YANG YANG

🏠 *Website* Elio-yang.github.io
✉ *Official* yangyang@virginia.edu
🐙 *GitHub* Github.com/Elio-yang
🏠 *Address* **Rice Hall, 85 Engineer's Way, Charlottesville, VA, 22904**

EDUCATION

🏛 **Jilin University, Changchun, China**   *Sept. 2019 – Jul. 2023*
🎓 **B.S. in Computer Science and Technology**
GPA: 3.69/4.0
Rank: 9%
Thesis: *The Design and Implementation of Binary Code Analysis Framework for NVIDIA GPU.* [Score: 95/100]
Advisor: Prof. Jingweijia Tan

🏛 **University of Virginia, Charlottesville, USA**   *Aug. 2023 – Present*
🎓 **Ph.D. in Computer Science**
Interests: GPU · Cryptography · Security · Memory
Advisor: Prof. Adwait Jog

PUBLICATION

Facilitating Profile Guided Compiler Optimization with Machine Learning.
Yang Yang, Xueying Wang, Guangli Li[*]
SRC@CGO'23 [[Poster]](Poster)

- Achieving an average of $1.03\times$ and $1.95\times$ speedups on representative real-world applications and *Polybench* benchmark suite over the baseline (i.e., the programs without PGO), respectively.

- The performance of our machine learning-aided PGO is very close to the classic PGO ($1.05\times$ and $1.97\times$ speedups over the baseline) while reducing 58.3% and 94.8% optimization costs.

RESEARCH EXPERIENCE

Insight Computer Architecture Lab   *Aug. 2023 – Present*
University of Virginia, Charlottesville, Virginia, USA
Advisor: Prof. Adwait Jog
Research on: GPU Memory & GPU Security
What We Do:

- Exploring the memory and storage system (e.g., direct GPU communication to NVMe SSDs).

- Characterizing and simulating confidential computing on GPUs and how to make it more efficient.

- Exploring how to utilize advanced cryptography (e.g. LWE, FHE, PIR and MPC) with GPUs and how to make them faster.

- Exploiting the feasibility for using CXL in GPU and how to make it more secure.

Emerging Technology Enabled Computer Architecture Lab   *Feb. 2022 – Jul. 2023*
Jilin University, Changchun, Jilin, P.R.China
Research Assistant, Advisor: Prof. Jingweijia Tan
Research on: GPU Architecture & Reliability & Energy Efficiency & Accelerator
What We Do:

- Explored the process variation of MCM-GPUs based on FinFET and state-of-the-art chiplet technology.

- Exploited the potential of FPGA for building open-sourced GPU like Vortex.

- Implemented a L̲ow-L̲evel A̲nalysis and M̲odeling framework for NVIDIA Ampere GPU.

- Applied deep learning techniques for accurate power modeling.

- Examined the power-level effect of the instruction control flag when generating the SASS.

Institute of Computing Technology, Chinese Academy of Science, Beijing, P.R.China
Research Assistant, Advisor: Prof. Guangli Li
Research on: Compiler & Programming Systems & Deep Learning
Project: Facilitating Profile-Guided Compiler Optimization with Graph Neural Networks

- **Proposed a branch predictor using XGBoost based on static features.**

- **Explore the speedup sensibility of different programs towards different feature design.**

- **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

Languages	**C/C++ · Assembly · Python · Go**
Frameworks	**CUDA · Pytorch · LLVM**
Software	**LINUX · LaTeX · Markdown · GNU compiler (gcc, *etc.*) · GPGPU-Sim · Varius-TC · Z3 Solver**