

2024 IEEE 32nd Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM 2024)

**Orlando, Florida, USA
5-8 May 2024**



**IEEE Catalog Number: CFP24054-POD
ISBN: 979-8-3503-7244-1**

**Copyright © 2024 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:	CFP24054-POD
ISBN (Print-On-Demand):	979-8-3503-7244-1
ISBN (Online):	979-8-3503-7243-4
ISSN:	2576-2613

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

CURRAN ASSOCIATES INC.
proceedings
.com

2024 IEEE 32nd Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM) **FCCM 2024**

Table of Contents

Message from General Chairs	xii
Message from Program Chairs	xiii
Organizing Committee	xiv
Program Committee	xv
Sponsors	xvii

Applications

Bandwidth Efficient Homomorphic Encrypted Discrete Fourier Transform Acceleration on FPGA 1 <i>Zhihan Xu (University of Southern California), Yang Yang (University of Southern California), Rajgopal Kannan (DEVCOM Army Research Office), and Viktor K. Prasanna (University of Southern California)</i>	
High Throughput Massive MIMO Signal Decoding Using Multi-Level Tree Search on FPGAs 13 <i>Mohamed Hassan (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), Hatem Ltaief (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), and Suhaib A. Fahmy (King Abdullah University of Science and Technology (KAUST), Saudi Arabia)</i>	
PCQ: Parallel Compact Quantum Circuit Simulation 24 <i>Shuang Liang (Imperial College London, UK), Yuncheng Lu (Imperial College London, UK), Ce Guo (Imperial College London, UK), Wayne Luk (Imperial College London, UK), and Paul H. J. Kelly (Imperial College London, UK)</i>	

Arithmetic

HiHiSpMV: Sparse Matrix Vector Multiplication with Hierarchical Row Reductions on FPGAs with High Bandwidth Memory 32 <i>Abdul Rehman Tareen (Paderborn University, Germany), Marius Meyer (Paderborn University, Germany), Christian Plessl (Paderborn University, Germany), and Tobias Kenter (Paderborn University, Germany)</i>	
--	--

Efficient 8-bit Matrix Multiplication on Intel Agilex-5 FPGAs	43
<i>Sergey Gribok (Intel Corporation, Unites States) and Bogdan Pasca (Intel Corporation, France)</i>	
Efficient Approaches for GEMM Acceleration on Leading AI-Optimized FPGAs	54
<i>Endri Taka (The University of Texas at Austin, USA), Dimitrios Gourounas (The University of Texas at Austin, USA), Andreas Gerstlauer (The University of Texas at Austin, USA), Diana Marculescu (The University of Texas at Austin, USA), and Aman Arora (Arizona State University, USA)</i>	

Machine Learning I

GCV-Turbo: End-to-end Acceleration of GNN-Based Computer Vision Tasks on FPGA	66
<i>Bingyi Zhang (University of Southern California), Rajgopal Kannan (DEVCOM Army Research Office), Carl Busart (DEVCOM Army Research Lab), and Viktor Prasanna (University of Southern California)</i>	
Accelerating ViT Inference on FPGA through Static and Dynamic Pruning	78
<i>Dhruv Parikh (University of Southern California), Shouyi Li (University of Southern California), Bingyi Zhang (University of Southern California), Rajgopal Kannan (DEVCOM Army Research Office), Carl Busart (DEVCOM Army Research Lab), and Viktor Prasanna (University of Southern California)</i>	
LAMPS: A Layer-Wised Mixed-Precision-and-Sparsity Accelerator for NAS-Optimized CNNs on FPGA	90
<i>Shuxin Yang (Southern University of Science and Technology, China), Chenchen Ding (Southern University of Science and Technology, China), Mingqiang Huang (Southern University of Science and Technology, China), Kai Li (Southern University of Science and Technology, China), Chenghao Li (Southern University of Science and Technology, China), Zikun Wei (Southern University of Science and Technology, China), Sixiao Huang (Southern University of Science and Technology, China), Jingyao Dong (Southern University of Science and Technology, China), Liuyang Zhang (Southern University of Science and Technology, China), and Hao Yu (Southern University of Science and Technology, China)</i>	
LGBM2VHDL: Mapping of LightGBM Models to FPGA	97
<i>Tomáš Martínek (Brno University of Technology, Czech Republic), Jan Kořenek (CESNET, z. s. p. o., Czech Republic), and Tomáš Čejka (CESNET, z. s. p. o., Czech Republic)</i>	

Compilation & CAD Tool

LightningSimV2: Faster and Scalable Simulation for High-Level Synthesis via Graph Compilation and Optimization	104
<i>Rishov Sarkar (Georgia Institute of Technology)</i>	

A Data-Driven, Congestion-Aware and Open-Source Timing-Driven FPGA Placer Accelerated by GPUs	115
<i>Zhili Xiong (The University of Texas at Austin, USA), Rachel Selina Rajarathnam (The University of Texas at Austin, USA), and David Z. Pan (The University of Texas at Austin, USA)</i>	
Synthesis of LUT Networks for Random-Looking Dense Functions with Don't Cares — Towards Efficient FPGA Implementation of DNN	126
<i>Yukio Miyasaka (University of California, USA), Alan Mishchenko (University of California, USA), John Wawrzynek (University of California, USA), and Nicholas J. Fraser (AMD Research and Advanced Development, Ireland)</i>	
A Routability-Driven Ultrascale FPGA Macro Placer with Complex Design Constraints	133
<i>Qin Luo (The Chinese University of Hong Kong), Xinshi Zang (The Chinese University of Hong Kong), Qijing Wang (The Chinese University of Hong Kong), Fangzhou Wang (The Chinese University of Hong Kong), Evangeline F.Y. Young (The Chinese University of Hong Kong), and Martin D.F. Wong (Hong Kong Baptist University)</i>	

Architecture and CGRA

HardCilk: Cilk-Like Task Parallelism for FPGAs	140
<i>Mohamed Shahawy (Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland), Canberk Sönmez (Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland), Cemalettin Belentepe (Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland), and Paolo Ienne (Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland)</i>	
Mapping Enumeration for Multi-context CGRAs Using Zero-Suppressed Binary Decision Diagrams....	151
<i>Rami Beidas (University of Toronto, Canada) and Jason H. Anderson (University of Toronto, Canada)</i>	
MPC-Wrapper: Fully Harnessing the Potential of Samsung Aquabolt-XL HBM2-PIM on FPGAs	162
<i>Jinwoo Choi (Yonsei University), Yeonan Ha (Yonsei University), Hanna Cha (Yonsei University), Seil Lee (Yonsei University), Sungchul Lee (Yonsei University), Joungwoo Lee (Yonsei University), Shin-haeng Kang (Samsung Electronics), Bongjun Kim (Samsung Advanced Institute of Technology), Hanwoong Jung (Samsung Advanced Institute of Technology), Hanjun Kim (Yonsei University), and Youngsok Kim (Yonsei University)</i>	
FINESSD: Near-Storage Feature Selection with Mutual Information for Resource-Limited FPGAs..	173
<i>Nikolaos Kyparissas (The University of Manchester), Gavin Brown (The University of Manchester), and Mikel Luján (The University of Manchester)</i>	

Machine Learning 2

SMOF: Streaming Modern CNNs on FPGAs with Smart Off-Chip Eviction	185
<i>Petros Toupas (Imperial College London; Information Technologies Institute, Centre of Research and Technology Hellas), Zhewen Yu (Imperial College London), Christos-Savvas Bouganis (Imperial College London), and Dimitrios Tzovaras (Information Technologies Institute, Centre of Research and Technology Hellas)</i>	
MRH-GCN: An Efficient GCN Accelerator for Multi-Relation Heterogeneous Graph	197
<i>Wenlu Peng (Sun Yat-sen University, China), Jianjun Chen (Sun Yat-sen University, China), Wenjin Huang (Sun Yat-sen University, China), and Yihua Huang (Sun Yat-sen University, China)</i>	
A Novel FPGA Accelerator of R(2+1)D	204
<i>Dehao Xiang (Sun Yat-sen University, China), Chenyang Li (Sun Yat-sen University, China), Wenjin Huang (Sun Yat-sen University, China), and Yihua Huang (Sun Yat-sen University, China)</i>	

Poster Session A

A Near-Sensor Image Processing Accelerator for Low-end FPGA Design	211
<i>Peter Mbuja (University of Florida, U.S.A), Zhaoqi Wang (University of Florida, U.S.A), Maximillian Panoff (University of Florida, U.S.A), and Christophe Bobda (University of Florida, U.S.A)</i>	
if-ZKP: Intel FPGA-Based Acceleration of Zero Knowledge Proofs	212
<i>Shahzad Ahmad Butt (Intel, USA), Benjamin Reynolds (Intel, USA), Veeraraghavan Ramamurthy (Intel, USA), Xiao Xiao (Intel, USA), Chu Pohrong (Intel, USA), Setareh Sharifian (Intel, USA), Sergey Gribok (Intel, USA), and Bogdan Pasca (Intel, USA)</i>	
FPGA Benchmark for Unrolled DNNs with Fine-Grained Sparsity and Mixed Precision	213
<i>Xilai Dai (Cornell University), Yuzong Chen (Cornell University), and Mohamed Abdelfattah (Cornell University)</i>	
SDAcc: A Stable Diffusion Accelerator on FPGA via Software-Hardware Co-Design	214
<i>Hao Zhou (Fudan University, China), Yang Liu (Fudan University, China), Hongji Wang (Fudan University, China), Enhao Tang (Fuzhou University, China), Shun Li (Fuzhou University, China), Yifan Zhang (Fudan University, China), and Kun Wang (Fudan university, China)</i>	
Learned Index Acceleration with FPGAs: A SMART Approach	215
<i>Geetesh More (University of New Brunswick), Suprio Ray (University of New Brunswick), and Kenneth B. Kent (University of New Brunswick)</i>	
SpGCN: An FPGA-Based Graph Convolutional Network Accelerator for Sparse Graphs	216
<i>Xiangzhi Xu (Sun Yat-sen University, China), Qi Liu (Sun Yat-sen University, China), Wenjin Huang (Sun Yat-sen University, China), WenLu Peng (Sun Yat-sen University, China), and Yihua Huang (Sun Yat-sen University, China)</i>	
HBMorphic: FHE Acceleration via HBM-Enabled Recursive Karatsuba Multiplier on FPGA	217
<i>Hassan Nassar (Karlsruhe Institute of Technology, Germany), Lars Bauer (Karlsruhe Institute of Technology, Germany), and Jörg Henkel (Karlsruhe Institute of Technology, Germany)</i>	

Poster Session B

OS4C: An Open-Source SR-IOV System for SmartNIC-Based Cloud Platforms	218
<i>Scott Smith (UIUC), Yuan Ma (UIUC), Marissa Lanz (UIUC), Bill Dai (UIUC), Martin Ohmacht (IBM), Bharat Sukhwani (IBM), Hubertus Franke (IBM), Volodymyr Kindratenko (UIUC), and Deming Chen (UIUC)</i>	
High-Performance Reconfigurable Accelerator for Knowledge Graph Reasoning	219
<i>Hanning Chen (University of California, Irvine), Ali Zakeri (University of California, Irvine), Yang Ni (University of California, Irvine), Fei Wen (Texas A&M University), Behnam Khaleghi (Qualcomm), Hugo Lapepie (Cisco Systems), and Mohsen Imani (University of California, Irvine)</i>	
BitSys: Bitwise Systolic Array Architecture for Multi-Precision Quantized Hardware Accelerators	220
<i>Yuhao Liu (Dresden University of Technology, Germany; ScaDS.AI Dresden/Leipzig, Germany), Salim Ullah (Ruhr University Bochum, Germany), and Akash Kumar (Ruhr University Bochum, Germany; ScaDS.AI Dresden/Leipzig, Germany)</i>	
RingTK: A Ring, Parallel and High Performance Top-K Sorter on FPGA	221
<i>Huawen Liang (University of Science and Technology of China, China), Qizhe Wu (University of Science and Technology of China, China), Wei Yuan (University of Science and Technology of China, China), Teng Tian (University of Science and Technology of China, China), and Xi Jin (University of Science and Technology of China, China)</i>	
Stay Flexible: A High-Performance FPGA NPU Overlay for Graph Neural Networks	222
<i>Taikun Zhang (University of Toronto, Canada; Intel Corporation), Andrew Boutros (University of Toronto, Canada), Sergey Gribok (Intel Corporation), Kwadwo Boateng (Intel Corporation), and Vaughn Betz (University of Toronto, Canada)</i>	
The BRAM is the Limit: Shattering Myths, Shaping Standards, and Building Scalable PIM Accelerators	223
<i>Arafat Kabir (University of Arkansas), Tendayi Kamucheka (University of Arkansas), Nathaniel Fredricks (University of Arkansas), Joel Mandebi (Advanced Micro Devices, Inc. (AMD)), Jason Bakos (University of South Carolina), Miaoqing Huang (University of Arkansas), and David Andrews (University of Arkansas)</i>	
Etna: MLIR-Based System-Level Design and Optimization for Transparent Application Execution on CPU-FPGA Nodes	224
<i>Stephanie Soldavini (Politecnico di Milano, Italy), Felix Suchert (TU Dresden, Germany), Serena Curzel (Politecnico di Milano, Italy), Michele Fiorito (Politecnico di Milano, Italy), Karl Friebe (TU Dresden, Germany), Fabrizio Ferrandi (Politecnico di Milano, Italy), Radim Cmar (Sygic, Slovakia), Jeronimo Castrillon (TU Dresden, Germany), and Christian Pilato (Politecnico di Milano, Italy)</i>	

PhD Forum Session

Ph.D. Project: Optimizing the Data Traffic for Large Graph Processing on FPGA via a Stateful Approach	225
<i>Qianyu Cheng (University of Science and Technology of China, China), Teng Wang (University of Science and Technology of China, China), and Chao Wang (University of Science and Technology of China, China)</i>	
PhD Project: Reconfigurable Network on Chip Architecture Through Topology Obfuscation For Protecting SoC Against Reverse Engineering	227
<i>Dipal Halder (University of Florida, Florida) and Sandip Ray (University of Florida, Florida)</i>	
Ph.D. Project: Field-Programmable Computing for Bayesian Network Structure Learning	229
<i>Ryota Miyagi (The University of Tokyo, Japan) and Hideki Takase (The University of Tokyo, Japan)</i>	
Ph.D. Project: Achieving Low-Latency Acceleration on Multi-FPGA for GPT Application	231
<i>ZhenDong Zheng (University of Science and Technology of China), Teng Wang (University of Science and Technology of China), and Chao Wang (University of Science and Technology of China)</i>	
Ph.D. Project: Investigating oneAPI Performance & Scalability using Feedforward FFT Architectures	233
<i>James Bickerstaff (University of Pittsburgh, USA) and Alan D. George (University of Pittsburgh, USA)</i>	
Ph.D. Project: Accelerated Machine Learning for On-Orbit OPIR Target Detection	235
<i>Daniel C. Stumpp (University of Pittsburgh, USA) and Alan D. George (University of Pittsburgh, USA)</i>	
Ph.D. Project: A Compiler-Driven Approach to HW/SW Co-Design of Deep-Learning Accelerators	237
<i>Tendayi Kamucheka (University of Arkansas, Arkansas) and David Andrews (University of Arkansas, Arkansas)</i>	
Ph.D. Project: ManAge: A Tool for Timing Characterization of FPGAs	239
<i>Bardia Babaei (The University of Manchester, UK) and Dirk Koch (The University of Manchester, UK; Heidelberg University, DE)</i>	
Ph.D. Project: Compiler-Driven Hardware/Software Co-Design for Embedded AI	241
<i>Patrick Schmidt (Karlsruhe Institute of Technology) and Jürgen Becker (Karlsruhe Institute of Technology)</i>	
Ph.D. Project Investigating Chiplet Interfaces for Efficient Near-Sensor Computing in Visual On-Device Intelligence	243
<i>Peter Mbua (University of Florida, Florida) and Christophe Bobda (University of Florida, Florida)</i>	
Ph.D. Project: Modernizing High-Level Hardware Design Workflows	245
<i>Rishov Sarkar (Georgia Institute of Technology)</i>	
Ph.D. Project - IsoFPGA - A Novel CMOS Galvanic Isolation Technique for Remote Physical Attacks in Multi-Tenant Cloud FPGA	247
<i>Muhammed Kawser Ahmed (University of Florida, USA) and Christophe Bobda (University of Florida, USA)</i>	
Ph.D. Project: Systems-on-Chip to Implement Zero-Trust Architectures	249
<i>Abigail Butka (University of Florida, Florida) and Christophe Bobda (University of Florida, Florida)</i>	

Author Index **251**