

蓝耘科技 (871169.NQ) 智算云平台发布深度 AI 赋能，加速重构算力增长新曲线

2025 年 01 月 05 日

——新三板公司研究报告

| | |
|--------------|------------|
| 日期 | 2025/1/3 |
| 当前股价(元) | 12.80 |
| 一年最高最低(元) | 16.72/8.20 |
| 总市值(亿元) | 11.83 |
| 流通市值(亿元) | 9.78 |
| 总股本(亿股) | 0.92 |
| 流通股本(亿股) | 0.76 |
| 近 3 个月换手率(%) | 23.52 |

诸海滨 (分析师)

zhuhaibin@kysec.cn

证书编号: S0790522080007

北交所研究团队

● GPU 算力解决方案为基础，算力云服务为核心

蓝耘科技是一家专注于为企业级规模的 GPU 加速工作负载提供算力云服务的公司。公司的主营业务为 IT 系统集成业务、技术服务业务、GPU 算力云服务业务，2024H1 算力云解决方案实现收入 4.11 万元，同比增长 380.63%，GPU 算力云服务业务实现收入 1.25 亿元，同比增长 139.74%。2023 年实现营收 4.09 亿元、归母净利润 5128.17 万元；2020-2023 年 3 年营收 CAGR 为 5.54%，归母净利润 CAGR 达到 15.75%，2023 年毛利率/净利率分别为 25.79%/12.38%。2024H1 实现营收 5.36 亿元，归母净利润 4827.93 万元。

● AI 算力需求扩张，云计算市场迎来新一轮增长

我国智能算力规模高速增长，预计 2026 年智能算力规模突破 ZFLOPS 量级，增长速度远超通用基础算力。2022-2027 年期间，预计中国智能算力规模年复合增长率达 33.9%。随着近年来 AI 技术的飞跃发展，以国产数据库、大数据、专有云、操作系统为代表的产品在垂直产业深入应用，特别是在政务、金融、交通、教育等关键基础设施领域，为中国云计算厂商带来了全新市场机遇。2023 年中国云计算市场规模达 6165 亿元，同比增长 35.5%，大幅高于全球增速。预计到 2024 年，云计算市场规模将达 8378 亿元，同比增长 35.9%。

● 发布新品智算云平台深度赋能 AI 研发，与战略客户签署战略合作协议加速合作

2024 年 11 月，公司正式发布了蓝耘智算云平台——元生代。平台基于 Kubernetes 原生云设计，依托行业领先的灵活基础设施和大规模 GPU 算力资源，实现从数据准备、代码开发、模型训练到推理部署等全场景覆盖，高效赋能用户 AI 研发全流程。同年 11 月，蓝耘科技与杭钢数科、浙文互联、京源云计算、浙大科创、拱墅国投分别签署战略合作协议。截至 2023 年 12 月 31 日，公司正在申请专利 14 项，目前拥有计算机软件著作权 125 项。技术创新方面，公司研发费用呈增长态势。

● 蓝耘科技同行可比公司 PE2023 均值为 73.0X

蓝耘科技同行可比公司 PE2023 均值为 73.0X。蓝耘科技专注于为企业级规模的 GPU 加速工作负载提供算力云服务，服务于高校科研、生命科学、人工智能、自动驾驶、工业设计、智慧城市等领域，主要业务包括 IT 系统集成业务和 GPU 算力云服务业务两大模块，GPU 算力云服务业务自 2021 年以来高速增长，目前与产业方广泛合作并通过融资租赁等方式加速 AI 时代布局，有望拓展算力版图，不断提升产业地位，在 LLM 等需求驱动下前景广阔。可比公司 PE 2023 均值 73.0X，蓝耘科技 PE 2023 为 21.0X，估值相对较低，同时毛利率高于可比公司平均，具备一定可比优势。

● 风险提示：算力需求波动、产业政策变化、技术迭代风险

相关研究报告

《专注 GPU 算力服务，立足 AI 模型算法驱动盈利增长——新三板公司研究报告》-2024.5.21

目录

| | |
|---|----|
| 1、 看点：智算云平台深度赋能，签署战略合作协议加速合作..... | 4 |
| 2、 行业：AI 算力需求扩张，云计算市场迎来新一轮增长..... | 9 |
| 2.1、 产业链：算力网络平台位于产业中游，包括数据中心、云运算等..... | 9 |
| 2.2、 算力市场：预计 2026 年智能算力规模突破 ZFLOPS 量级..... | 11 |
| 2.3、 智算服务市场：数字化转型加速，市场规模保持增长趋势..... | 12 |
| 2.4、 云计算市场：预计 2024 年市场规模将达 8378 亿元..... | 14 |
| 3、 业务：GPU 算力解决方案为基础，算力云服务为核心..... | 17 |
| 3.1、 产品：GPU 算力云业务占比持续提升，2024H1 毛利率达 39.16%..... | 17 |
| 3.2、 商业模式：自建智算云平台，服务高性能计算需求行业龙头企业..... | 18 |
| 3.3、 财务：2024H1 实现营收 5.36 亿元，归母净利润 4827.93 万元..... | 20 |
| 4、 估值对比：同行可比公司 PE2023 均值为 73.0X..... | 22 |
| 5、 风险提示..... | 22 |

图表目录

| | |
|---|----|
| 图 1： 蓝耘算力云平台是一个现代化的、基于 Kubernetes 的云平台..... | 4 |
| 图 2： 2024 年 11 月，公司正式发布了蓝耘智算云平台——元生代..... | 5 |
| 图 3： 蓝耘北京智算中心是蓝耘在算力基础设施建设方面的又一重要里程碑..... | 5 |
| 图 4： 蓝耘科技提供具备优势的英伟达 GPU 算力资源解决方案..... | 6 |
| 图 5： 蓝耘科技利用裸金属 K8S 部署实现推理服务的性能优势..... | 7 |
| 图 6： 2023 年研发费用率为 3.91%..... | 7 |
| 图 7： AI 算力产业链一览..... | 9 |
| 图 8： 上游芯片/器件领域代表性企业有中兴微电子、ARM、英特尔、英伟达等..... | 10 |
| 图 9： 整体来看，北京、上海、广东、江苏、湖北、四川等地产业链完善..... | 10 |
| 图 10： 智能计算是指面向人工智能应用的高性能计算..... | 11 |
| 图 11： 2022 年我国算力结构为通用算力：智能算力：超级算力=40%:59%:1%..... | 11 |
| 图 12： 预计到 2027 年中国智能算力规模达 1117.4EFLOPS..... | 12 |
| 图 13： 智算服务，是指以 GPU、FPGA、ASIC 等 AI 专用算力为主的基础设施服务..... | 13 |
| 图 14： 2023 年智算服务市场以 Non-GenAIaaS 为主，目前占整体的 40.2%..... | 13 |
| 图 15： 预计 2024 年中国智算服务市场规模将超过 200 亿元..... | 14 |
| 图 16： 从产业结构来看，我国目前云计算产业链结构完备..... | 15 |
| 图 17： 云计算产业最终服务于互联网、政府、金融等广大传统行业与个人用户..... | 15 |
| 图 18： 2023 年中国公有云市场规模 4562 亿元，同比增长 40.1%..... | 16 |
| 图 19： 预计到 2024 年，我国云计算市场规模将达 8378 亿元，同比增长 35.9%..... | 16 |
| 图 20： 未来十年，云计算将进入全新发展阶段..... | 17 |
| 图 21： GPU 算力云服务业务营收规模保持稳健增长，经营质效稳步提升..... | 18 |
| 图 22： 2024H1 算力云服务收入占比为 23.33%..... | 18 |
| 图 23： 2024H1 GPU 算力云服务业务营收达到 1.25 亿元..... | 18 |
| 图 24： 2024H1 实现营收 5.36 亿元..... | 20 |
| 图 25： 2024H1 实现归母净利润 4827.93 万元..... | 20 |
| 图 26： 2024H1 毛利率为 18.44%..... | 21 |
| 图 27： 2021-2022 算力云服务毛利率高于系统集成业务..... | 21 |

| | |
|---|----|
| 图 28: 2024H1 GPU 算力云服务毛利为 39.16% | 21 |
| 图 29: 2024H1 公司期间费用率为 8.35% | 21 |
| 图 30: 2024H1 净利率为 8.87% | 21 |
| 表 1: 截至 2023 年 12 月 31 日, 公司正在申请专利 14 项 | 7 |
| 表 2: 2023 年公司研发支出前五名的研发项目共研发支出 1205.28 万元 | 8 |
| 表 3: 2023 年蓝耘科技向前五大客户销售金额为 1.78 亿元 | 19 |
| 表 4: 同行业可比公司 PE2023 均值为 73.0X | 22 |

1、看点：智算云平台深度赋能，签署战略合作协议加速合作

蓝耘科技依托多年实践积累的技术经验，以智算中心“算力基座”为支点，持续扩充底层算力资源池，面对 AI 新时代下日益高涨的高频算力需求，通过聚焦于以算力解决方案为基础、GPU 算力云服务为核心的业务体系，为 AI 训练、推理、视觉特效和渲染及教科研等计算密集型场景用户，提供可按需付费的高效便捷的算力云服务，助力用户研发和科研提速增效。2024 年 11 月，蓝耘科技与杭钢数科、浙文互联、京源云计算、浙大科创、拱墅国投分别签署战略合作协议，为全面深化合作、共谋发展拉开序幕。

公司 GPU 算力云服务主要针对大模型、AIGC、高校科研机构、生命科学、自动驾驶、影视动漫、工业仿真、智慧城市等行业客户对 GPU 算力的需求，通过算力云平台为客户提供远程 GPU 算力，客户通过包年包月或按需付费的灵活方式预付费给蓝耘科技。蓝耘算力云平台基于行业领先的灵活的基础设施及丰富的 GPU 算力资源，从裸金属、容器到 Kubernetes，可以向用户提供集 GPU 池化资源、高性能网络、高性能存储等基础设施和资源调度、开发环境、运行监控、运维服务于一体的全栈服务，这些资源能够根据用户的需求动态分配，满足客户在 AI 模型构建、训练和推理业务全流程中对 GPU 算力灵活性及多样化的需求，并为教科研用户科研创新加速。

蓝耘算力云平台是一个现代化的、基于 Kubernetes 的云平台，专为大规模 GPU 加速工作负载而构建。蓝耘旨在为工程师和创新者提供无与伦比的计算解决方案，其速度可比传统云服务提供商快 35 倍，成本降低 30%。

图1：蓝耘算力云平台是一个现代化的、基于 Kubernetes 的云平台



资料来源：公司官网

2024 年 11 月，公司正式发布了蓝耘智算云平台——元生代。元生代具备智算算力调度，应用市场，AI 协作开发三大模块。裸金属调度，能够给用户更多的自由发挥空间，容器调度则依托方便快捷的调度能力，帮助纳管伙伴闲置的算力资源。

元生代 AI 协作开发模块充分考虑了 AI 开发的团队协作场景。整个平台分为前台、中台和后台；前台为开发工程师集成了常用的开发套件、存储调用、镜像仓库及高灵活度的资源调度，为了进一步提高训练效率，还做了包括计算网络轨道调优、

NUMA 亲和性调度、分布式缓存等等一些列后台优化，甚至可实现将自定义任务运行于同一组 Leaf 交换机上。中台主要给用户的运维团队提供集群基础设施级别的监控指标，让高端玩家更快更好的追踪训练任务各个阶段的资源使用水位线，进一步提高训练效率。同时可以更直观的看到每个项目的实际资源使用率，更迅速的做出资源的合理分配，避免团队资源的浪费。后台主要为非技术人员提供运营和财务等相关的功能。

图2：2024年11月，公司正式发布了蓝耘智算云平台——元生代



资料来源：公司发布会 PPT

蓝耘北京智算中心是蓝耘在算力基础设施建设方面的又一重要里程碑，该项目将规划部署多台单机柜 48kW 液冷机柜，并设立主机房、辅助区域、办公区及配套设施等多个功能区。此外，还将同步建设室外给排水、电气、暖通空调以及智能化系统等配套工程。计划到 2025 年，该智算中心将正式建成并投入运营。

图3：蓝耘北京智算中心是蓝耘在算力基础设施建设方面的又一重要里程碑



资料来源：公司公众号

蓝耘的 GPU 智算云服务具体包括三方面：1) GPU 虚拟云主机，面向 AI 训练、图形处理、科学计算等领域，为客户提供专享高性能 GPU 虚拟云主机，可大幅提升图形处理和计算能力；2) HPC 平台，为生命科学、CAE 仿真、物理化学、气象环境、石油勘探等领域的科研界提供 CPU、GPU 算力服务；3) 裸金属 GPU 服务器，提供高性能、资源独享、安全隔离的专属弹性裸金属服务器，同时具备云资源的敏捷性、灵活性和高可用性，蓝耘提供完备的设备管理权限及运维服务。

GPU 算力资源整合及优质云平台服务是蓝耘科技在 AI 时代的核心壁垒。蓝耘 GPU 算力云服务基于 Kubernetes 原生云设计，搭载高可配置高可用的英伟达高端 GPU 系列，支持按需灵活计费、轻松调整存储容量及扩展容量，协同网络架构水平扩展、系统调优及应用排错，可实现现有工作负载轻松迁移、快速部署和弹性拓展，平滑应对突发流量高峰，满足 AI 训练开发、机器学习、大规模数据处理、分析等计算密集型任务场景中对 GPU 算力灵活性及多样化的需求，也为集中涌现的 AI 推理场景应用构筑强大的计算力支撑底座。

图4：蓝耘科技提供具备优势的英伟达 GPU 算力资源解决方案



资料来源：公司官网

具体技术方案上，蓝耘科技积累了多年经验，通过硬件层、平台架构层、软件及数据层三大方面实现 AI 算力服务的优化升级。

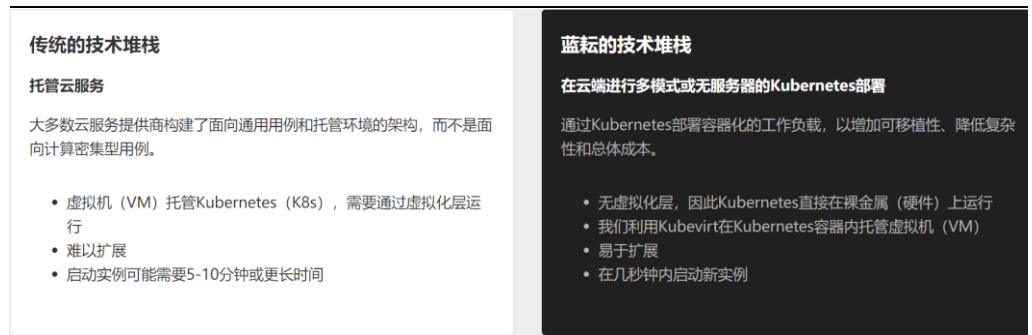
1) **硬件层**：基础设施硬件架构是 AI 算力优化要跨越的第一道“天堑”，蓝耘通过服务器配置调优、国际领先的 AI 加速卡、高速网卡、IB 卡搭载和管理网络、计算网络、存储网络的“三网分离”模式，以及针对 AI 大模型训练场景计算网络量身打造的专用 1:1 收敛的 Infiniband 网络，保障集群内任意 GPU 节点之间都能实现无阻塞的 RDMA 快速通信，为用户带来高带宽、低延迟的优异性能体验；在存储方面，蓝耘基于多年 HPC 系统集成运维、优化经验及 NVIDIA 专业团队技术支持和大模型客户的实际使用诉求，在 GPU 算力集群中优化部署高性能并行文件系统，可为 AI 大模型训练提供高达 100GB/s 的吞吐能力，充分满足 AI 训练、推理、微调及高性能计算场景中对计算、网络、存储等环节灵活性和多样化的需求。

2) **架构层**：平台架构层也是突破 AI 算力升级壁垒的重要一环。蓝耘自研的裸

金属调度平台，针对 AI 大模型训练场景优化了架构设计和系统镜像，配合全生命周期的管理，让大规模裸金属集群部署变得更加简单易用。蓝耘基于 Kubernetes 平台二次开发的容器调度平台，针对大模型训练过程中驱动、GPU 硬件、物理服务器硬件等故障导致的训练停止痛点，拥有无人工干预即可自动拉起失败任务的强大“自愈”功能，大大降低了大模型训练用户的运维难度。

3) 软件层：就软件栈和预训练数据层面而言，蓝耘 AI 大模型训练平台面向深度学习等开发场景，可实现计算资源统一分配调度、训练数据集中管理并加速，通过预置集成开箱即用的国内外主流大模型训练框架、SDK、加速器、预训练数据及容器镜像等，轻松实现集群环境快速部署，高效支撑 AI 创新研发。

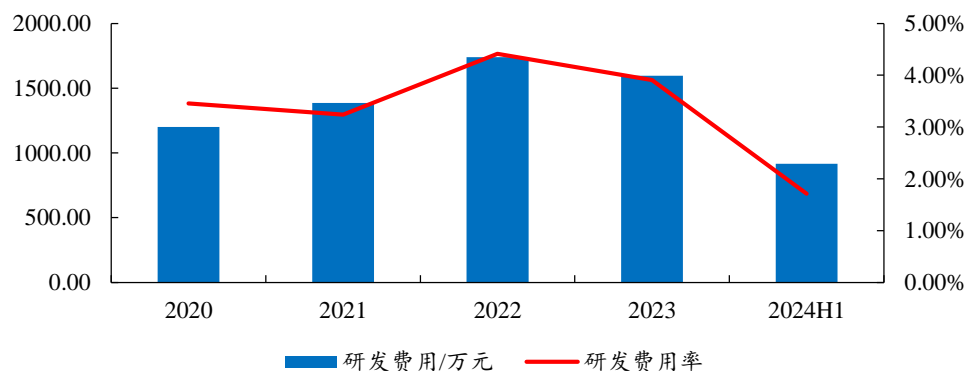
图5：蓝耘科技利用裸金属 K8S 部署实现推理服务的性能优势



资料来源：公司官网

技术创新方面，公司研发费用呈增长态势，自 2018 年的 288 万元上升至 2022 年的 1700 万元以上，2022-2024H1，研发费用分别为 1738.51 万元、1595.56 万元、916.20 万元，研发费用维持稳定。公司获得北京市“专精特新”中小企业、北京市专精特新“小巨人”企业、国家高新技术企业等称号。

图6：2023 年研发费用率为 3.91%



数据来源：Wind、开源证券研究所

公司的研发模式是以自主研发和外包研发的模式进行。截至 2023 年 12 月 31 日，公司正在申请专利 14 项，目前拥有计算机软件著作权 125 项。

表1：截至 2023 年 12 月 31 日，公司正在申请专利 14 项

| 专利类型 | 申请号/专利号 | 发明名称 | 申请日 |
|------|---------------|------------------------|------------|
| 发明专利 | 2023114259538 | 一种动态调整 CPU 和 GPU 缓存的方法 | 2023/10/31 |

| 专利类型 | 申请号/专利号 | 发明名称 | 申请日 |
|------|---------------|----------------------------|------------|
| 发明专利 | 2023111637099 | 一种云计算系统中的资源分配方法 | 2023/9/11 |
| 发明专利 | 2023111637084 | 一种高安全性的云计算方法 | 2023/9/11 |
| 发明专利 | 2023110276591 | 一种集群 GPU 资源调度方法 | 2023/8/16 |
| 发明专利 | 2023110276587 | 基于 CPU 与 GPU 协作的计算方法 | 2023/8/16 |
| 发明专利 | 2022116734457 | 企业终端智能安全管理技术平台 | 2022/12/26 |
| 发明专利 | 2022116744660 | gpu 虚拟化实施方法 | 2022/12/26 |
| 发明专利 | 2022116740693 | 多维并行调度系统 | 2022/12/26 |
| 发明专利 | 2022103107787 | 一种基于云计算移动网络资源管理的 HPC 云计算系统 | 2022/3/28 |
| 发明专利 | 2021103233234 | 一种云计算系统中的资源分配方法 | 2021/3/26 |
| 发明专利 | 2021103233249 | 一种高安全性的云计算方法 | 2021/3/26 |
| 发明专利 | 2021102330415 | 一种基于 CPU 与 GPU 协作的计算方法 | 2021/2/23 |
| 发明专利 | 202110233042X | 一种云计算环境中 GPU 资源调度系统 | 2021/2/23 |
| 发明专利 | 2021102330434 | 一种集群 GPU 资源调度系统 | 2021/2/23 |

数据来源：公司年报、开源证券研究所

2023 年，蓝耘科技共有研发项目 10 个，主要围绕 AI 模型算法开发、AI 模型预训练、AI 推理等场景的平台和工具的研发。并且开始在平台之上开始构建 AIGC 等相关应用。为用户从开发、训练到应用部署提供全流程支撑。研发的平台和工具降低了 AI 技术人员的使用门槛，也降低了应用落地的门槛，提高了应用部署的效率。提升了公司在行业内的竞争力。

表2：2023 年公司研发支出前五名的研发项目共研发支出 1205.28 万元

| 序号 | 研发项目名称 | 总研发支出金额（万元） |
|----|-------------------|-------------|
| 1 | GPU 算力云平台项目 | 357.89 |
| 2 | 冷冻电镜数据处理服务平台 | 281.40 |
| 3 | 面向在线服务的 AIGC 系统开发 | 210.64 |
| 4 | 人工智能博弈算法开发 | 201.06 |
| 5 | 云原生大模型训练平台 | 154.28 |
| | 合计 | 1205.28 |

数据来源：公司年报、开源证券研究所

2、行业：AI 算力需求扩张，云计算市场迎来新一轮增长

2.1、产业链：算力网络平台位于产业中游，包括数据中心、云运算等

AI 算力产业链上游为算力基础硬件设施，主要包括元器件、ICT 基础设施、其他硬件设备等。中游为算力网络与平台，上游硬件设备及基础设施共同组成数据中心、算力网络等，提供 IDC 服务、云计算服务、以及各类算力网络服务等。产业链下游则为应用场景与用户。

图7：AI 算力产业链一览



资料来源：前瞻产业研究院

AI 算力产业链上游芯片/器件领域代表性企业有中兴微电子、ARM、英特尔、英伟达等，服务器领域代表企业有浪潮电子、航天信息、曙光信息、烽火通信等，交换机领域主要企业有华为、锐捷网络、中兴通讯等，光模块领域主要企业有天孚通信、光迅科技、中际旭创、华工正源等。中游算力网络与平台以电信服务商及互联网企业布局较多。下游算力行业用户主要有公众用户和政企用户两大类，广泛应用于互联网、金融、公共事业、电信等领域。

图8：上游芯片/器件领域代表性企业有中兴微电子、ARM、英特尔、英伟达等



资料来源：前瞻产业研究院

AI算力产业链整体来看，北京、上海、广东、江苏、湖北、四川等地产业链完善。北京上中游企业密集，辐射整个华北地区，中游代表性企业有中国移动、百度、360等；华东地区以江浙沪为中心，上游企业云集，有剑桥科技、威固信息，中游有数据港、阿里、网宿科技等。华中地区以武汉为中心，中游企业有中电云计算、天楚云、三峡云、光谷云等；华南地区以深圳为中心，主要有华为、腾讯、金蝶等企业。

图9：整体来看，北京、上海、广东、江苏、湖北、四川等地产业链完善

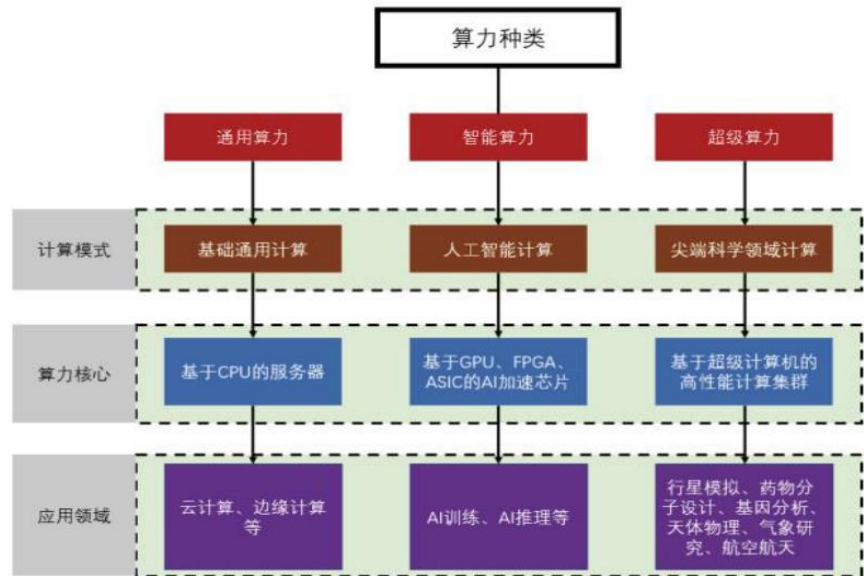


资料来源：前瞻产业研究院

2.2、算力市场：预计 2026 年智能算力规模突破 ZFLOPS 量级

智能计算是指面向人工智能应用(包括训练、推理、衍生应用场景等)的高性能计算。算力可分为:通用算力、智能算力和超级算力,分别对应三种计算模式:基础计算、智能计算和超级计算。不同应用场景下所需的计算精度不同,通常会采用不同种类的算力。目前,随着人工智能的快速发展以及国家层面的政策助推,智能算力规模和占比越来越大,算力需求逐步从通用计算转向智能计算。

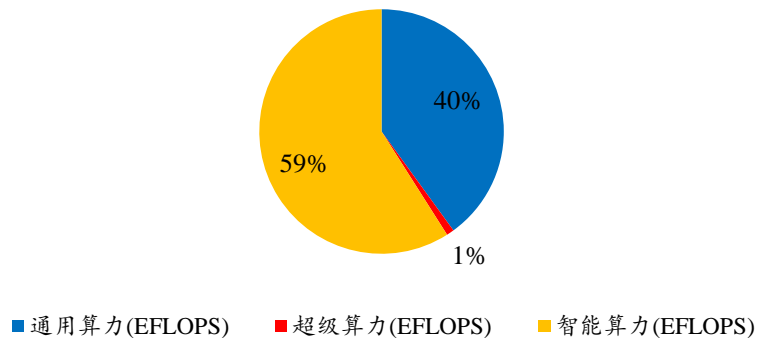
图10: 智能计算是指面向人工智能应用的高性能计算



资料来源:《2024 智能算力产业发展白皮书》

当前,我国智能算力占比已经超过通用算力,成为整体算力增长的主要驱动力。我国智能算力占总体算力的比重增长迅速,据中国信通院《中国算力发展指数白皮书(2023)》数据,2017-2022 年我国算力内部结构中智能算力占比持续增加,2022 年我国算力结构为通用算力:智能算力:超级算力=40%:59%:1%,智能算力占比已较大幅度超过通用算力占比,逐步成为总体算力增长的主要驱动力。据中国信通院预计,未来五年,全球算力规模增速超过 50%。

图11: 2022 年我国算力结构为通用算力:智能算力:超级算力=40%:59%:1%



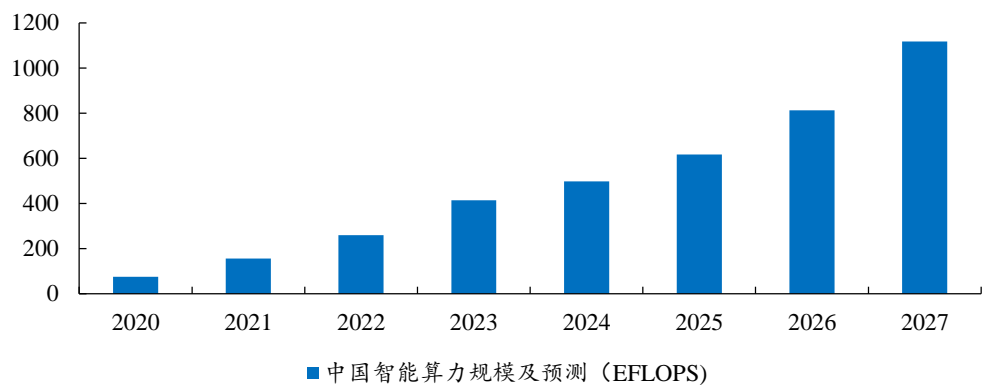
数据来源:《2024 智能算力产业发展白皮书》、中国信通院、开源证券研究所

我国智能算力规模高速增长,预计 2026 年智能算力规模突破 ZFLOPS 量级,

增长速度远超通用基础算力。在大模型取得突破、应用场景广泛开拓与深入发展的背景下，智能算力需求将在未来几年迎来井喷。随着数据量高速增长、数据要素化进程推进、同时算力模型复杂度日益提升，智能算力作为释放数据价值的必要工具，其需求与规模快速增长。openAI 分析显示，从 2012 年以来，最大规模的 AI 模型训练中所需要的计算量，每 3.5 个月便翻倍一次。相比于摩尔定律 2 年的倍增期，算力需求具有远高于芯片承载计算量的增长速度。

根据 IDC 和浪潮信息发布的《2023-2024 年中国人工智能计算能力发展评估报告》，预计到 2027 年通用算力规模将达到 117.3EFLOPS，智能算力规模达 1117.4EFLOPS；2022-2027 年期间，预计中国智能算力规模年复合增长率达 33.9%，同期通用算力规模年复合增长率为 16.6%。

图12：预计到 2027 年中国智能算力规模达 1117.4EFLOPS



数据来源：《2023-2024 年中国人工智能统计算力发展评估报告》、开源证券研究所

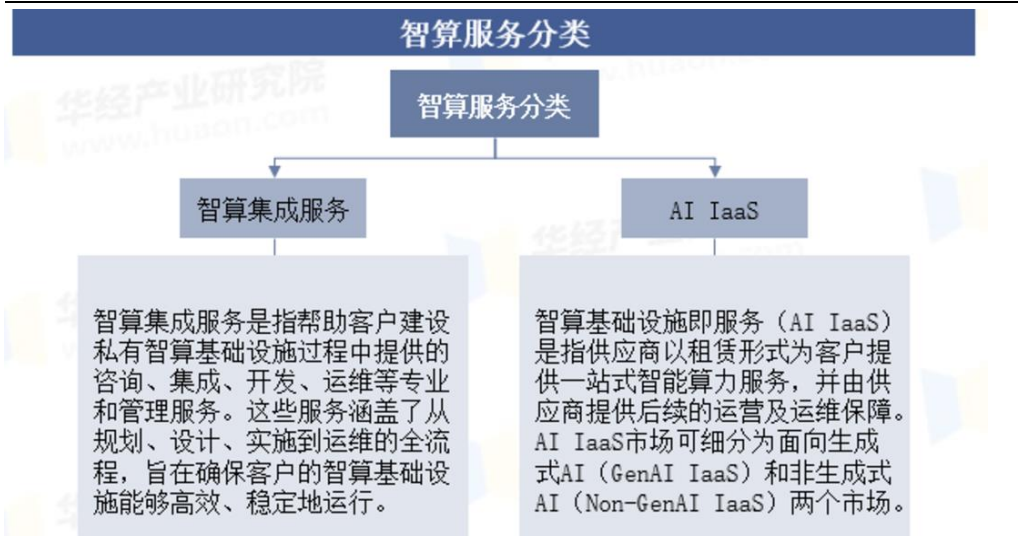
2.3、智算服务市场：数字化转型加速，市场规模保持增长趋势

智算服务，是指以 GPU、FPGA、ASIC 等 AI 专用算力为主的基础设施服务，主要包括智算集成服务、智算基础设施即服务 (AIaaS)。

智算集成服务是指帮助客户建设私有智算基础设施过程中提供的咨询、集成、开发、运维等专业和管理服务。

智算基础设施即服务 (AIaaS) 则是以租赁形式为客户提供一站式智能算力服务，这种服务允许客户根据自己的需求，灵活地使用和扩展智能计算资源，而无需承担高昂的基础设施建设和运维成本。

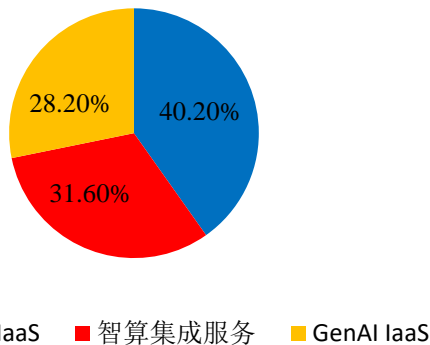
图13: 智算服务, 是指以 GPU、FPGA、ASIC 等 AI 专用算力为主的基础设施服务



资料来源: 华经产业研究院

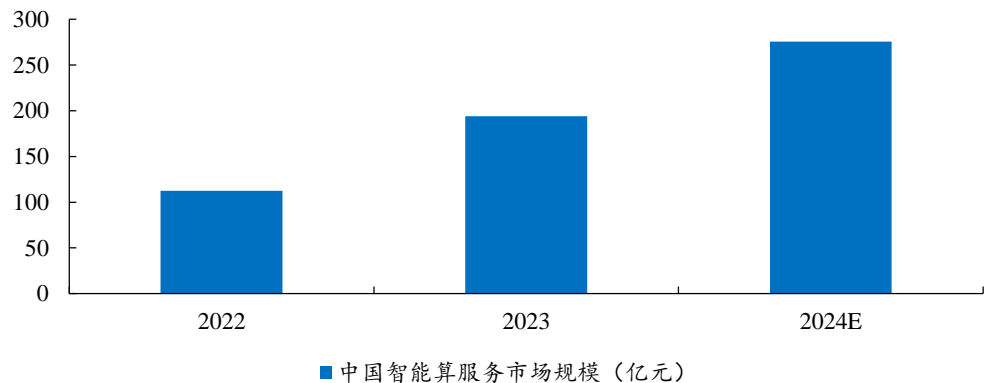
智算服务市场主要由智算集成服务和智算基础设施即服务 (AI IaaS) 两大板块构成。其中, 智算集成服务市场贡献了主要增量, 而 AI IaaS 市场则进一步细分为面向生成式人工智能的 GenAI IaaS 和面向非生成式人工智能的 Non-GenAI IaaS, 伴随着头部 GenAI 产品的逐步落地, 智能算力需求增长。2023 年智算服务市场以 Non-GenAI IaaS 为主, 目前占整体的 40.2%。智算集成服务和 GenAI IaaS 分别为 31.6% 和 28.2%。

图14: 2023 年智算服务市场以 Non-GenAI IaaS 为主, 目前占整体的 40.2%



数据来源: IDC、中商产业研究院、开源证券研究所

IDC 数据显示, 近两年来中国智算服务市场规模保持增长趋势, 2023 年市场规模约为 194.2 亿元, 同比增长 72.49%。中商产业研究院预测, 2024 年市场规模将超过 200 亿元。

图15：预计 2024 年中国智算服务市场规模将超过 200 亿元


数据来源：IDC、中商产业研究院、开源证券研究所

未来，随着云计算、大数据、深度学习等技术的不断发展，智算能力有望大幅提升，创新生态不断优化，同时，数字化转型的加速和智能化应用的普及，有望使得从政府到企业，对智算服务的需求不断扩大。

在“东数西算”工程和各类促进智能计算中心建设的政策引导下，电信运营商、互联网企业、设备制造商等产业链上下游企业将加强合作，共同推动智算服务产业的发展，各地智算中心的建设和运营有望促进区域间的产业协同和资源共享，智算中心的区域布局更加合理，形成东中西部协同发展的格局。

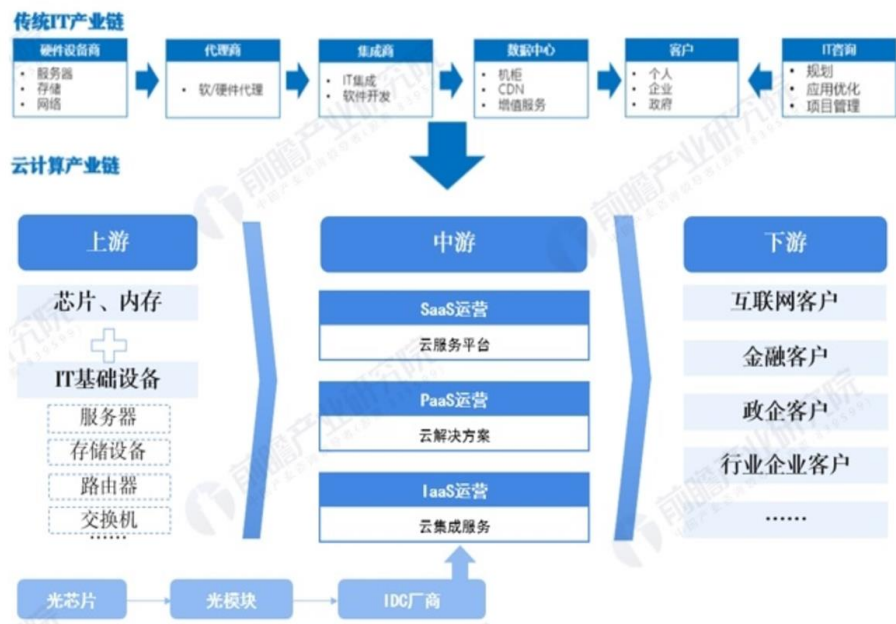
此外，智算服务或将从单一的算力提供向提供综合解决方案转变，包括算法、数据处理和智能应用等一体化服务。

2.4、云计算市场：预计 2024 年市场规模将达 8378 亿元

云计算(CloudComputing)是一种按使用量付费的模式，这种模式提供可用的、便捷的、按需的网络访问，进入可配置的计算资源共享池(资源包括网络，服务器，存储，应用软件，服务)，这些资源能够被快速提供，只需投入很少的管理工作，或服务供应商进行很少的交互。

从产业结构来看，我国目前云计算产业链结构完备，上游为核心硬件(芯片、内存等)、IT 基础设备(服务器、存储设备、网络设备)，中游为 IaaS、PaaS、SaaS 运营，下游为各类使用云的企业或组织。

图16: 从产业结构来看, 我国目前云计算产业链结构完备



资料来源: 前瞻产业研究院

云计算产业链的核心是云服务厂商, 海内外主要的厂商有亚马逊、微软、谷歌、Facebook、苹果、阿里、腾讯等互联网转型企业, 提供弹性计算、网络、存储、应用等服务。互联网数据中心(IDC)厂商为之提供基础的机房、设备、水电等资源。基础设备提供商将服务器、路由器、交换机等设备出售给 IDC 厂商或直接出售给云服务商, 其中服务器是基础网络的核心构成, 大约占到硬件成本的 60%-70%。CPU、BMC、GPU、内存接口芯片、交换机芯片等是基础设备的重要构成。光模块是实现数据通信的重要光学器件, 广泛用于数据中心, 光芯片是其中的核心硬件。云计算产业最终服务于互联网、政府、金融等广大传统行业与个人用户。

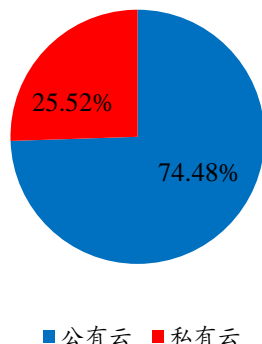
图17: 云计算产业最终服务于互联网、政府、金融等广大传统行业与个人用户



资料来源: 前瞻产业研究院

分类来看, 2023 年中国公有云市场规模 4562 亿元, 同比增长 40.1%, 占国内云计算市场的比重为 74.48%; 私有云市场规模 1563 亿元, 同比增长 20.8%, 占国内云计算市场的比重为 25.52%。

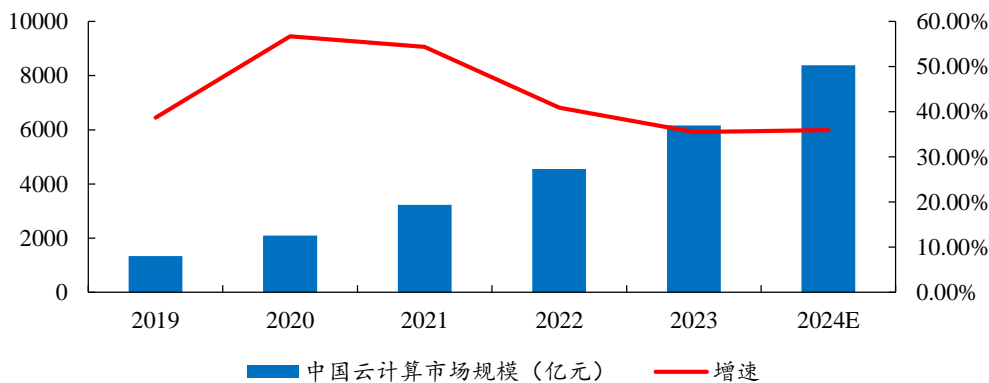
图18: 2023年中国公有云市场规模4562亿元, 同比增长40.1%



数据来源: 中国通信院、中商产业研究院、开源证券研究所

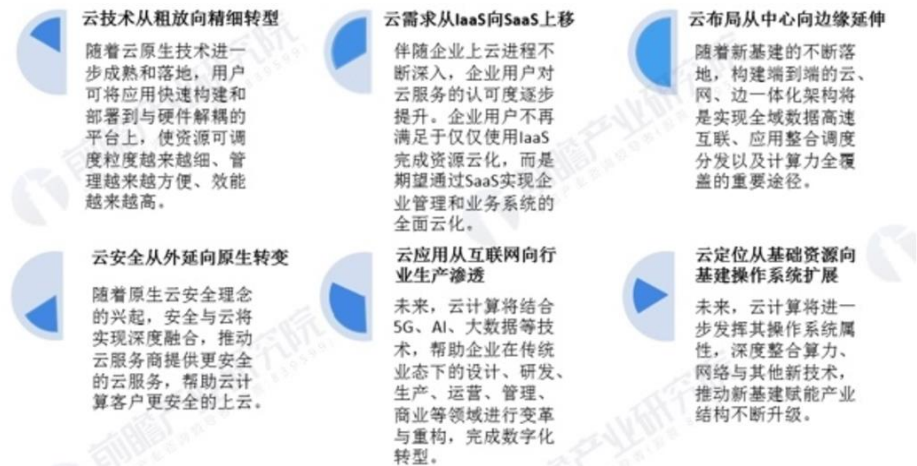
随着近年来 AI 技术的飞跃发展, 以国产数据库、大数据、专有云、操作系统为代表的产品在垂直产业深入应用, 特别是在政务、金融、交通、教育等关键基础设施领域, 为中国云计算厂商带来了全新市场机遇。2023 年中国云计算市场规模达 6165 亿元, 同比增长 35.5%, 大幅高于全球增速。预计到 2024 年, 云计算市场规模将达 8378 亿元, 同比增长 35.9%。

图19: 预计到2024年, 我国云计算市场规模将达8378亿元, 同比增长35.9%



数据来源: 中国通信院、中商产业研究院、开源证券研究所

2020 年以来, 无论是如火如荼的“新基建”、稳步推进的企业数字化转型, 还是突如其来的疫情, 都将云计算发展推向了一个新的高度。未来十年, 云计算将进入全新发展阶段。

图20：未来十年，云计算将进入全新发展阶段


资料来源：前瞻产业研究院

3、业务：GPU 算力解决方案为基础，算力云服务为核心

蓝耘科技是一家专注于为企业级规模的 GPU 加速工作负载提供算力云服务的公司，基于灵活的基础设施与大规模的 GPU 算力资源，为客户提供开放、高性能、高性价比的算力云服务。

公司成立于 2004 年，总部位于北京，在上海、广州、成都、厦门等地设有分子公司，于 2017 年挂牌新三板。公司拥有专业技术团队，从基础架构建设到云计算，从实施到售后，均有丰富的经验；已成为国家高新技术企业，北京市专精特新小巨人企业。

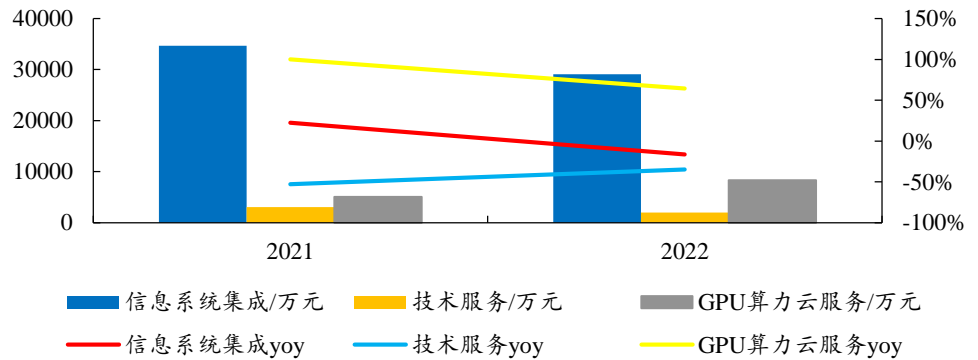
发展至今，蓝耘形成了以 GPU 算力解决方案为基础，GPU 算力云服务为核心的业务版图，为相关高校、科研机构、生命科学、人工智能、自动驾驶、工业设计、智慧城市等领域有高性能计算需求的客户，提供低成本高质量的 GPU 算力云服务，满足客户业务应用和 GPU 算力按需扩展的双重需求。蓝耘自 2019 年起积极布局算力云服务，着手 GPU 算力资源池构建，至今已形成以 GPU 算力解决方案为基础、GPU 算力云服务为核心的业务布局，基于灵活的基础设施、丰富的 GPU 算力资源和技术运营经验优势，依托智算中心“绿色基座”，致力于为 AI 训练、推理、视觉特效和渲染及教科研等计算密集型场景用户

3.1、产品：GPU 算力云业务占比持续提升，2024H1 毛利率达 39.16%

公司的主营业务为 IT 系统集成业务、技术服务业务、GPU 算力云服务业务，其中 GPU 算力云服务业务自 2021 年以来高速增长，近年来占比不断上升，2023 年公司算力云解决方案、GPU 算力云服务业务稳步推进，其中算力云解决方案收入占比为 66.14%，GPU 算力云服务收入占比从 2022 年的 21.16%稳步提升至 33.86%，公司业务结构持续优化，GPU 算力云服务业务营收规模保持稳健增长，经营质效稳步提升，且毛利率达到 39.63%，高于公司传统的系统集成业务。2024H1 算力云解决方案实现收入 4.11 万元，同比增长 380.63%，GPU 算力云服务业务实现收入 1.25 亿元，同比增长 139.74%。

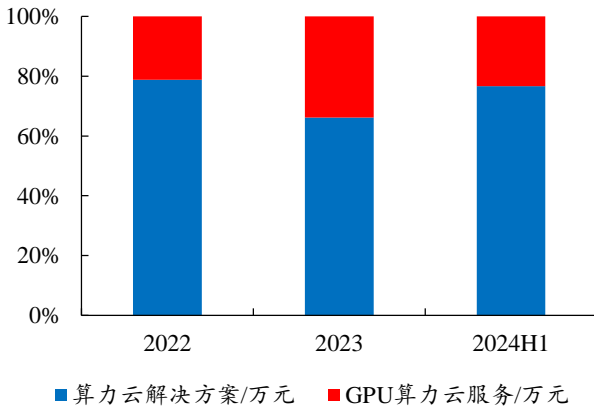
从业务内容来看，IT 系统集成业务主要产品与服务包括企业私有云解决方案、容灾备份解决方案、高性能计算 HPC 解决方案以及分布式存储系统解决方案；技术服务业务主要包括 IT 系统运维服务，软件系统开发服务；GPU 算力云服务主要是基于蓝耘 GPU 云 IAAS 基础设施，根据客户业务对 CPU 及 GPU 算力的需求，通过算力云平台为客户提供远程 CPU 或 GPU 多核并行算力，客户按需购买算力核时或卡时预付费给蓝耘科技。蓝耘算力云平台可以向用户提供虚拟化的计算资源、存储资源、网络资源等，这些资源能够根据用户的需求动态的分配。

图21: GPU 算力云服务业务营收规模保持稳健增长，经营质效稳步提升



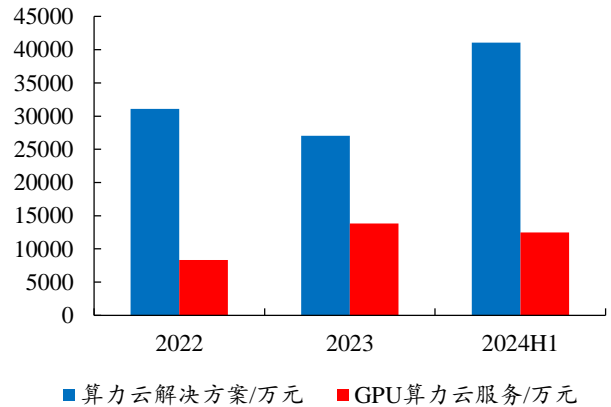
数据来源：Wind、公司财报、开源证券研究所

图22: 2024H1 算力云服务收入占比为 23.33%



数据来源：Wind、公司财报、开源证券研究所

图23: 2024H1 GPU 算力云服务业务营收达到 1.25 亿元



数据来源：Wind、公司财报、开源证券研究所

3.2、商业模式：自建智算云平台，服务高性能计算需求行业龙头企业

蓝耘算力云业务以 IaaS+Pass 为基础架构，依托数据中心，交付给用户高效、灵活、可扩展的基础设施资源，用户无需购买、维护硬件设备和相关系统软件，就可以直接在蓝耘云平台上构建自己的平台和应用。蓝耘算力云平台可以向用户提供虚拟化的计算资源、存储资源、网络资源等。这些资源能够根据用户的需求动态的分配。支撑该服务的技术体系包括虚拟化技术和相关资源动态管理与调度技术。

蓝耘 GPU 算力云业务目标用户包括个人用户与企业用户，个人用户的算力云服

务应用场景包括机器学习、计算机视觉、自然语言处理等 AI 领域的培训与学习，以及高校相关专业学生的毕业论文设计、科研项目基础研究等，涉及目标院系包括计算机、人工智能、物理化学、生命科学、环境科学、地理勘测、航空航天等。企业用户的算力云服务应用场景主要集中在自动驾驶、AI 大模型、生命科学、工业设计以及教科研领域的大型科研项目等。

2022-2023 年，蓝耘科技向前五大客户销售金额分别为 1.46 亿元、1.78 亿元，前五大客户收入占比分别为 37.05%、43.57%。

核心客户方面，公司系统集成业务的核心客户以数字化服务、数据服务、网络通信等领域的服务商和集成商龙头企业为主。2022 年，第一、第四和第五大客户，主要交易内容为服务器（配件）及安装服务，不直接提供算力服务；而 2022 年 GPU 算力云服务的主要客户包括**第二大客户（软件和信息技术服务商）、第三大客户（科技推广和应用服务商，致力于开发全球最大的中文搜索引擎之一的 ROI 媒体平台）**这类具有较大 GPU 算力加速需求的厂商。

表3：2023 年蓝耘科技向前五大客户销售金额为 1.78 亿元

| 客户排名 | 收入/万元 | 销售占比 |
|------|----------|--------|
| 2023 | | |
| 客户 1 | 4778.76 | 11.70% |
| 客户 2 | 4341.58 | 10.63% |
| 客户 3 | 3146.11 | 7.70% |
| 客户 4 | 2851.65 | 6.98% |
| 客户 5 | 2680.93 | 6.56% |
| 合计 | 17799.03 | 43.57% |
| 2022 | | |
| 客户 1 | 4675.75 | 11.87% |
| 客户 2 | 2796.88 | 7.10% |
| 客户 3 | 2680.93 | 6.80% |
| 客户 4 | 2607.19 | 6.62% |
| 客户 5 | 1838.46 | 4.67% |
| 合计 | 14599.22 | 37.05% |

数据来源：公司年报、开源证券研究所

经营模式方面，蓝耘目前正围绕 GPU 算力打造“系统集成及服务”+“算力资源”+“云服务”为一体的综合业务版图。

业务流程来看，算力云服务平台是基于 Kubernetes 原生云设计，专为大规模 GPU 加速工作量身打造，搭载高可配置高可用的英伟达高端 GPU 系列，支持按需灵活计费、轻松调整存储容量及扩展容量，客户可直接线上付费购买算力资源、进行服务器租借并实现定制化服务等等；蓝耘科技以自建云平台为主，需要采购各类应用和基础软件。此外，公司目前虽然系统集成业务收入占比最高，但承接的系统集成业务主要也是 GPU 服务器相关的 IT 系统集成，获取订单后由公司售前及商务人员配合根据客户需求进行拆单和选型，之后根据公司库存情况将需进行采购的部件向供应商下单采购。最后 IT 硬件到达公司仓库或客户现场后，由公司实施人员进行集成，集成部分工作较为简单，主要为软硬件的连接与调试；公司把集成中的服务分成

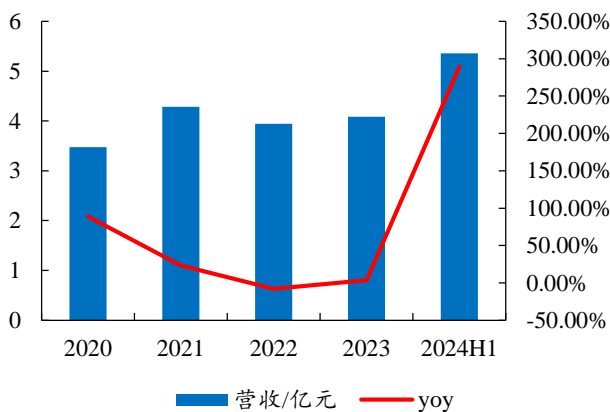
标准服务，并通过视频或现场指导外包服务团队按标准提供服务的方式完成集成，最后完成客户交付。公司业务过程中向不同客户提供的都是高度相似的类标准化服务，供应链体系和商务体系较为完善，在人员配备(1-2人)上无需较大投入。

采购项目方面，公司采购的主要为 IT 系统集成业务和 GPU 算力云服务业务项目所需的计算机信息系统相关的软件、硬件产品（如服务器、交换机、小型机管理控制台）、技术服务等。

3.3、财务：2024H1 实现营收 5.36 亿元，归母净利润 4827.93 万元

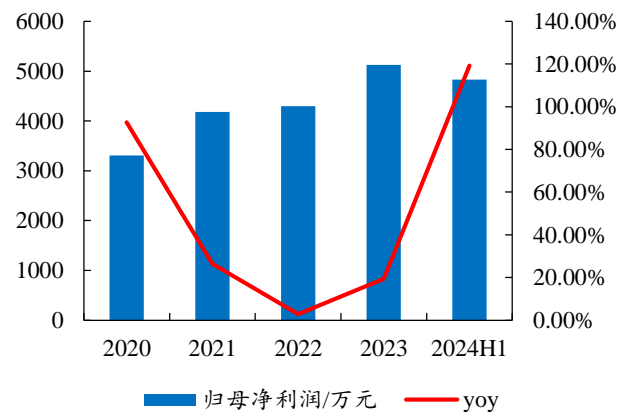
公司营收和盈利规模稳定增长，成长性良好。受益算力建设需求，公司业务快速增长，2023 年实现营收 4.09 亿元、归母净利润 5128.17 万元。公司算力云解决方案、GPU 算力云服务业务稳步推进，蓝耘智算云平台于 2022 年 4 月正式上线，基于蓝耘云基础设施，面向高校、科研院所、企事业单位等高性能计算需求的客户，满足客户业务应用和 GPU 算力按需扩展的双重需求。另外 2024 年 4 月公司签订重大销售合同，合同金额为 1.04 亿元。2024H1 实现营收 5.36 亿元，归母净利润 4827.93 万元。

图24：2024H1 实现营收 5.36 亿元



数据来源：Wind、开源证券研究所

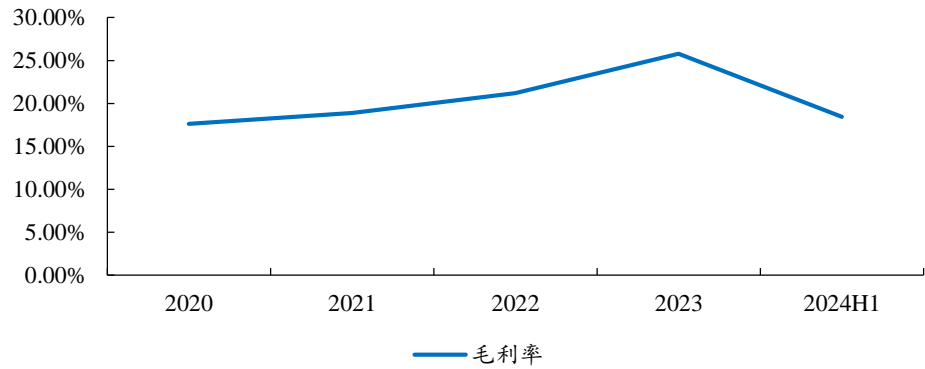
图25：2024H1 实现归母净利润 4827.93 万元



数据来源：Wind、开源证券研究所

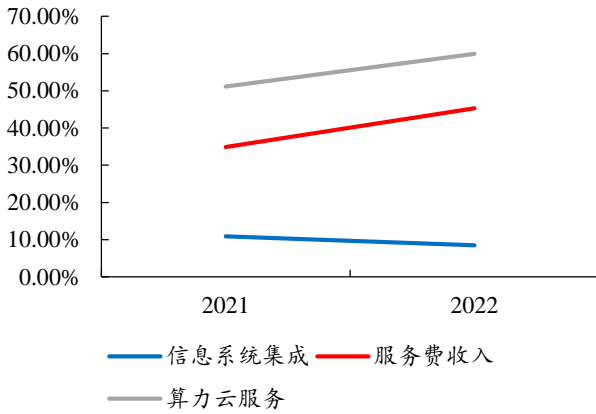
2021 年公司业务开始结构性优化，着力发展高毛利率的 GPU 算力云业务，推动公司盈利能力快速上升，2023-2024H1 毛利率分别为 25.79%、18.44%。

图26：2024H1 毛利率为 18.44%



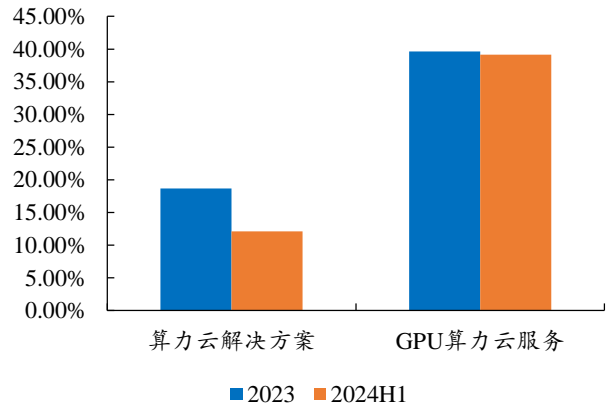
数据来源：Wind、开源证券研究所

图27：2021-2022 算力云服务毛利率高于系统集成业务



数据来源：公司年报、开源证券研究所

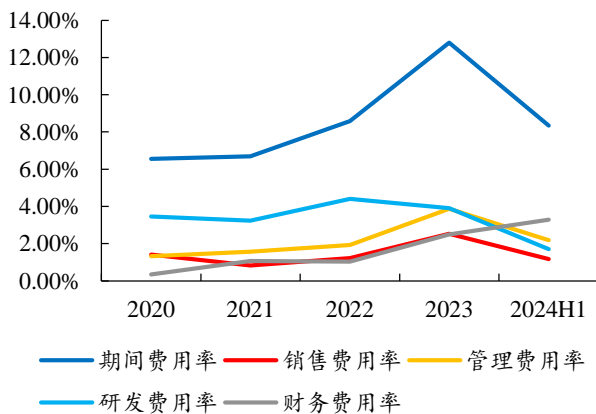
图28：2024H1 GPU 算力云服务毛利为 39.16%



数据来源：Wind、开源证券研究所

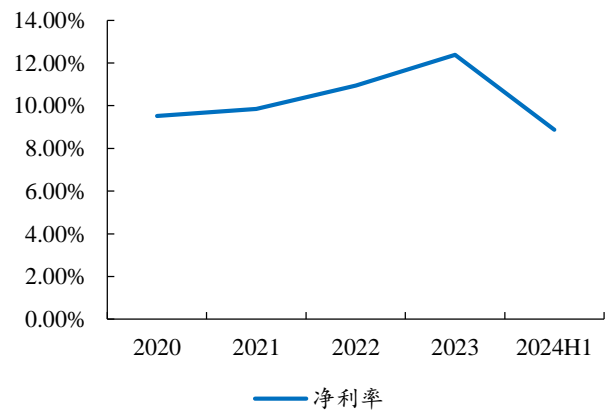
期间费用率呈先上升后下降态势，2020-2024H1，期间费用率分别为 6.55%、6.70%、8.59%、12.80%、8.35%，主要原因为公司业务规模扩大，营销费用及人员薪酬费用上升。净利率呈现波动趋势，2023-2024H1 净利率分别为 12.38%、8.87%。

图29：2024H1 公司期间费用率为 8.35%



数据来源：Wind、开源证券研究所

图30：2024H1 净利率为 8.87%



数据来源：Wind、开源证券研究所

4、估值对比：同行可比公司 PE2023 均值为 73.0X

蓝耘科技专注于为企业级规模的 GPU 加速工作负载提供算力云服务，服务于高校科研、生命科学、人工智能、自动驾驶、工业设计、智慧城市等领域，主要业务包括 IT 系统集成业务和 GPU 算力云服务业务两大模块，GPU 算力云服务业务自 2021 年以来高速增长，目前与产业方广泛合作并通过融资租赁等方式加速 AI 时代布局，有望拓展算力版图，不断提升产业地位，在 LLM 等需求驱动下前景广阔。而 2024 年 11 月，公司正式发布了蓝耘智算云平台——元生代，实现从数据准备、代码开发、模型训练到推理部署等全场景覆盖，更进一步打通用户 AI 研发全流程，会有较大的收入增长空间。

目前蓝耘科技可比公司 PE 2023 均值 73.0X，蓝耘科技 PE 2023 为 21.00X，估值相对较低，同时毛利率高于可比公司平均，具备一定可比优势。

表4：同行业可比公司 PE2023 均值为 73.0X

| 公司名称 | 股票代码 | PE TTM | 市值/亿元 | 2023 年归母净利润/百万元 | 2023 营业收入/亿元 | 2023PE | 2023 毛利率 |
|------|-----------|---------|--------|-----------------|--------------|---------|----------|
| 鸿博股份 | 002229.SZ | -87.77 | 73.88 | -54.44 | 6.20 | -123.95 | 24.95 |
| 并行科技 | 839493.BJ | -158.71 | 40.50 | -80.50 | 4.96 | -50.09 | 25.63 |
| 云赛智联 | 600602.SH | 100.88 | 167.77 | 192.91 | 52.64 | 101.52 | 17.99 |
| 浪潮信息 | 000977.SZ | 32.68 | 748.58 | 1782.78 | 658.67 | 44.49 | 10.04 |
| | 均值 | 66.78 | 257.68 | 987.85 | 180.61 | 73.01 | 19.65 |
| 蓝耘科技 | 871169.NQ | 13.89 | 10.77 | 51.28 | 4.09 | 21.00 | 25.79 |

数据来源：Wind、开源证券研究所 注：数据截至 20241223，PE2023、PE TTM、2023 年归母净利润/万元均值计算剔除了负值

5、风险提示

算力需求波动、产业政策变化、技术迭代风险

特别声明

《证券期货投资者适当性管理办法》、《证券经营机构投资者适当性管理实施指引（试行）》已于2017年7月1日起正式实施。根据上述规定，开源证券评定此研报的风险等级为R4（中高风险），因此通过公共平台推送的研报其适用的投资者类别仅限定为专业投资者及风险承受能力为C4、C5的普通投资者。若您并非专业投资者及风险承受能力为C4、C5的普通投资者，请取消阅读，请勿收藏、接收或使用本研报中的任何信息。

因此受限于访问权限的设置，若给您造成不便，烦请见谅！感谢您给予的理解与配合。

分析师承诺

负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映分析人员的个人观点。负责准备本报告的分析师获取报酬的评判因素包括研究的质量和准确性、客户的反馈、竞争性因素以及开源证券股份有限公司的整体收益。所有研究分析师或工作人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中具体的推荐意见或观点有直接或间接的联系。

股票投资评级说明

| | 评级 | 说明 |
|------|------------------|-----------------------|
| 证券评级 | 买入（Buy） | 预计相对强于市场表现 20%以上； |
| | 增持（outperform） | 预计相对强于市场表现 5%~20%； |
| | 中性（Neutral） | 预计相对市场表现在-5%~+5%之间波动； |
| | 减持（underperform） | 预计相对弱于市场表现 5%以下。 |
| 行业评级 | 看好（overweight） | 预计行业超越整体市场表现； |
| | 中性（Neutral） | 预计行业与整体市场表现基本持平； |
| | 看淡（underperform） | 预计行业弱于整体市场表现。 |

备注：评级标准为以报告日后的 6~12 个月内，证券相对于市场基准指数的涨跌幅表现，其中 A 股基准指数为沪深 300 指数、港股基准指数为恒生指数、新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）、美股基准指数为标普 500 或纳斯达克综合指数。我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。

分析、估值方法的局限性说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

法律声明

开源证券股份有限公司是经中国证监会批准设立的证券经营机构，已具备证券投资咨询业务资格。

本报告仅供开源证券股份有限公司（以下简称“本公司”）的机构或个人客户（以下简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。本报告是发送给开源证券客户的，属于商业秘密材料，只有开源证券客户才能参考或使用，如接收人并非开源证券客户，请及时退回并删除。

本报告是基于本公司认为可靠的已公开信息，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他金融工具的邀请或向人做出邀请。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。客户应当考虑到本公司可能存在可能影响本报告客观性的利益冲突，不应视本报告为做出投资决策的唯一因素。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。本公司未确保本报告充分考虑到个别客户特殊的投资目标、财务状况或需要。本公司建议客户应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。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。若本报告的接收人非本公司的客户，应在基于本报告做出任何投资决定或就本报告要求任何解释前咨询独立投资顾问。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的开源证券网站以外的地址或超级链接，开源证券不对其内容负责。本报告提供这些地址或超级链接的目的纯粹是为了客户使用方便，链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

开源证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。开源证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本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。

开源证券研究所

上海

地址：上海市浦东新区世纪大道1788号陆家嘴金控广场1号楼3层
邮编：200120
邮箱：research@kysec.cn

深圳

地址：深圳市福田区金田路2030号卓越世纪中心1号楼45层
邮编：518000
邮箱：research@kysec.cn

北京

地址：北京市西城区西直门外大街18号金贸大厦C2座9层
邮编：100044
邮箱：research@kysec.cn

西安

地址：西安市高新区锦业路1号都市之门B座5层
邮编：710065
邮箱：research@kysec.cn