



买入 (首次)

所属行业: 电子/半导体
当前价格(元): 153.70

证券分析师

陈海进

资格编号: S0120521120001

邮箱: chenhj3@tebon.com.cn

研究助理

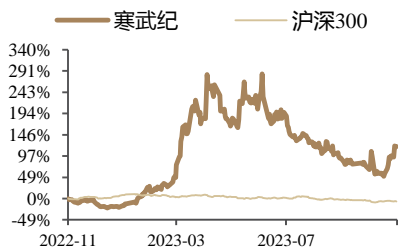
徐巡

邮箱: xuxun@tebon.com.cn

谢文嘉

邮箱: xiewj3@tebon.com.cn

市场表现



沪深 300 对比	1M	2M	3M
绝对涨幅(%)	29.82	13.42	-0.56
相对涨幅(%)	31.31	17.10	5.89

资料来源: 德邦研究所, 聚源数据

相关研究

寒武纪 (688256.SH): 国产 AI 龙头, 云边终端三位协同

投资要点

- 公司概况:** 全球知名 AI 芯片新星, 构建“云、边、终”多位一体新生态。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。2022 年公司实现营收 7.29 亿元, 同比增长 1.11%, 受益于云端产品思元 290、370 系列成功导入多家头部客户。边缘端和 IP 授权产品线营收下滑较为严重。22 年公司期间费用率总体上升, 研发费用率上涨最为显著。22 年公司实现归母净利润-12.57 亿元, 较上年同期扩大 4.32 亿元, 主要系研发费用的增长和部分云边端产品销量不及预期。
- 行业格局:** AI 浪潮拉动算力需求, 人工智能芯片前景广阔。目前 AIGC、高端服务器、高性能计算、5G 等应用快速发展, 这些应用离不开高制程算力芯片的底层支撑。根据我们测算, 在理想条件下通过租用线上算力完成一次参数量为 1750 亿的 GPT-3 模型训练需要 1024 块 A100 80G, 时长约 25 天, 训练成本约为 93 万美元。但实际成本和时间往往会更长全球 GPU 呈现“一超一强”的竞争格局, 根据 IDC 数据, 2021 英伟达在企业级 GPU 市场中占比 91.4%, AMD 占比 8.5%。目前中国加速计算服务器市场需求主要被 GPU 所承接。但 AI 智能芯片定制化程度更高, 未来有望凭借更灵活的架构设计和更高的算力, 成为市场主流。
- 公司业务:** 云端芯片为基, 云边终端三位协同。公司产品线涵盖云端、边缘端和 IP 授权及软件几大板块, 各产品线持续高端化迭代。其中, 云端智能芯片为业务核心, 同时延伸发展训练整机与智能计算集群, 主力产品思元 370 采用 7nm 先进工艺和 Chiplet 技术, 较前代产品算力提升显著。边缘产品思元 220 已进入生命周期后期, 受主力客户需求变化明显。终端 IP 授权及软件为公司早期产品, 目前逐渐成为业务中间形态, 服务于公司云边端的生态建设。
- 投资建议:** 公司为国内 AI 芯片龙头, “云、边、终端”全链路协同覆盖。随着 AI 芯片市场规模的快速增长, 公司有望依托完备的云、边、终端全面产品布局 and AI 生态建设的持续推进, 逐步抢占市场份额。我们预计公司将在 2023 年至 2025 年实现收入 11.26/15.38/20.44 亿元, 归母净利润-8.7/-7.2/-4.9 亿元, 对应当前 PS 估值 57/42/31 倍。首次覆盖, 给予“买入”评级。
- 风险提示:** AI 需求不及预期风险, 客户集中度较高风险, 供应链稳定相关风险。

股票数据

总股本(百万股):	416.59
流通 A 股(百万股):	266.45
52 周内股价区间(元):	54.56-267.00
总市值(百万元):	63,301.53
总资产(百万元):	6,581.32
每股净资产(元):	13.94

资料来源: 公司公告

主要财务数据及预测

	2021	2022	2023E	2024E	2025E
营业收入(百万元)	721	729	1,126	1,538	2,044
(+/-)YOY(%)	57.1%	1.1%	54.5%	36.6%	32.9%
净利润(百万元)	-825	-1,257	-869	-722	-489
(+/-)YOY(%)	-89.9%	-52.3%	30.9%	16.9%	-32.2%
全面摊薄 EPS(元)	-2.06	-3.14	-2.08	-1.73	-1.17
毛利率(%)	62.4%	65.8%	66.8%	67.8%	68.7%
净资产收益率(%)	-14.0%	-25.9%	-21.4%	-21.5%	-16.9%

资料来源: 公司年报 (2021-2022), 德邦研究所

备注: 净利润为归属母公司所有者的净利润

内容目录

1. 全球知名 AI 芯片新星，营收稳健增长.....	5
1.1. 中国 AI 芯片龙头，产品矩阵扩充迅速	5
1.2. 实控人研发经验丰富，公司与中科院计算所关系密切	5
1.3. 营收稳健增长，研发投入占比持续提升	6
2. 行业格局：AI 浪潮拉动算力需求，人工智能芯片前景广阔	8
2.1. 大模型需海量算力支撑，AI 服务器需求快速提升	8
2.2. 人工智能芯片定向针对 AI 领域，发展潜力广阔	9
3. 公司业务：云端芯片为基，云边终端三位协同	11
3.1. 云端智能芯片及加速卡：先进制程+先进封装，助力业务彰显技术实力	12
3.2. 边缘智能芯片及加速卡：立足高算力领域，新品推出激发新老客户放量	15
3.3. 终端 IP 授权及软件：服务生态建设，构建完备业务矩阵	16
4. 盈利预测与投资建议	17
5. 风险提示	18

图表目录

图 1: 公司产品矩阵迅速扩充.....	5
图 2: 公司前五大股东股权占比.....	6
图 3: 公司营业收入情况.....	6
图 4: 公司营收结构情况.....	6
图 5: 公司毛利率情况.....	7
图 6: 公司期间费用率情况.....	7
图 7: 公司研发投入情况.....	7
图 8: 公司归母净利润情况.....	8
图 9: GPT-3 训练时间及成本测算.....	8
图 10: 2022~2026 年 AI 服务器出货量 (单位: 万台).....	9
图 11: 2021 年企业级 GPU 市场份额.....	9
图 12: 中国智能算力规模及预测.....	9
图 13: 中国人工智能服务器工作负载及预测.....	9
图 14: GPU 与人工智能芯片对比.....	10
图 15: 中国加速计算服务器市场规模及增速.....	10
图 16: 2022 年中国加速服务器芯片类型市场份额占比.....	10
图 17: 英伟达 A100 与 H100 算力性能情况.....	11
图 18: 寒武纪人工智能芯片与谷歌 TPU 对比.....	11
图 19: 云、边、终端三类产品性能与应用场景对比.....	12
图 20: 公司主要产品类型与推出时间.....	12
图 21: 公司主要产品类型与推出时间.....	13
图 22: 公司客户情况与竞争对手情况.....	13
图 23: 公司云端智能芯片性能对比.....	14
图 24: 公司思元 370 产品市场竞争优势.....	14
图 25: 寒武纪首颗训练芯片思元 290.....	15
图 26: 寒武纪智能加速卡 MLU290-M5.....	15
图 27: 边缘智能物联网市场竞争格局.....	15
图 28: 思元 220 芯片的具体情况.....	16
图 29: 思元 220 系列 MLU220-M.2 边缘端智能加速卡规格参数.....	16
图 30: 公司终端智能处理器 IP 产品种类.....	17
图 31: Cambricon Neuware 框架结构.....	17
图 32: 寒武纪营收及毛利率预测 (单位: 亿元).....	18

图 33: 可比公司估值 (2023/10/29)18

表 1: 公司三大产品线简介5

1. 全球知名 AI 芯片新星，营收稳健增长

1.1. 中国 AI 芯片龙头，产品矩阵扩充迅速

寒武纪是中国最具代表性的智能芯片厂商之一。中科寒武纪科技股份有限公司成立于 2016 年，总部位于北京，是智能芯片领域全球知名的新兴公司。2021 年 3 月，公司上榜《EETimes》评选的“AI 芯片公司(AI CHIP) TOP 10”榜单。

公司专注于人工智能芯片产品的研发与技术创新。公司的主营业务是应用于各类云服务器、边缘计算设备、终端设备中人工智能核心芯片的研发、设计和销售，为客户提供丰富的芯片产品与系统软件解决方案。自公司成立以来，产品矩阵迅速扩充，从寒武纪 1A 处理器开始，陆续推出了思元、玄思等芯片/加速卡/加速器产品，发展至目前形成云端产品线、边缘产品线、IP 授权及软件三条产品线的产品矩阵。公司产品广泛应用于服务器厂商和产业公司，面向互联网、金融、交通、能源、电力和制造等领域的复杂 AI 应用场景提供充裕算力，推动人工智能赋能产业升级。

图 1：公司产品矩阵迅速扩充



资料来源：公司 2022 年年报，德邦研究所

表 1：公司三大产品线简介

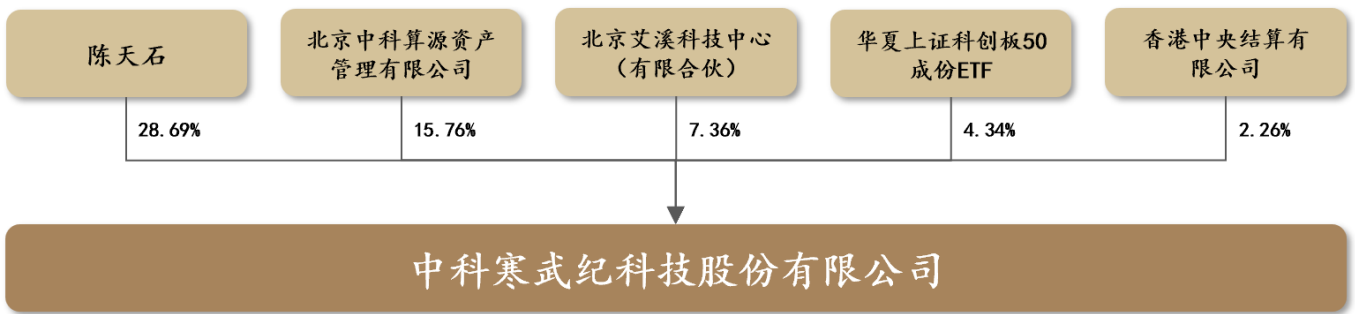
产品线	产品类型	产品特点
云端产品线	云端智能芯片及加速卡	为云计算和数据中心场景下的人工智能应用程序提供高计算密度、高能效的硬件计算资源，支撑该类场景下复杂度和数据吞吐量高速增长的人工智能处理任务
	训练整机	提供计算集群中的单体训练服务器
边缘产品线	边缘智能芯片及加速卡	有效弥补终端设备计算能力不足的劣势，缓解云计算场景下数据隐私、带宽与延时等潜在问题。
IP 授权及软件	IP 授权	将公司研发的智能处理器 IP 等知识产权授权给客户在其产品中使用
	基础系统软件平台	打破了不同场景之间的软件开发壁垒，兼具灵活性和可扩展性的优势，无须繁琐的移植即可让同一人工智能应用程序便捷高效地运行在公司云边缘端系列化芯片与处理器产品之上

资料来源：公司 2022 年年报，德邦研究所

1.2. 实控人研发经验丰富，公司与中科院计算所关系密切

截至 23 年三季报，公司实控人为公司董事长、总经理陈天石，持股比例为 28.69%，系公司第一大股东。陈天石博士曾在中科院计算所担任研究员（正高级职称），在人工智能及处理器芯片领域从事基础科研工作十余年，积累了坚实的理论功底及研发经验。中科院计算技术研究所独资公司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%。

图 2：公司前五大股东股权占比



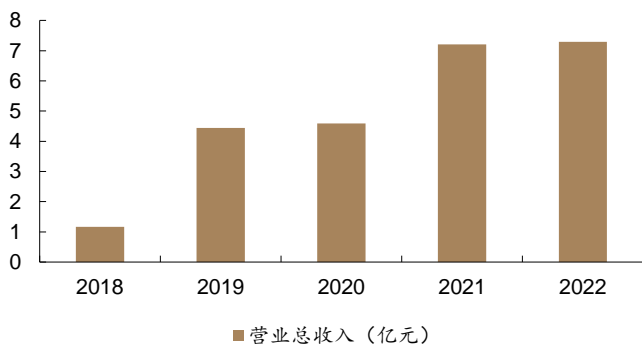
资料来源：公司 2023 年三季报，德邦研究所

1.3. 营收稳健增长，研发投入占比持续提升

公司营收稳健增长。在国内外经济形势均面临下行压力，公司坚持自主研发、提升产品成熟度、拓宽潜在市场、打造生态建设，营收实现稳健增长。2022 年公司实现营业收入 7.29 亿元，较上年同期增长 1.11%。

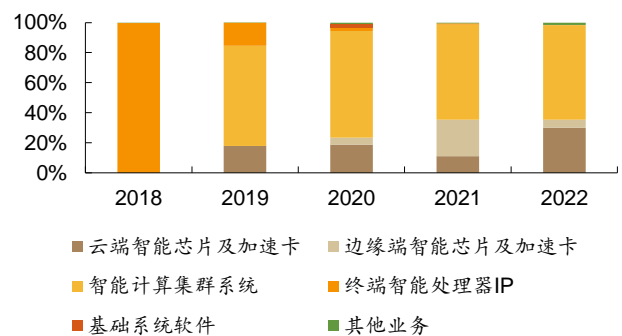
从营收结构来看，云端产品线收入大幅增长，边缘产品线、IP 授权及软件产品线收入下滑。2022 年，公司云端产品线单独对外销售部分贡献 21,944.89 万元的收入，占主营业务收入总额的 30.37%，思元 290、思元 370 系列加速卡和训练整机产品销量规模迅速上升，促使 2022 年云端产品线收入较上年同期增长 173.52%。因客户进行库存调控，未按预期进度下单，2022 年公司边缘产品线收入较上年同期下降 78.40%。公司成功中标南京智能计算中心项目（二、三期）等项目，保持智能计算集群系统业务开展的可持续性，2022 年公司智能计算集群系统业务收入较上年同期增长 0.64%。2022 年公司 IP 授权及软件产品线收入较上年同期减少 83.44%，一方面由于基础系统软件平台产品主要集成在智能计算集群系统中销售；另一方面由于 IP 授权业务前期客户按出货量为标准进行结算的收入随客户相关产品生命周期结束而减少。

图 3：公司营业收入情况



资料来源：公司公告，Wind，德邦研究所

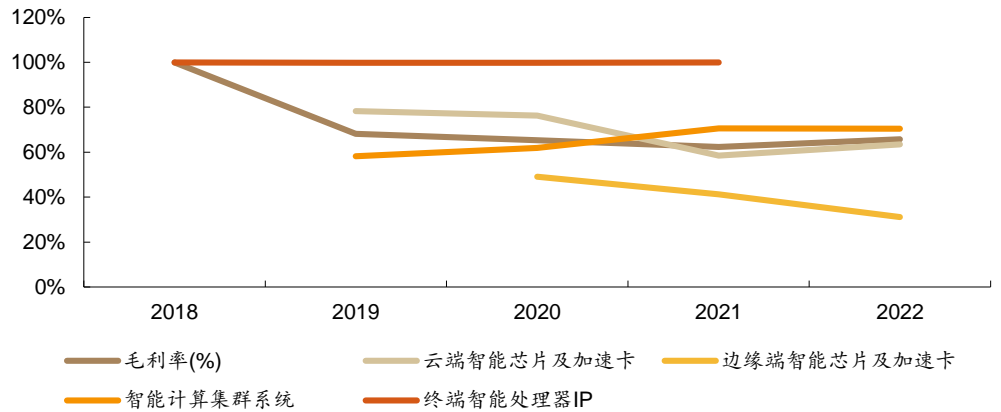
图 4：公司营收结构情况



资料来源：公司公告，Wind，德邦研究所

毛利率处于较高水平。2022 年公司主营业务综合毛利率为 65.85%，较上年增加 3.46 个百分点。主要原因有：(1) 云端产品线思元 290、思元 370 系列产品在多家头部企业完成产品导入，带动本期云端产品线业务收入的大幅增长；(2) 毛利率相对较低的边缘端产品在报告期内的收入占比下降。

图 5：公司毛利率情况

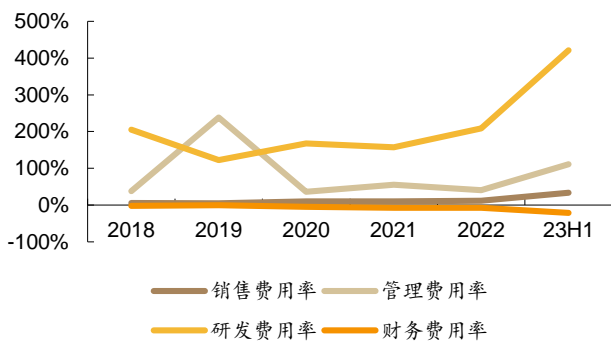


资料来源：公司公告，Wind，德邦研究所

期间费用率总体上升。2022 年公司销售费用率为 11.35%，较上年同期增加 1.44 个百分点，主要是销售人员增加导致职工薪酬增加、公司积极发力智能芯片市场推广及生态建设所致；管理费用率为 41%，系股份支付费用较高；财务费用率为-7.12%，主要系公司银行存款较多，银行存款利息收入较高所致。

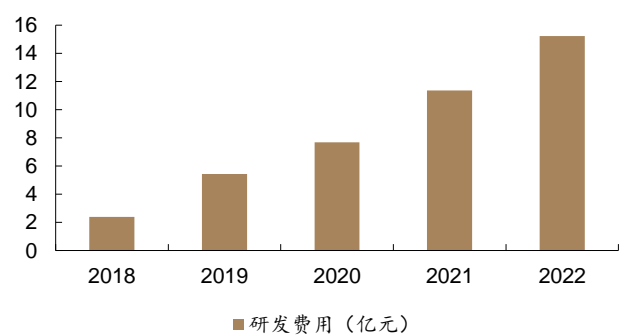
加注研发投入，研发属性强。高质量的研发投入是芯片行业实现长远发展的坚实基础，公司大力加注研发投入，2022 年研发投入总额为 15.23 亿元，较上年同期增加 3.87 亿元，同比增长 34.11%。主要原因系：(1) 半导体行业人才薪酬水平升高；(2) 新产品流片等相关费用增加；(3) 公司根据研发需求购置 IP、EDA 等无形资产以及研发设备等，导致无形资产摊销费用及固定资产折旧费用较上年同期增加。

图 6：公司期间费用率情况



资料来源：公司公告，Wind，德邦研究所

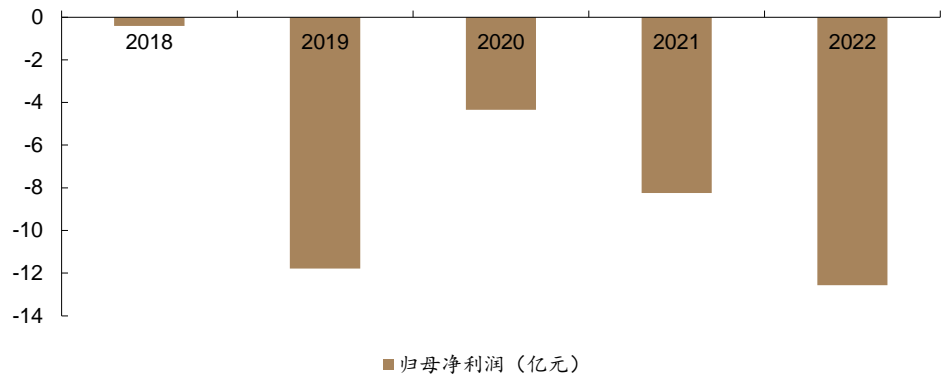
图 7：公司研发投入情况



资料来源：公司公告，Wind，德邦研究所

盈利端承压，亏损同比扩大。2022 年公司实现归母净利润-12.57 亿元，亏损金额较上年同期扩大 4.32 亿元，主要原因有：(1) 为确保产品高质量迭代，公司持续加大研发投入，研发费用增加；(2) 公司增加战略备货、处于生命周期末期的云端产品销量减少以及边缘智能芯片产品销量不及预期，导致存货及库龄增加，从而使 2022 年资产减值损失较上年同期有显著增加；(3) 公司对个别大额应收账款进行单项计提，从而使报告期信用减值损失较上年同期有显著增加。

图 8：公司归母净利润情况



资料来源：公司公告，Wind，德邦研究所

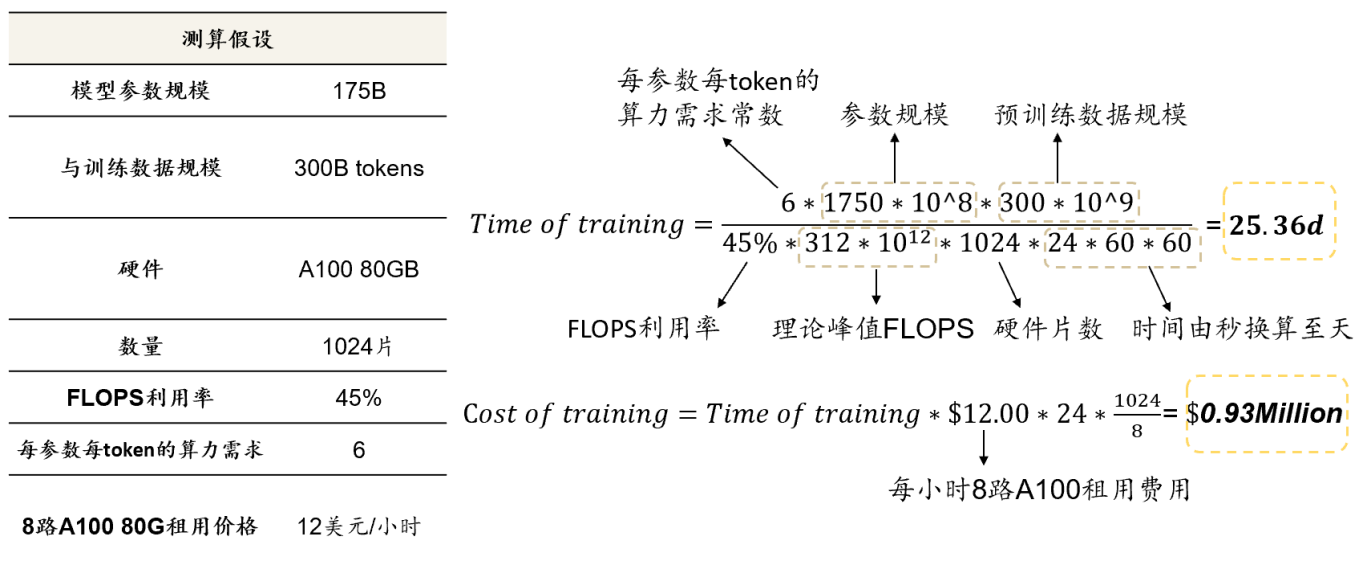
2. 行业格局：AI 浪潮拉动算力需求，人工智能芯片前景广阔

2.1. 大模型需海量算力支撑，AI 服务器需求快速提升

AIGC 需求兴起，产业发展依赖高制程芯片的底层算力支持。目前 AIGC、高端服务器、高性能计算、5G 等应用快速发展，这些应用离不开高制程算力芯片的底层支撑。以 AIGC 大模型为例，GPT-3 的模型规模达 1750 亿个参数，使用单块英伟达 V100GPU 进行训练需要 355 年，对算力芯片提出了海量的需求。

而以 1024 块英伟达 A100 芯片测算，GPT-3 的训练时长为 25 天以上。根据我们测算，在理想条件下通过租用线上算力完成一次参数量为 1750 亿的 GPT-3 模型训练需要 1024 块 A100 80G，时长约 25 天，训练成本约为 93 万美元。但实际成本和时间往往会更长，比如模型几乎不可能一次顺利完成训练等等。在训练算力需求之外，推理算力需求也会随着模型的商用流行度以及吸引的活跃用户数持续增加。我们认为，数量快速提升的大模型训练需求和大模型推理需求有望大幅推动以 GPU 为主的算力芯片市场规模提升。

图 9：GPT-3 训练时间及成本测算

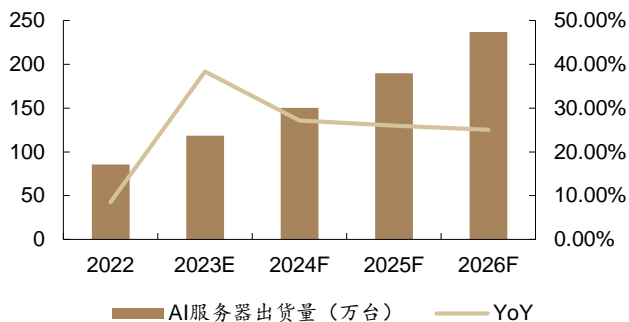


资料来源：Deepak Narayanan, Mohammad Shoeybi, et al. [Efficient Large-Scale Language Model Training on GPU Clusters Using Megatron-LM] arXiv:2104.04473, NVIDIA 官网, Lambdalabs 官网, 新智元公众号, CIBA 新经济公众号, 量子位公众号等, 德邦研究所

全球 AI 服务器出货迅速，AI 芯片市场主要为英伟达与 AMD 所占据。AIGC

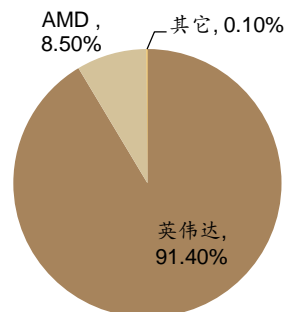
的发展带动 AI 服务器迅速增长，TrendForce 集邦咨询预计 23 年 AI 服务器出货量约 120 万台，同比+38.4%，占整体服务器出货量的比约为 9%，2022~2026 年 AI 服务器出货量 CAGR 将达 22%，而 AI 芯片 2023 年出货量将成长 46%。GPU 作为数据并行处理的核心，是 AI 服务器的核心增量。全球 GPU 呈现“一超一强”的竞争格局，根据 IDC 数据，2021 英伟达在企业级 GPU 市场中占比 91.4%，AMD 占比 8.5%。

图 10：2022~2026 年 AI 服务器出货量（单位：万台）



资料来源：TrendForce 集邦咨询，德邦研究所

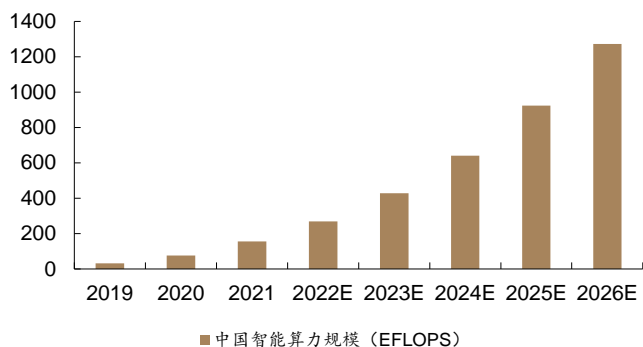
图 11：2021 年企业级 GPU 市场份额



资料来源：IDC，国际电子商情，德邦研究所

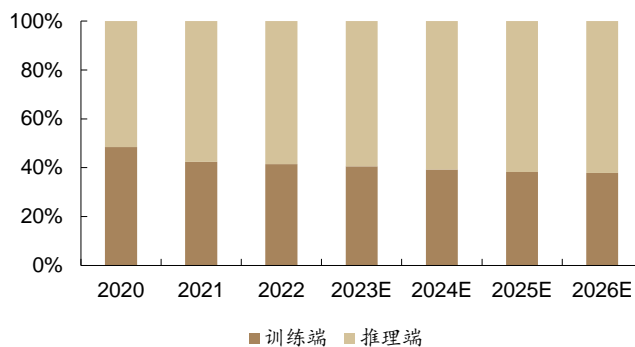
中国智能算力规模亦快速增长，CAGR 达 48%。根据 IDC 预测，2022 年中国智能算力规模将达到 268 EFLOPS，预计 2026 年将进入每秒十万亿亿次浮点计算 (ZFLOPS) 级别，达到 1271 EFLOPS，2022-2026 年 CAGR48%。伴随企业 AI 应用成熟度逐步递增，企业将把算力更多从训练端转移到推理端负载上，这意味着 AI 模型将逐步进入广泛投产模式，据 IDC 数据，2021 年中国数据中心用于推理的服务器的市场份额占比已经过半，达到 57.6%，预计到 2026 年，用于推理的工作负载将达到 62.2%。

图 12：中国智能算力规模及预测



资料来源：IDC，浪潮信息，《2022-2023 中国人工智能算力发展评估报告》，赣州科技公众号，德邦研究所

图 13：中国人工智能服务器工作负载及预测



资料来源：IDC，浪潮信息，《2022-2023 中国人工智能算力发展评估报告》，赣州科技公众号，德邦研究所

2.2. 人工智能芯片定向针对 AI 领域，发展潜力广阔

人工智能所需的算力芯片主要分为两类，一类是以 GPU、CPU 为主的传统芯片，一类是专门针对人工智能领域的人工智能芯片。在人工智能数十年的发展历程中，传统芯片曾长期为其提供底层计算能力。这些传统芯片包括 CPU、GPU 等，它们在设计之初并非面向人工智能领域，但可通过灵活通用的指令集或可重构的硬件单元覆盖人工智能程序底层所需的基本运算操作，从功能上可以满足人工智能应用的需求，但在芯片架构、性能、能效等方面并不能适应人工智能技术与应用的快速发展。而智能芯片是专门针对人工智能领域设计的芯片，包括通用型智能芯片与专用型智能芯片两种类型。

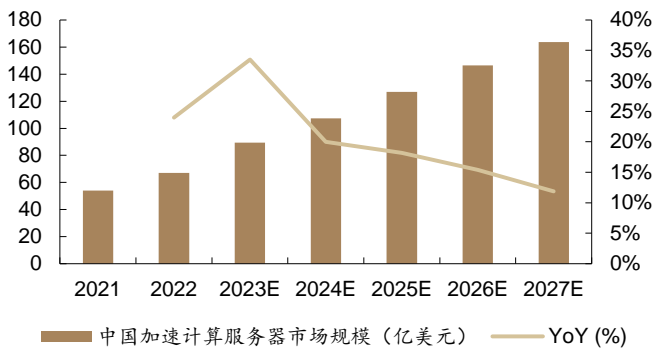
图 14: GPU 与人工智能芯片对比

芯片类型		云边缘普遍使用的处理芯片类型与代际	技术原理	技术优势与技术局限性	市场需求情况	在云边缘训练与推理场景的市场渗透率
传统芯片	GPU	(1) 云端主流产品为AMD和Nvidia产品, 主流工艺为7/12/16nm (2) 边缘端或终端主流产品为SoC集成的GPUIP, 主流厂商包括ARM、Imagination等	(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 (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	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	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
智能芯片	通用型智能芯片	(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 (Ascend310/910)、Google (TPUV1/V2/V3、TPUEdge) 等 (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	(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架构, 从而广泛支持多样化的人工智能算法和应用 (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 (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	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	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)	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。	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	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	应用细分市场需求大且分散, 成本敏感	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

资料来源: 寒武纪招股说明书, 德邦研究所

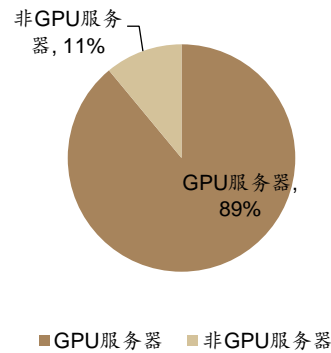
目前中国加速计算服务器市场需求主要被 GPU 所承接。根据 IDC 数据, 2022 年中国加速服务器市场规模达到 67 亿美元, 同比增长 24%, 其中 GPU 服务器依然是主导地位, 占据 89% 的市场份额。在 AI 加速应用部署的背景下, 2027 年中国加速服务器市场规模有望达到 164 亿美元, 2022-2027 年 CAGR20%。且由于 GPU 仍为中国主流 AI 加速芯片, 服务器 GPU 市场规模同样有望在 AI 的推动下持续增长。

图 15: 中国加速计算服务器市场规模及增速



资料来源: IDC, 德邦研究所

图 16: 2022 年中国加速服务器芯片类型市场份额占比



资料来源: IDC, 德邦研究所

GPU 为产业链最大增量, AI 趋势下单服务器需求量大幅提升。一般的普通服务器仅会配备单卡或双卡, 而 AI 服务器由于需要承担大量的计算, 一般配置四块或以上的 GPU 或人工智能芯片。且 AI 大模型在训练与推理时的计算量巨大, 中低端的处理器无法满足其运算需求。如在英伟达 DGX H100 中, 其配备 8 个 NVIDIA H100 GPU, 总 GPU 显存高达 640GB; 每个 GPU 配备 18 个 NVIDIA NVLink, GPU 之间的双向带宽高达 900GB/s。若以每个 NVIDIA H100 GPU 单价 4 万美元测算, DGX H100 的 GPU 价值量为 32 万美元, 为 AI 服务器中的最大增量。

图 17: 英伟达 A100 与 H100 算力性能情况

	A100	A100 稀疏	H100 SXM5 ¹	H100 SXM5 ¹ 稀疏	H100 SXM5 ¹ 相比 A100 的提速
FP8 Tensor Core	NA	NA	2000 TFLOPS	4000 TFLOPS	6.4 倍 (相比于 A100 FP16)
FP16	78 TFLOPS	NA	120 TFLOPS	NA	1.5 倍
FP16 Tensor Core	312 TFLOPS	624 TFLOPS	1000 TFLOPS	2000 TFLOPS	3.2 倍
BF16 Tensor Core	312 TFLOPS	624 TFLOPS	1000 TFLOPS	2000 TFLOPS	3.2 倍
FP32	19.5 TFLOPS	NA	60 TFLOPS	NA	3.1 倍
TF32 Tensor Core	156 TFLOPS	312 TFLOPS	500 TFLOPS	1000 TFLOPS	3.2 倍
FP64	9.7 TFLOPS	NA	30 TFLOPS	NA	3.1 倍
FP64 Tensor Core	19.5 TFLOPS	NA	60 TFLOPS	NA	3.1 倍
INT8 Tensor Core	624 TOPS	1248 TOPS	2000 TFLOPS	4000 TFLOPS	3.2 倍

资料来源: 英伟达官网, 德邦研究所

AI 智能芯片定制化程度更高, 未来有望凭借更灵活的架构设计和更高的算力, 成为市场主流。相比英伟达的 GPU, 寒武纪人工智能芯片更类似谷歌的 TPU。寒武纪产品与谷歌的 TPU, 都是通过对人工智能领域的计算特征和访存特征进行分析和抽象, 设计出的通用型智能芯片, 指令集、运算器架构和存储层次都非常适合智能算法, 从而在智能应用上的能效超过了传统 CPU、GPU。但两者在架构设计上, 仍有不同。

图 18: 寒武纪人工智能芯片与谷歌 TPU 对比

	相同点	不同点 (处理器架构)
寒武纪人工智能芯片	从技术原理上看, 两者都是通过对人工智能领域的计算特征和访存特征进行分析和抽象, 设计出的通用型智能芯片。指令集、运算器架构和存储层次都非常适合智能算法, 从而在智能应用上的能效超过了传统 CPU、GPU。	Google TPU 的核心是经典的脉动阵列机技术, 脉动阵列本身对于卷积类运算的效率较高, 但是对于相对低频的部分运算操作 (如全连接运算、激活运算) 的效率不高。
谷歌 TPU 产品		对于后者, Google TPU 引入了额外的硬件单元作为补充。 处理器架构: 运算、算数逻辑运算, 并在处理器中分别通过高维张量计算部件、向量计算部件、传统算术逻辑计算部件予以处理。 高维张量计算部件可高效支持卷积运算、全连接运算, 而向量计算部件则可以支持激活等运算, 传统算术逻辑计算部件则可以支持分支跳转等。

资料来源: 寒武纪招股说明书, 德邦研究所

3. 公司业务: 云端芯片为基, 云边终端三位协同

公司产品涉及“云、边、终端”, 全链路垂直协同布局。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 19：云、边、终端三类产品性能与应用场景对比

应用场景	芯片需求	典型计算能力	典型功耗	典型应用领域
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
终端	低功耗、高效率、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

资料来源：公司招股说明书，德邦研究所

注：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。

公司产品布局完备，持续像高端化迭代。公司面向云、边、端三大场景分别研发了三种类型的芯片产品，分别为云端智能芯片及加速卡、边缘智能芯片及加速卡、IP 授权及软件。公司产品由早期推出的“终端智能处理器 IP”逐渐向“云端智能芯片及加速卡”迭代，产品持续高端化迭代。

图 20：公司主要产品类型与推出时间

产品线	产品类型	寒武纪主要产品	推出时间
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
		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
IP 授权及软件	终端智能处理器IP	寒武纪1A处理器	2016年
		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

资料来源：公司 2022 年年度报告，德邦研究所

3.1. 云端智能芯片及加速卡：先进制程+先进封装，助力业务彰显技术实力

云端智能芯片为业务核心，延伸发展训练整机与智能计算集群。公司云端产品线的产品形态主要包括云端智能芯片和加速卡、训练整机及智能计算集群，上述不同形态的产品均以公司研发的云端智能芯片为核心。

- **云端智能芯片：**云端智能芯片对应的加速卡产品系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 PCIE 接口、OAM 接口）的板卡硬件产品。加速卡产品是云服务器、数据中心中进行人工智能处理的基本单元，主要为云计算和数据中心场景下的人工智能应用程序提供高密度、高效率的硬件计算资源，加速卡的核心算力来源是公司自研的云端智能芯片。
- **训练整机：**训练整机系面向云端训练场景，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。

- **智能计算集群系：**智能计算集群系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。能为人工智能应用部署技术能力相对较弱的客户提供软硬件整体解决方案。

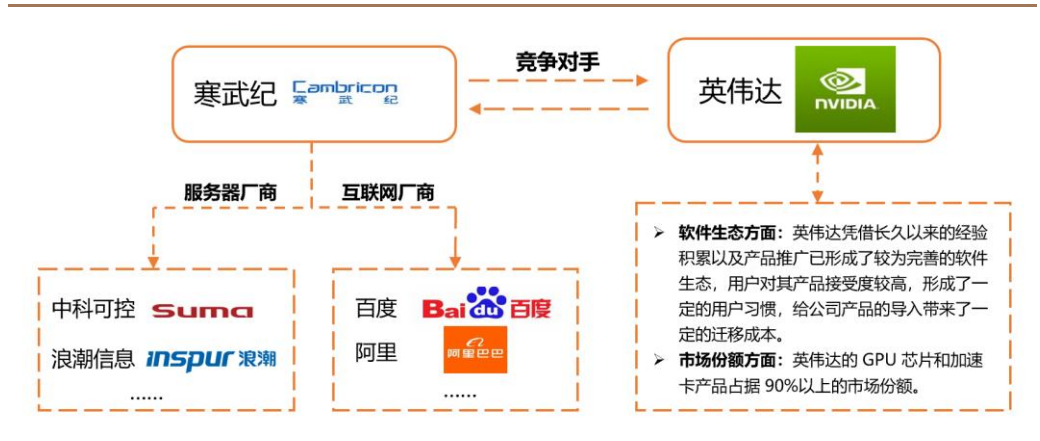
图 21：公司主要产品类型与推出时间

产品形态	产品图示	主要客户	应用场景
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
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
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

资料来源：《2022 年度向特定对象发行 A 股股票募集说明书》，德邦研究所

公司云端产品客户主要为服务器厂商与互联网厂商，竞争对手英伟达实力强大。目前，公司云端产品主要应用于互联网、金融等领域，主要客户为中科可控、浪潮信息等服务器厂商及阿里、百度等互联网公司。公司的云端智能芯片和加速卡的主要竞争对手为英伟达。在软件生态方面，英伟达凭借长久以来的经验积累以及产品推广已形成了较为完善的软件生态，一定程度上提高了公司产品的导入成本。目前，英伟达的 GPU 芯片和加速卡产品占据 90% 以上的市场份额。

图 22：公司客户情况与竞争对手情况



资料来源：《2022 年度向特定对象发行 A 股股票募集说明书》，德邦研究所

公司思元 290 与思元 370 采用 7nm 工艺设计，分别负责推算一体与云端推理两大方向。公司思元 370 为推算一体芯片，思元 290 为云端训练芯片及加速卡产品。训练端的思元 290 配备 HBM2，算力最高可达 512TOPS (INT8)、64TOPS (CINIT32)。而凭借 7nm 制程工艺和寒武纪最新智能芯片架构 MLUarch03，思元 370 峰值算力可达 256TOPS(INT8)，是寒武纪第二代产品思元 270 算力的 2 倍。

图 23：公司云端智能芯片性能对比

产品系列	推出时间	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) 芯片及云端智能加速卡	2018年	-	-	-	-	-	-	-
思元270 (MLU270) 芯片及云端智能加速卡	2019年	MLU270-S4智能加速卡	-	INT16, INT8, INT4, FP32, FP16	128TOPS(INT8) 256TOPS(INT4) 64TOPS(INT16)	DDR4,ECC	16GB	70w
		MLU270-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) 芯片及云端智能加速卡	2020年	MLU290-M5智能加速卡	7nm	-	自适应精度训练算力 512 TOPS (INT8) 256 TOPS (INT16) 64 TOPS (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	MLU370-S4智能加速卡	7nm	FP32、FP16、BF16、INT16、INT8、INT4	192 TOPS (INT8) 96 TOPS (INT16) 72 TFLOPS (FP16) 72 TFLOPS (BF16) 18 TFLOPS (FP32)	LPDDR5	24GB	75W
		MLU370-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
		MLU370-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

资料来源：公司官网，公司公告，德邦研究所

相比友商产品，公司思元 370 竞争优势显著，首次采用 **Chiplet** 工艺。思元 370 实测算力性能能与主流 GPU 比肩。且思元 370 芯片支持 LPDDR5 内存，高带宽且低功耗，可在板卡有限的功耗范围内给 AI 芯片分配更多的能源，输出更高的算力。

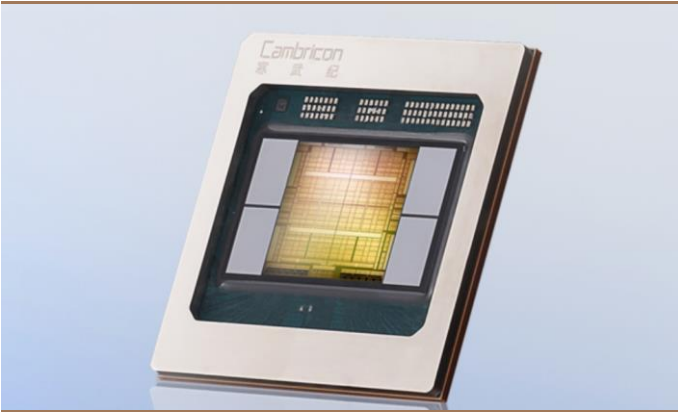
图 24：公司思元 370 产品市场竞争优势

竞争优势	具体功能
芯片算力和实测性能、能效提升	凭借7nm制程工艺和寒武纪最新智能芯片架构MLUarch03，思元370峰值算力可达256TOPS(INT8)，是寒武纪第二代产品思元270算力的2倍。同时，思元370芯片支持LPDDR5内存，高带宽且低功耗，可在板卡有限的功耗范围内给AI芯片分配更多的能源，输出更高的算力。相较于峰值算力的提升，思元370在实测性能和能效方面也具有一定优势。以ResNet-50为例，MLU370-S4加速卡（半高半长）实测性能为同尺寸主流GPU的2倍；MLU370-X4加速卡（全高全长）实测性能与同尺寸主流GPU相当，能效则大幅领先。
产品配置灵活，满足不同应用场景需求	寒武纪首次采用chiplet技术，通过不同芯粒组合规格多样化的产品，为用户提供适用不同场景的高性价比AI芯片。
推动“云边端一体、训推一体、软硬件协同”的统一智能生态建设	为了加快用户端到端业务落地的速度，减少模型训练研发到模型部署之间的繁琐流程，寒武纪的统一基础软件平台CambriconNeuware整合了训练和推理的全部底层软件栈，将MagicMind和深度学习框架Tensorflow、Pytorch深度融合，实现训推一体，有效提升开发部署效率，降低用户成本。

资料来源：公司公开投资者调研纪要，Wind，德邦研究所

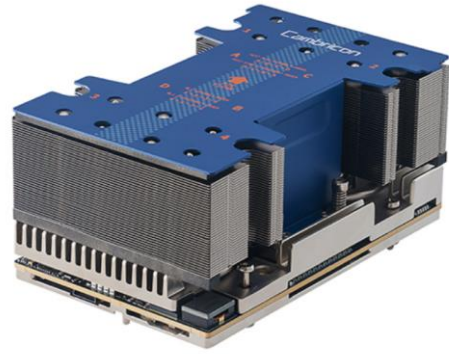
训练端 290 芯片持续拓展，新品 590 在研待发布。思元 290 智能芯片是寒武纪的首颗训练芯片，采用台积电 7nm 先进制程工艺，集成 460 亿个晶体管，支持 MLUV02 扩展架构，全面支持 AI 训练、推理或混合型人工智能计算加速任务。随着搭载思元 290 芯片的云端智能芯片及加速卡、训练整机产品的市场拓展和落地，云端产品线产品形成协同，大幅提升了云端产品的市场销售潜力。同时，根据寒武纪官方公众号，目前在研全新一代云端智能训练芯片思元 590，采用 MLUarch05 全新架构，实测训练性能较在售产品有了显著提升，它提供了更大的内存容量和更高的内存带宽，其 PCIE 接口也较上代实现了升级，有望接力 290 成为业务增长点。

图 25: 寒武纪首颗训练芯片思元 290



资料来源: 公司官网, 德邦研究所

图 26: 寒武纪智能加速卡 MLU290-M5



资料来源: 公司官网, 德邦研究所

3.2. 边缘智能芯片及加速卡: 立足高算力领域, 新品推出激发新老客户放量

边缘计算是对云端计算的有效补充。边缘计算是近年来兴起的一种新型计算范式, 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。

公司发力国内智能物联网领域, 在高算力市场取得较大份额。根据 ABI Research 预计, 全球边缘智能芯片市场规模将从 2019 年的 26 亿美元增长到 2024 年的 76 亿美元, 年化复合增长率达到 23.9%。公司当前已有产品涉足的主要领域为智能物联网, 目前该领域各类边缘智能芯片的国内市场规模约为 6-8 亿美元。在边缘计算市场中的智能物联网领域, 除公司外, 目前主要的边缘芯片供应商包括英伟达、Novatek、华为海思及瑞芯微等厂商。公司当前主要在智能物联网中的较高算力市场占有较大份额。


图 27: 边缘智能物联网市场竞争格局

公司	特点
NOVATEK	基于其成本优势, 在低算力视觉感知方向增长较快
华为海思	依靠其在芯片行业的长期积累。在中小算力市场获得了较大份额, 但后续受产品迭代的影响, 市场份额有所降低
英伟达	依赖其产品的通用性优势和软件生态优势, 较早就占领了智能物联网市场较大份额, 但由于其中小算力产品迭代较慢且性价比优势不明显, 早期取得的市场份额逐渐降低
瑞芯微	基于其传统智能广告屏及Linux/Android软件平台等方面的优势
寒武纪	当前主要在智能物联网中的较高算力市场占有较大份额

资料来源: 《寒武纪向特定对象发行股票申请文件的第二轮审核问询函的回复》, 德邦研究所

公司边缘芯片主要为思元 220, 累计出货百万片。公司于 2019 年 11 月推出了边缘智能芯片思元 220 及相应的 M.2 加速卡。思元 220 的推出, 标志着公司已经具备了从终端、边缘端到云端完整产品线。目前, 思元 220 累计销量突破百万片。

图 28: 思元 220 芯片的具体情况

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>

资料来源: 招股说明书, 德邦研究所

思元 220 销售依赖客户 A, 且已进入生命周期后期, 影响 22 年边缘产品收入。根据《关于中科寒武纪科技股份有限公司向特定对象发行股票申请文件的第二轮审核问询函的回复(修订稿)》, 公司自边缘端产品 220 量产后即开始与客户 A 展开合作, 其为行业头部客户, 因此公司边缘产品线收入主要来自公司 A, 2022 年公司边缘业务收入下滑主要受公司 A 需求变化影响, 此外公司边缘端产品思元 220 已经进入产品生命后期, 新产品尚未推出, 亦影响了 2022 年边缘端收入。随着公司 A 逐渐消化库存, 预计将继续采购思元 220, 保守预测其 2023 年内仍有 20 万片的采购需求, 对应采购金额约为 3500 万元。

图 29: 思元 220 系列 MLU220-M.2 边缘端智能加速卡规格参数

参数	规格
型号	MLU220-M.2
AI性能	8TOPS (INT8)
内存	LPDDR4x64bits
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+Mkey (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

资料来源: 公司官网, 德邦研究所

公司新边缘产品即将推出, 公司 A 与其它客户有望同步放量。根据《关于中科寒武纪科技股份有限公司向特定对象发行股票申请文件的第二轮审核问询函的回复(修订稿)》, 公司目前在研的边缘端产品计划于 2023 年下半年推向市场, 多家客户对相关产品有较强采购意向, 上述客户在技术交流阶段对产品规格和性能表示满意, 部分客户启动了产品立项工作。公司相关产品预计在 2023 年可贡献 840 万收入。预计主要客户的产品导入开发工作在 2023 年基本完成, 2024 年开始规模出货, 当年相关产品的收入将持续提升。

3.3. 终端 IP 授权及软件: 服务生态建设, 构建完备业务矩阵

终端智能 IP 为公司早期产品, 目前逐渐成为业务中间形态。终端智能处理器是终端设备中支撑人工智能处理运算的核心器件, 为了提升性能降低功耗, 同时节省成本, 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。随着公司云边端产品线的丰富, 终端智能处理器 IP 授权逐渐成为公司业务发展的一个中间形态, 更多服务于自有芯片处理器核心或生态建设拓展。已集成于超过 1 亿台智能手机及其他智能终端设备中。

图 30：公司终端智能处理器 IP 产品种类

产品型号及推出时间	技术指标	产品特点
寒武纪1A (2016)	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)	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)	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)	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)	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)	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。

资料来源：公司招股说明书，德邦研究所

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。

在软件层面，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，可让同一人工智能应用程序便捷高效地运行在公司云边端系列化芯片与处理器产品之上。

图 31：Cambricon Neuware 框架结构



资料来源：公司招股说明书，德邦研究所

4. 盈利预测与投资建议

公司为国内 AI 芯片龙头，“云、边、终端”全链路协同覆盖。公司云端业务的主力产品思远 290 和 370 采用了 7nm 先进工艺设计，370 系列首次采用 Chiplet 技术，有望在算力和良率水平提升的同时大幅降低成本；公司边缘端业务的产品为思远 220 系列，累计销量已达百万规模，随着主力客户去库存进程的完成，有望带动公司边缘端营收的回暖；终端智能处理器 IP 授权业务逐渐转为公司的存量业务，有望服务于公司整体云边端产线的生态建设。

我们预计公司云端产品线业务 2023-2025 年营收有望增长 137%/56%/48%，实现营收 5.20/8.10/12.00 亿元，毛利率分别为 68%/71%/73%；边缘产品线业务 2023-2025 年营收有望增长 84%/57%/27%，实现营收 0.70/1.10/1.40 亿元，毛利率分别为 37%/39%/40%；智能计算集群系统业务 2023-2025 年营收有望增

长 13%/15%/13%，实现营收 5.20/6.00/6.80 亿元，毛利率分别为 70%/69%/67%。IP 授权及软件业务 2023-2025 年营收有望增长 0%/100%/200%，实现营收 0.01/0.02/0.06 亿元，毛利率分别为 100%/100%/100%。

图 32：寒武纪营收及毛利率预测（单位：亿元）

	2018	2019	2020	2021	2022	2023E	2024E	2025E
营收	1.17	4.44	4.59	7.21	7.29	11.26	15.38	20.44
云端产品线		0.79	0.86	0.80	2.19	5.20	8.10	12.00
边缘产品线			0.21	1.75	0.38	0.70	1.10	1.40
智能计算集群系统		2.96	3.26	4.56	4.59	5.20	6.00	6.80
IP授权及软件	1.17	0.69	0.12	0.07	0.01	0.01	0.02	0.06
其它业务			0.14	0.03	0.12	0.15	0.16	0.18
营收同比		279%	3%	57%	1%	54%	37%	33%
云端产品线			9%	-7%	174%	137%	56%	48%
边缘产品线				733%	-78%	84%	57%	27%
智能计算集群系统			10%	40%	1%	13%	15%	13%
IP授权及软件		-41%	-83%	-42%	-86%	0%	100%	200%
其它业务				-79%	300%	25%	7%	13%
毛利率	100%	68%	65%	62%	66%	67%	68%	69%
云端产品线		78%	76%	59%	63%	68%	71%	73%
边缘产品线			49%	41%	31%	37%	39%	40%
智能计算集群系统		58%	62%	71%	70%	70%	69%	67%
IP授权及软件	100%	100%	100%	100%	100%	100%	100%	100%
其它业务	93%	0%	79%	67%	42%	50%	54%	63%

资料来源：Wind，德邦研究所预计

结合寒武纪的主营业务，我们选取英伟达作为海外可比公司，选取海光信息、龙芯中科和景嘉微作为国内可比公司。英伟达为全球 AI 计算领域的绝对龙头企业；海光信息主营业务为应用于服务器、工作站的高端处理器，龙芯中科专注于面向信息和工控系统的处理器及配套芯片研发，景嘉微为国内 GPU 厂商，成功研发多款图形处理芯片，均处于 AI 算力芯片赛道，与公司较为可比。我们预计公司将在 2023 年至 2025 年实现收入 11.26/15.38/20.44 亿元，归母净利润-8.7/-7.2/-4.9 亿元，对应当前 PS 估值 57/42/31 倍。首次覆盖，给予“买入”评级。

图 33：可比公司估值（2023/11/17）

公司简称	市值	营收				净利润				PE				PS			
		2022A	2023E	2024E	2025E	2022A	2023E	2024E	2025E	2022A	2023E	2024E	2025E	2022A	2023E	2024E	2025E
英伟达	12,222	270	-	-	-	44	-	-	-	278	-	-	-	45	-	-	-
海光信息	1,537	50.9	64.1	87.7	115.7	8.0	12.0	16.6	22.9	191	129	92	67	30	24	18	13
龙芯中科	380	7.3	9.4	15.2	22.2	0.5	0.2	2.0	3.5	735	2,318	192	110	52	41	25	17
景嘉微	390	11.4	13.2	19.4	27.1	2.9	2.7	4.1	5.5	135	143	95	71	34	29	20	14
平均										335	863	127	83	40	31	21	15
寒武纪	640	7.3	11.3	15.4	20.4	-12.6	-8.7	-7.2	-4.9	-	-	-	-	85	57	42	31

资料来源：Wind，德邦研究所预测

注 1：海光信息、龙芯中科、景嘉微盈利预测采用 Wind 一致预期，寒武纪盈利预测采用德邦预测数据。

注 2：英伟达市值、营收、净利润的单位为亿美元；海光信息、龙芯中科、景嘉微、寒武纪的市值、营收、净利润的单位为亿人民币。

5. 风险提示

AI 需求不及预期风险。目前 AI 产业仍处于发展初期，未来技术变革与需求变化仍不确定，若 AI 产业的需求不及预期，可能影响公司产品的价格与需求量。

客户集中度较高风险。2020 年、2021 年和 2022 年，公司前五大客户的销售金额合计占营业收入比例分别为 82.11%、88.60%和 84.94%，客户集中度较

高。若公司主要客户对公司产品的采购量大幅降低或者公司未能继续维持与主要客户的合作关系，将给公司业绩带来显著不利影响。

供应链稳定相关风险。2022 年 12 月 15 日，美国商务部工业和安全局（BIS）以国家安全和外交利益为由，将公司及部分子公司列入“实体清单”。目前，该事项对公司研发造成的影响程度尚在有限范围内，公司的核心技术来自于寒武纪的自主研发，拥有自主知识产权，不会对公司产品研发及核心竞争力产生重大不利影响。长期来看，切换新供应商将产生一定成本，将可能对公司经营业绩产生不利影响。

财务报表分析和预测

主要财务指标	2022	2023E	2024E	2025E
每股指标(元)				
每股收益	-3.14	-2.08	-1.73	-1.17
每股净资产	11.65	9.73	8.07	6.95
每股经营现金流	-3.19	-2.82	-2.40	-1.88
每股股利	0.00	0.00	0.00	0.00
价值评估(倍)				
P/E	—	—	—	—
P/B	4.68	15.79	19.04	22.13
P/S	84.50	56.87	41.63	31.33
EV/EBITDA	-17.80	-94.67	-140.98	-349.57
股息率%	0.0%	0.0%	0.0%	0.0%
盈利能力指标(%)				
毛利率	65.8%	66.8%	67.8%	68.7%
净利润率	-172.4%	-77.1%	-46.9%	-23.9%
净资产收益率	-25.9%	-21.4%	-21.5%	-16.9%
资产回报率	-21.8%	-12.9%	-9.1%	-5.6%
投资回报率	-29.0%	-14.5%	-9.3%	-4.9%
盈利增长(%)				
营业收入增长率	1.1%	54.5%	36.6%	32.9%
EBIT 增长率	-51.5%	38.5%	23.7%	-41.3%
净利润增长率	-52.3%	30.9%	16.9%	-32.2%
偿债能力指标				
资产负债率	14.4%	38.9%	57.3%	67.1%
流动比率	7.8	2.0	1.5	1.2
速动比率	7.2	1.8	1.4	1.2
现金比率	4.6	1.1	0.8	0.7
经营效率指标				
应收帐款周转天数	383.3	310.0	280.0	250.0
存货周转天数	419.8	320.0	250.0	210.0
总资产周转率	0.1	0.2	0.2	0.2
固定资产周转率	3.1	4.3	5.4	6.8

现金流量表(百万元)	2022	2023E	2024E	2025E
净利润	-1,257	-869	-722	-489
少数股东损益	-68	-36	-30	-20
非现金支出	680	280	257	234
非经营收益	-87	-100	-71	-70
营运资金变动	-598	-451	-435	-437
经营活动现金流	-1,330	-1,175	-1,000	-783
资产	-284	-251	-211	-192
投资	1,591	-154	-178	-168
其他	-530	141	198	264
投资活动现金流	777	-264	-192	-96
债权募资	0	2,017	1,898	1,324
股权募资	221	16	0	0
其他	-123	-293	-127	-195
融资活动现金流	99	1,739	1,771	1,129
现金净流量	-454	302	579	250

备注：表中计算估值指标的收盘价日期为 11 月 17 日
 资料来源：公司年报 (2021-2022)，德邦研究所

利润表(百万元)	2022	2023E	2024E	2025E
营业总收入	729	1,126	1,538	2,044
营业成本	250	374	495	639
毛利率%	65.8%	66.8%	67.8%	68.7%
营业税金及附加	3	5	6	8
营业税金率%	0.4%	0.4%	0.4%	0.4%
营业费用	83	135	185	245
营业费用率%	11.3%	12.0%	12.0%	12.0%
管理费用	297	394	431	491
管理费用率%	40.7%	35.0%	28.0%	24.0%
研发费用	1,523	1,644	1,846	2,044
研发费用率%	208.9%	146.0%	120.0%	100.0%
EBIT	-1,463	-899	-686	-403
财务费用	-52	-34	33	85
财务费用率%	-7.1%	-3.0%	2.2%	4.2%
资产减值损失	-196	-50	-30	-20
投资收益	91	146	200	266
营业利润	-1,324	-902	-749	-508
营业外收支	2	0	0	0
利润总额	-1,323	-902	-749	-508
EBITDA	-1,096	-669	-459	-188
所得税	2	3	2	2
有效所得税率%	-0.2%	-0.3%	-0.3%	-0.3%
少数股东损益	-68	-36	-30	-20
归属母公司所有者净利润	-1,257	-869	-722	-489

资产负债表(百万元)	2022	2023E	2024E	2025E
货币资金	2,467	2,769	3,348	3,598
应收账款及应收票据	766	956	1,180	1,400
存货	287	328	339	368
其它流动资产	735	963	1,189	1,421
流动资产合计	4,255	5,016	6,056	6,787
长期股权投资	222	296	395	483
固定资产	233	264	287	302
在建工程	26	27	28	29
无形资产	291	331	311	271
非流动资产合计	1,515	1,696	1,860	2,007
资产总计	5,770	6,712	7,916	8,794
短期借款	0	2,027	3,524	4,846
应付票据及应付账款	230	225	228	245
预收账款	0	0	0	0
其它流动负债	312	312	335	361
流动负债合计	542	2,565	4,087	5,453
长期借款	0	0	0	0
其它长期负债	288	44	448	450
非流动负债合计	288	44	448	450
负债总计	831	2,609	4,534	5,902
实收资本	401	417	417	417
普通股股东权益	4,855	4,054	3,363	2,893
少数股东权益	85	49	19	-2
负债和所有者权益合计	5,770	6,712	7,916	8,794

信息披露

分析师与研究助理简介

陈海进，德邦证券所长助理兼 TMT 组组长、电子首席分析师，8 年以上电子行业研究经验，曾任职于民生证券、方正证券、中欧基金等，南开大学国际经济研究所硕士。电子行业全领域覆盖。

徐巡，电子行业研究助理，上海交通大学经济学硕士，2 年以上电子行业研究经验，曾任职于凯盛研究院，覆盖 IC 设计、半导体设备与制造等领域。

谢文嘉，电子行业研究助理，香港大学硕士，覆盖模拟 IC 设计、数字 IC 设计等领域。

分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的职业态度，独立、客观地出具本报告。本报告所采用的数据和信息均来自市场公开信息，本人不保证该等信息的准确性或完整性。分析逻辑基于作者的职业理解，清晰准确地反映了作者的研究观点，结论不受任何第三方的授意或影响，特此声明。

投资评级说明

1. 投资评级的比较和评级标准： 以报告发布后的 6 个月内的市场表现为比较标准，报告发布日后 6 个月内的公司股价（或行业指数）的涨跌幅相对同期市场基准指数的涨跌幅；	类别	评级	说明
2. 市场基准指数的比较标准： A 股市场以上证综指或深证成指为基准；香港市场以恒生指数为基准；美国市场以标普 500 或纳斯达克综合指数为基准。	股票投资评级	买入	相对强于市场表现 20%以上；
		增持	相对强于市场表现 5%~20%；
		中性	相对市场表现在-5%~+5%之间波动；
		减持	相对弱于市场表现 5%以下。
	行业投资评级	优于大市	预期行业整体回报高于基准指数整体水平 10%以上；
		中性	预期行业整体回报介于基准指数整体水平-10%与 10%之间；
		弱于大市	预期行业整体回报低于基准指数整体水平 10%以下。

法律声明

本报告仅供德邦证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

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。

市场有风险，投资需谨慎。本报告所载的信息、材料及结论只提供特定客户作参考，不构成投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。客户应考虑本报告中的任何意见或建议是否符合其特定状况。在法律许可的情况下，德邦证券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并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经德邦证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。如欲引用或转载本文内容，务必联络德邦证券研究所并获得许可，并需注明出处为德邦证券研究所，且不得对本文进行有悖原意的引用和删改。

根据中国证监会核发的经营证券业务许可，德邦证券股份有限公司的经营经营范围包括证券投资咨询业务。