Columbia University   The Fu Foundation School of Engineering and Applied Science   Contact Us

Steven Nowick invited to present work on asynchronous on-chip networks at two national study groups

Like   Share   ⟨ 0 ⟩ | Share | | Tweet ⟨ 1 ⟩ | g+1 ⟨ 0 ⟩

Professor [Steven Nowick](#) was invited to join two leading national study groups sponsored by government agencies to define challenges and future research directions in two research frontier areas.

In March 2015, he attended the [NSF Workshop on Ultra-Low Latency Wireless Networks](#) in Phoenix, Arizona, a two-day national study group, on the future of wireless technology to achieve extreme low-latency communication, both for macro-level networks and micro-level on-chip networks. He participated in drafting the final white paper report, which documents the outcomes of the workshop and defines the challenges and opportunities of the area as well as the directions to achieve future breakthroughs. The report is expected to be used by the NSF for defining its future funding initiatives.

In August 2014, he participated in the [NSF/DARPA/DOE/NASA Workshop on System-on-Chip Design for High-Performance Computing ("SoC for HPC")](#) in Denver, Colorado, and gave an invited talk on his research on networks on chip (NoCs). This two-day national study group, sponsored jointly by NSF, DARPA, DOE, NASA, and Sandia and Lawrence Berkeley National Laboratories, was focused on the future of designing cost-effective high-performance parallel computers for both Big Data and consumer applications. The workshop had 35 attendees, with only 10 invited from academia. The outcome of this workshop is expected to shape the future landscape of high-performance computing.

Steven Nowick recently participated in two government-sponsored workshops on future technologies, one on wireless networks and one on high-performance parallel computing. While the two fields may seem to have little in common, the overlap is asynchronous communication, which enables components to operate independently of one another, making it easier to construct complex systems. For networks on chip, it enables more cost-effective design of larger chips; for high-performance parallel computing, it means using lower-priced, off-the-shelf components to achieve the same level of performance currently requiring expensive custom design.

Nowick's research includes a recent focus on *networks on chip* (NoCs), which replace the traditional bus with an on-chip network, allowing for ease of assembly by providing a standardized backbone that components can plug into and independently communicate with one another. As chips become larger and more complex, with more components—some chips today contain several billion transistors—on-chip networks allow for managing high-performance and reliable communication among many more components than would be possible with a traditional bus.

While still relatively new, the concept of on-chip networks is where chip design is headed, especially as demand grows for more energy-efficient and faster chips. Nowick is already looking at the next step in designing on-chip networks, particularly at the type of communications that occur over the network. Most networks on chips employ the synchronous method of communication, where a central clock coordinates the flow of data

About Steven Nowick

[Steven Nowick](#) is a professor of Computer Science and Electrical Engineering (and by courtesy, Electrical Engineering) at Columbia University, and the co-founder and former chair of the Computer Engineering Program.

His main research area is on designing methodologies and CAD tools for synthesis and optimization of asynchronous and mixed-timing (i.e., [GALS](#)) digital systems. Current projects include scalable networks on chip (NoCs) for shared-memory parallel processors and embedded systems, ultra-low energy digital systems, Internet-of-Things, and low-power and robust global communication.

He is an IEEE Fellow and a senior member of the ACM and has received numerous awards—among them, an Alfred P. Sloan Research Fellowship, NSF CAREER and RIA Awards, Columbia Engineering School Alumni Distinguished Faculty Teaching Awards—and his papers have won Best Paper Awards at the IEEE International Conference on Computer Design (1991, 2012) and the IEEE Async Symposium (2000). He holds 12 issued US patents.

Active in the engineering community, Nowick is a co-founder of the IEEE "Async" Symposia series, and served as its Program Committee Co-Chair and General Co-Chair. He was also Program Chair of the IEEE/ACM International Workshop on Logic and Synthesis (IWLS) and has served as a sub-committee/track chair for several leading international conference program committees: *ACM/IEEE Design Automation Conference*, *ACM/IEEE Design*, *Automation and Test in Europe*

and synchronizes the operations of multiple components, ensuring they communicate at the same frequency and in lockstep with one another. Synchronized circuits have worked well for so long because well-regulated chips are easy to design, test, and debug. But as chips continually get larger with increasing numbers and diversity of components, imposing unified synchronization is becoming unmanageable.

Instead Nowick and others are investigating *asynchronous communication* to dispense with the central fixed-rate clock and allow components to communicate with one another as needed and at their own rate. Removing the synchronous requirement makes it easier to again expand the number of components and build more even more energy-efficient and high-performance parallel computers. Already some leading companies are going in this direction (the recent IBM TrueNorth neuromorphic computer being a notable example [2014]).

Asynchronous communications has other major benefits besides. It saves power and energy since components, when not busy, are not activated at every clock cycle to remain synchronized.

Because of his research into asynchronous communication, Nowick was invited to participate in two government-sponsored workshops, one for wireless networks and one for high-performance parallel computing. Each was a two-day event where approximately 35 leading experts from industry and academia were invited to meet and help define the challenges and opportunities within the specific area. The ultimate goal was to help guide future funding and research initiatives for US government agencies.

Asynchronous communication for networks

The wireless workshop, held this past March, was sponsored by the National Science Foundation (NSF) and was entitled Workshop on Ultra-Low Latency Wireless Networks. While the NSF has previously hosted workshops on wireless networks, this was the first time such a workshop explored both macro-level networks (such as Ethernet) and micro-level on-chip networks, i.e, networks on chips. The idea was that those working in the relatively new area of on-chip networks could learn from those working on macro-level networks, and vice versa.

One potential borrowing from macro-level networks is the use of antennas to enable wireless networks on chips. Putting micro-antennas on a chip may seem a wild idea now, but it has several obvious benefits: it reduces the amount of wiring, freeing up valuable real estate on the chip while removing a source of heat dissipation; at the same time, it avoids the problem of overloading wires during periods of heavy processing.

(Wireless communication via antennas will also work in 3D chips to enable communications between layers.)

Nowick was at the workshop to discuss his research on asynchronous communication and how it contributes to ease of assembling large networks on chips. His presentation, one of many directions explored by a cross section of industry and academia people invited to the workshop, grew out of his work designing on-chip networks, but the concept of allowing components to operate independently has application for almost any complex system. In fact, he had made a similar presentation for an entirely different audience at a previous workshop not long before.

Asynchronous processing for high-performance computing

In August 2014, Nowick attended The System-on-Chip Design for High-Performance Computing workshop, which was sponsored jointly by the NSF, DARPA, DOE, NASA, and Sandia and Lawrence Berkeley National Laboratories, all of which are faced with analyzing massive amounts of scientific data, from astronomy data to nuclear, weather, social networks, and oceanography data. Processing the amount of data seen by these agencies is possible only through continued advances in high-performance computing and being able to more efficiently parallelize tasks among many processing clusters. Asynchronous communication has obvious application in handling the complexity of how data moves among multiple clusters.

Conference, and *IEEE International Conference in Computer Design*.

He is currently an associate editor of *IEEE Design & Test Magazine*, *IEEE Transactions on VLSI Systems*, and *ACM Journal on Emerging Technologies in Computer Systems*. He was also a guest co-editor of the *Proceedings of the IEEE* (vol. 87:2, Feb. 1999).

Nowick received his PhD in Computer Science from Stanford University in 1993, and his BA from Yale University.

But while people know how to build massively parallel computers, today it takes expensive custom design and special tools. Because very few companies have data at such high scale, the whole domain of extremely intensive parallel computing is currently low-volume and can't take advantage of the economies of scale seen in consumer electronics where costs can be amortized over high volume.

One purpose of the workshop, which covered multiple topics and was attended by Qualcomm's VP of Technology, Intel's director of future processor development, and NASA/JPL's manager of autonomous systems and flight computing, was to investigate the possibility of using low-cost commodity components (e.g., processors, memory, accelerators, and multimedia processing) while also borrowing the standardized workflow used in consumer electronics.

Borrowing the methods and tools used in consumer electronics would help rein in costs for high-performance computing, but it will not be easy. The whole multibillion-dollar chip industry is built on synchronous communication. CAD systems and other tools are designed for synchronous systems, students are trained on these tools, and those who design the tools are used to working with synchronous systems. Shifting to asynchronous systems will require a different mindset and a considerable financial commitment by companies that have spent years building and tweaking workflows built around synchronous communication.

Nowick's presentation on asynchronous communications included a possible compromise that combines the synchronous and asynchronous models into a single system, one that uses standard synchronous components connected through an asynchronous network capable of integrating components operating at different clock rates. This hybrid method, termed *globally asynchronous, locally synchronous systems* (GALS), has the benefits of asynchronous communication—scalability, modularity, low energy, and ease of assembly—with the advantage of the commodity pricing that exists on the consumer, synchronous side.

Though the two workshops covered different domains, the issue of increasing complexity is the same, and Nowick's message resonated at both; asynchronous communication, by allowing very different components to be connected together, provides scalability and ease of assembly while managing the increasing complexity. It's what's needed both for engineering bigger chips and for cost-effectively assembling high-performance and low-power computer systems.

Posted 7/27/2015

Like   Share   〈 0 | Share | **Tweet** 〈 1   g+1 〈 0