



电子行业研究

买入（维持评级）
行业深度研究

证券研究报告

电子组

分析师：樊志远（执业 S1130518070003）

fanzhiyuan@gjzq.com.cn

AI 应用带动推理算力需求，看好 ASIC 行业厂商发展

投资逻辑

我们认为 AI 推理降本有望带动 AI 应用的爆发，ASIC 作为高性价比的算力方案，有望充分受益推理算力需求增长。从 chatgpt3 发布以来，大语言模型的推理的成本以指数级别下降，单美元可生成的 token 数量持续增长，据 A16z Infrastructure 测算，推理成本每年降低幅度约 10 倍。模型推理的算法升级落地，通过包括强化学习、MOE 等方法提升模型推理能力。我们认为：模型的推理成本快速降低，有望带动应用的爆发，而应用爆发将带动更多的推理算力需求，同时推理算法的迭代也带动更多算力需求。ASIC 相比 GPU 可针对业务场景做定制开发，我们认为大量的推理算力以及降本诉求将有效带动 ASIC 的需求增长。目前北美云厂商积极布局 ASIC，已经实现规模化生产的 ASIC 包括谷歌的 TPU、亚马逊的 Trainium 与 Inferentia、微软的 Maia 以及 Meta 的 MTIA，其中谷歌和亚马逊的产品在 2025 年已经开始实现较大规模出货。我们认为其他云厂商 ASIC 在成熟度逐渐提升后，未来也具备较大的增长空间。Openai 等厂商也在积极布局 ASIC。从数量来看，TPU 是目前 ASIC 当中最为主要的产品之一。根据 Digitimes 测算，2023 年 TPU 在 ASIC 的市占率约 71%，2024 年全球 ASIC 出货达到 345 万颗，其中 TPU 占比提升至 74%。根据 Omidia 测算，2024 年 TPU 的销售额在 60~90 亿美元。亚马逊 Trainium2 已经推出，2024、2025 年出货增长迅猛。根据 Trendforce，AWS 的出货成长力道强劲，24 年年增率突破 200%。预计 2025 年 AWS 出货量将成长 70% 以上，并更聚焦往 Trainium 芯片发展，投入 AWS 公有云基础设施及电商平台等 AI 应用。Meta ASIC MTIA 已经推出第二代，目前主要用于模型推理，Meta 预计 26 年将有用用于训练的 ASIC 推出。除了北美四大 CSP 以外，我们预计未来领先的模型厂商如 openai 等也将开发 ASIC。根据《经济日报》，openai 预计将采用台积电 3nm 以及 A16 制程生产 ASIC，预计 openai 的 ASIC 将在 26 年年底进入量产。

我们看好 ASIC 设计服务公司、以太网白盒交换机厂商与以太网交换芯片厂商、AEC 厂商、PCB 厂商。1) ASIC 设计服务公司：目前除美系厂商外，台系厂商也已拿到北美云厂商 ASIC 项目，并有望在未来扩大份额。我们认为，ASIC 在推理算力占比提升的大背景下，需求增长的趋势明确，但需要关注行业竞争格局的变化，需要密切关注每个芯片项目的获得者。建议关注导入北美云厂商的 ASIC 设计服务公司。在美国制裁加剧情况下，我们认为国内云厂商 ASIC 项目有望采用国内设计服务企业，国内相关企业有望受益。2) 以太网白盒交换机厂商与以太网交换芯片厂商：由于 ASIC 采用以太网组网，我们认为 ASIC 的放量有望带动以太网在 AI 组网的渗透率提升以及需求。白盒以太网交换机方案趋势明确，主要厂商 Arista 在 10G 及以上的数据中心交换机市场的出货量以及营收市占率都已经超过思科。未来来看，我们认为云厂商采用自研系统采购白牌交换机的比例有望逐渐扩大。建议关注：Arista 以及白牌交换机厂商。以太网交换芯片是以太网交换机的核心部件，建议关注商用以太网交换芯片厂商。3) AEC 厂商：以太网白盒交换机厂商与以太网交换芯片厂商，我们认为 Credo 具有一定先发优势，看好 Credo，建议关注其他在 AEC 布局的厂商。4) PCB 厂商：我们认为 ASIC 兴起将带动产业链显著升级，AI 对算力 PCB 提出更高需求，同时数据量增长带动基础设施升级。另外 ASIC 芯片采用多 die 形式，显著提升对载板的需求。建议关注在高速通信领域具备较强高多层板与 HDI 板的厂商，以及具备封装基板能力的厂商。

投资建议

目前北美 CSP 已经积极布局 ASIC，另外模型厂商如 Openai 也在积极布局 ASIC。ASIC 有望迎来高速增长。我们看好 ASIC 设计服务公司、以太网白盒交换机厂商与以太网交换芯片厂商、AEC 厂商、PCB 厂商。

风险提示

AI 进展不及预期；行业竞争加剧；下游需求不及预期。



内容目录

一、AI 应用有望带动推理算力需求，云厂商持续加码 ASIC.....	4
1.1 推理降本趋势明显，ASIC 方案性价比凸显.....	4
1.2 谷歌、亚马逊加码 ASIC，其他 CSP 有望跟进.....	5
二、看好 ASIC 设计服务公司、以太网、AEC、PCB 相关厂商.....	7
2.1 ASIC 设计服务公司业绩进入放量期，关注 ASIC 项目 design wins 变化.....	7
2.2 ASIC 采用以太网组网，以太网交换芯片、白盒交换机厂商有望受益.....	10
2.3 ASIC 带动 AEC 需求加速，云厂商开始导入 ASIC.....	13
2.4 PCB：ASIC 的单位 PCB 耗用量更高，多 die 合封带来封装基板市场需求扩容.....	15
三、投资建议.....	18
四、风险提示.....	18

图表目录

图表 1：大模型推理成本按照指数下降，每年降低幅度约 10 倍.....	4
图表 2：Deepseek 依靠推理端算法优化可以实现较好智能程度（单位：%）.....	5
图表 3：ASIC 相比 GPU 牺牲灵活性，但提高效率.....	5
图表 4：v6e 在算力上有明显提升.....	6
图表 5：24 年全球 ASIC 出货 345 万颗，谷歌市占率 74%.....	6
图表 6：亚马逊 Trainium2 包括 Trainium2 与 Trainium2 Ultra 两个版本.....	6
图表 7：Meta、微软已经分别推出 ASIC MTIA、Maia.....	7
图表 8：定制 ASIC 业务的技术关键在于关键 IP 与先进封装技术.....	7
图表 9：博通 25 年将实现三家 ASIC 客户产品量产.....	8
图表 10：博通 AI 收入占半导体收入比例快速提升.....	8
图表 11：Marvell 已经获得三家客户的 ASIC 项目.....	8
图表 12：Marvell 预计 28 年定制化芯片 TAM 达到 429 亿美元（单位：十亿美元）.....	8
图表 13：世芯营收高速增长，24 年同比+70.5%.....	9
图表 14：世芯 7nm 及更先进制程项目收入占比超 90%.....	9
图表 15：联发科在 Computex 2024 上展示的 ASIC 设计方案.....	9
图表 16：创意电子 24 年实现营收 250.44 亿新台币，同比-4.6%.....	10
图表 17：北美 CSP 引入多家 ASIC 供应商.....	10
图表 18：白盒交换机解耦上层应用与底层硬件.....	11
图表 19：开放与解耦是数据中心网络发展方向.....	11
图表 20：Arista 数据中心交换机按营收市占率超过思科.....	11
图表 21：Arista 数据中心交换机按端口数市占率超过思科.....	11



图表 22: Arista 数据中心高速以太网 (100G/200G/400G) 端口数出货量已超思科两倍 (单位: 百万)	12
图表 23: 交换机主要芯片为交换芯片	12
图表 24: 交换机当中芯片成本占比达 32%	12
图表 25: DAC 直径、所占空间更大, 排线困难	13
图表 26: AEC 直径、所占空间更小	13
图表 27: AEC 在短距离高速率传输优势较 DAC 和光通信方案明显	13
图表 28: AOC、DAC、AEC 28 年市场有望达 28 亿美元, AEC 增速最快	14
图表 29: Credo 客户拓展顺利, FY25Q2 三家头部企业成为 Credo 营收占比超 10% 客户	14
图表 30: 主要厂商 AEC 产品进展	14
图表 31: 2023~2028 年全球 PCB 细分领域复合增速预测	15
图表 32: 2023~2028 年全球 PCB 细分领域复合增速预测	16
图表 33: 亚马逊 Trainium2 服务器示意图	16
图表 34: 谷歌自研 AI ASIC 芯片 TPU	16
图表 35: 亚马逊自研 AI ASIC 芯片 Trainium2	17
图表 36: 全球封装市场收入 (十亿美元)	17
图表 37: 全球封装晶圆产能耗用 (千片)	17
图表 38: 全球先进封装市场按工艺分布 (十亿美元)	17
图表 39: 2024~2029 年全球先进封装各工艺复合增速	17
图表 40: 载板在 FCBGA 封装制造成本中占据 50% 的比例	18

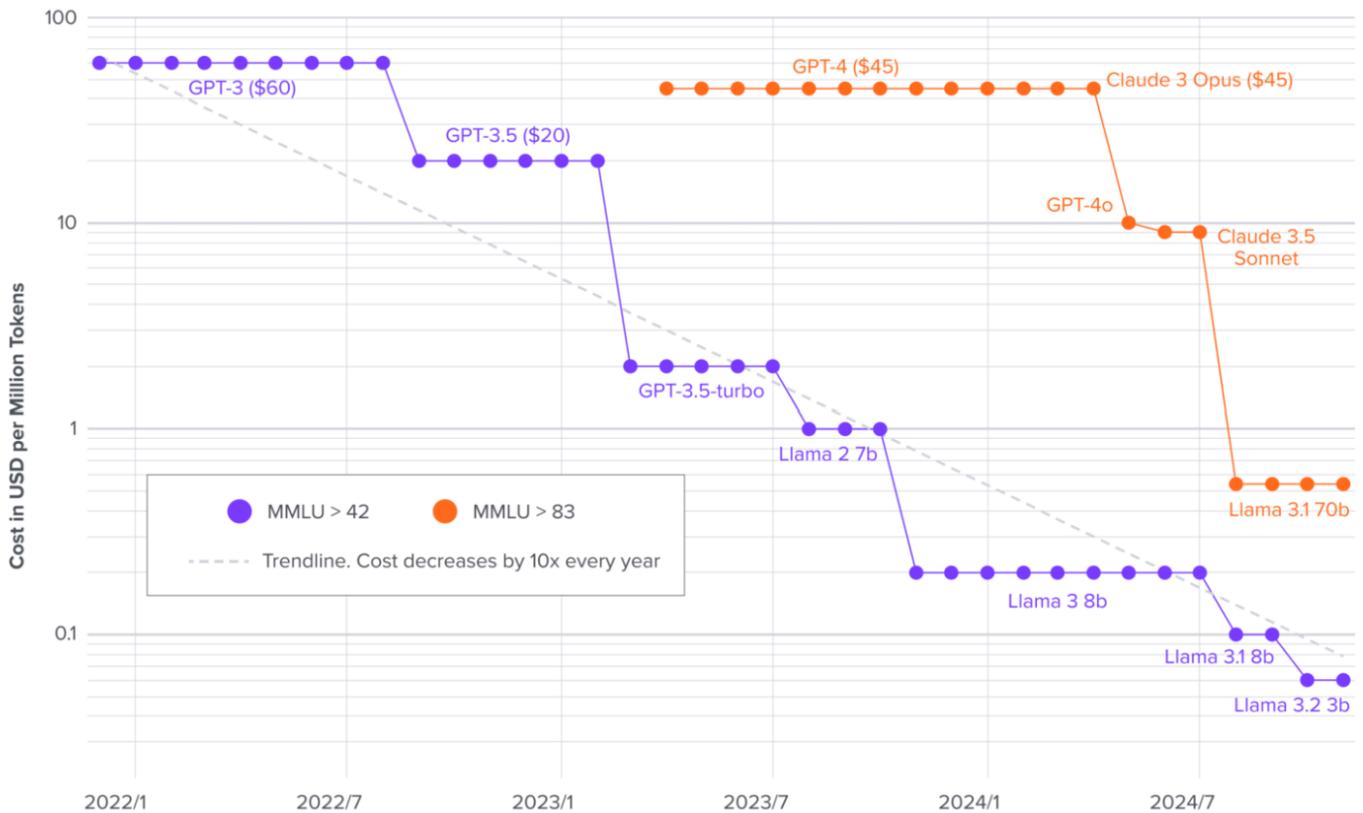


一、AI 应用有望带动推理算力需求，云厂商持续加码 ASIC

1.1 推理降本趋势明显，ASIC 方案性价比凸显

从 chatgpt3 发布以来，大语言模型的推理的成本以指数级别下降，单美元可以生成的 token 数量持续增长。同时模型的能力也持续增加。相较于最早的 chat gpt3，目前的主流模型都可以通过更低的推理成本达到更高的智能水平。在 2021 年 1 月，GPT3 是唯一可以达到 MMLU 42 分的大语言模型，当时百万 token 的成本在 60 美元，截至 2024 年 11 月，由 together.ai 提供的 Llama 3.2B 可以同样达到 MMLU 42 分的水平，但百万 token 的成本以及降低到 0.06 美元。而可以达到 MMLU 83 分的大语言模型中，Llama 3.1 70B 截至 2024 年 11 月百万 token 的成本已经小于 1 美元。根据 A16z Infrastructure 测算，推理成本每年降低幅度约 10 倍。

图表1：大模型推理成本按照指数下降，每年降低幅度约 10 倍

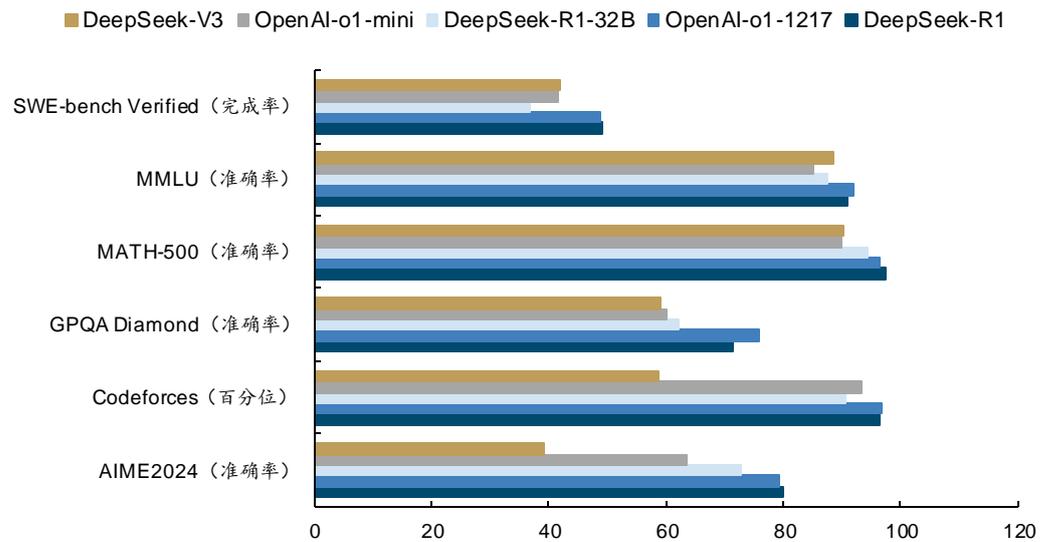


来源：A16z Infrastructure，国金证券研究所

另外我们注意到，目前模型的能力提升，除了模型的预训练以外，推理的算法升级也不断落地，通过包括强化学习、MOE 等方法提升模型推理能力。我们认为：模型的推理成本快速降低，有望带动应用的爆发，而应用爆发将带动更多的推理算力需求，同时推理算法的迭代也带动更多算力需求。我们认为大量的推理算力以及降本诉求将有效带动 ASIC 的需求增长。



图表2: Deepseek 依靠推理端算法优化可以实现较好智能程度 (单位: %)



来源:《DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning》, 国金证券研究所

ASIC 相比 GPU, 主要优势在于性价比。ASIC 采用定制化设计, 可以针对云厂商的业务, 以及模型做定制开发, 将其中常用的算子直接固化到硬件当中, 可以大幅提升运算效率, 同时降低功耗。另外 GPU 主流厂商如英伟达产品具有较高毛利率, 采用 ASIC 在单价上也有望降低。

图表3: ASIC 相比 GPU 牺牲灵活性, 但提高效率

	ASIC	GPU
效率	较高	中等
灵活性	低	高
初始成本	高	可变

来源: Bitdeer 网站, 国金证券研究所

1.2 谷歌、亚马逊加码 ASIC, 其他 CSP 有望跟进

目前北美云厂商积极布局 ASIC, 已经实现规模化生产的 ASIC 包括谷歌的 TPU、亚马逊的 Trainium 与 Inferentia、微软的 Maia 以及 Meta 的 MTIA, 其中谷歌和亚马逊的产品在 2025 年已经开始实现较大规模出货。我们认为其他云厂商 ASIC 在成熟度逐渐提升后, 未来也具备较大的增长空间。另外 Openai 等厂商也在积极布局 ASIC, 未来有望进入量产。

谷歌 TPU 在 24 年已经推出第六代产品当中用于推理的 v6e 产品, 与 TPU v5e 相比, Trillium TPU 每一芯片峰值计算效能提升了 4.7 倍, 为相当惊人的表现。TPU V6e 将 HBM 的容量与带宽提升 1 倍, 芯片间互连网络带宽也提高了一倍。此外, v6e 还配备了第三代 SparseCore, 这是处理超大嵌体的专用加速器, 常用于处理进阶排名与推荐工作负载。V6e 可以更快速地训练下一代基础模型, 并以较短的延迟时间与较低成本提供模型服务。与 TPU v5e 相比, TPU V6E 的能源效率高出 67%。

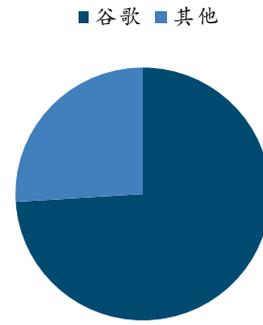
从数量来看, TPU 是目前 ASIC 当中最为主要的产品之一。根据 Digitimes 测算, 2023 年 TPU 在 ASIC 的市占率约 71%, 2024 年全球 ASIC 出货达到 345 万颗, 其中 TPU 占比提升至 74%。根据 Omidia 测算, 2024 年 TPU 的销售额在 60~90 亿美元。



图表4: v6e 在算力上有明显提升

	v4	v5e	v5p	v6e
pod 芯片数	4096	256	8960	256
fp16 算力 (Tflops)	275	197	459	926
HBM (GB)	32	16	95	32
HBM 带宽 (GB/s)	1228	820	2765	1640
芯片互联带宽 (Gb/s)	2400	1600	4800	3584

图表5: 24 年全球 ASIC 出货 345 万颗, 谷歌市占率 74%



来源: 谷歌网站, 国金证券研究所

来源: Digitimes, 国金证券研究所

亚马逊 Trainium2 已经推出, 2024、2025 年出货增长迅猛。根据 Trendforce, AWS 的出货成长力道强劲, 24 年年增率突破 200%。预计 2025 年 AWS 出货量将成长 70% 以上, 并更聚焦往 Trainium 芯片发展, 投入 AWS 公有云基础设施及电商平台等 AI 应用。

AWS Trainium2 芯片的效能比第一代 Trainium 提升高达 4 倍。以 Trainium2 为基础的 Amazon EC2 Trn2 实例专为生成式 AI 而建置, 是用于训练和部署具有数千亿到数万亿以上参数的模型的最强大 EC2 实例。Trn2 实例的价格效能比目前一代 GPU 型 EC2 P5e 和 P5en 实例更好 30-40%。Trn2 实例配备 16 个 Trainium2 芯片, 这些芯片透过 NeuronLink 实现互连。Trn2 UltraServer 是全新的 EC2 产品, 非常适合需要比独立 EC2 实例所能提供更多内存和内存带宽的最大型模型。UltraServer 设计使用 NeuronLink 将四个 Trn2 实例中的 64 个 Trainium2 芯片联机至一个节点中。对于推理而言, UltraServer 可协助提供业界领先的响应时间, 进而创造出最佳的实时体验。对于训练而言, 与独立实例相比, UltraServers 会透过更快的协同通讯来提高模型平行性的模型训练速度和效率。

目前包括 Adobe、AI 新创公司 Poolside、数据平台服务 Databricks 以及高通都通过 Trainium2 处理器训练其 AI 模型, 其中, 高通在云端计算 AI 模型后再将其传送至边缘端。另外苹果也在采用亚马逊 ASIC 芯片提供的服务, 应用于 Siri, Apple Maps 和 Apple Music。苹果使用亚马逊的 Inferentia 和 Graviton 芯片来服务搜索服务。

图表6: 亚马逊 Trainium2 包括 Trainium2 与 Trainium2 Ultra 两个版本

	Trainium2	Trainium2 Ultra
bf16 算力 TFLOPS	667	667
FP8 算力 TFLOPS	1299	1299
HBM 容量 GB	96	96
HBM 带宽 GB/s	2900	2900
功耗	500	500
散热	风冷	风冷
Scale up 技术	NeronLinkv3	NeronLinkv3
Scale up 带宽	512	640
Scale up 规模	16	64
Scale out 技术	以太网	以太网
Scale out 带宽	最高 800	200

来源: Semianalysis, 国金证券研究所

Meta ASIC MTIA 已经推出第二代, 目前主要用于模型推理, Meta 预计 26 年将有用用于训练的 ASIC 推出。MTIA 使用台积电 5nm 工艺制造, 具有 90W 的热设计功耗 (TDP), 显著降低了功耗需求, 使其在数据中心的更易管理。同时, 该处理器采用 16 通道 LPDDR5 内存, 配备 128GB 的内存配置, 为高效数据处理提供了强有力的支撑。根据 Meta 24Q4 业绩会, 目前 MTIA 主要用于模型推理以及推荐类任务, 预计 26 年 MTIA 将用于训练当中。



微软也已经推出 MAIA 100 ASIC,采用 5nm 制程, 64GB HBM2E。考虑微软 Azure 云具有较大体量, 同时微软具备 Copilot 等终端 AI 应用场景, 我们预计未来微软 ASIC 也将发力。

图表7: Meta、微软已经分别推出 ASIC MTIA、Maia

	META MTIA v2	微软 Maia 100
制程	5nm	5nm
fp16 算力 TFLOPS		177
fp8 算力 TFLOPS		354
显存容量	128GB LPDDR5	64GB HBM2E
显存带宽 GB/s	204.8GB/s	1.8TB/s
互联带宽	8*32GB/s	8*32GB/s

来源: Meta 网站、hotchips 网站, 微软, 国金证券研究所

除了北美四大 CSP 以外, 我们预计未来领先的模型厂商如 openai、deepseek 等也将开发 ASIC。根据《经济日报》, openai 预计将采用台积电 3nm 以及 A16 制程生产 ASIC, 目前已经向台积电预定 A16 产能, 预计 openai 的 ASIC 将在 26 年年底进入量产。OpenAI、甲骨文和软银合作建立星际之门项目, 预计投资 5000 亿美元用于人工智能基础设施建设, 考虑星际之门项目的高投资, 我们预计 openai 的 ASIC 未来也将有较大规模出货量。

二、看好 ASIC 设计服务公司、以太网、AEC、PCB 相关厂商

2.1 ASIC 设计服务公司业绩进入放量期, 关注 ASIC 项目 design wins 变化

目前云厂商所采用的设计服务公司主要包括美系厂商博通、Marvell, 以及台系厂商世芯、创意电子, 未来来看联发科也有望在 ASIC 获得一定份额。在美国制裁加剧情况下, 我们认为未来国内云厂商有望更多采用本土设计企业进行 ASIC 设计。

设计服务公司主要承担芯片的前道设计、后道封装设计, 并且对 ASIC 进行量产。设计服务公司的主要壁垒主要在关键 IP (如 HBM PHY 的 IP、高速 Serdes IP 等) 以及先进封装的经验。

图表8: 定制 ASIC 业务的技术关键在于关键 IP 与先进封装技术



优化的数据基础IP



高速SerDes互联



先进多芯片封装技术

来源: Marvell 网站, 国金证券研究所

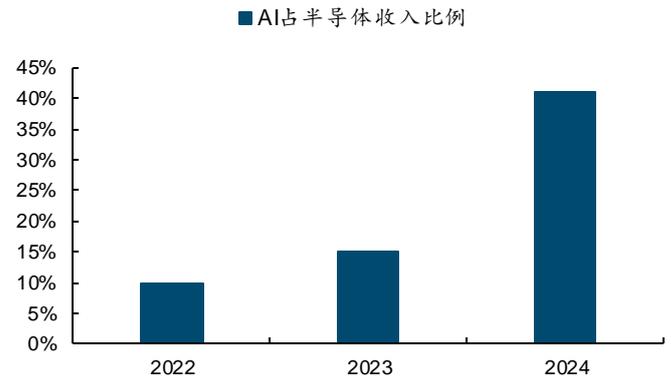
博通是目前最主要 ASIC 设计服务公司, 已经获得五个客户, 25 年预计有三个客户的 ASIC 都将量产。FY24 博通 AI 收入为 122 亿美元, 同比增长 220%, 达到博通半导体收入占比的 41%。根据博通 24 年 3 月的投资者交流会议, 博通已经获得三个 ASIC 客户, 根据博通 24 年 12 月业绩会, 博通 ASIC 获得了额外两个客户, 并且预计 27 年 AI 的 SAM 将达到 600~900 亿美元。



图表9: 博通 25 年将实现三家 ASIC 客户产品量产

	2024	2025	2026
客户 1	量产	量产	量产
客户 2	量产	量产	量产
客户 3		量产	量产

图表10: 博通 AI 收入占半导体收入比例快速提升



来源: 博通网站, 国金证券研究所

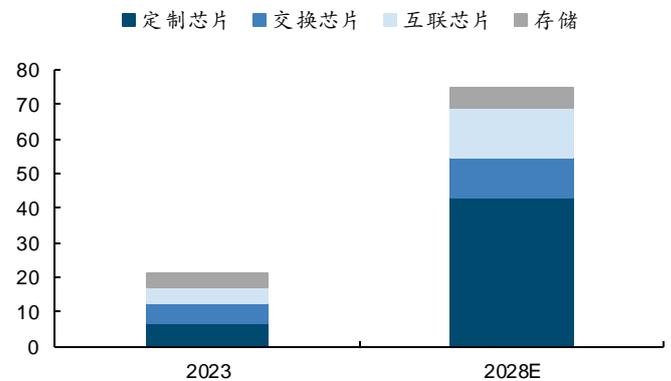
来源: 博通网站, 国金证券研究所

Marvell 是美国重要的 ASIC 设计服务厂商之一, 目前已经获得两个客户的 AI ASIC 项目, 另外获得了一家客户的 arm CPU 的项目。公司指引 FY25、FY26 AI 收入分别为 15、25 亿美元, 其中定制化芯片营收分别为 5、10 亿美元, 主要来自定制化 AI 芯片。根据 Marvell 测算, 2023 年 Marvell 在数据中心市场的 TAM 为 210 亿美元, 其中定制化芯片市场为 66 亿美元, 预计 2028 年 Marvell 在数据中心的 TAM 为 750 亿美元, 其中定制化芯片市场为 429 亿美元, 23~28 年定制化芯片的 CAGR 为 45%。

图表11: Marvell 已经获得三家客户的 ASIC 项目

	产品进展
客户 A	AI 训练 ASIC 24 年已经量产, 推理 ASIC 25 年计划量产
客户 B	Arm CPU 24 年已经量产
客户 C	AI ASIC 预计 26 年量产

图表12: Marvell 预计 28 年定制化芯片 TAM 达到 429 亿美元 (单位: 十亿美元)



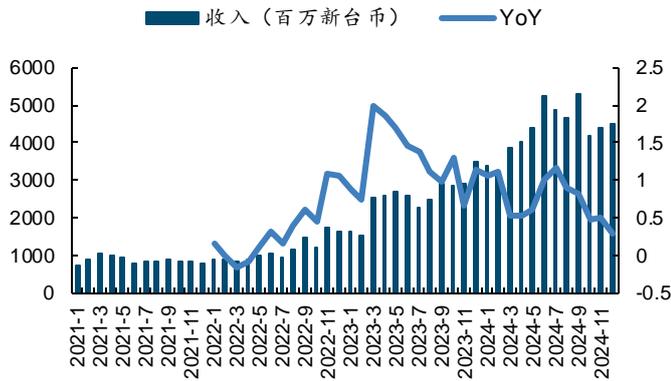
来源: Marvell 网站, 国金证券研究所

来源: Marvell 网站, 国金证券研究所

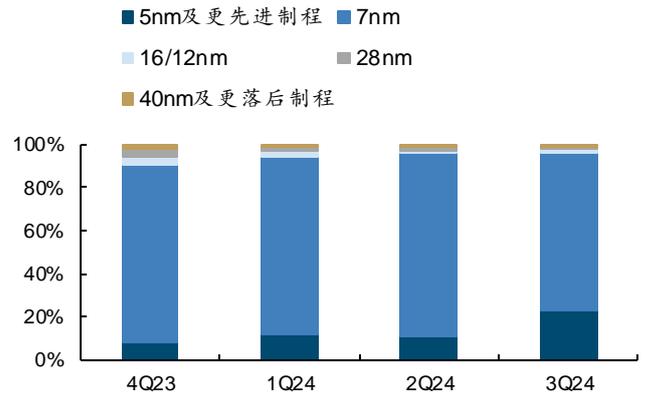
台系 ASIC 设计服务公司主要包括世芯、创意电子以及联发科, 世芯有望获得亚马逊 ASIC 项目, 联发科有望获得谷歌 ASIC 项目。根据《经济日报》, 亚马逊 AWS 于 12 月 3 日发表新一代 AI 自研芯片 Trainium3, 将于 2025 年底上市, 其中项目设计服务由世芯负责, 并采用 3nm 制程。世芯 23、24 年业绩大增, 24 年收入合计为 520.39 亿新台币, 同比+70.53%。主要受惠亚马逊 7nm 制程 AI 芯片项目, 目前亚马逊 AWS 对其业绩营收贡献度超过一半, 是举足轻重的大客户。不过因第 4 季 7nm 制程 AI 芯片进入出货尾声, 导致世芯营收较第 3 季下滑。



图表13: 世芯营收高速增长, 24年同比+70.5%



图表14: 世芯 7nm 及更先进制程项目收入占比超 90%

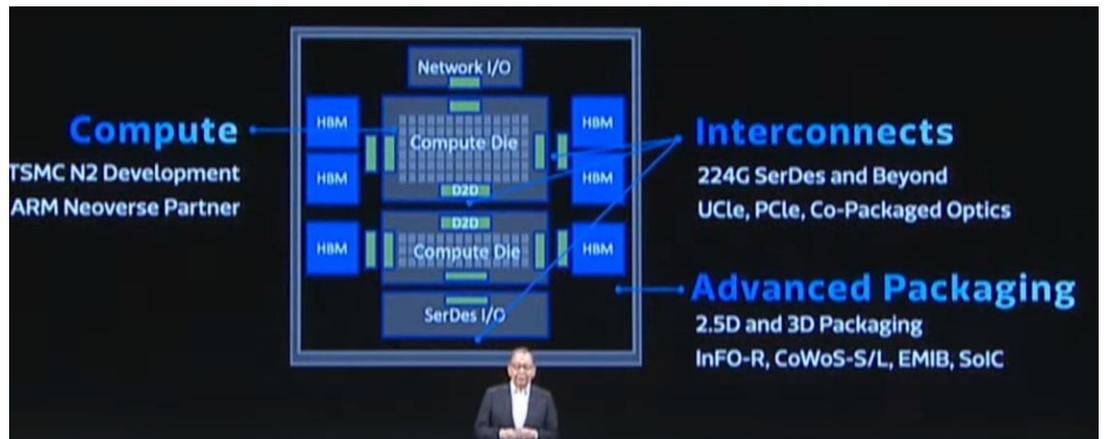


来源: 世芯网站, 国金证券研究所

来源: 世芯网站, 国金证券研究所

联发科积极布局 ASIC, 未来有望进入谷歌。根据《中时新闻网》, 联发科与 Google 推理 TPUv7 SerDes 的新案, 预期将带动 2025 至 2026 年的营运成长。联发科在 ASIC 领域已经具备一定技术积累, 包括 224G SERDES、PCIe 5.0/6.0/7.0、800G/1.6T Retimer、HBM 解决方案、CoWoS-S/CoWoS-L 先进封装能力。

图表15: 联发科在 Computex 2024 上展示的 ASIC 设计方案

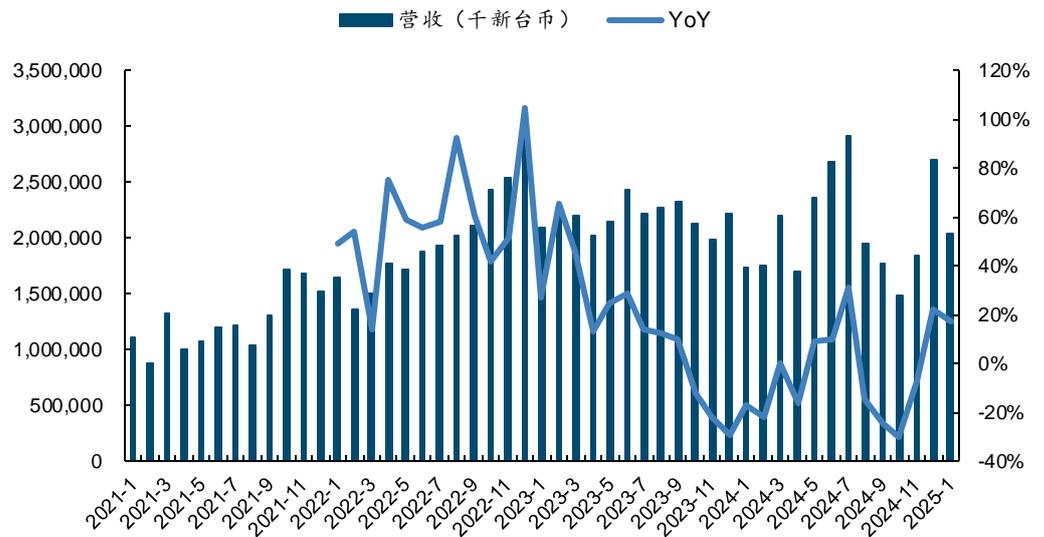


来源: Computex 网站, 国金证券研究所

创意电子在 HBM PHY 芯片具有较多设计项目, 另外在 AI ASIC 与微软合作。创意电子 HBM3E IP 方案在技术上领先市场, 并已通过台积电先进制程技术的硅验证, 包括 N7 / N6、N5 / N4P、N3E / N3P 制程。此外, 创意电子的 IP 亦通过所有主流 HBM3 供应商的硅验证, 确保其解决方案的兼容性与稳定性, 并且积极与 HBM 记忆体供应商如美光科技合作, 致力于开发下一代 HBM4 IP, 以应对未来更多元化的应用需求。在 AI ASIC 项目当中, 公司主要与微软合作, 目前主要产品包括 AI ASIC MAIA 100 以及 arm CPU Cobalt。根据《中时新闻》, 今年出货仍以第一代为主, Maia2 要等到 2026 年首季才放量, CPU 部分将持续出货原先的 Cobalt 100, 预计 25 年第四季结束出货。



图表16: 创意电子 24 年实现营收 250.44 亿新台币, 同比-4.6%



来源: 创意电子网站, 国金证券研究所

我们认为, ASIC 在推理算力占比提升的大背景下, 需求增长的趋势明确, 但需要关注行业竞争格局的变化。目前北美云厂商在 ASIC 领域都开始逐步引入多家设计服务公司, 需要密切关注每个芯片项目的获得者。考虑联发科在谷歌进展较缓慢, 需要 tpu v7 才有望实现产品导入, 以及博通在 meta XPU 目前仍然为唯一供应商, 我们认为博通在 ASIC 设计服务公司当中最为受益。而世芯在亚马逊有望重新拿到 ASIC 订单, 考虑亚马逊的较大 ASIC 需求, 26 年也有望具备较大增长动能。对于 ASIC 设计服务公司, 我们建议关注: 博通、世芯、联发科、创意电子、Marvell。

图表17: 北美 CSP 引入多家 ASIC 供应商

CSP	产品	供应商
谷歌	TPU	博通、联发科
	CPU	Marvell、创意电子
亚马逊	XPU (训练、推理)	Marvell、世芯
	消费型	世芯、创意电子
微软	XPU (Maia)	Marvell、创意电子
	CPU (Cobalt)	创意电子
Meta	XPU	博通
	消费型	联发科、创意电子

来源: 中时新闻, 国金证券研究所

我们认为 A 股公司当中翱捷科技、芯原股份有望受益国内 ASIC 趋势。翱捷科技能够按照客户需求提供芯片架构定义、芯片设计、封装测试、芯片量产、配套软件开发等各环节的部分或全部服务。可满足客户高端系列产品对开发更高性能产品的需求, 以及对供应链安全方面的本地化需求。翱捷科技 24H1 实现芯片定制业务收入 2.26 亿元, 较 23H1 的 1.18 亿元接近翻倍, 客户群体包括人工智能算法企业、互联网企业、大数据企业等。

芯原股份可以为客户提供从芯片定义到设计流片的全流程设计服务 (含软件设计服务), 还提供经过验证的平台化芯片解决方案以帮助客户快速完成芯片设计。此外, 芯原还可帮助客户委外完成晶圆制造、封装和测试等, 并提供优质的生产管理服务。目前芯原股份可以芯片设计公司和 IDM、系统厂商、大型互联网公司、云服务提供商在内的各种客户提供高效经济的半导体产品替代解决方案, 其终端应用可覆盖消费电子、汽车电子、计算机及周边、工业、数据处理和物联网等广泛的行业应用市场。

2.2 ASIC 采用以太网组网, 以太网交换芯片、白盒交换机厂商有望受益

由于 Infiniband 主要厂商 Mellanox 被英伟达收购, ASIC 主要采用公开的通信协议以太

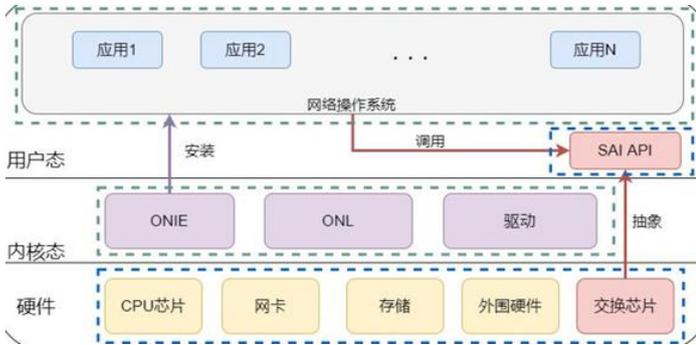


网进行组网。我们认为 ASIC 将带动白盒交换机需求，以及以太网通信芯片需求。

白盒交换机将硬件与软件分离，下游数据中心客户可选择为交换机安装外部操作系统或在交换机厂商已提供开放式操作系统基础上开发上层应用软件，并实现对交换机的统一部署与维护，极大提高数据中心运维效率。

图表18: 白盒交换机解耦上层应用与底层硬件

图表19: 开放与解耦是数据中心网络发展方向



来源:《白盒交换机技术白皮书》，国金证券研究所

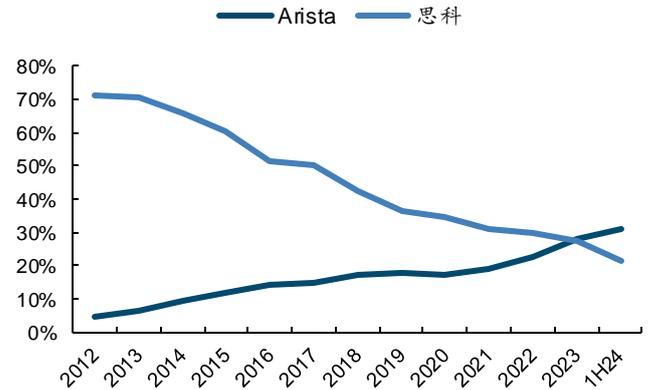
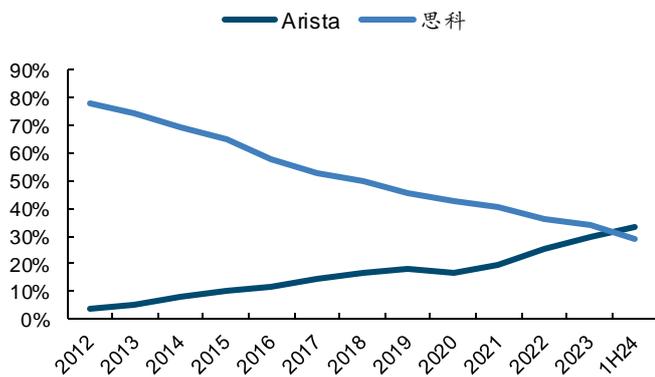
来源: C114 通信网，国金证券研究所

交换机软硬件解耦趋势下，传统黑盒方案厂商思科市占率下降迅速。从规模的角度看，Arista 在数据中心 10G 以上速率的交换机市场份额在过去 12 年中提升了 26.4%，已经在 24 年上半年反超思科；从端口数量的角度看，Arista 市场份额在近三年提升较快，并于 2023 年完成对于思科市场份额的反超。

而在数据中心高速以太网交换机市场 (100G/200G/400G)，Arista 于近三年持续保持行业领先地位，端口出货量不断拉大与思科的差距，2023 年，Arista 份额达到 45%，是思科的两倍之多。

图表20: Arista 数据中心交换机按营收市占率超过思科

图表21: Arista 数据中心交换机按端口数市占率超过思科

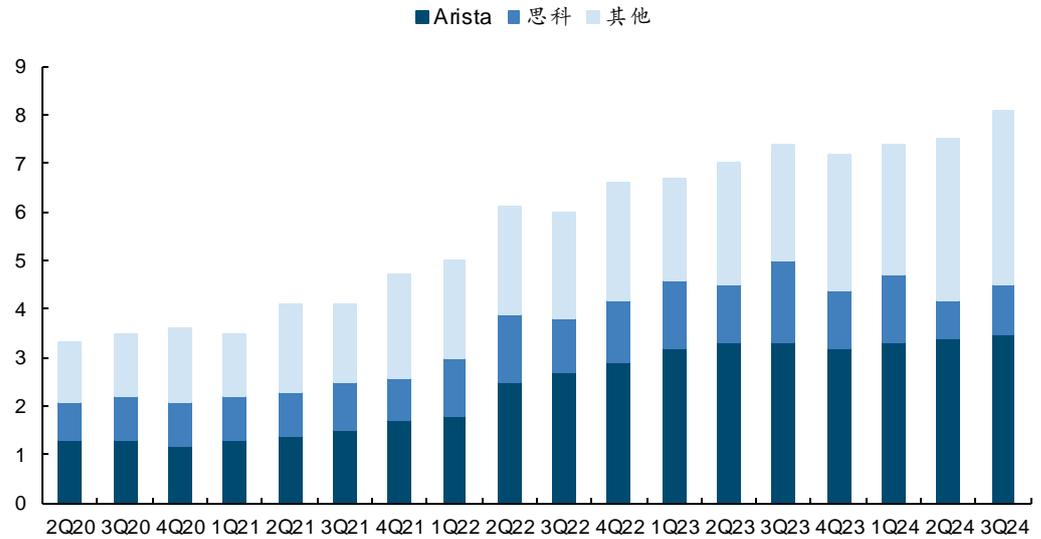


来源: Arista 网站，国金证券研究所

来源: Arista 网站，国金证券研究所



图表22: Arista 数据中心高速以太网 (100G/200G/400G) 端口数出货量已超思科两倍 (单位: 百万)



来源: Arista 网站, 国金证券研究所

白盒交换机厂商当中, OEM 厂商如 Arista 凭借软件的较高壁垒, 以自研操作系统+硬件的产品形态进行出售, 具备高毛利率, 白牌交换机厂商则主要是进行硬件组装, 毛利率较低。白牌交换机厂商较为依赖云厂商客户自身操作系统的能力。短期来看, Arista 凭借具备良好生态的成熟的自研操作系统有望继续维持当前份额, 长期看, 出于成本考虑我们认为云厂商采用自研系统采购白牌交换机的比例有望逐渐扩大。建议关注:Arista、天弘科技、智邦。

以太网交换机通常由以太网交换芯片、CPU、PHY、PCB、接口/端口子系统等组成, 其中以太网交换芯片作为以太网交换设备中最核心的部件, 主要用于交换处理大量数据和报文转发, 是针对数据传输和网络应用进行专门优化的芯片。以太网交换芯片内部的逻辑通路由数百个特性集合组成, 在协同工作的同时保持极高的数据处理能力, 因此其架构实现具有复杂性。CPU 是用来管理登录、协议交互的控制的通用芯片, PHY 用于处理电接口的物理层数据。部分以太网交换芯片还将进行协议控制管理的 CPU 和处理电接口物理数据的 PHY 集成在以太网交换芯片, 进一步提高了以太网交换芯片的技术密集度和产品复杂度。

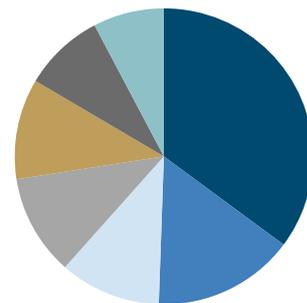
以 400G 交换机 N9510-64D 为例, 产品搭载的主要芯片包括英特尔 Xeon D-1627 CPU、博通 Tomahawk4 交换芯片、8GB DDR4 内存以及 240GB 闪存。根据中商产业研究院, 交换机主要由芯片、光器件、插接件、阻容器件、壳体、PCB 等资源组成。其中芯片成本占比最高, 达 32%。其次分别为光器件、插接件、阻容器件、壳体、PCB, 占比分别为 14%、10%、10%、8%、7%。

图表23: 交换机主要芯片为交换芯片

端口	64*400G QSFP-DD
CPU	Intel Xeon D-1627
内存	DDR4 8GB
闪存	240GB
交换芯片	BCM56990 Tomahawk4

图表24: 交换机当中芯片成本占比达 32%

■ 芯片 ■ 光器件 ■ 插接件 ■ 阻容器件 ■ 壳体 ■ PCB ■ 其他



来源: FS 网站, 国金证券研究所

来源: 中商产业研究院, 国金证券研究所



目前交换芯片主要分为商用芯片，以及部分交换机厂商的自用芯片。其中商用芯片厂商包括博通、Marvell、英伟达等，自用厂商则以思科、华为为主，自研芯片用于自己交换机。我们认为，随着白盒厂商如 Arista、天弘、智邦等厂商在数据以太网交换机市场快速增长，有望带动商用芯片需求快速增长，博通、Marvell 等商用芯片厂商有望充分受益。建议关注：博通、Marvell。

2.3 ASIC 带动 AEC 需求加速，云厂商开始导入 ASIC

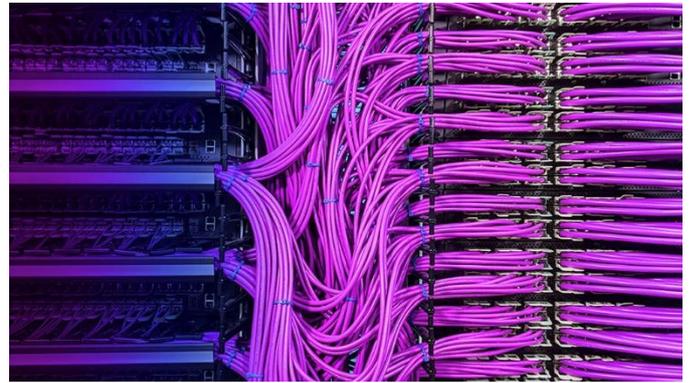
我们认为 ASIC 需求增长有望带动 AEC 需求，目前 AEC 领先厂商如 Credo 已经在北美 CSP 实现批量出货。根据亚马逊，支持 2.5 米传输的 100G 速率 DAC 外径为 6.7 毫米，而支持 2.5 米传输的 400G 速率的 DAC 外径达到了 11 毫米，使得云厂商在排列数据线时困难较大。另外，DAC 的更大的外径也意味着更大的弯曲半径，使得整个机架的占地面积和使用的空间更大。

目前高速铜连接的创新解决方案为有源电缆 AEC。AEC 相比 DAC，在铜缆两端加装芯片，进行信号恢复，可以减少高速信号在铜线传输时产生的损耗和衰减。因此 AEC 外径较传统 DAC 更小，所占用空间也更低。在大规模 AI 集群的搭建当中，我们认为由于 AI 集群的互联密度较云计算明显提升，AEC 具备更小的外径的特点更加适合大规模组网的布线。另外在短距离传输当中，AEC 与采用光模块、光纤的光通信方案相比，AEC 在短距离传输具备低成本、低能耗、低维护成本的优势。

图表25: DAC 直径、所占空间更大，排线困难



图表26: AEC 直径、所占空间更小



来源：亚马逊网站，国金证券研究所

来源：Credo 网站，国金证券研究所

图表27: AEC 在短距离高速率传输优势较 DAC 和光通信方案明显

	DAC	AEC	光通信方案
速率	400G 及以下	目前最高量产产品达 1.6T	目前最高量产产品达 800G
传输距离	一般 7 米以内	一般 7 米以内	可以支持长距离传输，数据中心主流产品 SR 支持 300 米传输，DR 支持 500 米传输
外径	大，400G 2.5 米是 11 毫米	较小，400G 3 米是 6.8 毫米	小，400G 速率仅 3 毫米
占用空间	大，400G 时是 AEC 方案的 4 倍	较小	小，400G 时是 AEC 的 1/4
成本	低，400G 时是 AEC 的一半左右	较低	高，400G 时是 AEC 方案的两倍左右
功耗	低，400G 时是 AEC 的 1/4 不到	中	高，400G 时是 AEC 方案的两倍以上

来源：Credo 网站，国金证券研究所整理

我们认为未来随着数据中心网络传输速率不断提升，在短距离传输当中 DAC 将面临较大困难，AEC 等创新连接方式有望对 DAC 实现替代。根据 Lightcounting23 年 12 月的测算，AOC、DAC 和 AEC 市场 23 年为 12 亿美元，2028 年有望达到 28 亿美元，其中 AOC、DAC、AEC 市场 23~28 年 CAGR 分别为 15%、25%、45%。



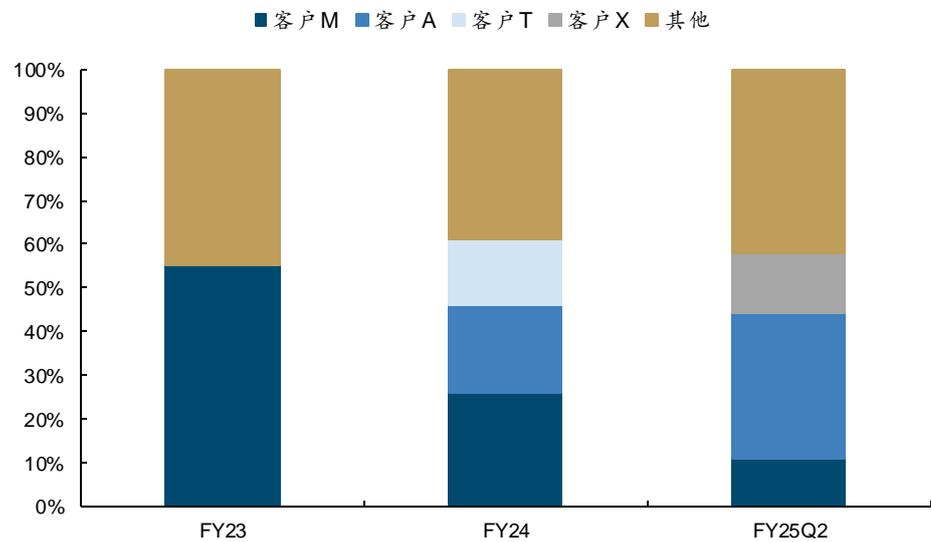
图表28: AOC、DAC、AEC 28 年市场有望达 28 亿美元, AEC 增速最快



来源: Lightcounting, 国金证券研究所

目前 Credo 是 AEC 行业的主要厂商, 部分北美头部云厂商已经导入 Credo 的 AEC 产品。目前 Credo 已经与北美头部云厂商都有接触。微软与亚马逊已经产生销售额, 公司与 XAI 也进行深度合作。公司 FY25Q2 (24.8~24.10) 有三家客户收入占比超过 10%, 其中微软收入占比达到 11%, 公司第二家 AEC 大型客户收入占比达到 33%, 另外公司本季度有一家新的大型客户导入, 收入占比达到 14%。

图表29: Credo 客户拓展顺利, FY25Q2 三家头部企业成为 Credo 营收占比超 10% 客户



来源: Credo 公告, 国金证券研究所

AEC 的关键在于铜缆两端的芯片, 我们认为具备高速 DSP 设计能力的公司具备切入 AEC 市场能力, 包括 Marvell、Maxlinear、Astera Lab 等。我们认为 Credo 具备一定先发优势, 其他厂商未来也有望实现终端客户导入。我们看好 Credo, 建议关注: Astera Lab、Marvell。

图表30: 主要厂商 AEC 产品进展

公司	AEC 产品进展
Credo	产品包括 400G~1.6T AEC 产品
Marvell	产品包括 400G~1.6T AEC 所用的 DSP 芯片
Maxlinear	推出了使用 Maxlinear DSP 的 AEC 设计工具



Astera Lab | 推出了 200G~800G AEC 产品

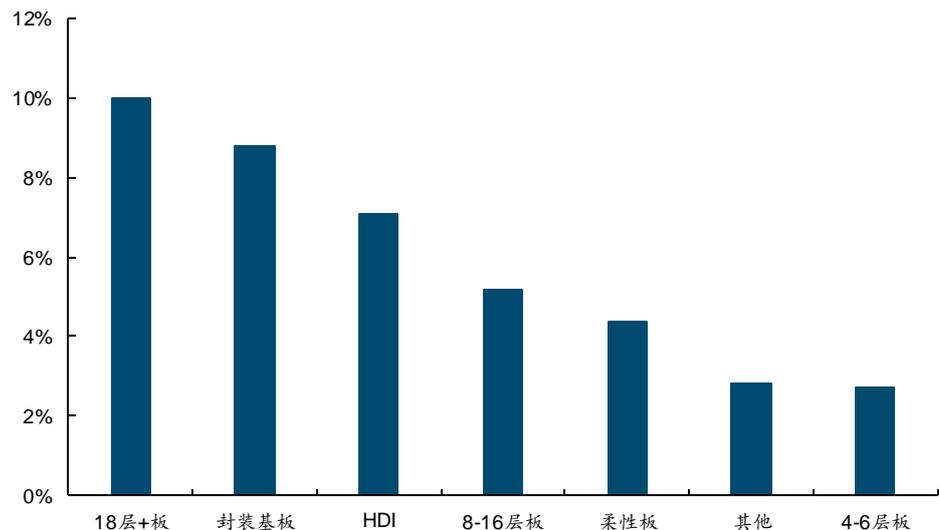
来源：各公司网站，国金证券研究所

国内连接器厂商目前也积极布局 AEC 产品线。立讯精密进展较快，立讯精密构建了柜内互连、柜间互连、服务器、交换机、基站射频等完整解决方案服务体系，送样测试与量产产品均收获海内外头部客户的充分认可与高度评价。目前立讯精密已经推出 DAC、ACC、AEC 不同产品方案，传输总带宽可以达到 400G、800G、1.6T。瑞可达、博创科技也在积极布局 AEC 产品线。瑞可达 24 年 7 月 29 日在投资者互动平台表示，AEC 高速组件是未来重点布局和发展的产品。博创科技 24 年 5 月 23 日在投资者互动平台表示，800G 高速有源铜缆已具备量产条件。我们认为立讯精密作为国内连接器龙头厂商之一，与海外客户具有较良好关系，有望充分受益。我们看好：立讯精密，建议关注：瑞可达、博创科技。

2.4 PCB: ASIC 的单位 PCB 耗用量更高，多 die 合封带来封装基板市场需求扩容

PCB 作为电子元器件之母，产品设计升级必然会导致 PCB 产品升级，ASIC 兴起也不例外，我们认为 ASIC 兴起将带动产业链显著升级。我们认为 ASIC 芯片将显著带动高速通信需求发展，而高速通信对于 PCB 的要求是显著提升的，根据世运电路引用的 Prismark 统计，2023~2028 年复合增长较高的细分领域为 18 层+板、封装基板、HDI，复合增速分别达到 10.0%、8.8%、7.1%，其中 18 层+板和 HDI 主要市场增量来自于高速通信领域 PCB 板级需求，而封装基板则来自于高速通信领域载板级需求。

图表31: 2023~2028 年全球 PCB 细分领域复合增速预测

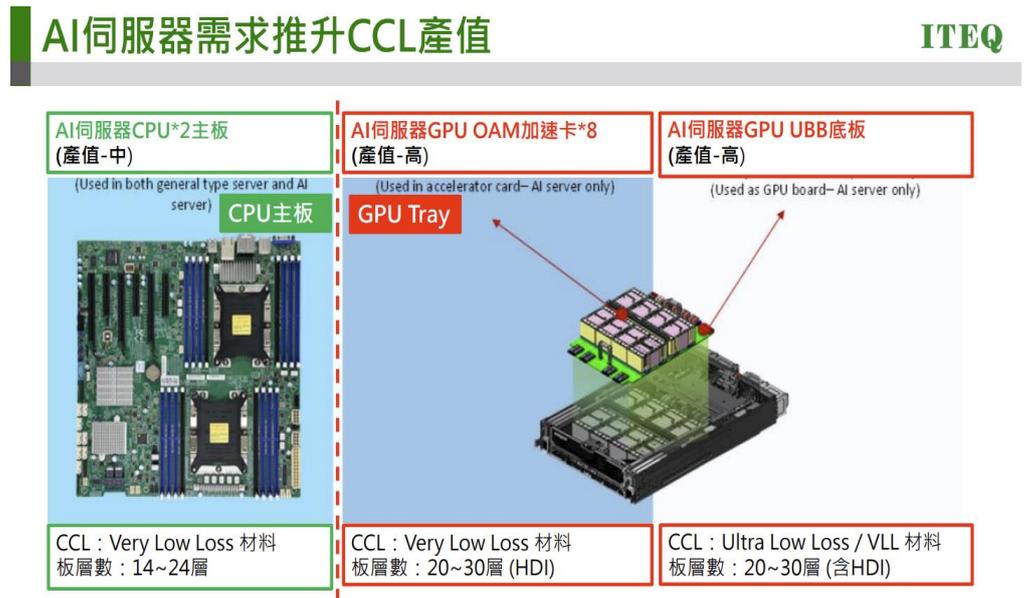


来源：世运电路公告，Prismark，国金证券研究所

- 1) ASIC AI 服务器驱动的 PCB 升级主要来自两个方面的因素，其一为 AI 对算力 PCB 提出了更高要求，其二数据量持续增长带来基础建设持续升级，AI 服务器相对传统普通服务器新增了 GPU 板组，而 GPU 对连接带宽要求高导致所需的 PCB 和 CCL 要求均有提升，一方面 PCB 层数从以往的 14~24 层提升至 20~30 层，同时 AI 领域开始加大对 HDI 这一 PCB 行业传统技术的应用，另一方面对于 CCL 行业，ASIC AI 服务器 CCL 等级将从原先的 Very low loss 提升至 Ultra low loss 或 Super ultra low loss，HDI 工艺也对 CCL 材料的 CTE 等指标提出了更高要求。



图表32: 2023~2028 年全球 PCB 细分领域复合增速预测



13

来源: 联茂官网, 国金证券研究所

图表33: 亚马逊 Trainium2 服务器示意图



图表34: 谷歌自研 AI ASIC 芯片 TPU



来源: 亚马逊官网, 国金证券研究所

来源: 谷歌官网, 国金证券研究所

- 2) ASIC 芯片采用多 die 形式, 显著提升对载板的需求。AI 计算芯片为了达到更多的浮点计算功能, 往往会采用 Chipllet 的形式将多个 die 封装在一起, 如亚马逊 Trainium2 芯片采用双 die 的架构。在 ASIC 这类高端芯片的带动下, 先进封装市场规模不断扩容, 根据 Yole 预测, 先进封装市场在 2024~2029 年间复合增长率将达到 10.7%, 至 2029 年市场规模将达到 699 亿美元, 规模将超过传统封装市场, 其中 Embedded Die (ED) 和 2.5D/3D 成为未来增长最快的封装工艺。多 die 合封的架构会使得承载的载板的面积显著增加、数据量增大也会使得整个载板的抗膨胀要求提升, 根据 yole 数据, 以 FCBGA 为代表的先进封装的成本结构中有 50%来自封装基板, 也正因如此, 先进封装的发展带动了封装基板显著增长, 从 2017 年以来封装基板的成长速度显著高于其他 PCB 板类型。

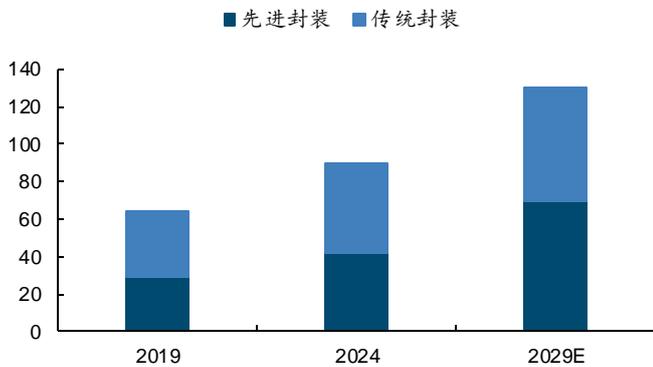


图表35: 亚马逊自研 AI ASIC 芯片 Trainium2



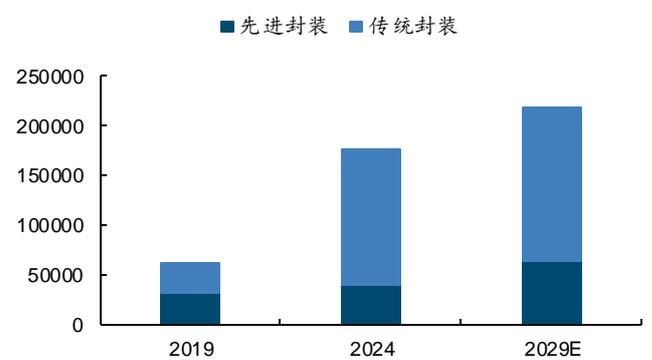
来源: 亚马逊官网, 国金证券研究所

图表36: 全球封装市场收入 (十亿美元)



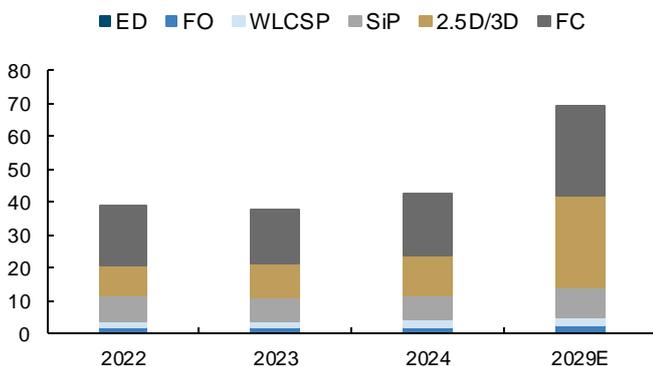
来源: Yole, 国金证券研究所

图表37: 全球封装晶圆产能耗用 (千片)



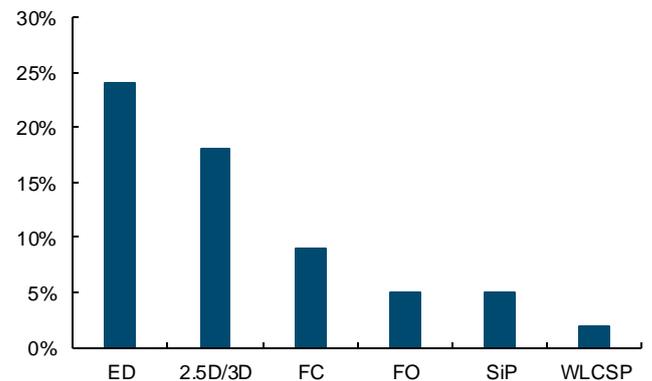
来源: Yole, 国金证券研究所

图表38: 全球先进封装市场按工艺分布 (十亿美元)



来源: Yole, 国金证券研究所

图表39: 2024~2029 年全球先进封装各工艺复合增速

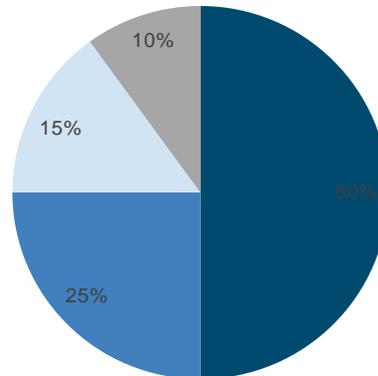


来源: Yole, 国金证券研究所



图表40：载板在FCBGA封装制造成本中占据50%的比例

■ 封装基板 ■ 组装工艺成本 ■ 组装材料成本 ■ Bumping



来源：Yole，国金证券研究所

三、投资建议

我们认为 Deepseek 等低成本高能力的开源模型的出现，有望带动 AI 应用爆发，带动推理算力的需求高速增长。ASIC 作为具备较高性价比的产品，有望在推理端得到广泛应用。目前北美 CSP 已经积极布局 ASIC，谷歌 TPU、亚马逊 Trainium 已经有较大规模出货，另外模型厂商如 Openai 也在积极布局 ASIC。ASIC 有望迎来高速增长。

我们看好 ASIC 设计服务公司、以太网白盒交换机厂商与以太网交换芯片厂商、AEC 厂商、PCB 厂商。

在 ASIC 设计服务领域，主要参与厂商包括博通、Marvell、世芯、联发科、创意电子等，目前北美云厂商 ASIC 开始逐步引入多个供应商，需要密切跟踪竞争格局的变化，建议关注：博通、世芯、联发科、翱捷科技、芯原股份、创意电子、Marvell。

在以太网交换机与以太网交换芯片领域，我们认为由于 ASIC 采用以太网组网，ASIC 的放量有望带动以太网在 AI 组网的渗透率和需求提升，我们看好白盒交换机厂商与以太网交换芯片厂商，建议关注：Arista、博通、天弘科技、智邦、Marvell。

我们认为 ASIC 有望带动 AEC 需求，目前北美 CSP 已经开始采用 AEC，我们认为 Credo 在 AEC 市场具备一定先发优势，我们看好：Credo、立讯精密，建议关注：Astera Lab、Marvell、瑞可达、博创科技。

我们认为 ASIC 兴起将带动 PCB 市场空间提升，主要体现在高速通信领域所用高多层高速板和 HDI 板，同时多 die 合封有望增加封装基板市场需求，建议关注：沪电股份、生益科技、联瑞新材、兴森科技等。

四、风险提示

AI 进展不及预期：如果 AI 进展不及预期，可能导致模型的应用需求不及预期、使用客户数量不及预期，影响 ASIC 的需求，导致相关公司业绩不及预期。

行业竞争加剧：AI 行业趋势明显，布局 AI 的厂商数量持续增多，将导致行业竞争加剧，影响相关公司业绩。

下游需求不及预期：ASIC 终端客户为云厂商，如果云厂商资本开支收缩、需求不及预期，将造成相关公司出货不及预期，导致相关公司业绩不及预期。



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。

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	北京	深圳
电话：021-80234211	电话：010-85950438	电话：0755-86695353
邮箱：researchsh@gjzq.com.cn	邮箱：researchbj@gjzq.com.cn	邮箱：researchsz@gjzq.com.cn
邮编：201204	邮编：100005	邮编：518000
地址：上海浦东新区芳甸路 1088 号 紫竹国际大厦 5 楼	地址：北京市东城区建国内大街 26 号 新闻大厦 8 层南侧	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