

DTCC

数 / 造 / 未 / 来

第十二届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2021



2021 年 10 月 18 日 - 20 日 | 北京国际会议中心





数 / 造 / 未 / 来
第十二届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2021

百度云原生数据库实践

邱学达-百度基础架构部

DTCC
2021



北京国际会议中心

2021/10/18-10/20



ChinaUnix.net

ITPUB



01

GaiaDB简介

02

发展历程

03

难点与优势

03

未来规划



什么是GaiaDB

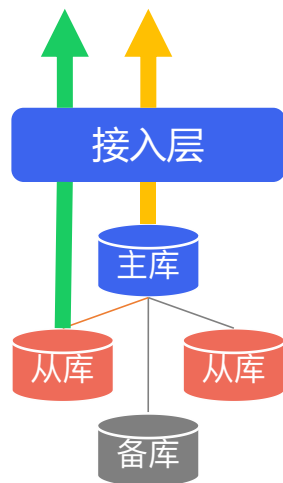
- 全兼容：完全兼容Mysql生态
- 易扩展：秒级扩缩容，资源按需分配
- 易维护：云原生一键部署，故障自恢复
- 强一致：RPO=0，低延迟
- 多地域：原生支持跨地域部署，金融级数据安全

- 支撑公司内部与云上多个业务，用户侧0改造
- 解决大型业务拆库困扰，线上单库最大规模百TB级
- 解决业务性能瓶颈，帮助业务最高性能提升10倍+
- 数据安全性已得到金融客户认可



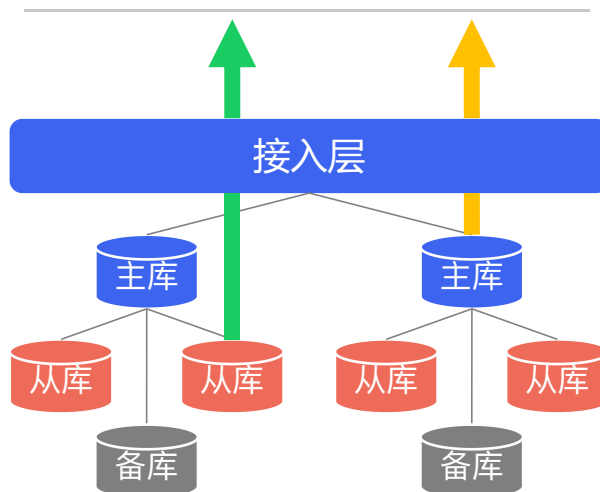
发展历程

传统主从复制



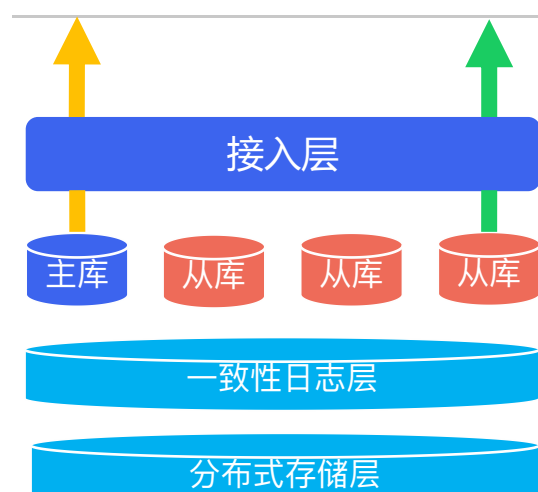
- 具备初步扩展能力

分片架构



- 扩展能力增强

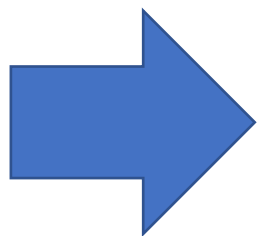
云原生架构



- 存储计算分离，高度扩展性
- 冷热数据分离，性能成本可兼顾
- 高可靠易维护，解放用户
- 插件化设计，二次开发简单安全

业务痛点

- 弹性痛点：新功能发布当天意外登上热搜，扩容需要数小时？
- 资源痛点：新功能开通了超大规格实例，结果大量资源闲置？
- 成本痛点：使用SSD成本太高，HDD查询太慢？
- 维护痛点：被拆表/迁移/扩缩容/修数据填满，没有时间做架构提升/成本优化？



数据搬运代价很高

成本与性能兼顾很难

数据安全性保证，代价很高且很难



DTCC 2021
第十二届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2021



数 / 造 / 未 / 来



IT168.com

ChinaUnix.net

ITPUB



多维度高弹性架构

- 计算资源弹性?
 - 计算节点横向+纵向弹性
- 写吞吐高?
 - 写性能只与日志层相关，好钢花在刀刃上
- 容量增长快?
 - 存储层按需分配，透明扩容，共享存储池利用率高
- 成本要求高?
 - 支持接入各类冷热介质/分层存储，介质只影响读延迟





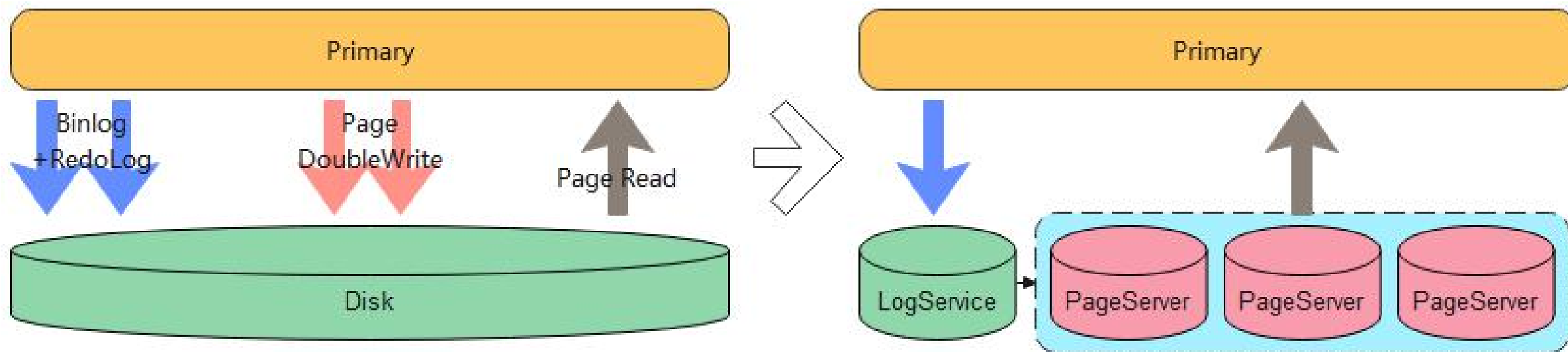
内部实践收益

- 性能最高提升10倍+
 - 突破单机磁盘吞吐瓶颈
 - 主库单核能效提升
 - 最大实例规格提升
 - 解决拆表写放大问题
 - RDMA+AEF硬件加速
- 成本最高节省90%
 - 资源利用率
 - 副本数降低





存算分离



- 主库负载减轻，脏页刷新/副本同步剥离至存储层，单核效能提升2倍以上
- 读写分离及优化
 - 写优化：IO merge、滑动窗口、高性能网络与介质
 - 读优化：并行读取、多分片/多副本负载均衡、分层存储



兼容性

- 完全兼容Mysql生态：
 - Binlog
 - DDL
 - 外键
 - 存储过程
 - 触发器
 - Etc..
- 事务
 - 原生单机事务,低延迟
 - 支持RU/RC/RR隔离级别





故障恢复

- 快速启动
 - 主库故障不影响数据一致性，无需恢复数据
 - 主库无状态，无需拷贝数据
 - 主库Term机制，无脑裂隐患
- 快速切换
 - 接入层视图感知能力



高性能日志服务

- 日志落盘区
 - 高可靠-Raft多副本
 - 高吞吐-滑动窗口
 - 低延迟-低于同步复制
 - 实时归档
- 日志分发
 - ForceFetch
 - LongPolling



自研存储

- 分布式存储，支持百TB以上容量
 - 全局强一致读，MVCC
 - 元信息全缓存，单跳网络读取
 - 分片存储，读性能高，长尾稳定
 - 可配多副本，兼顾成本与可靠性
 - 高可用，单副本故障不影响性能，自动Peer恢复



备份恢复

- 逻辑备份
 - 100%兼容现有工具
- 物理备份
 - 多分片并行备份，百TB备份时间<1h
 - 支持恢复至任意时间点
 - 手动/周期性自动备份
- Binlog/Redolog归档跨地域容灾



成本

- 资源利用率高
 - 存储池+按需分配
- 副本数少
 - 固定数据副本，不随从库增加
 - 无需备库
- 透明压缩引擎
 - 不消耗计算层CPU
- 冷存储引擎
 - 归档类/历史库



私有化能力

- 部署运维简单
 - K8S一键部署
 - 故障自恢复
- 硬件需求低
 - 支持通用X86/ARM Server，不依赖特殊硬件/网络环境
 - 内置存储层，无需外部提供分布式存储
- 监控完备
 - 内置图形化监控工具
 - 支持Prometheus等开源监控系统
 - 支持本地文件/Http等自定义采集方式



多地域

- 支持多地域、多可用区部署
 - 同地域读取，延迟低
 - 单可用区故障，不影响读写
 - 主地域故障，快速恢复服务





持续探索

- 多点写入
- 计算下推
- 列式存储
- 软硬结合





THANKS