

赵军旺

✉ zhjwpku@gmail.com · ☎ (+86) 132-6996-5827 · 🌐 zhjwpku

🎓 教育背景

北京大学	2013 年 9 月 – 2016 年 7 月
硕士 系统结构	
南京理工大学	2009 年 9 月 – 2013 年 7 月
学士 计算机科学与技术	

🏢 工作经历

星环科技	基础架构产品部
资深数据库内核研发	2023 年 7 月至今
阿里巴巴	阿里云数据库 OLAP 产品部
技术专家	2021 年 5 月 – 2023 年 6 月
华为	IT 产品线 FusionStorage 项目群
高级工程师	2018 年 4 月 – 2021 年 5 月

📁 项目经历

Spacture & Hippo	2023 年 7 月 – 至今
负责 Spacture 时空库及 Hippo 向量库内核研发。	
<ul style="list-style-type: none">Spacture 选型 HA、备份工具及插件，支持单机版和 MPP 分布式数据库的统一部署Hippo 支持 DiskANN 索引并优化性能，性能提升 30%Hippo 支持 JSON 类型开发 HippoBench，支持多种开源数据集，统一了 POC 调优、性能优化等场景使用的工具，显著提升了团队工作效率	
AnalyticDB PostgreSQL 高可用及同城容灾	2022 年 6 月 – 2023 年 6 月
改进 GPDB 的 FTS 逻辑，适配多副本架构，确保高可用部署。	
<ul style="list-style-type: none">修改 FTS 逻辑适配多副本架构，其中一个备升主后其它备要去和新主建立 replication设计基于实例节点独立 agent 的 HA 解决方案（类似 Patroni）	
AnalyticDB PostgreSQL Iceberg table	2022 年 1 月 – 2022 年 6 月
Apache Iceberg 是一个开源大数据表格式，旨在为数据湖提供一个高效的元数据管理框架。ADBPG 通过 Iceberg 通用表格式融入大数据生态。	
<ul style="list-style-type: none">实现 iceberg-c sdk 及 iceberg-fdw查询 OSS 上的 Iceberg Table，打破数据孤岛，实现数据湖与数据仓库的无缝连接未来规划: 像 Snowflake 一样将 Iceberg 视为本地表进行查询写入（适配 table access method）	
AnalyticDB 存储引擎支持 Z-order 排序	2021 年 5 月 – 2022 年 1 月
AnalyticDB 是阿里巴巴自主研发的海量数据实时高并发在线分析云计算服务，可以在毫秒级针对千亿级数据进行即时的多维分析透视和业务探索。	
<ul style="list-style-type: none">设计开发 Z-order 索引，build 过程中根据 z-value 对数据进行排序并写入 PartitionEngine	

- 根据统计信息将不同列在多维查询分析时，提高 `data skipping` 的概率，进而过滤掉不必要的数据块来提升查询速度
- 实现效果: 查询条件不带 `leading column` 时，性能优于 `cluster index`

TaurusDB 内核研发

2020 年 5 月 – 2021 年 5 月

TaurusDB for MySQL 是一款遵循 *LOG IS THE DATABASE* 的云原生数据库解决方案

- 使用 `mmap` 优化 `binlog` 性能，提高写入性能，`Use mmap for binlog`
- 参与 *TaurusDB for MySQL* 的 `Storage Abstraction Layer/Log Store` 模块开发，日志通过 `PLOG` 写入 `FusionStorage`

FusionStorage OBS Index Layer

2018 年 4 月 – 2020 年 5 月

OBS Index Layer 用于存储对象存储的元数据，基于 *MongoDB*，存储引擎使用 *mongo-rocks* 将数据持久化到 *FusionStorage*。

- 适配 *MongoDB transport layer* 到 `xnet` 库 (`RoCE`)，提高吞吐量并减少延时 (`bypass kernel`)
- 将 `mongo-rocks` 存储引擎对接到 `FusionStorage` 的 `PLOG` 接口
- 针对热点分析，通过合理的批处理、索引优化和参数调整等措施，提高写入、查询和 `Compaction` 性能

🔗 开源贡献

- `Patroni`
 - Abstract MPP handler and decouple it from DCS
- `pgvectorBench`
 - A lightweight, fast, flexible, and easy-to-use benchmarking tool dedicated to `pgvector`
- `pg_cron`
 - add possibility to schedule jobs on the last day of month
 - PRs contributed to `pg_cron`

🔧 技能栈

- 编程语言: `C/C++` > `Rust` > `Python` > `Java`
- 并发编程: 熟悉事件驱动模型及网络编程
- 常用工具: `GDB/Perf/Valgrind/gperftools`

⚙️ 其他

- 博客: <https://zhjwpku.com>
- 公众号: PostgreSQL 开源生态
- 语言: CET-6 VET/VEWT
- 论文阅读笔记: paper notes
- GSoC: 两次 Google Summer of Code (X.Org, Ceph) 项目经历