




金融业数据库创新发展报告

(2023)

金融信息化研究所 (FITI)

2023年10月



编制委员会

主任:

潘润红

副主任:

黄程林、庄文君

编委会成员（排名不分先后，按姓氏拼音排序）：

蔡仕志、陈志刚、戴 涛、杜志明、方 兴、冯 程、冯文忠、
郭志军、侯 敏、黄海明、李永海、马波勇、马晓煦、苏光牛、
王丽静、王新明、王义成、隗 华、巫建刚、肖静静、邢 磊、
徐晓剑、徐 翥、杨 阳、俞 枫、张俊喜、张 玮、张晓强、
赵 培

编写组成员（排名不分先后，按姓氏拼音排序）：

白军奎、鲍思佳、毕斓馨、曹平国、陈伟红、程 静、从平平、
崔安颀、管文琦、郭 茁、何 格、黄晨晨、黄 韦、李 凯、
李 倩、李瑞超、李文彬、李昕哲、林 海、刘 畅、刘传友、
卢 强、骆君柱、骆 毅、吕伟初、苏 强、隋东辉、孙存福、
孙腾腾、孙 伟、唐思源、王 枫、王富国、王 辉、王鹏冲、
王荣鑫、王帅强、王文清、王 栩、王 瑜、王子健、伍 华、
夏文勇、肖淑男、严 恒、颜 龙、杨 锐、杨征涛、张俊成、
张子鉴、赵义斌、周日明、邹 鹏

主编单位:

北京金融信息化研究所
中国工商银行股份有限公司
交通银行股份有限公司
平安银行股份有限公司
国泰君安证券股份有限公司
中国太平洋保险(集团)股份有限公司
泰康保险集团股份有限公司
华为云计算技术有限公司
中兴通讯股份有限公司
武汉达梦数据库股份有限公司
天津南大通用数据技术股份有限公司
腾讯云计算(北京)有限责任公司
北京科蓝软件系统股份有限公司

参编单位:

中国农业银行股份有限公司
中国银行股份有限公司
中国建设银行股份有限公司
中国邮政储蓄银行股份有限公司
中信银行股份有限公司
招商银行股份有限公司

华夏银行股份有限公司
北京银行股份有限公司
上海证券交易所
中信证券股份有限公司
华泰证券股份有限公司
国信证券股份有限公司
中国人寿保险股份有限公司
蚂蚁科技集团股份有限公司
阿里云计算有限公司
北京东方国信科技股份有限公司

支持单位:

平凯星辰（北京）科技有限公司

摘 要

2023年2月，习近平总书记在中共中央政治局第三次集体学习会议时强调，“基础软件要从底层和源头抓起，打好科学仪器、操作系统和基础软件攻坚战”。数据库作为构建金融信息系统的关键环节，成为当前金融领域创新及数字化转型的一个难点和焦点。近年来，金融机构不断加大我国成熟数据库产品应用力度，引入不同类型数据库产品在多种业务场景实现应用，在支持金融业务创新发展的同时，逐步提升金融业数据库安全可控能力。

但由于我国数据库产品还存在需要弥补的不足，随着数据库改造逐渐深入到核心系统，带来了选型难度高、实施运维难度大等挑战，亟需加大产用联合攻坚、创新应用模式、研发生态工具，推动我国数据库由“能用”向“好用、易用”迈进。同时，数据能力建设对数据仓库、非关系型数据库也提出了新需求，需要加大数据仓库技术创新，分类推动不同非关系型数据库应用发展，促进产品成熟落地。此外，金融业大量采用了开源数据库，面临开源协议、安全漏洞、知识产权、代码感染、开源停服断供等风险，亟需通过加大开源数据库治理、安全防范、建设行业开源社区等措施进行风险防范。

本报告从行业应用情况入手，围绕当前业界关注的金融核心系统数据库应用、数据能力提升和开源数据库风险防控三个领域，总结发展成效经验，剖析问题与风险，提出发展建议，展望发展趋势，为金融业数据库创新发展提供参考借鉴。

目录

1. 概述	1
2. 金融业数据库应用现状	2
2.1 金融业数据库应用总体情况	2
2.2 我国数据库产品在金融业应用取得积极成效	5
3. 金融业数据库应用重点领域分析	7
3.1 核心系统数据库应用分析	7
3.1.1 核心系统对数据库应用提出更为严格需求	8
3.1.2 当前核心系统数据库应用面临多重挑战	9
3.1.3 围绕关键环节推动核心系统数据库转型升级	11
3.2 数据库支持数据能力建设分析	14
3.2.1 不同类型数据库助力数据能力建设取得积极成效	14
3.2.2 金融业数字化快速发展对数据库提出新需求	15
3.2.3 分类推动各类数据库应用创新促进数据能力提升	17
3.3 开源数据库应用风险分析	18
3.3.1 开源数据库快速发展并在金融业得到广泛应用	19
3.3.2 金融业开源数据库面临诸多风险	20
3.3.3 多措并举防范开源数据库应用风险	23
4. 金融业数据库应用展望	25
4.1 我国主流数据库产品成熟度将进一步提升	25
4.2 金融核心系统数据库转型升级将稳步推进	26
4.3 新技术与业务场景将推动金融数据库创新发展	26
4.4 金融业开源数据库应用风险防范仍然不容忽视	27
附录：金融业数据库优秀应用案例集	28

1. 概述

国家的战略部署及金融科技发展相关的多个专项规划，都对关键硬件信息技术的突破应用、构建自立自强的数字技术创新体系提出了明确目标和要求。金融机构在数字化转型创新发展中，纷纷加快了信息技术的迭代升级，不断引入包括数据库在内新的技术产品，并持续扩大应用广度和深度，金融业信息技术供应链安全整体水平不断提升。

作为核心基础软件的数据库，始终是金融业关注的焦点，也是难点。近年来，传统集中式、新兴分布式数据库，关系型、非关系型数据库，联机事务处理型(OLTP)、联机分析处理型(OLAP)、混合负载型(HTAP)数据库，商用数据库、开源数据库等在不同业务场景中实现创新应用，推动构建新一代金融 IT 基础设施，在确保安全稳定的同时，有力支持金融业务创新发展。特别是我国成熟的数据库技术产品在金融业实现了广泛应用，并逐步推动在核心系统应用落地，取得显著成效。

但金融业数据库应用还面临一些挑战。一是随着我国数据库产品应用向核心系统深入，其性能、功能及稳定性与国际商用主流数据库产品还存在一定差距，尚不能完全满足金融业核心系统应用需求；二是数据仓库、非关系型数据库对金融业数据能力建设有着重要作用，随着金融业数字化转型深入推进，对非关系型数据库的需求越来越多，对数据仓库功能、性能、扩展性和安全

性提出更高要求，亟需提升产品能力满足金融业应用需求；三是开源数据库在金融业应用广泛，但存在开源协议、安全漏洞、知识产权、代码感染、开源停服断供、政策不确定性、掌控及服务能力不足等诸多风险。

为推动金融数据库创新发展，以应对面临的挑战和问题，金融信息化研究所组织各方力量，在深入调研、交流研讨基础上，编制《金融业数据库创新发展报告》（2023），对金融业数据库技术应用进行现状梳理、问题分析、建言献策及成果展示，为管理决策提供参考，为金融机构创新发展提供支持，为我国数据库产业发展提供助力，促进金融科技稳步发展和金融数字化转型深入推进。

2. 金融业数据库应用现状

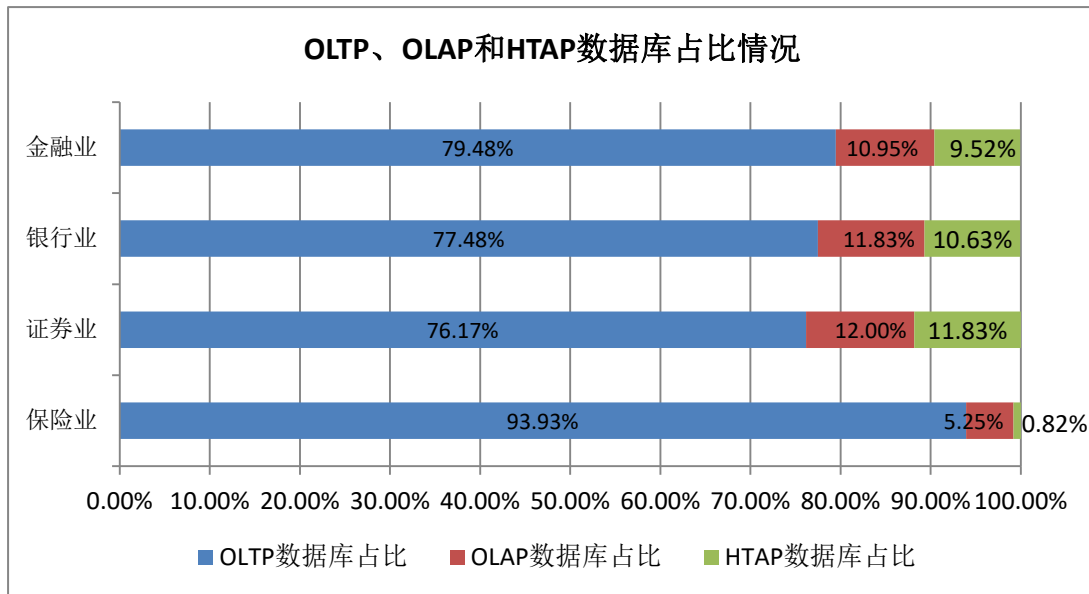
金融业始终走在 IT 技术应用和发展变革的前列，根据创新发展和安全可控的需要，金融机构普遍采取“先外围、后核心”的策略推进我国数据库产品的应用落地。随着多种类型的数据库产品在金融领域加快应用，促进了金融业务创新，同时也有力支持了我国数据库产业的发展。

2.1 金融业数据库应用总体情况

一是集中式数据库应用占比较高，分布式数据库应用呈现增长趋势。从产品架构看，金融业数据库呈现集中式和分布式并存发展态势。其中，集中式数据库以其较强的功能黏性、优秀的系

统稳定性、良好的软硬适配能力，目前在金融业的应用仍占据绝大多数份额。但随着金融业数字化转型的不断深入，分布式数据库因具备依托通用硬件、弹性扩展、内置高可用等特征，可更有效支持海量、高并发、高吞吐量的新型金融业务应用系统，在金融业的应用占比相较 2022 年调研结果实现了 5.76% 的增长。

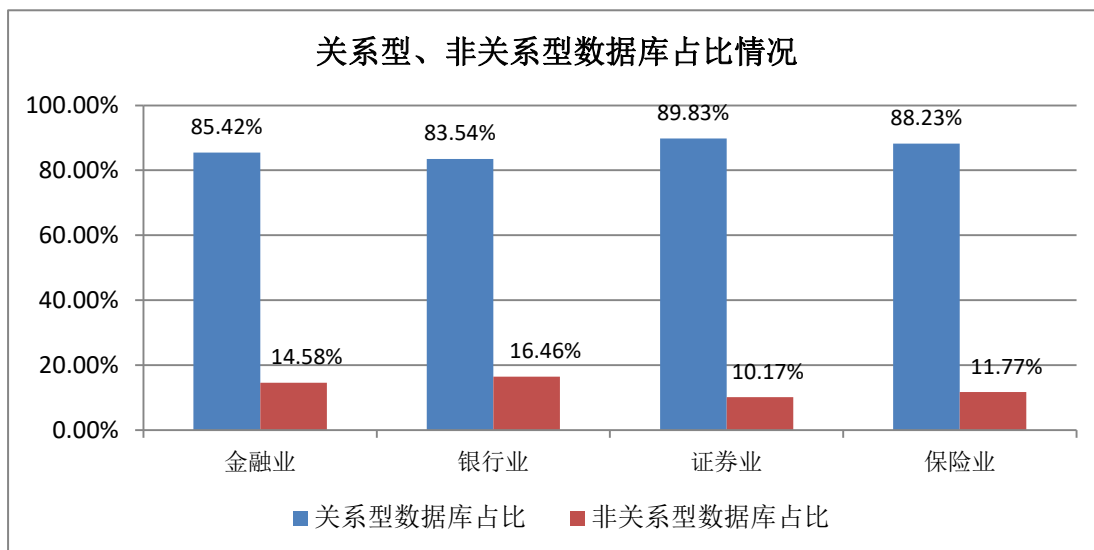
二是 OLTP 数据库应用占比较高，OLAP 和 HTAP 应用需求不断增加。金融业务系统的数据处理分为联机事务处理（OLTP）和联机分析处理（OLAP）两类。面向客户交易类、业务办理等系统通常选择 OLTP 类数据库，而报表类、分析类系统通常选择 OLAP 类数据库。随着金融业数字化进程加快，海量数据让 OLTP 和 OLAP 数据库的边界越来越模糊并不断融合，在同一个系统中同时需要 OLTP 和 OLAP 能力，即 HTAP 数据库的需求越来越多。通常 HTAP 主要应用分为以 OLTP 为主或以 OLAP 为主的两种 HTAP 场景。目前在金融业非融合的 OLTP 数据库占比仍然较高，为 76.83%。同时，不同细分行业中，银行业和证券业应用 OLAP 和 HTAP 数据库占比相对较高。详细情况如图 1 所示。



数据来源：金融信息化研究所

图 1 金融业 OLTP、OLAP、HTAP 数据库占比情况示意图

三是非关系型数据库在金融业加快探索实践。金融业具有客户量大、业务场景复杂等特点，需要存储处理多种类型数据，如庞大的用户、账户、交易、清算、产品、行情等结构化数据和海量的图像、视频、语音等非结构化数据。金融业务需要挖掘海量基础数据所蕴含的丰富信息资源，如隐藏的用户偏好、消费习惯、交易习惯、社会关系等，为非关系型数据库提供了丰富的应用场景，加快了非关系型数据库创新应用。目前，非关系型数据库在金融业应用占比已达到 14.58%。相对证券业和保险业，银行业应用非关系型数据库占比较高。另外，非关系型数据库中键值数据库和文档数据库应用最广泛，二者之和占有非关系型数据库比例超过 60%；图数据库、向量数据库、时序数据库开始探索应用，其中时序数据库在证券业首先开始应用，占比为 4.8%。金融业非关系型数据库应用占比情况如图 2 所示。



数据来源：金融信息化研究所

图 2 金融业关系型、非关系型数据库占比情况示意图

四是开源数据库在金融业得到广泛应用。与闭源商业数据库相比，开源数据库具有源码公开、使用成本低、获取途径广、对外开放、功能丰富等特点。这些优点使得开源软件得到广大开发人员的青睐，使用人员可在原有代码基础上进行业务适配修改，活跃的社区支持也为日益复杂的业务需求贡献了越来越多的解决方案，从而使得开源数据库在金融业实现了广泛应用。目前，约 90% 的金融机构都应用了开源数据库支撑业务发展，其中主要集中在一般业务系统和管理类系统。

2.2 我国数据库产品在金融业应用取得积极成效

一是我国数据库产品在金融业应用稳妥推进、持续增长。金融业始终走在 IT 技术应用和发展变革的前列，根据创新发展和安全可控的需要，金融机构都在以不同的速度和深度尝试应用我

国数据库产品支持业务发展。为确保安全生产、顺利实施，金融机构普遍采取“先外围、后核心”的策略推进我国数据库产品的应用落地，不断拓展应用广度和深度。调研显示，与 2022 年相比，目前应用我国数据库产品的金融机构数量占比增长约 10%，其中，证券业增幅明显。

二是我国数据库在金融业应用优势不断显现。金融业应用我国数据库产品，可针对特殊业务场景采取定制化改造，增强业务系统的服务能力，加快业务系统的响应速度，提升业务系统的吞吐量，本地化服务优势明显。同时，我国数据库产品在金融业的应用，一定程度上降低了企业成本，实现了降本增效。相比国外数据库高昂的购买费用和后续的技术服务成本，我国数据库产品后续成本投入相对稳定、成本可控。

三是金融业与数据库产业互相促进、发展共赢的成效显著。金融机构在应用我国数据库产品时持续进行产品能力测试，识别产品能力上的缺陷并不断反馈给数据库厂商，推动数据库产品能力提升。而且金融业拥有丰富的业务场景，我国数据库产品在大量的实践应用以及上下游产业的共同发展中，整体数据库能力、兼容性能力都有明显提高。随着我国数据库产品成熟度的不断提升，加快推动了新一代金融 IT 基础设施建设，不断夯实金融业数字化创新发展的基石。

3. 金融业数据库应用重点领域分析

随着我国数据库产品在金融业的不断推广应用，数据库改造逐渐深入到核心系统，金融机构普遍认为我国数据库产品在核心系统应用还有较大提升空间。当前，数字金融快速发展，数据仓库和非关系型数据库有力支持了金融机构开展数据挖掘和分析，推动数据能力建设，助力金融业务创新。另外，根据调研统计，约 90%的金融机构使用开源数据库支撑业务发展，而开源数据库应用面临着开源停用、协议风险、安全漏洞等多方面的风险，且大部分金融机构仅具备开源数据库的日常管理运维能力，仅少量大型金融机构具备代码级维护及二次开发能力。为此，本章主要从业界关注的核心系统数据库应用、数据仓库和非关系型数据库支撑数据能力建设、开源数据库应用三个重点方向，从深入应用、加大创新、防范风险三个视角来分析金融业数据库应用需求、面临的挑战，并提出下一步发展建议。

3.1 核心系统数据库应用分析

核心系统被喻为金融机构 IT 建设皇冠上的明珠。因其高可用、高可靠，高并发、高实时、高扩展、高安全等特性鲜明，要求提供 7x24 小时不间断服务能力，在信息系统中具有十分重要的地位，事关金融业务实现、金融安全稳定。因此，进行核心系统数据库转型升级牵一发而动全身，需要全面分析转型需求、面临的挑战，总结经验，推动转型实施。

3.1.1 核心系统对数据库应用提出更为严格需求

一是功能方面，核心系统新引入的数据库除了支持传统数据类型、对象类型、SQL 以及 PL/SQL 等基本功能外，特别需要在驱动、语法、架构等方面实现与现有国际主流数据库的兼容。且核心系统需要与交易、结算、风控等不同系统进行无缝集成，所以，需要核心系统数据库支持常用的数据接口和标准协议。同时，对于业务量大、性能要求高的核心系统，需要数据库适配新一代分布式、共享存储集群等架构体系。

二是在性能方面，核心系统对性能的极致要求，需要数据库具备极强的高并发、高速读写能力，以及满足低延迟、强实时性要求，并具备灵活的扩展性。核心系统数据库交易处理能力需稳定支持 1 至 2 万 TPS，峰值支持 2 至 4 万 TPS。核心系统的 OLTP 场景要求数据库响应时间能够达到毫秒甚至亚毫秒级别，OLAP 场景要求数据库响应时间能够达到秒甚至毫秒级别。此外，在分布式场景下，对数据库弹性扩展能力提出更高要求，实现自动在线横向扩展，扩展后性能呈线性增加，无明显损耗，且根据业务量的变化实现对数据库资源的弹性扩/缩容，以应对业务快速增长和海量数据处理需求。

三是鲁棒性方面，核心系统数据库需要具备稳定性、高可靠性和高可用性。在负载、数据量或者其它环境因素急剧变化时，核心系统数据库依然能提供稳定的、可预期的服务能力。核心系

统数据库需要满足金融核心业务对数据 100%可靠、99.999%以上的可用性极致要求。同时灾备能力要求 RTO 为秒级，RPO 为 0，且要求核心系统数据库能在线修复或者能自检测并自愈，保证金融服务的实时性和连续性。

四是安全方面，核心系统的数据不仅包括业务数据，还涉及到客户的敏感信息，其数据库在安全方面要具备“3A+E”（认证：Authorization、授权：Authentication、审计：Audit、加密：Encryption）的能力。通过提供数据库层的身份认证、访问授权、行为审计，保证正确的人、在正确的时间、以正确的方式访问正确的数据，并留下访问痕迹以供事后审计。随着数据资产化进程的加快，越来越强调精细颗粒度的访问授权。同时，为了防止数据泄露、失窃，多层次、多场景、基于商用密码加密算法的支持，对于多层次、灵活颗粒度的数据脱敏、数据遮蔽的支持也必不可少。

3.1.2 当前核心系统数据库应用面临多重挑战

一是我国数据库产品在核心系统应用还存在多方面能力不足。与国际主流数据产品相比，我国数据库产品在 SQL 优化器、全局事务管理、并发控制、列存和行列混存、共享集群、多读多写等多个关键核心技术上仍然需要打磨和进一步提升；在执行计划、资源利用、任务并行、算法、IO、存算分离等方面与国际商用主流数据库产品仍有一定差距。关键技术能力欠缺不利于我国

数据库产品在核心系统的顺利落地应用。另外，核心系统数据库需要有强大的技术支持服务体系，以及时解决运行过程中出现的问题，目前我国数据库相对国际商用主流数据库在该方面也存在不足。

二是我国数据库产品灾备和数据同步水平还难以完全满足核心系统需求。根据核心业务系统多活容灾建设需求，我国数据库产品在异地 RPO、故障切换和灾难恢复时间等灾备能力与核心系统要求还有一定距离。同时，在采用新旧系统并轨运行模式，确保核心系统转型的平稳过渡，实现核心系统异地容灾，以及在核心系统数据库作为周边系统为数仓、数据中台、客服系统等提供实时数据时，都需要进行数据同步，这些对数据库的数据同步能力提出了很高要求。

三是核心系统数据库创新实施及运维管理难度大。首先，由于核心系统安全稳定运行是金融机构安全生产的关键，在数据库创新实践中选择新引入的数据库要慎之又慎，加之目前我国数据库产品与现有国际商用主流数据库产品还有一定差距，且数量多、技术路线各异、发展的不平衡等多种因素，给核心系统数据库选型带来很大压力。其次，核心系统已建设多年，在数据处理逻辑方面较多依赖数据库本身，通过数据库的存储过程、函数等实现业务规则中的复杂逻辑计算，而这部分程序可移植性差，一般需经过改造才可以在新的异构数据库上运行，这些改造在切流后容

易给系统带来功能和性能上的风险。再次，为确保业务连续性，核心系统数据库创新实践通常采取新老系统并行策略，相关的开发、测试涉及新老两套系统，为确保新老系统的功能一致性，极大增加了开发测试难度。此外，现有数据库中存量数据大，迁移工具有限，导致迁移周期长，且由于技术队伍的不稳定、掌握新技术产品的人员有限，进一步加大了核心系统数据库创新实施难度。最后，由于运维的整体自动化、智能化水平有限，不同类型技术产品还难以做到统一运维管理，核心系统引入新的数据库，特别是分布式数据库架构复杂、集群规模庞大、技术多样化、运维管理复杂，进一步增加了运维保障难度，为核心系统安全稳定运行带来隐患。

3.1.3 围绕关键环节推动核心系统数据库转型升级

一是加大产用联合创新力度，基于核心系统应用场景打磨优化产品，提升服务能力。数据库厂商在自我研发、创新同时，还应进一步加强与金融机构的联合创新力度，充分利用在核心应用中复杂、高压、苛刻的场景，对产品进行反复打磨和优化，补齐在核心技术、灾备能力、技术服务能力的短板，满足核心系统数据库创新实践需求。同时数据库厂商需要与操作系统，芯片等数据库依赖的关键软硬件生态厂商展加大容性适配，并进一步依托我国在存储、计算、网络方面已经取得的产业优势和成熟经验，联合创新，实现数据库产品性能、可用性、集约性的突破。例如

在高可用高可靠方面,通过主备架构、多主架构、共享存储集群、分布式集群等架构,实现数据库的高可用性和容灾能力,设置自动故障切换和数据同步机制,确保系统在发生故障时能够快速切换到备用节点,并提供持续的服务。此外,产业侧需要提升数据库服务能力,不仅要提升原厂服务能力,同时也要通过原厂的培训加大数据库认证学习,为用户团队培养更多的数据库人才。

二是安全与应用并重,合理选型,高效推动核心系统数据库创新实践。由于金融业核心系统属于国家关键信息基础设施,核心系统在选择新引入数据库时,首先要考虑金融安全问题,其次考虑业务需求、自身技术实力、成本及数据库产品特性,综合评估后进行合理选型。针对传统核心系统与数据库采取的紧耦合模式带来的改造难度风险,核心系统数据库创新实践需秉承多品牌的策略,避免形成对单一数据库技术和产品的绑定。同时,为了提高应用研发和改造效率,降低后续产品收敛升级改造的成本,需要在技术路线、研发标准方面进行统一约束规范,尽量选择对业务侵入性比较小的数据库,最大限度实现应用与数据库解耦。

三是加大生态工具的研发应用,支持核心系统数据库创新实践。在数据库迁移方面,梳理核心系统常用数据库的特性,分析新引入数据库的实现差异,开发自动化迁移工具,周密设计迁移规则和数据库处理逻辑的对等转换模式,对于源数据库的复杂特性和巨量存储过程提供自动化迁移能力,以降低迁移改造人工成

本。在数据同步复制方面，优化异构数据库增量数据复制工具，在新老系统并行阶段，通过数据同步工具进行业务高峰期增量归档数据在异构数据库间的双向复制，实现新老系统业务数据的准实时一致，确保故障场景下能及时回切，提升对外服务的连续性。在数据库开发方面，通过研发友好的功能强大的客户端工具支持数据库设计、开发，提升数据库开发者的工作效率。在测试方面，研发覆盖单元测试、功能测试、性能测试、生产验证和测试管理过程的自动化测试工具链，降低测试人力投入和测试复杂度，提升测试效率。在数据库运维方面，通过完备的数据库运行状态检查、监控、备份、恢复接口和工具，为核心系统稳定运行、数据安全提供保障。

四是加强核心系统数据库创新实践的统筹管理、经验总结，逐步形成指导行业实践的方法论。金融机构应将核心系统数据库创新实践，纳入整体“一把手”重大工程规划中，统筹推进、重点关注，确保资源投入、精心设计、稳步实施。同时建立并行运行机制，明确专业分工，将基础技术和应用技术分开，安排专人进行版本管理和程序发布工作，减少核心业务应用系统开发人员转型难度。另外，在进行核心系统数据库升级优化时，金融机构要充分总结经验，编写部署方案、技术方案、数据库迁移技术指引、数据库迁移测试白皮书、各类工具使用手册等涵盖数据库升级优化全过程的指导手册，形成整套的系统性技术资产、解决方案和方法论，以指导后续核心系统数据库全面升级优化。

3.2 数据库支持数据能力建设分析

在数字化转型发展中，金融机构需要对海量、多维的数据进行存储管理、分析，挖掘有效信息支持业务发展和经营决策，离不开数据仓库、非关系型数据库的底层技术支撑，从而对数据仓库的迭代升级、非关系型数据库应用探索创新提出了更高要求。因此，需要总结不同类型数据库支持数据能力建设的现状，分析创新发展需求，并提出下一步发展建议。

3.2.1 不同类型数据库助力数据能力建设取得积极成效

一是数据仓库高效支持金融机构数字化经营。由于数据仓库可对异构源数据进行有效集成，面向数据分析场景，支持全局信息共享和决策分析处理，从而受到金融机构的青睐。金融机构通过建设数据仓库进一步加强对不同业务部门数据的集中存储管理，为数据存储、处理、质量管控和安全管控等提供支持，满足快速增长的海量数据处理、高可用、弹性扩展、动态负载管理、融合分析等应用需求，并基于相关算法、模型、工具，支持数据价值挖掘、分析应用，在精细化客户管理、风险管理、精准营销、智能化决策等不同业务场景中发挥高效作用。

二是非关系型数据库为金融机构数据能力建设提供新助力。近年来，非关系型数据库在海量、复杂关系、多维的数据管理和分析处理中因为查询速度快、高可用、高可扩展性等优势，在金

融业实现了较快的创新应用。其中键值数据库因数据类型丰富、高吞吐量、低时延而被大家熟悉，在海量并发场景可为业务提供极致的访问体验。图数据库由于能更好表达数据之间的复杂关联关系，构建包含实体、事件、概念之间关系的图模型，进行复杂的路径分析、社区发现、链接预测等，相比关系数据库能够更好地展现和处理这类数据，在反欺诈、合规风控、投资信贷决策等场景进行了较多应用探索。向量数据库针对机器学习和深度学习中数据的向量表示形式，能够实现快速查找和检索，在量化交易、智能风控，个性化投资组合管理以及智能客服、数字人、情感分析、新闻事件挖掘等自然语言处理的不同金融场景中开展应用探索。总体来看，非关系数据库在海量文件存储、灵活快速查询、大数据统计分析、分析决策报表、实时分析、高速缓存、影像图片数据管理、指标标签管理等领域发挥越来越重要的作用。

3.2.2 金融业数字化快速发展对数据库提出新需求

一是金融业数字化转型深入推进对数据仓库的功能、性能、扩展性和安全性提出更高要求。在功能方面，要求承载数据仓库的数据库系统要支持更大规模的数据存储管理、服务时效性、混合负载能力等。其中存储范围上要容纳海量的内外部结构化数据，还要支持对数据湖等系统中存储的非结构化数据和半结构化数据进行有效管理和高效访问。数据服务要支持批量服务和实时服务，特别是对分布式架构下 OLTP 与 OLAP 业务融合越来越多的场

景，对混合负载（HTAP）能力提出更高要求。在性能方面，金融企业级数据仓库处理的数据量巨大、查询条件复杂、服务的业务类型和业务人员众多，对查询、响应效率，高并发、批量加工作业时间窗口等提出了更高的要求。在扩展性方面，随着金融数据量不断爆发式增长，数据的存储、计算需求会随之快速增长，是否具备便捷的扩展、伸缩能力成为金融机构对数据仓库的刚性需求。在安全性方面，要具备数据丢失或损坏时的恢复能力、对敏感数据的保护能力、以及对人为有意或无意的误操作的隔离能力，确保数据和系统的安全。

二是数字金融快速发展对非关系型数据库提出更多需求。随着数字金融快速发展，数据规模迅速增长，金融机构对海量数据的深度分析、事务间的复杂关联分析、数据随时间的变化分析越发重要；人工智能和深度学习技术和应用的迅速发展，使科学计算中的高维向量数据、影音/图片/文档等多媒体的非结构化数据大幅度增加，存储管理这些多模态信息需求也快速增加，都对非关系型数据库提出更多的应用需求。其中，键值数据库需要解决海量并发中热 key 带来的性能挑战，权衡性能和数据一致性的影响，进而提供满足不同业务场景的高性能方案。图数据库需要适应目前金融业应用中标准化程度不高、复制性不强的实际，进行灵活创新应用；同时，不同应用场景对不同架构的图数据库需求差异较大，要求图数据库具备集中式处理和分布式加工两类处理模式的优势，以适应复杂的业务需求。而向量数据库要降低处理

金融数据高维特征对性能带来的影响、平衡成本与性能要求，满足数据处理和查询的高实时性要求，对敏感数据进行安全防护，确保数据安全和隐私保护。

3.2.3 分类推动各类数据库应用创新促进数据能力提升

一是加大数据仓库技术创新，不断提升金融机构数据能力。通过利用 MPP 架构、列存储、智能索引、向量化计算等多种技术，提升在大数据量、多表关联复杂计算的能力，提升数据吞吐量和查询计算效率，减少业务决策的停顿等待时间，优化查询能力。利用湖仓一体架构、存算分离架构，满足结构化、非结构化数据存储和计算的多源融合需求，打通多种数据库之间的壁垒，支持构建统一的数据分析平台，满足大数据量、高并发的数据查询请求，为不同的业务弹性分配所需算力，提升数据吞吐量、并发能力。

二是利用 HTAP 技术助力混合负载类业务系统建设。OLTP 与 OLAP 并存是金融业应用系统常见的场景。在传统 OLTP 类型数据库中，虽能保证高并发读写下的数据强一致，但是在多表关联、大数据量下的数据分析场景中表现稍显薄弱，尤其是在分布式数据库场景下。通过在 OLTP 分布式数据库上发展的 HTAP 关键技术，实现一套引擎同时支撑业务系统运行和分析决策场景，避免在传统架构中，在线与离线数据库之间大量的数据交互，为混合负载

场景的应用系统开发提供了便利，大幅提升面向复杂查询场景的处理能力。

三是分类推动不同非关系型数据库应用发展。目前键值数据库在金融业的应用较多，在应用时可针对不同的数据规模和业务场景，合理选择分布式集群和读写分离架构，以满足海量并发场景的处理能力和业务访问体验需求，同时加大键值数据库的高可用架构建设，为业务稳定、连续运行提供强有力的保障。对正在快速应用图数据库，注重高效的图数据处理能力、大规模图数据分析能力、可视化和全生命周期的管理能力的提升；使用标准的、符合未来发展方向的图查询语言，采用支持 ISO GQL 标准语言的图数据库产品，提升图数据库的标准化水平；针对不同阶段和业务场景，合理选择单机架构或分布式架构图数据库，进行合理的资源投入和架构设计。对于向量数据库，针对高维数据，选择合适的向量索引方法，以优化数据的查询性能；对于高维向量数据，进行必要的降维或特征选择，以减少数据存储和处理的复杂性，并提高数据处理效率；考虑使用并行计算、分布式集群部署、压缩技术，以满足金融数据大规模处理和实时查询的需求，减少高维向量数据的存储成本。

3.3 开源数据库应用风险分析

经过多年发展、应用，开源数据库在金融业广泛应用于各类业务、管理系统，已成为金融业信息系统的重要组成部分。随着

关键信息基础设施领域数据安全问题被提升到国家安全战略高度，在金融业应用国外开源软件存在安全隐患的问题已经逐步得到重视。目前，开源数据库应用除了通用的开源协议风险、安全漏洞风险、知识产权及代码感染风险外，还面临新的开源停用、断供风险，政策的不确定性风险及掌控和服务能力不足风险等，需要全面分析，加强防范。

3.3.1 开源数据库快速发展并在金融业得到广泛应用

上世纪90年代，随着MySQL 1.0版本和PostgreSQL的Stable版本的发布，开启了快速发展和广泛应用的潮流和趋势。随即国内开源数据库也迅速跟进，阿里、腾讯和华为等分别推出了国内开源数据库，并逐渐成为国内互联网行业的主流数据库。而TiDB、OceanBase、IvorySQL、PolarDB等项目相继推出，越来越多的企业和机构开始应用开源数据库。

我国金融业积极拥抱开源技术，纷纷加入主要开源组织，积极参与开源社区互动，并主动组织实施了多个有影响力的开源项目，取得良好效果。其中MySQL应用最为广泛和深入，应用生态也较为完善。以邮储银行为代表PostgreSQL应用也取得较好效果。近年来，华为、阿里、蚂蚁、腾讯、平凯星辰等我国主要数据库厂商推出的开源数据库在金融行业的应用力度逐步加大，为金融机构有效防控单一供应商风险和日益复杂的供应链安全风险提供了选择。

3.3.2 金融业开源数据库面临诸多风险

一是**开源协议风险**。开源产品通常需要根据特定的开源协议进行发布和分发。不同开源协议的要求和限制差别较大，包括但不限于使用、修改、复制和分发等，给金融机构的选择、使用和分发带来困扰。而且开源产品可能嵌套其它开源产品的部分或全部代码，而这些开源产品可能遵循不同的开源协议，对于用户而言，这是一个极复杂而且隐蔽的风险。同时，不同开源许可协议之间可能存在混合或者不兼容情况，开源协议可以由开源项目背后的开源基金会或者商业机构根据其商业意图进行转换，例如知名数据库厂商 mongodb 将 AGPL 协议变更为更为严格的 SSPL 协议，在其他一些开源项目中还可以观察到由开源直接转为闭源的事件发生，开源许可证的变更也为金融机构判断评估适用性增加了难度。违反许可证的条款造成违约行为，易导致金融机构面临知识产权侵权等风险，可能会被权利所有人提起专利诉讼并收取费用，影响较大的案例可能导致金融机构商誉受损。

二是**安全漏洞风险**。安全漏洞的发现、披露和解决速度是衡量软件安全的一个重要指标。开源数据库是由世界范围内的个人自愿贡献代码，由于错误的代码实现或设计缺陷、开源社区维护度不足、使用者未能及时更新软件版本等原因，可能存在漏洞和安全隐患，导致金融机构在使用开源数据库过程中面临数据泄露、身份验证问题、拒绝服务攻击等多种安全威胁。开源数据库往往

不会经历太多实际的业务性测试，遇到问题时可能无法得到及时的修复，解决风险需花费大量的时间。对于金融业来说，特别是对于很多中小金融机构，内部开源技术治理还不够成熟，缺乏配套的专业人员和工具，而开源数据库已经将源代码尤其是内核代码开放，对黑客而言开源数据库毫无秘密可守，容易遇到恶意攻击事件，引发数据泄漏问题，无法从根本上保障金融数据安全。

三是知识产权侵权、代码感染风险。开源数据库细分为两种：一种是源代码由国内数据库厂商完全自研后选择开源模式进行市场推广；一种是国内数据库厂商封装了国外开源数据库内核代码，在此基础上二次开发的代码必须遵守国外开源协议限制，将二次开发的源代码公开，事实上出现了代码感染。虽然这两种数据库都称之为开源数据库，但实际上存在一定差异。前者具备完整的数据库知识产权，是掌握数据库核心技术和知识产权后，选择了开源的商业模式。后者则会被国外开源软件体系限制，存在知识产权、代码感染风险，应用到金融机构关键业务系统可能会埋下安全隐患，引发数据安全问题。

四是开源停服、断供风险突出。当前，开源软件供应链形势愈加复杂和多样化。一旦开源数据库项目停止开发和维护，金融机构将无法升级数据库系统和获取新功能，亦无法得到此开源数据库相关的技术支持，甚至可能因为缺乏维护导致安全漏洞无法及时修复，会严重影响存量信息系统的运行维护，并带来大量的

系统更新需求和额外的成本及风险。目前，接近 90%的金融机构在一般业务和管理类系统应用使用了开源 MySQL，普遍面临 MySQL 5.7 版本的停服问题，导致金融机构被迫转型升级至更高版本，面临较大的升级改造压力。此外，开源数据库供应链风险不断增大，一旦国家之间的贸易管制政策蔓延到开源项目，金融机构托管在海外的开源代码资产将面临冻结风险，也会面临复杂的司法纠纷等问题。

五是政策不确定性风险。相对闭源的专利技术，开源数据库难以通过传统的专利和知识产权保护手段来证明自己的创新性和价值，使得开源数据库在获得资金、政策支持和竞争优势方面受到限制，也导致开源数据库在认定中面临难以评估和认定的困境。目前，金融机构大量使用开源数据库，开源数据库能否认定为符合政策要求对金融机构在数据库产品技术选型、升级迭代策略等方面具有较大影响，导致金融机构在数据库领域的规划布局存在较大的不确定性，也面临较大的技术路线选择纠错风险。

六是掌控和服务能力不足风险。不同于商业闭源数据库可提供友好且易于使用的界面和工具，开源数据库需要金融机构具备较强的掌控能力，开发运维团队需要理解和掌握数据库系统的原理、架构、配置和管理等方面的知识，具备一定的自主解决问题能力，才能确保开源数据库的顺利应用，而大量中小金融机构显然难以满足掌控能力要求。同时，对开源社区的贡献往往属于开

发者的个人行为，并非持续性的机构行为，绝大部分机构依然侧重于关注软件的技术研发维护能力，缺乏主动参与开源工作和贡献开源生态的意识和动力，导致金融机构面对开源数据库缺陷与问题时大多处于被动的局面，未能形成良好的产用双方技术共建能力。另外，相比商业数据库提供的全面技术支持和售后服务，开源数据库服务通常依赖于社区支持和开发者社区的帮助，而社区响应的及时性和质量并不一定能够满足金融机构的需求。

3.3.3 多措并举防范开源数据库应用风险

一是加大开源数据库治理，降低开源数据库使用风险。金融机构需要设置专人进行开源数据库治理，对开源数据库进行合理合法管控，确保合规使用开源数据库。在选择开源数据库时，金融机构首先要考虑其是否安全可控，然后全面考虑其版本稳定性、社区活跃度和更新频度，对相关工具和插件等开展研究，判断开源数据库项目的可持续性和发展前景，积极发展和应用自主开源项目，避免停机、断供带来的风险。同时，在引入开源数据库前，金融机构应进行充分测试和评估，通过搭建测试环境，模拟实际应用场景，进行全场景、全量、重压的测评，评估数据库的性能、稳定性和可靠性，减少由于技术局限性和不确定性带来的风险。另外，配备专业的法律团队或与外部专业法律机构合作，对开源数据库的开源协议进行分析、提供专业的法律支持，也可以考虑与开源软件社区或相应的权威机构进行合作，以获取有关开源协

议和风险管理的专业建议和支持，从源头上降低可能存在的知识产权风险。

二是提升开源数据库安全防范水平，确保安全生产。金融机构应加快建立软件成分清单生成与使用规范，标准化软件成分和软件成分可视化流程，保证组件透明理念的落实。同时，金融机构应配备专业的安全检测工具、漏洞扫描工具和安全人员，对开源数据库进行定期检查，发现潜在的漏洞和安全风险，并进行妥当评估、处置和回顾，结合安全可信白名单机制、风险预警与情报收集机制，保证内部环境安全。充分利用第三方代码检测和第三方安全评估机构的力量，获得开源数据库代码自主率、独立的安全建议和解决方案。通过使用强密码、多因素身份验证和访问控制列表等方法，限制对开源数据库的访问权限，避免未经授权的访问。通过安全教育和培训，提高开发人员和管理员的安全意识，了解并学习常见的安全威胁、攻击方式及防范措施，提高对安全问题的警惕性。

三是推动金融开源数据库社区建设，全面提升对开源数据库的掌控和服务能力。在行业主管部门的指导下，由权威的第三方行业组织推动建设金融开源数据库社区，加强产学研用多方协作，充分调动行业机构力量，引导生态链各方积极参与开源社区。同时，依托社区对开源数据库进行测试评估、漏洞风险通告，组织开展开源数据库的技术交流、研讨，并通过开源社区博客、视频

教程、在线文档等资源，提升金融机构对开源数据库的掌控能力。另外，要特别重视鼓励金融机构积极参与开源社区的活动，如报告漏洞、提出建议、参与开发、与开发者合作解决问题以及贡献代码等，获取更多有关开源数据库的信息，助力金融机构开源数据库技术人才培养，并获取社区对开源数据库使用的更好服务。

4. 金融业数据库应用展望

未来三至五年是金融业数据库创新发展的关键时期，也是高成长时期，无论从广度、还是深度上都将迎来巨大的突破和发展，我国数据库产品也将随着大量应用验证不断成熟，实现从“可用”向“好用”转变。随着技术演进变化，数字化、智能化普及，不同类型数据库应用创新力度不断加大，数据库领域基于分布式、云原生、人工智能等技术加持，进一步加快能力提升，更高效支持“数字金融”发展。

4.1 我国主流数据库产品成熟度将进一步提升

随着核心系统数据库转型加速，金融机构会持续加大对数据库产品的投入力度，为产业侧提供开发和优化数据库产品的动力和资源，尽快弥补我国数据库产品在功能、性能、异地灾备、数据同步等核心技术能力的差距，丰富配套工具，降低实施运维难度，推动我国数据库产品的成熟和稳定。同时，基于核心系统特性及其他非数据库技术要求，也将倒逼数据库服务能力提升，支

持更为便捷、高效的数据库应用开发、运维及安全防护，推动我国数据库由“能用”向“好用、易用”转变。另外，基于丰富的金融数据库应用实践，不断完善金融数据库应用相关标准规范，提升产品的标准化水平和快速复制能力，适应金融业快速发展的需要。

4.2 金融核心系统数据库转型升级将稳步推进

金融业数据库转型模式由政策驱动转向自觉行动。在主管部门的直接指导、推动下，金融业完成了多轮转型试点，实现了从“0”到“1”的突破。且随着国际形势愈加的复杂多变，全行业实现了从上至下理念的转变、认识的提升，确保关键软硬件技术供应链安全稳定已成为上下共识。经过近几年积累，我国数据库技术产品供给能力明显提升，应用生态不断完善，大量试点应用为金融核心系统数据库转型提供了宝贵的可借鉴经验，金融机构必将直面核心系统数据库转型的“硬骨头”。同时，主管部门将根据需要适时进行指导，提供数据库等关键软硬件技术应用的政策支持。在各方的共同推动下，我国数据库产品必将在金融核心系统实现快速深入应用、突破发展，切实提升安全可控水平。

4.3 新技术与业务场景将推动金融数据库创新发展

随着数字化转型深入推进，金融数据库创新将迎来新一轮高潮。数据仓库、数据湖、大数据等技术融合应用创新，不断提高

支持数据能力建设的水平。不同类型的非关系型数据库加快应用创新，适应数字经济时代海量、多维数据处理及不同细分场景的应用要求。同时，数据库也将逐步从分布式向云原生转型，提供超大并发支持能力和更强大弹性伸缩能力。基于人工智能的数据库 AI 优化器、AI 分析引擎，AI 自治运维系统等，将全面提升数据库的查询、交易处理速度，提供更智能、准确的业务洞察和风险评估能力，实现高效的自动化数据库运维。

4.4 金融业开源数据库应用风险防范仍然不容忽视

针对多年积累的存量开源数据库，金融机构还需进一步摸清家底，形成开源数据库底层和源头清单，对开源数据库进行合规治理，根据自身情况适时考虑是否收缩技术栈。同时，重点关注断供、停服风险，并持续加强协议风险、安全漏洞风险防范，确保安全生产。对于开源数据库使用的政策不确定性、掌控和服务能力不足的风险，需要加强政产学研用的合作力度，充分交流研讨，多方协同，形成应用共识，并通过加大开源社区、应用生态建设，确保开源数据库在金融业的顺利应用。

随着管理部门的大力支持和推动，行业标准规范的及时跟进和发布，数据库产业的巨大投入和快速迭代，金融机构的积极响应和快速推广，科研院所对数据库人才的培养，以及政产学研用联合攻关，将推动金融业数据库创新向深入发展。行则必至，做则必成，金融业数据库创新发展只待时日、花开有时。

附录：金融业数据库优秀应用案例集

一、政策性银行及开发性金融机构.....	1
案例一：国家开发银行基于 CirroData 和 BEH 的大数据平台建设应用实践.....	1
二、国有大型银行.....	5
案例二：工商银行分布式数据库 GaussDB 应用实践.....	5
案例三：农业银行大数据平台 GBase 8a 应用实践.....	13
案例四：中国银行数据仓库 GBase 8a 应用实践.....	16
案例五：建设银行核心系统 GoldenDB 应用实践.....	16
案例六：交通银行核心系统 OceanBase 应用实践.....	23
案例七：邮储银行新一代个人业务核心系统 GaussDB 应用实践.....	26
三、股份制银行.....	29
案例八：平安银行信用卡核心系统 TDSQL 应用实践.....	29
案例九：光大银行新一代信用卡核心系统 GoldenDB 应用实践.....	33
四、城市商业银行.....	36
案例十：北京银行分布式核心系统 TiDB 应用实践.....	36
案例十一：北京银行分布式核心核算引擎 OceanBase 应用实践.....	39
案例十二：北京银行基于 CirroData 和 BEH 的湖仓一体建设应用实践.....	41
案例十三：贵阳银行信用卡管理平台 SUNDB 应用实践.....	46
案例十四：江西银行企业网银系统 SUNDB 应用实践.....	50
案例十五：湖北银行新核心业务系统中达梦柔性迁移双活应用实践.....	55
案例十六：四川银行金融核心业务反洗钱系统 GBase 8s 应用实践.....	58
案例十七：秦皇岛银行分布式核心系统 TDSQL 应用实践.....	61
案例十八：梅州客商银行新核心业务系统中达梦高容灾可用性应用实践.....	65
案例十九：某银行核心业务领域 GBase 8c 多模多态分布式数据库应用实践.....	68

五、农村商业银行	72
案例二十：广东农信数据库联合实验室企业网银和银企直连 SUNDB 应用实践.....	72
案例二十一：浙江农商银行“丰收互联”TDSQL 存算分离应用实践..	75
六、证券机构	79
案例二十二：中信建投 OTC 核心业务系统中达梦集中式 HTAP 应用实践.....	79
案例二十三：国泰君安证券新一代分布式核心交易系统 GoldenDB 应用实践.....	82
案例二十四：申万宏源证券阿里云数据仓库 AnalyticDB 应用实践...	86
七、保险机构	90
案例二十五：中国人寿数据库迁移实践.....	90
案例二十六：中国太平洋保险核心系统 OceanBase 应用实践.....	96

金融信息

一、政策性银行及开发性金融机构

案例一：国家开发银行基于 CirroData 和 BEH 的大数据平台建设应用实践

一、应用场景

国家开发银行在分析系统领域使用了 TD、Oracle 等多种数据库，来满足各类场景要求，但在很多场景中仍受限较多，要么功能不能满足，要么性能不能满足，为此经过多个实际应用系统的实用验证，最终采用东方国信的 CirroData 和 BEH 构建了全行级的大数据平台，支撑全行的数据应用。应用场景的示范举例：

- 数据整合：基于大数据平台建设实现业务历史数据的集中和历史保留，并能方便地提供查询；
- 数据查询：查询平台从 TeraData 平滑迁移到 CirroData，并提升了查询效率；
- 高性能处理：主动探索系统基于传统技术和大数据技术无法达标，通过 CirroData 实现业务需要功能，且做到 3 秒返回结果；
- 功能提升：支持百亿级记录表快速、关联查询，使查询效率低、无法查询出结果的性能获得显著提高，可在分钟级获得查询结果；

目前，累计部署服务器超过 300 台，有效存储空间超过 10PB，

且使用常规的高性价比服务器，综合成本得到大幅降低。

二、总体方案

大数据平台整体规划如下图所示：

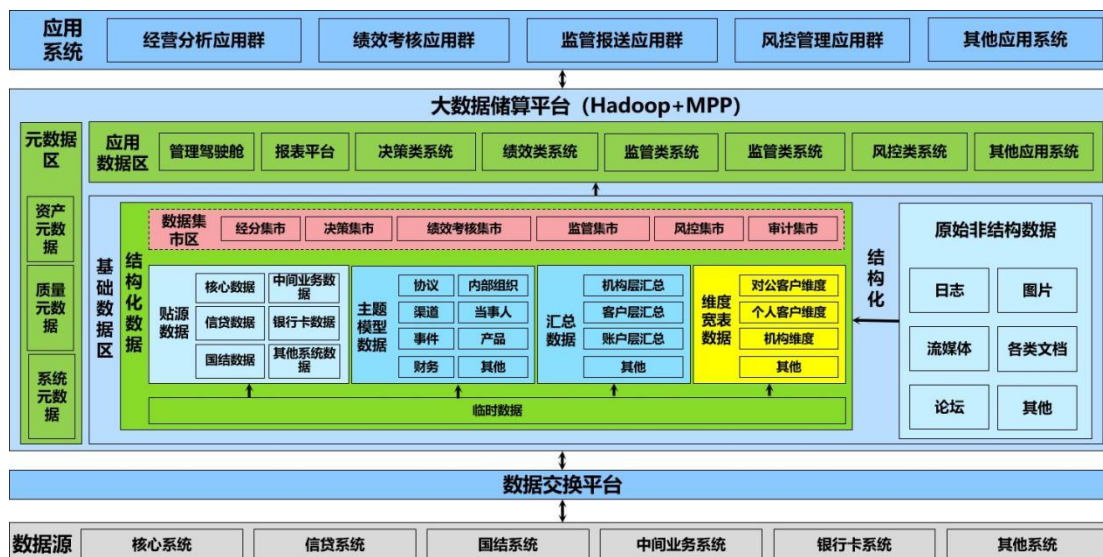


图 1 数据架构图

整个计算平台分为三部分：

1. 首先是基础数据区，里面包含了结构化数据和非结构化数据。结构化数据中包含了临时数据区，为全平台中统一的临时数据；贴源数据区为按系统保留的每日的原始数据；主题模型数据区为传统数据仓库的模型数据；汇总数据区为传统数据仓库的汇总层数据；维度宽表数据区为统一的企业级维度模型数据。贴源数据、主题模型数据、汇总数据、维度宽表数据是全行级的公共数据。其上面数据集市区存储多个数据集市，存储面向某个业务条线或业务板块公共的数据。这些构成了全行的基础数据。

2. 基础数据区上面是应用数据区，里面存放应用系统专用的数据。

3. 元数据区，存放数据平台的统一的元数据信息。

在技术上，结构化数据存储统一采用新一代 MPP 数据库存储，非结构化数据目前采用大数据技术存储，即 HDFS + ES + Hbase 等组件进行存储，一般大文件可直接放到 HDFS 上，小文件建议存到 Hbase 中，未来还可引入 OSS 对象存储。而数据的采集、处理和特定计算可依托 Hadoop 生态组件实现，从而实现新一代 MPP 数据库和 Hadoop 生态的有机结合。

三、难点问题及应对举措

该银行采用新一代 MPP 数据库 CirroData 和 Hadoop 发行版 BEH 产品，并在此基础上开发了全套完整配套产品和解决方案，解决了自身的如下痛点问题：

- 统一部署：CirroData 和 BEH 可以统一部署，使用一套分布式文件系统，可以按需部署，避免了传统的数据库和 Hadoop 必须独立部署，数据底层共享，大大减少处理环节，增加数据共享能力；
- 集中统一存储：避免了企业内数据在多个组件中冗余存放，以解决不同场景问题，结构化数据主要存储在 CirroData 中，非结构化数据主要存储在大数据组件中，在企业内形成标准化规范，并支持大部分数据应用场景，大大简化了企业内数据管理的复杂度和冗余存放；
- 开发统一化：由于大部分操作尤其 90%以上是对结构化数据处理，均可基于 CirroData 产品完成，因此大部分开发

人员只需掌握这一个产品，且以 SQL 操作为主，大大降低企业的学习成本；

- 完善迁移工具：基于 CirroData 和 BEH 开发了数据迁移、数据核对、程序迁移、调度迁移等配套产品，实现应用系统的快速迁移；
- 成本降低：CirroData 支持 X86 服务器和多种我国设备产品，并有多个实际案例，整体使用成本大大降低；
- 配套解决方案：除了产品以外，东方国信湖仓一体解决方案包含咨询规划能力、专业开发实施、专业运维服务和配套产品，提供全面服务，保证了湖仓一体平台的持续建设。

四、应用成效及经验

从使用 TeraData、Oracle 等多个传统数据库产品，到集中使用了 BEH 和 CirroData 产品，建设了统一的储算平台，已经落地了数据湖区、仓库区、集市区、应用区、实时区、非结构化区等，在此基础上新建或迁移了主动探索等几十个相关应用系统。

大数据平台项目的建设逐步统一技术路线，标准化了开发流程和规范，提升了面向业务的响应时效，满足了之前不能实现或实现困难的系统和功能。

二、国有大型银行

案例二：工商银行分布式数据库 GaussDB 应用实践

一、应用场景

工商银行经过技术攻关与创新实践，充分验证了传统集中式数据库向分布式数据库转型的可行性，为大型商业银行核心系统安全可控转型走出了宽阔的道路。未来，工商银行将持续围绕分布式数据库的创新运用，协同华为持续开展联合创新和技术攻关，夯实工行数字基建基础支撑，形成金融行业传统集中式数据库转型最佳实践，通过技术沉淀和转型实践经验总结，识别行业共性需求，形成高效可控低成本的数据库平滑转型技术方案及配套工具，为金融同业提供转型的良好借鉴，助力中小金融机构加快自身数据库转型进程，共建金融科技新生态，推动金融业实现高水平科技自立自强。

二、总体方案

（一）项目实施模式

对于大型 Oracle 应用系统，一般会采用双轨并行的模式。具体实施方式如下：

1. 存量数据迁移阶段：采用 DRS 数据复制工具，完成存量数据迁移。

2. 技术验证阶段：以技术验证为主，将业务联机、批量双写，切流部分查询交易，观察 GaussDB 数据库运行情况。

双库并行期间将非双写的表做单向增量数据同步。

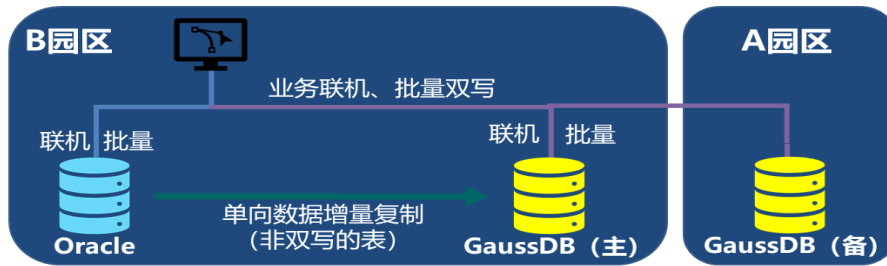


图 1 双库并行，数据增量同步

3. 全功能切流阶段：GaussDB 承接业务日常联机和批量数据处理，通过灰度引流，按试点维度（机构或客户）逐步推广，并可实现试点维度 GaussDB 和 Oracle 库快速互切。

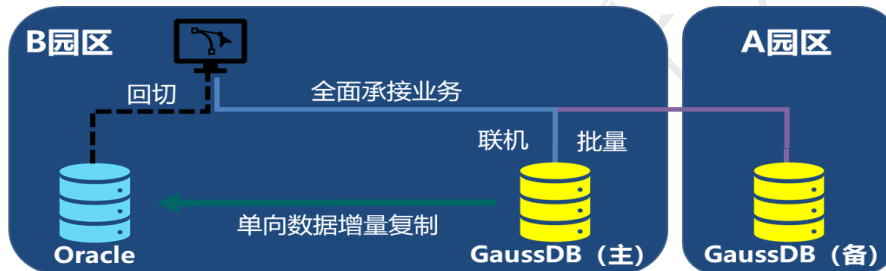


图 2 全面切流，GaussDB、Oracle 快速互切

4. 单轨运行阶段：在 GaussDB 全面承接业务并稳定运行半年后，将采用单轨运行的方式，并下线老的 Oracle 数据库。

应用采用 Java 拦截器特性进行双写。

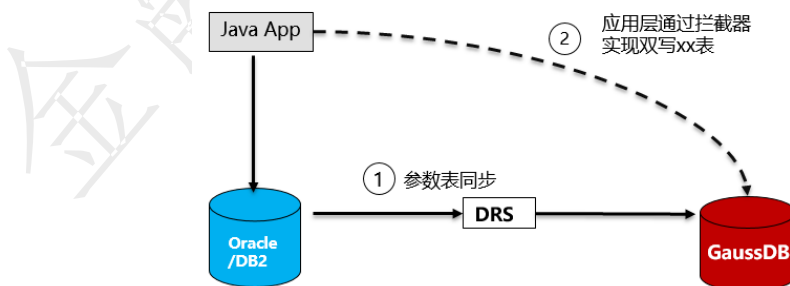


图 3 Java 拦截器双写

该方案利用 Java 拦截器特性，适用于 Java 类应用。应用侧需要改造做双写表，利用 DRS 继续同步应用相关参数表，由于 GaussDB 的数据有延迟，比对的数据以成功写入 GaussDB 的部分

为基准，反向与 DB2/Oracle 的数据进行校验。Java 拦截器 Interceptor 原理：拦截器是动态拦截 Action 调用的对象。它提供了一种机制使开发者可以定义在一个 Action 执行的前后执行的代码，也可以在一个 Action 执行前阻止其执行。应用通过拦截器，对业务的 Controller 请求进行拦截，执行双写逻辑。

（二）实施周期

整体项目周期约 5 个月，关键里程碑节点如下：



图 4 实施周期

（三）数据迁移

对原系统数据采用全量数据+增量数据迁移的策略。使用 DRS 数据复制服务。对于全量数据迁移，通过 select 的方式将源库的全量数据导出；调用 GaussDB 的 copy 接口批量写入目标库；可对单表根据主键进行进一步分片，每个分片作为独立的同步单元，并发执行同步效率 100MB/s。对于增量数据迁移，通过 Oracle LogMiner 解析数据库日志，DRS 将抓取到的数据落盘存储，读取落盘数据，重构成对应的 SQL 语句在目标库回放。

在数据迁移的过程中，主要遇到的问题有两个：1. 性能达不到性能要求。2. 增量迁移表需要有唯一索引。

针对第一个问题，一种方式是采用 LogMiner 的方式，但需要解决抓取 lob 不全，不支持无主键表等问题。

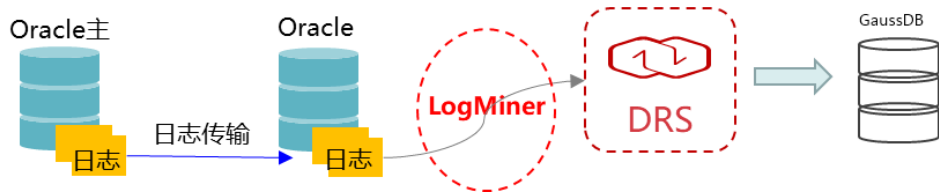


图 5 LogMiner 方案

一种方式是采用 Xstream 的方式，该方式需要 Oracle ogg license，在推广上有一定局限性。



图 6 Xstream 方案

还有一种是通过 CDC 到 kafka，再到 GaussDB 的方式，这种方式技术栈较复杂，并且链路较长，稳定性待验证。

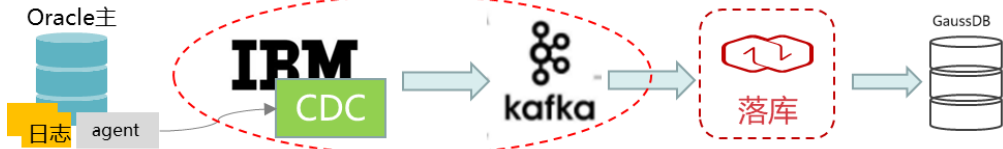


图 7 CDC、Kafka 方案

最终选择第一种方案进行实施，版本做了针对性的优化，满足同步性能。

针对第二个问题，通过增加主键或唯一键的表；早批文件加载临时表，可以不同步；纯技术备份表进行 DROP 三种方式进行解决。

三、难点问题及应对举措

从传统数据库往分布式数据库转型的过程中，面临如下痛点：

1. 高可用挑战大: 业界常用的数据库方案只能部署在跨园区的单 Region 中, 不能满足高等级应用的高可用要求。

2. 性能容量挑战大: 大型业务系统高峰期 TPS 近万/秒, 单库几十 T, 对基础设施及数据库的性能容量提出了较大的挑战。

3. 扩展性不足: 传统的集中式数据库由于架构限制, 往往采用一主一备, 即使是 Oracle RAC, 由于共享存储的限制, 也最多只能扩展 2 到 4 个节点, 单集群的容量容易出现瓶颈。

4. 安全性有待提升: 近年来随着大众对保护个人信息意识的提升, 以及各国家地区纷纷出台数据安全相关的政策法规, 如何保护数据安全隐私的问题受到重大的关注, 同时也是工业界与学术界研究的热点问题。

5. 数据同步挑战大: 大型业务系统高峰期超过几百 G/小时数据归档量, 对新旧系统双库并行期间的数据同步效率要求极高。

6. 数据库对象迁移成本高: Oracle 应用存量存储过程量大, 重构成本高, 特别是大型业务系统, 普遍使用多个 Oracle 库, 存储过程行数达到千万级, 业界常见的迁移工具的存储过程自动化迁移率 70%左右, 大型 Oracle 应用转型应用改造成本非常高。

可通过如下关键举措应对转型过程中面临的痛点问题:

1. 提升服务连续性: 对标主机 AB 站点双活方案, 对 GaussDB 承载大型业务系统 Oracle 数据库转型的部署方案开展联合创新, 实现同城双集群+Dorado 存算分离的部署架构。

2. 降低迁移开发成本: 建设完善的自动化迁移工具, 通过内

核兼容（或提供对等能力）+工具自动化转换，形成低成本、风险可控的应用迁移技术方案。

3. 提升测试效率：建设 GaussDB 自动化测试配套工具，支撑转型应用存量 Oracle 存储过程测试案例向 GaussDB 迁移并实现自动化测试，降低测试人力投入，提升测试效率。

4. 提升数据复制效率：优化数据复制工具，满足灰度切流双库并行期间大型业务系统高峰期的数据归档量的同步效率。

四、应用成效及经验

工商银行 Oracle 数据库转型项目，自 2019 年引入 GaussDB 以来，已在包括实物贵金属、中间业务系统等近 40 个业务系统试点上线，覆盖办公系统、一般业务系统和关键业务系统各类典型业务场景，初步形成一套涵盖工行主要商用交易型数据库（Oracle、DB2、SQL Server）的转型方案。

转型成效 1：联合创新，夯实数据库核心承载能力

一是数据可靠性高。基于全闪存集中存储实现计算与存储分离，软硬协同，具备 PB 级海量数据存储能力和企业级高可靠能力。二是系统可用性高。具备同城双园区+异地园区部署能力，园区内故障场景 RPO=0、RTO<60 秒，同城园区级故障场景 RPO=0、RTO<180 秒，同城双园区故障场景切换到异地园区 RPO<60 秒、RTO<600 秒。三是集群性能高。同城主备集群间采用磁盘级复制实现增量日志强同步，日志同步效率提升一倍以上，降低了主备同步对主集群性能的影响，基于 2 路国芯服务器最小规模主备集

群部署，TPMC 达到 68 万（约 2.5 万 TPS）。四是服务连续性能高。具备业务不中断前提下主备集群数据库版本轮换升级和应用版本灰度升级能力，满足了核心应用 7*24 小时服务连续性要求。

转型成效 2：应用创新，锤炼数据库平滑迁移能力

工商银行聚焦传统数据库与应用耦合度高的难点，联合华为进行突破，构建了整套的自动化工具链：

一是实现跨异构数据库的自动迁移。针对部分传统集中式数据库特有的数据库对象、高级特性和非标准 SQL 语法，通过建设配套的自动化数据库迁移工具，提前评估和规划迁移工作进程，识别迁移风险，再通过工具自动化进行语法转换和逻辑校验，降低迁移成本、控制迁移风险，快速低成本地实现从商业专用平台向开放创新平台的迁移，自动迁移成功率和编译通过率均可达 95% 以上。

二是实现异构数据库转型全过程的自动化测试。工商银行建设了覆盖单元测试、功能测试、性能测试和测试管理等研发测试全过程的自动化测试工具链，一方面通过 SQL 解析、分支预测等技术实现技术测试自动化，另一面复用存量业务测试资产实现业务功能测试自动化，整体自动化测试覆盖率可达 80%。

三是实现试运行阶段生产环境测试验证的系统级解决方案。研发交易录放工具，在试运行阶段，先在旧系统抓取流量，然后在新系统分别进行一致性流量回放和性能回放实现功能和性能

的验证：首先，一致性回放将抓取到的 SQL 按源库的执行顺序以事务为单位在新系统进行回放，实现业务功能全覆盖测试，保证新旧系统功能完全对等；第二，性能回放把抓取到的 SQL 按照一定规则分发进行多线程并发回放，以接近实际生产业务压力的速度回放到新系统，进行性能、可用性及可靠性测试，确保新旧系统可完整承载业务压力。

四是实现新旧系统并行阶段数据一致性的系统级解决方案。优化异构数据库增量数据复制工具，在双库并行阶段，新旧系统均有业务流量，通过数据复制工具进行业务高峰期增量归档数据在异构数据库间的双向复制，实现新旧系统业务数据的准实时一致，确保故障场景下能及时回切，提升对外服务的连续性。

转型成效 3：资产沉淀，形成普适性的一体化解决方案

工商银行协同华为公司组建涵盖数据库技术研究、研发、测试等方向跨开发中心多地的攻关团队，在传统集中式数据库转型实践中，充分总结经验，编写了转型部署方案、转型技术方案、数据库迁移技术指引、数据库迁移测试白皮书、各类工具使用手册等涵盖数据库转型全过程的指导手册，形成了整套的系统性技术资产和解决方案，开拓了传统集中式数据库转型工作的新思路、新方法，初步形成了一套无需整体重构 Oracle 存储过程逻辑，低成本、高效可控的转型技术方案。

案例三：农业银行大数据平台 GBase 8a 应用实践

一、应用场景

中国农业银行大数据平台是自主搭建的全行级数据采集、加工、服务一体化系统。数据来源于核心、信贷、财会、风险管控等近百个源系统，经过清洗转换、拼接整合、汇总加工后形成高价值密度的数据模型，面向全行各大业务领域和场景，提供大数据深度分析和挖掘，全面支撑资产负债、运营管理、风险管理、数据报送等各项业务领域应用，为业务提质增效提供有力支撑。

二、总体方案

中国农业银行大数据平台基于 GBase 8a MPP Cluster（以下简称 MPP 数据库）+ Hadoop 技术栈建立创新混搭式基础软件架构，自下而上分为数据处理层、模型指标层、数据集市层以及分析展示层。

分析展示层	数据报送 报表展示 AI分析 BI分析
数据集市层	MPP数据库 各业务领域数据集市
模型指标层	MPP数据库 加工 汇总 主备集群同步
数据处理层	Hadoop 抽取 清洗转换 拉链加工

图 1 大数据平台架构图

数据处理层使用 Hadoop 集群接收并处理各类业务源系统的结构化数据文件，完成数据的清洗转换以及拉链表的加工；模型

指标层、数据集市层采用 MPP 数据库构建，形成了企业级统一数据视图，面向特定业务领域提供定制化数据集合；各应用在分析展示层利用数据集合完成对数据的挖掘、分析和展示。MPP 数据库和 Hadoop 混搭式的总体架构，根据目标数据特点分层处理，充分发挥各自技术优势，高效应对海量数据的管理和流转。

三、难点问题及应对举措

（一）大规模数据需求加剧集群资源竞争。

采用 MPP 数据库虚拟集群技术对多个虚拟子集群进行统一管理，各虚拟集群在整个物理集群范围内独立运行，实现数据隔离、资源隔离、故障隔离；同时各虚拟集群共享统一的入口，支持跨虚拟集群数据访问和权限管理，实现多租户业务的整体调度，维护全局大规模数据视图。

（二）MPP 数据库在特定服务器上性能表现不及预期。

根据目标服务器 NUMA（Non-Uniform Memory Access，非一致性内存访问）架构特点，采用数据库计算实例与 NUMA 内核绑定技术，在单服务器上部署多个 MPP 数据库运行实例。多实例部署可充分发挥 NUMA 服务器的性能，降低大数据计算场景下跨 NUMA 节点内存访问带来的性能损耗，大幅提升大数据平台的计算性能。

（三）海量数据对面向业务连续性的容灾要求不断提升。

为配备高等级跨地域容灾能力，在基础环境方面，使用磁盘 RAID 技术、冗余电源、网卡绑定、多机架高可用等保障运行环

境稳定；在集群级高可用方面，创新性地提出和实现支持 PB 级海量数据增量同步的异步双活解决方案，并成功应用于系统灾备、机房搬迁等场景中，实现大数据平台同城机房间、异地机房间的超大规模数据同步、主备集群切换，保障业务稳定运行。

（四）内外部数据管理对数据安全保障提出更高要求。

平台利用 MPP 数据库安全功能特性，进行数据加密与安全管
理。一是对重点应用中高安全性要求的表，在建表时按需使用加
密关键字，对表级或列级数据进行加密；二是基于 MPP 数据库用
户认证和权限管理能力，结合行内用户安全管理基础设施，搭建
可视化工具对集群级、库级、表级、字段级等权限进行管控，实
现数据安全分级管理。

四、应用成效及经验

中国农业银行大数据平台已平稳运行近十年，系统部署架构、
高可用方案、数据安全方案等不断创新发展，为同业推进大数据
体系建设提供了有效借鉴。目前已部署 MPP 数据库集群近百套，
节点总规模超四千台，为总量达几十 PB 的海量数据提供管理和
服务；T+1 双活机制利用主集群闲暇时段进行数据高效同步，充
分保障了集群级高可用性，是金融机构应用 MPP 数据库进行海量
数据处理的优秀实践案例。

案例四：中国银行数据仓库 GBase 8a 应用实践

一、应用场景

中国银行以数字化转型为契机，构建“三横两纵一线”的新型数字资产运营服务体系，“三横”是搭建集团统一数据平台，以“数据+分析+展现”的三层架构，为数据资产的共享、分析应用、服务提供和价值创造提供全面、敏捷、精细的能力支撑，数据仓库是数据层的重要组成部分。

中国银行使用 GBase 8a MPP Cluster 集群承载数据仓库、风险数据集市、审计系统、SAS 模型管理平台、监控标准化数据报送平台、报表系统等数十套应用，累计部署千余台服务器，管理数十 PB 数据量。中国银行在 2020 年启动 GBase 8a MPP Cluster 环境的全面改造，包括硬件服务器、操作系统等。

二、总体方案

中国银行 GBase 8a MPP Cluster 环境的创新改造分为新建环境和已有环境的迁移改造。

首期全栈创新改造基于新建的数据湖仓项目进行，使用 GBase 8a MPP Cluster 作为数据仓库面向海量数据分析型应用领域，进行数据复杂分析和作业加工；Hadoop 集群作为数据湖负责数据汇聚、贴源数据清洗和沉淀历史数据资产。数据湖仓项目中共部署 GBase 8a MPP Cluster 集群近 400 节点，使用 190 多台海光芯片服务器和麒麟操作系统，采用多实例部署方式每台

海光芯片服务器上部署两个 GBase 8a 计算节点实例。

对基于 Intel 服务器的 GBase 8a MPP Cluster 集群，通过服务器迁移、集群升级方式完成创新应用环境的升级。将 GBase 8a MPP Cluster 的数据文件通过备份恢复方式迁移到海光芯片服务器上，之后对海光芯片服务器上的 GBase 8a MPP Cluster 集群进行版本升级和创新应用平台优化配置。

三、难点问题及应对举措

中国银行创新改造过程中面临如下几方面的问题，通过银行与 GBASE 厂商、芯片及服务器厂商、操作系统厂商紧密合作制定解决方案并最终成功解决问题。

1. 自主可控芯片单核计算性能低，GBase 8a MPP Cluster 对 SQL 执行的每个阶段都可以多线程并行，充分发挥海光等芯片的多核特性提升性能。

2. NUMA 架构服务器跨 NUMA Node 内存访问性能损耗大，采用多实例部署方式，在一个海光芯片服务器上部署两个计算节点实例，每个节点实例绑定一组相邻的 NUMA Node，降低跨 NUMA Node 内存访问带来的性能损耗，同时提高了服务器资源的利用率；在每个芯片服务器上部署两个计算实例，达到与 Intel 服务器同等的性能。

四、应用成效及经验

GBase 8a MPP Cluster 助力中国银行完成了全部 Intel 服务器环境下 GBase 8a MPP Cluster 集群环境的全栈创新升级。

通过产品的适配优化，确保基于创新应用平台的大数据查询分析性能不下降。

1. 提升性能，实现我国芯片平台与 Intel 服务器的一比一创新升级

GBase 8a MPP Cluster 针对创新应用平台进行了全面的优化改造，通过多实例部署、编译参数优化、硬件加速优化提升 GBase 8a 在 NUMA 架构的自主可控芯片服务器整体性能，相比于优化前性能大幅度提升。

2. 提升适配性，满足全栈创新应用平台要求和异构混搭需求

GBase 8a MPP Cluster 实现了对全栈创新应用平台的适配，支持海光芯片、鲲鹏芯片、飞腾芯片、申威芯片等服务器，支持银河麒麟、统信 UOS、中科方德等操作系统；支持异构平台混搭，满足用户对于多种异构的自主可控服务器的使用需求、Intel 服务器利旧的使用需求、异构平台双活的需求。

3. 满足高可用能力要求

GBase 8a MPP Cluster 的管理集群、数据集群都采用多活架构，无单点故障；支持在线节点替换，优化在线节点替换性能大幅度提升；在创新应用环境建设初期，偶发的服务器故障不影响集群运行，不中断业务，故障节点处理更平滑。

案例五：建设银行核心系统 GoldenDB 应用实践

一、应用场景

建设银行现有对私核心采用 GoldenDB 分布式数据库，实现 Z 系列主机为代表的现有对私核心系统进行分布式改造。新核心承接承载 8 亿用户信息、存款、贷款、转账等核心关键应用。2021 年 3 月实现全网并网运行，2022 年 12 月完成十个省分公司对私客户在 GoldenDB 分布式数据库投产，系统投产至今运行稳定。

二、总体方案

建设银行分布式系统总体技术架构包含应用路由、服务集成代理、配置中心、分布式数据库等主要组件。其中分布式数据库组件使用 GoldenDB 分布式数据库。各组件构成有机整体，实现大型系统灵活路由和高并发处理能力。具体如下：

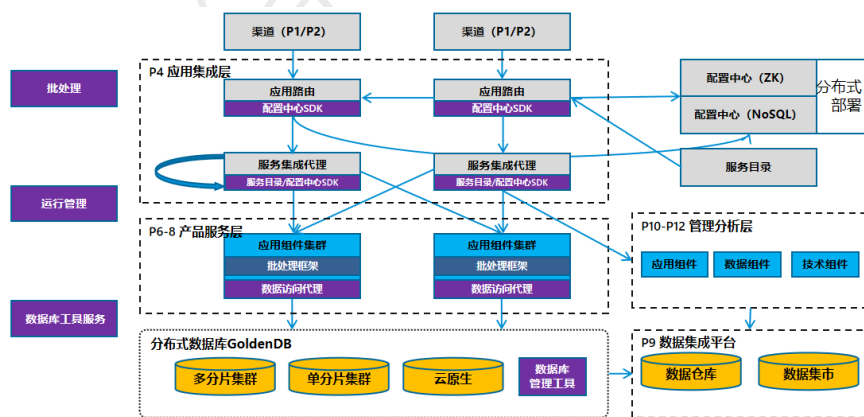


图 1 应用总体技术架构

应用路由负责交易请求的跨平台路由，查询配置中心接口，如属于本地 Region 的交易请求，发往本地应用集成代理层就地处理，否则将请求转发至交易归属 Region 的应用路由，后者接

收到请求后会将请求转发给与其同 Region 的服务集成代理处理。

服务集成代理负责跨系统事务一致性处理。服务集成代理会将单笔转账交易拆解成一笔转出交易和一笔转入交易，每笔交易都会同时准备好相应的冲正交易数据。此外，在系统投产过程中，一支业务的部分数据还在主机平台，部分数据在分布式平台，这种情况也由服务集成代理跨平台协调一致。

产品服务层提供银行业务的具体服务，包含各产品服务，是整个信息系统的主要业务组件，包括主机平台、C 平台和 JAVA 平台。开放平台上引入 GoldenDB 分布式数据库，承载原来主机平台上的数据库服务。

GoldenDB 分布式数据库提供大规模数据存储和高并发数据库访问能力。GoldenDB 实现数据多分片、弹性扩容和分布式事务强一致读写，通过数据多副本、快同步、分组管理、备份恢复等技术，实现大规模数据库的高可用、高可靠，以及多地多中心容灾能力。

全网对私核心有 8 亿用户数据，百 T 级别数据规模，并发处理要求高，系统采用大单元机制，在一套 GoldenDB 集群下创建四个租户集群(见下图 C1 到 C4)，每个租户有 25 个数据分片，分别部署在四个数据中心。同一数据中心内的请求均由同一数据库租户处理，无需应用层协调分布式事务，这显著降低跨服务集成代理跨子系统处理的业务吞吐量，减少资源压力。

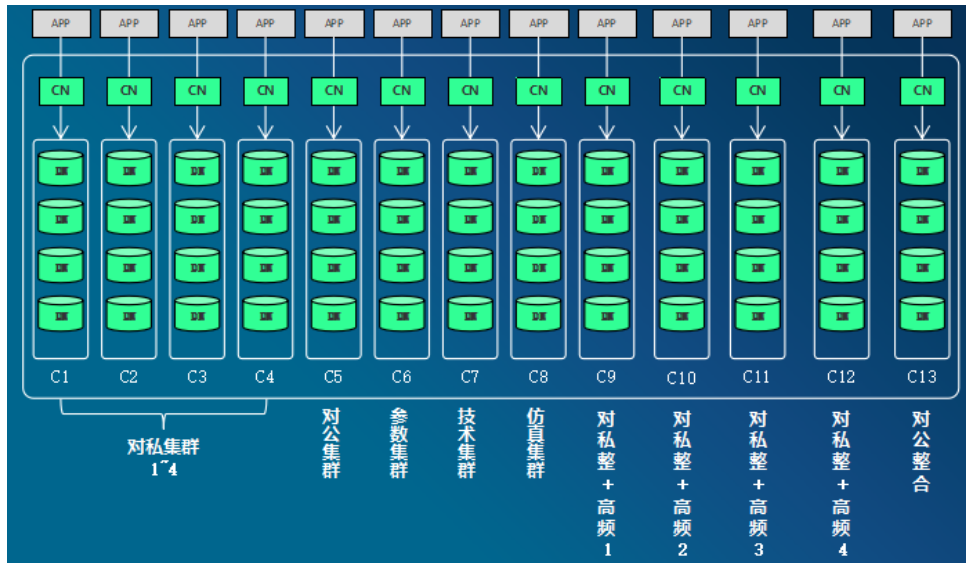


图 2 对私核心业务数据库系统框架

三、难点问题及应对举措

降低应用分布式处理压力。引入 GoldenDB 分布式数据库前，原方案采用细粒度单元设计，平均每个分行为一到两个单元。在该方案下服务集成代理需要协调数十套单元间事务一致性，既要负责处理跨分行之间的事务操作，同时要负责处理同一家分行跨多个服务单元的事务操作，应用的分布式事务调度压力过大。综合决策后，基于 GoldenDB 数据库调整为大单元设计，即将全国用户按区域分成四个大片区，片区内用户的转账等分布式事务交给数据库处理，简化后服务集成代理的分布式处理压力大大减少。

简化配置中心数据管理。项目初始设计方案采用集中式数据库进行小单元设计，由于数据库单元数量过多，需要在配置中心上配置大量业务处理单元的定位信息和路由数据，维护和变更复杂。通过 GoldenDB 的多租户管理、租户内分布式强一致处理能力，将全网用户集中到四个大单元，大大简化了业务处理单元

的定位信息和路由数据，路由信息未来十年无需修改。

四、应用成效

建设银行于 2021 年 3 月基于 GoldenDB 实现对私核心全国并网运行，到 2022 年 12 月已有十个省分公司对私客户在 GoldenDB 分布式数据库投产，对私核心请求达 8000+TPS，系统投产至今运行稳定。本项目实现两地 4AZ 多地多活部署、性能容量线性扩展，解决了应用从主机平滑迁移到分布式数据库的难题，满足业务未来发展需求，是国内大型银行 Z 系列主机核心应用系统多地多活分布式改造的典型实践案例，可为其他大中型银行提供借鉴参考。

案例六：交通银行核心系统 OceanBase 应用实践

一、应用场景

贷记卡作为满足人民美好生活需要的金融载体，在畅通市场循环、产业循环中持续发挥重要作用，且客户群体广泛，具有示范代表意义。交通银行贷记卡核心系统支撑着交行上亿客户交易，从大型主机下移到 x86 服务器，数据库升级至我国原生分布式数据库 OceanBase，实现了贷记卡核心系统的高性能、高扩展、高可用。

二、总体方案

交通银行核心贷记卡系统采用单元化多机房多活整体解决方案，OceanBase 数据库提供两地四中心五副本+主备库高可用方案，备集群部署在两地，提供城市级的容灾方案。系统采用单元化微服务架构，控制故障影响范围，减少故障爆炸半径。

新一代分布式云计算平台是交通银行分布式信息系统建设关键“底座”，也是未来分布式技术框架的核心。结合贷记卡系统重构建设，新云平台按照全行一体化协同、一体化运维、一体化分布式技术栈的建设要求，对城市 A 中心、同城中心、城市 B 中心和测试云等多个环境进行规划和建设。

交通银行的架构方案采用阿里云+Sofa+OceanBase 单元化多机房多活整体解决方案。OceanBase 提供两地四中心五副本+主备库高可用方案。OceanBase 采用同城三机房容灾部署架构+异

地主备集群架构。其中 OceanBase 主集群采用“2+2+1”五副本方式部署，通过 Paxos 协议保证同城三机房数据强一致性；同时上海同城机房多活，可按照流量比例进行调拨。

比较特殊的是，因为应用系统只是简单的主备架构，不能够像 OceanBase 一样是多地多中心，为了防止主机房出现双断，OceanBase 除了在某城市有一个备份集群，还在主机房增加了一个备份集群。这种部署也显示了 OceanBase 的灵活性。

三、难点问题及应对举措

交通银行原贷记卡核心系统使用 IBM 大型机+CICS+DB2 for zOS 的技术体系，整体架构稳定但技术栈封闭、很难水平扩展，且 IT 投入巨大，无法满足快速增长的业务量和快速响应各种新的业务需求。此外，一是容灾标准高。交通银行贷记卡核心系统支撑着交行上亿客户交易，需要满足 7x24 小时持续服务，高可用容灾要求达到 5 级。二是建设成本高。原有核心系统基于传统大机和 DB2 数据库的封闭模式运行架构，IT 建设成本高昂。三是备机房资源浪费。近年来随着业务并发量的不断增加，数据库系统处理能力不足的问题凸显。冷备机房随时待命但不提供数据服务，资源利用率低。

交通银行采用 OceanBase 作为其分布式信息系统建设关键“底座”和核心，帮助交通银行整体 IT 体系向新一代云平台、分布式架构全面转型。

四、应用成效及经验

交通银行贷记卡核心系统“大机下移分布式”，降低了主机 mips 的消耗，每年节约 IT 成本数千万元。贷记卡核心系统实现两地四中心 5 级容灾，满足 7x24 小时服务要求。采用单元化微服务架构，控制故障影响范围，减少故障爆炸半径。采用跨中心分布式单元化批量运行框架，实现核心系统双中心同时跑批、多单元并行处理，解决单元化文件拆分合并、跨机房文件传输等难点，提升资源利用率和批处理效率。

交通银行贷记卡系统属于 A 类核心业务系统，其上云实现了国有大行核心系统从 IBM 大机集中式架构，向云上分布式单元化金融级架构的技术转型。目标支撑上亿卡量，千万级别的日均交易量，并通过“同城双活+异地灾备”的两地三中心容灾架构，确保核心业务 RPO 为 0。生产环境性能测试已达到金融类交易 6000TPS+非金融类交易 20000TPS 的目标值。

案例七：邮储银行新一代个人业务核心系统 GaussDB 应用实践

一、应用场景

核心系统是银行 IT 系统建设的重中之重，是一家银行的大脑和心脏。其中，统一查询系统是核心系统中的重要组成部分，该系统承载着柜员和客户关于历史交易明细、账户信息等多个场景的查询以及实时收支分析的任务。系统拥有百 TB 级的海量历史数据、千亿级单表数据量，高峰期需面对上万的用户并发量，规模庞大、场景复杂、数据量巨大的业务特点对数据库大容量、高并发、高可用、高性能提出了极高的要求，这些要求，对任何传统商业数据库来说都是巨大挑战。

二、总体方案

（一）项目实施模式

双轨。旧核心和新核心进行并跑，逐步将旧核心数据迁移到新核心，直到完全替代。

（二）实施周期

2019 年 9 月：新一代个人新核心建设启动

2021 年 4 月：技术平台投产

2021 年 7 月：运维平台投产

2021 年 10 月：首批业务投产（国际汇款）

2022 年 4 月：新核心全功能投产

（三）数据迁移

原系统数据通过离线 CSV 文件导入的方式导入新系统，通过白名单方式进行逐批次的客户迁移，进行路由策略配置。

三、难点问题及应对举措

从传统数据库往分布式数据库转型的过程中，面临如下痛点问题：

1. 高可用挑战大：业界常用的数据库方案只能部署在跨园区的单 Region 中，不能满足高等级应用的高可用要求。

2. 性能容量挑战大：大型业务系统高峰期 TPS 近万/秒，单库几十 T，对基础设施及数据库的性能容量提出了较大的挑战。

3. 扩展性不足：传统的集中式数据库由于架构限制，往往采用一主一备，即使是 Oracle RAC，由于共享存储的限制，也最多只能扩展 2 到 4 个节点，单集群的容量容易出现瓶颈。

4. 安全性有待提升：近年来随着大众对保护个人信息意识的提升，以及各国家地区纷纷出台数据安全相关的政策法规，如何保护数据安全隐私的问题受到重大的关注，同时也是工业界与学术界研究的热点问题。

5. 数据同步挑战大：大型业务系统高峰期超过几百 G/小时数据归档量，对新旧系统双库并行期间的数据同步效率要求极高。

华为云 GaussDB 采用行业先进的全并行分布式架构，通过多个节点并行分担系统压力，提供极致吞吐量；还拥有超大存储容量，支持事务的强一致性；在数据保护方面，提供两地三中心的

容灾方案和多层级冗余保障数据的实时安全，实现系统无任何单点故障。

四、应用成效及经验

系统上线后可支撑海量交易、弹性伸缩、金融核心级高可靠和高可用，可具备为全行 6.5 亿个人客户、4 万个网点提供日均 20 亿笔，峰值 6.7 万笔/秒的交易处理能力。

邮储银行新一代个人业务核心系统全面投产上线，使邮储银行科技金融再添一部新引擎，将为邮储银行提供源源不断的科技动能，加速建设“一流大型零售银行”。同时鲲鹏、openGauss 开源数据库与 GaussDB 分布式云数据库将继续全力支持邮储银行数字化升级，共同为同业积累经验，探明路径，树立标杆。

该系统同时采用企业级业务建模和分布式微服务架构，是中国银行业金融科技关键技术可控的重大实践。业务处理性能 5 倍提升，由原来的 1 万笔/秒提升到 5.5 万笔/秒，支取响应提升 25%，查询响应提升 27%。支持 500T 超大数据存储，10 年超长数据查询范围。

三、股份制银行

案例八：平安银行信用卡核心系统 TDSQL 应用实践

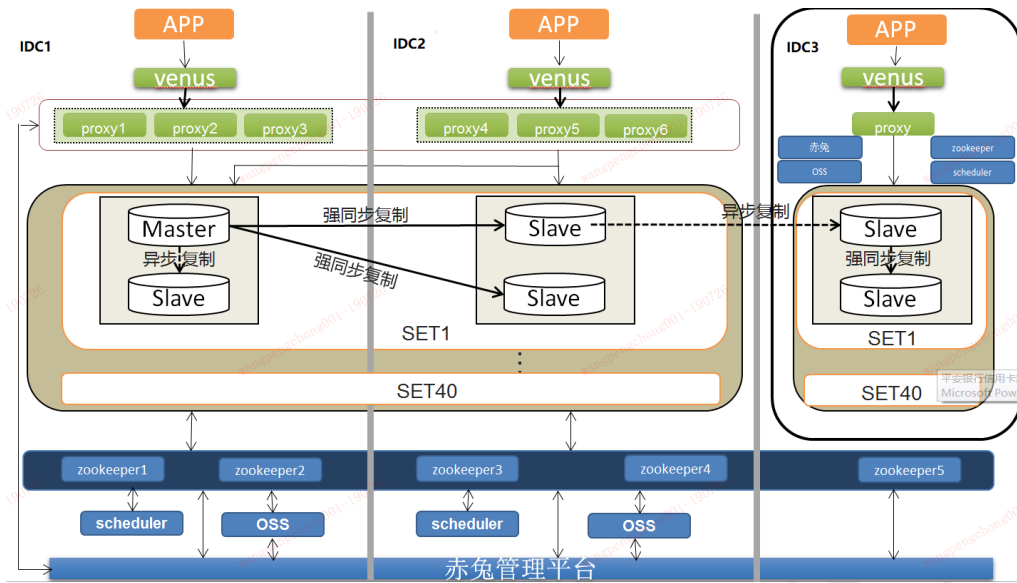
一、应用场景

平安银行早期的系统建设以 IOE 架构为主，近年来随着业务快速发展，系统的数据量和负载爆发增长，平安银行系统升级改造时选择了具有弹性扩容、高可用、高性能等优势分布式数据库，可以有效支持海量、高并发、高吞吐量的新型金融业务应用系统。其中平安银行信用卡新核心是我们探索应用我国分布式数据库并应用核心业务系统的具有里程碑意义的一步。

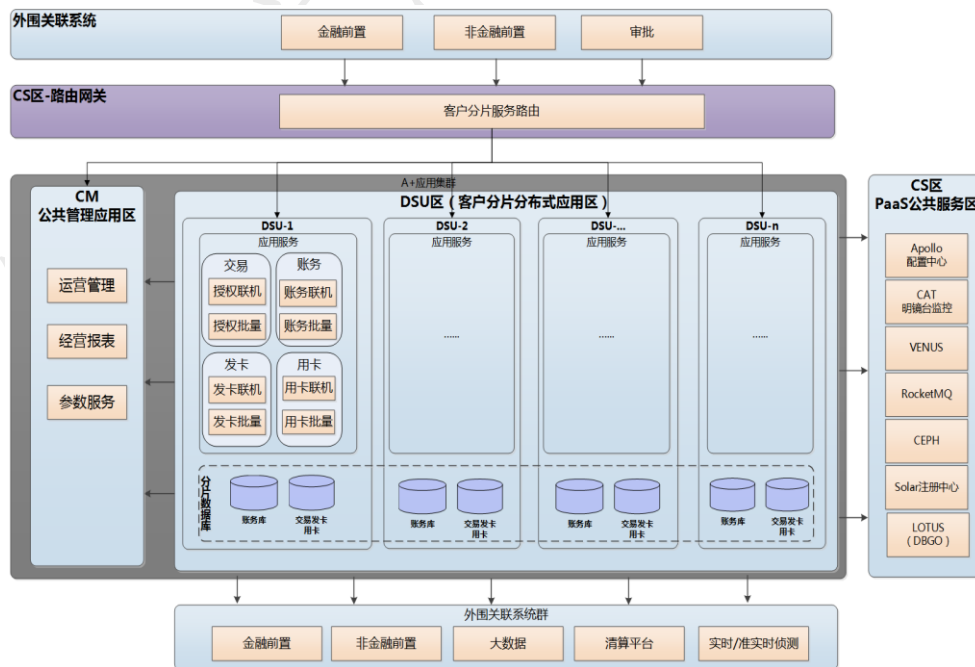
二、总体方案

平安银行信用卡新核心选用腾讯云企业级分布式数据库 TDSQL，构建了一套能够快速实现业务交付且灵活富有弹性的核心系统来替代以传统集中式数据库为主的老核心系统，同时建设了一整套企业级全栈式服务化的技术中台。

数据库采用典型的两地三中心、一主五备的部署架构，实现了同城多活、异地容灾、一键切换等需求，提高系统的并发处理、隔离故障和灵活扩展能力，具备读写分离、自动分片、主备强一致、实时监控、实时冷备等功能；在运维层面，基于 TDSQL 建设一套全链路的分布式运维工具。



信用卡新核心系统采用单元化 DSU 分布式架构，基于私有云和 PaaS 平台建设，应用微服务化，拆分解耦。DSU 整体设计逻辑采用按客户维度进行分片，由 GNS 去负责解析，完成分片，由 DLS 实现分片的路由，分片内部实现自包含，客户的业务均可以在单个分片完成，包括交易授权、用卡业务，也包括批量业务。



三、难点问题及应对举措

分布式单元化架构的采用会带来的主要问题:

1. 分布式架构建设和改进的工作量大, 需要配套建设的分布式工具也很多。

2. 过度单元化可能会带来很多问题, 比如上下游业务不均衡带来的小业务被动单元化, 以及单元化带来的应用层和数据库层的资源浪费等。

应对措施:

1. 平安银行信用卡新核心单元化架构基于专门打造的分布式 PaaS 平台, 应用在进行单元化分布式改造时尽可能只关注业务逻辑的实现, 而平台架构、数据库和各种工具交给 PaaS 层。

2. 数据操作尽可能规避分布式事务, 以保证系统的健壮性。在业务逻辑上, 尽可能实现单元内业务自包含。如果一定要进行分布式事务, 考虑在框架层支持, 如果框架层也不能支持时由业务保证。

3. 在 DSU 之外, 为了满足聚合查询、分析、归档需求, 建设一套 Sharding 版的 TDSQL, 用于实现聚合的查询, 支持全量、增量以及实时的数据同步; 同时为了完成数据归档, 建设了一套 Hadoop 集群作为归档数据平台。

四、应用成效及经验

平安银行信用卡新核心是由大型机集中式架构迁移到 PC 服务器分布式架构系统的成功案例, 业务系统完全自主可控。在成本方面, 根据实际的测算, 以 5 年为周期, 新核心系统相比老

系统成本节约近 70%，节省费用超 10 亿。在可靠性方面，通过可视化分布式调度平台，支持金融级高可用，同城双活异地灾备，故障场景下秒级切换。在性能方面，能够支持 10 万+业务作业统一处理、跨地域分布式调度以及可视化管理，满足金融级核心系统的要求，支持 10 万 TPS 交易高并发，相比老系统的交易处理能力提升了数十倍，日终业务批处理小时级时效。

2020 年上线至今，已经稳定运行 3 年，经历了各种节日或者业务促销，以及各种各样的软、硬件故障，总体运行情况符合预期。同时，经过多年的探索和实践，平安银行已经基本掌握了一整套适用于各种业务场景的分布式架构，在分布式数据库应用场景上也将不断扩展、深入，加大我国数据库的应用力度、加快原有系统升级改造，推动金融业数据库创新向深入发展。

案例九：光大银行新一代信用卡核心系统 GoldenDB 应用实践

一、应用场景

光大银行原有信用卡核心系统受限于传统主机核心系统的业务与应用架构，在整体业务规模加速增长的趋势下，难以有效支撑银行信用卡业务快速、灵活、弹性与创新的发展需要，为此基于 GoldenDB 分布式数据库建设新一代信用卡综合管理系统。新一代综合业务管理系统面向全网用户实现账户管理、信用卡申请处理、信用卡授权模块、交易管理、分期支付、会计处理、国际卡组织清算等信用卡业务服务。

二、总体方案

光大银行信用卡新一代综合业务管理系统使用单元化设计、多地多活等先进设计理念，将原有的核心业务按照客户维度划分为六个业务单元、一个灰度单元和一个全局单元。六个业务单元分别部署在不同的数据中心接管业务，并且在其他数据中心具有灾备能力。本地多中心能够实现数据的实时一致性复制，保证本地任何一个数据中心与主中心都能实现 $RP0=0$ ，灾备中心可以实现准实时灾备。该技术架构可以实现业务按照单元化进行灵活处理，控制故障半径，数据中心资源使用率高，支持灰度升级，在极端情况下一个数据中心即可承载全部业务。

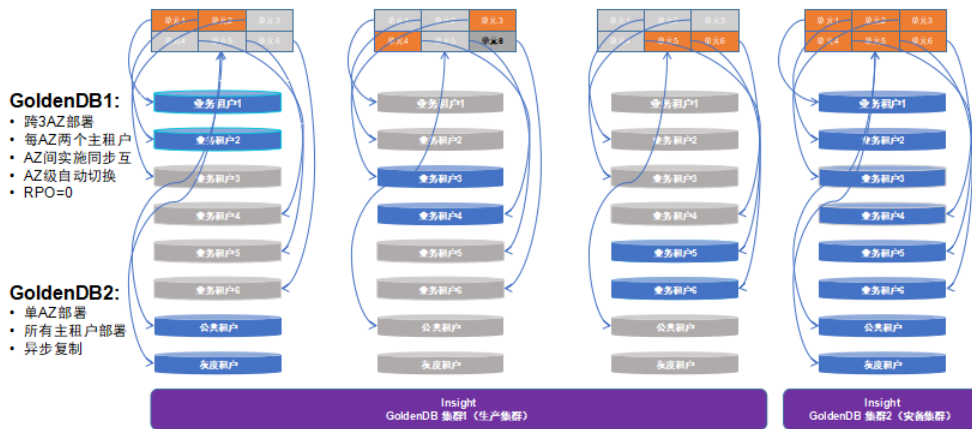


图 1 信用卡新一代综合业务管理系统架构

三、难点问题及应对举措

新应用迭代投产前如何进行生产试运行验证，是单元化方案的关键技术挑战。项目组通过多种参考方案对比和行内论证，最终使用灰度客户数据，基于 GoldenDB 多租户实现灰度单元方案，创建独立的 GoldenDB 租户单元，将银行职员作为受控灰度验证的客户群，存储于灰度租户单元中。每次应用新版本发布前，先在该灰度单元进行灰度发布，验证没有问题后再发布到其他业务服务单元上。灰度单元设计具有灵活的可用性，故障不会影响其他业务服务单元的运行和服务质量；同时用于灰度验证的都是受控客户，从内部的银行职员开始验证，整个灰度发布过程更加可控；该方案灰度用户总是在灰度单元里，每次灰度发布时无需数据搬迁，管理机制更加简单一致。

四、应用成效

光大银行信用卡新一代综合业务管理系统于 2023 年 3 月上线，支持 1 亿客户量、2 亿卡量，支持日常金额交易 1000tps，

电商活动高峰期 5000tps，查询维护类交易 6000tps，授权交易平均响应时延小于 60ms，查询维护类交易响应时间小于 50ms，每日主批运行时间小于 2 小时。系统提供有效的安全保密措施，确保系统和数据资源的安全，应用和数据库均具备良好的在线扩容能力，极大提升了该银行信用卡业务处理性能，很好支持信用卡业务的需求和信用卡核心技术创新的持续演进，提升以数据为驱动的数字转型升级能力。

四、城市商业银行

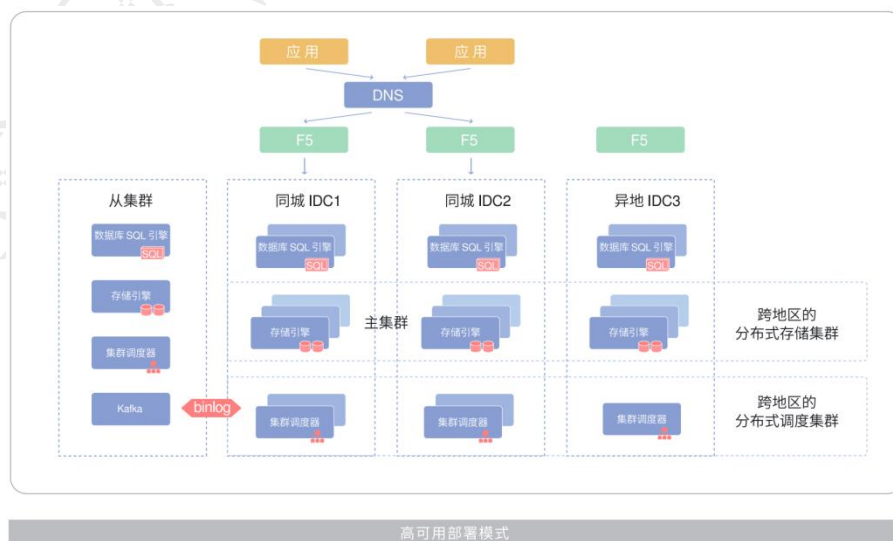
案例十：北京银行分布式核心系统 TiDB 应用实践

一、应用场景

面对海量数据、高并发的挑战，北京银行分布式核心系统采用“微服务架构+分布式数据库”的建设方案，构建起一套支持高并发、高可用、可横向扩展的分布式核心系统解决方案。对接网联支付清算平台、银联无卡快捷支付平台、金融服务互联平台、网贷业务平台等多个核心金融业务场景，实现了将分布式数据库解决方案应用于银行核心类业务场景。

二、总体方案

两地三中心部署 TiDB 集群，采用主从的多活架构，主集群作为生产集群承担日常的生产服务，主从之间采用 Kafka 同步 Binlog 的形式进行数据同步。



北京银行首先在网联支付清算平台和银联无卡快捷支付系统引入 TiDB 分布式数据库，以便更好地迎接互联网金融带来的大数据量和高并发的挑战。系统投产之后，已经成功应对两次双十一挑战。

北京银行 IT 团队进行多次线上的运维操作，包括版本升级、打补丁等，利用 TiDB 分布式数据库的多副本特性实现“运维零中断”的操作。随着系统升级，北京银行的网联业务链，包括上游的手机银行到网联、银联无卡快捷支付业务中台，到后台的金融日历、查询服务都已经进行了分布式架构的升级，完成了与 TiDB 的对接。

三、难点问题及应对举措

随着业务的快速增长，数据库成为基础设施链条里面最大的瓶颈。在大规模数据和高并发场景下，需进行数据库分布式改造。北京银行使用 TiDB 分布式数据库系统承载业务，解决在大数据量高并发压力下的处理性能瓶颈。利用分布式系统的敏捷扩展能力提升交付运营效率，实现在线弹性资源伸缩，支持跨数据中心双活的数据库系统架构，并利用行列共存的方式，解决 OLTP 和 OLAP 场景不同的业务需求。

四、应用成效及经验

北京银行分布式核心系统建设项目荣获 2020 年度亚洲银行家“中国最佳核心银行技术实施”大奖，从四个方面全面提升北京银行的金融服务能力：

1、提升系统性能

通过分离处理功能、分散处理压力、扩展处理能力等措施，保障海量数据、高并发的业务场景对接，大幅提升交易处理效率。

2、满足安全要求

基于一致性算法保证交易数据的强一致性，依托数据日志的备份恢复能力，提升数据的可追溯性，满足监管要求，提高自动化运维能力。

3、具备在线横向扩展能力

5小时内实现5亿条数据的在线扩缩容，整个过程对业务系统无侵入，保证了在线业务的连续性。

4、提升金融服务能力

分布式核心系统具备产品快速创新、差异化服务等能力，助力商业银行构建以客户体验为中心、产品丰富、核算分离的核心业务系统。

案例十一：北京银行分布式核心核算引擎 OceanBase 应用实践

一、应用场景

针对数字化转型重点任务，北京银行分布式核心系统采用分布式数据库升级改造建设方案，其中核算引擎系统作为承载北京银行全行级的、统一的、标准化的、可灵活配置核算规则的会计核算平台，是北京银行分布式数据库升级改造的重点系统之一。

二、分布式改造优化方案

核算引擎分布式数据库升级改造项目，采用将主任务拆分为多个子任务，多 POD 并行调用方案，经过负载均衡服务，发送至分布式数据库多副本多服务器内，充分利用分布式数据库架构，调用所有数据节点计算资源，并行处理大规模计算任务。

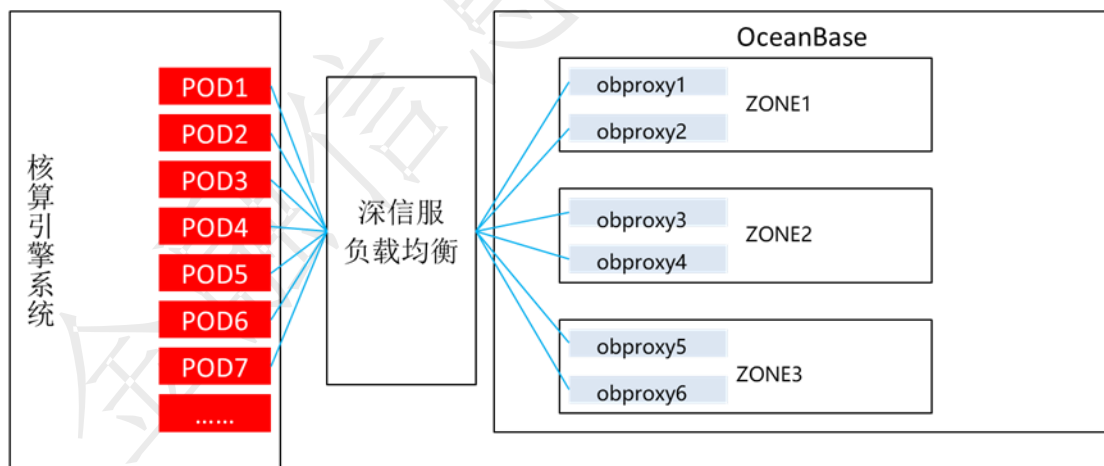


图 1 北京银行分布式核心核算引擎系统架构图

三、难点问题及应对举措

随着业务的快速增长，银行内部的日终、月终汇总计算面临着海量数据计算需求的艰难挑战，分布式数据库的优势在于大规模数据和高并发场景下的 OLTP 类业务处理，但 OLAP 类海量数据

计算仍与集中式数据库存在性能差距。为应对分布式数据库升级改造后的性能下降，项目组针对分布式数据库架构特点，采用串行任务拆分为并行执行、源表复制至目标端场景改造为表组绑定分区、单笔循环改为批量提交等策略，解决了分布式数据库数据分散导致 AP 类计算请求性能底下的问题，同等数据量级下执行效率甚至超过了原集中式数据库。

四、应用成效及经验

北京银行分布式核心核算引擎 OceanBase 应用实践，弥补了北京银行 OLAP 系统进行分布式数据库升级改造的技术短板，成为了分布式数据库开发设计、性能调优的优秀案例，从三个方面全面提升北京银行的金融服务能力：

1. 提升系统性能

充分利用分布式数据库特征，分散处理压力，针对海量数据计算需求，提升处理效率。

2. 具备在线横向扩展能力

分布式数据库多节点计算，针对业务发展导致的数据量增长场景，快速扩充服务器资源。

3. 提升金融服务能力

分布式数据库自带的多副本架构，具备两地三中心五副本等容灾能力，为北京银行节能增效，数字化转型提供有效支撑。

案例十二：北京银行基于 CirroData 和 BEH 的湖仓一体建设应用实践

一、应用场景

北京银行在分析系统领域使用了 GP、Netezza、Oracle 等多种数据库，来满足各类场景要求，且还不能很好的满足业务的功能需求和性能需求。为此经过全方位的充分验证测试，最终采用东方国信的 CirroData 和 BEH 进行了全面升级优化，建设了真正意义上湖仓一体的大数据云平台。大数据云平台实现了如下重要功能：

- 数据查询支持：查询平台从 Netezza 平滑迁移到 CirroData，并提升了查询效率；
- 数据服务支持：每天调用量超过几千万次，响应时间为 100 毫秒以内，并发支持 1000 以上；
- 高性能处理支持：公检法系统迁移到 CirroData 后，处理性能提高接近 100 倍，原来 1 个多小时的任务缩小为 1 分钟。

目前，累计部署服务器 500 台以上，有效存储空间超过 7PB，且使用常规的高性价比服务器，综合成本得到大幅降低。

二、总体方案

北京银行大数据云平台整体规划如下：

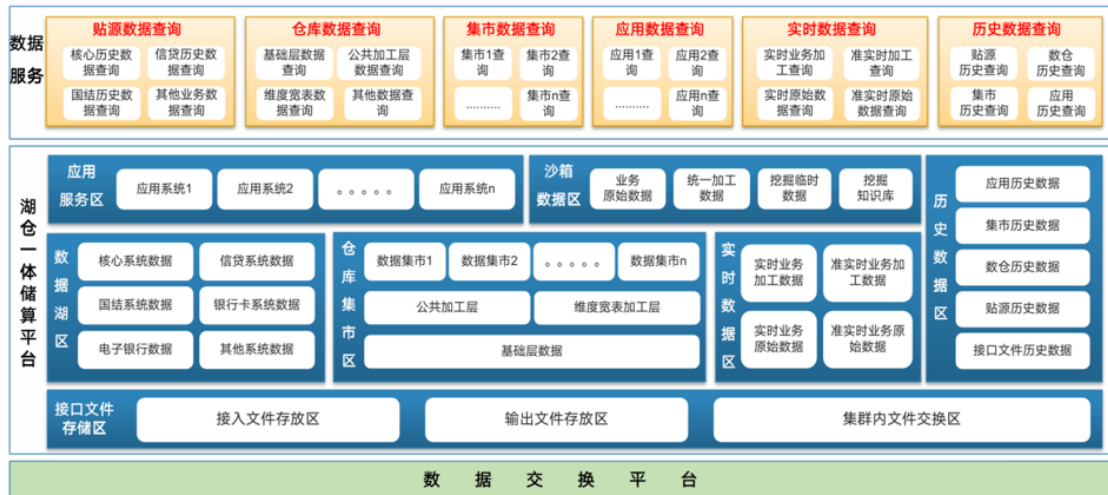


图 1 部署规划图

根据业务需求划分多个集群，其中数据湖区单独部署一个集群，部署 CirroData，存放贴源数据。仓库集市区单独部署一个物理集群，在此集群上部署两个独立的 CirroData 集群，分别存放数据仓库和数据集市。实时数据区单独部署一个集群，采用 BEH 和 CirroData 混合部署，满足各类实时场景需求。应用服务器单独部署一个集群，采用 BEH 和 CirroData 混合部署，满足各类应用场景实际需求。沙箱区单独部署一个集群，采用 BEH 和 CirroData 混合部署，满足各类实时场景需求。以上集群均采用常规高性价比服务器。历史区单独部署一个集群，主要部署 CirroData，存放历史数据，采用低计算高存储的服务器，降低存储成本。另外非结构化数据平台单独一个集群，主要部署 BEH，存放非结构化数据。

对于数据湖和数据仓库、数据集市的存储规划如下图所示：

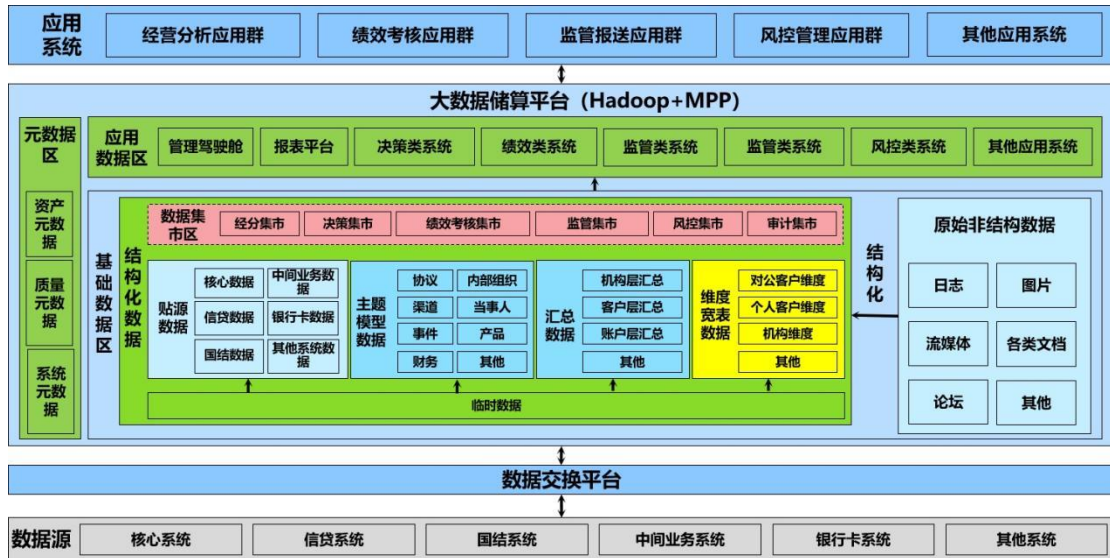


图 2 数据架构图

计算平台分为三部分：

1. 首先是基础数据区，里面包含了结构化数据和非结构化数据。结构化的贴源数据对应传统的 ODS，加上原始的非结构化数据对应常规概念的数据湖；主题模型数据和汇总数据对应传统的数据仓库；数据集市区对应传统的多个数据集市；维度宽表数据对应数据中台的企业级维度模型。

2. 基础数据区上面是应用数据区，里面存放应用系统专用的数据。

3. 还有元数据区，存放数据平台的统一的元数据信息。

在技术上，结构化数据存储统一采用新一代 MPP 数据库存储，非结构化数据目前仍建议采用大数据技术存储，如 HDFS + ES + Hbase 等组件进行存储，一般大文件可直接放到 HDFS 上，小文件建议存到 Hbase 中，未来还可引入 OSS 对象存储。而数据的采集、处理和特定计算可依托 Hadoop 生态组件实现，从而实现新一代 MPP 数据库和 Hadoop 生态的有机结合。

三、难点问题及应对举措

北京银行采用新一代 MPP 数据库 CirroData 和 Hadoop 发行版 BEH 产品，并在此基础上开发了全套完整配套产品和解决方案，解决了自身的如下痛点问题：

- 统一部署：CirroData 和 BEH 可以统一部署，使用一套分布式文件系统，可以按需部署，避免了传统的数据库和 Hadoop 必须独立部署，数据底层共享，大大减少处理环节，增加数据共享能力；
- 集中统一存储：避免了企业内数据在多个组件中冗余存放，以解决不同场景问题，结构化数据主要存储在 CirroData 中，非结构化数据主要存储在大数据组件中，且确定 ES、Solr 或 OSS 等，在企业内形成标准化规范，而 CirroData 的处理性能和联机查询性能也保证了企业内的大部分应用场景，只是针对特定数据的特定应用场景才需要使用特定组件进行数据缓存，而且是单向处理，大大简化了企业内数据管理的复杂度和冗余存放；
- 开发统一化：由于大部分操作尤其 90%以上是对结构化数据处理，均可基于 CirroData 产品完成，因此大部分开发人员只需掌握这一个产品，且以 SQL 操作为主，大大降低了企业的学习成本；
- 成本降低：CirroData 支持 X86 服务器和多种我国设备产品，并有多个实际案例，整体使用成本大大降低，完全可

以最优性价比在全企业推广，支持多种场景；

配套解决方案：除了产品以外，东方国信湖仓一体解决方案包含咨询规划能力、专业开发实施、专业运维服务和配套产品，提供全面服务，保证了湖仓一体平台的持续建设。

四、应用成效及经验

从使用 Oracle、Netezza、GP 等多个传统数据库产品，到集中使用了 BEH 和 CirroData 产品，建设了统一的数据湖，已经落地了数据湖区、仓库集市区、应用数据区、沙箱数据区、历史数据区、实时数据区等，建设了数据沙箱平台、数据服务平台、实时数据平台、非结构化数据平台、数据开发平台、运营监控平台、对公数据集市等多个数据集市，新建或迁移了公检法查询等几十个相关应用系统。

数据湖项目的建设逐步统一技术路线，标准化了开发流程和规范，提升了面向业务的响应时效，满足了之前不能实现或实现困难的系统和功能。

案例十三：贵阳银行信用卡管理平台 SUNDB 应用实践

一、应用场景

贵阳银行信用卡管理平台项目结合贵阳银行信用卡业务发展方向，采用信息技术解决方案，包括操作系统、中间件以及数据库等关键技术，建设了信用卡管理平台。在保证信用卡档案管理线上化安全性的基础上，平台整合各信用卡业务平台，支持贵阳银行信用卡档案的统一管理。减少信用卡部日常人工操作，降低人力成本及操作风险；最终实现对档案的长期保存、智能查询和科学管理。

贵阳银行总行信用卡部、管理行以及其下属的分支行，共计约 200 多家机构使用信用卡管理平台进行档案相关的管理工作。使用的功能和场景包括总行档案管理员、档案操作员每日查看业务办理流水，包括客户申请卡片、投诉、寄递、客户信息修改、账务类、核销等不同类型的业务，然后查看流水关联的影像，包括申请表、身份证信息、收入证明、房产证明、电核录音、主卡人职务证明材料等不同的材料，并对差错档案下发资料缺失、或者资料与实际不符等补录任务，分支行再在管理行的监督下进行档案的补录以及上传，最后由总行审批无误后进行档案的归档，以及对档案进行查询、下载以及调阅等应用。

二、总体方案

贵阳银行采用云上鲲鹏服务器 RH220T(芯片 HUAWEI Kunpeng

920 5220)、统信 UOS1020E 操作系统、科蓝 SUNDB 5.0 数据库、东方通 TongWeb7.0.4.4 中间件的技术路线。系统采用前后端分离技术，按照模块化开发、通过轻量化的 RPC 框架进行通信。

贵阳银行信用卡管理系统支持分布式部署和支持集群部署方式，提供灵活的可扩展能力。支持负载均衡处理机制，有效避免单点故障，以满足灾备部署要求。系统支持生产服务器与备份服务器同时投产的双活模式，灾备建设能支持双活中心部署的运营模式，同时本系统采用前端业务分离部署，根据需求通过网络对松散耦合的粗粒度应用组件进行分布式部署、组合和使用，使用服务的方式暴露接口，可高效快速的集成其他微服务系统，减少风险、人员成本和系统压力，从而保证了部署的易操作性、可复制性以及可参考性。最终在完成信息安全和创新目标的同时，实现小而自治，高内聚、低耦合，从而降低系统的复杂性，大幅降低系统建设、运维、升级带来的风险和成本。

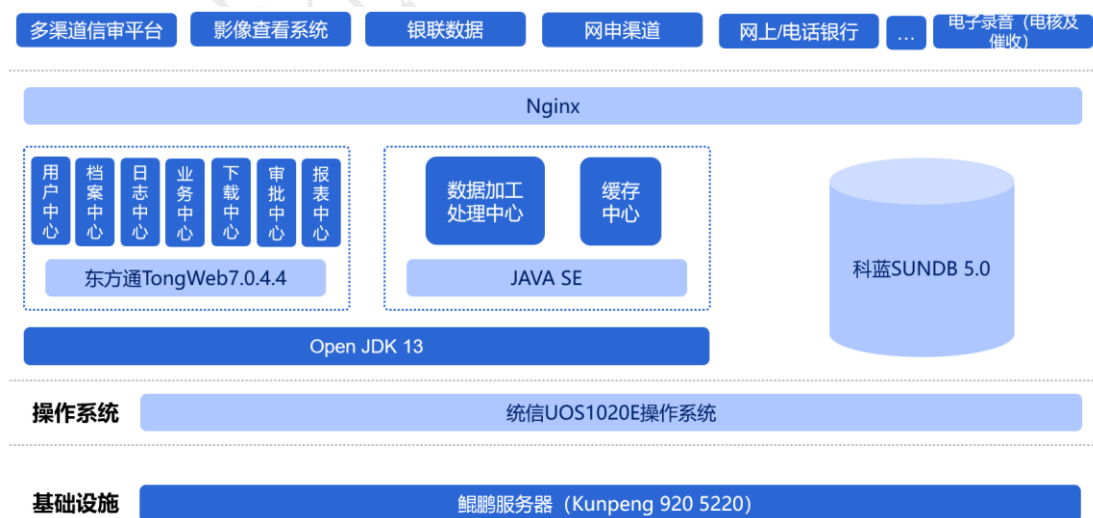


图 1 总体方案图

三、难点问题及应对举措

贵阳银行信用卡部现状，目前档案管理流程最主要的痛点在于纸质档案收集繁琐，审批周期长，且不利于对档案管理这类保密性工作进行业务监督。因此建设本系统最主要的目的是解决安全可控问题，并解决业务问题。

1. 安全可控问题，基于创新技术软硬件底层架构和全周期生态体系，解决核心技术关键环节安全可控的问题。

2. 业务上档案收集、审批流程长等问题，通过优化流程、信息化采集等技术手段解决。

3. 大字段数据类型转换问题，源数据使用特定的大字段，迁移 SUNDB 时面临接口不兼容情形，为避免较多的代码修改，由 SUNDB 开发兼容性接口解决。

四、应用成效及经验

贵阳银行基于信息安全基础软硬件、SUNDB 数据库构建的信用卡管理系统，完美地承载了信用卡管理业务的需求，促进了数字化业务发展，增强了数字化业务竞争实力。

贵阳银行信用卡业务系统自建设之初，就对系统的可靠性、安全性和系统性均提出了极为严格的要求。科蓝软件采用 SUNDB 数据库针和微服务方式管理，能够根据业务需要动态横向扩展，保证了系统的高性能的灵活扩展性。此外还部署了实时同步系统，保证了数据灾备的同时，也提供了实时全量备份。

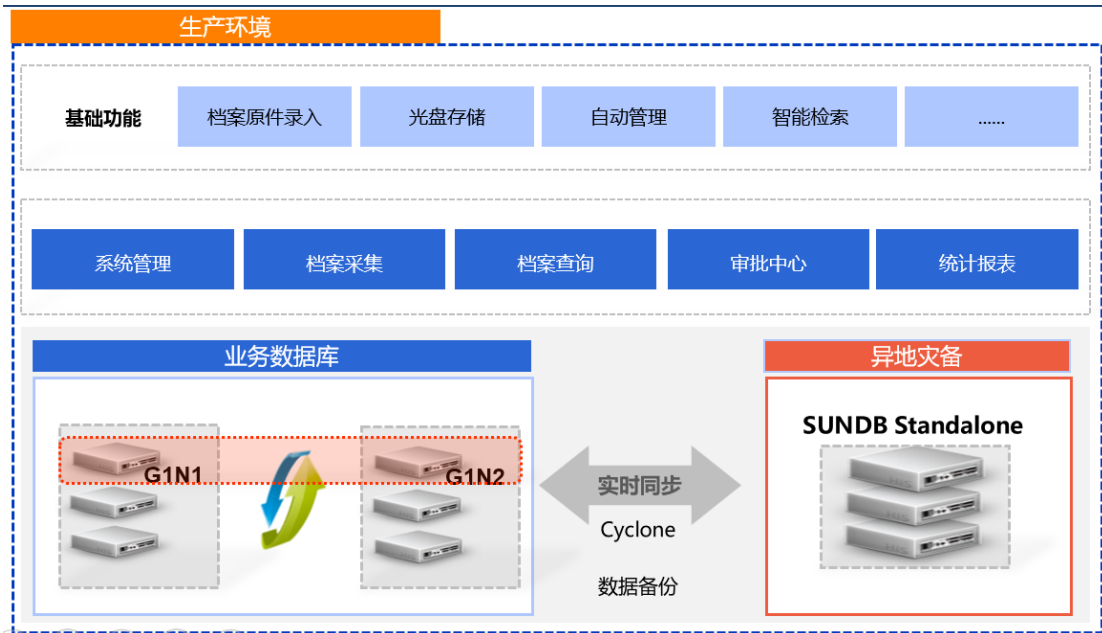


图 2 数据库架构图

金融信息化研究

案例十四：江西银行企业网银系统 SUNDB 应用实践

一、应用场景

江西银行选择科蓝软件对企业网银系统进行技术架构升级以及业务重构，联合建设基于全栈创新技术的企业银行系统，提高金融产品研发效率，提升客户业务处理能力。

江西银行企业网银通过对传统单体应用的电子渠道系统进行了深度分析与重构，建设基于分布式微服务架构的中台化电子渠道统一平台，具有显著的业务效果。首先是增强了系统的并发能力和横向扩展能力。其次是前端引入客户端的热更新及灰度发布能力，实现客户端无感更新，且功能布局可通过后管进行动态调整，随时随地高效地响应客户需求变化，提升客户体验。

江西银行基于我国基础软硬件，采用虚拟化技术、分布式微服务架构、SUNDB 分布式数据库，并按照分布式数据库特点进行数据模型设计，充分发挥分布式数据库的计算、存储的优势，支撑用户高并发、高性能的需求，达到应用、存储的主机下移的目标，并真正实现了基础软硬件的自主可控，同时避免了不必要的硬件资源浪费。

二、总体方案

江西银行企业网银系统采用科蓝“双鱼座”中台，支持分布式微服务；数据库采用科蓝自主知识产权的 SUNDB 分布式数据库；操作系统和 CPU 也采用我国的成熟产品。整合 PC 网银、APP 客

户端、现金管理用户体系，打造企业互联网统一平台，并需支持 IPV6 和全栈商用密码算法。



图 1 总体方案图

企业网银系统是行内第一个创新技术的关键系统，采用了微服务架构、分布式数据库等技术，同时也是业内基于全栈创新技术的关键类系统的案例之一。

1. 基础软硬件创新，包括服务器、操作系统、数据库、中间件等全面创新，摆脱对国外厂商的依赖，产品服务的可持续性不受外部环境因素影响。

2. 通过使用 SUNDB 分布式数据库实现计算下沉，实现数据存储层的线性水平扩展能力。

3. 使用 SUNDB 分布式数据库进行业务数据的存储，并按照分布式数据库特点进行数据模型设计，确保发挥分布式数据库的最大优势。

4. 通过应用的分布式微服务架构，实现计算能力的下移，并且能够实现线性扩展、稳定性运行、性能提升和高可用，避免对

主机计算能力的依赖。

5. 构建开放敏捷的全新技术体系框架，满足未来的 IT 发展和业务需求，满足数字化转型所需要的弹性、敏捷和协同共享等。

三、难点问题及应对举措

江西银行升级改造过程中面临如下几方面的问题，通过江西银行与 SUNDB 厂商、服务器厂商、操作系统厂商的紧密合作，完成问题解决方案和问题解决。

1. 数据同步问题，基于小型机的 DB2 与基于创新 CPU 的 SUNDB 解析数据的方式不同(大小端格式问题)，小端序系统和大端序系统通信时会发生数据解析错误，因此在发送数据前，要将数据转换为统一的格式。

2. 分布式数据模型设计问题，传统数据库在进行数据模型设计时可以按照数据库设计范式，分布式数据库因为将数据进行了分布，业务操作会涉及到跨节点的分布式事务，系统性能受到很大的挑战，为保障系统的整体性能，减少分布式事务、同时减少应用的改造是分布式数据库数据模型设计的最大问题。因此通过梳理业务数据特点及归纳，依据分布式数据库的分区能力制定数据模型设计规范，同时进行实际测试验证，确保分布式数据模型的设计合理、高效。

3. 我国服务器产品初始化问题，我国服务器产品的应用场景相对较少，尤其是作为最为重要的数据库服务器，所以服务器在初始化时对内存的初始化通常为小页且不能手动调整，而数据库

通常采用大页。对于采用小页的数据库，通常会有性能上问题。通过联系厂商，反馈问题，升级版本解决。

4. 我国操作系统安装包问题，我国操作系统的生态还不够成熟，对于一些在 centos 上较为常用的软件包、lib 等，在操作系统上存在找不到或版本不匹配的情形，通过与厂家的反馈得到解决。

5. 开发语言、框架、接口等问题，SUNDB 支持主流的开发框架和标准的数据库接口，但在开发过程中难免遇到使用较为小众、新兴的开发框架与接口的问题。评估后，通过开发解决。

四、应用成效及经验

江西银行基于我国基础软硬件产品，科蓝采用虚拟化技术、分布式微服务架构、SUNDB 分布式数据库，并按照分布式数据库特点进行数据模型设计，充分发挥分布式数据库的计算、存储的优势，支撑用户高并发、高性能的需求，达到应用、存储的主机下移的目标，并真正实现了基础软硬件的安全可控，同时避免了不必要的硬件资源浪费。

1. 性能提升

主机下移改造的业务应用系统功能满足业务要求，整体系统可靠性达到 99.99%以上，满足金融行业业务双活、性能提升、稳定运行的要求；

2. 创新基础软硬件

创新基础软硬件包括：服务器、操作系统、数据库、中间件

等，全面符合创新要求，摆脱对国外厂商的依赖，产品服务的可持续性不受外部环境因素影响。

3. 技术先进性

构建开放敏捷的全新技术体系框架，满足未来的 IT 发展和业务需求，满足数字化转型所需要的弹性、敏捷和协同共享等。使用 SUNDB 分布式数据库进行业务数据的存储，并按照分布式数据库特点进行数据模型设计，确保发挥分布式数据库的最大优势。

金融信息化研究

案例十五：湖北银行新核心业务系统中达梦柔性迁移双活应用实践

一、应用场景

对标某些银行核心业务系统 Oracle RAC + DataGuard 高可用架构升级改造场景，某些银行核心业务应用场景主要是 OLTP 交易型+报表分析型混合业务场景，具有交易复杂度高、数据强一致性要求高、响应时间延时低、高可用性，以及交易型和分析型混合业务场景等特点，此项目银行核心业务系统主要涉及到柜面系统、网银系统、支付系统、银行卡系统、信贷系统、风险管理系统、报表系统等。整体系统处理能力大约在 2000-5000TPS，响应时间在 100 毫秒以内，以及需要满足同城两中心或异地三中心高可用容灾架构。

二、总体方案

本项目在选型阶段，采用先课题研究，后应用实施；先实验室场景，后生产实测的路线，有序稳步推进。在进行数据库选型阶段，除了考虑常规技术标准，湖北银行结合自身业务需求、应用场景进行综合考量，着重考察不同维度最终选择达梦数据库，具体考察内容包括银行业务发展要求、降低项目风险与产品集成开发难度、技术服务能力、开发运维成本压力、数据库安全可靠、原创、银行业监管要求等。

关于实施和上线，湖北银行新核心业务系统主要分为两个阶

段，第一阶段是 2019 年实现了 Oracle 和达梦双中心双活运行，达梦数据库承担了包括手机银行、柜面、微信、支付宝在内新核心系统的 80%读业务，并且已在总线 ESB 系统的流水库和配置库完全使用达梦数据库，投入实际生产。第二阶段是 2023 年准备完成从达梦集中式读写分离集群向达梦共享存储集群转换，对等原 Oracle RAC + DataGuard 高可用架构，迁移到达梦共享存储集群 DSC+数据守护集群 Datawatch，实现第一阶段同城两中心的容灾高可用架构。

三、难点问题及应对举措

本项目中银行核心系统交互面广、层次结构复杂，数据库迁移风险难度大，如何实现循序渐进迁移改造、如何能够出现任何问题，立刻回退到原 Oracle 数据库，保障业务持续性，这是湖北银行新核心业务系统迁移中突出的问题。

本项目先通过实验室课题研究，后生产实测方式，有序稳步推进，首先，基于达梦全面支持 ANSI SQL 标准，快速实现数据从 ORACLE 向 DM 的移植，实现了低成本的应用改造。其次，本项目采用柔性替代的双中心双活解决方案，应用双写模式，实现了原 Oracle 和达梦的双活，达梦数据库分阶段逐渐分担 Oracle 业务，第一阶段已经完成 80%生产系统读业务在达梦数据库运行，第二阶段，准备完成数据库集群从达梦读写分离集群向达梦共享存储集群的升级，同时完成基于达梦数据库的同城两中心容灾部署，最终银行核心业务系统实现安全稳定的柔性迁移改造。

四、应用成效及经验

作为第一批在核心交易系统中采用我国数据库产品的湖北银行，本项目获得银监会银行科技发展奖二类奖。新核心业务系统的上线，对湖北银行在对优化客户体验、提升服务效率、防控操作风险等方面，具有重要作用：

1. 经实测数据的验证，达梦作业效率达 2000 TPS 以上，与 ORACLE 保持一致，提升了新核心系统的可用性和并发处理能力。新核心系统投产上线后，湖北银行个人开立账户只需 90 秒，大幅优化客户体验、提升服务效率。

2. 达梦数据库架构实现应用“双活”，可无缝切换数据库，确保系统运行安全。做到撤换对业务而言“无感知”，对新核心业务连续性“零影响”。

3. 达梦帮助提升核心业务系统安全性。达梦数据库拥有自主知识产权，帮助用户减轻信息安全隐患，避免了知识产权纠纷。

4. 达梦帮助提升核心系统经济可行性。在同等处理能力的情况下，大幅降低成本，使湖北银行拥有新的市场选择。在硬件方面，从 IOE 架构向 X86 服务器转型，节约高额硬件支出成本。

5. 在人力成本和开发成本方面，由于达梦数据库全面支持 ANSI SQL 标准和主流编程语言接口，能较好的嵌入业务应用，二次开发成本大幅降低。

同时，通过与达梦合作，推动湖北银行在我国数据库领域、金融科技领域关键技术知识储备和人才培养。

案例十六：四川银行金融核心业务反洗钱系统 GBase 8s 应用实践

一、应用场景

四川银行作为四川省首家也是唯一一家省级法人银行，是在成渝地区双城经济圈上升为国家战略的背景下，未来西部经济的发展最重要的金融支持单位。反洗钱工作是维护金融体系的稳健运行，维护公平公正的市场经济秩序的客观要求，对打击腐败等违法犯罪具有重要意义。反洗钱系统基本覆盖全行各种交易数据，对历史海量数据进行存储和计算的能力于数据库是很大考验。反洗钱系统以 OLAP+OLTP 使用场景，每天进行大额、风险评级、可疑案例分析，并对命中结果上报人民银行。

二、总体方案

该系统搭载全栈创新软硬件平台，采用华为鲲鹏服务器，中标麒麟操作系统，数据库为 GBase 8s 企业版并采用 GBase 8s 两节点 HAC 高可用主备集群解决方案。当数据库主节点出现故障情况下，备节点可以在 1 分钟内自动切换为主机，切换对应用系统完全透明，提升了集群灾备高可用能力，同时借助读写分离模式，HAC 备节点分担了主节点的读负载，保证了主节点的资源利用，满足 RPO=0，RTO<60s 的高可用能力。

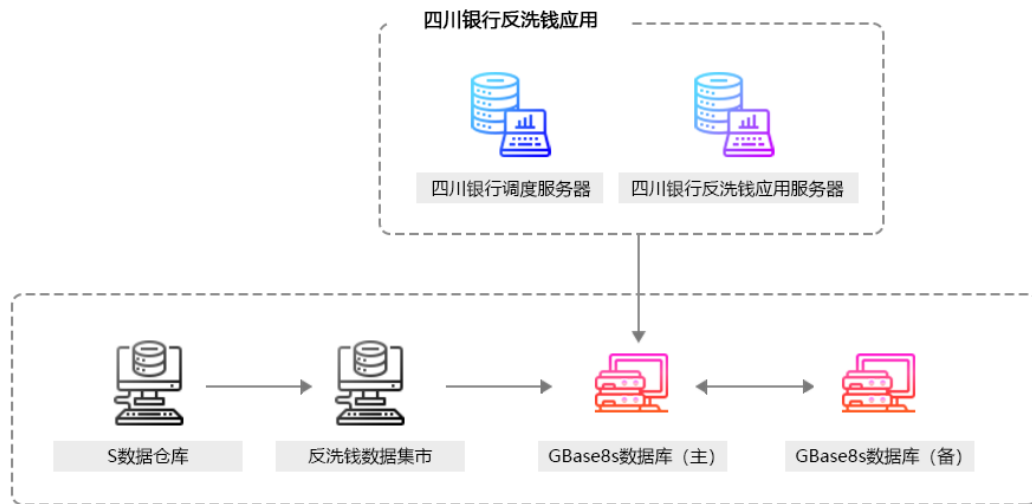


图 1 四川银行反洗钱系统逻辑架构图

在系统研发和改造期间，南大通用安排了专业的售后团队和研发团队在四川银行现场驻场提供最强有力的技术保障，在项目整体适配过程中出现过一些功能上的问题。南大通用通过快速迭代产品补丁的方式进行问题迅速应对。在项目期间先后升级了 4 个补丁版本，最终提供保证了功能、性能、稳定性的版本，满足了上线要求。并在 2023 年 3 月 29 日正式上线投产，实现了反洗钱系统在全栈创新环境下平稳高效的运行。目前业务数据量在 1T 左右，预估近 5 年总数据量持续增长至 3T。

三、难点问题及应对举措

1. 处理语法兼容问题：原系统 Oracle 迁移到 GBase8s，异构数据库存在语法和数据类型上的差异，需要调整 SQL、函数的创建及使用方式。

2. 批量任务 SQL 优化：处理复杂 SQL 执行效率上的优化，调整合理的索引创建方式。在性能压测阶段，出现了同样 SQL 语句在 Oracle 和 GBase8s 中运行时间差异大的问题。经过技术专家

分析定位其核心原因为两款数据库产品的查询优化器算法上差异导致。通过优化 SQL 写法，在保证结果集正确一致的前提下，优化后的 SQL 执行效率有了显著的提升，单个处理任务涉及 6 条上亿级大表的多表关联查询 SQL，通过优化调整，运行耗时从单次 12 个小时提升至 3 小时左右，超越原有系统处理效率，超过整体系统的性能标准预期。

四、应用成效及经验

1. 提升系统性能

GBase 8s 的引入提高了用户反洗钱工作的运行效率，降低了运营管理的难度，通过优化调整，运行耗时从单次 12 个小时提升至 3 小时左右，超越原有系统处理效率。

2. 提升金融服务能力

GBase 8s 在四川银行反洗钱系统的成功实践，是 GBASE 南大通用在金融领域扎根做实的又一有力印证，同时也为金融机构实现创新升级改造和 IT 系统的迭代演进的统一发展提供了典范。

3. 满足合规性要求

基于强一致特性和高可用技术保障了反洗钱业务的合规性，极大增强了数据存储和计算的准确和连贯。

案例十七：秦皇岛银行分布式核心系统 TDSQL 应用实践

一、应用场景

秦皇岛银行成立于 1998 年，由秦皇岛市 21 家城市信用社和联社组建而成，是市属唯一国有控股法人金融企业。秦皇岛银行紧跟金融科技发展趋势，顺应现代银行变革潮流，着力打造区域性精品银行重大战略部署，为数字化转型奠定坚实基础，依据秦皇岛银行信息科技三年战略发展规划，于 2021 年 4 月启动了新一代分布式核心系统建设。此次项目建设历时 15 个月，涉及 1 个核心系统建设和 63 个外围系统同步升级改造，规模大，范围广，难度高前所未有的。

二、总体方案

新一代分布式核心系统采用分布式微服务架构，包括存款、贷款、产品、运营、客户、核算引擎等微服务，每个微服务相对独立运行和能够对外提供服务。新系统首次使用腾讯云分布式数据库 TDSQL，同时采用了网关熔断、应用熔断、优雅停机等技术进一步提高系统的稳定性。

新一代分布式核心系统建设采用基于双中心双活架构设计，新系统实现了注册中心、Redis、联机应用、批量处理及运维管理的双活架构部署，对于应用系统来说，各个节点都支持双中心双活部署模式。

本次项目建设实现了核心系统及相关配合改造系统的全局

流水号、统一错误码、统一日志等架构规范，落实了超时处理、参数同步、统一对账等技术要求，为秦皇岛银行后续系统建设打下了坚实的基础。

在分布式数据库 TDSQL 集成方面，秦皇岛银行采用了逐级逐层实施的创新路线：基础功能与适配要求分析（业务系统函数、语义 SQL、单表数据与压力承担、指标计算与设计）；集成架构设计（数据库架构选择，涉及灾备、双活、主备等设计）；系统兼容改造（业务系统语法部分调整；TDSQL 参数与解析层应用调整）；版本验证（基于联机业务、批量业务、EOD 业务、数据传输与同步业务验证以及事务处理能力、索引能力）；基础架构验证（高可用、延迟、全链路等）；安全架构验证（数据环境迁移、加解密安全与效率、角色权限管控等）。

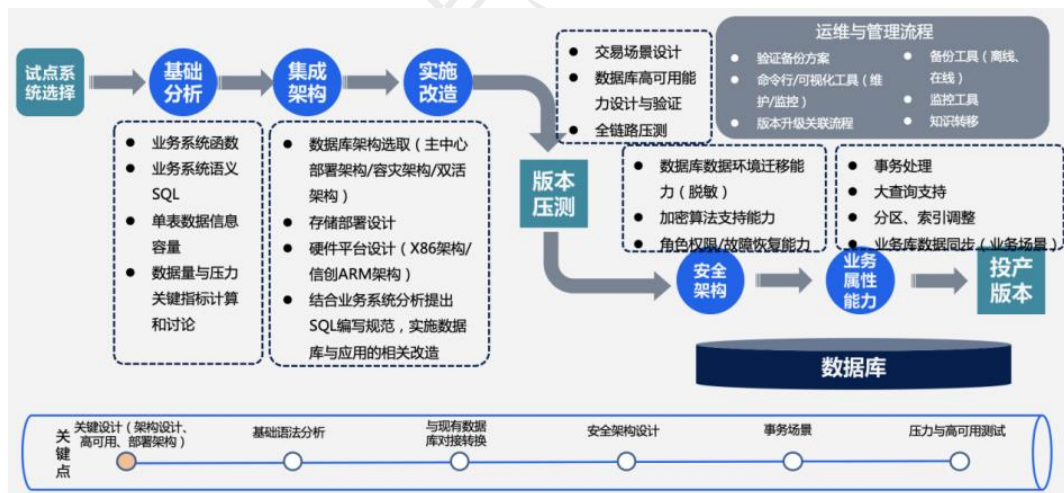


圖 1 秦皇島銀行核心系統分布式數據庫建設示意圖

基於逐級逐層實施的創新路線，秦皇島銀行通過關鍵系統選擇，基礎和高階語法分析、高可用架構、連續性設計與驗證、業務場景驗證等分階段分步驟的建設方法，既保障了秦皇島銀行系統建設週期和特點，又在風險可控條件下進行了前瞻探索和實踐

验证，从而确保本项目的成功落地。

三、难点问题及应对举措

在构建分布式核心系统的过程中，秦皇岛银行面临着诸多挑战，如何实现传统数据库海量数据向我国数据库产品的无损迁移是此次系统建设的难点所在。这是秦皇岛银行有史以来规模最大、范围最广、难度最高的一次系统升级，同时也是河北省首家在传统核心业务场景下使用我国数据库产品的城商行。

随着国际局势和疫情变化，秦皇岛银行发展既要满足科技创新要求，又要保持稳定基础的关键技术自主可控的路线。基于此，秦皇岛银行于 2021 年 4 月启动分布式核心系统建设，在 15 个月内完成 1 个核心系统和 63 个外围系统更新换代，提升运营基础、服务能力、产品工具、风险防控、创新引擎等多领域支撑能力。

四、应用成效及经验

新一代分布式核心系统是河北省首家在核心系统应用我国数据库产品 TDSQL 的案例。新系统同时采用了网关熔断、应用熔断、优雅停机、同城双活、弹性扩缩容等技术进一步提高系统的稳定性。

新一代分布式核心系统建设不仅满足了原有核心系统全部功能覆盖及新业务需求全面增强两大策略要求，而且同时实现了统一参数管理、业务流程优化、风险防控加强、多账户管理、弹性组织支持、精细化管理、差异化定价、产品快速创新、IT 系统加固等 9 大目标。

新一代分布式核心系统整体处理能力可达到 2000TPS，支持每日亿级交易量，动账类交易平均响应时间在 300 毫秒之内，查询类交易平均交易响应时间在 100 毫秒之内，日终批量时间平日缩短至 30 分钟左右，季度结息日缩短至 40 分钟左右，10 万笔社保批量代发 96 秒完成。

新一代分布式核心系统建设是秦皇岛银行改革发展的标志性工程，秦皇岛银行将依托新核心建设成果，全面提升数字化创新能力，运用金融科技手段为客户提供更加优质便捷的金融服务，更好的发挥金融保障作用，助力“六稳”“六保”，护航实体经济稳增长。

案例十八：梅州客商银行新核心业务系统中达梦高容灾可用性应用实践

一、应用场景

本项目对标集中式 Oracle RAC+DataGuard 高可用架构的升级改造的金融核心业务系统，其应用场景主要具有交易复杂度高、数据强一致性要求高、响应时间延时低、高可用性等特点，如在线事务交易型 OLTP 业务场景或大量 OLTP+少量 OLAP 混合新型业务场景等，如交易型和报表业务的银行核心业务系统。

二、总体方案

选型中考察 8 家数据库厂家不同类型数据库产品，包括集中式、分布式和云数据库，结合了自身业务需求和应用场景，着重考察数据库安全可控原创性、银行业监管要求、银行业务发展要求、开发运维成本压力、技术服务能力、降低项目风险、与 Oracle 产品兼容性、产品集成开发难度等多个维度进行数据库选型，最终选择达梦数据库作为唯一一家入围的数据库厂商参与梅州客商银行新核心系统改造项目建设。

关于实施和上线，根据梅州客商银行的具体实施安排，分为两个阶段：2019 年 1 月-2020 年 8 月，完成部分核心业务系统双轨运行；第二阶段：2023 年 3 月至今，实现了部分核心业务系统单轨运行和高可用部署，目前仍在不断快速扩大单轨运行范围。整体实施工作中，主要完成了开发阶段的模块深度解耦、不同阶

段功能和性能测试、不同阶段的迁移实施、双轨运行和单轨运行等。

三、难点问题及应对举措

针对银行核心业务系统中大量的 Oracle RAC + DataGuard 架构，我国数据库产品真正实现对标国外此产品的研发难度比较大，周期比较长，以及是否满足修改少量代码的兼容性，如何实现安全稳定的升级改造，如何沿用之前成熟的技术储备实现迁移改造之后的开发、测试和运维等问题，一直是金融核心业务系统比较关注的问题。

通过梅州客商银行核心业务系统的升级改造项目，探索式的解决了类似的一些问题：对标 Oracle RAC + DataGuard 高可用架构，达梦数据库共享存储集群 DSC 和数据守护集群 DataWatch 的组合架构可以实现对等国外商业数据库能力；同种类型数据库升级改造，更好实现了驱动、语法、架构型态的兼容性，带来更少的业务侵入性；以及大量实践升级改造中打磨的双轨迁移相关工具，如数据实时同步工具等，保障升级改造的安全稳定；其次，基于对标 OracleRAC+DG 的数据库的良好兼容性和技术延续性，实现了复用积累的运维管理技术，使用了之前成熟的运维习惯，以及从集中式到集中式的相同类型数据库升级改造，降低了后期测试、开发的难度，避免了两种或多种测试、开发、运维的双倍的成本压力。

四、应用成效及经验

梅州客商银行核心业务系统迁移改造项目，通过创新软件、本土化管理工具、完善的用户手册、以及高效原厂服务等保障了系统单轨安全稳定运行，具体成效及经验如下：

1. 通过部署我国共享存储集群+集中式主备的异地两中心高可用架构，单轨稳定运行半年多。

2. 集中式到集中式的同种对标数据库的升级改造，共享存储集群+集中式组合架构，实现了架构、语法、存储结构等不同层面的良好兼容性，对应用侵入性比较少，同时在确保安全稳定升级改造的同时降低了后期开发、测试和运维管理等成本。

通过此项目积累了我国数据库产品技术储备，制定了类似《基于达梦数据库的信息系统运维技术规范》等一系列技术资源储备，其次逐渐培养出我国数据库人才队伍，金融用户的技术团队很多获得了达梦数据库管理员认证等。

案例十九：某银行核心业务领域 GBase 8c 多模多态分布式数据库应用实践

一、应用场景

随着移动互联网的兴起，互联网金融蓬勃发展，各商业银行在基于自身优势的基础上，逐步结合移动互联网技术和业务模式探索互联网支付、消费金融和财富管理等业务领域的互联网化，寻求具备云原生、分布式以及全栈创新升级能力的可落地解决方案。2022年，南大通用 GBase 8c 多模多态数据库助力银行科技创新，逐步实现了该银行对公/零售信贷业务、风控业务、未来银行、互联网中台等多个业务系统的分布式探索和创新升级。

二、总体方案

为达到安全可控的建设目标，本项目搭载全栈安全可控软硬件平台，采用华为鲲鹏服务器、中标麒麟操作系统和东方通中间件，数据库则采用了南大通用 GBase 8c 多模多态数据库。根据客户的需求情况，项目最终基于 GBase 8c 的分布式部署形态打造了针对性解决方案。业务系统通过 GBase 8c 的 CN(协调节点)访问集群，核心业务数据水平分片，当数据主分片出现故障时，同城备数据分片可以在 30 秒内完成主分片切换，异地备数据分片可在 60 秒内完成主分片切换。切换过程对应用系统完全透明，业务中断时间短，数据无损失，用户感知好，同时具备了负载均衡、读写分离等效果，既提升整条链路的数据访问并发能力，又

满足了整体业务的高可用特性。

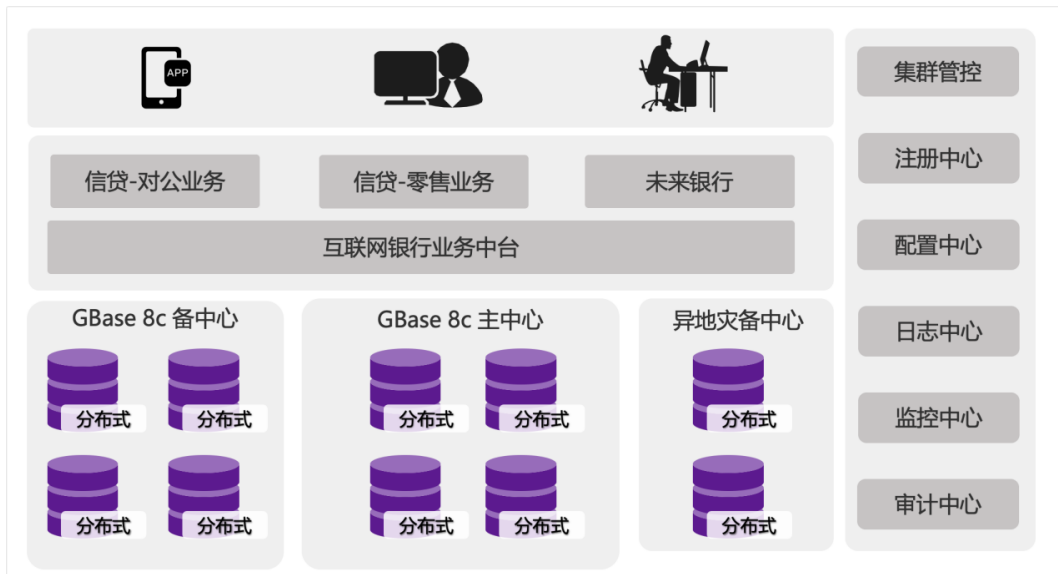


图 1 某银行分布式平台逻辑架构图

三、难点问题及应对措施

在本项目建设期间，用户的核心建设目标，同时也是分布式数据库在金融核心领域应用上的难点问题，主要包括：

- 亿行级表响应时间要求为毫秒级，以满足平台业务高峰时期的处理能力；
- 系统故障时必须具备秒级自动快速切换的能力，无需手动干预；
- 数据库必须具备平滑扩展的能力，以便于支撑未来数据量的巨大增长；
- 必须满足 7×24 小时的业务连续性需求，并且平滑升级业务不中断；
- 要求数据库对安全可控硬件平台的全适配能力。

为实现用户的核心诉求，GBase 技术实施团队从产品特性和

部署方式两个角度来满足客户核心诉求：

- 与应用系统开发团队一起，分析业务流程，对全部核心业务数据进行数据分布规划，并且针对分布式的部署方式进行存储过程等数据库对象的定向优化，满足性能要求；
- 设计并实施了两地三中心容灾部署方案，同城两中心，异地灾备中心，并且通过对内部高可用策略的配置和优化确保高可用稳定可靠；
- 从 CPU、服务器、操作系统、中间件到数据库进行全链路适配，并在本项目场景下充分测试，实现了全栈安全可控的综合解决方案。

四、应用成效及经验

本项目已完成 GBase 8c 集群的初步应用，汇聚了该银行互联网中台业务、对公/零售信贷业务、风控业务等系统的全业务数据，从安装部署、数据迁移、在线升级、在线扩/缩容、应用切换，再到故障演练、压力测试、性能调优，南大通用提供了全程技术支持，最终保障上述各业务系统的应用开发、迁移和平滑上线。

1. 显著提升了原有系统的性能与高可用能力。

整体数据层解决方案创新升级后，GBase 8c 以分布式的架构同时在业务性能、并发吞吐量、响应时间及高可用方面实现了全面的提升，在满足当下性能需求的同时，未来也可通过水平扩容的方式来继续扩展对业务负载的承载能力，达到客户预期，获

得客户好评。

2. 实现了分布式数据库、微服务架构等创新技术在银行核心系统的落地。

为促进行业技术架构发展，起到积极的表率作用，GBase 8c 产品和技术支持团队在该银行丰富的业务场景基础上，配合完成了对微服务架构、分布式数据库、金融云应用等关键基础设施进行独立可控的创新研究，为稳步实现银行数字化、智能化、生态化建设目标做出技术探索与保障。

3. 帮助客户打造了完善的金融安全可控生态环境。

GBase 8c 数据库支持国内所有安全可控服务器、操作系统和存储设备，100%满足银行业客户对安全可控的要求，为客户核心业务系统的数据安全提供有力保障。在为用户提供更好性价比产品的同时，通过更全面的安全可靠能力向用户提供产品定制性开发、优化等服务，使产品更加成熟和满足实际业务需要，在保证业务系统处理效率的基础上，提供了良好的业务拓展性，也提高了金融高密级业务系统的安全性。

五、农村商业银行

案例二十：广东农信数据库联合实验室企业网银和银企直连 SUNDB 应用实践

一、应用场景

广东省农村信用社联合社与科蓝软件成立数据库联合实验室，旨在探索安全可控的数据库，助力银行的数字化转型。

广东省农村信用社联合社选取银企直连应用场景，基于科蓝软件 SUNDB 分布式数据库，以提升银企直连的功能完备性、对接便利性和多渠道协同能力为联合验证的范畴。目标是全面提升银企直连服务能力。建立对公服务渠道间的协同机制。实现企业网银、企业手机银行、银企直连和开放银行渠道间审批流程上的协同，提升直连对接效率。

二、总体方案

广东省农村信用社联合社在该项目中，基于我国的服务器产品和操作系统，中间件采用科蓝 PE 平台，数据库使用了科蓝完全自主知识产权的 SUNDB 分布式数据库。平台内部采用统一、规范的 Context 传递参数，将银企直连和企业网银进行整合统一。

为确保应用的正常运行和平稳过度，降低应用适配改造风险，数据库采用 DB2 和 SUNDB 分布式数据库并行。将部分相对独立模块的业务数据落地在科蓝的 SUNDB 数据库，其余部分数据还是保留在原 DB2 数据库中。

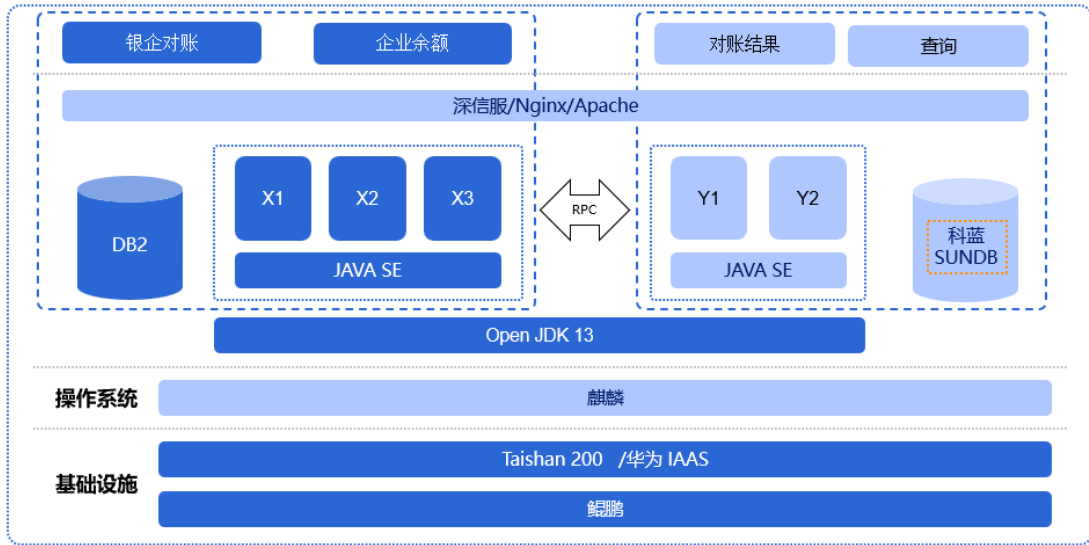


图 1 总体方案图

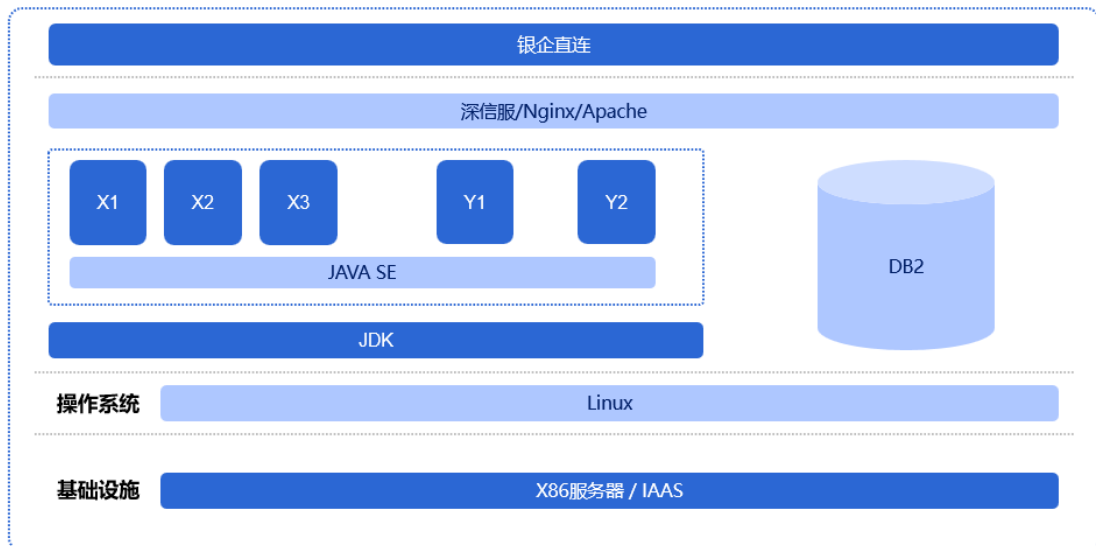


图 2 原始方案图

三、难点问题及应对举措

由于 DB2 和 SUNDB 数据库同时并行，数据分别保存在这两个库中，针对这两个数据库中的表有关联操作的时候，无法直接在数据库 SQL 层面进行关联，只能通过应用层面来实现。

1. 针对一个业务流程中有同时操作两个库表的场景，我们在应用代码层面通过回调方式保证事务的一致性。

2. 对原来直接多表联查（现在跨库多表联查）的，同时注入

DB2 和 SUNDB 数据库的数据源，在 SQL 层面进行拆分，查询之后再行组装，减少前端前台的应用改造。

3. 数据库性能、存储瓶颈和成本。采用传统集中式数据库 DB2，存在性能和容量的瓶颈上限问题，并且高度依赖高端设备，一般需要小型机和大型存储来保证数据库的性能和系统的稳定，采购成本、维护成本都比较高。通过使用 SUNDB 分布式数据库实现计算下沉、减少应用与数据库的交互次数，缩短业务处理链路，最终实现业务处理性能的优化。通过使用引入 SUNDB 分布式数据库，实现数据存储层的线性水平扩展支撑能力。解决存储瓶颈和成本问题。

四、应用成效及经验

采用基于国家信息安全要求的 ARM、X86 等创新 CPU 架构整机、操作系统的基础软硬件，银企直连通过科蓝 PE 平台产品和科蓝 SUNDB 数据库的结合。完成了银企直连由独立应用到与企业网银结合的平滑升级，丰富了银企直联能力，实现了企业网银、企业手机银行、银企直联跨渠道间审批流程上的协同，提升客户对接效率和对接体验，降低对接和服务成本，拓展了银企直联服务客群。

案例二十一：浙江农商银行“丰收互联”TDSQL 存算分离应用实践

一、应用场景

浙江农商联合银行是全省规模最大、网点覆盖最广、服务人员最多的金融机构，并打造丰收互联 APP 线上金融服务产品，从金融扩展到政务、民生、生活、商户等领域，将金融服务深入到每一个人的生活场景中。到 2023 年 3 月，月活跃人数超过 700 万。

浙江农商联合银行采用 TDSQL 承载“丰收互联”APP 应用，通过分布式数据库超强扩展性应对快速增长的访问需求。为应对日益增长的建设、运维成本，浙江农商联合银行联合 TDSQL 和华为存储，启动“丰收互联”系统存算分离改造。

二、总体方案

浙江农商联合银行采用 TDSQL 多主多备+存算分离方案，采用华为 OceanStor Dorado 全闪存共享存储存放所有数据。

数据库服务器集群包含 8 台服务器，包含 2 个 TDSQL 主节点，6 个备节点，其中主节点可提供读写，备节点提供只读服务。由于“丰收互联”查询类业务居多，2 主 6 备可提供最优性能。

8 台数据库节点数据全部存储在华为 OceanStor Dorado 全闪存共享存储上。OceanStor Dorado 可提供 RAID、双活等高可用保障能力，采用全 NVMe SSD 替代现网 SAS SSD，提供高性能

读写服务；此外，OceanStor Dorado 可提供慢盘隔离、故障预测等高效运维能力，确保业务系统稳定高效运行。

三、难点问题及应对举措

浙江农商联合银行改造初期采用基于服务器本地盘存储的存算一体架构，并很快面临三大问题。通过与华为存储、腾讯 TDSQL 紧密合作，逐步明确解决思路并解决问题。

一是业务连续性降级。业务连续性是金融行业的生命线。根据《银行业信息系统灾难恢复规划》要求，AB 类业务遭遇灾难时，业务中断需要低于 15 分钟，且数据 0 丢失。但由于服务器可靠性不足，故障后本地盘数据无法访问，分布式数据库只能通过多个服务器上存放多套副本提升可靠性，难以兼顾业务快速恢复和副本间数据 0 丢失，造成核心系统业务连续性降级。

二是能耗上升。双碳目标提出后，中共中央对金融业绿色低碳发展工作提出具体要求。但分布式数据库采用多副本，数据量急剧膨胀数倍，加上服务器盘无法共享导致被迫使用小容量盘，现网硬盘数量、服务器数量骤增，能耗快速上升，难以满足金融行业绿色低碳发展要求，也导致运营成本上升。

三是运维管理困难。服务器本地盘数量众多，当硬盘进入亚健康状态，服务器可能进入“半死不活”状态，即业务继续运行，但响应时延飙升；这种状态下，数据库不会触发主从切换，但整系统性能严重下降，需要人工发现并手动处理。由于服务器缺乏故障定位能力，管理人员需要反复进行插拔盘对比测试，甚至替

换掉所有硬盘才能定位亚健康盘，导致业务受影响时间长、运维成本高。

四、应用成效及经验

浙江农商联合银行采用基于华为 OceanStor Dorado 全闪存存储和腾讯 TDSQL 数据库的存算分离架构方案，彰显三大优势：

首先，企业级共享存储大幅提升数据库可靠性。相比于服务器，企业级存储架构可靠性更高，比服务器高出两个数量级，且不会因为服务器故障而导致数据丢失。当分布式数据库节点故障，只需要在健康节点拉起新的数据库实例即可完成补从，无需全量恢复数据，集群恢复时间从数小时降低至数分钟，且无需担心数据一致性问题。

其次，企业级共享存储大幅提升资源利用率。由于企业级存储大幅提升数据可用性，分布式数据库无需多副本，可降低至两副本甚至单副本，存储资源利用率大幅提升；另一方面，企业级存储每 U 空间容纳盘数比服务器多 50%~200%，且资源可供多个服务器共享，可大幅减少服务器数量，提升服务器资源利用率。

第三，企业级共享存储大幅提升运维管理效率。企业级存储可为 NVMe SSD 提供 RAID 保护，3 盘故障甚至整个硬盘框故障业务无影响，大幅降低运维事故发生风险；企业级存储也能提供硬盘故障的预测、告警及修复建议，大幅降低因硬盘亚健康问题带来的系统卡顿的几率以及故障定位及消除难度，运维效率大幅提升。

该方案通过节点故障、存储控制器故障、接口卡故障、三盘同时故障等多个极端故障条件测试，业务无中断；在“丰收互联”现网测试中，存算分离架构方案在查询内场景下平均性能提升50%。能效方面，该方案资源利用率提升30%+，成本降低40%+，并极大的降低能源消耗，达成绿色节能的目标。

金融信息化研究所

六、证券机构

案例二十二：中信建投 OTC 核心业务系统中达梦集中式 HTAP 应用实践

一、应用场景

中信建投 OTC 核心业务系统，应用场景主要特点是高并发 OLTP 交易业务为主与类 OLAP 业务（实时清算等）交织的 HTAP 新型混合业务场景，具体高并发查询、复杂事务逻辑的业务处理、大批量数据的 Merge 和 Update 高耗时业务、历史数据高耗时查询、以及高频率跑批业务数据分析等多种复杂业务交织的 HTAP 类型应用场景。

二、总体方案

证券核心业务系统选型中满足我国软硬件能力基础上，结合证券核心业务系统中 OLTP 为主和 OLAP 为辅的交织的 HTAP 混合业务场景特点，选择底层 ARM 芯片服务器，通过对三家不同数据库厂商的数据产品的性能、兼容性等能力的选型测试等，最终选择集中式主备架构的达梦数据守护集群 DataWatch 来处理证券核心业务系统的 HTAP 混合业务场景。

实施和上线过程中，最初从 2022 年 9 月开始选型测试，2023 年 3 月开始针对达梦数据库进行多轮功能和性能测试、进行分阶段的性能优化、故障切换测试、以及双轨迁移的数据实时同步工具测试等，目前已经完成了核心业务系统的部分模块的迁移上线，

还在快速稳定推进后续模块的迁移上线中。

三、难点问题及应对举措

针对证券核心业务系统中 HTAP 的多种业务场景，新型混合应用场景中众多业务特点的性能优化内容的侧重点不同，同一份数据不同的使用方式，需要平衡 OLTP 和 OLAP 之间的优化矛盾，如 OLTP 业务中为了提升性能创建索引，而 OLAP 业务上百万行表的索引反而影响性能等。在具体实施中，需要根据具体业务特点，有针对性的进行性能优化。

其次，HTAP 新型混合业务场景，资源隔离和限制尤其重要，针对此问题，此项目通过特定 CPU 架构，操作系统和数据库相互配合完成性能优化，以及通过资源绑定等技术，实现更好的资源限制和一定程度的资源隔离，保障了证券核心业务系统 HTAP 混合业务场景中 OLTP 和 OLAP 的稳定运行。

四、应用成效及经验

证券核心业务系统中 HTAP 新型混合业务场景，我国集中式数据库，经过软硬件优化和资源限制隔离等，优化后的核心业务系统实现了如下应用成效和经验：

1. 针对证券核心业务 OLTP 和 OLAP 交织的 HTAP 新型混合业务场景，我国集中式数据库满足了众多复杂业务需求，包括高并发查询、复杂事务处理、大批量数据的 Merge 和 Update 业务、历史数据高耗时查询、以及高频率跑批业务数据分析等。

2. 达梦数据守护集群实现了 OLTP 业务整体性能平均提升 75%

左右，OLAP 业务整体性能平均提升 90%左右，优化后整体平均性能与之前 Oracle 性能基本持平。

3. HTAP 新型混合业务场景的应用层面，通过规划不同业务逻辑先后顺序、交织出现 OLTP 和 OLAP，尽量减少了 OLTP 和 OLAP 业务同时操作，让复杂的问题变得相对简单。

证券核心业务系统在我国集中式数据库中的实践应用，提升了数据库的安全可靠能力。

金融信息化研究所

案例二十三：国泰君安证券新一代分布式核心交易系统 GoldenDB 应用实践

一、应用场景

当前证券行业核心交易系统架构多以应用加分库方案构建多套核心节点部署架构，面临着核心架构扩展性差、系统耦合度高、部署时间长等问题，整体架构过于复杂容易带来传导性风险，在信息安全保护的背景下，需要寻找可替代的 GCH 产品。国泰君安采用 GoldenDB 分布式数据库建设新一代分布式核心交易系统，有效解决上述技术难题，支持千万级全量交易客户，覆盖融资融券、现货、ETF、港股通等全交易业务品类，支撑 1500 万+账户的容量需求，并且支持全 GCH 平台部署。

二、总体方案

新一代分布式核心交易系统替换了此前 12 组 SQLServer 数据库分库分表构成的传统核心架构，同时支持低延时交易、横向扩展、异地灾备能力。整个系统包含交易、风控、行情等多个子系统，GoldenDB 承接来自于各个子系统的数据处理能力。

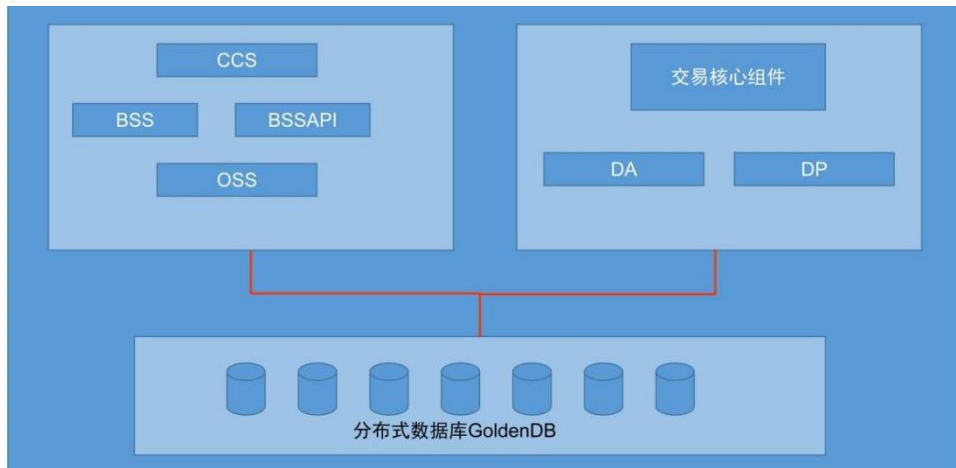


图 1 新一代核心交易系统架构

国泰君安交易平台与分布式数据库 GoldenDB 形成两条主线连接：一条是低延时交易组件通过 DA、DP 与 GoldenDB 进行交互，保证整个核心交易全流程；另一条是 BOS 运维/运营管理等系统与 GoldenDB 做各种对接复杂查询、修改等工作。上层分布式业务（含核心交易场景的数据上下场、复杂 SQL 查询等多种业务场景）将请求发送给 GoldenDB 的统一接口，由统一接口层下发处理，GoldenDB 承担两条路线下的语法树生成、路由分发、语句拆解、解析落盘等操作。此过程对应用透明，业务层面只需要做业务逻辑处理。

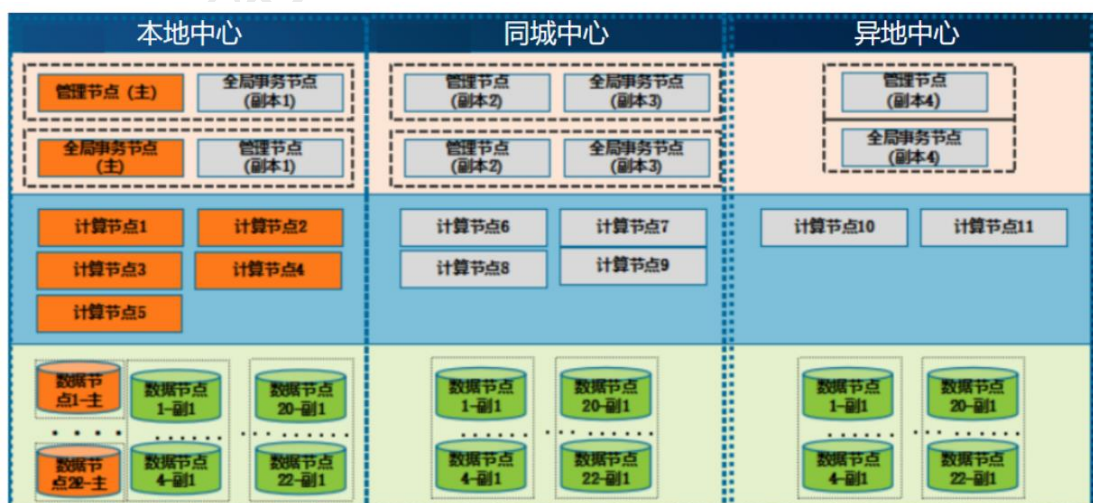


图 2 两地三中心高可用容灾架构

系统各业务组件及 GoldenDB 数据库基于两地三中心容灾架构部署，组件和数据库发生故障时可自动完成切换，具备机房级、城市级高可用能力。

三、难点问题及应对举措

通过 GoldenDB 分布式数据库 HTAP 能力解决证券复杂查询和联机交易混合处理难题。证券应用中 OLAP 类业务的复杂统计查询 SQL 开销高、耗时过长，直接影响 OLTP 类业务，导致后者卡顿。针对 OLAP 业务影响 OLTP 业务的问题，GoldenDB 引入 MPP 数据分析引擎，专门处理部分 OLAP 类复杂 SQL 语句，在 SQL 上加注解或者连接不同的计算层节点，将部分分析类 SQL 与 OLTP 业务在物理层面上隔离开，确保部分耗时过长的 SQL 不会影响到正常业务。

针对表关联时出现的耗时过长的的问题，例如分片表之间的关联，分片表与普通表之间的关联，会涉及到不同节点间的数据读取、传输与计算，性能表现不佳。通过对全量 SQL 的梳理、将部分配置数据表和基础信息表修改为全局表、合理设计索引和关联方式等措施，将 SQL 强制下推到底层数据节点，减少跨节点的数据交互，减少响应时延，有效解决了性能问题。

四、应用成效

历经一年多时间的努力，国泰君安基于 GoldenDB 分布式数据库完成新一代分布式核心交易系统在生产环境中的全栈软硬件适配验证，于 2022 年 8 月完成投产上线，实现 1500 万账户下

DP 整体持续写入性能 30 万行/秒、DA 持续写入性能 500 万行/秒的目标，助力新一代核心交易系统的能力提升。顺利完成投产目标后，为了适应证券行业核心系统的海量数据支撑需求，GoldenDB 进一步实现了高并发压力处理情况下实现低资源使用率、采用新传输协议提升大批量数据短时间快速上场效率、优化 HTAP 引擎满足超复杂 SQL 场景下并行计算下的低时延要求等系列技术革新。

国泰君安基于 GoldenDB 数据库实现的新一代分布式核心交易系统为证券行业核心系统首个案例，该方案可作为全行业推广，切实推动证券行业信息技术创新生态建设。

案例二十四：申万宏源证券阿里云数据仓库 AnalyticDB 应用实践

一、应用场景

国内领先的券商企业申万宏源证券成功完成了商业数据仓库 Teradata 向阿里云云原生数据仓库 AnalyticDB 的平滑升级，这一升级不仅实现了大数据分析、大赢家官方 APP、报表分析等多种复合业务应用场景的支持，还在确保安全可控的前提下，实现了成本和风险的双重降低，显著提升了工作效率和数据处理能力。

二、总体方案

项目总体迁移工作按照“统一规划，阶段实施，平稳有序，风险可控”的原则实施，分两个阶段完成原数据仓库到新数据仓库的迁移，更高效地为业务提供数据存储、计算、数据服务能力，满足企业未来业务快速发展的需要，项目采取分期实施。

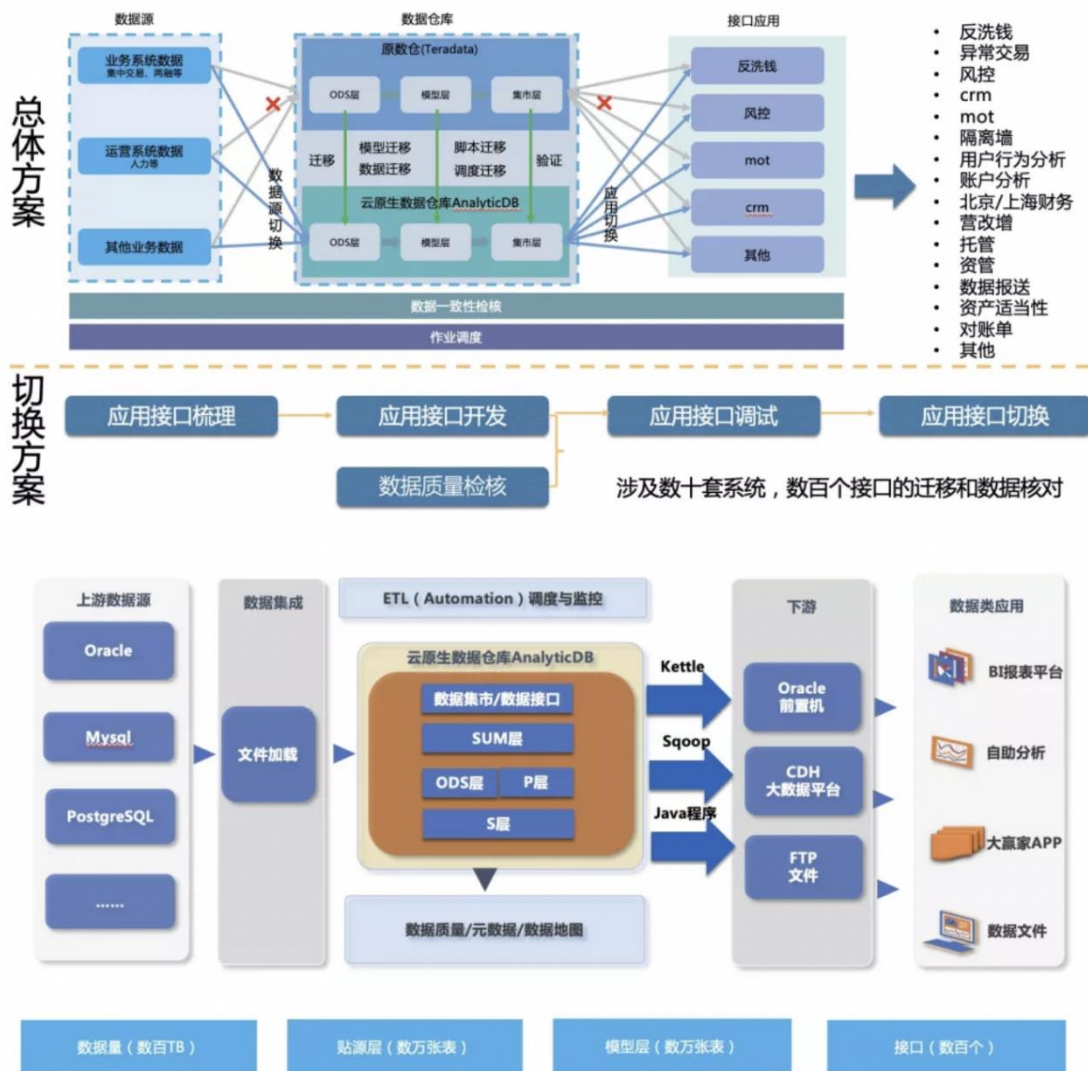


图 1 迁移方案架构图

整体方案中，降低实施风险和提高平滑度，对原有作业调度平台进行适配改造支持 AnalyticDB。利用阿里云提供的自动化迁移工具，将 Teradata 数据仓库迁移到 AnalyticDB，包括数据库对象和数据、用户以及权限体系；针对主要是实时数据文件集成，需要参照 Teradata 平台的实时文件集成程序提供公共程序，将 CDC 服务器上的实时数据文件集成到 AnalyticDB 平台的 AUD 审计表和贴源表。ETL 脚本迁移和接口文件导出程序迁移，参照 Teradata 平台模式提供公共导出程序，从 ADB PG 导出数据文件

对下游提供数据服务。

通过申万宏源的案例，阿里云针对 TD 的迁移，提供了一套完成的迁移实施方案，包括数据迁移、模型迁移、脚本迁移和数据校验。

三、难点问题及应对举措

本次实施涉及从 Teradata 数据仓库向阿里 AnalyticDB 数据仓库的迁移。在整体方案中，需要关注以下几点难点及应对措施。

首先是原有作业调度平台的适配改造。需要充分评估原平台的定制化开发情况，对其进行必要的技术改造，以适配 AnalyticDB 的功能和性能需求。在改造后确保作业调度与新数据仓库能够紧密联动。其次是 Teradata 数据仓库中的对象、数据、用户和权限的全面迁移。利用阿里云提供的数据库迁移工具，进行 Teradata 自动化迁移，迁移后验证权限和数据一致性，确保平滑过渡。另外实时数据文件集成程序需要重新设计实现。参考 Teradata 的设计思路，在 AnalyticDB 实现同类的实时集成组件，保证实时数据迁入。同时既有的 ETL 脚本和接口程序需要改造。保留 Teradata 中优化的脚本设计思路，在 AnalyticDB 实现同功能程序，经验证后使用。最后还需全面校验迁移后的数据完整一致性，按阿里云的标准流程，采用必要的程序或查询对比数据，以确保迁移质量。

通过以上关键措施，可以确保 Teradata 到 AnalyticDB 的平稳迁移，降低实施风险。

四、应用成效

通过采用阿里云云原生数据仓库 AnalyticDB 进行升级，申万宏源证券取得了显著的应用成效和宝贵的经验。

首先，AnalyticDB 的实时处理能力和高可扩展性大幅提升了数据仓库系统的性能，使申万宏源能够应对日益庞大的市场数据要求。数据处理性能提升了 1 倍以上，并能够支持数万张核心表，为申万宏源的运营和业务发展提供了可靠的基础设施环境。

其次，基于 AnalyticDB 等阿里云产品构建的数据仓库成为申万宏源数字化转型的重要底座。这个数据仓库不仅满足了申万宏源对大规模数据实时处理的需求，还提供了高度可靠的数据存储和计算能力。数据仓库为申万宏源构建智能化的风险管理、投资分析和决策支持系统提供了强大支持。

在实施过程中，申万宏源充分发挥了阿里云的先进技术和强大资源的优势，通过与阿里云团队的紧密合作，成功完成了数据模型从原数据仓库平滑迁移至新平台的转换。同时，在确保数据安全可控的前提下，申万宏源实现了成本和风险的双重降低，并显著提升了工作效率和数据处理能力。

申万宏源证券借助阿里云云原生数据仓库 AnalyticDB 的应用成效和经验是显著的。它不仅提升了数据仓库系统的性能，支持了大规模数据处理和分析，为数字化转型奠定了坚实基础。

七、保险机构

案例二十五：中国人寿数据库迁移实践

中国人寿保险股份有限公司数据库迁移方案融合了分层解耦的技术架构、多路并进的数据架构和标准统一的实施步骤。通过一年的艰苦努力，完成包括核心业务在内的全部业务系统数据库创新升级，实现从传统主备架构到分布式架构的技术革新，探索出可复制的大型金融企业数据库升级路径，带动了信息技术创新产业的快速发展。

一、应用场景

数据库迁移解决方案从全面应用技术革新出发，自主建立了一套分层分步实施的方法路径，实现数据库层的平滑迁移、企业技术架构的转型升级。经过实践验证，该方案具有高度的普适性，可用于常规单体系统，也可用于关联复杂的大型系统。

1. 填补经验空白：首创分层实施策略，配套归纳总结从评估、改造、迁移到运维的全流程实施路径，方案成熟可复制且工具完善，避免其他单位从 0 到 1 的探索投入。

2. 加速实施进程：一年时间完成数百个关键业务数据库迁移改造，帮助企业高速稳定实现数据库技术创新升级。

3. 适应多种场景：通过分层分步、以点带面的方法论，结合高效完善的工具栈，全面满足具有业务条线众多、关联调用复杂、用户访问量大等特征的企业要求。

4. 保证业务连续性：通过灵活的技术架构和完善的应急预案，保证数据库迁移过程中，业务不停、数据不断。

二、总体方案

整体方案分为技术架构、数据架构和实施步骤三个方面。

（一）技术架构

以分层解耦为主要特点。

应用系统按照基础硬件、操作系统、数据库、中间件、应用程序和客户端等分层，每个层次均可独立开展创新升级。对于数据库层而言，其重点是解除数据库与操作系统、中间件之间的耦合，避免数据库层“过重”导致难以升级。一是通过三层架构、平台化改造，简化应用和数据库开发模式；二是通过建立 PaaS 平台和应用微服务化改造，进一步解耦应用和数据库，使应用小型化、不再局限于使用一个大型的数据库；三是通过 DevOps、自动化测试等研发体系，极大提升了开发测试、系统集成和应用交付的速度。

（二）数据架构

以注重供应链安全、多技术路线齐头并进为原则。

根据业务特性和产品特点，对多类数据库产品进行应用场景化选型，形成多条技术路线齐头并进的格局。

1. 交易类：核心业务、销售支持、客户服务、经营管理等交易类应用系统，使用具备支持高事务处理及高可用能力的数据库产品，应用轻量化，不限定单一数据库产品。

2. 分析类：商务智能、监管报送、财务精算等分析类应用系统，使用具备海量数据处理及高可用能力的数据库产品。

3. 实时数据交换：交易类与分析类应用之间基于实时数据同步框架实现数据的流通，从而形成了基于新架构数据库的完整的企业级数据存储、流转和消费解决方案。

（三）实施步骤

以标准统一和风险可控为指导。

建立数据库迁移实施的统一规范和标准，遵循评估-实现-控制-分析改进的科学方法论，开展有序迁移，大幅降低单个系统的迁移风险。

整个实施步骤包括兼容性评估、负载评估、适配改造、存量增量迁移、反向回流、数据验证、持续监控和在线压测等 8 大环节，以及规范细化的 98 个子任务，结合系统上线前 5 倍峰值压测，保证了迁移风险最小化。

三、应用成效及经验

（一）技术优势

1. 自主程度：自主可控比例 100%，数据库使用达梦、Oceanbase、PolarDB、高斯 200 等；芯片使用海光、鲲鹏。

2. 降本增效：基于分布式架构的并行计算能力，资源使用效能和业务计算效率均提升数倍，数据库、服务器、存储等成本投入年均节省上亿元。

3. 容灾可靠：全面应用分布式技术架构，支持三地五中心部

署，满足不同级别、不同粒度的容灾要求，数据高可靠 RPO=0，业务高可用 RT0<5 分钟。

4. 风险可控：整体方案成熟可靠，迁移前评估充分，迁移中全面检查且高效可回退，迁移后监控完善且保障充分。

（二）应用效果

该方案实现公司仅用一年时间就全面完成核心交易、个险销售、多元销售、客户服务、经营管理、技术数据、办公管理等 7 大领域数百个数据库的优化改造，业务处理平稳连续，用户体验稳中有升，也入选了工信部组织的“2021 年数字技术融合创新应用优秀实践案例”。

1. 成本方面

1) 服务器设备选择基于通用硬件的机架式服务器作为机房上架标准，以分布式架构的容错能力来降低对高性能高可靠性设备的需求，年均节省上亿元。

2) 存储方面得益于新技术的数据高压缩比，迁移后核心业务数据容量下降 78%，大幅提升了存储的使用效能。

2. 性能方面

数据库层实现了弹性管理、敏捷交付、高效计算等特性。

1) 业务处理方面

联机交易、数据分析等核心指标迁移前后对比，普遍实行提升。其中，在线分析效率提升 2 倍，离线分析效率提升近 6 倍。

2) 资源管理方面

敏捷交付：数据库以资源池的方式管理和分配，提供统一便捷的管控平台，资源交付效率提升 5 倍以上。

弹性伸缩：数据库产品基于云原生架构，具备横向及纵向的弹性扩缩容能力，对应用透明且数据自动均衡，通过资源池的统一调度，应对业务冲刺等突发场景的伸缩效率提升 10 倍以上。

3. 运维体系方面

产品的运维能力建设已经成熟，满足金融级业务对易管理、高可用、高可靠等业务连续性运营的需要。

1) 容灾可靠性：灾备架构由主备模式提升为支持三地五中心部署，满足设备级、机房级、城市级等多级故障下的高可靠性保障。

2) 数据安全性：业务数据多副本存储，在分布式环境满足数据强一致性要求，RPO=0。

3) 业务连续性：数据库采用分布式架构，节点异常自动切换，RTO<5 分钟，生产事件处理时长从小时级缩短到分钟级。

4) 数据时效性：各数据库产品均具备数据流同步能力，且满足秒级时效性要求，支持企业数据高效流动。

(三) 示范效应

数据库迁移解决方案是一套充分验证、大量落地、完整成熟的方案。对于金融保险同业的信息技术创新应用都具有极大的参考和推广价值，通过与产业侧协同共建，有力带动了产品的快速成熟和生态的快速发展。

1. 全面升级

填补了数据库创新升级大规模实施经验的空白，基于大批量、高强度的应用实践，确定了 OceanBase-海光、达梦-海光、PolarDB-鲲鹏、高斯 200-鲲鹏的组合模式，以实际效果证明了新架构产品的可用及好用，探索出一条科学有效、安全弹性、可复制、可推广的成功路径。

2. 生态共建

促进新架构数据库的完善，帮助其产品从功能、性能、稳定性和周边生态方面发现和解决数百个问题和场景需求，公司提出产品优化需求数量占相关数据库总需求的 25%。

3. 产金共赢

该方案是金融行业与数据库产业携手共赢的典型案列，既完成了金融行业科技能力重塑再出发，也实现了相关产业场景孵化促成熟。

一方面，公司以此为契机实现了技术自主掌控和架构全面革新，通过科技全员参与，培养了数百名的新架构数据库专家；另一方面，数据库厂商获得了“金融首创、行业规模首屈一指”等具有标杆意义的实践经验，实施方案日趋成熟，产品能力更加完善，产品生态不断丰富，行业影响显著提升。

案例二十六：中国太平洋保险核心系统 OceanBase 应用实践

一、应用场景

中国太平洋保险 OceanBase 数据库支持多种保险业务，典型应用系统有核心客服系统、核心寿险理赔系统和保监会稽核接口系统。

客服系统涵盖子公司业务服务入口，为超过 2000 坐席提供系统服务，对接周边系统 200 余个。同时，系统提供 7*24 小时服务，系统可用性要求全年 99.9% 以上，因此是太保最复杂、最核心的系统之一。

寿险理赔系统是典型的 OLTP 系统，承担寿险保单理赔全流程功能，应用主要特点是以交易为主的作业流程，业务访问量大。

寿险保监会稽核接口系统是典型的 HTAP 系统，按照保监会稽核接口要求，后台批处理生成各类业务信息，并上报给保监会稽核检查。

二、总体方案

中国太平洋保险从 2021 年初启动分布式数据库的调研工作，从功能、性能、易用性、完整性、可移植性、可靠性、扩展性、安全性等指标进行综合评估，最终选择 OceanBase 数据库。

太保集团与 OceanBase 共同确定改造目标是“将传统商用数据库主备架构，升级为分布式架构，重点破除数据库与操作系统、中间件之间的强耦合”。

太平洋保险根据系统特点规划为“以交易为主”（如核心寿险理赔系统是典型的 OLTP 系统）和“以批处理计算为主”（如寿险保监会稽核接口系统是典型的 HTAP 系统）的两类数据库集群，分别管理分布式数据库资源。为避免集群过大造成的合并开销，集群规模多采用 2-2-2，最大不超过 6-6-6，租户的 Primary Zone 交错分布在不同 Zone 服务器上，以充分利用服务器资源，保障交易集群的业务响应性能。

三、难点问题及应对举措

太平洋保险 Oracle 数据库的应用存量、使用复杂度高。在代码不改动或者少改动，并保障业务系统稳定安全的前提下，进行分布式数据库的升级改造，难度极大。迁移的核心系统“P17 客户服务系统”，业务场景繁琐，又是 7*24 小时服务平台，对高并发和业务稳定性要求极高。该系统历史悠久，使用多个技术框架，涉及众多上下游系统接口。同时，系统与 Oracle 产品耦合度高，存储过程总代码量近百万行。除数据库外，配套使用的周边产品对 Oracle 也有深度依赖，适配改造复杂度极高。

OceanBase 数据库对 Oracle 语法的强兼容特性，在迁移过程中起到了重大作用，兼容度高，降低了应用改造工作量。与此同时，利用自研的数据库应用改造评估工具“指南针”，自动识别改造项，并给出改造建议，缩短了人工识别改造点时间。

四、应用成效及经验

截至 2022 年，太平洋保险已有包括核心应用系统在内的 80

余个应用系统使用了 OceanBase 数据库方案，覆盖各种保险业务、运营及客户服务的环节。

应用系统利用数据库系统的无损容灾和异地多活技术，将 Paxos 一致性算法运用于数据库主备副本之间的同步，在多数派副本正常的情况下，保证数据零丢失，服务不中断。自投产上线以来，持续平稳运行，广泛服务于数千名柜面人员、百万业务人员和亿级外部客户，不仅为用户提供了高效稳定的应用体验，在提升业务效率、节省存储空间、降低采购和运维成本等方面成效显著。

其中，“P17 客户服务系统”，实现业务成功交易率不低于 99.99%、交易总平均响应时间小于 1 秒、整体并发量 1500 人等预期指标。寿险保监会稽核接口系统，利用 OceanBase 实现单一引擎支持高性能混合负载（HTAP）应用，并通过基于时间片的混合负载调度技术，解决混合负载的资源隔离问题，相同计算资源的批处理操作时间节省 62%，监管报送批量场景性能提升 3 倍。寿险营销员系统，利用 OceanBase 水平扩展的分布式事务处理技术能力，将数据均匀分布到多台机器，保证分布式事务特性的前提下，解决单机计算资源不足无法满足需求的问题。智能决策服务平台和寿险统一承保平台，利用 OceanBase 单集群内同时兼容 MySQL 和 Oracle 两种主流数据库生态，智能决策服务平台（MySQL 租户）与寿险统一承保平台（Oracle 租户）共用集群，实现分散业务系统的整合，最大程度降低现有系统的扩展和迁移改造成

本，保护技术无形资产。

太平洋保险 OceanBase 分布式数据库方案投产后，数据库软件维保费用大幅降低，业务系统高性能、高稳定性地运行。利用 OceanBase 的高级压缩技术，存储容量平均节省 80%以上，分析型数据加工处理能力提升 10 倍，每年可节省设备投入数亿元。升级后的应用系统，构建起全面实时数据处理和服务能力；弹性扩缩容、处理速度、数据加工能力均实现大幅提升，为太平洋保险的数字化转型奠定了坚实的技术基础。

FITI 金科智库，一揽子解决金科难题
更多详情，欢迎前往金融信息化研究所公众号



中国金电集团



金融信息化研究所



FTPC 金科智库

网站: www.fitipbc.com

邮箱: yjs@icfcc.com