
Barcelona, Spain
12-16 March 2016
## Table of Contents

### Session 1A – Hardware Accelerators

- **Memristive Boltzmann Machine: A Hardware Accelerator for Combinatorial Optimization and Deep Learning** ................................................................. 1
  Mahdi Nazm Bojnordi and Engin Ipek (University of Rochester)

- **TABLA: A Unified Template-based Architecture for Accelerating Statistical Machine Learning** ............ 14
  Divya Mahajan, Jongse Park, Emmanuel Amaro, Hardik Sharma, Amir Yazdanbaksh, Joon Kim and Hadi Esmailzadeh (Georgia Institute of Technology)

- **Pushing the Limits of Accelerator Efficiency While Retaining General-Purpose Programmability** .......... 27
  Tony Nowatzki, Vinay Gangadhar, Karthikeyan Sankaralingam (University of Wisconsin – Madison) and Greg Wright (Qualcomm)

### Session 1B – Mobile/IoT

- **A Low Power Software-Defined-Radio Baseband Processor for the Internet of Things** ....................... 40
  Yajing Chen, Shengshuo Lu, Hun-Seok Kim, David Blaauw, Ronald G. Dreslinski and Trevor Mudge (University of Michigan)

- **Improving Smartphone User Experience by Balancing Performance and Energy with Probabilistic QoS Guarantee** .................................................................................................................. 52
  Benjamin Gaudette, Carole-Jean Wu and Sarma Vrudhula (Arizona State University)

- **Mobile CPU’s Rise to Power: Quantifying the Impact of Generational Mobile CPU Design Trends on Performance, Energy, and User Satisfaction** ................................................................. 64
  Matthew Halpern, Yuhao Zhu and Vijay Janapa Reddi (UT Austin)

### Session 2A – Non-volatile Memories

- **Atomic Persistence for SCM with a Non-intrusive Backend Controller** ................................................. 77
  Kshitij Doshi (Intel Corporation), Ellis Giles and Peter Varman (Rice University)

- **CompEx: Compression-Expansion Coding for Energy, Latency, and Lifetime Improvements in MLC/TLC NVM** .................................................................................................................. 90
  Poovaiah M. Palangappa and Kartik Mohanram (University of Pittsburgh)

- **A Low-Power Hybrid Reconfigurable Architecture For Resistive Random-Access Memories** ............... 102
  Miguel Angel Lastras Montañó, Amirali Ghofrani and Kwang-Ting Cheng (UCSB)

### Session 2B – Reconfigurable Architectures

- **A Performance Analysis Framework for Optimizing OpenCL Applications on FPGAs** ............................. 114
  Zeke Wang, Bingsheng He (Nanyang Technological University), Wei Zhang (HKUST) and Shunning Jiang (Nanyang Technological University)
• **HRL: Efficient and Flexible Reconfigurable Logic for Near-Data Processing** ................................................. 126
  Mingyu Gao and Christos Kozyrakis (Stanford University)

• **Software Transparent Dynamic Binary Translation for Coarse-Grain Reconfigurable Architectures** .......... 138
  Matthew A. Watkins (Lafayette College), Anthony Carno (Bucknell University) and Tony Nowatzki (University of Wisconsin-Madison)

**Session 3A – GPUs**

• **Core Tunneling: Variation-Aware Voltage Noise Mitigation in GPUs** ................................................................. 151
  Renji Thomas, Kristin Barber, Naser Sedaghati, Li Zhou and Radu Teodorescu (The Ohio State University)

• **Warped-Preexecution: A GPU Pre-execution Approach for Improving Latency Hiding** ......................... 163
  Keunsoo Kim, Sangpil Lee, Myung Kuk Yoon (Yonsei University), Gunjae Koo (University of Southern California), Won Woo Ro (Yonsei University) and Murali Annavaram (University of Southern California)

• **Approximating Warps with Intra-warp Operand Value Similarity** .............................................................. 176
  Daniel Wong (University of California, Riverside), Nam Sung Kim (University of Illinois at Urbana–Champaign) and Murali Annavaram (University of Southern California)

• **A Case for Toggle-Aware Compression for GPU Systems** ............................................................................. 188
  Gennady Pekhimenko (CMU), Evgeny Bolotin (NVIDIA), Nandita Vijaykumar, Onur Mutlu, Todd C. Mowry (CMU) and Stephen W. Keckler (NVIDIA / UT-Austin)

**Session 3B – Caches**

• **Minimal Disturbance Placement and Promotion** ........................................................................................................... 201
  Elvira Teran, Daniel A. Jiménez (Texas A&M), Zhe Wang (Intel Labs) and Yingying Tian (AMD)

• **Revisiting Virtual L1 Caches: A Practical Design Using Dynamic Synonym Remapping** ....................... 212
  Hongil Yoon and Gurindar S. Sohi (University of Wisconsin–Madison)

• **Modeling Cache Performance Beyond LRU** .............................................................................................................. 225
  Nathan Beckmann and Daniel Sanchez (MIT)

• **Efficient Footprint Caching for Tagless DRAM Caches** ..................................................................................... 237
  Hakbeom Jang (Sungkyunkwan University), Yongjun Lee (Sungkyunkwan University and Samsung Electronics), Jongwon Kim (Sungkyunkwan University), Youngsok Kim, Jangwoo Kim (POSTECH), Jinkyu Jeong and Jae W. Lee (Sungkyunkwan University)

**Session 4A – Coherence and Consistency**

• **SCsafe: Logging Sequential Consistency Violations Continuously and Precisely** ........................................... 249
  Yuelu Duan, Josep Torrellas (University of Illinois) and David Koufaty (Intel Corporation)

• **LASER: Light, Accurate Sharing dEtection and Repair** ....................................................................................... 261
  Liang Luo, Akshitha Sriraman, Brooke Fugate (University of Pennsylvania), Shiliang Hu, Gilles Pokam, Chris J. Newburn (Intel) and Joseph Devietti (University of Pennsylvania)
• Improving GPU Hardware Transactional Memory Performance via Conflict and Contention Reduction ................................................................................................................................. 274
  Sui Chen and Lu Peng (Louisiana State University)

• PleaseTM: Enabling Transaction Conflict Management in Requester-wins Hardware Transactional Memory ........................................................................................................................................... 285
  Sunjae Park, Milos Prvulovic (Georgia Institute of Technology) and Christopher J Hughes (Intel)

Session 4B – Interconnects
• Efficient Synthetic Traffic Models for Large, Complex SoCs ................................................................................................................................................................................................. 297
  Jieming Yin, Onur Kayiran, Matthew Poremba (AMD Research), Natalie Enright Jerger (AMD Research, University of Toronto) and Gabriel H. Loh (AMD Research)

• DVFS for NoCs in CMPs: A Thread Voting Approach ................................................................................................................................................................................................. 309
  Yuan Yao and Zhonghai Lu (KTH Royal Institute of Technology, Sweden)

• SLaC: Stage Laser Control for a Flattened Butterfly Network .................................................................................................................................................................................. 321
  Yigit Demir (Intel) and Nikos Hardavellas (Northwestern University)

• The Runahead Network-On-Chip ................................................................................................................................................................................................. 333
  Zimo Li, Joshua San Miguel (University of Toronto) and Natalie Enright Jerger (University of Toronto/AMD)

Session 5A – GPGPUs
• Towards High Performance Paged Memory for GPUs ................................................................................................................................................................................................. 345
  Tianhao Zheng (The University of Texas at Austin & NVIDIA), David Nellans, Arslan Zulfiqar, Mark Stephenson (NVIDIA) and Stephen W Keckler (NVIDIA / UT-Austin)

• Simultaneous Multikernel GPU: Multi-tasking Throughput Processors via Fine-Grained Sharing .......... 358
  Zhenning Wang (Shanghai Jiao Tong University), Jun Yang, Rami Melhem, Bruce Childers, Youtao Zhang (University of Pittsburgh) and Minyi Guo (Shanghai Jiao Tong University)

• iPAWS : Instruction-Issue Pattern-based Adaptive Warp Scheduling for GPGPUs ........................................ 370
  Minseok Lee (KAIST), Gwangsun Kim (KAIST / NVIDIA), John Kim (KAIST), Woong Seo, Yeongon Cho and Soojung Ryu (Samsung Electronics)

Session 5B – Security
• Lattice Priority Scheduling: Low-Overhead Timing Channel Protection for a Shared Memory Controller .................................................................................................................................................. 382
  Andrew Ferraiuolo, Yao Wang (Cornell University), Danfeng Zhang (Penn State University), Andrew C. Myers and Ed Suh (Cornell University)

• A Complete Key Recovery Timing Attack on a GPU ................................................................................................................................. 394
  Zhen Hang Jiang, Yunsi Fei and David Kaeli (Northeastern University)
• **CATalyst: Defeating Last Level Cache Side Channel Attacks in Cloud Computing** ........................................ 406
  Fangfei Liu (Princeton University), Qian Ge (NICTA and UNSW), Yuval Yarom (University of Adelaide and NICTA), Frank Mckeen, Carlos Rozas (Intel), Gernot Heiser (NICTA and UNSW) and Ruby B. Lee (Princeton University)

Session 6A – Large-Scale Systems

• **Predicting the Memory Bandwidth and Optimal Core Allocations for Multi-threaded Applications on Large scale NUMA Machines** ................................................................................................ 419
  Wei Wang, Jack W. Davidson and Mary Lou Soffa (University of Virginia)

• **A Market Approach for Handling Power Emergencies in Multi-Tenant Data Center** ........................................ 432
  Mohammad A. Islam (UC Riverside), Xiaqi Ren (Caltech), Shaolei Ren (UC Riverside), Adam Wierman (Caltech) and Xiaorui Wang (The Ohio State University)

• **SizeCap: Efficiently Handling Power Surges in Fuel Cell Powered Data Centers** ........................................ 444
  Yang Li (Carnegie Mellon University), Di Wang (Microsoft Corporation), Saugata Ghose (Carnegie Mellon University), Jie Liu, Sriram Govindan, Sean James, Eric Peterson, John Siegler (Microsoft Corporation), Rachata Ausavarungrirun and Onur Mutlu (Carnegie Mellon University)

Session 6B – Potpourri

• **MaPU: A Novel Mathematical Computing Architecture** .................................................................................. 457
  Donglin Wang, Shaolin Xie, Zhiwei Zhang, Xueliang Du, Lei Wang, Zijun Liu, Xiao Lin, Jie Hao, Chen Lin, Hong Ma, Zhonghua Pu, Guangxin Ding, Wenqin Sun, Fabiao Zhou, Weili Ren, Huijuan Wang, Mengchen Zhu, Lipeng Yang, NuoZhou Xiao, Qian Cui, Xingang Wang, Ruoshan Guo, Xiaqin Wang (Chinese Academy of Science, Institute of Automation), Leizu Yin (Spreadtrum Comm), Tao Wang and Yongyong Yang (Huawei)

• **Best-Offset Hardware Prefetching** ............................................................................................................... 469
  Pierre Michaud (Inria)

• **DUANG: Fast and Lightweight Page Migration in Asymmetric Memory Systems** ................................... 481
  Hao Wang (University of Wisconsin-Madison), Jie Zhang (Yonsei University), Gieseo Park (UT-Dallas), Sharmila Shridhar (University of Wisconsin-Madison), Myoungsoo Jung (Yonsei University) and Nam Sung Kim (University of Illinois-Urbana-Champaign)

Session 7A – Industry Session

• **Selective GPU Caches to Eliminate CPU–GPU HW Cache Coherence** ....................................................... 494
  Neha Agarwal (University of Michigan), David Nellans, Eiman Ebrahimi (NVIDIA), Thomas F. Wenisch (University of Michigan), John Danskin, and Stephen W. Keckler (NVIDIA)

• **Venice: Exploring Server Architectures for Effective Resource Sharing** ..................................................... 507
  Jianbo Dong, Rui Hou (Institute of Computing Technology, Chinese Academy of Sciences), Michael Huang (University of Rochester), Tao Jiang, Boyan Zhao (Institute of Computing Technology, Chinese Academy of Sciences), Sally A. McKee (Chalmers University of Technology), Haibin Wang, Xiaosong Cui (Huawei Technologies Co., Ltd) and Lixin Zhang (Institute of Computing Technology, Chinese Academy of Sciences)
• A Large-Scale Study of Soft-Errors on GPUs in the Field ................................................................. 519
  Bin Nie (College of William and Mary), Devesh Tiwari, Saurabh Gupta (Oak Ridge National Laboratory), Evgenia
  Smirni (College of William and Mary) and James H. Rogers (Oak Ridge National Laboratory)

• Design and Implementation of A Mobile Storage Leveraging the DRAM Interface .......................... 531
  Sungyong Seo, Youngjin Cho, Youngkwang Yoo, Otae Bae, Jaegun Park, Heehyun Nam, Sunmi Lee, Yongmyung
  Lee, Seungdo Chae, Moonsoo Kwon, Jin-Hyeok Choi, Sangyeun Cho, Jaeheon Jeong and Duckhyun Chang
  (Samsung Electronics Co., Ltd.)

Session 7B – Memory Technology

• Restore Truncation for Performance Improvement in Future DRAM Systems .................................. 543
  Xianwei Zhang, Youtao Zhang, Bruce R. Childers (Computer Science Department, University of Pittsburgh) and
  Jun Yang (Electrical and Computer Engineering Department, University of Pittsburgh)

• Parity Helix: Efficient Protection for Single-Dimensional Faults
  in Multi-dimensional Memory Systems .................................................................................................. 555
  Xun Jian, Rakesh Kumar (University of Illinois at Urbana Champaign) and Vilas Sridharan (AMD)

• Low-Cost Inter-Linked Subarrays (LISA): Enabling Fast Inter-Subarray
  Data Movement in DRAM ....................................................................................................................... 568
  Kevin K. Chang (Carnegie Mellon University), Prashant J. Nair (Georgia Institute of Technology), Saugata Ghose,
  Donghyuk Lee (Carnegie Mellon University), Moinuddin K. Qureshi (Georgia Institute of Technology) and Onur
  Mutlu (Carnegie Mellon University)

• ChargeCache: Reducing DRAM Latency by Exploiting Row Access Locality .................................. 581
  Hasan Hassan (Carnegie Mellon University, TOBB University of Economics & Technology), Gennady
  Pekhimenko, Nandita Vijaykumar, Vivek Seshadri, Donghyuk Lee (Carnegie Mellon University), Oguz Ergin
  (TOBB University of Economics & Technology) and Onur Mutlu (Carnegie Mellon University)

Session 8 – Modeling and Testing

• Amdahl’s Law for Lifetime Reliability Scaling in Heterogeneous Multicore Processors .................. 594
  William J. Song, Saibal Mukhopadhyay and Sudhakar Yalamanchili (Georgia Tech)

• LiveSim: Going Live with Microarchitecture Simulation ................................................................. 606
  Sina Hassani, Gabriel Southern and Jose Renau (UC Santa Cruz)

• McVerSi: A Test Generation Framework for Fast Memory Consistency Verification in Simulation ...... 618
  Marco Elver and Vijay Nagarajan (University of Edinburgh)

Session 9A – Caches and TLB

• Energy-Efficient Address Translation ........................................................................................................ 631
  Vasileios Karakostas (Barcelona Supercomputing Center, Universitat Politècnica de Catalunya), Jayneel Gandhi
  (University of Wisconsin – Madison), Adrián Cristal (Barcelona Supercomputing Center, Universitat Politècnica de
  Catalunya, IIIA-CSIC), Mark D. Hill (University of Wisconsin – Madison), Kathryn S. McKinley (Microsoft
  Research), Mario Nemirovsky (Barcelona Supercomputing Center, ICREA), Michael M. Swift (University of
  Wisconsin – Madison) and Osman Unsal (Barcelona Supercomputing Center)
• **RADAR: Runtime-Assisted Dead Region Management for Last-Level Caches** ........................................... 644
  Madhavan Manivannan, Vassilis Papaefstathiou, Miquel Pericas and Per Stenstrom (Chalmers University of Technology)

• **Cache QoS: From Concept to Reality in the Intel Xeon E5-2600 v3 Server Processor Family** .................. 657
  Andrew Herdrich, Edwin Verplanke, Chris Gianos, Ramesh Illikkal, Ronak Singhal, Ravi Iyer, and Priya Autee (Intel)

Session 9B – Microarchitecture

• **Symbiotic Job Scheduling on the IBM POWER8** .......................................................................................... 669
  Josué Feliu (Universitat Politècnica de València), Stijn Eyerman (Ghent University), Julio Sahuquillo and Salvador Petit (Universitat Politècnica de València)

• **ScalCore: Designing a Core for Voltage Scalability** .................................................................................. 681
  Bhargava Gopireddy (University of Illinois), Choungki Song (University of Wisconsin), Josep Torrellas, Nam Sung Kim (University of Illinois), Aditya Agrawal (Nvidia) and Asit Mishra (Intel)

• **Cost Effective Physical Register Sharing** ................................................................................................... 694
  Arthur Perais and André Seznec (INRIA)

Author Index ....................................................................................................................................................... 707