## **2018 IEEE 36th International Conference on Computer Design** (ICCD 2018)

Orlando, Florida, USA 7-10 October 2018



**IEEE Catalog Number:** 

CFP18ICD-POD **ISBN**: 978-1-5386-8478-8

### Copyright © 2018 by the Institute of Electrical and Electronics Engineers, Inc. All Rights Reserved

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

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.

 IEEE Catalog Number:
 CFP18ICD-POD

 ISBN (Print-On-Demand):
 978-1-5386-8478-8

 ISBN (Online):
 978-1-5386-8477-1

ISSN: 1063-6404

#### **Additional Copies of This Publication Are Available From:**

Curran Associates, Inc 57 Morehouse Lane Red Hook, NY 12571 USA Phone: (845) 758-0400

Fax: (845) 758-2633

E-mail: curran@proceedings.com Web: www.proceedings.com



# 2018 IEEE 36th International Conference on Computer Design ICCD 2018

### **Table of Contents**

| Welcome Message from the General Chair xvi Organizing Committee xvii  |
|---|
| Welcome Message from the Program Chair xviii  |
| Program Committee xix   |
| Keynotes  |
| Special Panels xxvi   |
| Tutorials xxxiii  |
| Session 1: Best Papers Session  |
| Composable Template Attacks Using Templates for Individual Architectural Components   |
| Thermal-Aware 3D Symmetrical Buffered Clock Tree Synthesis  |
| Low-Overhead Microarchitectural Patching for Multicore Memory Subsystems  |
| Power Grab in Aggressively Provisioned Data Centers: What is the Risk and What Can Be Done About It   |
| Xiaofeng Hou (Shanghai Jiao Tong University), Luoyao Hao (Shanghai<br>Jiao Tong University), Chao Li (Shanghai Jiao Tong University), Quan<br>Chen (Shanghai Jiao Tong University), Wenli Zheng (Shanghai Jiao Tong<br>University), and Minyi Guo (Shanghai Jiao Tong University) |
| Session 2A: SSD   |
| Pensieve: a Machine Learning Assisted SSD Layer for Extending the Lifetime  |

| Selective Compression Scheme for Read Performance Improvement on Flash Devices  | . 43 |
|---|------|
| OSPADA: One-Shot Programming Aware Data Allocation Policy to Improve 3D NAND Flash Read Performance  Fei Wu (Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, China), Zuo Lu (Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, China), You Zhou (Wuhan National Laboratory for Optoelectronics & School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China), Xubin He (Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA), Zhihu Tan (School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China), and Changsheng Xie (School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China)  Optical Science Science and Technology, Wuhan, China) |      |
| Cap: Exploiting Data Correlations to Improve the Performance and Endurance of SSD RAID  | . 59 |
| Session 2B: Side Channels   |      |
| A Timing Side-Channel Attack on a Mobile GPU  Elmira Karimi (Northeastern University), Zhen Hang Jiang (Northeastern  University), Yunsi Fei (Northeastern University), and David Kaeli  (Northeastern University)  | 67   |
| Analysis of Row Hammer Attack on STTRAM   | . 75 |
| Machine Learning on the Thermal Side-Channel: Analysis of Accelerator-Rich Architectures  | . 83 |
|   |      |

### **Session 3A: Security and Capability**

| CheriRTOS: A Capability Model for Embedded Devices  Hongyan Xia (University of Cambridge, UK), Jonathan Woodruff (University of Cambridge, UK), Hadrien Barral (École normale supérieure, France), Lawrence Esswood (University of Cambridge, UK), Alexandre Joannou (University of Cambridge, UK), Robert Kovacsics (University of Cambridge, UK), David Chisnall (University of Cambridge, UK), Michael Roe (University of Cambridge, UK), Brooks Davis (SRI International, US), Edward Napierala (University of Cambridge), John Baldwin (SRI International, US), Khilan Gudka (University of Cambridge, UK), Peter G. Neumann (SRI International, US), Alexander Richardson (University of Cambridge), Simon W. Moore (University of Cambridge, UK), and Robert N. M. Watson (University of Cambridge, UK) | 92  |
|--|-----|
| ReadPRO: Read Prioritization Scheduling in ORAM for Efficient Obfuscation in Main Memories1  Joydeep Rakshit (University of Pittsburgh) and Kartik Mohanram  (University)  | 100 |
| SGXlinger: A New Side-Channel Attack Vector Based on Interrupt Latency Against Enclave   |     |
| Execution  | 108 |
| Breaking the Oblivious-RAM Bandwidth Wall  | 115 |
| Session 3B: Microarchitecture  |     |
| Rearranging Random Issue Queue with High IPC and Short Delay Shinji Sakai (Nagoya University), Taishi Suenaga (Nagoya University), Ryota Shioya (The University of Tokyo), and Hideki Ando (Nagoya University)   | 123 |
| Array Tracking Prefetcher for Indirect Accesses  | 132 |
| Dynamically Disabling Way-prediction to Reduce Instruction Replay  | 140 |
| Analysis and Characterization of Ultra Low Power Branch Predictors   | 144 |
| OldSpot: A Pre-RTL Model for Fine-Grained Aging and Lifetime Optimization  | 148 |

| SPF: Selective Pipeline Flush  | 152 |
|--|-----|
| Session 4A: Logic and Circuit Design 1   |     |
| Power-Efficient ReRAM-Aware CNN Model Generation   | 156 |
| R-Accelerator: A Reconfigurable Accelerator with RRAM Based Logic Contraction and Resource Optimization for Application Specific Computing   | 163 |
| 3D Crosspoint Memory as a Parallel Architecture for Computing Network Reachability   | 171 |
| Dynamic Computing in Memory (DCIM) in Resistive Crossbar Arrays  | 179 |
| Low Area-Delay Complexity Digit-Level Parallel-In Serial-Out Multiplier Over GF(2m) Based on Overlap-Free Karatsuba Algorithm  | 187 |
| Session 4B: Design Automation  |     |
| Software and Hardware Techniques for Reducing the Impact of Quantization Errors in Memristor Crossbar Arrays  Baogang Zhang (University of Central Florida) and Rickard Ewetz (University of Central Florida)  | 195 |
| Trading Off Temperature Guardbands via Adaptive Approximations  Behzad Boroujerdian (university of Texas at Austin), Hussam Amrouch (Karlsruhe Institute of Technology), Jörg Henkel (Karlsruhe Institute of Technology), and Andreas Gerstlauer (Karlsruhe Institute of Technology) | 202 |
| Lattice-Traversing Design Space Exploration for High Level Synthesis  Lorenzo Ferretti (Università della Svizzera italiana, Switzerland), Giovanni Ansaloni (Università della Svizzera italiana, Switzerland), and Laura Pozzi (Università della Svizzera italiana, Switzerland)     | 210 |

### **Session 5A: Novel Architectures**

| Systems   | 218 |
|---|-----|
| Jinhang Choi (Pennsylvania State University), Jack Sampson<br>(Pennsylvania State University), and Vijaykrishnan Narayanan<br>(Pennsylvania State University)   | 210 |
| Puppet: Energy Efficient Task Mapping For Storage-Less and Converter-Less Solar-Powered Non-Volatile Sensor Nodes  Yue Xu (Chongqing University, China), Hyung Gyu Lee (Daegu University, Republic of Korea), Xianzhang Chen (Chongqing University, China), Bo Peng (Southwest University of Science and Technology, China), Duo Liu (Chongqing University, China), and Liang Liang (Chongqing university, China) | 226 |
| SYNCVIBE: Fast and Secure Device Pairing through Physical Vibration on Commodity  |     |
| Smartphones  Kyuin Lee (University of Wisconsin-Madison, USA), Vijay Raghunathan (Purdue University, USA), Anand Raghunathan (Purdue University, USA), and Younghyun Kim (University of Wisconsin-Madison, USA)   | 234 |
| FPGA Virtualization in Cloud-Based Infrastructures Over Virtio  | 242 |
| Forca: Fast and Atomic Remote Direct Access to Persistent Memory  | 246 |
| Session 5B: Memory 1  |     |
| CART: Cache Access Reordering Tree for Efficient Cache and Memory Accesses in GPUs<br>Yongbin Gu (Oregon State University) and Lizhong Chen (Oregon State<br>University)  | 250 |
| ArchSampler: Architecture-Aware Memory Sampling Library for In-Memory Applications  Jian Zhou (University of Central Florida) and Jun Wang (University of Central Florida)  | 258 |
| PIM-TGAN: A Processing-in-Memory Accelerator for Ternary Generative Adversarial Networks Adnan Siraj Rakin (University of Central Florida), Shaahin Angizi (University of Central Flroida), Zhezhi He (University of Central Florida), and Deliang Fan (University of Central Florida)  | 266 |
| Path Prefetching: Accelerating Index Searches for In-Memory Databases  Shuo Li (National University of Defense Technology, China), Zhiguang  Chen (Sun Yat-sen University, China), Nong Xiao (Sun Yat-sen  University, China), and Guangyu Sun (Peking University, China)   | 274 |

| Reducing Inter-Application Interferences in Integrated CPU-GPU Heterogeneous Architecture 278  Hao Wen (Virginia Commonwealth University) and Wei Zhang (Virginia  Commonwealth University) |
|---|
| Session 6A: Memory 2  |
| Solar-DRAM: Reducing DRAM Access Latency by Exploiting the Variation in Local Bitlines  |
| Scalable and Efficient Virtual Memory Sharing in Heterogeneous SoCs with TLB Prefetching and MMU-Aware DMA Engine   |
| DR DRAM: Accelerating Memory-Read-Intensive Applications  |
| Puzzle Memory: Multifractional Partitioned Heterogeneous Memory Scheme  |
| Session 6B: Logic and Circuit Design 2  |
| Synchronization of Ring-Based Resonant Standing Wave Oscillators for 3D Clocking Applications   |
| Generalized Tree Architecture for Efficient Successive-Cancellation Polar Decoding  |
| Parameterized Posit Arithmetic Hardware Generator   |

| BGIM: Bit-Grained Instant-on Memory Cell for Sleep Power Critical Mobile Applications  | 42 |
|--|----|
| Autonomous Temperature Management through Selective Control of Exact-Approximate Tiles 34<br>Siyuan Xu (The University of Texas at Dallas, TX, USA) and Benjamin<br>Carrion Schafer (The University of Texas at Dallas, TX, USA)   | 46 |
| Session 7A: Accelerators and GPUs  |    |
| Automatic Mapping of the Sum-Product Network Inference Problem to FPGA-Based Accelerators 3: Lukas Sommer (TU Darmstadt, Germany), Julian Oppermann (TU Darmstadt, Germany), Alejandro Molina (TU Darmstadt, Germany), Carsten Binnig (TU Darmstadt, Germany), Kristian Kersting (TU Darmstadt, Germany), and Andreas Koch (TU Darmstadt, Germany) | 50 |
| BLPP: Improving the Performance of GPGPUs with Heterogeneous Memory through Bandwidth- and Latency-Aware Page Placement  | 58 |
| General IDS Acceleration for High-Speed Networks   | 66 |
| Scalable Multi-Queue Data Transfer Scheme for FPGA-Based Multi-Accelerators  | 74 |
| Session 7B: Potpouri 1   |    |
| Characterizing 3D Charge Trap NAND Flash: Observations, Analyses and Applications  | 81 |
| A Plain-Text Incremental Compression (PIC) Technique with Fast Lookup Ability  | 89 |
| Towards Efficient Microarchitecture Design of Simultaneous Localization and Mapping in  Augmented Reality Era  | 97 |

| Training Neural Networks with Low Precision Dynamic Fixed-Point  | 405 |
|--|-----|
| Decentralized Collaborative Power Management through Multi-Device Knowledge Sharing  | 409 |
| Session 8A: NVM  |     |
| Breeze: User-Level Access to Non-Volatile Main Memories for Legacy Software  | 413 |
| R-Cache: A Highly Set-Associative In-Package Cache Using Memristive Arrays   | 423 |
| A Highly Non-Volatile Memory Scalable and Efficient File System  Fan Yang (SKLSDE Lab, Beihang University, China; Beijing Advanced Innovation Center for Big Data and Brain Computing, Beijing, China), Junbin Kangy (Beijing Advanced Innovation Center for Big Data and Brain Computing), Shuai Ma (Beijing Advanced Innovation Center for Big Data and Brain Computing), and Jinpeng Huai (Beijing Advanced Innovation Center for Big Data and Brain Computing) | 431 |
| NVCool: When Non-Volatile Caches Meet Cold Boot Attacks  Xiang Pan (The Ohio State University), Anys Bacha (University of Michigan), Spencer Rudolph (The Ohio State University), Li Zhou (The Ohio State University), Yinqian Zhang (The Ohio State University), and Radu Teodorescu (The Ohio State University)  | 439 |
| Session 8B: Test and Verification  |     |
| Guiding RTL Test Generation Using Relevant Potential Invariants  | 449 |
| Back-End Layout Reflection for Test Chip Design  | 456 |
| How Multi-Threshold Designs Can Protect Analog IPs  Abdullah Ash- Saki (Pennsylvania State University, USA) and Swaroop  Ghosh (Pennsylvania State University, USA)  | 464 |

| Optimization of Mutant Space for RTL Test Generation  | 472 |
|---|-----|
| A Reliability Study on CNNs for Critical Embedded Systems  Mohamed A. Neggaz (LAMIH, Polytechnic University Hauts-De-France), Ihsen Alouani (IEMN, Polytechnic University Hauts-De-France), Pablo R. Lorenzo (The Silesian University of Technology, Poland), and Smail Niar (LAMIH, Polytechnic University Hauts-De-France)  | 476 |
| Session 9A: Network on Chip and Synchronization   |     |
| DEC-NoC: An Approximate Framework Based on Dynamic Error Control with Applications to Energy-Efficient NoCs  Yuechen Chen (The George Washington University), Md Farhadur Reza (The George Washington University), and Ahmad Louri (The George Washington University)   | 480 |
| RETUNES: Reliable and Energy-Efficient Network-on-Chip Architecture   | 488 |
| Accelerating Synchronization in Graph Analytics Using Moving Compute to Data Model on Tilera TILE-Gx72  Halit Dogan (University of Connecticut), Masab Ahmad (University of Connecticut), Jose Joao (Arm Research), and Omer Khan (University of Connecticut)   | 496 |
| Eca-Router: On Achieving Endpoint Congestion Aware Switch Allocation in the On-Chip Network  Cunlu Li (National University of Defense Technology, China), Dezun Dong (National University of Defense Technology, China), and Xiangke Liao (National University of Defense Technology, China)  | 506 |
| Accurate Performance Bounds Calculation for Dynamic Voltage-Freq Islands in Best Effort NoCs  Dara Rahmati (Institute for Research in Fundamental Sciences (IPM)), Sobhan Masoudi (Islamic Azad University, Central Tehran Branch), Ahmad Khonsari (University of Tehran, Institute for Research in Fundamental Sciences (IPM)), and Reza Sabbaghi-Nadooshan (Islamic Azad University, Central Tehran Branch) | 510 |
| Session 9B: Potpouri 2  |     |
| Design and Evaluation of a PVT Variation-Resistant TRNG Circuit   | 514 |
| Hardware-Based Probabilistic Threat Detection and Estimation for Embedded Systems   | 522 |

| Reverse Engineering of Split Manufactured Sequential Circuits Using Satisfiability Checking   | 530 |
|---|-----|
| Minimizing Thermal Variation in Heterogeneous HPC Systems with FPGA Nodes   | 537 |
| Session 10A: File System and Cloud  |     |
| A Compact AES Hardware Implementation Secure Against 1st-Order Side-Channel Attacks   | 545 |
| PFCG: Improving the Restore Performance of Package Datasets in Deduplication Systems  Chunxue Zuo (Wuhan National Lab for Optoelectronics, Huazhong University of Science and Technology), Fang Wang (Wuhan National Lab for Optoelectronics, Huazhong University of Science and Technology), Ping Huang (Department of Computer and Information Sciences, Temple University, USA), Yuchong Hu (Wuhan National Lab for Optoelectronics, Huazhong University of Science and Technology), Dan Feng (Wuhan National Lab for Optoelectronics, Huazhong University of Science and Technology), and Yucheng Zhang (Wuhan National Lab for Optoelectronics, Huazhong University of Science and Technology) | 553 |
| OME: An Optimized Modeling Engine for Disk Failure Prediction in Heterogeneous Datacenter Yanwen Xie (Huazhong University of Science and Technology, China), Dan Feng (Huazhong University of Science and Technology, China), Fang Wang (Huazhong University of Science and Technology, China), Xinyan Zhang (Huazhong University of Science and Technology, China), Jizhong Han (Chinese Academy of Sciences, China), and Xuehai Tang (Chinese Academy of Sciences, China)   | 561 |
| Enabling Accurate Performance Isolation on Hybrid Storage Devices in Cloud Environment  | 565 |

| LEA: A Lazy Eviction Algorithm for SSD Cache in Cloud Block Storage   |    |
|---|----|
| Session 10B: FPGA and Machine Learning  |    |
| 2000.002  |    |
| Fine-Grained Parallel Routing for FPGAs with Selective Expansion  | 7  |
| DEEP: Dedicated Energy-Efficient Approximation for Dynamically Reconfigurable   |    |
| Architectures   | 37 |
| Siyuan Xu (The University of Texas at Dallas, TX, USA) and Benjamin<br>Carrion Schafer (The University of Texas at Dallas, TX, USA) |    |
| Load Balance-Aware Multi-Core Parallel Routing for Large-Scale FPGAs  | )5 |
| Using Machine Learning to Predict Path-Based Slack from Graph-Based Timing Analysis   | )3 |
| Analysis of Row Hammer Attack on STTRAM   | A  |
| Author Index 61   | 3  |