

Exploiting New Interconnect Technologies in On-Chip Communication

John Kim, *Member, IEEE*, Kiyong Choi, *Senior Member, IEEE*, and Gabriel Loh, *Senior Member, IEEE*

Abstract—The continuing scaling of transistors has increased the number of cores available in current processors, and the number of cores is expected to continue to increase. In such manycore processors, the communication between cores with the on-chip interconnect is becoming a challenge as it not only must provide low latency and high bandwidth but also needs to be cost-effective in terms of power consumption. The communication challenge is not only within a single chip but providing high bandwidth to the increasing number of cores from off-chip memory is also a challenge. The conventional metal interconnect is limited, especially for global communication, and can not scale efficiently. In this paper, we investigate alternative interconnect technologies that can be exploited to address the communication challenges in future manycore processor. We provide an overview of the different technologies that are available and then, investigate how these interconnect technologies impact the architecture of the on-chip communication and the system design.

Index Terms— Integrated circuit interconnections, multiprocessor interconnection, parallel architectures.

I. INTRODUCTION

ACCORDING to the International Technology Roadmap for Semiconductors (ITRS) [1], transistors will continue to scale but conventional metal wires will not scale efficiently in terms of latency and energy. The ITRS roadmap suggests that a new interconnect paradigm is needed to continue the scaling of transistors and the number of cores. The interconnect must not only provide high performance, which includes lower latency and higher bandwidth, but also needs to be very *cost-efficient*—i.e., minimize performance per cost where cost can be energy or area.

There has been significant amount of work done on improving the efficiency and performance of wires [2]. For local communication, conventional electrical wires will likely be very cost-efficient. However, with the increasing number of components being integrated into a single chip and larger chip size, the distance of on-chip communication will increase and

Manuscript received May 17, 2012; accepted May 17, 2012. Date of current version June 07, 2012. This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0003579).

J. Kim is with the Department of Computer Science, Korea Advanced Institute of Technology (KAIST), Daejeon 305-701, Korea.

K. Choi is with the School of Electrical Engineering and Computer Science, Seoul National University, Seoul 151-742, Korea.

G. Loh is with AMD Research, Advanced Micro Devices, Inc., Bellevue, WA 98007 USA.

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/JETCAS.2012.2201031

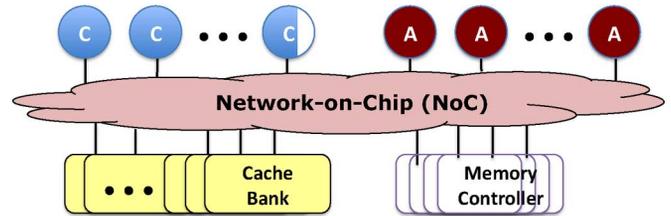


Fig. 1. Block diagram of manycore processors, consisting of processing cores (C), an on-chip cache, memory controllers, and different accelerators (A).

conventional wires will not scale efficiently, especially for global communication. As a result, one of the challenges to enable the scalability of future manycore processors is the on-chip communication [3].

To interconnect all the on-chip components together, the network-on-chip (NoC) [4], [5] approach has been proposed to design communication subsystem in system-on-chip (SoC) and manycore processors. A high-level block diagram of a system using NoC is shown in Fig. 1. The system consists of a large number of cores and a memory system (including different levels of the cache) and is connected to off-chip memory through the on-chip memory controllers. Future manycore processors can also be heterogeneous and contain different types of cores such as accelerators. Over the past ten years, there has been a substantial amount of research on different aspects of NoC [6], including architecture, circuits, and systems that employ NoC, but many of these works have used conventional, electrical signaling for communication. With the limitations of conventional electrical signaling¹ and its impact on the scalability of future manycore processors, there has been a recent increase in interest on NoC research using alternative interconnect technologies. This work explores how these new interconnect technologies can be exploited for on-chip communication.

The rest of the paper is organized as follows. In Section II, to avoid the limitations of conventional electrical signaling, we provide a brief overview of new alternative interconnect technologies and how they can be leveraged for on-chip communication. Given the availability of these new interconnect technologies, we discuss the main components that impact the design of an on-chip communication in Section III and the on-chip network architecture including topology, routing, and flow control in Section IV. Section V present a discussion on other relevant issues in designing an on-chip network including their impact on the system design.

¹In this work, we define conventional electrical signaling as channels that are designed with *RC* wires and repeaters.

II. INTERCONNECT TECHNOLOGIES

In this section, we provide an overview of some of the new interconnect technologies that have been recently proposed. Different technologies present different benefits and challenges. Based on these technologies, we describe how these technologies can be used for on-chip communication in the following sections.

A. Nanophotonics

Photonic interconnects have been widely used in long-haul interconnect and, recently, are being used in large-scale supercomputers and datacenters. The continued advances in photonic technology [7]–[10] have resulted in the decrease of CMOS-compatible device sizes and have become comparable to electrical components. These advances have enabled photonic interconnects for communication within a chip and chip-to-chip communication by providing more energy-efficient on-chip global communication with higher bandwidth and lower latency compared with electrical signaling [11].

The main components of on-chip nanophotonic communication include a light source, the waveguide where the light is routed, a modulator for electrical to optical (E/O) signal conversion, and a detector for the optical to electrical (O/E) conversion. For the modulators and the detectors, micro-ring resonator-based technology is commonly used. Each of these components can be built in a CMOS-compatible process to reduce the cost of integrating nanophotonics with other logic. The light source often leverages an off-chip laser and is coupled into on-chip waveguides. The waveguides, consisting of a high refractive index material as the core and a low refractive index material that form the cladding, are the “channels” used to transmit the light and guide the light. The ring resonators only couple a specific wavelength from the power waveguides while Germanium-based modulators are used as detectors. Dense wavelength division multiplexing (DWDM) allows multiple independent signals to share the same optical waveguide to increase the overall throughput.

Despite the benefits of nanophotonics, the technology has some disadvantages. For example, it is difficult to implement control logic or buffers in optics and thus, the control logic needs to be implemented with electrical signaling. The overhead of using on-chip nanophotonics can result in higher static power because of the micro-rings. The wavelength of each micro-rings can also drift as temperature changes. To avoid this drift, trimming process is commonly used which provides constantly heats the rings [12], [13]. Since the trimming process is done regardless of the activity on the optical channels, they introduce static power overhead. In addition, crossing of the waveguides results in signal losses and thus, waveguide crossing needs to be minimized or avoided.

B. RF/Wireless

The radio-frequency interconnect (RF-I) [14], [15] has been proposed for off-chip as well as on-chip communication. RF-I is projected to scale better than traditional *RC* wires [16] and reduce latency while providing high aggregate bandwidth. Unlike conventional *RC* wires that require charging or discharging the entire length of the wire, an electromagnetic carrier wave is sent

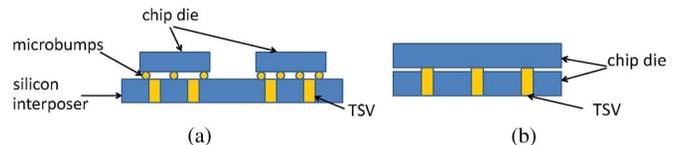


Fig. 2. Logical Diagram of 2.5D and 3D stacking. (a) 2.5D stacking. (b) 3D stacking.

along the transmission line in RF-I with the data modulated on to the carrier wave.

Similar to nanophotonics, the bandwidth efficiency of RF-I can be improved by sending multiple data simultaneously using multi-band RF, which requires multiple transmitters or mixers on the sender side to convert a data stream into a specific frequency band and multiple receivers on the receiving side to down-convert each signal. Wireless differs from RF as the communication medium is free space and also differs from other alternative interconnects as the *channel* does not need to be physically laid out and thus, is not limited by the interconnect routing. Wireless communication can be over different frequency ranges, as summarized in [17], and they impact the physical layer design, including the on-chip antennas and the wireless transceivers [18]. Challenges in wireless interconnect includes minimizing interference, and the cost of wireless links is proportional to the distance in terms of communication, as higher energy is needed for lower distance wireless communication.

C. 3D Integration

The idea of stacking multiple dies, i.e., 3D die-stacking or 3D integration, is not a new one. However, recent advances over the past several years in critical technologies such as through-silicon via (TSV) manufacturing, wafer thinning, wafer/die bonding, micro-bump construction, and other areas have moved the technology to being imminently practical for wide-spread, high-volume manufacturing [19]–[21]. Memory vendors are already showcasing advanced die-stacked prototypes involving multiple layers of DRAM [22] and even mixed DRAM+logic stacks [23]. Academic prototypes are advancing in complexity, with large many-core systems die-stacked with memory [24], [25].

Die-stacking technology has two primary incarnations. The first, and simpler, is called 2.5D stacking or horizontal stacking [26]. In this approach, multiple silicon chips are stacked side-by-side on a passive silicon interposer, as illustrated in Fig. 2(a). This approach simplifies the manufacturing process by only requiring simpler-to-implement “micro-bumps” at the interface between the chips and the interposer. The TSVs penetrate the interposer to provide power, ground, and IO connectivity to the external C4 bumps. Note that the interposer layer only contains a few metal layers to provide chip-to-chip and chip-to-TSV routing, but it does not support any transistors. The lack of devices simplifies the usage of TSVs because issues such as TSV impact on device performance (for example, TSV etching can disrupt the crystal structure of the silicon, thereby impacting carrier mobility in the devices) can be avoided. The second, and perhaps more iconic, approach

is vertical 3D stacking where multiple active chips (i.e., with transistors) are stacked on top of each other, as shown in Fig. 2(b) [27]. Vertical stacking may make potential use of one (or more) of several layer-to-layer interconnect technologies, which may include direct electrical connections in the form of TSVs or micro-bumps, capacitively-coupled circuits, and inductively-coupled circuits, each with different area, power, latency tradeoffs, and system implications.

Die-stacking technology provides the ability to integrate chips that would conventionally be placed in two separate packages into the same stack (whether through 2.5D or 3D integration). This provides direct benefits from the replacement of costly off-chip interconnects with the in-package interconnects. Even conventional single-chip systems may be repartitioned across multiple die-stacked chips, converting long, global, 2D on-chip routes into shorter 3D interconnects. Fundamentally, all such exploitations of die stacking take longer wires and makes them shorter, thus eliminating RC in the circuit which translates into lower latencies and lower power. Apart from the electrical benefits, die stacking will likely be a key enabling technology for other emerging interconnect technologies as well. On-chip RF/wireless interconnects may benefit from the use of a stacked layer with a logic process and metal stack better suited to the analog drivers, receivers, and transmission lines for on-chip RF signaling. Likewise, die-stacking can enable easier integration of photonic components without having to directly make the photonic devices compatible with the leading-edge logic process technology.

III. ON-CHIP COMMUNICATION COMPONENTS

The two main components that make up any on-chip communication system are the channels and the router microarchitecture. These two components significantly impact NoC architecture, and we discuss how new interconnect technologies impact these components.

A. Network Channel

The cost of the network channel and interconnect technology are an important factor in determining the optimal topology. This was illustrated with the migration towards high-radix networks in large-scale, off-chip networks [28] and has a similar impact on on-chip network topologies. We compare the cost of two interconnect technologies as the length of the channel increases in Fig. 3(a). Cost can be measured in terms of energy per bit or capital cost per bit, and we assume the bandwidth is held constant as the channel length increases. The y -intercepts of these lines represent the overhead of the interconnect technologies and reflect the cost of the transmitters/receivers, or if the cost is measured in terms of energy, it is the static energy of the network channel. The slope represents the change in cost as the channel length changes. A smaller slope represents the cost of the channel for a particular interconnect technology is less sensitive to the channel length. The intersection of the two lines represents the trade-off between these two technologies. In Fig. 3(a), for channel length smaller than l_c , technology A is more cost-efficient while for lengths greater than l_c , technology B is more cost-efficient. For these two technologies, a heterogeneous network based on the combination of the two intercon-

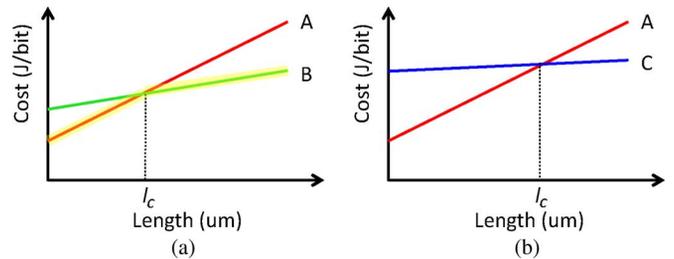


Fig. 3. Cost comparison of two interconnect technologies.

nects [shown with the highlighted line in Fig. 3(a)] can represent a more optimized topology that exploits the benefit of both interconnect technologies.

In this plot, for simplicity, we assumed that the cost of the technologies vary linearly with distance, but that might not necessarily be the case for some technologies. Some interconnect will realize a step-function behavior while other interconnect technologies have a flatter line with a slope approaching zero—i.e., once the overhead cost of the channel is paid for, the cost is relatively insensitive to the channel length. Fig. 3(b) shows such an example with nearly zero-slope interconnect technology but with a higher y -intercept. Compared with Fig. 3(a), the intersection point is moved substantially to the right (i.e., longer channel distance). Thus, the cost of the channels, in particular how it changes with the channel length, needs to be properly considered in the NoC architecture.

B. Channel Layout

In addition to the channel and its cost, another significant impact on the topology design is the packaging constraints of the interconnection network [29]. For large-scale networks, the packaging constraints of the router chip, backplane, cables, etc. impact the topology design, but for on-chip networks, the most important packaging constraint is how the channels are laid out within a 2D planar layout. Using 3D integration changes the packaging constraint with the added dimension, but the constraints within each level of 3D IC are still dependent on the planar layout. In most NoC channels based on conventional electrical signaling, the physical channel layout often corresponds to the logical channel that connects different router nodes together, and the length of the physical channel impacts the characteristics of the channel. If multiple logical channels are needed in the network, multiple physical channels (or wires) must be laid out to provide connectivity. However, one key difference between some of the advanced interconnect technologies and conventional electrical-signaling is the logical channel mapping to the physical channels, as the physical channel does not necessarily correspond to a particular logical channel connectivity. In Fig. 4, an example of a physical channel layout is shown with a physical channel consisting of various logical channels. By having multiple logical channels share a physical channel, it enables a scalable topology that could not be designed with conventional electrical wires.

Nanophotonics can leverage such advantages of a global physical channel to implement multiple logical channels. The physical channel can be laid out in a circular ring or a snake-like shape to implement different crossbar organiza-

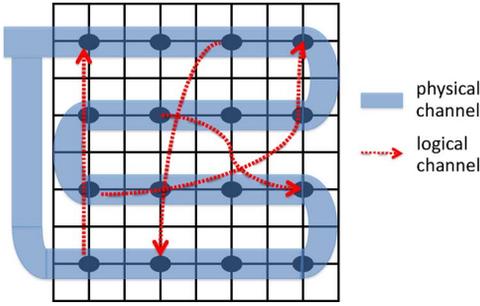


Fig. 4. Physical channel and logical channel mapping using advanced interconnect technology.

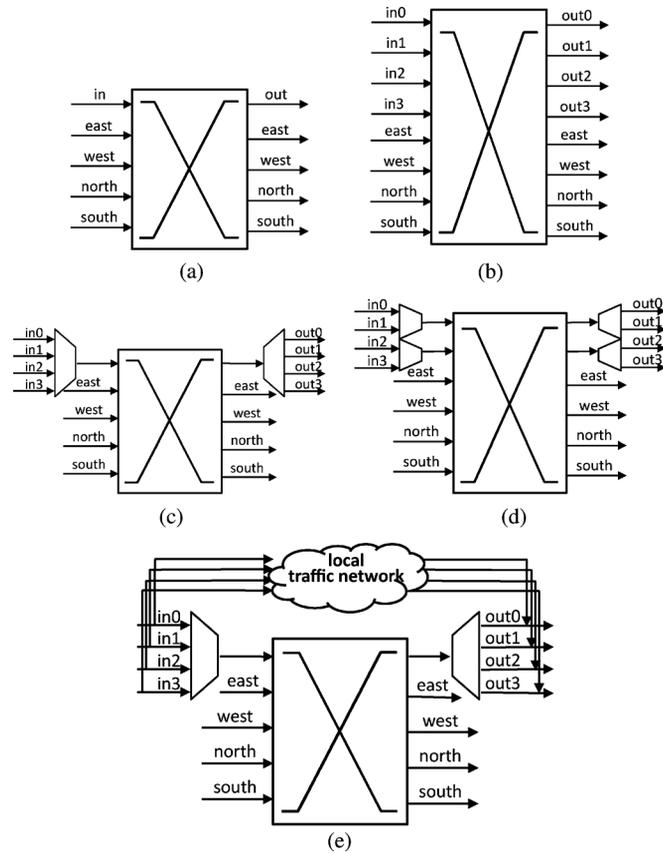


Fig. 5. Four-way concentration implementation in a 2D mesh network: (a) no concentration; (b) integrated concentration; (c) external concentration; (d) hybrid concentration; (e) external concentration with separate local traffic network.

tions including multiple-write-single-read (MWSR) [30] or a single-write-multiple-read (SWMR) [31], [32]. These two implementations differ on how the crossbar arbitration is done, but all provide high connectivity using a shared physical channel. Radio Frequency interconnect (RF-I) [33] also leverages a similar structure shown in Fig. 4 to provide additional logical channels to a conventional 2D mesh topology.

C. Concentration

One common component that impacts the scalability of the network is the *concentration* or sharing of network resources

among multiple terminal² nodes [34]. An example of a four-way concentration for a 2-D mesh topology is shown in Fig. 8(b) and results in a concentrated mesh (CMESH) [35]. With four-way concentration, the number of routers in the topology is reduced by a factor of 4 and also reduces the hop count since the number of intermediate routers decreases. In general, for a network with a concentration degree of C , the number of routers required can be reduced by a factor of C , thereby reducing the network cost. Concentration can be applied to other on-chip network topologies—for example, the flattened butterfly [36] example shown in Fig. 8(c) is another topology which uses a four-way concentration.

Concentration can be implemented in different ways and two examples include *external* concentration or *integrated* concentration [37]. External concentration adds a concentrator (or a mux) to the input of the router and a distributor (or a demux) to the output to share the network injection/ejection bandwidth, while an integrated concentrator integrates additional ports into the router. The integrated concentration approach requires increasing the router radix by $C - 1$, while external concentration does not require any changes to the router. The two implementations present different trade-offs: the external concentration minimizes the router complexity but limits the amount of bandwidth available to the terminal nodes while the integrated concentration increases the router complexity but provides more bandwidth to the terminal nodes. Integrated concentration can also increase overall performance by enabling multiple terminal nodes to use the network bandwidth simultaneously—e.g., in0 can use the East output while in1 can use the North output. For illustration, the external concentration was shown with a mux, but other techniques can be used to implement external concentration—for example, a local bus [38] can be used to aggregate local traffic before connecting to the on-chip network.

Although network cost can be reduced with external concentration, one performance bottleneck is for *local concentrated traffic*—i.e., traffic between the nodes being concentrated together or sharing the network channel bandwidth. If the local traffic uses the router, then the amount of bandwidth available for local traffic is reduced by a factor of C . This bottleneck can be avoided by either increasing the number of concentrators and creating a *hybrid* concentration [Fig. 5(d)] or by having a separate *local* network for local concentrated traffic [Fig. 5(e)]. This local network cannot be used for communicating with any other nodes in the network but only the nodes sharing the concentrator into the network. The concentrator (mux) and the local communication network are drawn separately in the figure but can be physically shared. The local network can be implemented with any architecture including a bus or a crossbar switch.

D. Router Microarchitecture

In a multi-stage network, a router is the building block of the network and the per-hop router delay impacts overall network latency. The main components of a router microarchitecture that impact its delay include the datapath components, such as the buffers and crossbar switch, and the control logic, which

²Terminal nodes are defined as any components that communicate through the on-chip network.

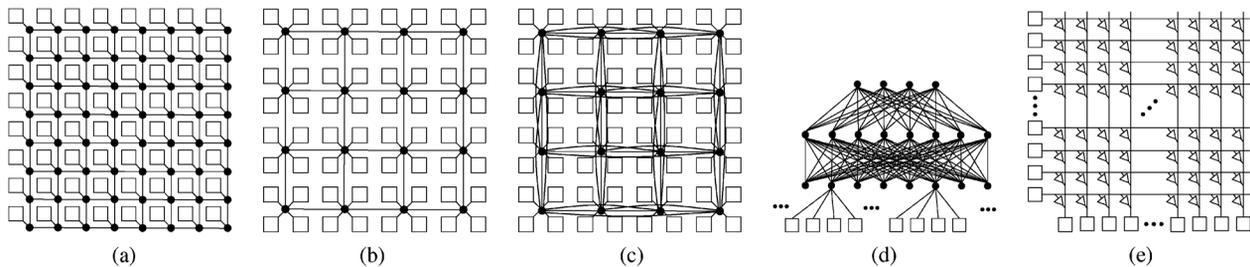


Fig. 8. Examples of different topologies for on-chip networks. (a) Mesh. (b) Concentrated mesh. (c) Flattened butterfly. (d) Fattree. (e) Crossbar.

E. Alternatives to Packet-Switched Network-on-Chip

The previous sections described different components of an network-on-chip and different issues that need to be considered when advanced interconnect technology is leveraged. The NoC assumed a packet-switched network as messages are converted to packets and injected into the network. However, such a packet-switched network introduces overhead, which includes area and power consumed by the buffers and the switch within the router. A very different approach to providing on-chip communication is avoiding packet-switched networks and leveraging a simple bus structure. Borkar [52] argued that with the large amount of wires available on-chip, a packet-switched network is not necessary and a bus-based interconnection network should be designed. A scalable bus design for a 64-node system was described [53] where a multiple-segment broadcast bus was used to create a bus-based NoC that was shown to be more energy efficient than a packet-switched network. Oh *et al.* [54] extended the idea of using a global bus by leveraging the benefits of transmission lines to provide high throughput while reducing global latency. The transmission line is not used a global bus but as a shared medium for point-to-point communication. Although these alternatives have been shown to improve the cost of on-chip communication, one of their assumptions is that global traffic is not necessarily high compared to local traffic. If there is significant amount of global traffic for a given application, these alternatives might not be appropriate. It is also not clear how scalable these approaches will be if the number of cores continue to increase. However, these alternatives can be combined with packet-switched networks to create a heterogeneous NoC (Section IV-A2) to provide additional scalability.

IV. NETWORK-ON-CHIP ARCHITECTURE

Based on the components described in the previous section, we discuss how these components impact the design of the NoC architecture, including the topology, routing, and the flow control.

A. Topology

Topology defines how the channels and the routers are connected in an interconnection network and determines the performance bounds—including zero-load latency and network throughput [34]. Examples of different on-chip network topologies are shown in Fig. 8. The topologies can be characterized by the router *radix* or the number of ports in the router and the network diameter that corresponds to the maximum hop count

between any two nodes. As the radix increases, the network diameter decreases, which can increase performance by reducing the hop count and network latency. However, increasing the radix can also increase router complexity. In comparison, as the radix decreases, the router microarchitecture is simplified but also increases the network diameter. The radix also impacts the channel length in the network. For example, a low-radix 2D mesh network keeps the channel length minimal as routers are connected only to neighboring routers. However, high-radix topologies requires longer channels to provide higher connectivity—e.g., the 2D flattened butterfly [36] requires long channels that are proportional to the single dimension of the chip and a crossbar requires longer channels. The channel lengths has significant impact on the channel cost, as described earlier in Section III-A, and the channel cost impacts the optimal NoC topology. In this section, we describe different topologies proposed that exploits the benefits of new interconnect technologies. In this work, we define a *homogeneous* NoC as a network where only a single interconnect technology is used while a *heterogeneous* NoC is defined as a network using two or more different interconnect technologies.

1) *Homogeneous Topology*: Based on different NoC topologies that have been proposed, only a few works rely strictly on advanced interconnect technologies for on-chip communication to create a homogeneous NoC topology. Corona [30] is one example of a homogeneous network as it leverages a global, nanophotonic crossbar. However, Corona uses a four-way concentration in the architecture and although the details are not clear, it is unlikely that local nodes communicate through the optical crossbar but more likely that they use some form of conventional electrical signaling for local communication. Instead of a global crossbar, Joshi *et al.* [55] propose a Clos network for global communication using silicon-photonics to reduce the network power while a fattree-based optical NoC (FONoC) [56] has also been proposed. Different topologies that leverage 3D integration [57] can also be considered homogeneous topologies since the TSVs are often treated as very short wires, enabling conventional 3D topologies such as a 3D mesh or hybrid bus-NoC approach where the TSVs are used as a bus that interconnects all the layers [42].

2) *Hierarchical, Heterogeneous Topology*: A heterogeneous network combines the benefits of different interconnect technologies and can be classified as an hierarchical network or a flat network (Fig. 9). A hierarchical heterogeneous network can consist of two different types of networks—a local network that interconnects some number of neighboring or local nodes and

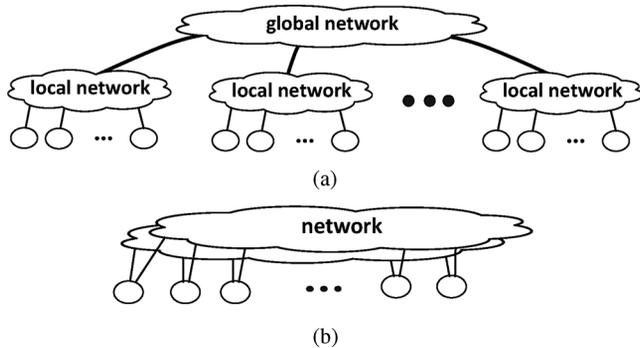


Fig. 9. Different heterogeneous network implementations using different interconnect technologies. (a) Hierarchical, heterogeneous network. (b) Flat, heterogeneous network.

a global network that interconnects the local network together for global communication.³ As described earlier, most of the interconnect technologies that have been proposed are often efficient for global communication and the overhead does not make them feasible for short-range communication. Because of this difference in cost based on channel length, a hierarchical heterogeneous topology can be created where one interconnect technology is used for the local network while another interconnect technology is used for the global network [Fig. 9(a)].

Both Firefly [32] and the opto-electrical crossbar [58] use an electrical mesh for local communication while a global optical crossbar is used for global communication, but these two networks differ in how the crossbar is implemented. Leveraging the benefits of plasmonic components in nanophotonics architecture, a hybrid flattened butterfly topology was proposed where the long channels that span the entire chip use plasmonic/photonic channels [59]. Another example is a radio frequency interconnect (RF-I) network where a conventional 2D mesh network is used as the baseline network, but this network is overlaid with an RF network that adds *shortcuts* to the network [33]. The shortcuts are express channels [60] that provide high bandwidth channels between different router nodes in the network while reducing the latency. WCube [61] is another hybrid architecture that combines wireless technology with a baseline concentrated mesh topology using electrical signaling. The wireless connectivity is similar to a hypercube, but instead of having dedicated channels connect the wireless routers, each wireless router has a single wireless transmitter and multiple receivers. The wireless router is shared among all the nodes within a cluster of nodes. Ganguly *et al.* [62] proposed a hierarchical structure that consists of clusters where the intra-cluster is connected using a wired interconnect. Within each cluster, a *hub* router that is used to communicate with other clusters. The neighboring clusters are connected using conventional wires while wireless interconnect is used to interconnect multihop clusters.

3) *Flat, Heterogeneous Topology*: In comparison, a *flat* network uses a network for the communication between all the nodes, but multiple networks can exist in parallel [Fig. 9(b)]—with each network possibly used for a different

³Concentration, described earlier in Section III-C, does provide a level of hierarchy as the nodes connected to the same concentrator create an initial hierarchy level of the network.

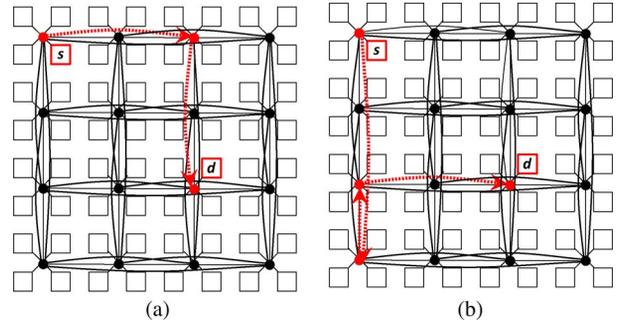


Fig. 10. Routing examples on the flattened butterfly topology from source (s) to destination (d). (a) Minimal routing. (b) Nonminimal routing.

purpose. This type of network can be viewed as a channel-sliced network [34] since the bandwidth from the nodes are sliced into multiple parallel networks. However, by leveraging the characteristics of the different interconnect technologies, the parallel networks are often used for different purposes. For example, research from Columbia [63] uses both an electrical network and an optical network layer using 3D stacking to create a hybrid circuit-switched network using a 2D torus topology. Since optical processing is difficult, the electrical network is used to setup a circuit, as well as to tear down the circuits, while the optical network is used to transmit the data. ATAC [64] architecture also uses a 2D electrical mesh extended with a global optical crossbar. However, in our classification, they differ from both Firefly [32] and the opto-electrical crossbar [58], where a cluster or group of nodes are connected with the 2D mesh network. In comparison, ATAC has a 2D mesh network (ENET) for the entire node and additionally have an optical crossbar (ONET) that provides efficient global communication.

The different topologies described in this section present different trade-offs in cost and performance. The optimal topology for any given interconnect technology will be impacted by not only the cost of the interconnect but also the communication characteristics of the workload on the manycore processors. For example, if the manycore processor architecture results in a significant local traffic, a hierarchical heterogeneous topology might be more appropriate while if there is uniform global access (e.g., a shared last-level cache distributed across the chip), either a homogeneous topology or a flat heterogeneous topology might better support such a communication pattern.

B. Routing

The routing algorithm determines the path a packet takes from its source to its destination. Routing algorithms can be classified into different algorithms, including minimal and non-minimal routing. Minimal deterministic routing is the simplest routing algorithm to implement and is commonly used, but the performance can be limited. In some topologies, nonminimal routing is critical to improve the throughput of the network in adversarial traffic patterns. An example of non-minimal routing is shown in Fig. 10 for a flattened butterfly topology. Nonminimal routing exploits the path diversity in the topology, and although it increases the hop count, overall latency can be minimized if there is congestion in the minimal path.

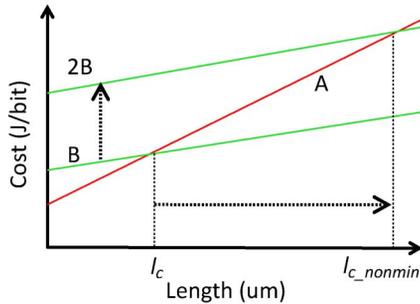


Fig. 11. Cost comparison when nonminimal routing introduces another hop count.

Given the different paths between minimal and nonminimal routes, an adaptive routing algorithm is needed to determine between minimal and nonminimal paths with the adaptive decision often determined by the network congestion [65]. However, it is not clear if nonminimal routing can be done energy-efficiently, especially if the nonminimal routing leverages an advanced interconnect that has higher overhead cost. To illustrate this impact, the cost comparison plot shown earlier in Fig. 3(a) is modified by doubling the cost of interconnect B, from the **B** line to **2B** line, as shown in Fig. 11. We assume non-minimal routing doubles the hop count and results in the cost being doubled. The increase in the cost results in a shift in the intersection point: the original intersection point of length l_c is moved right to l_{c_nonmin} , and the benefit of the new interconnect is only realized at much higher channel length.

As a result, the adaptivity or flexibility needs to be introduced into the routing in different ways such that the global channel is only traversed once. An example of such approach is the adaptive shortcuts [66] for the RF-overlaid topology [33]. Since the traffic pattern can vary depending on the workload, an adaptive shortcut allocates bandwidth differently based on the communication pattern. The RF physical channel is mapped to different logical channels and enables reconfiguring the topology via frequency-band reassignment, thereby providing the benefits of adaptive routing without having to pay the cost of traversing extra channels. Adaptive routing is also more complex than deterministic or oblivious routing, but the high bandwidth of alternative interconnect technology can be exploited and leverage oblivious routing. Using nanophotonics, an oblivious routing based on the wavelength to determine the path was used [67] to create an all-optical data communication for a 2D torus network.

The additional cost of multi-hop routing using a global interconnect can also be a problem if the global interconnect is used as point-to-point channels in a multi-stage topology as the benefit can be reduced because of the overhead. For example, a Clos topology [68] is a multi-stage network that provides high performance across any traffic pattern but requires a packet to traverse multiple point-to-point channels. To avoid this, Joshi *et al.* [55] proposed a photonic middle router such that a packet only needs to traverse a single point-to-point channel with nanophotonics while still providing the path diversity of a Clos topology.

C. Flow Control

Flow control determines how the network resources, primarily network channels and buffers, are allocated. Packets are partitioned into one or more *flits* (flow control digits) that are the unit of flow control. A simpler flow control is bufferless flow control [69] where buffers are removed, and when contention for an output occurs, one of the packets is either deflected [40] or dropped and re-transmitted [41]. However, both of these approaches increase the number of channels that need to be traversed and are likely not appropriate global interconnect technologies. Another bufferless approach is circuit switching, where the path from the source to the destination is reserved and avoid any contention in the network when transmitting the data. Circuit switching has been proposed with optics [47], [48] to provide high bandwidth and low latency between the source and destination and overcome the limitation of providing optical buffers. As long as the cost of setting up the circuit can be amortized by the usage, circuit switching can be a cost-efficient alternative.

In comparison to bufferless flow control, *buffered* flow control, such as virtual-channel (VC) buffered flow control [70], is commonly used in many NoC architectures. VCs partition each input buffer into multiple lanes. Buffered flow control requires proper buffer management to avoid buffer overflow through credit-based flow control or on/off flow control [34]. Although many interconnect technologies provide high bandwidth, *flits* need to be properly sized in order to avoid any bandwidth fragmentation. For example, many of the optical NoCs have assumed that a flit is equal to a packet size, which can be a cache line. Although this simplifies flow control, it is not the most efficient method—e.g., if the channel width is 256 bytes but some of the control packets are only 32 bytes, the remaining 224 bytes would not be efficiently utilized.

Flow control must also guarantee that deadlock does not occur in the network. Deadlock can occur because of routing deadlock or high level protocol deadlock and can be handled either with deadlock avoidance or deadlock detection and recovery. Deadlock avoidance often requires additional resources (such as VCs) while deadlock recovery requires, first, detecting deadlock has occurred and then, recovering from it in an efficient manner. Although deadlock avoidance has been commonly used, if deadlock occur infrequently and can be recovered efficiently (fast), deadlock recovery can be a better option. In RF-I [33], deadlock recovery was used instead of deadlock avoidance but requires an extra escape channel. It remains to be seen if a low-latency high-bandwidth global interconnect can be used to simplify the flow control.

To improve the performance of on-chip networks through better flow control, express virtual channel (EVC) [71] has been proposed where intermediate routers are bypassed. EVC enable packets to avoid the latency of traversing intermediate routers. For short-channels such as a 2D mesh topology, such flow control can be done efficiently but it is likely not appropriate for a global interconnect. However, a global interconnect can provide an alternative way of providing the benefits of EVC, such as the NOCHI architecture [72], which leverages low-swing multi-drop wires for flow control signals to overcome some limitations of EVC. In addition, instead of adding *virtual* express

channels such as EVC, the alternate global interconnect technologies can also be exploited by adding *physical* express channels [60]. This can be done either through the topology, such as the flattened butterfly topology, or application-specific long links can be added [73]. Physical express channels can also be leveraged in conjunction with 3D stacking [74], or wireless channels can be added such as iWISE (Inter-router Wireless Scalable Express Channels) [75]. With the alternative interconnect technologies, careful analysis is needed to evaluate the trade-off between physical and virtual express topologies [76].

V. SYSTEM IMPLICATIONS

In this section, we discuss how advanced interconnect technologies impact other aspects of system design including process-memory communication and cache coherence. In addition, we discuss how alternative interconnect technologies can enable other capabilities to future manycore processors.

A. Processor-Memory Communication

Most of the research that has been discussed so far focused on on-chip communications—i.e., communications that occur within a single-chip manycore processor. However, another significant challenge in manycore processors is overcoming the memory bandwidth wall [77]—how to meet the increased bandwidth demands of the increasing number of cores and supply the bandwidth from the off-chip main memory. Using electrical signaling, there can be more than an order of magnitude increase in cost (energy per bit) of going off-chip to main memory. As a result, different studies have been proposed that leverage advanced interconnect techniques not only for on-chip communication but also for processor-memory communication.

Batten *et al.* [58] use an opto-electrical global crossbar using monolithic silicon photonics, but optics is extended to the processor-to-DRAM network to increase the off-chip bandwidth while increasing the energy efficiency of off-chip accesses. Hendry *et al.* [78] also extend optics to off-chip access by extending photonic on-chip circuit switching to off-chip memory access and provide end-to-end communication between the cores and the DRAM modules through circuits. Multiband Radio Frequency Interconnect (MRF-I) is used in a tree-based multi-DIMM memory system architecture to increase the scalability as the RF interconnect is leveraged to provide point-to-point channels in the DIMM tree architecture [79]. 3D integrated technology also provides an efficient DRAM interface as the distance to memory is significantly reduced, and the stacking of memory using TSVs has been proposed [80]. 3D stacking by itself can provide advantages, but it can also be combined with other technologies. Corona [30] explored the idea of extending a manycore processor connected with an optical crossbar with 3D memory stacking, which was also connected optically.

3D stacking often uses TSV to interconnect the different layers but alternative methods can be used to communicate between stacked chips. While TSV still uses wires to communicate between the different layers, wireless methods to communicate between the different stacked chips have been recently proposed that include capacitive and inductive coupling. These technologies can improve the off-chip memory

bandwidth while improving the efficiency of off-chip communication. Capacitive coupling (or proximity communication [81]–[83]) improves the bandwidth/area significantly compared with I/O through ball bonding by using two chips that are placed face-to-face, separated by only a few microns. The capacitive coupling can be used for processor-to-memory communication or can be used to enable a scalable, multi-chip modules but requires that the chips be physically touching one another and is limited to adjacent chips. In comparison, inductive coupling [84], [85] is able to couple links between more than two chips, although the inductor diameter needs to be increased as the number of chips increases. Miura *et al.* [86] describes how DRAM chips can be stacked using the inductive-coupling interface to provide high bandwidth to GPU at low energy.

In addition to simply replacing electrical channels with an alternative interconnect, the interconnect can also impact the design of the DRAM memory system. Beamer *et al.* [87] extend photonics into the DRAM chip to design the photonically-interconnected DRAM (PIDRAM) chip while nanophotonics was used to connect 3D-stacked memory by leveraging a separate interface die within 3D stacked dies [88].

B. Cache Coherence and Other Features

In a shared-memory many-core system, providing scalable cache coherency is a significant challenge [89]. Two approaches are snoop-based cache coherence and directory-based cache coherence, each with a different trade-off. A bus-based network can easily enable snoop-based cache coherence protocol but the scalability of a conventional bus is often limited. Kirman *et al.* [31] leverage an optical bus to provide a scalable bus interconnect and support snoop-based cache coherence. The ATAC [64] architecture also uses nanophotonics to enable a fast, efficient broadcast to implement a new directory-based cache coherence protocol to increase the scalability.

Most of the prior research that leverages new interconnect technologies has attempted to improve the performance (such as latency or bandwidth), reduce the cost (e.g., power), and/or improve the efficiency in terms of Joules per bits. However, these prior works also have assumed the network to be a *dumb* network whose main goal is simply to transport bits from one location to another location. In comparison, intelligence can be added to the network to create a *smart* network and use the interconnect for something other than simply transporting bits. For example, in-network coherence [90], [91] has been proposed using conventional, electrical signaling where cache coherence protocol can be implemented in the network itself by embedding directories in the routers. Since coherence information is added to the network, the round-trip latency required in a directory-based cache coherence is avoided. The low-latency of global communication with alternative interconnect technologies enable fast exchange of information among the different nodes in the network and presents new opportunities to create a more smart network.

In addition, the advanced interconnect technologies can provide additional capabilities. For example, in addition to using the nanophotonic for global data communication or arbitration as discussed earlier, the nanophotonic can be leveraged to build a race-free cache coherence protocol [92]. In this work, the light

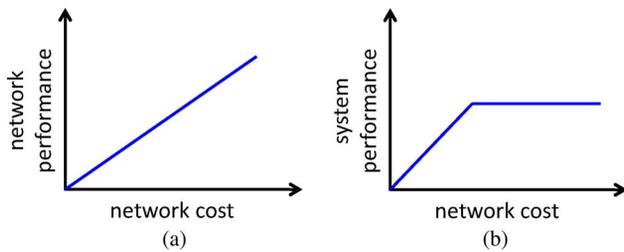


Fig. 12. Impact of on-chip interconnect on network and system performance.

pulse is used to represent a mutex and the low latency of the optics enable atomic coherence to be provided. Nanophotonics can also be leveraged to implement a barrier in a multi-thread workload [93]. TLSync [94] also implements a barrier network using transmission lines as different radio frequency spectrums are used to support multiple barrier networks.

New interconnect technologies can also enable features that might not have been available using conventional electrical signaling technology. For example, 3D stacking is leveraged to assist in debugging and testing by adding another layer connected through 3D stacking that is provided as an option for software developers [95]. One of the biggest challenge of future manycore processors is parallel programming, and within parallel programming, the difficulty is often in the communication aspect between the different cores. An *ideal* network for parallel programming is one where not only high bandwidth and low latency are provided but minimal overhead is introduced for communicating between any two components. One recent approach to provide such communication is a fully-connected topology with direct links between any two nodes: directly connected arbitration-free photonic crossbar (DCAF) [96]. DCAF provides a fully connected topology using photonics, but this is also enabled by 3D stacking, by using a separate layer for photonic channels and using photonic vias and grating couplers [97].

With the increase in the different types of components integrated into a single chip or with different numbers of programs running simultaneously, some form of quality-of-service (QoS) will be needed as different components or workloads will have different resource requirements. Different QoS mechanisms have been proposed for electrical signaling baseline networks [98]–[100]; however, any advanced interconnect can provide additional advantages. For example, GSF [98] requires a barrier network, and the performance impact of supporting QoS is impacted by the barrier network latency. Advanced interconnect technology can potentially improve overall performance through a fast barrier network. In addition, the high bandwidth and low latency of nanophotonics can be leveraged to provide QoS as well [101], [102].

C. Network and System Evaluation

The interconnect performance metrics are relevant, including latency and bandwidth, but how the interconnect impact overall system performance and cost is more important since an on-chip interconnect is not used by itself but integrated with other components. For example, the performance of an NoC-only can be represented with a plot shown in Fig. 12(a)—more resources

(or increased network cost with additional network bandwidth) will likely improve the network performance, such as providing higher throughput. However, for the same x-axis, if we measure the overall system performance, the performance will saturate at some point [Fig. 12(b)]—i.e., additional network resources do not continue to improve overall system performance. Thus, the network must be evaluated appropriately within the system to understand the impact of the interconnect on overall system and cost. The plots in Fig. 12 also assumed a constant noninterconnect component (i.e., fixed number of cores) and plotted performance as the interconnect resource was varied. However, in power-constrained manycore processors, the constraints will be different since as more power is consumed by the network, the power consumed in the cores will need to be reduced which can negatively impact overall performance. Thus, a properly balanced network-on-chip design is needed.

VI. SUMMARY

This paper has provided an overview of new interconnect technologies and their impact on the design of on-chip communication. As interconnect technologies continue to evolve, their impact on on-chip communication will continue to change. Although these technologies present some significant advantages, in order to fully exploit the benefits of these new interconnect technologies, more research is needed to overcome the challenges presented by these technologies and properly incorporate them in the circuit and system design of future systems.

REFERENCES

- [1] International Technology Roadmap for Semiconductors(ITRS) 2009.
- [2] R. Ho, K. W. Mai, S. Member, and M. A. Horowitz, "The future of wires," *Proc. IEEE*, vol. 89, no. 4, pp. 490–504, Apr. 2001.
- [3] J. D. Owens, W. J. Dally, R. Ho, D. N. Jayasimha, S. W. Keckler, and L.-S. Peh, "Research challenges for on-chip interconnection networks," *IEEE Micro*, vol. 27, no. 5, pp. 96–108, Sep./Oct. 2007.
- [4] W. J. Dally and B. Towles, "Route packets, not wires: On-chip interconnection networks," in *Proc. Design Automat. Conf. (DAC)*, Las Vegas, NV, Jun. 2001, pp. 684–689.
- [5] L. Benini and G. De Micheli, "Networks on chips: A new soc paradigm," *IEEE Computer*, vol. 35, no. 1, pp. 70–78, Jan. 2002.
- [6] N. D. E. Jerger and L.-S. Peh, *On-Chip Networks*, ser. Synthesis Lectures on Computer Architecture. San Rafael, CA: Morgan Claypool, 2009.
- [7] V. R. Almeida, C. A. Barrios, R. R. Panepucci, M. Lipson, M. A. Foster, D. G. Ouzounov, and A. L. Gaeta, "All-optical switching on a silicon chip," *Opt. Lett.*, vol. 29, no. 24, pp. 2867–2869, Dec. 2004.
- [8] K. Preston, P. Dong, B. Schmidt, and M. Lipson, "High-speed all-optical modulation using polycrystalline silicon microring resonators," *Appl. Phys. Lett.*, vol. 92, no. 15, p. 151104, 2008.
- [9] T. Woodward and A. Krishnamoorthy, "1-Gb/s integrated optical detectors and receivers in commercial CMOS technologies," *IEEE J. Select. Topics Quantum Electron.*, vol. 5, no. 2, pp. 146–156, Mar./Apr. 1999.
- [10] H. Park, Y. hao Kuo, A. W. Fang, R. Jones, O. Cohen, M. J. Paniccia, and J. E. Bowers, "A hybrid algalain-silicon evanescent preamplifier and photodetector," *Opt. Express*, vol. 15, no. 21, pp. 13 539–13 546, Oct. 2007.
- [11] G. Chen, H. Chen, M. Haurylau, N. Nelson, P. M. Fauchet, E. G. Friedman, and D. Albonese, "Predictions of CMOS compatible on-chip optical interconnect," in *Proc. 2005 Int. Workshop System Level Interconnect Prediction*, New York, 2005, pp. 13–20.
- [12] M. Lipson, "Compact electro-optic modulators on a silicon chip," *IEEE J. Sel. Topics Quantum Electron.*, vol. 12, no. 6, pp. 1520–1526, Nov.–Dec. 2006.
- [13] C. Nitta, M. K. Farrens, and V. Akella, "Addressing system-level trimming issues in on-chip nanophotonic networks," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, San Antonio, TX, 2011, pp. 122–131.

- [14] M. Chang, V. Roychowdhury, L. Zhang, H. Shin, and Y. Qian, "RF/wireless interconnect for inter- and intra-chip communications," *Proc. IEEE*, vol. 89, no. 4, pp. 456–466, Apr. 2001.
- [15] J. Ko, J. Kim, Z. Xu, Q. Gu, C. Chien, and M. Chang, "An RF/baseband FDMA-interconnect transceiver for reconfigurable multiple access chip-to-chip communication," in *Dig. Tech. Papers Int. Solid-State Circuits Conf.*, Feb. 2005, vol. 1, pp. 338–602.
- [16] M. C. F. Chang, J. Cong, A. Kaplan, C. Liu, M. Naik, J. Premkumar, G. Reinman, E. Socher, and S. W. Tam, "Power reduction of CMP communication networks via RF-interconnects," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, Lake Como, Italy, 2008, pp. 376–387.
- [17] S. Deb, A. Ganguly, P. Pande, D. Heo, and B. Belzer, "Wireless NOC as interconnection backbone for multicore chips: Promises and challenges," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, Jun. 2012.
- [18] J.-J. Lin, H.-T. Wu, Y. Su, L. Gao, A. Sugavanam, J. Brewer, and K. O, "Communication using antennas fabricated in silicon integrated circuits," *IEEE J. Solid-State Circuits*, vol. 42, no. 8, pp. 1678–1687, Aug. 2007.
- [19] S. Das, A. Fan, K.-N. Chen, C. S. Tan, N. Checka, and R. Reif, "Technology, performance, and computer-aided design of three-dimensional integrated circuits," in *Proc. 2004 Int. Symp. Physical Design*, New York, 2004, pp. 108–115, ACM.
- [20] P. Morrow, M. Kobrinisky, S. Ramanathan, C. M. Park, M. Harnes, V. Ramachandrarao, H. Park, G. Kloster, S. List, and S. Kim, "Wafer-level 3D interconnects via Cu bonding," in *Proc. 21st Adv. Metallizat. Conf.*, 2004.
- [21] W. R. Davis, J. Wilson, S. Mick, J. Xu, H. Hua, C. Mineo, A. M. Sule, M. Steer, and P. D. Franzon, "Demystifying 3D ICS: The pros and cons of going vertical," *IEEE Design Test Computers*, vol. 22, no. 6, pp. 498–510, Nov. 2005.
- [22] J.-S. Kim, C. S. Oh, H. Lee, D. Lee, H.-R. Hwang, S. Hwang, B. Na, J. Moon, J.-G. Kim, H. Park, J.-W. Ryu, K. Park, S.-K. Kang, S.-Y. Kim, H. Kim, J.-M. Bang, H. Cho, M. Jang, C. Han, J.-B. Lee, K. Kyung, J.-S. Choi, and Y.-H. Jun, "A 1.2 v 12.8 gb/s 2 gb mobile wide-i/o dram with 4×128 i/os using tsv-based stacking," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)*, Feb. 2011, pp. 496–498.
- [23] J. T. Pawlowski, "Hybrid memory cube: Breakthrough dram performance with a fundamentally re-architected dram subsystem," in *Proc. Hot Chips 23*, Stanford, CA, 2011.
- [24] D. H. Kim, K. Athikulwongse, M. Healy, M. Hossain, M. Jung, I. Khorosh, G. Kumar, Y.-J. Lee, D. Lewis, T.-W. Lin, C. Liu, S. Panth, M. Pathak, M. Ren, G. Shen, T. Song, D. H. Woo, X. Zhao, J. Kim, H. Choi, G. Loh, H.-H. Lee, and S. K. Lim, "3D-maps: 3D massively parallel processor with stacked memory," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)*, Feb. 2012, pp. 188–190.
- [25] D. Fick, R. G. Dreslinski, B. Giridhar, G. Kim, S. Seo, M. Fojtik, S. Satpathy, Y. Lee, D. Kim, N. Liu, M. Wiecekowsky, G. Chen, T. Mudge, D. Sylvester, and D. Blaauw, "Centip3de: A 3930 dmips/w configurable near-threshold 3D stacked system with 64 arm cortex-m3 cores," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)*, Feb. 2012, pp. 190–192.
- [26] Y. S. Deng and W. Maly, "2.5d system integration: A design driven system implementation schema," in *Proc. 2004 Asia South Pacific Design Automat. Conf.*, Piscataway, NJ, 2004, pp. 450–455.
- [27] B. Black, D. W. Nelson, C. Webb, and N. Samra, "3D processing technology and its impact on ia32 microprocessors," in *Proc. IEEE Int. Conf. Comput. Design*, Washington, DC, 2004, pp. 316–318.
- [28] J. Kim, W. J. Dally, B. Towles, and A. Gupta, "Microarchitecture of a high-radix router," presented at the Int. Symp. Comput. Archit. (ISCA), Madison, WI, Jun. 2005.
- [29] W. J. Dally, "Performance analysis of k-ary n-cube interconnection networks," *IEEE Trans. Computers*, vol. 39, no. 6, pp. 775–785, 1990.
- [30] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. P. Jouppi, M. Fiorentino, A. Davis, N. L. Binkert, R. G. Beausoleil, and J. H. Ahn, "Corona: System implications of emerging nanophotonic technology," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, Beijing, China, 2008, pp. 153–164.
- [31] N. Kirman, M. Kirman, R. K. Dokania, J. F. Martinez, A. B. Apsel, M. A. Watkins, and D. H. Albonese, "Leveraging optical technology in future bus-based chip multiprocessors," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, Orlando, FL, 2006, pp. 492–503.
- [32] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang, and A. Choudhary, "Firefly: Illuminating future network-on-chip with nanophotonics," presented at the Proc. Int. Symp. Comput. Archit. (ISCA), Austin, TX, 2009.
- [33] M. F. Chang, J. Cong, A. Kaplan, M. Naik, G. Reinman, E. Socher, and S.-W. Tam, "Cmp network-on-chip overlaid with multi-band RF-interconnect," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, Salt Lake City, UT, Feb. 2008, pp. 191–202.
- [34] W. J. Dally and T. B. *Principles and Practices of Interconnection Networks*. Waltham, MA: Morgan Kaufmann, 2004.
- [35] J. Balfour and W. J. Dally, "Design tradeoffs for tiled CMP on-chip networks," in *Proc. Int. Conf. Supercomput. (ICS)*, Carns, Queensland, Australia, 2006, pp. 187–198.
- [36] J. Kim, J. Balfour, and W. J. Dally, "Flattened butterfly topology for on-chip networks," in *IEEE/ACM Int. Symp. Microarchitecture (MICRO)*, Chicago, IL, Dec. 2007.
- [37] P. Kumar, Y. Pan, J. Kim, G. Memik, and A. Choudhary, "Exploring concentration and channel slicing in on-chip network router," in *IEEE Int. Symp. Network-on-Chip (NOCS)*, San Diego, CA, May 2009, pp. 276–285.
- [38] R. Das, S. Eachempati, A. Mishra, V. Narayanan, and C. Das, "Design and evaluation of a hierarchical on-chip interconnect for next-generation CMPs," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, Raleigh, NC, Feb. 2009, pp. 175–186.
- [39] J. Kim, "Low-cost router microarchitecture for on-chip networks," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, Dec. 2009, pp. 255–266.
- [40] T. Moscibroda and O. Mutlu, "A case for bufferless routing in on-chip networks," in *ISCA*, Austin, TX, 2009, pp. 196–207.
- [41] M. Hayenga, N. E. Jerger, and M. Lipasti, "Scarab: A single cycle adaptive routing and bufferless network," in *MICRO*, 2009, pp. 244–254.
- [42] J. Kim, C. Nicopoulos, D. Park, R. Das, Y. Xie, V. Narayanan, M. S. Yousif, and C. R. Das, "A novel dimensionally-decomposed router for on-chip communication in 3D architectures," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, San Diego, CA, Jun. 2007, pp. 138–149.
- [43] K. Puttaswamy and G. H. Loh, "Thermal herding: Microarchitecture techniques for controlling hotspots in high-performance 3D-integrated processors," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, Washington, DC, 2007, pp. 193–204.
- [44] D. Park, S. Eachempati, R. Das, A. K. Mishra, Y. Xie, N. Vijaykrishnan, and C. R. Das, "MIRA: A multi-layered on-chip interconnect router architecture," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, Beijing, China, 2008, pp. 251–261.
- [45] A. Shacham, B. Lee, A. Biberman, K. Bergman, and L. Carloni, "Photonic NOC for DMA communications in chip multiprocessors," in *Proc. Hot Interconnects*, Palo Alto, CA, Aug. 2007, pp. 29–36.
- [46] F. Xia, L. Sekaric, and Y. Vlasov, "Ultra-compact optical buffers on a silicon chip," *Nature Photon.*, vol. 1, no. 1, pp. 65–71, Dec. 2006.
- [47] A. Shacham, K. Bergman, and L. P. Carloni, "On the design of a photonic network-on-chip," in *IEEE Int. Symp. Network-on-Chip (NOCS)*, Princeton, NJ, 2007, pp. 53–64.
- [48] A. Shacham, K. Bergman, and L. Carloni, "Photonic networks-on-chip for future generations of chip multiprocessors," *IEEE Trans. Computers*, vol. 57, no. 9, pp. 1246–1260, Sep. 2008.
- [49] M. J. Cianchetti, J. C. Kerekes, and D. H. Albonese, "Phastlane: A rapid transit optical routing network," presented at the Proc. Int. Symp. Comput. Archit. (ISCA), Austin, TX, 2009.
- [50] D. Vantrease, N. L. Binkert, R. Schreiber, and M. H. Lipasti, "Light speed arbitration and flow control for nanophotonic interconnects," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, New York, 2009, pp. 304–315.
- [51] Y. Pan, J. Kim, and G. Memik, "Flexishare: Channel sharing for an energy-efficient nanophotonic crossbar," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, Bangalore, India, Jan. 2010, pp. 1–12.
- [52] S. Borkar, "Networks for multi-core chips—A contrarian view, ispld keynote," presented at the Int. Symp. Low Power Electron. Design, Portland, OR, Aug. 2007.
- [53] A. N. Udipi, N. Muralimanohar, and R. Balasubramonian, "Towards scalable, energy-efficient, bus-based on-chip networks," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, Bangalore, India, Jan. 2010, pp. 1–12.
- [54] A. Carpenter, J. Hu, J. Xu, M. Huang, and H. Wu, "A case for globally shared-medium on-chip interconnect," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, San Jose, CA, 2011, pp. 271–282.
- [55] A. Joshi, C. Batten, Y.-J. Kwon, S. Beamer, I. Shamim, K. Asanovic, and V. Stojanovic, "Silicon-photonic CLOS networks for global on-chip communication," in *IEEE Int. Symp. Network-on-Chip (NOCS)*, San Diego, CA, 2009, pp. 124–133.
- [56] H. Gu, J. Xu, and W. Zhang, "A low-power fat tree-based optical network-on-chip for multiprocessor system-on-chip," in *Design, Automat. Test Eur. (DATE) Conf.*, Nice, France, 2009, pp. 3–8.

- [57] V. Pavlidis and E. Friedman, "3D topologies for networks-on-chip," *IEEE Trans. Very Large Scale Integrat. (VLSI) Syst.*, vol. 15, no. 10, pp. 1081–1090, Oct. 2007.
- [58] C. Batten, A. Joshi, J. Orcutt, A. Khilo, B. Moss, C. Holzwarth, M. Popovic, H. Li, H. Smith, J. Hoyt, F. Kartner, R. Ram, V. Stojanovic, and K. Asanovic, "Building manycore processor-to-dram networks with monolithic silicon photonics," in *Proc. Hot Interconnects*, Stanford, CA, 2008, pp. 21–30.
- [59] H. Wassel, D. Dai, L. Theogarajan, J. Dionne, M. Tiwari, J. Valamehr, F. Chong, and T. Sherwood, "Opportunities and challenges of using plasmonic components in nanophotonic architectures," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, Jun. 2012.
- [60] W. J. Dally, "Express cubes: Improving the performance of k-ary n-cube interconnection networks," *IEEE Trans. Computers*, vol. 40, no. 9, pp. 1016–1023, Sep. 1991.
- [61] S.-B. Lee, S.-W. Tam, I. Pefkianakis, S. Lu, M. F. Chang, C. Guo, G. Reinman, C. Peng, M. Naik, L. Zhang, and J. Cong, "A scalable micro wireless interconnect structure for CMPS," in *Proc. Networking*, Beijing, China, 2009, pp. 217–228.
- [62] A. Ganguly, K. Chang, S. Deb, P. Pande, B. Belzer, and C. Teuscher, "Scalable hybrid wireless network-on-chip architectures for multicore systems," *IEEE Trans. Computers*, vol. 60, no. 10, pp. 1485–1502, Oct. 2011.
- [63] A. Shacham, K. Bergman, and L. P. Carloni, "The case for low-power photonic networks-on-chip," in *Proc. Design Automat. Conf. (DAC)*, San Diego, CA, 2007, pp. 132–135.
- [64] G. Kurian, J. E. Miller, J. Psota, J. Eastep, J. Liu, J. Michel, L. C. Kimerling, and A. Agarwal, "ATAC: A 1000-core cache-coherent processor with on-chip optical network," in *Proc. 19th Int. Conf. Parallel Archit. Compilat. Techn.*, Vienna, Austria, 2010, pp. 477–488.
- [65] Singh, "Load-balanced routing in interconnection networks," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2005.
- [66] M.-C. F. Chang, J. Cong, A. Kaplan, C. Liu, M. Naik, J. Premkumar, G. Reinman, E. Socher, and S.-W. Tam, "Power reduction of CMP communication networks via RF-interconnects," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, Lake Como, Italy, Nov. 2008, pp. 376–387.
- [67] N. Kirman and J. F. Martinez, in *Proc. Archit. Support Programm. Languages Operat. Syst. (ASPLoS)*, Pittsburgh, PA, 2010, pp. 15–28.
- [68] C. Clos, "A study of non-blocking switching networks," *Bell Syst. Tech. J.*, vol. 32, no. 2, pp. 406–424, Mar. 1953.
- [69] P. Baran, "On distributed communications networks," *IEEE Trans. Prof. Tech. Group Commun. Syst.*, Jan. 1964.
- [70] W. J. Dally, "Virtual-channel flow control," *IEEE Trans. Parallel Distrib. Syst.*, vol. 3, no. 2, pp. 194–205, 1992.
- [71] A. Kumar, L. S. Peh, P. Kundu, and N. K. Jha, "Express virtual channels: Towards the ideal interconnection fabric," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, San Diego, CA, Jun. 2007, pp. 150–161.
- [72] T. Krishna, A. Kumar, P. Chiang, M. Erez, and L.-S. Peh, "NOC with near-ideal express virtual channels using global-line communication," in *Proc. Hot Interconnects*, Stanford, CA, 2008, pp. 11–20.
- [73] U. Ogras and R. Marculescu, "'It's a small world after all': NOC performance optimization via long-range link insertion," *IEEE Trans. Very Large Scale (VLSI) Syst.*, vol. 14, no. 7, pp. 693–706, Jul. 2006.
- [74] Y. Xu, Y. Du, B. Zhao, X. Zhou, Y. Zhang, and J. Yang, "A low-radix and low-diameter 3D interconnection network design," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, Raleigh, NC, 2009, pp. 30–42.
- [75] D. DiTomaso, A. Kodi, S. Kaya, and D. Matolak, "iwise: Inter-router wireless scalable express channels for network-on-chips (NOCS) architecture," in *IEEE 19th Annu. Symp. High Performance Interconnects (HOTI)*, Santa Clara, CA, Aug. 2011, pp. 11–18.
- [76] C.-H. O. Chen, N. Agarwal, T. Krishna, K.-H. Koo, L.-S. Peh, and K. C. Saraswat, "Physical vs. virtual express topologies with low-swing links for future many-core NOCS," in *Proc. 2010 4th ACM/IEEE Int. Symp. Networks-on-Chip*, Washington, DC, 2010, pp. 173–180.
- [77] W. A. Wulf and S. A. McKee, "Hitting the memory wall: Implications of the obvious," *SIGARCH Comput. Archit. News*, vol. 23, no. 1, pp. 20–24, Mar. 1995.
- [78] G. Hendry, E. Robinson, V. Gleyzer, J. Chan, L. Carloni, N. Bliss, and K. Bergman, "Circuit-switched memory access in photonic interconnection networks for high-performance embedded computing," New Orleans, LA, Nov. 2010, pp. 1–12.
- [79] K. Therdsteerakudki, G.-S. Byun, J. Ir, G. Reinman, J. Cong, and M. Chang, "The dimm tree architecture: A high bandwidth and scalable memory system," in *IEEE 29th Int. Conf. Comput. Design (ICCD)*, Oct. 2011, pp. 388–395.
- [80] T. Kgil, S. D'Souza, A. Saidi, N. Binkert, R. Dreslinski, T. Mudge, S. Reinhardt, and K. Flautner, "Picoserver: Using 3D stacking technology to enable a compact energy efficient chip multiprocessor," in *Proc. 12th Int. Conf. Archit. Support Programm. Languages Operat. Syst.*, New York, 2006, pp. 117–128.
- [81] R. Drost, R. Hopkins, R. Ho, and I. Sutherland, "Proximity communication," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1529–1535, Sep. 2004.
- [82] D. Hopkins, A. Chow, R. Bosnyak, B. Coates, J. Ebergen, S. Fairbanks, J. Gainsley, R. Ho, J. Lexau, F. Liu, T. Ono, J. Schauer, I. Sutherland, and R. Drost, "Circuit techniques to enable 430 gb/s/mm² proximity communication," in *IEEE Int. Solid-State Circuits Conf.*, 2007, pp. 368–369.
- [83] R. Drost, H. R. , and S. I. , "Proximity communication," in *Proc. Custom IC Conf. (CICC)*, 2003, pp. 469–472.
- [84] D. Mizoguchi, Y. Yusof, N. Miura, T. Sakura, and T. Kuroda, "A 1.2 gb/s/pin wireless superconnect based on inductive inter-chip signaling (IIS)," in *IEEE Int. Solid-State Circuits Conf.. Dig. Tech. Papers*, Feb. 2004, vol. 1, pp. 142–517.
- [85] N. Miura, H. Ishikuro, K. Niitsu, T. Sakurai, and T. Kuroda, "A 0.14 pj/b inductive-coupling transeiver with digitally-controlled precise pulse shaping," *IEEE J. Solid-State Circuits*, vol. 43, no. 1, pp. 285–291, Jan. 2008.
- [86] N. Miura, M. Saito, and T. Kuroda, "A 1TB/s 1pJ/b 6.4mm²/TB/s QDR inductive-coupled interface between > 65nm CMOS logic and emulated 100 nm DRAM," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, to be published.
- [87] S. Beamer, C. Sun, Y.-J. Kwon, A. Joshi, C. Batten, V. Stojanović, and K. Asanovic, "Re-architecting dram memory systems with monolithically integrated silicon photonics," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, Saint-Malo, France, 2010, pp. 129–140.
- [88] A. N. Udipi, N. Muralimanohar, R. Balasubramonian, A. Davis, and N. P. Jouppi, "Combining memory and a controller with photonics through 3D-stacking to enable scalable and energy-efficient systems," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, San Jose, CA, 2011, pp. 425–436.
- [89] D. J. Sorin, M. D. Hill, and D. A. Wood, *A Primer on Memory Consistency and Cache Coherence*, ser. Synthesis Lectures on Computer Architecture. San Rafael, CA: Morgan Claypool, 2011.
- [90] N. Easley, L.-S. Peh, and L. Shang, "In-network cache coherence," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, Orlando, FL, 2006, pp. 321–332.
- [91] N. Agarwal, L.-S. Peh, and N. K. Jha, "In-network coherence filtering: Snoopy coherence without broadcasts," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, New York, 2009, pp. 232–243.
- [92] D. Vantrease, M. H. Lipasti, and N. Binkert, "Atomic coherence: Leveraging nanophotonics to build race-free cache coherence protocols," in *Int. Symp. High-Performance Comput. Archit. (HPCA)*, Feb. 2011, pp. 132–143.
- [93] N. Binkert, A. Davis, M. H. Lipasti, R. Schreiber, and D. Vantrease, "Nanophotonic barriers," presented at the Workshop Photon. Interconnects Comput. Architect. Held Conjunction With 42nd Annu. IEEE/ACM Int. Symp. Microarchit. (MICRO-42), New York, Dec. 2009.
- [94] J. Oh, M. Prvulovic, and A. G. Zajic, "Tlsync: Support for multiple fast barriers using on-chip transmission lines," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, San Jose, CA, 2011, pp. 105–116.
- [95] S. Mysore, B. Agrawal, N. Srivastava, S.-C. Lin, K. Banerjee, and T. Sherwood, "Introspective 3D chips," in *Proc. 12th Int. Conf. Archit. Support Programm. Languages Operat. Syst.*, San Jose, CA, 2006, pp. 264–273.
- [96] C. Nitta, M. Farrens, and V. Akella, "Dcaf—A directly connected arbitration-free photonic crossbar for energy-efficient high performance computing," in *Proc. Int. Symp. Parallel Distributed Process. (IPDPS)*, Shanghai, China, May 2012.
- [97] G. Maire, L. Vivien, G. Sattler, A. Kazmierczak, B. Sanchez, K. B. Gylfason, A. Griol, D. Marris-Morini, E. Cassan, D. Giannone, H. Sohlström, and D. Hill, "High efficiency silicon nitride surface grating couplers," *Opt. Exp.*, vol. 16, no. 1, pp. 328–333, Jan. 2008.
- [98] J. Lee, M. Ng, and K. Asanovic, "Globally-synchronized frames for guaranteed quality-of-service in on-chip networks," in *Proc. Int. Symp. Comput. Archit. (ISCA)*, Beijing, China, 2008, pp. 89–100.
- [99] B. Grot, S. W. Keckler, and O. Mutlu, "Preemptive virtual clock: A flexible, efficient, and cost-effective QOS scheme for networks-on-chip," in *IEEE/ACM Int. Symp. Microarchit. (MICRO)*, New York, 2009, pp. 163–174.

- [100] J. Ouyang and Y. Xie, "LOFT: A high performance network-on-chip providing quality-of-service support," in *Proc. 2010 43rd Annu. IEEE/ACM Int. Symp. Microarchit.*, 2010, pp. 409–420.
- [101] Y. Pan, J. Kim, and G. Memik, "Featherweight: Low-cost optical arbitration with QOS support," presented at the IEEE/ACM Int. Symp. Microarchit. (MICRO), Porte Alegre, Brazil, Dec. 2011.
- [102] J. Ouyang and Y. Xie, "Enabling quality-of-service in nanophotonic network-on-chip," in *Proc. 16th Asia South Pacific Design Automat. Conf.*, 2011, pp. 351–356.



John Kim (M'09) received the B.S. and M.Eng. degree in electrical engineering from Cornell University, Ithaca, NY, in 1997 and 1998, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2008.

He is currently an Assistant Professor in the Department of Computer Science at the Korea Advanced Institute of Science and Technology (KAIST), Seoul, Korea with joint appointment in the Web Science and Technology Division at KAIST. He spent several years working on the design of

different microprocessors at Motorola and Intel. His research interests includes multicore architecture, interconnection networks, and datacenter architecture.

Dr. Kim is a member ACM.



Kiyong Choi (M'88–SM'08) received the B.S. degree in electronics engineering from Seoul National University, Seoul, Korea, in 1978, the M.S. degree in electrical and electronics engineering from Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1980, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1989.

From 1989 to 1991, he was with Cadence Design Systems, Inc. In 1991, he joined the faculty of the School of Electrical Engineering and Computer Science, Seoul National University. His primary research interests include various aspects of computer-aided electronic systems design including embedded systems design, high-level synthesis, and low-power systems design. He is also interested in computer architecture and especially in configurable and reconfigurable architecture design. He was as an Associate Editor for the ACM *Transactions on Design Automation of Electronic Systems* (2003–2008). He was a Guest Editor for the ACM *Transactions on Embedded Computing Systems* (2008).

Dr. Choi is currently, he is on the editorial board of the IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS. He was Technical Program Co-Chair (2003) and General Co-Chair (2004) of ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED), Technical Program Co-Chair (2006) and General Co-Chair (2007) of the International Conference on Hardware/Software Codesign and System Synthesis CODES+ISSS, and Technical Program Vice Co-Chair (2007) and Technical Program Co-Chair (2008) of the Asia and South Pacific Design Automation Conference.



Gabriel Loh (M'05–SM'09) received the B.E. degree in electrical engineering from the Cooper Union, New York, and the M.S. and Ph.D. degrees in computer science from Yale University, New Haven, CT.

He is a Principal Researcher at Advanced Micro Devices (AMD). He was also a tenured Associate Professor in the College of Computing at the Georgia Institute of Technology, a visiting researcher at Microsoft Research, and a senior researcher at Intel Corporation. His research interests include computer architecture, processor microarchitecture, emerging

technologies and 3D die stacking.

Dr. Loh is a senior member of the ACM.