仅供内部参考，请勿外传

华泰证券 HUATAI SECURITIES

行业加速分化，AI 成熟度成分化根源

--计算机行业 26 年中报总结

华泰研究

2026 年 9 月 01 日 | 中国内地　　专题研究

26 年中报总结：营收盈利增速边际放缓，行业分化逐渐显现

2026 年 H1 计算机板块业绩增速边际出现回落，营业收入同比增速中位数下降至 2.5%，低位窄幅波动特征显著；扣非净利润同比增速中位数-4.3%，同比转负；经营性净现金流净额中位数同比增速-6.3%，反映行业供需趋于平衡，业内竞争下回款滞后、营运资本承压，挤占利润空间。从结构上看，毛利率延续下行态势，研发费用率略有下降，管理、销售投入强度持续提高，期间费用率抬头，此外信用减值损失占比回升，资产减值损失强度保持上升态势。

大模型：长程 Agent、开放生态与商业化协同演进

大模型产业正由单点能力竞争转向长程 Agent、开放生态与商业化协同演进。海外模型持续强化长任务、工具调用及模型-Harness 协同能力，国内厂商则通过后训练 Scaling、稀疏及线性注意力、超长上下文和推理系统优化加快追赶；与此同时，开放模型竞争已由权重发布延伸至许可证、Infra、框架及真实 Token 调用，模型能力提升也开始在 API 使用、ARR 及企业服务收入中得到验证。我们认为，后续行业竞争重点将进一步转向长任务稳定性、单位有效任务成本及模型能力向真实商业需求的转化效率。

全球软件 FDE 加速，看好国内软件厂商业绩放量

AI 时代软件竞争从标准化产品逐步转向“产品能力+行业 Know-how+深度交付”并重，全球软件厂商同步加速 FDE 转型，国内软件厂商的 Know-how 优势及客户粘性优势正加速凸显。从中报来看，由于闭源模型叙事走弱，美股软件进入估值修复期，带动全球软件估值预期改善。同时，国内 AI 应用代表企业 26H1 普遍实现营收加速及盈利改善，实现估值+业绩共振。我们认为，FDE 模式放大国内软件厂商的优势，正加速推动软件厂商 AI 业务由试点验证向正式合同转化，看好国内软件厂商全年业绩放量。

算力租赁：收入利润高增，资产端积极扩张

算力租赁板块延续高景气，我们选取 20 家算力租赁相关公司作为样本，研究样本 26H1 营业收入中位数同比增速达到 40.1%，反映前期投入的算力资源陆续完成交付上架，叠加下游 AI 算力需求持续增长，收入保持较快放量；同期归母净利润中位数同比增速达到 11%，主要受益于资产利用率提升、高端算力价格支撑及固定成本摊薄。资产端看，2026H1 样本公司在建工程、固定资产、使用权资产及存货平均规模分别同比增长 92.7%、65.9%、128.9%和 64.8%，后续有较多资源具备转化为可计费产能的空间。展望后续，我们认为伴随国内算力需求的高景气，算力租赁业绩弹性有望持续释放。

AI4S：收入分化显著，大额 BD 驱动行业商业化提速

2026 年上半年 AI4S 代表性港股标的业绩分化显著，行业步入商业化早期，大额 BD 授权成为业绩重要变量，盈利拐点尚未全面到来。收入端，企业增长梯队明显分化：英矽智能受益礼来 BD 首付款及研发服务交付，营收同比增长 287%；剂泰科技在低基数与 BD 项目集中确认下营收增速高达 13399%；德适依靠高毛利模型许可业务实现收入 21%稳健增长；晶泰控股受去年 DoveTree 合作带来的高基数影响，营收同比下滑 24%。利润端仅英矽智能实现盈利，主要受益于大额 BD，其余三家仍持续亏损。费用方面行业销售费用整体可控，但研发、管理费用刚性较强，营收规模是摊薄费用的重要抓手。当前 BD 带来的业绩脉冲具备阶段性，企业长期盈利仍有赖常态化订单落地与内部费用管控。

风险提示：下游需求波动，AI 迭代不及预期，AI4S 赛道毛利率分化。

科技	增持（维持）
计算机	增持（维持）

郭雅丽　研究员
SAC No. S0570515060003　guoyali@htsc.com
SFC No. BQB164　+(86) 21 3847 6016

范映蕊　研究员
SAC No. S0570521060004　fanyirui@htsc.com
SFC No. BWD469　+(86) 21 2897 2228

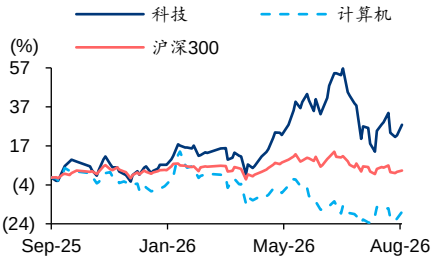
袁泽世*，PhD　研究员
SAC No. S0570524090001　yuanzeshi@htsc.com
+(86) 21 2897 2228

岳铂雄*　研究员
SAC No. S0570524080004　yueboxiong@htsc.com
+(86) 21 3847 6087

王浩天*　联系人
SAC No. S0570125010006　wanghaotian@htsc.com
+(86) 21 2897 2228

徐诚伟　联系人
SAC No. S0570125070089　xuchengwei@htsc.com
SFC No. BXO453　+(86) 21 2897 2228

行业走势图



资料来源：Wind，华泰研究

仅供内部参考，请勿外传

正文目录

2026 年中报总结：营收盈利增速边际放缓，分化趋势显现 4

整体业绩：营收增速放缓，利润承压 4

经营效率：毛利率承压，现金流下滑及人员扩张趋缓 5

大模型：长程 Agent、开放生态与商业化兑现 7

海外模型：长程任务成为能力升级主线 7

国内模型：后训练与架构效率同步突破 8

调用与 ARR：能力提升开始跨过可付费阈值 10

开放生态：竞争从权重发布延伸至全栈扩散 11

业绩变化：企业服务驱动国产模型收入放量 12

AI 应用：全球软件 FDE 加速，看好国内软件公司业绩放量 14

海外：闭源模型叙事趋缓，美股软件估值快速修复 14

国内：AI 商业化加速，AI 业务驱动营收稳步加速 15

展望：FDE 模型成为全球共识，看好国内软件公司业绩放量 17

算力租赁：收入保持高增，毛利率修复带动盈利弹性释放 18

样本公司收入利润及资产端情况 18

算力租赁板块收入利润快速兑现 19

算力租赁板块资产端进入快速扩展周期 20

AI4S：收入高速增长，大额 BD 驱动行业商业化提速 22

风险提示 23

图表目录

图表 1： 近年计算机板块营业收入同比增速（中位数法） 4

图表 2： 近年计算机板块扣非净利润同比增速（中位数法） 4

图表 3： 2026 年 H1 计算机板块营收增速前十企业 4

图表 4： 2026 年 H1 计算机板块扣非净利润增速前十企业 4

图表 5： 2026 年 H1 不同收入增速区间的公司数量统计 5

图表 6： 近年计算机板块毛利率情况（中位数法） 5

图表 7： 近年计算机板块销售、管理、研发费用率情况（中位数法） 5

图表 8： 近年计算机板块资产减值损失、信用减值损失占收入比情况（中位数法） 6

图表 9： 近年计算机企业经营性净现金流净额同比增速情况（中位数法） 6

图表 10： 近年计算机板块员工数量同比增速情况（中位数法） 6

图表 11： 海外主要模型第二季度更新总结 8

图表 12： 国内主要模型本季度更新总结 10

图表 13： 主要模型厂商近期年化收入指标 11

图表 14： 二季度代表性开放模型及生态层次 12

图表 15： MiniMax 分业务收入及盈利变化（千美元） 13

仅供内部参考，请勿外传

图表 16： OpenAI（绿色折线）、Anthropic（粉色折线）ARR 统计 14
图表 17： 中报业绩期（7 月底开始）IGV 指数修复 14
图表 18： AI 应用代表公司 26H1 营收基本呈修复趋势 16
图表 19： AI 转型带动软件公司盈利能力改善 16
图表 20： AI 应用代表公司关于 FDE 的核心表述 17
图表 21： 算力租赁研究样本公司收入利润情况 18
图表 22： 算力租赁研究样本公司资产情况 18
图表 23： 2022-2026H1 算力租赁板块营收及增速情况（中位数法） 19
图表 24： 2022-2026H1 算力租赁板块净利润及增速情况（中位数法） 19
图表 25： 2026H1 算力租赁板块营业收入增速统计 19
图表 26： 2022H1-2026H1 算力租赁板块毛利率情况（中位数法） 20
图表 27： 2022H1-2026H1 算力租赁板块费用率情况（中位数法） 20
图表 28： 2022-2026H1 算力租赁板块固定资产及增速（平均数法） 20
图表 29： 2022-2026H1 算力租赁板块使用权资产及增速（平均数法） 20
图表 30： 2022-2026H1 算力租赁板块在建工程及增速（平均数法） 21
图表 31： 2022H1-2026H1 算力租赁板块存货及增速（平均数法） 21
图表 32： AI4S 研究样本公司收入利润情况 22
图表 33： AI4S 研究样本公司毛利率及费用率情况 23

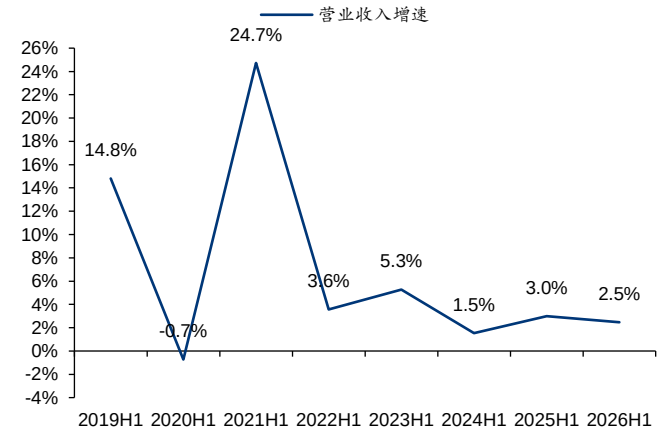
仅供内部参考，请勿外传

2026 年中报总结：营收盈利增速边际放缓，分化趋势显现

整体业绩：营收增速放缓，利润承压

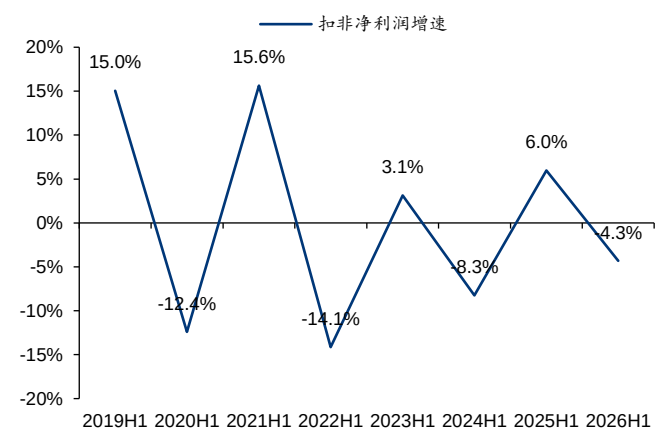
2026 年 H1 计算机板块营收同比增速略有回落，扣非净利润增速转负。剔除缺失值、极端值后，我们选取计算机板块共包含 330 家公司的样本，用中位数法进行分析。2026 年 H1 计算机板块营业收入同比增速中位数为 2.48%，收入端维持稳定增长态势。2026 年 H1 扣非净利润同比增速中位数为-4.32%，相较于 25 年 H1 转负。从个股表现看，收入增速前十公司均实现 50%以上增长，其中行云科技（+497.11%）、宏景科技（+156.43%）、直真科技（+150.07%）领跑，扣非净利润增速前十公司增幅均超过 350%，增长受益于业务结构调整与控费出清等。

图表1：近年计算机板块营业收入同比增速（中位数法）



资料来源：iFind，华泰研究

图表2：近年计算机板块扣非净利润同比增速（中位数法）



资料来源：iFind，华泰研究

图表3：2026 年 H1 计算机板块营收增速前十企业

证券代码	证券简称	收入增速	市值（亿元）
300209 CH	行云科技	497.11%	311.37
301396 CH	宏景科技	156.43%	378.80
003007 CH	直真科技	150.07%	46.66
300042 CH	朗科科技	101.76%	79.36
300270 CH	中威电子	90.98%	34.46
301208 CH	中亦科技	74.71%	36.59
300678 CH	中科信息	65.10%	65.50
300536 CH	农尚环境	61.37%	26.57
001339 CH	智微智能	56.64%	252.57
605118 CH	力鼎光电	54.39%	122.71

资料来源：iFind，华泰研究

图表4：2026 年 H1 计算机板块扣非净利润增速前十企业

证券代码	证券简称	扣非净利润增速	市值（亿元）
300277 CH	汽轮科技	71539.63%	180.93
600288 CH	大恒科技	3242.08%	49.40
301638 CH	南网数字	2735.30%	715.74
300302 CH	同有科技	1462.62%	146.42
688787 CH	海天瑞声	813.52%	56.72
600654 CH	中安科	623.42%	103.29
300042 CH	朗科科技	595.42%	79.36
301095 CH	广立微	572.41%	149.77
002990 CH	盛视科技	566.84%	122.55
300235 CH	方直科技	385.86%	27.42

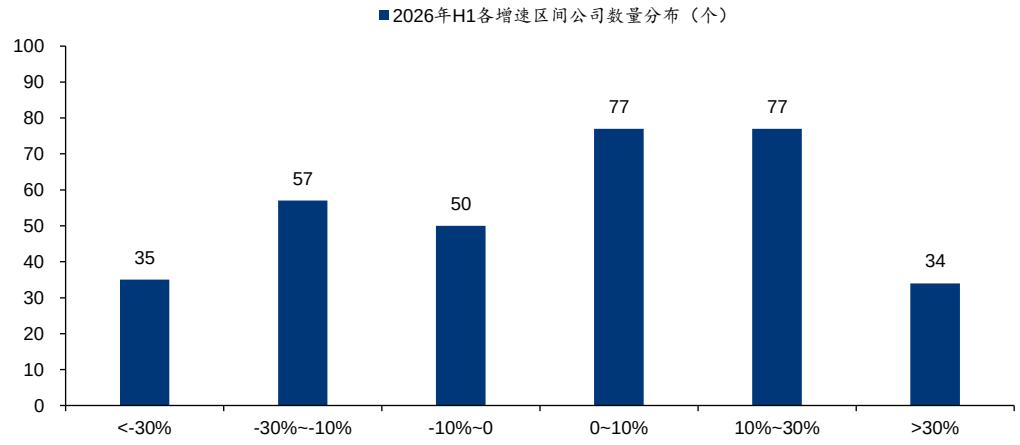
资料来源：iFind，华泰研究

2026 年 H1 计算机板块收入增速呈纺锤形分布。0～10%、10%～30%增速区间公司数量最多（均为 77 家），-30%～-10%其次（57 家），三者合计占比超过六成；负增长区间公司合计 142 家，占比约 43.03%，相比去年同期（43.54%）略有下降，但仍有超过四成的公司面临收入下滑压力；增速低于-30%（35 家）与高于 30%（34 家）的公司数量占比均超 10%，行业分化逐渐显现。

人在草木间

仅供内部参考，请勿外传

图表5：2026 年 H1 不同收入增速区间的公司数量统计

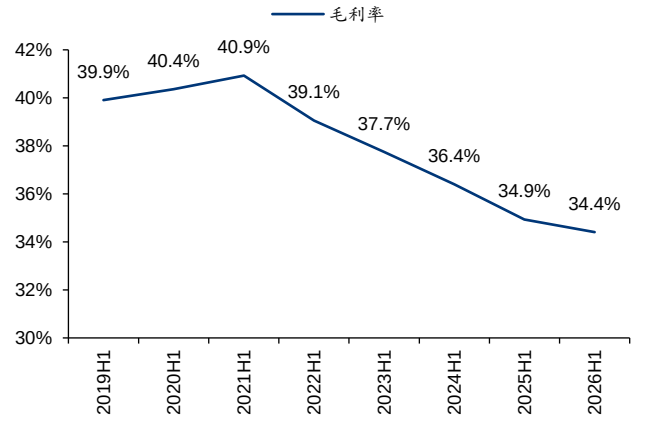


资料来源：iFind、华泰研究

经营效率：毛利率承压，现金流下滑及人员扩张趋缓

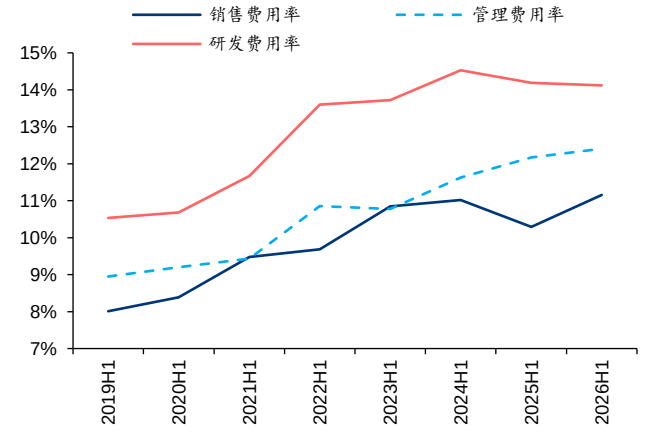
板块毛利率持续承压，期间费用率结构转差。2026 年 H1 板块毛利率为 34.4%，较 2025 年 H1 下降 0.5 个百分点。期间费用率方面，总体较 2025 年 H1 上升 1.03 个百分点，出现结构性抬升：销售费用率为 11.16%，同比上升 0.86 个百分点；管理费用率为 12.41%，同比上升 0.24 个百分点；研发费用率为 14.12%，同比下降 0.07 个百分点，研发强度略有下滑。

图表6：近年计算机板块毛利率情况（中位数法）



资料来源：iFind，华泰研究

图表7：近年计算机板块销售、管理、研发费用率情况（中位数法）

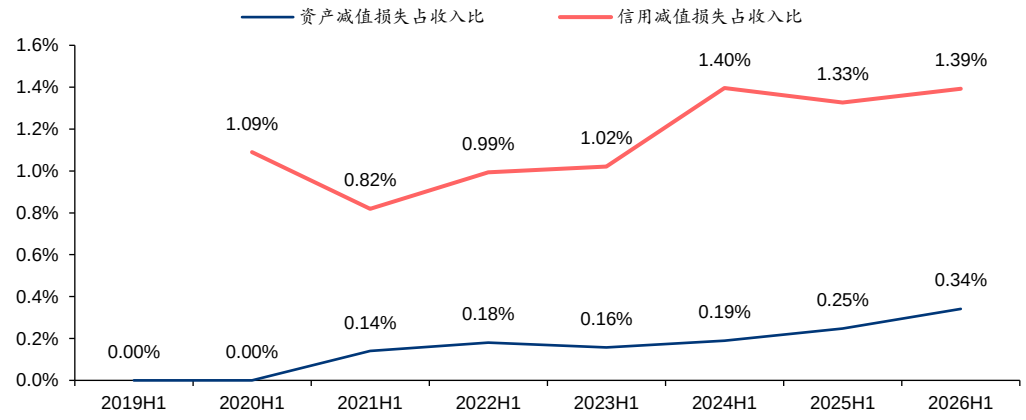


资料来源：iFind，华泰研究

资产减值压力持续增加，信用减值压力维持高位。资产减值损失占收入比逐年上升，在 2026 年 H1 达 0.34%，相较 2025 年 H1 扩大 0.09 个百分点，反映存货、固定资产等减值风险加速出清。信用减值损失占收入比维持在高位，2026 年 H1 达 1.39%，相对 2025 年 H1 增长 0.06 个百分点，相对 2024 年 H1 下降 0.01 个百分点，具备潜在冲回弹性。

仅供内部参考，请勿外传

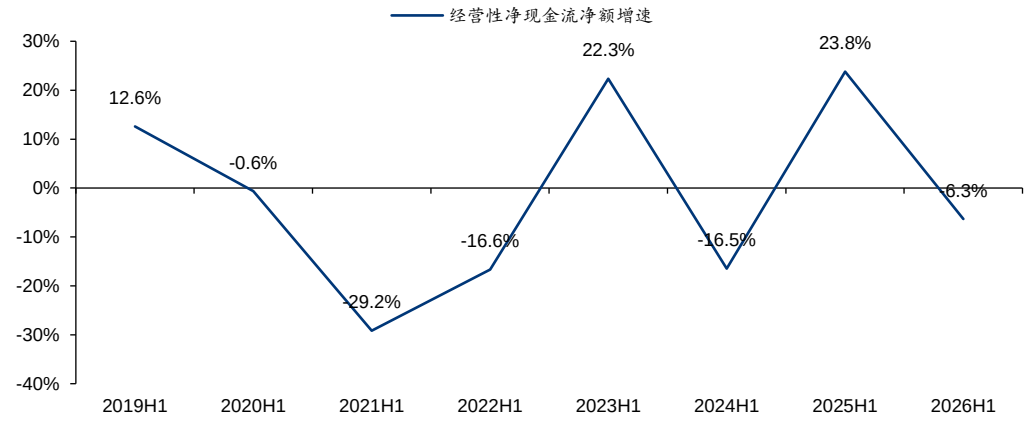
图表8：近年计算机板块资产减值损失、信用减值损失占收入比情况（中位数法）



资料来源：iFind、华泰研究

现金流边际回落，区间震荡明显。2026 年 H1，板块经营性现金流净额同比增速中位数为-6.3%，增速相较于 2025 年 H1 的 23.8%降幅明显。从时间上看，2023 年 H1 增速为 22.3%，2024 年 H1 则降至-16.5%，震荡明显，产业链上下游博弈加剧。

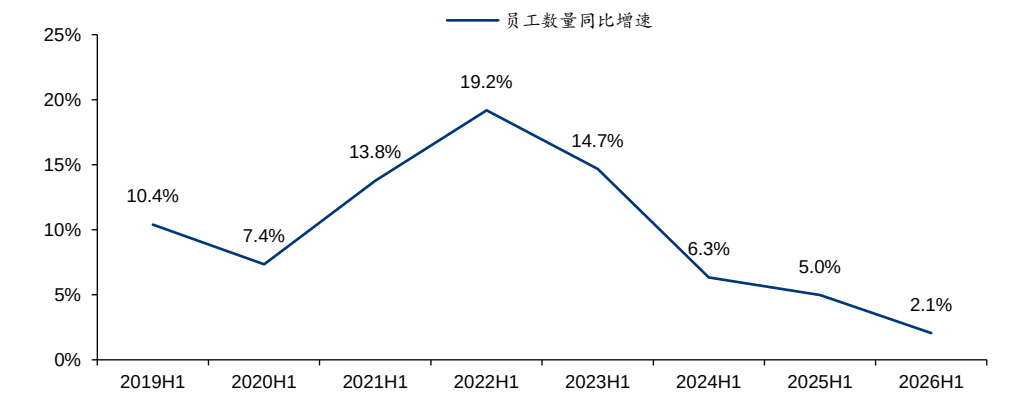
图表9：近年计算机企业经营性净现金流净额同比增速情况（中位数法）



资料来源：iFind、华泰研究

员工数量增速边际持续放缓。2026 年 H1 板块员工同比增速为 2.1%，相对 2025 年 H1 下降 2.9 个百分点。整体上，伴随 AI 应用的持续普及和控费转型的推进，行业的扩张节奏逐渐放缓，个人效率的提升成为后续发展的重点。2026 年 H1 增速持续下降，行业整体人员规模趋于稳定的态势仍在延续。

图表10：近年计算机板块员工数量同比增速情况（中位数法）



资料来源：iFind、华泰研究

仅供内部参考，请勿外传

大模型：长程 Agent、开放生态与商业化兑现

我们认为，二季度模型层出现六项结构性变化。1）长程 Agent 成为能力升级主线，模型价值越来越取决于持续规划、工具调用、错误修正和结果验证；2）后训练 Scaling 的重要性上升，同一预训练底座通过任务环境扩展和强化学习仍可获得显著能力增量；3）国内创新开始深入注意力、连接结构、KV Cache、Kernel 及 Serving 软件栈，能力追赶与推理效率改善同步发生；4）开放生态从权重下载扩展至许可证、Infra、框架、渠道和真实 Token 份额；5）OpenAI、Anthropic 与智谱、MiniMax 等公司的年化收入快速上升，为模型能力跨过可付费阈值提供了商业验证；6）开放模型追赶周期持续缩短，闭源模型在长任务稳定性、深层专业任务、Harness 产品化和算力正反馈上仍保持优势。

海外模型：长程任务成为能力升级主线

长程任务成为海外模型的共同主线。Anthropic、OpenAI、Google、Meta 和 xAI 二季度发布的模型均将能力重点放在 Agent Coding、工具调用、复杂工作流及长时间运行上，传统数学、知识问答等单点能力的权重边际下降。据 METR 2026 年 1 月 29 日发布的《Time Horizon 1.1》及截至 5 月 8 日更新的 Task-Completion Time Horizons 页面，2023 年以来模型 50%任务完成时长的翻倍周期约为 130.8 天，2024 年以来约为 88.6 天；截至 3 月 20 日，Claude Opus 4.6 的 50%任务完成时长约为 12 小时。据 Anthropic 2026 年 7 月 24 日官方 Blog《Introducing Claude Opus 5》，7 月发布的 Opus 5 继续强化长程多步骤任务、自我验证和迭代执行能力。当模型能够稳定执行数小时任务后，竞争重点会转向过程规划、工具协同、错误修正和结果验证，长任务完成率也更接近真实生产力。

Anthropic 继续强化长任务与受控能力释放。据 Anthropic 2026 年 6 月 9 日官方 Blog《Claude Fable 5 and Claude Mythos 5》，6 月发布的 Claude Fable 5 和 Mythos 5 重点提升长文本、长任务和 Agent Coding，其中 Mythos 5 仅向 Glasswing 合作伙伴开放；面向大众的 Fable 5 在网络安全、生物化学和模型蒸馏等方向设置能力限制，触发相关问题时切换至 Opus 4.8，平均触发率低于 5%。据 Anthropic 2026 年 7 月 24 日官方 Blog《Introducing Claude Opus 5》及同日发布的《Claude Opus 5 System Card》，7 月发布的 Opus 5 进一步改善编码、工具使用、自我检查和长任务能力，但系统卡同时指出，其尚未跨过自动化 AI 研发门槛，Opus 5 与 Mythos 5 的 AI 研发加速指标也基本相当。第二季度 Anthropic 的增量主要集中在长任务可靠性、受控能力释放和模型效率。

GPT-5.6 兼顾能力上限与 Token 效率。据 OpenAI 2026 年 7 月 9 日官方 Blog《GPT-5.6: Frontier intelligence that scales with your ambition》，OpenAI 于 7 月全面发布 GPT-5.6 系列，分为 Sol、Terra 和 Luna 三档，并为高复杂度任务提供 max 和 ultra 推理模式，其中 ultra 默认协调 4 个 Agent 并行探索、检查和修订方案；GPT-5.6 Sol 在 Agents' Last Exam 得分 53.6，较 Claude Fable 5 高 13.1 分，在 Artificial Analysis Coding Agent Index 中得分 80，较 Fable 5 高 2.8 分，同时输出 Token 不足后者一半、运行时间不足一半。

OpenAI 内部使用数据进一步验证了 Agent 需求的弹性。据 OpenAI 2026 年 7 月 9 日官方 Blog《GPT-5.6: Frontier intelligence that scales with your ambition》，过去 6 个月，用于内部编码推理的研究算力份额增长 100 倍，Agent Token 用量增长约 22 倍；GPT-5.6 测试期内，活跃研究人员日均输出 Token 较 GPT-5.5 历史峰值提高一倍以上。我们认为，能力提升会同时扩大单个任务的复杂度、并发 Agent 数量和持续运行时间，Token 需求可能由“人发起一次问答”转向“Agent 连续执行任务”。模型效率因而需要以单位有效任务的 Token 和时间评价，单纯压低每百万 Token 价格无法完整衡量经济性。

仅供内部参考，请勿外传

Google 将模型效率嵌入 Agent 能力升级。据 Google 2026 年 7 月 21 日官方 Blog《Introducing Gemini 3.6 Flash, 3.5 Flash-Lite, and 3.5 Flash Cyber》，Gemini 3.6 Flash 相较 3.5 Flash 减少 17%的输出 Token，在部分 DeepSWE 任务中 Token 消耗下降幅度可达 65%，DeepSWE 得分由 37%提升至 49%，MLE Bench 由 49.7%提升至 63.9%，OSWorld 由 78.4%提升至 83.0%；Gemini 3.5 Flash-Lite 输出速度达到 350Token/秒，Terminal-Bench 2.1 由 31%提升至 54%，长上下文得分由 60.1%提升至 72.2%，并支持调节思考等级及电脑操作；同期推出的 Flash Cyber 聚焦漏洞发现、验证和修复，并通过多模型协作形成安全报告。据 Google 2026 年 8 月 13 日官方 Blog《Introducing Gemini 3.7 Flash》及 ARC Prize 同日发布的 Gemini 3.7 Flash 评测结果，Gemini 3.7 Flash 较 3.6 Flash 每百万 Token 便宜 50%，ARC-AGI-2 得分 84.6%，单项任务成本为 0.25 美元。

Google 同步探索更轻量的开放生成架构。据 Google Developers Blog 2026 年 6 月 3 日《Gemma 4 12B: The Developer Guide》，Gemma 4 12B 采用统一、无编码器的多模态架构，以轻量嵌入模块替代视觉编码器，并将原始音频直接投射至 Token 空间；模型采用 Apache 2.0 许可，约 16GB 显存即可本地运行。据 Google Developers Blog 2026 年 6 月 10 日《DiffusionGemma: The Developer Guide》，DiffusionGemma 将文本生成由逐 Token 自回归扩展为对整段内容进行扩散式迭代细化，可在生成过程中同步起草、纠错和填补空白，生成速度超过 1,000Token/秒，并可在 18GB 消费级显卡运行。两条路径分别降低多模态编码和序列生成的系统负担，开放模型也开始从能力追赶转向架构创新和本地实时部署。

Meta 与 xAI 集中补强真实软件工程能力。据 Meta 2026 年 8 月 5 日官方发布的《Introducing Muse Code and Muse Spark 1.2》，Muse Spark 1.2 针对多文件重构、长时间调试和完整代码库生成进行训练，通过规划、目标条件和上下文压缩维持数小时任务方向，同时支持异步及并行工具调用，并将上下文扩展至 100 万 Token。据 SpaceXAI 2026 年 7 月 16 日官方 Blog《Introducing Grok 4.5》，Grok 4.5 在 Terminal-Bench 2.1 得分 83.3%，接近 GPT-5.5 xhigh 的 83.4%；在 SWE-Bench Pro 得分 64.7%，低于 Fable 5 的 80.4%，但平均输出 15,954Token，约为 Opus 4.8 max 的四分之一。据 SpaceXAI 2026 年 8 月 12 日官方 Blog《Introducing Grok 4.6》，Grok 4.6 进一步聚焦长程 Agent，Artificial Analysis Intelligence Index 得分 61，与 GPT-5.6 Sol 同为 61；CursorBench 3.2 得分 69.9%，DeepSWE 1.1 得分 65.9%。

图表11：海外主要模型第二季度更新总结

厂商	代表模型	本季度主要变化	关键数据
Anthropic	Claude Fable 5 / Mythos 5 / Opus 5	长任务、Agent Coding 与能力治理。	Fable 5 受限能力平均触发率低于 5%；Opus 5 强化工具使用和自我检查。
OpenAI	GPT-5.6 Sol / Terra / Luna	多 Agent 协作、Token 效率与网络安全同步推进。	Sol 在 Agents' Last Exam 得 53.6，Coding Agent Index 得 80；内部 Agent Token 过去 6 个月增长约 22 倍。
Google	Gemini 3.6 / 3.7 Flash；Gemma 4	Flash 系列提效并强化 Agent；Gemma 推进轻量开放多模态。	3.6 Flash 输出 Token 减少 17%；部分 DeepSWE 任务 Token 消耗降幅可达 65%；3.7 Flash 价格低 50%。
Meta	Muse Spark 1.2 / Muse Glimmer	多文件重构、长调试及并行工具调用。	Muse Spark 1.2 支持 100 万 Token 上下文；Muse Glimmer 为 30B 密集模型。
xAI	Grok 4.5 / 4.6	真实软件工程与长程 Agent 继续迭代。	Grok 4.5 在 Terminal-Bench 2.1 得 83.3%，SWE-Bench Pro 得 64.7%，平均输出 15,954Token。

资料来源：各公司官网、华泰研究

国内模型：后训练与架构效率同步突破

后训练成为国产模型能力提升的重要抓手。据 Z.ai 2026 年 8 月 14 日官方 Blog《GLM-5.3: Frontier Coding with Emergent Cyber Capabilities》及 DeepSeek API 2026 年 7 月 31 日《DeepSeek-V4-Flash 更新》，GLM-5.3 沿用 GLM-5.2 的基座模型和参数规模，DeepSeek-V4-Flash 正式版也与预览版保持相同架构和规模，两者的主要增量均来自更长周期任务环境、Agent 轨迹和强化学习。

仅供内部参考，请勿外传

GLM-5.3 验证了后训练 Scaling 的有效性。据 Z.ai 2026 年 8 月 14 日官方 Blog《GLM-5.3: Frontier Coding with Emergent Cyber Capabilities》，该模型沿用 GLM-5.2 基座模型，通过扩大长程任务环境、丰富环境类型及增加后训练计算提升能力，并采用 IndexShare、单次 Rollout 异步优化 SAO 及新一代 Slime 强化学习框架；Z.ai Code Bench 较 GLM-5.2 提升约 50%，Terminal-Bench 3.0 由 4.6 提升至 28.3，DeepSWE v1.1 由 46.2 提升至 66.9，Agents' Last Exam 由 23.8 提升至 28.5；在 Z.ai Code Bench High 档设置中得分 31.4%，高于 Claude Opus 4.8 同档设置的 29.5%，平均每项任务约使用 5 万 Token，后者约为 12 万 Token。

GLM-5.3-Flash 把开放、效率与多模态结合到同一模型。据 Z.ai 2026 年 8 月 26 日官方 Blog《GLM-5.3-Flash: Frontier Intelligence, Flash Cost》，该模型采用 320B 总参数、18B 激活参数的 MoE 架构，使用 30 万亿 Token 多模态预训练语料，是 GLM-5 系列首个原生多模态模型，并将稀疏注意力与线性注意力、流形约束超连接 mHC 及 IndexPool 结合；相较 GLM-5.3，其 Attention 计算量降低 3 倍、KV Cache 降低 4.4 倍，IndexPool 通过加权池化将索引器 4 个缓存向量压缩为 1 个，进一步降低 100 万 Token 上下文下的时延和内存开销；在同一国产芯片硬件上通过 Kernel 和 Serving 软件栈优化，端到端服务性能提升 3 倍。据 Z.ai 上述官方 Blog 及 OpenRouter 2026 年 8 月 26 日前后公开的模型使用榜数据，模型先以 Ox Alpha 匿名上线接受真实流量测试，6 天处理约 42 万亿 Token，随后公开身份并开放权重，形成“真实调用验证—系统优化—权重发布”的新路径。

MiniMax M3 以稀疏注意力支撑超长任务。据 MiniMax 2026 年 6 月 1 日官方 Blog《MiniMax M3: Frontier Coding, 1M Context, Native Multimodality — All in One Model》，M3 采用 428B 总参数、23B 激活参数的 MoE 架构，从预训练 Step 0 开始混合文本、图像等多模态数据，并支持 100 万 Token 上下文；MiniMax Sparse Attention 使 100 万上下文下每 Token 计算量约为上代模型的 1/20，Prefill 和 Decode 阶段分别实现超过 9 倍和 15 倍加速。该篇官方 Blog 展示的长任务案例显示，M3 连续运行接近 12 小时复现论文结果，产出 18 次 Commit 和 23 张实验图；另一项 Kernel 优化任务连续执行约 24 小时，完成 147 次 Benchmark 提交和 1,959 次工具调用。

Kimi K3 以更大 MoE 规模推进开放模型前沿。据 Kimi 2026 年 7 月 16 日官方 Blog《Kimi K3: Open Frontier Intelligence》及 7 月 27 日发布的同名技术报告，K3 总参数达到 2.8 万亿、激活参数 104B，设置 896 个专家、每 Token 激活 16 个专家，支持 100 万 Token 上下文和原生多模态理解。据 Kimi 2026 年 7 月 27 日技术报告《Kimi K3: Open Frontier Intelligence》，K3 采用 Kimi Delta Attention，将百万 Token 解码加速 6.3 倍，并以 Attention Residuals 替代传统残差连接，训练效率提升约 25%。据 Arena.ai 2026 年 7 月 16 日 Frontend Code Arena 公开榜单，K3 以 1,679 分位列榜单首位，两两比较胜率 76%，高于 Fable 5 的 63%和 GPT-5.6 Sol 的 58%。

Kimi K3 同时揭示了超大开放模型的部署约束。据 Artificial Analysis 截至 2026 年 8 月 30 日 Kimi K3 模型页面的模型规模测算，K3 权重本身约需 1,570GB 显存，尚未计入 KV Cache，无法放入单台 Hopper GPU 节点；据 Kimi 2026 年 7 月 27 日技术报告《Kimi K3: Open Frontier Intelligence》，896 个专家还需要 WideEP 将权重分布到多张 GPU，并在每层进行 Token 分发和汇总，Kimi Delta Attention 可以显著降低 KV Cache 传输，但对权重存储、专家并行和节点间通信仍提出较高要求。因此，开放权重降低了获取门槛，生产级部署成本仍可能把大量用户推向云端托管和按量 API，开放模型的商业化与推理基础设施将长期绑定。

仅供内部参考，请勿外传

DeepSeek 围绕 V4 底座形成 Flash、Pro 分层。据 DeepSeek API 2026 年 7 月 31 日更新日志《DeepSeek-V4-Flash 更新》，DeepSeek-V4-Flash 正式版与预览版保持 284B 总参数、13B 激活参数的 MoE 架构，支持 100 万 Token 上下文，升级重点集中在编程、工具调用和 Agent 任务，并原生支持 Responses API；同一更新日志显示，其多项 Agent Benchmark 已超过此前的 V4-Pro Preview，体现轻量模型通过后训练反超旧版大模型的可能。据 DeepSeek 2026 年 8 月 13 日官方公告《DeepSeek-V4-Pro 正式版上线》，V4 Pro 采用 1.6T 总参数、49B 激活参数并以 MIT 许可开放权重，进一步加强规划、自测和专业知识任务；据 Artificial Analysis 截至 2026 年 8 月 30 日更新的《LLM Leaderboard》，其智力指数为 53，较 4 月版本提升 8 分。

生成式多模态模型继续向开放生态扩散。据 MiniMax 2026 年 7 月 31 日官方 Blog《MiniMax H3: An Open Model Breaking the Boundaries Between Tasks and Modalities》，MiniMax 公开 MiniMax-H3 模型权重，并随后发布覆盖量化、本地部署、LoRA 及多 GPU 生产栈的生态集成。据 Arena.ai 2026 年 8 月 13 日发布的 Video Edit Arena 榜单，H3 得分 1,390，领先第二名 32 分；同期多图生视频公开榜单显示其位列第二，并成为排名靠前的开放权重视频模型。据 MiniMax 2026 年 8 月公开的 H3 生态集成信息，模型发布后衍生接近 300 个社区模型，ComfyUI 版本下载量接近 700 万，说明视频模型开放后可通过量化、工作流和微调快速扩展应用边界。

图表12： 国内主要模型本季度更新总结

厂商	代表模型	技术/产品主线	关键数据
Z.ai	GLM-5.3 / GLM-5.3-Flash	后训练 Scaling、网络安全、稀疏+线性注意力及 Serving 优化。	Flash 为 320B 总参数/18B 激活；30 万亿 Token 预训练；Attention 计算量降低 3 倍、KV Cache 降低 4.4 倍。
MiniMax	M3 / H3	稀疏注意力支撑超长任务，开放生态扩展至视频。	M3 为 428B 总参数/23B 激活，支持 100 万 Token 上下文；Prefill/Decode 加速超过 9 倍/15 倍。
Kimi	K3	更大 MoE、Delta Attention 与开放 Infra。	总参数 2.8 万亿、激活 104B、896 个专家；100 万 Token 上下文；百万 Token 解码加速 6.3 倍。
DeepSeek	V4 Flash / Pro / Vision	形成轻量、高性能与视觉分层，并强化 Agent 和 Responses API。	Flash 为 284B/13B 激活；Pro 为 1.6T/49B 激活；均围绕 100 万 Token 上下文扩展。

资料来源：各公司官网、华泰研究

调用与 ARR：能力提升开始跨过可付费阈值

海外前沿模型的收入增速与 Agent 能力升级同步加快。据 Epoch AI 2026 年 8 月 26 日更新的《Data on AI Companies》及 8 月 27 日发布的《An update on AI's most important number》，Anthropic 年化收入运行率由 2025 年底约 90 亿美元升至 2026 年 2 月 140 亿美元、4 月超过 300 亿美元、5 月超过 470 亿美元；据 Reuters 2026 年 8 月 17 日报道《Anthropic revenue run rate tops $65 billion, source says》，其 7 月底年化收入运行率达到 650 亿美元。据 Epoch AI 上述数据集，OpenAI 年化收入运行率由 2025 年底约 214 亿美元升至 2 月底 250 亿美元；据 Bloomberg 2026 年 8 月 13 日报道《OpenAI's Revenue Run Rate Tops $40 Billion Ahead of IPO》，其 8 月年化收入运行率超过 400 亿美元。据 Epoch AI 2026 年 8 月 27 日《An update on AI's most important number》，两家公司合计年化收入运行率由年初约 300 亿美元升至约 1,050 亿美元，8 个月约增长 2.5 倍、规模达到年初的 3.5 倍。至 2026 年 8 月，两家公司在高基数下仍维持高增速，说明编码 Agent 与复杂知识工作流可能已经跨过新的能力—付费阈值。

Anthropic 的增长更直接体现模型能力、Harness 与按量计费的联动。据 SemiAnalysis 2026 年 7 月 8 日发布的《Anthropic 3Q26 Profit Over $1B: The Anthropic IPO Financials Sneak Peak》，Claude Code 把模型与工具链组合成可持续完成软件工程任务的产品，并推动 Anthropic 形成以企业 API 和使用量计费为主的收入结构；该报道估算约 75%—85% 的收入来自 usage-based API，消费订阅仅约占 ARR 的 5%。按 Token 计费下，同一客户将模型用于更多代码库、更长任务和更多并发 Agent 即可继续扩大收入，模型推理效率、吞吐和任务成功率会直接影响 ARR 与毛利率。OpenAI 历史收入更偏消费订阅，固定月费下高 Token 用户可能压低单位经济性，因此两家公司的收入质量和模型使用深度不能简单等同。

仅供内部参考，请勿外传

国内模型的商业验证开始出现，但财务披露仍明显少于海外。据 MiniMax 2026 年 8 月 26 日 2026 年中期业绩电话会及财新网 8 月 27 日报道《MiniMax 上半年收入 1.17 亿美元超 2025 全年 8 月 ARR 达 8 亿美元》，公司 8 月 ARR 达到 8 亿美元，CEO 闫俊杰表示增长主要由模型能力改善驱动；公司在同次业绩沟通中披露，过去 6 个月推出 7 个大模型，企业及开发者客户数量较 6 个月前增长超过 5 倍。据 Bloomberg 2026 年 7 月 20 日报道《Moonshot Narrowed the US's Evaporating Lead in AI》，月之暗面 ARR 由 4 月超过 2 亿美元升至 6 月 3 亿美元，K3 发布后日销售额增长约 6 倍。据 The Information 2026 年 7 月 14 日报道《DeepSeek's Annualized Revenue Nears $500 Million, Boosting Fundraise and IPO Plans》，DeepSeek 7 月年化收入运行率约 4 亿—5 亿美元，收入主要来自模型 API。

智谱已形成 ARR、API 用户与开发者产品同步扩张的商业化信号。据《证券日报》2026 年 7 月 17 日报道《智谱年度经常性收入达到 10 亿美元》，截至 7 月智谱 ARR 达到 10 亿美元。据《晚点 LatePost》2026 年 8 月 10 日报道《晚点独家丨智谱 API 用户数近 700 万，新启用超 5 万块国产算力芯片》，智谱 2026 年年内 ARR 增长 15 倍；MaaS 开放平台注册用户近 700 万，较 7 月初增加约 200 万，其中包括 2.3 万家企业客户，ZCode 上线 1 个月用户突破 100 万，同期已启用超 5 万块国产算力芯片。我们认为，个人用户、企业客户与算力投入同步增长，为 Coding 产品与 API 调用扩张提供了商业化验证。

图表13：主要模型厂商近期年化收入指标

公司	时点	指标	规模	披露性质与说明
Anthropic	2026 年 7 月底	年化收入运行率	>650 亿美元	据 Reuters 2026 年 8 月 17 日报道《Anthropic revenue run rate tops $65 billion, source says》；5 月超过 470 亿美元亦见该报道。
OpenAI	2026 年 8 月	年化收入运行率	>400 亿美元	据 Bloomberg 2026 年 8 月 13 日报道《OpenAI's Revenue Run Rate Tops $40 Billion Ahead of IPO》；Epoch AI 2026 年 8 月 26 日《Data on AI Companies》同步收录。
MiniMax	2026 年 8 月	ARR	8 亿美元	据 MiniMax 2026 年 8 月 26 日 2026 年中期业绩电话会及财新网 8 月 27 日报道《MiniMax 上半年收入 1.17 亿美元超 2025 全年 8 月 ARR 达 8 亿美元》。
月之暗面	2026 年 6 月	ARR	3 亿美元	据 Bloomberg 2026 年 7 月 20 日报道《Moonshot Narrowed the US's Evaporating Lead in AI》；4 月超过 2 亿美元、6 月 3 亿美元。
DeepSeek	2026 年 7 月	年化收入运行率	4 亿—5 亿美元	据 The Information 2026 年 7 月 14 日报道《DeepSeek's Annualized Revenue Nears $500 Million, Boosting Fundraise and IPO Plans》；主要来自模型 API。
智谱 AI	2026 年 7 月	ARR	10 亿美元	据《证券日报》2026 年 7 月 17 日报道《智谱年度经常性收入达到 10 亿美元》。

资料来源：路透社、彭博、The Information、各公司官网、证券日报、华泰研究

开放生态：竞争从权重发布延伸至全栈扩散

“开放模型”的竞争内涵明显扩大。第一层是权重可下载和本地部署；第二层是许可证决定商业使用边界；第三层是注意力内核、训练方法、推理框架与硬件适配决定部署效率；第四层是聚合平台、云渠道、衍生模型和开发者工具决定真实使用规模。二季度国内模型在后两层的推进尤其明显：MiniMax 开放 MSA 内核并获得 SGLang Day-0 支持，Kimi 同步开放关键 Infra 技术，DeepSeek 采用 MIT 许可并推出插件化 Harness，GLM-5.3-Flash 则以匿名真实流量完成服务验证后再开放权重。开放程度需要按这四层拆解，单看是否能在 Hugging Face 下载会遗漏主要变化。

仅供内部参考，请勿外传

图表14：二季度代表性开放模型及生态层次

模型	开放内容	许可证/部署	生态证据
Gemma 4 12B	模型权重与统一多模态架构	Apache 2.0；约 16GB 显存本地运行	据 Google Developers Blog 2026 年 6 月 3 日《Gemma 4 12B: The Developer Guide》；面向消费级本地多模态。
Muse Glimmer	30B 密集模型权重	可本地运行	据 Meta Superintelligence Lab 2026 年 8 月 10 日《Muse Glimmer Model Card》；补充 Meta 轻量开放模型层。
GLM-5.3-Flash	权重、稀疏+线性注意力架构	320B-A18B；100 万 Token 上下文	据 Z.ai 2026 年 8 月 26 日《GLM-5.3-Flash: Frontier Intelligence, Flash Cost》及 OpenRouter 同期公开榜单；匿名上线 6 天处理约 42T Token。
MiniMax M3	权重、MSA 内核与技术论文	自定义许可；小规模商业使用设报备与署名条件	据 MiniMax 2026 年 6 月 1 日《MiniMax M3: Frontier Coding, 1M Context, Native Multimodality — All in One Model》；SGLang Day-0 支持 NVIDIA 与 AMD。
Kimi K3	权重、技术报告、关键 Infra	2.8T；生产部署需要机架级高带宽系统	据 Kimi 2026 年 7 月 16 日《Kimi K3: Open Frontier Intelligence》及 7 月 27 日同名技术报告；开放内容覆盖模型与基础设施。
DeepSeek V4 Pro	模型权重	MIT；1.6T 总参/49B 激活	据 DeepSeek 2026 年 4 月 24 日《DeepSeek-V4 预览版：迈入百万上下文普惠时代》及 8 月 13 日《DeepSeek-V4-Pro 正式版上线》；与 Flash、Vision 及插件化 Harness 形成分层。
MiniMax H3/Music3	视频与音乐模型权重	覆盖本地量化、LoRA 和多 GPU 部署	据 MiniMax 2026 年 7 月 31 日《MiniMax H3: An Open Model Breaking the Boundaries Between Tasks and Modalities》及截至 8 月 30 日 Hugging Face、ComfyUI 公开页面统计；H3 衍生近 300 个模型，ComfyUI 下载近 700 万。

资料来源：各公司官网、华泰研究

开放模型已开始获得真实 Token 份额。据 Vercel CEO Guillermo Rauch 2026 年 8 月 22 日在 X 发布的 AI Gateway 实时流量数据，开放模型 Token 占整体 Token 的份额在两个月内由 6 月的 28%升至 62%；该指标反映 Vercel AI Gateway 平台的生产调用结构，显示开发者在实际调用中提高了开放模型使用比例。据 SemiAnalysis 2026 年 8 月 21 日发布的《Are Open Models Catching Up?》，开放模型服务商 Fireworks 在 8 月的日处理量超过 40 万亿 Token，约为 3 月底 OpenAI API 流量的两倍。另据 Z.ai 2026 年 8 月 26 日官方 Blog《GLM-5.3-Flash: Frontier Intelligence, Flash Cost》及 OpenRouter 同期公开的模型使用榜数据，GLM-5.3-Flash 以 Ox Alpha 身份上线后 6 天处理约 42 万亿 Token 并登顶 OpenRouter 使用榜；SemiAnalysis 2026 年 8 月 21 日的报道《Are Open Models Catching Up?》同时称其 8 月服务规模达到每日 100 万亿 Token，且全部由国产芯片承载。结合 H3 的衍生模型和下载量，开放生态的衡量标准已经由“是否发布”扩展至实际调用、社区改造和多硬件部署。

开放权重并不天然对应低部署成本。据 Artificial Analysis 截至 2026 年 8 月 30 日 Kimi K3 模型页面的模型规模测算，Kimi K3 需要约 1.57TB 显存保存权重，并依赖高带宽专家并行。据 MiniMax 2026 年 6 月 1 日《MiniMax M3: Frontier Coding, 1M Context, Native Multimodality - All in One Model》、Kimi 2026 年 7 月 27 日《Kimi K3: Open Frontier Intelligence》、Google Developers Blog 上述 Gemma 4 12B 公告及 DeepSeek 2026 年 8 月 13 日《DeepSeek-V4-Pro 正式版上线》披露的许可证信息，MiniMax M3、Kimi K3 等模型采用自定义许可证，商业规模扩大后还可能触发报备、署名或单独授权，而 Gemma 4 与 DeepSeek V4 Pro 分别采用 Apache 2.0 和 MIT 许可，商业边界相对更宽。开放模型在 API 端可以依靠多家服务商竞争压低价格，企业若选择自建，仍要承担推理优化、并发调度、硬件折旧和运维成本。因此，开放模型对闭源厂商的压力首先体现为 Token 价格、渠道议价和开发者入口，距离模型层全面商品化仍有较长路径。

业绩变化：企业服务驱动国产模型收入放量

MiniMax 2026H1 收入增长较快，企业服务成为主要增量。据 MiniMax2026 年 8 月 26 日发布的中期业绩公告，2026H1 公司实现收入 1.17 亿美元，同比增长 283.1%，已超过 2025 年全年的 0.79 亿美元；其中 AI 原生产品收入 0.43 亿美元，同比增长 100.9%，占收入 36.6%；开放平台及其他基于 AI 的企业服务收入 0.74 亿美元，同比增长 703.1%，收入占比由 2025H1 的 30.3%提升至 63.4%。公司在业绩会披露，开放平台及其他基于 AI 的企业服务收入增长主要来自付费个人用户及企业用户数量增长、API 调用量增加以及 Token 计划推广，收入结构进一步向 API 及企业工作负载倾斜。

仅供内部参考，请勿外传

MiniMax 毛利率改善，研发投入仍维持较高强度。据 MiniMax 2026 年 8 月 26 日发布的中期业绩，2026H1 公司实现毛利 2,081.3 万美元，同比增长 464.8%，毛利率由 2025H1 的 12.1%提升至 17.9%，公司披露主要受到基础设施效率提升推动；同期研发开支 2.97 亿美元，同比增长 138.8%，低于收入增速。期内亏损 3.58 亿美元，同比收窄 11.0%；经调整净亏损 2.93 亿美元，同比扩大 111.2%，主要由于研发投入持续增长。此外，中国内地以外地区收入 7,082.8 万美元，占总收入 60.8%。我们认为，企业服务占比提升与毛利率改善共同验证 API 调用规模扩张的商业化价值，但研发投入和经调整亏损仍处于较高水平。

图表15： MiniMax 分业务收入及盈利变化（千美元）

项目	截至 12 月 31 日止年度			截至 6 月 30 日止六个月	
	2023	2024	2025	25H1	26H1
AI 原生产品					
收入	**758**	**21,805**	**53,075**	**21,223**	**42,644**
yoy		2,776.6%	143.4%		100.9%
%收入	21.9%	71.4%	67.2%	69.7%	36.6%
毛利	(2,882)	(1,776)			
毛利率	-380.2%	-8.1%			
AI 企业服务					
收入	**2,702**	**8,718**	**25,963**	**9,206**	**73,929**
yoy		222.6%	197.8%		703.1%
%收入	78.1%	28.6%	32.8%	30.3%	63.4%
毛利	2,028	5,514			
毛利率	75.1%	63.2%			
合计					
总收入	**3,460**	**30,523**	**79,038**	**30,429**	**116,573**
yoy		782.2%	158.9%		283.1%
总毛利	(854)	3,738	20,079	3,685	20,813
yoy		转正	437.2%		464.8%
总毛利率	**-24.7%**	**12.2%**	**25.4%**	**12.1%**	**17.9%**

资料来源：Wind、华泰研究

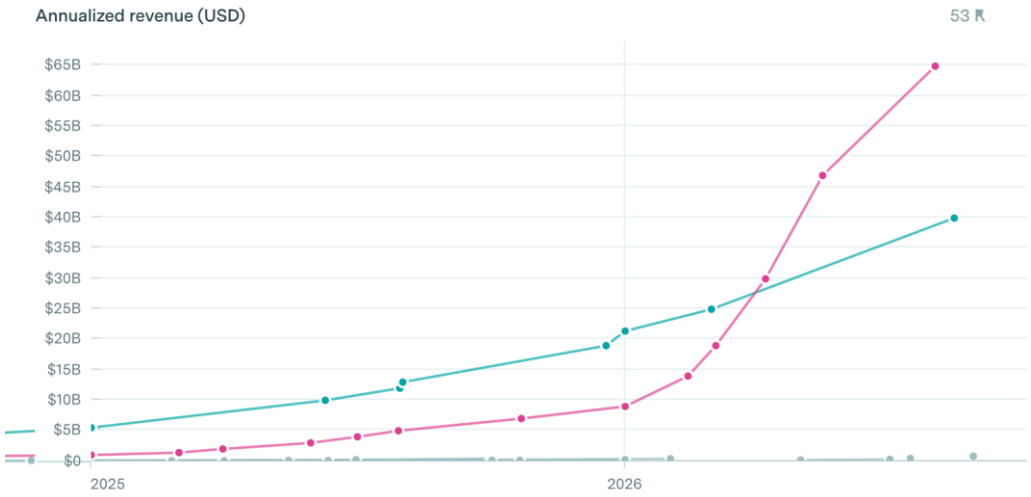
仅供内部参考，请勿外传

AI 应用：全球软件 FDE 加速，看好国内软件公司业绩放量

海外：闭源模型叙事趋缓，美股软件估值快速修复

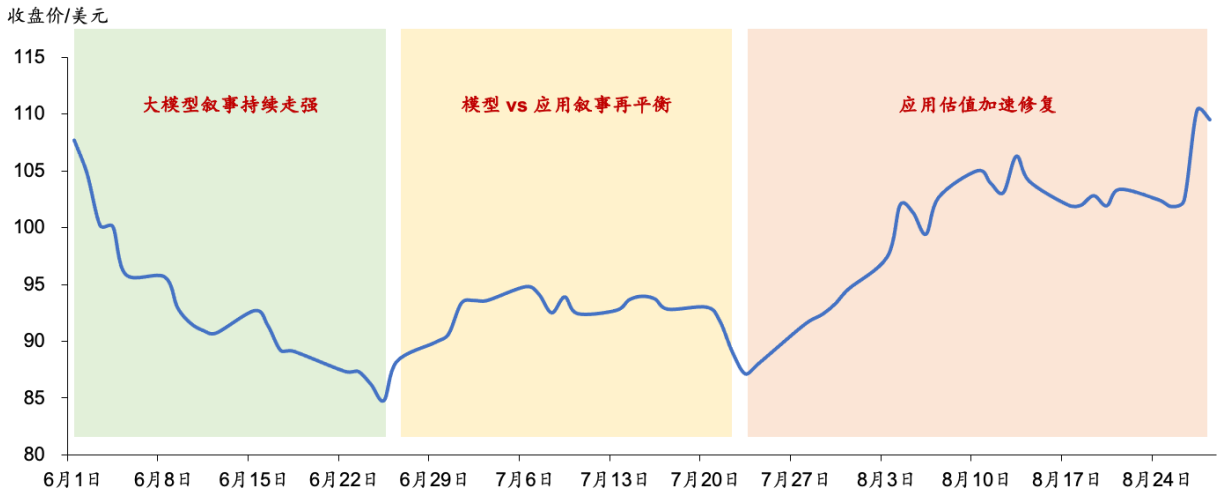
闭源模型 ARR 增速趋缓，美股软件估值压力暂缓，7 月底实现估值+业绩共振修复。25 年底，海外 Anthropic ARR 走强，带动“模型吞噬软件”悲观预期加重，市场担心通用大模型对软件短期业绩及长期成长空间造成侵蚀，导致美股软件估值持续承压。26 年 6-7 月，市场开始讨论大模型 ARR 是否存在增速放缓风险，美股软件进入估值压力暂缓的过渡期。26 年 7 月下旬，美股软件进入二季报业绩披露期，Microsoft、Palantir、Cloudflare、Salesforce 等美股软件业绩超 VisibleAlpha 一致预期，同时叠加模型预期相对走弱，美股软件进入估值+业绩的共振修复期，美股 IGV 指数快速上涨，截至 26 年 8 月 28 日，IGV 指数较 7 月 24 日已上涨约 24%。

图表16：OpenAI（绿色折线）、Anthropic（粉色折线）ARR 统计



资料来源：EPOCH AI 官网、华泰研究

图表17：中报业绩期（7 月底开始）IGV 指数修复



资料来源：Wind、华泰研究

仅供内部参考，请勿外传

国内：AI 商业化加速，AI 业务驱动营收稳步加速

国内软件 AI 商业化进展提速，驱动 H1 营收整体稳步加速。根据各公司披露的中报业绩，港股、A 股 AI 应用典型代表的 AI 商业化均实现同比较快增长。同时 AI 业务正逐步成为部分公司的新增长驱动，带动公司整体营收呈现稳步修复趋势。

1）金蝶国际（港股）：26H1 总营收增速为 13.56%（去年同期为 11.24%），AI 业务驱动营收稳步加速。公司 AI 原生产品目前已经形成“合同—收入—订阅”的完整验证链条，26H1 公司 AI 原生产品合同金额 6.05 亿元，同比增长 159.2%；26H1 公司 AI 原生产品收入 2.96 亿元，同比增长 189.0%，其中订阅收入 1.46 亿元，同比增长 282.9%。

2）迅策（港股）：26H1 总营收增速为 388.78%（去年同期为-29.98%），营收大幅增长主要受益于 AI 业务快速放量。公司 26H1 完成 Token 商业模式的初步验证，Token 业务收入占总收入超 10%；同时公司已形成覆盖“算力—数据—Token—模型—应用”完整布局，公司 ARPU 由 25H1 的 164 万元提高至 556 万元，客户留存率超过 90%。

3）迈富时（港股）：26H1 总营收增速为 111.18%（去年同期为 25.54%），公司加速从 AI 应用平台升级为“全栈 Token 工厂”，带动营收加速增长。26H1 公司 AI 应用业务营收 11.28 亿元，同比增长 123.7%，占总收入 57.5%；26H1 公司付费用户 27,269 家，同比增长 25.9%，月均 ARPU 同比增长 80%至 8,124 元。

4）滴普科技（港股）：26H1 总营收增速为 114.97%（去年同期为 118.36%），营收维持较高增速主要受益于 AI 业务持续推进。公司将 FastAGI 升级为 DeepexiOS 企业级 AI 操作系统，26H1 公司 DeepexiOS 业务营收 2.26 亿元，同比增长 209.2%，占总收入 79.6%，毛利率 57.2%，同比提升约 2pct，AI 业务已经成为公司主体收入来源。

5）汉得信息（A 股）：26H1 总营收增速为 6.07%（去年同期为 3.54%），公司营收稳步修复主要得益于客户 AI 需求持续放量。公司围绕“得·灵”构建企业 AI 体系，推出“灵猿—大圣”AI 中台及 DataClaw“灵舟”企业数字员工平台，连接 ERP、CRM、MES 并执行长流程任务。26H1 公司企业级 AI 智能化应用收入约 2.2 亿元，同比增长超 100%；其中 Q1 约 0.81 亿元、Q2 约 1.4 亿元，Q2 环比增长超过 70%。

6）鼎捷数智（A 股）：26H1 总营收增速为 5.37%（去年同期为 4.08%），公司加速推动 AI 产品升级，发布企业智能运行空间 EIOSpace，覆盖 9 个业务域、100 项制造业高频场景；26H1 公司 AI 业务签约金额约为 2 亿元，上半年签约金额已与 25 年全年签约金额持平，公司 AI 业务加速实现从试点验证到规模化落地的关键跨越。

7）能科科技（A 股）：26H1 总营收增速为 0.29%（去年同期为 4.59%），公司总体营收增速放缓主要由于公司主动收缩软件系统业务，目前 AI 业务已成为公司营收占比最高的业务板块，26H1 公司 AI 产品及服务收入 2.94 亿元，同比增长 39.15%，占总收入近 40%；AI 业务毛利率 53.40%，同比提升 6.03pct；26H1 公司 AI 业务新增客户 20 余家，累计客户超过 100 家，AI 业务加速规模化复制；

8）用友网络（A 股）：26H1 总营收增速为 4.08%（去年同期为-5.89%），AI 产品商业化加速带动总体营收修复。26H1AI 相关合同签约金额 9.06 亿元，半年签约规模达 25 年全年的 54%；26H1 公司 AI 相关收入 4.64 亿元，其中 Q1 约 1.37 亿元、Q2 约 3.27 亿元，Q2 环比增长超过 139%，AI 业务落地节奏显著加快。

仅供内部参考，请勿外传

9）泛微网络（A 股）：26H1 总营收增速为 5.04%（去年同期为-8.79%），公司营收修复主要受益于 AI 大模型加速向泛办公场景延伸，客户 AI 需求加速放量。26H1 公司持续推进协同办公产品 AI 升级，泛微数智大脑 Xiaoe.AI 已推出基于 Agentic 架构的新版本，可通过自然语言完成信息采集、文档处理及日常办公等任务。同时公司加强与腾讯的生态合作，泛微事并然项目管理平台已完成与腾讯智能体 WorkBuddy 对接。

10）金山办公（A 股）：26H1 总营收增速为 24.69%（去年同期为 10.12%），公司营收增长加速主要受益于 AI 大模型加速向泛办公场景延伸，公司卡位优势持续凸显。受益于个人及企业端 AI 需求增长，26H1 公司 WPS 个人、WPS 365 业务均快速增长，营收分别为 20.14、4.97 亿元，分别同比增长 15.20%、60.84%。产品端，26H1 公司推出全新版灵犀（2C 产品）、WPS Comate（2B 产品），旨在通过产品功能迭代强化原有的入口粘性。

图表18：AI 应用代表公司 26H1 营收基本呈修复趋势

公司	25H1营收同比增速（%）	26H1营收同比增速（%）	对比
金蝶国际	11.24	13.56	加速
迅策	-29.98	388.78	加速
迈富时	25.54	111.18	加速
滴普科技	118.36	114.97	减速
汉得信息	3.54	6.07	加速
鼎捷数智	4.08	5.37	加速
能科科技	4.59	0.29	减速
用友网络	-5.89	4.08	加速
泛微网络	-8.79	5.04	加速
金山办公	10.12	24.69	加速

资料来源：Wind、华泰研究

AI 战略转型同步驱动利润改善，后续有望持续推进。AI 转型背景下，国内软件企业正通过 FDE（Forward Deployed Engineer，驻场部署工程师）模型加速提升业务质量，AI 业务带动多数公司毛利率、费用率稳步改善。**1）毛利率：**受益于 AI 应用产品稳步放量，国内软件公司毛利率呈现改善趋势，以用友网络为例，公司 26H1 毛利率同比提升 2.56pct；**2）费用率：**受益于 AI 内部提效推进，国内软件公司费用管控效果初显，以金蝶国际为例，公司销售/管理/研发费率分别同比收窄 2.89、0.69、1.97pct。看好软件公司长期盈利能力改善。

图表19：AI 转型带动软件公司盈利能力改善

	净利率		毛利率		销售费率		管理费率		研发费率	
公司	25H1	26H1	25H1	26H1	25H1	26H1	25H1	26H1	25H1	26H1
金蝶国际	-3.42	1.33	65.63	67.74	39.59	36.70	10.84	10.15	24.29	22.32
迅策	-54.59	6.93	66.66	60.07	7.67	6.22	17.38	8.08	84.96	41.01
迈富时	3.86	10.34	48.56	45.19	16.96	11.37	10.19	4.58	9.06	16.59
滴普科技	-233.32	-11.24	55.04	56.46	37.33	23.13	110.15	12.70	44.09	25.40
汉得信息	5.55	5.57	33.97	35.11	8.21	7.96	13.53	11.54	8.10	7.23
鼎捷数智	4.25	16.59	58.01	55.48	30.66	28.44	11.16	10.55	8.20	6.64
能科科技	21.63	17.95	52.07	53.22	5.52	7.05	6.02	6.72	15.58	16.52
用友网络	-28.57	-25.43	48.33	50.89	29.28	26.36	16.50	16.07	31.41	31.50
泛微网络	8.20	13.52	93.88	94.24	66.77	63.07	4.12	5.88	18.71	16.84
金山办公	27.85	75.99	84.88	84.90	18.19	18.04	8.32	8.36	36.07	34.76

注：红色标注表示当期数值大于去年同期，；绿色标注表示当期数值小于去年同期

资料来源：Wind、华泰研究

仅供内部参考，请勿外传

展望：FDE 模型成为全球共识，看好国内软件公司业绩放量

全球软件厂商加速 FDE 转型，国内软件 Know-how 优势及客户粘性优势有望凸显，看好全年业绩放量。AI 时代，软件产品竞争从标准化产品逐步转向“产品能力+行业 Know-how+深度交付”并重，全球软件厂商同步加速 FDE 转型。FDE 要求软件厂商具备将模型能力嵌入具体业务流程、打通存量数据并持续迭代应用的能力，国内软件公司长期服务于本土大型企业和政企客户，具备贴近客户现场、理解业务流程、具备二次开发能力的关键能力，其 Know-how 优势及客户粘性优势有望凸显。我们认为，FDE 模式将放大国内软件厂商在行业理解和客户基础上的传统优势，同时推动 AI 订单由 POC（Proof of Concept，概念验证）试点加速向正式合同转化，看好国内软件厂商全年营收及利润加速进入放量阶段。

图表20： AI 应用代表公司关于 FDE 的核心表述

公司	FDE相关描述
迅策	•2026年7月19日，公司与青岛鑫宸科创签署战略合作备忘录，计划探索垂直行业数据Token化治理、推动被投企业AI能力升级，并基于FDE模式打造Token商业化标杆落地案例。
汉得信息	•将FDE作为AI工程化实施理念，形成“灵策AI实施方法论”，覆盖战略规划、方案设计、落地实施及运行优化的完整生命周期，目标是使AI项目不仅能够上线，还能持续产生业务价值。
能科科技	•设立AI融合创新中心并组建FDE组织。工程师长期派驻客户业务一线，挖掘隐性需求，负责数据接入、模型适配、系统集成和现场调优，并将共性需求反馈至研发端，形成“一线洞察—产品进化—规模复制”的闭环。
用友网络	•FDE团队长期驻场客户，负责数据及AI项目端到端落地，包括数据治理、定制功能开发、AI能力验证，并将业务痛点和落地经验反馈给产品团队。公司将传统软件套件实施的SDE与智能及数据前沿服务的FDE并行布局。
金山办公	•以FDE理念升级WPS 365企业级交付模式，加强产品、技术、售前、交付和运营协同。针对制造、金融、医药、政务、教育等行业，根据客户业务流程、组织架构和数据环境制定差异化部署方案，将WPS 365与既有业务系统及组织流程深度融合。

资料来源：Wind、华泰研究

仅供内部参考，请勿外传

算力租赁：收入保持高增，毛利率修复带动盈利弹性释放

样本公司收入利润及资产端情况

我们选取行业 20 家算力租赁公司作为我们的研究样本，所选样本的收入利润情况及资产端情况如下：

图表21：算力租赁研究样本公司收入利润情况

公司	代码	营业收入（亿元）	营收 yoy	归母净利润（亿元）	利润 yoy	算租收入（亿元）	占比	算租 yoy
协创数据	300857 CH	125.23	153.29%	18.31	334.07%	32.16	25.68%	163.39%
宏景科技	301396 CH	18.43	156.43%	3.05	405.50%	16.89	91.63%	134.91%
东阳光	600673 CH	93.76	31.61%	4.33	-30.91%	2.89	3.08%	-
行云科技	300209 CH	2.54	497.11%	0.14	464.11%	0.67	26.36%	-
盈峰环境	000967 CH	78.72	21.80%	4.84	21.77%	6.86	8.71%	-
赛意信息	300687 CH	10.33	14.43%	0.53	325.85%	-	-	-
杰创智能	301248 CH	3.43	8.72%	-0.05	-125.00%	1.72	50.11%	-
中贝通信	603220 CH	13.58	-14.32%	0.08	-82.50%	4.07	29.98%	37.04%
盛视科技	002990 CH	7.34	33.73%	1.07	342.91%	3.55	48.35%	-
智微智能	001339 CH	30.50	56.64%	6.76	272.88%	10.30	33.77%	245.64%
鸿博股份	002229 CH	1.87	-76.39%	-0.46	-195.60%	0.52	27.88%	-
润建股份	002929 CH	37.56	-22.83%	0.05	-71.33%	5.19	13.82%	50.43%
利通电子	603629 CH	21.01	38.58%	7.05	1265.74%	12.74	60.63%	162.14%
天阳科技	300872 CH	9.64	-3.29%	0.54	10.20%	-	-	-
中威电子	300270 CH	0.72	90.98%	0.08	120.80%	-	-	-
金刚光伏	300093 CH	0.91	-35.18%	-1.11	60.54%	0.28	30.91%	-
高乐股份	002348 CH	1.65	25.29%	-0.29	-3.73%	-	-	-
莲花控股	600186 CH	19.14	18.10%	2.50	53.58%	0.54	2.82%	-21.74%
先河环保	300137 CH	3.53	-15.52%	-0.20	-179.87%	-	-	-
蜂助手	301382 CH	13.39	36.14%	0.77	3.29%	-	-	-

资料来源：公司半年报，华泰研究

图表22：算力租赁研究样本公司资产情况

公司	代码	预付账款（亿元）	预付账款 qoq	在建工程（亿元）	在建工程 qoq	固定资产（亿元）	固定资产 qoq	使用权资产（亿元）	使用权资产 qoq	存货（亿元）	存货 qoq
协创数据	300857 CH	17.57	70.92%	8.99	-85.12%	220.24	95.79%	0.60	127.14%	53.52	25.61%
宏景科技	301396 CH	3.70	318.17%	17.26	13.93%	60.73	217.48%	32.20	6.16%	13.25	117.99%
东阳光	600673 CH	4.35	-40.76%	23.42	70.66%	111.94	41.08%	0.81	98.54%	20.26	-2.43%
行云科技	300209 CH	0.96	-57.03%	19.92	-	0.09	397.70%	0.29	301.82%	0.59	-5.82%
盈峰环境	000967 CH	1.40	-6.57%	3.94	5.74%	26.81	4.42%	52.89	24.45%	14.48	15.95%
赛意信息	300687 CH	4.36	117.19%	2.92	15.65%	2.85	7.77%	7.57	3045.87%	2.76	2.77%
杰创智能	301248 CH	0.45	-75.68%	5.06	-53.90%	20.79	108.77%	0.04	347.87%	3.54	-25.16%
中贝通信	603220 CH	1.00	203.86%	6.11	71.61%	28.14	23.72%	3.83	-10.15%	7.93	80.79%
盛视科技	002990 CH	14.97	6473.52%	0.00	-	4.84	46.71%	0.07	0.71%	10.39	309.82%
智微智能	001339 CH	2.81	-83.13%	0.01	1.85%	5.73	-0.78%	0.49	-1.58%	47.67	41.35%
鸿博股份	002229 CH	0.47	168.56%	0.11	-	5.06	-5.57%	0.89	-7.45%	1.00	-26.80%
润建股份	002929 CH	24.40	530.88%	0.91	12.08%	10.01	-2.78%	0.70	-7.19%	20.52	10.69%
利通电子	603629 CH	5.46	658.12%	0.18	2720.31%	13.16	-10.34%	14.45	-3.76%	7.20	-73.53%
天阳科技	300872 CH	0.90	21.89%	-	-	2.75	56.46%	0.06	86.51%	6.83	-0.15%
中威电子	300270 CH	0.05	64.30%	-	-	0.49	-15.60%	0.02	1.35%	0.18	-16.29%
金刚光伏	300093 CH	0.03	-69.88%	3.51	172.64%	20.66	-3.99%	0.02	-	0.13	-56.98%
高乐股份	002348 CH	0.07	-1.08%	0.01	-55.51%	4.23	-0.83%	-	-	0.52	31.29%
莲花控股	600186 CH	1.27	-54.44%	-	-	5.55	-8.99%	0.17	-21.12%	7.24	168.18%
先河环保	300137 CH	1.06	156.55%	2.09	-	3.57	72.68%	0.17	114.03%	2.18	-0.61%
蜂助手	301382 CH	11.43	-4.23%	1.80	29.44%	2.28	-1.18%	0.02	-18.73%	0.28	-10.65%

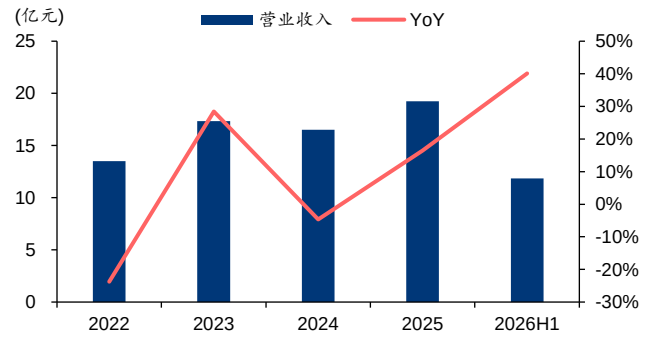
资料来源：公司半年报，华泰研究

仅供内部参考，请勿外传

算力租赁板块收入利润快速兑现

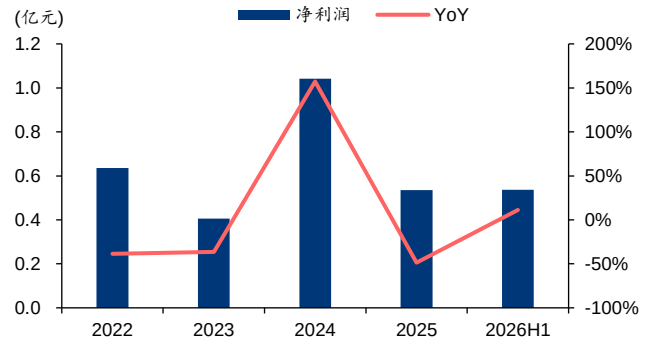
算力租赁板块收入端率先进入加速兑现期，利润端同步修复但弹性暂弱于收入。2022-2025 年板块收入增速经历较大波动，2025 年恢复至 16.5%的正增长后，2026H1 收入中位数同比进一步加速至 40.1%，反映前期投入的算力资源逐步完成上架并进入收入确认阶段，板块景气由资产扩张向经营兑现传导。相比之下，净利润在 2022-2025 年波动更为明显，2026H1 归母净利润中位数同比转为 11%的正增长，显示盈利端已走出前期调整。结合毛利率同比明显回升而销售、管理及研发费用率均有所抬升，我们认为，当前算力利用率提升及高端算力价格支撑已开始改善项目毛利，伴随资产端的逐步扩张，更大的利润弹性有望在后续季度逐步释放。

图表23：2022-2026H1 算力租赁板块营收及增速情况（中位数法）



注：2022-2025 为年报数据。
资料来源：iFinD，华泰研究

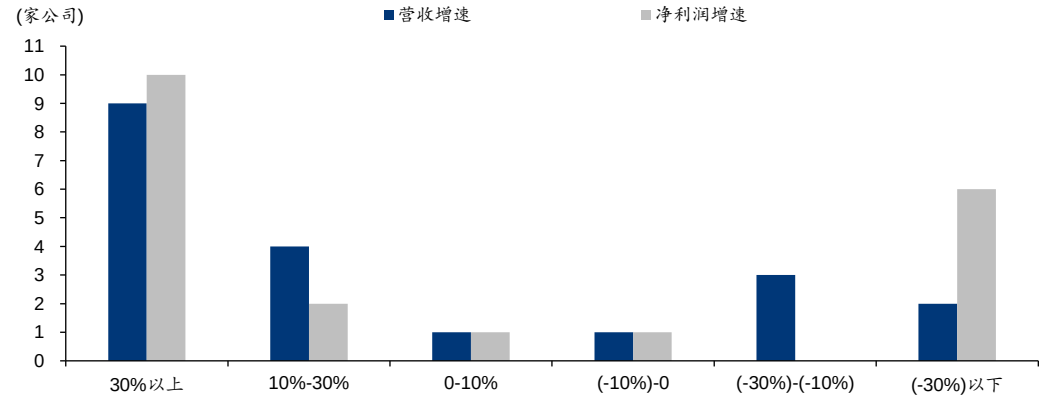
图表24：2022-2026H1 算力租赁板块净利润及增速情况（中位数法）



注：2022-2025 为年报数据。
资料来源：iFinD，华泰研究

2026H1 算力租赁板块收入端保持较快增长，利润端分化更为显著。我们选取 20 家算力租赁相关公司作为样本，2026H1，其中 14 家公司实现同比正增长，占比 70%；收入增速超过 30%的公司达到 9 家，10%-30%的公司为 4 家，两者合计占比 65%，显示随着前期算力资源逐步上架、客户需求持续兑现，板块收入仍处于较快放量阶段。13 家公司实现利润同比正增长，其中 10 家公司增速超过 30%。我们认为，头部公司随着算力资产利用率提升及毛利率改善，收入向利润的转化效率增强。

图表25：2026H1 算力租赁板块营业收入增速统计

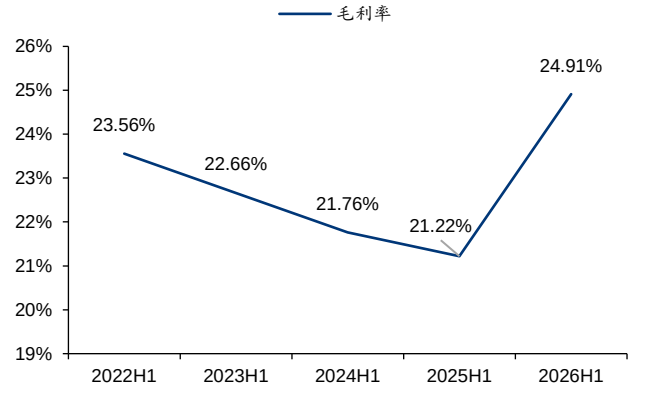


资料来源：iFinD，华泰研究

毛利率修复，强化经营杠杆。2026H1 样本公司毛利率中位数同比提升 3.69pct 至 24.91%，结束 2022H1 以来的持续下行趋势；同期销售、管理及研发费用率分别升至 19.43%/6.16%/4.06%，均较 2025H1 有所抬升。我们认为，这一变化主要受益于前期集中投放的算力设备逐步进入稳定运营阶段，上架率和利用率提升使折旧、机房及运维等固定或半固定成本得到更充分摊薄，带动单位算力成本下降；同时，高端算力供给相对紧张亦对租赁价格形成一定支撑，进一步改善项目毛利率水平。

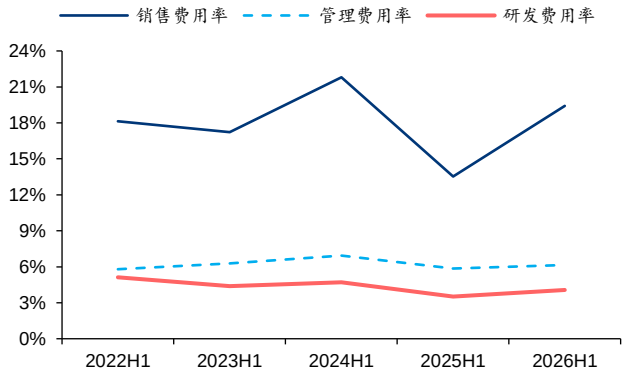
仅供内部参考，请勿外传

图表26：2022H1-2026H1 算力租赁板块毛利率情况（中位数法）



资料来源：iFinD，华泰研究

图表27：2022H1-2026H1 算力租赁板块费用率情况（中位数法）



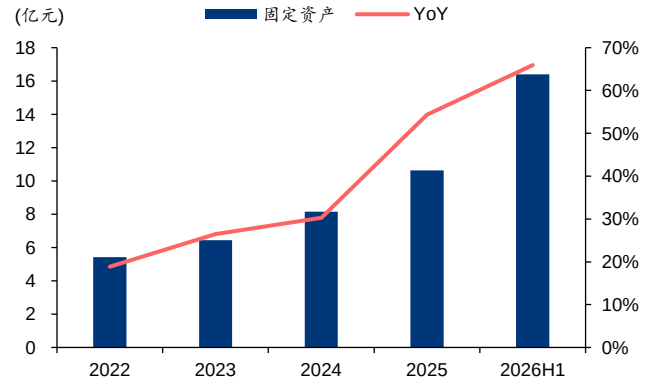
资料来源：iFinD，华泰研究

算力租赁板块资产端进入快速扩展周期

资产端整体保持高景气，算力租赁板块仍处于较强的产能建设与资源扩张周期。2026H1 样本公司在建工程、固定资产、使用权资产及存货平均规模分别达到 4.81/16.40/5.76/11.02 亿元，同比增速分别为 92.7%/65.9%/128.9%/64.8%，较此前年度整体明显加速。其中，在建工程接近翻倍，反映新增机房、机柜及供配电等基础设施建设持续推进；固定资产仍保持较高增速，说明前期项目正同步完成验收、转固和上架；存货高增则显示服务器、GPU 等设备采购及待交付资源明显增加。整体来看，板块并非单一处于建设期，而是呈现“设备采购、项目建设、资产转固同步推进”的扩张特征。

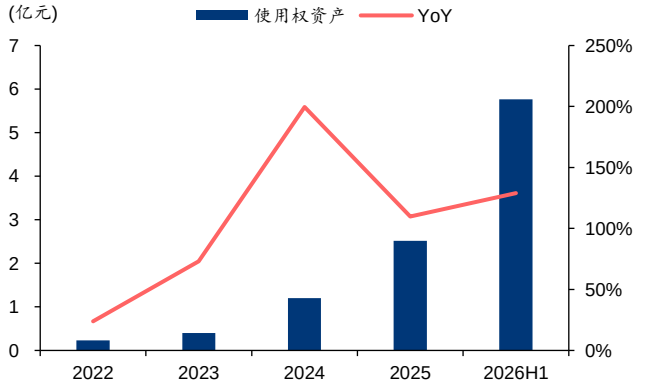
从资产结构看，本轮算力扩张同时体现重资产投入与租赁扩容两条路径，新增供给仍在快速形成。固定资产自 2022 年以来持续较快增长，表明直接采购设备和建设基础设施仍是主要扩产方式；使用权资产则连续高速增长，2026H1 同比增速达到 128.9%，反映企业也在更多通过租赁 IDC 机房、机柜及设备提升资源获取和部署效率。与此同时，在建工程和存货增速仍处高位，意味着仍有较多算力资源停留在建设、交付及上架前环节，后续仍具备进一步转化为可计费产能的空间。我们认为，当前板块的关键已逐步从“扩张了多少资产”转向新增资产能否快速上架、提升利用率并形成有效收入贡献，资产周转和单位资产产出将成为判断本轮扩张质量的核心。

图表28：2022-2026H1 算力租赁板块固定资产及增速（平均数法）



资料来源：iFinD，华泰研究

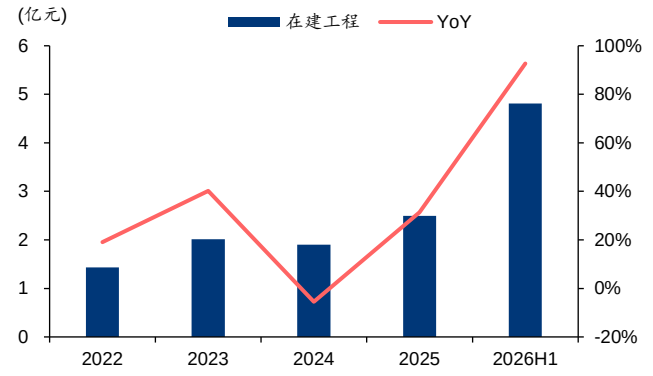
图表29：2022-2026H1 算力租赁板块使用权资产及增速（平均数法）



资料来源：iFinD，华泰研究

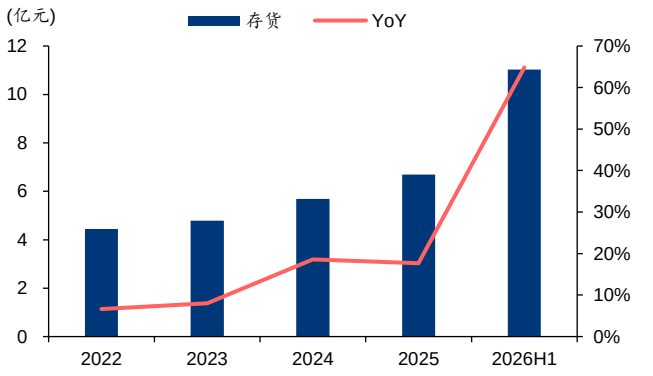
仅供内部参考，请勿外传

图表30：2022-2026H1 算力租赁板块在建工程及增速（平均数法）



资料来源：iFinD，华泰研究

图表31：2022H1-2026H1 算力租赁板块存货及增速（平均数法）



资料来源：iFinD，华泰研究

仅供内部参考，请勿外传

AI4S：收入高速增长，大额 BD 驱动行业商业化提速

我们选取行业 4 家 AI4S 公司作为我们的研究样本，所选样本的收入利润情况如下：

图表32： AI4S 研究样本公司收入利润情况

公司	代码	2026H1 收入（亿元）	2026H1 收入增速	2026H1 归母净利润（亿元）	2026H1 归母净利润增速
晶泰控股	2228 HK	3.94	-24%	-2.52	-585%
英矽智能	3696 HK	7.24	287%	2.42	307%
剂泰科技 - P	7666 HK	1.54	13399%	-1.53	20%
德适 - B	2526 HK	1.09	21%	-0.56	-4070%

资料来源：Wind、华泰研究

收入端：从 2026 年上半年四家 AI4S 样本企业营收表现来看，行业内公司收入增长呈现出显著分化的格局，各家企业商业化落地节奏差异较大。四家样本企业 2026H1 营收规模跨度较大，英矽智能以 7.24 亿元的营业收入居于首位，大幅领先其余三家企业；晶泰控股实现营收 3.94 亿元；剂泰科技-P、德适-B 营收体量相对偏小，分别录得 1.54 亿元、1.09 亿元。

从收入增速维度拆解，行业增长梯队分化特征尤为突出。剂泰科技-P 营收增速高达 13399%，增速表现亮眼，该超高增速主要源于前期收入基数较低，伴随着项目订单逐步落地、商业化客户持续拓展，公司上半年业务规模实现跨越式扩张。英矽智能营收增速达到 287%，在较高收入体量之上依旧维持高速增长，商业化成果持续兑现，客户认可度与项目交付能力稳步提升，增长质量相对优异。德适-B 收入增速为 21%，营收保持稳健的正向增长，商业化进程平稳推进。晶泰控股营收同比下滑 24%，主要由于去年上半年与 DoveTree 合作确认 5100 万美元收入款导致基数较高。

整体来看，当前 AI4S 行业尚处在商业化早期阶段，不同企业之间营收增长率存在一定分化。头部企业已跑出高速增长曲线，部分新兴企业依托低基数实现超高速增长。行业尚未形成统一稳定的增长范式，后续各家企业商业化转化能力、客户拓展进度，将会成为拉开营收差距的核心变量，行业分化态势或将延续。

利润端：2026 年上半年四家 AI4S 样本公司归母净利润表现分化明显，盈利兑现进度出现分水岭，行业整体仍处于盈利探索阶段。四家企业中仅英矽智能实现正向盈利，其余三家企业依旧维持亏损状态，反映出 AI 药物研发赛道前期研发投入高、盈利周期长的行业属性，大额 BD 首付款对利润的提振具备较强的阶段性特征。

具体来看，英矽智能 2026H1 实现归母净利润 2.42 亿元，归母净利润增速达 307%。本期利润大幅向好核心受益于礼来大额 BD 首付款收入确认，高额一次性授权收益大幅增厚当期利润。在营收高速增长的带动下成功实现盈利，利润端增长势头强劲。公司商业化收入规模不断扩大，规模效应逐步显现，收入增长有效覆盖各项运营与研发成本，成为样本企业中率先完成由亏转盈的标的。晶泰控股归母净利润亏损 2.52 亿元，净利润同比增速-585%，亏损幅度大幅扩大，主要是由于人员、算力、研发等成本费用刚性支出，对短期利润表现产生一定压力。剂泰科技-P 上半年归母净利润亏损 1.53 亿元，亏损同比收窄 20%，管线研发、团队扩张等持续投入，短期仍处于亏损状态。德适-B 归母净利润亏损 0.56 亿元，净利润增速-4070%，主要是由于算力投入及上市相关开支同比增加所致。

综合而言，当前 AI4S 行业盈利拐点尚未全面到来，绝大多数企业依旧处于投入期，盈利企业仍属少数。大额 BD 收益可以带来阶段性利润拐点，后续需要观察常态化商业化订单落地情况，以及企业能否持续降本增效、放大持续性商业化收入规模，是决定能否扭亏为盈的关键。

仅供内部参考，请勿外传

AI4S 样本公司毛利率及费用率情况如下：

图表33：AI4S 研究样本公司毛利率及费用率情况

公司	代码	2026H1 毛利率	2026H1 销售费用率	2026H1 管理费用率	2026H1 研发费用率
晶泰控股	2228 HK	57%	12.19%	53.89%	93.45%
英矽智能	3696 HK	90%	5.99%	11.83%	46.40%
剂泰科技 - P	7666 HK	99%	13.78%	62.31%	111.72%
德适 - B	2526 HK	74%	14.66%	34.76%	58.99%

资料来源：Wind、华泰研究

2026 年上半年四家 AI4S 样本企业毛利率整体处于较高区间，行业轻资产属性凸显，公司间毛利率水平分化明显，主要由于各家收入结构、业务模式存在显著差异。四家公司的毛利率均高于 55%，剂泰科技-P 毛利率高达 99%，居于样本首位；英矽智能紧随其后，毛利率达到 90%；德适-B 毛利率为 74%；晶泰控股毛利率录得 57%。

从业务结构角度拆分，剂泰科技-P 的高毛利率主要来自 BD 授权首付款收入，管线对外授权业务几乎无对应的增量履约成本，拉动整体毛利率抬升；英矽智能 90%的高毛利率一方面受益于礼来大额 BD 授权收入入账，另一方面其 AI 药物研发服务依托自研算法平台，算力、人员等固定成本摊销完毕后，增量项目边际成本较低，服务业务盈利质量优异。德适-B 毛利率 74%处于中等水平，上半年高毛利的医疗影像大模型 IMEDImage 模型许可业务收入占比大幅提升，收入结构进一步优化。晶泰控股毛利率 57%，主要受收入结构与项目交付结构影响，AI4S 智慧解决方案尚处于高速扩张阶段，毛利率尚未完全稳定

整体来看，AI4S 赛道具备高毛利的天然属性，大额 BD 授权收入能够显著抬升当期毛利率水平。毛利率分化本质是收入结构差异的体现，侧重管线授权 BD 交易的企业短期毛利率更高。展望未来，随着商业化订单规模持续扩大，固定成本不断摊薄，服务类业务毛利率有望迎来修复；而以一次性首付款为主的 BD 收入带来的高毛利具备阶段性特征，长期毛利率走势仍取决于常态化业务结构的变化。

风险提示

下游需求波动。若宏观经济、行业景气度等因素导致下游客户预算收缩或项目推进节奏放缓，可能影响相关企业订单获取及收入确认，导致业绩增长不及预期。

AI 技术迭代不及预期。若大模型能力提升、算法优化及 AI 应用落地进展不及预期，可能影响 AI 产业应用拓展及商业化进程，导致行业增长不及预期。

AI4S 赛道毛利率分化。AI4S 不同细分领域在产品形态、项目交付模式及客户结构等方面存在差异，若低毛利业务占比提升或行业竞争加剧，可能导致相关企业毛利率承压。

仅供内部参考，请勿外传

免责声明

分析师声明

本人，郭雅丽、范昳蕊、袁泽世、岳铂雄，兹证明本报告所表达的观点准确地反映了分析师对标的证券或发行人的个人意见；彼以往、现在或未来并无就其研究报告所提供的具体建议或所表达的意见直接或间接收取任何报酬。请注意，标*的人员并非香港证券及期货事务监察委员会的注册持牌人，不可在香港从事受监管活动。

一般声明及披露

本报告由华泰证券股份有限公司或其关联机构制作，华泰证券股份有限公司和其关联机构统称为“华泰证券”(华泰证券股份有限公司已具备中国证监会批准的证券投资咨询业务资格) 。本报告所载资料是仅供接收人的严格保密资料。本报告仅供华泰证券及其客户和其关联机构使用。华泰证券不因接收人收到本报告而视其为客户。

本报告基于华泰证券认为可靠的、已公开的信息编制，但华泰证券对该等信息的准确性及完整性不作任何保证。

本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，华泰证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。以往表现并不能指引未来，未来回报并不能得到保证，并存在损失本金的可能。华泰证券不保证本报告所含信息保持在最新状态。华泰证券对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

华泰证券（华泰证券（美国）有限公司除外）不是 FINRA 的注册会员，其研究分析师亦没有注册为 FINRA 的研究分析师/不具有 FINRA 分析师的注册资格。

华泰证券力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，华泰证券及作者均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现，过往的业绩表现不应作为日后回报的预示。华泰证券不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。

华泰证券及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，华泰证券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，为该公司提供投资银行、财务顾问或者金融产品等相关服务或向该公司招揽业务。

华泰证券的销售人员、交易人员或其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。华泰证券没有将此意见及建议向报告所有接收者进行更新的义务。华泰证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。投资者应当考虑到华泰证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。有关该方面的具体披露请参照本报告尾部。

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、发布的机构或人员，也并非意图发送、发布给因可得到、使用本报告的行为而使华泰证券违反或受制于当地法律或监管规则的机构或人员。

本报告版权仅为华泰证券所有。未经华泰证券书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人(无论整份或部分)等任何形式侵犯华泰证券版权。如征得华泰证券同意进行引用、刊发的，需在允许的范围内使用，并需在使用前获取独立的法律意见，以确定该引用、刊发符合当地适用法规的要求，同时注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。华泰证券保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为华泰证券的商标、服务标记及标记。

中国香港

本报告由华泰证券股份有限公司或其关联机构制作,在香港由华泰金融控股（香港）有限公司向符合《证券及期货条例》及其附属法律规定的机构投资者和专业投资者的客户进行分发。华泰金融控股（香港）有限公司受香港证券及期货事务监察委员会监管，是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。在香港获得本报告的人员若有任何有关本报告的问题,请与华泰金融控股（香港）有限公司联系。

仅供内部参考，请勿外传

香港-重要监管披露

- 华泰金融控股（香港）有限公司的雇员或其关联人士没有担任本报告中提及的公司或发行人的高级人员。
- 有关重要的披露信息，请参华泰金融控股（香港）有限公司的网页 https://www.htsc.com.hk/stock_disclosure

其他信息请参见下方 **“美国-重要监管披露”**。

美国

在美国本报告由华泰证券（美国）有限公司向符合美国监管规定的机构投资者进行发表与分发。华泰证券（美国）有限公司是美国注册经纪商和美国金融业监管局（FINRA）的注册会员。对于其在美国分发的研究报告，华泰证券（美国）有限公司根据《1934 年证券交易法》（修订版）第 15a-6 条规定以及美国证券交易委员会人员解释，对本研究报告内容负责。华泰证券（美国）有限公司联营公司的分析师不具有美国金融监管（FINRA）分析师的注册资格，可能不属于华泰证券（美国）有限公司的关联人员，因此可能不受 FINRA 关于分析师与标的公司沟通、公开露面和所持交易证券的限制。华泰证券（美国）有限公司是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。任何直接从华泰证券（美国）有限公司收到此报告并希望就本报告所述任何证券进行交易的人士，应通过华泰证券（美国）有限公司进行交易。

美国-重要监管披露

- 分析师郭雅丽、范映蕊、袁泽世、岳铂雄本人及相关人士并不担任本报告所提及的标的证券或发行人的高级人员、董事或顾问。分析师及相关人士与本报告所提及的标的证券或发行人并无任何相关财务利益。本披露中所提及的“相关人士”包括 FINRA 定义下分析师的家庭成员。分析师根据华泰证券的整体收入和盈利能力获得薪酬，包括源自公司投资银行业务的收入。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或不时会以自身或代理形式向客户出售及购买华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或其高级管理层、董事和雇员可能会持有本报告中所提到的任何证券（或任何相关投资）头寸，并可能不时进行增持或减持该证券（或投资）。因此，投资者应该意识到可能存在利益冲突。

新加坡

华泰证券（新加坡）有限公司持有新加坡金融管理局颁发的资本市场服务许可证，可从事资本市场产品交易，包括证券、集体投资计划中的单位、交易所交易的衍生品合约和场外衍生品合约，并且是《财务顾问法》规定的豁免财务顾问，就投资产品向他人提供建议，包括发布或公布研究分析或研究报告。华泰证券（新加坡）有限公司可能会根据《财务顾问条例》第 32C 条的规定分发其在华泰证券内的外国附属公司各自制作的信息/研究。本报告仅供认可投资者、专家投资者或机构投资者使用，华泰证券（新加坡）有限公司不对本报告内容承担法律责任。如果您是非预期接收者，请您立即通知并直接将本报告返回给华泰证券（新加坡）有限公司。本报告的新加坡接收者应联系您的华泰证券（新加坡）有限公司关系经理或客户主管，了解来自或与所分发的信息相关的事宜。

评级说明

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力（含此期间的股息回报）相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，台湾市场基准为台湾加权指数，日本市场基准为日经 225 指数，新加坡市场基准为海峡时报指数，韩国市场基准为韩国有价证券指数，英国市场基准为富时 100 指数，德国市场基准为 DAX 指数），具体如下：

行业评级

增持：预计行业股票指数超越基准
中性：预计行业股票指数基本与基准持平
减持：预计行业股票指数明显弱于基准

公司评级

买入：预计股价超越基准 15%以上
增持：预计股价超越基准 5%~15%
持有：预计股价相对基准波动在-15%~5%之间
卖出：预计股价弱于基准 15%以上
暂停评级：已暂停评级、目标价及预测，以遵守适用法规及/或公司政策
无评级：股票不在常规研究覆盖范围内。投资者不应期待华泰提供该等证券及/或公司相关的持续或补充信息

仅供内部参考，请勿外传

法律实体披露

中国：华泰证券股份有限公司具有中国证监会核准的“证券投资咨询”业务资格，经营许可证编号为：91320000704041011J

香港：华泰金融控股（香港）有限公司具有香港证监会核准的“就证券提供意见”业务资格，经营许可证编号为：AOK809

美国：华泰证券（美国）有限公司为美国金融业监管局（FINRA）成员，具有在美国开展经纪交易商业务的资格，经营业务许可编号为：CRD#:298809/SEC#:8-70231

新加坡：华泰证券（新加坡）有限公司具有新加坡金融管理局颁发的资本市场服务许可证，并且是豁免财务顾问，经营许可证编号为：202233398E

华泰证券股份有限公司

南京
南京市建邺区江东中路 228 号华泰证券广场 1 号楼/邮政编码：210019

电话：86 25 83389999/传真：86 25 83387521
电子邮件：ht-rd@htsc.com

北京
北京市西城区太平桥大街丰盛胡同 28 号太平洋保险大厦 A 座 18 层/邮政编码：100032
电话：86 10 63211166/传真：86 10 63211275
电子邮件：ht-rd@htsc.com

深圳
深圳市福田区益田路 5999 号基金大厦 10 楼/邮政编码：518017
电话：86 755 82493932/传真：86 755 82492062
电子邮件：ht-rd@htsc.com

上海
上海市浦东新区东方路 18 号保利广场 E 栋 23 楼/邮政编码：200120
电话：86 21 28972098/传真：86 21 28972068
电子邮件：ht-rd@htsc.com

华泰金融控股（香港）有限公司
香港中环皇后大道中 99 号中环中心 53 楼
电话：+852-3658-6000/传真：+852-2567-6123
电子邮件：research@htsc.com
http://www.htsc.com.hk

华泰证券（美国）有限公司
美国纽约公园大道 280 号 21 楼东（纽约 10017）
电话：+212-763-8160/传真：+917-725-9702
电子邮件: Huatai@htsc-us.com
http://www.htsc-us.com

华泰证券（新加坡）有限公司
滨海湾金融中心 1 号大厦, #08-02, 新加坡 018981
电话：+65 68603600
传真：+65 65091183
https://www.htsc.com.sg

©版权所有2026年华泰证券股份有限公司