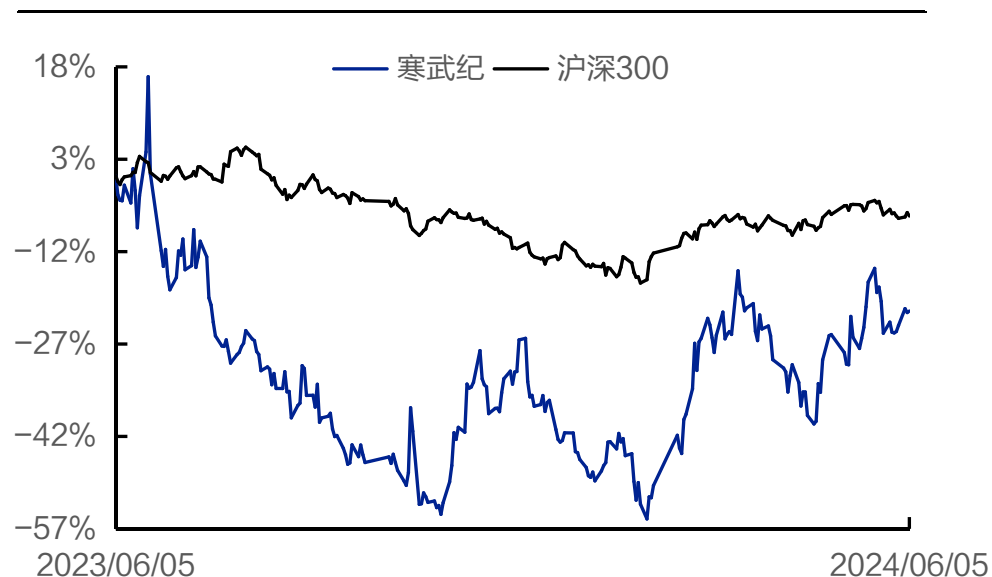


寒武纪(688256)公司报告： 全球AI芯片先行者，迎来多重成长驱动

评级：增持(首次覆盖)

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
寒武纪	5.1%	2.9%	-21.8%
沪深300	-0.3%	0.8%	-6.5%

市场数据

2024/06/05

当前价格 (元)	180.10
52周价格区间 (元)	95.85-269.50
总市值 (百万)	75,028.66
流通市值 (百万)	75,028.66
总股本 (万股)	41,659.45
流通股本 (万股)	41,659.45
日均成交额 (百万)	1,054.11
近一月换手 (%)	0.00

◆ **公司是全球智能芯片的先行者，将迎来大模型产业快速发展、AI芯片国产化以及公司新品迭代等多重成长驱动。**

◆ **公司概况：云边端一体智能芯片厂商，股权激励彰显成长信心**

公司为智能芯片领域全球知名的新兴公司，能提供云边端一体、软硬件协同、训练推理融合系列化智能芯片产品和平台化基础系统软件。2023年，公司营收7.09亿元，其中智能计算集群营收6.05亿元、同比+31.85%，主要参与台州、沈阳算力基础设施建设。2023年，公司制定股权激励计划，目标值为2024年收入 \geq 11亿元，2024-2025年累计营收 \geq 26亿元，2024-2026年累计营收 \geq 46亿元，彰显公司中长期成长信心。

◆ **AI芯片行业：海内外大模型持续发展，AI芯片国产化进程加快**

1) **需求端：海内外大模迭代，CSP与运营商投入力度加大。**2024年5月，OpenAI官宣GPT-4o多模态大模型，相同硬件条件下GPT-4o推理速度是GPT-4 Turbo两倍；OpenAI GPT-4o输入价格约0.035元/千Tokens，输出约0.1元/千Tokens。**大模型价格降低，或推动大模型厂商头部集中与AI应用加速发展。****CSP与运营商需求较高：**2024年，预期微软、谷歌、AWS、Meta的AI服务器量分别占全球采购量的20%、17%、16%、11%；2024年，中国移动启动7994台AI服务器集采招标。

2) **供给端：全球AI芯片出货量提升，国产化进程提速。**2024年，预期全球AI芯片市场规模671亿美元，同比+26%。2023年，中国加速芯片市场规模达到近140万张，推理芯片占比67%；中国本土品牌出货量超20万张，占比14%。

◆ **公司产品：云端芯片为智能计算核心，边+端侧丰富产品矩阵**

1) **云端：**为核心智能计算能力的来源，芯片包括思元290、思元370等，2022年9月公司表示思元590芯片在研待发布；公司也纵向扩展品类到训练整机（玄思1000与1001）及智能计算系统。

2) **边缘侧：**主要为思元220芯片，截至2023年底，累计出货百万片；2020年边缘芯片量产后与行业头部客户A合作，主要收入来源于公司A。

3) **IP授权及软件：**IP授权是将公司研发的智能处理器IP等知识产权授权给客户，公司也为云边端全系列智能芯片与处理器产品提供统一的平台级基础系统软件 Cambricon Neuware。

◆ **盈利预测与投资评级：**我们预计公司2024~2026年分别实现营收14.00/24.93/34.84亿元，归母净利润-4.41/0.18/2.50亿元，EPS为-1.06/0.04/0.60元，对应PS为53.59/30.10/21.53X，首次覆盖，给予“增持”评级。

◆ **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、公司业绩不及预期、公司新品研发不及预期。

预测指标	2023A	2024E	2025E	2026E
营业收入（百万元）	709	1400	2493	3484
增长率(%)	-3	97	78	40
归母净利润（百万元）	-848	-441	18	250
增长率(%)	32	48	104	1264
摊薄每股收益（元）	-2.04	-1.06	0.04	0.60
ROE(%)	-15	-8	0	5
P/E	—	—	4089.67	299.90
P/B	9.95	14.25	14.20	13.56
P/S	79.26	53.59	30.10	21.53
EV/EBITDA	-80.23	-420.78	232.27	137.57

资料来源：Wind资讯、国海证券研究所

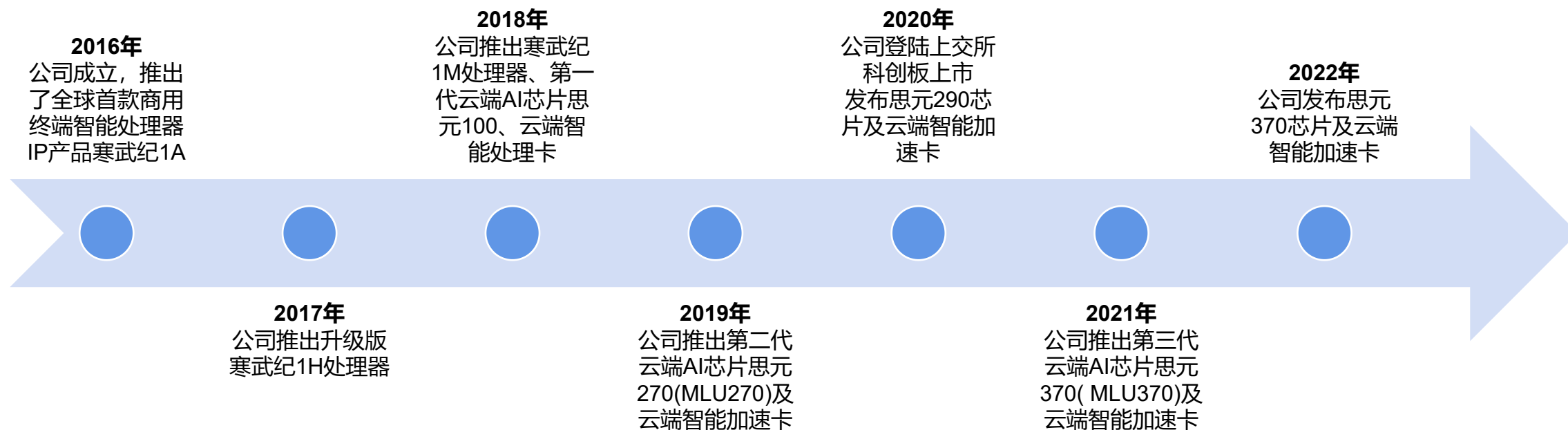
第一章 公司介绍

国内领先的人工智能芯片企业

1.1 公司介绍：国内领先的人工智能芯片企业

- 公司是智能芯片领域全球知名的新兴公司，能提供云边端一体、软硬件协同、训练推理融合、具备统一生态的系列化智能芯片产品和平台化基础系统软件。2021年，公司上榜《EETimes》评选的“AI 芯片公司(AI CHIP) TOP 10”榜单。公司智能芯片及加速卡持续在互联网、运营商、金融、能源等多个重点行业持续落地，并在业界前沿的大模型领域以及搜索、广告推荐等领域取得了长足的进步。
- 公司产品矩阵持续丰富，可满足云、边、端不同规模的人工智能计算需求。2016年成立，从寒武纪1A终端智能处理器开始，持续拓展产品线至思元云端智能芯片、玄思训练整机，思元边缘侧智能加速卡，以及相关配套软件设施等。

图：公司历史沿革



1.1 公司介绍：产品横向拓展至云边端，纵向持续升级性能

- 公司成立以来，持续横向拓宽产品类别，纵向升级迭代产品性能。
- IP产品：2016年，公司推出寒武纪1A智能处理器IP产品，随后升级1H、1M两代产品；
- 云端：2018年，公司扩充产品线至云端AI芯片，发布思元100芯片系列，随后持续迭代至思元270、思元290芯片和思元370，性能不断提升；
 - ✓ 训练整机：2020、2022年分别发布玄思1000、1001智能加速器；
- 边缘侧：2019年，发布思元220芯片及边缘智能加速卡产品。

图：公司产品矩阵

全算力布局

云边端一体

统一的基础软件系统平台

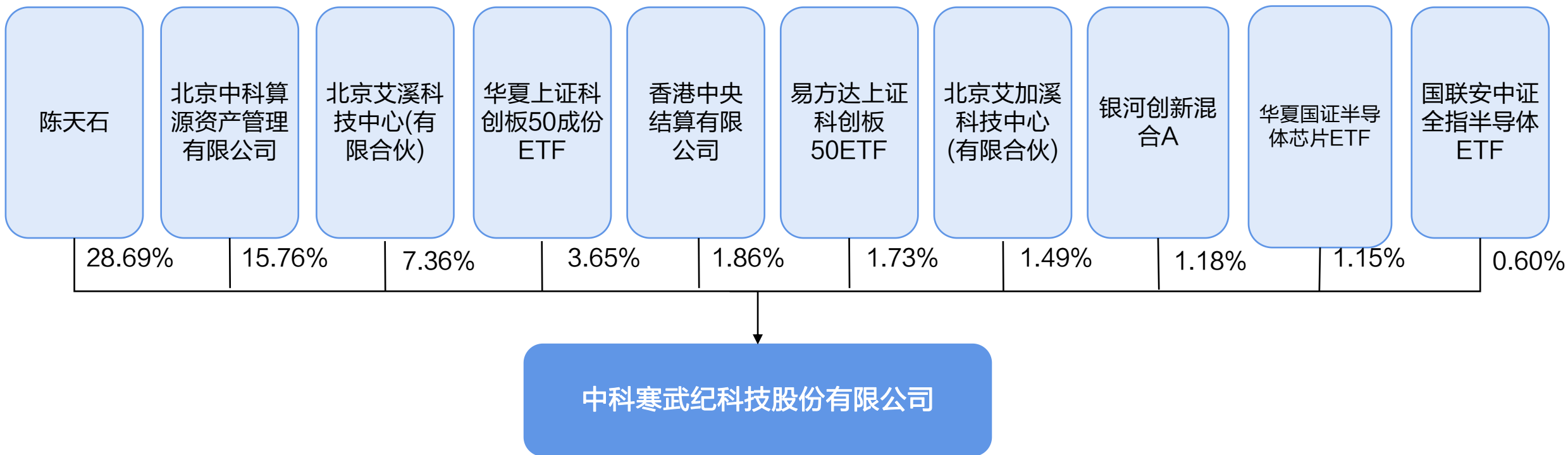
训练推理融合



1.2 公司架构：股权结构稳定，董事长为实际控制人

- 公司股权结构较为稳定。截至2024年3月，公司董事长、总经理陈天石先生是公司的第一大股东和实际控制人、直接持有公司股份28.69%，曾在中科院计算所担任研究员、博士生导师。北京中科算源资产管理有限公司为公司第二大股东、持股比例达15.76%，是中科院计算技术研究所全资设立的资产经营管理公司，具有国有法人性质。

图：公司股权架构图（截至2024年3月）



1.2 公司架构：管理层技术背景深厚，股权激励目标2024年营收 ≥ 11亿

- 公司管理层拥有深厚技术背景与从业经验。董事长、总经理陈天石为中国科学技术大学博士，曾就职于中科院计算所。副总经理王在，为中国科学技术大学计算机应用技术博士，曾于中科院计算所进行科研。公司高层学历较高，具备深厚从业经验与技术知识。
- 公司公告2023年限制性股票激励计划（草案），目标值为2024年营收 ≥ 11亿元。2023年11月，公司制定股权激励计划，涉及激励对象共计715人，约占公司全部职工人数的56.5%。根据本次激励考核目标，2024年收入应不低于11亿元，2024-2025年累计营收应不低于26亿元，2024-2026年累计营收应不低于46亿元，彰显公司对中长期成长信心。

图：公司管理层拥有深厚技术背景和丰富从业经验

姓名	职位	工作经历
陈天石	董事长 总经理	中国科学技术大学博士。2010年7月至2019年9月就职于中科院计算所，历任副研究员及硕士生导师、研究员及博士生导师。2016年3月创立寒武纪，现任中科寒武纪科技股份有限公司董事长、总经理。
王在	董事 副总经理	中国科学技术大学计算机应用技术博士。2016年至2018年就职于中科院计算所从事科研工作。2016年作为寒武纪创始团队成员加入寒武纪，现任中科寒武纪科技股份有限公司董事、副总经理。
刘立群	董事 副总经理	清华大学科技哲学硕士。2002年起任中科寒武纪科技股份有限公司副总经理；2016年至今就职于国投创业投资管理公司历任董事总经理、副总经理。现任中科寒武纪科技股份有限公司董事。
张尧	副总经理	中国科学院计算机系统结构硕士。2015年至2016年任北京小米松果电子有限公司高级工程师。2016年至今，就职于中科寒武纪科技股份有限公司，现任中科寒武纪科技股份有限公司副总经理、芯片部高级总监。
叶湫尹	副总经理 财务负责人 董事会秘书	北京大学硕士。2012年至2016年就职于中国高新投资集团公司曾高级投资经理，曾任国投创业投资管理有限公司投资副总裁。2019年加入寒武纪，现任中科寒武纪科技股份有限公司董事、副总经理、财务负责人及董事会秘书。

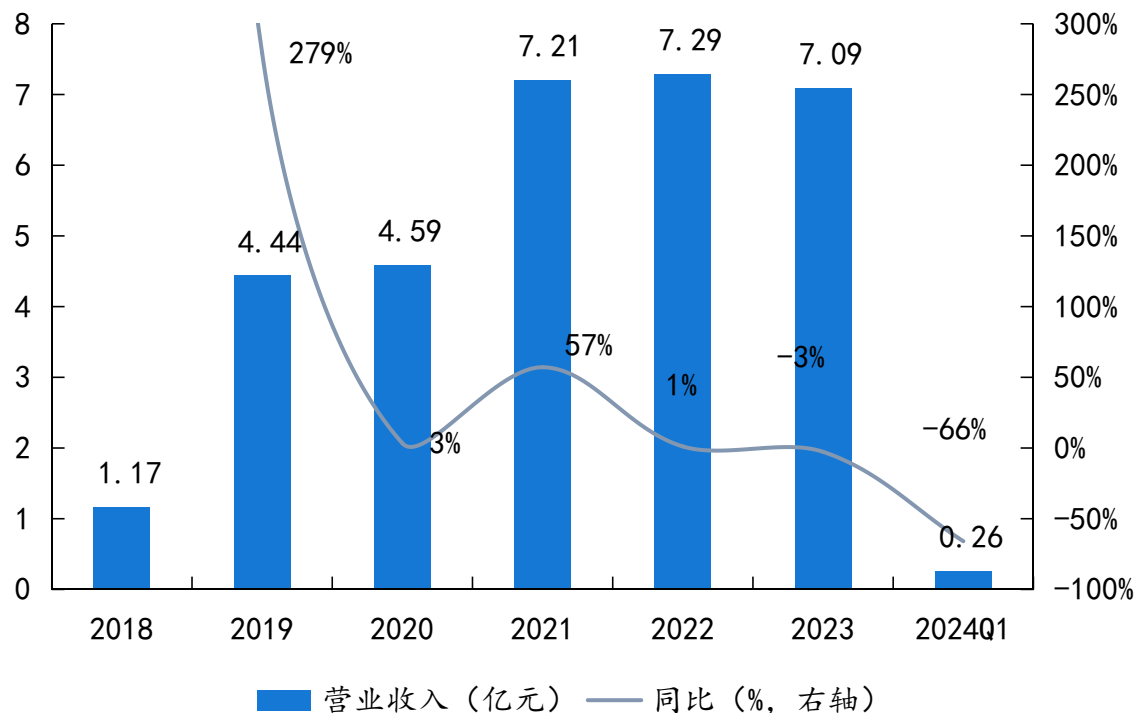
图：2024-2026会计年度考核目标

归属期	目标值	触发值
	公司层面归属系数 100%	公司层面归属系数 80%
第一个归属期	2024年营业收入值 不低于11亿元	2024年营业收入值 不低于8.8亿元
第二个归属期	2024-2025年累计 营业收入值不低于 26亿元	2024-2025年累计 营业收入值不低于 20.8亿元
第三个归属期	2024-2026年累计 营业收入值不低于 46亿元	2024-2026年累计 营业收入值不低于 36.8亿元

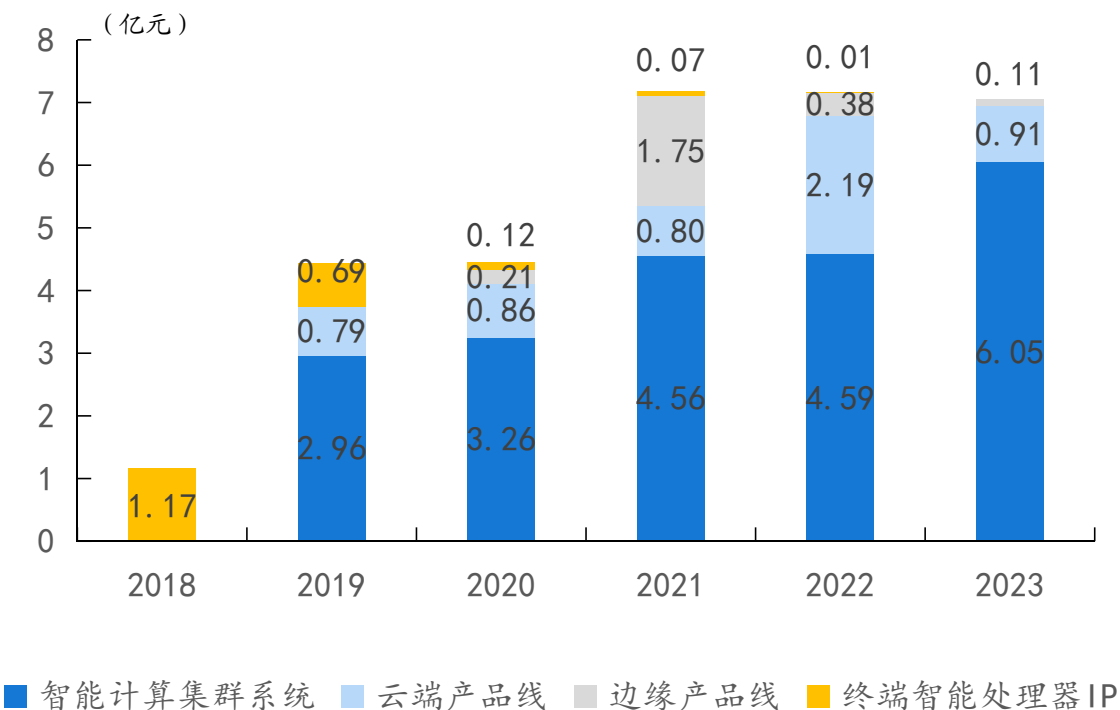
1.3 公司财务：2021-2023年营收平稳增长，云端产品为主要动能

- **营收端：**2018-2023年，公司营收CAGR为43.39%，增速较快。2018年，公司营收主要来自智能处理器IP产品；2019-2020年，公司云端产品线与智能计算集群产品持续增长，成为营收增长的主力；2021年，公司发布边缘端产品思元220系列，带来当年边缘产品线营收1.75亿元、增长快速；2022-2023年，智能计算集群系统的营收增长较快。

图：2018-2024Q1 公司营收及同比



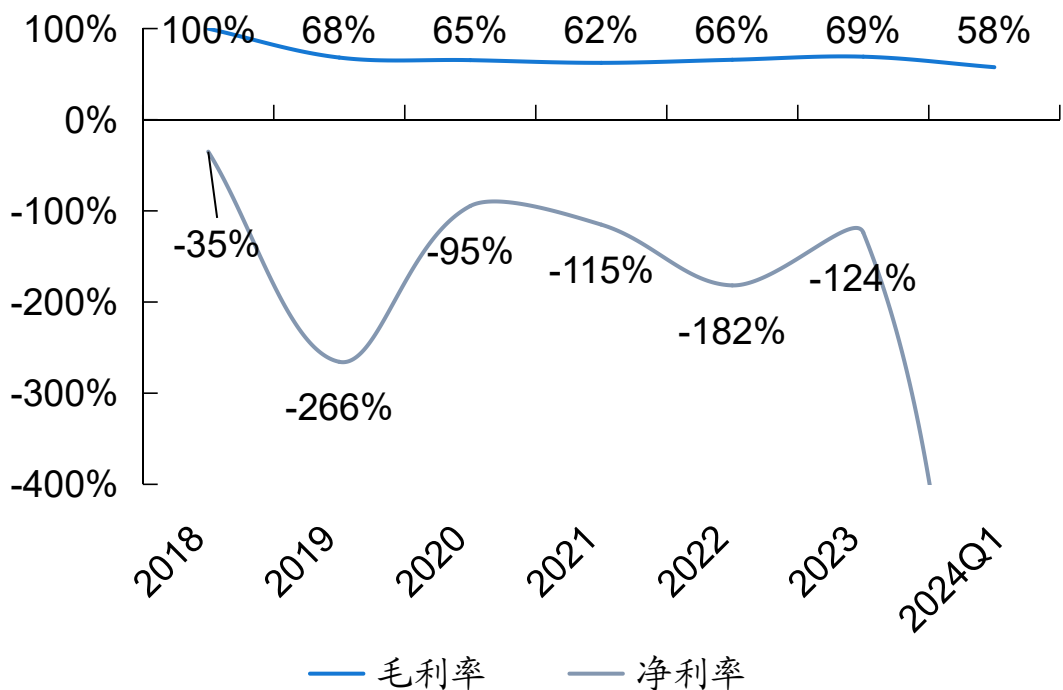
图：2018-2023年 公司分产品营收情况



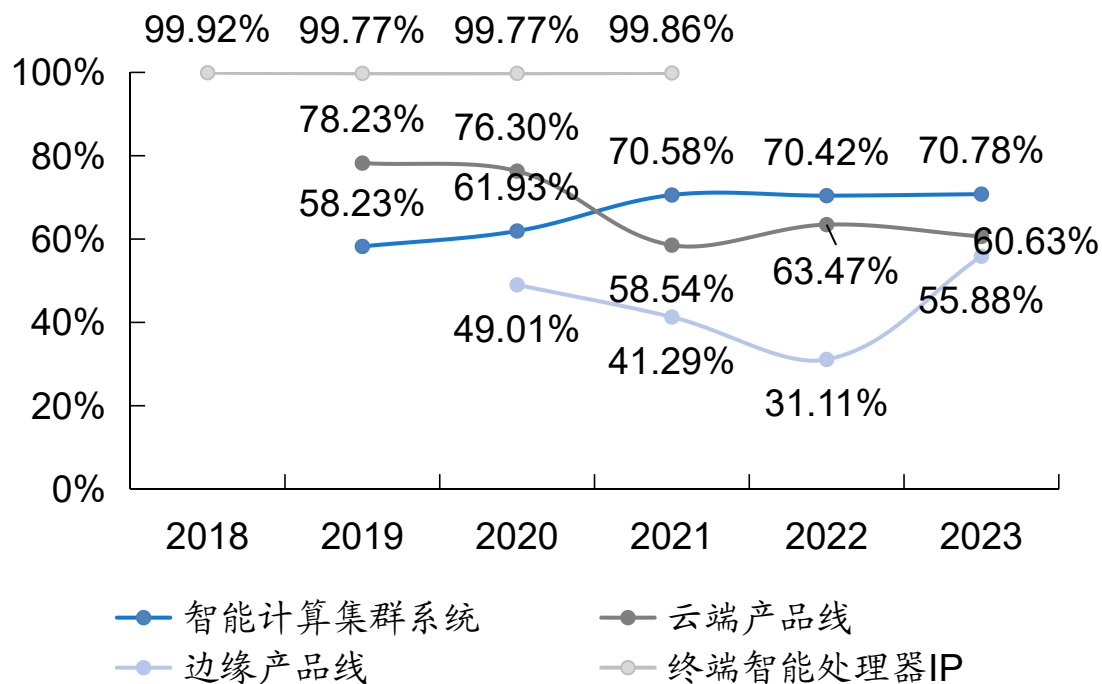
1.3 公司财务：毛利率保持相对稳定，产品结构影响较大

- 2019-2023年，公司毛利率保持相对平稳：2018年，公司营收来自于终端智能处理器IP，故毛利率较高；2019年-2023年，公司毛利率保持在62-69%，其中2021年，毛利率较低的边缘产品线占比提升，并且云端产品线毛利率有所下降，导致毛利率下滑至62%，但后续公司边缘端产品线占比降低，并且云端产品线毛利率回升至60%上下，以及毛利率70%+的智能计算集群系统业务营收持续增长，带来公司毛利率持续提升。
- 2023年，公司智能计算集群系统业务毛利率为70.78%，参与台州、沈阳两地的算力基础设施建设项目，营收达到6.05亿元、占比高达85%。

图：2018-2024Q1公司毛利率与净利率情况



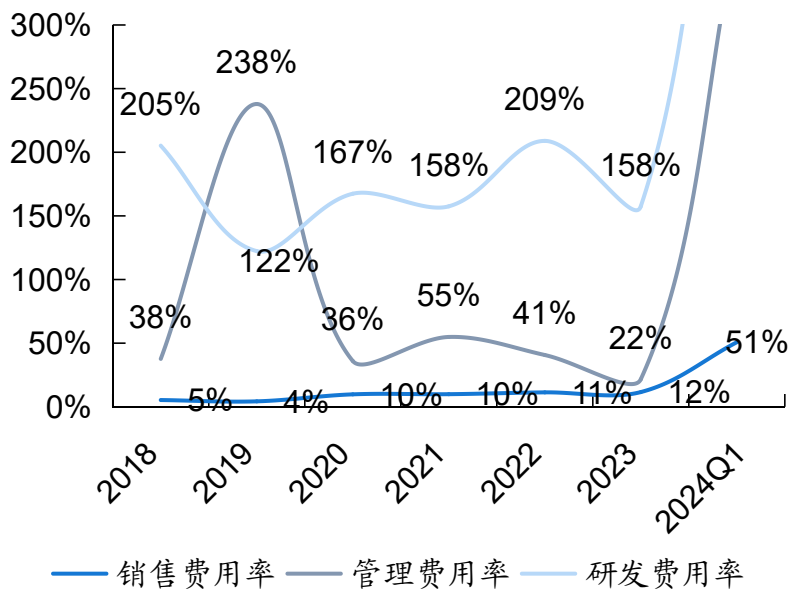
图：2018-2023年公司分产品毛利率情况



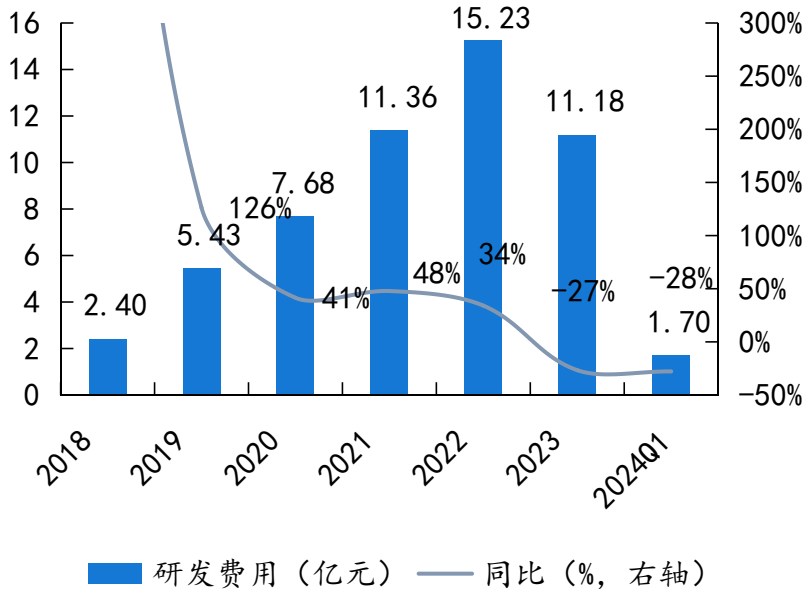
1.3 公司财务：期间费用率总体上升，2018-2022年研发投入持续增长

- **销售费用率**：2019-2023年，公司销售费用率持续增长。2022年，销售费用率为11.35%、同比+1.44pct，主要系公司完善销售体系、拓展销售渠道，并且积极发力智能芯片市场推广与生态建设；
- **管理费用率**：2019年，公司管理费用率达到238%，主要系股份支付金额较高；2023年管理费用率22%、同比减少19pct，主要系分摊股份支付费用同比减少所致；
- **研发费用率**：2019-2022年，公司研发费用率持续增长，2023年呈下降趋势。2022年，公司研发费用率209%、同比+51pct，主要系公司研发人员薪酬增长，测试化验加工费增加（新产品流片等费用增长），无形资产摊销与固定资产折旧增长（为研发需求购置了IP、EDA等无形资产以及研发设备等）。2023年，公司研发费用11.18亿元、同比-27%，主要系2022年12月15日，BIS将公司与部分子公司列入“实体清单”，受此影响，公司调整战略，陆续暂停部分预期毛利率较低的研发项目。

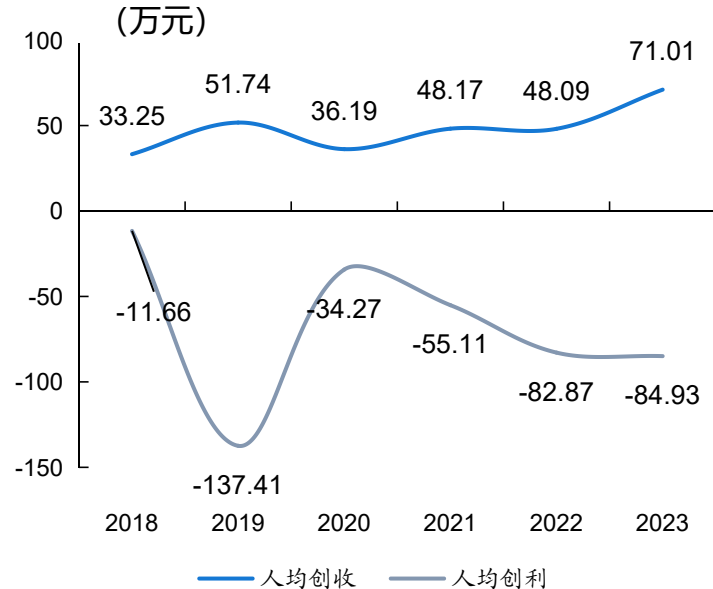
图：2018-2024Q1 公司期间费用率情况



图：2018-2024Q1 公司研发费用情况



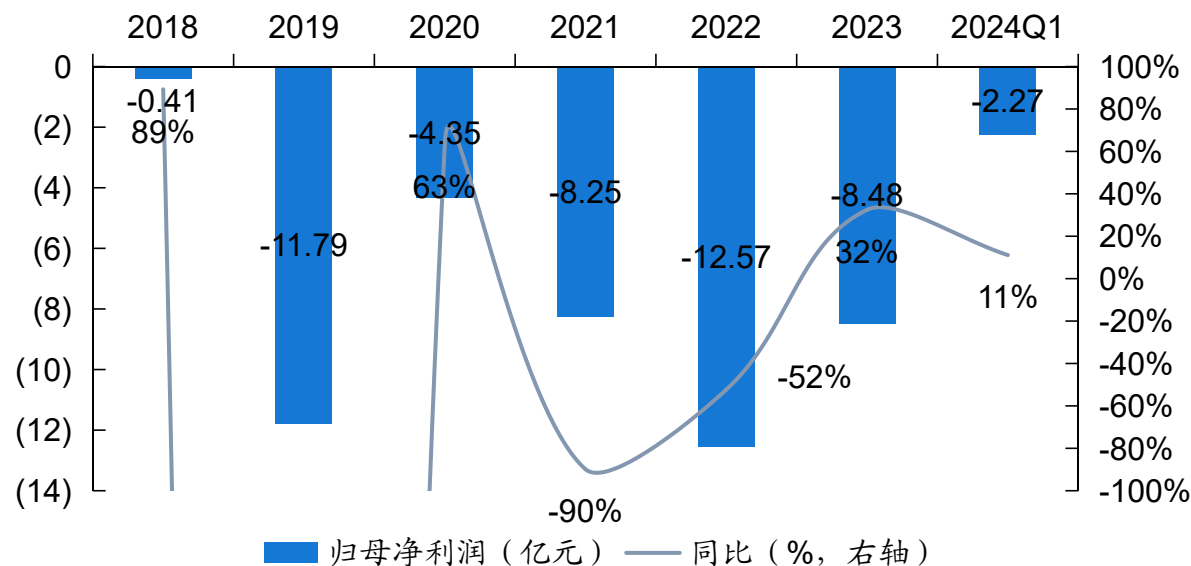
图：2018-2023年 公司人均创收创利



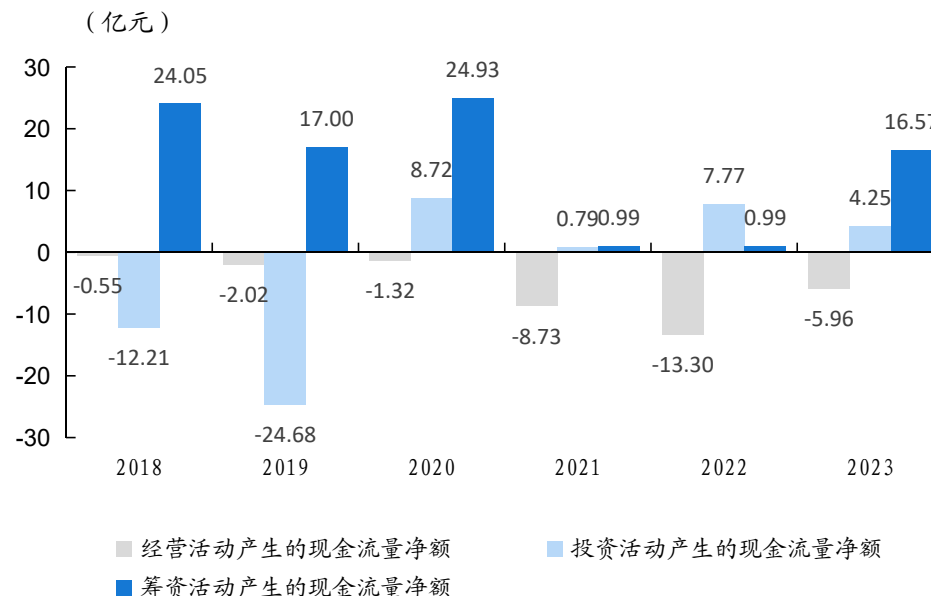
1.3 公司财务：利润与现金流承压，关注供应链+业务拓展

- **2018-2024Q1，公司净利润承压。**2023年，公司归母净利润为-8.48亿元、减亏4.09亿元，主要系被列入实体清单后，公司供应链呈现压力，并且研发费用、资产减值损失、信用减值损失增长，代理公司净利润承压。我们预期，随着公司供应链好转，并且调整销售策略，优先服务毛利较高、信用较好的客户，利润有望持续好转。
- **2018年-2023年，公司经营活动现金流量净额为负。**2023年，公司经营性现金流量净额为-5.96亿元、同比+7.34亿元，主要系2023年销售回款较上年同期增加3.76亿元，以及采购支出较上年同期减少2.92亿元所致。我们认为，预计未来随着公司业务规模扩大、盈利能力提升，现金流量状况有望得到改善。

图：2018-2024Q1 公司归母净利润及同比



图：2018-2023年 公司现金流情况



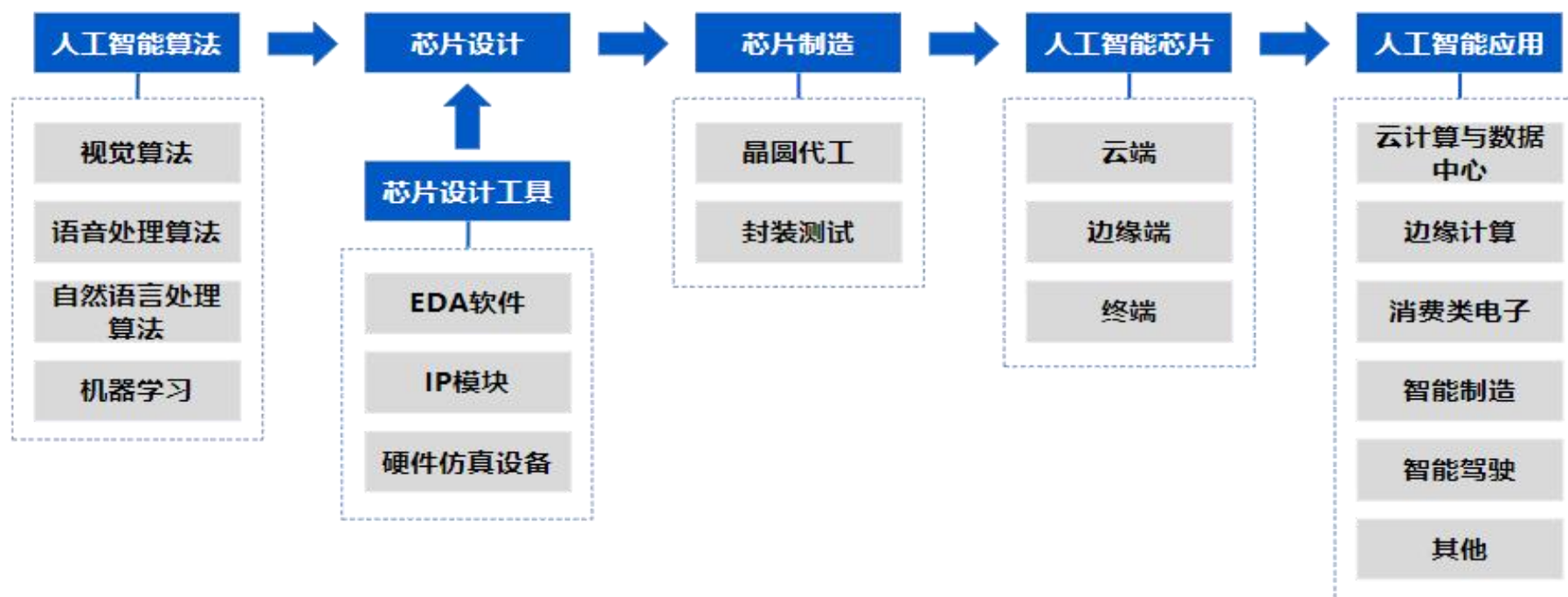
第二章 AI芯片行业

大模型推动AI算力需求高增，云边端芯片持续发展

2.1 人工智能产业链：算法+硬件+应用

- 人工智能产业链包含人工智能算法、芯片设计制造、人工智能应用侧等环节。
 - ✓ **算法**：覆盖广泛，包括视觉算法、语音处理算法、自然语言处理算法以及各类机器学习方法。
 - ✓ **芯片**：人工智能芯片行业的核心为芯片设计和芯片制造，AI芯片分为面向云端、边缘端、终端等品类。
 - ✓ **应用**：当前我国人工智能芯片行业的下游应用场景主要聚集在云计算与数据中心、边缘计算、消费类电子、智能制造、智能驾驶、智慧金融、智能教育等领域。

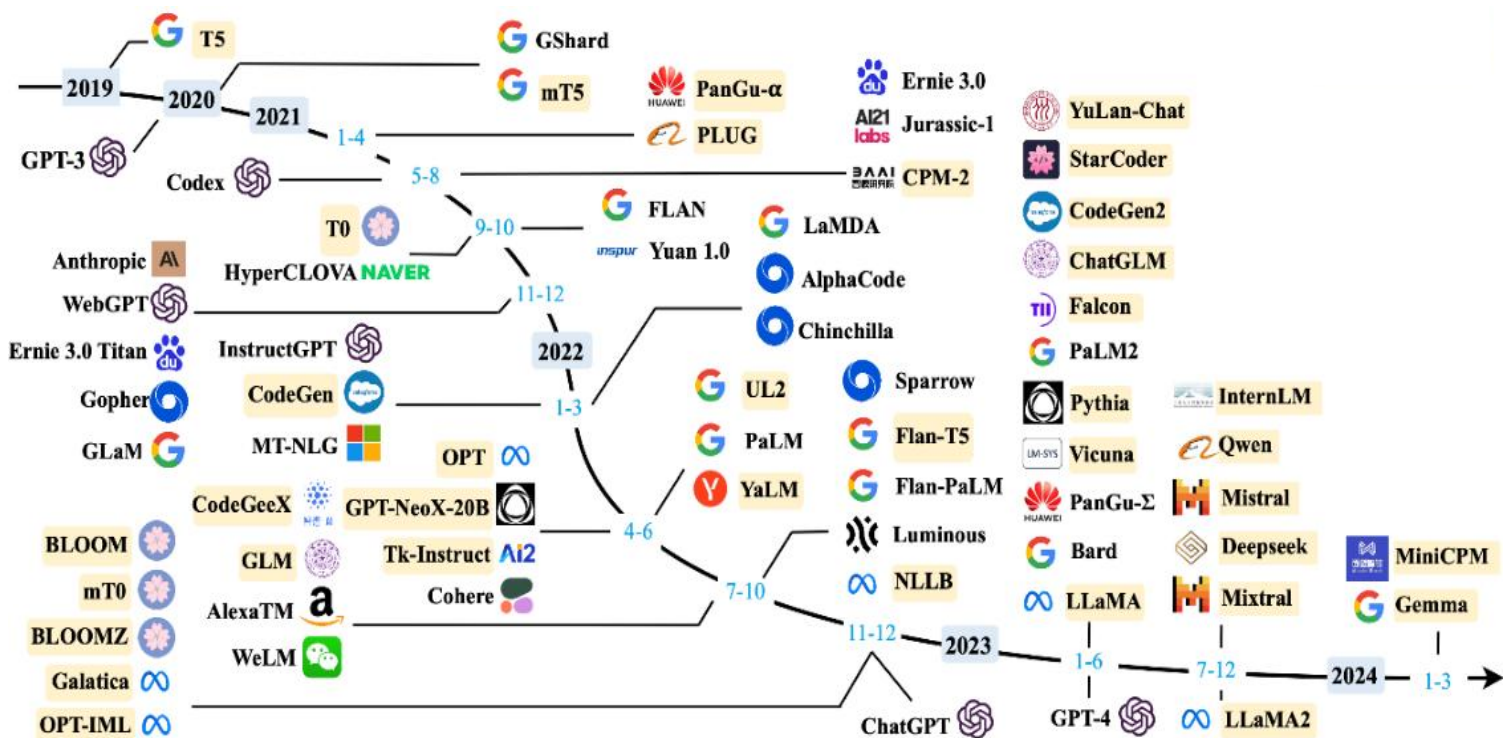
图：人工智能产业链



2.2 需求端：大模型多模态能力迭代，推动算力需求提升

- OpenAI持续迭代大模型，推动新一轮技术创新与产业变革。2024年2月，OpenAI发布Sora多模态文生视频大模型，从“生成视频的长度”、“镜头切换与画面流畅度”、“使用的基础模型架构”三个方面均远超现有同类产品水平。5月，OpenAI官宣GPT-4o多模态大模型，相同硬件条件下GPT-4o的推理速度是GPT-4 Turbo的两倍，且完全免费；GPT-4o也具备更好的图文影音识别能力、更自然的交互融合能力、与更好的理解能力，有望成为效率更高的人工智能助理。
- 目前，AI大模型实现了从“文生文”到“文生图”、“文生视频”等多模态能力的发展，带来了算力需求显著提升。

图：全球大模型发展时间线



表：OpenAI大模型迭代情况

版本	GPT	GPT-2	GPT-3	GPT-4
发布时间	2018.6	2019.2	2020.5	2023.3
发布者	OpenAI	OpenAI	OpenAI	OpenAI
参数数量	1.17亿	15亿	1750亿	预计万亿
数据	5.7GB 文本	40GB 文本	45TB 文本	20万亿 Tokens

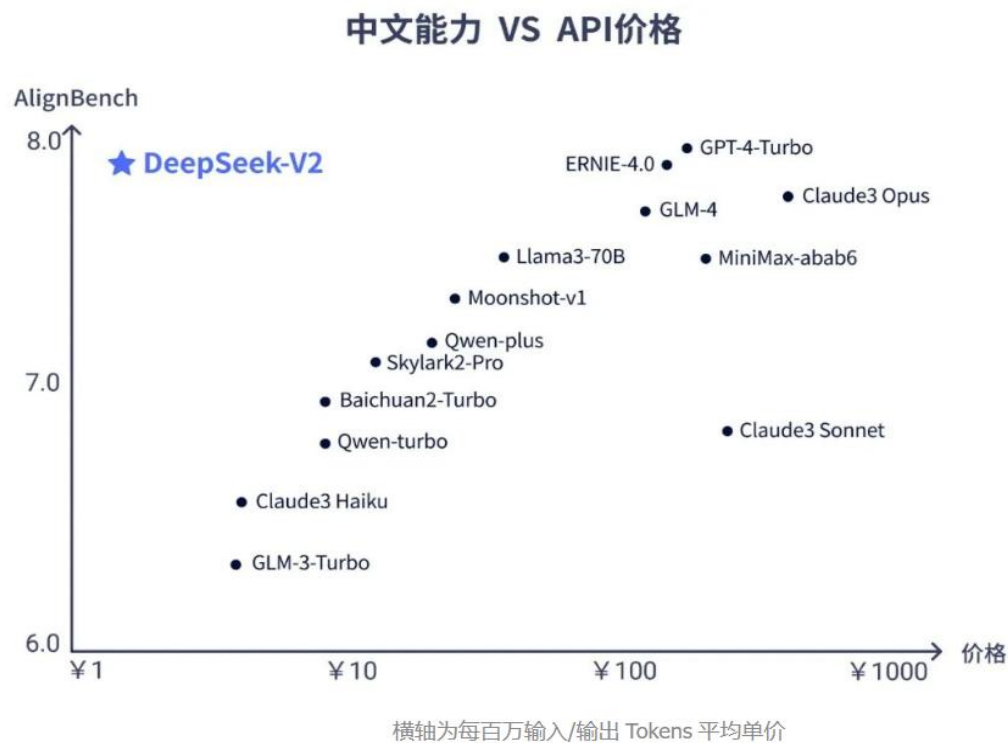
2.2 需求端：大模型降价或引头部集中，推动AI应用发展提速

- **大模型价格战打响，从强调技术迭代到强调性价比。** 国外大模型中，OpenAI GPT-4o输入价格约为0.035 元/千Tokens，输出约为0.1元/千Tokens。国内大模型中，5月15日，字节刚官宣豆包大模型，输入/输出价格0.0008元/千tokens，比行业价格低了99.3%；5月21日，阿里云宣布通义千问主力模型Qwen-Long直降97%（输入/输出价格0.0005元/千tokens），约为GPT-4价格的1/400。
- **1) 大模型更趋头部集中趋势：** 价格战打响，或意味着大模型行业整合的开始，效率不高的公司或被淘汰。**2) AI应用发展提速：** 大模型API接口成本降低，或有助于企业调用基座大模型进行推理，AI应用发展有望提速。

表：OpenAI大模型迭代情况

公司	大模型	输入价格 (元/千Tokens)	输出价格 (元/千Tokens)
Open AI	GPT-4o	0.035	0.1
百川智能	Baichuan2-Turbo	0.008	0.008
DeepSeek	DeepSeek v2	0.001	0.002
阿里	Qwen-Long	0.0005	0.0005
文心一言	ERNIE-4.0-8K	0.12	0.12
字节跳动	豆包	0.0008	0.0008
智谱AI	GLM-3-Turbo	0.001	0.001

图：国内外大模型性价比盘点



2.2 需求端：国内运营商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

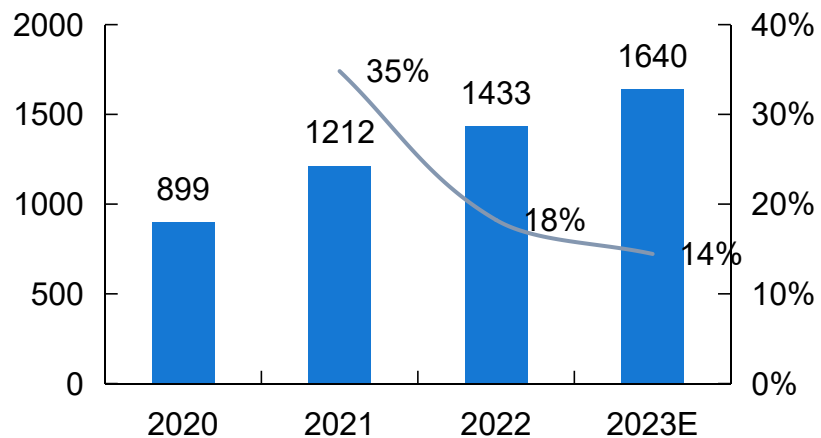
标包	产品名称	需求数量	中标人数与份额
标包1	人工智能服务器	7994台	<ul style="list-style-type: none"> ①昆仑技术：190.97亿元，份额21.05%； ②华坤振宇：190.97亿元，份额17.54%； ③宝德计算机：190.97亿元，份额15.79%； ④百信：191.36亿元，份额14.04%； ⑤长江计算：191.17亿元，份额12.28%； ⑥神州鲲泰：191.17亿元，份额10.53%； ⑦湘江鲲鹏：190.58亿元，份额9.77%
标包2	交换机	60台	<ul style="list-style-type: none"> 中标人1家，中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

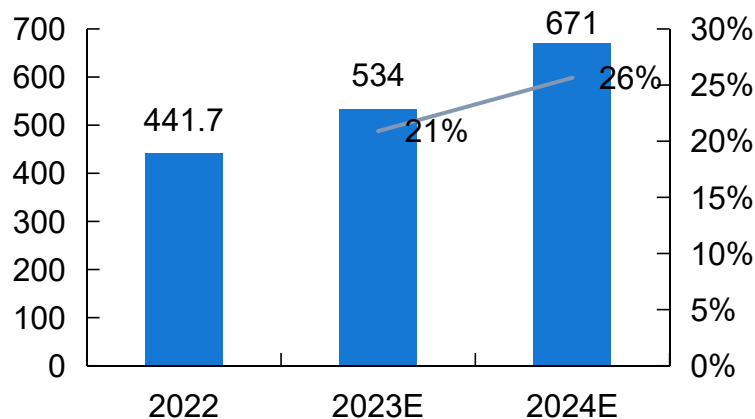
2.3.1 供给端：全球AI服务器+芯片市场规模快速提升

图：2020-2023E 全球AI芯片数量统计



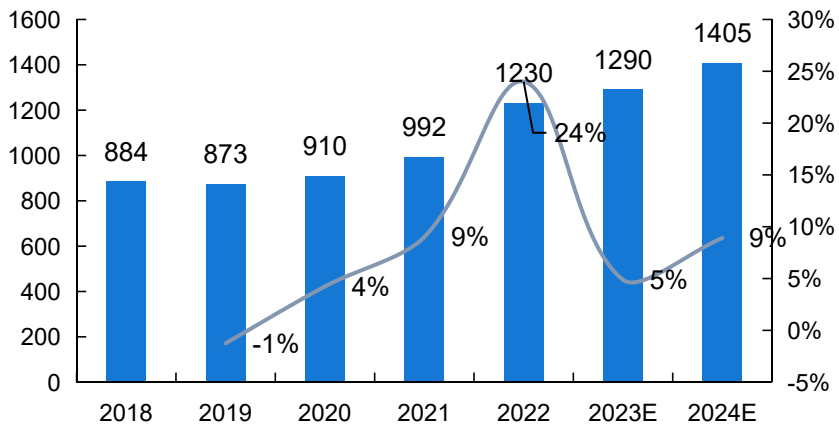
■ 全球AI芯片数量统计 (万套) — 同比 (%) (右轴)

图：2022-2024E 全球AI芯片市场规模



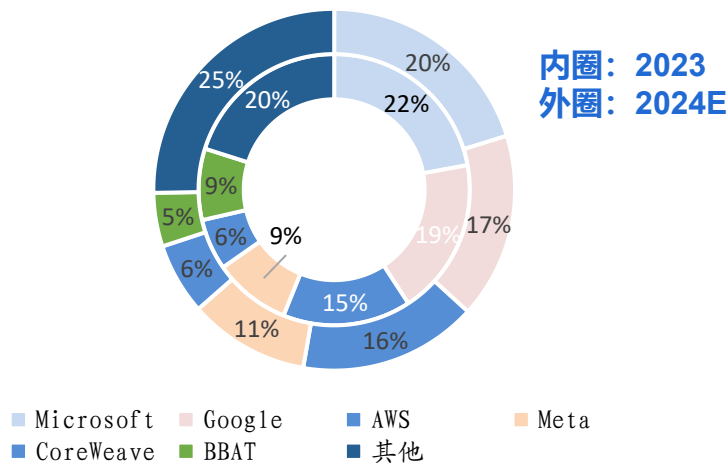
■ 全球AI芯片市场规模预测 (亿美元) — 同比 (%) (右轴)

图：全球AI服务器市场规模



■ 全球服务器市场规模 (亿美元) — 同比 (%) (右轴)

图：2023-2024E 全球AI服务器采购量占比

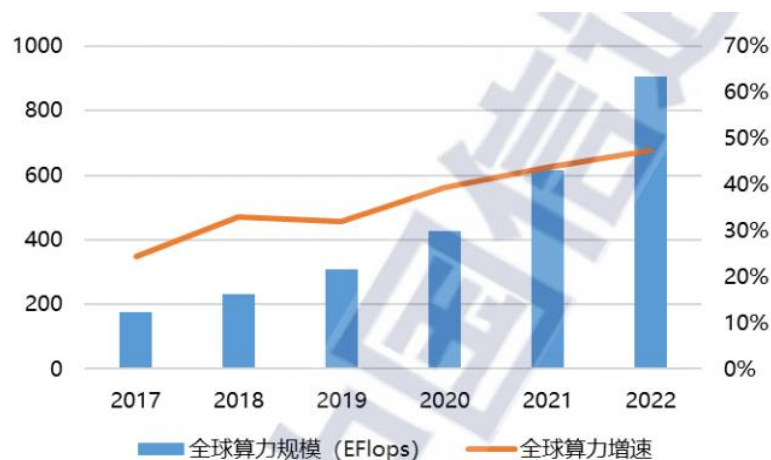


■ Microsoft ■ Google ■ AWS ■ Meta
■ CoreWeave ■ BBAT ■ 其他

全球算力市场：

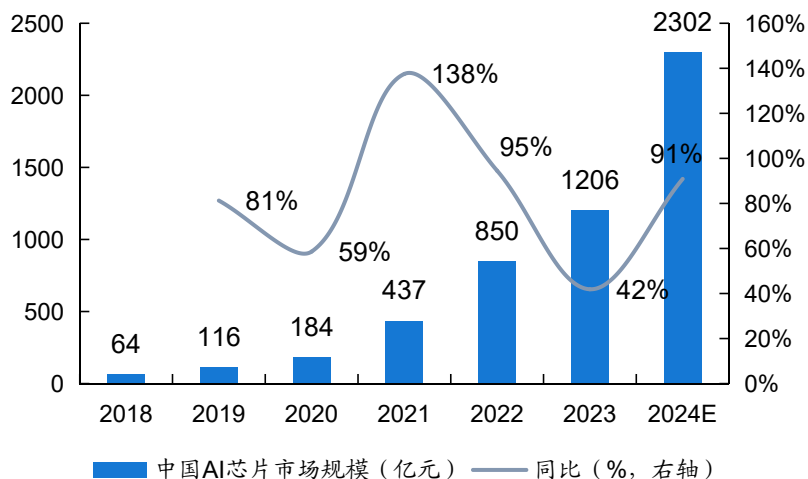
- 2024年，Gartner预期全球AI芯片市场规模671亿美元，同比增长26%。
- 2024年，TrendForce预期微软、谷歌、AWS、Meta的AI服务器量分别占全球采购量的20%、17%、16%、11%。

图：2017-2022年全球算力规模及增速

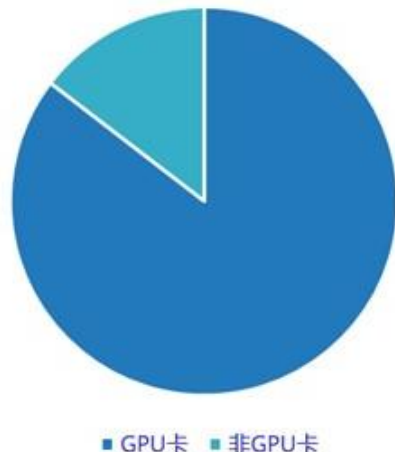


2.3.1 供给端：国内AI算力市场持续景气，推理占比持续提升

图：2018-2024年中国AI芯片市场规模预测

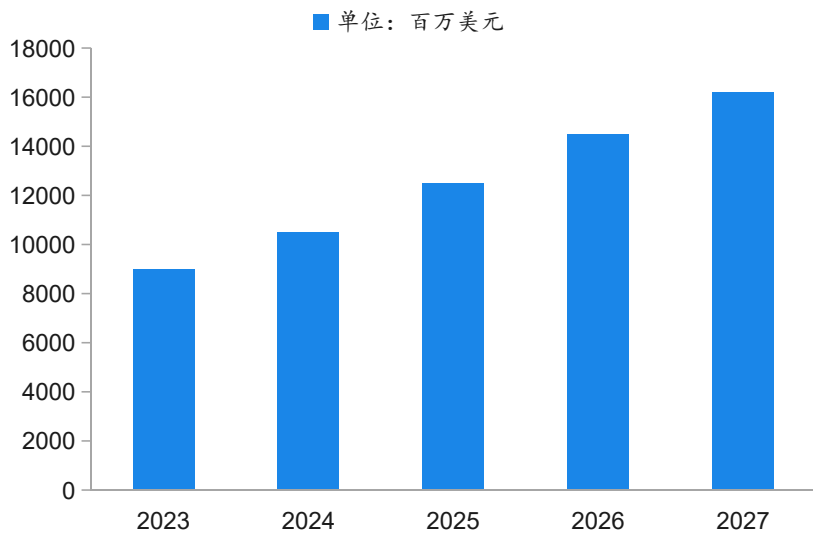


图：2023H2 中国人工智能芯片市场份额

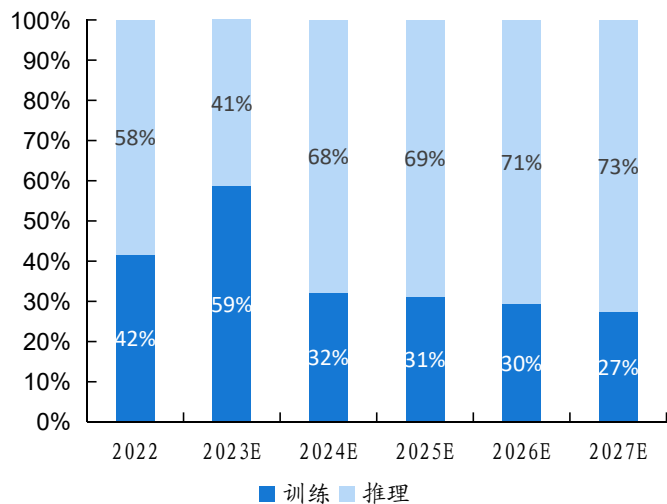


- 中国加速芯片市场持续增长，2023年国产化比例达14%。据IDC，2023年，中国加速芯片的市场规模达到近140万张。从技术角度来看，GPU卡占据85%的市场份额；从品牌角度来看，中国本土人工智能芯片品牌的出货量已超过20万张，约占整个市场份额的14%。

图：中国加速服务器市场规模



图：2022-2027E 中国AI服务器工作负载预测



- 推理占比持续提升。2023年，用于推理的人工智能芯片占据了67%的市场份额。据IDC，预期2023-2027年，推理AI服务器工作负载占比从41%提升至73%左右。

- 2023年中国AI服务器增长迅速。IDC数据显示，2023全年中国加速服务器市场规模达到94亿美元，同比2022年增长104%。其中GPU服务器依然是主导地位，占据92%的市场份额，达到87亿美元。

2.3.2 AI芯片：面向云、边、端侧，下游场景丰富

- 人工智能技术在云端（云）、边缘端（边）和终端（端）设备中均有广泛应用，但都需要由核心芯片提供计算能力支撑。云、边、端三种场景对于芯片的运算能力和功耗等特性有着不同要求，单一品类的智能芯片难以满足实际应用的需求

图：面向云、边、端侧的芯片情况

应用场景	芯片需求	典型计算能力	典型功耗	典型应用领域
终端	低功耗、高能效、推理任务为主、成本敏感、硬件产品形态众多	<8TOPS	<5瓦	各类消费类电子、物联网产品等
云端	高性能、高计算密度、兼有推理和训练任务、单价高、硬件产品形态少	>30TOPS	>50瓦	云计算数据中心、企业私有云等
边缘端	对功耗、性能、尺寸的要求常介于终端与云端之间、推理任务为主、多用于插电设备、硬件产品形态相对较少	5TOPS至30TOPS	4瓦至15瓦	智能制造、智能家居、智能零售、智慧交通、智慧金融、智慧医疗、智能驾驶等众多应用领域

注1：云、边、端应用场景尚无标准划分界限，上表为公司基于自主研发技术体系划分。

注2：TOPS为计量运算能力的基本单位，1TOPS代表每秒处理一万亿次人工智能基本运算。

2.3.2 AI芯片：包括GPU、ASIC、通用智能芯片等

- 人工智能芯片包括传统芯片与智能芯片，传统芯片以GPU为主；智能芯片是面向人工智能领域而专门设计的芯片，性能和能效优势主要集中于智能应用，不支持双精度浮点运算、图形渲染类运算、无线通信类信号处理运算等。在人工智能领域，智能芯片的优势明显，可替代GPU等传统芯片；在通用计算和图形渲染等人工智能以外的其他领域，智能芯片无法替代CPU、GPU等传统芯片，存在局限性。

图：多种人工智能芯片对比情况

	芯片	技术原理	技术特点	技术优劣势	市场需求情况	处理器类型与代表厂商	市场渗透率
传统芯片	GPU	<p>(1) GPU的基本原理为：通过简化控制单元并集成大规模的并行运算单元，实现对图形渲染等并行任务的良好支持</p> <p>(2) 具体对于智能训练和推理应用，通过GPU的向量等指令组合出训练或推理需要的运算操作，从而实现对智能算法的支持</p>	技术成熟，通用性较好，擅长数据级并行处理，为图形处理、科学计算等传统任务提供了良好的硬件支持	峰值运算性能高，但整体能耗较高；在云端具备成熟的应用开发生态，但在终端生态尚不成熟	图形渲染和科学计算市场需求大且稳定，但在人工智能领域面临通用型智能芯片的挑战	<p>(1) 云端主流产品为AMD和Nvidia产品，主流工艺为7/12/16nm</p> <p>(2) 边缘端或终端主流产品为SoC集成的GPU IP，主流厂商包括ARM、Imagination等</p>	在人工智能领域，GPU多用于服务器与数据中心，是目前渗透率最高且最主流的芯片类型，在终端应用较少
	FPGA	<p>(1) FPGA的基本原理是：通过集成大量的可重构逻辑单元阵列，可支持硬件架构的重构，从而灵活支持不同的人工智能模型</p> <p>(2) FPGA一般仅用于智能推理，其技术原理通常是将对应的智能算法通过硬件编程的方式实现在FPGA上，从而支持推理任务</p>	技术成熟，包含充裕的可重构逻辑单元阵列，硬件编程开发和调试门槛较高	可通过硬件重构方式灵活实现适合于人工智能应用的架构，但成本和能效与主流技术路径仍有差距	芯片原型验证市场需求大且稳定	主流厂商包括Xilinx和Intel等。其中，Xilinx FPGA产品包括Virtex、Kintex等系列；Intel FPGA产品包括Agilex、Stratix等系列	在芯片原型验证与仿真中有着广泛应用，在人工智能芯片市场渗透率相对较低
智能芯片	通用型智能芯片	<p>(1) 通用型智能芯片的原理是：通过对各类智能应用和算法的计算和访存特点进行抽取和抽象，定义出一套适用于智能算法且相对灵活的指令集和处理器架构，从而广泛支持多样化的人工智能算法和应用</p> <p>(2) 智能芯片的指令通常与人工智能算法中的关键运算操作相匹配</p> <p>(3) 在具体的训练和推理应用中，对于关键运算操作，智能芯片指令可直接支持，从而实现高效的训练和推理</p>	相关技术持续发展，全新指令集完备高效，可覆盖各类智能算法所需的基本运算操作	性能、功耗比较传统芯片优势明显，可适应各种场景和规模的人工智能计算需求	人工智能市场需求潜力大，未来将成为该市场主流产品	<p>(1) 云端和边缘通用型智能芯片处于应用推广期，主要厂商和产品为寒武纪（思元100/270/220）、华为海思（Ascend 310/910）、Google（TPU V1/V2/V3、TPU EDGE）等</p> <p>(2) 终端通用型智能处理器多集成于手机SoC等芯片中，已实现大规模应用，主要厂商和产品为华为海思（麒麟970/980/990）等</p>	在云端、边缘端和消费类电子终端都开始出现广泛应用，渗透率将逐渐提升
	专用型智能芯片 (ASIC)	专用型智能芯片的原理是：针对面向特定的、具体的、相对单一的人工智能应用专门设计的芯片，具体实现方法为在架构层面对特定智能算法作硬化支持，多用于推理任务	相关技术持续发展，在架构层面对特定智能算法作硬化支持，指令集简单或指令完全固化	成本相对较低，软件栈相对简单，设计和生产周期短，但通用性较差	应用细分市场市场需求大且分散，成本敏感	目前主要应用于终端，主要形态为行业专用SoC，较多集中于语音处理领域	常用于在低功耗、成本敏感的终端上支撑特定的智能应用，在云端、边缘端等场景渗透率相对较低

第三章 公司产品

云端产品增长快速，边+端侧丰富产品矩阵

3.1 公司产品：云边端横向拓宽产品，纵向升级迭代产品性能




图：公司主要产品线、推出时间

产品线	产品类型	主要产品	推出时间	具体内容
云端产品线	云端智能芯片及加速卡	思元 100 (MLU100) 芯片及云端智能加速卡	2018年	1) 云端智能芯片及加速卡： 云服务器、数据中心等进行人工智能处理的核心器件 2) 整机： 由公司自研的训练服务器产品
		思元 270 (MLU270) 芯片及云端智能加速卡	2019年	
		思元 290 (MLU290) 芯片及云端智能加速卡	2020年	
		思元 370 (MLU370) 芯片及云端智能加速卡	2021年 2022年	
	训练整机	玄思 1000 智能加速器	2020年	
		玄思 1001 智能加速器	2022年	
边缘产品线	边缘智能芯片及加速卡	思元 220 (MLU220) 芯片及边缘智能加速卡	2019年	边缘计算式在终端和云端之间的设备上配备适度的计算能力
IP 授权及软件	终端智能处理器 IP	寒武纪 1A 处理器	2016年	1) IP授权： 将智能处理器IP等知识产权授权给客户在其产品中使用 2) 基础系统软件平台： 公司为云边端全系列智能芯片与处理器产品提供统一的平台级基础系统软件（包含软件开发工具链等）
		寒武纪 1H 处理器	2017年	
		寒武纪 1M 处理器	2018年	
	基础系统软件平台	寒武纪基础软件开发平台（适用于公司所有芯片与处理器产品）	持续研发和升级，以适配新的芯片	

3.2 云端产品：为核心智能计算能力的来源

- 云端智能芯片为核心智能计算能力的来源，纵向扩展品类至训练整机及智能计算集群系统。
- 1) 云端智能芯片：包括云端训练芯片（思元290）、云端推理芯片（思元100、思元270及思元370）。
- 2) 训练整机：包括玄思1000、1001系列，其中玄思1000训练整机于2021年批量出货。
- 3) 智能计算集群系统：收入主要来自1-2个规模较大的城市智能计算中心，2019年主要系西安沣东、珠海横琴项目；2020年主要系江苏南京、无锡数据湖项目；2021年主要系江苏昆山项目；2022年主要系江苏南京项目（二期）；2023年参与台州、沈阳两地的算力基础设施建设项目。

图：公司主要产品线、推出时间及营业收入

产品形态	产品图示	主要客户	应用场景	具体内容
云端智能芯片和加速卡		服务器厂商、互联网、金融、电信运营商等	智能服务器；互联网领域的视觉、语音、自然语言处理；运营商数据中心	以云端智能芯片为核心芯片，配备DRAM芯片颗粒等外围芯片的PCB电路板，与散热片等结构件组合后形成的可直接插入服务器标准接口（如PCI-E接口、OAM接口）的板卡硬件产品。
训练整机		服务器厂商、智能算法企业等	互联网领域的视觉、语音、自然语言处理；算法企业；数据中心	面向云端训练场景，为公司云端训练芯片及加速卡配套设计整机主板，并将多块加速卡通过标准接口集成于主板上，同时配套电源机箱等部件形成的整机产品。主要面向有一定技术基础的商业客户群体。
智能计算集群		城市数据中心等	以云计算的方式向有智能计算需求的分散客户提供算力资源	<ul style="list-style-type: none"> • 将公司自研的加速卡或训练整机产品与合作伙伴提供的服务器设备、网络设备与存储设备结合，并配备公司的集群管理软件组成的数据中心集群。 • 公司还为人工智能应用部署技术能力相对薄弱的客户搭建成熟专业的应用管理平台，提供系统运维、资源调度、应用管理等功能。

3.2 云端产品：芯片产品纵向持续迭代性能

产品系列	推出时间	批量出货时间	细分产品	制程	计算精度支持	峰值算力	内存类型	内存容量	最大热功耗
思元100 (MLU100) 芯片	2018年	2019年	-	-	-	-	-	-	-
思元270 (MLU270) 芯片	2019年	2019年	MLU 270-S4 智能加速卡	-	INT16, INT8, INT4, FP32, FP16	128TOPS(INT8) 256TOPS(INT4), 64TOPS(INT16)	DDR4, ECC	16GB	70W
			MLU 270-F4 智能加速卡	-	INT16, INT8, INT4, FP32, FP16	128TOPS(INT8), 256TOPS(INT4) 64TOPS(INT16)	DDR4, ECC	16GB	150W
思元290 (MLU290) 芯片	2020年	2021年	MLU 290-M5 智能加速卡	7nm	-	自适应精度训练算力 512 TOPS(INT8), 256TOPS(INT16), 64TOPS(CINT32)	HBM2高带 宽内存	32GB	350W
思元370 (MLU370) 芯片 及云端智能加速卡	2021年	2022年	MLU 370-S4 智能加速卡	7nm	FP32, FP16, BF16, INT16, INT8, INT4	192TOPS(INT8), 96 TOPS(INT16) 72 TFLOPS (FP16), 72 TFLOPS (BF16), 18 TFLOPS (FP32)	LPDDR5	24GB	75W
			MLU 370-X4 智能加速卡	7nm	FP32, FP16, BF16, INT16, INT8, INT4	256 TOPS(INT8), 128 TOPS (INT16), 96 TFLOPS (FP16), 96 TFLOPS (BF16), 24 TFLOPS (FP32)	LPDDR5	24GB	150W
			MLU 370-X8 智能加速卡	7nm	FP32, FP16, BF16, INT16, INT8, INT4	256 TOPS(INT8), 128 TOPS (INT16), 96 TFLOPS (FP16), 96 TFLOPS (BF16), 24 TFLOPS (FP32)	LPDDR5	48GB	250W

3.2 云端产品：客户主要为服务器厂商与互联网厂商

- 公司云端产品客户主要为服务器厂商与互联网厂商，也拓展金融、运营商、能源等客户。公司云端产品线前五大客户，包括中科可控、浪潮信息等服务器厂商及阿里、百度等互联网公司，目前公司已完成服务器厂商适配并与主要服务器厂商建立长期深入合作关系。公司逐步拓展金融、通信运营商、能源等行业客户，也发展智慧粮仓、智慧矿山、智慧物流等场景的业务落地。

图：公司与多行业的客户进行技术适配和商务洽谈

客户行业	具体客户	具体内容
互联网企业	阿里、腾讯 百度	<ul style="list-style-type: none">• 思元 370 芯片及加速卡与数家头部互联网企业在视觉、语音、图文识别等场景下完成适配工作后，进入了批量销售环节，2022年实现收入突破。• 大模型领域，公司与国内头部的算法公司形成紧密的技术和产品合作；公司也与智象未来、百川智能等客户进行适配。
金融	工商银行 建设银行	<ul style="list-style-type: none">• 2022年，公司持续扩展云端产品在多家头部银行的业务，其中 MLU370-X8 训练产品已纳入某头部银行的采购框架并于报告期内实现了小批量采购。• 2023年，公司与多家大型银行进行了语言大模型的测试，并正在推动金融领域的应用落地。目前交通银行已选择嵌入公司智能加速卡的 GPU 服务器为指定选型。
服务器厂商	中科可控 浪潮信息等	<ul style="list-style-type: none">• 已应用到国内主流服务器厂商的产品中，并已实现量产出货。
头部运营商	...	<ul style="list-style-type: none">• 寒武纪依托集采入围，继续赋能运营商集团及下属省专公司常用的人工智能业务。未来，寒武纪将持续助力三大运营商共同赋能更多业务场景的人工智能应用落地，向“AI+”延伸拓展。
智能电网	南方电网 国电通	<ul style="list-style-type: none">• ...
轨交行业	国铁集团 恒远智能	<ul style="list-style-type: none">• ...

3.2.1 思元370：产品成熟度持续提升

- 思元370产品成熟度提升。2022年3月21日，寒武纪正式发布新款训练加速卡MLU370-X8。MLU370-X8搭载双芯片四芯粒思元370，集成寒武纪MLU-Link™多芯互联技术，主要面向训练任务，在业界应用广泛的YOLOv3、Transformer等训练任务中，8卡计算系统的并行性能平均达到350W RTX GPU的155%，并已实现商业化部署。与此前发布的 MLU370-S4、MLU370-X4 相比，MLU370-X8定位为训练加速卡。

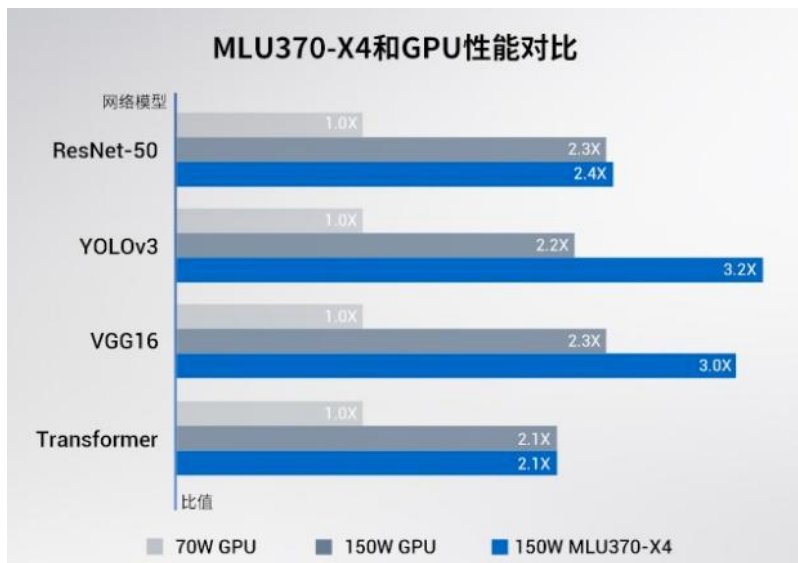
图：思元370芯片性能

	MLU370-S4	MLU370-X4	MLU370-X8
面向市场	机器视觉/推理任务	互联网行业等推理任务，或训推一体场景	训练任务
市场特点	整机计算密度较高，单卡算力需求适中	单卡算力需求较高	单卡算力需求较高，互联带宽需求高
工艺制程	7nm	7nm	7nm
内存容量	24GB	24GB	48GB
最大热功耗	75W	150W	250W
算力	192 TOPS (INT8) 96 TOPS (INT16) 72 TFLOPS (FP16) 72 TFLOPS (BF16) 18 TFLOPS (FP32)	256 TOPS (INT8) 128 TOPS (INT16) 96 TFLOPS (FP16) 96 TFLOPS (BF16) 24 TFLOPS (FP32)	256 TOPS (INT8) 128 TOPS (INT16) 96 TFLOPS (FP16) 96 TFLOPS (BF16) 24 TFLOPS (FP32)
互联带宽	307.2 GB/s	307.2 GB/s	614.4 GB/s

图：MLU370-S4/S8智能加速卡



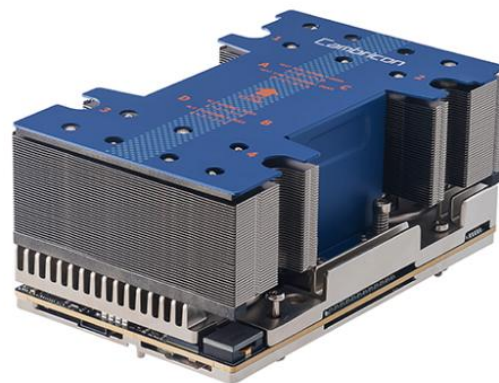
图：MLU370-X4和GPU性能对比



3.2.2 思元290+590：完善训练产品布局

- 训练端290芯片完善训练产品布局。思元290是寒武纪首颗训练芯片，采用MLUv02扩展架构，使用台积电7nm先进制程工艺，集成460亿个晶体管，采用HBM2内存提供AI训练所需的高内存带宽，全面支持AI训练、推理或混合型人工智能计算加速任务。
- 新品思元590芯片在研待发布，性能持续提升。2022年9月1-3日，在2022世界人工智能大会，寒武纪董事长、总经理陈天石博士透露在研新品信息，包括全新一代云端智能训练芯片思元590。其采用MLUarch05全新架构，实测训练性能较在售产品有显著提升，它提供了更大内存容量和更高内存带宽，PCIe接口也较上代实现升级。

图：MLU 290 – M5加速卡产品



搭载思元290
采用OAM设计
支持单机八卡机内互联
多机多卡机间互联

图：MLU290 的 MLUv02 架构进行了多项扩展



图：思元270与思元290理论峰值性能对比



3.3 边缘端：思元220累计出货百万片

- 边缘计算是一种新型计算范式，在终端和云端之间的设备上配备适度的计算能力。一方面可有效弥补终端设备计算能力不足的劣势，另一方面可缓解云计算场景下数据安全、隐私保护、带宽与延时等潜在问题，可应用于智能制造、智能零售、智能教育、智能家居、智能电网等众多领域。
- 公司边缘芯片主要为思元220，截至2023年累计出货百万片。2019年11月，公司正式发布思元220边缘智能芯片及相应的M.2加速卡；截至2023年底，思元220累计销量突破百万片。公司自2020年边缘端产品量产后即开始与行业头部客户公司A展开合作，其对边缘产品需求较高，因此公司边缘产品收入主要来自公司A。
- 1) 短期：公司边缘产品线业务对客户A存在一定依赖。2022年，因客户自身需进行库存调控消化，导致公司边缘产品线在主要客户处的收入有所下降；2) 长期：随着边缘计算应用场景（如工控机、边缘服务器、机器人、智能电视等方向）的不断发展，边缘计算整体市场空间将不断扩大，公司将进一步开拓更为丰富的客户资源，少数客户依赖将有所降低。

图：公司边缘智能芯片思元 220 及相应的 M.2 加速卡

产品型号	产品概况	产品展示
思元220 (MLU220)	<ul style="list-style-type: none"> 思元220使用公司自研的MLUv02指令集，面向人工智能边缘推理任务 基于台积电16nm先进工艺制造，芯片面积94.8mm²，集成了丰富的视频图像编解码硬件单元和外设接口 在1GHz的主频下，理论峰值性能为32TOPS (INT4)、16TOPS (INT8)、8TOPS (INT16)，芯片典型功耗小于10瓦 在8.25瓦的M.2加速卡整体功耗限制下，理论峰值性能为16TOPS (INT4)、8TOPS (INT8)、4TOPS (INT16) 	<p>芯片</p> <p>加速卡</p>

图：思元 220 系列 MLU220-M.2 边缘端智能加速卡规格

参数	规格
型号	MLU220-M.2
AI 性能	8TOPS (INT8)
内存	LPDDR4x 64 bits
编解码能力	H.264, HEVC (H.265), VP8, VP9;
图片解码	JPEG, 最大图片分辨率8192 × 8192
接口规格	M.2 2280, B+M key (PCIe3.0 × 2)
结构尺寸	长80mm, 宽22mm, 高7.2mm (无散热) / 21.3mm (带散热)
功耗	8.25W
散热方式	被动散热

3.4 IP授权及软件：已集成超1亿台智能手机及其他智能终端设备

- IP授权及软件业务包括 IP 授权和基础系统软件平台。IP 授权是将公司研发的智能处理器 IP 等知识产权授权给客户在其产品中使用，终端智能处理器通常不是以独立芯片的形式存在，而是作为一个模块集成于终端设备的 SoC 芯片当中。公司的终端智能处理器 IP 产品主要有 1A、1H 和 1M 系列，已集成于超过 1 亿台智能手机及其他智能终端设备中。

表：公司 IP 产品技术指标与特点情况

产品型号	技术指标	产品特点
寒武纪1A (2016)	<ul style="list-style-type: none"> • 1GHz主频下，非稀疏理论峰值性能0.5TOPS (FP16)，稀疏等效理论峰值性能2TOPS (FP16) 	<ul style="list-style-type: none"> • 全球首款商用终端智能处理器IP产品，可支持视觉、语音和自然语言处理等消费电子领域的人工智能应用；根据客户的公开宣传信息，搭载寒武纪1A的某旗舰手机芯片在人工智能应用上达到了4核CPU25倍以上的性能和50倍以上的能效，采用该手机芯片的旗舰手机产品每分钟可识别2,005张图片。
寒武纪1H (2017)	<ul style="list-style-type: none"> • 子型号寒武纪1H8：1GHz主频下，非稀疏理论峰值性能1TOPS (INT8) • 子型号寒武纪1H16：1GHz主频下，非稀疏理论峰值性能0.5TOPS (FP16) 或 1TOPS (INT8)，稀疏等效理论峰值性能2TOPS (FP16) 	<ul style="list-style-type: none"> • 寒武纪1H的功耗和面积等指标较上一代产品有显著提升，支持双核模式，并增加了对8位定点 (INT8) 人工智能运算的支持。根据客户的公开宣传信息，搭载寒武纪1H的某旗舰手机芯片，每分钟可识别4,500张图片，是上一代产品的2.2倍。
寒武纪1M (2018)	<ul style="list-style-type: none"> • 提供三种尺寸的配置，在1GHz主频下，理论峰值性能分别为2TOPS (INT8)、4TOPS (INT8)、8TOPS (INT8) 	<ul style="list-style-type: none"> • 寒武纪1M针对7nm等先进工艺作了专门优化，进一步提升了处理器性能和能效；提供不同性能档位的处理器配置，支持多核模式；在业界率先支持定点化训练，可在终端支持人工智能训练任务。

注 1：非稀疏理论峰值性能代表处理非稀疏深度学习模型的理论最高性能，稀疏等效理论峰值性能代表处理稀疏深度学习模型的等效理论最高性能。

注 2：INT8 代表 8 位定点运算，FP16 代表 16 位浮点运算。

3.4 IP授权及软件：基础系统软件平台持续改善

- **基础系统软件平台：**Cambricon Neuware 是公司打造云边端统一的人工智能开发生态的核心部件。公司为云边端全系列智能芯片与处理器产品提供统一的平台级基础系统软件 Cambricon Neuware（包含软件开发工具链等），打破了不同场景之间的软件开发壁垒，兼具高性能、灵活性和可扩展性的优势。
 - ✓ **推理加速引擎MagicMind：**2022 年公司持续完善了推理加速引擎 MagicMind 及其周边生态。在功能上，算子支持数量持续增加；在模型部署优化上，支持了 QAT模型直接部署，并完整支持 Dynamic Shape等特性；在性能上，MagicMind 在多个领域的典型模型上，均取得不弱于同档位友商产品的性能表现。
 - ✓ **训练软件平台：**研发和改进工作持续进行。在功能上支撑公司新的硬件平台，支撑新的PyTorch版本，算子覆盖度达到 80%，TensorFlow算子数量及交付网络模型均增加，支持了包括 GPT 类语言模型在内的多种主流模型的分布式训练。

图：Cambricon NeuWare 软件栈



图：寒武纪训练软件平台



图：推理加速引擎MagicMind



第四章 盈利预测

4.1 盈利预测及投资建议

- 公司为国内领先人工智能芯片企业，未来有望受益大模型持续推进带来人工智能芯片需求。我们预测公司2024-2026年营业收入分别为14.00/24.93/34.84亿元，分别同比+97%/+78%/+40%。分业务看：
 - 1) **智能计算集群系统**：公司云端产品产能性能较优，或将受益于国内智算产业发展，实现营收增长，我们预计此业务2024-2026年营业收入分别为12.09/21.76/30.47亿元，分别同比+100%/+80%/+40%，毛利率保持相对稳定在72%。
 - 2) **云端产品线**：受益于国内大模型产业发展，以及训推卡需求提升，我们预计此业务2024-2026年营业收入分别为1.72/2.93/4.10亿元，分别同比+90%/+70%/+40%，毛利率保持相对稳定、为62%左右。
 - 3) **边缘产品线**：受益AI大模型升级发展，我们预计此业务2024-2026年营收分别为0.14/0.18/0.21亿元，同比+30%/+25%/+18%，毛利率相对稳定在56%。
 - 4) **IP授权及软件**：我们认为此业务因客户自研以及更新换代等原因将维持低位。

表：寒武纪盈利预测

收入（百万元）	2022	2023	2024E	2025E	2026E
智能计算集群系统	459	605	1209	2176	3047
增速	1%	32%	100%	80%	40%
毛利率	70%	71%	72%	72%	72%
云端产品线	220	91	172	293	410
增速	174%	-59%	90%	70%	40%
毛利率	63%	61%	62%	62%	62%
边缘产品线	38	11	14	18	21
增速	-78%	-71%	30%	25%	18%
毛利率	31%	56%	56%	56%	56%
IP授权及软件	1	0	0	0	0
增速	-83%	-79%	30%	24%	17%
毛利率	100%	100%	100%	100%	100%
其他业务	12	3	4	6	7
增速	278%	-73%	40%	30%	20%
毛利率	36%	47%	47%	47%	47%
收入总计	729	709	1400	2493	3484
增速	1%	-3%	97%	78%	40%
毛利率	65%	69%	70%	70%	71%

4.1 盈利预测及投资建议

- 公司为智能芯片领域全球知名的新兴公司，能提供云边端一体、软硬件协同、训练推理融合、具备统一生态的系列化智能芯片产品和平台化基础系统软件，有望受益于大模型发展带来的算力需求，以及公司产品不断迭代。我们预计公司2024~2026年分别实现营收14.00/24.93/34.84亿元，对应PS为53.59/30.10/21.53X，首次覆盖，给予“增持”评级。

表：公司费用率与归母净利润预测

预测指标	2023A	2024E	2025E	2026E
营业收入（百万元）	709	1400	2493	3484
归母净利润（百万元）	-848	-441	18	250
摊薄每股收益（元）	-2.04	-1.06	0.04	0.60
估值	2023A	2024E	2025E	2026E
市盈率（倍）	—	—	4,089.67	299.90
市销率（倍）	79.26	53.59	30.10	21.53

表：可比公司估值表（2024年6月5日）

股票代码	股票名称	收盘价	营收（亿元）				PS		
		6月5日	2023	2024E	2025E	2026E	2024E	2025E	2026E
688041.SH	海光信息	70.87	60.12	81.79	110.42	147.05	20.14	14.92	11.20
688047.SH	龙芯中科	98.51	5.06	7.51	10.45	13.59	52.61	37.81	29.07
300474.SZ	景嘉微	67.57	7.13	12.37	18.24	22.34	25.05	16.99	13.87
688256.SH	寒武纪-U	180.10	7.09	14.00	24.93	34.84	53.59	30.10	21.53

寒武纪盈利预测表

证券代码： 688256

股价： 180.10

投资评级： 增持(首次覆盖)

日期： 20240605

资产负债表 (百万元)	2023A	2024E	2025E	2026E	利润表 (百万元)	2023A	2024E	2025E	2026E	每股指标与估值	2023A	2024E	2025E	2026E
现金及现金等价物	4654	3631	2745	2325	营业收入	709	1400	2493	3484	每股指标				
应收款项	644	1223	2353	3165	营业成本	219	419	739	1020	EPS	-2.07	-1.06	0.04	0.60
存货净额	99	372	614	739	营业税金及附加	4	6	12	18	BVPS	13.56	12.64	12.68	13.28
其他流动资产	250	324	548	844	销售费用	82	126	199	244	估值				
流动资产合计	5648	5550	6260	7073	管理费用	154	294	499	610	P/E	—	—	4089.7	299.9
固定资产	142	147	90	26	财务费用	-45	0	0	0	P/B	10.0	14.3	14.2	13.6
在建工程	109	124	146	172	其他费用/(-收入)	1118	1750	2468	3380	P/S	79.3	53.6	30.1	21.5
无形资产及其他	290	277	249	259	营业利润	-876	-454	18	256	财务指标	2023A	2024E	2025E	2026E
长期股权投资	230	240	252	261	营业外净收支	1	1	1	1	盈利能力				
资产总计	6418	6337	6996	7791	利润总额	-875	-453	19	258	ROE	-15%	-8%	0%	5%
短期借款	0	0	0	0	所得税费用	3	2	0	-1	毛利率	69%	70%	70%	71%
应付款项	237	391	724	1019	净利润	-878	-455	19	259	期间费率	27%	30%	28%	25%
合同负债	0	0	1	1	少数股东损益	-30	-14	1	9	销售净利率	-120%	-31%	1%	7%
其他流动负债	226	413	719	960	归属于母公司净利润	-848	-441	18	250	成长能力				
流动负债合计	463	804	1444	1980	现金流量表 (百万元)	2023A	2024E	2025E	2026E	收入增长率	-3%	97%	78%	40%
长期借款及应付债券	0	0	0	0	经营活动现金流	-596	-925	-990	-644	利润增长率	32%	48%	104%	1,264%
其他长期负债	0	0	0	0	净利润	-848	-441	18	250	营运能力				
长期负债合计	225	202	202	202	少数股东损益	-30	-14	1	9	总资产周转率	0.12	0.22	0.37	0.47
负债合计	689	1006	1646	2182	折旧摊销	330	283	292	271	应收账款周转率	1.01	1.50	1.39	1.26
股本	417	417	417	417	公允价值变动	0	0	0	0	存货周转率	1.13	1.78	1.50	1.51
股东权益	5730	5331	5350	5609	营运资金变动	-249	-570	-955	-697	偿债能力				
负债和股东权益总计	6418	6337	6996	7791	投资活动现金流	425	563	105	224	资产负债率	11%	16%	24%	28%
					资本支出	-100	-279	-229	-244	流动比	12.19	6.90	4.33	3.57
					长期投资	-85	670	-12	-10	速动比	11.49	6.10	3.60	2.84
					其他	610	172	346	477					
					筹资活动现金流	1657	19	0	0					
					债务融资	0	-18	0	0					
					权益融资	1804	56	0	0					
					其它	-147	-19	0	0					
					现金净增加额	1486	-343	-885	-420					

- 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) **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公司的业绩不及预期。
- 6) **公司新品研发不及预期**：公司新品研发是公司未来持续经营的重要支撑，若研发进度以及产品性能不及预期，或增加公司未来经营以及业绩释放不及预期的风险。

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R4，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