

2025年 头豹行业词条报告

报告提供的任何内容(包括但不限于数据、文字、图表、图像等)均系头豹研究院独有的高度机密性文件(在报告中另行标明出处者除外)。未经头豹研究院事先书面许可,任何人不得以任何方式擅自复制、再造、传播、出版、引用、改编、汇编本报告内容,若有违反上述约定的行为发生,头豹研究院保留采取法律措施、追究相关人员责任的权利。头豹研究院开展的所有商业活动均使用"头豹研究院"或"头豹"的商号、商标,头豹研究院无任何前述名称之外的其他分支机构,也未授权或聘用其他任何第三方代表头豹研究院开展商业活动。

中央处理器 川 性能跃迁与行业应用的深度剖析 头豹词条报告系列



赵歆禹

2025-01-14 ◇未经平台授权,禁止转载

摘要 GPU行业专注于研发与生产图形处理芯片,加速图像创建与精准输出,NVIDIA与AMD领航市场。技术壁垒高,需跨领域积累深厚经验,持续创新是关键。政策推动显著, 如中英政府强化AI研究合作,推动国产GPU企业发展。市场需求旺盛,微型计算机需求增长及AI算力井喷带动市场规模扩张。未来,桌面端需求平稳增长,算力端需求将维 持井喷态势,推动GPU行业市场规模持续增长。

|行业定义

GPU行业,即图形处理单元制造领域,专注于研发与生产专用的图形处理芯片(GPU)。GPU作为现代计算体系的核心部件,其核心职能在 于加速图像在帧缓冲区中的创建,并迅速将图像数据输出至显示设备,确保视觉信息的精准呈现。在个人电脑、工作站、游戏主机及便携设备等 多种平台上,GPU均扮演着至关重要的角色。GPU行业深耕GPU研发,加速图像创建与精准输出。NVIDIA与AMD两大巨头领航市场,产品广泛 覆盖娱乐、专业及并行计算领域。凭借卓越并行处理能力,GPU在AI、HPC等大规模计算任务中表现突出。

| 行业分类

按照产品用途的分类方式,中央处理器GPU行业可以分为如下类别:

独立GPU

独立GPU适用于图形密集型场景,此类GPU配备了独立的显存,确保图形处理任务的高效执行与数据处理的独立性。市场上的知名产品,如英伟 达的GeForce RTX系列与AMD的Radeon系列,均属于独立GPU的范畴,它们在图形渲染、游戏性能及专业图形设计领域展现出了卓越的实力。

集成GPU

集成GPU以其成本效益、低能耗以及较小的物理尺寸,在入门级计算设备、便携式设备及某些特定的嵌入式应用中仍占有一席之地。例如,英伟 达的GeForce 6100系列与AMD-Ati的X1250系列,便是集成GPU领域的典型代表。

独立GPU

独立GPU适用于图形密集型场景,此类GPU配备了独立的显存,确保图形处理任务的高效执行与数据处理的独立性。市场上的知名产品, 如英伟达的GeForce RTX系列与AMD的Radeon系列,均属于独立GPU的范畴,它们在图形渲染、游戏性能及专业图形设计领域展现出了 卓越的实力。

集成GPU

集成GPU以其成本效益、低能耗以及较小的物理尺寸,在入门级计算设备、便携式设备及某些特定的嵌入式应用中仍占有一席之地。例 如,英伟达的GeForce 6100系列与AMD-Ati的X1250系列,便是集成GPU领域的典型代表。

| 行业特征

中央处理器GPU的行业特征包括:技术壁垒高,创新引领发展;政策推动显著,行业发展迅速;市场需求旺盛、多元化应用驱动增长

☑ 横跨多技术领域,更新频率快,造就高技术壁垒

GPU行业技术壁垒高,开发挑战大。成功开发出顶尖图形GPU需跨越多技术领域,积累深厚流片经验,精通底层技术、算法架构,并具备 跨平台整合能力。同时,行业发展的关键是持续推动软硬件创新,适应市场需求与技术演进。英伟达GPU架构每两年就要对硬件升级改

进,其研发成本从2018年的28亿美元上升到2024年的86亿美元。此般巨额的研发投入显现出GPU行业对于研发的巨大需求。因此,GPU 行业的技术壁垒不仅在于技术积累,更在于持续技术创新与突破所需的资源要求。

2 资源投入巨大,政策推动为行业主导

GPU行业发展显著受政策驱动。英国政府计划斥资1亿英镑采购高性能AI芯片,强化AI研究能力,深化与英伟达等全球科技巨头的合作。中国政府亦高度重视CPU、GPU等关键产品研发,其中国产GPU企业景嘉微电子在政策的扶持下积极推进无锡项目落地,预计带来产值超50亿元。由此,政策支持成为GPU行业发展的重要推动力,为企业提供土地、资金、人才等便利支持。

3 市场需求旺盛、多元化应用驱动增长

GPU产业市场需求极为旺盛,应用场景日益多元化。根据工信部数据,2022年国内微型计算机产量高达4亿台,且从2019年至2022年,国内微型计算机产量同比增速大部分时间保持正增长,直接带动了桌面端GPU的增长。同时,在人工智能领域,大模型的迭代如GPT-4的参数从GPT-3的1,750亿翻了10倍至17,600亿,需要大量的GPU芯片进行推理训练。目前,国内的大模型如阿里云通义千问也在不断推进迭代升级,对GPU的推理训练能力提出了更高要求。这些多元化的应用需求共同推动了GPU产业的持续增长。

|发展历程

阶段1:萌芽期 (1970-1999年)

80年代,Commodore Amiga电脑引领映像显示功能,IBM 8514图形系统推动2D显示技术发展。1999年,NVIDIA发布GeForce 256。这一时期各大厂致力于开发能实现GPU功能的芯片,NVIDIA发布GeForce 256,正式定义GPU概念标志着GPU的正式诞生。

阶段2:启动期 (2000-2015年)

NVIDIA与ATI(后被AMD收购)两大巨头开始角逐市场,NVIDIA发布GeForce 3,而ATI则推出了Direct3D 9.0,两者共同推动了GPU技术的不断进步。2011年,NVIDIA发布TESLA GPU。2014年4月,景嘉微成功研发出国内首款高性能、低功耗GPU芯片—JM5400。这标志着GPU在通用计算及超级计算机领域的崛起,逐渐取代CPU成为主角,同时中国国产原创GPU开始起步。

阶段3: 高速发展期 (2016-至今)

AI技术推动算力芯片崛起,NVIDIA与AMD快速迭代算力GPU,NVIDIA从P系列发展至GB200,并规划GB300及后续系列;AMD则推出MI6、MI8、MI25至MI300系列;2019年,华为发布910系列算力GPU;2020年,寒武纪上市;2023年10月美国对中国算力GPU发起制裁。AI迅速发展,算力芯片登上历史舞台,GPU相较于CPU的显著优势使之成为了算力芯片中的主角。以NVIDIA、AMD、GOOGLE为代表的大厂迅速迭代其算力芯片产品,算力GPU为GPU行业带来数十倍增长空间;华为910B芯片成为当之无愧的国产算力GPU领军产品,寒武纪则成为国产AI算力GPU上市领军公司,来自美国的国际制裁为国产算力GPU公司承接NVIDIA、AMD、GOOGLE等公司在国内的市场份额提供了机遇。

萌芽期・1970-01-01~1999-01-01

80年代,Commodore Amiga电脑引领映像显示功能,IBM 8514图形系统推动2D显示技术发展。1999年,NVIDIA发布GeForce 256。

这一时期各大厂致力于开发能实现GPU功能的芯片,NVIDIA发布GeForce 256,正式定义GPU概念标志着GPU的正式诞生。

启动期 • 2000-01-01~2015-01-01

NVIDIA与ATI(后被AMD收购)两大巨头开始角逐市场,NVIDIA发布GeForce 3,而ATI则推出了Direct3D 9.0,两者共同推动了GPU技术的不断进步。2011年,NVIDIA发布TESLA GPU。2014年4月,景嘉微成功研发出国内首款高性能、低功耗GPU芯片—JM5400

标志着GPU在通用计算及超级计算机领域的崛起,逐渐取代CPU成为主角,同时中国国产原创GPU开始起步。

高速发展期 • 2016-01-01~2024-01-01

AI技术推动算力芯片崛起,NVIDIA与AMD快速迭代算力GPU,NVIDIA从P系列发展至GB200,并规划GB300及后续系列;AMD则推出MI6、MI8、MI25至MI300系列;2019年,华为发布910系列算力GPU;2020年,寒武纪上市;2023年10月美国对中国算力GPU发起制裁。

AI迅速发展,算力芯片登上历史舞台,GPU相较于CPU的显著优势使之成为了算力芯片中的主角。以NVIDIA、AMD、GOOGLE为代表的大厂迅速迭代其算力芯片产品,算力GPU为GPU行业带来数十倍增长空间;华为910B芯片成为当之无愧的国产算力GPU领军产品,寒武纪则成为国产AI算力GPU上市领军公司,来自美国的国际制裁为国产算力GPU公司承接NVIDIA、AMD、GOOGLE等公司在国内的市场份额提供了机遇。

|产业链分析

中央处理器GPU发展现状

GPU行业产业链上游为晶圆、光刻机及EDA/IP提供商,中游为GPU设计制造提供商,下游为互联网、游戏、消费电子、智能汽车等领域。 其中上游主要为半导体硬件生产加工商,为中游提供晶圆、光刻机及EDA/IP等关键材料与技术支持,确保GPU设计制造的高质量和效率。中游则 反馈市场需求、技术趋势及产品质量性能评估,引导上游进行针对性研发与改进。同时中游根据下游需求,提供定制化GPU芯片及解决方案,并 辅以技术支持与服务。下游互联网、游戏等厂商向中游GPU设计制造商反馈市场需求变化及产品应用效果,助力中游优化GPU产品设计、提升产 品性能,共同推动GPU产业的创新与发展。

中央处理器GPU行业产业链主要有以下核心研究观点:

上游晶圆制造产能有望复苏,光刻机供应紧张将有所缓解。

2023年全球晶圆出货量下降14%,营业额下降10%,预计2024年全球硅晶圆的出货量将缩减至2%,总量将达到12,174百万平方英寸。后续随着晶圆需求逐渐从当前的下行周期中复苏,预计2025年全球晶圆出货量将强劲反弹,增幅预计将达到10%,总量有望达到13,328百万平方英寸。 ASML光刻机全球范围内供不应求,订单量巨大。由于中国市场订单数量激增导致ASML无法在短时间内满足所有客户需求,2023年ASML的光刻机分完成率低于50%。预计到2024年ASML的订单完成率将会有所改善,届时有望更好应对市场需求。

全球GPU设计公司英伟达一家独大,中国市场华为有望领跑。

国际市场上,英伟达过去几年GPU出货量市占率超90%,形成一家独大局面,而占据第二的AMD市占率不到10%。伴随国际制裁背景下,国内竞争格局也在逐渐重构。中国市场上,目前华为的昇腾系列GPU已经处于国内领先地位,昇腾910基于华为自主研发的达芬奇架构设计,INT8算力达640TOPS,单卡算力与英伟达A100相当。此外昇腾芯片还支持广泛的软件生态,专为昇腾AI处理器优化搭建的CANN深度学习平台可以最大限度地发挥出昇腾芯片的计算潜力。华为昇腾算力芯片凭借先进的架构设计、卓越的性能表现、高效的能耗管理和丰富的软件生态系统,有望领跑中国GPU算力芯片市场。

GPU需求旺盛, AI大模型和智能汽车成为GPU需求新增长点。

GPU设计商下游多元应用需求不断提升,推动GPU需求持续增长,其中AI大模型和智能汽车提供了新的增长点。大模型的迭代升级需要大量GPU 芯片进行推理训练,以Meta Llama 3.1-405B为例,参数约4050亿,使用1.6万张NVIDIA H100 GPU,在15.6T token数据上预训练需54天,期间 经历466次工作中断。国内大模型厂商众多,大模型领域的GPU需求依旧庞大,如阿里云通义大模型从2023年4月的1.0版升级到2024年5月的2.5版,模型性能在中文语料理解方面逐步追上GPT-4 Turbo,但后续迭代升级仍有空间。英伟达智能驾驶和专业设计GPU收入从2019年的7亿和12亿美元增至2023年的10亿和15亿美元。智能汽车对图像处理要求很高,对高端GPU的需求也在不断增加。

申央处理器GPU产业链上游分析 ● 中央处理器GPU产业



产业链上游分析

<u>晶圆制造产能出现波动,预计未来产能或将有效增长。</u>

晶圆制造的产能波动对中游GPU设计公司产生了显著的影响。2023年全球晶圆的出货量出现了明显的下降,降幅达到了14%,总量降至120亿平方英寸。晶圆的营业额也出现了10%的下降,降至123亿美元。与前几年相比,2019年的出货量为118亿平方英寸,而2022年的出货量则为147亿平方英寸,显示出了一定的波动性。预计2024年全球硅晶圆的出货量将有所下降,降幅预计为2%,总量将达到12,174百万平方英寸。随着晶圆需求逐渐从当前的下行周期中复苏,未来全球晶圆出货量有望回暖突破。

光刻机供货紧张状态有望缓解,中国EDA行业国产替代发展迅速。

GPU上游的核心在于半导体制造,其中光刻机和EDA领域至关重要。光刻机不仅是精确绘制芯片电路图案的核心工具,更直接决定芯片的性能、集成度及良率。ASML作为光刻机领域行业龙头,全球范围内光刻机订单量巨大。2023年ASML的光刻机订单完成率低于50%,中国市场的订单数量激增导致ASML无法在短时间内满足所有客户需求。2024年ASML的订单完成率有所改善,有望更好应对市场需求。EDA(电子设计自动化)软件同样是半导体产业链中不可或缺的一部分,涵盖了从概念到成品整个过程的设计、验证和优化。随着技术封锁的压力增大,中国对于EDA工具的自主可控需求日益迫切,促使国产EDA替代进程的加速发展。据中国半导体行业协会预测,2025年中国的EDA市场规模将达到184.9亿元人民币(约合25亿美元),届时将占全球EDA市场的18.1%。

🌵 中央处理器GPU产业链中游分析

品牌端

GPU设计制造提供商

中游厂商

図 深圳市海思半导体有限公司
 図 上海壁の科技股份有限公司
 図 上海壁の科技股份有限公司
 図 上海壁の科技股份有限公司
 図 上海水芯集成电路股份有限公司
 図 上海水芯集成电路股份有限公司
 図 上海水芯集成电路股份有限公司
 図 上海水芯集成电路股份有限公司
 図 産尔线程智能科技(北京)有限责任公司

产业链中游分析

中游GPU设计公司研发投入巨大,对研发人才需求极大。

中游GPU设计公司研发投入巨大,且需要大量高端研发人才。中科寒武纪2023年的研发投入达到了11.18亿元,占其营业收入的比例高达157.53%,目前公司研发团队有752人,占员工总人数的75.28%,78.46%以上研发技术人员拥有硕士及以上学历;海光信息的研发投入为28.10亿元,同比增长35.93%,占营业收入的46.74%,海光拥有1,641名研发技术人员,占员工总数的91.68%。GPU设计公司研发投入大是由技术进步、市场需求变化及竞争对手挑战等多方面压力所致。GPU芯片对图形渲染和计算任务,以及接口速率、带宽、存储、封装技术等方面都有极高的要求,给设计公司带来极大挑战。同时下游应用端对于GPU芯片性能要求的不断提高也促使GPU设计公司加大研发投入力度,优化现有微架构,探索新技术和方法,从而实现更高的GPU并行处理能力和效率。

国际竞争格局一家独大,中国市场上华为有望领跑。

国际市场上,英伟达过去几年GPU出货量市占率超90%,形成一家独大局面,而占据第二的AMD市占率不到10%。伴随国际制裁背景下,国内竞争格局也在逐渐重构。中国市场上,目前华为的昇腾系列GPU已经处于国内领先地位,昇腾910基于华为自主研发的达芬奇架构设计,INT8算力达640TOPS,单卡算力与英伟达A100相当。此外昇腾芯片还支持广泛的软件生态,专为昇腾AI处理器优化搭建的CANN深度学习平台可以最大限度地发挥出昇腾芯片的计算潜力。华为昇腾算力芯片凭借先进的架构设计、卓越的性能表现、高效的能耗管理和丰富的软件生态系统,有望领跑中国GPU算力芯片市场。

◎ 中央处理器GPU产业链下游分析

渠道端及终端客户 互联网、游戏、消费电子、智能汽车等领域 渠道端 ☑ 互联网企业(图像处理,AI大模型,云计算需求) ◎ 消费电子行业(智能手机,电脑,ARAVR) ◎ 游戏开发(图形处理,物理运算)

产业链下游分析

下游多元应用需求提升,推动GPU需求持续增长。

英伟达游戏专业GPU应用收入从2019年的55亿美元增至2023年的104亿美元,游戏行业不断发展也需要更多的高端桌面GPU。全球人工智能硬件市场(服务器)规模预计从2022年的195亿美元增长到2026年的347亿美元,年复合增长率17.3%;2023年中国人工智能服务器出货量超30万台,2027年预期超70万台,年复合增长率预计为21.6%。GPU在人工智能服务器中通过其强大的并行计算能力、高内存带宽、能效比和灵活性,成为加速AI计算的关键组件,随着服务器出货量增加,对GPU的需求也将不断增加。

<u>AI大模型、智能汽车成为GPU新兴应用领域,市场前景广阔。</u>

AI大模型、智能汽车是目前GPU的新兴应用领域。AI大模型的迭代升级需要大量GPU芯片进行推理训练,以Meta Llama 3.1-405B为例,参数约4050亿,使用1.6万张NVIDIA H100 GPU,在15.6T token数据上预训练需54天,期间经历466次工作中断。国内大模型厂

商众多,大模型领域的GPU需求依旧庞大,如阿里云通义大模型从2023年4月的1.0版升级到2024年5月的2.5版,模型性能在中文语料理解方面逐步追上GPT-4 Turbo,但后续迭代升级仍有空间。智能汽车方面,英伟达智能驾驶和专业设计GPU收入从2019年的7亿和12亿美元增至2023年的10亿和15亿美元。智能汽车对图像处理要求很高,对高端GPU的需求也在不断增加。英伟达智能驾驶和专业设计GPU收入从2019年的7亿和12亿美元增至2023年的10亿和15亿美元。智能汽车对图像处理要求很高,对高端GPU的需求不断增加。中国厂商智能汽车收入也在大幅增长,华为2023年智能汽车收入达47.37亿元人民币,同比增长128.1%。

|行业规模

中央处理器GPU行业规模的概况

2018年—2023年,中央处理器GPU行业市场规模由952.61亿元增长至4,329.30亿元,期间年复合增长率35.36%。预计2024年—2028年,中央处理器GPU行业市场规模由6,451.15亿元增长至31,556.67亿元,期间年复合增长率48.72%。

中央处理器GPU行业市场规模历史变化的原因如下:

23年以前微型计算机需求增长拉动GPU市场规模扩张

根据工信部数据,2022年中国微型计算机产量约为4.34亿台,2019-2022年同比增速分别为13.2%、42.3%、22.3%、-8.3%,微型计算机设备产量增加需要更多的GPU,带动了桌面端GPU厂商收入增长。数据显示,全球桌面PC和笔记本电脑出货量在2019-2021年间一直保持较快增长速度,最高季度出货量高达9,000万台,智能设备对GPU的需求较为刚性。

23年AI算力需求井喷式增长,带动GPU市场规模急速扩张

2023年AI算力井喷推动了GPU市场规模急速扩张。根据OpenAI官网数据,GPT-4模型的参数量从GPT-3的1,750亿激增至17,600亿,而中国市场以阿里云推出的通义千问为例,2023年10月推出的通义2.0版本参数量仅达到千亿级别,相较于GPT-4仍存在一定差距,后续迭代无疑将继续维持甚至增加对于高性能GPU的需求。GPU供应商方面,英伟达2023年度数据中心收入达到了474.05亿美元,同比暴增215%,占据了英伟达总收入的77.8%,而2022年这一比例仅有55.9%,大模型对算力芯片需求的急剧增加带动了GPU厂商收入的急剧增加。

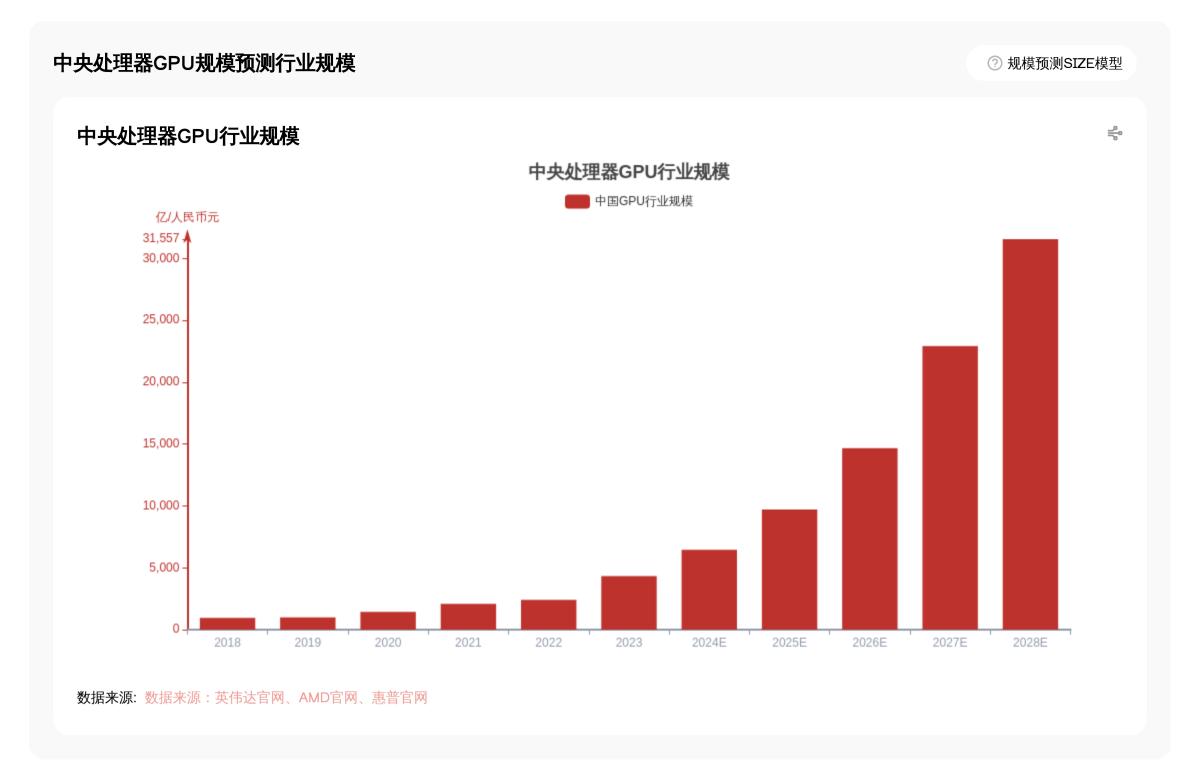
中央处理器GPU行业市场规模未来变化的原因主要包括:

各大厂均桌面端需求保持增长

中国桌面端需求稳中有进,特别是在大型桌游市场中表现得尤为明显。近期国产3A游戏《黑神话:悟空》在Steam平台上的销量已达到2000万份,游戏总收入超过了9.61亿美元(约合人民币67.9亿元),与此同时,对高性能GPU电脑设备有需求的客户端游戏整体市场份额也持续提升,2023年、2024上半年分别同比增长8%、2.56%。高端桌面游戏需求的不断增加使得用户对高性能GPU电脑设备的需求不断增加。

未来中国算力需求仍将维持井喷态势

从政策端来看,中国对算力发展有明确规划,根据2023年工信部等六部门发布的《算力基础设施高质量发展行动计划》,预计2025年算力规模将超过300EFlops,其中智能算力占比将达到35%;从需求端来看,国内互联网厂商AI资本支出方面,阿里巴巴、腾讯和百度上半年的资本支出合计为500亿元人民币(70亿美元),而去年同期为230亿元人民币,这些互联网厂商表示,资本支出重点方向是购买与为AI的大型语言模型训练提供支持相关的处理器和基础设施,包括自己公司的AI大模型和其他模型。此外,国内外算力芯片大厂也对算力需求进行了展望,华为预计到2030年,全球通用算力需求将达到3.3ZFLOPS,是2020年的10倍;全球AI算力需求将达到105ZFLOPS,是2020年的500倍;英伟达预计AI芯片和系统规模将达3,000亿美元;AMD预计2027年AI加速器市场达4,000亿美元。



|政策梳理

	政策名称	颁布主体	生效日期	影响
	《工业和信息化部等七部门关于推动未来产业创新发展的实施意见》	工信部等七部门	2024-01-01	8
政策内容	做优信息服务产品。发展下一代操作系统,构筑安全可靠的数字底座。推广开源技术,建设开源社区,构建开源生态体系。探索以区块链为核心技术、以数据为关键 要素,构建下一代互联网创新应用和数字化生态。面向新一代移动信息网络、类脑智能等加快软件产品研发,鼓励新产品示范应用,激发信息服务潜能。			
政策解读	该政策通过推动下一代操作系统发展、构筑安全可靠的数字基础、推广开源技术及建设开源社区等措施,为国产GPU及相关信息技术产业营造了良好的生态环境。 特别是在鼓励面向新一代移动信息网络、类脑智能等领域的软件产品研发与应用方面,不仅有助于加速国产GPU技术的创新和市场化进程,还促进了上下游产业链 的协同发展,提升了整个信息技术服务业的自主可控能力和国际竞争力。这对于打破国外技术垄断、保障国家安全以及推动经济高质量发展具有重要意义。			
政策性质	鼓励性政策			

	政策名称	颁布主体	生效日期	影响		
	《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					
政策内容	到2025年底,普惠易用、绿色安全的综合算力基础设施体系初步成型,东西部算力协同调度机制逐步完善,通用算力、智能算力、超级算力等多元算力加速集聚, 国家枢纽节点地区各类新增算力占全国新增算力的60%以上,国家枢纽节点算力资源使用率显著超过全国平均水平。					
政策解读	该政策通过构建普惠易用、绿色安全的综合算力基础设施体系,促进东西部算力资源的优化配置与高效协同,对国产GPU的发展具有重要推动作用。政策目标明确,至2025年底,国家枢纽节点地区的算力资源将占据全国新增算力的大部分份额,这不仅为国产GPU提供了广阔的市场空间,也加速了其在高性能计算、人工智能等领域的应用与技术创新,有助于提升国产GPU的市场竞争力和国际影响力。					
政策性质	鼓励性政策					

	政策名称 颁布主体		生效日期	影响 10
	《算力基础设施高质量发展行动计 划》	工信部等六部门	2023-10-01	10
政策内容	预计2025年,在计算力方面,算力规模将超过300EFlops,其中智能算力占比将达到35%(105EFlops)。重点在西部算力枢纽及人工智能发展基础较好地区集约开 展智算中心建设,逐步合理提升智能算力占比。			
政策解读	该政策从全国角度为算力基础设施发展指明了方向,而GPU为代表的算力芯片建设是算力基础设施建设中的关键一环,有助于明确中国算力GPU行业政策力度和市 场规模。			
政策性质	鼓励性政策			

	政策名称	颁布主体	生效日期	影响	
	《数字中国建设整体布局规划》	国务院	2023-02-01	8	
政策内容	系统优化算力基础设施布局,促进东西部算力高效互补和协同联动,引导通用数据中心、超算中心、智能计算中心、边缘数据中心等合理梯次布局。整体提升应用基础设施水平,加强传统基础设施数字化、智能化改造。				
政策解读	该政策通过优化算力基础设施布局,促进东西部算力资源的高效互补与协同联动,对国产GPU的发展具有重要意义。它不仅为国产GPU提供了更为广阔的应用场景和市场机会,还通过加强传统基础设施的数字化、智能化改造,推动了对高性能计算和智能计算需求的增长,从而加速了国产GPU在技术创新和市场拓展方面的步伐,增强了国内企业在国际竞争中的地位。				
政策性质	鼓励性政策				

	政策名称 颁布主体		生效日期 影响	
	《国务院关于印发新时期促进集成 电路产业和软件产业高质量发展若 干政策的通知》	国务院	2020-07-01	9
政策内容	线宽小于28纳米(含),经营15年以上,1-10年免征企业所得税。线宽小于65纳米(含),经营15年以上,1-5年免征所得税,6-10年按25%征税。线宽小于130纳 米(含),经营10年以上,1-2年免征所得税,3-5年按25%征税。线宽小于130纳米(含)的企业纳税年度亏损准予结转,不超过10年。			
政策解读	显著降低了企业的财务负担,激发创新活力与投资热情,有助于加速国产GPU技术的研发进程,提升产品竞争力,促进整个半导体产业链的发展完善。			
政策性质	鼓励性政策			

| 竞争格局

中央处理器GPU概况

中国中央处理器GPU竞争格局目前呈现英伟达一家独大,华为逐步追上的态势。

行业主要分为三个梯队,第一梯队是英伟达和华为,华为产品逐渐超越英伟达中国定制版芯片;第二梯队包括寒武纪、海光、燧原、壁仞,这些公司产品参数、良率、盈利能力都弱于英伟达和华为,但都各有特色,如寒武纪已上市且算力参数良好,海光已上市且产品适配军用,燧原、壁仞已谋划上市且陆续有产品迭代推出;第三梯队包括天数智芯、摩尔现场、沐曦、登临、阿里平头哥等公司,这些公司的芯片仅停留于发布会,未有大厂等客户表示有相关送样检测等信息流出。

中央处理器GPU行业竞争格局的形成主要包括以下原因:

英伟达GPU产品技术领先,在国际制裁背景下仍能推出中国市场专供芯片。

英伟达不断推出革命性的GPU架构,其AI芯片Blackwell GPU物理尺寸上拥有当前最大的GPU面积,集成了高达2,080亿个晶体管,AI性能达到了20petaflops(20,000TOPS),与前代产品相比性能显著提升。此外英伟达的CUDA库已成为主流深度学习中GPU并行加速的基础,其芯片在视觉处理方面的能力尤为突出。国际制裁背景下,2023年初英伟达针对中国市场推出专供芯片H800。即使在算力受限的情况下,H800在INT8下的算力仍能达到3,958TOPS,远超华为昇腾910B的640TOPS,此时英伟达算力GPU在中国市场依旧保持着压倒性的优势。

国家政策支持背景下,华为等芯片厂商持续迭代发展算力GPU。

2023年工信部等六部门印发《算力基础设施高质量发展行动计划》,目标2025年算力规模超300EFLOPS,智能算力占比达35%,政策上还在持续加大对算力企业的金融支持。中国算力GPU发展方面,2023年11月H800受限后,英伟达转供H20芯片,华为910B算力达640TOPS,超过H20的296TOPS,有望打破英伟达算力GPU在中国市场的主导地位。

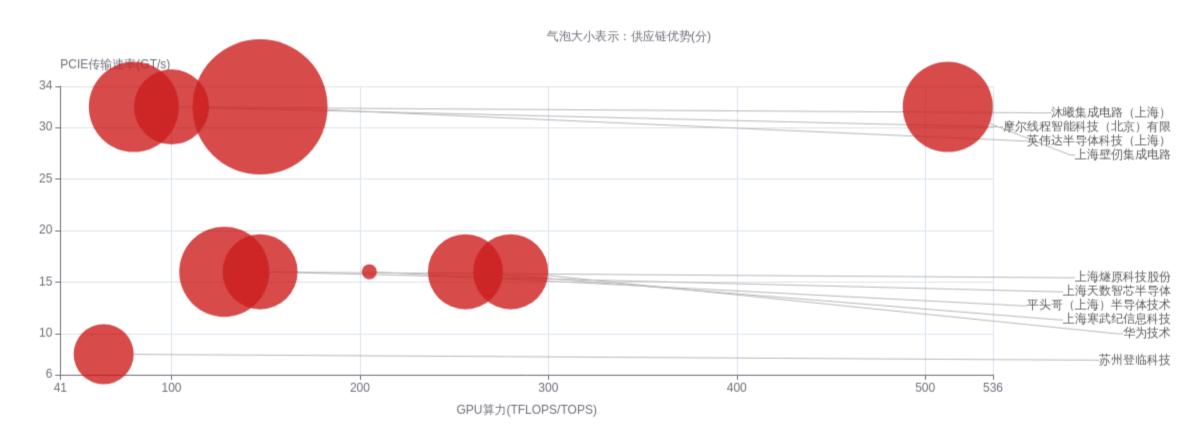
中央处理器GPU行业竞争格局的变化主要有以下几方面原因:

由于国际制裁,英伟达只能向中国出口H20,这一芯片性能已经被华为超越,后续陆续将被其他国内厂商超越。

2023年10月,美国出口管理条例提出规定,禁止向中国出口高算力芯片,H800被禁,英伟达只能提供H20芯片。目前H20已被华为昇腾910B超越,华为昇腾910B在INT8下算力达640TOPS,超过H20的296TOPS。华为910C已于2024年下半年小规模送往英伟达客户测试,预计将与H100正面竞争,进一步扩大中国市场优势。寒武纪芯片已获阿里等互联网大厂测试,性能得到认可,海光传闻也送到大厂进行测试但出现流片情况。若后续禁令得不到放松,H20还将陆续被其他国内厂商超越。

<u>中国半导体行业协会出台最新芯片指导意见,桌面端GPU竞争格局有望重构。</u>

2024年12月3日,中国半导体行业协会出台最新芯片指导意见,认为美国芯片产品不再安全、不再可靠,中国相关行业将不得不谨慎采购美国芯片。由于英伟达算力端GPU H20已被华为超越,故此这一政策主要利好桌面端GPU竞争格局重构。国产桌面端GPU领先厂商兆芯的产品已适配联想开天、同方、紫光等多款桌面整机,有望成为国产桌面端GPU的"华为";军工&通用两手抓的景嘉微拥有多款自研GPU,全面支持国产CPU、操作系统和固件,可广泛应用于PC、服务器和图形工作站等计算机设备,满足地理信息系统、三维测绘、三维制图、媒体处理、辅助设计、显示渲染等高性能显示需求的同时也能有效保障信息数据的安全。



上市公司速览



|企业分析

4 华为技术有限公司

■ 公司信息				
企业状态	存续	注册资本	4084113.182万人民币	
企业总部	深圳市	行业	计算机、通信和其他电子设备制造业	
法人	赵明路	统一社 会 信用 代码	914403001922038216	
企业类型	有限责任公司(自然人投资或控股的法人独资)	成立时间	1987-09-15	
品牌名称	华为技术有限公司	经营范围	程控交换机、传输设备、数据通信设备、宽带多媒体设备、电源、无线通信设备、微电子产品、软件、系统集成工程、计算机及配套设备、终端设备及相关通信信息产品、数据中心机房基础设施及配套产品(含供配电、空调制冷设备、智能管理监控等)的开发、生产、销售、技术服务、工程安装、维修、咨询、代理、租赁;信息系统设计、集成、运行维护;集成电路设计、研发;统一通信及协作类产品,服务器及配套软硬件产品,存储设备及相关软件的研发、生产、销售;无线数据产品(不含限制项目)的研发、生产、销售;无线数据产品(不含限制项目)的研发、生产、销售;通信站点机房基础设施及通信配套设备(含通信站点、通信机房、通信电源、机柜、天线、通信线缆、配电、智能管理监控、锂电及储能系统等)的研发、生产、销售;能源科学技术研究及能源相关产品的研发、生产、销售;大数据产品、物联网及通信相关领域产品的研发、生产、销售;汽车零部件及智能系统的研发、生产、销售及服务;建筑工程;设计、制作、发布、代理各类广告;通信设备租赁(不含限制项目);企业管理咨询(不含限制项目);进出口业务;国内商业、物资供销业业务(不含专营、专控、专卖商品);对外经济技术合作业务;房屋租赁业务(持许可经营证);以及其他法律法规不禁止的经营活动(依法须经批准的项目,经相关部门批准后方可开展经营活动)。^增值电信业务经营。	

华为技术有限公司竞争优势

• 竞争优势

| 附录

法律声明

^

权利归属:头豹上关于页面内容的补充说明、描述,以及其中包含的头豹标识、版面设计、排版方式、文本、图片、图形等,相关知识产权归头豹所有,均受著作权法、商标法及 其它法律保护。

尊重原创:头豹上发布的内容(包括但不限于页面中呈现的数据、文字、图表、图像等),著作权均归发布者所有。头豹有权但无义务对用户发布的内容进行审核,有权根据相关证据结合法律法规对侵权信息进行处理。头豹不对发布者发布内容的知识产权权属进行保证,并且尊重权利人的知识产权及其他合法权益。如果权利人认为头豹平台上发布者发布的内容侵犯自身的知识产权及其他合法权益,可依法向头豹(联系邮箱:support@leadleo.com)发出书面说明,并应提供具有证明效力的证据材料。头豹在书面审核相关材料后,有权根据《中华人民共和国侵权责任法》等法律法规删除相关内容,并依法保留相关数据。

内容使用:未经发布方及头豹事先书面许可,任何人不得以任何方式直接或间接地复制、再造、传播、出版、引用、改编、汇编上述内容,或用于任何商业目的。任何第三方如需转载、引用或基于任何商业目的使用本页面上的任何内容(包括但不限于数据、文字、图表、图像等),可根据页面相关的指引进行授权操作;或联系头豹取得相应授权,联系邮箱:support@leadleo.com。

合作维权:头豹已获得发布方的授权,如果任何第三方侵犯了发布方相关的权利,发布方或将授权头豹或其指定的代理人代表头豹自身或发布方对该第三方提出警告、投诉、发起 诉讼、进行上诉,或谈判和解,或在认为必要的情况下参与共同维权。

完整性:以上声明和本页内容以及本平台所有内容(包括但不限于文字、图片、图表、视频、数据)构成不可分割的部分,在未详细阅读并认可本声明所有条款的前提下,请勿对本页面以及头豹所有内容做任何形式的浏览、点击、引用或下载。

■商务合作

阅读全部原创报告和百万数据

会员账号

募投可研、尽调、IRPR等研究咨询

定制报告/词条

定制公司的第一本

白皮书

内容授权商用、上市

招股书引用

企业产品宣传

市场地位确认

丰富简历履历,报名

云实习课程

头豹研究院

咨询/合作

网址: www.leadleo.com

电话: 13544093853 (游先生)

电话: 13080197867 (李先生)

深圳市华润置地大厦E座4105室



