Chuyue Zhang

Distributed Systems Engineer

⚲ Markham, ON, Canada   ✉ zhangchuyue.zhang@mail.utoronto.com   ☎ 14373438066

in Chuyue Zhang   ⌨ IronDumpling

Skills

Languages: C++, Modern C++ (11/14/17), C#, Java, Go, Python, Rust, SQL

Databases & Data Platforms: PostgreSQL, MySQL, MongoDB, Redis, Cassandra, DynamoDB, ClickHouse; Query Optimization, Partitioning & Indexing Strategies, Execution Planning; Data Modeling (OLTP/OLAP), Columnar Storage, Materialized Views, Paruqet; Data Warehousing (Snowflake), PySpark

Distributed Systems: ACID Transactions & Distributed Transactions (2PC); CAP Theorem, Sharding & Rebalancing, Replication Topologies; Scalability & Fault Tolerance, Consensus Algorithms (Raft, Paxos); Distributed Caching (Redis Cluster); Message Queues (Kafka); RDMA / Remote Direct Memory Access

Web & API Development: Node.js, Express.js, Django, .NET Core, RESTful APIs, Microservices, MVC Frameworks

DevOps & Cloud: Docker, Kubernetes, AWS S3, Microsoft Azure, Jenkins, Git, Shell Scripting

Additional: Linux Kernel, Parallel Programming, Performance Benchmarking (TPC-H/TPC-C/TPC-DS/sysbench), TCP/IP, Socket Programming, LLVM, Deep Learning, Generative Learning, AI Agent, TypeScript, React.js, GPU Programming

Experience

Huawei Canada, Distributed Database Team (DDSM) · **Software Engineer (Full-time)**   Markham, ON   Jul 2024 – Jul 2025

- **Accelerated** TPC-H query runtime by **4.7×** (28 s → 6 s, 1-node cluster) and **6.2×** (28 s → 4.5 s, 3-node cluster) on the GaussDB, a PostgreSQL-based database.
- **Optimized** Parquet storage of the GaussDB, boosting sysbench point-select TPS by **60%** and reducing 3-node TPC-H performance to **5.8 s**.
- **Extended** the distributed query system utilizing a share-everything architecture on the GaussDB, efficiently supporting **SF-300** scale TPC-H workloads.
- **Designed** and **implemented** a 3-node Jenkins CI pipeline, completing **1,500+** automated runs, enabling **40+** engineers to validate code changes rapidly.

C-MORE Lab, Robotics Team · **Machine Learning Researcher (Intern)**   Toronto, ON   May 2023 – Dec 2023

- **Engineered** a pipeline for converting 2D maps to 3D Gazebo environments using ROS and XML plugins, automating batch simulations for multi-robot scenarios.
- **Developed** reinforcement learning–based exploration algorithms, improving multi-robot coverage efficiency by **45%**.
- **Applied** Bayesian Optimization for robot parameter tuning, reducing trial count by **50%** in real-world deployment scenarios.

Intel Corp., Programmable Solutions Group (PSG) · **Software Engineer (Intern)**   Toronto, ON   May 2022 – Jun 2023

- **Streamlined** chip analysis tools by combining PostgreSQL query tuning and Pandas preprocessing, reducing end-to-end analysis latency by **85%**.
- **Implemented** Django-powered dynamic query features, enabling **20+** engineers to customize chip data visualizations.
- **Managed** 10 M+ Quartus chip database with PostgreSQL and **integrated** it into internal CI/CD regression pipelines

Projects

Dynamo/Cassandra-Style Distributed Key-Value Store   distributed-storage-service ☑

- **Designed** a peer-to-peer, masterless KV store with consistent hashing, virtual-node support, and 3-replica redundancy to ensure high availability and minimal data movement during node join/leave.
- **Implemented** dynamic cluster membership, failure detection, and eventual consistency model, mirroring architecture of Cassandra's replication strategies.
- **Benchmarked** performance under 10K QPS with <2ms read latency and near-linear throughput scaling.
- **Technologies:** Java, TCP/IP sockets, Consistent Hashing, Virtual Nodes, Failure Detection, Fault Tolerance

ClickHouse-HE Distributed Query on Homomorphic Encrypted Database   clickhouse-he-encrypted-distributed-query ☑

- **Extended** the distributed query system that enables secure filtering, aggregation, and join operations over homomorphically encrypted data.
- **Built** a secure middleware and query planner to route encrypted queries across multiple shards, ensuring confidentiality through the entire query lifecycle.
- **Technologies:** C++, ClickHouse, HElib, Homomorphic Encryption, Distributed Query, Privacy-Preserving Computation

Enhance Redis Performance with RDMA & VMA   redis-rdma-vma ☑

- **Extended** Redis 7.2.4 with one-sided RDMA reads for snapshot replication, cutting full resync time by **>10×** with zero code changes.
- **Integrated** NVIDIA VMA user-space TCP offload to bypass the Linux kernel, boosting single-node throughput by **66%** and cluster throughput by **30–125%**.
- **Benchmarking** across 1 GB–20 GB datasets and 1–9 node clusters, demonstrating linear scalability and identifying CPU vs. network bottlenecks.
- **Technologies:** Redis, RDMA (verbs), NVIDIA VMA, Linux socket bypass, perf, C/C++

SmallC Compiler with LLVM   LLVM-small-C-Compiler ☑

- **Developed** a C-subset compiler featuring AST construction, symbol table, semantic analysis, LLVM IR generation, and IR optimizations.
- **Validated** correctness against 90+ SmallC test cases with 100% pass rate.
- **Technologies:** C++, LLVM, ANTLR4, Compiler Internals

Anime Goods E-Commerce Platform   anime-goods-e-commerce ☑

- **Engineered** a Node.js/Express backend with PostgreSQL for order, inventory, and user management, supporting 5k daily transactions.
- **Integrated** Redis caching and JWT-based auth to reduce average API response from 120 ms to 40 ms.
- **Technologies:** Express.js, PostgreSQL, React, TypeScript

Education

MEng   **University of Toronto**, Computer Engineer   Sept 2024 – Dec 2025
- Expected Dec 2025

BASc   **University of Toronto**, Computer Engineer   Sept 2019 – May 2024
- **GPA:** 3.76/4.0
- **Graduate with Honours**
- **Dean's Honours List** (**5** Semesters)