

分析师：唐月
登记编码：S0730512030001
tangyue@ccnew.com 021-50586737

AI 时代下，算力产业的发展趋势及河南概况

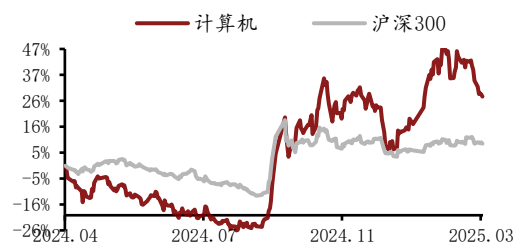
——计算机行业深度分析

证券研究报告-行业深度分析

强于大势(维持)

计算机相对沪深 300 指数表现

发布日期：2025 年 03 月 31 日



资料来源：中原证券，聚源

相关报告

《计算机行业月报：国内算力投入明显加快，平台企业借势积极入局》 2025-03-13

《计算机行业深度分析：三大要素齐发力，AI 应用步入全面加速期》 2025-03-07

《计算机行业月报：DeepSeek-R1 带来 AI 技术突破，持续关注 GPT-5 的推出进程》 2025-02-13

联系人：李智

电话：0371-65585629

地址：郑州郑东新区商务外环路 10 号 18 楼

邮编：上海浦东新区世纪大道 1788 号 T1 座 22 楼

投资要点：

- 算力包括了计算、传输、存储三个方面的能力。算力基础设施是数字基础设施的重要组成部分，是数字经济时代的新型生产力。
- 从国家算力网络总体规划来看，算力产业作为典型的高耗能产业，伴随信息技术的迅猛发展，其总体需求持续快速攀升。然而，我国在算力需求和电力供给方面，存在显著的区域错配问题。为应对这一挑战，随着新基建的推进，全国一体化数据中心建设正式启动。我国适时实施“东数西算”工程，确立了算力网络的国家枢纽节点，旨在引导东部算力需求有序向西部转移。在此基础上，我国进一步提出建设“一体化算力网络”，进一步推动全国范围内算力资源的高效协同。
- 从算力的产业链来看，算力产业主要载体是数据中心，其中服务器作为计算力的承载者，是数据中心的核​​心设备。从服务器的成本构成来看，芯片处于了核心地位，在 AI 大发展的背景下，GPU 又成为芯片需求的主导。
- 2023 年以来随着大模型的快速发展，算力爆发性增长，同时随着 OpenAI 推理模型 o1 的发布和 DeepSeek-R1 的开源，AI 应用加速落地，带来了 Scaling 向后训练和推理阶段转移，继续支撑全球科技厂商的智算投入。AI 服务器、液冷、国产 AI 芯片成为我国未来智算产业的三大核心增长点。
- 从总体发展情况来看，河南目前运载力水平位居全国第 10，相对领先。在算力方面，河南算力位列全国前 10，同时保持了较高的算力利用率。河南形成了“一核四极多点”的算力区域架构，主要数据中心规划高度集中在省会郑州。随着 AI 的大发展，河南省对于算力和数据中心建设热度持续提升，并提出了积极的算力规划，致力于 2025 年算力规模步入全国第一梯队。
- 从河南算力企业布局情况来看，河南目前在上游服务器和硬件端优势最为明显，拥有国产服务器出货量第一的超聚变，紫光建设智慧计算终端的全球总部；豫信电科作为国资方代表全方位推动河南算力产业建设；龙芯中科的入驻，为河南补齐了在芯片端短板；在下游应用端，拥有了 APUS、科大讯飞等厂商加速 AI 应用和落地；同时作为比亚迪未来最大制造基地，也有望在智驾方案全面普及的过程中受益。

风险提示：国际局势的不确定性；下游企业削减开支。

并购家 免费行业报告网

IPOIPO.CN 每日更新 不用注册会员 无广告 永久免费

内容目录

1. 算力的定义	5
2. 国家算力网络的总体规划	6
2.1. 算力需求与电力供给存在区域错配	6
2.2. 新基建的开启和“东数西算”工程的实施	7
2.3. 全国算力资源的高效协同，开启“一体化算力网络”建设	9
3. 算力产业链	12
3.1. 芯片：GPU 成为主导	14
3.1.1. CPU：X86 主导，ARM 等其他架构加速发展	15
3.1.2. GPU：AI 业务带动需求爆发式增长	16
3.2. 服务器：中国形成了独立于全球的竞争格局	18
3.3. 数据中心：运营商占据主导地位，厂商差异明显	19
3.4. 云计算：巨头云集，马太效应明显	21
3.5. 大模型：市场格局还处于高度变化时期，未来产业生态有望进一步集中	22
4. AI 拉动了算力需求的增长	24
4.1. 大模型技术下，算力需求爆发性增长	24
4.2. 推理模型将 Scaling 向后训练和推理阶段转移，应用落地加大推理需求	25
4.3. 截止 2024 年底，我国智能算力建设领先于国家整体规划水平	27
4.4. AI 驱动了算力建设热情，全球科技厂商仍在积极加大对算力的资本投入	28
4.5. 智算增长推动三大细分市场加速发展	31
4.5.1. AI 服务器仍将持续快速增长，ASIC 市场有望迎来更快增长势头	31
4.5.2. 液冷即将迎来规模化应用	32
4.5.3. 国产 AI 芯片迎来持续突破	34
5. 河南积极布局算力产业，力争进入全国第一梯队	35
5.1. 河南运载力较为领先，排名全国第 10	35
5.2. 河南在用算力位列全国前 10，算力利用率较高	36
5.3. 随着 AI 大发展，河南省政策对于算力和数据中心建设热度持续提升	39
5.4. 河南推出积极的算力规划，致力于 2025 年算力规模步入全国第一梯队	44
5.5. 河南的算力区域架构及重点数据中心	45
5.5.1. “一核四极多点”的算力区域架构	45
5.5.2. 河南主要数据中心项目的分布及建设情况	46
5.5.3. 超算领域，河南拥有国家超算互联网核心节点，地位特殊	48
5.6. 河南算力产业核心企业	48
5.6.1. 超聚变：河南算力产业核心企业，国产服务器出货量第一	48
5.6.2. 豫信电科：国资代表，全方位推动河南算力产业建设	49
5.6.3. 龙芯中科：完善了河南芯片板块布局，助力信创发展	50
5.6.4. 紫光股份：计划在河南投资 100 亿，建设智慧计算终端全球总部基地	51
5.6.5. APUS：实现了河南省大模型备案零的突破	52
5.6.6. 科大讯飞：与农大联合发布弘农大模型，加速 AI 大模型在河南的应用	53
5.6.7. 比亚迪：河南建设最大制造基地，全面普及智驾方案有望催生较大的算力需求	53
6. 风险提示	54

图表目录

图 1：2023 年全球算力规模	5
------------------------	---

图 2: 数字中国建设的整体框架与算力基础设施的关系	6
图 3: 2016 年我国数据中心业务的区域分布情况	7
图 4: 2024 年我国十大电力装机省份份额	7
图 5: 8 大算力枢纽节点及 10 个国家数据中心集群的分布情况	8
图 6: 数据中心设备	13
图 7: 算力产业链	14
图 8: 不同类型服务器的成本构成	15
图 9: 2021 年和 2022 年全球数据中心 CPU 市场收入份额	16
图 10: 2021Q1 至 2024Q4 英特尔、AMD、英伟达的数据中心业务收入（亿美元）及增速	16
图 13: 2023 年我国数据中心服务商市场厂商份额	20
图 14: 2023 年我国第三方数据中心服务商市场厂商份额	20
图 29: 2024H1 中国公有云 IaaS 厂商市场份额	22
图 30: 2024H1 中国公有云 PaaS 厂商市场份额	22
图 20: 推理模型与语言模型在数学、科学、编码方面的能力对比	25
图 21: DeepSeek-V3 在模型训练不同阶段的开销	26
图 22: DeepSeek-R1 的训练流程	26
图 23: DeepSeek-R1-Zero 在训练中 AIME 精度提升情况	27
图 24: DeepSeek-R1-Zero 在训练中的平均响应时长	27
图 25: 模型 Scaling 的新范式	27
图 29: 通用算力行业应用情况	28
图 30: 智能算力行业应用情况	28
图 27: 2021Q1-2024Q4 主要芯片厂商数据中心业务收入（亿美元）	29
图 29: 阿里、腾讯、百度的资本开支（亿元）	29
图 30: 5 大科技厂商资本投入（百万美元）	29
图 29: 2018H1-2024H1 我国 AI 服务器市场规模及增速	31
图 30: 2024-2028 中国 AI 服务器市场预测（百万美元）	31
图 31: 液冷技术与数据中心能耗指标关系图	33
图 32: 2022 年我国液冷服务器市场厂商份额	34
图 33: 2023 年我国液冷服务器市场厂商份额	34
图 34: 2022Q1 至 2024Q4 英伟达分区域收入（亿美元）及占比	34
图 35: 2023H1-2024H1 我国 AI 服务器芯片国产化占比	35
图 41: 国家新一代人工智能公共算力开放创新平台的区域分布情况（个）	39
图 42: 各地算力规划中关于总算力的相关目标（EFLOPS）	44
图 43: 各地算力规划中关于智能算力的相关目标（EFLOPS）	44
图 45: 2021-2024 超聚变收入及增速（亿元）	49
图 46: 龙芯中科处理器路线图	51
图 47: 比亚迪天神之眼技术矩阵	54
表 1: 三大算力概念及用途	5
表 2: 东数西算工程的区域划定及区域定位	8
表 3: 国家数据局组建方案	9
表 4: 一体化算力网实施意见中的相关算力建设目标	10
表 5: 近年来算力领域的相关政策	10
表 6: 英伟达、AMD 人工智能芯片参数对比	17
表 7: 2023 年企业 IDC 相关业务对比（亿元）	20

表 8：2024H1 企业 IDC 相关业务对比（亿元）	20
表 9：OpenAI 大模型产品竞品推出情况.....	23
表 10：OpenAI 大模型产品	24
表 11：大模型领域相关算力相关投入.....	30
表 12：运力分指数体系	36
表 13：河南省算力产业相关政策.....	39
表 14：河南省算力相关目标	43

算力包括了计算、传输、存储三个方面的能力。根据《算力基础设施高质量发展行动计划》的定义，算力是集信息计算力、网络运载力、数据存储力于一体的新型生产力，主要通过算力基础设施向社会提供服务。

(1) 计算力 (CP)：是衡量数据中心计算能力的综合指标，包括了通用计算能力、智能计算能力和超级计算能力三个类型。其中通用计算力占比最高，智能计算力作为 AI 发展的重要支撑，占比在快速提升过程中，而超级计算能力占比相对较小。

表 1: 三大算力概念及用途

类型	概念	主要芯片类型	应用场景
通用算力	最基础的计算能力，用于执行常见的计算任务	CPU	日常办公、网页浏览、文件处理等基础计算任务
智能算力	提供高强度的 SI 计算能力，用 AI 服务器提供算力	GPU/ASIC/FPGA、CPU	图像识别、自然语言处理、模型训练推理、自动驾驶等
超算算力	面向计算复杂或数据密集型的科学计算问题，由超级计算机等高性能计算集群提供算力	高性能 CPU、协处理器	科学研究、气象预测、资源勘探、生命科学、航空航天等

资料来源：信通院，中原证券研究所

截止 2023 年底，我国算力位于全球第 2，占比 26%，仅落后于美国（占比 32%），远远领先于其他国家。从算力结构来看，通用算力仍然占比最大。

图 1: 2023 年全球算力规模



资料来源：《中国算力发展报告（2024）》，中原证券研究所

(2) 运载力 (NP)：是衡量网络传输调度能力的综合指标，涉及数据中心内部和数据中心之间的网络传输。

(3) 存储力 (SP): 是衡量数据中心数据存储能力的综合指标, 包括额存储阵列等外置存储设备和服务器内置存储设备。

从数字中国建设的整体框架来看，算力基础设施是数字基础设施（还包括网络基础设施、应用基础设施）的重要组成部分，是数字经济时代的新型生产力。

图 2：数字中国建设的整体框架与算力基础设施的关系



资料来源：《数字中国建设整体布局规划》，中原证券研究所

2. 国家算力网络的总体规划

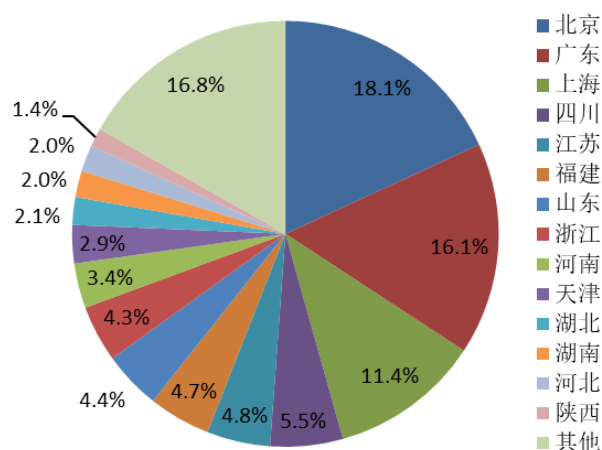
2.1. 算力需求与电力供给存在区域错配

我国面临了来自算力需求和电力供给方面区域错配。

由于东部经济发达省份的人口和政企用户密度都远高于中西部地区，对算力需求也更大，因而也成为了数据中心业务需求最旺盛的区域。根据 IDC 数据，2016 年全国数据中心分布来看，仅北上广三地数据中心占比就达到了 45.6%，数据中心上架率也远高于其他区域。随着北上广对新建数据中心审批趋严，数据中心建设需求逐渐转移到其周边区域。

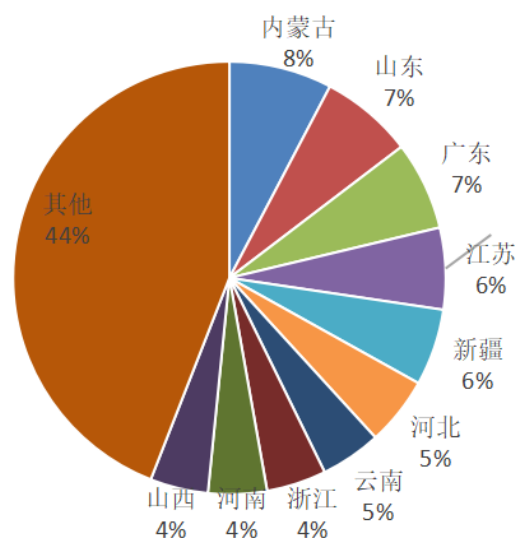
而从电力供应的角度来看，我国煤炭、油气资源主要分布在西部和北部，西高东低的地势导致河流流向自西向东，在这一自然资源禀赋下，我国通过西电东送和北煤南运来弥补资源的错配和用电缺口。根据中电联数据，2024 年全国十大电力装机省份中有 5 个都是中西部区域。

图 3：2016 年我国数据中心业务的区域分布情况



资料来源：IDC 图，中原证券研究所

图 4：2024 年我国十大电力装机省份份额



资料来源：国家能源局，中电联，中国电力知库，风电头条，国际太阳能光伏网，中原证券研究所

算力产业是一个高耗能产业，同时随着信息技术的高速发展，总体需求仍在快速增长。

2021 年，全国数据中心耗电量为 2166 亿千瓦时，占全社会用电量 2.6%，为同期三峡电站累计发电量 1036.49 亿千瓦时的两倍。2022 年全国数据中心耗电量达到 2700 亿千瓦时，同比增长 25%，占全社会用电量约 3%。IDC 预测 2024 年全国数据中心用电量占全社会用电量比重将提升至 5%。

因而，在算力产业快速发展的趋势下，东部电力缺口进一步加大，为后续国家“东数西算”发展规划的出台埋下了伏笔。

2.2. 新基建的开启和“东数西算”工程的实施

以建设数据强国和新基建为契机，全国一体化数据中心建设拉开帷幕。2020 年 3 月，习近平总书记在政治局常务委员会上强调，要“加快 5G 网络、数据中心等新型基础设施建设进度”，将数据中心上升为数字经济时代区域和国家竞争力的核心战略资源。2020 年 5 月，发改委明确将“实施全国一体化大数据中心建设重大工程，布局 10 个左右区域级数据中心集群和智能计算中心”纳入新基建发展年度重点任务。

确立了算力网络的国家枢纽节点，开始实施“东数西算”工程。2020 年 12 月，发改委、中央网信办、工信部、能源局 4 部门联合发布了《关于加快构建全国一体化大数据中心协同创新体系的指导意见》，提出了“京津冀、长三角、粤港澳大湾区、成渝等重点区域，以及部分能源丰富、气候适宜的地区布局大数据中心国家枢纽节点”的布局方案。2021 年 5 月，4 部门又联合发布了《全国一体化大数据中心协同创新体系算力枢纽实施方案》，确定了在京津冀、长三角、粤港澳大湾区、成渝，以及贵州、内蒙古、甘肃、宁夏等地布局建设全国一体化算力网络国家枢纽节点，同时还提出在国家枢纽节点之间进一步打通网络传输通道，加

快实施“东数西算”工程。而后在 2021 年 12 月到 2022 年 2 月，东数西算涉及的 8 个国家算力枢纽节点和 10 个国家数据中心集群全部获批。至此，八个国家枢纽节点全部确立，被认为标志着“东数西算”工程的全面启动

图 5：8 大算力枢纽节点及 10 个国家数据中心集群的分布情况



资料来源：百度百科，中原证券研究所

从东数西算工程的区域划定来看，京津冀、长三角、粤港澳大湾区、成渝代表的是经济发达区域和算力主要需求方，本地优先主要满足实时性算力需求，贵州、内蒙古、甘肃、宁夏能源丰富、气候适宜，将主要承载东部的数据备份和非实时算力的应用算力需求。

表 2：东数西算工程的区域划定及区域定位

八大枢纽节点	10 大集群	起步区边界（21 个片区）	区域定位
京津冀枢纽	张家口集群	河北省张家口市怀来县、张北县、宣化区	围绕数据中心集群，抓紧优化算力布局，积极承接北京等地 实时性 算力需求，引导 温冷业务向西部迁移 ，构建辐射华北、东北乃至全国的 实时性 算力中心。
长三角枢纽	长三角生态绿色一体化发展示范区集群	上海市青浦区、江苏省苏州市吴江区、浙江省嘉兴市嘉善县	围绕两个数据中心集群，抓紧优化算力布局，积极承接长三角中心城市 实时性 算力需求，引导 温冷业务向西部迁移 ，构建长三角地区算力资源“一体协同、辐射全域”的发展格局
	芜湖集群	芜湖市鸠江区、弋江区、无为市	
粤港澳枢纽	韶关集群	韶关高新区	围绕韶关数据中心集群，抓紧优化算力布局，积极承接广州、深圳等地 实时性 算力需求，引导 温冷业务向西部迁移 ，构建辐射华南乃至全国的实时性算力中心
成渝枢纽	天府集群	成都市双流区、郫都区、简阳市	围绕两个数据中心集群，抓紧优化算力布局，平衡好城市与城市周边的算力资源部署，做好与“东数西算”衔接。
	重庆集群	重庆市两江新区水土新城、西部（重庆）科学城璧山片区、重庆经济技术开发区	
内蒙古枢纽	和林格尔集群	和林格尔新区、集宁大数据产业园	充分发挥集群与京津冀毗邻的区位优势，为 京津冀高实时性 算

			力需求提供支援，为长三角等区域提供非实时算力保障
宁夏枢纽	中卫集群	中卫工业园西部云基地	要充分发挥区域可再生能源富集的优势，积极承接东部算力需求，引导数据中心走高效、清洁、集约、循环的绿色发展道路。
甘肃枢纽	庆阳集群	庆阳西峰数据信息产业聚集区	要尊重市场规律、注重发展质量，打造以绿色、集约、安全为特色的数据中心集群，重点服务京津冀、长三角、粤港澳大湾区等区域的算力需求
贵州枢纽	贵安集群	贵安新区贵安电子信息产业园	抓紧优化存量，提升资源利用效率，以支持长三角、粤港澳大湾区等为主，积极承接东部地区算力需求

资料来源：政府网站，中原证券研究所

东数西算工程的实施，带来了显著的成效。根据信通院 2023 年 12 月公众号文章，通过“东数西算”发展战略，我国中西部算力设施占比已经达到 41%，占比持续提升。根据国家数据局披露的数据，截止 2024 年 3 月底，10 个国家数据中心集群算力总规模超过 146 万标准机架，整体上架率为 62.72%，较 2022 年提升 4 个百分点；东西部枢纽节点间网络时延已基本满足 20 毫秒（ms）要求；数据中心绿电占比超过全国平均水平，部分先进数据中心绿电使用率达到 80% 左右，新建数据中心 PUE（电能利用效率）最低降至 1.10。

2.3. 全国算力资源的高效协同，开启“一体化算力网络”建设

2023 年成立独立机构开展国家数据的管理工作。2023 年 3 月，在新一轮的国家机构改革中，在国家发展和改革委员会下组建了国家数据局，并在 2023 年 10 月正式挂牌。配合国家网络强国和数字中国发展战略，国家数据局的成立，加速了数据管理顶层制度设计，也重塑了从中央到地方的数据工作的管理架构。

表 3：国家数据局组建方案

变化	具体职责	职责划出部门
划入职责	拟订数字中国建设方案、协调推动公共服务和社会治理信息化、协调促进智慧城市建设、协调国家重要信息资源开发利用与共享、推动信息资源跨行业跨部门互联互通等	中央网络安全和信息化委员会办公室
	统筹推进数字经济发展、组织实施国家大数据战略、推进数据要素基础制度建设、推进数字基础设施布局建设等	国家发展和改革委员会

资料来源：政府网站，中原证券研究所

在“东数西算”工程的基础上，我国又提出了建设“一体化算力网络”，进一步推动全国范围内算力资源的高效协同。2023 年 12 月，国改委、国家数据局等 5 个部门联合印发了《关于深入实施“东数西算”工程加快构建全国一体化算力网的实施意见》（以下简称“实施意见”），部署了统筹通用算力、智能算力、超级算力协同计算，统筹东中西地区及大中小城市协同布局，统筹算力、数据、算法协同应用，统筹算力和绿色电力协同建设，统筹算力和安全协同保障等五项重点任务。

相对于 2020 年的《关于加快构建全国一体化大数据中心协同创新体系的指导意见》而言，实施意见对 2025 年的发展目标进一步细化到了算力布局、时延网络、绿色低碳、普惠成

本等多个方面。可以看到除了数据中心建设布局以外，一体化算力网络还需要实现“算网”结合和“算能”结合。

表 4：一体化算力网实施意见中的相关算力建设目标

方向	实施意见中 2025 年目标	其他政策中关于 2025 的建设目标
算力布局	国家枢纽节点新增算力占比超 60%，西部枢纽算力使用率显著提升	算力规模超过 300 EFLOPS，智能算力占比达到 35%。整体上上架率不低于 60%，国家枢纽节点算力资源使用率显著超过全国平均水平。 存储总量超过 1800EB，先进存储容量占比达到 30% 以上，重点行业核心数据、重要数据灾备覆盖率达到 100%。
时延网络	1ms 城市网、5ms 区域网、20ms 跨枢纽网	国家枢纽节点数据中心集群间基本实现不高于理论时延 1.5 倍的直连网络传输，重点应用场所光传送网覆盖率达到 80%，骨干网、城域网全面支持 IPv6，SRv6 等创新技术使用占比达到 40%
绿色低碳	算力电力双向协同机制初步形成，枢纽新建数据中心绿电占比超 80%	平均电能利用效率 PUE 降至 1.5 以下（新建及改扩建大型和超大型数据中心电能利用效率降至 1.25 以内，国家枢纽节点数据中心项目电能利用效率不得高于 1.2），可再生能源利用率年均增长 10%
普惠成本	枢纽间网络传输费用 “大幅降低”	

资料来源：政府网站，中原证券研究所

在后续的政策落地方面，2024 年 7 月，发改委、工信部、能源局、国家数据局共同出台了《数据中心绿色低碳发展专项行动计划》，来推动数据中心的绿色低碳发展。2024 年 8 月，发改委、能源局、国家数据局共同出台了《加快构建新型电力系统行动方案（2024～2027 年）》，推动“实施一批算力与电力协同项目”。同时，工信部等 11 个部门联合发布《关于推动新型信息基础设施协调发展有关事项的通知》进一步聚焦骨干网络设施、算力基础设施的统筹规划布局，实现算网、算能协同。

表 5：近年来算力领域的相关政策

时间	部门	政策	简介
2016.10.9	中央政治局第三十六次集体学习		以推行电子政务、建设新型智慧城市等为抓手，以数据集中和共享为途径，建设全国一体化的国家大数据中心，推进技术融合、业务融合、数据融合，实现跨层级、跨地域、跨系统、跨部门、跨业务的协同管理和服务。
2019.10	中国共产党第十九届中央委员会第四次全体会议		正式将数据列为新的生产要素
2020.3.4	中共中央政治局常务委员会		习总书记提出“加快 5G 网络、数据中心等新型基础设施建设进度”
2020.4.10	中共中央、国务院	《关于构建更加完善的要素市场化配置体制机制的意见》	中央第一份关于要素市场化配置的文件。文件中分类提出了土地、劳动力、资本、技术、数据五个要素领域改革的方向，其中“数据”作为一种新型的生产要素首次被列入。
2020.5	发改委	《关于 2019 年国民经济和社会发展计划执行情况	明确将“实施全国一体化大数据中心建设重大工程，布局 10 个左右区域级数据中心集群和智能计算中心”纳入新基建发展年度

		与 2020 年国民经济和社会发展计划草案的报告》	重点任务
2020.12.23	发改委、中央网信办、工信部、能源局	《关于加快构建全国一体化大数据中心协同创新体系的指导意见》	加强全国一体化大数据中心顶层设计。到 2025 年，全国范围内数据中心形成布局合理、绿色集约的基础设施一体化格局。
2021.5.24	发改委、中央网信办、工信部、能源局	《全国一体化大数据中心协同创新体系算力枢纽实施方案》	在京津冀、长三角、粤港澳大湾区、成渝，以及贵州、内蒙古、甘肃、宁夏等地布局建设全国一体化算力网络国家枢纽节点。国家枢纽节点之间进一步打通网络传输通道，加快实施“东数西算”工程，提升跨区域算力调度水平。
2021.7.4	工信部	《新型数据中心发展三年行动计划（2021-2023）》	到 2021 年底，全国数据中心平均利用率力争提升到 55% 以上，总算力超过 120 EFLOPS，新建大型及以上数据中心 PUE 降低到 1.35 以下。 到 2023 年底，全国数据中心机架规模年均增速保持在 20% 左右，平均利用率力争提升到 60% 以上，总算力超过 200 EFLOPS，高性能算力占比达到 10%。国家枢纽节点算力规模占比超过 70%。新建大型及以上数据中心 PUE 降低到 1.3 以下，严寒和寒冷地区力争降低到 1.25 以下。国家枢纽节点内数据中心端到端网络单向时延原则上小于 20 毫秒。
2021.7.30	国务院	《关键信息基础设施安全保护条例》	为加快提升关键信息基础设施安全保护能力，提出关键信息基础设施的具体范围和安全保护办法
2021.11.30	工信部	《“十四五”大数据产业发展规划》	加快构建全国一体化大数据中心体系，推进国家工业互联网大数据中心建设，强化算力统筹智能调度，建设若干国家枢纽节点和大数据中心集群。建设高性能计算集群，合理部署超级计算中心。
2021.12.29	发改委、中央网信办、工信部、能源局	《关于同意宁夏回族自治区启动建设全国一体化算力网络国家枢纽节点的复函》等	批复宁夏（中卫）、甘肃（庆阳）、内蒙古（和林格尔）、贵州（贵安）建设全国一体化算力网络国家枢纽节点
2022.1.12	国务院	《“十四五”数字经济发展规划》	推进云网协同和算网融合发展。加快构建算力、算法、数据、应用资源协同的全国一体化大数据中心体系。
2022.2.16	发改委、中央网信办、工信部、能源局	《关于同意京津冀地区启动建设全国一体化算力网络国家枢纽节点的复函》等	批复京津冀（张家口）、长三角（长三角、芜湖）、成渝（天府、重庆）、粤港澳大湾区（韶关）建设全国一体化算力网络国家枢纽节点
2022.12.19	中共中央、国务院	《关于构建数据基础制度更好发挥数据要素作用的意见》	从数据产权、流通交易、收益分配、安全治理等方面构建数据基础制度，提出 20 条政策举措。
2023.2.27	中共中央、国务院	《数字中国建设整体布局规划》	构建了数字中国建设的“2522”整体框架，提出将数字中国建设工作情况作为对有关党政领导干部考核评价的参考。
2023.3.16	中共中央、国务院	《党和国家机构改革方案》	组建国家数据局
2023.10.8	工信部、中央网信办、教育部、卫健委、央行、国资委	《算力基础设施高质量发展行动计划》	从算力力、运载力、存储力以及应用赋能四个方面提出了到 2025 年发展量化指标。
2023.10.25		国家数据局挂牌	
2023.12.29	发改委、国家数据局、	《关于深入实施“东数西	在产业生态、算力供给、网络传输、业务调度、系统运营、技术

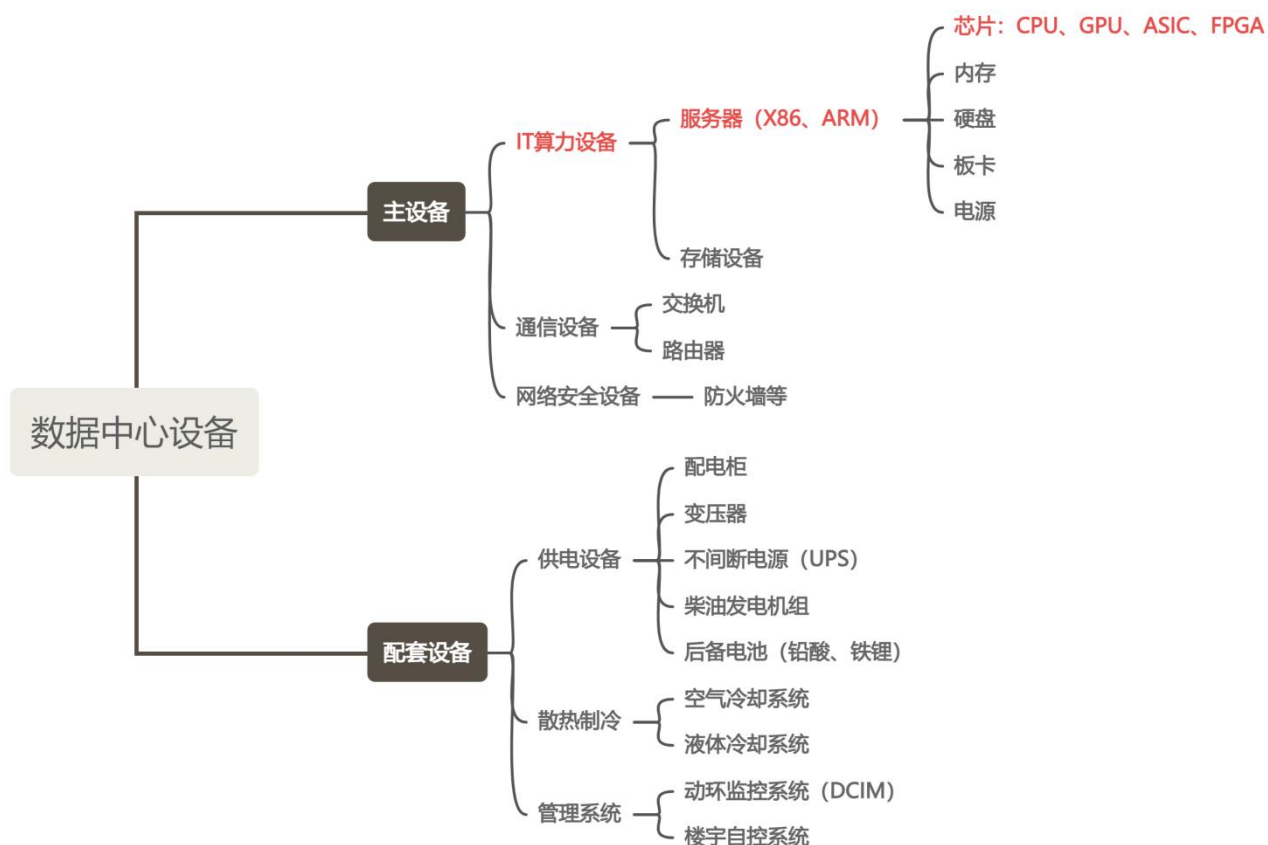
	中央网信办、工信部、能源局	算”工程加快构建全国一体化算力网的实施意见》	创新等方面部署，旨在建设联网调度、普惠易用、绿色安全的全国一体化算力体系。
2024.3.5	十四届全国人大二次会议	政府工作报告	适度超前建设数字基础设施，加快形成全国一体化算力体系，培育算力产业生态。
2024.7.21	二十届三中全会	《中共中央关于进一步全面深化改革 推进中国式现代化的决定》	建设和运营国家数据基础设施，促进数据共享
2024.7.3	发改委、工信部、能源局、国家数据局	《数据中心绿色低碳发展专项行动计划》	到 2025 年底，全国数据中心布局更加合理，整体上架率不低于 60%，平均电能利用效率 PUE 降至 1.5 以下，可再生能源利用率年均增长 10%，平均单位算力能效和碳效显著提高。
2024.8.6	发改委、能源局、国家数据局	《加快构建新型电力系统行动方案（2024~2027 年）》	明确加快推进新型电力系统建设部署，实施一批算力与电力协同项目，进一步完善算力体系建设，统筹数据中心发展需求和新能源资源禀赋，开展算力和电力基础设施协同规划布局。
2024.8.19	工信部、中央网信办、教育部、财政部、自然资源部、住建部、农业部、卫健委、央行、国资委、中国国家铁路集团	《关于推动新型信息基础设施协调发展有关事项的通知》	提出统筹规划骨干网络设施，鼓励网络与算力设施协同发展。引导大型、超大型数据中心、智能计算中心、超算中心在枢纽节点部署。支持算力基础设施与能源、水资源协调发展。支持企业和机构建设面向行业应用的标准化公共数据集。
2024.12.31	发改委、国家数据局、工信部	《国家数据基础设施建设指引》	到 2029 年，基本建成国家数据基础设施主体结构，初步形成横向联通、纵向贯通、协调有力的国家数据基础设施基本格局，构建协同联动、规模流通、高效利用、规范可信的数据流通利用体系，协同构筑数据基础设施技术和产业良好生态，国家数据基础设施建设和运营体制机制基本建立。
2025.2.13			经济观察报报道国资委正在筹备 国家数据集团 这家新央企，其将承担起整合、运营和开发国家级数据资源的重任。

资料来源：政府网站，中原证券研究所

3. 算力产业链

算力产业主要载体是数据中心。在数据中心的设备构成方面，主要设备包括了服务器、存储设备、交换机、路由器、防火墙等，在配套设备方面主要包括了供配电、散热制冷、管理系统三个大类。这其中，**服务器是计算力的主要载体，也是数据中心的核**心设备。

图 6：数据中心设备



资料来源：中原证券研究所

算力产业链上游主要包括芯片、服务器、网络设备、基础软件、网络安全等环节，中游包括数据中心（IDC）、云计算服务，下游包括了大模型、互联网、软件、政企行业及个人客户等环节。

图 7：算力产业链

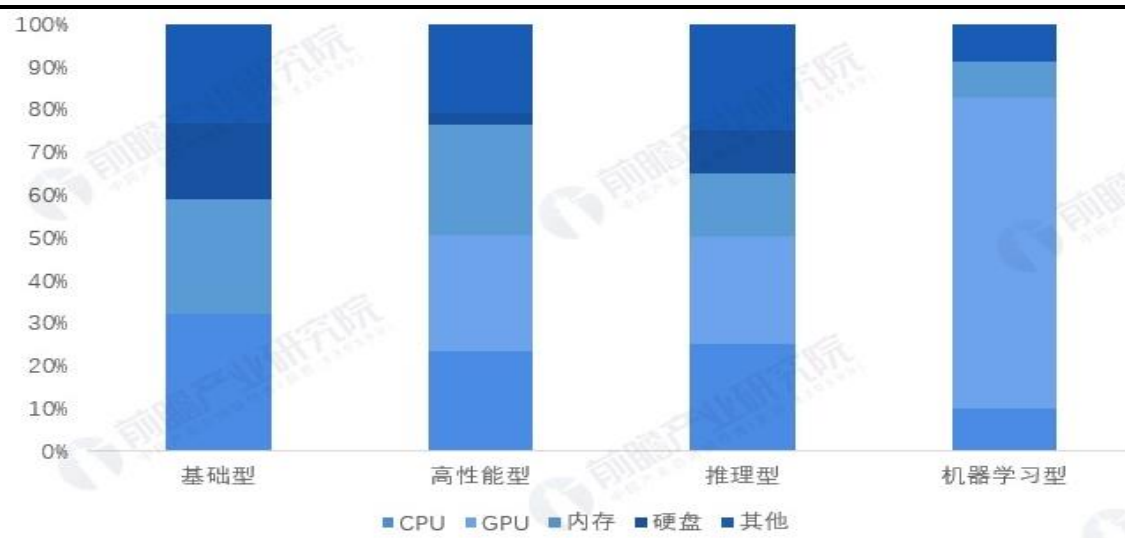


资料来源：中原证券研究所

3.1. 芯片：GPU 成为主导

从服务器的成本构成来看，芯片占据了核心地位。在基础型服务器中，芯片成本占比超过 30%，而在高性能服务器中占比超过 60%。随着 AI 的发展，GPU 在 AI 服务器中占比大幅提升，其中最昂贵的机器学习型服务器中，芯片成本占比超过了 80%。

图 8：不同类型服务器的成本构成



资料来源：前瞻产业研究院，中原证券研究所

3.1.1. CPU：X86 主导，ARM 等其他架构加速发展

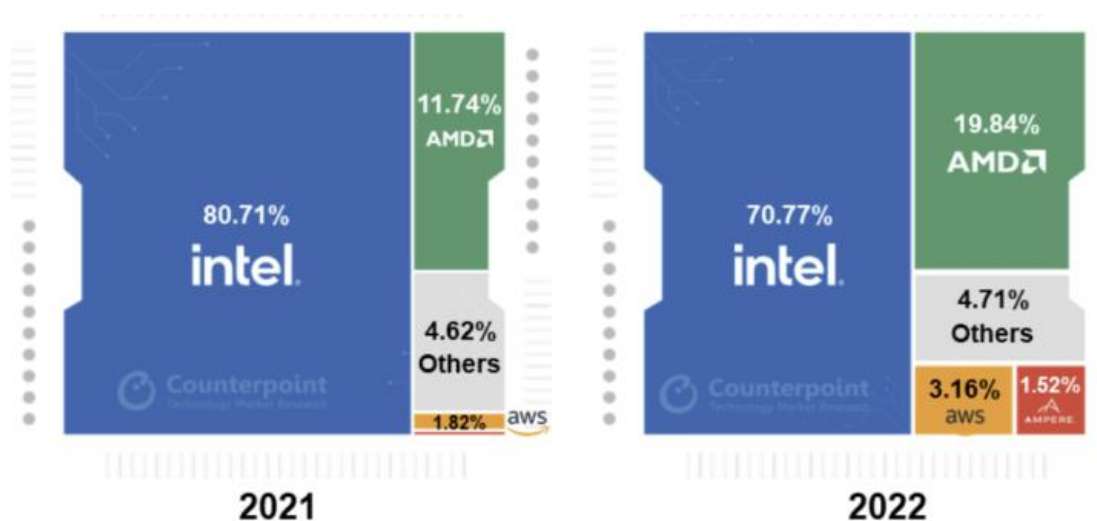
在通用服务器中，CPU（中央处理器，Central Processing Unit）是最核心部件，完成了指令处理、数据运算、控制协调、存储管理等众多功能。

传统 CPU 都是以 X86 架构芯片为主导，主要芯片厂商为 Intel 和 AMD；此外国内在信创市场的带动下，有华为海思（鲲鹏芯片，ARM 架构）、海光（X86 架构）、飞腾（ARM 架构）、龙芯（LoongArch 架构）等多家国产厂商快速成长；英伟达在 2023 年上半年推出了服务器 CPU，而后通过 2 个 GPU 捆绑 1 个 CPU 做成的 GB200 超级芯片，带动了自家 Grace CPU（ARM 架构）的销售，同时英伟达即将在 2026 年推出自主架构的 Vera CPU（Olympus 架构）；由于 CPU 在数据中心成本中占比较大，云厂商也进行了 CPU 的自研工作，2022 年 AWS（ARM 架构）的市场份额已经达到了 3.16%；以多核为特色的 Arm 服务器厂商 Ampere，在 2024 年 5 月推出了 256 核的 CPU，给希望提升整体性能的客户带来了更多的选择。

从市场份额来看，英特尔仍然占据主导地位（2022 年市占率 70.77%），但是 AMD 的竞争优势却在持续提升。

从架构角度来看，到 2022 年 X86 CPU 的市场份额仍然超过 90%，但 ARM 等其他架构的市场份额仍在持续提升，主要动力包括了 AWS 和阿里云等的自研 CPU、Ampere、华为海思、英伟达等。

图 9：2021 年和 2022 年全球数据中心 CPU 市场收入份额

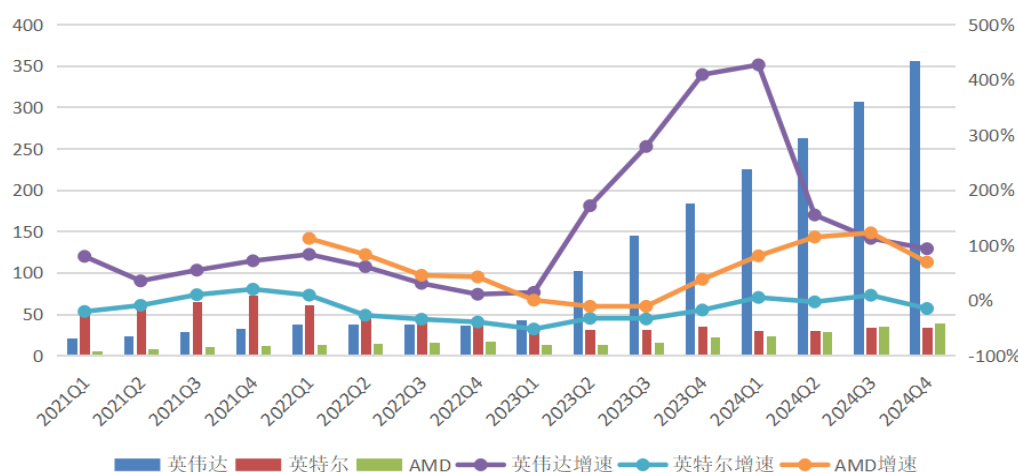


资料来源：Counterpoint Analysis，中原证券研究所

3.1.2. GPU：AI 业务带动需求爆发式增长

GPU 起源于图形处理需求。由于人们对图形处理的需求持续增大，特别是随着游戏画面的持续提升，CPU 对于图像处理的能力已经无法满足要求，就催生出擅长并行计算的 GPU。而在 AI 快速发展的同时，人们开始将 GPU 擅长并行计算的特性应用与 AI 计算，与此同时 GPU 厂商英伟达也成为了 AI 需求增长最大受益方。

图 10：2021Q1 至 2024Q4 英特尔、AMD、英伟达的数据中心业务收入（亿美元）及增速



资料来源：，中原证券研究所（英伟达财报时段较以上周期晚一个月）

2023 年 AI 大模型训练需求暴增，英伟达产能难以满足市场需求，同时大量芯片厂商也加大了在 GPU 的投入步伐。全球范围来看，英伟达仍然是 GPU 市场的主导；AMD 通过 MI300X、MI325X 性能上较好地对标了英伟达，带动了数据中心业务的快速增长；英特尔也推出了 GPU 芯片，但影响力有限；中国市场由于受到了美国 AI 芯片制裁，华为海思（昇腾）、海光、寒武纪、昆仑芯、登临科技、壁仞、沐曦、摩尔线程等一批国产厂商迎来了发展

机会；此外大模型厂商、云厂商也都在积极进行 AI 芯片自研，从而也带动了 ASIC 芯片的快速增长。

在硬件性能之外，软件生态成为了 GPU 厂商发展的关键壁垒。以英伟达来看，其所有的硬件和解决方案都拥有统一的底层软件 CUDA 和运行库 CUDA-X。随着 GPU 计算性能的逐步强大，英伟达在 2016 年推出了 CUDA，让 GPU 可以用于图像计算以外的目的，也大幅提升了 GPU 的计算性能。自 CUDA 推出以来，英伟达在将 GPU 转化为更通用的计算工具上，累计投入了将近 100 亿美元，全球主流的深度学习框架都基于 CUDA 进行加速，生态壁垒强大。

表 6：英伟达、AMD 人工智能芯片参数对比

	H100 SXM	H200 SXM	B200	B300		MI300X	MI325X	MI355X
厂商	英伟达	英伟达	英伟达	英伟达	英伟达	AMD	AMD	AMD
发布时间	2022.3	2023.11	2024.3	2025.3	2026 下半年	2023.6	2024.6	2025
架构	Hopper	Hopper	Blackwell	Blackwell Ultra	Rubin	CDNA 3	CDNA 3	CDNA 4
工艺	4nm	4nm	4nm	4nm	3nm	5nm	5nm	3nm
晶体管数量	800 亿	800 亿	2080 亿			1530 亿	1530 亿	
最大功率	700W	700W	1000W	1400W		750W	1000W	
显存容量	80GB	141GB	192GB	288GB	288GB	192GB	256GB	288GB
显存带宽	3.35TB/s	4.8TB/s	8TB/s	8TB/s	13TB/s	5.3TB/s	6TB/s	8TB/s
显存颗粒	HBM3	HBM3e	HBM3e 8hi	HBM3e 12hi	HBM4	HBM3	HBM3e	HBM3e 12hi
传输带宽	900GB/s	900GB/s	1800GB/s	1800GB/s		896GB/s		
FP64	34 TFLOPS	34 TFLOPS						
FP64 Tensor Core	67 TFLOPS	67 TFLOPS						
FP32	67 TFLOPS	67 TFLOPS						
TF32						653.7 TFLOPS		
TF32 Tensor Core	989 TFLOPS	989 TFLOPS	2.2 PFLOPS					
BFLOAT16	989.4 TFLOPS					1307.4 TFLOPS		
BFLOAT16 Tensor Core	1979 TFLOPS	1979 TFLOPS						
FP16	989.4 TFLOPS					1307.4 TFLOPS	1.3 PFLOPS	2.3 PFLOPS
FP16 Tenor Core	1979 TFLOPS	1979 TFLOPS	4.5 PFLOPS					
FP8	1978.9 TFLOPS					2614.9 TFLOPS	2.6 PFLOPS	4.6 PFLOPS
FP8 Tensor Core	3958 TFLOPS	3958 TFLOPS	9 PFLOPS					

INT8	1978.92TOPS					2614.92 TOPS		
INT8 Tensor Core	3958 TOPS	3958 TOPS	9 PFLOPS					
FP6	-	-	9 FLOPS			-		9.2 PFLOPs
FP4	-	-	18 PFLOPS	27 PFLOPS	50PFLOPS	-		9.2 PFLOPs

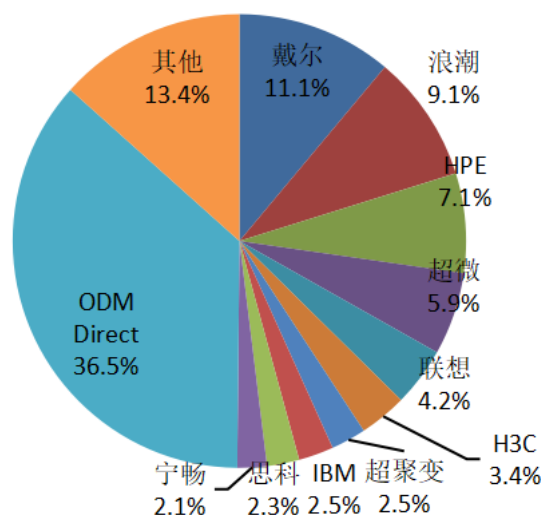
资料来源：英伟达，AMD，中原证券研究所

3.2. 服务器：中国形成了独立于全球的竞争格局

从全球服务器产业发展来看大致可以分为几个阶段：（1）集中式架构下，昂贵的 IBM 大型机、小型机这类专用服务器主导高端客户需求；（2）分布式架构对集中式架构的取代，需求端采用了云的形式实现了 X86 服务器集群对于大型机和小型机的替代；（3）随着大模型在 2023 年以后快速成长，AI 服务器作为新的高价值服务器（训练服务器普遍在 200 万元量级）需求爆发，与此同时传统服务器（数万元量级）需求下滑，同时英伟达通过在 GPU 的垄断地位，推出了 GB200 NVL72 这一类的高性能 AI 整机架服务器。

从全球服务器市场格局来看，头部的品牌厂商主要包括了戴尔、浪潮、HPE、超微、联想、新华三、超聚变、IBM、思科、宁畅等，其中浪潮、联想、超聚变、宁畅都是国产厂商；随着云厂商的壮大，ODM 白牌服务器厂商的业务占比也持续增大。

图 11：2023Q3 全球服务器市场厂商营收份额

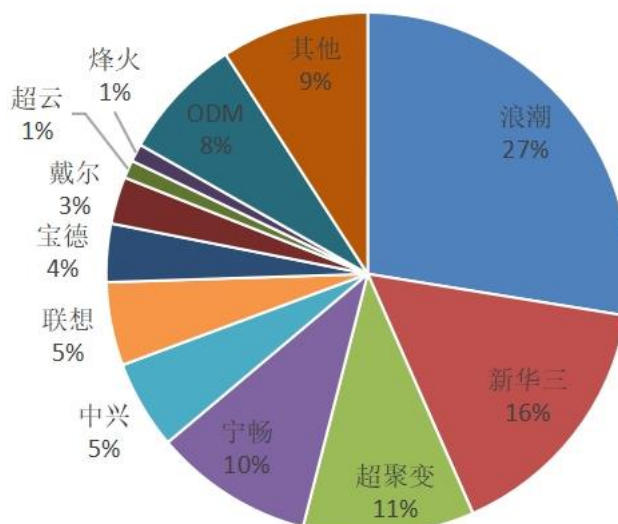


资料来源：IDC，中原证券研究所

同中国的互联网、云计算等市场一样，我们的服务器市场也呈现出了独立于全球的竞争格局，本地厂商占据了绝对的主导。戴尔、HPE 这类全球头部品牌服务器厂商在国内的份额逐年递减；浪潮依靠早年深度绑定互联网行业客户，多年稳居行业第一；由于服务器市场份额大、布局门槛较低，新华三（原华为与 3Com 布局企业网络的子公司）、超聚变（原华为服

务器业务子公司）、中兴、烽火这类原网络设备厂商都在积极进行服务器业务布局，其中超聚变、新华三位居行业前三；擅长超算和液冷等技术的老牌服务器厂商中科曙光，同时还拥有了子公司海光信息，占据独特优势；联想在收购了 IBM X86 服务器业务以后，在国内市场整体份额下滑；由于国内服务器市场整体竞争激烈，ODM 厂商份额明显小于海外。

图 12：2023 年中国 X86 服务器市场厂商份额（按销售额）



资料来源：IDC，中原证券研究所

总体来看，服务器生产环节整体毛利率偏低，门槛也较低。一方面服务器中的核心部件 CPU、GPU、内存、硬盘等占据了绝大部分的核心价值，赚走了大量的利润，另一方面面临互联网、运营商等大客户利润挤压，面临了来自同行和 ODM 厂商较为激烈的竞争。

3.3. 数据中心：运营商占据主导地位，厂商差异明显

2023 年，我国数据中心服务（IDC）市场规模 1250.1 亿元，同比下滑 3.4%。由于公有云 IaaS 增速的放缓（从 2022 年 21% 下滑到 10.2%）及规模以上互联网厂商增收不增利，数据中心服务的需求不及预期。

从市场竞争格局来看，2023 年我国的 IDC 市场是以三大运营商为主导的，占据了 56.9% 的市场份额，较上一年继续下滑了 3.3 PCT。运营商通过考核，将 IDC 业务向云计算牵引，加速向云业务转型。

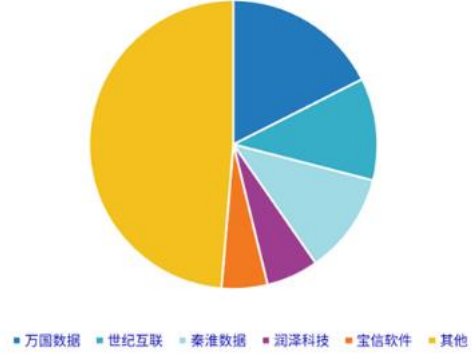
第三方 IDC 厂商中，排名靠前的厂商分别为万国数据、世纪互联、秦淮数据、润泽科技、宝信软件，2023 年前 5 大厂商共计占有 51.3% 的市场份额，依托核心地段资源优势以及更优质的经营能力，集中度持续提升。

图 13：2023 年我国数据中心服务商市场厂商份额



资料来源：IDC，中原证券研究所

图 14：2023 年我国第三方数据中心服务商市场厂商份额



资料来源：IDC，中原证券研究所

从厂商数据财务来看，厂商的毛利率差异巨大，从 17%到 50%不等。这也反映出了不同厂商在竞争能力，发展策略等方面的差异。第三方数据中心的头部厂商中，万国数据和世纪互联虽然规模优势明显，但是都面临较为严重的亏损状况。而服务于字节的秦淮和润泽都取得了较好的盈利能力和增长趋势。宝信软件借助前期高价值的上海数据中心储备，取得了行业最高的毛利水平。数据港和光环新网增速较为缓慢，奥飞数据仍然处于快速增长期。

表 7：2023 年企业 IDC 相关业务对比（亿元）

公司	相关业务	2023 年							2022 年
		收入	增速	净利润	增速	毛利率	毛利额	人员增速	收入
万国数据		100	7%	-43	-238%	19%	19	7%	93
世纪互联		74	5%	-26	-241%	17%	13	-8%	71
秦淮数据*						42%			46
润泽科技		44	60%	17.6	47%	49%	21	25%	27
宝信软件	服务外包	35	2%	-	-	50%	17.5	-	34
光环新网	IDC 及其增值服务	22	6%			36%	7.9		21
数据港		15	6%	1.2	5%	29%	4.5	-2%	15
奥飞数据		13	22%	1.4	-15%	27%	3.6	20%	11

资料来源：上市公司公告，，中原证券研究所（*秦淮退市，故取 2022 年数据）

2024H1，润泽科技实现了收入翻倍的增长，同时毛利率有一定下滑，但是依然处于领先地位；万国数据和世纪互联虽然保持了稳定的增长，但是依然处于亏损状态；奥飞数据在上半年也实现了 59%的增长，但毛利率较上年同期下滑了约 4 个百分点；宝信软件、光环新网、数据港保持了业务的相对稳定。

表 8：2024H1 企业 IDC 相关业务对比（亿元）

公司	相关业务	2024H1						2024 年		
		收入	增速	净利润	增速	毛利率	毛利额	收入	增速	毛利率
万国数据		55	12%	-6	19%	22%	12	100	7%	19%
世纪互联		39	7%	-1.2	18%	21%	8	74	5%	17%
润泽科技		36	112%	10	38%	35%	12	44	60%	49%
宝信软件	服务外包	16	-1%	-	-	49%	8	35	2%	50%

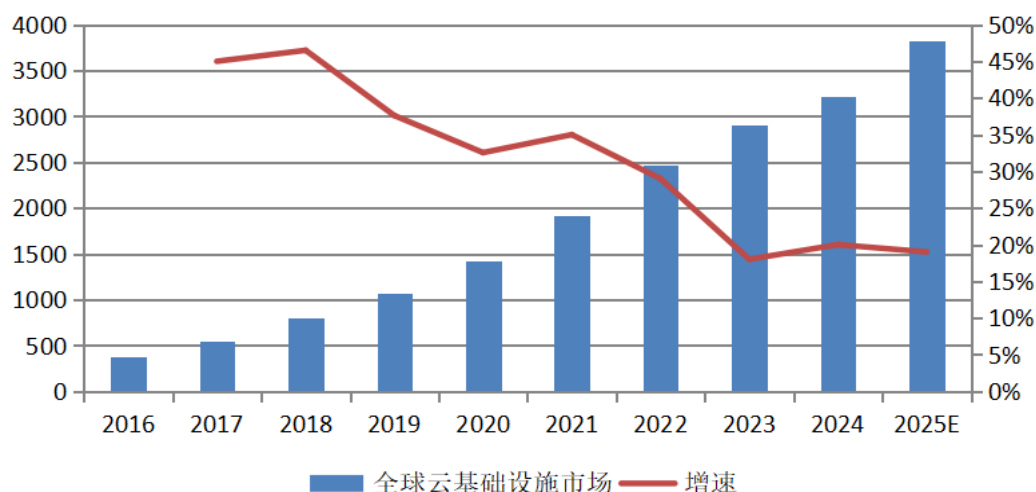
光环新网	IDC 及其增值服务	11	1%	-	-	35%	3.7	22	6%	36%
奥飞数据		11	59%	0.8	3%	24%	2.5	13	22%	27%
数据港		8	4%	0.7	3%	27%	2.1	15	6%	29%

资料来源：上市公司公告， ，中原证券研究所

3.4. 云计算：巨头云集，马太效应明显

根据 Canalys 的数据，2024 年全球云基础设施服务支出市场规模 3213 亿美元，同比增长 20%。随着 AI 需求的持续升温，让全球云计算业务保持了较高的增长速度，预计 2025 年仍将获得 19% 的增速。

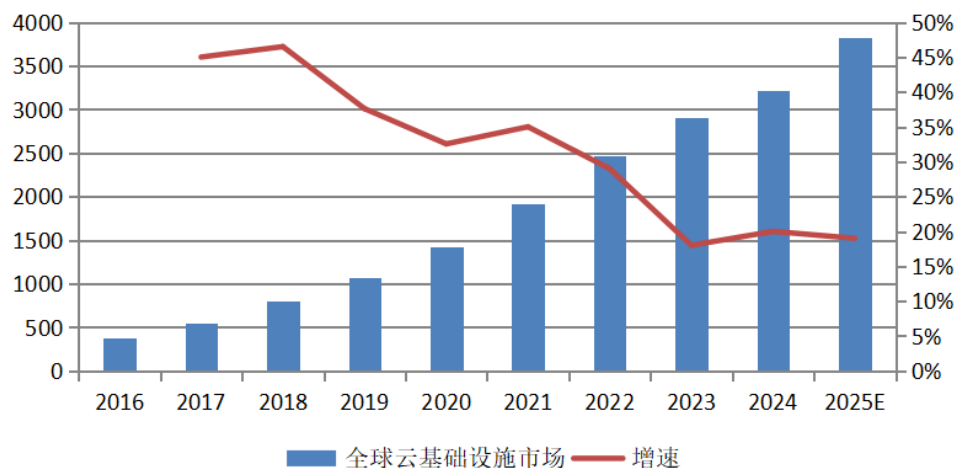
图 15：2016-2025E 全球云基础设施服务支出（亿美元）



资料来源：Canalys，中原证券研究所

云计算市场总体马太效应非常强。从全球市场来看，亚马逊、微软、谷歌稳居头部位置，2024Q4 三家合计占据了云计算市场 64% 的市场份额，全都是全球头部的科技厂商。

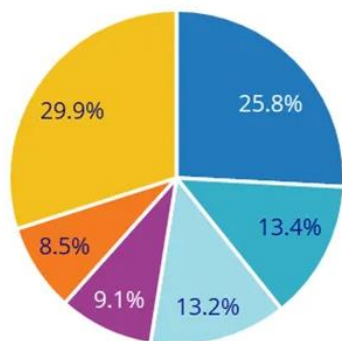
图 16：2024Q2 年全球云基础设施服务支出市场份额（亿美元）



资料来源：Canalys，中原证券研究所

不同于全球云计算市场，中国形成了相对独立的云计算竞争格局。行业龙头为阿里云，近年来增速放缓较为明显；华为云和腾讯云也较早布局了云计算市场；AWS 作为全球龙头，在国内始终保持了前五大厂商地位；2020 年以后，三大运营商发力云计算市场，成为了政府数据重要的承载者，目前也成为市场头部厂商；金山云借助小米系相关业务加快成长，也是市场的重要参与者。

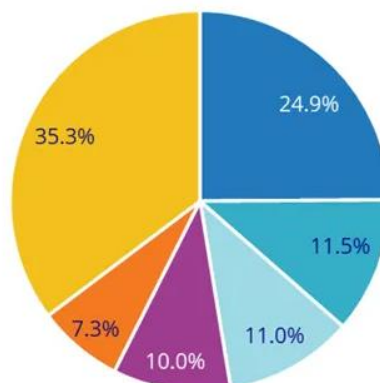
图 29：2024H1 中国公有云 IaaS 厂商市场份额



■ 阿里巴巴 ■ 华为 ■ 中国电信 ■ 中国移动 ■ 腾讯 ■ 其他

资料来源：IDC，中原证券研究所

图 30：2024H1 中国公有云 PaaS 厂商市场份额

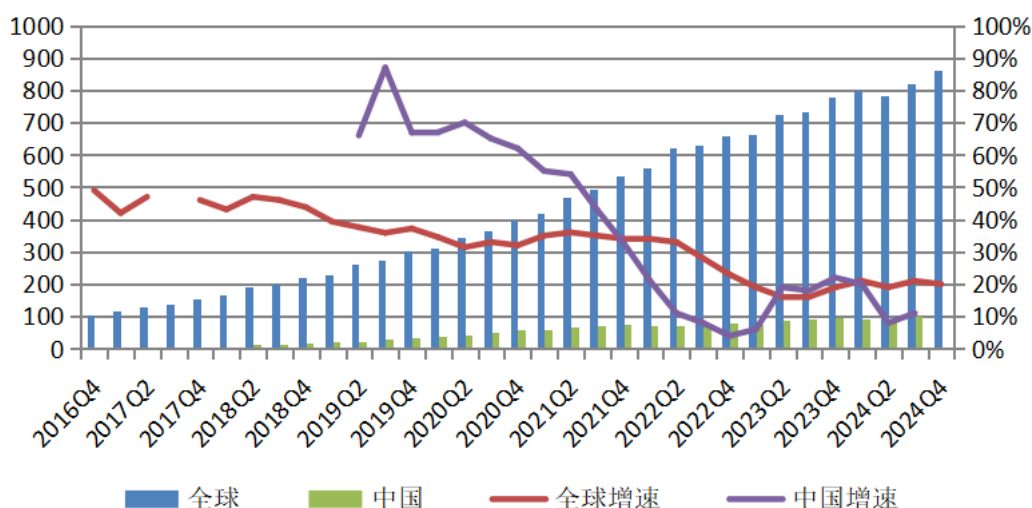


■ 阿里巴巴 ■ 腾讯云 ■ 华为 ■ AWS ■ 中国电信 ■ 其他

资料来源：IDC，中原证券研究所

从全球与中国云市场增速对比来看，2021Q4 以后，市场迎来拐点，中国云计算市场增速整体落后于全球水平（除 2023Q2-2023Q4），我们预计随着 DeepSeek-R1 的开源和国内平台经济的加速发展，后续中国的云计算发展速度有望再次超过全球平均水平。

图 17：2016Q4-2024Q4 年全球与中国云基础设施服务支出市场增速对比（亿美元）



资料来源：Canalys，中原证券研究所

3.5. 大模型：市场格局还处于高度变化时期，未来产业生态有望进一步集中

大模型市场格局还处于高度变化的时期。OpenAI 在引领了早期大模型市场，成为了行业标杆。但是在最近的 2 年多时间中，Claude、谷歌、Meta、xAI 也都在紧跟 OpenAI 的步伐；国内厂商虽然布局较晚，但是与海外的差距持续缩小，字节跳动的豆包实现国内应用市场的后来居上；DeepSeek 在发布并开源 DeepSeek-R1，追平了海外大模型水平，并在依靠低推理成本，开启了国内大模型厂商用技术弥补算力不足的新路径；阿里积极进行模型开源，总体模型实力较强，是 iPhone 在国内模型服务器供应商；腾讯拥有庞大生态基础和微信入口优势，未来在 AI 布局能力可期。

表 9: OpenAI 大模型产品竞品推出情况

对标模型	时间	说明	厂商	产品	推出时间	备注
GPT-4o	2024.5.14	多模态交互模型	谷歌	Project Astra	2024.5.15	
				Gemini Live	2024.8.14	
				Gemini 2.0 Pro	2025.2.5	
			Anthropic	Claude 3.5 Sonnet	2024.6.21	
GPT-4.5	2025.2.28	较 4o 有更高情商	xAI	Grok-2	2024.8.14	
				Grok-3	2025.2.18	
			智谱	GLM-4-Plus	2024.8.29	
			阿里	Qwen-Max 升级版	2024.9.19	
				Qwen2.5-Max	2025.1.29	
			字节	豆包通用模型 pro	2024.12.18	价格为 GPT-4o 的 1/8
			幻方	DeepSeek-V3	2024.12.26	开源
o1	2024.9.13	推理模型	Anthropic	Hunyuan-TurboS	2025.2.27	
				升级版 Claude 3.5 Sonnet	2024.10.23	
			谷歌	Claude 3.7 Sonnet	2025.2.24	首款混合推理模型
				Gemini 2.0 Flash Thinking	2024.12.20	
o3-mini	2025.1.31	o1 的下一代产品	幻方	DeepSeek-R1-Lite-Preview	2024.11.21	对标 o1-preview
				DeepSeek-R1	2025.1.20	开源
			阿里	QwQ-32B-Preview	2024.11.28	对标 o1-mini 和 o1-preview 开源
				QwQ-32B	2025.3.6	开源
			智谱	GLM-Zero-Preview	2024.12.31	对标 o1-preview
			xAI	Grok-3 Reasoning Beta	2025.2.18	
				Grok-3 mini Reasoning	2025.2.18	

资料来源：OpenAI，界面新闻，机器之心 Pro，腾讯科技，中原证券研究所

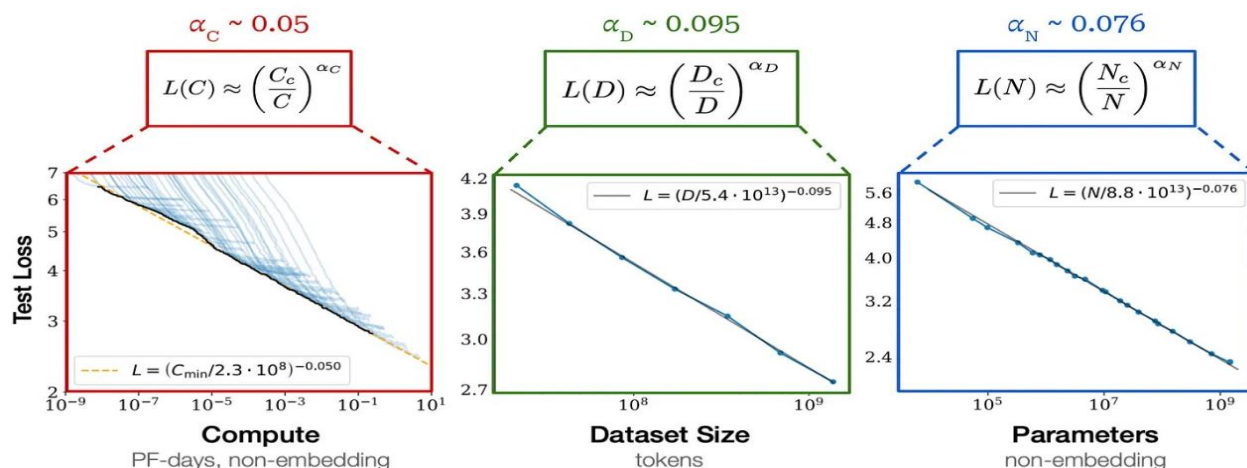
2025 年模型生态或将进一步向头部厂商集中，或将看到更多的大模型厂商向行业模型转型。从大模型市场格局来看，一方面，行业标杆 OpenAI 逐渐从开源走向闭源，意图构建技术壁垒；另一方面，Meta、阿里积极加入开源模型构建，但是在模型先进性方面落后于闭源模型。DeepSeek-R1 的开源，在模型能力上比肩业界头部模型，在成本上形成了明显的优势，打破闭源模型厂商的技术壁垒，推动了 AI 的技术平权，并加速了大模型应用的落地。

4. AI 拉动了算力需求的增长

4.1. 大模型技术下，算力需求爆发性增长

2022 年 11 月，OpenAI 推出的 ChatGPT 火爆全网，由此拉开了大模型加速发展的序幕。OpenAI 的成功源于其利用了 Transformer 的所遵循的 Scaling Law（规模效应），在增大模型参数数量、训练数据、用于训练是计算量的同时，带来系统能力稳步的增长。甚至在模型参数达到一定量级，模型能力会得到显著提升，表明较大模型有涌现能力。

图 18：参数、数据和计算量与 LLM 性能的关系（在 WebText2 上的测试损失）



资料来源：Scaling laws for neural language models，机器之心，中原证券研究所

因而，我们看到 OpenAI 在模型下一代模型中，不论参数量还是数据集都有显著的增长。GPT、GPT-2、GPT-3、GPT-4 的模型（最大）参数分别为 1.17 亿、15 亿、1750 亿、1.8 万亿，训练数据也从 GPT 约 7000 本书的 BooksCorpus，增加到 GPT-2 的互联网数据抓取的 WebText，到 GPT-3 约大 10 倍于 WebText 的 CommonCrawl。

表 10：OpenAI 大模型产品

模型	时间	参数	上下文限制	训练数据截止时间	公开的程度	备注
GPT	2018.6	1.17 亿			开源（技术细节）	
GPT-2	2019.2	15 亿			开源（源代码）	
GPT-3	2020.5	1750 亿		2019.10	论文	
GPT-3.5	2022.3	-	4K/16K	2021.6		2022.11 ChatGPT 基于 GPT 3.5 推出
GPT-4	2023.3	1.8 万亿	8K/32K	2021.9	测试结果	2023.7 向所有开发者开放使用权限
GPT-4 Turbo	2023.11	-	128K	2023.4		速度是 GPT4 的 5 倍，价格降低 2/3
Sora	2024.2					首个文生视频模型
GPT-4o	2024.5					具有了强大的多模态交互能力 对用户免费，API 速度较 GPT-4 Turbo 提高 2 倍，价格降低 50%
o1	2024.9					更擅长解决科学、编码、数学和类似领域的复杂问题

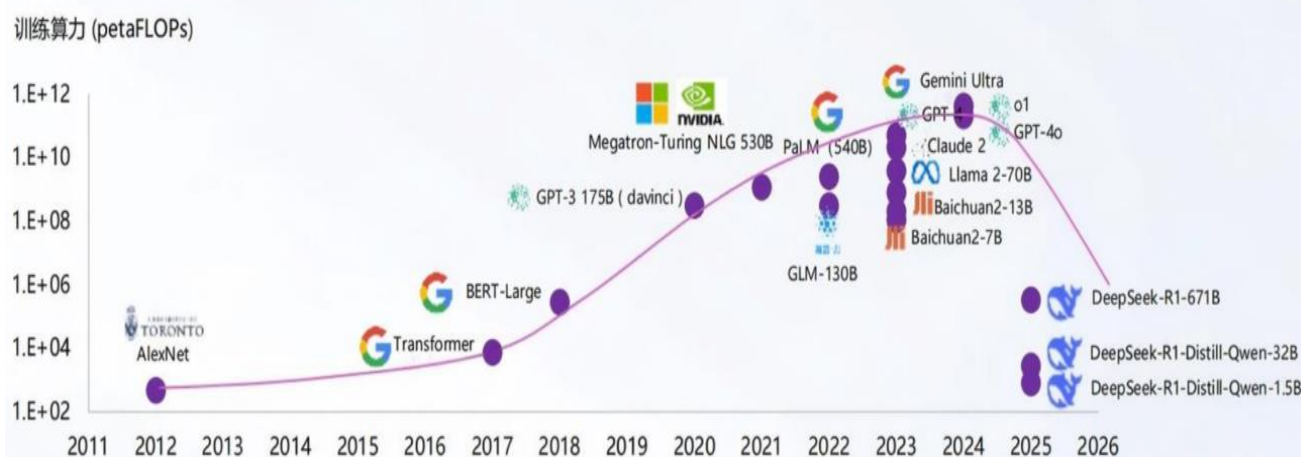
o3-mini	2025.1.31					
GPT-4.5	2025.2.28					具有更高情商

资料来源：OpenAI，中原证券研究所

4.2. 推理模型将 Scaling 向后训练和推理阶段转移，应用落地加大推理需求

随着模型持续增大，我们看到 **Scaling Law** 失效的风险。一方面，高质量数据已经作为训练数据，面临增长瓶颈，另一方面，头部模型厂商在更大参数的新一代模型的推进方面进度不及预期。对于前者，人类试图通过合成数据提高数据供给量，对于后者而言，也有 DeepSeek-V3 这样的模型，探讨出了更多节省算力的工程改进方案，谷歌的 DiLoCo 也试图通过分布式训练解决模型持续增长的问题，为更大参数规模的模型预训练创造了可能性。

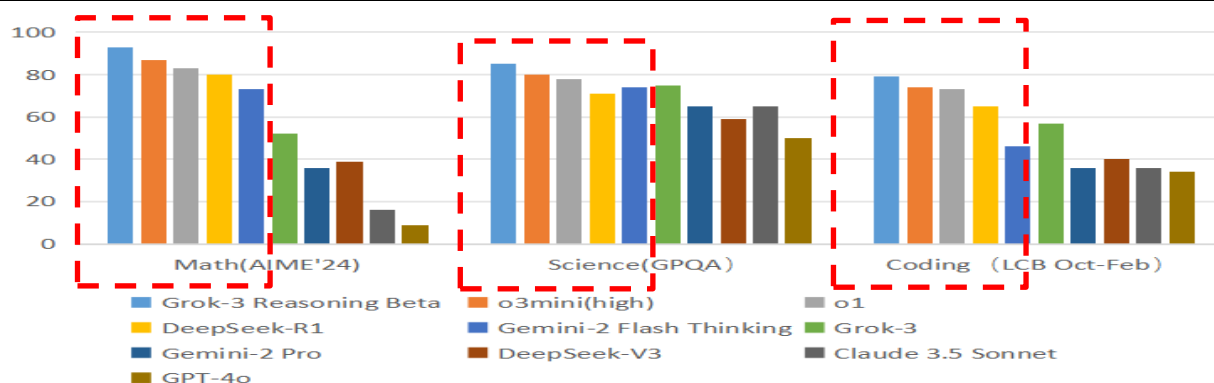
图 19：主流大模型训练算力需求



资料来源：厦门大学，中原证券研究所

o1 为代表的推理模型的出现，探索出了大模型能力提升的新路径。此前的大模型在逻辑推理能力方面能力不足，导致科学界对于 LLM 的技术路线能否在提高语言能力的同时，高效提高逻辑能力和推理能力存在较大质疑，而 OpenAI 的 **o1** 有效提升了大模型解决科学、编码、数学和类似领域的复杂问题。

图 20：推理模型与语言模型在数学、科学、编码方面的能力对比



资料来源：xAI，DeepTech 深科技，中原证券研究所

推理模型改变了原有大模型在预训练上 **Scaling** 的范式，将算力开销进一步向后训练和推理阶段转移。以 DeepSeek-V3 为例，我们看到其在预训练阶段的训练时间为 266.4 万 H800 GPU 小时，文本延伸需要 11.9 万 GPU 小时，而后训练只需要 0.5 万 GPU 小时。

图 21：DeepSeek-V3 在模型训练不同阶段的开销

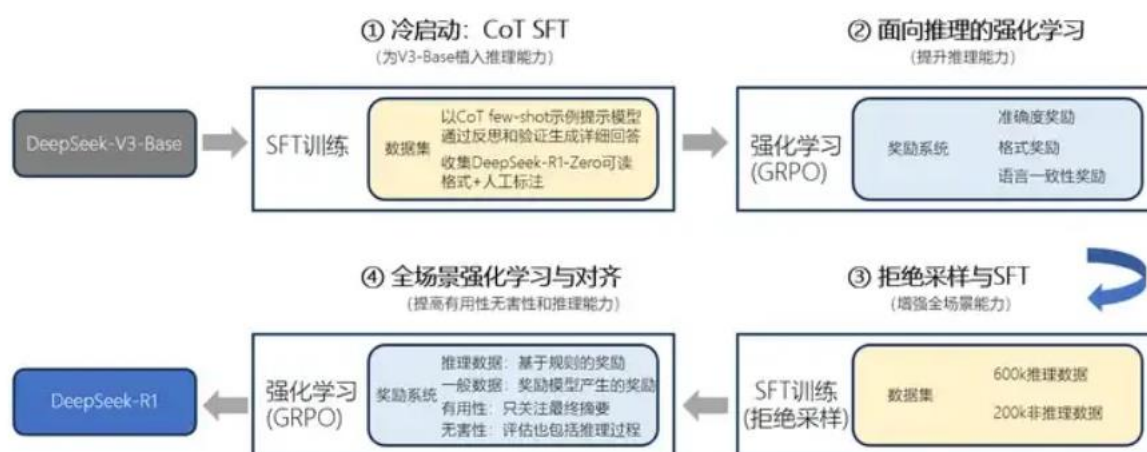
Training Costs	Pre-Training	Context Extension	Post-Training	Total
in H800 GPU Hours	2664K	119K	5K	2788K
in USD	\$5.328M	\$0.238M	\$0.01M	\$5.576M

Table 1 | Training costs of DeepSeek-V3, assuming the rental price of H800 is \$2 per GPU hour.

资料来源：DeepSeek，中原证券研究所

推理模型通过对基础模型进行强化学习等后训练过程来提升逻辑推理能力。DeepSeek-R1 的训练过程来看，DeepSeek 使用 DeepSeek-V3-Base 作为基础模型，通过强化学习（RL）训练出了 DeepSeek-R1-Zero，又在 R1-Zero 训练方法的基础上加入了多阶段的训练，解决可读性差和语言混合等问题。

图 22：DeepSeek-R1 的训练流程

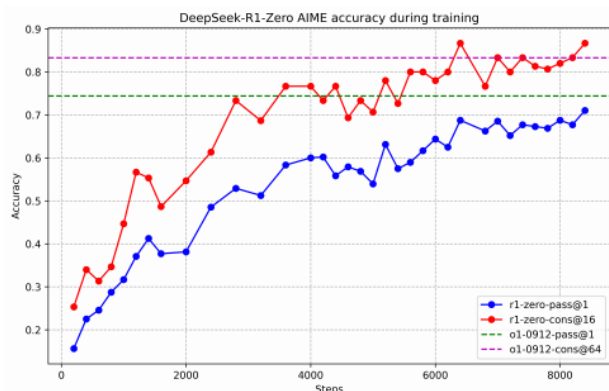


资料来源：DeepSeek，中存算，中原证券研究所

在 DeepSeek-R1-Zero 的训练中，通过增加强化学习的训练迭代步数，其思考的时间在持续增加，AIME 精度也在持续提升。

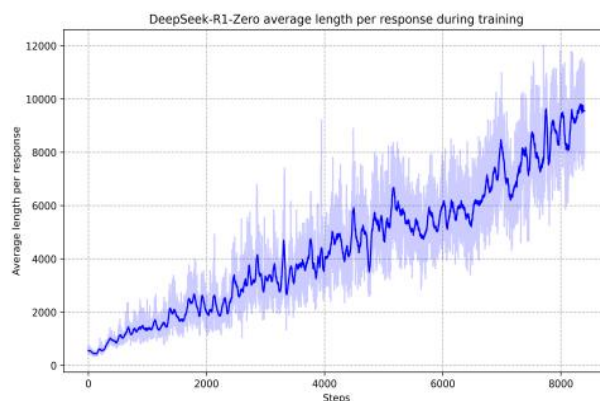
推理模型使用的推理策略在很大程度上基于思维链，模型用它们来思考问题、将问题分解成更简单的步骤、尝试各种解决问题的方法，也可以纠正自己的错误。而思维链的生成也导致了模型在推理过程中的算力大幅增长。

图 23: DeepSeek-R1-Zero 在训练中 AIME 精度提升情况



资料来源: DeepSeek, 中原证券研究所

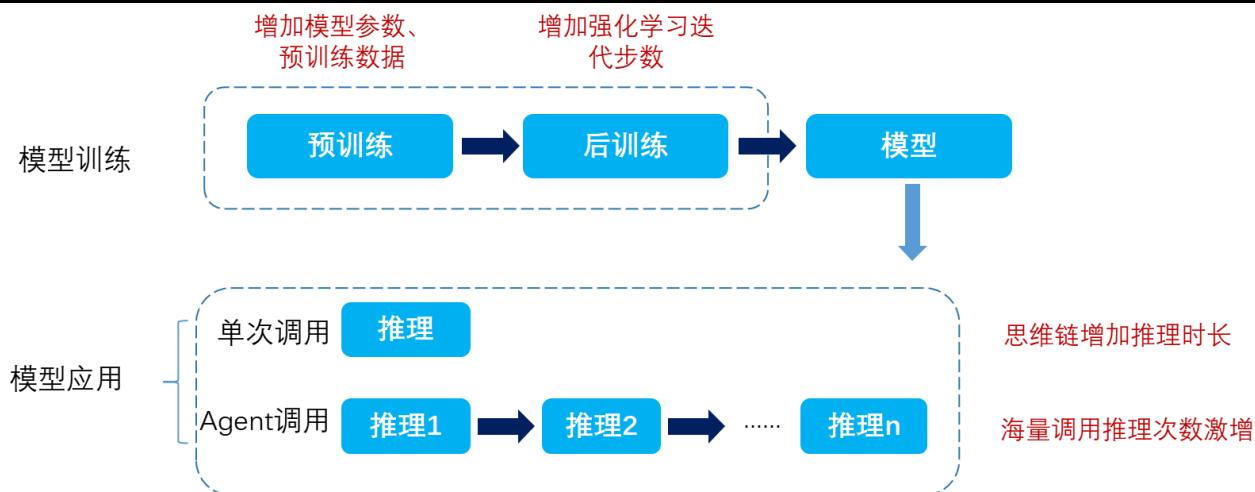
图 24: DeepSeek-R1-Zero 在训练中的平均响应时长



资料来源: DeepSeek, 中原证券研究所

同时, DeepSeek - R1 发布后, 大模型逻辑推理能力的提升、大模型成本显著下降、开源阵营在性能上实现超越, AI 应用集齐全面落地的三大关键要素。随着 AI 应用进入爆发期, 应用端的推理需求将成为下一阶段拉动算力增长的主要驱动力。

图 25: 模型 Scaling 的新范式



资料来源: 中原证券研究所

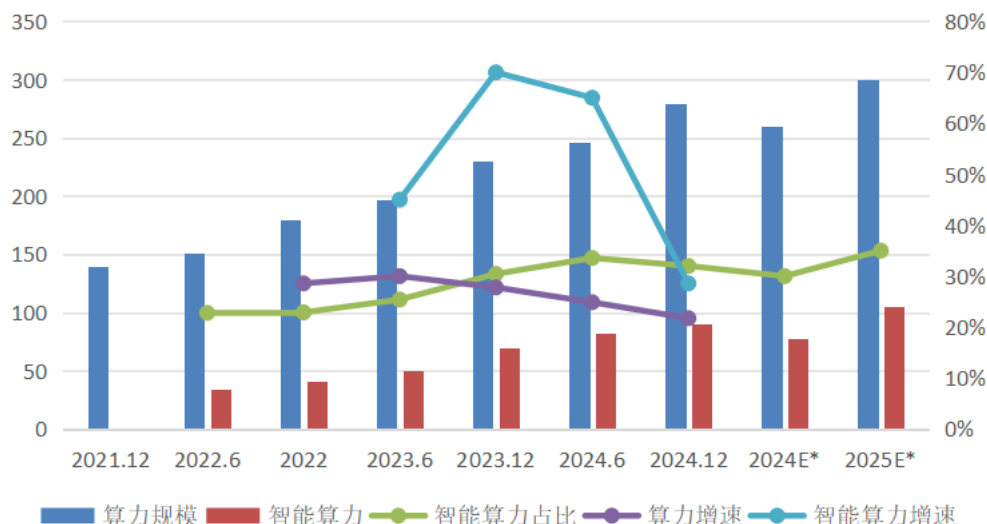
4.3. 截止 2024 年底, 我国智能算力建设领先于国家整体规划水平

在全国总体算力发展目标方面, 我们设定了 2025 年的发展目标。根据 2023 年 9 月工信部等六部门联合印发的《算力基础设施高质量发展行动计划》, 计划在 2025 年算力规模超过 300EFLOPS (EFLOPS 是指每秒百亿亿次浮点运算次数), 智能算力占比达到 35%, 存储总量 1800EB, 其中先进存储容量占比 30%, 国家枢纽节点数据中心集群间网络时延达标率 80%。

从完成情况来看, 我们在智算力建设明显领先于规划水平。根据工信部数据, 截止 2024 年底, 我国算力总规模达 280EFLOPS (FP32), 其中智能算力 90EFLOPS (FP32), 占比达 32%。对比《算力基础设施高质量发展行动计划》给出的 2024 参考值来看, 算力总规模超

出 20EFLOPS，其中 12EFLOPS 来自智能算力的贡献，智算占比也高于参考值 30%。2024 年的存储总量 1580EB，超出参考值 80EB，先进存储占比 28%，与参考指标相当。国家枢纽节点算力中心集群间网络时延达标率超 75%，与参考指标相当。

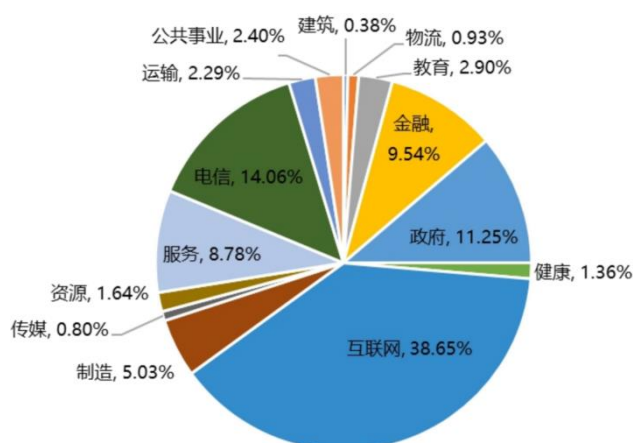
图 26：2021-2025E 我国算力规模及结构（EFLOPS，FP32）



资料来源：《算力基础设施高质量发展行动计划》，工信部，信通院，中原证券研究所
(*指《算力基础设施高质量发展行动计划》中的算力基础设施高质量发展指标)

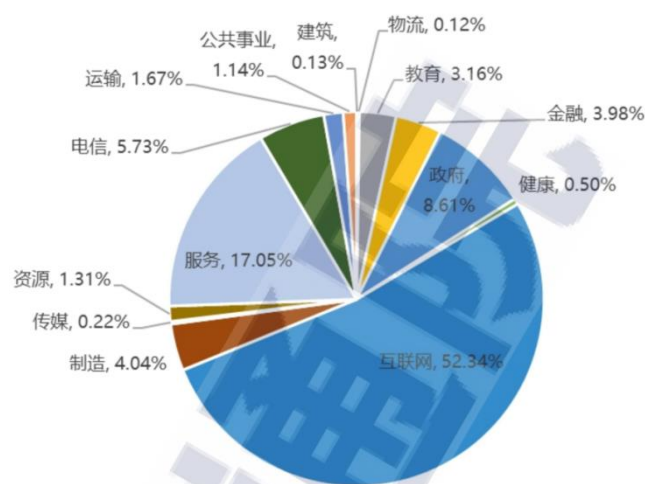
从算力的行业分布来看，智能算力更大受到了互联网、服务行业的需求驱动，两者需求占比分别为 52.34%和 17.05%，远高于通用算力中 38.65%和 8.78%的占比；教育行业占比 3.16%，也较通用算力中 2.9%的占比更高；作为传统需求较大的行业，电信、政府、金融、制造、公共事业的智算需求都有待进一步发掘。

图 29：通用算力行业应用情况



资料来源：信通院，中原证券研究所

图 30：智能算力行业应用情况

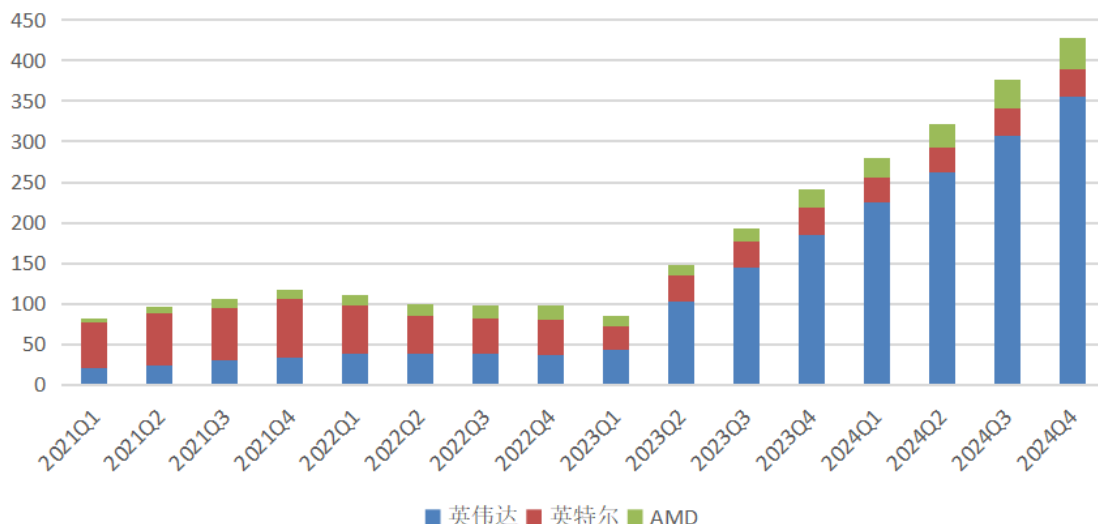


资料来源：信通院，中原证券研究所

4.4. AI 驱动了算力建设热情，全球科技厂商仍在积极加大对算力的资本投入

AI 需求带来了数据中心芯片收入的爆发式增长。受到 AI 技术突破的带动，2023Q2 开始，英伟达数据中心收入开始持续上涨。截止 2024Q4，英伟达、AMD、英特尔三大主要芯片厂商的数据中心收入较 2022Q4 增长了 342%。

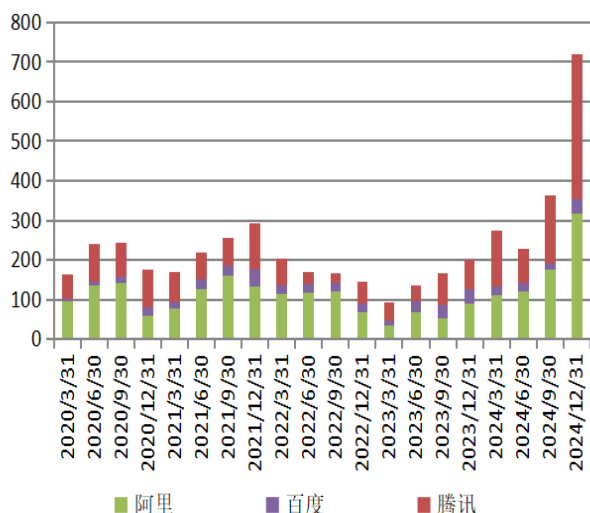
图 27：2021Q1-2024Q4 主要芯片厂商数据中心业务收入（亿美元）



资料来源：上市公司财报，中原证券研究所（英伟达财报较正常季度延后 1 个月）

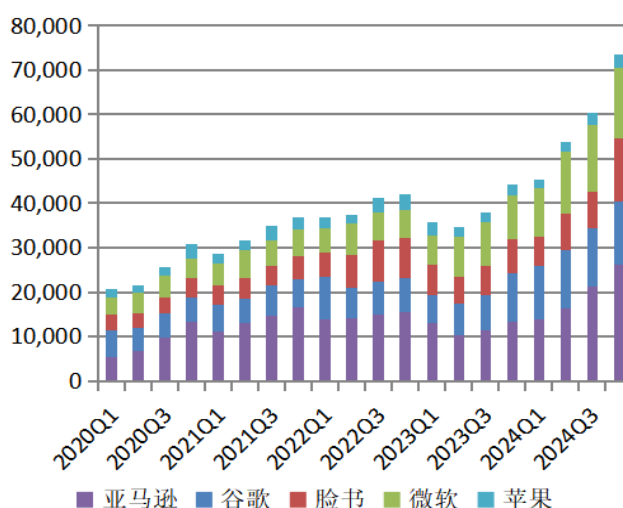
而从 2023 年以来国内外科技厂商资本投入来看，国内厂商虽然前期投入进程慢于海外，但是在 2024 年下半年加速趋势都非常明显。美国 5 大科技厂商（亚马逊、谷歌、Meta、微软、苹果）早在 2023Q4 就超过了前期高点，而国内 3 大厂商（阿里、百度、腾讯）直到 2024Q3 才超过前期高点，但是在 2024Q4 获得了同比 260% 和环比 99% 的增长。

图 29：阿里、腾讯、百度的资本开支（亿元）



资料来源：上市公司财报，中原证券研究所

图 30：5 大科技厂商资本投入（百万美元）



资料来源：，中原证券研究所

在 2025 年资本开支方面：

海外：亚马逊计划投资 1000 亿美元，谷歌为 750 亿美元，Meta 为 600-650 亿美元，都较上一年有了较为明显的提升，微软也提出了 2025 财年（6 月结束）800 亿的数据中心投资

目标。考虑到除了微软以外，其他科技厂商都是在 DeepSeek-R1 发布以后才公布的投资目标，这表明 2025 年全球在 AI 领域的投入仍将呈现快速增长趋势。

国内：阿里未来三年将投入超过 3800 亿元，用于建设云和 AI 硬件基础设施，总额超过去十年的总和。腾讯计划在 2025 年进一步增加资本支出，并预计资本支出将占收入的低两位数百分比。中国移动 2025 年计划资本投入算力 373 亿元，同比增长 0.54%，智算从 29.2EFLOPS（FP16）增长到 34EFLOPS，通算达到 8.9 EFLOPS（FP32），推理投资不设上限。中国联通 2025 年在总资本开支下滑 10% 的基础上，算力投资同比增长 28%。

表 11：大模型领域相关算力相关投入

公司	时间	详情
微软	2024.3.31	与 OpenAI 正计划建设一个名为“星际之门”的特殊数据中心，为 OpenAI 提供强大的计算支持，计划最快在 2028 年启动该项目，并且 2030 年之前会进一步扩建，所需要的电力可能高达 5 千兆瓦。“星级之门”超算会配备百万专用的服务器芯片，项目成本预计高达 1150 亿美元 。
	2025.1.4	将在本财年（6 月结束）投资 800 多亿美元 建设人工智能负载数据中心，以训练人工智能模型，并部署人工智能和基于云的应用程序。
OpenAI	2025.1.21	美国总统特朗普在白宫宣布，由美国甲骨文公司、OpenAI 和软银共同出资投建，三家公司组成的联合企业将投资 5000 亿美元，在美建设名为“星际之门”的 AI 基础设施。该项目计划在美国建设数据中心，以支持人工智能技术发展。
亚马逊	2024.10.29	亚马逊正在核能上押下重注，包括在美国三个州投资超过 520 亿美元 ，以帮助推动其庞大数据中心的扩张。
	2024.12.16	AWS 计划追加投资约 100 亿美元，在俄亥俄州扩建数据中心基础设施。新数据中心将包含计算机服务器、数据存储驱动、网络设备和其他形式的技术基础设施，用于支持包括 AI 和机器学习在内的云计算。
	2025.2.6	预计将在 2025 年投入超过 1000 亿美元，高于 2024 年的 830 亿美元。
谷歌	2024.12.11	谷歌与能源公司 Intersect Power 和投资公司 TPG Rise Climate 合作，计划在本十年内投资 200 亿美元 在美国建设多个“工业园区”，这些园区将配备可再生能源发电设施，以支持数据中心运营。第一个园区预计将在 2026 年部分投入运营，并于 2027 年全面完工。
	2025.2.4	计划在 2025 年投资约 750 亿美元。
Meta	2025.1.24	计划在 2025 年投入 600 亿至 650 亿美元，用于 AI 相关的资本支出，计划在 2025 年上线 1GW 的算力，预期到年底时能够拥有 130 万块 GPU。
xAI	2024.9.3	xAI 仅用 122 天就建成了 10 万张 H100 的 Colossus 集群，未来还会扩展到 15 万张 H100 和 5 万张 H200。
	2024.12.5	xAI 承诺将其超级计算机项目 Colossus 的算力扩大十倍，未来将搭载 100 万块 GPU ，从而对抗谷歌、OpenAI 和 Anthropic 等竞争对手。
小米	2024.12.26	小米集团正在积极构建自己的 GPU 万卡集群，以加大对 AI 大模型的投入。
字节	2024.3.1	截止 2023 年 9 月，字节跳动已经建立超过一万张的英伟达 Ampere 架构 GPU 集群，目前正在建设 Hopper 架构的集群。
	2024.12.31	Information 报道，字节跳动计划在 2025 年斥资 70 亿美元购买英伟达芯片，该消息随后被字节跳动发言人否定。
	2025.1.22	据财联社报道，已为 2025 年在中国购买人工智能芯片编列了 400 亿人民币（55 亿美元）的预算，是去年支出的两倍。公司还计划在海外投资约 68 亿美元，采用英

		伟达芯片加强基础模型训练能力。
法国	2025.2.9	未来几年，法国将投资超 1090 亿欧元用于建设 AI。
阿里	2025.2.24	未来三年，阿里将投入超过 3800 亿元，用于建设云和 AI 硬件基础设施，总额超过去十年的总和。
联通	2025.3.18	2025 年在总资本开支下滑 10% 的基础上，算力投资同比增长 28%。
腾讯	2025.3.19	计划在 2025 年进一步增加资本支出，并预计资本支出将占收入的低两位数百分比。
移动	2025.3.21	2025 年计划资本投入算力 373 亿元，智算规模（FP16）突破 34EFLOPS，通算达到 8.9 EFLOPS（FP32）。

资料来源：财联社，科创板日报，福布斯中国，第一财经，IT 之家，新智元，金融界，The Information，同花顺，中原证券研究所

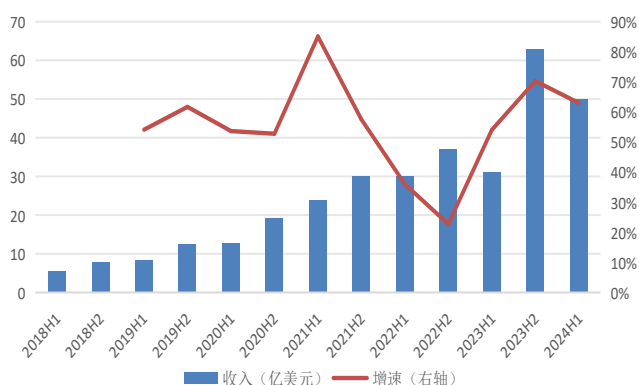
英伟达在 2025 年 GTC 上表示，美国前四大云服务供应商 2024 年总计采购 130 万颗 Hopper 系列芯片，而 2025 年 3 月总计采购已经 360 万颗 Blackwell 架构芯片，需求持续增长。

4.5. 智算增长推动三大细分市场加速发展

4.5.1. AI 服务器仍将持续快速增长，ASIC 市场有望迎来更快增长势头

由于大模型高速发展，AI 服务器市场持续快速增长。根据 IDC 数据，2024H1 我国 AI 服务器市场规模 50 亿美元，较 2023H1 增长 63%。IDC 预测 2024 中国 AI 服务器市场规模将达到 190 亿美元，同比增长 87%。IDC 预测，预计到 2028 年，AI 服务器市场规模将达到 550 亿美元，2023-2028 年年复合增速 40.2%。

图 29：2018H1-2024H1 我国 AI 服务器市场规模及增速



资料来源：IDC，中原证券研究所

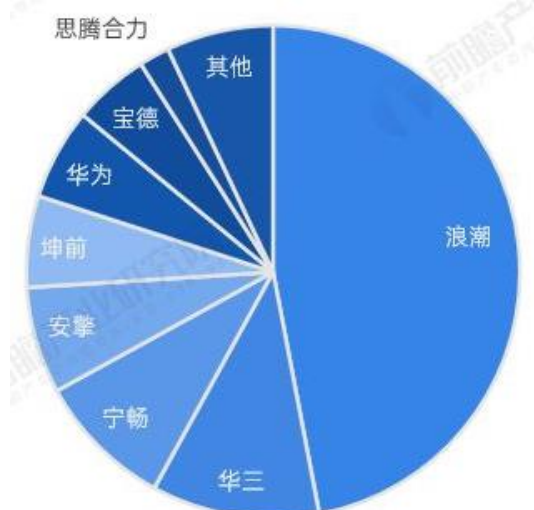
图 30：2024-2028 中国 AI 服务器市场预测 (百万美元)



资料来源：IDC，中原证券研究所

从厂商销售额角度看，2023 年浪潮、新华三、宁畅在国内 AI 服务器市场位居前三，占据了近 70% 的市场份额；从服务器出货台数角度看浪潮、坤前、新华三位居前三名，占有超过 50% 以上的市场份额。由于 IDC 未公布 AI 厂商 2023 年具体份额，我们以 2022 年数值作为参考，可以看到浪潮遥遥领先其他厂商，前 8 大厂商还包括新华三、宁畅、安擎、坤前、华为（超聚变）、宝德、思腾合力等。

图 28：2022 年中国加速计算服务器市场厂商份额分布



资料来源：IDC，前瞻产业研究院，中原证券研究所

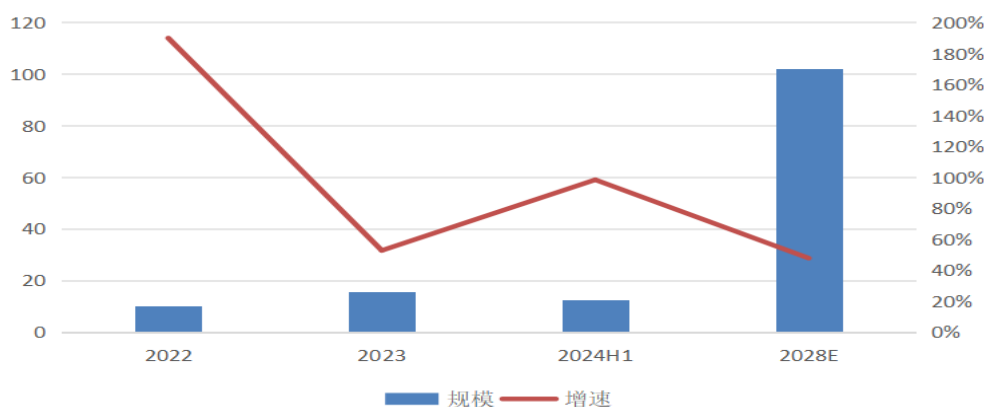
从 AI 服务器下游行业的来看，2023 年互联网作为最大的采购行业，占整体加速服务器市场近 60% 的份额。

IDC 预测 GPU 仍将在 AI 芯片中占据主导，但是 2028 年中国 AI 服务器市场中，ASIC 加速服务器的市场占比将接近 40%。由于 GPU 的供应不足，同时互联网厂商出于降本和更好适配的考虑，也将加大自研 ASIC 芯片，利好于 ASIC 的发展。

4.5.2. 液冷即将迎来规模化应用

液冷服务器将保持持续高增长。根据 IDC 数据，2023 年我国液冷服务器市场规模 15.5 亿美元，同比增长 52.6%。2024H1，我国液冷市场规模 12.6 亿美元，与 2023 年同期相比增长了 98.3%，加速趋势明显。根据 IDC 预测，2023-2028 年，我国液冷服务器市场年复合增速将达到 47.6%，还将保持快速增长势头，2028 年规模将达到 102 亿美元。

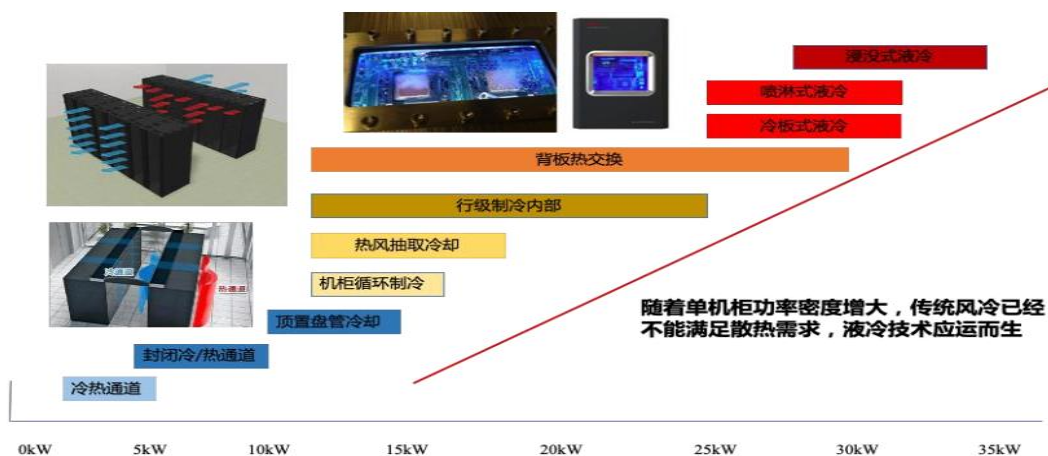
图 29：2022-2028 年我国液冷服务器市场预测（亿美元）



资料来源：IDC，中原证券研究所

随着生成式人工智能的发展，智算需求爆发式增长，同时由于智算中心高密度属性，风冷散热难以满足耗能指标要求，AI 计算、液冷、整机柜的结合将成为未来的发展趋势，液冷行业也迎来了刚需放量的元年。

图 30：单机柜密度和冷却方式



资料来源：赛迪，中原证券研究所

2025 年运营商有望开启液冷技术的规模化应用。互联网行业作为液冷技术积极的拥抱者，2024 年在液冷服务器市场上的份额从上年同期的 54% 提高至 60%。电信运营商也计划在 2025 年开展液冷的规模化应用。

图 31：液冷技术与数据中心能耗指标关系图

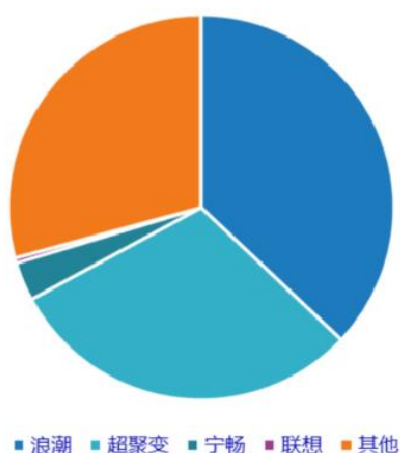


资料来源：《电信运营商液冷技术白皮书（2023 年）》，中原证券研究所

从目前液冷发展来看，核心驱动因素为服务器的功率。当芯片性能提升的同时，芯片耗电量相应提升，单机柜的功率也随之提升。由于液体的比热容远高于气体，单位温度变化所需吸收/放出的热量远大于空气，使散热效率得到极大提升，降低风扇、空调等风冷方面的能源消耗，极大降低数据中心 PUE 值。根据 IT 时报的报道，曙光数创测算下，当单个机架功率超过 10kW 以后，采用冷板液冷架构的单 kW 制冷成本将低于传统的风冷架构。当前新建机架功率 20kW 起步，并向 60kW 普及，对液冷服务器将形成刚性需求。

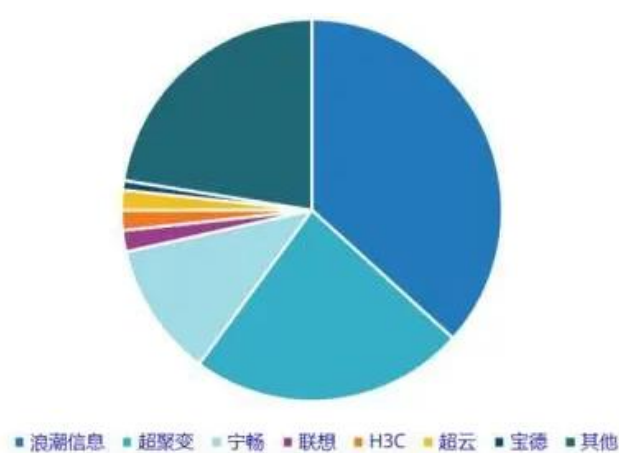
截止 2023 年液冷服务器前 3 大厂商分别为浪潮信息、超聚变、宁畅，市场份额集中度较高。

图 32：2022 年我国液冷服务器市场厂商份额



资料来源：IDC，中原证券研究所

图 33：2023 年我国液冷服务器市场厂商份额

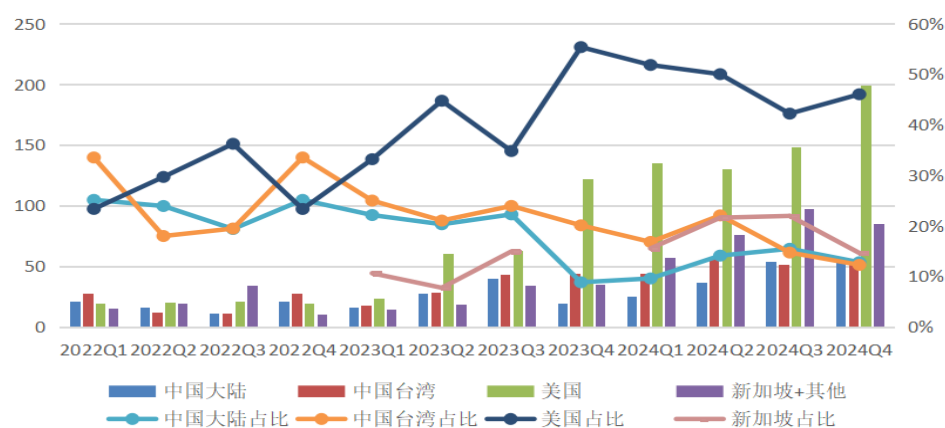


资料来源：IDC，中原证券研究所

4.5.3. 国产 AI 芯片迎来持续突破

随着美国对中国科技企业的制裁延续，制裁的核心逐步集中在芯片领域，特别是海外 AI 芯片采购和国产 AI 芯片的生产制造。2022 年 8 月，美国禁止了英伟达 A100 和即将推出的 H100 向中国和俄罗斯的所有出口实施一项新的许可要求。2022 年 10 月 7 日，美国商务部工业和安全局（BIS）正式宣布了其对中国超算和半导体制造物项实施的出口管制新规，全面加大了对超算和芯片制造领域的限制措施，对我国务高端芯片领域供给产生了全面而深远的影响。2023 年 10 月 17 日，BIS 发布了针对芯片的出口禁令新规，禁运 AI 芯片进一步扩大到英伟达 A100、H100、A800、H800、L40、L40S、RTX 4090 以及集成这些高性能计算卡的 DGX/HGX 系统。2024 年 12 月，BIS 发布了《出口管理条例(EAR)》的修订说明，加强了对 HBM、华为产业链、先进制造、国产设备厂商、EDA 的出口限制。2025 年 3 月，BIS 又将浪潮集团子公司及宁畅等企业列入实体清单，加大了对服务器厂商的制裁。此外，美国还加大对新加坡、马来西亚等地的 AI 芯片监管，以防这些产品流向中国。

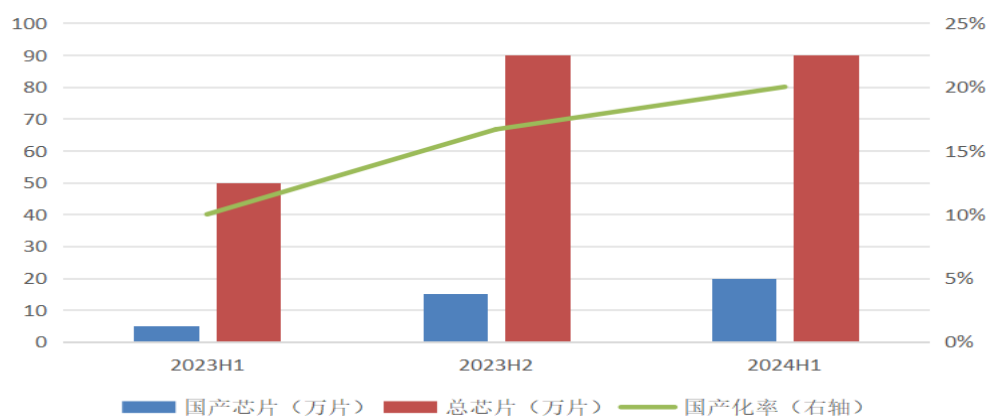
图 34：2022Q1 至 2024Q4 英伟达分区域收入（亿美元）及占比



资料来源：英伟达，中原证券研究所（英伟达财报时段较以上周期晚一个月）

受制于芯片供应，当前国内 AI 芯片的国产化率持续提升，但是在产业格局上仍有较大的不确定性。根据 IDC 数据，国内 AI 服务器市场中，国产化芯片比例持续提升，从 2023 年上半年的 10%，提升到下半年的 17%，2024 年上半年这一比例又进一步提升到了 20%，这意味着国产芯片在持续发展壮大。但是与此同时，美国从 2024 年 11 月以来，对我国 AI 芯片限制进一步加紧，增加了未来的国内芯片供应上的不确定性。

图 35：2023H1-2024H1 我国 AI 服务器芯片国产化占比



资料来源：IDC，中原证券研究所

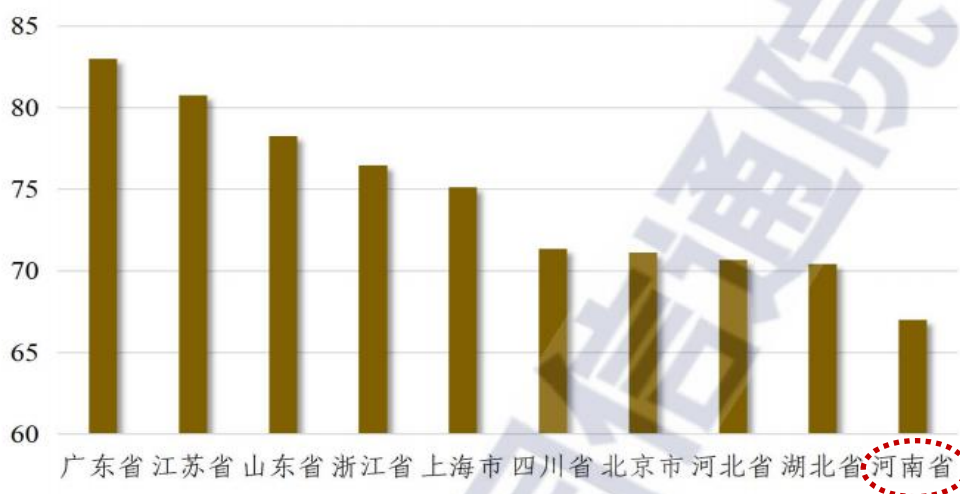
随着 DeepSeek-R1 的开源，我们看到国产模型找到了在算力不足的情况下，通过软件弥补的可行路径。同时，各家国产 AI 芯片都在积极适配 DeepSeek 的部署，进一步加快了国产芯片在推理端应用的进程。此外，2025 年 3 月，蚂蚁集团成功利用阿里巴巴、华为等国产芯片训练出了 2900 亿参数的大模型“Ling-Plus”，其性能与英伟达 H800 芯片训练的同类模型相当，且训练成本降低 20%，成为了国产芯片在 AI 训练领域又一重要突破的标志性事件。

5. 河南积极布局算力产业，力争进入全国第一梯队

5.1. 河南运载力较为领先，排名全国第 10

从算力、运载力、存储力三个方面来看，河南在运载力方面实力最强，在省级行政区中排名第 10。

图 36：省级行政区运力分指数



资料来源：《综合算力评价研究报告（2024）》，中原证券研究所

在运力的三大细分项中，河南在算间网络排名第 6，在入算网络排名第 11 位，而在算内网络方面相对落后。郑州成为了国内第 9 个简称的国家级互联网骨干直联点，到 2020 年底，郑州国家级互联网骨干直联点总带宽达到 1360G，居全国第 3 位，河南互联网网内、网间平均时延均居全国第 2 位。根据《综合算力评价研究报告（2024）》，河南到算力枢纽的光层平均时延在 7.52ms，光层时延最大值 12.7ms，都排名全国第 1。同时郑州有望在 2025 年完成国家超算互联网核心节点的建设，有望进一步提升算力传输能力。

表 12：运力分指数体系

细分项	河南排名	细分指标
算间网络	6	国家级互联网骨干直联点数、省际出口带宽、单位面积长途光缆长度、省内高速光传输网络端口数、重点站点全光交换（OXC）部署率、城域出口 IP 带宽、省内高速 IP 路由器网络端口数
入算网络	11	互联网带宽接入端口数、单位面积接入网钢缆线路长度、千兆光网覆盖率、综合接入节点 OTN 覆盖率、综合接入节点 IP 专线覆盖率、高速 IP 专线用户数、互联网专线用户数、固定带宽平均下载速度、移动带宽平均下载速度
算内网络		算力中心网络出口带宽、算力中心单机架带宽、算力中心省级骨干网络接入、算力中心城域网接入、算力中心出口光纤路由可靠性、算内高性能网络技术应用、省内智能化网络技术应用

资料来源：2024 中国算力大会，中原证券研究所

5.2. 河南在用算力位列全国前 10，算力利用率较高

根据信通院的《综合算力评价研究报告（2024）》和《中国综合算力指数（2024 年）解读》，河南在算力整体实力上未能进入全国前十，但是在在用算力规模、在建算力规模、算力质效（特别是上架率）方面相对领先。

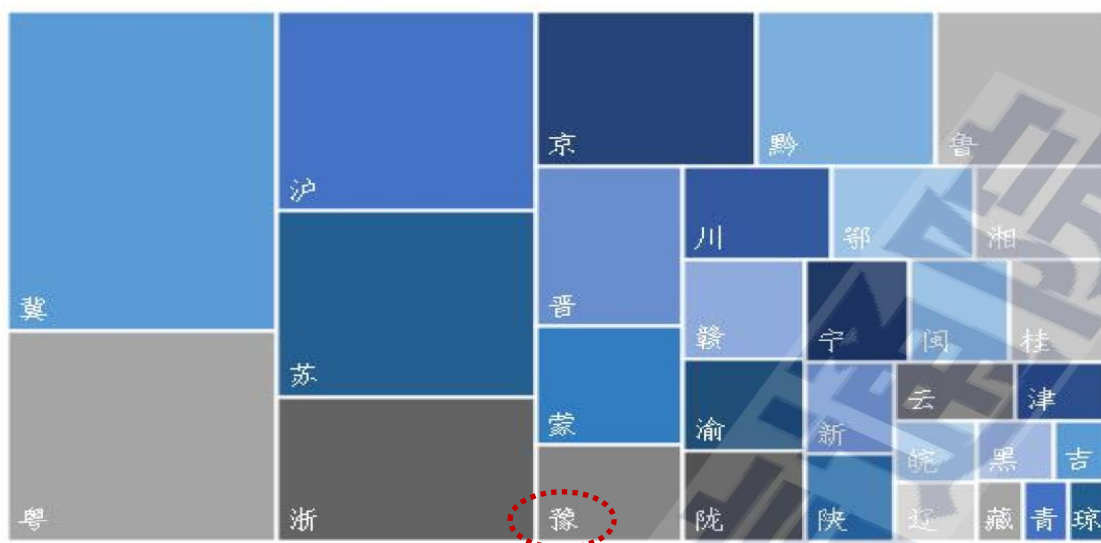
图 37：算力分指标体系



资料来源：《综合算力评价研究报告（2024）》和《中国综合算力指数（2024 年）解读》，中原证券研究所

在用算力方面，河南省位于全国第 11 名，位居河北、广东、上海、江苏、浙江、北京、贵州、山东、山西、内蒙古之后，排名较为靠前。位于河南前列的省级区域，仅有山西不是位于 8 大枢纽节点区域。

图 38：我国在用算力规模分布



资料来源：综合算力评价研究报告（2024），中原证券研究所

在建算力方面，河南同样位于全国第 11，位列河北、山西、广东、江苏、上海、湖南、宁夏、浙江、湖北、北京之后，这其中包括山西、湖南、湖北都是非 8 大枢纽节点区域。

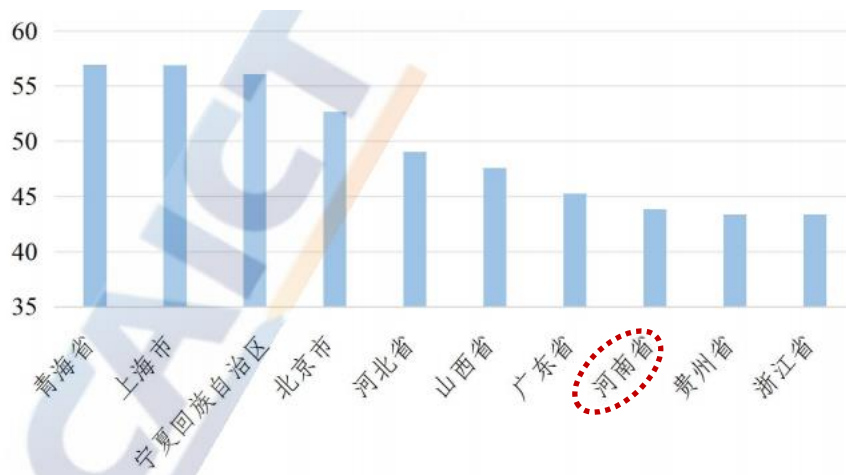
图 39：我国在建算力规模分布



资料来源：综合算力评价研究报告（2024），中原证券研究所

算力质效方面，河南跻身全国第 8，通过积极推动算力业务发展，加快数据信息产业落地，有效地提升了算力质效水平。其中在上架率方面，河南高达 70%以上，仅次于山西、河北、青海、宁夏、海南，位列全国第 6；在 PUE 方面，河南未能进入前十，优势并不明显。

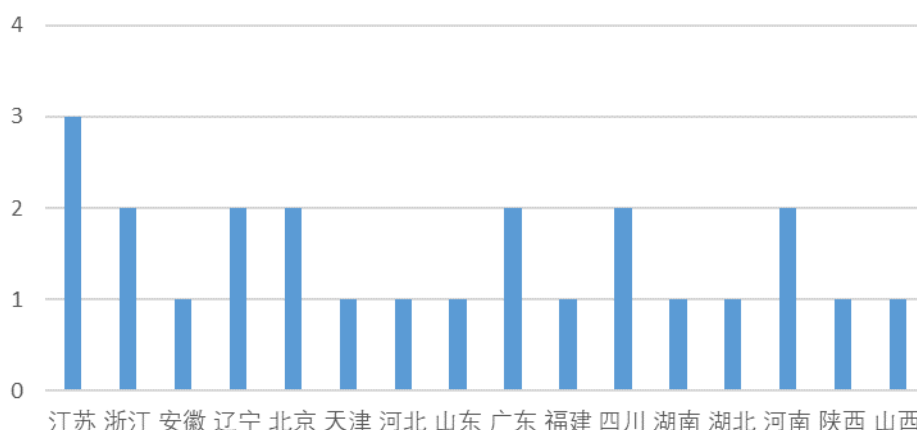
图 40：省级区域算力质效水平对比



资料来源：综合算力评价研究报告（2024），中原证券研究所

此外，河南在 2023 年选出的 25 个国家新一代人工智能公共算力开放创新平台中占据了两席，将对数据中心建设和利用带来积极的影响。

图 41：国家新一代人工智能公共算力开放创新平台的区域分布情况（个）



资料来源：科技部，中原证券

5.3. 随着 AI 大发展，河南省政策对于算力和数据中心建设热度持续提升

通过梳理河南省算力产业相关政策我们可以看到：

（1）从 2021 年 4 月出台的《河南省推进新型基础设施建设行动计划（2021-2023 年）》和 2022 年 2 月出台的《河南省“十四五”新型基础设施建设规划》，我们可以看到在此前的新基建中，政府对于信息技术相关新基建的重点更加偏向网络建设。而后在人工智能大发展和算力建设加快的大背景下，2024 年 11 月省政府又出台的《河南省算力基础设施发展规划（2024—2026 年）》，这是河南的首个算力基础设施的专项发展规划，表明新基建是重心进一步向数据中心偏移。

（2）通过对比河南省历年政府工作报告，我们看到省政府算力产业布局重心方面，2023 年更多在上游制造端，2024 年向基建端偏移，2025 年进一步向应用端偏移，对算力产业的重视方向进一步向下游延伸。

表 13：河南省算力产业相关政策

时间	部门	政策	简介
2020.4.3	河南省发改委、工信厅	《河南省鲲鹏计算产业发展规划》	到 2020 年，形成年产“Huanghe”服务器 35 万台、PC 机 75 万台、平板电脑 20 万台的生产能力；软硬件产品和服务在政务系统全面推广，鲲鹏产品适配认证平台和标准化认证体系初步建成；培养 2000 名鲲鹏技术应用型人才，校企共建 10 所鲲鹏产业学院，完成 100 家软件开发业务产品的适配测试。 2022 年，形成年产“Huanghe”服务器 50 万台、PC 机 100 万台、平板电脑 50 万台的生产能力，实现重点行业领域应用全覆盖，集聚一批产业链上下游企业，初步建成国内重要的鲲鹏软件研发集聚地，力争品牌影响力进入全国第一方阵。 2025 年，“Huanghe”成为我国信息技术自主可控知名品牌，培育 2~3 家行业领军企业，以研发设计、生产制造、创新集成、人才培养为核心的鲲鹏产业生态不断完善，形成具有国际影响力的千亿级鲲鹏计算产业集群。
2021.3.23	河南省发改委	《2021 年河南省数	推进许昌鲲鹏硬件生产基地建设，拓展产品应用市场，培育壮大“Huanghe”

		字经济发展工作方案》	品牌，推动黄河鲲鹏生态政务应用。加快建设中原鲲鹏生态创新中心，完善中原鲲鹏生态创新中心鲲鹏产品适配认证平台和标准化认证体系，提供重点行业系统化应用鲲鹏解决方案，完成 400 个左右软件产品适配认证。推动企业和高校联动合作，完善鲲鹏人才认证体系，培养鲲鹏产业生态人才 3000 人以上。加快中国长城（郑州）自主创新基地、紫光智慧计算终端全球总部基地等重大项目建设。 推动河南讯飞、寒武纪、中原动力等企业研发落地一批人工智能项目，拓展智能工业检测、智能机器人、智能语音应用市场。结合新型智慧城市试点市建设，打造一批“智能+”应用场景。支持郑州市争创国家新一代人工智能创新发展试验区。
2021.4.13	河南省政府	《河南省国民经济和社会发展第十四个五年规划和二〇三五年远景目标纲要》	做大做强电子制造产业，提升智能终端、计算机和服务器、汽车电子、光电子芯片等整机制造和零部件加工水平。 实施鲲鹏计算产业生态系列项目
2021.4.17	河南省政府	《河南省推进新型基础设施建设行动计划（2021-2023 年）》	统筹布局建设数据中心，重点支持郑州、洛阳市加快建设大型绿色数据中心，积极争取全国一体化大数据中心国家级枢纽节点在我省布局。加快实施省大数据中心项目。提升国家超级计算郑州中心效能，围绕精准医学、生物育种、环境治理等领域培育一批超算应用。实施大型数据中心绿色化改造，推动现有数据中心改造后电能利用效率不高于 1.4，新建数据中心电能利用效率不高于 1.3。
2021.12.7	科技部	《科技部关于支持郑州建设国家新一代人工智能创新发展试验区的函》	开展智能物流、智能制造、智慧农业等领域技术集成和创新应用，培育智能传感器、智能网联汽车等产业集群
2021.12.28	河南省第十三届人民代表大会常务委员会第二十九次会议	《河南省数字经济促进条例》	省人民政府及有关部门应当统筹推进数据中心、智能计算中心、超级计算中心等基础设施建设；支持数据中心优化建设和升级改造，提升资源利用水平和运行效率；推动云计算、边缘计算等多元计算协同发展，构建高效协同的数据处理体系。设区的市、县级人民政府可以根据需要，按照国家产业政策统筹建设本区域数据中心、边缘计算中心等算力基础设施，在土地、电力保障等方面给予支持。
2022.2.9	河南发改委	《河南省“十四五”新型基础设施建设规划》	加快推进“中原智谷”建设，搭建中原昇腾人工智能生态创新中心等创新平台，集聚人工智能龙头企业和领军人才团队，建设国家新一代人工智能创新发展试验区。 加快推动总规模 300P 算力的中原人工智能计算中心建设，搭建省级算力共享服务平台，打造人工智能共享数据集，为全省人工智能创新提供算力支撑。
2022.2.16	河南省政府	《河南省“十四五”数字经济和信息化发展规划》	黄河鲲鹏计算产业。郑州中原鲲鹏生态创新中心、许昌鲲鹏制造基地、新乡鲲鹏软件园快速发展，许昌制造基地已具备年产“Huanghe”服务器 36 万台、PC 机 75 万台、主板 25 万片的能力，成为华为鲲鹏国内重要生产基地。 优化升级 5G、千兆光纤、移动互联网、卫星互联网等通信网络基础设施，统筹布局以数据中心、边缘计算中心、人工智能计算中心为核心的算力基础设施和新技术设施，加快推进传统基础设施智能化升级，前瞻布局创新基础设施。
2022.3.19	河南省发改委	《2022 年河南省数字经济发展工作方案》	全省服务器、PC 机产能分别达到 150 万台、300 万台。 开展“双千兆”建设，实施一批千兆宽带应用试点，推动 5G 网络由规模建设、广泛覆盖向按需建设、深度覆盖发展，新建 5G 基站 4 万个。扩容提升郑州国家级互联网骨干直联点等关键枢纽设施，规划建设国家新型互联网交换中心。 统筹布局数据中心、智能计算中心、边缘计算中心，加快实施中原大数据中心二期、中国移动网络云郑州大区中心、郑州人工智能计算中心等项目。加快自

			主可控、全栈国产化的“黄河云”建设，推动省内重点国有企业、中小企业上云。提升 黄河、紫光、曙光、浪潮、长城 等重点企业产业化能力。
2022.5.19	郑州市政府	《郑州国家新一代人工智能创新发展试验区建设实施方案》	到 2025 年，人工智能治理体系初步建立，人工智能创新能力、产业发展、与实体经济发展深度融合等方面达到国内一流水平，推动人工智能成为城市转型发展的重要引领力量。
2022.9.21	河南省政府办公厅	《河南省大数据产业发展行动计划（2022—2025 年）》	积极融入国家“东数西算”工程建设，构建全省一体化大数据中心协同创新体系，推动郑州、洛阳培育超大型绿色数据中心集群，支持相关地方适度建设新型数据中心，打造“两核多点”发展格局。加快推进国家（郑州）数据枢纽港等重点项目建设，争取国家工业互联网大数据分中心在我省布局。聚焦 VR/AR（虚拟现实/增强现实）、5G、车联网、智能制造等领域，建设一批边缘数据中心。到 2025 年，全省数据中心机架数达到 15 万，新建大型、超大型数据中心电能使用效率降至 1.3 以下。
2023.1.14	河南省第十四届人民代表大会第一次会议	《2023 年河南省政府工作报告》	2022 年 超聚变服务器产值突破 230 亿元 ，郑州市获批建设国家区块链发展先导区，5G 基站总数超过 15 万个。新增智能工厂 185 个、上云企业 3.6 万家。数字经济对经济增长的贡献率超过 70%。 提出“着力数字化绿色化转型”。 延伸电子信息产业链条，加快打造万亿级新一代信息技术产业高地。到 2025 年形成传统产业为基础、新兴产业为支柱、未来产业为先导的先进制造业体系，制造业增加值占生产总值比重超过 30%，战略性新兴产业占规模以上工业增加值比重超过 30%。
2023.2.1	河南省数字经济发展领导小组	《2023 年河南省数字经济发展工作方案》	新型基础设施建设完成年度投资 500 亿元，打造 5G 精品网络，5G 基站总数突破 18 万个，数据中心机架数超 9 万架，全省物联网终端用户突破 1 亿户。电子信息制造业营业收入力争突破 8000 亿元，先进计算、软件产业规模均超过 500 亿元，新一代人工智能、网络安全、智能传感器等产业集群能级显著提升。 数字技术和实体经济深度融合，全省工业云平台应用率达到 52%，农业生产信息化发展水平超过 30%。 深化“一件事一次办”“跨省通办”，完成 30 项以上改革事项“一件事一次办”，新增 50 项“跨省通办”事项、200 项“全豫通办”事项，建成一批智慧县城、数字乡村试点示范项目。
2023 年底		《河南省新一代人工智能产业链培育行动方案(2023—2025 年)》	在全面梳理全省人工智能产业链现状、关键领域和薄弱环节的基础上，明确了推动算法创新发展、突破发展智能硬件、提升软件支撑能力、拓展行业应用场景 4 个主攻方向。 到 2025 年，全省人工智能技术创新能力和产业竞争力显著提升，部分关键技术、算法研究和产品装备达到国内先进水平，产业规模突破 1000 亿元，有效赋能经济社会数字化转型，成为全国具有影响力的人工智能产业高地和创新应用示范区。 在供给端，引进超聚变，培育黄河鲲鹏，吸引浪潮、长城、曙光等头部企业入驻河南。
2024.1.28	河南省第十四届人民代表大会第二次会议	《2024 年河南省政府工作报告》	2023 年 国家超算互联网核心节点项目启动 ，龙芯中科中原总部基地建成投用。加快新型基础设施建设，力争 5G 基站达到 21.6 万个、数据中心标准机架数达到 20 万架，新建 10 个左右省级工业互联网平台。
2024.3.15	河南省数字经济发展领导小组	《2024 年河南省数字经济发展工作方案》	2024 年，全省数字经济持续快速发展，增速保持在 10% 以上，总量规模和发展质量稳步提升。5G 基站总数达到 21.6 万个，数据中心标准机架数达到 20 万架。
2024.4.26	河南省制造强省	《2024 年河南省数	数字基础设施进一步升级。5G 基站总数突破 21.6 万个，5G 网络覆盖广度深

	建设领导小组办公室	数字化转型战略工作方案》	度持续拓展，争创 5G 行业虚拟专网先导区。郑州互联网骨干直联点总带宽扩容至 3600G，全国重要的信息通信枢纽和信息集散中心地位持续巩固。推动郑州、洛阳等加快建设超大型绿色数据中心集群。加快中国移动（河南）数据中心、中国联通中原数据基地等超大型数据中心建设，服务器机架数力争达到 20 万架。开展数据中心节能监察、诊断服务和能效对标工作，培育一批省级绿色数据中心。统筹布局政务云计算中心，持续推动老旧算力设施转型升级。推进国家新一代人工智能公共算力开放创新平台——“嵩山”人工智能公共算力开放创新平台一期建设，谋划二期 300PFP16 算力资源与软件生态建设。推进超算创新生态在气象环保、精准医学、国土资源、生物育种等多领域的应用示范落地，建强国家超算平台。
2024.10.29	河南省政府办公厅	《河南省推动“人工智能+”行动计划（2024-2026 年）》	制定算力基础设施发展规划，统筹布局智能算力中心，加快国家超算互联网核心节点工程、郑州人工智能计算中心等建设，推进人工智能公共算力开放创新平台建设。谋划建设省级算力调度平台，推动算力一网化、统筹一体化、调度一站式，为创新主体提供泛在便捷、优质普惠的算力支持。加强“算力券”等政策工具引导，进一步降低算力使用成本。 成立省人工智能行业应用工作专班。
2024.11.14	河南省政府办公厅	《河南省算力基础设施发展规划（2024—2026 年）》	到 2026 年年底，形成布局合理、绿色低碳、高效集约、安全可靠的算力基础设施发展格局， 全省算力基础设施标准机架数达到 35 万架，平均利用率达到 70% 以上，算力规模超过 120EFLOPS，智算、超算等高性能算力占比超过 90% ，形成 10 个以上算力规模达到 E 级的大型算力中心。建成投用一批多元算力资源有机协同的融合算力中心，在智能制造、智慧交通、智慧农业、智慧医疗等领域形成一批算力赋能的典型应用场景，算力服务普惠易用水平显著提升。
2025.1.18	河南省第十四届人民代表大会第三次会议	《2025 年河南省政府工作报告》	2024 年超聚变国产化服务器销售额居全国首位。 国家数据要素综合试验区获批建设，河南空港智算中心揭牌投运。 2025 年深入推进“人工智能+”行动，打造数字化宏大应用场景，实现规上制造企业智能应用场景全覆盖，赋能传统产业焕新升级、新兴产业发展壮大。提升省算力调度服务平台功能，构建全省一体化算力网，推进全省信创云网建设，开展公共数据授权运营，做强郑州数据交易中心。推动平台经济健康发展。加快 5G-A 商用部署，争创国家“5G+工业互联网”融合应用先导区。加快郑州、鹤壁、新乡中小企业数字化转型试点城市建设。
2025.3.24	河南省人工智能行业应用工作专班办公室	《河南省 2025 年推进“人工智能+”行动工作要点》	建成国家超算互联网核心节点，推进河南空港智算中心、中原算谷等智算中心建设。强化“嵩山”“中原”人工智能公共算力开放创新平台功能，提升一站式人工智能算力服务能力。到 2025 年底，全省算力规模突破 94EFLOPS ， 进入全国第一梯队。

资料来源：政府网站，中原证券研究所

（3）我们将河南省在算力建设方面发展目标做对比，可以看到算力产业推进有明显加快的趋势。以算力中心机架数指标来看，在 2021 年的数字经济十四五规划中，河南省提出算力中心总规模到 2025 年 15 万架的目标，政府工作报告 2024 年目标已经提升到 20 万架，又在 2024 年 11 月提出了 2026 年建设 35 万架的发展目标。另外作为算网结合的重要指标，河南在 2021 年 12 月的数字经济十四五规划中提出互联网骨干直联点总带宽 2000G 的目标，2022 年 1 月进一步提高到 2400G，2024 年 3 月提出 2024 年将该指标提高到 3000G，2024 年 4 月又将当年目标提升到 3600G，并在 11 月提出 2026 年达到 4000G。同时，政策中对于数据中心的建设指标也进一步增加，全面细化了需要完成的目标和政策导向。

表 14：河南省算力相关目标

类别	指标	单位	2020 年	2023 年	2024	2025 年	2026 年
计算力	算力规模	EFLOPS				94 (2025)	120 (2024.11)
	算力中心总规模 (2.5kW 标准机架)	万架	3	9 (2023)	20 (2024.1)	15 (2021)	35 (2024.11)
	智算规模						100 (2024.11)
	E 级以上算力中心	个					10 (2024.11)
	高性能算力占比	%					90 (2024.11)
运载力	5G 基站总数	万个	4.5	16.8 (2021)	21.6 (2024.1)	20 (2022.1) 18 (2022.9)	26 (2024.11)
	新建 5G 基站	万个			2.9 (2024.3)		
	互联网骨干直连点总带宽	G	1360		3000 (2024.3) 3600 (2024.4)	2000 (2021.12) 2400 (2022.1)	4000 (2024.11)
	5G 用户普及率	%	19.7			60	
	光纤入户用户占宽带用户的比重	%	96.6			99	
	互联网省际出口带宽	G	52800			130000	
	月户均移动数据流量	G	7.3			35	
	物联网终端连接数	万个	6656			30000 (2022)	
	新增 10G-PON 及以上端口	万个			10	110 (2022)	
	省际出口	个				28 (2022)	
	千兆带宽用户数	万户				900 (2022)	
	移动网络 IPv6 占比	%				70 (2022)	
	存储总量	EB					5 (2024.11)
应用赋能	算力应用推广案例	个					100 (2024.11)
	算力中心平均利用率	%					70 (2024.11)
	工业互联网标识解析二级节点	个	3			10 (2022)	
	省级工业互联网平台	个	25			100 (2022)	
	中小学智慧校园覆盖率	%	-			30 (2022)	
	数字化医院	个	13			30 (2022)	
	人工智能产业规模	亿				1000 (2023)	
绿色安全	算力中心平均 PUE (电源使用效率)						<1.3 (2024.11)
	大型、超大型数据中心 PUE					<1.3 (2022)	
	算力中心国产芯片使用占比	%					60 (2024.11)

资料来源：河南省政府，中原证券研究所（括号中为目标提出时间）

5.4. 河南推出积极的算力规划，致力于 2025 年算力规模步入全国第一梯队

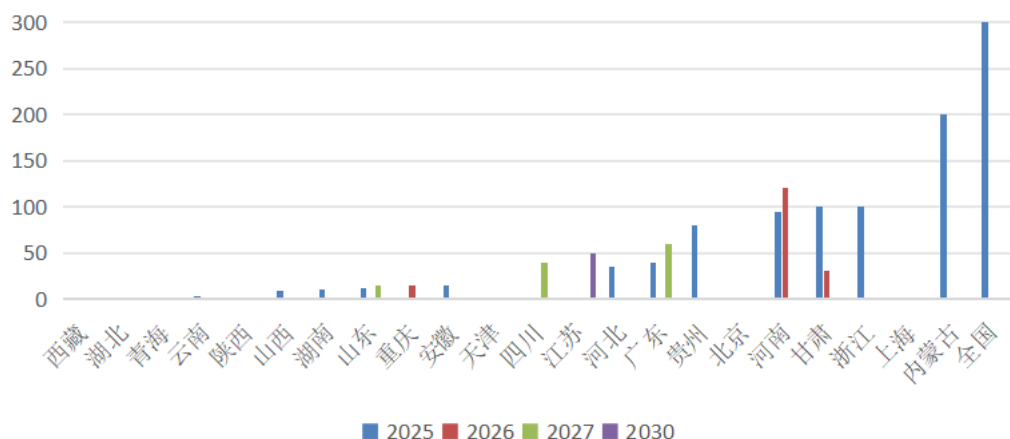
2024 年 11 月河南省出台了《河南省算力基础设施发展规划（2024—2026 年）》，并提出了在 2026 年全省算力规模超过 120EFLOPS、智算规模达到 100EFlops 以上的发展目标，同时规划智算、超算等高性能算力占比超过 90%。

2025 年 3 月河南省出台了《河南省 2025 年推进“人工智能+”行动工作要点》，提出到 2025 年底，全省算力规模突破 94EFLOPS，进入全国第一梯队。

对比全国来看，河南目前不论总算力、智算算力目标都名列前茅，也意图将算力和人工智能作为新质生产力来带动河南的发展。

从 2025 年算力目标来看，河南的 94EFLOPS 目前仅次于位于内蒙古 200EFLOPS、上海 100EFLOPS（智算）、甘肃（庆阳）100EFLOPS、浙江 100EFLOPS，位列全国第 5。

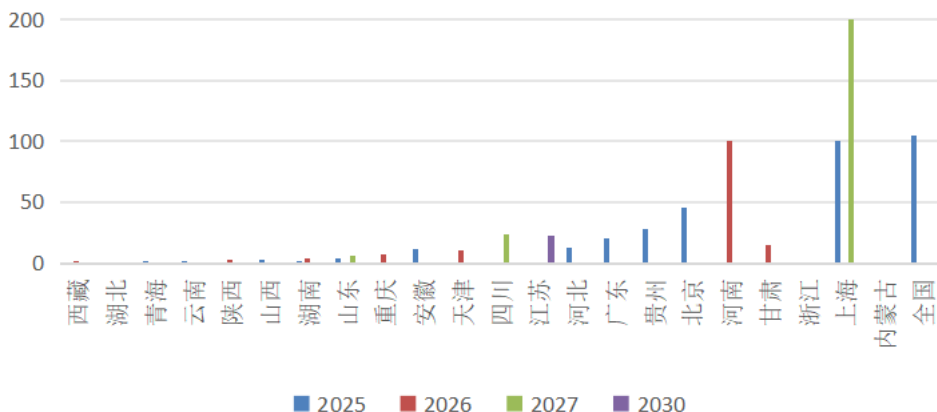
图 42：各地算力规划中关于总算力的相关目标（EFLOPS）



资料来源：政府网站，中原证券研究所

从智算目标来看，目前有明确目标的地区中，河南 2026 年 100EFLOPS 目标，仅次于上海（2025 年 100EFLOPS 和 2027 年 200EFLOPS）。

图 43：各地算力规划中关于智能算力的相关目标（EFLOPS）



资料来源：政府网站，中原证券研究所

5.5. 河南的算力区域架构及重点数据中心

5.5.1. “一核四极多点”的算力区域架构

河南在算力规划方面整体采用的是以郑州市（含航空港区）为核心，洛阳、鹤壁、商丘、信阳等市为区域增长极，各地多点配套发展的“一核四极多点”总体布局。

1. “一核”：结合国家新一代人工智能创新发展试验区、国家数据基础设施试点建设，依托郑州网络枢纽地位和算力资源，以算力网引领发展，加快智算中心布局，构建以国家超算互联网核心节点为主的算力中心集群，把郑州市（含航空港区）打造成为保障省内、辐射中部地区乃至全国的综合性多功能算力枢纽中心。

2. “四极”：支持洛阳、鹤壁、商丘、信阳等网络基础好、数字经济发展快、区位优势明显的城市打造省内区域级算力增长极，充分利用郑州市（含航空港区）和当地算力资源，面向省内周边区域和周边省份临近区域开展算力服务和算力调度。

洛阳：依托制造业、网络优势和较好算力基础，建设成为省内算力副中心和豫西区域算力增长极，辐射服务晋南、陕东等邻近省份区域。

鹤壁：依托数字经济发展基础和能源比较优势，建设成为豫北区域算力增长极，辐射服务冀南等邻近省份区域。

商丘：依托枢纽区位优势，建设成为豫东区域算力增长极，辐射服务鲁西南、皖北等邻近省份区域。

信阳：依托电子信息产业、云灾备建设基础和绿色能源优势，建设成为豫南区域算力增长极，辐射服务鄂北、皖西南等省份区域。

3. “多点”：支持有需求、有条件的地方部署边缘计算中心，打造城市计算节点。

图 44：河南算力产业的区域布局



资料来源：政府网站，中原证券研究所

2025年3月，许昌市发布了《许昌市加快推动人工智能产业发展实施方案(2025-2030年)》，提出了到2030年算力规模突破2000P、人工智能核心产业及相关产业规模突破500亿元、位居全省人工智能产业规模及产业竞争力第一方阵的发展目标。

5.5.2. 河南主要数据中心项目的分布及建设情况

河南省在2025和2026年的算力建设将全面加速，结合各区域的数据中心布局情况来看，河南绝大部分算力都集中在郑州。

河南空港智算中心：位于郑州航空港区豫发锦荣信息科技园，总算力规划10万P，将成为中部最大万卡算力集群。一期规划算力30000P，首批2000P已经于2024年11月投产，第二阶段10000P算力集群项目预计于2025年一季度全面投运。截止2025年3月，已吸引了新华三、科大讯飞、腾讯云等40余家创新企业入驻。

郑州中原空港行业智算中心：位于郑州航空港区，总建筑面积1.4万平方米，主要建设航空港区行业智算中心及其他配套设施等，规划建设算力约5000P，计划在2025年开工建设。其中一期1000P、二期4000P。

郑州人工智能计算中心：位于郑州航空港区，2024年3月举行揭牌仪式。作为中部地区首个由政府主导、**联通**投资建设运营的智算算力集群，总体规划集群规模1000P，2600个机柜，总建筑面积4.6万平方米。2024年7月，一期点亮400P。郑州是中国联通全国25个AIDC重点发展的城市节点之一，目前奇虎360的智脑大模型、奇安信のQ-GPT大模型、阿帕斯的天燕大模型等已经在中国联通中原数据基地投用。

中国电信中部数据中心：位于郑州航空港，总建筑面积18万平方米，总投资约34亿元，可容纳约2.8万个机柜。一期于2021年10月开工，2023年10月竣工验收。

中国移动（河南郑州）航空港区数据中心：位于郑州市航空港区，2016年3月正式开工建设，2017年年底一期工程全部建成投运，运行着超过4万台设备，二期2.1万架业务机架、15万台服务器。网络云承载河南、湖北、湖南三省核心网络业务，曾被工信部等六部委评为“国家绿色数据中心”。

中原算谷（中原大数据中心）：位于郑州郑东开发区，由**豫信电科旗下河南数字中原数据有限公司**于2020年2月启动建设，并于2024年5月揭牌，是河南省最大的第三方国资数据中心。项目占地74.5亩，建筑面积10.2万平方米，总投资超百亿元，规划算力超万P，整体规划机柜超10000架。目前一期、二期已经建设完成并合计交付机柜6000个，已经建成算力1000P，采取“自投+合作”模式，先后落地人工智能视觉渲染算力资源池和白沙智算中心两个示范性项目，项目三期已经启动，2025年底项目全部建成。项目通过了数据中心基础设施运维评价增强级（L3级），达到国家“绿色数据中心”标准。

郑州高新智算基地：位于郑州高新区主建成区内，规划 **200P** 算力规模，一期已建成投用 160P 算力，同时配备通用算力、高性能存储系统和网络系统，集算力生产供应、数据开放共享、智慧生态建设和产业创新聚集四大功能于一体。

360 智算中心：落地郑州高新区中原数据基地，全部由 **360** 公司自身投入。项目总投资规模约在 10 亿元，总计超过 **5 万台服务器**，**2000P** 的智能算力。一期 2024 年 2 月建成，正常投产运行。安全大模型已经落地存储服务器 30000 台，智能算力 1000P。

中国移动（河南郑州）数据中心项目：位于郑州市高新区，总建筑面积 12.6 万平方米，容纳服务器机架 **1.69 万架**、**15 万台服务器**，项目于 2016 年 3 月开工，总计投资 45 亿元，计划在 2025 年建设完成。

信阳人工智能计算中心：位于信阳市羊山新区，为信阳市政府与**中国电信**河南分公司联合打造项目，2023 年 10 月开工建设，项目建设金额约为 17 亿元。总建筑面积 9.46 万平方米，数据中心总共规划建设 **4000 个机架**用于建设云数据中心和智算中心，计划在 2025 年竣工。

中国移动（洛阳）数据中心：位于洛阳洛龙区，该项目一期于 2016 年建成投用，建筑面积 8300 平方米，拥有 **899 个业务机柜**，装机容量 **1.2 万台服务器**。项目二期于 2022 年 7 月完工，总建筑面积 2.22 万平方米建设机柜 **4000 架**预计装机容量达到 **5 万台服务器**，建成后成为豫西地区规模最大的大数据基础设施。

中科星云（鹤壁）城市算力中心：中科星云（鹤壁）产业园项目的一个组成部分（还包括数据方舱生产基地、设计创新中试基地、人工智能研究院等），总投资 9.5 亿元。

京东（鹤壁）算力中心：位于鹤壁市示范区京东智能制造产业园，由京东云和鹤壁市政府合作共建，是京东私有云在中原地区的唯一运营中心，于 2024 年 9 月建成。

京东云（华中）数据中心：位于商丘市宁陵县，租赁标准化厂房建筑面积 12.6 万平方米，算力 **340P**，主要建设厂房、信息平台服务中心等，计划在 2025 年开工建设。

商丘市梁园算力中心：位于商丘市梁园区，总建筑面积 10 万平方米，算力 **5000P**，主要包括算力机楼、互联网数据中心楼、数据储存楼及配套楼。

中原人工智能计算中心：位于许昌市，是基于华为鲲鹏+昇腾生态建设的中原首个大规模人工智能算力集群项目，也是继深圳和武汉之后全国第三家建成并投用的人工智能计算中心，为完善河南省黄河鲲鹏计算产业生态而提前谋划布局的“智算”数字基建核心底座，2023 年入选国家新一代人工智能公共算力开放创新平台（筹）。2021 年 10 月建成投用，一期建设 100P 人工智能算力，已建成投用，使用率超过 90%，二期建设 300P 人工智能算力，目前累

计投资约 5 亿元，现阶段人工智能算力规模 150P，在华为全国布局建设的 28 个人工智能计算中心中，中原人工智能计算中心综合实力全国第三，是河南省数据要素入表第一家。

5.5.3. 超算领域，河南拥有国家超算互联网核心节点，地位特殊

河南省在超算方面的建设成就主要包括：

国家超级计算郑州中心（河南省超级计算中心），位于郑州高新区郑州大学国家大学科技园内，是科技部于 2020 年 11 月批复同意通过验收并正式纳入序列管理的全国第七家国家超级计算中心，也是河南省首个国家级重大科研基础设施，由郑州大学负责建设、管理、运行和服务工作。配备了主机“嵩山”超级计算机，理论峰值计算能力 100PFLOPS，存储容量 100PB，采用绿色节能的浸没式相变液冷冷却技术，PUE 值低于 1.04。

2023 年，国家超级计算郑州中心获科技部批复筹建国家新一代人工智能公共算力开放创新平台（筹）——“嵩山”人工智能公共算力开放创新平台。这一平台将通过大数据、云计算等技术实现资源池的共享和优化利用，推动“超算+智算”融合式发展。

国家超算互联网核心节点：随着人工智能等技术的快速发展，全社会对算力提出更高要求，算力中心亟需突破现有单体运营模式。为此，2023 年 4 月科技部高新司在天津组织召开国家超算互联网工作启动会，并发起成立国家超算互联网联合体。2023 年 10 月，国家超算互联网核心节点在河南郑州正式启动建设。根据《2024 年河南省数字经济发展工作方案》，由河南省科学院负责建设的“国家超算互联网核心节点”项目，总计投资 80 亿元，在 2023 年 10 月开工，计划于 2025 年 12 月完工。

5.6. 河南算力产业核心企业

5.6.1. 超聚变：河南算力产业核心企业，国产服务器出货量第一

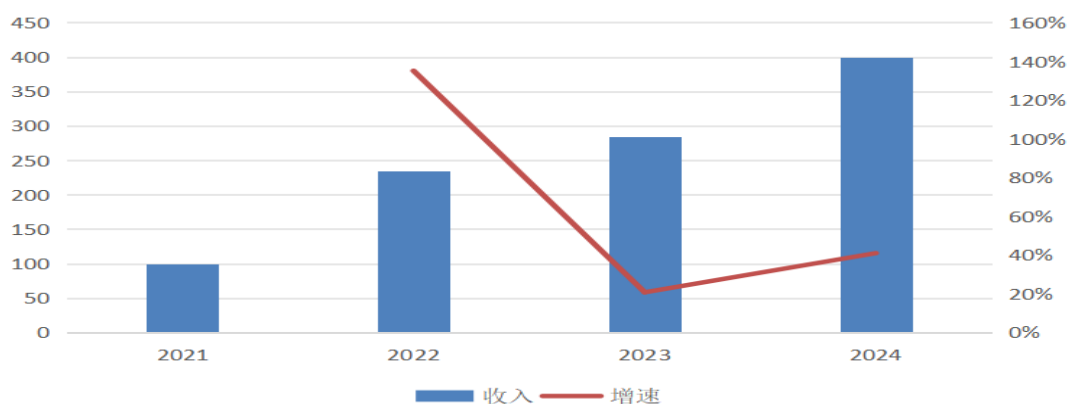
超聚变源自华为 X86 服务器业务。为了应对美国禁令的影响，2021 年 9 月成立华为全资子公司超聚变，2021 年 11 月华为完成了对超聚变的资产剥离，正式落地河南。在落户河南以后，超聚变就作为河南数字经济领域优势企业，获得河南省政府重点关注。在河南省的全力配合和支持下，超聚变也在快速成长，从选址到投产用时仅 55 天，就实现了郑州产线的首台服务器正式下线，同时 2022 年及之后业务快速恢复，重回国内服务器市场份额第二位。

超聚变总部位于河南省郑州市郑东新区，超聚变研发中心及总部基地 2025 年 2 月投入使用，集研发、生产及全球供应链管理功能，其在全球部署 9 个研发中心、6 大供应中心、6 个全球技术服务中心。

超聚变 2023 年及 2024 收入增速行业领先。2023 年超聚变实现总营收 283.8 亿元，产量 30 万台，居国内服务器市场份额第二位。对比头部服务器厂商来看，超聚变是行业头部企业

中少数实现两位数收入增长的公司。2024 年收入超过 400 亿，同比增速超过 40%。同时在外市场三年复合增速超 50%，在互联网和运营商行业三年分别获得 10 倍和 3 倍增长。

图 45：2021-2024 超聚变收入及增速（亿元）



资料来源：大河报，大河立方，超聚变，中原证券研究所

超聚变在国产服务器和液冷服务器领域优势突出。超聚变在 2023Q3 位居全球服务器市场第七，2024 年位居中国服务器市场出货量和销售额第二，AI 服务器市场位居第一。在细分赛道上，超聚变在鲲鹏加昇腾总产量第一，昇腾服务器产量第一，国产服务器赛道上优势明显；在 2023 年中国液冷市场位居第二，同时连续 2 年在标准液冷服务器赛道位居行业第一。超聚变参与了首个液冷整机柜服务器领域的行业标准制定，同时相关案例入选了工信部的“先进计算赋能新质生产力典型应用案例”。

当前行业整体在向着对超聚变有利的方向发展。（1）美科技封锁下，国产化进程加速，超聚变作为华为最大的合作伙伴最为受益。（2）AI 发展步入推理阶段，对芯片性能要求降低，有助于国产芯片推广。（3）我国智算实现了超预期增长，其中河南省规划全国领先。（4）液冷 2025 年有望步入规模化应用阶段，超聚变作为行业头部企业优势明显。

5.6.2. 豫信电科：国资代表，全方位推动河南算力产业建设

豫信电子科技集团有限公司（简称豫信电科）的前身为河南信息产业投资有限公司，成立于 2018 年 2 月。2022 年，经省政府批准，以河南信息产业投资有限公司为基础，将正数网络技术有限公司并入，整合全省信息产业资源，政府配套出资组建。

豫信电科为省管骨干企业，是河南省统一负责数字政府建设及公共数据一级开发、推动产业数智化转型、培育集成电路产业的省属功能性主体、省管重要骨干企业，注册资本 105.99 亿元，资产规模 290 亿元，在管基金总规模 137 亿元，带动数字经济相关产业投资超 600 亿元。

在数字政府建设方面，按照“一局、一中心、一集团”协同联动机制，推动全省数字政府一体化统建工作。搭建全省“一朵云”，实现省级 74 个部门的 559 个政务信息系统上云部

署，推动整合全省“1+19”政务云平台。打造“一张网”承载服务：建设河南政务服务网和豫事办 APP，助力河南网上政务服务能力提升至“非常高”水平。构建“一道墙”安全防范，推动安全技术措施与项目的同规划、同建设、同使用，上线“一道墙”安全运营支撑平台。

在电子信息制造方面，主导超聚变服务器项目落地，建设 5 条 SMT 生产代工线、7 条整机组装线，最高年产高性能服务器 60 万台。主建成立了黄河科技集团，生产“HUANGHE”牌、安装“鲲鹏”芯片和“欧拉”操作系统服务器及 PC 机。激发产业“链式”效应，发起省电子信息产业联盟，汉威科技、中电科 27 所、中航光电等 160 余家企事业单位入盟。

在网络及信息基础设施方面，建设河南最大的第三方数据中心——中原大数据中心，提升算力资源集中调度和综合服务能力。

在企业数字化转型方面，牵头组建省国资国企信息和数智化发展中心，构建“一舱三横三纵”的穿透式监管体系，打造漯河食品云、钨钼产业数字供应链平台、高标准农田等项目，助力传统产业加速数字化转型。

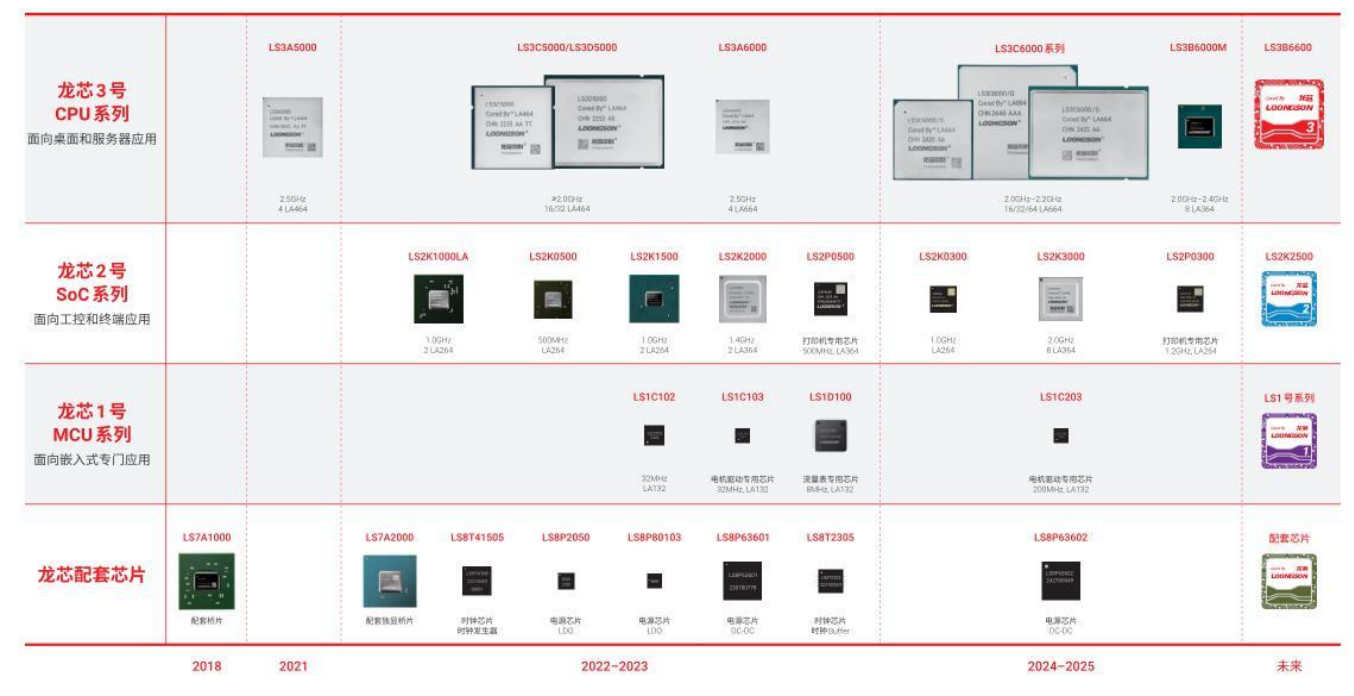
在数据要素价值化方面，组建河南数据集团，完成全省首例数据资产入表、在数据要素价值化领域的关键环节实现“零突破”，同步聚焦教育、医疗、金融、物流等加快培育河南省数据要素市场。

2025 年 3 月，豫信电科完成了 5 亿元科创债的发行，将通过发债聚焦人工智能、第三代半导体材料等领域重点布局，力争在三至五年内建成具有全国影响力的电子信息产业投资运营平台，也有望以此助力河南算力产业的发展。

5.6.3. 龙芯中科：完善了河南芯片板块布局，助力信创发展

龙芯中科（688047）是国内 CPU 领军企业，于 2001 年开始研发龙芯 CPU，拥有完全自主的龙芯架构（Loongson Architecture），起源于中科院计算所。2002 年 8 月研发出我国首枚拥有自主知识产权的通用高性能微处理芯片“龙芯一号”，2015 年 3 月我国发射了首枚使用“龙芯”的北斗卫星。2023 年 8 月，新一代四核处理器龙芯 3A6000 流片成功，总体性能与 Intel 公司 2020 年上市的第 10 代酷睿四核处理器相当。

图 46：龙芯中科处理器路线图



资料来源：龙芯中科，中原证券研究所

龙芯中科于 2022 年 6 月上市，2024 年实现收入 5.07 亿元，净利润-6.24 亿元。2024 年其传统优势工控市场停滞影响仍存在，工控类芯片营收大幅下降；电子政务市场开始回暖，信息化类芯片收入大幅增加；芯片类产品营收同比有较大幅度增长的同时，公司主动减少解决方案类业务，解决方案类业务营收同比有较大幅度下降。

龙芯中科在河南郑州、鹤壁、许昌三地进行了业务布局，建有芯片封装基地。2022 年 8 月，龙芯中科与河南省人民政府达成“1+3”战略合作，将围绕河南省信创产业，布局郑州中原总部、鹤壁生产基地、许昌整机协同基地三大产业集群。2023 年 10 月，龙芯中科芯片封装基地项目投产仪式在鹤壁科创新城举行，成为了龙芯的首个芯片封装基地。2023 年 11 月，龙芯中科中原总部基地暨中豫芯科揭牌仪式在郑州举行。龙芯中科中原总部基地建在郑州航空港区，主要建设研发创新中心、生态适配中心、信创展示中心，将围绕国产信息化、工控以及龙芯的 MCU、工控芯片、打印机芯片进行研发和市场推广。

5.6.4. 紫光股份：计划在河南投资 100 亿，建设智慧计算终端全球总部基地

紫光股份（000938）主营信息电子产业，是国内 ICT 领军企业，是国内排名前三的服务器企业。公司的核心业务基本覆盖 IT 服务的重要领域：硬件方面提供智能网络设备、存储系统、全系列服务器等为主的面向未来计算架构的先进装备。公司结合全球信息产业的发展趋势及自身优势业务的特点，将公司战略聚焦于 IT 服务领域，致力于打造一条完整而强大的“云—网—端”产业链，向云计算、移动互联网和大数据处理等信息技术的行业应用领域全面深入，并成为集现代信息系统研发、建设、运营、维护于一体的全产业链服务提供商。

紫光股份 2016 年完成了对新华三 51% 股权的收购，并在 2024 年 9 月继续收购了其 30% 股权。新华三脱胎于华为企业网络业务，也是紫光股份最具核心价值的资产。2023 年紫光股份收入 773.08 亿元，归母净利润 21.03 亿元，其中新华三收入 519.39 亿元，同比增长 4.27%，实现净利润 34.11 亿元。

紫光在河南成立了子公司，主要布局在郑州，将河南作为新华三大数据板块的全国总部，同时将河南作为其智慧计算终端的全球总部。2017 年 3 月，紫光股份旗下新华三集团的全资子公司——新华三大数据技术有限公司——在郑州高新区成立，成为新华三集团大数据板块的全国总部，主要负责承接集团的大数据以及人工智能的发展战略。2020 年 6 月，紫光股份在郑州高新区投资设立的项目公司——紫光计算机科技有限公司（简称：紫光计算机），注册成立。2020 年 8 月，紫光计算机全球研发中心总部入驻郑州高新区天健湖智联网产业园。2023 年 11 月，郑州高新区与新华三签约合作共建郑州图灵小镇，成为其“1+1+N2”智能算力服务框架中与多地共建 AI 产业基地的重要组成部分，推动其私域大模型百业灵犀的落地。

根据《2024 年河南省数字经济发展工作方案》，紫光股份总计投资 100 亿元的智慧计算终端全球总部基地项目，于 2021 年 3 月开工，计划 2030 年 12 月建设完成，其中 2024 年计划投资 12 亿元。该项目主要建设内容为：商用计算终端设计研发、生产制造中心，年产台式机、笔记本电脑、图形工作站、显示器、智能大屏等 450 万台。与此同时，该项目对应的紫光超级智能工厂项目也是 2024 年河南省列入“实施数字产业提质行动”的数字经济重大事项。

5.6.5. APUS：实现了河南省大模型备案零的突破

郑州阿帕斯科技有限公司是麒麟合盛网络技术股份有限公司（APUS）的全资子公司。APUS 于 2014 年成立，后于 2019 年在郑州建立全球第二总部。

在大模型训练方面，2024 年 APUS 发布了开源的千亿参数规模 MoE 架构大模型 APUS AI 4.0，可在消费级 GPU 上部署。此外，APUS 完成了北京 1000 亿参数和河南 2100 亿参数规模的通用大模型双备案，实现了河南省大模型备案牌照“零”的突破，使河南成为全国第 7 个拥有自有通用大模型备案的省份。

在垂直行业应用方面，APUS 训练了医疗、政务、电商、传媒、中医药、教育、司法、农科等 11 个垂直领域大模型，为相关产业赋能。

在智慧医疗方面，APUS 与省儿童医院合作的医疗大模型上线至今累计为 110 万病患进行诊治，项目入选了“中央网信办 2024 年度中国创新发展十大典型案例”“2024 全球数字经济大会医疗大模型典型案例”等。

在中医药方面，APUS 与河南工业大学联合训练的中医药大模型，执业医师考试达到

80.3 分，这一成绩在国内外同领域大模型中具有领先地位。

在政务方面，APUS 与河南省网信办合作打造的网信大模型，累计辟谣次数 6.7 万次，准确率达到 99%。

5.6.6. 科大讯飞：与农大联合发布弘农大模型，加速 AI 大模型在河南的应用

科大讯飞（002230）是国内智能语言领域龙头企业，专注于人工智能技术研究和应用，是首批国家新一代人工智能开放创新平台、语音及语言信息处理国家工程研究中心、国家 863 计划成果产业化基地及国家智能语音高新技术产业化基地，拥有国产自主可控的“星火大模型”。

2020 年，科大讯飞与河南省政府签订战略合作，并设立科大讯飞中原总部，成立河南讯飞人工智能科技有限公司、人工智能产业加速中心、人工智能创新研究院三大核心主体，并逐步引进和推广人工智能技术在教育、医疗等领域的相关应用。2021 年，经省委、省政府批准，由科大讯飞牵头组建了河南省人工智能产业研究院。2023 年 3 月，郑州市与科大讯飞签署了战略合作协议，共同推动人工智能产业高质量发展。

2024 年 7 月，科大讯飞中原总部与河南农业大学联合发布弘农大模型，在河南率先打造国产可控农业领域大模型。弘农大模型基于讯飞星火大模型和河南农业大学权威专家数据库，实现精准全面、统筹多场景的智能感知、分析、预警、决策和控制能力，有效赋能育种、种植、养殖及经营管理、人才培养等工作。

5.6.7. 比亚迪：河南建设最大制造基地，全面普及智驾方案有望催生较大的算力需求

比亚迪（002594）是新能源汽车行业的领航者。2024 年，比亚迪营收 7771 亿元，同比增长 29%，净利润 402.5 亿元，同比增长 34%，全年实现汽车销量 425 万辆，同比增长 41%。

比亚迪在河南建设其全球最大制造基地。2021 年比亚迪落户郑州航空港，确定落地后 37 天内实现开工，17 个月顺利投产，创造了新能源汽车整车制造的“河南速度”。2024 年下半年，比亚迪又与郑州航空港签约新能源零部件扩产、全球模具中心、新能源汽车动力系统零部件扩产、新能源乘用车及零部件三期、新型动力电池三期共五个项目。项目全部投产后，将形成比亚迪全球最大的制造基地。

2024 年 2 月，比亚迪发布了其整车智能化战略，其“天神之眼”高阶智驾系统将覆盖比亚迪全系车型，正式拉开了 2025 年全民智驾的序幕。考虑到自动驾驶是人工智能领域的重要应用以及自动驾驶对时延高要求，我们预计智驾车型的全面普及也对河南算力需求起到较强的拉动作用。

图 47：比亚迪天神之眼技术矩阵



资料来源：比亚迪，百度百科，中原证券研究所

6. 风险提示

国际局势的不确定性；下游企业削减开支。

行业投资评级

强于大市：未来 6 个月内行业指数相对沪深 300 涨幅 10% 以上；

同步大市：未来 6 个月内行业指数相对沪深 300 涨幅-10% 至 10% 之间；

弱于大市：未来 6 个月内行业指数相对沪深 300 跌幅 10% 以上。

公司投资评级

买入：未来 6 个月内公司相对沪深 300 涨幅 15% 以上；

增持：未来 6 个月内公司相对沪深 300 涨幅 5% 至 15%；

谨慎增持：未来 6 个月内公司相对沪深 300 涨幅-10% 至 5%；

减持：未来 6 个月内公司相对沪深 300 涨幅-15% 至 -10%；

卖出：未来 6 个月内公司相对沪深 300 跌幅 15% 以上。

证券分析师承诺

本报告署名分析师具有中国证券业协会授予的证券分析师执业资格，本人任职符合监管机构相关合规要求。本人基于认真审慎的职业态度、专业严谨的研究方法与分析逻辑，独立、客观的制作本报告。本报告准确的反映了本人的研究观点，本人对报告内容和观点负责，保证报告信息来源合法合规。

重要声明

中原证券股份有限公司具备证券投资咨询业务资格。本报告由中原证券股份有限公司（以下简称“本公司”）制作并仅向本公司客户发布，本公司不会因任何机构或个人接收到本报告而视其为本公司的当然客户。

本报告中的信息均来源于已公开的资料，本公司对这些信息的准确性及完整性不作任何保证，也不保证所含的信息不会发生任何变更。本报告中的推测、预测、评估、建议均为报告发布日的判断，本报告中的证券或投资标的的价格、价值及投资带来的收益可能会波动，过往的业绩表现也不应当作为未来证券或投资标的表现的依据和担保。报告中的信息或所表达的意见并不构成所述证券买卖的出价或征价。本报告所含观点和建议并未考虑投资者的具体投资目标、财务状况以及特殊需求，任何时候不应视为对特定投资者关于特定证券或投资标的的推荐。

本报告具有专业性，仅供专业投资者和合格投资者参考。根据《证券期货投资者适当性管理办法》相关规定，本报告作为资讯类服务属于低风险（R1）等级，普通投资者应在投资顾问指导下谨慎使用。

本报告版权归本公司所有，未经本公司书面授权，任何机构、个人不得刊载、转发本报告或本报告任何部分，不得以任何侵犯本公司版权的其他方式使用。未经授权的刊载、转发，本公司不承担任何刊载、转发责任。获得本公司书面授权的刊载、转发、引用，须在本公司允许的范围内使用，并注明报告出处、发布人、发布日期，提示使用本报告的风险。

若本公司客户（以下简称“该客户”）向第三方发送本报告，则由该客户独自为其发送行为负责，提醒通过该种途径获得本报告的投资者注意，本公司不对通过该种途径获得本报告所引起的任何损失承担任何责任。

特别声明

在合法合规的前提下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问等各种服务。本公司资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告意见或者建议不一致的投资决策。投资者应当考虑到潜在的利益冲突，勿将本报告作为投资或者其他决定的唯一信赖依据。