

詹晔康

zhanyekang@foxmail.com | (+86) 18271842976 | github.com/YekangZhan |

教育背景

华中科技大学

武汉光电国家研究中心 信息存储系统教育部重点实验室 计算机系统结构 博士
· **研究方向:** 存储系统, 文件系统, 存储 I/O 栈, 动态 I/O 调度, 新型存储器
· **已获荣誉:** 华为奖学金 (2%), 校三好研究生 (5%), 科技创新奖学金 (5%)

湖北 武汉

09/2021 – 至今

华中农业大学

信息学院 计算机科学与技术 本科
· **已获荣誉:** 保研排名 2/132 (1.5%), 校三好学生 (5%), 优秀共青团员 (5%)

湖北 武汉

09/2017 – 06/2021

实习经历

腾讯科技(深圳)有限公司

广东 深圳

技术工程事业群(TEG) 存储引擎组 (Research)

04/2023 – 03/2024

- 研发基于新型非易失内存(NVM)的存储系统。
- 设计并开发了一种高性能 NVM 文件系统(约 13000 行 C++代码)和一种高效的 Userspace 和 Kernelspace 的元数据异步交互机制,比起 SOTA **实现了至多 1.96X 的写入性能**。基于该工作进一步提出的文件系统架构论文已被发表到 FAST25。

技术工程事业群(TEG) 文件存储组

04/2025 – 06/2025

- 研发腾讯云第三代分布式文件系统 HiFS。
- 针对 HiFS 元数据路径中 RocksDB WAL 文件更新给底层文件系统 BlueFS 带来的**双写问题**, 设计了一种对 RocksDB 透明的新 WAL 写入机制: 将 WAL 文件的元数据更新包装为一种特殊格式的 WAL Record, 追加在 WAL 文件数据更新之后, 从而消除双写问题, **减少了至多 49% 的写入延迟**。该方案无需修改 RocksDB 源码。
- 在基于同步架构的 HiFS 元数据路径中, RocksDB WAL 文件严格串行的写缓存和写盘操作是制约系统并发性能的关键因素。针对该问题, 在上述新 WAL 写入机制的基础上进一步设计了一种**异步无锁化写入架构**, 实现并行的前台写缓存与后台写盘操作。该方案在上述新写入机制的基础上, **进一步提升了至多 16.02X 的写入性能**。

主要成果

Rethinking the Request-to-IO Transformation Process of File Systems for Full Utilization of High-Bandwidth SSDs
(高性能文件系统与异构存储, **华中科技大学首篇文件系统方向顶会论文**) **FAST25, CCF-A, 一作**

Rearchitecting Buffered I/O in the Era of High-Bandwidth SSDs
(内核存储 I/O 栈优化与高性能存储) **FAST26, CCF-A, 一作**

AIS: An Active Idleness I/O Scheduler to Reduce Buffer-Exhausted Degradation of Solid-State Drives
(I/O 调度与固态硬盘性能优化) **TACO24, CCF-A, 一作**

RomeFS: A CXL-SSD Aware File System Exploiting Synergy of Memory-Block Dual Paths
(CXL 互连技术与新型文件系统, **首篇基于 CXL-SSD 的文件系统论文**) **SoCC24, CCF-B, 一作**

HBtree: A Heterogeneous B+ tree with Multi-granularity for Hybrid NVM-SSD Storage
(高效索引结构与异构存储) **IEEE NAS22, 一作**

SchInFS: A File System Integrating Functions of the Block I/O Scheduler for ZNS SSDs
(ZNS-SSD 与新型文件系统) **ICCD24, CCF-B**

HeteroGNN: A Heterogeneous Stage Division Based GNN Training Framework to Maximize CPU-GPU Parallelism
(GNN 训练加速与细粒度任务调度) **ICME25, CCF-B**

一种基于设备感知的请求划分的数据存储系统 **发明专利, 案卷号 AJ2510226**

科研项目

OrchFS: 利用异构的低延迟设备最大化高带宽 SSD 性能的高速文件系统 (国家重点研发计划 No.2022YFB2804302)

- **存储栈 I/O 性能瓶颈识别**: 结合理论与实验分析发现, 现有存储系统难以充分利用高带宽 SSD 性能。其根本原因在于: 1) 昂贵的 I/O 对齐, 2) 高页缓存软件开销, 3) 不充足的 I/O 并发性。
- **最大化存储性能**: 设计了对齐优先且最小化碎片的运行时请求划分机制和文件系统并行 I/O 处理机制去最大化设备性能。
- **实验效果**: 比起 SOTA 和相关的产品级文件系统, 该文件系统实现了至多 29.76 倍的写性能提升与 6.79 倍的读性能提升。
- **已有成果**: 文件系统代码已开源: <https://github.com/YekangZhan/OrchFS>, 约 17000 行。论文已发表在 FAST25。

WSBuffer: 利用高带宽存储性能强化内核文件系统 buffered I/O 访问模式 (国家重点研发计划 No.2024YFB4505105)

- **最小化缓存的高效数据访问机制**: 引入了一种新的内存页结构 scrap-page。在写路径上使用 scrap-page 去自适应地部分缓冲用户写入, 并主动利用 SSDs 的高带宽来直接处理前台的用户写入, 从而让 buffered I/O 能够直接受益于 SSD 带宽。与此同时将 page cache 从繁琐的写处理中解放出来, 使其专注于所擅长的读处理, 从而让 buffered I/O 的读取更加高效。
- **机会主义的两阶段 dirty-page flushing**: 动态地选择负载较低的 SSD 执行脏页下刷, 在第一阶段执行大读去填充未满足的 scrap-page, 在第二阶段执行大写去将已满的 scrap-page 写回 SSDs。两阶段的操作能够以页粒度并行执行。
- **并发的页管理机制**: 针对传统 page cache 中自由页插入, 干净页删除, 和脏页状态维护等操作导致的高锁竞争开销问题, 引入了一种并发页管理机制去分别管理 page cache 和 WSBuffer, 显著减少锁竞争。
- **实验效果**: 比起 SOTA 和相关的产品级文件系统, 该系统实现了至多 3.91 倍的吞吐量提升和 82.80 倍的延迟改善, 并在图计算和 HPC 应用程序中节省了至多 99.66% 的内存使用。
- **已有成果**: 在 Linux kernel 6.8 的 XFS 中实现了 WSBuffer 架构, 核心代码量超 4000 行。论文已发表在 FAST26。

RomeFS: 基于 CXL-SSD 独特双路访问特性的新型文件系统架构 (国家自然科学基金 No.62172175)

- **CXL-SSD 双数据路径访问特性分析**: CXL.mem 路径支持字节寻址, 提供低延迟的数据访问, 但大型数据访问会导致其性能降级。CXL.io 路径为块数据路径, 性能相对稳定但访问延迟高。
- **协同的双路数据访问文件系统架构**: 在文件系统中透明地主动利用 CXL.mem 路径去处理所有的元数据操作与事务性操作, 使用 CXL.mem 和 CXL.io 路径去协同处理文件数据操作。
- **实验效果**: 比起 SOTA 和相关的产品级文件系统, 该文件系统在广泛的负载评估下实现了至多 14.24 倍的性能提升。
- **已有成果**: 实现了约 14000 行代码的文件系统原型和约 3000 行代码的 CXL-SSD 模拟器。论文已发表在 SoCC24。

AIS: 一种最大化固态硬盘内部缓存恢复的主动空闲调度器 (国家自然科学基金 No.61821003)

- **固态硬盘缓存耗尽性能降级模型**: 分析了固态硬盘因内部缓存耗尽而导致的性能降级问题, 并建立了缓存恢复模型。经过测试验证, 该模型可以很好地用于三星, 西数和铠侠等主流 SSD 厂商的多款 SSD。
- **主动空闲调度器**: 在该模型的基础上设计了一个 SSD 性能降级预测器, 并实现了一个块层调度器去通过预测器信息来动态调整 SSD 负载, 从而缓解 SSD 性能降级问题。
- **实验效果**: 在真实世界的云存储 trace 下, 该调度器分别改善了 SSD 的平均 I/O 延迟和尾延迟至多 29.3% 和 78.7%。比起相关的 SOTA 调度器, 该调度器在 LevelDB 和图计算应用上提升了至多 51.9% 的应用程序性能。
- **已有成果**: 实现了主动空闲调度器和一个自动化的 SSD 模型特征参数提取工具。论文已发表在 TACO24。

竞赛奖项

第五届中国大学生程序设计竞赛(CCPC)湘潭赛站 **银牌**

第五届中国大学生程序设计竞赛(CCPC)秦皇岛站 **铜牌**

第四十三届国际大学生程序设计竞赛(ICPC)亚洲区域赛青岛赛区 **铜牌**

第四十五届国际大学生程序设计竞赛(ICPC)亚洲区域赛昆明赛区 **铜牌**

第十届蓝桥杯全国软件和信息技术专业人才大赛 C/C++A 组 全国总决赛 **二等奖 (1.5%)**

第十届蓝桥杯全国软件和信息技术专业人才大赛湖北赛区 C/C++ A 组 **一等奖 (10%)**

第八届 ASC 世界大学生超级计算机竞赛 **二等奖**

CCF 计算机软件能力认证(CSP) **364 分 (全国前 1.07%)**