONI area is 0.3 mm^2, and its vertical projection can be considered as the reserved space for placing the TSV array. An additional 0.15 mm^2 should be reserved for electronic components, thus leading to a total ENI area

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ORTÍN OBÓN *et al.*: CONTRASTING LASER POWER REQUIREMENTS OF WRONoC TOPOLOGIES   5

of 0.45 mm². This can be considered as the minimum ENI value for any kind of 16 × 16 WRONoC topologies (both FbONoCs and ORings), since it accounts for just the interface circuits and the TSV array, although minor variations may arise when considering the concentrated or the sparse nature of the TSV array for FbONoC or ring topologies, respectively.

For ORing topologies, Section IX-B2 will prove that for most configurations, hubs turn out to be larger than the minimum ENI size of 0.45 mm². When this is the case, we assume that: 1) the floorplanning fence for the ENI is enlarged to match the footprint of the vertically aligned ONI; 2) TSVs are placed in the ENI fence as dictated by the ONI layout in Fig. 4; and 3) P&R of ENI electronic circuits is performed at reduced row utilization within the enlarged ENI fence.

V. WRONoC PHYSICAL MAPPING FLOW

This paper aims at accounting for the physical mapping phase when assessing static power efficiency of WRONoC topologies. For this purpose, a customized cross-layer synthesis methodology is defined, where logical design properties and physical mapping options are tightly intertwined.

The input to the synthesis methodology is the designer specification of a tentative floorplanning and minimum area requirements for the electronic layer: processor cores, caches, memory macros, I/O peripherals, electrical interconnection fabric, and ENIs. This results in a lower bound for chip area. The main motivation for starting the physical mapping flow this way is that array fabrics of homogeneous processing and memory tiles lend themselves to a straightforward regular floorplanning for the electronic layer. Therefore, the designer may want to build the whole system around it, and check whether the optical plane can be inferred accordingly while meeting the relative interdependences. When this is the case, the electronic layer is not modified and the final chip area is the same as the minimal area of the electronic layer. However, when this is not the case (e.g., ONI is larger than ENI, or more interhub spacing is required), chip area is increased. The proposed synthesis methodology does not break the layout regularity of the electronic layer, but preserves it at the cost of enlarging blocks and lowering area utilization inside them.

Chip area overhead is typically the side effect of searching for more static-power efficient ONoC implementations on the optical plane, such as increasing the SDM degree of ORings, or decreasing the partitioning granularity of FbONoCs. The proposed design methodology returns a number of alternative physical design options for each optical topology under test, spanning a different tradeoff point between chip area and total static power. The designer may then decide to save chip area while accepting a less power-efficient WRONoC design, or the other way around.

Although sharing the above design philosophy, FbONoCs and ORings need for customized steps in their synthesis methodology, due to their complementary characteristics in terms of connectivity pattern versus hub intricacy.

A. FbONoC Physical Mapping Flow

In FbONoCs, the placement and the routing of the topology take place outside the ONIs, and the spacing among ONIs
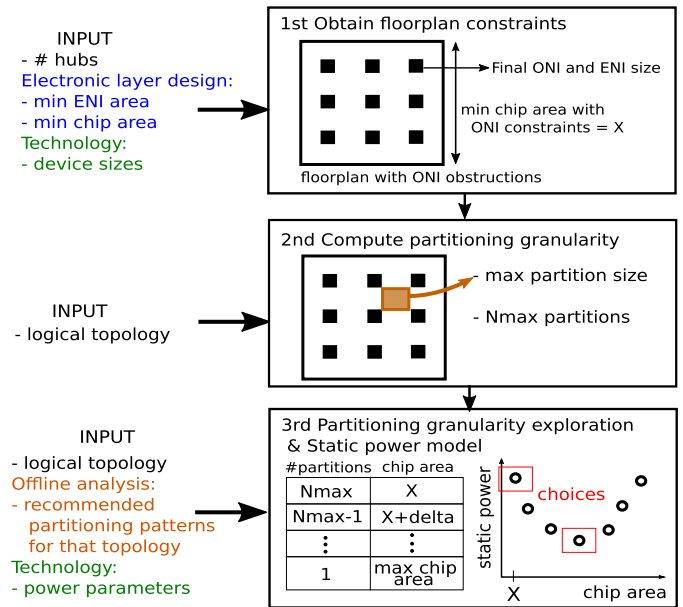

Fig. 5.   Synthesis methodology for FbONoC topologies.

gives rise to a constraint on the area of the largest physical partition that can be implemented on the optical plane. The basic idea behind the proposed physical partitioning strategy is that an FbONoC topology comes with a regular connectivity pattern; therefore, physical partitions should coincide with chunks of the logical topology for better design predictability.

The first step of the synthesis methodology, as shown in Fig. 5, shows how the floorplanning constraints are given by the chip size, limiting the physical topology spread, and by the ONI size (see Section IV-A), dictating floorplanning obstructions, both specified by the designer.

Next, the partitioning granularity has to be computed. Intuitively, the number of partitions should be as small as possible, since disaggregating the logical topology ends up in physical design overhead (e.g., additional crossings for inter-partition connectivity and intersections with the PDN). The maximum physical partition size will ultimately determine the maximum number of partitions that will be inferred (*Nmax*).

Once the partitioning granularity is determined, knowledge of the target logical topology and of its recommended partitioning patterns becomes key. This is the outcome of an offline exploration, which is performed in this paper for the lambda router and snake topologies (see Section VI). In the last step, the target partitioning pattern and the granularity are applied to the logical topology, thus resulting in a layout solution featuring a specific static power requirement. Other solutions, potentially more power-efficient, may be explored by increasing the chip area and thus the interhub spacing, allowing to fit larger fewer partitions.

B. Optical Ring Physical Mapping Flow

The key issue to lay out an optical ring consists of addressing the possible mismatches between ENI and ONI size. In fact, this latter depends on the ring configuration in terms of SDM and WDM degrees. Clearly, layout design and architecture configuration are tightly intertwined, and are
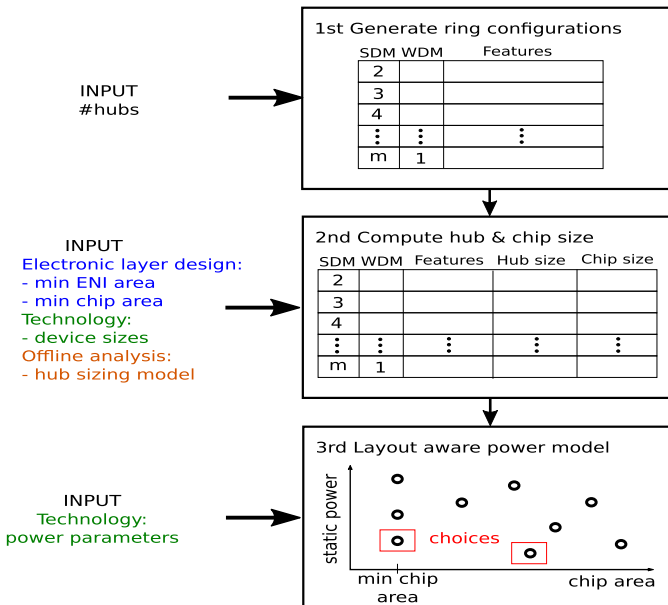
Fig. 6.   Synthesis methodology for ring configurations.

comprehensively addressed in the ring synthesis methodology shown in Fig. 6.

As a first step, we propose a synthesis algorithm that optimizes the number of wavelengths to be used to deliver contention-free all-to-all connectivity on top of a preassigned number of physical ring waveguides (see Section VII-A). Therefore, it is possible to build up a table associating the minimum WDM degree for a given SDM degree, where the latter is varied from its minimum (two waveguides) to its maximum value (until the WDM degree saturates at one wavelength).

Then, a parametric analytical model is applied to compute the hub size for each ring configuration, by following the layout design guidelines in Section IV-B. If the resulting size fits within the minimum ENI area specified by the designer, the configuration can be implemented with the given minimum chip size. If it is bigger, the ENI fence will have to be enlarged (see vertical alignment dependences assumed in Section IV), thus increasing chip size accordingly. After this step, all the ring configurations will be associated with their feasible chip area.

Finally, a layout-aware power modeling framework is applied to compute the static power consumption of the ring configurations. Such configurations will then be placed on a (static power-total chip area) 2-D plot, where the designer will be able to select the desired tradeoff point.

Interestingly, both physical mapping flows end up in the same kind of design curve, positioning physical design options in the power-area optimization space. The only difference is that such design options correspond to different partitioning granularities for FbONoC topologies, and to different WDM-SDM configurations for ORing ones.

VI. SYMMETRY-BASED FbONoC PARTITIONING

This section searches for the most power-efficient partitioning pattern for FbONoC topologies under test, which is an input for their physical mapping flow.

A. Partitioning Patterns

The state of the art in physical mapping of FbONoC topologies consists of monolithically placing the logical schemes in Fig. 1(a) and (b) into the section delimited by a large fence in the middle of the optical layer [29]. We refer to the resulting design as the single box (SB) pattern. Unfortunately, this technique lacks physical design flexibility. Therefore, we present several partitioning patterns that retain the symmetry of the the topologies.

1) Central Partition Criteria: We extract a central partition from the topology and interconnect the remaining I/O partitions through it. Using this criteria, we develop two partitioning patterns: asymmetric central partition (AsyCP), with a different number of PSEs across partitions [Figs. 7(a) and 8(a)], and symmetric central partition (SyCP), with the same number of PSEs in all partitions [Figs. 7(b) and 8(b)]. The SyCP pattern can also be scaled in order to obtain smaller yet fairly homogeneous partitions [scaled central partition (SCP)] in Fig. 7(c) for the lambda router, not applicable to the snake due to its triangular shape.

2) Tiling Criteria (TLG): We create partitions with a uniform number of 2×2 PSEs, where each partition or tile connects either initiators or targets. Figs. 7(d) and 8(c) show the tiling patterns for the lambda router (with rectangular tiles) and snake (with triangular tiles), respectively. This partitioning mechanism can be scaled down to obtain smaller partitions, as shown in Figs. 7(e) and 8(d) (SLTLG).

3) Horizontal Striping Criteria: Since communication actors are both initiators and targets, we create partitions that are connected to both the initiator and target side of such actors, in order to exploit their physical proximity. The resulting patterns, with each partition connected to four nodes, are shown in Figs. 7(f) and 8(e). Due to their different geometrical shapes, in the lambda router, the partitions have the shape of horizontal stripes, while in the snake, they appear as bent stripes (folded horizontal striping).

B. Physical Layouts

From Section V-A, a partition must fit within the space among hubs. Fig. 9 shows the physical layouts for all the proposed partitioning patterns for the lambda router topology. Similarly, snake layouts are generated, but omitted for lack of space. Floorplan area was set to 16 mm $times$ 16 mm, so that even the most demanding SB solution could easily fit. The layouts are composed of two overlapped networks.

1) Communication Network: It connects the initiator side of the hubs with the partitions (red links), the reception side of the hubs with the partitions (black links), and the partitions among each other (yellow links). These physical layouts were implemented manually to minimize the number of waveguide crossings and the propagation distance, which are indirect indicators of energy efficiency.

2) Power Distribution Network: It is implemented as a perfect binary tree to minimize the number of splitters (blue links).
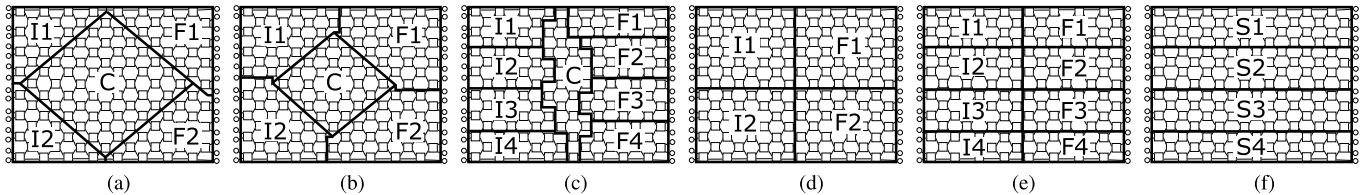
Fig. 7. Partitioning patterns of the lambda router topology. (a) AsyCP. (b) syCP. (c) SCP. (d) TLG. (e) STLG. (f) HS.
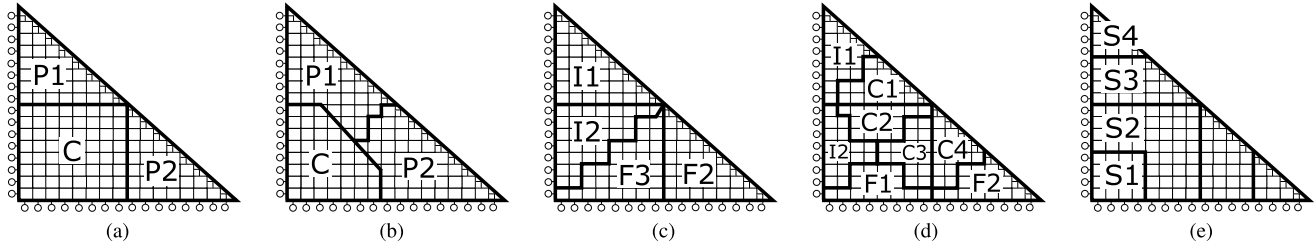

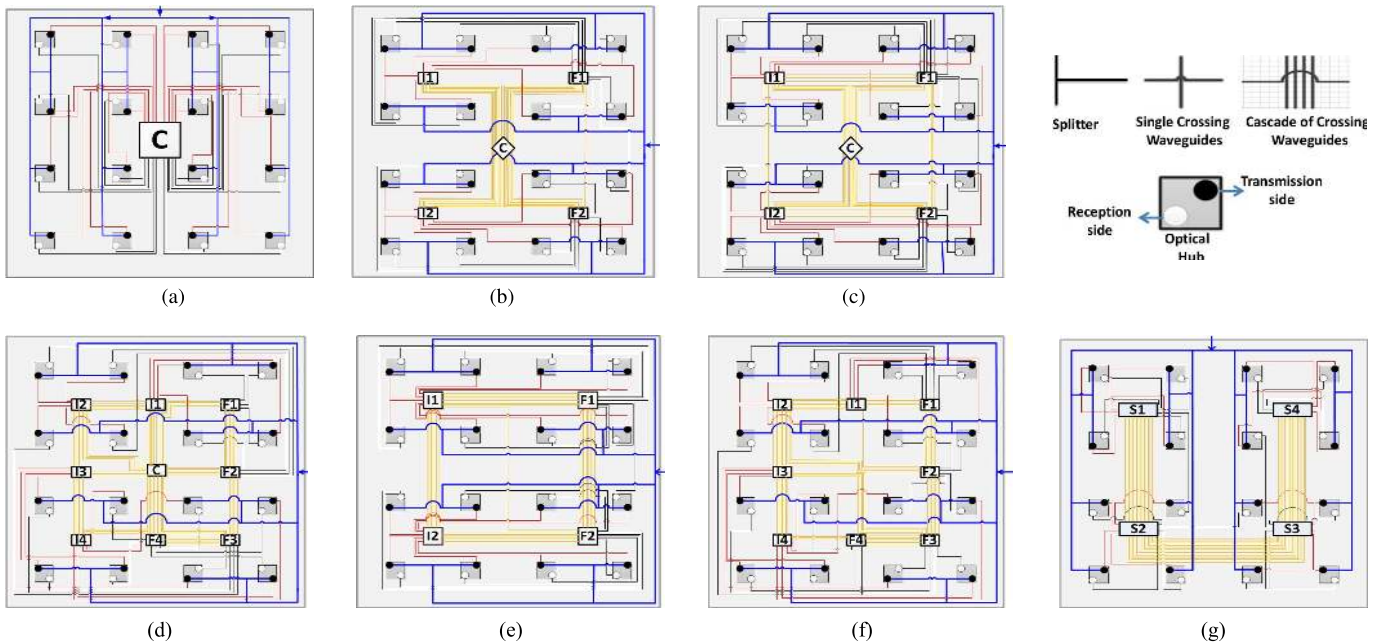Fig. 8. Partitioning patterns of the snake topology. (a) AsyCP. (b) syCP. (c) TLG. (d) STLG. (e) FHS.

Fig. 9. Physical layout of the proposed partitioning patterns for the lambda router topology, corresponding to the logical layouts shown in Fig. 7. Note that the same partition labels introduced in that the figure is reused now to indicate where each partition has been manually placed. (a) SB. (b) AsyCP. (c) SyCP. (d) SCP. (e) TLG. (f) STLG. (g) HS.

VII. RING SYNTHESIS ALGORITHM

This section describes the algorithm developed to generate optical ring designs that corresponds to the first step in the ring synthesis flow presented in Section IV-A. We then introduce the PDN used to bring the laser into every hub of the optical ring.

A. Generating the Optical Ring Communication Matrices

The waveguide-wavelength pair to implement each communication on an optical ring has to be chosen to ensure that the same wavelength will never be used on the same waveguide section twice, thus avoiding interferences.

Our mechanism to build the ring communication matrix is detailed in Algorithm 1. We share the same basic design ideas with the most relevant previous work on the topic [19]: a wavelength can be used to implement several communications on the same waveguide, and we alternate clockwise and counterclockwise waveguides. The latter gives flexibility to implement comunications using the shortest path between nodes, which involves reduced insertion loss and maximizes the reuse of waveguides and wavelengths. Every waveguide is assigned a unique propagation direction in order to simplify the hub design.

As an input, our algorithm needs the number of waveguides of the ring, a maximum number of wavelengths, and the connectivity matrix. The latter allows us to indicate which specific nodes we want the ring to connect and, therefore, to customize the ring for partial connectivity. The output consists of two communication matrices: one for the waveguides and one for the wavelengths that will have to be used for each communication. For a given number of waveguides, which may be limited by the place-and-route constraints, the algorithm generates the ring design with minimal number of wavelengths.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8   IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS

Algorithm 1 Generate Optical Ring Communication Matrices

1: **Input Data:** *num_waveguides, max_num_wavelengths,*
        connectivity_matrix
2: **Output Data:** *waveguide_matrix, wavelength_matrix*
3: $ring \leftarrow generate\_ring(num\_waveguides, 0\ wavelengths)$
4: $used\_wavelengths \leftarrow 0$
5: **for** $communications\ from\ connectivity\_matrix\ COM$ **do**
6:       ▷ First try to reuse a wavelength to set the
            communication on the short path.
7:       $success \leftarrow false$
8:       **for** $used\_wavelengths\ wl$ **do**
9:           $wg \leftarrow ring.get\_free\_wg\_short\_path(COM, wl)$
10:          **if** wg exists **then**
11:              $store\_communication\_short\_path(COM,$
                    $waveguide\_matrix, wavelength\_matrix, wl, wg)$
12:              $ring.store\_use\_short\_path(COM, wl, wg)$
13:              $success \leftarrow true$
14:              **break**
15:          **end if**
16:      **end for**
17:
18:      ▷ If it did not work, try adding a new wavelength
19:      **if** $NOT\ success$ & &
            $used\_wavelengths < max\_num\_wavelenghts$
               then
20:          $ring.add\_wavelength()$
21:          $used\_wavelengths + +$
22:          $store\_communication\_short\_path(COM,$
                $waveguide\_matrix, wavelength\_matrix, new\ wl, first\ wg)$
23:          $ring.store\_use\_short\_path(COM, new\ wl, first\ wg)$
24:          $success \leftarrow true$
25:      **end if**
26:
27:      ▷ If we could not add more wavelengths, try setting
            the communication on the long path
28:      **for** $used\_wavelengths\ wl$ **do**
29:          $wg \leftarrow ring.get\_free\_wg\_long\_path(COM, wl)$
30:          **if** wg exists **then**
31:              $store\_communication\_long\_path(COM,$
                    $waveguide\_matrix, wavelength\_matrix, wl, wg)$
32:              $ring.store\_use\_long\_path(COM, wl, wg)$
33:              $success \leftarrow true$
34:              **break**
35:          **end if**
36:      **end for**
37:
38:      **if** $NOT\ success$ **then**
39:          $ERROR : Unable\ to\ generate\ ring$
40:          **break**
41:      **end if**
42: **end for**

For each communication that needs to be implemented in the ring (loop in line 5 of the pseudocode), the algorithm first tries to set the connection on the minimal path between the two nodes reusing a wavelength already present in the design (lines 8–16). If that is not possible because some of the required ring sections are not free in any waveguide with any of the existing wavelengths, a new wavelength will be added to set the communication (lines 19–25). If the maximum number of wavelengths had already been reached, then the algorithm will try to set the communication on the nonminimal path, going around the ring in the other direction (lines 28–36). If it is not possible to do that either, the algorithm will finish its execution unable to generate the ring design with the given input (lines 39–41). The complexity of the algorithm is polynomial: $O(n^3)$, n being the number of nodes. The polynomial complexity guarantees that the algorithm will scale efficiently as we increase the number of nodes.

The first difference of the algorithm in [17] with respect to ours is that they fix the number of wavelengths to use and utilize all of them in the same waveguide until it is not possible to set any of the remaining communications, at which point
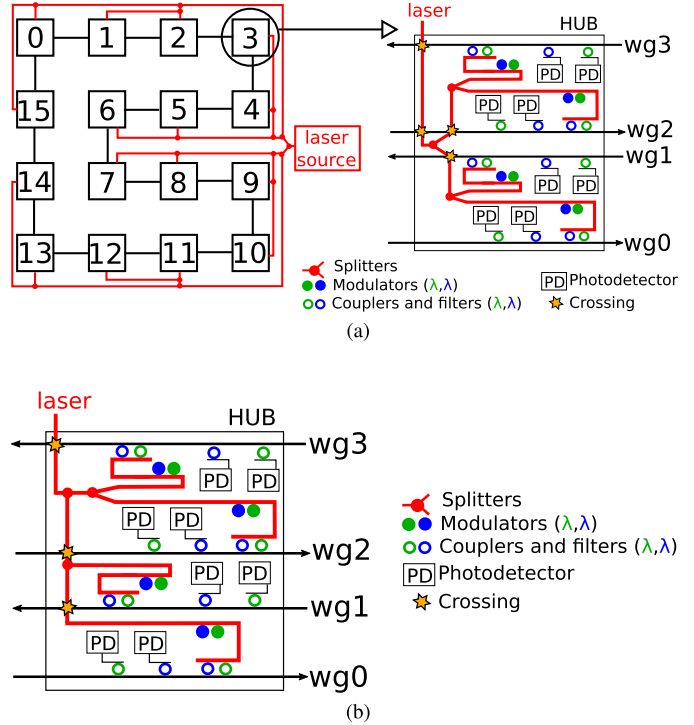

(a)

(b)

Fig. 10.   PDN for an optical ring. (a) Perfect binary tree PDN in a 16-node ring with four waveguides, with details of the optical power distribution to all the ring waveguides. (b) Optimized PDN inside the hub to reduce the number of crossings.

they add a new waveguide. The drawback is that using up all the sections of the first waveguide before adding a new one forces the algorithm to use nonminimal paths for several communications that could have found a shorter path on a second waveguide. Having longer paths has a negative impact on laser power requirements, because it increases the number of crossings and the propagation loss. To avoid this problem, we fix the number of waveguides of our ring and reuse the same wavelength on all of them as much as we can before adding a new one, always trying to set communications on the shortest path first.

A very important detail not mentioned in [17] is the order in which communications are set. Setting long-path communications first allows shorter communications to fill the gaps left on the ring by the longer communications, and generally yields better results.

B. Power Distribution Network

Unlike FbONoC topologies, inside every ring hub, the laser needs to reach all the ring waveguides. This is achieved by using a hierarchical perfect binary tree (both at the top level and at the hub level), which ends up generating unintended crossings inside hubs [Fig. 10(a)]. We optimize the consequent insertion loss degradation by strategically repositioning the splitters in order to minimize the number of crossings. The optimized PDN layout for four waveguides is presented in Fig. 10(b). Note that this optimization will have a higher impact on rings with a higher number of waveguides. Equivalent designs for any number of waveguides are used in the rest of this paper.

TABLE I

TECHNOLOGY PARAMETERS

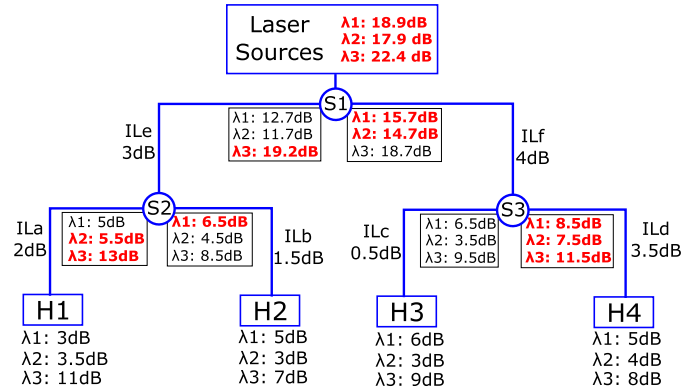| Photonic components	Device Information
Propagation loss	0.274 dB/cm [41]
Bend	0.005 dB [13]
Crossing	0.05 dB [33]
Splitter	0.2 dB [33]
MRR Drop loss	1 dB [33]
MRR Passing loss	0.005 dB [33]
Receiver	1 dB [33]
Modulator	1 dB [33]
Coupling efficiency	90% [18]
Laser efficiency	20% [16]
Receiver sensitivity	-20 dBm [33]

Fig. 11. Example of the laser distribution network to bring three wavelength carriers to four nodes. The insertion loss of the optical paths fed by each wavelength channel has been precalculated for every node. At every splitter, the total insertion loss incurred by each wavelength channel on each branch has been annotated, and the worst case has been highlighted in red. An insertion loss of 0.2 dB has been considered for the splitters.

VIII. LASER POWER MODEL

This section describes the methodology we have followed to calculate the static power required by any WRONoC topology. Contributors to static power are given by the total laser power, the thermal tuning of microring resonators, and by modulators and receivers. In this paper, we focus only on laser power as the key differentiator among the physical designs under test, the other contributors being equivalent. In fact, the number of transmitters and receivers is dictated by the wavelength routing methodology, not by the topology. Also, when microring resonator-level thermal tuning is applied [4], [35], the number of physical devices to control is the same for ORings and FbONoCs. In addition to the common devices, ORings make use of couplers, while FbONoCs make use of switching microring resonators. It can be easily demonstrated that their number is in both cases equal to $N(N-1)$, where N is the number of initiators/targets. As a result, the laser power is the only differing contribution to static power across the designs under test, since it is a function of the layout design efficiency.

We calculate insertion loss based on the parameters in Table I, assuming aggressive crossing and propagation loss parameters from the literature. We then compute total laser power to guarantee that, after the insertion loss along every path, the received optical power at photodetectors matches the target receiver sensitivity.

In order to feed multiple optical paths (each one originating at a different hub) with the same wavelength carrier coming from the same laser source, we need to use splitters. Ideally, we would need selective splitters to apply the required splitting ratio to each individual wavelength carrier, in order to bring the exact optical power needed by every hub. Unfortunately, we would need as many PDNs as wavelength channels. This approach is clearly impractical due to the wiring intricacy that would arise. As a consequence, we consider a unified PDN for all the wavelength channels, which necessarily gives rise to some form of power equalization across them at PDN splitters.

As an example, let us consider the case of a PDN built with 3-dB splitters (i.e., 50% splitting ratio), and let us appreciate the equalization effect on the example in Fig. 11. This example represents a system with four hubs and three wavelength channels, for a generic WRONoC topology without SDM. For each hub, we have annotated the insertion loss of the paths that use the different wavelength channels and start at that hub. We have to calculate power independently for each wavelength, traversing the laser distribution tree starting from the leaves and working our way up to the root. We show here how to calculate the power for $\lambda1$. We start by adding the insertion loss in hub H1 (i.e., 3 dB) to the insertion loss of the path to reach the splitter S2, ILa, resulting in a total of 5 dB. We repeat the process for hub H2, obtaining a loss of 6.5 dB. Since we are assuming a splitting ratio of 50%, we need to select the worst case and use it for both branches in order to guarantee that enough power reaches the two hubs. In this case, we assume a loss of 6.5 dB for both branches; this means that we will waste optical power in the branch with the lower power requirement. Then, we add 3 dB, which corresponds to the 50% splitting ratio, the power corresponding to the loss of the splitter itself (0.2 dB in our case), and the path to the next splitter (3 dB of ILe). Here, we repeat the same process of choosing the worst case between the splitter branches. When reaching the PDN root, the worst case insertion loss for $\lambda1$ is finally derived, which can be easily converted into a static power requirement for the associated optical source by fulfilling the receiver sensitivity requirement. To get the total laser power requirement of a topology, we simply need to add the power for all the wavelength channels.

Considering different splitting ratios at each splitter requires a different equalization approach (e.g., per-hub instead of per-wavelength channel), and we have experimentally verified that it is not necessarily the most power-efficient approach: it depends on the exact value of the insertion losses for the optical paths at hand. For instance, on the example in Fig. 11, the per-hub equalization provides worse laser power requirements. Also, highly unbalanced splitting ratios come with manufacturing issues due to the unavoidable variability of process parameters. We therefore consider the most common and reliable splitting ratio (50%) across all topologies, and we compare the resulting power efficiency with that of an ideal PDN (no insertion losses and no equalization) bringing only the needed amount of laser power to the hubs. This allows us to evaluate the power distribution intricacy for each topology,
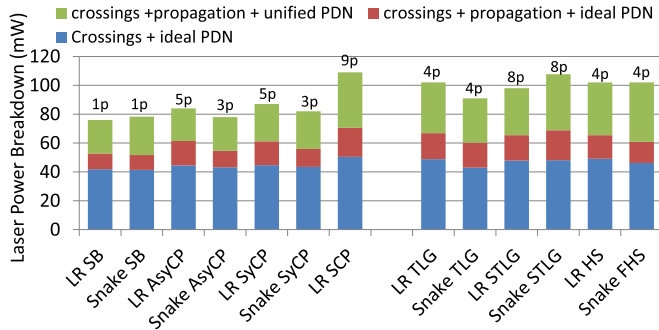
Fig. 12. Laser power requirements for the filter-based topologies with all the partitioning configurations. The number of partitions of each configuration is annotated above each bar. Chip size is 16 mm × 16 mm.

TABLE II

MINIMUM CHIP SIZE TO PLACE AND ROUTE THE LAMBDA ROUTER AND SNAKE WITH EACH PARTITIONING PATTERN

Partition pattern	lambda router minimum die size	Snake minimum die size
Single Box	8.8 x 8.8 mm	10.4 x 10.4 mm
Asymmetric Central Partition	7.6 x 7.6 mm	7 x 7 mm
Symmetric Central Partition	6.8 x 6.8 mm	7.4 x 7.4 mm
Scaled Central Partition	6 x 6 mm	-
Tiling	7.2 x 7.2 mm	7.4 x 7.4 mm
Scaled Tiling	6 x 6 mm	6.4 x 6.4 mm
Stripes	10 x 10 mm	9.6 x 9.6 mm

but also to highlight whether custom-tailored PDNs for each topology might potentially reverse the power balance achieved with 3-dB splitters.

IX. EXPERIMENTAL RESULTS

We present the power-area tradeoff for the considered topologies and compare their static power subject to die area and P&R constraints.

A. Physical Partitioning of Filter-Based Topologies

Fig. 12 presents the laser power requirements of the lambda router and snake topologies laid out with all the partitioning patterns considered in this paper. A conservative chip size of 16 mm × 16 mm is considered, so that all topology configurations match the available interhub spacing. We notice that the lambda router and snake are quite equivalent, even though the snake is a structurally unbalanced topology. This is because our physical design methodology combines the efficient paths of the snake with the lossiest paths of the PND, thus leveling the insertion loss. Looking at the breakdown, we observe that most of the contribution to the laser power comes from the waveguide crossings and the equalization effect of the unified PDN.

Clearly, configurations with more partitions involve higher power consumption, up to 43% and 35% for the lambda router and snake, respectively, with respect to the SB pattern. This is because more waveguides are required to connect all the partitions with each other, thus increasing the number of crossings and the propagation distance of optical paths.

Tiling for the lambda router is an exception though. It exhibits a high inefficiency from the ground up (e.g., with four partitions), due to lots of crossings between the PDN and the communication network. However, the situation improves by scaling up the tiled layout to eight partitions (see STLG bar). This is because partitions can be placed closer to the hubs they connect. Finally, the striping pattern does not outperform the tiling one, since the partially open square of interpartition links comes at the cost of more waveguides per square edge, thus making the interaction with the PDN and the I/O links equally important. Overall, a central partition is a good idea to lay out the lambda router with up to five partitions, while a more aggressive unfolding of the topology is better assisted by the tiling pattern. For the snake, the conclusion is similar.

These results clearly suggest to consider the physical design configuration with the smallest number of partitions that fits with the chip area requirement. Table II includes the minimum chip area required to place and route the lambda router and snake with each partitioning pattern. These results follow directly from the layouts shown in Fig. 9. We find that an SB layout, eliminating interpartition connections, can be inferred with a minimum chip area of 8.8 mm × 8.8 mm for the lambda router. With a larger number of partitions, the minimum chip area of 6 mm × 6 mm can be achieved with the same topology through the SCP and scaled tiling configurations. The snake has different area requirements for the same partitioning patterns, in general higher than the lambda router.

Interestingly, the striping pattern for the lambda router turns out to be more area hungry (and obviously less power efficient) than the SB one, because each stripe encompasses a large number of optical switches and waveguides. Therefore, we will not consider this pattern for our future experiments. We will use the above results to select physical design options compatible with chip area requirements in Sections IX-C and IX-D.

B. WDM-SDM Tradeoff in Optical Ring Design

We run our ring synthesis algorithm to obtain optical rings that connect a varying number of hubs, and demonstrate that our designs have fewer number of wavelengths and/or waveguides than previous proposals, hence potentially resulting in more power-efficient design points. Then, we derive the fundamental implication of the achieved WDM-SDM tradeoff over the hub area for layout-aware ring design.

1) Generation of Efficient ORing Configurations: Fig. 13 reports the number of wavelengths that are needed to deliver conflict-free all-to-all connectivity over rings with different number of nodes and waveguides. Fig. 13 also reports the comparison with the available data from the algorithm proposed by Le Beux *et al.* [17].

As expected, as we increase the number of nodes to be interconnected, more waveguides and/or wavelengths are needed to implement all the connections. Above all, increasing the number of waveguides allows us to reduce the number of required wavelength channels, because each one of them can be reused for more communications across the available waveguides. The most significant result is that, given the same number of waveguides, our algorithm is able to build the ring with fewer wavelength channels than [17] in all cases,
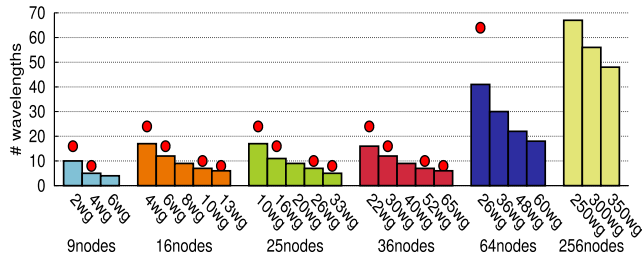
Fig. 13. Our ring synthesis algorithm at work for different numbers of waveguides and different network sizes. The red dots represent the results from the algorithm in [17]. (note that these numbers have been extracted from a graph and there may be small imprecisions).
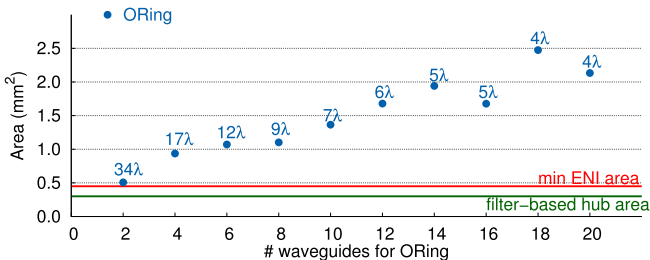

Fig. 14. Hub area for optical rings that connect 16 nodes with different numbers of waveguides and wavelengths. Hub area for filter-based topologies and minimum ENI area are included for comparison.

the difference becoming more prominent as the system size scales up. As a result, we will derive more compact hubs and more power-efficient design points than those achievable from the existing literature, hence improving the competitiveness of ORing configurations.

2) Hub Area for ORing Configurations: Fig. 14 shows the hub area for several ring configurations as well as for filter-based topologies, and reports the minimum ENI area as a reference, as explained in Section IV-A. We include only configurations with an even number of waveguides, because having the same number of clockwise and counterclockwise waveguides allows to build more balanced and power-efficient desings.

The hub for the optical ring with two waveguides roughly matches the minimum ENI area. As the number of waveguides increases, the hub area consistently grows. The reason for this is twofold.

- Increasing the number of waveguides causes the height of the hub to increase.
- As the number of waveguides increases, the total number of wavelengths decreases on a similar proportion. However, the hub width is determined by the maximum number of communications that start and finish on each waveguide, and this decreases only marginally.

The hub for a wavelength-routed optical ring will always require a larger area on the optical plane than the corresponding network interface in the electronic layer, thus causing a potential source of inefficiency for 3-D stacked designs. In contrast, FbONoC topologies require a hub that fits within the vertically projected ENI fence.
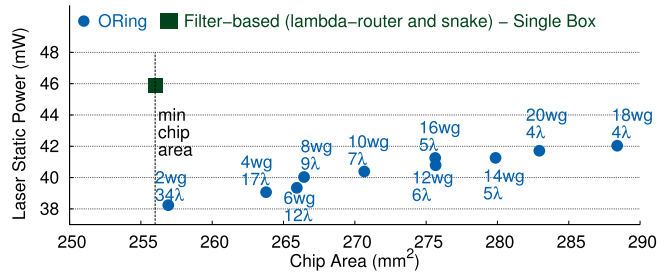

Fig. 15. Power for the optical ring and filter-based topologies with a baseline chip size of 16 mm × 16 mm, without considering the PDN.

C. Laser Power Versus Die Size For 16 Nodes, Without Power Distribution Network

We first compare the laser power requirements of the optical ring versus filter-based topologies without including the PDN, as explained in Section VIII. We follow the custom-tailored synthesis methodologies highlighted in Figs. 5 and 6, using the previous results as inputs to the flows: recommended partitioning patterns from Section IX-A for FbONoC topologies, and WDM-SDM configurations, in addition to hub area, from Section IX-B for ORing design. We initially consider a conservative chip size requirement of 16 mm × 16 mm provided by the electronic layer designer, for the connectivity of 16 hubs with each other. The minimum ENI area is specified to be again 0.45 mm². Both physical mapping flows end up in a laser power-chip area design plot, which is jointly shown in Fig. 15. The area is the same for the electronic and optical layers; 16 mm × 16 mm is the required chip area for the electronic layer, and has been enlarged when necessary for the components in the optical layer to fit.

For improved readability, only the best lambda router and snake configurations are reported, which are also apparently overlapped due to the scale on the *y*-axis. In fact, with the specified chip area, the filter-based topologies can be efficiently placed as an SB; in contrast, the larger hub area of rings forces us to increase the ENI area on the electronic plane and, hence, the chip area. Overall, the ring with two waveguides has 17% lower power consumption than the lambda router and snake with only a marginal increase in area. This is because its straightforward conversion from the logical to the physical layout generates a simple design with reduced insertion loss.

As we increase the number of waveguides of the optical ring, we can implement the communications with fewer wavelengths. However, this comes at the cost of a fast degradation of chip area. In terms of total laser power requirement, it is reduced as an effect of the lower number of laser sources, but increases because of more waveguide crossings at hubs, and of longer propagation distance. In practice, Fig. 15 shows that laser power consistently increases under the prevailing effect of the more intricate hub wiring.

Fig. 16 shows the laser power requirements of the topologies under test, when scaling the baseline chip size down to 8 mm × 8 mm. From Table II, the SB layout becomes infeasible, while the AsyCP pattern should be chosen as the most power-efficient solution (see Fig. 12). It is, however,
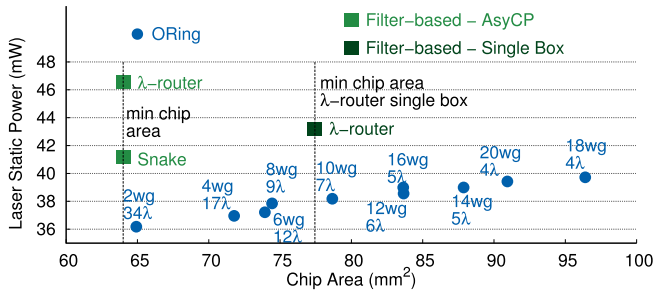
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

12

IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS

Fig. 16.   Power for the optical ring and filter-based topologies with a baseline chip size of 8 mm × 8 mm, without considering the PDN.
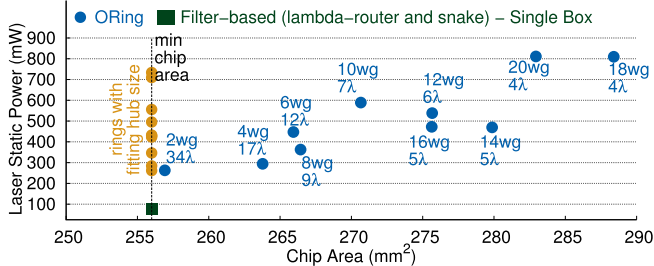

Fig. 17.   Power for the optical ring and filter-based topologies with a baseline chip size of 16 mm × 16 mm, including the PDN. We also include the same ring configurations under the (unrealistic) assumption that the hubs fit the minimum chip area requirement (orange circles).

not worth it to increase the area of filter-based topologies beyond the specification in order to restore the feasibility of SB layouts, since we would achieve only 1.3% power savings with an area overhead of 44%. Using the AsyCP pattern, the snake features three partitions, while the lambda router requires four, which explains the lower power consumption of the former. Again, the ring is more power efficient than FbONoC topologies, with a reduction of 13% and only 1% area overhead.

D. Laser Power Versus Die Size for 16 Nodes Including the Power Distribution Network

Fig. 17 shows laser power requirements for a chip size of 16 mm × 16 mm when including the real PDN (i.e., its crossings, propagation distance, bends, and equalization effects). Fig. 17 can be directly compared with Fig. 15, which lacks the PDN contribution. This time, the filter-based topologies consume between 70% and 90% less power than the optical ring. The big difference comes from the splitters, which send the same amount of optical power toward every branch, even though this is not strictly needed.

The effect of the splitters is not so harmful in filter-based topologies for a twofold reason. First, the PDN tree is not as deep as in the ring (the tree leaves are the hubs themselves, instead of the several ring waveguides inside each hub). Second, all wavelength channels are used at every hub to feed just as many optical paths.

In contrast, the ring has a deeper PDN tree, hence potentially resulting in a larger number of power equalization points, and a nonuniform utilization of wavelength channels across hubs. In practice, this means that the PDN will bring a wavelength channel to hubs and/or waveguides inside the hubs where that wavelength carrier is not actually used. That is why total

TABLE III
LASER POWER FOR THE TOPOLOGIES UNDER TEST WITH VARYING TECHNOLOGY PARAMETERS

Technology	Snake	λ-router	Best ORing 2 waves	Ratio Best FbONoC / Ring
Crossing loss = 0.05 dB Propag. loss=0.274dB/cm (Aggressive technology - this paper)	78 mW	76 mW	263 mW	-71%
Crossing loss = 0.15 dB Propagation loss=1 dB/cm (Conservative technology)	389 mW	375 mW	591 mW	-36%
Crossing loss = 0.05 dB Propagation loss=1.5dB/cm (Propag. loss-dominated)	290 mW	290 mW	872 mW	-67%
Crossing loss = 0.5 dB Propag. loss=0.274dB/cm (Crossing loss-dominated)	4.27 W	3.43 W	442 mW	7.75X

laser power increases so sharply as we increase the number of waveguides of the ring.

Finally, we demonstrate that this laser power overhead for the ring comes mainly from the waste of bringing optical power to locations where it is not needed, and not from the bigger chip size. Fig. 17 includes the power consumption of the same ring configurations assuming the hubs always fit the vertically projected ENI fence, which is not the case in practice. Power consumption is slightly lower because propagation distances are shorter, and differences are more pronounced for configurations with larger chip sizes. However, even assuming this ideal engineering of ring hubs, the gap with the laser power requirement of filter-based topologies is still significant.

In order to generalize the results in Fig. 17, we report in Table III the laser power requirements of the topologies under test when varying technology parameters. We notice that the optical ring is more static power efficient only when the weight of the crossing losses is disproportionate.

E. Static Power Breakdown

Fig. 18 shows the total static power for the best configuration of each of the studied topologies in a 16 mm × 16 mm chip, with a breakdown of laser, thermal tuning, and transmitter and receiver contributions. For thermal tuning, we consider 1-μW heating power per ring per kelvin, and 20k tuning range [4]. Also, static power is 0.025 mW per transmitter, and 0.05 mW per receiver [2]. Laser power is the main contributor to total static power, corresponding to 54% and 89% in the ring without and with the PDN, respectively. As we have explained in Section VIII, thermal tuning and transmitter and receiver power are the same for every topology. Therefore, including them uniformly increases the power for all topologies. The ring has 9.7% lower total static power consumption than filter-based topologies if the PDN is not included. In contrast, when the PDN is included, filter-based topologies consume 63% less total static power than the ring.

F. Crosstalk Analysis

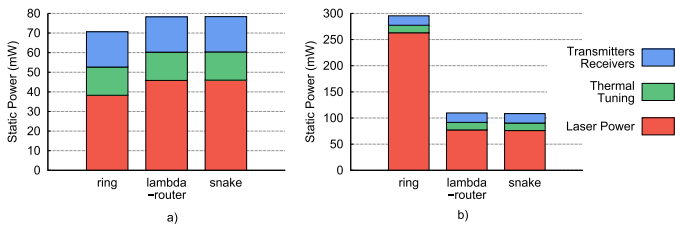We compute first order crosstalk noise for all the network components by using the same methodology and crosstalk

Fig. 18. Total static power for the optical ring and filter-based topologies with a baseline chip size of 16 mm × 16 mm. (a) Without PDN. (b) With PDN.

TABLE IV
MINIMUM, MAXIMUM, AND AVERAGE SNR FOR THE BEST FbONoC AND RING TOPOLOGIES IN FIG. 17

Topology	min	avg	max
Lambda-router	8.7 dB	12.0 dB	18.8 dB
Ring	-0.73 dB	6.4 dB	18.7 dB
Optimized ring	11.9 dB	16.0 dB	24.1 dB

coefficients as in [40]. We select the actual values for wavelengths and microring resonator radii to guarantee that no routing faults take place in the WRONoC [41].

Table IV shows the minimum, maximum, and average signal-to-noise ratios (SNR) for the best FbONoC and Oring designs with a chip size of 16 mm × 16 mm. Surprisingly, even though the ring has a simple physical design with few crossings and no PSEs, the SNR is better in the lambda router. Again, this is a side effect of the inefficient PDN, since it brings power on all wavelength channels to every node and waveguide, even though they are not used at that specific location. That wasted optical power filters through the couplers as first-order crosstalk noise, and is captured by on-resonance filters at the reception side of downstream hubs.

A straightforward optimization applies to the ORing: filtering off the unused channels before the coupler array at each node. The SNR improvement is apparent in the last row of Table IV (33% better than lambda router on average), which, however, comes with an overhead of about 5% in chip size (to make room for the additional filters), and 9.9% in total static power without the PDN (2.4% with the PDN).

X. CONCLUSION

The comprehensive layout-aware topology comparison framework reported in this paper points out that ring topologies can take advantage of their efficient and predictable routing pattern by instantiating a low number of waveguides (2 in a 16×16 network). The potential power gap with respect to FbONoC topologies ranges between 12% and 17%. However, routing the PDN reverses the laser power balance, mainly because of the equalization effect of power requirements across optical paths that a unified distribution network imposes. The inefficient PDN has negative side effects on the SNR of ring optical paths too. While simple optimizations can restore superior noise immunity over FbONoC topologies, they come at a nonnegligible area and static power overhead.

Another milestone achieved by this paper consists of the systematic definition and vertical integration of the physical mapping steps for wavelength-routed topologies. This paper lays the groundwork for the full automation of this flow,

for which there are existing gaps. In particular, for filter-based topologies, partitioning of the connectivity pattern, together with placement and routing of partitions, needs to be automated. However, the key take-away from this paper is that some form of PSE clustering easily leads to more power-efficient layouts than fully distributed approaches. For ring-based topologies, routing of the PDN is left for future automation. However, we stress that emphasis should be given to the equalization effect that unified PDNs induce rather than to the physical routing of PDN waveguides itself. In both cases, the state of the art and methodologies should be upgraded to gain visibility of layout effects.

REFERENCES

[1] D. A. B. Miller, "Rationale and challenges for optical interconnects to electronic chips," *Proc. IEEE*, vol. 88, no. 6, pp. 728–749, Jun. 2000.

[2] S. Beamer *et al.*, "Re-architecting DRAM memory systems with monolithically integrated silicon photonics," in *Proc. Int. Symp. Comput. Archit.*, Saint-Malo, France, Jun. 2010, pp. 129–140.

[3] G. Kurian *et al.*, "ATAC: A 1000-core cache-coherent processor with on-chip optical network," in *Proc. Int. Conf. Parallel Archit. Compil. Techn.*, Minneapolis, MN, USA, Sep. 2010, pp. 477–488.

[4] A. Joshi *et al.*, "Silicon-photonic Clos networks for global on-chip communication," in *Proc. Int. Symp. Netw.-Chip*, San Diego, CA, USA, May 2009, pp. 124–133.

[5] J. Chan and K. Bergman, "Photonic interconnection network architectures using wavelength-selective spatial routing for chip-scale communications," *J. Opt. Commun. Netw.*, vol. 4, no. 3, pp. 189–201, Mar. 2012.

[6] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Opt. Netw. Mag.*, vol. 1, no. 1, pp. 47–60, Jan. 2000.

[7] L. Ramini, M. Tala, and D. Bertozzi, "Exploring communication protocols for optical networks-on-chip based on ring topologies," in *Asia Commun. Photon. Conf. OSA Tech. Dig.*, 2014, pp. 1–8, paper ATh3A.165, doi: 10.1364/ACPC.2014.ATh3A.165.

[8] W. Bogaerts *et al.*, "Silicon microring resonators," *Laser Photon. Rev.*, vol. 6, no. 1, pp. 47–73, 2012.

[9] N. S. Droz *et al.*, "Optical 4×4 hitless silicon router for optical networks-on-chip (NoC)," *Opt. Exp.*, vol. 16, no. 20, pp. 15915–15922, 2008, doi: 10.1364/OE.16.015915.

[10] B. Stern *et al.*, "On-chip mode-division multiplexing switch," *Optica*, vol. 2, no. 6, pp. 530–535, 2015, doi: 10.1364/OPTICA.2.000530.

[11] G. Van der Plas *et al.*, "Design issues and considerations for low-cost 3-D TSV IC technology," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 293–307, Jan. 2011.

[12] J. Chan, G. Hendry, A. Biberman, and K. Bergman, "Architectural exploration of chip-scale photonic interconnection network designs using physical-layer analysis," *J. Lightw. Technol.*, vol. 28, no. 9, pp. 1305–1315, May 1, 2010.

[13] A. Scandurra and I. O'Connor, "Scalable CMOS-compatible photonic routing topologies for versatile networks on chip," in *Proc. 1st Workshop Netw.-Chip Archit. (NoCArc)*, 2008, pp. 44–50.

[14] L. Ramini, P. Grani, S. Bartolini, and D. Bertozzi, "Contrasting wavelength-routed optical NoC topologies for power-efficient 3D-stacked multicore processors using physical-layer analysis," in *Proc. Design, Autom. Test Eur. Conf. Exhibit.*, Mar. 2013, pp. 1589–1594.

[15] X. Tan, M. Yang, L. Zhang, Y. Jiang, and J. Yang, "On a scalable, non-blocking optical router for photonic networks-on-chip designs," in *Proc. Symp. Photon. Optoelectron. (SOPO)*, May 2011, pp. 1–4.

[16] S. Koohi, M. Abdollahi, and S. Hessabi, "All-optical wavelength-routed NoC based on a novel hierarchical topology," in *Proc. NOCS*, Pittsburgh, PA, USA, 2011, pp. 97–104.

[17] S. Le Beux, J. Trajkovic, I. O'Connor, G. Nicolescu, G. Bois, and P. Paulin, "Optical ring network-on-chip (ORNoC): Architecture and design methodology," in *Proc. Design, Autom. Test Eur. Conf. Exhibit. (DATE)*, Mar. 2011, pp. 1–6.

[18] G. Hendry, J. Chan, L. P. Carloni, and K. Bergman, "VANDAL: A tool for the design specification of nanophotonic networks," in *Proc. DATE*, 2011, pp. 1–6.

[19] C. Chen, T. Zhang, P. Contu, J. Klamkin, A. K. Coskun, and A. Joshi, "Sharing and placement of on-chip laser sources in silicon-photonic NoCs," in *Proc. IEEE/ACM Int. Symp. Netw.-Chip (NoCS)*, Sep. 2014, pp. 88–95.

[20] C. Condrat, P. Kalla, and S. Blair, "Channel routing for integrated optics," in *Proc. SLIP*, 2013, pp. 1–8.

[21] D. Ding, Y. Zhang, H. Huang, R. T. Chen, and D. Z. Pan, "O-router: An optical routing framework for low power on-chip silicon nano-photonic integration," in *Proc. DAC*, 2009, pp. 264–269.

[22] D. Ding, B. Yu, and D. Z. Pan, "GLOW: A global router for low-power thermal-reliable interconnect synthesis using photonic wavelength multiplexing," in *Proc. ASPDAC*, 2012, pp. 621–626.

[23] A. K. Coskun *et al.*, "Cross-layer floorplan optimization for silicon photonic NoCs in many-core systems," in *Proc. DATE*, 2016, pp. 1309–1314.

[24] S. A. Le Beux, I. O'Connor, G. Nicolescu, G. Bois, and P. Paulin, "Reduction methods for adopting network on chip topologies to 3D architectures," *J. Microprocess. Microsyst.*, vol. 37, no. 1, pp. 87–98, 2013.

[25] 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," in *Proc. Int. Symp. Netw.-Chip (NOCS)*, May 2012, pp. 185–192.

[26] A. Boos, L. Ramini, U. Schlichtmann, and D. Bertozzi, "PROTON: An automatic place-and-route tool for optical networks-on-chip," in *Proc. ICCAD*, 2013, pp. 138–145.

[27] Y. Ye *et al.*, "3-D mesh-based optical network-on-chip for multiprocessor system-on-chip," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 32, no. 4, pp. 584–596, Apr. 2013.

[28] M. Tala, M. Castellari, M. Balboni, and D. Bertozzi, "Populating and exploring the design space of wavelength-routed optical network-on-chip topologies by leveraging the add-drop filtering primitive," in *Proc. IEEE NOCS*, Aug./Sep. 2016, pp. 1–8.

[29] S. Le Beux, H. Li, G. Nicolescu, J. Trajkovic, and I. O'Connor, "Optical crossbars on chip, a comparative study based on worst-case losses," *Concurrency Comput. Pract. Exper.*, vol. 26, no. 15, pp. 2492–2503, Oct. 2014.

[30] S. Beamer *et al.*, "Re-architecting dram memory systems with monolithically integrated silicon photonics," in *Proc. Int. Symp. Comput. Archit.*, Jun. 2010, pp. 129–140.

[31] S. Le Beux, J. Trajkovic, I. O'Connor, and G. Nicolescu, "Layout guidelines for 3D architectures including optical ring network-on-chip (ORNoC)," in *Proc. Int. Conf. VLSI Syst.-Chip (VLSI-SoC)*, Oct. 2011, pp. 242–247.

[32] P. Grani and S. Bartolini, "Design options for optical ring interconnect in future client devices," *ACM J. Emerg. Technol. Comput. Syst.*, vol. 10, no. 4, p. 30, May 2014.

[33] M. J. R. Heck and J. E. Bowers, "Energy efficient and energy proportional optical interconnects for multi-core processors: Driving the need for on-chip sources," *IEEE J. Sel. Topics Quantum Electron.*, vol. 20, no. 4, Jul./Aug. 2014, Art. no. 8201012.

[34] C. Batten, A. Joshi, V. Stojanovic, and K. Asanovic, "Designing chip-level nanophotonic interconnection networks," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 2, pp. 137–153, Jun. 2012.

[35] K. Padmaraju, D. F. Logan, T. Shiraishi, J. J. Ackert, A. P. Knights, and K. Bergman, "Wavelength locking and thermally stabilizing microring resonators using dithering signals," *J. Lightw. Technol.*, vol. 32, no. 3, pp. 505–512, Feb. 1, 2014.

[36] M. Ortín-Obón, L. Ramini, D. Bertozzi, and V. Viñals-Yufera, "Capturing the sensitivity of optical network quality metrics to its network interface parameters," *J. Concurrency Comput., Pract. Exper.*, vol. 26, no. 15, pp. 2504–2517, Jul. 2014.

[37] M. Nikdast *et al.*, "Crosstalk noise in WDM-based optical networks-on-chip: A formal study and comparison," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 11, pp. 2552–2565, Nov. 2015.

[38] A. Peano, L. Ramini, M. Gavanelli, M. Nonato, and D. Bertozzi, "Design technology for fault-free and maximally-parallel wavelength-routed optical networks-on-chip," in *Proc. 35th Int. Conf. Comput.-Aided Design (ICCAD)*, 2016, p. 3.

Marta Ortín-Obón received the M.S. and Ph.D. degrees in computer science engineering from the University of Zaragoza, Zaragoza, Spain, in 2012 and 2016, respectively.

She has been a Visiting Researcher with the University of Ferrara, Ferrara, Italy, and Intel Mobile Communications, Munich, Germany. She currently holds a post-doctoral position with the University of Zaragoza. Her current research interests include memory hierarchies and networks-on-chip, both electronic and optical.

Mahdi Tala is currently pursuing the Ph.D. degree with the University of Ferrara, Ferrara, Italy.

He also holds a research assistant position with the Integrated Optics and Photonics Laboratory, University of Ferrara. His current research interests include exploring synthesis methodologies for optical networks-on-chip, bringing physical-layer and layout awareness into the front-end design steps.

Luca Ramini received the Ph.D. degree in electrical engineering from the University of Ferrara, Ferrara, Italy, in 2014.

He was a Visiting Researcher with Columbia University, New York City, NY, USA, in 2011, and holds a post-doctoral position with the University of Ferrara and a Contract Professor with the University of Verona, Verona, Italy, from 2014 to 2016. He has been a Technical Leader of system-level crossbenchmarking efforts between optical interconnects and their electrical counterparts. He is currently a Silicon Photonics Designer with STMicroelectronics, Italy. His current research interests include emerging interconnection technologies for high speed and low power systems-on-chip with major emphasis on the design of silicon-photonic devices and optical networks.

Víctor Viñals-Yufera received the M.S. degree in telecommunications and the Ph.D. degree in computer science from the Universitat Politècnica de Catalunya, Barcelona, Spain, in 1982 and 1987, respectively.

He was an Associate Professor with the Facultat d'Informàtica de Barcelona, Barcelona, from 1983 to 1988. He is currently a Full Professor with the Informática e Ingeniería de Sistemas Department, University of Zaragoza, Zaragoza, Spain. He also belongs to the Computer Architecture Group and the I3A Institue, University of Zaragoza. His current research interests include processor microarchitecture, memory hierarchy, and parallel computer architecture.

Dr. Viñals is a member of the ACM, the IEEE Computer Society, and HiPEAC.

Davide Bertozzi received the Ph.D. degree in electrical engineering from the University of Bologna, Bologna, Italy, in 2003.

He holds an assistant professor position with the University of Ferrara, Ferrara, Italy, since 2004, where he leads the MPSoC Research Group. He has a been a Visiting Researcher at international academic institutions, such as Stanford University, Stanford, CA, USA, and large semiconductor companies, such as NEC America Labs (USA), NXP Semiconductors (Holland), STMicroelectronics (Italy), and Samsung Electronics (Korea). His current research interests include all aspects of on-chip communication. He has been actively involved in many EU-funded initiatives (projects: Galaxy, NaNoC, and vIrtical).

Dr. Bertozzi is a HiPEAC Member.