Editorial

Networks-on-Chip: Emerging Research Topics and Novel Ideas

Davide Bertozzi,[1] Shashi Kumar,[2] and Maurizio Palesi[3]

[1] Engineering Department, University of Ferrara, 44100 Ferrara, Italy
[2] Department of Electronics and Computer Engineering, School of Engineering, Jönköping University, 55111 Jönköping, Sweden
[3] Department of Information Engineering and Telecommunication, University of Catania, 95125 Catania, Italy

Networks-on-chip (NoCs) are being devoted intensive research efforts by R&D institutions all around the word, and it is our pleasure to host in this special issue the latest contributions on key design issues at different levels of abstraction, namely, physical link design, architecture design and optimization, performance and power characterization, and design technology. Applying the networking concept to on-chip communication is part of the breakthrough solutions urged by the advances in silicon manufacturing technology (which keeps scaling well beyond 100 nm), by increased time-to-market pressures and by the growing computing requirements of current and future embedded applications. In fact, scalable computation horsepower has been traditionally provided through an increase of clock frequency of monolithic processor cores at each technology node. This trend is, however, running into the barriers of nanoscale technologies, such as heating levels beyond the capability of state-of-the-art packaging and cooling technologies, limited scaling of memory access times and the von Neumann bottleneck. These limitations are being increasingly overcome by breaking up functions into concurrent tasks, assigning them to parallel computational units and operating them at a lower frequency than monolithic cores. This approach paves the way for energy-efficient massively parallel chip-level computation architectures, which are at the core of multiprocessor system-on-chip (MPSoC) technology.

This trend has profound implications on the communication architecture as well, since the communication requirements of an increasing number of processor cores have to be accommodated by the system interconnect. In contrast, state-of-the-art interconnect fabrics will soon incur severe scalability limitations. The International Technology Roadmap for Semiconductors foresees that they will represent the limiting factor for performance and power consumption in next generation SoCs. In the last few years, a number of advances in on-chip interconnect architectures have tried to relieve the limitations of the communication sub-system. First, more parallel topologies have been proposed to increase the amount of delivered bandwidth, such as partial or full crossbars. However, scalability limitations of crossbar-based interconnection fabrics are well known, and they will not be a long-term solution. Second, new communication protocols have been developed, aiming at a more effective exploitation of the available bandwidth. AMBA 3.0 AXI and the open-core protocol (OCP) are examples thereof. Interestingly, these latest protocols provide support for point-to-point communication only (e.g., an IP core with a bus or directly with another IP core) and do not provide any specification on the interconnect fabric, which can (almost) freely evolve in the direction of a higher communication parallelism.

In a short span of seven years, networks-on-chip (NoCs) have been recognized as the most important alternative for the design of modular and scalable communication architectures, providing inherent support to the integration of heterogeneous cores through standard socket interfaces. Not only NoCs relieve system-level integration issues, but are also suitable to deal with the challenges of nanoscale technology. The degradation of the RC propagation delay of signals across global wires is in fact making multiclock cycle signal propagation come true. At the same time, design predictability of global chip-wide structures (like some state-of-the-art system interconnects) is increasingly jeopardized. Through an aggressive path segmentation, NoCs loosen the delay bottleneck of on-chip interconnects and improve design predictability.

Unfortunately, area and power overheads incurred by current NoC prototypes remain still significant in spite of the performance benefits, calling for further research efforts to make this solution more mature and viable from an

industrial viewpoint. Recently, there have been a few books edited describing various aspects of NoC design and implementation. Almost all important international conferences related to electronic design and to its automation have special sessions focusing on NoC-based MPSoC design. This special issue serves the purpose of collecting papers proposing innovative and even exotic solutions to NoC design.

We received twenty-nine submissions, four of which were invited, and a total of ten were accepted. The high level of competition has led to the selection of top-level contributions from renowned academic institutions and industries, covering issues at different levels of the design process. In particular, the topics covered by this special issue include physical link design, NoC architecture design and optimization, power modeling and performance evaluation, mapping and routing strategies. Interestingly, we have challenges associated with nanoscale designs (signal integrity and process variability) addressed at different levels of abstraction, from physical- to system-level design.

NoC physical links are responsible for actual signal propagation among routers and/or network interfaces. The choice of a physical link technology (e.g., serial versus parallel, synchronous versus asynchronous, voltage versus current-mode signaling) has deep implications on even system-level performance, area and power figures due to the relevance of wiring for NoC designs. Moreover, it is at this level that the fundamental challenges of nanoscale technologies have to be primarily tackled. In "Variation-tolerant and low-power source-synchronous multicycle on-chip interconnect scheme," the authors M. Ghoneima et al. address the problem of on-chip communication links whose length makes it impossible to reach the destination in a single clock cycle. They present a variation-tolerant low-power source synchronous multicycle interconnect scheme which is proven to be tolerant to process variations and energy-effective when compared to a traditional pipelined link.

In "High-performance long NoC link using delay-insensitive current mode signaling," the authors E. Nigussie et al. present a high-performance long NoC link implementation based on multilevel current mode signaling and delay-insensitive 1-of-4 encoding. They show that current mode signaling reduces the communication latency of long wires significantly compared to voltage mode signaling, making it possible to achieve high throughput without pipelining and/or using repeaters. In "Online reconfigurable self-timed links for fault tolerant NoC" the authors T. Lehtonen et al. tackle the problem of reliable system design by proposing link structures for NoCs that have properties for effectively tolerating transient, intermittent, and permanent errors. They show how considerable enhancements in fault tolerance can be achieved at the cost of performance and area, and with only a slight increase in power consumption.

At a higher level of abstraction, the behavior of on-chip interconnects, as well as of other NoC building blocks such as switches and network interfaces, can be captured by means of high-level yet accurate models.

In "Area and power modeling for networks-on-chip with layout awareness" the authors P. Meloni et al. present a

flow to devise analytical models of area occupation and power consumption of NoC switches for a reference architecture. Such models are parameterized on several architectural, synthesis-related, and traffic variables.

While a lot of prior work focuses on the switch network, protocol interactions between NoC and IP cores should be carefully considered since they introduce message dependencies that affect deadlock properties of the MPSoC as a whole. In "Avoiding message-dependent deadlock in network-based systems on chip" the authors A. Hansson et al. analyse message-dependent deadlock, survey possible solutions, and show that deadlock avoidance, in the presence of higher-level protocols, poses a serious challenge for many current NoC architectures. Finally, the authors evaluate the solutions qualitatively, and for a number of designs they quantify the area cost for the two most economical solutions, namely, strict ordering and end-to-end flow control.

Developing NoC-based systems tailored to a particular application domain is crucial for achieving high-performance, energy-efficient customized solutions. The effectiveness of this approach largely depends on the availability of a design methodology that, starting from a high-level application specification, derives an optimized NoC configuration with respect to different design objectives. Design choices include topology mapping, routing, and assignment of network channel capacity.

In "A unified approach to mapping and routing on a network on chip for both best-effort and guaranteed service traffic," the authors A. Hansson et al. address the problem of spatial mapping of cores and routing of the communication between cores on NoCs. They present a unified single-objective algorithm which couples path selection, mapping of cores and time-division multiplexing time-slot allocation to minimise the network required to meet the constraints of the application. The application of the algorithm to an MPEG decoder SoC results in a quite significant reduction of area, power dissipation, and worst-case latency over a traditional multistep approach.

In "A method for routing packets across multiple paths in NoCs with in-order delivery and fault-tolerance gaurantees," the authors S. Murali et al. present a multipath routing strategy that guarantees in-order packet delivery for NoCs. The strategy is based on the idea of routing packets on partially nonintersecting paths and rebuilding packet order at path reconvergent nodes. The authors present a design methodology that uses the routing strategy to optimally spread the traffic in the NoC to minimize the network bandwidth needs and power consumption. They also integrate support for tolerance against transient and permanent failures in the NoC links in the methodology by utilizing spatial and temporal redundancy for transporting packets.

In "Network delays and link capacities in application-specific wormhole NoCs," the authors Guz et al. consider a NoC-based application-specific SoC scenario, where information traffic is heterogeneous and delay requirements may largely vary. In this context, the individual capacity assignment for each link in the NoC is required. The authors present an analytical delay model for virtual-channeled

wormhole networks with nonuniform links and apply the analysis in devising an efficient capacity allocation algorithm which assigns link capacities such that packet delay requirements for each flow are satisfied.

Finally, the special issue reports two contributions that take a more radical approach to NoC design, based on novel architectures or concepts that represent revolutionary solutions with respect to the common design practice.

In "Comparison of a ring on-chip network and a code-division multiple-access on-chip network," the authors Wang and Nurmi discuss advantages and disadvantages of applying CDMA techniques for on-chip communication, consisting of the multiplexing of data transfers in code domain instead of in time domain. A comparison with a bidirectional ring connection scheme is performed in terms of network structure, data transfer principle, network node design, asynchronous design, and performance.

In "Stochastic communication: a new paradigm for fault-tolerant networks-on-chip," the authors P. Bogdan et al. introduce a novel communication paradigm for SoCs, called stochastic communication, which allows to relax the requirement of 100% correctness for devices and interconnects by providing a high degree of system-level fault-tolerance NoCs. Using this communication scheme, authors show how a large percentage of data upsets, packet losses due to buffers overflow, and severe levels of synchronization failures can be tolerated, while providing high levels of performance.

We sincerely hope you will enjoy this special issue and that it will inspire further research in this very important area of electronic system design.

We would like to thank all authors who submitted papers to this special issue as well as the authors of the invited papers. Special thanks go to the referees for their time and diligence during the review process and for providing us with high-quality reviews. In conclusion, we would like to thank Bernard Courtois, Editor-in-Chief of VLSI Design, for offering us the opportunity to bring about this special issue.

Davide Bertozzi
Shashi Kumar
Maurizio Palesi