

Andrew Ho

github.com/anmho
andrewho.io

EDUCATION

University of California, Irvine

Grad. June 2025

Bachelors in Computer Science

GPA 3.7

Courses: Efficient ML Computing, Data Structures, Algorithms, Search Engines, Distributed Systems, Operating Systems, Databases, Parallel Computing

SKILLS

Languages: Go, Python, TypeScript, Java, SQL, C, C++

Tools: Kubernetes, gRPC, Docker, ArgoCD, GraphQL, Git, PySpark, PyTorch, pandas, numpy

Cloud Services/Databases: AWS, GCP, Kubernetes, Stripe, DynamoDB, Kafka, Cassandra, PostgreSQL, MySQL

EXPERIENCE

Snap Inc.

Aug. 2025 - Present

Software Engineer - AI Inference Caching Infra

Los Angeles, CA

- Building heavyweight model inference caching infrastructure with C++ and storage backends like Valkey, Redis, and KeyDB.

Amazon Web Services

May. 2025 - July 2025

Software Development Engineer Intern - FSx

New York, NY

- Supported 100% YoY product growth by building scalable infrastructure to manage petabyte-scale file systems.
- Designed and implemented a distributed “orphanator” system in Java using AWS S3, EBS, and Simple Workflow Service (SWF) to detect and delete orphaned snapshot shards, reducing storage bloat and improving cost efficiency

Tesla

Sep. 2024 - Jan. 2025

Software Engineer Intern

Palo Alto, CA

- Developed aerodynamics tools with Go, React, Kafka, and Spark to streamline the vehicle engineering lifecycle.
- Integrated StarCCM+ CFD simulation data streams with Kafka and Spark streaming to provide real-time updates on simulation jobs, enabling faster decision-making and accelerating the engineering design loop.

Snap Inc.

Jun. 2024 - Sep. 2024

Software Engineer Intern - Messaging

Los Angeles, CA

- Developed new features for the Snapchat messaging core service using *Go*, *gRPC*, *DynamoDB* and *AWS*
- Built chat wallpaper feed bump and notification feature, resulting in a \$600,000+ increase in Snapchat+ subscriptions.
- Collaborated with Meta to create rich URL previews for *Instagram* chat links using oEmbed Graph API.

Snap Inc.

Jun. 2023 – Oct. 2023

Software Engineer Intern - Monetization

Los Angeles, CA

- Designed and implemented multi-dimensional range partitioning for *Apache Druid* analytics database cluster.
- Created *Apache Beam* & *Spark (PySpark)* pipelines to support range partitioning, speeding up queries by 35%.
- Deployed *Apache Kafka* cluster using *Kubernetes* to process 80TB/day analytics metrics stream from GCS

AI Hardware Acceleration Researcher

Apr. 2024 – Jun. 2024

Donald Bren School of Information and Computer Science

Irvine, CA

- Researched low-latency LLM/transformer optimization strategies for hardware acceleration under Prof. Thomas Yeh, focusing on performance-per-watt tradeoffs.

NASA

May. 2024 – Jul. 2024

Machine Learning Engineer Intern

Los Angeles, CA

- Developed high-performance systems in C++, CUDA, and PyTorch to support NASA’s Lucy rover autonomy.

PROJECTS

Push Notification Scheduler | *Go*, *AWS Lambda*, *DynamoDB*, *AWS Event Scheduler*, *AWS CDK*, *Expo*

- Built a serverless multi-tenant microservice for sending push notifications to app users and scheduling one-time delayed push notification execution for React-Native mobile apps.

Distributed Key-Value Store | *Go*, *gRPC*, *Docker*, *Kubernetes*

- Developed multithreaded distributed key-value store cluster in *Go*, using *B trees*, *LSM trees*, and *bloom filters*
- Implemented hot-cold system using LRU eviction and compaction using Raft replication.