

| 证券研究报告 |

数据库专题系列一： 挑战与机遇并存，数据库行业渐成百家争鸣之势

2024.9.6

分析师：闻学臣
执业证书编号：S0740519090007

分析师：苏仪
执业证书编号：S0740520060001

研究助理：王雪晴
Email: wangxq03@zts.com.cn

- **从全球视角来看**，目前全球数据库产品数量整体分布以非关系型及混合型数据库为主；在流行程度方面，关系型数据库占据主导地位，根据DB-engines的数据显示，Oracle、MySQL、Microsoft SQL Server稳居前三，同时随着业务的多样化，多模数据库将成为新趋势；在市场份额方面，“云+数据库”战略成为技术发展新趋势，AWS以及微软占领市场前二。
- **从国内视角来看**，我国数据库产品数量呈现以关系型为主，非关系型数据库为辅的局面；在流行程度方面，PolarDB、OceanBase、openGauss热度持续攀升，排名前三；在市场份额方面，国内数据库市场仍被Oracle、Microsoft等国外厂商占据较多市场份额。
- **挑战孕育着机遇**。预计到2025年中国数据总量预计将跃居全球第一，占比有望达到27%以上，数据量的爆发式增长意味着将拉动数据库行业的需求空间。在本地部署模式下，国内关系型数据库市场格局中，海外厂商仍占据较多市场份额，在政策、新技术等因素的推动下，国内数据库厂商的市场规模及竞争力有望快速攀升。
- **投资建议**：数据库作为计算机系统的核心基础软件，同时也是信创产业的关键环节，我们认为在政策、新技术等因素的扶持下，国内数据库产业有望持续受益。当前时点，我们持续看好国内数据库产业的投资机遇，建议关注达梦数据、星环科技、太极股份（人大金仓）。
- **风险提示**：技术发展不及预期；行业竞争加剧风险；政策落地不及预期；研究报告使用的公开资料可能存在信息滞后或更新不及时的风险等。



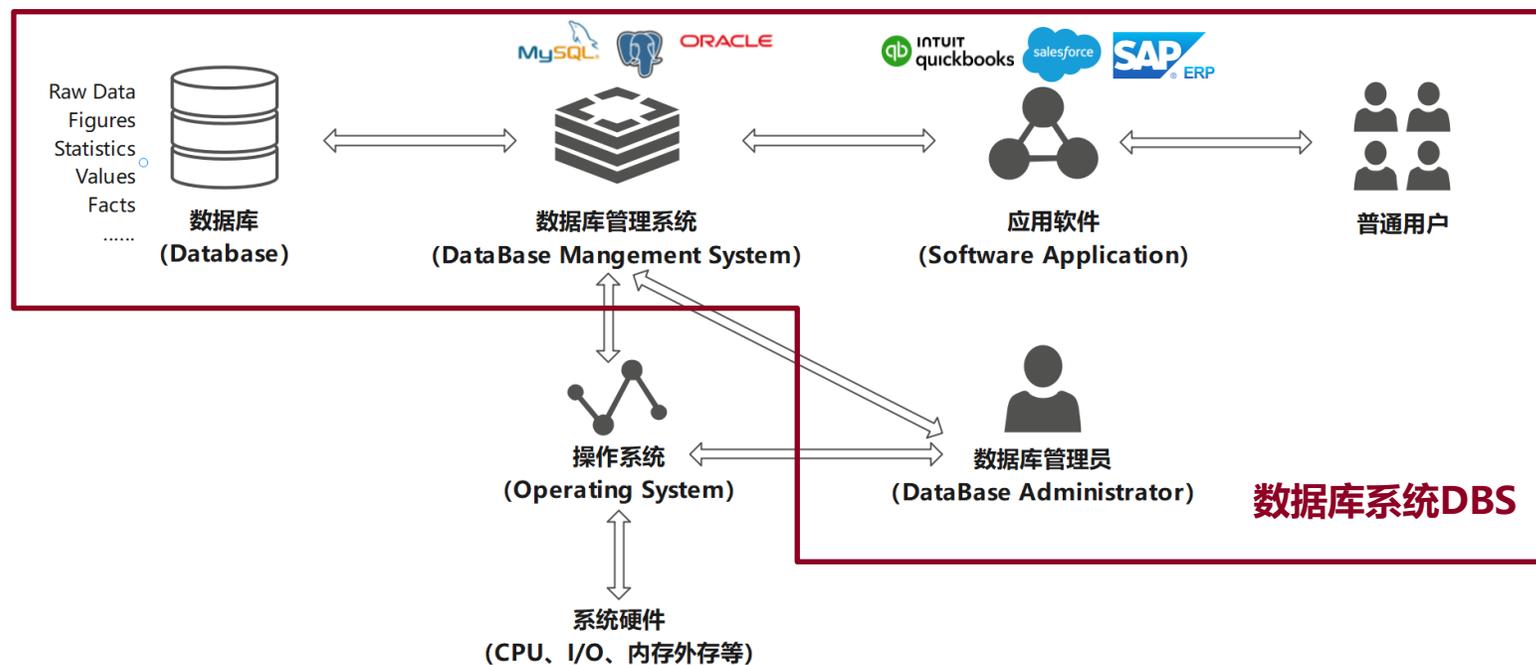
1

数据库：管理数据资源的基础软件

1.1 数据库系统：数据、管理/应用软件与终端用户的统称

- **数据库 (Database, DB)** 数据库是结构化信息或数据的有序集合，一般以电子形式存储在计算机系统中。由数据库管理系统 (Database Management System, DBMS) 控制。DBMS能够与最终用户、应用程序和数据库本身进行交互以捕获和分析数据。广义上可将数据、数据库管理系统、应用软件以及对应的普通用户和管理用户统称为数据库系统 (Database System, DBS)。

图表：数据库系统结构示意图

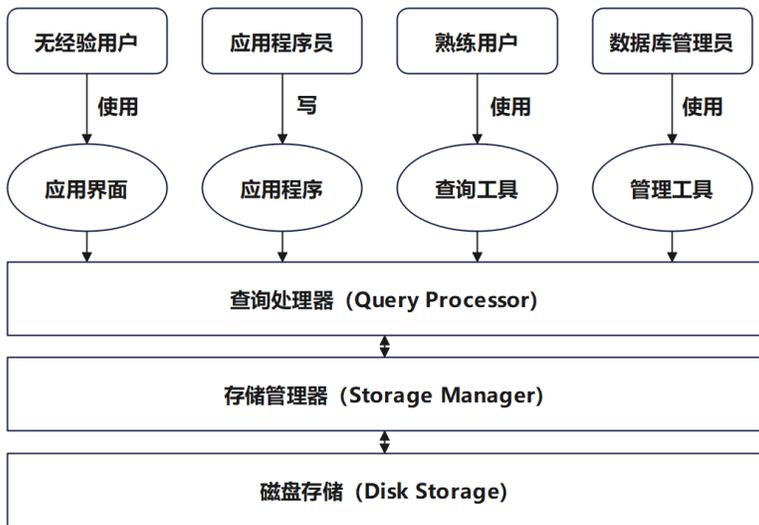


资料来源：Towards Data Science, Collidu, 中泰证券研究所

1.2 数据库软件的定位：计算机系统的核心基础软件

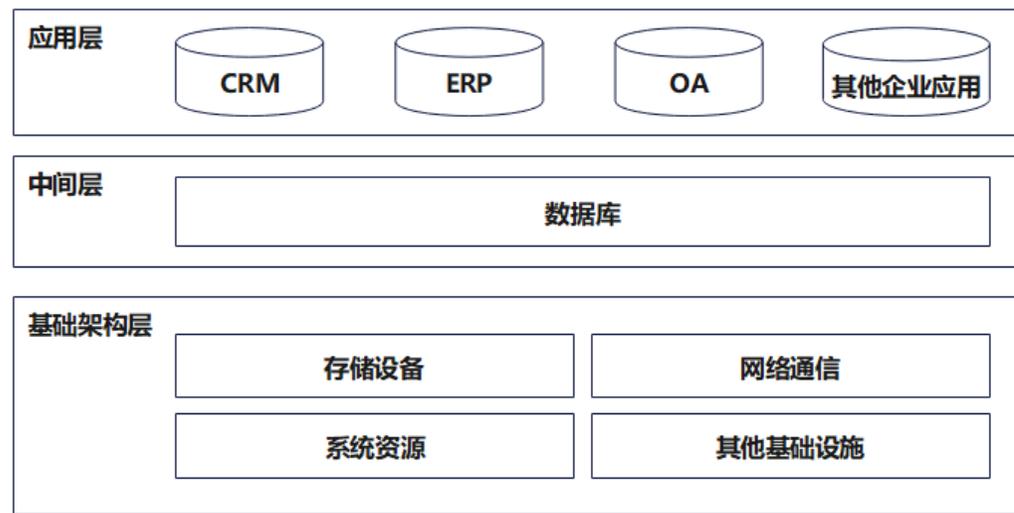
- 数据库作为沟通企业应用软件和底层存储计算资源的纽带：
- 数据库为上层应用提供高效的数据管理和操作功能。数据库提供了易于使用的接口和查询语言，使得应用程序可以轻松地与数据库进行交互。通过数据库连接，上层应用可以实现数据的存储、查询、更新、检索和删除工作，满足应用程序对数据的需求。数据库为底层基础资源提供统一的数据访问接口。通过连接底层基础资源，数据库能够直接与存储设备、网络通信和系统资源进行交互。这种连接能力使得数据库能够有效地管理数据的物理存储，实现数据的持久性和可靠性，同时也为应用程序提供了高度的可扩展性和可定制性。

图表：数据库管理软件的功能架构



资料来源：《数据库系统概念》，中泰证券研究所

图表：数据库连接上下游资源

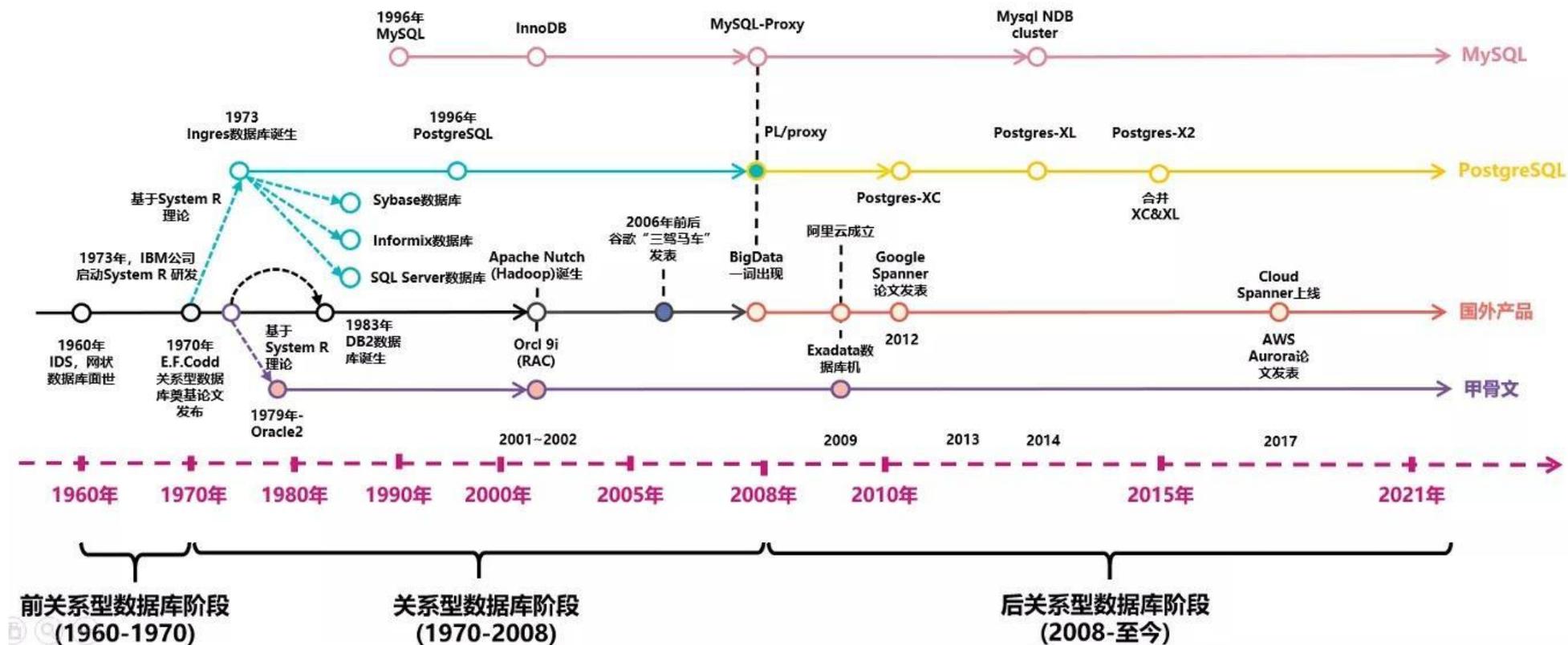


资料来源：信通院云计算开源产业联盟，中泰证券研究所

1.3 数据库发展历史：主要分为三个阶段

■ 数据库发展历程主要可以分为前关系型阶段、关系型阶段和后关系型阶段。

图表：数据库发展历史

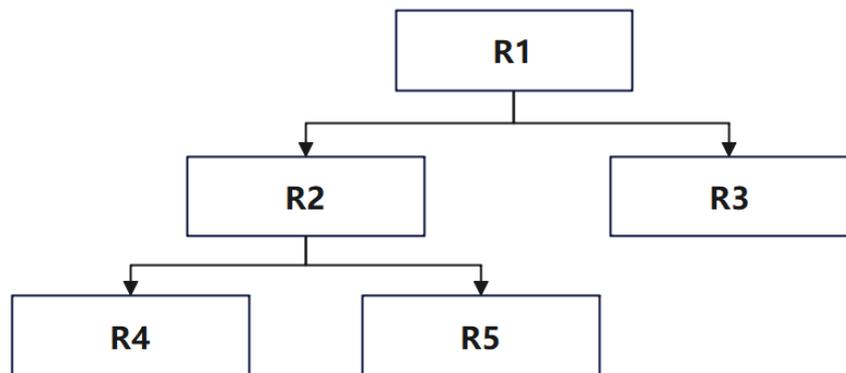


资料来源：信通院，中泰证券研究所

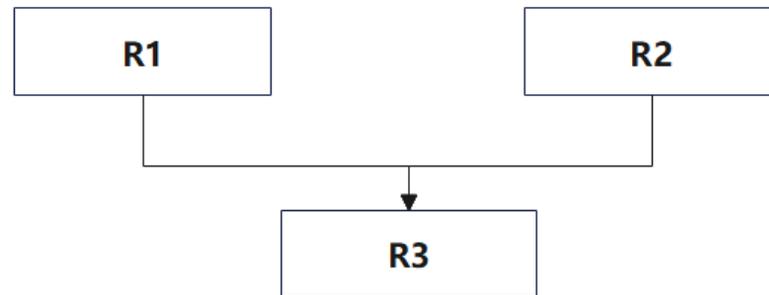
1.3.1 前关系型阶段（1960-1970）：网状、层次数据库出现

- 为解决“阿波罗登月”计划处理庞大数据量的需求，北美航空公司(NAA)开发出 GUAM(Generalized Update Access Method)软件。其设计思想是将多个小组件构成较大组件，最终组成完整产品。这是一种倒置树的结构，也被称之为层次结构。随后 IBM 加入 NAA，将 GUAM 发展成为 IMS(Information Management System)系统并发布于 1968 年，成为最早商品化的层次 DBMS。
- 1963 年，通用电气公司的 Charles Bachman 等人开发出世界上第一个数据库管理系统（以下简称 DBMS）也是第一个网状 DBMS——集成数据存储（Integrated Data Store, IDS）。

图表：层次模型示意



图表：网状模型示意



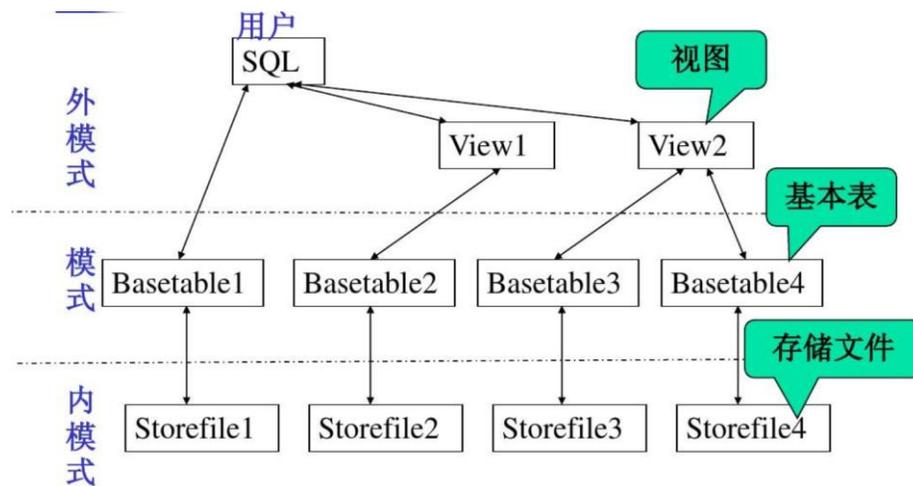
1.3.2 关系型数据库阶段（1970-2008）：SQL成为主流语言，RDB大规模应用

■ 1974年，Ingres原型诞生，为后续大量基于其源码开发的PostgreSQL、Sybase、Informix和Tandem等著名产品打下坚实基础。1977年，Oracle前身SDL成立。1978年，SDL发布Oracle第一个版本。

20世纪80年代，关系型数据库进入商业化时代。1980年，关系型数据库公司RTI成立并销售Ingres。1983年，IBM发布Database2(DB2) for MVS，标志DB2正式诞生。1984年，Sybase公司成立。1985年，Informix发布第一款产品。1986年，美国国家标准局(ANSI)数据库委员会批准SQL作为数据库语言的美国标准并公布标准SQL文本。

■ 1987年，国际标准化组织(ISO)也做出了同样决定，对SQL进行标准化规范并不断更新，使得SQL成为关系型数据库的主流语言。20世纪90年代，Access、PostgreSQL和MySQL相继发布。关系型数据库理论得到了充分的完善、扩展和应用。

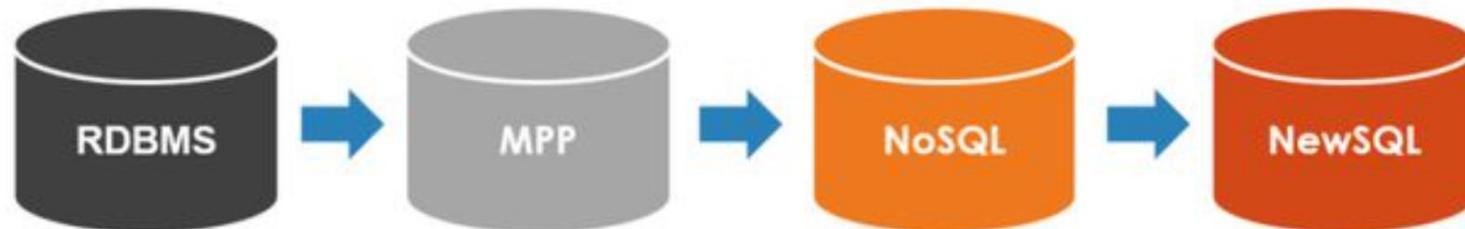
图表：SQL支持关系型数据库



1.3.3 后关系型阶段（2008至今）：模型拓展与架构解耦并存

- 随着信息技术及互联网不断进步，数据量呈现爆发式增长，各行业领域对数据库技术提出了更多需求，数据模型不断丰富、技术架构逐渐解耦，一部分数据库走向分布式、多模处理、存算分离的方向演进。
- 为了解决大规模数据集合和多种数据类型带来的挑战，NoSQL 数据库应运而生，其访问速度快，适宜处理互联网时代容量大、多样性高、流动性强的数据。

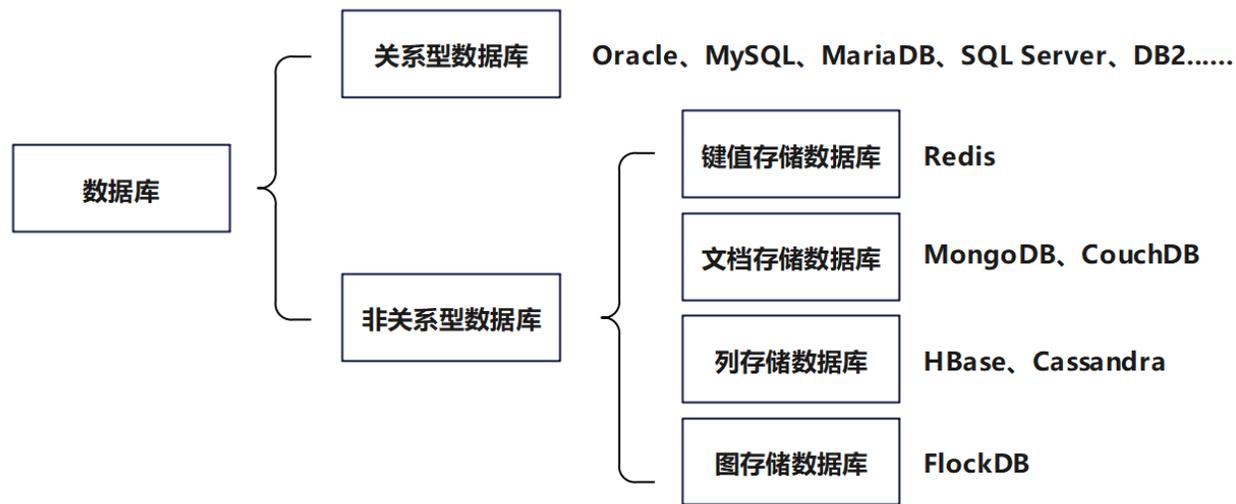
图表：数据库模式发展历程



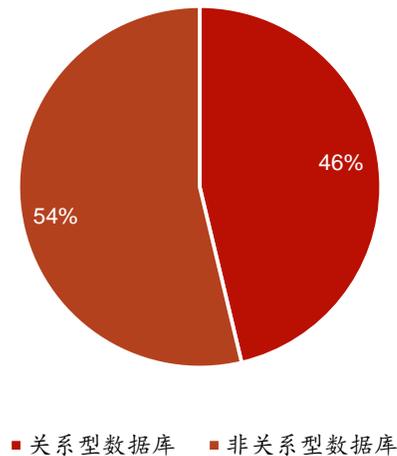
1.4.1 根据数据模型分类：关系型Vs非关系型

- **关系型数据库 (Relational Database, RDB)** 以预定义的关系组织数据，将数据存储在一个或多个由列和行构成的表 (tables) 中，表之间通过特定的关系相互连接。截止2024年6月，全球关系型数据库共有330款，占比46%。
- **非关系型数据库 (Not Only SQL, NoSQL)** 与传统的关系型数据库不同，数据库不使用固定的表结构、主键和外键关系，而是提供更灵活的数据模型以适应不同的应用场景。其类型包括文档数据库 (Document Store)、键值数据库 (Key-Value Store)、列族数据库 (Column Family Store)、图数据库 (Graph Database) 等。截止2024年6月，全球非关系型数据库产品共有383款，占比54%。

图表：关系型与非关系型数据库分类



图表：关系型与非关系型数据库占比 (截止2024年6月)



资料来源：Github，中泰证券研究所

资料来源：CCSA TC601，中泰证券研究所

1.4.1 关系型数据库：依靠关系建立二维表格，满足ACID规则要求

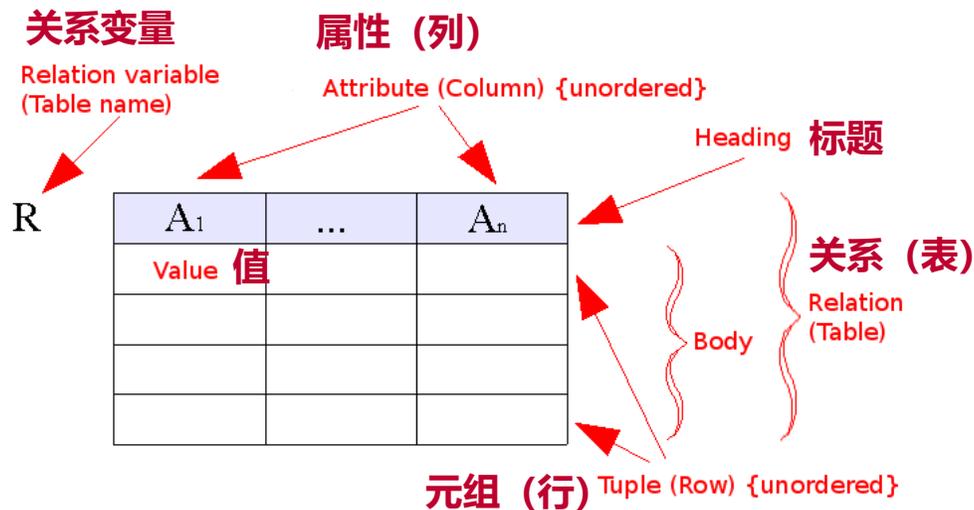
- 关系型数据库以行和列的形式存储数据。关系型数据库这一系列的行和列被称为表，一组表组成了数据库。表是以行（元组Tuple）和列（属性Attribute）的形式组织起来的数据的集合。
- 关系型数据库采用结构化查询语言（即SQL）来对数据库进行查询，也称SQL数据库。能够支持数据库的CRUD（增加，查询，更新，删除）操作。强调ACID规则：原子性（Atomicity）、一致性（Consistency）、隔离性（Isolation）、持久性（Durability），可以满足对事务性要求较高或者需要进行复杂数据查询的数据操作，而且可以充分满足数据库操作的高性能和操作稳定性的要求。

图表：关系型数据库的一对一关系示意

Agents					
AgentID	AgentFirstName	AgentLastName	DateOfHire	AgentHomePhone	<<其他列>>
1	William	Thompson	1997-05-15	555-2681	...
2	Scott	Bishop	1998-02-05	555-2666	...
3	Carol	Viescas	1997-11-19	555-2571	...

Compensation			
AgentID	Salary	CommissionRate	<<其他列>>
1	\$35,000.00	4.00%	...
2	\$27,000.00	4.00%	...
3	\$30,000.00	5.00%	...

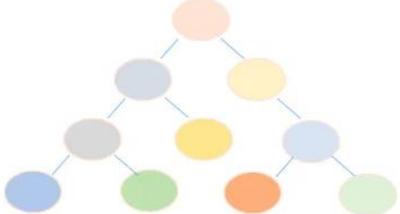
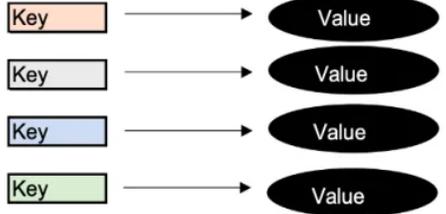
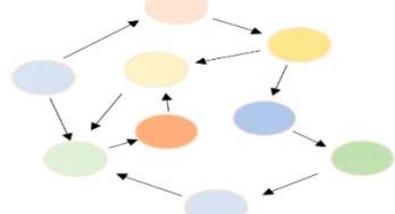
图表：关系型数据库的表（Table）示意



1.4.1 非关系型数据库：适用于大数据量场景，较灵活、易扩展

- 非关系型数据库（NoSQL）不保证关系数据的ACID特性，一般支持BASE原则（基本可用、最终一致性等），适用于大数据量、高性能的场景。代表性的数据库有文档数据库、键值数据库、列式数据库和图数据库四种类型，代表性的MongoDB、Redis、Cassandra和Neo4J等。
- 大多数非关系型数据库具有灵活的数据模型，可以轻松的处理半结构化和非结构化数据；一般设计为分布式系统，能够处理大规模数据和高并发访问，较容易进行水平扩展。

图表：非关系型数据库分类

	文档数据库	键值数据库	列式数据库	图数据库
典型应用场景	Web应用	内容缓存（处理大量数据的高访问负载）	分布式文件系统	社交网络、推荐系统等
优势	数据结构要求不严格，表结构可变，不需要像关系型数据库一样预先定义表结构。	查找速度快、可扩展性强、高性能	查找速度快、可扩展性强、更容易分布式部署	可应用图结构相关算法如最短路径寻址
缺点	查询性能不高，且缺乏统一的查询语法	数据无结构化	功能相对局限	可能需要对整个图进行计算获取信息，不易分布式部署
代表产品	MongoDB	Redis	Cassandra, Hbase	Neo4J
数据模型				

资料来源：腾讯云、CSDN，中泰证券研究所

1.4.2 根据业务负载特性分类：OLTP、OLAP和HTAP数据库

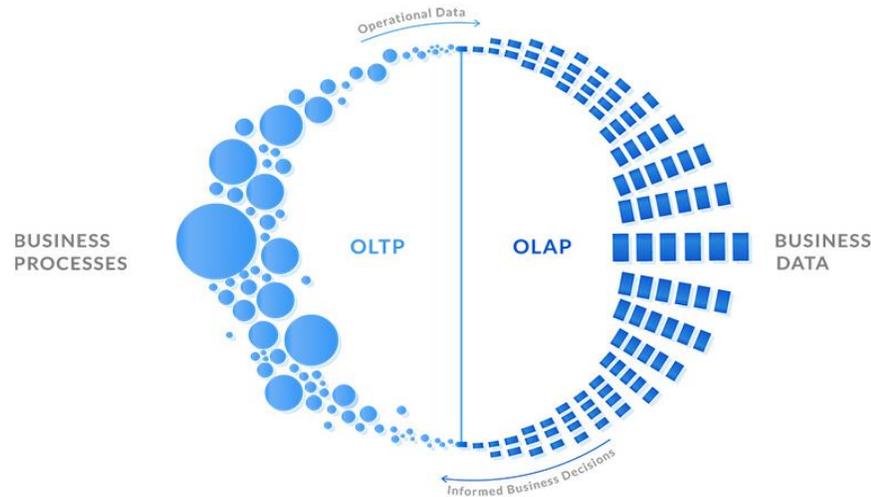
- **OLAP (On-Line Analytical Processing)**：联机分析处理，OLAP是数据仓库系统的主要应用，支持复杂的分析操作，侧重决策支持，并且提供直观易懂的查询结果。系统性能与每秒执行的Transaction以及Execute SQL的数量等有关。它的主要功能包括查询、分析、预测、数据挖掘等，为用户提供灵活的数据分析和快速决策支持。
- **OLTP (on-line transaction processing)**：联机事务处理，传统的关系型数据库的主要应用，主要是基本的、日常的事务处理。OLTP系统通常用于处理企业的日常交易数据，例如订单处理、库存管理、银行交易等。它的主要功能是支持事务和实时数据处理，为用户提供高效的交易处理服务。
- **HTAP (Hybrid Transaction and Analytical Processing)**：也称混合型关系数据库，是能同时提供OLTP和OLAP支持的混合关系型数据库。

图表：OLAP和OLTP的特点

	OLTP	OLAP
面向用户	操作人员，底层管理人员	决策人员、高级管理人员
功能	实时数据处理和高并发事务处理	复杂数据分析和决策
时间需求	实时性要求高，通常在毫秒级	时间要求不严格
空间需求	通常较小，MB到GB级	通常较大，GB到PB级
DB设计	面向应用，如银行业、零售业	面向主题，如销售、库存
DB类型	一般是关系型数据库	非关系型数据库更适用
数据	当前最新数据，一般二维存储	多维、大规模、聚合的历史数据

资料来源：CSDN，中泰证券研究所

图表：OLTP和OLAP的差异



资料来源：SAP，中泰证券研究所

1.4.3 根据架构分类：分布式和集中式

- 按照部署方式，数据库可以分为集中式数据库和分布式数据库两类。
- 集中式架构的数据库的软硬件价格昂贵，更适合传统的、对可靠性需求较高的业务；分布式架构硬件平台兼容性强，能适合互联网等行业公司灵活、高并发的业务场景，但在安全性上有一定隐患。

图表：分布式和集中式数据库的特点

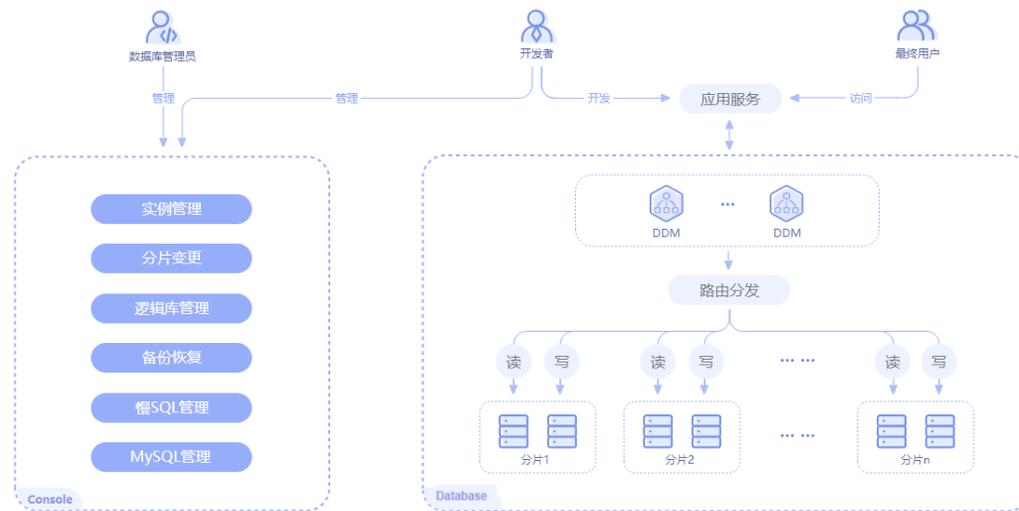
	集中式架构	分布式架构
价格成本	软硬件价格昂贵，商用集中架构设备基本被IBM、Oracle、EMC三家公司垄断采购和后续运维的持续投入大；	成本合理可控，云平台降低、分摊研发投入，边际成本下降迅速
兼容性	限制较多，硬件平台兼容性差	灵活方便，硬件平台兼容性强
扩展伸缩	适应传统业务特性，对项目规划要求较高，业务支撑能力有设计上限	互联网业务特性，能够灵活支撑高并发交易
可靠性	可靠性高	可靠性较高
安全性	安全性高	需要通信网络实现分布控制，安全性存在隐患

资料来源：51CTO，中泰证券研究所

1.4.3 分布式数据库实现方法：中间件分布式与原生分布式

- 分布式数据库解决方案可以分为分布式数据库中间件和原生分布式数据库。
- 分布式数据库中间件是架构在多个传统单点数据库系统上的中间层解决方案，通过将数据分拆到不同的数据库节点上，利用中间件来管理和访问各个数据库中的数据，通常需要用户参与到数据分拆和节点管理过程中。
- 原生分布式数据库是指从架构设计、底层存储和查询处理均面向分布式数据管理需求，数据库集群作为一个整体对外提供服务，用户无需关注集群内部的实现细节。从实现方式来看，可以先进行分布式存储，再叠加数据库能力；也可以基于分布式数据库理论实现的原生分布式数据库，先部署TP单机数据库引擎，再叠加分布式能力。

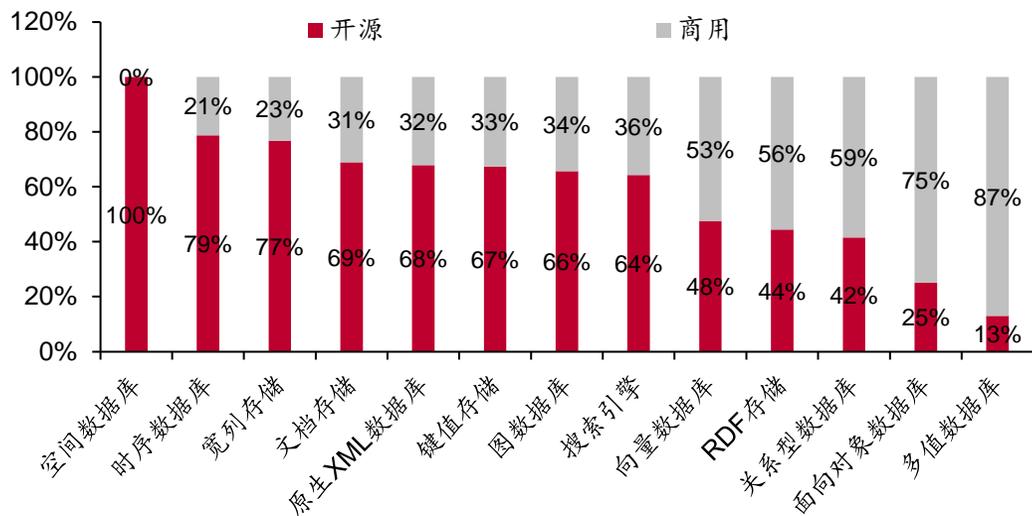
图表：华为DDM分布式数据库中间件



1.4.4 根据商业模式分类：开源数据库与商用（闭源）数据库

■ 数据库主要分为开源和商用（闭源）两种。截止2024年7月，关系型数据库产品中约有41.5%为开源数据库，58.5%为闭源商用数据库。空间数据库、时序数据库等模型架构的数据库中开源占比较多，多值数据库、面向对象数据库、关系型数据库等类别中闭源商用占比较多。

图表：开源Vs商用数据库占比（2024年7月统计）



图表：排名前五的开源和闭源数据库（2024年7月统计）

排名	商业数据库	分数	全部排名	开源数据库	分数	全部排名
1	Oracle	1240	1	MySQL	1039	2
2	Microsoft SQL Server	808	3	PostgreSQL	639	4
3	Snowflake	137	7	MongoDB	430	5
4	IBM Db2	124	9	Redis	157	6
5	Microsoft Access	101	11	Elasticsearch	131	8

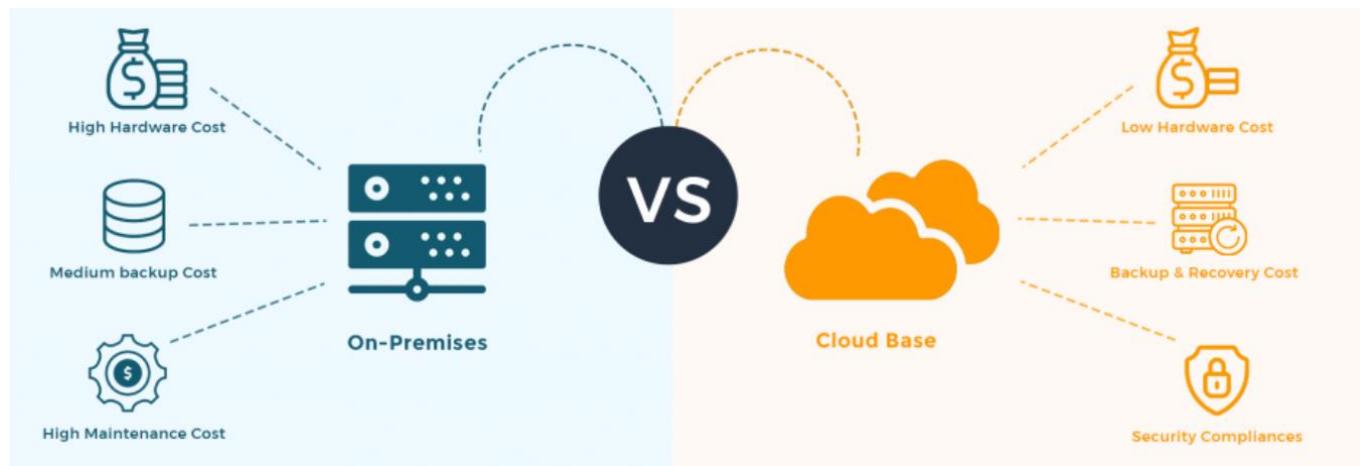
资料来源：DB Engines，中泰证券研究所

资料来源：DB Engines，中泰证券研究所

1.4.5 根据部署方式分类：云数据库与本地数据库

- 根据部署方式不同，数据库可以分为云数据库和本地数据库。
- 云数据库是一种基于云计算平台提供的数据库服务，它允许用户在云端存储、管理和处理数据。传统的普通数据库通常需要在本地进行部署和维护，需要购买和管理相应的硬件和软件。而云数据库则部署在云环境中，用户只需通过互联网进行访问即可，无需关心底层硬件和软件的配置和维护。
- 云数据库具有出色的可扩展性。企业可以根据业务需求随时增加或减少存储容量，而无需进行繁琐的硬件升级或数据迁移。而普通数据库在扩展时需要经历繁琐的升级和数据迁移过程，对时间和资源的消耗较大。

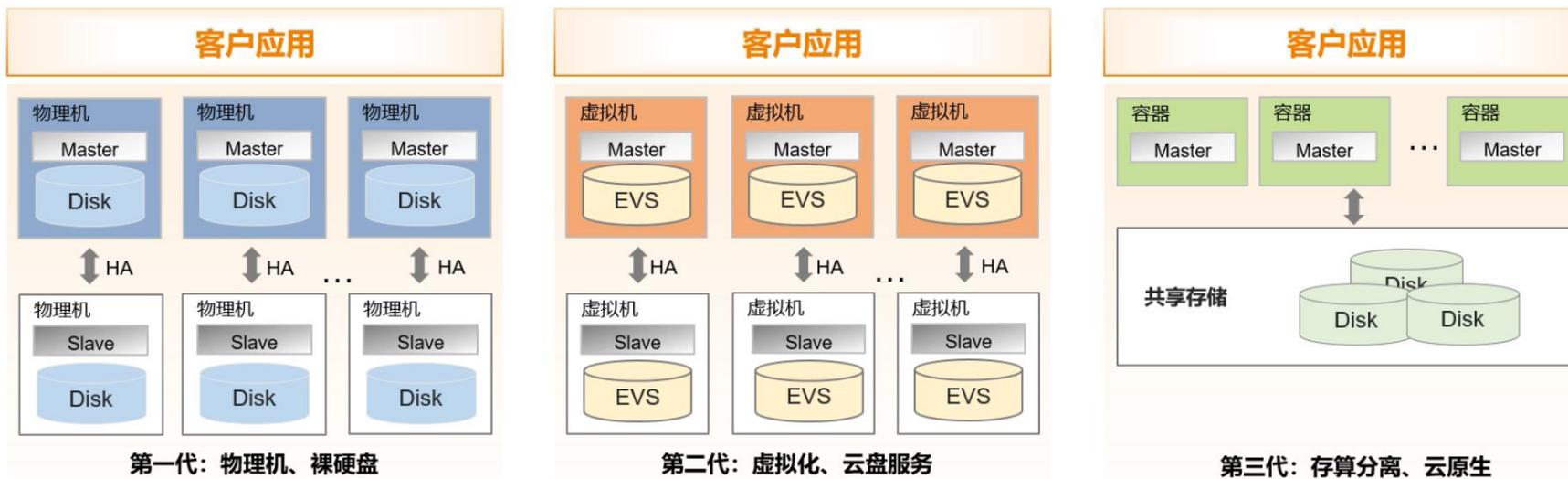
图表：本地数据库与云数据库的区别



1.5 数据库正走向云原生：存算分离+多模数据正成为数据库系统设计的趋势

- 云原生正在成为数据库系统设计的趋势，能够充分利用云平台的池化资源，适应云平台的基础设施。与传统数据中心采用的物理服务器、物理网络、物理存储设施不同，云计算环境下普遍采用基于容器化、微服务、Serverless 等理念进行基础设施的架构，以云化运行环境为前提，在设计上通常结合分布式技术并普遍采用计算——存储分离、日志即数据的设计思想，能够灵活调动资源进行扩缩容，可进一步实现资源池化、弹性变配、集约运维等能力，以应对更便捷、更低门槛实现云上数字化转型与升级的挑战。

图表：云原生数据库演进



1.6 数据湖：大数据时代的又一存储选择，湖仓一体或成未来趋势

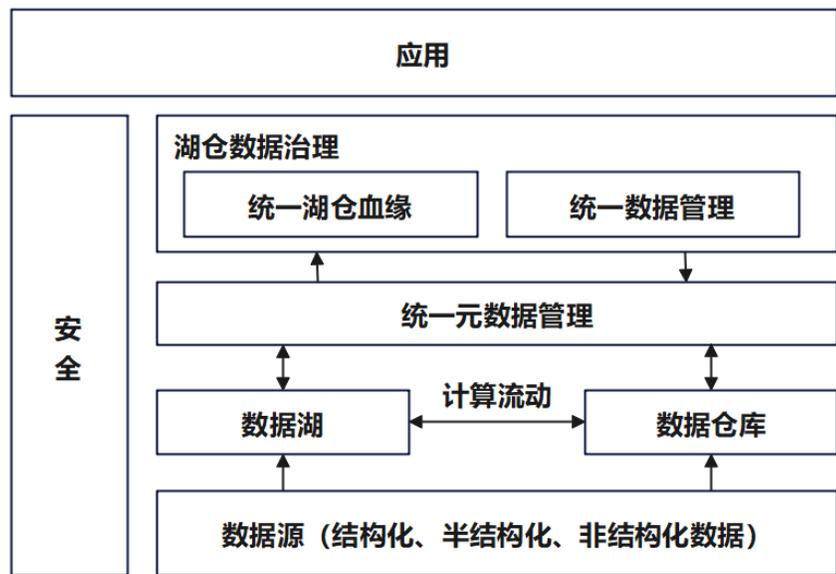
- 数据湖可存储结构化和非结构化数据,是一种面向大规模、多来源、高度多样化数据的组织方法。数据湖将数据管理的流程简化为数据入湖和数据分析两个阶段,能够支持各种类型数据的统一存储。数据分析则以读取型 (schema on read) 形式,极大提升分析效率。
- 湖仓一体技术或成为未来普遍应用的数据治理架构。通过融合数据湖与数据仓库的优势,可以形成一体化、开放式数据处理平台,使得数据处理平台底层支持多数据类型统一存储,实现数据在数据湖、数据仓库之间无缝调度和管理,并使得上层通过统一接口进行访问查询和分析。

图表：数据平台的发展历程



资料来源：亿信华辰软件，中泰证券研究所

图表：湖仓一体数据库架构设计

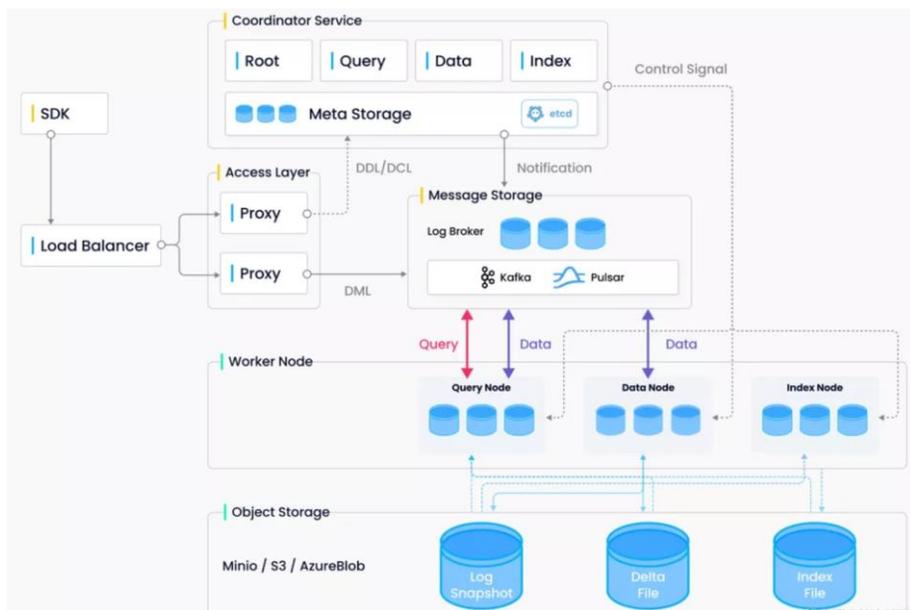


资料来源：亿信华辰软件，中泰证券研究所

1.7 AI与数据库融合，RAG向量数据库迎来发展机遇

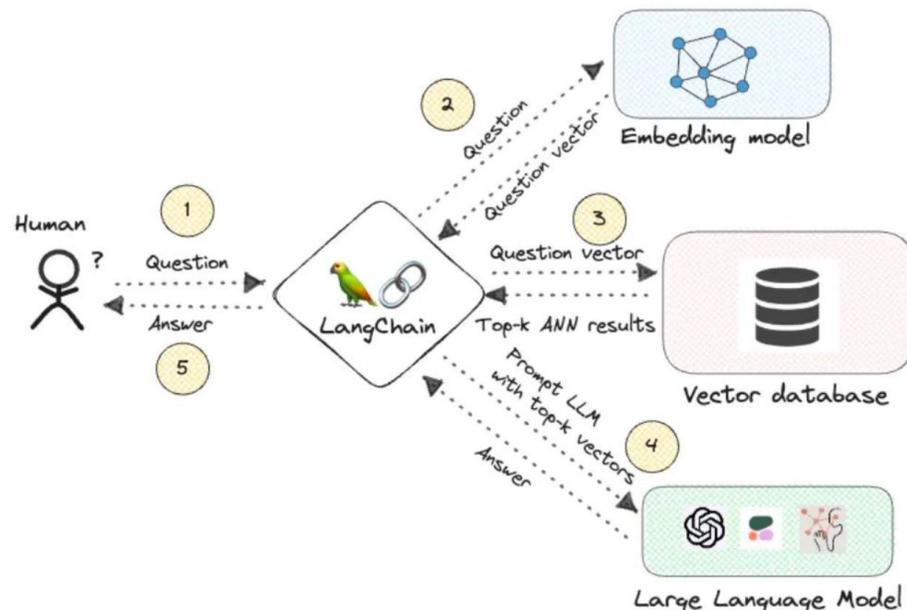
- 人工智能技术快速发展，为数据库与AI深度融合带来新机遇。生成式AI在数据库结构设计、架构设计、数据分析挖掘等方面可以不同程度简化人员操作，提高开发、运维、分析的效率；多模态数据存储和计算的需求随着大语言模型出现而剧增，向量数据库的重要程度不断提升。
- Zilliz 开源了全球最流行的向量数据库 Milvus，并于 2022 年正式发布 Zilliz cloud，一款公有云全托管的向量检索 SaaS 产品。腾讯，阿里等云厂商也纷纷跟进，发布了自己的向量数据库产品。

图表：Milvus系统架构



资料来源：51CTO，中泰证券研究所

图表：向量数据库在大模型中的应用



资料来源：51CTO，中泰证券研究所



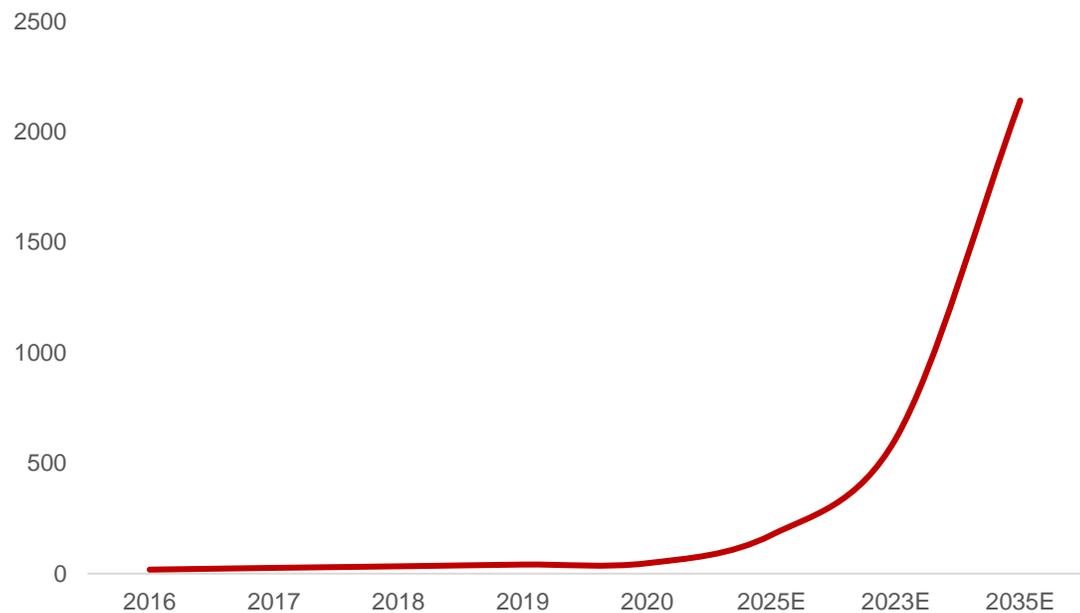
2

数据库产品百花齐放，
国产替代正当时

2.1 数据规模快速增长，价值与日俱增

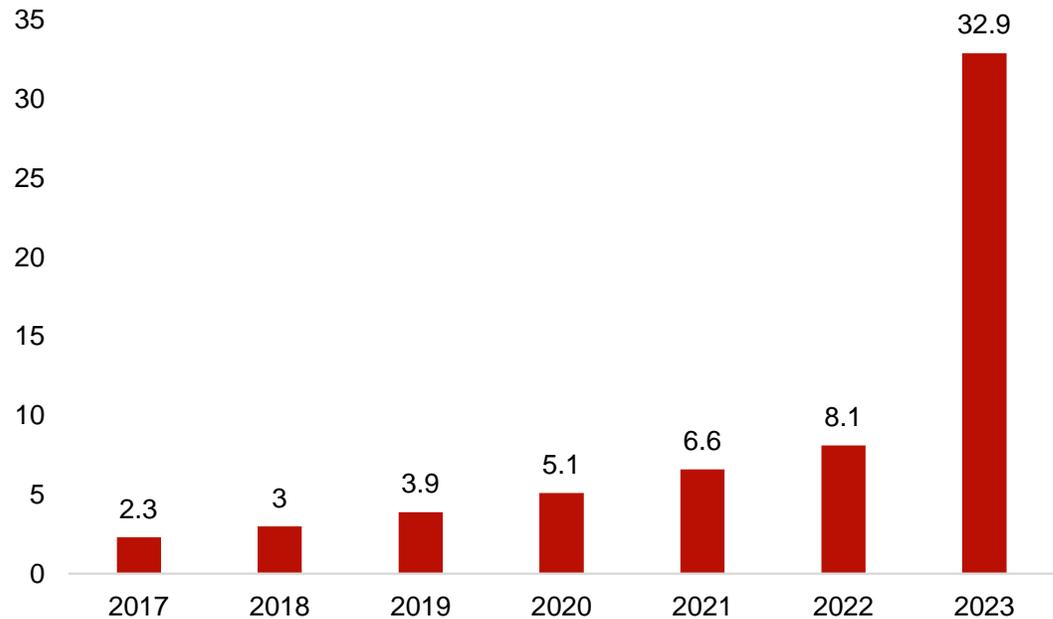
- 根据艾瑞咨询的报告显示，到2035年，全球生产和存储的数据总量将达到2142ZB（泽字节），涵盖交通、零售、工业、教育、卫生、政务等社会经济各个领域。
- 我国数据资源总量快速提升，对数据资源的管理需求将会不断增长。全国2023年数据生产总量达到32.85泽字节（ZB），约相当于1000多万个中国国家图书馆的数字资源总量。预计到2025年中国数据总量预计将跃居全球第一，占比有望达到27%以上。

图表：全球数据量（ZB）



资料来源：艾瑞咨询，中泰证券研究所

图表：2017-2023年我国数据产量（ZB）

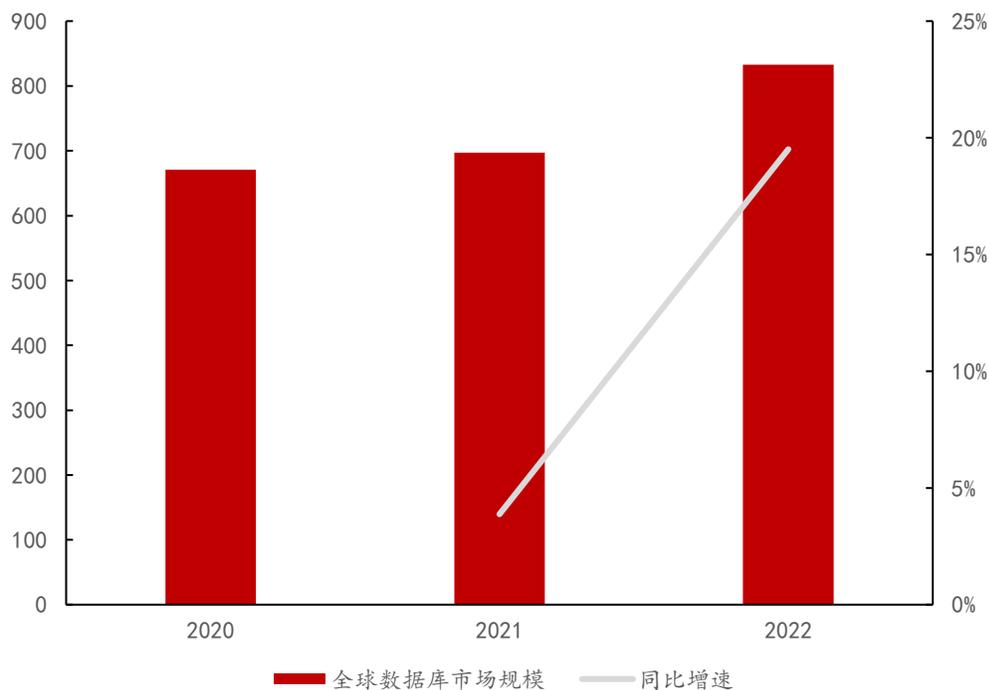


资料来源：艾瑞咨询，中泰证券研究所（其中2017-2022年为网信办调研发布，2023年为国家数据局、网信办多部门进行数据资源调查公布，口径有所差异）

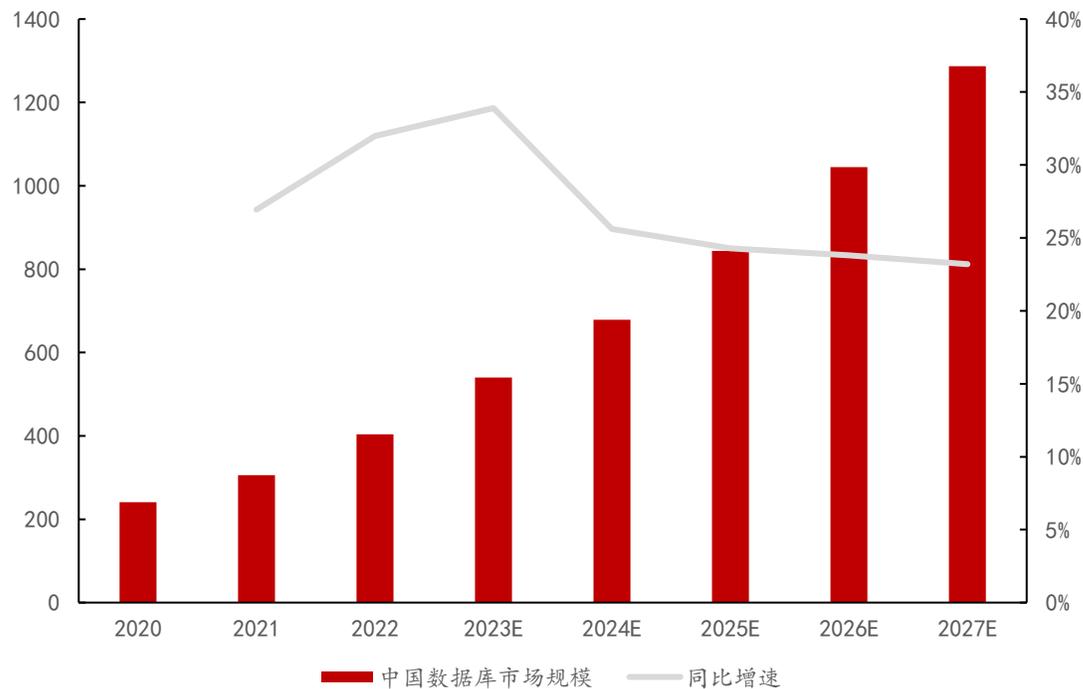
2.1 数据库市场处于高速发展阶段，我国数据库产业欣欣向荣

- 据 CCSA TC601 测算，2022 年全球数据库市场规模为 833 亿美元，同比增长19.5%。其中中国数据库市场规模为 59.7 亿美元（约合 403.6 亿元人民币），占全球市场的7.2%；到2027年，中国数据库市场总规模有望增至1286.8亿元。

图表：全球数据库市场规模



图表：中国数据库市场规模



资料来源：中国信通院、中泰证券研究所 单位：亿美元

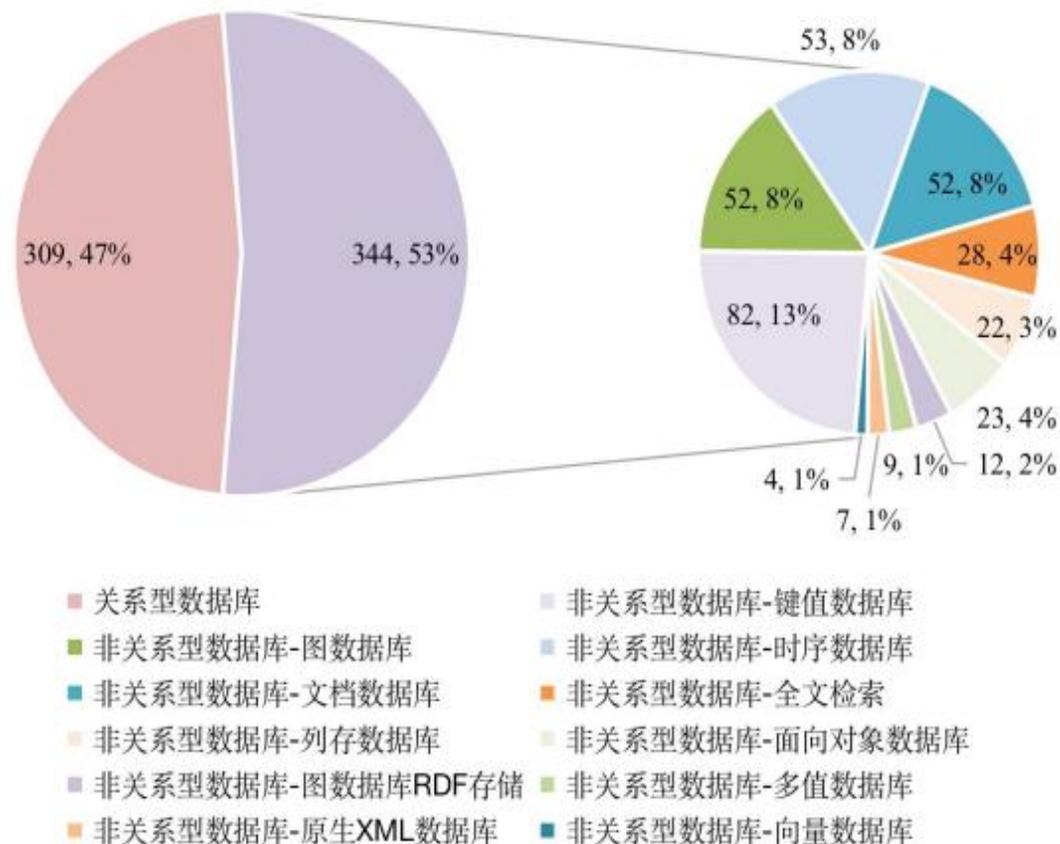
资料来源：中国信通院、中泰证券研究所 单位：亿元

2.2.1 全球视角：全球数据库产品类型分布多样

- 全球数据库产品数量整体分布呈现以非关系型及混合型数据库为主。（截止至2023年6月）

- 全球数据库产品中，关系型数据库达 309 个，非关系型数据库有达344 个，占比分别为 47.3%和 52.7%。
- 非关系型数据库中，键值型数据库 82 个、时序数据库 53 个、图数据库 52 个，在非关系数据库中分别占比 23.8%、15.4%和 15.1%。

图表：全球数据库产品类型分布



2.2.2 全球视角：流程度-国外数据库受追捧，关系型数据库占主导地位

- 根据DB-engines的数据显示，Oracle、MySQL、Microsoft SQL Server深受欢迎，稳居前三。
- 关系模型占主导地位。从数据库结构来看，关系模型因其特有的原子性、一致性、隔离性和持久性优势，成为了当代主流的数据模型。
- 多模数据库成为新趋势。随着业务的多样化，应用对于数据的多类型处理能力提出了更高的要求。从表格中可以看到，主流数据库系统打上了Multi-model的能力标签。

图表：全球数据库流行度情况

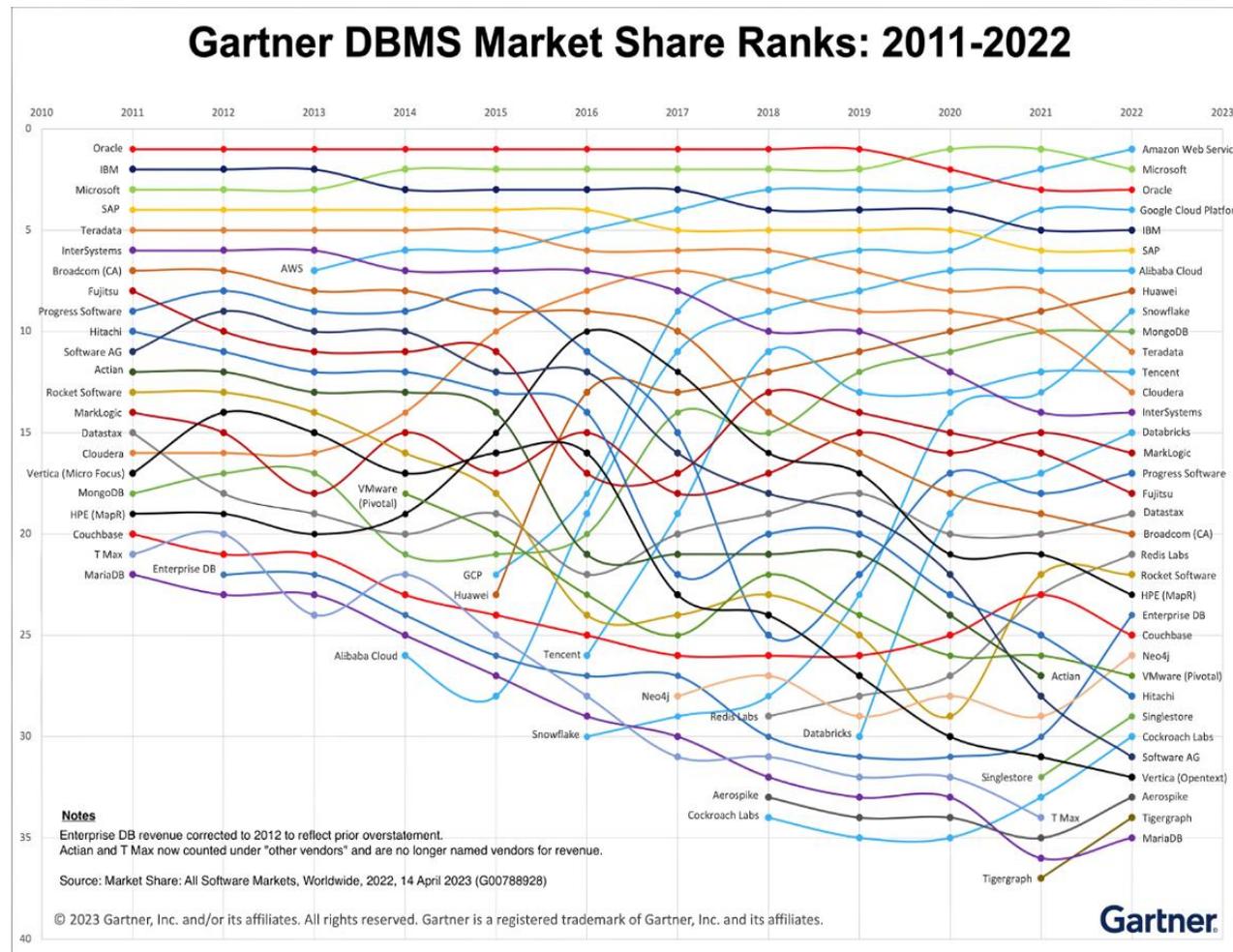
	排名			数据库	数据结构
	7月 2024	6月 2024	7月 2023		
	1	1	1	Oracle	Relational, Multi-modle
	2	2	2	MySQL	Relational, Multi-modle
	3	3	3	Microsoft SQL Server	Relational, Multi-modle
	4	4	4	PostgreSQL	Relational, Multi-modle
	5	5	5	MongoDB	Document, Multi-modle
	6	6	6	Redis	Key-value, Multi-modle
↑	7	8	11	Snowflake	Relational
↓	8	7	8	Elasticsearch	Search engine, Multi-modle
↓	9	9	7	IBM Db2	Relational, Multi-modle
↓	10	10	10	SQLite	Relational
↑	11	11	9	Microsoft Access	Relational
↓	12	12	12	Apache Cassandra	Wide column, Multi-modle
↑	13	14	14	Splunk	Search engine
↓	14	13	13	MariaDB	Relational, Multi-modle
↑	15	15	18	Databricks	Multi-modle
↓	16	16	15	Microsoft Azure SQL Database	Relational, Multi-modle
↓	17	17	16	Amazon DynamoDB	Multi-modle
↑	18	19	20	Google BigQuery	Relational
↓	19	18	17	Apache Hive	Relational
↑	20	20	21	FileMaker	Relational
↑	21	21	22	Neo4j	Gtaph
↓	22	22	19	Teradata	Relational, Multi-modle
	23	23	23	SAP HANA	Relational, Multi-modle
	24	24	24	Apache Solr	Search engine, Multi-modle
	25	25	25	SAP Adaptive Server	Relational, Multi-modle
↑	26	26	27	Apache Hbase	Wide column
↓	27	27	26	Microsoft Azure Cosmos DB	Multi-modle
	28	28	28	InfluxDB	Time Series, Multi-modle
	29	29	29	PostGIS	Spatial, Multi-modle
↑	30	30	31	Firebird	Relational

2.2.3 全球视角：市场份额-新技术掀起市场格局洗牌

- 据 Gartner发布的2022年数据库市场份额报告来看，回顾2011年至2022年数据库厂商的市场份额排名，我们可以看到几大变化：

- 在新兴技术的影响下，市场领导者的地位发生了更迭。“云+数据库”战略成为技术发展新趋势，AWS以及微软占领市场前二；
- 中国数据库厂商快速发展。在国产化替代以及云技术趋势下，阿里云、华为以及腾讯排名快速上升；
- 非关系型数据库重视度提升。图数据库受到认可，2022年Neo4j、TigerGraph 排名有所上升。

图表：2022年数据库市场份额情况

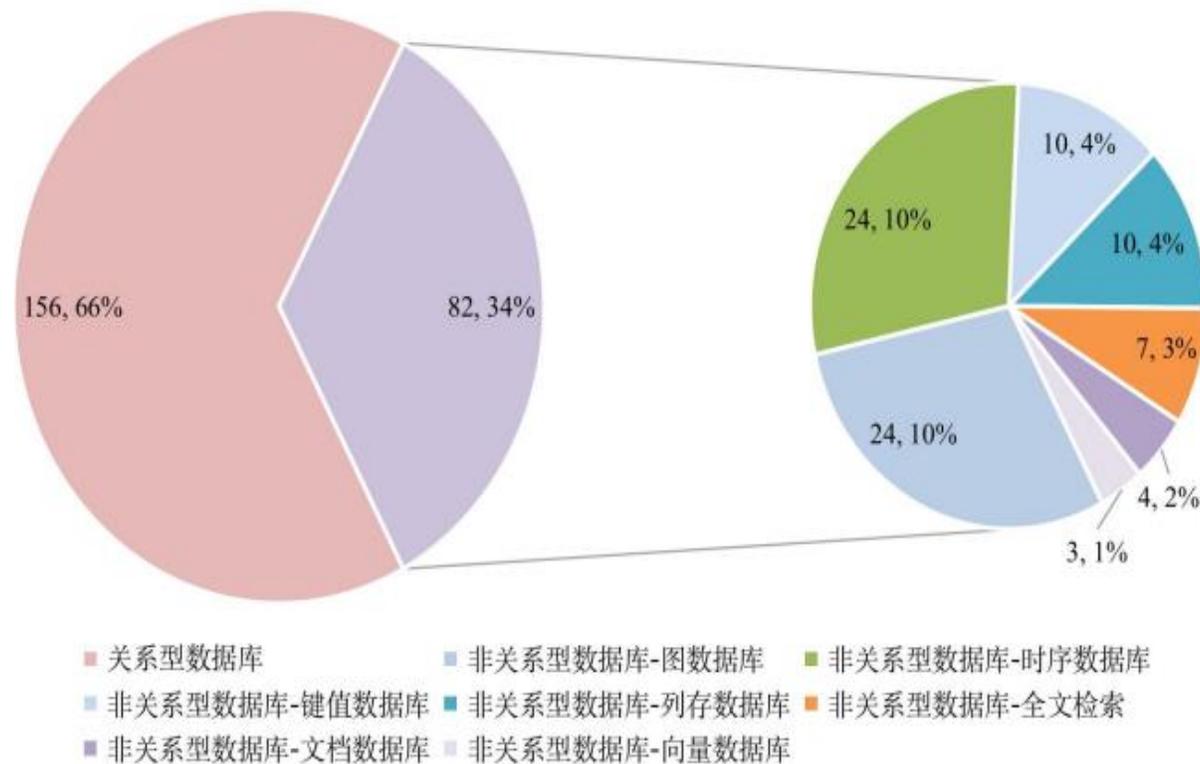


2.3.1 国内视角：国内数据库产品类型以关系型为主

- 我国数据库产品数量呈现以关系型为主，非关系型数据库为辅的局面。

- 依据CCSA TC601的数据显示，截止至2023年6月，国内的关系型数据库 156 个，非关系型数据库有 82 个，占比分别为 65.5%和 34.5%。
- 在国内的非关系型数据库市场中，图数据库和时序数据库各有24个产品，分别占据了29.3%的市场份额。键值数据库和列存数据库各有10个产品，各自占据市场份额的12.2%。

图表：中国数据库产品类型分布

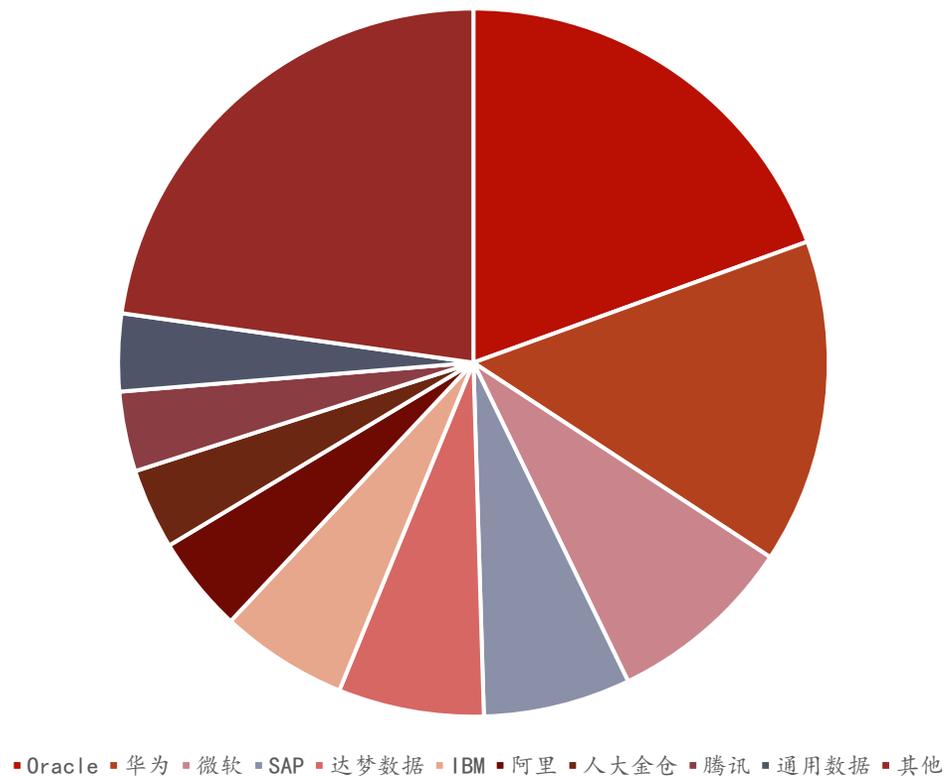


资料来源：CCSA TC601、中泰证券研究所

2.3.2 国内视角：市场份额-海外厂家仍占主导，国产替代未来可期

- 在本地部署模式下，国内关系型数据库市场格局中，海外厂商仍占据较多市场份额，国产化替代加速进行中。
 - Oracle 以19.45%的份额占主导地位，微软以8.57%排名第三位；
 - 国产数据库正在发力，华为、达梦数据、阿里、人大金仓等排名靠前，不断推动国产数据库事业的进步。
- 市场竞争愈发激烈，前五名厂商份额达**56.16%**。

图表：2022年中国关系型数据库市场份额-本地部署



资料来源：IDC、达梦数据招股书、中泰证券研究所

2.3.3 国内视角：流程度-国产数据库厂商势如破竹

- 各大国产数据库厂商在不同领域表现势如破竹。

- PolarDB在专利、案例以及论文方面的数据表现优异居于榜首；
- OceanBase拿下榜眼，6月份，OceanBase进行了一次关键商业模式转变，公司将更加独立地面向市场竞争；
- openGauss拿下探花，openGauss社区推出“结队计划”，加强社区与伙伴、客户之间的连接；
- 达梦数据库7月跃升两位，且已成功登陆科创板；
- PingCAP坚持国际化，在国际上取得了亮眼的表现；
- 人大金仓屡获订单，展现强大实力。

图表：2024年7月中国数据库排行榜

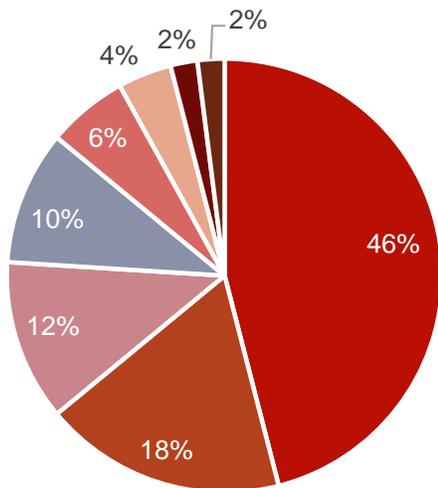
排行	上月	半年前	名称	模型	数据处理	部署方式	商业模式	专利	论文	案例	资质	书籍	得分
1	↑ 2	↑ 2	PolarDB	关系型	HP	☒ ☒ ☒	☒ ☒	592	73	74	14	2	812.78
2	↓ 1	↓ 1	OceanBase	关系型	HP	☒ ☒	☒ ☒	151	26	43	14	1	670.68
3	↑ 4	3	openGauss	关系型	TP	☒	☒	573	11	31	7	6	602.19
4	↑ 5	4	TiDB	关系型	HP	☒ ☒	☒ ☒	40	59	18	8	1	592.32
5	↓↓ 3	5	人大金仓	关系型	TP AP	☒ ☒ ☒	☒	534	1	25	12	3	578.58
6	↑↑ 8	6	达梦数据库	关系型	TP	☒ ☒ ☒	☒	518	0	19	8	11	559.06
7	↓ 6	7	GBASE	关系型	AP TP	☒ ☒ ☒	☒	219	0	46	13	0	493.53
8	↓ 7	8	GaussDB	关系型	HP	☒ ☒	☒	679	22	19	15	4	427.58
9	9	↑ 10	GoldenDB	关系型	HP	☒ ☒	☒	697	80	73	11	2	360.34
10	10	↓ 9	TDSQL	关系型	HP	☒ ☒	☒ ☒	136	19	16	12	0	236.46

资料来源：墨天轮、中泰证券研究所

2.4.1 信创：作为核心系统，央国企必须于2027年完成国产替代

■ 国产信创替代正在走向行业阶段。“十四五”规划中明确指出到2025年行政办公及电子政务系统要全部完成国产化替代。2022年9月，国资委发布79号文提到要求国央企落实信息化系统的信创国产化改造，要求央企、国企、地方国企全面落实信创国产化，并要求2027年全部完成国产化替代。目前国产替代最领先的为金融领域；从下游行业分布来看，金融行业占比最大（约为46%），其次为政府、电信、互联网等行业。

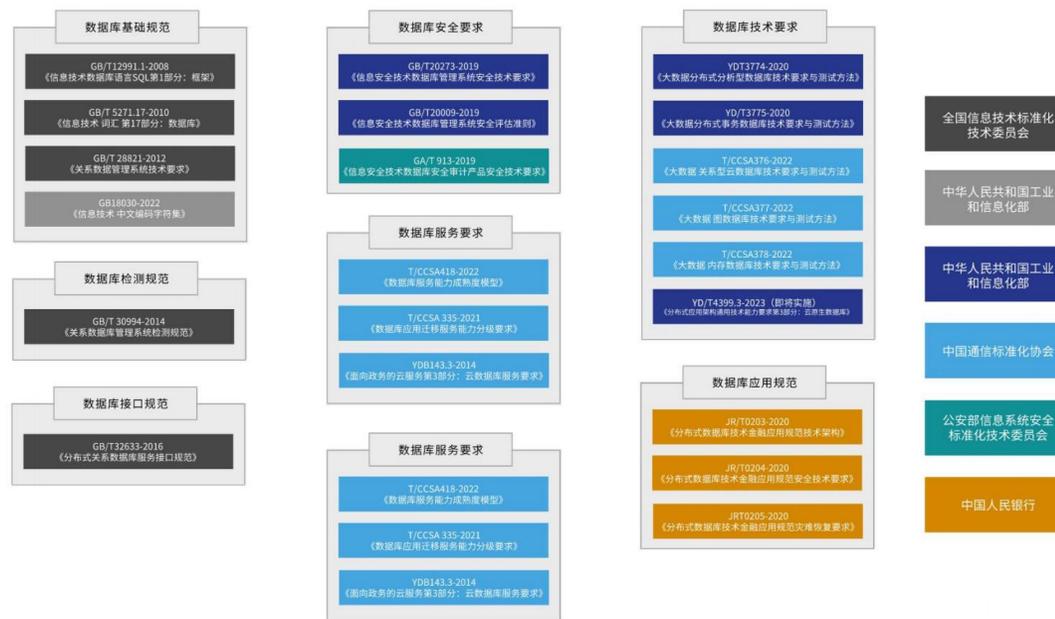
图表：2023年国产数据库行业应用分布



■ 金融 ■ 政府 ■ 电信 ■ 互联网 ■ 交通 ■ 能源 ■ 教育 ■ 医疗

资料来源：亿欧智库，中泰证券研究所

图表：国产数据库标准分类及归口部门



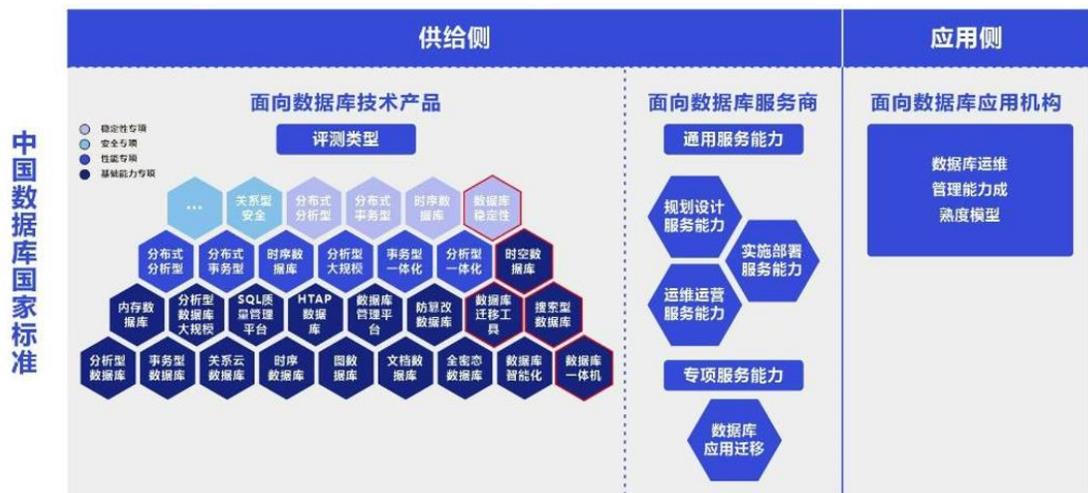
资料来源：墨天轮，中泰证券研究所

2.4.2 国产替代面临的困难：技术壁垒高、选型困难、生态护城河难以逾越

- 国产数据库性能较海外大厂有所差距。数据库产品拥有极高的技术know-how壁垒，目前部分国产数据库性能不稳定、不达标、生态兼容差，替换核心系统有较大压力，甚至无法替代。
- 国产数据库类别多，技术路线不一，竞争格局分散。面对近300种国产数据库，客户选型周期长、难度大、风险高，采购、迁移、实施、运维等环节累计综合成本可能比替换前更高。
- 国产数据库系统在生态构建上较海外大厂还有差距。从产品角度看，成熟的数据库产品通常拥有完善的配套生态工具；从整体行业角度看，行业标准建立、行业应用集成、开源生态、人才培养、硬件适配等诸多生态问题难以解决。

图表：数据库国家标准（2023年6月）

图表：阿里云数据库生态工具



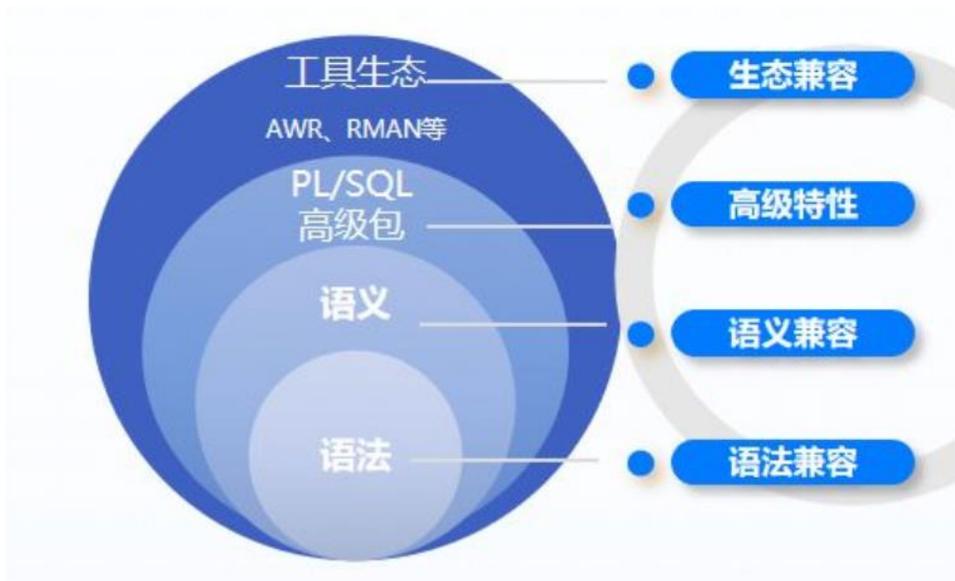
资料来源：CCSA TC601，中泰证券研究所

资料来源：阿里云，中泰证券研究所

2.4.3 国产数据库企业的破局之道：兼容Oracle、MySQL等语法，降低迁移成本

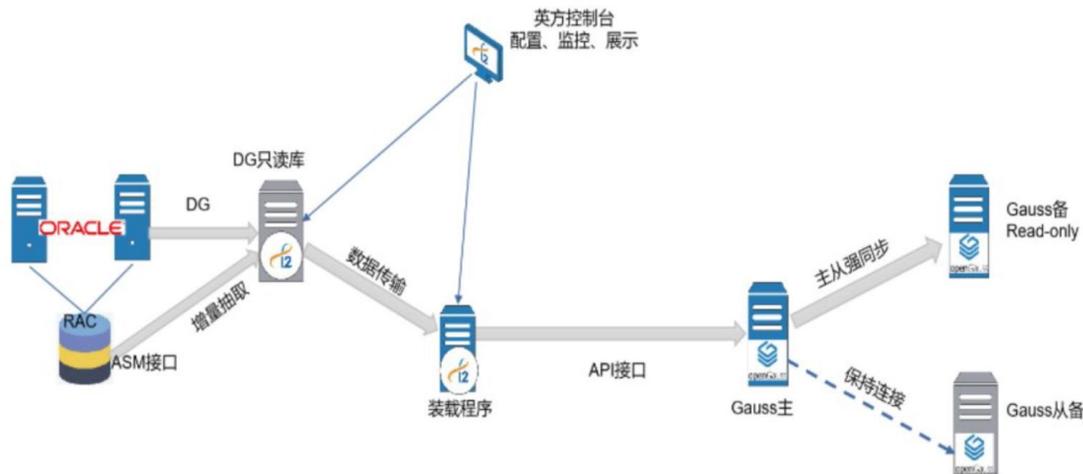
- 由于大批量的应用是基于Oracle数据库构建的，国产数据库厂商就要提供在驱动、内核、资源、架构等层面与Oracle兼容的能力，以降低用户的替换成本。如华为GaussDB实现了对Oracle和MySQL常用语法的高度兼容以及一站式的迁移解决方案；人大金仓KingbaseES能够提供Oracle、MySQL和SQL Server三大兼容模式。

图表：Oracle的四层生态兼容



资料来源：CSDN，中泰证券研究所

图表：某国有银行Oracle到GaussDB的数据复制项目



资料来源：英方云，中泰证券研究所



3

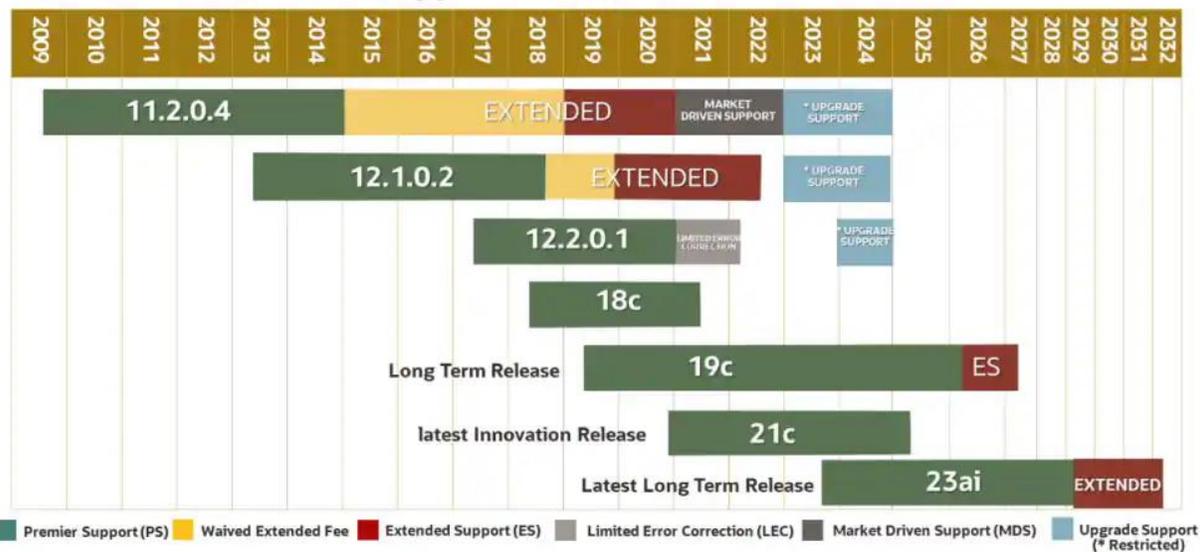
国内外数据库厂商蓬勃发展

3.1.1 海外厂商：Oracle-全球数据库管理龙头企业

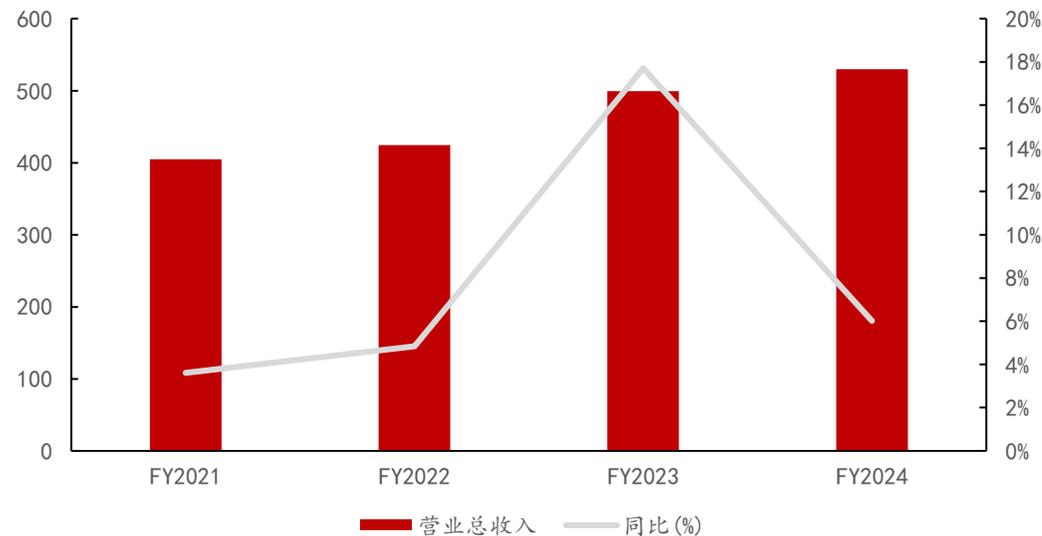
- **Oracle公司是全球数据库管理龙头企业。** Oracle以数据库业务起家，通过不断的战略收购，逐渐发展为提供全面解决方案的全球领先的信息管理软件及服务厂商，公司在1986年成功在纳斯达克上市，目前已拥有超40万客户。
 - Oracle数据库自最初的V2版到至今的23ai版，产品持续进行创新和升级。Oracle Database 23ai推出的新功能AI Vector Search，使新的和现有的应用能够以语义方式搜索数据，并将数据的含义与数据库中已存储的业务数据相结合。
 - 公司FY2024收入高达530亿美元，同比增速达6%；公司拥有15.9万员工，为全球客户服务。

图表：Oracle 各版本support时间线

Database Releases and Support Timelines



图表：Oracle营业收入情况



3.1.2 海外厂商：IBM-关系型数据库领航者

- **关系型数据库领航者。**IBM成立于1911年，是全球最大的信息技术和业务解决方案公司。1970年，IBM实验室的EdgarFrankCodd发表了一篇题为《大型共享数据库数据的关系模型》论文，为关系型数据库技术奠定了理论基础；1973年，IBM启动验证关系型数据库系统的项目System R；1983年发布商业版数据库 DB2。
- IBM Db2 是一种云原生数据库，基于数十年的专业知识构建，为客户的任务关键型数据、分析和 AI 驱动应用程序及操作带来数据治理与安全性、低延迟事务和连续可用性。Db2支持大规模的事务和分析工作负载，提供了保持应用程序运行、存储和查询任何内容并简化开发的单一平台。
- Db2 11.5提供Db2 Community Edition、Db2 Community Edition、Db2 Advanced Edition三种版本，面向不同客户以及开发人员和合作伙伴社区。

图表：新的 Db2 版本及其相应的旧版本

先前版本	新版本	
Express-C Edition、Developer Community Edition	Db2 Community Edition	为企业提供全面的数据库解决方案，并提供许可指标以促进混合云部署。此解决方案提供数据仓库功能、事务功能和分析功能。可作为永久软件许可证和按月订阅两种形式
Advanced Workgroup Server Edition、Workgroup Server Edition、Direct Standard Edition	Db2 Standard Edition	面向小型或中型企业的高性能数据库，具有混合云部署的许可；部署大小受到限制
Advanced Enterprise Server Edition、Enterprise Server Edition、Developer Edition、Direct Advanced Edition	Db2 Advanced Edition	免费提供 Db2 的所有核心功能；面向开发人员和合作伙伴社区的入门级 Db2 数据服务器

3.1.3 海外厂商：Microsoft-数据库行业的领军企业之一

- 微软作为全球领先的软件、服务、设备和解决方案供应商，自1975年成立以来，一直致力于帮助个人和企业用户全面发挥科技潜能。
- Microsoft SQL Server是Microsoft开发设计的一个关系数据库智能管理系统，目前是世界主流数据库之一。
- SQL Server 2022提供Enterprise、Standard、Web、开发人员、Express Edition五个版本。相较于SQL Server 2019，SQL Server 2022有许多性能改进，包括增强的查询处理能力、更高的并发性能、更快的索引重建速度等；同时在安全方面，引入了如动态数据脱敏、敏感数据探测和分类、数据保护和加密等新的安全功能。

图表：SQL Server 2022 版本

版本	定义
Enterprise	作为高级产品/服务，SQL Server Enterprise Edition 提供了全面的高端数据中心功能，具有极高的性能和无限虚拟化，还具有端到端商业智能，可为任务关键工作负载和最终用户访问数据见解提供高服务级别
Standard	SQL Server Standard 版提供了基本数据管理和商业智能数据库，使部门和小型企业能够顺利运行其应用程序并支持将常用开发工具用于内部部署和云部署，有助于以最少的 IT 资源获得高效的数据库管理
Web	对于 Web 主机托管服务提供商（包括在 Azure 上的 IaaS 上选择 Web 版）和 Web VAP 而言，SQL Server Web 版本是一项总拥有成本较低的选择，可针对从小规模到大规模 Web 资产等内容提供可伸缩性、经济性和可管理性能
开发人员	SQL Server Developer 版支持开发人员基于 SQL Server 构建任意类型的应用程序。它包括 Enterprise 版的所有功能，但有许可限制，只能用作开发和测试系统，而不能用作生产服务器。SQL Server Developer 是构建和测试应用程序的人员的理想之选
Express Edition	SQL Server Express Edition 是入门级的免费数据库，是学习和构建桌面及小型服务器数据驱动应用程序的理想选择。它是独立软件供应商、开发人员和热衷于构建客户端应用程序的人员的最佳选择。如果您需要使用更高级的数据库功能，则可以将 SQL Server Express 无缝升级到其他更高端的 SQL Server 版本。SQL Server Express LocalDB 是 Express 版本的一种轻型版本，该版本具备所有可编程性功能，在用户模式下运行，并且具有快速零配置安装和必备组件要求较少的特点

3.1.4 海外厂商：AWS-全球领先的云数据库服务提供商

- AWS 提供云数据库服务超过10年，基于丰富的产品和配套工具，可满足用户不同数据类型和业务场景的各种需求。其云数据库具备高度可伸缩、完全托管、兼容性强、安全性高、性能优越、成本效益高、丰富的数据分析工具等特点。
- Amazon RDS是一个托管式服务的集合，可以简化在云中设置、运营和扩展数据库的过程。提供八种热门引擎以供选择：Amazon Aurora PostgreSQL 兼容版、Amazon Aurora MySQL 兼容版、RDS for PostgreSQL、RDS for MySQL、RDS for MariaDB、RDS for SQL Server、RDS for Oracle 和 RDS for Db2。

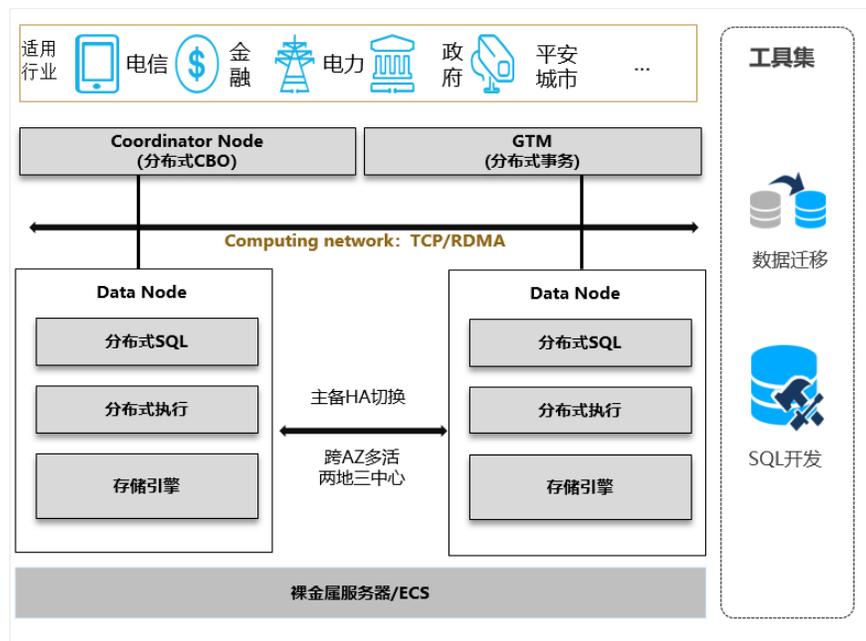
图表：AWS 提供最齐全的专用数据库

数据库类型	使用案例	AWS服务
关系	传统应用、企业资源规划 (ERP)、客户关系管理 (CRM)、电子商务	 Amazon Aurora  Amazon RDS  Amazon Redshift
键值	高流量 Web 应用程序、电子商务系统、游戏应用程序	 Amazon DynamoDB
内存中	缓存、会话管理、游戏排行榜、地理空间应用程序	 Amazon ElastiCache  Amazon MemoryDB for Redis
文档	内容管理、目录、用户配置文件	 Amazon DocumentDB (兼容 MongoDB)
宽列	用于设备维护、队列管理和路线优化的大规模工业应用程序	 Amazon Keyspaces
图形	欺诈检测、社交网络、建议引擎	 Amazon Neptune
时间序列	物联网 (IoT) 应用程序、开发运维、工业遥测	 Amazon Timestream
分类账	系统记录、供应链、注册、银行事务	 Amazon Ledger Database Services (QLDB)

3.2.1 国内厂商：华为-云数据库GaussDB

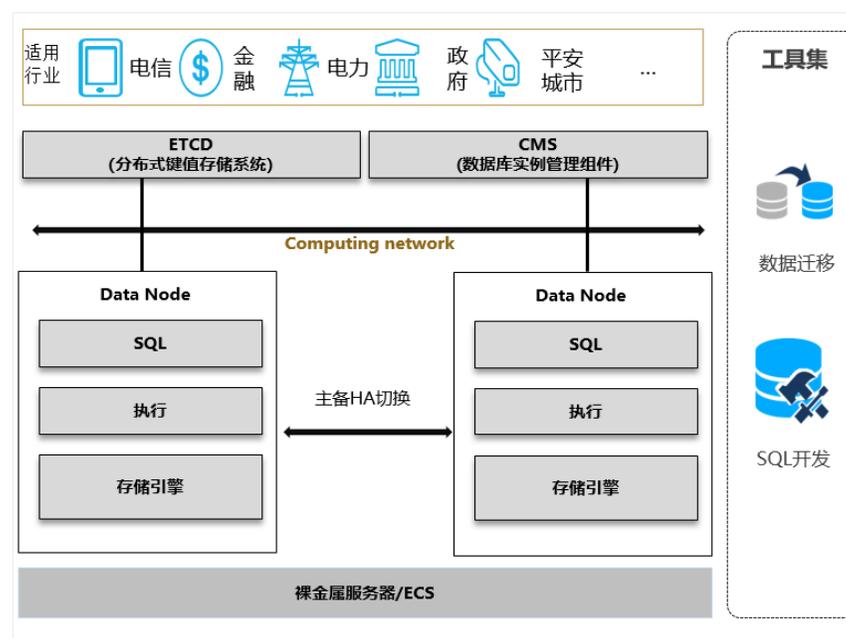
- 华为自主创新研发的分布式关系型数据库GaussDB支持分布式事务，同城跨AZ部署，支持1000+的扩展能力，PB级海量存储。GaussDB还具备多项关键能力，包括高可用性、高可靠性、高安全性，以及弹性的资源扩展能力。
 - 目前GaussDB支持分布式版和主备版实例。分布式形态支撑较大的数据量，且提供了横向扩展的能力，可以通过扩容的方式提高实例的数据容量和并发能力；主备版适用于数据量较小，且长期来看数据不会大幅度增长，对数据的可靠性，以及业务的可用性有一定诉求的场景。

图表：GaussDB分布式形态整体架构



资料来源：华为官网、中泰证券研究所

图表：GaussDB主备版实例整体架构

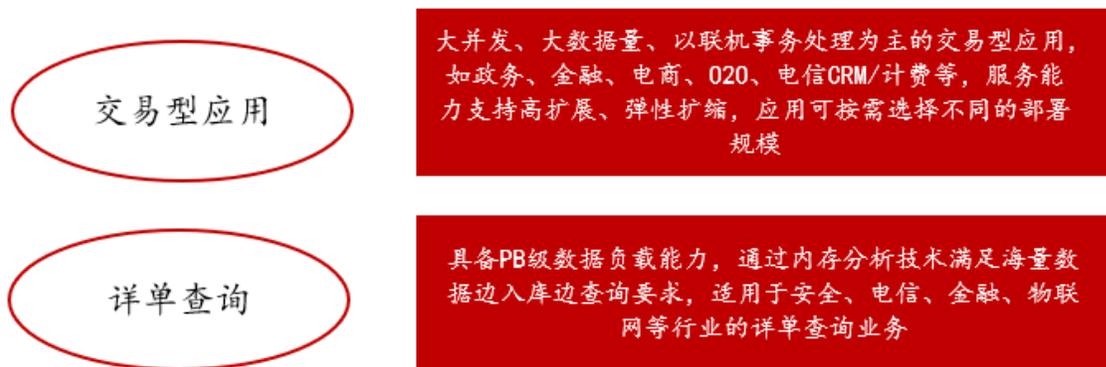


资料来源：华为官网、中泰证券研究所

3.2.1 国内厂商：华为-云数据库GaussDB

- GaussDB适用于大并发、大数据量、以联机事务处理为主的交易型应用，同时具备PB级数据负载能力，适用于安全、电信、金融、物联网等行业的详单查询业务。
- 基于20余年战略投入、软硬全栈协同所研发，GaussDB具备高安全、高性能、高可用、高弹性以及高智能等技术特点。

图表：GaussDB应用场景



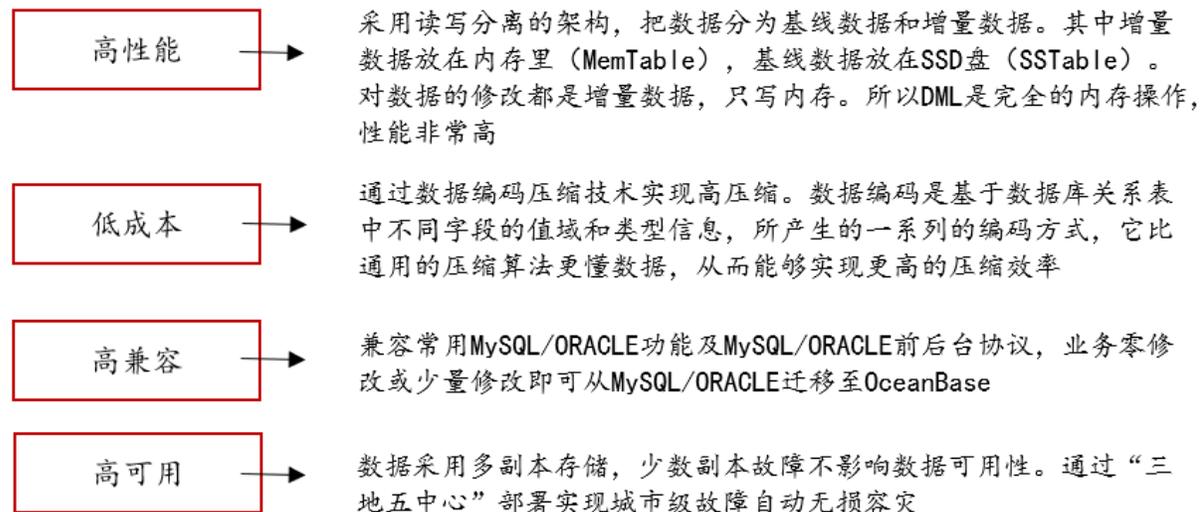
图表：GaussDB技术特点



3.2.2 国内厂商：阿里-分布式关系型数据库OceanBase

- OceanBase是由蚂蚁集团、阿里巴巴完全自主研发的分布式关系型数据库。OceanBase具有数据强一致、高可用、高性能、在线扩展、高度兼容SQL标准和主流关系型数据库、低成本等特点。
- OceanBase包括标准版（云盘）、标准版（本地盘）、标准版（历史库）和标准版（Key-Value）四种不同的系列。
- OceanBase目前可主要运用在交易支付、会员系统的多维度查询以及批处理系统等应用场景，已在南京银行、网商银行、支付宝、淘宝网等成功应用。

图表：OceanBase产品优势



图表：OceanBase产品系列

产品系列	特点
标准版（云盘）	覆盖规格和部署方案最全的系列；可以灵活进行套餐规格扩容和存储空间扩容
标准版（本地盘）	主要面向高稳定性需求场景，本地盘存储具有 10 时延低、性能高的特点，没有时延的长尾问题；标准版（本地盘）仅支持套餐规格扩容
标准版（历史库）	为节约成本，可选择购买标准版（历史库），将长期不使用的数据迁移至以文件形式存储的廉价存储设备上
标准版（Key-Value）	是在 OceanBase 关系型数据库（关系型+SQL）的基础上发展而来，为用户提供更加简单高效的非关系型 API 接口，支持多种非关系型数据模型的存储与处理，充分发挥 OceanBase 一体化架构的优势，满足用户不同场景下数据的多样性需求

3.2.2 国内厂商：阿里-云原生数据库 PolarDB

- PolarDB是阿里巴巴自研的新一代云原生数据库，在计算存储分离架构下，利用了软硬件结合的优势，提供具备极致弹性、高性能、海量存储、安全可靠的数据库服务。
- 目前100%兼容MySQL和PostgreSQL生态，高度兼容Oracle语法。
- PolarDB共有三个引擎，分别为PolarDB MySQL版、PolarDB PostgreSQL版、PolarDB分布式版。
- 目前已在心动网络、数云、赛盒科技、特步、启迪公交等成功应用。

图表：PolarDB

引擎	数据库生态	兼容性	产品架构	输出形态	应用场景
PolarDB MySQL版	MySQL	100%兼容MySQL	Share Storage, 计算存储分离	公共云、专有云企业版	MySQL生态下的云原生数据库
PolarDB PostgreSQL版	PostgreSQL、Oracle	100%兼容 PostgreSQL、高度兼容Oracle	Share Storage, 计算存储分离	公共云、专有云企业版、DBStack	PostgreSQL生态下的云原生数据库
PolarDB分布式版	MySQL	标准版100%兼容MySQL、企业版高度兼容MySQL	Share Nothing, 分布式	公共云、专有云企业版、DBStack	MySQL生态下的大规模数据、超高并发应用

资料来源：阿里云、中泰证券研究所

3.2.3 国内厂商：达梦数据-国产数据库市场领军企业

- 达梦数据正式成立于2000年，是国内领先的数据库产品开发服务商，数据库基础软件产业发展的关键推动者。达梦数据40余年磨一剑，坚持原始创新、独立研发的技术路线，2022年实现中国数据库管理系统国产数据库市场占有率排名第一。
- 目前公司已经成功地为国家发展改革委员会、外交部等各级党政机关，以及国家电网、中航信、中国建设银行、交通银行、中国人寿等重要客户实现国产数据库管理系统的应用，服务领域广泛，产品实力过硬。

图表：达梦数据库产品



资料来源：公司官网、中泰证券研究所

图表：达梦数据客户案例



资料来源：公司官网、中泰证券研究所

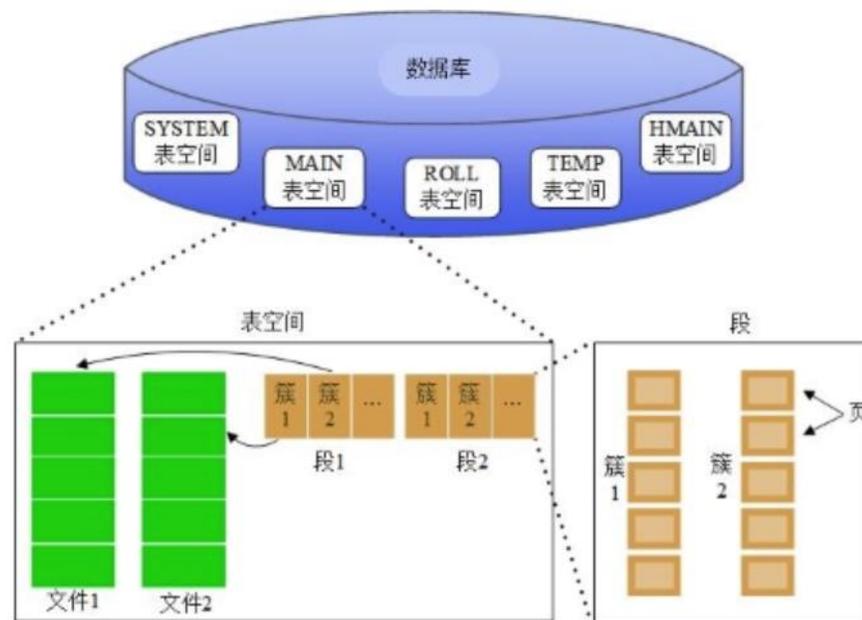
3.2.3 国内厂商：达梦数据-国产数据库市场领军企业

- 在 DM 数据库内部，所有的数据文件组合在一起被划分到一个或者多个表空间中，所有的数据库内部对象都存放在表空间中。且表空间被划分为段、簇和页，使得 DM 数据库能够更加高效地控制磁盘空间的利用率。
- 达梦数据库管理系统DM8：新一代大型通用关系型数据库，全面支持 SQL 标准和主流编程语言接口/开发框架。行列融合存储技术，在兼顾 OLAP 和 OLTP 的同时，满足 HTAP 混合应用场景。

图表：DM8 产品的构成

版本	特点
DM Standard Edition 标准版	DM 标准版是为政府部门、中小型企业及互联网 / 内部网应用提供的数据库管理和分析平台。它拥有数据库管理、安全管理、开发支持等所需的基本功能
DM Enterprise Edition 企业版	DM 企业版是伸缩性良好、功能齐全的数据库，无论是用于驱动网站、打包应用程序，还是联机事务处理、决策分析或数据仓库应用，DM 企业版都能作为专业的服务平台。DM 企业版支持多 CPU，支持 PB 级海量数据存储和大量的并发用户，并为高端应用提供了数据守护、大规模并行处理 MPP、共享存储集群、分布计算集群等高可靠性、高性能和高扩展的数据库解决方案，完全能够支撑各类企业应用
DM Security Edition 安全版	DM 安全版拥有企业版的所有功能，并重点加强了其安全特性，引入强制访问控制功能，采用数据库管理员 (DBA)、数据库审计员 (AUDITOR)、数据库安全员 (SSO)、数据库对象操作员 (SYSDBO) 四权分立安全机制，支持 KERBEROS、操作系统用户等多种身份鉴别与验证，支持透明、半透明等存储加密方式以及审计控制、通信加密等辅助安全手段，使 DM 安全级别达到国家信息技术安全评估准则 ISO/IEC 15408 EAL4+ 级水平，适合于对安全要求较高的政府或企业选用

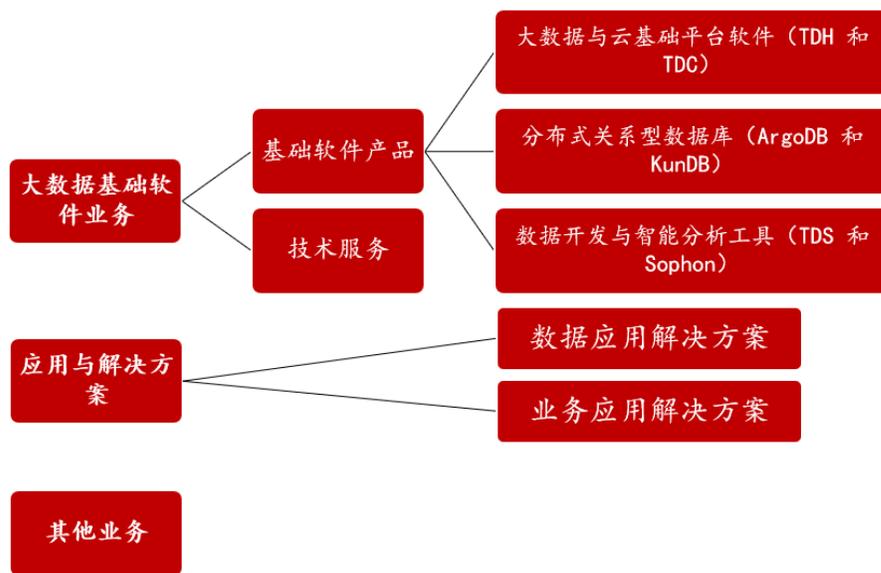
图表：DM 逻辑存储结构



3.2.4 国内厂商：星环科技-企业级大数据基础软件开发商

- 星环科技是国内领先的大数据基础软件开发商，主要提供大数据与云基础平台、分布式关系型数据库、数据开发与智能分析工具三大类基础软件产品及配套服务，助力客户实现数字化转型。
- 公司以大数据基础平台软件为切入点，围绕数据的集成、存储、治理、建模、分析、挖掘和流通等逐步发布相关新的软件产品来完善公司的大数据产品矩阵。

图表：公司产品分布情况



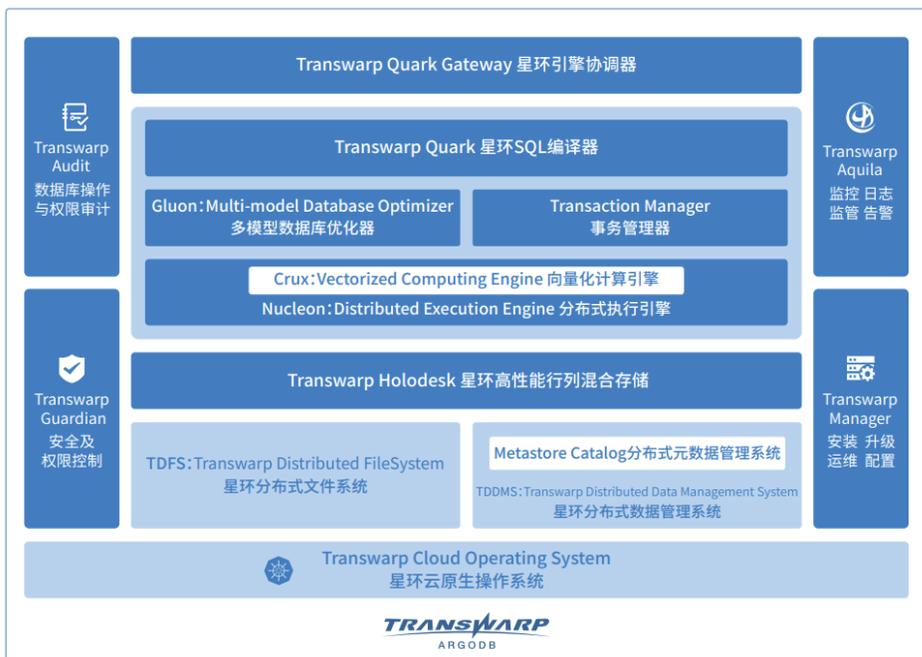
图表：星环科技产品体系发展历程



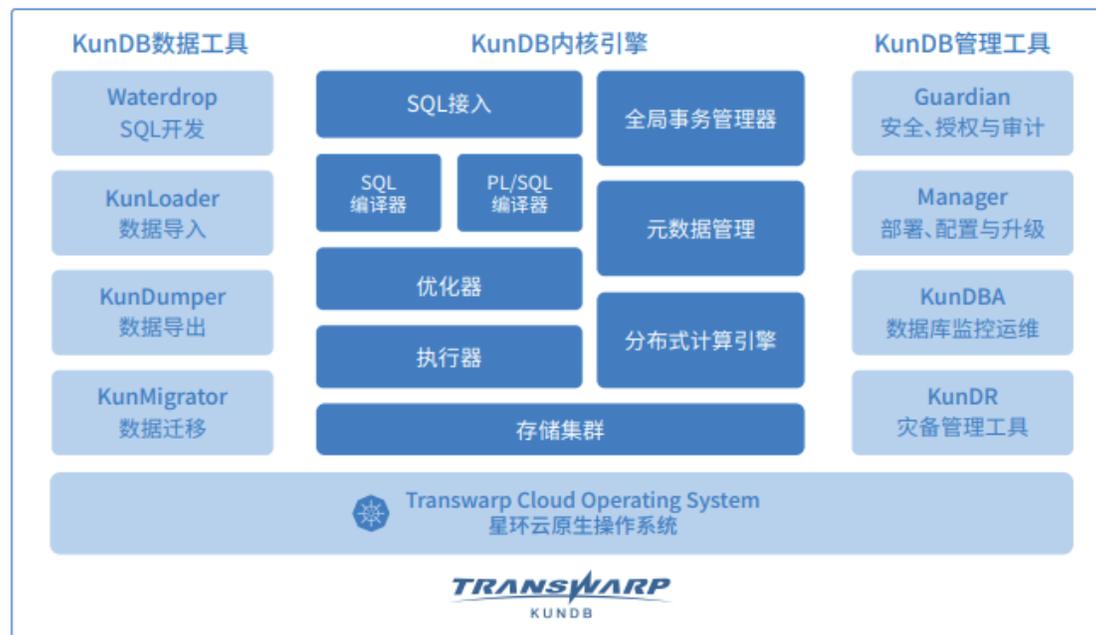
3.2.4 国内厂商：星环科技-企业级大数据基础软件开发商

- 公司在基础软件产品领域积累多年，采用新一代的分布式技术，自主研发了分布式分析型数据库 ArgoDB 和分布式交易型数据库 KunDB。KunDB与ArgoDB均是面向云原生架构设计的新一代数据库，根据工信部电子第五研究所代码扫描测试报告，ArgoDB代码自主率（行数）超过90%。
- ArgoDB和KunDB在银行、运营商、能源等行业领域均有落地案例，公司持续通过市场拓展形成了广泛的客户基础。

图表：ArgoDB



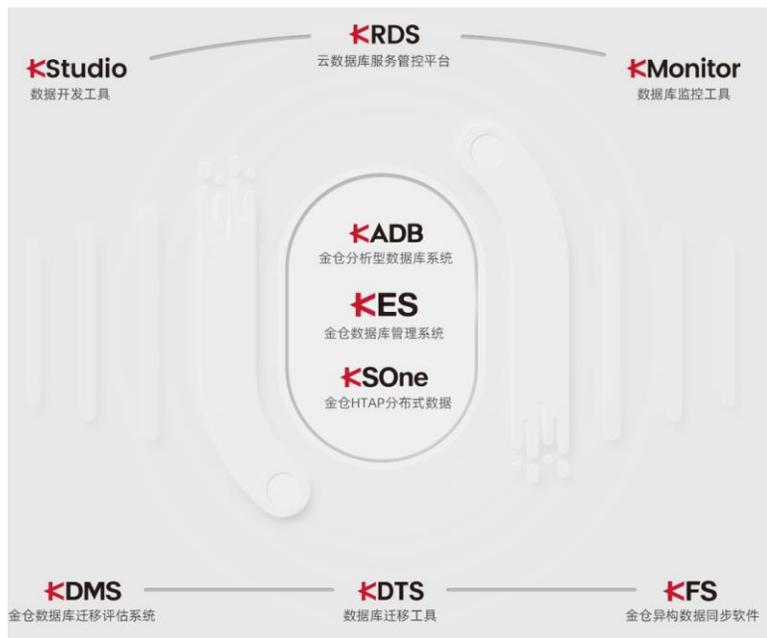
图表：KunDB



3.2.5 国内厂商：人大金仓-KingbaseES

- 人大金仓专注数据库领域20余载，其数据库产品、服务及解决方案体系，广泛服务于电子政务、国防军工、能源、金融、电信等60余个重点行业和关键领域，累计装机部署超百万套。
- 目前数据库产品主要包括KingbaseES、KSSOne 以及KingbaseAnalyticsDB。

图表：人大金仓数据库产品与服务



资料来源：公司官网、中泰证券研究所

图表：人大金仓数据库产品

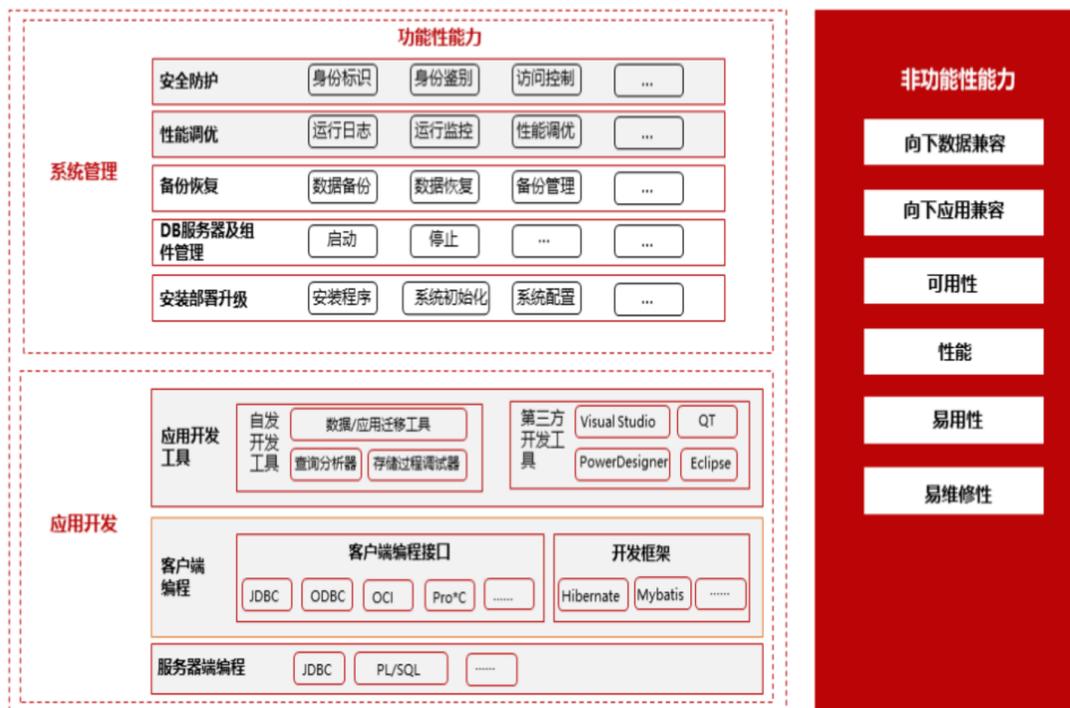


资料来源：公司官网、中泰证券研究所

3.2.5 国内厂商：人大金仓-KingbaseES

- 核心产品金仓数据库管理系统 KingbaseES是企业级关系型数据库管理系统，主要面向政府、金融、电信、电力等行业，具有大型通用、“三高”（高可靠、高性能、高安全）、“三易”（易管理、易使用、易扩展）、运行稳定等特点，是唯一入选国家自主创新产品目录的数据库产品，也是国家级、省部级实际项目中应用最广泛的国产数据库产品。

图表：KingbaseES整体能力结构



图表：KingbaseES产品优势



3.2.6 国内厂商：PingCAP-分布式关系型数据库 TiDB

- PingCAP 成立于 2015 年，是一家企业级开源分布式数据库厂商。PingCAP 公司自主设计、研发的开源分布式关系型数据库 TiDB 是一款同时支持在线事务处理与在线分析处理 (HTAP) 的融合型分布式数据库产品，具备水平扩容或者缩容、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 MySQL 生态等重要特性。

图表：TiDB



资料来源：公司官网、中泰证券研究所

图表：PingCAP 用户

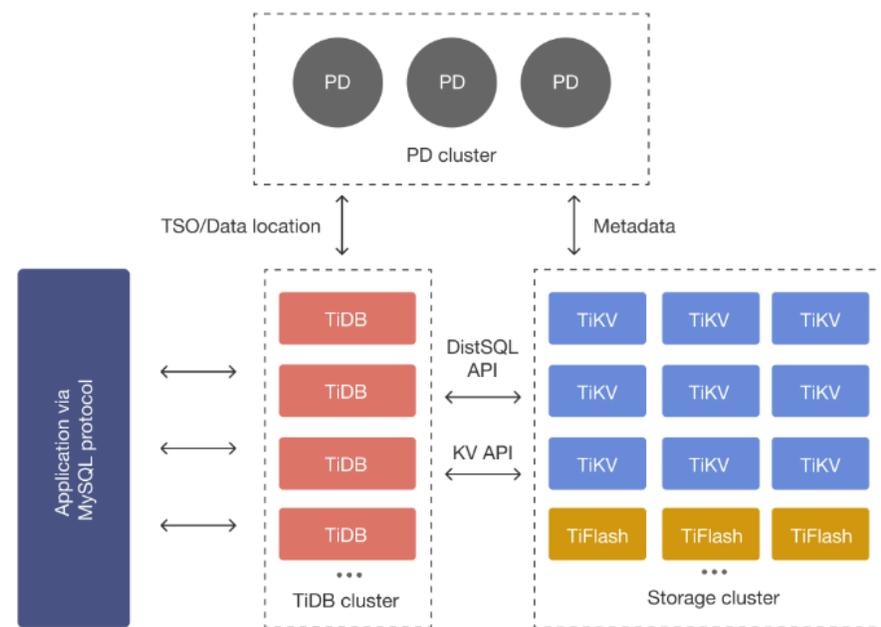


资料来源：公司官网、中泰证券研究所

3.2.6 国内厂商：PingCAP-分布式关系型数据库 TiDB

- 在内核设计上，TiDB 分布式数据库将整体架构拆分成了多个模块，各模块之间互相通信，组成完整的 TiDB 系统。
- 与传统的单机数据库相比，TiDB 具有以下优势：
 - 纯分布式架构，拥有良好的扩展性，支持弹性的扩缩容；
 - 支持 SQL，对外暴露 MySQL 的网络协议，并兼容大多数 MySQL 的语法，在大多数场景下可以直接替换 MySQL；
 - 默认支持高可用，在少数副本失效的情况下，数据库本身能够自动进行数据修复和故障转移，对业务透明；
 - 支持 ACID 事务；
 - 具有丰富的工具链生态，覆盖数据迁移、同步、备份等多种场景。

图表：TiDB架构图



资料来源：公司官网、中泰证券研究所



4

投资建议与风险提示

投资建议

- 数据库作为计算机系统的核心基础软件，同时也是信创产业的关键环节，我们认为在政策、新技术等因素的扶持下，国内数据库产业有望持续受益。当前时点，我们持续看好国内数据库产业的投资机遇，建议关注达梦数据、星环科技、太极股份（人大金仓）。

风险提示

- 技术发展不及预期；
- 行业竞争加剧风险；
- 政策落地不及预期；
- 研究报告使用的公开资料可能存在信息滞后或更新不及时的风险。

重要声明

- 中泰证券股份有限公司（以下简称“本公司”）具有中国证券监督管理委员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。
- 本报告基于本公司及其研究人员认为可信的公开资料或实地调研资料，反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响。本公司力求但不保证这些信息的准确性和完整性，且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，可能会随时调整。本公司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。本报告所载的资料、工具、意见、信息及推测只提供给客户作参考之用，不构成任何投资、法律、会计或税务的最终操作建议，本公司不就报告中的内容对最终操作建议做出任何担保。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。
- 市场有风险，投资需谨慎。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。
- 投资者应注意，在法律允许的情况下，本公司及其本公司的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并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。本公司及其本公司的关联机构或个人可能在本报告公开发布之前已经使用或了解其中的信息。
- 本报告版权归“中泰证券股份有限公司”所有。事先未经本公司书面授权，任何机构和个人，不得对本报告进行任何形式的翻版、发布、复制、转载、刊登、篡改，且不得对本报告进行有悖原意的删节或修改。