

电子行业深度报告

从云端算力国产化到端侧 AI 爆发, 电子行业的戴维斯双击时刻——电子行业 2026 年投资策略

2025 年 12 月 10 日

证券分析师 陈海进

执业证书: S0600525020001

chenhj@dwzq.com.cn

证券分析师 谢文嘉

执业证书: S0600525120001

xiewenjia@dwzq.com.cn

研究助理 解承堯

执业证书: S0600125020001

xiechy@dwzq.com.cn

研究助理 李雅文

执业证书: S0600125020002

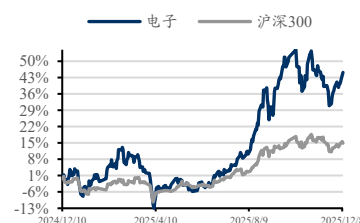
liyew@dwzq.com.cn

增持（维持）

投资要点

- **云端算力芯片: AI CapEx 大力投入, 看好国产算力业绩释放。**全球 CSP 资本开支延续季度上行趋势, 25Q3 海外四大 CSP (谷歌、亚马逊、微软、Meta) 的资本开支合计 979 亿美元, 环比+10%。国内算力在 Token 需求 (如字节等) 逼近海外巨头而供给不足的背景下, 追赶空间广阔, 建议重点关注寒武纪、海光信息等。此外华为昇腾作为国产核心底座, 率先推行“超节点”战略 (Atlas900/950 SuperPoD), 通过万卡级无收敛互联实现性能对标英伟达 GB200。
- **端侧算力芯片: 看好海外大模型带动产业链公司受益, 以及 NPU 技术路径创新开拓新应用场景。**(1) 海外链: 从海外科技巨头来看, 端侧 AI 的战略地位正快速抬升并进入全面落地阶段。谷歌将模型集成到搜索、Gmail、Android 等核心产品中, 并结合其既有硬件终端布局, 形成从云端到端侧的体系化协同。整体看, 海外链端侧 AI 需求的确性增强, 有望驱动 SoC 厂商在新一代轻量化大模型终端中实现结构性突破。(2) NPU 协处理器: 端侧模型升级催生硬件架构向专用协处理器演进。以瑞芯微为代表的厂商正率先推出面向端侧 AI 的协处理器创新方案, 端侧算力协处理器系列芯片内置大算力 NPU 和高带宽嵌入式 DRAM, 能够更好地满足端侧模型部署的算力、存力、运力三者动态平衡需求。
- **存储: 板块本身为强周期板块, 本轮周期自 25Q2 持续上行, 有望持续至 26 年全年。**参考 DRAM 指数, 2025 年 9-11 月合计上涨 101%, 而 NAND 指数同期亦上涨 79%。目前各 CSP 厂为保证服务器的正常出货, 均加大对存储产品的采购, 且由于其对存储价格不敏感, 推动存储价格涨幅持续超预期, 造就“超级周期”。25Q3 前五大存储厂企业级存储营收合计逾 65.4 亿美元, 环比增长 28%。中国字节、阿里等服务器厂商亦持续上修其 AI Capex 预期, 带动企业级存储需求提升。在此趋势下, 国产厂商积极布局企业级存储, 受益 (1) 国产 CSP 厂商资本开支上修 (2) 国产企业级存储份额提升双重逻辑。
- **模拟: 板块汽车需求持续增长, 但今年预计仍有降价压力。**工业领域去库结束, 目前处于复苏初期。同时, 我们认为中美围绕关税的博弈、对抗、和解或为 26 年的主旋律。后续关注成熟制程反倾销等政策的落地, 有望极大改善模拟行业供给格局。26 年伴随 AI 应用落地, 我们认为围绕 AI 诞生的新兴模拟料号将有广阔发展机遇。核心看好, Drmos、微泵液冷等模拟芯片发展机遇。
- **晶圆制造: 晶圆制造资本开支迈入新台阶, 先进制程与封装加速突围。**展望 2026 年, 国内 Fab 厂将迎来存储与先进逻辑的双重“扩产大年”, 强力支撑晶圆代工景气度维持高位。在此背景下, 半导体设备板块有望演绎“ $\beta + \alpha$ ”共振行情, 建议关注受益于扩产红利的行业龙头及精智达等具备技术兑现逻辑的成长标的。与此同时, 在 AI 算力爆发与外部封锁倒逼下, 产业链正加速构建“前道光刻机自主可控+后道先进封装性能跃升”的双轮驱动格局: 以芯上微装、茂莱光学为代表的企业在光刻机整机及核心零部件领域加速破局, 而盛合晶微等厂商则依托 2.5D/3D 先进封装技术筑牢算力底座, 共同重构国产高性能芯片供应链的坚实壁垒。

行业走势



相关研究

《Credo 营收超预期&Marvell 收购 Celestial AI, 催化光铜走强》

2025-12-08

《阿里 Q3 验证 AI 需求高景气, DeepSeek 与谷歌 TPU 引领软硬件进阶》

2025-11-30

■ 消费电子：AI 驱动终端交互变革，手机换机与 AR 新品爆发共振。1)

AI 手机方面，端侧 AI 正从单点助手向跨应用操作系统的 OS Agent 形态跃迁，手机作为算、存、交互兼备的成熟载体将率先受益；其中苹果有望通过 OS Agent 升级引领体验变革，强力驱动 iPhone 15 Pro 以前的存量用户开启换机周期，重点看好立讯精密、蓝思科技等果链龙头及折叠屏相关标的。2) AR 眼镜方面，AR 眼镜产业正迈向关键拐点，2026 年预计将成为 AI 智能眼镜放量元年与 AR 产品力质变之年，随着 Meta、苹果、三星等巨头新品密集发布，Micro-LED、光波导及 SiC 材料等光学显示增量赛道将迎来明确的投资机遇。

■ PCB/CCL：AI 算力驱动 PCB/CCL 市场，架构升级和 M9 材料应用市场规模量价齐升。全球云厂商进入资本开支扩张期，谷歌、亚马逊、微软等海外巨头正采取“前置建设算力”战略，通过创纪录的资本投入驱动 Gemini 3 等领先模型的爆发式增长。PCB/CCL 方面，英伟达机柜架构从 GB200/300 迈向 Rubin/Kyber，通过引入正交背板以及 M9/PTFE 等极低损耗材料，使单机柜 PCB 总价值量实现成倍增长，推动 2026 年 AI PCB 市场规模迈向 600 亿元，同比增长 229.8%；ASIC 方面，谷歌 TPUv8 或将采用 M9 材料，单板价值量约是上代的 1.5 倍，成为 AI PCB 市场另一主要增长来源。上游核心材料市场亦迎来爆发，为满足 224Gbps 的超高速率需求，石英布需求激增，预计到 2027 年石英布市场需求将达约 99 亿元，同比 2026 年 30 亿元增长 230%。同时 HVLP4 铜箔供需结构性短缺，加速国内厂商的进口替代进程，共同驱动 CCL 产业链实现量价齐升。

■ 相关标的：

云端算力芯片：寒武纪、盛科通信、澜起科技、海光信息、中兴通讯等。

端侧算力芯片：晶晨股份、瑞芯微、恒玄科技、乐鑫科技等。

PHY&SerDes 芯片：龙迅股份、裕太微等。

存储：兆易创新、江波龙、德明利、佰维存储、普冉股份、香农芯创。

模拟：思瑞浦、圣邦股份、杰华特、帝奥微、雅创电子。

设备：精智达、中科飞测、芯源微、北方华创、中微公司、华海清科、盛美上海。

代工：中芯国际、华虹半导体、晶合集成。

光刻机和封测：茂莱光学、福光股份、福晶科技、汇成真空、阿石创、强力新材。

消费电子：立讯精密、领益智造、蓝思科技、信维通信、天岳先进等。

PCB/CCL：胜宏科技、沪电股份、深南电路、景旺电子、生益科技、南亚新材、生益电子、威尔高等。

铜缆/铜连接：沃尔核材、兆龙互连、华丰科技、立讯精密、鸿腾精密等。

光芯片/光器件：长芯博创、源杰科技、仕佳光子、太辰光、长光华芯等。

液冷：英维克、思泉新材、申菱环境、高澜股份等。

服务器代工：工业富联、华勤技术。

■ 风险提示：AI 算力基础设施建设放缓，行业竞争加剧与价格战风险，上游原材料价格波动风险，汇率波动风险，AI 创新与应用落地不及预期。

内容目录

1. 算力芯片	6
1.1. 云端算力芯片	6
1.1.1. GPU & ASIC: AI CapEx 大力投入, 看好国产算力业绩释放	6
1.1.2. Switch 芯片: 超节点产业趋势下, 看好国产厂商逆袭	6
1.1.3. 华为 Ascend: 全栈协同构筑国产 AI 算力底座, 引领核心基础设施价值重估	7
1.2. 端侧算力芯片	10
1.2.1. 海外大厂 AI 端侧布局下, 看好海外链 SoC 龙头受益	10
1.2.2. 端侧独立 NPU 元年, 看好泛 AIoT 市场加速 AI 渗透	11
1.2.3. PHY&SerDes 芯片: 国产替代细分赛道, 诞生“小而美”优质标的	12
2. 存储: 企业级存储发力, 26 年迎接“超级周期”	13
2.1. 本轮周期有何不同?	13
2.2. 国产厂商机遇何在?	15
3. 模拟: 关注供需格局优化, AI 新品持续拓展	16
3.1. 供需格局持续优化, 价格战压制逐渐解除	16
3.2. 2026 年迎业绩拐点, 新品类再填动力	16
4. 晶圆制造: 资本开支迈入新台阶, 先进制程与封装加速突围	18
4.1. 半导体设备: Fab 厂资本开支增速显著, 设备享受高景气	18
4.2. 代工: 先进逻辑扩产上一台阶, 晶圆代工景气维持	19
4.3. 光刻机和先进封装: 国产 AI 芯片供应链的重构与崛起	19
4.3.1. 光刻机: 自主可控迫在眉睫, 整机与零部件加速突围	20
4.3.2. 先进封装: 后摩尔时代的算力基石, 国产封测迈向高端化	24
5. 消费电子: 科技巨头引领端侧创新大年	27
5.1. 端侧交互范式重构: 模型能力跃迁, 统一入口加速成形	27
5.1.1. 端侧模型进展: 从技术演进到生态加速	27
5.1.2. 端侧 AI 统一入口: 从单点助手到跨应用操作系统	30
5.2. 手机: OS Agent 启动, iPhone 先行	31
5.2.1. 苹果端侧智能生态加速成型	31
5.2.2. OS Agent 有望掀起存量换机浪潮	32
5.2.3. 硬件创新与 AI 迭代同步, 核心硬件更新公司迎来量价齐升	34
5.3. 硬件形态更新: 眼镜是大厂必争之地	36
5.3.1. 大厂的硬件布局, 各类产品之中聚焦眼镜	36
5.3.2. 明年 AI 眼镜增加显示有望带动场景升级打开出货量空间	36
5.3.3. AI 眼镜核心硬件 光学、SoC 等	37
6. PCB: AI 算力基建加码, CAPEX 超级周期驱动 PCB 量价齐升	39
6.1. 全球资本开支共振架构升级, AI PCB 迎来量价齐升新周期	39
6.1.1. 全球 AI 巨头资本开支高增, 算力“前置建设”引爆基础设施超级周期	39
6.1.2. 机柜架构升级带动 PCB 价值量成倍增长, 2026 年市场规模迈向 600 亿元	40
6.2. CCL&上游材料: AI PCB 驱动 CCL 行业量价齐升, 上游材料供不应求	42
6.2.1. M9: Rubin/Kyber 驱动 M9 材料升级, 224G 速率开启 PCB 极低损耗时代	42
6.2.2. 铜箔、玻布: 石英布与 HVLP4 铜箔市场需求持续旺盛, 国产替代加速	44
7. 风险提示	46

图表目录

图 1:	23Q1-25Q3 海外头部 CSP 资本支出（亿美元）	6
图 2:	23Q1-25Q3 国内头部 CSP 资本支出（亿元）	6
图 3:	华为 384 超节点	7
图 4:	中科曙光 scaleX640 超节点	7
图 5:	华为 Ascend 芯片	8
图 6:	华为 Atlas 900 A3 SuperPoD	9
图 7:	华为 Atlas 950 Super PoD	9
图 8:	基于灵衢的超节点参考架构	10
图 9:	Meta 与 Ray-ban 合作开发 AI 眼镜	11
图 10:	Gemini 端侧机器人模型性能大幅领先	11
图 11:	瑞芯微协处理器 RK1820 结构图	12
图 12:	瑞芯微协处理器产品规划	12
图 13:	中国乘用车摄像头安装情况	13
图 14:	存储原厂季度营收变化（亿美元）	14
图 15:	DXI 指数	14
图 16:	DRAM 与 NAND 指数	14
图 17:	SSD 现货价趋势（美元）	15
图 18:	UFS 现货价趋势（美元）	15
图 19:	25Q3 前五大品牌厂商营收排名	16
图 20:	2024 年中国企业级固态硬盘市场份额	16
图 21:	H100 中 Drmos 的分布	17
图 22:	南芯科技：自主研发的 190Vpp 压电微泵液冷驱动芯片 SC3601	18
图 23:	ASML 分地区营收占比	20
图 24:	海外对华半导体管制措施	20
图 25:	上海微电子 SSX600 系列光刻机	21
图 26:	蔡司主要 EUV 光学部件	22
图 27:	茂莱光学生产的光学器件及光学系统	22
图 28:	汇成真空电子束蒸发高精密度光学镀膜机	23
图 29:	福晶科技生产的晶体、光学元件及激光器件	24
图 30:	芯粒多芯片集成封装的部分代表性技术平台和芯片产品	24
图 31:	2024 年全球前十大封测企业排名	25
图 32:	2019-2029 年中国大陆集成电路封测行业市场规模	25
图 33:	盛合晶微 2.5D 封装平台	26
图 34:	CoWoS、CoPoS、CoWoP 封装类型的比较	27
图 35:	2025 年端侧大模型发布里程碑事件梳理	28
图 36:	内存读取所需功耗远高于计算	29
图 37:	多重模型压缩技术协同驱动端侧模型高效落地	29
图 38:	AI 赋能下，同质化产品/服务可实现差异化体验	29
图 39:	多方玩家布局端侧 AI 统一入口	31
图 40:	苹果 iOS 26 AI 功能梳理	32
图 41:	下一代 Apple Intelligence 赋予屏幕感知能力	32

图 42:	下一代将具有跨应用执行能力.....	32
图 43:	Apple 丰富的软件生态是系统级 AI 的基础.....	33
图 44:	A 系列芯片 NPU 算力 (TOPS)	34
图 45:	M1 芯片 NPU 能力 (TOPS)	34
图 46:	折叠屏在苹果以外的高端手机渗透率测算.....	35
图 47:	折叠屏的核心成本提升在屏幕和铰链.....	35
图 48:	AI 眼镜催化一张图	36
图 49:	2023 年和 2024 年 AR 眼镜的使用场景	37
图 50:	Google Project Astra 展示的 AR 场景	37
图 51:	AR 光学显示方案构成	38
图 52:	AR 各光机方案对比	38
图 53:	几何阵列光波导原理图.....	39
图 54:	表面浮雕衍射光波导原理图.....	39
图 55:	三种材料的折射率比较.....	39
图 56:	三种材料的重量比较.....	39
图 57:	海外云厂商 CAPEX 情况 (亿美元)	40
图 58:	AI PCB 市场规模测算.....	42
图 59:	英伟达 Kyber 机柜正交背板.....	43
图 60:	各厂商材料型号情况.....	43
图 61:	石英布市场空间弹性测算.....	44

表 1:	FT 高速测试机对比.....	19
------	-----------------	----

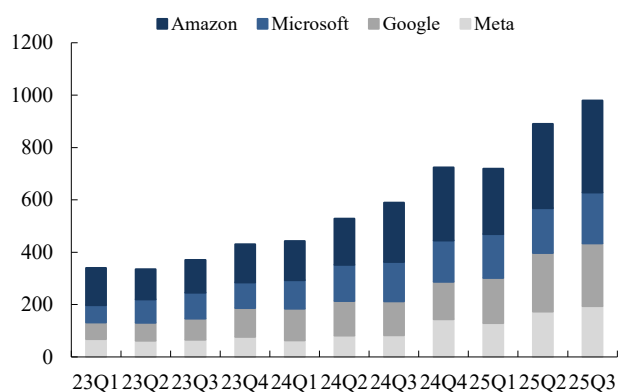
1. 算力芯片

1.1. 云端算力芯片

1.1.1. GPU & ASIC: AI CapEx 大力投入，看好国产算力业绩释放

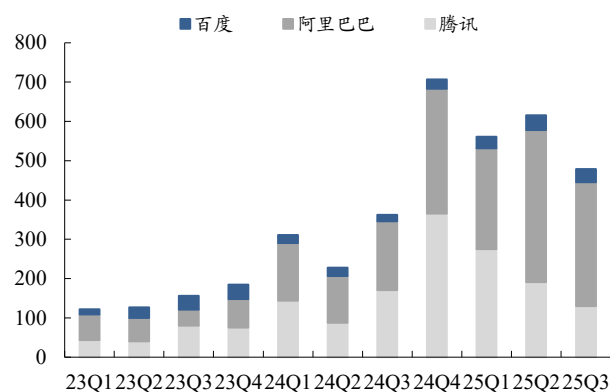
全球 CSP 资本开支持续抬升，国内需求空间广阔。随着 Token 调用量持续攀升，在全球 AI 推理需求快速增长的驱动下，海外 CSP 进一步加大算力基础设施投入，AI 推理相关资本开支持续上行。25Q3 海外四大 CSP 的资本开支合计 979 亿美元，qoq+10%，延续了季度间上升的趋势。相较之下，国内整体算力资本开支仍处于追赶阶段。尽管国内 CSP 与海外巨头在投入总量上仍存在一定差距，但从 Token 调用量与业务规模来看，字节等头部厂商已接近谷歌体量。我们认为国内需求端的算力消耗快速增长，而当前供给侧投入相对不足，为未来数年资本开支大幅提升奠定了扎实空间。

图1：23Q1-25Q3 海外头部 CSP 资本支出（亿美元）



数据来源：Bloomberg，东吴证券研究所

图2：23Q1-25Q3 国内头部 CSP 资本支出（亿元）



数据来源：Bloomberg，Wind，东吴证券研究所

算力产业链自主化进程加速，看好国产算力业绩释放。中芯国际先进制程扩产稳步推进，公司近几年产能利用率维持在 90%左右，表明订单需求与扩产节奏匹配度较高，市场开拓进度良好。整体来看，国产先进制程扩产稳步推进叠加产业链自主可控进展加速，将显著增强国内算力产业的供给保障能力。在 AI 推理和训练需求持续提升的背景下，国产算力厂商有望充分受益，业绩释放可期。**建议关注：寒武纪、海光信息等。**

1.1.2. Switch 芯片：超节点产业趋势下，看好国产厂商逆袭

推理时代性能瓶颈转移：从“算力”到“运力”。小批量、低精度、MoE 的推理任务对计算本身要求不高，但在分布式部署下，多卡之间的通信延迟往往成为效率瓶颈。为此，行业正在把互联方式从机柜之间的 Scale-Out 网络转向机柜内部的 Scale-Up 网络（NVLink、UALink、PCIe 等），利用更短的传输距离实现更高带宽和更低延迟，从而提升整体吞吐。在 Scale-Up 网络架构中，AI 服务器内部使用的 switch 芯片大致可分 CPU-

GPU 连接与 GPU-GPU 互联两类。前者主要负责把 CPU 的有限上行通道扩展出来，连接到多块 GPU、网卡以及存储设备。GPU-GPU switch 芯片则直接服务于 GPU 之间的大规模数据面通信，是 GPU 集群实现高带宽、低延迟的核心。

国产自研超节点崭露头角，“运力”成为国产算力“赶超”海外龙头的关键一环。25 年下半年是国产超节点方案陆续进入公众视野的阶段，无论是互联网厂商、交换机厂商、GPU 自研厂商，各家均有亮眼产品陆续发布。我们看好接下来国产方案百花齐放，一方面，以华为、曙光为代表的全栈自研路径已有重磅方案发布，另一方面，看好第三方 Switch 芯片厂商绑定互联网大厂客户做进终端方案。核心在于产品顺利研发推进，以及导入验证节奏。随着国产算力逐渐进入放量期，国产超节点产业链有望迎来更高确定性的增长机遇。建议关注：盛科通信、万通发展、澜起科技、海光信息、中兴通讯等。

图3：华为 384 超节点



数据来源：新浪财经，东吴证券研究所

图4：中科曙光 scaleX640 超节点








数据来源：中科曙光公众号，东吴证券研究所

1.1.3. 华为 Ascend：全栈协同构筑国产 AI 算力底座，引领核心基础设施价值重估

华为 Ascend 系列处理器持续迭代，致力于提供行业领先的 AI 算力。当前路线图显示，从 2025 年第一季度的 Ascend 910C 开始，该系列处理器在微架构、算力和互联带宽上实现跨越式升级。作为迭代主力，Ascend 950PR（2026 年 Q1 出货）和 Ascend 950DT（2026 年 Q4 出货）均采用 SIMD/SIMT 微架构，并将互联带宽从 910C 的 784 GB/s 大幅提升至 2 TB/s，算力提升至 1 PFLOPS FP8/FP4。在内存方面，Ascend 950DT 配置 144 GB 内存和 4 TB/s 带宽，并开始导入 HiZQ 2.0 自研 HBM 技术。随后的 Ascend 960（2027 年 Q4 出货）进一步将算力翻倍至 2 PFLOPS FP8/4 PFLOPS FP4，互联带宽达到 2.2 TB/s，内存容量和带宽也大幅提升至 288 GB 和 9.6 TB/s。到 Ascend 970（2028 年 Q4 出货），算力再次翻倍至 4 PFLOPS FP8/8 PFLOPS FP4，互联带宽达到 4 TB/s。

图5：华为 Ascend 芯片

					
	Ascend 910C	Ascend 950PR	Ascend 950DT	Ascend 960	Ascend 970
	2025 Q1	2026 Q1	2026 Q4	2027 Q4	2028 Q4
Microarchitecture 微架构	SIMD	SIMD/SIMT		SIMD/SIMT	SIMD/SIMT
Data formats 数值类型	FP32/HF32/FP16/ BF16/INT8	FP32/HF32/FP16/BF16/FP8/ /MXFP8/HiF8/MXFP4		FP32/HF32/FP16/BF16/FP8/ MXFP8 HiF8/MXFP4/HiF4	FP32/HF32/FP16/BF16/FP8/ MXFP8/HiF8/MXFP4/HiF4
Interconnect bandwidth 互联带宽	784 GB/s	2 TB/s		2.2 TB/s	4 TB/s
Computing power 算力	800 TFLOPS FP16	1 PFLOPS FP8, 2 PFLOPS FP4		2 PFLOPS FP8, 4 PFLOPS FP4	4 PFLOPS FP8, 8 PFLOPS FP4
Memory 内存	128 GB, 3.2 TB/s	Ascend 950DT: 144 GB, 4 TB/s Ascend 950PR: 128 GB, 1.6 TB/s		288 GB, 9.6 TB/s	288 GB, 14.4TB/s
Proprietary HBM 自研HBM		HiBL 1.0	HiZQ 2.0		

数据来源：新智元，东吴证券研究所

随着大模型参数规模持续指数级增长，传统单柜算力已难以满足训练推理效率要求。为实现算力集约化、降低通信损耗并提升集群利用率，华为率先提出“超节点”战略，通过将多颗 Ascend NPU 和鲲鹏 CPU 高效连接，使集群在调度逻辑上如同一台超级计算机运行。这一模式显著改善了 AI 计算资源碎片化问题，并成为 Cloud Matrix 战略落地的核心形态。

当前华为主推的 Ascend 384 超节点（Atlas 900 A3 SuperPoD）基于 Matrix Link 高速互联架构，实现 384 颗 NPU 与 192 颗 CPU 全对等通信。该架构支持“一卡一专家”MoE 推理模式，使单卡吞吐量达到 2300 tokens/s，同时通过训推一体的调度机制，使算力利用率提升超过 50%。截至 2025 年 9 月，该型超节点已部署超 300 套，服务超过 1300 家客户，并可按板卡级扩展至 10 万卡规模的 Atlas 900 Super Cluster。与英伟达 GB200 NVL72 机柜相比，其总算力接近 2 倍，总内存容量和内存带宽分别达到 3.6 倍和 2.1 倍。

在此基础上，华为面向下一代 AGI 算力需求推出 Atlas 950 SuperPoD。该产品通过器件级升级、晶间互联技术迭代和光电融合方案，引入递归直连拓扑（UB-Mesh），实现从单板到机架间 NPU “无收敛全互联”，可按 64 卡步长线性扩展至 8192 卡，并结合零线缆电互联与液冷浮动盲插技术显著提升系统可靠性和可维护性。Atlas 950 Super Cluster 由 64 个超节点构成，总计搭载超过 52 万片 Ascend 950DT 芯片，FP8 算力可达 524 EFLOPS，节点规模超越 xAI Colossus。

图6： 华为 Atlas 900 A3 SuperPoD



数据来源：CSDN，东吴证券研究所

图7： 华为 Atlas 950 Super PoD

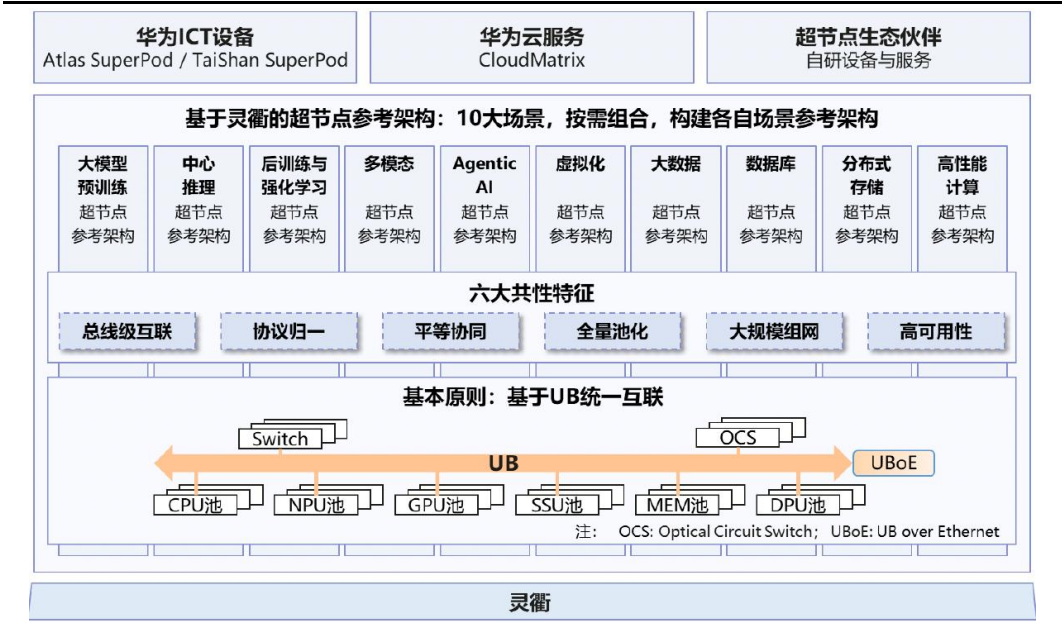


数据来源：华为，东吴证券研究所

硬件演进之外，华为也同步在软件架构层面推出“灵衢”作为 AI 时代数据中心参考设计。灵衢实现 I/O、内存及计算资源在总线级的统一互联（百纳秒级时延、TB/s 带宽），支持 CPU、NPU、GPU 等多类型算力的无中心化协同，并通过全量池化和 UBoE 协议扩展实现百万卡级高效调度。该架构具备高弹性、高可用特性，是 Atlas 超算平台从单柜向大规模矩阵型集群演进的核心软件支撑。

总体来看，华为超节点的迭代，不仅体现硬件算力密度和互联能力的跨越式提升，更结合“灵衢”平台实现软件定义资源、算力智能调度和训推一体化运行，为大参数模型、MoE 模型及 AGI 探索提供面向未来的基础设施底座。这一体系也为其进一步参与全球 AI 基础设施竞争奠定战略优势。

图8：基于灵衢的超节点参考架构



数据来源：《基于灵衢的超节点参考架构白皮书》，东吴证券研究所

在华为昇腾超节点持续推进算力集约化、互联架构不断升级的大背景下，高速连接器正成为 AI 基础设施体系中不可或缺的底层支撑环节。背板连接器作为大型服务器和超算系统的关键组件，承担高速差分信号传输及高电流承载功能，其对传输速率、信号完整性及空间密度要求极为严苛。

华丰科技在高速互联领域具备较强技术储备，是国产替代中最具确定性的连接器龙头之一。公司 112G 高速线模组已实现稳定供货，224G 产品完成客户验证，性能达量产要求，正加速导入华为、浪潮、超聚变、曙光等设备厂及阿里、腾讯、字节等互联网客户。伴随明年新一代算力平台放量、机柜架构复杂度提升，高速线模组单机价值量及单价均有望上升，公司有望实现量价齐升，业绩进入高增周期。结合国产 AI 服务器扩张及华为超节点战略推进，华丰科技有望成为算力互联系统国产化核心受益标的。

1.2. 端侧算力芯片

1.2.1. 海外大厂 AI 端侧布局下，看好海外链 SoC 龙头受益

从海外科技巨头来看，端侧 AI 的战略地位正快速抬升并进入全面落地阶段。2025 年 11 月 18 日，Gemini 3 正式发布，在“人类终极考试”中得分 37.5%，远超前代及竞品。11 月 21 日，Nano Banana 2.0 Pro 随后发布，具备世界知识和逻辑推理能力，能精准识别坐标、时间等信息并生成符合现实场景的图像，还能生成结构化信息图，展现对知识的准确理解和视觉化表达能力。谷歌将模型集成到搜索、Gmail、Android 等核心产

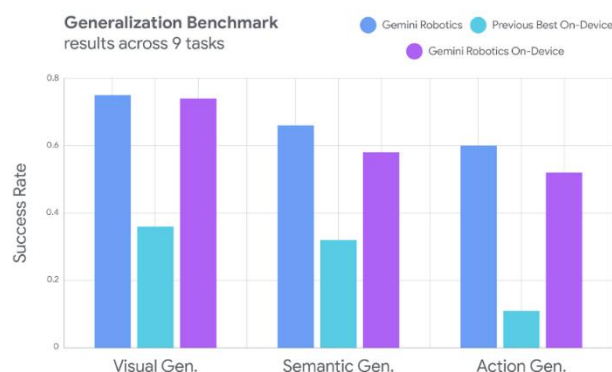
品中，并结合其既有硬件终端布局，形成从云端到端侧的体系化协同。随着模型能力与终端设备的深度融合，其 AI 在端侧的渗透率有望加速提升。**Meta 方面**，管理层在业绩会上将 AI 眼镜视为下一代计算平台核心形态。**OpenAI 也已完成端侧能力的战略补位**：gpt-oss-20B 模型可在 16GB 内存的端侧设备上运行，公司也已正式启动面向消费级 AI 设备的产品化布局。总体来看，端侧 AI 正逐步迈向大规模商业化部署，有望驱动未来 SoC 需求的结构性增长。

图9：Meta 与 Ray-ban 合作开发 AI 眼镜



数据来源：Reuters，东吴证券研究所

图10：Gemini 端侧机器人模型性能大幅领先



数据来源：DeepMind 官网，东吴证券研究所

海外端侧 AI 已进入实质性落地阶段，建议关注海外链 SoC 相关厂商。晶晨股份在谷歌智能家居生态中持续深化绑定关系。根据公司 25Q3 财报，晶晨已与谷歌合作推出多款适配其端侧大模型 Gemini 的新品，包括智能音箱、可视门铃及室内外摄像头等，推动谷歌智能家居全面向嵌入端侧大模型能力的下一代产品升级。**恒玄科技方面**，公司面向智能眼镜场景的 BES6000 系列异构 SoC 正在加速推进；该产品需同时满足高性能与低功耗的系统要求，公司 25Q3 公告显示，BES6000 研发进展顺利，预计 26H1 进入送样阶段。整体看，海外链端侧 AI 需求的确定性增强，有望驱动 SoC 厂商在新一代轻量化大模型终端中实现结构性突破。

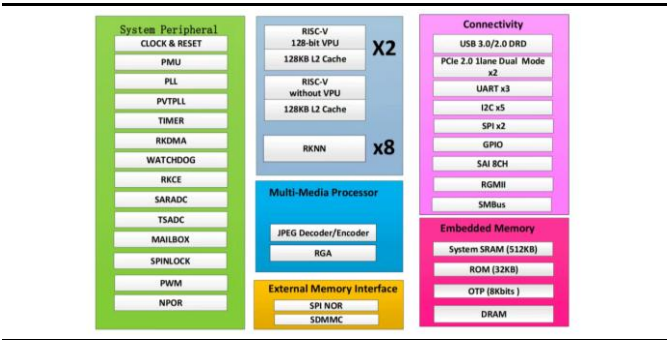
1.2.2. 端侧独立 NPU 元年，看好泛 AIoT 市场加速 AI 渗透

端侧模型升级催生硬件架构向专用协处理器演进。随着 AIGC 和大模型在终端侧的渗透加速，端侧 AI 的算力需求呈现出不同的结构性变化：一方面，模型推理从云侧部分迁移到端侧，带来持续提升的本地计算、带宽与存储需求；另一方面，端侧设备对功耗、成本与空间更为敏感，通用 CPU/GPU 已难以在能效、并行度及实时性之间取得平衡。因此，以瑞芯微为代表的厂商正率先推出面向端侧 AI 的协处理器创新方案，端侧算力协处理器系列芯片内置大算力 NPU 和高带宽嵌入式 DRAM，能够较好地满足端侧模型部署的算力、存力、运力三者动态平衡需求；同时具有较高的灵活性和可扩展性，通过与主处理器的灵活搭配使用，以较小的硬件升级成本满足端侧 AIoT 不同场景的产品

迭代与创新需求，以及相同场景不同定位的产品差异化算力扩展需求，大大节省终端客户的时间成本和开发周期。

瑞芯微：协处理器布局完善、卡位优势明确。在端侧独立 NPU 的产业趋势下，瑞芯微已提前完成端侧协处理器产品线的战略布局，并在性能、生态和产品组合上形成明显卡位优势。公司已推出 RK182X 系列专用协处理器，能够承载 3B–7B LLM 的推理需求，可与主处理器灵活组合，提升终端整机的 AI 性能密度。RK182X 已在车载座舱、智能家居、会议终端、教育设备、机器人、机顶盒和边缘网关等场景中落地。公司同步规划下一代高效能协处理器 RK1860，瞄准更大参数规模与更复杂的多模态场景，进一步提升终端产品的差异化能力。

图 11：瑞芯微协处理器 RK1820 结构图



数据来源：公司官网，东吴证券研究所

图 12：瑞芯微协处理器产品规划

产品	核心参数
RK182X	多核 RISC-V 协处理器，支持 3B/7B 级 LLM/VLM，内置 2.5GB/5GB 超高带宽 DRAM，提供 PCIe2.0/USB3.0/以太网等接口，适用于端侧轻量级大模型推理。
RK1860	下一代协处理器，60–80 TOPS，支持 1.5B–13B 主流端侧大模型，带宽 1TB+，采用 3D 架构与近存计算，接口丰富（PCIe、UCIe），易于搭配主流 AP 做多片联接与算力升级。

数据来源：瑞芯微开发者大会，东吴证券研究所

1.2.3. PHY&SerDes 芯片：国产替代细分赛道，诞生“小而美”优质标的

国产整车厂正在促成汽车产业智能化渗透。随着汽车产业智能化、电动化与网联化加速演进，汽车传感器作为感知系统的核心部件，正成为推动技术变革与产业升级的关键要素。据佐思汽车研究，25H1 中国乘用车整体摄像头安装量达 5239.6 万颗，同比增长 34.9%，主要受比亚迪、小米汽车、小鹏拉动；其中新势力摄像头安装量最高，达 3072.7 万颗。

看好 26 年国产车载 SerDes 进入导入放量期。车载 SerDes 用于大量高清图像和数据的高速传输，凭借低延迟和高带宽特性成为 ADAS 和智能驾驶系统中不可或缺的组件。据 QYResearch，2023 年 ADI（Maxim）和 TI 占有全球大约 92% 的市场份额。我们统计国产参与者包括一些通信 IC 和模拟 IC 厂商，部分领先厂商的产品经过前期的研发积累已逐步进入客户导入阶段。**龙迅股份**针对汽车市场对于视频长距离传输和超高清视频显示需求开发的车载 SerDes 芯片组目前处于全面市场推广阶段，公司把握新兴市场发展机遇，已将 SerDes 芯片组拓展进入 eBike、摄像云台、无人机等新业务领域，我们看好公司 2026 年持续取得突破。**裕太微**具备 MIPI-APHY 和 HSMT 两套协议开发能力和经验，自研基于 HSMT 协议的车载高速视频传输芯片预计于 2025 年年底前问世，可支持 200 万至 800 万像素摄像头的图像传输需求。

图13：中国乘用车摄像头安装情况

2025年H1中国新车前视摄像头安装量为981.3万颗，同比增长38.1%，主要受比亚迪、小米汽车拉动；周视安装量达976.6万颗，同比增长106.6%；环视安装量为2341.6万颗，同比增长29.3%；其他摄像头看，夜视摄像头安装量同比增长最快，主要受方程豹、吉利银河拉动。

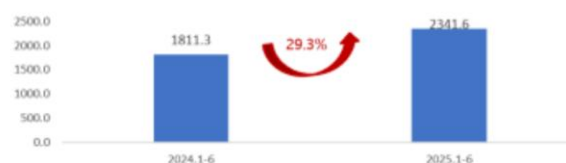
图：2025年中国乘用车前视摄像头安装量（万颗）及同比



图：2025年中国乘用车周视摄像头安装量（万颗）及同比



图：2025年中国乘用车环视摄像头安装量（万颗）及同比



表：2025年中国乘用车其他摄像头安装量（万颗）及同比

其它摄像头安装量	2024.1-6	2025.1-6	同比
座舱	248.0	323.0	30.2%
夜视	0.2	1.2	386.5%
行车	175.3	163.1	-6.9%
倒车	428.4	383.5	-10.5%
流媒体	23.6	44.1	86.7%
盲点	8.5	4.8	-43.6%
AR实景导航	0.9	1.0	10.6%
电子外后视镜	0.0	0.5	-
其他	4.7	19.0	306.8%

数据来源：佐思汽车研究，东吴证券研究所

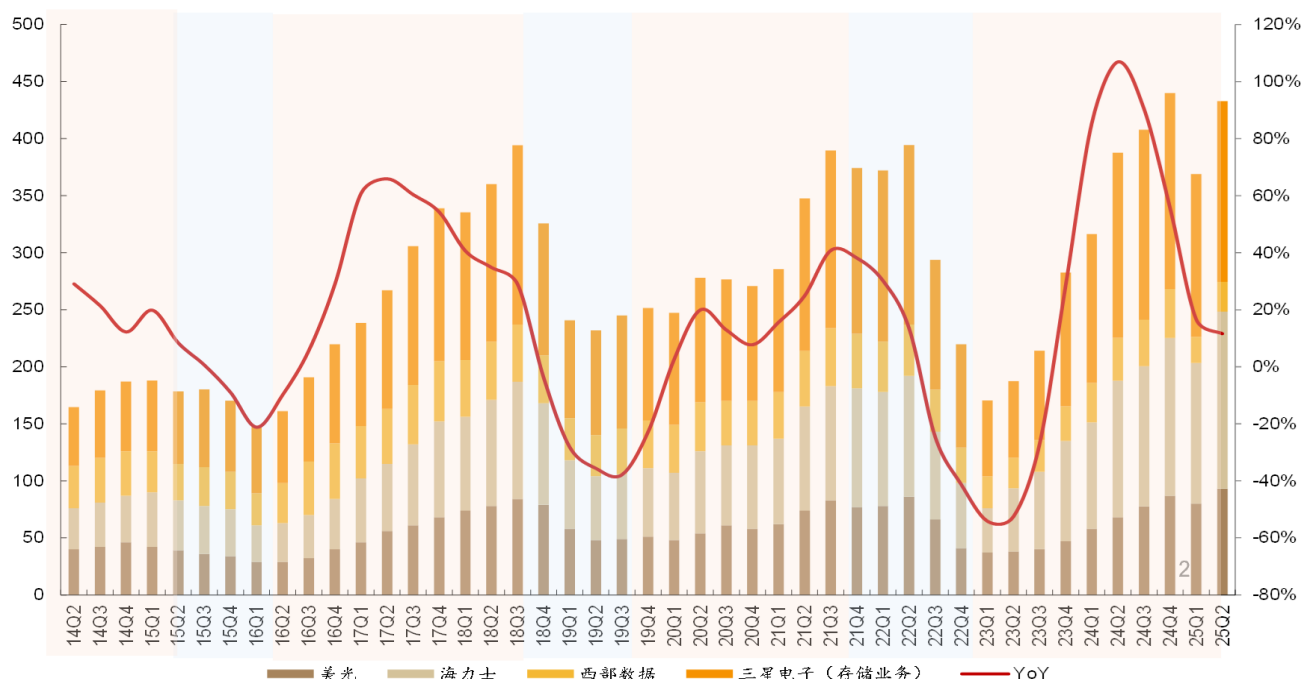
注：左下“图：2025年中国乘用车周视摄像头安装量（万颗）及同比”中蓝色柱体为24H1，橙色柱体为25H1

2. 存储：企业级存储发力，26年迎接“超级周期”

2.1. 本轮周期有何不同？

存储板块本身为强周期板块，本轮周期自25Q2持续上行，有望持续至26年全年。存储由于产品标准化程度高，大宗属性强，且供给格局集中。因此具较强周期性。本轮周期自25Q2开始上行，9月之前更多为转产导致DD4涨价，9月之后由于HDD供给短缺，AI推理带来的存储需求快速导入SSD，推动SSD价格快速上涨。由于存储供给短缺可能导致AI服务器出货下修，因此各下游AI客户对存储价格上涨不敏感，推动存储价格涨幅持续超预期，造就“超级周期”。

图14: 存储原厂季度营收变化 (亿美元)

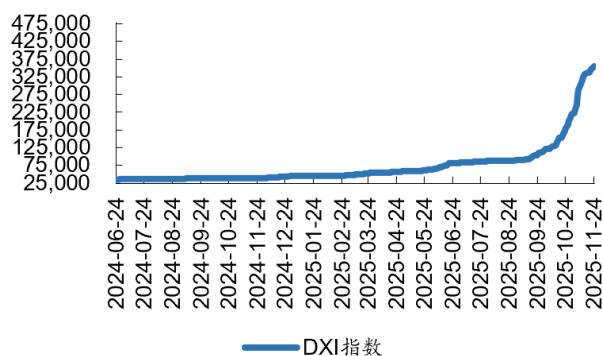


数据来源: iFinD、各公司公告、东吴证券研究所

注: 图中 YoY 为“原厂合计季度营收增速”

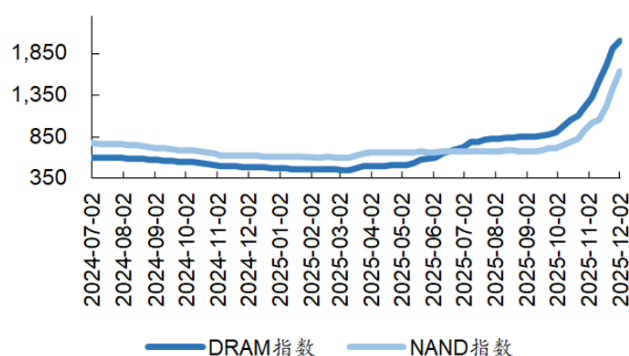
Q3 存储现货价格持续上涨，中国台湾存储厂指引乐观。参考 DRAM 指数，其 25 年 9-11 月期间合计上涨 101%，而 NAND 指数同期亦上涨 79%。存储模组大厂威刚董事长陈立白 11 月 24 日指出，AI 需求存储器的强度远超过历史上任何时期，DRAM、NAND 快闪存储器双缺持续升温，缺货深度与时间皆创 20 年来最严重，合约价将持续上涨至少两至三季以上，2026 年上半年 DRAM 与 NAND 快闪存储器仍将呈现全面缺货态势。对于市场疑虑客户可能重复下单（double booking）的问题，陈立白认为，各家终端产品大厂包括美国、韩国、中国等，各家董事长、总经理都亲自出马与供应商谈货源，市场完全没有“double booking”的状况，因为“下 2 张订单、3 张订单都不一定能拿到货”，供应商拥有绝对议价权，多数客户连 1 年约都签不到，只能“每月谈”。

图15: DXI 指数



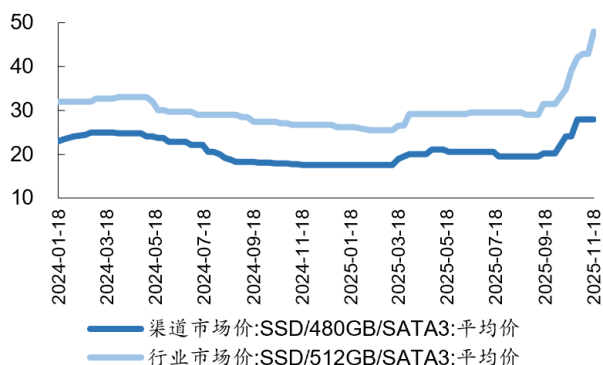
数据来源: iFinD、东吴证券研究所

图16: DRAM 与 NAND 指数



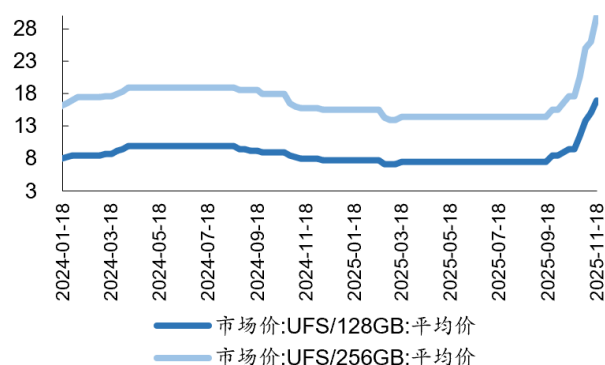
数据来源: iFinD、东吴证券研究所

图17: SSD 现货价趋势 (美元)



数据来源: iFinD、东吴证券研究所

图18: UFS 现货价趋势 (美元)



数据来源: iFinD、东吴证券研究所

2.2. 国产厂商机遇何在？

在此趋势下，国产厂商积极布局企业级存储，受益（1）国产 CSP 厂商资本开支上修（2）国产企业级存储份额提升双重逻辑。

25Q3 企业级 SSD 市场迎来显著成长。根据 TrendForce 集邦咨询最新调查，受益于 AI 需求快速从训练端外溢至推理端，以及北美云端服务厂商同步扩张 AI 基础设施与通用型 Server 建设，25Q3 Enterprise SSD 出货量与价格强势上扬，前五大品牌厂合计营收环比+28%，达 65.4 亿美元，创今年新高。TrendForce 集邦咨询表示，25Q4 市场氛围从“复苏”转向“抢料”，主要因 NAND Flash 供应商有鉴于过往的市场波动，对产能扩张采取审慎态度，将导致 Enterprise SSD 产能增长幅度远低于需求增速。为避免 SSD 短缺延误昂贵的 AI Server 建置进度，CSP 买家近期积极建立库存。在供不应求的卖方格局下，预计 25Q4 Enterprise SSD 平均合约价将季增逾 25%，产业营收有望再创新高。

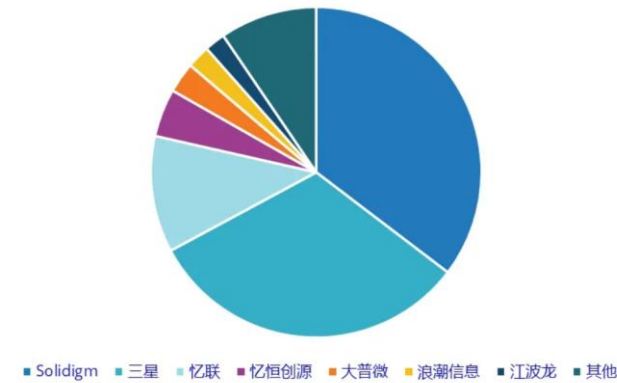
同时，政策推动国产化率提升，本土存储企业通过大容量产品和主控芯片等技术创新，以及产业链和服务本地化，缩小与国际品牌的差距，在份额上有较大幅度的提升。

图19：25Q3 前五大品牌厂商营收排名

Rankings	Company	Revenue (US\$M)		Market Share (%)	
		3Q25	QoQ (%)	3Q25	2Q25
1	Samsung	2,441.9	28.6%	35.1%	34.6%
2	SK Group (SK hynix + Solidigm)	1,861.0	27.3%	26.8%	26.7%
3	Micron	991.0	26.3%	14.3%	14.3%
4	Kioxia	978.4	30.4%	14.1%	13.7%
5	SanDisk	269.0	26.3%	3.9%	3.9%
Total of Top 5		6,541.3	28.0%	94.2%	95.1%

数据来源：Trendforce、东吴证券研究所

图20：2024 年中国企业级固态硬盘市场份额



数据来源：IDC、东吴证券研究所

3. 模拟：关注供需格局优化，AI 新品持续拓展

3.1. 供需格局持续优化，价格战压制逐渐解除

需求端：汽车持续增长，工业去库结束。模拟板块汽车需求持续增长，但今年预计仍有降价压力。工业领域去库结束，目前处于复苏初期。通信领域（如基站）经历 22-24 年去库，25 年开始有补库需求，26 年持续成长。除汽车外，各领域、各料号产品线价格今年预计企稳。部分料号交期变长。

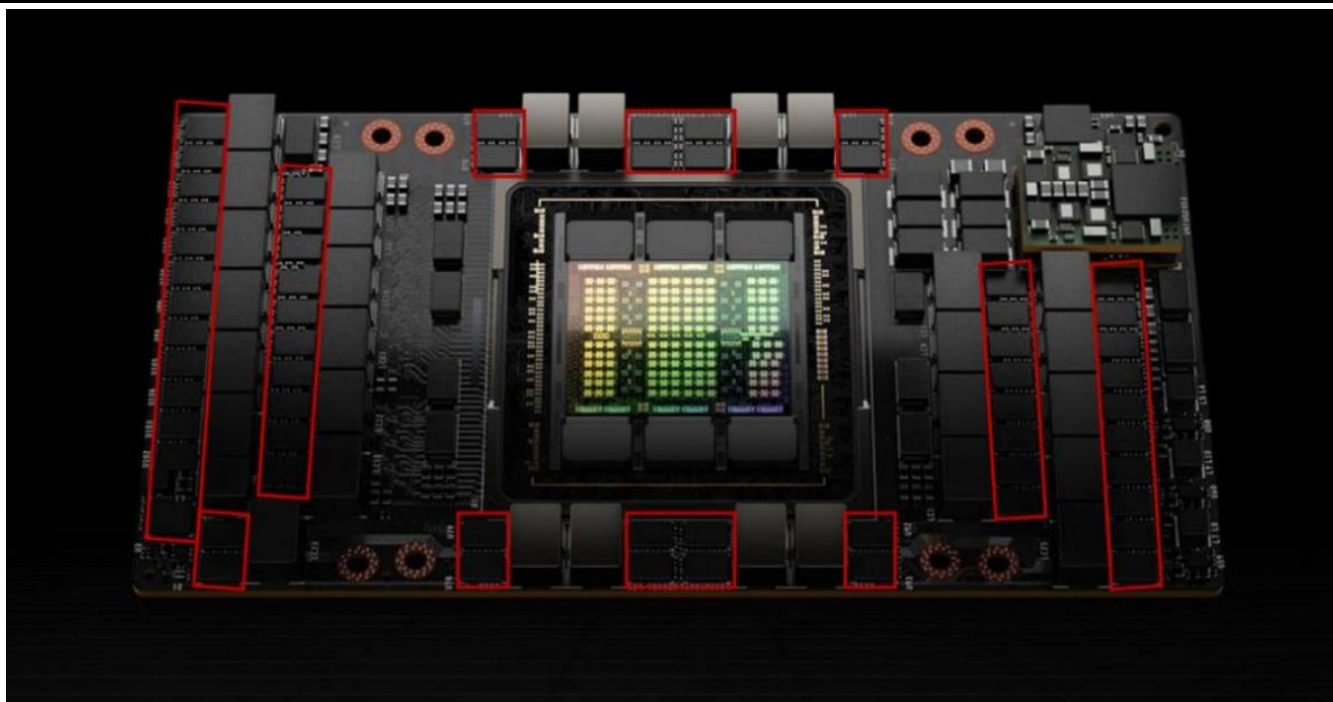
同时，我们认为中美围绕关税的博弈、对抗、和解或为 26 年的主旋律。后续关注成熟制程反倾销等相关政策，仍有落地空间，极大改善模拟行业供给格局。价格为模拟板块最核心的压制，无论是 TI 价格战趋缓亦或是成熟制程反倾销潜在出台预期，都将极大优化模拟行业生态。25 年模拟需求、供给、库存、价格、竞争格局全面向好，乐观看待模拟行业 25 年复苏趋势。思瑞浦等汽车工业占比高标的有望核心受益。

3.2. 2026 年迎业绩拐点，新品类再填动力

26 年伴随 AI 应用落地，我们认为围绕 AI 诞生的新兴模拟料号将有广阔发展机遇。核心看好，Drmos、微泵液冷等模拟芯片发展机遇。

（1）**Drmos：**跟随国产算力芯片出货&功耗双重提升，26 年进入加速出货阶段。GPU 的功耗提升，对 DrMOS 数量的需求增加。随着大数据、云计算、人工智能的兴起，需要更大算力、更大功耗的 GPU/ASIC 的支持，同时也对负载电源提出了更高的要求。英伟达 A100 功耗为 300W，DrMOS 至少需要 16 颗。H100 功耗为 700W，需要 40~50 颗 DrMOS。若按单颗 1 美金测算，单卡价值量约 300 人民币。B100 功耗为 700W，B200 功耗为 1000W，B300 功耗为 1400W，DrMOS 的需求数量有望继续增加。

图21: H100 中 Drmos 的分布



数据来源：英伟达、东吴证券研究所

(2) **微泵液冷：功耗提升，散热需求增加，端侧主动散热新方案。**AI 技术爆发式迭代的驱动下，算力芯片及超轻薄终端的性能瓶颈日益凸显，液冷散热驱动凭借主动温控强势崛起，通过冷却液循环带走热量，散热效率较被动方案提升 3 倍以上。例如，艾为电子发布的超低功耗高压 180Vpp 压电微泵液冷驱动产品，通过高压 180Vpp 和中低频振动（10~5000HZ）驱动微通道内冷却介质实现超低功耗、超小体积、超高背压流量以及超静音散热。

而从竞争格局看，目前全球仅 Boris、艾为、南芯有所布局，且艾为南芯均在 23 年投入项目研发，今年加大研发重视程度及优先级，竞争格局良好。

图22：南芯科技：自主研发的 190Vpp 压电微泵液冷驱动芯片 SC3601



数据来源：南芯科技、东吴证券研究所

4. 晶圆制造：资本开支迈入新台阶，先进制程与封装加速突围

4.1. 半导体设备：Fab 厂资本开支增速显著，设备享受高景气

两存扩产再上一个台阶，294 层的 NAND 和 DDR5 的 DRAM 都来到了跟国际水平相当的最新产品代际，帮助两存迎来新一轮扩产浪潮。长江存储与长鑫存储 26 年产能持续扩张，在存储需求缺口逐步暴露的背景下，未来产能还有望进一步扩大。

先进逻辑扩产增长势头尤为显著，2026 年有望成为先进制程的“扩产大年”。中芯南方、中芯东方、永芯、华力、ICRD 等核心主体均在加速产能建设，预计明年行业整体扩产规模将实现爆发式增长，有力支撑国产先进制程产业链的快速突围。

成熟类设备公司享受 β 行情。北方华创、中微公司、拓荆科技、华海清科、盛美上海等在各自领域份额相对稳定，将充分享受先进存储和先进逻辑扩产大年的 β 。同时，北创的 CCP、中微的 on-on 和 CUBS，将为两个企业带来额外的增量。

成长类设备公司享受 $\alpha + \beta$ 行情。在优质 β 之下，亦有部分优质企业可以兑现自己的 α ，这也是我们 26 年首推的设备公司：精智达、中科飞测、芯源微，此三家公司将于 26 年分别在存储+逻辑测试机、明场、Track 三个品类上迎来技术兑现和国产突破，其中精智达更早一步已经实现了突破，明年将迎来放量元年。在戴维斯双击的帮助下，以上三家公司的预期涨幅有望大幅跑赢板块。

表1: FT 高速测试机对比

参数项	精智达 FT 测试机	泰瑞达 epic	爱德万 T5503
测试对象	DRAM、HBM	LPDDR5-DRAM、高速 NAND 闪存等	LPDDR5-DRAM、高速 NAND 闪存等
测试速率	9 Gbps	8 Gbps	8 Gbps
通道数	20000-30000	10000-20000	10000-20000
并行测试能力	600-700	300-400	300-400
主要应用领域	高性能存储器的封装后功能测试	存储器件的晶圆和封装测试	存储器件的晶圆和封装测试
市场定位	面向 DRAM 和 HBM 制造商	高端存储器件测试	高端存储器件测试
价格范围	250 万美金	300 万美金	300 万美金

数据来源：公司官网，东吴证券研究所

4.2. 代工：先进逻辑扩产上一台阶，晶圆代工景气维持

先进制程扩产储备丰厚，晶圆代工催化不断。目前国内先进制程尤其是 7nm 及以下供给严重不足，这也是制约两大代工厂该业务放量的主要因素，26 年先进扩产丰厚，晶圆代工景气将继续维持。

一线代工厂核心受益于先进制程扩产红利。展望 2026 年，中芯系体内先进制程产能建设持续推进，叠加前期扩产产能的逐步释放，预计公司先进制程业务营收量级将实现可观增长。此外，随着先进制程良率的稳步提升，相关产线的盈利能力有望改善，并表的少数股东损益表现亦值得期待。

二线代工厂有独立逻辑兑现，同样值得关注。晶合集成为首的二线代工厂 26 年有望实现独立逻辑，55/40nm 预计稼动率饱满，28nm 也将在 26 年开始迎来放量。此外，公司有望与长鑫存储达成合作，从成熟品类的代工做起，逐步渗透到更先进代际产品上，潜在市场空间较大，值得关注。

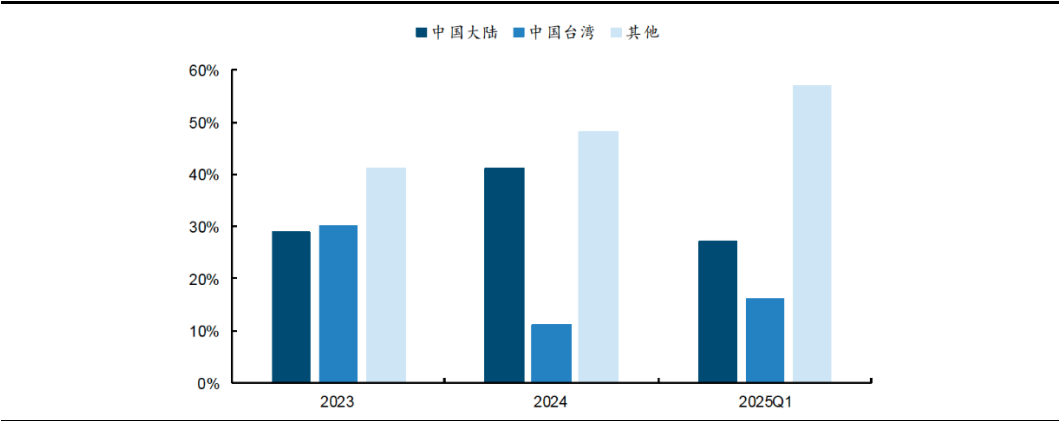
4.3. 光刻机和先进封装：国产 AI 芯片供应链的重构与崛起

在 AI 算力爆发与地缘政治博弈的双重背景下，半导体产业链正加速构建“前道自主可控+后道性能跃升”的双轮驱动格局。一方面，面对美日荷对先进光刻机封锁的不断加码，中国作为全球最大半导体设备市场，正倒逼产业链从整机（如芯上微装、上海微电子）到核心零部件（光源、光学、镀膜等）实现全方位协同突围；另一方面，在后摩尔时代，以 2.5D/3D 为代表的先进封装技术成为延续摩尔定律、提升算力密度的关键基石，全球封测产能持续向亚洲集聚，国内先进封装市场 2024-2029 年 CAGR 达 14.4%，结构性优化趋势显著。两者共同构筑了国产高性能芯片产业化的坚实底座。

4.3.1. 光刻机：自主可控迫在眉睫，整机与零部件加速突围

中国大陆稳居全球最大半导体设备市场，也是 ASML 最核心的客户来源。2023 年中国大陆贡献了 ASML 总营收的 29%；2024 年，受益于晶圆厂积极扩产及超额备货策略，这一比例爆发式增长至 41%。尽管 2025 年一季度受出口管制等外部因素影响，营收占比回落至 27%，但展望后续，随着国内晶圆厂扩产计划的持续推进，中国市场对光刻机的需求预计将依然维持强劲态势。

图23：ASML 分地区营收占比



数据来源：ASML 投资者调研纪要，东吴证券研究所

美日荷持续加码光刻机管制，加速国产替代进程。全球先进半导体设备市场主要被美日荷垄断，光刻机领域更是由 ASML、尼康（Nikon）和佳能（Canon）三巨头统治。由于国内技术相对落后，高精尖设备长期依赖进口。美国自 2018 起便施压限制高端光刻机对华出口，随后相继出台“1007 新规”并联合日荷实施产业封锁，限制先进制程设备流入中国。在封锁不断升级的背景下，突破以光刻机为代表的“卡脖子”关键设备，加速国产化替代已迫在眉睫。

图24：海外对华半导体管制措施

时间	国家	具体管制情况
2022 年 4 月	美国	美国政府提议与韩国、日本、中国台湾地区建立“芯片四方联盟”（CHIP4），试图将中国大陆排除在全球半导体供应链联盟之外。
2022 年 7 月	美国	美国商务部禁止 ASML、应用材料等企业向中国出口 14nm 及以下先进制程设备，并限制美国公民参与中国芯片研发。
2023 年 1 月	美日荷	美日荷秘密达成三方协议，限制向中国出口 DUV 光刻机及零部件。荷兰 ASML 随后停止向中国出口部分 DUV 设备。
2023 年 1 月	日本	日本宣布自 2023 年 7 月起，对 23 种先进半导体制造设备实施出口管制，包括光掩模镀膜设备、光掩模检测设备、光刻步进器以及符合氟化物（ArF）DUV 性能标准或更高水平的扫描仪设备等。
2024 年 9 月	美国	美国商务部更新量子计算与半导体出口管制，限制中国获取先进光刻机。
2025 年 1 月	美国	发布“全球 AI 管控新规”，将全球国家分为三级管控区，中国被列为最高风险等级，全面禁止 AI 芯片和模型对华出口。
2025 年 5 月	美国	BIS 向 Synopsys、Cadence 和西门子 EDA 这三家全球前三的 EDA 厂商发出通知，要求他们停止向中国提供服务。

数据来源：国合中心、路透社，东吴证券研究所

在政策的强力驱动下，国内光刻机产业正加速突围。近年来国产光刻机产业化成果

斐然：2016 年上海微电子 90nm ArF 设备实现出货；2020 年华卓精科打破 ASML 双工件台技术垄断；2025 年哈尔滨工业大学成功研制 13.5nm EUV 光源，中科院上海光机所的全固态深紫外光源系统更是将芯片工艺推向 3 纳米理论极限。总体而言，国内光刻机技术持续进阶，国产替代空间十分广阔。

芯上微装交付第 500 台设备，携手盛合晶微强化 AI 算力底座。2025 年 8 月 8 日，国产步进光刻机领军企业芯上微装举办第 500 台设备交付仪式，充分验证了其自主创新实力。作为公司拳头产品，该款先进封装光刻机凭借高分辨率、高套刻精度及优异的翘曲/厚胶处理能力，完美适配 Flip-chip、Fan-out、2.5D/3D 等技术，2024 年已占据全球 35%、国内 90% 的市场份额。此次交付客户为全球领先的晶圆级先进封装厂商盛合晶微，后者致力于通过异构集成技术提升 GPU、AI 芯片等高性能产品的算力与能效，双方合作将进一步巩固国产 AI 芯片产业链。

图25：上海微电子 SSX600 系列光刻机



数据来源：上海微电子官网，东吴证券研究所

光学系统为光刻机中技术壁垒最高的零部件，也是国产化攻坚中难度最高的“深水区”。不同于 DUV 的透射系统，EUV 光刻机必须采用原子级平整度的离轴反射系统，其镀膜厚度误差需控制在 0.025nm 以内，面型精度误差低于 0.25nm，被誉为“宇宙中最光滑的人造结构”，对真空环境、材料选择及超精密加工工艺提出了极限要求。尽管国内企业如茂莱光学已在 i-line 光刻机光学器件上实现应用突破，但客观来看，在面型精度（PV 值）及表面光洁度等核心指标上，国产技术距离蔡司的 EUV 级水准仍有显著差距。超精密光学部件的国产化任重而道远，极高的技术壁垒也意味着巨大的追赶空间。

图26：蔡司主要 EUV 光学部件

光学元件	作用	图示
投影光学元件	光学 EUV 系统由投影光学元件组成，六个镜子是世界上最精确的反射镜，用于将纳米范围内的掩模结构成像到光刻胶涂层晶圆上。需要大约 20000 个重达 2 吨的单个零件。	
镜块	镜块使晶圆能够与掩模和投影光学器件精确对准，以实现晶圆曝光。尽管晶圆扫描仪存在热和高动态载荷，但镜块几乎保持完美稳定。	

数据来源：蔡司官网，东吴证券研究所

茂莱光学作为精密光学综合解决方案提供商，产品线覆盖深紫外 DUV 至远红外全谱段，广泛应用于半导体光刻与检测、生命科学及 AR/VR 等前沿领域，并已进入 Camtek、KLA 等全球半导体巨头的供应链。尽管全球超精密光学市场长期被蔡司、尼康和佳能垄断，且国内核心零部件尚高度依赖进口，但在国内光刻机市场规模扩张及存量设备维护升级需求迫切的背景下，茂莱光学凭借在双面非球面、异形非球面及离轴非球面透镜等高端产品上的精密制造能力，有望在国产替代的浪潮中加速突围，抢占更多市场份额。

图27：茂莱光学生产的光学器件及光学系统

类别	产品	图示	应用
光学器件	透镜：球面透镜、非球面透镜、柱面镜、胶合透镜等 平片：多光谱滤光片、荧光滤光片、反射镜、相位延迟窗口等 棱镜：胶合棱镜、异形棱镜等		广泛运用于航空航天、生物医疗等光学系统中
光学镜头	显微物镜系列、机器视觉镜头、成像镜头和监测镜头		应用于半导体检测设备、基因测序显微设备和 3D 扫描、光电传感、航天监测及激光雷达等领域
光学系统	激光干涉系统、明场显微系统、高功率 DUV 激光扩束整形系统 光显微系统、体视显微系统、3D 扫描模组、生物识别光学模组、AR/VR 光学测量模组及检测设备		应用于半导体量检测设备中的晶圆三维形貌量测、封装缺陷 2D/3D 检测、晶圆缺陷检测；基因检测、病理检验、眼科手术、模等医疗仪器和设备中；AR/VR 的性能测量

数据来源：公司官网，东吴证券研究所

汇成真空立足于真空镀膜与成膜工艺核心技术，全面掌握蒸发、磁控溅射及离子镀膜等技术体系，产品广泛覆盖消费电子、半导体（含光刻掩模版、传感器）、汽车及精密光学等领域，并为客户提供从生产工艺支持到运维改造的全套增值服务，成功卡位核心客户资源。在此坚实基础基础上，公司进一步向高端前沿领域延伸，于 2025 年 5 月 22 日与浙江大学极端光学技术与仪器全国重点实验室签署战略合作协议，共建“联合实验室”；双方将针对半导体检测用剪切干涉仪、大数值孔径显微物镜及先进制程光刻关键光学元件等

项目展开深度合作。我们认为，这一产学研合作将加速科研成果转化，有力推动波长光电在光刻机及半导体检测核心零部件领域的战略布局。

图28：汇成真空电子束蒸发高精密光学镀膜机



数据来源：公司官网，东吴证券研究所

波长光电为国内精密光学元件及组件的核心供应商，长期深耕工业激光加工与红外热成像领域，产品广泛覆盖激光、红外及消费级光学等赛道，并成功进入华工科技、大族激光、高德红外以及 IPG、FLIR 等国内外行业龙头的供应链。在此坚实业务基础上，公司于 2025 年 5 月 22 日与浙江大学极端光学技术与仪器全国重点实验室签署战略合作协议，共建“联合实验室”，双方将聚焦半导体检测用剪切干涉仪、大数值孔径显微物镜及先进制程光刻关键光学元件等高端项目展开实质性合作；我们认为，这一产学研深度融合将促进科研成果转化，加速公司在光刻机核心零部件领域的战略布局与突破。

福晶科技作为光电元器件龙头，主营业务稳健增长且核心产品市占率领先，2024 年非线性光学晶体、激光晶体及精密光学元件营收分别达 2.35 亿元、1.51 亿元和 3.11 亿元，同比分别增长 14.89%、8.63% 和 24.18%，广泛应用于激光器、光通信、AR/VR 及激光雷达等领域。在夯实主业的同时，公司通过持有至期光子 45.79% 股份积极布局高端赛道，至期光子聚焦纳米及亚纳米级超精密光学元件与复杂光机组件研发，瞄准半导体量测等国家重大战略装备需求，致力于突破关键技术瓶颈，加速提供高品质的高端光学元器件国产替代解决方案。

图29：福晶科技生产的晶体、光学元件及激光器件

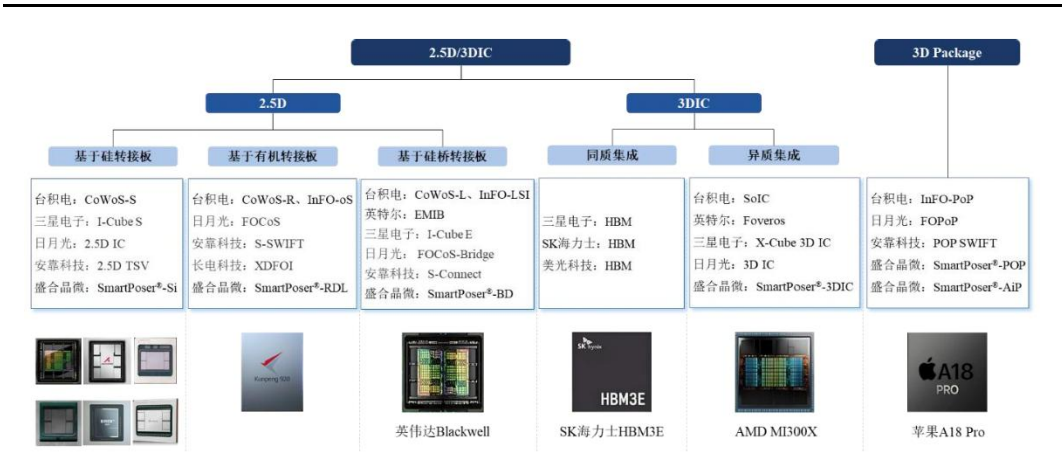
类别	产品	图示	用途
晶体	非线性光学晶体、激光晶体、双折射晶体、磁光晶体、声光及电光晶体、闪烁晶体等		<p>主要用途：作为固体激光器的工作物质、非线性频率转换、磁光材料、电光材料等</p> <p>主要细分市场：固体激光器、光纤激光器等</p>
精密光学元件	非球面透镜、球面透镜、柱面透镜、反射镜、窗口片、棱镜、波片、偏振镜、分光镜、光栅等		<p>主要用途：应用于激光器谐振腔、准直聚焦、光路传输、光束整形、偏振转换、分光合束等</p> <p>主要细分市场：固体激光器、光纤激光器、光通讯、AR/VR、激光雷达、半导体设备、光学检测设备、分析仪器、生命科学等</p>
激光器件	磁光器件、声光器件、电光器件、驱动器、光开关、光学镜头（扫描场镜、扩束镜）、光纤传输器件等		<p>主要用途：光纤与固体激光器的声光调制器、电光调制器、Q开关、隔离器等</p> <p>主要细分市场：固体激光器、光纤激光器、光通讯等</p>

数据来源：公司年报，东吴证券研究所

4.3.2. 先进封装：后摩尔时代的算力基石，国产封测迈向高端化

先进封装技术正向高密度互联演进，主要包括 2.5D 和 3D 集成两大核心路径。2.5D 集成主要利用转接板（分为硅转接板、有机转接板及硅桥转接板三类）作为中介介质，实现多颗芯片的高密度水平互联；而 3D 集成（3DIC）则通过微凸块或混合键合等技术，达成芯片间的垂直堆叠与互联，支持同质与异质芯片的系统级整合。在此基础上发展的 3D Package（新型扇出型封装）进一步融合了重布线与凸块等水平及垂直互联工艺，不仅大幅缩减封装体积，更有效缩短信号路径、降低噪声延迟并优化散热性能，凭借其高集成度与超薄特性，已成为高端消费电子及 5G 毫米波通信等领域的关键技术方案。

图30：芯粒多芯片集成封装的部分代表性技术平台和芯片产品



数据来源：灼识咨询，东吴证券研究所

集成电路封测产能已逐步从欧美向亚洲转移，目前形成了中国大陆、中国台湾与美

国“三足鼎立”的竞争格局。行业集中度较高，2024 年全球前三大封测企业占据了约 50% 的市场份额；其中，中国厂商处于优势地位，在全球前十大封测企业中，中国大陆与中国台湾分别有 4 家和 3 家企业。

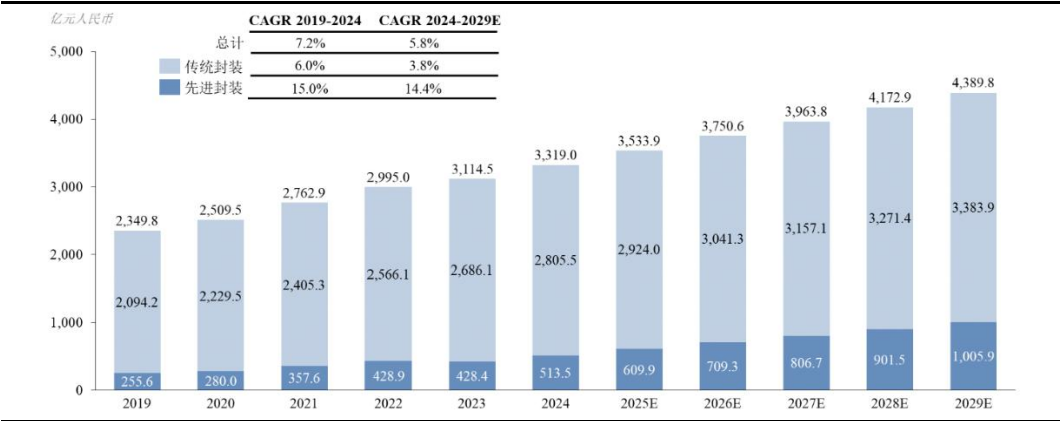
图31：2024 年全球前十大封测企业排名

排名	公司名称	国家/地区	2024 年市场份额	2023 年市场份额	2022 年市场份额
1	日月光	中国台湾	23.7%	24.8%	27.7%
2	安靠科技	美国	15.0%	16.2%	16.0%
3	长电科技	中国大陆	11.3%	10.4%	11.2%
4	通富微电	中国大陆	7.8%	7.8%	7.0%
5	力成科技	中国台湾	4.9%	5.2%	5.2%
6	华天科技	中国大陆	3.8%	3.2%	3.0%
7	京元电子	中国台湾	2.7%	2.6%	2.8%
8	联合科技	新加坡	2.0%	2.0%	1.7%
9	韩亚微	韩国	1.7%	1.4%	1.3%
10	盛合晶微	中国大陆	1.6%	1.1%	0.6%

数据来源：Gartner，东吴证券研究所

中国大陆封测行业呈现出由长电、通富、华天三大巨头领跑的竞争格局，其余新兴企业多深耕细分领域。受益于政策与需求双轮驱动，中国大陆封测市场规模由 2019 年的 2349.8 亿元稳步增长至 2024 年的 3319 亿元（CAGR 为 7.2%）。尽管当前仍以传统封装为主（2024 年先进封装占比约 15.5%），但在全球产业转移背景下，未来市场将持续扩容，预计 2029 年总规模将达 4389.8 亿元。更重要的是，行业结构正加速优化，2024-2029 年先进封装预计将以 14.4% 的复合增速（远超传统封装的 3.8%）成为增长核心引擎，推动其占比在 2029 年提升至 22.9%。

图32：2019-2029 年中国大陆集成电路封测行业市场规模



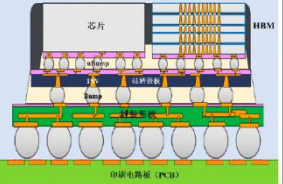
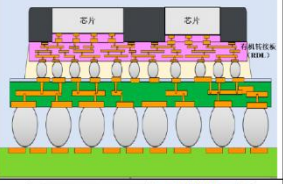
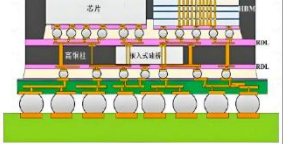
数据来源：中国半导体行业协会，灼识咨询，东吴证券研究所

盛合晶微已构建起覆盖硅转接板（TSV）、有机转接板（扇外型 RDL）及硅桥转接板等所有主流方案的全流程 2.5D 技术体系，并具备成熟的量产能力。作为中国大陆在芯粒（Chiplet）及多芯片集成封装领域起步最早、技术最先进、布局最完善且生产规模最大的企业之一，公司在关键前沿领域已实现实质性突破，展现出追赶全球顶尖企业的

强大竞争力。公司芯粒多芯片集成封装业务营收占比从2022年的仅5.32%迅速攀升2024年的44.39%，并于2025年上半年进一步扩大至56.24%，公司先进封装业务需求持续强劲，技术实力已位于国内领先水平。

作为国产COWOS技术龙头的盛合晶微，上市融资后，有望开启先进封装的扩产新周期，建议重点关注盛合晶微产业链芯源微、强力新材等，同时建议关注其他先进封装企业甬矽电子、长电科技、通富微电等。

图33：盛合晶微2.5D封装平台

技术平台	平台简介	平台研发及产业化进展	平台图示
SmartPoser®-Si	结构： 基于硅通孔转接板（“硅转接板”）的2.5D技术平台，多颗芯片并列放置在硅转接板顶部，通过微凸块（uBump）和硅转接板中的布线实现芯片间的高密度互联，硅转接板通过硅通孔实现上下层间的互联 特点： 利用硅转接板中微米级线宽线距及多层水平互联的高密度布线优势，可将原本单一SoC芯片内部的部分互联转移到硅转接板中，实现集成芯片的效果	已实现大规模量产，是中国大陆量产最早、生产规模最大的基于硅转接板的2.5D技术平台之一	
SmartPoser®-RDL	结构： 基于扇出型重布线层（“有机转接板”）的2.5D技术平台，多颗芯片并列放置在有机转接板顶部，通过有机转接板中的布线实现高密度互联，有机转接板通过多层重布线实现上下层间的互联 特点： 利用有机转接板中微米级线宽线距及多层重布线的高密度布线优势，可将原本单一SoC芯片内部的部分互联转移到有机转接板中，实现集成芯片的效果	已进入小量试产	
SmartPoser®-BD	结构： 基于嵌入式硅桥（“硅桥转接板”）的2.5D技术平台，多颗芯片并列放置在扇出型重布线层顶部，并在扇出型重布线层中局部内嵌硅桥，通过微凸块和硅桥中的布线实现芯片间的高密度互联，硅桥转接板通过高铜柱以及硅桥中的硅通孔实现垂直方向的互联 特点： 可视作硅转接板和有机转接板的结合，在降低硅转接板高成本的同时，能够解决有机转接板互联密度低、垂直供电弱的问题，可以兼顾性能和成本	已完成全流程验证和可靠性验证，并开展更大尺寸的结构优化	

数据来源：盛合晶微招股说明书，东吴证券研究所

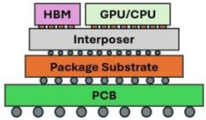
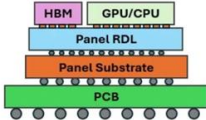
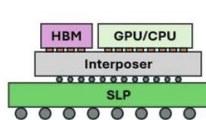
CoWoP 区别于其他芯片封装形式，直接去除中间基板对 PCB 精度提出极高要求。 CoWoS、CoPoS、CoWoP 代表了先进封装技术演进的不同阶段，各自具有独特的结构和应用。CoWoS 目前是高端人工智能服务器的主要解决方案，其核心是将 CPU/GPU 和 HBM 集成在硅中介层上，然后将其安装在 ABF 封装基板上，支持 CoWoS-L 的线宽/线距达 5-8 微米，CoWoS-S 的线宽/线距达 10-15 微米，具备成熟的技术及高频、高速互连能力。CoPoS 进一步引入了具有 RDL 的结构的面板级 ABF 基板，使得在方形或者大面积封装上成为可能，从而提高生产效率并降低成本。线宽/线距为 8-15 微米，CoPoS 目前处于早期研发阶段。

CoWoP 完全去除 ABF 基板，直接将芯片与中介层安装在高精度 PCB 上，例如 SLP，理论上实现了最短的信号路径、最低的成本并改善了热管理的灵活性。其要求线宽/线距精度 15-20 微米，甚至 10 微米以下。目前处于验证阶段，可能用于 Rubin Ultra 平台。

CoWoP 封装方案对 PCB 板各方面工艺提出极高要求，直接提升 PCB 整体价值量。 1) 对于板面平整度要求高，延伸出对于玻纤布的 Low-CTE 要求，目前石英布拥有极佳的电性能与较低的 CTE 值；2) 线宽/线距精度高，系统板必须承担此前封装基板的高密度布线功能，线宽/线距必须达到 15-20 微米或更低，同时保持严格的板面平整度和尺寸稳定性，延伸出 mSAP 工艺；3) 热机械可靠性，直接安装在 PCB 上的裸芯片和中介层

面临由于热循环与 CTE 不匹配而加剧的翘曲风险；4) 制造良率挑战，传统 PCB 工厂需要升级到接近先进封装标准的洁净室，并且装配良率必须很高。CoWoP 所需要的 PCB 对于材料、加工工艺都提出极高要求，加工设备也需使用高端设备，PCB 整体价值量有望进一步提高。

图34：CoWoS、CoPoS、CoWoP 封装类型的比较

类型	CoWoS (Chip on Wafer on Substrate)	CoPoS (Chip on Panel on Substrate)	CoWoP (Chip on Wafer on PCB)
架构			
核心概念	包含HBM、GPU/CPU、中介层、封装基板、印刷电路板	包含HBM、GPU/CPU、面板重布线层、面板基板、印刷电路板	包含HBM、GPU/CPU、中介层、类基板印刷电路板
基板类型	ABF基板（传统封装基板）	面板级ABF基板	高精度类基板印刷电路板
线宽/线距	CoWoS-L: 5-8μm CoWoS-R: 8-12μm CoWoS-S: 10-15μm	面板级: 8-15μm	15-20μm，全面量产需要小于10μm
优势	技术成熟；支持HBM堆叠和高速互连	生产效率更高，单位成本更低，适合大面积设计	结构简化，信号路径最短，热设计灵活性更好，理论上成本最低
挑战	基板尺寸限制；翘曲风险	面板级设备仍在开发中；需要生产转换	对印刷电路板精度要求极高，热机械可靠性问题，良率风险高
量产时间线	已实现量产（H100/H200、人工智能/高性能计算）	处于早期研发阶段，计划被英伟达采用	处于早期研发阶段，计划被英伟达采用
供应链影响	依赖ABF基板；台积电CoWoS产能是关键瓶颈	推动面板级封装设备和材料升级	需要高精度印刷电路板、超薄铜箔、类基板印刷电路板和改良型半加成法工艺作为关键支撑

数据来源：SemiVision Research，东吴证券研究所

5. 消费电子：科技巨头引领端侧创新大年

5.1. 端侧交互范式重构：模型能力跃迁，统一入口加速成形

5.1.1. 端侧模型进展：从技术演进到生态加速

端侧模型可以解决隐私、离线可用性、低延迟等核心问题，成为推动消费电子产业周期的核心变量。AI 加速渗透消费电子终端，而智能终端对实时性、隐私性和成本功耗要求更高，端侧模型成为推动 AI 在端侧落地的关键路径。端侧模型的成熟度成为 AI 能否深入渗透终端，从而推动新一轮消费电子产业周期的核心变量。回顾 2025 年以来，端侧大模型在技术路径、生态布局以及产品形态均取得明显进展，下面将分别展开介绍。

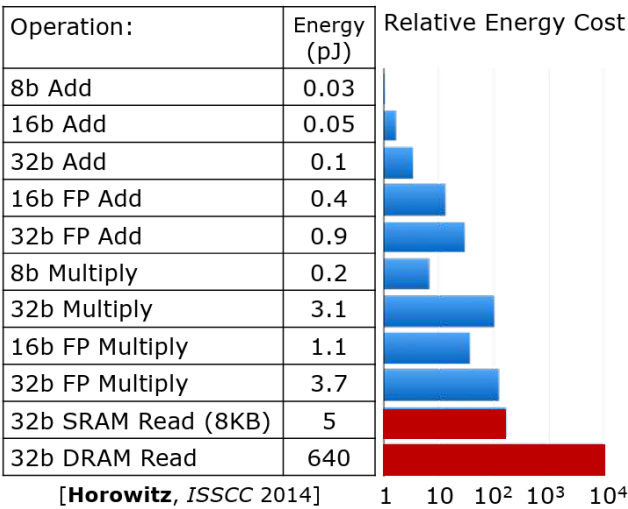
图35： 2025 年端侧大模型发布里程碑事件梳理

厂商	发布时间	发布模型型号	参数规模	技术特点
DeepSeek	2025年1月	DeepSeek-R1 蒸馏版本	1.5-70B	基于旗舰模型蒸馏微调得到
OpenAI	2025年8月	gpt-oss-20b	20B	MoE架构，强化推理与工具调用能力
Google	2025年3月	Gemma 3	1-27B	支持图文多模态输入
	2025年5月	Gemma 3n	5/8B	降低内存占用
	2025年6月	Gemini Robotics On-Device		端侧VLA模型
Microsoft	2024年12月	Phi-4	14B Dense	语言处理、数学和逻辑推理
	2025年2月	Phi-4-mini	3.8B	新增工具调用
	2025年2月	Phi-4-multimodal	5.6B	图像理解、语音理解与生成；多语言理解
	2025年4月	Phi-4-reasoning/Phi-4-reasoning-plus/Phi-4-Mini-Reasoning	14/14/3.8B	优化推理能力
	2025年7月	Phi-4-mini-flash-reasoning	3.8B	加快推理速度
阿里巴巴	2025年4月	Qwen 3系列	0.6-32B稠密模型	基于旗舰模型蒸馏微调得到
	2025年8月	Qwen3-4B-Instruct-2507和Qwen3-4B-Thinking-2507	4B	256K tokens上下文理解能力，理解和推理能力增强
	2025年10月	Qwen3-VL-4B/8B	4B/8B	同时提升视觉与文本能力，实现多模态能力不降级
面壁智能	2025年6月	MiniCPM4	8B/0.5B	MoE架构文本基座模型
	2025年8月	MiniCPM-V 4.5	8B	端到端视觉推理、图文对话、跨 30+ 语言多模态能力
	2025年9月	MiniCPM4.1	8B	引入可训练稀疏注意力；支持“深推理模式 vs 非推理模式”可切换
小米	2025年4月	MiMo-7B 推理模型	7B	数学推理、代码生成等
	2025年8月	MiDashengLM-7B	7B	统一理解语音、环境声与音乐的跨领域能力
苹果	2025年7月	Apple Intelligence 第二代基座模型	3B	键值缓存共享架构降低TTFT和内存需求；模型权重压缩至2-bit
	2025年7月	Beyond Sensor Data		用 25亿小时、16.2 万人可穿戴“行为信号”训练健康基础模型

数据来源：各公司官网，TechCommunity，Huggingface，东吴证券研究所

技术层面看：推理叠加模型压缩技术推动小模型高效落地。此前业界主流观点认为模型智能程度与参数规模正相关，因此模型参数规模持续指数级扩张。具体到端侧，硬件内存与带宽提升滞后于参数量提升，数据搬运成本成为主要瓶颈，从而限制了大模型在端侧的部署，进而影响模型的智能程度、用户体验以及用户的付费意愿。**2025 年 1 月，DeepSeek 通过“延长思考时间即可提升输出质量”的新 Scaling Law 及其蒸馏小模型展示了小参数也能凭借强化推理能力脱颖而出的可行路径**，随后微软在 3 月推出面向 Copilot+ PC 的 DeepSeek R1 蒸馏模型，端侧模型渗透开始加速。复盘 2025 年端侧模型迭代可见，业内通过量化、稀疏化与蒸馏等压缩手段，将大模型能力下沉至端侧，同时在推理阶段引入“快/慢双模式”以在时延与效果之间动态权衡。具体看，Phi 等小模型主要依托教师模型的思维链蒸馏获得能力强化；低比特量化成为降低存储与算力开销的标配。同时，MoE、稀疏注意力、KV cache 优化、以及多 Token 预测（MTP）等架构设计进一步压缩计算量和解码时延。面壁、千问等厂商亦通过同一模型内置深推理/快速响应切换机制，实现端侧场景下的自适应算力分配。

图36：内存读取所需功耗远高于计算



数据来源：MIT，东吴证券研究所

图37：多重模型压缩技术协同驱动端侧模型高效落地

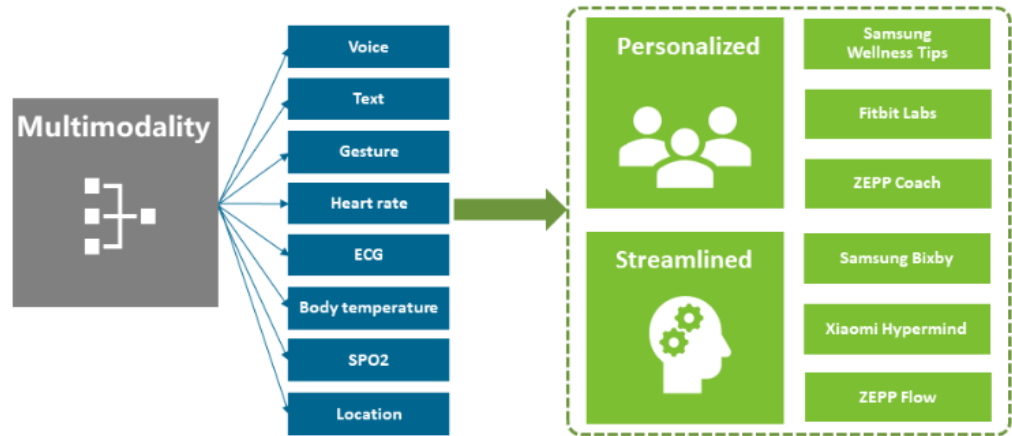


数据来源：机器之心，面壁智能，东吴证券研究所

生态层面看：多方玩家加速布局端侧模型，鉴于端侧相较云端具备更强的用户黏性，且本地智能能够为同类服务形成显著差异化，未来将有更多厂商持续加码。从玩家构成来看，我们认为端侧模型参与者可大致分为互联网生态巨头、专业大模型厂商以及硬件厂商三类，2025 年各家普遍强化了端侧模型的产品线。具体来看，OpenAI 于 8 月推出 gpt-oss-20b，这一端侧友好模型对其构建“全场景 AI 覆盖”具有战略意义；Meta Llama、阿里千问、谷歌等互联网与 AI 头部厂商则在既有端侧布局基础上进一步迭代与补强。

各家部署端侧模型的共同目标在于增强自身业务能力：微软于 2025 年 5 月向 Edge 浏览器开放端侧模型能力，有望显著提升浏览器侧的原生智能服务能力；小米在今年发布的 MiMo 与 MiDashengLM 系列则分别强化自身生态链产品的智能程度与语音交互体验。展望未来，端侧智能与用户日常生活具有天然高频接触特点，其粘性高于云端形态；又因 AI 模型的赋能可以为原本同质化的产品和服务打造差异化体验，未来将有更多厂商加码参与。

图38：AI 赋能下，同质化产品/服务可实现差异化体验



数据来源：Canalys，东吴证券研究所

从产品形态演进看，端侧智能的中长期主线将同时指向多模态虚拟智能与具身智能两条路径。我们认为按是否与现实世界发生物理交互，智能终端可划分为虚拟智能与具身智能两大类：前者主要覆盖手机、PC、可穿戴等数字化终端，后者对应机器人等需要在真实环境中执行动作的物理载体。复盘 2025 年行业进展，大多数厂商仍以虚拟智能为主要落点，沿“基础模型（文本能力）→推理模型（推理增强）→多模态模型（语音/视觉/文本端到端理解与生成）”的路径持续迭代。相较之下，具身智能在多模态理解的基础上还必须把感知与语言转化为可执行动作；同时受限于高质量动作标注数据获取成本高、安全试错代价大等问题，研发难度显著高于虚拟智能。当前仅少数头部厂商推进具身端侧模型布局，例如 Google 推出的 Gemini Robotics On-Device 采用 VLA 框架在机器人本体离线运行，但整体仍处于早期探索阶段。若上述训练瓶颈逐步突破，具身智能有望打开更广阔的端侧应用空间，成为下一阶段端侧 AI 渗透的重要增量方向。

5.1.2. 端侧 AI 统一入口：从单点助手到跨应用操作系统

从产业共识来看，端侧 AI 被视为下一代人机交互方式的革新。AMD 在白皮书中指出，AI 从长期看将改变用户与终端的交互方式。谷歌也表达过“将 Gemini 打造为通用人工智能代理”的愿景。要实现这种“系统级 AI”的统一入口形态，模型本身也在完成智能体能力（Agentic application）的补齐，我们认为主要包括推理规划、工具调用与多功能编排、记忆管理和多模态感知能力。2025 年以来，OpenAI 在 DevDay 发布 Agents SDK 等 agent 开发套件并披露其平台已拥有超 400 万开发者；Anthropic 等亦推出 Claude Agent SDK 等 agentic 开发框架的开源方案。从工具链和脚手架的完善情况看，我们认为智能体能力已基本达到可以在企业与消费端场景中大规模落地和推广的阶段。

在此基础上，行业已围绕“端侧 AI 统一入口”展开布局与竞争。从当前进展看，主要可分为应用侧与终端侧两类路径。应用侧玩家从 App 入口切入，借助自身生态或开放生态构建统一界面。典型如微软 Copilot Agents 与阿里千问 App，均以 AI 为统一入口，将自家生态中的文档、办公、搜索、服务等能力进行整合；而 OpenAI 等厂商则通过在 ChatGPT 中纳入第三方应用，试图以开放生态的方式扩大自身的入口优势，实现跨应用任务的统一调度。终端侧玩家则依托系统级能力，从硬件与 OS 层面推动跨 App 集成。多数厂商以内置应用为主要载体实现系统级协同；同时，也出现了更深层的生态合作，例如三星联合谷歌生态应用实现跨 App 操作，华为则在 HarmonyOS 6 中通过小艺与支付宝健康等服务联动，完成端侧跨应用服务集成。展望未来，我们认为端侧 AI 的交互形态将沿着统一入口的方向持续演进，并将覆盖手机、PC、可穿戴、车机等多终端。

图39：多方玩家布局端侧 AI 统一入口

厂商	产品	核心功能
应用侧		
OpenAI	Apps inside ChatGPT	允许第三方App直接运行在ChatGPT对话里，ChatGPT 逐步变成人-应用交互的操作系统
Google	Gemini App	支持任务式 agent，可打开网页并执行如筛选、查询等复杂操作。
AWS	Q Business（暂定）	面向B端，旨在提供一个“单一界面”，可访问各类业务应用程序（如 Asana、Excel），并自动完成操作，如关单或发邮件，实现代理式工作流。
Microsoft	Copilot Agents	Copilot 可在 Word、Excel、PowerPoint 等应用中直接生成内容，并支持跨应用操作。
阿里巴巴	千问APP	能通过图片识别商品，并在淘宝、1688、闲鱼等平台查找替代品，一键跳转下单。
智谱	AutoGLM 2.0	通过“智能体手机/电脑”，用户一句话即可让 AutoGLM 云端自动操作美团、京东、小红书、抖音等几十个应用（如订外卖、买机票、查房源）。执行期间用户仍可自由使用其他 App。
终端侧		
Samsung	Galaxy AI	一句语音指令即可在多个应用间完成多项任务。支持 Google 应用、三星原生应用以及 WhatsApp、Spotify 等第三方应用。
Apple	Apple Intelligence	自动理解用户意图并跨 App 完成动作，例如：发送特定草稿邮件、增强照片后将其加入备忘录的某条笔记。
华为	小艺智能体	HarmonyOS 6允许Xiaoyi跨应用协调，无需用户逐个打开 App。例如，健康咨询只需用户说：“我最近血压波动，请通过支付宝健康管家检查我应该挂号哪个科室。”Xiaoyi会直接将请求传递给蚂蚁集团构建的医疗AI系统，呈现相关专家，并安排预约。

数据来源：各公司官网，TheVerge，AppsFlyer，机器之心，IPO 早知道，东吴证券研究所

5.2. 手机：OS Agent 启动，iPhone 先行

5.2.1. 苹果端侧智能生态加速成型

端侧基础模型向开发者社区开放、智能生态加速成型。iOS 26 在 Apple Intelligence 加持下，已经形成一套完整的 AI 能力体系。从文本写作、跨语言沟通、视觉理解，到系统自动化、个性化内容管理，再到开发者可直接调用的端侧基础模型框架，苹果正将 AI 深度嵌入操作系统的每一层结构。在具体能力上，iOS 26 的写作工具显著降低用户的时间成本；视觉智能可对相机画面与截图进行识别；叠加 Spotlight 的智能意图推理，使 iOS 已然具备个人助手的素质和能力。此外，Genmoji、Image Playground 等轻量生成式功能使创作门槛持续下降。Siri 依托 GPT-5 的复杂问答能力，也使系统具备更强的外部知识补全能力。端侧模型向开发者社区开源意味着 AI 走向生态社区，有望催生更多基于本地推理的创新应用。我们认为 iOS 26 的全面智能化升级有望刺激用户换机意愿，并为苹果长期的 AI 生态战略奠定坚实基础。

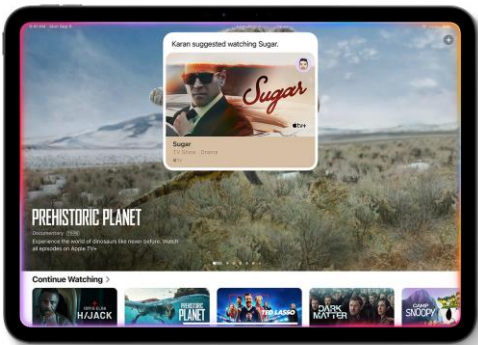
图40：苹果 iOS 26 AI 功能梳理

类别	功能
文本写作	写作工具： 在输入框处提供改写、润色、总结、校对、改变语气/风格、提炼要点等。 通知智能摘要： 对通知做摘要、聚合 短信/邮件智能功能： 对话关键点总结；重要邮件置顶等
跨语言能力	实时翻译： Messages、通话、FaceTime 实时双向翻译
视觉智能	相机： 用相机对准现实世界做识别、提问、翻译、总结 屏幕/截图： 截图后可问答；对高亮图像进行跨App搜索；识别图片上的事件并添加到日历
创作	Genmoji： 文本描述生成表情；多emoji混合 Image Playground： 图像生成/风格化
系统自动化	Spotlight： 动作意图推理
个人内容理解	照片智能搜索： 自然语言找照片/视频、识别情境、对象与文字 照片编辑： 清理画面杂物；生成记忆影片
Siri	更好的上下文理解；复杂问题集成GPT-5
开发者能力	Foundation Models framework： 开发者可直接调用端侧 LLM

数据来源：苹果官网，MacRumors，东吴证券研究所

OS Agent 将构成苹果核心 AI 能力,通过屏幕理解、个人知识库与跨应用操作能力，实现具备上下文记忆的全局智能交互模式。全局屏幕理解方面，系统将通过视觉智能实时解析用户当前界面内容，并可直接将识别结果转化为结构化操作，如从海报自动生成日程、对当前页面进行摘要、或基于屏幕出现的物品触发跨应用搜索等。其次，借助深度个人上下文建模，Siri 能够在本地设备的全域数据中检索用户请求的信息，而无需用户明确指定信息来源，实现如“查找某位联系人推荐的播客”或“查询家人航班信息”等复杂语义请求。通过跨应用编排，系统可在用户无需提供明确路径的情况下，自动完成跨 App 的多步骤操作，例如发送先前编辑的邮件草稿或处理照片并写入指定笔记。这些能力将通过后续系统更新逐步释放，构成苹果未来的核心 AI 能力。

图41：下一代 Apple Intelligence 赋予屏幕感知能力



数据来源：Apple 官网，东吴证券研究所

图42：下一代将具有跨应用执行能力



数据来源：Apple 官网，东吴证券研究所

5.2.2. OS Agent 有望掀起存量换机浪潮

苹果的用户理解、硬件产品矩阵、软件生态和变现能力都领跑行业，整合数据和资

源的空间最大。系统级 AI 能力主要在于用户信息的调用、支配、处理和执行，而苹果在个人数据和个人资料的智能管理和使用上优势明显。

1) 丰富的大量个人数据使苹果端侧 AI 模型的用户理解度高。核心产品 iPhone 在 2024 年获得 18% 的市场份额，在手机市场覆盖最多的高端机用户人群。苹果的大量第一方应用存储着大量用户信息、行为历史记录等，如财务信息、健康数据、图库信息、社交信息等，这是洞察用户需求的核心要件。

2) 丰富的硬件产品矩阵使得苹果 AI 生态再强化。苹果拥有 iPhone、iPad、Macbook、Airpods、Watch、HomePod 等丰富的产品矩阵。丰富的产品矩阵能够获取全面的传感器数据，控制各种个人设备，协同配合完成任务。

3) 丰富的软件生态优势。苹果有大量第一方应用，并对第三方应用有巨大赋能。这让系统级 AI 不仅能够对 App 内部进行操作控制，AI 还能更好的完成跨 App 的任务。同时 App Store 拥有强大的市场地位和高付费能力用户，在 Apple AI 生态的推进过程中，开发者的高积极性将进一步扩大苹果 AI 的优势。

图43: Apple 丰富的软件生态是系统级 AI 的基础

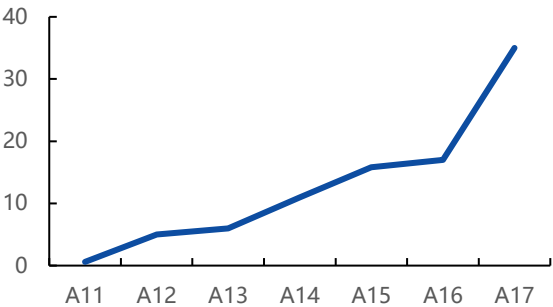


数据来源：WWDC，东吴证券研究所

iPhone 存量基数大，AI 开启新一轮换机周期。AI 功能仅适配 iPhone15 Pro 以上机型。iOS18 系统升级可以适用于 iPhoneX 以上的所有机型，但 Apple Intelligence 功能仅适用在搭载 A17 Pro 芯片的 iPhone 15 Pro 及以上机型，而在 iPad 和 Mac 产品上，只有搭载 M1 或往后芯片的机型才能支持。影响大模型能否在端侧运行的制约主要在算力和运存。

1) A14 系列及以上芯片即可满足 11TOPS 算力需求。目前支持苹果 AI 功能的芯片中，M1 芯片的 NPU 算力最低，为 11TOPS。横向对比来看，A14 及往后的芯片已经具备 11TOPS 算力。算力并非制约 AI 功能更新的核心要素。

图44: A 系列芯片 NPU 算力 (TOPS)



数据来源: cpu-monkey, 东吴证券研究所

图45: M1 芯片 NPU 能力 (TOPS)

	M1	M1Pro	M1 Max
制程和晶体管数量	5nm, 160亿	5nm, 337亿	5nm, 570亿
CPU	8核, 4*高性能核心+4*高效能核心	10核, 8*高性能核心+2*高效能核心	10核, 8*高性能核心+2*高效能核心
GPU	8核	16核	32核
NPU	16核神经网络引擎, 11TOPS	16核神经网络引擎, 11TOPS	16核神经网络引擎, 11TOPS
支持内存	16GB	32GB	64GB
内存带宽	68GB/s	200GB/s	400GB/s
媒体处理引擎	支持H.264、HEVC硬件加速, 视频编解码引擎	支持H.264、HEVC、ProRes RAW硬件加速, 视频编解码引擎、ProRes编解码引擎	

数据来源: Apple 官网, 东吴证券研究所

2) 内存或为苹果 AI 适配性的关键因素: 模型推理过程中需要占用较大的运行内存, 较小的内存会严重拖慢推理速度甚至无法完成推理。目前主流的 70 亿参数模型约需要占用 4GB 的内存空间, 需要采用至少 8GB 的 LPDDR5x。从现有 iPhone 的内存配置来看, 仅有 iPhone 15 Pro 以后的苹果手机能够达到 8GB 内存, 因而 8GB 内存可能就是当下苹果的 3B 端侧模型所需的最小内存数。

iPhone 15 Pro 以前的机型无法使用端侧 AI 功能, 这部分用户基数较大。增加 OS Agent 之后有望带动这部分存量用户大规模换机。

考虑到苹果 AI 落地场景成熟、具备系统级 AI 能力、端侧模型能力领先, 我们预估苹果 AI 将为存量用户换机提供巨大动能, 从而引发新一轮换机潮。

1) 果链白马受益于 OS Agent 落地手机, 迎来放量大年。相关标的: 立讯精密、领益智造、蓝思科技等

2) AI 需求带动手机硬件革命, 看好电池、散热环节有望核心受益于 AI 手机放量, 迎来量价齐升大年。相关标的: 珠海冠宇、信维通信、苏州天脉。

5.2.3. 硬件创新与 AI 迭代同步, 核心硬件更新公司迎来量价齐升

2026 年苹果折叠屏手机问世, 未来有望打开超 5000 万的放量空间。除了 AI 生态的更新, 苹果硬件产品也在 2026 年迎来变革大年, 其中潜力最大的就是苹果折叠屏。24 年全球折叠屏市场出货量约 1530 万部。目前仅安卓系&鸿蒙系手机推出了折叠屏, 且售价普遍在 600 美元之上。我们测算 2024 年, 600 美元以上的非苹果的高端手机中, 折叠屏的渗透率约有 19%。

图46: 折叠屏在苹果以外的高端手机渗透率测算

	2020	2021	2022	2023	2024
全球折叠屏手机出货量 (百万台)	2.8	8.9	12.0	14.6	15.3
600美金以上高端机销量 (百万台)					305
iPhone年销量 (百万台)	207.2	230.1	232.2	229.2	225.9
600美金以上非苹果的高端机销量 (百万台)					79.1
折叠屏在600美金以上苹果以外的手机渗透率					19%

数据来源: Canalys, Counterpoint, 东吴证券研究所

苹果折叠屏弹性测算: 若按未来苹果手机每年 2.5 亿部左右出货量水平。由于苹果手机售价除了少量 SE 系列外, 基本都在 600 美元以上, 因此我们参考安卓高端机中 19% 的折叠屏渗透率, 给苹果折叠屏 20% 的渗透率假设, 那么到 2030 年, **苹果折叠屏有望打开约 5000w 部放量空间, 是过去安卓&鸿蒙折叠屏市场的三倍**。若考虑苹果对折叠屏生态优化的潜力有望带动更多用户使用折叠屏, 给苹果折叠更高的 30% 渗透率, 则苹果有望在 2030 年打开 7500w 部折叠屏市场空间。潜力巨大!

图47: 折叠屏的核心成本提升在屏幕和铰链

	Galaxy Fold	占比	Galaxy S9+	占比
显示屏 / 触控模组	\$219	34%	\$79	21%
机械 / 机电部件	\$88	14%	\$30	8%
内存	\$79	12%	\$57	15%
应用处理器	\$71	11%	\$67	18%
摄像头	\$49	8%	\$38	10%
电源管理芯片	\$11	2%	\$9	2%
蓝牙 / 无线局域网	\$7	1%	\$7	2%
射频 / 功率放大器 / 前端	\$21	3%	\$19	5%
传感器	\$7	1%	\$6	1%
电池组	\$9	1%	\$5	1%
包装配件	\$19	3%	\$16	4%
其他组件	\$58	9%	\$45	12%
总物料清单 (BOM) 成本	\$637	-	\$376	-
(无补贴) 零售价	\$1,800	-	\$840	-
扣除 BOM 后的利润率 (%)	65%	-	55%	-

数据来源: CGS-CIMB, 头豹产业研究院, 东吴证券研究所

UTG 盖板玻璃和铰链模组相关环节最为受益。参考三星的折叠屏和直板机的 Bom 拆分来看, 折叠屏核心的价值增量在显示屏模组和机械/机电部件 (核心为铰链), 较直板机均有翻倍增量。2026 年苹果折叠屏上量有望将折叠屏市场打开三倍向上空间, 看好 UTG 盖板玻璃和铰链模组相关公司受益于苹果折叠放量迎来量价齐升大年。

UTG 盖板玻璃相关公司: 蓝思科技。

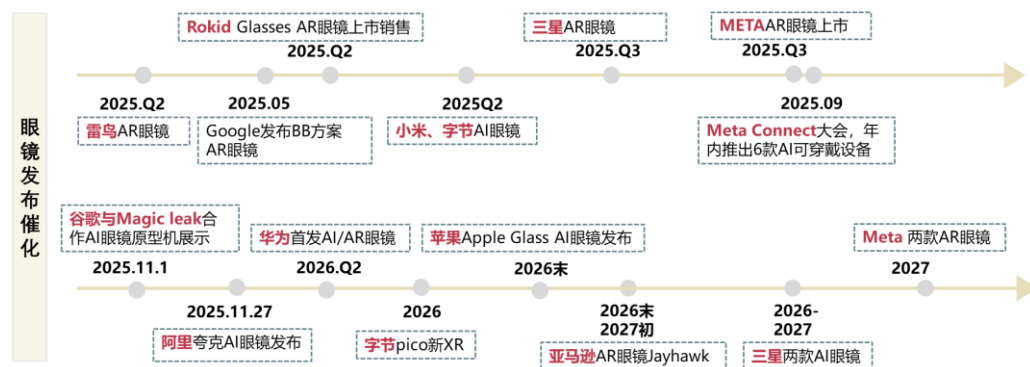
铰链相关公司: 铂力特、华曙高科、领益智造、宜安科技、统联精密、东睦股份、精研科技等。

5.3. 硬件形态更新：眼镜是大厂必争之地

5.3.1. 大厂的硬件布局，各类产品之中聚焦眼镜

终端厂商密集下海布局 AI 眼镜，AI 眼镜爆发在即。Ray-Ban Meta2 爆火之下，大厂纷纷布局 AI 眼镜。2026 年，Meta、三星、苹果、小米等多家厂商都将发布 AI 智能眼镜或 AI+AR 眼镜。大厂打样，白牌跟随，AI 眼镜有望快速放量。

图48：AI 眼镜催化一张图



数据来源：钛媒体，映维网，艾邦 VR 产业资讯，VR 陀螺，智东西，三次方 AIRX，MicroDisplay，Xreal，SHARGE 闪极，ARVR 星球，超时空视角，新浪 VR，仙瞬科技，莫界科技，东吴证券研究所

1) AI 眼镜：轻量化+价格亲民的 AI 智能眼镜有望在 2026 延续放量之年。CES2025 上，不带显示的 AI 智能眼镜重量普遍 < 40g。同时，AI 眼镜售价进一步下探，雷鸟 AI 眼镜售价降低至 1800 元，闪极 AI 眼镜价格降低至 999 元。AI 眼镜赛道将在 2026 年延续“从 1 到 100”的放量之年。预估 AI 眼镜 26 年达到 500w 出货量，28 年实现 5000w 放量。

2) AR 眼镜：具备成像显示的 AR 眼镜有望在 2026 迎来产品力的拐点之年。据 Wellsenn XR，2023 年和 2024 年全球 AR 销量均为 50w 台。AR 赛道的年出货量增速缓慢，我们分析，应用场景简单、硬件技术不成熟、以及高昂的价格，是制约 AR 眼镜放量的核心原因。

5.3.2. 明年 AI 眼镜增加显示有望带动场景升级打开出货量空间

我们推测，AR 眼镜的卡点有望在未来 3~5 年被逐个击破。

短期：2025~2026 年，观测 AI+AR 实现应用场景突破，从而到达产品力拐点。

模型视频理解/生成能力升级，端侧场景落地加速。Meta Live AI、豆包视频模型相继迭代，2025 年 5 月的豆包视频模型中，用户可以“和豆包打视频电话”。豆包可以根据摄像头实时拍摄到的现实场景，回答用户的提问。2024 年 12 月 Meta Rayban2 推出的

LiveAI 功能也与此相似。AR 眼镜具备最短的模型唤醒链路和视频的输入输出能力，有望成为模型视频理解能力提升最为受益的载体。

大厂陆续下海 AR 眼镜，AR 眼镜有望依托大厂生态进一步实现场景落地。从公开信息看，Amazon、三星、Meta、Google 等均将在 2025~2026 年发布 AR 眼镜，这些厂商具备丰富的软硬件生态，有望与 AR 眼镜行成联动，进一步延展 AR 眼镜的应用场景。

图49：2023 年和 2024 年 AR 眼镜的使用场景



数据来源：京东，东吴证券研究所

图50：Google Project Astra 展示的 AR 场景



数据来源：Google，东吴证券研究所

中长期：大厂入局+光学显示方案成熟带动 AR 眼镜硬件能力达到可放量水平。2025 年，Meta、Rokid、雷鸟、三星均将尝试将光学显示环节加入 AI 眼镜。我们预估随着光波导、Micro-LED 等技术逐渐成熟，光学模组重量和成本下降。应用场景更广的 AI+AR 眼镜有望走向 C 端。预估到 2029 年，AR 眼镜有望达 3000 万出货量。

5.3.3. AI 眼镜核心硬件 光学、SoC 等

AI 眼镜核心硬件受益方向：1) **SoC：**AI 眼镜中性能、续航和重量的平衡是核心痛点，SoC 厂商将持续受益于眼镜端低功耗 SoC 需求。2) **ODM：**短期受益于功能简单的 AI 智能眼镜的量增，中长期受益于高端 AI+AR 眼镜的 ASP 提升。

光学显示方向会是 AR 赛道的核心增量方向。光学模组是 AR 眼镜中的核心难点和增量环节。同时拆分 AR 眼镜的 BOM 成本看，光波导和光引擎两个核心的零组件在 AR 眼镜中价值量各占 30%，是核心的价值增量环节。

图51：AR 光学显示方案构成



数据来源：雷鸟，东吴证券研究所

1) 光引擎环节关注终局环节 Micro-LED。Micro-LED 因为其具备高亮度、高分辨率、小体积、低重量和低功耗，成为 AR 眼镜理想的光引擎方案。现阶段 Micro-LED 技术尚未完全成熟，还存在价格高、亮度、分辨率较低、全彩化难度大等问题，预估未来到 2028 年，随着这些难点问题逐步攻破，Micro-LED 将被陆续用于 AR 眼镜，打开近 300 亿元市场空间。

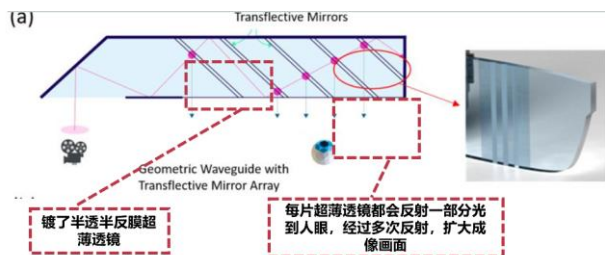
图52：AR 各光机方案对比

显示方案	LCoS	DLP	LBS	Micro-OLED	Micro-LED
图示					
显示原理	反射式液晶显示	数字微镜阵列	外部激光光源	有机自发光	无机自发光
响应时间	ms(毫秒)	μs(微秒)	ns(纳秒)	μs(微秒)	ns(纳秒)
对比度	1,000:1	2,500:1	2,000:1	100,000:1	1,000,000:1
亮度	根据背光源亮度决定，普遍 >10,000nit	根据背光源亮度决定，普遍 >20,000nit	100,000nit	1,000-6,000nit	100,000nit(全彩), 10,000,000nit(单色)
器件结构	复杂	复杂	复杂	简单	简单
工作温度	10°-70°	-40°-90°	较宽	-50°-70°	-100°-120°
功耗	高	中等	中等	低	低
光机体积	5-6cc	4cc	0.5-1cc	0.5-1cc	<0.5cc
寿命	10万小时	10万小时	10万小时	<1万小时	>10万小时
技术成熟度	制造技术较为成熟，成本较为低廉	制造技术成熟	技术成熟	成本高，制备工艺技术要求高	各参数表现优异，技术难度较大
产业化进展	规模量产，高性价比	规模量产，成本中等	成本高	小规模量产，未到达高良率	小规模试产

数据来源：VR 陀螺，东吴证券研究所

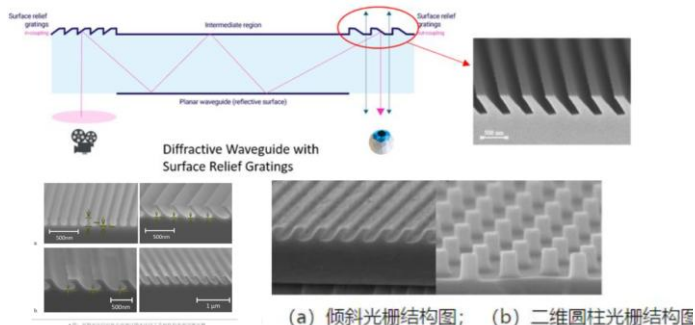
2) 光波导环节关注几何阵列光波导和表面浮雕衍射光波导两条技术路径。从成像效果看，几何阵列光波导因为采用折反射原理，光效更高，色彩均匀性更好。但劣势在于容易出现鬼影和杂散光等现象。从量产降本空间看，几何阵列光波导基于光学冷加工工艺，步骤繁多，镀膜和胶合环节的良率损耗大，因而整体良率提升和成本下降的难度较大。

图53：几何阵列光波导原理图



数据来源：VR 陀螺，东吴证券研究所

图54：表面浮雕衍射光波导原理图



数据来源：VR 陀螺，东吴证券研究所

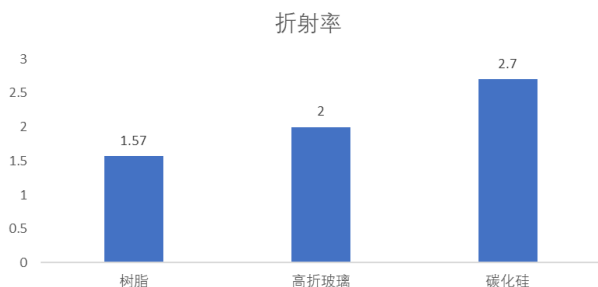
SiC 波导具备大 FOV、全彩化、轻量化等优势，有望成为衍射光波导的终局材料路径。

减少彩虹纹：碳化硅（2.7 折射率）折射率大，光栅的周期小，让环境光的衍射角度变大，减少彩虹纹。

FOV 做大：FOV 可以做到 70° ，普通玻璃基衍射光波导普遍在 $40\sim 50^\circ$ 。

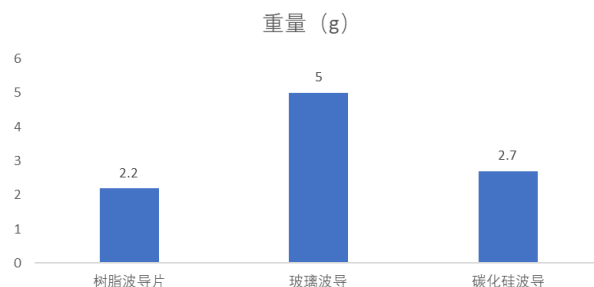
减重：不用做多层波导方案，并且 SiC 的密度更低，因而相比玻璃波导，重量更轻。Meta 从 2019 年开始尝试碳化硅方案，其 24 年 9 月发布的原型机 Orion 采用碳化硅波导，FOV 达到 70° ，其光机模组厚度降至 5mm 以下，重量减轻 40%，光效提升 85%。

图55：三种材料的折射率比较



数据来源：行家说三代半，东吴证券研究所

图56：三种材料的重量比较



数据来源：鲲游官网，东吴证券研究所

6. PCB: AI 算力基建加码，CAPEX 超级周期驱动 PCB 量价齐升

6.1. 全球资本开支共振架构升级，AI PCB 迎来量价齐升新周期

6.1.1. 全球 AI 巨头资本开支高增，算力“前置建设”引爆基础设施超级周期

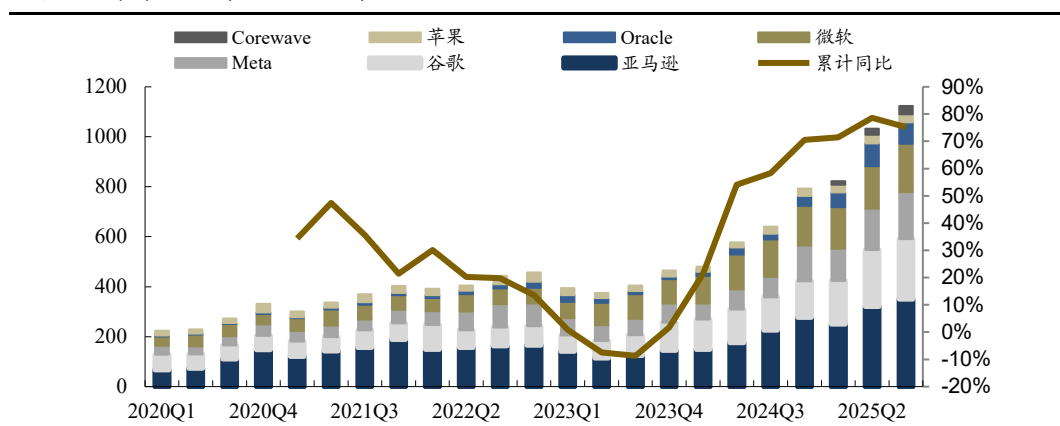
当前全球 AI 产业正从技术验证期加速迈向规模化产业化阶段，算力、模型与应用的联动推动产业链进入新一轮结构性扩张周期。过去一年，以生成式 AI 为核心的技术突破带动搜索、广告、电商、办公、视频等核心业务模式重塑，AI 渗透率快速提升。各

大科技公司普遍采取“前置投入、长期吸收”战略，资本开支由支持性投入转为核心资源配置方向。随着大模型训练成本随参数量呈指数级上升、推理规模化推广至企业级与消费级场景，AI 基础设施建设进入高密度扩张期，全球算力需求预计在 2025-2026 年持续翻倍增长。数据中心、电力系统、冷却、网络互联、AI 芯片等领域全面进入投资高峰，产业链已步入具有长期确定性的战略建设周期。

北美四大云服务厂商正共同开启 AI 基础设施竞争的资本密集期。谷歌依托“算力一模型一应用”一体化策略，Gemini 3 Pro 与 Nano Banana Pro 推动业务多元化扩张，2025 年资本开支有望延续 910-930 亿美元高位并在 2026 年仍具上调空间。微软通过 Azure 和自研 AI 平台强化算力供给，部署全球首个大规模 GB300 集群，计划未来两年将数据中心面积翻倍，2025 年资本开支有望继续上修，2026 年维持高增趋势。亚马逊在 Trainium 芯片、电力基础设施和 Rainier 集群上加码投资，2025 年预计支出约 1250 亿美元，2026 年有望保持高位扩张。Meta 则提前建设算力以支持“个人超级智能”战略，预计 2026 年资本开支或突破 800-850 亿美元，加速 AI 内容与商业应用孵化。四巨头在 AI 基础设施层面同步加速，构成全球算力扩张的核心驱动力。

基于此判断，AI 算力建设进入高增长阶段，将直接拉动上游核心硬件投资，尤其是 PCB/CCL 领域。AI 服务器、GPU 高速互联、电力转换模块等环节对高端 PCB 的层数、散热与可靠性要求显著提高，预计 2025-2026 年 AI 相关 PCB 市场需求将伴随算力扩张实现倍增。随着 1.6T 光模块、先进封装、HVDC 电源系统快速放量，高端覆铜板、铜连接等领域企业将充分受益。我们建议积极关注具备高端产能、算力服务器卡板布局及供应头部云厂商能力的 PCB 产业链公司，认为 AI 时代的资本开支爆发将转化为上游硬件端最具确定性的投资机会。

图57：海外云厂商 CAPEX 情况（亿美元）



数据来源：Bloomberg，东吴证券研究所

6.1.2. 机柜架构升级带动 PCB 价值量成倍增长，2026 年市场规模迈向 600 亿元

英伟达的机柜架构升级显著推动了 PCB 的量价齐升。在当前的 GB200/300 NVL72 机柜中，计算板（Bianca board）采用 22 层 HDI，交换板为 26 层通孔板，PCB 材料已

使用了高性能的 M8 等级。迈入下一代 Rubin 系列机柜后，PCB 设计和规格发生了重大飞跃：首先，Rubin NVL144 机柜新增了 Midplane，CPX-CX9 网卡模组，而计算板与交换板也做了重大升级，极大地提升了单个 PCB 的价值量。另外，Rubin Ultra NVL576 (Kyber) 机柜引入了革命性的正交背板方案，替代了铜缆连接并大幅提升芯片密度。由于数据传输速率要求超过 224Gbps，PCB 材料必须升级到 M9 或 PTFE 等更高等级的极低损耗材料，同时层数和工艺要求也极高。Kyber 机柜所需的 4 块正交背板，结合更高规格的计算板，将使单机柜的 PCB 总价值量实现成倍增长。

市场空间方面，2026 年 AI PCB 市场规模将达 604 亿元，同比增长 229.8%。GB200 NVL72 机柜中计算板由 5 阶 22 层 HDI 组成，整机柜使用 36 块，单板价值量约在 4000 元，交换板为 26 层通孔板，单板价值量 8000 元，对应单机柜 PCB 价值量约 21.6 万元，GB300 整体变化不大，总体 PCB 价值量与 GB200 相当。Rubin NVL144 机柜计算板为 5 阶 24 层 HDI，使用 M8 材料，交换板为 M9 级别 34 层通孔板，新增 CPX-CX9 模组板（M8 级别 5 阶 22 层 HDI）与 Midplane（M9 级别 44 层通孔板），NVL144 全系机柜均使用 Midplane 替换部分内部铜缆连接，对应整机柜 PCB 价值量（含 CPX 机架）预估提升至约 65.7 万元。下一代 Kyber 机柜将采用正交背板，柜内 GPU 数量提升至 144 个，整机柜 PCB 价值量或将进一步成倍增长。

ASIC 方面，谷歌 TPUv6/v7/v8 均为 64 卡机柜，其中 v6e/v6p/v7e/v8 的 PCB 制程分别为 M6 级 22 层板/M7 级 22 层板/M8 级 36 层板/M9 级多层板（待定），对应的单张 PCB 价值量预计为 4500/7000/16000/25000 元。

图58: AI PCB 市场规模测算

	GB200	GB300	Rubin	AWS TR2	AWS TR3	Google TPUv6e	Google TPUv6p	Google TPUv7e	Google TPUv8	Meta MTIA2	Meta MTIA3	AMD MI3XX	AMD MI4XX
单机柜使用量（块）													
OAM	36	36	36									8	
UBB				32	36	16	16	16	16	16	32	1	18
交换板	9	9	9							8	8		9
CPX-CX9模组			36										
正交背板													
正交中板			18										
PCB价值量（元/unit）													
OAM	4000	4000	7000									800	
UBB				11000	11000	4500	7000	16000	25000	6000	12000	7000	14000
交换板	8000	8000	11000							10000	20000		10000
CPX-CX9模组			3000										
正交背板													
正交中板			5000										
单机柜芯片数量	72	72	72	64	72	64	64	64	64	16	32	8	72
PCB价值量（元/GPU）	3000	3000	7625	5500	5500	1125	1750	4000	6250	11000	17000	1675	4750
单机柜PCB价值量（元）	216000	216000	549000	352000	396000	72000	112000	256000	400000	176000	544000	13400	342000
出货量（万颗）													
2025E	160	56		110		116	64			13		45	25
2026E		420	210	10	230	100	210	30	50	8	28	42	65
AI PCB TAM（亿元）													
2025E	48.00	16.80		60.50		13.05	11.20			14.08		7.54	11.88
合计													183.04
2026E		126	160.125	5.5	126.5	11.25	36.75	12	31.25	8.8	47.6	7.035	30.88
合计													603.69

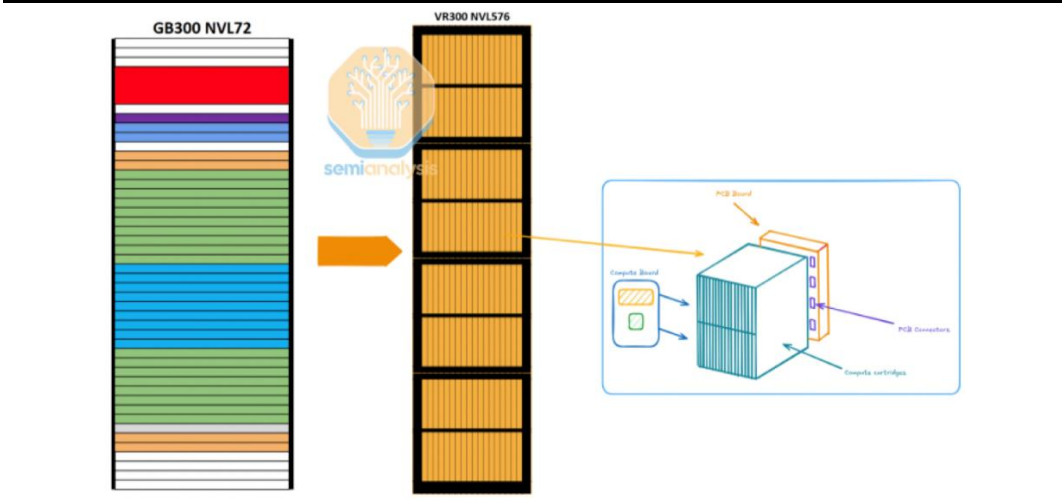
数据来源：根据市场公开信息测算，东吴证券研究所

6.2. CCL&上游材料：AI PCB 驱动 CCL 行业量价齐升，上游材料供不应求

6.2.1. M9: Rubin/Kyber 驱动 M9 材料升级，224G 速率开启 PCB 极低损耗时代

正交背板核心在高速率信号传输低损耗，对 PCB 材料提出更高要求。今年 GTC 大会上，黄仁勋展示了将于 2027 年下半年推出的 Rubin Ultra NVL576 产品，该产品使用了英伟达全新的 Kyber 机柜架构。每个机架包含 4 个 Canister，每个 Canister 包含 18 个计算板、6 个交换板以及 2 个机柜管理板。而正交背板取代了铜缆背板，为 Canister 中的计算板和交换板提供扩展链路。英伟达 Rubin Ultra 芯片使用 224G SerDes，传输速率较 B200/B300 进一步提高，正交背板树脂材料、玻纤布需要进一步升级以满足低损耗需求，目前 M9 及 PTFE 材料能满足低损耗要求，PTFE 具备更好的电性能，但加工难度极大，目前国内头部 PCB 公司正在进行产品验证。

图59：英伟达 Kyber 机柜正交背板



数据来源：SemiAnalysis，东吴证券研究所

随着数据中心对带宽需求持续提高，单通道 SerDes 速率升级到 224Gbps 已成为必然趋势，这使得 M9 及以上等级的极低损耗树脂材料的应用成为刚需。目前，在英伟达下一代 Rubin/Rubin Ultra 机柜架构中，M9 材料将正式投入使用：其中，Rubin 的 Midplane 和 CPX 模组，以及 Rubin Ultra 的正交背板，因信号传输距离和高密度要求，将主要采用 M9+Q 布材料。ASIC 体系在成本效益和供给考量下，仍将以 M8 材料为主。

图60：各厂商材料型号情况

光面	100GE x 64 (28Gbps)	400GE x 32 or 64 (56Gbps)	800GE x 64 (112Gbps)	1.6TE (224Gbps)
Ethernet	100GE x 64 (28Gbps)	400GE x 32 or 64 (56Gbps)	800GE x 64 (112Gbps)	1.6TE (224Gbps)
PCIe	PCIe 5.0 (32GT/s)	PCIe 6.0 (64GT/s)	PCIe 7.0 (128GT/s?) 及以上	
CCL	M6	M7	M8	M9
	Very Low Loss	Ultra Low Loss	Extreme Low Loss	Beyond
台光电 (Dk/Df)	EM-891 (3.2/45)	EM-890K (2.8/24)	EM-892K (2.8/17)	EM-896K2 (2.8/08)
	EM-528 (3.5/58)	EM-626K (2.9/25)	EM-892K2 (2.8/13)	EM-896K3 (2.8/07)
		EM-626 (3.1/38)		
		EM-891K (3.0/33)		
		EM-890 (3.0/36)		
台耀科技 (Dk/Df)	TU-883C (3.39/45)	TU-933 (3.4/25)	TU-943SN/SR (3.15/13)	TU-953Q (3.0/<.07)
		TU-883A (3.44/32)	TU-943HN/HR (3.15/12)	
联茂电子 (Dk/Df)	IT-968G (3.59/50)	IT-988G (3.76/35)	IT-988GSE (3.3/26)	IT-988GSE2 (3.05/11)
		IT-988GL (3.63/35)	IT-988GLSE (3.3/26)	IT-988GSE3 (2.93/07)
斗山 (Dk/Df)	DS-7409DV (T)	DS-7409DJG (N)	DS-7409DQN (3.07/15)	DS-7409DYQ (2.91/07)
	DS-7409DJG	DS-7409DV (NT)	DS-7409DQN2 (3.02/12)	DS-7409DYL2 (2.99/10)
松下 (Dk/Df)	Megtron6 (3.36/29)	Megtron7 (3.2/21)	Megtron8 (3.1/12)	Megtron9? (3.1/07)
			Megtron8S (3.19/12)	
生益科技 (Dk/Df)	Synamic 6 (3.68/36)	Synamic 8GN (3.28/19)	Synamic 9GN (3.34/14)	Synamic 10GQ (?/<.07)
		Synamic 6N (3.35/21)	Synamic 9GN2 (?/11)	
铜箔	(A-) RTF3/HVLP2	(A-) RTF4/HVLP2-4	(A-) RTF5/ HVLP3-5	
金居 (Rz,um)	RG312 (2.1 or below)	RG313 (1.9 or below)	RG315 (1.7 or below)	
	HVLP2 (1.5 or below)	HVLP3 (1.0 or below)	HVLP5 (0.5 or below)	
CCP (Rz,um)	RTF-35 (1.7 or below)	RTF-45 (1.5 or below)	RTF-55 (1.3 or below)	
			SI3-VSP/SF-VSP (0.4)	
三井 (Rz,um)	MLS-G4 (1.5)	MLS-G5 (0.8)		
	HS2-VSP (1.0)	SI-VSP (0.6)		
Circuit Foil (acquired by Defu)	BF-NN-(HT)	BFL-NN-Z	BFL-NF-Z	
		BFL-NX-Y		
古河	F1X-WS	FOX-WS	FOTW-WS	
	T9S-SV	DSV-A	T9DA-DSV	
Fukuda (福田)		T9DA-SV		
玻纤布	E Glass	NE/L Glass	NER/L2 Glass	Q Glass
Dk	6.6	4.7	4.5	3.4
Df	0.006	0.003	0.001	0.0007
CTE	5.4-5.5	3.1-3.8	3.0-3.5	<1.0

数据来源：各公司官网，东吴证券研究所

6.2.2. 铜箔、玻布：石英布与 HVLP4 铜箔市场需求持续旺盛，国产替代加速

石英布：高速通信升级下的市场爆发。电子玻纤布是高性能电路板中的关键增强材料，其中石英布以其综合表现最出色，代表了高速数据通信需求向第三代低介电常数材料升级的最高方向。随着 AI 服务器和交换机对低介电常数材料需求的激增，石英布的市场规模预计将迎来爆发式增长，其市场空间预估从 2026 年的约 30 亿元，高速增长至 2027 年的 99 亿元，同比增长 230%。

图61：石英布市场空间弹性测算

	2025年	2026年	2027年
M9材料--Rubin正交中板			
正交中板对应PCB市场规模（亿美元）		6	9
电子布价值量占PCB比例		12%	12%
正交中板对应石英布市场空间（亿人民币）		5	7
石英布价格（元/米）		300	300
石英布市场空间（万平米）		216	308
M9材料--Rubin正交背板			
正交背板对应PCB市场规模（亿美元）		4	31
电子布价值量占PCB比例		12%	12%
正交背板对应石英布市场空间（亿人民币）		3	26
石英布价格（元/米）		300	300
石英布市场空间（万平米）		144	1083
M9材料--交换机			
交换机对应PCB市场规模（亿美元）		6	19
电子布价值量占PCB比例		16%	16%
交换机对应石英布市场空间（亿人民币）		7	22
石英布价格（元/米）		300	300
石英布市场空间（万平米）		274	912
M9材料--其他			
采用M9的PCB市场规模（亿美元）		21	56
电子布价值量占PCB比例		12%	12%
M9中电子布市场空间（亿人民币）		18	47
石英布市场空间占比假设		50%	50%
二代布市场空间（万平米）		822	2141
石英布市场空间（万平米）		377	981
石英布合计市场空间（亿元）		30	99

数据来源：根据市场公开信息测算，东吴证券研究所

铜箔方面，预计 2026 年全球 HVLP4 铜箔总需求量将达到 8600 吨，主要由英伟达与亚马逊两大算力巨头的高端硬件放量驱动。其中，明年 Rubin 机柜预计出货 4 万台、单柜 HVLP4 用量 100kg，英伟达侧总需求约为 4000 吨；明年 Trainium 3 芯片预计出货 230 万颗（对应 115 万个 UBB，单 UBB 用量 4kg），亚马逊侧总需求约为 4600 吨。

在全球科技巨头掀起前所未有的 AI 资本开支热潮下，以谷歌 Gemini 3 和亚马逊 Trainium2 为代表的算力需求正加速传导至上游硬件供应链，推动其结构性重构。英伟达机柜架构的代际升级是这一趋势的核心体现：从 GB200/300 迈向 Rubin/Kyber 架构，通过引入 Midplane 和正交背板，极大地提高了 PCB 的层数、阶数和技术复杂度，预计将使单机柜 PCB 总价值量实现成倍增长，推动 AI PCB 市场规模至 2026 年达到 600 亿元。为满足 224Gbps 等超高速率传输需求，核心材料必须升级，M9/PTFE 极低损耗树脂材料将正式用于 Rubin 的正交背板和 Midplane，同时石英布等高端玻纤布和 HVLP4

铜箔的需求也迎来爆发性增长。总体而言，AI 军备竞赛已全面打开硬件产业链量价齐升的新成长周期。我们推荐产业链相关标的：1）PCB：胜宏科技、景旺电子；2）CCL：生益科技；3）电子布：菲力华；4）铜箔：铜冠铜箔。

7. 风险提示

AI 算力基础设施建设放缓。当前行业高景气度很大程度上由 AI 算力需求驱动。若下游 AI 大模型商业化落地不及预期，导致云厂商削减资本开支，将可能向上传导，导致先进制程与 CoWoS 等先进封装产能利用率下滑，进而影响上游设备需求。

行业竞争加剧与价格战风险。随着部分成熟制程及消费电子细分赛道产能扩充，市场竞争可能进一步加剧。若相关公司无法持续保持技术领先或成本优势，可能面临价格战导致的毛利率下滑及市场份额丢失的风险。

上游原材料价格波动风险。电子制造环节对金、铜、硅片及特种气体等原材料依赖度较高，若上游大宗商品或核心材料价格大幅上涨且无法有效向下游传导，将压缩中游制造环节的盈利空间。

汇率波动风险。半导体行业供应链全球化程度高，涉及大量元器件进口及终端产品出口，若汇率出现大幅波动，将直接影响相关公司的汇兑损益、原材料成本及出口产品的全球竞争力。

AI 创新与应用落地不及预期。AI 大模型及端侧 AI（AI 手机、AIPC）是驱动本轮电子行业成长的关键动力，如果 AI 商业化路径受阻、云厂商资本开支放缓或消费者对 AI 硬件换机意愿不强，则可能拖累产业链整体的复苏节奏。

免责声明

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

本研究报告仅供东吴证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。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。

在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

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，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

买入：预期未来 6 个月个股涨跌幅相对基准在 15%以上；

增持：预期未来 6 个月个股涨跌幅相对基准介于 5%与 15%之间；

中性：预期未来 6 个月个股涨跌幅相对基准介于-5%与 5%之间；

减持：预期未来 6 个月个股涨跌幅相对基准介于-15%与-5%之间；

卖出：预期未来 6 个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

增持：预期未来 6 个月内，行业指数相对强于基准 5%以上；

中性：预期未来 6 个月内，行业指数相对基准-5%与 5%；

减持：预期未来 6 个月内，行业指数相对弱于基准 5%以上。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，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。

东吴证券研究所
苏州工业园区星阳街 5 号
邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>