2019 IEEE 37th International Conference on Computer Design (ICCD 2019)

Abu Dhabi, United Arab Emirates 17 – 20 November 2019



IEEE Catalog Number: CFP19ICD-POD ISBN:

978-1-7281-1215-2

Copyright \odot 2019 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:
 CFP19ICD-POD

 ISBN (Print-On-Demand):
 978-1-7281-1215-2

 ISBN (Online):
 978-1-5386-6648-7

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



2019 IEEE 37th International Conference on Computer Design (ICCD) ICCD 2019

Table of Contents

Message from General Chairs and Program Chair	
Organizing Committee	xx
Program Committee	xxi
Sub Reviewers	xxv
Sponsors	xxvii
Session 1A: Optimized Design Methodology	
Reading Between the Dies: Cross-SLR Covert Channels on Multi-Tenant Cloud FPGAs Ilias Giechaskiel (University of Oxford, United Kingdom), Kasper Rasmussen (University of Oxford, United Kingdom), and Jakub Szefer (Yale University, USA)	1
IPSA: Integer Programming via Sparse Approximation for Efficient Test-Chip Design Qicheng Huang (Carnegie Mellon University), Chenlei Fang (Carnegie Mellon University), Zeye Liu (Carnegie Mellon University), Ruizhou Ding (Carnegie Mellon University), and R. D. Shawn Blanton (Carnegie Mellon University)	11
Stealthy Rootkits in Smart Grid Controllers	20
Efficient Scalable Three Operand Multiplier Over GF(2^m) Based on Novel Decomposition Chiou-Yng Lee (Lunghwa University of Science & Technology) and Jiafeng Xie (Villanova University)	Strategy29
Session 1B: Accelerators and Machine Learning I	
An FPGA Implementation of Stochastic Computing-Based LSTM	38

Process Variation Mitigation on Convolutional Neural Network Accelerator Architecture	7
Characterizing On-Chip Traffic Patterns in General-Purpose GPUs: A Deep Learning Approach	5
AccUDNN: A GPU Memory Efficient Accelerator for Training Ultra-Deep Neural Networks	5
Session 2A: Architectural Advances for IoT Applications	
Session 2A: Architectural Advances for IoT Applications A Distributed Scheme for Accelerating Semantic Video Segmentation on An Embedded Cluster	3
A Distributed Scheme for Accelerating Semantic Video Segmentation on An Embedded Cluster	
A Distributed Scheme for Accelerating Semantic Video Segmentation on An Embedded Cluster	2

Session 2B: Architecture & Compilers II

Bandwidth-Aware Last-Level Caching: Efficiently Coordinating Off-Chip Read and Write Bandwidth
To Update or Not To Update?: Bandwidth-Efficient Intelligent Replacement Policies for DRAM Caches 119 Vinson Young (Georgia Institute of Technology) and Moinuddin K. Qureshi (Georgia Institute of Technology)
Compiler-Assisted Selection of Hardware Acceleration Candidates from Application Source Code
A New Traffic Offloading Method with Slow Switching Optical Device in Exascale Computer
Session 3A: Advances in the Design and Implementation of Neural Networks
ReRAM Crossbar-Based Analog Computing Architecture for Naive Bayesian Engine
HBUNN - Hybrid Binary-Unary Neural Network: Realizing a Complete CNN on an FPGA
Hardware Acceleration of Multilayer Perceptron Based on Inter-Layer Optimization

NR-MPA: Non-Recovery Compression Based Multi-Path Packet-Connected-Circuit Architecture of Convolution Neural Networks Accelerator
A Memory-Access-Efficient Adaptive Implementation of kNN on FPGA through HLS
Session 3B: Storage Systems
Adaptive Write Interference Management with Efficient Mapping for Shingled Recording Disks
RepEC-Duet: Ensure High Reliability and Performance for Deduplicated and Delta-Compressed Storage Systems
Lifelong Disk Failure Prediction via GAN-Based Anomaly Detection
An Efficient and Flexible Metadata Management Layer for Local File Systems
Session 4A: Innovation on Safety and Security for Robust Real-Time Systems
Security-Driven Codesign with Weakly-Hard Constraints for Real-Time Embedded Systems

Software Timing Analysis for Complex Hardware with Survivability and Risk Analysis Sergi Vilardell (Universitat Politecnica de Catalunya, Barcelona Supercomputing Center), Isabel Serra (Centre de Recerca Matematica, Barcelona Supercomputing Center), Jaume Abella (Barcelona Supercomputing Center), Joan Del Castillo (Universitat Autonoma de Barcelona), and Francisco J. Cazorla (Barcelona Supercomputing Center)	. 227
Integrating Cyber-Attack Defense Techniques into Real-Time Cyber-Physical Systems Xiaochen Hao (Northeastern University, China), Mingsong Lv (Northeastern University, China), Jiesheng Zheng (Northeastern University, China), Zhengkui Zhang (Northeastern University, China), and Wang Yi (Uppsala University, Sweden)	. 237
Astraea: Self-Balancing Federated Learning for Improving Classification Accuracy of Mobile Deep Learning Applications	. 246
A Comparative Analysis on the Impact of Bank Contention in STT-MRAM and SRAM Based LLCs	. 255
NVDL-Cache: Narrow-Width Value Aware Variable Delay Low-Power Data Cache Nezam Rohbani (Institute for Research in Fundamental Sciences (IPM)), Tapas Kumar Maiti (Hiroshima University), Dondee Navarro (Hiroshima University), Mitiko Miura-Mattausch (Hiroshima University), Hans Jürgen Mattausch (Hiroshima University), and Hirotaka Takatsuka (Mie Fujitsu Semiconductor Limited)	. 264
SpecLock: Speculative Lock Forwarding	. 273
Fine-Grained Management of Thread Blocks for Irregular Applications	. 283

Session 5a: IoT Covert Channel Attacks in Air-Gapped Networks
**
Red Teaming a Multi-Colored Bluetooth Bulb
Bridging The Gap: Data Exfiltration In Highly Secured Environments Using Bluetooth IoTs
Covert Data Exfiltration Using Light and Power Channels
Session 5B: Miscellaneous topics in Test, Verification, and Security
Post-Model Validation of Victim DRAM Caches
Revisiting Capacitor-Based Trojan Design
Formal Modeling and Verification of NAND Flash Memory Supporting Advanced Operations
SCARF: Detecting Side-Channel Attacks at Real-Time using Low-Level Hardware Features
Session 6a: Hardware and Software Implementations for Efficient Post Quantum Cryptography
SaberX4: High-Throughput Software Implementation of Saber Key Encapsulation Mechanism
Advances and Challenges of Rank Metric Cryptography Implementations

(Technology Innovation Institute), Rusydi Makarim (Technology Innovation Institute), Marc Manzano (Technology Innovation Institute), Chiara Marcolla (Technology Innovation Institute), and Victor Mateu

(Technology Innovation Institute)

Flexible NTT Accelerators for RLWE Lattice-Based Cryptography
PlaidML-HE: Acceleration of Deep Learning Kernels to Compute on Encrypted Data
Power, Area, Speed, and Security (PASS) Trade-Offs of NIST PQC Signature Candidates Using a C to ASIC Design Flow
Session 6B: Architecture & Compilers I
TicToc: Enabling Bandwidth-Efficient DRAM Caching for Both Hits and Misses in Hybrid Memory Systems . 341 Vinson Young (Georgia Institute of Technology), Zeshan A. Chishti (Intel), and Moinuddin K. Qureshi (Georgia Institute of Technology)
Static Function Prefetching for Efficient Code Management on Scratchpad Memory
Low Power Design through Frequency-Optimized Runtime Micro-Architectural Adaptation
HiNUMA: NUMA-Aware Data Placement and Migration in Hybrid Memory Systems 367 Zhuohui Duan (Huazhong University of Science and Technology, China), Haikun Liu (Huazhong University of Science and Technology, China), Xiaofei Liao (Huazhong University of Science and Technology, China), Hai Jin (Huazhong University of Science and Technology, China), Wenbin Jiang (Huazhong University of Science and Technology, China), and Yu Zhang (Huazhong University of Science and Technology, China)
Session 7A: Advances in Power, Thermal and Timing Aware Optimization
FPGA Energy Efficiency by Leveraging Thermal Margin Behnam Khaleghi (University of California San Diego), Sahand Salamat (University of California San Diego), Mohsen Imani (University of California San Diego), and Tajana Rosing (University of California San Diego)

A Memetic Algorithm Based PVT Variation-Aware Robust Transistor Sizing Scheme for Power-Delay Optimal Digital Standard Cell Design
CSM-NN: Current Source Model Based Logic Circuit Simulation - A Neural Network Approach
Exploiting the Benefits of High-Level Synthesis for Thermal-Aware VLSI Design
Efficient Linear System Solution Techniques in the Simulation of Large Dense Mutually Inductive Circuits
System-Level Optimization of Network-on-Chips for Heterogeneous 3D System-on-Chips
Lennart Bamberg (Universität Bremen, Germany), Alberto García Oritz (Universität Bremen, Germany), and Thilo Pionteck (Otto-von-Guericke-Universität Magdeburg, Germany)
(Universität Bremen, Germany), and Thilo Pionteck
(Universität Bremen, Germany), and Thilo Pionteck (Otto-von-Guericke-Universität Magdeburg, Germany)
(Universität Bremen, Germany), and Thilo Pionteck (Otto-von-Guericke-Universität Magdeburg, Germany) Session 7B: EDA for Non-Logic Issues and Non-CMOS Technologies Constraint-Programmed Initial Sizing of Analog Operational Amplifiers 413 Inga Abel (Technical University of Munich), Maximilian Neuner (Technical University of Munich), and Helmut Graeb (Technical
(Universität Bremen, Germany), and Thilo Pionteck (Otto-von-Guericke-Universität Magdeburg, Germany) Session 7B: EDA for Non-Logic Issues and Non-CMOS Technologies Constraint-Programmed Initial Sizing of Analog Operational Amplifiers

qCG: A Low-Power Multi-Domain SFQ Logic Design and Verification Framework Shahin Nazarian (University of Southern California, USA), Arash Fayyazi (University of Southern California, USA), and Massoud Pedram (University of Southern California, USA)	446
Session 8a: Robust Hardware Design with Machine Learning	
A Learning-Based Framework for Automatic Parameterized Verification Yongjian Li (Chinese Academy of Sciences, Institute of Software, The State Key Laboratory of Computer Science, Beijing, China), Jialun Cao (Chinese Academy of Sciences, Institute of Software, The State Key Laboratory of Computer Science, Beijing, China), and Jun Pang (University of Luxembourg, FSTC & SnT, Esch-sur-Alzette, Luxembourg)	. 450
Learning-Based Diversity Estimation: Leveraging the Power of High-Level Synthesis to Mitigate Common-Mode Failure Farah Naz Taher (The University of Texas at Dallas, USA), Anjana Balachandran (The Hong Kong Polytechnic University), and Benjamin Carrion Schafer (The University of Texas at Dallas, USA)	460
Soft Error Resilience in Chip Multiprocessor Cache using a Markov Model Based Re-usability Predictor Avishek Choudhury (New Alipore College) and Biplab K Sikdar (IIEST Shibpur)	468
CHASE: A Customized Time Series Machine Learning Approach for Hardware-Based Stealthy Malware Detection Hossein Sayadi (California State University, Long Beach, USA), Yifeng Gao (George Mason University, USA), Hosein Mohammadi Makrani (University of California, Davis, USA), Sai Manoj P D (George Mason University, USA), Avesta Sasan (George Mason University, USA), Jessica Lin (George Mason University, USA), Setareh Rafatirad (George Mason University, USA), and Houman Homayoun (University of California, Davis, USA)	N/A
Session 8B: Accelerators and Machine Learning II	
ReNEW: Enhancing Lifetime for ReRAM Crossbar Based Neural Network Accelerators Wen Wen (University of Pittsburgh), Youtao Zhang (University of Pittsburgh), and Jun Yang (University of Pittsburgh)	487
Ebird: Elastic Batch for Improving Responsiveness and Throughput of Deep Learning Services Weihao Cui (Shanghai Jiao Tong University, China), Mengze Wei (Shanghai Jiao Tong University, China), Quan Chen (Shanghai Jiao Tong University, China), Xiaoxin Tang (Shanghai University of Finance and Economics, China), Jingwen Leng (Shanghai Jiao Tong University, China), Li Li (Shanghai Jiao Tong University, China), and Mingyi Guo (Shanghai Jiao Tong University, China)	. 497

When Deep Learning Meets the Edge: Auto-Masking Deep Neural Networks for Efficient Machine Learning on Edge Devices	06
A High-Performance Processing-in-Memory Accelerator for Inline Data Deduplication	15
Session 9A: Innovative Circuits for Improved Power Consumption and Performance Characteristics	
Low Power Design of Runtime Reconfigurable FPGAs through Contexts Approximations	24
Applying Swarm Intelligence to Distributed On-Chip Power Management	32
FeMAT: Exploring In-Memory Processing in Multifunctional FeFET-Based Memory Array	41
Threshold Logic in a Flash	50
Session 9B: Hardware Security	
Adaptive Masking: a Dynamic Trade-off between Energy Consumption and Hardware Security	59
Cyclic Beneš Network Based Logic Encryption for Mitigating SAT-Based Attacks	67

PUF-RLA: A PUF-Based Reliable and Lightweight Authentication Protocol Employing Binary String Shuffling	76
AdapTimer: Hardware/Software Collaborative Timer Resistant to Flush-Based Cache Attacks on ARM-FPGA Embedded SoC	35
Session 10A: Miscellaneous Topics in Computer Systems I	
Architectural and Cost Implications of the 5G Edge NFV Systems	94
Imbalance-Aware Scheduler for Fast and Secure Ring ORAM Data Retrieval)4
SPA-SSD: Exploit Heterogeneity and Parallelism of 3D SLC-TLC Hybrid SSD to Improve Write Performance 61 Wenhui Zhang (Huazhong University of Science and Technology), Qiang Cao (Huazhong University of Science and Technology), Hong Jiang (University of Texas at Arlington), Jie Yao (Huazhong University of Science and Technology), Yuanyuan Dong (Alibaba Group.), and Puyuan Yang (Alibaba Group.)	13
Mitigating Application Diversity for Allocating a Unified ACC-Rich Platform 62 Jinghan Zhang (Northeastern University, USA), Hamed Tabkhi (University of North Carolina at Charlotte, USA), and Gunar Schirner (Northeastern University, USA)	22
VNet: A Versatile Network for Efficient Real-Time Semantic Segmentation 62 Ning Lin (nstitute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Hang Lu (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jingliang Gao (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Shunjie Qiao (University of Hong Kong), and Xiaowei Li (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences)	26
Session 10B: Machine Learning Techniques for Innovative Energy-Efficient Solutions	
Energy Prediction for Cache Tuning in Embedded Systems	30

Power Management of Wireless Sensor Nodes with Coordinated Distributed Reinforcement Learning
Dynamic Optimization of Battery Health in IoT Networks
Reduced Precision for Energy Efficient FPGA
BRASIL: A High-Integrity GPGPU Toolchain for Automotive Systems
Session 11A: Miscellaneous Topics in Computer Systems II
Balancing Performance and Energy Efficiency of ONoC by Using Adaptive Bandwidth 664 Mingzhe Zhang (State Key Laboratory of Computer Architecture, ICT, CAS, Beijing, China), Lunkai Zhang (Intel Corporation), Frederic T. Chong (University of Chicago), and Zhiyong Liu (State Key Laboratory of Computer Architecture, ICT, CAS, Beijing, China)
Valler: Threshold Voltage Distribution Aware LLR Optimization to Improve LDPC Decoding Performance for 3D TLC NAND Flash
Reinforce Memory Error Protection by Breaking DRAM Disturbance Correlation Within ECC Words
Archivist: A Machine Learning Assisted Data Placement Mechanism for Hybrid Storage Systems
Detecting and Predicting Performance Degradation Caused by Impaired Cache Isolation

Session 11B: Processor and Memory Architectures

A Case for Software-Based Adaptive Routing in NUMA Systems	684
Value Speculation through Equality Prediction	694
A Novel Convolutional Neural Network Accelerator That Enables Fully-Pipelined Execution of Layers Donghyun Kang (Seoul National University, Republic of Korea), Jintaek Kang (Seoul National University, Republic of Korea), Hyungdal Kwon (Samsung Electronics, Republic of Korea), Hyunsik Park (Samsung Electronics, Republic of Korea), and Soonhoi Ha (Seoul National University, Republic of Korea)	698
Freeflow Core: Enhancing Performance of In-Order Cores with Energy Efficiency Raj Kumar Choudhary (Indian Institute of Technology Bombay), Newton Singh (Indian Institute of Technology Bombay), Harideep Nair (Indian Institute of Technology Bombay), Rishabh Rawat (Indian Institute of Technology Bombay), and Virendra Singh (Indian Institute of Technology Bombay)	702
Author Index	707