

互联网资本开支持续扩大，AI服务器景气度逐季提升 ——AI算力月度跟踪（202405）

评级：推荐(维持)

刘熹(证券分析师)

S0350523040001

liux10@ghzq.com.cn

最近一年走势



相对沪深300表现

表现	1M	3M	12M
计算机	1.5%	7.8%	-24.5%
沪深300	3.2%	9.3%	-7.1%

相关报告

《计算机事件点评：运营商AI服务器订单超预期，国产化不断提速（推荐）*计算机*刘熹》——2024-04-22

《计算机行业动态研究：阶跃星辰发布Step-2万亿参数模型，AI应用+算力景气上行（推荐）*计算机*刘熹》——2024-03-26

《AI算力月度跟踪（202403）：NVIDIA B200再创算力奇迹，液冷、光模块持续革新（推荐）*计算机*刘熹》——2024-03-24

《计算机行业事件点评：NVIDIA H20或将与国产算力同步高增（推荐）*计算机*刘熹》——2024-03-03

《计算机行业月报：英伟达业绩再超预期，国产算力需求再扩大（推荐）*计算机*刘熹》——2024-02-28

《计算机事件点评：央企AI专题会：推动AI生态“国家队”加快布局（推荐）*计算机*刘熹》——2024-02-22

《AI算力“卖水人”系列专题（1）：2024年互联网AI资本开支持续提升（推荐）*计算机*刘熹》——2024-02-19

核心观点：北美CSP厂商资本开支持续大幅增长，NV GB200等芯片迭代+供给提速，台股服务器厂商普遍展望2024年AI服务器将实现三位数增长，液冷散热方案占比持续提升，算力行业景气度持续上行。

- **AI芯片：2024年9月GB200有望出货，供给端持续提速。** 1) **海外：**①**英伟达：**广达表示2024年9月，NVIDIA GB200及B100等产品有望出货；NVIDIA H100交付周期缩短。②**AMD：**截至2024Q1，MI300系列加速器的累计收入已经超过10亿美元，预计2024年可超40亿美元。③**Intel：**Gaudi 3训推能力优于H100，预期2024Q3批量交付。④**自研芯片：**谷歌公开自研Arm CPU，Meta发布新代AI芯片MTIA。2) **国产：**2023Q4海思处理器出货量达680万片。
- **服务器：预期2024年全球AI服务器三位数增长，中国市场国产服务器占比提升。** 1) **台股：**2024年4月，广达、纬创、英业达、纬颖等营收呈增长趋势，主要系服务器营收增长，预计2024年，通用服务器全年同比持平或微增，AI服务器将三位数增长；2) **内地：**“昇腾+鲲鹏”生态逐步完善，华鲲振宇/神州鲲泰等整机伙伴新品频发；华勤技术表示NV L20服务器已出货，H20服务器预期Q2批量出货。2024年，中国移动启动7994台AI服务器集采招标，国产AI服务器需求提升。
- **零部件：液冷散热渗透提升，2024年1-4月信骅营收同比+59.28%。** 1) **散热：**据奇鋌，2024Q1营运优于预期，受益进入英伟达3D VC供应链。展望2024年，3DVC产品有望显著放量，液冷散热将于2024H2小批出货、2025年快速成长。2) **主板：**AI服务器推升主板出货量提升，叠加第一季主板传统小旺季效益，华擎/技嘉等厂商2024Q1营收同比增长。展望2024年，预期AI服务器为主要拉货动能。3) **BMC：**信骅1-4月营收同比+59.28%，预期2024年AI推理服务器需求提升，通用服务器也会回升。
- **产业链：1) CoWoS：**据财联社，英伟达GB200预估2025年出货量或破百万颗，台积电提升2024全年CoWoS产能，预估至年底每月产能将逼近4万，同比提升逾150%。2) **资本开支：**FY2024Q3（截至2024.3.31），微软资本开支110亿美元，同比+66%，表示下一季度资本开支将环比大幅提升；2024Q1，谷歌资本开支120亿美元、同比+90%左右，预计全年的季度资本开支将等于或高于一季度水平。

- **投资建议：**大模型训推带动 AI算力需求增长，算力产业链中的AI芯片、服务器整机及零组件、光模块、IDC等环节有望持续受益。维持对计算机行业“推荐”评级。
- **相关公司**
 - 1) **服务器整机：**浪潮信息、工业富联、中科曙光、华勤技术、紫光股份、中国长城、高新发展、神州数码、烽火通信、拓维信息、软通动力、纬创、广达、英业达、纬颖、超微电脑。
 - 2) **服务器组件：**①**AI芯片：**海光信息、寒武纪、龙芯中科、景嘉微、英伟达、AMD、Intel；②**散热：**飞荣达、曙光数创、英维克、同方股份、申菱环境、高澜股份、奇鋆科技、双鸿、健策、VERTIV；③**主板：**沪电股份、深南电路、胜宏科技、技嘉、华擎。
 - 3) **光模块：**天孚通信、中际旭创、新易盛、光迅科技、华工科技。
 - 4) **数据中心：**奥飞数据、光环新网、宝信软件、数据港、电科数字。
- **风险提示：**宏观经济影响下游需求、大模型产业发展不及预期、市场竞争加剧、中美博弈加剧、相关公司业绩不及预期等，各公司并不具备完全可比性，对标的相关资料和数据仅供参考。

第一章 AI芯片

GB200预计Q3出货，供给端持续提速

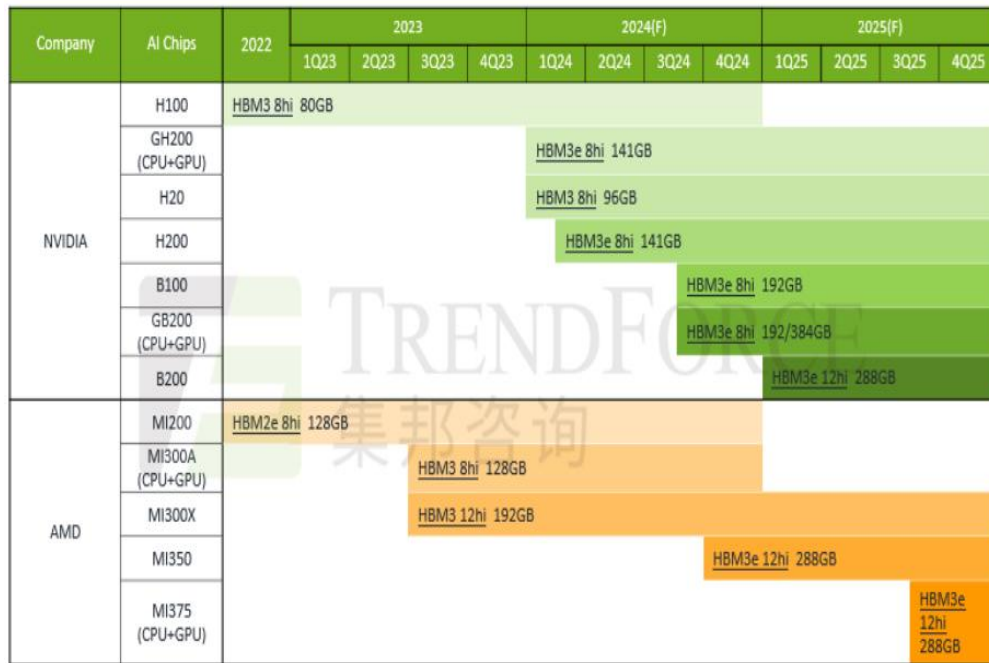
1.1.1 英伟达：GB200 GPU有望2024年9月出货，供给提速

- 2025年，GB200出货预期超百万颗、占英伟达高端GPU近4-5成。4月16日，TrendForce集邦咨询指出，GB200的前一代为GH200，皆为CPU+GPU方案，主要搭载NVIDIA Grace CPU及H200 GPU，但以GH200出货量估仅占整体NVIDIA高端GPU约5%。目前供应链对NVIDIA GB200寄予厚望，预估2025年出货量有机会突破百万颗，占NVIDIA高端GPU近4~5成。
- 2024年9月，GB200及B100等产品有望出货，供给提速。NVIDIA计划2024H2推出GB200及B100等产品，据中国台湾代工厂广达，搭载英伟达GB200的服务器正开发中，GB200伺服器需求量很大，预计9月份量产出货，目前看来下半年晶片缺料有望舒缓。

表：英伟达多款AI芯片性能情况

型号	GB200	H200 SXM	HGX H20	A100 80GB PCIe	..
FP4	40 petaflops	-	-	-	-
FP6/FP8	20 petaflops	-	-	-	-
FP16	10 petaflops	1979 TFLOPS	148 TFLOPS	312 TFLOPS	1513 TFLOPS
FP32/TF32	5 petaflops	67 TFLOPS	44 TFLOPS	19.5 TFLOPS	51 TFLOPS
FP64	90 teraflops	34 TFLOPS	1 TFLOPS	9.7 TFLOPS	26 TFLOPS
显存容量	384GB	141GB HBM3e	96GB HBM3	80GB HBM2e	80GB HBM3
内存带宽	16TB/s	4.8 TB/s	4.0TB/s	1935GB/s	2.0TB/s
互联技术	3.6TB/s	NVLink 900GB/s	NVLink 900GB/s	NVLink 600GB/s	NVLink 600GB/s
最大功耗	up to 2700W	700W	400W	300 W	300~350W

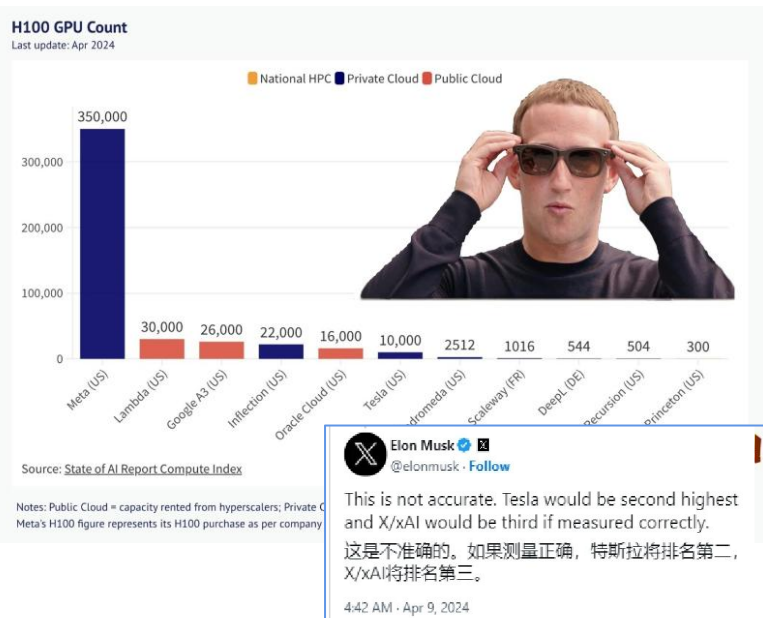
图：NVIDIA及AMD AI芯片发展进程及HBM规格比较



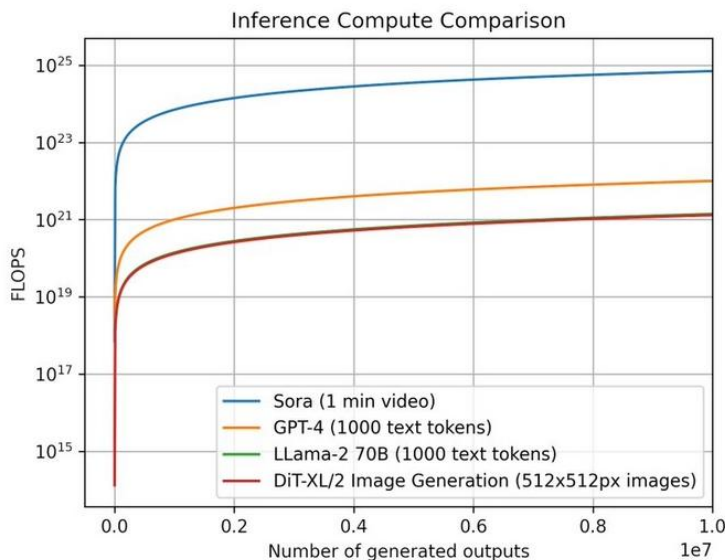
1.1.2 英伟达：微软与特斯拉等厂商算力需求持续旺盛

- **特斯拉与xAI的AI芯片投资力度较大。**据4月9日 IT之家，X 平台“The Technology Brother”表示Meta 公司已经囤积了约35万块 H100 GPU，全球第一。随后，马斯克回应“如果计算正确，特斯拉应该是第二名，xAI 将是第三名”。这或许意味着特斯拉目前可能拥有 3 万到 35 万颗的 H100 GPU，而 xAI 则拥有大约 2.6 万到 3 万颗。
- **微软和OpenAI计划投资1000亿美元建造“星际之门”AI超级计算机。**据The Information，微软、OpenAI正在进行一项分为五个阶段的超级计算机建设项目，拟动用的投资将超过1150亿美元。在微软和OpenAI的算力开发第四、第五阶段，都会有一台标志性的超算，第四阶段超算预期将在2026年投入使用，而第五阶段的超算星际之门（Stargate），微软高层希望最快能够在2028年推出这款超算。
- **OpenAI部署Sora在峰值期需要72万块H100 GPU。**据机构Factorial Funds报告，OpenAI部署Sora在峰值期需要72万片英伟达 H100 AI加速卡，如果按照每片H100 3万美金成本计算，72万片需要216亿美元。Factorial Funds 预估训练 Sora 1 个月，需要使用 4200 到 10500 片英伟达 H100 AI 加速卡，如果使用单片 H100 加速卡，生成 1 分钟视频需要渲染 12 分钟。

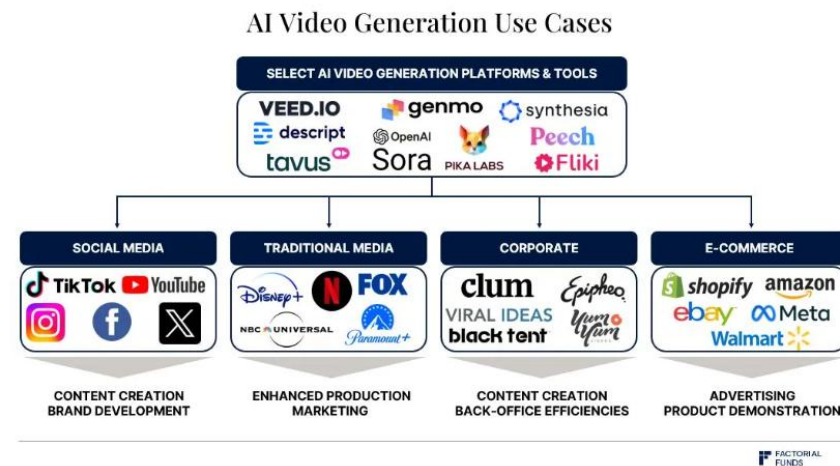
图：海外厂商AI芯片量预期情况



图：不同模型的推理算力消耗情况



图：AI视频生成用例



1.1.3 供给端：H100交货时间缩短至2-3个月

- **英伟达H100交货时间缩短至2-3个月、供应瓶颈逐步缓解。**4月11日，据 Digitimes ，台积电台湾地区总经理 Terence Liao 表示，过去几个月英伟达 H100 的交货等待时间大幅缩短，从最初的 3-4 个月降至目前的 2-3 个月（8-12 周）。服务器代工厂商也透露，相比 2023 年几乎无法买到英伟达 H100 的情况，目前供应瓶颈正在逐步缓解。
- **厂商库存H100转售与云厂商开放H100云端租赁服务，或为交货时间缩短的原因。**据IT之家，2-3 个月的交货等待时间是英伟达 H100 GPU 史上最短的。自 2024 年初以来，H100 GPU 交货等待时间一直在显著缩短，导致这种情况变化的部分原因可能是，一些公司持有过剩的 H100 GPU 库存，并且为了降低闲置库存的高额维护成本而进行转售。此外，亚马逊云科技 (AWS) 方便了用户通过云端租赁英伟达 H100 GPU，这也帮助缓解了部分对 H100 的需求压力。

1.2 AMD：2024Q1数据中心增速较快，Q2业绩指引符合预期

- **2024Q1，AMD收入符合预期。**2024Q1，公司收入54.73亿美元，同比增长2%，环比下跌11%；毛利润25.60亿美元，同比增长9%，环比下跌12%；毛利率47%，同比增长3个百分点，环比持平。净利润1.23亿美元，同比增长188%，环比下跌82%；每股收益0.07美元，同比增长178%，环比下跌83%。
- **分业务来看，数据中心业务同比增长，游戏业务与嵌入式业务承压。**2024Q1，AMD数据中心业务实现收入 23.37 亿美元，同比上升 80%，主要是因为公司 AMD Instinct GPU 销量和第四代 EPYC（霄龙）CPU 销量的提升。其中，MI300系列加速器的累计收入已经超过10亿美元，预计最终可超40亿美元。游戏业务因为显卡销量低迷、半定制产品下滑；嵌入式业务因为客户库存影响，收入下滑46%至8.46亿美元。
- **业绩指引：**2024 年第二季度预期收入 54-60 亿美元和 non-GAAP 毛利率 53% 左右。

图：2024Q1 AMD业绩情况

\$ in millions, except per share data and %	Q1'24	Q1'23	Y/Y	Q4'23	Q/Q
Revenue	\$5,473	\$5,353	Up 2%	\$6,168	Down 11%
Gross Profit	\$2,861	\$2,675	Up 7%	\$3,133	Down 9%
Gross Margin	52%	50%	Up 2 pts	51%	Up 1 ppt
Operating Expenses	\$1,741	\$1,587	Up 10%	\$1,727	Flat
Operating Expense/Revenue %	32%	30%	Up 2 pts	28%	Up 4 pts
Operating Income	\$1,133	\$1,098	Up 3%	\$1,412	Down 20%
Operating Margin	21%	21%	Flat	23%	Down 2 pts
Net Income	\$1,013	\$970	Up 4%	\$1,249	Down 19%
Earnings Per Share	\$0.62	\$0.60	Up 3%	\$0.77	Down 19%

1. See Appendices for GAAP to Non-GAAP reconciliation

Q1 2024 FINANCIAL RESULTS - APRIL 30, 2024

AMD together we advance

图：2024Q1 AMD的数据中心业务情况

(\$ in millions)	Q1'24	Q1'23	Y/Y	Q4'23	Q/Q
Data Center					
Net Revenue	\$2,337	\$1,295	Up 80%	\$2,282	Up 2%
Operating Income	\$541	\$148	Up 266%	\$666	Down 19%
Client					
Net Revenue	\$1,368	\$739	Up 85%	\$1,461	Down 6%
Operating Income (Loss)	\$86	\$(172)	Up 150%	\$55	Up 56%
Gaming					
Net Revenue	\$922	\$1,757	Down 48%	\$1,368	Down 33%
Operating Income	\$151	\$314	Down 52%	\$224	Down 33%
Embedded					
Net Revenue	\$846	\$1,562	Down 46%	\$1,057	Down 20%
Operating Income	\$342	\$798	Down 57%	\$461	Down 26%

1.2 AMD：面向AIPC，推出锐龙PRO 8040/8000系列处理器

- 4月16日，AMD公布全新AI PC处理器锐龙PRO 8040/8000系列。作为业界率先在X86处理器内整合人工智能引擎（NPU）的公司，AMD此次推出锐龙PRO 8040系列移动处理器性能出色，可以降低功耗，保障续航时间；面向商用台式机产品，AMD推出由AI驱动的产品、锐龙 PRO 8000系列处理器。
- 锐龙PRO 8040/8000系列集成CPU+GPU+NPU，性能出色。基于目前最先进的4nm制造工艺、Zen 4 CPU架构、RDNA3 GPU架构、XDNA NPU架构的强大组合，并加入AMD PRO技术组合，同时还可支持目前最快的Wi-Fi 7和蓝牙5.4无线连接标准。

图：锐龙PRO 8040/8000系列处理器



图：锐龙PRO 8040/8000系列处理器性能介绍



1.3 Intel: 2024Q1业绩符合预期, Q2营收指引125-135亿美元

- **2024Q1, Intel业绩基本符合预期。**2024Q1, 公司总营收127亿美元, 同比增长8.6%, 净利润亏损3.8亿美元, 同比去年亏损大幅收窄, 两项数据基本符合市场预期。
- **分业务来看, 2024Q1公司产品部门总营收持续增长。**公司业务线包括晶圆代工、产品部门和其他业务。2024Q1, Intel产品部门总营收119亿美元, 同比+17%。其中, 客户端计算事业部营收75亿美元, 同比+31%; 数据中心和人工智能事业部营收30亿美元, 同比+5%; 网络及边缘营收14亿美元, 同比-8%。
- **业绩展望: 2024Q2, 预期公司营收在125-135亿美元, 同比接近持平; non-GAAP下毛利率43.5%, 同比+3.7pct; non-GAAP EPS 0.10美元, 同比下降0.03美元。**

图: 2024Q1, Intel 营收情况

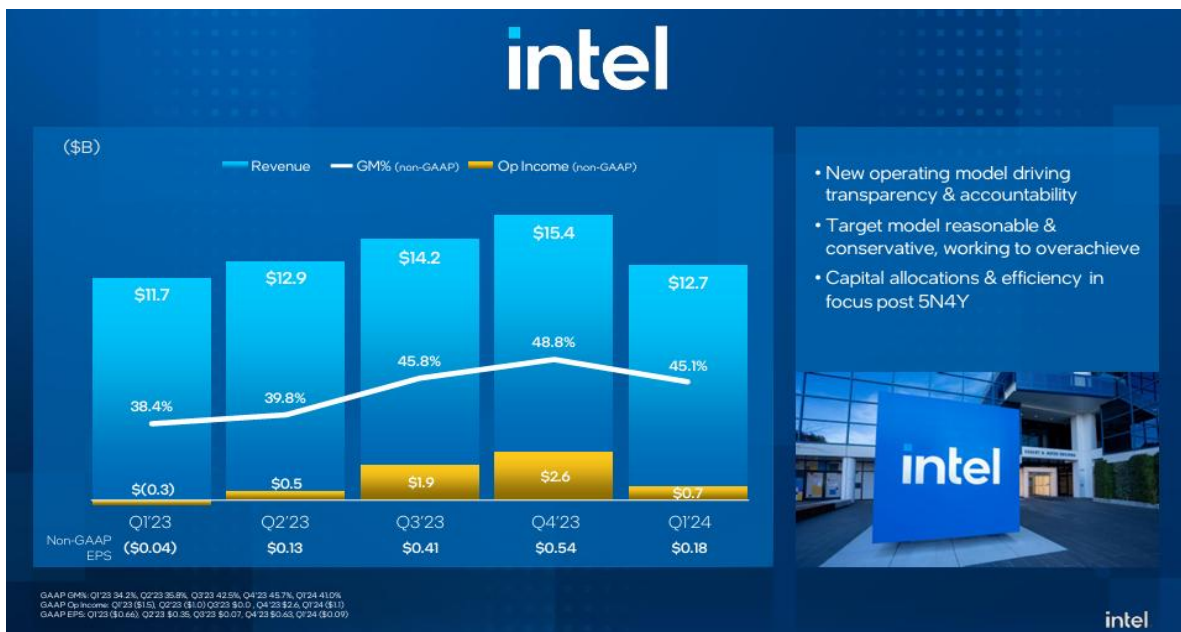
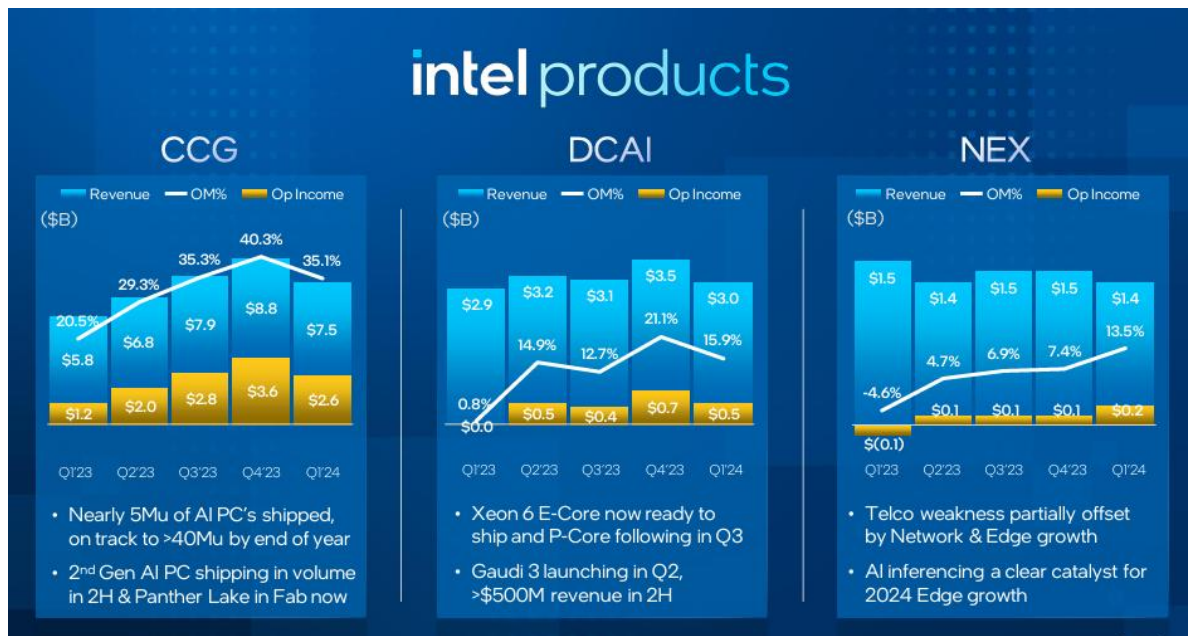


图: 2024Q1, Intel 产品部门营收情况



1.3 Intel: Gaudi 3正式推出, BF16工作负载性能为前代四倍

- 英特尔正式推出了全新的Gaudi 3。4月9日, 英特尔召开“Intel Vision 2024”大会, 发布Gaudi 3, 采用台积电5nm工艺, 为两个计算集群的Chiplet设计, 而非Gaudi 2采用的单个集群的方案。Gaudi 3提供64个第五代张量处理核心和8个矩阵计算引擎, 并搭载128GB速率达3.7TB/s的HBM内存和96MB速率达12.8TB/s的SRAM。连接方面支持24条的200GbE符合RoCE标准的以太网总线以及最高16条PCIe 5.0总线。
- Gaudi3的BF16性能为Gaudi2的4倍。相比Gaudi 2, Gaudi 3在BF16工作负载方面的性能将是Gaudi 2的四倍, FP8性能也将是Gaudi 2两倍, 网络性能是Gaudi 2的两倍 (Gaudi 2是24个内置的100 GbE RoCE Nic), HBM容量是Gaudi 2的1.5倍。

图: Intel Gaudi 3 产品设计

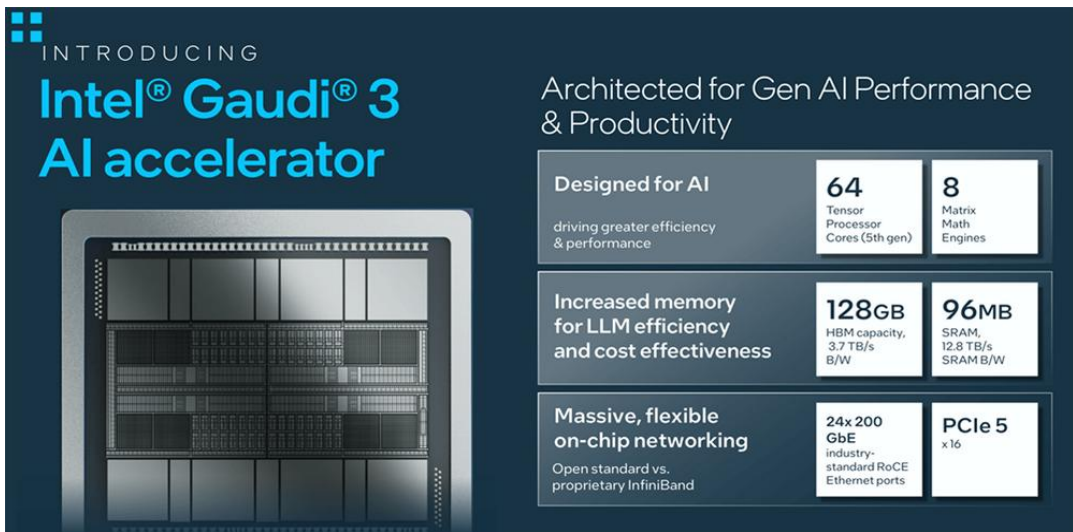
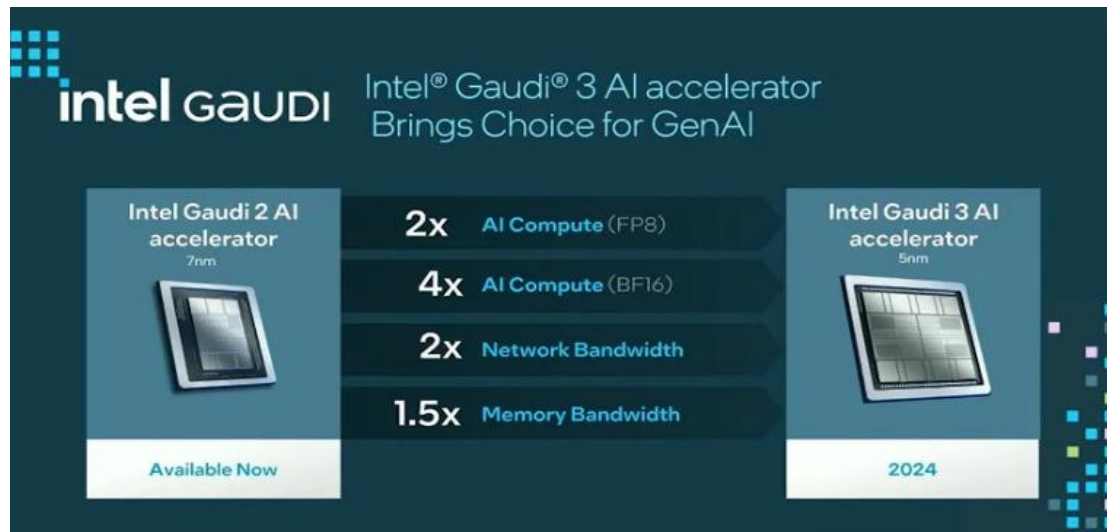


图: Intel Gaudi 3 架构



1.3 Intel: Gaudi 3 训推能力优于H100, 预期2024Q3批量交付

- 相比H100, 英特尔Gaudi 3训练时间最高快1.7倍、推理时间最高快4倍。英特尔Gaudi 3与英伟达H100在相同节点数量下, 相关大模型训练时间对比上最高快1.7倍, 其中, LLAMA2 70 亿参数对比有1.5倍于H100的优势, LLAMA2 130 亿参数最高有1.7倍优势, GPT3 1750亿参数有1.4倍优势。在大模型推理速度表现上, Gaudi 3相比H100平均快了1.5倍, 最高快4倍。
- Gaudi 3具备风冷与液冷版本, 将在2024Q3批量交付。英特尔在 Vision 2024 上也公布了 Gaudi 3 生产节点, 2024 年一季度将率先推出风冷版样品, 二季度推出液冷版样品, 并在今年第三、第四季度分别批量交付风冷版和液冷版。Gaudi 3 将由戴尔、惠与、联想和超微四家 OEM 提供。

图: Intel Gaudi 3产品训练能力

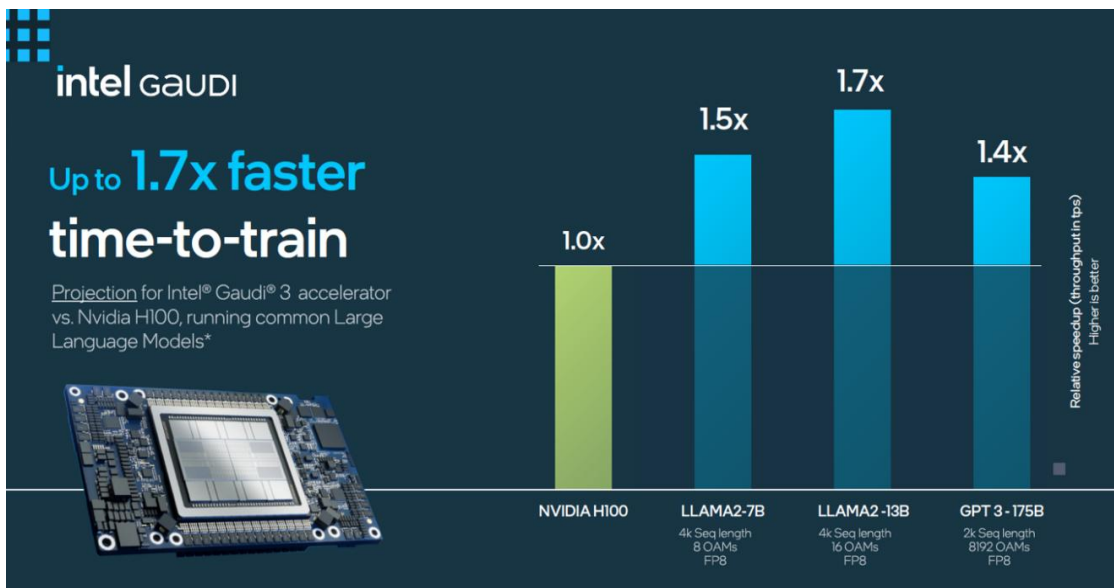
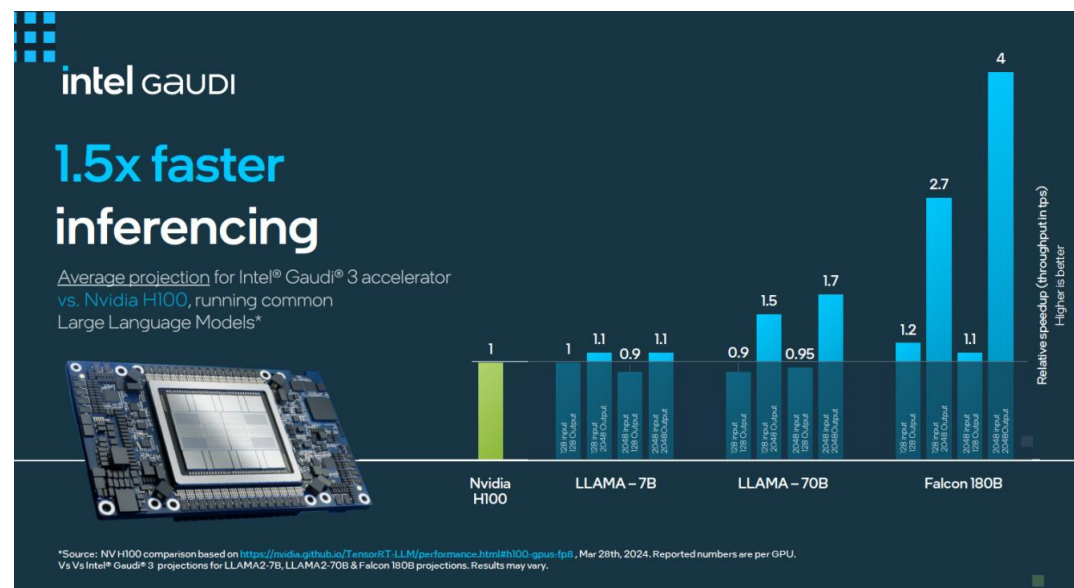


图: Intel Gaudi 3 产品推理能力



1.3 Intel: 推出2款面向中国市场的Gaudi 3加速卡

- 英特尔准备推出两款专供中国市场的 Gaudi 3 AI 加速器芯片。具体包括两种硬件形态加速卡：
 - ✓ **HL-328**: OAM兼容夹层卡 (Mezzanine Card) , 预计将于今年6月24日推出;
 - ✓ **HL-388**: PCIe加速卡, 预计将于今年9月24日推出。
 - ✓ **性能**: 均为搭载128GB HBM2e VRAM、3.7TB/s 带宽、96MB 缓存、PCIe 5.0 x16 接口以及解码标准, TDP均为450瓦。
 - ✓ 基于内核数量、工作频率、TDP等参数估算, 相比Gaudi 3国际版, “中国特供版” HL-328芯片性能或降低约92%左右。

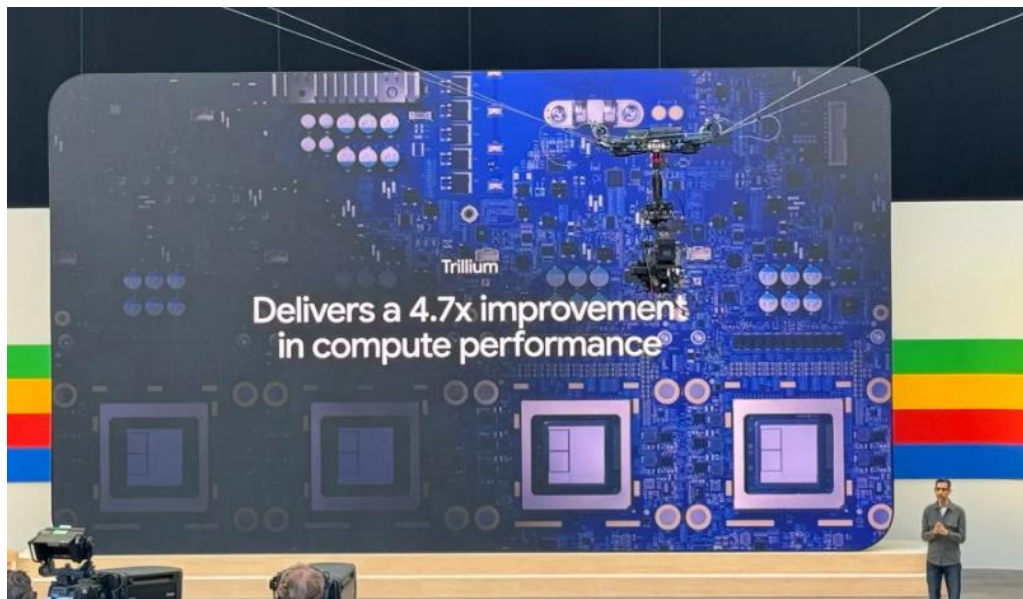
图: Intel Gaudi 3 产品推理能力

Family	Intel® Gaudi® 3-OAM			Intel® Gaudi® 3-PCIe	
	HL-325L	HL-328	HL-335	HL-338	HL-388
Part Number	HL-325L	HL-328	HL-335	HL-338	HL-388
Export Market	Non-PRC	PRC (*)	Non-PRC	Non-PRC	PRC (*)
Availability	Mar '24	Jun '24	Oct '24	Sep '24	Sep '24
Form Factor	OAM	OAM	OAM	Dual-slot PCIe, Full Height, 10.5" Length	
Cooling	Air	Air	Liquid (1P or 2P)	Air	Air
TDP [W]	900	450	900	600	450
Scale-up support	via x8 Baseboard (HLB-325)			x4 Top-Board (HLTB-304)	
Scale-out support	via Baseboard (6x OSFP800) or via Host-Nic			via Host-Nic or via dual QSFP-112 on PCIe card	
HBM Capacity	128GB				
Peak HBM Bandwidth	3.7TB/sec				
HBM Interface and Type	1024-bit x 8 stacks HBM2e				
Last-Level Cache Capacity	96MB				
Host Interface	PCIe Gen5 x16				
Media Decoders	HEVC/H.265, AVC/H.264, VP9, JPEG				

1.4 自研芯片：谷歌发布第6代TPU，Meta发布新代AI芯片

- **谷歌发布第6代Trillium TPU：性能提升4.7倍、内存带宽翻番、节能 67%。**5月15日，谷歌公在今天召开的 I / O 2024 开发者大会上，宣布了第六代数据中心 AI 芯片 Tensor 处理器单元（TPU）--Trillium，并表示将于今年晚些时候推出交付。谷歌通过扩大芯片的矩阵乘法单元（MXU）和提高整体时钟速度，达到每块芯片的计算性能将比第五代提升 4.7 倍，还将内存带宽提高了一倍。
- **Meta公布全新自研MTIA芯片、性能为上一代三倍。**4月10日，Meta公布“下一代MTIA”，专注于推理任务。新款 MTIA 芯片采用台积电的5nm工艺，拥有更多的处理核心，将拥有256MB的片上内存，频率为1.3GHz。Meta的早期测试结果显示，通过测试两种芯片的“四个关键模型”的性能，新芯片的性能是一代版本的三倍。在硬件方面，Meta开发大型机架式系统支持下一代芯片，最多可容纳72个加速器，可将芯片时钟频率从初代800 MHz提高至1.35GHz，并以90瓦功率运行，而初代设计的功耗为25瓦。

图：谷歌发布第6代Trillium TPU



图：Meta为MTIA开发的大型机架式系统



1.5 国产芯片：经营情况向好，推进产业生态适配

- 国产芯片经营情况向好，推进互联网/金融等领域生态适配。据Canalys《全球智能手机市场总览》，2023Q4海思处理器出货量达680万片，同比+5121%；收入达到70亿美元。海光DCU实现LLaMa、GPT、Bloom、紫东太初等大模型全面应用，寒武纪表示云端产品主要应用于互联网、金融等领域；燧原科技表示与无问芯穹合作在湖北宜昌智算中心打造千卡集群案例等。

图：国产芯片全面突围，多家厂商完成适配

公司	具体内容
华为海思	<ul style="list-style-type: none">• 据Canalys《全球智能手机市场总览》，2023Q4海思处理器出货量达680万片，同比+5121%；收入达到70亿美元，同比暴涨24471%。
海光信息	<ul style="list-style-type: none">• 2024年4月25日，公司发布2024年一季报，实现营收15.92亿元，同比+37.09%；归母净利润2.89亿元，同比+20.53%。• 据2023年报，海光DCU能够支持全精度模型训练，实现LLaMa、GPT、Bloom、ChatGLM、悟道、紫东太初等为代表的大模型的全面应用，与国内包括文心一言等大模型全面适配，达到国内领先水平。公司表示，NVIDIA CUDA EULA中相关调整不会对DCU相关产品产生影响，DCU硬件及软件栈产品依然会延续自主研发路线进行演进。
寒武纪	<ul style="list-style-type: none">• 公司表示，云端产品主要应用于互联网、金融等领域，其中互联网客户在云端智能计算市场占据大部分市场空间。在互联网行业的客户拓展已经取得了一定进展，例如与阿里、百度等头部互联网企业的多个业务部门进行了密切交流并已实现产品导入，在视觉、语音等场景的适配性能表现超出客户预期，部分场景已经形成一定规模收入。
摩尔线程	<ul style="list-style-type: none">• 据4月19日摩尔线程公众号，摩尔线程与国内技术领先的企业级云服务商与数字化解决方案提供商青云科技达成战略合作。双方将以摩尔线程全功能GPU、青云科技AI智算产品及服务的深度融合为创新驱动，聚焦国产化人工智能创新应用和数字化建设，共同推动人工智能与政企市场数字化应用深度融合，共建AGI时代的国产智能算力繁荣生态，助力数字经济发展大局。

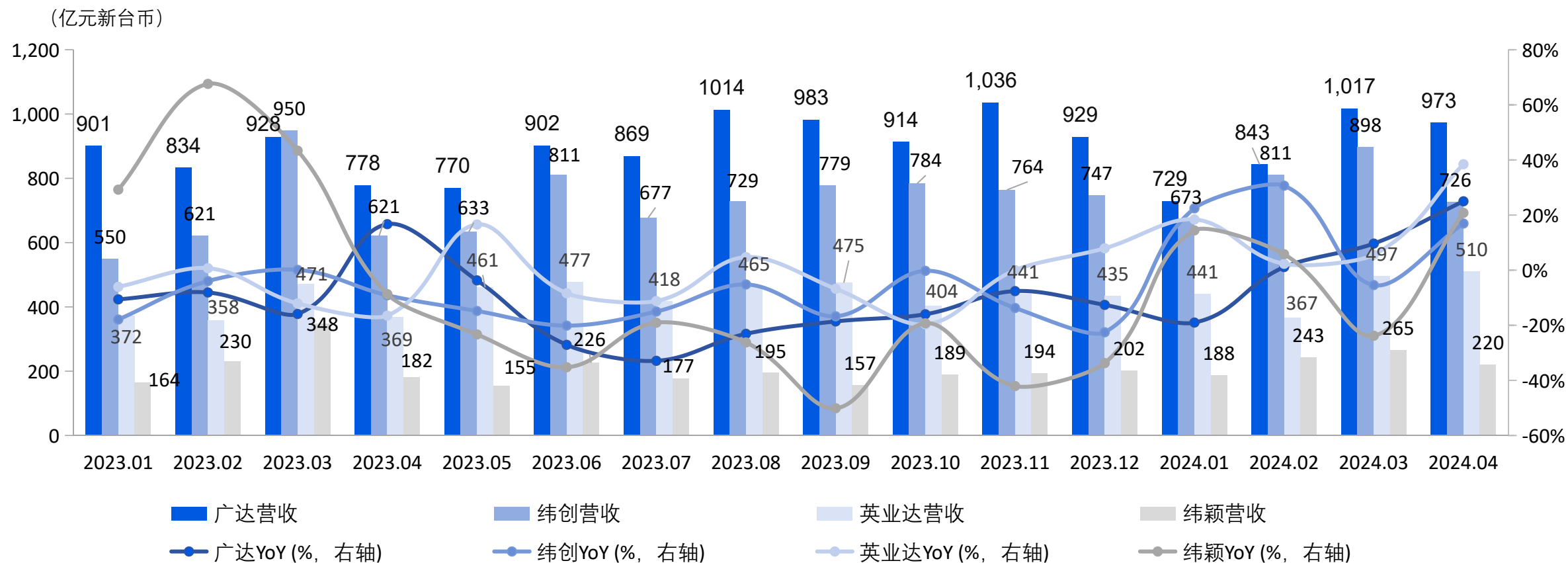
第二章 AI服务器

AI服务器逐季增长，H20预计Q2批量出货

2.1.1 台股服务器：受益于AI服务器增长，通用+PC维持稳定

- 2024年4月，广达、纬创、英业达、纬颖、鸿海营收分别为973、726、220、510、5109亿元新台币，同比+25%、+17%、+21%、+39%、+19%，均呈现同比增长趋势，AI服务器增长动能强劲。

图：中国台湾服务器公司月度收入及同比情况



2.1.1 台股服务器：2024Q2，AI服务器有望增长，至Q3逐步放量

- **展望2024Q2：**1) **PC：**2024Q2，纬创表示PC有望小幅增长，广达表示4月出货量较3月下滑，英业达表示第二季笔电出货或将相对第一季下移；2) **通用服务器：**英业达表示服务器或呈季度环比单位数增长，我们认为，主要系2023年需求触底、库存消耗后需求回暖，并且2024Q2处于英伟达GPU产品交替之际，存在资本开支部分回流通用服务器的情况；3) **AI服务器：**2024Q2，各厂商预期AI服务器营收有望增长，营收从Q2至Q3逐步放量。

表：2024Q2中国台湾ODM厂商对服务器与PC的展望情况

产品	2024Q2 预期
AI服务器	<ul style="list-style-type: none">• 随上游缺料情况缓解，预期Q2的AI服务器拉货动能转强，但英伟达等AI芯片产品代际更换或会带来一定影响。• 1) 广达：随着零组件缺料情况改善，预期二季度起AI服务器出货将温和成长，受惠超大规模数据中心客户积极建设。• 2) 纬创：2024Q2，公司预期AI 伺服器有望两位数成长。
通用服务器	<ul style="list-style-type: none">• 英业达表示，服务器出货在第二季将较第一季有个位数成长的表现，未来也将逐季成长。
PC	<ul style="list-style-type: none">• 纬创表示PC有望小幅增长，广达表示4月出货量将较3月下滑，英业达表示第二季笔电出货或将相对第一季下移。• 1) 广达：由于3月逢季底拉货时期，垫高PC出货基期，广达表示，4月出货量将较3月下滑。• 2) 纬创：桌机同样有望两位数成长，NB 与显示器第二季呈现小幅成长。• 3) 英业达：笔电出货量因提前至第一季出货，再加上商务专案往后递延至第三季，因此预期第二季笔电出货将较第一季下修

2.1.1 台股服务器：展望2024年，AI服务器预期三位数增长

- 展望2024年全年：1) PC：各厂商预期PC维持平稳或微增；2) 通用服务器：预期全年同比持平或微增；3) AI服务器：多厂商预测三位数增长，占营收比例提升。

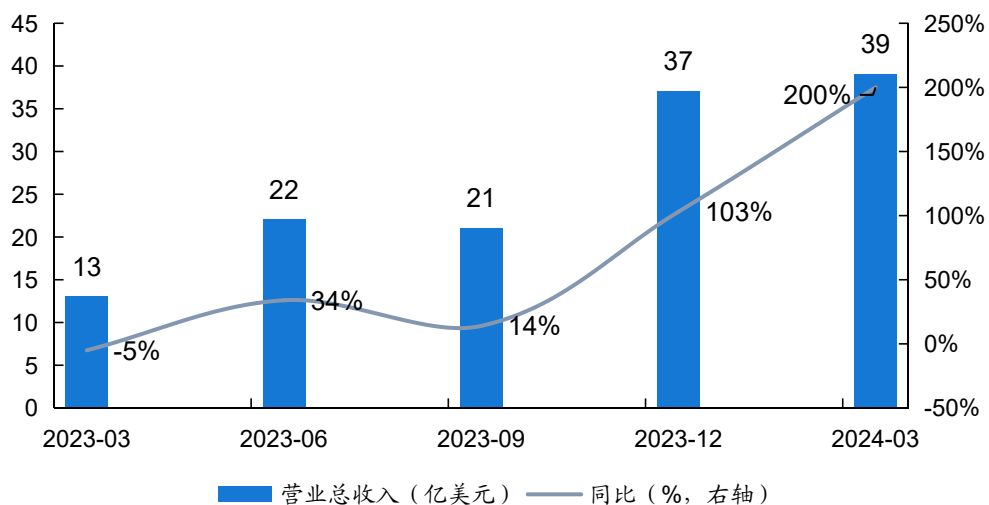
表：2024年，中国台湾ODM厂商对服务器与PC的展望情况

产品	2024年 全年指引情况
AI服务器	<ul style="list-style-type: none"> ● 纬创、英业达等厂商预期全年AI服务器呈现三位数增长。鸿海、广达、纬颖等厂商表示AI服务器营收占比持续提升。 ● 1) 鸿海：1) 人工智能服务器业务在整体服务器营收占比将达到四成以上（2023年约为三成）；2) 2024年GPU模组或超三位数年增幅度；3) 目标AI服务器机架级产品市占率四成。 ● 2) 广达：预计下半年AI服务器有望迎来出货爆发期，已经拿到了云端服务供应商（CSP）订单，以及非CSP厂的新客户订单，预计2024年AI服务器占服务器整体收入比超过50%。部分AI服务器客户项目已看到2025年出货量持续攀升。 ● 3) 纬创：预期维持AI服务器将有三位数成长的表现，就全年来看，下半年AI产品业绩会优于上半年。客户对AI产品需求量持续增加，法人预期，GPU伺服器相关业务上半年不淡、下半年则再加温，由于产品组合持续优化，有助于毛利率向上。 ● 4) 纬颖：随着新产品研发完毕量产出货，AI服务器占比将由20%起跳、逐季提升。 ● 5) 英业达：2024年，全年则可望有双位数成长，AI伺服器出货将有倍增表现。2023年AI伺服器营收占比约5%至6%，2024年将大增至超过一成
通用服务器	<ul style="list-style-type: none"> ● 多厂商预期通用服务器收入全年呈现持平或微幅增长。 ● 1) 广达：通用服务器已看到“隧道尽头的光”，预期需求很快就会回来。2024年服务器表现将是所有产品线最佳，营收会有强劲的双位数成长。 ● 2) 纬创：企业服务器/台式机出货量预期为低个位数增长，略好于行业增长、主要系市场份额或增长。 ● 3) 纬颖：由于AI伺服器比重拉升，预期2024年通用伺服器比重下降，不过绝对金额预估较2023年持平至微幅成长。 ● 4) 英业达：伺服器出货量将持平至小幅成长。
PC	<ul style="list-style-type: none"> ● 多厂商预期笔电产品与2023年持平，或个位数增长。 ● 1) 广达：2024Q1的AI PC出货量相对有限，相关应用仍在起步阶段，预期商业机型将先行，消费性机型跟进，受益于PC、Windows10换机潮等动能驱动，对未来几年产品升级趋势乐观看待。 ● 2) 鸿海：2024年智能消费者(iPhone)和计算(台式机、NB、iPad)的收入将与上年同期持平，好于此前预期的同比下降。 ● 3) 广达：全年NB出货量持平指导。NB的收入贡献在2023年为40%-45%，在2024年下降到低30%。 ● 4) 纬创：随着2023年产业触底，预期2024年会小幅成长、成长幅度落在0-5%。 ● 5) 英业达：预期与PC市场同步、有个位数成长。

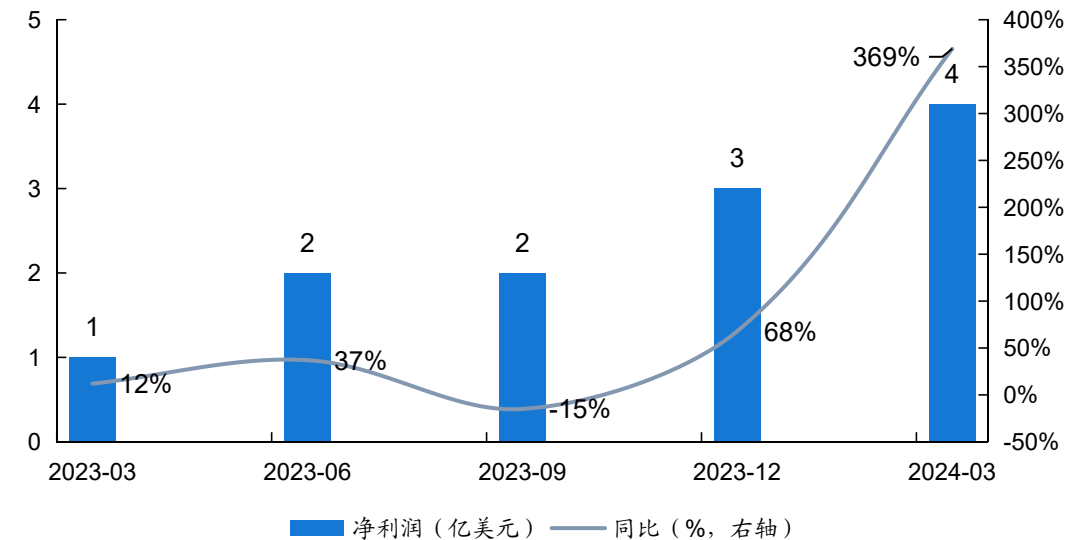
2.1.2 超微电脑：受益AI浪潮，上调FY2024营收指引147-151亿美元

- **FY2024Q3，公司调整后EPS高于预期。**4月30日，超微电脑公布FY2024Q3财报，总营收为38.5亿美元，同比增长200.8%；毛利率为15.5%，高于2024财年第二财季的15.4%，低于去年同期的17.6%，符合预期；调整后EPS为6.65美元，同比增长308%，高于预期5.78美元。
- **业绩指引方面，超微电脑预计第四财季营收51亿至55亿美元，高于预期47.3亿美元；调整后每股收益区间为7.62至8.42美元。**对于2024财年全年（截止今年6月30日），公司预计全财年营收147亿至151亿美元，高于公司原本预计143亿至147亿美元。
- **公司的快速发展，主要得益于AI机架规模即插即用（PnP）解决方案的强劲需求，以及团队开发创新的DLC设计的能力，使公司在AI基础设施市场的领导地位得到了扩展。**公司目前客户基础正在强劲增长。展望未来，公司将继续致力于开发先进的解决方案，包括完全量产的DLC，公司也预计将持续提升市场份额。

图：分季度公司总营收情况



图：分季度公司净利润情况



- **英伟达L20/H20方面：**
- **浪潮信息：**据5月15日公司公告，公司表示对于H20，2024年Q1尚未形成规模销售，预计Q2、Q3可能有一定量产交付，目前对业务整体的影响暂无法明确预估。2023年，整体上看，公司通用服务器占比下降，人工智能服务器占比提升。在各方面正常情况下，公司预计2024年全年整体的产品结构将与2023年类似。
- **华勤技术：**4月2日，华勤技术表示，AI服务器业务发展迅猛、增长迅速，2024Q1已经量产出货最新的NV L20 GPU AI服务器产品，NV H20 GPU平台产品也在完成互联网客户的适配和集群的灰度测试，正在导入中，预计Q2可形成批量出货。预计2024年高性能计算数据产品业务依然保持高速增长。

2.2 内地服务器：运营商AI服务器大单落地，国产算力需求提升

- 2024年，中国移动启动7994台AI服务器集采招标。2024年4月18日，中国移动采购与招标网发布公告：启动2024年至2025年新型智算中心采购招标，将采购7994台AI服务器及配套产品、60台白盒交换机。
- 2024年5月16日，中国移动发布《中国移动2024年至2025年新型智算中心采购中标候选人公示》，其中昆仑技术、华鲲振宇、宝德计算机、长江计算、神州鲲泰等7家中标标包1（7994台AI服务器）。

表：中国移动2024年至2025年新型智算中心采购项目标包划分

标包	产品名称	需求数量	中标人数与份额
标包1	人工智能服务器	7994台	<ul style="list-style-type: none"> ①昆仑技术：190.97亿元，份额21.05%； ②华坤振宇：190.97亿元，份额17.54%； ③宝德计算机：190.97亿元，份额15.79%； ④百信：191.36亿元，份额14.04%； ⑤长江计算：191.17亿元，份额12.28%； ⑥神州鲲泰：191.17亿元，份额10.53%； ⑦湘江鲲鹏：190.58亿元，份额9.77%
标包2	交换机	60台	<ul style="list-style-type: none"> 中标人1家，中标份额为100%。

表：中国移动2023年至2024年新型智算中心(试验网)采购项目标包

标段	产品名称	单位	需求数量	中标候选人（投标报价均为不含税金额）
标包1	高性能无损交换机	台	324	-
标包1	机架式PC服务器	套	2	-
标包1	通用AI训练服务器(扣卡液冷)	台	256	-
标包1	管理软件	许可	2050	-
标包2	通用AI训练服务器(扣卡液冷)	台	374	-
标包3	通用AI训练服务器(PCIe风冷)	台	86	-
标包4	特定场景AI训练服务器(PCIe风冷)	台	52	新华三技术有限公司，投标报价1472.38万元，中标份额70%；烽火通信科技股份有限公司，投标报价1438.97万元，中标份额30%
标包5	通用AI推理服务器(PCIe风冷)	台	16	中兴通讯股份有限公司，投标报价362.99万元，中标份额100%
标包6	特定场景AI推理服务器(PCIe风冷)	台	64	新华三技术有限公司，投标报价1066.11万元，中标份额70%；河南昆仑技术有限公司，投标报价1014.96万元，中标份额30%
标包11	特定场景AI训练服务器(扣卡液冷)	台	356	河南昆仑技术有限公司，投标报价49095.61万元，中标256台；四川华鲲振宇智能科技有限责任公司，投标报价49049.16万元，中标100台
标包12	特定场景AI训练服务器(扣卡风冷)	台	106	河南昆仑技术有限公司，投标报价247372.15万元，中标份额40.96%；四川华鲲振宇智能科技有限责任公司，投标报价247372.14万元，中标份额30.08%；烽火通信科技股份有限公司，投标报价247372.92万元，中标份额20.48%；神州数码(中国)有限公司，投标报价247372.28万元，中标份额8.48%
	特定场景AI训练服务器(扣卡液冷)	台	1144	

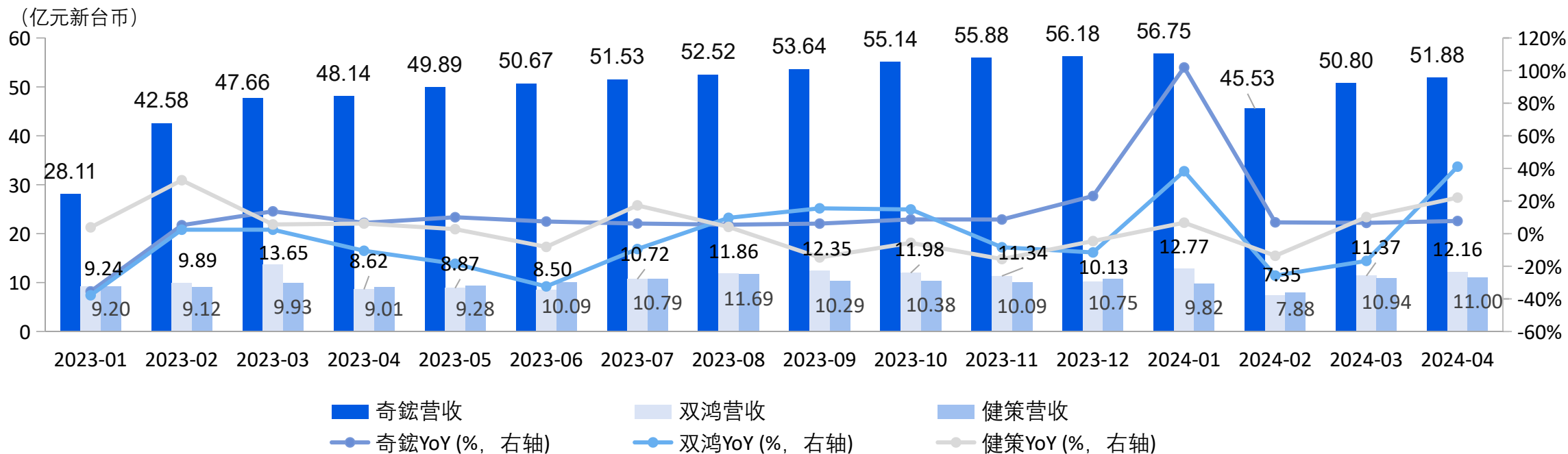
第三章 算力产业链

液冷散热迎倍数增长，信骅BMC数据亮眼

3.1.1 台股散热：AI服务器需求增长打开散产品热成长空间

- 1) 奇鋌：2024年4月，营收51.88亿元新台币，同比+7.77%，环比明显回温。奇鋌预期GP 水冷板部分客户在今年第四季就可望量产，奇鋌初估去年贡献营收仅2%的液冷散热，今年占比可达5-10%，另基于现有订单看来，明年应可创造倍增营收。展望未来，奇鋌表示，第一季因有农历春节，会是全年的谷底，第二季营收和毛利率都会较首季微幅成长，第三季则可望有明显的成长，全年可望朝「季季高」的方向前进。
- 2) 双鸿：2024年4月，营收12.16亿元新台币，同比+41.09%、环比+6.95%；累计前四月营收43.65亿元新台币，同比+5.44%。AI伺服器需求畅旺下，水冷产品出货成长乐观。双鸿目前水冷板月产能为30万组、CDU及分歧管每月则能生产2~3000组，足以因应客户需求。
- 3) 健策：2024年4月，营收10.99亿元，环比+0.5%、同比+22.1%；前四月累计营收39.63亿元，同比+6.4%。健策积极卡位AI伺服器商机，并与辉达（NVIDIA）、超微（AMD）、英特尔等晶片大厂保持密切合作关系，预料今年起均热片出货量有望增长。

图：中国台湾散热厂商月度营收情况



3.1.1 台股散热：2024年，预期3D VC实现放量，液冷板小量出货

- **展望2024Q2：**1) **奇鋌：**2024年第一季高于预期，第二季预期和第一季相当。第二季，AI客户去化零组件库存，但车用等整体需求增加，整体表现会与第1季持平。2) **双鸿：**伺服器接单和水冷需求都见升温，新款NB采用的散热模组陆续出货，虽然VGA第二季表现偏淡，但对第二季展望仍偏正向，法人预期季度环比增幅可望上看1成。
- **展望2024年：**1) **3D VC：**根据奇鋌，2024年用于NVIDIA H100与ASIC伺服器的3D VC模组将显著放量，通用伺服器对3D VC的需求日益提升；2) **液冷板：**奇鋌表示，液冷散热将在2024年下半年起小量出货，并于2025年呈现快速成长；双鸿认为，在新一代GPU带动下，2024H2水冷需求高于风冷。英伟达GB200采用液冷方案，或进一步推动冷板式渗透率提升。

表：2024年，中国台湾散热厂商的展望情况

产品	内容
3D VC	<ul style="list-style-type: none">● 奇鋌：目前奇鋌的3D VC良率已高达90至95%，最高可解到1,000瓦的热能，也因为良率在去年已获得大幅提升，因此未来越南二厂也会开始生产3D VC，借此扩大产能，抢攻AI商机。由于2024年用于NVIDIA H100与ASIC伺服器的3D VC模组将显著放量，加上通用伺服器对3D VC的需求也日益高，预期2024年3D VC模组将显著放量。
液冷板	<ul style="list-style-type: none">● 1) 奇鋌：液冷散热将在今年下半年起小量出货，并于2025年呈现快速成长。奇鋌的水冷板已有先行者优势，包括机柜、风扇、Pump等都能自制，浸没式技术则还在研究阶段。● 2) 双鸿：积极锁定水冷散热商机，目前已加大和客户共同开发的力度，并且已有不少水冷关键零组件开始陆续出货，目标是能提供客户完整的水冷散热解决方案。公司认为2024H2在新一代AIGPU带动下，冷需求将高于气冷。展望双鸿全年，今年的水冷产品占营收比重有望逐季向上，直到第四季占比有望达10%，全年水冷营收占比约中高个位数。预计到2025年，双鸿乐观预期水冷营收占比有望达15%~20%，带动产品组合持续转佳。

3.1.2 内地散热：电信液冷万卡集群运营，液冷产品落地高校与运营商

- 飞荣达、曙光数创、英维克等液冷厂商产品渗透率提升。曙光数创液冷数据中心解决方案在深圳某重要项目成功落地，也适配华南理工大学液冷数据中心；2023年，曙光数创冷板液冷数据中心相关产品营收1.90亿元、同比+430.66%。此外，英维克第一份额中标中国电信弹性DC舱集采项目，含液冷层DC舱，液冷产品持续落地。

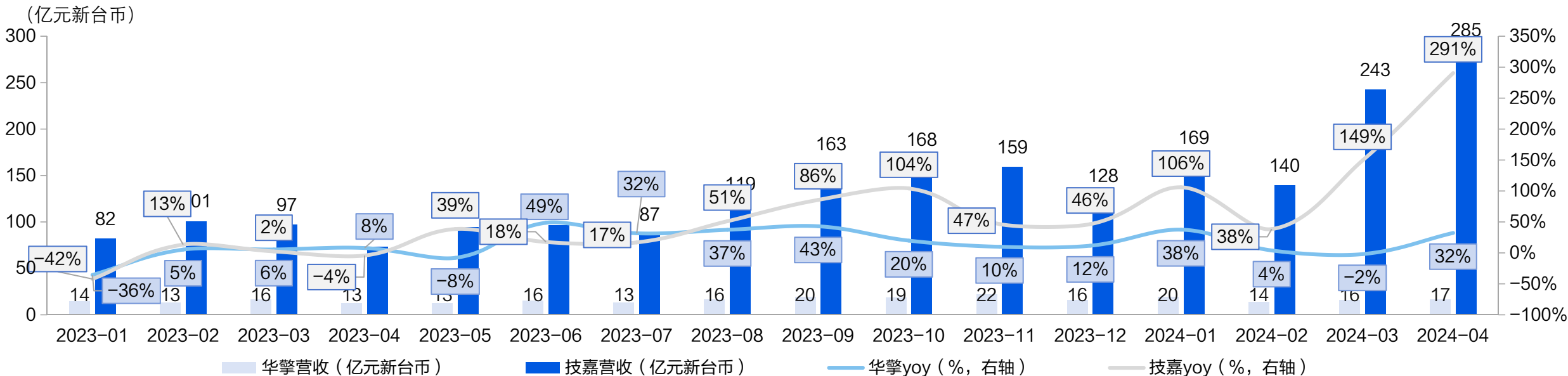
表：各公司散热业务更新

公司	详细内容
飞荣达	<ul style="list-style-type: none">• 整体营收稳步上升。2023年，公司实现营收43.46亿元，同比+5.37%；归母净利润1.03亿元，同比+7.31%；2024Q1，公司实现收入9.82亿元，同比+27.05%；归母净利润0.24亿元，同比扭亏为盈。服务器相关业务开展较为顺利，公司拥有单相液冷模组、两相液冷模组、3D-VC散热模组、轴流风扇及特种散热器等产品。
曙光数创	<ul style="list-style-type: none">• 曙光数创液冷数据中心解决方案在深圳某重要项目成功落地。采用3台高换热量CDU（冷却液分配单元）作为核心设备，每台CDU的换热能力高达1500kW，配套DN200大管径不锈钢二次侧管路，提高冷却液的流量和流速，增强换热效率。• 首个存算一栈式液冷数据中心落地高校。华南理工大学液冷数据中心选用了曙光数创风液混冷先进数据中心解决方案，包含6台300kW CDU，20台液冷机柜；其存储设备采用曙光数创的冷板式液冷技术进行冷却，PUE最低可降至1.1以下。• 2023年，公司实现营收6.50亿元，同比+25.63%；归母净利润1.04亿元，同比-10.63%。冷板液冷数据中心基础设施产品：2023年，营收1.90亿元、同比+430.66%，营收占比29.22%，2022年占比为6.92%；毛利率20.88%、同比-9.24pct。营收快速增长主要系公司认为未来3-5年冷板液冷产品迎来快速发展期，积极布局冷板液冷市场。
英维克	<ul style="list-style-type: none">• 2023年实现营业收入35.29亿元，同比增长20.72%，归母净利润3.44亿元，同比增长22.74%。2024Q1实现营收7.46亿元，同比增长41.36%，归母净利润0.62亿元，同比增长146.93%• 英维克第一份额中标中国电信弹性DC舱集采项目，含液冷层DC舱。4月1日公司发布公告称，英维克第一份额中标中国电信弹性DC舱（2024-2025年）集中采购项目，合计采购弹性DC舱3200套，共涉及5种DC舱模型，初期预计建设323套液冷层DC舱。• 全变频+干冷器自然冷却，英维克储能液冷系统节能全面升级。4月11-13日第十二届储能国际峰会暨展览会上，英维克应用自然冷却技术的储能液冷机组受到客户广泛关注。该产品将自然冷却技术引入储能温控领域，让储能系统能在低温条件下自动切换自然冷却模式，全年能效比AEER最高可达4.62。

3.2.1 主板：厂商营收增长点主要来与AI服务器需求提升

- AI服务器需求旺盛，推升主板出货量提升，叠加第一季主板传统小旺季效益，各厂商2024Q1营收同比增长。
- 1) 华擎：2024年4月，公司营收16.68亿元新台币，同比+32.25%。展望未来，公司表示第2季仍处于传统主机板淡季，预估营收将季减5%；但是AI服务器方面，公司表示超微（AMD）的MI300晶片已经在第1季出货，华擎正在与现有客户讨论规格及生产，由于AI伺服器设计与开发需要两到三季，预估华擎的AI伺服器将在第2季末到第3季量产，届时可望带动业绩明显增长。
- 2) 技嘉：2024年4月，技嘉营收285.24亿元新台币，同比+291%，主要是AI伺服器出货升温所带动。展望未来，预期AI伺服器出货动能将持续，由于技嘉伺服器产品线齐全，且与晶片大厂关系密切，今年在客户需求推升下，加上GPU供给有望逐步舒缓，今年伺服器业务将保持强劲成长幅度。公司表示，第2季步入板卡传统出货淡季，整体板卡出货将较为平淡，不过AI伺服器动能仍正向，将支撑第2季营运表现不看淡。

图：中国台湾主板厂商营收及增速情况



3.2.1 主板：展望2024年，AI服务器营收增长，主机板持平+显示卡增长

- **展望2024年：**1) **服务器：AI服务器：**预期受益于AI服务器的拉货动能，AI服务器相关营收持续增长，但需关注英伟达新款GB200相关产品的供应资格认证情况；**通用服务器：**据技嘉，通用服务器今年拉货已见回暖。2) **主机板：**PC出货预期优于2023年低谷，有望持平或微增。3) **显示卡：**预期需求维持，或受益于厂商自身新品出货动能，带来营收提升。

表：2024Q2及2024年，中国台湾主板厂商的展望情况

	2024年Q2 展望
预期	<ul style="list-style-type: none"> ● 1) AI服务器：华擎旗下高阶AI伺服器产品预期于第二季起陆续出货。华擎的AI伺服器能够涵盖L6(主机板)、L10(整机系统组装)的产品，并预期将以L10 产品出货为主。 ● 2) 通用服务器：预期去年遭到AI影响的通用服务器需求今年将恢复成长，不过动能仍主要来自AI 伺服器，去年AI 伺服器尚无明显贡献公司营收，预计今年第2 季有机会开始出货。整体而言，华擎去年伺服器营收占比下滑，今年占比将提升。
分产品	2024年展望
服务器	<ul style="list-style-type: none"> ● 技嘉表示，2024年：1) 通用型伺服器拉货已见回温；2) AI服务器：公司全年AI伺服器业务的营收比重，将可达其整体伺服器的两到三成以上。AI伺服器出货的关键在芯片，2023年下半年以来的AI需求可以延续到2024年上半年，至于今年下半年因为有英伟达AI晶片的新旧平台转换问题，因此，还需要观察，主要是新旧平台转换涉及到英伟达的芯片分配与客户的选择问题。
主机板	<ul style="list-style-type: none"> ● 华擎表示：2024Q4预期发布新品，在缺乏新品刺激下，今年看法持平。 ● 技嘉表示：PC方面需求预期好于2023年低谷，出货落在1,100万到1,500万片。公司表示厂商都要靠拢AI PC。
显示卡	<ul style="list-style-type: none"> ● 华擎：主要为AMD 平台，预期在AMD 市占率持续扩大之下，有机会同步带动华擎显示卡优于去年表现； ● 技嘉：出货约在350万片

3.2.2 主板：产能持续爬坡，2024Q1实现增长

- 主板厂商产能持续爬坡，2024Q1实现增长。沪电股份表示，泰国生产基地目前处于建设期，预期其在2024年第四季度实现量产。深南电路表示，BT类封装基板保持稳定批量生产，FC-BGA封装基板各阶产品对应的产线验证导入、送样认证等工作得到有序推进。

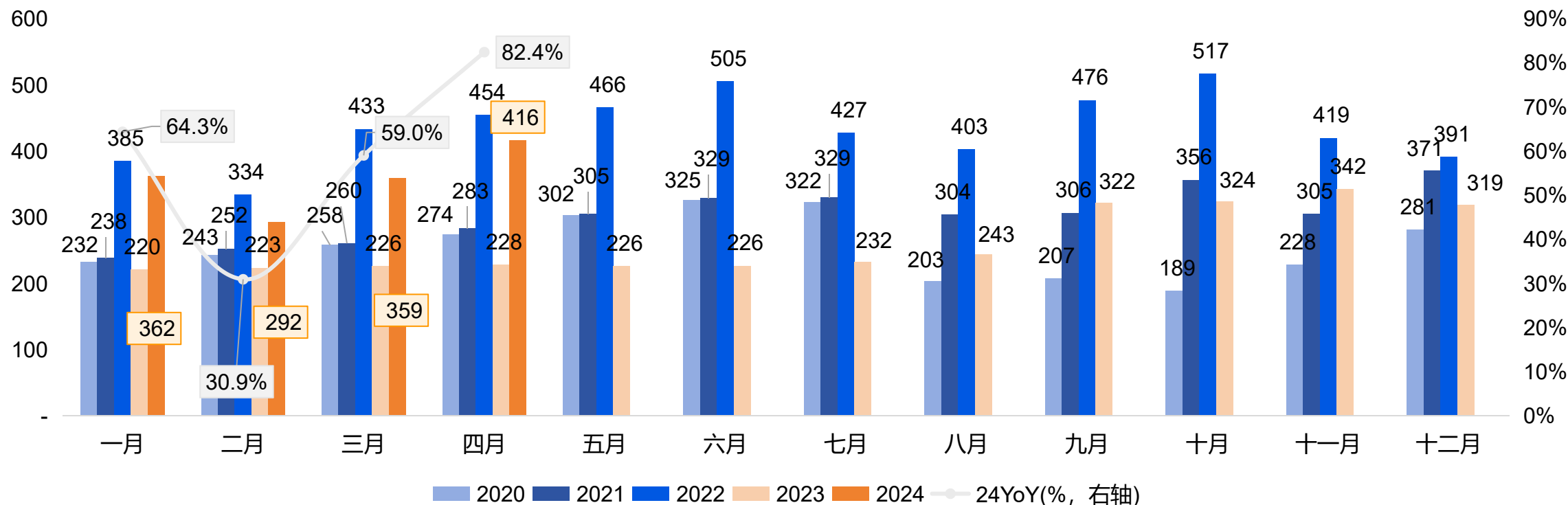
图：基板厂商产能持续爬坡，2024Q1恢复增长

公司	具体内容
沪电股份	<ul style="list-style-type: none"> • 2024Q1,公司实现营业收入约25.84亿元，同比增长38.34%，归母净利润约5.15亿元，同比增长157.03%。 • 4月3日，公司发布公告，泰国生产基地目前处于建设期，预期其在2024年第四季度实现量产 • 4月26日，公司发布公告表示，公司也在整合生产和管理资源，并相应针对性的扩充产能。2024年初公司决议投资约5.1亿元人民币，实施面向算力网络的高密高速互连印制电路板生产线技改项目（下称“技改项目”）。技改项目实施后将提高公司面向算力网络相关产品的HDI阶数、层数，提升相关产品的品质稳定性及可靠性，提升生产线数字化、智能化水平。
深南电路	<ul style="list-style-type: none"> • 2024Q1，公司实现营业收入39.61亿元，同比+42.24%，归母净利润3.80亿元，同比+83.88%，PCB、封装基板、电子装联三项业务营业收入同比均有提升。 • 4月16日公司接受机构调研表示，BT类封装基板保持稳定批量生产，FC-BGA封装基板各阶产品对应的产线验证导入、送样认证等工作得到有序推进。

3.3 信骅：BMC新品渗透率+ASP提升，AI服务器需求增长

- 2024年4月，信骅单月营收达4.159亿元，环比+16%、同比+82%。法人表示，通用服务器市况历经一年库存调整，BMC晶片需求开始回温。信骅受惠BMC新品AST2600出货有望随著新平台的渗透率提升，推升营运动能，带动整体平均销售单价(ASP)提升。生成式AI应用热潮兴起，AI服务器相关需求逐渐增加，信骅成功搭上NV AI服务器商机，今年出货动能将可望大幅成长。法人指出，辉达AI服务器今年订单仍满手，让信骅BMC订单动能也看增，预期今年相关出货动能可望逐季攀升。

图：信骅月度营收情况（亿元新台币）



3.3 BMC：增长点来自AI服务器与VoIP芯片推动

- **展望2024年：**1) **库存：**历经库存调整后，今年市场需求将出现延续性回温。2) **服务器：**全球伺服器出货量将呈现平稳，通用服务器或见回升；AI推理服务器需求将快速提升；AI伺服器对资本支出的侵蚀现象将减缓，且AI伺服器对BMC的需求将会提高；3) **VoIP芯片：**预期全球VoIP芯片的出货量将以15.9%的复合年增长率成长；作为1G AVoIP市场的领导者，信骅将持续拓展AVoIP的新市场。

表：2024Q2及2024年，中国台湾BMC相关厂商的展望情况

	2024年Q2 展望
整体	<ul style="list-style-type: none"> ● 展望2024Q2，信骅预期营收约10.58~11.34亿元新台币、同比+55.6%~66.8%，毛利率达64.0%~65.5%。
分产品	2024年展望
整体	<ul style="list-style-type: none"> ● 信骅表示，历经库存调整后，今年市场需求将出现延续性回温，加上英特尔新伺服器平台加持，以及AI伺服器所需管理晶片数量成长，mini BMC客户数量也增加带动，信骅今年业绩料重回成长轨迹，营运可望逐季成长。
云端企业	<ul style="list-style-type: none"> ● 信骅表示，全球伺服器出货量将呈现平稳，AI推理服务器需求将快速提升；AI伺服器对资本支出的侵蚀现象将减缓，且AI伺服器对BMC的需求将会提高；最大化每台服务器内含的信骅芯片价值(每台服务器美元含量)；服务器主板模组化(模块化)。 ● 信骅表示，英伟达AI伺服器今年订单较满，让信骅BMC订单动能也看增，预期今年相关出货动能可望逐季攀升。
智慧AV解决方案	<ul style="list-style-type: none"> ● 信骅预计AVOIP在2024年将维持其双位数的增长率；酷博乐360迅速布署到智慧工厂以及其他应用场景。信骅及旗下子公司酷博乐先前宣布携手鸿海科技集团子公司鸿佰科技，以及超大规模AI 3D建模技术供应商VIZZIO，引进创新智慧巡检解决方案至鸿佰科技桃园南青厂，实现远端即时沉浸式管理。

3.4 光模块：OFC 2024新品频发，光通信全产业链加速升级

- **OFC 2024新品频发，光通信全产业链加速升级。** OFC 2024大会上，完成了从800G、1.6T等高端产品到相关配套组件全产业链的升级：新易盛展出业界第一个基于200G VCSEL的1.6T产品；天孚通信展示了适配800G/1.6T的光引擎产品；光迅科技联合思科正式发布1.6T硅光模块。据LightCounting，预计2024年光模块销售额将增长40%，2025年增长20%以上，2026年-2027年保持两位数增长。2024Q1，中际旭创、新易盛、天孚通讯等公司归母净利润均呈现倍数同比增长。

表：OFC 2024新品频发，光通信全产业链加速升级

公司	内容
中际旭创	<ul style="list-style-type: none">• 2024Q1，公司实现营业收入48.43亿元，同比增长163.59%；实现归母净利润10.09亿元，同比增长303.84%。• 800G：预计800G产品在订单可见度的二季度的出货量会保持环比增长。
新易盛	<ul style="list-style-type: none">• 2024Q1，公司实现营业总收入11.13亿元，同比增长85.41%；归母净利润3.25亿元，同比增长200.96%• 212Gb/s VCSEL：OFC 2024上，新易盛展示了业界首款基于212 Gb/s VCSEL技术搭建的光模块传输链路，验证了VCSEL仍是下一代数据中心和 AI/ML 集群短距离高速互连的一种竞争技术方案。
天孚通讯	<ul style="list-style-type: none">• 2024Q1，公司实现营收7.3 亿元，同比增长155%，实现归母净利润2.8 亿元，同比增长203%• 800G/1.6T：OFC 2024大会上，天孚通信重点展示了为1.6T/800G光模块配套应用的Mux TOSA、Demux POSA、Lensed FAU等光引擎产品和解决方案。
光迅科技	<ul style="list-style-type: none">• 1.6T：OFC 2024大会上，光迅科技联合思科正式发布全球领先的1.6T硅光模块，并率先在展会现场公开演示。• 800G：4月30日公司表示，目前公司200G\400G\800G硅光芯片已具备批量能力，正逐步应用于公司的产品中。1.6T 光模块属于下一代的产品，预计最快2024年下半年正式商用，目前公司1.6T产品正在客户验证阶段。

第四章

CoWoS产能快速增长，CSP厂商资本开支提升

4.1 CoWoS：台积电订单需求强劲，Blackwell出货有望提速

- 台积电订单需求强劲，产能大幅增长。AI芯片需求提升，英伟达对CoWoS需求或增长三倍。据4月16日财联社，英伟达GB200预估2025年出货量有机会突破百万颗，将耗费更多CoWoS产能，台积电提升2024全年CoWoS产能需求，推动持续建厂扩产，日本、美国新工厂启动，预估至年底每月产能将逼近4万，相较2023年总产能提升逾150%。

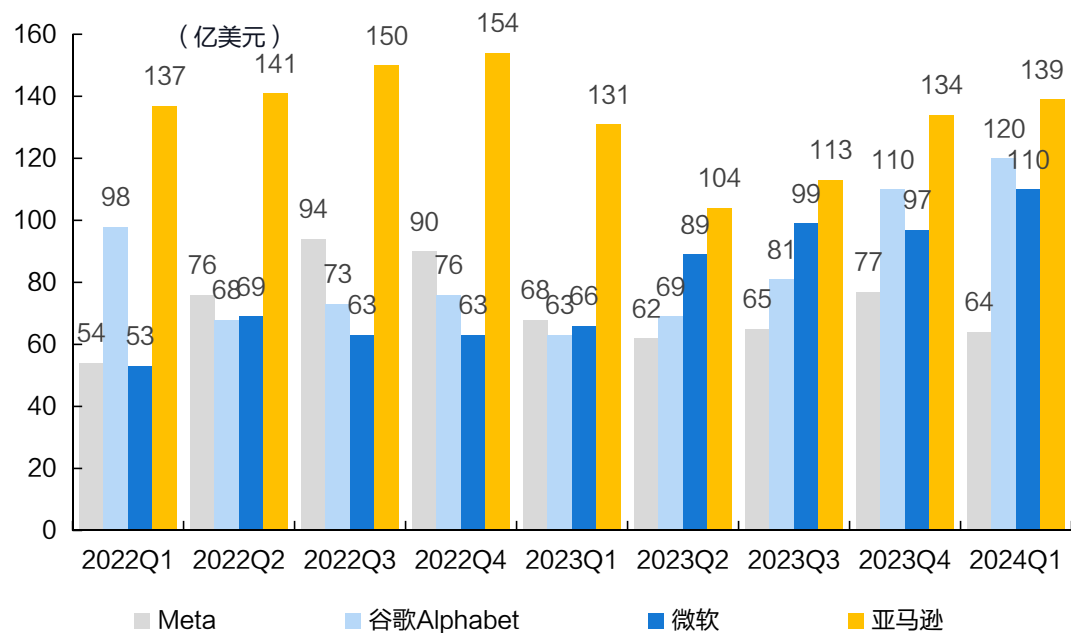
表：台积电订单需求强劲，产能大幅增长

分类	具体内容
产能	<ul style="list-style-type: none">• 据4月16日财联社，TrendForce指出，英伟达GB200预估2025年出货量有机会突破百万颗，将耗费更多CoWoS产能，台积电提升2024全年CoWoS产能需求，预估至年底每月产能将逼近4万，相较2023年总产能提升逾150%。
制程	<ul style="list-style-type: none">• 据Counterpoint，预期台积电3nm旗舰智能手机应用将在2024H2成长，但2nm制程量产将延迟至2026年底。
扩产	<ul style="list-style-type: none">• 据4月8日财联社，美国商务部与台积电达成初步协议，台积电将在亚利桑那州凤凰城建立第三家工厂，总投资增加至650亿美元。台积电表示，台积电在亚利桑那州的第一家晶圆厂将于2025H1开始生产，采用4nm技术；第三家晶圆厂将采用2nm或更先进的工艺生产芯片，将于本十年末开始生产。• 据5月13日财联社，台积电未来3年将是扩产再扩产的盛况，预估2024年第4季至2026年，龙潭、竹南、中科、南科及日前发布建厂的嘉义AP7厂将陆续启动。

4.2 资本开支：海外CSP大厂云收入高增，且上修资本开支指引

- 各大CSP厂商推动AI商业化落地：**
 - 1) 微软：**据Microsoft start, FY24Q3智能云业务同比增长21%，Azure 和其他云服务收入增长31%，AI 贡献7%增长，上一季度为6%；
 - 2) 谷歌：**据澎湃新闻, 2024Q1, 谷歌云收入95.7亿美元, 同比增长28.4%；
 - 3) Meta：**据Meta业绩会, 公司发布由 Llama 3 模型支持的 Meta AI 新版本, Facebook feed 上约 30% 的帖子和 Instagram 上超过 50% 的内容由人工智能推荐系统提供；
 - 4) 亚马逊：**据财联社, 2024Q1, AWS收入250亿美元, 同比增长 17%, 超出预期的241亿美元。
- 展望2024年，Meta、Alphabet（谷歌）、微软、亚马逊等厂商均增加资本开支指引，加大力度发展人工智能。**微软预计FY2024Q4资本开支环比大幅提升，谷歌表示全年季度资本开支将等于或高于一季度水平，Meta预计2024年全年资本支出在350-400亿美元之间（原指引300-370亿美元），亚马逊表示资本支出将增长来支持AWS的增长，包括生成式人工智能。

图：2022-2024Q1 各厂商资本性开支



表：各CSP厂商2024年资本开支指引

分类	具体内容
Microsoft	<ul style="list-style-type: none"> FY2024Q3 (截至2024.3.31), 资本开支110亿美元, 同比+66%。公司表示下一季度资本开支将环比大幅提升, 并且2025财年的资本开支仍将高于2024财年
Alphabet (谷歌)	<ul style="list-style-type: none"> 2024Q1, 资本开支120亿美元、同比+90%左右。公司预计, 全年的季度资本开支将等于或高于一季度水平
Meta	<ul style="list-style-type: none"> Meta预计2024年全年资本支出在350-400亿美元之间 (原指引300-370亿美元), 并预计明年资本支出将继续增加。
亚马逊	<ul style="list-style-type: none"> 2024Q1, 资本开支139亿美元, 同比+7%。公司预计, 2024Q1会是全年资本开支的低点, 资本支出将在2024年“有意义地增加”, 主要用于支持AWS的增长, 包括生成式人工智能

第五章 投资建议与风险提示

- 大模型训推带动 AI算力需求增长，算力产业链中的AI芯片、服务器整机及零组件、光模块、IDC等环节有望持续受益。维持对计算机行业“推荐”评级。

- 相关公司

- 1) 服务器整机：浪潮信息、工业富联、中科曙光、华勤技术、紫光股份、中国长城、高新发展、神州数码、烽火通信、拓维信息、软通动力、纬创、广达、英业达、纬颖、超微电脑。
- 2) 服务器组件：①AI芯片：海光信息、寒武纪、龙芯中科、景嘉微、英伟达、AMD、Intel；②散热：飞荣达、曙光数创、英维克、同方股份、申菱环境、高澜股份、奇鋆科技、双鸿、健策、VERTIV；③主板：沪电股份、深南电路、胜宏科技、技嘉、华擎。
- 3) 光模块：天孚通信、中际旭创、新易盛、光迅科技、华工科技。
- 4) 数据中心：奥飞数据、光环新网、宝信软件、数据港、电科数字。

5.1 投资建议与相关公司

- ☆主板 (0.3%)
 - 内地: 沪电股份、胜宏科技、深南电路、生益科技
 - 中国台湾: 泰安电脑(神达)、技嘉、华擎、华硕、英业达、纬创、研华
 - 美国: Intel、Supermicro
- CPU (1.9%)
 - 内地: 华为海思、海光信息、飞腾信息、龙芯中科
 - 海外: Intel、AMD
- 电源模块 (0.4%)
 - 内地: 中国长城、欧陆通、杰华特、泰嘉股份
- 光模块
 - 内地: 中际旭创、新易盛、光迅科技、天孚通信
- 灵活 I/O 卡
 - 智能网卡 (4.1%)
 - 内地: 星云智联、大禹智芯、云豹智能
 - 中国台湾: 研华、凌华、其阳
 - 美国: Mellanox (NVIDIA)

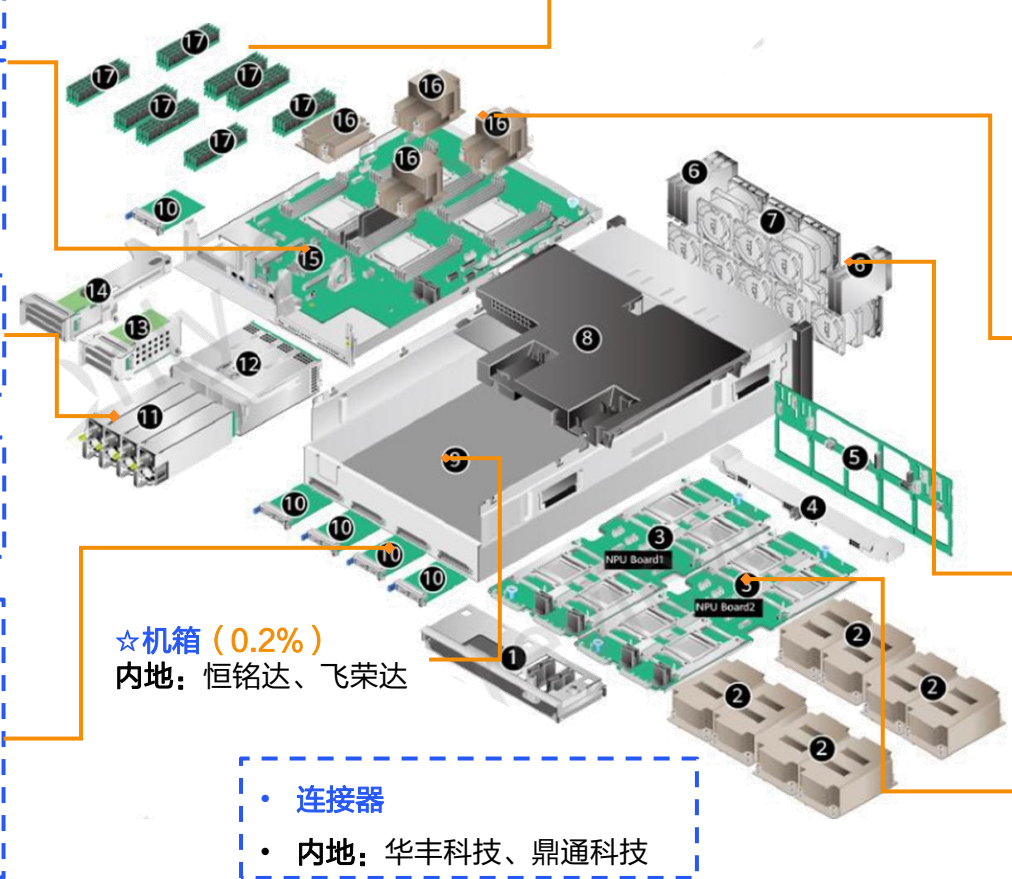
- DIMM
 - 内存 DRAM (2.9%)
 - 内地: 澜起科技、聚辰股份、江波龙、长鑫存储
 - 海外: 三星、SK海力士、美光

- ☆组装测试 (0.6%)
 - 整机厂商ODM: 1) 内地: 浪潮信息、工业富联、中科曙光、紫光股份、中国长城、华勤技术、神州数码、拓维信息、烽火通信、软通动力; 2) 中国台湾: 广达、纬创、英业达、纬颖; 3) 美股: 戴尔、超微电脑
 - BIOS/BMC: 1) 内地: 卓易信息; 2) 中国台湾: 新唐、系微、信骅

- ☆散热 (0.2%)
 - 内地: 曙光数创、飞荣达、中航光电、英维克、同飞股份、申菱环境、高澜股份、依米康
 - 中国台湾: 奇鋌、双鸿、健策、建准、高力

- 存储 NAND (1.3%)
 - 公司: 长江存储、兆易创新、佰维存储、朗科科技
 - 海外: 三星、西部数据、铠侠、SK海力士、美光

- AI芯片-GPU、NPU等 (72.5%)
 - 内地: 华为海思、海光信息、寒武纪、龙芯中科
 - 海外: NVIDIA、AMD、Intel



☆机箱 (0.2%)
内地: 恒铭达、飞荣达

- 连接器
 - 内地: 华丰科技、鼎通科技

注: 蓝字为零部件, 括号中橙色数字为价值量占比。

上图展示以华为 Atlas 800 训练服务器为例 (鲲鹏 920 * 4, 昇腾910 * 8), 其中CPU集成在主板上, NPU集成在NPU板上; 具体价值量数字对标Nvidia DGX H100服务器

- 1) **宏观经济影响下游需求**：宏观经济环境下行，将影响客户对信息化基础设施的采购需求；
- 2) **大模型产业发展不及预期**：行业的核心驱动力是人工智能大语言模型的训练和推理对算力的需求，如果大模型行业发展不及预期将会影响AI算力的相关需求；
- 3) **市场竞争加剧**：IT 产品和服务行业是成熟且完全竞争的行业，新进入者可能加剧整个行业的竞争态势；
- 4) **中美博弈加剧**：国际形势持续不明朗，美国不断通过“实体清单”等方式对中国企业实施打压，若中美紧张形势进一步升级，将可能导致中国半导体供应链供应受到影响；
- 5) **相关公司业绩不及预期**：市场环境变化、公司治理情况变化、其他非主营业务经营不及预期等原因或将导致相关公司的整体业绩不及预期。
- 6) **各公司并不具备完全可比性，对标的相关资料和数据仅供参考。**

计算机小组介绍

刘熹，计算机行业首席分析师，上海交通大学硕士，多年计算机行业研究经验，致力于做前瞻性深度研究，挖掘投资机会。新浪金麒麟新锐分析师、Wind金牌分析师团队核心成员。

分析师承诺

刘熹，本报告中的分析师均具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立，客观的出具本报告。本报告清晰准确的反映了分析师本人的研究观点。分析师本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收取到任何形式的补偿。

国海证券投资评级标准

行业投资评级

推荐：行业基本面向好，行业指数领先沪深300指数；
中性：行业基本面稳定，行业指数跟随沪深300指数；
回避：行业基本面向淡，行业指数落后沪深300指数。

股票投资评级

买入：相对沪深300 指数涨幅20%以上；
增持：相对沪深300 指数涨幅介于10%~20%之间；
中性：相对沪深300 指数涨幅介于-10%~10%之间；
卖出：相对沪深300 指数跌幅10%以上。

免责声明

本报告的风险等级定级为R3，仅供符合国海证券股份有限公司（简称“本公司”）投资者适当性管理要求的客户（简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。客户及/或投资者应当认识到有关本报告的短信提示、电话推荐等只是研究观点的简要沟通，需以本公司的完整报告为准，本公司接受客户的后续问询。

本公司具有中国证监会许可的证券投资咨询业务资格。本报告中的信息均来源于公开资料及合法获得的相关内部外部报告资料，本公司对这些信息的准确性及完整性不作任何保证，也不保证其中的信息已做最新变更，也不保证相关的建议不会发生任何变更。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。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。报告中的内容和意见仅供参考，在任何情况下，本报告中所表达的意见并不构成对所述证券买卖的出价和征价。本公司及其本公司员工对使用本报告及其内容所引发的任何直接或间接损失概不负责。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露义务。

风险提示

市场有风险，投资需谨慎。投资者不应将本报告为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向本公司或其他专业人士咨询并谨慎决策。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息。本报告不构成本公司向该机构之客户提供的投资建议。

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。本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

郑重声明

本报告版权归国海证券所有。未经本公司的明确书面特别授权或协议约定，除法律规定的情况外，任何人不得对本报告的任何内容进行发布、复制、编辑、改编、转载、播放、展示或以其他方式非法使用本报告的部分或者全部内容，否则均构成对本公司版权的侵害，本公司有权依法追究其法律责任。

国海证券 · 研究所 · 计算机研究团队

心怀家国，洞悉四海



国海研究上海

上海市黄浦区绿地外滩中心C1栋
国海证券大厦

邮编：200023

电话：021-61981300

国海研究深圳

深圳市福田区竹子林四路光大银
行大厦28F

邮编：518041

电话：0755-83706353

国海研究北京

北京市海淀区西直门外大街168
号腾达大厦25F

邮编：100044

电话：010-88576597