Suhail Basalama

Ph.D. Student, University of California - Los Angeles
basalamasuhail@gmail.com — Personal Website — LinkedIn — Google Scholar

RESEARCH INTERESTS

- Hardware acceleration for machine learning
- Electronic design automation (EDA)
- Algorithm/hardware co-design
- High-level synthesis (HLS)
- Domain-specific and customized computing
- VLSI design methodologies
- Hardware compilation and synthesis
- Reconfigurable computing

EDUCATION

University of California, Los Angeles — September 2020 — Current
Ph.D. in Computer Science, Cumulative GPA: 3.9 /4.0 — Advisor: Prof. Jason Cong

University of California, Los Angeles — September 2020 — December 2024
M.S. in Computer Science, Cumulative GPA: 3.9 /4.0 — Advisor: Prof. Jason Cong

University of Arkansas, Fayetteville — August 2017 — December 2019
B.S. in Computer Engineering, Cumulative GPA: 4.0 /4.0

University of Arkansas, Fayetteville — August 2017 — May 2020
B.A. in Political Science, Cumulative GPA: 4.0 /4.0

ACADEMIC EXPERIENCE

University of California, Los Angeles, *VAST Lab* — June 2020 — Present
Graduate Student Researcher — Advisor: Prof. Jason Cong

Stream-HLS: An automatic end-to-end MLIR-based compiler that converts sequential multi-kernel applications into optimized parallel dataflow architectures with streaming capabilities

- Developed multiple MLIR passes for the global optimization of multi-kernel applications
- Constructed a comprehensive design space beyond those of prior works
- Devised a global scheduling approach using mixed-integer nonlinear programming (MINLP)
- Evaluated Stream-HLS against SoTA frameworks (5.42× - 2270.25× geo. mean speedups)
- Accepted as a conference paper at ***FPGA 2025*** (**Best Paper Candidate**)

Odyssey: A comprehensive design space exploration framework for systolic arrays using a combination of mathematical programming, an evolutionary algorithm, and a padding-based heuristic

- Analyzed the impact of choosing divisor vs. non-divisor tiling factors in systolic array designs
- Constructed a comprehensive design space (larger than $O(2^{40})$) for systolic arrays
- Designed a near-optimal heuristic algorithm for pruning the design space and reducing search time
- Published as a conference paper at ***DAC 2023***

FlexCNN: An HLS framework for composing CNN accelerators on FPGA with versatile systolic arrays

- Integrated a versatile systolic arrays to support transposed and dilated convolution layers
- Improved the front-end of the framework to take ONNX CNN models as inputs
- Implemented VGG16, U-Net, and E-Net with comparable or better performance than prior works
- Published as a journal paper at ***TRETS 2023***

Versatile Systolic Arrays (VSA): An HLS versatile systolic array for efficient processing of transposed and dilated convolution on FPGAs.

- Implemented a decomposition of transposed and dilated convolutions as a versatile systolic array
- Automated the process of generating new VSA designs based on the target CNN
- Published as a poster at ***FCCM 2022***

University of Arkansas, Computer Systems Design Laboratory — January 2019 — May 2020
Undergraduate Student Researcher — Advisor: Prof. David Andrews

SPAR-2 Project: Four different SIMD serial systolic array processor architectures on FPGAs for accelerating machine learning workloads

- Built a MicroBlaze System-on-Chip (SoC) with separate BRAM for instructions
- Wrote an Instruction Sequencer Module in Verilog and packaged it as an AXI4 IP core
- Manipulated Booth's and Modified Booth's algorithms to reduce the shift operations
- Published as conference papers at ***ICDIS 2020*** and ***FPL 2021***

University of Arkansas, Smart Embedded Systems Lab — June 2018 – January 2019
Undergraduate Student Researcher — Advisor: Prof. Christophe Bobda

ARLO Robot: An autonomous ground vehicle built using the Parallax Arlo Robot System and the Digilent Zybo Z7-20 FPGA

- Built a Zynq-7000 SoC with UART IP and a custom Ultrasonic Sensor IP I implemented
- Deployed a Petalinux system on the SoC and configured the kernel with the needed modules

Cryptocurrency Wallet: A hardware-based wallet with end-to-end AES encryption for cryptocurrency on the Lattice iCE40 Ultra Wearable Development Platform

- Helped develop a UART-BlueTooth communication interface between the FPGAs and a smartphone
- Helped implement the AES encryption algorithm in Verilog on FPGA

Professional Experience

Micron, Catapult-MLIR — June 2023 — September 2023
HLS Intern Researcher — Managers: Dr. Nihar Athreyas and Dr. Lorenzo Ferretti

Developed an MLIR infrastructure for automatic code transformation for ASIC designs using Catapult-Siemens high-level synthesis

Microsoft, Brainwave Project — June 2021 — September 2021
Silicon/Firmware Intern Researcher — Managers: Dr. Steven K. Reinhardt and Dr. Gabriel Weisz

Created an interactive performance visualization/profiling tool that analyses RTL waveforms from the command-line and produces a visualization of the task dependencies and their run-times traces

PUBLICATIONS

8. [**FPGA'25**] **Suhail Basalama**, and Jason Cong. "Stream-HLS: Towards Automatic Dataflow Acceleration." In *Proceedings of the 2025 ACM/SIGDA International Symposium on Field Programmable Gate Arrays (FPGA)*, 2025. **(FPGA'25 Best Paper Candidate)**. Open source: https://github.com/UCLA-VAST/Stream-HLS

7. [**ASPLOS'24**] Neha Prakriya, Yuze Chi, **Suhail Basalama**, Linghao Song, and Jason Cong. "TAPA-CS: Enabling Scalable Accelerator Design on Distributed HBM-FPGAs." In *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), Volume 3*, pp. 966-980. 2024. Open source: https://github.com/UCLA-VAST/TAPA-CS

6. [**FPGA'24**] Linghao Song, Licheng Guo, **Suhail Basalama**, Yuze Chi, Robert F. Lucas, and Jason Cong. "Callipepla: Stream Centric Instruction Set and Mixed Precision for Accelerating Conjugate Gradient Solver." In *Proceedings of the 2023 ACM/SIGDA International Symposium on Field Programmable Gate Arrays (FPGA)*, pp. 247-258. 2023. Open source: https://github.com/UCLA-VAST/Callipepla

5. [**DAC'23**] **Suhail Basalama**, Jie Wang, and Jason Cong. "A Comprehensive Automated Exploration Framework for Systolic Array Designs." In *2023 60th ACM/IEEE Design Automation Conference (DAC)*, pp. 1-6. IEEE, 2023. Open source: https://github.com/UCLA-VAST/Odyssey

4. [**TRETS'23**] **Suhail Basalama**, Atefeh Sohrabizadeh, Jie Wang, Licheng Guo, and Jason Cong. "FlexCNN: An end-to-end Framework for Composing CNN Accelerators on FPGA." *ACM Transactions on Reconfigurable Technology and Systems (TRETS) 16*, no. 2 (2023): 1-32. https://github.com/UCLA-VAST/FlexCNN

3. [**FCCM'22**] **Suhail Basalama**, Atefeh Sohrabizadeh, Jie Wang, and Jason Cong. "A Versatile Systolic Array for Transposed and Dilated Convolution on FPGA." In *2022 IEEE 30th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pp. 1-2. IEEE, 2022.

2. [**FPL'21**] Atiyehsadat Panahi, **Suhail Basalama**, Ange-Thierry Ishimwe, Joel Mandebi Mbongue, and David Andrews. "A Customizable Domain-specific Memory-centric FPGA Overlay for Machine Learning Applications." In *2021 31st International Conference on Field-Programmable Logic and Applications (FPL)*, pp. 24-27. IEEE, 2021. Open Source: https://github.com/SuhailB/ArrayProcessor

1. [**ICDIS'20**] **Suhail Basalama**, Atiyehsadat Panahi, Ange-Thierry Ishimwe, and David Andrews. "SPAR-2: A SIMD Processor Array for Machine Learning in IoT Devices." In *2020 3rd International Conference on Data Intelligence and Security (ICDIS)*, pp. 141-147. IEEE, 2020. Open Source: https://github.com/SuhailB/ArrayProcessor

TEACHING Experience

University of California - Los Angeles, Digital Design CS M152A March 2023 — June 2023
Teaching Assistant Lecturer: Prof. Majid Sarrafzadeh

University of California - Los Angeles, Digital Design CS M152A March 2022 — June 2022
Teaching Assistant Lecturer: Prof. Majid Sarrafzadeh

University of Arkansas, Digital Design CSCE 2114 August 2019 — December 2019
Teaching Assistant Lecturer: Prof. Patrick Parkerson

- Taught around 30 students in two lab sections the fundamentals of digital and hardware design including:
  1. Number Representation, Combinational/Sequential Circuits, Logic Functions, Flip-Flops,
  2. Registers, Counters, Lookup Tables, Hardware Description Languages (Verilog and VHDL), Microcontrollers
- Held office hours to assist students with the class or lab material
- Graded and corrected students' assignments, quizzes, lab reports, and exams

SELECTED COURSES

Master's Courses

- Advanced Computer Architecture
- VLSI Design Automation
- Big Data Application
- Linear Programming
- Types and Programming Languages
- Deep Learning
- Reinforcement Learning

Bachelor's Courses

- Computer Architecture
- System Synthesis and Modeling
- Algorithms and Data Structures
- Embedded System
- Compiler Design
- Artificial Intelligence
- Wearable and Ubiquitous Computing

Additional Courses

- Ethics and the Professions

AWARDS

2025 Best Paper Candidate at International Symposium on Field-Programmable Gate Arrays 2025
2019 The Foundation for the International Exchange of Students Scholarship at UARK
2019 Dr. Henry M. Alexander Memorial Award
2019 Rosecrans, Sr Endowed Memorial Scholarship
2018 The Charles D. Brock Scholarship by the College of Engineering at UARK
2018 The Foundation for the International Exchange of Students Scholarship at UARK
2017 The John and Marie Lavallard International Scholarship at the University of Arkansas
2017 The University of Arkansas Transfer Student Scholarship
2014 The Silver Medal Representing Yemen in the Third Gulf Mathematics Olympiads in Oman
2013 The Top-Ten Student Ministerial Scholarship from the Yemeni government
2013 Ranked 9th among 218,964 Yemeni students in the National High-school Exams ***(Top 0.000041%)***

TALKS

- July 23, 2024 Member Talk at SpatialML: "End-to-End CNN Acceleration and Systolic Array Design Space Exploration for FPGAs."
  YouTube: https://youtu.be/CXH6Uu_gTDA?si=xXTuQASid7F6ich5.

TECHNICAL SKILLS

- **Design Skills**: Performance Modeling and Profiling, Design Space Exploration, System-on-Chip, IP Packaging, Static Timing Analysis, Algorithms, Finite State Machines
- **Languages**: C, C++, Python, High-level synthesis (HLS), Verilog, VHDL, Java, Tcl, Bash
- **CAD Tools**: Vitis HLS, Vivado Xilinx/AMD, Catapult HLS, Modelsim, Synopsys Design Compiler
- **Technologies**: FPGAs (Xilinx/AMD, Altera Intel, Lattice), Raspberry Pi, Arduino
- **Operating** Systems: Linux, Windows, macOS
- **Other** Skills: Git, SVN, LaTeX, Microsoft Office Suite

REFERENCES

Prof. Jason Cong, Distinguished Professor
Director, Center for Customizable Domain-Specific Computing
Director, VLSI Architecture, Synthesis, and Technology (VAST) Laboratory
CS and ECE Departments
University of California - Los Angeles, Los Angeles, California, USA
Phone: (310) 206-2775
E-mail: cong@cs.ucla.edu
Scholar Profiles: Personal Page — Google Scholar

Prof. David Andrews, Professor
Director, Computer Systems Design Laboratory
EECS Department
University of Arkansas, Fayetteville, Arkansas, USA
Phone: (479) 575-6529
E-mail: dandrews@uark.edu
Scholar Profiles: Personal Page — Google Scholar

Prof. Christophe Bobda, Professor
Director, SmartSystems Lab
Citi Endowed Professor in Advanced Technologies
ECE Department
University of Florida, Gainseville, Florida, USA
Phone: (352) 294-2024
E-mail: cbobda@ece.ufl.edu
Scholar Profiles: Personal Page — Google Scholar