ArchSummit
全球架构师峰会 2017

主办方 Geekbang. 极客邦科技 InfoQ

MySQL数据库架构的演化观察

余锋（褚霸） V1.4
阿里巴巴研究员

SPEAKER
INTRODUCE

余锋（褚霸）

阿里巴巴研究员

2013年起负责阿里云计算公司的数据库相关业务，他的研究领域包括云计算、边缘计算和软硬件协同设计。

2010年加入阿里前，在中国互联网基础软件领域工作了12年，从事过互联网、游戏、通讯业务等行业，作为早期员工服务过网易、迅雷等公司。

作为资深工程师在中间件、数据库、存储系统和硬件等技术领域都有突出的造诣，有超过20年的系统软件编码功底和丰富的大规模复杂集群系统的构建和演进经验。

TABLE OF

CONTENTS 大纲

- 架构演进的动力
- 架构演进的路径
- 技术社区的变化

用户视角需求

- 成本合理，只为使用付钱，维护费用少
- 使用方便，尊重当下投资，满足业务未来发展需求
- 业务适应面广， 最好ONE SIZE FITS ALL
- 平台开放性，不被lock-in

架构为业务服务

- 面向如政府、银行、IoT、高并发互联网场景一揽子解决方案
- QPS/$ 可靠性、成本和性能的平衡
- 稳定性第一、不惹事（业务连续）
- 数据闭环和生态，尊重用户熟悉的姿势，提供选择的能力
- 用户行为不可预测带来挑战，系统自证清白能力

业务关键词

易用 成本 安全 生态 性能

可靠

交付 兼容 隔离 合规

可用

平滑

架构演进的推动力

- 云化普及、基础设施地位
- 大数据、IoT和AI带来的机会
- 硬件迭代：CPU（128核）| 网络(10us级别）| IO(10us级别) | 超大内存（T级别）| NvRAM
- 基础软件技术突破

架构演化路径

- 数据结构：算法围绕数据转
- 业务架构：计算围绕着数据库
- ###MySQL在DBRanking排名第二，在业务架构举足轻重，通用数据库
- 层次化，每层可插拔（提供选择，同样界面加快演化）
  - 数据分层和不同的SLA
  - 不同引擎，提供不同性能、成本述求
    - 基础数据结构： B+ tree、LSM tree、Fractal tree(分形树)、Succinct等 | NVRAM数据结构
    - 定位：通用 OR 专用引擎 | 聚焦成本 OR 性能

架构演化路径-续

- ###数据库（资源消耗和智力密集性基础软件） = SQL + 存储 + 计算
- 存储层面
  - 主路径 -> 共享存储（实时），安全和容灾能力
  - 互通 -> 文件存储（类似AWS Athena)
  - 异构数据冗余和可靠性保证
- 计算层面
  - 类似Oracle数据库能力
  - 混合计算（OLTP和OLAP、搜索、Cache）
  - 实时化能力，Lambda数据流动和事件触发（存储过程）

架构演化路径-续

- SQL层面：
  - 屏蔽后端变化（软硬件升级、灾难、平滑扩缩容）
  - SQL路由负载均衡/读写分离
- 运营和生态支撑
  - 可视化闭环系统，提高研发和运营效率
  - 异构数据互通（如PG、MongoDB互通）| 异构系统互通（如Search互通）
  - 数据仓库/BI支撑等生态能力
- 服务
  - 专家服务和体检 <= 事前 | 诊断系统 <= 事中 | 救援服务 <= 事后

架构突破

- 计算存储分离
  - 计算 <---网络---> 存储
- 好处：减少维度
- 计算维度 => CPU + 内存
- 存储维度 => IO + 网络
- 网络是最大瓶颈，解决手段：
  - 提高网络速度和带宽<-硬件红利
  - 减少网络交互次数和加大信息密度<-软件红利
  - 多AZ，加大并行度和容灾<-IDC红利

架构突破-续

- 单机架构逐层拆解：
- | SQL |
- | 事务 |
- | Cache |
- | 存储 |
- 关键词offload
  - 数据库核心逻辑下沉到存储：利用集群分布和闲时资源细水长流算
  - 再下沉到硬件层：FPGA、ASIC、盘载计算资源

架构突破-续

- 解决数据副本、数据变形、Schema演化即刻计算的代价
- 解决高并发、高扩展(毛刺和突发）
- 解决隔离问题，从存储颗粒级别、设备、存储、引擎、事务、SQL、虚拟化整链条互不干扰
- 解决扩展问题:
  - 计算能力： 一主多读 -> 读写分离 -> Multi Master
  - 存储能力： 提高引擎存储效率（新数据结构，压缩）-> 共享存储 -> 引擎逻辑下沉的共享存储
- 解决SaaS化二次多租户问题
- 衡量指标：
  - 架构保证业务平滑
  - 秒开秒关，秒扩秒缩，数据保持温热

规模化带来的挑战

- 无需提前容量规划、秒计费、应对突发和临时需求
- 辅助运营 Advisor/CloudDBA
- 可时光回溯的问题追踪定位系统
- 全自动化下的业务顺滑和风险规避

技术社区的变化

- 商业驱动技术变化
- 数据库厂商版权收紧推迟开源
- Cloud Native设计理念、混合云趋势化

THANK YOU

如有需求，欢迎至 [讲师交流会议室] 与我们的讲师进一步交流

ArchSummit
全 球 架 构 师 峰 会 2017