

Congrui Yin

✉ yin00486@umn.edu | 🌐 github.com/JerryYin777 | 🌐 jerryyin777.github.io | 📄 Google Scholar

Education

University of Minnesota Twins Cities	2024/01 - 2025/06 (Expected)
Candidate in Bachelor of Arts in Computer Science	Minneapolis, MN, USA
Nanchang University	2021/09 - 2023/12
B. Eng in Artificial Intelligence	Nanchang, China
<ul style="list-style-type: none">• GPA 93.80/100, Rank 1/149, National Scholarship 2019 (Top 0.2% national-wide)• A+ Courses: Operating Systems, Computer Architecture, Computer Networks, and 28 others	

Work Experience

RisingWave Labs	2021/08 - 2022/07
Database System R&D Intern	Shanghai, China
<ul style="list-style-type: none">• Top contributor of RisingWave as of 2023/03. RisingWave is a database system with PostgreSQL-compatible interface that incrementally maintains materialized views. Worked on the development of almost all components related to stream computing and state store.• Streaming Index Joins: Designed shared state and streaming index in RisingWave; implemented index lookup join executor; implemented delta join DAG optimizer transformations; implemented distributed delta join scheduler• Performance Improvement: Conducted intensive benchmarks and analyzed performance issues. Fixed bugs, proposed strategies, and led cross-team collaboration which improved the system throughput by 10x in a 3-month period• Developer Experience. Initiated the RiseDev development tool to start a RisingWave cluster with one command, which is deeply integrated into the development workflow across unit testing, integration testing, and benchmarking. Built RisingWave Streaming Dashboard that shows cluster metrics and visualizes streaming query plans in one place• Mentoring. Mentored database kernel interns and helped their integration into the team. Maintained overview documents of the database kernel to facilitate knowledge transfer and help new hires learn about the system.	
ByteDance, Ltd.	2021/06 - 2021/08
Storage System R&D Intern, TerarkDB Team	Beijing, China
<ul style="list-style-type: none">• Co-Optimized TerarkDB and ZenFS. Implemented Zone-aware Garbage Collection in TerarkDB and WAL-Aware Zone Allocator in ZenFS, which reduced 3-4x of space amplification and improved the p999 tail latency by 100x	
PingCAP, Inc.	2020/08 - 2021/01
Storage System R&D Intern	Shanghai, China
<ul style="list-style-type: none">• Built LSM-based storage engine AgateDB from ground-up. Inspired by WiscKey and BadgerDB, AgateDB separates large values from LSM tree into value log, so as to reduce write amplification.	

Open-Source Contributions

cmu-db/bustub 🌟	2022/08 - Now
<ul style="list-style-type: none">• Lead the development of the BusTub database system. Added SQL support/query processing layer to the system.• Redesigned course projects to help students better understand the concepts and apply them to real-world scenarios. Developed leaderboard tests to challenge advanced students and enable further study.	
RisingLight Community 🌟	2022/01 - Now
<ul style="list-style-type: none">• Lead the development of RisingLight, an OLAP database system for educational purpose.	
TiKV Community 🌟	2020/05 - Now
<ul style="list-style-type: none">• Maintains TiKV Coprocessor, the push-down execution framework of TiDB. Mentored community members to contribute features (e.g. new data types, plugin system) in the LFX Mentorship. 🌟 🌟	

Personal Projects

	5.3k followers 🌟 skyzh
<ul style="list-style-type: none">• mini-lsm (★1k) Build a simple LSM-Tree storage system in Rust in a week• type-exercise-in-rust (★1k) Learn Rust generics by implementing a vectorized expression evaluation framework	