

Wuwen Wang

+65 8654-8320 | wuwenwcn@gmail.com | [GitHub](#) | [LinkedIn](#)

EDUCATION

Carnegie Mellon University

Master of Science in Electrical and Computer Engineering

Dec. 2020

GPA: 3.87/4.00

University of Illinois at Urbana-Champaign

Bachelor of Science with Honors in Computer Engineering

May 2019

GPA: 3.67/4.00

TECHNICAL SKILLS

Programming Languages: C++, SQL, Java, Rust, C, Scala, Python, Bash

Tech Stack and Tools: LLVM, CLANG, CMake, GDB, MySQL, PostgreSQL, NoSQL, Spark, GCP, AWS, HDFS, Docker, CI/CD, Perf, OpenSSL

WORK EXPERIENCE

TikTok/ByteDance Inc.

Software Engineer, C++ Database Developer

Dec. 2022 - Present

Singapore

- Contributed to Bytehouse's development, a cloud-native OLAP database forked from ClickHouse, featuring compute and storage separation, with code primarily in C++17. ([ByteHouse](#))
- Engineered a MySQL compatibility framework for Bytehouse, introducing a dialect for seamless MySQL migration, thus simplifying user adoption and improving experience.
- Designed and implemented over 100 MySQL-specific functions, with frameworks for DDL/DML/DQL support, case-insensitivity and implicit conversion, thereby improving system functionality and data handling efficiency.
- Developed a SQL-driven Python UDF feature using a hybrid architecture with in-database processing and Python executors in side-car Docker, resulting in a 6x performance enhancement through zero-copy shared memory techniques and detailed permission management.
- Enhanced Bytehouse's integrity and reliability by stabilizing the CI pipeline, promptly addressing critical bugs, and providing on-call and PoC support.
- Provided mentorship to two interns and engaged in the hiring process through candidate interviews. Actively contributed to establishing the team's development cycle and enhancing cross-team collaborations.

Swiss Federal Institute of Technology Lausanne/EPFL

Research Assistant/Ph.D. student, supervised by Prof. Anastasia Ailamaki

Sept. 2021 - Aug. 2022

Lausanne, Switzerland

- Contributed to the Proteus just-in-time executor written in C++ 20. ([Proteus](#))
- Implemented a GPU-accelerated transactional system that benefits from the massive parallelism of GPU, achieving 2.1x throughput than CPU-based execution engine.
- Extended the YCSB benchmark to evaluate the transactional system's throughput, scalability, and adaptivity.
- Improved performance bottlenecks in GPU transaction processing such as thread synchronization overhead.
- Investigated a transaction admission system with efficiency, usability and throughput, by pushing the execution logic from database to application through code injection during compilation.

StarTree Inc.

Software Development Engineer Intern

Apr. 2021 - Aug. 2021

Mountain View, CA, USA

- Worked as a developer for Apache Pinot, a real-time distributed OLAP datastore written in Java. ([Apache Pinot](#))
- Programmed an efficient top-k implementation in Pinot's aggregator using quick-select algorithm to achieve an optimal runtime and evaluated against the priority-queue based approach.
- Reduced network overhead, memory consumption, and aggregation latency by embedding trimmers in Pinot's distributed workers to push down the result filter.

RESEARCH EXPERIENCE

Carnegie Mellon Database Group

Jan. 2020 - Mar. 2021

Research Assistant, supervised by Prof. Andy Pavlo

Pittsburgh, PA, USA

- Contributed to the NoisePage in-memory database management system written in C++17. ([NoisePage](#))
- Implemented parallel CREATE INDEX in the execution engine to support runtime code generation for index population using LLVM compiler infrastructure, yielding a speedup in analytical query execution.
- Ran experiments on TPC-C benchmark to investigate the effectiveness of indexes, analyze the latency of index creation, and indicate that the self-driving model worked correctly on DBMS.
- Built a compilation manager to optimize the query compilation process by supporting an adaptive execution mode that smoothly replaces interpreted code with byte-code asynchronously generated in the background.
- Implemented SQL LIKE operator as a runtime generated predicate for SELECT statement and sequential scan.
- Added support for constraint creation and enforcement for UNIQUE, Primary Key, and Foreign Key, in a Postgres convention to ensure the code compatibility.

PROJECTS

ByConity | *Open-source Project* | C++

Dec. 2022 – Present

- Served as Core maintainer of ByConity, the open-source version of Bytehouse, leading community collaboration, and ensuring project stability and growth. ([ByConity](#))
- Engaged actively in the ByConity community, resolving issues, contributing to roadmap and developments, influencing both the current and future directions of the project.

ProjectDV | *Personal Project* | Rust

Sept. 2023 - Present

- Collaborated with a team of four on a geo-location-based application for Android and iOS, enabling users to visualize past journey routes on a map interface.
- Implemented an efficient caching layer in the application to enhance map rendering speed, incorporating timestamp validation for data integrity and freshness.

HONORS

The Dean's List of College of Engineering

Recipient

University of Illinois at Urbana-Champaign

Earl J. Eckel Scholarship

Recipient

University of Illinois at Urbana-Champaign

PUBLICATIONS

1. Lin Ma, William Zhang, Jie Jiao, **Wuwen Wang**, Matthew Butrovich, Wan Shen Lim, Prashanth Menon, and Andrew Pavlo. "MB2: Decomposed Behavior Modeling for Self-Driving Database Management Systems". In: *Proceedings of the 2021 International Conference on Management of Data*. SIGMOD '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1248–1261. isbn: 9781450383431. doi: 10.1145/3448016.3457276.
2. Ling Zhang, Matthew Butrovich, Tianyu Li, Yash Nannapaneni, Andrew Pavlo, John Rollinson, Huanchen Zhang, Ambarish Balakumar, Daniel Biales, Ziqi Dong, Emmanuel Eppinger, Jordi Gonz'alez, Wan Shen Lim, Jianqiao Liu, Prashanth Menon, Soumil Mukherjee, Tanuj Nayak, Amadou Latyr Ngom, Jeff Niu, Deepayan Patra, Poojita Raj, Stephanie Wang, **Wuwen Wang**, Yao-Tin Yu, and William Zhang. "Everything is a Transaction: Unifying Logical Concurrency Control and Physical Data Structure Maintenance in Database Management Systems". In: *Conference on Innovative Data Systems Research*. 2021. <https://api.semanticscholar.org/CorpusID:228096089>.
3. Liu, C., Das, A., **Wang, W.**, Kuchemann, S., Kenesei, P., Maass, C. R. E. (2019). "Shear-band cavities and strain hardening in a metallic glass revealed with phase-contrast x-ray tomography". *Scripta Materialia*, 170, 29-33. <https://doi.org/10.1016/j.scriptamat.2019.05.021>