SPECIAL ISSUE PAPER

Capturing the sensitivity of optical network quality metrics to its network interface parameters

Marta Ortín-Obón[1,*,†], Luca Ramini[2], Víctor Viñals[1] and Davide Bertozzi[2]

[1]*DIIS, University of Zaragoza, Zaragoza, Spain*
[2]*ENDIF, University of Ferrara, Ferrara, Italy*

SUMMARY

Optical networks-on-chip (ONoCs) are gaining momentum as a way to improve energy consumption and bandwidth scalability in the next generation multicore and many-core systems. Although many valuable research works have investigated their properties, the vast majority of them lack an accurate exploration of the network interface architecture required to support optical communications on the silicon chip. The complexity of this architecture is especially critical for a specific kind of ONoCs: the wavelength-routed ones. These are capable of delivering contention-free all-to-all connectivity without the need for path reservation, unlike space-routed ONoCs. From a logical viewpoint, they can be considered as full nonblocking crossbars; thus, the control complexity is implemented at the network interfaces. To our knowledge, this paper proposes the first complete network interface architecture for wavelength-routed optical NoCs, by coping with the intricacy of networking issues such as flow control, buffering strategy, deadlock avoidance, serialization, and above all, their codesign in a complete architecture. The evaluation methodology spans from area and energy analysis via actual synthesis runs in 40-nm technology to RTL-equivalent (register-transfer level) SystemC modelling of the network architecture and aims at verifying whether the projected benefits of ONoCs versus their electrical counterparts are still preserved when the complexity of their network interface is considered in the analysis. Copyright © 2014 John Wiley & Sons, Ltd.

1. INTRODUCTION

Optical interconnect technology is nowadays mainstream in data centres to overcome the bandwidth density and the communication power concern. At a smaller scale, there are lots of ongoing development activities to exploit it for chip-to-chip communication [1], especially in the context of the network-in-package paradigm [2]. The current research frontier for on-chip interconnection networks consists of assessing the feasibility of the optical interconnect technology by exploiting the recent remarkable advances of silicon photonics [3]. The literature on this topic is becoming quite rich, mainly projecting superior bandwidth, latency and energy with respect to electrical wires beyond a critical length [4]. This benefits are extended to on-chip communication architectures, either as standalone optical networks (ONoCs) [5] or as hybrid interconnect fabrics [6]. Nonetheless, projected quality metrics are overly optimistic for a number of reasons extensively discussed in [7], including optimistic technology assumptions, use of logical topology designs instead of physical ones and overlooking static power. A big approximation of many projected results is the lack of a complete and accurate network interface architecture for driving on-chip optical communication, which may account for a large fraction of the overall network complexity. This is especially

---

*Correspondence to: Marta Ortín-Obón, gaz-DIIS-i3A, University of Zaragoza, Zaragoza, Spain.
†E-mail: ortin.marta@unizar.es

true for a particular category of optical networks-on-chip: the wavelength-routed ones (WRONoCs). These networks deliver contention-free global connectivity without the need for arbitration or routing. They achieve this goal by replicating the amount of wavelengths and associating each of them with a different and nonconflicting optical routing path. Despite the limited scalability, these networks are attractive for specific application domains, where performance predictability and ultralow latency communications are a must (e.g. data centre applications [8]).

Wavelength-routed optical NoCs can be conceptualized as nonblocking full crossbars. Therefore, all the complexity of the control architecture is located at the boundary of the interconnect fabric. To our knowledge, no complete network interface (NI) architecture has been reported so far in the open literature, with the exception of NIs for space-routed ONoCs. However, these are conceptually simpler because of the intuitive conversion of electrical bit parallelism into optical wavelength parallelism [9]. In contrast, WRONoCs rely on serialization or on a limited bit parallelism, which questions the achievement of performance goals. Even neglecting this difference, the NI design for an optical medium is a nontrivial task because of the interdependent issues that come to the forefront: end-to-end flow control, buffer sizing, clock resynchronization and serialization ratio. This paper takes on the challenge of designing and characterizing the complete NI architecture for emerging WRONoCs, in an attempt to validate whether (and to what extent) the projected benefits of optical NoCs over their electrical counterpart are still preserved with the NI in the picture. The distinctive feature of this work is the completeness of the architecture, including both initiator and target side. Especially, the digital architecture to enable optical NoC operation has been designed out of state-of-the-art basic building blocks (e.g. mesochronous synchronizers and dual-clock FIFOs), thus reflecting realistic quality metrics. The system-level requirements of a target multicore processor with cache-coherent memory architecture have a large impact on the interface footprint. Finally, for the optical and opto-electronic components, we have used a consistent set of static and dynamic power values from the same literature source [7, 10].

Our evaluation methodology consists of two steps: First, we synthesize and characterize latency and power for all the architecture components on a low power industrial 40-nm technology; second, we set up a complete SystemC-based simulation infrastructure (for both the optical and electronic parts) with RTL-equivalent accuracy, thus enabling to capture fine grained performance effects associated with the microarchitecture.

2. RELATED WORK

Early ONoC evaluation studies rely on coarse, high-level models or unrealistic traffic patterns [11–14], while more recent ones come up with complete end-to-end evaluations using real application workloads [15–18] or more accurate optical network models [17, 19–21]. Looking in retrospect, early results have been only partially confirmed, nonetheless showing the potential of ONoCs for on-chip communication. For instance, with an aggressive electrical baseline technology, it became more difficult to make a strong case for purely on-chip nanophotonic networks [18]. However, even in this case, it was still possible to show significant potential in using seamless intrachip/interchip nanophotonic links. Moreover, other works (such as [7]) related network energy to total system energy, thus making the point for fast interconnect fabrics capable of cutting down the static energy of non-network components, although they are themselves not energy efficient.

The refinement of comparative analysis frameworks is far from stabilizing. In fact, other missing aspects are progressively coming to the forefront as the ONoC research concept strives to become an industry-relevant technology. So far, the NI architecture has been overlooked in most evaluation frameworks or, in the best case, only considered in the early stage of design. Some pioneer works account for the NI in their network analysis for wavelength-routed optical networks [7, 15, 22–24] or space-routed ONoCs [9]. In every case, they suffer from one of the following weaknesses: First, they model NI components only at behavioural level [24], or they target only the more abstract level of formalization of interface specification [23]; second, they consider only the signal driving section of the NI, basically up to the (de)serializers. This way, higher level network architecture design issues such as flow control, synchronization or buffering are overlooked.

This work makes a relevant step forward with respect to NI design practice for emerging optical NoCs, thus enabling their future inclusion in network or system-wide evaluation frameworks. The distinctive features of our approach are architecture completeness, comparison with electrical inter-face counterparts, physical synthesis of digital components, RTL-equivalent SystemC modelling for microarchitectural performance characterization and analysis of the impact of NI parameters on global network quality metrics.

3. BACKGROUND ON WRONoCs

Wavelength-routed optical NoCs rely on the principle of wavelength-selective routing. As it is conceptually showed in Figure 1, every initiator can communicate with every target at the same time using different wavelengths. For instance, initiator I1 uses wavelengths 1, 2, 3 and 4 to reach targets 1, 2, 3 and 4, respectively. The topology connectivity pattern is chosen to ensure that wavelengths will never interfere with each other on the network optical paths. This way, all initiators can communicate with the same target by using different wavelengths. WRONoCs support contention-free all-to-all communication with a modulation speed of 10 Gbps/wavelength.

Our NI can work with any WRONoC topology. Without lack of generality, we model a wavelength-routed ring inspired by Le Beux *et al.* [25] implemented on an optical layer vertically stacked on top of the baseline electronic layer.

4. TARGET ARCHITECTURE

During the design of the NI, we consider a high-impact system requirement: message-dependent deadlock avoidance. Message-dependent deadlock arises from the interactions and dependencies created at network endpoints between different message types [26, 27]. Figure 2 shows the dependence between a request and response at the NI. In a complete system, the combination of these effects may lead to cyclic dependencies. Message-dependent deadlocks, once they occur, block resources at both network endpoints and inside the network indefinitely, even if an algorithm is used to avoid routing-dependent deadlocks in the network-on-chip.

When we apply these considerations to WRONoCs, the problem gets simplified by the fact that there is no buffering inside the network, which means messages do not stop along the path and, there-fore, cannot get blocked. Therefore, the ONoC automatically satisfies the consumption assumption,
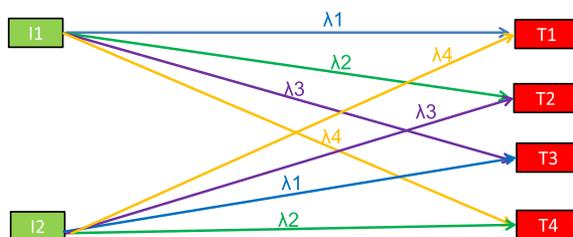

Figure 1. Logical view of wavelength-selective routing. All initiators can communicate with all the targets simultaneously by using nonconflicting wavelengths.
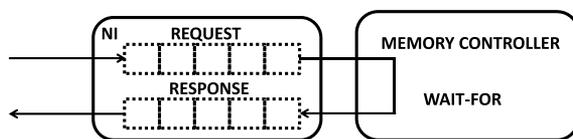

Figure 2. Dependence between a request and response at the NI.

which is a necessary (but not sufficient) condition for deadlock avoidance [27]. To enforce the sufficient condition, we must break the dependency cycles at the boundaries of the NoC by allocating a different buffer for each kind of message in the NI. This has direct implications on the buffering architecture of our NI (that is, on the number of virtual channels), depending on the communication protocol the WRONoC needs to support.

As a consequence, we make an assumption on a target system architecture. Without lack of generality, we focus on a homogeneous chip multiprocessor with 16 cores, similar to the Tilera architecture [28]. Each core has a private L1 cache and a bank of the shared distributed L2 cache, both connected to a common NI through a crossbar. The system has directory-based coherence managed with a MESI protocol (with states Modified, Exclusive, Shared and Invalid). By analysing the dependency chains of the protocol and deadlock-free buffer sharing opportunities, we came up with a requirement of three virtual channels (VCs) for deadlock avoidance. Proof is omitted for lack of space.

5. NI ARCHITECTURE

This section presents, to the best of our knowledge, the first complete NI architecture for wavelength-routed optical networks, as depicted in Figure 3. As a consequence, the objective is not to present the best possible design point but rather to start considering the basic components, and deriving guidelines about which ones deserve the most intensive optimization effort. Clearly, ONoCs move most of their control logic to the NIs, which should therefore not be oversimplified with abstract models.

To avoid message-dependent deadlock, every NI needs separate buffering resources (VCs) for each one of the three message classes of the MESI protocol. This should be combined with the requirements of wavelength routing: each initiator needs an output for each possible target, and each target needs an input for each possible initiator. As a result, in an initial version of the NI, each initiator came with three FIFOs for each potential target and each target, with three FIFOs for
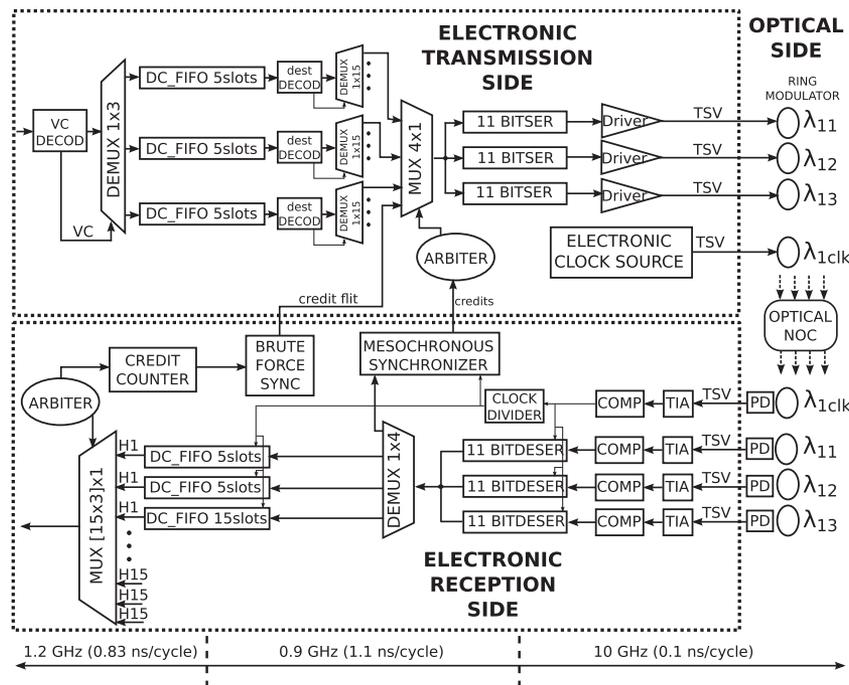

Figure 3. Optical network interface architecture for wavelength-routed optical NoCs. It supports communications with 15 destinations and sources (16 nodes with no self-communication) and 32 bits/flit, has three virtual channels and 3-bit parallelism. Each device runs at the frequency indicated at the bottom of the diagram.

each potential initiator. In a more energy-efficient version of the NI (Figure 3), the transmission side reuses the same three FIFOs for all destinations, and flits are dispatched to different paths afterwards (all the logic components after the 1×15 demultiplexers are replicated for each destination). For each one of those paths, there is an arbiter that grants access to the ONoC and keeps count of the empty slots at the reception buffers. Therefore, it is still possible to manage flow control independently for each initiator-target pair while sharing the transmission side buffers. This energy optimization will not cause relevant latency degradations, because the nodes produce packets sequentially (not in parallel for every destination) and the network is lightly loaded.

All the FIFOs at both the transmission and the reception side must be dual-clock FIFOs (DC FIFOs) to move data between the processor frequency domain (we assume 1.2 GHz) and the one used inside the NI. As hereafter explained, the latter depends on bit parallelism. We used the DC FIFO architecture presented in [29]. To size the DC FIFOs, we considered the size of the packets that would use each of the VCs: control packets need two flits, while data packets need 21 flits assuming flits are 32-bits long. The FIFO depth will be assessed in the experimental results, as well as the flit width. The minimum size for the DC FIFO to achieve perfect throughput is five slots [29], so all the VCs in the transmission side have been sized this way. For the reception side, we sized the data VC based on the round-trip latency in order to allow uninterrupted communications, ending up with 15-slot DC FIFOs. However, for the control VCs, we decided to keep small 5-slot DC FIFOs because they can already fit two complete packets, and we do not expect to send many back-to-back control packets with the target cache-coherence protocol.

After flits are sent to the appropriate path depending on their destination, they need to be translated into a 10-GHz bit stream in order to be transmitted through the optical NoC. This serialization process is parallelized to some extent to increase bandwidth and reduce latency; 3-bit parallelism means that three serializers of 11 bits each work in parallel to serialize the 32 bits of a flit, resulting on a bandwidth of 30 Gbps. The bit parallelism determines the frequency inside the optical NI: 1.1 ns (0.1*number of bits) is needed to serialize a flit with 3-bit parallelism, but only 0.8 ns is needed with 4-bit parallelism. In turn, this also impacts the size of the reception DC FIFO based on round-trip latency, which increases from 15 to 17 slots when moving from 3- to 4-bit parallelism.

Another key issue to be considered in the NI is the resynchronization of received optical pulses with the clock signal of the electronic receiver. In this paper, we assume source-synchronous communication, which implies that each point-to-point communication requires a strobe signal to be transmitted along with the data on a separate wavelength. With the current technology, this seems to be the most realistic solution, even considering the promising research effort that is currently being devoted to transmitting clock signals across an optical medium [30]. The source-synchronous clock is then used at the reception side of the NI to drive the deserializers and, after a clock divider, the front-end of the DC FIFOs. We assume that a form of clock gating is implemented, so when no data is transmitted, the optical clock signal is gated.

Another typically overlooked issue is the backpressure mechanism. We opt for credit-based flow control because it does not rely on timing assumptions, and credit tokens can reuse the existing communication paths. Besides, the low dynamic power of ONoCs can easily tolerate the signaling overhead of this flow control strategy. Credits are generated at the reception side of the NI when a flit leaves the DC FIFO (at the processor frequency) and forwarded to the transmission side so that they can be sent back to the source (at the NI frequency). In order to change from one frequency domain to the other, we opted for synchronizing the valid bits with a brute force synchronizer. For the scheme to work, the credit flit data must be stable for three NI cycles; during that time, credits are accumulated in credit counters. As soon as the credit flit arrives at the transmission side, it has priority over the flits from the VCs. The mandatory waiting time guarantees VCs will not suffer from starvation. To make a better use of the 32 bits of a flit, credits for all VCs of the same destination are sent together in the same credit flit. When credits arrive at the reception side of the source NI, they need to go through a mesochronous synchronizer to adapt the frequency derived from the received clock to the local NI frequency. Dedicated FIFOs for each source are needed at the reception side of the NIs to support this credit-based flow control. This is a clear candidate for future optimizations.

6. BASELINE ELECTRONIC NoC

The baseline electronic switch architecture is the consolidated ×pipesLite architecture [31], which represents an ultralow complexity design point. The ×pipesLite architecture is an input-buffered switch, with logic-based distributed routing and wormhole flow control. Every 32-bit switch includes three VCs to avoid message-dependent deadlock, with five slots each. VCs are implemented by replicating the basic switch architecture without VC capabilities three times, based on the approach presented in [32]. This way, the shorter critical path enables area optimizations during logic synthesis that offset the apparently larger amount of used resources. This technique results in higher maximum operating speed, better performance and smaller area. It takes one cycle to traverse the switch and one cycle to traverse each link.

The NI consists of two parts [18]. The first one is a packetizer, which acts as protocol converter from the IP-core protocol to the network one. This block is also required for the ONoC; therefore, it is not considered in this comparison framework and is not showed in Figure 3 either. The second one is the buffering stage. In order to preserve the generality of the design and support cores with different operating frequencies that access an electronic NoC (ENoC) with fixed common frequency, DC FIFOs have been included at the electronic NIs, similar to the ONoC NI design. However, in this case, all DC FIFOs have five slots at both initiator and target side, because round-trip latency does not require larger buffers for maximum throughput operation.

7. EVALUATION

This section characterizes the most important network quality metrics for the electro-optical NI: latency, throughput, static power and energy-per-bit. Results for an ENoC configured with typical parameters in [31] are also included. This aims to set the bases for a future comprehensive cross-benchmarking study, which is out of the scope of this paper.

7.1. Methodology

To obtain accurate latency results, we implemented detailed RTL models of the optical and electronic NIs and NoCs using SystemC [33]. We instantiated a 4×4 2D mesh for the ENoC and a similar system connected by the optical ring for the ONoC. The network-wide focus, well beyond the NI, aims at relating NI quality metrics to network ones. Delay values for the optical ring have been back annotated from physical-layer analysis results [34] and have been differentiated on a per-path basis. We omit from this implementation the logic to divide the packet into flits and reassemble it afterwards, because it is common for both the electronic and optical systems.

For power modelling, every electronic component has been synthesized, placed and routed using a low-power 40-nm industrial technology library. Power metrics have been calculated by back annotating the switching activity of block internal nets and then importing waveforms in the *PrimeTime* tool. We have applied clock gating to achieve realistic static power values. Energy-per-bit has been computed by assuming 50% switching activity. For the fast developing optical technology, we consider a coherent set of both conservative and aggressive values (obtained in [7, 10]). The photonic components and values are listed in Table I. Table II sums up the static power and energy-per-bit for all the electronic and optical devices (all DC FIFOs, independently of size and frequency, are reported to have the same static power as a consequence of clock gating). These values are only realistic under the assumption of low network contention, which reflects the typical operating condition of cache-coherent multicore processors [18].

7.2. NI latency breakdown

Figure 4 presents the latency breakdown for the NI components and the ONoC, obtained from our accurate RTL-equivalent simulations. The difference between data and control flits comes from the different number of flits in control and data messages and the different size of the corresponding virtual channel buffers. We clearly see that the latency of the network is negligible, but it requires

Table I. Photonic components and parameters with their values with aggressive and conservative technologies.

Parameter	Cons. tech.	Aggr. tech.
Coupler loss	0.46 dB	0.46 dB
Modulator insertion loss	4.0 dB	4.0 dB
Photodetector loss	1.0 dB	1.0 dB
Filter drop loss	1.0 dB	1.0 dB
Through ring loss	0.0001 dB/ring	0.0001 dB/ring
Propagation loss	1.5 dB/cm	1.5 dB/cm
Bending loss	0.0005 dB	0.0005 dB
Crossing loss	0.52 dB	0.18 dB
Wall-plug laser efficiency	8%	20%
Thermal tuning	20 uW/ring	20 uW/ring
Transmitter (dyn. energy)	50 fJ/bit	20 fJ/bit
Transmitter (fixed energy)	0.100 mW	0.025 mW
Receiver (dyn. energy)	25 fJ/bit	10 fJ/bit
Receiver (fixed energy)	0.150 mW	0.050 mW

Table II. Static power and dynamic energy of electronic and optical devices.

Hardware components	3-bit parallelism			4-bit parallelism		
	Count per NI	Static power (mW)	Dynamic energy (fJ/bit)	Count per NI	Static power (mW)	Dynamic energy (fJ/bit)
DC_FIFO 5slots (TX)	3	0.12	10.65	3	0.12	12.72
DC_FIFO 5slots (RX)	30	0.12	8.54	30	0.12	10.2
DC_FIFO 15-17 slots	15	0.12	26.50	15	0.12	31.65
DEMUX1 × 3	1	0.000725	0.92	1	0.000725	0.92
DEMUX1 × 15	3	0.0021	25.21	3	0.0021	25.21
DEMUX1 × 4	15	0.00056	6.72	15	0.00056	6.72
MUX4 × 1 + ARB	15	0.08	0.36	15	0.11	0.49
MUX45 × 1 + ARB	1	0.9	5.09	1	0.9	5.09
SERIALIZER	45	0.0475	9.41	60	0.0417	2.63
DESERIALIZER	45	0.0289	7.74	60	0.0281	6.12
MESO-SYNCHRONIZER	45	0.041	8.00	45	0.0565	11.1
COUNTER 2bits	45	0.01482	1.014	45	0.01482	1.014
BRUTE FORCE SYNC	15	0.004234	1.4	15	0.00503	1.66
CLOCK DIVIDER	15	0.01172	0.6	15	0.0139	0.714
TSV	120	/	2.50	150	/	2.50
TRANSMITTER aggressive	60	0.025	20	75	0.025	20
TRANSMITTER conserv.	60	0.100	50	75	0.100	50
RECEIVER aggressive	60	0.050	10	75	0.050	10
RECEIVER conservative	60	0.150	25	75	0.150	25
THERMAL TUNING /RING 20K	180	0.020	/	225	0.020	/
LASER POWER aggressive	/	0.0421	/	/	0.0525	/
LASER POWER realistic	/	0.308	/	/	0.385	/
E-SWITCH (3VCs)	/	17.9	193	/	17.9	193

support from a time-consuming NI. Inside the NI, the DC FIFOs are the components with the largest latency.

7.3. Transaction latency

We simulate the most common traffic patterns generated by an MESI coherence protocol in our RTL models without any contention. The increased accuracy of our analysis stems from the fact that our packet injectors and ejectors model actual transactions of the protocol, as well as their
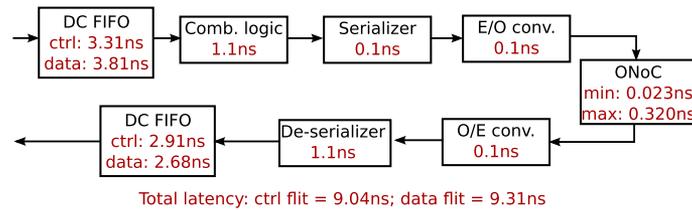
Figure 4. Latency breakdown to send a control or data flit through the network interface (with 32 bits/flit and 3-bit parallelism) and the optical ring.

Table III. Message patterns generated by the MESI coherence protocol.

id	Event	Sequence of messages
P1a	L1 miss	1. Request from L1 to L2 2. Data reply from L2 to L1 3. ACK from L1 to L2
P1b/c	L1 write miss, 1/2 sharers	1. Request from L1 to L2 2. L2 sends data reply and invalidates 1/2 sharers 3. Sharers sends ACK to L1 req. 4. ACK from L1 to L2
P2a	L1 needs upgrade to write	1. Request from L1 to L2 2. ACK reply from L2 to L1 3. ACK from L1 to L2
P2b/c	L1 needs upgrade to write, 1/2 sharers	1. Request from L1 to L2 2. ACK reply from L2 to L1 and invalidates 1/2 sharers 3. Sharers send ACK to L1 req. 4. ACK from L1 to L2
P3	L1 write miss, another owner	1. Request from L1 to L2 2. L2 forwards request to owner 3. Owner sends data to L1 4. ACK from L1 to L2
P4	L1 read miss, another owner	1. Request from L1 to L2 2. L2 forwards request to owner 3. Owner sends data to L1 and L2 4. ACK from L1 to L2
P5	L1 replacement	1. Writeback from L1 to L2 2. ACK from L2 to L1

interdependencies. Table III describes the analysed compound transactions, and Figure 5 presents the zero-load latency results. For the ENoC, we depict latencies obtained from the communication of nodes that are separated by the mesh minimal, maximum and average distance. For the ONoC, we present an average of several occurrences of each pattern, because latency changes slightly depending on the distance between the rising edges of the processor and internal NI clocks at the moment the transaction begins. The messages included in these patterns amount to an average 99.9% of the total network traffic, as we observed from full-system simulations of realistic parallel benchmarks from PARSEC and SPLASH2 and multiprogrammed workloads built with SPEC applications (we only exclude communication with the memory controllers). Therefore, they are a very good indicator of the network latency improvements we can expect from the optical network, including its (non-negligible) NI overhead.

We observe that in all the patterns except the last one, the ONoCs either beat or obtain equal results to the ENoC with all path lengths. As opposed to the ENoC, most of the latency of the ONoC
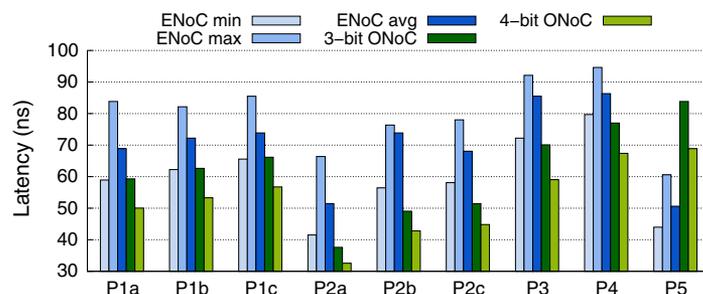
Figure 5. Zero-load latency of the most common communication patterns in the MESI protocol (as explained in Table III). For the ENoC, we include minimum, maximum and average paths.
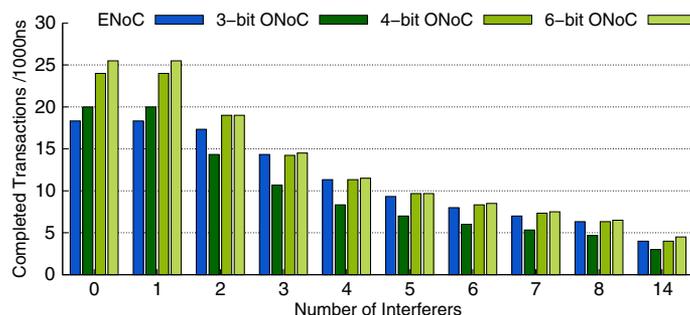

Figure 6. Number of completed transactions per 1000 ns between two nodes as the number of interferers increases.

is spent in the NI, which is needed to support the low latency optical communication. The tendency changes in pattern 5 because the replacement packet is using a VC designed for control to transmit data and the smaller FIFO cannot store enough flits to support the round-trip latency. However, this messages are only 7.4% of the total network traffic.

7.4. Throughput

In this section, we test the behaviour of the electronic and optical networks under contention. To do that, we focus only on requests and data replies. We leave the ACKs out because they are not in the critical path of the communications. We pick a node to be the main L1 and another node to be the L2, and count the number of completed transactions per second. Then, we gradually insert congestion into the network by having all the other nodes sending requests to the same L2 cache and keep monitoring the transactions just between the main L1 and the L2. All the L1s support only one outstanding transaction and inject a new request as soon as they receive the reply.

Figure 6 presents the results for the ENoC and the 3-, 4- and 6-bit parallelism ONoCs. Without contention, more transactions get completed in the optical NoC because their latency is lower. Including only one interferer does not affect results because all networks have enough bandwidth to support two concurrent L1 requestors at maximum throughput. As we keep increasing the number of interferers, the throughput for the 3-bit parallelism ONoC drops much faster than for the ENoC. This is because the former can eject a maximum of 30 Gbps, while the latter transmits flits at 38.4 Gbps. For this reason, the replies need to wait much longer until they can be transmitted. However, when considering the ONoC with 4-bit parallelism, which has a bandwidth of 40 Gbps, we see results comparable and even superior to those of the ENoC. At 6-bit parallelism, the increased bandwidth (60 Gbps) only gives the ONoC a slight advantage because performance is limited by the core frequency. This advantage is not enough to justify the increase in static power (as it will be documented later).

Table IV. Buffer sizes explored for the three VCs at each side of the NI.

id	Transmission side	Reception side
A	3, 3, 3	3, 3, 3
B	3, 3, 5	3, 3, 5
C	5, 5, 5	5, 5, 5
D	5, 5, 5	5, 5, 15
E	5, 5, 22	5, 5, 15
F	10, 10, 44	10, 10, 44

Note that the actual capacity of the DC FIFOs is one flit less than the number of slots, one slot is needed for the correct management of the DC FIFO.
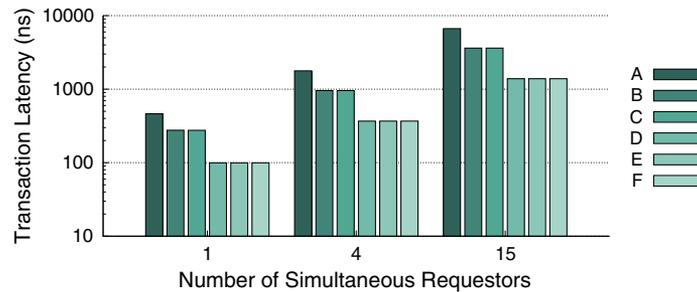

Figure 7. Average request–reply transaction latency with random uniform traffic testing various buffer sizes (as detailed in Table IV).

7.5. Buffer size exploration

In this section, we analyse the effect of modifying the buffering of the optical NI. We fix the bit parallelism at 3 and explore all the buffer size combinations detailed in Table IV. Using the same request-reply pattern as in the previous section with a maximum of four outstanding requests per node, we analyse how buffer size in the NI affects transaction latency. Results are depicted in Figure 7.

In case A, the minimum buffering has a very negative impact on performance because data packets are stalled waiting for credits from the reception side FIFOs, which can only store two flits (for a correct management of the DC FIFO, one slot is always left empty). This effect is slightly mitigated when we increase the buffer size for this VC to five slots in case B. Even though the DC FIFOs can achieve perfect throughput, backpressure is still preventing faster communications. We do not see any difference by increasing the size of control VCs in case C because the bottleneck is in the data VC. However, in case D, the reception side has been sized based on the round-trip latency, and we achieve the maximum throughput possible. The larger buffers in cases E and F do not show any further improvements because the network is already using up all the bandwidth.

7.6. Flit width exploration

In this section, we analyse the effect of modifying the flit width and the bit parallelism of the optical NI. Figure 8 presents the transaction latency for several combinations using a request–reply pattern, while increasing the injection rate. Four nodes that act as L1 caches send requests to the same L2 cache, with a maximum of four outstanding requests each. While the injection rate is low, there is enough time between requests to service all of them without degradation of the latency. When the injection rate increases, requests start accumulating at the L2 until a saturation point is reached. At that point, the number of outstanding requests is maximum, and we obtain the transaction latency under congestion. This saturation point arrives later with higher bit parallelism, and the maximum latency is lower because each transaction requires less time.
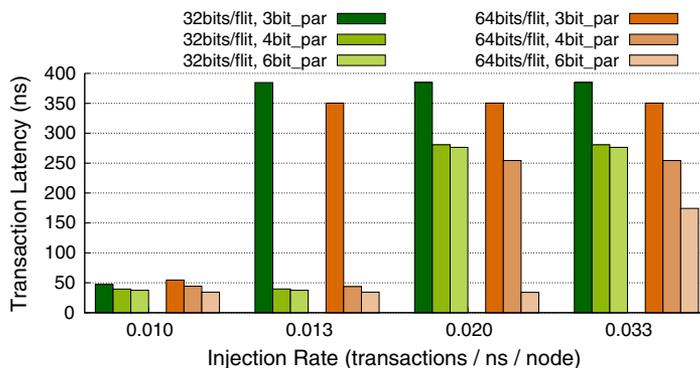
Figure 8. Average request–reply transaction latency with random uniform traffic testing several flit width and bit parallelism combinations.
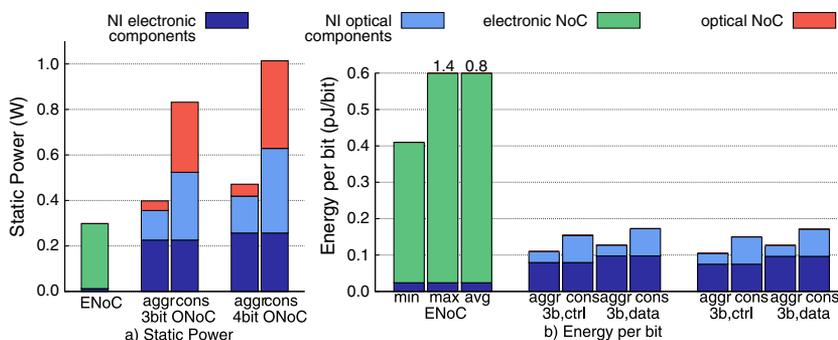

Figure 9. Static power and energy-per-bit of the NIs and the electronic and optical NoCs. For the optical NoC, we show results with 32 bits/flit and 3- and 4-bit parallelism and two sets of parameters: aggressive and conservative. For the energy of the ENoC, we include minimum, maximum and average path lengths.

If we compare results without congestion for 32- and 64-bit flits, we see that, for a given bit parallelism, latency is shorter with 32-bit flits. This is because, even though 64-bit flit packets have less flits, it takes longer to serialize all their bits for optical transmission. However, this trend is reversed under congestion because the packet serialization latency that determines the length of the queue at the L2 is shorter. Compared with the latency for 32-bit flits with 3-bit parallelism under no contention, 64-bit flits with 4- and 6-bit parallelism are 6.4% and 27.3% faster, respectively.

7.7. *Power and energy-per-bit*

Figure 9 depicts the static power and (dynamic) energy-per-bit for the ENoC versus the 3- and 4-bit parallelism ONoCs. We do not consider ONoCs with less than 3-bit parallelism because the bandwidth of the optical paths would be too low or ONoCs with more than 4-bit parallelism, because the static power becomes unacceptable (we can see a clear trend in Figure 9). We present a breakdown of the contributions of the NIs and NoCs. For the NI, we also separate the electronic components from the optical (and analogic) ones. The optical NoC is solely composed of laser power, so it has no impact on dynamic energy. In computing total power figures, we consider two sets of parameters for optical interconnect technology, corresponding to its high maturity (*aggressive* parameters) and to its low maturity (*conservative* parameters).

We observe that the electronic switches dominate the static power, accounting for 95.8% of the total. However, this trend is reversed in the ONoC, with a contribution of only 10.6% and 11.8% for the aggressive technology with 3- and 4-bit parallelism, respectively. It is worth highlighting that most of the static power of the electronic components in the NI comes from the DC FIFOs. Also, the savings in execution time of the ONoC versus the ENoC may compensate the higher static power
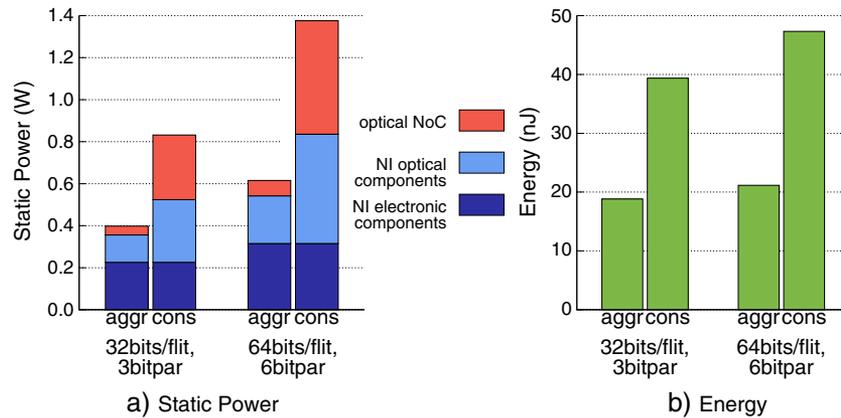
Figure 10. Static power and network energy comparing 32-bit flits with 3-bit parallelism versus 64-bit flits with 6-bit parallelism. We include two sets of parameters for the optical technology: aggressive and conservative. The network energy shows the energy needed to complete a request–reply transaction under no contention.

and result in overall energy reductions. This is especially true when we consider the power of the system as a whole, as claimed by Kurian *et al.* [18].

For energy-per-bit, we include minimum, maximum and average-length paths for the ENoC and specific values for control and data packets for the ONoC (which change because of the different size of the reception DC FIFOs). We clearly see that the ONoC has significantly lower energy-per-bit than the ENoC, which confirms the trend observed in the previous literature. Apart from that, we still see how the main contributor for the ENoC energy is the NoC, while the NI carries all the complexity for the ONoC.

Figure 10 shows how static power and energy change when modifying the flit width. We focus on the ONoC with 32-bit flits and 3-bit parallelism, and the ONoC with 64-bit flits and 6-bit parallelism. In both cases, it takes 1.1 ns to serialize the bits of the flit (with 64 bits, the flit width is double but so is the bit parallelism), but the latter needs less flits to transmit each packet. Therefore, it has better performance (as we showed in Section 7.6), and it will be interesting to explore the tradeoff with power and energy.

Static power for the 64-bit flit ONoC is 1.64 times larger than for the 32-flit ONoC with conservative technology, because of the larger number and size of electronic and optical components needed to support the increased flit width and bit parallelism. With aggressive technology, this factor is reduced to 1.54.

In order to check if the increased power is compensated by the reduced latency, we calculated the static energy burnt by the ONoC to complete a request–reply transaction under no contention. We consider only the contribution of the static power. This is by far the largest percentage of the total energy in optically enabled real cache-coherent systems, which typically experience very low traffic loads [18]. The 64-bit flit ONoC consumes 17% and 11% more energy than the 32-bit flit ONoC with the conservative and aggressive technologies, respectively. Therefore, we note that the improved transaction latency is not able to compensate or reverse the energy trend. However, when introducing IP-core static power into the picture, the conclusion may significantly change depending on the potential of the enhanced parallelism to cut down on system execution time. This is left for future work.

8. CONCLUSIONS

This paper presents an accurate design of NIs for WRONoCs, captures the effect on the most important network quality metrics and sets the scene for further improvements of comparative ONoC analysis.

Regarding latency, the ONoC is always faster than its electronic counterpart even considering the NI, thus preserving the primary goal of a WRONoC. The behaviour under contention depends mainly on the available bandwidth of the interconnect technologies under test. For the WRONoC, such bandwidth can be modulated by tuning the bit parallelism and adjusting buffer size to flow control requirements for maximum throughput operation. Similar tuning knobs do exist for ENoCs, namely, flit width and buffer sizes. Therefore, the ultimate question is whether such tuning knobs are energy efficient in comparative terms, which depends on the sensitivity of system performance to such knobs for the application at hand. This is left for future work.

When we consider power figures, we notice that, while switches are the main contributors in ENoCs, the NI has the largest share in ONoCs. For static power, this contribution is in the same order of magnitude than that from laser sources with conservative optical technology parameters. However, by further improving the optical technology, the role of the NI becomes dominant, thus making it the main target for future optimizations. Finally, the ONoC preserves its superior dynamic power properties over its ENoC counterpart, even in the presence of the NI.

This paper shows that the NI architecture should not be overlooked for realistic ONoC assessments and comes up with new insights not provided by earlier photonic network evaluations. The most important one is that NI optimizations perhaps have higher priority over the relentless search for ultra-low-loss optical devices.

REFERENCES

1. Koka P, McCracken MO, Schwetman H, Zheng X, Ho R, Krishnamoorthy AV. Silicon-photonic network architectures for scalable, power-efficient multi-chip systems. *Proceedings of the 37th Annual International Symposium on Computer Architecture*, ISCA '10, ACM, New York, NY, USA, 2010; 117–128. DOI: 10.1145/1815961.1815977.
2. Arakawa Y, Nakamura T, Urino Y, Fujita T. Silicon photonics for next generation system integration platform. *IEEE Communications Magazine* 2013March; **51**(3):72–77. DOI: 10.1109/MCOM.2013.6476868.
3. Galland C, Novack A, Liu Y, Ding R, Gould M, Baehr-Jones T, Li Q, Yang Y, Ma Y, Zhang Y, Padmaraju K, Bergmen K, Lim AE-J, Lo G-Q, Hochberg M. A CMOS-compatible silicon photonic platform for high-speed integrated opto-electronics. *Proceedings of Integrated Photonics: Materials, Devices, and Applications*, Grenoble, France, 2013; 87670G.
4. Kapur P, Saraswat KC. Optical interconnects for future high performance integrated circuits. *Physica E: Low-dimensional Systems and Nanostructures* 2003; **16**:620 –627. DOI: 10.1016/S1386-9477(02)00686-0. (Available from: http://www.sciencedirect.com/science/article/pii/S1386947702006860, Accessed on 26 June 2014).
5. Koohi S, Abdollahi M, Hessabi S. All-optical wavelength-routed NoC based on a novel hierarchical topology. *International Symposium on Networks on Chip (NoCS)*, Pittsburgh, PA, 2011; 97–104.
6. Chan J, Hendry G, Biberman A, Bergman K. Architectural design exploration of chip-scale photonic interconnection networks using physical-layer analysis. *2010 Conference on (OFC/NFOEC) Optical Fiber Communication (OFC), Collocated National Fiber Optic Engineers Conference*, San Diego, CA, USA, 2010; 1–3.
7. Batten C, Joshi A, Stojanovic V, Asanovic K. Designing chip-level nanophotonic interconnection networks. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 2012; **2**:137–153. DOI: 10.1109/JETCAS.2012.2193932.
8. Kapoor R, Porter G, Tewari M, Voelker GM, Vahdat A. Chronos: predictable low latency for data center applications. *Procs of the Third ACM Symposium on Cloud Comp.*, ACM, New York, NY, USA, 2012; 9:1–9:14. DOI: 10.1145/2391229.2391238.
9. Hendry G, Chan J, Kamil S, Oliker L, Shalf J, Carloni LP, Bergman K. Silicon nanophotonic network-on-chip using TDM arbitration. *Annual Symposium on High Performance Interconnects (HOTI)*, Mountain View, CA, 2010; 88–95. DOI: 10.1109/HOTI.2010.12.
10. Beamer S, Sun C, Kwon Y-J, Joshi A, Batten C, Stojanovic V, Asanovic K. Re-architecting dram memory systems with monolithically integrated silicon photonics. *ISCA*, ACM, New York, NY, USA, 2010; 129–140. (Available from: http://dblp.uni-trier.de/db/conf/isca/isca2010.html#BeamerSKJBSA10, Accessed on 26 June 2014).
11. Pan Y, Kumar P, Kim J, Memik G, Zhang Y, Choudhary A. Firefly: illuminating future network-on-chip with nanophotonics. *Proceedings of the International Symposium on Computer Architecture*, ISCA '09, ACM, New York, NY, USA, 2009; 429–440. DOI: 10.1145/1555754.1555808.

12. Vantrease D, Binkert N, Schreiber R, Lipasti MH. Light speed arbitration and flow control for nanophotonic interconnects. *42nd Annual IEEE/ACM International Symposium on Microarchitecture, 2009. Micro-42*, New York, NY, 2009; 304–315.

13. Joshi A, Batten C, Kwon Y-J, Beamer S, Shamim I, Asanovic K, Stojanovic V. Silicon-photonic clos networks for global on-chip communication. *International Symposium on Networks-on-Chip*, San Diego, CA, 2009; 124–133.

14. Pan Y, Kim J, Memik G. Flexishare: channel sharing for an energy-efficient nanophotonic crossbar. *International Symposium on High Performance Computer Architecture*, Bangalore, 2010; 1–12.

15. Morris R, Kodi AK, Louri A. Dynamic reconfiguration of 3d photonic networks-on-chip for maximizing performance and improving fault tolerance. *Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO-45, IEEE Computer Society, Washington, DC, USA, 2012; 282–293. DOI: 10.1109/MICRO.2012.34.

16. Bartolini S, Grani P. Co-tuning of a hybrid electronic-optical network for reducing energy consumption in embedded CMPs. *Proceedings of the First International Workshop on Many-Core Embedded Systems*, MES '13, ACM, New York, NY, USA, 2013; 9–16. DOI: 10.1145/2489068.2489070.

17. Koka P, McCracken MO, Schwetman H, Chen CHO, Zheng X, Ho R, Raj K, Krishnamoorthy AV. A microarchitectural analysis of switched photonic multi-chip interconnects. *Proceedings of the 39th Annual International Symposium on Computer Architecture*, ISCA '12, IEEE Computer Society, Washington, DC, USA, 2012; 153–164. (Available from: http://dl.acm.org/citation.cfm?id=2337159.2337177, Accessed on 26 June 2014).

18. Kurian G, Sun C, Chen CHO, Miller JE, Michel J, Wei L, Antoniadis DA, Peh LS, Kimerling L, Stojanovic V, Agarwal A. Cross-layer energy and performance evaluation of a nanophotonic manycore processor system using real application workloads. *International Parallel Distributed Processing Symposium (IPDPS)*, Shanghai, 2012; 1117–1130. DOI: 10.1109/IPDPS.2012.103.

19. Ramini L, Grani P, Bartolini S, Bertozzi D. Contrasting wavelength-routed optical NoC topologies for power-efficient 3d-stacked multicore processors using physical-layer analysis. *Proceedings of the Conference on Design, Automation and Test in Europe*, DATE '13, EDA Consortium, San Jose, CA, USA, 2013; 1589–1594. (Available from: http://dl.acm.org/citation.cfm?id=2485288.2485666, Accessed on 26 June 2014).

20. Bahirat S, Pasricha S. Meteor: hybrid photonic ring-mesh network-on-chip for multicore architectures. *ACM Transactions on Embedded Computing Systems* March 2014; **13**(3s):116:1–116:33. DOI: 10.1145/2567940.

21. Chan J, Hendry G, Biberman A, Bergman K. Architectural exploration of chip-scale photonic interconnection network designs using physical-layer analysis. *Journal of Lightwave Technology* 2010; **28**:1305–1315. DOI: 10.1109/JLT.2010.2044231.

22. O'Connor I, Van Thourhout D, Scandurra A. Wavelength division multiplexed photonic layer on CMOS. *Proceedings of the 2012 Interconnection Network Architecture: On-Chip, Multi-Chip Workshop*, INA-OCMC '12, ACM, New York, NY, USA, 2012; 33–36. DOI: 10.1145/2107763.2107772.

23. Biere M, Gheorghe L, Nicolescu G, O'Connor I, Wainer G. Towards the high-level design of optical networkson-chip. Formalization of opto-electrical interfaces. *International Conference on Electronics, Circuits and Systems*, Marrakech, 2007; 427–430. DOI: 10.1109/ICECS.2007.4511021.

24. O'Connor I, Brière M, Drouard E, Kazmierczak A, Tissafi-Drissi F, Navarro D, Mieyeville F, Dambre J, Stroobandt D, Fedeli J-M, Lisik Z, Gaffiot F. Towards reconfigurable optical networks-on-chip. *RECO SoC*, Montpellier, France, 2005; 121–128.

25. Le Beux S, Trajkovic J, O'Connor I, Nicolescu G, Bois G, Paulin P. Optical ring network-on-chip (ORNoC): architecture and design methodology. *Design, Automation Test in Europe Conference Exhibition (DATE)*, Grenoble, France, 2011; 1–6. DOI: 10.1109/DATE.2011.5763134.

26. Dally W, Towles B. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 2003.

27. Hansson A, Goossens K, Rdulescu A. Avoiding message-dependent deadlock in network-based systems on chip. *VLSI Design* 2007; **2007**:Article ID 95859, 10 pages.

28. TileraCorporation. Tile-Gx8016 specification. (Available from: http://www.tilera.com/sites/default/files/productbriefs/Tile-Gx-8016-SB011-03.pdf, Accessed on 26 June 2014).

29. Strano A, Ludovici D, Bertozzi D. A library of dual-clock FIFOs for cost-effective and flexible MPSoC design. *International Conference on Embedded Computer Systems (SAMOS)*, Samos, 2010; 20–27. DOI: 10.1109/ICSAMOS.2010.5642098.

30. Leu J, Stojanovic V. Injection-locked clock receiver for monolithic optical link in 45nm SOI. *2011 IEEE Asian Solid State Circuits Conference (A-SSCC)*, Jeju, 2011; 149–152. DOI: 10.1109/ASSCC.2011.6123624.

31. Stergiou S, Angiolini F, Carta S, Raffo L, Bertozzi D, De Micheli G. xpipes Lite: a synthesis oriented design library for networks on chips. *Design, Automation and Test in Europe*, Vol. 2, Munich, Germany, 2005; 1188–1193. doi:10.1109/DATE.2005.1.

32. Gilabert F, Gomez ME, Medardoni S, Bertozzi D. Improved utilization of NoC channel bandwidth by switch replication for cost-effective multi-processor systems-on-chip. *International Symposium on Networks-on-Chip (NOCS)*, Grenoble, France, 2010; 165–172. DOI: 10.1109/NOCS.2010.25.

33. Grotker T, Liao S, Martin G, Swan S. *System Design with SystemC*. Kluwer Academic Publishers: Hinghman, MA, USA, 2002.

34. Chan J, Hendry G, Biberman A, Bergman K, Carloni LP. PhoenixSim: a simulator for physical-layer analysis of chip-scale photonic interconnection networks. *Proceedings of the Conference on Design, Automation and Test in Europe*, DATE '10, European Design and Automation Association, 3001 Leuven, Belgium, Belgium, 2010; 691–696. (Available from: http://dl.acm.org/citation.cfm?id=1870926.1871093, Accessed on 26 June 2014).