EDUCATION

m University of Virginia, Charlottesville, USA

Aug. 2023 - Present

Ph.D. in Computer Science

GPA: 3.91/4.0

Topics: GPU × {Cryptography, Trusted-Computing, Memory, CXL}

Advisor: Prof. Adwait Jog

ill Jilin University, Changchun, China

Sept. 2019 – Jul. 2023

B.S. in Computer Science

GPA: 3.69/4.0

Thesis: The Design and Implementation of Binary Code Analysis Framework for NVIDIA GPU.

Advisor: Prof. Jingweijia TAN

PUBLICATIONS

[C2] (ISCA'25, To Appear)

NetCrafter: Tailoring Network Traffic for Non-Uniform Bandwidth Multi-GPU Systems

Amel Fatima, Yang Yang, Yifan Sun, Rachata Ausavarungnirun, Adwait Jog

In the Proceedings of International Symposium on Computer Architecture, Tokyo, Japen, June 2025

[C1] (ISPASS'25, To Appear)

Dissecting Performance Overheads of Confidential Computing on GPU-based Systems

Yang Yang, Mohammad Sonji, Adwait Jog

In the Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), Ghent, Belgium, May 2025

[P1] (Poster, SRC@CGO'23, Pre-PhD)

Facilitating Profile Guided Compiler Optimization with Machine Learning

Yang Yang, Xueying Wang, Guangli Li

Student Research Competition of IEEE/ACM International Symposium on Code Generation and Optimization (CGO), Montreal, Canada, February 2023

RESEARCH EXPERIENCE

Insight Computer Architecture Lab

Aug. 2023 – Present

University of Virginia, Charlottesville, Virginia, USA

Advisor: Prof. Adwait Jog

Focus:

- GPU memory and storage system (e.g., UVM, CXL and GPU controlled communication).
- ${
 m TDX}$ -based confidential computing on GPUs.
- Advanced cryptography (e.g. FHE, PIR and MPC) with GPUs.
- Counter-mode encryption for GPU systems.

Emerging Technology Enabled Computer Architecture Lab

Feb. 2022 – Jul. 2023

Jilin University, Changchun, Jilin, P.R.China

Advisor: Prof. Jingweijia Tan

 $\textbf{Topics} \colon \mathsf{GPU} \times \{\mathsf{PTX/SASS}, \, \mathsf{Reliability}, \, \mathsf{Energy} \, \, \mathsf{Efficiency}\}$

Focus:

- Process variation of FinFET and chiplet based MCM-GPUs.
- SASS level analysing and modeling framework for NVIDIA Ampere GPUs.
- Learning techniques for GPU power modeling.
- · Instruction level under-voltage reliability of GPUs

State Key Laboratory of Processor

Jul. 2022 - Sept. 2023

Institute of Computing Technology, Chinese Academy of Science, Beijing, P.R.China

Advisor: Prof. Guangli Li

Topics: Compiler & Programming Systems

Focus: Facilitating Profile-Guided Compiler Optimization with Graph Neural Networks

- Proposed a branch predictor using XGBoost based on static features at compile time.
- Utilize GNNs to build predictive profile-guided optimization framework and integrated it into LLVM.
- Released a new dataset for graph-related static analysis tasks.

TEACHING EXPERIENCE

24 Fall @ UVA, TA for CS: 6354 Computer Architecture

SKILLS

Software

Languages C/C++ · Assembly · Python · Go Frameworks CUDA · Pytorch · LLVM · TDX

å LINUX · LATEX · Markdown · GNU compiler (gcc, etc.) · GPGPU-Sim · Varius-TC · Z3 Solver

RELEVANT COURSEWORK

CS6354 Computer Architecture Digital computer organization.

CS6501 Program Analysis and ML Formally verify properties of NNs using Z3 solver.

CS6501 Software-Defined Networking P4 and programmable switch, implemented an NN-based packet classification inside OvS.

CS6501 Modern Computer Architectures GPU, quantum and superconducting architectures.

 $\textbf{CS6501 Cloud System Reliability} \ \texttt{Explored remote in-memory object integrity when bit-flip happens}.$