

TOP PICKS FROM THE COMPUTER ARCHITECTURE CONFERENCES OF 2009



Trevor Mudge
University of Michigan

..... This special issue is the seventh in an important tradition in the computer architecture community: *IEEE Micro's* Top Picks from the Computer Architecture Conferences. This tradition, started by *IEEE Micro's* then editor in chief, Pradip Bose, provides a means for sharing a sample of the best papers published in computer architecture during the past year with the *IEEE Micro* readership and researchers in the computer architecture community.

As might be expected, it is difficult to select only 13 from the many high-quality papers that have already been distinguished by their publication in the proceedings of our field's top conferences. This year was no different. In any review process there is a degree of uncertainty, but, that aside, this selection of articles is representative of the leading research in the computer architecture field, and the articles in Top Picks remain one of the best introductions to the current state of research in the field.

The review process

For each submission, we requested, in addition to a copy of the conference paper, a three-page summary that highlighted the work's novelty and argued its relevance for architects and designers of current- and future-generation computing systems. We did not limit from which conferences the papers came, but the

overwhelming majority came from the International Symposium on Computer Architecture (ISCA), the International Symposium on Microarchitecture (MICRO), the International Symposium on High-Performance Computer Architecture (HPCA), and the International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS). We received 91 submissions, which is a notable increase over recent years and a sign that there continues to be a great deal of activity in this field.

To review the submissions, we assembled the following program committee of 31 highly respected computer architects from both industry and academia:

- Murali Annavaram, University of Southern California
- Doug Burger, Microsoft Research
- John Carter, IBM
- Luis Ceze, University of Washington
- Nathan Clark, Georgia Institute of Technology
- Tom Conte, Georgia Institute of Technology
- Pradeep Dubey, Intel
- Lieven Eeckhout, Ghent University
- Joel Emer, Intel
- Krisztián Flautner, ARM
- Dirk Grunwald, University of Colorado
- Mark Hill, University of Wisconsin

Top picks of 2009

Caching, coherence, and multiprocessing

- “iCFP: Tolerating All-Level Cache Misses in In-Order Processors,” by Andrew Hilton, Santosh Nagarakatte, and Amir Roth
- “Near-Optimal Cache Block Placement with Reactive Nonuniform Cache Architectures,” by Nikos Hardavellas, Michael Ferdman, Babak Falsafi, and Anastasia Ailamaki
- “A Task-Centric Memory Model for Scalable Accelerator Architectures,” by John H. Kelm, Daniel R. Johnson, Steven S. Lumetta, Sanjay J. Patel, and Matthew I. Frank
- “DMP: Deterministic Shared-Memory Multiprocessing,” by Joseph Devietti, Brandon Lucia, Luis Ceze, and Mark Oskin

Practical prefetching

- “Making Address-Correlated Prefetching Practical,” by Thomas F. Wenisch, Michael Ferdman, Anastasia Ailamaki, Babak Falsafi, and Andreas Moshovos

Accelerating critical sections

- “Accelerating Critical Section Execution with Asymmetric Multi-core Architectures,” by M. Aater Suleman, Onur Mutlu, Moinuddin K. Qureshi, and Yale N. Patt

Performance evaluation

- “Per-Thread Cycle Accounting,” by Stijn Eyerman and Lieven Eeckhout

Application domain specific computing

- “AnySP: Anytime Anywhere Anyway Signal Processing,” by Mark Woh, Sangwon Seo, Scott Mahlke, Trevor Mudge, Chaitali Chakrabarti, and Krisztián Flautner

Security

- “Gate-Level Information-Flow Tracking for Secure Architectures,” by Mohit Tiwari, Xun Li, Hassan M.G. Wassel, Bitu Mazloom, Shashidhar Mysore, Frederic T. Chong, and Timothy Sherwood

Controlling voltage variation

- “Predicting Voltage Droops Using Recurring Program and Micro-architectural Event Activity,” Vijay Janapa Reddi, Meeta Gupta, Glenn Holloway, Michael D. Smith, Gu-Yeon Wei, and David Brooks

Sensor architectures

- “Architectural Implications of Nanoscale-Integrated Sensing and Computing,” by Constantin Pistol, Wutichai Chongchitmate, Christopher Dwyer, and Alvin R. Lebeck

Nonvolatile memory

- “Gordon: An Improved Architecture for Data-Intensive Applications,” by Adrian M. Caulfield, Laura M. Grupp, and Steven Swanson
- “Phase-Change Technology and the Future of Main Memory,” by Benjamin C. Lee, Ping Zhou, Jun Yang, Youtao Zhang, Bo Zhao, Engin Ipek, Onur Mutlu, and Doug Burger

- Mary Jane Irwin, Pennsylvania State University
- Lizy John, University of Texas
- Alvin Lebeck, Duke University
- Hsien-Hsin Lee, Georgia Institute of Technology
- Charles Lefurgy, IBM
- Mikko Lipasti, University of Wisconsin
- José Martínez, Cornell University
- Andreas Moshovos, University of Toronto
- Trevor Mudge, University of Michigan
- Onur Mutlu, Carnegie Mellon University
- Ravi Nair, IBM
- Sanjay Patel, University of Illinois
- Yale Patt, University of Texas
- Steven K. Reinhardt, AMD
- Eric Rotenberg, North Carolina State University
- Valentina Salapura, IBM
- André Seznec, INRIA/IRISA

- Josep Torrellas, University of Illinois
- Tom Wenisch, University of Michigan

Each program committee member was responsible for reviewing 12–16 papers. This was a considerable amount of work, and I would like to go on record with my thanks for their valuable contributions. Alvin Lebeck handled separately those papers in which I was a coauthor. This year we revived the idea of having the program committee meet as a group to discuss the reviews and make the final selections. More than 75 percent of the committee members attended the daylong meeting on October 17, 2009.

The articles

The first three articles examine various aspects of caching, coherence, and multiprocessing. In “iCFP: Tolerating All-Level Cache Misses in In-Order Processors,” Hilton et al. (see the sidebar for a full listing of authors)

propose a set of simple extensions to an in-order processor that allows it to tolerate load misses at any cache level. In "Near-Optimal Cache Block Placement with Reactive Nonuniform Cache Architectures," Hardavellas et al. achieve near-optimal cache block placement by classifying blocks online and placing data close to the cores that use them. Finally, in "A Task-Centric Memory Model for Scalable Accelerator Architectures," Kelm et al. provide a protocol that allows highly parallel chip multiprocessors with caches to support a coherent single address space view of memory without requiring coherence in hardware.

The next article also concerns multiprocessors. In "DMP: Deterministic Shared-Memory Multiprocessing," Devietti et al. describe a multiprocessor architecture that removes the nondeterminism from shared memory parallel programming, while providing performance competitive with today's nondeterministic multiprocessor designs.

The topic of prefetching is covered in the next article. Despite a decade of research demonstrating its efficacy, address-correlated prefetching has remained impractical and has not been implemented in a shipping processor. In "Making Address-Correlated Prefetching Practical," Wenisch et al. introduce new mechanisms for a practical storage-, latency-, and bandwidth-efficient design.

Thread serialization continues to be a central challenge in parallel processing, limiting parallel application scalability and performance. In "Accelerating Critical Section Execution with Asymmetric Multicore Architectures," Suleman et al. introduce a new execution paradigm and architecture that reduces thread serialization due to critical sections by executing critical sections on the faster cores in an asymmetric multicore architecture.

Evaluation of the effects of multithreading continues to be a thorny problem. In "Per-Thread Cycle Accounting," Eyerhan and Eeckhout present a novel performance counter architecture that monitors the progress of individual threads on a multithreaded processor by quantifying the performance impact of sharing resources.

Application-domain-specific processing is one of the areas in which parallelism has the most immediate payoff. In "AnySP: Anytime Anywhere Anyway Signal Processing," Woh et al. demonstrate that a low-power programmable convergent architecture can satisfy the requirements of both 4G wireless and video decoding by employing parallelism.

Tiwari et al. present their work in hardware support for security in "Gate-Level Information-Flow Tracking for Secure Architectures." They describe a new method for building secure systems in which all information flows within a machine can be accounted for, starting from a simple logic gate and building up to an entire microprocessor.

Supply voltage fluctuations within a microprocessor are becoming larger as technology nodes scale toward smaller feature sizes. This scaling requires advances in voltage noise tolerance. In "Predicting Voltage Droops Using Recurring Program and Microarchitectural Event Activity," Reddi et al. propose a voltage emergency predictor that learns to anticipate dangerous voltage fluctuations based on prior activity, so that recurring large voltage fluctuations can be smoothed away effectively.

In "Architectural Implications of Nanoscale-Integrated Sensing and Computing," Pistol et al. present a less conventional theme. They explore nanoscale sensor processors that integrate computation and sensing to support applications that seek to understand molecular-scale phenomena.

Finally, the last two articles explore the use of nonvolatile memories, specifically flash and phase change memory (PCM). Flash has expanded from MP3 players and memory sticks to disk replacement in the form of solid-state drives (SSDs). Its pending widespread use promises to be one of the most important recent developments in computer systems and architecture. In "Gordon: An Improved Architecture for Data-Intensive Applications," Caulfield et al. create fast, reenergy-efficient data-centric computing by tightly integrating low-power computation with specialized flash memory storage arrays. Moreover, there are technologies on the horizon that provide the nonvolatility of flash

but offer more attractive scalability, performance, power, and endurance. These advantages could enable nonvolatility not only in storage, but also in main memory. The last article combines two papers on PCM. In "Phase Change Technology and the Future of Main Memory," Lee et al. make a first attempt to embrace scalable, nonvolatile PCM technology into main memory, proposing architectural designs that make PCM competitive with DRAM in performance, power, and lifetime.

We hope you enjoy reading these articles. Don't forget that they are abridged versions, so we encourage you to also go back and read the originals of those articles you find most interesting. MICRO

Acknowledgments

In addition to the program committee members, I would like to thank Geoff Blake who managed the submission/reviewing website. I would also like to thank Ron Dreslinski who, with Geoff, provided

support on the day of the program committee meeting—Geoff and Ron are members of my group at Michigan.

Dave Albonesi, our editor in chief, has been a constant help—thanks Dave. Last, but by no means least, I would like to thank *IEEE Micro* staff: Joan Taylor, Robin Baldwin, Debby Mosher, Alkenia Winston, and Ed Zintel.

Trevor Mudge is the first Bredt Family Professor of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor. His research interests include computer architecture, CAD, and compilers. Mudge has a PhD in computer science from the University of Illinois, Urbana-Champaign. He is a member of the ACM, Institution of Engineering and Technology, and British Computer Society, and is a Fellow of the IEEE.



Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.

ADVERTISER INFORMATION

JANUARY/FEBRUARY 2010 • IEEE MICRO

Advertising Personnel

Marion Delaney
IEEE Media, Advertising Dir.
Phone: +1 415 863 4717
Email: md.ieeemedia@ieee.org

Marian Anderson
Sr. Advertising Coordinator
Phone: +1 714 821 8380
Fax: +1 714 821 4010
Email: manderson@computer.org

Sandy Brown
Sr. Business Development Mgr.
Phone: +1 714 821 8380
Fax: +1 714 821 4010
Email: sb.ieeemedia@ieee.org

Advertising Sales Representatives

Recruitment:

Mid Atlantic
Lisa Rinaldo
Phone: +1 732 772 0160
Fax: +1 732 772 0164
Email: lr.ieeemedia@ieee.org

New England
John Restchack
Phone: +1 212 419 7578
Fax: +1 212 419 7589
Email: j.restchack@ieee.org

Southeast
Thomas M. Flynn
Phone: +1 770 645 2944
Fax: +1 770 993 4423
Email: flyntom@mindspring.com

Midwest/Southwest
Darcy Giovingo
Phone: +1 847 498 4520
Fax: +1 847 498 5911
Email: dg.ieeemedia@ieee.org

Northwest/Southern CA
Tim Matteson
Phone: +1 310 836 4064
Fax: +1 310 836 4067
Email: tm.ieeemedia@ieee.org

Japan
Tim Matteson
Phone: +1 310 836 4064
Fax: +1 310 836 4067
Email: tm.ieeemedia@ieee.org

Europe
Heleen Vodegel
Phone: +44 1875 825700
Fax: +44 1875 825701
Email: impress@impressmedia.com

Product:

US East
Dawn Becker
Phone: +1 732 772 0160
Fax: +1 732 772 0164
Email: db.ieeemedia@ieee.org

US Central
Darcy Giovingo
Phone: +1 847 498 4520
Fax: +1 847 498 5911
Email: dg.ieeemedia@ieee.org

US West
Lynne Stickrod
Phone: +1 415 931 9782
Fax: +1 415 931 9782
Email: ls.ieeemedia@ieee.org

Europe
Sven Anacker
Phone: +49 202 27169 11
Fax: +49 202 27169 20
Email: sanacker@intermediapartners.de