

数据库行业深度：商业模式及发展趋势分析、竞争格局及相关公司深度梳理

近年来，中央出台多项信创相关政策，大力支持信创产业持续发展，努力实现国产替代。数据库作为信息系统的核心和信创基础软件的重要部分，将迎来重大发展机遇。据 Gartner 报告，2021 年全球数据库管理系统市场接近 800 亿美元，其中关系型数据库占比达到 80%，是全球数据库的主流。据中国信通院预测，我国数据库市场规模 2025 年将达到 688 亿元，增长迅速。

2022年全球数据库产业图谱



数据库的发展受哪些因素驱动？数据库的分类及技术溯源有哪些？从国内外数据库的发展历程结合我国发展现状来看国产数据库有哪些值得学习的地方？数据库行业的商业模式及竞争格局是怎样的？市场空间有多大？未来有怎样的发展趋势？相关公司又有哪些？下面我们进行一一梳理，希望对大家了解数据库行业有所启发。

目录

一、驱动因素	2
二、行业概览	3
三、商业模式	9
四、竞争格局及发展现状	10
五、市场空间	13
六、公司梳理	14
七、发展趋势	22
八、参考资料	25

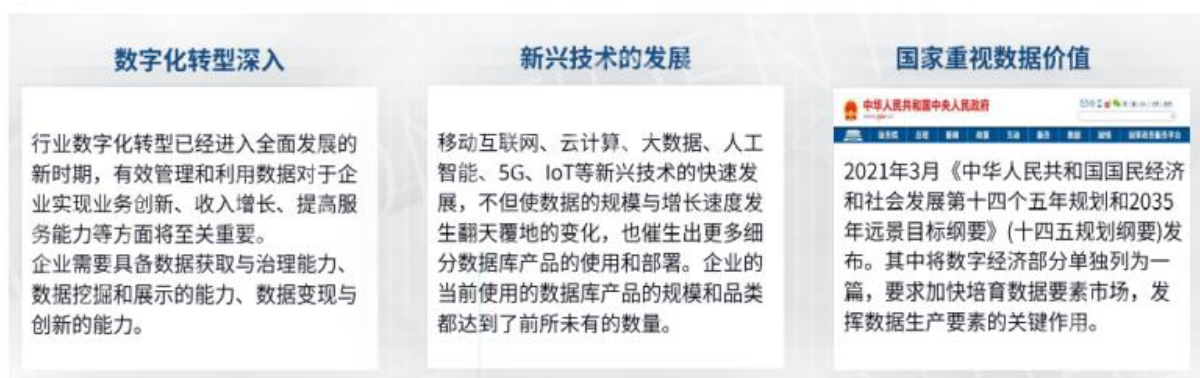
一、驱动因素

国产数据库产业黄金期加速到来，主要受四方面因素影响：

1. 核心驱动力

数字经济的发展，数据量的高增、数据结构和来源多元化、以及部署环境的日益复杂等是数据库产业的核心驱动力。IDC 预计，2018-2025 年，中国数据量复合增速达 30%，2025 年将达到 48.6ZB，成为全球第一。届时，全球数据量将达到 175ZB。

宏观环境促使企业更加重视数据库管理工具



2. 政策驱动

在信创风口下，国产数据库迎来发展新机遇。2020 年 8 月，国务院印发《新时期促进集成电路产业和软件产业高质量发展的若干政策》，同年，重点领域的信创集采加速启动，促进国产软硬件企业加速迭代。数据库等核心产品在政策驱动下，逐渐从“能用”过渡到“好用”。

2022年党政及行业信创政策密集推出

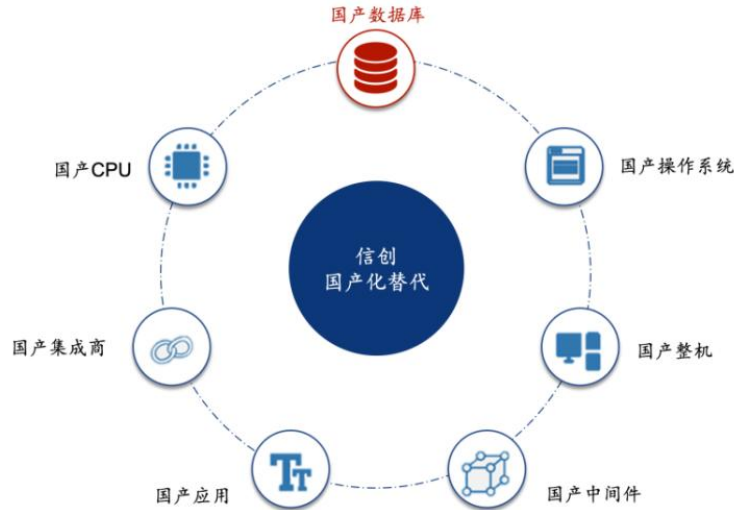
行业	机构	政策	发布时间
党政	发改委	《“十四五”推进国家政务信息化规划》	2022.1.6
	国务院	《“十四五”数字经济发展规划》	2022.1.12
	中国人民银行	《金融科技发展规划（2022-2025）》	2022.1.4
	银保监会	《银行业保险业数字化转型的指导意见》	2022.1.26
金融	中国人民银行、监管总局、银保监会	《金融标准化“十四五”发展规划》	2022.2.9
		《中国银联金融信息技术应用创新产品能力评估指引》	2022.3.29
	中国人民银行	《金融稳定法》	2022.4.6
	证监会	《证券期货业网络安全管理办法（征求意见稿）》	2022.4.29
交通	交通运输部、科技部	《交通领域科技创新中长期发展规划纲要（2021-2035年）	
	交通运输部、科技部	《“十四五”交通领域科技创新规划》	2022.3.25
教育	教育部	《义务教育信息科技课程标准》	2022.4.11
医疗	安徽省财政厅、安徽省卫健委	《关于规范公立医疗机构政府采购进口产品有关事项通知》	2022.4.1

资料来源：智研咨询，申万宏源研究；注：表格内仅为2022H1部分相关政策

3. 国产替代驱动

国产数据库作为信创的关键环节，随国产化替代深入推进而受到高度重视，产品、技术均实现快速发展；2013 年，棱镜门事件爆发，信息安全的自主可控成为国家需求，国内企业积极响应国家需求，自主研发数据库系统。国产数据库作为国产化替代的重要环节，在我国信创产业政策的指引下实现加速发展。

图表25：国产数据库是国产化替代的关键环节



资料来源：《国产数据库的崛起》，赵伟（2020）、华泰研究

4. 资本驱动

2019 年开始资本大量进入中国数据库行业，其中 2021 年来更是井喷式爆发，融资次数达到了 20 多次，融资额度超过 30 亿元人民币。

近期中国数据库厂商融资盘点

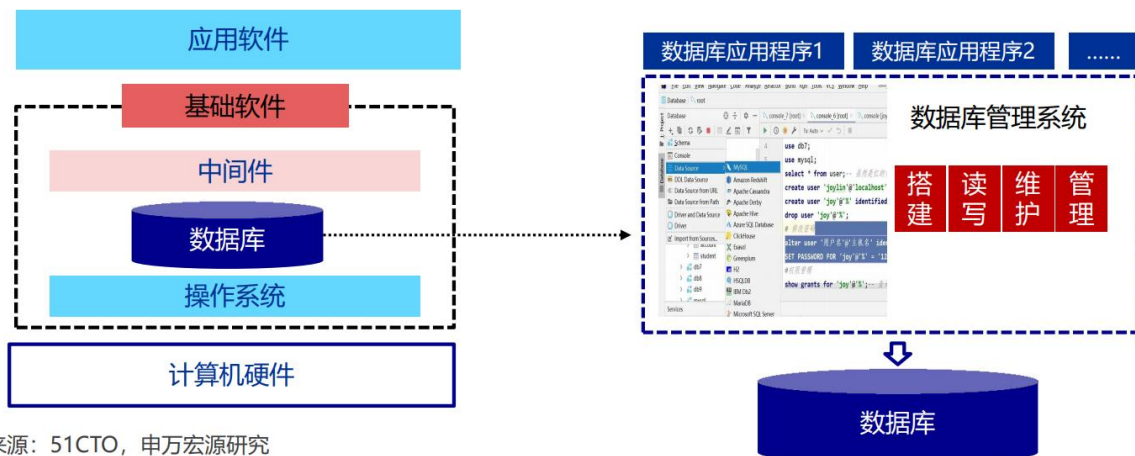
时间	公司名称	轮次	金额	投资方	数据库名称
2022/3/22	天云数据	D轮	数亿元人民币	融溢资本、绿地创投、京国创新辉投资	Hubble
2022/1/4	SphereEx	Pre-A	近千万美元	嘉御资本、红杉中国种子基金、初心资本、指数创投	ShardingSphere
2021/12/16	人大金仓	战略融资	近2亿人民币	太极股份	KingBase
2021/12/9	智史科技	B轮	1亿人民币	方广资本、凯泰资本、亿联凯泰基金、朗玛峰创投	DolphinDB
2021/12/6	创邻科技	A++轮	过亿人民币	同创伟业、达晨财智	Galaxybase
2021/11/29	四维纵横	A轮	1亿元	东方富海、某头部云厂商	MatrixDB
2021/11/19	拜贝思云计算科技	天使轮	300万美元	经纬创投、黄东旭	Bytebase
2021/10/29	矩阵起源	战略融资	数千万美元	钟鼎资本、五源资本、险峰K2VC、基石资本	MatrixOne
2021/9/28	聚云位智	B轮	近亿人民币	达晨财智(领投)、朗玛峰创投、中翔资本	LinkoopDB
2021/8/25	偶数科技	B+轮	2亿人民币	腾讯投资(领投)、红杉资本中国、红点中国...	OushuDB
2021/7/20	PingCAP	E轮	数亿美元	红杉资本中国(领投) GIC新加坡政府投资公司、BAI资本...	TiDB
2021/7/12	睿帆科技	A轮	5000万人民币	东方通(领投)、沅扬资本	Snowball
2021/7/8	人大金仓	战略融资	近亿人民币	太极股份、电科研发、南威软件、东华软件	KingBase
2021/6/22	中科知道	天使轮	1200万人民币	泰岳梧桐资本	PandaDB
2021/6/8	矩阵起源	天使轮	千万级美元	五源资本、险峰K2VC、源来资本、微光创投	MatrixOne
2021/5/24	涛思数据	B轮	4700万美元	经纬中国(领投)、GGV纪源资本、红杉资本中国...	TDengine
2021/5/9	易鲸捷	战略投资	5769万人民币	中国软件	EsgynDB

二、行业概览

1. 定义

数据库是信息系统运行的关键基础。从定义来看，数据库是按照一定的数据结构组织、存储和管理数据的仓库，在计算机中一般由一个或者一组文件构成；从本质来看，计算机解决的是数据计算和数据处理问题，数据库则是计算机应用系统中的专门管理数据资源的系统；从架构上看，数据库作为计算机三大基础软件（操作系统、数据库、中间件）之一，向下可充分发挥硬件算力，向上支撑上层的应用需求，是信息系统高效运行的关键基础。

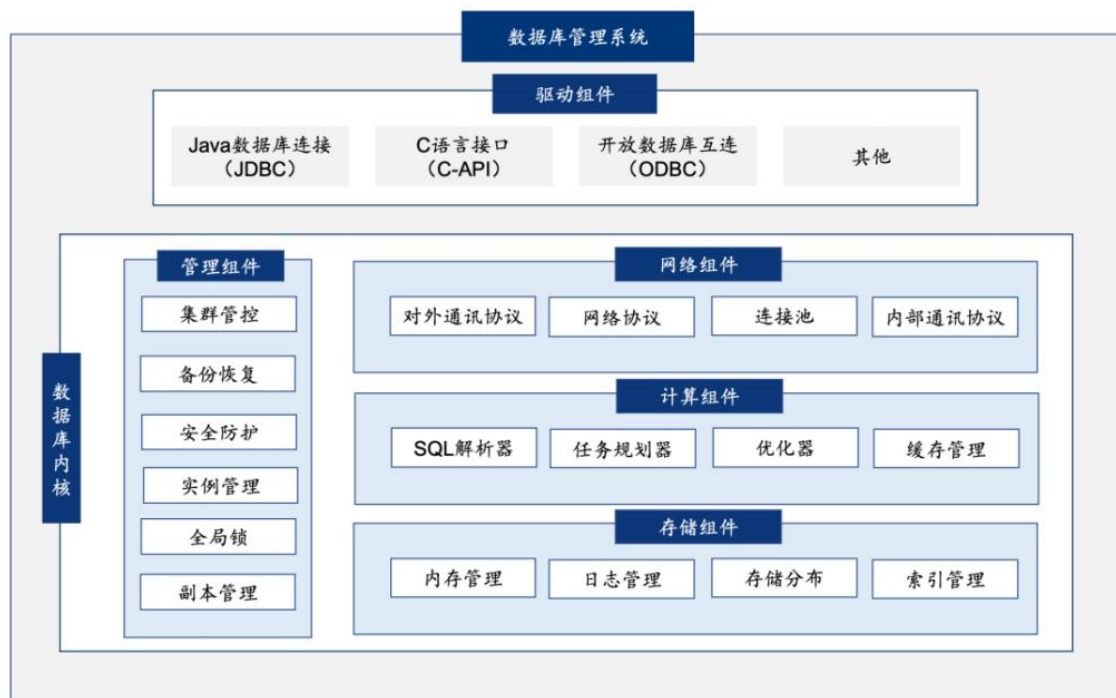
数据库是三大基础软件之一



资料来源：51CTO，申万宏源研究

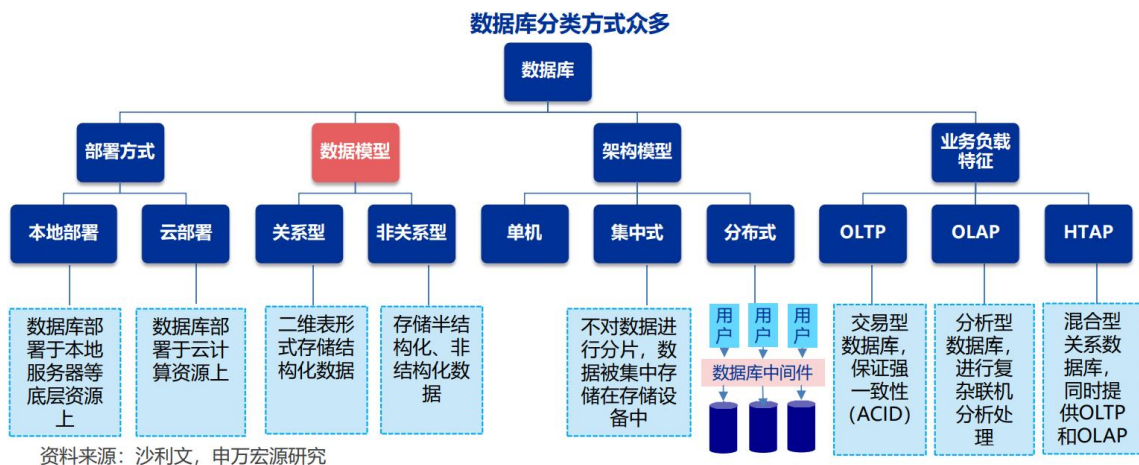
数据库软件的核心是数据库管理系统。数据库作为基础软件，可面向多种应用，被多个用户、程序共享，其中数据库管理系统（Data Base Management System，DBMS）负责搭建、处理、维护数据库的数据及数据间逻辑关系，由外部组件集与内核组件集共同组成：1）外部组件集：以数据库配套的独立支撑软件为主，例如数据库驱动；2）内核组件集：一般可以分为管理组件、网络组件、计算组件、存储组件四大模块。

图表2：数据库管理系统总体架构图



资料来源：《中国数据库发展研究报告》，中国信通院（2021）、华泰研究

2. 分类



根据数据模型、架构模型、业务负载特征的不同，数据库可划分为不同类别。

（1）根据数据模型类型：分为关系型和非关系型

数据模型决定了数据库的逻辑结构以及数据的存储，组织和操作方式，是最重要的要素。关系型数据库又称为 SQL 数据库，它建立在关系模型基础上，取代了层次模型和网络模型，以行和列的形式存储数据，以便于用户理解。常见的 Oracle、MySQL 等数据库均为关系型数据库。关系型数据库以二维表格形式存储结构化数据；ACID 特性保证数据库稳定可靠，关系型是传统行业业务应用的最佳匹配；关系型数据库市场仍占据市场的主流地位，2021 年占比约 80%。

非关系型数据库也称作 NoSQL 数据库，采用不同于关系数据“行列”组织的数据模型，数据结构类型复杂，是对传统关系型数据库的拓展与补充。搜索引擎数据库就属于非关系型数据库。

ACID和SQL是关系型数据库核心特征

	关系型	非关系型
数据存储	关系表	数据集（键值JSON文档/哈希表/其他）
模式结构	结构化，提前定义表结构	动态调整模式，非结构化
扩展方式	纵向扩展，提高处理能力	横向扩展，增加分布式节点
数据查询	标准通用的查询语言SQL	非标准非结构化的查询语言
关键特性	ACID	CAP、BASE
主要优势	结构化、事务处理、易于维护使用	扩展性、灵活调整、大数据分析
主要劣势	扩展性、高并发场景、大数据分析	事务支持较弱、标准不统一

（2）根据架构模型：可分为集中式数据库、分布式数据库

集中式数据库指将信息存储、维护在单个位置的数据库；分布式架构逐渐成熟，主要解决集中式架构扩展性差的问题。

图表8：集中式与分布式数据库对比

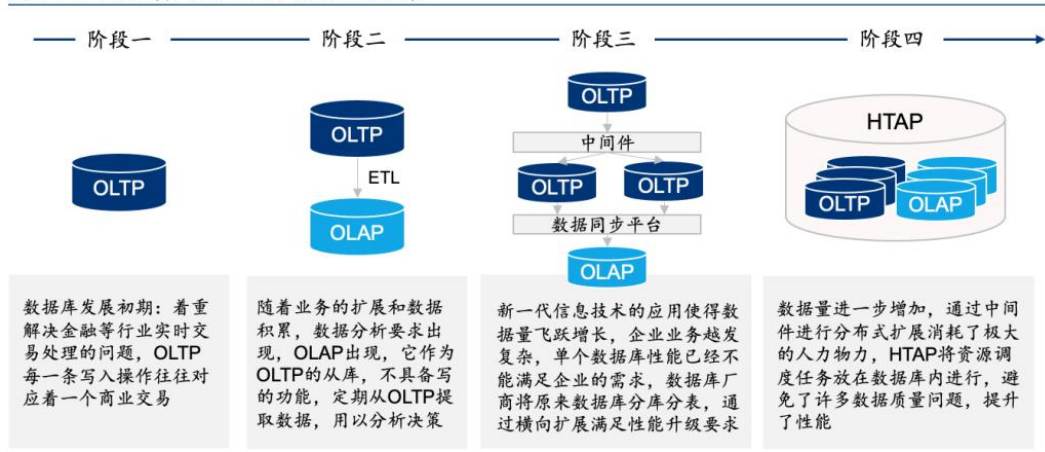
	集中式数据库	分布式数据库
架构	完全共享 (Shared-everything)	分库分表 共享存储 (Shared-storage) 无共享 (Shared-nothing)
硬件	一般依赖于特殊硬件	通用硬件
扩展性	较差	可大规模水平扩展
可用性	难以跨地域容灾	多地多中心
可维护性	简单集中管控	分布式管控
成熟度	高	较高

资料来源：《集中式数据库与分布式数据库》，墨天轮（2022）、华泰研究

（3）根据业务负载特性：可分为 OLAP 数据库、OLTP 数据库、HTAP 数据库。

OLAP 为分析型数据库，OLTP 系统为事务型数据库。HTAP（Hybrid Transaction and Analysis Processing）数据库为融合 OLTP 与 OLAP 业务的处理系统。

图表10：数据库产品朝 HTAP 方向融合发展



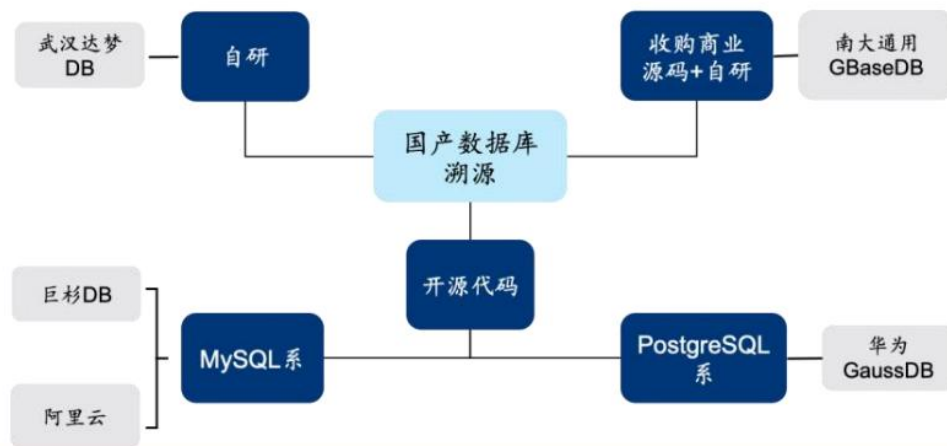
资料来源：《中国数据库行业研究报告》，艾瑞咨询（2021）、华泰研究

3. 技术溯源

根据源代码来源不同，国产数据库代码来源可分为三大类：

- 1) 基于开源代码研发：MySQL 系的巨杉数据库、阿里云数据库，PostgreSQL 系的华为 Gauss 数据库；
- 2) 收购商业源码+自研：南大通用收购 IBM 的 Informix 数据库源码；
- 3) 独立自研：以武汉达梦数据库为典型代表。

图表31：国产数据库技术溯源



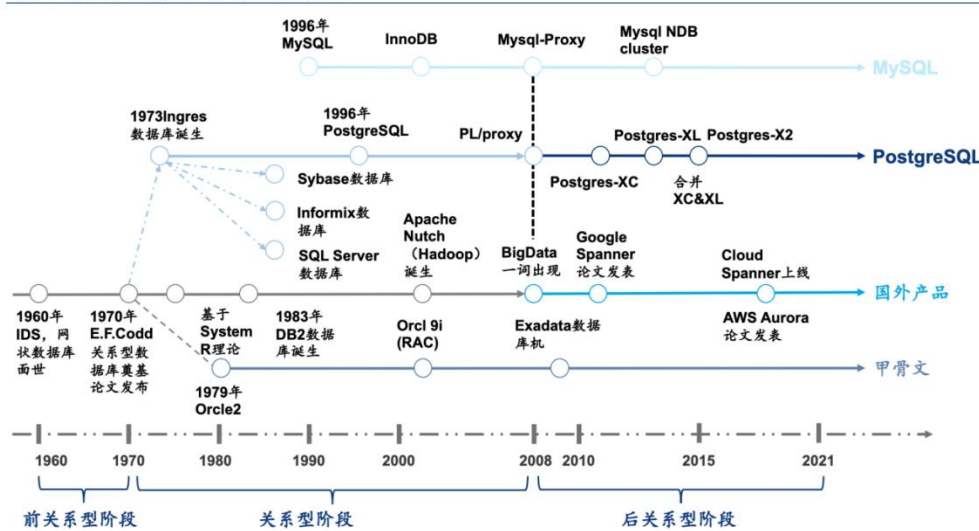
资料来源：人大金仓官网、华泰研究

4. 发展历程

(1) 国际

自 20 世纪 60 年代以来，数据库行业随信息技术发展而快速演变，数据库的发展历经前关系型、关系型和后关系型三大阶段，现阶段的技术路线更加多元化。

图表3：数据库产业发展进入后关系型数据库阶段



资料来源：《中国数据库发展研究报告》，中国信通院（2021）、华泰研究

（2）国内

海外垄断期（1978-2000年） SQL体系初建立，美国三巨头（Oracle, MySQL, SQLServer）相继初露峥嵘，并逐步统治全球市场。1977年国内召开第一次数据库年会，改革开放引领浪潮，积极引入国外技术，以三大巨头为代表的海外数据库于90年代席卷中国。

国产萌芽期（2000-2009年） “十五”计划期间，国家863计划设立“数据库重大专项”，国内研究所与大学积极投入数据库研究，IT社区逐步兴起；1999年，中国首个数据库“人大金仓 KingbaseES 数据库系统”诞生，随后武汉达梦数据库、神通数据库系统等国产数据库系统相继面世，打破了Oracle、IBM的市场垄断格局。

快速发展期（2009年至今） 2009年，阿里巴巴成立阿里云，开始研发自己的数据库产品AliSQL；随后，华为、腾讯等企业相继加入自主研发队伍，推出自有数据库产品；云计算时代与开源社区的兴起，国产数据库实现弯道超车，基于NoSQL等新技术的数据库产品加速面世。2019年，国外厂商神话被打破，蚂蚁集团的Ocean Base数据库成功登顶世界上最权威的数据库评测机构TPC（国际事务处理性能委员会）排行榜榜首。国产数据库逊于海外数据库已成为过往，国产自研数据库进入百花齐放阶段，近年来随着国产化替代深入推进，软件厂商、集成商、运营商纷纷入局数据库市场，我国数据库产业加速发展。

图3：国产数据库发展历程



数据来源：墨天轮，东方证券研究所绘制

三、商业模式

数据库产品包含开源数据库、商业数据库两种商业模式。

1. 开源数据库

开源数据库供用户免费使用，主要商业模式包括三种：

（1）完全开源式

借助基金会完全托管，以 HBase 为典型代表。

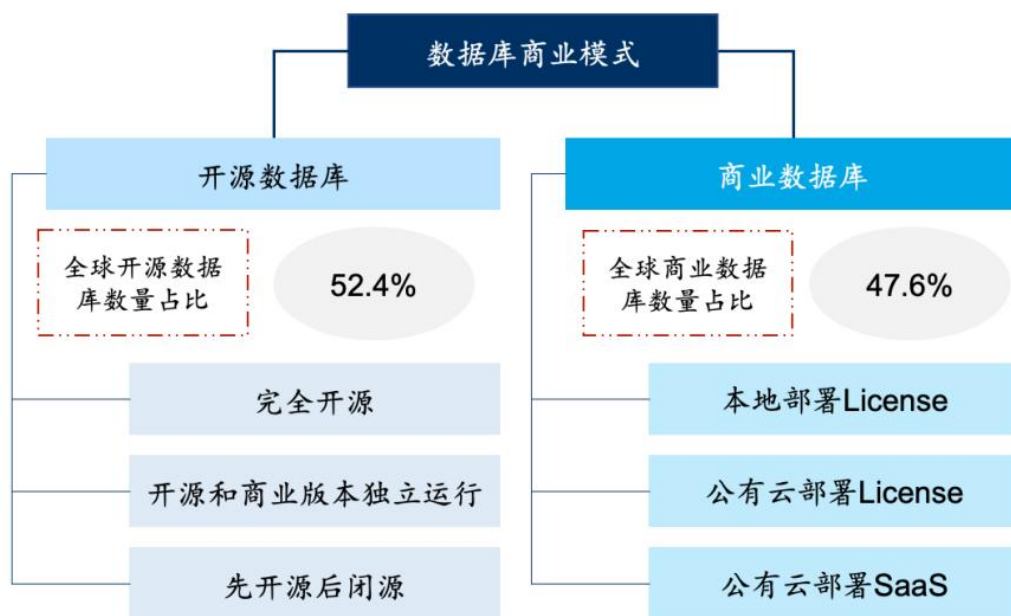
（2）开源版本和商业版本分别运营

通过运营开源版本社区积累人才、应用场景、市场品牌等，再通过售卖商业版本获取利润，以 TiDB 为典型代表。

（3）先开源后闭源

运作前期通过开源社区优化产品与品牌，之后停止社区的维护将产品闭源商业化，以 Neo4j 为典型代表。

图表22：数据库商业模式：开源数据库、商业数据库



资料来源：Oracle 官网、Snowflake 官网、《数据库发展研究报告》，中国信通院（2021）、华泰研究

2. 商业数据库

商业数据库收费方式主要有两种：

（1）本地部署

本地部署的商业数据库，一般采用 License 订阅方式收费，一般按用户数或者按 CPU 数定价，按年订阅，以 Oracle 为典型代表；

（2）公有云部署

公有云部署的商业数据库，除了 License 收费还可采用 SaaS 收费方式，用户可按用量付费，按年/季度/月收费，以 Snowflake 为典型代表。

四、竞争格局及发展现状

1.全球数据库产业呈现海外厂商主导的典型特征

相较海外厂商来说，国内数据库研究起步较晚，大约是 20 世纪 90 年代改革开放以后才开始进入萌芽阶段，该阶段数据库研究主要源自国家的相关研究计划或者大学科研需求，缺乏实际的业务场景驱动，因此数据库技术发展较为缓慢。从全球市场来看，以 Microsoft、Oracle 为代表的海外数据库厂商仍占据绝大部分市场份额，处于绝对领导地位。

图表20：2017-2021 年全球数据库市场份额排名变化

2017		2018		2019		2020		2021	
厂商	份额	厂商	份额	厂商	份额	厂商	份额	厂商	份额
Oracle	36.1%	Oracle	31.1%	Oracle	27.4%	Microsoft	24.3%	Microsoft	24.0%
Microsoft	21.5%	Microsoft	23.6%	Microsoft	24.7%	Oracle	23.8%	AWS	23.9%
IBM	12.7%	AWS	13.5%	AWS	17.1%	AWS	20.6%	Oracle	20.6%
AWS	9.2%	IBM	10.4%	IBM	8.8%	IBM	6.8%	Google	6.5%
SAP	7.4%	SAP	6.9%	SAP	6.5%	SAP	5.6%	IBM	5.6%

注：绿色表示排名上升、红色表示排名下降

资料来源：《全球数据库市场份额报告》，Gartner（2021）、华泰研究

2.国产厂商分类

国产数据库参与厂商包括传统厂商、初创厂商、云厂商和跨界厂商四类。

图表30：国产数据库产业图谱



资料来源：《中国数据库行业研究报告》，艾瑞咨询（2021）、华泰研究

- 1) 传统厂商：以达梦数据、人大金仓、南大通用等企业为代表的传统数据库厂商，是我国最早参与数据库研发、应用的企业，在数据库领域具有深厚的技术沉淀。
- 2) 初创厂商：以巨杉、Ping CAP、偶数科技、星环科技等企业为代表的初创厂商，依托于新兴的数据库技术与需求，近年来实现蓬勃发展。
- 3) 云厂商：以阿里巴巴、腾讯等企业为代表的云数据库厂商，得益于互联网业务和开源技术的快速发展，积极搭建自研云数据库平台，迅速占领云数据库市场份额。
- 4) 跨界厂商：以中兴、浪潮等企业为代表的跨界厂商，积极布局企业级数据库市场。

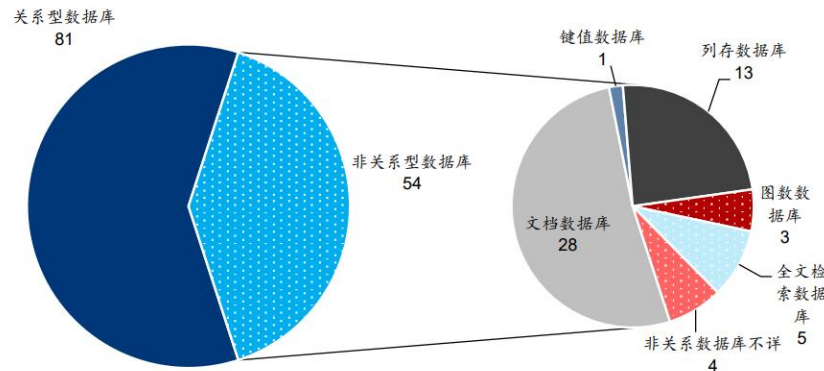
3. 我国数据库发展现状

经过多年技术研发和经验积累，国产数据库市场份额逐年提升。以人大金仓、南大通用、神舟通用为代表的国产数据库厂商自 1999 创建以来不断发力，逐步打破了海外厂商的垄断局面，同时数据库初创厂商、云厂商等也加速发力，快速替换海外数据库厂商的市场份额。

（1）关系型数据库占主流地位，市场竞争格局初显

关系型数据库占主流，多基于 MySQL 和 PostgreSQL 二次开发而来。关系数据库作为传统数据库产品，诞生时间早于非关系型数据库，产品体系更加成熟，在我国数据库市场中占据主流地位。

图表32：关系型数据库为国产数据库主流产品

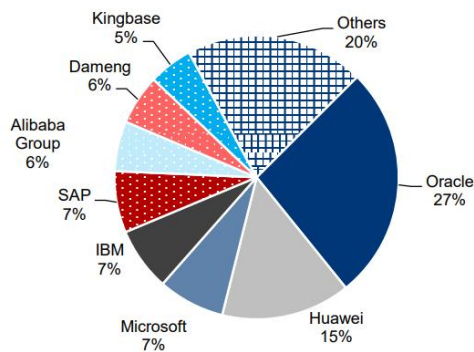


资料来源：《数据库发展研究报告》，中国信通院（2021）、华泰研究

（2）云数据库产品优势逐步显现，国产厂商掌握场景红利

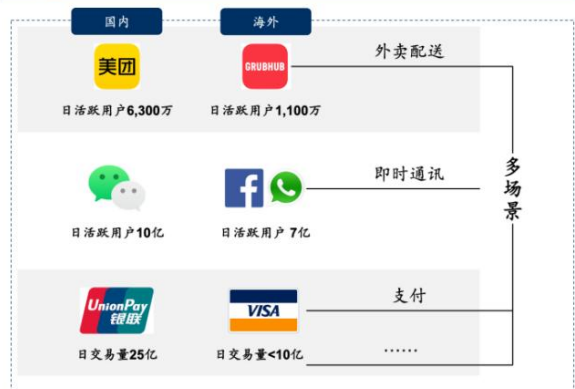
国产云数据库厂商的快速发展主要得益于我国的互联网场景红利，过去十年互联网及移动互联网在中国快速发展，国产厂商享受了高人口基数与丰富软件应用场景带来的场景红利，积极开拓面向海量、高并发数据的云数据库产品，并凭借产品优势与本土化业务理解迅速占领国内市场。

图表33：2021H2 中国公有云部署关系型数据库市场份额



资料来源：IDC、华泰研究

图表34：国产数据库厂商掌握场景红利



资料来源：《2021年中国分布式数据库市场报告》，沙利文（2021）、华泰研究

（3）初创厂商不断涌现，数据库产业体系日益完善

国产厂商多处在发展初期，具备较大成长空间。与老牌海外数据库厂商 50 余年的发展历史相比，多数国产数据库厂商的成立时长在 10 年以内，在人才储备与技术积累等方面，仍与海外厂商存在较大差距，大部分厂商仍然拥有较大的成长空间。

1）初创厂商多基于新兴数据库技术，关注非关系型数据库细分赛道

初创厂商纷纷涌入非关系型数据库细分赛道，有望带动基于新技术架构的非关系数据库快速发展，据艾瑞咨询预测，2020-2025 年以初创厂商为代表的新兴数据库市场规模有望实现高于 10 倍的增长。

2）国产数据库的支撑体系不断完善

数据库支撑体系包括学术组织、行业支撑组织、人才培养体系三大类。

图表38：国产数据库支撑体系日益完善

数据库学术会议		
数据库领域顶级会议 VLDB、SIGMOD、ICDE	高校 清华大学、中国科学院大学、 华中科技大学、浙江大学等	企业 阿里巴巴、华为、腾讯、百 度、武汉达梦、人大金仓等
数据库行业支撑组织		
数据库技术爱好者自发的用户组织 基于DBA的ACDU、面向Oracle用户的ACOUG、 面向MySQL用户的ACMUG、面向PostgreSQL 用户的中国开源软件推进联盟PostgreSQL分会		官方背景的研究组织 中国计算机学会（CCF）数据库专业委员会、通 信标准化协会大数据技术标准推进委员会 （CCSA TC601）
特定数据库产品讨论的官方技术社区 阿里云开发社区、华为云openGauss社区、 PinCAP AskTUG社区、PostgreSQL中文社区等		第三方技术社区 ITPUB、墨天轮、DBAplus
数据库人才培养体系		
教材	专业课程	教学服务
数据库厂商 PingCAP University、 Oceanbase学院、阿里云大学 AUCP、巨杉大学等	竞赛大赛	考试认证
专业培训机构 思墨学院、新炬学院、盘古云 课堂等	创新中心	生态人 才市场
高校 西北工业大学、清华大学、 武汉大学、中国人民大学等		

资料来源：《中国分布式数据库市场研究报告》，沙利文（2021）、华泰研究

3）学术研究与行业发展齐头并进，学术影响逐步提升

阿里巴巴、华为、腾讯、蚂蚁金服等企业和清华大学、香港科技大学、北京大学、香港中文大学、香港大学、浙江大学等高校论文纷纷入选三大顶会，显示我国数据库学术水平国际影响力不断扩大。

4）国内厂商重视人才培养，加快人才培养体系建立

数据库厂商积极搭建官方社区、开源社区、第三方数据库，与全国开发者积极进行技术讨论与交流，加快技术发展，如阿里云开发者社区、PostgreSQL中文社区、墨天轮等。

5）数据库开源实现人才积累，贡献长效技术动力

数据库开源对于实现人才积累、生态拓展、产品技术迭代具有重要意义，据 Github 预测，中国有望于 2030 年超越美国，成为全球最大的开发者来源。

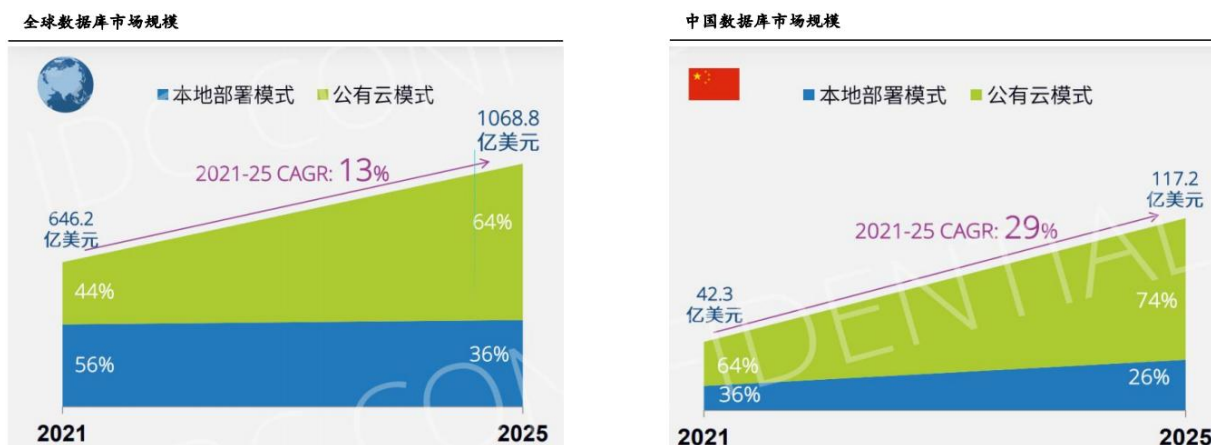
五、市场空间

1. 全球

2021 年，全球数据库管理软件市场规模为 646.2 亿美元，预计 2025 年将达到 1068.8 亿美元，2021-2025 年 CAGR 为 13%，其中，公有云模式的数据库占比将从 2021 年的 44%提升至 2025 年的 64%。

2. 中国

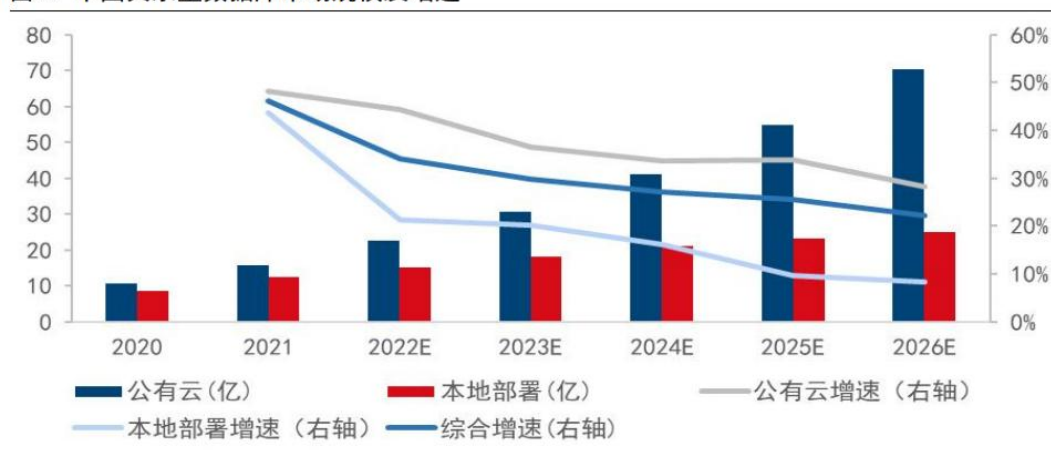
2021年，中国数据库管理软件市场规模为42.3亿美元，预计2025年将达到117.2亿美元，2021-2025年CAGR为29%，其中，公有云模式的数据库占比将从2021年的64%提升至2025年的74%。未来随我国数据库市场规模扩大与国产数据库替代加速，国产数据库有望同步扩大存量与增量市场，国产数据库市场空间前景广阔。



3. 关系型数据库及公有云数据库

目前全球范围内，数据库主要市场份额仍集中在微软、亚马逊、甲骨文、IBM、Google等几大巨头中，剩余公司数据库市场份额较小。从国内市场看，在本地部署关系型数据库方面，华为已超过部分海外巨头，阿里、达梦等国内厂商在市场份额上同样有所斩获。数据库领域国产化格局初步形成，同时具备走出去的长期潜力。根据IDC预测，2022-2026年中国关系型数据库综合市场规模年复合增速达到27.7%，其中本地部署数据库年复合增速达到15.0%，公有云数据库年复合增速达到35.2%。数字经济持续推进“用数上云”，国产公有云数据库仍有份额抬升潜力，值得持续关注。

图66：中国关系型数据库市场规模及增速



资料来源：IDC，国信证券经济研究所整理

六、公司梳理

1. 各自优势

（1）海外厂商向云化方向深度迈进，SaaS 模式助推新兴厂商加速发展

分布式云数据库成为海外厂商的发展共识。通过总结 Oracle、Microsoft、SAP、MongoDB 等海外传统数据库厂商以及 AWS、Snowflake、data bricks 等新兴云数据库厂商的发展路径、技术演进以及现有产品体系，可以看到分布式云数据库已成为海外厂商的发展共识，各类云托管、云原生数据库不断面世，加速数据库产品云化转型。传统数据库厂商基于原有业务优势，以云托管数据库为主要发展方向，支持数据库产品在公有云、私有云、混合云环境下部署；新兴厂商多选取云原生路线，Snowflake、data bricks 等厂商更是推出专门面向公有云环境的数据仓（分析型数据库）服务，SaaS 模式助力新兴厂商业绩高增。

（2）开源、分布式两大技术趋势下，国产数据库有望弯道超车

开源构建数据库产业正反馈循环，主流数据库产品陆续拥抱开源。数据库开源旨在通过开放核心代码，运营开源社区以实现多方共赢，1）对于社区开发者来说，开源数据库提供免费的数据产品，成本优势突出；2）对于数据库厂商来，通过开源社区可快速获得产品反馈，加快产品开发、提升产品质量；3）对于生态伙伴来说，开源产品的生态影响力更强，可以更好地实现产品兼容。通过梳理国产数据库厂商的发展历程，可以看到众多国产厂商最早都基于开源数据库构建底层代码，同时近年来主流数据库产品纷纷拥抱开源，以加速产品技术升级，构建友好的数据库生态。典型代表有巨杉数据库 Sequoia DB、华为 Gauss 数据库、阿里云 Polar DB for PostgreSQL 数据库、蚂蚁集团 Ocean Base 数据库。

分布式数据库时代加速到来，国产厂商加速分布式产品布局。分布式数据库可有效解决现有互联网环境下海量数据、高并发计算的数据需求，成为国产数据库厂商的产品发展共识。传统国产数据库厂商重点关注分布式关系型数据库，主要选择分库分表+中间件、原生分布式两类技术路线；新兴国产数据库厂商广泛布局分布式关系型数据库、NoSQL 数据库、New SQL 数据库等多类型分布式数据库，大量非关系型数据库厂商实现弯道超车。

2. 海外厂商

海外厂商：云数据库是主要方向，传统龙头与新兴厂商并驾齐驱。根据 DB-Engines 发布的数据库流行度最新排行（2022 年 10 月），全球数据库排名前五名依次为：Oracle、MySQL、Microsoft SQL Server、PostgreSQL、MongoDB，新兴数据库厂商 Snowflake、Data bricks 分别位列 13、20 名。

3. 国产厂商

根据墨天轮数据（截至 2022 年 11 月 17 日），共有 246 个数据库参与墨天轮数据库流行度排行，其中 TiDB、Ocean Base、达梦、open Gauss、Gauss DB 等数据库排名靠前。

图11：国内数据库市场规模及增速

排行	上月	半年前	名称	模型	属性	三方评测	生态	专利	论文	得分	上月	半年前
1	1	↑ 2	TiDB +	关系型	混合云原生	高可用	生态丰富	15	23	580.67	-31.78	-6.22
2	2	↑ 3	OceanBase +	关系型	混合云原生	高可用	生态丰富	137	17	574.56	-9.56	+11.45
3	3	↑↑ 5	达梦 +	关系型	混合云原生	高可用	生态丰富	381	0	497.01	-69.11	+34.23
4	4	↓↓↓ 1	openGauss +	关系型	混合云原生	高可用	生态丰富	562	65	486.17	-47.16	-102.85
5	↑ 6	↓ 4	GaussDB +	关系型	混合云原生	高可用	生态丰富	562	65	402.19	-33.20	-91.98
6	↓ 5	6	PolarDB +	关系型	混合云原生	高可用	生态丰富	512	26	388.53	-47.77	-10.83
7	7	↑ 8	人大金仓 +	关系型	混合云原生	高可用	生态丰富	232	0	387.93	-43.44	+52.21
8	↑ 9	↓ 7	GBase +	关系型	混合云原生	高可用	生态丰富	152	0	259.13	-16.04	-125.77
9	↓ 8	9	TDSQL +	关系型	混合云原生	高可用	生态丰富	39	10	251.47	-28.23	-17.68
10	10	10	AnalyticDB +	关系型	混合云原生	高可用	生态丰富	480	28	172.02	-20.50	-22.03

资料来源：墨天轮，民生证券研究院（截至 2022 年 11 月 17 日）

（1）达梦数据：背靠中国电子，主攻混合型数据库 HTAP

背靠 CEC 中国电子，主攻混合型数据库 HTAP。武汉达梦成立于 2000 年，是中国电子信息产业集团（CEC）旗下基础软件企业，专注于为客户提供以数据库软件、集群软件、云计算、大数据平台为代表的全栈数据产品及解决方案，产品涵盖数据库、云计算、大数据三大类。在数据库领域，武汉达梦主攻混合型数据库 HTAP，旨在用一种数据库模式处理客户所有数据库需求，适合业务广、数据量大的综合型客户使用，目前已掌握数据管理与数据分析领域的核心前沿技术，拥有全部源代码，具有完全自主知识产权。达梦数据库产品目前已广泛应用于金融、电力、航空、通信、电子政务等 30 多个行业领域。达梦主推透明分布式数据库（DMTDD）技术。新一代分布式数据库 DMDPC 具备极致的横向扩展能力。新一代分布式数据库 DMDPC 满足金融、电网、交通、工业互联网等多场景需求。

（2）人大金仓：背靠 CETC 中国电子科技集团，老牌数据库

背靠 CETC 中国电子科技集团，是成立最早的国产数据库企业。以数据存储计算为核心，提供完善的产品体系。三大数据库产品各具优势，使用多类型应用场景。人大金仓致力于为政务、能源、国防、金融、电信等国家企事业单位提供数据解决方案。

图表72：人大金仓主要应用场景及客户概览



资料来源：人大金仓官网，华泰研究

(3) 阿里 Ocean Base：金融级分布式关系数据库，核心代码开源

发端于阿里内部，2021 年正式开源。OceanBase 是由蚂蚁金服、阿里巴巴完全自主研发的分布式关系型数据库，始创于 2010 年，应用于支付宝全部核心业务以及阿里巴巴淘宝业务，从 2017 年开始，开始服务外部客户。OceanBase 是金融级分布式关系数据库，具备在线水平扩展能力。OceanBase 从 3.0 版本起核心代码正式开源，打造开源社区，实现生态共赢。OceanBase 客户主要集中在银行、保险证券、政企、互联网。

图表76: OceanBase 主要应用场景及客户概览

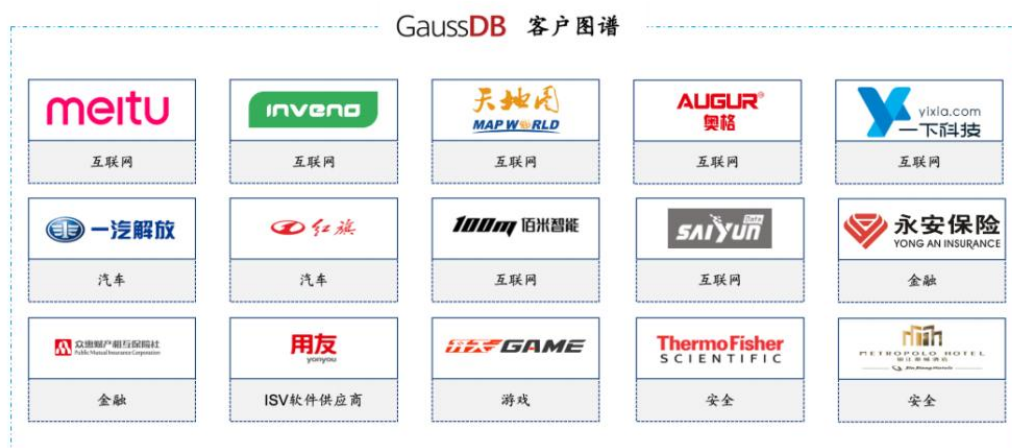


资料来源: OceanBase 官网、华泰研究

(4) 华为高斯: AI 原生数据库, 拥抱开源生态

AI 原生&支持异构计算。华为 Gauss DB 是一个企业级 AI-Native 分布式数据库, 可为超大规模数据管理提供高性价比的通用计算平台, 也用于支撑各类数据仓库系统、BI 系统和决策支持系统, 为上层应用的决策分析提供服务。Gauss DB 包含三大产品线: 关系型数据库、非关系数据库、数据库生态工具/中间件。openGauss 拥抱开源。GaussDB 适应互联网、金融、游戏、汽车等多业务场景。

图表80: GaussDB 主要客户及所属行业



资料来源: 华为官网、华泰研究

(5) 巨杉数据库: 自研金融级分布式数据库的独立厂商

专注分布式数据库技术研发的自研数据库独立厂商。从商用首发时间看，巨杉数据库是国内最早进行商业化布局的分布式数据库。

自研内核具有领先性，支持跨结构化、半结构化、非结构化的多模数据处理。湖仓一体平台为数据联通提供解决方案。巨杉主要为金融客户提供数据库产品，业务延伸至政府、能源、电信等多领域。助力数据库信创生态、支撑体系建设。

图表84：巨杉数据库客户概览

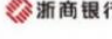
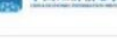


资料来源：巨杉数据库官网、华泰研究

(6) PingCAP:开源分布式关系型数据库

TiDB 为 PingCAP 自主研发的一栈式实时 HTAP 数据库。TiDB 适应多应用场景。“开源社区”助力“开源商业化”。PingCAP 拥有丰富的开源社区活动，形成开源社区和自身研发的有效协同。通过开源及免费策略，快速扩展开发者及技术粉丝用户群体，实现未来商业变现机会的长期积累。

图表86: TiDB 数据库国内用户概览

 Square	 Streak	 pony.ai	 UNEX	 SJ	 dailymotion	 Opera Ads	 ninjavan
 BIGO	 Shopee	 LMS	 bookshow	 U-NEXT	 colopl	 CAPCOM	 istyle
 SAKURA internet	 株式会社ネットスチーズ	 SHAREit	 agora	 中国平安	 中国银行	 中国农业发展银行	 中国人寿
 Bank 中国光大银行	 中国银联	 陆金所	 北京银行	 浦发银行	 WeBank	 招商银行	 广发银行
 浙商银行	 杭州银行	 亿联银行	 万达普惠	 CREDIT	 量化派	 广发证券	 安信证券
 国信证券	 LEXIN 乐信	 翼支付	 易宝支付	 贝壳金服	 马上消费	 拉卡拉	 360 数科
 PING++	 ziroom 自如	 中国移动	 中国电信	 China unicom	 中国南方电网	 人民网	 凤凰网
 吉林祥云	 中国经济信息社	 中移铁通	 ENN 新奥	 Haier	 Lenovo	 oppo	 XPENG
 理想	 SF EXPRESS	 中通快递	 G7 汇通天下	 快成	 J&T 极兔速递	 Tencent 腾讯	 美团
 58同城	 iQIYI 爱奇艺	 知乎	 丰巢	 Sogou 搜狗	 momo	 脉脉	 360

资料来源: TiDB 官网、华泰研究

(7) 星环科技:专注企业级大数据服务, 布局分布式数据库

重视技术研发, 核心研发团队曾供职于英特尔。市场地位领先, 大数据平台管理领域的佼佼者。布局分布式数据库, 搭建数据服务闭环。Argo DB 满足数据多模、联邦计算需求。与传统的 MPP 数据库相比, Argo DB 具有存算解耦结构, 存储和计算独立按需扩容, 具备更强的可扩展性、一致性、平滑适配云原生架构等优势。Kun DB 支持高并发交易场景的核心业务。星环科技数据库服务能源、金融、政府、交通等多业务场景。积累大量优质客户, 信创领域不断发力。

图表90：星环科技数据库主要应用场景及客户概览



资料来源：星环科技官网、华泰研究

(8) TDSQL：企业级分布式数据库

TDSQL 是腾讯云企业级分布式数据库。2020 年 12 月，腾讯云将原有的 TDSQL、T Base、CynosDB 三大产品线统一升级为“腾讯云企业级分布式数据库 TDSQL”。

(9) Polar DB：领先的云原生数据库

云原生关系型数据库 Polar DB 是阿里巴巴自主研发的下一代云原生关系型数据库。100%兼容 MySQL、PostgreSQL、高度兼容 Oracle 语法。计算能力最高可扩展至 1000 核以上，存储容量最高可达 100TB。经过阿里巴巴双十一活动的最佳实践，让用户既享受到开源的灵活性与价格的优惠，又享受到商业数据库的高性能和安全性。

(10) 云和恩墨：企业级关系型数据库 MogDB

MogDB 是云和恩墨基于 open Gauss 开源数据库进行完善增强的企业发行版。云和恩墨创立于 2011 年，产品和服务主要涵盖数据管理（数据库基础软件、数据库云化管理平台、数据技术服务）、承载（分布式存储、数据持续保护）、加工（应用开发质量管控、数据模型管控、数字化转型咨询）和应用（数据服务化管理平台、数据智能）等领域。

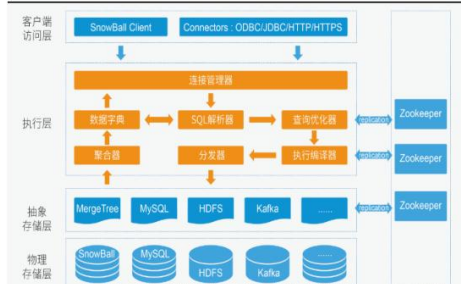
(11) 睿帆科技：分布式分析型数据库雪球 DB

睿帆科技成立于 2015 年，以大数据及人工智能为核心技术，专注于中国大数据服务市场。睿帆科技是中国移动最大的大数据处理供应商，为全国近一半的省份提供数据处理支持。分布式分析型数据库雪球 DB 是一款睿帆科技自主研发的基于 PB 级数据在线高并发极速即席查询的联机分析处理 (OLAP) MPP 列式数据库管理系统。

赛帆科技业务遍布全国10+个省市自治区



赛帆科技雪球数据库架构



赛帆科技雪球数据库产品特点



（12）万里开源：万里数据库拥有“万里前程”

冉冉升起的数据库“小巨人”：专注于国产、自主可控数据库及 Linux 操作系统产品研发，处于行业领先。掌控底层核心代码，已经广泛应用于多个行业。金融、电信运营商、能源等重要行业已实现应用。

（13）海量数据：高斯生态核心力量，核心客户验证自身实力

公司的国产化数据库产品主要包括 VastbaseG100、VastbaseE100 和配套智能工具。与华为进行多项合作，在高斯开源厂商中是除华为外贡献最大厂商，有望伴随巨头成长做大做强。

（14）南大通用：深耕行业多年，金融等领域获得大量核心客户

南大通用深耕数据库多年，为上万家企业提供服务。广泛展开对外合作，金融等领域有所作为。公司主要数据库产品包括 GBase8a 分析型数据库、基于共享存储的数据库集群 GBase8s、多模多态分布式数据库 GBase8c、极速内存数据库集群管理系统等。

（15）拓尔思：国内搜索型数据库龙头，国产替代下长期机遇可期

拓尔思作为中文全文检索技术的开创者和引领者，早在上世纪 90 年代就开始深耕搜索型数据库领域。目前公司的 TRS 海贝大数据管理系统（以下简称海贝）是拓尔思自主研发的一款搜索型数据库，适用于数字、文本、地理位置、结构化数据、非结构化数据等所有数据类型，为大数据应用提供高效的数据存储、全文检索、分析统计等数据管理服务。

七、发展趋势

1. 分布式数据库成为增量市场

（1）全球数据量激增，分布式数据库机遇显现

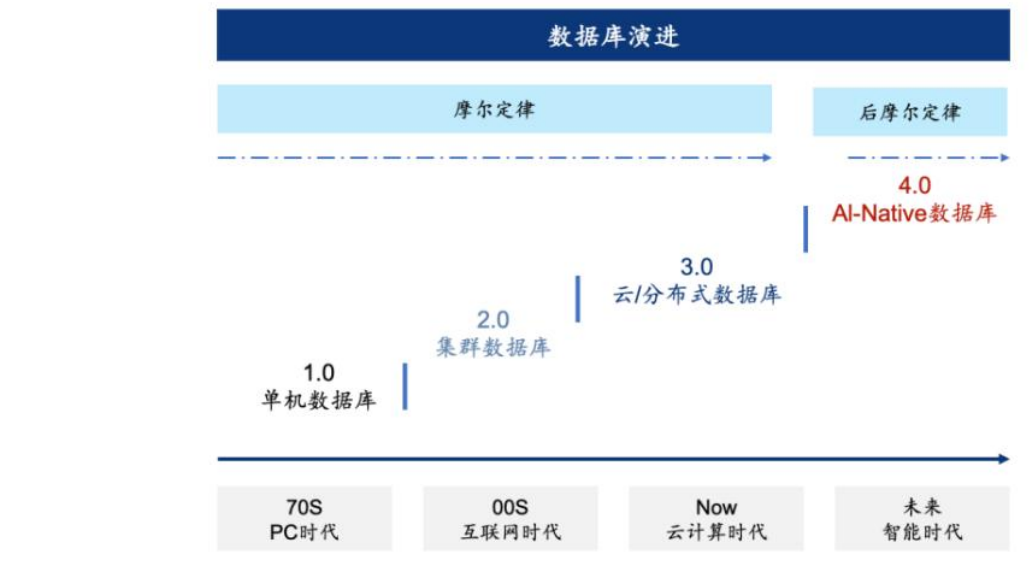
随着智能移动手机普及、云计算的兴起以及互联网的快速发展，全球数据量不断攀升，随着数据量上升，全球数据库规模快速增长，据 Gartner 预测，2024 年全球数据库市场规模有望突破千亿美元，2018-2024 年复合增长率为 13.78%，同时大数据分析、高并发计算、非结构化/异构数据处理的需求日益显现，分布式数据库兼具拓展性与高可用性，或将成为数据库行业新的增量市场。

（2）在非结构化数据与高并发计算需求的推动下，分布式数据时代加速到来

在计算、数据的双重变革下，我们认为，数据库行业的发展演进也将经历重要的技术变革。在传统计算环境和数据类型方面，传统的关系型数据库依然发挥着重要的作用；但面向未来新的计算场景与数据需

求，数据库产品亟需进行技术升级以适应下游需求的变化，通过统计国内外典型数据库厂商的技术路线及产品体系变化，我们判断，数据库行业已进入以云数据库、分布式数据库为代表的 3.0 阶段。

图表17：数据库发展进入以分布式数据库为代表的 3.0 时期



资料来源：《国产数据库的崛起》，赵伟（2020）、华泰研究

2. 新型数据处理技术应运而生

数据库技术近年来得到了迅速发展，新型硬件、新型交叉学科技术、新型数据模型、安全隐私等因素对数据库技术的发展起到了至关重要的作用，产生了很多的新型数据处理技术，主要包括基于新型硬件的数据管理、智能数据管理、多模数据管理、数据安全和隐私保护。

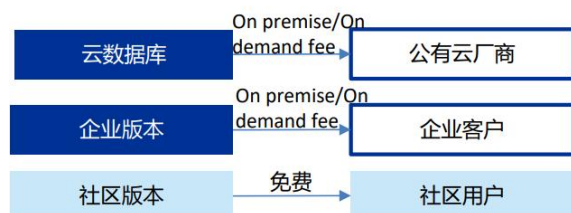
数据库行业未来发展趋势

新型硬件	多模数据管理	智能化	隐私计算
数据库在基础硬件和上层软件之间起到了“承上启下”的作用，向下发挥硬件算力，向上支撑上层应用。底层硬件技术决定了数据存取、并发处理等处理性能的物理极限，而上层软件系统也需要通过优化数据库架构和算法的设计，以提高软硬件契合度、最大化硬件利用效率。近年来，各类高性能处理器、非易失存储器、支持 RDMA 的高速网络等新硬件技术取得了重大突破，给数据管理系统带来新的机遇与挑战。	多模态数据统一管理即是指针对多种模态的数据进行统一管理。多模态数据主要包括结构化数据、半结构化数据和非结构化数据，例如关系、键值、图、XML/JSON 文档等。传统的数据库都是针对单个特定模态数据管理设计与实现的。目前很多领域都呈现多模态数据统一管理的重大需求，特别是在制造业领域。而针对制造业大数据管理的主要挑战在于多模态数据管理。目前，多模态数据统一管理已成为世界前沿科技热点。	近年来，随着 AI（人工智能）技术的发展，学术界和工业界开始探索利用人工智能方法解决数据库管理、优化、运维等问题，因此 AI 赋能的智能数据库系统应运而生。借鉴人工智能技术，可以实现包括数据库的自优化、自管理、自监控、自诊断、自恢复等在内的多维度等高度自治功能。从功能角度看，智能数据库系统也称为自治数据库系统；从技术角度看，智能数据库系统也可被称为 AI 赋能的数据库系统。	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，例如中国的《中华人民共和国数据安全法》与欧盟的《通用数据保护条例》(General Data Protection Regulation, GDPR) 等。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。涉及方向有：全密态数据处理、数据安全多方计算、防篡改数据处理、数据隐私计算四个研究方向。

3. 开源趋势推动非信创数据库发展

开源数据库流行度和数量超过商业数据库，正成为技术与市场变革的新引擎。开源数据库拥有更低的前期许可和持续维护成本，普及更迅速。通过运营开源社区快速获得反馈并加快产品开发、提升产品质量，同时反哺社区开发者及独立软件开发商（ISV）等生态伙伴，能够达到多方共赢。云、ICT 厂商及初创企业陆续开源旗下数据库产品。从 2015 年起，**PingCAP、Ocean Base、巨杉、涛思数据、Nebula Graph** 等创业公司第一时间推出了开源版本，随后**华为、阿里云、腾讯云、百度**等大厂也相继宣布开源，国内开源大军不断壮大。

开源商业模式主要向企业客户及云厂商收费



2021年全球开源数据库数量高于商业数据库



www.swsresearch.com 资料来源：DB-Engine，阿里云，华为云数据库，申万宏源研究

4. 搜索型数据库新兴力量可期，湖仓一体引领发展潮流

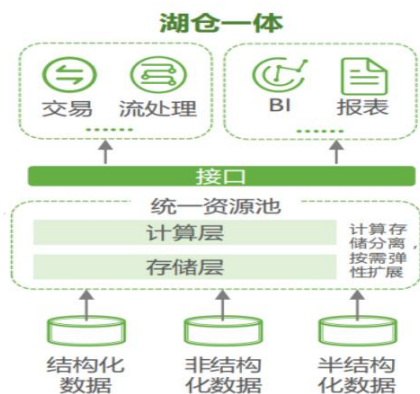
（1）搜索型数据库：自主化势在必行，国产厂商勇担重任

国产数据库浪潮已起，大多数组织都通过关系型数据库很好地管理与利用了结构化数据，但是很多业务的重要见解都隐藏在非结构化数据中。随着我国数字化转型的蓬勃发展，非结构化数据的处理需要依赖基于全文检索的搜索引擎技术，而目前市场上开放源代码的领先的全文搜索引擎主要是 **Solr** 和 **Elastic Search**（以下简称 **ES**）。国外危机频发，搜索型数据库自主势在必行。1）**ES** 数据泄露事件频发，给国内各行业用户敲响了数据安全的警钟。2）开源软件和安全工具受到国外出口法律限制。3）开源协议变更也带来商业风险。国产厂商奋起发力，**TRS 海贝** 大数据管理系统实现从底层到上层的全国产化。

（2）湖仓一体：打通数仓和数据湖，实现优势融合

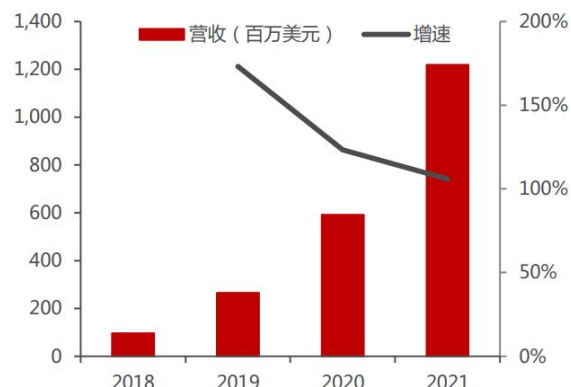
数据湖、数据仓库均有自身缺点。湖仓一体是一种新型的开放式架构，打通了数据仓库和数据湖，将数据仓库的高性能及管理能力和数据湖的灵活性融合了起来，底层支持多种数据类型并存，能实现数据间的相互共享，上层可以通过统一封装的接口进行访问，可同时支持实时查询和分析，为企业进行数据治理带来了更多的便利性。以 **Snowflake、Data bricks** 为代表的新一代“湖仓一体”数据库厂商快速崛起。

图9：湖仓一体架构示意图



资料来源：艾瑞咨询，民生证券研究院

图10：Snowflake 营收及变化情况



资料来源：wind，民生证券研究院

八、参考资料

1. 华泰证券-科技行业专题研究：信创正当时，国产数据库发展提速
2. 华西证券-计算机行业信创系列专题报告（五）：数据库，信创的关键环节
3. 东方证券-计算机行业深度报告：国产数据库百花齐放，搜索引擎数据库风口已至
4. 民生证券-计算机行业信创新时代系列深度报告数据库篇：谁是少年中国“甲骨文”
5. 国信证券-计算机行业大数据系列专题（2）：国产数据库百花齐放，崛起正当时
6. 申万宏源-信创数据库行业深度：替代能力被低估的核心软件
7. 中信证券-计算机行业“构筑中国科技基石”系列报告 23：数据库，关键三问深度解读
8. 国信证券-2023 年度策略：布局 A 股高质量投资

免责声明：以上内容仅供学习交流，不构成投资建议。