Chi-Wei (Tomy) Hsieh

✉ cv@tomy.me 📞 (412) 584-1843 🌐 https://tomy.me ⬤ tomy0000000 in tomy0000000

EDUCATION

Carnegie Mellon University   Pittsburgh, PA
Master of Software Engineering - Scalable Systems   Graduating December 2024
- Cloud Computing, Intro to Database Systems, Distributed Systems, Introduction to Computer Systems

National Chung Hsing University   Taichung, Taiwan
Bachelor of Science in Computer Science   June 2022
- Honorable Mention for Graduate Capstone Project

WORK EXPERIENCE

WeRide   San Jose, CA
Software Engineer Internship   May 2024 - August 2024
- Tripled the CI/CD pipeline's task capacity by revamping system architecture for optimal performance.
- Slashed deployment time by 80% by innovating new features and orchestrating chaining deployment workflows in Go.
- Fortified the company's cloud infrastructure by deploying sidecar service containers segregated in Kubernetes.
- Expanded the user base by 30% by designing and implementing highly intuitive and user-friendly features.

Intel   Taipei, Taiwan
Software Engineer Internship   July 2021 - December 2021
- Engineered and sustained a Django full-stack website for dynamic test report visualization.
- Accelerated the bug detection and resolution process by 70% by deploying an automated validation routine.
- Maximized system utilization 80% by integrating time series models into a FastAPI backend platform.
- Enhanced model accuracy by 12% by enforcing rigorous data validation in the data pipeline using Pydantic.

SKILLS

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

PROJECTS

BusTub - RDBMS Implementation with C++17
- Implemented the Presto's dense layout HyperLogLog algorithm for fast cardinality estimation in large datasets.
- Devised a thread-safe buffer pool manager and LRU-K replacement policy and a disk scheduler to optimize memory management and disk I/O efficiency.
- Developed a concurrent B+Tree, supporting efficient search, insertion, deletion, and in-order iteration with proper concurrency control mechanisms.

Twitter PageRank Recommendation System
- Architected a large-scale PageRank Spark application to analyze Twitter graph data on Azure HDInsight and Databricks.
- Designed a PySpark ETL pipeline that processed over 1TB and loaded data to the MySQL database on AWS.

Real-Time Ride Matching and Ad Targeting System
- Implemented high-throughput real-time data pipelines on AWS EMR, ensuring seamless stream processing at scale.
- Enhanced driver-matching and ad-personalizing by optimizing Kafka and Samza streaming performance.

Tubee - Automated YouTube Subscription Platform
- Constructed the project from scratch with Flask, PostgreSQL, and Docker.
- Processed over 1,000 new videos daily and delivered personalized recommendations to users faster than official app.

tw-invoice + K
- Created and Engineered an SDK for exchanging data with Taiwan MoF e-Invoice platform, accelerated application development by incorporating built-in data-validating features.
- Architected a spending tracking system with FastAPI, PostgreSQL, and Docker, enabling users to manage and analyze their spending habits.