

# Haoyang Zhang

SECOND-YEAR PH.D. STUDENT, COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

✉ zhang402@illinois.edu | 🏠 hieronzhang.github.io | 📧 HieronZhang | 🐦 @HaoyangZhang19

## Education

---

### University of Illinois Urbana-Champaign

Urbana, IL, U.S.

PH.D. COMPUTER SCIENCE

Aug. 2022 - Present

- Advisor: prof. Jian Huang
- Research Interest: computer architecture and system software, with a focus on exploiting HW-SW co-design and building novel memory/storage systems to support AI infrastructure/platforms

### University of Michigan

Ann Arbor, MI, U.S.

B.S.E. COMPUTER SCIENCE (DUAL DEGREE)

Aug. 2020 - Apr. 2022

- Advisor: prof. Baris Kasikci

### Shanghai Jiao Tong University

Shanghai, China

B.S.E. ELECTRICAL AND COMPUTER ENGINEERING (DUAL DEGREE)

Sep. 2018 - Aug. 2022

- Advisor: prof. Weikang Qian

## Publications

---

*G10: Enabling An Efficient Unified GPU Memory and Storage Architecture with Smart Tensor Migrations.*  
**Haoyang Zhang\***, Yirui Eric Zhou\*, Yuqi Xue, Yiqi Liu, Jian Huang. (MICRO 2023).

*Debugging in the Brave New World of Reconfigurable Hardware.*

Jiacheng Ma, Gefei Zuo, Kevin Loughlin, **Haoyang Zhang**, Andrew Quinn, Baris Kasikci. (ASPLOS 2022).

*Scheduling Information-Guided Efficient High-Level Synthesis Design Space Exploration.*

Xingyue Qian, Jian Shi, Li Shi, **Haoyang Zhang**, Lijian Bian, Weikang Qian. (ICCD 2022).

## Research Experience

---

### PlatformX Group - University of Illinois Urbana-Champaign

Urbana, IL, U.S.

ADVISOR: JIAN HUANG

Aug.2022 - Present

- **Project: G10: Enabling An Efficient Unified GPU Memory and Storage Architecture with Smart Tensor Migrations (Published in MICRO'23)**

We present a unified GPU memory and storage architecture driven by the fact that DNN workloads are highly predictable. G10 integrates the host memory, GPU memory, and flash memory into a unified memory space, to scale the GPU memory capacity while enabling transparent data migrations. G10 utilizes compiler techniques to characterize the tensor behaviors in DNN workloads to schedule data migrations in advance by considering the available bandwidth of flash memory, host memory, and interconnections.

- **Project: A Survey on CXL-based Memory/Storage Systems**

Read and analyze recent research papers on CXL-based memory/storage systems. Understand potential opportunities and trade-offs brought by CXL techniques.

### Efeslab - University of Michigan

Ann Arbor, MI, U.S.

ADVISOR: BARIS KASIKCI

Apr.2021 - Apr.2022

- **Project: Reconfigurable Hardware Debugging (Published in ASPLOS'22)**

Studied bugs in existing FPGA designs and produced a testbed to reliably reproduce each bug. Helped build a novel collection of hybrid static/dynamic program analysis and monitoring tools for debugging FPGA designs. Implemented the simulator for SignalCat, a recorder and printer on actual hardware deployments.

### CELAB - University of Michigan

Ann Arbor, MI, U.S.

ADVISOR: TODD AUSTIN

Sept. 2021 - Dec. 2021

- **Project: A Spectre Attack Targeting Compiler-based Protections. (Also EECS 573 Course Project)**

Found a way to subvert compiler-based LFENCE Spectre protections by identifying and jumping over the speculation barrier code inserted by the compiler, basically by exploiting the partial-tagged and partial-data BTB to conduct a special kind of intra-process, out-of-place Spectre attack.

## Emerging Computing Technology Laboratory – Shanghai Jiao Tong University

Shanghai, China

ADVISOR: WEIKANG QIAN

Jan. 2021 – Aug. 2022

- **Project: MiniHLS: A Simple High-Level Synthesis Tool (Published in ICCD'22)**

Helped to build an HLS tool based on LLVM, which can compile C functions to Verilog modules, having multiple features including loop optimization, tcl file parsing, etc. Test the HLS output on FPGA, do post-synthesis debugging/on-board verification, and hence modify the source code.

## Industry Experience

---

### T-head Division, Alibaba Cloud

Hangzhou, China

MENTOR: YUNHAI SHANG

Jul. 2021 - Aug. 2021

- **Position: Software Research & Development Intern**

- **Project:** Implement the optimization for the jpeg library for RISC-V vector processors.

## Teaching

---

### SHANGHAI JIAO TONG UNIVERSITY

SU. 2022 **Computer Architecture (VE470)**, Teaching Assistant

SU. 2020 **Honors Physics (VP160)**, Teaching Assistant

## Selected Awards and Honors

---

2021 **Roger King Scholarship**, University of Michigan

2021 **University Honors**, University of Michigan

2021, 2020 **Dean's List**, University of Michigan

2020, 2019 **Undergraduate Academic Excellence Scholarship**, Shanghai Jiao Tong University

2020 **Student Development Scholarship**, Shanghai Jiao Tong University

2019 **Silver Medal**, University Physics Competition

## Selected Projects

---

### 3-Way Superscalar R10K-Style Out-of-Order Processor Core

Ann Arbor, U.S.

CAPSTONE. ADVISOR: JON BEAUMONT

Feb. 2021 - April. 2021

- Build a 3-way Superscalar R10K-Style Processor with a few advanced features: Non-blocking L1 Data cache, Dynamic branch prediction, Instruction prefetching, etc.
- Implement our processor using SystemVerilog. Analyze and optimize the performance of our processor.

### An Optimized Compiler for Decaf Language

Ann Arbor, U.S.

ADVISOR: LINGJIA TANG

Jan. 2021 - April. 2021

- Implement a compiler with source language Decaf and target language MIPS using flex (lexical analysis), yacc/bison (syntax analysis), and C/C++ (other parts).
- Implemented a few IR optimizations including Dead Code Elimination, Common Subexpression Elimination, etc.

## Skills

---

Language **C/C++, HLS, Verilog/SystemVerilog, CUDA, Yacc, Murphi, Python, MATLAB, TeX**

ISAs **RISC-V, X86, LEGv8, MIPS, LC2K**

EDA Tools **Synopsys VCS, Verilator, Xilinx Vivado**

## Community Services

---

### SERVICE

2019 **SJTU Student Science and Technology Innovation Association**, Consultant

2019 **Ji Young Volunteers Association**, Minister

### VOLUNTEERING

2019 **Yunnan San He Junior High School Volunteer Teaching Team**, Team leader