






Chi-Wei (Tomy) Hsieh

 cv@tomy.me  (412) 584-1843  https://tomy.me  tomy0000000  tomy0000000

Software Engineer specialized in scalable backend, DevOps, and distributed systems for production-grade applications.

EDUCATION

- Carnegie Mellon University**
Master of Software Engineering - Scalable Systems
• Coursework: Cloud Computing, Intro to Database Systems, Introduction to Computer Systems
- Pittsburgh, PA
December 2024
- National Chung Hsing University**
Bachelor of Science in Computer Science
• Honorable Mention for Graduate Capstone Project: *Question Generation Quality Enhancement based on BART*
- Taichung, Taiwan
June 2022

WORK EXPERIENCE

- WeRide**
Software Engineer
- San Jose, CA
February 2025
- Designed and deployed a Kubernetes-based Cloud IDE for 500+ engineers across multiple regions, enabling on-demand CPU/memory/GPU provisioning and strict data localization.
 - Optimized GPU usage by ~40% via just-in-time mounting/unmounting, reducing idle costs and boosting utilization.
 - Secured the monorepo by restricting access to authenticated IDE instances, mitigating exfiltration risks while streamlining developer workflows.
- WeRide**
Software Engineer Internship
- San Jose, CA
May 2024 - August 2024
- Revamped a large-scale CI/CD pipeline architecture, tripling task capacity and slashing deployment time by 80%.
 - Developed automated deployment workflows in Go, improving reliability of feature rollouts across multi-cloud environments.
 - Bolstered cloud security via sidecar containerization, enhancing data isolation and compliance with internal standards.
- Intel**
Software Engineer Internship
- Taipei, Taiwan
July 2021 - December 2021
- Engineered and sustained a Django-based full-stack solution to visualize test reports, cutting manual analysis time by 50%.
 - Accelerated bug detection and resolution by 70% using automated Python validation routines, driving faster release cycles.
 - Boosted hardware test capacity by 80% through time series models in a FastAPI backend for real-time analytics.

SKILLS

Programming Languages	Python, Go, C/C++, Java, Shell/Bash, JavaScript, SQL
Data	MySQL, Postgres, MongoDB, Kafka, Hadoop, Spark, Samza
Web & Frameworks	Flask, Django, FastAPI, Gin, React, Next.js
Cloud, DevOps	AWS, Google Cloud, Azure, Linux/Unix, Docker, Kubernetes, Helm, Ansible, Terraform

PROJECTS

- BusTub - RDBMS Implementation with C++17**
 - Implemented Presto's dense layout HyperLogLog for fast cardinality estimation over large datasets.
 - Developed a thread-safe buffer pool manager with LRU-K replacement and a disk scheduler to optimize memory management and disk I/O efficiency.
 - Built a concurrent B+Tree supporting efficient search, insertion, deletion, and in-order iteration with proper concurrency control mechanisms.
- Real-Time Ride Matching and Ad Targeting System**
 - Implemented high-throughput real-time data pipelines on AWS EMR with Kafka and Samza, ensuring near-zero-latency stream processing.
 - Optimized driver-matching and personalized ad-targeting algorithms, boosting throughput and lowering cost of operations.
- tw-invoice + K**
 - Developed a single source of truth for personal finances system with FastAPI, PostgreSQL, Next.js and Docker Compose, automating invoice retrieval, and streamlining multi-country expense management.
 - Implemented multi-currency & i18n with UTC-based data storage and Decimal fields for reliable cross-border, multi-account expense tracking.
 - Ensured high reliability and scalability via containerized architecture, CI/CD pipelines, and Pydantic data validation.