2012 Sixth IEEE/ACM International Symposium on Networks-on-Chip

# Engineering a Bandwidth-Scalable Optical Layer for a 3D Multi-core Processor with Awareness of Layout Constraints

Luca Ramini
Engineering Department
University of Ferrara
Ferrara, 44122 Italy
Email: luca.ramini@unife.it

Davide Bertozzi
Engineering Department
University of Ferrara
Ferrara, 44122 Italy
Email: davide.bertozzi@unife.it

Luca P. Carloni
Computer Science Department
Columbia University
New York, 10027 USA
Email: luca@cs.columbia.edu

***Abstract*—The performance of future chip multi-processors will only scale with the number of integrated cores if there is a corresponding increase in memory access efficiency. The focus of this paper on a 3D-stacked wavelength-routed optical layer for high bandwidth and low latency processor-memory communication goes in this direction and complements ongoing efforts on photonically integrated bandwidth-rich DRAM devices. This target environment dictates layout constraints that make the difference in discriminating between alternative design choices of the optical layer. This paper assesses network partitioning options and bandwidth scalability techniques with deep technology and layout awareness, the main contribution lying in the characterization and precise quantification of such interaction effects between the technology platform, the layout constraints and the network-level quality metrics of a passive optical NoC.**

## I. Introduction

The performance of future multi-core processors will only scale with the number of integrated cores if there is a corresponding increase in memory bandwidth. For this purpose, silicon photonic technology is being investigated as a way to improve pin bandwidth density and power of DRAM memory devices [12].

In parallel, it is necessary to innovate the memory access architecture on the processor side. Processor-memory communication in chip multiprocessors (CMPs) and high-performance multi-core embedded systems is typically accommodated via an on-chip electronic network, which provides larger communication parallelism and higher aggregate bandwidth when compared to shared bus and multi-layer interconnect solutions. However, there is a clear gap between performance of such electronic NoCs and that of the high-bandwidth density, data-rate transparent and distance- independent off-chip optical links. The only way to bridge this gap is to bring the photonic interconnect technology deeper into the chip.

In recent years, a significant amount of work has been done on exploring various optical NoC topologies (e.g., [19], [20], [30], [29], [23], [24]. Many of these rely on active optical devices [7], where a dual electronic NoC is required to establish and manage optical paths across broadband optical switches. Setting up and tearing down optical circuits, however, are time consuming tasks that call for expensive architectural support such as the management of dropped requests or the capability of adaptive routing. Above all, applications may have different types of memory access requirements, ranging from the high-bandwidth processor-memory communication for streaming media decoding to the latency-constrained communications of those control applications for which response time is the key metric [15]. Active optical NoCs are not the best match for this latter kind of applications. Another set of works envisions fully optical interconnect solutions such as [14], however their efficiency depends on the availability of optical devices with stringent power consumption and signal loss features that are hardly achievable by current silicon photonic technology.

The above shortcomings motivate the main idea of this paper of using passive photonic NoCs (PPNoCs) for processor-memory communication. In PPNoCs, the route followed by a packet depends solely on the wavelength of its carrier signal (wavelength routing), and not on the information either contained or traveling along with it. In this way, the expensive O-E/E-O conversions required by some all-optical approaches like [11] are not needed. Also, if the routing pattern is set at design time and the wavelength employed for a source-destination pair is invariant for that pair, it does not depend on ongoing transmissions by other nodes and no time is lost in routing/arbitration. This is an appealing property for a processor-memory network in mixed criticality systems. In order to make silicon photonic technology affordable also for more cost-constrained multi-core systems (e.g., in the high-end embedded computing domain), we envision its implementation through a separate, vertically stacked optical layer with no integrated electronic devices. The 3D-stacking approach is the reference solution for cost-effective integration of heterogeneous technologies [26].

Although PPNoCs have been studied before in the literature, their use for processor-memory communication is challenged by a number of constraints that are typically overlooked: addressing these constraints is the main contribution of this paper. First, a processor-memory network poses specific layout constraints to the placement of the electro-optical network

978-0-7695-4677-3/12 $26.00 © 2012 IEEE
DOI 10.1109/NOCS.2012.29

interfaces and hence of the gateways to the optical layer. Moreover, memory controllers are typically distributed around the processor chip periphery to remove centralized communication bottlenecks from the on-chip network. Such layout constraints radically question the practical feasibility of appealing logic connectivity schemes proposed for passive NoCs and make the design of their associated physical topology mandatory. This latter incurs a hardly-predictable number of additional waveguide crossings depending on the specific layout constraints of the design at hand. The resulting unexpected insertion-loss may offset the theoretical properties of the topology and change relative comparison results among them. We borrow logic schemes for passive NoC topologies from [20] and derive the corresponding physical topologies when accounting for the layout constraints of a processor-memory network. We then quantify insertion-loss deviations and gain practical insights.

Second, network traffic is inherently heterogeneous and comprises several traffic classes. For a processor-memory network, it is possible to discriminate memory requests from responses and both of them from core-to-core communications. Awareness of such traffic classes inspired us alternative PPNoCs partitioning options, and led us to assess their layout implications. We contrast a global connectivity solution with multiple network partitions associated with different traffic classes. Again, we bring layout awareness to this comparative evaluation.

Third, the engineered PPNoC solution should be conceived with scalability in mind from the ground up, not as an afterthought. While scalability with the number of network nodes has been addressed in the literature (e.g., see the hierarchical approach in [23]), bandwidth scalability under a fixed number of network gateways and memory controllers is a largely unexplored topic. Yet, this is a key concern for those systems scaling to a larger number of cores. Bandwidth provided by photonically integrated DRAMs will most likely enable to preserve the number of memory controllers across a few device generations. Similarly, the need to amortize electro-optical conversion will cause the aggregation factor of processor cores around an optical gateway to increase, without immediately reverting to more gateways. In this context, the optical network must serve a larger number of memory access requests from its gateways.

Two known techniques can provide such bandwidth scalability: spatial parallelism (SPM) and broadband passive switching (BPS). Many works choose one technique over the other based mostly on qualitative reasons, and some of them end up selecting inefficient solutions. Instead we present the first quantitative comparative analysis between SPM and BPS for passive optical NoCs.

Combining the above contributions yields the complete engineering of a passive optical layer stacked on top of a multicore processor for high-bandwidth and low-latency processor-memory communication. This additional paper contribution stems directly from the previous one: layout awareness drives the key choices for planning the entire optical layer of a 3D multicore processor, thus providing well-grounded design guidelines. In this direction, we engineer wavelength reuse strategies and low cost implementation techniques.

In order to preserve technology-awareness in the analysis in spite of the focus at the network level we rely on a SystemC modeling and simulation environment where routing functionality is merged with FDTD-derived technology annotations in the models of the optical devices.

## II. Related Work

Shacham et al. [13] propose a circuit-switched on-chip photonic network with reconfigurable broadband optical switches. Circuit management via a dual electronic NoC may cause unpredictable communication latencies. Cianchetti et al. [11] propose another switch-based on-chip photonic network. It uses source-based routing and reconfigurable optical switches to route data. Switch setup is performed by converting the optical control signals that travel along the data to electrical form, and setting up the switch accordingly. Unlike the works above, Vantrease et al. [14] propose a fully optical solution. The large number of components, especially for high node counts, makes the viability of this architecture highly dependent on its ability to rein in the power consumption and signal losses of optical components, which will be heavily dependent on the maturity and efficiency of the optical technology employed. An approach with milder technology assumptions comes from [10] where a fully passive optical NoC is suggested. While the authors discuss bandwidth scalability options, they chose to use spatial parallelism without fully motivating the reason behind this decision. The key contribution of this paper is to provide a quantitative comparison between bandwidth scalability techniques for passive optical NoCs.

In literature there are many works on passive optical topologies. In particular, Connor et al. in [19] is one of the first examples. A basic element such as an add-drop filter was used to build the 4x4 network defined as lambda-router on which wavelength routing was applied. Scandurra and Connor presented a scalable and fully connected Optical NoC topology for multiple cores and heterogeneous systems-on-chip [20]. A full 8x8 ONoC topology is proposed and analyzed in two communication scenarios such as total and grouped connectivity. Grouped connectivity makes it clear that if total connectivity is not required, significant reductions in complexity can be achieved. Differently from these works, we provide total connectivity while capturing the layout implications on the physical topology.

Le Beux et al. in [21] further refine the topology. In particular, a single 8x8 interconnection network has been transformed into two optical sub-networks to reduce the number of crossings on the critical path. However, the number of optical sources stays the same. In contrast, we exploit network partitioning as a way to reuse optical sources. Also, the layout constraints assumed in [21] do not match those posed by processor-memory networks.

Layout design rules for a 3D environment have been analyzed in [22], even considering a variable number of optical network interfaces, waveguides and electronic layers

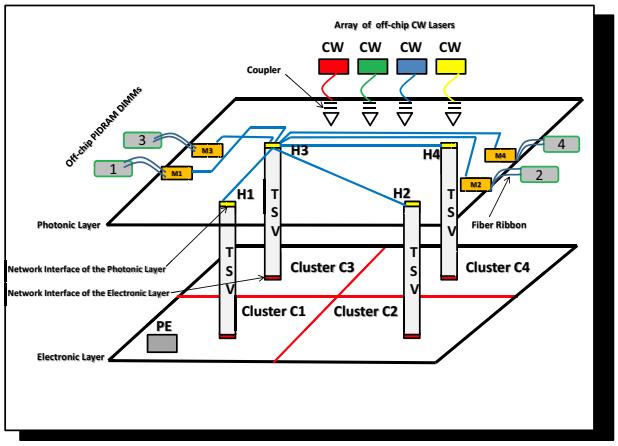



Fig. 1. The proposed 3D-Architecture

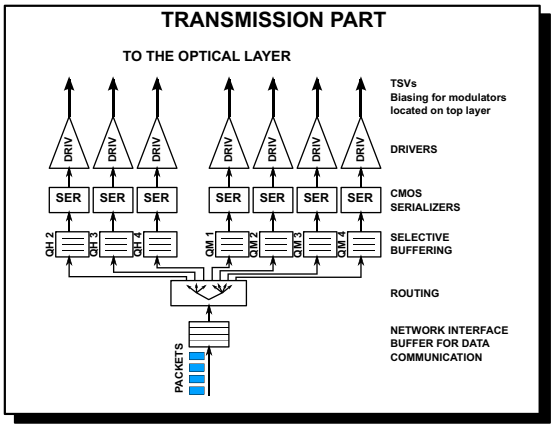



Fig. 2. Electronic Network Interface: Transmission Part

respectively. This analysis is limited to a ring topology and has the typical floorplanning of a regular homogeneous multi-core system. Therefore, they ignore the intricacy of connecting central hubs to the chip periphery, which severely challenges the physical topology for non-ring structures. The ring topology was improved in [24] by upgrading it to the Spidergon topology for all-optical wavelength routing. Despite constant node degree of four, network diameter of the Spidergon topology limits the efficiency of large scale optical NoCs. The drawbacks of Spidergon have been overcome in [23], where a two dimensional hierarchical expansion of ring topology has been proposed. This is an effective way of providing scalability to large scale optical networks. This paper is complementary to our work, which is focused on bandwidth scalability for a fixed number of topology nodes and leaves scalability issues with the number of nodes for future work.

With respect to the illustrated literature, our work quantifies to which extent layout effects and placement constraints in a realistic processor-memory communication setting cause the insertion-loss of the physical topology to deviate from its logic one.

## III. Target 3D-Architecture

In this section we describe the 3D-architecture of a multi-core processor including the passive optical layer for processor-memory communication whose detailed design is presented in section V.

As illustrated in Fig.1, an electronic layer positioned at the bottom of this 3D-integrated system consists of an array fabric of homogeneous processor cores. Similar architectures are already available in the market, e.g. the Tilera family of multi-core processors [1], which currently features arrays of 16, 36, 64 and 100 cores.

We consider an electronic layer that consists of 64 cores connected by an electronic NoC with a 2D mesh topology. We assume that cores are grouped into 4 clusters $C_i$ of 16 cores each. Every cluster has its own access to the optical layer which is vertically stacked on top of the electronic layer. We refer to the network interfaces between the electronic and the optical layer as the *hubs* $H_i$ and the number of cores inside each cluster as the $aggregationfactor$. This factor is design- and technology- dependent, since the cost (power and latency) for domain crossing dictates the most convenient boundary between the electronic and the optical NoC for cost-effective long range communication. Identifying this boundary is outside the scope of this paper. The optical layer accommodates three kinds of communications:

(a) between a pair of clusters;

(b) from a cluster to a memory controller of an off-chip DRAM DIMM;

(c) from a memory controller to a cluster.

We assume that the optical power is provided by an array of off-chip continuous wave (CW) lasers and that multi-wavelength signals are coupled into the chip and brought to the initiators for modulation. As demonstrated later on, the same array of CW lasers can be shared by all the initiators. For the given system configuration, 4 lasers are sufficient since every initiator modulates the same 4 wavelengths. This way, we are able to connect 8 initiators (4 hubs, 4 memory controllers) with 8 targets (the target interface of the same 4 hubs and 4 controllers).

The microarchitecture of memory controllers depends on the specific implementation of the memory sub-system. As an example, in [12] optical command, read and write busses connect the controller to the off-chip photonically integrated DRAM (PIDRAM) DIMMs via a fiber ribbon. In any case, the memory controllers are typically placed all around the chip. Their specific location depends on the position of the PIDRAM DIMMs on the board and can be optimized to reduce contention (hot spots) in the on chip interconnect fabric. We assume that memory controllers are located pairwise at opposite positions of the chip like in the architecture in [1] thus reflecting a common industrial practice (see Fig.1). The specific topology of the passive optical NoC is discussed in section V.

The electro-optical network interface (NI) resides partly in the electronic layer and partly in the optical one. Fig.2 shows

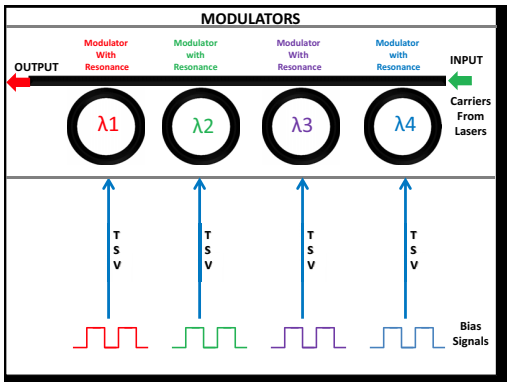



Fig. 3. Array of modulators in the optical layer

the transmission part that is implemented in the electronic layer. Packets coming from the cluster's electronic NOC are buffered at the network interface front-end. Based on their destination, they are stored in distinct buffers (in our target system, there are 7 buffers associated with the other clusters and with the memory controllers). A serializer reads packets from the buffers and feeds them to the drivers.

We assume that drivers are directly connected to the through-silicon vias (TSVs) and through them to the modulators on the optical layer. The latest technological developments about 3D-integration enable TSVs with a pitch of 5um x 5um and therefore a large TSV integration density (up to 160K TSVs in a 10mmx10mm die). As reported in [2] [5], the TSVs can deliver high-speed transmission from 1 Gbit/s to 10 Gbit/s. This performance motivates our choice of using them to provide the biasing signal to the optical modulators in the optical plane (see Fig.3). The rationale behind this choice is to avoid integrating electronic devices in the optical layer. In turn this enables low-cost fabrication of this layer, a key requirement to make silicon photonics affordable in the future also for the embedded multi-core computing domain. In line with current technology, we assume modulation rates of 10 Gbit/sec for each wavelength. Therefore, the injection rate of every hub peaks at 40 Gbit/sec. In the optical layer we use passive, wavelength-routed networks: every destination-specific buffer in the electronic NI is associated with a different wavelength in the modulation array. By complementing this with a network made up of add-drop filters, contention-free optical communication is achieved with no latency overhead for arbitration, routing or circuit setup.

The reception part is specular. In the optical layer an array of add-drop filters for each hub feeds photodiodes that convert the optical signal back into the electrical domain. The photodiodes' outputs are conveyed to the transimpedance amplifiers in the electronic layer by means of TSVs. Again, we opt for not placing the electronic devices in the optical layer. Digital comparators and de-serializers complete the domain conversion. Buffers are associated with packet source and from here on the electronic network interface functions come into play (i.e., association of memory responses with memory requests, packetization for the electronic NoC).

## IV. Technology-Aware Network-level Simulation Framework

In order to design a passive optical layer for processor-memory communication, it is necessary to explore interconnect solutions at the network level. For optical NoCs, this calls for a simulation environment with modeling capability of the routing functionality (which is the key feature of a wavelength-routed network) and of wavelength division multiplexing. However, in optical NoCs the characteristics of the technology platform cannot be completely abstracted away because they determine the viability of a logic scheme. Finding this out during layout design is too late and an effective design iteration is not possible any more. Unlike electronic NoC design, where such an iteration may simply consist of a new link inference technique (e.g., insertion of a pipeline stage on a slow link), in optical NoCs the number of waveguide crossings in the physical topology may be so large that the initial logic scheme may have to be entirely replaced by a more technology-friendly one. We propose a SystemC modeling and simulation environment as a good candidate to accommodate multiple requirements: efficient network-level simulation, support for technology annotations, compatibility with industrial frameworks for system-level design of the electronic part. Our modeling strategy is based on four main steps:

(a) Finite-Difference Time-Domain (FDTD) simulations of basic building blocks of optical NoCs, including straight and bending waveguides, 1x2 and 2x2 photonic switching elements (PSEs).

(b) derivation of the equations of the analytical governing models describing the optical behaviour of PSEs.

(c) back-annotation of the analytical model in SystemC and validation with FDTD. Accuracy validation of abstract models was proven in [18]: the mean-squared error is below 2% across the entire optical spectrum (1500-1600nm).

(d) modeling of higher-order switching structures in SystemC by means of a compositional approach. FDTD simulations of such higher-order structures are not practical because require very long simulation times. A SystemC modeling framework is reported in [27], where separate channels are used to model wavelength and power information of optical signals. In [28], a new SystemC class is created to manage analog signals transmitted between modules. We leverage the existing port-interface-channel constructs of SystemC, thus making the top-level view of an optical NoC look like the same of an electronic NoC: the difference lies just in module implementations and in the data types exchanged through the pre-defined SystemC channels.

The optical link model is at the core of our SystemC modeling framework. The sc_signal channel is instanced with a data type modeling the relevant features of an optical link: logic value, optical wavelength and signal amplitude. The optical wavelength is used by the router model for routing decisions, while the signal amplitude is used to preserve technology awareness inside the same router model. In fact, the analytical

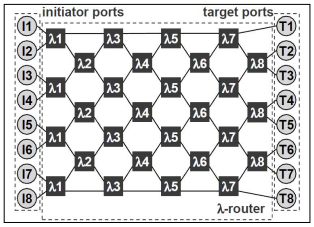



Fig. 4. Logic scheme of the 8x8 PPNoC topology from [20].

model outputs insertion-losses and crosstalk noise that affect the value of such amplitude. Back-annotated losses from FDTD simulations include waveguide crossing loss, drop-into-a-ring-loss, waveguide propagation and bending loss. By comparing the logic value with the signal amplitude the opto-electronic receiver reads out of the optical link, we get bit-error rate indications. The link model can support WDM by simply extending the user-defined data type to represent multiple wavelengths (and associated logic values and signal amplitudes) which might be propagating at the same time across an optical link.

## V. Passive optical NoC design

The most straightforward solution to interconnect the hubs with each other and with the memory controllers consists of an 8x8 passive optical NoC. We use the NoC topology proposed by O'Connor et al. [20] (see Fig.4), which we will hereafter refer to as the 8x8 PPNoC. In order to interconnect 8 initiators with 8 targets, the topology instantiates 8 stages of 4 and 3 add-drop filters. The only difference with respect to the original scheme is that we replace their 2x2 add-drop filters with a 2x2 photonic switching element [7]. This consists of 2 micro-ring resonators and two straight waveguides in orthogonal position. This choice eases layout design while leaving the routing functionality unchanged. Unfortunately, the appealing logic scheme of the topology does not match the actual floorplan in real-life systems. This is subject to specific constraints which may lead to a physical topology radically different from the logic one. In our target architecture, location of the hubs is dictated by the position of the electro-optical network interface in the electronic layer, which in turn depends on the aggregation factor. As a consequence, we have four hubs that are located along a square in the middle of the optical layer. In contrast, memory controllers are placed pairwise in opposite directions at the boundary of the chip.

Since the layout of the optical NoC must satisfy these physical constraints, it necessarily results in a larger number of waveguide crossings than in the logic scheme. Such crossings increase the total insertion-loss and hence the power of the optical laser which is required by the optical signal to stay above the minimum detection threshold at the photodiodes.

Such layout constraints, which are sometimes overlooked by theoretical studies of optical NoC topologies, are instead the main target of this paper. In Fig.5 we report the actual layout of the 8x8 PPNoC that we obtained given the layout constraints of our target system. It should be observed that each hub and

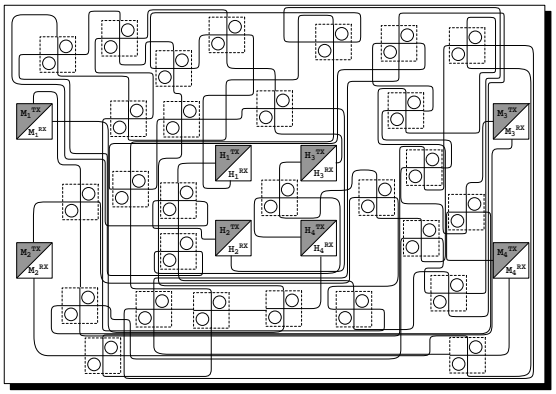
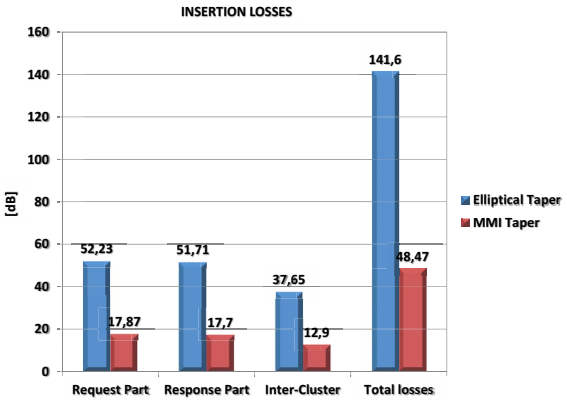


Fig. 5. Layout of Global 8x8 PPNoC

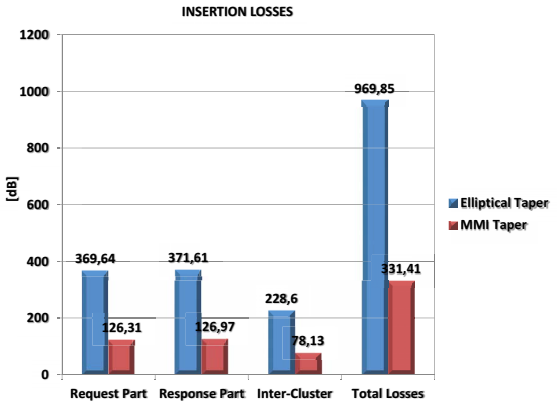
each memory controller is both initiator and target for the network, which should be therefore folded accordingly.

We adopted design guidelines similar to those followed for the layout of fat-tree topologies in electronic NoCs [25]. As a result, we were able to minimize waveguide crossings, homogeneously exploit the floorplan and avoid intricate routes for the silicon waveguides. The deviation of the physical topology from the logic one can be quantified by the insertion-loss critical path. It grows from 7 crossings in the logic scheme to 64 in the physical one. This effect is then reflected in the total insertion-loss of the topology.

Fig.6(a) and Fig.6(b) report total losses for the logic and physical topology, respectively. Such worst case losses account for all the wavelengths, which are therefore assumed to be used by all initiators at the same time. We consider both elliptical tapers [31] at waveguide crossings and MMI (Multimode Interference) tapers [32], two common physical implementation options. The experiments have been carried out by means of the SystemC modeling framework of optical devices illustrated in section IV. In Fig.6(b), the breakdown of the losses into request path (i.e., memory access requests from the hubs to the memory controllers), response path and inter-cluster communication allows us to assess the quality of the layout. In fact, request and response paths show similar losses denoting the symmetric layout of the topology. Inter-cluster losses are lower due to the physical proximity of the hubs. Notice the significant amount of total losses and the relevant savings obtained with MMI tapers, which are however not capable of cutting down losses below 48 dB in the ideal case. The picture becomes worse when we move to the real layout in Fig.6(b), where even with MMI tapers the total losses are more than 7 times higher, achieving 331 dB. This is clearly unacceptable for power-efficient implementations. This experimental insight suggests that other solutions should be researched. One idea is to partition the global PPNoC into three sub-networks, each dedicated to a different traffic class. By scaling down the same topology to 4 initiators and 4 targets, we derived the network for memory access requests. Similarly, we designed the network for memory responses, which features initiators and targets at flipped positions in the layout. Finally, we opted for a different topology for inter-cluster communication, where




(a) Global 8x8 PPNoC with ideal layout




(b) Global 8x8 PPNoC with real layout

Fig. 6. Insertion-loss analysis for the 8x8 PPNoC.

a scheme better matching the hub-delimited square shape of the available floorplan is needed. For this purpose, we selected the 4x4 GWOR topology [8]. This is a scalable and non-blocking passive optical router design using micro-ring resonators with four bidirectional ports located in the North, West, South and East directions. Also, two horizontal and two vertical waveguides are used, which is a valuable property for a squared floorplan.

The resulting layout is shown in Fig.7. It is evident that this layout is much less intricate with a lower number of additional crossings even for the 4x4 request and response PPNoCs. More precisely, the request PPNoC has the same number of crossings both in the logic and in the physical topology (3), while the response one features just a couple of additional crossings due to the opposite placement of initiators and targets on the optical layer. The improvement of the partitioned network with respect to the global network is quantified in Fig.8. Regardless of the specific taper configuration, the total insertion-losses are reduced by 21x in the partitioned PPNoC. Another advantage of partitioned networks which should not be overlooked is the reduction of the number of CW lasers. In the 8x8 PPNoC, every initiator modulates the same 8 wavelengths, thus requiring 8 external laser sources. With the

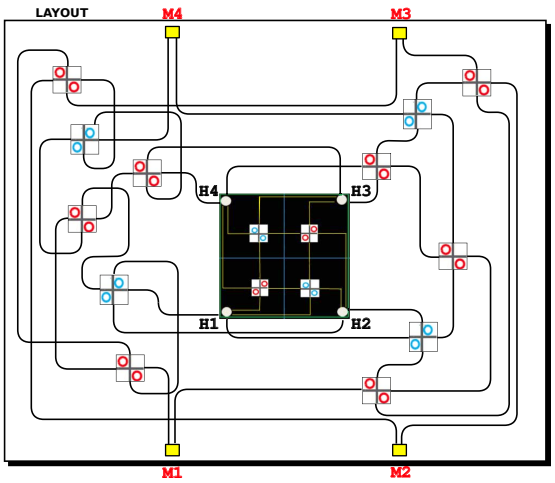



Fig. 7. Layout of the optical layer with network partitioning

partitioned solution, wavelengths can be reused across the multiple networks, thus only 4 CW lasers are needed. The partitioned architecture is assumed in the following sections.

## VI. Bandwidth Scalability in Passive Networks

In our target system, we have so far assumed an aggregation factor of 16 resulting in 4 clusters for the 64 core system. Successive generations of the same system will integrate more cores. The number of hubs, however, does not necessarily need to be increased accordingly to amortize the cost for electro-optical conversion and for the optical NoC infrastructure support (e.g., laser sources, distribution network of the optical power). Instead, bandwidth scalability techniques to increase the peak injection rate of the hubs could be implemented.

The same considerations hold for the number of memory controllers, which could stay the same for a few generations. In fact, congruent multiples in memory bandwidth could come from the deeper integration of silicon photonics into the DRAM chip before reverting to multiple DRAM memory channels. Identifying the new aggregation factor that justifies the increase in the number of network hubs and of memory controllers remains as future work.

In contrast, we engineer the optical NoC for bandwidth scalability, so that the peak bandwidth can be increased to accommodate the memory traffic that the hubs aggregate from a larger number of cores. With the assumptions made in section III, the peak injection rate of each hub is 40 Gbit/sec. A cost-effective way to augment this network's bandwidth is to embed multiple virtual networks in the same set of waveguides, using spare wavelengths which may be available depending on the maturity of the technology. One possibility is to employ the technique proposed by Small et al. [9], which essentially places several wavelengths in the resonance band of a microring resonator. In that case, it is possible to route multiple bits of a message in parallel with little extra hardware: at each node, multiple modulators/detectors must tap separately on each of these wavelengths in order to inject/extract the bits of information; however the only

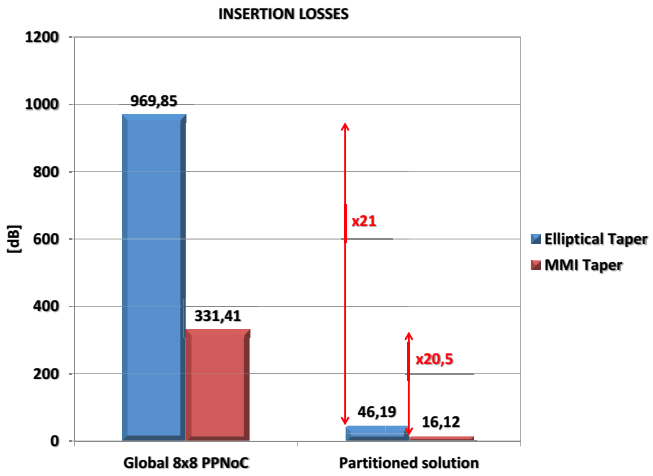



Fig. 8. Global 8x8 PPNoC vs Partitioned solution

change for the routers and filters is necessary to broaden the resonance band of their microrings, in order to correctly route such wavelength bundles. This solution, which we call *broadband passive switching, (BPS)*. obviously requires a larger number of off-chip sources to provide optical power to the new wavelengths.

Another way to achieve higher network bandwidth is simply to replicate the network (*spatial parallelism, SPM*). Notice that all replicated networks must be laid out in a way that minimizes waveguide crossings, which are a significant source of optical power losses. Multiple physical networks can be used to transfer more bits of the same message, or alternatively, more messages. With respect to BPS, SPM uses the same additional number of modulators and detectors but on different waveguides, while in BPS they increase the degree of wavelength division multiplexing on the same waveguides. Also, unlike BPS the number of laser sources stays the same, since the further replicated networks share the same wavelengths with the baseline ones. The output of the optical sources is just split among the individual network partitions. This advantage is compensated by the need to provide enough optical power for each wavelength to feed all network partitions. In principle, the total power provided by the optical source sub-system should be more or less the same, since in all cases the networks are (virtually or physically) replicated: the insertion-loss comes either from new wavelengths on the same waveguides (BPS) or from the same wavelengths on new waveguides (SPM). We experimentally measured the total insertion-loss of PBS and SPM to be around 12 dB, with only minor (less than 0.3dB) differences associated with the slightly different response of switching elements as a function of the signal wavelength.

The two solutions differ for the actual layout implementation, which has a critical impact on power losses, particularly for SPM. To quantify such effect we designed the layout of the SPM solution and accounted for the additional waveguide crossings in the technology-annotated SystemC simulation.

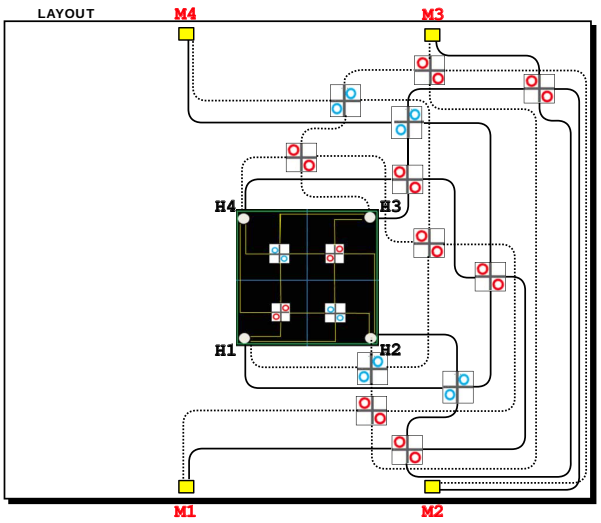



Fig. 9. Layout of the Request Passive Network with Spatial Parallelism

For the sake of clarity, Fig.9 reports only the layout for the replicated request PPNoC. Similar considerations apply for the response and the inter-cluster networks. For BPS, the layout remains the same as in Fig.7.

Total insertion-losses across all wavelengths/network partitions in the bandwidth-scaled request PPNoC are reported in Fig.10. These losses are not comparable with those in Fig.8 since the new plot refers to an injection rate from each hub that has been doubled and now peaks at 80 Gbit/sec.

The plot clearly shows that only BPS preserves the nominal insertion-loss of around 12 dB, while it grows up to 3x in SPM because of the waveguide crossings that the real layout constraints impose. Even with the MMI taper optimization, SPM is not able to go below 39 dB of total insertion-loss.

This behaviour is reflected into the critical path of the two solutions: SPM has a critical path insertion-loss which is 4 times larger than BPS.

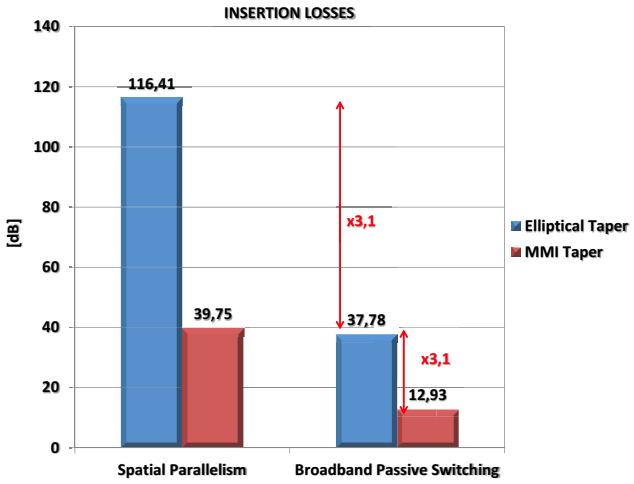



Fig. 10. Contrasting SPM vs BPS: total insertion-loss

## VII. Conclusion

In this paper we have quantified the deviation between quality metrics of logic topology as opposed to physical ones. This latter stems from the mapping of the logic connectivity scheme onto the real layout subject to placement constraints of communication actors and their network interfaces. As a case study, we consider processor-memory communication in a 3D system. We came up with three sets of results. First, insertion losses in the physical topology were one order of magnitude larger than expected because of the high number of waveguide crossings needed to lay it out. Second, we provided well-grounded results supporting optical NoC partitioning instead of full connectivity, materializing around 20x lower insertion losses as well as an effective reuse of wavelengths and off-chip laser sources. Third, we compared SPM with BPS as bandwidth scalability techniques. Real layout constraints heavily penalized SPM, since the additional waveguide crossings made insertion-losses 3x larger than in the nominal case. In contrast, BPS preserved such nominal values at the cost of more optical sources. The above practical insights enabled us to engineer a complete optical layer for high-bandwidth, low-latency and low-cost processor-memory communication in a 3D multi-core system.

## Acknowledgements

This work has been partially supported by the PHOTONICA project (under the "FIRB-Futuro in Ricerca" program, funded by the Italian Government) and by the National Science Foundation (under Award number: 5-25083).

## References

[1] D. Wentzlaff et al., "On-Chip Interconnection Architecture of the Tile Processor", IEEE Micro, vol.27, n.5, pp. 15-31, November 2007.

[2] J.U. Knickerbocker et al., "Three-dimensional silicon integration", IBM Journal of Research and Development, vol.52, n.6,pp. 553-569, November 2008.

[3] T. Karnik, D.Somasekhar, S.Borkar, "Microprocessor system applications and challenges for through-silicon-via-based three-dimensional integration", IET: Computer and Digital Techniques, vol. 5, n.3, pp.205-212, May 2011.

[4] M. Koyanagi et al., "Future system-on-silicon LSI chips", IEEE Micro, vol.18, n.4, pp.17-22, August 1998.

[5] J. So Pak, C. Ryu, J. Kim, "Multi-Stacked Through-Silicon-Via Effects on Signal Integrity and Power Integrity for Application of 3-Dimensional Stacked-Chip-Package", XXIX General Assembly of International Union of Radio Science (URSIGA 2008), August 2008.

[6] H. Wang et al., "Nanophotonic Optical Interconnection Network Architecture for On-Chip and Off-Chip Communications", OFC/NFOEC'08: Conference on Optical Fiber communication/National Fiber Optic Engineers, May 2008.

[7] A. Shacham, K. Bergman, L P. Carloni, "On the Design of a Photonic Network-on-chip", NOCS'07: International Symposium on Networks-on-Chip, May 2007.

[8] X. Tan et al., "On a Scalable, Non-Blocking Optical Router for Photonic Networks-on-Chip Designs", Photonics and Optoelectronics (SOPO), May 2011.

[9] B.A. Small, B.G. Lee, K.Bergman, Q.Xu and M.Lipson, "Multiple-wavelength integrated photonic networks based on microring resonator devices", Optical Networking Journal, vol.6, n.2, pp.112-120, 2007.

[10] N. Kirman and J. F. Martinez, "A Power-efficient All-optical On-chip Interconnect Using Wavelength-based Oblivious Routing", ASPLOS'10: International Conference on Architectural Support for Programming Languages and Operating Systems, March 2010.

[11] M. J. Cianchetti, J. C. Kerekes, and D. H. Albonesi "Phastlane: A rapid transit optical routing network", ISCA'09: International Symposium on Computer Architecture, June 2009.

[12] S. Beamer et al., "Re-Architecting DRAM Memory Systems with Monolithically Integrated Silicon Photonics", ISCA'10: International Symposium on Computer Architecture, June 2010.

[13] A. Shacham, K. Bergman, and L P. Carloni "Photonic Networks-on-Chip for Future Generations of Chip Multiprocessors", IEEE Trans. on Computers, vol.57, n.9, pp. 1246-1260, September 2008.

[14] D. Vantrease et al., "Corona: System implications of emerging nanophotonic technology", ISCA'08: International Symposium on Computer Architecture, June 2008.

[15] B. Akesson et al., "Memory Controllers for High-Performance and Real-Time MPSoCs Requirements, Architectures, and Future Trends", CODES+ISSS'11: International Conference on Hardware/Software Codesign and System Synthesis, October 2011.

[16] C. Batten et al., "Building manycore processor-to-DRAM networks with monolithic CMOS silicon photonics", HOTI'08: IEEE Symposium on High Performance Interconnects, September 2008.

[17] A. Hadke et al., "OCDIMM: Scaling the DRAM memory wall using WDM based optical interconnects", HOTI'08: High-Performance Interconnects Symposium, August 2008.

[18] A. Parini, L. Ramini, G. Bellanca and D. Bertozzi "Abstract Modelling of Switching Elements for Optical Networks-on-Chip with Technology Platform Awareness", INA-OCMC'11:International Workshop on Interconnection Network Architecture on Chip, Multi-Chip, January 2011.

[19] I. O'Connor et al., "Towards Reconfigurable Optical Networks on Chip", ReCoSoC 2005, pp.121-128.

[20] A. Scandurra and I.O'Connor, "Scalable CMOS-compatible photonic routing topologies for versatile networks on chip", Network on Chip Architecture, 2008.

[21] S. Le Beux et al., "Multi-Optical Network-on-Chip for Large Scale MP-SoC", IEEE embedded systems letters, vol.2, n.3, pp. 77-80, September 2010.

[22] S. Le Beux, J.Trajkovic, I.O'Connor and G.Nicolescu, "Layout Guidelines for 3D Architectures including Optical Ring Network-on-Chip (ORNoC)", VLSI-SoC'11: International Conference on VLSI and System-on-Chip, October 2011.

[23] S. Koohi, M.Abdollahi, S.Hessabi, "All-Optical Wavelength-Routed NoC based on a Novel Hierarchical Topology", NOCS'11: International Symposium on Networks-on-Chip, May 2011.

[24] S. Koohi, S.Hessabi,S.J.B.Yoo, "An Optical Wavelength Switching Architecture for a High-Performance Low-Power Photonic Network on Chip", Advanced Information Networking and Applications (WAINA), 2011 IEEE Workshops of International Conference, March 2011.

[25] D. Ludovici et al., "Assessing Fat-Tree Topologies for Regular Network-on-Chip Design under Nanoscale Technology Constraints", DATE'09: Conference on Design, Automation and Test in Europe, April 2009.

[26] Y. Xie, "Processor Architecture Design Using 3D Integration Technology", VLSID'10: International Conference on VLSI Design, January 2010.

[27] E. Drouard et al., "Optical Network on chip Multi-Domain modelling using SystemC", FDL 2004, pp.123-135.

[28] M. Briere et al., "Heterogeneous Modelling of an Optical Network-on-Chip with SystemC", IEEE Int. Workshop on Rapid System Prototyping, pp.10-16, 2005.

[29] Y. Pan et al., "Firefly:Illuminating future network-on-chip with nanophotonics", ISCA'10: International Symposium on Computer Architecture, June 2009.

[30] A. Joshi et al., "Silicon-photonic Clos networks for global on-chip communication" NOCS'09: International Symposium on Networks-on-Chip, May 2009.

[31] N. Sherwood-Droz et al., "Optical 4x4 hitless silicon router for optical Networks-on-Chip (NoC)", Opt. Expr., vol. 16, n. 20, pp. 15915-15922, 2008.

[32] H. Chen and A.W Poon, "Low-Loss Multimode-Interference-Based Crossings for Silicon Wire Waveguides", Photonics Technology Letters,IEEE., vol. 18, n. 21, pp. 2260-2262, 2006.