

DTCC

2017第八届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2017

面向未来的数据库技术的思考

张瑞 @ 阿里巴巴

数据驱动·价值发现 | 北京·国际会议中心

SequeMedia
感知传媒

IT168

IT-PUB

ChinaUnix

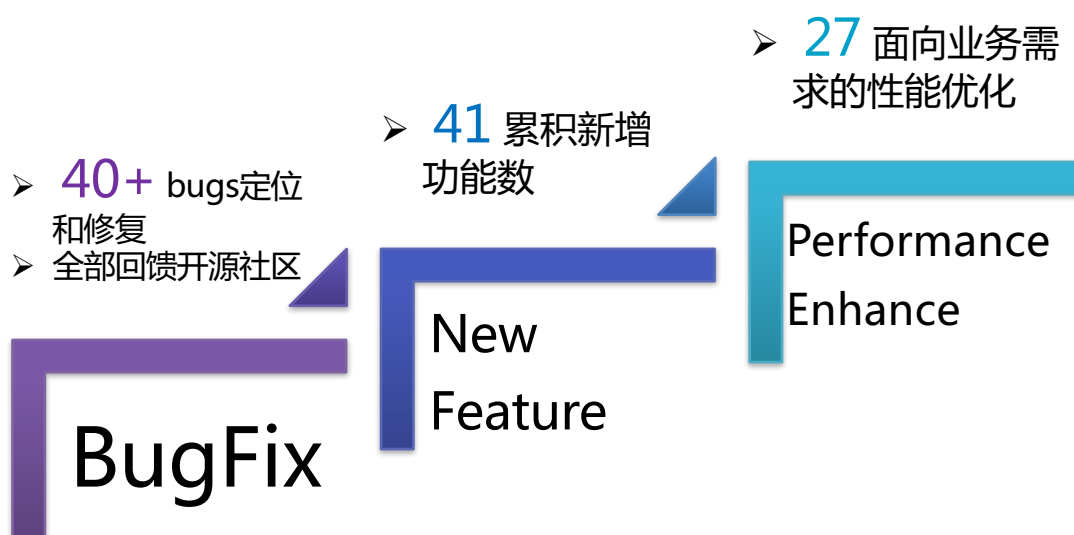
关于我

- 2005年加入阿里巴巴数据库技术团队，推动并见证了阿里数据库技术发展的变革历程
- 我们正在思考阿里下一代数据库技术的发展方向
- 分享我们的成果和思考，抛砖引玉

大纲

- 数据库内核技术创新
- 数据库实现弹性调度
- 从自动化走向智能化
- 走过的路 未来的路

AliSQL—过去五年





阿里场景下数据库技术的思考

□ 阿里的场景分析

✓ 交易、支付型应用

- 特点一：持续可用、数据强一致 → 高可用 强一致
- 特点二：数据量大、重要程度高 → 存储成本
- 特点三：数据有明显的生命周期特性，冷热数据特点鲜明 → 数据生命周期管理
- 特点四：交易、库存，支付等业务，操作逻辑简单，高性能 → 极致性能

✓ 有一个超级热点：双11

- 双11需要大量的机器资源，需要极致弹性伸缩的能力

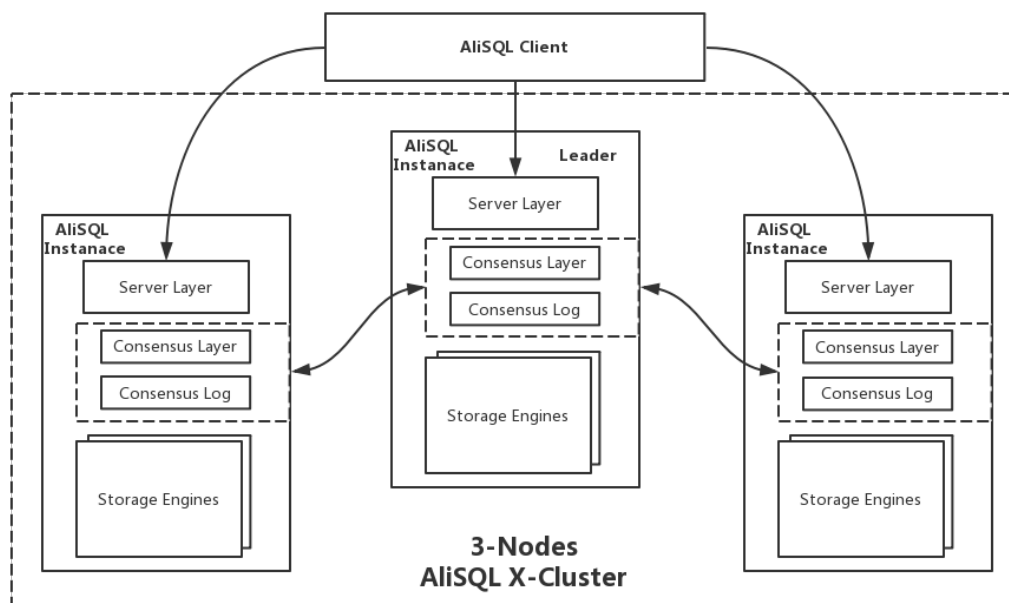
✓ DBA成为业务发展瓶颈

- 阿里场景有海量的性能数据，如何做到数据库自诊断自优化



数据库内核技术创新

AliSQL X-Cluster

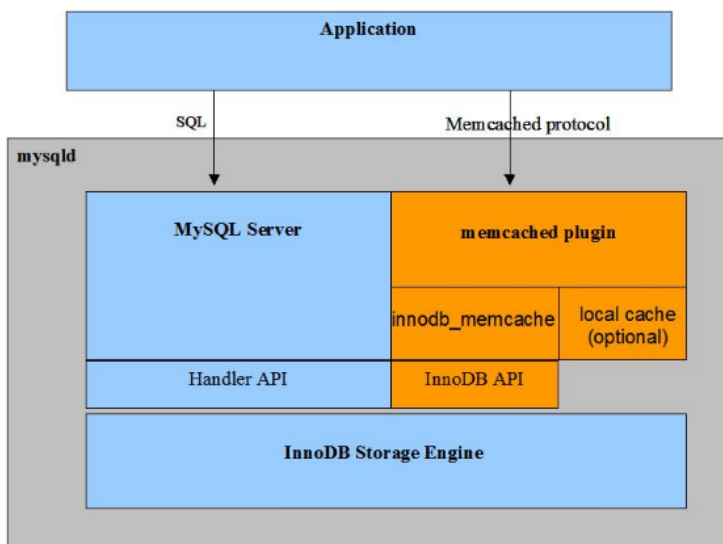




数据库和Paxos协议的关联

- 服务的本质，就是Server状态机中的一个状态变更，通过Paxos来保证状态变更全局一致
- Paxos基本功能
 - ✓ 系统各节点状态强一致
 - ✓ 系统少数派宕机，不影响可用性（持续可用）
 - ✓ 系统自封闭，零外部系统依赖
- 作为大型系统的Build Block
 - ✓ Spanner、MegaStore、CockroachDB.....
- 作为基础系统对外提供服务
 - ✓ 分布式锁服务、配置中心（元数据管理）、集群管理（Leader Election）...
 - ✓ Zookeeper、ETCD ...

AliSQL - X-KV



什么是X-KV：

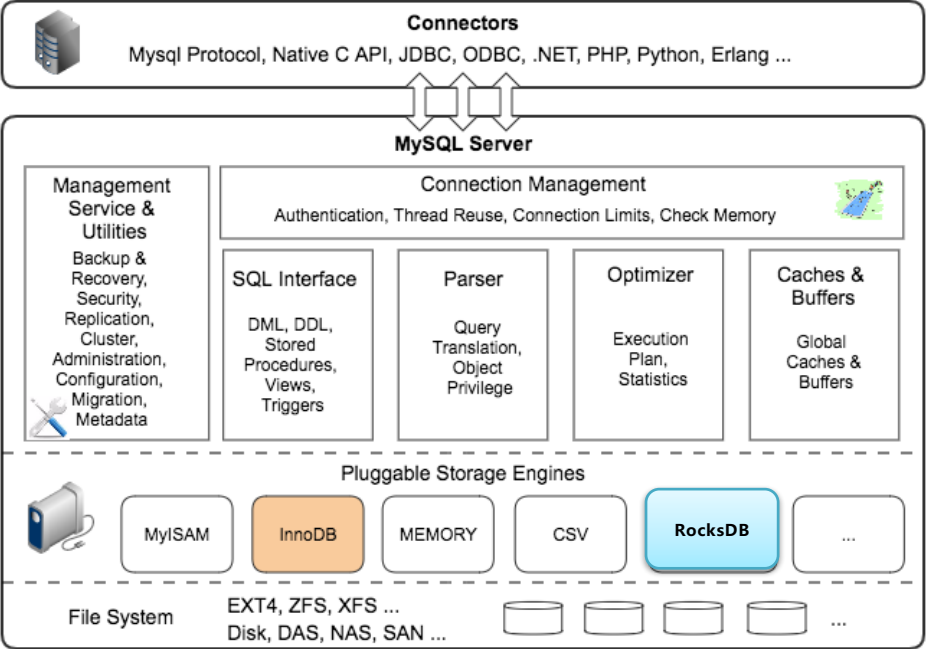
- 基于Memcached plugin
- 支持更多数据类型
- 支持非唯一索引，组合索引
- 协议层优化

X-KV优势：

- 超高读取性能
- 数据强一致（一份数据）
- 减少应用响应时间（降低成本）



AliSQL多引擎架构



LSM Tree vs B+ Tree

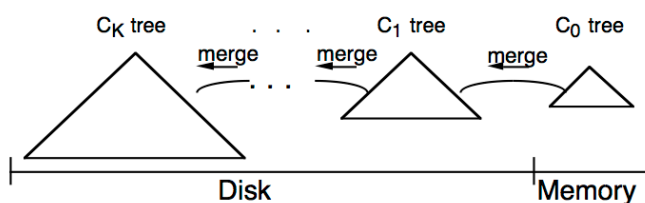
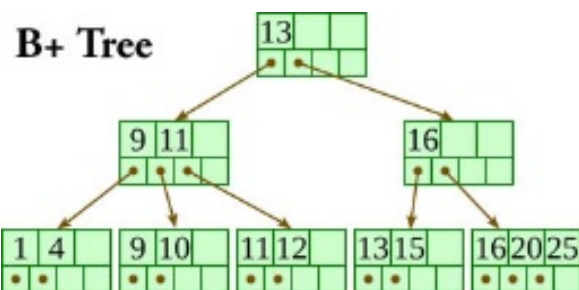


Figure 3.1. An LSM-tree of $K+1$ components



原理分析：LSM Tree vs B+ Tree

- ✓ 顺序写入 vs 随机写入
- ✓ 内存更新、定期合并 vs 实时更新

LSM Tree优点：

- ✓ 高压缩率，写入优化，SSD友好

LSM Tree缺点：

- ✓ 读取性能偏弱



数据库实现弹性调度



数据库实现弹性调度

□ 实现弹性调度的两大基础条件

- ✓ 容器化
- ✓ 计算存储分离

□ DB容器化

- ✓ 支持物理机，VM，Docker
- ✓ 性能：容器性能与物理机持平

□ 存储计算分离

- ✓ 新技术发展：25G网络，RDMA等技术让大规模存储计算分离成为可能
- ✓ 数据库优化：减少网络IO，变离散IO为顺序IO
- ✓ 存储成本：共享存储池，提升存储利用率
- ✓ 计算成本：一主一备 -> 多主一备



从自动化走向智能化

CloudDBA进展与方向

□ 大量的数据

- ✓ 采集每一条运行的SQL信息
- ✓ 秒级监控指标采集
- ✓ 应用端错误日志
- ✓ 响应时间采集

□ 存储和计算能力

- ✓ 海量数据存储
- ✓ 计算平台

□ 我们的目标

- ✓ 在未来两三年，CloudDBA可以实现DBA的大部分诊断和优化工作

□ 未来方向

- ✓ 机器学习



走过的路 未来的路

- 工欲善其事 必先利其器
 - ✓ 利用多种数据存储技术解决问题
- 建设两个平台
 - ✓ 数据库支撑平台
 - ✓ 数据库服务平台
- 不要相信那些已经过期的神话
 - ✓ 软硬件技术快速发展让一切变为可能
- 自动化系统的悖论
 - ✓ 随着自动化程度的提升，伴随而来的人的能力降低

结束语

谢谢大家



We must know , we will know
我们必须知道 , 我们终将知道
- 大卫·希尔伯特