

Guest Editors' Introduction

TOP PICKS FROM THE 2008 COMPUTER ARCHITECTURE CONFERENCES



..... We are pleased to introduce *IEEE Micro's* sixth annual Top Picks from Computer Architecture Conferences special issue. This issue has become an important tradition in the computer architecture community, and one of the most important forms of recognition of research excellence. It is also an opportunity for those who don't follow the research conferences closely to get a quick snapshot of some of the most influential research in this field all in one place.

Working with *IEEE Micro* Editor in Chief Dave Albonesei, we decided to make a subtle change to the selection criteria this year—we simply sought to identify high-impact publications. This was not so much intended to weaken the former emphasis on industry impact as much as to include publications that might have a deep impact on future research directions even when we don't expect to see an industrial implementation in exactly the form proposed.

Selecting the articles

Not surprisingly, choosing 12 articles from all of the computer architecture papers published this year was a difficult task. We received 76 submissions as a result of the call for papers. Each submission included both the original paper and a three-page summary that highlighted the work's novelty and potential impact.

The most important part of the process, from our perspective, was selecting a highly qualified program committee of experts from academia and industry to review and ultimately select the top papers. We greatly appreciated their hard work and dedication to doing things right. Committee members were:

- Sarita Adve, University of Illinois, Urbana-Champaign
- Luiz Barroso, Google
- Ras Bodik, University of California, Berkeley
- Pradip Bose, IBM
- David Brooks, Harvard University
- Doug Burger, Microsoft
- Derek Chiou, University of Texas at Austin
- Jose Duato, Universidad Politécnica de Valencia
- Sandhya Dwarkadas, University of Rochester
- Lieven Eeckhout, Ghent University
- Babak Falsafi, Ecole Polytechnique Fédérale de Lausanne
- Paolo Faraboschi, HP
- Sudhanva Gurusurthi, University of Virginia
- James Hoe, Carnegie Mellon University
- Ravi Iyer, Intel
- Steve Keckler, University of Texas at Austin

Joel Emer
Intel

Dean Tullsen
University of California,
San Diego

- Gabe Loh, Georgia Tech
- Scott Mahlke, University of Michigan
- Srilatha Manne, AMD
- Milo Martin, University of Pennsylvania
- Onur Mutlu, Microsoft
- Ravi Nair, IBM
- Li-Shiuan Peh, Princeton University
- Yiannakis Sazeides, University of Cyprus
- Jim Smith, University of Wisconsin, Madison, and Google
- Steven Swanson, University of California, San Diego
- Craig Zilles, University of Illinois, Urbana-Champaign

Five program committee members reviewed each paper before we met to vote on the top papers. This year, we experimented by conducting the meeting virtually over a conference call. To aid the process, we used some shared interactive documents through Google Docs to track meeting progress, manage the comings and goings of people with conflicts of interest with particular papers, and even to expedite the voting process.

We initially limited voting on each paper to committee members who had reviewed the paper. In cases where there wasn't an overwhelming consensus in that vote, we took a vote of the entire committee. Because of the difficulty in selecting only 12 papers, that process resulted in each paper receiving a designation of accept, reject, or accept-if-room. At the end of the meeting we employed an "instant runoff" vote to choose the final papers, as we were able to accommodate fewer than half of the last group. There is no question that a number of excellent articles were among those that could not be accommodated in this issue.

We handled conflicts of interest by requiring authors or those with professional connections to authors to leave the call for the discussion of the paper (either by setting the phone down or by hanging up, depending on the nature of the conflict). The outcome of program committee member papers was hidden until after the meeting.

The articles

Last year's ISCA conference is well represented among the selected articles, but we also have articles originally appearing

in Micro, ASPLOS, ICCD, and even Siggraph.

The articles appearing in this special issue highlight the following key trends in current computer architecture design and research:

- the increased importance of thread-level parallelism;
- the increasing complexity of software and the need for architectures that improve programmability, analyzability, and correctness;
- the increasing urgency of controlling chip-wide power; and
- the emerging problem of process variation.

The first five articles concern parallel architectures. The first article, by Seiler et al. (see the sidebar for the full listing of article titles and authors), describes Larrabee, a unique many-core processor for vision and graphic processing from Intel. It combines the high parallelism of GPUs with the programmability and generality of general-purpose processors. The second article, by Mutlu and Moscibroda, shows that we must reconsider the design of memory controllers in light of new highly parallel architectures. The authors account for interference between threads' access streams to preserve memory-level parallelism and fairness. The third article, by Kim et al., introduces a highly effective new interconnection network topology that exploits recent trends toward high-radix routers to reduce the network's interconnect and cabling costs. The fourth article, by Lim et al., examines the increasingly important area of large-scale warehouse computing environments. It shows that there are significant potential gains from designing the CPU and system specifically for such workloads. The final article in this group, by Gurumurthi, Sankar, and Stan, examines a different kind of parallelism. It proposes and describes novel changes to the architecture of disk drives that would allow multiple requests to be serviced simultaneously on the same device, allowing higher throughput without the cost of adding disks just for performance.

The next three articles recognize the increased difficulty of writing efficient and

GUEST EDITORS' INTRODUCTION

Top Picks of 2008

Parallelism

- "Larrabee: A Many-Core x86 Architecture for Visual Computing," by Larry Seiler, Doug Carmean, Eric Sprangle, Tom Forsyth, Michael Abrash, Pradeep Dubey, Stephen Junkins, Adam Lake, Jeremy Sugerman, Robert Cavin, Roger Espasa, Ed Grochowski, Toni Juan, and Pat Hanrahan
- "Parallelism-Aware Batch Scheduling: Enabling High-Performance and Fair Shared Memory Controllers," by Onur Mutlu and Thomas Moscibroda
- "Cost-Efficient Dragonfly Topology for Large-Scale Systems," by John Kim, William Dally, Steve Scott, and Dennis Abts
- "Server Designs for Warehouse-Computing Environments," by Kevin Lim, Parthasarathy Ranganathan, Jichuan Chang, Chandrakant Patel, Trevor Mudge, and Steven K. Reinhardt
- "Using Intradisk Parallelism to Build Energy-Efficient Storage Systems," by Sudhanva Gurumurthi, Mircea R. Stan, and Sriram Sankar

Architectural support for software systems

- "Flexible Hardware Acceleration for Instruction-Grain Lifeguards," by Shimin Chen, Michael Kozuch, Phillip B. Gibbons, Michael Ryan, Theodoros Strigkos, Todd C. Mowry, Olatunji,

Ruwase, Evangelos Vlachos, Babak Falsafi, and Vijaya Ramachandran

- "Atom-Aid: Detecting and Surviving Atomicity Violations," by Brandon Lucia, Joseph Devietti, Luis Ceze, and Karin Strauss
- "SoftSig: Software-Exposed Hardware Signatures for Code Analysis and Optimization," by James Tuck, Wonsun Ahn, Josep Torrellas, and Luis Ceze

Low power

- "Trading Off Cache Capacity for Low-Voltage Operation," by Chris Wilkerson, Hongliang Gao, Alaa R. Alameldeen, Zeshan Chishti, Muhammad Khellah, and Shih-Lien Lu
- "Mixed-Signal Approximate Computation: A Neural Predictor Case Study," by Renée St. Amant, Daniel A. Jiménez, and Doug Burger

Process variation

- "Temperature Variation Characterization and Thermal Management of Multicore Architectures," by Eren Kursun and Chen-Yong Cher
- "Revival: A Variation-Tolerant Architecture Using Voltage Interpolation and Variable Latency," by Xiaoyao Liang, Gu-Yeon Wei, and David Brooks

correct programs for today's complex parallel systems. The first of these articles, by Chen et al., presents an effective attack on the important problem of the slowdown associated with program monitors. The authors present a set of general mechanisms for accelerating program-tracking lifeguards, such as memory safety checking, taint checking, and race detection. The second article in this section, by Lucia et al., addresses the challenging issue of atomicity violations by implicitly creating atomicity blocks not specified by the programmer and thus reducing the probability of atomicity violations automatically. Finally, the article by Tuck et al. defines an elegant mechanism that gives programs the ability to dynamically create signatures for sets of addresses, and some operations with addresses and address signatures with the aim of providing efficient support for code introspection and optimization techniques.

The third set of articles seeks to reduce the power consumed by traditional architectural features. The first of these articles, by Wilkerson et al., considers the critical

problem of designing caches that allow for aggressive voltage scaling. The primary challenge here is to identify and account for regions of the cache that fail at low voltages. The second article in this group, by St. Amant, Jiménez, and Burger, present the novel idea of creating a branch predictor using a low-power analog circuit rather than a conventional digital design.

The final pair of articles addresses the problem of process variation, which is an issue of ever-increasing importance as semiconductor features continue to scale down. The first of these articles, by Kursun and Cher, provides deep insight into the interplay of variation, dynamic power, and thermal management by using thermal readings to infer on-chip variation and incorporating that information into the chip's thermal and power management. The second article, by Liang, Wei, and Brooks, shows impressive results in addressing the issue of process variability after fabrication by using two synergistic fine-grained tuning techniques: voltage interpolation and variable latency.

We hope you enjoy and learn from these articles. We also remind you that each article was abridged to fit in this issue, so we encourage you to also go back and read the originals of those articles you find most interesting. Please let us know if you have any feedback to improve future Top Picks issues. MICRO

Acknowledgments

This special issue of *IEEE Micro* wouldn't have been possible without the outstanding efforts of many people. University of California, San Diego, graduate student Leo Porter modified and managed the submission and review software, and put in countless hours to keep the entire process running smoothly. We would like to thank Dave Albonesi, editor in chief of *IEEE Micro*, for his ever-prompt support and sage advice. We also received much valuable advice from Pradip Bose, the initiator of the original Top Picks issue. We very much appreciated the efforts of the *IEEE Micro* staff, especially Robin Baldwin and Joan Taylor, who provided comprehensive and professional support. Finally, we particularly want to thank the program committee members for the time and effort they put into the review process and all of the authors for their excellent submissions.

Joel Emer is an Intel Fellow and director of microarchitecture research at Intel. His current research interests include memory hierarchy design, processor reliability, FPGA-based computation, and performance modeling. Emer has a PhD in electrical engineering from the University of Illinois at Urbana-Champaign. He is a Fellow of both the ACM and the IEEE.

Dean Tullsen is a professor in the Computer Science and Engineering Department at the University of California, San Diego. His research is on the architecture of high-performance processors, including multicore and multithreaded processors. Tullsen has a PhD in computer science and engineering from the University of Washington. He is a Fellow of the IEEE and a member of the ACM.

Direct questions and comments about this article to Dean Tullsen, CSE Dept., UCSD, 9500 Gilman Dr., La Jolla, CA 92093-0404; tullsen@cs.ucsd.edu.

For more information on this or any other computing topic, please visit our Digital Library at <http://computer.org/csdl>.

Call for Papers | General Interest

IEEE *Micro* seeks general-interest submissions for publication in upcoming issues. These works should discuss the design, performance, or application of microcomputer and microprocessor systems. Of special interest are articles on performance evaluation and workload characterization. Summaries of work in progress and descriptions of recently completed works are most welcome, as are tutorials. *Micro* does not accept previously published material.

Check our author center (www.computer.org/micro/author.htm) for word, figure, and reference limits. All submissions pass through peer review consistent with other professional-level technical publications, and editing for clarity, readability, and conciseness. Contact *IEEE Micro* at micro-ma@computer.org with any questions.