

云计算蓝皮书

(2024 年)

中国信息通信研究院

2024年7月

版权声明

本蓝皮书版权属于中国信息通信研究院，并受法律保护。转载、摘编或利用其它方式使用本蓝皮书文字或者观点的，应注明“来源：中国信息通信研究院”。违反上述声明者，本院将追究其相关法律责任。

更名声明

原“集智”白皮书更名为“集智”蓝皮书。“集智”蓝皮书将继续秉承原有的编撰理念和高质量标准，致力于提供有价值的信息和洞见。



前 言

习近平总书记在中共中央政治局第十一次集体学习时强调，加快发展新质生产力，扎实推进高质量发展。2024年政府工作报告提出，适度超前建设数字基础设施，加快形成全国一体化算力体系。云计算作为数字时代的新型基础设施，是整合科技创新资源，加快形成新质生产力的关键要素和构建现代化产业体系的底座支撑。深化推进云计算和实体经济融合升级，是技术革命性突破、生产要素创新性配置、产业深度转型升级的重要路径，是助力新型工业化发展、加快中国式现代化进程的催化剂。

过去一年，人工智能产业呈现井喷式发展，充分释放出以云计算为中枢构建的数字化智能化实体的竞争力。云计算作为大模型的底层算力支撑，进入到“人工智能+”的新发展阶段，进一步影响人类生产生活方式和全球产业格局，并呈现以下新发展特点：

一是各国持续提升云计算战略价值，安全发展成为欧美关注重点。美国持续加强云计算行业应用，同时出台政策、规则关注云安全和云主权，避免技术外溢，以巩固其在云计算领域的领先地位。欧盟发布宣言、计划，将云计算服务的使用率定为提升国家竞争力的关键绩效指标，确认云计算作为国家科技创新和数字经济时代的重要支撑，特别强调云数据安全。东南亚和中东地区等国发布数字路线图、法案，注重境外资源引入。日韩侧重扩大云计算投资，推动本土云服务扩张和智能化发展。我国持续加强云计算基础设施建

设和行业应用，助力产业高质量发展。

二是大模型推动云计算产业开启新一轮增长，我国市场保持较高活力。2023年，全球云计算市场规模为5864亿美元，同比增长19.4%，在生成式AI、大模型的算力与应用需求刺激下，云计算市场将保持长期稳定增长，预计2027年全球云计算市场将突破万亿美元。2023年，我国云计算市场规模达6165亿元，同比增长35.5%，仍保持较高活力，AI原生带来的云技术革新和企业战略调整，正带动我国云计算开启新一轮增长，预计2027年我国云计算市场将突破2.1万亿元。

三是云计算技术“人工智能+”融合趋势明显，赋能多产业加快形成新质生产力。在“智转数改”的新需求下，企业上云用云需求不断深化，对应用现代化能力、稳定性保障能力、云原生安全能力、云成本优化能力、垂直类应用能力以及云算融合能力等方面要求不断提升，带动相关技术创新发展，特别是云计算与智算的加速融合，推动人工智能技术发展和应用快速革新。“云+AI”作为新一代人工智能发展的驱动力量，正在重构云服务产业格局，加速形成新质生产力。

在此背景下，中国信息通信研究院继《云计算白皮书(2012年)》后第10次发布云计算蓝皮书，本蓝皮书立足产业新发展、新变化、新需求，聚焦新形势下全球云计算发展重点，梳理我国云计算整体发展特点，首次提出中国云计算发展评价体系，并对云计算助力人工智能发展进行了论述，最后给出发展建议和展望。

目 录

一、全球云计算发展概述.....	1
(一) 各国持续深化云计算战略, 安全和智能化成发展重点.....	1
(二) 云计算市场持续稳定增长, 行业巨头保持领先地位.....	3
(三) 云计算技术融合价值凸显, 成为新技术创新发展的土壤.....	6
二、我国云计算发展概述.....	14
(一) 中央与地方协同推进云计算, 助力产业高质量发展.....	14
(二) 云计算市场迎来新一轮增长, AI 加速产业格局重新洗牌.....	16
(三) 各行业上云积极性持续攀升, 梯队分层效应依旧明显.....	19
三、中国云计算发展评价.....	24
(一) 云计算发展评价体系.....	24
(二) 云资源供需发展水平评价.....	25
(三) 云服务应用发展水平评价.....	27
(四) 云技术融合发展水平评价.....	31
四、云计算与智算加速融合, 开启智能化新纪元.....	34
(一) AIaaS 夯实云计算技术底座, 持续提升智能软硬件效能.....	35
(二) AIPaaS 聚焦智算服务开发能力, 成为 AI 工程化实践最优解.....	38
(三) MaaS 定义云上服务新范式, 赋能 AI 创新与效率提升.....	41
(四) AISaaS 推动 AI 应用落地实践, 加速企业数字化转型进程.....	43
五、建议与展望.....	46
(一) 发展建议.....	46
(二) 发展展望.....	47
附件一 数据来源.....	50
附件二 计算方法.....	52
附件三 计算口径.....	53

图 目 录

图 1 全球云计算市场规模及增速（亿美元）.....	4
图 2 2023 年全球各区域云计算市场规模占比.....	5
图 3 2023 年全球主要厂商云计算业务营收（亿美元）.....	6
图 4 中国云计算市场规模及增速（亿元）.....	17
图 5 中国云计算细分领域市场规模及增速（亿元）.....	18
图 6 2023 年中国公有云 IaaS 厂商市场占比.....	19
图 7 云计算发展指数体系.....	25
图 8 2023 年 31 个省（自治区、直辖市）供云量指数.....	26
图 9 2023 年 31 个省（自治区、直辖市）用云量指数.....	27
图 10 2019-2023 年典型行业用云热力图.....	28
图 11 2023 年行业用云量占比.....	29
图 12 2023 年不同规模企业上云率.....	30
图 13 2023 年 31 个省（自治区、直辖市）IaaS 服务体验指数.....	31
图 14 2023 年 31 个省（自治区、直辖市）算云融合指数.....	33
图 15 数据服务云化部署比例.....	34
图 16 智能计算服务云化部署比例.....	34
图 17 智能云体系架构图.....	35

表 目 录

表 1 指标体系与计算口径.....	53
--------------------	----

一、全球云计算发展概述

（一）各国持续深化云计算战略，安全和智能化成发展重点

美国持续加强云计算行业应用，同时出台政策、规则关注云安全发展，避免技术外溢，以巩固其在云计算领域的领先地位。一方面，美国不断加强云安全防范。2023 年 3 月，美国白宫发布《国家网络安全战略》，提出要提高包括云基础设施在内的关键设施的网络安全和弹性，利用云服务实现更好、更经济的大规模网络安全实践，加速向云服务迁移，提升整个联邦政府的网络安全态势。2024 年 2 月，美国云安全工作组发布安全策略文件，提出优化政府云安全性的措施，建议国会通过立法进一步规范云服务安全，提倡白宫管理和预算办公室将其云智能指南更新为“云安全”指南，并增强网络指标。另一方面，美国持续关注云服务的安全管控。2024 年 1 月，美国商务部产业安全局发布关于《“采取额外措施应对重大恶意网络活动方面的国家紧急情况”拟议规则的通知》，以加强美国云计算服务商对外国客户的身份识别和报告义务，拟议规定提供美国“基础设施即服务”的云服务商、经销商须核验外国人客户身份信息，并在特定情形下向美国商务部报告外国人客户的详细身份信息和人工智能大模型训练活动情况。

欧盟发布宣言、计划，将云计算服务的使用率定为提升国家竞争力的关键绩效指标，确认云计算作为国家科技创新和数字经济时

代的重要支撑。2023 年 11 月，欧洲贸易协会组织发布《欧洲 2030: 数字强国—提升欧洲科技领导力和复原力的 20 种解决方案》宣言文件，提出一项关键绩效指标，以衡量建设有竞争力的欧洲方面取得的进展，指标包括：75%的欧盟公司应使用云计算服务、大数据和人工智能。2023 年 12 月，欧盟委员会批准一项 12 亿欧元的国家援助计划“欧洲共同利益重要计划——下一代云基础设施和服务”，以支持欧洲先进云计算和边缘计算技术的研究、开发和首次工业部署。此外，欧盟还在持续关注云主权。2023 年 5 月，欧洲网络安全认证小组对新的《欧盟云服务网络安全认证计划草案》展开审查，要求云厂商提供的云服务必须在欧盟运营和维护，所有云服务的客户数据都必须在欧盟存储和处理，欧盟法律优先于有关云服务提供商的非欧盟法律，欧盟以外的任何实体不得直接或间接、单独或共同对申请云服务认证的云服务提供商拥有有效控制。

东南亚和中东地区国家发布数字路线图、法案，将投资和发展云计算作为国家长期战略。2024 年 2 月，泰国数字经济与社会部发布 2024 年数字路线图，致力于通过专注云计算、人工智能、数字劳动力以及网络安全等领域来提高在全球数字经济中的竞争力，其中云优先政策提倡广泛采用云系统进行数字基础设施改造，提高政府服务质量。2023 年 11 月，越南国会正式批准《新电信法》，放宽了对“数据中心服务”“云计算服务”和“基于互联网的媒体服务”的监管，对国外企业在越南直接开展数据中心和云计算服务几乎没有

施加限制，表明了其积极拥抱云计算的态度。2023 年 5 月，新加坡智慧国家和数字政府办公室与谷歌云联合宣布推出人工智能政府云集群，通过云平台加速新加坡公共部门对人工智能技术的采用。2023 年 4 月，沙特宣布在新设立的四个经济特区中，单独设立了“云计算经济特区”，专项鼓励云计算服务企业入驻，力求打造沙特科技创新新高地。

韩国和日本不断加大云计算投资力度，持续推动本土云服务扩张和智能化发展。2023 年 9 月，韩国提出“制定第五个科学技术基本规划的方向（2023-2027 年）”，并发布“第三个云计算基本计划”，鼓励公共部门优先部署私有云，促进公共部门的数字化创新，加快软件行业的云部署进程，加强云产业竞争力，打造云产业永续发展的生态圈。2024 年 2 月，韩国科技信息通信部宣布，将投资 9150 万美元来推动本国云计算产业智能化发展。2024 年 1 月，日本政府接受亚马逊云在其本土投资的 150 亿美元，旨在推动日本企业的生成式人工智能可用性和亚马逊生成式人工智能的客户采用率，促进日本云计算产业的智能化发展。

（二）云计算市场持续稳定增长，行业巨头保持领先地位

从整体来看，全球云计算市场进入稳定增长阶段。Gartner 数据显示¹，2023 年以 IaaS、PaaS、SaaS 为代表的全球云计算市场规模达到 5864 亿美元，增速 19.4%。参考过去几年的市场规模数据，可

¹ 《Forecast: Public Cloud Services, Worldwide, 2021-2027, 2Q23 Update》Gartner 2024 年 4 月

见由疫情带来的波动已回落到正常水平，云计算仍然是市场高稳定增长的不可或缺的技术。未来几年，随着云计算与生成式 AI、大模型、算力的深度融合，市场将以 18.6% 的年复合率增长，预计到 2027 年全球云计算市场将突破万亿美元。

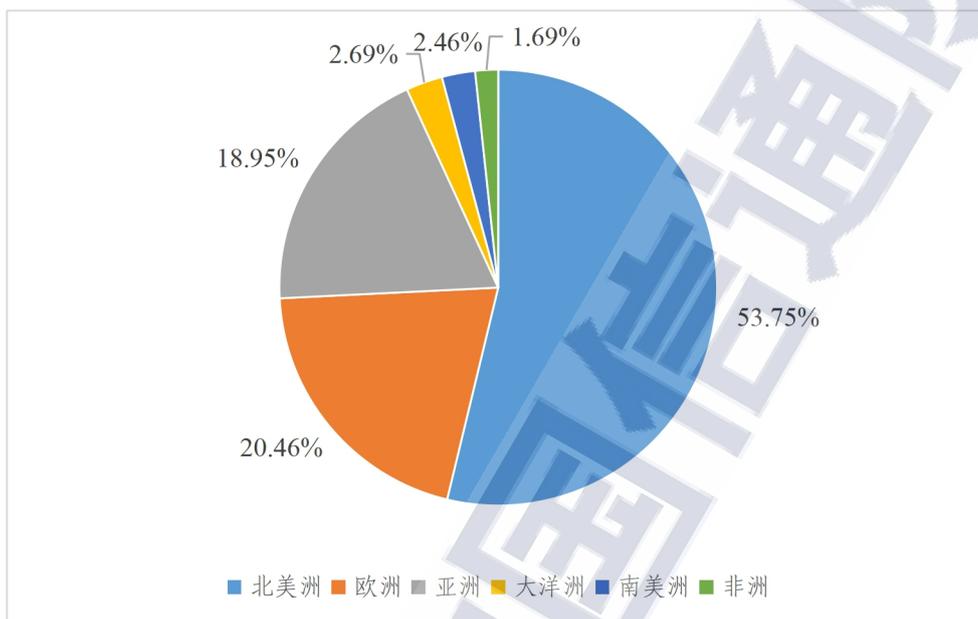


来源：Gartner，2024 年 4 月

图 1 全球云计算市场规模及增速（亿美元）

从区域层面来看，亚洲市场潜力巨大，中东地区发展迎来热潮。受人工智能技术发展带动底层云计算需求增长，2023 年北美洲云计算市场规模占比 53.75%，保持领先地位。得益于部分国家的利好政策等因素，亚洲云计算市场实现快速发展，占比与欧洲的差距缩小到两个百分点以内。大洋洲和南美洲与去年持平，整体占比变化不大，非洲则实现了 3 倍增长至 1.69%。此外，中东地区的发展尤其突出，有望迎来发展热潮，阿联酋计划投资 400 亿美元，加速云计

算在内的先进技术研究，亚马逊云计算服务等头部企业也加码沙特、埃及等中东地区新型基础设施建设，这些都将持续推动亚洲、非洲的云计算规模扩大。

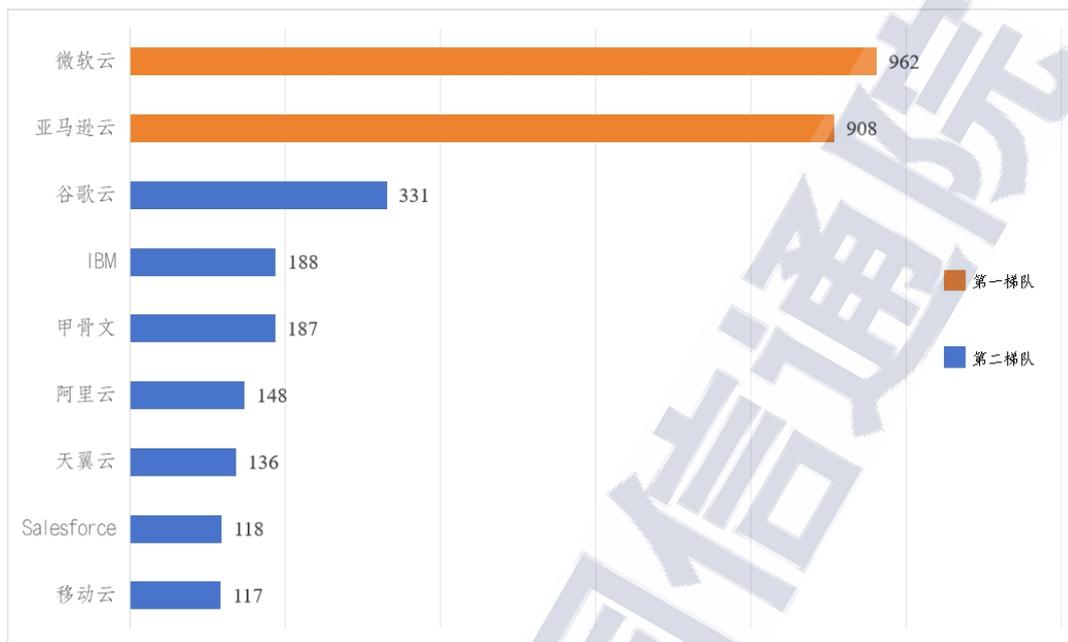


来源：Gartner，2024 年 4 月

图 2 2023 年全球各区域云计算市场规模占比

从服务商层面来看，云计算巨头持续领跑全球，第二梯队深耕细分技术领域保持追赶。2023 年，微软云和亚马逊云凭借 AI 云产品改造和市场布局优势，分别以 962.13 亿美元和 907.57 亿美元的营收稳居全球市场前两名。第二梯队服务商保持积极追赶态势，但受市场布局先手影响，营收增长有限，转向业务整合，通过深化主营业务技术优势，拉高领域技术壁垒，推动其业务增长。例如，谷歌云基于开源技术的领导地位大力发展 AI 工具、IBM 关注混合云与 AI 解决方案咨询、天翼云和移动云借助运营商资源优势聚焦数字化

转型方案。



来源：公开资料整理

图 3 2023 年全球主要厂商云计算业务营收（亿美元）

（三）云计算技术融合价值凸显，成为新技术创新发展的土壤

经过多年发展，云计算技术已日臻成熟，成为承载各类新技术创新发展的关键基础设施，并与多个领域融合衍生出新的技术方向。

“云+应用”方面，应用现代化促进技术革新和产业需求相互融合，推动企业数字化升级转型。云原生作为新型基础设施，其资源平台建设已日趋成熟，随着新技术栈与平台架构的不断升级，上层应用生态面临革新，应用现代化应运而生，能够从供给方式、生产模式、生产工具等方面助力企业提升生产效率和质量。当前，应用现代化已进入规模实践期，并呈现以下特点：**一是技术架构方面，**

分布式云原生有力支撑多云应用的一体化管理诉求。出于对数据产权、安全合规与业务稳定等因素的考量，多云策略成为应用部署的重要选择。分布式云原生方案具备统一的资源纳管、运维管理、安全保障、算力调度与成本控制能力，为多云环境下应用的全生命周期管理提供底层架构支持。二是应用架构方面，事件驱动为应用提供高效灵活的架构模式。传统应用架构的“请求-响应”模式已难以应对大规模、强实时、高可用的现代化业务需求，事件驱动架构则凭借其松耦合、分布式、异步执行、可扩展性等特点，能够有效降低应用的开发与维护难度，增强应用的容错能力。三是数据架构方面，云原生融合大数据技术加速释放应用的数据价值。数字时代各行业的迅速发展引发数据爆炸式增长，大数据云原生架构凭借弹性伸缩、统筹资源调度以及智能化运维等方面的优势，能够实现低成本且快速的数据汇聚、处理与分析。例如，亚马逊云 EMR、谷歌云 Dataproc 和阿里云 MaxCompute 等云原生大数据产品能够快速部署、管理和扩展 Hadoop、Spark 集群，满足各类数据处理需求。

“云+管理”方面，云成本治理与云原生、大数据、SaaS 等技术结合，引领企业迈向精细化运营管理阶段。Flexera《2024 年云现状调查报告》显示，84% 以上的企业面临云成本管理挑战，管理云支出已经连续两年成为企业用云面临的最大的挑战。FinOps 可提供有效应对复杂云环境的成本管理框架，同时提升业务敏捷性与成本效率的动态平衡。当前，云成本治理基于 FinOps 理念由资源监控过渡

至精细化治理阶段，实现方式不断升级。一是云成本治理基于对云原生应用组件的资源监控，实现成本的精细化、动态化管理。云原生技术以微服务架构、容器化、持续交付等核心要素重新定义应用程序的构建、部署和管理方式，云原生应用的资源颗粒度更为精细，云成本治理能够对每一个应用组件的资源消耗进行精确计量，实现资源利用的最优化。例如，谷歌云依托 Kubernetes Service 提供高度灵活的资源管理、自动扩展和成本分析工具，使企业在云环境中实现资源的快速部署、智能监控和成本优化；阿里云借助容器服务和专为云原生设计的服务网格构建容器编排和微服务管理能力，实现云原生平台资源成本的动态优化。二是云成本治理利用大数据和 AI 算法建模，助力企业精准预测成本趋势。大数据分析结合 AI 算法能够快速识别出成本使用的异常模式，帮助云成本治理实现从被动响应到主动预测和策略驱动的转变，提升企业对成本控制的能力和效率。例如，亚马逊云利用 Cost Explorer 工具深度分析历史消费数据，为用户生成高精度的费用预测模型与优化建议；腾讯云的成本优化平台利用机器学习模型预测成本趋势，帮助企业精准预估未来开支并及时定位资源浪费。三是云成本治理利用 SaaS 化工具重塑服务形态，简化企业成本控制复杂度。随着 SaaS 化云成本治理解决方案与成本优化工具的兴起，企业能够即时享受自动化预算监控、资源调度优化和费用预测分析等成本优化服务，显著降低企业部署难度，提升运营效率和成本效益。例如，谷歌云采用 Cloud Billing、Cloud

Cost Management 等工具提供集中化成本监控、预算分析和优化建议，促进企业成本治理与云效率的双重提升；华为云的一站式成本管理平台集成了成本收集、使用分析、监控预警等功能，企业能够轻松实现云端资源消耗的全景概览，从而大幅简化云成本处理流程。

“云+行业”方面，行业云平台基于行业技术方案构建，具有共享化、场景化和市场化特征，成为大型企业建云新模式。Gartner 在 2023 年、2024 年均将行业云平台列为十大战略技术趋势之一，并预测到 2027 年超过 70% 的企业将采用行业云平台加速业务创新。行业云平台融合垂直领域解决方案，是专为行业内企业提供业务创新和数字转型的云平台。架构方面，行业云平台通过组装式应用实现场景级能力输出，推出共享可复用的生态产品。建设方通过建设生态场景级的可复用基础框架，建立统一管理的共享服务标准和数据标准，形成产品场景的标准化和业务共享服务能力，从而促进生态产品的高质量开发和能力复用。如工商银行金融生态云构建开放的金融生态平台，通过应用组装实现优势产品的标准化封装和输出，推出丰富的金融服务解决方案。应用方面，行业云平台具备生态 SaaS 应用场景持续扩展能力，提高企业云上业务创新的敏捷性和灵活性。行业云平台建设单位聚焦业务逻辑解决方案，通过自建或联合伙伴，孵化聚焦行业痛点的生态 SaaS 应用，为行业内中小企业数字化转型提供原动力。Salesforce 面向电商零售行业，基于云计算整合从购物、订单履行到客户服务所有商务领域的一站式 SaaS 服务，依托

Salesforce 生态伙伴打造零售电商行业生态服务能力。中化集团化工行业云整合产业链资源，聚焦行业领域各类场景，能够为化工行业企业提供云计算垂类场景服务。运营方面，行业云平台构建完整的云资源交易链路，功能迭代呈现市场化特征。行业云构建公有云级别的资源交易全链路体系，支持租户进行云产品下单支付使用和账单结算等能力，通过计量计费、账单结算、伙伴厂商管理等运营技术实现云资源资产化管理模型。例如，上证云通过实现统一门户、产品运营管理平台与经营分析、应用商城管理平台等能力，面向证券、基金等行业企业用户建立市场化体系。

“云+稳定性”方面，云上系统稳定性融合多种技术，推动“稳保”进入主动化与透明化新阶段。近年来，云服务平台及大型互联网平台运行事故频发，凸显系统间互联与依赖性增强所带来的挑战。保障系统稳定性呈现出新策略、新防御和新观测的发展趋势。一是大模型技术引领运维策略的革新。大模型在运维领域的实践，能够加速故障识别与自我修复过程，优化资源调度防止过载，以及提供智能化决策支持。国内云厂商在大模型智能运维方面已开展积极探索，并在智能问答、辅助诊断、故障总结等场景形成了实践案例；国外厂商则尝试将运维大模型用于故障的理解和内容生成。二是混沌工程为系统稳定性构筑了坚实的前置防线。混沌工程作为一种通过设计并执行一系列实验，帮助发现系统技术架构与运营流程、隐藏风险与薄弱环节的技术，起初主要在运维团队内部获得了较高的

关注度。随着混沌工程的实践价值得到了更广泛的认可，其应用不再局限于运维阶段，而是扩展到了测试团队乃至整个软件开发生命周期中。在国内，腾讯云、京东云等厂商纷纷构建自己的混沌工程体系，并融入产品研发上线运营全流程；在国外，以美国的 Gremlin 公司为例，其提供的混沌工程平台已被多家跨国企业采用。三是端到端可观测性成为复杂系统故障发现和定位的关键。端到端的可观测性技术通过采集、分析整个服务链路上的数据，实现对系统状态的实时监控与精准故障定位。如国内多家云服务商正探索构建端到端的监控体系：贯穿业务、应用程序、云服务、基础设施层乃至硬件设备，旨在实现全链路监控能力的深度覆盖。全球范围内，Datadog、Splunk 等公司提供的全链路可观测性解决方案，亦在众多行业得到广泛应用。

“云+安全”方面，云原生安全呈原生化、一体化、智能化发展，加速企业云原生化转型进程。根据云原生计算基金会 CNCF 统计，2023 年 66% 的用户正在或已在生产环境中使用云原生，同时安全也是云原生用户面临的最大挑战。2023 年 Gartner 技术成熟度曲线显示，云原生安全技术正处于期望膨胀期，云上新增安全风险正推动云安全向云原生安全持续演化。国内外云原生安全创新技术和产品不断涌现，并呈现出原生化、一体化、智能化的特点。原生化从设计到运行，实现云平台和安全的深度融合。云原生安全的原生化一是要求将安全内嵌入云服务之中，提供高效的安全防护措施；二是

要求安全产品充分云原生化，提供灵活敏捷的安全服务。在云原生安全发展初期，安全技术提供商结合安全左移理念，更关注开发安全，聚焦于将容器安全、镜像安全、网络微隔离等工具与云环境深度融合。而当前安全原生已覆盖应用的开发、部署、运行，与 DevSecOps 理念相结合，确保每个阶段都内嵌安全措施，包括腾讯云、华为云、阿里云、小佑科技、青藤、亚马逊云、JFrog、Argus 在内的国内外厂商均推出了相关产品。一体化整合单点工具和安全平台，实现应用全流程一体化防护。随着云原生技术被应用到更多核心业务中后，以单点工具为主的云原生安全防护手段显然已经无法满足安全要求。其一，安全防护体系的木桶效应明显，单个安全产品的短板会影响整体安全防护；其二，云原生环境安全孤岛和整体复杂性提高，攻击盲点扩大，企业协同运维管理难度增加。因此，云原生安全开始从过去使用单点安全工具，或将多个解决方案组合解决安全问题的模式，向全流程一体化防护的模式转变。国内外安全厂商纷纷推出一体化云原生应用保护平台，例如 Palo Alto 推出的 Prisma Cloud、谷歌云推出的 Google Cloud Platform 等。智能化利用人工智能技术，提高安全防护的效率和准确性。随着 AI 技术的不断进步，云原生安全正朝着智能化加速发展。安全技术提供商一方面重点发展自身安全大模型，在安全产品中集成智能问答能力；另一方面侧重于将 AI 技术应用于安全运营中，实现智能研判和智能响应。在 2024 RSAC 大会中，包括 RSA Security 在内的十余家主要供应商

和 50 多家初创公司，已经发布了 AI 驱动的安全产品。

“云+算网”方面，算网云调度操作系统推动算、网、云协同发展，加速构建高效互联的算力互联网。随着人工智能大模型快速发展，数据传输从下行转向上行，应用任务对算力资源需求指数级上升，单一资源池难以匹配任务需求，算力互联互通成为主要目标，催生协同计算的算网云调度操作系统，推动不同通信主体、算力资源之间的网络互联，业务互通，加速构建传统互联网之上增强的算力互联网。当前，算网云调度操作系统呈现以下几个特点：**算力资源层面，算力标识定义算力资源，促进资源汇聚寻址。**算力互联网中计算任务在应用层基于算力标识匹配合适的资源和路径。算网云调度操作系统基于算力标识自动注册上报，汇聚算力资源信息，实现算力的描述、定位及寻址。例如，中国移动“百川”系统基于算力标识已接入超过 1.9EFLOPS 的社会算力，和自有算力形成互补。**调度层面，操作系统屏蔽底层异构，提供智算等场景下的任务封装，支持跨域调度，提供丰富的调度策略。**算网云调度操作系统提供基于任务封装的增删改查、数据卷挂载等能力，提供跨域、跨服务商的任务调度能力，提供 GPU 独占和共享调度等丰富的调度策略。例如，CNCF 开源算网云操作系统 KubeWharf 基于云原生技术及核心调度能力，支持大规模多租集群跨域调度和机器学习编排调度等场景。**网络传输层面，高性能传输协议 RDMA 解决大数据量远距离传输的网络传输问题。**算网云调度操作系统基于 RDMA 能够提供卡间、

DPU 间的大数量远距离、低时延传输，也可以提供实时准确的卡间、服务器间数据交换监测。

二、我国云计算发展概述

（一）中央与地方协同推进云计算，助力产业高质量发展

2024 年 1 月 31 日，习近平总书记在主持中共中央政治局第十一次集体学习时强调，“发展新质生产力是推动高质量发展的内在要求和重要着力点”，“必须继续做好创新这篇大文章，推动新质生产力加快发展”。新时代新征程，以中国式现代化全面推进强国建设、民族复兴伟业，实现新型工业化是关键任务。深化推进云计算与实体经济融合升级是落实新型工业化大会精神“把建设制造强国同发展数字经济、产业信息化有机结合”，“大力推进数字技术与实体经济深度融合”的关键抓手，是技术革命性突破、生产要素创新性配置、产业深度转型升级的重要路径。

国家层面，政策支持力度不断加大，持续加强云计算基础设施建设和行业应用。基础设施建设方面，2024 年政府工作报告指出发挥好政府投资的带动放大效应，重点支持科技创新、新型基础设施、节能减排降碳。2023 年 10 月，工信部等六部门联合印发《算力基础设施高质量发展行动计划》，提出“推动以云服务方式整合算力资源，充分发挥云计算资源弹性调度优势”，促进多方算力互联互通。2024 年 1 月，工信部在《云计算综合标准化体系建设指南》中，明确优先制定云计算创新技术产品、新型服务应用和重要缺失领域的关键

标准。行业应用方面，2023 年 3 月，中共中央办公厅、国务院办公厅印发了《关于进一步完善医疗卫生服务体系的意见》，提出要加快推进云计算等在医疗卫生领域中的应用，加强健康医疗大数据共享交换与保障体系建设。2023 年 11 月，《工业和信息化部关于健全中小企业公共服务体系的指导意见》提出，要充分运用云计算等技术，建立贯通国家、省、市、县各级中小企业公共服务机构的全国中小企业服务“一张网”，与政务服务平台加强互联互通。2023 年 12 月，工信部等八部门联合发布《关于加快传统制造业转型升级的指导意见》，提出大力推进企业智改数转网联，立足不同产业特点和差异化需求，加快云计算等信息技术与制造全过程、全要素深度融合。

地方层面，各地积极贯彻落实国家战略部署，结合区域特点和发展需求，推动云计算技术创新和产业智能化应用。一方面，各省市通过政策支持和资金投入，大力推动云计算关键技术的研发与创新。例如，北京、山东等地采用“揭榜挂帅”和“赛马制”等方式，推动云计算重大研究项目和关键技术攻关；浙江省聚焦云计算等五大战略领域，重塑“互联网+”战略科技力量体系架构，加快打造一批重大战略成果；广东省致力于形成城市级分布式云计算能力，促进云计算与边缘计算的协同发展。另一方面，各地政府结合本地产业特点和发展需求，积极推动产业数字化转型和智能化升级，促进云计算技术的实际应用。例如，北京市大力推进智能制造，通过“新智造 100”工程和企业数字化转型，构建先进的算力中心和 5G 基站，

推动产业数字化和上云应用；天津市加快新型工业化建设，推动智能制造和数字化转型，鼓励中小企业“上云上平台”和数字赋能，建设智能工厂和数字化车间；山东省聚焦智能化和集群化，实施制造业数字化转型提标行动，推动中小企业“上云用数赋智”，加快智能制造和先进产业集群建设。此外，各地还在积极探索大模型和人工智能技术在云计算中的应用。例如，深圳市通过建设城市级和企业级智能算力平台，推动智能算力云服务，提升产业集聚水平；湖北省加快打造智能算力平台，推动大数据和云计算中心建设，支持公共数据资源开放共享，进一步提升人工智能技术在云计算中的应用水平。

（二）云计算市场迎来新一轮增长，AI 加速产业格局重新洗牌

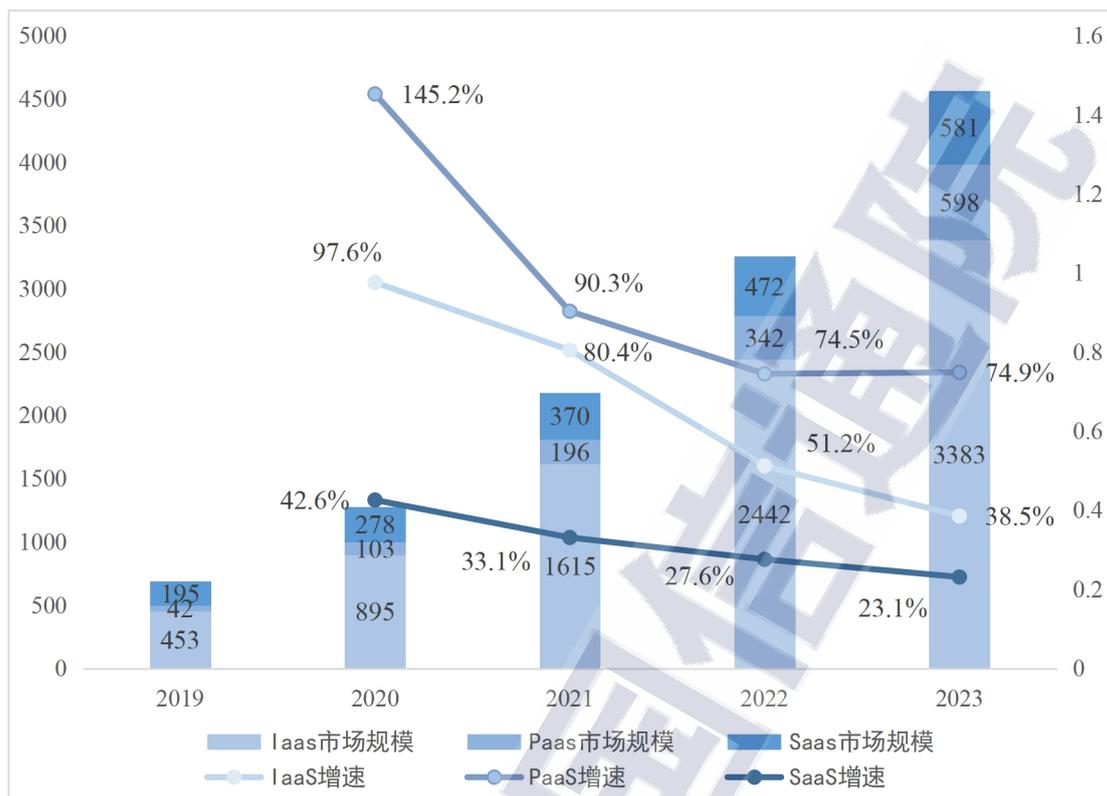
从整体来看，我国云计算市场保持较高活力。据中国信通院统计，2023 年我国云计算市场规模达 6165 亿元，较 2022 年增长 35.5%，大幅高于全球增速。其中，公有云市场规模 4562 亿元，同比增长 40.1%；私有云市场规模 1563 亿元，同比增长 20.8%。随着 AI 原生带来的云计算技术革新以及大模型规模化应用落地，我国云计算产业发展将迎来新一轮增长曲线，预计到 2027 年我国云计算市场规模将超过 2.1 万亿元。



来源：中国信息通信研究院，2024 年 5 月

图 4 中国云计算市场规模及增速（亿元）

从细分领域来看，AI 推动市场增长点向 PaaS、SaaS 上移。2023 年我国 IaaS 市场增速达到 38.5%，市场总额达 3383 亿元，电信运营商在 IaaS 领域的市场份额稳步上升，为整体市场营收持续提供增长动力；PaaS 市场达 598 亿，同比增长 74.9%，得益于公有云出海业务及 AI 发展的需求，预计 PaaS 领域产品将持续增加；SaaS 市场渗透率逐年提升，2023 年市场总额达到 581 亿元，增长率为 23.1%。未来，随着 AI 大模型进入商业落地阶段，预计大量中小型创新企业和投资公司涌入 SaaS 领域，商业化应用将全面发展。



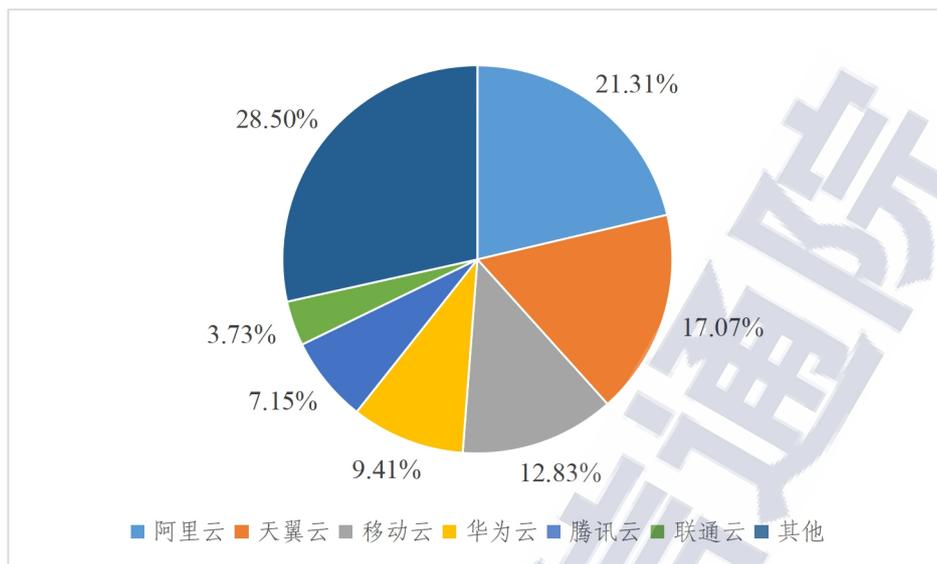
来源：中国信息通信研究院，2024 年 5 月

图 5 中国云计算细分领域市场规模及增速（亿元）

从厂商层面来看，产业升级带来新机遇，市场格局或面临重新洗牌。据中国信通院调查统计²，阿里云、天翼云、移动云、华为云、腾讯云、联通云占据中国公有云 IaaS 市场份额前六³；公有云 PaaS 方面，阿里云、百度云、华为云、腾讯云、天翼云、移动云处于领先地位。受生成式 AI 和大模型等新技术影响，云计算市场面临新机遇，中腰部厂商借助科技平台优势强势发力，快速推动原有云业务智能化转型从而抢占市场，厂商格局或借此机遇进行洗牌。

² 市场规模为 2023 年全年数据统计，主要依据企业财报、人员访谈、可信云评估、历史数据等得出。对于市场数据不明确的领域，只发布头部企业整体情况，不做具体排名。

³ 因为 IaaS 和 CDN 是两种业态，需要分别获得互联网资源协作服务业务牌照和内容分发网络业务牌照，所有 IaaS 不包括 CDN 收入，只统计计算、存储、网络等基础资源服务收入。



来源：中国信息通信研究院，2024 年 5 月

图 6 2023 年中国公有云 IaaS 厂商市场占比

（三）各行业上云积极性持续攀升，梯队分层效应依旧明显

去年，《云计算白皮书（2023 年）》全面分析了我国各行业上云水平参差不齐的现状，并以成熟期、成长期、探索期三个梯队形式梳理了各行业应用发展情况。今年，本蓝皮书从各梯队中选取政务、交通、能源等三个具有代表性行业，详细介绍其上云发展特点及成果，以期发挥标杆示范作用，促进云计算应用普及深化，加快推进产业高质量发展。

政务方面。政务云从“量的增长”向“质的提升”全面升级，不断迸发新活力。政务云作为数字政府建设的关键基础设施，正步入融合升级与智能跃进的新时期。整体来看，政务云发展趋势呈现出四大显著特征：**一是政策环境优化升级，促进政务“一体化”建**

设提速。面向新阶段，各省市先后对政务云建设发展提出了新要求、明确新规范。河南省以“一云统构”为基座，“一云纳管”为中枢，“一云聚数”为目标，构建形成河南省一体化政务云体系。江苏省采用“一核多点”总体布局和“三横四纵”整体架构，统筹建设全省政务“一朵云”，各级部门单位和县（市、区）不再单独建设或外租云资源。

二是新技术融合创新，释放政务云核心价值。越来越多的政务应用将基于云原生进行搭建，可以更好的满足业务轻量、敏捷、高度自动化需求，保障政府数字化应用的“广度”和“韧性”。此外，算力强化与 AI 大模型的深度融合，为政务云解锁新潜能，实现政务云的高效运营和智能化发展，有力提升政务云对上层业务的赋能作用。

三是场景应用多元化，赋能服务深度与广度提升。政务云服务场景日益多元化，全面覆盖政府治理与民生服务多个维度。在政务服务领域，政务云支撑智慧政务大厅、在线审批系统等应用，极大简化行政流程，提升政府办公效率；在社会治理方面，环保监测和城市综合治理通过集成云计算与物联网技术，实现实时监控与智能分析，助力构建绿色智慧城市；此外，结合区块链技术的政务云在监管、审计等场景的应用，进一步增强了政务活动的透明度与信任度。

四是促进开放合作，构建生态产业新格局。政务云市场正迎来多元竞合的新时期，为政务需求打造的云服务商与大型模型制造商相继入场，给市场带来新活力，促进产业链条上下游深度融合与协同发展，为政府数字化转型提供强大的技术支撑和生态环境。

交通方面。交通云重心从基础设施向应用层转移，水陆空部分细分行业进入发展快车道。交通云作为数字交通建设关键，在推动智慧交通场景落地、支撑产业创新发展方面发挥重要作用。整体来看，交通行业上云用云主要呈现以下三个特点：一是上云模式方面，交通行业由资源上云向应用上云过渡，PaaS、SaaS 等高阶服务渗透率逐步提高。机场云、城轨云等细分行业由传统虚拟机、一体机为主的基础资源池化模式逐步迈向容器化、微服务等云原生架构。物流行业核心系统如运输、仓储、配送等应用基本采用云原生架构，例如申通快递云原生已覆盖运输、仓储、流通加工、配送、信息服务等环节，顺丰快递通过云原生大数据平台打造智慧物流供应链，实现整体物流系统的网络化、自动化、可视化。二是技术创新方面，新兴技术加速升级，驱动交通云底座建设再上新台阶。近年来，车路协同、数字孪生、5G、物联网等新兴技术加速升级，有效驱动了云计算产品的深度使用，为新兴技术发展提供肥沃土壤，从而助力提升交通出行体验和货物运输效率。如腾讯提出了融合“实时孪生+AI+交通 OS”分布式交通云解决方案，促进新兴技术+云+交通的广泛应用，百度推出交通行业大模型，增强云计算在数据融通、算力资源调度和交通个性化服务等方面的能力建设。三是应用场景方面，水陆空云上应用多点开花，公路、民航、物流领域进入发展快车道。公路领域，山东、四川等地基于云计算落地智能化交管系统和车路一体化协同，有效提高公路交通安全性和效率；民航领域如

深圳、上海等地机场集团加快“机场一朵云”建设，针对航班运行、旅客服务等方面加快基础设施云化改造，持续提升旅客便捷高效的出行体验；物流领域如京东物流、菜鸟等企业积极探索并推动智能仓储、自动化分拣等方面的云上应用实践，不断提高物流效率和准确性。

能源方面。能源云聚焦底座规划和建设能力，并积极探索将云应用在能源生产等业务环节。当前，能源行业上云用云尚处于夯实基础阶段，主要有四大特点：**一是能源行业转型升级政策持续加码，提升云平台建设与应用水平已成为共识。**国家能源局《2024 年能源工作指导意见》指出，深入实施创新驱动发展战略，聚焦高端化、数字化、智能化，加强能源科技自主创新，提升能源产业链供应链自主可控水平，促进新质生产力发展。2024 年 3 月国资委党委扩大会议指出，加快建设新能源基础设施，推进电网基础设施智能化改造和智能微电网建设，着力提高电网对清洁能源的接纳、配置和调控能力。随着国家在能源行业上云相关的政策逐渐落实，能源产业升级将进一步推动能源行业上云用云水平。**二是能源云建设从基础设施云化改造，向综合能源服务云平台转变。**能源产业具备环节多、链条长、应用范围广的特点，通过对发电厂、输电线路、变电站炼化系统、供应系统等能源基础设施进行云化改造，能源企业可以充分挖掘能源数据价值，加强产业链各环节协同，并结合数据建模、人工智能等云上应用对外提供服务。例如，中国华能、上海电气等

企业通过构建能源综合服务平台，开展数据管理、能源金融、资源勘探等多个数据服务，实现了能源领域资源与数据的流通。三是能源云“龙头效应”明显，电力、油气领域应用落地进程较快。能源产业作为国家支柱产业，呈现重点领域大型央国企牵头上云，其他领域中小企业积极跟进上云的特点。国家电网、中国海油等头部能源企业目前已经开展探索能源云的二次开发与应用，并形成面向碳排放计算、能源资源分布、能源接入流程上云等落地应用。其他如水力、煤矿等行业在能源生产、监管维护等基础业务领域具备应用覆盖，目前正在持续探索能源深度用云场景。四是“云计算+人工智能”模式逐步应用于能源业务生产各环节，成为落实“双碳”目标重要抓手。企业在能源勘探、碳排放监测过程中积累了大量生产数据，通过对数据进行整合与处理，可以围绕碳排放的降低与平衡进行人工智能训练，从能源认知分析、能源生产优化、能源预测维护等多方面探索“双碳”战略的进一步落实方式。例如，中国海油依托人工智能云平台将岩层、油气产能数据进行评估与分析，提升能源勘探效率减少碳排放。山东能源成立人工智能联合创新中心，发布油气行业的 AI 大模型，应用场景涵盖采煤、焦化、安监等应用场景，推动煤矿生产迈向智能化管理，有效降低了企业的碳排放。

三、中国云计算发展评价

（一）云计算发展评价体系

过去十多年，在政策引导和产业界各方共同努力下，我国云计算发展已从概念期进入到创新活跃、广泛普及、应用繁荣的新阶段，取得了积极进展。市场规模从“十二五”期间的几十亿元增长至现在的数千亿元，技术创新持续突破，部分指标达到国际先进水平，行业应用加速向政务、金融、制造、交通、能源等领域渗透，成为加速我国新型工业化、推动形成新质生产力的重要支撑。

今年恰逢中国信通院第 10 次发布云计算蓝皮书，为客观全面地梳理我国各地区各行业云计算发展水平，中国信通院云计算团队基于多年来对云产业的研究，形成了云计算发展评价体系，该体系包含云资源供需发展水平、云服务应用发展水平、云技术融合发展水平，旨在衡量一个地区或行业的云计算供需水平、应用深度、技术融合等情况。

其中，云资源供需发展水平评价包括供云量、用云量两个维度，展现我国云计算资源供需现状；云服务应用发展水平评价包含行业应用发展水平、企业上云发展水平、云服务体验水平三个维度，反映云服务应用落地能力和应用服务深度；云技术融合发展水平评价包括算云融合水平和数智云化水平，主要反映通用计算、智算、超算等算力云化进程以及新技术云化赋能水平。

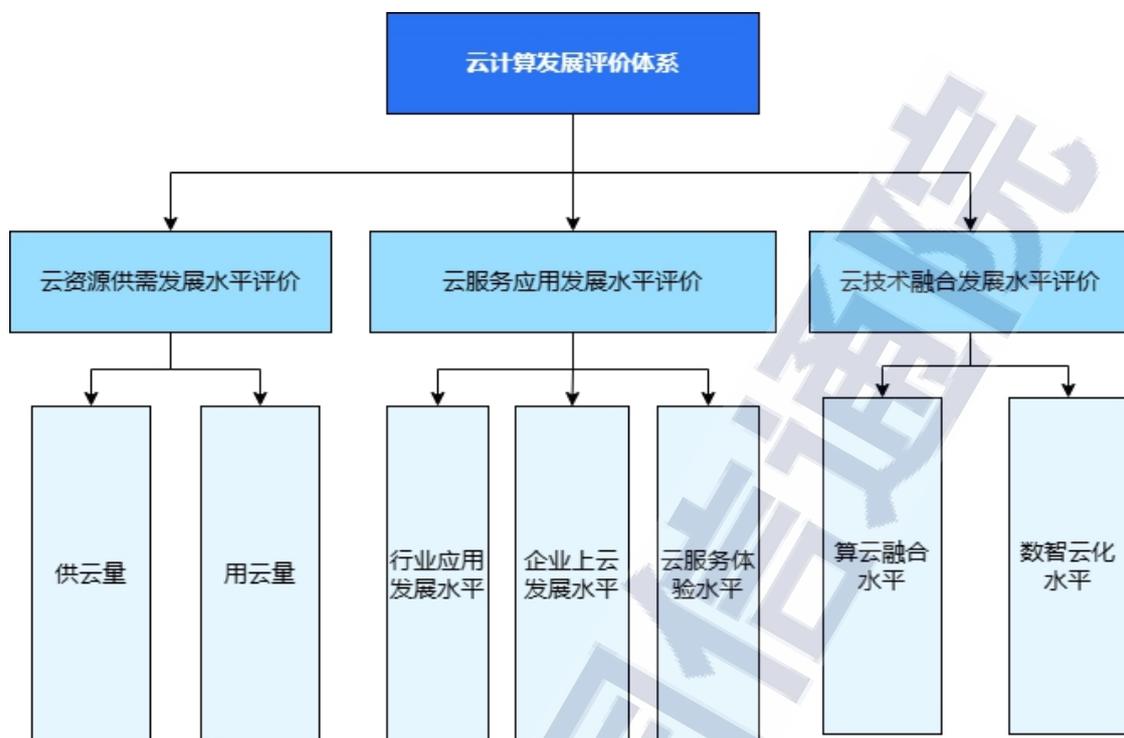


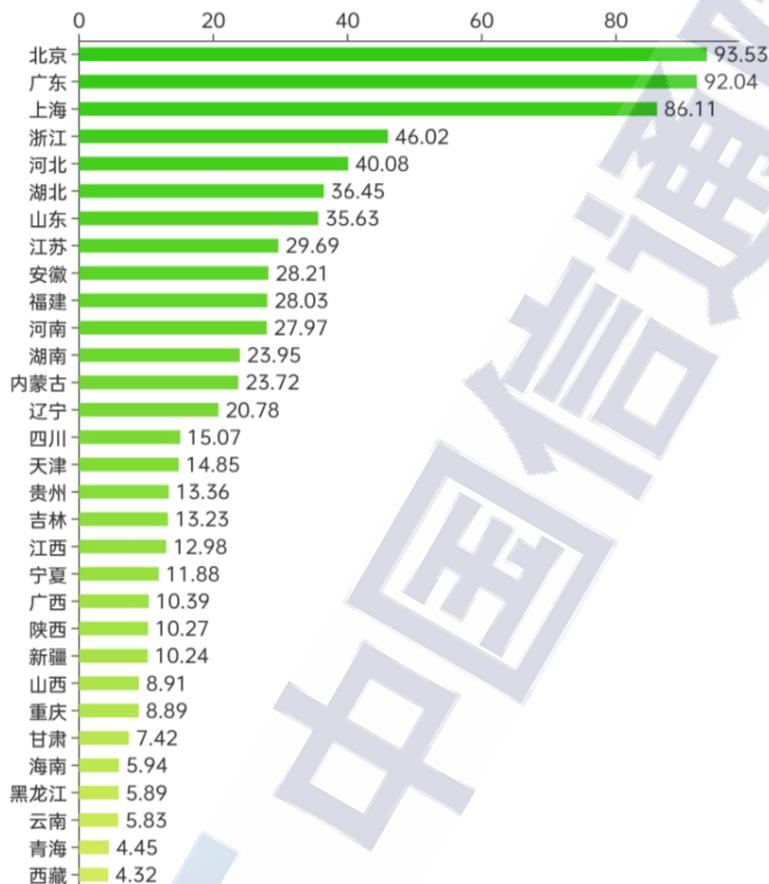
图 7 云计算发展指数体系

（二）云资源供需发展水平评价

1. 供云量

供云量指云服务商在各地区能够提供的云计算供给规模，反映各地区云计算供给水平。供云量评价综合考虑云服务商的云服务覆盖范围、IaaS/PaaS/SaaS 服务规模、云服务价格等因素。根据中国信通院测算，我国 31 个省(自治区、直辖市)云计算对用户可供给 vCPU 核数超过 10 亿个，云存储规模超 50EB。从区域上看，中东部地区供云量领跑全国，目前各大云服务商在中东部地区的云计算供给量占整体公有云资源的 60%左右；西部地区整体处于发展和蓄能阶段，

如贵州数据中心资源储备量大，但云化程度较低，其整体供云量处于中游水平，未来结合政策将有巨大发展空间。



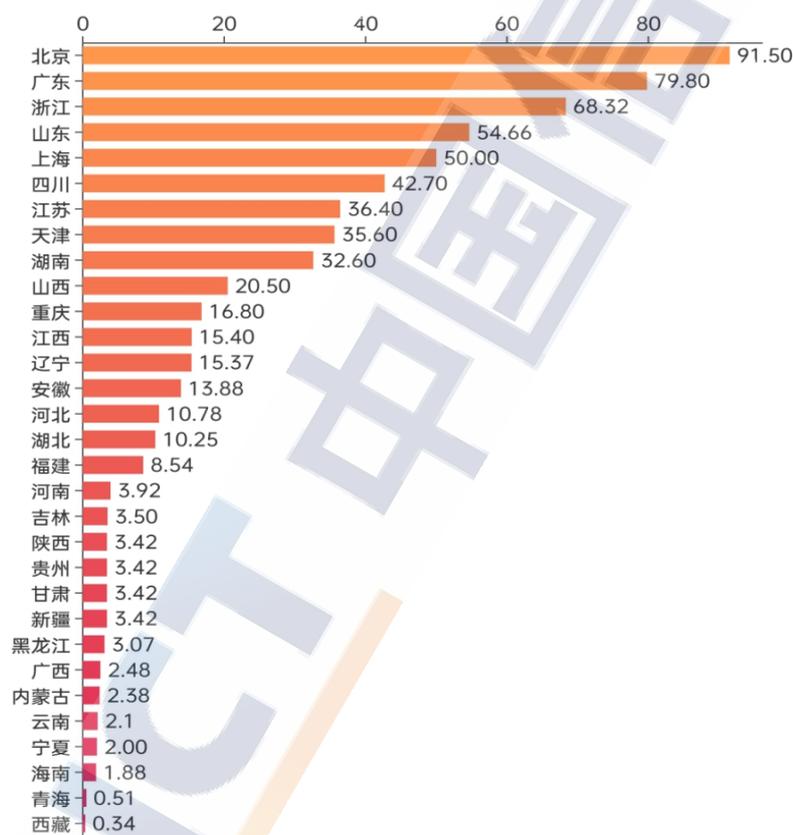
数据来源：中国信通院，2024 年 5 月

图 8 2023 年 31 个省（自治区、直辖市）供云量指数

2. 用云量

用云量指各地区使用云计算相关产品或服务的情况，反映各地区上云用云的程度。用云量已成各省数字经济发展晴雨表和数字经济发展活力的指标，各省用云水平深度与数字经济发展成正相关。用云量排名前十的省份与数字经济发展较好的省份基本一致，随着

省市用云量提高，云计算所拉动的产业数字化增长随之升高。从地域分布看，全国大部分省份已完成初步上云，用云量第一梯队主要集中在京津冀、长三角、珠三角等地，东南地区用云量明显高于中西部地区。结合各地区供云量数据来看，我国部分省份存在云资源供需不平衡的情况，例如贵州、宁夏、内蒙古等省份的供云量远高于其用云量，川渝地区的用云量则高于其自身供云量。



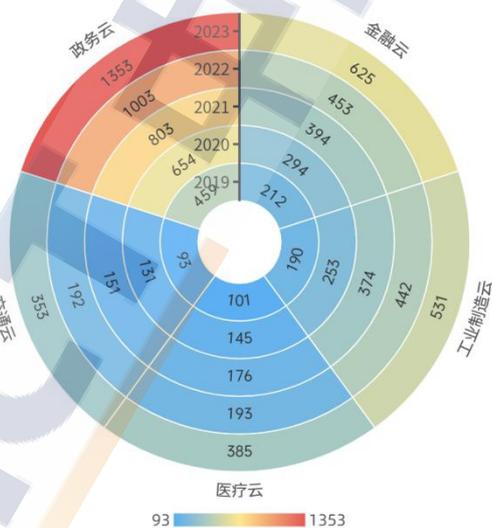
数据来源：中国信通院，2024 年 5 月

图 9 2023 年 31 个省（自治区、直辖市）用云量指数

（三）云服务应用发展水平评价

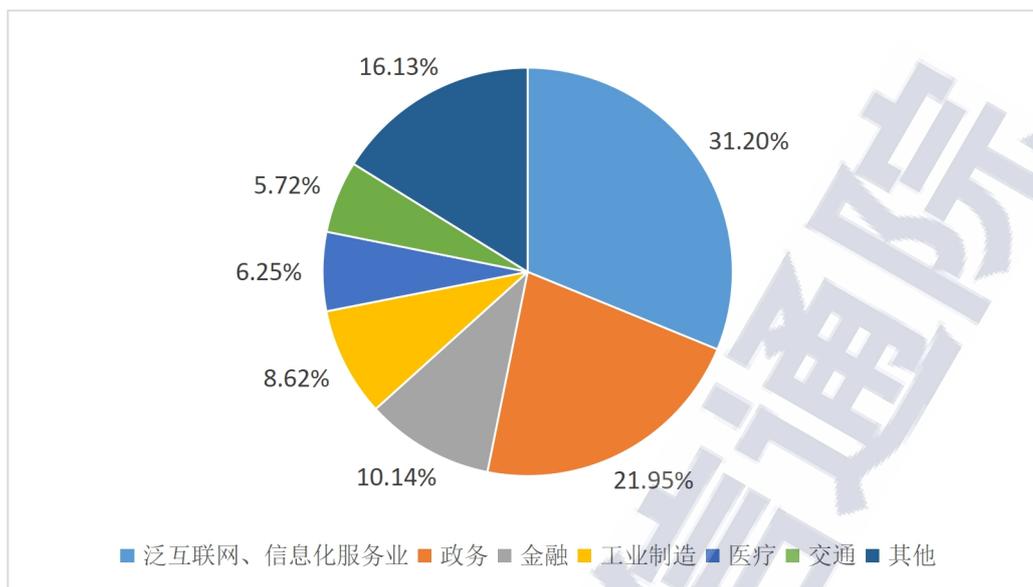
1. 行业应用发展水平

行业应用发展水平通过行业过去五年间云计算市场规模，横向对比各行业间的云化热度，纵向对比行业用云量占比，从而反映各行业企业上云用云的程度。从行业用云热力值来看，政务云占据领头地位，逐年跨级增长。金融云、工业制造云暂处“温带”，在行业政策推动下，上云用云积极性将大幅度提高。医疗、交通行业发展速度相对较慢，随着医疗云行业应用落地及汽车云概念兴起，未来发展空间巨大。从行业用云量占比看，以泛互联网、电商为代表的互联网原生行业数字化、信息化程度较为领先，占市场总额的 1/3 左右。以政务、金融、工业制造为代表的传统行业呈追赶状态，随着“上云用数赋智”理念深入，用云量规模将不断扩张。



数据来源：根据公开数据整理

图 10 2019-2023 年典型行业用云热力图

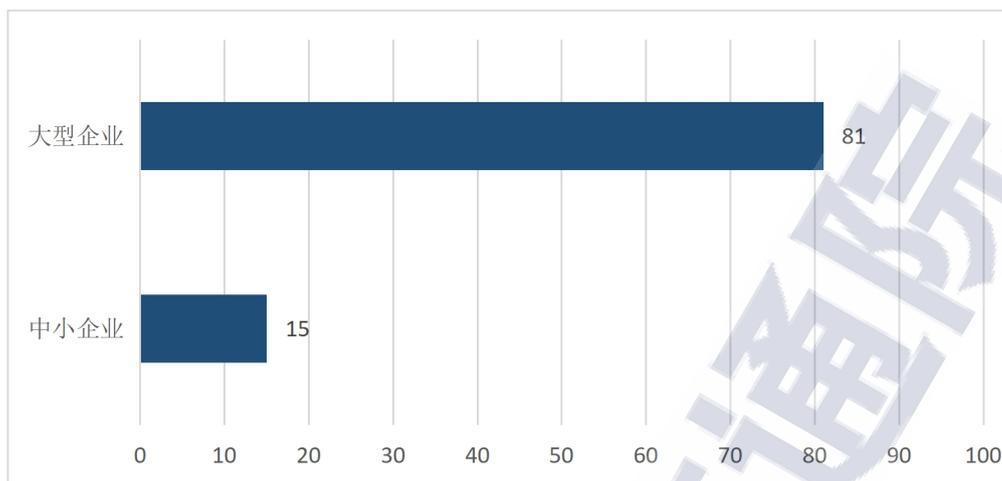


数据来源：根据公开数据整理

图 11 2023 年行业用云量占比

2. 企业上云发展水平

企业上云发展水平通过企业信息系统的上云率来体现，反映不同规模企业应用云计算的程度。大型企业通常自身盈利能力较强，IT 基础设施投入较大，已从“资源上云”向“深度用云”阶段发展，各类核心业务系统已逐步迁移到云端。目前我国大型企业上云率超过 80%，正在成为提升企业应用发展水平的主力军。中小企业数量众多遍布各行各业，具有资金流转快、技术人才需求高、上云周期短等特点，主要通过 SaaS 模式快速上云，如协同办公、客户管理、财税费控、电子合同等。我国中小企业接近 5000 万户，是经济发展的重要力量，目前整体上云率在 15%左右，未来发展空间巨大。



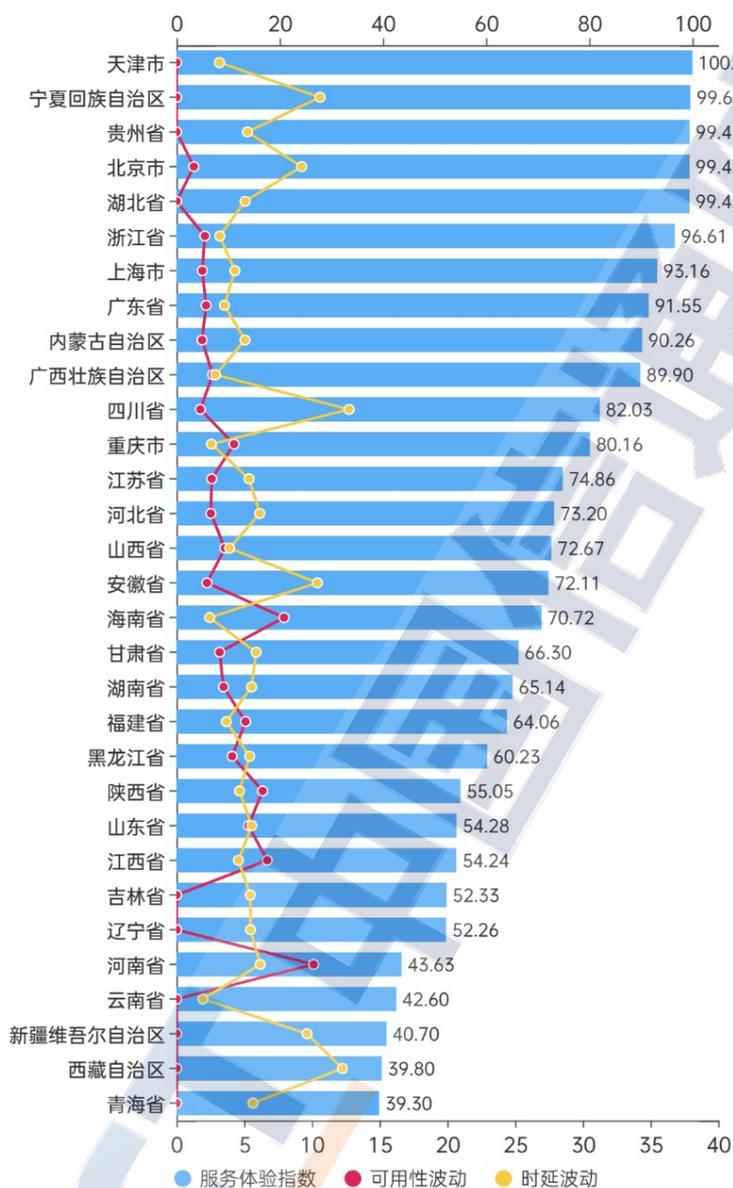
数据来源：根据公开数据整理

图 12 2023 年不同规模企业上云率

3. 云服务体验水平

云服务体验水平指用户使用云服务时感知到的服务体验质量。

根据中国信通院可信云监测,我国东部地区 IaaS 服务质量水平较高,整体运行情况较为稳定,2023 年平均服务可用性达 98.43%;西南、中部地区公有云 IaaS 服务质量在可用性、访问时延等方面仍有优化空间,服务质量相比于京津冀、长三角、粤港澳三大经济区低 20%左右。2023 年,我国 SaaS 服务的年度可用性高于 99.5%,服务平均响应时间波动较大,服务稳定性体验方面仍有较高提升空间;其中,企业协作、财务管理、点播直播等企业级 SaaS 服务的整体运行情况表现较好。



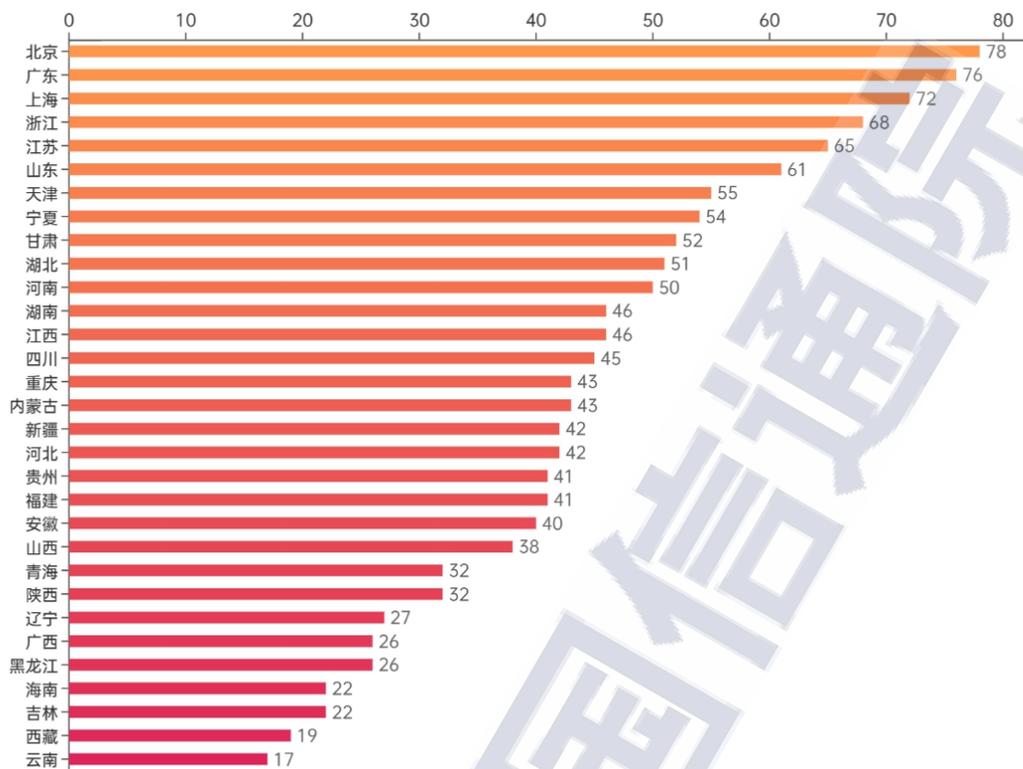
数据来源：根据公开数据整理

图 13 2023 年 31 个省（自治区、直辖市）IaaS 服务体验指数

（四）云技术融合发展水平评价

1. 算云融合水平

算云融合水平指通用算力基础设施融合虚拟化、云原生等云计算技术实现算力及其关联资源的云服务化输出，反映各区域的算力云化程度。指数综合考虑各区域算力规模、算力云化比例及云化资源利用率，对 31 个省份展开客观评价。整体上，我国各省份算力规模及算力云化程度差异大，云化势头强劲，发展空间巨大。以北京、广东、上海、浙江为代表的省份处于第一梯队，算力建设布局普遍较早，算力云化程度高。此类省份对算力规划布局早，云化比例均超过 65%，算力建设经验丰富且已经形成了规模化的算力基础，同时大力开展云计算、算力服务相关技术与落地，形成了体系化的算力云化输出，算云融合指数高。以宁夏、甘肃为代表的省份处于第二梯队，算力建设布局较晚，但算力云化程度较高。这些省份算力规划和建设虽然较北上广等起步晚，但及时拥抱算力等新技术，在建设过程中即与云计算深度融合，算云融合程度较高。以贵州、内蒙古为代表的省份处于第三梯队，算力建设布局较晚，算力云化程度相对较低。这些省份基于地理、环境等优势完成了较大规模的算力资源建设，但整体云化比例偏低，在算云融合方面仍需要持续投入，以实现算力资源的高效利用与价值转化。



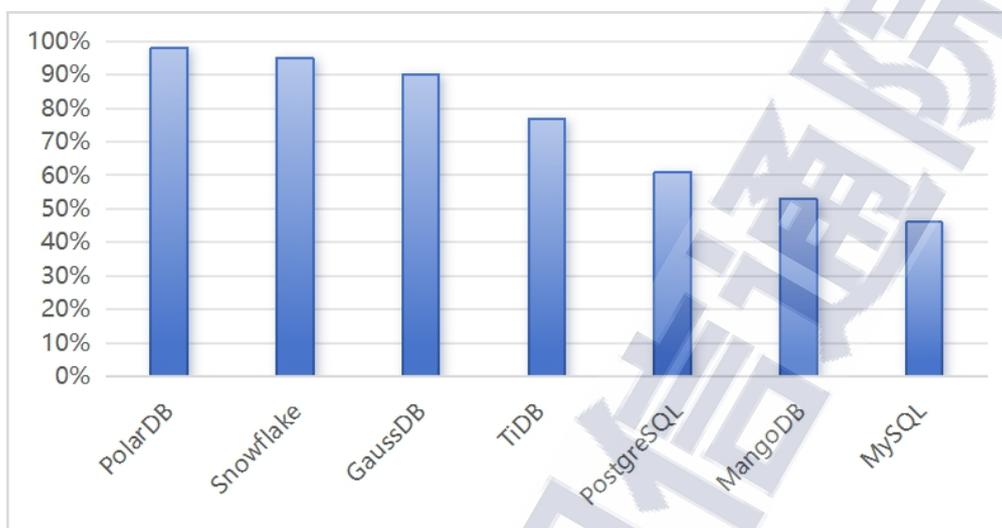
数据来源：中国信通院，2024 年 5 月

图 14 2023 年 31 个省（自治区、直辖市）算云融合指数

2. 数智云化水平

数智云化水平指数据服务与智能计算服务融合虚拟化、云原生等云计算技术实现在不同场景的应用和服务部署的云化程度，指标包括数据服务云化部署比例、智能计算服务云化部署比例。在数据服务云化部署方面，云原生数据库由于其高弹性、高可用等优势逐渐成为用户数据库服务的主流选择，云化部署比例极高，如 PolarDB、Snowflake、GaussDB、TiDB 等，均超过 70%。在智能计算服务云化部署方面，图形渲染、AI 推理等场景云化部署比例较高，均超过 25%，

生物科学、科学计算等场景因受其计算范式所限，多使用本地部署的智能算力，云化部署比例在 10%-20%之间。



数据来源：中国信通院，2024 年 5 月

图 15 数据服务云化部署比例



数据来源：中国信通院，2024 年 5 月

图 16 智能计算服务云化部署比例

四、云计算与智算加速融合，开启智能化新纪元

当前，全球人工智能技术迅猛发展，已经成为世界科技强国重

点布局的关键赛道。随着智能算力逐渐成为算力结构的主要组成，传统的通用云计算正加速与智算融合，升级成为可服务于人工智能技术和应用发展的**智能云**。智能云通过对大规模异构智算资源的融合与调度，能够屏蔽各种底层复杂的计算资源、兼容多种芯片架构和开源框架，提供丰富的云计算工具，提高算力资源利用率，保障各种 AI 模型算法在智能云平台上实现高效便捷地运行。从体系架构来看，智能云通常包括**智能云基础设施服务 AIaaS**、**智能云平台服务 AIPaaS**、**大模型服务 MaaS**、**智能云应用服务 AISaaS** 等，具体如图 17 所示。智能云作为新一代人工智能发展的驱动力量，正在重构云服务产业格局，加速形成新质生产力。

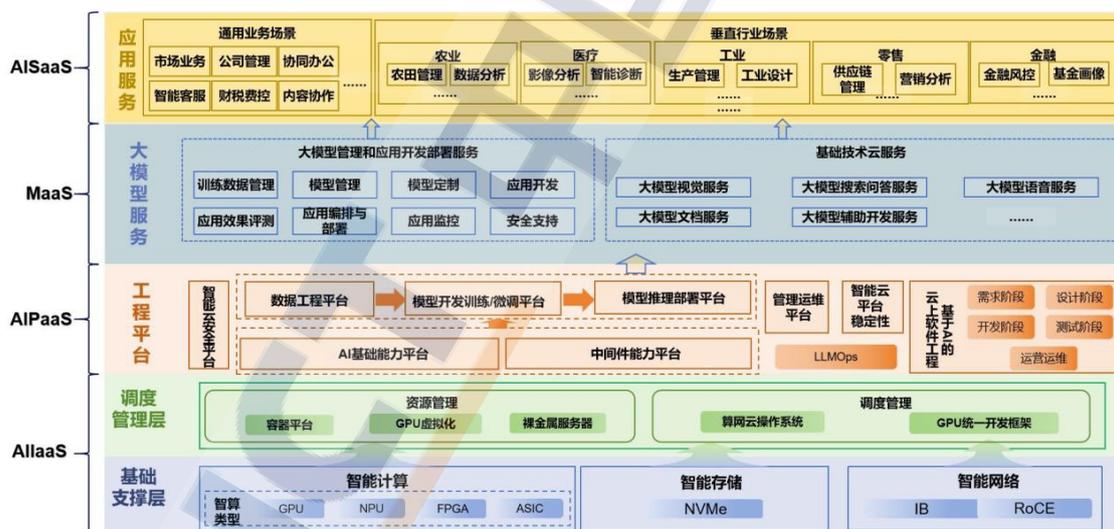


图 17 智能云体系架构图

（一）AIaaS 夯实云计算技术底座，持续提升智能软硬件效能

AIaaS 以智算资源为核心，依托对智算资源的融合提供智能化

云服务，并使用调度管理工具充分发挥智能云资源池能力。在传统模式的 IaaS 中，云资源池以通用计算为主体，随着全球 AI 大模型的井喷式发展，智能算力需求增速远超芯片性能提升和产能扩张速度的上限，其中以 GPU（Graphics Processing Unit，图形处理器）为代表的芯片成为提供智能算力的主力军，借助云计算能够实现零散智算资源集中与纳管的优势，各大云厂商纷纷在智算领域进行布局，形成千卡、万卡智能云集群，AIaaS 成为云计算新范式--以云服务的方式提供可便捷获取的智能算力。云计算与智算资源融合形成的智能云能够为大模型训练和推理提供充足的算力资源，已经成为促进人工智能领域发展的坚实技术底座。

AIaaS 提供智算基础支撑，为智能云提供坚实的底层资源。一是智能计算能够提供丰富多样的智能云资源池。区别于传统的依靠 CPU 提供算力，在 AIaaS 中主要依靠 GPU、NPU（Neural Processing Unit，神经网络处理单元）、TPU（Tensor Processing Unit，张量处理单元）等新型芯片提供智算资源，例如谷歌云通过其云服务平台 Google Cloud 将 TPU 作为服务提供给外部用户，使得开发者和企业能够利用 TPU 的强大计算能力进行机器学习模型的训练和推理。二是智能网络能够提供适用于智算场景的组网选择。区别于传统的以太网组网选择，智算场景对云资源池的网络提出了高带宽、低延时和高可靠的要求，以 IB（InfiniBand，无限带宽）、RoCE（Remote Direct Memory Access over Converged Ethernet，基于聚合以太网的远程内

存直接访问技术)为代表的高性能远程直接内存访问(Remote Direct Memory Access, RDMA)网络成为智能网络的首选,例如火山引擎在高性能计算集群内使用 RDMA 技术实现高速网络互联,腾讯云的高性能计算集群以高性能云服务器为节点,通过 RDMA 互联,大幅提升网络性能。三是智能存储能够提供适用于智算场景的存储能力。目前云计算的后端存储系统多为基于开源的软件存储协议,主要使用 Ceph(一种开源的分布式存储系统),iSCSI(Internet Small Computer System Interface,互联网小型计算机系统接口)或者厂商自研的存储平台。由于目前使用的 Ceph 和 iSCSI 在协议上的限制,在高性能块存储的场景下,需要在存储的后端和前端通过基于 SPDK(Storage Performance Development Kit,存储性能开发工具包)和 DPDK(Data Plane Development Kit,数据平面开发套件)的用户态框架来提升带宽性能,在新型的块存储上,各大厂商正在向 NVMe(Non-Volatile Memory Express,非易失性快速存储器)接口演进。

AIaaS 调度管理层提供资源管理功能,助力智能云提效升级。

一是资源管理能够提升智能云多类型资源的调度分配能力。在 AIaaS 中存在多种资源管理模块,包括容器平台、虚拟化平台和裸金属服务器等,各个资源管理模块使用智能调度和管理功能自动分配和优化资源调度,实现提高整个系统工作效率的目标。二是基础软件创新优化智能云调度管理能力。GPU 作为智能云的核心计算资源,在对外提供云服务时仍存在异构化的特点,为满足用户的不同

需求，在 AIaaS 中部署统一的 GPU 开发框架有助于消除支撑层智算资源异构化带来的弊端，加速 AI 应用落地。在调度管理层面，使用算网云操作系统完成智算资源的调度管理，充分发挥智算集群的工程化能力，实现大规模芯片资源高效协同工作。三是关键技术突破提升智能云资源利用效率。虚拟化作为云计算领域的核心技术之一，能够使计算资源更加灵活和高效。在 AIaaS 中，GPU 虚拟化技术能够通过虚拟化管理器实现对 GPU 资源的抽象和管理，提升智能云对计算资源尤其是智算资源的高效使用。例如微软云和谷歌云都提供了 GPU 虚拟化的实现方案，允许将单个物理 GPU 分为多个虚拟 GPU，供不同的虚拟机使用；火山引擎具备容器共享功能，实现 GPU 算力与显存的灵活切分与隔离，提升 GPU 的使用率。

（二）AIPaaS 聚焦智算服务开发能力，成为 AI 工程化实践最优解

随着大模型技术的火热以及云计算技术的发展，打造智能算力、通用算法和开发平台一体化的新型智能云工程平台成为云服务商竞相发力的方向。智能云工程平台在传统的云服务基础上融合智算资源，更加注重平台在智算服务开发能力的丰富和创新，为人工智能的发展提供更加高性能的计算、更加智能的运维、更加稳定的平台性能和更加灵活的应用生成能力。

AIPaaS 提供丰富的 AI 工具组件和模型及算法仓库，让用户模块化、定制化进行人工智能开发。AIPaaS 通过融合强大的智算资源，提

供高效的开发工具、通用模型仓库、私域模型仓库以及集成高质量的数据集，为满足人工智能从零开始训练大规模模型、处理复杂计算任务的需求提供了强有力的支撑。平台丰富的算法库、数据集和开发工具，通过自动化的流程管理和优化，降低了人工智能应用的开发门槛。平台内置的通用大模型和私域模型，还可以实现模型的开箱即用，提高开发效率。例如腾讯云 TI 平台、火山云方舟平台，支持多种编程语言、框架和工具，满足不同开发者的需求，提供包含各种预训练模型的通用模型仓库，可以供开发者直接使用或作为迁移学习的起点。这些平台还提供丰富的数据集，包括公开数据集和定制数据集，帮助开发者进行模型训练和测试。

AIPaaS 是一站式模型开发训练推理部署平台，成为大模型时代应用生成的加速器。随着模型规模的不断增大和复杂度的不断提高，其对计算资源和开发效率的要求也越来越高。AIPaaS 在智算资源的加持下，提供一站式的大模型开发部署流程，包括数据预处理、模型训练、模型评估、模型部署等。这些流程可以通过自动化工具进行高效执行，减少了开发者的重复劳动，提高了开发效率。这种集中化的资源管理、自动化的流程执行、高效的协作方式、弹性扩展的能力，为大模型时代的人工智能应用生成提供了强有力的支持。智能云工程平台继承了传统通用云平台开发和部署流程的优势，支持多用户协作开发，不同的开发者可以同时在上进行工作，通过版本控制等工具进行代码和数据的共享和协作，这种高效的协作方

式可以加快模型的迭代和优化速度，提高模型的性能。例如阿里云灵骏、天翼云慧聚以及移动云九天智算，这些智能云工程平台提供了从数据到模型应用的全流程开发服务，通过自动化的流程管理和优化，突破原有 AI 能力构建效率低、成本高、迭代混乱等限制，构建通用 AI 中台，加速模型快速生成和应用落地。

AIPaaS 提供安全稳定的运行环境，为人工智能模型开发提供支撑和保障。 AIPaaS 通过多层次的安全保障和稳定性措施，确保了 AI 模型从训练到部署的整个过程都处在一个安全、可靠的环境中。在安全方面，智能云工程平台全面考虑了数据安全、模型安全、通信安全、平台安全、网络安全和算法安全等多个维度。通过数据加密、访问控制、模型鉴权、算法认证、入侵检测与防御等手段，防止了未经授权的访问和恶意攻击。同时，平台还对 AI 模型进行了严格的安全防护，防止了模型被篡改或滥用。这些安全措施共同构成了一个全方位的安全体系，为 AI 模型的开发和应用提供了坚实的保障。在稳定性方面，面对大规模集群训练中断频繁，恢复时间长，重复训练等困境，智能云工程平台通过容灾建设、可观测性和混沌工程等手段，确保了平台的稳定性和可靠性。平台具备强大的容错和恢复能力，能够在发生故障或异常情况时迅速恢复服务，保障业务的连续性。同时，平台提供了全面的监控和日志记录功能，使管理员能够实时了解平台的运行状态和性能表现，及时发现潜在问题并进行处理。此外，平台还通过混沌工程等手段模拟各种故障和异常情况，

测试平台的容错能力和恢复能力，进一步提高平台的稳定性和可靠性。

（三）MaaS 定义云上服务新范式，赋能 AI 创新与效率提升

MaaS 将复杂的 AI 模型应用开发能力进行标准化封装，为用户提供了一个便捷化的模型服务平台。在传统模式下，大模型的训练、推理和部署通常需要庞大的底层算力、丰富的训练数据以及先进的模型算法支撑。此外，数据管理、模型安全、应用编排、性能监控等环节也需要大数据、云原生、网络安全等复杂技术的支撑。这些因素不仅增加了企业的技术投入和运营成本，也限制了 AI 技术的广泛应用。然而，MaaS 的出现彻底改变了这一局面，它在 AIPaaS 基础上通过标准化封装，深度整合计算资源、模型算法资源以及数据管理、模型安全、应用编排、性能监控等能力，形成了面向应用开发的覆盖训练、推理到部署的全生命周期模型服务方案，用户只需通过 API 和 SDK 即可进行模型服务的调用和使用。这种一站式的服务模式为用户提供了便捷化的模型服务体验，极大地降低了用户管理模型和开发应用的门槛，使得 AI 技术成为每个企业都可以轻松利用的工具。

MaaS 提供通用模型管理和垂域模型训练服务，助力企业推动 AI 产品创新。在 MaaS 平台中，常见的通用大模型和 AI 算法可以被接入纳管，面向企业不同的 AI 大模型使用需求和垂域模型的训练需求，可以通过预设的模型和算法智能调整参数、优化模型性能、自动部署模

型到生产环境、实时监控模型运行状态和更新微调模型，使得用户无需过多关注包括 AIPaaS 层在内的底层大模型训练、部署等技术细节，只需专注于业务核心逻辑和应用场景，大大提高了垂域模型训练的效率。为实现通用大模型便捷使用和垂域模型的高效训练，国内外云厂商纷纷推出 MaaS 模型管理平台，例如 Azure 的 Azure AI Studio, Google 的 Vertex AI Studio 以及阿里云的百炼大模型服务平台、百度智能云的千帆大模型平台，使得用户可根据需要灵活构建和管理模型，提高 AI 产品的上线速度，更好地满足市场需求。

MaaS 提供基于模型的全流程应用开发工具链，助推 AI 应用价值落地。AI 应用开发涉及了数据处理、AI 模型、分布式计算、部署推理等一系列复杂的技术栈，而 MaaS 提供了全流程的 AI 应用开发能力，覆盖数据管理、模型训练、模型微调、编排部署等，帮助下游用户更加高效、便捷地实现大模型的应用开发。为实现 AI 应用价值落地，国内外各大厂商都推出了全流程应用开发的 MaaS 平台，包括亚马逊云的 Amazon Bedrock、腾讯云 TI 平台、华为云 ModelArts 等。这些平台覆盖了数据管理、模型训练、模型调优、模型评估、推理部署、提示工程、测试编排等完整的 AI 应用开发工具链，并逐步在政府、金融、制造、能源、医疗、零售等多个行业中进行应用开发，助力 AI 应用价值落地。

MaaS 提供丰富的基础 AI 云服务，赋予了企业定制开发 AI 应用的能力，支撑各行业生产力变革。随着 ChatGPT、Gemini 等 AI 大模型

的涌现,人工智能 2.0 时代已经来临,各行业都在积极探索更加智能化、自动化的 AI 大模型应用落地。在政府、金融、制造等行业中,智能办公、文档写作等应用需求日益增长,而智能交互、应用商店则成为零售、电商、文娱等行业中的关键需求。针对这些 AI 应用需求, MaaS 提供了丰富的基础 AI 云服务,赋予了企业定制开发 AI 应用的能力,企业只需要选择需要的 AI 应用类型和功能、上传企业专有训练数据,即可获得企业专属的 AI 应用,进一步降低了企业开发专属 AI 应用的难度和工作量。国内外厂商已推出多种 MaaS 基础 AI 云服务,例如亚马逊云的 Amazon Transcribe、微软云的 Copilot 代码助手、谷歌云的 Document AI 等智能办公产品,阿里云的 AI 数字人、华为云的智能客服等智能交互产品。此外,腾讯云、百度智能云、讯飞星火等服务商也相继推出大模型应用商店,为企业业务拓展和 AI 创新需求提供了更多可能性,进而实现企业数字化转型和行业生产力变革。

（四）AISaaS 推动 AI 应用落地实践，加速企业数字化转型进程

AISaaS 推动 AI 技术深入行业企业,助力用户快速应用 AI 服务。SaaS 是云技术的综合体现和应用形式,通过结合云与 AI 技术,面向企业与各行业需求开发,包装形成开箱即用的 AISaaS,为用户提供便捷的 AI 服务入口。SaaS 与 AI 的深度融合不仅加速 AI 本身的创新和训练过程,而且极大地拓展了 AI 技术的应用边界。一是 SaaS 模式助力 AI 理解业务细节。AI 技术实现大范围应用,需要深入了

解行业业务的运作机制、业务流程和痛点难点。与深入各个业务领域的 SaaS 紧密合作，有助于 AI 能够深度融入业务需要，解决实际问题。二是 SaaS 模式推动用户快速接入 AI。SaaS 即开即用的业务模式帮助快速部署落地生产级的可用 AI 能力，避免为搭建 AI 基础设施和组建研发团队投入大量成本，并且 SaaS 模式按需使用、按量付费的收费方式也降低了用户的智能化成本。三是 SaaS 模式提供高质量业务数据提升模型质量。SaaS 服务商通过积累的大量业务数据训练 AI，使其提供的人工智能服务能够根据用户需求和市场趋势持续迭代更新，使用户在没有充足数据训练的情况下，也能享受经过完善训练后的 SaaS 化智能服务。

AISaaS 在通用业务场景中展现广泛应用。在市场业务方面，通过 AI 增强的客户关系管理平台分析客户行为数据，预测客户需求，为用户定制个性化营销策略；通过分析销售数据识别市场趋势，针对性配置销售团队资源投入，分析客户质量，制定对应销售策略，提高转化率；通过 AI 驱动的市场营销自动化 SaaS 划分消费者兴趣点，执行定制化营销活动，提高营销行为的投资回报率。在公司管理方面，人力资源管理 SaaS 通过结合 AI 技术辅助建立简历筛选、智能匹配职位，预测员工离职风险，以及个性化培训建议等功能，提升企业人力部门的工作效率和效果；财税管理 SaaS 结合 AI 技术进行自动化账务处理，发票识别、财务预测、风险评估等重复性财会工作，减少人工错误，进而提高财务部门的准确性和速度。在协

同办公方面，企业员工通过 AIGC 增强的办公系统提高工作效率，如智能会议纪要、智能助手、文本生成，自动化审批等；项目管理者通过 AI 分析进行资源分配、工作量预测和进度跟踪，提高项目管理的效率和精确度，确保按时完成项目目标。

AISaaS 赋能垂直行业快速转型，适应市场变化。面向垂直行业的 SaaS 服务可以对特定行业的需求提供高度定制化的解决方案，结合大模型等 AI 技术后，能够更好地理解行业特有的数据特征和业务需求，从而提供更加精准的服务。**农业领域**，SaaS 化物联网设备管理与 AI 算法分析，极大提升了作物种植、畜牧和水产等子行业的种植和养殖效率，推动农业管理的现代化和精细化。如某农业数字化 SaaS 平台整合物联网设备收集的农田环境和作物生长等数据，结合 AI 算法进行分析，为农业企业提供作物管理建议、病虫害预警和精准施肥指导等服务。**医疗领域**，医疗领域大模型辅助医生在复杂诊断任务中作出更准确的判断，当前已有多家医疗行业 SaaS 服务商提供医疗影像 SaaS 服务用于存储、统计和标注医学影像数据，在接入影像分析专项的 AI 后，医学影像 SaaS 的分析能力得到显著增强，有效解决基层医疗机构的服务能力不足，提高患者就医体验。**工业领域**，将生产、运维、质量监测等 SaaS 接入大模型，对生产过程中的各种参数进行精细调整以提高生产效率，同时结合计算机视觉自动检测产品缺陷，提高产品质量控制的准确性和效率。如某云服务商开发了针对制造质检的 SaaS 方案，对生产过程中的产品进行拍照

分析，检测产品质量，如尺寸、缺陷等，实现对传统人工质检的替代。**零售领域**，通过结合智能分析的供应链管理 SaaS，预测市场需求变动，优化采购计划，实现供应链的快速响应和灵活调整，降低物流成本，提升整体供应链效率。如某大型零售公司的补货系统，通过分析销售数据、节假日、天气预测等因素调整订单内容和频率，确保库存在客户满意度和周转率之间实现平衡。

五、建议与展望

（一）发展建议

云计算是信息技术发展和服务模式创新的集中体现，是信息化发展的重大变革和必然趋势，是信息时代国际竞争的制高点和经济发展新动能的助燃剂。做大做强我国云计算产业，加快推动云计算的普及应用，提升核心技术能力，有助于加快形成新质生产力。

资源供给方面。在云资源供给基础上，各地应加强与技术突破形成合力，多层面均衡基础资源供需关系。**一方面**，增加智算中心、超算中心等云基础设施节点建设，同时降低现有方案损耗，优化现有算力和云存储资源供给效力。**另一方面**，从技术侧突破跨地域跨形态资源调度难关，打破资源“混而不合”的痒点，形成各区域、各省份间资源互联互通网，完成分级分层的资源调度方案和供给策略制定，促进地区数字经济快速发展。

服务水平方面。**一方面**，持续提高云应用与服务的易用性和丰富度，深挖制造、交通、医疗、农业等传统行业需求，针对性升级

迭代云上应用和服务，加速千行百业深度上云。另一方面，引导云计算与人工智能等新一代信息技术加速融合、积极创新，推动云应用与服务在弹性、便捷、普惠的基础上，进一步向高效化、个性化、智能化演进，提高用户使用体验。

技术融合方面。一是推动算力建设起步晚、云化程度较低的省份深化云技术融合，加大技术、资金和人才的投入，引导企业发挥好云计算弹性按需的能力特点，利用云技术提高已建成算力资源的利用率，充分发挥算云融合赋能企业数智转型的作用；二是提高各场景的数智云化水平，尤其是云化部署比例较低的科学计算等场景，应针对性地加快向智算云服务转型，充分利用云计算资源弹性调度的优势，纳管 CPU、GPU 等多种异构芯片，形成千卡、万卡规模化集群，支撑人工智能应用场景对智能算力的需求，赋能千行百业。

（二）发展展望

随着数字中国建设持续深化，产业变革不断加速，云计算作为人工智能、大数据等数字技术产业底层助力的重要性将日益凸显，正在成为持续推进新一代科技革命和产业变革的关键技术引擎，加速新质生产力发展，构筑国家竞争新优势。未来数年，云计算在提升人工智能算力、算效等方面的价值将被接续发掘，并深度影响社会发展、产业变革、人类生产生活。聚焦云计算产业，将呈现如下趋势：

战略上，全球主要国家本土云计算战略仍将继续升级，全球云

计算开放度或将收紧。随着人工智能算力成为全球性紧缺战略资源，云计算的算、调一体属性将天然使其成为解决算力供应问题的最优方式，从而决定人工智能竞争胜负的关键。美欧等国一方面将持续提升云计算在本土领域的战略价值，推升云计算对其政务、工业、制造等行业的价值供给，另一方面出于国家安全和战略优势考虑，将不断收紧对外国厂商和外国用户的政策包容度，这一趋势预计在当前国际形势下会得以延续。我国将持续聚焦云计算赋能价值，助力企业数字化智能化发展。

路径上，云计算产业关注点不断上移，将更加侧重提升易用、安全、稳定、优化等精细化管理水平。经过十余年发展，云计算产业发展已至成熟阶段，主要表现为市场增速稳定、核心技术成熟、竞争格局稳定、产业链条完整。云计算底层技术已基本完善，产业发展特征推动其发展重点上移，应用现代化、稳定性保障、云原生安全、云成本优化、行业云平台等云计算“工程实施”层面的内容将愈发成为今后发展的核心关注点。**模式上，“云+AI”服务模式创新发展，将开启云计算产业智能化新纪元。**技术侧，云算加速融合，算网云调度操作系统推动算力、网络、云计算协同发展，加速高效互联的算力互联网体系构建，算力标识、高性能传输协议 RDMA 等核心技术将成为创新发展方向。**服务侧，人工智能技术演进正加速算力结构变革，智能算力成为未来算力主要竞争点，“计算能力+AI服务能力”的智能云计算将成为关键，智能云服务技术和应用发展**

成为趋势。应用侧，MaaS、AISaaS 以标准化、便捷化等为特征的云计算基础设施封装 AI 能力，将极大地推进人工智能大模式的落地应用，成为未来数年的发展重点。

“九万里风鹏正举，五千年云鹤长鸣！”中国云计算人将接续奋斗，砥砺前行，以推动云计算产业融合升级，助力新质生产力与新型工业化发展，为中国式现代化和中华民族伟大复兴贡献创新力量！



附件一 数据来源

本报告从我国云资源供需、云服务应用、云技术融合三大层面，对云计算综合发展水平进行量化评价。本报告除明确时间数据外，其他数据统计截止时间为 2023 年底，各维度的数据来源于工信部、国家统计局、中国信通院、各省份统计部门及公开数据整理。

云资源供需发展水平方面。供云量包含的计算、存储、网络硬件规模为云服务商各可用区的 CPU 核数、硬盘容量以及网络出口带宽，超分能力为厂商虚拟化能力，资源价值为服务商对云资源的公开定价，数据来源于工信部统计数据、云服务商调研访谈以及互联网公开数据；用云量包含的各省用云规模为该企业 IaaS、PaaS、SaaS 用云规模，数据来自各省份统计年鉴、统计公报、经济普查、政策采购数据、招投标数据等，企业上云数量来源于工信部统计数据、政府网站公开数据、中国信通院整理。

云服务应用发展水平方面。行业云应用水平评价包括近五年该行业云计算市场规模以及各行业用云情况，数据来源于工信部统计数据、中国信通院统计数据、行业机构公开报告以及互联网公开数据等；企业上云发展水平评价包括企业信息系统上云率等数据，数据来源于工信部统计数据、企业年报、行业报告以及中国信通院调研数据等；云服务体验水平评价包括云服务可用性监测数据、访问时延监测数据、各地方网络质量数据以及云服务商客户服务能力情况数据，数据来源于工信部监测统计数据、中国信通院监测数据、

云服务商披露数据等。

云技术融合发展水平方面。算云融合水平层面，各区域算力规模、算力云化比例展水平及云化资源利用率来源于工信部、国家统计局及各省份统计部门相关数据，云化算力使用体验来源于中国信通院统计整理。数智云化水平层面，数据服务云化部署比例为 PolarDB、Snowflake 等数据库服务的云化部署规模与其总部署规模的比值，数据来源于企业披露信息、网上公开资料和中国信通院整理；智能计算服务云化部署比例为 AI 推理、训练等智能计算服务的云化部署规模与其总部署规模的比值，数据来源于工信部统计和中国信通院整理。

附件二 计算方法

计算方法：**指标数据的无量纲化、标准化处理**，采用极差标准化法，参考每项指标的最大值、最小值，对各指标的观察值进行无量纲化处理，消除量纲、变量和指标大小的影响。**确定指标权重**，基于 AHP 层次分析法，结合定量和定性分析，通过专家经验判断三级指标的相对重要程度，形成三级指标的相对权重。**计算评价结果**，根据三级指标的无量纲标准化观察值和对应的权重，形成各个二级指标的评价结果。

计算结果说明：本计算方法中得到的评价数值范围为 0-100，得分越高，说明该地区/行业/领域对应的指标能力越强。

附件三 计算口径

表 1 指标体系与计算口径

一级指标	二级指标	三级指标	计算口径
云资源供需发展水平	供云量	计算硬件规模	各可用区物理 CPU 总核数、内存容量、GPU 核数以及显存容量
		存储硬件规模	各可用区在用硬盘容量
		网络规模	各可用区网络带宽
		超分能力	各云服务商的 CPU、内存、存储、网络、GPU 等物理资源的虚拟化能力
	云资源价值	各云服务商对云资源的定价	
	用云量	基础设施用云规模	各省份企业在公有云、私有云、混合云、虚拟化设施等方面的支出
		平台建设用云规模	各省份企业在云数据库、云平台、云中间件、云管理等 PaaS 方面的支出
		应用服务用云规模	各省份企业在云 SaaS 服务、云软件等 SaaS 方面的支出
企业上云用云数量		各省份企业使用云服务的数量	
云服务应用发展水平	行业应用发展水平	行业应用范围	根据当前年度各行业行业云规模量选取统计范围
		时间切片	统计数据的时间范围
		行业热力	时间切片内各行业云市场规模
		行业用云占比	当前年度各行业行业云规模量
	企业上云发展水平	信息系统上云率	企业核心以及支撑业务系统的上云情况
	云服务体验水平	可用性	云服务所在可用区的月可用性
		时延	云服务所在可用区的访问响应时延
		网络质量	云服务所在可用区的网络访问质量情况
客户服务能力		云服务商具备的向用户提供的客户支持、服务保障、服务设施等能力	
云技术融合发展水平	算云融合水平	通用算力规模	各地区算力规模= \sum 各地区通用数据中心及私有化部署服务器算力规模
		通用算力云化比例	各地区通用算力云化规模/各地区通用算力规模
		云化资源利用率	各地区云化通用算力的利用情况
	数智云化水平	数据服务云化部署比	PolarDB、Snowflake 等数据库服务的云化

		例	部署规模/总部署规模
		智能计算服务云化部署比例	AI 推理、训练等智能计算服务的云化部署规模/总部署规模

CAICT 中国信通院

中国信息通信研究院

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-62300072

传真：010-62304980

网址：www.caict.ac.cn

