

**2024 IEEE 35th International
Conference on
Application-specific Systems,
Architectures and Processors
(ASAP 2024)**

**Hong Kong
24-26 July 2024**



**IEEE Catalog Number: CFP24063-POD
ISBN: 979-8-3503-4964-1**

**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:	CFP24063-POD
ISBN (Print-On-Demand):	979-8-3503-4964-1
ISBN (Online):	979-8-3503-4963-4
ISSN:	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

CURRAN ASSOCIATES INC.
proceedings
.com

2024 IEEE 35th International Conference on Application- specific Systems, Architectures and Processors (ASAP) **ASAP 2024**

Table of Contents

Preface	xi
Message from the Chairs	xii
Organizing Committee	xiv
Program Committee	xvi
Reviewers	xviii
Sponsors	xx

2024 IEEE 35th International Conference on Application-Specific Systems, Architectures and Processors

<p>Multi-Level Prototyping of a Vertical Vector AI Processing System 1</p> <p><i>Frederik Kautz (Cadence Design Systems, Germany; Cadence Design Systems GmbH, Germany), Sven Gesper (Chair for Chip Design for Embedded Computing, Technische Universitaet Braunschweig, Germany), Gia Bao Thieu (Chair for Chip Design for Embedded Computing, Technische Universitaet Braunschweig, Germany), Hans-Martin Bluethgen (Cadence Design Systems, Germany), Holger Blume (Institute of Microelectronic Systems, Leibniz University Hannover, Germany), and Guillermo Payá-Vayá (Chair for Chip Design for Embedded Computing, Technische Universitaet Braunschweig, Germany)</i></p> <p>SpDCache: Region-Based Reduction Cache for Outer-Product Sparse Matrix Kernels 3</p> <p><i>Valentin Isaac-Chassande (Université Grenoble Alpes, CEA, LIST, France), Adrian Evans (Université Grenoble Alpes, CEA, LIST, France), Yves Durand (Université Grenoble Alpes, CEA, LIST, France), and Frédéric Rousseau (Université Grenoble Alpes, CNRS, Grenoble INP, TIMA, France)</i></p> <p>CoNAX: Towards Comprehensive Co-Design Neural Architecture Search Using HW Abstractions ... 8</p> <p><i>Yannick Braatz (Robert Bosch GmbH, Germany), Taha Soliman (Robert Bosch GmbH, Germany), Shubham Rai (Robert Bosch GmbH, Germany), Dennis Sebastian Rieber (Robert Bosch GmbH, Germany), and Oliver Bringmann (Eberhard Karls Universität Tübingen, Germany)</i></p>	<p>1</p> <p>3</p> <p>8</p>
---	----------------------------

An End-to-End Agile Design Framework to Improve Energy Efficiency on CGRAs	17
<i>Jiangnan Li (Fudan University, China), Yazhou Yan (Fudan University, China), Guowei Zhu (Fudan University, China), Wenbo Yin (Fudan University, China), and Lingli Wang (Fudan University, China)</i>	
Fault-Tolerant DAG Scheduling with Runtime Reconfiguration on Multicore Real-Time Systems	19
<i>Yuanhai Zhang (Sun Yat-sen University, China), Shuai Zhao (Sun Yat-sen University, China), Gang Chen (Sun Yat-sen University, China), and Kai Huang (Sun Yat-sen University, China)</i>	
Design Space Exploration of Semantic Segmentation CNN SalsaNext for Constrained Architectures	28
<i>Oliver Renke (Leibniz Universität Hannover, Germany), Christoph Riggers (Leibniz Universität Hannover, Germany), Jens Karrenbauer (Leibniz Universität Hannover, Germany), and Holger Blume (Leibniz Universität Hannover, Germany)</i>	
Memory Access Acceleration Through Architecture Design for Edge SoCs	30
<i>Hongbin Wang (South China University of Technology, China; Allwinner Technology, China), Zou Yi (South China University of Technology, China), Guohua Wen (South China University of Technology, China), and Junfeng Hu (South China University of Technology, China)</i>	
ZeroVex: A Scalable and High-Performance RISC-V Vector Processor Core for Embedded Systems..	32
<i>Tenghao Zhao (Tsinghua University, China) and Zhaohui Ye (Tsinghua University, China)</i>	
MSCA: A Multi-Grained Sparse Convolution Accelerator for DNN Training	34
<i>Yingchang Mao (Tianjin University, China), Qiang Liu (Tianjin University, China), and Ray C.C. Cheung (City University of Hong Kong)</i>	
BitShare: An Efficient Precision-Scalable Accelerator with Combining-Like-Terms GEMM	36
<i>Yasong Cao (National University of Defense Technology, China), Mei Wen (National University of Defense Technology, China), Junzhong Shen (National University of Defense Technology, China), and Zhongxing Li (National University of Defense Technology, China)</i>	
Configurable Loop Shuffling via Instruction Set Extensions	45
<i>Songqiao Cui (KU Leuven, Belgium) and Josep Balasch (KU Leuven, Belgium)</i>	
Deoxys: Defensive Approximate Computing for Secure Graph Neural Networks	54
<i>Haoran Su (George Washington University, USA) and Nan Wu (George Washington University, USA)</i>	
A Framework for Generating Accelerators for Homomorphic Encryption Operations on FPGAs	61
<i>Yang Yang (University of Southern California), Rajgopal Kannan (DEVCOM Army Research Office), and Viktor Prasanna (University of Southern California)</i>	
Real-Time Order Book Building and Snapshot Generating for High Frequency Trading on FPGA	71
<i>Yao Liu (China Securities Co., Ltd, China), Shiyang Chen (China Securities Co., Ltd, China), Long Ma (China Securities Co., Ltd, China), Guolong Yang (China Securities Co., Ltd, China), and Kun Wan (China Securities Co., Ltd, China)</i>	

A FPGA-HBM-Based Hardware Streaming Accelerator for GNN Sampling	77
<i>Yuchen Gui (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), Huawei Liang (University of Science and Technology of China, China), Xiaotian Wang (University of Science and Technology of China, China), and Xi Jin (University of Science and Technology of China, China)</i>	
A DRL-Based Multi-Priority Task Division Scheduling Strategy in IIoT	79
<i>Haifeng Sun (Southwest University of Science and Technology, China) and Yunfeng Deng (Southwest University of Science and Technology, China)</i>	
TwinStep Network (TwNet): a Neuron-Centric Architecture Achieving Rapid Training	88
<i>Zhenyu Xu (Clemson University, USA), Miaoxiang Yu (Clemson University, USA), Jillian Cai (Clemson University, USA), Qing Yang (University of Rhode Island, USA), and Tao Wei (Clemson University, USA)</i>	
Analysis and Optimization of Block LU Decomposition for Execution on Tightly Coupled Processor Arrays	97
<i>Dominik Walter (Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany), Thomas Adamtschuk (Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany), Frank Hannig (Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany), and Jürgen Teich (Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany)</i>	
Accelerating MRI Uncertainty Estimation with Mask-Based Bayesian Neural Network	107
<i>Zehuan Zhang (Imperial College London, London), Matej Genci (Imperial College London, London), Hongxiang Fan (Imperial College London, London), Andreas Wetscherek (Imperial College London, London), and Wayne Luk (Imperial College London, London)</i>	
Out-of-Order and Recursive RAS: A Return Address Stack Design On High Performance Processor	116
<i>Yude Fang (National University of Defense Technology, China), Junhui Wang (National University of Defense Technology, China), Libo Huang (National University of Defense Technology, China), Yongwen Wang (National University of Defense Technology, China), and Weixia Xu (National University of Defense Technology, China)</i>	
SPARKLE: A 1,024-Core/16,384-Thread Single FPGA Many-Core RISC-V Barrel Processor Overlay	118
<i>Riadh Ben Abdelhamid (Heidelberg University, Germany), Vladislav Valek (Heidelberg University, Germany), and Dirk Koch (Heidelberg University, Germany)</i>	
Design Space Exploration of FFT Accelerators for IEEE 802.11ax Using High-Level Synthesis	120
<i>Uyong Lee (Konkuk University, South Korea), Yeji Park (LG Electronics, South Korea), Junsu Heo (Konkuk University, South Korea), Sungkyung Park (Pusan National University, South Korea), and Chester Sungchung Park (Konkuk University, South Korea)</i>	

Sparm: A Sparse Matrix Multiplication Accelerator Supporting Multiple Dataflows	122
<i>Shengbai Luo (National University of Defense Technology, China), Bo Wang (National University of Defense Technology, China), Yihao Shi (National University of Defense Technology, China), Xueyi Zhang (National University of Defense Technology, China), Qingshan Xue (National University of Defense Technology, China), and Sheng Ma (National University of Defense Technology, China)</i>	
Extending the RISC-V Instruction Set for High Performance Data Compression Hardware Acceleration	131
<i>Junzhe Huang (National University of Defense Technology, China), Qiang Dou (Phytium Technology Co., Ltd., China), and Li Shen (National University of Defense Technology, China)</i>	
MDCRA: A Reconfigurable Accelerator Framework for Multiple Dataflow Lanes	133
<i>Shaoyang Sun (Fudan University, China), Boyin Jin (Fudan University, China), Jiahang Lou (Fudan University, China), Jiangnan Li (Fudan University, China), Yuhang Cao (Fudan University, China), Jingyuan Li (Fudan University, China), Chen Shen (Fudan University, China), Yuan Dai (Fudan University, China), Wenbo Yin (Fudan University, China), Wai-Shing Luk (Fudan University, China), and Lingli Wang (Fudan University, China)</i>	
RO-SVD: A Reconfigurable Hardware Copyright Protection Framework for AIGC Applications	135
<i>Zhuoheng Ran (City University of Hong Kong, China), Abdelgawad Muhammad A.A. (City University of Hong Kong, China), Zekai Zhang (City University of Hong Kong, China), Ray C.C. Cheung (City University of Hong Kong, China), and Hong Yan (City University of Hong Kong, China)</i>	
Analyzing GPU Energy Consumption in Data Movement and Storage	143
<i>Paul Delestrac (LIRMM, University of Montpellier, Montpellier, France), Jonathan Miquel (LIRMM, University of Montpellier, Montpellier, France), Debjyoti Bhattacharjee (IMEC, Leuven, Belgium), Diksha Moolchandani (IMEC, Leuven, Belgium), Francky Catthoor (IMEC, KU Leuven, Leuven, Belgium), Lionel Torres (LIRMM, University of Montpellier, CNRS, Montpellier, France), and David Novo (LIRMM, University of Montpellier, CNRS, Montpellier, France)</i>	
k-way In-Place Merge by CPU-GPU Cooperative Processing	152
<i>Shinya Miura (Tokyo Institute of Technology, Japan), Qiong Chang (Tokyo Institute of Technology, Japan), and Jun Miyazaki (Tokyo Institute of Technology, Japan)</i>	
CSIFA: A Configurable SRAM-Based In-Memory FFT Accelerator	161
<i>Yiyang Lin (South China University of Technology, China), Yi Zou (South China University of Technology, China), and Yanfeng Yang (South China University of Technology, China)</i>	
A Convolutional Spiking Neural Network Accelerator with the Sparsity-Aware Memory and Compressed Weights	163
<i>Hanqing Liu (Peking University Shenzhen Graduate School, China), Xiaole Cui (Peking University Shenzhen Graduate School, China), Sunrui Zhang (Peking University Shenzhen Graduate School, China), Mingqi Yin (Peking University Shenzhen Graduate School, China), Yuanyuan Jiang (Peking University, China), and Xiaoxin Cui (Peking University, China)</i>	

Spatzformer: An Efficient Reconfigurable Dual-Core RISC-V V Cluster for Mixed Scalar-Vector Workloads	172
<i>Matteo Perotti (ETH Zurich, Switzerland), Michele Raeber (ETH Zurich, Switzerland), Mattia Sinigaglia (Università di Bologna, Italy), Matheus Cavalcante (ETH Zurich, Switzerland), Davide Rossi (Università di Bologna, Italy), and Luca Benini (ETH Zurich, Switzerland; Università di Bologna, Italy)</i>	
SLICE Matrix: A Memory Access Scheduling Policy for Multicore Network Processors	174
<i>Shipeng Yue (Tsinghua University, China), Honghao Liang (Sun Yat-sen University, China), Xinpeng Xing (Sun Yat-sen University, China), and Haigang Feng (Tsinghua University, China)</i>	
Enhancing a Hearing Aid Processor with ISA Extensions Supporting Flexible Fixed-Point Formats	176
<i>Jens Karrenbauer (Leibniz University Hannover, Germany), Sven Schönewald (Leibniz University Hannover, Germany), Simon Klein (Leibniz University Hannover, Germany), and Holger Blume (Leibniz University Hannover, Germany)</i>	
MLIR-to-CGRA: A Versatile MLIR-Based Compiler Framework for CGRAs	184
<i>Tianyi Yu (University of Toronto, Canada), Omar Ragheb (University of Toronto, Canada), Stephen Wicklund (University of Toronto, Canada), and Jason Anderson (University of Toronto, Canada)</i>	
Design of Light-Weight Encryption Algorithm Based on RISC-V Platform	193
<i>Zekai Zhang (City University of Hong Kong, China) and Ray Chak Chung Cheung (City University of Hong Kong, China)</i>	
Raising Compute Density of Molecular Dynamics Simulation Through Approximate Memoization	195
<i>Salim Khemira (University of Toronto, Canada), Xinyuan Wang (University of Toronto, Canada), Anh Nguyen (Fujitsu Laboratories, Japan), Yutaka Tamiya (Fujitsu Laboratories, Japan), Makoto Taiji (RIKEN Center for Biosystems Dynamics Research, Japan), Takahide Yoshikawa (Fujitsu Laboratories, Japan), and Jason Anderson (University of Toronto, Canada)</i>	
Leveraging MLIR for Efficient Irregular-Shaped CGRA Overlay Design	204
<i>Mohamed Bouaziz (King Abdullah University of Science and Technology, Saudi Arabia) and Suhaib A. Fahmy (King Abdullah University of Science and Technology, Saudi Arabia)</i>	
xTern: Energy-Efficient Ternary Neural Network Inference on RISC-V-Based Edge Systems	206
<i>Georg Rutishauser (ETH Zurich, Switzerland), Joan Mihali (ETH Zurich, Switzerland), Moritz Scherer (ETH Zurich, Switzerland), and Luca Benini (ETH Zurich, Switzerland; Università di Bologna, Italy)</i>	
LLM Based End-to-end Branch Predictor Optimization Generator	214
<i>Qingchen Zhai (University of Chinese Academy of Sciences), Zhiwei Zhang (Institute of Automation, Chinese Academy of Sciences), and Ruozhou Xiao (Institute of Automation, Chinese Academy of Sciences)</i>	
Multiplier Design Addressing Area-Delay Trade-offs by Using DSP and Logic Resources on FPGAs	217
<i>Andreas Böttcher (Fulda University of Applied Sciences, Germany) and Martin Kumm (Fulda University of Applied Sciences, Germany)</i>	

A Full-System Approach to Multi-Valued Logic Design	226
<i>Annina Gutermann (Institute of Information Processing Technology, Karlsruhe Institute of Technology, Germany) and Juergen Becker (Institute of Information Processing Technology, Karlsruhe Institute of Technology, Germany)</i>	
CHIME: Energy-Efficient STT-RAM-Based Concurrent Hierarchical In-Memory Processing	228
<i>Dhruv Gajaria (University of Arizona, USA), Tosiron Adegbiya (University of Arizona, USA), and Kevin Gomez (University of Massachusetts, USA)</i>	
ISP2DLA: Automated Deep Learning Accelerator Design for On-Sensor Image Signal Processing ..	237
<i>Dong-eon Won (Hanyang University, South Korea), Yeeun Kim (Hanyang University, South Korea), Janghwan Lee (Hanyang University, South Korea), Minjae Lee (Hanyang University, South Korea), Jonghyun Bae (SK hynix Inc., South Korea), Jongjoo Park (SK hynix Inc., South Korea), Jeongyong Song (SK hynix Inc., South Korea), and Jungwook Choi (Hanyang University, South Korea)</i>	
Voltage Range Evaluation of an Optically Reconfigurable Gate Array VLSI	239
<i>Yuki Shimamura (Okayama University, Japan), Minoru Watanabe (Okayama University, Japan), and Nobuya Watanabe (Okayama University, Japan)</i>	
Lightweight Extension of RISC-V Core for NTT-Like Algorithms	241
<i>Mateusz Wyrzywalski (AGH University of Krakow, Poland) and Robert Szczygiel (AGH University of Krakow, Poland)</i>	
A LLC-Friendly LSM-Tree	243
<i>Tao Cai (Jiangsu University, China), Jianfei Dai (Jiangsu University, China), Dejiao Niu (Jiangsu University, China), Lei Li (Jiangsu University, China), Zeyu Huang (Jiangsu University, China), and Qiangqiang Ni (Jiangsu University, China)</i>	
Research on High-Efficiency Asynchronous Superscalar Processors	245
<i>Kangli Zhao (Engineering Lanzhou University, China), Anping He (Engineering Lanzhou University, China), and Di Zhao (Institute of Computing Technolog, Chinese Academy of Science, China)</i>	
Design of High-Performance while Energy-Efficient Microprocessor with Novel Asynchronous Techniques (PhD Forum Paper)	247
<i>Xiqin Tang (University of Chinese Academy of Sciences, China) and Delong Shang (Nanjing Institute of Intelligent Technology, China)</i>	
Author Index	249