

中国数据库的时与势



阿里巴巴集团副总裁 李飞飞

数据库、操作系统和中间件并列为三大基础软件，无论是在银行存取款，还是进行健康码查询，我们的日常应用和企业业务背后都离不开数据库。可以说，没有数据库，就难以构建数字化底座。

在国际权威机构Gartner公布的2021年全球云数据库魔力象限评估结果中，作为中国科技公司代表，阿里云蝉联了“领导者”（LEADERS）象限，意味着阿里云数据库综合实力已稳居全球第一阵营。

这既是对阿里云数据库实力的肯定，同时也证明了阿里云围绕云原生数据库新赛道进行前瞻布局 and 自研创新的战略是正确的。

今年入选Gartner全球云数据库魔力象限的企业共有20家，中国2家，占比10%。

这与中国巨大的市场空间和产业应用前景并不相称。在我看来，未来一定会有更多中国厂商出现在这份榜单上，并且不断向着“领导者”象限迈进。

之所以有这样的判断，是因为云计算的出现，让全球数据库市场格局迎来了40年以来的最大拐点。而这一次，时与势站在中国厂商一边。

数据库的40年大变局

过去的40多年，数据库技术一直在创新与迭代，经历了不亚于移动通信技术从1G到5G的巨大跨越。

上世纪80年代初，数据库系统逐步走进信息技术舞台的中央。2000年前后，大数据技术兴起；再到2010年后，云计算热度持续升温，云原生、分布式等技术不断发展。展望未来，数据库技术会因为云计算实现技术上的极大跃迁。

随着企业业务全面向数字化、在线化、智能化演进，企业面临指数级递增的海量存储产生的需求和挑战，包括并不限于诸如业务热点和突发流量带来的挑战。企业不仅需要降本增效，还需要进行数据分析、数据洞察，从而产生可指导行动的智能决策，传统的商业数据库已经难以满足和响应快速增长的业务诉求。

第一，系统面临持续扩容的挑战。传统数据库是在冯·诺依曼架构下发展而来，其核心特征是计算、存储等资源的紧耦合。当业务需要的资源超过底层系统所能提供的容量后，需要对现有系统进行灵活弹性的扩容。

第二，系统始终会存在宕机的风险。如果支撑核心在线业务的数据库出了问题，将导致业务系统无法正常运行。数据库需要高可用，确保无论底层资源发生任何问题，数据库都可以实现不同服务器、不同虚拟机、甚至是不同可用区和地域之间的切换，保证上层业务系统对底层硬件资源的错误或者风险的无感知容错。

面对这两个核心挑战，云原生数据库利用虚拟化技术将计算、存储等资源进行池化，将存储计算分离。打个比方，以前

创建独立的数据库，好比每家每户各打一口水井。如今的云原生数据库，则是把这些水井底部的水资源连接起来，表面上还是一口口独立的水井，但是水井的底部已经连在一起，形成了一个隐形的池塘，能实现随时按需用水。

在云原生数据库系统中，我们还可以通过AI和机器学习技术实现智能自动化的运维过程，快速调动资源，匹配业务需求，需要多少资源就扩展多少资源。就像自来水系统，用户可对水量进行实时灵活的控制。

概而言之，从传统数据库到云原生数据库，最大的技术变革是资源池化与资源解耦，以及由此而带来的弹性、高可用、容器化部署、智能化运维等核心能力。

双11是全球最大的超级数字工程之一，是阿里巴巴技术的练兵场。2020年双11零点的订单峰值58万笔/秒，这种瞬间爆发的场景，如果不是利用云原生的技术，仅依赖传统技术根本无法满足这种高并发、弹性、高可用的要求。

2021年是数字化转型的普及年，千行百业的业务加速上云进程。即使是单一客户，也拥有多样化的复杂业务以及丰富的业务负载特性，单一产品和架构已很难满足客户需要。阿里云数据库全面拥抱云原生，首次从客户场景视角，提出了“一站式全链路数据管理与服务”的理念，希望通过触手可及、简单易用、安全可靠的云数据库，让数据无缝的自由流动，满足企业多样化的业务诉求。

阿里云所打造的云原生数据库2.0——一站式全链路数据管理与服务，将云计算领域非常重要的云原生技术和分布式技术合二为一：上层是分布式，下层以云原生的方式实现，每个分区都可以享受云原生带来弹性、高可用的能力，同时上面有分布式带来的水平拓展的能力，解决高并发可能带来的瓶颈问题。在这个基础上，打造面向不同应用场景和需求的专属数据库和数据仓库系统，做到专库专用；同时，利用HTAP、在离线一体化、仓库一体、多模数据处理等核心技术，实现数据处理与分析的融合，以及数据在不同应用场景和系统之间的自由流

转。最后，通过在线数据管理平台实现一站式全链路数据管理与服务。

展望未来，云原生与分布式一体化、在离线一体化、HTAP混合负载查询与处理、物联网及多模数据融合处理与分析、安全可信与隐私保护、智能化运维与调优、机器学习和AI负载以及智能化算子与应用支持、新型硬件的适配和优化将成为云原生数据库重点发力和突破的技术方向。

中国数据库厂商的时与势

数据库作为应用型技术，先发优势和生态建设非常重要，如果没有技术上的突破性创新，后来者想要超越，可能性很小。在云时代的滚滚洪流之下，云计算已成为数据库发展的新赛道，中国数据库企业迎来了绝佳的变革机遇期。

据Gartner预测，到2022年，云数据库管理系统（DBMS）的收入将占DBMS市场总收入的50%，比2020年预测的提前了一年。数据库继续向云迈进，全球数据库产业结构正在加速重构。

云计算作为一种全新的科技服务，对数据库的研发、使用等方面带来了彻底的变革，也让全球厂商在数据库赛道上站在了同一起跑线上。中国有着众多的数字用户和独特的业务场景。没有成功应对双11流量洪峰、“健康码”等场景的数据库厂商，很难知道如何解决海量数据、超高并发交易洪峰等实际业务问题。放眼全球，中国厂商独享这样的机会。这一次，中国厂商乘“云”而上，拥有了比肩国外数据库的技术和创新能力。

以阿里云自研云原生数据库PolarDB为例，2017年才启动自主研发，但到今天很多功能，如内存、计算与存储三层解耦架构实现秒级弹性、多主多写、基于内存池化的列存索引支持HTAP等已经是全球首创或业内领先的技术，创新步伐已经领先国外同行。

数字经济的蓬勃发展，必然推动中国数据库市场的快速增长。中国信息通信研究院今年6月发布的《数据库发展研究报告（2021年）》指出，我国数据库产业进

入重大发展机遇期，预计到2025年，中国数据库市场总规模将达到688亿元，市场年复合增长率（CAGR）为23.4%，全球占比达到12.3%左右。

中国的数据库企业能够挺进并蝉联Gartner魔力象限的领导者地位，是几代人不懈努力的成果。数据库行业取得今天的成绩，与国家对基础软件产品的扶持，政策的引导和加持是分不开的。如果说，过去中国数据库技术打不开市场，有很多客观原因，但是今天中国乃至海外的企业，已经开始大规模接受中国数据库技术，并且金融、政务等核心应用开始尝试用云数据库大规模替代传统数据库。随着科技自立自强重大战略的实施，这种全面替代会是一个确定性的趋势，越来越多的企业将使用中国数据库技术。

今天，中国提供数据库产品的厂商已超过80家，其中一些企业已经成为新锐的独角兽厂商。虽然国内数据库企业单一厂商的人员规模还不足以与全球巨头企业抗衡，但是产业人才已经具备了一定的规模并形成了梯队。这为中国数据库产业的稳步发展打下了坚实的人才基础。

此外，通过产学研的携手创新，在数据库领域VLDB、SIGMOD和ICDE三大顶级会议的论文中，中国厂商的论文数量也呈现了逐年上升趋势且部分论文提出的数据库技术创新甚至获得全球同行的认可，这也是中国数据库企业乘势而起的标志之一。

“因天之时，就地之势，依人之利，则所向者无敌，所击者万全矣。”在产业大变革时代，深刻认识和正确把握我们所面临的时与势，有利于中国数据库厂商看清方向，找到快速成长的路径。我相信，时与势也必定转化为中国数据库产业快速发展的动力，让我们在这样关键的信息技术领域拥有应有的一席之地。

本文作者系阿里巴巴集团副总裁、阿里云数据库产品事业部负责人、达摩院首席数据库科学家，IEEE Fellow、ACM杰出科学家。

接种新冠病毒疫苗

利己利家利国

新冠病毒疫苗

2019-2020

中宣部宣教局 国家卫生健康委员会宣传司 中国疾病预防控制中心 中国健康教育中心

公益广告