

中国数据库研究报告

©2022.12 iResearch Inc.



中国数据库市场规模：据艾瑞统计，2021年中国数据库市场总规模达286.8亿元，同比增长16.1%。中国数据库市场虽受疫情影响，但整体稳步增长，由多方面因素促成：1) 信创利好；2) 需求驱动；3) 国产厂商厚积薄发；4) 国内用户对基础软件的IT支出，尤其是国产数据库的付费意愿逐年提升。



在数据库市场供给侧：分布式架构的技术特性趋于融合，云化部署大势所趋，使用门槛不断降低；数字化业务场景不断丰富，业务需求驱动HTAP技术成为重要的技术底座；在开源数据库领域，国产开源时间不长，整体表现尚有提升空间；国产数据库借信创东风，迎来了前所未有的发展机会。



在数据库市场需求侧：在数字化转型浪潮下，企业用户的数据已初具规模，业务场景和数据结构越来越复杂，对于数据分析的需求日趋强烈。数据库部署方式呈现多元化，其中分布式架构和云部署的渗透率显著提升。随着国产数据库的技术创新以及产品性能提升，国产产品的头部效应明显，用户体验仍有待进一步完善。



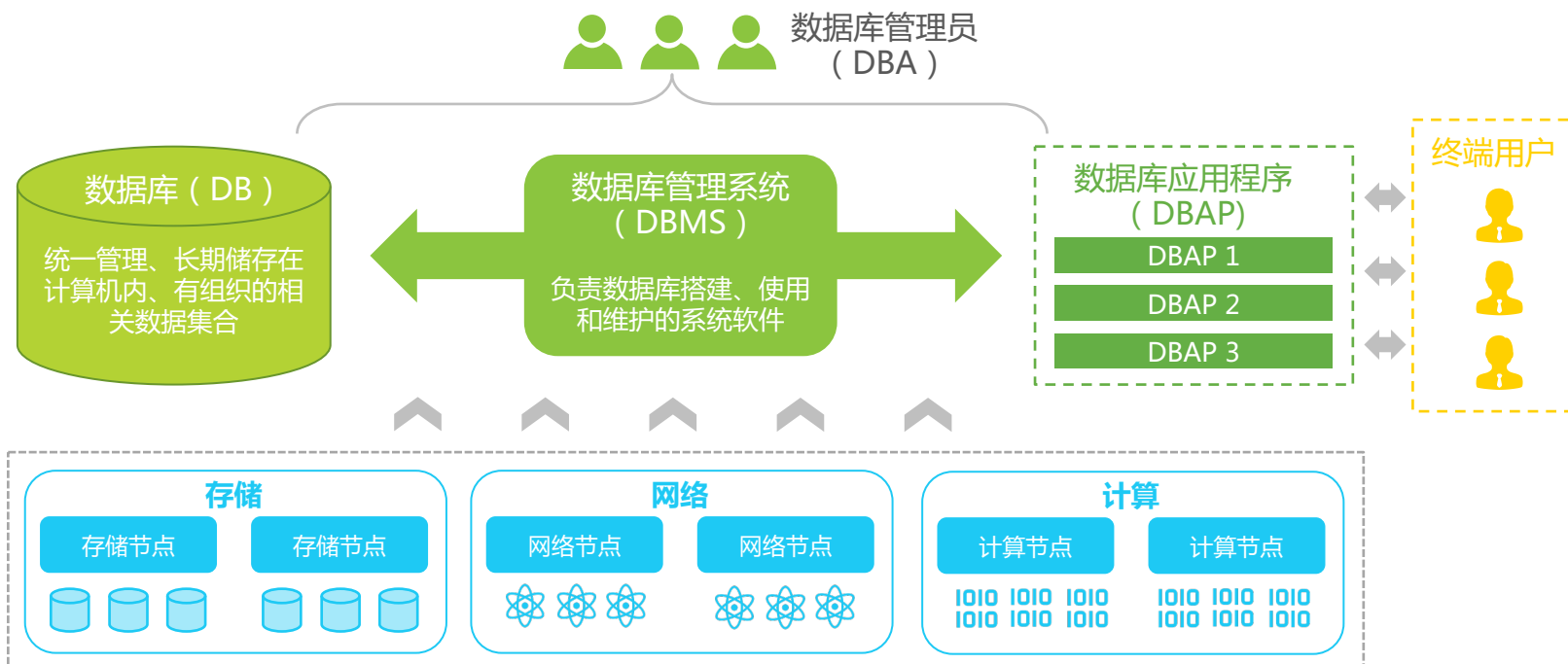
数据库行业趋势洞察：数据库出海成为国产厂商的全新增量逻辑，欧美线路和东南亚线路是国产厂商出海的优选项；云原生进入2.0时代，实现了云资源对数据库的细粒度资源拆解和优化编排，云托管数据库进一步向原生式数据库推演；数据库与大数据技术的边界不断模糊，呈现融合趋势。

数据库概述	1
数据库供给侧分析	2
数据库需求侧分析	3
典型企业案例	4
行业趋势洞察	5

数据库是由DBMS统一管理的，有组织、共享的数据集合

数据是描述事物的符号记录，具有多种表现形式，包括文字、图形、图像、声音和语言等。数据库是相互有联关系的若干数据的集合，特点是数据间联系密切、冗余度小、独立性较高、易扩展，并且可为各类用户共享。数据库管理系统（DBMS）是负责数据库搭建、使用和维护的系统软件，通过组织、索引、查询、修改数据库文件，实现数据定义、组织、存储、管理以及数据库操作、运行、维护等主要功能。围绕DBMS构建包含硬件操作系统、数据库管理系统、数据库管理员以及相关机制配套在内的数据库系统（DBS），才能有组织、动态地存储大量相关数据，方便多用户访问。

数据库系统架构

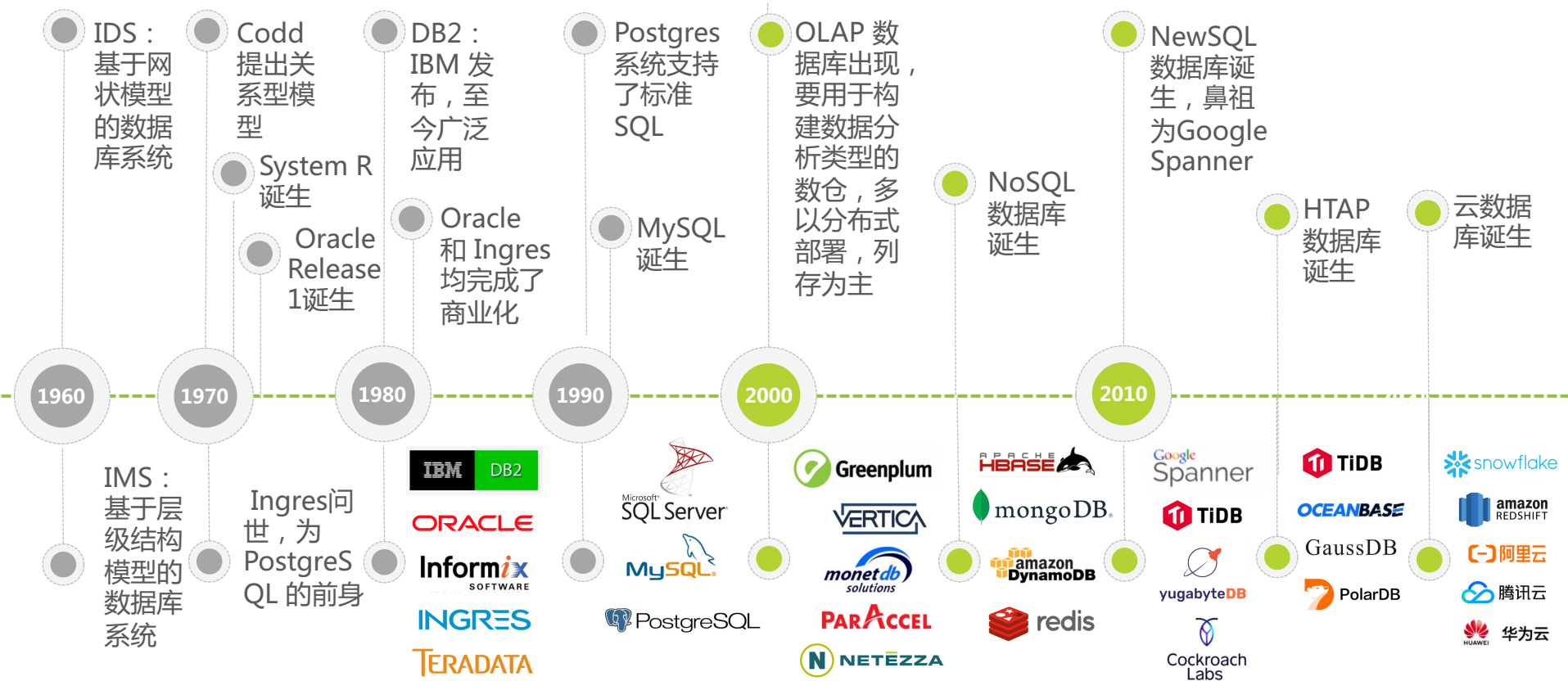


来源：公开资料，艾瑞咨询研究院整理及绘制。

发展历程

硬件技术、应用场景等快速变化，行业格局百花齐放

数据库行业发展历程



来源：公开资料，艾瑞咨询研究院整理及绘制。

分类

数据库按照不同维度，分类方法多种多样

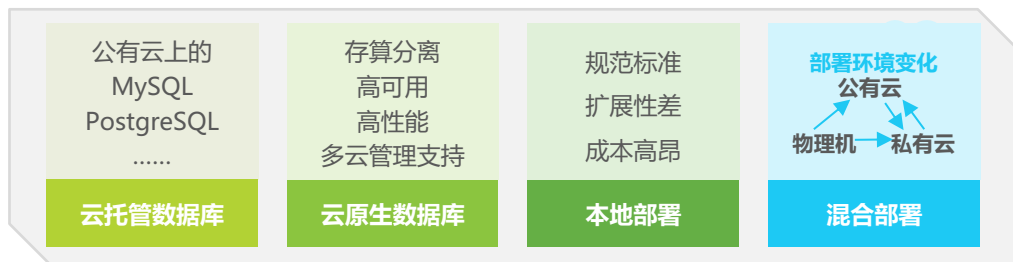
按数据结构分类



按架构分类



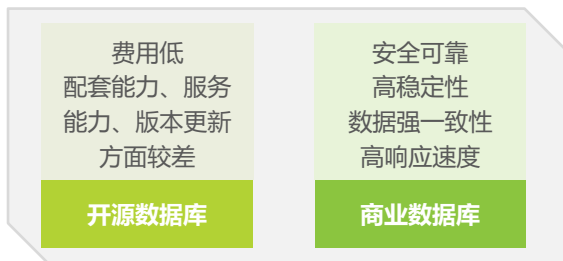
按部署模式分类



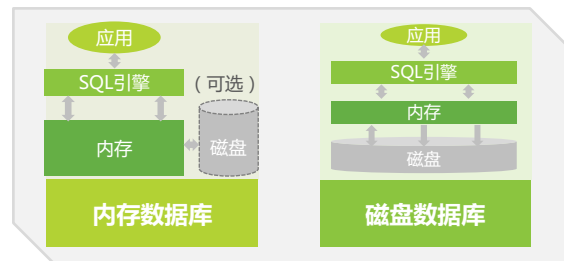
按功能分类



按商业模式分类



按存储介质分类



来源：艾瑞咨询研究院根据公开资料整理及绘制。

数据库概述

1

数据库供给侧分析

2

数据库需求侧分析

3

典型企业案例

4

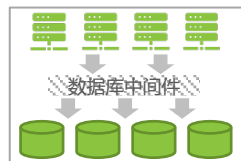
行业趋势洞察

5

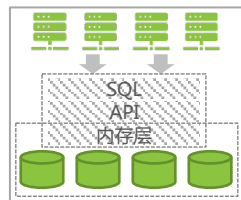
技术特性趋于融合，云化部署大势所趋，使用门槛不断降低

分布式数据库是通过计算机网络将物理分散的多个数据库单元连接组成的逻辑上统一的数据库。目前分布式数据库的技术路线分为三类：分库分表+中间件方案、共享存储分布式数据库、去中心化的分布式数据库。当下分布式数据库市场产品繁多，但其技术特性逐渐趋同，从两个角度分析：其一，系统架构。存储采用开源或者自研内核，用一主多从的数据副本保证数据完整性，计算通过 SQL 语句解析分发和结果汇聚，管控用于保证各模块的高可用和故障切换等管理功能。其二，开源和自研两条研发线路。开源主要基于 MySQL 和 PostgreSQL，难点在于源代码改造的技术难度大；自研周期长且技术要求高，国内完全自研产品屈指可数，如 TiDB 和 OceanBase 等。未来分布式数据库的核心竞争要素为分布式事务的稳定性、产品全面性以及生态工具的完备性。

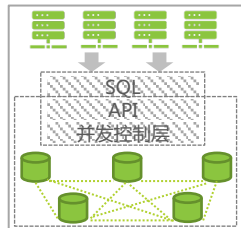
分布式数据库技术线路分类



- **方案：**分库分表+中间件，下层单机数据库提供存储，其上封装中间层补足分布式能力
- **优势：**兼容性好，学习成本低，用户门槛低
- **劣势：**代码修改成本高，计算节点扩展性、全局事务能力及高可用不佳，机器冗余度高



- **方案：**共享存储分布式，共享存储集群，提供动态扩容的分布式高性能存储，计算节点独立
- **优势：**兼容性好，应用无需改造；事务性能好，读写响应快；共享存储一致性高
- **劣势：**依赖共享存储系统，扩展性有限，难以做到跨地域高可用，架构可改造性较低



- **方案：**去中心化的分布式，各节点独立计算和存储功能，且不共享数据（shared-nothing），架构存算分离，通过共识算法保证副本高可用
- **优势：**架构高度解耦，按需扩展，无数量和规模限制，数据一致性的安全机制更好，高可用高兼容
- **劣势：**对硬件要求较高，上下游生态适配度仍需进一步完善

来源：艾瑞咨询研究院根据公开资料整理及绘制。

分布式数据库发展方向

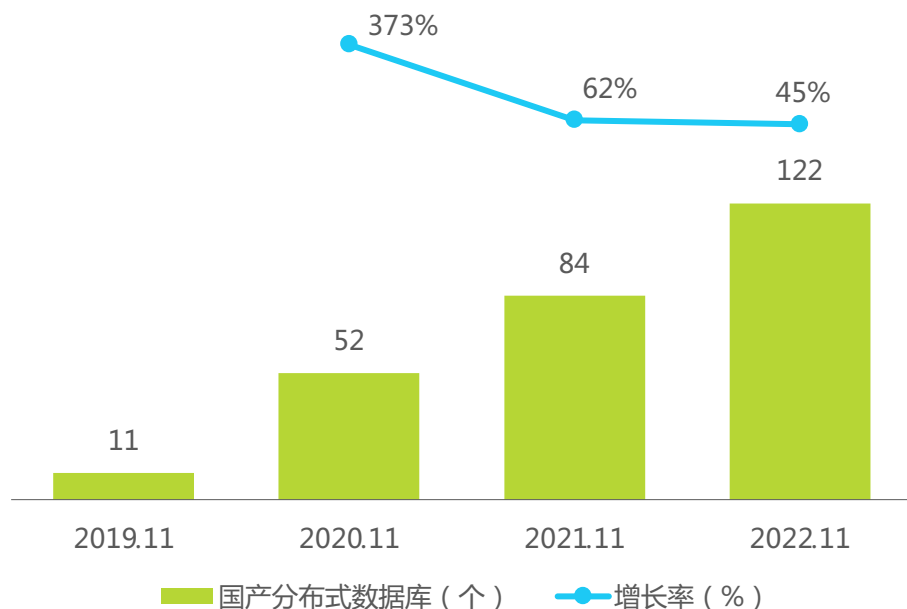
- 1 技术特性融合，能力补齐**
分布式数据库的不同技术路线逐渐融合，取长补短，丰富产品能力，集中体现于两方面：第一，分布式事务的一致性提升，全局一致性协议保护加强；第二，分布式场景下数据分片智能化，多副本细粒度控制能力得到加强。
- 2 云化与云原生设计**
分布式数据库与云环境加速融合。一方面，分布式的架构优势可将传统部署与公有云资源结合，解决私有部署灵活度低、成本高等问题；另一方面，云环境可低成本加强分布式架构的高容灾和算力扩展，融合云基础资源实现云原生。
- 3 软硬件结合，异构芯片混合部署**
以多核 CPU、异构计算（如 DSP、GPU、ASIC、FPGA）持久化内存等为代表新硬件，为分布式数据库提供丰富的想象空间。通过异构芯片的混合部署，让分布式数据库在不同层级上提供锁定模式的安全机制，提供分离模式的性能。
- 4 单体分布式一体化数据库架构**
在单机部署下实现分布式数据库的完整功能，从技术上突破分布式数据库的单机性能瓶颈，极大降低分布式数据库的部署门槛。无论用户自身规模大小，采用单机部署抑或水平扩展部署，一套数据库系统便可满足核心系统要求。

来源：艾瑞咨询研究院根据公开资料整理及绘制。

加强底层故障隐患感知，提升异常处置能力，实现开箱即用

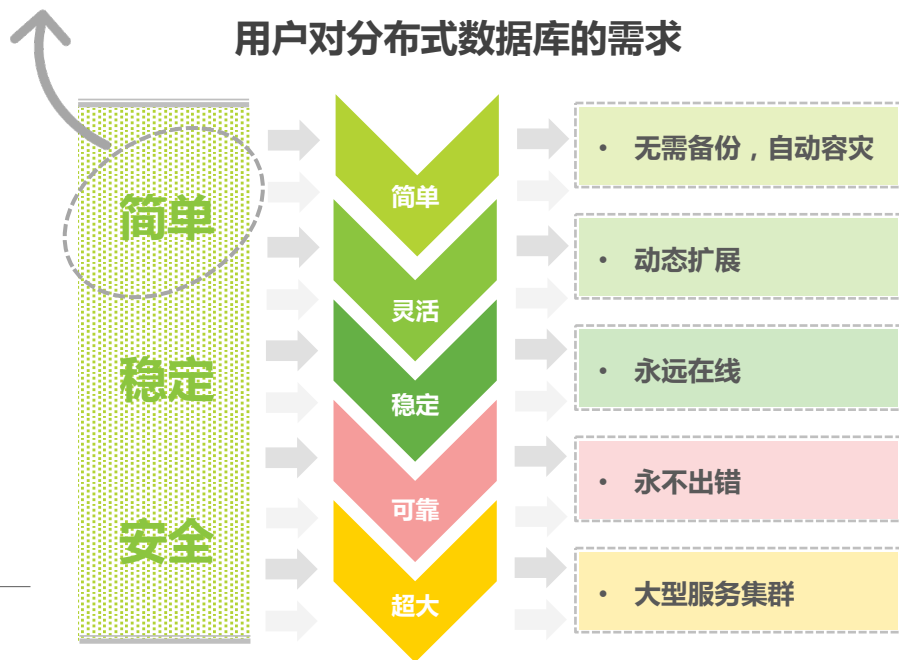
墨天轮数据显示，国产分布式数据库自2019年以来保持高速增长，热度持续高涨。而通过用户调研了解到，面对分布式数据库的故障时，运维人员往往束手无策。主要原因是相较集中式数据库，分布式数据库对 IT 基础设施可靠性要求更高了，其核心代码主要覆盖了 SQL 实现和数据存储，而未能自动感知各种对数据库稳定性、性能、并发能力有重大影响的故障隐患，也难以在代码层面对此类问题进行处理，从而实现故障自动规避。未来分布式数据库应提升基础数据探测和分析能力，随时针对出现的异常隐患提前处置，实现无需运维人员过多干预的高效自治运行，让分布式数据库从 IT 工程化产品变为真正开箱即用的通用型产品。

2019-2022年国产分布式数据库规模及增速



来源：墨天轮，艾瑞咨询研究院整理及绘制。

用户对分布式数据库的需求

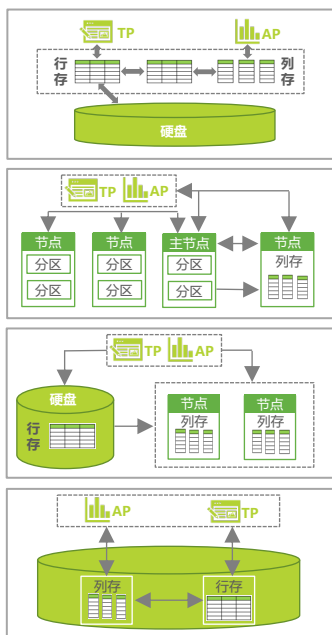


来源：艾瑞咨询研究院根据公开资料整理及绘制。

业务驱动HTAP技术成为重要底座，与分布式技术天然适配

随着业务系统接入数据源复杂性提升，混合负载需求愈发强烈，数据库技术正向多源异构、高实时并发、多 SQL 标准接口的方向演进。过去，HTAP 数据库的技术架构以扩展主行存技术，在行存基础上加列存的方式为主，典型代表有 SQL Server, Oracle 和 L-store 等。随着分布式技术不断成熟，天然适配 HTAP 数据库，于是开启了分布式的架构实现，满足了高并发需求，典型代表有 SingleStore、MySQL Heatwave 和 Greenplum。我们认为未来 HTAP 仍应是在 OLTP 数据库的基础上，引入原生分布式架构和低成本存储引擎以扩展大数据量 OLAP 能力。最后，为更好支持 OLAP 的数据开发和建模能力，提升实时分析，HTAP 支持物化视图和外部表，并与各类数据开发工具和 BI 工具适配对接。

现阶段 HTAP 数据库主要架构



主行存储+内存中列存储

- 特征:** 主行存为TP负载，内存列存为AP负载，行存经内存优化后高效将增量存储并至列存
- 优势:** TP和AP高吞吐，数据新鲜度高
- 劣势:** AP扩展性差，负载隔离性差

分布式行存储+列存储副本

- 特征:** 依赖分布式架构，主存储为行存，部分从节点作为列存服务器为AP查询加速
- 优势:** 高扩展性，负载隔离性高
- 劣势:** 数据新鲜度低

磁盘行存储+分布式列存储

- 特征:** RDBMS为TP负载，列数据从行存提取，热数据留在IMCS，冷数据移至磁盘
- 优势:** TP和AP高吞吐，高扩展，负载隔离性好
- 劣势:** 数据新鲜度低

主列存储+增量行存储

- 特征:** 主列存为AP负载，增量行存处理TP负载
- 优势:** AP高吞吐，数据新鲜度高
- 劣势:** TP扩展性差，负载隔离性差

HTAP 面临的主要挑战



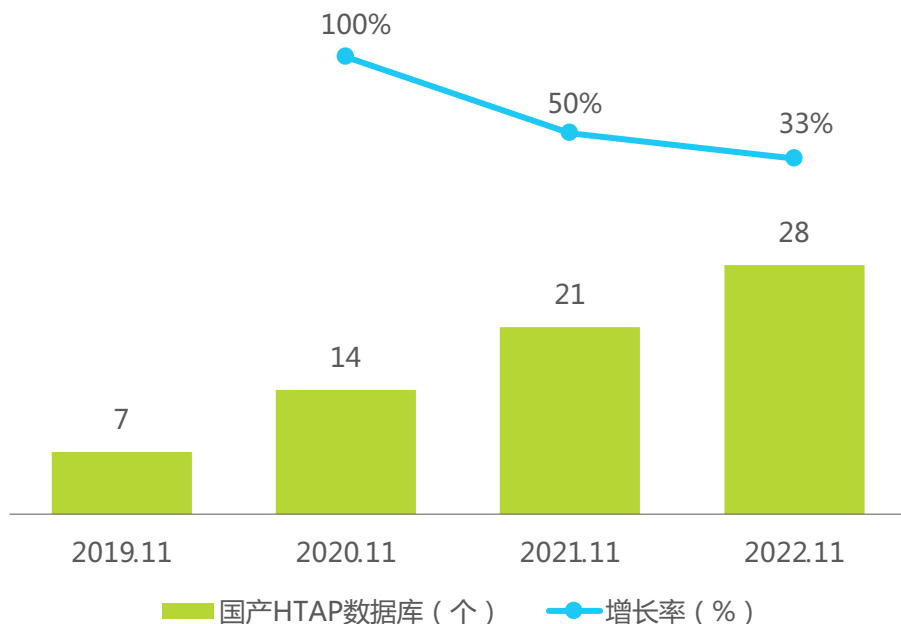
来源：公开资料，艾瑞咨询研究院整理及绘制。

来源：公开资料，艾瑞咨询研究院整理及绘制。

运用内置流处理能力进一步加强AP和TP间的融合

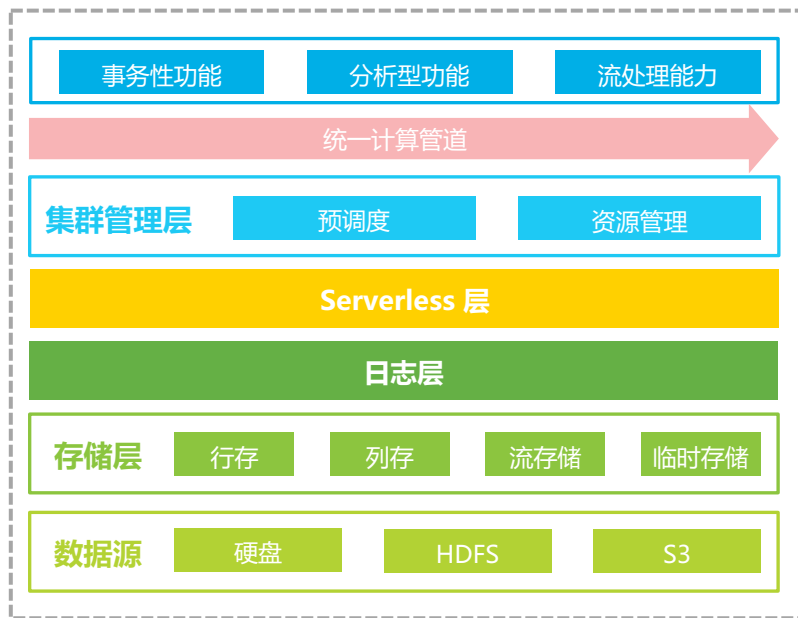
墨天轮数据显示，自2020年以来国产 HTAP 数据库发展迅猛，几乎成为新兴数据库的必选项。Gartner 也指出 HTAP 已经成为全球范围内新一代数据库的入场筹码，HTAP 能力成为数据库重要选项。随着数据技术的不断创新，未来 HTAP 数据库会进一步加强 AP 和 TP 之间的连接和融合，在数据库内部实现 AP 和 TP 之间的内置流（Streaming）处理能力。通过将 ETL 工作内置于数据库当中，让 HTAP 数据库同时具备 AP、TP 和流（Streaming）能力。用户可以按需创建各类表，运用流处理能力将表连接，从而获得简单、便捷的数据处理能力。数据库技术的进一步融合会打破当下数据栈的割裂状态，HTAP、流批一体、湖仓一体等技术趋势最终会让数据集中在简单易用、安全可靠、高性价比的数据平台。

2019-2022年国产 HTAP 数据库规模及增速



来源：墨天轮，艾瑞咨询研究院整理及绘制。

基于内置流处理能力的 HTAP 架构

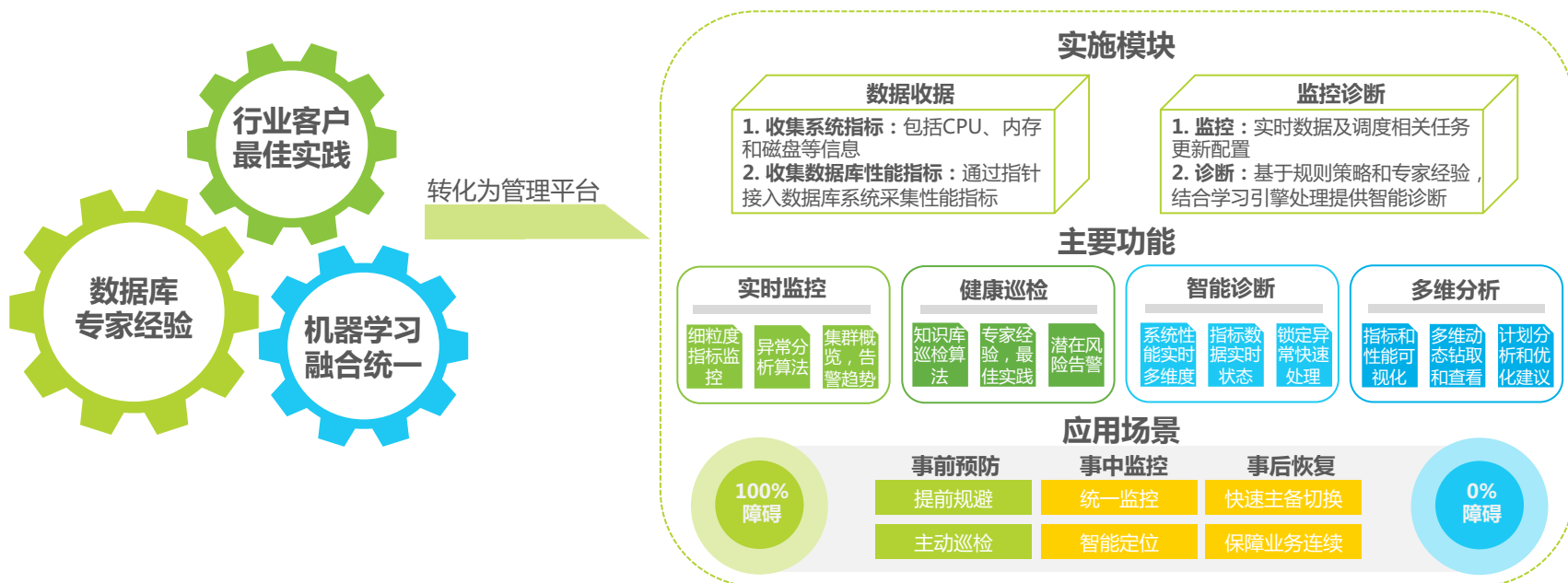


来源：艾瑞咨询研究院根据公开资料整理及绘制。

融合智能技术实现复杂数据库环境的管理自治

DT 时代数据库技术架构和运行环境日趋复杂，种类从单一产品转变为混合型商业数据库和开源数据库组合，依靠人工运维显得捉襟见肘。作为智能化的数据库周边工具，数据库管理平台将机器学习与数据管理在功能上融合统一，利用机器学习增强系统设计开发，以标准化、自动化、智能化的方式提供实时监控、健康巡检、智能诊断、多维分析等数据库管理服务。数据库管理平台的本质是数据库管理经验的代码化，核心方法论是云资源池化、分层解耦以及服务化，实现手段是基于微服务、分布式等云技术实现多元混合数据库环境的统一管理，目标是实现海量数据高安全、高可用、高性能的运维要求，助力数字化转型。

智能化的数据库管理生态

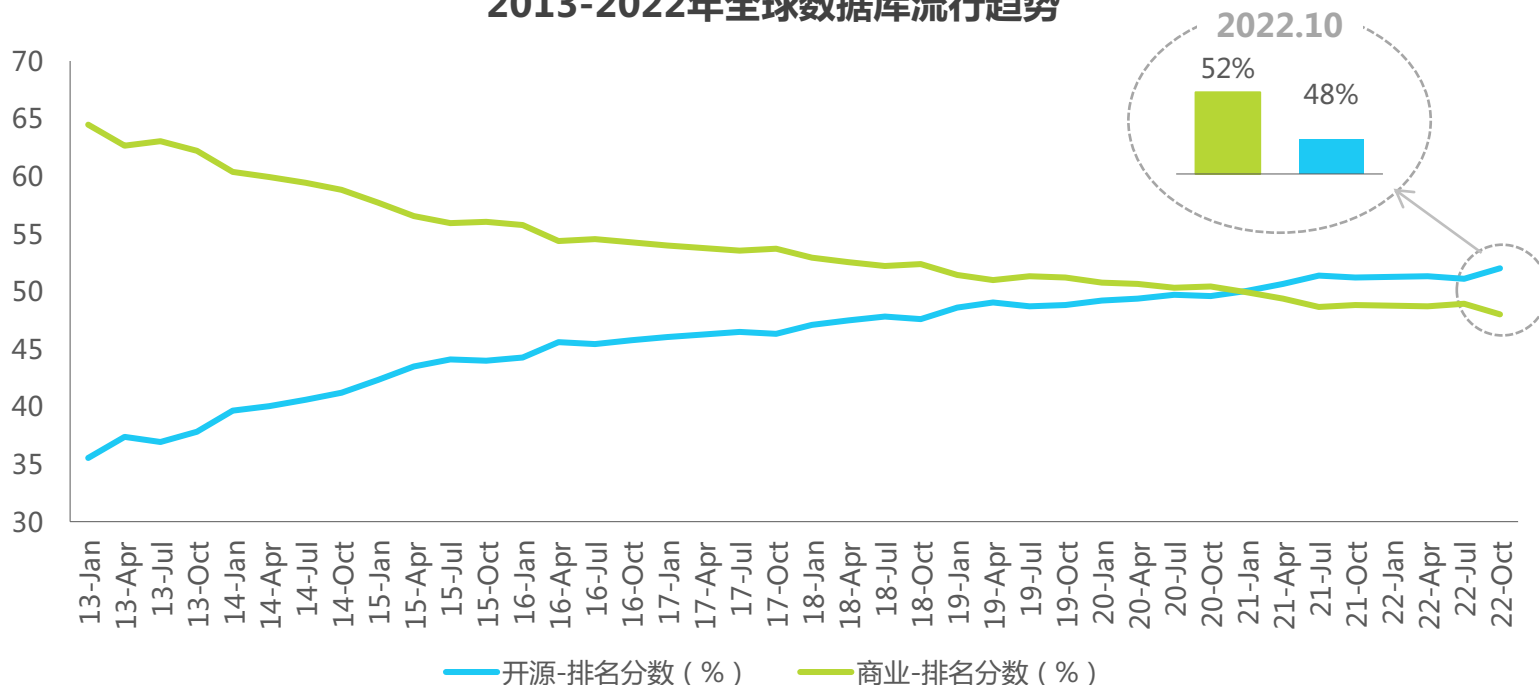


来源：艾瑞咨询研究院自主研究及绘制。

热度持续，成为国产数据库走向国际化的重要途径

开源即开放源代码，用户可在源代码的基础上完成学习和修改。DB-Engines数据显示，截至2021年1月，开源数据库的全球部署首次超过商业数据库。此后，开源数据库凭借在成本、产品丰富度、社区活跃性等方面的突出优势，持续获得了全球的高关注度。开源成为驱动数据库技术和市场变革的重要力量。在此趋势下，开源成为数据库项目冷启动的最佳方式，也是构建国产自研数据库生态、加速技术迭代、让国产数据库走向国际化的重要途径。

2013-2022年全球数据库流行趋势

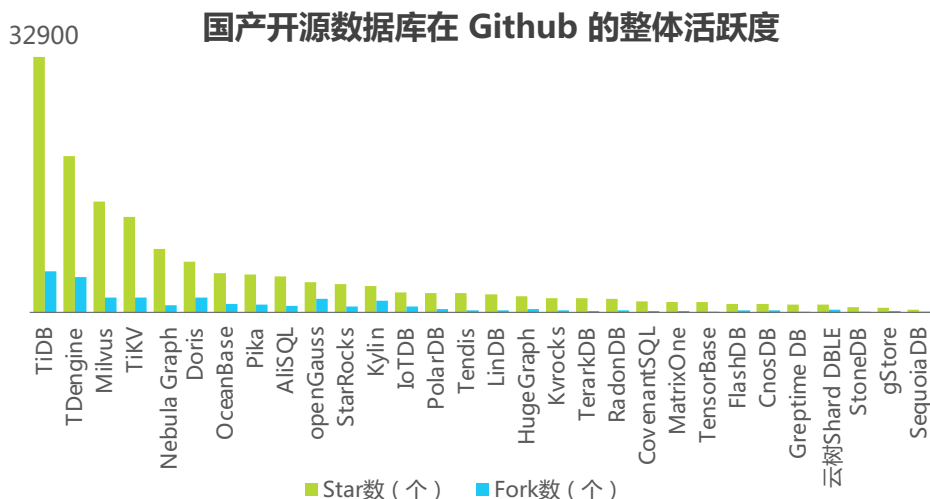


注释：1、DB-Engines排名是按照当前流行程度的排名，较实际使用情况具有一定的超前性，具体指标包括网站上系统提及的频次、Google trends、IT论坛上系统讨论的频率、提及系统的工作机会数量、专业网络中提及系统的配置文件数和系统在社交网络中的提及次数；

2、开源、商业-排名分数(%)为相对值。

来源：DB-Engines，艾瑞咨询研究院自主研究及绘制。

国产数据库开源时间不长，整体表现有待提升



近年来国内数据库厂商逐步推进开源策略。但因国内数据库开源时间不长，运营有待加强，整体表现有较大提升空间。以 TiDB 为例，从开发者、使用者及代码角度分析，其开发者活跃、用户众多、代码变化较大，平均2-3年完成一次软件重构，侧面印证了上述特征。

国产数据库开源的必要性和发展趋势



必要性 发展趋势

获取市场信任

基于国产数据库行业现状，只有选择开源，彻底打开“黑盒”，才能真正获取用户信任，打开市场局面

开发方式被重塑

开发者的用户体验得到高度关注，开发者的门槛不断降低，只需具备前端能力，其他都由全链路的可伸缩服务替代

构建成熟的数据库生态

单纯依靠个别技术或软件很难构建成熟、完善的产品，只有选择开源，依赖社区生态推进，才能打磨出色的数据库产品

Serverless

基于云原生的 Serverless 生态和技术。各链路 Serverless 解决方案串联成全栈 Serverless 体验

满足业务场景敏捷性

闭源数据库厂商的视角相对滞后，产品反馈链条较长，只有选择开源，才能及时捕获业务场景的敏捷变化

从 SQL 到 API

将数据库用户定位至应用开发者。简化 API 封装和使用，拉近 SQL 到实际应用的距离，让应用开发友好性最大化

国产开源数据库在 Github 的活跃度细分情况

项目	开发活跃度			使用者维度			代码维度		贡献者维度
	active	merge	open	active	close	new	新增	删除	数量
TiDB	604	392	212	599	271	328	123161	24239	1146
OceanBase	13	4	9	45	9	36	125473	28777	74
openGauss	44	-	-	198	1	1	-	-	121
TDengine	283	246	37	65	36	29	25943	9269	140
PolarDB	5	5	0	4	1	3	2908	368	41

来源：Github，艾瑞咨询研究院根据公开资料自主研究及绘制。

来源：艾瑞咨询研究院根据公开资料自主研究及绘制。

中国数据库产业图谱

中国数据库产业图谱

国内厂商

传统厂商



云厂商和互联网厂商



新兴厂商



跨界厂商



国外厂商

商业数据库



开源数据库



数据库管理

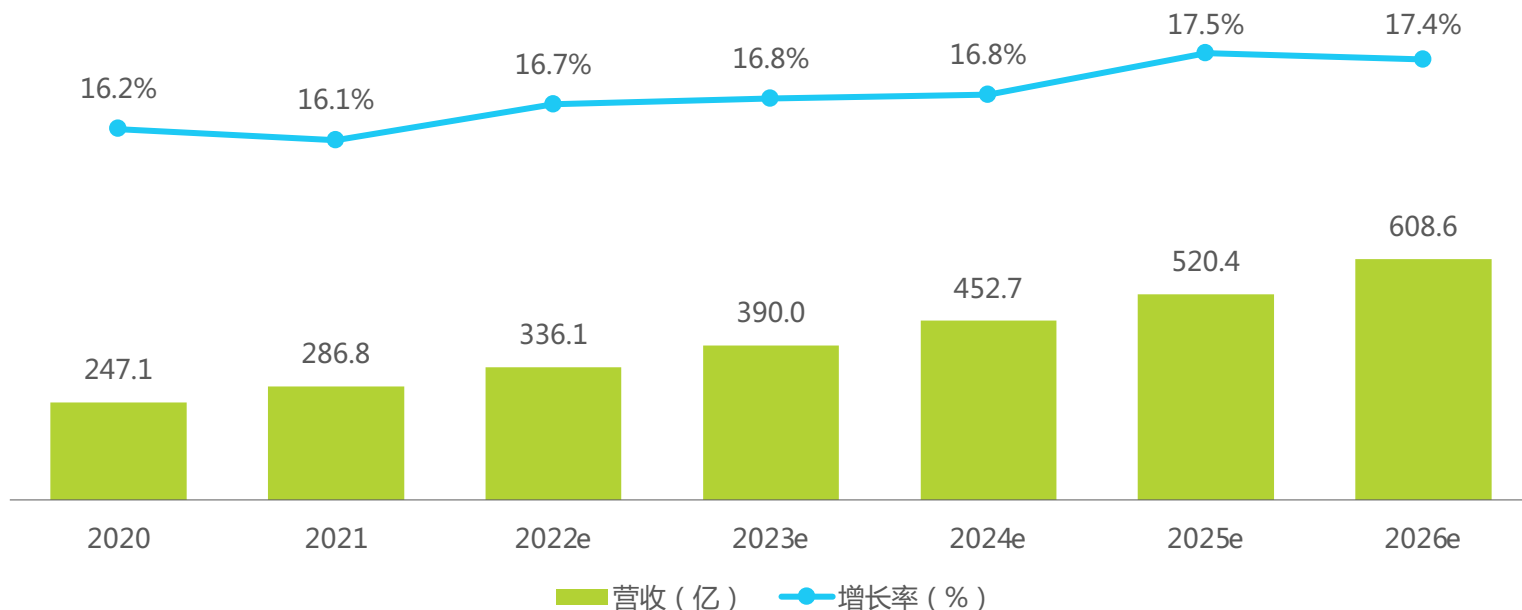


来源：公开资料，专家访谈，艾瑞咨询研究院自主研究及绘制。
注：图谱仅展示部分典型厂商，图谱展示顺序不代表市场份额排名。

受政策驱动和需求催化，2021年数据库市场规模达286.8亿

据艾瑞统计，2021年中国数据库市场总规模达286.8亿，较2020年增长16.1%，CAGR（2021-2026e）达13.4%。中国数据库市场虽受疫情影响，但整体稳步增长，由多方面因素促成：1）信创利好，国家对国产数据库的支持力度大；2）需求催化，数字化业务场景带动数据库多元化发展；3）供给侧厂商厚积薄发，技术创新涌现，产品性能显著提升；4）国内用户对基础软件的IT支出和国产数据库的付费意愿逐年提升。

2020-2026中国数据库市场规模及增速



注释：市场规模统计口径为国内外厂商在中国数据库销售的营收。其中包括DBMS基础软件的收入，必要配套工具的收入（数据迁移、数据备份等工具），项目定制化开发、实施、运维等服务的收入，数据库软硬一体机的收入；不包括单独售卖的硬件的收入，单独售卖的大数据平台的收入、单独售卖中间件及应用软件的收入。

注释：此处市场规模中包含云厂商托管开源数据库（MySQL、PostgreSQL、MongoDB、Redis等）所得的收入。

来源：根据公开资料、企业访谈，结合艾瑞统计模型核算。

发展特点（一）

现阶段集中式数据库仍占据主流，分布式有待场景需求拉动

据艾瑞调研，若剔除数仓（绝大多数数仓采用分布式部署），集中式数据库的市场份额接近80%，而分布式部署不足20%。从技术角度看，虽然集中式数据库的水平扩展有限，但优点是足够简单和易用，在小规模数据量的场景下性能表现优异，而绝大多数业务场景的数据量并未大到必须使用分布式架构承载。分布式数据库的特征则恰好相反。从应用角度看，分布式部署更受行业头部用户青睐，其中信创环境下的国产化替代风潮起到了很大的推动作用。但运行场景多以非核心业务为主，用户的核心业务仍通过集中式数据库运行。因此，我们认为未来分布式数据库的发展一定是由业务场景的需求拉动。

2021年中国数据库市场份额：集中式vs分布式



集中式数据库

- 从传统行业的整体应用情况看（互联网行业自身基本采用分布式部署），数据库部署类型以集中式为主，产品类型以 Oracle 为主。
- 从传统行业的细分情况看，金融行业对数据安全可靠的要求极高，容错率极低，对集中式数据库高度依赖，部署占比80%左右；政务行业受信创驱动，国产化替代率高，集中式占比60%左右；零售行业头部企业的集中式占比与政务行业接近，在60%左右。



分布式数据库

- 从传统行业的整体应用情况看，现阶段分布式部署的占比较低，主要集中于互联网行业本身，占比20%左右。产品以 MySQL 和 PostgreSQL 为代表的各类开源数据库为主。
- 当前分布式数据库在传统行业的渗透率有待进一步提升，但随着数字化业务场景的深入，未来市场空间较为广阔。从使用热度上看，TiDB、OceanBase、GaussDB、PolarDB 等国产数据库产品在全球和中国的热度都经年不少。

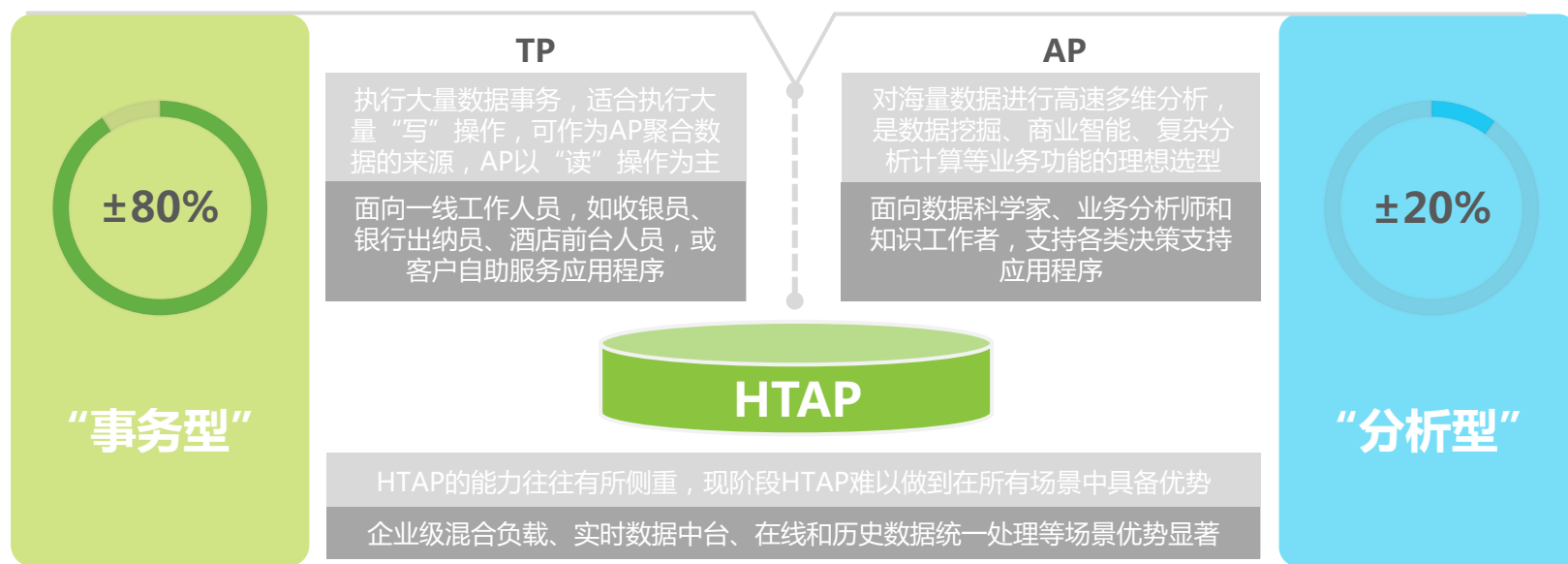
来源：专家访谈、公开资料，艾瑞咨询研究院自主研究及绘制。

发展特点（二）

事务型数据库增速放缓，分析型和HTAP数据库不断涌现

从收入规模来看，事务型数据库的市场份额占主导地位，分析型数据库占比不足20%。但预计未来3-5年内，分析型数据库的市场份额会有较为显著的提升，原因如下：各行业的数字化进程已取得了初步成果，企业核心业务的信息化过程主要是运用关系型数据库承载的事务类业务，因此当下事务型数据库增速已有所放缓。相反随着业务场景的丰富和数据复杂度的提升，诸如市场推广和营销分析类的分析型场景剧增，各类分析模式和分析方法不断涌现，最终都需要具体的分析型数据库执行。此外，HTAP 理念也受到了热捧，但根据调研了解，HTAP 技术线路和产品实现尚未达成行业共识。

2021年中国数据库市场份额：事务型vs分析性



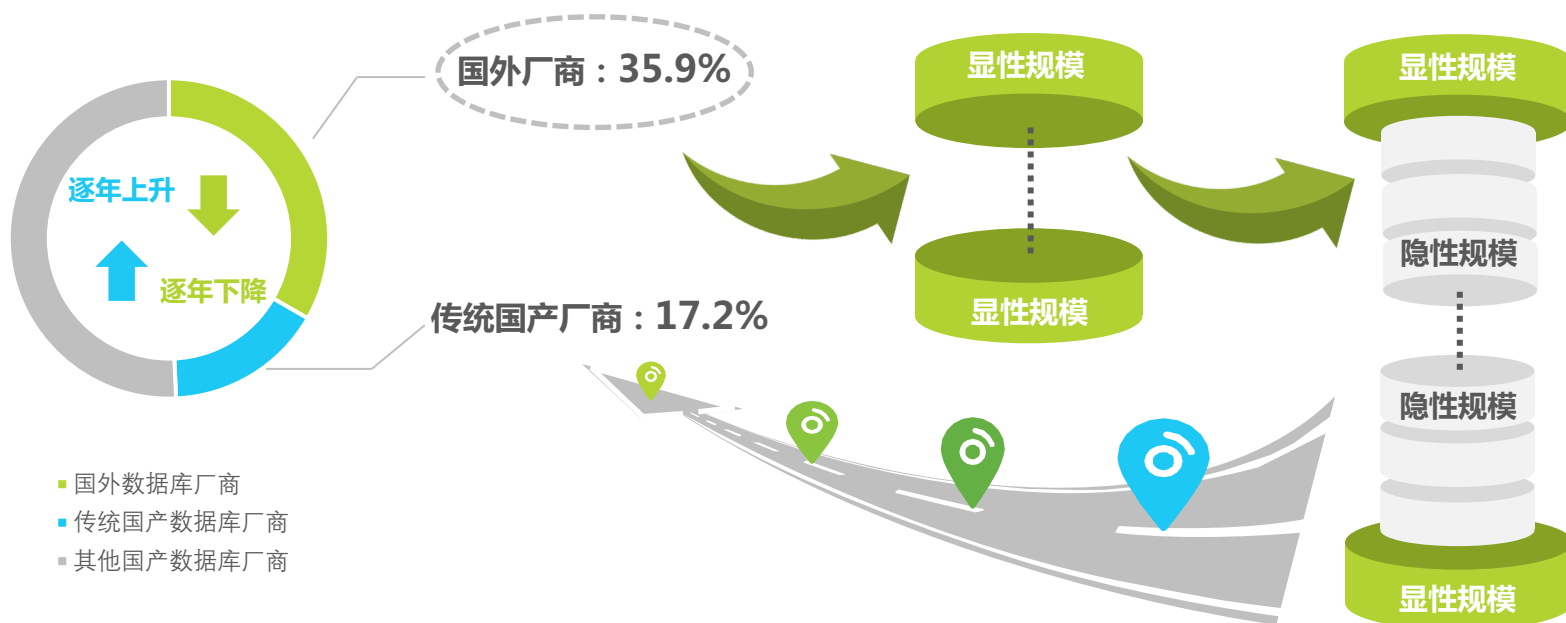
来源：专家访谈、公开资料，艾瑞咨询研究院自主研究及绘制。

发展特点（三）

信创为国产数据库厂商提供发展沃土，未来市场空间可期

在国内数据库市场，传统的商业数据库有两类玩家：其一，以 Oracle 为代表的国外厂商；其二，以达梦、人大金仓、南大通用、神舟通用为代表的国产厂商。二者的此消彼长一方面源于国产厂商技术实力的不断增强，另一方面也反映了信创东风下，国家政策对以数据库为核心的基础软件领域的大力扶持。在此，我们希望通过分析国外厂商在国内数据库市场的商业模式，以反映未来国产化替代过程中，国产数据库厂商的市场空间之巨：据艾瑞调研了解，一直以来国外厂商对国内用户的限制较低，导致在国内数据库市场的隐性规模尚未浮出水面。随着国产替代进程的不断深入，国外厂商的隐性市场份额被逐步释放，国产厂商潜在的市场空间可期。

2021年中国数据库市场份额：国产vs国外



来源：专家访谈、公开资料，艾瑞咨询研究院自主研究及绘制。。

数据库概述

1

数据库供给侧分析

2

数据库需求侧分析

3

典型企业案例

4

行业趋势洞察

5

中国数据库用户使用情况调研说明

01 研究目的及内容

为全面了解我国各行业数据库应用情况，深入分析当前我国数据库用户的真实需求和使用痛点，从需求侧洞察产业发展现状及未来趋势，以推动我国数据库产业健康有序发展。

本次研究采用综合定性和定量问卷调研的研究方法，面向国内不同行业数据库使用单位中的CTO、CIO、架构师、工程师、数据库管理员（DBA）、产品经理等角色，围绕企业业务特征、数据结构特征、数据使用场景、数据库产品选型要素及使用情况等内容展开分析，以期加深对行业及用户的了解。

01 研究目的及内容

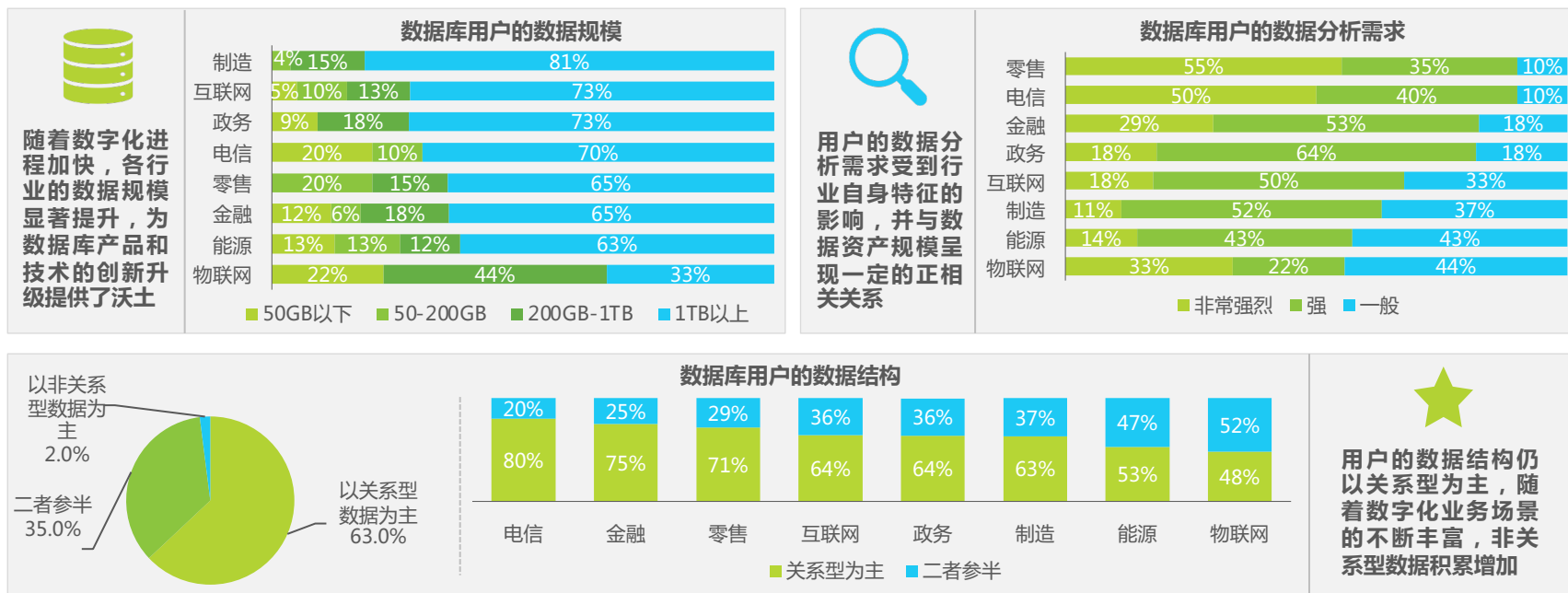
调研概况	描述
样本来源	艾瑞 iClick 在线调研社区
调研时间	2022.11
覆盖地区	全国
有效样本数量	问卷调研国内数据库用户172家

数据初具规模，数据结构向非关系型倾斜，分析需求增加

随着国内数字化转型节奏加快以及国产替代浪潮涌动，基础软件领域迎来了高速发展。数据库作为信息系统核心软件以及信创的关键环节，迎来产业黄金期，数据库产品的国内渗透率达到前所未有的高度。随着国内数据库用户不断增加，用户画像逐渐清晰，描绘出符合中国特色的数据库用户特征。

数据库用户的数据资产积累已初具规模，尤其是互联网、制造、政务、电信等领域多半已达到TB级别。用户的数据结构仍以关系型数据为主，但非关系型数据的比例有明显提升。用户的数据分析需求存在较大的行业偏差，互联网、金融、电信、零售、政务等领域需求较强烈，而能源、制造、物联网等领域相对较弱。

2022年中国数据库用户画像



来源：艾瑞咨询2022年数据库使用情况调研问卷数据。
样本：N=172；于2022年11月通过艾瑞 iClick 在线调研社区获得。

角色定位

数据库使用单位中不同角色对产品的考虑要素有所差异

数据库使用单位中各职能角色的考虑要素



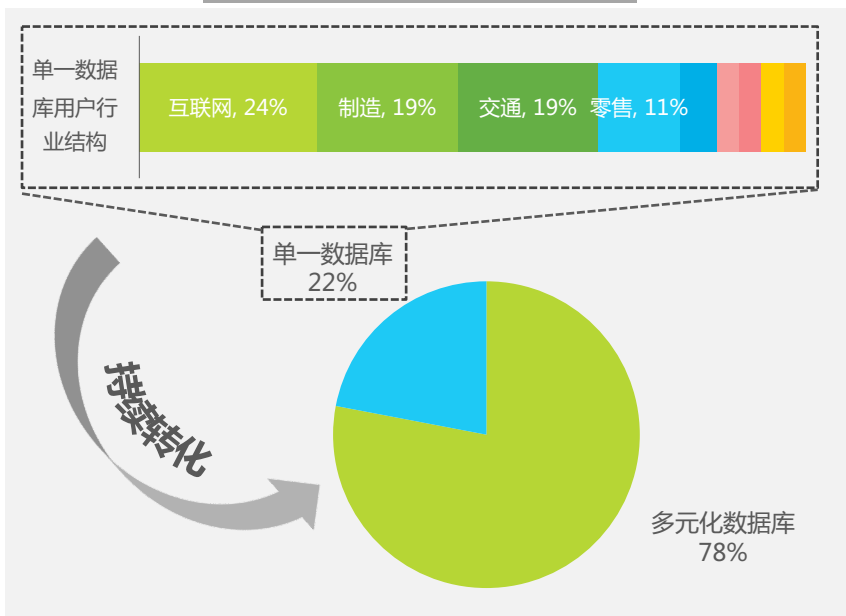
来源：艾瑞咨询2022年数据库使用情况调研问卷数据。
样本：N=172；于2022年11月通过艾瑞 iClick 在线调研社区获得。

数据库趋于多元化部署，分布式架构和云部署渗透率提升

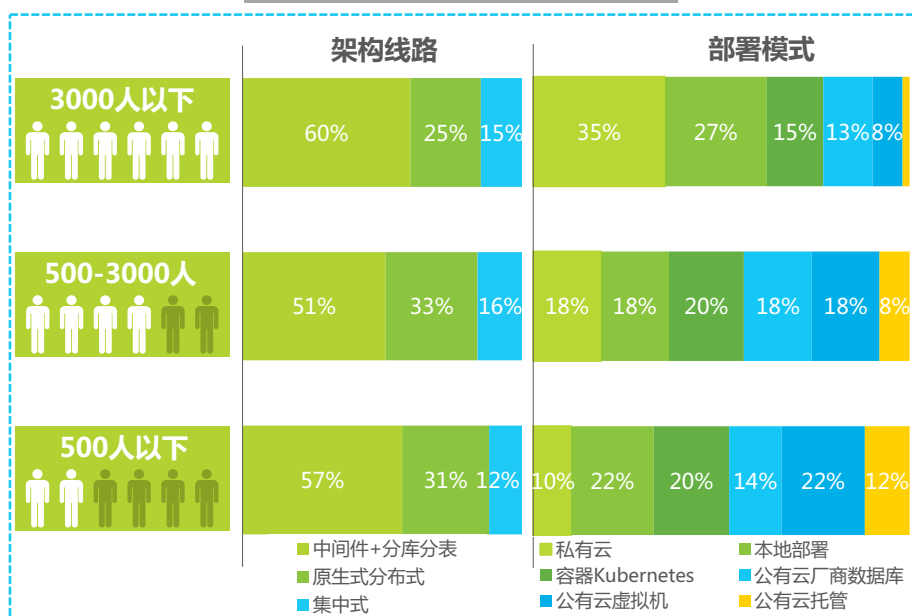
近年来，数据复杂度以及对数据的海量、多模、异构、弹性等要求不断提高，单一数据库难以满足当下的用数需求，数据库的多元化趋势显著。从调研数据可见，架构师和开发人员基于前端的数据使用需求呈现多样化，在研发规划之初便选择不同的数据库以满足开发需求。随着数据库技术的更新迭代，用户的选择灵活度更大，除功能、安全等传统因素之外，可从架构、研发、运维等多个维度考虑。在架构选型上，分布式架构的整体渗透率已超80%，其中仍以中间件+分库分表的实现方式为主。从部署模式看，云数据库的使用率接近80%，其中公有云占比60-70%。

数据库用户的产品部署特征

数据库用户的多元化部署



不同规模数据库用户的产品部署



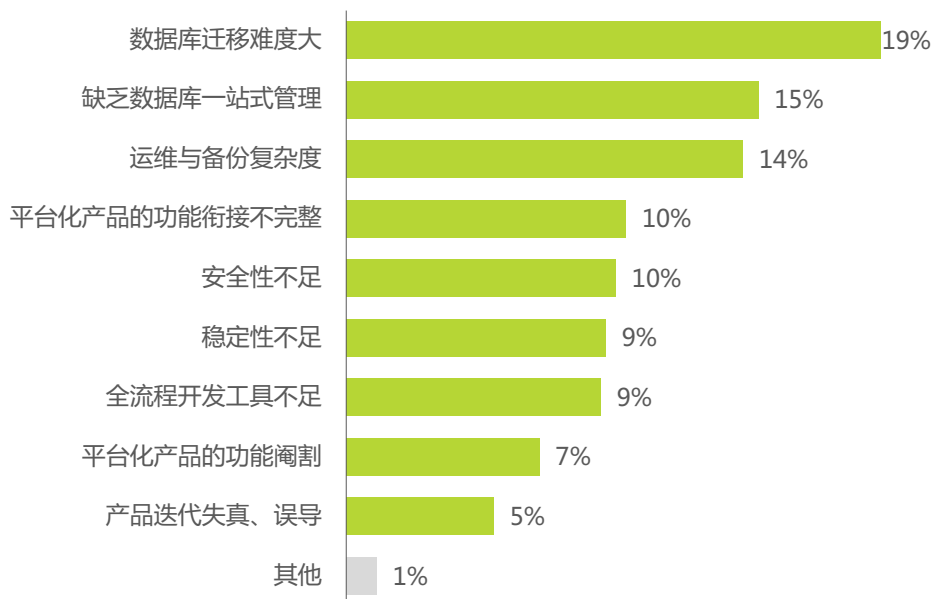
来源：艾瑞咨询2022年数据库使用情况调研问卷数据。N=172

注释：问题：请问企业当前使用的数据库有哪些？请问企业使用分布式数据库的情况？请问企业使用云数据库的情况？

国产产品渗透率的头部效应明显，用户体验有待进一步完善

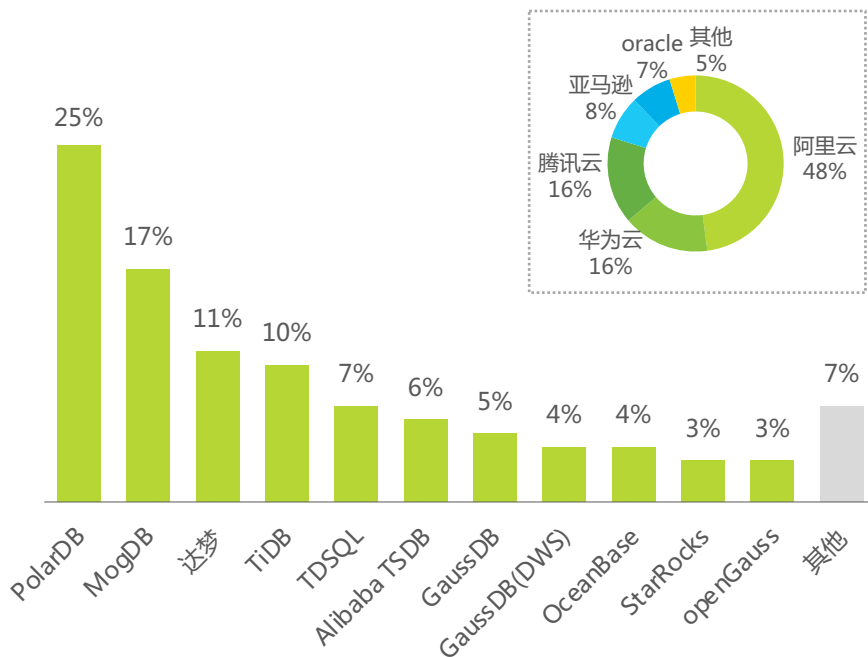
在调查数据库用户的使用体验时，我们收获的反馈相对集中：第一，数据库迁移难度大，占比19.2%，难点聚焦于迁移过程中的兼容性、数据安全、停机时间、数据校验和性能保证等。第二，缺乏数据库的一站式管理，占比14.8%，当前市场上的一站式管理平台以管控类为主，主要面向DBA等专业技术人员，而对开发、业务、运维等其他角色的门槛偏高。第三，运维与备份的复杂度高，占比14.3%，随着技术路线的灵活度和复杂度提升，运维与备份和难度势必增加。具体到国产数据库的市场渗透情况，PolarDB 排名第一，MogDB 排名第二，达梦排名第三，TiDB排名第四。在云厂商中，阿里云的市场份额排名第一，华为云和腾讯云并列第二，亚马逊排名第三。

数据库用户的使用体验



来源：艾瑞咨询2022年数据库使用情况调研问卷数据。N=172
注释：问题：请问企业使用数据库的体验如何，有哪些痛点？

国产数据库的市场渗透情况

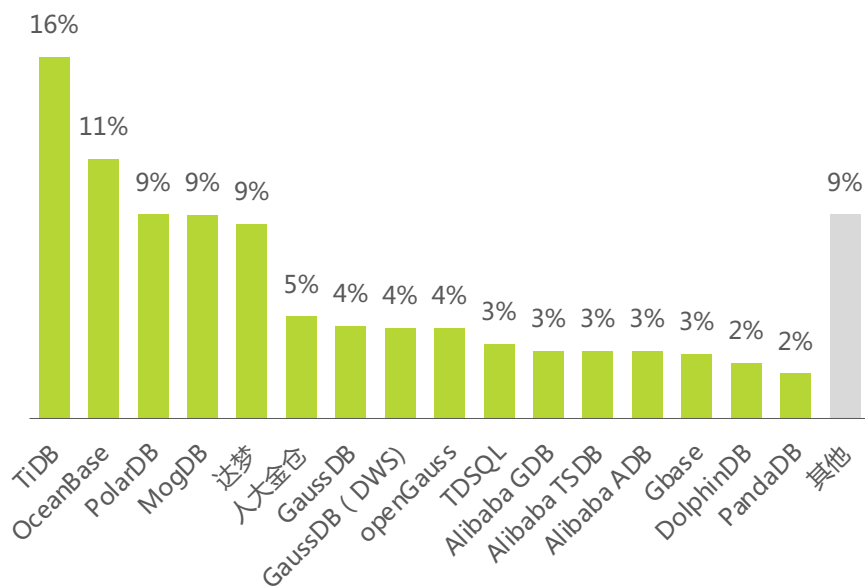


来源：艾瑞咨询2022年数据库使用情况调研问卷数据。N=172
注释：问题：请问企业使用的国产数据库有哪些？云数据库使用者选择了哪些云厂商？

深化多元化部署，技术选型更加注重安全稳定与服务能力

在调研数据库用户的未来部署规划时，我们发现超过60%的用户计划未来使用更多类型的数据库产品，以满足日趋复杂的数字化业务场景。在产品架构选择上，以单机或集中式架构为主的数据库用户中，超过50%计划未来选用分布式数据库，其中计划采用原生式和中间件+分库分表的比例接近1：1。而在产品部署形态上的调研结果较为有趣，当前采用本地化部署形态的数据库用户中，超过60%的用户表示未来上云的意愿不强。究其原因，主要是市场环境 with 体制因素导致的国内外云计算市场的显著差异，具体体现为传统企业上云以项目制、定制化解决方案为主，以混合云带动公有云发展，理论发展前景大，但实际增速明显低于海外。

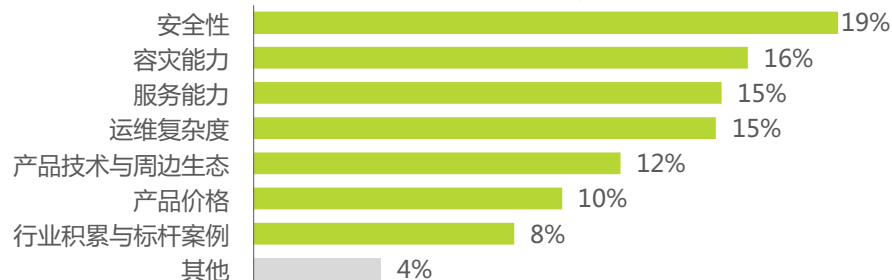
用户未来规划中的国产数据库热度汇总



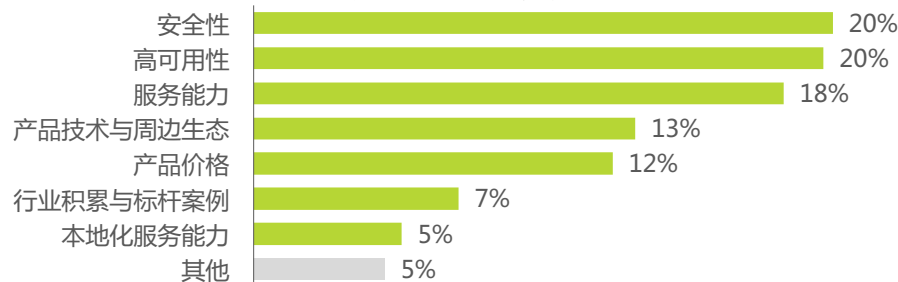
来源：艾瑞咨询2022年数据库使用情况调研问卷数据。N=172

注释：问题：在企业计划使用更多数据库的前提下，请问企业未来计划使用或感兴趣的数据库有哪些？

分布式数据库选型要素



云数据库选型要素



来源：艾瑞咨询2022年数据库使用情况调研问卷数据。N=172

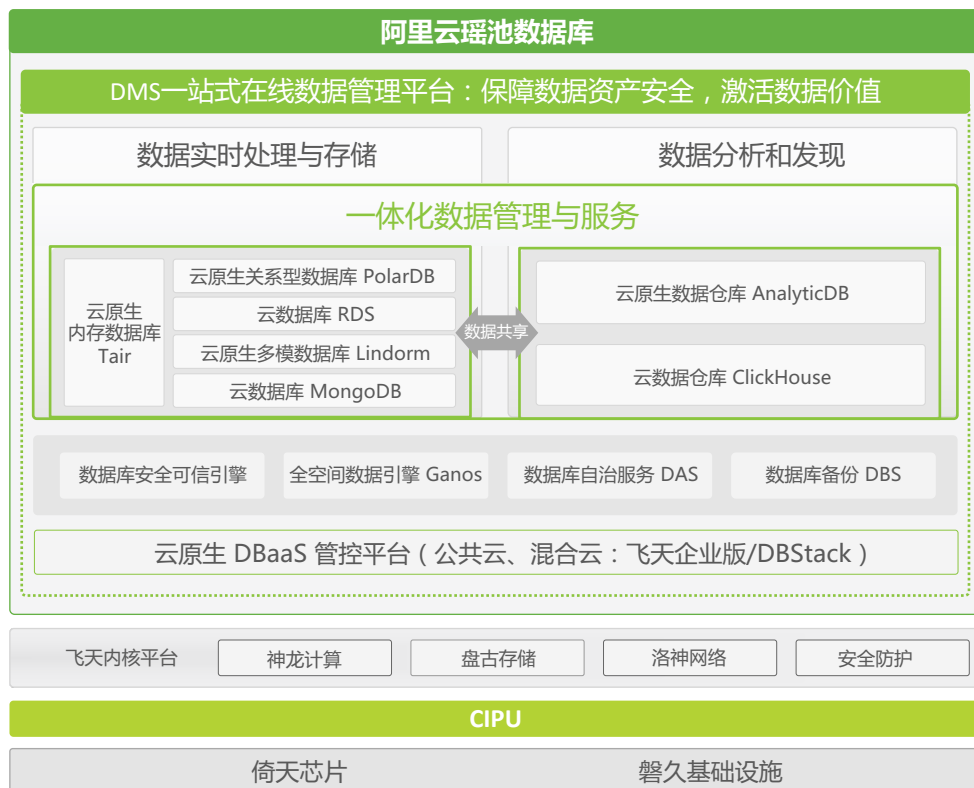
注释：问题：企业在分布式数据库/云数据库选型时的主要考虑因素？

数据库概述	1
数据库供给侧分析	2
数据库需求侧分析	3
典型企业案例	4
行业趋势洞察	5

云原生数据库2.0——云原生一站式数据管理与服务

阿里云瑶池数据库涵括关系型数据库、分析型数据库、NoSQL数据库、数据库生态工具等版块，包含PolarDB、AnalyticDB、Lindorm、RDS、MongoDB、DMS等产品家族，为企业提供覆盖实时处理与存储、分析和发现、数据开发与治理的云原生一站式数据管理与服务。

阿里云瑶池数据库架构体系



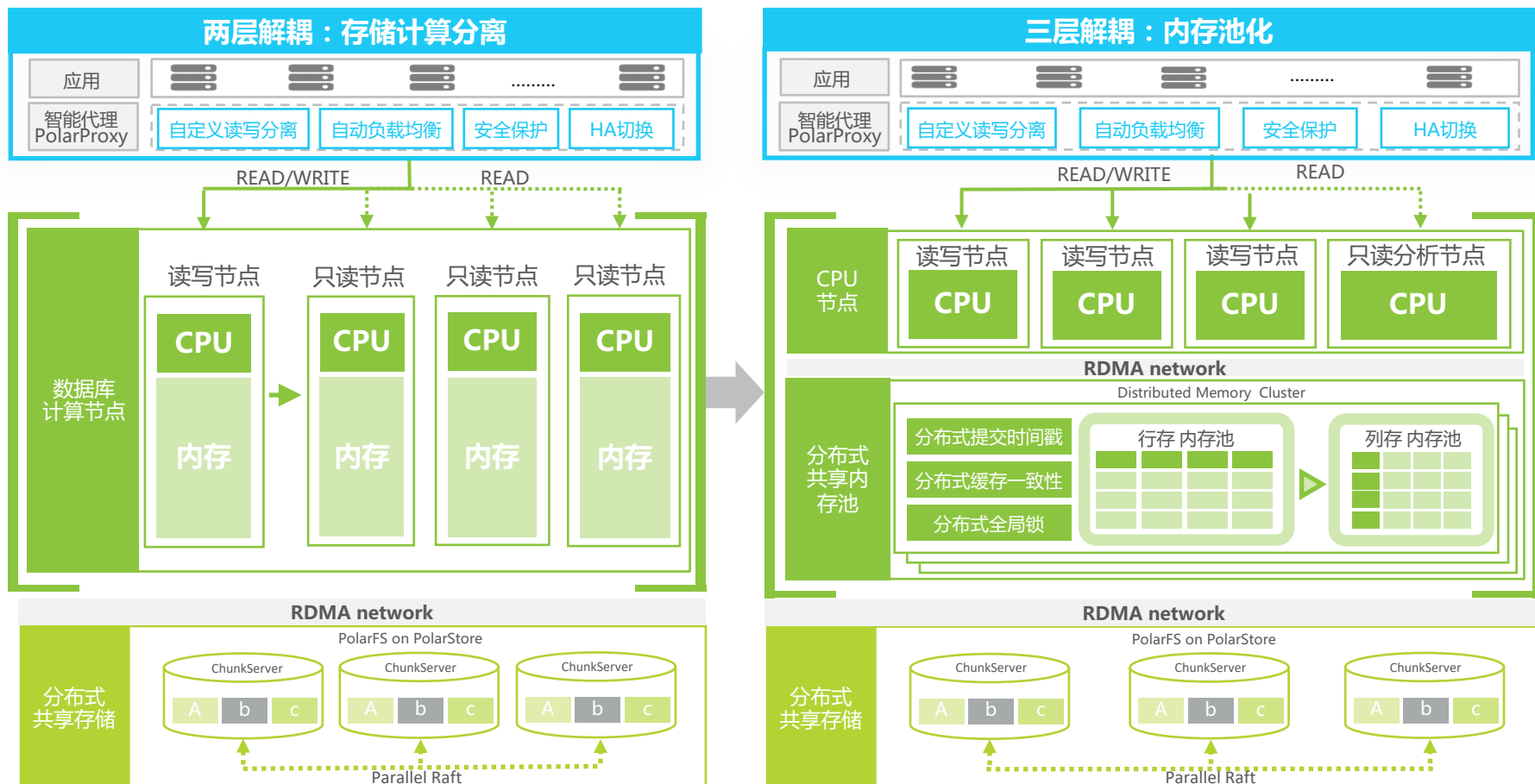
主要用户及技术领先性



来源：阿里云，艾瑞咨询研究院整理及绘制。

云原生数据库PolarDB：业内首次实现“三层解耦”架构

“三层解耦”架构演进

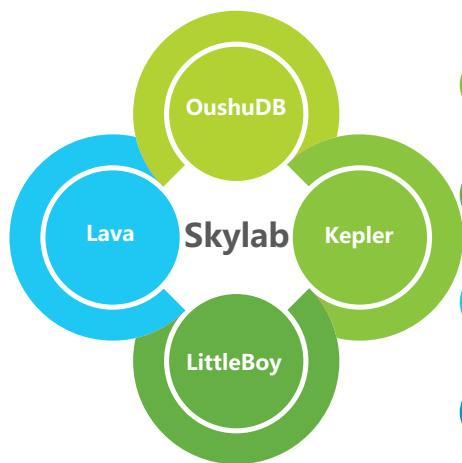


来源：阿里云，艾瑞咨询研究院整理及绘制。

打造 Skylab 产品体系，致力于云原生数据技术的突破创新

偶数科技专注于云数据平台产品和解决方案，由 Apache HAWQ 数据库项目创始人发起成立，有多位核心 Apache Committer。偶数科技搭建起云数据平台 Skylab 产品体系，包含：云数据库 OushuDB、自动化机器学习平台 LittleBoy、数据分析与应用平台 Kepler 以及数据管理平台 Lava。Skylab 实现了云原生实时湖仓一体六大创新能力，攻克了集群规模和并发受限、非结构化数据无法整合、建模路径冗长、数据一致性弱等难题，打破数据孤岛、优化数据供给能效、降低运维成本和技术门槛。偶数科技是国家级专精特新“小巨人”企业，入选福布斯中国企业科技 50 强，Gartner Cool Vendor 和 IDC Innovator。

Skylab 产品体系及 ANCHOR 特性



- 支持多类型数据：**
支持关系表、文本、图像、视频等结构化数据和非结构化数据存储。
- 云原生：**
适合云环境，自由增减计算和存储资源，按用量计费，节约成本。
- 数据一致性：**
通过完善的事务机制，保障不同用户同时查询和更新同一份数据时的一致性。
- 超高并发：**
支持数十万用户使用复杂分析查询并发访问同一份数据。
- 一份数据：**
所有用户（BI 用户、数据科学家等）可以共享同一份数据，避免数据孤岛。
- 实时 T+0：**
全量数据 T+0 流处理和实时按需查询，基于数据的事前预测、事中判断和事后分析。

平台及行业解决方案



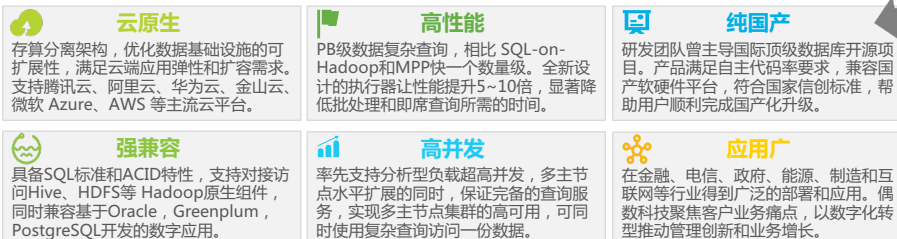
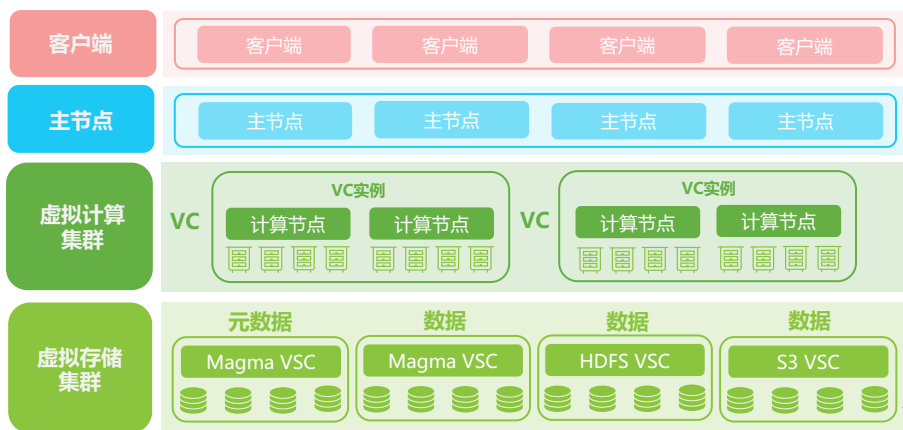
来源：偶数科技，艾瑞咨询研究院整理及绘制。

来源：偶数科技，艾瑞咨询研究院整理及绘制。

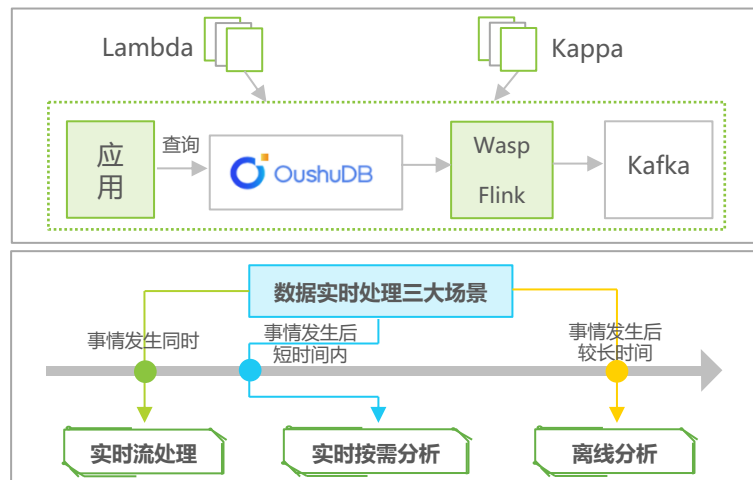
OushuDB 分布式云原生数据库，自主研发，信创可用

OushuDB 是由偶数科技自主研发打造的，新一代云原生分布式数据库，采用存算分离架构，按需弹性，支持公有云与私有云。OushuDB 遵循 ANSI-SQL 标准，支持完整事务特性，具备极速执行器，提供 PB 级数据交互式查询能力，以及对主要 BI 工具的描述性分析支持和高级机器学习功能。在实时数据处理方面，基于 OushuDB，偶数科技提出了满足实时流处理、实时按需分析和离线分析的 Omega 架构，率先实现了实时湖仓 ANCHOR 能力，在简化数据架构，提升数据处理效率，避免数据孤岛的同时，极大提升了企业实时数据分析能力。

OushuDB 的产品架构与主要特点



Omega 架构与实时数据处理场景



偶数科技主要企业用户



来源：偶数科技，艾瑞咨询研究院整理及绘制。

来源：偶数科技，艾瑞咨询研究院整理及绘制。

开源超融合能力打破数据边界，全面降低数智化转型门槛

矩阵起源是一家专注于为企业用户提供简捷强大的数据操作系统的数据基础软件公司。公司创始团队来自腾讯云、Snowflake等国内外一流的互联网企业、软件公司、数字化企业和开源社区，核心团队为产品、研发、解决方案、生态和开源社区等领域的专家，在分布式架构、数据库、云计算、大数据及人工智能等领域积累了丰富经验。作为一家数据智能领域的创新企业，矩阵起源虽成立时间不久，但凭借领先的技术创新和工程能力，自主研发超融合异构云原生数据库MatrixOne。产品提供超融合引擎、异构云原生、极致性能等核心能力，帮助用户充分释放数据潜力和创新力，在行业市场已拥有一定的品牌影响力，并获得“2022中国信科种子独角兽企业”、“大湾区高成长性企业TOP50”等行业荣誉。

MatrixOne 的产品架构



MatrixOne 的核心能力



极简开发运维

- **极简开发运维**：统一的开发协议、数据治理和安全管理，开箱即用服务多种数据应用，大幅降低开发、运维及运营的成本。



变ETL为ELT

- **变ETL为ELT**：实时数据更新即可见，摆脱冗余ETL和各种复杂依赖，解决数据不一致问题，有效提高数据应用开发效率。



极致分析性能

- **极致分析性能**：在资源隔离的基础上支持OLTP和OLAP负载的独立拓展，复杂查询性能提升3~5倍，增加业务敏捷度。



跨云高性价比

- **跨云高性价比**：避免厂商绑定，实现数据集跨云跨地域的统一管理与极致伸缩，优化混合基础设施部署性价比。

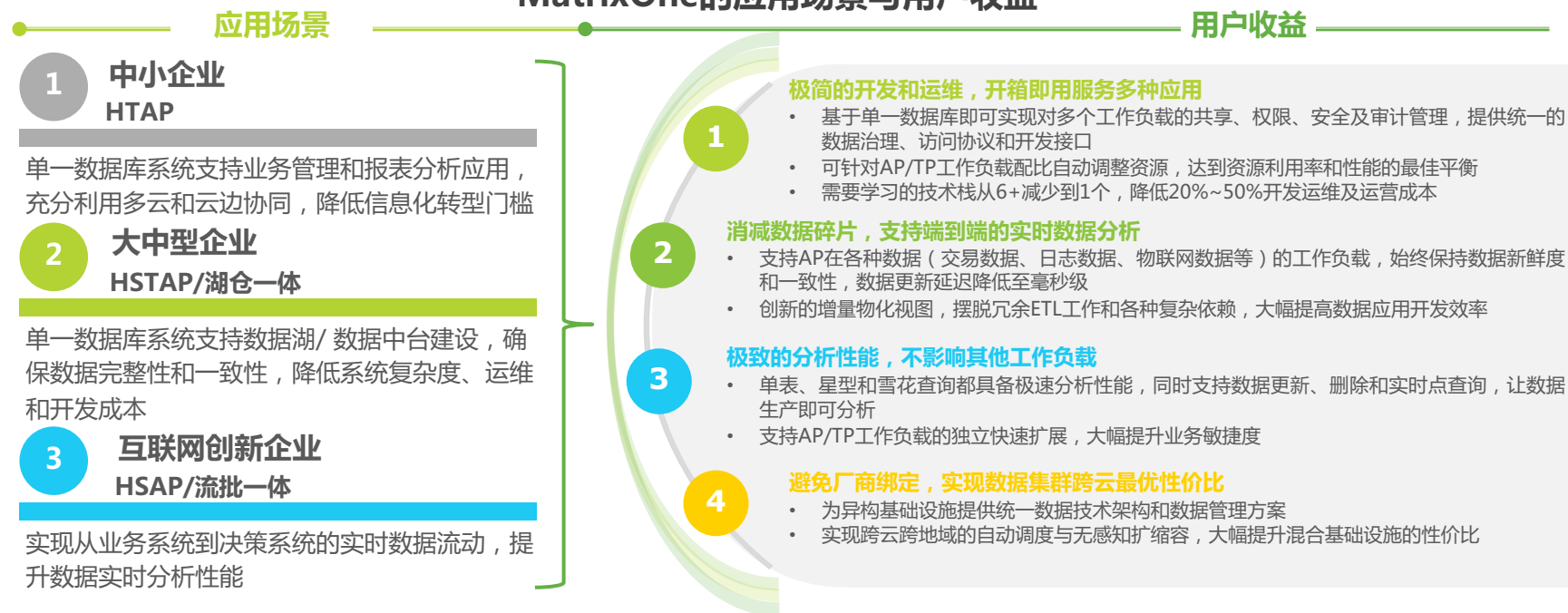
来源：矩阵起源，艾瑞咨询研究院整理及绘制。

来源：矩阵起源，艾瑞咨询研究院整理及绘制。

灵活支持事务、分析和流计算的超融合异构云原生数据库

MatrixOne借助自主研发的分布式存算框架，实现一套数据库系统同时灵活支持事务、分析和流计算的数据管理和应用，并支持公有云、私有云、数据中心和边缘节点的部署和使用。MatrixOne基于S3优化的存算分离架构，让用户的云原生/本地部署都有高可扩展、高可用、高弹性的性能体验。通过全新设计和开发的TAE存储引擎，实现一个引擎同时支持事务和分析功能，并保证数据分析的高性能和低延迟。此外，MatrixOne内置的流引擎允许直接、无缝地访问表，对数据源进行流化处理，避免冗余的数据复制、转换和移动，并通过基于因式计算查询引擎的增量物化视图技术，高效处理工作负载。基于矩阵起源的核心技术与创新能力，MatrixOne让数据应用开发更加敏捷，运维更加简化，同时保证了性能极致。

MatrixOne的应用场景与用户收益

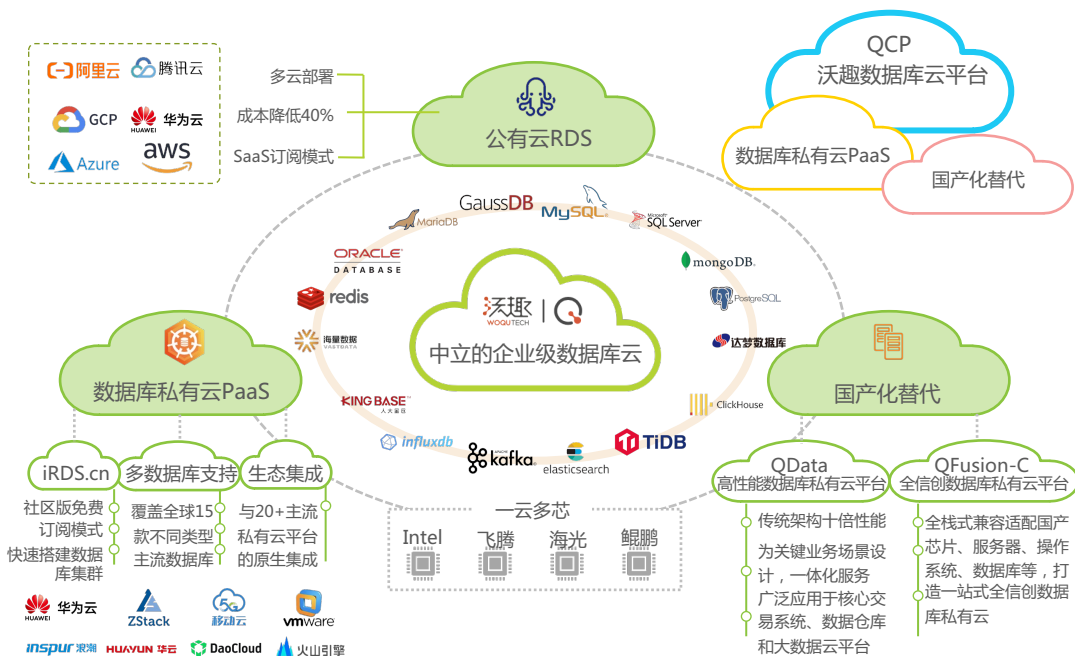


来源：矩阵起源，艾瑞咨询研究院整理及绘制。

中立的企业级数据库云，让数据库基础设施更简单

沃趣科技创建于2012年，创始团队为原阿里巴巴去“IOE”时期数据库团队骨干。多年来结合数据库多元化、国产化及云化趋势，为企业用户提供最佳的数据库基础设施创新。目前，沃趣科技的产品版图覆盖数据库国产化替代、私有云PaaS和公有云RDS三大体系，支持国内外主流数据库，提供企业级全生命周期的数据库管理能力。除基础平台外，沃趣还独立研发了DTS、DBS和DMS等数据库工具，降低用户数据迁移、备份和管理的成本。至2022年，沃趣已累计服务3000+企业客户与40+行业，是国家级高新技术企业、浙江省级企业研究院单位、浙江省信创联盟理事单位、杭州独角兽企业。

沃趣科技的核心产品体系



来源：沃趣科技，艾瑞咨询研究院整理及绘制。

沃趣科技的典型案例

证券行业

- 解决金融机构核心交易数据库集中化难题，为核心交易业务基础平台架构的开放和可控发展提供可靠保障，提升核心系统交易处理能力5~20倍，TCO降低70%，实现水平云化扩展。
- 基于国产数据库、操作系统、芯片提供全信创数据库一体化服务。

运营商

- 构建云原生数据库底层基座，为运营商提供安全、稳定、可靠的开放数据库云平台解决方案。
- 完成多类型数据库云部署，覆盖14个数据中心、12个省级中心，目前已稳定运行了6w+实例，4000+部署节点。

制造行业

- 满足技术无感化需求，对上端业务屏蔽底层数据库复杂技术实现，提升资源池化能力，覆盖数据库服务多个应用场景。
- 指数级提升业务效率，支撑制造业进行更多更广的信息化建设，助力其达到最佳营收增长。

教育与医疗

- 承载医院核心HIS系统，覆盖备份容灾等多业务数据库，医疗行业多数据库全生命周期管理，助力“云上”医院数字化进阶。
- 基于统一数据库服务能力支撑，助力智慧校园IT架构整合，保障核心数据系统稳定运行。

来源：沃趣科技，艾瑞咨询研究院整理及绘制。

数据库概述	1
数据库供给侧分析	2
数据库需求侧分析	3
典型企业案例	4
行业趋势洞察	5

国产数据库的全新增量逻辑，风险与机会并存

随着基础软件领域成为资本热土，仅2021年新成立数据库公司超30家，新兴厂商核心团队普遍来自大厂，技术扎实，经验丰富，产品各具特色，性能较佳，并且资金充裕。但另一方面，国内数据库市场的用户侧呈现如下特点：第一，互联网用户多采用开源或自研数据库，极少使用新兴的产品；第二，政企类用户具备完善的产品和服务生态，新兴厂商难切入；第三，缺乏互联网+新“风口”带动市场增量空间；第四，大量中小企业存续不稳定，对数据库需求不强烈。数据库厂商纷纷提出出海战略，创业公司自成立之初便定位国际化，但普遍面临信任度、监管、地理位置等障碍，而上云、开源、新技术敏感度成为除基础技术以外，国产数据库顺利出海的关键要素。

数据库出海是必然趋势

与生俱来的全球化基因

数据库的底层语言为全球通用的 SQL 语句，性能优异的产品理应在全球化市场中占有一席之地。

数据库上云降低出海门槛

打破本地部署模式下，数据库海外交付的物理阻碍，借助云厂商的原生式 IaaS 能力，快速搭建海外渠道。

数据库开源，内外兼修

一方面开源社区加快产品研发，提升产品质量；另一方面，开源有利于降低海外用户对于中国数据库的顾虑，友好性提升。

国产原生分布式技术具备先进性

为国产数据库提升可靠性、满足日趋复杂的用数场景提供了全新的“解题思路”，加快了赶超国际一流数据库的步伐。

来源：公开资料，艾瑞咨询研究院整理及绘制。

国产数据库的出海线路

➤ 欧美线路

• 数据库行业制高点

国际数据库巨头齐聚，是商业化难度最大、技术趋势最为领先、最具挑战性的市场，但也是最有说服力的市场。

• 欧美打标杆，日韩做变现

在世界顶级市场拿下标杆案例，是国产数据库出海的重要意义，之后进击日韩基础软件市场，已成为比较理想的发展路径。

➤ 东南亚线路

• 潜力巨大，蓝海市场

相对中国，现阶段东南亚的互联网水平整体较为滞后，处于互联网业务激增、传统业务数字化转型的初级阶段。

• 商业化路径更快

经国内数字化转型打磨多年的国产数据库产品与东南亚市场现阶段用数需求契合。

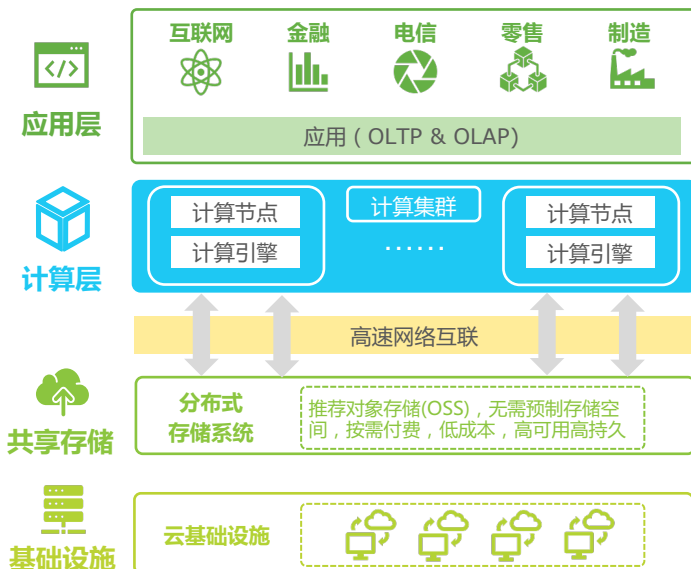
来源：公开资料，艾瑞咨询研究院整理及绘制。

趋势二：云原生

上层需求驱动，底层技术支撑，存算分离，优化资源编排

数据库云托管解放了底层硬件的管理成本和计算资源约束，但底层存算一体的资源绑定导致无效成本和资源浪费。云原生数据库实现了云资源对数据库的细粒度资源拆解，符合业务的资源编排。尽管各家厂商的架构设计和创新思路不尽相同，但可总结为“软硬兼施”。在软件层面，计算层打包 SQL 语句解析、物理计划执行、事务处理等，共享存储层存放事务日志和数据存储，并通过分布式技术保障高可用和一致性，最后二者采用高速网络互联，通过数据传输协议或其他技术提升 I/O 性能。此外，云原生数据库的分层架构还需结合新硬件技术的特性进行重构，如运用可扩展 CPU 和高主频内核进行算力优化，运用持久内存 (PMEM) 重构二级缓存以提升 I/O 密集型场景下的读写性能。未来数据库将进入从硬件平台到架构层再到应用层的全栈优化时代。

云原生数据库的整体架构和主要特点

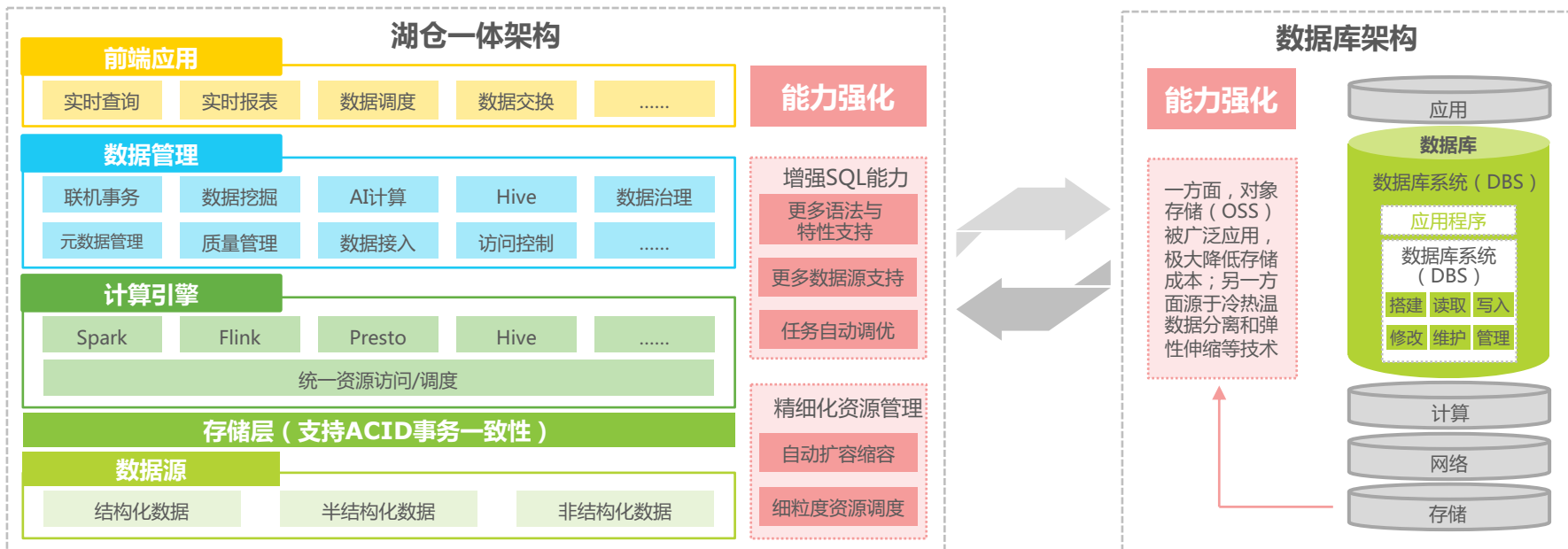


趋势三：湖仓一体

数据库与大数据技术边界不断模糊，呈现融合趋势

以 Hadoop 为代表的大数据技术以低成本的方式提供海量数据的开发和运行处理，但缺点是结构化能力较差，早期不支持 SQL 语句解析。随着大数据技术发展，基于 Hadoop 的数仓工具 Hive 开始提供 SQL 查询功能，但整体性能较差。随后支持 SQL 查询的工具（如 Spark SQL）开始涌现，大数据技术的分析处理能力提升，结构化趋势明显。另一方面，数据库天然具备强结构化能力，尤其关系型数据库原生支持 SQL 查询，但早期存储成本高，对海量数据的支持能力不足。但随着对象存储（OSS）被广泛应用，其低成本、高可用、数据持久及按需付费的特点，让数据库产品能够应对轻体量的大数据场景。由此，随着数据技术的发展创新，数据库与大数据技术的边界不断模糊，二者互相延伸。我们认为传统大数据技术有着更为丰富的生态，但是对技术要求较高，相比较而言，从仓向湖的延伸路线，更适合传统企业和中小企业。

湖仓一体下数据库与大数据技术的相互延伸



来源：公开资料，艾瑞咨询研究院整理及绘制。

艾瑞新经济产业研究解决方案



行业咨询

- 市 场 进 入 为企业提供市场进入机会扫描，可行性分析及路径规划
- 竞 争 策 略 为企业提供竞争策略制定，帮助企业构建长期竞争壁垒



投资研究

- IPO行业顾问 为企业提供上市招股书编撰及相关工作流程中的行业顾问服务
- 募 投 为企业提供融资、上市中的募投报告撰写及咨询服务
- 商业尽职调查 为投资机构提供拟投标的所在行业的基本面研究、标的项目的机会收益风险等方面的深度调查
- 投后战略咨询 为投资机构提供投后项目的跟踪评估，包括盈利能力、风险情况、行业竞对表现、未来战略等方向。协助投资机构为投后项目公司的长期经营增长提供咨询服务

关于艾瑞




艾瑞咨询是中国新经济与产业数字化洞察研究咨询服务领域的领导品牌，为客户提供专业的行业分析、数据洞察、市场研究、战略咨询及数字化解决方案，助力客户提升认知水平、盈利能力和综合竞争力。

自2002年成立至今，累计发布超过3000份行业研究报告，在互联网、新经济领域的研究覆盖能力处于行业领先水平。

如今，艾瑞咨询一直致力于通过科技与数据手段，并结合外部数据、客户反馈数据、内部运营数据等全域数据的收集与分析，提升客户的商业决策效率。并通过系统的数字产业、产业数据化研究及全面的供应商选择，帮助客户制定数字化战略以及落地数字化解决方案，提升客户运营效率。

未来，艾瑞咨询将持续深耕商业决策服务领域，致力于成为解决商业决策问题的顶级服务机构。

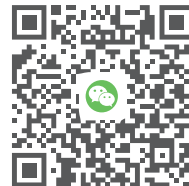
联系我们 Contact Us

 400 - 026 - 2099

 ask@iresearch.com.cn



企 业 微 信



微 信 公 众 号

法律声明

版权声明

本报告为艾瑞咨询制作，其版权归属艾瑞咨询，没有经过艾瑞咨询的书面许可，任何组织和个人不得以任何形式复制、传播或输出中华人民共和国境外。任何未经授权使用本报告的相关商业行为都将违反《中华人民共和国著作权法》和其他法律法规以及有关国际公约的规定。

免责条款

本报告中行业数据及相关市场预测主要为公司研究员采用桌面研究、行业访谈、市场调查及其他研究方法，部分文字和数据采集于公开信息，并且结合艾瑞监测产品数据，通过艾瑞统计预测模型估算获得；企业数据主要为访谈获得，艾瑞咨询对该等信息的准确性、完整性或可靠性作尽最大努力的追求，但不作任何保证。在任何情况下，本报告中的信息或所表述的观点均不构成任何建议。

本报告中发布的调研数据采用样本调研方法，其数据结果受到样本的影响。由于调研方法及样本的限制，调查资料收集范围的限制，该数据仅代表调研时间和人群的基本状况，仅服务于当前的调研目的，为市场和客户提供基本参考。受研究方法和数据获取资源的限制，本报告只提供给用户作为市场参考资料，本公司对该报告的数据和观点不承担法律责任。

合作说明

该报告案例章节包含部分企业的商业展示，旨在体现行业发展状况，供各界参考。

为商业决策赋能

EMPOWER BUSINESS DECISIONS

