

证券研究报告
行业评级 强大于市
2022年10月



数据库行业发展报告

证券分析师：

钱劲宇 执业证书编号：S0210522050004

请务必阅读报告末页的重要声明

核心观点:

- 数据库应用的发展趋势：**混合事务分析处理 (HTAP)**

针对OLTP和OLAP单独的数据应用需求，如果由各自专有系统处理，一般来说性能将比混合通用系统性能高一到两个数量级，但未来整合的趋势更明显，客户需要可以满足混合需求的通用系统，HTAP基于一体化架构高度融合，在事务支持和数据时效上更有保证，同时减少专有系统数据迁移处理下的同步成本。

- 数据库架构的发展趋势：**分布式架构**

分布式数据库继承了传统单机数据库的核心特性，同时还拥有分布式系统的处理能力，具有可水平扩展、高性能、高可用、混合负载、更安全等优点。原生分布式架构数据库直接改变底层结构，因此性能不受底层数据库制约，也天然契合云原生理念。

- 数据库部署的发展趋势：**云原生部署**

云原生数据库结合云原生与分布式特点，最大限度实现资源池化、弹性变配、超高并发等能力，可实现随时随地的多前端访问，提供云服务的计算节点，并且能够灵活及时调动资源扩缩容，助力企业降本增效，国内云原生数据库虽起步晚，但发展迅速。

投资建议：国内数据库相关企业国内处于早期快速发展阶段，建议关注相关拟上市或筹备上市公司。

风险提示：技术发展不及预期、相关公司上市情况不及预期。

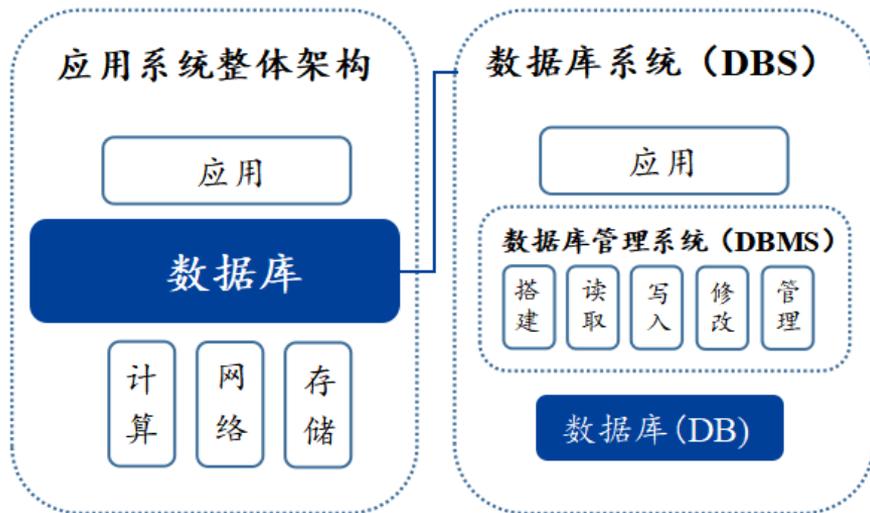
Part 1 数据库行业介绍及发展趋势 P04-P13

Part 2 国外数据库市场主要玩家 P14-P18

Part 3 国内数据库市场及竞争格局 P19-P32

数据库(DBS)是一个按数据结构来存储和管理数据的计算机软件系统。

数据库管理系统(DBMS)是数据库系统(DBS)的核心组成部分，主要完成对数据库的操作与管理功能，实现数据库对象的创建、数据库存储数据的查询、添加、修改与删除操作和数据库的用户管理、权限管理等



关系型
数据库

NoSQL
数据库

NewSQL
数据库

- **关系型数据库:** 存储格式可直观反映实体间的关系，和常见的表格类似。我国数据库产品数量分布仍以关系型数据库为主，占比60%左右。
- **NoSQL数据库:** 分布式、非关系型、不保证遵循ACID原则。随着数据量及数据结构复杂度提升，大量的NoSQL数据库出于简化结构、避免冗余目的被设计，支持海量数据的存储和高并发读写。
- **NewSQL数据库:** 现代关系型DBMS。分布式架构基因，兼具关系型数据库的ACID保证和NoSQL的强可扩展性。如 Google Spanner、Amazon Aurora。

ACID—为了保证数据库一致性，事务必须满足以下四个原则：
原子性(Atomic): 事务中任何操作的失败会导致整个事务的失败
一致性(Consistent): 事务结束后系统状态都是一致的
隔离性(Isolated): 并发执行的事务彼此操作互不影响
持久性(Durable): 事务完成后所做的改动都会被持久化

数据库常见有三种分类：

- 按模型可分为：关系型、非关系型（包括键值型、文档型、图、列簇型、时序、空间等）；
- 按架构可分为：单机、集中式、分布式；
- 按部署形态可分为：本地部署、云部署。

据中国信通院数据，截止2021年6月，我国数据库产品共有135款，其中关系型数据库81个，占比60%。

关系型数据库	按部署分类	按架构分类		
	本地部署 需要大量维护成本	集中式	单机	共享所有计算资源（CPU、RAM、Disk）和数据。不支持高并发场景，扩展难、数据迁移难、运维难。
云部署 云部署与分布式架构天然高度契合。	分布式	分布式中间件+单机	Shared-disk ：通过数据在逻辑层的线性切割实现性能和存储的扩展。能力受单机数据库自身能力的制约。	
		分布式存储	Shared-nothing ：非对称计算节点+共享存储(存储和计算未分开)，如Netezza/Teradata/Redshift均采用此架构。	
		原生分布式数据库	Multi-cluster, Shared data ：对称计算节点提供对等的读写服务(存储和计算彻底分离)，如Snowflake/TiDB均采用此架构。	

关系型数据库的不足：

- **高并发读写性能低**：关系型数据库的性能和访问承载能力，在面向单一数据节点的企业级应用时代无可挑剔。但随着数据量和节点急剧提升，已很难继续支撑巨大的规模系统，甚至成为应用系统的瓶颈；
- **不适合储存半结构和非结构类型数据**：二维表格数据模型不能有效地处理多维数据，不能有效处理互联网应用中半结构化和非结构化的海量数据，如Web页面、电子邮件、音频、视频等；
- **容量有限**：当一个应用系统的用户量和访问量与日俱增的时候，传统的关系型数据库却没有办法简单地通过添加更多的硬件和服务节点来扩展性能和负载能力。

NoSQL的出现并非为了取代关系型数据库，而是指Not Only SQL，提供了在SQL之外的另一种选择。

NoSQL 数据库	键值型数据库	支持高扩展、高并发，但数据无结构化，只被当做字符串或二进制数据，只能通过键来查询值。国外有Redis，国内有TcaplusDB,Tendis；
	宽列存储数据库	适用分布式数据存储与管理，以列族式存储，高扩展，查找简单快速，但不支持事务强一致性。国外有HBase,Cassandra，国内有GeminiDB,iBASE；
	文档数据库	适用存储面向文档或近似半结构化的数据，数据结构灵活，可以根据 value 构建索引，缺乏统一查询语法。国外有MongoDB、CouchDB，国内目前还没有；
	图形数据库	适用于社交网络，专注构建关系图谱，支持复杂图形算法，但复杂性高只支持一定数据规模。国外有Neo4j，国内有TGDB,Alibaba GDB,GeaGraph；

数据库诞生于20C60s，前后经历了网状数据库、层次数据库、关系型数据库、云数据库、云原生数据库。

- **网状数据库**：1964年，通用电气公司发布世界上第一个网状数据库系统IDS(Integrated Data Storage集成数据存储)，IDS是网状数据库，奠定了数据库发展的基础。
- **层次数据库**：1968年，IBM发布世界上第一个层次数据库系统IMS (Information Management System信息管理系统)，是第一个大型商用的数据库系统。
- **关系型数据库**：1978年，第一个关系型数据库Oracle 1.0 诞生，后续几十年中关系型数据库一直占据市场主导地位，直至移动互联网时代到来。
- **NoSQL数据库**：随着WEB 2.0网站兴起，NoSQL数据库应运而生，解决了大规模数据集合和多种数据类型的处理挑战。

New SQL既有No SQL相同的可扩展性，又保证了事务的ACID特性，既基于传统关系模型保留了SQL作为查询语言，又天生具有分布式架构基因（传统的SQL架构设计基因中是没有分布式的，而New SQL生于云时代天生就是分布式架构）。

- NewSQL 在SQL数据库的基础上集成了NoSQL强大的可扩展性；
- NewSQL 在NoSQL的发展中增强了对SQL、OLTP的支持；



数据库的两大应用场景：OLTP和OLAP

	联机事务处理/数据库 OLTP (On-Line Transaction Processing)	联机分析处理/数据仓库 OLAP (On-Line Analytical Processing)
应用场景	面向交易的处理系统 (业务系统)	分析驱动, 面向信息的分析过程
典型应用	银行交易系统	数据仓库
特征	业务在数据库联机的日常操作, 通常是对记录进行查询、修改, 用户关心快速响应、数据的安全性、完整性和并发支持的用户数等	一般针对某些主题的历史数据进行分析, 支持管理决策
数据量	每次交易涉及的数据量很小; 对响应时间要求非常高; 总体数据量相对较小。	每次查询涉及数据量很大(常涉及多表联结); 响应时间与具体查询有很大关系; 总体数据量相对较大。
数据	来源于系统中交易本身产生的数据	不产生数据, 数据来源于生产系统操作数据
使用人员	操作人员	管理人员
用户数量	极多	较少
交互载体	SQL为交互载体	以SQL为主要载体, 也支持其他语言交互
设计重点	尽量避免冗余, 为捕获数据而设计	有意引入冗余, 为分析数据而设计

2014年，Gartner提出 HTAP(Hybrid Transaction/Analytical Processing)混合事务分析处理的概念。

HTAP基于一体化架构，是高度融合的一个系统，在技术上更难，但对事务的支持和数据的时效上更有保证，同时减少数据在链条流动存在的多份数据冗余保存的负担，节省了技术、人力、时间和数据同步成本。

而如果用“OLTP+OLAP”模式，两套系统让迁移适配成本大幅上升，且为系统后续的开发和运维带来难处，同时也没法保证对事务的支持能力和数据的时效。OLAP本身不生产数据，需通过对OLTP操作数据的转化清洗迁移后，才可以进行专题需求分析。



集成数据平台 (HTAP) 节省了将OLTP的数据变化“以log日志的形式暴露、通过消息队列解耦传输、再被后端ETL拉取”，然后同步到OLAP的繁琐步骤。

针对OLTP和OLAP单独的数据应用需求，如果由各自的专有系统处理，一般来说性能将比混合通用系统性能高一到两个数量级，但就软硬件技术发展现状和当前需求来看，未来整合的趋势更为明显。除了极少数企业需要使用专有系统来实现其特殊的需求，集成数据平台(HTAP)将能满足绝大多数用户的场景。

从早期的单机关系型数据库到NoSQL再到如今的NewSQL，数据库领域不管是场景还是技术都发生了巨大的变化。在当下云原生时代，任何软件系统拥有分布式能力似乎成了标配。

- 分布式数据库继承了传统单机数据库的核心特性，同时还拥有分布式系统的处理能力。
- 分布式数据库可水平扩展、高性能、高可用、混合负载、更安全。

分布式数据库出现的原因：

性能可水平扩展：

数字化转型过程中，数据量、数据结构复杂度均呈几何级增长，传统数据库出现性能瓶颈。

分析能力出色：

分布式数据库的混合负载能力可大幅度提升分析时效性，减少数据冗余，灵活性大大提高。

政策战略扶持：

多项政策明确重点布局数据库分布式转型、应用创新战略，扶持国产数据库“换道超车”。

分布式数据库正以猛烈的趋势追赶，而传统数据库指数在慢慢下滑，且得益于国内互联网行业的快速发展，国外的数据库系统远不能满足需要，由此催生了一批走在世界前沿的国产分布式数据库。越来越多的国企、政府单位、银行证券这些传统行业也开始关注分布式数据库，去O道路正在加速。

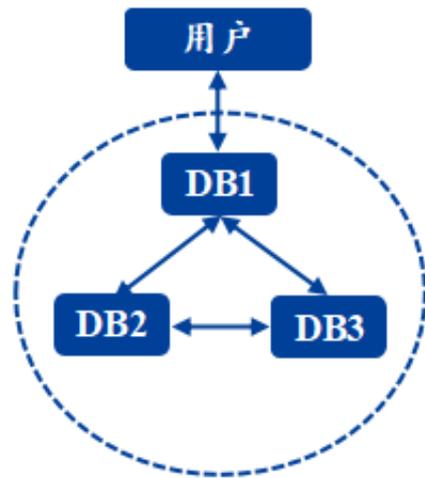
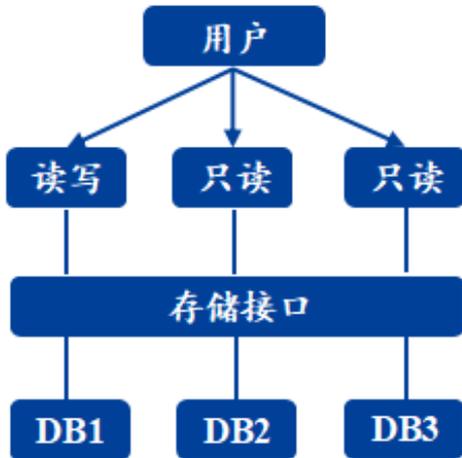
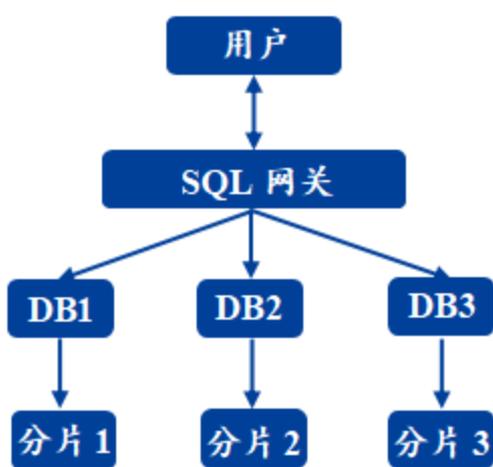
数据库架构：分布式架构的三条技术路线

1、 分布式中间件+单机数据库：通过数据在逻辑层的线性切割实现性能和存储容量的扩展，但能力受单机数据库自身制约。且移动扩展困难、数据迁移困难、运维难度高。

2、 分布式存储：大部分公有云数据库都属此类，以共享存储和非对称计算节点实现数据一致性和读写分流，有限解决了扩展性问题。如GaussDB(for MySQL)、TDSQL-C、SequoiaDB等。

3、 原生分布式数据库：各计算节点提供对等的读写服务，TiDB、OceanBase、ZnBase等。

原生分布式架构直接改变底层结构，因此性能不受底层数据库制约，也天然契合云原生理念。



据Gartner 2021年数据库产品提供商排名，Microsoft 凭借云数据库的后发优势，抢占了Oracle占据十年的榜首，前十名中有四家以云服务为主的企业，分别为 Microsoft/Amazon/Google/阿里云。Gartner预测到2023年，全球75%的数据库将完成云平台的迁徙，仅有 5%的数据保持在原本的本地部署。

据中国信通院，2020 年中国公有云数据库市场规模为 107.68 亿元，后续5年公有云数据库市场年复合增长率将达到36.1%，预计到2025年，中国公有云数据库市场总规模将达到 503.31 亿元。

- **传统数据库的云部署**：云与数据库的集合，减少了数据库参数的重复配置，具有快速部署、高扩展、高可用、可迁移、易运维和资源隔离等特点。
- **云原生数据库**：随时随地的多前端访问，提供云服务的计算节点，并且能够灵活及时调动资源扩缩容，助力企业降本增效。

传统数据库的云部署

基于云资源部署的传统数据库。

云原生数据库

基于容器化、微服务等设计的存算分离架构。

云原生数据库结合云原生与分布式特点，最大限度实现资源池化、弹性变配、超高并发等能力。

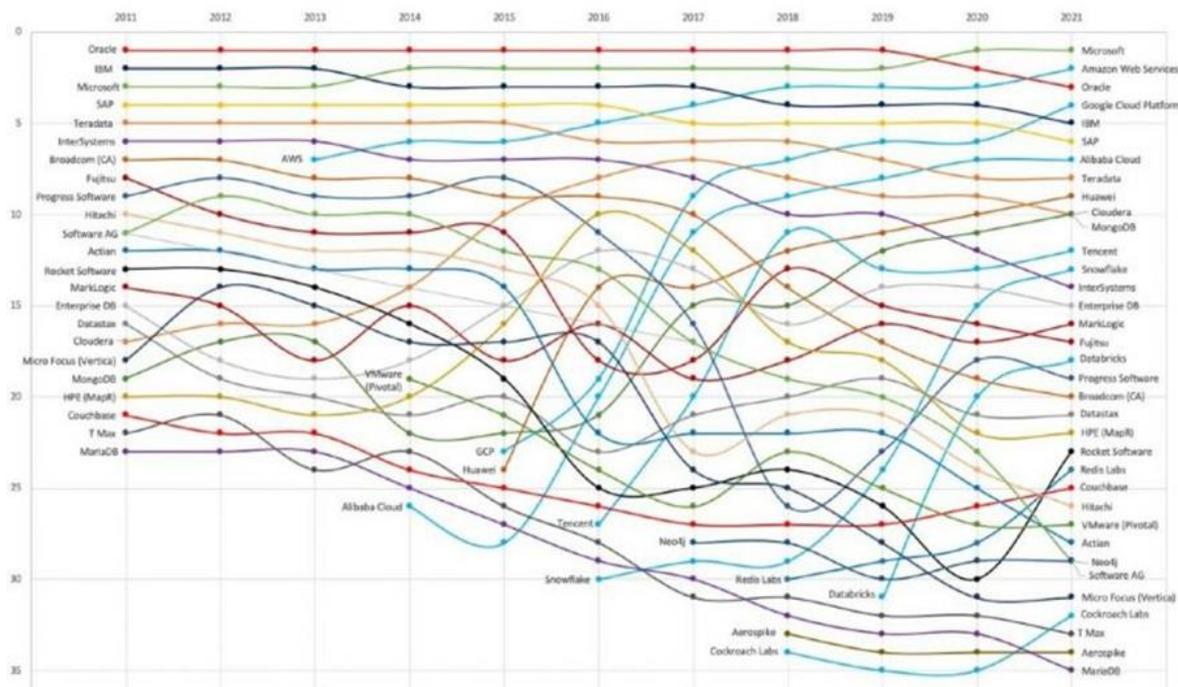
云原生数据库起源于Amazon，后受到国内厂商的广泛关注。据中国信通院对国内企业调研，受访者认为云原生数据库前四分别是AWS的Aurora、华为云GaussDB、阿里云PolarDB和腾讯云TDSQL-C，可以看出国内云原生数据库虽起步晚，但发展较快。

据 Gartner 在2011-2021十年期间的全球数据库市场份额变化，云厂商在数据库领域的竞争力不断提高，“云+数据库”已成为行业新的竞争力。

目前市场份额显著增长的厂商，都以云作为核心竞争力和创新方向。2021年Oracle 已经连续两年市场地位下降，跌至第三位，而 Microsoft、Amazon AWS 都是凭借数据库上云超越 Oracle，Google 也通过云数据库进入排名前。中国的三大云厂商阿里云、华为云、腾讯云同样获得数据库的高速增长。

在传统数据库厂商中，市场份额正在下降的厂商都是因为不具备云优势，例如 IBM、SAP、Teradata、Cloudera 等。

Gartner DBMS Market Share Ranks: 2011-2021



Microsoft: 2020-2021蝉联数据库市场份额榜首

Gartner 2021年数据库市场份额排名显示：微软2020-2021连续两年蝉联榜首；AWS排名上升一位超越Oracle位列第二；Oracle 降至全球第三。

DB-Engines 2022年10月排行榜显示：微软两款数据库受欢迎度前十一SQL Server(第三)和Access(第九)。

- 1、**Access**：1992年推出的Access与专业数据库比简单易用，但它又只支持有限的使用范围。从2013年起，Access就已不再推新并不断削减功能，如dBASE导入旧格式功能、数据迁移至SQL Server的升级向导等；
- 2、**SQL Server**：至今为止已发布30+个版本，最新的Microsoft SQL Server 2019具有三大特性：
 - **核心引擎**：SQL Server是HTAP混合负载能力的行业引领者，早在2017年的版本中就已实现了在单一引擎中对 OLTP 和 OLAP 工作负载的同时支撑；
 - **数据虚拟化**：为维护在企业整体架构中的核心地位，数据库必须支持多模型多范式，针对客观存在的异构数据源，在传统建立ETL通道外，数据虚拟化是更先进的理念。
 - **大数据集群**：创造性地将Hadoop和Spark等开源大数据技术组件直接纳入SQL Server，将SQL Server完全容器化并以Kubernetes为基础架构实现底层计算资源的编排和管理

数据库产品	特点	适用场景/客户
Microsoft SQL Server	整合封装：如大数据集群的一体化架构	本地部署的大型客户
Amazon Aurora	计算存储分离：每个数据组件单独的功能	云部署的中小企业客户

从云数据库到云原生数据库：近几年随着云基础设施的高速发展，同时具有按需扩展、按需付费优异特性的云数据库获得中小企业及互联网客户的青睐。然后云数据库本质不是针对云场景或云环境来设计或构建的，只是用到了云的资源，因此存在一些天然的问题，比如存储空间浪费、计算资源浪费、较大的恢复时间目标和数据滞后、系统性能受限、网络带宽消耗大等，从而限制了业务发展。云原生数据库应运而生。

- 2012年，亚马逊云科技推出首个云原生数据库DynamoDB，也是业界第一个真正意义上云原生数据库；
- 2014年，亚马逊云科技又推出云原生的关系型数据库Aurora，成为亚马逊历史上用户增速最快的云服务；
- 为了进一步简化客户在创建、维护和扩展数据库方面的工作，亚马逊云科技推出多种具有Serverless功能的数据库，让数据库的扩展性及自动伸缩容量达到新的高度，其中Aurora Serverless V2可以在几分之一秒内将数据库工作负载从数百个事务扩展到数十万个事务，与按照峰值负载配置容量的成本相比，最多可节省 90% 的数据库成本。

亚马逊云科技	SQL数据库	Aurora、Relational Database Service
	NoSQL数据库	DynamoDB、DocumentDB、Neptune
	Serverless数据库	Aurora serverless、DynamoDB、Timestream、Keyspaces、QLDB

Gartner发布的2021年全球数据库管理系统市场报告显示，亚马逊云科技在全球数据库市场的份额由2020年的20.6%增长到2021年的23.9%，跃居全球数据库细分市场第二位，2021年增长率高达42.3%。截至2022年9月底，亚马逊云科技已帮助超65万个数据库迁移服务至亚马逊云科技。

- 1970年，IBM研究员发表论文提出关系数据模型的概念，奠定了关系数据模型的理论基础。
- 1978年，Larry Ellison在为中情局做数据项目时发现了关系数据库的商机。同年Oracle 1.0诞生，Oracle成为第一家销售RDBMS软件的公司，在1982年一年内年收入就达到250万美元。

397 systems in ranking, October 2022

Rank			DBMS	Database Model	Score		
Oct 2022	Sep 2022	Oct 2021			Oct 2022	Sep 2022	Oct 2021
1.	1.	1.	Oracle +	Relational, Multi-model	1236.37	-1.88	-33.98
2.	2.	2.	MySQL +	Relational, Multi-model	1205.38	-7.09	-14.39
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model	924.68	-1.62	-45.93
4.	4.	4.	PostgreSQL +	Relational, Multi-model	622.72	+2.26	+35.75
5.	5.	5.	MongoDB +	Document, Multi-model	486.23	-3.40	-7.32
6.	6.	6.	Redis +	Key-value, Multi-model	183.38	+1.91	+12.03
7.	7.	↑ 8.	Elasticsearch	Search engine, Multi-model	151.07	-0.37	-7.19
8.	8.	↓ 7.	IBM Db2	Relational, Multi-model	149.66	-1.73	-16.30
9.	9.	↑ 11.	Microsoft Access	Relational	138.17	-1.87	+21.79
10.	10.	↓ 9.	SQLite +	Relational	137.80	-1.02	+8.43
11.	11.	↓ 10.	Cassandra +	Wide column	117.95	-1.17	-1.33
12.	12.	12.	MariaDB +	Relational, Multi-model	109.31	-0.85	+6.71
13.	13.	↑ 18.	Snowflake +	Relational	106.72	+3.22	+48.46
14.	14.	↓ 13.	Splunk	Search engine	94.66	+0.60	+4.04
15.	15.	↑ 16.	Amazon DynamoDB +	Multi-model	88.35	+0.93	+11.80
16.	16.	↓ 15.	Microsoft Azure SQL Database	Relational, Multi-model	84.96	+0.54	+5.24

22年10月最新DB-Engines的统计排名显示，最受欢迎前二仍是甲骨文Oracle和MySQL数据库，“去IOE”已经提出很多年，但Oracle数据库的地位在很多场景仍无可撼动。

排名主要依据该数据库在Google、Bing等搜索引擎上的数目，专业技术讨论中提到次数，专业网站中使用频率等影响力指标。

Gartner 2021年数据库市场份额排名显示：

- 微软：2020-2021年蝉联榜首
- AWS：上升一位，位列第二
- Oracle：降至全球第三

《2022年10月DB-Engines的全球数据库统计排名》

图片来源<https://db-engines.com/en/ranking>

在传统数据库领域，海外厂商起步较早，包括 Oracle、IBM、Microsoft 等，通过早期不断的技术积累和产品迭代主导了传统数据库市场，具有先发优势。IBM 在1973年就启动 SystemR 项目来研发关系型数据库，1983年发布商业版数据库 DB2。DB2代表了一组完整的数据管理系统，包括可在云环境中使用的服务器。早期 DB2 支持关系数据库模型，但现在很多产品已经扩展为支持关系和非关系结构。IBM 在2019年为 DB2 制定的目标是成为帮助增强认知应用程序的人工智能数据库，在数据库产品中融入 AI 功能，帮助用户跨本地和多云环境实现结构化数据和非结构化数据管理，并加速他们的 AI 应用程序开发。Gartner 在2021年云数据库管理系统魔力象限中再次将 IBM 提名为领导者。

IBM 作为传统巨头在数据库领域历经几十年而不衰，我们认为主要有以下四点原因：

- **起步较早**：海外数据库厂商起步早于国内厂商20-30年。IBM 在上世纪70年代已经开始研发数据库，先发优势带来了领先技术和客户粘性，因此海外巨头厂商至今仍然占据国内数据库市场较大份额。IBM 在过去几十年的产品开发和 Service 中建立起强大的品牌效应和客户声誉，在数据库领域意义重大。
- **迭代迅速**：巨头数据库厂商积极进行产品迭代，产生较大的用户粘性，帮助厂商维持市场份额优势。DB2 每年进行一或两次重大更新，并根据需要进行增量修复、性能优化，吸引已有用户持续使用这款产品。
- **技术领先**：产品性能是数据库最重要的指标之一，DB2 产品具有全球领先的技术。能够通过快速数据检索以及磁盘空间压缩使事务响应时间缩短30%，同时在高安全性、持续可用性、自动化管理和调优、可伸缩性等特点上优势出众。
- **生态完善**：作为传统数据库巨头，IBM 的 IT 布局非常完备，开发产品涉及多个领域且商品化程度高。注重技术创新和边界拓展，近年来在 AI、混合云、量子计算以及安全领域不断创新。

云原生时代下，大数据的爆发带来多元化需求，传统数据库厂商难以满足各类复杂场景下的业务需求，一大批独立厂商、初创厂商快速崛起。独立厂商 Snowflake 于2012年成立，专注于数据仓库的开发，主要应用于分析类场景。Snowflake 云数据平台基于公有云实现智能基础设施、产品弹性和优化的存储。核心产品数据仓库建立在云原生架构上，创造存储、计算以及管理服务分离的架构，摆脱了传统大厂数仓架构带来的局限性，能够满足客户对数据仓库并发性、易用性、扩展性的诉求，确保客户在各种分析场景下安全访问数据，同时支持在 Amazon AWS、Microsoft Azure、Google Cloud 三大公有云进行灵活部署，符合未来多云部署的发展趋势。



WHY SNOWFLAKE FOR DATA WAREHOUSING

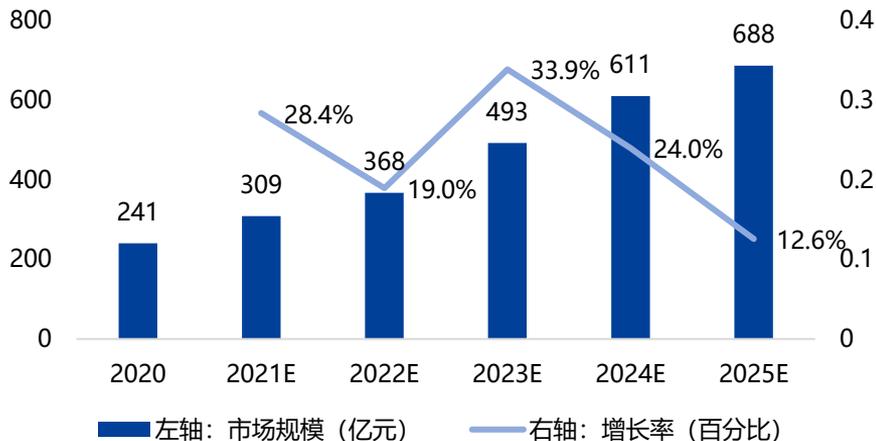
SECURE ACCESS TO DATA	ANALYTICS FOR ALL	SCALE WITH DEMAND
 Governed and immediately queryable source for data, including JSON and XML, with near-unlimited, low-cost storage.	 From analysts to data scientists to developers, meet the needs of different skill sets with tight integrations with the leading business intelligence and advanced analytics tools.	 Scale Snowflake up and down, on the fly or automatically, with per-second pricing and near-zero administration, across three major cloud providers.

中国信通院测算，2020 年全球数据库市场规模671亿美元，其中中国数据库市场占全球 5.2%，约35亿美元。

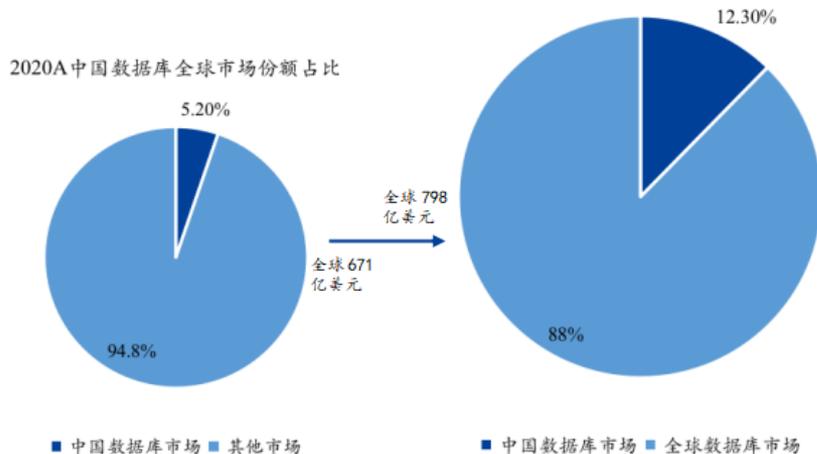
预计到 2025年，全球数据库市场规模将达到798亿美元。中国 IT 总支出在全球占比为12.3%。

预计2025年中国数据库市场在全球的占比接近中国IT支出在全球的占比，因此2025年中国数据库市场总规模将达到 688 亿元，年复合增长率（CAGR） 23.4%。

中国数据库市场规模及增速



2025E中国数据库全球市场份额占比



全球数据库玩家众多。海外来看，传统数据库Oracle、IBM DB2、Microsoft SQL Server等长期占据了to B领域的重要行业，拥有众多企业级用户。海外开源数据库MySQL、PostgreSQL等通过开源数据为用户提供了—条开放通路，走向更广阔的应用。国内来看，达梦、人大金仓等是老牌数据库厂商，PingCAP 是初创厂商的代表，凭借开源数据库 TiDB 在市场中占据—定份额。

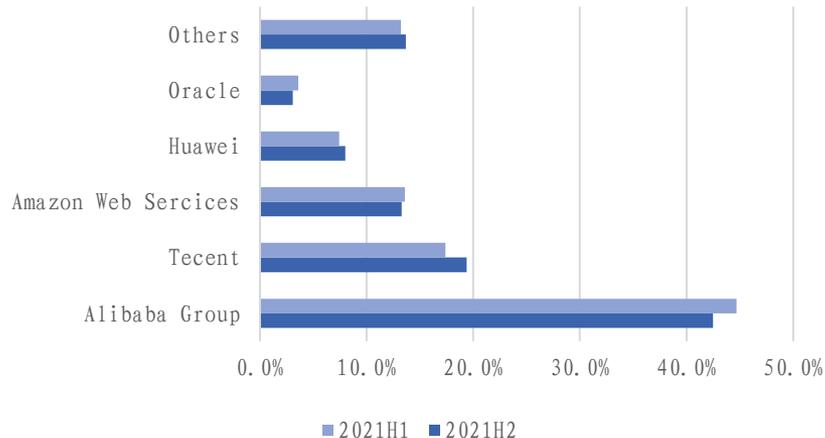
近年来，云计算的兴起也吸引全球众多云厂商成为数据库玩家。海外云原生数据库以AWS和Google Cloud为代表；国内阿里云的OceanBase和PolarDB，华为云的openGauss和GaussDB都乘云而上。



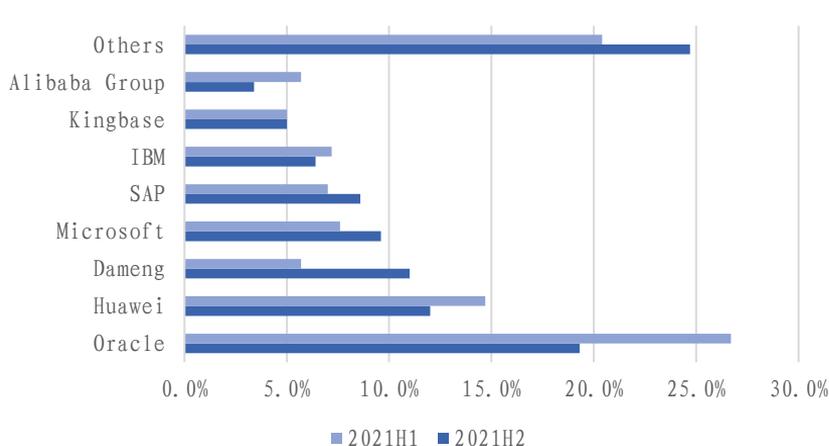
根据 IDC 数据，2021年下半年国内数据库厂商在关系型数据库市场的市占率继续扩大。其中公有云模式市场前三有国内三家厂商，两家海外厂商份额均下降，腾讯、华为市场份额都扩大。在本地部署模式市场，国内厂商正在快速追赶 Oracle、IBM 等海外厂商。达梦市场份额扩大接近一倍；人大金仓过去一年在政府行业获得大量订单；华为在政企、金融行业也获得突破。总体国内厂商的份额不断扩大。

国内，达梦、人大金仓等是老牌数据库厂商，PingCAP 是初创厂商代表，凭借开源数据库TiDB在市场中占据一定份额。国内阿里云OceanBase和PolarDB，华为云openGauss 和GaussDB则纷纷乘云而上。

中国关系型数据库软件市场份额——公有云模式



中国关系型数据库软件市场份额——本地部署模式

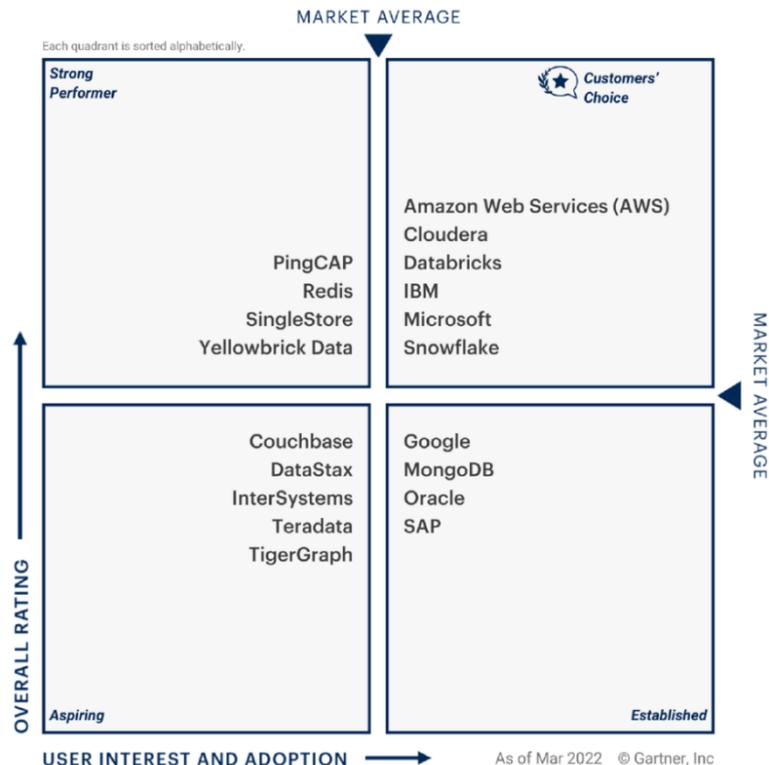


PingCAP: 不同于老牌厂商的定位

受到国产替代化的影响，PingCAP、Sequoia、易鲸捷等初创厂商纷纷涌入数据库市场，国产厂商阵营日渐强大。2022年，PingCAP 入选 Gartner 云数据库领域的 Peer Insights “Voice of the Customer”，成为唯一入选的中国企业级开源分布式数据库厂商，并在 Strong Performer 象限位列第一。公司旗下的 TiDB 数据库获得 94% 全球用户的高度认可。

TiDB 是开源分布式关系型数据库，同时支持在线事务处理与在线分析处理（HTAP），具备水平扩容或缩容、金融级高可用、实时 HTAP、云原生的分布式数据库、兼容 MySQL 协议和生态等重要特性，运维成本极低。TiDB 技术创新能力处于世界领先水平，每年会进行超过 50% 的代码更新，在 GitHub 总计获得超过32000颗星，目前已服务于全球超过3000家企业。TiDB 6.0的发布以及 TiDB Cloud 的商用化标志着 TiDB 数据库和云正在不断融合，加速释放数据价值。

Gartner Peer Insights “Voice of the Customer”
Cloud Database Management Systems



作为国内初创厂商的主要代表，PingCAP 的发展战略相比老牌玩家有四点优势：

- **全球化**：公司成立开始即定位全球化，通过全球化路线构建社区并使用英文创作代码，为打造国际化企业奠基。
- **开源数据库**：公司的竞争壁垒不在于代码，而是快速构建强大的生态。数据扩展性最重要的门槛是使用人数、迭代和计算速度。针对开源，TiDB 每年会进行超过 50% 的代码迭代。
- **互联网为市场切入点**：互联网是当前面临数据挑战和压力最大的行业，同时互联网技术也是目前最先进的技术，只有利用未来的技术才能抢占未来的市场。公司2015年成立后，充分借助了2015-2020年互联网行业高速发展的红利，利用互联网客户使用 TiDB 数据库的反馈不断积累场景经验、更迭产品技术。国内头部用户带来的场景经验成公司在海外市场的竞争优势。
- **选择 MySQL 协议兼容**：MySQL 在中国市场最为流行，对于业务或代码基于 MySQL 生态的用户，TiDB 数据库和 MySQL 协议兼容能够大大降低他们的迁移和学习成本。

6月15日，PingCAP 与阿里云达成合作，集双方技术优势打造的云数据库 TiDB 上线阿里，使TiDB成为全球横跨 AWS、Google Cloud、阿里云三朵云的数据库，这一合作代表了开源数据库和云服务不断融合的大趋势。PingCAP 的未来发展方向仍然以分布式数据库为主，将云原生技术融入数据库底层架构，助力更多企业数字化转型。同时，TiDB 数据库的定位基于 HTAP，降低用户的程序写作门槛，继续革新数据库的易用性，打造更开放、更通用、更大生态的 TiDB 数据库。

OceanBase 是全球唯一在 TPC-C 和 TPC-H 测试上都刷新了世界纪录的国产原生分布式数据库，实现 100% 自研，没有基于任何开源数据库模块，兼顾分布式架构的扩展性与集中式架构的性能优势，用一套引擎同时支持 OLTP 和 OLAP 的混合负载，具备数据强一致、高扩展、高可用、高性价比、高度兼容 Oracle/MySQL、稳定可靠等特征，不断用技术降低企业使用数据库的门槛。OceanBase 经历开源到闭源，2021年6月再次开源，以开放化的方式让软件本身和上下游技术以及应用生态对接，共建分布式数据库的未来。

同时，OceanBase 也实现了跨云架构统一。2022年8月，OceanBase Cloud 面向全球开服，以 OceanBase 数据库为核心，结合强大的运维管控平台和工具体系，以底层的阿里云、AWS 云基础设施为支持，为客户提供弹性扩展、卓越性能、主流兼容的高性价比的数据库云服务。

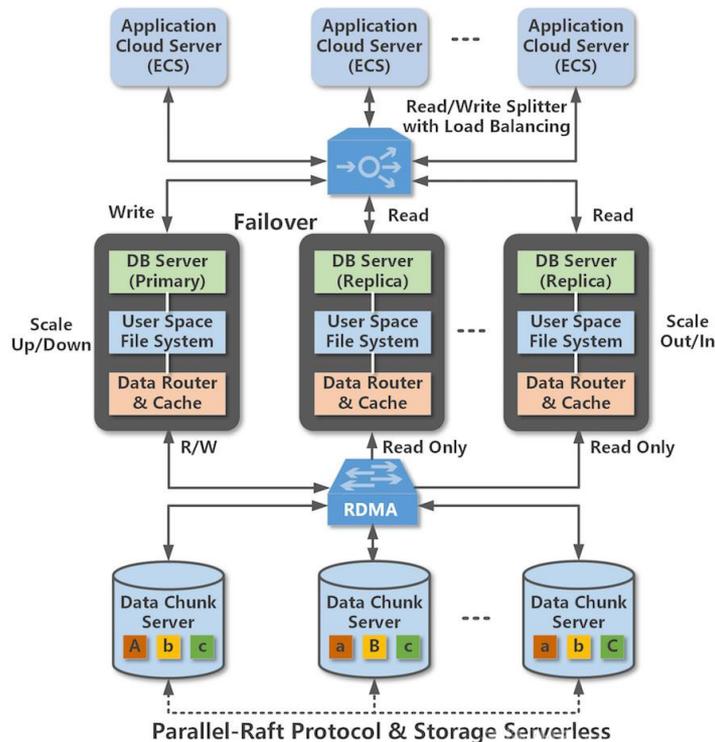
特点	原生分布式	高兼容	HTAP	稳定可靠	自主研发	高性价比
技术支持	自研一体化架构突破高性能和高可用，实现应用无限扩展和服务永远在线	Oracle/MySQL 平滑迁移快速、最小成本搬迁应用与数据	一份数据既能做事务处理又能实时分析，通过 HTAP 助力拓展更多可能	全量数据校验真正实现数据强一致，数据不丢失，“三地五中心”实现城市级容灾 RT0<30s	12 年完全自主研发，代码级可控，大规模金融核心场景 9 年可靠性验证	基于 LSM-Tree 的高压缩引擎平衡了“性能”和“压缩”的瓶颈，有效降低存储成本 70% - 90%，原生多租户，资源隔离按需使用

据 Gartner 预测，2023年全球75%的数据库将在云平台上，云发展趋势不可抵挡。2015年 PolarDB 只是阿里云的一个探索产品，而今天 PolarDB 成为阿里云数据库众多产品中的王牌，见证了阿里云在云原生数据库领域的自研创新突围。阿里云已经在2020-2021年连续两年进入 Gartner 云数据库的领导者象限。

传统数据库采用单节点架构，存储和计算绑定在一起。而分布式最大的特点是采用解耦合的共享存储，通过 RDMA 网络形成分布式共享存储，可以实现一写多读共享存储的模式。因此，PolarDB 采用存储与计算分离的设计理念，数据库计算节点和存储节点之间采用高速网络互联，并通过 RDMA 协议进行数据传输，I/O 性能不再成为瓶颈，让跨区域、跨共享区的复杂度降低。

在云数据库这一新赛道上，中国的云原生数据库在很多技术创新领域已经处于领先地位。目前 PolarDB 已经实现的计算、内存与存储资源的三层解耦、多主多写、HTAP、Serverless 等功能，皆为全球领先技术。未来，数据库领域的核心发展方向是云原生+分布式，中国云数据库必将更上一层楼。

阿里云PolarDB数据库架构图



据华为 GIV 报告显示，全球数据量到2025年将增长至180ZB，AI应用率也将达到80%。在未来数据量不断增长的背景下，华为面向全球发布人工智能原生（AI-Native）数据库 GaussDB，广泛应用于金融、运营商、政府、能源、医疗、制造、交通等多个行业，金融行业成为最主要的用户。对数据安全和可靠性要求最为严苛的金融行业，GaussDB 在业内领先的金融级分布式数据库可以满足其高可用的诉求。一方面，构建 openGauss 开源社区，打造开放生态；另一方面，首创的 Switch Turbo 技术在数据中心出现故障后实现秒级切换，该项指标是同类产品的1/500，实现 RTO≈0 的诉求。因此，GaussDB 既拥有商业数据库的稳定可靠性，又拥有开源数据库的灵活性和低成本。

相比其他数据库，GaussDB 也有独特的亮点：

- GaussDB 支持 X86、ARM、GPU、NPU 等异构架构，支持行存储和列存储，可增加节点实现存储、查询及加载性能的线性扩展。可以看出，GaussDB 未来的目标市场不只是国产服务器市场，可能还包括未来5G 带来更多的计算应用场景。
- GaussDB 是业界第一个人工智能原生态（AI-Native）数据库，首次将 AI 技术融入分布式数据库的全生命周期，实现了一定的自运维、自管理、自调优、故障自诊断和自愈能力。通过多元算力+DFV存储+RDMA 的技术支持，实现国内唯一软硬全栈与数据库调优，调优性能比业界提升60%以上。
- GaussDB 可以建立在华为高度耦合的 IT 架构底层生态上，包括网络设备、芯片、操作系统、云计算、人工智能等。

GaussDB 结合数据库云化改造技术，大幅优化传统数据库，针对不同的使用场景打造了更高可用、更高可靠、更高安全、更高性能、即开即用、便捷运维、弹性伸缩的数据库服务，目前在金融政企等主要行业已经拥有超过2500家标杆大客户。

分类	产品	使用场景	性能容量规格
关系型数据库	GaussDB(for openGauss)	应用于金融、电信、政府等行业关键核心系统，高性能场景	1.最大支持256分片1000+节点扩展能力；2.有效存储最大4PB
	GaussDB(for MySQL)	中大型企业生产系统（高性能，大数据容量），例如金融、互联网等	1.单节点最高可达150万QPS、50万TPS；2.支持并行查询，NDP算子下推，提升复杂查询效率，TPCH 标准SQL测试，性能提升3-20倍；3.单实例最高支持128TB存储
非关系型数据库	GaussDB(for Mongo)	应用于游戏（装备、道具）、泛互联网场景	1.支持最大12个节点集群；2.支持最大96TB数据存储；3.同等资源情况下性能可达到社区3倍
	GaussDB(for Cassandra)	泛互联网日志数据存储（并发写入量大，存储容量高）、工业互联网数据存储（写入规模大、存储容量大）	1.支持最大12个节点集群；2.支持最大96TB数据存储；3.同等资源情况下性能可达到社区3倍
	GaussDB(for Influx)	工业互联网时序数据、用户银行流水数据、物联网数据存储（时序）	1.同等资源情况下性能可达到社区5倍；2.压缩比为10:1；3.扩展比大于80%；4.最大支持12节点集群
	GaussDB(for Redis)	Key-Value存储模式，可用于互联网场景	1.最大支持12个节点集群；2.同等规格成本低于自建4倍；3.数据压缩比为自建4倍；4.p99时延达到社区2倍

达梦数据库是老牌厂商的主要代表之一，2000年成立至今，二十余年的发展几乎构筑了一部国产数据库的成长史。其主要产品线从通用关系型数据库产品向全栈数据产品延伸，目前各类数据库产品可覆盖数据全生命周期。二十多年自主研发技术的历练既是与国外巨头不断交锋的过程，也是让国内客户逐渐认可的过程。最新版本 DM8 是国内第一款实现数据共享存储集群的数据库，采用全新的体系架构，在保证大型通用的基础上融合分布式、弹性计算与云计算的优势，对灵活性、易用性、可靠性、高安全性等方面进行了大规模改进，支持超大规模并发事务处理和事务-分析混合型业务处理（兼顾 OLTP 和 OLAP），动态分配计算资源以降低成本。受益于多年自研技术积累和国产替代的热潮，目前达梦数据库的多款产品已经实现对标 Oracle，未来将继续赶超国外巨头厂商。

达梦
DM达梦数据库管理系统DM8
DMDSC磁盘共享集群
DMMPP分布式并行集群
DM MGBase非结构化数据库
DMTDD透明分布式数据库
DMETL数据抽取、转换、加载工具
DMHS异构数据同步工具
DM Designer数据仓库建模工具
DM OLAP联机分析服务
DM DataMiner数据挖掘工具

VS

Oracle
Oracle Database Software 20C
Oracle RAC
Oracle Exadata
Oracle NoSQL
Cloud
Oracle ODI
Oracle Golden Gate
Oracle Data Designer
Oracle Data Warehouse
Oracle Intelligence BI

达梦数据库产品线完备，通过基础数据库软件、面向不同应用场景的数据库集群软件以及云计算系列软件、大数据系列软件等产品形成以数据库管理系统为核心的产品生态体系，覆盖从数据交换、数据储存、数据治理到数据分析的全产业链，可为客户提供企业级的全栈数据解决方案和技术服务。

达梦数据库在国产数据库中市场占有率持续领先，凭借多年的市场积累，市场销售额保持快速增长，在国内处于领先地位。根据赛迪顾问发布的报告，达梦数据库在2019-2021年中国国产数据库管理软件市场占有率排名均位列第一。

2019-2021年中国国产数据库管理软件市场占有率排名

厂商	2021	2020	2019
达梦数据	1	1	1
人大金仓	2	2	5
优炫软件	4	3	4
通用数据	3	4	2
神州通用	5	5	3

注：赛迪顾问报告中，因公有云厂商（如阿里云、华为、腾讯等）数据库产品的经营模式差异较大，因此未做排名统计

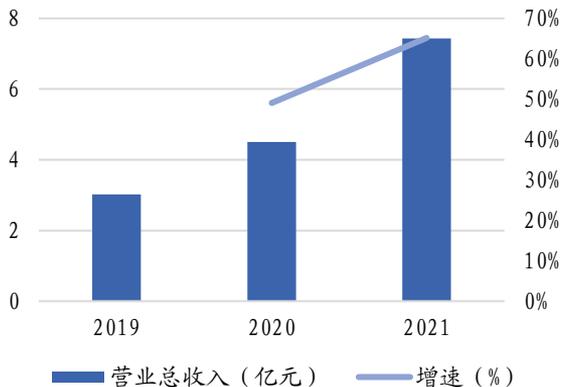
达梦数据库软件产品生态链



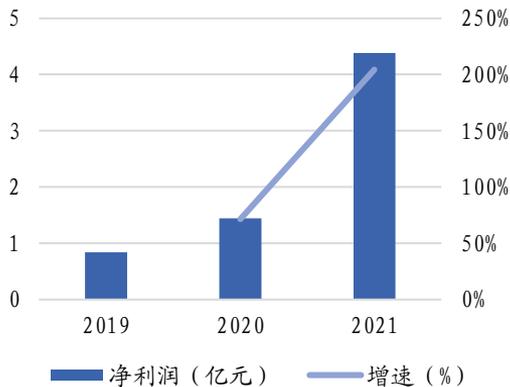
近年来达梦数据业绩高涨，2019-2021年营业收入分别为3.02/4.50/7.43亿元，增速持续增长，三年营业收入的年均复合增长率为56.95%。三年净利润分别为0.84/1.44/4.38亿，2021年增速高达204.17%，整体盈利水平较高。其中，公司主营业务毛利率三年分别高达80.32%/90.27%/90.60%，毛利率持续增长主要系2020年高毛利率的软件产品使用授权业务在公司主营业务结构中的比重有较大幅度提高。

To G市场的大中型政企用户是公司获得高增长的关键，以2021年为例，在达梦核心业务软件产品授权中，政府机关贡献收入最多，较去年增加1.4亿元，增幅达63%。另一方面，政策利好为公司提供大量政府补贴。2019-2021年，达梦每年获得的政策补助优惠(合并税收、政府补助)占收入比例达到17.6%/20.5%/32.75%。

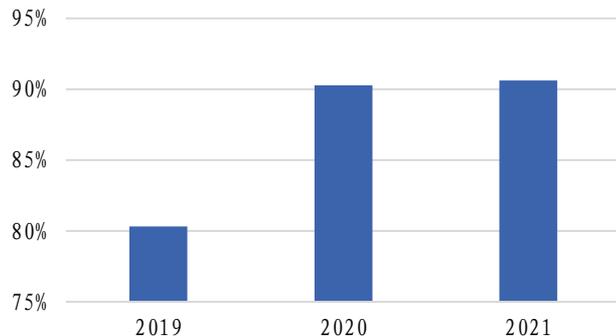
2019-2021年营业总收入及增速



2019-2021年净利润及增速



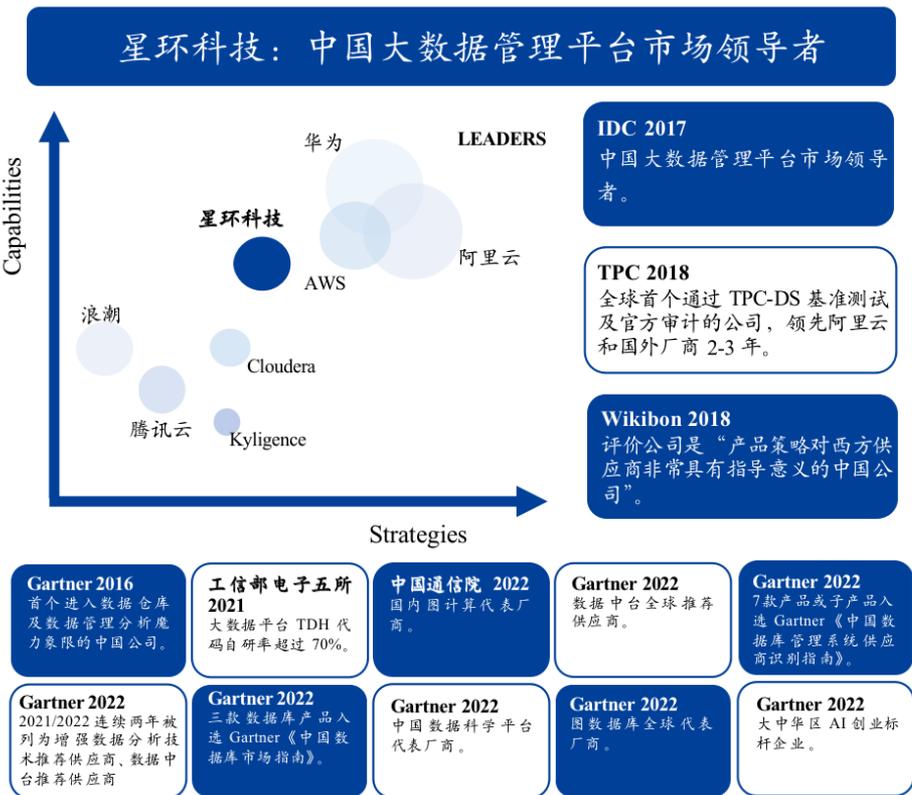
2019-2021年主营业务毛利率 (%)



星环科技是国内大数据基础软件的领导者，公司致力于打造企业级大数据基础软件，围绕数据的集成、存储、治理、建模、分析、挖掘和流通等数据全生命周期提供基础软件与服务，构建明日数据世界。公司以自主研发为核心竞争力，自2013年创建以来已建立了多个产品系列：一站式大数据基础平台TDH、分布式分析型数据库ArgoDB及交易型数据库KunDB、基于容器的智能数据云平台TDC、大数据开发工具TDS、智能分析工具Sophon和超融合大数据一体机TxData Appliance 等，产品线完备。

产品	使用场景	产品核心竞争力
星环大数据基础平台 Transwarp Data Hub	数据仓库、实时流处理等技术场景	采用创新的多模型技术架构，能够支持9种存储引擎、10种存储模型；统一的数据管理，统一的SQL语言；能够实现高性能、低成本。
分布式分析型数据库 Transwarp ArgoDB	在线分析处理、数仓数集、AETP、联邦计算等技术场景	支持标准SQL语法，提供多模分析、实时数据处理、存算解耦、混合负载、数据联邦、异构服务器混合部署等先进技术能力，可以满足数据仓库、实时数据仓库、数据集市、OLAP、AETP、联邦计算等各种需求。
分布式交易型数据库KunDB	高并发读写、交易分析混合处理、国产化替代、服务化SOA等技术场景；应用于金融交易、电信计费、电子政务、医疗信息系统等行业场景	支持标准SQL并兼容MySQL、Oracle方言及生态，支持Oracle PL/SQL愈发，并且能提供高可用和分布式两种部署模式，高并发事务处理场景性能卓越；支持国产CPU和服务器等自主可控等硬件平台，能够确保数据信息的安全。
星环大数据开发工具 Transwarp Data Studio	应用于大型企业中数据的资产化与业务化	大数据开发与治理一站式平台，TDS与TDH深度融合与优化，能够实现对其他数据库活数据平台的统一管理；采用分布式架构设计，可支持最大PB级别的数据平台建设。

星环科技凭借过硬的自主研发实力及产品多次获得全球著名咨询机构Gartner、IDC的肯定。



- 2016年星环科技成为中国首个进入Gartner数据仓库及数据管理解决方案魔力象限的厂商，且被评为最具前瞻性的远见者；
- 2021年起，星环科技蝉联Gartner增强分析技术中国推荐供应商；
- 2022年，星环科技入选Gartner数据中台领域全球推荐供应商。
- 2017年被IDC评为中国大数据市场领导者；
- 2020年再次被IDC评为中国大数据管理平台领导者；
- 2021年上半年，根据IDC报告，星环科技在中国大数据平台软件的市场份额排名第七；
- IDC发布《MarketScape：中国大数据管理平台厂商评估，2020》，公司综合能力排名市场第四。

投资建议：

国内数据库相关软件企业国内处于早期快速发展阶段。据中国信通院测算，2020年中国数据库市场占全球5.2%，市场规模约35亿美元。预计到2025年，中国数据库市场将占全球数据库市场12.3%，市场规模将达到688亿元，年复合增长率（CAGR）达23.4%。

国内目前的主营业务为数据库的相关公司（除头部互联网公司及华为等巨头企业），主要分为学院派和创业派，学院派始发于科研院校如武汉达梦和人大金仓等，创业派结合了资本和技术革命者，催生了一系列的新兴数据库企业如PingCAP和星环科技等。

考虑国产数据库行业仍处于早期高速发展阶段，主要的相关公司还处在拟上市或发力上市阶段，因此建议关注相关拟上市或筹备上市公司。

风险提示:

- 技术发展不及预期:

国产数据库(大数据)行业及相关产品仍处在早期高速发展阶段,并在分布式架构等新兴技术方向出现具有国际性竞争优势的产品和厂商,但如果技术优势不能保持,或将影响国产数据库在国内渗透率的提升。

- 相关公司上市情况不及预期:

国内主营数据库的相关厂商相处于初创阶段公司较多,均处于拟上市或筹备上市阶段,考虑或存在上市不利风险。

分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

一般声明

华福证券有限责任公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。本报告仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告的信息均来源于本公司认为可信的公开资料，该等公开资料的准确性及完整性由其发布者负责，本公司及其研究人员对该等信息不作任何保证。本报告中的资料、意见及预测仅反映本公司于发布本报告当日的判断，之后可能会随情况的变化而调整。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。本公司不保证本报告所含信息及资料保持在最新状态，对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

在任何情况下，本报告所载的信息或所做出的任何建议、意见及推测并不构成所述证券买卖的出价或询价，也不构成对所述金融产品、产品发行或管理人作出任何形式的保证。在任何情况下，本公司仅承诺以勤勉的职业态度，独立、客观地出具本报告以供投资者参考，但不就本报告中的任何内容对任何投资做出任何形式的承诺或担保。投资者应自行决策，自担投资风险。

本报告版权归“华福证券有限责任公司”所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。未经授权的转载，本公司不承担任何转载责任。

特别声明

投资者应注意，在法律许可的情况下，本公司及其本公司的关联机构可能会持有本报告中涉及的公司所发行的证券并进行交易，也可能为这些公司正在提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。投资者请勿将本报告视为投资或其他决定的唯一参考依据。

投资评级声明

类别	评级	评级说明
公司评级	买入	未来6个月内，个股相对市场基准指数涨幅在20%以上
	持有	未来6个月内，个股相对市场基准指数涨幅介于10%与20%之间
	中性	未来6个月内，个股相对市场基准指数涨幅介于-10%与10%之间
	回避	未来6个月内，个股相对市场基准指数涨幅介于-20%与-10%之间
	卖出	未来6个月内，个股相对市场基准指数涨幅在-20%以下
行业评级	强于大市	未来6个月内，行业整体回报高于市场基准指数5%以上
	跟随大市	未来6个月内，行业整体回报介于市场基准指数-5%与5%之间
	弱于大市	未来6个月内，行业整体回报低于市场基准指数-5%以下

备注：评级标准为报告发布日后的6~12个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。其中，A股市场以沪深300指数为基准；香港市场以恒生指数为基准；美股市场以标普500指数或纳斯达克综合指数为基准（另有说明的除外）。

诚信专业 发现价值

联系方式

华福证券研究所 上海

公司地址：上海市浦东新区滨江大道5129号陆家嘴滨江中心N1幢

邮编：200120

邮箱：hfyjs@hfzq.com.cn

