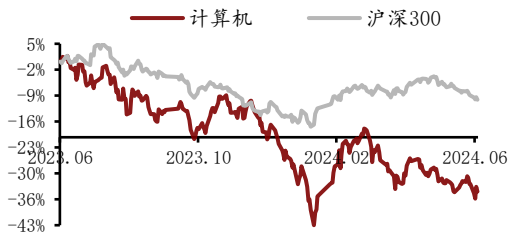


证券研究报告-行业半年度策略

同步大市(下调)

计算机相对沪深 300 指数表现

发布日期：2024 年 07 月 01 日



资料来源：中原证券研究所，聚源

相关报告

《计算机行业专题研究：数字经济产业链分析之河南概况》 2024-05-29

《计算机行业专题研究：政策聚焦新质生产力，2024 人工智能将加速发展》 2024-03-06

《计算机行业专题研究：Sora 发布后，重新审视 AI 的发展走向》 2024-02-23

投资要点：

- 给予行业同步大市的投资评级。在全球经济景气度下滑的大背景下，我们注意到计算机行业虽然需求保持相对的稳定性，费用得到较好的控制，但是在收入质量方面仍然存在隐忧，或持续对业绩形成压力。在这样的发展背景下，行业不同赛道甚至是赛道中的不同企业之间的景气度差异或将进一步加大。
- 2024 年上半年，计算机行业跌幅居前，当前计算机行业估值明显低于中长期均值水平，处于近 3 年估值较低的位置，整体风险得到较好的释放，下半年可以重点聚焦的方向包括：

(1) 人工智能：作为当前创新推动的核心焦点，人工智能领域孕育了巨大的发展机遇。在 OpenAI 的 Sora、GPT-4o 发布的同时，开源模型也在加速壮大，对我国大模型产业进步起到了积极的推动作用。加速芯片领域英伟达持续引领，垄断趋势明显，但是国内需求在美国禁令之下，形成了对国产芯片厂商难得的发展机遇。在 OpenAI 暂停中国 API 接口、国内大模型的降价潮的影响下，国内大模型将获得更好的应用。推荐个股金山办公 (688111)。

(2) 算力建设：大模型加速发展，算力建设重要性进一步凸显。在“适度超前建设”、智算比例持续提升的大趋势下，我们看到 2024 年运营商在加快算力布局、国内外科技大厂加大资本投入、地方算力建设相关政策集中发布，也将给产业链带来更多发展机遇。AI 计算、液冷、整机柜的结合将成为未来的发展趋势，液冷行业也迎来了刚需放量的元年。推荐个股紫光股份 (000938)。

(3) 国产替代：在禁令加剧的大趋势下，我们看到芯片的国产化比例在持续提升、运营商集采国产化比例逐年增长、地方积极推动算力的国产化，工业领域更新换代也将推动国产化发展进程。此外，华为在算力芯片和纯血鸿蒙的动向，也将成为 2024 值得重点关注的方向。推荐个股中望软件 (688083)、中控技术 (688777)、索辰科技 (688507)。

风险提示：国际局势的不确定性；下游企业削减开支；金融风险的释放。

内容目录

1. 2024 年行业费用压力有望缓解，但仍需关注毛利率水平	4
1.1. 行业数据：2024 年行业收入增速略有下滑，盈利能力略有改善	4
1.2. 上市公司财报数据：2024 年费用压力有望缓解，但收入质量还需重点关注	4
1.2.1. 年报数据	4
1.2.2. 一季报数据	6
1.3. 高景气赛道：IC 设计、云服务与大数据服务	7
2. 人工智能加速发展，行业变革持续深入	8
2.1. 谷歌、OpenAI 竞争大模型第一梯队，推动大模型加速发展	8
2.2. 开源大模型队伍加速壮大，对我国大模型发展起到了积极推动作用	9
2.3. 英伟达改写行业运行逻辑，成为市场最大赢家	10
2.4. 美国禁令趋势下，国产芯片阵营获得难得的发展机遇	14
2.5. OpenAI 暂停中国 API 接口，国内大模型将获得更好应用	15
3. 2024 年算力需求增长趋势明显	16
3.1. 我国近期算力规划及总体完成情况	16
3.2. 2024 年算力建设加速迹象明显	17
3.2.1. 2024 年运营商加快算力布局	17
3.2.2. 国内外科技大商加大资本投入	18
3.2.3. 地方算力建设相关政策集中发布	18
3.3. 液冷服务器将迎来快速发展机遇	20
4. 国产化推行持续深入，纯血鸿蒙商用值得期待	21
4.1. 禁令加剧，华为、AI、芯片是核心聚焦点	21
4.2. 国产化推进持续深入	22
4.2.1. 芯片环节国产化比例持续提升	22
4.2.2. 运营商采购国产化率呈上升趋势	23
4.2.3. 地方算力国产化比例要求成为发展指引	24
4.2.4. 工业领域的更新换代，有望加速工业领域软件的国产化进程	24
4.3. 华为引领下的国产突围	26
4.3.1. 华为算力芯片布局加快趋势明显	27
4.3.2. 纯血鸿蒙与盘古大模型结合，打造系统级原生智能	28
5. 投资策略	29
5.1. 行情回顾：2023 年人工智能、华为概念带动 TMT 领涨	29
5.2. 估值：明显低于中长期均值水平，处于近三年估值较低的位置	31
5.3. 行业观点与投资建议	31
6. 风险提示	32

图表目录

图 1：2016-2024 年我国软件业务收入及增速（月度累计值）	4
图 2：2016-2024 年我国软件业务利润总额及增速（月度累计值）	4
图 3：中信计算机行业营收及增速（亿元）	5
图 4：中信计算机行业净利润及增速（亿元）	5
图 5：中信计算机行业毛利及增速（亿元）	5
图 6：中信计算机行业信用减值和资产减值（亿元）	5
图 7：中信计算机行业员工数量及增速	6

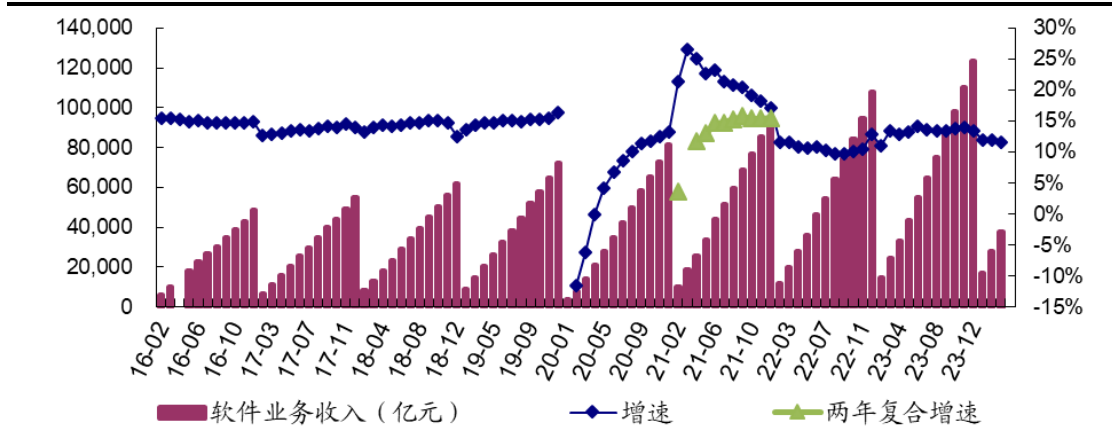
图 8: 中信计算机行业三大费用及增速 (亿元)	6
图 9: 中信计算机行业 Q1 营收及增速 (亿元)	6
图 10: 中信计算机行业 Q1 净利润及增速 (亿元)	6
图 11: 中信计算机行业毛利及增速 (亿元)	7
图 12: 中信计算机行业信用减值和资产减值 (亿元)	7
图 13: 2024 年 1-4 月和 2023 年 1-4 月我国软件业务子行业增速对比	7
图 14: 2021.1-12 至 2024.1-4 我国电子产品产量增速 (累计同比)	8
图 15: 2016Q1-2024Q1 三大芯片厂商数据中心业务收入及增速 (亿美元)	11
图 16: 英伟达 B200 芯片	13
图 17: 英伟达 GB200 超级芯片	13
图 18: 一个 Blackwell 计算节点	13
图 19: 英伟达 DGX GB200 NVL72	13
图 20: 英伟达 DGX SuperPOD	14
图 21: 2023 年-2028 年中国加速计算服务器市场预测 (百万元)	15
图 22: 斯坦福大模型榜单 HELM MMLU (2024 年 6 月 11 日)	16
图 23: 2022.6-2025E 我国算力规模及结构 (EFLOPS)	17
图 24: 2020-2024 年中国移动资本开支增速及 5G、算力的开支增速	18
图 25: 2020Q1-2024Q1 BAT 的资本开支 (亿元)	18
图 26: 19Q1-24Q1 5 大科技厂商资本投入 (百万美元)	18
图 27: 2022 年中国液冷服务器规模及增速 (亿美元)	20
图 28: 单机柜密度和冷却方式	20
图 29: 2022 年我国液冷服务器市场厂商份额	21
图 30: 2023 年我国液冷服务器市场厂商份额	21
图 31: 2015Q4-2024Q1 全球智能手机市场厂商市场份额	21
图 32: 2022Q1-2024Q1 英伟达单季度收入的区域分布 (亿美元)	22
图 33: 2021-2024 年我国集成电路对进口依赖度和出口占比	23
图 34: 2022Q2-2024Q1 全球智能手机操作系统份额	28
图 35: 2022Q2-2024Q1 中国智能手机操作系统份额	28
图 36: 2024 年中信一级子行业涨跌幅 (截止 2024.6.27)	29
图 37: 近三年计算机行业相关概念年度涨跌幅 (% , 2024.6.27)	31
图 38: 近 10 年申万计算机行业估值水平 (截止 2024.6.27)	31
表 1: OpenAI 大模型产品	8
表 2: 谷歌大模型产品	9
表 3: 主要开源模型对比	10
表 4: 英伟达、AMD 人工智能芯片参数对比	11
表 5: 近期省级算力基础设施相关规划	19
表 6: 中国电信及中国移动服务器集采情况及国产化比例	23
表 7: 近期省级算力基础设施相关规划	25
表 8: 近年来华为核心产品发布的情况	26
表 9: 服务器相关业务财务对比 (亿元)	28

1. 2024 年行业费用压力有望缓解，但仍需关注毛利率水平

1.1. 行业数据：2024 年行业收入增速略有下滑，盈利能力略有改善

相比于 2023 年来看，我国软件产业在 2024 年的增速略有回落。根据工信部数据，2024 年 1-4 月软件业务收入 3.79 万亿元，同比增长 11.6%，较 2023 年全年 13.4% 的增速回落了 1.8 PCT。

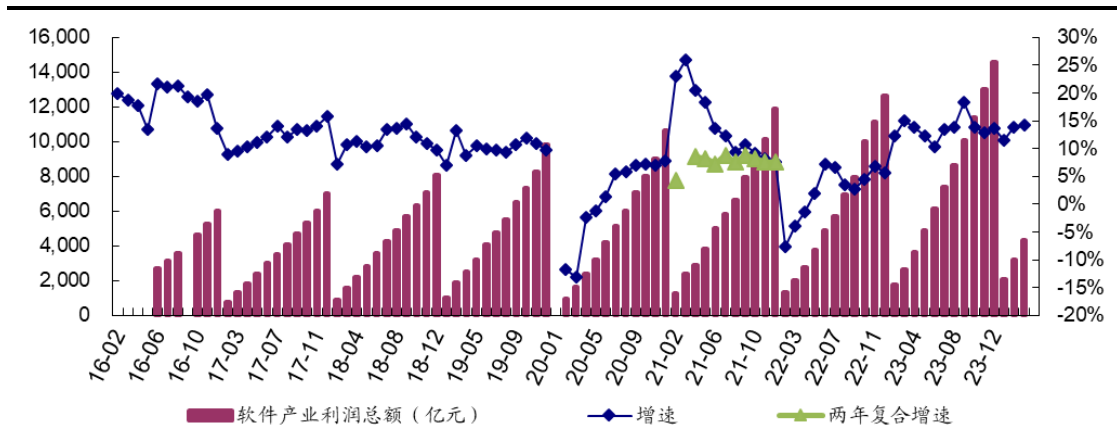
图 1：2016-2024 年我国软件业务收入及增速（月度累计值）



资料来源：工信部，中原证券研究所

利润方面，软件利润增速继续保持较好的稳定性，且盈利水平获得改善。2024 年 1-4 月，软件业务利润总额 4314 亿元，同比增长 14.3%，较 2023 年全年 13.6% 的增速水平提升了 0.7 PCT，且高于同期收入增速 2.7 PCT，盈利能力获得改善。

图 2：2016-2024 年我国软件业务利润总额及增速（月度累计值）



资料来源：工信部，中原证券研究所

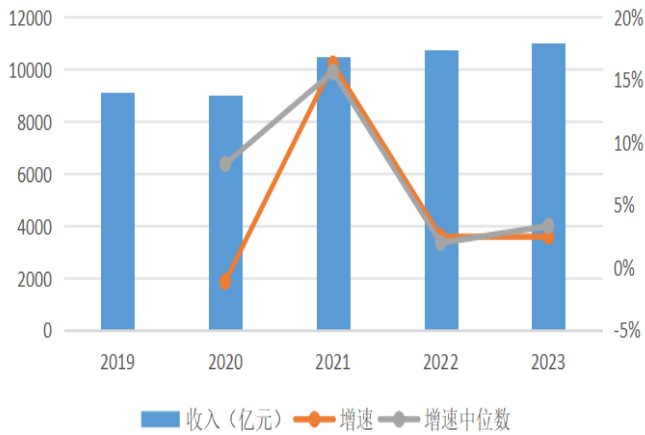
1.2. 上市公司财报数据：2024 年费用压力有望缓解，但收入质量还需重点关注

1.2.1. 年报数据

从中信计算机行业数据来看，2023 年行业总计收入 1.58 万亿。考虑到工业富联占比过

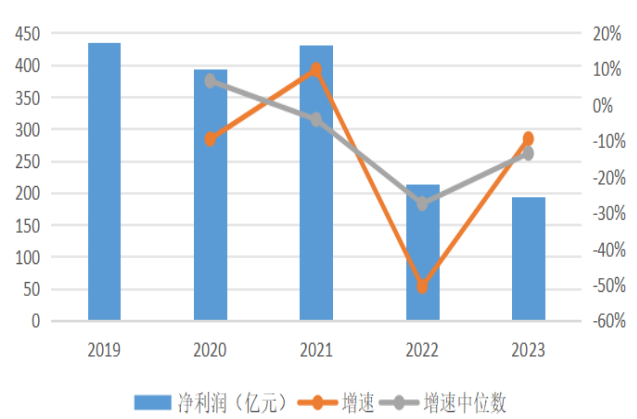
高，将其剔除后，行业总计收入 1.10 万亿，增速在 2022 年见底 (2.45%) 后略有回升至 2.49%。去除富士康后，行业全年盈利 193.17 亿元，净利润增速从 2022 年的-50.54% 回升到 -9.47%，净利润总额仅为 2019 年峰值的 44%。

图 3：中信计算机行业营收及增速（亿元）



资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

图 4：中信计算机行业净利润及增速（亿元）

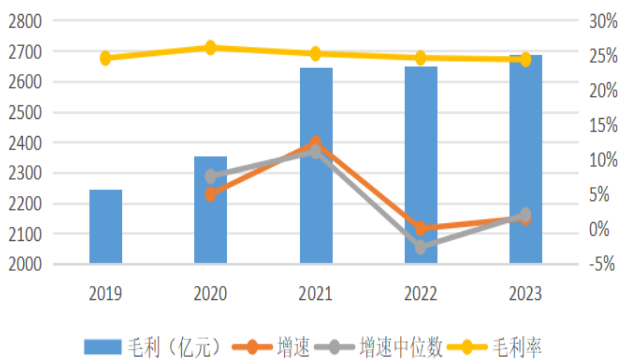


资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

从行业毛利情况来看，2023 年行业上市公司（除工业富联）毛利总计 2689 亿元，同比增速 1.54%，较上年回升了 1.50 PCT，但是连续三年低于收入增速水平。2023 年总体毛利率 24.36%，较 2019 年 24.54% 的毛利率水平来看，总体相差不大。

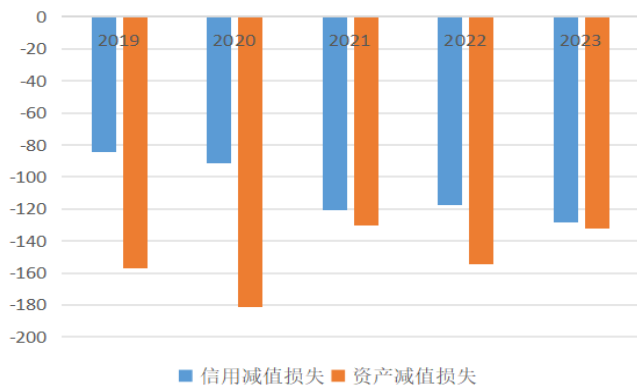
此外，行业上市公司（除工业富联）2023 年信用减值和资产减值分别为 128.48 亿元和 132.54 亿元，两者总计较上年减少了 11.09 亿元。

图 5：中信计算机行业毛利及增速（亿元）



资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

图 6：中信计算机行业信用减值和资产减值（亿元）



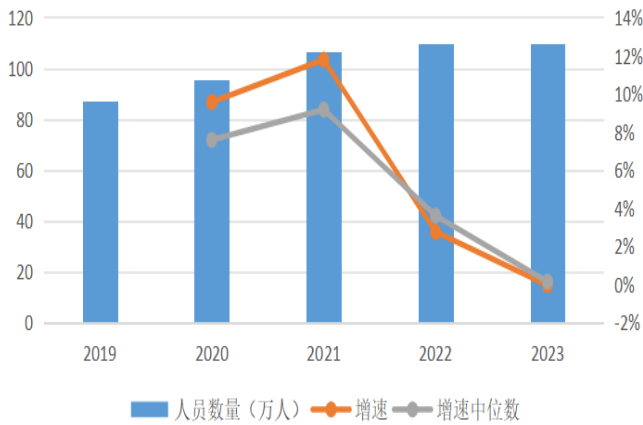
资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

行业 2023 年增收不增利主要源于费用的增长。2023 年行业（除工业富联）销售费用、管理费用、研发费用同比分别增长了 8.12%、5.77%、4.38%，三大费用合计增长 5.97%，三大费用合计占比收入的 21.37%，较上年提升了 0.70 PCT。

2023 年行业费用增长的背后，我们认为主要有 4 大因素驱动：（1）2022 年疫情因素导

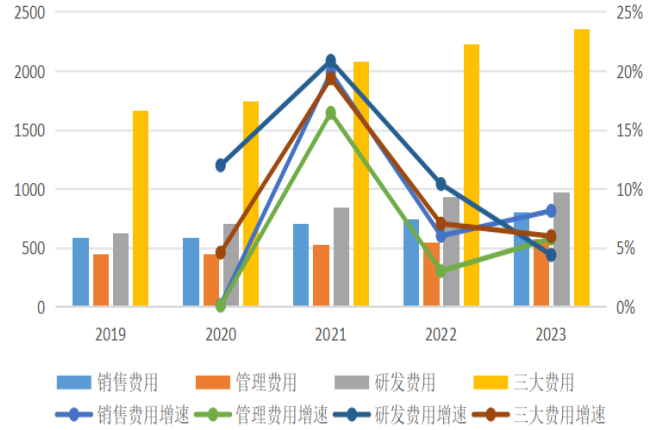
致业务无法正常开展，从而导致当期费用基数的减少；(2) 2022 年行业人员增速 3%，相关费用增长会延迟体现在 2023 年；(3) ChatGPT 为代表的 AI 大模型给行业带来巨大变革，也促进了企业加大研发投入；(4) 2023 年企业人员缩减，带来解除劳动合同相关的赔付费用。

图 7：中信计算机行业员工数量及增速



资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

图 8：中信计算机行业三大费用及增速（亿元）



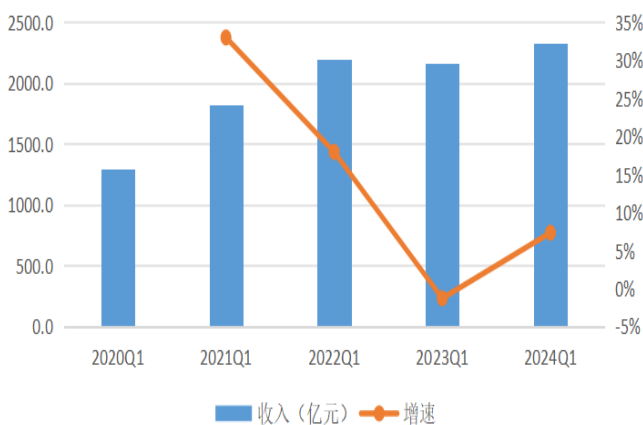
资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

考虑到 2023 年行业上市公司（除工业富联）员工数量已经呈现下滑迹象，我们认为 2024 年行业费用投入有望保持下滑或相对持平的水平。

1.2.2. 一季报数据

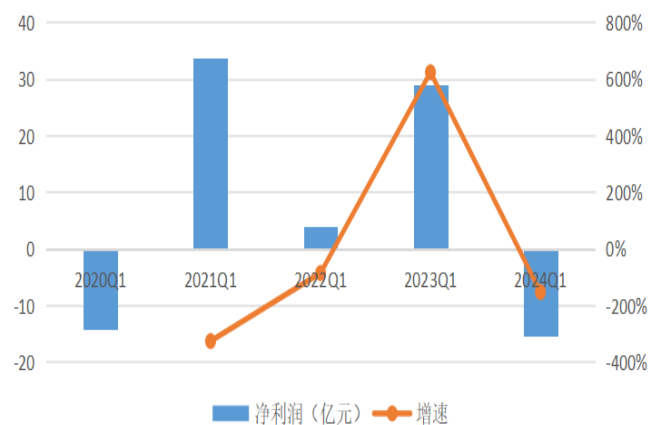
2024Q1 来看，行业（除工业富联）收入 2324.6 亿元，同比增长 7.34%，较上年同期改善明显，但同期净利润-15.50 亿元，继 2020Q1 后再次步入亏损状态。

图 9：中信计算机行业 Q1 营收及增速（亿元）



资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

图 10：中信计算机行业 Q1 净利润及增速（亿元）

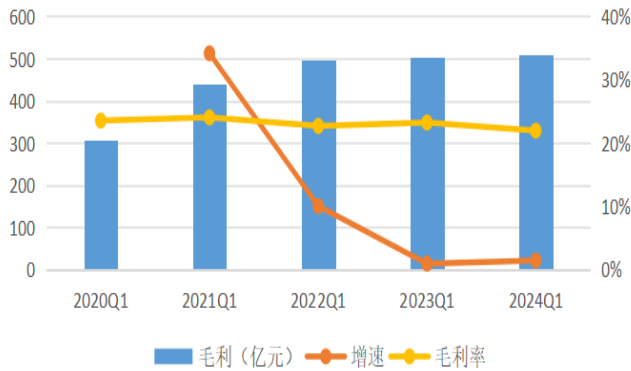


资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

行业增收不增利的背后，行业毛利率下滑问题影响较为明显。Q1 行业毛利额增长仅

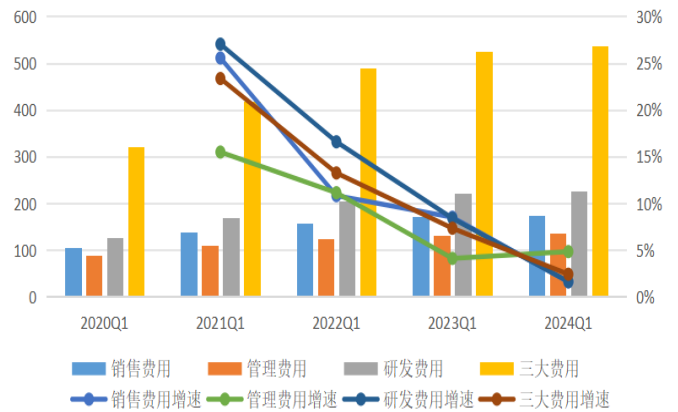
1.44%，毛利率从上年的 23.23% 下降到了 21.96%，下降了 1.28 PCT。同期，销售、管理、研发三大费用同比增长 2.40%，较上年下滑了 4.93 PCT，但仍然高于毛利增速水平。三大费用占营收比重从上年的 24.26% 下滑到了 23.15%。以上数据说明由于经济形势对下游需求的影响，行业目前收入增长质量不高，虽然费用增速已经放缓，但是厂商为了保持盈利的稳定和改善，不排除会采取进一步的控费措施。

图 11：中信计算机行业毛利及增速（亿元）



资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

图 12：中信计算机行业信用减值和资产减值（亿元）



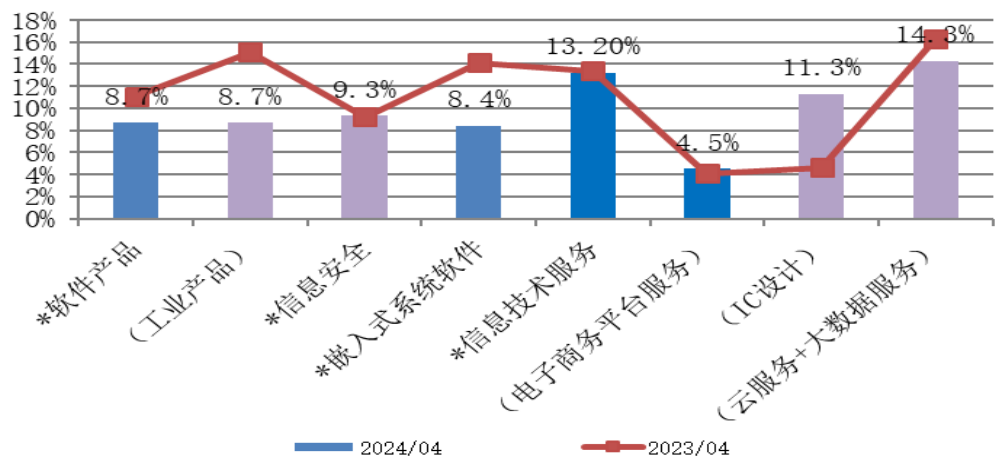
资料来源：Wind，上市公司公告，中原证券研究所（除工业富联）

1.3. 高景气赛道：IC 设计、云服务与大数据服务

从细分行业数据来看，2024 年 1-4 月软件行业细分子行业增速普遍低于上一年，其中值得重点关注的赛道为：

(1) 云服务与大数据服务：1-4 月累计收入同比增长 14.3%，领先其他细分子行业。由于 AI 大模型的快速发展，我们看到行业从 2023 年年初以来都是处于高景气状态。

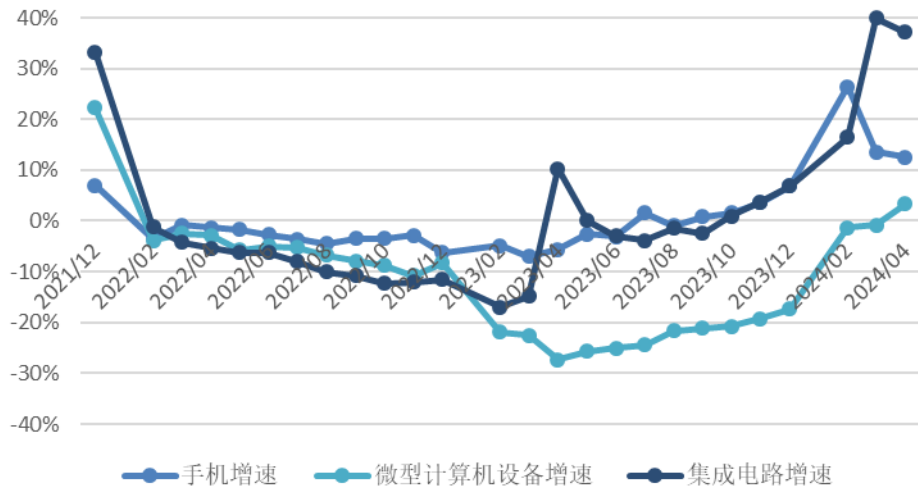
图 13：2024 年 1-4 月和 2023 年 1-4 月我国软件业务子行业增速对比



资料来源：工信部，中原证券研究所（蓝色柱体为一级子行业，紫色柱体为二级子行业）

(2) IC 设计：增速从上一年度的 4.6% 提升到了 11.3%，同比改善最为明显。由于电子产业的全面复苏，相关产业链的景气度也获得持续提升。

图 14：2021.1-12 至 2024.1-4 我国电子产品产量增速（累计同比）



资料来源：工信部，中原证券研究所

2. 人工智能加速发展，行业变革持续深入

2.1. 谷歌、OpenAI 竞争大模型第一梯队，推动大模型加速发展

OpenAI 在本轮大模型技术的演进中持续处于引领者地位。2023 年初，ChatGPT 火遍全网，让人们开始思考 AI 改变将给人类生活带来的改变；2023 年 3 月 GPT-4 的惊艳亮相，科技界开始全面拥抱 AI 大模型；2023 年 11 月 GPT-4 Turbo 的性能提升和价格下降，让人们看到了大模型应用落地发展路径；2024 年 2 月 Sora 出现后，大模型对真实世界物理逻辑关系的理解能力超出了人们的想象，推动了市场对具身智能产业的关注，也让人们看到了通用人工智能的曙光；2024 年 5 月，OpenAI 推出了 GPT-4o，具有强大的多模态交互能力，可以实现文本、音频和图像的任意组合输入和实时推理相应。

表 1：OpenAI 大模型产品

模型	时间	参数	上下文限制	训练数据截止时间	公开的程度	备注
GPT-1	2018.6	1.17 亿			开源（技术细节）	
GPT-2	2019.2	15 亿			开源（源代码）	
GPT-3	2020.5	1750 亿		2019.10	论文	
GPT-3.5	2022.3	-	4K/16K	2021.6		2022.11 ChatGPT 基于 GPT 3.5 推出
GPT-4	2023.3	1.8 万亿	8K/32K	2021.9	测试结果	2023.7 向所有开发者开放使用权限
GPT-4 Turbo	2023.11	-	128K	2023.4		速度是 GPT4 的 5 倍，价格降低 2/3
Sora	2024.2					首个文生视频模型
GPT-4o	2024.5					具有了强大的多模态交互能力 对用户免费,API 速度较 GPT-4 Turbo

资料来源：OpenAI，中原证券研究所

作为昔日 AI 领域的领军者，谷歌早在 2017 年就提出了 Transformer 模型，而在 OpenAI 的 ChatGPT 出现后，谷歌一直呈现追赶状态。直到 2024 年 5 月，谷歌 AI 全家桶的推出，意味着谷歌再次回到了头部厂商的队列中。

表 2：谷歌大模型产品

模型	时间	参数	上下文限制	公开的程度	备注
PaLM	2022.4.	5400 亿			Pathways 架构，语言大模型，预训练数据 7800 亿 token
ViT	2023.2.20	220 亿			计算机视觉模型，参数扩充到 220 亿
PaLM-E	2023.3.7	5620 亿			PaLM 和 ViT 的结合
PaLM 2	2023.5.11	3400 亿			语言大模型，拥有从小到大四种尺寸，分别为 Gecko、Otter、Bison 和 Unicorn，预训练数据 3.6 万亿 token
Gemini 1.0	2023.12.6				多模态大模型，推出 Ultra、Pro 和 Nano 三个参数版本
Gemini 1.0 Ultra			32K		
Gemini 1.0 Pro			32K		
Gemini 1.0 Nano					面向端侧
Gemini 1.5	2024.2.15		1M		基于 Transformer 和 MoE 架构
Gemini 1.5 Pro			1M		
Gemma	2024.2.21	20 亿/70 亿		开源	预训练数据集 2 万亿/6 万亿
Gemini 1.5 Flash	2024.5.15		1M		轻量版
Gemini 1.5 Pro 进阶版	2024.5.15		2M		
Gemma 2	2024.6		270 亿	开源	

资料来源：谷歌，中原证券研究所

除了 OpenAI 和谷歌以外，我们还看到 OpenAI 灵魂人物 Iliya 在离开 OpenAI 后创建了新公司 SSI，给未来大模型的头部厂商的竞争带来新的变量。

2.2. 开源大模型队伍加速壮大，对我国大模型发展起到了积极推动作用

2024 年，我们看到加入开源模型发布的厂商队伍迅速扩大，同时新模型发布的频率也明显加快。在无法成功超越 OpenAI 和谷歌的情况下，Meta 选择了开源 Llama 的路线来壮大自身队伍；从 OpenAI 出走的团队创立了 Anthropic，并发布了 Claude 模型；马斯克的 xAI 发布了 Grok-1 模型，同时有望进一步开源 Grok-1.5；中国厂商队伍中，阿里在 2 月开源了 Qwen1.5，在 6 月又将模型更新到了 Qwen2，模型整体能力突出。

除此以外，值得重点关注的包括：

(1) 谷歌在 2 月发布的 Gemma，有望在 6 月更新到 Gemma 2。谷歌加入开源路线，

可以视为对 Meta 等厂商的积极应战。

(2) 英伟达在 6 月发布了 Nemotron-4，参数量高达 3400 亿，构建了高质量合成数据生成的完整流程，使得模型训练不再完全依赖真实世界的的数据，成为了具有里程碑意义的事件。

表 3: 主要开源模型对比

模型	时间	公司	参数	上下文限制	预训练数据集	
Llama	2023.2.24	Meta	70 亿/ 130 亿/ 330 亿/ 650 亿	2K	1 万亿/1.4 万亿	
Llama 2	2023.7.19	Meta	70 亿/130 亿/700 亿	4K	2 万亿	
Mixtral 8x7B	2023.12.8	Mistral AI	450 亿	32K		MOE(8)架构
Qwen1.5	2024.2.6	阿里	5 亿/18 亿/40 亿/70 亿 /140 亿/720 亿	32K		
Gemma	2024.2.21	谷歌	20 亿/70 亿		2 万亿/6 万亿	
Claude 3 sonnet	2024.3.6	Anthropic		200K		本系列中平衡性能和速度的模型
Claude 3 Haiku	2024.3.13	Anthropic		200K		本系列中速度最快的模型
Grok-1	2024.3.18	xAI	3140 亿	8.19K		MOE(8)架构
DBRX	2024.3.28	Databricks	1320 亿	32K	12 万亿	MOE(16)架构，平均激活 360 亿个参数，没有公开数据细节
Mixtral 8x22B	2024.4.11	Mistral AI	1760 亿	65K		sMOE(8)架构
Claude 3 sonnet	2024.4.17	Anthropic		200K		
Llama 3	2024.4.18	Meta	80 亿/ 700 亿	8K	15 万亿	
Qwen 2	2024.6.7	阿里	4.9 亿/15.4 亿/70.7 亿 /574.1 亿/727.1 亿	32K-128K		
Nemotron-4	2024.6.14	英伟达	3400 亿	4K	9 万亿	构建了高质量合成数据生成的完整流程
Grok-1.5		xAI		128K		
Gemma 2	2024.6		270 亿			
Llama 3			4000 亿			

资料来源：新智元，凤凰网，澎湃，中原证券研究所

开源模型加速了大模型技术的普及，成为了可以抗衡闭源大模型的有效武器，也对我国大模型发展起到了积极的推动作用。同时，开源大模型的普及，也对云厂商的销售形成了有利推动。英伟达这一大参数开源模型的发布，有望进一步拉动大模型的算力需求，利好自身发展。

2.3. 英伟达改写行业运行逻辑，成为市场最大赢家

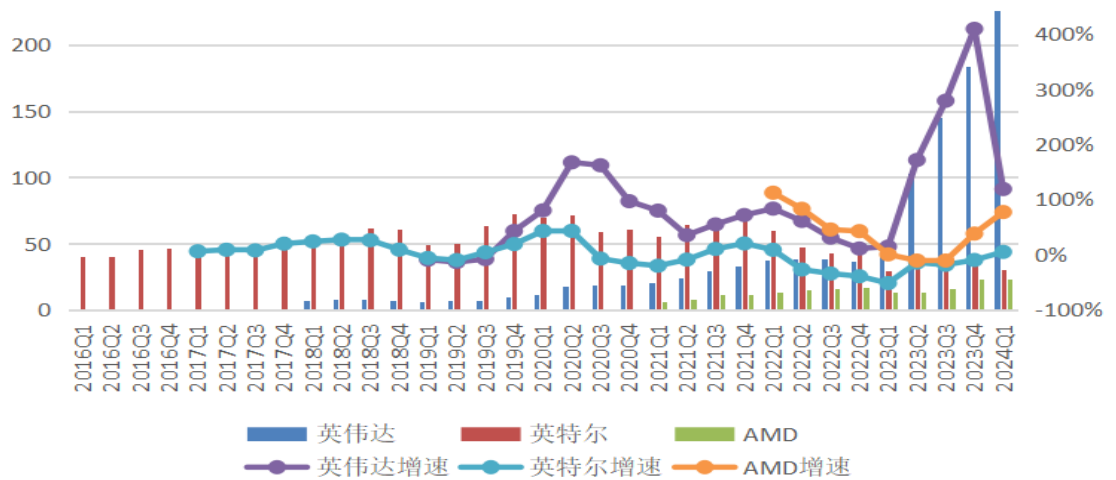
当前，英伟达依靠 AI 的发展，已经成为了全球市值最高的公司和市场最大的赢家，同时

我们也看到英伟达作为 AI 产业链上的核心企业，正在改写行业的发展轨迹，主要包括：

(1) GPU 替代 CPU，主导了算力的新需求。

随着人工智能的发展，以 8 个 GPU、4 个 GPU 为典型配置的 AI 服务器需求大幅增长，导致 AI 算力需求带动下，GPU 的需求量大幅增长，而以 CPU 为主导的传统服务器需求受到挤压。作为 GPU 的领导厂商，英伟达在数据中心业务上的收入从 2023Q1 的 42.84 亿美元增长到了 2024Q1 的 225.63 亿美元，在 2024Q1 仍然实现了同比 119% 和环比 23% 的增速。而作为 CPU 的典型厂商，2024Q1 英特尔和 AMD 的数据中心业务仅实现了 30.4 亿美元和 23.37 亿美元，这意味着 GPU 市场规模目前已经是 CPU 的数倍。

图 15：2016Q1-2024Q1 三大芯片厂商数据中心业务收入及增速（亿美元）



资料来源：上市公司财报，中原证券研究所（英伟达财报较正常季度延后 1 个月）

(2) 英伟达依靠 CUDA 的生态捆绑，实现一家独大。

AI 快速发展导致 GPU 紧缺，创造了巨大的市场红利，也让更多厂家加入到加速芯片的研发和生产的队伍中。这其中既有云计算厂商，也有像 AMD 这样的老牌芯片厂商。但时至今日，英伟达仍然占据了绝对的主导地位，也赚取了大部分 AI 市场的利润。

2023 年 6 月，AMD 发布了 MI300X，并在 12 月上市，性能赶超同期市售的 H100。但是我们却看到受制于英伟达在运算平台 CUDA 上的多年投入和 CUDA 建立的强大生态，MI300X 上市后对市场影响仍然相对有限，2024Q1 收入增长低于市场预期。同时，随着英伟达在 2024 年 3 月发布了 B200，性能超越 MI300X，重新夺回王者宝座。

表 4：英伟达、AMD 人工智能芯片参数对比

	A100 SXM	H100 SXM	H200 SXM	B200		MI300X	MI325X
厂商	英伟达	英伟达	英伟达	英伟达	英伟达	AMD	AMD
发布时间	2020.3	2022.3	2023.11	2024.3	2025	2023.6	2024.6
架构	Ampere	Hopper	Hopper	Blackwell	Blackwell Ultra	CDNA 3	CDNA 3
工艺	7nm	4nm	4nm	4nm		5nm	3nm

晶体管数量	540亿	800 亿	800 亿	2080 亿		1530 亿	
最大功率	400W	700W	700W	1000W		750W	
显存容量	80GB	80GB	141GB	192GB	800GB	192GB	288GB
显存带宽	2039GB/s	3.35TB/s	4.8TB/s	8TB/s		5.3TB/s	6TB/s
显存颗粒	HBM2e	HBM3	HBM3e	HBM3e		HBM3	HBM3e
传输速率	600GB/s	900GB/s	900GB/s	1800GB/s		896GB/s	
FP64	9.7 TFLOPS	34 TFLOPS	34 TFLOPS				
FP64 Tensor Core	19.5 TFLOPS	67 TFLOPS	67 TFLOPS				
FP32	19.5 TFLOPS	67 TFLOPS	67 TFLOPS				
TF32	312TFIOPS					653.7 TFLOPS	
TF32 Tensor Core		989 TFLOPS	989 TFLOPS				
BFLOAT16		989.4 TFLOPS				1307.4 TFLOPS	
BFLOAT16 Tensor Core	624 TFLOPS	1979 TFLOPS	1979 TFLOPS				
FP16		989.4 TFLOPS				1307.4 TFLOPS	1.3 PFLOPS
FP16 Tensor Core	624 TFLOPS	1979 TFLOPS	1979 TFLOPS				
FP8		1978.9 TFLOPS				2614.9 TFLOPS	2.6 PFLOPS
FP8 Tensor Core		3958 TFLOPS	3958 TFLOPS	10 PFLOPS			
INT8		1978.92 TOPS				2614.92 TOPS	
INT8 Tensor Core	1248 TOPS	3958 TOPS	3958 TOPS				
FP6	-	-	-	10PFLOPS		-	
FP4	-	-	-	20PFLOPS		-	

资料来源：英伟达，AMD，中原证券研究所

(3) 英伟达产品供应从芯片到板卡到整体算力解决方案，不断解构原有的产业体系，话语权持续增大。

2024年3月18日，英伟达推出了新架构 Blackwell 及该架构芯片 B200，以及一系列新产品，其中包括：

a、芯片：B200 由 2 片 4nm 芯片并排摆放，通过超高速片上互联（10TB/s），是英伟达首个采用 MCM（多芯片封装）设计的 GPU。

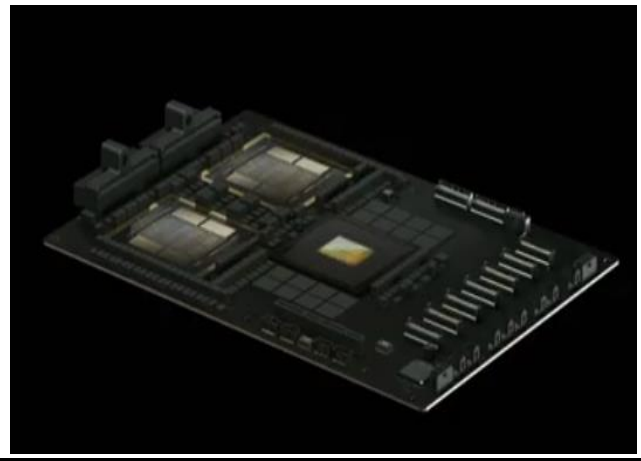
b、超级芯片：由 2 个 B200 GPU 和 1 个 Grace CPU 组成一个 GB200 超级芯片。

图 16: 英伟达 B200 芯片



资料来源: 英伟达, 中原证券研究所

图 17: 英伟达 GB200 超级芯片



资料来源: 英伟达, 中原证券研究所

c、服务器: 2 个 GB 200 超级芯片组成一个 Blackwell 计算节点。

d、整机柜服务器: 18 个计算节点加上 NVLink Switch 的支持, 组成 36 CPU+ 72 GPU 的“虚拟 GPU”——DGX GB200 NVL72。

图 18: 一个 Blackwell 计算节点



资料来源: 英伟达, 中原证券研究所

图 19: 英伟达 DGX GB200 NVL72



资料来源: 英伟达, 中原证券研究所

e、AI 超级计算机: 英伟达为训练和推理万亿参数生成式 AI 模型而设计的数据中心解决方案 DGX SuperPOD, 可以通过 NVIDIA Quantum InfiniBand 连接多个机架, 实现数万个 GB200 超级芯片的扩展 (8 个或以上 DGX GB200 构成)。

图 20：英伟达 DGX SuperPOD



资料来源：英伟达，中原证券研究所

我们看到此前数据中心的主流方案是 x86 CPU 与 8 张 GPU 的组合，GPU 侧做了复杂的网络和 NVLink 互联。而英伟达借助互联带宽上的优势，推出自家 CPU+GPU 的私有化方案，并且在 DGX 方案中也剥离了 x86 CPU。这意味着计算机形态又一次重大变迁，经历过数十年从昂贵的 IBM 大型机走向廉价的 x86 处理器、Windows 操作系统为主导的计算机大规模普及，又在大模型推动下，重新回到大型机时代昂贵而奢侈的私有化解决方案。这一方案或将进一步提高 AI 产业的布局门槛，降低下游的投入产出比。

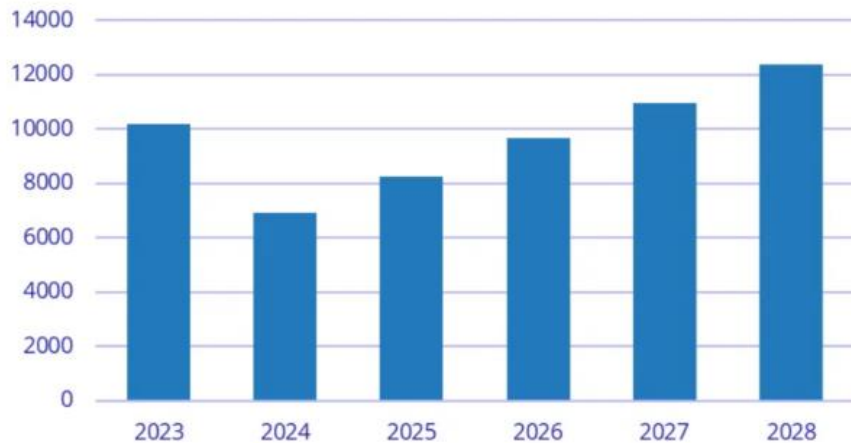
（4）参与开源大模型发布，加速算力需求增长。

2024 年 6 月 14 日宣布开源模型 Nemotron-4 340B，这一模型不但性能超越 Mixtral 8x22B、Claude sonnet、Llama3 70B、Qwen 2 等市场主流开源模型，同时在训练数据生成方式方面实现了重要突破。Nemotron-4 340B 构建了一个高质量数据合成的完整流程，他包括基础模型 Base、指令模型 Instruct 和奖励模型 Reward 的 3 个模型中，指令模型的训练是在 98% 的合成数据上完成的。至此我们也可以看到，英伟达的业务领域也进一步延伸到了 AI 模型算法层，通过推出大参数的开源模型，助推智能算力需求的增长，倒逼行业购买更贵的机器。

2.4. 美国禁令趋势下，国产芯片阵营获得难得的发展机遇

在英伟达成为全球 AI 产业焦点的过程中，美国不断加大对中国芯片出口控制，英伟达针对中国市场推出的 A800、H800 也从 2023 年 10 月开始受到了新的禁令的限制，对中国 AI 芯片的供给造成了较大的负面影响。为此，IDC 甚至预测，到 2028 年中国加速服务器市场规模将达到 124 亿美元（65.7 万台），2023-2028 年复合增速仅 5.7%，而对于 2024 年，IDC 预测加速计算服务器市场规模下降。

图 21: 2023 年-2028 年中国加速计算服务器市场预测 (百万元)



资料来源: IDC, 中原证券研究所

但是与此同时, 我们也看到随着 2023 年 8 月华为 Mate60 系列的正式发售, 华为芯片的生产问题获得了突破。2023 年 10 月, 科大讯飞表示昇腾 910B 能力已经基本做到可对标英伟达 A100。此外, 华为拥有了对标 CUDA 的 CANN、AI 计算框架昇思、盘古大模型, 形成了较英伟达更完整的产业布局。

2024 年以来, 英伟达在加快芯片发布频率的同时, 也采取了一系列的反制措施, 对竞争对手的发展进行阻碍, 包括 (1) 2024 年 3 月, 英伟达在 CUDA 11.6 的用户许可中明确表示, 禁止其他硬件平台上通过翻译层运行 CUDA; (2) 2024 年 5 月, 据路透社报道, 英伟达大幅削减供应中国市场的 H20 人工智能芯片价格。

英伟达的缺位, 给以华为为代表的国产 AI 芯片厂商提供了难得的发展机遇, 让国内市场成为了国产芯片发展的小气候。根据英伟达的数据, 中国大陆地区的收入约占比在 20% 以上, 有足够的市场体量培养出与英伟达相抗衡的芯片企业。随着互联网厂商开始积极试用和采购国产芯片, 国产生态环境有望逐步建立, 最终帮助国产芯片和下游厂商突破 CUDA 的封锁。

2.5. OpenAI 暂停中国 API 接口, 国内大模型将获得更好应用

2024 年 6 月 25 日, OpenAI 突然宣布停止向中国提供 API 服务。而在此前, OpenAI 虽然没有对中国用户开放 API 服务, 但是国内用户往往选择代理之类的方式绕过区域限制。但是随着中国用户开始收到 OpenAI 的邮件通知, 这意味着 OpenAI 将采取更加严格的检测来限定国内客户的调用。

OpenAI 的这一举动将利好于国内大模型厂商。6 月 25 日, 阿里云百炼宣布将为 OpenAI API 用户提供替代方案, 并为中国开发者提供 220 万免费 tokens 和专属迁移服务。

从模型价格来看, 国产厂商价格优势明显。通义千问 GPT4 级主力模型 Qwen-plus 在阿里云百炼上的调用价格为 0.004 元/千 tokens, 仅为 GPT-4 的 50 分之一。其他国产厂商也都

在 5 月以来大模型降价潮中，纷纷推出了降价或免费方案，为大模型应用场景落地起到了积极的作用。

从模型性能来看，国产模型较 OpenAI 的差距也在逐步缩小。根据斯坦福大模型榜单，在前十名大模型产品中，有两家中国大模型厂商，其中来自阿里巴巴的 Qwen2 Instruct (72B) 与 GPT-4 (0613) 取得了相同的分数。

图 22：斯坦福大模型榜单 HELM MMLU (2024 年 6 月 11 日)

Model	MMLU All Subjects - EM
Claude 3 Opus (20240229)	0.846
GPT-4o (2024-05-13)	0.842
Gemini 1.5 Pro (001)	0.827
GPT-4 (0613)	0.824
Qwen2 Instruct (72B)	0.824
GPT-4 Turbo (2024-04-09)	0.813
Gemini 1.5 Pro (0409 preview)	0.81
GPT-4 Turbo (1106 preview)	0.796
Llama 3 (70B)	0.793
Yi Large (Preview)	0.793

资料来源：斯坦福大学基础模型研究中心，IT之家，中原证券研究所

3. 2024 年算力需求增长趋势明显

3.1. 我国近期算力规划及总体完成情况

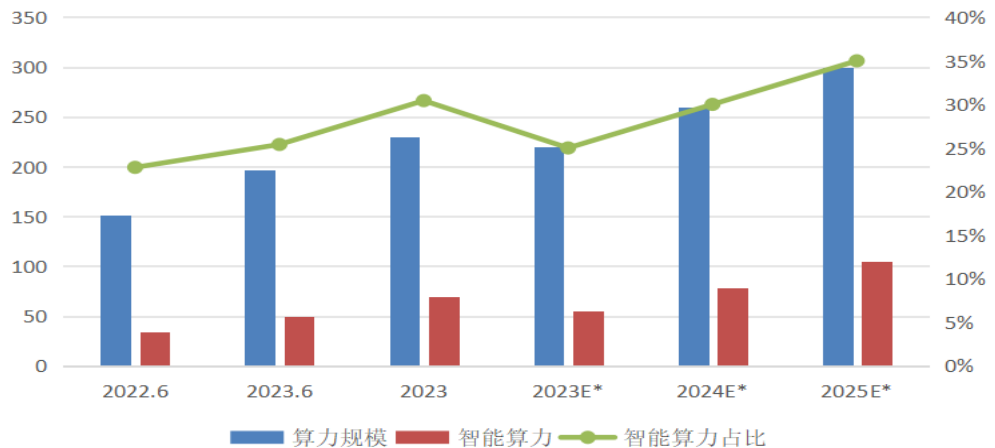
我国当前正在推动的全国一体化算力网，包括了通用算力、智能算力、超级算力 3 个类型的算力一体化建设。根据《中国综合算力指数 (2023 年)》，截止 2023 年 6 月底，我国算力总规模达到 197EFLOPS，比上年同期增长 30%，其中智能算力规模占整体算力规模的比例提高到 25.4%，智能算力规模同比增长 45%。

根据 2023 年 9 月工信部等六部门联合印发的《算力基础设施高质量发展行动计划》，计划在 2025 年算力规模超过 300EFLOPS (EFLOPS 是指每秒百亿亿次浮点运算次数)，智能

算力占比达到 35%，东西部算力平衡协调发展。

智能算力加速发展趋势明显，有望提前完成 2025 年全国的发展目标。根据工信部数据，2023 年我国在用数据中心机架总规模超过 810 万标准机架，算力总规模达到 230 EFLOPS，智能算力 70EFLOPS，智算比例达到 30%，智能算力增速达到 70%，对比《算力基础设施高质量发展行动计划》中的参考指标和 2023 年 6 月的行业数据来看，整体算力建设发展超预期，其中智能算力建设更是呈现出了加速趋势。

图 23：2022.6-2025E 我国算力规模及结构（EFLOPS）



资料来源：《算力基础设施高质量发展行动计划》，《中国综合算力指数（2023 年）》，工信部，中原证券研究所（*《算力基础设施高质量发展行动计划》中的算力基础设施高质量发展指标）

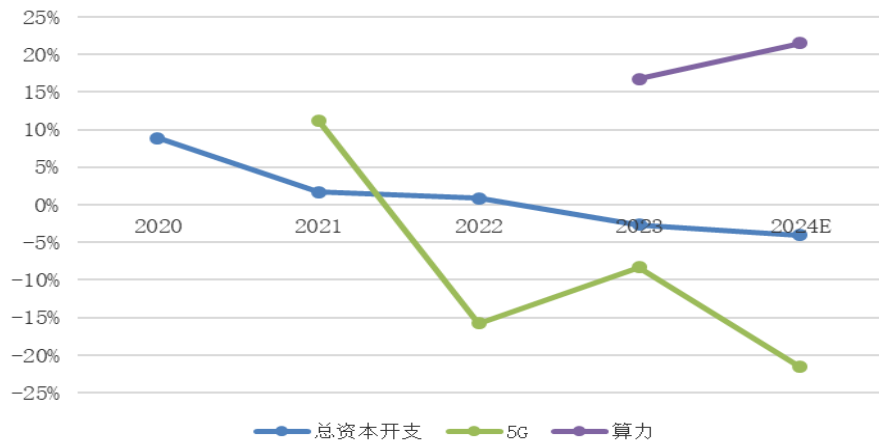
根据 2024 年 3 月的政府工作报告，我国计划“适度超前建设数字基础设施，加快形成全国一体化算力体系，培育算力产业生态”，全国的算力建设有望进一步加快。

3.2. 2024 年算力建设加速迹象明显

3.2.1. 2024 年运营商加快算力布局

作为国内算力建设的重要参与方，我们看到三大运营商 2024 年的资本开支计划都呈现出下滑态势，但是在积极加大算力建设的投入。随着 5G 建设渡过高峰期，移动 2024 年计划算力投入 1730 亿元，下滑 4%，但是在算力网络投资方面计划投入 475 亿元，同比增长 21.5%。电信计划 960 亿元的资本开支中 180 亿元投入到云与算力，计划智算算力从 2023 年的 11 EFLOPS 增长到 21 EFLOPS。联通也提出算网数智投资适度超前，加快布局。

图 24：2020-2024 年中国移动资本开支增速及 5G、算力的开支增速

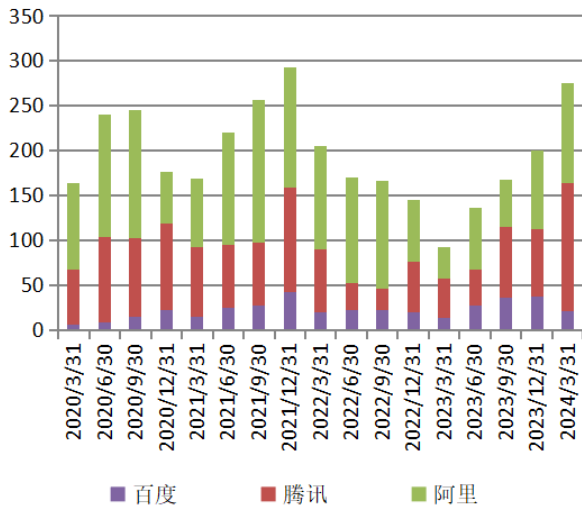


资料来源：移动云公众号，中原证券研究所

3.2.2. 国内外科技大商加大资本投入

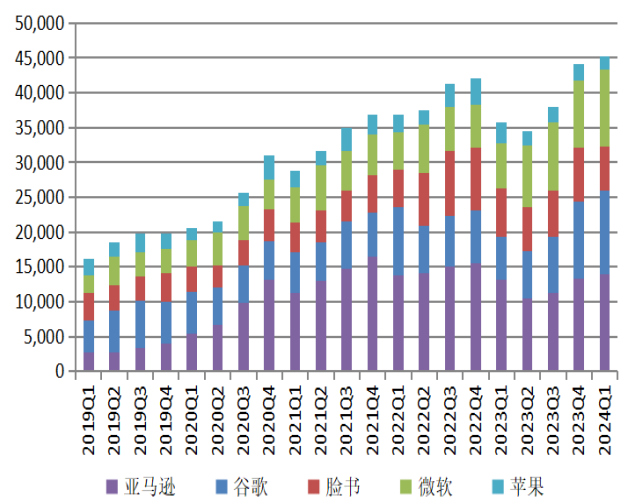
由于人工智能带来的科技变革，国内外科技巨头也都在加大资本投入，积极进行算力建设。国内来看，BAT（百度、腾讯、阿里）三家的资本开支在 2023Q1 见底后逐季提升，2024Q1 已经接近 2021Q4 的峰值。国外来看，亚马逊、谷歌、苹果、微软、脸书为代表的 5 大科技厂商的总资本开支也呈现逐季提升趋势，已经在 2024Q1 创下历史新高。

图 25：2020Q1-2024Q1 BAT 的资本开支（亿元）



资料来源：上市公司财报，中原证券研究所

图 26：19Q1-24Q1 5 大科技厂商资本投入（百万美元）



资料来源：上市公司财报，Wind，中原证券研究所

3.2.3. 地方算力建设相关政策集中发布

在 2023 年《算力基础设施高质量发展行动计划》、《关于深入实施“东数西算”工程加快构建全国一体化算力网的实施意见》聚焦算力基础设施，2024 年政府工作报告提出“适度超前建设数字基础设施”的基础上，2024 年各地有关算力基础设施建设的规划政策也在接连出

台。

表 5: 近期省级算力基础设施相关规划

时间	区域	政策	规划
2023.12.15	安徽	《安徽省数字基础设施建设发展三年行动方案(2023-2025年)》	预期 2025 年数据中心机架规模不少于 45 万架, 数据中心平均上架率不少于 65%, 算力规模 14.6 EFLOPS, 智能算力规模超过 12 EFLOPS。 同时要求 2025 年新建大型及以上数据中心电能利用率低于 1.3。
2024.1.25	山西	《山西省算力基础设施高质量发展实施方案》	到 2025 年山西省算力规模超过 9EFLOPS, 智能算力占比达到 35%, 建成 4 个以上的智能计算中心; 存储总量超过 36EB, 先进存储容量占比达到 30%以上, 重点行业核心数据、重要数据灾备覆盖率达到 100%; 重点应用场所光传送网(OTN)覆盖率达到 80%, 算力全光调度 OXC 节点占比达到 80%, 骨干网、城域网全面支持 IPv6、SRv6 等创新技术, 使用占比达到 40%, 实现省内城市内 1 毫秒、城市间 3 毫秒、到京津冀枢纽节点 5 毫秒、到其他枢纽节点 20 毫秒的运力网络。
2024.2.2	贵州	《贵州省算力基础设施高质量发展行动计划(2024-2025年)》	到 2025 年, 计算供给均衡合理, 总算力规模提升到 80 EFLOPS, 超总算力与智总算力占比达到 35%。存储保障安全可靠, 存储总量超过 60EB, 全省先进存储容量占比达到 30%以上, 重点行业的核心数据、重要数据的灾备存储覆盖率达到 100%。算力设施绿色集约, 加强对存量数据中心的改造提升, 推动数据中心采用高效节能冷却系统, 全省新建大型以上数据中心 PUE 值低于 1.2。
2024.3.22	上海	《上海市智能算力基础设施高质量发展“算力浦江”智算行动实施方案(2024-2025年)》	到 2025 年, 本市智能算力规模超过 30EFlops, 占比达到总算力的 50%以上。算力网络节点间单向网络时延控制在 1 毫秒以内。智算中心内先进存储容量占比达到 50%以上。本市新建智算中心国产算力芯片使用占比超过 50%, 国产存储使用占比超过 50%, 服务具有国际影响力的通用及垂直行业大模型设计应用企业超过 10 家。本市新建智算中心 PUE 值达到 1.25 以下, 存量改造智算中心 PUE 值达到 1.4 以下。智算中心内绿色能源使用占比超过 20%, 液冷机柜数量占比超过 50%。
2024.3.22	广东	《广东省算力基础设施高质量发展行动暨“粤算”行动计划(2024-2025年)》	到 2025 年, 在算力方面, 算力规模达到 38EFLOPS, 智能算力占比达到 50%。建成智能计算中心 10 个。在存储力方面, 存储总量超过 260EB, 先进存储容量占比达到 30%以上, 重点行业核心数据、重要数据灾备覆盖率达到 100%。力争到 2025 年底, 新增国产化算力占比达到 70%。
2024.4.23	江苏	《江苏省算力基础设施发展专项规划》	算力基础设施方面, 到 2030 年, 全省数据中心机架规模达 120 万标准机架, 全省在用总算力超过 50EFLOPS, 智能算力占比进一步提升, 占比超过 45%, 智算中心数量突破 20 个, 算力算效水平显著提高; 存储总量超过 500EB, 进一步提升先进存储容量占比至 40%以上。到 2030 年, 全国一体化算力网络长三角国家枢纽节点数据中心集群平均 PUE(电能利用效率)低于 1.25, 太阳能、风能、冷能等可再生能源利用率达到 25%以上。基础软硬件和云操作系统实现自主可控, 新建算力中心国产算力芯片使用占比达 70%以上、国产存储使用占比达 90%以上。
2024.4.24	北京	《北京市算力基础设施建设实施方案(2024—2027年)》	改变智算建设“小、散”局面, 集中建设一批智算单一集群, 到 2025 年, 本市智算供给规模达到 45EFLOPS。到 2027 年, 实现智算基础设施软硬件产品全栈自主可控, 整体性能达到国内领先水平, 具备 100%自主可控智算中心建设能力。本市新建和改扩建智算中心 PUE 值一般不超过 1.25, 年能耗超过 3 万吨标煤的大规模先进智算中心 PUE 值一般不超过 1.15。推进本市存量数据中心升级改造, 到规划期末所有存量数据中心 PUE 值均不高于 1.35。
2024.5.7	甘肃	《甘肃算力基础设施高质量发展三年行动计划(2024—2026年)》	到 2026 年, 算力方面, 全省算力规模超过 30EFlops; 运载力方面, 实现庆阳数据中心集群到兰州时延达到 3ms 以内, 到成渝枢纽时延达到 8ms 以内, 到京津冀枢纽时延达到 10ms 以内, 到长三角枢纽时延达到 15ms 以内, 到粤港澳大湾区枢纽时延达到 15ms 以内; 存储力方面, 存储总量超过 65EB, 先进存储容量占比达到 32%以上, 重点行业核心数据、重要数据灾备覆盖率达到 100%。

2024.5.27	山东	《山东省算力基础设施高质量发展行动方案》	到 2025 年，全省算力基础设施发展总体水平进入全国前列，打造绿色低碳高质量发展先行区算力底座：全省数据中心在用标准机架总数达到 45 万架，总算力达到 12.5 EFLOPS，智能算力占比达到 35%；存力规模达到 65EB，先进存储占比达到 35% 以上。
-----------	----	----------------------	---

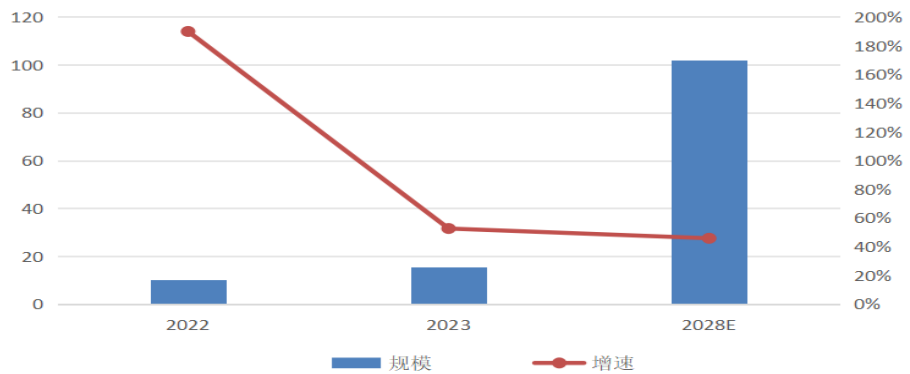
资料来源：地方政府网站，中原证券研究所

从已经出台的省级的算力规划中，我们可以看到上海市、广东省、安徽省等地对于智算比例的要求都明显高于全国总体目标水平。

3.3. 液冷服务器将迎来快速发展机遇

液冷服务器将迎来高速发展期。根据 IDC 数据，2023 年我国液冷服务器市场规模 15.5 亿美元，同比增长 52.6%，预计到 2028 年将增长到 102 亿美元，复合增长率 45.8%，这一规模已经快与 IDC 预测同期 124 亿美元的加速服务器市场规模相当。

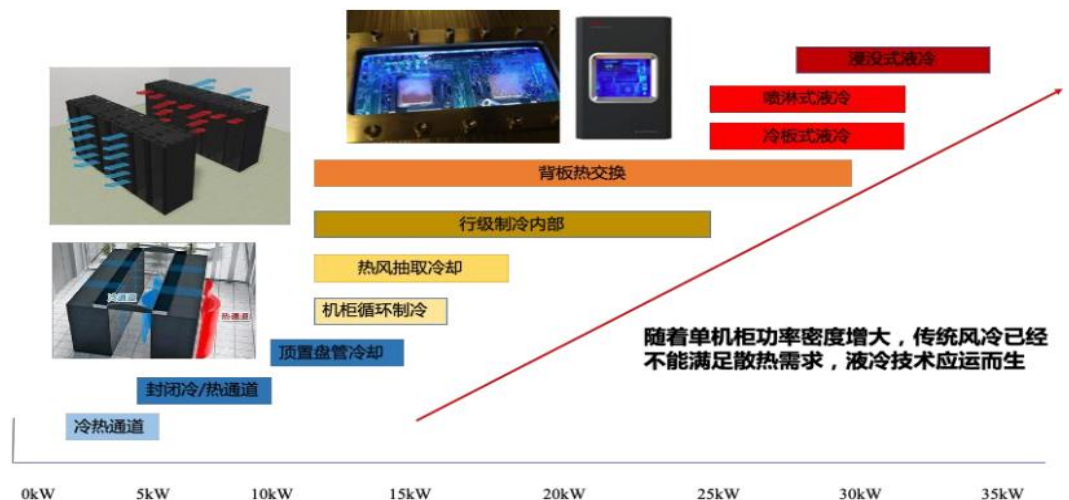
图 27：2022 年中国液冷服务器规模及增速（亿美元）



资料来源：IDC，中原证券研究所

随着生成式人工智能的发展，智算需求爆发式增长，同时由于智算中心高密度属性，风冷散热难以满足耗能指标要求，AI 计算、液冷、整机柜的结合将成为未来的发展趋势，液冷行业也迎来了刚需放量的元年。

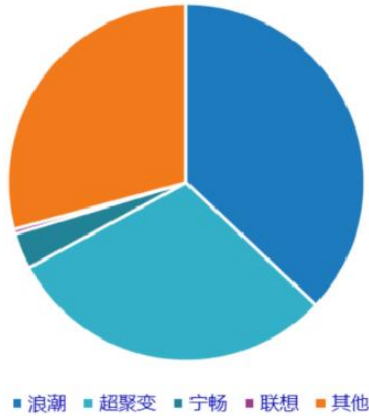
图 28：单机柜密度和冷却方式



资料来源：赛迪，中原证券研究所

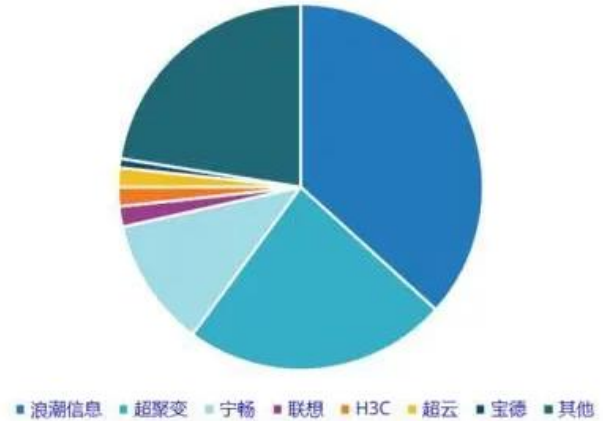
从液冷服务器的竞争格局来看，目前浪潮、超聚变、宁畅占据了70%以上的市场份额，居于绝对的主导地位，同时也有更多的厂商加入到液冷服务器的竞争中来。

图 29：2022 年我国液冷服务器市场厂商份额



资料来源：IDC，中原证券研究所

图 30：2023 年我国液冷服务器市场厂商份额



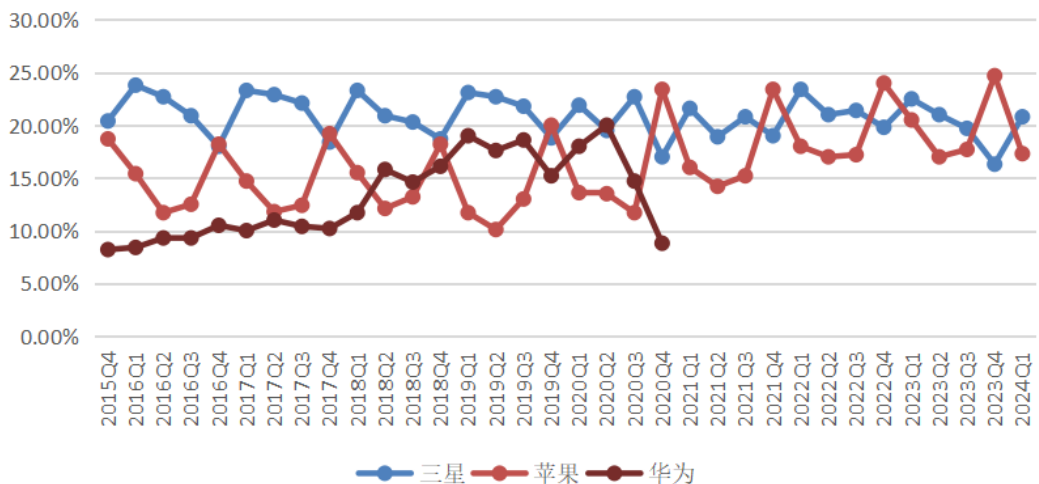
资料来源：IDC，中原证券研究所

4. 国产化推行持续深入，纯血鸿蒙商用值得期待

4.1. 禁令加剧，华为、AI、芯片是核心聚焦点

在美国禁令的影响下，华为曾面临了严重的芯片生产问题。在 2021-2022 年期间，华为的旗舰手机发布频率从一年 2 次，减少到 1 年一次；手机芯片从自家的麒麟芯片变为了高通的产品；基带芯片也从 5G 退回到了 4G 产品。与此同时，华为从 2020Q2 全球手机市场第一，到 2021Q1 直接跌出了前五大手机厂商行列。华为的消费者业务板块原本是公司规模最大、增速最高的业务，在禁令出台后，2021 年直接下滑了 49.6%，拖累整体收入下滑 28.6%，2022 年消费者业务继续下滑了 11.9%。

图 31：2015Q4-2024Q1 全球智能手机市场厂商市场份额



资料来源：IDC，中原证券研究所

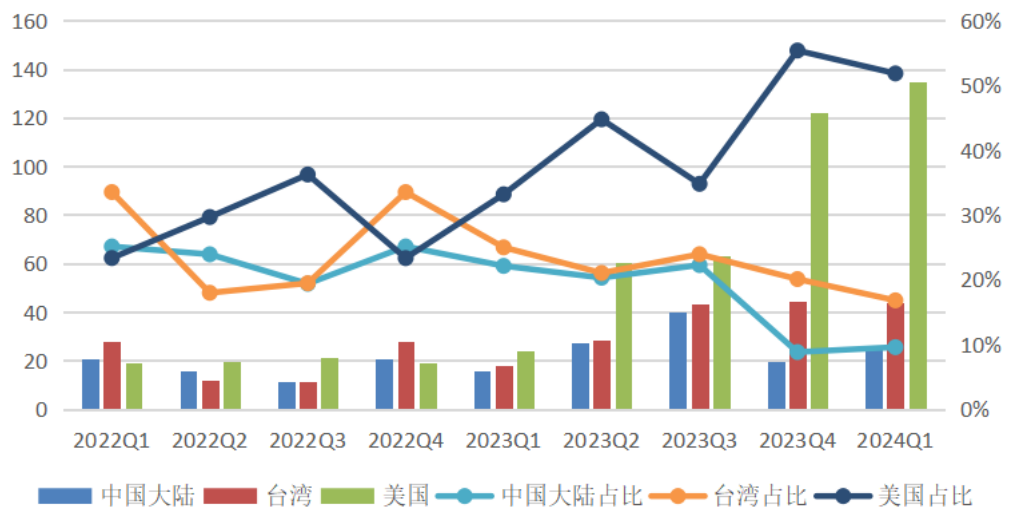
美国对 AI 芯片相关的限制也在持续扩大。2022 年 8 月，美国禁止了英伟达 A100 和即将推出的 H100 向中国和俄罗斯的所有出口实施一项新的许可要求。

2022 年 10 月 7 日，美国商务部工业和安全局（BIS）正式宣布了其对中国超算和半导体制造物项实施的出口管制新规，全面加大了对超算和芯片制造领域的限制措施，对 28 家实体清单公司的限制进一步扩大，同时还有 31 家实体被其加入未经核实名单（UVL）。本次新规对我国务高端芯片领域供给产生了全面而深远的影响，使得产业界对维持正常市场供应的期望最终破灭。

为了避免美国的限制，英伟达分别在 2022 年 11 月和 2023 年 3 月发布了 A100、H100 的中国特供版本 A800 和 H800。而 2023 年 10 月 17 日，美国商务部工业和安全局（BIS）发布了针对芯片的出口禁令新规，对于中国半导体的制裁进一步升级。受管制的包括但不限于英伟达 A100、H100、A800、H800、L40、L40S、RTX 4090 以及集成这些高性能计算卡的 DGX/HGX 系统。

根据英伟达财报数据，随着最新禁令的出台，2024Q1 英伟达来自美国的收入占比进一步提升到了 52%，而与此同时其来自中国大陆的收入从 2023Q3 的 40 亿美元降低到了 24.91 亿美元。

图 32：2022Q1-2024Q1 英伟达单季度收入的区域分布（亿美元）



资料来源：公司公告，Wind，中原证券研究所（英伟达财报较正常季度延后 1 个月）

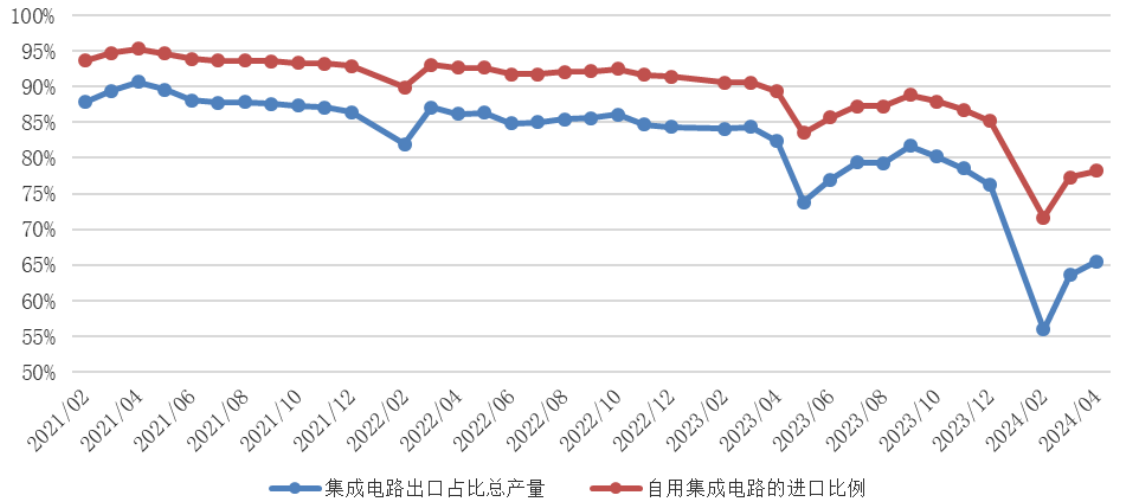
4.2. 国产化推进持续深入

4.2.1. 芯片环节国产化比例持续提升

在应对禁令的同时，国产芯片也在快速成长，带动我国自主可控方面能力的持续提升。

根据工信部和海关总署的数据进行测算，我们看到 2021 年以来，我国自用的集成电路中，来自进口的比例呈现持续下滑的趋势，已经从 2021 年前 2 个月的 94% 下滑到了 2024 年前 4 个月份的 78%，同时我们看到 2024 年这一指标下滑的趋势是非常明显的。

图 33：2021-2024 年我国集成电路对进口依赖度和出口占比



资料来源：工信部，海关总署，中原证券研究所

4.2.2. 运营商采购国产化率呈上升趋势

运营商采购侧的采购来看，国产服务器的采购比例也呈现逐年提升趋势。2020 年 5 月电信服务器集采中，国产芯片服务器占比为 20%，到了 2021 年 11 月的电信集采的国产化率进一步提升至 27%。而与此同时，移动于 2021 年 10 月启动的集采和 2022 年 3 月的补充采购合并来看，国产化占比达到了 44%，而后移动在 2021 年 4 月到 11 月启动的第二批“2021—2022 年 PC 服务器集中采购”中，国产化率小幅提升到了 47%。2023 年 10 月，电信启动了 AI 服务器采购，采购金额超过 80 亿元，其国产化比例也达到了 47%。与此同时移动在 2023 年 9 月至 2024 年 1 月进行的新型智算中心集采中（试验网），国产化比例达到了 70%（预估值），在 2024 年 4 月进行的新型智算中心集采中，服务器国产化比例已经达到了 100%，同时采购金额也超过 190 亿元，再次刷新了行业内 AI 服务器集采的记录。而在 2024 年 2 月到 3 月启动的 2024 年移动 PC 服务器集采中，我们预估国产化服务器的集采占比也到达了 60% 以上。

表 6：中国电信及中国移动服务器集采情况及国产化比例

招标时间	项目	总服务器数量	国产服务器数量*	国产化率*	国产化标包占比*
2024.4	中国移动 2024 年至 2025 年新型智算中心集采项目	7994	7994	100%	1/1
2024.2-3	中国移动 2024 年 PC 服务器产品集中采购项目	259122	155517	60%	11/22
2023.9-2024.1	中国移动 2023 年至 2024 年新型智算中心（试验网）采购	2454	1722	70%	4/8

2023.10	中国电信 AI 算力服务器（2023-2024 年）集采	4175	1977	47%	2/4
2022.4-11	中国移动 2021—2022 年 PC 服务器集中采购第二批次（标包 1-9+补充）	92571	43958	47%	8/10
2021.10-2022.3	中国移动 2021-2022 年 PC 服务器集中采购（第 1 批次+补充采购）	204696	90055	44%	13/23
2021.11	中国电信 2021-2022 年服务器集中采购项目	200000	53401	27%	1/7
2020.5	中国电信服务器（2020 年）集中采购项目	56314	11185	20%	1/8

资料来源：中国电信，中国移动，中原证券研究所（*估计比例）

4.2.3. 地方算力国产化比例要求成为发展指引

3 月以后出台的省级算力规划中，大都包括了较为严格的国产化比例的要求，具体包括：

（1）上海：提出到 2025 年，“新建智算中心国产算力芯片使用占比超过 50%，国产存储使用占比超过 50%”。

（2）广东：提出“力争到 2025 年底，新增国产化算力占比达到 70%”。

（3）江苏：提出到 2030 年，“基础软硬件和云操作系统实现自主可控，新建算力中心国产算力芯片使用占比达 70% 以上、国产存储使用占比达 90% 以上”。

（4）北京：提出“到 2027 年，实现智算基础设施软硬件产品全栈自主可控，整体性能达到国内领先水平，具备 100% 自主可控智算中心建设能力”。

以上地区作为全国经济最发达的区域，在算力国产化的推动方面也有望实现引领。其中在国产化比例方面，我们看到其要求的比例都在 50% 以上，北京甚至要求“具备 100% 自主可控智算中心建设能力”，这意味着提供国产算力的厂商能力也将持续提升，并有望借助国内市场成长为具备国际竞争力的企业。

4.2.4. 工业领域的更新换代，有望加速工业领域软件的国产化进程

2024 年 4 月 9 日，工信部联合发改委、财政部、央行、税务局、市监局、金融监管总局共同发布了《推动工业领域设备更新实施方案》（以下简称《方案》），来推动工业领域大规模的设备更新。

《方案》提出“到 2027 年，工业领域设备投资规模较 2023 年增长 25% 以上，规模以上工业企业数字化研发设计工具普及率、关键工序数控化率分别超过 90%、75%，工业大省大市和重点园区规上工业企业数字化改造全覆盖，重点行业能效基准水平以下产能基本退出、主要用能设备能效基本达到节能水平，本质安全水平明显提升，创新产品加快推广应用，先进产能比重持续提高。”

《方案》还强调了“坚持软硬件一体化更新。主动适应和引领新一轮科技革命和产业变

革，积极推进新一代信息技术赋能新型工业化，在推动硬件设备更新的同时，注重软件系统迭代升级和创新应用”，利好工业软件及工业领域配套基础软件的替换，也有望推动工业领域软件的国产化进程。

《方案》出台后，上海、广东、浙江等 20 余个省（自治区、直辖市）陆续出台了配套政策和任务清单，明确了财税金融方面的配套支持政策，成为了各地工业设备更新的主要发力点。

表 7：近期省级算力基础设施相关规划

时间	区域	政策	规划
2024.4.2	浙江	《浙江省推动大规模设备更新和消费品以旧换新若干举措》	到 2027 年，全省工业、能源、建筑、交通、农业等领域设备投资规模较 2023 年增长 30%以上，重点行业主要用能设备能效基本达到先进水平， 规模以上工业企业数字化研发设计工具普及率、关键工序数控化率分别超过 93%、75%。 完成工业企业数字化改造。 支持企业在设计、生产、仓储和运营等环节的数字化改造，到 2027 年规上工业企业数字化改造实现全覆盖。支持企业对装备、生产线、车间等进行智能化升级，到 2027 年新增未来工厂 80 家、数字化车间（智能工厂）600 个（家）。 严格落实设备更新和技术改造贷款贴息政策。鼓励国有企业加快设备更新。 聚焦各领域重大技术装备研发和生产的关键共性技术及科学问题， 通过“揭榜挂帅”“赛马”等机制开展重大项目攻关，到 2027 年组织实施重大科技项目 200 个以上，取得重大国产化替代成果 100 项。提升省实验室、省技术创新中心、省制造业创新中心等高能级创新平台建设质效，推进概念验证中心、中试基地建设。深入实施制造业首台（套）提升工程，到 2027 年新培育认定首台（套）产品 1000 项以上，通过保险补偿、应用奖励等方式支持首台（套）产品的推广应用。
2024.4.13	广东	《广东省人民政府关于印发广东省推动大规模设备更新和消费品以旧换新实施方案的通知》	到 2027 年，工业、能源、农业、建筑、交通、教育、文旅、医疗等领域设备投资规模较 2023 年增长 25%以上；重点行业主要用能设备能效基本达到节能水平，环保绩效达到 A 级水平的产能比例大幅提升， 规模以上工业企业数字化研发设计工具普及率、关键工序数控化率分别超过 90%、75% 加强智能制造软硬件产品设备推广，加快工业互联网建设和普及应用。
2024.4.19	北京	《进一步加强重点领域设备购置与更新改造贷款贴息的实施方案》	固定资产投资对符合条件的重点领域设备更新改造项目给予 2.5 个百分点的贴息，期限 2 年。2024 年度支持科技创新、战略性新兴产业、未来产业、先进制造业和现代服务业“两业”融合发展、新型基础设施、社会投资公共服务、文化旅游、市政基础设施、农业等 9 大领域 42 个细分领域设备购置与更新改造
2024.4.31	山东	《山东省推动工业领域设备更新实施方案》	到 2025 年，全省高新技术产业产值占规模以上工业总产值比重达到 53%左右， 制造业数字化水平明显提升，信息化与工业化融合发展指数保持全国前列 ，工业重点领域达到能效标杆水平的产能比例达到 35%。 到 2027 年，全省工业领域设备投资规模较 2023 年增长 28%以上， 规模以上工业企业数字化研发设计工具普及率、关键工序数控化率分别超过 92%、75%，规模以上工业企业数字化转型实现基本全覆盖 ，重点行业低于能效基准水平的全部完成更新改造、主要用能设备能效基本达到节能水平，本质安全水平明显提升，创新产品加快推广应用，先进产能比重明显提高。
2024.5.31	上海	《上海市推动工业领域大规模设备更新和创新产品扩大应用的专项行动》	到 2027 年，实现本市工业领域设备更新规模累计达 2000 亿元，设备投资规模较 2023 年增长 25%以上， 规模以上工业企业数字化研发设计工具普及率、关键工序数控化率分别超过 95%、77%，规上工业企业数字化改造全覆盖 ，重点行业能效基准水平以下产能基本退出、主要用能设备能效基本达到节能水平，本质安全水平明显提

			<p>升，市级创新产品推荐目录突破 1500 个，先进产能比重持续提高。</p> <p>基础软件和工业软件研发与推广。聚焦重点行业，对基础软件和工业软件等重大研发项目，放宽新增投资额度及支持比例，支持比例最高为项目新增投资的 30%，支持金额原则上不高于 1 亿元。鼓励以用促研，对采购使用本市重点软件产品的用户单位，按照软件销售金额予以一定补贴。</p> <p>智能工厂领航计划。加快高水平智能工厂建设，到 2027 年培育不少于 20 家标杆性智能工厂、300 家示范性智能工厂，加快推动全市规上企业开展智能化改造。</p>
2024.5.31	河南	《河南省推动工业领域设备更新实施方案》	<p>到 2027 年，每年实施技术改造项目 3000 个左右，有效益的投资和有潜能的需求水平显著提升，工业领域设备投资规模较 2023 年增长 25%以上，规模以上工业企业数字化研发设计工具普及率、关键工序数控化率分别超过 90%、75%，建成 500 个智能工厂、500 个绿色工厂。重点行业能效基准水平以下产能基本退出、主要用能设备能效基本达到节能水平，本质安全水平明显提升，创新产品加快推广应用，先进产能比重持续提高。</p>
2024.6.5	江苏	《江苏省推动工业领域设备更新实施方案》	<p>到 2025 年，全省工业领域设备投资规模较 2023 年增长 15%以上，重点行业落后低效设备更新基本完成，规模以上工业企业数字化研发设计工具普及率超过 90%，关键工序数控化率达到 70%左右，重点用能行业能效基准水平以下产能基本退出；</p> <p>到 2027 年，全省工业领域设备投资规模较 2023 年增长 30%左右，规模以上工业企业关键工序数控化率超过 75%，推动规模以上工业企业数字化改造实现全覆盖，重点行业先进产能比重大幅提升，主要用能设备能效基本达到节能水平。</p> <p>到 2027 年，新增智能车间 3000 个、智能工厂 1000 家。推进设备网络化联接，全省数字化生产设备联网率达到 60%左右。</p>

资料来源：地方政府网站，中原证券研究所

4.3. 华为引领下的国产突围

近年来，华为全面构建自身的核心能力，以一己之力对标多个 IT 核心领域，成为了国产厂商当之无愧的领军者。软件层面，产品涉及到了最底层的编译器、编译语言，核心软件层的鸿蒙操作系统、GuassDB 数据库实现了突围，MetaERP、EDA 软件通过多年攻关摆脱了对海外产品的依赖。芯片层面，拥有自研的达芬奇架构，芯片产品全面覆盖了手机芯片、基带芯片、CPU 芯片、GPU 芯片。在人工智能领域，华为从底层的昇腾芯片、CANN 异构计算架构、昇思 AI 计算框架到自有的盘古大模型实现了完整的产品覆盖，作为国内最具竞争实力人工智能芯片厂商，带领我国人工智能产业实现了突围。

表 8: 近年来华为核心产品发布的情况

	自有产品	对标产品/厂商	进展
编译器	方舟		2019 年 8 月上线，开源
编译语言	仓颉		2022 年 11 月发布
操作系统	鸿蒙	安卓	<p>2019 年 8 月发布，2020 年 9 月发布鸿蒙 2.0。2021 年 8 月鸿蒙 2.0 正式上线。</p> <p>2022 年 7 月发布鸿蒙 3.0。2023 年 8 月发布鸿蒙 4.0、HarmonyOS Next 开发者预览版本。2024 年 6 月 HarmonyOS Next 进入开发者和先锋用户 Beta 测试阶段，2024Q4 将开启大规模商用。</p> <p>全场景，微内核，分布式</p>

服务器操作系统	欧拉	Windows、Linux	2021年9月推出OpenEuler，开源 2022年10月，OpenEuler22.09正式版发布，实现欧拉与鸿蒙的互通
数据库	GuassDB	Oracle	2019年5月发布，2023年6月发布新一代分布式产品 开源
ERP	MetaERP	SAP、Oracle	2023年4月发布
移动服务	HMS	GMS	2020年1月发布HMS Core 4.0，2020年3月在P40系列上线 2021年10月，发布HMS Core 6
AI计算框架	昇思 MindSpore	PyTorch、TensorFlow、 飞桨	2019年8月发布，开源
EDA		新思科技、楷 橙电子	2023年2月宣布华为芯片设计EDA工具团队联合国内EDA企业，共同打造了14nm以上工艺所需EDA工具，基本实现了14nm以上EDA工具的国产化，2023年将完成工具全面验证 2024年3月发布面向中小规模IC设计企业的半导体EDA仿真基础设施解决方案
异构计算架构	CANN	英伟达 CUDA	2018年10月，发布CANN 1.0，2021年12月发布CANN 5.0 2023年底发布CANN 6.0
大模型	盘古	GPT、文心一 言	2021年4月正式发布盘古大模型 2022年4月发布盘古大模型2.0 2023年7月发布盘古大模型3.0 2024年6月发布盘古大模型5.0
芯片架构	达芬奇	Arm	基于ARM架构，是华为自研的面向AI计算特征的计算架构
手机芯片	麒麟	高通、苹果	2018年发布麒麟980，全球首款7nm手机芯片 2019年9月发布麒麟990和麒麟990 5G，7nm，全球首款商用旗舰5G SoC芯片 2020年10月发布麒麟9000芯片，全球第一颗5nm芯片 2023年8月，搭载麒麟9000S的华为Mate60 Pro发布
基带芯片	巴龙	高通	2019年1月发布巴龙5000，7nm，首款单芯片多模的5G芯片
CPU芯片	鲲鹏	英特尔、AMD	2019年1月发布鲲鹏920，是第三代芯片 2023年11月联合19家伙伴共同发布了19款鲲鹏一体机产品
GPU芯片	昇腾	英伟达、AMD	采用达芬奇架构，2019年8月发布昇腾910 2023年10月，科大讯飞表示昇腾910B能力已经基本做到可对标英伟达A100

资料来源：华为，CNMO手机中国，百度百科，科大讯飞公告，中原证券研究所

4.3.1. 华为算力芯片布局加快趋势明显

2023年华为在国产服务器芯片的布局方面成效显著，2024有望持续加速。

2019年华为开始宣布硬件开放、软件开源，华为作为业界唯一同时拥有“CPU、NPU、存储控制、网络互连、智能管理”5大关键芯片的厂商，利用自己的硬件能力，对外提供主板、SSD、网卡、RAID卡、Atlas模组和板卡，优先支持合作伙伴发展服务器等计算产品。在这样的背景下，国内成长出了宝德自强®鲲鹏、湘江鲲鹏兆瀚、神州鲲泰、Huanghe等一系列服务器领域的自主品牌。2023年以超聚变、清华同方、华鲲振宇、神州鲲泰（神州数码）、湘江鲲鹏为代表的服务器厂商，作为华为芯片的合作伙伴，获得了较快的出货量增长。

表 9: 服务器相关业务财务对比 (亿元)

公司	相关业务	2023 年							2022 年
		收入	增速	净利润	增速	毛利率	毛利额	人员增速	收入
工业富联	云计算	1943	-8.5%	-	-	5%	99	-	2124
浪潮信息	-	659	-5%	17.8	-15%	10%	66.1	-4%	695
联想集团	基础设施解决方案	\$86	-4%						\$90
新华三	计算存储	328	7%	-	-	14%	45.5	-	307
超聚变	-	284	21%	-	-	-	-	-	235
中科曙光	IT 设备业务	127	12%	18.4	9%	22%	28	10%	114
中兴通讯	服务器及存储								100
清华同方	(软通动力持股)								46
华鲲振宇									34
神州数码	自主品牌业务	38	49%	-	-	10%	4	-	26
拓维信息	计算机、通信和其他电子设备制造业	17.8	121%			12%	2.2		8.1
湘江鲲鹏	(拓维信息持股)	15.6		0.005					

资料来源：上市公司公告，Wind，大河网，中原证券研究所

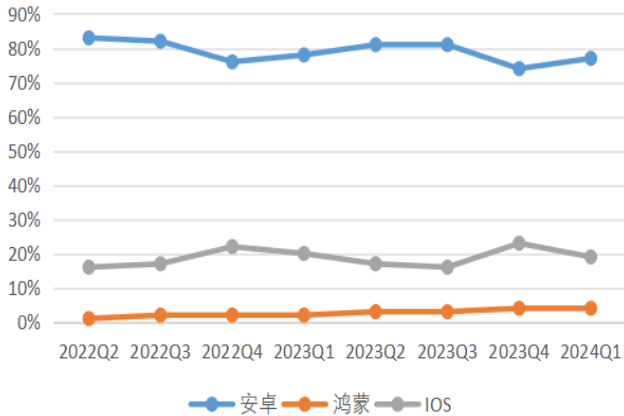
2024 年来看，华为的加速芯片在国产厂商中性能最为优异，也是运营商集采 AI 服务器的主力芯片。

4.3.2. 纯血鸿蒙与盘古大模型结合，打造系统级原生智能

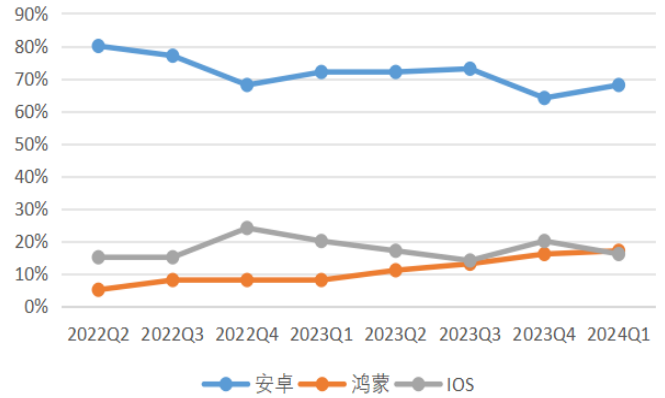
不兼容安卓的鸿蒙 NEXT 即将商用。2023 年 8 月，华为发布了鸿蒙 NEXT 开发者预览版本。2023 年 9 月，华为在全场景发布会上宣布鸿蒙原生应用全面启动。2024 年 6 月 21 日，华为在开发者大会 2024 上宣布鸿蒙 NEXT 正式进入开发者和先锋用户 Beta 测试阶段。到 2024Q4，前期参与 Beta 测试的设备将升级为正式版鸿蒙 NEXT 系统，鸿蒙 NEXT 系统将开启大规模商用阶段。

在市场占有率方面，鸿蒙生态设备数已经突破 9 亿台，拥有 254 万开发者。根据 CounterPoint 数据，2024Q1，华为鸿蒙系统在中国份额为 17%，超过了苹果 IOS 16% 份额，成为中国的第二大智能手机操作系统。同期，鸿蒙在全球智能手机操作系统份额为 4%，较上年 2% 的份额也有明显增幅。

图 34: 2022Q2-2024Q1 全球智能手机操作系统份额
图 35: 2022Q2-2024Q1 中国智能手机操作系统份额



资料来源: counterpoint, IT之家, 中原证券研究所



资料来源: counterpoint, IT之家, 中原证券研究所

在性能方面，鸿蒙 NEXT 超越了 Linux 内核，整体性能提升了 10.7%，将给设备带来 30%的性能提升，同时功耗下降 20%。

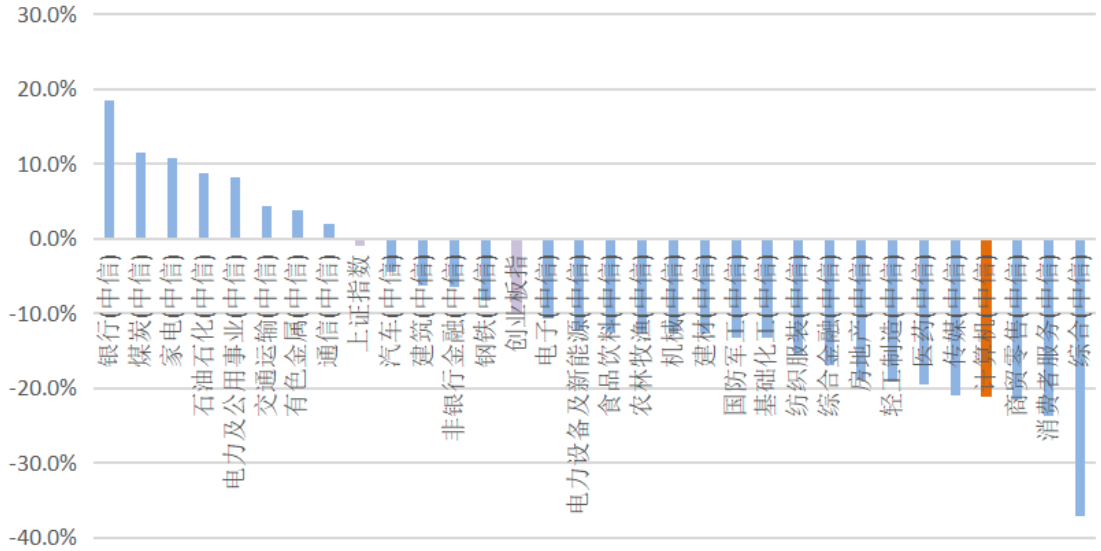
在产品先进性方面，鸿蒙 NEXT 是国内首个系统级原生智能。此前苹果发布了自研的 Apple Intelligence，并在其系统生态中集成了 GPT-4o，从而实现了大模型在与系统底层深度结合。而鸿蒙 NEXT 也首次将 AI 能力融入系统，结合自家的盘古大模型，带来 AIGC 图像生成、AIGC 声音修复、AI 图像识别等功能。考虑到鸿蒙是万物互联的全场景分布式操作系统，可以实现跨终端的链接，鸿蒙 NEXT 的 AI 能力又可以得到进一步地放大。

5. 投资策略

5.1. 行情回顾：2023 年人工智能、华为概念带动 TMT 领涨

2024 年计算机行业跌幅居前。截止 6 月 27 日，中信计算机指数下跌 21.23%，跑输大盘 20.25 PCT，跑输创业板指数 11.28 个百分点，在 30 个中信一级行业中排名第 27，在 TMT 行业中表现最弱。

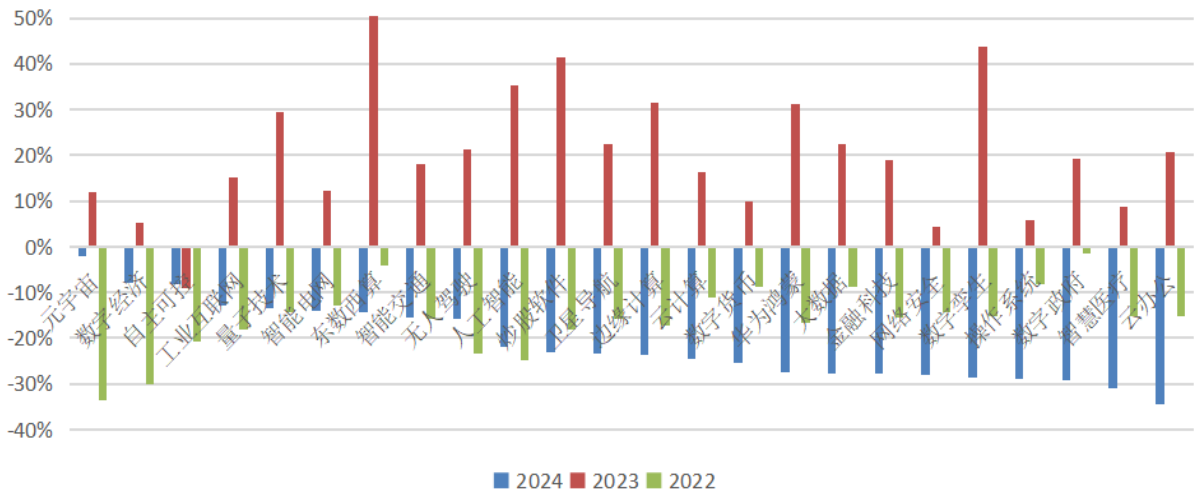
图 36：2024 年中信一级子行业涨跌幅（截止 2024.6.27）



资料来源: Wind, 中原证券研究所

从概念指数表现来看，2024 年计算机行业几乎所有的概念指数均处于下跌态势，其中元宇宙-1.92%、数字经济-7.78%、自主可控-8.28%相对抗跌，都是上一年表现较弱的板块，跌幅居前的为云办公 34.50%、智慧医疗 31.01%、数字政府 29.20%，更多受到了行业景气度的影响。

图 37: 近三年计算机行业相关概念年度涨跌幅(%, 2024.6.27)

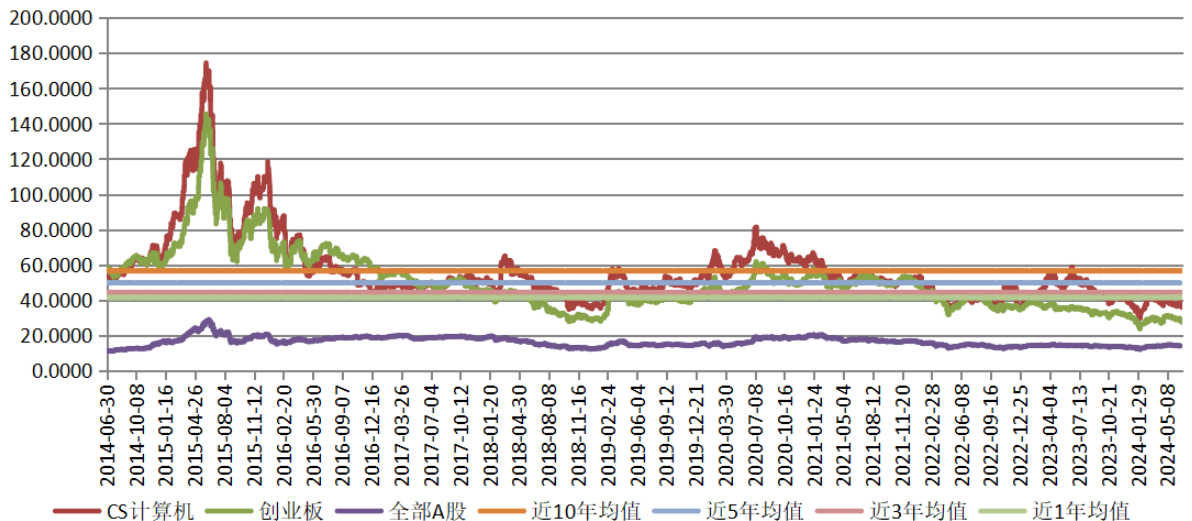


资料来源: Wind, 中原证券研究所

5.2. 估值: 明显低于中长期均值水平, 处于近三年估值较低的位置

行业的估值明显低于中长期均值水平, 处于近三年估值较低的位置。根据 Wind 数据, 2024 年 6 月 27 日中信计算机行业 TTM 整体法 (剔除负值) 估值为 36.74 倍, 同期 A 股估值 13.66 倍创业板估值 27.82 倍。行业近 1 年、3 年、5 年、10 年的平均估值分别为 41.40 倍、44.40 倍、49.66 倍、56.47 倍, 较近 3 年最高点 58.79 倍和最低点 29.70 倍, 分别有 60.00% 的上涨空间和 19.19% 的下跌空间。

图 38: 近 10 年申万计算机行业估值水平 (截止 2024. 6. 27)



资料来源: Wind, 中原证券研究所

5.3. 行业观点与投资建议

给予行业同步大市的投资评级。在全球经济景气度下滑的大背景下, 我们注意到计算机

行业虽然需求保持相对的稳定性，费用得到较好的控制，但是在收入质量方面仍然存在隐忧，或持续对业绩形成压力。在这样的发展背景下，行业不同赛道甚至是赛道中的不同企业之间的景气度差异或将进一步加大。

2024 年上半年，计算机行业跌幅居前，当前计算机行业估值明显低于中长期均值水平，处于近 3 年估值较低的位置，整体风险得到较好的释放，下半年可以重点聚焦的方向包括：

(1) 人工智能：作为当前创新推动的核心焦点，人工智能领域孕育了巨大的发展机遇。在 OpenAI 的 Sora、GPT-4o 发布的同时，开源模型也在加速壮大，对我国大模型产业进步起到了积极的推动作用。加速芯片领域英伟达持续引领，垄断趋势明显，但是国内需求在美国禁令之下，形成了对国产芯片厂商难得的发展机遇。在 OpenAI 暂停中国 API 接口、国内大模型的降价潮的影响下，国内大模型将获得更好的应用。**推荐个股金山办公（688111）。**

(2) 算力建设：大模型加速发展，算力建设重要性进一步凸显。在“适度超前建设”、智算比例持续提升的大趋势下，我们看到 2024 年运营商在加快算力布局、国内外科技大厂加大资本投入、地方算力建设相关政策集中发布，也将给产业链带来更多发展机遇。AI 计算、液冷、整机柜的结合将成为未来的发展趋势，液冷行业也迎来了刚需放量的元年。**推荐个股紫光股份（000938）。**

(3) 国产替代：在禁令加剧的大趋势下，我们看到芯片的国产化比例在持续提升、运营商集采国产化比例逐年增长、地方积极推动算力的国产化，工业领域更新换代也将推动国产化发展进程。此外，华为在算力芯片和纯血鸿蒙的动向，也将成为 2024 值得重点关注的方向。**推荐个股中望软件（688083）、中控技术（688777）、索辰科技（688507）。**

6. 风险提示

国际局势的不确定性；下游企业削减开支；金融风险的释放。

行业投资评级

强于大市：未来 6 个月内行业指数相对沪深 300 涨幅 10% 以上；

同步大市：未来 6 个月内行业指数相对沪深 300 涨幅 -10% 至 10% 之间；

弱于大市：未来 6 个月内行业指数相对沪深 300 跌幅 10% 以上。

公司投资评级

买入：未来 6 个月内公司相对沪深 300 涨幅 15% 以上；

增持：未来 6 个月内公司相对沪深 300 涨幅 5% 至 15%；

谨慎增持：未来 6 个月内公司相对沪深 300 涨幅 -10% 至 5%；

减持：未来 6 个月内公司相对沪深 300 涨幅 -15% 至 -10%；

卖出：未来 6 个月内公司相对沪深 300 跌幅 15% 以上。

证券分析师承诺

本报告署名分析师具有中国证券业协会授予的证券分析师执业资格，本人任职符合监管机构相关合规要求。本人基于认真审慎的职业态度、专业严谨的研究方法与分析逻辑，独立、客观的制作本报告。本报告准确的反映了本人的研究观点，本人对报告内容和观点负责，保证报告信息来源合法合规。

重要声明

中原证券股份有限公司具备证券投资咨询业务资格。本报告由中原证券股份有限公司（以下简称“本公司”）制作并仅向本公司客户发布，本公司不会因任何机构或个人接收到本报告而视其为本公司的当然客户。

本报告中的信息均来源于已公开的资料，本公司对这些信息的准确性及完整性不作任何保证，也不保证所含的信息不会发生任何变更。本报告中的推测、预测、评估、建议均为报告发布日的判断，本报告中的证券或投资标的价格、价值及投资带来的收益可能会波动，过往的业绩表现也不应当作为未来证券或投资标的表现的依据和担保。报告中的信息或所表达的意见并不构成所述证券买卖的出价或征价。本报告所含观点和建议并未考虑投资者的具体投资目标、财务状况以及特殊需求，任何时候不应视为对特定投资者关于特定证券或投资标的的推荐。

本报告具有专业性，仅供专业投资者和合格投资者参考。根据《证券期货投资者适当性管理办法》相关规定，本报告作为资讯类服务属于低风险（R1）等级，普通投资者应在投资顾问指导下谨慎使用。

本报告版权归本公司所有，未经本公司书面授权，任何机构、个人不得刊载、转发本报告或本报告任何部分，不得以任何侵犯本公司版权的其他方式使用。未经授权的刊载、转发，本公司不承担任何刊载、转发责任。获得本公司书面授权的刊载、转发、引用，须在本公司允许的范围内使用，并注明报告出处、发布人、发布日期，提示使用本报告的风险。

若本公司客户（以下简称“该客户”）向第三方发送本报告，则由该客户独自为其发送行为负责，提醒通过该种途径获得本报告的投资者注意，本公司不对通过该种途径获得本报告所引起的任何损失承担任何责任。

特别声明

在合法合规的前提下，本公司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、财务顾问等各种服务。本公司资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告意见或者建议不一致的投资决策。投资者应当考虑到潜在的利益冲突，勿将本报告作为投资或者其他决定的唯一信赖依据。