

2025年中国AI算力基础设施发展趋势洞察

易观企业数字化中心

中国AI算力基础设施现状总览 – 国家顶层支持、地方自主建设、智算超前布局



算力基础设施是国家新基建的关键承载

- 算力建设是国家新基建战略的重要组成。2022年“东数西算”工程启动，在京津冀、长三角、粤港澳大湾区、成渝、内蒙古、贵州、甘肃、宁夏8地建设国家算力枢纽节点，规划10大数据中心集群。
- 顶层设计加速，包括《数字中国建设规划》《东数西算实施意见》等政策夯实算力发展的总体格局。



已建成投产算力规模全球第二

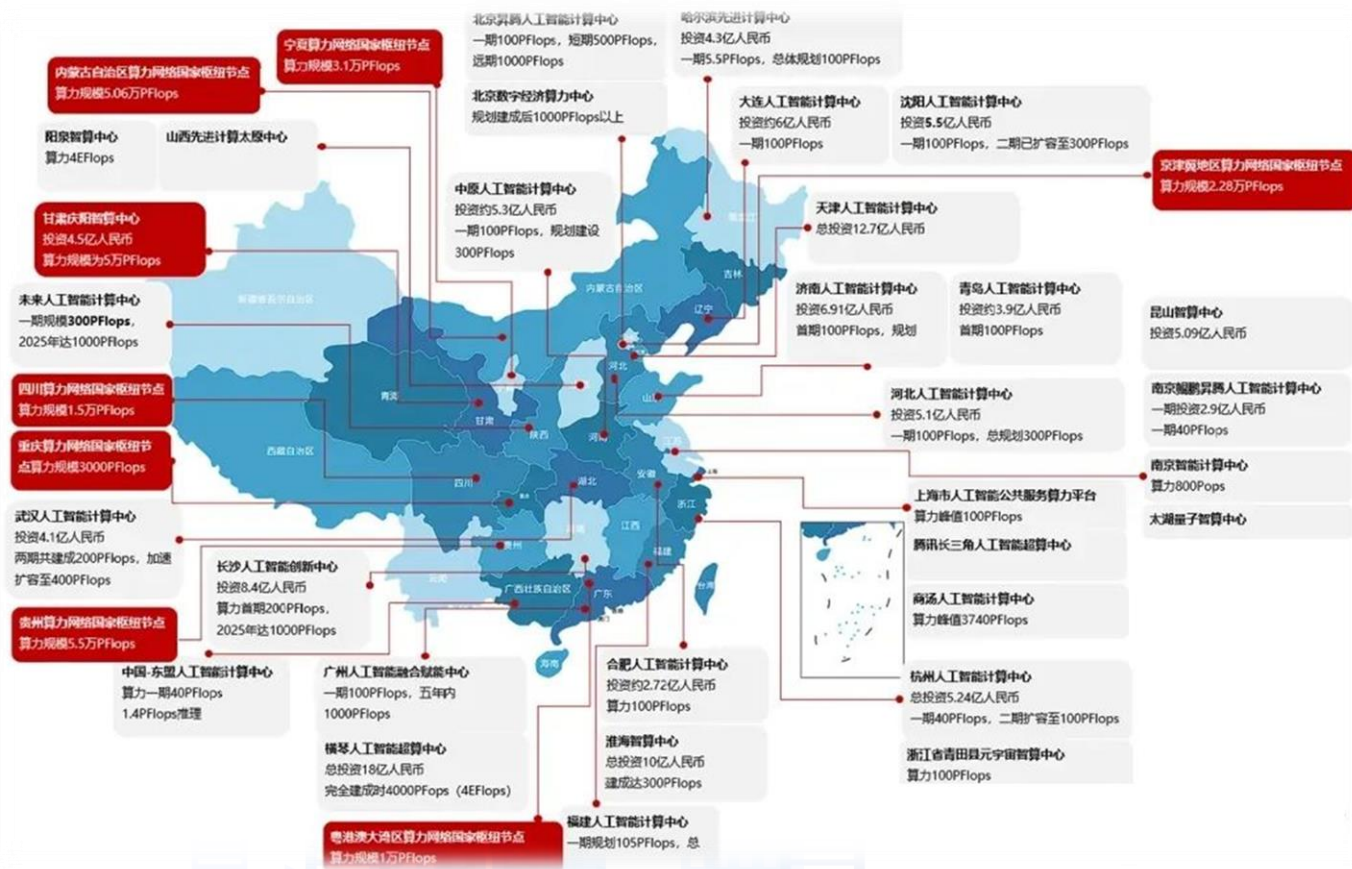
- 截至2024年，全国在用算力中心机架总规模达830万标准机架。**全国算力总规模超过280 EFLOPS**，居全球第二位。
- 智能算力规模达90EFLOPS，占比超过30%，比2019年增长近13倍，年均增速约90%，占全国算力比重接近30%。



东数西算+八大枢纽组成多层算力网络

- “全国一张网”格局加速形成，东部需求旺盛，西部能源充足成本低，形成“核心枢纽引领、区域梯队协同”的多层算力网络。
- 北京、上海、广州、重庆等地建成面向AI的大型智算中心，为本地产业提供算力支撑，全国算力年均增速约30%。

国内主要省市智算中心建设情况



数据来源：公开信息

激发科技与创新活力

中国AI算力基础设施发展环境

— 政策释放、产业支撑、技术趋于成熟，商业闭环形成

政策：算力是AI行业政策红利释放最充分的赛道

- **国家层面政策红利密集释放。**以“东数西算”大战略为脉络，《算力基础设施高质量发展行动计划》等政策夯实顶层设计。
- **地方跟进快速且具体。**北京、上海、浙江、广东等多地发布算力发展行动方案，明确未来几年智算建设目标。
- **着力支持算力高质量创新发展。**包括新型基础设施、绿色节能和自主创新等方向，为产业提供长期稳定导向。

产业：新经济和新质产业亟需夯实智能化底座

- **大模型引爆算力增量需求。**2023年以来生成式AI引爆算力需求，大模型训练与推理带动算力消耗激增。
- **新经济&新质产业是算力需求主力。**智能算力主要应用于互联网行业，占比逾50%，互联网企业大规模投入大模型训练制造、金融、政务等传统行业的智能化升级提速，垂直场景AI应用多点开花。
- **AI价值外溢推动算力需求β升高。**各领域对AI算力的需求正从探索期进入增长期，多元共振的需求结构提供持续动力。

科技：芯片及调度、热利用等技术落地趋于成熟

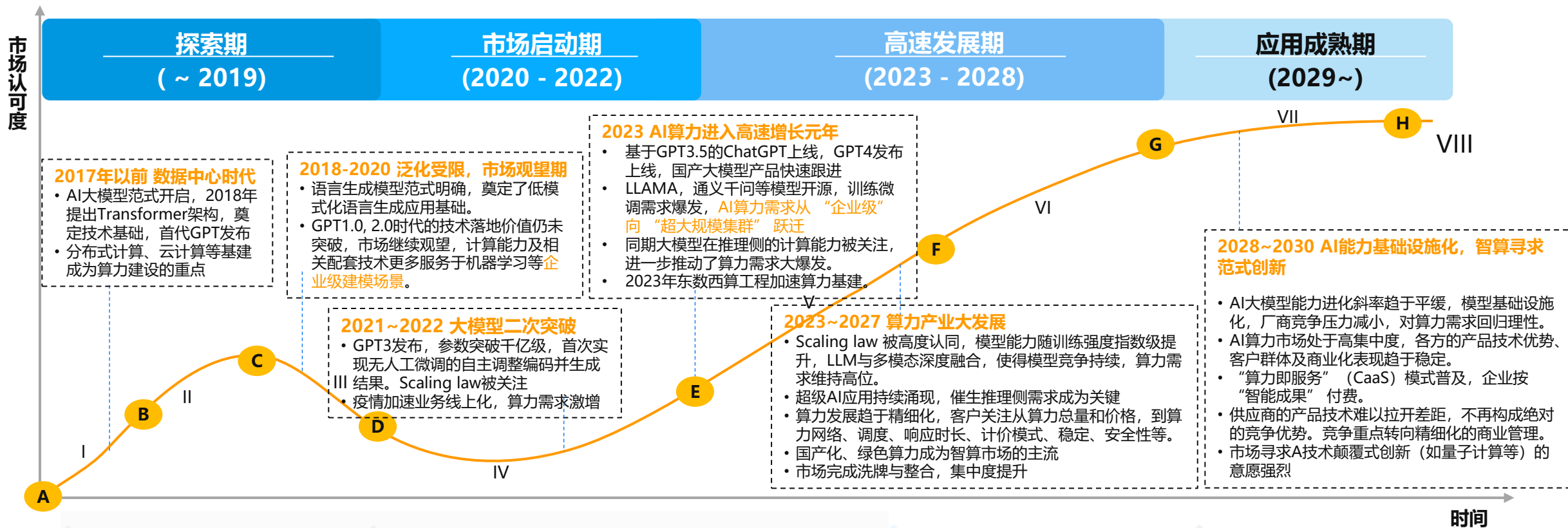
- **AI算力能力跃升。**AI芯片和算力架构创新层出不穷，GPU、NPU等硬件性能不断突破，为智算提供澎湃动力。
- **液冷技术和热利用创新突破。**液冷、浸没式冷却等散热技术逐步应用于数据中心，单机柜功率密度提高的同时能效显著优化。
- **算力调度软硬件协同发展，**高速互联网络 and 软件定义技术使算力资源池化、细粒度调度成为可能。
- **“算力-算法-应用”全栈优化趋势明显。**在软硬件厂商的协同攻关下，算力利用率和AI模型训练效率均得到有效提升。

商业化：供应商已经跑通基于算力的商业模式

- **算力服务市场进入快速成长期，**云服务商、运营商、IDC厂商等纷纷布局智算业务算力供给由过去“重资产投入”转向“平台化服务+生态协同”的运营模式。
- **算力普惠相关新模式推动商业化加速。**按需租赁、算力交易等新商业模式兴起，降低了中小企业获取AI算力的门槛。算力普惠成为大势所趋。
- **产业链分工化发展。**头部算力服务商构筑资源壁垒，中小型服务商侧重特色领域，产业生态走向专业化、集约化经营。

中国AI算力基础设施发展进程 – 处在高速发展阶段的前中期

中国AI算力基础设施发展成熟度曲线（基于易观AMC模型）



AI算力基础设施发展核心驱动因素 – 五大维度动因共振

大模型迭代（参数规模迈向万亿）

大模型参数量级从十亿级跃升到千亿、万亿级，训练需要海量算力支撑。中国大模型数量占全球36%，仅次于美国。模型规模与能力增长带来对算力的指数级需求增长，是智算扩容的第一驱动力。

政策与资本联动

政策支持与资本投入为算力产业提供持续驱动力。“东数西算”等工程撬动各方投资，仅八大算力枢纽相关项目就带动超2000亿元投资。以国资为核心的风险投资和产业基金聚焦AI芯片、算力服务等赛道，推动产业链上下游协同发展。

产业应用规模化

AI应用正从互联网向金融、制造、医疗等实体产业渗透。各行业的专用AI模型和智能应用不断落地，带动算力供给从满足单一领域向多行业拓展。工业、能源等场景化智算需求增势明显，成为算力增长新动能。

长尾算力潜力释放

部分区域算力利用率不高的问题催生了共享和调度新模式。分布式算力调度技术日益成熟，将原本孤立的闲置GPU/TPU算力通过云平台租赁给有需求用户，实现“云上算力池”模式，提高设备利用率，满足中小企业按需需求。

云化调度技术推动

容器化、弹性伸缩、分布式并行等技术在AI算力领域应用，使“算力像水电一样即取即用”成为可能。调度软件可实时在不同节点间分配负载，标准化计量计费体系逐步建立，降低了使用门槛，扩大了潜在用户群体。



以上海为例，“AI+产业”先进地区已经打造出AI算力集群和生态优势

上海AI算力建设主要指标处于国内领先地位

根据《中国综合算力指数》，上海在算力发展水平方面位居全国领先行列。截至2024年，上海全市算力总规模估计已达数十EFLOPS量级，智能算力占比持续提升，人工智能算力方面的投入强度、算力密度等关键指标均处于国内第一梯队。

AI智算政策体系超前布局，对智算云出台专项规划

上海率先推出智算云产业专项政策。2025年3月发布《上海市促进智算云产业创新发展实施意见（2025-2027年）》，设定目标：到2027年，全市智算产业规模力争突破2000亿元，智算算力规模达到200 EFLOPS，自主可控算力占比超70%。同时打造若干综合型智算云平台 and 一批垂直型平台，培育1-2家战略型龙头企业和20家高成长企业。

智算中心集群建设起步早，进展快，规格高

已建成和在建多座大型智算中心，形成区域算力集群优势。在浦东张江、临港新片区等地，先后布局人工智能公共算力服务平台和超大规模智算中心。临港智算园区在2024年启动建设15,000卡GPU的智能算力中心，包含国内规模最大的单池万卡液冷算力集群，采用了国产高性能AI芯片服务器，为上海提供顶尖AI算力供给能力。

围绕AI+算力搭建了以数据为核心的配套生态链

上海聚拢了完备的AI算力产业生态。拥有国家级人工智能实验室等科研力量，孕育了“1+3”基础大模型体系（“书生·浦语”等）。成立全国首家人工智能语料公司“库帕思”，专注于高质量训练数据供给。设立总规模近千亿元的多支人工智能产业基金，聚集了商汤、华为、阿里等头部企业，形成了模型、数据、算力“三位一体”的完整生态链条。

上海大型智算中心建设情况总览

上海典型智算中心

- 浦东张江人工智能公共算力服务平台
- 临港新片区15,000卡GPU智算中心
- 腾讯、阿里云、商汤共建的AIDC

上海智算2027目标

- 产业规模: 2000亿元
- 算力规模: 200 EFLOPS
- 自主可控比例: >70%



上海智算生态示例

基础模型

“书生·浦语”等“1+3”大模型体系

数据支持

“库帕思”等AI语料专业服务

资本支持

千亿级AI产业基金群

企业集聚

商汤、华为、阿里等头部企业

2025年中国AI算力基础设施六大关键趋势

供给侧 – 算力基座升级



①自主可控算力

国产芯片和软硬件生态建设加速，算力自主权建立



②绿色算力

强调能效和清洁能源，液冷技术加速落地，推进算力碳中和



③跨区域算力互联

打造全国算力网络，实现异地算力协同和调度

需求侧 – 服务模式革新



④智算云平台"双轨制"

通用型和垂直行业型AI算力平台并进，满足多元用户需求



⑤LLM+知识库激发算力新需求

大模型与领域知识结合，催生垂直领域的新型算力需求



⑥算力云化加速，普惠发展

算力以云服务形式供给的模式全面普及，算力租赁市场高速增长

趋势一：自主可控算力

🏰 国产算力战略升级

在地缘政治和安全背景下，中国将算力自主可控上升到前所未有的高度。关键核心芯片和算力基础设施去“卡脖子”化成为战略目标。基于国家《算力基础设施高质量发展行动计划》，先进地区设立明确的国产化指标，比如目标2027年智算规模中≥70%由国产算力提供，力求摆脱对进口高端GPU/CPU的依赖。

🧠 AI芯片国产化提速，智算云全面适配

本土厂商推出高性能AI芯片，性能逐步逼近国际领先水平。比如华为昇腾系列芯片已在鹏城云脑等超算中规模部署，寒武纪推出思元系列AI芯片，壁仞科技的BR100通用GPU部分实现对A100级别芯片的国产替代。同时，以九章智算云为代表的新一代云智算平台已经实现了对国产化主流GPU/NPU芯片的充分适配。

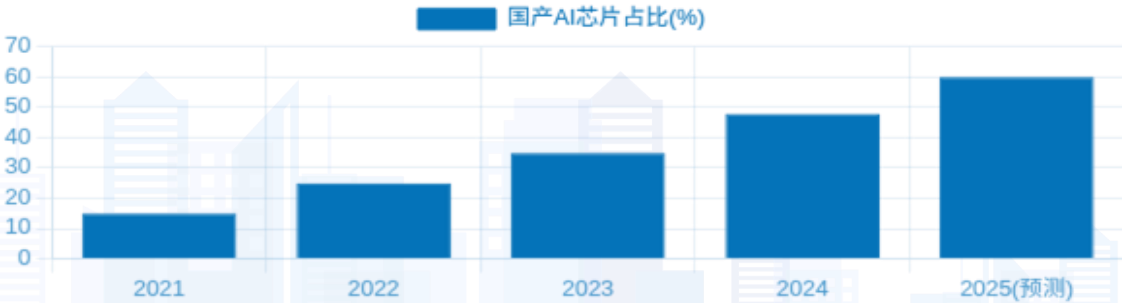
🔧 软硬件生态自主可控

自主AI软硬件生态正在构建完善。国内服务器厂商研发的AI整机柜系统在智算中心大量投入使用，操作系统、中间件、安全方案等均在推进国产化替代，实现从芯片、服务器到云平台的全栈自主。

📍 先进地区引领国产智算集群建设（以上海为例）

先行地区已率先开展大规模国产算力示范应用。比如上海临港建设万卡规模的国产AI算力池，采用国产GPU和液冷技术，建成后将成为全国最大的国产智算集群之一。这类项目为国产算力设备提供宝贵实践经验，也为其他地区树立标杆。

自主可控算力架构与代表厂商



趋势二：绿色算力



双碳约束下，低碳算力成新建项目标配

"碳达峰、碳中和"目标倒逼算力设施走向绿色低碳。工信部明确要求新建大型数据中心PUE应降至1.3以下，业内超大型智算中心更以 $PUE \leq 1.2$ 为设计标准。各地审批新算力项目时将能效和碳排放作为重要考核指标，绿色算力已从倡导走向硬性指标约束。



先进液冷散热与节能技术投产

为降低PUE，数据中心纷纷采用液冷、浸没式冷却等先进散热技术替代传统风冷。液冷技术可大幅提升散热效率，使高功率AI训练服务器保持低温。采用液冷方案的大型AI算力中心平均PUE可低至1.2左右，相比传统机房的1.5以上大幅改善。同时，新一代AI芯片在性能提升的同时单位能耗持续下降。



清洁能源与能源管理

越来越多算力中心接入清洁能源供电以降低碳足迹。西部枢纽引入水电、风电、光伏等可再生能源；东部通过绿电交易购买外送清洁电力。上海等地推进算力中心与电网协调，利用峰谷电价和储能设施优化用电。部分智算园区已实现50%以上用电来自清洁能源。



绿色标准体系

国家正在推进制定绿色算力评估标准和认证体系，包括PUE、碳排放因子、可再生能源利用率等指标。标准体系的完善将指导企业投资更绿色技术，并为政府政策提供依据。绿色算力不仅是环保责任，也通过降低能耗成本提高算力运营的经济性。

绿色算力先进技术路线



液冷技术

冷板直接接触芯片，散热效率提升40%，可支持高密度部署



浸没式冷却

服务器整体浸入绝缘液体，散热均匀，无噪音，节能30%+



高能效比AI芯片

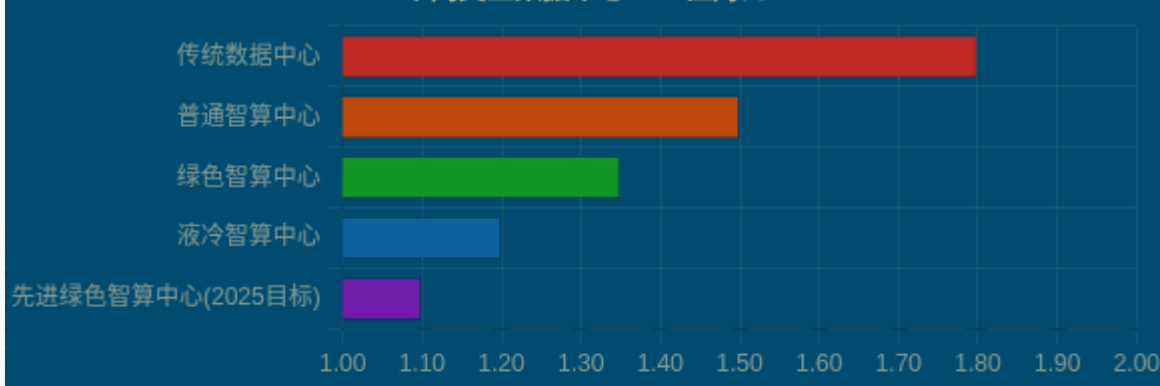
新一代国产AI芯片单位算力能耗较上代降低25%，提升资源利用率



智能调度系统

AI优化任务分配与资源使用，提高设备利用率，减少能源浪费

不同类型数据中心PUE值对比



趋势三：跨区域算力互联

"东数西算"深化发展

- 跨区域算力互联是提升全国算力资源利用效率的必然趋势。"东数西算"工程全面实施后，下一步重点在于枢纽节点互联互通。
- 2025年国家级骨干直联点数量将进一步增加（目前已达26个），骨干网总带宽持续扩展，形成全国范围算力自由流动格局。

统一标准与调度模式

- 跨区互联要发挥效益，需要统一的技术标准和调度体系。国家正推动算力网络标准规范制定和国家、区域算力调度平台建设，包括算力度量标准、接口协议等，确保不同区域的算力设施兼容对接。阿里云、华为云、九章云极等主流的市场化智算云平台厂商也在积极参与，成为国家“东数西算”战略的支撑平台。

东西部算力"外溢"与"回流"

- 跨区算力互联将打通算力供需余缺。西部富余的低成本绿色算力可"外溢"支援东部，夜间则反向利用东部低谷期算力支持西部，实现算力时空互济。如宁夏中卫数据中心的算力已通过高速网络服务北京、上海的AI企业，这种"回流"模式既提高西部收益，又满足东部需求。



统一标准实时监控智能调度跨区互联

趋势四：AI智算云平台"双轨制"

综合型平台崛起

- 由大型云服务商或科技龙头打造，提供大模型开发全链条的一站式服务。集成数据处理、模型训练、微调、推理部署、监控运维等完整功能，平台汇聚主流预训练模型、多模态支持、标准API接口，被视为AI时代的"操作系统"，大幅降低企业AI研发门槛。

垂直型平台并进

- 专注于服务特定行业AI需求，由行业龙头企业与算力服务商联合搭建，提供行业定制化模型即服务（MaaS）。
- 主要特征是内置行业知识库、专用算法模型和开发工具链，针对行业特点优化，显著降低行业用户采用AI的技术门槛。

双轨并行生态

- "综合+垂直"双轨制平台将长期并存、协同发展。综合型平台擅长提供通用基础能力，垂直型平台深入行业场景。
- 两类平台融合的趋势是存在的 - 综合平台输出通用AI能力给垂直平台调用，垂直平台的行业最佳实践反哺综合平台功能优化，多层次满足各类用户需求。



趋势五：大模型+知识库激发算力新需求

大模型走向“有知识”

- 通过引入知识图谱、企业数据库、行业文档等外部知识，大模型能获取事实依据和专业背景，显著提升回答准确性和专业性。这种LLM+知识的结合赋予模型“可解释性”和“可信度”，为大模型应用拓宽了边界，计算需求增加。

通过RAG或预训练微调，实现知识增强

- 知识增强型大模型通常通过RAG(检索增强生成)技术实现：在模型生成回答前，先从知识库检索相关资料，再融合到回答中。
- 随着检索-生成架构日趋成熟，各行业也在构建专属知识库。一些模型在训练阶段就引入知识蒸馏，将知识库内容融入参数，使模型自带领域常识。

泛模型开发者激发算力中长尾需求

- 知识库引入使模型参数量和计算复杂度提高，需要实时检索、匹配、推理，增加计算开销。根据易观测算，企业部署LLM+中大型知识库所需的智算能力相当于基础LLM的2~3倍。
- 各行业知识库训练同样需要大量算力，其中相当一部分算力用户是非专业工程师，对算力的需求是碎片的、基础的、即买即用的，随着垂直行业大模型应用在金融、医疗、制造等行业爆发，“模型+知识”方案将催生对高性能智算中心的新需求。

知识增强型大模型工作原理

- 1 用户查询输入**
“公司新产品上市所需营销预算的汇报审批流程与注意事项是什么？”
- 2 知识库检索**
从企业政策文档、FAQ中检索相关内容
- 3 知识融合与上下文增强**
将检索结果融入大模型输入上下文
- 4 大模型生成回答**
基于检索知识生成准确、可靠的回答

知识库规模与算力需求关系



趋势六：算力云化加速，普惠发展

算力即服务（CaaS）成为算力使用的主流方式

算力云化意味着将计算资源以服务形式在线提供。2024年智能算力服务市场规模增速达到80%，越来越多中小企业和开发者通过云端租赁获得高端算力，而非一次性投入巨资购置设备。各大云厂商也推出AI算力包、按小时GPU租赁等产品，企业可以在高峰时租用高端GPU，低谷时立即释放资源节约开支，削峰填谷。

分散算力在云端整合，标准化计费

算力云化将有效整合原本分散的资源。高校、科研院所的计算节点在闲置期可通过云平台对外提供服务。在这个模式下，算力使用的标准化界定和计费规则将统一，除了国家在推动算力互联规则建设，优势企业也在推动标准化落地实践，九章云极开创性提出的“1度算力”基于对标准化计量计费，确保不同来源算力可以公平交易，提高全社会算力的平均利用率。

算力普惠化

算力集成调度技术的进步，大幅降低了中小企业和个人开发者的获取门槛，进一步促进了算力普惠化发展。比如九章智算云的Serverless架构，支持企业不需自建机房和维护硬件，也不必关注IaaS层的技术实现逻辑，即可获得世界一流算力，引领了算力普惠的范式创新。

云边端协同拓展智算云应用场景

在工业、金融等场景，算力云化也面临多租户安全隔离、网络带宽瓶颈、数据跨地域合规等问题。随着5G、云网融合及边缘计算发展，云化将延伸到网络边缘乃至终端，形成云-边-端一体算力供给体系，解决场景安全问题。

算力云服务模式



高性能GPU云服务器
AI训练与推理专用实例

显著降价↓



按量计费算力
按PFLOPS时/卡时/度等单位计费

弹性扩展



碎片算力整合
闲置资源统一调度

利用率↑

(亿元)

中国智算云服务市场增长趋势(2022-2027)



利益相关方建议 – 明晰定位规划，共建高效、开放、可持续AI智算生态

政府层面

明确定位和规划：加强顶层规划和区域统筹布局，制定区域AI算力网络发展路线图；

制定算力互联的标准体系和测评体系：推进全国算力统筹调度平台建设，对绿色算力、自主可控技术研发给予资金和税收优惠支持

公共数据开放：开放公共场景数据，推动政务、医疗等释放算力需求

企业（算力供需双方）

供给方：提供高品质算力解决方案，建设易用、高性能的算力云平台；在特定行业领域构建数据能力和专业认知，开发定制化AI算力解决方案

需求方：将算力视为关键生产要素纳入战略规划，提前布局；优先利用云上算力进行AI研发，以租赁代替自建降低成本；

园区与枢纽

系统规划：打造智算产业集群，完善生态基础配套设施，避免散点式发展

提供方便可用算力资源：比如部署绿色能源、高速网络直联、可靠制冷创新园区运营模式——算力先租后买、按效果付费等机制

注重产品选型，基于本地算力发展要求，与具备技术领先性的智算云平台开展合作，避免一哄而上导致资源闲置，确保高效利用

生态协同

高校与科研机构：和企业联合攻关算力关键技术，共享成果；推进高校和企业合作培养“智算工程师”等专业化人才。

行业组织：联合芯片、服务器、云平台等企业共同参与算力标准制定产业联盟，避免内卷式竞争。推动形成“芯-云-用”紧密衔接的创新链和价值链

关于易观

易观创立于2000年，是中国数智化领域专业的科技和 market 分析机构，创立之初是全球知名第三方IT分析机构高德纳（Gartner）在华唯一合作伙伴。我们致力于帮助各行业客户和科技厂商在数字化战略、数字化业务以及数字技术应用等方面分析评估现状，明确转型方向，做出正确决策，完成面向数智化时代的转型。经过24年同科技、金融、零售、制造等行业客户的深入互动和持续服务，易观积累了超过8000份独有的分析成果，内容涵盖数字经济和数字技术应用全链条中的业务场景、技术厂商、产品数据，行业案例等、并拥有“易观实力矩阵（Power Matrix）”，“应用成熟度曲线（AMC）”等多个易观自主知识产权的分析模型。



官网网址: www.analysys.cn

客户热线: 4006-010-231

扫描二维码，获取易观最新资讯

易观2025年AI Infra系列研究，敬请关注

- 《2025年中国AI云智算平台市场专题研究》（四大区域篇）
- 《2026年中国AI基础设施发展洞察》
- 《2026年中国产业数字化发展趋势洞察》

欢迎与易观企业数字化中心接洽

张澄宇
(报告主笔)

易观合伙人&企业数字化中心负责人
zhangchengyu@analysys.com.cn
15120031028 (电话&微信)

任 洁

商务负责人
renjie@analysys.com.cn
13681118925 (电话&微信)

相关交流请扫码添加作者

