Shicong Nie

13026339196 | niebayes@gmail.com | github.com/niebayes

Summary

- Open source enthusiast participating in Google Summer of Code with Jina AI and Open Source Promotion Plan with Apache Doris.
- Passionate about technology, participated in PingCAP TinyKV Bootcamp and OceanBase miniob Database Competition.
- Interested in distributed storage, implemented Multi-Raft, Multi-Paxos, and LSM-DB.
- Strong self-motivation, studied renowned open courses such as MIT 6.824, CMU 15-445, MIT 6.S081, Harvard CS265.
- Enthusiast in systems programming, familiar with distributed systems, database systems, and operating systems.
- Skilled in fundamental data structures and algorithms, familiar with programming languages such as C/C++, Go, Python, Rust.
- Interested in generative AI, multimodal AI, Neural Search, and AI infrastructure, actively learning related technologies.
- Research experience in robot localization, multi-sensor fusion, and 3D vision, have a strong foundation in mathematics.

Education

Wuhan University of Technology — Wuhan, China
Master degree, Major in Mechanical Engineering — *Sep. 2018 – Jul. 2021*

Wuhan University of Technology — Wuhan, China
Bachelor degree, Major in Mechanical Engineering — *Seq. 2014 – Jul. 2018*

Projects

Highly-Available Distributed Key-Value Database based on Multi-Raft
Raft, Percolator, Go

- Consensus Layer: Implemented the Raft protocol supporting leader election, log replication, and log compaction. Implemented optimizations such as prevote, automatic step down, async apply, and flow control.
- Service Layer: Supported single-point configuration change and region split. Utilized a scheduler for automatic region scheduling to achieve load balancing. Implemented read index to improve throughput.
- Transaction Layer: Implemented MVCC based on the Percolator protocol.

Write-Optimized Key-Value Database based on LSM Tree
LSM tree, Rust

- Supported insert, update, delete, and scan operations.
- Implemented the memtable backed by a B-tree and designed the file format, Bloom filter, and sparse index for the sstable.
- Designed and implemented a unified iterator interface for efficient data retrieval.
- Supported leveled, tiered, and hybrid compaction strategies.
- Implemented crash recovery using Write-Ahead Logging mechanism.

Open Source Projects

Google Summer of Code - Jina AI — Apr. 2023 - Present
Contributor

- Enhanced the stateful executor feature based on Raft protocol by wrapping relevant Go code into Python bindings using cgo library.
- Implemented optimizations such as follower read and provided consistency mode options for service deployment in the stateful executor feature.

- Developed a highly available and high-throughput neural search service using the Jina framework and ANNLite library.

Open Source Promotion Plan - Apache Doris May. 2023 - Present

Contributor

- Added SQL parsing and execution support for functions like.
- Included information about Segment Compaction in the Load Profile.

Other Projects

bustub | *Database Storage, C++*

- Implemented buffer pool manager, B+ tree index, extendable hash index, lock manager, and deadlock prevention for the bustub database.

miniob | *Database SQL, C++*

- Added parsing and execution capabilities for SQL statements such as multi-row insert, multi-col update, cross join, inner join, order by, group by, having, sub-select, and aggregation to the miniob database.

xv6 | *Operating System Kernel, C*

- Added features like mmap, kernel thread, copy-on-write, lazy allocation, and compensated round-robin scheduling to the xv6 operating system.

balancebeam | *Proxy Server, Rust*

- A proxy server with features such as health checks, thread pool, connection pool, cache management, traffic control, and load balancing.

deet | *Debugger, Rust*

- A debugger with common debugging commands like breakpoints, next, continue, and backtrace.

Technical Skills

Programming Languages: C/C++, Go, Python, Rust
Frameworks or Libraries: Jina, DocArray, Apache Doris, Apache Spark, LevelDB, Hashicorp Raft, etcd Raft
Developer Tools: Linux, Git, Docker, CMake, VS Code, Vim, LaTeX
Languages: Mandarin (Native), English (Fluent)