

**2025 IEEE 36th International  
Conference on  
Application-specific Systems,  
Architectures and Processors  
(ASAP 2025)**

**Vancouver, British Columbia, Canada  
28-30 July 2025**



IEEE Catalog Number: CFP25063-POD  
ISBN: 979-8-3315-9553-1

**Copyright © 2025 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:    | CFP25063-POD      |
| ISBN (Print-On-Demand): | 979-8-3315-9553-1 |
| ISBN (Online):          | 979-8-3315-9552-4 |
| 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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2025 IEEE 36th International Conference on Application- specific Systems, Architectures and Processors (ASAP) **ASAP 2025**

## Table of Contents

|                                   |       |
|-----------------------------------|-------|
| Message from General Chairs ..... | x     |
| Organizing Committee .....        | xii   |
| Program Committee .....           | xiii  |
| Keynotes .....                    | xv    |
| Reviewers .....                   | xvii  |
| Sponsors .....                    | xviii |

### Scalable Systems and Secure Acceleration

|   |    |
|---|----|
| AIRES: Accelerating Out-of-Core GCNs via Algorithm-System Co-Design .....   | 1  |
| <i>Shakya Jayakody (University of Central Florida), Youpeng Zhao (University of Central Florida), and Jun Wang (University of Central Florida)</i>  |    |
| A Dynamic Allocation Scheme for Adaptive Shared-Memory Mapping on Kilo-Core RV Clusters for Attention-Based Model Deployment .....  | 9  |
| <i>Bowen Wang (ETH Zürich, Switzerland), Marco Bertuletti (ETH Zürich, Switzerland), Yichao Zhang (ETH Zürich, Switzerland), Victor J.B. Jung (ETH Zürich, Switzerland), and Luca Benini (ETH Zürich, Switzerland; Università di Bologna, Italy)</i>  |    |
| CRYPTONITE: Scalable Accelerator Design for Cryptographic Primitives and Algorithms .....   | 17 |
| <i>Karthikeya Sharma Maheswaran (Georgia Institute of Technology, USA), Camille Bossut (Georgia Institute of Technology, USA), Andy Wanna (Georgia Institute of Technology, USA), Qirun Zhang (Georgia Institute of Technology, USA), and Cong Hao (Georgia Institute of Technology, USA)</i> |    |
| FAV-NSS: An HIL Framework for Accelerating Validation of Automotive Network Security Strategies .....   | 25 |
| <i>Changhong Li (Trinity College Dublin, Ireland), Shashwat Khandelwal (Trinity College Dublin, Ireland), and Shreejith Shanker (Trinity College Dublin, Ireland)</i>   |    |

## Design Exploration and Emerging Hardware

- PRDSE: A Prior-Driven Design Space Exploration Method ..... 33  
*Junda Zhu (Northwestern Polytechnical University, China), Xiaoya Fan (Northwestern Polytechnical University, China), Jianfeng An (Northwestern Polytechnical University, China), and Kaijie Feng (Northwestern Polytechnical University, China)*
- Trimming Down Large Spiking Vision Transformers via Heterogeneous Quantization Search ..... 41  
*Boxun Xu (UC Santa Barbara, USA), Yufei Song (UC Santa Barbara, USA), and Peng Li (UC Santa Barbara, USA)*
- X-pSRAM: A Photonic SRAM with Embedded XOR Logic for Ultra-Fast In-Memory Computing .... 49  
*Md Abdullah Al Kaiser (University of Wisconsin-Madison, USA), Sugeet Sunder (USC Information Sciences Institute, USA), Ajey P. Jacob (USC Information Sciences Institute, USA), and Akhilesh R. Jaiswal (University of Wisconsin-Madison, USA)*

## RISC-V and Custom Architectures

- Labidus: RISC-V Overlay with Streaming Asynchronous Custom Instructions ..... 57  
*Gongjin Sun (University of California, Irvine, USA), Seongyoung Kang (University of California, Irvine, USA), Jane He (University of California, Irvine, USA), Se-Min Lim (University of California, Irvine, USA), and Sang-Woo Jun (University of California, Irvine, USA)*
- Ahead of Time Generation for GPSA Protection in RISC-V Embedded Cores ..... 65  
*Louis Savary (Univ Rennes), Simon Rokicki (Univ Rennes), and Steven Derrien (UBO)*
- SCAL: An Open-Source Scalable Core Adaptation Layer for Interfacing RISC-V ISA Extensions ..... 73  
*Brindusa Mihaela Damian Kosterhon (Technical University of Darmstadt, Germany), Florian Meisel (Technical University of Darmstadt, Germany), and Andreas Koch (Technical University of Darmstadt, Germany)*
- MXDOTP: A RISC-V ISA Extension for Enabling Microscaling (MX) Floating-Point Dot Products .... 81  
*Gamze İslamoğlu (ETH Zurich, Switzerland), Luca Bertaccini (ETH Zurich, Switzerland), Arpan Suravi Prasad (ETH Zurich, Switzerland), Francesco Conti (University of Bologna, Italy), Angelo Garofalo (University of Bologna, Italy), and Luca Benini (ETH Zurich, Switzerland)*

## LLMs, Transformers, and Accelerators

- ReaLLM: A Trace-Driven Framework for Rapid Simulation of Large-Scale LLM Inference ..... 85  
*Huwan Peng (University of Washington, USA), Scott Davidson (University of Washington, USA), C.-J. Richard Shi (University of Washington, USA), and Michael Taylor (University of Washington, USA)*

|   |     |
|---|-----|
| METAL: A Memory-Efficient Transformer Architecture for Long-Context Inference on FPGA .....   | 93  |
| <i>Zicheng He (University of California, Los Angeles), Shaoqiang Lu (Ningbo Institute of Digital Twin, Eastern Institute of Technology, China), Tiandong Zhao (University of California, Los Angeles), Jinlong Yan (Ningbo Institute of Digital Twin, Eastern Institute of Technology, China), Chen Wu (Ningbo Institute of Digital Twin, Eastern Institute of Technology, China), and Lei He (University of California, Los Angeles)</i>   |     |
| An MLA-LLM Hardware Acceleration with Tensor-Train Decomposition on Group Vector Systolic Accelerator .....   | 101 |
| <i>Sixiao Huang (Southern University of Science and Technology, China), Tintin Wang (Southern University of Science and Technology, China), Keyao Jiang (Southern University of Science and Technology, China), Ao Shen (Southern University of Science and Technology, China), Kai Li (Southern University of Science and Technology, China), Ang Li (Southern University of Science and Technology, China), Mingqiang Huang (Southern University of Science and Technology, China), and Hao Yu (Southern University of Science and Technology, China)</i> |     |
| SpiRec: Soft-Logic Architecture Exploration of Reconfigurable Systems for Spiking Neural Networks .....   | 109 |
| <i>Xunqin Lai (Eindhoven University of Technology, Netherlands), Federico Corradi (Eindhoven University of Technology, Netherlands), and Siva Satyendra Sahoo (IMEC, Belgium)</i>   |     |

## ASAP since the Millennium: A Selection of Most Representative Papers

|  |     |
|--|-----|
| Looking back 25 years to the HP Labs ASIC synthesis project, PICO .....                                | 117 |
| <i>Rob Schreiber (Cerebras Systems, Inc), Shail Aditya (NVIDIA Corp), and Vinod Kathail (AMD, Inc)</i> |     |
| The Rise and Fall of Automatic Instruction-Set Extensions .....  | 119 |
| <i>Paolo Ienne (Ecole Polytechnique Federale de Lausanne)</i>  |     |
| Table-based Polynomials for Fast Hardware Function Evaluation .....                                    | 121 |
| <i>Florent de Dinechin (INSA Lyon-Inria, France)</i>   |     |

## Special Session: Reconfigurable Edge Computing

|   |     |
|---|-----|
| Managing Computation Offloading from Edge Devices to a Reconfigurable Edge Cloud .....  | 123 |
| <i>Zhehang Zhang (University of Massachusetts, USA), Bharadwaj Madabhushi (University of Massachusetts, USA), Sandip Kundu (University of Massachusetts, USA), and Russell Tessier (University of Massachusetts, USA)</i> |     |
| Bio-Inspired Event Cameras for Robust Edge System in Challenging Environments .....   | 131 |
| <i>Zhaoqi Wang (University of Florida, U.S.), Wade A. Fortney (University of Florida, U.S.), and Christophe Bobda (University of Florida, U.S.)</i>   |     |
| ELLIE: Energy-Efficient LLM Inference at the Edge via Prefill-Decode Splitting .....  | 139 |
| <i>Haoyang Fan (University of Southern California), Yi-Chien Lin (University of Southern California), and Viktor Prasanna (University of Southern California)</i>   |     |

## Special Session: Architectures for Sustainable Security

|   |     |
|---|-----|
| Differential Power Analysis on Low-Energy Keystream Generating Hardware: Full-State Recovery in DIZY Implementation .....   | 147 |
| <i>Martin Schmid (University of Passau, Germany) and Elif Bilge Kavun (Barkhausen Institut, Germany; Dresden University of Technology, Germany)</i>   |     |
| Architectures for Sustainable Security in the Computing Continuum .....   | 155 |
| <i>Francesco Regazzoni (University of Amsterdam and Università della Svizzera italiana)</i>   |     |
| Unified FPGA Design of Kyber and Dilithium with Provable Fault Tolerance .....  | 158 |
| <i>Siddhartha Chowdhury (Indian Institute of Technology Kharagpur, India), Nimish Mishra (Indian Institute of Technology Kharagpur, India), Sarani Bhattacharya (Indian Institute of Technology Kharagpur, India), and Debdeep Mukhopadhyay (Indian Institute of Technology Kharagpur, India)</i> |     |

## Posters

|  |     |
|--|-----|
| PAMA: Large-Scale GNN Acceleration with Pre-Aggregation in Multi-Node Architecture .....   | 166 |
| <i>Weiwei Wu (Tsinghua University, China), Fengbin Tu (The Hong Kong University of Science and Technology, China), Shaojun Wei (Tsinghua University, China), Xiangyu Li (Tsinghua University, China), Yang Hu (Tsinghua University, China), and Shouyi Yin (Tsinghua University, China; International Innovation Center of Tsinghua University, China)</i> |     |
| High Performance Soft Multi-Processor Arrays .....   | 168 |
| <i>Martin Langhammer (Altera Corporation), Gregg Baeckler (Altera Corporation), and Kim Bozman (Altera Corporation)</i>  |     |
| FPGA-Based Acceleration for EMT Simulation of Electrical Distribution Network .....  | 170 |
| <i>Zhenrui Wang (Texas A&amp;M University, USA), Jiang Hu (Texas A&amp;M University, USA), and Weiping Shi (Texas A&amp;M University, USA)</i>   |     |
| Noise Reduction in Hearing-Aid Processors: Traditional Methods vs. Neural Networks .....   | 172 |
| <i>Simon Klein (Leibniz University Hannover, Germany), Lando Rossol (Leibniz University Hannover, Germany), Finn Venema (Leibniz University Hannover, Germany), Sven Schoenewald (Leibniz University Hannover, Germany), Jens Karrenbauer (Leibniz University Hannover, Germany), and Holger Blume (Leibniz University Hannover, Germany)</i>              |     |
| An Application Mapping Algorithm for Mesh-Based Hybrid Optical Network-on-Chip Architecture .....  | 174 |
| <i>Sucharita Samanta (BITS Pilani, India), Nandan Surani (BITS Pilani, India), Devarsh Patel (BITS Pilani, India), Vikas Dubey (BITS Pilani, India), and Kanchan Manna (BITS Pilani, India)</i>  |     |
| Permuting Accumulation Order for Low-Precision Machine Learning .....  | 176 |
| <i>Ebby Samson (Imperial College London, UK), Tony Liu (Imperial College London, UK), Wayne Luk (Imperial College London, UK), and George A. Constantinides (Imperial College London, UK)</i>  |     |

|   |            |
|---|------------|
| Exploration Framework for IDS Optimization on FPGA .....  | 178        |
| <i>Kévin Druart (Thales DMS / Lab-STICC, France), David Espes (Lab-STICC, France), Catherine Dezan (Lab-STICC, France), and Alain Deturche (Thales DMS, France)</i>   |            |
| Scalable Malware Detection Framework using Performance Counters and Gradient Boosting .....   | 180        |
| <i>Chutitep Woralert (Clarkson University, USA), Chen Liu (Clarkson University, USA), and Zander Blasingame (Clarkson University, USA)</i>  |            |
| High-Performance, Area-Efficient and Predictable Matrix Multiplication for ASIC .....   | 182        |
| <i>Liliia Almeeva (Universtiy of Rostock, Germany), Alexander Lehnert (Universtiy of Rostock, Germany), Ralf Müller (Friedrich-Alexander-Universität Erlangen-Nurnberg, Germany), and Marc Reichenbach (Universtiy of Rostock, Germany)</i> |            |
| Integrated Security Mechanisms for Weight Protection in Memristive Crossbar Arrays .....  | 184        |
| <i>Muhammad Faheemur Rahman (University of Massachusetts Amherst) and Wayne Burleson (University of Massachusetts Amherst)</i>  |            |
| <b>Author Index .....</b>   | <b>187</b> |