

Honglin Cao

Toronto, ON h45cao@uwaterloo.ca 647-939-8018 v2ark.com linkedin.com/in/v2ark github.com/V2arK

Education

University of Waterloo, Bachelor of Computer Science

Sep 2020 – Aug 2025

- President's Scholarship of Distinction, Faculty Cumulative Average: **93%** / GPA: **4.0**

Experience

Platform Software Engineer — CentML - Toronto, ON

Sep 2024 – Aug 2025

- Led the development of a local development environment mirroring the actual server setup implemented in **Infrastructure as Code** on **Pulumi** with **Python**, interacting with **Kubernetes**, **Docker**, **Knative**, and **AWS / Google cloud**. Setup and troubleshoot tools such as **LocalStack**, **Grafana**, **Prometheus**, **Istio** and **Minikube** with **Helm** and adjust them to resolve critical issues related to billing, monitoring, deployment, and database mocking, ensuring identical API interactions and deployment processes with the production environment.
- Designed and implemented new **APIs** integrated with container deployment, billing, and user storage on **PostgreSQL** in **Python**, while ensured upgrade / downgrade path with **SQLAlchemy**, adhering to modern safety standards to protect against malicious users.
- Automated key processes, including releasing **public API Clients** and non-root **Container Images** for user deployment and control plane services upon platform releases with **Github Actions**, designed and built **integration tests** in **Python** utilizing **pytest**, **jwt**, **WorkOS**, ensuring operations without intervention and providing up-to-date user experiences.
- Contributes on projects including **llama-stack**, **huggingface_hub**, **litellm** and **huggingface.js** in **Python** and **JavaScript**, alongside with comprehensive **documents** to promote our platform and provide alternative ways to interact with our serverless endpoints.

Distributed Database Engineer — Huawei - Markham, ON

Jan 2022 – Jan 2024

- Designed an **RPC** protocol over **TCP** and **RDMA** in **C** to eliminate size limits, enabling **crash recovery** messages on multi-node **GaussDB** configurations.
- Quantified database performance with **perf**, **gstack**, **vmstat/iostat**, **CPU Flame Graphs**, and **jTPCC**. Automated the process as a program with GUI using **Bash**, **Python**, and **HTML/CSS/PHP**.
- **Standardized** automated **TPC-C** benchmark on single-node, physical and logical multi-node **GaussDB** configurations with templates in **Groovy**, **Bash**, **Python**, **Java**, **GitLab CI**, and **Jenkins** across **ARM** and **x86** environments, maintained and adapted them to suit rapid development goals.
- **Managed** servers to suit developers' needs; **troubleshoot** issues ranging from faulty link negotiation settings to low performance caused by unoptimized **sysctl** settings, **allocated** and **set up** working environments for developers, and **negotiated** with headquarters for resources needed across the teams.

Projects

Fluid Simulation — C++, OpenGL

Apr 2024

- A **Weakly Compressible Smooth Particle Hydrodynamics** simulation with **Rasterization** on **OpenGL shaders**.

Pet Health Monitor — Python

Jan 2024

- Fine-tuned **YOLO-v8** on personal datasets, achieving **98%** accuracy on validation, status detection within **200 ms** on low-power **IoT** devices.

Skills

Languages: C++, C, Python, Go, Java, C#, SQL, Bash, Groovy, HTML, CSS, R, ~~TeX~~^{LaTeX}, JavaScript, PHP, Kotlin ...

Tools: Docker, Kubernetes, Jenkins, GitLab, VS Code, Postman, PostgreSQL, GDB, Valgrind, OpenGL, CUDA ...