

# 中国软件根技术发展白皮书


White Paper on software root technology development in China

( 数据库册 )

2023版



**CSIA** 中国软件行业协会  
China Software Industry Association

 **openGauss**

中国软件行业协会 编

# 中国软件根技术发展白皮书

（数据库册）

中国软件行业协会

## 目 录

前言 .....	1
一、数据库概述 .....	3
（一）数据库相关概念 .....	3
（二）数据库分类 .....	5
（三）数据库架构标准 .....	6
（四）数据库产业链 .....	8
二、全球数据库竞争格局和发展态势 .....	9
（一）全球数据库行业竞争格局 .....	9
（二）主流数据库厂商产品分析 .....	10
1.Oracle .....	10
2.IBM .....	11
3.Microsoft .....	12
4.Amazon .....	12
5.Google .....	13
（三）发展现状 .....	14
1. 数据库是数字经济发展的基础技术和设施 .....	14
2. 美国数据库厂商长期垄断全球数据库市场 .....	14
3. 数据库技术路径分化带来市场分裂与竞合 .....	15
4. “开源 + 云”已经成为数据库革新的关键力量 .....	16
（四）全球数据库发展新态势 .....	16
1. 数据库和硬件协同成为新兴竞争力量 .....	16
2. 数据库领域的投资热情正在逐年升温 .....	16
3. 中国数据库迎来技术积累爆发增长期 .....	17
4. 技术收敛是国产数据库突围有效路径 .....	18

三、国产数据库发展状况 .....	19
（一）发展现状 .....	19
（二）市场规模 .....	20
（三）国产数据库主要厂商产品分析 .....	21
1. 达梦数据库 .....	22
2. 南大通用 .....	22
3. 神舟通用 .....	23
4. 人大金仓 .....	23
5. 华为公司 .....	24
6. 阿里云 .....	24
7. 腾讯云 .....	25
8. 平凯星辰 .....	25
9. 云和恩墨 .....	26
10. 海量数据 .....	27
（四）国产数据库在重点行业应用分析 .....	27
（五）国产数据库厂商积极布局生态建设 .....	28
四、主流数据库技术发展分析 .....	30
（一）关系型数据库 .....	30
1. 技术特点 .....	30
2. 主流产品 .....	31
3. 发展趋势 .....	31
（二）非关系型数据库 .....	31
1. 技术特点 .....	31
2. 主流产品 .....	32
3. 发展趋势 .....	33
五、国产数据库发展面临的形势、问题及机遇 .....	36
（一）面临的主要形势 .....	36

(二) 存在的主要问题 .....	38
1. 产品种类繁多，竞争能力不足 .....	38
2. 技术基础薄弱，专业人才匮乏 .....	38
3. 安全性能存忧，用户使用存虑 .....	38
4. 产品标准各异，缺乏行业标准 .....	39
5. 产业资源分散，生态建设滞后 .....	40
(三) 发展机遇 .....	40
1. 利好政策出台释放巨大红利 .....	40
2. 数字化激发数据库市场潜力 .....	41
3. 数据库国产化替代空间巨大 .....	41
4. 行业技术实力积累整体增强 .....	41
六、国产数据库发展基本判断 .....	43
(一) 国产数据库技术路线收敛势在必行 .....	43
(二) 国产数据库厂商应重视用户的需求 .....	43
(三) 存算分离 + 多读多写成为发展新趋势 .....	43
(四) 开源是国产数据库发展的有效路径 .....	44
七、建立以 openGauss 为代表的国产数据库体系 .....	45
(一) 内核创新研究，打造坚实数据底座 .....	45
(二) 联合产业资源，共建资源池化架构 .....	46
(三) 满足行业诉求，打造关基行业标杆 .....	47
(四) 打造开源社区，共建技术开发生态 .....	48
(五) 走进高校课堂，培养专业技术人才 .....	49
八、发展对策建议 .....	50
(一) 引导数据库技术路线收敛提升竞争能力 .....	50
(二) 强化存算分离 + 共享存储标准体系建设 .....	50
(三) 出台专项政策支持首台套和规模化应用 .....	50
(四) 强化数据库技术创新集聚全产业链资源 .....	51
(五) 集中支持开源社区加大产业链生态建设 .....	51



## 前 言

党的二十大报告明确提出，坚持创新在我国现代化建设全局中的核心地位，加快实现高水平科技自立自强，加快建设科技强国，并对完善科技创新体系、加快实施创新驱动发展战略等作出了系列重要部署。

当前，在错综复杂的国际环境和新技术快速迭代的背景下，基础软件根技术已经成为国家战略科技竞争的制高点。要从根本上摆脱对单一技术供应链的依赖，必须坚持原始技术创新，坚定不移地摸索出一条独立自主的软件根技术产业体系发展之路。数据库是业界公认的最复杂、跨技术领域最多的基础软件，随着数字化的深入推进，数据库在技术架构、数据管理、数据安全创新，以及数据科学基础研究、生态建设、“产学研用”深入合作等方面都面临着巨大的挑战。

现阶段，我国数据库行业市场前景广阔，整体呈现出稳步上升的发展势头，经过多年的技术沉淀和市场洗礼，正在经历从“引进”到“自研”，由“可用”向“好用”进阶的关键节点，这为国产数据库发展提供了良好的成长土壤。国产数据库要实现突破式发展，需要与行业发展变化需求相协同，也需要产业界形成合力持续推动原始创新，还需要构建起完善健康的产业生态体系。

我们倡导国产数据库技术收敛路线，鼓励产业上下游生态链资源联合起来共同打造具有国际竞争力的国产数据库技术、产品和解决方案，推动我国数据库产业高质量发展。



## 一、数据库概述

数字化正在改变世界万物，而数字世界的底层，则是庞大的不断产生、汇集、运算的数据。实现数字化的前提是数据化，而数据的可存、可用、可管、可流动、可分析，都离不开数据库。

数据库作为信息系统的核心，在计算机中扮演着承上启下的作用，向下实现调动计算、网络、存储等基础资源，向上完成支撑各种应用软件需求，是计算功能得以实现的基础设施底座。数据库的发展不仅影响着计算机技术的发展进程，也推动着经济社会和数字中国的持续进步。

### （一）数据库相关概念

**数据（Data）**：数据库中存储的基本对象，是描述实物的符号记录。描述实物的符号可以是数字，也可以是文字、图形、图像、声音、语言等，数据有多重表现形式，它们都可以经过数字化后存入计算机。

**数据库（DB）**：长期存储在计算机内的、有组织的、可共享的、统一管理的大量数据的集合。数据库中的数据按一定的数据模型组织、描述、存储，其特点是冗余度小，具有较高的数据独立性、共享性和易扩展性。

<b>专栏 1：数据模型是数据库系统的核心和基础</b>
数据模型是对现实世界数据特征的抽象。通俗地讲数据模型就是对现实世界的模拟映射。



→数据模型分类:

**概念模型**，也称信息模型。按用户的观点来对数据和信息建模，用于数据库设计。

**逻辑模型**: 包括网状模型、层次模型、关系模型、面向对象数据模型、对象关系模型、半结构化数据模型等。

**物理模型**: 对数据最底层的抽象，描述数据在系统内部的表示方法和存取方法、或者在磁盘或磁带上的存储方式和存取方法。

→数据模型的组成要素:

严格定义的一组概念的集合。精确描述了系统的静态特性、动态特性和完整性约束条件。

数据库管理系统 (DBMS) : 一种用于管理、维护和操纵数据库的大型软件系统。它通过提供统一的管理和控制，保证数据库的安全性和完整性，满足多个应用程序和用户在同一时刻或不同时刻建立、修改和查询数据库。DBMS 主要由管理模块、控制模块和各种数据库操作语言组成。

专栏 2: 数据库管理系统的主要特性

**独立性**: 通过 DBMS，不同应用程序和用户可以在同一数据库下操作，各数据之间互相独立，互不干扰。

**共享性**: 数据库系统可以进行多用户共享，多个用户或应用程序能够同时访问同一数据库，进行数据的查询和修改。

**集中性**: 通过 DBMS，可以将不同的数据集中到同一数据库中进行管理和维护，提高了数据的管理效率和数据间的联系性。

**安全性**: 通过 DBMS，可以实现对数据的安全保护，数据访问控制和数据备份等功能。

**一致性**: 通过 DBMS，不论是在单用户还是多用户的情况下，数据的一致性能得到保障。

## （二）数据库分类

历经 70 年的不断发展，数据库已经衍生出多种技术路径和方向。为了更清楚地掌握数据库之间的特性差异，数据库产品可根据不同标准进行分类，常见的分类依据包括数据模型、架构模型、业务负载特征、部署方式、适配硬件架构等，其中数据结构类型是最常用的分类标准。

- 根据数据模型分类：关系型数据库、非关系型数据库。
- 根据架构模型分类：集中式数据库、分布式数据库。
- 根据业务负载特性分类：事务型数据库（OLTP）、分析型数据库（OLAP）、混合型数据库（HTAP）。

现阶段，最常见的数据库模型是关系型数据库和非关系型数据库，国内超过 90% 的数据库市场选择使用关系型数据库，故本文将以数据结构为重点对数据库进行研究。

**关系型数据库（RDB）：**一种采用关系模型来组织数据的数据库，而关系模型是由二维表及其联系组成的数据组织。在关系型数据库中，对数据的操作几乎全部建立在一个或多个关系表格上，通过对这些关联表的表格进行分类、合并、连接或选取等运算来实现数据的管理。

专栏 3：关系型数据库的优缺点
<p>→主要优点：</p> <p>使用表结构，格式一致，易于维护；使用通用的 SQL 语言操作，使用方便，可用于复杂查询；数据存储在磁盘中，安全性高。</p> <p>→主要缺点：</p> <p>读写性能比较差，不能满足海量数据的高效率读写；不节省空间，</p>

建立在关系模型上，要遵循某些规则，比如数据中某字段值即使为空仍要分配空间；固定的表结构，灵活度较低。

非关系型数据库（NOSQL）：一种数据结构化存储方法的集合，可以是文档或者键值对等。作为关系型数据库的一个补充，能在特定场景和特点问题下发挥高效率和高性能。非关系型数据库可以分为 **key-value** 型（针对高性能并发读写场景）、文档型（针对海量数据访问场景）、列式数据库、图形数据库四种类型。

#### 专栏 4：非关系型数据库的优缺点

##### →主要优点：

存储数据的格式可以是 **key-value** 形式、文档形式、图片形式等，使用灵活，应用场景广泛；速度快，效率高，NoSQL 可以使用硬盘或者随机存储器作为载体，而关系型数据库只能使用硬盘；海量数据的维护和处理非常轻松；具有扩展简单、高并发、高稳定性、成本低廉的优势；可以实现数据的分布式处理。

##### →主要缺点：

暂时不提供 SQL 支持，学习和使用成本较高；没有事务处理，没有保证数据的完整性和安全性。适合处理海量数据，但是不一定安全；功能没有关系型数据库完善。

### （三）数据库架构标准

产业数字化加速产生了诸如图、流、时序和地图空间等多种数据类型的海量数据，计算机体系架构正从以 CPU 为中心的架构向多样性算力协同的对等架构演进。在多模数据和多样性算力双轮驱动下，数据库架构需要与时俱进，有效利用多样

性算力，进行资源的集约化管理和调度，实现多模数据的高效处理和数据价值挖掘。

数据库组件化架构标准以 SQL 标准作为应用接入的标准界面，并进一步往下延伸，分层定义分发、执行、存储引擎的交互接口，实现数据库架构的可组装可演进，打造面向未来的数据库组件化技术生态。

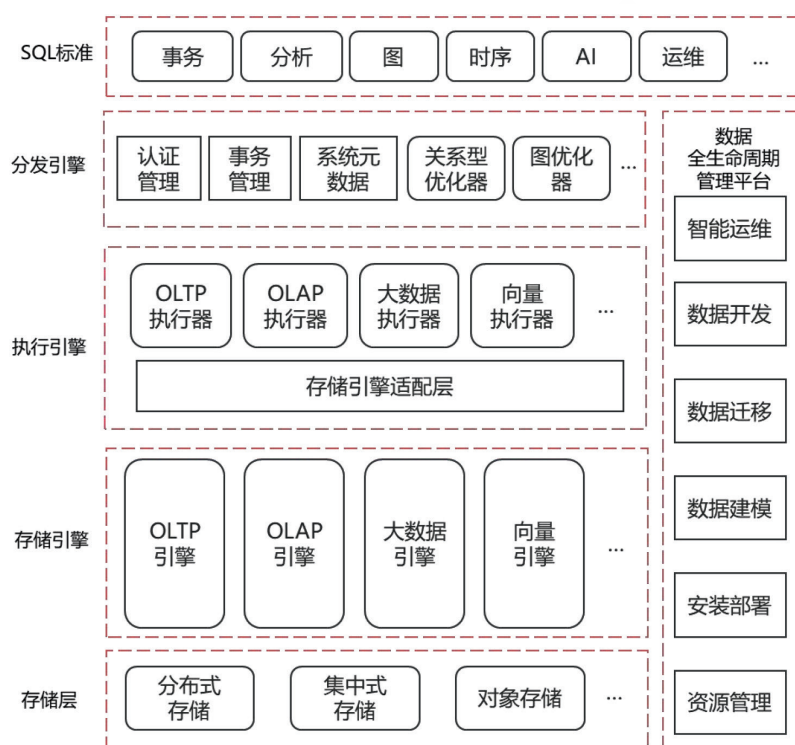


图 1 数据库组件化架构标准

图 1 所示的数据库组件化架构标准中，SQL 标准定义了全场景 SQL 标准；分发引擎实现 SQL 语句到执行计划的生成；执行引擎实现高效数据处理；存储引擎实现对多模数据的管理；存储层实现数据的透明化企业级存储能力；数据全生命周期管

理平台实现智能运维、集群管理等公共能力。

组件化架构标准在每一层都提供了扩展协同接口，以支持多样性的 SQL 接口、多模优化器、多模执行器、多模存储引擎。同时在部署形态上，每一层既可以单机部署，也可以集群部署，从而实现每一层基于业务负载的资源弹性伸缩。

#### （四）数据库产业链

数据库与芯片、操作系统并列为全球技术三大件，也是企业 IT 系统必不可少的核心技术。从图 2 所示的数据库上下游产业链图谱中可以看到，产业链上游主要是以 CPU 厂商、服务器厂商、网络设备厂商、存储厂商组成的网络和硬件厂商；中游主要是以数据库、操作系统和中间件等组成的基础软件厂商；下游主要是以各行业的集成商、开发商组成的应用开发厂商。

从产业链发展视角来看，数据库处在产业链中游枢纽位置，向上承接各类硬件、网络服务资源，向下提供数据存储、管理与分析服务，因此，数据库的地位和作用突显重要。

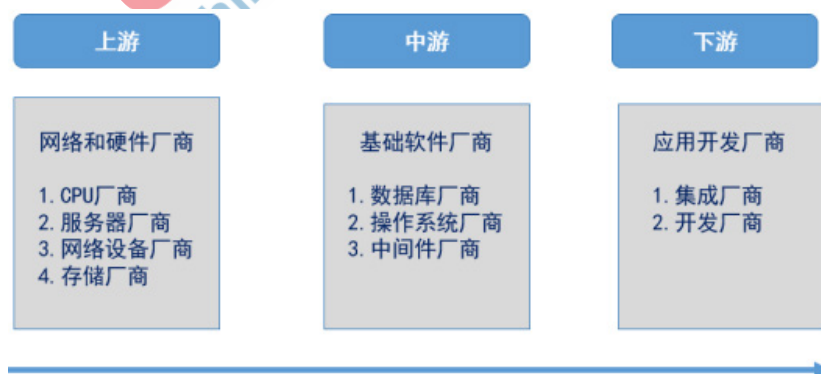


图 2 数据库上下游产业链图谱



## 二、全球数据库竞争格局和发展态势

数据库最早起步于 1964 年，先后经历了两次发展热潮。20 世纪 80 年代随着关系型数据库的理论突破和技术创新，全球数据库行业迎来第一次发展热潮。2000 年以来，随着移动互联网的发展兴起，数据库的应用场景越来越丰富，推动了全球数据库在 2010-2019 年进入第二次发展高峰期，大批新兴企业竞相涌现，全球有将近 50% 的数据库企业相继在这一时期成立。进入云智能时代，全球数据库产业正在进入第三次爆发式发展周期，整体呈现百花齐放、百家争鸣的发展态势。

### （一）全球数据库行业竞争格局

全球数据库行业主要有 Oracle、IBM、Microsoft、Amazon、Google 等主流厂商。关系型数据库依然是现阶段市场成熟度最高、应用范围最广的数据库产品，占据全球约 80% 的市场份额，Oracle、IBM、Microsoft 是行业的佼佼者。

近年来，随着云数据库产品异军突起，Amazon、Google 等国际云数据库厂商快速崛起，市场份额快速增长。此外，在商业化和开源并行发展的背景下，也涌现出 MongoDB、Snowflake 等一批以技术起家的数据库独立厂商，在各自的强势领域深耕细作，占据了一席之地。

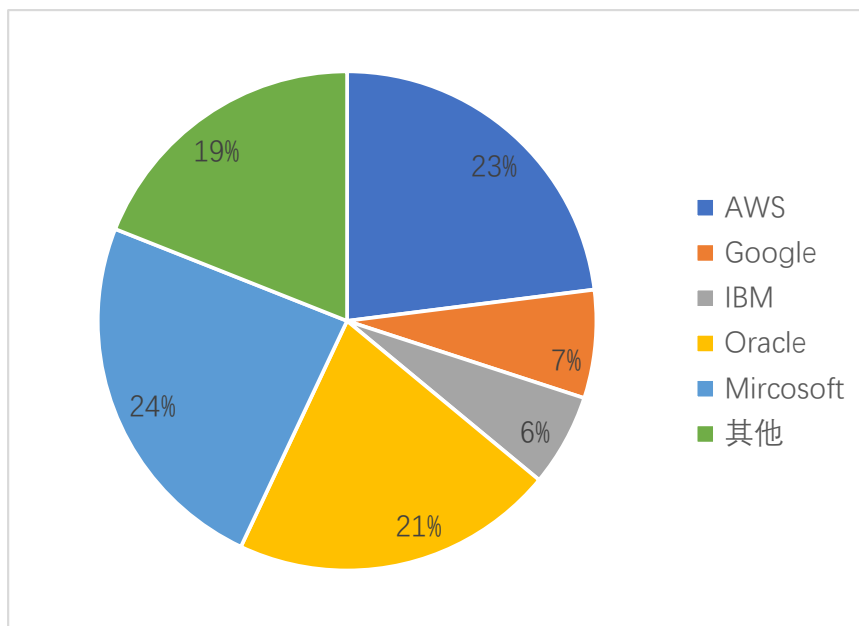


图 3 2022 年全球数据库竞争格局

## (二) 主流数据库厂商产品分析

### 1. Oracle

甲骨文(Oracle)是一家总部位于美国加州的跨国科技公司，是全球知名的数据库管理系统供应商，作为全球领袖级别的企业级数据管理解决方案供应商，Oracle 提供的数据库产品和系统已经逐渐成为业界标准，应用领域和范围十分广泛。Oracle 旗下的数据库产品覆盖了不同的领域，包括企业级应用系统、小型应用程序、云环境等，产品形态也多样化，适合不同类型的应用需求。

主要产品：Oracle Database、Oracle MySQL、Oracle NoSQL Database、Oracle Berkeley DB、Oracle Rdb、Oracle TimesTen 等。

主要应用领域：甲骨文数据库产品长期以来一直是全球市

场份额最大的产品之一。甲骨文数据库以其稳定性、安全性、易用性和强大的性能著称，被广泛应用于全球各行各业。

专栏 5：甲骨文发展脉络与启示
<p>1977 年：拉里·埃里森、鲍勃·明顿和爱德华·奥茨共同创建了甲骨文公司。专注于开发和销售数据库管理系统。</p> <p>1986 年：甲骨文公司在纳斯达克上市，成为公众持股的公司。</p> <p>1990 年：甲骨文开始转型，从一个主要提供数据库管理系统的公司变为一个提供全方位企业级解决方案的公司。甲骨文发布了第一个 ERP 产品，使其开始在企业应用软件领域占据一席之地。</p> <p>2004 年：甲骨文提出并开始实施并购策略，先后收购了 PeopleSoft、Siebel Systems 等公司，进一步扩展了产品和服务的范围。</p> <p>2010 年：甲骨文完成对 Sun Microsystems 的收购，标志着甲骨文进入了硬件领域，并开始提供硬件和软件一体化的解决方案。</p> <p>2012 年：甲骨文发布了其首个公有云服务，开始转型为云服务提供商。</p> <p>近年来，甲骨文在云服务、人工智能和大数据等新兴领域进行了大量的研发和投资，以适应不断变化的市场需求。</p> <p>启示：甲骨文一直在不断地调整和改变自身的发展策略，以适应不断变化的市场环境。甲骨文始终保持着数据库技术创新优势，积极应对各种挑战，从而保持了其在全球科技领域的领导地位。</p>

2.IBM

IBM 公司是全球领先的计算机科技公司，也是数据库领域的佼佼者。IBM 数据库是一款功能强大、安全可靠、适应性强的数据库产品，广泛应用于多个行业和领域中，可以为企业带来高效、稳定、成本优化的数据管理能力。

主要产品：IBM DB2、IBM Informix、IBM Cloudant 等。

主要应用领域：（1）企业级应用，支持企业级高吞吐量的应用开发，如金融交易系统、数据仓库等；（2）物联网应用，支持物联网的数据存储和分析处理，如传感器数据、智能家居、工业自动化等；（3）移动应用，支持移动应用的后台数据存储和处理，如社交应用、移动支付等；（4）云环境，支持云环境下的多租户存储和处理，如 PaaS、SaaS 等云服务模式。

### 3.Microsoft

微软公司（Microsoft）是一家美国跨国科技企业，主要以研发、制造、授权和提供广泛的电脑软件服务业务为主。随着互联网和数字化时代的到来，数据管理越来越成为各个企业的重要课题。微软公司顺势推出了一系列高效稳定的数据库软件，助力企业通过从海量数据中获得有价值的信息，增强企业决策的数据支撑能力。

主要产品：SQL Server、Azure SQL、Azure Cosmos DB、Access 等。

主要应用领域：微软数据库软件作为高效稳定的数据管理工具，广泛应用于银行、电力、通讯、零售等行业，为企业在数据管理领域提供着强大的支持。

### 4.Amazon

近 10 年以来，全球数据库市场加速变革，云数据库尤其是云原生数据库成为整个数据库市场的关键变量。亚马逊公司（Amazon）作为全球云原生数据库的领导者，其基于云原生数据库，构建起涵盖数据存储、查询、数据分析、机器学习、

商业智能、编目与治理的端到端的数据战略。

主要产品：Amazon Aurora Serverless v2、Amazon DynamoDB、Amazon Timestream、Amazon QLDB、Amazon Keyspace、Amazon Neptune Serverless 等。

主要应用领域：AWS 数据库系统具有高可用性、高安全性、易于管理等优势，在传统应用、ERP、CRM、电子商务、游戏、工业、社交网络、物联网、银行等各种应用场景中得到了广泛的应用。

#### 专栏 6：亚马逊成为全球数据库管理系统领导者

在 Gartner 公布的 2022 年全球数据库管理系统的市场份额排名中，作为纯云厂商的亚马逊云科技，超越了老牌传统数据库厂商甲骨文和微软，首次位居第一，市场份额达到 25.3%。

亚马逊云科技已经连续 8 年位居 Gartner 云数据库管理系统领导者象限，更是在愿景完备性和执行能力两个维度都超出所有对手。

### 5.Google

谷歌公司（Google）是全球最大的互联网公司之一，谷歌数据库为高性能和大规模的数据处理、存储和管理提供了出色的解决方案。

主要产品：Google Cloud SQL、Google Cloud Spanner、AlloyDB for PostgreSQL、BigQuery 等。

主要应用领域：广泛应用在金融、IT、生物医药、物流、电商零售、物联网、政府等领域。



### （三）发展现状

#### 1. 数据库是数字经济发展的基础技术和设施

近年来，数字技术加速融入经济社会发展的各个领域和全过程，进入数字经济时代的重要特征是生产力的全面数字化，产生的数据量也呈现爆炸式增长，数据要素价值日益显现，海量数据需要领先且强大的数字基础设施存储与管理，数据库则扮演了核心角色。只有建立起强大的数据技术软件体系，才能满足数字经济发展的需要，而数据库为数字中国建设和数字经济社会的稳定及健康发展提供了基础技术和设施保障。

#### 2. 美国数据库厂商长期垄断全球数据库市场

全球数据库市场主要由传统 IT 巨头、云计算厂商、独立数据库厂商三大势力进行角逐，传统 IT 巨头（Oracle、IBM、Microsoft）多在关系模型领域深耕，整体占据市场主导地位，同时随着云计算、大数据的快速发展使得需求有了进一步的爆发，推动了云数据库厂商（Amazon、Google）的快速崛起。美国 Oracle、IBM、Microsoft、Amazon、Google 五家科技巨头公司联合占据了全球数据库 75% 以上的市场份额，全球数据库行业呈现美国一家独大的局面。在国内，Oracle、Microsoft 等国际厂商也占据了超过 65% 的数据库市场份额，与国际市场面临着相似的境况。

### 专栏 7：全球数据库市场呈现三雄争霸局面

在传统数据库领域，由于国外数据库厂商起步较早，同时在发展过程中不断进行战略布局和产品迭代，已经形成了先发优势，代表性企业有 Oracle、IBM、Microsoft 等。

伴随云计算、大数据等技术的快速发展，互联网巨头在数据库市场大放异彩，云计算巨头的数据库产品借势快速发展，代表性企业有亚马逊、谷歌、华为、阿里、腾讯等。

在商业化和开源并行发展的同时，也涌现出一批技术起家的数据库独立厂商，如 MongoDB、Snowflake 等公司。

总体来看，传统 IT 巨头在关系型 OLTP 领域强势，云计算厂商以及独立数据库厂商在非关系型领域以及 OLAP 领域更具竞争力。云计算厂商更多的是依靠云计算、大数据的快速发展崛起，同时，独立数据库公司更加依赖开源来快速提升产品影响力和市场认知度，同时社区的力量也可以让公司的产品快速迭代升级。

### 3. 数据库技术路径分化带来市场分裂与竞合

过去 10 年，全球有近 300 家数据库企业相继成立，而全球数据库产业绝大部分市场份额主要被美国科技巨头占据，其他 90% 以上的数据库企业生存艰难。这些数据库企业中，一部分基于其他商业或开源数据库做二次开发的数据库产品，一部分通过自研开发数据库产品，还有一部分企业在一些垂直细分行业领域深耕，但由于市场发展空间有限，生存处境比较艰难。未来绝大部分数据库企业可能会逐渐消亡或被头部公司吞并收购，只有那些迎合市场发展需求，保持技术领先优势、生态完善的企业才能存活下来，数据库市场将有可能从分散逐步收敛走向集中化。

#### **4. “开源 + 云”已经成为数据库革新的关键力量**

进入云时代，全球数据库行业正在经历大变局，而驱动这场大变局的关键因素，主要来自数据技术供需两侧的双向巨变。在需求侧，数字化在全行业的加速推进为数据库产业带来了持续的增长动力；在供给侧，“开源 + 云”成为改变数据库技术的内聚力量，正在把过去 20 余年分散多元的数据库技术栈融合起来。未来借助“开源 + 云”的力量，将会涌现出一批真正能够解决企业数据管理瓶颈的服务企业，重塑全球数据库产业发展格局。

#### **（四）全球数据库发展新态势**

##### **1. 数据库和硬件协同成为新兴竞争力量**

随着新型硬件成本逐渐降低，充分利用新兴硬件资源提升数据库性能是未来数据库发展的重要方向之一。在需求侧，随着数据量暴增和实时性的要求越来越高，数据库围绕处理器、内存、存储和网络四个硬件架构方向不断突破和创新。新兴硬件技术将塑造未来数据库的发展趋势，向着更高的性能、更好的可扩展性和更低的总体成本发展。数据库和新兴硬件之间的融合创新，将共同推动机器学习、人工智能、大数据分析、大语言模型等领域技术的发展进步。

##### **2. 数据库领域的投资热情正在逐年升温**

2022 年全球数据库市场规模超过 975 亿美元，同比增长 14%。中国软协预测，未来 3-5 年内，全球数据库市场将保持 10%-15% 年复合增长率，预计到 2025 年市场规模有望突破 1250 亿美元。

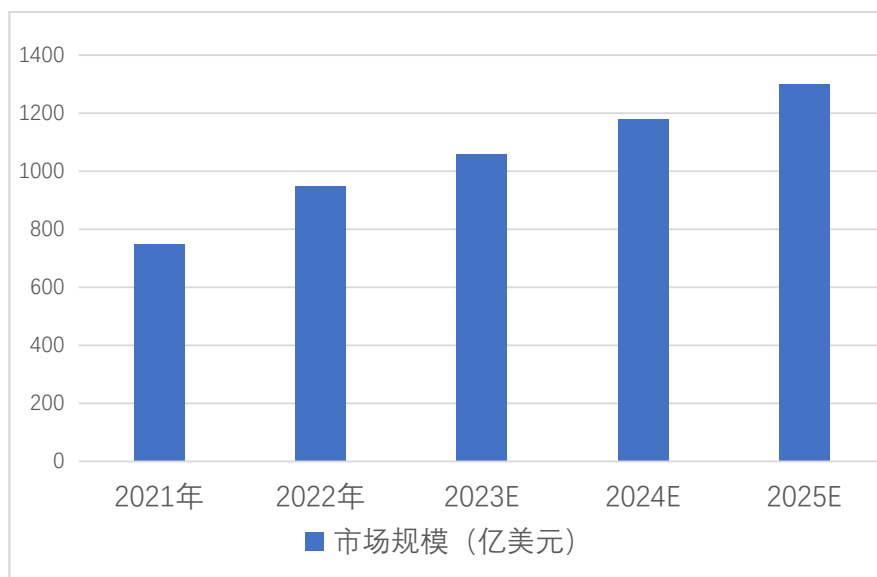


图 4 2021-2025 年全球数据库市场规模预测

据不完全统计，2021-2022 年共有 32 家开源商业公司获得融资且有多家公司在短时间内获得多轮融资，其中单轮融资数达到 1 亿美元的企业数量占比接近 20%，达到 1000 万美元的企业数量占比超过 50%。其中，数据库领域成为开源融资界的宠儿。2022 年数据库商业公司获得融资的数量占比接近所有行业融资额的 25%，明显高于其他行业公司，这说明市场对数据库行业的投资热情仍然很高。

### 3. 中国数据库迎来技术积累爆发增长期

当前，中国数据库市场迎来黄金发展期，国内拥有一个蓬勃发展的市场，大市场带来了丰富的应用场景，为国内数据库技术实力的提升提供了优势基础。另外，开源模式为技术创新的扩散应用提供了良好的成长土壤，也为全球经济的转型发

展提供了广阔的创新思路。我国在全球开源体系中占据重要地位，国内在数据库、人工智能、操作系统领域的开源项目日益活跃，这跟中国互联网行业的市场特点和当前技术发展紧密相关。以华为 openGauss 为代表的国产数据库相继开源，国内的数据库开源社区、开源人才、开源组织、开源项目的数量和质量正在持续提升，加速推动了国产数据库技术的迭代升级发展。

#### 4. 技术收敛是国产数据库突围有效路径

经过近十年的快速发展，国产数据库已经具备应用在关键基础设施行业的技术储备。随着政策利好、数字化转型大潮以及市场教育逐渐成熟，国内企业的 IT 支出提升，推动着数据库行业迎来发展机遇。但是国内厂商众多、技术路线分散导致数据库市场严重内卷，对于客户来说也会出现难以选择厂商的问题。预计未来 5-10 年，我国数据库市场将会产生激烈竞争，最终可能突围留下技术成熟度高、安全可控、生态布局完善的少数几家头部企业，其它更多的数据库企业将会逐渐融入到头部企业生态圈，实现良性循环发展。



### 三、国产数据库发展状况

国产数据库在过去三十多年里一直处于追赶状态，经过最近十年的细心打磨沉淀，国产数据库已经取得了显著进步，涌现出了一批较为成熟的国产数据库产品，并与国外数据库产品形成竞争态势，国产数据库市场规模也呈现出快速增长的趋势。

#### （一）发展现状

数据库是数字基础设施的根技术，在数据库发展的历史进程中，前五十年数据库世界的主角都是美国科技厂商。自 2000 年以来，我国开始陆续出现商业数据库厂商，特别是国产数据库软件在过去几年中快速成长，呈现出持续向好的发展势头。

——**技术创新不断突破。**国产数据库软件注重技术创新，国产数据库开发不断引入最新的技术和算法，提高数据库的性能和功能。例如，一些国产数据库产品采用了列存储、分布式架构、内存计算、存算分离、软硬协同等先进技术，以满足大数据、实时分析等新兴应用需求。

——**产品功能更加丰富。**国产数据库不断添加新功能和特性，以满足不同行业和用户的需求。国产数据库软件除了基本的 SQL 支持，还具备数据处理和分析功能。

——**商业应用进程加速。**国产数据库在商业化方面也有了积极的进展。一些公司提供了商业许可证的产品版本，并提供了技术支持、培训和咨询等服务，以满足企业用户的需求。

——**开源模式广泛兴起。**国产数据库中的一些产品越来越

重视开源模式，吸引了大量的社区贡献者和用户参与，形成了活跃的开发和生态系统。如华为公司开发的 openGauss 企业级开源关系数据库，受到业内的广泛关注和使用。

总体来看，随着国家不断提高对信息安全和自主可控的重视程度，在国家政策的驱动下，国内数据库企业更加重视自主研发创新能力，相继推出了众多具有自主知识产权、性能完备的数据库产品，在国内数据库市场已经占据了一席之地。

## （二）市场规模

随着数字经济的兴起，各领域数字化转型深入推进，千行百业对数据库的应用需求显著增加，促进了我国数据库市场的繁荣发展。根据中国软协市场调查数据显示，2022 年我国数据库市场规模约 364.5 亿元，其中国产数据库市场规模约 27.4 亿元，占比为 7.5%。中国软协预测 2023-2025 年我国数据库市场将保持高速增长态势，年平均增速为 35%，预计 2025 年市场规模有望达到 900 亿元。

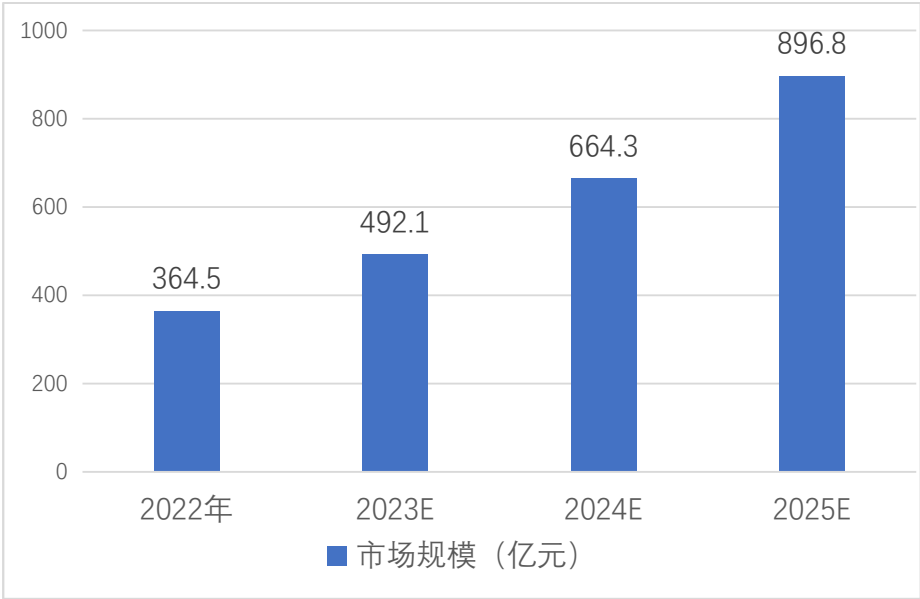


图 5 2022-2025 年我国数据库市场规模预测

（三）国产数据库主要厂商产品分析

据不完全统计，国内现有数据库企业数量超过 250 家，数据库产品种类超过 280 个，本文选取了具有行业代表性的 10 家企业进行重点分析。



图 6 国产数据库产业全景图谱

## 1. 达梦数据库

武汉达梦数据库股份有限公司成立于 2000 年，是国内领先的数据库产品开发服务商，国内数据库基础软件产业发展的关键推动者。公司为客户提供各类数据库软件及集群软件、云计算与大数据等一系列数据库产品及相关技术服务，致力于成为国际顶尖的全栈数据产品及解决方案提供商。达梦公司拥有数据管理与数据分析领域的核心前沿技术，拥有主要产品全部核心源代码的自主知识产权。

主要产品：DM8、DMDPC、GDMBASE、DMCDM、启云数据库等。

主要应用领域：服务于包括建设银行、中国人保、国家电网、中国航信、中国移动、中国烟草等在内的知名用户，成功应用于金融、能源、航空、通信、政府、医疗、教育等多个领域。

## 2. 南大通用

天津南大通用数据技术股份有限公司成立于 2004 年，自成立以来始终坚持自主创新，产品的核心技术及底层代码自主可控，构建了覆盖数据管理全生命周期，包括分析型、事务型、分布式事务型、云原生数据仓库等全技术栈的数据产品体系及服务解决方案，在国内数据库市场具有较高的品牌知名度。

主要产品：GBase 8a、GBase 8s、GBase 8c、GBase XDM、GBase 8d 等。

主要应用领域：南大通用自主研发的 GBase 系列数据库产品及服务范围覆盖海外 34 个国家，全国 32 个省级行政区域，

为金融、电信、政务、能源、交通等百余个行业上万家用户提供产品和服务，目前已建立节点超过 35000 个，管理数据总量超过 400PB。

### 3. 神舟通用

天津神舟通用数据技术有限公司隶属于中国航天科技集团，是北京神舟航天软件技术股份有限公司控股子公司，是国内从事数据库、数据挖掘分析产品研发的专业公司，是国内最具影响力的基础软件企业之一。神舟通用公司获得了国家核高基科技重大专项重点支持，是核高基专项的牵头承担单位。

主要产品：神通关系型数据库、神通 KStore 海量数据管理系统等。

主要应用领域：神通数据库已经广泛应用于政府、军队、军工、金融、电信、航天、邮政、能源、互联网、交通运输、制造业、教育、烟草、税务、审计等各个领域的涉密信息系统和信息系统安全等级要求较高的核心信息系统等领域。在国民经济支柱行业，尤其是对于具有自主可信需求的领域，神通数据库也为其提供了可靠的数据存储管理支持。

### 4. 人大金仓

北京人大金仓信息技术股份有限公司成立于 1999 年，是较早成立的拥有自主知识产权的国产数据库企业，也是中国电子科技集团成员企业。公司专注数据库领域 20 余载，具备出色的数据库产品研发及服务能力。曾先后承担国家“863”、



电子发展基金、信息安全专项、国家重点研发计划、“核高基”等重大课题研究。

主要产品：KES、KADB、KFS 等。

主要应用领域：人大金仓具备国内先进的数据库产品、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。2022 年，金仓数据库管理系统 KES 入选国务院国资委发布十项国有企业数字技术典型成果。

## 5. 华为公司

华为公司是全球领先的数字基础设施和智能终端服务商。openGauss 是华为公司自主研发的一款全面友好开放，携手伙伴共同打造的企业级开源关系型数据库。openGauss 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。openGauss 深度融合华为在数据库领域多年的研发经验，结合企业级场景需求，持续构建竞争力特性。

主要产品：openGauss、GaussDB 等。

主要应用领域：华为数据库产品广泛应用于 IoT、电子商务、电子政务、移动游戏、金融保险、互联网、工业制造、智慧城市、智能家居、车联网等行业领域。

## 6. 阿里云

阿里云是全球领先的云计算及人工智能科技公司，为 200 多个国家和地区的企业、开发者和政府机构提供服务。阿里云

拥有强大且丰富的云数据库产品家族，涵盖关系型数据库、非关系型数据库、数据仓库、数据库生态工具四大版块，可以为企业数据生产和集成、实时处理、分析与发现、开发与管理提供全链路生命周期的服务。

主要产品：PolarDB、RDS 等。

主要应用领域：阿里云数据库已经应用于多个行业，支持电商、金融、游戏、媒体、人工智能、物联网等多种场景，为数据处理、存储和管理提供多种解决方案，满足不同业务的需求。

## 7. 腾讯云

腾讯云是腾讯旗下云计算服务公司，提供全球领先的云计算、大数据、人工智能等技术产品与服务。腾讯云数据库为企业完善的关系型数据库、非关系型数据库、分析型数据库和数据库生态工具。用户可以通过产品选择和组合搭建，实现高可靠、高可用性、高性能等数据库需求。

主要产品：TDSQL-C、TencentDB for MySQL、TDSQL for PostgreSQL、TencentDB for Redis 等。

主要应用领域：腾讯云数据库服务超过 2000+ 的金融政企，行业覆盖银行、保险、证券、政务、互联网金融等各个领域。其中，TDSQL 目前已成功打造粤省事、深圳地铁、中国银行、富途证券、阅文集团、游族网络、腾讯游戏等多个项目方案。

## 8. 平凯星辰

北京平凯星辰科技发展有限公司（PingCAP）是业界领先的企业级开源分布式数据库企业，提供开源分布式数据库产品、

解决方案与咨询、技术支持与培训认证服务，致力于为全球行业用户提供稳定高效、安全可靠、开放兼容的新型数据服务平台，解放企业生产力，加速企业数字化转型升级。

主要产品：TiDB。

主要应用领域：TiDB 在全球已拥有超过 3000 家企业级用户，其中海外用户超过 1000 家，覆盖美国、法国、挪威、爱沙尼亚、日本、澳大利亚、新西兰、新加坡、印度、越南、秘鲁等国家和地区，涉及政府、金融、电信、能源、制造业、教育、信息安全、互联网、物流等各个行业，产品用户数在国内处于领先地位。其中，TiDB 在金融、电信、互联网、物流、高端制造业等对数据库要求较高的领域广受欢迎。

## 9. 云和思墨

云和思墨（北京）信息技术有限公司创立于 2011 年，是中国领先的数据库系统软件提供商，在数据管理、承载、加工和应用领域为各个组织提供可信赖的产品、服务和解决方案。云和思墨以数据库为核心，向下提供数据库一体机平台、向上提供数据库云管平台，形成了端到端的数据库解决方案。此外，公司打造的墨天轮社区已成为国内最大的数据库生态社区。

主要产品：MogDB、Uqbar、zCloud、zData 等。

主要应用领域：云和思墨累计直接服务金融、保险、通信、能源、政务、制造、交通、医疗、商贸 8 大关键行业的 1000 多家企业级客户，特别是在电信、金融等领域拥有广泛的成功案例。

## 10. 海量数据

北京海量数据技术股份有限公司成立于 2007 年，是国内以数据库为主营业务的上海主板上市公司。海量数据始终坚持自主研发和创新，拥有北京、杭州、广州三大研发基地，团队技术人员占比超过 65%。

主要产品：Vastbase G100、Vastbase E100、AtlasDB 等。

主要应用领域：海量数据库可支持绝大部分业务应用场景、开发环境和 IT 架构，目前广泛应用于政务、制造、金融、通信、能源、交通、国防、军工、医疗、教育等多个重点行业。特别是在制造行业的细分领域中拥有较高的市场占有率。

### （四）国产数据库在重点行业应用分析

近几年以来，国产数据库在技术创新、市场竞争和用户认可度等方面取得了显著进步，国产 OLAP 和非关系型数据库逐渐成为国内众多企业和机构的首选数据库解决方案，并且在党政军、金融、电信、能源、电力、医疗、教育、交通、公共事业等领域和关键行业取得了重要应用成就。同时，一些国产数据库厂商开始瞄准走向国际市场，积极主动地参与到全球竞争中。通过技术创新、合作伙伴关系和市场营销等多种方式，争取更多的国际用户和业务机会，提升在全球数据库市场的竞争力。

分行业来看，金融、电信、政府、制造、交通五大行业占据数据库服务市场的份额超过 85%。各行业的数据库服务市场份额比例分别为金融 24.3%、电信 19.6%、政府 17.2%、制造 15.3%、交通 9.8%。

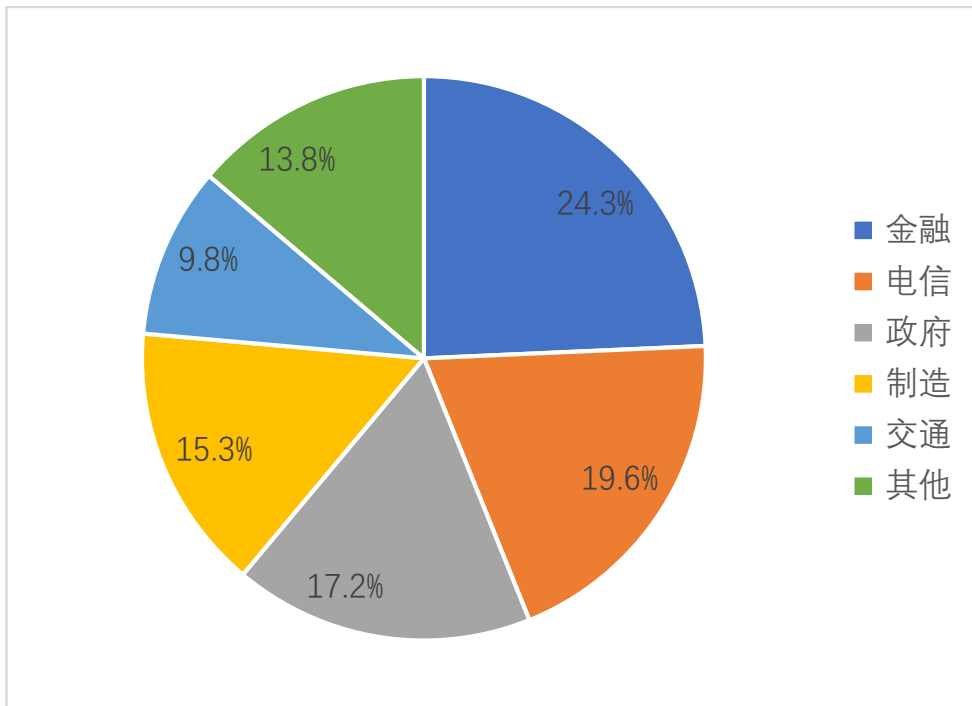


图 7 各行业的数据库服务市场份额

未来几年，随着我国在基础软件研究领域的不断进步，国产数据库技术将会在更多的行业领域发挥关键性作用，支撑千行百业实现数字化变革和高质量发展。

### （五）国产数据库厂商积极布局生态建设

数据库作为基础软件，需要完成芯片、服务器、操作系统、中间件的适配工作，ISV、集成商、二次开发商、IT 咨询公司都是数据库厂商生态伙伴体系中的重要参与者。现阶段，企业自建生态难度较大，兼容生态成为主要选择。生态伙伴赋能企业不仅能够快速实现业务扩张，还能最大程度地减少成本支出，使得企业能将有限的人员和资金投入到了核心技术和核心产品研发中。

随着国产数据库技术的快速进步，同时借助生态链的构建和逐步完善，国产数据库企业不断发力、崛起，涌现出了一批极具创新力的优秀产品，极大地推动了国产数据库应用的遍地开花。华为 openGauss、阿里云 PolarDB for PostgreSQL、PingCAP TiDB 等开源社区的发展驶入快车道，开源模式正在重塑国产数据库生态新格局，也成为未来数据库领域发展的重要驱动力。

**专栏 8：国产数据库公司发展的战略重点在于生态扩展**

做数据库一定要重视生态建设，自从 X86 替换小型机，DB2 的市场份额逐渐下降；软硬件分离的趋势，导致 Intel+ 微软的操作系统得以普遍应用，在此生态下才有 Oracle、MySQL 发展壮大的机会。

合作伙伴生态是 Oracle 早期占领中国市场的核心要素之一，早期 Oracle 中国区 90% 以上收入来自 2000 多个合作伙伴。



## 四、主流数据库技术发展分析

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，是数据库的核心要素，数据模型的变迁史是数据库技术架构升级的主线。关系型和非关系型数据库是业界使用最多最普遍的主流分类方式，故本文主要从数据结构分类的角度对数据库技术的发展进行深入分析。

### （一）关系型数据库

#### 1. 技术特点

关系型数据库具备 ACID<sup>[1]</sup> 特性，是常见的主流数据库类型。关系型数据库在维护数据库完整性、数据一致性方面优势突出，适用于对数据安全性及事务支持要求较高的应用场景。

1) 可扩展性能良好。关系型数据库可以轻松地进行垂直和水平扩展，包括添加更多的 CPU、内存、磁盘等。同时，关系型数据库还支持分区，可以将数据库水平分成多个部分，从而实现更高的吞吐量。

2) 兼容性出色。由于 SQL 是关系型数据库的核心查询语言，因此基于 SQL 的关系型数据库可以无缝地与其他数据库软件（如 BI 报表、ETL 工具等）集成。这样，用户可以将数据频繁地从一个环境转移到另一个环境，同时也方便了开发人员在多个数据库之间进行操作和查询。

---

ACID: 即原子性 (Atomicity)、一致性 (Consistency)、隔离性 (Isolation)、持久性 (Durability)。

3) 成熟和稳定性高。关系型数据库走过了 40 多年的发展和完善过程，已经变得越来越成熟和稳定。与数据库配套的相关软件和工具逐渐趋于完善，有大量的模板和文档可供使用，同时也有一个庞大和信息充沛的社区来维护、支持。

4) 数据一致性和可靠性好。关系型数据库支持事务，可以在多个操作之间创建可回滚的模式，同时也支持记录锁和表锁，这使得关系型数据库在保证数据的一致性和可靠性时更加出色。

## 2. 主流产品

市场上主流的关系型数据库包括 Oracle、MySQL、IBM DB2、Amazon RDS、Microsoft SQL Server、Microsoft Access、PostgreSQL 等多个产品种类，每种数据库的语法、功能和特性也各具特色。

## 3. 发展趋势

关系型数据库仍然在许多方面具有非常重要的功能作用，尤其是在企业应用场景中，关系型数据库是最流行的解决方案之一。未来，随着产生的数据量的不断增长和数据流程的加速，关系型数据库的优越性将得到更加充分的体现。同时，关系型数据库也将继续不断地适应新的技术和需求，保持持续发展和创新。

## (二) 非关系型数据库

### 1. 技术特点

区别于关系型数据库，NoSQL 数据库遵守 BASE<sup>[2]</sup> 模型。

BASE：即基本可用（Basically Available）、软状态（Soft state）、最终一致性（Eventual consistency）。

NoSQL 数据库具有易扩展、大数据量、高性能、灵活和高可用的特点。

1) 易扩展。NoSQL 数据库种类繁多，但都有一个共同的特点，即不具有关系数据库的关系型特性。数据之间无关系，这样就非常容易扩展。无形之间，在架构的层面上带来了可扩展的能力。

2) 大数据量、高性能。NoSQL 数据库都具有非常高的读写性能，尤其在大数据量下，同样表现优秀。这得益于它的无关系性，数据库的结构简单。一般 MySQL 使用 Query Cache。NoSQL 数据库的 Cache 是记录级的，是一种细粒度的 Cache，所以 NoSQL 数据库在这个层面上来说性能就要高很多。

3) 数据类型灵活。NoSQL 数据库无须事先为要存储的数据建立字段，随时可以存储自定义的数据格式。

4) 高可用。NoSQL 数据库在不太影响性能的情况下，就可以方便地实现高可用的架构。比如 Cassandra、HBase 模型，通过复制模型也能实现高可用。

## 2. 主流产品

市场上常见的 NoSQL 数据库包括键值型数据库、列族数据库、文档数据库、图数据库、时序数据库、空间数据库等。

1) 键值型数据库：适用于内容缓存，如会话、配置文件、参数等。其扩展性高，灵活性好，大量操作时性能高，但数据无结构化，查询方法单一。典型产品：MemcacheDB、Berkeley DB、Redis、Apache Ignite、Riak 等。

2) 列族数据库：适用于分布式数据存储与管理，将同一列数据存在一起，可扩展性强，查找速度快，复杂性低，但功能局限。典型产品：Hbase、Cassandra、Hypertable 等。

3) 文档数据库：适用于存储文档数据，数据结构灵活，但缺乏统一查询语法。典型产品：MongoDB、Apache CouchDB、ArangoDB、Couchbase、Cosmos DB、IBM Domino、MarkLogic、OrientDB 等。

4) 图数据库：适用于图像数据、社交网络、推荐系统，专注构建关系图谱，支持复杂的图形算法，但只能支持一定的数据规模。典型产品：Neo4J、FlockDB、AllegroGraph 等。

5) 时序数据库：为包含个体、时间、状态信息的实时流数据进行存储、检索和管理数据库，适合于物联网、性能监控服务。典型产品：InfluxDB、Kdb+、Prometheus、Graphite、TimescaleDB、阿里云 TSDB 等。

6) 空间数据库：针对二维或多维的空间数据进行存储和管理的数据库，适合于地图服务和时空分析场景。典型产品：Oracle Spatial、PostGIS、Microsoft SQL Server Spatial 等。

### 3. 发展趋势

NoSQL 数据库扩展性好，储存模式简单，查询速度更快。随着云计算的发展和大数据时代的到来，关系型数据库越来越无法满足行业发展需要，NoSQL 数据库将成为云时代数据库领域的重要增长引擎。

表 关系型数据库与 NoSQL 数据库特点对比分析

	关系型数据库	NoSQL 数据库
模式	预定义的模式	没有预定义的模式
查询语言	结构化查询语言 (SQL)	没有声明性查询语言
一致性	严格的一致性	最终一致性
事务	支持	不支持
理论基础	ACID	CAP, BASE
扩展	纵向扩展	横向扩展 (分布式)

### (三) 数据库技术架构演进

数据库随底层 IT 架构、数据特征、企业需求等因素的变化及发展而呈现出渐进式的演进特征，涌现出新的数据模型、产品架构、负载需求以及部署方式，数据模型变迁、分析型需求崛起、分布式架构的演进以及开源模式的流行等要素对推动数据库架构的不断革新产生重要影响。

#### ——数据模型的变迁

数据模型作为数据描述的主要手段，是连接真实世界和逻辑世界的桥梁，数据模型的演进是贯穿整个技术架构升级的主线，带来了从层次、网状、关系型，再到 NoSQL、NewSQL 的变迁。

#### ——分析型需求的兴起

伴随着数据库在各行业领域的逐渐渗透，场景应用累积的大量数据仅靠原来的业务支持系统已经无法完成信息价值的提取，因此，面向分析需求的 OLAP 技术路径从 OLTP 中独立出来，并预示着大数据时代的到来。

## ——分布式架构的演进

从 Scale Up 到 Scale out，分布式架构的诞生使得数据库得以实现横向扩容以及更好的并行处理能力。随着技术的不断发展和应用场景的不断丰富、变化，分布式架构也在不断进行演进和完善，未来可能向智能化、自动化及可信任的分布式架构演进。

## ——开源思潮的流行

开源作为一种高效开放共享的软件开发组织模式，从 MySQL 到 PostgreSQL，开源生态圈的加速建立也在技术浪潮中影响着数据库厂家技术路线的选择。

## ——数据库未来的思考

数据库面临着技术收敛与新兴技术潮流下的分化发展考验。一方面，经历几十年的发展，数据库技术不断分化、丰富，然而随着企业数字化进程的推进，部署多套系统带来了数据一致性与运维管理等方面的挑战，因此出现了技术收敛的趋势，对应多模、数据湖仓等发展方向。另一方面，数据库持续与新兴的 IT 技术结合以进一步提升性能，主要对应原生分布式、云数据库、智能自治、结合新硬件等发展方向。



## 五、国产数据库发展面临的形势、问题及机遇

### （一）面临的主要形势

国际方面：中美两国之间的贸易摩擦和制裁加剧，美国在科技领域对中国的打压制裁事件频发。

当前我国数字经济的发展逐步走向深化应用、规模发展、普惠共享的新阶段，需要逐步摆脱对国外根技术的严重依赖，以确保我国的产业升级不受制于人，数据库是需要重点关注的基础软件根技术领域。

#### 专栏 9：甲骨文数据库断供俄罗斯启示

长期以来，全球流行的 Oracle、SQL Server、DB2 等数据库全部都被美国所掌握。随着俄乌战争的爆发，美国众多企业也开始配合美国政府的行动，对俄罗斯进行制裁。截至 2023 年 11 月，陆续已有百余家欧美企业对俄罗斯进行制裁，或暂停在俄罗斯的业务，或退出俄罗斯市场。美国最大的数据库公司甲骨文也对俄罗斯进行制裁，暂停在俄罗斯的一切业务，这给俄罗斯银行、能源等领域造成了较大的影响。

启示：在当前国际贸易摩擦加剧、国际局势多变的情况下，数据库软件作为最重要的 IT 基础设施之一，我国必须做好防范，要加快实现数据库国产化和自主可控。

国内方面：在信创政策持续催化下，国产数据库发展窗口已经打开。

从 2016 年信创概念被首次提出，到 2020 年信创运动的逐渐兴起，再到 2022 年逐步深化发展，数据库等核心基础软件产品，在相关政策的驱动下，逐渐从“能用”向“好用”转变。

国产数据库厂商获得产品应用以迭代升级的机会，有望加速进步。党政、金融、电信、石油、电力、交通等行业出台相关政策，将推动其他关联行业国产化替代进程。

专栏 10：数据库为什么实现国产化
<p>国家战略高度：数据库是信息化建设的基础设施之一，国家需要加强对数据库软件的自主研发和掌控，以保障国家信息化建设的安全和可持续发展。</p> <p>技术革新驱动：国产数据库软件的研发和生产可以促进国内技术的与时俱进，进一步提高国内企业的竞争力。</p> <p>数据安全形势：关键行业的数据涉及到国家安全和经济安全，如果使用国外数据库软件，可能会存在数据泄露和被窃取的风险。</p> <p>自主可控需要：使用国产数据库软件可以保证数据的自主可控，避免因技术原因无法对数据进行有效管理和控制。</p>

**市场方面：国内数据库市场各类型国产品牌已经崛起。**

根据调查数据统计，排名前五的国产数据库厂商的市场份额从 2018 年的 27.1% 上升到 2022 年的 55.4%；而前五名国际厂商的市场份额则从 2018 年的 57.3% 下降到 2022 年的 27.3%。国产数据库产品与国际厂商展开竞争与替代，已经越来越广泛地被使用到关键行业和企业核心系统中，特别是在分布式、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。

2023 年随着疫情管控的放开，国内关系型数据库市场格局正在重塑，数据库厂商的业务开展加速步入正轨，企业数据库建设项目广泛开展，基于构建市场新格局角度，2024 年将是国产数据库实现突破的关键转折之年。

## （二）存在的主要问题

我国是全球数据体量最大，应用类型最丰富的国家之一。各种新业态、新场景正在国内蓬勃涌现，庞大的数据要素市场亟待释放。数据库产业作为数据价值的载体，当前还存在市场碎片化、技术创新迭代滞后、产业生态不健全等问题。

### 1. 产品种类繁多，竞争能力不足

我国国产数据库市场存在严重的碎片化现象，虽然厂商数量众多，但“小、散、弱”的问题十分突出，无法形成对国外数据库巨头的竞争优势。据统计，我国数据库厂商数量已超过 250 家，国产数据库产品超过 300 种。而大多数厂商的数据库产品都是来自于开源项目，这些企业研发的数据库产品缺乏核心技术，在国内数据库市场竞争力较弱，发展空间比较有限。

### 2. 技术基础薄弱，专业人才匮乏

数据库技术门槛高、研发周期长、工程要求高，做出一款好的数据库产品需要大量的专业技术人才长期进行研发。但由于国产数据库的研发历史较短，技术成熟度欠缺，与国外产品相比还存在许多缺陷和差距。同时，经过调研发现国内数据库研发人才储备比较匮乏，研发人数占比过低，超过 60% 的厂商研发人员不足百人，跟国际主流数据库厂商对比差距明显。专业人才储备不足，严重制约了数据库技术的研发和创新，导致产品市场竞争力不足。

### 3. 安全性能存忧，用户使用存虑

国内目前很多代码自主化率并不高，大部分国产数据库都

是在现有的开源基础上进行修改，这就导致自主性较低，发展空间受限。开源代码的根社区位于海外，如果欧美对我国科技产业进行限制打压，开源数据库产品的安全性势必受到严重威胁。特别是央国企用户对于替代使用国产数据库普遍存在担忧，怕担责任而不想也不敢尝试选择使用国产数据库产品。在进行数据库迁移时也面临配套工具不完善、数据库不稳定等难题，一般来说，一家企业完成数据库的国产替代需要花上 2-3 年的时间，在这过程中需要评估改造难度。

<p><b>专栏 11：以 MySQL 构建的数据库生态存在安全风险</b></p> <p>国内有大量的 MySQL 部署和基于 MySQL 提供运维服务的商业公司。但是，使用国外开源数据库面临风险，特别是 MySQL 所采用的 GPL 协议无法做到自主可控。MySQL 使用的 GPL v2 协议具有传染性，基于开源 MySQL 以及二次开发的国产数据库也必须开源和免费。GPLv2 协议规定任何修改或基于 GPL 软件二次开发的代码都要履行开源义务，使用不慎很容易导致私有代码被动开源。</p> <p>此外，美国商务部工业和安全局（BIS）发布的《信息安全控制：网络安全物项》，就美国实体对相关国家及地区分享网络安全漏洞设立许可申请，限制中国等获得网络安全漏洞的能力。这意味着国内企业获取 MySQL 安全漏洞需要向美国申请，不能及时获取 MySQL 安全漏洞，存在数据库生态安全风险。</p>
---

4. 产品标准各异，缺乏行业标准

国内数据库发展正处在百花齐放、百家争鸣阶段，市场上出现了不同的架构、不同的技术路线、不同的编程接口。由于各家公司的数据库产品标准各异，导致系统与系统之间产生兼容性问题，基于对维护国产数据库市场的长期稳定发展的考量，

国产数据库行业亟需制定规范统一的行业标准规范。

### **5. 产业资源分散，生态建设滞后**

国产数据库厂商普遍缺乏对产品的长远发展蓝图和规划，没有清晰的产品路线图，也未建立起长期持续发展的机制。由于当前国内数据库产业研发缺乏明确的分工，各厂商将有限的研发人才投入到数据库技术堆栈的各个领域，导致各个数据库厂商在研发过程中重复研发全栈技术，造成了大量的人力资源浪费。

数据库系统生态建设需要研发针对新应用场景的评测方法和工具，需要构建多方积极参与、密切互动、持续成长的社区，需要培育包括用户、产业链厂商、研发梯队的社区队伍。国产数据库系统厂商起步相对较晚，发展时间较短，做好生态建设也是难点问题。

### **（三）发展机遇**

数据库是数字经济高质量发展的基础设施，保障数据库行业健康发展，对于促进数字经济发展与实体经济深度融合具有重要的底层支撑作用。

#### **1. 利好政策出台释放巨大红利**

近年来，国家十分重视国产数据库产业发展，围绕信息技术、政务民生、金融、交通物流、工业制造、农业农村、气象、知识产权等多个领域相继出台了专门的支持政策。借助政策东风，国产数据库厂商奋发有为，市场版图快速扩张。另外，作为信创领域基础软件的重要组成部分，数据库的国产化替代也



在加速进行中，GaussDB、PolarDB、DM8、GBase、TiDB 等数据库产品是近年来备受关注的—批国产数据库产品，在各自的领域和场景下都有着广泛的应用。

## **2. 数字化激发数据库市场潜力**

近年来，随着数字经济的蓬勃发展和数字化转型的深入推进，各行各业数字化转型对国产数据库的需求显著增加，产业生态也在逐步建立和完善布局，行业发展基础已初步搭建完成，市场环境的变动也给国产数据库厂商带来了诸多利好，国产数据库的应用领域范围 and 市场份额有望随之进一步扩大，未来增长空间值得期待。

## **3. 数据库国产化替代空间巨大**

很长一段时间以来，海外数据库厂商在我国占据了绝大部分的市场份额。伴随着我国数字经济占比的逐年提高、数字基础设施规模能级的大幅提升、数据资源体系的加快构建，各行各业对数字信息安全的重视程度不断提高，国产化的替代意愿正在增强。同时国产数据库产品在价格、性能、功能等方面都取得长足进步，基本具备了与国外产品同台竞争的能力，应用领域也不再限于党政、金融领域，开始广泛拓展到电信、交通等诸多新场景、新消费、新服务领域，行业发展基础已初具规模，国产数据库市场有望迎来新的发展窗口。

## **4. 行业技术实力积累整体增强**

在 2013 年之前，我国大多数公司主要使用美国企业的数据库工具，其他国产数据库也受到外国相关技术的支撑。近年



来，随着互联网的兴起和快速发展，多种类型的数据呈爆发式增长，各种创新业务场景层出不穷，进而促进了供给端厂商技术和产品架构的持续创新。以华为公司为代表的一批国产数据库厂商，重视基础研究与根技术的创新投入，坚持自主独立研发，拥有对关键基础设施行业应用的技术和人才储备，国产数据库厂商的整体技术实力得到了显著提升。



## 六、国产数据库发展基本判断

### （一）国产数据库技术路线收敛势在必行

通过对行业用户、数据库厂商、专家学者的深入调研，结合对国内外操作系统发展进程的对比研究分析，我们认为，在当前错综复杂的国际形势下，资源过度分散、无效内耗竞争是国产数据库难以与国外数据库竞争、大规模应用的重要原因。现阶段有必要集中国家力量、集合行业资源，重点支持少数几家国内领先的数据库根社区推动技术收敛，短期内快速推进国产化替代、强化生态建设，持续推动中国标准和产业链出海参与全球竞争。

### （二）国产数据库厂商应重视用户的需求

现阶段很多客户使用的数据库还是以国外产品为主，换用国产数据库产品则会面临各种迁移和兼容性问题，客户在进行应用迁移时普遍不愿花费额外的财力、人力进行应用改造，只想实现简单便捷的无缝迁移，以减少数据迁移和运维成本。客户场景端时效性要求倒逼数据库厂商不断提升产品兼容性能，提供异构数据库的平滑迁移工具或者解决方案。

### （三）存算分离 + 多读多写成为发展新趋势

随着数字化转型的深入，很多行业用户存在着大量云与互联网业务场景，业务形态逐渐变化，业务动态性增强，这对基础设施资源的灵活性和利用率提出了更高要求。在业务场景、硬件技术快速发展等多重因素的驱动下，国产数据库厂商正逐步

解决当前存算一体架构存在的问题。因此，新型存算分离架构正在迅速兴起，未来 2-3 年有望成为基础设施领域新的关注点，一些国产数据库厂商将在这一领域进行优化探索，数据库由“存算一体”向“存算分离”、由一写多读向多读多写架构演进。

#### （四）开源是国产数据库发展的有效路径

开源是当今基础软件在全球范围内取得成功的最优路径，开源软件商业化能力得到市场的正向反馈，因此开源将成为国产数据库厂商吸引用户、盘活生态和促进技术发展的重要手段。通过将底层技术开源，可以吸引更多专业人员加入开发、测试、维护，这不仅可以降低厂商售后和服务的成本，还可以进一步提高数据库软件的质量，同时有利于构建生态。华为公司的 openGauss 有望成为国内主流的数据库技术路线之一，基于 openGauss 的商业版公司有望脱颖而出。

## 七、建立以 openGauss 为代表的国产数据库体系

### （一）内核创新研究，打造坚实数据底座

openGauss 是华为公司全球 11 个研究所的上千名数据库专家研究开发的一款数据库产品，采用技术自立、发展独立、长期演进的技术路线。openGauss 围绕内核高性能、高安全、高可用、高智能四大竞争力，构筑领先的企业级开源数据库。

1) 高性能：基于 NUMA-Aware 架构优化，多版本并发控制等技术大幅提升单机性能；自研 In-place Update 引擎，采用原位更新和 Undo 空间统一分配机制，节约空间利用率。

2) 高可用：多层次可用性设计，提供同 AZ 内 RTO<10s 的业务恢复能力；引入 Paxos 分布式一致性协议，实现多地多中心的高可用能力；采用并行流水线解码技术，支撑基于 openGauss 的单元化多活容灾架构业界领先。

3) 高安全：率先在开源数据库中提供可商用的全密态能力，实现一次加密，全程保密，对敏感数据提供强大的保护；实现库内原生区块链算法和多方共识技术，对于数据修改可追溯，防抵赖，实现数据真实可信。

4) 高智能：提供原生 AI 能力，DBMind 为系统性能调优、故障预测、系统运维提供高效助力；基于 AI 的优化器 ABO 实现智能基数估计，贝叶斯网络算法大幅提升多列基数估计准确度；DB4AI 内置几十种常见的 ML 算法和模型，并通过 SQL 语法对外提供 AI 能力，使用户无障碍获得 AI 训练和推

理能力，实现会 SQL 就会用 AI。

## （二）联合产业资源，共建资源池化架构

数据库组件化架构标准已经成为业界的普遍共识。openGauss 推出了基于资源池化架构的数据库组件 DataPod，为数据库组件化架构标准提供了一个全栈优化、软硬协同的参考实现。



图 8 openGauss 资源池化架构图

资源池化架构由计算池化层、内存池化层、存储池化层组成，资源池化架构每层的功能划分如下：

1) 存储池化：对应组件化架构标准的存储层，支持多种存储，如分布式存储、企业存储、对象存储，实现一份数据服务于多种计算。通过高效裸设备访问，元数据共享，实现数据库原生的文件系统；通过 SQL 算子卸载的 NDP 技术，大幅提升了 SQL 处理效率、消减网络 I/O 流量。

2) 内存池化: 对应组件化架构标准的存储引擎, 实现计算节点间内存的互联, 通过同步事务信息和数据库缓存, 实现多节点下的多版本快照一致性读能力, 结合 RoCE 和 SCM 等硬件, 实现极致的 Commit 加速和大容量内存访问等能力。

3) 计算池化: 对应组件化架构标准的执行引擎和分发引擎, 支持多样性算力, 基于 X86、鲲鹏等算力, 为应用提供 TP 行存加速、AP 列存加速、AI 训练推理、向量数据库等全方位的数据服务, 实现多模融合, 使用最佳的引擎满足不同业务处理诉求。

同时, openGauss 推出数据全生命周期管理工具 DataKit, 实现包括资源管理、安装部署、数据建模、数据迁移、数据开发、智能运维的全场景数据管理功能。openGauss DataPod 和 openGauss DataKit 为组件化架构标准提供了参考实现, 通过软硬协同、全栈优化, 全新打造了一个高性能、高可靠、高安全的数智融合处理平台。

### (三) 满足行业诉求, 打造关基行业标杆

openGauss 聚焦核心场景与客户联合创新, 在多行业获得广泛商用, 目前已累积部署超过 3 万套。其中 2022 年部署超过 2 万套, 占非云新增市场的 10%。在金融行业, 中国邮政储蓄银行、民生银行都已经在核心系统中规模商用。在电信行业, 中国移动不但选择使用 openGauss 数据库, 实现分钟级割接, 整体性能倍增, 同时还基于 openGauss 社区版做了更多的创新。



#### 专栏 12: openGauss 典型场景应用案例

2019 年底，邮储银行立项开始新核心建设，并在 2020 年中选型 openGauss，通过 3 年的开发与建设，于 2022 年实现原旧核心业务和用户全量平滑迁移到基于 openGauss 的新核心上，这是全球首家大型商业银行中，基于通用硬件和开源数据库开发的个人核心交易系统。

新核心建成后，有效支撑了 6.5 亿用户，18 亿帐户 4 万个网点。性能与体验大幅提升，如柜面交易时间下降 50%，而交易记录查询从 1 年提升到 8 年。

#### （四）打造开源社区，共建技术开发生态

openGauss 社区是业界首个既面向 DBV、ISV 等乙方单位，又面向金融、运营商等行业甲方客户的联合社区，三大电信运营商、七大头部银行、头部央国企加入社区理事会运作，共建、共享、共治，实现产业链合力共建根社区。

当前社区用户吸引贡献人数超过 5500 人，下载量超过 223 万，395 家单位会员。社区成立 24 个 SIG 组，涵盖 SQL 引擎、存储引擎、安全、OM、工具、AI 等数据库主要模块。

华为公司积极构建起国产数据库生态，与鲲鹏计算生态形成优势互补，为国内数据库产业贡献力量，推动产业整体向更高水平发展。

#### 专栏 13: openGauss 提供三种产品形态，供行业客户选择

##### →商用发行版

由行业知名数据库厂商(DBV)基于 openGauss 社区版做二次开发，各个 DBV 立足对行业发展的认知和对客户的理解，把长期积累下来的应用实践融入自己的发行版中，形成差异化数据库产品，为行业客户

提供多样性选择。如云和恩墨的 MogDB，海量数据的 Vastbase G100，神舟通用数据库 openGauss 版。

#### →用户自用版

对于能力型客户，需要敏捷支持自身业务的情况，社区会支持客户基于社区版研发适合自身业务需求的数据库版本，可以与自身的业务深度融合，更高效支撑业务发展和商业目标达成。如中国移动的 CMDB、中国电信的 TeleDB for openGauss 以及中国联通的 CUDB 都是此种模式的合作成果。

#### →社区发行版

对于只关注数据库性能、安全、可用性等基础能力的客户，如果同时拥有运维团队和能力，可以选择完全免费的社区发行版。也可以购买第三方 DBV 的服务，在成本可控的情况下，可以获得强大的数据库能力。

### （五）走进高校课堂，培养专业技术人才

openGauss 依托智能基座，联合教育部辐射全国 2000 所高校，进课堂，进实验室，成为高校研究创新的新动力，并与产业界合作，高校学生可以选择进入企业实习，理论与实践结合，实现产教融合、协同育人。

同时 openGauss 已进入教育部的计算机等级考试，让高校学生不但能学 openGauss，还能从官方渠道获得认证，实现学习、认证、实习端到端打通，扩大学生就业和企业求才的喇叭口，实现三方共赢。

openGauss 与中国软件行业协会深度合作，建立起了一套完整的职业培训认证体系，从 openGauss 工程师到 openGauss 专家培训，广纳社会数据库人才，为企业提供高质量的数据人才输入。

## 八、发展对策建议

在数字化快速发展的时代，发展国产数据库离不开政府部门的统筹支持，同时也需要数据库厂商在技术、市场、生态等方面联合起来协同推进发展。

### （一）引导数据库技术路线收敛提升竞争能力

在当前复杂的国际形势和环境下，为进一步加快国产数据库快速突围，改变当前国产数据库厂商“小、散、乱”的局面，凝聚国内数据库行业力量的关键是“技术收敛”。对数据库技术路线的支持应当坚持主流少数的原则。要立根铸魂，重点支持 2-3 个技术路线，遴选较为成熟、创新能力较强、市场占有率较高的国产数据库产品，从上游根社区开始做技术收敛统一，减少碎片化和低水平重复开发，共同发挥国内各数据库厂商的技术创新能力，将创新成果持续汇聚在国内数据库原生根社区，为我国数据库产业发展增添新动能。

### （二）强化存算分离 + 共享存储标准体系建设

新的存算分离数据库架构已成为趋势，在此基础上，政府和行业机构需要统一制定新的标准规范，包括存算分离 + 共享存储架构中数据库与存储的接口标准，建立业界认可统一的技术规范，统一数据库厂商、计算及存储厂商的调用框架，更好地指引数据库产业链发展方向，形成产业合力。

### （三）出台专项政策支持首台套和规模化应用

数据库行业不能完全依靠市场的无序竞争，更需要政府出

台指导性、专项政策。国家和地方各级政府要主动牵头，制定系统完善的落地规划和免责机制，指导各行业和企业逐步使用国产数据库，促进行业用户“愿用、敢用、用好”，通过 3-5 年逐步完成国产化迁移替代。同时要建立良好的服务机制，推广国产数据库在金融、能源等重点行业的规模化应用，然后逐步将国产数据库拓展应用到更多的领域和场景中，获取更多的市场份额。

#### **（四）强化数据库技术创新集聚全产业链资源**

技术创新是源动力，国产数据库要坚持长期稳定的研发投入，不断进行技术升级，深耕数据库根技术原始创新，将自主安全可控作为重要目标，真正打造具有自主知识产权、安全可控的数据库内核根技术和开发工具，具备可独立演进能力，推出适应时代潮流、更加安全高效、性能稳定的国产数据库品牌。同时要汇聚上下游产业链资源，在国内快速建立起更加专业的技术支撑体系，打造更加完善的数据库服务工具和解决方案，帮助客户更好地使用国产数据库产品。

#### **（五）集中支持开源社区加大产业链生态建设**

全力扶持国产开源数据库发展，提供能够促使其快速成长的沃土。一是支持以 openGauss 为代表的开源社区发展，使之成为具有国际影响力的数据库根社区；二是建立健全开源社区运营模式，激励开发者加入社区，鼓励数据库厂商积极参与开源社区共建，贡献开源力量，凝聚产业力量，构筑面向全球的产业竞争力；三是加大与行业领袖的合作，构建完善的国产数

据库应用生态；四是充分发挥行业协会的桥梁纽带作用，促进政府、企业与市场的有效衔接；五是鼓励国产数据库企业开拓海外市场，在成熟的欧美市场打造标杆客户，到东南亚国家布局覆盖新兴市场，扩大国产数据库在全球的市场份额和影响力。

