



计算机行业研究

买入（维持评级）
行业研究

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：赵伟博（执业 S1130525120006）
 zhaoweibo@gjzq.com.cn

 分析师：陈芷婧（执业 S1130525120008）
 chenzhijing@gjzq.com.cn

 分析师：郑元昊（执业 S1130525120004）
 zhengyuanhao@gjzq.com.cn

 分析师：李可夫（执业 S1130525120009）
 likefu@gjzq.com.cn

 联系人：孙恺祈
 sunkaiqi@gjzq.com.cn

计算机行业 2026 年度投资策略：两仪生四象，超级时代

行业观点

- **2026 年 AI 演绎路径更加清晰，人工智能太极生两仪，模型加速迭代与 AI 超级载体出现奠定基调。**1) 算力基础设施变革、算法细节创新不断驱动模型仍在快速迭代。全球顶尖大模型持续突破，谷歌 Gemini 3 在基础推理与多模态能力上实现跃迁式提升，最关键的指标 screen understanding 准确度从上一代 11.4% 提升到 72.7% 意味着 AI 手机爆发成为必然，同时具身智能模型有可能在三年内能理解自然语言指令，变成场景内多用途的机器人；算法创新仍然频繁：国产开源模型 DeepSeek-V3.2 Speciale 在推理性能上逼近顶尖闭源模型，其稀疏注意力（DSA）、高比例后训练和大规模合成数据等方面的创新，证明算法与训练范式仍具巨大突破空间，预训练与后训练 Scaling Law 远未触及瓶颈；目前模型训练仍以英伟达 Hopper 系列与谷歌 TPU v3、v5 为主，随着英伟达 Blackwell 集群与谷歌 TPU v7 年底投入使用，2026 年大模型迭代仍然精彩。2) AI 超级载体出现：手机本身是人们工作与生活信息流的集大成者，全能的 AI 助手出现将全面重构所有人的工作与生活。字节跳动于 2025 年 12 月发布豆包 AI 手机助手技术预览版，全能智能助理实现 AI 对手机的跨应用自主操作，虽然尚有不足之处但已经展现全能潜力，形成数据飞轮，开始快速迭代。GUI 技术突破扰乱了 API 时代形成的护城河，且 AI 手机是其他 AI 终端爆发的前置条件，预计全球顶尖大厂将极致投入到这场新入口之战。3) **两仪生四象：算力为基、空天时代、具身智能、AI 应用精彩不断。**
- **四象一：一切仍然指向算力，算力/存力/超节点景气加速，国内算力斜率极高。**1) Scaling Law 从 pre Training 向 pre/Post/Test-Time 三维度转变，叠加 Interleaved Thinking 的新推理范式，训练和推理算力均有望指数级爆发。2) AI 手机驱动入口之战，没有厂商能独善其身，腾讯、华为近期投入信号明显。预计全球大厂投入容忍度将极高，目前统计主流大厂现金流无压力。3) 国内算力：国产 GPU 性能追平 H20，上市/增发融资叠加产能突破，国产 GPU 厂商“融资-研发-量产”的闭环跑通，26 年国产替代加速；4) 存储：多模态渗透与 KV Cache 的广泛应用，存储正经历从“供过于求”向“供不应求”的转变，价格与景气度持续上行；5) 超节点：超节点通过高速互联弥补贷款不足、高时延等问题，已成为 AI 基础设施新常态，国产 GPU 厂商加速布局。
- **四象二：空天时代，AI 的锋刃。**1) 海外端，英伟达、谷歌、SpaceX 等机构纷纷投入太空算力建设，且 SpaceX 有望成就史上最大 IPO，估值或高达 1.5 万亿美元，是全球算力 6000 亿美金市场未来超万亿美金市场中增长斜率最高的部分。2) 政策端：近期支持政策频出，“十五五”规划中重视航空航天发展，2025 年国家航天局印发行动规划并随后成立专职监管机构商业航天司，近日商业火箭企业适用科创板第五套上市标准等重磅政策不断。3) 国内端，我们积极规划多个万星低轨星座并已规划三体计算星座与北京太空数据中心参与太空算力角逐。4) 我国可回收火箭技术突破正在临点，“国家队”与民企积极良性竞争，关键瓶颈正在逐渐突破。
- **四象三：具身智能有望变得精准与有用，实现放量的基础。**1) 特斯拉 optimusV3 临近发布，预计为机器人小脑发展的集大成者，各类任务精准度快速提升。2) 模型大脑侧，Gemini3.0pro 意味着大模型多模态理解能力跃升，未来机器人能理解多种人类指令，结合场景完成泛化任务从不可能变得可能，这是具身智能真正放量的基础。3) 从产业信号看，特斯拉人形机器人有望迎量产元年，成为一切的起点。
- **四象四：AI 应用破土而生，大模型/AI Infra/高增长/高壁垒垂类软件值得关注。**1) 模型迭代、应用深化、无论是 C 端还是 B 端 token 调用进一步爆发。2) 计算机行业持仓在历史低位，具备交易与反弹条件。3) 计算机行业收入、利润增速于 2025 年初步见底，在 AI 的驱动下存在重回增长轨道可能，建议重点关注四类确定性较高的投资方向：一是大模型及超级入口，头部基模厂商高增的 ARR 与 Token 消耗量证明，基于个性化与生态化构建的个人及企业智能体已广泛铺开；二是 AI Infra，作为连接模型与落地的关键中间层，其直接决定 AI 的使用成本与效果上限，Databricks 与 Snowflake 的高增表现印证该环节的高价值属性；三是高增长垂类软件：动态漫、广告、医疗等垂类 AI 都展现着快速增长潜力；四是高壁垒垂类软件，依托行业 Know-how、专有数据、复杂流程及合规壁垒，此类应用能将大模型作为其核心能力的放大器，在避免被颠覆的同时实现价值跃升。
- **其他重点细分：1) 3D 打印：当创新遇到物理瓶颈，3D 打印开启序章。**中国 3D 打印市场重回较快增长，2034 年全球市场规模有望突破 1145 亿美元。我们认为随着 AI、空天、3C、机器人、汽车等产业不断突破创新边界，传统制造工艺已临近上限，难以满足散热、轻量化等实际需要，亟需 3D 打印技术突破传统上限，多维需求爆发有望加速产业发展。2) **金融 IT：降息周期全球资本活跃，跨境支付重塑金融生态。**2026 年美联储降息预期有望推升资本市场交易活跃度；“十五五”明确资本市场高质量发展，资本市场 IT 有望迎来新的建设浪潮；全球跨境支付创新加速变革，为数字人民币更大力度推广应用打开空间。3) **智能驾驶：正式进入产业化阶段。**①25 年 9 月，“中国 L2 强标”进入公开征求意见阶段。25 年 12 月，工信部正式公布我国首批 L3 级有条件自动驾驶车型准入许可，标志着我国 L3 级自动驾驶从测试阶段正式迈入商业化应用。②年末，小马智行第七代 Robotaxi 率先在广州实现了城市级单车 UE 模型转正，中美 Robotaxi 车队或处于爆发前夜。③以鸿蒙智行为代表的智能驾驶车型不断迭代，市场份额快速上升。

风险提示

- 行业竞争加剧的风险；技术迭代不及预期的风险；特定行业下游资本开支周期性波动的风险。



投资建议

■ 相关标的:

- 1) 算力: 寒武纪、东阳光、中际旭创、海光信息、协创数据、新易盛、神州数码、华丰科技、中芯国际、华虹半导体、兆易创新、大普微、长川科技、中微公司、源杰科技、品高股份、胜宏科技、兴森科技、中科曙光、禾盛新材、天普股份、有方科技、沪电股份、浪潮信息、东山精密、亿田智能、奥飞数据、云赛智联、瑞晟智能、科华数据、潍柴重机、金山云。
- 2) 空天: 西部材料、顺灏股份、航天动力、飞沃科技、普天科技、迈为股份、航天电子、复旦微电、超捷股份、蓝思科技、宇晶股份、斯瑞新材、广联航空、上海瀚讯、臻镭科技、航天宏图、中科星图、海格通信、航天机电、高华科技、陕西华达、航天电器、电科数字、佳缘科技、盟升电子、震有科技、通宇通讯、信维通信、铂力特等。
- 3) AI 应用: 阿里巴巴、腾讯控股、Minimax、智谱、金山办公、合合信息、星环科技、百度、金蝶国际、深信服、汉得信息、鼎捷数智、用友网络、税友股份、万兴科技、福昕软件、科大讯飞、卓易信息、泛微网络、美年健康、润达医疗、致远互联、石基信息、麦迪科技、恒为科技、美图公司、阜博集团、海天瑞声、汇量科技、焦点科技、阅文集团、中文在线、德才股份、虹软科技、金桥信息等。
- 4) 具身智能: 斯菱智驱、福赛科技、科森科技、新泉股份、恒立液压、三花智控、拓普集团、泛亚微透、浙江荣泰、领益智造、新益昌、蓝思科技、海康威视、大华股份。
- 5) 3D 打印: 华曙高科、银邦股份、飞沃科技、铂力特。
- 6) 金融 IT: 同花顺、东方财富、指南针、财富趋势、京北方、新大陆、宇信科技、新国都、拉卡拉。
- 7) 自动驾驶: 江淮汽车、赛力斯、小鹏汽车、禾赛、地平线机器人、世运电路。



内容目录

一、AI 演绎路径更加清晰.....	6
1.1 全球顶尖模型快速迭代	6
1.2 字节 AI 手机发布，AI 超级载体出现	8
二、四象一：一切仍指向算力，2026 国产算力爆发之年	11
2.1 一切仍然指向算力，Scaling Law 从训练侧向推理侧拓展	11
2.2 国产 GPU：产能+性能持续突破，26 年加速爆发	12
2.3 入口之战已打响，大厂持续加大投入.....	14
2.4 存储：AI 训练/推理持续拉动存储需求	15
2.5 超节点：产能+性能持续突破，26 年加速爆发	16
三、四象二：火箭瓶颈突破在即，太空算力打开应用空间	17
3.1 海外巨头竞逐太空算力，打开万亿美金市场	17
3.2 我国近期空天支持政策频出，行业发展有望加速.....	19
3.3 我国积极布局大型星座与太空算力.....	20
3.4 我国火箭发射瓶颈有望逐渐突破.....	21
四、四象三：具身智能有望变得精准与有用，实现放量的基础.....	23
五、四象四：AI 应用破土而生，大模型/AI Infra/高增长/高壁垒垂类软件值得关注	26
5.1 大模型：已广泛应用，向个性化与生态化深入发展.....	26
5.2 AI Infra：获取卖铲子的确定收益，直接奠定应用成色.....	28
5.3 高增长场景应用：广告营销、动态漫、医疗等有望跑通飞轮	29
5.4 高壁垒垂类应用：具备数据与 workflow 护城河，对客价值更为显著	31
六、3D 打印：当创新遇到物理瓶颈，3D 打印开启序章.....	32
七、金融 IT：降息周期全球资本活跃，跨境支付重构金融生态	35
八、智能驾驶：L2 强标落地，L3 元年，Robotaxi 单车 UE 转正	37
风险提示	43

图表目录

图表 1： Gemini 3 在多项评分基准大幅领先	6
图表 2： Hopper (H100) 与 Blackwell (B200/GB300) 对比.....	8
图表 3： 豆包手机助手执行用户任务.....	9
图表 4： API Agent 和 GUI Agent 在完成"3 月 8 日下午 4:00 在 Google 日历上安排 1 小时会议"任务时的区别	10
图表 5： API Agent 与 GUI Agent 在关键维度上的比较	10
图表 6： 人工智能太极生两仪，两仪生四象.....	11



图表 7: 三大 Scaling 定律	11
图表 8: Interleaved Thinking 推理范式	12
图表 9: 2020-2028 年中国智能算力规模及预测	12
图表 10: 2024 年中国 AI 加速芯片市场竞争格局	12
图表 11: 腾讯云宣布全面适配主流国产芯片	13
图表 12: 国产通用 GPU 从“可用”向“好用”升级	13
图表 13: OpenAI 与 Anthropic 收入曲线	14
图表 14: 微信淘宝等超级 APP 封禁豆包手机	14
图表 15: 豆包手机扩充 VIVO、联想、传音等合作商	14
图表 16: 北美科技巨头最新季度财务状况	15
图表 17: 大模型对训练、推理的存储需求	15
图表 18: 存储现货价格变动情况	16
图表 19: 存储合约价变动情况	16
图表 20: 主流并行计算方式及算力资源扩展策略	16
图表 21: Scale-up 网络关键技术分层分析	16
图表 22: 昇腾 384 超节点与曙光 AI 超集群对比	16
图表 23: 美国初创科技公司 Starcloud	17
图表 24: 马斯克认为运行人工智能成本最低的方式将在太空进行	18
图表 25: 2025 年以来我国低轨卫星互联网组网明显提速	20
图表 26: 北京太空数据中心建设工作推进大会进行	21
图表 27: 朱雀三号首飞的技术突破	22
图表 28: 长征十二号甲遥一运载火箭飞行试验任务	23
图表 29: 空天时代临点将至	23
图表 30: 2027 年中国具身智能市场规模有望突破 1.25 万亿元	24
图表 31: 2025 年具身智能投资分布中人形机器人为第一大细分赛道	24
图表 32: Optimus V2.5 无需连接电线，可自主完成行走	25
图表 33: ScreenShot-Pro 评测的部分数据集，包含各类桌面应用和场景	25
图表 34: Nano Banana Pro 图像生成：创建一个信息图，展示如何制作香茅茶	25
图表 35: 软件行业营收及增长率	26
图表 36: 软件行业归母净利润及增长率	26
图表 37: OpenAI 与 Anthropic 收入曲线	26
图表 38: 个人大模型使用场景及工作流	27
图表 39: OpenAI AgentKit 使用案例	28
图表 40: Databricks AI 平台	29
图表 41: Snowflake Cortex AI	29



图表 42:	Applovin 业务模式	30
图表 43:	Applovin 营收及利润	30
图表 44:	抖音漫剧市场商业化指标	31
图表 45:	Palantir AIP 系统架构	32
图表 46:	Palantir 收入及利润 (亿美元)	32
图表 47:	2024 年全球 3D 打印市场规模已达 219 亿美元	32
图表 48:	2025 年国内 3D 打印行业投融资总额预计达 84 亿元	32
图表 49:	微通道液冷板制造中 3D 技术路线前景更好	33
图表 50:	流道拓扑优化后可以显著提升散热性能	33
图表 51:	发动机成本占火箭总成本 50%	34
图表 52:	第三代猛禽发动机通过 3D 打印工艺, 将前两代的多零部件整合为一	34
图表 53:	3D 打印制造轻量化弹性体为 IRON 机器人塑造“肌肉”	34
图表 54:	Kengoro 机器人使用金属 3D 打印打造多孔框架为全身电机降温	34
图表 55:	超薄均热板能够高效消除局部热点, 保障高功率柔性电子设备稳定运行	35
图表 56:	OPPO Find N5 采用 3D 打印打造铰链、翼板及外传轴中框	35
图表 57:	3D 打印汽车零部件在保障强度的同时能够有限减重	35
图表 58:	保时捷提供定制 3D 打印座椅	35
图表 59:	过去 24 个月, 全球央行多次降息	36
图表 60:	“十五五”规划明确资本市场“高质量发展”核心主线	36
图表 61:	2024 年全球跨境支付规模增长至 194.6 万亿美元	37
图表 62:	三大平台协同解决了传统跨境支付的效率与成本痛点	37
图表 63:	新旧国标对比: 适用对象由 M1 扩展至 N1 类汽车, AEBS 系统激活速度区间及误检&漏检要求升级	38
图表 64:	欧洲、美国、日本均通过强制法规要求安装 AEBS 系统	38
图表 65:	AEBS 激活速度区间扩展后需要更远的探测距离	39
图表 66:	新国标新增多项误响应试验, 3D 毫米波雷达在技术层面易造成误触发	40
图表 67:	L2 强标对复杂环境下的多目标状态识别提出较高要求	40
图表 68:	梅赛德斯 L3 级 DRIVE PILOT 系统配备一颗激光雷达	41
图表 69:	梅赛德斯 L3 级 DRIVE PILOT 系统设置 10 秒响应接管时间	41
图表 70:	国内车企密集公布 L3 进展, 已公布车型均标配 4-5 颗激光雷达	41
图表 71:	国内 Robotaxi 落地车型搭载 7-10 颗激光雷达	42
图表 72:	2025 年 11 月鸿蒙智行单月交付破 8 万辆	43
图表 73:	鸿蒙智行用时 43 个月累计交付量破百万辆	43



一、AI 演绎路径更加清晰

1.1 全球顶尖模型模型快速迭代

2025 年 11 月 19 日谷歌 Gemini 3 发布，跃进式的得分提升、强大的多模态理解力、更加多样化的 UI、惊艳的前端能力，让大模型的进展迈出显著一步。

基础思考能力上, Humanity's Last Exam 是衡量 AI 能否解决人类顶尖难题的终极试金石。在 Gemini 3 之前，Gemini 2.5 Pro 的得分是 21.6%，Claude Sonnet 4.5 仅为 13.7%。Gemini 3 Pro 交出的答卷是 37.5%（无工具）和 45.8%（带工具）。ARC-AGI-2 测试，旨在衡量模型处理从未见过的新颖推理任务的能力，而非死记硬背。Gemini 3 Pro 拿下了 31.1% 的分数，而 GPT-5.1 仅为 17.6%，Gemini 2.5 Pro 甚至只有 4.9%。这意味着它开始展现出一种接近人类的流体智力，能够在没有大量训练数据覆盖的领域进行抽象推理。

在多模态领域 Gemini 3 尤其惊艳，MMMU-Pro 的 81.0%，CharXiv Reasoning 的 81.4% 都超越了对手。而在理解截图的测试 ScreenSpot-Pro 的 72.7% 中，Gemini 3 的得分是 Claude Sonnet 4.5 的两倍，GPT-5.1 的二十倍。这对于构建能够真正理解和操作图形界面的 AI 代理至关重要。

图表1: Gemini 3 在多项评分基准大幅领先

Benchmark	Description		Gemini 3 Pro	Gemini 2.5 Pro	Claude Sonnet 4.5	GPT-5.1
Humanity's Last Exam	Academic reasoning	No tools	37.5%	21.6%	13.7%	26.5%
		With search and code execution	45.8%	—	—	—
ARC-AGI-2	Visual reasoning puzzles	ARC Prize Verified	31.1%	4.9%	13.6%	17.6%
GPQA Diamond	Scientific knowledge	No tools	91.9%	86.4%	83.4%	88.1%
AIME 2025	Mathematics	No tools	95.0%	88.0%	87.0%	94.0%
		With code execution	100%	—	100%	—
MathArena Apex	Challenging Math Contest problems		23.4%	0.5%	1.6%	1.0%
MMMU-Pro	Multimodal understanding and reasoning		81.0%	68.0%	68.0%	76.0%
ScreenSpot-Pro	Screen understanding		72.7%	11.4%	36.2%	3.5%
CharXiv Reasoning	Information synthesis from complex charts		81.4%	69.6%	68.5%	69.5%
OmniDocBench 1.5	OCR	Overall Edit Distance, lower is better	0.115	0.145	0.145	0.147
Video-MMMU	Knowledge acquisition from videos		87.6%	83.6%	77.8%	80.4%
LiveCodeBench Pro	Competitive coding problems from Codeforces, ICPC, and IOI	Elo Rating, higher is better	2,439	1,775	1,418	2,243
Terminal-Bench 2.0	Agentic terminal coding	Terminus-2 agent	54.2%	32.6%	42.8%	47.6%
SWE-Bench Verified	Agentic coding	Single attempt	76.2%	59.6%	77.2%	76.3%
t2-bench	Agentic tool use		85.4%	54.9%	84.7%	80.2%
Vending-Bench 2	Long-horizon agentic tasks	Net worth (mean), higher is better	\$5,478.16	\$573.64	\$3,838.74	\$1,473.43
FACTS Benchmark Suite	Held out internal grounding, parametric, MM, and search retrieval benchmarks		70.5%	63.4%	50.4%	50.8%
SimpleQA Verified	Parametric knowledge		72.1%	54.5%	29.3%	34.9%
MMMLU	Multilingual Q&A		91.8%	89.5%	89.1%	91.0%
Global PIQA	Commonsense reasoning across 100 Languages and Cultures		93.4%	91.5%	90.1%	90.9%
MRCR v2 (8-needle)	Long context performance	128k (average)	77.0%	58.0%	47.1%	61.6%
		1M (pointwise)	26.3%	16.4%	not supported	not supported

来源：腾讯科技，国金证券研究所

2025 年 12 月 11 日 OpenAI 发布 GPT-5.2，为专业工作设计的顶尖模型。

GPT-5.2 旨在为人们带来更多经济价值。该模型在制作电子表格、设计演示文稿、编写代码、识别图像、理解长文本上下文、使用工具以及处理复杂的多步骤项目方面表现更佳。

GDPval 评测是一项覆盖 44 个职业、用于衡量明确知识型工作任务的评估。在 GDPval 测试中，模型尝试完成定义明确的知识型工作，内容涵盖美国 GDP 贡献度最高的 9 个行业中的 44 种职业。任务要求生成真实的工作成果，例如销售演示文稿、会计表格、急诊排班表、制造业图表或短视频。根据人类专家评审的结果，GPT-5.2 Thinking 在 70.7% 的高难度知识型工作任务上，表现优于行业顶尖专家，或与其持平。这些任务包括制作演示文稿、电子表格以及其他专业产出。GPT-5.2 Thinking 完成任务的速度大约是专家的 3 倍，而成本只有大约



1%。

12月1日 DeepSeek-V3.2 正式版发布，国产模型引领开源创新。

在公开的推理类 Benchmark 测试中，DeepSeek-V3.2 达到了 GPT-5 的水平，仅略低于 Gemini-3.0-Pro；DeepSeek-V3.2-Special 是 DeepSeek-V3.2 的长思考增强版，在主流推理基准测试上的性能表现媲美 Gemini-3.0-Pro，成功斩获 IMO 2025（国际数学奥林匹克）、CMO 2025（中国数学奥林匹克）、ICPC World Finals 2025（国际大学生程序设计竞赛全球总决赛）及 IOI 2025（国际信息学奥林匹克）金牌。其中，ICPC 与 IOI 成绩分别达到了人类选手第二名与第十名的水平。

DeepSeek V3.2 系列模型提出多项创新：延续使用 DSA 稀疏注意力机制、加码后训练、大规模合成数据。

DSA 全称 DeepSeek Sparse Attention，旨在降低计算复杂度，同时在上下文场景中还能保持模型性能。举个例子，当你面前有一本超厚的书（假设页数上万页），现在需要从这本书里查询特定信息，要求只能一页一页地看，且需要记忆看过的内容，这就是传统 AI 模型处理长文本时的困境——计算量随文本长度呈平方级增长。而 DSA 的作用如同将这种“大海捞针”的方式转为“精准定位”，不是每页都要看，而是能够只关注与目标信息最相关的几页，大幅提升了模型记忆与理解效率。

此外，DeepSeek 意识到开源模型的后训练投入普遍不足，限制了模型任务表现，于是采用更激进的方式，将后训练算力预算调整到超过预训练成本的 10%。

DeepSeek 这次用到了大规模智能体数据合成，根本原因在于缺乏足够多样化的真实训练环境，无法让智能体具备任务泛化能力。从测试评估结果看，使用合成数据使模型泛化能力提升显著。

我们认为模型的迭代还未到瓶颈，2026 年模型的进展会更精彩：

1) 靠 Scaling law 提升模型能力的路径依旧有效。

DeepSeek V3.2 后训练规模扩展到预训练规模 10%，预计未来比例还会提高，而且 DeepSeek 团队在论文表示由于训练算力有限，DeepSeek-V3.2 的世界知识广度还是落后于 Gemini 3 pro 这样的顶尖闭源模型，团队计划未来进一步扩大预训练规模。同时 DeepSeek 大量使用合成数据有效说明不用担心数据会遇到瓶颈。

2) DeepSeek-V3.2 提出的 DSA 机制展示出强大算法创新能力，不必担心大模型技术创新已经到达瓶颈。

3) 大模型训练的硬件基础升级。

英伟达的 Hopper 架构正在转向 Blackwell 架构，Blackwell 相比前代在单卡算力、显存带宽、显存容量、以及集群互联都大幅提升，这对大模型训练的意义一方面是加速和降低成本，另一方面是可以使用更大的 Batch Size（模型更新学习内容前一次性处理的训练样本数量），这对训练稳定性有帮助，更大的 Batch Size 能够更准确地估计整个数据集的梯度，从而使学习过程更加稳定，而较小的批次则会产生噪声过大且特征过于明显的信号，这可能导致模型的学习路径出现不稳定的跳跃。



图表2: Hopper (H100) 与 Blackwell (B200/GB300) 对比

特征	NVIDIA HOPPER (H100)	NVIDIA BLACKWELL (B200/GB300)
建筑概论	Hopper (H100 Tensor Core, 2022)	Blackwell (B200/GB300 Tensor Core, 2024/25)
每个GPU的SM数量	约80个SM (GH100)	每个芯片约 80 个 SM (B200), 2 个芯片共封装 (nm)
晶体管 (总数)	~80B (整体式) (nm)	~208B (2x芯片组) (nm)
进程节点	台积电 4N (定制 5nm)	台积电 4NP (新型 4nm 变体) (nm)
芯片包装	单GPU芯片	双GPU芯片组, 带10 TB/s NVLink芯片间链路 (nm)
每个SM的共享/L1内存	256 KB (可配置)	128 KB (更小, 统一) (nm)
统一的 L2 缓存	总计 50 MB (分块)	65 MB 单体式 (nm)
指令缓存	每个SM的L0缓存	更大的统一 L0 (减少指令获取延迟)
HBM 容量	80 GB HBM3 (nm)	192 GB HBM3e (B200) / 288 GB (GB300) (nm) (nm)
HBM带宽	3.35 TB/s (nm)	8.0 TB/s (nm)
NVLink (芯片间)	NVLink4 - 900 GB/s	NVLink5 - 1.8 TB/s (nm)
NVSwitch (DGX pod 中的链接)	~50 TB/s (在 16 路 GPU 的 DGX A100 上)	每个 72-GPU NVL72 pod 的处理能力为 130 TB/s (nm)
张量核心精度 (密集)	FP16/BF16、FP8、INT8	FP16/BF16、FP8、FP6、FP4 (nm) (nm)
FP8/密集型 TFLOPS	峰值约为 4-5 PF	~9 PF (nm) (每个 GPU)
FP4/密集型 TFLOPS	(不支持)	~18 PF (nm) (每个 GPU, 新模式)
FP64 TFLOPS (张量)	60 TF (张量) (nm)	较低 (弱化) (nm)
峰值功率 (TGP)	约700瓦 (nm)	约 1,200-1,400 瓦 (nm)

来源: intuitionlabs.ai, 国金证券研究所

谷歌也在2025年11月宣布其第七代TPU Ironwood将在未来几周内全面上市。Ironwood专为应对最严苛的工作负载而打造: 从大规模模型训练和复杂的强化学习到高容量、低延迟的 AI 推理和模型服务, 它都能轻松胜任。与 TPU v5p 相比, Ironwood 的峰值性能提升了 10 倍; 与 TPU v6e (Trillium) 相比, 其单芯片在训练和推理工作负载下的性能均提升了 4 倍以上, 使其成为谷歌迄今为止性能最强大、能效最高的定制芯片。我们认为谷歌的芯片升级也会对模型迭代带来显著促进作用。

1.2 字节 AI 手机发布, AI 超级载体出现

手机本身是人们工作与生活信息流的集大成者, 全能的 AI 手机助手出现将全面重构所有人的工作与生活。

2025 年 12 月字节跳动在中兴合作的工程样机 nubia M153 上发布了豆包手机助手技术预览版。目前该版本已面向开发者和科技爱好者少量发售, 售价 3499 元。用户不仅可以通过语音、侧边键或豆包 Ola Friend 耳机直接唤醒豆包, 实现交互以及调用语音通话、视频通话和屏幕共享等功能。此外, 豆包手机助手可以通过 AI 直接操控手机功能, 用户输入查票订票、商品下单、批量下载文件等信息, 豆包手机助手可在多个 App 之间跳转, 帮用户完成多个平台搜索同款商品、比价、选择最便宜的商品下单等操作。

我们认为豆包手机助手是 AI 在 C 端从单一的语音助手进化为真正会行动的助理的重要里程碑, 不仅能理解用户意图, 还能跨应用自主执行复杂任务, 体现了 AI 在真实场景中的巨大进展, 也预示着手机交互方式和用户效率将被系统性重塑。



图表3：豆包手机助手执行用户任务



来源：字节跳动，界面新闻，国金证券研究所

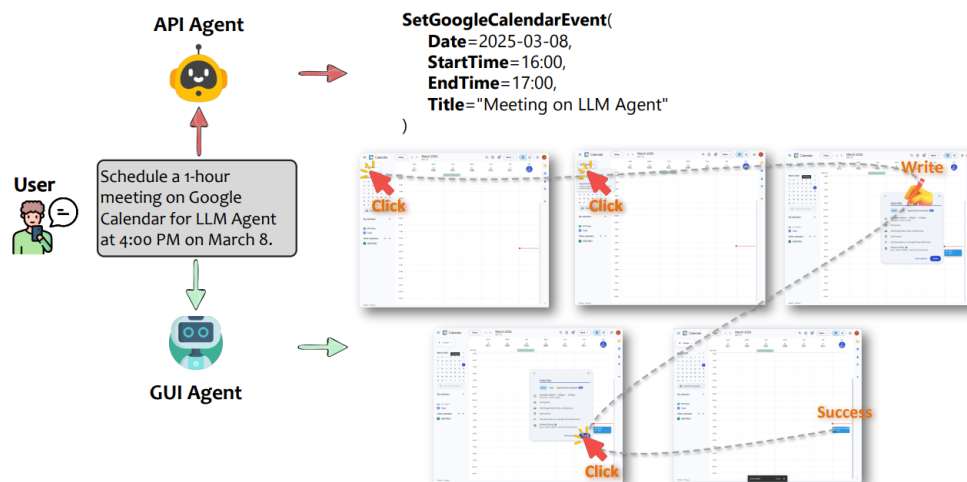
基于图形用户界面（GUI）的触控是移动互联网时代的核心交互方式，但也存在操作流程复杂的问题以及 App 间信息孤岛的问题。如今大模型正推动移动端交互革命：用户从复杂的点击跳转中解放，只需下达自然语言指令，AI Agent 即可跨越应用边界完成任务。针对如何打破应用壁垒，行业目前形成了两套演进逻辑：

API Agent 路线：通过标准化协议让开发者主动接入。如苹果通过 App Intents 框架让开发者可以将应用接入 Apple Intelligence。

GUI Agent 路线：利用多模态模型模拟人类视觉和操作。这种方式无需开发者配合，凭借通用视觉能力直接操作各类 App，如豆包手机助手。



图表4: API Agent 和 GUI Agent 在完成"3月8日下午4:00在Google日历上安排1小时会议"任务时的区别



来源:《API Agents vs. GUI Agents: Divergence and Convergence》, Chaoyun Zhang 等, 国金证券研究所

在效率、隐私、可靠性等方面 API Agent 更有优势,但是 GUI Agent 的最核心优势在于灵活通用,不需要应用厂商主动适配,是打通碎片化应用生态的利器。预计模型厂商、应用厂商和手机厂商为了抢夺流量入口,会在两种路线选择上进行激烈的博弈。

图表5: API Agent 与 GUI Agent 在关键维度上的比较

维度	API Agent	GUI Agent
模态	依赖于文本的 API 调用	依赖屏幕截图或辅助功能树
可靠性	通常较高,具有定义明确的端点	较低,受视觉解析和布局变化影响
效率	单次调用即可完成复杂任务	需要多次类似人类用户的操作步骤
可用性	局限于已发布或预定义的 API	可操作任何可见的 UI 元素
灵活性	受现有 API 的限制	对新功能或未公开功能具有高度适应性
安全性	可通过细粒度的端点控制进行管理	风险较高,因为具有对 UI 元素的广泛访问权限
可维护性	只要 API 版本保持稳定,维护较易	在 UI 重构或改版时容易失效
透明度	通常是隐藏的、后端驱动的	逐步进行,视觉上可追溯
类人交互	纯程序化交互	模拟在屏幕上的真实用户操作

来源:《API Agents vs. GUI Agents: Divergence and Convergence》, Chaoyun Zhang 等, 国金证券研究所

我们认为 AI 手机是其他 AI 终端爆发的前置条件,预计全球顶尖大厂将极致投入到这场新入口之战。

两仪已定,四象初生:算力为基、空天时代、具身智能、AI 应用精彩不断。模型的快速迭代与 AI 手机聚集数据流并拥有智能会在 2026 年推动 4 个领域明确的投资机会。1)算力的争夺更加激烈,且上限非常高。2)空天时代成为算力领域斜率最高的方向。3)机器人可能变得能听懂话,且结合环境执行多种任务。4)AI 应用渗透率临近甜蜜点(5%~20%为新科技渗透率甜蜜区)。



图表6：人工智能太极生两仪，两仪生四象



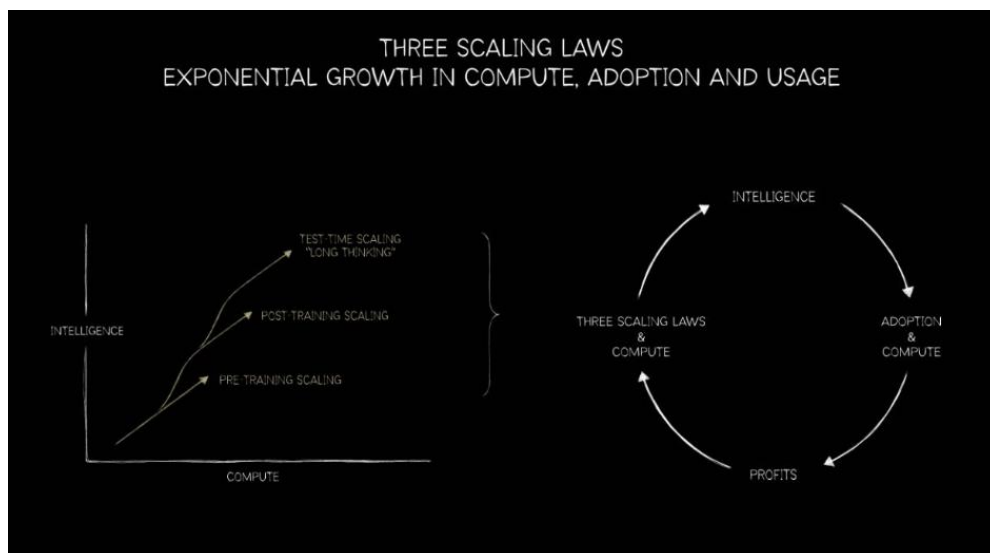
来源：国金证券研究所绘制

二、四象一：一切仍指向算力，2026 国产算力爆发之年

2.1 一切仍然指向算力，Scaling Law 从训练侧向推理侧拓展

大模型三大 **Scaling** 定律：预训练（pre-training）/后训练（post-training）/推理深度思考（test-time long thinking）。算力与应用循环促进：算力 Scaling—>智能提升（Intelligence）—>应用的广泛采用（Adoption）—>经济效益（Profits）—>算力 Scaling。GPTo1 之后，模型 **Scaling law** 从单一的预训练（pre-training）向三大 **Scaling** 转变。1）模型在回应用户之前产生很长的内部思考链，思考时间越长，答案的质量就越高；2）Multi-agent 进一步提升模型性能，但其 Token 消耗量往往达到对话聊天的数倍。

图表7：三大 Scaling 定律



来源：英伟达官网，国金证券研究所

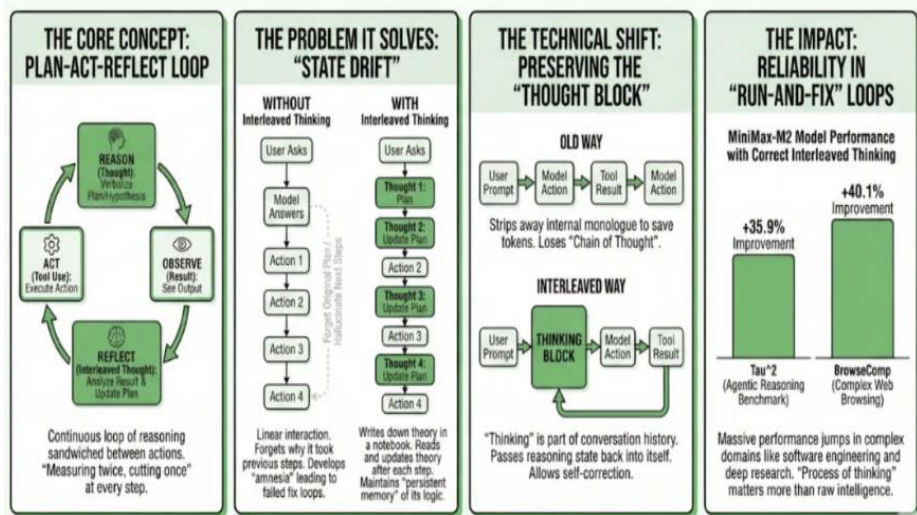
训练层面，预训练重启，后训练加码。预训练方面，DeepSeek 团队在论文中坦言，受训练算力约束，DeepSeek-V3.2 的世界知识广度仍落后于领先闭源模型（如 Gemini 3 Pro），后续将通过扩大预训练规模补齐能力上限。此外，V3.2 往往需要更多 token 才能逼近 Gemini-3.0-Pro 等模型的输出质量，未来将聚焦提升推理链的“智能密度”，以更少 token 达到同等效果；后训练投入强度提升：DeepSeek 认为开源模型后训练投入普遍不足并制约任



务表现，因此将后训练算力预算上调至超过预训练成本的 10%。

推理方面，MiniMax M2、DeepSeek V3.2 等国产大模型，把“思考 → 行动 → 观察 → 再思考”的动态循环模式（Interleaved Thinking）融入推理流程，正成为 Agent 模型标配。我们认为，**国内 Interleaved Thinking 范式渗透，能够显著提高 Agent 的准确性和规划能力，2026 年 Token 消耗与推理需求有望加速爆发。**

图表8: Interleaved Thinking 推理范式

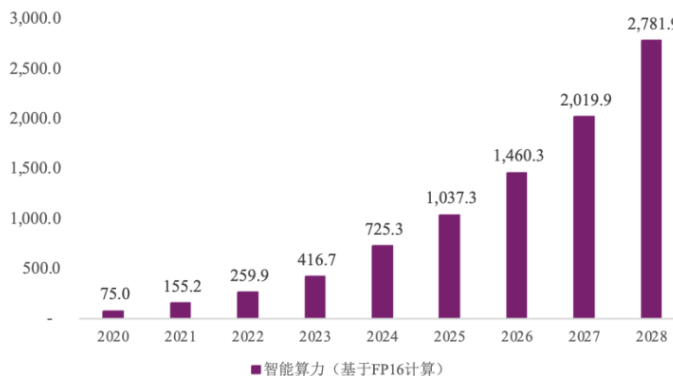


来源：英伟达官网，国金证券研究所

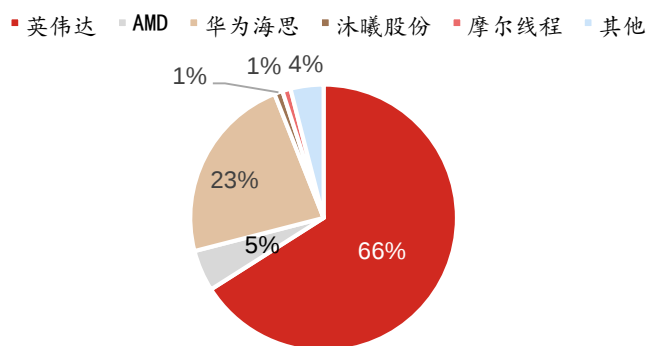
2.2 国产 GPU：产能+性能持续突破，26 年加速爆发

智算中心持续扩容，国产替代加速。根据 IDC 数据，2020 年中国智能算力规模为 75.0 EFLOPS，到 2028 年预计将达到 2,781.9 EFLOPS，预计 2020-2028 年复合增长率达到 57.1%。随着地缘政治紧张局势推动中国企业寻求本地替代方案，以及国产芯片技术的稳步提升，国内云服务提供商正在加速构建异构环境（将不同类型的芯片结合使用，如 CPU、GPU、国内替代芯片）。例如，腾讯云已将其异构计算平台全面兼容所有主流国产芯片，以满足内部开发和客户对 AI 算力的需求。根据 Bernstein Research，2024 年中国 AI 加速芯片市场中，英伟达、AMD 市场份额分别为 66%、5%，合计占比达 71%。但受益于国产替代趋势及供应链安全需求，国内计算芯片公司正迅速提升。其中，华为海思/沐曦/摩尔市场份额已分别达到 23%/1%/1%。

图表9: 2020-2028 年中国智能算力规模及预测



图表10: 2024 年中国 AI 加速芯片市场竞争格局



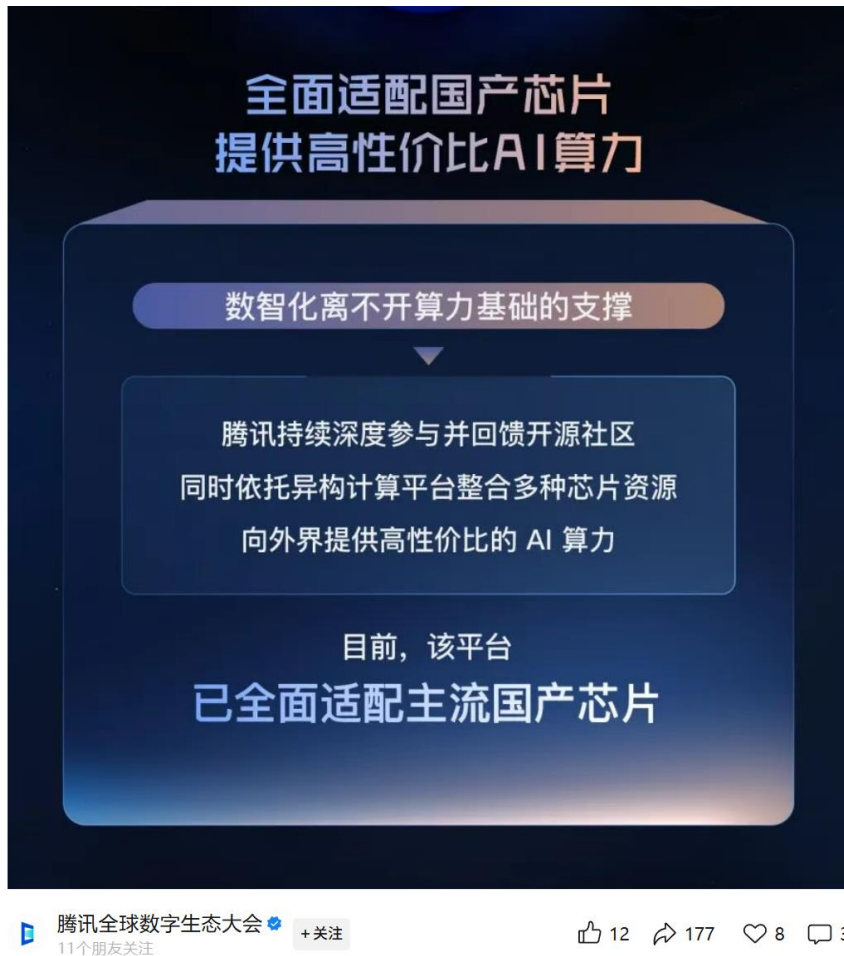
来源：沐曦招股说明书，国金证券研究所

来源：沐曦招股说明书，国金证券研究所

CSP 厂商加速适配，助力国产芯片生态建设。英特尔的 X86 生态、英伟达的 CUDA 生态之所以难以撼动，核心在于形成了“芯片-软件-应用”的闭环。而当前国产阵营中，华为昇腾、阿里平头哥、壁仞科技等芯片厂商各有技术路线，生态分散问题显著。腾讯集团高级执行副总裁汤道生在交流中坦言，不同参数规模的 AI 模型需要适配不同芯片配置，当前只能通过与多家厂商合作实现场景覆盖。**百度、阿里等企业加速适配国产芯片，推动“芯片-模型-应用”闭环形成。**



图表11: 腾讯云宣布全面适配主流国产芯片



来源：腾讯全球数字生态大会，国金证券研究所

国产通用 GPU 从“可用”向“好用”升级。国产 GPU 在性能指标、软件生态、应用适配等方面与 NV 最先进一代仍有差距，但已基本追平 H20、A100 等，且在本地化服务、政策支持、成本控制等方面具备优势。随着资本持续注入，国产企业有望在细分场景实现突破，逐步扩大市场份额。**1) 算力指标上：**国内多数头部企业主流在售产品的 FP16/BF16 在 100-300 TFLOPS 左右，处于英伟达 A100 产品阶段，少数厂商通过先进封装等方式实现接近英伟达 H100 产品的算力，为国内最先进水平；**2) 显存方面：**国内企业结合自身产品特点，分别选择 HBM2e、HBM2、GDDR 等显存类型，显存带宽在 0.5-2 TB/s 左右。

图表12: 国产通用 GPU 从“可用”向“好用”升级

参数 / 厂商	平头哥	NV		华为	壁仞
型号	PPU	A800	H20	昇腾 910B	104P
显存容量	96G	80G	96G	64G	32G
显存类型	HBM2e	HBM2e	HBM3	HBM2	HBM2e
片间带宽 (GB/s)	700	400	900	392	256
PCIe	5.0 × 16	4.0 × 16	5.0 × 16	4.0 × 16	5.0 × 16
功耗 (W)	400	400	550*	350	300

来源：芯东西微信公众平台，国金证券研究所

供给侧：中芯国际作为中国集成电路领导者，产能/利用率持续提升。11 月 13 日，中芯国际披露 2025 年第三季度财报，2025Q3 中芯国际营收 171.62 亿元，环比增长 6.9%，毛利率 25.5%；月产能为 102.28 万片(折合 8 英寸)，同比增加产能约 13.85 万片(折合 8 英寸)；产能利用率达 95.8%，环比增长 3.3 个百分点。中芯国际产能全球第三，约为台积电的三分之一。公司资本支出维持高位，未来产能或将继续增加。

图表13: OpenAI 与 Anthropic 收入曲线



来源：TrendForce，国金证券研究所

2.3 入口之战已打响，大厂持续加大投入

入口不再仅限于手机，而是演变为“OS 级智能体”与“超级 APP 层面较量。1) 超级 APP 的入口之争已经打响。2025 年 12 月 24 日，字节跳动旗下 AI 应用豆包宣布日均活跃用户数 (DAU) 突破 1 亿；千问 App 近期持续扩大投流，截止 12 月 10 日 (公测 23 天) 月活已突破 3000 万，下成全球增长最快 AI 应用。阿里巴巴在内部沟通中提及，千问 C 端事业群的首要目标是将千问打造成为一款超级 APP，成为 AI 时代用户的第一入口；2) “入口之战”从流量分层推到了系统权限层，已成必争之地。豆包绕开传统接口，用“看屏幕+代操作”的方式在现有生态之上再搭一层“AI 操作系统”，AIOS 直接触碰微信、支付宝等超级 App 的商业命门，传统 App 时代的游戏规则面临挑战，微信、淘宝等超级 App 陆续封禁豆包手机权限。但值得注意的是，华为、小米、荣耀、OPPO、vivo 等品牌均在旗舰机型里注入 AI Agent 能力，模型厂商仍有望通过开源或广泛结盟来绕过 APP 厂商的封锁。

图表14: 微信淘宝等超级APP封禁豆包手机

豆包手机刚上线，就被微信淘宝“拉黑”了

中国企业家杂志 · 2025年12月04日 15:36

字节为何敢于打破行业默契？

“ 电商比价、浏览微信，字节为何敢于打破行业默契？ ”

字节的硬件野心，从未加以掩饰。但当它踏入了手机和超级App们的领地，甚至希望重新分配移动互联网的控制权时，很快撞上了阻力墙。

12月1日，豆包发布手机助手，在中兴nubia M153系列手机上，获得了操作系统层面的高权限。用户通过唤醒“豆包”，可以完成跨应用比价、购物、撰写报告等任务；也可以打开微信，自动回复消息，玩小程序游戏等。

图表15: 豆包手机扩充 VIVO、联想、传音等合作商

豆包手机扩充合作厂商 vivo、联想、传音在列

2025年12月19日 12:10 手机之家

据界面新闻今日报道，字节跳动正与 vivo、联想、传音等硬件厂商推进 AI 手机合作计划，拟通过预装 AIGC 插件获取用户入口。有 vivo 员工证实，字节跳动已与 vivo 确立合作意向，具体方案仍在商讨中。截至IT之家发稿，相关企业未对合作事宜作出回应。



来源: 36Kr, 国金证券研究所

来源：新浪财经，国金证券研究所

全球科技巨头军备赛，现金流仍支撑进一步加大投入。全球科技巨头对 AI 的投入进入军备竞赛阶段，但综合考虑各家的自由现金流以及在手的现金及现金等价物和短期投资看：北美科技巨头除 Meta 外，即使在假设人工智能（AI）的贡献对自由现金流的推动作用尚未完全显现的情况下，2026 年的资本支出压力预计仍处于可控范围。同时，我们认为 AI 需求的不确定性有望训练和推理相关的投入资本回报率 ROIC。吴泳铭表示从需求侧和供给侧两个维度来看：从需求侧，有两大核心驱动力支撑 ROIC 的提升，一是模型能力持续进阶，基座模型、视频生成模型及未来的全模态模型等，能力仍在不断突破，规模化法则的发展尚未触顶，谷歌 Gemini 3



人工智能模型的进展就印证了这一点。随着模型能力提升，AI 可承载的任务与适配的场景持续增多。二是 AI 任务在各行业的渗透率不断提高，两大动能叠加，未来三年内 AI 需求的增长具有高度确定性。

图表16：北美科技巨头最新季度财务状况

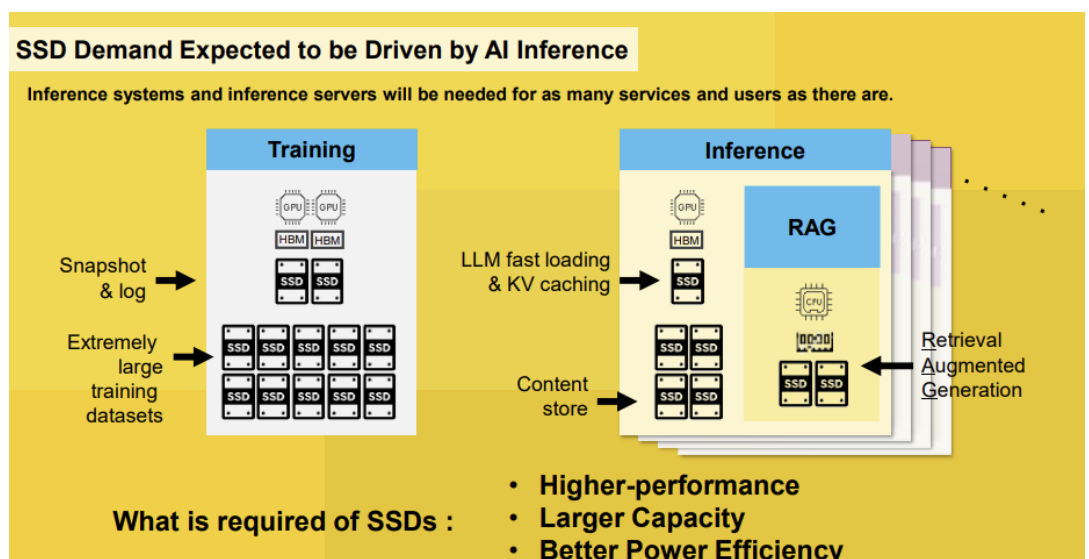
公司名称	经营性现金流 (OCF) (亿美元)	最新季度资本支出 (CapEx) (亿美元)	自由现金流 (FCF) (亿美元)	现金及现金等价物和 短期投资 (亿美元)
Google	484	240	244	985
Microsoft	451	349	102	1020
Meta	300	194	106	444
Apple	297	32	265	547
Amazon	1307(TTM)	342 (估算)	148 (TTM)	942

来源：富途，各公司官网，今日头条，国金证券研究所（注：亚马逊 OCF 和 FCF 为披露的 12 个月数据，其他未特别标注 TTM 的为单季度数据。）

2.4 存储：AI 训练/推理持续拉动存储需求

AI 大模型推理拉动存储需求快速增长。训练侧，SSD 不仅负责储存模型参数，包含不断更新的权重和偏差，而且可以创建检查点（check point），以定期保存 AI 模型训练进度，即使训练中断也能从特定点恢复；推理侧，主要包括三个方面：1）LLM fast loading & KV caching：快速装载模型相关内容并支持 KV 缓存，加速推理响应；2）Content store：推理侧需要内容存储作为在线供给（可理解为面向推理的内容/数据仓储）；3）RAG：检索增强生成把“检索”引入推理闭环，推理链路对存储访问更敏感。存储芯片市场正经历一场由 AI 驱动的、前所未有的全面缺货与涨价潮。进入第四季度非但没有趋缓，反而有进一步加剧的迹象。

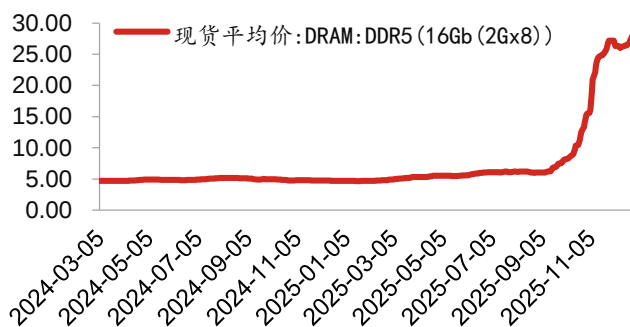
图表17：大模型对训练、推理的存储需求



来源：kioxia，国金证券研究所

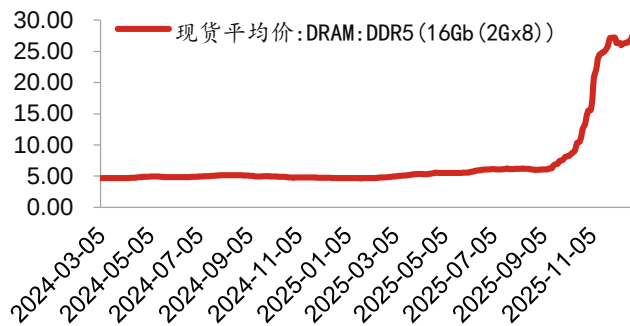


图表18: 存储现货价格变动情况



来源: Wind, 国金证券研究所

图表19: 存储合约价变动情况



来源: Wind, 国金证券研究所

2.5 超节点: 产能+性能持续突破, 26 年加速爆发

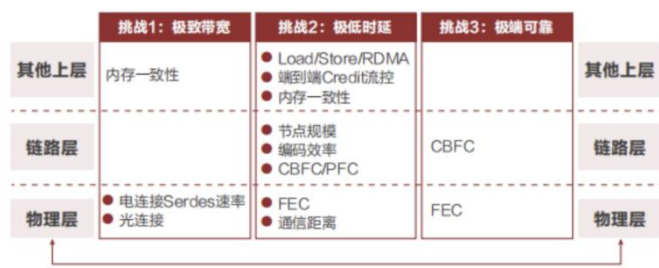
Scale-up 与 Scale Out 持续拓展。大模型需要采用多机多卡集群的方式进行训练, 主流的并行训练方式有数据并行、模型并行(张量并行、流水线并行)、专家并行等。张量并行和专家并行的并行计算方式, 每次迭代数据量达几百 GB, GPU 互联需要超高带宽和极低时延, 以压缩 GPU 之间的通信开销成本。随着大模型参数规模指数级增长, 单卡有效算力难以应对超大规模的算力需求, 高带宽、低延迟的卡间互连技术成为关键竞争要素, 对 GPU 集群互连拓扑结构也提出了更高要求, 驱动了纵向扩展(Scale-up)与横向扩展(Scale-out)网络架构的技术演进。

图表20: 主流并行计算方式及算力资源扩展策略

切分方式	通信操作	通信量(单卡)	对网络的需求
张量并行(TP)	AllReduce	百GB级别	(超)节点内高速互联
专家并行(EP)	All-to-All	百GB级别	(超)节点内高速互联
流水并行(PP)	Send/Recv	MB级别	节点间高速互联
数据并行(PP)	AllReduce	GB级别	节点间高速互联

来源: 通信世界网, 国金证券研究所

图表21: Scale-up 网络关键技术分层分析



来源: 通信世界网, 国金证券研究所

国产超节点陆续发布, 2025 年是超节点产业的“加速渗透年”。从华为昇腾 384 超节点真机在 2025 世界人工智能大会期间首秀, 到 12 月 18 日中科曙光 scaleX 万卡超集群真机亮相, 超节点已经加速从技术突破走向产品落地。华为、中科曙光、阿里、百度等头部企业均发布超节点产品, 中兴通讯、浪潮、新华三、超聚变等企业纷纷跟进推出解决方案, 形成全面竞争格局。

图表22: 昇腾 384 超节点与曙光 AI 超集群对比

对比维度	华为昇腾 384 超节点	曙光 AI 超集群系统
核心特点	超大带宽、超低时延、超强性能	超高性能、超高效率、超高可靠、全面开放
路径模式	“类大型机” / “类苹果” 的封闭垂直优化, 单一巨头企业主导	“类 Cluster” / “类安卓” 的开放跨层优化, 大量厂商分工协作
加速卡类型	支持华为昇腾 910C 加速卡	支持市面上多品牌 AI 加速卡
算力性能	单机柜 32 卡, 单机柜算力约 25PF, 384 超节点共 300PF (12 个计算柜)	单机柜 96 卡, 单机柜算力大于 100PF
网络性能	超节点内任意两卡间通信带宽相较于	采用层次化高速互连网络, 单机柜访存总带宽大于



	传统架构提升 15 倍, 单跳通信时延降低 10 倍, 是业界唯一突破 Decode 时延 15ms 的方案	180TB/s, 单机柜片间互连总带宽大于 50TB/s
效率优化	强调算力互联效率, 在通信占比更高的 MoE 模型上达到 3 倍以上性能提升。	强调系统多维效率提升, 千卡集群性能提升 2.3 倍, GPU 算效提升 55%, 能源使用效率 PUE 小于 1.12, 模型开发部署效率提高 4 倍。
可扩展性	最大 16 万卡 Scale-out 集群	支持百万卡超大规模集群扩展

来源: 芯语, 国金证券研究所

三、四象二: 火箭瓶颈突破在即, 太空算力打开应用空间

3.1 海外巨头竞逐太空算力, 打开万亿美金市场

英伟达首次将芯片送入太空, 开启太空算力时代。根据机器之心报道, 2025 年 11 月 2 日, 英伟达首次把 H100 GPU 送入了太空。此次测试飞行搭载于初创公司 Starcloud 的 Starcloud-1 卫星上, 是该公司雄心勃勃的计划的第一步, 该计划旨在将全球耗能巨大的数据处理基础设施迁移到太空。这项为期三年的任务将由 SpaceX 的猎鹰 9 号火箭发射升空。仅重 60 公斤的 Starcloud-1 卫星将在距离地球约 350 公里的超低轨道上运行。在那里, 它将接收来自美国 Capella 公司运营的合成孔径雷达(SAR)地球观测卫星群的数据, 实时处理这些数据, 并将信息传回地球。Starcloud 甚至预测在未来十年内, 几乎所有新建数据中心都将建在太空, 这完全是因为地面能源的限制。一颗功率更大的 100 千瓦卫星预计将于 2027 年入轨。Starcloud 公司认为, 到 2030 年代初, 它将在太空中拥有一个 40 兆瓦的数据中心, 其数据处理成本与地球上的数据中心相当。

图表23: 美国初创科技公司 Starcloud



来源: 机器之心, 国金证券研究所

谷歌披露“Project Suncatcher”, 进军太空算力。根据华尔街见闻, 谷歌披露了名为“Project Suncatcher”的计划, 旨在构建一个由太阳能驱动的太空数据中心原型。该项目并不是建立



一个单一的轨道巨石，而是由 81 颗搭载 AI 芯片的卫星组成的集群，它们将在太空中协同飞行并处理数据。作为该计划的第一步，谷歌将与卫星公司 Planet 合作，预计于 2027 年向近地轨道发射两颗原型卫星。此举的核心逻辑在于利用太空独特的环境优势——特别是太阳同步轨道提供的近乎恒定的太阳能，以及免除在地面建设所需的土地和水资源。对于投资者而言，这释放了一个明确的信号：尽管面临物理世界的限制，科技行业仍试图证明 AI 具备无限的可扩展性。如果 Gemini 等 AI 模型的查询能够在太空中处理并将结果传回地球，将彻底改变算力基础设施的成本结构与能源依赖。

马斯克目标在 4-5 年将通过星舰完成每年 100GW 的数据中心部署。根据财联社报道，2025 年 11 月 4 日马斯克表示，将扩大星链 V3 卫星规模，建设太空数据中心，目标在 4-5 年将通过星舰完成每年 100GW 的数据中心部署。此外，根据 IT 之家报道，马斯克估计，AI 算力未来可能需要 200 至 300 吉瓦甚至接近 1 太瓦的持续功率输出。作为对比，一座典型的核电站仅能提供约 1 吉瓦的持续电力，而整个美国的持续总发电量也仅为 490 吉瓦左右。因此，马斯克断言，在地球上建设如此规模的电厂来支持 AI 是“不可能的”，唯一的出路在于太空。

图表24：马斯克认为运行人工智能成本最低的方式将在太空进行



The AI Investor
@The_AI_Investor

订阅



Elon and Jensen share the vision that in less than five years, the lowest-cost way to run AI compute will be in space.

With minimal cooling needs, no storage requirements, cheaper solar power, and massive scale, space makes AI compute far cheaper.

埃隆和詹森都认为，在不到五年的时间内，运行人工智能计算成本最低的方式将在太空进行。

由于冷却需求极低、无需存储、太阳能更便宜且规模巨大，太空使得人工智能计算成本大大降低。

翻译帖子



来自 Philip Johnston

上午2:23 · 2025年11月20日 · 1万 查看

IT之家
www.ithome.com

来源：IT之家，国金证券研究所

马斯克旗下 SpaceX 目标估值高达 1.5 万亿美元。根据界面新闻报道，埃隆·马斯克旗下 SpaceX 正加速推进首次公开募股（IPO）计划，目标估值高达 1.5 万亿美元，拟募资超 300



亿美元。若计划实现，SpaceX 将超越石油巨头沙特阿美此前 290 亿美元的募资规模，成为史上最大规模 IPO。

我们认为，太空算力有望成为全球算力中斜率最陡峭细分赛道。根据科创板日报报道，英伟达在 2026 财年第二季度财务业绩电话会议表示“仅今年一年，全球在数据中心基础设施与计算领域的投资规模预计就将达到 6000 亿美元，这一数字在两年内几乎实现翻倍。”我们认为，由于美国科技巨头纷纷发力太空算力，太空算力有望成为全球算力中斜率最陡峭细分赛道。

3.2 我国近期空天支持政策频出，行业发展有望加速

2014 年我国鼓励民间资本参与空间基础设施建设，拉开空天时代序幕。2014 年 11 月，国务院发布《国务院关于创新重点领域投融资机制鼓励社会投资的指导意见》，其中明确提出“鼓励民间资本参与国家民用空间基础设施建设。完善民用遥感卫星数据政策，加强政府采购服务，鼓励民间资本研制、发射和运营商业遥感卫星，提供市场化、专业化服务。引导民间资本参与卫星导航地面应用系统建设。”随后在 2015 年，发改委发布《关于印发国家民用空间基础设施中长期发展规划(2015-2025 年)的通知》，提出“探索国家民用空间基础设施市场化、商业化发展新机制，支持和引导社会资本参与国家民用空间基础设施建设和应用开发，积极开展区域、产业化、国际化及科技发展等多层面的遥感、通信、导航综合应用示范，加强跨领域资源共享与信息综合服务能力，加速与物联网、云计算、大数据及其他新技术、新应用的融合，促进卫星应用产业可持续发展，提升我国空间基础设施全面支撑经济社会发展的水平和能力。”自此，我国空天时代发展拉开序幕。

商业航天 2024-2025 年连续被写入政府工作报告，受重视程度持续提升。2024 年政府工作报告中提出“积极打造生物制造、商业航天、低空经济等新增长引擎。”2025 年政府工作报告中提出“新培育一批国家级先进制造业集群，商业航天、北斗应用、新型储能等新兴产业快速发展。开展新技术新产品新场景大规模应用示范行动，推动商业航天、低空经济、深海科技等新兴产业安全健康发展。”商业航天受重视程度持续提升。

“十五五”规划中同样重视空天行业发展。2025 年 10 月国务院发布《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》，其中提出“培育壮大新兴产业和未来产业。着力打造新兴支柱产业。实施产业创新工程，一体推进创新设施建设、技术研究开发、产品迭代升级，加快新能源、新材料、航空航天、低空经济等战略性新兴产业集群发展。完善产业生态，实施新技术新产品新场景大规模应用示范行动，加快新兴产业规模化发展。”

国家航天局印发《国家航天局推进商业航天高质量发展安全发展行动计划（2025-2027 年）》，指引行业发展路径。2025 年 11 月新华社报道，国家航天局印发《国家航天局推进商业航天高质量发展安全发展行动计划（2025-2027 年）》，明确五方面 22 项重点举措，推进商业航天高质量发展和高水平安全。其中明确，“鼓励商业航天布局航天产业链相应环节”“重点支持商业航天主体开发新技术、新产品，挖掘应用新场景”“完善商业航天发展投融资体制机制，设立国家商业航天发展基金”等。在推动产业发展壮大方面，《行动计划》明确，推进产业结构优化，鼓励商业航天布局航天产业链相应环节，做强火箭制造，做大应用服务，拓展测控运营能力，推动商业航天产业高质量发展。支持拓展新业态。支持商业航天主体围绕太空资源开发利用、太空制造、在轨维护与服务、太空环境监测探测、空间碎片监测预警与减缓清除、太空旅游、太空生物制药等新领域，加强原始创新和关键核心技术攻关、系统开发和应用服务。

商业航天司成立，商业航天迎来专职监管机构。根据新华社 2025 年 11 月报道，国家航天局已于近期设立商业航天司，相关业务正在逐步开展，标志着我国商业航天产业迎来专职监管机构，未来将持续推动我国商业航天高质量发展，产业链有望全线受益。近年来，我国商业航天在政策牵引、技术突破与市场驱动下，以全产业链协同创新，完成了历史性的跨越。国家航天局相关负责人介绍，当前我国商业航天企业数量超 600 家，在确保安全的前提下逐步释放商业航天发展潜力。

商业航天有望获得更多资金支持，助力头部企业积极技术突破。根据财联社报道，2025 年 11 月 26 日上午，国家发改委、财政部等多部门联合发布国家创投基金启动相关政策；傍晚，商业火箭企业适用科创板第五套上市标准的专项细则正式落地，两项重磅政策将为商业航天产业发展注入强劲动力。1) 国家创投引导基金，在实施区域上，重点辐射京津冀、长三角、粤港澳大湾区三大经济核心带；在产业方向上，围绕集成电路、人工智能、生物制造、未来能源等前沿领域布局，大力支持发展我们国家的塔尖产业，商业航天也被明确纳入支持清单，凸显其国家硬科技产业布局中的核心地位。2) 科创板专项细则的出台，则是为商业航天企业登陆资本市场开辟了“绿色通道”。这是科创板第五套上市标准重启后，首次针对单一行业发布专项实施细则。此前，该标准已明确将人工智能、低空经济、商业航天等领域纳入支持范畴，而此次细则进一步聚焦商业火箭赛道，对企业的硬科技属性、技术优势、科研成果转化、资质认定及商业空间等核心指标作出明确界定。



3.3 我国积极布局大型星座与太空算力

我国规划多个万星以上星座，积极抢占稀缺轨道资源。根据东方中科微信公众号：1) 2021年4月28日中国星网成立，标志国有资源与民营资本共同推进和加速我国空间互联网建设，主导建设巨型低轨卫星工程“GW”星座计划，整合鸿雁星座、虹云星座等原有卫星计划，实现卫星网络全面布局。GW星座包含两个子星座GW-A59（6080颗）和GW-2（6912颗）。2) 我国第二大巨型低轨商业卫星星座—G60星座（又名：千帆星座）部署计划为在2027年前1296颗卫星提供全球网络覆盖，到2030年完成超1万颗卫星组网，为全球用户提供低延时、高速率、高可靠性的多网融合卫星互联网服务。3) 2025年5月24日ITU收到上海蓝箭鸿擎科技有限公司（简称：鸿擎科技）提交的申请文件，内容“鸿鹄三号”星座是继GW星座和G60星座计划之后，我国第三个超万颗卫星的巨型低轨卫星星座计划。据ITU规定，卫星轨道资源和频率资源采取“先申请先获得”原则。为真正拥有频段使用权，获得批准实体必须在限期七年内完成卫星发射和信号验证。如果超期，获得资源作废。

2025年以来我国低轨卫星组网速度明显提速，后续有望进一步加快组网节奏。千帆星座首次组网于2024年8月开始，GW星座于2024年底开始组网，通过统计发现2025年我国两大星座组网显著加速，截至2025年12月26日，GW星座2025年合计成功发射16组低轨卫星。我们认为，由于轨道资源稀缺，后续我国卫星组网速度有望进一步加速。

图表25：2025年以来我国低轨卫星互联网组网明显提速

事件	时间
千帆极轨01组卫星成功发射	2024/8/6
千帆极轨02组卫星成功发射	2024/10/15
千帆极轨03组卫星成功发射	2024/12/5
卫星互联网低轨01组卫星成功发射	2024/12/16
千帆极轨04组卫星成功发射	2025/1/23
卫星互联网低轨02组卫星成功发射	2025/2/11
“千帆星座”第五批卫星成功发射	2025/3/12
卫星互联网低轨03组卫星成功发射	2025/4/29
卫星互联网低轨04组卫星成功发射	2025/6/6
卫星互联网低轨05组卫星成功发射	2025/7/27
卫星互联网低轨06组卫星成功发射	2025/7/30
卫星互联网低轨07组卫星成功发射	2025/8/4
卫星互联网低轨08组卫星成功发射	2025/8/13
卫星互联网低轨09组卫星成功发射	2025/8/17
卫星互联网低轨10组卫星成功发射	2025/8/26
卫星互联网低轨11组卫星成功发射	2025/9/27
卫星互联网低轨12组卫星成功发射	2025/10/16
“千帆星座”第六批卫星成功发射	2025/10/17
卫星互联网低轨13组卫星成功发射	2025/11/10
卫星互联网低轨14组卫星成功发射	2025/12/6
卫星互联网低轨15组卫星成功发射	2025/12/9
卫星互联网低轨16组卫星成功发射	2025/12/12
卫星互联网低轨17组卫星成功发射	2025/12/26

来源：无锡博报，央广网，上海市国资委，新华网，工信部官网，新浪财经，国资委官网等，国金证券研究所整理

太空算力方面我国同样积极布局。1) 2025年5月14日，我国首个整轨互联太空计算卫星星座正式进入组网阶段，这也是中国“三体计算星座”的首次发射。“三体计算星座”是千星规模的太空计算基础设施，建成后总算力可达1000POPS（每秒百亿亿次运算）。其中，此次



发射的计算卫星核心载荷——星载智能计算机把卫星算力从T级提升到P级,实现了10—100倍的提升;星载高通量路由器作为星座天地网络互联的核心设备,将完成星间、星地和星内百G比特高速率、低延迟网络传输。2)根据IT之家报道,北京市支持研发的首个算力星座将发射首颗算力试验卫星——“辰光一号”已完成产品研制,正在开展总装试验。预计5年左右,首座太空算力中心将建成。根据规划,太空数据中心将部署在距地面700至800公里的晨昏轨道,建设过程分为“天数天算”“地数天算”“天基主算”三个阶段。项目采用模块化、标准化设计,通过在轨交会对接搭建平方公里级大型航天器,最终具备超大模型训练与推理能力。算力提升方面,规划明确:未来3年内,星座太空算力将达到1000P(每秒千万亿次浮点运算)量级;到2030年,算力有望提升至40万P,相当于当前我国所有地面数据中心的算力总和,首座太空数据中心将正式建成。届时,该中心将在6G通信、自动驾驶、气象预报等多个关键领域发挥核心支撑作用,为相关产业的技术突破与规模化应用注入强劲动力。

图表26: 北京太空数据中心建设工作推进大会进行



来源: 北京发布公众号, 国金证券研究所

3.4 我国火箭发射瓶颈有望逐渐突破

我国当前发射节奏较慢,核心矛盾在于运力不足。根据央广网报道,目前我国低轨卫星星座的发射节奏尚不能称为大规模组网,且面临“占频保轨”的压力——国际电信联盟规定,运营商须在申报7年内发射首颗卫星,9年内发射星座总数量的10%,12年内完成50%,14年内必须100%部署完毕。当前的核心矛盾在于卫星亟待发射,而火箭运力却严重不足。承担两大星座卫星发射任务的火箭包括长征十二号、长征八号甲、长征六号改等“国家队”火箭,但这些火箭还需兼顾国家其他航天任务,排期紧张;与此同时,商业航天公司尚无成熟的大运力火箭可用,导致整体发射进度不及预期。

可回收技术突破可大幅降低火箭发射成本。根据36氪报道,2015年,猎鹰9号火箭一级首次成功着陆回收,2016年首次实现海上平台的着陆回收,2017年实现火箭重复使用,2019年将首批60颗星链卫星发射入轨。根据经济观察网报道,美国SpaceX的“猎鹰九号”近地轨道运力约22.8吨,2024年单箭发射报价为6975万美元,折合每公斤约3000美元(约合人民币2.1万元/公斤),而国内商业火箭的发射报价多在每公斤6万元到15万元之间。猎鹰九号能够实现第一级火箭和整流罩的回收再复用:一子级约占火箭成本的67%,整流罩成本占比约10%。SpaceX已成功实现单枚火箭第一级重复使用22次,总回收复用次数超300次,火箭成本因此得以大幅摊薄。Starcloud公司联合创始人、CEO Philip Johnston表示,“我们预计每公斤发射成本约为500美元,达到盈亏平衡点。而使用星舰后,预计发射成本会更低”。



一旦星舰全面投入运营，其每公斤发射价格预计将在 10 美元到 150 美元之间。

我国可回收火箭技术突破正在临点，“国家队”与民企积极良性竞争。1) 2025 年 12 月 3 日，朱雀三号遥一运载火箭在东风商业航天创新试验区发射升空，按程序完成了飞行任务，火箭二级进入预定轨道。与此同时，本次任务开展了一子级垂直回收技术的飞行验证。根据飞行测量数据，火箭一子级在着陆段点火后出现异常，未实现在回收场坪的软着陆，残骸着陆于回收场坪边缘。2) 12 月 23 日，长征十二号甲遥一运载火箭在东风商业航天创新试验区发射升空，运载火箭二子级进入预定轨道，一子级未能成功回收，飞行试验任务获得基本成功。本次任务虽未实现预定的火箭一级回收目标，但是获取了火箭真实飞行状态下的关键工程数据，为后续发射、子级可靠回收奠定了重要基础。根据贝壳财经，首次发射就实现回收没有先例。猎鹰 9 号从首发到成功回收历经 5 年，经历了 20 余次试错爆炸，新格伦火箭历经发动机爆炸、储箱爆裂等多重挫折才实现入轨。我们认为，在当前政策支持下叠加国营企业与民企积极竞争，我国火箭发射瓶颈将得到快速突破。

图表27：朱雀三号首飞的技术突破

- | | |
|-------------|--|
| NO.1 | 在国内首次实现了全新总体布局的重复使用液氧甲烷运载火箭入轨飞行。突破动力、气动、飞控、结构等多专业耦合设计关键技术，对一子级创新采用燃箱在上/氧箱在下的总体布局和面对称边条翼、P型局部后掠栅格舵及流线型着陆腿外罩为典型特征的气动布局，可有效降低返回过程推进剂消耗量，提升重复使用模式下的运载能力。 |
| NO.2 | 在国内首次实现九机并联液氧甲烷动力系统的集成应用，突破了大流量推进剂稳定可靠输送技术，满足了全任务复杂剖面下发动机多次启动、深度变推需求，实现了滑行段失重状态大姿态机动下，低温推进剂有效管理和可靠沉底；解决了全过冷加注甲烷易冻结的难题，实现了射前2小时大流量全过冷加注。 |
| NO.3 | 在国内首创全新的高强度不锈钢/高性能激光焊接贮箱制造材料和工艺体系，自主开发全套不锈钢贮箱激光焊接工艺装备和产线，实现大直径超薄壁不锈钢贮箱制造成本较铝合金大幅降低80%，生产周期缩短40%。 |
| NO.4 | 在国内首次实现入轨级重复使用运载火箭高精度返回导航、制导与控制技术的飞行验证。突破复杂约束条件、大范围状态变化、强不确定环境下的在线轨迹优化、姿态稳定和高精度控制技术。 |
| NO.5 | 在国内首次实现重复使用运载火箭混合冗余分布式综合电子集成设计与应用，构建面向服务的集控制、测量、健康管理于一体的开放式系统。基于高速Powerlink实时以太网构建的高带宽、高可靠性实时通信平台，通信速率超100Mbps，较传统火箭1553B提升百倍。采用汽车电子先进的高可靠计算平台研制高算力返回控制计算机，较传统运载火箭算力提升10倍以上，打造火箭航电系统与高端先进制造业相融合的新质生产力。 |

来源：界面新闻，国金证券研究所



图表28：长征十二号甲遥一运载火箭飞行试验任务



来源：第一财经，国金证券研究所

我们认为，当前空天行业正处于从 1~10 的快速发展阶段，可回收火箭技术作为行业瓶颈有望得到快速突破，全球加速竞争太空算力为空天行业打开广阔应用市场，2026 年空天行业有望持续表现亮眼。

图表29：空天时代临点将至



来源：国金证券研究所绘制

四、四象三：具身智能有望变得精准与有用，实现放量的基础

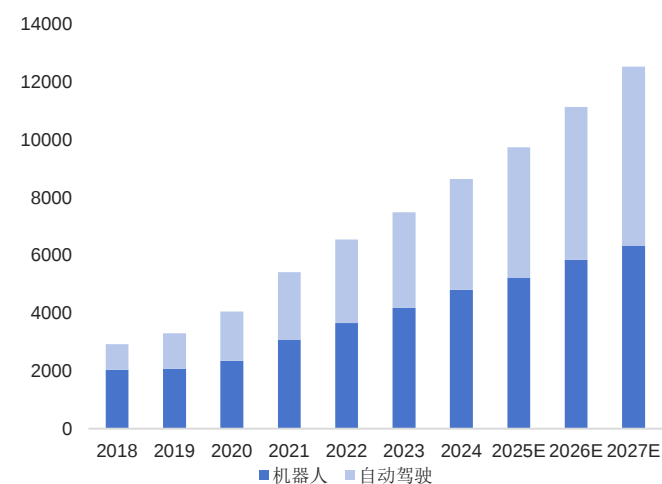
“十五五规划”提出“将具身智能打造成新的经济增长点”。政策不断提升“具身智能”战



略定位，2023 年《“十四五”机器人产业发展规划》首次提及具身智能，2025 年首次写入《政府工作报告》，二十届四中全会“十五五”规划中明确提出“将具身智能打造成新的经济增长点”。各地政府积极响应提出明确量化目标，加速产业落地：

- **北京市：**提出到 2027 年，突破百余项关键技术，产出不少于 10 项国际领先的软硬件产品；实现不少于 100 项规模化应用，量产总规模率先突破万台，培育千亿级产业集群。
- **上海市：**提出到 2027 年，实现具身模型、具身语料等方面核心算法与技术突破不少于 20 项，我市具身智能核心产业规模突破 500 亿元。
- **深圳市：**提出到 2027 年新增培育估值过百亿企业 10 家以上、营收超十亿企业 20 家以上，实现十亿级应用场景落地 50 个以上，关联产业规模达到 1000 亿元以上。

图表30：2027 年中国具身智能市场规模有望突破 1.25 万亿元



来源：亿欧智库，国金证券研究所

图表31：2025 年具身智能投资分布中人形机器人为第一大细分赛道

投资领域	金额占比	代表案例	政策杠杆
人形机器人	35%	智元机器人 B 轮融资超 50 亿元	北京千亿政府基金定向支持
核心零部件	28%	傲意灵巧手产能扩至 500 台/月	深圳百亿产业基金优先采购
具身大模型	22%	光轮智能合成数据平台估值翻倍	国家 AI 重大专项倾斜
应用生态	15%	太希外骨骼景区租赁模式落地	场景开放补贴制度

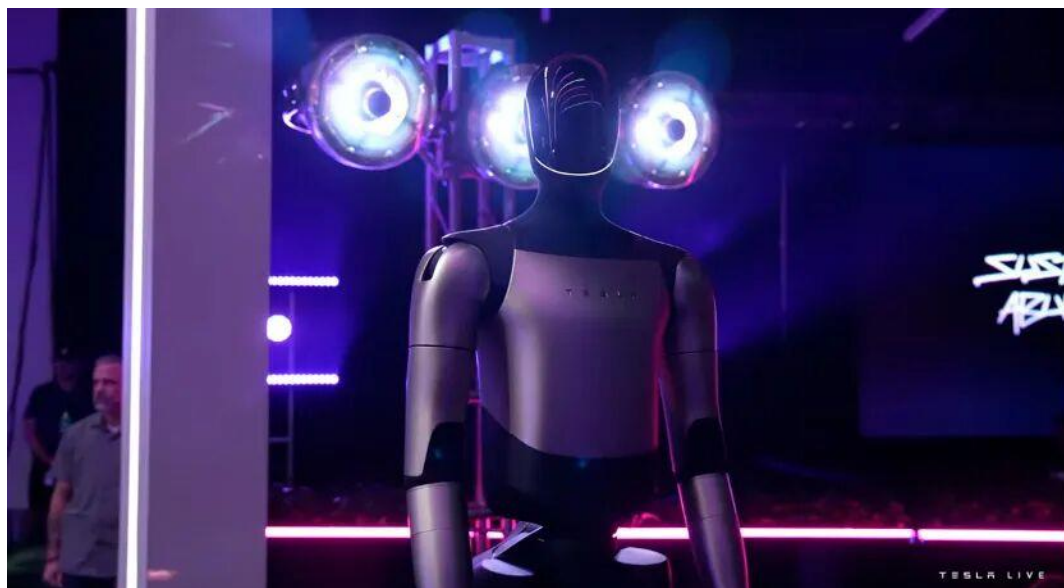
来源：中投产业研究院，国金证券研究所

特斯拉 Optimus V3 临近发布，预计为机器人小脑发展集大成者，各类任务精准度快速提升。马斯克明确“2026 年启动 Optimus V3 量产”，因此量产前 V3 的发布将会更近，目前根据马斯克公开表述，“Optimus V3 将在手部灵活性和 AI 大脑层面拥有巨大提升”，同时得益于多自由度运动系统和自适应学习能力，有望成为机器人小脑发展集大成者：

- 1) 手部灵活性上，**马斯克称“人类的手时极其精妙的进化成果。驱动人手的大部分肌肉其实在小臂上，人手大约有 27、28 个自由度。要让机器人成为真正的通用人形机器人，必须解决‘手’的问题”，“造机器人的根本难点在于硬件设计，手和前臂的工程难度最大。如果硬件问题解决了，软件方面可以直接利用大模型进展对机器人下达指令”。
- 2) 在 AI 大脑上，**马斯克确认“V3 已经启用了 Grok 语音助手”，Grok 与 Optimus 的集成或许能让机器人与周围环境和用户更自然地互动，能让 Optimus 对普通消费者更加友好。
- 3) 多自由度运动系统，**全身自由度达 200 以上，手部拥有 11-27 个自由度，可实现精准抓握与精细操作，如抓取 20 磅重物或完成叠衣服等家务、感知与算力系统。
- 4) 自适应学习能力，**比如通过端到端神经网络，仅用 2D 摄像头、手部触觉及压力传感器数据即可生成关节控制序列，能在分拣电池等任务中自我纠正失误。



图表32: Optimus V2.5 无需连接电线，可自主完成行走



来源：中国机器人网，国金证券研究所

模型大脑侧，Gemini 3.0Pro 实现多模态理解能力跃升，未来机器人结合场景泛化任务从不可能变得可能。Gemini3 拥有极强的多模态理解能力，在 ScreenShot-Pro 评测基准中 Gemini 3 Pro（得分 72.7%）大幅领先，是 Claude Sonnet 4.5 的两倍，GPT 5.1 的二十倍。空间推理能力的提升，使模型能够更好地服务于自动驾驶、机器人、XR 设备和智能终端，加速具身智能落地。Gemini 3 Pro 能预测运动轨迹、理解物体关系并判断任务进度，为下一代自动化系统奠定基础，也是通向实体机器人的关键一步。其本质在于“感知—决策—行动”的统一：模型学会在屏幕上识别按钮并操作，与机器人识别现实中的物体并完成抓取，在底层逻辑上是相通的，这一能力也可自然扩展到对真实设备和复杂操作界面的理解与控制。我们认为未来机器人能理解多种人类指令，结合场景完成泛化任务从不可能变得可能，这是具身智能真正放量的基础。

图表33: ScreenShot-Pro 评测的部分数据集，包含各类桌面应用和场景



来源：《ScreenSpot-Pro: GUI Grounding for Professional High-Resolution Computer Use》，国金证券研究所

图表34: Nano Banana Pro 图像生成：创建一个信息图，展示如何制作香茅茶



来源：谷歌官网，国金证券研究所

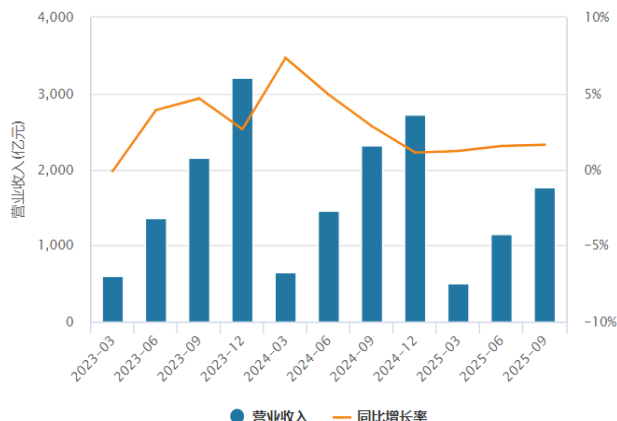
从产业信号看，特斯拉人形机器人有望迎量产元年，成为一切的起点。特斯拉 2025 年股东大会上，马斯克明确“2026 年启动 Optimus V3 量产，2027 年推出 V4，2028 年迭代至 V5；远期规划年产能 1000 万台，最终成本降至 2 万美元”，“公司计划首先在加州弗里蒙特工厂建立一条年产 100 万台 Optimus 的生产线，并在得克萨斯州超级工厂建设规模更大的产线，年产能预计达数千万台。”我们认为 2026 年随着 Optimus 量产，叠加大模型多模态理解能力的跃升，具身智能将迎来真正应用的元年，在部分简单场景真正发挥价值，成为一切的起点。



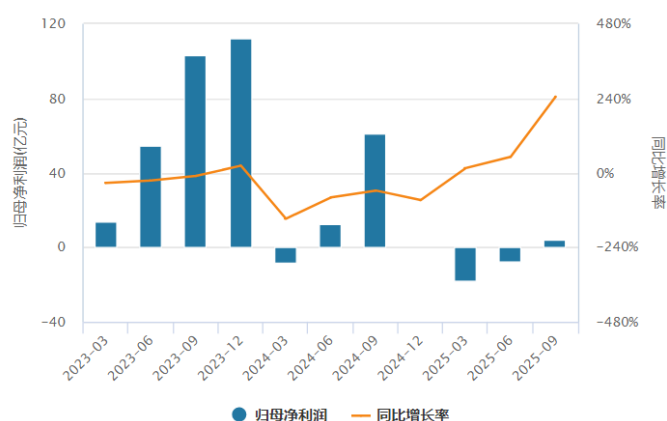
五、四象四：AI 应用破土而生，大模型/AI Infra/高增长/高壁垒垂类软件值得关注

计算机行业基本面已于 2025H2 见底回升，具备反弹条件。根据 Wind 软件指数(882250.WI)，可以看到，软件行业营收已于 2025H2 见底回升。其中 2025Q3，行业营业收入 1765.19 亿元，同比增长 1.61%，收入端全年维稳、缓慢修复；归母净利润 3.77 亿元，同比增长 244.56%，预计主要受益于各类降本增效措施。

图表35：软件行业营收及增长率



图表36：软件行业归母净利润及增长率



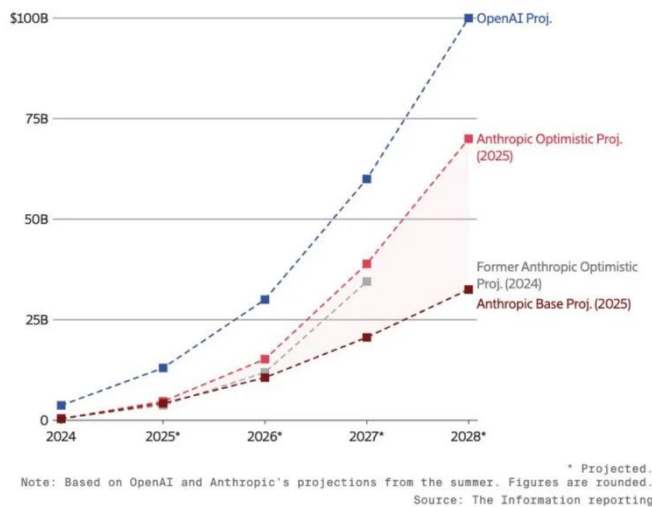
来源：Wind，国金证券研究所

来源：Wind，国金证券研究所

5.1 大模型：已广泛应用，向个性化与生态化深入发展

大模型本身即为最具代表性的 AI 应用，全球领军模型 ARR 已达百亿美金。1) 无论是 OpenAI 与 Gemini，亦或是 Deepseek 与 Qwen 等，大模型本身即为最具代表性的 AI 应用。2) 自 2022 年底以来，经过约 3 年发展，OpenAI ARR 预计 2025 年底将超过 200 亿美元，并计划到 2030 年实现“数千亿美元”级别的收入规模；Google Gemini 虽未直接披露 ARR，但其 token 使用量呈指数级别提升，2025 年 10 月已达到 1.3 千万亿个/月；Anthropic Claude ARR 在 2025 年 7 月已达到 50 亿美元，过去 7 个月翻了 5 倍。3) 国内来看，大模型使用量同样快速增长，以豆包为例，2025 年 12 月其日均 Tokens 使用量已突破 50 万亿。

图表37：OpenAI 与 Anthropic 收入曲线



来源：智东西，The Information，国金证券研究所

在使用大模型过程中，用户可将其与个人知识库及工作流深度结合。1) 大模型使用场景远不止对话及搜索，用户可利用大模型为自己设计和搭建的自动化、智能化工作流程，旨在提高效率、简化任务或实现特定目标。2) 根据日常的工作生活，工作流可包括信息收集与整理、内容创作与辅助、任务管理与提醒、财务规划与分析、健康管理、旅行规划等。相比通用 AI，自定义 Agent 可与使用者个人的知识库及工作流深度结合，通过多次使用与反馈，Agent 的个



性化及效益水平可持续提升。

图表38：个人大模型使用场景及 workflow

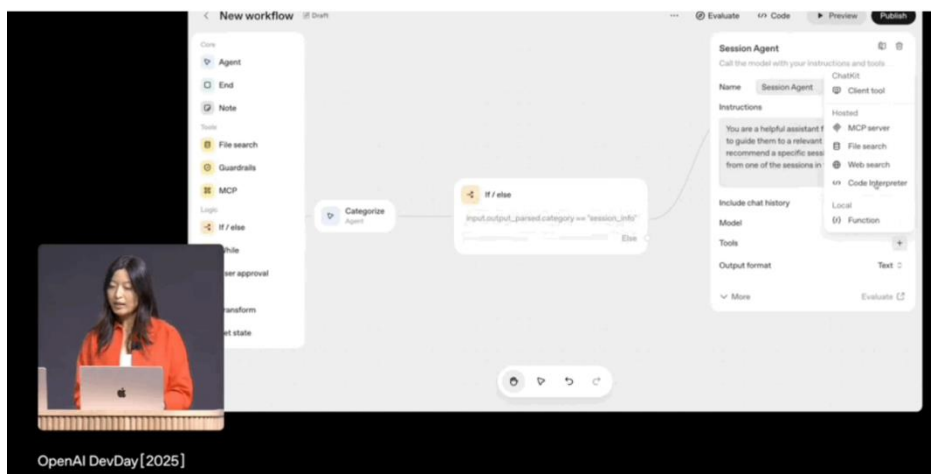
场景	内容	流程	调用工具
信息收集与整理	自动收集新闻、行业动态、学术论文等信息，并进行分类整理	Agent 通过网络搜索工具获取指定关键词的信息 → 对信息进行筛选和去重 → 按主题或日期分类存储 → 生成摘要或索引	可结合搜索引擎 API、RSS 订阅工具和文本处理模型实现
内容创作辅助	帮助撰写文章、文案、报告等，提供创意灵感、语法检查和内容优化	用户输入主题或关键词 → Agent 生成多个创意方向 → 选择方向后生成初稿 → 进行语法和逻辑检查 → 提供修改建议	大型语言模型（如 ChatGPT、豆包等）搭配文本编辑工具
任务管理与提醒	自动规划每日任务、设置提醒，并跟踪进度	用户输入任务清单 → Agent 根据优先级和时间安排任务 → 设置定时提醒 → 任务完成后记录结果 → 提供总结报告	结合日历应用、提醒工具和简单的逻辑判断模块
财务规划与分析	自动整理收支记录、生成财务报表，并提供理财建议	连接银行账户或导入财务数据 → Agent 分类统计收入和支出 → 生成月度/年度报表 → 分析消费趋势 → 提供节省开支或投资建议	财务数据接口、数据分析模型和可视化工具
健康管理	跟踪运动数据、饮食记录，并提供健康建议	连接运动设备或手动输入数据 → Agent 分析运动量和饮食结构 → 提供个性化运动计划和饮食建议 → 设置健康目标并跟踪进度	健康监测设备 API、营养数据库和健康模型
旅行规划	自动搜索目的地信息、预订酒店和机票，并生成旅行攻略	用户输入旅行时间和地点 → Agent 搜索景点、酒店和航班信息 → 根据预算和偏好筛选选项 → 生成行程安排 → 提供当地美食和文化推荐	旅行预订平台 API、地图服务和知识图谱

来源：国金证券研究所归纳整理

头部基模厂商已发布各类开发工具，助力基模更方便、更深入地部署。1) OpenAI: 2025 年 10 月，OpenAI 开发者日 2025 召开，发布 AgentKit 工具包，包含诸多模块，可以帮助开发者更高效设计工作流。例如通过 Agent Builder 模块，开发者无需从代码开始，只需添加并设置一系列节点即可用可视化直观的方式快速构建 Agent。OpenAI 研究员现场演示，在 8 分钟内用 AgentKit 为其开发者日官方网站构建 Agent。2) 腾讯: 微信最终会推出一个 AI 智能体，让用户在微信内部就可利用 AI 完成多项任务，能够理解用户的需求、意图和兴趣。他表示，微信拥有强大的通信和社交生态系统，还有购物支付等场景，几乎是用户的理想助手。但刘炽平也认为，目前这项愿景还处于“非常早期的阶段”；目前，微信正在更多地引入元宝能力，包括用 AI 来增强搜索等。同时，腾讯也在开发垂直领域的智能体能力，这些能力未来会逐步开放、整合，从而打造出微信智能体的理想蓝图。



图表39: OpenAI AgentKit 使用案例



来源：量子位，国金证券研究所

5.2 AI Infra: 获取卖铲子的确定收益，直接奠定应用成色

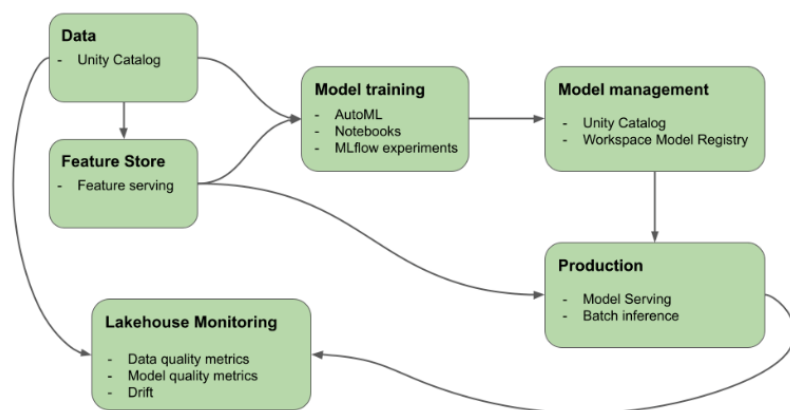
AI Infra 不直接生产模型和应用，但提供了基础的“水电煤”，直接奠定 AI 的“成本”与“上限”。1)算力调度: 与指挥交通类似,将庞大的计算任务分配给成千上万个处理器(GPU/TPU),并处理节点故障,确保训练不中断。2)数据吞吐: 在训练中,海量数据需要尽可能快地“喂”给 GPU,不能让 GPU 等数据(IO Wait),需要高性能的存储和网络;在推理中,高并发数据涌入,需要尽可能提高 CPU/GPU 计算效率,降低用户延迟。3)模型训练与推理支持: 提供环境及工具,让模型学得快,并在用户使用回得快。4)资源抽象与虚拟化: 开发者无需关心底层是何种 GPU,可直接通过软件层调用算力。

Databricks: “湖仓一体”全球大数据领军, AI 收入已达 10 亿美元。

- Databricks 成立于 2013 年,主要提供统一的数据与 AI 平台服务,帮助企业整合和处理大规模数据,用于数据工程、数据科学、机器学习与 AI 应用,也能为电商、金融、医疗等领域的企业提供数据服务。公司为“湖仓一体”数据库架构开创者,是数据智能领域代表性企业,也是 AI 时代重要的数据基础设施提供商。根据智东西报道,截至 2025 年 12 月,有超过 60% 的财富 500 强企业采用了 Databricks 的数据智能平台来管理数据,并将其与 AI 结合。
- 随着 AI 技术的全面渗透,“湖仓一体”模式正成为数据基础设施的核心选型。面对 AI 全生命周期中对多模态庞杂数据的需求,该架构不仅实现了统一纳管和实时处理,更利用云原生的弹性扩缩容机制达到了降本增效的目的。值得注意的是,它还针对 AI 场景进行了底层优化,例如内建向量搜索功能并对接 PyTorch 等开发生态,从而精简了数据预处理到模型应用的复杂流程。
- 12 月 16 日, Databricks 宣布完成超过 40 亿美元融资,投后估值 1340 亿美元。经营数据方面,公司披露其第三季度对应的年化营收规模超过 48 亿美元,同比增长超过 55%;同时称其 AI 相关产品与数据仓库业务的年化收入规模均超过 10 亿美元,并在过去 12 个月实现自由现金流为正。



图表40: Databricks AI 平台

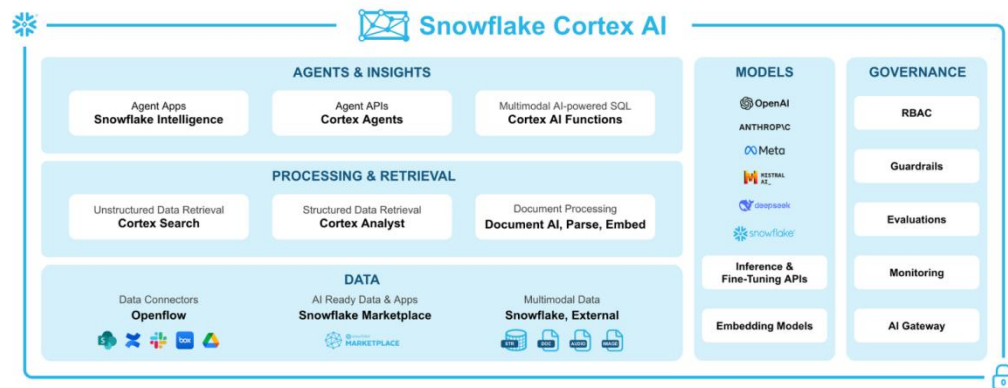


来源：微致资本，国金证券研究所

Snowflake: 全能型数据云平台，Cortex 有望成为企业 AI 开发支柱。

- Snowflake 成立于 13 年前，基于公有云构建了一个完整的数据和 AI 平台。其致力于帮助客户处理各种不同类型的数据，支持多种语言进行开发和数据建模，并为商业用户提供数据工程、数据分析、AI 以及数据共享方面的应用。截至 2025 年 11 月，Snowflake 在全球拥有超过 12,000 个企业级客户，其中超过 50% 的客户使用了其 AI 功能和产品。
- Snowpark: 允许开发者不用 SQL，而是用 Python、Java 等语言直接在 Snowflake 内部处理数据。此为机器学习的关键。
- Snowflake Cortex: 原生 AI 和机器学习套件，直接构建于 Snowflake 平台之中。它使用户能够利用 Snowflake 数据运行高级 AI 模型，而无需迁移数据或配置单独的基础设施。
- FY2026Q3, Snowflake 实现营业总收入 12.13 亿美元，净亏损 2.94 亿美元。其中，AI 业务年化营收超过 1 亿美元，比原计划提前约一个季度，主要得益于创新速度、跨部门的高效协作以及众多标杆客户的早期采用。

图表41: Snowflake Cortex AI



来源：Snowflake，国金证券研究所

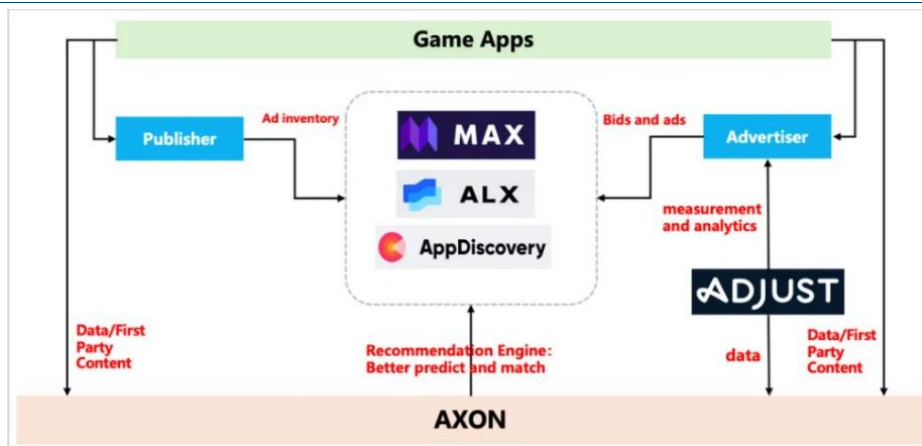
5.3 高增长场景应用：广告营销、动态漫、医疗等有望跑通飞轮

我们认为，大模型相比上代 AI，具备更强大的语义理解、数据分析及多模态处理能力。在部分场景中，已展现出更优秀的落地潜力，高增长态势已较为显著。

广告营销已成为 AIGC 商业化落地最具代表性的场景之一。1) 得益于该场景高频、海量的业务特性，基于 AIGC 的“数据飞轮”效应已成功跑通，实现了算法优化与投放效果的正向循环。不同于单纯的降本，AIGC 在此场景下能通过精准化投放与自动化素材生成，直接为广告主带来可量化的收入增量与 ROI（投资回报率）提升，有效激发了下游客户的付费意愿，推动该场景率先实现从技术验证向规模化应用的跨越。2) 以 Applovin 为例，其收入利润于近年高速增长。其业务包括 App 游戏、广告平台以及底层算法基建等三大块，而这三大块是能够通过互相推动来旋转的业务飞轮。

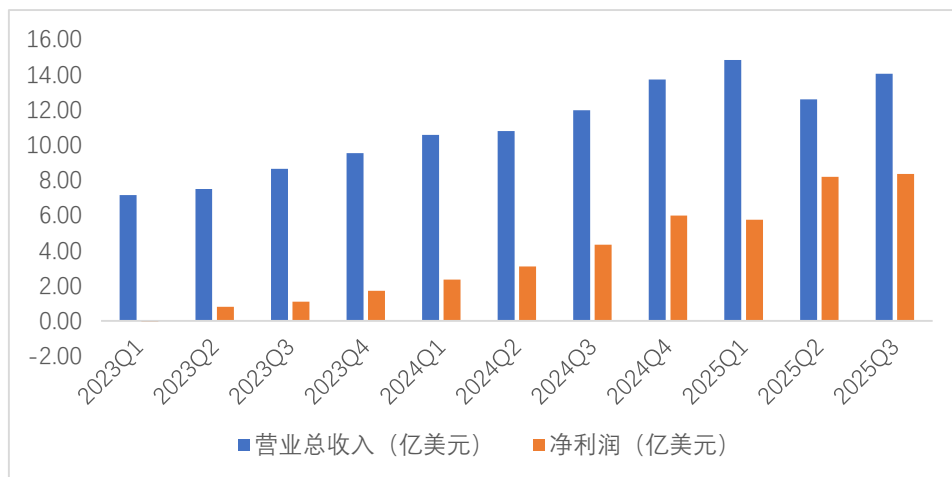


图表42: Applovin 业务模式



来源：华尔街见闻，国金证券研究所

图表43: Applovin 营收及利润



来源：Wind，国金证券研究所

多模态技术红利释放，动态漫已预定“爆款”。1) 随着 2025 年多模态大模型技术路径的收敛与突破，动态漫的视觉表现力实现了质的飞跃。在需求侧，其叙事节奏紧凑、情绪密度高，精准契合了当下用户在碎片化时间内的娱乐偏好；在供给侧，相较于传统动画或真人短剧，AI 赋能下的动态漫制作周期大幅缩短，具备低成本、批量化产出的显著优势。技术跃升叠加供需适配，正推动该赛道加速从边缘走向主流。2) 根据巨量引擎与火山引擎统计，2025 上半年，漫剧市场商业化高速增长。抖音端漫剧付费自然流：从 2024 上半年的 0，一路跑到 2025 年二季度的 300+万/天、2025 年 8 月的 1000+万/天；抖音端漫剧商业化投流：从 2024 下半年的 1 万/天，逐步上涨至春节的 20 万/天、2025 年二季度的 80 万/天，并在 2025 年 8 月破峰，达到 400 万/天。



图表44：抖音漫剧市场商业化指标



来源：短剧自习室，国金证券研究所

医疗 AI 能力全面跃迁，有望成为下一个高速增长的场景。1) 该场景拥有高质量的私有数据沉淀与稀缺的医学专家资源，为模型训练提供了坚实基础。通过将专家标注（RLHF）与既有的场景卡位优势相结合，AIGC 能够有效整合碎片化信息，实现从“数据要素”到“临床决策辅助”的完整业务闭环。这种深度融合不仅解决了医疗领域对专业性与准确性的极高要求，也助力拥有核心数据资产的厂商构建起难以复制的竞争护城河。2) 从案例来看，截至 2025 年 12 月 16 日，蚂蚁阿福下载量即飙升至苹果应用榜第三、月活用户规模突破 1500 万；美年健康或可凭借近 3 亿人次的医疗级体检数据积累、覆盖全国的近 600 家线下服务网络以及成熟的 AI 产品矩阵，成为核心受益方。

5.4 高壁垒垂类应用：具备数据与 workflow 护城河，对客价值更为显著

大模型的发展将加速“简单应用”的商品化与淘汰；而对于具备深厚行业 Know-how、专有数据资产、复杂流程整合能力及严格合规资质的应用而言，大模型非但不是威胁，反而将成为巩固其核心优势的关键工具。

“简单”应用一般具备如下特征：1) 依赖公开知识：核心输入输出基于互联网通用语料及共识信息；2) 验证门槛低：结果的优劣对错可被快速判定（如代码单元测试）；3) 责任风险弱：容错率高，回滚成本低（如图像生成、文案编辑）；4) 系统耦合度低：任务单一，不涉及跨系统协作或复杂的权限与审计流程。

而深层垂类应用，具备四类大模型难以攻克的壁垒：

- **深层行业 Knowhow：**特定行业依赖隐性经验、场景化判断及不成文规范。通用大模型难以低成本覆盖此类长尾且深度的领域知识。因此，能够将垂直领域业务逻辑转化为专用模型或 Agent 的企业，将凭借比通用方案更高的精准度与可靠性，构建起显著的竞争壁垒。
- **专有数据资产：**随着基础模型能力的普惠化，竞争重心将通过模型层下沉至数据层。拥有独特、高质量且难以通过公开渠道获取的闭环数据，才是核心差异化优势。在此场景下，大模型仅作为数据价值的放大器，数据所有者将占据价值链的主导地位。
- **复杂流程整合：**现代企业软件的价值在于对跨部门、多环节业务流程的深度编排（如集成财务、HR 及审批流的 ERP 系统）。这种深度嵌入工作流的特性带来了极高的客户粘性与替换成本。大模型难以独立完成需强耦合与可追责的复杂执行，更倾向于被集成至现有系统中以提升效率。
- **资质与合规壁垒：**在强监管领域，业务开展依赖明确的责任主体、执照资质及合规体系。虽然模型可辅助生产，但最终的法律责任与结果背书必须由具备资质的实体承担。这一刚性约束决定了此类企业在产业链中不可替代的地位及更高的议价能力。

以上因素同样决定了在与基模合作中，应用公司能占据更大价值。

Palantir：军工业务为基，展现复杂场景下的 Alpha

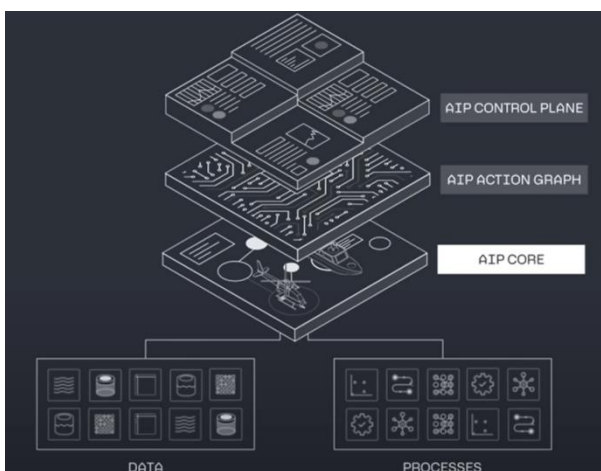
- Palantir 成立 2003 年，总部位于美国科罗拉多州丹佛，实际运作始于 2004 年，专注于提供大数据整合、可视化和分析平台，帮助组织从海量、复杂的数据中获取有价值的洞察。基于超强的数据挖掘分析能力，Palantir 成为美国军政机构的大数据软件核心供应商，并凭借发现麦道夫“庞氏骗局”和追踪本拉登两大传闻而名声大噪。在商业领域，通过与客户的深度合作帮助其构建数据基础设施，利用数据挖掘分析实现业务成功，如摩根



大通、默克集团、法国空客、克莱斯勒等。

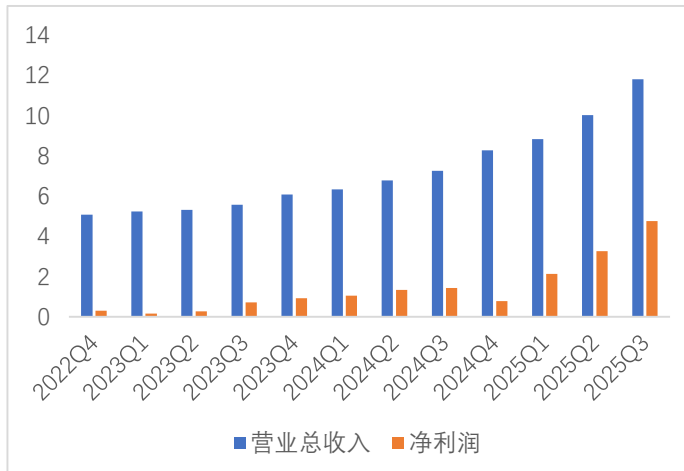
- 根据广电运通研究总院研究，与众多 SaaS 公司不同，Palantir 采用深度定制的方式与客户合作，每个实施项目都可能需要数月甚至数年时间。它不仅为客户提供软件平台，还深度参与客户的数据整合和决策流程，形成了一种“软件+咨询”的混合服务模式。这种模式使得 Palantir 能够解决客户最复杂的数据问题，虽然限制了客户基数的快速增长，但也建立了极高的客户粘性和切换成本。
- 产品包括 4 类：1) 面向政府和情报机构的 Palantir Gotham；2) 面向企业的 Palantir Foundry；3) 专门支持 Gotham 和 Foundry 的 Apollo 平台；4) 基于 Palantir 机器学习技术和大模型的 AIP 平台。其中，AIP 平台是专注于提供高级数据整合、分析和智能决策支持解决方案，提供了大型语言模型和先进人工智能技术的强大能力；无论是在机密系统还是战术边缘设备上，AIP 都能在私有网络中实现对 AI 解决方案的负责、有效和安全的部署，可以高效构建 AI 应用、持续改进 AI 驱动工作流的功能点、提供 Ontology SDK 与企业运营快速结合。

图表45: Palantir AIP 系统架构



来源：腾讯云，国金证券研究所

图表46: Palantir 收入及利润（亿美元）

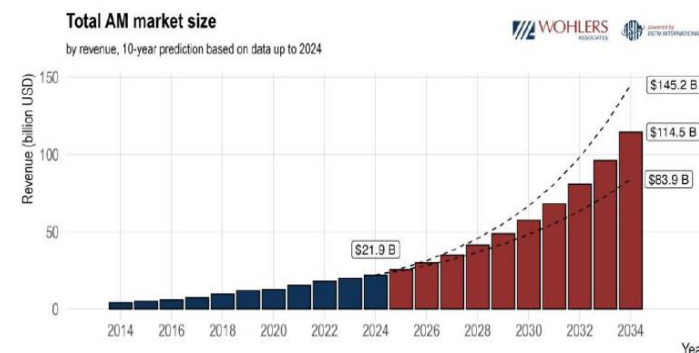


来源：Wind，国金证券研究所

六、3D 打印：当创新遇到物理瓶颈，3D 打印开启序章

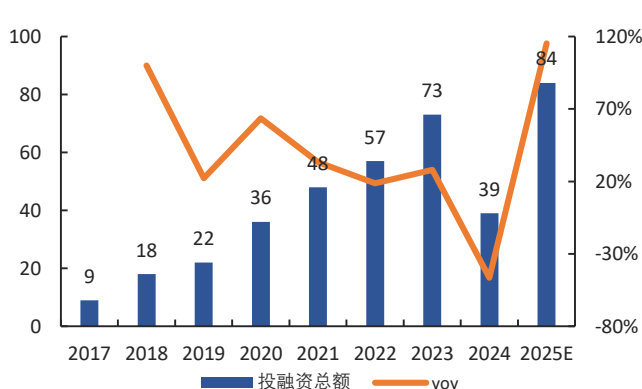
中国 3D 打印市场重回较快增长，2034 年全球市场规模有望突破 1145 亿美元。中国 3D 打印赛道投融资额重拾快速增长，2025 年全年行业投融资总额预计达 84 亿元，创历史新高，同比增长 115.4%，扭转了 2024 年同比下滑的局面。在航空航天、液冷、3C 及汽车等需求拉动下，全球 3D 打印市场将迎来快速增长，根据 WOHLERS 统计，2024 年全球 3D 打印市场规模已达 219 亿美元，预计未来 10 年将以 18% 的年复合增长率稳健向千亿美元量级迈进，预计 2034 年将突破 1145 亿美元。

图表47: 2024 年全球 3D 打印市场规模已达 219 亿美元



来源：Wohlers Report 2025，国金证券研究所

图表48: 2025 年国内 3D 打印行业投融资总额预计达 84 亿元



来源：南极熊 3D 打印公众号，国金证券研究所



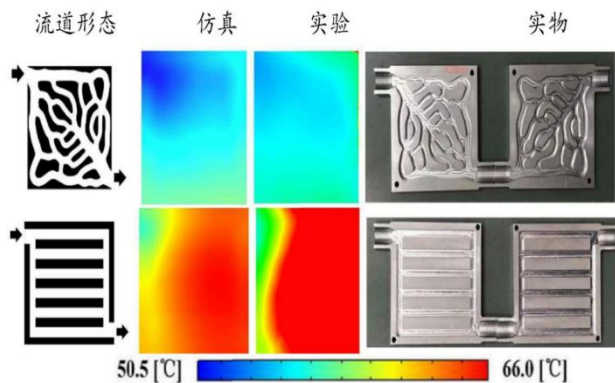
我们认为随着 AI、空天、3C、机器人、汽车等产业不断突破创新边界，传统制造工艺已临近上限，难以满足散热、轻量化等实际需要，亟需 3D 打印技术突破传统上限，多维需求爆发有望加速产业发展。

AI 大爆发，3D 打印赋能微通道解决散热痛点。随着 AI 芯片功率密度爆发式增长，具备极高换热效率的微通道液冷方案需求迫切，然而传统铲齿、CNC 及焊接工艺在制造微通道（深宽比大、几何结构复杂且直径低于 1mm 的）时，受限于加工原理无法实现精细复杂流道成型，又因焊接工序额外引入热阻存在泄漏风险，难以匹配 AI 芯片创新节奏，亟需 3D 打印助力突破制造上限。3D 打印彻底解放了流道设计束缚，其无需模具与焊接的一体化成型能力，可直接制造传统工艺无法企及的拓扑优化、仿生叶脉状等复杂流道，大幅增加换热面积并优化流体动力学性能，同时通过微米级精密制造，能从根源上消除焊接带来的热阻与泄漏隐患，提升结构强度，完美适配新一代 AI 芯片对极致散热性能的刚性需求。

图表49：微通道液冷板制造中 3D 技术路线前景更好

图表50：流道拓扑优化后可以显著提升散热性能

加工技术	原理	优势	局限性
微机械切削	通过微铣削或微车削进行加工，宽度取决于刀具主切削刃的长度	走刀速度快，热量不易聚集；加工工件的轮廓形状一致性好	加工尺寸受刀具直径限制，会对薄板造成一定程度变形
微电火花加工	利用高温火花放电对金属表面进行烧蚀以达到加工材料的目的	非接触式加工；材料去除率高	电极丝直径小，容易断裂；只能在极低走丝速度下实现高精度，加工效率低
微冲压	结合精密模具使用，通过冲压使材料变形	生产效率；成本低	主要加工高延展性材料；精密模具制造难度大
微液压成形	使用加压液体作为介质，将材料加工成理想的零件形状	效率高；能够制备出复杂形状的零件	依赖复杂专业的设备
激光加工	激光在短小时内以高峰值功率蒸发或烧蚀材料	灵活性高；表面质量好	加工精度取决于参数设置
聚焦等离子体蚀刻	通过离子束对表面进行轰击，达到去除表面原子的目的，或通过离子束轰击气相前驱体，使气相前驱体分解并沉积在表面	可以加工任何硬金属和非金属，无材料选择性；具有精确制造复杂三维微纳结构的能力	效率低；成本高；加工的基底尺寸有限；只能小批量生产
增材制造技术	以数字化模型为基础，按照计算机设定的路径，将金属粉末或丝材逐道熔凝堆积，最终制备成三维金属零件	具有较强的灵活性和可控性，能够制造传统方法所不能获得的复杂结构；减少材料浪费，高净成形	部分增材制造工艺制备的零件需要后处理；成形过程复杂，目前的研究工作不能精准预防缺陷的产生



来源：《复杂微通道热管理部件的应用与加工技术进展》，国金证券研究所

来源：《基于 COMSOL 液冷板拓扑优化设计》，国金证券研究所

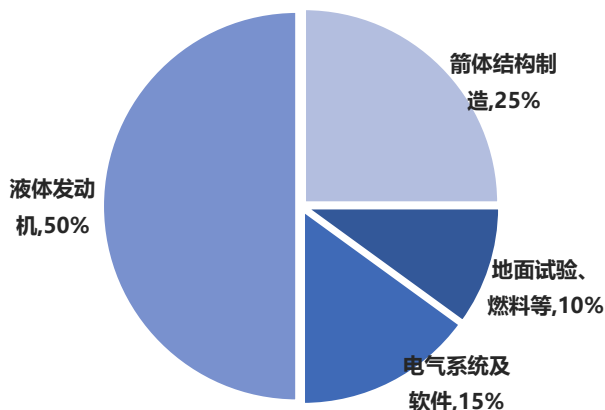
3D 打印降低火箭发动机生产成本，重塑空天时代。航空航天领域对零部件的强度、轻量化、复杂结构一体化要求极高，与 3D 打印特质天然耦合。以火箭发动机为代表，其制造成本占火箭总成本的比例高达 50%，为追求极致的性能和性价比，3D 打印广泛应用于该领域：

- 天兵科技的天龙二号采用的 3D 打印高压补燃发动机，接近 90%的部件通过增材制造完成，与传统工艺相比，该发动机制造周期缩短 70%-80%，成本与重量均降低 40%-50%。
- 星河动力的“苍穹”液氧/煤油发动机采用 3D 打印技术，实现了 4:1 的变推比和 50 次重复使用能力。飞而康科技与星河动力合作，利用华曙高科金属 3D 打印解决方案，在 4 个月内完成 30 多件火箭发动机零件的交付。
- 星际荣耀的双曲线二号验证火箭完成国内首次液体火箭全尺寸一子级垂直起降与重复使用飞行试验，铂力特为其研制了多个发动机关键零部件，有效帮助缩短生产周期，节约成本。

我们认为 3D 打印在航空航天领域，突破了传统工艺上限，以更低的成本，提供了一体化的复杂结构生产能力、更快速的研发节奏、更好的强度和轻量化表现，作为必要的生产环节，未来有望随着商业航天的快速爆发，迎来确定性的需求爆发。



图表51：发动机成本占火箭总成本 50%



来源：你好太空，国金证券研究所

图表52：第三代猛禽发动机通过 3D 打印工艺，将前两代的多零部件整合为一



来源：宇航总体技术，国金证券研究所

3D 打印通过拓扑优化，解决具身智能轻量化与散热双重约束。

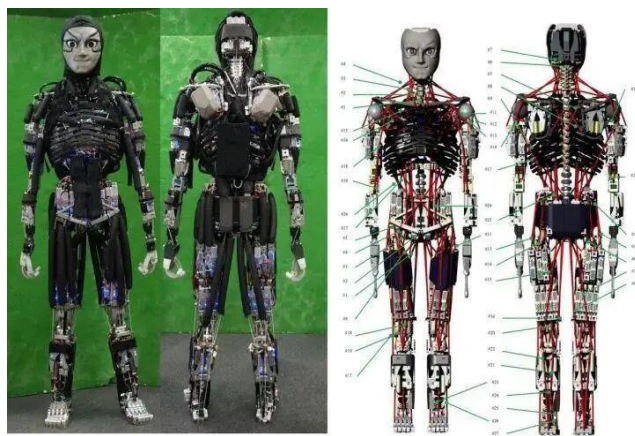
- 在轻量化方面，3D 打印凭借其对复杂几何形状的成型能力，能够通过拓扑优化设计和点阵晶格结构，制造出高强度且中空的骨骼与肢体部件，在大幅降低机器人自重、提升运动灵活性与续航能力的同时，通过一体化成型显著减少零部件数量与装配难度。
- 在散热管理层面，面对人形机器人关节内部空间狭小且电机高负荷运转产生的热聚集问题，传统散热手段往往难以奏效，而 3D 打印能够将热管理功能直接融入结构件设计，通过构建复杂的一体化导热壳体与优化内部热传导路径，极大增加了有效散热表面积并提升了热交换效率，有效解决了高集成度下的热聚集失效风险，为具身智能的高动态性能迭代奠定了物理基础。

图表53：3D 打印制造轻量化弹性体为 IRON 机器人塑造“肌肉”



来源：36 氪公众号，国金证券研究所

图表54：Kengoro 机器人使用金属 3D 打印打造多孔框架为全身电机降温

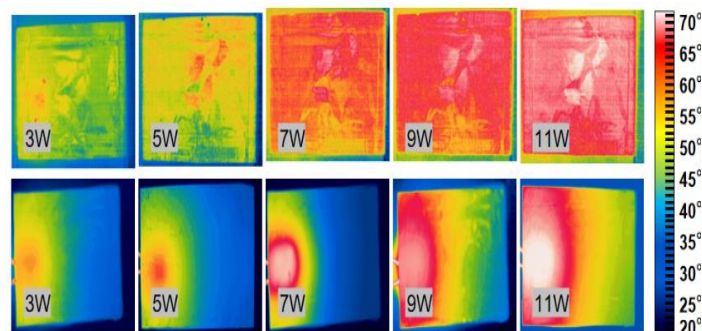


来源：车乾科技公众号，国金证券研究所

端侧 AI 时代，3D 打印赋能端侧热管理与定制化。随着 AI 大模型在手机及 AIPC 等移动端落地，芯片瞬时高算力引发的热流密度激增与设备极致轻薄化的空间限制形成矛盾，3D 打印技术能够制造出传统冲压或蚀刻工艺无法实现的、适应异形狭窄空间的超薄均热板，通过构建复杂的内部三维毛细结构大幅提升被动散热效率，保障端侧 AI 的稳定运行；另一方面，针对折叠屏手机铰链等关键结构件向高强度钛合金转型的趋势，3D 打印有效克服了钛合金材料传统 CNC 加工良率低、刀具磨损大且难以切削复杂内腔的难题，实现了铰链轴盖等精密部件的一体化成型与拓扑减重，在满足数十万次折叠寿命的同时为电池释放了宝贵的堆叠空间，成为高端 3C 产品迈向轻量化与耐用性平衡的必选方案。



图表55：超薄均热板能够高效消除局部热点，保障高功率柔性电子设备稳定运行



来源：《High-precision additive manufacturing enabled ultra-thin flexible vaporchamber for adaptive and efficient electronic cooling》，国金证券研究所

图表56：OPPO Find N5 采用 3D 打印打造铰链、翼板及外传轴中框



来源：铂力特市场营销公众号，国金证券研究所

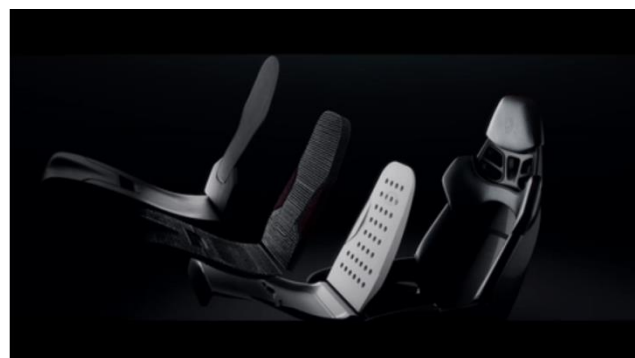
3D 打印帮助汽车突破轻量化瓶颈。在整车结构一体轻量化领域，传统冲压、铸造工艺受限于分体制造逻辑，需通过焊接组装多部件，存在重量冗余，又有连接点疲劳失效风险，而 3D 打印技术能够利用拓扑优化算法，将底盘副车架、控制臂等关键承力部件设计为仿生镂空或点阵结构并实现一体化成型，在大幅降低簧下质量的同时避免了多部件焊接带来的强度衰减；同时，针对高端车型日益增长的个性化需求，该技术彻底摆脱了内饰件与外壳制造对昂贵模具的依赖，能够直接制造出具有复杂晶格结构的透气座椅中底、特殊纹理仪表盘及异形车身外壳，不仅显著缩短了研发试制周期，更以低成本小批量生产模式满足了智能座舱对功能性与美学的双重定制需求。

图表57：3D 打印汽车零部件在保障强度的同时能够有限减重

零件名称	3D 打印优势
立柱	该零件基于拓扑优化创新设计的 3D 打印技术，传统铸造或机加工无法制造，能够保证极高的制造精度，显著提升轻量化表现，助力赛车成为中国大学生方程式汽车大赛 8 年来最轻赛车
轮毂	3D 打印降低零件重量，解决传统铸造因结构复杂导致的缺陷问题，提升成品率，保障零件性能稳定，满足设计要求
摇臂	采用高强铝 3D 打印。通过拓扑优化设计，相比传统制造方法减重 30%，力学性能完全满足应用要求

来源：铂力特招股说明书，国金证券研究所

图表58：保时捷提供定制 3D 打印座椅



来源：叁迪科技公众号，国金证券研究所

七、金融 IT：降息周期全球资本活跃，跨境支付重构金融生态

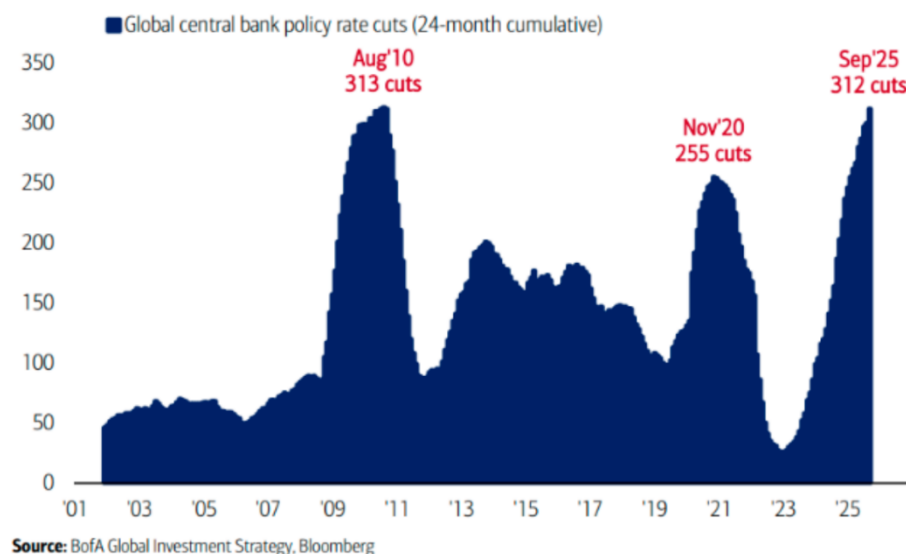
降息周期全球资产打开估值弹性，金融 IT 业绩有望修复。全球流动性迎来拐点，2025 宽松周期叠加 2026 降息预期，有望改善风险资产估值环境。据美国银行统计，过去两年全球央行降息频次已超 2008 年金融危机时期，虽然当前中性利率仍处高位，但机构预测美联储未来 12-18 个月仍有 100BP 的降息空间，中性利率有望回归至 3%-3.25% 区间。更为关键的边际变化在于，美联储会议纪要暗示量化紧缩或将于 12 月结束，甚至不排除在 2026 年重启量化宽松以对冲就业压力。我们认为流动性的实质性改善将有效降低资金成本并推升资产价格，



资本市场交易活跃度有望回升，利好交易系统及相关 IT 服务商的业绩修复。

图表59：过去 24 个月，全球央行多次降息

Chart 2: Number of global rate cuts past 24 months almost at post-GFC high
Global central bank policy rate cuts: 2-year cumulative



来源：BofA Global Investment Strategy, Bloomberg, 国金证券研究所

“十五五”明确资本市场高质量发展，资本市场 IT 迎来建设新浪潮。2025 年 10 月 28 日，证监会副主席李超在科创板相关仪式上明确表示，证监会将抓紧谋划“十五五”时期资本市场战略任务，核心围绕“防风险、强监管、促高质量发展”三大支柱展开。当前政策端强调以深化投融资综合改革为牵引，旨在持续增强资本市场的包容性与适配性，服务于金融强国建设目标。我们认为，随着新一轮资本市场改革，资本市场 IT 板块有望迎来新的系统性建设浪潮。

图表60：“十五五”规划明确资本市场“高质量发展”核心主线

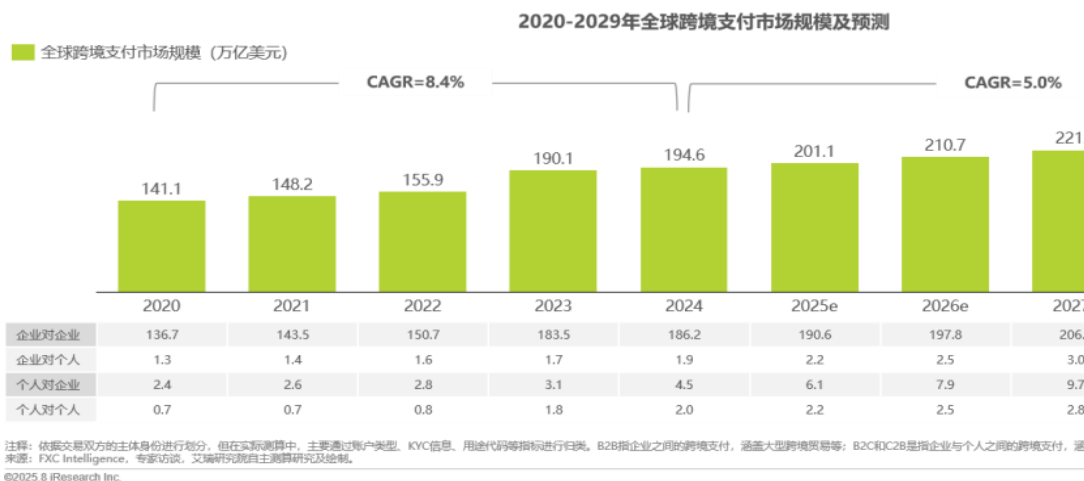


来源：证监会发布公众号，国金证券研究所

跨境支付规模稳步增长，全球监管框架加速演进。跨境支付规模持续增长，2020 至 2024 年全球跨境支付规模从 141.1 万亿美元增至 194.6 万亿美元，其中 B2B 业务增速趋稳，而受益于电商出海与数字服务扩张，以个人为参与方的跨境支付复合增速超 30%，实现快速增长。全球监管框架加速演进，针对反洗钱（AML）、反恐怖融资（CFT）及客户身份识别（KYC）的合规要求日益严苛，监管体系向模块化、实时化转型。覆盖范围已拓展至传统跨境支付、数字资产、加密资产等多元支付形态；监管机构也在积极引入实时风控、监管科技等工具，强化动态风险的识别与管理能力。



图表61：2024 年全球跨境支付规模增长至 194.6 万亿美元



来源：iresearch 网站，国金证券研究所

数字人民币再升级，进一步助力建设金融强国，推动人民币国际化。2026 年 1 月 1 日，央行出台《关于进一步加强数字人民币管理服务体系和相关金融基础设施建设的行动方案》，新一代数字人民币计量框架、管理体系、运行机制和生态体系将于正式启动实施。本轮数字人民币升级的目的，我们认为主要系：1) 为落实十五五规划要求的“加快建设金融强国”“稳步发展数字人民币”指示；2) 为应对全球虚拟资产/稳定币等挑战，积极拥抱区块链及智能合约技术，提高跨境支付效率。从行动方案细节看，最核心的变化有三点：

- **新增计息收益。**升级前的数字人民币本质是“数字现金”，拿在手里的现金而非存在银行的存款，无法产生利息收入，升级后的数字人民币本质是“数字存款”，银行需为用户钱包里的数字人民币计息付息，有利于激励用户使用意愿；同时，作为“数字存款”存在，受存款保险保护，安全保障更高。
- **加强智能合约功能。**智能合约能力的升级，为数币服务商提供更丰富的“武器库”，有利于创新供应链金融、绿色金融、预付资金管理、定向补贴等服务。
- **优化跨境体验。**升级版数字人民币将优化跨境支付体验，比如中国游客在老挝旅游时，可用数字人民币 App 直接扫码支付，无需兑换外币，便捷中国游客和老挝商户体验。

我们认为在全球支付创新背景下，数字人民币被赋予了新内涵，对内以数字存款计息为切入点吸引更多用户，对外可以便利跨境支付，整体有望进一步推动人民币国际化行稳致远。

图表62：三大平台协同解决了传统跨境支付的效率与成本痛点

平台名称	核心功能	解决痛点
跨境数字支付平台	支持法定数字货币跨境清算	传统跨境支付效率低、成本高
区块链服务平台	提供跨境交易转接及链上支付服务	区块链互操作性不足
数字资产平台	为金融级数字资产提供发行、托管等标准化服务	数字资产缺乏合规基础设施

来源：每经网，国金证券研究所

八、智能驾驶：L2 强标落地，L3 元年，Robotaxi 单车 UE 转正

＜L2 辅助驾驶：AEBS 系统或升级为强制性国标，“中国 L2 强标”征求意见稿发布。2025 年 2 月 28 日，《轻型汽车自动紧急制动系统技术要求及试验方法》完成起草，于 4 月 30 日进入公开征求意见阶段。该标准将代替现行国标 GB/T39901-2021《乘用车自动紧急制动系统（AEBS）性能要求及试验方法》，由“推荐性国标”升级为“强制性国标”。

对比现行国家标准，新国标要求 AEBS 在覆盖范围、激活速度和误检&漏检等方面全面升级。1) 新国标覆盖范围从 M1 类乘用车扩展至 N1 类轻型载货汽车；2) M1/N1 类汽车 AEBS 系统激活速度区间由 15km/h 以上分别改为 10~80km/h/10~60km/h，同时明确对行人和两轮车有效识别的激活速度区间为 20km/h~60km/h；3) 误响应试验新增右转跟车、静止自行车等复杂场景，提升车道内铁板误检难度，试验漏检率要求由 40%降至 10~20%。



图表63：新旧国标对比：适用对象由 M1 扩展至 N1 类汽车，AEBS 系统激活速度区间及误检&漏检要求升级

	现行国标 GB/T 39901-2021	新国标征求意见稿
文件名	《乘用车自动紧急制动系统（AEBS）性能要求及试验方法》	《轻型汽车自动紧急制动系统技术要求及试验方法》
适用对象	适用于 M1 类车辆	适用于 M1 及 N1 类汽车
标准性质	推荐性标准：适用于安装 AEBS 的 M1 车辆	强制性标准：全部 M1 及 N1 车型应满足文件规定的全部要求
车辆碰撞激活速度区间	15km/h 以上	M1：10km/h~80km/h N1：10km/h~60km/h
行人、自行车、摩托车碰撞激活速度区间	无	20km/h~60km/h
预警性能	不晚于紧急制动前 1s	车辆碰撞：不晚于紧急制动前 0.8s 其他：不晚于紧急制动
紧急制动性能	紧急制动不应在预计碰撞时间 3s 前开始	紧急制动后车辆的减速度绝对值的最大值不小于 5.0m/s ²
误响应试验	相邻车道车辆、车道内铁板误响应试验	相邻车道车辆、车道内铁板、跟车过程中车辆目标右转、经过同向运动的成年行人、经过对向静止自行车误响应试验
车道内铁板误响应试验	直径 600mm、厚度 10mm 的圆形铁板	2.4 m×3.7 m×25 mm 的长方形铁板
试验方法	实车测试	部分试验可通过仿真试验开展，至少 30%的试验通过场地开展
试验要求	预警及启动试验 5 次中至少 3 次满足规定；误响应试验 5 次中至少 4 次满足规定	预警及启动实验中车对车、车对行人试验通过次数占比不低于总次数的 90%；车对自行车、摩托车通过次数占比不低于总次数的 80%

来源：《乘用车自动紧急制动系统（AEBS）性能要求及试验方法》，《轻型汽车自动紧急制动系统技术要求及试验方法》，国金证券研究所

全球范围内，欧洲、美国、日本均已通过相关法规要求在不同程度上强制要求安装 AEBS 系统。1）24 年 4 月，美国发布 FMVSS 127，要求自 2029 年 9 月 1 日起，所有新生产的轻型车必须配备符合该标准的 AEBS 系统（目前 AEB 系统仍为推荐技术，NHTSA 的新车评估计划自 2018 年起将 AEBS 纳入测试）。2）21 年 4 月，欧盟发布 UN R152，要求自 2022 年 7 月 6 日起进入欧盟市场的新车型都需要配备满足 UN-R152 法规的 AEBS 功能。3）2020 年底，日本宣布本土新车型自 2021 年 11 月起，进口新车型 2024 年 6 月起，必须装配 AEBS 系统并符合 UN R152；对于已批准车型，本土新车需自 2025 年 12 月起、进口新车于 2026 年 6 月起，必须装配 AEBS 系统并符合 UN R152。

图表64：欧洲、美国、日本均通过强制法规要求安装 AEBS 系统

	最新法规	发布时间	适用范围	激活速度区间	实施要求
美国	FMVSS 127	2024 年 4 月 29 日	适用于轻型车辆(总质量 ≤4536kg)	车辆碰撞： 最高 145km/h 检测到行人： 最高 72km/h	自 2029 年 9 月 1 日起，所有新生产的轻型车必须配备符合该标准的 AEBS 系统(目前 AEBS 系统仍为推荐技术，NHTSA 的新车评估计划自 2018 年起将 AEBS 纳入测试，测试最大刹停速度为 40km/h)。
欧盟	UN R152	2021 年 4 月 2 日	M1 类和 N1 类车辆	10km/h~60km/h	自 2022 年 7 月 6 日起进入欧盟市场的新车型都需要配备满足 UN-R152 法规的 AEBS 功能。
日本	UN R152	2021 年 4 月 2 日	M1 类和 N1 类车辆	10km/h~60km/h	日本本土新车型自 2021 年 11 月起,进口新车型 2024 年 6 月起，必须装配 AEBS 系统并符合 UN R152；对于已批准车型，本土新车需自 2025 年 12 月起、进口新车于 2026 年 6 月起，必须装配 AEBS 系统并符合 UN R152。



图表64：欧洲、美国、日本均通过强制法规要求安装AEBS系统

来源：UNECE 官网，NHTSA 官网，Federal Register 官网，中汽研科技、中国汽研检测工程事业部、中科创达科技微信公众平台，国金证券研究所

AEBS 系统分为感知层（传感器）、决策层（算法/芯片）和执行层（线控制动）三大模块，核心是通过传感器实时监测前方环境，在判定碰撞风险超过阈值且驾驶员未采取制动措施时，依次触发声光警报、紧急制动响应机制。在技术路线方面，当前 AEBS 传感器方案主要是对激光雷达、毫米波雷达、摄像头的不同组合。随着新国标对激活速度区间和误检&漏检率的要求提高，激光雷达具有明显优势。

新国标要求 M1 类汽车的 AEBS 系统至少在 10km/h~80km/h 速度区间内处于激活状态，N1 类汽车激活速度区间为 20km/h~60km/h，并要求紧急制动后车辆减速度绝对值最大值不小于 5.0m/s^2 。相对于现行国标要求最低触发速度 $\geq 15\text{km/h}$ ，激活速度区间大幅扩展。实际运用中，行业也在不断追求更高的 AEBS 刹停速度，华为问界新 M7 目前升级至 120km/h 的刹停时速，理想 L6 可以实现雨夜面对前方事故车 120km/h 刹停。

对于 AEBS 激活速度区间的扩展实际上是对传感器探测距离的更高要求。以 5.0m/s^2 减速度计算，80km/h 速度行驶的汽车至少需要约 73m 的刹车距离，而当车速达到 130km/h 时，这一距离扩展至 169m。

图表65：AEBS 激活速度区间扩展后需要更远的探测距离

5m/s ² 减速度下												
车辆速度 (km/h)	20	30	40	50	60	70	80	90	100	110	120	130
反应距离 (m)	6	9	12	15	18	21	24	27	30	33	36	39
制动距离 (m)	3	7	12	19	28	38	49	63	77	93	111	130
刹车距离 (m)	9	16	24	34	46	59	73	90	107	126	147	169

来源：感光现象微信公众平台，国金证券研究所（注：上表刹车距离=反应距离+制动距离。）

据禾赛微信公众号披露，相比无激光雷达车型，装有激光雷达的车型，其 AEBS 速度上限具有约 50% 的明显提升：1）在无激光雷达情况下，AEBS 速度上限约为 85km/h，搭载激光雷达后提升至 140km/h；2）夜间行驶环境下，无激光雷达情况下，AEB 速度上限约为 80km/h，搭载激光雷达后提升至 120km/h。

同时，新国标新增对行人、两轮车的识别要求，试验漏检率要求由 40% 降至 10~20%，对传感器精度要求进一步提升。纯视觉方案中摄像头作为模拟人眼的视觉工具，在强光、夜间等复杂光照条件下性能显著降低，有国际论文指出夜间融合了激光雷达的平均感知精度是纯视觉的 3 倍。毫米波雷达随着探测距离的提升精度大幅降低，据雷军介绍，4D 毫米波雷达仅能识别 100m 的人，目前市场上的 4D 毫米波雷达每帧约输出 1,000 多点，而主流的 128 线激光雷达每帧输出可达十几万点，分辨率差异较大。

激光雷达能够在 200m 探测距离下保持厘米级精度。以速腾聚创的超长距数字化激光雷达 EM4 为例，其支持最远 600 米测距，为系统预留作出决策和响应的的时间将最高增加 70%，在提高安全性的同时，可显著提升 AEBS 系统覆盖的速度范围。

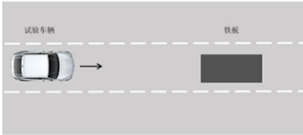
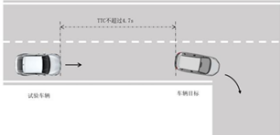
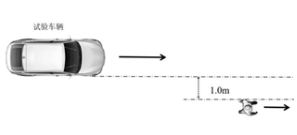
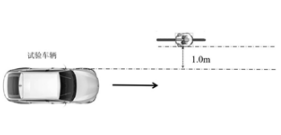
新国标在误响应控制上趋严，毫米波雷达恐难以满足要求。新国标增加了跟车过程中车辆目标右转、经过同向运动的成年行人、经过对向静止自行车误响应试验，车道内铁板误响应试验铁板厚度由 10mm 增加至 25mm。

3D 毫米波雷达将前方物体压缩至二维平面，不能判断物体轮廓，不具备高度信息识别能力。车道内铁板易被识别为障碍物目标。车道上例如自行车、桥洞等静止物体易触发毫米波回波形成检测点，和属于前车的毫米波触发误匹配，造成“幽灵刹车”，严重甚至造成后车追尾。激光雷达则通过点云图对环境进行三维重建，无需提前训练即可识别上述障碍物，增强汽车在复杂场景下的感知能力，降低误触发率。

4D 毫米波雷达增加了垂直角度检测，能够提供高度信息，但分辨率相比激光雷达仍有很大差距。目前 4D 毫米波雷达整体价格约在千元，相比目前已进入 200 美元时代的激光雷达成本优势并不明显。我们认为，在实际量产车辆中，4D 毫米波雷达应用较少，已错过上车窗口期。



图表66：新国标新增多项误响应试验，3D 毫米波雷达在技术层面易造成误触发

	车道内铁板误响应	跟车过程中车辆目标右转误响应	经过同向运动成年行人误响应	经过对象静止自行车误响应
毫米波雷达	不具备高度信息识别能力，易将铁板识别为障碍物目标，引起 AEB 系统误触发。	判定为前车减速，引起 AEB 系统误触发。	分别属于前车和行人/两轮车的毫米波检测被错误匹配，判定为前车突然刹停或减速，引起 AEB 系统误触发。	
激光雷达	通过点云图对环境进行三维重建，无需提前训练即可识别上述障碍物，增强汽车在复杂场景下的感知能力，降低误触发率。			
试验示意				

来源：《轻型汽车自动紧急制动系统技术要求及试验方法》，焉知汽车、深圳海冰科技有限公司、智能车参考微信公众平台，国金证券研究所

2025 年 9 月 17 日，工信部对《智能网联汽车 组合驾驶辅助系统安全要求》强制性国家标准公开征求意见。标准针对单车道、多车道、领航辅助等不同功能设置全方位安全技术要求。标准基于我国道路交通特征，构建了包括道路交叉口、施工区、环岛、隧道等交通环境的试验场景；为确保系统声明的各项能力均经过充分验证，要求至少通过场地试验、道路试验和文件检验验证系统符合相关安全要求等。

图表67：L2 强标对复杂环境下的多目标状态识别提出较高要求

触发条件类别		触发条件
环境	天气	雨、雪、雾
	光照	一天中的时间（清晨、白天、傍晚、夜晚）、光照变化（逆光、强光照、光照突然变化）
交互	道路类型	弯道、匝道、隧道、坡道
		道路交叉口、环岛
	道路标记	没有车道标记、断线、变浅的车道标记、多车道标记
	道路设施	施工区域、交通指示灯、道路交通标志
		龙门架、收费站
		限高架
	目标物类型	机动车、行人、摩托车、动物、临时障碍物（如纸箱等掉落物）、异形目标物（洒水车/路障车等异形车辆）、低对比度目标物
		自行车、电动二轮车、异形目标物（撑伞行人、披雨衣/斗篷电动二轮车）
	目标物状态	直立静止、运动、侧翻、斜置
	多目标交互	遮挡、横穿、并行、对向行驶、切入、切出、紧急制动
驾驶员误用		误干预、激活后误用、误激活

来源：《智能网联汽车 组合驾驶辅助系统安全要求》征求意见稿，国金证券研究所

L3 有条件的自动驾驶：“接管缓冲期+责任转移”，工信部公布首批 L3 准入许可。25 年 12 月，工信部正式公布我国首批 L3 级有条件自动驾驶车型准入许可，两款分别适配城市拥堵、高速路段的车型将在北京、重庆指定区域开展上路试点，标志着我国 L3 级自动驾驶从测试阶段迈入商业化应用的关键一步。

相比 L2 辅助驾驶，L3 自动驾驶允许驾驶员将手和视线从方向盘和道路上移开，碰撞责任或转移至主机厂。目前，海外已有成熟实践。梅赛德斯的 L3 级旗舰车型已在加利福尼亚州、内华达州和德国投入使用，预计到 2026 年，英国也将看到 L3 级车辆上路。



L3 系统需要在驾驶员不干预的情况下运行，需要更为周密的安全冗余设计。联合国《自动驾驶法规 R157》明确要求 L3 级系统需提供 10 秒接管缓冲期。梅赛德斯 L3 级车型 DRIVE PILOT 系统能够在 40 英里/小时（约合 64.4 公里/小时）以内启用，设置 10 秒的响应接管时间，10 秒内驾驶员未接管将启动紧急制动程序，这一目标需要超 200 米的探测距离，高度依赖激光雷达的超视距感知能力。DRIVE PILOT 系统在梅赛德斯 L2 级自动驾驶系统的基础上增加一颗激光雷达、一枚车轮摄像头、一枚后视镜摄像头，同时还有 V2X 模块、厘米级高精度定位模块和高精地图。

L2 级和 L3 级之间的关键区别在于责任的转移。梅赛德斯承诺，车主在启动 DRIVE PILOT 期间，如果发生车祸等安全事故，车主原则上不用负责，由主机厂承担全部责任。因而，我们认为，主机厂对于提高 L3 级别无人驾驶的安全性更为重视，以确保责任成本不会超过部署 L3 级系统的潜在收益，其中，高性能激光雷达不可或缺。

图表68：梅赛德斯 L3 级 DRIVE PILOT 系统配备一颗激光雷达

图表69：梅赛德斯 L3 级 DRIVE PILOT 系统设置 10 秒响应接管时间



来源：梅赛德斯官网，国金证券研究所

来源：梅赛德斯官网，国金证券研究所

国内车企对 L3 级自动驾驶的落地布局正在全面提速。1）政策上，《北京市自动驾驶汽车条例》、《武汉市智能网联汽车发展促进条例》、《深圳经济特区智能网联汽车管理条例》等法规的出台，对 L3 级自动驾驶车辆的上路流程、事故责任界定等关键方面进行了规定，推动了 L3 级自动驾驶商业化落地的进程。2）技术上，激光雷达环境感知愈发精准，安全冗余度达到较高水平，过去制约 L3 级普及的高成本瓶颈也逐渐被打破。

今年上半年，国内车企密集公布 L3 进展，L3 架构车型均标配多颗激光雷达。2025 年 4 月，华为发布了新一代 ADS 4.0 智能驾驶辅助系统，支持高速 L3 级自动驾驶，是中国首个高速 L3 商用解决方案。基于 ADS 4.0 系统的问界新 M9 和尊界 S800 按照 L3 智能驾驶架构设计，均配置 4 颗激光雷达。2025 年 3 月 3 日，吉利发布千里浩瀚安全高阶智能驾驶系统，千里浩瀚 H9 作为具有量产落地能力的 L3 解决方案，配置 5 颗激光雷达。

在 L3 责任转移问题上，国内车厂提出智驾险解决方案。华为针对问界全系新增智驾无忧服务权益（试行），场景包括智能泊车辅助、代客泊车辅助、遥控泊车辅助、车道巡航辅助及智驾领航辅助，权益时长为自权益激活后 12 个月，最高保障权益上限 500 万元。小鹏汽车推出智能辅助驾驶安心服务，宣称是全车系、全场景，支持五大头部保险公司，赔付不限次数，是全行业唯一支持 NGP 退出 5 秒内仍可生效的保险方案。

图表70：国内车企密集公布 L3 进展，已公布车型均标配 4-5 颗激光雷达

	L3 进展	发布时间	L3 架构车型	激光雷达数量
华为	发布新一代 ADS 4.0 智能驾驶辅助系统	2025 年 4 月 22 日	问界新 M9、尊界 S800	4
吉利	发布千里浩瀚安全高阶智能驾驶系统	2025 年 3 月 3 日	极氪千里浩瀚智驾 H9	5
奇瑞	推出“猎鹰智驾”系统，宣布将于 2026 年实现 L3 级技术上车	2025 年 3 月 18 日	/	/
广汽	发布智能科技品牌“星灵智行”，宣布将于今年第四季度量产并上市销售 L3 自动驾驶乘用车	2025 年 3 月 18 日	/	/

图表70: 国内车企密集公布L3进展, 已公布车型均标配4-5颗激光雷达

长安	明确 2026 年天枢智驾将实现全场景 L3 级自动驾驶	2025 年 2 月 9 日	/	/
----	------------------------------	----------------	---	---

来源：太平洋汽车、中国汽车报、赛博汽车、智能车参考、科技每日推送、汽车天涯微信公众平台，国金证券研究所

L4 自动驾驶：小马智行 Robotaxi UE 模型打平，中美主流厂商进展提速。 L4 为高度自动驾驶，要求车辆在特定场景内，无需人类驾驶员干预，独立、安全地应对所有突发状况。目前，中美两国已在武汉、北京、上海、凤凰城、旧金山等多个城市开展 Robotaxi 商业化运营。Robotaxi 对环境感知能力和安全冗余能力提出更高要求，目前主流 Robotaxi 车型均采用融合激光雷达的感知方案，单车搭载 7-10 颗激光雷达（Waymo 为 4 颗机械式环视方案）。

年末小马智行率先在广州实现了城市级单车盈利转正。在 12 月 16 日的媒体沟通会上，小马智行联合创始人、CFO 王皓俊披露，目前广州地区的 Robotaxi 日均订单已达到 23 单/辆，日均营收达到 299 元/辆（已扣除折扣和退款），这一水平支撑广州车队在 11 月实现了单车营收平衡，构建起一个可验证的单车盈利模型。小马智行第七代 Robotaxi 于 11 月起在北京、广州、深圳开启全无人驾驶商业化运营。第七代车型是车队快速扩张的主要支撑，过往第五代、第六代车型最大的问题在于“太贵”，每辆车的成本达到“大几十万级”，因此投放越多也就亏损越多。但是第七代车自动驾驶套件成本下降了 70% 以上，且明年成本相比今年能再下降 20%。

图表71: 国内 Robotaxi 落地车型搭载 7-10 颗激光雷达

	最新硬件平台	发布日期	合作车企	技术架构	搭载激光雷达数量
滴滴	滴滴自动驾驶全新一代硬件平台	2025年4月23日	广汽埃安	滴滴自动驾驶硬件平台4.0 2个声音传感器 4个4D毫米波雷达 6个近距补盲激光雷达 4个长距激光雷达 1个红外相机 1个8M像素长距相机 4个近距补盲鱼眼相机 1个Falcback前毫米波雷达 1个Falcback中距相机 1个红绿灯摄像头 6个8M像素中距相机 2个Falcback角毫米波雷达	10 (4颗长距激光雷达, 6颗近距补盲激光雷 达)
小马智行	第七代车规级自动驾驶软硬件系统方案	2025年4月23日	丰田、北汽、广汽埃安	XPENG XPILOT 7.0 4D成像毫米波雷达x1：探测距离550米，探测角度90°×15°，应用物体检测。 - 前向长距激光雷达x1：探测距离450米，探测角度10°×15°，应用物体检测。 - 侧向前距激光雷达x1：探测距离150米，探测角度10°×15°，应用物体检测。 - 后向中距激光雷达x1：探测距离120米，探测角度10°×15°，应用物体检测。 - 远端激光雷达x1：探测距离500米，探测角度5°×15°，应用物体检测。 - 侧前方激光雷达x4：探测距离200米，探测角度10°×15°，应用物体检测。 - 侧后方激光雷达x4：探测距离200米，探测角度10°×15°，应用物体检测。 - 广角镜头x4：探测范围更广，提升感知精度。 - 超广角镜头x4：探测范围更广，提升感知精度。 - 中距激光雷达x4：探测距离200米，探测角度10°×15°，应用物体检测。 - 补盲激光雷达x4：探测距离50米，探测角度10°×15°，应用物体检测。	9 (4颗中距激光雷达, 4颗补盲激光雷达)
百度Apollo	Apollo RT6	2022年7月21日	江铃新能源		8 (4颗主激光雷达, 4 颗补盲激光雷达)
文远知行	WeRide Sensor Suite 5.0	2022年6月7日	宇通、博世等	WeRide Sensor Suite 5.0 - MEMS LIDAR：固态激光雷达 - Integrated Antenna：一体化集成天线 - High Definition Camera：高清摄像头	7 (顶部尾翼4颗主激 光雷达, 3颗补盲激 光雷达)



图表71：国内 Robotaxi 落地车型搭载 7-10 颗激光雷达

Waymo	第六代传感器套件	2024 年 6 月	极氪		4
特斯拉	FSD 套件	承诺 6 月开启运营	/	/	/

来源：滴滴自动驾驶、百度 Apollo、小马智行、文远知行 WeRide、智能车参考微信公众平台，国金证券研究所

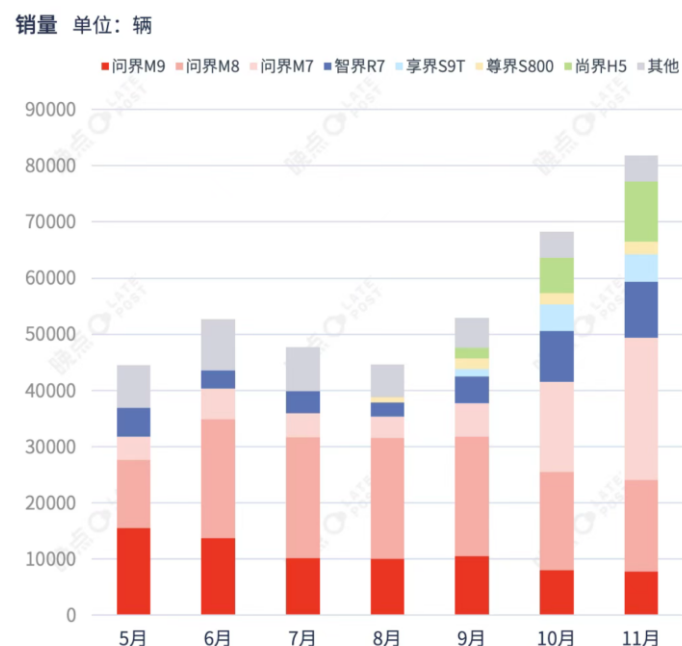
鸿蒙智行增长最为陡峭的一年，五大品牌几乎实现全市场画像覆盖。鸿蒙智行通过五大品牌覆盖了 15-100+万元价格带，各品牌在 2025 年承担了不同的市场分工：

- 问界：销量主力，全新问界 M7 上市 68 天交付量突破 4 万辆，位居售价 30 万元以上市场第一；问界 M8 上市 7 个月累计交付超 13 万辆，是售价 40 万元以上市场的交付冠军；问界 M9 上月交付超 7500 辆，继续保持售价 50 万元以上市场累计交付第一的位置。
- 尚界：定位入门级市场，11 月单月交付超 1 万辆。
- 智界：面向年轻群体，月交付量连续两个月破万。
- 享界：享界 S9 系列成为售价 30 万元级新能源豪华轿车销量冠军。
- 尊界：尊界 S800 连续三个月位居 70 万元以上豪华轿车销量第一。

从整体销量规模来看，鸿蒙智行达成首个“100 万辆”交付累计用时 43 个月。基于当前增速，余承东预测第二个 100 万辆的交付周期将缩短至 10 个月左右。

今年四季度，由于全国多地置换补贴政策相继退出，汽车市场的增长压力比往年更大，但鸿蒙智行依然保持销量增长。据晚点，截至 12 月 29 日 14 时，鸿蒙智行本月销量已达 83,241 辆，超过上个月的 81,864 辆，刷新历史月销纪录，年末强势收官。

图表72：2025 年 11 月鸿蒙智行单月交付破 8 万辆



来源：晚点 latepost 微信公众平台，国金证券研究所

图表73：鸿蒙智行用时 43 个月累计交付量破百万辆



来源：晚点 latepost 微信公众平台，国金证券研究所

风险提示

- 行业竞争加剧的风险：



在政策持续加码支持计算机行业发展的背景下，众多新兴玩家参与到市场竞争之中，若市场竞争进一步加剧，竞争优势偏弱的企业或面临出清，某些中低端品类的毛利率或受到一定程度影响。

■ 技术迭代不及预期的风险：

计算机行业技术开发需投入大量资源，如果相关厂商新品研发进程不及预期，表观层面将呈现出投入产出在较长时期的滞后特征。

■ 特定行业下游资本开支周期性波动的风险：

部分计算机公司系顺周期行业，下游资本开支波动与行业周期性相关性较强，或在个别年份对于上游软件厂商的营收表现产生扰动。



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。

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：北京市东城区建国门内大街 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