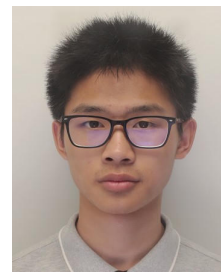


张宇

153-5928-3175 · zyuyz@stu.xmu.edu.cn · github.com/codezyu



我硕士就读于厦门大学先进存储实验室 (ASTL), 专注于存储优化和存储加速技术, 在读期间参与多个实验室项目, 对数据业务和存储系统有深入研究, 围绕存储系统软硬件结合性能优化展开, 学习能力强, 求知欲强烈, 有良好的团队协作能力。

教育背景

厦门大学 985

软件工程 硕士 信息学院

2023 年 09 月 - 2026 年 06 月

研究方向: 存储系统

厦门大学 985

软件工程 本科 卓越工程师班 信息学院

2019 年 09 月 - 2023 年 06 月

GPA: 3.5 / 4.0

科研经历

DPU 存储加速系统

2023 年 12 月 - 2025 年

问题: 以 RocksDB 为典型的持久存储算力依赖集中, 硬件配置要求高, 带来了额外数据搬移开销和大量 CPU 资源消耗。

- 设计: (1) 高性能服务开发: 基于 DPU 硬件特性重构 RocksDB 压缩流程, 将 DPU 作为 RocksDB 的模拟存储介质实现压缩任务硬件加速, 并借助 DPU 的网存一体能力加速主副本的数据同步, 缩短数据路径。(2) 分布式节点服务管理: 通过 DPU 集群化锁管理服务, 提供全局的重删服务。
- 贡献: 架构设计和代码实现, 论文编写投稿中 (第一作者)
- 优化结果: 基于 DPU 的存储系统吞吐量提升 30%+, 顺序写入情况下, 尾延迟降低了 1.5 倍, DPU 集群化锁管理服务 P99 延迟优化 1.64 倍。

基于 RDMA 的分离式存储系统

2024 年 05 月 - 2025 年

问题: 传统存储系统的计算和存储耦合, 导致扩展性受限。面向未来场景存储计算分离场景的 B+索引。

- 设计: (1) 核心算法研发: 建立计算, 存储分离的计算池和内存池, 实现基于单边 RDMA 的 B+树索引算法, 通过数据分片策略降低网络请求 25%。(2) 开发 QoS 调度算法: 基于 RDMA 优先级队列实现多级请求调度, 提升 KV 请求的响应速度, 减少在 RNIC 上数据的排队的时间。

实习经历

算法引擎部 (ML 存储引擎架构组)

快手

在组期间主要参与到快手推荐以及大模型存储服务研发中, 主要参与的工作有

- 参与持久 KV 存储引擎的维护工作, 面向快手推荐系统多种数据源如推荐, 商业化等等, 主要负责优化和维护工作, 比如主从数据同步等等。
- 参与维护开发 RDMA 数据存取框架, 以加速精排推理期间对 embedding 数据的拉取, 相比 TCP, 吞吐提升 93%, 整体延迟降低 50ms
- 参与内存 KV 优化升级到高密机器, 面向 embedding 数据存取, 配合 RDMA 通信框架, 吞吐提升 2 倍, 整体 CPU 节省至原来的一半

项目经历和技能

清华开源操作系统 rCore-OS 项目, 基于 QEMU 模拟实现的 RISC-V 架构上基于 Rust 实现的简易操作系统, 实现了内存管理, 文件系统, 进程管理, IO 设备管理等功能

编程语言: C/C++, Python, Rust, Java 等, 熟悉 Linux 以及分布式系统和相关框架, 英语读写能力良好 (英语六级: 520)