



金融分布式事务数据库白皮书

中国信息通信研究院云计算与大数据研究所

中国支付清算协会金融科技专业委员会

2018年4月

版权声明

本白皮书版权属于中国信息通信研究院云计算与大数据研究所和中国支付清算协会金融科技专业委员会共同所有，并受法律保护。转载、摘编或利用其它方式使用本白皮书文字或者观点的，应注明“来源：中国信息通信研究院云计算与大数据研究所、支付清算协会金融科技专业委员会”。违反上述声明者，本院将追究其相关法律责任。

牵头编写单位及人员：

中国信息通信研究院云计算与大数据研究所：魏凯、姜春宇、马鹏玮
中国支付清算协会金融科技专业委员会：丁华明、颜勇

参与编写单位：

中信银行：陈蓓、刘文涛、童庆峰
北京银行：马晓煦、于振华、怀文佳
建设银行：刘征宇、舒展、邢磊、白彧斐、刘海、安兴朝
光大银行：赵雪枫
民生银行：郝庆运
华为技术有限公司：叶涛、李海丰、朱松
中兴通讯股份有限公司：罗圣美、钱煜明、邹海丽、张丽、周日明
蚂蚁金融服务集团：蒋志勇、徐岩、李阳
PingCAP：余军、郑万波、房晓乐
广州巨杉软件开发有限公司：郝大为、孙伟、李方舟
上海热璞网络科技有限公司：金官丁、黄小慧，刘小成，魏晗清
腾讯云计算（北京）有限责任公司：潘安群、孙康、刘峰、苏强、吴月玲
上海爱可生信息技术股份有限公司：李恒、梁广涛、陈书俊、张翔、王松强

前 言

当前，金融信息化建设主要依托原有集中型 IT 架构进行维护扩展，系统规模及复杂程度呈指数级增长，各类瓶颈逐渐暴露，日益增长的数字金融需求同旧式的系统架构缺陷之间的矛盾愈加凸显。中国人民银行、中国银行保险监督管理委员会等金融监管部门逐渐推出分布式转型政策要求，金融企业开始兴起分布式转型浪潮。

本白皮书首先阐述了分布式事务数据库的发展状况与相关技术体系，之后以金融业务为背景，从金融业务对分布式事务数据库的能力要求、分布式事务数据库对现有金融业务影响、金融业分布式事务数据库迁移方法三方面进行分析论述，力求梳理金融业应用分布式事务数据库全流程概念及方法。白皮书结合当前应用现状及问题，提出了金融分布式事务数据库发展的策略建议。由于时间仓促，水平所限，错误和不足之处在所难免，欢迎各位读者批评指正，意见建议请发送至 mapengwei@caict.ac.cn。

目 录

版权声明.....	2
一、分布式事务数据库概述.....	1
(一) 信息系统向分布式转型的趋势.....	1
(二) 分布式数据库发展历程.....	2
(三) 分布式事务数据库的核心属性和能力.....	6
(四) 分布式事务数据库的技术架构.....	7
(五) 集中式事务数据库与分布式事务数据库对比.....	10
二、金融行业事务数据库现状概述.....	11
(一) 金融行业业务需求分析.....	11
(二) 金融数据库系统现状及痛点分析.....	12
(三) 分布式事务数据库成为转型必由之路.....	15
三、金融关键业务对分布式事务数据库的要求.....	16
(一) 基本功能.....	17
(二) 兼容能力.....	17
(三) 管理能力.....	18
(四) 高可用能力.....	18
(五) 扩展能力.....	19
(六) 安全能力.....	19
四、分布式事务数据库对金融业务设计的改变.....	19
(一) 业务系统组件化、服务化改造.....	20
(二) 数据库模式设计，满足扩展需求.....	20
(三) 数据库 SQL 能力评估及业务开发注意事项.....	21
(四) 系统容量评估及弹性方案.....	22
(五) 高可用及容灾方案.....	23
(六) 系统运维方案.....	24
五、金融分布式事务数据库迁移策略.....	25
(一) 指导原则.....	25
(二) 实施策略.....	26
(三) 典型问题和误区.....	29
六、结论与建议.....	31
(一) 建立技术标准体系.....	31
(二) 开展技术试验与测试.....	31
(三) 在金融机构开展试点.....	32
(四) 完善支撑配套体系.....	32

一、分布式事务数据库概述

（一）信息系统向分布式转型的趋势

互联网业务爆发式发展催生信息系统架构转型需求。电子商务的迅速发展引发全民购物狂欢浪潮，互联网金融让日常支付更加便捷，用户形成了从现金付款到数字货币结算的消费习惯。这为银行业、支付机构和金融监管机构带来海量的小额资金交易笔数，衍生出海量活跃用户、数据以及并发量的业务场景。据统计，淘宝网“双十一”交易量从2009年的0.52亿元增加到2017年1682亿元，8年时间增长了大约3000倍，而在2011-2014年“双十一”过程中，16家上市银行中8家网银支付出现拥堵现象。业务信息系统急需适应当前业务海量用户的数据服务和存储处理需求。

从集中式架构向分布式系统架构转型探索取得积极进展。当前大型企业如金融银行等，大都采用集中式的数据库架构，多采用 scale-up，即增加单机硬件性能，增加处理能力上限。而随之摩尔定律的逐渐失效，scale-up 方案出现瓶颈。新兴互联网企业因为成本和性能等原因，绝大部分开始了自主探索分布式数据库架构的道路。以支付宝为例，2013年全线下架 IBM 小机和 EMC 存储设备，改用廉价 X86 服务器，2014年使用自研数据库方案替换 Oracle 数据库，完成传统 IOE 集中式架构向分布式架构转型过程，并成功经过了 2015-2017 年三年的线上业务考验。据报道，相应软硬件成本下降大约 60%，而性能提升数十倍。故通过多年的实践经验和验证，分布式架构的数据库已经可以支撑当前互联网企业的大规模、高并发、多模式的业务类型，形成了

极具价值的参考经验。

金融监管机构相关文件指出了金融信息系统向分布式架构转型的方向。2017年6月，中国人民银行《中国金融业信息技术“十三五”发展规划》指出在传统金融机构在新技术应用和分布式架构转型方面进展较慢，提出在“十三五期间”，要“在巩固集中式架构安全稳定运行的基础上，综合成本、效率、资源等方面，以业务适用性为原则，研究分布式架构应用的可行性”。2016年7月，原中国银监会发布《中国银行业科技“十三五”发展规划监管指导意见(征求意见稿)》也明确提出，要“提升技术架构的精细化水平，基于开放性强、透明度高、适用面广的信息技术提升架构设计和管控能力，鼓励合规使用开源技术和正版软件，鼓励使用创新的分布式架构。”

(二) 分布式数据库发展历程

数据库(Database)是按照数据结构来组织、存储和管理数据的仓库，它产生于距今六十多年前，随着信息技术和市场的发展而不断演化。从分布式化角度，目前已经历经4个主要的技术发展阶段，即分析型系统与事务型系统分离阶段、硬件提升带动数据库能力阶段、非关系型数据库发展阶段、分布式事务数据库发展阶段。本白皮书同时预测未来分布式数据库将会朝着分析事务混合型技术阶段演进。

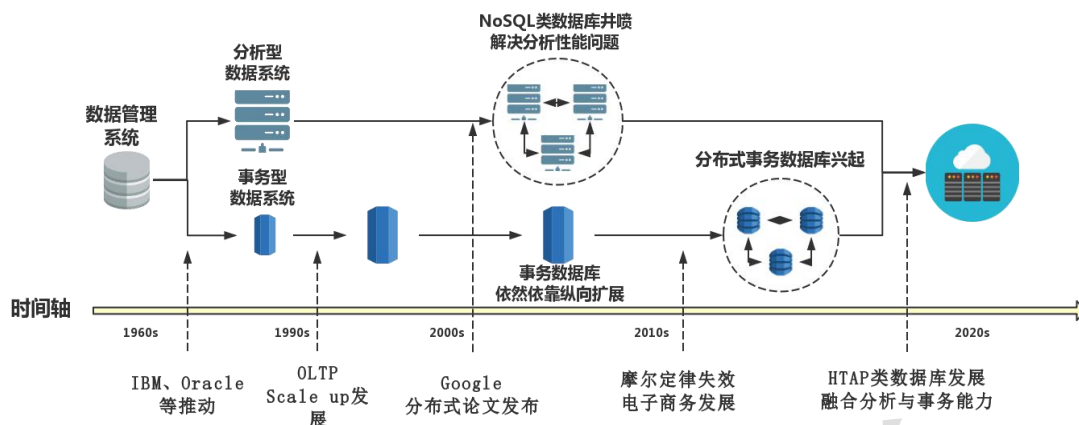


图 1 分布式数据库发展历程

首先在 1960 年代，数据管理系统已经得到了一定程度的应用。伴随着关系数据库理论和技术的诞生和普及，以最初的技术创始和发展者为主的 IBM 和 Oracle 公司及相关的数据库科学家和研发人员逐渐开始在各种数据库学术及产业会议上，使用联机事务处理 OLTP (Online Transaction Processing) 来定义关系数据库的主要应用场景，于此相配套产生了针对分析型业务场景的定义，即联机分析处理 OLAP (Online Analytical Processing)。这样的理论定义共识和产业标准公司，使得市场在应用关系数据库的时候，有了清晰的应用场景区分的定义。从而对于关系型数据库本身的技术分类和专业化，场景化方向发展，起到了巨大的推动作用。相应的 OLTP 与 OLAP 业务场景对比如下表所示。

表 1 OLTP 与 OLAP 对比

	OLTP	OLAP
时延要求	实时	大部分不需要实时
业务类型	高频的海量小事务	低频的大规模数据分析
保障能力	实时高可用能力	可以适当降低高可用特性

并发特性	大用户量、高并发	小用户量、低并发
热数据特性	热数据随机分布	热数据集中分布
索引特性	索引依赖	大部分不依赖
数据管理	需要生命周期管理	不需要生命周期管理

至此产业内形成了关系数据库体系和非关系数据库两种主要应用场景的清晰的划分。数据库理论体系、工程化及产业体系，在关系数据库的基础上，针对这两大场景，开始了半个世纪的理论和技术实践创新。并围绕着这两个场景，产生了细分领域和细分技术，使得关系数据库产业得到技术专业性的革命性的提升。因 IBM 和 Oracle 等公司在商业上的推动策略，早期绝大多数关系数据库管理系统均针对 OLTP 联机事务处理场景研发，典型的 OLTP 场景包括银行核心事务和交易，零售核心数据处理等。

其次从 1990 年代开始，随着互联网的崛起，海量数据的处理，海量数据的计算和分析需求越来越普遍，在各种计算机应用场景中，对于关系数据库系统不断提出新的挑战和要求，关系数据库系统技术面临着理论升级和技术升级的迫切需求。此时关系数据库主要依靠纵向提升数据处理硬件能力上限，即“scale up”方式，满足逐渐增高业务场景需求。

接下来 2000 年代，随着谷歌公司三篇围绕大数据相关的论文的发表以及以 Hadoop 技术为代表的非结构化大规模数据处理技术的崛起，在 OLAP 领域，数据库系统似乎找到了如何在海量数据，海量增长，性能，维度等多方面要求下，突破原有单机物理限制，迈向以分布式处理为核心思路的进化方向。这得益于对传统关系数据库的数

据模型的重新认识、挑战和取舍。在一段时间内，包括大数据在内的 NoSQL (Not Only SQL) 思想和技术大行其道。其解决大规模数据管理，操纵和处理的核心思路是：

- a) 利用 X86 架构的计算设备横向扩展，突破单设备物理限制；
- b) 妥协甚至抛弃关系型数据库管理系统的关系模型，聚焦在这对非结构化数据模型下的数据处理需求。

而在联机事务数据库 OLTP 数据库领域，利用产业界已有的分布式计算和存储体系，横向突破单机本身的计算和存储容量限制则成为一个非常有挑战性的课题。这些问题集中在：

- a) 遵从现有的关系数据模型遵从的前提下，在一个分布式架构中，如何实现标准数据库事务处理的原子性，一致性和完整性；
- b) 如何在分布式系统的 CAP (Consistency、Availability、Partition tolerance) 约束下，达成对业务一致性和可用性的平衡；
- c) 如何建立性能和存储这两个核心的 OLTP 数据库诉求的设计策略。

综上所述，分布式事务数据库的发展，是在传统联机事务关系数据库基础上，遇到了比联机分析处理数据库 OLAP 数据库更具有挑战性的来自理论和工程实践的要求。而业界在实现分布式事务数据库技术的历程中也经过了若干的技术演进阶段，采用了不同的思路 and 手段。

经过 10 多年的数据库学术界和产业界的共同努力，业界通过将分

布式计算这个独立学科与数据库学科的有机结合，交叉影响和推动。到今天为止，新一代分布式事务数据库无论在理论体系和工程实践上已逐步完善和健全，成为下一个 10 年数据库大发展的基石。

(三) 分布式事务数据库的核心属性和能力

目前，业界还没有分布式事务数据库（Distributed Transactional Database）的权威定义。本白皮书认为，分布式事务数据库应具备的属性和能力可概括如下：

(1) 3 大核心属性：

- a) 事务数据库属性。即数据库核心数据逻辑管理模型具备事务能力，遵从完整的 ACID（原子性、一致性、隔离性、持久性）属性要求；
- b) OLTP 联机交易能力属性。即能够完整处理联机事务处理负载的关系数据库管理系统；
- c) 分布式架构属性。计算和存储均通过分布式技术分布于不同的节点上。

(2) 6 个核心能力：

- a) 计算和存储能够基于 X86 等工业标准硬件实现；
- b) 遵从事务数据库的关系模型，完全满足事务的 ACID 的要求；
- c) 支持高并发 OLTP 工作负载；
- d) 原生支持标准的 SQL（标准查询语言）数据操纵界面；
- e) 自动化故障自愈与切换的高可用性保障；

f) 支持跨数据中心站点级别的集群弹性横向水平扩展能力。

(四) 分布式事务数据库的技术架构

业界在这几个方面通过各种尝试和技术手段，不断进行问题求解和进化。数据库理论界，陆续的提出新的数据管理模型，试图根本解决相关问题。而数据库产业界，通过各种工程化研发手段，在数据库的设计架构上寻求突破，综合过去 10 年内，在分布式事务数据库领域的学术和产业界工作，可以梳理出以下几种发展方向：

表 2 分布式事务数据库技术方案

一级分类	二级分类	三级分类	特征	典型产品
基于单机事务数据库	基于业务改造		业务层人为实现分布式相关特性，对业务层强侵入	基于 Oracle 的业务流程改造
	基于中间件改造	嵌入应用层的中间层实现分布式相关特性	对业务层一般侵入	开源 TDDL
		独立的中间件实现分布式相关特性	对业务层弱侵入	开源 MyCAT
基于新型数据库理论			新型计算存储算法及工程实现分布式相关特性，对业务层弱侵入	谷歌 Spanner

1、基于单机事务关系数据库的业务分布式事务改造方案

利用原有单机事务处理数据库的成熟度优势，通过在独立应用层面建立起数据分片和数据路由的规则，建立起一套复合型的分布式事务处理数据库的架构。此类方式需要 DBA (Database Administrator) 手动控制应用流量的分发和管理，业务层侵入较严重，但能够缓解单

一集中式数据库存储和计算能力不足的劣势。

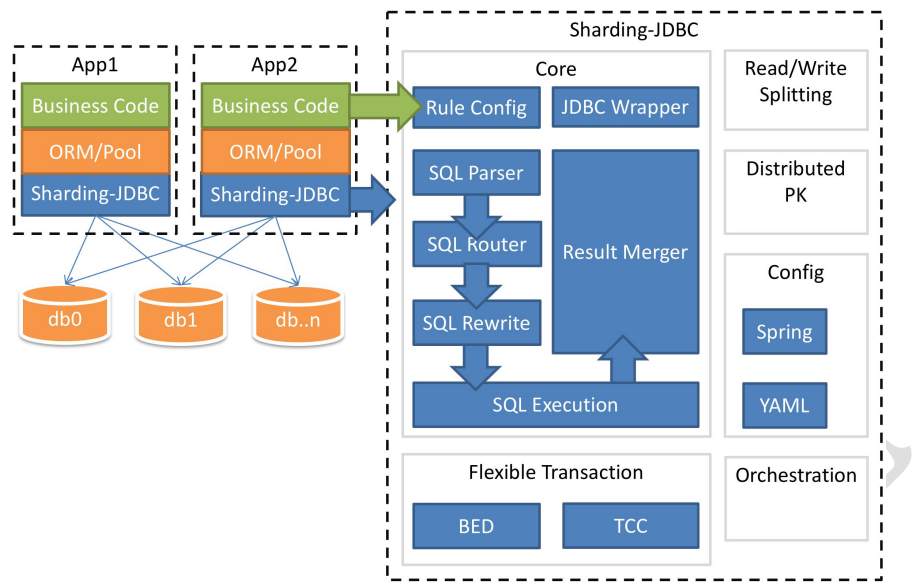


图 2 基于单机事务处理系统的分布式架构

2、基于单机事务关系数据库的分布式管理组件方案

将原有的单机事务数据库的引擎层设计架构进行改进，利用独立的分布式控制逻辑单元实现自动路由策略，同时通过一致性方面的配套理论和算法，解决分布式事务数据库的一致性问题。主要有嵌入式中间件方式与独立中间件方式两种实现方式。

嵌入式中间件方式，以 DAL（Data Access Layer，数据访问层）管理数据分片规则和路由策略。这里的 DAL 层并非独立的中间件，更近似于嵌入业务方的中间层模式，通常以 Jar 包方式提供给应用系统进行调用。这里以开源 TDDL 架构为例。

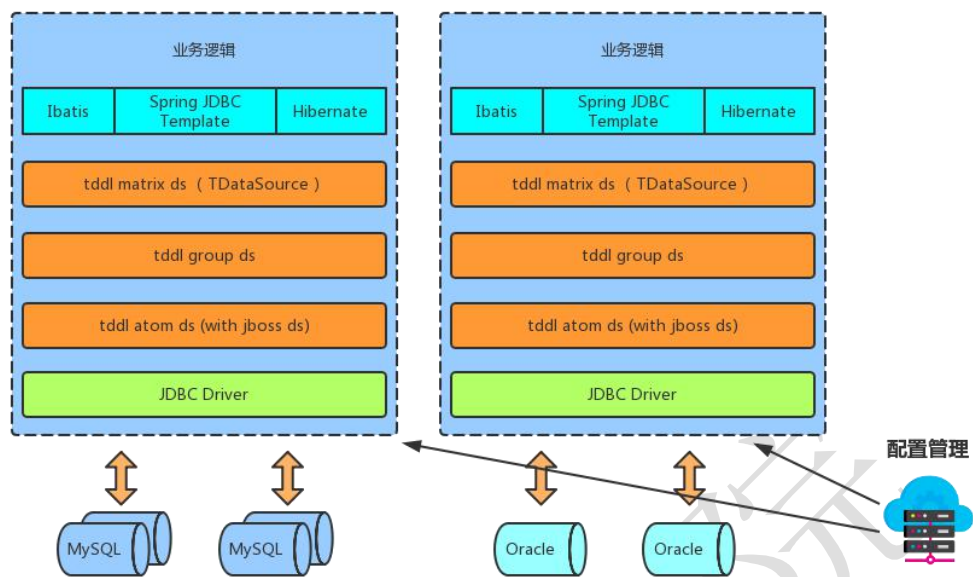


图 3 开源 TDDL 架构示意

独立中间件方式。以代理服务器机制管理数据分片规则和路由策略。中间件通常为独立逻辑或物理处理系统结构，与业务层和数据存储计算层隔离，聚焦数据分布管理功能。这里以开源 MyCAT 架构为例。

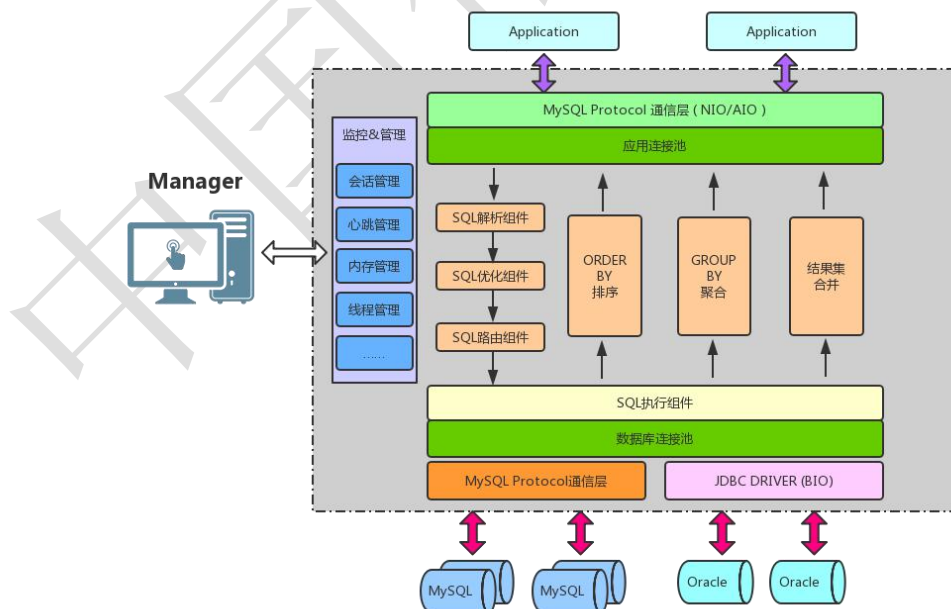


图 4 开源 MyCAT 架构示意

3、最新分布式数据库理论和工程实现架构

该类数据库通常由数据库驱动、计算节点集群、数据节点集群、全局管理节点四个部分组成。数据的多个副本按定制化或自动优化的策略存储在数据集群的多个节点上，数据节点之间不共享数据，各副本之间通过自动一致性算法实现同步。计算节点负责分布式查询处理和全局事务控制。全局管理节点是实现分布式事务一致性和维护分表分区、拓扑关系的必要部件。节点之间通过高速互联的网络通讯，完成对应用数据请求的快速处理和响应。抽象该类数据库架构图如下所示。

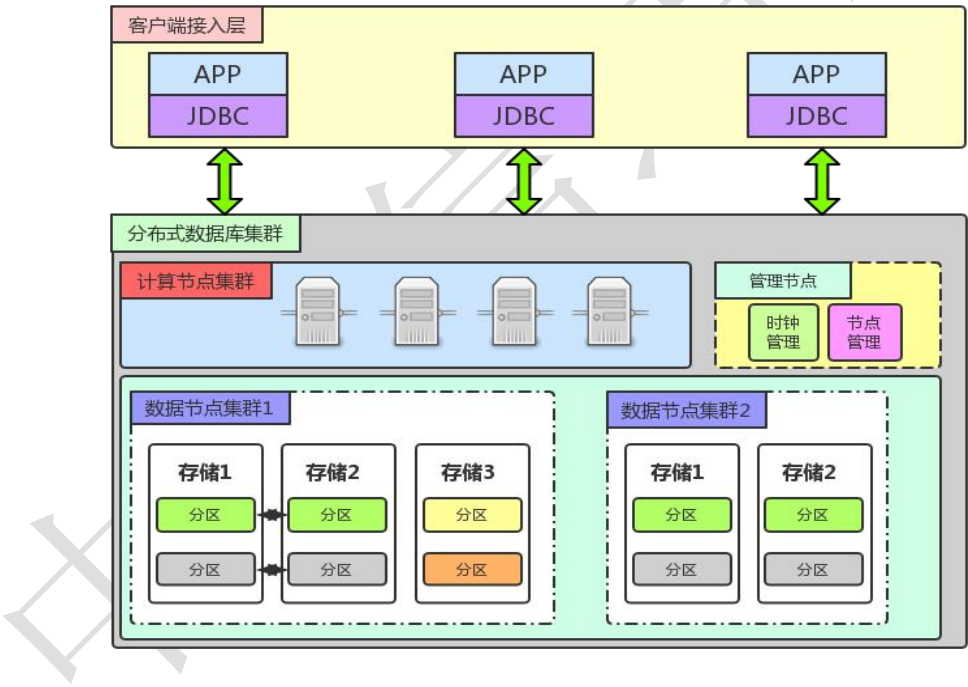


图 5 基于新一代理论的分布式架构

（五）集中式事务数据库与分布式事务数据库对比

此处从几个大致的方面对分布式事务数据库与集中式事务数据库进行一个大概的对比，从而明确分布式事务数据库相对现有集中式事务数据库的优势和区别。注明此处的分布式事务数据库指代能够实

现自动化分布式相关特性的技术方案，不包含手工实现分布式相关控制的技术方案。

表 2 集中式与分布式对比¹

	集中式事务数据库	分布式事务数据库
业务支撑能力		
经济性	软硬件价格昂贵	合理可控
灵活性/兼容性	应用新业务限制较多	灵活方便
扩展性/伸缩性	只能垂直扩展	可以大规模水平扩展
可用性、一致性和可靠性		
可用性	存在单点隐患	分布式备份
一致性/可靠性	单机事务可靠	强/最终 一致性保障
运维和故障恢复能力		
维护性	简单集中管理	需要实现自动化管控
业务恢复	异步备份切换	无感知自动切换

二、金融行业事务数据库现状概述

（一）金融行业业务需求分析

当前，随着普惠金融、数字金融的快速推进，数据能力已经成为金融单位在新时代业务能力的重要抓手。与此同时，移动互联网和电子支付业务的蓬勃发展，给金融行业的典型应用场景，例如核心账户与账务交易、在线支付/移动支付交易业务、实时交易监控与指标分析等，带来了新数据形态下金融系统能力需求。

一是金融行业的数据急剧增长，这对数据存储和管理提出了更高要求。以大型商业银行为例，通常拥有成百上千个业务系统以及上亿

¹ 引用蚂蚁金服团队观点

用户的海量数据，且数量呈现指数级增长，从 TB 级别增加到 PB 级别，未来很快就会增加至 EB 级别，这些都需要有效的管理以及实现实时访问。

二是金融行业对业务连续性能力更加重视。当前，金融行业对 IT 端安全可控要求不断提升，在确保数据不丢失的同时，还要做好数据的容灾备份以及核心数据系统的“双活”，即在两个数据中心实时进行备份，一旦失去一个中心，所有业务可以及时切换中心继续运行。

三是需要面临高并发业务和高用户量带来的系统压力。随着互联网金融业务的开展，金融行业用户量和用户自身应用权限不断增加，持续面临高并发业务和高用户量带来的系统压力。比如，过去银行主要通过柜面接收账单查询，现在则主要依托移动端用户自助查询，在某一峰值的操作量就可能超过过去一整天业务操作量的总和。

四对移动应用响应速度要求更快。在移动互联网用户应用需求迅速增加的背景下，用户对移动应用的响应速度有了更高要求，移动端响应必须做到秒级以下。除此之外，金融行业自身也需要提升数据技术能力和业务价值，增强自身运营管理能力。

五是技术层面的国产化需求。近年来，金融行业在技术方面正在积极进行国产化改造。金融业在稳定性和安全性上对数据产品要求近乎严苛，但海外产品和开源技术不能完全满足国内用户需求，所以亟需能满足核心系统需求的真正国产自主数据库产品。

（二）金融数据库系统现状及痛点分析

数据库系统作为 IT 基础架构的重要组建，在几十年的发展过程中，

已经成为上层金融业务架构高效建设与发展的支撑组件，以及顶层愿景和战略顺利实施的重要保障。在数据库技术生态和服务体系发展过程中，金融行业内逐渐形成以集中式架构事务数据库为主的格局。

首先在硬件方面，以 **IBM、HP** 为主的大型机和小型机拥有较大的市场份额。其中根据金融机构自身规模及业务的不同，各种硬件产品占比有较大差异。在国有大型银行中，以 **AS390** 为代表的大型机能够达到一定数量占比，数量在百台左右，用于保障现有业务的性能及稳定性。而在中等规模的股份制银行或城商行中，业务量相对国有大型银行较小，同时成本影响因子有所提升，故以 **AS400** 为代表的小型机为主要硬件机型，数量在千台以下级别，负责绝大部分金融业务，部分银行也会同时配置数量较少的大型机用于核心业务的处理。同时，各银行仍会使用一定比例的 **X86** 服务器，用作外围业务的支撑扩展。

其次在软件方面，目前各类业务主要运行在以 **Oracle、DB2** 为主的商用数据库软件系统上。据统计，各类数据库在金融体系内占比如下图所示。其中，**Oracle** 以其稳定性、功能黏性、服务保障体系完善性占据大比例市场份额，而 **DB2** 以一体机捆绑销售方式，也根植于银行现有数据库体系中。而开源 **MySQL、PostgreSQL** 数据库近些年开始服务于大量互联网核心业务或传统金融外围业务，正在加速扩张。同时随着大数据业务在金融领域的发展，**Greenplum、Teradata** 等 **MPP** 和数据仓库类系统也开始占据一定比例。

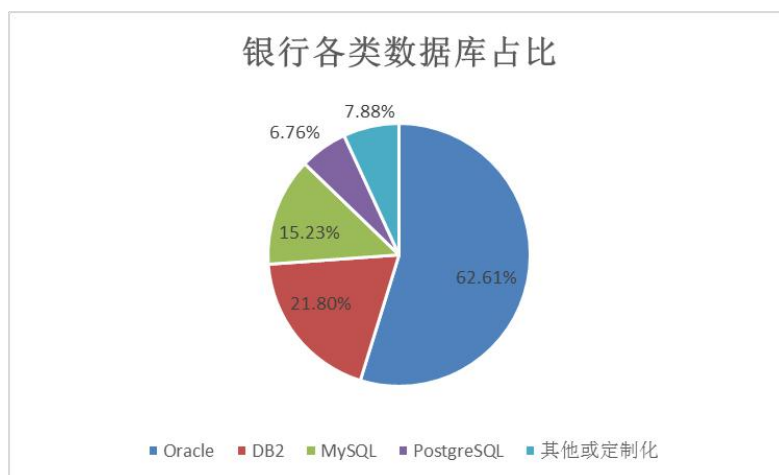


图6 银行各类数据库占比（数据来自信通院）

但由于新时代下金融业务的发展更新，数据库技术能力也面临多方面的发展需求。

第一，亟需提升管理弹性。当金融单位系统内数据量急剧增大时，系统需要弹性地扩容以应对PB级别以上的数据管理，这种弹性容量调整可以实现让所有数据保持在线。

第二，亟需提升处理性能。针对客户的实时需求，金融单位数据库系统需要满足高并发业务操作需求，实现海量数据超高性能读写以及实时访问查询。

第三，亟需升级可靠性能力。数据安全不仅仅是简单的备份，除了实现数据的持续高可用外，还应支持异地容灾甚至数据中心“双活”，进而保障数据安全。

第四，亟需满足多类型数据处理需求。在跨业务的融合中，亟需实现对多模数据的统一管理，从结构化数据到半结构化数据再到非结构化数据，进而实现不同类型的数据统一融合管理，从而大大提升系统效率。

第五，亟需增强开发运维能力。随着应用的增多，更需要强大数据库运维能力的支持，进行数据分区管理，实现业务有效隔离。同时，保持系统的弹性、兼容性，大大简化运维开发。

第六，亟需有能力支撑核心业务系统的自主可控产品。除了数据安全的要求和“技术先进”，对于核心业务，更重要的是对产品代码级具有控制力，这样可以保证产品针对用户共性需求不断进行迭代更新，也带来产品稳定性以及高效强力的技术支持。

（三）分布式事务数据库成为转型必由之路

随着互联网行业在自身业务场景的探索与实践，分布式事务数据库开始广泛应用在各种高并发，大数据量，联机事务工作负载（OLTP 及 OLTP/OLAP 混合负载）要求强一致性且需要强大横向扩展能力的业务场景。

分布式事务数据库采用多种模式实现数据的分散存储，将数据库压力分散到不同服务器上。与集中式事务数据库相比，分布式事务数据库可以均衡交易负载，并采用高并发的架构提升系统的交易处理能力，而其统一的资源管理机制也使得数据库的性能扩展不再是设备的替换式升级，而是通过增加存储或计算节点来实现弹性升级，极大地节约了升级成本。

概括应用分布式事务数据库的几大优势如下：

- a) **高性能：**分布式事务数据库能轻松面对海量数据和高并发的请求处理，做到数千以上并发、数万以上连接、数万 TPS、数十万 QPS 等；

-
- b) 弹性扩容：分布式事务数据库通过提供在线扩容、在线迁移等多种弹性扩容方式，应对业务变更及不定期扩容缩容要求；
 - c) 多中心：分布式事务数据库能够实现单个机房内数据同步并保持数据一致以及数据机房级容灾，主要借助分布式数据库的跨数据同步组件技术；
 - d) 低成本：分布式事务数据库无需引进昂贵的存储设备，国产标准化 x86 服务器，SATA、SSD 本地硬盘存储即可，同时对于网络要求带宽也适度降低；
 - e) 无绑定：当前分布式事务数据库在国内外蓬勃发展，相应的服务提供商数量正在快速增长，所以并未出现严重的服务厂商绑定现象，应用方选择权扩大。

根据上述分析，分布式事务数据库能够解决当前金融业数据库无法解决的问题，故金融业数据库向分布式转型已经成为必由之路。

三、金融关键业务对分布式事务数据库的要求

为满足金融关键业务场景的核心诉求，分布式事务数据库首先必须有准确的产品定位、可持续演进的整体架构、清晰的应用场景。当前业内尚未对分布式事务数据库的基准能力拥有统一认识，大部分企业仍然使用传统事务数据库能力范围对分布式事务数据库进行能力适配和验证，这种方式无法适应新环境下分布式事务数据库的选型、规划、应用。

中国信息通信研究院云计算与大数据研究所与支付清算协会共同成立金融分布式事务数据库工作组，联合金融机构、研究组织、产

业专家共同制定《大数据产品 分布式事务数据库 第一部分：技术要求》，梳理了分布式事务数据库 6 项核心能力要素，为分布式事务数据库的应用与发展提供了理论基础和依据。

（一）基本功能

首先，金融分布式事务数据库应该能够具备事务数据库系统的传统基本功能，从而在最大限度简化当前业务系统和业务人员的适配工作的同时，保障数据库业务的正常运转。同时，金融分布式事务数据库需要能够体现分布式基本能力，为性能扩展和新架构运维等做好技术准备。

基本功能指标包括（1）通用数据类型、操作符、字符集、函数、SQL 语法等的支持能力（2）数据分片操作、分区操作、分布式事务、隔离级别设定的支持能力（3）执行计划、表分区、索引的支持能力。

（二）兼容能力

兼容能力代表了金融分布式事务数据库与现有业务系统和通用数据库硬件、软件工具的衔接能力，这一项衡量了金融分布式事务数据库落地后对现有业务生态造成的影响大小。

兼容能力指标包括（1）支持 ODBC、JDBC 等通用连接方式的连接（2）数据类型隐式转化能力（3）传统数据库中异构数据全量迁移能力（4）能够使用其他数据库进行远端连接操作的外部表能力（5）X86 等主流硬件的支持能力。

（三）管理能力

完整而友好的管理能力能够为金融分布式事务数据库运维工作提供支撑，从而使相关数据库 DBA 管理人员、数据库业务负责人、相关应用支撑团队拥有对数据库的良好掌控能力。

管理能力指标包括（1）友好的安装部署、配置管理、升级能力（2）能够进行节点和数据的实时监控、统计分析，并以多种形式通知告警（3）能够具备无死锁或死锁处理能力（4）能够对系统节点和数据的元信息一致性检测（5）具备在线 DDL 命令、管理命令支持的能力（6）能够对各类日志进行管理、查看、处理（7）能够对数据以定时、增量、全量等多种方式进行备份和恢复工作（8）能够对集群内资源进行定制化隔离分组。

（四）高可用能力

金融业务的稳定性受到了监管机构的严格规定，所以金融分布式事务数据库作为核心业务的基础支撑设施，高可用能力必须得到保障，从而满足金融企业的业务和监管要求。

高可用能力包括（1）电源掉电、硬盘故障、网络闪断、机房失联等硬件事故下的数据库服务能力（2）主副中心数据备份和同步过程中故障下的数据库服务能力（3）CPU 资源过载、I/O 资源过载、内存资源过载、磁盘空间过载等操作系统软件层事故下的数据库服务能力（4）数据文件损坏、数据节点损坏、系统表损坏等数据库服务故障下的数据库服务能力（5）应用连接丢失、外围系统错误引发的

应用故障下的数据库服务能力。

（五）扩展能力

分布式事务数据库必须提供在线 **scale-out** 水平扩展能力。在客户业务扩容过程无明显感知的基础上，满足客户业务增长的需求，解决现有系统在磁盘容量、性能等方面逐渐出现的瓶颈问题。

扩展能力具体指标包括（1）读写业务的计算资源能够在集群中各服务器间进行自动均衡的能力（2）存储数据能够按照业务方定制化需求进行均衡的能力（3）集群的在线扩容能力，从而在不影响业务正常运行情况下解决资源使用瓶颈（4）集群的在线扩容、缩容能力，从而在不影响业务情况下达到资源的合理规划利用。

（六）安全能力

完整的安全能力能够使金融分布式事务数据库稳定承担业务负载，防止敏感数据泄露，杜绝误操作行为，从而防范金融风险的发生。

具体指标包括（1）对数据库中库级别、表级别数据操作的权限验证能力（2）对接入数据库用户身份认证能力，能够实现白名单、黑名单审核（3）对数据库内各类操作的审计回溯工作，从而能够定位分析关键问题（4）具备数据库流量控制能力，从而防止某突发业务流量造成数据库系统故障，影响其他业务的正常运行。

四、分布式事务数据库对金融业务设计的改变

金融业务架构从集中式架构转型为分布式架构具有重大意义。在转型过程中，底层数据库存储也从原先的单节点的商业数据库变成多

节点分布式数据库。相对于分布式系统中的其他组件，分布式数据库作为底层基础架构，具有强重要性和较大改造压力。

本章力图分析分布式数据库改造后将会对现有金融信息系统设计模式带来的变化，从而指导相关人员从业务设计和实际操作等方面完成意识认知和过渡。

（一）业务系统组件化、服务化改造

对业务系统进行分层解耦，确定应用层、服务层及数据层的边界，以适应系统弹性扩展的需求。对承上启下的服务层进行服务化改造，提供“去中心化”的服务调用，并以“弹性”扩展计算能力为主要衡量指标。

（二）数据库模式设计，满足扩展需求

在数据库模式设计的过程中，主要关注如下几个方面：

扩展性需求。随着业务的快速发展，数据量会急剧增大，在模式设计的时候要充分考虑，通常要对数据进行分区，比如按照时间做 Range 分区、按照记录的某个特征值做 Hash 分区。保持每个分区的大小适中，对系统的负载均衡、调度以及扩展性都有帮助。

性能需求。分布式数据库系统的节点数通常从几个到几百个不等，对数据的访问需要路由。在进行数据库模式设计的时候，需要结合数据常用的访问模式，来确定分区键，以便根据 SQL 语句就可以准确地路由，减少二次转发带来的性能损失。同时，尽管大多数的分布式数据库系统都支持分布式事务，但是相对于单机事务，性能有一定的

损耗，所以在模式设计的时候，也需要考虑不同表数据的亲和性，把可能会被一起访问的数据放在同一个节点上，减少分布式事务的比例。此外，由于多维度访问数据的需求，有可能需要创建全局索引，相对于单机数据库系统而言，在分布式数据库中维护全局索引会产生分布式事务，对性能也有较大影响，所以要平衡全局索引带来的查询性能提升和维护成本增大。

全局一致性快照。有的分布式事务数据库系统只实现了单个节点上的一致性快照，而没有实现全局的一致性快照，在一条语句中访问位于不同节点上的数据时，这些数据可能不属于同一个快照点。这种行为的一个后果是无法保证不同连接上操作的因果序，会给业务带来一定的困扰。为避免这种情况，在模式设计的时候告诉数据库系统相关联数据的亲和性，以便数据库系统在调度的时候将这些数据聚合起来。

（三）数据库 SQL 能力评估及业务开发注意事项

目前分布式事务数据库系统并未形成统一标准，产品之间 SQL 能力差异性比较大，而传统金融业务开发多基于 DB2、Oracle 等商业数据库系统，或多或少利用了商业数据库特性，这些特性在分布式事务数据库中有的是功能不支持，有的是性能上有损耗，需要在设计中加以考虑，并进行适当取舍。典型的特性如下：

Sequence（序列）支持。有些业务使用数据库的 sequence 功能来产生记录的唯一键，是否支持该功能以及在分布式系统中获取 sequence 的性能对系统有较大的影响。

强一致性事务及事务隔离级别支持。对于金融核心业务，必须有强一致性的事务支持。分布式关系数据库通常都支持 `read committed`（读已提交）隔离级别，如果有其他隔离级别的需求，需要在系统设计过程中结合数据库能力加以考虑。

复杂查询支持。相对于单节点数据，分布式数据库优化器生成最优计划面临更大挑战。对包含多表联接以及复杂子查询的语句，能否确定合适的联接顺序，将联接和计算更多的下降到每一个节点，减少节点间的数据传输，不仅仅是功能性问题，也是性能问题。

（四）系统容量评估及弹性方案

分布式事务数据库的一个重要特性是扩展性好，可以方便地对系统容量进行弹性伸缩。在系统设计的时候，要对系统容量进行评估，以便确定数据库系统的硬件资源需求。在容量评估过程中，除从系统总的 `TPS/QPS`（每秒事务能力/每秒查询能力）角度去估算以外，还要关注热点账户等关键特征，以便在资源和负载均衡等方面有一个更全面的考虑。在完成对日常场景的容量评估后，还需要对诸如“双十一大促”、“春节红包”、“秒杀”等特定事件触发的流量峰值规划弹性方案，上述场景大多是有计划的，可以提前进行容量规划和根据预估峰值流量进行扩容；当该特定事件结束后，系统要相应地进行缩容以便释放多余资源。扩容和缩容都应该是在线的，不影响正常的，需要根据业务特点（流水型或者状态型），结合分布式数据库的弹性能力，进行弹性方案的设计。

（五）高可用及容灾方案

和传统商业数据库系统依赖硬件设备高可用来达到系统高可用不同，分布式事务数据库运行在普通的硬件设备上，硬件故障和损坏是一个大概率事件。数据库系统的高可用依靠分布式数据库软件来实现。目前主流的高可用方案是多副本强一致方案，同一个数据的多个副本存放在不同的数据库节点上，这些数据库节点有可能存放在同一个机房中，也可能存放在同一个城市的不同机房中，甚至是不同城市的机房中，多个副本之间通过回放日志进行数据同步，事务对数据的修改必须在大多数的副本上生效才算成功。

系统设计过程中，要根据系统对高可用的要求选择相应的方案：

- 单机房多副本，副本存放在不同节点中。该方案可以应对单节点故障，事务提交操作没有跨机房日志同步操作，对事务响应时间影响最小。在单节点故障情况下，数据不丢失，原故障节点对外提供的服务有几十秒的影响；
- 同城三机房，每个机房一个数据副本。该方案可以应对单机房故障，事务提交操作需要在同城机房间进行日志同步，对事务响应时间有较小影响；
- 两地三机房，可以选择 3 个副本或者 5 个副本。该方案在应对单机房故障的基础上，如果存放单机房的的城市发生故障，数据也不会丢失；事务提交操作大概率在同城机房间进行日志同步，对事务响应时间有较小影响；但一旦同城的两个机房中一个机房发生故障，事务提交就需要跨城日志同步，对事务响应时间

有较大影响。

- 三地三机房，多选择 5 个副本。可以应对城市级故障。事务提交操作需要跨城进行日志同步，对事务响应时间有较大影响。

在考虑数据库系统的部署方案时，也要一并考虑业务系统的部署方案，比如对于可用性要求特别高的系统，分布式数据库系统采用三地三机房的方案，业务系统也需要相应部署在三个机房里，避免跨城数据库访问造成的性能损失；同时每一次事务提交需要跨城事务日志同步，对响应时间有一定的影响。

（六）系统运维方案

相对于集中式系统，分布式系统的运维复杂度更高，需要尽可能自动化和在线操作。两个常见的运维过程：系统升级和系统备份，都需要在线实施，不能影响业务的正常运行。为达到这个目的，业务系统和分布式数据库系统都需要支持在线升级。对集群数据库节点分批升级，在升级之前，利用多副本特点，将一部分节点的流量切换到位于其他节点的副本上，在升级完成后再切换回来；业务系统的在线升级过程也类似。

此外，对于分布式系统，问题跟踪也更加复杂。在业务系统设计过程中，需要单独考虑问题跟踪能力，比如为每一笔业务，生成一个单独的交易号，通过交易号串联整个业务流程，并持久化在数据库中，方便异常情况下的问题分析和解决。

五、金融分布式事务数据库迁移策略

（一）指导原则

结合金融行业业务特性和安全保障的要求，金融分布式事务数据库的迁移工作，建议按照以下的指导原则开展工作：

1、可持续发展原则

新一代分布式事务数据库的发展，在全球还处于实践起步阶段。金融行业在考虑和应用分布式事务数据库的过程中，在迁移阶段，要充分评估和考量分布式事务数据库架构的可持续性发展，迁移过程中要考虑到应用适配和调整策略符合中长期的发展要求。

典型的考量有：

- 数据模型和核心功能架构先进性是否具备可持续发展的能力；
- 数据库的工程化实现及运行架构是否具备可持续发展的能力；
- 分布式事务数据库的跨行业，跨区域开拓是否具备可持续发展能力；
- 分布式数据库的许可证模式，推广模式的可持续发展能力。

2、透明开放原则

在金融行业应用分布式事务数据库的过程中，迁移过程中，需要尽可能按照透明开放的原则来进行各种任务的规划和要求。尤其是对于分布式事务数据库的核心架构、应用接口、运维管理架构等应从规划、测试、维护、定制等多个方面，保证透明和开放度。

典型的考量有：

-
- 分布式事务数据库核心存储计算架构透明开放度；
 - 应用接口和 API （应用平台接口）规范的透明开放度；
 - 底层部署和运行架构的透明开放度；
 - 扩缩容及变更运维平台/工具的透明开放度；
 - 测试工具和用力的透明开放度；
 - 功能的可定制接口的透明开放度。

3、代价可控原则

在向分布式事务数据库迁移的过程中，一定会涉及到对业务应用的调整和优化，随着业务系统自身的复杂性差异，这种调整往往会伴随有时间、资金、风险和人员方面的成本存在。在选择迁移路径和方法的过程中，要尽可能选择总体迁移代价最优的迁移路径。

典型的考量有：

- 应用适配代价；
- 接口适配代价；
- 基础设施资源支撑代价；
- 运维管理基础设施调整代价；
- 日常运维管理代价；
- 变更和应急代价；
- 人员能力掌握和维持代价。

（二）实施策略

1、知识传递先行

不同于传统的事务数据库产品落地应用。分布式事务数据库在业界处于实践起步阶段。金融行业应用分布式事务数据库，需要在实施阶段，首先执行深入的知识传递和过渡。

核心的内容包括但不限于：

- 核心架构原理；
- 应用适配；
- 业务验证框架；
- 评价体系；
- 资源供给规划；
- 容量和性能管理；
- 实施和运行管理；
- 故障排查和应急；
- 日常运维管理手段。

2、测试评估价值体系的建立

在进行充分的知识传递之后，需要在实施正式启动前，建立有效的测试评估评价体系，对于分布式事务数据库的评估评价体系的建立，可以考虑以下几个维度进行充分验证和评估：

- 产品（或软件项目）生态评估；
- 产品（或软件项目）成熟度评估；
- 产品（或软件项目）关键功能和企业级特性评估；
- 产品（或软件项目）性能和可用性评估；

-
- 产品（或软件项目）服务与支持能力评估。

3、上线方案和上线实施

因为金融业务的强安全要求，所以金融企业在进行分布式事务数据库改造时，不可以贸然将所有业务进行全量替换。金融机构需要根据业务重要性、业务安全要求以及业务流量压力等维度，对自身现有存量及增量业务进行分级分类，按照一定顺序逐步验证分布式事务数据库能力，最终完成全量替换。常用的上线方案分为几种方案。

（1）先增量，后存量

对于金融机构内互联网跨界融合等增量业务，因为安全性要求较低，不会发生金融安全风险，可以先行替换试验。当新上线分布式事务数据库能够满足业务需求，并且相关业务人员已经适应相关操作方法和模式后，通过切割流量的方法逐步对存量的核心业务进行替换。在存量业务替换的同时，需要同时维护原系统的备份，作为灾备方案。当新系统稳定运行得到验证后，逐步下线原系统。

（2）按业务性能瓶颈顺序

分布式事务数据库会带来强大的吞吐/处理能力。金融机构可对当前各类业务性能瓶颈进行分析评估，目前哪项指标急需分布式事务数据库方案进行解决，按照一定的优先级顺序，逐轮进行流量分割处理，在缓解现有业务压力的同时，验证新系统的各项能力，层层递进。

（3）按业务不同的技术实现类型规划

金融机构内的业务纷繁复杂，各类业务对信息系统的技术要求并不相同。例如转账业务主要要求大并发的事务能力，批量计提结息业

务要求大规模数据的并行处理能力，日志分析要求实时大规模数据的分析能力。所以金融机构可以按照不同业务对技术的需求能力，安排改造顺序，从而逐步验证分布式事务数据库能力。

4、实施的阶段性规划

下面为分布式事务数据库的实施阶段参考：

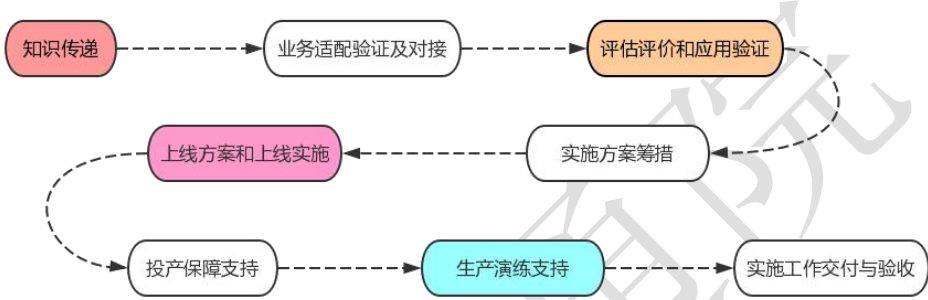


图 7 分布式事务数据库实施过程

（三）典型问题和误区

1、转型过程中的应用过度适配和改造

在实施分布式事务数据库的过程中，重要的一项工作为应用适配和调整优化。因分布式事务数据库的技术种类有差异，在某些技术架构中，为了适配分布式事务数据库的技术特性，比如分布式事务控制，分布式数据库表关联（即 JOIN），一致性控制等，会要求应用侧进行局部性调整。

通常从业务角度，无论是业务应用侧和分布式事务数据库侧，被视为作为一个整体进行统一的调整。为了达到更好的业务逻辑正确性与性能，数据规模（存量 / 增量）和数据库伸缩性之间的平衡，会采用对业务应用采用深入定制的方式。这样的策略和方式，最后能够达

到整体业务目标和技术目标，但这是以通过紧耦合业务应用和分布式事务数据库特性为代价的。

从技术角度来说，这样的策略路径会带来可持续性发展的风险。过度的应用侧改造调整，降低了灵活性，耦合度的增加对于整体扩展能力和运行维护成本控制都有负面影响。

2、使用旧管理策略约束分布式事务数据库

分布式事务数据库从底层的内部可用性机制到全局的对外业务连续性管理（BCM）和传统单机事务数据库差异非常大。在分布式事务数据库的应用场景中，单个节点的失效，网络的抖动/闪断/隔离，存储的失效都是分布式事务数据库架构设计范畴内的预设的故障场景。单个点失效和多点失效，在分布式事务数据库的场景中，采取的策略简单的重试，采取剔除和隔离故障（分布式系统中所谓的“牛群模式”）。而传统的单机事务数据库，在生产中的可用性管理，在传统高可用架构中，单个节点失效，需要尽快的进行复杂的修复工作和节点替换工作（分布式系统中所谓的“宠物模式”）。

从对分布式系统的可用性度量和管理上，应该以整个分布式事务数据库集群的对外业务可用性作为核心指标，内部单个节点的损坏，应采取更加符合分布式事务数据库机制的运维管理流程进行合理管理。

3、不重视基础网络环境建设

分布式事务数据库作为提供高吞吐，大并发，大数据量的事务数据库系统，本质是一套分布式计算与存储系统，分布式计算与存储系

统对于网络吞吐和网络延迟有非常高的依赖性和要求。无论是同城同数据中心的分布式事务数据库集群内部的网络通信，或者是跨数据中心的多个数据中心容灾/灾备或多活部署，都需要网络资源供给侧能够提供高吞吐，低延迟的网络。

六、结论与建议

金融业信息系统分布式转型是必然趋势，其中分布式事务数据库是架构转型的难点，一是技术门槛高，二是缺乏相应国际经验与成熟商用解决方案，三是从集中式向分布式转型的路径不明确，标准不统一，四是应用开发、安全保障、运维服务等配套体系还不成熟，生态系统需要进一步完善。为此，还需要金融用户单位、技术提供商、研究机构和学术界共同努力，在一下几个方面开展工作：

（一）建立技术标准体系

未来金融分布式事务数据库需要通过“新型产品技术要求”、“原有体系迁移规范”、“运维服务保障准则”三大类标准，设立转型前选择、转型中实施、转型后保障三层次门槛，以标准带产业，统一认知水准，实现业务共识，从而实现风险分摊。

（二）开展技术试验与测试

金融 IT 新架构的能力测试需要达到自动化、智能化、全面化，而当前业内测试方案仍以国际标准组织 TPC 研制的 TPC-C 物流仓储模型的性能测试为主，不符合金融特征，且功能较为单一。为有效推进金融架构转型，需要抽象金融业务特征，依托金融场景构造自动化测

试工具及体系，并依托测试工具组织数据库产品能力测试，做到使用金融特征衡量金融产品。

（三）在金融机构开展试点

第三方中立机构需要协助金融机构，以实际业务为背景，开展新型产品真实环境验证试点工作，为分布式事务数据库成功在金融机构应用提供新模式，同时总结梳理相关经验、共享成功案例，为后续全面推广应用打好基础。

（四）完善支撑配套体系

为形成金融单位和 IT 企业联动的新服务格局，发挥各自的技术与业务优势，解决金架构技术与金融应用适配问题，相应组织需要建立应用方与服务方融合新生态，打造互通平台，提高服务效率，填补金融业与 IT 业的连接洼地，共同完善应用开发、安全保障、运维服务等金融数据库业务配套生态体系，保障金融分布式转型平稳国度。

中国信息通信研究院云计算与大数据研究所

地 址：北京市海淀区花园北路 52 号

邮政编码：100191

联系电话：010-62304839

传 真：010-62304980

网 址：www.caict.ac.cn



中国信息通信研究院



大数据发展促进委员会

中国支付清算协会金融科技专业委员会

地 址：北京市西城区金融大街 20 号交通银行大厦 B 座 7 层

邮 编：100033

联系电话：010-88665181

传 真：010-88665185

网 址：www.pcac.org.cn