

The 47th Annual IEEE/ACM International Symposium on Microarchitecture

MICRO-47

13-17 December 2014, Cambridge, United Kingdom

























CONFERENCE INFORMATION

PAPERS BY SESSION

PAPERS BY AUTHOR

GETTING STARTED

TRADEMARKS

SEARCH

Published by



Proceedings

47th Annual IEEE/ACM International Symposium on Microarchitecture

MICRO 2014

Proceedings

47th Annual IEEE/ACM International Symposium on Microarchitecture

13–17 December 2014 Cambridge, United Kingdom





All rights reserved.

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries may photocopy beyond the limits of US copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Other copying, reprint, or republication requests should be addressed to: IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 133, Piscataway, NJ 08855-1331.

The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Computer Society, or the Institute of Electrical and Electronics Engineers, Inc.

IEEE Computer Society Order Number E5349 BMS Part Number CFP14071-ART ISBN 978-1-4799-6998-2 ISSN 1072-4451

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: +1 800 272 6657
Fax: +1 714 821 4641
http://computer.org/cspress
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
http://shop.ieee.org/store/
customer-service@ieee.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: +81 3 3408 3118
Fax: +81 3 3408 3553
tokyo.ofc@computer.org

Individual paper REPRINTS may be ordered at: <reprints@computer.org>

Editorial production by Lisa O'Conner Cover art production by Mark Bartosik





IEEE Computer Society
Conference Publishing Services (CPS)

http://www.computer.org/cps

2014 47th Annual IEEE/ACM International Symposium on Microarchitecture

MICRO 2014

Table of Contents

Message from the General Chair Message from the Program Co-Chairs Organizing Committee Program Committee External Reviewer Committee External Reviewers	X
	xi
	xiii
	xiv
	xvi
	xix
Keynote Abstracts	xx
Session 1A: Stacked DRAM	
CAMEO: A Two-Level Memory Organization with Capacity of Main Memory and Flexibility of Hardware-Managed Cache	1
Transparent Hardware Management of Stacked DRAM as Part of Memory	13
Unison Cache: A Scalable and Effective Die-Stacked DRAM Cache	25
Bi-Modal DRAM Cache: Improving Hit Rate, Hit Latency and Bandwidth	38
Citadel: Efficiently Protecting Stacked Memory from Large Granularity Failures Prashant J. Nair, David A. Roberts, and Moinuddin K. Qureshi	51
Session 1B: GPGPU and Data Parallel Architectures	
Locality-Aware Mapping of Nested Parallel Patterns on GPUs	63
Accelerating Irregular Algorithms on GPGPUs Using Fine-Grain Hardware Worklists Ji Yun Kim and Christopher Batten	75

Guoyang Chen, Bo Wu, Dong Li, and Xipeng Shen	88
Exploring the Design Space of SPMD Divergence Management on Data-Parallel	
Architectures	101
Yunsup Lee, Vinod Grover, Ronny Krashinsky, Mark Stephenson, Stephen W. Keckler, and Krste Asanović	
Managing GPU Concurrency in Heterogeneous Architectures	114
Onur Kayiran, Nachiappan Chidambaram Nachiappan, Adwait Jog,	
Rachata Ausavarungnirun, Mahmut T. Kandemir, Gabriel H. Loh, Onur Mutlu, and Chita R. Das	
Session 2A: Memory Systems, Scheduling, and Optimization	
Load Value Approximation	127
Joshua San Miguel, Mario Badr, and Natalie Enright Jerger	
Arbitrary Modulus Indexing	140
Jeffrey R. Diamond, Donald S. Fussell, and Stephen W. Keckler	
FIRM: Fair and High-Performance Memory Control for Persistent Memory Systems Jishen Zhao, Onur Mutlu, and Yuan Xie	153
Short-Circuiting Memory Traffic in Handheld Platforms	166
Praveen Yedlapalli, Nachiappan Chidambaram Nachiappan, Niranjan Soundararajan,	
Anand Sivasubramaniam, Mahmut T. Kandemir, and Chita R. Das	
Efficient Memory Virtualization: Reducing Dimensionality of Nested Page Walks	178
Jayneel Gandhi, Arkaprava Basu, Mark D. Hill, and Michael M. Swift	
Session 2B: Security	
Iso-X: A Flexible Architecture for Hardware-Managed Isolated Execution	190
Dmitry Evtyushkin, Jesse Elwell, Meltem Ozsoy, Dmitry Ponomarev, Nael Abu Ghazaleh, and Ryan Riley	
Random Fill Cache Architecture	203
Fangfei Liu and Ruby B. Lee	
CC-Hunter: Uncovering Covert Timing Channels on Shared Processor Hardware	216
Jie Chen and Guru Venkataramani	
Continuous, Low Overhead, Run-Time Validation of Program Executions	229
Erdem Aktas, Furat Afram, and Kanad Ghose	
A Practical Methodology for Measuring the Side-Channel Signal Available to	
the Attacker for Instruction-Level Events	242
Robert Callan, Alenka Zajić, and Milos Prvulovic	

Session 3A: Methodology, Modeling, and Tools	
RpStacks: Fast and Accurate Processor Design Space Exploration Using	
Representative Stall-Event Stacks	255
Jaewon Lee, Hanhwi Jang, and Jangwoo Kim	
GPUMech: GPU Performance Modeling Technique Based on Interval Analysis	268
Jen-Cheng Huang, Joo Hwan Lee, Hyesoon Kim, and Hsien-Hsin S. Lee	
PyMTL: A Unified Framework for Vertically Integrated Computer Architecture	
Research	280
Derek Lockhart, Gary Zibrat, and Christopher Batten	
Session 3B: Reliability and Fault Tolerance	
Calculating Architectural Vulnerability Factors for Spatial Multi-Bit Transient Faults	293
Mark Wilkening, Vilas Sridharan, Si Li, Fritz Previlon, Sudhanva Gurumurthi, and David R. Kaeli	
Using ECC Feedback to Guide Voltage Speculation in Low-Voltage Processors	306
Anys Bacha and Radu Teodorescu	
Harnessing Soft Computations for Low-Budget Fault Tolerance	319
Session 4A: TLB and Cache Optimization	
Skewed Compressed Caches	331
Somayeh Sardashti, André Seznec, and David A. Wood	
Adaptive Cache Management for Energy-Efficient GPU Computing	343
Xuhao Chen, Li-Wen Chang, Christopher I. Rodrigues, Jie Lv, Zhiying Wang, and Wen-Mei Hwu	
Futility Scaling: High-Associativity Cache Partitioning	356
Ruisheng Wang and Lizhong Chen	
Session 4B: Managing Voltage and Time	
Voltage Noise in Multi-Core Processors: Empirical Characterization and Optimization	
Opportunities	368
Ramon Bertran, Alper Buyuktosunoglu, Pradip Bose, Timothy J. Slegel, Gerard Salem,	
Sean Carey, Richard F. Rizzolo, and Thomas Strach	
Enabling Realistic Fine-Grain Voltage Scaling with Reconfigurable Power	
Distribution Networks	381
Waclaw Godycki, Christopher Torng, Ivan Bukreyev, Alyssa Apsel, and Christopher Batten	
Micro-Sliced Virtual Processors to Hide the Effect of Discontinuous CPU Availability	
for Consolidated Systems	394
Jeongseob Ahn, Chang Hyun Park, and Jaehyuk Huh	

Session 5A: Energy-Efficient Computation

SMiTe: Precise QoS Prediction on Real-System SMT Processors to Improve Utilization	
in Warehouse Scale Computers	406
Yunqi Zhang, Michael A. Laurenzano, Jason Mars, and Lingjia Tang	
A Front-End Execution Architecture for High Energy Efficiency	419
Ryota Shioya, Masahiro Goshima, and Hideki Ando	
Execution Drafting: Energy Efficiency through Computation Deduplication	432
Michael Mckeown, Jonathan Balkind, and David Wentzlaff	
PPEP: Online Performance, Power, and Energy Prediction Framework and DVFS	
Space Exploration	445
Bo Su, Junli Gu, Li Shen, Wei Huang, Joseph L. Greathouse, and Zhiying Wang	
Session 5B: Interconnects	
NoC Architectures for Silicon Interposer Systems: Why Pay for more Wires when	
you Can Get them (from your interposer) for Free?	458
Natalie Enright Jerger, Ajaykumar Kannan, Zimo Li, and Gabriel H. Loh	
Hi-Rise: A High-Radix Switch for 3D Integration with Single-Cycle Arbitration	471
Supreet Jeloka, Reetuparna Das, Ronald G. Dreslinski, Trevor Mudge, and David Blaauw	
Multi-GPU System Design with Memory Networks	484
Gwangsun Kim, Minseok Lee, Jiyun Jeong, and John Kim	
Dodec: Random-Link, Low-Radix On-Chip Networks	496
Haofan Yang, Jyoti Tripathi, Natalie Enright Jerger, and Dan Gibson	
Session 6A: Branch Prediction and Prefetching	
Wormhole: Wisely Predicting Multidimensional Branches	509
Jorge Albericio, Joshua San Miguel, Natalie Enright Jerger, and Andreas Moshovos	
Bias-Free Branch Predictor	521
Dibakar Gope and Mikko H. Lipasti	
Loop-Aware Memory Prefetching Using Code Block Working Sets	533
Adi Fuchs, Shie Mannor, Uri Weiser, and Yoav Etsion	
BuMP: Bulk Memory Access Prediction and Streaming	545
Stavros Volos Javier Picorel Bahak Falsafi and Boris Grot	

Session 6B: Compilation and Code Generation

Protean Code: Achieving Near-Free Online Code Transformations for Warehouse	
Scale Computers	558
Michael A. Laurenzano, Yunqi Zhang, Lingjia Tang, and Jason Mars	
Compiler Support for Optimizing Memory Bank-Level Parallelism	571
Architectural Specialization for Inter-Iteration Loop Dependence Patterns	583
Specializing Compiler Optimizations through Programmable Composition for Dense Matrix Computations	596
Session 7: Best Paper Nominees	
DaDianNao: A Machine-Learning Supercomputer	609
B-Fetch: Branch Prediction Directed Prefetching for Chip-Multiprocessors	623
Pipe Check: Specifying and Verifying Microarchitectural Enforcement of Memory	
Consistency Models Daniel Lustig, Michael Pellauer, and Margaret Martonosi	635
Equalizer: Dynamic Tuning of GPU Resources for Efficient Execution	647
COMP: Compiler Optimizations for Manycore Processors Linhai Song, Min Feng, Nishkam Ravi, Yi Yang, and Srimat Chakradhar	659
Author Index	672

Message from the General Chair

It is my pleasure to welcome all Micro47 delegates to Cambridge in the United Kingdom, which, as of this proceedings, still includes Scotland. I also have the honor of organizing Micro in The UK for the first time. Cambridge is a relatively small town located 50 miles north of London, and boasts of having the World's third oldest university along with a large concentration of high-tech companies in computers, electronics and biotechnology. The university has a great number of colleges (Kings, Trinity, St John etc.) in walking distance from the conference venue. In fact Micro47 will be held in one of these colleges: Robinson College. Cambridge offers visitors historic buildings and colleges, punting on the river Cam, a great variety of restaurants and the famous "real ale" in traditional English pubs.

On the technical side, Micro47 has a stimulating program with 53 full papers. I would like to thank the Program Co-chairs Thomas Wenisch and Emre Ozer for planning, organising and scheduling an exciting Micro47 technical program. This year, we have three outstanding invited speakers: Mike Muller, CTO of ARM, Trevor Mudge, University of Michigan and James Smith, University of Wisconsin-Madison. We also have a total of 7 workshops and tutorials scheduled on Sat and Sun. I thank the Workshops and Tutorials Chair Jason Mars for his diligent work.

On the social side, we planned a guided tour of Cambridge followed by punting on the river. You will sip your mulled wine under the blanket on the boat, and enjoy the great atmosphere of the ancient Cambridge colleges. As for the stories that the punt-guides will tell: feel free to be skeptical but some may actually have a grain of truth buried somewhere deep. This outing will be followed by a hot-topic debate in Cambridge Union Debating Chamber, using a traditional format. Finally, the day will be finished with a lavish dinner in Kings College.

I would like to thank the organising committee: Registration Chair Tim Jones for undertaking a great responsibility for handling the conference registration and accounts, Publication Chair Mike Ferdman for preparing the conference proceedings, Finance Chair Bruno Jansen for planning the conference budget, Publicity Chair Ron Dreslinski for reaching wide audience to announce Micro47, Submission Chair Anthony Gutierrez for managing the paper submission process, and Web Chair Marios Kleanthous for maintaining the conference website. Special thanks to Emre Ozer for shouldering most of the burden of keeping things on track day to day on my behalf.

I also thank the Micro Steering Committee Chair Rich Belgard and the entire Steering Committee for their guidance and advice.

Finally, I would like to express my gratitude to the corporate sponsors for their generous donations, and our society sponsors IEEE TC-uARCH and ACM SIGMICRO.

I am looking forward to seeing you in Cambridge!

Krisztian Flautner

The General Chair, MICRO-47

Message from the MICRO-47 Program Co-Chairs

Welcome to Cambridge and to MICRO-47! On behalf of the Program Committee, we are delighted to present you the technical program for MICRO-47.

The program includes 53 papers selected from 279 submissions. The program committee (PC) of 42 distinguished experts used a two-round review process. In the first round, all papers received 2 reviews from PC members and one review from an external expert. The 202 papers with at least one review with a positive score for overall merit were promoted to the second round. This year, for the first time, we notified authors of papers that did not advance to the second round shortly after these decisions were made, to provide more timely feedback. In the second round, papers received at least one review from a PC member and one review from an external expert, for a minimum total of 5 reviews per paper. Overall, 200 reviewers submitted 1,646 paper reviews.

The PC Co-Chairs personally selected all PC and external reviewers for each paper after careful examination of the papers. Papers for which only one of the Co-Chairs had a conflict-of-interest were handled by the other Co-Chair. For the 10 papers where both of us had conflicts, Milos Prvulovic made all reviewer selections and led any discussion at the PC meeting.

At the conclusion of the review process, authors were given the opportunity to respond to the review comments and scores via the rebuttal process. We made the author rebuttals available to all reviewers for a 10-day online discussion. During the online discussion, PC and external reviewers were required to enter a second "Post-Discussion" merit score and reach a preliminary recommendation to accept, reject, or discuss the paper further at the meeting. The online discussions lead to 17 accept recommendations, 92 reject recommendations, and 93 discuss recommendations. The "Post-Discussion" merit scores were used to design the PC meeting agenda. A total of 1,851 comments were entered during the online discussion.

The Program Committee met on August 29, 2014 at the Westin at Detroit Metro Airport. The PC considered the papers with "discuss" recommendations and discussed 63 of them in detail during the PC meeting. Where possible, decisions were made by discussing a paper with the full PC until a unanimous consensus was reached among the assigned reviewers. PC members with a conflict of interest on a paper were excused from the discussion. If unanimous consensus could not be reached, all non-conflicted PC members were asked to vote on the paper outcome. A small number of papers were conditionally accepted with shepherds to help address specific shortcomings in otherwise excellent pieces of research.

In an effort to improve transparency and capture key feedback from the online and PC meeting discussion, a brief "post-discussion summary" comment to the authors was entered for all papers considered in the online and PC discussion process, in addition to the reviews and scores. We hope that this information will help all authors improve their research, regardless of the outcome for their submission.

The final program includes papers that represent all aspects of our field: branch prediction and prefetching, stacked DRAM, reliability and fault tolerance, security, interconnects, compilation and code generation, energy-efficient computation, memory scheduling and optimization, methodology and modeling tools, TLB and cache optimizations, GPGPU and data parallel architectures, virtualization, and power management. In addition to the paper presentations, the program includes a

single-track session with brief presentations on the key idea in every paper (lightning talks) and a poster session. We hope that these two sessions will enable attendees to get exposure to the whole program and delve into the detailed contributions of the papers with their authors during the conference.

We are also particularly excited about the three keynote talks. The first keynote is by Mike Muller, Chief Technology Officer of ARM, Ltd. The second keynote is by Prof. Trevor Mudge, from the University of Michigan, and will address past and present predictions of the "End of Moore's Law". Finally, Prof. James Smith of the University of Wisconsin will discuss ongoing investigation of the brain's computational paradigm and how it relates to computer architecture research.

Finally, the program includes a moderated debate entitled "End of the Road for von Neumann Architecture?" which will be held in the traditional English debating chamber at the storied Cambridge Union. We also invite attendees to a social program including the tradition of "Punting on the Cam".

The MICRO-47 program is the result of the hard work of the PC members that sacrificed a significant portion of their free time this summer to provide service to our community. We are grateful for their professionalism in all aspects of their duties. We would also like to thank the members of the External Review Committee and additional external reviewers for their thorough reviews, often at a very short notice. We are particularly thankful to the PC members that diligently shepherded papers, working closely with the authors for weeks past the program selection. We would also like to recognize the PC members that volunteered to be session chairs during the conference.

Special thanks to Milos Prvulovic who handled all aspects of submissions for which the Co-chairs had conflicts. Finally, we are grateful to Anthony Gutierrez, our submission chair, for providing both authors and reviewers with a very smooth experience. He also handled preparations for the PC meeting and note-taking and time management during the meeting. He deserves a lot of credit for a well-organized and productive meeting.

We would also like to thank the members of the organizing committee, in particular Kris Flautner, the general chair, for the continuous support to all program committee activities and requests. We thank the past program and general chairs, especially Matt Farrens, Chris Nitta, and Christos Kozyrakis for guidance. We are grateful to all the members of the steering committee for their valuable advice, and in particular Rich Belgard and Onur Mutlu for fielding numerous questions.

Finally, we would like to thank all the authors that submitted papers to MICRO-47. The strong program is a reflection of a thriving research community with hardworking members with broad interests. We hope you all enjoy the technical program of MICRO-47 and the opportunity to exchange ideas and interact with colleagues in beautiful Cambridge, UK.

Emre Ozer Thomas F. Wenisch

Program Co-Chairs, MICRO-47

Organizing Committee

General Chair

Krisztian Flautner, ARM

Program Co-Chairs

Thomas F. Wenisch, University of Michigan

Emre Ozer, ARM

Workshop & Tutorials Chair

Jason Mars, University of Michigan

Publicity Chair

Ronald Dreslinski, *University of Michigan*

Registration Chair

Tim Jones, University of Cambridge

Finance Chair

Bruno Jansen, ARM

Publication Chair

Michael Ferdman, Stony Brook University

Submission Chair

Anthony Gutierrez, University of Michigan

Web Chair

Marios Kleanthous, Mesoyios College

Steering Committee

Richard Belgard (Chair), Consultant

David Albonesi, Cornell

Tom Conte, Georgia Tech

Kemal Ebcioglu, Global Supercomputing

Matthew Farrens, UC Davis

Scott Mahlke, University of Michigan

Bill Mangione-Smith, Consultant

Onur Mutlu, Carnegie Mellon University

Yale Patt, UT Austin

Milos Prvulovic, Georgia Tech

Program Committee

Abhishek Bhattacharjee, Rutgers

Ali Saidi, ARM

Andrew Hilton, Duke University

Avi Mendelson, Technion

Benoit Dupont de Dinechin, Kalray

Boris Grot, University of Edinburgh

Carole-Jean Wu, Arizona State University

Chris Wilkerson, Intel

Daniel Sanchez, MIT

David Brooks, Harvard

David Meisner, Facebook

David Wood, University of Wisconsin

Gabriel Loh, AMD

Hillery Hunter, IBM

Hsien-Hsin Lee, TSMC & Georgia Tech

Huiyang Zhou, North Carolina State University

Jangwoo Kim, POSTECH

Jared Smolens, Oracle

Karin Strauss, Microsoft

Kevin Lim, HP

Kim Hazelwood, Google

Luis Ceze, University of Washington

Mary Jane Irwin, Penn State

Michael Ferdman, Stony Brook

Mike O'Connor, Nvidia

Mike Shebanow, Samsung

Milo Martin, University of Pennsylvania

Milos Prvulovic, Georgia Tech

Murali Annavaram, USC

Onur Mutlu, Carnegie Mellon

Ozcan Ozturk, Bilkent University

Roger Isaac, Waveconnex

Ron Dreslinski, University of Michigan

Scott Mahlke, University of Michigan

Shubu Mukherjee, *Cavium*Simha Sethumadhavan, *Columbia*Thomas Pawlowski, *Micron*Trey Cain, *Qualcomm*Vijay Reddi, *University of Texas, Austin*Walid Najjar, *University of California, Riverside*Yanos Sazeides, *University of Cyprus*Yunji Chen, *ICT*

External Review Committee

Aamer Jaleel, Intel

Alaa Alameldeen, Intel

Alper Buyuktosunoglu, IBM

Andre Seznec, INRIA

Andreas Hansson, ARM

Andreas Moshovos, University of Toronto

Andrew DeOrio, University of Michigan

Arrvindh Shriraman, Simon Frasier University

Artur Klauser, Google

Babak Falsafi, EPFL

Ben Lee, Duke University

Benjamin Serebrin, Google

Brad Beckmann, AMD

Chita Das, Penn State University

Chris Batten, Cornell University

Chris Stewart, Ohio State University

Dan Sorin, *Duke University*

Daniel Jimenez, Texas A&M

David Penry, Brigham Young University

David Wentzlaff, Princeton

Debbie Marr, Intel

Dennis Abts, Google

Derek Chiou, University of Texas, Austin

Doug Burger, Microsoft

Edward Suh, Cornell University

Engin Ipek, *University of Rochester*

Eric Chung, Microsoft

Eric Rotenberg, NC State University

Guiseppe Desoli, ST Microelectronics

Hadi Esmaelizadeh, Georgia Tech University

Hyesoon Kim, Georgia Tech University

Jaehyuk Huh, KAIST

James Hoe, Carnegie Mellon University

James Laudon, Google

Jason Cong, UCLA

Jason Mars, University of Michigan

Jian Li, Huawei Technologies Co. Ltd.

Jichuan Chang, Google

Joe Devietti, University of Pennsylvania

John Davis, Consultant

John Kim, KAIST

Jose Renau, UCSC

Joseph Greathouse, AMD

Jung Ho Ahn, Seoul National University

Kanad Ghose, SUNY Binghampton

Karu Sankaralingam, University of Wisconsin

Kelly Shaw, University of Richmond

Koji Inoue, Kyushu University

Leslie Barnes, AMD

Lieven Eeckhout, Ghent University

Lingjia Tang, University of Michigan

Mahmut Kandemir, Penn State University

Mark Oskin, University of Washington

Martha Kim, Columbia University

Matt Farrens, University of California, Davis

Mattan Erez, UT Austin

Michael Huang, *University of Rochester*

Michael Marty, Google

Michel Dubois, USC

Moinuddin Qureshi, Georgia Tech University

Nam Sung Kim, University of Wisconsin

Natalie Enright Jerger, University of Toronto

Naveen Muralimanohar, HP Labs

Nikos Hardavellas, Northwestern University

Norm Jouppi, Google

Omer Khan, University of Connecticut

Paolo Faraboschi, HP Labs

Partha Ranganathan, Google

Paul Gratz, TAMU, Texas A&M

Per Stenstrom, Chalmers, Chalmers University

Philip Wells, Google

Radu Teodorescu, Ohio State University

Rajeev Balasubramonian, University of Utah

Rajiv Gupta, UC Riverside

Rakesh Kumar, UIUC

Reetuparna Das, University of Michigan

Resit Sendag, University of Rhode Island

Ricardo Bianchini, Rutgers University & Microsoft

Robert Colwell, DARPA

Robert Hundt, Google

Ronny Ronen, Intel

Ruby Lee, *Princeton University*

Rune Holm, ARM

Sangyeun Cho, University of Pittsburgh

Scott Rixner, Rice University

Shidhartha Das, ARM

Suzanne Rivoire, Sonoma State University

T N Vijaykumar, Purdue University

Tao Li, University of Florida

Tim Sherwood, *UCSB*

Trevor Mudge, University of Michigan

Wen-Mei Hwu, UIUC

Xiaorui Wang, *Ohio State University*

Xipeng Shen, College of William & Mary

Yuan Xie, Penn State University

Zhibin Yu, Shenzhen University

External Reviewers

Adrian Caulfield George Tziantzioulis Nick Kidd
Adrian Sampson Gu-Yeon Wei Prabal Dutta

Adrian Tang Guru Venkataramani Qi Guo

Akanksha Jain Hank Hoffman Raanan Sade Alexander Butuzov Helia Naeimi Rathijit Sen

Alexander Grosul Hyeran Jeon Richard Sampson

Andre Baixo Igor Markov Ronny Krashinsky

Andrew Targhetta Isaac Richter Satish Narayanasamy

Anirudh Badam Jack Sampson Saurabh Gupta Ben Ransford Jim Smith Shidhartha Das Ben Serebrin Joel Coburn Shih-Lien Lu Brad Beckmann John Demme Soojung Ryu **Brandon Holt** Karthik Ramani Sriram Govindan Brandon Lucia Ke Wang Steve Keckler

Bruce Jacob Kimish Patel Steve Pelley

Calvin Lin Krishna Malladi Steve Reinhardt

Chaitali Chakrabarti Laurent Morin Sudhakar Yalamanchili

Chao Li Leonid Yavits Sunjae Park

Ching-Kai Liang Liangjun Zhang Sylvain Collange

Chris Nitta Lihu Rappoport Tao Tong
Chrys Nicopoulos Lisa Hsu Tarek Taha
Dan Gibson Lizy John Tor Aamodt
Daniel Wong Lu Peng Wei Wu

David A. Roberts Mahdi Nazm Bojnordi Xiaochen Guo Martin Dimitrov Xiaochen Guo **Dmitry Ponomarev** Domagoj Babic Matt Horsnell Xuan Zhang **E** Lewis Matthew Fojtik Xuehai Oian Efi Rotem Mikko Lipasti Yan Solihin Faissal Sleiman Mohit Tiwari Yi Yang

Francois Irigoin Moty Mehalel Yoonseo Choi Fred Chong Naser Sedaghati Yuan Chou

Ganesh Dasika Neha Agarwal

Keynotes

From IoT to Services – Efficiency Matters

Mike Muller, Chief Technology Officer, ARM Holdings plc

The Internet of Things is driving developments from near threshold lowest power and energy sensors through to highest performance servers and networking infrastructure enabling cloud services. Despite the diversity of different products they all share the same common requirement namely efficiency. But more than just products it is the transformation of business models that make IoT an exciting opportunity.

Bio:

Mike was one of the founders of ARM. Before joining the Group, he was responsible for hardware strategy and the development of portable products at Acorn Computers. He was previously at Orbis Computers. At ARM he was VP, Marketing from 1992 to 1996 and EVP, Business Development until October 2000 when he was appointed Chief Technology Officer. In October 2001, he was appointed to the Board. He is also a non-executive director of Intelligent Energy an intellectual property company specializing in the development of modular, low carbon fuel cell systems for their partners.

The End of Moore's Law - Again

Trevor Mudge, Bredt Professor of Computer Engineering, the University of Michigan, Ann Arbor

Claims that Moore's Law is about to end occur with surprising regularity. Each time the claims have been proved wrong. Is it different this time? There are some indicators that it may be: Dennard scaling is breaking down and the skyrocketing cost of semiconductor fabs may stop progress before technical limits.

We will review some of the past claims that prematurely forecast Moore's Law's demise, and the developments that overcame those barriers. Those developments also resulted in significant changes to the criteria we use to design computers. We will discuss them too.

Discussion about the "end of Moore's Law" raises the question of physical limits for computer design in general. We will include them in the talk and show how they may shed light on the current claim that Moore's Law is ending.

We will review promising technologies for memories and die stacking that will allow us to move forward in the immediate future. Finally, we will discuss newer technologies that

Bio:

Trevor Mudge received the Ph.D. degrees in Computer Science from the University of Illinois, Urbana in 1977. Since then he has been on the faculty of the University of Michigan, Ann Arbor. In 2003 he was named the first Bredt Family Professor of Electrical Engineering and Computer Science. Previously he served a ten-year term as the Director of the Advanced Computer Architecture Laboratory, which is a group of eight faculty and about 60 graduate students. He is author of numerous papers on computer architecture, programming languages, VLSI design, and computer vision. He has also chaired 49 theses in these areas. His research interests include computer architecture, computer-aided design, and compilers.

In 2014 he was the recipient of ACM/IEEE CS Eckert-Mauchly Award "For pioneering contributions to low-power computer architecture and its interaction with technology." In addition to his position as a faculty member, he runs Idiot Savants, a chip design consultancy. Trevor Mudge is a Life Fellow of the IEEE, a member of the ACM, the IET, and the British Computer Society.

Investigating the Brain's Computational Paradigm

James E. Smith, Professor Emeritus, the University of Wisconsin-Madison

Understanding and implementing the brain's computational paradigm is the grand challenge facing computer researchers. Not only does it provide computational capabilities far beyond those of conventional computers, its energy efficiency is truly remarkable. The brain's neocortex is constructed of massively interconnected neurons that compute and communicate via voltage spikes, and a strong argument can be made that precise spike timing is an essential element of the paradigm. This argument leads to the seeming paradox that the way we naturally reason about computation is an obstacle to understanding the way our brain, itself, computes. I will describe a small set of biologically plausible computational elements based on precise spike timing. Then, through examples using these computational elements, I will illustrate some features of spike-based temporal computation and how it differs from other brain-inspired approaches. This is only one person's first step toward understanding the brain's computational paradigm, and as such it is one of many divergent theories. However, it illustrates important aspects of this grand challenge research area and serves as a case study in the way a computer researcher can participate in addressing the challenge.

Bio:

James E. Smith is Professor Emeritus in the Department of Electrical and Computer Engineering at the University of Wisconsin-Madison. He received his PhD from the University of Illinois in 1976. He then joined the faculty of the University of Wisconsin-Madison, teaching and conducting research — first in fault-tolerant computing, then in computer architecture. He has been involved in a number of computer research and development projects both as a faculty member at Wisconsin and in industry (Control Data Corporation, Astronautics Corporation, Cray Research, Google, and Intel).

Prof. Smith has made a number of significant contributions to the development of superscalar processors. These contributions include basic mechanisms for dynamic branch prediction and implementing precise traps. He has also studied vector processor architectures and worked on the development of innovative microarchitecture paradigms. He received the 1999 ACM/IEEE Eckert-Mauchly Award for these contributions.

More recently, Prof. Smith focused on the virtual machine abstraction as a technique for providing high performance and power efficiency through co-design and tight coupling of virtual machine hardware and software. He is co-author of a book on virtual machines. Currently, he is studying computational neuroscience at home along the Clark Fork near Missoula, Montana.