

Steven Zhou

✉ 647-838-7228 | 📩 stevenzhou2526@gmail.com | 🌐 steventzhou

Education

University of Waterloo

Honours Bachelor of **Computer Science** – Dean’s List

Waterloo, ON

Sep 2021 - Apr 2026

Courses: Operating Systems, Distributed Systems, Concurrency, Performance, Compilers, Databases, Artificial Intelligence

Skills

Languages C++, C, Python, Java, Rust, SQL

Frameworks Abseil, Boost, Spark, Pandas, NumPy, PyTorch, TensorFlow

Tools AWS, MySQL, PostgreSQL, GraphQL, Redis, CUDA, LLVM, MLIR, Clang, Bazel, Kubernetes, Docker, Jenkins, Git

Experience

OpenAI

San Francisco, CA

Member of Technical Staff - Database Systems - C++

Incoming 2026

- Building a **real-time search, analytics**, and **vector** database powering **retrieval-augmented generation** across research and products

Meta

Menlo Park, CA

Software Engineer Intern - Machine Learning Infrastructure - Python, C++

May 2025 - Aug 2025

- Optimized adaptive ads placement with **user sensitivity modeling** to increase annual revenue by **\$120 million** over Facebook surfaces
- **Auto-tuned parameters** with **Bayesian Optimization** over multi-treatment predictions to ensure safe and efficient **signal exploration**
- Launched a no-code workflow for **90 engineers** and **15 teams** to automatically train and generate code changes for model parameters
- Designed a scalable **MySQL** database to store deployed parameter tuning history and live traffic feedback for **synthetic data** generation

Snowflake

San Mateo, CA

Software Engineer Intern - Query Processing - C++

Jan 2025 - Apr 2025

- Researched and developed **distributed** and **performant** infrastructure to profile process level **memory allocation** and **thread** states
- Introduced an **asynchronous** worker pool to perform **concurrent system calls** to reduce resource contention by **spilling** buffered data
- Optimized the core query **task scheduler** to coordinate process resources adaptively and eliminate out-of-memory scenarios by **30%**
- Designed a **lightweight** topological traversal algorithm to enforce **runtime invariants** and maintain consistency across query plans

Snowflake

Bellevue, WA

Software Engineer Intern - SQL Compiler - C++, Java

May 2024 - Aug 2024

- Directed the compilation, execution, and rollout of **7** new range-based **SQL** window functions to increase annual revenue by **\$2 million**
- Introduced a compute node within the query plan to maximize **concurrent data streaming** throughput and reduce processing overhead
- Defined **scalable** partitioning heuristics for **trillions** of table rows to enable efficient data allocation over **distributed** compute nodes
- Enhanced null handling across **6 billion** daily queries by maintaining a verification bit to reduce **CPU** usage by **4%** during disk operations

BlackBerry

Waterloo, ON

Software Engineer Intern - Core Infrastructure - Python, C++

Jan 2024 - Apr 2024

- Shipped a **highly available** cloud-native **containerized** change management platform with **MySQL, Docker, AWS ECS**, and **Terraform**
- Maintained scalability through **parallelized** request handling, **streamed** API responses, and **Redis** caches to reduce database accesses

Bank of Montreal

Toronto, ON

Software Engineer Intern - Data Ingestion - Python

May 2022 – Aug 2022

- Integrated in-memory caching and **data encryption** by migrating **5TB** of enterprise data from **PostgreSQL** to SAP HANA and **AWS S3**
- Developed **ETL** pipelines to automate data analysis and streamline auction reporting for bi-weekly CTO meetings with **1500** attendees

Projects

CastorOS

- Operating system built in **C**, capable of **threading, scheduling, synchronization**, and **file system** operations through system calls

C Compiler

- Compiler written in **C++** to convert **C** code into **machine instructions** through tokenization, syntactic parsing, and semantic analysis