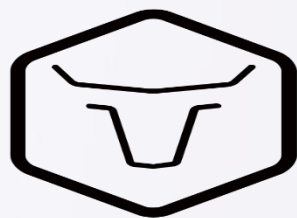




华为云数据库 Taurus 介绍

数据库服务产品部, Cloud BU

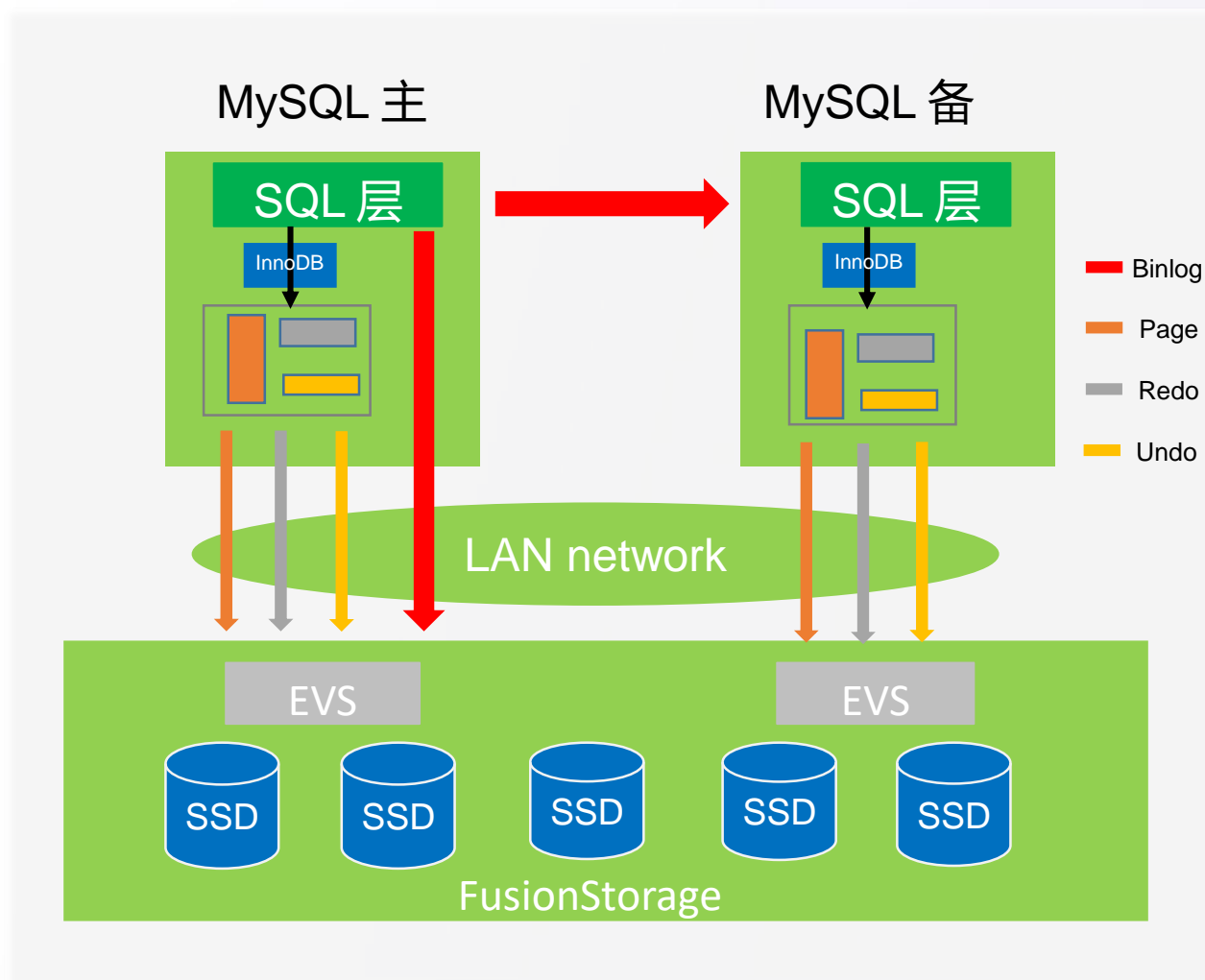




华为云数据库 Taurus

Taurus 是华为自研的最新一代企业级高扩展海量存储分布式数据库，完全兼容 MySQL。基于华为最新一代 DFV 存储，采用计算存储分离架构，支持1主15只读的高扩展性、128T的海量存储，无需分库分表，数据0丢失，成为企业级数据库解决方案。

RDS for MySQL 的当前架构



- MySQL 主备双机高可用, Binlog 复制(DRBD)
- VM+EVS /容器+本地盘
- 备库不可见, 可以添加最多5个只读实例
- HA模块负责主备切换
- 采用 extrabackup 物理备份/逻辑备份/快照备份

RDS for MySQL的瓶颈



浪费计算资源

备库往往不承担流量；复制线程会消耗主库/从库计算资源



浪费存储资源

EVS 是3副本冗余， $3*(N+1)$ 个副本，冗余严重



浪费网络资源

多种数据需要走存储网络进行刷盘，性能影响较大；写放大问题



添加只读很慢

往往需要拷贝数据，重放 binlog，尤其是采用本地盘方案



复制延迟问题

只读数据落后；主备倒换受到影响，难于保证 RTO

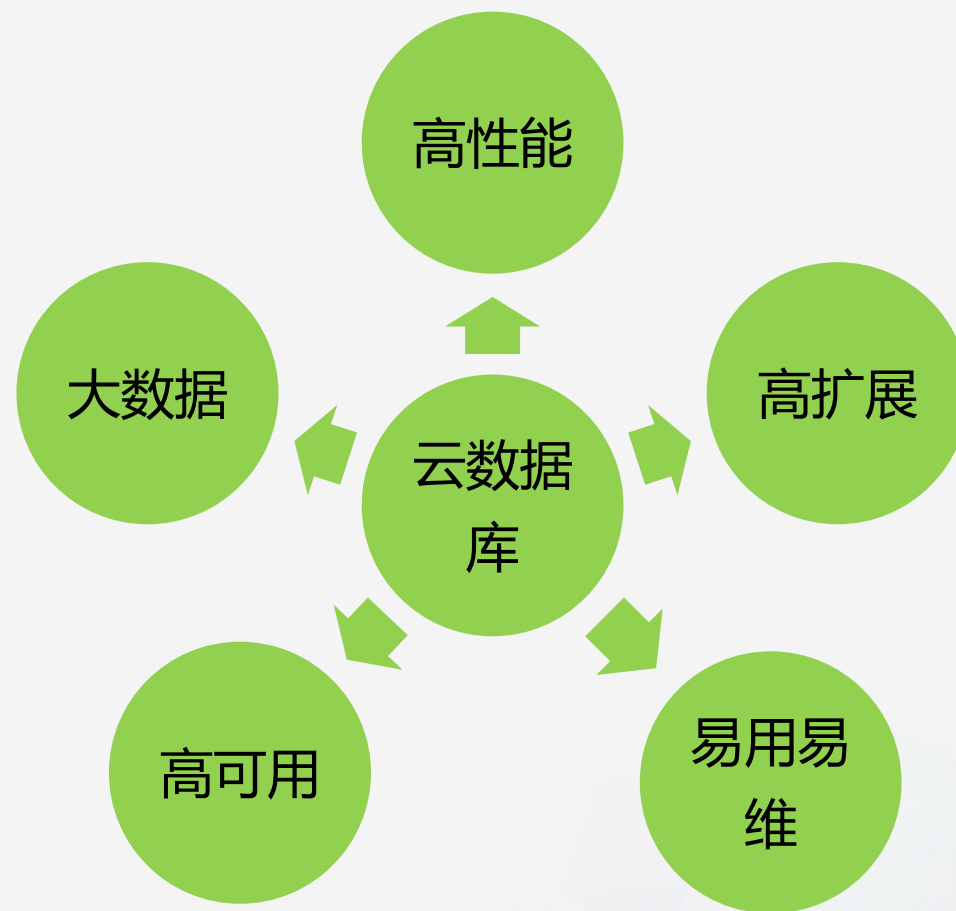


备份恢复慢

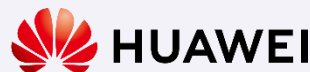
对于1T 以上的数据量都以小时计

来自客户的声音

TopN 需求和痛点	需求描述
兼容 MySQL	对于原有 MySQL 应用无需任何改造
海量数据存储	支持互联网业务的大数据量
分布式高扩展	自动化分库分表或者非分库分表，应用透明
无缝弹性扩容	分钟级弹性扩容，应用透明
强一致事务	支持分布式事务的强一致性
高可用	支持跨AZ高可用、跨 Region 容灾
高并发性能	支持大并发下的高性能
非中间件式架构	非 DDM 类方案（或者非 DRDS 那样）



华为 Cloud Native 数据库设计原则



解耦

- ✓ 计算与存储分离
- ✓ 主从解耦



近数据计算下推 (NDP)

- ✓ IO密集型负载下推到存储节点完成, 比如: redo处理, 页重构



充分利用云存储的能力

- ✓ 存储层实现独立容错和自愈服务
- ✓ 共享访问(单写多读)



发挥 SSD 的优势

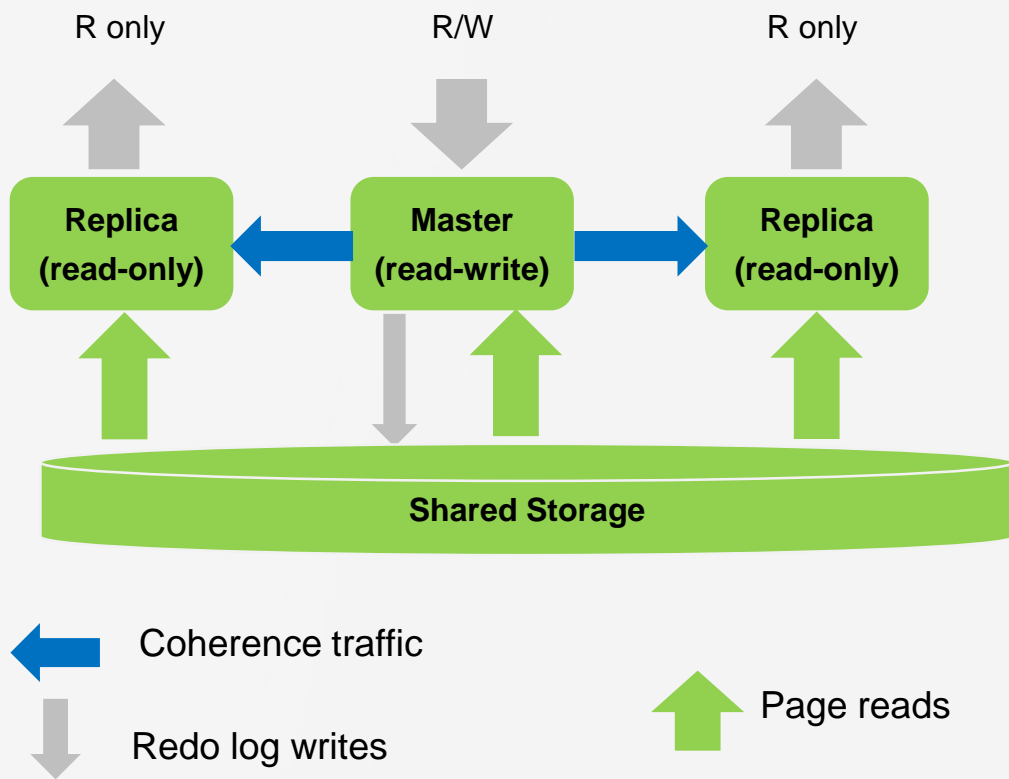
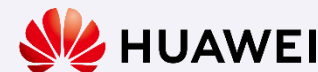
- ✓ 避免随机写带来的写放大, 减少磨损, 减小时延;
- ✓ 充分利用 SSD 的随机读性能



性能瓶颈已经从计算和存储转向网络

- ✓ 减少网络流量
- ✓ 采用新的网络技术和硬件, 比如: RDMA

Taurus 的架构和特点



◆ 极致可靠

数据0丢失，故障闪恢复，支持跨AZ 高可用。

◆ 多维扩展

计算节点双向扩展，横向扩展：支持1写15读横向扩展，纵向扩展：在线弹性扩容，按需计费

◆ 海量存储

单实例扩容数据达128T，无需分库分表，超低应用改造，业务急速上云。

◆ 创新自研

cloud-native 分布式数据库架构，基于华为最新一代 DFW 存储计算存储分离，保证扩展性价比；数据库逻辑下推存储，最小网络负载，极致性能。

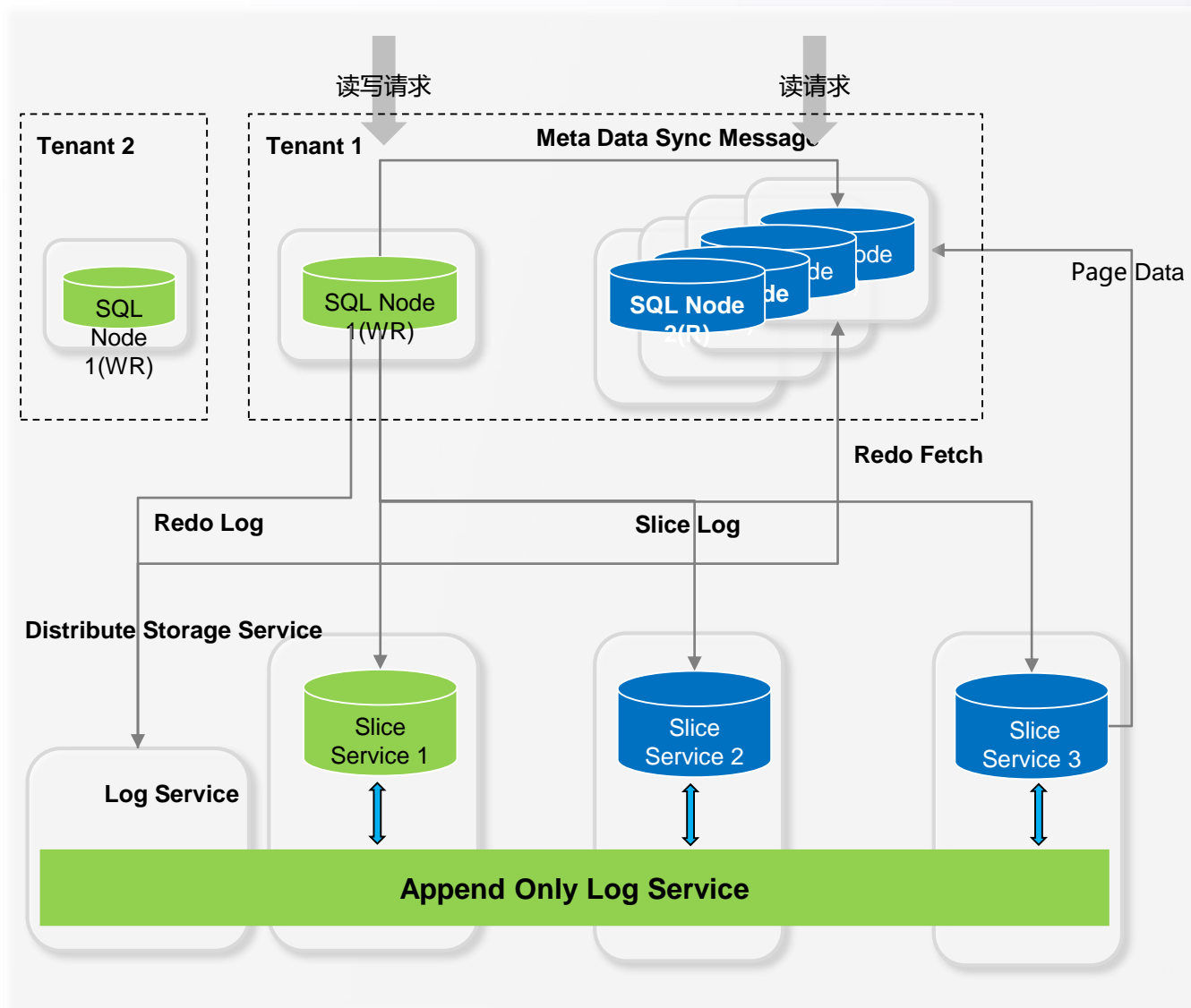
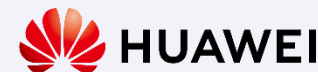
◆ 卓越性能

性能最高提升至原生 MySQL 的7倍，100%兼容 MySQL，业界领先

◆ 尖端硬件

基于 V5CPU+Optane DC SSD+RDMA 网络行业领先硬件组合，数据处理既稳又快

Taurus 更高性价比



共享 DFV 存储

与传统的 RDS for MySQL 相比，只有一份存储。添加一个只读节点时，只需添加一个计算节点，无需再额外购买存储。如果只读节点越多，节省的存储成本更多

Active-active 架构

与传统的 RDS for MySQL 相比，不再有备库的存在，所有的只读都是 active 状态，并且承担读流量，使得资源利用率更高

日志即数据架构

与传统的 RDS for MySQL 相比，不再需要刷 page，所有的更新操作仅记录日志，不再需要 double write。减少了宝贵的网络带宽

Taurus 扩展性、可靠性更高



横向扩展更快

与传统添加只读副本时，需要同步数据不同。Taurus 由于共享存储，只需添加计算节点，无论多大数据量，只需5分钟左右。



采用分布式存储

存储容量最大达128T，存储采用按需付费，不需要提前规划存储容量。减少用户成本。

01

02

03

04



主备倒换时间更快

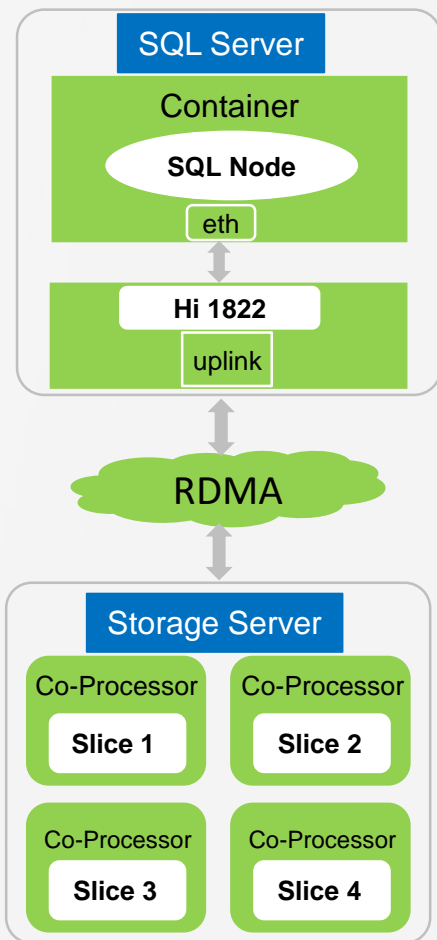
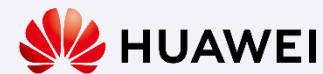
消除了Binlog复制延迟，RTO有保证。



数据库crash恢复更快

存储层在不断的异步、分布式对日志进行推进。

Taurus 性能更强



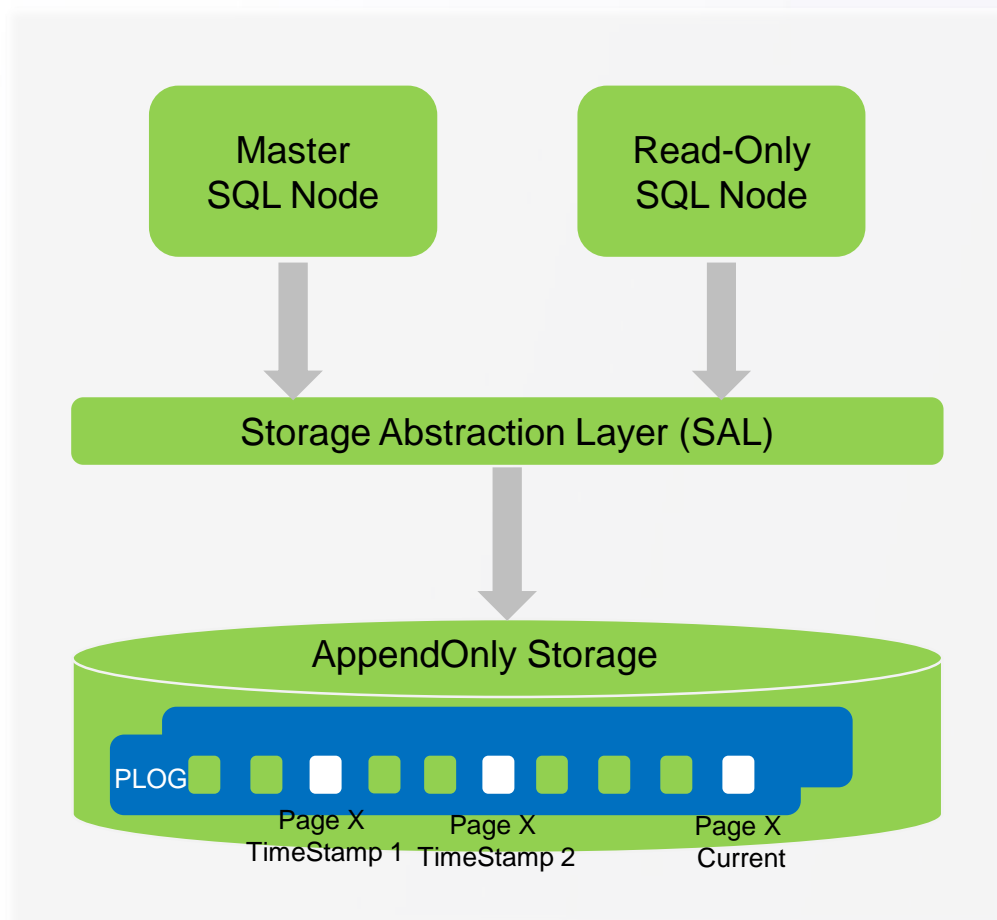
内核优化

- 去掉 double write
- Query Cache/Plan Cache 优化
- Innodb Lock Management 优化
- Audit Pluggin 效率优化
- 社区 Bug 修复

硬件提升

- 容器化
- Hi1822卸载
- NVMe SSD
- RDMA

Taurus 备份恢复更快



专为 Taurus 引擎定制的分布式存储系统，极大提升数据备份、恢复性能

01 强大的数据快照处理能力

AppendOnly vs. WriteInPlace, 数据天然按多时间点多副本存储，快照秒级生成，支持海量快照

02 任意时间点快速回滚

03 快速实例恢复

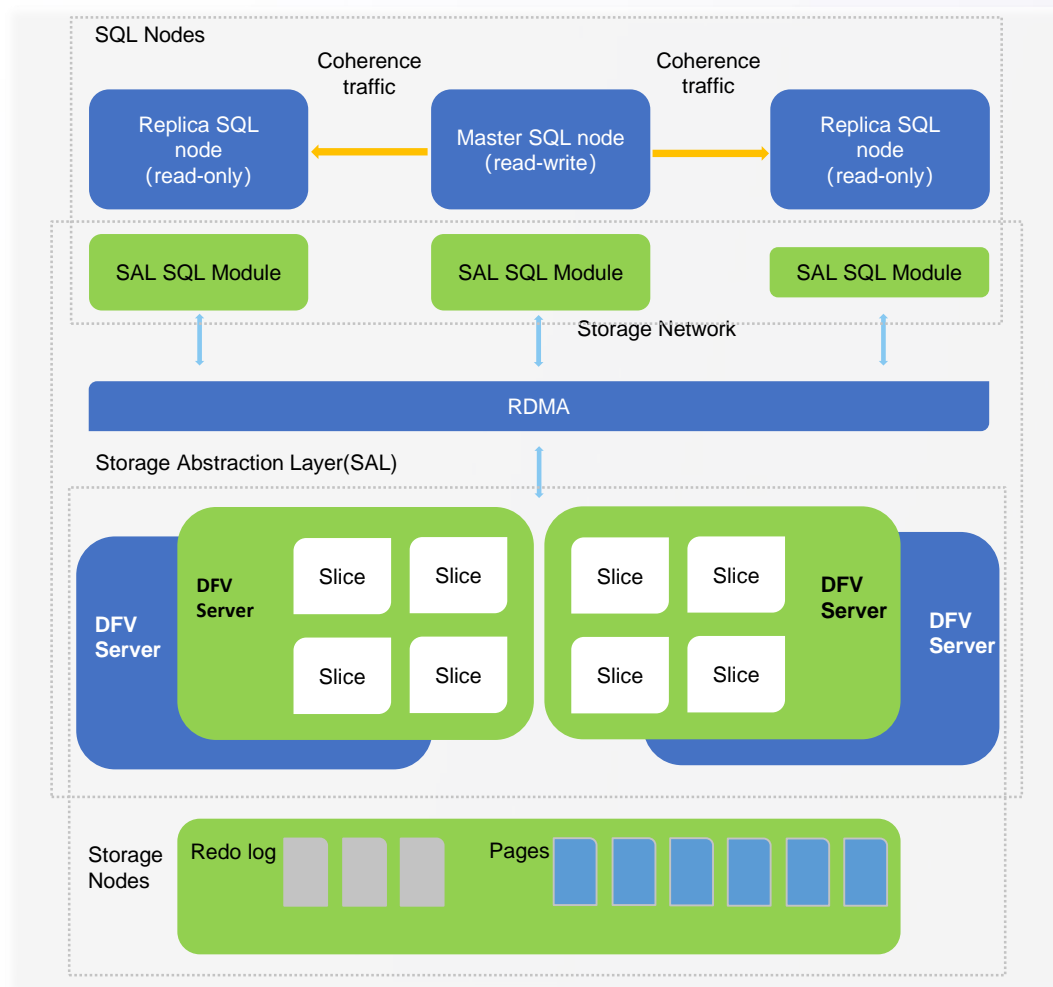
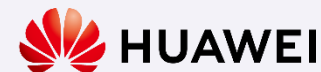
基于底层存储系统的多时间点特性，不需增量日志回放，可直接实现按时间点回滚

备份及恢复逻辑下沉到各存储节点，本地访问数据并直接与第三方存储系统交互，高并发高性能

04 快速实例恢复

通过异步数据拷贝+按需实时数据加载机制，Taurus 实例可在数分钟内达到完整功能可用

Taurus云数据库概览



主节点

- ✓ 处理所有的写
- ✓ 写 WAL 日志 (Redo)

01

只读实例

- ✓ 处理只读请求
- ✓ 快速升主
- ✓ 快速增/删

02

数据分区

- ✓ 数据页以 slice 组织, 分布在各个 slice 服务
- ✓ Slice 冗余保存
- ✓ 日志按 slice 分发

03

Slice 服务

- ✓ 单数据库可以分配多个 Slice, 访问隔离
- ✓ 存储和处理日志记录
- ✓ 维护和重构数据页
- ✓ 服务数据页读请求

04

SQL Nodes



SQL Node

Trans
action

Buffer
Pool

SQL
Parse

SQL Node

SQL Node

Storage Node

Slice Service

Log
Write

Log
Replay

Page
Manage
r

Log
Read

Crash
Recover
y

Log
Replicat
ion

Storage Node

Slice
Service

.....

Log
Replicat
ion

Storage Node

Slice
Service

.....

Log
Replicat
ion

管理客户端连接、解析 SQL 请求、分析执行计划并执行和管理事务隔离级别

一主多读

主库和只读松耦合

主库和只读通信少

快速 Failover

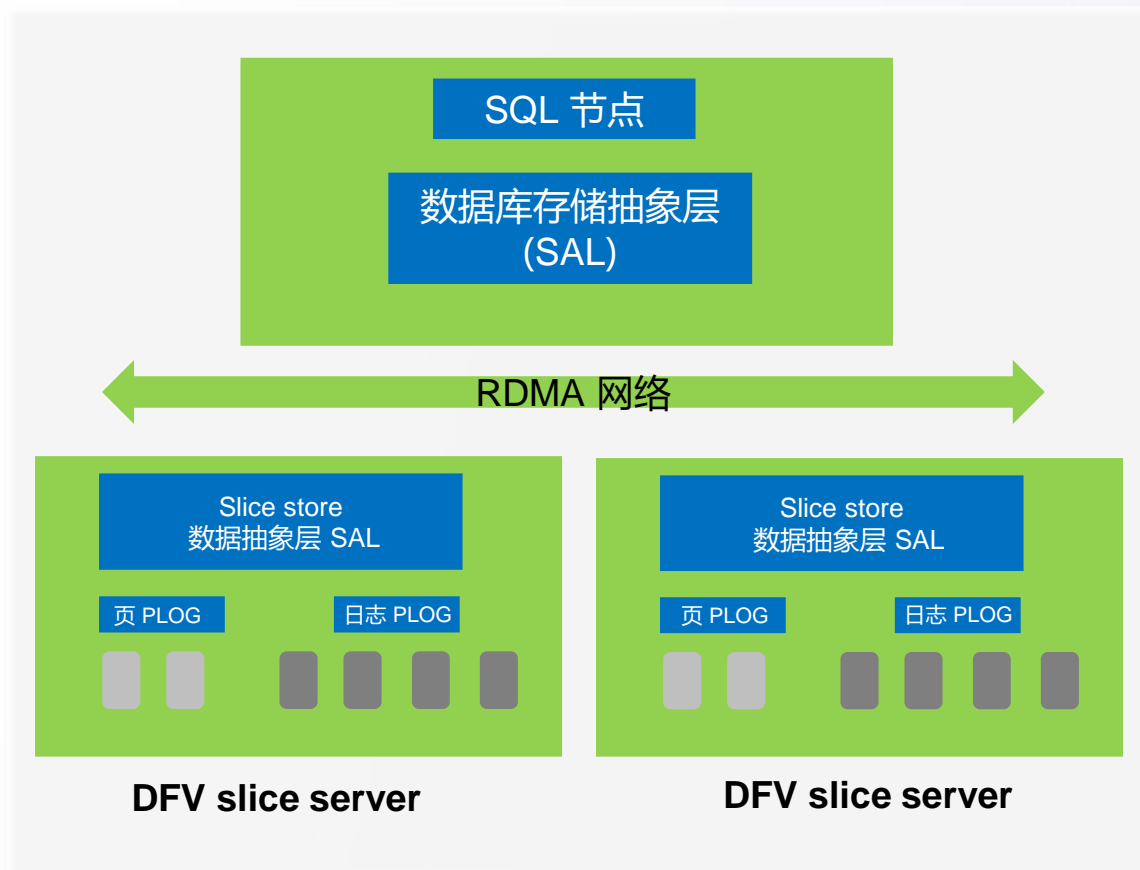
引入 HWSQL 内核中的一些特性

Query result cache

Query plan cache

Online DDL

存储抽象层 (SAL)



存储抽象层 (SAL)是逻辑层

将数据存储和 SQL 前端、事务、查询执行等进行隔离

由在 SQL 节点上执行的公共日志模块和存储节点上执行的 slice 节点组成

SAL 将所有数据页基于 {spaceID, pageID} 划分为 slice

横向扩展 - 随着数据库规模的增长, 可用资源 (存储、内存) 随着 slice 的创建按比例增长.

数据本地化 - 数据密集型操作是在存储节点上由 slice 服务执行

极致可靠：数据0丢失，故障闪恢复

金融行业：

监管诉求驱动，可靠性可用性是核心诉求

典型客户：



招商银行
CHINA MERCHANTS BANK



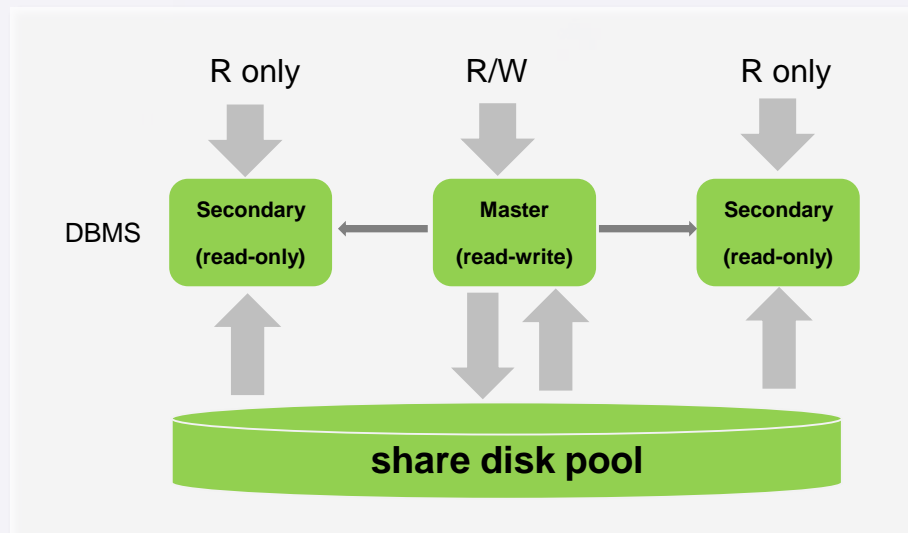
富途证券



行业痛点：

不能接受数据丢失，故障恢复秒级

传统MySQL不能支持此高要求



关键技术：计算存储分离，数据零丢失，强一致

Taurus采取计算存储分离架构，将数据持久化放入新一代存储中。

- 算节点故障秒级恢复
- 计算节点故障秒级恢复
- 存储3副本容错，单点故障0中断
- 支持跨AZ部署，跨Region容灾

多维扩展，海量存储，满足企业高速发展需求

金融行业：

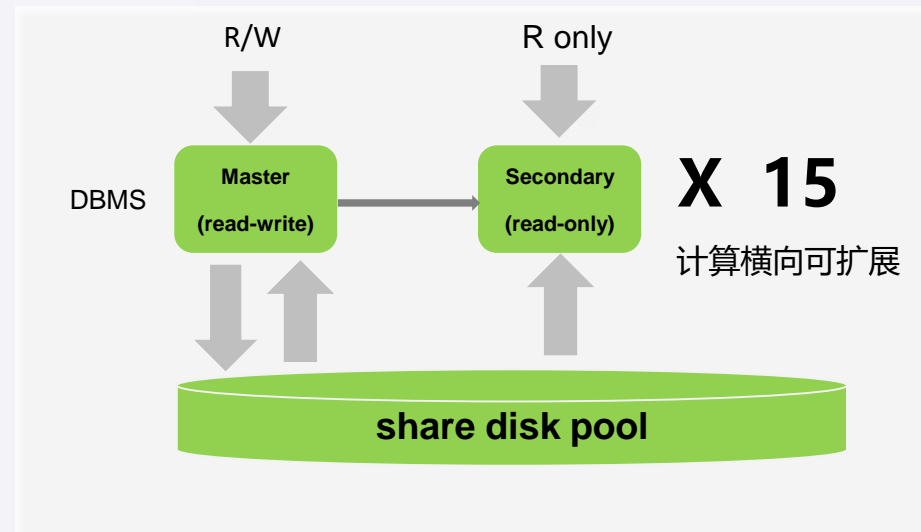
轻资产，快速扩容是其使用云数据库驱动力

典型客户：



行业痛点：

无法预测用户流量以及产生的数据量，业务高峰时客户体验会受到影响，甚至要停服扩容



Share disk 3副本可支撑上层15节点

计算节点双向扩展

- 基于云虚拟化，单节点可规格变更
- 支持1写15读，扩展比0.9

存储池化

- 最大支持128TB存储容量
- 计算节点扩容不会带来存储成本上升
- 存储按需计费，扩容不中断业务

强悍性能支撑业务拓展

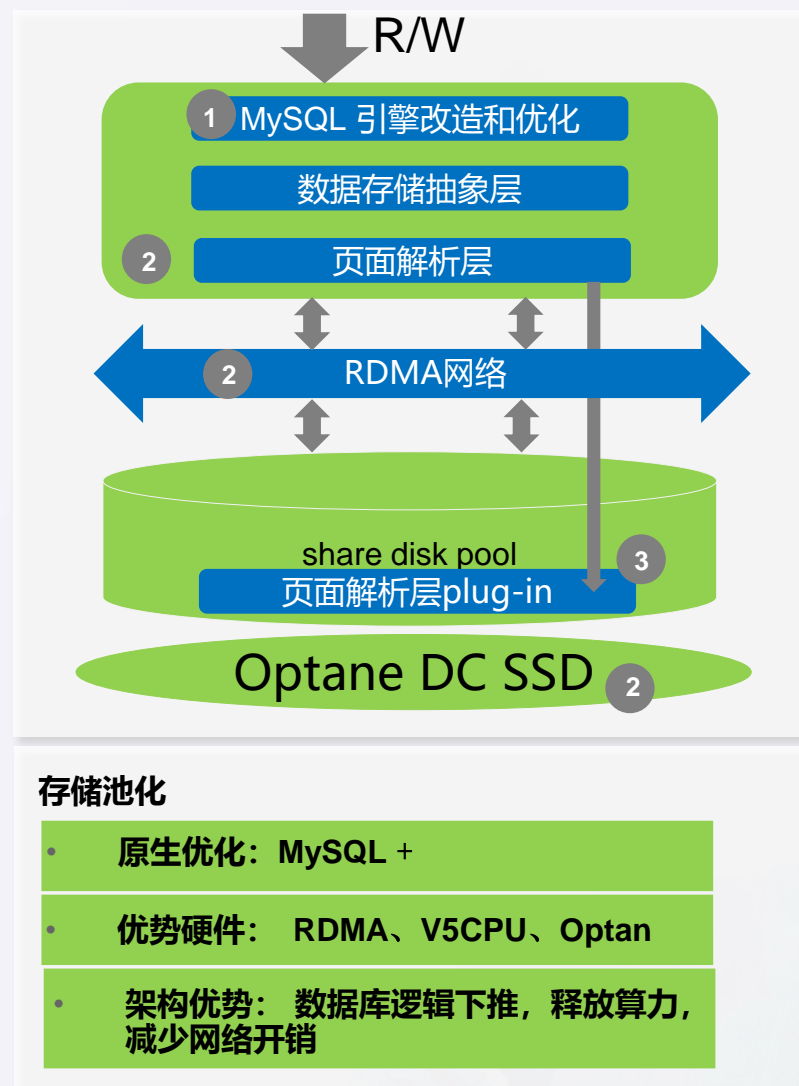
SaaS应用切入企业级市场、大型互联网公司、传统大企业：
高吞吐，高性价比

典型客户：



务痛点与客户诉求

- 业务庞大，吞吐量很高，开源库无法解决，采取分库分表等复杂化方案
- 企业客户一般偏好使用商用数据库（SQL Server、Oracle），license费用高



提供7倍于原生 MySQL 的性能

产品官网相关链接

产品主页: <https://www.huaweicloud.com/product/taurus.html>

帮助中心: <https://support.huaweicloud.com/productdesc-taurusdb/introduction.html>

FAQ: https://support.huaweicloud.com/taurusdb_faq/taurus_faq_0001.html



THANK YOU

