

On the Area and Energy Scalability of Wireless Network-on-Chip: A Model-Based Benchmarked Design Space Exploration

Sergi Abadal, *Student Member, IEEE*, Mario Iannazzo, Mario Nemirovsky, Albert Cabellos-Aparicio, Heekwan Lee, and Eduard Alarcón, *Member, IEEE*

Abstract—Networks-on-chip (NoCs) are emerging as the way to interconnect the processing cores and the memory within a chip multiprocessor. As recent years have seen a significant increase in the number of cores per chip, it is crucial to guarantee the scalability of NoCs in order to avoid communication to become the next performance bottleneck in multicore processors. Among other alternatives, the concept of wireless network-on-chip (WNoC) has been proposed, wherein on-chip antennas would provide native broadcast capabilities leading to enhanced network performance. Since energy consumption and chip area are the two primary constraints, this work is aimed to explore the area and energy implications of scaling a WNoC in terms of: 1) the number of cores within the chip, and 2) the capacity of each link in the network. To this end, an integral design space exploration is performed, covering implementation aspects (area and energy), communication aspects (link capacity), and network-level considerations (number of cores and network architecture). The study is entirely based upon analytical models, which will allow to benchmark the WNoC scalability against a baseline NoC. Eventually, this investigation will provide qualitative and quantitative guidelines for the design of future transceivers for wireless on-chip communication.

Index Terms—Area, design space exploration, emerging interconnect technologies, multicore processors, Network-on-chip, on-chip antennas, power, wireless network-on-chip, wireless transceivers.

I. INTRODUCTION

IN THE ever-changing world of microprocessor design, multicore architectures are currently the dominant trend for both conventional and high-performance computing. These architectures consist of the interconnection of several

Manuscript received September 19, 2013; revised February 21, 2014; accepted June 12, 2014; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor E. Ekici. Date of publication July 02, 2014; date of current version October 13, 2015. This work was supported by Samsung under the Global Research Outreach (GRO) Program and INTEL under the Doctoral Student Honor Program, and in part by the Spanish Ministry of Science and Innovation under Project RUE CSD2009-00046 (Consolider-Ingenio 2010).

S. Abadal, M. Iannazzo, A. Cabellos-Aparicio, and E. Alarcón are with the NaNoNetworking Center in Catalonia (N3Cat), Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (e-mail: abadal@ac.upc.edu; acabello@ac.upc.edu; mario.enrique.iannazzo@estudiant.upc.edu; eduard.alarcon@upc.edu).

M. Nemirovsky is an ICREA Senior Research Professor with the Barcelona Supercomputing Center (BSC), 08034 Barcelona, Spain (e-mail: mario.nemirovsky@bsc.es).

H. Lee is with the Samsung Advanced Institute of Technology (SAIT), Suwon 440-600, Korea (e-mail: heekwan.lee@samsung.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNET.2014.2332271

TABLE I
BASELINE NOC PARAMETERS

Parameter	Value	Unit
System		
Chip Area	400	mm ²
CMOS Technology Node	32	nm
Operation Frequency	5	GHz
Supply Voltage	1	V
Topology	Mesh	-
Number of Links	224	-
Link (per hop)		
Capacity	240	Gbps
Energy	540	fJ/bit
Area	0.009	mm ²
Static Power	3.8	mW
Router (per hop)		
Energy	220	fJ/bit
Area	0.11	mm ²
Static Power	64	mW

independent processors or *cores*, as well as of a multilevel cache to improve the memory throughput. Communication among these elements is required for the implementation of diverse signaling schemes essential for the correct operation of a multiprocessor and largely impacts upon the computation performance. As the number of cores within these processing systems increases, their communication needs rise dramatically, to the point of turning communication into the major performance bottleneck of current multicore architectures.

With the aim of coping with the increasing on-chip communication requirements, a common practice has been to replace traditional bus architectures with networks of on-chip wires and routers [1]. This approach, also referred to as network-on-chip (NoC), can be understood as the application of networking principles and methods upon a set of electrical interconnects [2], [3]. However, NoCs enabled by these interconnects present fundamental limitations that point toward a reduced scalability beyond several tens of cores. As thoroughly discussed in [4] and references therein, the available energy for interconnects will soon be under the 100-fJ/bit barrier and will not be enough to cover the requirements of electrical wires (Table I). Also, their decreasing multicast performance is foreseen to be a significant issue in future many-core architectures (more details in Section II).

As a consequence of such limited scalability, considerable research efforts have been directed toward extending the original concept of NoC to other interconnect technologies. Diverse

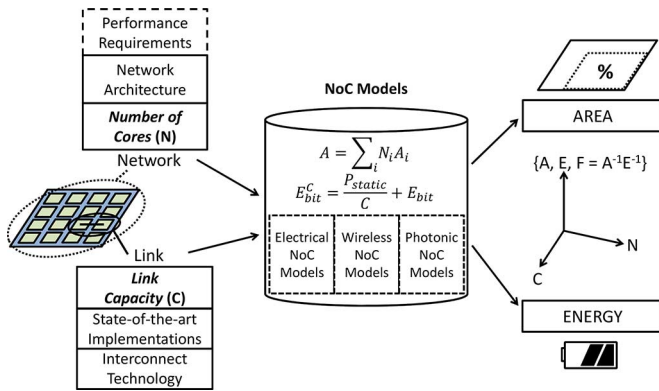


Fig. 1. Model-based approach employed in this work.

examples can be found in the literature, including the employment of vertical vias within stacked architectures [5], [6], of on-chip transmission lines for the transmission of modulated RF signals [7], or of nanophotonic interconnects enabling optical on-chip communication [4], [8]. Such emerging technologies may be used either to completely replace traditional NoCs [9], [10] or to follow a hybrid approach that leverages the capabilities of different types of interconnects [11], [12], targeting to ensure the scalability of on-chip networks beyond thousands of cores.

In line with the recent research trends, the possibility of implementing on-chip wireless communication by means of integrated antennas has been proposed [13]. The resulting wireless NoCs have garnered considerable interest from the community by virtue of, among others, their native broadcast and multicast capabilities [14]. Since the medium is shared among the cores, either multiplexing techniques or medium access control (MAC) protocols are required to achieve multiuser communication [15]. As a result, the concept of wireless NoC has been thus far analyzed in the form of specific network architectures and benchmarked employing traffic patterns from a set of standard applications. Alternatively, we aim to provide an interconnect-driven view of this research area by performing, as the main contribution, a circuit-oriented design space exploration of wireless NoC.

The employed methodology is summarized in Fig. 1. The investigation is entirely based on analytical models and compares how the area and energy consumption of wireless NoC scale as a function of the size and bandwidth requirements of the network for a given architecture. The results are then compared to that of a baseline electrical NoC (the interested reader will find data for a 64-core 48-bit instance in Table I) and of a selection of emerging alternatives. We expect that this design space exploration will allow for the identification of the scenarios wherein wireless NoC will potentially outperform other interconnect technologies. Furthermore, it will provide guidelines for the design not only of future transceivers and protocols for wireless on-chip communication, but also of network architectures that leverage different interconnect technologies.

The remainder of this paper is as follows. In Section II, we present a case study that will try to motivate the aim of this paper. In Section III, we review the state of the art of the wireless on-chip networking field. After introducing the analytical

framework and general assumptions in Section IV, the area and energy models for the different interconnect technologies are depicted in Sections V and VI, respectively. The results of the design space exploration are discussed in VII. Section VIII concludes the paper.

II. MOTIVATION

As the integration of a higher number of cores in the same chip is enabled, the general trend is to scale current multicore architectures and then to address the resulting increase in communication demands by means of enhanced on-chip networks. Provided that the architecture defines the characteristics of these communication demands, multicore processors have been designed taking into consideration the NoC capabilities. For instance, multicast has traditionally been a costly communication in chip environments and has been widely avoided. This tendency continues as in conventional NoCs, multicast messages are broken down into multiple unicast packets and generate large levels of contention. The work in [17] shows that conventional NoC latency and throughput suffer a degradation proportional to the multicast traffic intensity and reports significant reductions even for 1% of multicast traffic in a 4×4 mesh. It is expected that such impact will further increase in larger networks, as the number of destinations per message may potentially grow with the number of cores.

Even though on-chip multicast communications have been traditionally avoided, some architectural methods will need multicast in order to scale. For instance, cache coherency protocols normally avoid multicast by storing the state of each shared variable in a directory. This produces area and energy overheads proportional to the number of cores and may not be affordable in many-core systems. Instead, broadcast-based implementations do not store the state of each variable, but need to issue a broadcast for each coherence operation [18]. In this case, it is shown that improving the NoC multicast performance results in a significant reduction of both the interconnect power and execution time for a set of benchmark applications [19]. The introduction of an effective platform for the service of multicast messages would be highly beneficial in this context, but, more importantly, could open the door for new many-core architectures.

Aware of the importance of such an issue, explicit support for multicast communications within conventional electronic NoCs has been widely proposed for moderately sized multiprocessors [17], [19]–[22]. Still, the scalability of these solutions in terms of performance and cost has not been discussed in the literature. Fig. 2 plots the delay-throughput characteristic of a two-dimensional electrical mesh in the presence of broadcast traffic (as a particular case of multicast), showing a considerable performance deterioration as the number of cores is scaled. The results were obtained using the PhoenixSim framework [23], a cycle-accurate simulator that includes a wide variety of tools and methods for the evaluation of NoCs. In light of this, it remains unclear whether the aforementioned improvements will suffice to enable the use of traditional architectures in many-core processors.

Alternatively, the introduction of emergent interconnect technologies has opened a wide range of possibilities for

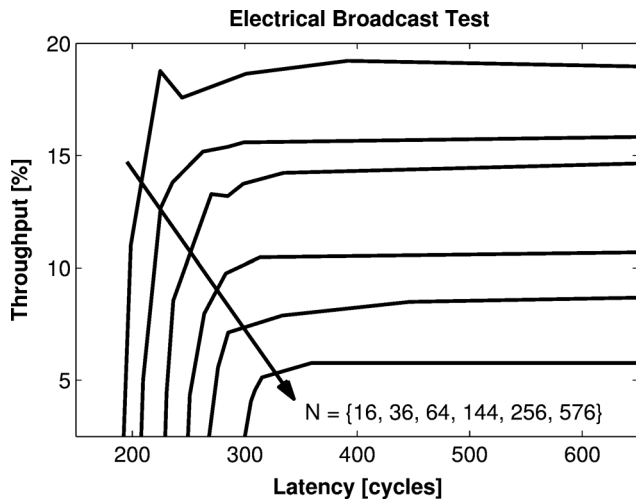


Fig. 2. Simulated delay-throughput characteristic of electrical meshed NoCs as a function of the number of nodes, considering pure broadcast traffic. Links are optimally repeated, with a link width of 64 at a clock frequency of 5 GHz; whereas routers implement unbalanced tree multicasting with a minimum routing latency per flit of 4 clock cycles. The throughput is in transmission, and it is expressed as a percentage of a link capacity.

cost-effective multicast on-chip communications. In 3-D NoCs, the reduced distance among cores both physically and in terms of number of hops inherently allows for an improved multicast performance. Also, the employment of one-to-all or all-to-all channels by means of global RF transmission lines and nanophotonic waveguides has been inspected [9], [11], [24]. In the case of wireless NoC, the native broadcast capabilities of such a technique show great promise toward implementing efficient architectural methods for many-core processors, as detailed and quantified in the following sections. It is important to note, though, that each of the aforementioned options presents its particular tradeoffs in terms of area, energy, and communication performance.

III. WIRELESS NETWORK-ON-CHIP

The constant improvement in the operating speeds of transistors has enabled the implementation of multigigahertz digital and RF circuits. In this context, the concept of on-chip antenna becomes a possibility since an antenna of a few millimeters in size is able to radiate at these frequencies [13]. Also, transceivers suited to the needs of the wireless chip communications have been developed: A wide variety of millimeter-wave implementations can be found in the literature covering many alternatives in terms of technology generation, modulation, or transceiver architecture [25]–[30]. For transmission ranges of up to a few centimeters, these provide high multigigabit data rates, and it is expected that these figures will keep increasing as technology evolves. A factor that aims to quantify the maturity of technology within this context is proposed in Section IV.

In light of the availability of both on-chip antennas and of appropriate transceivers, their employment to build wireless networks-on-chip (WNoCs) has been proposed. In this approach, information is radiated and propagates within the chip package following different propagation mechanisms [31]. Planar antennas can be used in spite of their typically low gain in the coplanar direction, in which case communication takes

place by means of space waves that are reflected upon the chip package. Alternatively, thanks to their potentially larger radiation efficiency in the chip plane, three-dimensional antennas could lead to achieving wireless communication through surface waves [32]. However, such antennas require complex microelectromechanical systems (MEMS) technologies for its fabrication.

As information may potentially reach any core regardless of its location, WNoC offers native broadcast capabilities, as well as the possibility of implementing flexible and one-hop communications. Multicast messages may actually be conveyed to the receivers in a few clock cycles, as opposed to in conventional NoCs. However, as the core density increases, the size of the millimeter-wave antennas may restrict the scope of WNoC to hybrid architectures wherein the wireless plane is employed to communicate clusters of cores. Although such a wireless backbone approach allows a reduction of the network diameter and has been shown to outperform conventional NoCs [33]–[36], its potential for broadcast-based communications is limited by the performance of the electrical edges of the network.

As further CMOS advancements push the operating frequencies toward the terahertz band [37], [38], the implementation of micrometer antennas becomes feasible. Moreover, novel planar antennas based on graphene promise to be able to radiate within this frequency band while being two orders of magnitude below, in size, of their metallic counterparts [39], [40]. In order to drive the antennas, transmitters and receivers for multigigabit communication at frequencies ranging from 0.1 to 0.4 THz have been already proposed [41]–[45]. Additionally, components reaching frequencies of 0.8 THz are under intense research [46]–[49], thus far leading to the apparition of transmitters and detectors for terahertz imaging and sensing [50]–[52].

Assuming a similar evolution than that of millimeter-wave transceivers, terahertz implementations could provide data rates of hundreds of gigabits per second at the chip scale. By virtue of this and the potentially reduced size of these terahertz systems, architectures implementing wireless communication at the core level can be envisaged [14] and will be considered throughout this work. In many-core processors, this approach will likely generate extremely high levels of contention when accessing the shared medium. Multiplexing techniques may not be suitable in this scenario due to the large number of channels required and the implications of this fact upon the complexity of the transceiver. Instead, a MAC protocol could arbitrate access to a single broadband channel and enable the development of broadcast-based WNoC architectures. In transmission, packets are serialized into bits and broadcast regardless of the number of intended destinations; whereas the receiver deserializes the incoming bits and then accepts or discards the packet after decoding its address. Buffer requirements for this process will be affordable as long as the packet rate after deserialization (C/L , where C is the link capacity in bits per second and L is the packet length in bits) is below the system clock frequency.

Since the bandwidth is shared among the nodes, the expected aggregated throughput of WNoC will be extremely low when compared to a wired NoC. In light of this, first uses of this broadcast-based platform may be restricted to

servicing a selection of control and signaling messages. These are latency-critical, often dense multicasts, and require lower bandwidths as they generally represent a small fraction of all the on-chip traffic. The approach is only feasible provided that this wireless control plane will complement a throughput-oriented wired NoC that will compensate for the low WNoC bandwidth by transporting the rest of the communication flows. Such a hybrid NoC could potentially reduce the latency of time-critical control messages while avoiding a deterioration of the wired NoC performance, potentially opening the door for new multiprocessor architectures.

IV. FRAMEWORK AND GENERAL CONSIDERATIONS

Given the stringent requirements of the on-chip communication scenario, in this work we explore the area and energy implications of scaling a WNoC system in terms of: 1) the number of effective receivers or network size, and 2) the capacity of a wireless link. The results of this implementation study are compared to that of representative examples of conventional and photonic NoC configurations, being aware of the main differences among them. For instance, since we assume that all nodes share a single broadband channel, the network capacity in WNoC is equal to a link capacity. Furthermore, the need of a MAC protocol implies that the effective network throughput in this case will be significantly lower than the network capacity. In contrast, the network capacity in wireline NoCs is the sum of the capacities of all the dedicated links that can simultaneously transmit data. Even though wireline NoCs will therefore yield much larger network throughput figures than WNoC for similar link capacities, the comparison will be performed at the link level. From a network throughput perspective, it remains unclear whether this large gap in nominal capacity will be compensated by the inherent difference in communication typology (i.e., local against global, unicast against broadcast). In future work, we will address this issue by investigating both the minimum wireless capacity requirements of different multiprocessor architectures, as well as the potential performance improvements of adding a wireless control plane.

A. Maturity Factor

On the one hand, the relation between the area/energy of a WNoC and its size in number of nodes can be easily described by means of simple models, as shown in Sections V-A and VI-A. On the other hand, it is not straightforward to assess how the area and energy of a wireless transceiver scale with its maximum achievable data rate due to the number of factors involved. Such a data rate, which is referred to as *link capacity* throughout the paper, depends on the transceiver bandwidth and the spectral efficiency of the selected modulation: In the transceiver design process, a given architecture is chosen in order to achieve the target bandwidth while implementing the selected modulation. On top of that, the frequency band wherein the communication will take place imposes additional requirements on some components of the transceiver, again depending on the architecture. Finally, the maturity of the employed technology should be taken into consideration, especially when reaching extremely high frequency bands.

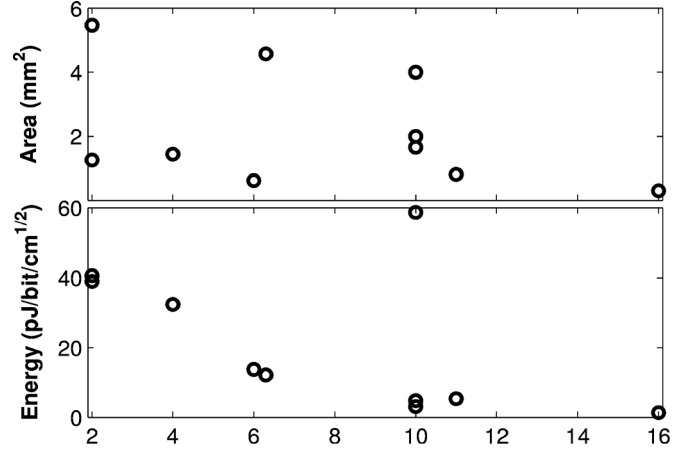


Fig. 3. Area and energy of state-of-the-art wireless transceivers [25]–[29], [41]–[44], [53] as a function of their data rate.

TABLE II
SUMMARY OF TRANSCEIVER SPECIFICATIONS

Technology	40 - 130 nm CMOS 130 - 250 nm SiGe BiCMOS
Transceiver Architecture	Impulse Radio (IR), Continuous Wave (CW)
Modulation	On-Off Keying (OOK), Amplitude Shift Keying (ASK), Phase Shift Keying (BPSK, QPSK), Frequency Shift Keying (FSK), Quadrature Amplitude Modulation (QAM)
Operation Frequency (f_c)	8 - 820 GHz
Transmission Range (d_{max})	1.4 - 210 cm
Data Rate (R)	2 - 18 Gbps

In order to extract a trend from the state of the art, a generally accepted approach is to represent the area or energy efficiency as a function of the data rate of different transceiver implementations, as done in Fig. 3. However, the tendency shown by such scatter plots is unclear and only covers a range between 2 and 18 Gb/s, rendering its extrapolation inadequate for the purpose of this work.

In light of the complexity of the analysis and of the heterogeneity of the state of the art in the field (see Table II), we will consider the following. Let us define the *Maturity Factor* as

$$M = S_E \cdot Q \quad [\text{b/s/Hz}] \quad (1)$$

where $S_E = \frac{R}{B}$ is the spectral efficiency of the employed modulation or data rate over the operation bandwidth, and $Q = \frac{B}{f_c}$ is the transceiver quality factor or its bandwidth over the operation frequency. Therefore

$$M = \frac{R}{f_c}. \quad (2)$$

In summary, the maturity factor tries to evaluate the efficiency of implementing a given modulation and bandwidth in order to yield a target data rate operating at a target frequency band. As technology matures, we expect highly optimized transceivers leading to increasing maturity factors, that is, higher data rates for similar area and energy values. For a transceiver at a given operation frequency and with certain area and energy efficiency figures, we will *a priori* assume a maturity value in order to extract a projected data rate. This way, a rough estimate of the

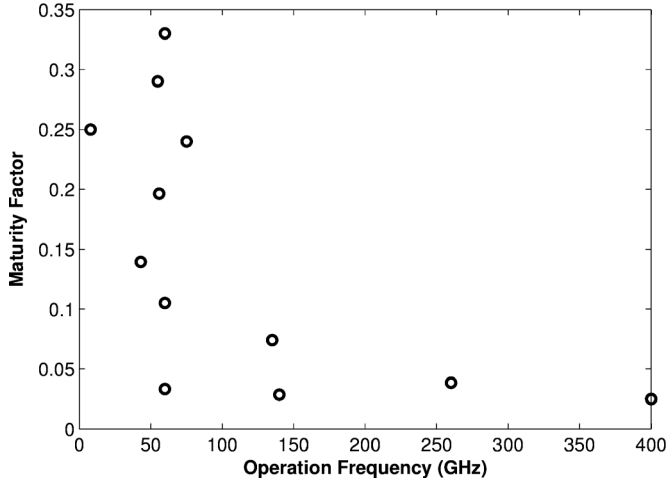


Fig. 4. Maturity factor as a function of the operation frequency of the transceiver for proposals [25]–[29], [41]–[44], [53].

area and energy efficiency of future wireless transceivers can be obtained.

Fig. 4 shows the maturity factor of several state-of-the-art transceivers [25]–[30], [41]–[44], [53] as a function of their frequency. We observe factors of up to 35% at the 60-GHz band followed by a decrease below 5% when reaching sub-terahertz frequencies. These values will be used throughout this work as reference guidelines indicating the maturity of wireless transceivers at a given frequency. We will consider that initial designs could achieve a maturity factor of up to 10%, while refined implementations may reach 20% and well-established transceivers could provide 30%. However, this rule of thumb may find exceptions as novel technologies are introduced. For instance, an impressive 100-Gb/s wireless transmission was recently accomplished using a photonic transmitter at a carrier frequency of 237.5 GHz, resulting in a maturity factor above 42% [54].

Eventually, the feasibility of the WNoC approach will be determined by the data rate requirements of the system. These could be met with current designs as transceivers with such performance have been already proposed [30]. Data rates up until 60 Gb/s may be achievable in the near future provided that either technologies at 100–300 GHz mature and reach a reasonable factor of 20%, or initial designs appear in the terahertz band. In order to reach speeds above 60 Gb/s, midterm efforts are required in order to raise the maturity of transceivers at in the terahertz band close to well-established levels.

V. AREA MODELS

While integration levels have been constantly increasing over the years, die sizes have practically stayed constant. Recently, 3-D stacking techniques have emerged, allowing the integration of devices in various vertically stacked layers. Still, the chip area is a finite resource that needs to be carefully managed: The area devoted to a given NoC will not be available for the core implementation, and vice versa.

In order to calculate the area overhead of an on-chip interconnection network, we will use the following general expression:

$$A = N_{\text{TX}}A_{\text{TX}} + N_{\text{RX}}A_{\text{RX}} + N_{\text{L}}A_{\text{L}} + N_{\text{R}}A_{\text{R}} \quad (3)$$

where N_i and A_i indicate the number of components of type i and its mean area occupancy, being the types divided in transmitters (TX), receivers (RX), links (L) and routers, switches, or other arbitration mechanisms (R). In the following, we will detail the analytical models that relate the number of components and their area to the number of nodes of the network and the targeted link capacity in wireless, electrical, and photonic NoCs. Note that the area figures will be independent of the traffic typology, as the considered NoCs are designed to support both unicast and multicast.

A. Wireless NoC Area Models

In the case of wireless on-chip communication, physical links are not needed in order to convey the information from the transmitter to the receiver. Moreover, switches or routers are not required if we assume one-hop communication. Therefore, the only components that occupy chip area are the antennas and the transceivers needed to modulate the data and to drive the signals to the antenna. We will assume one antenna and one transceiver per node, even though configurations with multiple antennas could be devised. Also, the analysis does not consider the area occupied by the logic required for the MAC protocol. For all this, (3) can be reduced to

$$A = N_{\text{TX}}A_{\text{TX}} + N_{\text{RX}}A_{\text{RX}} = N(A_{\text{ant}} + A_{\text{txrx}}) \quad (4)$$

where N is the number of nodes in the network, A_{ant} is the antenna area, and A_{txrx} is the transceiver area. The antenna and transceiver area will be mainly determined by the on-chip communication requirements. In order to achieve a given goal, the wireless plane must provide a certain effective network throughput that depends on the MAC protocol that arbitrates medium access and, more importantly, the data rate of each transceiver. Generally, higher data rates require higher bandwidths, which, in turn, require communication in higher frequency bands.

Such a tendency fortunately imposes a downscale on the antenna size. Due to the planar nature of a chip, we will consider the employment of patch antennas. The dimensions of such antennas are as follows: the width (W) is comparable to a wavelength λ , while the length L must be approximately $\lambda/2$. Therefore, for a given operation frequency f

$$A_{\text{ant}} \approx \frac{c_0^2}{2\epsilon_{\text{eff}}f^2} \quad (5)$$

where c_0 is the speed of light and ϵ_{eff} is the effective permittivity of the antenna. In order to fulfill the bandwidth requirements B at such a resonance frequency f_c , the antenna must yield a quality factor of $Q \approx \frac{B}{f_c}$. In order to simplify the analysis, we will consider that this quality factor will be achieved by means of techniques that do not largely affect the area occupied by the antenna, e.g., the quality factor in patch antennas is mainly determined by the distance between the patch and the ground plane.

In the case of the transceiver, the relation between the area and peak data rate is calculated as discussed in Section IV. A given maturity factor M is assumed so that the data rate requirement R can be achieved by operating at least at a frequency $f_c = \frac{R}{M}$. The area for such a transceiver can be extrapolated

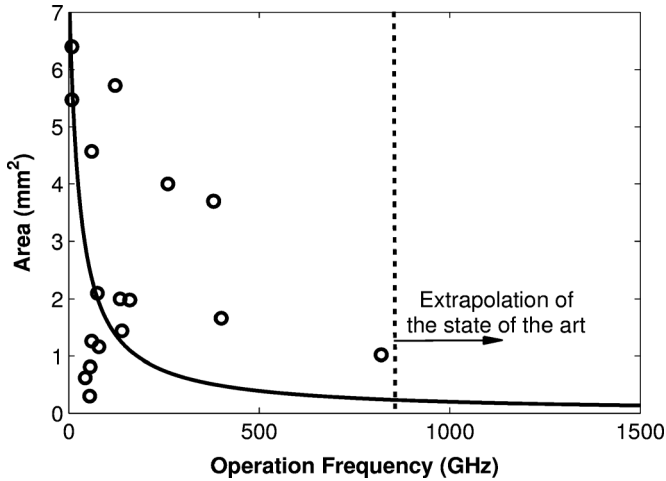


Fig. 5. Area of state-of-the-art wireless transceivers [25]–[29], [41]–[44], [50], [53], [55]–[59] as a function of their central frequency.

with data from the state of the art, which points toward a decrease in area when the frequency is upscaled (see Fig. 5). The reasons for the observed tendency may stem from the strong downsizing that is applied to the passive RF components of a transceiver when the operation frequency is increased. On the other hand, the scaling of active RF components remains unclear and should be inspected in future work with the aim of obtaining an accurate area scaling model for wireless on-chip transceivers. In this work, we will use a model obtained by applying fitting methods to the data represented in Fig. 5, which yielded the following equation:

$$A_{\text{txrx}} = \frac{206.1}{f_c + 27.22} \quad [\text{mm}^2] \quad (6)$$

wherein f_c is expressed in gigahertz. Rational fitting was chosen on the grounds that it delivers the most accurate result among the possible fittings and that it does not yield negative values for high frequencies. The weight of each data point is assigned in inverse proportion to the operation frequency, implying that implementations for well-established technologies at low frequencies are more representative than initial designs at the terahertz band. The resulting coefficient of determination, which evaluates the goodness of fit, is 0.68 (with 1 being an exact fit).

B. Electronic NoC Area Models

Two steps have been performed in order to calculate the area of an electronic NoC. First, the number of elements that constitute a given architecture can be easily derived by observing how its topology scales with the number of nodes. Once the topology is fixed, the area of each element can be calculated by means of simulation taking into consideration the topology and the target capacity. Our analysis has been performed by means of ORION, a widely recognized power-area simulator for on-chip interconnection networks [16].

Let us assume that each node has two line drivers, one for transmission (TX) and one for reception (RX). A typical line

driver accounts for an inverter and a D flip-flop, and ORION allows the user to calculate their area occupancy for a given technology node. In the case of the on-chip wires, ORION evaluates the number of repeaters needed for each link (L) based on its length (which is determined by the topology) and technology node. The area of each repeater is then calculated, added to the physical area of the wire, and multiplied by the number of parallel wires in a link, i.e., datapath width. Finally, the chip area of each router (R) is assessed by breaking the router down to the transistor level, calculating the number of transistors needed, and multiplying it by the size of a transistor for a given technology node. The final result will depend on parameters such as the number of ports, the size of the buffers, or the datapath width.

C. Photonic NoC Area Models

A photonic on-chip network essentially includes modulators, waveguides, switches, filters, and photodetectors. In the transmitting side (TX), we will assume that modulators are made of one active ring resonator, whereas receivers (RX) consist of a passive ring resonator-based filter and a photodetector. Switches can also be devised by employing ring resonators as building blocks [12]. Finally, we also consider that all ring resonators are of the same size. Given these assumptions, the area of a given architecture can be approximated as

$$A \approx N_{\text{ring}} A_{\text{ring}} + N_{\text{det}} A_{\text{det}} + \sum_i A_{\text{wg},i} \quad (7)$$

where N_{ring} and N_{det} are the number of ring resonators and photodetectors, respectively. $A_{\text{ring}} = W_{\text{ring}}^2$ is the area of each ring, or the square of its pitch, A_{det} is the photodetector area, and $A_{\text{wg},i}$ is the area of waveguide i . As in conventional electronic NoCs, the specific network architecture will determine the exact number of components as a function of the number of nodes and the target link capacity. The interested reader can find more details in [60], including a more detailed description of the architectures as well as the area and insertion loss values used in the analysis.

VI. ENERGY MODELS

The power consumed by any communication network can be classified in two main groups: static and dynamic. The static or zero-load power is the energy consumed independently of the traffic being served, whereas the dynamic power is a load-dependent component. Due to their distinct nature, static and dynamic powers are usually expressed in different units. Static power P_{static} is expressed in watts and gives insight about the energy that is consumed invariably through time to, for instance, maintain the circuitry active, whereas dynamic power E_{bit} is expressed in joules per bit and gives insight about the energy required to physically transmit one bit of data without errors from the transmitter to the intended receivers for a given interconnect technology.

As a rule of thumb, we will calculate the power consumed by a given on-chip network by using the following formula:

$$P = P_{\text{static}} + E_{\text{bit}} \cdot T \quad (8)$$

where T is the network throughput in bits per second. In a reverse process, we can also calculate the energy required to convey one bit of information from the transmitter to the intended receivers, operating at a given throughput

$$E_{\text{bit}}^T = \frac{P_{\text{static}}}{T} + E_{\text{bit}} \quad (9)$$

where the throughput T is ideally equivalent to the link capacity considering one transmission flow and no packet loss.

A. Wireless NoC Energy Models

Unlike in traditional wireless networks, the network nodes in a WNoC are integrated within the same platform and share the same power supply. Moreover, we will assume one shared channel and enough transmission power so that each wireless message is received by all the processing cores. In this context, the energy consumed in the transmission and reception of one bit is independent of whether the message is unicast or multicast and can be expressed as

$$E_{\text{bit},W}^T = E_{\text{bit}}^{\text{tx}} + N \cdot E_{\text{bit}}^{\text{rx}} \quad (10)$$

where $E_{\text{bit}}^{\text{tx}}$ and $E_{\text{bit}}^{\text{rx}}$ are the mean energy consumption in transmission and reception, respectively. Leakage currents of the $N - 1$ inactive transmitters, as well as the power consumed by the logic required to implement the MAC protocol are neglected. For a transceiver implementation with measured power in transmission P_{tx} and measured power in reception P_{rx} , both for a data rate R and a given transmission range, the equation above can be also expressed as

$$E_{\text{bit},W}^T = \frac{P_{\text{tx}} + N \cdot P_{\text{rx}}}{R} \quad (11)$$

It is important to remark that (11) expresses the energy per bit of a specific wireless transceiver yielding a data rate R . Since both metrics depend on several factors such as the selected modulation, the transceiver architecture, the transmission range, or the maturity of the employed technology, analytically obtaining a model that relates both the energy efficiency of wireless communication and its data rate is deemed highly challenging. Instead, as discussed in Section IV, we will assume a maturity factor M so that a target peak data rate R can be achieved by operating at least at a frequency $f_c = \frac{R}{M}$. We further consider that applying the MAC protocol, such a data rate will yield an effective network throughput that meets the communication requirements set by the multiprocessor.

This way, a generic trend can be extracted from the state of the art in wireless transceivers. Authors in [61] propose and discuss a figure of merit for wireless transceivers that encompasses both their energy efficiency E_{bit} and transmission range d_{max} by means of the following expression: $\Phi = \frac{E_{\text{bit}}}{\sqrt{d_{\text{max}}}}$. Fig. 6 shows how this figure of merit scales as a function of the frequency for implementations [25]–[29], [41]–[44], [53]. A similar fitting approach than that used in Section V-A provided the following relation:

$$\frac{E_{\text{bit}}}{\sqrt{d_{\text{max}}}} = \frac{1.41 \cdot 10^3}{f_c + 28.81} \quad [\text{pJ/bit/cm}^{1/2}] \quad (12)$$

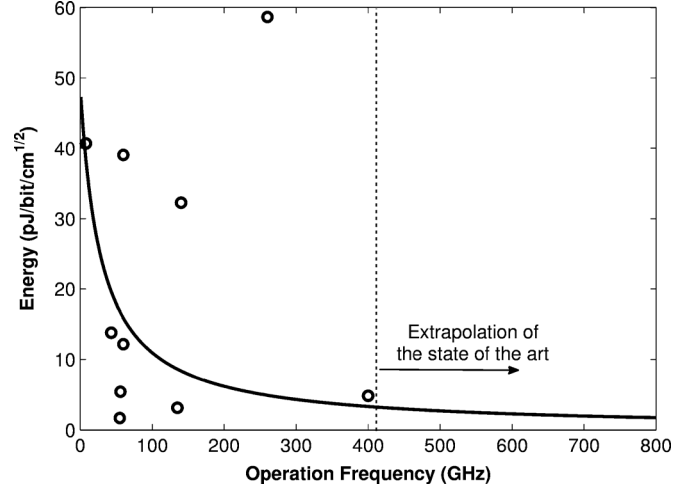


Fig. 6. Energy efficiency figure of merit of state-of-the-art wireless transceivers [25]–[29], [41]–[44], [53] as a function of their central frequency.

with a coefficient of determination of 0.65. In this case, $E_{\text{bit}} = E_{\text{bit}}^{\text{tx}} + E_{\text{bit}}^{\text{rx}}$ and f_c is expressed in gigahertz. Energy values can be extrapolated for frequencies beyond 400 GHz using the equation above.

The dependence on the transmission range is an important aspect to consider since, under the assumption that any transmitter should be able to reach any receiver, the nodes located at the chip edges will need a higher range than that of more centric nodes. This has two main implications: On the one hand, centric nodes need less transmission power to fulfill the sensitivity requirements at the chip edges. Therefore, the power amplifier can be tuned to consume less power. On the other hand, centric nodes receive transmissions with high power since the link budget is performed considering the worst case, that is, to reach the chip edges. In this case, the requirements for the low noise amplifiers are significantly relaxed. In our analysis, we will calculate which is the average energy per bit over all the on-chip transmitters following the aforementioned considerations with static power allocation.

Finally and unless noted, we will assume $E_{\text{bit}}^{\text{tx}} = E_{\text{bit}}^{\text{rx}} = E_{\text{bit}}/2$.

B. Electronic NoC Energy Models

Again, ORION is employed to determine both the static and dynamic power of an electronic NoC. In the former case, we will consider the power due to leakage currents in wires and routers. ORION breaks down these digital circuits to the transistor level and uses experimentally validated values for quiescent currents. In the latter case, ORION provides means to calculate the energy required to perform one hop within the network, which includes the energy required to: 1) transmit one bit of data through an on-chip wire of fixed length, and 2) read one bit of data from a router buffer, route it, and write it into the next router buffer.

Assuming a throughput T equal to the link capacity C , the energy per bit in an electronic NoC is

$$E_{\text{bit},E}^T = \frac{P_{\text{leakage}}}{C} + H \cdot E_{\text{b,hop}} \quad (13)$$

where P_{leakage} is the power due to leakage currents and $E_{\text{b,hop}}$ is the average energy required for one bit to perform one hop. H is the average distance between transmitter and receiver in terms of number of hops and solely depends on the network topology. For a 2-D mesh of N cores, $H_{\text{ucast}} = \frac{2\sqrt{N}}{3}$, whereas $H_{\text{bcast}} = N - 1$ considering a routing algorithm that minimizes the number of hops needed to deliver the message once to all the destinations.

C. Photonic NoC Energy Models

The power consumption in a photonic NoC is mainly driven by three components, namely, the laser power, the ring heating, and the energy required to perform the electrooptic (E/O) and optoelectric (O/E) conversions at the modulators and photodetectors, respectively.

Laser Power: Since integrating individual laser sources on a chip is currently unfeasible, it is generally accepted that light in a photonic NoC is supplied by an external multiwavelength source. This light is coupled, modulated, and then guided within the chip toward the intended receiver. In order to fulfill the sensitivity requirements at the receiver, the laser must transmit enough power to compensate for the losses incurred by the components found in the light path. Moreover and unless practical real-time laser management systems are made available [62], the laser power needs to be statically allocated to the worst-case scenario. In this context, a power budget analysis is performed following the expression:

$$P_{\text{laser}}(\text{dBW}) = S_{\text{RX}}(\text{dBW}) + \sum_i L_i(\text{dB}) \quad (14)$$

where P_{laser} is the electrical power consumed by the laser, S_{RX} is the receiver sensitivity, and L_i is the loss of component i , which includes both the laser and coupling efficiencies. The size and architecture of the network, as well as the target link capacity, will determine the number of components in the critical light path. The interested reader will find more details and a comparison of the laser power for different architectures in [60].

Ring Heating: Another source of static energy in photonic NoCs is the power needed to maintain ring resonators tuned to the desired frequency. Such components are extremely temperature-sensitive as small variations produce a shift in their resonant frequency. The power needed to keep ring resonators thermally tuned is

$$P_{\text{heat}} = N_{\text{ring}} \cdot P_{\text{ring}} \quad (15)$$

where N_{ring} is the number of ring modulators in the architecture, and P_{ring} is the power needed to maintain one ring finely tuned (see Table III). As commented in Section V-C, we will assume one ring per modulator and filter in all cases.

E/O and O/E Conversions: The dynamic power consumption in a photonic NoC is mainly due to the energy required to convert one electronic bit to light, and vice versa. In this case, we will consider fixed values demonstrated in the literature, which are shown in Table III. Similarly to wireless NoC, the energy required for the transmission and reception of one bit will depend on the number of k simultaneous receivers

$$E_{\text{bit}} = E_{\text{bit}}^{\text{tx}} + k \cdot E_{\text{bit}}^{\text{rx}} \quad (16)$$

TABLE III
PHOTONIC NOC PARAMETERS

Parameter	Value	Units	Ref.
Ring Losses	0.01-1	dB	[60]
Ring Area	64	μm^2	[60]
Ring Heating Power	26	$\mu\text{W}/\text{ring}$	[63]
Propagation Loss	0.5	dB/cm	[64]
Bending Loss	0.15	dB	[64]
Waveguide Pitch	2	μm	[64]
E/O Conversion	82	fJ/bit	[10]
O/E Conversion	50	fJ/bit	[10]
Photodetector Area	20	μm^2	[24]
Photodetector Sensitivity	-30	dBm	[60]

The parameter k is generally dependent on the photonic NoC architecture. Generally, point-to-point ($k = 1$) optical communication is implemented, and a separated broadcast channel ($k = N$) is employed for multireceiver transmissions [9]. Alternatively, a broadcast-based architecture would deliver any message to all the receivers, which would check the destination address and discard the message if necessary [24].

Assuming a throughput T equal to the link capacity C and using (14)–(16), the energy per bit in a photonic NoC is

$$E_{\text{bit},P}^T = \frac{P_{\text{laser}} + P_{\text{heat}}}{C} + E_{\text{bit}} \quad (17)$$

VII. BENCHMARKED DESIGN SPACE EXPLORATION

In this section, the results of the design space exploration are presented. We compare a small selection of architectures, namely:

- *EMesh*: which implements a conventional electrical mesh. We consider one 5-port router per core and bidirectional links connecting neighboring routers.
- *WMesh*: a WNoC-based architecture accounting for one communication unit (antenna and transceiver) per core. We assume that all cores share the same broadband channel and that a tailor-made MAC protocol arbitrates medium access.
- *OBUS*: a photonic bus arbitrated by means of an all-optical token-based scheme.
- *OXBar1*: an optical crossbar, wherein each core is tuned to a unique wavelength in transmission and broadcasts its messages to the rest of the cores. For more details on this architecture, see [60].
- *OXBar2*: another optical crossbar, wherein each core is associated to a unique data waveguide. Through this dedicated channel, a given core is able to receive data modulated by any of the other cores. For more details on this architecture, see [60].

Tables I and III show a summary of the technological parameters used in the study. The variable *number of cores* is swept between 4 and 1024, whereas the *link capacity* is scaled up to 250 Gb/s. Note that when the number of cores is increased, the network capacity remains constant in WMesh and grows proportionally to that increase in the rest of the alternatives.

A. Area

Fig. 7 shows the area-network size plane of the design space, corresponding to fixing the link capacity to a value of 80 Gb/s.

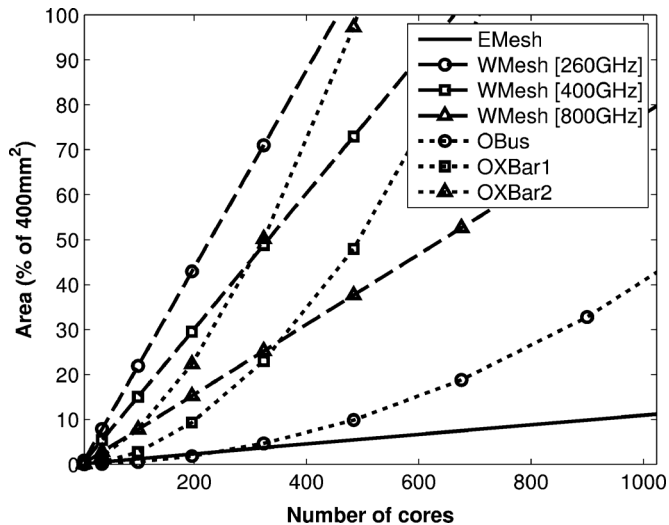


Fig. 7. Area scaling as a function of the number of cores for different interconnect technologies and architectures. The link capacity is set to 80 Gb/s.

The electrical and wireless options show a linear behavior, while photonic NoCs grow with the square of the number of cores due to the quadratic scaling in number of components [60].

In the WNoC case, three different operation frequencies have been chosen, namely 260, 400, and 800 GHz. Taking into account the targeted link capacity, such frequencies lead to maturity factors not exceeding 30%, in consonance with the values shown in the state of the art (see Fig. 4). From an area overhead perspective, high frequencies are beneficial since they entail lower area both for the antenna and the transceiver, according to the tendency pointed out in Section V-A. Nevertheless, the area occupation in most cases is higher than that of the electrical and photonic alternatives. Considerable transceiver area optimization is needed in order to enable size compatibility with massive multicore architectures: Reducing the area of an 800-GHz transceiver to 0.1 mm^2 would yield an overhead of 27% in a 1000-core processor. By employing graphene-based nano-antennas [39], [40], such an area overhead would be further reduced to a 25%.

Fig. 8 shows the area-capacity plane of the design space, corresponding to fixing the network size to a value of 256 nodes. It can be observed that both electronic and photonic NoCs show a linear growth of area with respect to the link capacity since higher bandwidth requirements are generally fulfilled by means of additional wires and circuitry.

In the wireless case, we consider different preset maturity factors and then scale the operation frequency in accordance with the link capacity objectives. Once the operation frequency is chosen, the area is calculated using the model presented in Section V-A. Such an approach explains the negative slope of the WNoC area plots: Higher bandwidth requirements imply an increase in the operation frequency, which in turn entails a reduction in the size of both the antenna and the transceiver. Due to this, it is expected that WNoC will be able to compete with the electrical and photonic alternatives at high link capacities due to the extremely high operation frequencies required for transmission. It is important to note, though, that such possibility is limited by the state of technology as it determines the maximum

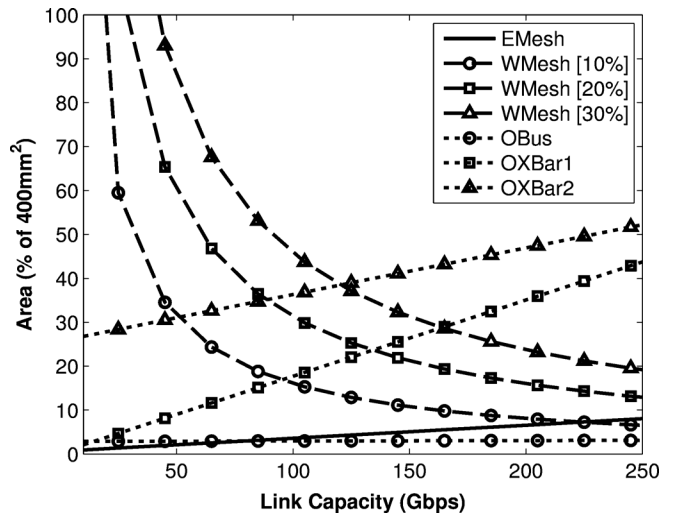


Fig. 8. Area scaling as a function of the link capacity for different interconnect technologies and architectures. The number of nodes is set to 256.

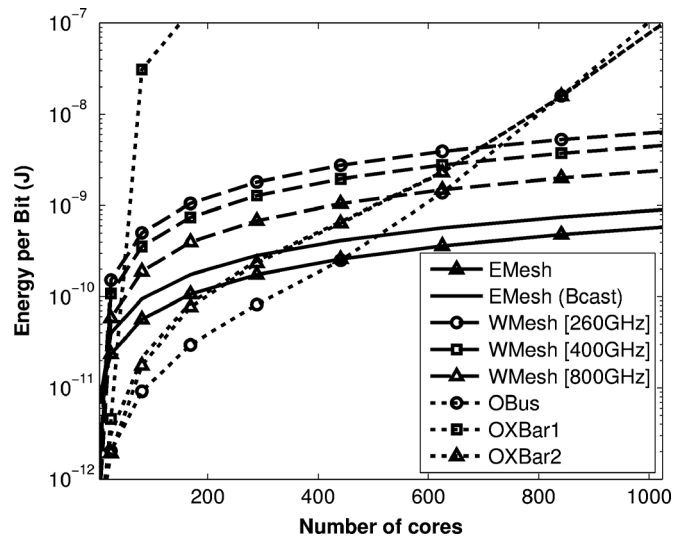


Fig. 9. Energy per bit scaling as a function of the network size for different interconnect technologies and architectures. The link capacity is set to 80 Gb/s.

frequency at which circuits can operate. This may also imply that higher maturity factors may need to be sought in order to increase the link capacity of a WNoC over a given value.

B. Energy

Fig. 9 shows the energy-network size plane of the design space for a fixed link capacity of 80 Gb/s. There are several aspects to be noted.

- In a conventional NoC, there is a considerable gap between the energy per bit in a unicast transmission and in a broadcast transmission. In both cases, conventional designs outperform wireless and photonic NoCs.
- WNoCs follow a similar trend than conventional NoCs, being the options working at higher frequencies closer to achieve an energy efficiency comparable to that of conventional NoCs, in accordance to the extrapolation proposed in Fig. 6.

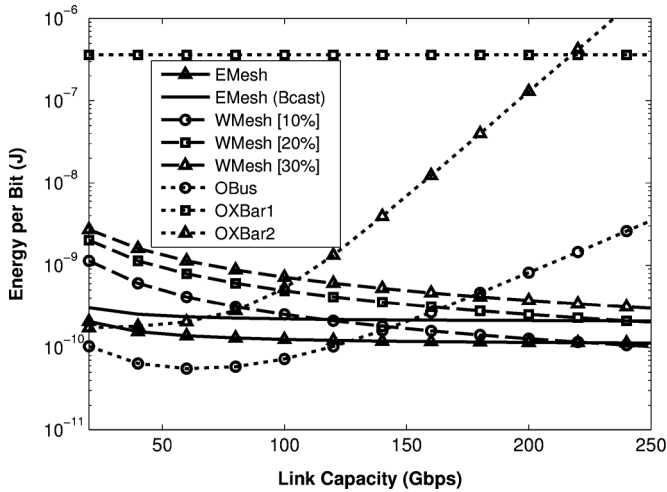


Fig. 10. Energy per bit scaling as a function of the link capacity for different interconnect technologies and architectures. The network size is fixed to 256 nodes.

- In a photonic NoC, the energy figures can be considered independent on whether the transmission is unicast or multicast by virtue of the extremely low energy needed for the O/E conversions. However and despite such potential for low energy transmissions, the photonic NoC configurations scale poorly due to their high laser power requirements, specially at high core counts [60].

Fig. 10 shows the energy-capacity plane of the design space. On the one hand, it is observed that conventional NoCs yield an energy efficiency that is almost invariant with respect to the link capacity. On the other hand, the energy efficiency of WNoCs not only improves with the link capacity, but also outperforms conventional NoCs at some point, provided that the trend observed in the state of the art continues in future transceivers (see Fig. 6). Finally, our results confirm that the different photonic NoC options do not scale well, as their efficiency substantially deteriorates for high link capacities. This is mainly due to the steep increase in number of components leading to an extremely high accumulated loss and, eventually, to unaffordable laser power requirements.

C. Area-Energy Figure of Merit

As seen in the previous sections, a given on-chip network may scale remarkably well in terms of area and perform poorly in terms of energy, or vice versa. In order to evaluate both the area and energy scalability of each solution, we propose the following figure of merit:

$$\text{FoM} = \frac{1}{A \cdot E_{\text{bit}}^T} \quad [\text{bits}/\text{J}/\text{mm}^2]. \quad (18)$$

Such a performance metric can be understood as the average number of bits that can be effectively transmitted for: 1) each consumed joule of energy, and 2) square millimeter of chip real estate. It is therefore an indicator of the joint energy and area efficiency of a given on-chip network. A large value of this figure of merit is desired.

On the one hand, Fig. 11 shows how the figure of merit scales as a function of the network size in number of cores.

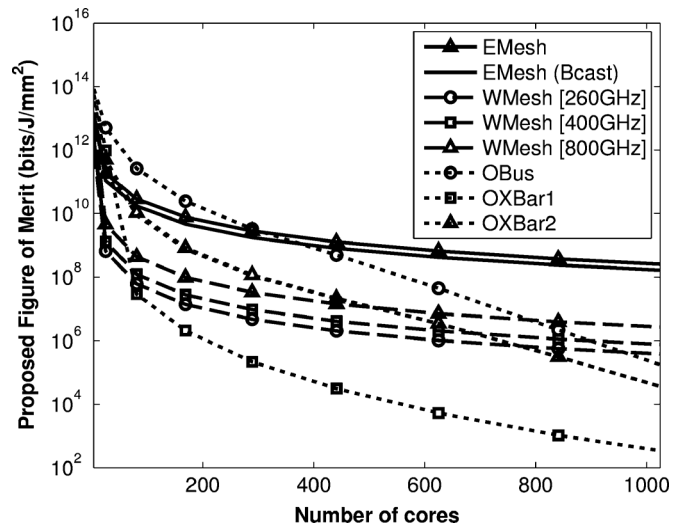


Fig. 11. Scaling of the proposed figure of merit (higher is better) as a function of the number of cores for different interconnect technologies and architectures. The link capacity is set to 80 Gb/s.

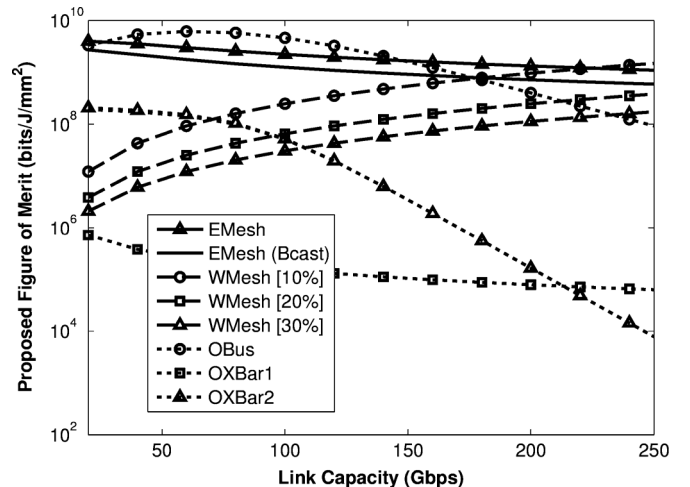


Fig. 12. Scaling of the proposed figure of merit (higher is better) as a function of the link capacity for different interconnect technologies and architectures. The network size is fixed to 256 nodes.

Again, electrical and wireless NoCs show a similar trend, while a rapid decrease of the figure of merit is observed in photonic NoCs. Overall, conventional NoC yields the best performance. On the other hand, Fig. 12 shows how the figure of merit scales as a function of the link capacity, in a network consisting of 256 cores. In this case, the analysis is slightly more complex. While it is clear that the optical crossbars scale poorly with the link capacity, the rest of the options yield similar performance. According to our analysis, the optical bus shows the best performance for low link capacities, whereas wireless NoCs could yield an improved efficiency for high link capacities if the scaling trends observed in the state of the art continue.

D. Discussion and Open Challenges

Results revealed in previous sections indicate that, in absolute terms, the baseline NoC performs remarkably better than its potential alternatives. However, it is important to note that the technologies employed for electrical on-chip wires and routers

TABLE IV
DOMINANT AREA AND ENERGY SCALABILITY TRENDS

Architecture	Area	Energy
EMesh	$O(NC)$	$O(N)$
WMesh	$O(N/C)$	$O(N/C)$
OBus	$O(N^2)$	$O(\alpha^N \beta^C)$
OXBar1	$O(N^2 C)$	$O(\gamma^N)$
OXBar2	$O(N^2 C)$	$O(\delta^N \epsilon^C)$

($\alpha, \beta, \gamma, \delta, \epsilon$ are constants)

are thus far much more optimized than nanophotonic or wireless chip-area technologies, which are still in their infancy and may substantially improve in the following years. In the specific case of WNoC, the efficiency of the communication could be improved at different levels of design.

- *At the transceiver level:* Unlike in traditional wireless systems, all the on-chip wireless transceivers share the same power supply, and therefore the energy per bit metric encompasses the energy consumed by transmitter and all the receivers within the transmission range—see (11). Thus far, we assumed $E_{\text{bit}}^{\text{tx}} = E_{\text{bit}}^{\text{rx}}$ in order to simplify the analysis. However, the ratio between such figures could be chosen in the transceiver design process. To this end, a model accounting for the tradeoffs between transceiver energy consumption, radiated power, and received power would enable the optimization of the energy efficiency.
- *At the circuit level:* In this work, we considered a heterogeneous set of transceivers implementing different modulations and aiming at different communication scenarios, which are not necessarily oriented to low area and low power. Novel and optimized circuit topologies could allow for a substantial improvement of the area and energy efficiencies in wireless chip communication.
- *At the technology level:* The performance of a given wireless transceiver is undeniably limited by the underlying technology. Generally, technological advancements lead to higher operation frequency, lower area, and potential for lower energy consumption. The trend set by current state-of-the-art transceivers will continue provided that the employed technologies evolve accordingly. However, the advent of a new technology bringing disruptive improvements, such as the graphene technology [65]–[68], may allow to go beyond the predicted performance.

For the sake of fairness, the comparison must account for the structural tendencies rather than for the absolute area and energy values. Table IV summarizes the trends obtained through the application of fitting methods to the area and energy plots. We can observe that *WMesh* offers a good area and energy scalability with respect the number of nodes and an excellent scalability with respect to the link capacity. From this, we can infer that the concept of WNoC is better suited to the case of high data rate requirements leading to a very high radiation frequency. Conversely, in small networks working at lower speeds, electrical and photonic interconnects are expected to offer improved area and energy efficiencies. It is important to remark that these results do not include the area and energy required by the circuits required to implement the MAC protocol. However, SD-MAC [69] represents the only MAC protocol

for WNoC implemented to date and consumes very low area and bit energy ($\sim 0.01 \text{ mm}^2$ and $\sim 70 \text{ pJ/packet}$ in $0.18 \mu\text{m}$ CMOS), suggesting that the impact of including the MAC protocol within the analysis is negligible in light of the results shown in this paper. This aspect will be further addressed in future work.

VIII. CONCLUSION

The area and energy scalability of WNoC in terms of: 1) the number of cores within a multiprocessor, and 2) the capacity of each link in the network has been analyzed and compared to those of conventional and optical NoCs. In support of this study, we modeled the area and energy efficiencies of high-speed transceivers by means of extrapolation with respect to the state of the art and proposed a figure of merit encompassing both metrics. Although it is shown that the baseline NoC outperforms the wireless and optical alternatives in absolute terms, such a comparison is implementation-dependent and does not reveal the fundamental scalability trends. A further analysis of the results shows that WNoC offers good scalability both in area and energy, especially with respect to the link capacity. This outcome confirms the feasibility of WNoC, which may take a central role in future multiprocessors given the rising importance of multicast communication in such scenario.

REFERENCES

- [1] W. Dally and B. Towles, "Route packets, not wires: On-chip interconnection networks," in *Proc. 38th IEEE Design Autom. Conf.*, 2001, pp. 684–689.
- [2] L. Benini and G. De Micheli, "Networks on chips: A new SoC paradigm," *Computer*, vol. 35, no. 1, pp. 70–78, 2002.
- [3] T. Bjerregaard and S. Mahadevan, "A survey of research and practices of network-on-chip," *Comput. Surveys*, vol. 38, no. 1, pp. 1–51, Jun. 2006.
- [4] D. A. B. Miller, "Device requirements for optical interconnects to silicon chips," *Proc. IEEE*, vol. 97, no. 7, pp. 1166–1185, Jul. 2009.
- [5] A. W. Topol *et al.*, "Three-dimensional integrated circuits," *IBM J. Res. Dev.*, vol. 50, no. 4, pp. 491–506, Jul. 2006.
- [6] B. S. Feero and P. P. Pande, "Networks-on-chip in a three-dimensional environment: A performance evaluation," *IEEE Trans. Comput.*, vol. 58, no. 1, pp. 32–45, Jan. 2009.
- [7] E. Socher and M.-C. F. Chang, "Can RF help CMOS processors?," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 104–111, Aug. 2007.
- [8] R. G. Beausoleil, P. J. Kuekes, G. S. Snider, S.-Y. Wang, and R. S. Williams, "Nanoelectronic and nanophotonic interconnect," *Proc. IEEE*, vol. 96, no. 2, pp. 230–247, Feb. 2008.
- [9] D. Vantrease *et al.*, "Corona: System implications of emerging nanophotonic technology," *Comput. Archit. News*, vol. 36, no. 3, pp. 153–164, 2008.
- [10] N. Kirman and J. F. Martínez, "A power-efficient all-optical on-chip interconnect using wavelength-based oblivious routing," *ACM SIGPLAN Notices*, vol. 45, no. 3, pp. 15–28, 2010.
- [11] M.-C. F. Chang, E. Socher, S.-W. Tam, J. Cong, and G. Reinman, "RF interconnects for communications on-chip," in *Proc. ISPD*, 2008, p. 78.
- [12] A. Shacham, K. Bergman, and L. P. Carloni, "Photonic networks-on-chip for future generations of chip multiprocessors," *IEEE Trans. Comput.*, vol. 57, no. 9, pp. 1246–1260, Sep. 2008.
- [13] K. K. O *et al.*, "On-chip antennas in silicon ICs and their application," *IEEE Trans. Electron Devices*, vol. 52, no. 7, pp. 1312–1323, 2005.
- [14] S. Abadal, E. Alarcón, M. C. Lemme, M. Nemirovsky, and A. Cabellós-Aparicio, "Graphene-enabled wireless communication for massive multicore architectures," *IEEE Commun. Mag.*, vol. 51, no. 11, pp. 137–143, Nov. 2013.
- [15] S. Deb, A. Ganguly, P. P. Pande, B. Belzer, and D. Heo, "Wireless NoC as interconnection backbone for multicore chips: Promises and challenges," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 2, pp. 228–239, Jun. 2012.

- [16] A. Kahng, B. Li, L. Peh, and K. Samadi, "Orion 2.0: A fast and accurate noc power and area model for early-stage design space exploration," in *Proc. Design, Autom. Test Eur.*, 2009, pp. 423–428.
- [17] N. E. Jerger, L.-S. Peh, and M. Lipasti, "Virtual circuit tree multicasting: A case for on-chip hardware multicast support," in *Proc. Int. Symp. Comput. Archit.*, Jun. 2008, pp. 229–240.
- [18] M. Lodde, J. Flich, and M. E. Acacio, "Heterogeneous NoC design for efficient broadcast-based coherence protocol support," in *Proc. IEEE/ACM 6th Int. Symp. Networks-on-Chip*, May 2012, pp. 59–66.
- [19] T. Krishna, L. Peh, B. Beckmann, and S. K. Reinhardt, "Towards the ideal on-chip fabric for 1-to-many and many-to-1 communication," in *Proc. 44th Annu. IEEE/ACM MICRO*, 2011, vol. 2, pp. 71–82.
- [20] S. Rodrigo, J. Flich, J. Duato, and M. Hummel, "Efficient unicast and multicast support for CMPs," in *Proc. 41st IEEE/ACM Int. Symp. Microarchit.*, Nov. 2008, pp. 364–375.
- [21] F. A. Samman, T. Hollstein, and M. Glesner, "Multicast parallel pipeline router architecture for network-on-chip," in *Proc. DATE*, 2008, pp. 1396–1401.
- [22] R. Manevich, I. Walter, I. Cidon, and A. Kolodny, "Best of both worlds: A bus enhanced NoC (BENoC)," in *Proc. 3rd ACM/IEEE Int. Symp. Networks-on-Chip*, 2009, pp. 173–182.
- [23] J. Chan, G. Hendry, A. Biberman, K. Bergman, and L. P. Carloni, "PhoenixSim: A simulator for physical-layer analysis of chip-scale photonic interconnection networks," in *Proc. DATE*, 2010, pp. 691–696.
- [24] G. Kurian *et al.*, "ATAC: A 1000-core cache-coherent processor with on-chip optical network," in *Proc. PACT*, 2010, pp. 477–488.
- [25] W.-H. Chen *et al.*, "A 6-Gb/s wireless inter-chip data link using 43-GHz transceivers and bond-wire antennas," *IEEE J. Solid-State Circuits*, vol. 44, no. 10, pp. 2711–2721, Oct. 2009.
- [26] H. Wang, M.-H. Hung, Y.-C. Yeh, and J. Lee, "A 60-GHz FSK transceiver with automatically-calibrated demodulator in 90-nm CMOS," in *Proc. IEEE VLSIC*, 2010, pp. 95–96.
- [27] K. Kawasaki *et al.*, "A millimeter-wave intra-connect solution," *IEEE J. Solid-State Circuits*, vol. 45, no. 12, pp. 2655–2666, Dec. 2010.
- [28] X. Yu *et al.*, "A wideband body-enabled millimeter-wave transceiver for wireless network-on-chip," in *Proc. 54th IEEE MWSCAS*, Aug. 2011, pp. 1–4.
- [29] K. Okada, K. Kondou, M. Miyahara, M. Shinagawa, and H. Asada, "Full four-channel 6.3-Gb/s 60-GHz CMOS TRX with low-power analog and digital baseband circuitry," *IEEE J. Solid-State Circuits*, vol. 48, no. 1, pp. 46–65, Jan. 2013.
- [30] S. Kawai *et al.*, "Direct-conversion transceiver in 65-nm CMOS," in *Proc. IEEE RFIC*, 2013, no. 1, pp. 137–140.
- [31] Y. P. Zhang, Z. M. Chen, and M. Sun, "Propagation mechanisms of radio waves over intra-chip channels with integrated antennas: Frequency-domain measurements and time-domain analysis," *IEEE Trans. Antennas Propag.*, vol. 55, no. 10, pp. 2900–2906, Oct. 2007.
- [32] P. Nenzi, F. Tripaldi, V. Varlamava, F. Palma, and M. Balucani, "On-chip THz 3D antennas," in *Proc. 62nd IEEE ECTC*, 2012, pp. 102–108.
- [33] A. Ganguly *et al.*, "Scalable hybrid wireless network-on-chip architectures for multi-core systems," *IEEE Trans. Comput.*, vol. 60, no. 10, pp. 1485–1502, Oct. 2010.
- [34] C. Wang, W.-H. Hu, and N. Bagherzadeh, "A wireless network-on-chip design for multicore platforms," in *Proc. 19th Int. Euromicro Conf. Parallel, Distrib. Netw.-Based Process.*, Feb. 2011, pp. 409–416.
- [35] S.-B. Lee *et al.*, "A scalable micro wireless interconnect structure for CMPs," in *Proc. ACM MobiCom*, 2009, p. 217.
- [36] D. Matolak *et al.*, "Wireless networks-on-chips: Architecture, wireless channel, and devices," *IEEE Wireless Commun.*, vol. 19, no. 5, pp. 58–65, Oct. 2012.
- [37] S. Sankaran *et al.*, "Towards terahertz operation of CMOS," in *Proc. Int. Solid-State Circuits Conf.*, 2009, pp. 202–204.
- [38] E. Seok, D. Shim, and C. Mao, "Progress and challenges towards terahertz CMOS integrated circuits," *IEEE J. Solid-State Circuits*, vol. 45, no. 8, pp. 1554–1564, Aug. 2010.
- [39] I. Llatser *et al.*, "Graphene-based nano-patch antenna for terahertz radiation," *Photon. Nanostruct., Fund. Appl.*, vol. 10, no. 4, pp. 353–358, 2012.
- [40] J. M. Jornet and I. F. Akyildiz, "Graphene-based plasmonic nano-antenna for terahertz band communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 685–694, Dec. 2013.
- [41] N. Ono *et al.*, "135 GHz 98 mW 10 Gbps ASK transmitter and receiver chipset in 40 nm CMOS," in *Proc. IEEE VLSIC*, 2012, pp. 50–51.
- [42] E. Laskin, P. Chevalier, B. Sautreuil, and S. Voinigescu, "A 140-GHz double-sideband transceiver with amplitude and frequency modulation operating over a few meters," in *Proc. IEEE BCTM*, 2009, pp. 178–181.
- [43] S. Hu *et al.*, "A SiGe BiCMOS TX/RX chipset with on-chip SIW antennas for Terahertz applications," *IEEE J. Solid-State Circuits*, vol. 47, no. 11, pp. 2654–2664, Nov. 2012.
- [44] J.-D. Park, S. Kang, S. Thyagarajan, E. Alon, and A. Niknejad, "A 260 GHz fully integrated CMOS transceiver for wireless chip-to-chip communication," in *Proc. IEEE VLSIC*, 2012, pp. 48–49.
- [45] B. Khamaisi, S. Jameson, E. Socher, and S. Member, "A 210 227 GHz transmitter with integrated on-chip antenna in 90 nm CMOS technology," *IEEE Trans. Terahertz Sci. Technol.*, vol. 3, no. 2, pp. 141–150, Mar. 2013.
- [46] A. Lisauskas, S. Boppel, M. Mundt, V. Krozer, and H. G. Roskos, "Subharmonic mixing with field-effect transistors: Theory and experiment at 639 GHz high above fT," *IEEE Sensors J.*, vol. 13, no. 1, pp. 124–132, Jan. 2013.
- [47] R. Han and E. Afshari, "A high-power broadband passive terahertz frequency doubler in CMOS," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 3, pp. 1150–1160, Mar. 2013.
- [48] F. Golcuk, O. D. Gurbuz, and G. M. Rebeiz, "A 0.390.44 THz 2 × 4 amplifier-quadrupler array with peak EIRP of 34 dBm," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 12, pp. 4483–4491, Dec. 2013.
- [49] H. Rucker, B. Heinemann, and A. Fox, "Half-terahertz SiGe BiCMOS technology," in *Proc. 12th IEEE SiRF*, 2012, pp. 133–136.
- [50] E. Öjefors, J. Grzyb, B. Heinemann, B. Tillack, and U. R. Pfeiffer, "A 820 GHz SiGe chipset for terahertz active imaging applications," in *Proc. IEEE ISSCC*, 2011, pp. 224–225.
- [51] R. A. Hadi, J. Grzyb, B. Heinemann, and U. R. Pfeiffer, "A terahertz detector array in a SiGe HBT technology," *IEEE J. Solid-State Circuits*, vol. 48, no. 9, pp. 2002–2010, Sep. 2013.
- [52] U. R. Pfeiffer, J. Grzyb, H. Sherry, A. Cathelin, and A. Kaiser, "Toward low-NEP room-temperature THz MOSFET direct detectors in CMOS technology," in *Proc. 38th IRMMW-THz*, Sep. 2013, pp. 1–2.
- [53] T. Abe, Y. Yuan, H. Ishikuro, and T. Kuroda, "A 2 Gb/s 150 mW UWB direct-conversion coherent transceiver with IQ-switching carrier recovery scheme," in *Proc. IEEE ISSCC*, 2012, pp. 442–444.
- [54] S. Koenig *et al.*, "Wireless sub-THz communication system with high data rate," *Nature Photon.*, vol. 7, no. 12, pp. 977–981, Oct. 2013.
- [55] L. Zhou *et al.*, "A 2 Gbps RF-correlation-based impuleradio UWB transceiver front-end in 130 nm CMOS," in *Proc. IEEE RFIC*, 2009, pp. 65–68.
- [56] I. Sarkas *et al.*, "An 18-Gb/s, direct QPSK modulation SiGe BiCMOS transceiver for last mile links in the 7080 GHz band," *IEEE J. Solid-State Circuits*, vol. 45, no. 10, pp. 1968–1980, Oct. 2010.
- [57] C. Wagner, H.-P. Forstner, G. Haider, A. Stelzer, and H. Jager, "A 79-GHz radar transceiver with switchable TX and LO feedthrough in SiGe," in *Proc. Bipolar/BiCMOS Circuits Technol. Meeting*, 2008, pp. 105–108.
- [58] I. Sarkas, J. Hasch, A. Balteanu, and S. Voinigescu, "A fundamental frequency 120-GHz SiGe BiCMOS distance sensor with integrated antenna," *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 3, pp. 795–812, Mar. 2012.
- [59] Y. Zhao, E. Ojefors, K. Aufinger, T. Meister, and U. Pfeiffer, "A 160-GHz subharmonic transmitter and receiver chipset in a SiGe HBT technology," *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 10, pp. 3286–3299, Oct. 2012.
- [60] S. Abadal *et al.*, "Area and laser power scalability analysis in photonic networks-on-chip," in *Proc. 17th ONDM*, 2013, pp. 131–136.
- [61] J. Gorisse, D. Morche, and J. Jantunen, "Wireless transceivers for gigabit-per-second communications," in *Proc. 10th IEEE Int. NEWCAS Conf.*, Jun. 2012, pp. 545–548.
- [62] C. Chen and A. Joshi, "Runtime management of laser power in silicon-photonic multibus NoC architecture," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, p. 3700713, Mar.–Apr. 2013.
- [63] J. Ahn *et al.*, "Devices and architectures for photonic chip-scale integration," *Appl. Phys. A*, vol. 95, no. 4, pp. 989–997, Feb. 2009.
- [64] J. Cardenas, C. Poitras, and J. Robinson, "Low loss etchless silicon photonic waveguides," *Opt. Exp.*, vol. 17, no. 6, pp. 4752–4757, 2009.
- [65] A. K. Geim and K. S. Novoselov, "The rise of graphene," *Nature Mater.*, vol. 6, no. 3, pp. 183–191, Mar. 2007.
- [66] Y. Wu *et al.*, "State-of-the-art graphene high-frequency electronics," *Nano Lett.*, vol. 12, no. 6, pp. 3062–3067, Jun. 2012.
- [67] Y. Wu, D. B. Farmer, F. Xia, and P. Avouris, "Graphene electronics: Materials, devices, and circuits," *Proc. IEEE*, vol. 101, no. 7, pp. 1620–1637, Jul. 2013.
- [68] S.-J. Han, A. V. Garcia, S. Oida, K. A. Jenkins, and W. Haensch, "Graphene radio frequency receiver integrated circuit," *Nature Commun.*, vol. 5, p. 3086, 2014.

- [69] D. Zhao and Y. Wang, "SD-MAC: Design and synthesis of a hardware-efficient collision-free QoS-aware MAC protocol for wireless network-on-chip," *IEEE Trans. Comput.*, vol. 57, no. 9, pp. 1230–1245, Sep. 2008.



Sergi Abadal (S'12) received the B.Sc. and M.Sc. degrees in telecommunication engineering from the Technical University of Catalunya (UPC), Barcelona, Spain, in 2010 and 2011, respectively, and is currently pursuing the Ph.D. degree in computer architecture at the NaNoNetworking Center in Catalunya, UPC.

From 2009 to 2010, he was a Visiting Researcher with the Broadband Wireless Networking Laboratory, Georgia Institute of Technology, Atlanta, GA, USA. His current research interests

are graphene-based wireless and nanophotonic communications for on-chip networks.

Mr. Abadal was awarded by INTEL within his Doctoral Student Honor Program in 2013.



Mario Iannazzo received the M.Sc. degree in electrical engineering from the Technical University of Catalunya, Barcelona, Spain, in 1998, and the M.A. degree in digital arts from the Pompeu Fabra University, Barcelona, Spain, in 2006, respectively, and is currently pursuing the Ph.D. degree in electrical engineering at the Technical University of Catalunya.

From 1998 to 2004, he was an AMS IC Design Engineer with Nokia Mobile Phones Oy, Oulu, Finland. From 2005 to 2006, he was an IC Patent Engineer with Oficina Ponti, Barcelona, Spain. From 2007 to

2009, he was an IC Consultant Engineer with Alten GmbH, Munich, Germany, working with Infineon Technologies AG as an RF IC Test Engineer. From 2010 to 2011, he was an AMS IC Design Engineer with Decawave Ltd., Dublin, Ireland. His current research interests include the areas of graphene transistor modeling, circuit, and transceiver design.



Mario Nemirovsky received the Telecommunications Engineering degree from the National University of La Plata, La Plata, Argentina, in 1980, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, CA, USA, in 1990

He was an Adjunct Professor with the University of California, Santa Barbara, from 1991 to 1998. After being Chief Architect in companies such as Apple, Inc., National Semiconductors, and General Motors (GM), he founded several renowned start-ups

including FlowStorm Networks, Xstream Logic, ConSentry Networks, and Miraveo. In 2007, he became an ICREA Senior Research Professor with the Barcelona Supercomputing Center (BSC), Barcelona, Spain. He holds more than 60 issued patents: He pioneered the concepts of massively multithreading (MMT) processing for the high-performance processor and the by now well-established simultaneous multithreading architecture (SMT). He also architected the GM engine control being used in all GM cars for over 20 years. His current research interests include multithreaded multicore systems, high-performance systems, network processors, and Big Data.



Albert Cabellos-Aparicio received the B.Sc., M.Sc., and Ph.D. degrees in computer science engineering from the Technical University of Catalunya, Barcelona, Spain, in 2001, 2005, and 2008, respectively.

He has also been an Assistant Professor with the Computer Architecture Department and Researcher with the Broadband Communications Group, Technical University of Catalunya, since 2005. In 2010, he joined the NaNoNetworking Center in Catalunya, where he is the Scientific Director. He is an Editor of

Nano Communication Networks and founder of the ACM NANOCOM conference, the IEEE MONACOM workshop, and the N3Summit. He has also founded the LISPmob open-source initiative along with Cisco. He has been a Visiting Researcher with Cisco Systems, San Jose, CA, USA, and Agilent Technologies, Santa Clara, CA, USA, and a Visiting Professor with the Royal Institute of Technology (KTH), Stockholm, Sweden, and the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. He has given more than 10 invited talks (MIT, Cisco, INTEL, MIET, Northeastern University, etc.) and coauthored more than 15 journal and 40 conference papers. His main research interests are future architectures for the Internet and nanoscale communications.



Heekwan Lee received the B.Sc. degree in electrical engineering from Yonsei University, Seoul, Korea, in 1996, and the M.A. degree in mathematics and M.Sc. and Ph.D. degrees in electrical engineering from the University of Southern California (USC), Los Angeles, CA, USA, in 1999, 2001, and 2005, respectively.

After his graduation, he joined the Samsung Advanced Institute of Technology, Suwon, Korea. Now he is working in DMC with Samsung Electronics. His current research interests include coding theory,

cryptography, and information theory and security.



Eduard Alarcón (S'96–M'01) received M.Sc. (national award) and Ph.D. degrees in electrical engineering from the Technical University of Catalunya (UPC), Barcelona, Spain, in 1995 and 2000, respectively

He became an Associate Professor with UPC in 2001 and has been a Visiting Professor with the University of Colorado at Boulder, Boulder, CO, USA, in 2003, and KTH, Stockholm, Sweden, in 2011. He has coauthored more than 250 scientific publications, four book chapters, and four patents and has been involved in different national, EU, and US R&D projects. Research interests include the areas of on-chip energy management circuits, energy harvesting and wireless energy transfer, and nanocommunications.

Dr. Alarcón was elected an IEEE Circuits and Systems (CAS) Society Distinguished Lecturer, and was elected a member of the IEEE CAS Board of Governors (2010–2013). He was TPC co-chair and a TPC member of 15 IEEE conferences. He has been an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS, IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS, *Journal of Low Power Electronics*, and *Nano Communication Networks*, as well as Co-Editor of four journals special issues and five conference special sessions. He is the recipient of the Best Paper Award at IEEE MWSCAS 1998.