2020 IEEE 31st International Conference on Application-specific Systems, **Architectures and Processors** (ASAP 2020)

Manchester, United Kingdom 6-8 July 2020



IEEE Catalog Number: CFP20063-POD **ISBN:**

978-1-7281-7279-8

Copyright © 2020 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.

P 2
3-1
3-1
50-

CFP20063-POD 978-1-7281-7279-8 978-1-7281-7147-0 2160-0511

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



2020 IEEE 31st International Conference on Applicationspecific Systems, Architectures and Processors (ASAP) ASAP 2020

Table of Contents

Message from the Conference Chairs x
Organizing Committee xii
Steering Committee xiii
Technical Program Committee xiv
External Referees xvi

Session 1: Heterogeneous Computing, Manycore Systems, Reconfigurable Accelerators

A Template-Based Framework for Exploring Coarse-Grained Reconfigurable Architectures .1 Artur Podobas (KTH Royal Institute of Technology, RIKEN Center for Computational Science), Kentaro Sano (RIKEN Center for Computational Science), and Satoshi Matsuoka (RIKEN Center for Computational Science, Tokyo Institute of Technology)
Accelerating Radiative Transfer Simulation with GPU-FPGA Cooperative Computation .9 Ryohei Kobayashi (University of Tsukuba), Norihisa Fujita (University of Tsukuba), Yoshiki Yamaguchi (University of Tsukuba), Taisuke Boku (University of Tsukuba), Kohji Yoshikawa (University of Tsukuba), Makito Abe (University of Tsukuba), and Masayuki Umemura (University of Tsukuba)
Termination Detection for Fine-Grained Message-Passing Architectures .17 Matthew Naylor (University of Cambridge), Simon W. Moore (University of Cambridge), Andrey Mokhov (Newcastle University and Jane Street), David Thomas (Imperial College London), Jonathan R. Beaumont (Imperial College London), Shane Fleming (Microsoft Research), Theodore Markettos (University of Cambridge), Thomas Bytheway (University of Cambridge), and Andrew Brown (University of Southampton)
Condensing an Overload of Parallel Computing Ingredients into a Single Architecture Recipe.25 Riadh Ben Abdelhamid (University of Tsukuba), Yoshiki Yamaguchi

(University of Tsukuba), and Taisuke Boku (University of Tsukuba)

FPGA-Based Network Microburst Analysis System with Flow Specification and Efficient Packet Capturing 29 Shuhei Yoshida (NTT Corporation), Yuta Ukon (NTT Corporation), Shoko Ohteru (NTT Corporation), Hiroyuki Uzawa (NTT Corporation), Namiko Ikeda (NTT Corporation), and Koyo Nitta (NTT Corporation)
FPGA-Accelerated Time Series Mining on Low-Power IoT Devices .33 Seongyoung Kang (Kookmin University), Jinyeong Moon (Florida State University), and Sang-Woo Jun (University of California, Irvine)
Session 2: Machine Learning and Acceleration of Neural Networks
Array Aware Training/Pruning: Methods for Efficient Forward Propagation on Array-Based Neural Network Accelerators .37 Krishna Teja Chitty-Venkata (Iowa State University) and Arun K. Somani (Iowa State University)
Design Space Exploration for Softmax Implementations .45 Zhigang Wei (The University of Texas at Austin), Aman Arora (The University of Texas at Austin), Pragenesh Patel (The University of Texas at Austin), and Lizy John (The University of Texas at Austin)
Hamamu: Specializing FPGAs for ML Applications by Adding Hard Matrix Multiplier Blocks .53 Aman Arora (The University of Texas at Austin), Zhigang Wei (The University of Texas at Austin), and Lizy K. John (The University of Texas at Austin)
Hardware Acceleration of Large Scale GCN Inference .61 Bingyi Zhang (University of Southern California), Hanqing Zeng (University of Southern California), and Viktor Prasanna (University of Southern California)
Training Neural Nets using only an Approximate Tableless LNS ALU .69 Mark Arnold (XLNS Research), Ed Chester (Goonhilly Earth Station), and Corey Johnson (XLNS Research)
Temporal Motionless Analysis of Video using CNN in MPSoC .73 Somdip Dey (University of Essex), Amit Kumar Singh (University of Essex), Dilip Kumar Prasad (UiT The Arctic University of Norway), and Klaus McDonald-Maier (University of Essex)
An Efficient Convolution Engine Based on the À-Trous Spatial Pyramid Pooling .77 Cristian Sestito (University of Calabria), Fanny Spagnolo (University of Calabria), Pasquale Corsonello (University of Calabria), and Stefania Perri (University of Calabria)
Fast and Accurate Training of Ensemble Models with FPGA-Based Switch .81 Jiuxi Meng (Imperial College London), Ce Guo (Imperial College London), Nadeen Gebara (Imperial College London), and Wayne Luk (Imperial College London)

vi

Session 3: Emerging Technologies and Neuromorphic Computing

Persistent Fault Analysis of Neural Networks on FPGA-Based Acceleration System .85 Dawen Xu (Chinese Academy of Sciences, Hefei University of Technology), Ziyang Zhu (Hefei University of Technology), Cheng Liu (Chinese Academy of Sciences), Ying Wang (Chinese Academy of Sciences), Huawei Li (Chinese Academy of Sciences), Lei Zhang (Chinese Academy of Sciences), and Kwang-Ting Cheng (Hong Kong University of Science and Technology)
A Parallel-Friendly Majority Gate to Accelerate In-Memory Computation .93 John Reuben (Friedrich-Alexander-Universität Erlangen-Nürnberg) and Stefan Pechmann (Universität Bayreuth)
A System for Generating Non-Uniform Random Variates using Graphene Field-Effect Transistors .101
Nathaniel J. Tye (University of Cambridge), James T. Meech (University of Cambridge), Bilgesu A. Bilgin (University of Cambridge), and Phillip Stanley-Marbell (University of Cambridge)
Efficient FeFET Crossbar Accelerator for Binary Neural Networks .109 Taha Soliman (Bosch GmbH), Ricardo Olivo (Fraunhofer IPMS), Tobias Kirchner (Bosch GmbH), Cecilia De la Parra (Bosch GmbH), Maximilian Lederer (Fraunhofer IPMS), Thomas Kämpfe (Fraunhofer IPMS), Andre Guntoro (Bosch GmbH), and Norbert Wehn (TU Kaiserslautern)
A Design Methodology for Post-Moore's Law Accelerators: The Case of a Photonic Neuromorphic Processor .113 Armin Mehrabian (The George Washington University), Volker J. Sorger (The George Washington University), and Tarek El-Ghazawi (The George Washington University)
Improved Side-Channel Resistance by Dynamic Fault-Injection Countermeasures .117 Jan Richter-Brockmann (Ruhr-Universität Bochum) and Tim Güneysu (Ruhr-Universität Bochum, DFKI)

Session 4: Computing in the Cloud and Datacenters

Architecture Support for FPGA Multi-Tenancy in the Cloud .125 Joel Mandebi Mbongue (University of Florida), Alex Shuping (University of Florida), Pankaj Bhowmik (University of Florida), and Christophe Bobda (University of Florida)
FPGAs in the Datacenters: the Case of Parallel Hybrid Super Scalar String Sample Sort .1.3 Mikhail Asiatici (EPFL), Damian Maiorano (Politecnico di Torino), and Paolo Ienne (EPFL)
SLATE: Managing Heterogeneous Cloud Functions .141 Jessica Vandebon (Imperial College London), Jose G. F. Coutinho (Imperial College London), Wayne Luk (Imperial College London), Eriko Nurvitadhi (Intel Corporation), and Mishali Naik (Intel Corporation)

Session 5: Approximate Computing and Computer Arithmetic

Reconfigurable Stream-Based Tensor Unit with Variable-Precision Posit Arithmetic .149 Nuno Neves (INESC-ID), Pedro Tomás (INESC-ID, Instituto Superior Técnico, Universidade de Lisboa), and Nuno Roma (INESC-ID, Instituto Superior Técnico, Universidade de Lisboa)
Anytime Floating-Point Addition and Multiplication – Concepts and Implementations .157 Marcel Brand (Friedrich-Alexander University Erlangen-Nurnberg), Michael Witterauf (Friedrich-Alexander University Erlangen-Nurnberg), Alberto Bosio (Ecole Centrale de Lyon, Institut des Nanotechnologies de Lyon), and Jürgen Teich (Friedrich-Alexander University Erlangen-Nurnberg)
BWOLF: Bit-Width Optimization for Statistical Divergence with Logarithmic Functions .165 Qian Xu (University of Maryland), Guowei Sun (University of Maryland), and Gang Qu (University of Maryland)
Efficient Floating-Point Implementation of the Probit Function on FPGAs .173 Mioara Joldes (CNRS, LAAS-CNRS) and Bogdan Pasca (Intel Corporation)
Combining Fixed-Point and SORN Arithmetic in a MIMO BPSK-Symbol Detection Architecture .183 Moritz Bärthel (University of Bremen), Jochen Rust (University of Bremen), and Steffen Paul (University of Bremen)
ParaHist: FPGA Implementation of Parallel Event-Based Histogram for Optical Flow Calculation .185 Mohammad Pivezhandi (Iowa State Uniersity), Phillip H. Jones (Iowa State Uniersity), and Joseph Zambreno (Iowa State Uniersity)

Session 6: Edge Computing

Optimizing Grouped Convolutions on Edge Devices 189. Perry Gibson (University of Glasgow), José Cano (University of Glasgow), Jack Turner (University of Edinburgh), Elliot J. Crowley (University of Edinburgh), Michael O'Boyle (University of Edinburgh), and Amos Storkey (University of Edinburgh)
Dynamic Sharing in Multi-Accelerators of Neural Networks on an FPGA Edge Device .197 Hsin-Yu Ting (University of California, Irvine), Tootiya Giyahchi (University of California, Irvine), Ardalan Amiri Sani (University of California, Irvine), and Eli Bozorgzadeh (University of California, Irvine)
A New Hardware Approach to Self-Organizing Maps 205 Leonardo A. Dias (Federal University of Rio Grande do Norte), Maria G. F. Coutinho (Federal University of Rio Grande do Norte), Elena Gaura (Coventry University), and Marcelo A. C. Fernandes (Federal University of Rio Grande do Norte)
Low-Cost DNN Hardware Accelerator for Wearable, High-Quality Cardiac Arrythmia Detection .213 Johnson Loh (Rheinisch-Westfälische Technische Hochschule Aachen), Jianan Wen (Rheinisch-Westfälische Technische Hochschule Aachen), and Tobias Gemmeke (Rheinisch-Westfälische Technische Hochschule Aachen)

Author Index 217