

# PROTON+: A Placement and Routing Tool for 3D Optical Networks-on-Chip with a Single Optical Layer

ANJA VON BEUNINGEN, Technische Universität München  
LUCA RAMINI, University of Ferrara  
DAVIDE BERTOZZI, University of Ferrara  
ULF SCHLICHTMANN, Technische Universität München

Optical Networks-on-Chip (ONoCs) are a promising technology to overcome the bottleneck of low bandwidth of electronic Networks-on-Chip. Recent research discusses power and performance benefits of ONoCs based on their system level design, while layout effects are typically overlooked. As a consequence, laser power requirements are inaccurately computed from the logic scheme but do not consider the layout. Within this contribution we propose PROTON+, a fast tool for placement and routing of 3D ONoCs minimizing the total laser power. Using our tool the required laser power of the system can be decreased by up to 94% compared to a state-of-the-art manually designed layout. In addition, with the help of our tool we study the physical design space of ONoC topologies. For this purpose, topology synthesis methods (e.g. global connectivity and network partitioning) as well as different objective function weights are analyzed in order to minimize the maximum insertion loss and ultimately system's laser power consumption. For the first time, we study optimal positions of memory controllers. A comparison of our algorithm to a state-of-the-art placer for electronic circuits shows the need for a different set of tools custom-tailored for the particular requirements of optical interconnects.

Categories and Subject Descriptors: J.6 [**Computer-Aided Engineering**]: Computer-Aided Design

General Terms: Algorithms, Design

Additional Key Words and Phrases: Optical Networks-on-Chip, Physical Design, Placement, Routing, 3D Architecture

## 1. INTRODUCTION

Networks-on-Chip are an attractive technology to increase the bandwidth of today's multicore systems. With shrinking size of device dimensions in nanoscale technologies, the performance of electrical interconnects of the global communication network degrade relative to the gate performances [Ho et al. 2001]. The usage of optical interconnections instead of metal wires can make unprecedented communication bandwidth available while minimizing dynamic power dissipation. Many research contributions in this field evaluate power and performance benefits at the system level [Vantrease et al. 2008][Shacham et al. 2008][Cianchetti et al. 2009][Koohi et al. 2011]. In the best case, system-level evaluations are based on the logic scheme [Chan et al. 2010] of optical Networks-on-Chip (ONoC) topologies. However, such accuracy is limited to capturing the relevant sources of optical power loss in the logic topologies. Unfortunately, logic connectivity schemes may turn out to be profoundly different from their physical implementations as an effect of placement and routing constraints, especially in a 3D setting. This aspect should not be overlooked, because the physical design significantly affects the laser power required by the optical network. The laser power should be tuned in such a way that the optical power reaching the photodetector exceeds the minimum detection threshold for error-free operation. Let the critical path of an ONoC be the optical path with highest insertion loss. Then, the required laser power of the system is essentially measured by the insertion loss of the critical path, which mainly depends on the number of the waveguide length and the number of waveguide crossings of the critical path. Because other losses such as drop loss do not depend on the layout or are negligible compared to the losses caused by waveguide crossings and long waveguides, the waveguide length and the number of crossings have to be minimized during placement and routing. The logic scheme alone cannot provide a good estimation of the waveguide length and number of crossings, because usually there are unrealistic floorplanning assumptions made, e.g. regarding the positions of initiators and targets of a physical path [Scandurra 2008]. As an effect, when laying out an optical NoC with realistic placement and routing constraints, unexpected

waveguide crossings and longer-than-expected waveguides show up, that would be impossible to be extrapolated from the logic schemes. Moreover, some logic schemes are optimized based on predefined positions of initiators and targets of the physical paths [Tan et al. 2011]. If these positions change, no conclusion about the required laser power can be given anymore. Before the first automatic placement and routing tools [Seo et al. 2005][Minz et al. 2007][Ding et al. 2009][Boos et al. 2013][Condrat et al. 2014] were proposed, the designers had to place and route ONoCs manually, which is time consuming and error prone. In this paper we present a placement and routing algorithm, which places the optical devices overlap-free inside the footprint area and routes the waveguide. Although first contributions about multiple optical layers and multi-layer optical devices have been presented in recent literature [Biberman et al. 2011b][Biberman et al. 2011a][Parini et al. 2014], we focus on systems with one optical layer to reduce manufacturing costs. All optical devices and waveguides are located on this single optical layer. Existing placement algorithms for VLSI design cannot easily be adapted to the optical interconnect, because they do not enable wire crossings. To the best of our knowledge, PROTON+ is the only automatic placement and routing tool for 3D ONoCs next to PROTON proposed in [Boos et al. 2013]. Our algorithm is an extension of PROTON with an improved crossing approximation function for the placement of PSEs and a new net order in the routing algorithm. In addition, we provide the definition of regions for the PSE placement. The distinctive features of the PROTON+ are the following:

- Flexibility: During placement and routing it is possible to minimize the propagation loss only, the crossing loss only or a combination of both, which gives the designer the possibility to tune the objective function for different use cases.
- Technology independence: the fast evolving silicon photonic technology can be reflected in the tool by simply updating few relevant parameters. In particular, as the insertion loss of waveguide crossings is improved by means of carefully engineered tapers, the proposed tool can predict implications on topology quality metrics.
- Design space exploration capability: the tool can be used to explore a large design space, which is out-of-reach for current manual layout design frameworks. First, the tool is not custom-tailored for specific connectivity patterns, therefore it lends itself to the comparative analysis of ONoC topologies. Second, topology synthesis techniques can be investigated, ranging from monolithic topologies delivering global connectivity to partitioned topologies that enable the reuse of laser sources across network partitions. Third, the locations of off-chip memory and memory controllers, the interfaces between the off-chip memory and the optical network, can be optimized based on the implications on the quality metrics of topologies. Fourth, the physical design space of topologies can be explored by exploiting the flexible tool objective function illustrated above.

Without lack of generality, the tool is applied to Wavelength-Routed Optical NoCs (WRONoCs), which are the most challenging benchmarks for a placement and routing algorithm, since their physical design differs significantly from their logic scheme. In particular, the tool is put to work with the most significant WRONoC topologies reported in the literature so far. The paper reports an extensive set of validation results for the new placement and routing tool:

- We compare automatically generated WRONoC topology layouts with known manual ones from the literature.
- We compare a couple of synthesis options for partitioned topologies, borrowing the concept of placement regions from CAD tools for VLSI design.
- We explore the implications of different mapping options of controllers for off-chip memory access across the chip boundary.
- We prove the key role of waveguide crossings for the minimization of the maximum insertion loss of a topology.

- We show the tool capability to deal with overly tight vs. overly loose placed initiators and targets.

The rest of the paper is structured as follows: In Section 2 the background and state-of-the-art of ONoCs is presented. Our placement and routing algorithm is proposed in Section 3. Experimental results are shown in Section 4, where we place and route different benchmarks, compare the results with manually designed layouts proposed in the literature and show the flexibility of our tool. Finally, a conclusion is drawn in Section 5.

## 2. STATE-OF-THE-ART AND BACKGROUND

Due to manufacturing challenges 3D-Stacking seems to be a good technology to integrate optical components into an electronic system.

### 2.1. Architecture

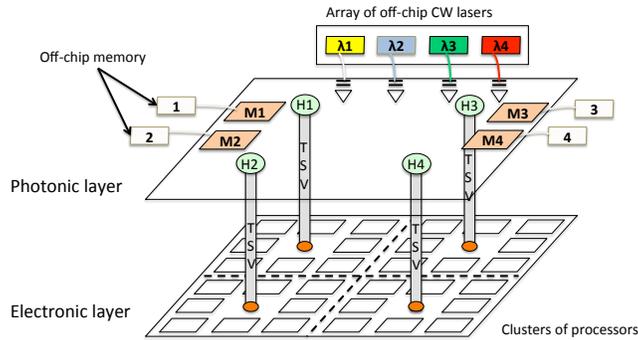


Fig. 1. Target architecture

Similar to the architecture proposed in [Wentzlaff et al. 2007] we consider a 3D system consisting of a lower, electronic layer and an upper, photonic layer, which are connected by arrays of through-silicon vias (TSVs) as can be seen in Figure 1. On the electronic layer a grid of 32 processors partitioned into 4 clusters is located. Each cluster consists of 8 cores and is connected to a network interface on the electronic layer, which guides the signals via one TSV array to the optical interface on the optical layer. We assume a core size and a die size of  $1.33mm \times 1.33mm$  and  $9mm \times 9mm$  respectively. Therefore cores occupy about  $57mm^2$  of the available  $81mm^2$ , the rest is being occupied by the on-chip electronic network (roughly  $3.5mm^2$  for the target system size), by I/O peripherals, by TSV arrays, and by the electronic circuits needed to drive the optical hubs [Ortín-Obón et al. 2014]. Three different kinds of communication have to be enabled: a) between clusters, b) from a cluster to an off-chip memory connected to the optical layer, and c) from a memory to a cluster. The memory controllers are located close to the boundary of the optical layer. Their position is determined by the location of the off-chip memory [Beamer et al. 2010]. In most of our experiments, they are located pairwise at the eastern and western boundary as can be seen in Figure 1. On top of the center point of each cluster a hub, which consists of the optical components of the interface between the electronic and the optical network, is positioned. To connect the 8 initiators (the output pin of the 4 hubs and 4 memory controllers) with the 8 targets (the input pin of the same 4 hubs and 4 memory controllers), a Wavelength-Routed Optical NoC (WRONoC) can be used. The above general scenario is compatible with different manufacturing technologies and architectural requirements that may arise in future photonically integrated systems. On one hand, no or very few electronics might be implemented in the optical plane, thus relieving the integration requirements with

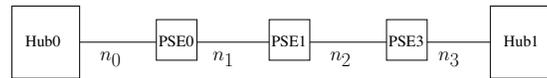


Fig. 2. Path  $p$  connects Hub0 and Hub1 and consists of four two-pin-connections  $n_0, \dots, n_3$ , which connect a hub and a PSE or two PSEs.



Fig. 3. The PSE is sensitive to wavelength  $\lambda_1$ . (a) A signal with wavelength not equal to  $\lambda_1$  crosses the PSE. (b) A signal with wavelength  $\lambda_1$  is coupled into the ring, intensified by resonance processes and decoupled into another waveguide.

photonic devices. From an architectural viewpoint, this corresponds to the case, where memory controllers reside on the processing plane and have dedicated or multiplexed hubs in the optical plane. Also, DRAM memory channels should be nodes of the optical network. On the other hand, the maturity of monolithic integration might enable a use case where memory controllers reside on the optical plane as well, in line with previously published architectures [Beamer et al. 2010][Udipi et al. 2011]. Therefore, without lack of generality, in what follows we assume that the photonic layer is 3D-integrated with a purely electrical layer designed using a traditional monolithic CMOS process. Architecture blocks are on such electronic layer, including serializers and deserializers. Then, we assume that electrical circuits for E/O and O/E conversions (e.g., drivers, photodetectors, transimpedance amplifiers, digital comparators) are split between the electronic and the optical layer based on the maturity of monolithically integrated optical technologies, and on performance/energy considerations (e.g., parasitic capacitance of the photodetector) that go beyond the scope of this paper.

## 2.2. Wavelength-Routed Optical NoCs

In Wavelength-Routed Optical NoCs the principle of Wavelength-Selective Routing is used. Let a path be the connection between one initiator and one target. Each path starting at the same initiator uses a specific wavelength to reach a specific target. Besides the initiator and target a path consists of several Photonic Switching Elements (PSEs) as can be seen in Figure 2. These are optical switches containing 2 microring resonators (MRRs) and either 2 input and 2 output ports (referred to as 2x2 optical switch) or one MRR and one input and 2 output ports (referred to as 1x2 optical switch). In our work we focus on passive PSEs. Depending on the wavelength of the signal entering the PSE, it is either coupled into the ring, intensified by interference and decoupled into the waveguide leaving the PSE at the port rotated by 90 degrees to the input port. This is referred to as drop function. Or the signal crosses the PSE. Both behaviors are shown in Figure 3. Active PSEs would need control devices, which activate or deactivate each PSE before sending a signal on a specific path. Using passive PSEs allows us to define a contention-free, fully connected network without the overhead of control devices. In addition, no time is spent on the (de-)activation of PSEs. In our experiments we assume a PSE size of  $70\mu\text{m} \times 70\mu\text{m}$  [Sherwood-Droz et al. 2008]. Since we have to connect 8 initiators and 8 targets in our use case, we need an 8x8 WRONoC. Below we present the most important 8x8 WRONoC topologies proposed in the literature

- The 8x8  $\lambda$ -Router [Scandurra 2008] is illustrated in Figure 4(a). The 8 initiators are shown on the left hand side, and the 8 targets can be seen on the right hand side.

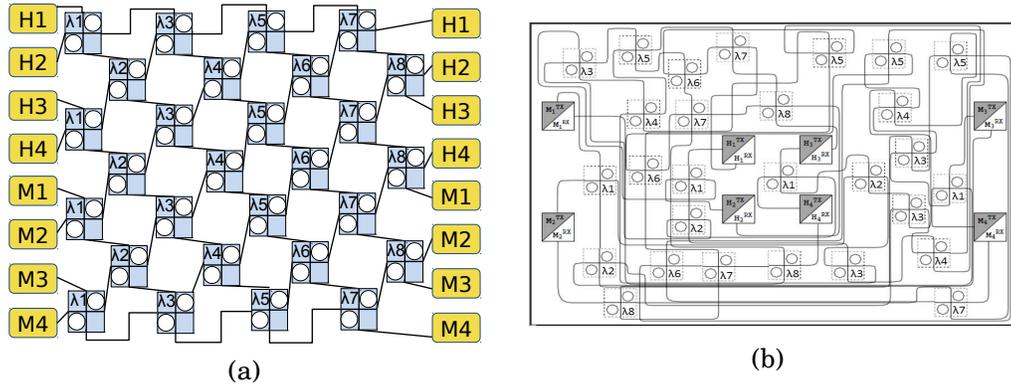


Fig. 4. (a) Logic scheme and (b) manually designed layout [Ramini et al. 2012] of the 8x8  $\lambda$ -Router

The initiators and targets are connected via 28 2x2 optical switches and 8 different wavelengths.

- The 8x8 Generic Wavelength-Routed Optical Router (GWOR) can be obtained based on the wavelength assignment of the 4x4 GWOR proposed in [Tan et al. 2011]. The 8x8 GWOR includes 24 2x2 optical switches and needs 7 different wavelengths. In contrast to the 8x8  $\lambda$ -Router, the unnecessary paths of self-communication are not provided.
- In contrast to the 8x8  $\lambda$ -Router and the 8x8 GWOR, the 8x8 Standard Crossbar [Gu et al. 2009] uses 64 1x2 optical switches. The 8x8 Standard Crossbar works with 8 different wavelengths.

### 2.3. Maximum Insertion Loss

With the help of the physical design the static power consumption of the optical NoC can be calculated. A measurement of the laser power required by the system is the maximum insertion loss over all paths  $p \in P$ , where one path connects an initiator with a target via several PSEs. The insertion loss mainly is determined by six different losses:

- The propagation loss occurs when the optical signal propagates through the waveguide due to light scattering at the waveguide sidewalls.
- The crossing loss occurs when the optical signal goes through any waveguide crossings.
- The bending loss occurs when the optical signal propagates along a waveguide bend of typically 90 degrees curvature.
- The drop loss describes the loss that occurs if the signal is coupled into the microring and changes its direction inside a PSE.
- The modulator loss occurs, when the optical signal is modulated at the source before being propagated into the network waveguide [Ophir and Bergman 2013].
- The photodetector loss occurs, when the optical signal is detected at the target and converted into the corresponding electrical counterpart [Beamer et al. 2010].

The sum of all six losses in one path  $p \in P$  is the insertion loss of path  $p$ . The maximum insertion loss over all paths is named the *maximum insertion loss*  $il_{max}$ , which determines the total laser power required by the ONoC. In our study we do not consider the modulator and photodetector loss, because they are independent of the layout. Because the drop loss does not depend on the placement and routing, and there is at least one drop per path, the drop loss is also not considered during layout synthesis. As can be seen in Table I the bending loss is very small compared to the propagation loss and crossing loss. Thus, it will be neglected in our placement algorithm. [Simi-](#)

Table I. Loss parameters published in [Chan et al. 2010]

Parameters	Value
Propagation loss [Chan et al. 2010]	1.5 dB/cm
Bending loss [Chan et al. 2010]	0.005 dB
Crossing loss	0.15 dB
Drop loss	0.5 dB

larly, other loss parameters that do affect quality metrics of optical NoC topologies have been omitted from Table I since they do not contribute to discriminate between alternative layout solutions (e.g., coupling loss of fibers to on-chip waveguides). During routing the number of bends is considered, when there are two or more possible routing paths between an initiator and a target suffering from the same approximated insertion loss, which is calculated from the number of crossings and the routing path length. Then, the path with lower number of bends is chosen. Last but not least, recent papers are reporting crossing losses as low as 0.04 dB [Liu et al. 2014]. However, this does not mean that we can omit the number of crossings from the objective function of placement and routing frameworks due to the following reasons:

- Low losses are typically traded off with a larger footprint.
- Crossings remain a source of crosstalk noise.
- Such low-losses are not constant over the bandwidth, where they may vary even by an order of magnitude.
- Process parameter variations may more than double the crossing loss.
- Propagation losses are an active area of improvements too, and losses as low as 0.274 dB/cm have been reported [Dong et al. 2010].

Hence, PROTON+ aims at practical relevance for technology assumptions that go beyond the parameter values reported in Table I, which are specified to enable concrete experiments.

#### 2.4. Logic Scheme vs. Physical Layout

Optical NoC topologies usually are proposed by their logic scheme in recent literature. Although the laser power consumption of the ONoC topologies often is calculated based on the logic scheme, only the physical design can be used to accurately calculate the laser power required by the system, because the logic scheme does not consider the real positions of hubs, which are determined by the electronic layer, and the real positions of memory controllers, which are determined by the placement of the off-chip memory. Figure 4(b) shows a manually created layout of the 8x8  $\lambda$ -Router proposed in [Ramini et al. 2012]. Please notice that the dimensions of the PSEs, hubs and memory controllers in Figure 4(b) do not match the real dimensions. In contrast to the logic scheme shown in Figure 4(a), the real positions of hubs and memory controllers were considered. Due to complexity reasons only the crossing loss was aimed to be minimized during the placement of PSEs and the routing of waveguides. Although the 8x8  $\lambda$ -Router seems to be a small topology without any crossings outside the PSEs, it is hard to determine an optimal layout manually, when the hubs are positioned in the center, and the memory controllers are positioned at the boundary of the chip. The highest number of crossings (considering the waveguide crossings within and outside the PSEs) within one path increases from 7 in the logic scheme to 64 in the physical layout. As a consequence, the maximum insertion loss of the real ONoC is significantly higher than the maximum insertion loss calculated based on the logic scheme. On one hand, the manually created layout can help to make more realistic assumptions about the laser power consumption of ONoCs. On the other hand, placing and routing manually is error-prone and time consuming, especially when the number of initiators and targets and therefore the number of PSEs scale up. Hence, algorithms are needed for the automatic physical design of ONoCs.

## 2.5. Previous Work about the Physical Design of ONoCs

Conventional placement and routing algorithms for the design of electronic ICs or NoCs can not be used for the design of ONoCs. In contrast to electronic ICs, only one layer is used for placing the optical components and routing the waveguides. In the electronic design several metal layers are used to avoid unintended crossings of the wires. The crossing of waveguides and thereby of light signals does not influence the behavior but leads to a higher laser power consumption of the system. Thus, new algorithms for the physical design of the optical layer of ONoCs are needed, which enable waveguide crossings but minimize them at the same time to minimize the system's laser power consumption. Since no routing algorithm for wire routing enables the functionality of crossings, new routing methods have to be developed. Because routing results strongly depend on the placement, the number of possible waveguide crossings should be minimized during the placement step as well. No algorithm for the placement of electronic modules considers or minimizes the number of wire crossings. Thus, new placement algorithms for ONoCs are needed too. The authors of [Seo et al. 2005] and [Seo and Chatterjee 2002] presented placement and routing algorithms for 2D-optoelectronic systems-on-packages (SoPs) and systems-on-chip (SoC) minimizing the delay of the SOP or SoC but did not consider the laser power consumption of the system. A routing algorithm for 3D optical SOPs is presented by [Minz et al. 2007], where the delay of the entire system is reduced with the help of optical waveguides. Some parts of nets in the critical path are selected to be rerouted as optical signals, but no PSE placement or laser power consumption of the system is considered. Condrat et al. [Condrat et al. 2014] presented a channel routing algorithm for optical interconnects but assumed the placement of all optical devices to be given. A routing algorithm for 3D optical NoCs without the placement of the PSEs was introduced in [Ding et al. 2009]. The synthesis of 3D hybrid nanophotonic-electric NoC architectures, which combine electrical NoCs with free-space nanophotonic interconnect NoCs, was conducted in [Bahirat and Pasricha 2014]. Since no waveguides are needed to interconnect one core to another, the routing problem was not required. A multi-stage hybrid NoC synthesis framework exploiting technologies resources to adapt to the application bandwidth requirements was instead explored in [Hammami and Hamwi 2013]. Here, NoCs were placed and routed automatically layer by layer on the electronic part while manually on the optical part. The logic synthesis for integrated optics was addressed in [Condrat et al. 2011], although no implication at layout level was analyzed. In [Boos et al. 2013] the first automatic placement and routing algorithm for 3D optoelectronic NoCs minimizing the total laser power of the system was presented. During placement waveguide crossings are predicted with the help of a Gaussian function. The algorithms of [Boos et al. 2013] are extended and improved within this contribution. The crossing prediction function is improved. The new function ensures symmetry, e.g. the probability of net 1 crossing net 2 is equal to the probability of net 2 crossing net 1. Moreover, the new feature, which allows to define placement regions, helps the designer to influence the resulting layouts. Furthermore, the net order in the routing algorithm is improved by considering the crossing prediction results of the placement step. For the first time, the algorithms are used to study the optimal position of hubs and memory controllers and to place and route network-partitioned topologies. In addition, we compare our results obtained by PROTON+ with a state-of-the-art placer and router to confirm the need of our placement and routing algorithm.

## 3. PLACEMENT AND ROUTING ALGORITHM

In this section we introduce our placement and routing algorithm used by PROTON+.

### 3.1. Selection of the objective function

One major pillar of any optimization framework consists of selecting a suitable objective function. Intuitively, the use of multiple wavelength channels in a wavelength-routed topology might suggest to minimize the sum (over the number of wavelength

channels) of the worst-case insertion losses among all the optical paths that switch the optical signal at a given wavelength. This way, each laser source could be tuned individually and the total laser power minimized. Unfortunately, this has deep and non-trivial implications on both the optical power generation mechanism and on the power distribution network. First, this would require independent and individually tunable DFB laser sources multiplexed onto the same input WDM signal for the optical NoC. However, viability of this approach is questionable for future photonic integration of high-end embedded systems, since multiwavelength laser sources are far more compact as off-chip laser sources [Heck and Bowers 2014]. They can directly generate a spectral comb, although they do not easily enable power tuning of wavelength channels individually. Second, bringing unmodulated optical signals to each hub with custom-tailored power levels for them (reflecting insertion loss of optical paths originating from them) requires a complex power distribution network, consisting of cascaded splitters with splitting ratios, which should be custom-tailored to the specific initiator and/or wavelength channel. In contrast, the most common splitter device is the one, which equally splits the input power into two halves. Overall, in this work we did not want to restrict the objective function to overly specific or still unconsolidated technology assumptions, and decided to go for an objective function that reflects the use of comb laser sources and of 3-dB power splitters. As a consequence, we target the minimization of the worst case insertion loss across all optical paths, regardless of the wavelength channel they use. We leave the extension to more specific and/or newer combinations of laser sources and power distribution networks as future work, although we anticipate that the minimization of power levels on an initiator- and/or wavelength channel-basis requires a multiobjective optimization framework. **One key finding of this paper is that the objective functions for the placement and routing of optical NoCs are tightly inter-related with the solutions chosen for the generation and distribution of optical power.**

Given the chip area, the dimensions and positions of hubs and memory controllers, the dimensions of PSEs and the connectivity (e.g. the list of paths), PROTON+ places all PSEs overlap-free inside the chip area and routes the waveguides on the same layer. In both steps the maximum insertion loss  $il_{max}(\mathbf{x}, \mathbf{r})$  over all paths is minimized:

$$\text{minimize } il_{max}(\mathbf{x}, \mathbf{r}) = \max_{p \in P} il_p(\mathbf{x}, \mathbf{r}) \quad (1)$$

where  $P$ ,  $\mathbf{x}$  and  $\mathbf{r}$  are the set of all paths, the positions of all PSEs and the positions of the waveguides, respectively. As discussed in Section 2.3, the insertion loss mainly depends on the propagation loss, the crossing loss, the bending loss, and the drop loss. Because the bending loss is very small compared to the crossing loss and propagation loss, it will not be considered during placement and routing. In addition, the drop loss of each path is independent of the physical design of the ONoC. Thus, it is a constant number in the objective function, which does not have to be considered during minimization. Minimizing the propagation loss and minimizing crossing loss can be contradictory objectives, e.g. sometime it is useful to allow the waveguide making a small detour, if a waveguide crossing can be avoided with the help of this detour. Hence, the weighted sum of propagation loss and crossing loss is minimized during placement and routing to find a trade-off of both.

$$il_{max}(\mathbf{x}, \mathbf{r}) = \max_{p \in P} (\alpha \cdot pl_p(\mathbf{x}, \mathbf{r}) + \beta \cdot cl_p(\mathbf{x}, \mathbf{r})), \quad (2)$$

where we assume  $\alpha, \beta \in \mathbb{R}_0^+$  and (without loss of generality)  $\alpha + \beta = 1$ . The parameters  $pl_p$  and  $cl_p$  describe the propagation loss and crossing loss of path  $p$ , respectively. The weights  $\alpha$  and  $\beta$  can be chosen by the designer to tune our tool. In general, the actual weights are calculated based on the physical loss parameters given in Table I. Depending on the choice of the weights another layout is produced by PROTON+. Due to the complexity of placement and routing, all PSEs are placed first, before the waveguides are routed in a second step.

### 3.2. Placement

Because the positions of the waveguides are not available during placement, the waveguide length and the number of waveguide crossings have to be approximated for each path. The waveguide length  $L_p(\mathbf{x}, \mathbf{r})$  of a path  $p$  is approximated using a quadratic net model. Let  $(x_i, y_i)$  be the position of module  $i$ , where a module is a PSE, hub or memory controller. The approximated waveguide length  $\tilde{L}(x_i, y_i)$  between module  $i$  and module  $j$  is defined as

$$\tilde{L}(x_i, y_i) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (3)$$

Consequently, the approximated waveguide length  $\tilde{L}_p(\mathbf{x})$  of path  $p$  is the sum of the approximated waveguide lengths over all two-pin-connections  $(i, j)$  in path  $p$

$$\tilde{L}_p(\mathbf{x}) = \sum_{(i,j) \in N_p} \tilde{L}(x_i, y_i), \quad (4)$$

where  $N_p$  is the number of all two-pin-connections in path  $p$ , and a two-pin-connection connects exactly two modules. In the optical topologies studied in this contribution all nets are two-pin-connections. In contrast to the placement of electronic modules, the number of possible waveguide crossings has to be considered during the placement of the optical devices to improve the routing results and thereby the maximum insertion loss. For the approximation of the number of waveguide crossings in path  $p$  a heuristic is used. Assume that net  $n_{ij} = (i, j)$  connects module  $i$  and module  $j$ , and net  $n_{rs} = (r, s)$  connects module  $r$  and module  $s$ . Let  $l_{ij}$  and  $l_{rs}$  be the straight lines connecting the center points of modules  $i$  and  $j$  and of modules  $r$  and  $s$ , respectively. For each pair of nets the probability of a crossing of their waveguides has to be estimated. If  $l_{ij}$  and  $l_{rs}$  are located close to each other, we assume that the probability of a crossing of their waveguides will be high. If the two lines are located far away from each other, we assume that the probability of a crossing of their waveguides will be low. For the mathematical formulation of this heuristic a Gaussian function is used. As shown in

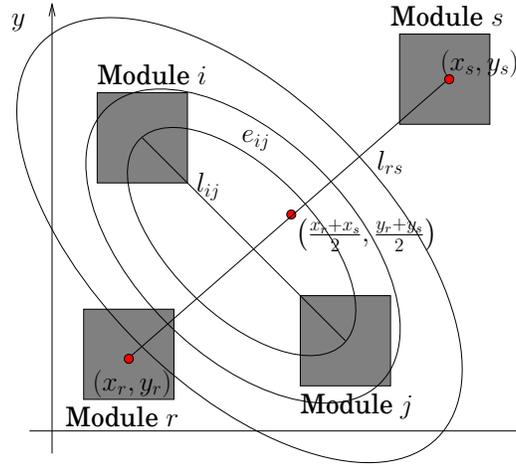


Fig. 5. Net  $n_{ij}$  between module  $i$  and module  $j$  is approximated by line  $l_{ij}$ . Around  $l_{ij}$  the ellipse  $e_{ij}$  is defined.

Figure 5, we define an ellipse  $e_{ij}$  around line  $l_{ij}$  with length  $\|(x_i, y_i) - (x_j, y_j)\|_2 = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ , where  $\|(x_i, y_i) - (x_j, y_j)\|_2$  is the Euclidean distance between the center points of module  $i$  and module  $j$ . The width of the ellipse is assumed to be

$\alpha \|(x_i, y_i) - (x_j, y_j)\|_2$  with  $0 < \alpha \leq 1$ . This ellipse can be interpreted as the level curve of the Gaussian function

$$f_{ij}(x, y) = ae^{b_{11}(x-x_c)^2 - 2b_{12}(x-x_c)(y-y_c) + b_{22}(y-y_c)^2}, \quad (5)$$

shown in Figure 6, where  $a$ ,  $b_{11}$ ,  $b_{12}$  and  $b_{22}$  are chosen as

$$a = 1 \quad (6)$$

$$d_1 = \frac{\ln\left(\frac{TOL}{a}\right)}{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (7)$$

$$d_2 = \frac{1}{\alpha^2} \frac{\ln\left(\frac{TOL}{a}\right)}{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (8)$$

$$x_{min} = y_i < y_j ? x_i : x_j \quad (9)$$

$$y_{min} = \min(y_i, y_j) \quad (10)$$

$$\beta = \arccos \left[ \frac{|x_{min} - x_c|}{\sqrt{(x_{min} - x_c)^2 + (y_{min} - y_c)^2}} \right] \quad (11)$$

$$b_{11} = d_1 \cos^2(\beta) + d_2 \sin^2(\beta) \quad (12)$$

$$b_{12} = (d_1 - d_2) \sin(\beta) \cos(\beta) \quad (13)$$

$$b_{22} = d_1 \sin^2(\beta) + d_2 \cos^2(\beta). \quad (14)$$

Those parameters are derived to ensure the following two conditions: The Gaussian function  $f_{ij}(x, y)$  reaches its maximum at the center point  $(x_c, y_c) = \left(\frac{x_i+x_j}{2}, \frac{y_i+y_j}{2}\right)$  of line  $l_{ij}$  with the function value  $f(x_c, y_c) = a = 1$ . The function values of the Gaussian function at the boundary of the ellipse and especially at the center points of modules  $i$  and  $j$  are  $TOL$ , where we assume a tolerance of  $TOL = \frac{a}{2}$  and  $\alpha = 1$  in our experiments. The three degrees of freedom  $a$ ,  $TOL$  and  $\alpha$  are chosen based on an experimental analysis. The probability of a crossing  $c_{ij}(x_r, y_r, x_s, y_s)$  of the waveguides of nets  $n_{ij}$  and  $n_{rs}$  is estimated by a weighted sum of the function values of the Gaussian function at the two end points and the center point of line  $l_{rs}$ . The weights are determined with the help of Simpson's rule [Press et al. 2007]:

$$c_{ij}(x_r, y_r, x_s, y_s) = \frac{\sqrt{(x_r - x_s)^2 + (y_r - y_s)^2}}{6} \left( f_{ij}(x_r, y_r) + 4f_{ij}\left(\frac{x_r + x_s}{2}, \frac{y_r + y_s}{2}\right) + f_{ij}(x_s, y_s) \right) \quad (15)$$

Because the probability  $c_{ij}(x_r, y_r, x_s, y_s)$  of a waveguide crossing of nets  $n_{ij}$  and  $n_{rs}$  should be equal to the probability  $c_{rs}(x_i, y_i, x_j, y_j)$  of a waveguide crossing of nets  $n_{rs}$  and  $n_{ij}$ , the final crossing probability is calculated as

$$c_{ijrs}(\mathbf{x}) = \sqrt{c_{ij}(x_r, y_r, x_s, y_s)c_{rs}(x_i, y_i, x_j, y_j)}. \quad (16)$$

In addition to the objective function, some constraints of the minimization problem (1) have to be formulated as almost everywhere differentiable functions. Let  $w_k$  and  $h_k$  be the width and the height of module  $k$ , respectively. The constraints

$$g_{ij}(\mathbf{x}) \geq 0 \quad \text{with} \quad g_{ij}(\mathbf{x}) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} - \sqrt{w_i^2 + h_i^2} - \sqrt{w_j^2 + h_j^2} \quad (17)$$

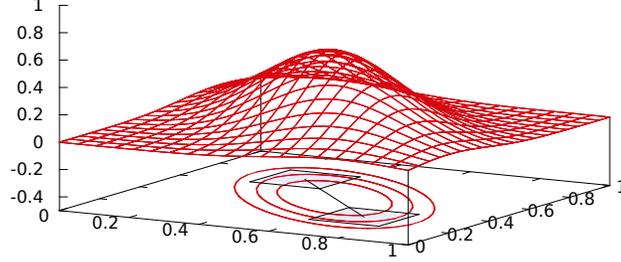


Fig. 6. The ellipse  $e_{ij}$  around  $l_{ij}$  is a contour line of a Gaussian function, whose function values are high close to  $l_{ij}$  and low elsewhere.

ensure a non-overlapping placement by defining circles around each module, which are not allowed to intersect each other. The equations

$$\frac{w_i}{2} \leq x_i \leq w_c - \frac{w_i}{2} \quad (18)$$

$$\frac{h_i}{2} \leq y_i \leq h_c - \frac{h_i}{2} \quad (19)$$

ensure that all PSEs are placed within the footprint area  $[0, w_c] \times [0, h_c]$  of the chip, where  $w_c$  and  $h_c$  are the width and the height of the optical layer. The placement problem is now formulated as

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} && \max_{p \in P} && \sum_{n_{ij}=(i,j) \in N_p} && \left[ \alpha((x_i - x_j)^2 + (y_i - y_j)^2) \right. \\ & && && && \left. + \beta \sum_{n_{rs} \in N, n_{rs} \neq n_{ij}} c_{ijrs}(\mathbf{x}) \right] \\ & \text{subject to} && && && \frac{w_i}{2} \leq x_i \leq w_c - \frac{w_i}{2} \quad \forall i = 1, \dots, m \\ & && && && \frac{h_i}{2} \leq y_i \leq h_c - \frac{h_i}{2} \quad \forall i = 1, \dots, m \\ & && && && g_{ij}(\mathbf{x}) \geq 0 \quad \forall i, j = 1, \dots, m, i \neq j, \end{aligned} \quad (20)$$

where  $m$  is the total number of PSEs, hubs and memory controllers. The problem (20) is solved by an interior point method proposed by Wächter et al. [Wächter and Biegler 2006]. This iterative method needs a convex and differentiable objective function as well as convex and differentiable constraints. As a concatenation of differentiable functions, the crossing estimation (16) and the constraint (17) are differentiable except for all points with

$$x_i = x_j \text{ and } y_i = y_j \text{ for any } i \neq j. \quad (21)$$

This equation corresponds to the situation, where the center points of module  $i$  and module  $j$  are equal. Hence, the two modules lie on top of each other. These points do not fulfill constraint (17) and thus are not valid as a solution of our optimization problem. In practice, when choosing a starting point with

$$x_i \neq x_j \text{ or } y_i \neq y_j \text{ for all } i \neq j \quad (22)$$



Fig. 7. Minimizing (a) propagation loss or (b) crossing loss can result in an increase of crossing loss or propagation loss, respectively.

the situation of equation (21) never occurred in our experiments.

### 3.3. Routing

After all PSEs are placed, the waveguides have to be routed. For simplification reasons we allow waveguides to run horizontally and vertically only. Because only a single layer is available for placing the PSEs and routing the waveguides, crossings of waveguides often are not avoidable. Since each crossing in a path increases the insertion loss of this path, the number of waveguide crossings has to be minimized during routing. In addition, the waveguide length should be as short as possible. On one hand, the minimization of propagation loss and the minimization of crossing loss can be contradictory constraints as illustrated in Figure 7. In Figure 7(a) propagation loss, e.g. the waveguide length, is minimized, which results in one crossing. In Figure 7(b) crossing loss is minimized, which results in no crossing but a significant longer waveguide length. On the other hand, a short waveguide length can decrease the number of crossings. If a lot of detours are taken, and the waveguides are excessively long, there exist more possibilities for crossings. Thus, the algorithm prefers short waveguides and avoids crossings only, if there are short detours available.

For routing we use a modified version of a maze router [Lee 1961]. Maze routers as well as all routing algorithms for the design of electronic systems do not allow crossings of wires to prevent unintended behavior of the system. We adapt the maze router to the routing of ONoCs by enabling but penalizing waveguide crossings to reduce the number of their occurrence.

First, all paths are split into nets, which are routed sequentially. In contrast to [Boos et al. 2013], where the shortest nets are routed first, and the longest nets are routed last, we use the crossing estimation calculated in the placement step to determine the net order. Nets with lowest probability of crossings with all other nets are routed first. By applying this strategy, nets with a probably high number of crossings can detour the other nets and avoid to be crossed by the nets with a low crossing probability. For routing, the chip is overlaid by a grid. In our experiments a grid bin size of  $9\mu\text{m} \times 9\mu\text{m}$  is used to ensure the minimum distance of  $5\mu\text{m}$  between two waveguides to avoid crosstalk. This grid bin size corresponds to a grid of  $1000 \times 1000$  bins. All grid bins containing obstacles, e.g. PSEs, hubs or memory controllers, are marked as not available. All pins are mapped to the next grid bin. Then, the nets are routed sequentially. As shown in Figure 8(a), a '0' is assigned to the bin 'S1' containing the starting pin of the first net. A '1' is assigned to the four neighbor bins of this starting bin. The neighbors of these neighbors get a '2' and so on. This strategy is continued until the bin 'T1' belonging to the target pin is reached. In the next step the algorithm determines the shortest way back from bin 'T1' to bin 'S1'. If there exists more than one shortest way back, at each junction our algorithm prefers the direction without a bend to minimize the bending loss. After the first net is routed all bins used by the first net are marked. In the original algorithm these bins are blocked and cannot be used by further nets. Since waveguide crossings have to be enabled in ONoCs but should be minimized, we mark them and penalize a further usage as shown in Figure 8(b). The algorithm proceeds with the next net until all nets are routed. For penalization

we use an adaptive term. As can be seen in Figure 9(a) a constant penalty term can increase the number of crossings. Here, a penalty term of 10 is chosen. After net  $n_1$  has been routed, the second net avoids a crossing with net  $n_1$  by making a detour. Because of this detour there is no possibility for net  $n_3$  to avoid at least a crossing with  $n_2$ . In addition, a crossing with net  $n_1$  can only be avoided by the acceptance of a long detour, which increases the propagation loss. Thus, the insertion loss especially in net  $n_3$  is high. In general, if the first nets to be routed already make long detours to avoid a crossing, a lot of bins are already blocked when routing the last nets, and the probability of crossings is high. On the other hand, if the penalization term is too small, crossings will not be avoided. Thus, a low penalization term for the first nets forces them to be routed as short as possible. Increasing the penalization term for later nets avoids crossings for these last nets, which have a high estimated number of crossings with our choice of net order. In Figure 9(b) a penalty term of 5 is chosen for net  $n_2$ , and a penalty term of 10 is chosen for net  $n_3$ . The usage of the adaptive term results in only one crossing (between  $n_1$  and  $n_2$ ). Table II shows an experimental study about the choice of the penalty term. The columns *start*, *before limit*, *limit*, and *after limit* show the penalty term for the first net, the term added to the starting penalty term for each further net  $k$  (with  $k = 2, \dots, \textit{limit}$ ), a limit, and the number added to the previous penalty term for each further net  $k$  (with  $k = \textit{limit} + 1, \dots, \#nets$ ), respectively. The parameters  $\#nets$ ,  $\beta$ , and  $k$  name the number of nets, the crossing loss weight of equation (2), and the number of the current net to be routed, respectively. The last three columns show the maximum insertion loss for three different topologies when minimizing both, propagation loss and crossing loss, where the parameters  $\alpha$  and  $\beta$  in equation (2) were chosen based on the loss parameters proposed in Table I. Increasing the start penalty term results in an increase of the maximum insertion loss for two out of three topologies. Thus, a start term of 5 is chosen for the first net. Adding a nonzero term for each net before the limit, forces the algorithm to make a lot of detours. Thus, the probability of waveguide crossings for further nets strongly increases, and therefore the maximum insertion loss is high. A similar observation can be made if the limit is set too low. Adding a constant term for each net after the limit, produces good results that can be improved by adding an adaptive term dependent on  $\beta$  and the number of nets that have already been routed. On average, the adaptive term produces a bit better results, because a smaller penalty term for the nets close to the limit forces them to make small detours. This prevents a blocking of too many bins and thereby further crossings with later routed nets. In our experiments we use the parameters that are marked gray in Table II, because on average they produce the best results for the three topologies under consideration. For each net our routing algorithm has a worst-case runtime of  $O(m_1 m_2)$ , where  $m_1 m_2$  is the number of grid bins. Because most modules are positioned close to each other, the runtime in practice is much faster. In all of our experiments a grid size of  $1000 \times 1000$  bins was suitable. If the number of nets increases strongly, the number of grid bins has to be increased as well. To ensure low crosstalk, the grid bin size should be larger than  $5.4\mu m \times 5.4\mu m$ , where  $5\mu m$  is the distance between waveguides to reduce crosstalk [Donzella et al. 2013], and the width of a waveguide is assumed to be  $400nm$ . In this paper we assume a fixed crossing footprint, since we view waveguide intersections as basic and proven "cells" provided by the foundry. In general, it would be possible to consider multiple intersection cells featuring different tapers, i.e., different trade-offs between crossing footprint and crossing loss. This is left for future work. In this paper, we consider two kinds of crossings: those that are inside photonic switching elements, and those that are outside and are targeted by the proposed crossing minimization framework.

### 3.4. Definition of Regions

To reduce the maximum insertion loss of ONoCs, the authors of [Ramini et al. 2012] proposed a network partitioning of the optical topology. In particular, the global topology with 8 initiators and 8 targets is partitioned into three subnetworks: (a) a memory request subnetwork, (b) a memory response subnetwork, and (c) a subnetwork for

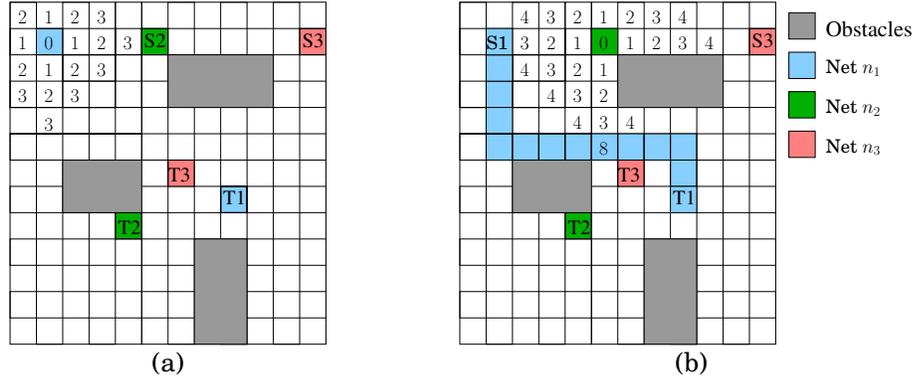


Fig. 8. (a) Routing is done by spreading wavelike over the chip area. (b) A penalty term (e.g. penalty term = 5) penalizes an already used bin.

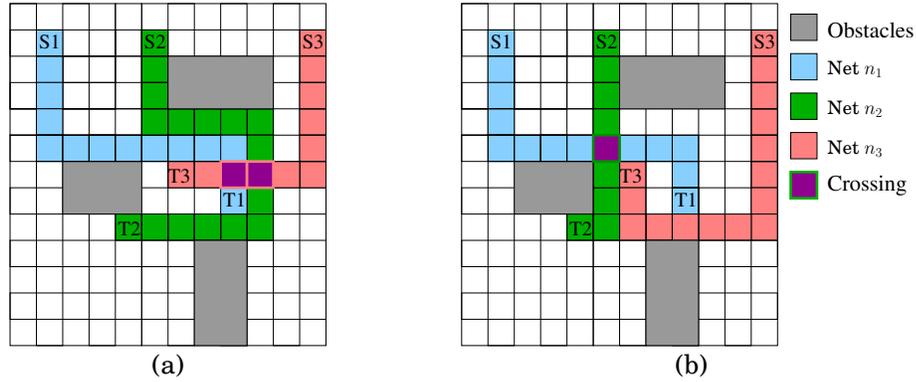


Fig. 9. A (a) constant penalty term can increase the number of crossings compared to an (b) adaptive term.

Table II. Experimental analysis for different penalty terms

parameters				maximum insertion loss		
start	before limit	limit	after limit	8x8 $\lambda$ -Router	8x8 GWOR	8x8 Standard Crossbar
5	0	0	0	8.6	8.8	8.7
20	0	0	0	8.3	9.3	8.7
50	0	0	0	8.3	9.2	8.9
100	0	0	0	9.0	9.2	9.1
5	5	0	5	9.2	9.0	13.1
5	20	0	20	9.2	13.1	14.9
5	50	0	50	9.2	13.1	14.9
5	100	0	100	9.2	13.1	14.9
5	0	$\max(0, \#nets - 10)$	$10k\beta$	8.4	8.4	8.5
5	0	$\max(0, \#nets - 20)$	$10k\beta$	10.0	9.8	8.8
5	0	$\max(0, \#nets - 30)$	$10k\beta$	9.2	12.1	9.6
5	0	$\max(0, \#nets - 40)$	$10k\beta$	9.2	12.1	9.6
5	0	$\max(0, \#nets - 50)$	$10k\beta$	9.2	13.1	12.2
5	0	$\max(0, \#nets - 60)$	$10k\beta$	9.2	13.1	12.5
5	0	$\max(0, \#nets - 10)$	1	8.6	8.8	8.7
5	0	$\max(0, \#nets - 10)$	5	8.6	8.8	8.7
5	0	$\max(0, \#nets - 10)$	$k\beta$	8.6	8.4	8.5
5	0	$\max(0, \#nets - 10)$	$5k\beta$	8.6	8.4	8.5
5	0	$\max(0, \#nets - 10)$	$20k\beta$	8.4	8.4	8.5
5	0	$\max(0, \#nets - 10)$	$50k\beta$	8.4	11.1	8.5
5	0	$\max(0, \#nets - 10)$	$100k\beta$	8.4	12.3	8.5

the inter-cluster communication. For the inter-cluster communication we use a 4x4 GWOR. The basic cell of a GWOR topology is a 4x4 optical crossbar that delivers an optical output to each cardinal point without any additional crossing loss other than those that are inside the PSEs. Therefore, such cell is the ideal solution for inter-cluster communication. For the memory request and response subnetworks a 4x4  $\lambda$ -Router is used, consisting of four initiators and four targets only. The inter-cluster communication is aimed to be placed inside the square, whose corners are the center points of the four hubs. All PSEs belonging to the request network should be placed left of this square, and all PSEs belonging to the response network should be placed right of it. This partitioning of the die area results in fewer conflicts between waveguides. To support this feature PROTON+ is able to deal with regions. For each PSE a specific region can be defined, where the PSE has to be placed. This will add new constraints to our minimization problem given in equation (20). The routing algorithm is not affected by the regions, because waveguides are allowed to be placed all over the chip. Since next to the number of waveguide crossings the waveguide length is minimized during routing, the waveguides are mainly routed in the area around the modules they are connected to and do not unnecessarily leave the defined regions their modules belong to.

### 3.5. Calculation of the Maximum Insertion Loss

After placement and routing, the maximum insertion loss of the ONoC has to be calculated. PROTON+ automatically counts the number of waveguide crossings in each path and determines the length of the waveguides. In addition, the number of crossings and drops inside the PSEs are determined. Using the loss parameters proposed in Table I, the propagation loss  $pl_p(\mathbf{x}, \mathbf{r})$ , the crossing loss  $cl_p(\mathbf{x}, \mathbf{r})$ , and the drop loss  $dl_p(\mathbf{x}, \mathbf{r})$  of each path  $p$  can be calculated as

$$pl_p(\mathbf{x}, \mathbf{r}) = 1.5 \frac{dB}{cm} \cdot L_p \quad (23)$$

$$cl_p(\mathbf{x}, \mathbf{r}) = 0.15dB \cdot C_p \quad (24)$$

$$dl_p(\mathbf{x}, \mathbf{r}) = 0.5dB \cdot D_p \quad (25)$$

$$bl_p(\mathbf{x}, \mathbf{r}) = 0.005dB \cdot B_p, \quad (26)$$

where  $L_p$ ,  $C_p$ ,  $D_p$  and  $B_p$  are the waveguide length in cm, the number of crossings, the number of drops, and the number of bends in path  $p$ , respectively. In PROTON+ the insertion loss of path  $p$  is approximated as the sum of propagation loss, crossing loss, and drop loss. Hence, the maximum insertion loss is calculated as

$$il_{max}(\mathbf{x}, \mathbf{r}) = \max_{p \in P} (pl_p(\mathbf{x}, \mathbf{r}) + cl_p(\mathbf{x}, \mathbf{r}) + dl_p(\mathbf{x}, \mathbf{r}) + bl_p(\mathbf{x}, \mathbf{r})). \quad (27)$$

If the technology changes in the future, only the loss parameters have to be adapted. The rest of PROTON+ is flexible and technology independent.

## 4. EXPERIMENTAL RESULTS

PROTON+ is implemented in C++. All experiments are performed on an Intel Core 2 Quad CPU with 8GB RAM running at 2.33GHz. The nonlinear optimization problem (20) is solved by IPOPT [Wächter and Biegler 2006], which is one of the leading libraries in nonlinear optimization. The maximum number of iterations is set to 100, which gives suitable results within an acceptable runtime for most benchmarks used in this section.

The rest of this section is structured as follows: In Section 4.1 we compare PROTON+ to manually created layouts regarding the required laser power of the resulting ONoC. In Section 4.2 we adapt Cadence Encounter for placement and routing of an optical NoC and discuss, if state-of-the-art placers can easily be adapted to the optical placement and routing problem. In Section 4.3 a study of different topologies is given, and the best topology is chosen. For the first time, we propose a study about the

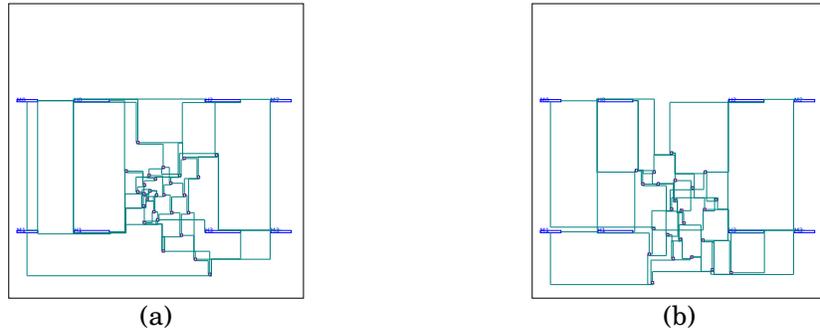


Fig. 10. Resulting placement for (a) the 8x8  $\lambda$ -Router and (b) the 8x8 GWOR using PROTON+ and minimizing crossing loss only

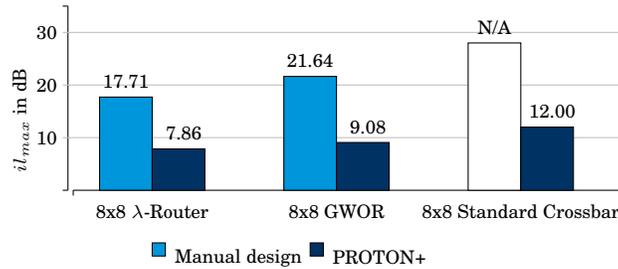


Fig. 11. Insertion loss of the manually and automatically designed layouts

positions of memory controllers in Section 4.4. In addition, we show that PROTON+ can handle loose as well as tight placed initiators and targets in Section 4.5. Finally, we study the scalability of PROTON+ in Section 4.6, and for the first time we discuss the layouts of a new type of topologies in Section 4.7, which are obtained by network partitioning.

#### 4.1. Manual Layout vs. Our Algorithm

In this section the manually designed layout of an 8x8  $\lambda$ -Router and an 8x8 GWOR are compared with the automatically created layouts of the same topologies obtained by PROTON+. We are interested in the required laser power of the system, which is the minimum laser power needed to guarantee a predefined bit error rate at the receivers. For the 8x8  $\lambda$ -Router 8 different wavelengths are needed, while the 8x8 GWOR works with 7 different wavelengths. We assume an equal laser power for all lasers and thus are interested in the path with maximum required laser power. A measurement for the laser power is the maximum insertion loss  $il_{max}$ . For the calculation of the insertion loss we assume the loss parameters given in Table I. Due to complexity reasons the manual layout published in [Ramini et al. 2012] and shown in Figure 4(b) minimizes only the maximum number of crossings over all paths. To ensure fairness we compare these results with layouts obtained by PROTON+ when only crossing loss is minimized. The die size in both cases is assumed to be  $9mm \times 9mm$ . The memory controllers are positioned pairwise at the boundary of the chip as illustrated in Figure 13(a). The results for the 8x8  $\lambda$ -Router and the 8x8 GWOR are shown in Figure 10. In contrast to the manual designs, our layouts are much denser, which results in lower propagation loss. In addition, for the 8x8  $\lambda$ -Router the number of crossings in the critical path decreased from 64 in the manual design to 27 in our layout.

Figure 11 shows the maximum insertion loss of the manually created layouts published in [Ramini et al. 2012] and of PROTON+. Due to complexity reasons there is no manually designed layout available for the 8x8 Standard Crossbar. Obviously, in

both cases the 8x8  $\lambda$ -Router is the topology with lowest insertion loss. Although the 8x8 GWOR has fewer nets and PSEs the predetermined positions of hubs and memory controllers is more disadvantageous for the GWOR than for the  $\lambda$ -Router. In particular, the maximum insertion loss of the 8x8  $\lambda$ -Router and of the 8x8 GWOR is 17.71 dB and 21.64 dB for the manually designed layout and 7.86 dB and 9.08 dB for the automatic generated layout. This is an improvement of more than 2x for both topologies. The benefit of this effect is strengthened when we calculate the required laser power of the system. In our experiments we assume that the worst insertion loss over all paths, e.g. the path with  $il_p = il_{max}$ , determines the required laser power. For the laser sources, we assume a laser-efficiency PLE of 20% and a coupling laser-link PCW of 90%. PLE corresponds to the laser efficiency for the conversion of electrons to photons. PCW is the coupling coefficient between the laser source and the optical waveguide. For the detectors a sensitivity of  $S = -17dBm$  and a BER of  $10^{-12}$  is assumed. Thus, given the maximum insertion loss  $il_{max}$  in dB, the optical laser power  $P_{opt}$  in  $mW$  is calculated as

$$P_{opt} = 10^{\left(\frac{il_{max}+S}{10}\right)}. \quad (28)$$

The electrical power can be obtained by dividing the optical power by the laser-efficiency PLE and the coupling laser-link PCW. Figure 12 shows the required laser power for the manually created layouts and for the layouts obtained with PROTON+. This is the laser power that is required by each hub in the assumption of providing the same worst-case laser power to all wavelength channels of that hub. We are reporting the laser power for each hub and not for the whole system, since we are not considering the problem of synthesizing a suitable power distribution network within this contribution. Because the optical laser power is monotonically increasing with the insertion loss, the required laser power of the 8x8  $\lambda$ -Router is lowest followed by the 8x8 GWOR and the 8x8 Standard Crossbar. In particular, we obtain 52.34 mW for the 8x8  $\lambda$ -Router and 113.19 mW for the 8x8 GWOR when the layouts are created manually. With PROTON+ we are able to decrease the required laser power to 5.42 mW for the 8x8  $\lambda$ -Router and 6.28 mW for the 8x8 GWOR. This is an improvement of 90% and 94% for the 8x8  $\lambda$ -Router and the 8x8 GWOR respectively. Although the 8x8 topologies seem to be small, and each path contains at most 7 crossings in the logic scheme, these results show that it is difficult to determine the optimal layout manually when the hubs are positioned in the center, and the memory controllers are placed close to the boundary of the chip. Based on an analysis of the layouts created by PROTON+, we assume that the optimal layout has an irregular structure, which is hard to be determined by the human brain. Placement and routing with PROTON+ takes a couple of minutes, while the manually created layouts took the designer more than one week per topology. Compared to the results published in [Boos et al. 2013], the number of crossings is reduced from 41 to 27 for the 8x8  $\lambda$ -Router and from 43 to 38 for the 8x8 GWOR due to the improved crossing approximation function and the changed net order during routing. The required laser power of the system can hardly be compared due to updated loss parameters (given in Table I). These results show that the quality of the layout significantly influences the maximum insertion loss and the system's laser power. Thus, placement and routing tools for ONoCs are indispensable.

#### 4.2. Cadence Encounter vs. Our Algorithm

In this section we apply the state-of-the-art placer and router Cadence Encounter to the placement and routing problem of ONoCs. Cadence Encounter does not enable any crossings, because crossings of metallic interconnections can lead to unintended behavior of the circuit. The easiest solution to deal with crossings is the usage of two metal layers for routing. On the first layer all wires are routed horizontally, and on the second layer the wires are routed vertically only. In that way, two wires will not run in parallel on top of each other. Vias are used for the interconnections between the two layers. Because each via means a direction change of the signal (between horizontal and vertical direction), a via in Cadence Encounter corresponds to a waveguide bend of

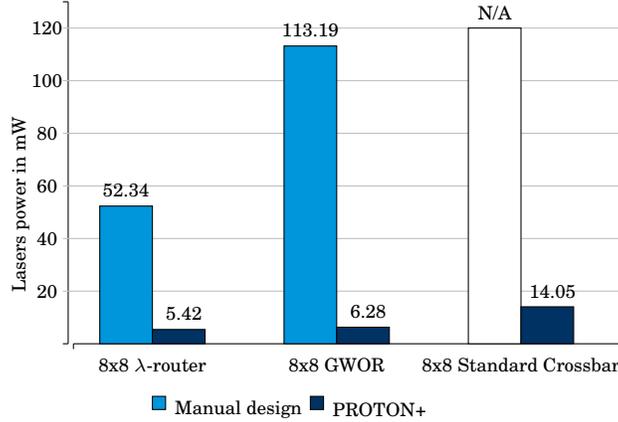


Fig. 12. Required laser power for each hub for the manually designed layouts and the automatic generated layouts

the ONoC. All hubs, memory controllers and PSEs are treated as macro modules and are allowed to be placed at each position inside the chip area. With the help of Cadence Encounter we place and route the 8x8  $\lambda$ -Router. Because Encounter does not support the option of crossing minimization, only the wire length is minimized in our experiment, which corresponds to a minimization of propagation loss. In addition, only the total wire length can be minimized by Encounter. In general, the total wire length is a poor approximation of the maximum wire length over all paths, especially if one path has an excessively high wire length, and the wire length of all other paths is small. After placement and routing, the number of waveguide crossings has to be counted manually, since this feature is not supported by Encounter. The path with the highest insertion loss has 74 crossings and a waveguide length of 26192  $\mu\text{m}$ , which results in a maximum insertion loss of 15.03 dB when bending loss is not considered. The average number of crossings per path is 39, while the average waveguide length is 18272  $\mu\text{m}$ . Compared to the 7.86 dB obtained with our algorithm, the maximum insertion loss increases by 91%. These results confirm the need for new algorithms for placement and routing of ONoCs, which enable but minimize the number of waveguide crossings. In addition, state-of-the-art placers and routers for electronic circuits usually minimize the total wire length and are not able to minimize the maximum wire length of a path. While the second concern can be disposed with comparatively low effort, the crossing approximation is a difficult problem that cannot be integrated into existing tools easily.

### 4.3. Best Topology Selection

As can be seen in equation (2), the weighted sum of propagation loss and crossing loss is minimized during placement and routing. PROTON+ can be calibrated with the help of the weights  $\alpha$  and  $\beta$ , which corresponds to a physical design space exploration. We place the 8x8  $\lambda$ -Router, the 8x8 GWOR and the 8x8 Standard Crossbar on a die with dimensions of 9mm  $\times$  9mm. The memory controllers are placed pairwise at the boundary of the chip as illustrated in Figure 13(a).

Table III shows the results obtained by the variation of  $\alpha$  and  $\beta$  from 0 to 1 with a step size of 0.1. The weights are given in the first two columns. Please notice, that  $\alpha + \beta = 1$  always has to be fulfilled. The maximum insertion loss is given in the column named  $il_{max}$ . The number of crossings and the waveguide length in  $\mu\text{m}$  of path  $p$  with maximum insertion loss is given in columns named  $C$  and  $L$  respectively. At last, the CPU time used by PROTON+ is presented in columns referred to as CPU. As can be seen in Table III, the insertion loss depends on the crossing loss as well as the propagation loss for the topologies under test. The maximum insertion loss is lowest

Table III. Results for variation of propagation and crossing loss weights

weights		8x8 $\lambda$ -Router				8x8 GWOR				8x8 Standard-Crossbar			
$\alpha$	$\beta$	$il_{max}$	C	L	CPU	$il_{max}$	C	L	CPU	$il_{max}$	C	L	CPU
1.0	0.0	16.6	93	12231	6.5	11.5	60	20160	7.0	12.6	67	11511	38.0
0.9	0.1	8.6	41	10413	99.5	8.4	38	13347	78.3	8.5	36	15255	674.0
0.8	0.2	7.8	32	14868	139.0	9.4	42	16938	92.6	14.2	63	26100	652.0
0.7	0.3	9.1	44	11781	117.0	7.9	35	13788	93.7	12.2	48	27972	621.7
0.6	0.4	8.3	32	17784	160.4	8.7	30	11448	93.1	8.1	34	15129	615.8
0.5	0.5	8.1	29	20070	139.5	8.7	30	19431	90.9	7.6	32	13500	613.4
0.4	0.6	8.3	37	12897	139.7	8.5	34	23121	91.2	9.5	45	12924	679.6
0.3	0.7	7.8	39	18144	149.0	8.4	28	16479	92.1	12.9	53	27315	611.3
0.2	0.8	8.4	37	14040	143.9	9.6	36	12492	90.9	10.8	46	19755	635.6
0.1	0.9	8.5	33	18810	141.0	9.4	34	16182	92.0	10.7	42	23886	621.5
0.0	1.0	7.9	27	20817	146.9	9.1	42	13680	95.0	12.0	63	11628	622.3

Table IV. Number of nets and PSEs of the topologies

topology	number of PSEs	number of two-pin-connections
8x8 $\lambda$ -Router	28	64
8x8 GWOR	24	56
8x8 Standard Crossbar	64	111

when both, crossing and waveguide length optimization, is considered. In particular, the lowest maximum insertion loss values are obtained for  $\alpha = 0.8$ ,  $\alpha = 0.7$  and  $\alpha = 0.5$  for the 8x8  $\lambda$ -Router, the 8x8 GWOR, and the 8x8 Standard Crossbar, respectively. A minimization of propagation loss only leads to a high number of crossings and thereby to highest maximum insertion losses. At the same time the minimization of the number of crossings only indirectly minimizes the waveguide length too. If the two modules of the net, which is used to define the major axis of the ellipse, are positioned far away from each other, the area of the ellipse is large. Thus, many points, e.g. the points inside the ellipse, are mapped to high function values, which increases the crossing probabilities with other nets. Hence, minimizing crossing loss only (and thereby minimizing propagation loss indirectly) obtains better results than when minimizing propagation loss only. Best results are obtained when considering both losses. Please notice that just a small shift of one PSE can already result in one or more additional crossing(s). If a PSE is shifted by just one bin in the routing grid, the algorithm could prefer a crossing instead of a detour for one net. Hence, the routing and the resulting number of crossings are very sensitive to the placement, and some variations in the number of crossings occur. In addition, the number of crossings  $C$  and the waveguide length  $L$  is given for the path with maximum insertion loss. This does not have to be the path with the maximum number of crossings or the maximum waveguide length. On average, the maximum insertion loss  $il_{max}$  of the 8x8  $\lambda$ -Router is smaller compared to the 8x8 GWOR and the 8x8 Standard Crossbar. Thus, in terms of maximum insertion loss the 8x8  $\lambda$ -Router is the topology of choice followed by the 8x8 GWOR and the 8x8 Standard Crossbar. Our algorithm acts fast on these three topologies. The 8x8 GWOR, which has the lowest number of two-pin-connections, is placed and routed in less than 2 minutes. For placing and routing the 8x8  $\lambda$ -Router PROTON+ needs less than 3 minutes. The 8x8 Standard Crossbar is placed and routed within 12 minutes. The runtime mainly depends on the number of PSEs to be placed and the number of nets to be routed, which are given in Table IV. Compared to the previous work published in [Boos et al. 2013], the runtime increases slightly because of the symmetric definition of the crossing function given in equation (16).

#### 4.4. Positions of memory controllers

Similar to the Tiler architecture proposed in [Wentzlaff et al. 2007], we assumed the hubs to be located at the center of the optical layer and the memory controllers to be positioned pairwise at the boundary of the optical layer in previous experiments. The positions of hubs are determined by the lower (electronic) layer, but the positions of the memory controllers are a degree of freedom. The designer has to decide, where to place

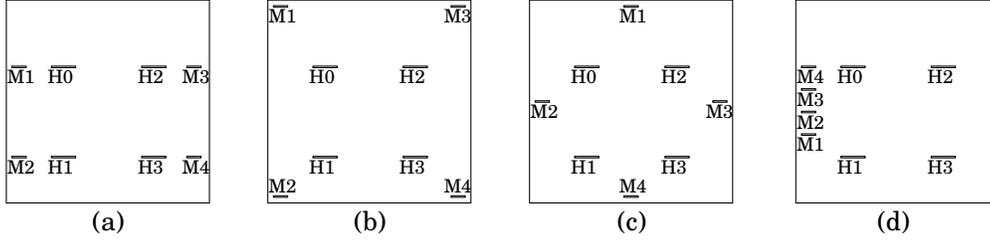


Fig. 13. Positions of hubs and memory controllers (a) pairwise, (b) at the corners, (c) with M1 north and (d) for benchmark *onside*.

Table V. Results for different positions of memory controllers

Weights		8x8 $\lambda$ -Router				8x8 GWOR				8x8 Standard-Crossbar			
objective function	positions	$il_{max}$	C	L	CPU	$il_{max}$	C	L	CPU	$il_{max}$	C	L	CPU
<i>pl</i>	pairwise	16.6	93	12231	6.5	11.5	60	11826	7.0	12.6	67	11511	38.0
<i>pl</i>	M1North	15.3	84	12690	6.4	8.8	43	10476	5.4	10.7	54	11583	35.0
<i>pl</i>	M3North	14.1	77	11178	6.1	9.8	49	11457	5.7	17.4	98	12681	35.3
<i>pl</i>	corners	16.0	84	12690	6.4	8.5	37	14661	7.7	18.9	105	15390	38.6
<i>pl</i>	onside	8.8	43	11034	5.9	10.9	57	10629	4.8	12.8	69	10620	33.4
$\alpha pl + \beta cl$	pairwise	8.6	41	10413	95.8	8.4	38	13014	77.1	8.5	36	15255	606.9
$\alpha pl + \beta cl$	M1North	8.2	35	14616	130.5	8.1	38	10971	88.5	10.0	50	10998	618.5
$\alpha pl + \beta cl$	M3North	9.0	41	14751	138.5	8.1	38	11016	88.2	10.9	43	23958	613.6
$\alpha pl + \beta cl$	corners	8.6	37	15048	90.0	9.1	41	14742	81.5	10.0	45	15579	615.5
$\alpha pl + \beta cl$	onside	6.6	27	12726	134.0	8.1	35	13806	79.0	13.0	40	41031	649.0
<i>cl</i>	pairwise	7.9	27	20817	146.9	9.1	42	13680	95.0	12.0	63	11628	622.3
<i>cl</i>	M1North	7.5	31	17748	135.4	8.0	36	12528	90.7	9.6	45	13032	624.7
<i>cl</i>	M3North	9.6	37	21843	142.0	10.5	33	31635	90.0	9.6	47	11799	629.9
<i>cl</i>	corners	7.8	29	18513	136.5	10.5	38	26937	93.4	10.6	51	13725	605.9
<i>cl</i>	onside	9.0	27	28386	134.9	11.3	51	19062	91.3	10.5	33	31743	601.6

them to further decrease the laser power consumption of the ONoC. In this section we study different positions of the memory controllers. Their positions should be restricted to the boundary of the layer, because all memory controllers are connected to an off-chip memory, and excessively long optical connections to the boundary will increase the risk of an on-chip waveguide crossing. We consider four different cases:

- The situation illustrated in Figure 13(a), where the four memory controllers are located pairwise at the boundary of the optical layer, is referred to as *pairwise*.
- The situation shown in Figure 13(b), where the four memory controllers are located at the four corners of the optical layer, is referred to as *corners*.
- The situation illustrated in Figure 13(c), where the four memory controllers are distributed to all four sides, and the memory controller M1 is placed in the northern position, is named *M1North*.
- The situation similar to *M1North*, where the four memory controllers are rotated counter clockwise compared to the situation *M1North*, is named *M3North*. The memory controller M3 is located in the north.
- The situation shown in Figure 13(d), where the four memory controllers are located at the left most side of the optical layer, is named *onside*.

The situations M1North and M3North are equal except for module swapping.

Table V shows the results of PROTON+ for the five different locations of the memory controllers and three different objective functions, where *pl* and *cl* are referred to as the propagation loss and crossing loss minimization, respectively. For the minimization of propagation and crossing loss  $\alpha pl + \beta cl$ , the weights were chosen corresponding to the parameters given in Table I. As indicated by the maximum insertion loss values and shown in Figure 14, the layouts for the situations M1North and M3North are not equal. The layout not only depends on the positions of the memory controllers but also on the positions of the hubs. Because the positions of the hubs do not change, the layout cannot be rotated or mirrored, and thus the new positions of memory controllers result in an entirely new layout. **For the 8x8  $\lambda$ -Router the best results are obtained for**

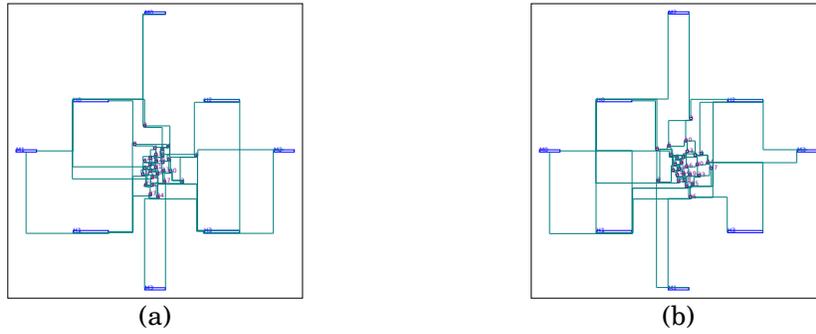


Fig. 14. Resulting placement for 8x8- $\lambda$ -Router when minimizing crossing loss of (a) M1North and (b) M3North.

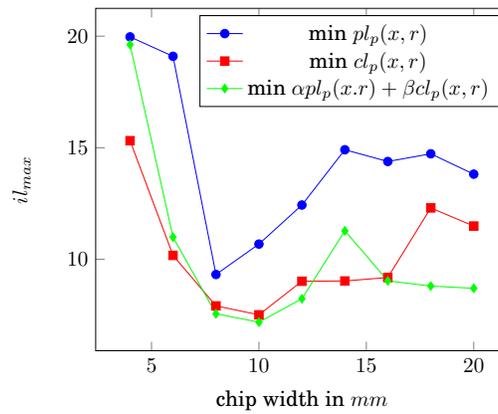


Fig. 15. Maximum insertion loss for different chip densities

the benchmark *onside* when minimizing both, crossing loss and propagation loss. For the 8x8 GWOR the benchmark M1North is the best choice, since its maximum insertion loss is lowest for these positions when minimizing crossing loss and propagation loss at once or minimizing crossing loss only. For the 8x8 Standard Crossbar the best result is obtained by the minimization of crossing and propagation loss as well as the pairwise position of memory controllers. For the 8x8  $\lambda$ -Router and the 8x8 GWOR the designers should avoid the situation M3North and corner respectively, while for the 8x8 Standard Crossbar all five possibilities should be considered.

#### 4.5. Loose vs. Tight

In this section we study the resulting layouts of PROTON+, when the die size increases. We consider the 8x8  $\lambda$ -Router since it has lowest maximum insertion loss as shown in Section 4.3. For all experiments we assume a square die area, while the chip width (and height) grows from 4mm to 20mm. The relative positions of hubs and memory controllers are kept, while they are shifted away from each other when the die size is increased. E.g. the positions of hubs and memory controllers are scaled dependent on the die width. The parameters  $\alpha$  and  $\beta$  are chosen based on the loss parameters given in Table I, and the memory controllers are placed pairwise at the boundary of the chip as illustrated in Figure 13(a). In Figure 15 the results are shown for three different objective functions: minimize propagation loss only, minimize crossing loss only and minimize both, propagation and crossing loss, simultaneously. Because minimizing propagation loss only results in high numbers of crossings, the maximum insertion

loss is highest using the propagation loss as the objective function. When minimizing the number of crossings the waveguide length indirectly is minimized too as explained in Section 4.3. For small die sizes the maximum insertion loss is highest. Due to the limited available space for PSEs and waveguides, all PSEs are placed dense, and the waveguides are located close to each other, which increases the number of crossings. When increasing the die size the maximum insertion loss decreases. After a certain die size is reached, the maximum insertion loss starts increasing again. For the 8x8  $\lambda$ -Router a die size of  $10mm \times 10mm$  is the best choice. When this certain die size is reached, PROTON+ already found a good solution for placement and routing, and the number of crossings is lowest. When minimizing both, this solution is just scaled with increasing die size, while the number of crossings is more or less constant. Due to the scaled positions of hubs and memory controllers, the waveguide length and thereby the propagation loss further increases, which results in an increase of the maximum insertion loss. The runtime of PROTON+ is independent of the die size. In fact, the CPU time only depends on the number of PSEs to be placed and the number of nets to be routed. Just for special cases it might happen that another die size leads to fewer iterations in the IPOPT library and thus lower runtime. The maximum number of iterations is restricted to 100 for all experiments. The mainly computation time is needed for the crossing approximation. Minimizing propagation loss needs only 5% of the time when propagation loss and crossing loss are minimized. Compared to the results published in [Boos et al. 2013], the runtime for the crossing loss minimization increased up to 61% due to the new crossing approximation function given in equation (16), which ensures symmetry between the crossings of two nets. Still, the runtime is acceptable for the automatic design of ONoCs.

#### 4.6. Scalability

In this section we show the feasibility of PROTON+ for large topologies. As a case study, the number of cores in the electronic layer is increased to 96, while preserving 8 cores per cluster. Consequently, 12 hubs are needed on the optical layer, while the number of memory controllers is kept. In total, we have 16 initiators and 16 targets to be connected. The topology shown in Figure 4(a) is scaled up to a 16x16  $\lambda$ -Router. The die size is increased to  $12mm \times 16mm$ . The results can be seen in Table VI, where the second row shows the maximum insertion loss in dB. The number of crossings and the waveguide length in  $\mu m$  of the path with highest insertion loss are shown in columns 3 and 4 respectively. Finally, the runtime in minutes is shown in the last column. These results were obtained by minimizing the crossing loss only. Compared to the 8x8  $\lambda$ -Router, the number of PSEs in the 16x16  $\lambda$ -Router increases by a factor of 4.3x, which results in an increase of the maximum insertion loss by 4.3x due to an increase of the number of crossings by a factor of 7.9x. Furthermore, the runtime increases by a factor of 173.7x. A layout of a 16x16  $\lambda$ -Router obtained with PROTON+ is shown in Figure 16. The large and equally distributed (blue) rectangles are the hubs and memory controllers. In contrast to these fixed modules, the PSEs are the small rectangles connected by lines, which illustrate the waveguides. **Already for 16x16 networks, a comparison with manual layouts is unaffordable due to error-prone and time-consuming manual design. A key contribution of PROTON+ is to materialize such a design point. For larger topologies, wavelength-routing may become impractical because of the proliferation of wavelength channels. As demonstrated in [Ramini et al. 2014], wavelength-arbitration may become mandatory to preserve power efficiency beyond 16x16 optical NoCs. Last but not least, a practically-relevant way of working around the large insertion loss penalty of high-radix topologies consists of network partitioning, which is effectively supported by the tool and hereafter explained in Section 3.4.**

#### 4.7. Network Partitioning

In this section we study the maximum insertion loss of the layouts obtained by network partitioning [Ramini et al. 2012]. Memory controllers are positioned at the boundary of the chip as shown in Figure 13(a). We place and route the three subnetworks on a

Table VI. Comparison of the 16x16  $\lambda$ -Router with an 8x8  $\lambda$ -Router

topology	$il_{max}$	C	L	CPU
8x8 $\lambda$ -Router	7.9	27	20817	2.4
16x16 $\lambda$ -Router	38.9	214	36832	416.9

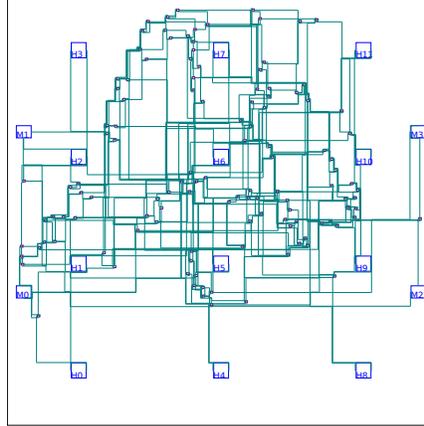

 Fig. 16. Resulting placement for 16x16  $\lambda$ -Router obtained by PROTON+

Table VII. Results for the network partitioning without defining regions

objective function	$il_{max}$	C	L	CPU
$pl$	7.1	33	10422	8.3
$\alpha pl + \beta cl$	5.8	23	11151	26.9
$cl$	6.8	22	19062	31.3

Table VIII. Results for the network partitioning when defining regions

objective function	$il_{max}$	C	L	CPU
$pl$	6.6	24	15327	7.1
$\alpha pl + \beta cl$	5.9	19	15687	30.5
$cl$	6.2	15	21915	30.2

die with size  $9mm \times 9mm$  without the definition of regions. Table VII gives the results for minimizing propagation loss ( $pl$ ), minimizing crossing loss ( $cl$ ), and minimizing a weighted sum of both ( $\alpha pl + \beta cl$ ). Again, the maximum insertion loss in dB is given in the second column named  $il_{max}$ , the number of crossings C and the waveguide length L in  $\mu m$  of the path with highest insertion loss are shown in columns three and four respectively. In the last column the CPU time in seconds is shown. PROTON+ is able to place and route all three circuits in less than 32 seconds, which is a runtime decrease of more than 78% compared to the 8x8  $\lambda$ -Router. In addition, the maximum insertion loss decreases from 7.8 dB (which was the best result obtained for the 8x8  $\lambda$ -Router in previous experiments with a pairwise location of the memory controllers) to 5.8 dB when minimizing the weighted sum of propagation loss and crossing loss. Due to a smaller number of PSEs to be placed and waveguides to be routed, the number of waveguide crossings and the waveguide length in the path with maximum insertion loss decreased. Because the runtime of PROTON+ is nonlinear dependent on the number of PSEs, the tool places the three subnetworks containing 16 PSEs in total much faster than the 8x8  $\lambda$ -Router containing 28 PSEs. The partitioning suggested in [Ramini et al. 2012] can be used for a various number of hubs and memory controllers. The runtime is lowest when all subnetworks contain an equal number of PSEs to be placed. Even if very small subnetworks can be placed by hand, an automatic tool for routing and placement is inevitably needed when the number of hubs and memory controllers scales up.

In Table VIII the results for the three subnetworks with the definition of regions are given. The regions are defined according to the suggestions given in [Ramini et al. 2012]. The maximum insertion loss slightly worsens with the definition of regions

when minimizing both, propagation loss and crossing loss. In particular, the number of crossings decreases, while the waveguide length increases. Thus, depending on the insertion loss parameters, which are given in Table I and might change for future technologies, it can be useful to define regions, especially when the crossing loss dominates the insertion loss. In addition, sometimes the designer wants to place some modules in specific areas on the chip. The runtime in our experiments is still less than 31 seconds. The region definition strongly depends on the positions of hubs and memory controllers. The best results are obtained when the regions do not intersect each other and contain all targets and initiators of the subnetworks to be placed in the regions. Then, waveguide crossings between subnetworks are avoided, and thereby the total number of crossings is decreased, which results in a lower maximum insertion loss. The complexity of determining an optimal region depends on the topologies to be placed and routed as well as on the positions of hubs and memory controllers. With the help of network partitioning and of PROTON+ including the region definition, even large optical topologies can be placed and routed fast, and the final systems consume less laser power.

## 5. CONCLUSION

We introduced PROTON+, a fast placement and routing algorithm for optical Networks-on-Chip. The algorithm iteratively places all PSEs overlap-free inside the chip-area and routes the waveguides to connect all optical devices. During placement and routing, the maximum insertion loss and thereby the total laser power required by the ONoC is minimized. Compared to manually designed state-of-the-art layouts, PROTON+ is able to reduce the laser power by up to 94%. All 8x8 topologies can be placed within 12 minutes. Furthermore, for the first time we gave a case study for the positions of memory controllers and placed and routed a network-partitioned topology automatically. In future work we plan to add a post-layout optimization step. Since most crossings occur around PSEs, a rotation of the PSEs can help to further reduce the number of crossings, which can further reduce the maximum insertion loss and thereby the required laser power of the ONoC.

## REFERENCES

- S. Bahirat and S. Pasricha. 2014. 3D HELIX: Design and Synthesis of Hybrid Nanophotonic Application-Specific 3D Network-On-Chip Architectures. *Workshop on Exploiting Silicon Photonics for Energy efficient Heterogeneous Parallel Architectures (SiPhotonics)* (2014).
- S. Beamer, C. Sun, Y.-J. Kwon, A. Joshi, C. Batten, V. Stojanović, and K. Asanović. 2010. Re-architecting DRAM Memory Systems with Monolithically Integrated Silicon Photonics. *SIGARCH Comput. Archit. News* 38, 3 (June 2010), 129–140.
- A. Biberman, K. Preston, G. Hendry, N. Sherwood-Droz, J. Chan, J. S. Levy, M. Lipson, and K. Bergman. 2011a. Photonic Network-on-chip Architectures Using Multilayer Deposited Silicon Materials for High-performance Chip Multiprocessors. *Journal on Emerging Technologies in Computing Systems (JETC)* 7, 2, Article 7 (Jul 2011), 7:1–7:25 pages.
- A. Biberman, N. Sherwood-Droz, Xiaoliang Zhu, M. Lipson, and K. Bergman. 2011b. High-speed data transmission in multi-layer deposited silicon photonics for advanced photonic networks-on-chip. In *Conference on Lasers and Electro-Optics (CLEO)*. 1–2.
- A. Boos, L. Ramini, U. Schlichtmann, and D. Bertozzi. 2013. PROTON: An automatic place-and-route tool for optical Networks-on-Chip. In *International Conference on Computer-Aided Design (ICCAD)*. 138–145.
- J. Chan, G. Hendry, A. Biberman, and K. Bergman. 2010. Architectural Exploration of Chip-Scale Photonic Interconnection Network Designs Using Physical-Layer Analysis. *Lightwave Technology, Journal of* 28, 9 (May 2010), 1305–1315.
- M. Cianchetti, J. Kerekes, and D. Albonesi. 2009. Phastlane: A Rapid Transit Optical Routing Network. *SIGARCH Comput. Archit. News* 37, 3 (June 2009), 441–450.
- C. Condrat, P. Kalla, and S. Blair. 2011. Logic Synthesis for Integrated Optics. In *Great Lakes Symposium on VLSI*. ACM, New York, NY, USA, 13–18.
- C. Condrat, P. Kalla, and S. Blair. 2014. Crossing-Aware Channel Routing for Integrated Optics. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 33, 6 (Jun 2014), 814–825.
- D. Ding, Y. Zhang, H. Huang, R. T. Chen, and D. Z. Pan. 2009. O-Router: An optical routing framework for low power on-chip silicon nano-photonics integration. In *Design Automation Conference (DAC)*. 264–269.

- Po Dong, Wei Qian, Shirong Liao, Hong Liang, Cheng-Chih Kung, Ning-Ning Feng, R. Shafiha, J. Fong, Dazeng Feng, A.V. Krishnamoorthy, and M. Asghari. 2010. Low loss silicon waveguides for application of optical interconnects. In *Photonics Society Summer Topical Meeting Series, 2010 IEEE*. 191–192.
- V. Donzella, S. Fard, and L. Chrostowski. 2013. Study of waveguide crosstalk in silicon photonics integrated circuits. *Proc. SPIE 8915, Photonics North* (Oct 2013), 89150Z.
- H. Gu, J. Xu, and W. Zhang. 2009. A low-power fat tree-based optical Network-On-Chip for multiprocessor system-on-chip. In *Design, Automation Test in Europe Conference Exhibition (DATE)*. 3–8.
- O. Hammami and K. Hamwi. 2013. MHYNESYS II: Multi-stage hybrid Network on chip synthesis for Next Generation 3D IC Manycore. In *International Symposium on Circuits and Systems (ISCAS)*. 325–328.
- M. Heck and J. Bowers. 2014. Energy Efficient and Energy Proportional Optical Interconnects for Multi-Core Processors: Driving the Need for On-Chip Sources. *Journal of Selected Topics in Quantum Electronics* 20, 4 (July 2014), 332–343.
- R. Ho, K.W. Mai, and M.A. Horowitz. 2001. The future of wires. *Proc. IEEE* 89, 4 (Apr 2001), 490–504.
- S. Koohi, M. Abdollahi, and S. Hessabi. 2011. All-optical wavelength-routed NoC based on a novel hierarchical topology. In *International Symposium on Networks on Chip (NoCS)*. 97–104.
- C. Lee. 1961. An Algorithm for Path Connections and Its Applications. *Transactions on Electronic Computers* EC-10, 3 (Sept 1961), 346–365.
- Y. Liu, J. Shainline, X. Zeng, and M. Popović. 2014. Ultra-low-loss CMOS-compatible waveguide crossing arrays based on multimode Bloch waves and imaginary coupling. *Opt. Lett.* 39, 2 (Jan 2014), 335–338.
- J. Minz, S. Thyagara, and S. Lim. 2007. Optical Routing for 3-D System-On-Package. *IEEE Transactions on Components and Packaging Technologies* 30, 4 (Dec 2007), 805–812.
- N. Ophir and K. Bergman. 2013. Analysis of high-bandwidth low-power microring links for off-chip interconnects. *Proc. SPIE 8628* (Mar 2013), 86280N–86280N–7.
- M. Ortín-Obón, L. Ramini, V. Viñals, and D. Bertozzi. 2014. Capturing the sensitivity of optical network quality metrics to its network interface parameters. *Concurrency and Computation: Practice and Experience* 26, 15 (2014), 2504–2517.
- A. Parini, G. Calò, G. Bellanca, and V. Petruzzelli. 2014. Vertical link solutions for multilayer optical-networks-on-chip topologies. *Optical and Quantum Electronics* 46, 3 (2014), 385–396.
- W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. 2007. *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (3 ed.). Cambridge University Press, New York, NY, USA.
- L. Ramini, D. Bertozzi, and L.P. Carloni. 2012. Engineering a Bandwidth-Scalable Optical Layer for a 3D Multi-core Processor with Awareness of Layout Constraints. In *International Symposium on Networks on Chip (NoCS)*. 185–192.
- Luca Ramini, Mahdi Tala, and Davide Bertozzi. 2014. Exploring Communication Protocols for Optical Networks-on-Chip based on Ring Topologies, In Asia Communications and Photonics Conference 2014. *Asia Communications and Photonics Conference 2014* (2014), ATH3A.165.
- A. Scandurra. 2008. Scalable CMOS-compatible photonic routing topologies for versatile networks on chip. In *Network on Chip Architecture*.
- C.-S. Seo and A. Chatterjee. 2002. A CAD tool for system-on-chip placement and routing with free-space optical interconnect. In *International Conference on Computer Design: VLSI in Computers and Processors*. 24–29.
- C.-S. Seo, A. Chatterjee, and N. M. Jokerst. 2005. Physical design of optoelectronic system-on-a-package: a CAD tool and algorithms. In *International Symposium on Quality of Electronic Design (ISQED)*. 567–572.
- A. Shacham, K. Bergman, and L.P. Carloni. 2008. Photonic Networks-on-Chip for Future Generations of Chip Multiprocessors. *IEEE Trans. Comput.* 57, 9 (Sept 2008), 1246–1260.
- N. Sherwood-Droz, H. Wang, L. Chen, B. Lee, A. Biberman, K. Bergman, and M. Lipson. 2008. Optical 4x4 Hitless Silicon Router for Optical Networks-on-Chip (NoC). *Opt. Express* 16, 20 (2008), 15915–15922.
- X. Tan, M. Yang, L. Zhang, Y. Jiang, and J. Yang. 2011. On a Scalable, Non-Blocking Optical Router for Photonic Networks-on-Chip Designs. In *Symposium on Photonics and Optoelectronics*. 1–4.
- A. Udipi, N. Muralimanohar, R. Balasubramonian, A. Davis, and N. Jouppi. 2011. Combining Memory and a Controller with Photonics Through 3D-stacking to Enable Scalable and Energy-efficient Systems. In *Proceedings of the 38th Annual International Symposium on Computer Architecture (ISCA '11)*. ACM, New York, NY, USA, 425–436.
- D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N.P. Jouppi, M. Fiorentino, A Davis, N. Binkert, R.G. Beausoleil, and J.H. Ahn. 2008. Corona: System Implications of Emerging Nanophotonic Technology. In *International Symposium on Computer Architecture*. 153–164.
- A. Wächter and L. Biegler. 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106, 1 (2006), 25–57.
- D. Wentzlaff, P. Griffin, H. Hoffmann, L. Bao, B. Edwards, C. Ramey, M. Mattina, C.-C. Miao, J. Brown III, and A. Agarwal. 2007. On-Chip Interconnection Architecture of the Tile Processor. *IEEE Micro* 27, 5 (2007), 15–31.