



F R O S T & S U L L I V A N

*60 Years of Growth, Innovation and Leadership*

中国关系型数据库产业发展和行业应用（2023）

——国产数据库助推产业创新，塑造数字化未来新格局

A Frost & Sullivan  
White Paper

FROST & SULLIVAN

沙利文

科壹資訊  
CCW RESEARCH

[www.frost.com](http://www.frost.com)

◆ 数据库产业概况	-----	2
• 数据库的概念及主要类型		
• 数据库技术发展趋势		
◆ 中国关系型数据库市场现状	-----	9
• 国产数据库品牌逐渐崛起及采取多种技术路线		
• 国内开源数据库根社区和根生态初步形成		
◆ 中国关系型数据库市场重点行业应用场景与特点	-----	15
• 金融及电信行业数据库应用场景与特点		
• 政府及制造行业数据库应用场景与特点		
◆ 国产数据库在重点行业应用态势分析	-----	20
• 数字化创新转型由一般系统转向核心		
• 集中式数据库在OLTP核心系统的作用不容忽视		
• 分布式数据库成为金融、电信等行业的新选择		
• 应用更关注数据安全		
• 兼容性、迁移适配方法论逐步成熟		
◆ 国产数据库行业应用深化面临的问题	-----	27
• 关键核心技术能力不足；行业应用迁移替代难度大		
• 产业生态力量布局分散；产业链风险问题急需解决		
• 数据库专业研发人才短缺		
◆ 对策建议	-----	32
• 强化原创性技术创新；深化行业应用牵引机制		
• 优化产业发展环境；保障产业链韧性安全		
• 加大研发人才培养力度		

# 章节一 数据库产业概况

## 数据库的概念与分类

### 关键发现

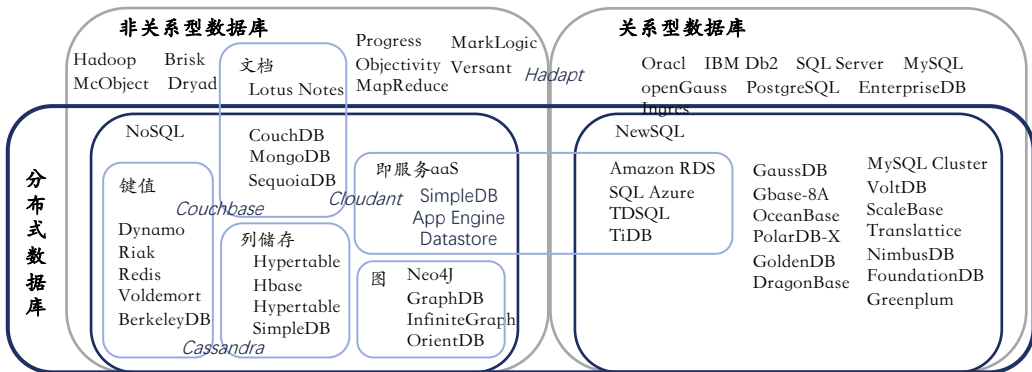
- 数据库 (Database) 是按照数据结构来组织、存储、管理, 并且可共享的数据集合软件, 是IT系统存储与计算的基础, 与芯片、操作系统共同组成IT系统的核心, 广泛服务于各类行业应用, 是软件产业生态体系构建的重要枢纽
- 按照不同的维度, 数据库有不同划分 (1) 按数据结构模型可分为关系型数据库和非关系型数据库; (2) 按网络架构可分为单机型数据库、存算分离性、集中式数据库和分布式数据库; (3) 云数据库, 是指被优化或部署到一个虚拟计算环境中的数据库

### 按数据结构模型可分为关系型数据库和非关系型数据库

**关系型数据库 (SQL)**, 采用了关系模型来组织数据, 以行和列的形式存储数据。关系型数据库的行和列的集合被称为表, 表的集合则组成数据库。其诞生40多年, 已形成较为成熟的产品体系, 代表产品如 Oracle、DB2、SQL Server、MySQL、PostgreSQL、openGauss等, 其优点是事务的一致性, 在金融等行业要求数据完整性、一致性较高的领域中广泛应用。

**非关系型数据库 (NoSQL)**, 泛指除关系型以外的数据库, 是对关系型数据库的一种补充。非关系型数据库的产生是为了解决大规模数据集合多重数据种类数据库的分类

带来的诸多挑战, 尤其是大数据应用难题, 其包括: 键值存储数据库 (Key-value), 典型产品有 Memcached、Redis 和 Ehcache; 列存储数据库, 典型产品有 Cassandra 和 HBase; 面向文档数据库, 典型产品有 MongoDB 和 CouchDB; 图数据库, 典型产品如 Neo4J、InforGrid; 时序数据库, 典型产品如 InfluxDB。非关系型数据库具有扩展性强、高并发读写、灵活的数据模型等特点, 广泛应用于数据量大的业务系统。但是也存在明显的短板, 如, 种类多、需要兼顾各类非关系型数据难度较大、无法对传统的数据类型 (关系型) 的应用进行升级等。



来源: 沙利文整理

## 数据库按结构模型分类

### 关系型数据库与非关系型数据库特点

名称	定义	典型应用场景	
关系型数据库	支持关系数据模型，由库、表、数据类型，记录组成，按元素关系组织的数据库。	事务型为主	
非关系型数据库	键值数据库	只能存储键和值，满足高性能需求。	嵌入式系统或高性能进程需求
	列存数据库	能容纳大量动态列。由于列名和记录键不是固定的，并且由于记录可能有数十亿列，因此宽列存储可以看作是二维键值存储。查找速度快；支持分布横向扩展；数据压缩率高	分析型业务场景
	文档数据库	面向文档的数据库系统，其特征在于它们的无模式数据组织	文件管理等
	图数据库	面向图的数据，将图结构中的数据表示为节点和边缘，表达节点之间关系。	社交、零售、金融风险、公共安全
	时序数据库	时序数据库是一个针对处理时间序列数据进行了优化：每个条目都与一个时间戳相关联的数据库管理系统。	物联网
实时数据库	结合实时处理技术，直接实时采集系统运行过程中的数据的数据库系统。	流控制、工业控制	

### 关系型数据库与非关系型数据库优缺点

类型	关系数据库	非关系数据库
优点	容易理解、使用方便、易于维护	无需经过SQL层的解析，读写性能很高、数据容易扩展、存储数据的格式多样化
缺点	处理高并发读写性能较差、表结构较为固定、不便于更改扩展	处理多种非关系型数据库具有挑战性、难以对传统关系型应用进行升级
代表产品	Oracle、MySQL、SQL Sever	MongoDB、CouchDB

关系型数据库优点在于数据结构清晰、提供ACID事务特性和丰富的管理工具，但在高并发读写和数据扩展方面表现较差。非关系型数据库具有高读写性能、灵活的数据结构和扩展性强的优点，但处理多种数据库和升级传统关系型应用方面存在挑战。

来源：沙利文整理

## 数据库按网络架构分类

单机型数据库、存算分离性、集中式数据库和分布式数据库特点

分类	应用介质	数据存储	应用场景
单机型	单机PC服务器，小型机，或大型机	本地硬盘	小规模企业
存算分离型	多个PC服务器，小型机，大型机	高速网络技术，整合存储外挂，磁盘阵列，分布式存储	中等规模数据应用
集中式数据库	多个PC服务器，小型机，大型机	共享磁盘，磁盘阵列，全局资源管理	数据是共享，主要用在企业级核心业务，满足高可用和性能要求
分布式数据库	多台服务器	无共享架构 (Shared Nothing)，不对其中的磁盘和内存进行实时同步	数据是分散存储的，面向互联网、电商等业务场景，关注点在于横向扩展能力、高可用和性能要求

**单机数据库**即是传统意义上的数据库，又称为完全共享型 (Shared Everything) 数据库。数据库管理软件部署于单台服务器上 (PC 服务器，小型机，大型机等)，使用本地磁盘存放数据。

**存算分离型数据库**是指数据库管理软件依旧部署到单台服务器上，但是存储部分利用高速网络技术，将存储外挂，利用磁盘阵列，分布式存储等设备将单台服务器的有限存储进行拓展。

**集中式数据库**是指采用集中式架构，将数据存储在大中型主机或小型机上进行集中管理，其操作系统，中间件，数据库等“基础软件”多为闭源商用系统，典型的

集中式架构是IOE (IBM, Oracle, EMC) 提供的计算设备、数据库技术和存储设备共同组成的系统，同时也是目前银行、电信等行业的主流应用模式。

**分布式数据库**是由若干个节点集合而成，它们通过网络联接在一起，每个节点都是一个独立的数据库系统，它们都拥有各自的数据库、中央处理机、存储，以及各自的局部数据库管理系统。分布式数据库具有高可用、高可靠、可扩展的特性，在Google、Amazon、Facebook、阿里巴巴、腾讯等互联网公司广泛应用。随着数据量大幅增长以及高并发环境下对数据处理能力的要求越来越高，分布式数据库也逐渐被金融行业关注和应用。

来源：沙利文整理

## 分布式数据库与传统数据库的区别

### 集中式数据库 vs 分布式数据库

	集中式数据库	分布式数据库
定义	仅在一个位置上 储存、定位和维护的数据库	分布于计算机网络 且逻辑上统一的数据库
可扩展性	横向扩展受限，支持纵向扩展	支持横向扩展
扩容弹性	大变更需停机	快速迭代、小时级投产、分钟级在线扩容
成本	扩展需要高配置硬件，集中式存储系统成本较高	利用廉价PC Server组建集群，成本相对较低，但Server数量需求较多
兼容性	无法进行非结构化大数据处理，硬件兼容能力弱	支持非结构化大数据处理，硬件兼容能力强
高并发能力	可以支持单表千万级数据量的存储，但是难以支撑密集的并发读写，存在容量与性能瓶颈	关联多个节点，减少单个节点数据量；实现并行计算，支持PB级数据量访问，以及百万级高并发
事务性	遵循ACID	遵循CAP、BASE，少数提供ACID能力
自治性	集中式控制、厂商标准、封闭	局部DBMS自治性、产业标准、开放
可用性	系统监控与发布部署有架构优势	容灾机制和故障恢复有架构优势

集中式与分布式数据库各有优缺点。近年来，分布式数据库架构发展比较迅速，首选，在经济性方面，分布式架构通常能够更有效地利用计算资源，通过水平扩展来降低成本，而集中式数据库往往需要昂贵的硬件和基础设施来支持其运行；其次，在安全自主性方面，分布式数据库往往设计有更严格的权限控制和数据隔离机制，能够更好地保护用户数据的安全和隐私；此外，分布式架构在灵活性和可伸缩性方面也具有显著优势，能够轻松应对不断变化的业务需求和数

据量增长。虽然集中式架构在可维护性和一致性方面可能更具优势，但分布式系统通过采用先进的分布式中间件和运维平台，也可以达到甚至超过集中式数据库的性能和可靠性水平。分布式中间件可以帮助管理和协调分布式系统中的各个节点，提供统一的服务接口和数据处理能力，从而简化开发和运维的复杂性。同时，运维平台可以对分布式系统进行全面的监控和管理，确保系统的稳定性和高效运行。

来源：沙利文整理

## 云数据库

云数据库是指被优化或部署到一个虚拟计算环境中的数据库，可以实现按需付费、按需扩展、高可用性以及存储整合等优势，其本质是将各类数据库技术与云平台技术结合，通过虚拟化、容器化或者裸金属等方式将数据库进行云化管理，以服务的形式对用户进行交付，而非传统的License交付。云数据库不仅提供WEB界面进行配置、操作数据库实例，还提供可靠的数据备份和恢复、完备的安全管理、完善的监控、轻松扩展等功能支持。相对于用户自建数据库，云数据库具有更经济、更专业、更高效、更可靠、简单易用等特点，使用户能更专注于核心业务。

数据库不同分类维度之间，可以互有交叉。一个数据库产品可同时是“关系型-决策型-分布式”，如Greenplum数据库、

DMMP数据库等。同一个分类维度之间也不是“非黑即白”，存在“跨界”产品。如混合事务-分析处理数据库（HTAP）同时具备事务型数据库和分析型数据库的能力，多类型数据库是可同时管理关系型、键值型、文档型等模型的数据库。

由于中国数据库产品呈现以关系型为主，非关系型数据库为辅的局面，关系型数据库在中国数据库总体市场中的占比超过60%，成为中国数据库创新发展的“主战场”，本报告的分析集中于关系型数据库领域。

### 云数据库特点



- 云数据库可弹性扩展，快速提供存储和计算资源，适应业务增长，无需大量人力和硬件投入
- 云数据库具备高可用和可靠性，通过分布式架构和冗余备份，确保故障时的数据安全和系统可用性，保障业务连续性
- 云数据库灵活可定制，提供丰富配置选项和参数调整，适应业务需求。用户可按需选择存储引擎、调整资源比例

# 数据库技术发展趋势

## 发展趋势

### 1 数据库架构走向存算分离、资源池化

存算分离架构将计算和存储资源充分解耦并实现资源池化，计算节点转为无状态节点，支持按需分配；外置的共享存储可以使用中心化的存储集群，提升存储的可靠性和扩展性。存算分离将存储资源和计算资源拆分为独立的模块进行建设，在资源利用率、存储资源高效共享、多场景灵活部署等方面具有显著优

势。此外，随着高性能盘、NVMe新协议、RDMA/NoF新型网络等新技术的大量涌现，为存算分离架构提供了更好的技术支撑，使得存算分离架构成为技术演进趋势上的选择。目前，AWS Aurora、阿里云PolarDB、openGauss等都不约而同采用了存算分离架构去提升数据库的整体能力。

### 2 混合事务与分析型数据库需求不断增长

在需要同时支持OLTP和OLAP场景，基于创新的计算存储框架，在同一份数据上在保证事务的同时可支持高效实时分析，省去费时的ETL过程，构建出混合事务分析处理的HTAP数据库。HTAP数据库采用事务处理、分析处理、数据同步、查询优化、资源调度等多种技术，实现

在事务型行存基础上，定期将增量数据合并到列存储中，用以满足分析型负载，并结合分布式调度技术实现并行化，进一步加速处理。目前各大厂商都在布局HTAP，如阿里云PolarDB、腾讯云TDSQL、openGauss等。

### 3 多模态数据库将成为流行

数据库的下一个模式会向多模方向发展，同时支持关系、KV、文档、图、时序等模式。多模数据库支持灵活的数据存储类型，将各种类型的数据进行集中存储、查询和处理，可以同时满足应用程序对于结构化、半结构化和非结构化数据的

统一管理需求。目前已经有多种原生多模数据库，如ArangoDB、OrientDB等，但也有传统的关系型数据库，如PostgreSQL、openGauss等，通过插件扩展支持时序、图、GIS等多种模态。

### 4 分布式数据库迎来快速发展机遇

分布式数据库是采用计算机网络将物理上分散的多个数据库单元连接起来组成的一个逻辑上统一的数据库，具有高并发/高吞吐读写、海量存储、弹性伸缩、高可靠性等特点。相较于传统数据库，分布式数据库具有更高的性能和可扩展

性，能够满足大规模数据处理和实时数据更新的需求。目前被广泛应用于互联网、金融、电商等领域。目前各大厂商数据库都具备分布式能力，如华为云GaussDB、奥星贝斯OceanBase、腾讯云TDSQL等。

来源：沙利文整理



## 数据库技术发展趋势

### 5 AI赋能数据库成为未来发展的重要方向

可以通过AI技术实现数据库的自优化、自监控、自调优、自诊断；另一方面可以实现库内AI训练，降低AI使用门槛。从赋能对象来看，人工智能与数据库的结合既可以体现在数据库系统自身的智能化，包括但不限于数据分布技术智能化、库内进行训练和推理操作、数据库自动诊断、容量预判等；也可以体现在数据库周边工具的智能化，能够在提升

管理效率、降低错误引入率、减少安全隐患的同时也大大降低了运营成本。目前学术界和工业界共识的研究重点是将机器学习与数据管理在功能上融合统一，来实现更高的查询和存储效率，自动化处理各种任务。目前业界openGauss在这一块处于领先地位，已经具有较完备的AI for DB和DB for AI能力。

### 6 软硬协同是数据库效能提升的关键

新兴硬件可以从计算、存储和传输三个层面赋能数据库。计算层面，借助众核CPU、GPU、FPGA、AI芯片等，可以实现包括但不限于多核并行优化、事务并发控制、查询加速、存储层计算卸载、数据压缩加速、工作负载迁移等能力；存储层面，随着NVMe的出现和发展，内存和外存的界限变得模糊，针对传统块存储设计的索引在NVMe中面临新的性能

挑战；传输层面，RDMA带来网络传输高性能表现和CPU卸载能力。为充分榨取软硬协同的系统性能，未来可能对数据库系统的架构设计带来颠覆性变化。

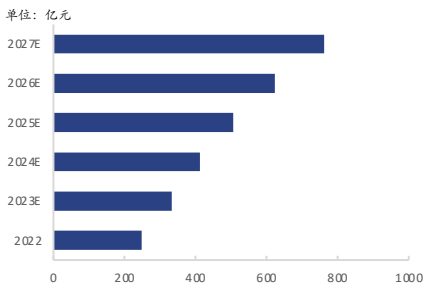
目前各数据库都在利用硬件进行加速，如：PostgreSQL利用GPU加速复杂查询、openGauss利用NVMe做二级缓存、PolarDB利用RDMA加速网络通讯等。

## 中国关系型数据库总体市场规模及市场行业份额

### 关键发现

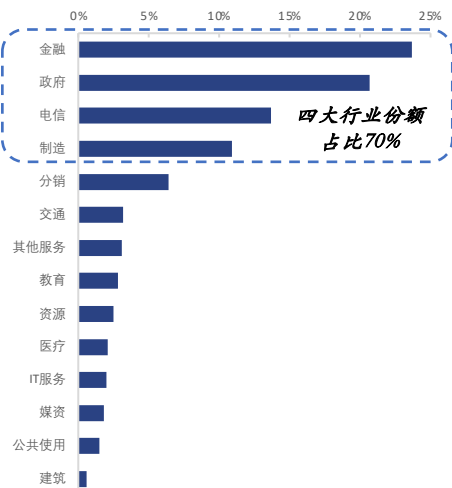
- “十四五”规划提出“加快数字化发展，建设数字中国”的总体规划，作为“数字中国”战略实施的重要一环，近年来中国关系型数据库产业迅猛发展，市场规模逐年高速增长，产业生态不断壮大，行业应用持续深化。根据弗若斯特沙利文预测，到2027年中国关系数据库市场将保持23%的复合增长率，预计到2027年将达762亿元
- 从行业领域市场来看，线下传统部署数据库市场份额主要集中在金融、政府、运营商和制造业四大行业，市场份额合计占比高达市场总份额的70%

### 中国关系型数据库总体市场规模



中国关系型数据库产业近年来呈现出迅猛的发展势头。市场规模逐年高速增长，产业生态不断壮大，行业应用持续深化。据弗若斯特沙利文预测，中国关系数据库市场将在2027年前保持23%的复合增长率，并有望达到762亿元的市场规模。这一趋势反映了中国在数字化转型中的坚定决心和实力，同时也为全球数据库市场带来了新的活力和机遇。

### 中国数据库市场份额（分行业统计）



从行业领域市场来看，线下传统部署数据库市场的份额分布呈现出一定的集中趋势。具体而言，金融、政府、运营商和制造业四大行业是这一市场的主要玩家，它们合计占据了市场总份额的70%。

金融行业由于其庞大的数据处理需求和高度的安全性要求，一直是数据库技术的重要应用领域。政府和运营商则因为其庞大的组织结构和复杂的数据管理需求，同样对数据库技术有着巨大的依赖。制造业在数字化转型的推动下，也开始大量采用数据库技术来提升生产效率和管理水平。

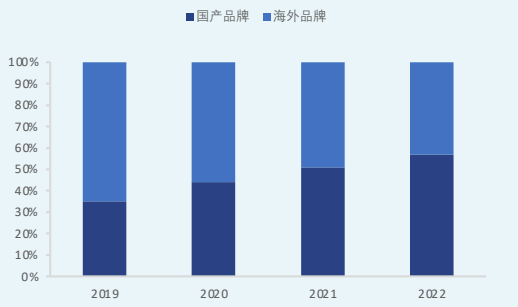
这四大行业的市场份额占比之高，不仅体现了数据库技术在这些领域的广泛应用，也预示着未来这些行业对数据库技术的需求将持续增长。

来源：沙利文整理

## 国产数据库品牌逐渐崛起

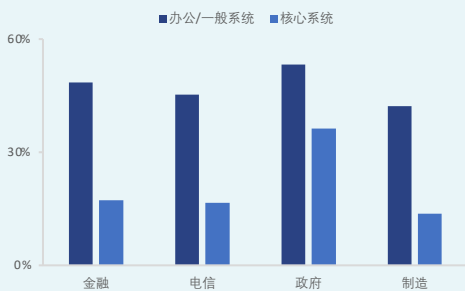
中国关系型数据库市场中本土品牌已经崛起，前五名国产厂商份额总计从2018年的27.1%上升到2022年的55.4%；反之，前五名国际厂商份额总计从2018年的57.3%下降到2022年的27.3%。这其中既有中国本土公有云数据库服务快速发展的因素，也有本土传统数据库厂商近年来持续发力、业务高速增长的原因。本土数据库产品已经越来越广泛地被使用到关键行业和企业核心系统中替换其它存量数据库，在分布式、云原生、HTAP等增量数据库市场，本土数据库品牌的优势也愈发明显。中国关系型数据库市场格局正在重塑，随着疫情管控的开放，数据库厂商的业务开展步入正轨，企业数据库建设项目广泛开展，对于树立市场新格局来讲2023年将是非常重要的。国产数据库产品经过20多年的发展，产品技术逐步走向成熟，国产数据库产业呈现出百花齐放、百家争鸣的局面。基础研究方面，国内数据库厂商加速追赶国际龙头，不断诞生创新理论，技术突破方面，实现了多个关键技术点上的重要突破，产业基本具备了面向关键技术研

中国区数据库市场行业份额（分行业统计）



发能力，正在形成正面竞争力；整个数据库产业呈现出百花齐放、百家争鸣的局面：整体以关系型数据库为主，非关系型数据库为辅，发展路线多样，分布式、云化等数据库产品热度持续高涨；同时，国内数据库采购量大幅提高，市场规模稳定增长，已经具备在各行业核心场景“商用”的基础。2022年国产数据库市场份额已经由2019年的35%，提升到了57%。

国产数据库重点行业应用国产品牌使用率  
(按核心/非核心场景统计)



如今，国产数据库已过了“会用”的阶段，正在走向成熟，并不断走向“能用、好用”

的阶段。对于数据库这样的“国产产品”，国人需要更加包容、更加开放、更加自信地使用，从而让其使用更加高效、稳定，加速成长。

随着中国各行业数字化经济的转型和深入发展，数据库市场需求持续增长，但也面临着数据安全、供应链安全、自主创新需求日益增强的挑战。国家对网络安全、信息安全、信息技术应用自主创新等政策的重视和推动进一步加强。在这样的背景下，中国数据库产业自主品牌发展趋势日益明显，国产数据库在中国金融、电信、政府、制造等行业的应用趋势日益明显，国产品牌份额也逐年提高。

来源：沙利文整理

## 国产数据库采取多种技术路线

中国数据库市场中，基于PostgreSQL（PG）和MySQL两大开源项目的国产数据库产品接近60%，产业整体发展以“跟随”模式为主。近年来，中国数据库产业发展迅速，涌现出一大批数据库产品。追溯国内主要数据库产品的技术来源，大致

可以分为**独立自研型**、**基于国内自主技术开源型**、**基于海外开源代码的二次开发型等技术路线**。国产数据库沿着三条技术路线不断深耕，衍生出不同类型的数据库产品。

### 独立自研路线

该类数据库产品发展之初采取纯自研开发方式，不采用开源代码，企业具备内核级代码能力，可保证技术独立演进，不受国际“断供”风险威胁。主要特点包括：一是产品自研度高，能够独立解决底层代码等核心技术问题；二是兼容传统优势数据库，很强兼容性，具备较强的替换能力

例如，OceanBase经国家工业信息安全发展研究中心源代码评估，文件级代码自有率为96.4%。除少量经典算法库调用外，代码均自研

### 风险分析

独立自研型产品因技术自主度强，不易受外部环境因素影响，缺点是产品技术积累不足、产品成熟度有待验证、迭代创新空间受限

### 典型产品

GaussDB  
达梦DM  
奥星贝斯OceanBase  
平凯星辰TiDB

由于中国数据库市场长期被国外产品占据，独立自研型数据库在市场竞争及替换迁移过程中，仍需投入精力兼容Oracle和MySQL，以此跟进国外产品最新特性等来抢占存量市场，影响其本身技术创新，制约产品迭代升级效率

### 基于国内自主技术开源、自主“根社区”和开源生态

这一技术路线的数据库产品基于国内自主技术研发与开源、构建自主“根技术”、“根社区”和开源生态，逐步形成了具有影响力的开源数据库技术路线和产业生态，能实现独立演进发展，不受“断供”风险威胁

比如，GaussDB和openGauss基于统一的数据库内核，发展分布式与集中式数据库架构路线，建立起自主技术与生态圈。国内基于openGauss已衍生出多款数据库产品，如海量VastBase、云和恩墨MogDB等

### 风险分析

该类数据库产品具有很强的技术自主能力，可持续自我演进，不易受到“断供”影响，是中国数据库产业发展和升级的重要力量。但自主演进型数据库产品需强大的技术投入和不断完善发展的生态，以保持其可持续发展

### 典型产品

华为  
openGauss

整体看国内数据库开源生态的发展还处于初级阶段，但部分国内自主开源产品与生态逐步走向成熟，并形成社区开源与DBV伙伴商业版本发行的生态与产业双轮驱动发展模式，后续依托自主生态与市场良性运转、产业共建与政策引导，逐步形成国内自主开源产业发展的主导与中坚力量

## 国产数据库采取多种技术路线

### 基于海外开源代码的二次开发路线（上游兼容型）

上游兼容型数据库产品主要分支于海外开源数据库或购买商业授权，通过二次开发成为开源或闭源的产品，但其技术发展仍依赖于或受限于上游产品，保持与上游开源社区代码的兼容性，这类产品大部分依然受上游开源许可协议（如GPL/FOSS等）所制约，无法形成自主创新、可控的开源知识产权或商业版权控制能力

这类产品拥有研发上市周期短、借助开源社区的资源技术支撑、与上游产品高度兼容等特点

除传统关系型数据库产品大量采用MySQL和PostgreSQL发展外，由于很多非关系型数据库概念来源于国外产品，国内在非关系型数据库方面依靠国外开源发展现象更为普遍。

### 风险分析

这类依托海外开源代码的数据库产品存在如下潜在风险：一是**知识产权风险**，二是**容易产生技术依赖**，三是**自主创新受限**

**知识产权风险**：若未遵循原产品的许可协议，可能面临法律纠纷和版权问题。如在商业用途中修改、使用了GPL协议的产品源代码，必须遵循开源规定；**容易产生技术依赖**：这类产品若在很大程度上依赖上游技术研发，未能完全掌握核心技术，只能被动兼容，若上游产品出现“断供”或安全漏洞，可能直接影响其稳定性和持续发展能力；**自主创新受限**：套壳产品的创新投入力度不足，将受上游产品限制，可能无法充分发挥其创新能力。

### 典型产品

MySQL系衍生产品及RDS服务

### 独立自研路线

技术独立演进

抗“断供”风险

兼容性强

技术积累不足

成熟度待验证

创新空间受限

特点分析

风险分析

典型产品

奥星贝斯OceanBase

### 基于自主技术开源、“根社区”和开源生态

独立技术根基

社区影响力

生态圈建设

技术投入需求大

生态待完善

市场运转挑战

华为openGauss

### 基于海外开源代码的二次开发路线（上游兼容型）

上游兼容型数据库

普遍依赖国外开源

自主创新受限

知识产权风险

容易产生技术依赖

自主创新受限

MySQL系衍生产品及RDS服务

## 国内开源数据库根社区和根生态初步形成

近年来开源数据库在中国获得广泛普及。2000年以后全球开源软件产业在全球得到了快速发展和应用普及，MySQL、PostgreSQL等开源数据库在全球数据库市场渐露头角到逐渐发展，开源数据库在中国数据库市场也逐渐得到行业应用。最先由开源爱好者、小型客户、教育、科研等客户测试、研究或一般业务系统应用，扩展到中小型企业IT系统、办公应用、电子政务等系统。2010年前后，开源数据库软件以其成本低、源代码开放、产品迭代和定制化开发快等特点，受到了互联网行业客户青睐。中国互联网公司强大的研发实力、广泛的应用场景、快速应用开发和业务推广的能力，推动了开源数据库产品和应用场景的成熟，完善了开源软件生态。互联网行业的示范好技术溢出效应，带动开源数据

库在政府、企业、金融、电信等行业应用探索和应用推广。当前开源数据库软件已经成为中国数据库产业创新发展的重要模式之一。

国内开源数据库早期主要是引入、追随海外MySQL、PostgreSQL等知名的开源数据库软件，基于海外社区的开源代码开发行业自用数据库版本，或发布商业发行版本。随着开源数据库技术和应用逐步成熟，中国开源数据库产业也由跟随、模仿，逐步走向引领创新和生态共建的模式，例如腾讯、阿里巴巴、百度、华为云等互联网企业先后基于MySQL、PostgreSQL等开源代码开发了云原生、分布式数据产品和RDS数据库云服务。

### 开源数据库的使用风险分析

#### 合规风险

与开源许可证协议、版权和出口管制相关的风险。例如，GPL V2协议的强制开源和传染风险

#### 断供风险

由于开源软件和开源托管平台可能受美国EAR出口管制条例的管控，如Oracle和GitHub、SourceForge均遵守这些条例

#### 安全风险

开源社区和用户不及时更新和升级软件漏洞，增加网络安全攻击风险；同时，国外漏洞管理要求导致国内厂商难以及时获取国外开源软件的漏洞信息

(1) **合规风险**：开源许可证协议、版权和出口管制等方面的合规风险。如GPL V2协议的强制开源和传染风险

(2) **断供风险**：如开源软件和开源托管平台受美国EAR出口管制条例管控风险，典型的MySQL的母公司Oracle会严格遵守美国EAR出口管制条例；GitHub、SourceForge等开源托管平台亦明确声明EAR条例。当前美国已对1300多家中国企业列入实体清单，限制高新技术供应

(3) **安全风险**：如开源社区和国内用户对软件漏洞更新、升级不及时，造成易受网络安全攻击风险；国外漏洞管理要求的约束控制，导致国内厂商不能及时获取国外开源软件的漏洞信息风险

来源：沙利文整理

## 国内开源数据库根社区和根生态初步形成

### 国产数据库的发展

#### 政策推动

为加快自主创新和自主知识产权的国产数据库发展，国家出台了相关发展政策

#### 自主开源根社区发展

基于自主知识产权和国内自主开源根社区的国产开源软件得到了快速发展，如openGauss、TiDB等开源数据库在国内多个行业的应用越来越广泛

#### 产业生态构建

国产开源数据库构建自主创新、自主知识产权生态，降低用户成本和风险，提供本地化支持

为加快中国自主创新、自主知识产权的国产数据库的产业发展，国家出台了国产数据库自主创新发展政策，基于自主知识产权、国内自主开源根社区的国产开源软件也得到了快速发展。如openGauss、TiDB等开源数据库在国内金融、电信、政府、制造等行业的应用越来越广泛。国产开源数据库基于国内自主开源根社区，构建自主创新、自主知识产权开源数据库社区、产品和产业生态，更好地支持和适应国内市场需求和特点，满足不同行业客户的业务需求和行业特性，降低用户的成本和风险，提供本地化的技术支持和服务，促进国内软件产业的发展和 innovation，培养本土的软件人才，提高软件质量和安全性，以及推动开源社区的交流和合作。

### openGauss的特性和影响

#### 技术层面

openGauss采用木兰宽松许可证v2发行，提供面向多核架构的极致性能、全链路的业务、数据安全、基于AI的调优和高效运维的能力

#### 社区和产业生态

截至2023年底，近600家企业和6200名开发者参与openGauss社区贡献。15家DBV合作伙伴或企业用户已发布基于openGauss的商业或自用版

openGauss是华为推出的基于国内自主根社区、自主创新、自主知识产权的新一代企业级开源数据库。它采用木兰宽松许可证v2发行，提供极致性能、全链路业务与数据安全、AI调优和高效运维能力。截至2023年底，近600家企业和6200名开发者参与openGauss社区贡献。15家DBV合作伙伴或企业用户已发布基于openGauss的商业或自用版。至2023年，openGauss及其DBV伙伴版累计完成6.1万套企业装机应用，2023年新增3.09万套openGauss装机部署：其中2.19万套为线下集中式部署，占2023年中国线下集中式数据库新增装机市场的21.9%，另外0.9万套为云数据库服务部署。

### 2023年openGauss数据库中国重点行业新增装机量

行业	金融	电信	政府	制造	能源	教育	其它	合计
装机量(套)	3400	8000	4200	350	700	550	4700	21900

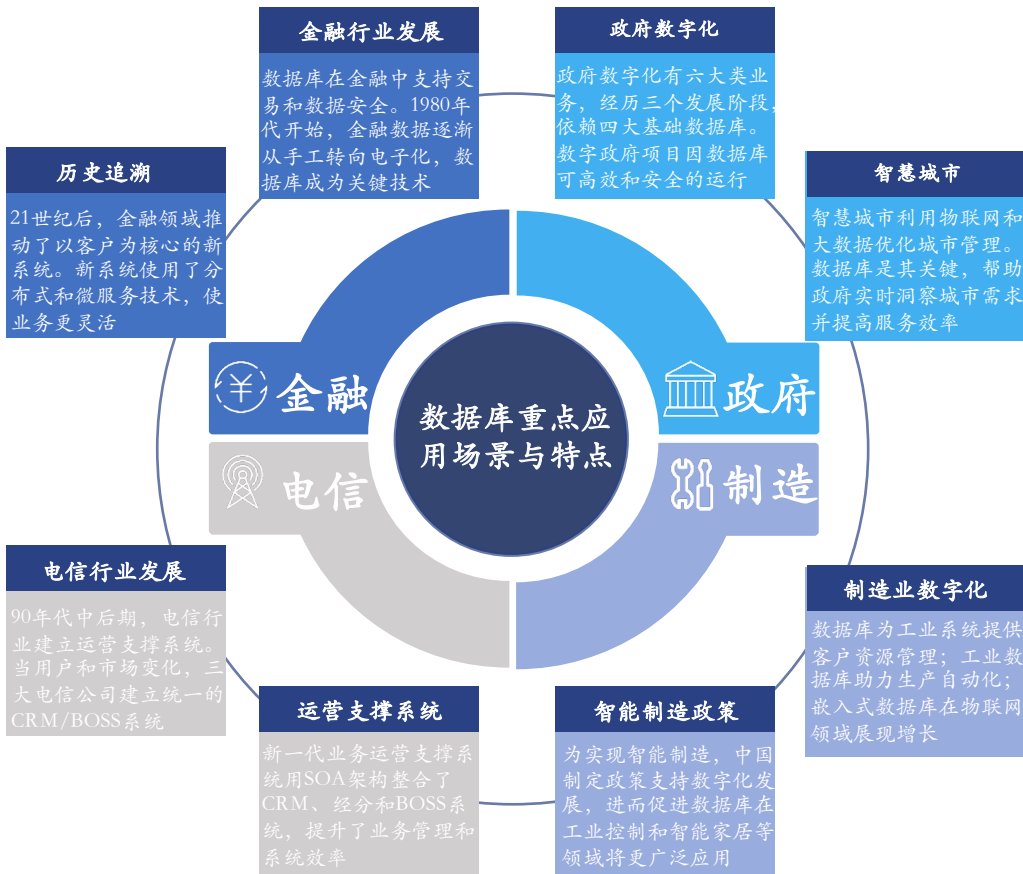
来源：沙利文整理

## 中国关系型数据库在四大关键行业中的应用与特点

### 关键发现

- （1）金融行业数据库的核心系统从集中式数据库架构转向面向客户为中心的分布式微服务架构，以更好地进行业务处理和数据管理；（2）电信行业数据库应用经历了从基本程控化到现代化的发展，进而实现新一代业务运营支撑系统的整合和优化；（3）政府行业数据库应用呈现多元化特点，涵盖数字政府、智慧城市等多个领域，成为支撑政府决策、城市管理和公共服务的关键基础设施；（4）制造业数据库在CRM/ERP系统、工业数据库和嵌入式数据库三大方面发挥核心作用，支撑企业的客户关系管理、内部资源优化和工业生产数据处理。

### 关系型数据库在四大关键行业中的应用与特点



来源：沙利文整理



## 金融行业数据库应用场景与特点

- 数据库是金融行业业务信息系统的核心数字化基础设施，承载与支撑融业务交易、交易与历史数据保存、数据安全、风险审计等关键的业务职能。金融行业也是中国数据库市场的最大行业用户，占据中国数据库市场总体20%以上份额。
- 数据集中以及新兴业务的快速发展，伴随着系统长期运行，逐渐暴露了一些问题。主要是核心系统越来越庞大，数据容量快速增长，业务与系统耦合越发严重，形成了所谓的“胖核心”，造成了建设、扩容及运营成本的快速增加，每次修改都感觉牵一发而动全身，而且开发新功能时会发现改动与评估内容很多，开发耗时越来越长，无法做到快速的响应业务变化，限制了业务的发展。

### 金融行业数据库应用发展历程

#### 1980-1990年代

金融业从现金交易、手工记账的时代，开启了电子化、信息化、电子交易的技术演进和现代化改造进程，数据库成为金融行业信息系统的核心，为金融行业的快速发展、战略转型和增强国际竞争力提供了有力的技术支持。

#### 1990-2000年代

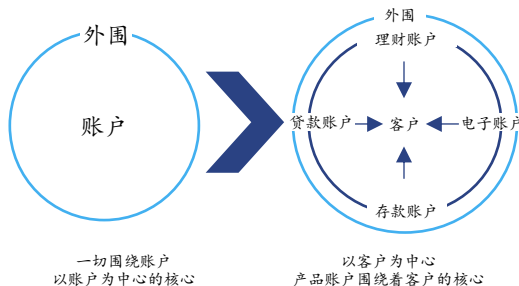
金融行业为满足金融业务快速发展，解决推动电子化、信息化初期建设阶段存在的分支行系统各自为政、数据与业务打通、共享等业务问题，开始建立统一的数据中心和数据库。在这个阶段，普遍选择引入业界成熟的IOE (IBM、Oracle、EMC) 模式，构建集中式的金融核心数据库系统。

#### 2000-2010年

中国社会生产力和经济快速发展，国内商业银行开始规模化经营的快速发展路线，营业网点高速扩张，加大拓展代收代缴、代售保险、基金等业务，新一代渠道建设。在这个阶段银行数据库系统的建设也得到了快速发展。

- 为了解决“胖核心”系统庞大耦合的问题，以国有大行为先导，提出了建设“新一代核心系统”的概念。
- “新一代核心系统”实现了从传统的“面向账户”为主转向以“面向客户为中心”的核心系统机构及业务转变。新核心在系统架构上采用分布式、微服务技术架构，通过搭建组件化、参数化、平台化的企业级系统架构平台，在高度抽象账户、产品、合约等银行业务基本要素基础上，采用“高内聚、松耦合”应用设计，实现交易、核算、支付、账户等业务处理解耦，支持产品灵活配置和业务快速扩展。

### 金融核心系统的业务演变示意图



来源：沙利文整理

# 电信行业数据库应用发展与现状

## 电信行业数据库应用发展历程

### 1987-1996年代

该时期是电信运营支撑系统的起步阶段，中国电信业发展迅猛，电话交换设备基本程控化。运营支撑系统基本以一个地城市本地网有一套系统。这一时期可以看作是正式进入运营支撑系统的建设阶段，但是系统的建设还不完整、不完备，也缺乏统一的规划。

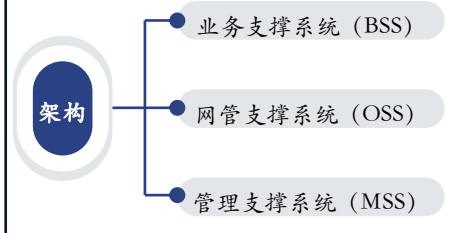
### 1997-2005年代

中国电信运营支撑系统进入了快速发展的轨道，初步形成了运营支撑系统整体框架的雏形。通信网络规模与用户规模井喷式高速发展，信息技术快速发展（包括计算机硬件技术的成熟，ORACLE、SYBASE、BB2、等数据库软件的普及等）也成为运营支撑系统开始腾飞的关键因素之一。

电信市场基本完成了从垄断到市场化经营的转变

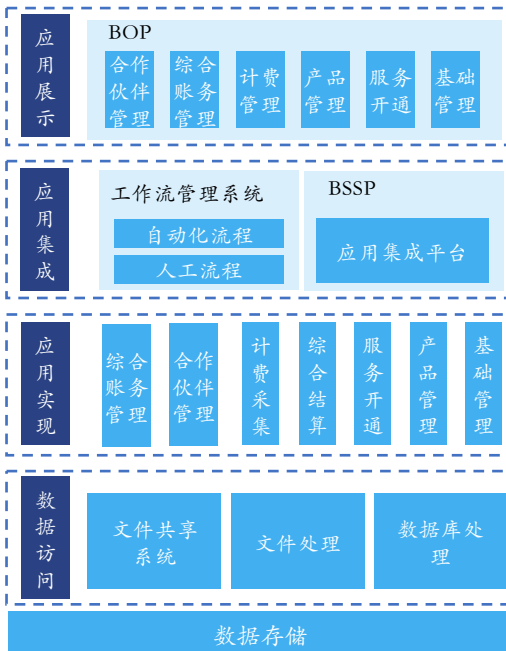
中国三大运营商相继建成了具有里程碑意义的集团总部+省级集中化的CRM/BOSS系统，建立集团客户统一数据资料、一本账、统一付费，支持跨省业务调度，实现一点受理，一点支付，全网服务能力。

初步形成电信运营支撑系统三大架构体系



- 新一代业务运营支撑系统（NGBOSS）是基于SOA架构理念设计，将CRM与经分和BOSS的拆分解耦，实现统一接口平台、全业务运营、全网资源管理、业务流程管理能力。是将原有BOSS系统的营业部分和客服系统进行整合成CRM系统，实现各渠道的一体化营销与服务能力，构具备营销服务一体化能力的CRM系统；将计费、帐务等模块从现有BOSS系统中拆分独立，避免因后台系统进行出帐等大规模运算时降低前台的营销服务质量，而计费、账务等核心数据库平台能够采用数据库一体机等更先进、优化的数据库平台，提高数据库系统的性能和效率。

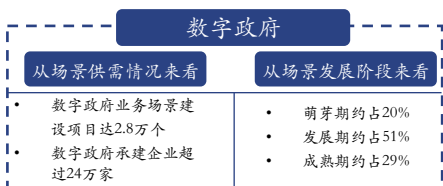
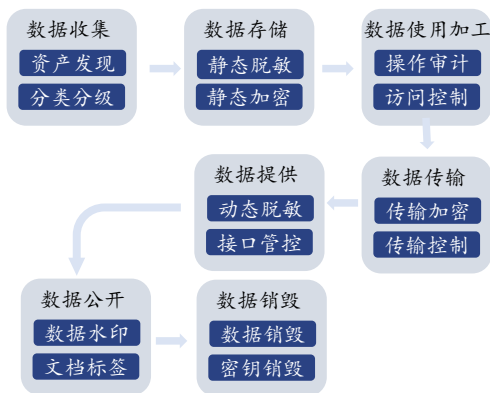
## 新一代业务运营支撑系统（NGBOSS）架构图



## 政府行业数据库应用场景与特点

- 中国政府行业数据库市场销售规模在全行业中排名第二，仅次于金融行业，占20.7%的行业市场份额。政府行业有广泛的数据库应用场景，从数字政府、智慧城市，到公安、人设、工商、税务等众多应用场景。
- **数字政府**：数字政府典型业务场景可梳理归纳为：经济调节类、市场监管类、社会管理类、公共服务类、环境保护类、政务运行类六大类、22类二级业务场景、89类三级细分场景。从场景供需情况看，“十四五”以来中国数字政府业务场景建设项目达2.8万个，数字政府承建企业超过24万家。从场景发展阶段看，从场景市场需求潜力与获取项目机会两大维度出发，将89类数字政府三级细分场景及3类综合场景的市场供需格局进一步划分为萌芽期、发展期和成熟期三个发展阶段，其中萌芽期约占20%，发展期约占51%，成熟期约占29%。数字政府需要四大基础数据库做为支撑：人口基础数据库、法人数据库、宏观经济数据库，自然资源与空间地理数据库。同时数据库在数字政政建设中也更有广泛的应用场景，发挥着重要的作用和意义。选择领先的数据库产品及解决方案，可保障数字政项目和应用系统的高性能、安全、可靠地运行，保证数据的完整性、一致性和准确性，防止数据的泄露和丢失，保护数据的安全性和隐私性。
- **智慧城市**：智慧城市是指在城市规划、设计、建设、管理、运营等领域中，通过物联网、云计算、大数据、空间地理信息集成等智能计算技术的应用，使得城市管理、教育、医疗、房地产、交通运输、公用事业和公众安全等城市组成的关键基础设施组件和服务更互联、高效和智能，从而为市民提供更美好的生活和工作服务、为企业创造更有利的商业发展环境、为政府赋能更高效的运营与管理机制。
- 智慧城市以信息技术为基础，利用信息技术、数据库及大数据分析，实现城市管理、服务、规划和创新的一种新型城市模式。智慧城市的核心是数据库，它可以存储、处理和共享各种类型的城市数据，为智慧城市的各个领域提供支撑和价值。通过数据库及数据分析，政府可以实时监测和分析城市的各项指标，如交通、环境、安全、能源等，及时发现和解决问题，优化资源配置，提升公共服务质量，增强城市的可持续发展能力。

### 智慧城市数据处理周期安全



来源：沙利文整理

## 制造业数据库应用场景与特点

- 中国制造业是国民经济的重要支柱，也是数字化转型的重点领域。数据库作为数据管理的核心技术，对制造业的数字化发展起着至关重要的作用。制造业数据库的应用场景主要包括以下几个方面：

### 1. CRM、ERP系统数据库：

- 数据库是CRM、ERP系统的核心，大型企业的CRM、ERP系统的数据库量庞大，业务流程复杂，对数据库的性能、稳定性、高可用性和安全性要求有非常高的要求。制造业数据库品牌以SQLServer、SAP HANA、DB2、Oracle为主，海外数据库市场占比约为80%。国产数据库占比20%。

### 2. 工业数据库：

- 工业数据库包含用于存储和管理工业领域的时序、实时数据库，以及用于设计、生产的CAD、CAM数据库。工业数据库能够高效地处理大量的时序、实时数据，以及用于设计、生产的CAD、CAM数据，为工业生产提供可靠的数据存储和访问功能。在工业自动化、物联网和大数据等领域，工业

数据库发挥着重要的作用。中国作为工业制造大国，也相继颁布多项政策文件，倡导智能制造，反复明确强调了工业数字化发展的大方向。工业数据库作为实现智能制造的基础软件，其发展同步受到政策推动。中国实时数据库市场规模将于2025年达到269亿元，CAGR为36.6%。

### 3. 嵌入式数据库：

- 是一种专为嵌入式系统设计的数据库管理系统。它具有微小内核、低安装系统包、低算力和内存要求，以及高效、稳定、可靠的特点，能够提供快速的数据存储和检索功能，为嵌入式设备的数据处理提供了强大的支持。嵌入式数据库在工业控制系统、移动设备、物联网设备、智能家居等系统都有广泛的应用。随着物联网、人工智能等技术的快速发展，国产嵌入式数据库将迎来更广阔的应用前景。未来在工业制造领域，国产嵌入式数据库将更加注重数据安全性、性能优化和智能化的发展，为各行各业提供更好的数据处理解决方案。

### 制造业应用场景特点



来源：沙利文整理

## 国产数据库在关键行业应用上有五大态势

### 关键发现

- (1) 在自主创新的趋势下，国内各关键行业数字化创新、转型正从一般应用系统向核心系统转型；(2) 集中式数据库如Oracle和DB2在中国关键行业中扮演着关键角色，同时，国产集中式数据库已能为关键行业提供稳定和可靠的解决方案；(3) 金融和电信等核心行业正在加速转向分布式数据库，以应对高并发、大数据和实时处理的需求；(4) 在数字化时代背景下，“数据安全”已成为中国国家战略核心，对各行业应用提出了更加严格的数据保护和和管理要求；(5) 数据库整体兼容性、迁移适配的案例越来越多，业务方法论逐步成熟

### 国产数据库在关键行业应用上有五大态势

#### 1 数字化创新升级由一般系统转向核心系统

**行业创新推动** 各行业普遍致力于自主创新，尤其是金融、能源、电信等核心行业正深入进行核心系统改造

**国家要求** 国家要求2025年行业完成创新改造，构建数字化新升级环境，2023年底形成中等规模解决方案

**金融业核心应用情况** 大银行已完成替换，城商行正在验证。替代周期长，国产技术面临集成和生态问题

#### 2 集中式数据库在OLTP核心系统的作用不容忽视

**主导地位** 以Oracle、DB2等为代表的集中式数据库在中国金融、电信等行业占主导地位，近几年，分布式数据库虽有所突破，但集中式数据库仍保持80%以上市场份额

**发展成熟** 具备支持金融应用的能力，迁移至同类型国产数据库难度小、成本低、风险可控

#### 3 分布式数据库成为金融、电信等行业新选择

**IT架构转型需求** 金融、电信等行业转型要求分布式应用系统，分布式数据库实现数据海量存储和高并发处理

**优势** 在性能、可靠性、成本和安全可控性方面，分布式数据库具备优势，满足金融业务的高要求

**行业应用** 在金融、电信等行业广泛应用，关键在主机下移、新平台建设、湖仓一体数据平台构建中发挥关键作用

#### 4 应用更关注数据安全

**数据安全战略** 中国正建立健全数据安全治理体系并提高保障能力，对数据流通的管理和应用提出更高的安全要求

**数据价值挖掘** 以数据为中心，整合计算、存储、网络和软件资源，加速数据共享和融合，是应对数据安全挑战的关键

#### 5 兼容性、迁移适配方法论逐步成熟

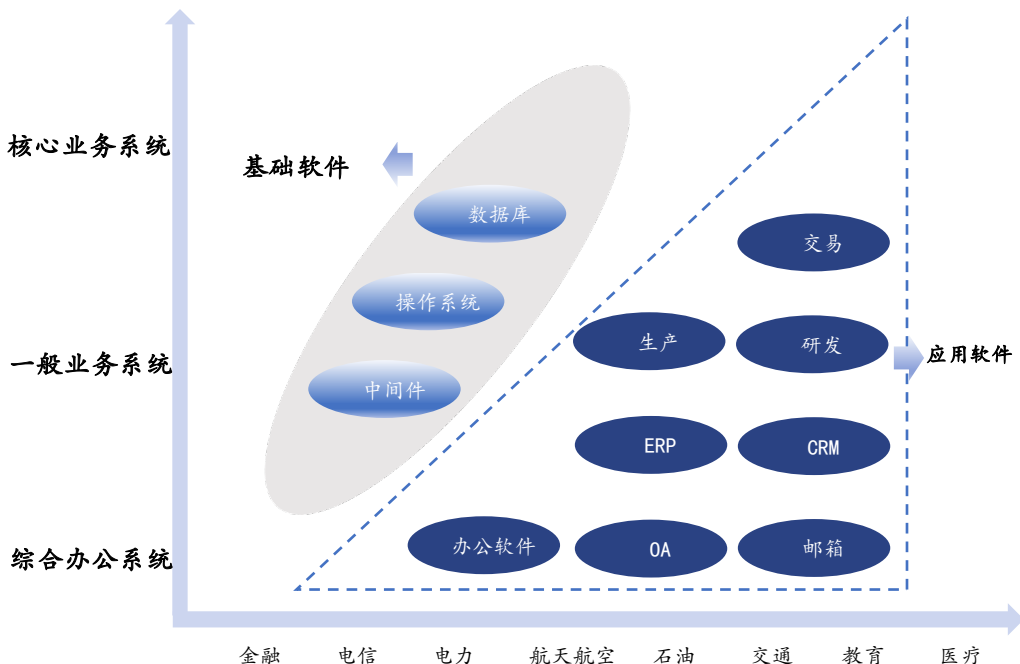
随着各行业客户迁移的陆续开展，整体的迁移工程案例越来越多，迁移方案论已经日趋成熟，整体过程分为5个步骤

迁移评估 应用适配 数据迁移 试运行 生产割接

## 数字化创新升级由一般系统转向核心

- 在各行业自主创新背景下，国内14个关键行业客户陆续都在开展应用系统创新改造，尤其是金融、能源、电信等重点行业，已进入深水区开展核心系统的改造。参考国家对行业的创新改造要求，在2025年，一般应用系统原则上完成创新改造，存量可双轨/异构兼容；关键业务系统，构建国产基础软硬件的迁移环境，初步形成替代解决方案，并在23年底要形成中等以上规模复杂场景解决方案。
- 以金融核心应用为例，承载着日均数亿级别复杂金融交易的运转，对稳定性和安全性有极高要求，当下大部分商业银行依旧维持集中式为主，大行基本已经完成方案验证和陆续替换进行中，城商行已启动方案验证。在替代的同时，并行结合未来业务需求，开展微服务技术架构重建。整体而言，核心系统对银行是命脉，不同客户技术架构不同，开发商不同。目前国产数据库、虚拟化、云原生容器等领域都在做优化，不同厂商之间的集成等生态问题是最大的挑战，替代仍需要比较长期的工作周期。

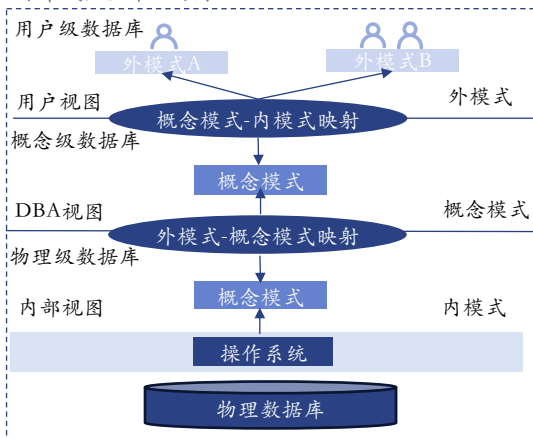
应用系统数据库发展趋势



## 集中式数据库在OLTP核心系统的作用不容忽视

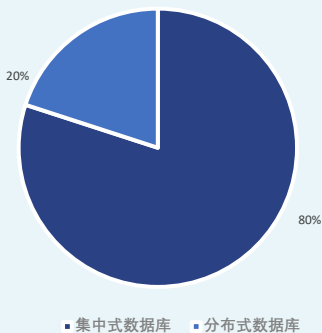
- 传统的集中式数据库在中国金融、电信等行业电子化、信息化建设过程中，以及数据大集中时代占据绝对主导地位，以 Oracle、DB2、MySQL、PostgreSQL 等为代表的集中式数据库，以高效的数据存储效率和优异的系统稳定性首先在中国金融业得到广泛应用，有力支持了金融、电信等行业数据中心和业务系统开发、建设、业务应用实践。
- 集中式数据库是数据库行业发展最早、技术最成熟的数据库技术，具有成熟的数据强一致性、严格的ACID 事务技术特性，支持多机主备或多活HA 集群和两地三中心高可用方案，能够更好地满足银行、保险、证券等金融业务交易的高复杂度、数据强一致性、高性能、低延时的金融业OLTP 业务场景。集中式数据库在中国传统线下部署数据库行业应用中一直处于主导地位，当前仍保持80%以上的份额。

集中式数据库结构图



- 国产集中式关系型数据库已经历了近30年发展，技术产品已相对成熟，已经在政企等不同领域得到较为充分的应用和打磨，国内主流的商业数据库产品和自主开源社区版本已经比较成熟，初步具备支持金融应用的能力，基本满足金融行业对可靠性、可用性、高性能、高安全性的需求。
- 国产集中式数据库厂商经过多年发展以及金融业应用探索，陆续研发数据库内核级兼容相关语法的评测工具、异构平台下的数据迁移工具、海量数据处理场景下的数据交换工具、数据治理工具以及决策支持类的大数据可视化或智能软件等，并得到实践检验。这些工具对于提高迁移效率、降低迁移成本和防范风险发挥重要作用。中国主要的集中式数据库产品对传统主流数据库进行大量兼容适配开发，同时对于开发人员和运维人员，其原有的技术、技能均可沿用，为中国数据库产品应用规避知识和技能壁垒。

集中式和分布式数据库市场份额占比情况

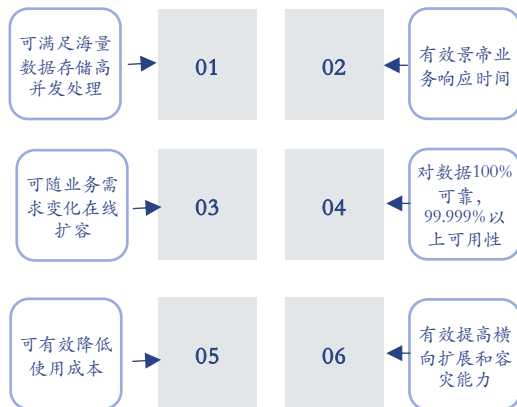


来源：沙利文整理

## 分布式数据库成为金融、电信等行业的新选择

- 金融、电信等行业IT架构转型加速分布式数据库应用现代金融、电信服务逐步从柜面向网上、移动响应式服务转变，为用户带来极大的便利。现代金融、电信业务系统的海量数据、高并发、实时处理等需求推动了新一代分布式应用系统建设。由于分布式数据库所具备计算与存储分离、弹性扩展、高可用等特征，契合新一代金融分布式体系建设要求，迅速实现推广应用。
- 分布式、云化是数据库未来发展趋势已成为多方共识，特别是分布式数据库已在金融、电信行业实现了突破应用，具有明显的优势。在整体性能和扩展性方面，分布式数据库利用普通PC服务器，通过横向扩展，实现存储容量和处理性能的线性提升。不仅满足应用系统PB级海量数据存储和百万级TPS/QPS高并发的处理要求，而且有效降低业务响应时间，提升客户满意度。同时，可随业务需求变化，进行系统在线扩容，避免硬件资源浪费，提高性价比。
- 在高可靠和高可用方面，分布式数据库采用多副本冗余存储，将数据副本存储在多台设备上，避免存储冗余不够导致的不稳定问题，保障数据可靠性。通过一致性复制协议，保证数据副本间的一致性。数据副本可跨可用区、跨数据中心或城市部署，满足可用区级、数据中心级或城市级的不同等级的容灾要求。在少数机构、个别数据中心故障或单城市发生灾难时，可实现数据不丢失、服务不中断。满足金融、电信等核心业务对数据100%可靠、99.999%以上可用性极致要求。
- 在成本和安全可控方面，分布式数据库基于开放PC服务器，取代数据库对高端小型机和大型机的依赖，硬件和维保资源投入显著降低。同时，数据存储在低成本本地磁盘，替代昂贵的高端集中式磁盘阵列，并利用高压缩比数据存储技术，可进一步降低存储成本。依托丰富的应用场景，中国数据库厂商在分布式数据库技术和产品上实现了弯道超车，实现了中国分布式数据库产品的深入应用和创新发 展，降低了系统对国外高端硬件的依赖，提升安全可控水平。
- 总之，现代分布式核心业务系统利用分布式数据库，以更低的成本，更多的硬件选择，实现更大规模的数据存储能力、更高的业务处理响应能力、更佳的横向扩展能力和更好的容灾能力，显著提升系统处理能力，保证业务连续性，提升客户满意度。

分布式数据库特征

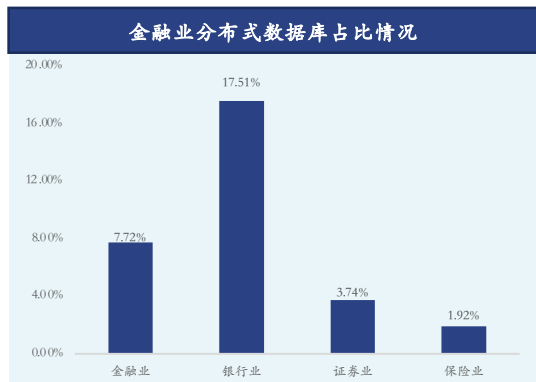


来源：沙利文整理



## 分布式数据库成为金融、电信等行业的新选择

- 分布式数据库在金融、电信行业用成效明显。分布式数据库已在中国金融、电信等不同领域逐步开展应用，涵盖不同类型的业务系统，总体占比达到7%，其中银行业超过了17%，证券业和保险业相对较低。
- 目前，分布式数据库在主机下移、新一代分布式平台建设以及湖仓一体的数据分析平台构建中加快应用，积累可供行业借鉴的实践经验，全面提升安全可控能力，助力金融、电信行业数字化转型和快速发展。



### 分布式数据库在关键系统 主机下移中发挥重要作用

传统银行、电信关键系统运行在集中式架构下，体系相对封闭、扩展能力差、建设成本高、迭代升级困难等。分布式数据库通过普通服务器堆叠，实现海量数据存储、高并发处理能力，通过多副本实现系统高可靠，满足核心系统容灾和业务连续性要求，有力支持核心系统架构向开放式架构转型，降低建设成本。



### 分布式数据库在构建新一代信息系统中作用明显

当前金融、电信等行业系统数据规模越来越庞大，联机交易和实时分析要求高，分布式数据库具有良好的扩展性，适合海量数据存储和处理，实现多元全量数据高效存储、联机处理与联机分析混合处理，适合金融、电信业务中对联机交易和联机分析有较强需求的场景，可用于实时风控、反欺诈等实时数据分析系统。



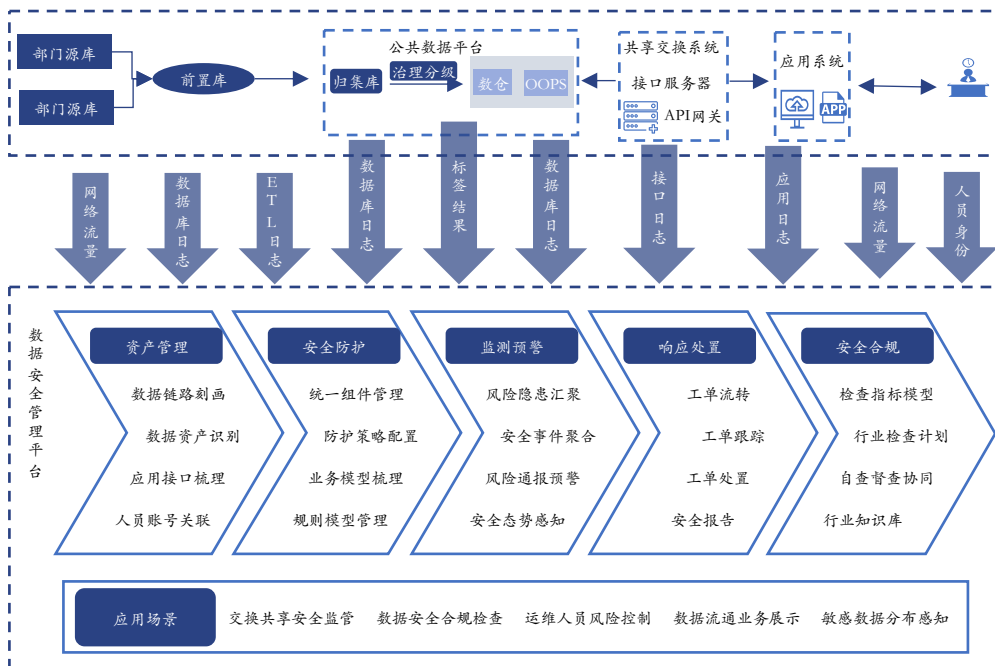
### 分布式数据库支持构建湖仓一体数据平台

将数据仓库的高性能、管理能力及数据湖的灵活性相互融合，支持多种数据类型并存，实现数据间相互共享，并通过统一封装接口进行访问。湖仓一体使数据入湖后可原地进行数据处理加工，避免数据多份冗余及流动导致的算力、网络及成本开销，同时有效支持开展数据治理。

## 应用更关注数据安全

- 各行业的应用不可避免都要跟数据对接，而随着数字化时代的来临，数据已经成为国家的生产要素，“数据安全”已上升到了中国国家安全战略高度，明确提出要“释放商业数据价值潜能，加快建立数据产权制度，开展数据资产计价研究，建立数据要素按价值贡献参与分配机制”。《数据安全法》坚持总体国家安全观，建立健全数据安全治理体系，提高数据安全保障能力。当前，中国数据要素市场发展尚处于起步阶段，对如何管理各应用中流通的数据提出了更高的安全要求。
- 以数据为中心，深度整合计算、存储、网络和软件资源，充分挖掘数据价值，加快数据共享和融合，对使数据“存得下、流得动、用的好”提出了更多的要求。以金融为例，关键数据安全成为持续影响金融科技企业最主要的网络安全风险，主要挑战有数据全生命周期的隐私保护问题，即系统无论在何种业务场景和环境下，数据在传输、运算以及存储的各个环节始终都要处于密文状态。当数据拥有者在客户端完成数据加密并发送给服务端后，在攻击者借助系统脆弱点窃取用户数据的状态下仍然无法获得有效的价值信息，从而起到保护数据隐私的能力。

数据安全治理流程图



来源：沙利文整理

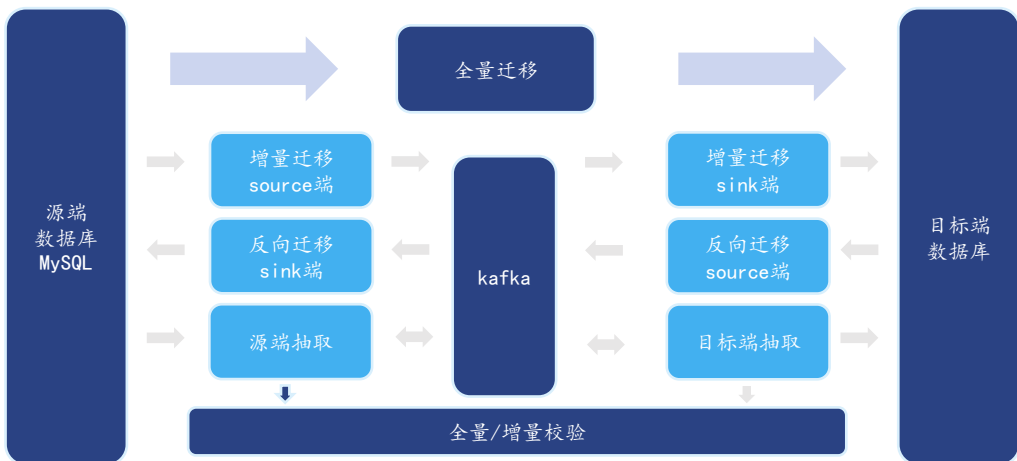
## 兼容性、迁移适配方法论逐步成熟

随着各行业客户迁移的陆续开展，整体的迁移工程案例越来越多，迁移方案论已经日趋成熟

整体过程分为迁移评估、应用适配、数据迁移、试运行和生产割接5个步骤，详情如下：

<b>目标</b>	评估源库现状，提供目标库选型和部署建议，评估目标库兼容性、改造工作量及迁移时长	基于迁移评估结果，适配修改和验证不兼容SQL，支持修改应用程序源文件	确定组网，实施对象和数据迁移	迁移流量到目标库，对目标库进行实际业务测试，修复故障、调优性能	启动反向迁移，业务割接到目标库
<b>关键动作</b>	源库画像：部署架构、容量、负载、复杂度、热点； 目标库选型和部署建议； 兼容性评估：对象、SQL兼容性评估，包括应用程序中的SQL； 迁移时长评估；	测试数据预置 修改/验证SQL 应用修改	组网部署 全量迁移 增量迁移	流量镜像 故障修复 性能调优	反向迁移 运维服务
<b>关键输出</b>	评估报告：源库画像、目标库选型和部署建议、兼容性评估、改造工作量、迁移时长	应用适配报告：适配列表、数据库配置文件	数据迁移报告	试运行测试报告：对比原数据库和新数据库的业务性能	数据库运行报告
<b>运行环境</b>	原数据库生产环境	测试环境	新数据库类生产环境+原数据库生产环境		生产环境

过程中也孵化出了各类迁移工具，整体架构图如下：



来源：沙利文整理

## 国产数据库行业在深度应用上面临着五个挑战

### 关键发现

- (1) 中国数据库产业在核心技术上面临原始创新不足、协同创新不足以及产品成熟度低的关键能力不足问题；(2) 用户的“路径依赖”、高端场景应用困境和跨厂商迁移难度是国产数据库行业应用迁移替代难度大的三大核心挑战；(3) 中国数据库企业面临企业规模差距悬殊和技术路线分散、同质化竞争等产业生态力量布局分散问题；(4) 国内基于国外开源数据库的衍生产品面临安全漏洞、开源协议约束和供应链安全等多重产业链风险；(5) 国内数据库行业面临核心人才缺失、待遇和培养模式滞后等问题

### 数据库应用深化面临的问题和挑战



## 关键核心技术能力不足

关键核心技术是国之重器，对保障国产数据库安全可控具有十分重要的意义。中国数据库产业面临部分“卡脖子”技术问题的根源主要在于：

- 一是原始创新能力不足。国产数据库多是基于MySQL和PostgreSQL等国外开源产品内核进行封装和增强，或者跟踪参照Oracle、IBM DB2等商业数据库，自主研发产品的基础研究投入不足，普遍缺乏技术亮点和自主发展能力。数据库技术创新更多体现于在国外既有技术基础上的集成创新，形成了在别人的架构之上进行开发的局面，鲜有对行业产生重大影响的原发性、基础性创新，数据库理论研究和创新亟需强化和突破。
- 二是协同创新能力不足。中国数据库企业呈现“多散杂”特征，与产业上下游厂商协同创新不足。国外Oracle、IBM等厂商不仅建立了深厚的技术壁

垒，更培养了完善的产业生态和上下游协同创新能力。如在金融领域关键核心业务，IBM硬件大型机与DB2数据库整体软硬件结合的解决方案，具备较强的故障冗余、故障监控、快速的容灾倒换等能力，据悉能保证十余年不出故障。目前国内数据库厂商基础软硬件解决方案尚难达到此种水平，上下游厂商协同研发创新能力还不够。

- 三是国产数据库产品成熟度不足。产品在功能完备性、稳定性、兼容性、扩展性等方面与国外主流产品差距巨大，难以支撑重点行业复杂业务场景应用。部分产品购买国外厂商源码授权，虽然加速了数据库产业发展起步，但是产品核心技术架构短期内无法完全消化掌握，更不可能加以调整，导致产品在版本更新、客户新需求响应等方面科技创新处于劣势，亟待突破。

### 关键核心技术能力分析



#### 原始创新能力不足

国产数据库依赖国外开源/商业内核，自主研发不足，缺乏技术亮点和自主发展能力。创新多为集成创新，缺乏原创性、基础性创新



#### 协同创新能力不足

国产数据库企业“多散杂”，与上下游厂商协同不足。Oracle、IBM等国外厂商技术壁垒深厚，产业生态完善，软硬件结合解决方案能力强



#### 产品成熟度不足

国产数据库与国外主流产品差距大，难以支撑重点行业应用。部分购买国外源码加速发展，但短期内难掌握核心技术，科技创新处于劣势，亟待突破

## 行业应用迁移替代难度大

- 行业用户存在“路径依赖”，切换动力意愿不足。数据库对于用户系统的黏性较强，产品更换、应用迁移难度大。且金融、电信等关键行业领域的核心业务系统对数据库一致性和可靠性的要求极高，如果简单通过改造和重建的方式向国产数据库迁移上层应用，时间和费用成本极高。一旦核心出现故障会导致全部业务的中断甚至瘫痪，用户单位难免需要承担责任。用户侧“自上而下”对核心业务进行替代的决心和容错机制不足，国产数据库难有“有武之地”。
- 二是高端场景应用难以破局，缺乏良性循环迭代环境。目前大部分国产数据库得不到核心业务场景的应用机会，替代一套数据库系统的成本很高，但场景的应用和技术的打磨是相辅相成的。金融、电信等行业核心关键领域、复杂度高的场景涉及核心业务、高价值等问题，客户担心替换后出现重大事故而不愿开放核心关键场景给国产

数据库厂商替换，导致厂商只能替换场景复杂度低、低价值的业务，不利于数据库产品在复杂业务场景中迭代升级。

- 三是跨厂商应用迁移困难，公共适配和测试资源匮乏。甲骨文等国外厂商在兼容性、标准接口、语义语法等方面设置排他性的“技术门槛”，跨厂商应用迁移难度大、成本高，国内仍然缺乏支持异构数据库之间数据平滑迁移的成型工具产品，应用迁移过程中业务架构和代码修改调整的失控风险极大，迁移后产品应用调优工作繁琐，对用户企业进行数字化创新升级造成巨大压力和难度。同时，数字化创新是一项系统推进的总体工程，国产数据库与芯片、操作系统、应用软件、网络设备等国产软硬件的配套使用过程中，对数据库与国产软硬件的兼容能力、无缝衔接能力以及自身容错能力要求很高，产品应用调优难度较大。

### 行业应用迁移替代分析

行业用户存在“路径依赖”		高端场景应用难以破局		跨厂商应用迁移困难	
01	用户系统黏性强	01	核心业务场景应用机会少	01	排他性技术门槛
02	迁移成本高	02	替代成本高	02	缺乏平滑迁移工具
03	一致性和可靠性要求极高	03	客户担忧事故风险	03	繁琐的调优工作
04	替代决心和容错机制不足	04	迭代升级受限	04	整体系统推进难度大

## 产业生态力量布局分散

- 一是企业“小、散、弱”尚未根本性改变。国内企业与甲骨文等国外龙头企业呈量级差距，缺少支柱型企业。甲骨文公司在2023年第一季度其营收就达到了816.04亿元，反观国内，传统数据库龙头达梦在2022年全年实现营收6.88亿元人民币，只相当于甲骨文单季度营收的1/118，企业“散、小、弱”现象突出，严重制约其在研发、运维、服务支持等方面的投入，尚难突破“低端锁定”困局。
- 二是技术路线分散、同质化竞争严重。目前国内数据库厂商产品数量众多，数据库厂商有两百多家，各类技术产品分叉演进，水平参差不齐，产业呈

碎片化发展，没有形成集聚发展效益。数据库产品研发的技术路线也各不相同，产业整体处于各自竞争阶段。技术力量的分散在一定程度上限制了数据库产业发展速度。即使排名第一的国产数据库企业市场份额也不足6%，仅为Oracle市场份额五分之一。当前行业用户侧部署多套系统带来的一致性与运维的需求，需进一步促使技术栈收敛，产品在行业通用性、跨平台兼容性、安全性需要进一步优化。下一步，中国数据库产业须进一步推动技术收敛，加快从“数量型”向“质量型”转变。

### 行业端国产替代供应商生态情况

项目	达梦数据	华为云	奥星贝斯	平凯星辰	阿里云	腾讯云	openGauss
典型产品	DMB	GaussDB	Oceanbase	TiDB	PolanDB	TDSQL	openGauss 及DBV伙伴 版本
典型行业	政府、能源、 交通	金融、电信、 政府	金融、电商、 互联网	金融、互联 网	线上教育、 在线游戏	金融、互联 网、政府	金融、电信、 政府、制造
产品架构	集中式	分布式	分布式	分布式	分布式	分布式	集中式
生态兼容	Oracle	MySQL PGSQL	MySQL为主、 Oracle	MySQL	MySQL	MySQL、 PostgreSQL、	MySQL PGSQL
技术特点	企业级功能 完备	支持分布式 事务强一致性	分布式高并 发处理	分布式事务 处理	云化按需服 务	事务-分析 混合处理	企业级 资源池化 多模多态





## 数据库专业研发人才短缺

- 一是内核研发人才缺口大。国内从事数据库底层核心设计与研发的高端人才极为匮乏，懂技术、懂架构、懂管理的数据库产品经理数量也相对不足，产业发展人才支撑能力亟待提升。
- 二是人才待遇难以保证。数据库研发周期长、回报慢，核心人才频遭互联网等其他行业“挖墙脚”，逐渐丧失“坐冷板凳”信心，人才流失现象严重，国产数据库发展面临“人才空心化”隐忧。
- 三是人才教育和培养模式落后。高校作为数据库人才培养的第一阵地，在课程教学上存在着理论课程内容依赖国外或开源数据库、教学案例与生动的实际应用脱节、缺少基于国产数据库的操作实践学习等痛点，传统人才培养理念和举措满足不了企业对数据库内核研发人才的需求，产教融合成为中国数据库人才培养模式的创新和必经之路。

### 数据库人才现状

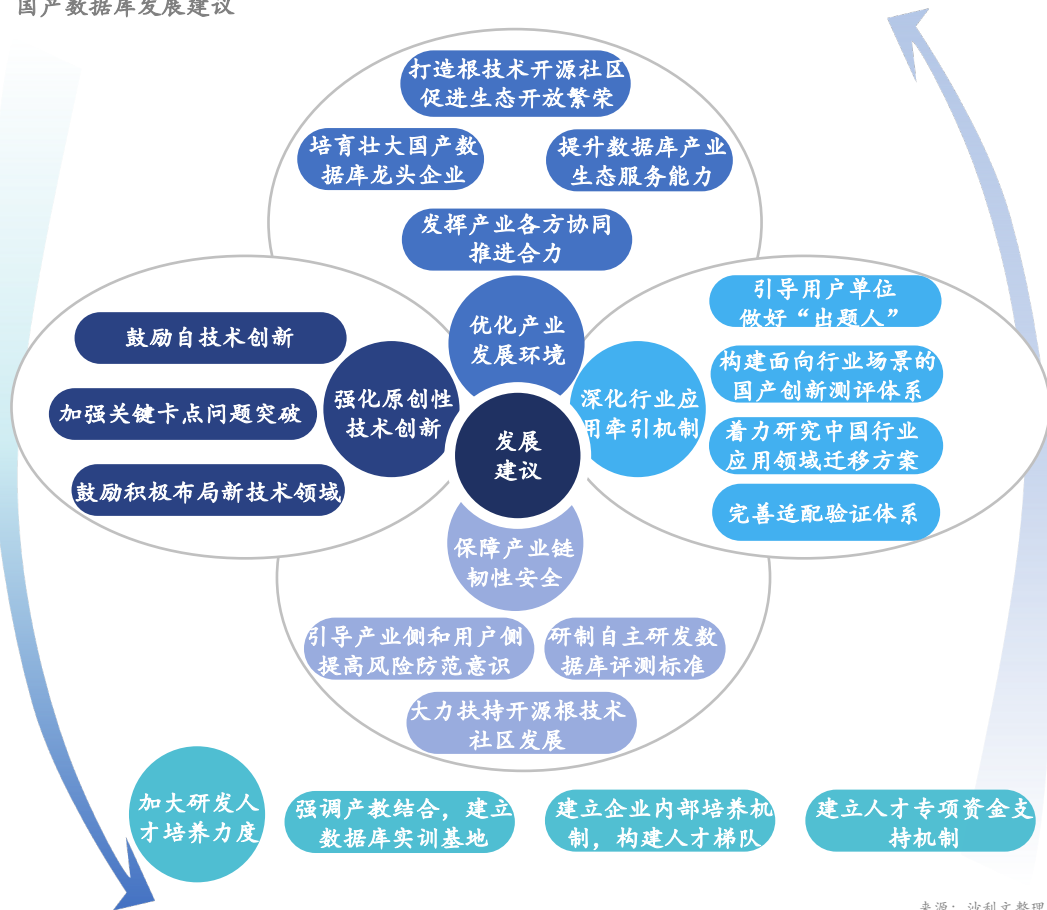


## 国产数据库高质量发展建议

### 关键发现

- (1) 强化原创性技术创新：鼓励自主技术创新，加强关键卡点问题突破，积极布局新技术领域；(2) 深化行业应用牵引机制：引导用户单位明确发展要求，构建面向行业场景的国产创新测评体系；(3) 优化产业发展环境：打造根技术开源社区，培育壮大龙头企业，提升生态服务能力，发挥产业各方协同推进合力，促进数据库产业规范化发展和生态构建；(4) 保障产业链韧性安全：提高风险防范意识，完善风险防范措施；(5) 加大研发人才培养力度：加强数据库人才培养，需产教结合，建立企业内部培养机制

### 国产数据库发展建议



来源：沙利文整理

## 强化原创性技术创新

- **一是鼓励自主创新：**抓住国际产品相继停服和国家大力发展基础软件的历史机遇，以数据库企业为创新主体，加强原创性理论研究和攻关，确保产品本身具备有效应对断供停服等事件的实力，进一步提升产品自主创新水平。
- **二是加强关键卡点问题突破：**聚焦国产数据库功能不完备、性能不足、兼容性差等共性问题，充分提炼不同行业场景、不同应用阶段的需求，推动并发控制、数据缓存、读写分离、行列整合、高性能集群等数据库关键技术研究，解决技术瓶颈问题，大力推进自研数据库加速赶超。
- **三是鼓励积极布局新技术领域：**聚焦分布式、云原生、图数据库等长板优势技术领域，推动国产数据库“补短板”和“拓长板”，加快打磨优秀数据库产品，形成技术优势。深化数据库与人工智能、隐私计算、云计算等技术的融合，推动技术升级和新一代产品布局，促进新技术新模式发展。

强化原创性技术创新示意图

### 1 鼓励自主创新

#### 历史机遇

随着国际产品相继停服和国家对基础软件的大力发展，国产数据库企业面临着一个独特的历史机遇

#### 推动自主创新

鼓励数据库企业作为创新主体，加强原创性理论研究和攻关，以确保产品具备有效应对断供停服等事件的实力

### 2 加强关键卡点问题突破

#### 国产数据库问题

国产数据库功能、性能、兼容性不足，需提炼跨行业需求，深入研究关键技术，加速赶超

#### 推进关键技术研究

充分提炼不同行业场景、不同应用阶段的需求，推动并发控制、数据缓存、读写分离、行列整合、高性能集群等数据库关键技术研究

### 3 鼓励积极布局新技术领域

#### 鼓励积极布局新技术领域

利用分布式、云原生、图数据库等技术优势，完善并扩展国产数据库，加速优秀产品成型，深化与AI、隐私计算、云计算的融合

#### 技术战略

“补短板”与“拓长板”  
加快优秀产品打磨  
深化技术融合

来源：沙利文整理

## 深化行业应用牵引机制

- 一是引导用户单位做好“出题人”。产业链上下游各方力量应加深协作，加强供需对接交流，协力优化数据库产业供给结构。行业用户单位应充分发挥对数据库产业发展的牵引作用，依托各行业不同类型应用场景，明确数据库应用深化的发展要求。坚持问题导向，开放关键典型场景，引导产业侧开展数据库产品的关键核心技术攻关，不断提升数据库产品供给质量，更好地匹配行业用户需求。
- 二是构建面向行业场景的国产创新测评体系。研制面向行业应用场景的数据库测评模型，积极开展数据库产品测试验证，形成满足各行业场景应用需求的数据库产品榜单，超越国际 TPC-C 单一性能排名，为行业用户单位提供强有力的选型依据，推动数据库向行业关键核心系统的深度融合应用，带动国产数据库产业高质量发展。
- 三是着力研究中国行业应用领域迁移方案。鼓励重点行业用户、数据库厂商、科研院所等第三方机构协力推动国产数据库实战应用标准体系的建设，组织研制面向行业应用的国产数据库接口开发、工具生态、灾备、运维、文档要求等系列标准，编制迁移适配、数据库一站式管理等实战指南。瞄准核心系统的应用需求，筛选将上层应用从国外数据库迁移到中国自主基础软硬件环境的实践方案，评估其可行性和运行效果，以支撑存量核心业务系统的平滑迁移切换，推动中国数据库向重点行业深化应用。
- 四是完善适配验证体系。依托国家数据库适配验证中心，系统开展数据库与各类软硬件适配验证，完善大规模应用的系统性配套工作。加快基础性、关键技术、典型应用等产业亟需标准研制，统一数据格式、接口标准。

深化行业应用牵引机制示意图



### 引导用户单位做好“出题人”



产业链各方应加深协作，优化供给结构。行业用户应发挥牵引作用，明确发展要求。坚持问题导向，引导产业侧进行技术攻关，提升数据库产品供给质量，匹配用户需求



### 着力研究中国行业应用领域迁移方案



鼓励行业用户、数据库厂商、科研机构合作，推动国产数据库实战应用标准建设，制定系列标准和实战指南。筛选国外数据库迁移方案，评估可行性，支持平滑切换



### 构建面向行业场景的国产创新测评体系



研制行业应用场景的数据库测评模型，进行产品测试验证，形成满足各行业需求的数据库榜单，超越国际单一性能排名，为行业用户提供选型依据，推动深度融合应用



### 完善适配验证体系



依托国家数据库适配验证中心，开展数据库与软硬件适配验证，完善大规模应用配套工作。加快研制产业亟需标准，统一数据格式、接口标准

来源：沙利文整理

## 优化产业发展环境

- 一是打造根技术开源社区，促进生态开放繁荣。坚持根技术开源开放策略，建立国内原生的数据库根社区，充分发挥开源模式协同创新优势，集聚产业资源，发挥社区合力，拓展应用生态，在市场竞争机制下促进根技术路线有效收敛，加快培育富有竞争力的数据库明星开源项目和产业生态。
- 二是培育壮大国产数据库龙头企业。引导行业龙头企业通过兼并、重组等方式发展壮大，牵引数据库产业实现强链补链固链，建立国际化创新网络，培育形成一批具有全球影响力的数据库行业国家队。
- 三是提升数据库产业生态服务能力。搭建一体化、集约化的数据库公共服务平台，打通信息整合、成果转化、科研合作、投融资等产业发展关键环节，促进信息资源共建共享、提升产业服务效率。推动专业化服务机构建设，强化面向数据库领域的规划咨询、质量保障等支撑能力，实现数据库产业配套服务的供需精准匹配。
- 四是发挥产业各方协同推进合力。成立中国数据库产业联盟、中国数据库产业发展论坛等组织，充分发挥数据库相关资源平台、行业协会、产业联盟等组织作用，制定相关产业发展公约，引导形成以产品服务能力为导向的良性市场竞争环境，促进产业规范化发展。组织数据库、支持数据库及上下游相关企业、用户单位、科研院所等生态各方共同举办数据库品牌峰会、专业论坛、创新大赛等多样化生态活动，推动技术生态融合创新，促进产业生态加速构建。

### 优化产业发展环境示意图

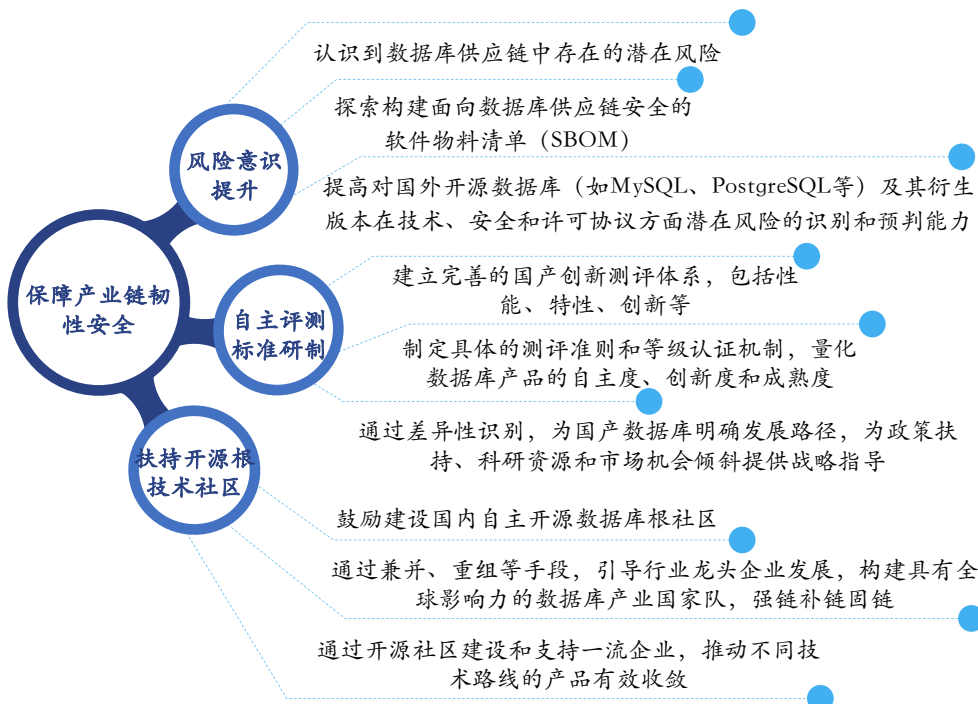


来源：沙利文整理

## 保障产业链韧性安全

- 一是引导产业侧和用户侧提高风险防范意识。探索构建面向数据库供应链安全的软件物料清单（SBOM），提升识别和预判MySQL、PostgreSQL等国外开源及其衍生版潜在的技术、安全、许可协议等方面风险的能力，全面完善风险防范措施。
- 二是研制自主研发数据库评测标准。基于国产创新测评体系识别真正能自主演进的技术路线，通过相关测评准则和等级认证机制，识别出各国产数据库产品的自主度、创新度和成熟度的差异性，确立国产数据库的根技术发展路径，为后续的政策扶持、科研资源和市场机会倾斜打下基础。
- 三是大力扶持开源根技术社区发展。鼓励建设国内自主开源数据库根社区，或基于国外开源数据库形成可自我演进的开源分支，扶持培育对标世界一流水平的数据库龙头企业和产品，促进不同技术路线产品的有效收敛，提高中国数据库的技术、产品与产业竞争力，加速促进中国数据库产业高质量发展。

保障产业链韧性安全示意图



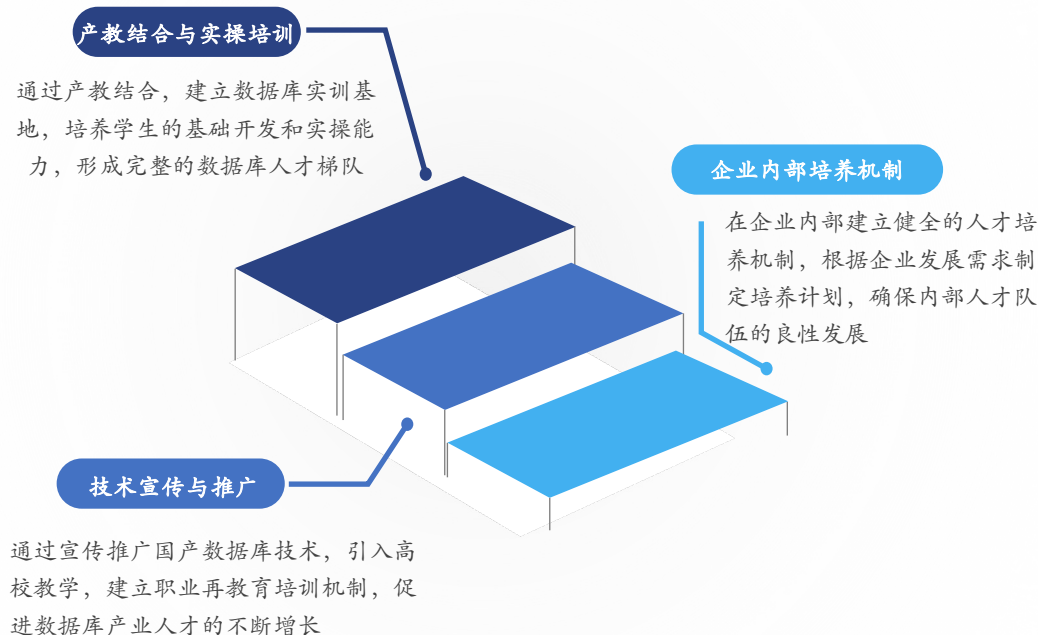
来源：沙利文整理

## 加大研发人才培养力度

- 加强数据库人才培育，一是重视产教结合，积极探索企业和高校联合培养新模式，探索建立数据库实训基地，注重高校人才基础开发能力和实操能力的培养，培育建立数据库人才梯队。二是企业内部要建立培养机制，企业应根据自身发展需求，制定人才培养机制。三是加强国产数据库技术的宣传和推广。一方面，注重国产数据库技术培训教材的编制，让国产数据库走进高校。同时，建立完善职业再教

育培训机制，进一步促进数据库产业人才增长。四是制定实施人才专项资金，对到国产数据库企业就业的应届毕业生给予住房补贴、人才补贴等，提高研发人才福利待遇，做好数据库人才的持续有效供给。

加大研发人才培养力度示意图



来源：沙利文整理

## 法律声明

- ◆ 本报告著作权归沙利文所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得沙利文同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“弗若斯特沙利文”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。在法律许可的情况下，沙利文可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。本报告所指的公司或投资标的的价值、价格及投资收入可升可跌。
- ◆ 本报告的部分信息来源于公开资料，沙利文对该等信息的准确性、完整性或可靠性不做任何保证。本文所载的资料、意见及推测仅反映沙利文于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，沙利文可发出与本文所载资料、意见及推测不一致的报告和文章。沙利文不保证本报告所含信息保持在最新状态。同时，沙利文对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。