

行业投资评级

强于大市|维持

行业基本情况

收盘点位	4656.97
52周最高	5592.36
52周最低	3303.08

行业相对指数表现（相对值）



资料来源：聚源，中邮证券研究所

研究所

分析师:孙业亮
SAC 登记编号:S1340522110002
Email:sunyeliant@cnpsec.com
研究助理:丁子惠
SAC 登记编号:S1340123020006
Email:dingzihui@cnpsec.com
研究助理:常雨婷
SAC 登记编号:S1340123040019
Email:changyuting@cnpsec.com

近期研究报告

《通用人工智能迎政策利好，产业发展步入新纪元》 - 2023.05.16

国产 AI 芯片的创业裂变

● AICG 拉动算力需求，国产芯片把握机遇乘风而上

AI 进入“大模型”时代，算力需求陡升。自 2018 年 GPT-1.0 模型首次发布以来，OpenAI 不断迭代模型，近期发布了 GPT-4.0 模型，它拥有更大的参数量、更长的迭代时间和更高的准确性。随着数据不断增长和算法复杂度提高，AI 模型迭代推动算力需求爆发。

英伟达业绩超预期，再次强化 AI 算力逻辑。5 月 24 日英伟达发布 2024 一季报财报，其中数据中心营收创下历史新高，同比+14%/环比+18%，主要原因为 GPU 需求随着 AI 发展水涨船高。

中国 AI 芯片市场呈现显著增长动力，众多企业“小试牛刀”。据数据显示，2021 年我国 AI 芯片市场规模达到 427 亿元，同比增长 124%。预计到 2023 年，市场规模将进一步扩大至 1206 亿元。

● 突围国外巨头围猎，国产替代初迎曙光

国产 GPU 在性能方面与国际领先水平的差距在加速缩小。在图形渲染 GPU 方面，景嘉微的 JM9 系列图形处理芯片性能与英伟达 GeForce GTX 1050 相近。虽然景嘉微的产品与国际尖端 GPU 存在着极大的差距，但作为一颗由中国企业完全独立研发、采用正向设计、具有自主知识产权的 GPU，已经有长足的进步。

ASIC 国产替代正当时，FPGA 发展机遇大。ASIC 市场未形成垄断，国产厂商仍存在可观的机会抢占市场。另一方面，与其他 AI 芯片相比，FPGA 芯片开发门槛较高，前期需要投入大量资源，存在壁垒。中国 FPGA 厂商仍处于起步阶段，但是随着 5G、AI 产业化带来的巨大需求，国内 FPGA 企业技术有望不断实现突破，迎来快速成长。

● AI 芯片迎政策利好，把握内、外双重机遇

近年来国家高度关注人工智能芯片产业的发展，发布一系列产业支持政策，为人工智能芯片行业建立了优良的政策环境，促进行业的发展。2021 年，《“十四五”规划纲要和 2035 年远景目标纲要》指出，“十四五”期间，我国新一代人工智能产业将聚焦高端芯片等关键领域。从国家战略高度为人工智能芯片行业建立了优良的政策环境。各地方也根据各自的背景与条件，发布促进和扶持人工智能产业发展的方案方针。

内驱动力：国家教育政策加大人才培养，为 AI 行业筑底强基。在 2018-2021 年，超过 300 所高校开设了人工智能专业；部分企业也开始与高校进行合作，以产学研合作教学模式共同培养综合能力突出的优质人才。

外驱动力：资本持续进入国内 AI 芯片市场，为 AI 行业添砖加瓦。截止 2022 年 1 月，2021 年中国人工智能芯片相关领域融资事件共计 92 起，总金额约 300 亿人民币。企业方面，中国的科技巨头如百度、阿里巴巴、华为等都在人工智能芯片领域进行了大量的投资。

此外，还涌现出了许多专注于人工智能芯片研发的创业公司，如 Horizon Robotics、Cambricon 等。

● 投资建议

算力是 AI 发展的基础底座。受益于 AI 大模型的突破与发展，AI 芯片的需求随之水涨船高，我们看好在人工智能时代及国产替代主题共振下的 AI 芯片投资机遇。建议关注 (1) AI 芯片：海光信息、寒武纪、景嘉微、中科曙光等；(2) 与 AI 芯片创业公司合作：首都在线、浪潮信息、优刻得、光环新网、同方股份等。

● 风险提示

研发进度不及预期；实体清单导致的供应链风险；产品迭代不及预期；商业落地不及预期；行业竞争加剧等风险。

重点公司盈利预测与投资评级

代码	简称	投资评级	收盘价 (元)	总市值 (亿元)	EPS (元)		PE (倍)	
					2023E	2024E	2023E	2024E
688041.SH	海光信息	买入	81.00	1882.71	0.55	0.82	91.55	61.49
688256.SH	寒武纪	未评级	194.50	806.44	-1.96	-1.35	-99.16	-144.53
300474.SZ	景嘉微	未评级	79.81	377.52	0.89	1.27	93.01	65.55
300846.SZ	首都在线	未评级	14.25	66.52	0.16	0.37	89.06	38.51
000977.SZ	浪潮信息	未评级	39.92	584.31	1.82	2.27	21.93	17.58
603019.SH	中科曙光	未评级	47.20	691.01	1.38	1.78	34.28	26.48
688158.SH	优刻得	未评级	17.01	77.07	-0.52	-0.33	-32.71	-51.55
300383.SZ	光环新网	未评级	10.84	194.86	0.35	0.44	30.88	24.86
600100.SH	同方股份	未评级	7.94	266.01	-	-	-	-

资料来源：iFinD，中邮证券研究所（注：未评级公司盈利预测来自 iFinD 机构的一致预测）

目录

1 AI 芯片——人工智能的基石	6
1.1 AI 大模型推动算力需求爆发	6
1.2 高算力需求催生 AI 芯片兴起	8
1.3 GPU 为当前 AI 芯片领导者	10
2 突围国外巨头围猎，国产替代初迎曙光	14
2.1 国内外 AI 芯片差距较大，科技巨头展开布局	14
2.2 人才+资本双驱动，AI 芯片国产化有望加速	19
3 国内 AI 芯片上市公司开启国产替代征程	22
3.1 海光信息：国产 CPU+DCU 龙头	22
3.2 寒武纪：ASIC 路线先行者	23
3.3 景嘉微：军用为基，民用突破	25
4 创业公司纷纷入局，与多方建立合作	27
4.1 沐曦科技	27
4.2 燧原科技	28
4.3 壁仞科技	29
4.4 摩尔线程	31
5 风险提示	32

图表目录

图表 1: 全球人工智能支出、数字化转型支出及 GDP 增长趋势预测.....	6
图表 2: 中国智能算力规模及预测.....	7
图表 3: 中国通用算力规模及预测.....	7
图表 4: 2022 年中国 AI 算力发展城市排行.....	8
图表 5: AI 芯片技术架构.....	8
图表 6: AI 与半导体芯片的发展进程对照.....	9
图表 7: 云-边-端协同示意图.....	10
图表 8: AI 芯片设计流程与设计优势对比.....	11
图表 9: 冯·诺依曼、近存计算、存内计算架构对比.....	12
图表 10: 全球人工智能服务器 GPU、ASIC 和 FPGA 芯片搭载率.....	13
图表 11: 中国人工智能芯片市场规模占比.....	13
图表 12: 2017-2023 年中国 AI 芯片市场规模预测趋势图.....	14
图表 13: AI 芯片产业图谱.....	15
图表 14: 国内 AI 厂商成立时间表.....	16
图表 15: 2019-2022 年全球 GPU 芯片市场份额.....	16
图表 16: 2021 年全球独立 GPU 芯片市场份额.....	16
图表 17: JM9 系列性能对比.....	17
图表 18: 海光 DCU 与国际主流产品性能对比.....	17
图表 19: 全球 FPGA 芯片市场份额.....	18
图表 20: 海光与国际主流产品性能对比.....	18
图表 21: AI 相关论文发表和海外人才回国情况.....	20
图表 22: 中国人工智能技术方向岗位供需情况.....	20
图表 23: 十四五期间中国人工智能芯片产业上市公司-AI 芯片业务规划.....	20
图表 24: 中国人工智能芯片行业融资整体情况.....	22
图表 25: 中国人工智能芯片行业投融资轮次情况.....	22
图表 26: 海光与国际主流产品性能对比.....	22
图表 27: 海光与国产服务器 CPU 产品性能对比.....	23
图表 28: 寒武纪产品研发年历.....	24
图表 29: 端云一体解决方案.....	24
图表 30: 寒武纪主流产品情况.....	25
图表 31: 景嘉微发展历程.....	25
图表 32: 机载航电系统图形显控示意图.....	26
图表 33: 景嘉微 GPU 产品性能情况.....	26
图表 34: 沐曦科技合作情况.....	27
图表 35: 燧原科技合作情况.....	29

图表 36: 壁仞科技合作情况.....	30
图表 37: 摩尔线程合作情况.....	31

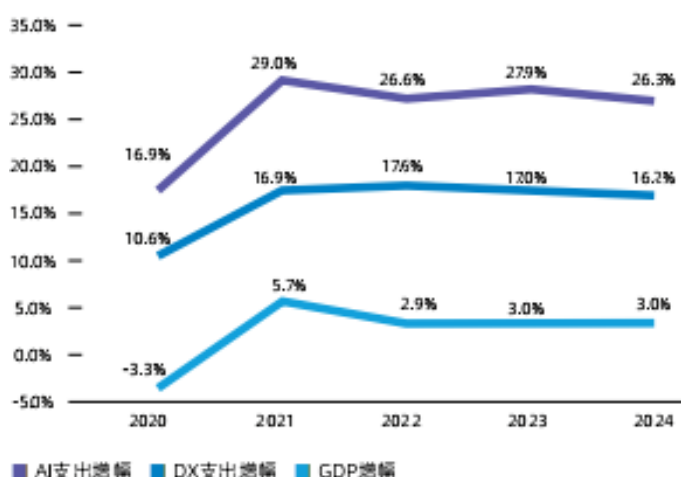
1 AI 芯片——人工智能的基石

1.1 AI 大模型推动算力需求爆发

AI 已经进入了“大模型”时代。自 2018 年 GPT-1.0 模型首次发布以来,OpenAI 不断迭代模型,最近发布了 GPT-4.0 模型,它拥有更大的参数量、更长的迭代时间和更高的准确性。随着人工智能的快速发展,算力的提升也被加速推动。人工智能非常依赖于相关基础设施,包括计算、储存和网络等。随着数据不断增长和算法复杂度提高,人工智能对算力提出了更高的要求。全球人工智能算力基础设施产业正在加速发展,为人工智能技术在更广泛场景中落地创造可能。

释放算力的价值对国家整体经济发展将发挥推动作用。算力指数每提高 1 点,数字经济和 GDP 将分别增长 3.5‰和 1.8‰。可见,国家算力指数越高,对经济的拉动作用越强。在业界,人工智能的应用产生了很多需求,其中最直接的赛道是企业数字化转型。据 IDC 统计,全球范围内,企业在人工智能(AI)市场的技术投资从 2019 年的 612.4 亿美元增长至 2021 年的 924.0 亿美元,预计到 2022 年(同比)将增长 26.6%至 1,170.0 亿美元,并有望到 2025 年突破 2,000 亿美元,增幅高于企业数字化转型(DX)支出整体增幅。

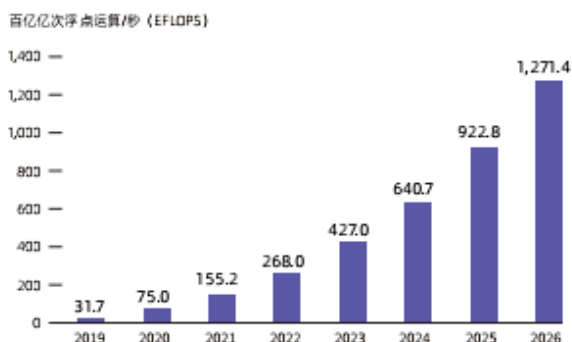
图表1: 全球人工智能支出、数字化转型支出及 GDP 增长趋势预测



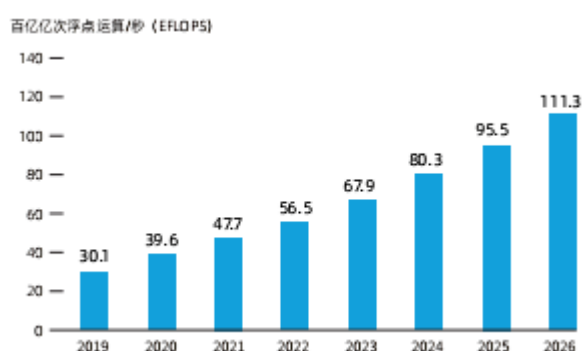
资料来源: IDC, 世界银行, 中邮证券研究所

中国的智能算力规模正在快速增长。算力是实现人工智能产业化的核心力量,其发展对人工智能技术的进步和行业应用起着决定性作用。随着人工智能向多场景化、规模化、融合化等高应用阶段发展,数据体量呈现出急剧增长态势,算法

模型的参数量也呈指数级增加，以加速计算为核心的算力中心规模将不断扩大。在科技兴国政策驱动下，人工智能在提升中国核心竞争力的重要支撑作用得以确立。结合新基建、数字经济等利好政策，中国人工智能市场稳中有进，凸显在数字经济时代中技术的力量。据 IDC 预测，到 2022 年，中国的智能算力规模将达到 268.0 EFLOPS，超过通用算力规模。预计到 2026 年，智能算力规模将进入 ZFLOPS 级别，达到 1,271.4 EFLOPS。在 2021-2026 年期间，中国的智能算力规模年复合增长率预计将达到 52.3%，而同期通用算力规模的年复合增长率为 18.5%。根据《IDC 中国服务器市场季度跟踪报告》及 CPU 双精度 (FP64) 运算能力数据，测算了中国通用算力规模。2021 年中国通用算力规模达 47.7EFLOPS，预计到 2026 年通用算力规模将达到 111.3 EFLOPS。2021-2026 年期间，预计中国智能算力规模年复合增长率达 52.3%，同期通用算力规模年复合增长率为 18.5%。

图表2：中国智能算力规模及预测


资料来源：IDC，中邮证券研究所

图表3：中国通用算力规模及预测


资料来源：IDC，中邮证券研究所

中国人工智能的发展正在向城市覆盖面更广泛的地区扩展。未来预计将涌现出更多具有城市特色的人工智能示范区，为产业发展树立标杆。根据 2022 年中国人工智能城市排行榜显示，北京排名第一，杭州、深圳、上海和广州分别位居第二到第五名，这些城市均为发展较好、经济水平高的一、二线城市，而天津、成都和南京等城市则进入了前十名。不同地区都在提升城市算力基础设施水平的同时，也推动产业智能化的发展以适应自身的发展需求。

图表4：2022 年中国 AI 算力发展城市排行



资料来源：IDC，世界银行，中邮证券研究所

1.2 高算力需求催生 AI 芯片兴起

面向人工智能领域的芯片均被称为 AI 芯片。广义而言，AI 芯片指的是专门用于处理人工智能应用中大量计算任务的模块，除了以 GPU、FPGA、ASIC 为代表的 AI 加速芯片，还有比较前沿性的研究，例如：类脑芯片、可重构通用 AI 芯片等。狭义的 AI 芯片指的是针对人工智能算法做了特殊加速设计的芯片。“无芯片不 AI”，以 AI 芯片为载体实现的算力是人工智能发展水平的重要衡量标准，发展更注重超速运算能力的 AI 芯片成为推动人工智能产业爆发的关键核心要素之一，其快速发展对人工智能技术的进步和行业应用起到了决定性的作用。

图表5：AI 芯片技术架构

产品类别	简要说明
GPU	全称为图形处理器，在计算方面具有高效的并行性，用于图像处理。GPU 芯片因海量数据并行运算能力，被最先引入深度学习。功耗比较低。

FPGA	全称为现场可编程门阵列，是一种集成大量基本门电路及存储器的芯片， 最大特点为可编程 。具有明显能耗优势，有低延时和高吞吐的特性。
ASIC	全称为专用集成电路，是 专用定制芯片 ，为实现特定要求而定制 的芯片 。在功耗、可靠性、体积方面均有优势。缺点是不能拓展应用。
类脑芯片	参考人脑神经元结构和人脑感知认知方式来设计的芯片。

资料来源：与非网，中邮证券研究所

AI 芯片的兴起源于 AI 对算力的高需求。算力是实现 AI 产业化的核心力量，随着人工智能应用场景的不断拓展和数据规模的急剧增长，算法模型的参数量呈指数级增加，特别是深度学习成为当前 AI 研究和运用的主流方式，加速计算成为不可或缺的需求。虽然 CPU 可以拿来执行 AI 算法，但因为内部有大量其他逻辑，而这些逻辑对于目前的 AI 算法来说是完全用不上的，所以，使用 CPU 并不能达到最优的性价比。因此，AI 芯片应运而生，它们被设计出来适配海量并行计算能力，进而可以加速 AI 计算。

图表6：AI 与半导体芯片的发展进程对照



资料来源：亿欧智库，中邮证券研究所

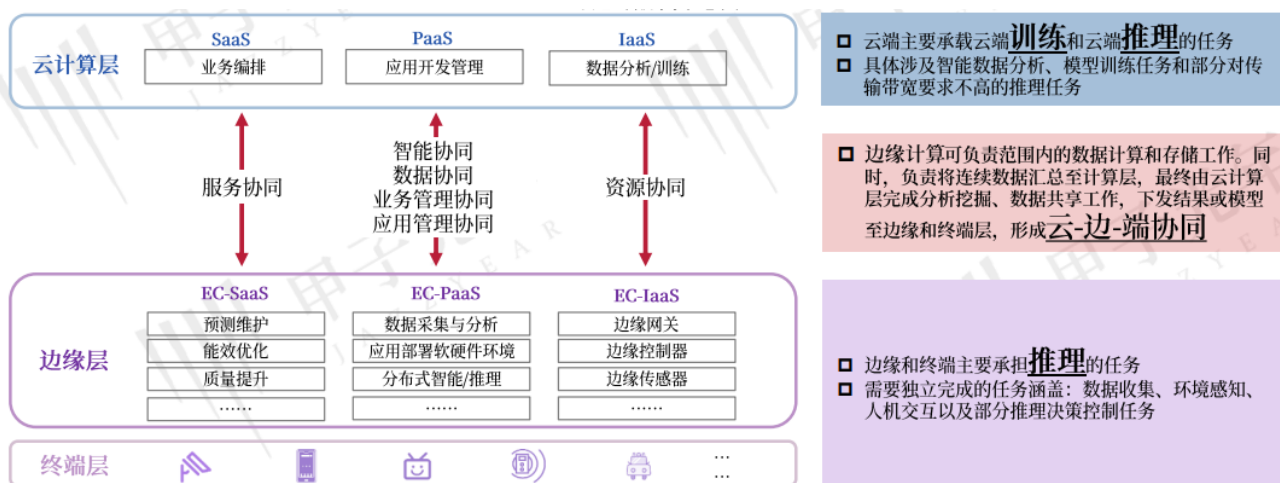
AI 芯片的维度。以部署位置（云端、终端）和承担任务（训练、推理）为横纵坐标，可以清晰的划分出 AI 芯片的市场领域。

根据部署的位置不同，AI 芯片可以分为：云 AI 芯片、端 AI 芯片。云端即数据中心，在深度学习的训练阶段需要极大的数据量和大运算量，单一处理器无法独立完成，因此训练环节只能在云端实现。云 AI 芯片的特点是性能强大、能够同时支持大量运算、并且能够灵活地支持图片、语音、视频等不同 AI 应用。基

于云 AI 芯片的技术，能够让各种智能设备和云端服务器进行快速的连接，并且连接能够保持最大的稳定。终端即手机、安防摄像头、汽车、智能家居设备、各种 IoT 设备等执行边缘计算的智能设备。终端的数量庞大，而且需求差异较大。终端 AI 芯片的特点是体积小、耗电少，而且性能不需要特别强大，通常只需要支持一两种 AI 能力。

根据承担任务的不同，AI 芯片可以分成：训练芯片、推理芯片。训练芯片主要用于构建神经网络模型，即用大量标记过的数据来“训练”相应的系统，使之可以适应特定的功能。训练需要极高的计算性能，需要较高的精度，需要能处理海量的数据，需要有一定的通用性，以便完成各种各样的学习任务。对于训练芯片来说，更注重绝对的计算能力；推理芯片主要用于利用神经网络模型进行推理预测，指利用训练好的模型，使用新数据推理出各种结论。即借助现有神经网络模型进行运算，利用新的输入数据来一次性获得正确结论的过程。推断芯片更注重综合指标，需要全面考虑单位能耗算力、时延、成本等等。以部署位置（云端、终端）和承担任务（训练、推理）为横纵坐标，可以清晰的划分出 AI 芯片的市场领域。

图表7：云-边-端协同示意图



资料来源：甲子光年，中邮证券研究所

1.3 GPU 为当前 AI 芯片领导者

在短期中，GPU 仍延续 AI 芯片的领导地位，FPGA 增长较快。GPU 短期将延续 AI 芯片的领导地位。目前 GPU 是市场上用于 AI 计算最成熟应用最广泛的通用

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

国产AI芯片的创业裂变.pdf

请登录 <https://shgis.com/post/1400.html> 下载完整文档。

手机端请扫码查看：

