

聂师聪

电话: 13026339196 | 邮箱: niebayes@gmail.com
微信: nsc-sunflower | 个人网站: <https://github.com/niebayes>

个人简介

- 开源爱好者, 正在参加 Jina AI 的谷歌开源之夏、Apache Doris 的中科院开源之夏项目
- 热爱技术, 参加过 PingCAP tinykv 训练营和 OceanBase miniob 数据库大赛
- 对分布式存储感兴趣, 实现过 Multi-Raft、Multi-Paxos、LSM-DB
- 自驱能力强, 学习过 MIT 6.824、CMU 15-445、MIT 6.S081、Harvard CS265 等知名公开课
- 系统编程爱好者, 熟悉分布式系统、数据库系统、操作系统
- 掌握基础的数据结构和算法, 熟悉 C/C++、Go、Python、Rust 等编程语言的使用
- 对生成式 AI、多模态 AI、神经搜索、AI 基础设施很感兴趣, 正在积极学习相关技术
- 有机器人自主定位、多传感器融合、三维视觉方面的研究经验, 有良好的数学基础

教育经历

| | |
|-------------------------|-------------------|
| 武汉理工大学 - 机械工程 硕士 | 2018.09 - 2021.07 |
| 武汉理工大学 - 机械设计制造及其自动化 本科 | 2014.09 - 2018.07 |

个人项目

基于 Multi-Raft 的高可用分布式键值数据库

- 共识层: 使用 Raft 协议, 支持领导选举、日志复制、日志压缩; 实现了 prevote、automatic step down、async apply、流量控制等优化
- 服务层: 支持单点配置变更、region split; 使用 scheduler 自动调度 region 以实现负载均衡; 实现了 read index 优化
- 事务层: 实现了基于 Percolator 协议的 MVCC 并发控制

基于 LSM-Tree 的单机 key-value 数据库

- 支持单点插入、单点更新、单点删除、区间查询等功能
- 实现了基于 B 树的 memtable, 设计并实现了 sstable 的文件格式、布隆过滤器、稀疏索引
- 设计并实现了统一的迭代器读取接口
- 支持 leveled、tiered、hybrid compaction 策略
- 支持基于 WAL 机制的故障恢复

其它项目

- bustub: 为 bustub 数据库实现缓冲池管理器、B+ 树索引、可扩展哈希索引、锁管理器、死锁预防
- miniob: 为 miniob 数据库添加对 multi-row insert、multi-col update、cross join、inner join、order by、group by、having、sub-select、aggregation 等 SQL 语句的解析和执行功能
- xv6: 为 xv6 操作系统添加 mmap、kernel thread、copy-on-write、lazy allocation、compensated round-robin scheduling 等功能
- balancebeam: 一个具备健康检查、线程池、连接池、缓存管理、流量控制、负载均衡等功能的代理服务器
- deet: 一个具备 breakpoint、next、continue、backtrace 等常用调试命令的 debugger

开源经历

| | |
|---|--------------|
| 谷歌开源之夏 - Jina AI | 2023.04 - 至今 |
| <ul style="list-style-type: none">• 完善基于 Raft 的 stateful executor feature, 使用 cgo 库将相关的 Go 代码包装为 Python binding• 为 stateful executor feature 添加 follower read 等优化, 为服务部署提供 consistency mode 选项• 实现基于 Jina 框架和 ANNLite 库的高可用、高吞吐神经搜索服务 | |
| 中科院开源之夏 - Apache Doris | 2023.05 - 至今 |
| <ul style="list-style-type: none">• 支持对 array_contains_all 等函数的 SQL 解析和执行• 为 Load Profile 添加关于 Segment Compaction 的信息 | |

技能

- 编程语言: C/C++、Go、Python、Rust
- 框架或库: Jina、DocArray、Apache Doris、Apache Spark、LevelDB、Hashicorp Raft、etcd Raft
- 工具: Linux、Git、Docker、CMake、VS Code、Vim、LaTeX
- 语言: 普通话(母语)、英语(流畅)