

分析师：唐月
登记编码：S0730512030001
tangyue@ccnew.com 021-50586737

美国制裁密集出台，全面围堵对我国的 AI 算力供应能力

——计算机行业专题

证券研究报告-行业专题报告

同步大市(维持)

计算机相对沪深 300 指数表现

发布日期：2025 年 01 月 24 日



资料来源：中原证券，聚源

相关报告

《计算机行业月报：AI 模型成果密集发布，制裁或将全面升级》 2025-01-10

《计算机行业深度分析：经济工作会议聚焦人工智能+，地方发力产业建设》 2024-12-20

《计算机行业月报：AI 应用潜力亟待释放，禁令下国产化持续深入》 2024-12-09

联系人：李智

电话：0371-65585629

地址：郑州郑东新区商务外环路 10 号 18 楼

邮编：上海浦东新区世纪大道 1788 号 T1 座 22 楼

投资要点：

在美国政府换届前期，美国制裁方案接连出台，已经形成了对我国 AI 算力供应全面围堵之势。我们希望通过梳理制裁政策出台的情况和制裁的行业背景，对后续国内 AI 发展做出预判。

- 从当前形势来看，AI 作为当下科技布局最重要的发展方向，已经成为了中美科技企业争相投入的方向。我国目前在大模型领域较美国的差距在逐步缩小，同时在自动驾驶、机器人、无人机、摄像头等 AI 应用方向上形成领先优势，导致美国在 AI 相关领域对我国的制裁持续收紧，以期通过硬件限制来形成模型上的代差。
- 结合全球 1 月 13 日的全球 AI 管控新规来看，美国一方面在进行全球 GPU 供应方面的管控，另一方面通过代工环节限制竞争对手的供应。同时，2024 年底以来的制裁措施，以打补丁的方式，意图形成对我国更加有效的技术封锁。
- 我们的突破取决于禁令执行层面的严格程度和上游芯片代工、光刻机等卡脖子环节的技术突破能力。在 2022 年和 2023 年的两轮制裁后，我国国产芯片替代的能力持续提升，产业届在国产替代方面形成了较好的共识。在应对制裁的过程中，以华为为代表的厂商也形成了较好的突围能力，为后续应对持续严苛的制裁措施积累了丰富的经验。
- 同时，幻方给国内大模型打样之后，也为国内模型发展探索出一条发展的新路径。在 2025 年 1 月 20 日举办的专家、企业家、科教文体代表座谈会中，幻方创始人梁文锋也作为代表出席了会议，表明了中央对于幻方路径的认可和对人工智能的关注，有望在后续政府工作报告中出台更多有利于 AI 行业发展的政策。
- 2025 年 1 月 20 日美国新一届政府上台，我们看到新政府在与中方的各项关系方面表现得更为缓和，虽然 AI 仍将是双方后续关注的焦点，但是我们认为在制裁执行和新的制裁措施下达方面仍具有缓和空间。

风险提示：国际局势的不确定性；上游企业削减开支。

内容目录

1. 美国对我国科技领域制裁全面升级	3
1.1. 面向 HBM、设备厂商、EDA 等，制裁全面化	3
1.2. AI 领域制裁成为美国反制手段之一	3
1.3. 制裁范围或进一步扩大	4
1.4. 全面管控 AI 芯片和模型的全球应用	4
1.5. 限制网联汽车在美销售	6
1.6. 限制 16nm 和 14nm 芯片代工	6
2. 我国开始更加积极地反制	8
2.1. 三中提出健全反制裁机制	8
2.2. 2024 年 12 月以来我们采取的反制措施	8
3. 海内外 AI 投入持续加大，中美 AI 大模型差距持续缩小	9
3.1. 我国在 AI 领域差距持续缩短	9
3.2. 海内外算力投入持续加大，智算成为发展的核心	11
3.3. 2024 年 AI 芯片国产化比例快速提升	12
4. 总结与预判	14
5. 风险提示	15

图表目录

图 1: 美国芯片限制的区域划分	5
图 2: 2022Q1-2024Q3 英伟达单季度收入的区域分布 (亿美元)	6
图 3: 2021 年中芯国际不同制程的收入结构	7
图 4: 2024Q4 台积电不同制程的收入占比	7
图 5: 2024Q4 台积电来自不同平台的收入占比	8
图 6: 2024Q4 台积电来自不同平台的收入增速	8
图 7: 2024 年 10 月 AI 应用榜单 (中国)	10
图 8: DeepSeek-R1 与 o1 在调用成本差距 (美元/百万 token)	11
图 9: 阿里、腾讯、百度的资本开支 (亿元)	12
图 10: 5 大科技厂商资本投入 (百万美元)	12
图 11: 2021-2024E 超聚变收入及增速 (亿元)	13
图 12: 2023H1-2024H1 我国 AI 服务器芯片国产化占比	14
表 1: OpenAI 大模型产品竞品推出情况	10
表 2: 大模型厂商算力相关投入	11
表 3: 2024 年企业服务器相关业务数据对比 (亿元)	13

在美国政府换届前期，美国制裁方案接连出台，已经形成了对我国 AI 算力供应全面围堵之势。我们希望通过梳理制裁政策出台的情况和制裁的行业背景，对后续国内 AI 发展做出预判。

1. 美国对我国科技领域制裁全面升级

1.1. 面向 HBM、设备厂商、EDA 等，制裁全面化

2024 年 12 月 2 日，美国 BIS 发布了《出口管理条例(EAR)》的修订说明，在修订半导体相关的出口管制规则的同时，将 140 个中国实体列入“实体清单”，本次制裁或将对我国的 AI 和芯片产业产生较大的影响。

本次管制的核心内容包括：（1）将 140 家公司列入实体清单，并修改对 14 家企业（标注“脚注 5”）供应含美技术外国产品的限制要求，同时从验证最终用户（VEU）“白名单”计划中删除 3 家中国企业；（2）针对高带宽内存（HBM）增加新的管制措施；（3）扩展部分半导体制造设备和相关物品的管制品类，适用于外国直接产品规则。

本次管制与计算机行业相关的要点有：

（1）对于 HBM 的出口限制，对 AI 芯片影响较大；

（2）包括华为产业链、先进制造、科研院所等多个方面的实体受到了脚注 5 的限制。

（3）国产设备厂商受到了新增限制，将受到严格审查和供应链审查。

（4）首次提出对先进制造的软件进行限制，此外 EDA 企业也进入实体清单，华大九天及其子公司都成为了禁令限制的对象。

1.2. AI 领域制裁成为美国反制手段之一

美国商务部下属工业与安全局（BIS）于 2025 年 1 月 3 日修订《出口管理条例》（EAR），将中国（11）、缅甸（1）和巴基斯坦（1）三个国家共 13 个实体加入实体清单，自 2025 年 1 月 6 日起生效。11 个清单当中，大部分实体与超音速飞机产业链、AI 技术等高科技领域方向有关，包括长春光机所、上海光机所、季华实验室、鹏程实验室、南京施密特光学仪器有限公司等。

美国财政部在 2025 年 1 月 3 日发布声明称，已对永信至诚实施制裁，将其列入 SDN 清单（特别指定国民清单），被控为国家网络攻击提供便利。

永信至诚是数字安全“测试评估赛道”领跑者，网络靶场和网络安全人才建设领军者，国家级专精特新“小巨人”企业。公司自主研发的网络靶场核心技术，获北京市科技进步奖一等奖（2020 年）、国家科技进步奖二等奖（2024 年），属网络空间安全领域的硬科技。公司首

创“数字风洞”测试评估产品技术体系，为用户在数字化、智能化转型中面临的网络安全、数据安全等问题提供了切实有效的解决方案。此次制裁成为了对公司能力和行业地位的背书。

我们认为本轮美国的制裁更多是针对中国 2025 年 1 月 2 日制裁行动的反制裁，总体影响较为有限。

1.3. 制裁范围或进一步扩大

2024 年 12 月 17 日，美国国防部宣布已于 2024 年 12 月 13 日将中微半导体设备（上海）股份有限公司和 IDG 资本从中国军事公司清单（CMC 清单或 1260H 清单）中移除，这意味着两家公司不再正式被认定与中国军方存在直接联系。其中中微公司于 2024 年 1 月底被列入 CMC 清单，并从 8 月开始将美国国防部告上了法庭，取得了对抗制裁良好的反击效果。

当地时间 2025 年 1 月 6 日，美国国防部发布最新版本“中国涉军企业”（CMC）将新一批中国公司列入黑名单，其中新增腾讯控股、宁德时代、移远通信(603236.SS)、长鑫存储、商汤科技、无人机制造商道通科技 Autel Robotics (688208.SS)等。同时五角大楼取消了 6 家据称不再符合指定要求的公司，包括人工智能公司旷视科技、中国铁建、深圳市科思科技、中国海防、中国建筑和中国电信。

总体来看，国防部 CMC 清单对企业影响总体影响有限，且通过诉讼方式已经有多家公司成功移除。但是腾讯首次进入美国制裁名单，表明接下来美国制裁将不仅仅局限于限制我国军事领域、关键技术环节等领域。

1.4. 全面管控 AI 芯片和模型的全球应用

2024 年 12 月 20 日消息，美国商务部近期要求英伟达调查过去一年该公司产品是如何流入中国的。据悉，一些客户复制了从超微公司购买的嵌有英伟达芯片的服务器的序列号，并将其连接到他们可以访问的其他服务器上。在某些情况下，走私者甚至会更改服务器操作系统中的序列号。

2025 年 1 月 13 日，BIS 正式发布了全球 AI 管控新规，首次全面规范了 AI 芯片和模型在全球范围内的管控，对出口限制进行了全面的升级。

新规将全球国家和地区分为 3 个等级进行管控，包括：

Tier 1 的美国及其少数盟友（包括美国及德国、荷兰、日本、韩国等 18 个同盟），可以不受限制获取美国芯片，总部在该区域的企业在 Tier1 地区以外的算力不得超过总和的

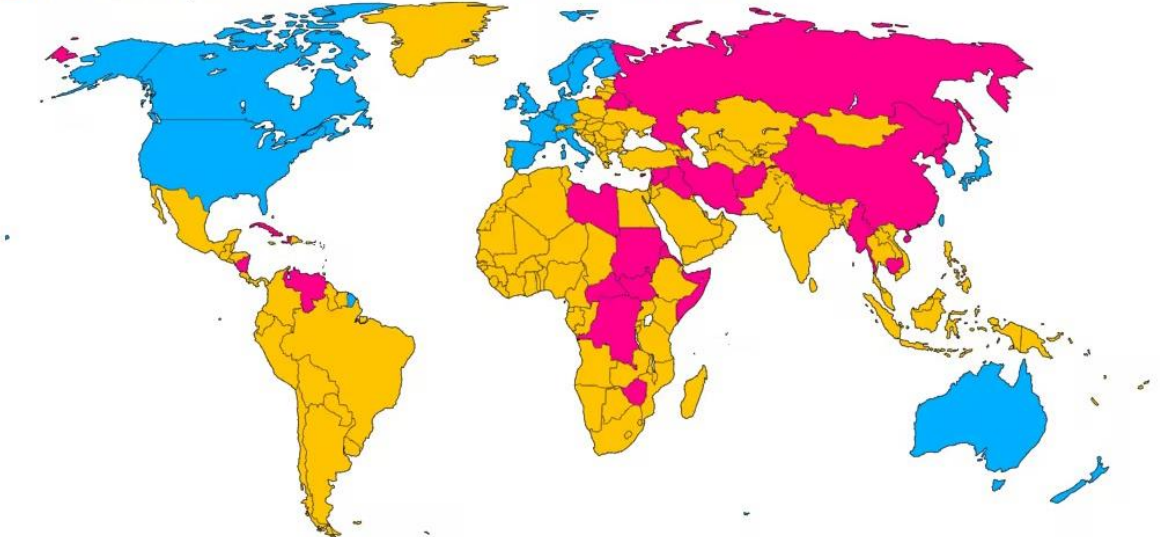
25%，在单一 Tier 2 区域的算力不超过 7%，总部在美国的企业需保证 50%的算力在美国境内。

Tier 2 的国家和地区（包括了全球多数国家和地区），面临以国为单位的总算力限制，2025-2027 年期间每个国家只能获得 5 万块 H100 等效 GPU（这些国家通过申请 NVEU 资格可以获批到最高 32 万块 H100 的等效算力），以保证这些国家在训练模型是所用的计算集群落后美国 12 个月或一代。企业的高性能闭源模型部署在 Tier 2 区域的部署需要获得美国政府许可。

Tier 3 的 24 个国家和地区（包括中国、俄罗斯、白俄罗斯、伊朗、中国澳门等美国武器禁运的国家和地区），数据中心将被全面禁止进口芯片，企业不能在这些区域部署高性能的闭源模型。

图 1：美国芯片限制的区域划分

■ Tier 1 (Most permissive) ■ Tier 2 ■ Tier 3 (Most restrictive)



Source: Bloomberg reporting

Note: Mapped data show level of restrictions on chip shipments for distinct markets

资料来源：彭博，新智元，中原证券研究所

如果卖家订购不超过 1700 个 H100 将不用计入上限，只需向政府进行通报，无需获得许可，这意味着大学、医疗机构、研究组织等低风险出货不受到以上限制。

新规将在 120 天后生效，将对全球 AI 算力格局产生深远影响，将由新一届政府决定是否继续执行。

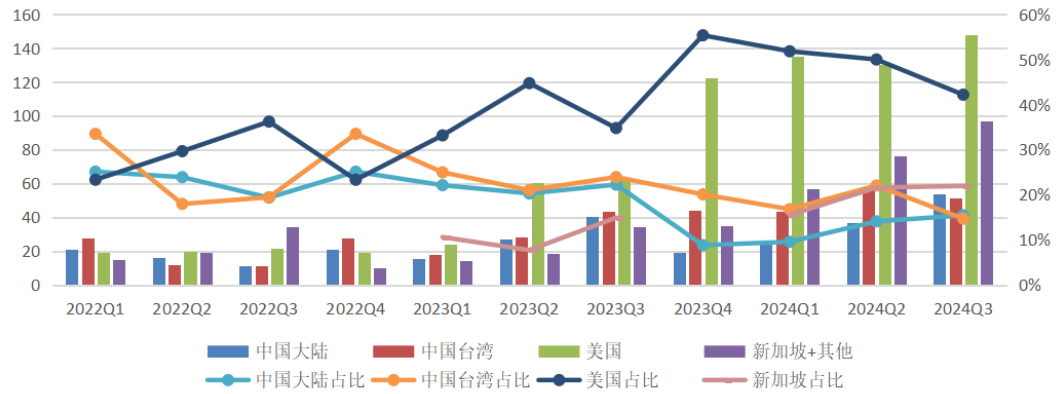
从英伟达最新财报来看，2024Q3 其收入中，42%来自美国，22%来自新加坡，来自中国大陆和中国台湾收入占比分别为 15%。其中新加坡是 Tier 2 区域，如果新规实施，我们认为出口到新加坡的算力芯片将受到较大的影响。

结合此前新闻，2024 年 12 月 20 日消息，美国商务部近期要求英伟达调查过去一年该公司产品是如何通过非正规渠道流入中国。因而我们认为在新规的全面管控下，通过非合法渠

道满足的国内算力需求也将同样受到影响。

考虑到全球大部分地区都处于受限区域，这意味着这些区域将释放出大量的 AI 算力需求。我国国产 AI 芯片厂商如果能解决代工问题，将在全球范围内迎来一个广阔的市场空间。

图 2：2022Q1-2024Q3 英伟达单季度收入的区域分布（亿美元）



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

1.5. 限制网联汽车在美销售

当地时间 2025 年 1 月 14 日，美国政府通过白宫官网正式公布了针对网联汽车的最终限制规定，以规避中国和俄罗斯利用美国网联汽车供应链所带来的美国国家安全风险。最终限制规定禁止美国销售和进口中国产和俄罗斯产网联汽车软硬件系统及网联汽车，包括车辆互联系系统（VCS），或通过蓝牙、蜂窝网络、卫星连接和 Wi-Fi 模块将车辆与外部世界连接起来的系统和零部件，以及自动驾驶系统（ADS），使用 VCS 和 ADS 软件的网联汽车，以及 VCS 硬件设备。对软件和网联汽车的限制规定将于 2027 车型年起生效，对硬件的限制规定将于 2030 车型年起生效（对于没有车型年份的车辆，硬件禁令从 2029 年 1 月 1 日起生效）。上述最终规定适用于乘用车，而针对商用网联车辆供应，美国商务部也在计划制定中。

2024 年 9 月 23 日，美国商务部工业和安全局就发布了该制裁措施的拟议规则，并计划在 2025 年 1 月 20 日前最终敲定，因而在市场的预期之内。

当地时间 2025 年 1 月 16 日，美国贸易代表办公室发布了对华海事、物流和造船业 301 调查报告。这次调查于 2024 年 4 月 17 日宣布开启。

我们看到以上两项制裁都是在预先框架下推进的，更多是换届前政府加速推动制裁落地的结果。但是网联汽车作为 AI 应用的重要领域，围堵趋势可以看做限制我国 AI 发展的一个重要的表现。

1.6. 限制 16nm 和 14nm 芯片代工

当地时间 2025 年 1 月 15 日，美国商务部工业和安全局（BIS）发布两项新规。

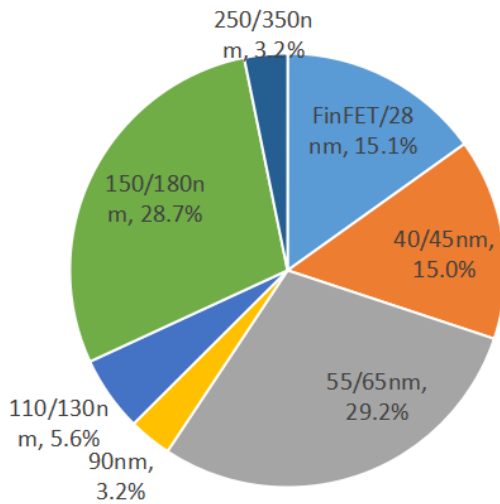
(1) 新规将对 16nm 及以下工艺或采用非平面晶体管结构生产的逻辑电路进行出口管制，对 DRAM 会实施更加严格的管制，而对于晶体管数量少于 300 亿个并由一家值得信赖的公司封装的先进芯片不会受到限制。

(2) 14 家中国实体和 2 家新加坡实体列入实体清单，其中中国实体为 13 家智算芯片企业算能科技及其子公司和封测企业渠梁电子。同时，BIS 将 11 家中国实体列入实体清单，理由是 10 家实体（智谱及其子公司）开发和整合了先进的 AI 研究，哈勃投资的科益虹源参与了中国先进节点制造设施的光刻技术开发。本次实体清单企业涉及到了大模型、芯片、封测、光刻机等产业环节。

2024 年 11 月，台积电和三星已经开始停止向中国大陆客户供应 7nm 及更先进工艺的 AI 芯片，本次新规对先进制程 AI 芯片的代工环节限制进一步加大到 14nm 和 16nm，美国还将加强对台积电、三星、英特尔等晶圆代工厂的尽职调查，并防止其流向中国。

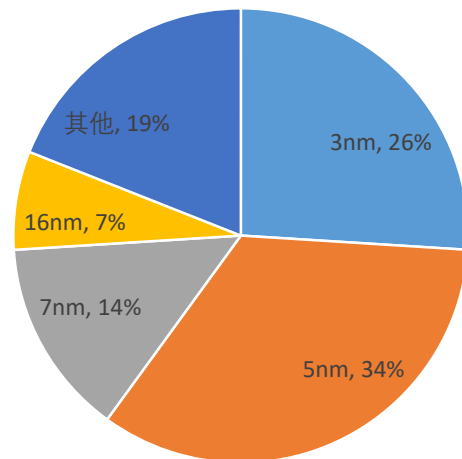
从 2024Q4 台积电财报来看，其收入中有 74% 来自 7nm 及以下制程，另有 16nm 制程收入占比 7%，对应本次新增的制程范围。对于国内代工能力来说，根据中芯国际财报，其在 2021 年 14nm 和 28nm 制程的收入占比 15.1%，具有 14nm 代工能力，但是更多产能集中在成熟支撑领域（中芯国际后续不在年报中披露各制程收入比例数据，此处仅作为参考）。

图 3：2021 年中芯国际不同制程的收入结构



资料来源：中芯国际公告，中原证券研究所

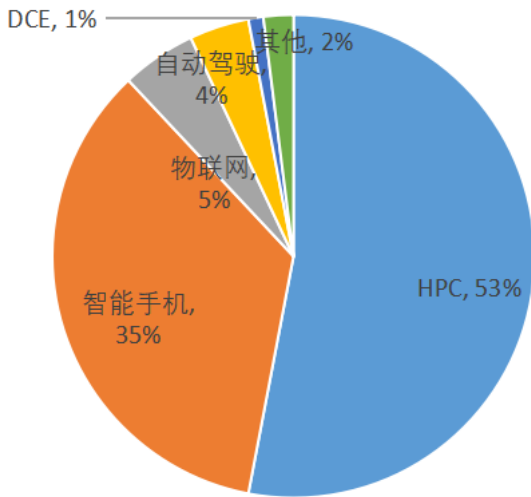
图 4：2024Q4 台积电不同制程的收入占比



资料来源：台积电，中原证券研究所

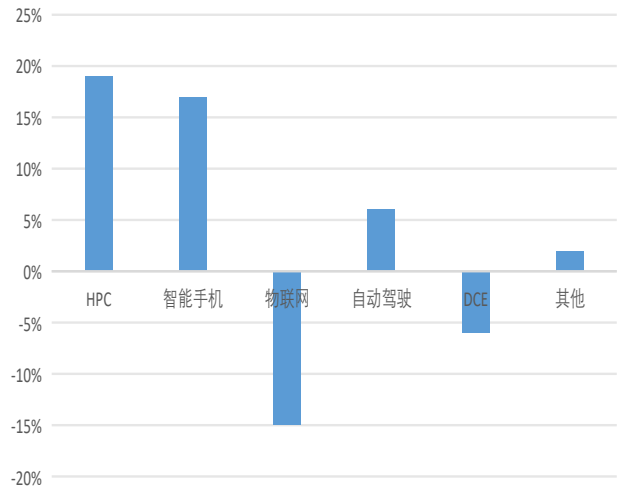
按平台来计算，2024Q4 台积电来自 HPC（高性能计算）的收入占比已经达到 53%，这部分收入在 2024Q4 同比增长了 19%，需求仍然非常强劲。

图 5：2024Q4 台积电来自不同平台的收入占比



资料来源：台积电，中原证券研究所

图 6：2024Q4 台积电来自不同平台的收入增速



资料来源：台积电，中原证券研究所

2. 我国开始更加积极地反制

2.1. 三中提出健全反制裁机制

在 2024 年 7 月 21 日二十届三中全会的相关决定中，我们看到第十三部分关于“推进国家安全体系和能力现代化”的内容中提到：“健全反制裁、反干涉、反‘长臂管辖’机制”，可以视为是层面面对制裁手段的政策升级。

在此后，我们面对的制裁采取了更加积极的反制。以无人机为例，2024 年 6 月 14 日，美国众议院通过一项无人机法案，依图禁止大疆无人机进入美国市场，但由于大疆无人机强大的市场覆盖率，美国参议院在 2024 年 7 月 11 日最新发布的国防授权法案中删除了关于大疆的禁令。然而，2024 年 7 月 31 日，我国商务部、海关总署、中央军委武装发展部联合发布公告，将于 2024 年 9 月 1 日开始对无人机出口进行管制。

2.2. 2024 年 12 月以来我们采取的反制措施

2024 年 12 月 12 月 3 日，中国互联网协会、中国半导体行业协会、中国汽车工业协会、中国通信企业协会相继发声，建议中国企业谨慎采购美国芯片。

2024 年 12 月 12 月 3 日，我国决定加强有关两用物项对美国出口管制，包括禁止两用物项对美国军事用户或军事用途出口，严控镓、锗、锑、超硬材料、石墨相关两用物项对美国出口等。

2024 年 12 月 12 月 5 日，外交部发起了对美国 13 家军工企业级 6 名企业高管的制裁。

2024 年 12 月 12 月 5 日，财政部发布《关于政府采购领域本国产品标准及实施政策有关

事项的通知（征求意见稿）》，提出“对本国产品报价给予 20% 的价格补贴”。

2024 年 12 月 9 日，国家市场监督管理总局消息称，因英伟达公司涉嫌违反《中华人民共和国反垄断法》及《市场监管总局关于附加限制性条件批准英伟达公司收购迈络思科技有限公司股权案反垄断审查决定的公告》（市场监管总局公告〔2020〕第 16 号），市场监管总局依法对英伟达公司开展立案调查。

以上反制措施，可以视为 BIS 将 140 个中国实体列入“实体清单”的反制措施。

2025 年 1 月 2 日，商务部还 10 家美国军工列入**不可靠实体清单**，同时商务部将 28 家美国军工企业列入**出口管制清单**，禁止向其出口双用途物品。在此次制裁中，我国的角色更具主动性，对美国主要军工企业或将产生较大的负面影响，因而美国也对我国 AI 相关产业发布了新的制裁。

2025 年 1 月 14 日，商务部发布公告，将参与对台湾地区军售的 7 家美国企业列入不可靠实体清单。这是 2025 年以来，商务部第三次针对美国军工企业出台限制措施。

2025 年 1 月 16 日，商务部就国内有关芯片产业反映自美进口成熟制程芯片低价冲击国内市场事答记者问。商务部称，国内有关芯片产业反映，一段时间以来，拜登政府对芯片行业给予了大量补贴，美企业因此获得了不公平竞争优势，并对华低价出口相关成熟制程芯片产品，损害了中国国内产业的合法权益。对此商务部表示，调查机关将按照中国相关法律法规，遵循世贸组织规则进行审查，并将依法启动调查。（消息来源：新快报）**这成为了针对美国政府换届前一系列制裁措施的反制手段。**

针对本次调查，当日中国机电商会、中国半导体行业协会也纷纷表态支持。

3. 海内外 AI 投入持续加大，中美 AI 大模型差距持续缩小

3.1. 我国在 AI 领域差距持续缩短

从目前人工智能的发展来看，我们与 OpenAI、谷歌、Anthropic 等全球头部大模型公司差距基本控制在半年以内。相比 2023 年来看，国内大模型在模型能力、产品价格、算力需求等方面都取得了都亮点十足，同时国内的头部模型还在积极选择开源方式，积极推动大模型技术加快发展。

表 1: OpenAI 大模型产品竞品推出情况

对标模型	时间	参数	厂商	产品	推出时间	备注
GPT-4o	2024.5.14	多模态交互模型	谷歌	Project Astra	2024.5.15	
				Gemini Live	2024.8.14	
			Anthropic	Claude 3.5 Sonnet	2024.6.21	
			智谱	GLM-4-Plus	2024.8.29	
			阿里	Qwen-Max 升级版	2024.9.19	
			字节	豆包通用模型 pro	2024.12.18	价格为 GPT-4o 的 1/8
		幻方	DeepSeek-V3	2024.12.26	开源	
o1	2024.9.13	推理模型	Anthropic	升级版 Claude 3.5 Sonnet	2024.10.23	
			谷歌	Gemini 2.0 Flash Thinking	2024.12.20	
			幻方	DeepSeek-R1-Lite-Preview	2024.11.21	对标 o1-preview
				DeepSeek-R1	2025.1.20	开源
			智谱	GLM-Zero-Preview	2024.12.31	对标 o1-preview
o3	2025.1	o1 的下一代产品				

资料来源: OpenAI, 界面新闻, 机器之心 Pro, 腾讯科技, 中原证券研究所

豆包大模型推出明显落后于行业, 但是却成为目前国内用户量第一的模型产品, 说明了大模型领域目前并不具备明显的先发优势。反而对于后发的厂商来说, 模型研发方向更加明确, 节约了试错成本, 同时算力采购成本也更低。

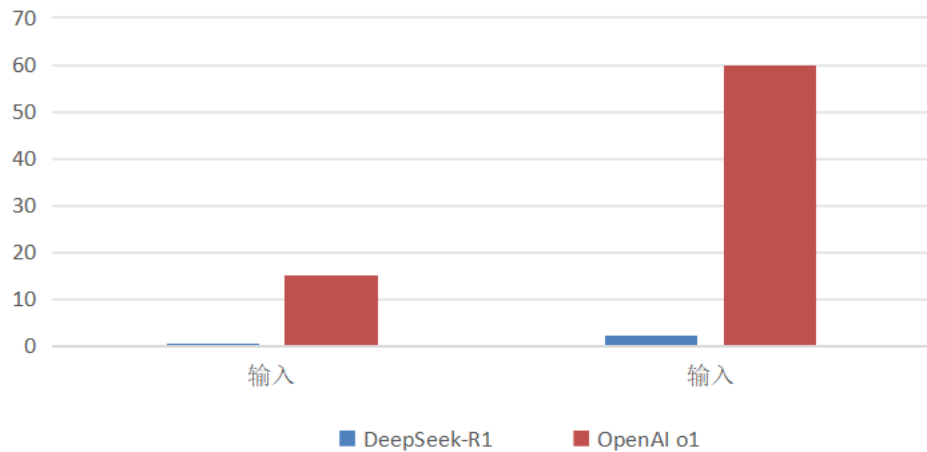
图 7: 2024 年 10 月 AI 应用榜单 (中国)



资料来源: 奇异因子, 中原证券研究所

幻方作为国内一家量化私募, 在国内算力供应和自身资金投入有限的情况下, 2024 年 12 月发布的 DeepSeek-V3 在性能上看齐了 GPT-4o, 同时模型训练成本仅为其 1/20。2025 年 1 月发布的 DeepSeek-R1, 又在性能上看齐了 o1, 同时 API 调用成本仅为其 1/30。此外, 幻方还将两大模型开源, 使得开源大模型在性能上看齐了闭源大模型水平。

图 8: DeepSeek-R1 与 o1 在调用成本差距 (美元/百万 token)



资料来源: DeepTech 深科技, 中原证券研究所

3.2. 海内外算力投入持续加大, 智算成为发展的核心

目前人工智能领域已经成为了科技巨头集中投入的方向, 我们看到巨头们都拟定了向数据中心或核电等基础设施方面的大规模投入计划, 投资的热度还在持续升温的过程中。2025 年 1 月 21 日, 甲骨文、OpenAI、软银宣布将共同出资 5000 亿美元建设名为“星际之门”的 AI 基础设施项目, 将 AI 基建力度拉升到了新的高度。

表 2: 大模型厂商算力相关投入

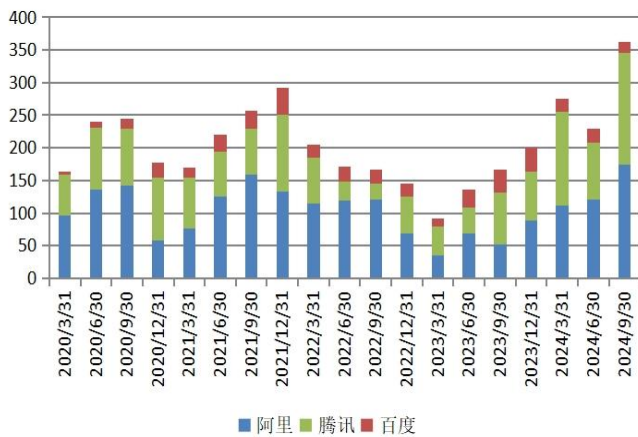
公司	时间	详情
微软	2024.3.31	与 OpenAI 正计划建设一个名为“星际之门”的特殊数据中心, 为 OpenAI 提供强大的计算支持, 计划最快在 2028 年启动该项目, 并且 2030 年之前会进一步扩建, 所需要的电力可能高达 5 千兆瓦。“星级之门”超算会配备百万专用的服务器芯片, 项目成本预计高达 1150 亿美元 。
	2025.1.4	将在本财年 (6 月结束) 投资 800 多亿美元 建设人工智能负载数据中心, 以训练人工智能模型, 并部署人工智能和基于云的应用程序。
OpenAI	2025.1.21	美国总统特朗普在白宫宣布, 由美国甲骨文公司、OpenAI 和软银共同出资投建, 三家公司组成的联合企业将投资 5000 亿美元, 在美建设名为“星际之门”的 AI 基础设施。该项目计划在美国建设数据中心, 以支持人工智能技术发展。
亚马逊	2024.10.29	亚马逊正在核能上押下重注, 包括在美国三个州投资超过 520 亿美元 , 以帮助推动其庞大数据中心的扩张。
	2024.12.16	AWS 计划追加投资约 100 亿美元, 在俄亥俄州扩建数据中心基础设施。新数据中心将包含计算机服务器、数据存储驱动、网络设备和其他形式的技术基础设施, 用于支持包括 AI 和机器学习在内的云计算。
谷歌	2024.12.11	谷歌与能源公司 Intersect Power 和投资公司 TPG Rise Climate 合作, 计划在本十年内投资 200 亿美元 在美国建设多个“工业园区”, 这些园区将配备可再生能源发电设施, 以支持数据中心运营。第一个园区预计将在 2026 年部分投入运营, 并于 2027 年全面完工。
xAI	2024.9.3	xAI 仅用 122 天就建成了 10 万张 H100 的 Colossus 集群, 未来还会扩展到 15 万张 H100 和 5 万张 H200。

	2024.12.5	xAI 承诺将其超级计算机项目 Colossus 的算力扩大十倍，未来将搭载 100 万块 GPU ，从而对抗谷歌、OpenAI 和 Anthropic 等竞争对手。
小米	2024.12.26	小米集团正在积极构建自己的 GPU 万卡集群，以加大对 AI 大模型的投入。
字节	2024.3.1	截止 2023 年 9 月，字节跳动已经建立超过一万张的英伟达 Ampere 架构 GPU 集群，目前正在建设 Hopper 架构的集群。
	2024.12.31	Information 报道，字节跳动计划在 2025 年斥资 70 亿美元 购买英伟达芯片，该消息随后被字节跳动发言人否定。
	2025.1.22	英国《金融时报》消息，字节跳动计划在 2025 年投资超过 120 亿美元 用于 AI 基础设施建设，其中 400 亿元人民币（约 55 亿美元）将用于购买中国 AI 芯片，是 2024 年的两倍。公司还计划在海外投资约 68 亿美元，以使用先进的 Nvidia 芯片来增强基础模型训练能力。该消息随后被字节跳动发言人否定。

资料来源：科创板日报，福布斯中国，第一财经，IT之家，新智元，金融界，中国新闻网，The Information，金融时报，同花顺，中原证券研究所

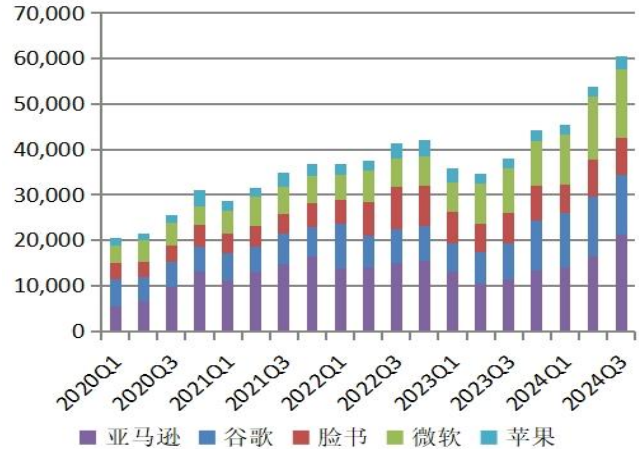
国内外的头部科技厂商的资本开支也在持续创出新高。 国外来看，亚马逊、谷歌、苹果、微软、脸书为代表的 5 大科技厂商的总资本开支继续呈现提升趋势，已经在 2024Q3 达到了 604.28 亿美元，同比增长了 59%；国内来看，2024Q3 阿里、腾讯、百度 3 家资本开支总计 362.22 亿元，同比增长 117.2%。

图 9：阿里、腾讯、百度的资本开支（亿元）



资料来源：上市公司财报，中原证券研究所

图 10：5 大科技厂商资本投入（百万美元）

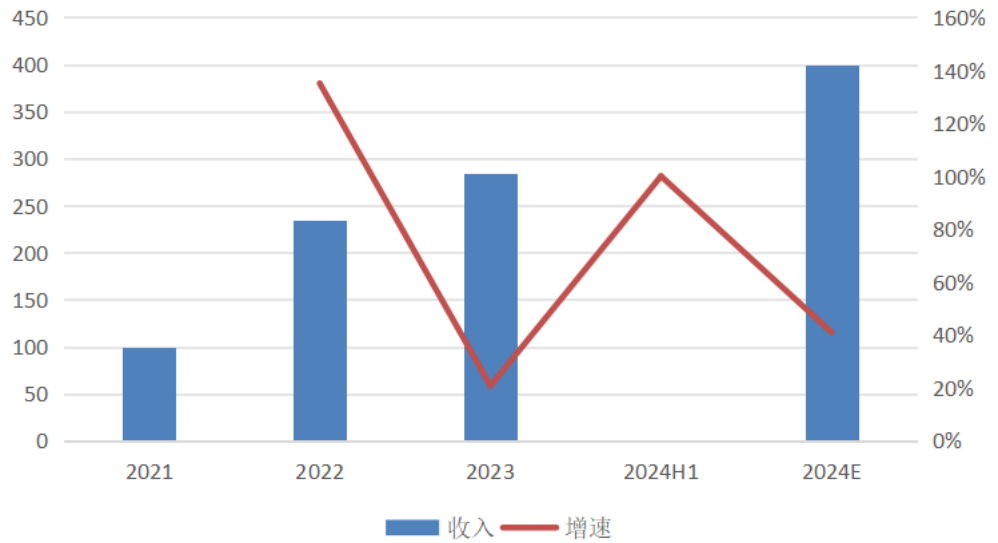


资料来源：Wind，中原证券研究所

3.3. 2024 年 AI 芯片国产化比例快速提升

超聚变作为华为昇腾加鲲鹏总产量第一的合作伙伴，2024 年实现收入的高速增长。作为 2021 年 9 月华为成立全资子公司，超聚变主要从事华为 X86 服务器业务，并在 11 月完成股权转让。根据华为计算官方视频号的信息，我们看到在 2024 年华为合作伙伴大会上，昆仑技术（超聚变子公司）COO 介绍，在过去一年当中，昆仑（及超聚变子公司）取得了鲲鹏加昇腾总的产量排名第一，以及昇腾排名第一，获得了昇腾的战略级合作伙伴以及鲲鹏的领先级伙伴的殊荣。2024 上半年，超聚变营收同比增长超过 100%，公司预计 2024 年收入达到 400 亿以上。

图 11: 2021-2024E 超聚变收入及增速 (亿元)



资料来源: 大河报, 大河立方, 超聚变, 中原证券研究所

对比同行来看, 超聚变在上半年收入增速继续保持了在头部企业中的领先趋势, 且同为华为合作伙伴的同方、神州泰坤、湘江鲲鹏也都实现了较快的增长, 这表明 2024 年华为服务器生态在国产替代方面取得了优异的成果。

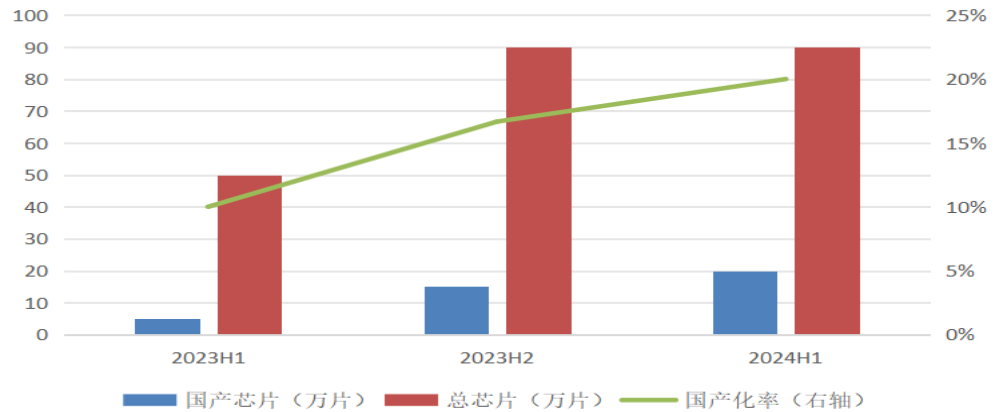
表 3: 2024 年企业服务器相关业务数据对比 (亿元)

公司	相关业务	2024H1						2024.1-9			
		收入	增速	净利润	增速	毛利率	毛利	收入	增速	收入两年增速	毛利两年增速
工业富联	云计算		60%	-	-				71%		
	-AI 服务器		230%						228%		
浪潮信息	-	421	69%	6.0	91%	7.7%	33	831	72%	58%	-12%
联想集团	基础设施解决方案	406	44%								
超聚变			100%								
紫光股份	-	380	5%	10	-2%	19%	72	588	7%	9%	-9%
	ICT 基础设施及服务	262	5%			24%	64				
	新华三	264	6%	18	0%			403	7%	10%	
中科曙光	-	57	6%	5.6	3%	26%	15	80	4%	9%	21%
	IT 设备业务	51	4%	-	-	14%	6				
软通动力	计算产品与数字基础设施服务	39				6%	2.5				
	同方计算机	15		-0.67							
神州数码	自主品牌业务(旧)	23.7	57%	-	-						
	自有品牌产品(新)	26.6	46%	0.44		10.5%					
	-信创业务	21.9	73%			12.1%					
	-神州泰坤人工智能服务器	5.6	273%								
拓维信息	计算机、通信和其他电子设备制造业	8.7	89%			10%	0.9				
	湘江鲲鹏	7.0	243%	0.007	100%						

资料来源: 上市公司公告, Wind, 超聚变, 大河立方, 中原证券研究所

国内 AI 服务器市场中，国产化芯片比例也在持续提升。根据 IDC 数据，AI 服务器的国产芯片占比从 2023 年上半年的 10%，提升到下半年的 17%，2024 年上半年这一比例又进一步提升到了 20%。

图 12: 2023H1-2024H1 我国 AI 服务器芯片国产化占比



资料来源：IDC，中原证券研究所

4. 总结与预判

从当前形势来看，AI 作为当下科技布局最重要的发展方向，已经成为了中美科技企业争相投入的方向。我国目前在大模型领域较美国的差距在逐步缩小，同时在自动驾驶、机器人、无人机、摄像头等 AI 应用方向上形成领先优势，导致美国在 AI 相关领域对我国的制裁持续收紧，以期通过硬件限制来形成模型上的代差。

结合全球 1 月 13 日的全球 AI 管控新规来看，美国一方面在进行全球 GPU 供应方面的管控，另一方面通过代工环节限制竞争对手的供应。同时，2024 年底以来的制裁措施，以打补丁的方式，意图形成对我国更加有效的技术封锁。

我们的突破取决于禁令执行层面的严格程度和上游芯片代工、光刻机等卡脖子环节的技术突破能力。在 2022 年和 2023 年的两轮制裁后，我国国产芯片替代的能力持续提升，产业届在国产替代方面形成了较好的共识。在应对制裁的过程中，以华为为代表的厂商也形成了较好的突围能力，为后续应对持续严苛的制裁措施积累了丰富的经验。

同时，幻方给国内大模型打样之后，也为国内模型发展探索出一条发展的新路径。在 2025 年 1 月 20 日举办的专家、企业家、科教文体代表座谈会中，幻方创始人梁文锋也作为代表出席了会议，表明了中央对于幻方路径的认可和对人工智能的关注，有望在后续政府工作报告中出台更多有利于 AI 行业发展的政策。

2025 年 1 月 20 日美国新一届政府上台，我们看到新政府在与中方的各项关系方面表现得更为缓和，虽然 AI 仍将是双方后续关注的焦点，但是我们认为在制裁执行和新的制裁措施

下达方面仍具有缓和空间。

5. 风险提示

国际局势的不确定性；上游企业削减开支。

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。