Research Directions for 21st Century Computer Systems

ASPLOS 2013 Panel

Mark D. Hill

Computer Sciences Department
University of Wisconsin-Madison
markhill@cs.wisc.edu

Abstract

Four recent efforts call out architectural challenges and opportunities up and down the software/hardware stack. This panel will discuss, *"What should the community do to facilitate, transcend, or refute these partially overlapping visions?"* The panel is chaired by Mark D. Hill with other panel members not finalized for the ASPLOS'13 proceedings.

Categories and Subject Descriptors C. [Computer Systems Organization], D. [Software].

General Terms Algorithms, Measurement, Performance, Design, Economics, Reliability, Security, Languages, Verification.

Keywords computer systems, architecture, programming methods, performance, energy, new technology.

1. Four Studies

This panel will discuss four recent studies.

1.1 NAS Report "The Future of Computing Performance: Game Over or Next Level?"

Fuller led software and hardware researchers Barroso, Colwell, Dally, Dobberpuhl, Dubey, Hill, Horowitz, Kirk, Lam, McKinley, Moore, and Yelick to a 2011 report [1]. It found that: (a) Computer hardware has transitioned to multicore, (b) Dennard scaling of CMOS has broken down, (c) Parallelism and locality must be exploited by software, and (d) Chip power will soon limit multicore scaling. It recommended enhanced research investment in:

• Algorithms to exploit parallel processing,
• Programming methods to enable efficient use of parallel systems,
• Long-term efforts on rethinking of the canonical computing "stack,"
• Parallel architectures driven by applications,
• Making computer systems more power efficient,
• Cooperation & innovation of open interfaces for parallel programming,
• Tools and methods to transform legacy apps to parallel systems, and
• Increased emphasis on parallelism in computer science education.

1.2 CCC Workshops "Advancing Computer Architecture Research (ACAR)"

Oskin and Torrellas led two two-day 2010 workshops with dozen participants. *ACAR-1 Failure is not an Option: Popular Parallel Programming* [2] advocated research advancing:

• Data centers and large-scale systems,
• Architectures to enhance programmability,
• Hardware-software co-design and asymmetry, and
• Domain specific languages.

ACAR-2 Laying a New Foundation for IT: Computer Architecture for 2025 and Beyond [3] advocated:

• Revisiting the boundary between hardware and software,
• Adding new capabilities for computer architectures, especially for security and reliability, and
• Exploitation of emerging technologies.

1.3 ISAT Workshop "Advancing Computer Systems without Technology Progress"

Hill and Kozyrakis led 48 software and hardware researchers in a two-day 2012 workshop [4] whose premises were that (a) CMOS transistors will soon stop getting "better," (b) Post-CMOS technologies are not ready, and (c) Computer system superiority is central to US security, government, education, commerce, etc. The workshop participants were asked, "How to advance computer systems without (significant) technology progress?" and found:

• Even with stalling technology progress, two decades of Moore's Law-like performance per unit energy gains (~1000x) are possible
• By wringing out inefficiencies used to harvest Moore's Law,
• Including via (a) hardware and software specialization and co-design, (b) reducing software bloat, and (c) ex-ploitation of lower-precision or approximate computing.

ASPLOS'13, March 16–20, 2013, Houston, Texas, USA.
ACM 978-1-4503-1870-9/13/03.

1.4 CCC White Paper "21st Century Computer Architecture"

Hill led Adve, Ceze, Irwin, Kaeli, Martonosi, Torrellas, Wenisch, Wood, Yelick and others in a 2012 community white paper [5]. It found that: (a) Semiconductor technology and computer architecture (writ broadly) were key—intentionally invisible—enablers for past information and communication technology (ICT) innovations. (b) Semiconductor technology is facing significant challenges, as new smaller transistors are not necessarily "better" due to power, reliability, interconnect and one-time engineering costs and concerns. (c) Compelling future ICT visions exist—e.g., personalized medicine, terrorist network analysis, and telepresence—that often require processing "big data" using distributed designs, working within form-factor constraints, and reconciling rapid deployment with efficient operation. It recommended that computer systems architecture to enable 21st Century ICT innovation:

• *Architecture as Infrastructure: From Sensors to Clouds.* With computers as a pillar of societal infrastructure, research must examine computation and communication in context—from sensors to data centers—considering human-centric design goals including improved programmability, privacy, security, and availability.

• *Energy First.* More energy-efficient systems—to make up for stalled transistor energy improvement—will re-quire 100x parallelism, specialization (beyond GPUs), and cross-layer design to manage communication even with big data.

• *New Technologies.* Emerging technologies provide challenges and opportunities: Flash and phase change memory for storage, photonic and 3D chip stacking for communication, and near-threshold CMOS and exotic technologies in computation.

• *Cross-cutting.* ICT researchers must develop richer inter-layer interfaces to empower change via further ab-stracting parallelism, heterogeneity, communication, security, reliability, etc.

References

[1] S. Fuller and L. Millett, "The Future of Computing Performance: Game Over or Next Level?," The National Academy Press, 2011 (http://books.nap.edu/openbook.php?record_id=12980&page=R1).

[2] J. Torrellas and M. Oskin, "Failure is not an Option: Popular Parallel Programming," Workshop on Advancing Computer Architecture Research, August 2010 (http://www.cra.org/ccc/docs/ACAR_Report_Popular-Parallel-Programming.pdf).

[3] M. Oskin and J. Torrellas, "Laying a New Foundation for IT: Computer Architecture for 2025 and Beyond," Workshop on Advancing Computer Architecture Research, September 2010 (http://www.cra.org/ccc/docs/ACAR2-Report.pdf).

[4] M. Hill and C. Kozyrakis, "Advancing Computer Systems without Technology Progress," ISAT Outbrief, April 17-18, 2012, of DARPA/ISAT Workshop, March 26-27, 2012 (http://www.cs.wisc.edu/~markhill/papers/isat2012_ACSWTP.pdf).

[5] A community white paper, "21st Century Computer Architecture," Computing Community Consortium, May 25, 2012 (http://cra.org/ccc/docs/init/21stcenturyarchitecturewhitepaper.pdf).