YUAN LI

Address: George Washington University SEH 5750, 800 22nd Street NW, Washington, DC 20052 Email: liyuan5859@gmail.com **Telephone:** +1 (202) 710–3327

Website: https://gwuyuan.github.io/mysite

RESEARCH INTERESTS

Emerging Hardware Architecture and Technology for High-Performance and Energy-Efficient Computing Systems including:

- AI/ML Hardware Accelerator
- Network-on-Chip (NoC)
- Heterogeneous Manycore Architecture
- Chiplet-based Heterogeneous Integration
- Silicon Photonics
- Non-Volatile Memory (NVM)

EDUCATION

George Washington University

Ph.D. in Computer Engineering

Department of Electrical and Computer Engineering

Washington, D.C. August 2017 - December 2022

University of Newcastle upon Tyne

M.S. in Microelectronics

School of Electrical, Electronic and Computer Engineering

Newcastle upon Tyne, U.K. September 2010 – December 2011

University of Science and Technology of China (USTC)

B.S. in Physics

Department of Physics

Hefei, China September 2006 – June 2010

PUBLICATIONS

My research deliverables lead to 7 first-authored publications and 2 co-authored publications in premier computer architecture, circuit, and EDA conferences and journals (HPCA, PACT, DAC, DATE, TPDS, TCAS, and TSUSC).

Conference Papers

- [C1] Yuan Li, Ahmed Louri, and Avinash Karanth. "A Silicon Photonic Multi-DNN Accelerator." to appear in IEEE/ACM International Conference on Parallel Architectures and Compilation Techniques (PACT), 2023.
- [C2] Yuan Li, Ahmed Louri, and Avinash Karanth. "Efficient Multicast Communication in Silicon Photonics Enhanced DNN Acceleration." in Proceedings of the IEEE Photonics Summer Topicals Meeting Series (SUM), 2023. [Link]
- [C3] Yuan Li, Ahmed Louri, and Avinash Karanth. "SPACX: Silicon Photonics-based Chiplet Accelerator for DNN Inference." in Proceedings of the IEEE International Symposium on High-Performance Computer Architecture (HPCA), 2022. [Link]
- [C4] Ke Wang, Hao Zheng, Yuan Li, Jiajun Li, and Ahmed Louri. "AGAPE: Anomaly Detection with Generative Adversarial Network for Improved Performance, Energy, and Security in Manycore Systems." in Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE), 2022. [Link]
- [C5] Yuan Li, Ahmed Louri, and Avinash Karanth. "Scaling Deep Learning Inference with Chiplet-based Architecture and Photonic Interconnects." in Proceedings of the ACM/IEEE Design Automation Conference (DAC), 2021. [Link]

Journal Papers

- [J1] Yuan Li, Ke Wang, Hao Zheng, Ahmed Louri, and Avinash Karanth. "ASCEND: A Scalable and Energy-Efficient Deep Neural Network Accelerator with Photonic Interconnects." in IEEE Transactions on Circuits and Systems I (TCAS-I), 2022. [Link]
- [J2] Yuan Li, Ahmed Louri, and Avinash Karanth. "SPRINT: A High-Performance, Energy-Efficient, and Scalable Chiplet-based Accelerator with Photonic Interconnects for CNN Inference." in IEEE Transactions on Parallel and Distributed Systems (TPDS), 2021. [Link]
- [J3] Ke Wang, Hao Zheng, Yuan Li, and Ahmed Louri. "SecureNoC: A Learning-Enabled, High-Performance, and Secure On-Chip Communication Framework Design." in IEEE Transactions on Sustainable Computing (TSUSC), 2021. [Link]
- [J4] **Yuan Li** and Ahmed Louri. "ALPHA: A Learning-Enabled High-Performance Network-on-Chip Router Design for Heterogeneous Manycore Architectures." in IEEE Transactions on Sustainable Computing (TSUSC), 2020. [Link]

Patents

[P1] Yuan Li and Ahmed Louri. "SPACX: A Hardware and Algorithm Co-Optimized Photonic Deep Neural Network Computing Architecture." U.S. Provisional Patent, No. 63/456,255, 2023

HONORS AND AWARDS

2021
2019
2010
2006

RESEARCH EXPERIENCE

Silicon Photonics and Chiplet-based Deep Neural Network (DNN) Hardware Accelerators

2020 - 2022

- Explored the use of silicon photonics technology to overcome the communication challenges in DNN hardware accelerators
- Designed and simulated photonic interconnection networks that adapt to various general and tailored dataflows for high-performance and energy-efficient communication in DNN hardware accelerators, with a collection of simulators and tools including SCALE-Sim, Timeloop, MAESTRO, DRAMSim, CACTI, DSENT, and Synopsys Design Compiler
- Delivered two conference papers (HPCA' 22 and DAC' 21) and two journal papers (TCAS-I and TPDS)

DRAM and NVM Integration and Management in Chiplet-based Systems

2019 - 2020

- Developed a data exchange mechanism between DRAM and NVM stacks and corresponding hardware modifications targeting
 actively moving memory pages with high access count yet short access time frame to DRAM stacks
- · Simulated the data exchange mechanism and hardware architecture with Gem5-GPU, DRAMSim, CACTI, and DSENT
- Won the NSF I-Corps Site Grant Award.

Efficient NoC for Accelerator-Rich Heterogeneous Manycore Systems

2018 - 2019

- Developed an NoC router microarchitecture and an artificial neural network (ANN) based mechanism to alleviate local and global contention for high-throughput and low-latency communication
- Simulated the router microarchitecture and ANN mechanism with Gem5, Gem5-GPU, DSENT, and Synopsys Design Compiler
- Delivered one journal paper (*TSUSC*)

Intelligent Hardware Trojan (HT) Detection in Secure Network-on-Chip Architectures

2021 - 2022

- Participated in developing runtime accurate HT detection modules in NoC architectures using multilayer perceptron (MLP)
 and generative adversarial network (GAN) models
- Participated in power and area evaluation of HT detection modules with DSENT and Synopsys Design Compiler
- Delivered one conference paper (*DATE'* 22) and one journal paper (*TSUSC*)

TEACHING EXPERIENCE

Graduate Teaching Assistant

2020, 2021

- Data Structures and Algorithms (GWU ECE 1125.30)
- Computer Architecture and Design (GWU ECE 4535.81 / 6005.81)

PRESENTATIONS

[1] "SPACX: Silicon Photonics-based Chiplet Accelerator for DNN Inference", 28th IEEE International Symposium on High-Performance Computer Architecture (HPCA), Virtual Conference, April 2022.

[2] "Scaling Deep Learning Inference with Chiplet-based Architecture and Photonic Interconnects", 58th ACM/IEEE Design Automation Conference (DAC), San Francisco, CA, December 2021.

SERVICE AND TECHNICAL REVIEWING

Conference Reviewer/Sub-Reviewer • ACM/IEEE International Symposium on Computer Architecture (ISCA)	2019, 2022
• IEEE International Symposium on High-Performance Computer Architecture (HPCA)	2020, 2021
• IEEE/ACM International Symposium on Networks-on-Chip (NOCS)	2018, 2019
IEEE International Conference on Computer Design (ICCD)	2018
r In	

Journal Reviewer

• IEEE Transactions on Parallel and Distributed Systems (TPDS) 2023

• IEEE Transactions on Emerging Topics in Computing (TETC)

2018

RELEVANT TECHNICAL COURSEWORK

Computer Architecture Modeling

- Processor microarchitecture and memory hierarchy modeling with SimpleScalar, Gem5, Gem5-GPU, CACTI, DRAMSim, etc.
- NoC modeling with BookSim, Garnet, DSENT, OptiSPICE, etc.
- DNN hardware accelerator modeling with Timeloop, MAESTRO, SCALE-Sim, etc.

Embedded System Design

- Data acquisition & processing and motion control system development with NI LabVIEW and NI PXI & CompactRIO
- Data acquisition & processing and waveform generation system development with Microchip PIC18 controllers and MPLAB X IDE

Printed Circuit Board (PCB) Design

· Signal processing, thermal control, and fiber-optic communication PCBs design with NI Multisim and Altium Designer

Design for Testability (DFT)

MIPS processor scan flops synthesis and ATPG with Synopsys Design Vision and Synopsys TetraMAX

HDL Programming and Synthesis

- Digital filter implementation with VHDL and FLEX EPF10K70 FPGA
- MIPS processor design and synthesis with Verilog, Cadence SimVision, Synopsys Design Vision, and Cyclone IV EP4CE115 FPGA
- Virtual channel NoC router design and synthesis with Verilog, Cadence SimVision, and Synopsys Design Vision

Integrated Circuit Layout Design

- D flip-flop layout design with Cadence Virtuoso
- MIPS processor layout design and routing with Cadence Virtuoso and Cadence Encounter

Device Simulation and Fabrication

- MOS transistor scaling and simulation with Synopsys TSUPREM-4
- Device fabrication training with 0.16 μ m process technology at INEX Microtechnology Ltd.

SKILLS

Programming Languages

• C/C++, Python, Visual Basic, MATLAB, Assembly, Verilog/VHDL

Design & Modeling Tools

- Gem5, Gem5-GPU, SimpleScalar, Timeloop, MAESTRO, SCALE-Sim, DRAMSim, BookSim, Garnet, CACTI, DSENT, OptiSPICE
- LabVIEW, MPLAB X IDE, NI Multisim, Altium Designer
- Synopsys TetraMAX, Altera Quartus II, Cadence SimVision, Cadence Virtuoso/Encounter, Synopsys Design Vision, Synopsys TSUPREM-4