

PieCloudDB 云原生数仓虚拟化的 演进与规划

郭罡

拓数派创始合伙人、拓数派首席技术官（CTO）

CONTENT

目录

—

- 01 PieCloudDB 简介
- 02 PieCloudDB 的诞生之路
- 03 PieCloudDB 的产品规划

π CloudDB

- 一个功能齐全的云原生分布式 OLAP 数据库
 - 云原生数据计算平台的一个基座
 - “数据计算，只为新发现”
- 数据爆炸的时代：数据流动、数据计算 产生价值



生产部

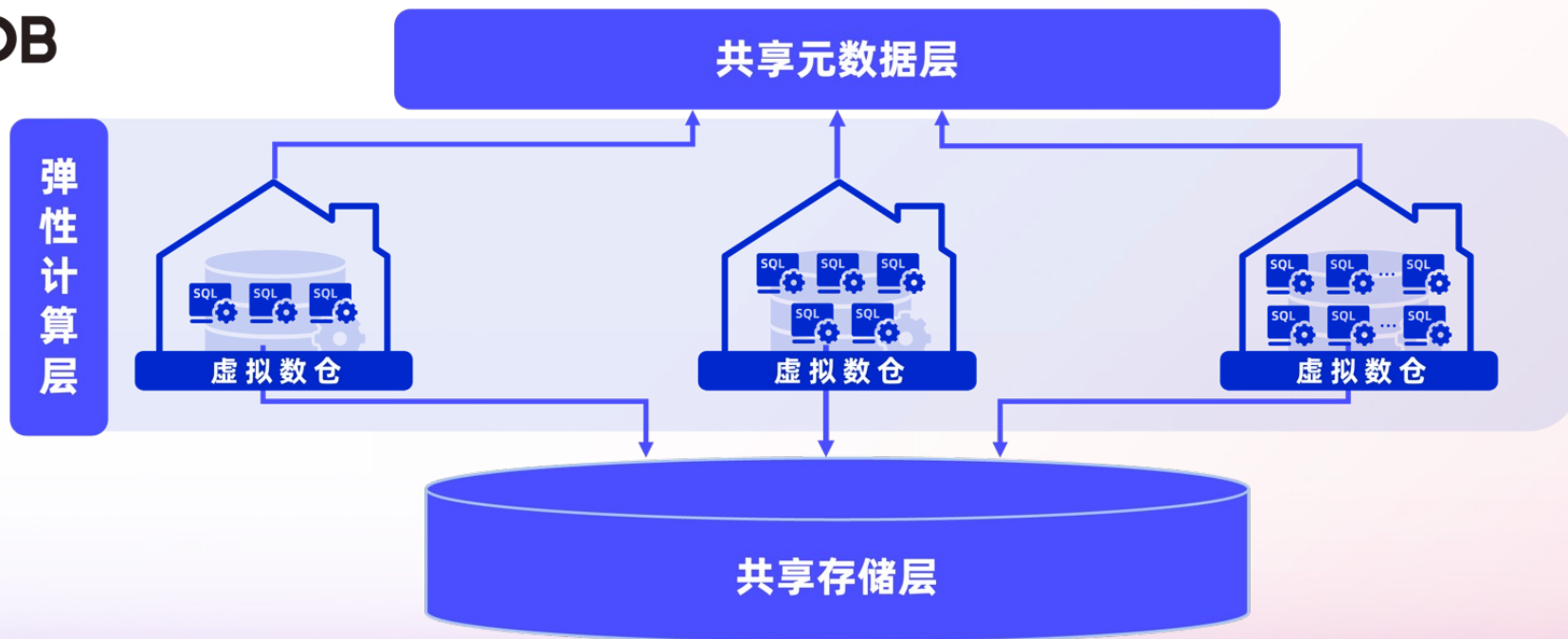


销售部



财务部

πCloudDB



云平台

对比传统分布式MPP架构

- 弹性计算：计算随时启停、存储计算按使用付费、快速调整资源.
- 打破数据孤岛：避免不同集群数据不一致和数据冗余，“一份数据”，流通产生更大价值.
- 慢？缓存，IO优化（并发/异步读写，预读等），计算优化.....
- 真正的云原生
- 接受度在上升：参考云计算的趋势

重要特点

- eMPP (弹性MPP): 伸缩(集群横向、集群纵向、多集群)
- 完备的数据库功能: SQL标准、ACID、隔离级别等
- 友好的接口: webSQL, JDBC/ODBC, Data Frame等 – Not Only SQL
- Postgres 生态: 各种开源组件、接口等
- 安全可靠: 从连接到数据存储全链路加密、高可用设计
- 数据处理: 结构化、半结构化、无结构化
- 智能云平台

发行版本

- 社区版：免费使用
- 企业版：企业级服务支持
- 云上云版：阿里云上线（可以免费注册试用 <https://app.pieclouddb.com/> ）

CONTENT

目录

—

01 PieCloudDB 简介

02 PieCloudDB 的诞生之路

03 PieCloudDB 的产品规划

- 第一步：单机 Postgres 版本上实现存算分离，同时云原生平台开启研发
- 第二步：分布式化实现，同时开启生态相关开发工作

- 设计原则
 - 同一租户多集群共享
 - 元数据和集群分离
 - 元数据尽可能精简
 - 本质上是一个轻量级的 OLTP 系统

- 设计

- 借助于 FoundationDB 的串行化事务模拟轻量级锁
- 分布式锁来避免必要的并发冲突
- KV 存储的排序性来实现 btree 索引
- 充分利用缓存系统来实现加速和减轻 FoundationDB 访问
- 没有 WAL
-

- 设计原则

- 远程(S3/HDFS/NAS) 访问友好: data skipping、Cache ...
- OLAP 友好: 行列混存、极速访问、中等尺寸...
- 计算引擎加速友好: SIMD、Cache line、SIMT ...

.....

- 简墨 (JANM)格式
 - 行列混存
 - Apache Arrow 转换更友好
 - 更灵活的实现 (特大字段等)
 - 通常的优化：辅助数据、压缩...

- 当前 COW 方式模式：MVCC 实现更清晰
- 开源格式(如Parquet等) 访问：部分格式已支持
- 分布式文件系统中立：抽象访问
- 湖？仓？都是“Shared Storage”，会统一治理
- 加速访问
 - 数据缓存
 - Data Skipping
 - 预聚集
 -

- Coordinator才有事务 - 不需要2PC
- Coordinator才访问元数据
- 执行节点Catalog来源：本地（固定不变的）、Dispatch
- 足够灵活：将来的计算引擎的基底

- 原则

- 性能
- 稳定性
- 功能增强和补全
- 生态特性拓展

- 聚集下推
- 海量数据更新
- 备份功能
- 统计数据自动更新
- Spill到分布式存储
- 异常处理
- 安全功能:TDE等
- 数据流动: kafka, flink, COPY等...
-

CONTENT

目录

—

01 PieCloudDB 简介

02 PieCloudDB 的诞生之路

03 PieCloudDB 的产品规划

- 元数据

- 解耦 FoundationDB（可替换）
- 解耦分层服务：降低复杂度、降低学习曲线、提高稳定性
- 性能优化：tps and scalability

- 用户数据

- 性能：（稀疏）索引、MergeOnRead 访问、Bloom filter.....
- 更多按需的存储引擎
- 分布式缓存和调度

- 优化器

- PieCloudDB 优化器相对已经较为完善
- 但还需要针对云原生环境下的功能完善和创新
 - 比如 Lateral, Subplan 优化等

- 计算引擎

- 第一步是 SIMD 执行引擎，各种性能优化技术开发
- 第二步是 Pipeline 执行引擎
- 第三步是各种条件下的计算资源隔离调度，用户真正的 serverless
- 异构计算数据更长远的将来

- 数据流动:

- 流式数据
- 外部数据源或者引擎（包括开源数据湖格式）
- 读取PieCloudDB数据
- 数据共享、数据分支等
- 生态工具（BI、ETL/ELT等）
- 数据治理

- 数据库管理

- 数据分析:

- AI 与数据科学
 - Data Frame
 - PL
- GIS
-

πCloudDB



加入PieCloudDB技术群

获得更多技术干货

关注我们，随时获得资讯

OpenPie | π CloudDB



关注OpenPie公众号

获得更多资讯



关注PieCloudDB公众号

获得产品动态

自主 · 创新 · 引领

THANKS