

国产 HTAP 数据库 在金融规模化数据服务场景的应用





国产 HTAP 数据库 在金融规模化数据服务场景的应用



某头部银行核心领域多个系统有约 500TB、跨度十余年的联机交易数据。依托国产 HTAP 数据库 PB 级扩展、金融级高可用、实时 HTAP 和分布式强一致等关键优势,构建了企业级一栈式综合数据服务系统。该系统提供高并发联机访问、大数据量批处理和高业务连续性保障,通过一套数据库系统取代原有的 Oracle、MongoDB、Hive 等多种数据栈,极大简化了整体数据架构,增强了业务敏捷性。这一举措为金融机构实现技术自主、业务创新和降本增效提供了有力支撑。

关键词



HTAP 数据库



分布式技术



综合数据服务系统



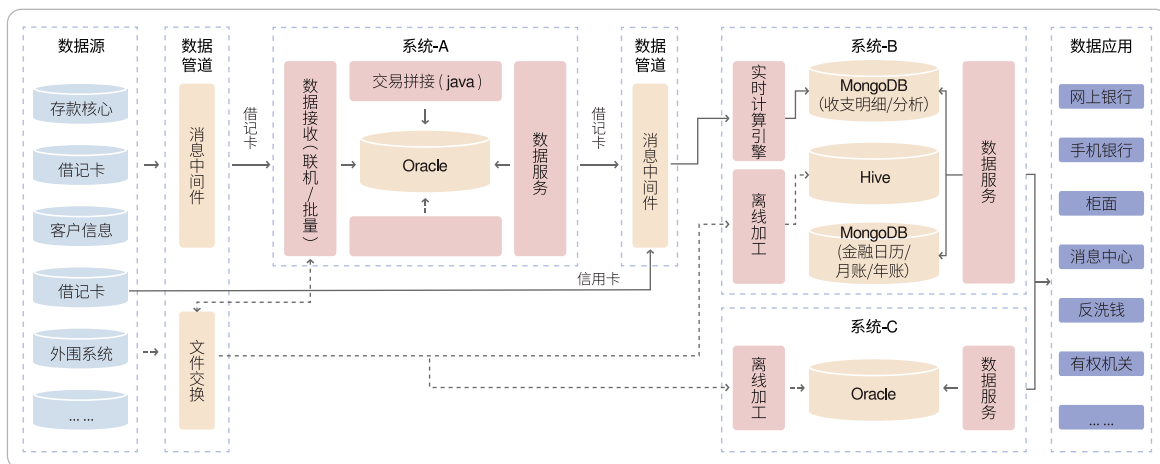
实时数仓

在新一代数字技术推动下,全球正在加速迈进数字经济时代,数据已成为关键生产要素。金融行业的数据技术栈过去高度依赖于国外厂商。然而,随着数字化转型步伐的加快和分布式、HTAP (hybrid transactional analytical processing, 混合事务与分析处理)、云原生技术的不断演进,国产数据库在分布式系统架构、OLTP/OLAP 混合负载、高可用领域等方面有了长足的进步,其稳定性、可靠性和安全性得到验证,已具备成熟且可持续发展的生态体系。本文以某头部银行客户交易综合查询场景为例,阐述国产 HTAP 数据库在金融规模化数据服务场景的应用实践

业务现状和痛点分析

对于银行核心交易领域服务属性最强、服务类型最综合的多维度交易信息查询、实时收支分析类业务来讲,原先的竖井式架构(如图 1 所示)极大地制约了业务发展,暴露出查询周期和跨度偏短、数据完整度/准确度不足、统计口径不一、数据时效性不高等短板。为了更好地响应业务要求,该头部银行采用国产 HTAP 数据库打造面向全球客户的企业级一栈式综合数据服务系统。

(图 1) 原有基于多套数据栈的系统架构



一栈式综合数据服务系统建设

建设思路与技术路线选择

该头部银行遵循企业级架构的设计原则,以信息整合“一个中心”、多维刻画“一套标签”、全景展示“一张视图”、信息标准“一套规范”的建设思路,通过业务架构设计,打破部门壁垒,消除业务及流程竖井,构建企业级数据模型。数据采集范围覆盖全渠道、全产品,提升高价值交易型数据的完整性、准确性、扩展性和对外服务能力。

混合事务与分析处理 (hybrid transactional analytical processing, HTAP) 技术是一种基于一站式架构同时处理事务请求与查询分析请求的技术。HTAP 基于一站式架构混合处理 OLTP 和 OLAP 的负载,省去了传统架构中的 ETL 过程,使得企业可以在数据密集型应用中分析最新事务数据,从而做出及时且准确的决策。例如,大型银行和金融系统可以在处理高并发交易事务的同时检测出异常的欺诈交易,既提供了高质量交易服务,又保障了交易安全。[1]

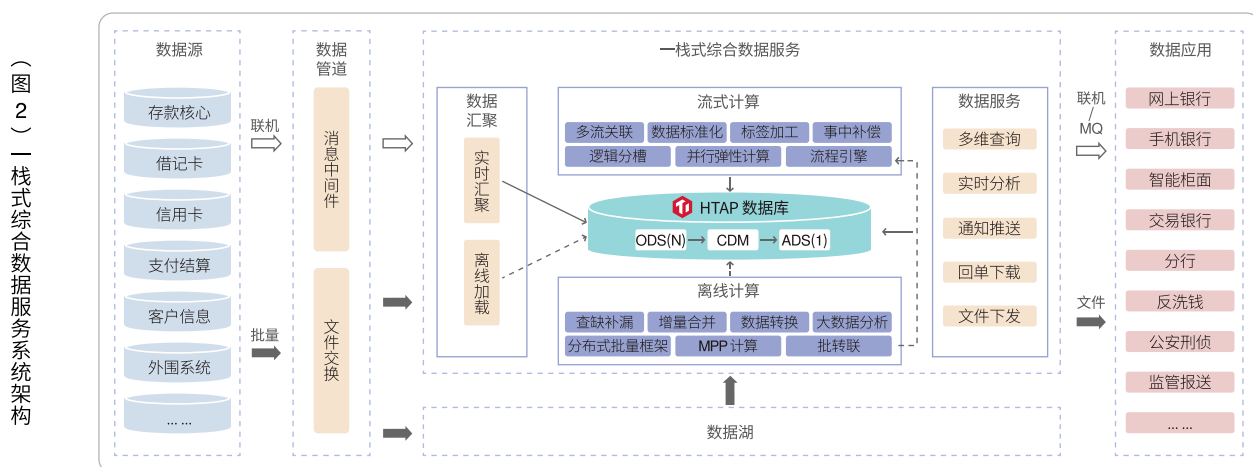
经过对国产单机 MySQL 数据库、分库分表数据库和分布式 HTAP 数据库进行全面的对比验证后(如表 1 所示),该头部银行选择了 TiDB 构建一栈式综合数据服务系统。

	某单机 MySQL 数据库	某分库分表数据库	分布式 HTAP 数据库 TiDB
架构维度	<ul style="list-style-type: none"> 需通过数据访问代理、分片键实现跨库数据路由 需通过索引库实现“全局二级索引” 	<ul style="list-style-type: none"> 通过数据访问代理、分片键实现跨分片数据路由 需通过索引表实现“全局二级索引” 	<ul style="list-style-type: none"> 存算分离架构, 数据自动分片, Raft 共识协议保障一致性 [2] 支持业务透明的“全局二级索引”
数据分片	跨库访问必须设计表级别的分片键, 业务侵入性大	跨分片访问必须设计表级别的分片键, 业务侵入性大	无需绑定分片键, 按需设计主键、索引
跨分片事务	应用侧实现	由 GTM 实现	Percolator 模型, 去中心化
多维查询	<ul style="list-style-type: none"> 分片键维度性能最优, 非分片键需要经索引库转换 关联键、where 条件中都包含分片键才可 join 下推, 否则会退化成数据节点笛卡尔积关联 	<ul style="list-style-type: none"> 分片键维度性能最优, 非分片键需要经索引表转换 关联键、where 条件中都包含分片键才可 join 下推, 否则会退化成数据节点笛卡尔积关联 	<ul style="list-style-type: none"> 支持多维度灵活关联, 多种 join 算法 筛选度高的可使用 TiKV, 数据量大、筛选度低的可使用 TiFlash
跨分片批处理/分析	不支持, 通过汇聚库实现	跨分片访问易产生数据节点间的笛卡尔积, 不适合	支持, 可根据场景选择 TiKV 或 TiFlash 实现
容量弹性扩展	<ul style="list-style-type: none"> hash 分片策略会导致数据重分布 range 分片策略需借助单元化工具实现单元扩容 	<ul style="list-style-type: none"> 目前已有基于一致性 hash 的分片机制 支持相对友好的 range 分区机制 	默认基于 range 分片的策略、以及 region 级细粒度的调度能力, 使弹性伸缩更为优雅, 业务影响可控

表 1: 国产数据库的综合能力对比

一栈式综合数据服务系统架构

系统采用典型的实时数仓架构,通过一种数据库整合替代了原先 Oracle、MongoDB、Hive 多种异构产品,并通过两地三中心高可用架构,确保了高等级应用系统的业务连续性,同时支持应用双活。系统架构整体包括数据汇聚、数据加工、数据服务和存储四大业务模块(如图 2 所示)。



数据汇聚层

- 基于 HTAP 数据库的弹性扩展和高并发写入能力,支持基于消息中间件的实时汇聚、基于文件交换的离线加载,覆盖众多上游系统的多种类型数据。

数据加工层

- 具备流式计算、批量处理的混合处理能力。流式计算基于 HTAP 数据库的动态数据路由、原生分布式事务能力,通过流程引擎、无锁化逻辑分槽等机制并行对多个源表进行关联、数据标准化、标签加工处理,应用无需关注底层数据分布和跨分片事务,只需按业务要求编写业务 SQL 即可。离线计算基于 HTAP 数据库的行列混合、MPP 能力,实现了日均百 GB 增量数据的快速导入,以及几十亿级数据的聚合分析、指标结果高并发访问,避免了多套异构数据库加 ETL 传统模式的局限性、复杂性,极大简化了 HTAP 场景的技术架构。

数据服务层

- 利用 HTAP 数据库的多维度数据访问能力,向全行客户提供不同维度、灵活条件的高并发 T+0 数据查询、分析、推送和数据下发等服务。

数据存储层

- 基于 HTAP 数据库实现了数据的统一存储、冷热分离,兼顾了处理性能和 SLA 等级,实现了更优的成本控制。

应用成效

一栈式综合数据服务系统基于企业级架构规划,采用国产 HTAP 数据库、操作系统和 IaaS 平台,在对接业务规模、系统资源规模、数据规模和应用规模等方面都具有示范效应:

- **对接业务规模:**系统共对接了近百个上下游业务系统,数据覆盖全行近 230 个业务产品、超 3,000 个交易场景;
- **系统资源规模:**系统集群资源约 2,700 个应用虚拟节点、300 个数据库物理节点;
- **数据规模:**迁移了原有多个系统近 500TB 的单副本存量数据,新系统多副本数据规模近 PB 级,最大数据表达上千亿行记录;
- **应用规模:**日均接收和加工亿级增量数据,联机写入 TPS 约 10,000,日均通知推送约 3.5 亿笔,核心加工链路耗时控制在 5 秒内;日均联机查询交易量超千万笔,峰值 QPS 超 50,000,应用侧平均响应耗时在 100ms 内;日均加载上千个批量文本,最大离线分析规模为基于 40 亿行原始明细产生约 4 亿行指标结果数据,并为多个下游系统提供数据下传和推送服务,包括反洗钱、有权机关等监管类场景。

自 2023 年 7 月底上线投产,新系统整体运行平稳、对外切换停机窗口短,实现了于对客户无感的平滑迁移,同时各项服务指标较原先都有所提升。新系统整合了多个技术栈,整体维护成本更低,更有利于业务的快速迭代。与原来最大访问周期小于 5 年、跨度小于半年的查询方式相比,新系统能够支撑超 10 年的永久查询和单次跨度 5 年的大范围查询及实时分析,且交易数据完整性、准确性更高,数据统计口径更为统一。

技术创新点

该项目的成功落地是前瞻技术在头部银行重要业务系统规模化应用的典范,其先进性主要体现在以下几个方面:

- **规模化的企业级数据应用能力:**实现上百个业务系统数百 TB 数据的整合与供给,支撑灵活多变的数据消费场景,提升内部、外部、生态伙伴的多元化数据服务效率。
- **更为精细化的降本增效:**基于同城双活架构,双机房同时具备读写能力,在吞吐能力、弹性伸缩能力、部署密度、资源使用率方面较原有模式均有较大提升,实现灵活高效的资源调配和精细化的降本增效。
- **高效集约的企业级运营能力:**支持运营需求共享,建立线上化、自动化、智能化的新型运营模式,在改善体验、提高效率、集中管控风险等方面实现飞跃。
- **实时数据服务提升客户体验:**通过各类服务和访问入口的整合高效地支持定制化、差异化服务和精准营销,支撑超过 10 年的数据查询,提供实时、一致的客户体验。
- **依托分布式和 HTAP 实现整体架构跃迁:**兼顾海量数据存储、大规模多源汇聚、多样化数据实时分析和多维度数据访问,大幅提升了业务的敏捷性;一套创新的 HTAP 数据库替换了 Oracle、MongoDB、Hive 三套数据库体系,验证了 MongoDB 替换方案的可行性,实现了整体架构的跃迁。

参考文献

- [1] 张超, 李国良, 冯建华, 张金涛. HTAP 数据库关键技术综述. 软件学报, 2023, 34(2): 761-785. <http://www.jos.org.cn/1000-9825/6713.htm>
- [2] Huang DX, Liu Q, Cui Q, et al. TiDB: A raft-based HTAP database. Proc. of the VLDB Endowment, 2020, 13(12): 3072-3084.



平凯星辰(北京)科技有限公司



北京 | 上海 | 杭州 | 成都 | 深圳 | 广州 | 硅谷 | 东京 | 新加坡
咨询热线: 4006790886 www.pingcap.com



立即咨询
TiDB 企业版



PingCAP
WeChat

  是平凯星辰(北京)科技有限公司的商标或者注册商标。在本手册中以及本手册描述的产品中,出现
的其他商标、产品名称、服务名称以及公司名称,由其各自的所有人拥有。©平凯星辰(北京)科技有限公司, 2023。

平凯星辰(北京)科技有限公司保留一切权利。未经事先书面许可,严禁复制、改编、汇编、传播或翻译本资料全部或部分內容。