

YIWEI YANG

✉ vickiegypt@ucsc.edu · 🌐 asplos.dev 🐙 vickiegypt

EDUCATION

UC Santa Cruz, Ph.D. Student 08/2022 – 06/2028

- Major: Computer Science, advised by Andrew Quinn. 22 Fall TA of Computer Architecture

ShanghaiTech University, Undergraduate 09/2018 – 06/2022

- Major: Computer Science, finished Compiler, Network, Database, OS, CA, Convex, RL, Parallel Computing. 21& 22 Spring TA of Compiler

WORK EXPERIENCE

Jump Trading, Shanghai, China 07/2020 – 09/2020




(Linux Team) Production Engineer Intern

- High Frequency Trade Order Book simulation applying Linear Regression Method.
- Designed a user interface to automate the core affinity of jobs.
- Designed an eBPF exporter of GPFS full OSS lifetime traces for better reporting the bottleneck.

RESEARCH EXPERIENCE

Storage Systems Research Center, UC Santa Cruz 08/2022 – Present

(Graduate Research) Assistant

- Data-driven CXL.mem allocation, prefetching, and replacement policies Operating System. 
- Make Hardware Software Co-design for on CXL.cache data movement 
- Make universal migration based on WebAssembly and PTX ^{MV}_{VM}
- Make cross kernel-userspace eBPF observation for distributed system 

SELECTED PUBLICATIONS

”eGPU: Extending eBPF Programmability and Observability to GPUs” HCDS25 **Yiwei Yang**, Tong Yu, Yusheng Zheng, Andrew Quinn.

”wBPF: Efficient Edge-Case Observability for CXL Pooling systems via eBPF” HCDS25 Yusheng Zheng, Tong Yu, **Yiwei Yang**, Andrew Quinn.

”Extending Applications Safely and Efficiently” OSDI25 Yusheng Zheng, Tong Yu, **Yiwei Yang**, Yanpeng Hu, Xiaozheng Lai, Dan Williams, Andrew Quinn.

”Transparent and Efficient Live Migration across Heterogeneous Hosts with Wharf” Preprint **Yiwei Yang**, Aibo Hu, Yusheng Zheng, Brian Zhao, Xinqi Zhang, Andrew Quinn

”CXLMemSim: A pure software simulated CXL.mem for performance characterization.” Yarch23 **Yiwei Yang** Pooneh Safayanikoo, Jiacheng Ma, Tanvir Ahmad Khan, Andrew Quinn.

SELECTED PORTFOLIOS

TVA <https://github.com/Sluglab/TVA> Exploiting Control-flow Enforcement Technology for Sound and Precise Static Binary Disassembly

MVVM <https://github.com/Multi-V-VM/MVVM> An LLM agent debugging playground

ZLUDA <https://github.com/Multi-V-VM/hetGPU> Live Migration middleware for Heterogeneous GPU

SKILLS

- **Programming Languages:** not limited to any specific language, and experienced in Python/C++/Rust, comfortable with Golang/C/Java/Scala/TypeScript (in random order).
- **System:** Specialist in Compiler & Performance Analysis, familiar with LLVM, MLIR, Gem5, WASM, gdb, eBPF, ptx, qemu, chisel, Linux mm & observability subsystem.
- **Machine Learning:** familiar with general knowledge of machine & reinforce learning, interested of System for LLM for Sys.

MISCELLANEOUS

- Interests: HPC, Architecture, Security, JIT, CXL, LLM4Sys4LLM, etc.
- Lead GeekPie_HPC *Ranked 2*, SC-SCC21. *Ranked 4*, ISC22. Advise Not-Slow-Slug *Ranked 2*, ISC23.
- Advise Kiki Zhang, Zhen Cao to be Ph.D. students.
- Google Summer of Code Mentor for BPFTIME