2026 年 11 月 07 日

计算机行业研究

买入（维持评级）

行业专题研究报告

证券研究报告

计算机组
分析师：刘高畅（执业 S1130525120005） liugaochang@gjzq.com.cn
分析师：陈芷婧（执业 S1130525120008） chenzhijing@gjzq.com.cn
分析师：鲍淑娴（执业 S1130526020002） baoshuxian@gjzq.com.cn

算力过于悲观

投资逻辑：

民主党多项 AI 监管主张集中升温，资本市场已提前计提悲观预期。11 月 3 日美国中期选举后，民主党在 AI 监管及数据中心建设方面的政策主张引发市场担忧。1）得州候任州长 Gina Hinojosa 提出，上任后拟暂停数据中心审批，直至相关居民保护立法出台；2）桑德斯再次呼吁暂停全国数据中心建设，此前提出的 S.4214 法案仍处于国会审议阶段；3）民主党众议员 Yassamin Ansari 提出暂停先进 AI 开发、加强行业监管等主张。11 月 5 日纳斯达克指数大跌 5.52%，市场对 AI 资本开支及算力产业增长前景的悲观预期明显升温。既有地方限制与新增监管主张相互叠加，市场已提前计提限制进一步扩大、持续时间延长的潜在影响，当前对算力产业的发展前景过于悲观。

基模性能、经济性及泛化能力持续突破，AI 应用渗透才刚刚开始。基础模型仍在持续迭代，性能提升、成本下降与泛化能力增强同步推进。Anthropic 于 10 月发布 Claude Haiku 5.5，平均运行成本较前代降低约 75%，知识工作、电脑操作等任务能力显著提升。随着模型能力边界不断拓宽、应用成本持续降低，AI 正由简单问答向复杂任务执行和真实生产流程加速渗透，催生出更加多元的应用形态，如企业 AI、AI 漫剧、AI4S、物理 AI 等，长期发展空间广阔，有望持续催生新的训练及推理算力需求。技术产业的中长期发展并不以短期政策意志为转移。

海内外算力供需持续紧张，国内高端有效算力尤其稀缺。算力租赁价格进一步验证供需关系。海外方面，Ornn 数据显示，截至 11 月 5 日，H100、H200、B200 按需租赁价格较 30 日前分别上涨 10.21%、27.22%、24.97%。国内方面，高端算力供给紧张更值得关注：据上海钢联 11 月 6 日周报，H100 包月、按需租赁价格指数较上周分别上涨 9.36%、6.99%；阿里云、腾讯云年内相继上调部分 AI 算力服务价格，部分算租企业合同亦出现提价及扩量。我们认为，当前国内高端有效算力呈现结构性紧缺，需求快速增长与供给扩张受限之间的矛盾仍然突出。从产业链看，关键设备、FAB 先进制造产能及高性能存储是最上游的重要供应瓶颈，芯片及云平台则是当前旺盛算力需求的有限承接者。伴随 AI 需求持续扩张，上游供给能力及下游优质算力资源的稀缺价值有望进一步凸显。

相关标的

国内算力：寒武纪、海光信息、长鑫科技、东阳光、中芯国际、华虹半导体、中科曙光、摩尔线程、华勤技术、利通电子、行云科技、智微智能、亿田智能、天数智芯、沐曦股份、壁仞科技、协创数据、禾盛新材、华丰科技、京基智农、浪潮信息、胜蓝股份、盛视科技、神州数码、润泽科技、大位科技、科华数据、潍柴重机、欧陆通、杰创智能、奥尼电子等。

海外算力：工业富联、中际旭创、东山精密、胜宏科技、方正科技、麦格米特、先导基电、光智科技、AXTI、新易盛、蓝思科技（元拾科技）、创世纪、海鸥股份、天孚通信、光迅科技、源杰科技、江海股份、生益科技、深南电路、广合科技、奥士康、世运电路、景旺电子、鹏鼎控股、沪电股份、华盛昌、联讯仪器、唯科科技、领益智造（立敏达）、中钨高新、英维克、飞龙股份、冰轮环境、德福科技、南亚新材、火炬电子、三环集团、海星股份、天岳先进、云南锗业、力量钻石、沃尔德、四方达等；英伟达、Alphabet、台积电、Lumentum、SK 海力士、英特尔、Bloom Energy、美光、闪迪。

大模型&云：智谱、MiniMax、阿里巴巴、腾讯控股、金山云、百度集团、优刻得、网宿科技等。

AI 应用：深信服、德才股份、迈富时、汇量科技、卓易信息、鼎捷数智、达梦数据、中控技术、金蝶国际、用友网络、海致科技集团、汉朔科技、金山办公、合合信息、福昕软件、金斯瑞生物科技、晶泰控股-P、英矽智能等。

风险提示

数据中心政策及出口管制收紧风险；AI 技术迭代及商业化不及预期风险；上游扩产与集群交付不及预期风险；算力租赁价格及利用率波动风险；行业竞争加剧风险。。

内容目录

一、政策主张尚未全面落地，资本市场已提前计提 3
二、模型能力与经济性持续改善，应用渗透仍有广阔空间 3
2.1 性能提升叠加成本下降，更多任务进入可商业化区间 3
2.2 企业 AI 进入核心经营流程，模型厂商与业务系统加速融合 5
2.3 AI 漫剧供需同步扩张，视频模型升级改善生产经济性 5
2.4 尖端模型泛化能力提升，物理 AI 延伸真实环境交互 7
2.5 AI4S 进入设计与实验迭代，科学计算需求持续拓展 7
三、海内外租金反映资源稀缺，国内高端有效算力尤为紧缺 9
3.1 海外多型号租金近月上涨，Neocloud 需求持续旺盛 9
3.2 国内高端 GPU 价格上行，云厂商提价验证有效供给紧张 9
3.3 算租公司合同出现提价与扩容，交付资源的商业价值提升 11
3.4 设备、FAB 与存储约束扩产，上游能力决定放量速度 11
3.5 芯片与云加快放量承接需求，但高端供给仍紧缺 12
四、相关标的 13
风险提示 13

图表目录

图表 1：GovInfo 所列 S.4214 仍为参议院提出稿，最后行动为交委员会审议 3
图表 2：Claude Haiku 5.5 官方跨任务评测结果 4
图表 3：SAP Joule Studio 支持按业务事件触发补货 Agent 5
图表 4：2026 年 9 月短剧视频各类应用月活跃用户规模 TOP10 6
图表 5：Cosmos Framework 覆盖数据、训练、适配与推理优化 7
图表 6：Claude 参与抗体研发：目标分解、设计优化与候选筛选 8
图表 7：Ornn 多型号 GPU 按需租赁价格近月上行 9
图表 8：上海钢联国内外主要 GPU 算力租赁价格周度变化 10
图表 9：部分算租合同的收费机制及调整情况 11
图表 10：TrendForce 预计 2026Q4 DRAM 与 NAND 合约价继续上涨 12
图表 11：国产芯片公司 2026 年前三季度收入与归母净利润 12
图表 12：主要云厂商业务进展与算力供给投入 13

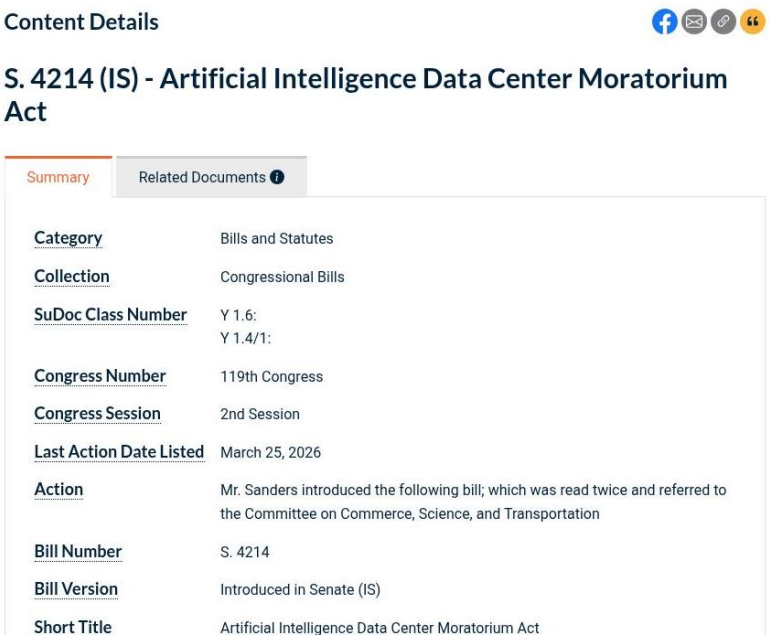
一、政策主张尚未全面落地，资本市场已提前计提

美国数据中心监管争议主要源于算力建设加速带来的电力、水资源及居民成本压力。得州是重要建设区域，此前跟踪的待接网数据中心项目申请规模已超过 400GW，并非已建成容量。据路透社 11 月 4 日报道，Abbott 自 8 月起收紧数据中心接网；9 月 21 日，得州州长办公室正式要求环境质量委员会暂停相关许可，等待电网等审计完成，并要求数据中心自行承担相关成本、保护电网和水资源。

11 月 3 日美国中期选举后，民主党胜选强化了市场对监管持续收紧的预期。得州候任州长 Gina Hinojosa 于 10 月 22 日承诺，就任后暂停数据中心审批，直至州议会通过居民保护立法，并提出保障社区用水、避免电费转嫁及取消相关税收补贴等主张。她将于 2027 年 1 月 19 日就任。与 Abbott 以完成审计为条件的既有限制相比，新主张可能使审批暂停持续更长时间，相关新增安排尚未实施。

联邦层面的 AI 监管主张也在选举后重新受到关注。桑德斯于 3 月 25 日提出 S.4214《人工智能数据中心暂停法案》，拟暂停相关建设，直至出台公众保护立法；9 月 23 日又提出 S.5493《禁止人工超级智能法案》，Casar 次日提出众议院版本 H.R.10538，Ansari 为共同提案人。后两项法案拟设立联邦 AI 监管机构，在监管机制建立前暂停先进 AI 开发，并禁止人工超级智能开发与部署。11 月 4 日，Ansari 再次呼吁将上述方向列为议程重点。截至 11 月 6 日，三项法案所列版本均停留在提出及交委员会审议阶段，尚未成为生效法律。

图表 1：GovInfo 所列 S.4214 仍为参议院提出稿，最后行动为交委员会审议



来源：美国政府出版局 GovInfo，查阅日期 2026 年 11 月 6 日，国金证券研究所

资本市场已提前反映对 AI 产业增长前景的担忧。据路透社收盘报道，11 月 5 日纳斯达克综合指数下跌 5.52%，标普 500 指数下跌 4.36%，费城半导体指数跌逾 12%，英伟达跌逾 9%，CoreWeave 跌近 10%。本轮调整叠加了部分企业业绩表现及监管预期变化等因素。数据中心审批与 AI 开发限制，可能改变未来资本开支、项目投运及设备采购节奏，加剧市场对算力增长持续性的担忧。

市场已提前计提限制范围进一步扩大、持续时间延长的潜在影响。既有地方措施与选举后升温的新增主张相互叠加，但新增安排尚未全面实施，联邦法案也尚未生效。与此同时，模型能力持续进步、AI 应用加速扩张，训练及推理需求仍在增长，当前市场对算力产业的中长期发展前景过于悲观。

二、模型能力与经济性持续改善，应用渗透仍有广阔空间

2.1 性能提升叠加成本下降，更多任务进入可商业化区间

基础模型的性能与经济性仍在同步改善。Anthropic 于 10 月 7 日发布 Claude Haiku 5.5，适用于摘要、数据库查询、实时客服及编程子任务等高频场景，平均运行成本较 Haiku 4.5 降低约 75%。官方评测显示，其 GDPval-AA v2.1 知识工作得分由前代 735 提升至 1620，OSWorld 2.1 离线子集电脑操作成绩由 15.7%提升至 72.4%。

前沿能力提升与能力下沉同步推进。OpenAI 发布的 GPT-6 Astra 可执行多步骤专业工作流，在通过多视图生成 CAD 代码重建三维物体的 BenchCAD 评测中，几何重叠得分达到 95.9%，GPT-5.6 Sol 为 83.3%；GPT-6.1 Sol 则在 Agent 编程、电脑操作和专业工作等任务上接近 Astra 水平，标准输入、输出 Token 价格均为 Astra 的五分之一，每百万 Token 分别为 2 美元、10 美元。

图表 2：Claude Haiku 5.5 官方跨任务评测结果

	Haiku 5.5	Haiku 4.5	GPT-6 Luna	Sonnet 5.5 For reference
Knowledge work GDPval-AA v2.1	**1620**	735	1437	1840
Knowledge work AA-Briefcase v1.1	**1578**	614	1336	1824
Computer use OSWorld 2.1	**72.4%** Offline subset	15.7% Offline subset	48.9% Offline subset	83.9% Offline subset
Multidisciplinary reasoning Humanity's Last Exam	**45.9%** no tools	10.2% no tools	—	56.9% no tools
	57.4% with tools	18.7% with tools	—	64.5% with tools
Agentic coding Terminal-Bench 4.0	**39.2%**	0.0%	16.4%	70.6%
Agentic coding FrontierCode 1.1 (Main)	**46.4%**	—	42.4%	52.1% Xhigh
Visual reasoning Chartography	**46.4%** no tools	6.4% no tools	29.1% no tools	61.6% no tools

For details on how we run our evaluations, see the Haiku 5.5 System Card.

来源：Anthropic《Introducing Claude Haiku 5.5》，国金证券研究所

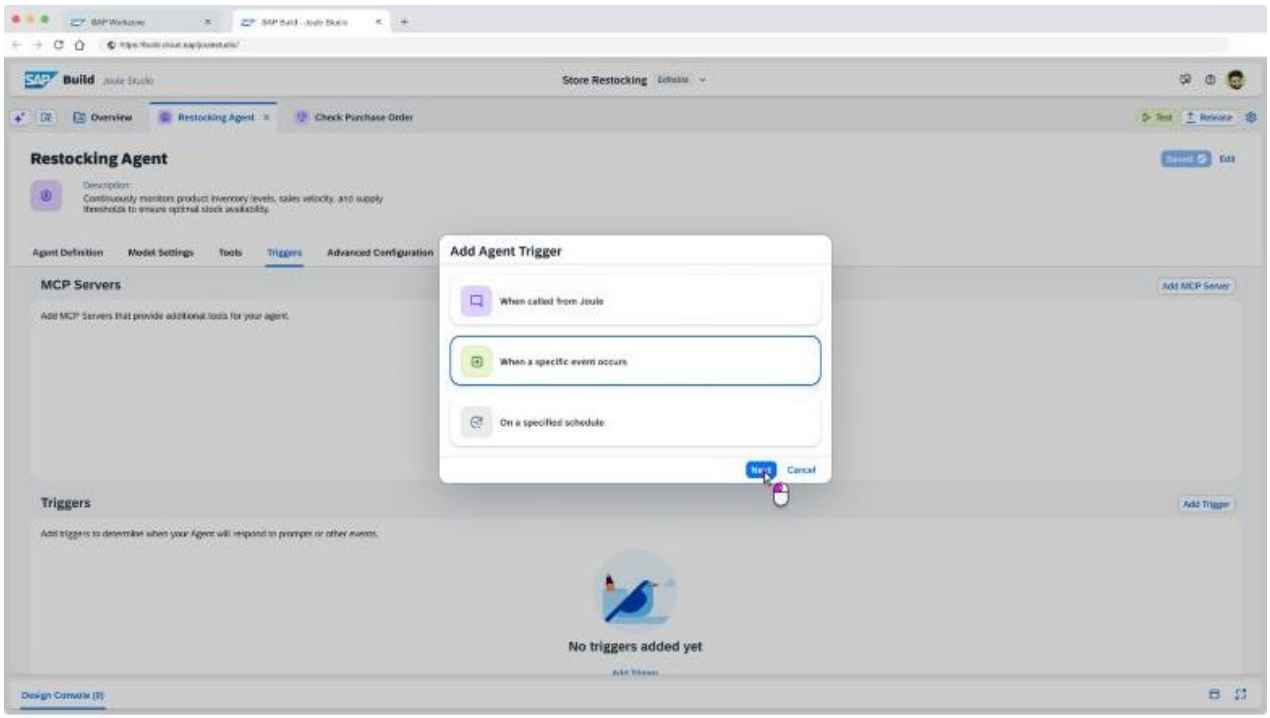
模型能力和经济性同步提升，Coding 之外的 AI 应用正在加速渗透、百花齐放。企业经营、视频制作、真实环境交互和科学研发等场景，正随着模型更好用、使用成本更低而加快商业化。

2.2 企业 AI 进入核心经营流程，模型厂商与业务系统加速融合

企业 AI 正在从员工辅助工具进入 ERP 等核心经营系统。SAP 于 10 月 26 日发布的三季度 AI 更新披露，已有 4.5 万家云客户使用其 AI 解决方案。Joule 结合业务上下文、用户权限和专业 Agent 执行任务，应用包括开具发票、识别异常、处理过账错误、发现供应链短缺及协调维护行动。Joule Studio 支持为自定义 Agent 配置触发条件，使 AI 能够嵌入具体业务流程。

模型厂商正通过合作嵌入企业业务系统。1）Anthropic 与 Salesforce 共同推出的 Salesforce in Claude 于 9 月 15 日进入测试版，提供 37 项销售技能，覆盖客户研究、会前准备、销售管线复盘及 CRM 更新，已有 7000 名 Salesforce 销售人员使用；2）OpenAI 与 ServiceNow 于 10 月 28 日宣布扩大合作，将业务上下文和工作流操作引入 ChatGPT，ServiceNow MCP 插件当日已在 OpenAI 市场上线，并计划继续扩展应用开发、安全等能力。

图表 3：SAP Joule Studio 支持按业务事件触发补货 Agent



来源：SAP《AI Release Highlights Q3 2026》官方产品界面，国金证券研究所

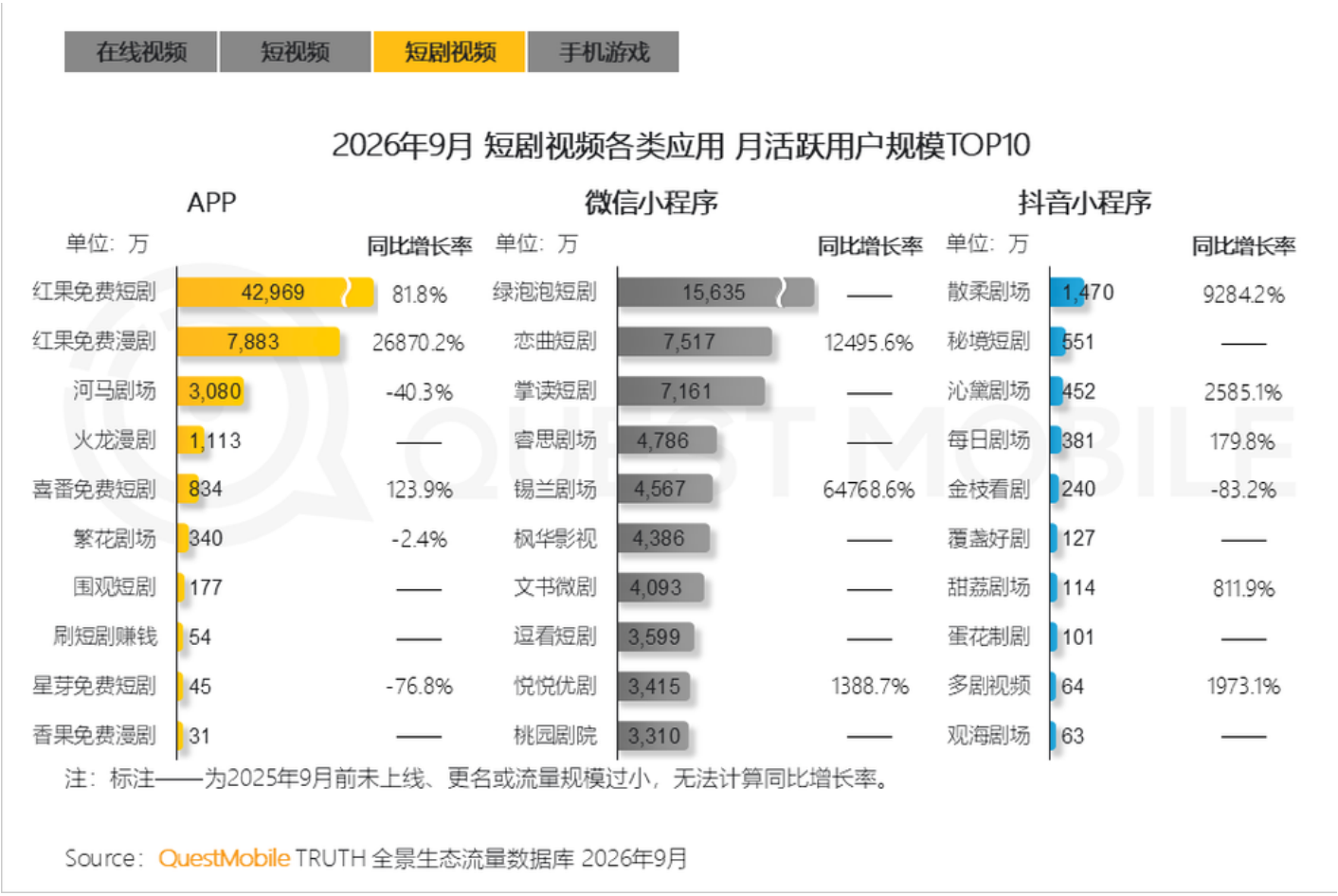
企业 AI 的算力增量来自业务流程中的持续调用。采购申请、订单变化、财务单据和销售跟进，都能触发模型查询、分析与执行，随着 AI 覆盖更多部门和经营环节，有望形成高频、持续的推理需求。

2.3 AI 漫剧供需同步扩张，视频模型升级改善生产经济性

AI 漫剧的用户与内容供给仍在快速扩张。QuestMobile 数据显示，2026 年 9 月按月度全景流量去重口径，红果免费短剧用户规模约 4.30 亿，超过同口径豆包的约 4.19 亿；红果免费漫剧 APP 月活跃用户为 7883 万。DataEye 研究院数据显示，9 月抖音原生端 AI 剧漫剧新剧数为 5.18 万部、环比增长 10.7%，新剧播放增量为 2093.04 亿、环比增长 27%。

视频生成正在由单镜头演示走向连续叙事与专业制作。字节跳动 7 月 31 日发布 Seedance 2.5，单次生成时长由 15 秒延长至 30 秒，支持多轮延长、最多 30 张图片参考和时间戳级定向编辑；可灵 4.0 于 10 月 15 日上线，改善高速动作、复杂运镜与连续镜头稳定性，支持原生 30 秒生成、最多 15 项参考[illegible]关键帧输入。

图表 4：2026 年 9 月短剧视频各类应用月活跃用户规模 TOP10



来源：QuestMobile《2026 中国移动互联网秋季大报告》，国金证券研究所

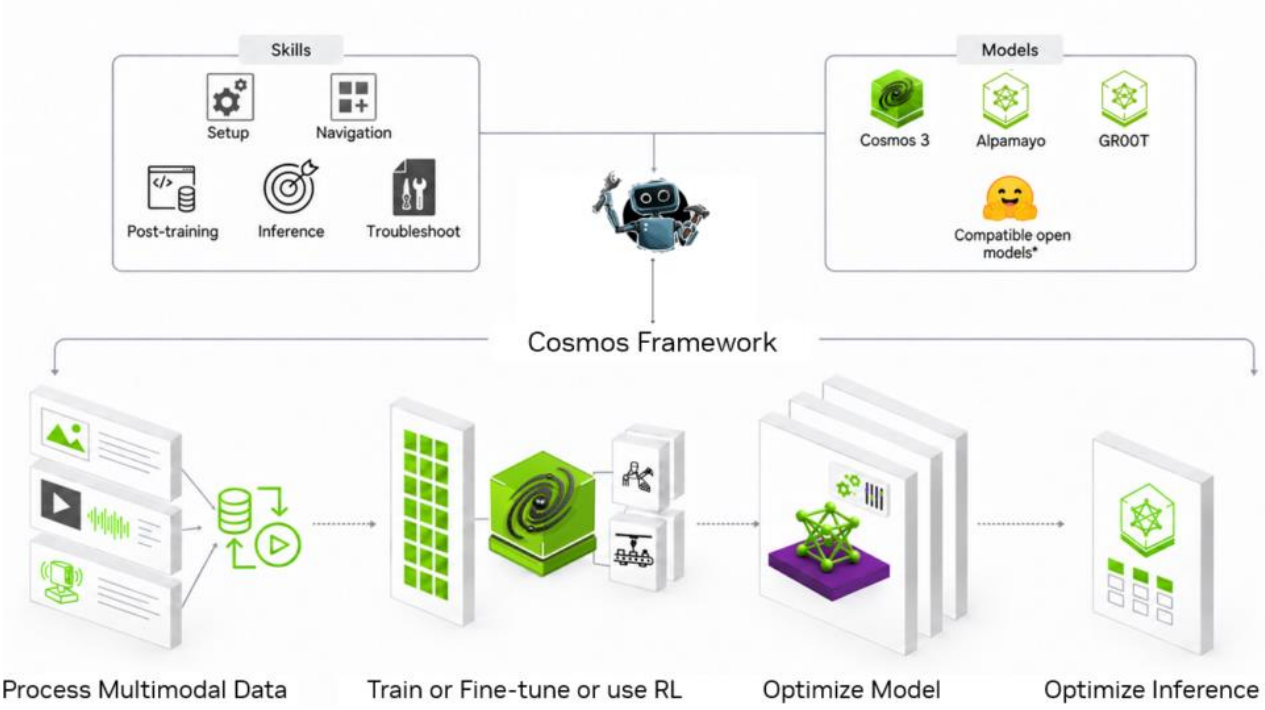
视频模型升级正在转化为付费需求。快手 2026Q2 业绩公告披露，可灵 AI 季度收入超过 8.5 亿元，同比增长超过 200%。更稳定的角色、更连贯的镜头和更方便的编辑，有助于减少试错和返工，推动 AI 漫剧规模化生产。随着内容产量增加、画质和叙事要求提高，视频生成及多轮修改带来的算力需求有望继续增长。

2.4 尖端模型泛化能力提升，物理 AI 延伸真实环境交互

Astra 等尖端模型的任务理解、规划和工具调用能力持续提升，正在为 AI 向物理世界泛化提供基础。模型由处理文字扩展至理解空间、操作专业软件和执行多步骤任务，有助于机器人将通用能力迁移到不同环境，完成识别、决策与行动。真实场景更丰富、任务链条更长，也将带来更多训练和推理需求。

NVIDIA 于 10 月 21 日扩展 Cosmos 平台，为机器人和自动驾驶提供数据生成与训练工具。ANYbotics 已使用相关工具生成超过 1500 段合成视频，训练工业巡检机器人识别液体泄漏等环境变化，将模型能力用于具体的真实环境任务。

图表 5：Cosmos Framework 覆盖数据、训练、适配与推理优化



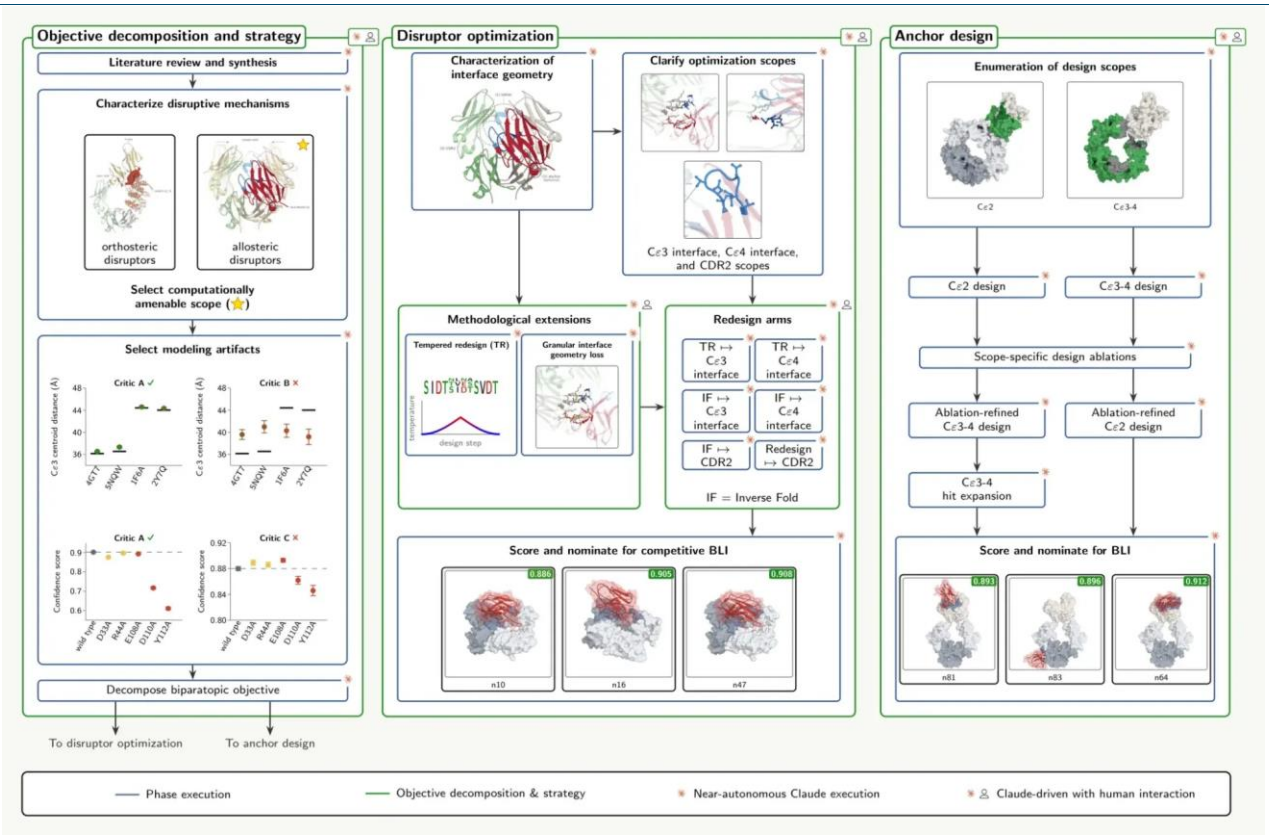
来源：NVIDIA《NVIDIA Cosmos Opens New Tools to Train, Evaluate and Deploy World Models》，国金证券研究所

2.5 AI4S 进入设计与实验迭代，科学计算需求持续拓展

AI4S 正在由科研知识辅助进入分子设计、计算筛选与实验迭代。Anthropic 于 10 月 27 日披露与 Divergence Labs 合作的抗体研发案例，Claude 结合分子设计和结构预测工具生成约 3.5 万个设计，再筛选候选进入实验测试，显示模型已开始参与科研流程。

Twist Bioscience 参与的蝎毒蛋白结合物项目更直观地体现了科研计算需求。Anthropic 于 10 月 20 日披露，Claude 研究模型在 64 小时内调用计算工具，计算用量约 1 万 GPU 小时；Adaptyv Bio 与 Twist 随后对 211 个候选进行实验测试，其中 17 个达到预设结合标准。模型推理、专业工具计算和实验反馈相互衔接，推动研发向多轮迭代延伸。

图表 6：Claude 参与抗体研发：目标分解、设计优化与候选筛选



来源：Anthropic《Early progress on accelerating drug development with Claude》，国金证券研究所

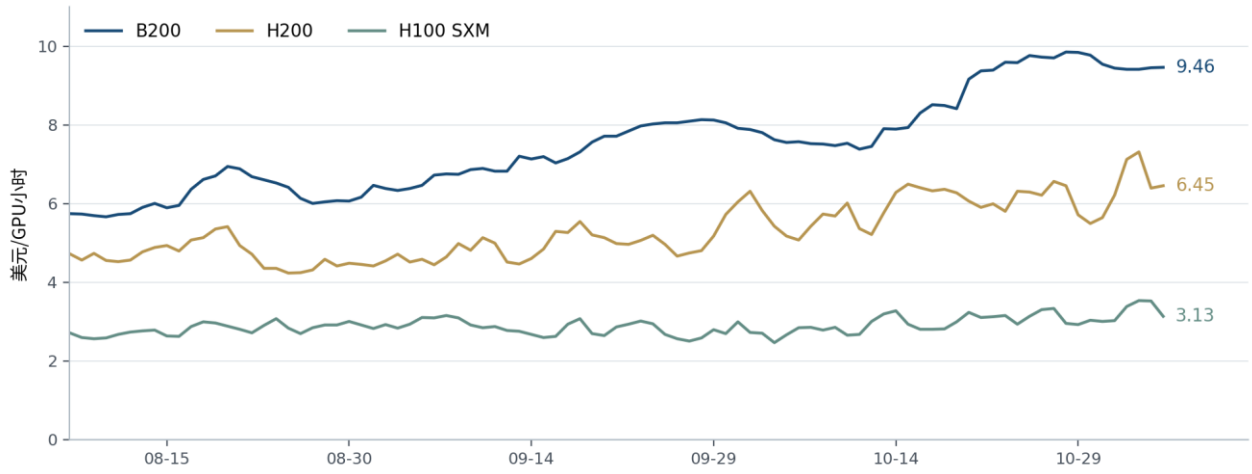
企业 AI、AI 漫剧、物理 AI 和 AI4S 均处于渗透早期。随着模型能力与经济性改善，更多经营、内容和科研任务将进入可商业化区间，应用规模及算力需求仍有广阔增长空间。技术产业的中长期发展并不以短期政策意志为转移。

三、海内外租金反映资源稀缺，国内高端有效算力尤为紧缺

3.1 海外多型号租金近月上涨，Neocloud 需求持续旺盛

海外多代 GPU 价格共同反映算力服务仍处于涨价周期。Ornn Compute Price Index 显示，截至 11 月 5 日，H100、H200、B200 按需租赁价格分别为 3.13、6.45、9.46 美元/GPU 小时，较 30 日前分别上涨 10.21%、27.22%、24.97%。价格上涨同时出现在不同代际及显存配置的产品，高性能算力需求仍有较强支撑。

图表 7：Ornn 多型号 GPU 按需租赁价格近月上行



来源：Ornn Compute Price Index；数据截至 2026 年 11 月 5 日，国金证券研究所

Nebius 年内两次调价进一步验证 Neocloud 需求旺盛。公司官方更新日志显示，首轮调价披露、6 月 1 日生效，涉及 H100、H200、B200、B300 等产品；9 月 17 日路透社报道，公司将自 10 月 1 日起再次提高 GPU 按需租金，涨幅约 17%—21%。目前官网 H100、H200、B200、B300 按需价格分别为 4.50、5.40、8.50、9.50 美元/GPU 小时。

涨价背后是合同规模与客户付费能力同步增强。Nebius 在 8 月 12 日股东信中披露，二季度 AI 云收入为 5.75 亿美元、同比增长 514%，签下四笔平均总合同价值超过 10 亿美元的大单；约 70%的当季合同包含预付款，可覆盖相关资本开支的 50%—60%，旧代 GPU 定价亦较一季度提高超过 30%。新客户、大额合同与预付需求共同支撑资源扩张，紧缺算力的议价能力持续提升。

3.2 国内高端 GPU 价格上行，云厂商提价验证有效供给紧张

国内高端 GPU 价格呈现结构性上行。上海钢联 Mysteel 11 月 6 日周报显示，H100 80G 包月租赁价格指数为 19.62，较上周五上涨 9.36%；按需指数为 19.14，上涨 6.99%。H20 96G 按需指数为 7.64，上涨 1.60%，包月指数为 6.37、持平；A100 40G 包月指数下降 8.13%。高端资源涨价与部分型号回落并存，供需紧张集中于具体性能及计费资源。

图表 8：上海钢联国内外主要 GPU 算力租赁价格周度变化

市场	GPU型号	租期	10月30日	11月6日	周涨跌幅	走势
国内	H100 80G	包月	17.94	19.62	9.36%	↑ 上涨
国内	H100 80G	按需	17.89	19.14	6.99%	↑ 上涨
国内	H20 96G	按需	7.52	7.64	1.60%	↑ 上涨
国内	A800 80G	按需	7.03	7.07	0.57%	↑ 上涨
国内	A100 40G	包月	2.46	2.26	-8.13%	↓ 下跌
国内	A100 80G	按需	9.01	8.84	-1.89%	↓ 下跌
国内	H800 80G	按需	12.66	12.38	-2.21%	↓ 下跌
国内	PRO 6000 96G	包月	6.19	5.98	-3.39%	↓ 下跌
国内	RTX 5090 32G	按需	3	2.96	-1.33%	↓ 下跌
国内	RTX 4090 D 24G	按需	1.92	1.88	-2.08%	↓ 下跌
国内	RTX 4090 24G	包月	1.77	1.77	0.00%	→ 持平
国际	RTX 5090 32G	按需	0.61	0.68	11.48%	↑ 上涨
国际	A100 80G	按需	1.67	1.76	5.39%	↑ 上涨
国际	H200 141G	按需	4.59	4.6	0.22%	↑ 上涨
国际	AMD MI300X 192G	按需	3.01	3.03	0.66%	↑ 上涨
国际	B200 192G	按需	8.03	7.64	-4.86%	↓ 下跌
国际	H100 80G	按需	3.21	3.14	-2.18%	↓ 下跌

来源：上海钢联 Mysteel 2026 年 11 月 6 日周报；价格指数及计费口径按源站列示，国金证券研究所

国内云厂商的正式调价提供了供需与成本变化的直接证据。1）阿里云 3 月 18 日发布公告，自 4 月 18 日起调整 AI 算力及 CPFS 智算版等服务价格，所列 AI 算力卡相关服务上调 5%—34%，CPFS 智算版上调 30%。公告将调价原因归于全球 AI 需求增长、供应链涨价及核心硬件采购成本上升，既有订单在当前计费周期内不受影响。2）腾讯云 4 月 9 日发布正式调价公告，自 5 月 9 日起将公告所列 AI 算力相关产品服务刊例价上调 5%，TKE 原生节点及 EMR 相关产品服务亦上调 5%。公告解释原因为全球 AI 需求激增、核心硬件等供应链成本大幅上涨，已购订单在当前计费周期内不受影响，续费执行新价格。

国内紧缺的是能够稳定承接大模型训练和复杂推理的高端有效算力。芯片性能、显存、集群互联和软件适配共同决定资源的可用性，供给扩张还需要完成组网和部署。高端 GPU 租金与云服务价格上行，反映了需求快速增长与有效供给释放较慢之间的矛盾，已有可交付算力资源的稀缺价值更加突出。

3.3 算租公司合同出现提价与扩容，交付资源的商业价值提升

行云科技 7 月 30 日披露，与 VB 客户的合作规模由 128 个单元扩展至 256 个单元，已交付 8 个单元及其余 248 个单元的每单元月服务费分别提高 21.21%、36.36%。10 月 12 日，公司再次公告调整同一合作，在 256 个单元、5 年服务期基础上，合同总额由 30.528 亿元提高至 33.5689 亿元，增幅约 9.96%，固定服务费与固定 Token 分成的结构同步调整。两次合同变化体现了合作规模与收费机制的持续演进，具体条款见图表 9。

图表 9：部分算租合同的收费机制及调整情况

公司 / 披露时间	合同性质	价格及收费变化	供需观察角度
行云科技 / 7 月 30 日	VB 客户扩容及调整收费	已交付单元月费+21.21%；其余单元月固定算力费+36.36%	单价与规模同时调整
行云科技 / 10 月 12 日	同一客户重新签约	剩余单元固定服务费-11.11%；固定 Token 分成+210%	两项共同构成月度服务费
利通电子 / 10 月 15 日	长期战略合作客户	协议期内一般为固定价格	新增订单及续约另行定价

来源：行云科技、利通电子公司公告；7 月条款为当次调整，后续以新合同为准，国金证券研究所

其他算租公司的合同与定价安排亦在变化。1）阳光股份 10 月 21 日披露，对尚未交付的服务器租赁服务提高单价并增加数量，含税服务费总额由 6.33 亿元增至 11.0448 亿元；2）利通电子 10 月 15 日披露，长期合作客户在约定期限内一般执行固定价格，超出期限后可结合市场供需及资源配置调整；3）赛意信息 9 月 16 日披露，两份 60 个月算力服务合同的合计含税金额由 64.5 亿元提高至 67.2 亿元，服务期限及交付排期等核心条款保持不变。

3.4 设备、FAB 与存储约束扩产，上游能力决定放量速度

设备端的交付约束已有龙头公司明确表述。ASML 在 10 月 14 日发布的 2026Q3 业绩材料中表示，客户持续提高中短期设备需求预期，主要由 AI 投资规模和持续时间的判断驱动；部分客户有意为 2029 年甚至 2030 年下单。公司预计，按照当前规划，2027 年和 2028 年的供给仍不足以完全满足客户需求，并研究进一步提高 2028 年产能。

FAB 端的扩产与高负荷运行并存。中芯国际 2026Q2 业绩公告显示，期末月产能由一季度的 107.825 万片增至 109.65 万片，均为折合 8 英寸标准逻辑口径，产能利用率为 93.7%；二季度资本开支 18.36 亿美元。公司表示，将调配现有产能并加快新增产能验证，以缓解产业链供应约束。制造供给正在增加，但高端 AI 芯片还依赖特定工艺平台及其量产爬坡，整体晶圆产能扩张无法直接等同于高端算力供给释放。

存储供给紧张同时反映在长期协议与价格预期中。SK 海力士 10 月 28 日披露，已与 14 家主要客户签署长期供应协议，覆盖 2027 年及 2028 年总产能的 50%以上，先进工艺难度和新设施建设周期限制供给扩张。TrendForce 于 9 月 30 日预计，2026Q4 传统 DRAM 合约价环比上涨 10%—15%，含 HBM 的 DRAM 综合价格上涨 15%—20%，NAND Flash 合约价上涨 15%—20%。

国内产业链亦有直接反馈。协创数据 11 月 3 日公开投资者关系活动记录披露，HBM 占用产能使 DDR 供给紧张，近期 DDR 报价提高 30%且获客户接受，部分国产 GPU 也受到 HBM 供应约束。

图表 10：TrendForce 预计 2026Q4 DRAM 与 NAND 合约价继续上涨

DRAM and NAND Flash Contract Price Projections, 3Q26-4Q26

	3Q26E	4Q26F
Total DRAM	Conventional DRAM: up 13~18% HBM Blended: up 8~13%	Conventional DRAM: up 10~15% HBM Blended: up 15~20%
Total NAND Flash	up 18~23%	up 15~20%

Source: TrendForce, Sept. 2026

TrendForce

来源：TrendForce，2026 年 9 月；4Q26F 为预测值，国金证券研究所

设备交付、FAB 工艺与产能爬坡、高性能存储供应共同决定高端算力的放量速度。扩产投入需要经过建设、设备安装、工艺验证和量产爬坡，需求增长难以在短期内换来同比例的有效供给增加。

3.5 芯片与云加快放量承接需求，但高端供给仍紧缺

国产芯片与云平台正在加快放量，持续增加可交付的有效算力。芯片企业将有限的制造和存储资源转化为训练、推理产品，云平台进一步完成网络、存储、调度及模型适配，将芯片能力转化为客户能够持续调用的服务。两者共同缓解算力缺口，并在旺盛需求中扩大业务规模。

国产芯片公司的收入增长已体现需求承接。2026 年前三季度，寒武纪收入 95.26 亿元、同比增长 106.76%，归母净利润 37.15 亿元、同比增长 131.49%；海光信息收入 149.55 亿元、同比增长 57.58%，归母净利润 28.47 亿元、同比增长 45.12%；摩尔线程收入 32.89 亿元、同比增长 319.15%，归母净利润 0.38 亿元、实现扭亏。

图表 11：国产芯片公司 2026 年前三季度收入与归母净利润

公司	前三季度收入（亿元）	收入同比	前三季度归母净利润（亿元）	利润同比
寒武纪	95.26	106.76%	37.15	131.49%
海光信息	149.55	57.58%	28.47	45.12%
摩尔线程	32.89	319.15%	0.38	扭亏为盈

来源：寒武纪、海光信息、摩尔线程 2026 年第三季度报告；公司整体口径，国金证券研究所

云厂商亦在持续扩大资源投入。2026Q2，阿里巴巴 AI 云与算力服务收入 484.37 亿元、同比增长 45%，集团资本开支 676.78 亿元、同比增长 75%；腾讯集团资本开支 527.84 亿元、同比增长 176%，AI 服务需求、国际拓展及定价改善推动云服务收入增长；金山云公有云收入同比增长 45.1%，当季资本开支 33 亿元，包含租赁资本化资产。

图表 12：主要云厂商业务进展与算力供给投入

厂商 / 期间	业务增长或进展	供给投入与建设
阿里巴巴 / 2026Q2	AI 云与算力服务收入 484.37 亿元，同比增长 45%	集团资本开支 676.78 亿元，同比增长 75%
腾讯控股 / 2026Q2	AI 服务需求、国际拓展和定价改善推动云服务收入增长	集团资本开支 527.84 亿元，同比增长 176%
天翼云 / 2026H1	收入 618 亿元，同比增长 7.8%	已建成 16 个百 MW 级园区；自有+接入智算规模超 118.8EFLOPS
金山云 / 2026Q2	公有云收入 23.58 亿元，同比增长 45.1%；AI 云账单收入同比增长 82%	资本开支 33 亿元，含租赁资本化资产
优刻得 / 2026Q1—Q3	收入 15.91 亿元，同比增长 29.67%；AI 相关收入占比超 40%	购建长期资产现金支出 6.15 亿元；智算中心资源投产上架
青云科技 / 2026H1	完善 MaaS、Token 资源管理及多模态模型服务计费	智算平台支持 1 小时内部署及异构算力自动化部署

来源：阿里巴巴、腾讯控股、金山云季度业绩，中国电信、青云科技半年报，优刻得三季报；期间按表中列示，国金证券研究所

国产芯片放量与云平台扩建正在缓解算力缺口，但上游设备、先进制造和高性能存储仍制约高端供给扩张。芯片与云持续承接旺盛需求，优质可交付算力资源的稀缺价值依然突出。

四、相关标的

国内算力：寒武纪、海光信息、长鑫科技、东阳光、中芯国际、华虹半导体、中科曙光、摩尔线程、华勤技术、利通电子、行云科技、智微智能、亿田智能、天数智芯、沐曦股份、壁仞科技、协创数据、禾盛新材、华丰科技、京基智农、浪潮信息、胜蓝股份、盛视科技、神州数码、润泽科技、大位科技、科华数据、潍柴重机、欧陆通、杰创智能、奥尼电子等。

海外算力：工业富联、中际旭创、东山精密、胜宏科技、方正科技、麦格米特、先导基电、光智科技、AXTI、新易盛、蓝思科技（元拾科技）、创世纪、海鸥股份、天孚通信、光迅科技、源杰科技、江海股份、生益科技、深南电路、广合科技、奥士康、世运电路、景旺电子、鹏鼎控股、沪电股份、华盛昌、联讯仪器、唯科科技、领益智造（立敏达）、中钨高新、英维克、飞龙股份、冰轮环境、德福科技、南亚新材、火炬电子、三环集团、海星股份、天岳先进、云南锗业、力量钻石、沃尔德、四方达等；英伟达、Alphabet、台积电、Lumentum、SK 海力士、英特尔、Bloom Energy、美光、闪迪。

大模型&云：智谱、MiniMax、阿里巴巴、腾讯控股、金山云、百度集团、优刻得、网宿科技等。

AI 应用：深信服、德才股份、迈富时、汇量科技、卓易信息、鼎捷数智、达梦数据、中控技术、金蝶国际、用友网络、海致科技集团、汉朔科技、金山办公、合合信息、福昕软件、金斯瑞生物科技、晶泰控股-P、英矽智能等。

风险提示

数据中心政策及出口管制收紧风险：美国联邦及地方数据中心监管、项目审批、电力成本分担与芯片出口政策发生变化，可能影响建设进度、设备采购和跨境交付。

AI 技术迭代及商业化不及预期风险：模型能力改善和应用渗透速度低于预期，客户付费意愿、企业 IT 预算及项目回报不足，可能影响新增算力需求。

上游扩产与集群交付不及预期风险：关键设备、制造、存储和封装供给不足，或工艺爬坡、软件适配及系统验证进度低于预期，可能影响可用资源释放。

算力租赁价格及利用率波动风险：不同 GPU 型号供需变化、技术迭代及租赁期限差异，可能引发资源价格、利用率和运营回报波动。

行业竞争加剧风险：芯片、云与应用市场竞争加剧，可能影响企业定价、客户获取成本和盈利能力。

行业投资评级的说明：

买入：预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上；

增持：预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%；

中性：预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%；

减持：预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。

特别声明：

国金证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级（含 C3 级）的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海
电话：021-80234211
邮箱：researchsh@gjzq.com.cn
邮编：201204
地址：上海浦东新区芳甸路 1088 号紫竹国际大厦 5 楼

北京
电话：010-85950438
邮箱：researchbj@gjzq.com.cn
邮编：100005
地址：北京市东城区建内大街 26 号新闻大厦 8 层南侧

深圳
电话：0755-86695353
邮箱：researchsz@gjzq.com.cn
邮编：518000
地址：深圳市福田区金田路 2028 号皇岗商务中心 18 楼 1806

【小程序】
国金证券研究服务

【公众号】
国金证券研究