📞 🗞 (+86)132-8866-2339 | 💌 tomatofaq@gmail.com | 🞧 github.com/TOMATOFGY | 🏫 TOMATOFGY's Blog

北京大学・软件与微电子学院・网络安全・在读硕士研究生・(GPA 3.53/4.00, rank 4/20) 2021/09 - 2024/7 (预计) 新加坡国立大学·Computing School·暑期实习 2019/06 - 2019/08

北京邮电大学・计算机学院・计算机科学与技术・学士・(GPA 91/100, rank 12/404, top 3.0%) 2017/06 - 2021/06

实习经历

易方达 · 中后台条线 · DBA

2023/07 - 至今

管理业务所涉及的各类数据库、优化数据库管理系统性能、保证系统高可用。

字节跳动 · Data · 数据平台 · 分析型数据库 · C++ 研发实习生

2023/03 - 2023/07

参与字节旗下分析型数据库 ByConity 存储层的优化工作。

· 参与 ByConity 冷存温存、监控链路功能的设计与实现。

商汤科技 · 存储系统与技术部 · 缓存与数据加速组 · 存储系统开发实习生

2022/06 - 2022/11

参与商汤内部缓存服务的开发。其基于一个支持多访问协议的,基于 nvme 的,键值分离的分布式键值数据库。

- 为该缓存服务设计了 POSIX 接口,提高了服务易用性。用户通过 FUSE 实现的接口,以访问常见文件系统的方式 访问访存中的数据,从而实现无需修改业务代码无感利用缓存。
- 为用户态文件系统设计了系统调用劫持机制,节省访问开销。实现了绕过 libfuse 直接访问缓存服务的功能。
- 优化了云使用场景下的读写效率。为 S3 等存储服务建立了基于该缓存系统的 shadow filesystem。访问存储服务 的行为将被重定向自缓存中获取数据,从而实现了读写速率上 100x 的提升。

微软亚洲研究院 · 创新工程组 (IEG) · 软件开发实习生 🙈

2021/08 - 2022/06

负责一个用户量为十万级的应用的后端与 iOS 端的开发。

• 深度参与项目重构的设计与实现。提高了产品迭代效率,解决了各类技术问题 30+。

获奖情况

蚂蚁集团 · 2022 OceanBase 数据库大赛 ② 季军 (决赛 rank 4/50,初赛 rank 11/1180)

2022/10 - 2023/01

初赛内容为拓展 miniob 的功能,使其成为功能齐备的数据库。

• 负责数据库页缓冲区、索引等存储层功能的开发。

决赛内容为在给定计算资源限制下,为 OceanBase 优化旁路导入功能。

- 负责旁路导入中文件解析、压缩算法、归并排序算法、csv2sstable 算法等模块的设计与实现。
- 充分使用调优技术捕获大小优化点加以优化。
- 最终优化性能相较于 OceanBase 原 batch insert 方案实现了 10x 的提升。 知乎 知乎

项目经历

Sourcetrail Golang Indexer 开源贡献 🞧

2021/01 - 2021/06

为源码阅读软件 Sourcetrail 提供了对 Golang 语言的支持。

 利用程序静态分析技术,分析并生成 Golang 项目的函数间调用图、函数内控制流图、类继承关系及模块间依赖 关系,并将上述数据绑定到 Sourcetrail 中,使用户能以可视化的形式与调用图等数据结构进行交互。

自制操作系统内核 课程设计

2020/03 - 2020/06

负责实现一个拥有中断机制、进程调度、文件系统等常见功能的小型内核。

考察了uCore@thu, rCore@thu, xv6@MIT, BlogOS@Phil-opp等常见的开源操作系统.

• 负责特权级转换功能的实现;负责中断机制的实现。包含中断屏蔽、二级中页断等基本功能;负责内核内存管 理功能的实现。包括 sv39 页表机制的实现、基于 BuddySystem 的内存管理;负责文件系统的实现。仿照 Linux 的 虚拟文件系统架构,基于 Ext2 实现了一个简易的文件系统。

技能

- 语言: C/C++, Shell, Rust, Golang, HTML/CSS/JavaScript, Python, SQL, Swift, Obj-C, VHDL
- 工具&框架&产品: CMake; Git; GDB, Perf, Flamegraph; Docker; MySQL, Clickhouse, Redis, Memcached, LevelDB
- 外语: TOEFL: 102; CET-6: 559