Tommy Zhao

(legal name: Kaida Zhao)
tzhao451@gmail.com
1(312)375-8295
Personal, LinkedIn, GitHub
Chicago, IL
Open to relocation
US citizen

About Me

Summary

Hard-working, passionate, and dedicated software engineer with a graduate degree and 7 years of professional experience that intersects software and hardware for high performant and cost-efficient computing.
Great teamwork to write clean and scalable code for fast and robust software systems.
Committed to continuously learning and teaching to keep up with current technologies.

Skills

Programming Languages: assembly, C, C++, CUDA, Java, JavaScript, MATLAB, Python, Solidity, SQL, TypeScript, and more
Programming Knowledge: AWS tools, build systems (make/bazel), code reviews, compilers, databases, datacenters, production software releases, version control (svn/git/mercurial), scalability, scheduling, and resource monitoring
Programming Paradigms: high performance parallel programming (GPUs and multicores), distributed systems, blockchain programming, OOP, and functional programming

Education

M.S. in Computer Science from University of Wisconsin-Madison in 2015
B.S. in Electrical and Computer Engineering with Honors and with minors in Computer Science, Math, and Biology from University of Illinois-Chicago in 2014

Work Experience

Backend Software Engineer, Aperture Finance, *Mountain View, CA* — Mar 2024 – Mar 2025
- Wrote block-chain smart contracts and automation SDK for decentralized finance liquidity management.
- Implemented and maintained swapping, increasing/decreasing liquidity, reinvesting, rebalancing, staking, and market making.
- Implemented solvers for optimal swap routing for best quotes.

Software Engineer, Google, *SF Bay Area, CA*

Ads budgeting infrastructure team — Apr 2020 – Mar 2023
- Reduced system wide resource utilization by 80% by leveraging global sharding framework, reconfiguring worker resources, and reducing extraction frequency.
- Migrated ads budgeting infrastructure to use new hierarchical budgeting entities, which enables more budgeting features.
- Owned and on-call for budget extractor and monetizer, vital systems for ads, which is ~80% of Google's revenue.

Hardware/software codesign for web search team — July 2018 – Apr 2020
- Evaluated the performance of a deep machine learning NLP transformer model (BERT) for web search.
- Evaluated fleet-wide resource utilization for all web search clusters to identify bottlenecks and trends, and made hardware configuration recommendations.

CPU Diagnostics Engineer, Advanced Micro Devices (AMD), *Austin, TX* — Aug 2017 – July 2018
- Found, reported, and root caused post-silicon hardware bugs by generating instructions to stress microarchitectural features.

Various: Lecturer, Research Assistant, Teaching Assistant, Embedded Engineer, Automation Software Engineer, SDET
- Prepared and presented lectures for a class of 132 students.
- Researched programmable accelerators, fair resource distribution with connectivity constraints, and hardware security and IC piracy.
- Wrote scripts to automate grading programming assignments.
- Implemented, tested, debugged, and released a real-time operating system (VxWorks) for custom hardware.
- Wrote, ran, and maintained automated test cases to find bugs.

Projects

Financial Portfolio Visualizer (devpost)
- Upload an excel, csv, or text to visualize live (during market hours only) portfolio performance from various time intervals.
- Try it yourself here. Example: cvs generates this portfolio viz.

UniswapV4 Backtester - hackathon winner & Uniswap Foundation 1st place prize for research and integration (showcase, slides)
- Evaluates UniswapV4 trading strategies with specified hooks using real, historical data for traders and liquidity providers.
- Extracts and replays liquidity events in specified time interval.

Bonsai, a distributed data collection and storage system for data processing (pdf)
- Scalable data collection in a ring structure with a circular buffer for storing data and a token for synchronization logic.
- Real-time data processing in a tree structure by multicasting map requests to backends and reducing results towards the frontend.