

## “算海计划”共建超大单体智算集群，运营效率+算力利用率双提升

### ——北交所公司深度报告

## 并行科技 (839493.BJ)

2024年08月26日

投资评级：买入（上调）

诸海滨（分析师）

zhuhaibin@kysec.cn

证书编号：S0790522080007

赵昊（分析师）

zhaohao@kysec.cn

证书编号：S0790522080002

日期	2024/8/26
当前股价(元)	36.20
一年最高最低(元)	80.93/32.99
总市值(亿元)	21.08
流通市值(亿元)	9.41
总股本(亿股)	0.58
流通股本(亿股)	0.26
近3个月换手率(%)	70.26

北交所研究团队

#### ● “算海计划”共建超大单体智算集群，2024年上半年业绩扭亏为盈

并行科技是国内领先的超算云服务和算力运营服务提供商，主要服务包括并行通用超算云、并行行业云、并行AI云等。并行科技合作内蒙古算力基地于2024年5月15日正式上线运营，共同筹建面向模型训练的大规模单体集群。通过多年积累，公司具备分布式超算集群、算力资源网络、AI算力调优等领域的领先技术，参与多项国家重点研发计划。近年来，公司业绩持续修复，2021Q3-2024Q2，各季度营收均保持同比正增长，2024Q1公司实现扭亏为盈，2024H1归母净利润达423万元，2024H1毛利率回升至34.06%。我们上调2024-2026年盈利预测，预计2024-2026年归母净利润为-0.15/0.15/0.54(原：-0.37/0.07/0.45)亿元，对应EPS分别为-0.26/0.25/0.92元/股，对应当前股价的PS分别为3.0/2.3/1.8倍，看好公司AI算力云及行业云释放增量，上调“买入”评级。

#### ● 市场对超大规模训练需求供不应求，公司擅长围绕用户需求制定方案

近年来，公司获取算力资源方式由以外购为主逐步转向以共建为主，逐步扩充自有算力资产池，2023年11月至今，公司采购算力资产金额共计不超过4.62亿元，其中设备类采购金额约3.52亿元。在软件方面，公司推出Paramon应用运行特征收集器和分析器，可应用于高性能计算、大数据和人工智能领域。从行业来看，近期AI大模型发展较快，xAI已开始“在孟菲斯超级集群”上进行训练，2024-2025年各模态的AI将在各赛道全面走向深化成熟阶段，训练、推理等算力需求进一步提升，2023年度公司AI云实现营业收入1.1亿元，同比增长263%，公司擅长围绕用户需求制定方案，下游大模型领域的发展将为公司打开业绩新增量。

#### ● 行业云+AI云双引擎持续驱动，绘制公司第二增长曲线

并行行业云在并行通用超算云的基础上增加了业务系统对接、提供专线链路等物理层、系统层、应用层的专业化服务等以满足用户需求，提供整体解决方案，如气象海洋超算云解决方案、生命科学行业云解决方案、高端制造行业云解决方案等，公司服务了不少自动驾驶领域的科研类客户。2020年-2023年行业云业务CAGR为138.32%，2023年，行业云业务收入达到1.32亿元，占比已达到32%。

#### ● 风险提示：行业需求不及预期、新业务开拓受阻、客户合作风险

### 相关研究报告

《超算云龙头布局AI算力第二曲线，技术与算力资源巩固领先壁垒——北交所首次覆盖报告》-2024.4.21

### 财务摘要和估值指标

指标	2022A	2023A	2024E	2025E	2026E
营业收入(百万元)	313	496	701	918	1,184
YOY(%)	42.1	58.5	41.3	31.0	29.1
归母净利润(百万元)	-114	-80	-15	15	54
YOY(%)	-39.8	29.5	81.4	198.1	264.9
毛利率(%)	28.3	25.6	27.5	27.7	27.5
净利率(%)	-36.5	-16.2	-2.1	1.6	4.5
ROE(%)	-87.2	-22.4	-4.4	4.1	13.1
EPS(摊薄/元)	-1.96	-1.38	-0.26	0.25	0.92
P/S(倍)	6.7	4.3	3.0	2.3	1.8
P/B(倍)	17.3	6.3	6.6	6.3	5.4

数据来源：聚源、开源证券研究所

## 目 录

1、以“算海计划”共建超大单体智算集群，加速算力资源采购.....	4
2、市场对超大规模训练需求供不应求，公司擅长围绕用户需求制定方案.....	8
2.1、上游：英伟达为中国研制 B20 芯片，燧原科技等国产企业实现量产.....	9
2.2、下游：xAI 在“孟菲斯超级集群”训练，2023 公司 AI 云营收过亿.....	11
2.3、中游：“东数西算”等规划颁布，华为云、华南数谷等智算中心投产.....	13
3、行业云+AI 云双引擎持续驱动，绘制第二增长曲线.....	16
3.1、并行行业云近三年 CAGR 138%，解决气象海洋等多个行业痛点.....	16
3.2、并行 AI 云近三年 CAGR 253%，服务人工智能、高性能计算领域.....	19
4、运营效率+算力利用率双提升，2024H1 业绩持续修复已扭亏为盈.....	21
5、盈利预测与投资建议.....	23
6、风险提示.....	24
附：财务预测摘要.....	25

## 图表目录

图 1：公司产品包括超算云、智算云、超算行业云等六种产品.....	4
图 2：根据云服务模式，超算云可分为 IaaS、PaaS 及 SaaS 三类.....	4
图 3：合作内蒙古新型算力基地在大规模部署、自然气候等方面具有优势.....	5
图 4：获取算力资源方式由以外购为主逐步转向以共建为主，是公司设备类采购规模提升的主要原因.....	7
图 5：Paramon 应用运行特征收集器运营管理能力先进.....	7
图 6：Paratune 应用运行特征分析器用于优化程序性能.....	7
图 7：超算服务在业务超算等领域具备优势.....	8
图 8：AI 产业链上游是算力基础设施，下游包括互联网、金融、公共事业等领域.....	9
图 9：B20 芯片基于的 Blackwell 性能优于 Hopper 架构.....	10
图 10：马斯克亲自参与孟菲斯超级集群建设.....	11
图 11：马斯克自研 Dojo AI 加速器超级计算机集群.....	11
图 12：从文字到视频的各类大模型逐步发展，预计将在 2024 年进一步实现成熟应用.....	12
图 13：AI 大模型需要大算力，训练时长与模型参数量、训练数据量成正比.....	12
图 14：预计 2024 年中国 AIGC 核心市场规模将增长至 471.7 亿元.....	13
图 15：2021 年超算服务市场规模约为 196.6 亿元，预计在 2025 年达到 466 亿元.....	16
图 16：并行科技行业云在超算资源、资源弹性、资源调度程度、安全性方面具有优势.....	17
图 17：行业云 SaaS 化解决方案为用户提供底层基础设施、PaaS API 等服务.....	18
图 18：2023 年行业云收入为 1.32 亿元（+77.44%）.....	19
图 19：2023 年行业云收入占超算云收入比例为 31.89%.....	19
图 20：并行 AI 云以 GPU 为核心资源池.....	20
图 21：2023 年 AI 云业务收入为 1.11 亿元（+262.70%）.....	21
图 22：2023 年 AI 云业务收入占超算云营收 26.91%.....	21
图 23：2021Q3-2024Q2 营收呈现上升趋势.....	21
图 24：2021-2024Q2 营收保持同比增长 单位：万元.....	21
图 25：2024 年 Q2 归母净利润达 173 万元，同比上涨 111%.....	22
图 26：2024Q2 毛利率为 33.86%，Q2 净利率在 Q1 由负转正后小幅上升.....	22
图 27：2024 年 H1 期间费用率为 32.14%，较 2021 年缩减一半以上.....	23

---

图 28: 2019-2021 年公司经营活动现金净流量逐年提升, 2021-2023 年保持稳定 .....	23
表 1: 2023 年 11 月以来, 公司购买算力资产金额共计不超过 4.62 亿元 .....	5
表 2: 国内超 20 家 GPU 领域厂商发展成果显著 .....	10
表 3: 国家和各级地方颁布多项政策, 支持算力行业发展 .....	13
表 4: 浪潮集团、阿里等国内 AI 算力企业在全国多地兴建智算中心。 .....	15
表 5: 并行行业云为客户提供硬件资源、专线连接、灵活的计费模式等服务 .....	17
表 6: 并行行业云针对气象海洋、生命科学等不同行业制定解决方案 .....	18
表 7: 并行 AI 云具有高性价比、资源丰富、易用高效、服务贴心等优点 .....	20
表 8: 可比公司 PS 2024E 均值为 3.6X .....	24

## 1、以“算海计划”共建超大单体智算集群，加速算力资源采购

并行科技成立于2007年，是国内领先的超算云服务和算力运营服务提供商，聚焦应用真实运行性能，致力于为用户提供全面、安全、易用、高性价比的超算云服务。从产品业务角度来看，公司的主要产品包括超算云、智算云、超算行业云、设计仿真云、运维服务、软件产品六类产品服务。

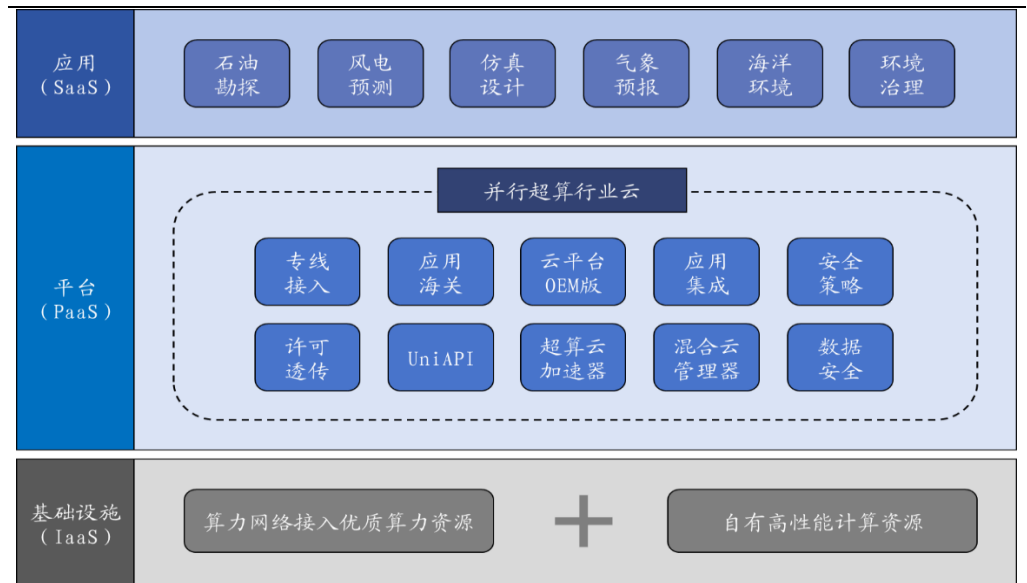
图1：公司产品包括超算云、智算云、超算行业云等六种产品



资料来源：公司官网、开源证券研究所

其中超算云服务为公司的主要业务，根据云服务模式，超算云可分为IaaS、PaaS及SaaS三类：IaaS将超级计算的基础设施提供给用户，PaaS将运行在超算云基础设施之上的软件开发和运行平台提供给用户，SaaS向用户交付运行在超算云基础设施之上完整且可以直接使用的软件应用。其中IaaS是PaaS和SaaS的基础，包括外购算力资源和自有算力资源两大类。

图2：根据云服务模式，超算云可分为IaaS、PaaS及SaaS三类



资料来源：公司招股说明书、开源证券研究所

并行科技合作内蒙古算力基地于2024年5月15日正式上线运营，并向算力建设合作伙伴发起号召：共同筹建面向模型训练的大规模单体集群。并行科技合作内蒙古新型算力基地，规划建设4000个20kW高功率智算机柜，最大可以支持单一集群6万卡。“算海计划”，寓意纳入所有江河湖泽，海纳百川。

合作内蒙古算力基地具有诸多优势：在大规模部署方面，可部署 4000 个 20kW 机柜；自然气候方面，基地所处地区年平均气温仅为 4.3°C，较低气温有助于算力基地提升散热效率、提高设备性能；同时，合作内蒙古算力基地是绿色低碳算力基地，电费单价仅为 0.262 元，有效降低算力设备运行成本；公司还拥有智能运维管理服务，专家团队 7×24 小时在线提供服务。

**图3：合作内蒙古新型算力基地在大规模部署、自然气候等方面具有优势**



资料来源：并行科技官网、开源证券研究所

近期，公司频繁进行相关资源采购，公司算力资源和相关配套 IT 设施采购主要分为设备类采购和资源类采购。设备类采购主要包括针对共建集群和部分行业云项目所需的设备采购，包括服务器、存储、网络设备等；资源类采购包括并行超算云服务开展所需的 CPU 算力资源、GPU 算力资源、互联网带宽等相关资源。2023 年 11 月至今，公司采购算力资产金额共计不超过 4.62 亿元。其中设备类采购金额约为 3.52 亿元（服务器 3.51 亿元，存储设备 100 万元），资源类采购（GPU 算力资源）约为 1.1 亿元。

**表1：2023 年 11 月以来，公司购买算力资产金额共计不超过 4.62 亿元**

序号	公告日期	交易对象	标的	数量	金额（万元）
1	2023/11/16	北京誉成云创科技有限公司、中科云达（北京）科技有限公司	AI 算力服务器(NVIDIARTX4090 型号 GPU 显卡)	92 台（736 张）	687
	2023/11/16	紫光华山科技有限公司	A800 型号显卡 AI 算力服务器	22 台	
2	2023/12/27	山东中创软件工程股份有限公司	AI 算力服务器	28 台	649.18
	2023/12/27	安擎计算机信息股份有限公司、联创万通（北京）科技有限公司、南京坤前计算机科技有限公司	AI 算力服务器	175 台	不超过 4000
	2023/12/27	北京神州数码有限公司	存储设备	4 套	不超过 100
	2024/4/10	深圳市君和信息技术有限公司	采购 CPU 服务器		628.8
3	2024/4/10	内蒙古新东吉泰科技有限责任公司	采购 AI 算力服务器		2480
	2024/4/10	深圳市四通科技控股有限公司	采购 AI 算力服务器		1444
4	2024/4/15	深圳市四通科技控股有限公司、联创万通（北京）科技有限公司	采购 AI 算力服务器		不超过 9500
5	2024/6/7	海兰云（海南）数据中心科技有限公司	采购 GPU 算力资源		不超过 11000
	2024/7/2	北京国科欣翼科技有限公司	采购 AI 算力服务器		2900
6	2024/7/2	山东正云信息科技有限公司、北京有为信通科技发展有限公司、北京国科欣翼科技有限公司	采购 AI 算力服务器		不超过 12790

资料来源：公司公告、开源证券研究所

近年来，公司采购规模有所提升，采购结构出现变化，其主要原因包括算力获取途径改变、潜在客户增加、资金实力提高和冗余算力需求增加，主要影响包括服务水平提高、生产成本降低和风险降低。

#### ➤ 主要原因

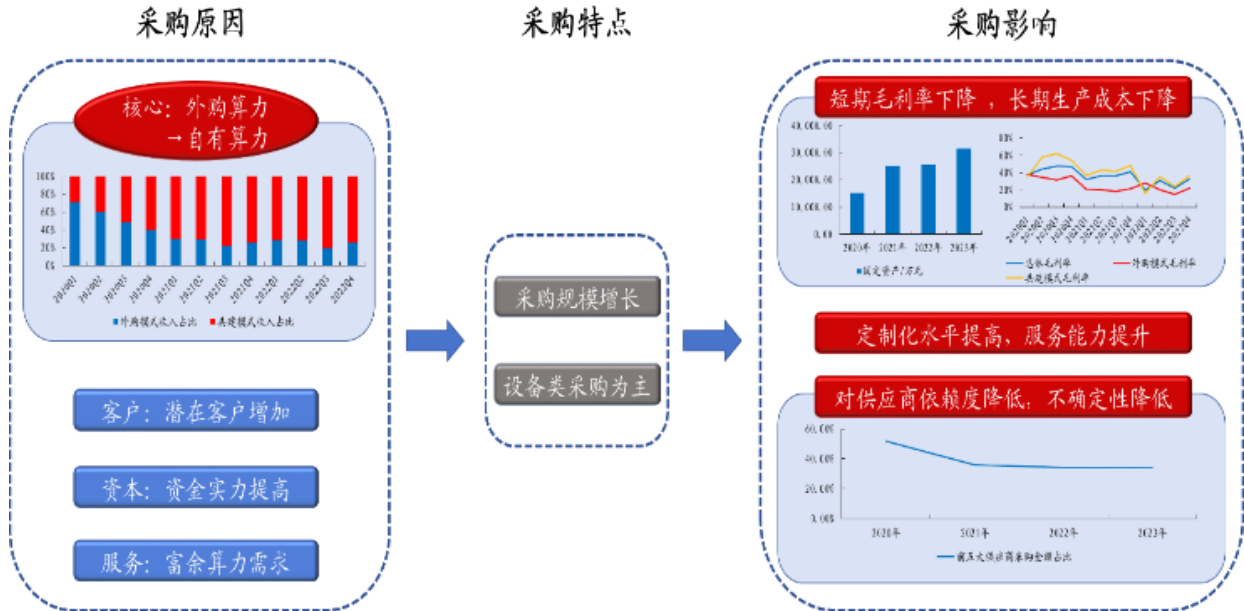
- **核心因素：**获取算力资源方式由以外购为主逐步转向以共建为主。外购模式是指公司与上游算力供应商签订算力资源采购协议，通过对接调试将其算力资源纳入公司超算云平台，并根据下游客户的需求量提供服务的形式；共建模式是指公司采购算力服务器、存储及网络设施等，租赁和使用数据中心相关设施形成计算集群，并为下游客户提供公司自有算力的形式。近年来，公司的算力资源由外购逐步转变为自有，2021-2022年，共建模式的收入占比稳定高于70%，公司算力设备需求增加，推动了设备采购金额的增长。
- **客户：**潜在客户增加推动公司算力设备需求增长。公司用户类型主要分为三类，包括高等院校、科研机构和企业用户，2023年度高校、科研机构和企业用户收入占比大致为40%、30%、30%。其中企业客户行业类型主要包括航空航天、汽车、人工智能、智能制造、能源气象等，随着企业客户AI算力需求的增长和公司品牌价值的提升，预计公司企业客户将有所拓展，贡献收入的占比将有所上升。
- **资金：**公司资金实力的提高为算力设备储备提供物质条件。2023年公司在北交所上市，募集资金总额达3.34亿元，净额达2.87亿元，资金实力有较大提升，为公司购买AI算力服务器等算力资源提供资金支持。
- **服务：**储备富余算力资源需求推动算力设备采购。从用户使用体验角度，公司需要保证一定富余的算力资源以及时提供替代资源，减少宕机、排队等情形对用户体验的影响，因此需要在布局满足收入增长所需的算力资源的基础上，布局更多的算力基础设施。

#### ➤ 主要影响

- **成本：**短期内折旧增加带来毛利率下降，长期促使生产成本下降。公司加大对算力设备相关固定资产采购以丰富自有算力资源池，2023年固定资产达3.1亿元，短期内固定资产折旧成本增加导致毛利率有所下滑。但长期来看，利用自有算力提供服务相比于外购第三方算力资源更具成本优势，共建模式毛利率高于外购模式毛利率，外购模式向共建模式转变将带来成本下降。
- **服务：**自有算力资源的增加将提高公司为客户提供服务的质量。一方面，相较于外购模式获取的算力资源，共建模式的自有算力资源可根据用户计算需求提供定制化服务。另一方面，富余算力资源可及时为客户提供替代资源，减少宕机、排队等问题的发生，提高客户体验感。
- **风险：**共建模式将减轻直接外购算力资源的不确定性。在直接采购模式下，公司算力资源的供应商为各大超算中心，各大超算中心与公司存在部分客户的重叠，具有一定竞争关系。同时，在当前强化网络和数据安全监管的背景下，各大超算中心算力资源的开放权限可能会根据法律法规及政策要求发生变化。相关业务的潜在竞争及算力资源开

放权限变动的不确定性，将增加公司风险。通过购买算力设备增加自有算力资源，将减少公司对供应商的依赖，2021年公司由外购算力转变为主为自有算力为主，前五大供应商采购金额占比为35.87%，同比下降15.67pcts，2022、2023年保持下降趋势。自有算力在一定程度上化解直接外购算力资源带来的风险。

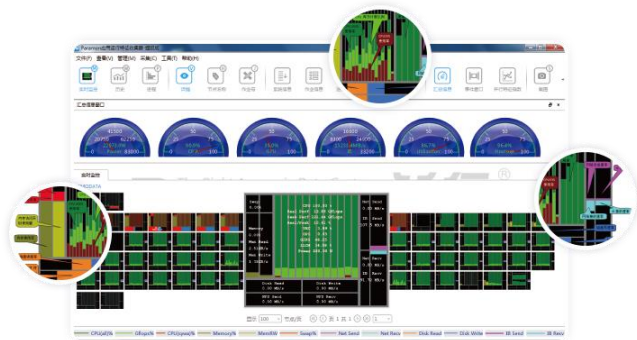
图4：获取算力资源方式由以外购为主逐步转向以共建为主，是公司设备类采购规模提升的主要原因



资料来源：公司公告、公司2023年年报及招股说明书、公司第一轮问询函回复等、开源证券研究所

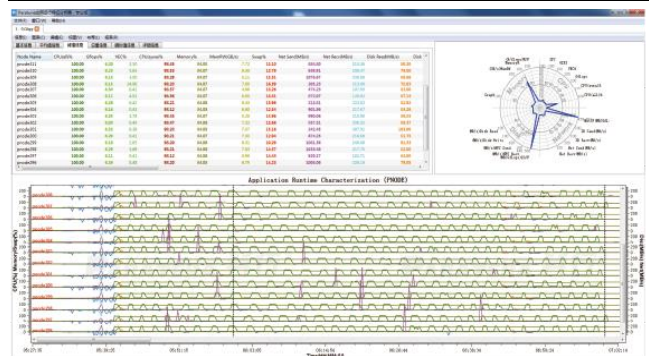
除购买硬件设施外，公司还推出了 Paramon 应用运行特征收集器和 Paratune 应用运行特征分析器两款软件产品。Paramon 广泛应用于高性能计算、大数据和人工智能领域。收集器可秒级监控集群和应用的系统级、微架构级以及函数级等性能指标，智能分析性能异常，实时告警。Paratune 分析器用于分析优化程序性能，尤其针对大规模并行计算程序，通过多节点间系统级和微架构级等性能指标定性与定量关联分析，帮助程序开发者快速定位程序性能瓶颈点，优化编程提高程序性能。

图5：Paramon 应用运行特征收集器运营管理能力先进



资料来源：公司官网

图6：Paratune 应用运行特征分析器用于优化程序性能






资料来源：公司官网

## 2、市场对超大规模训练需求供不应求，公司擅长围绕用户需求制定方案

超级计算，也称之为高性能计算，指利用并行工作的多台计算机系统的集中式计算资源，处理极端复杂的或数据密集型的问题。从算力资源的需求看，高性能计算可以分为尖端超算、通用超算、业务超算和人工智能超算四大类。在超算服务未来增长中，预计业务超算与人工智能超算将带来主要增长。

图7：超算服务在业务超算等领域具备优势

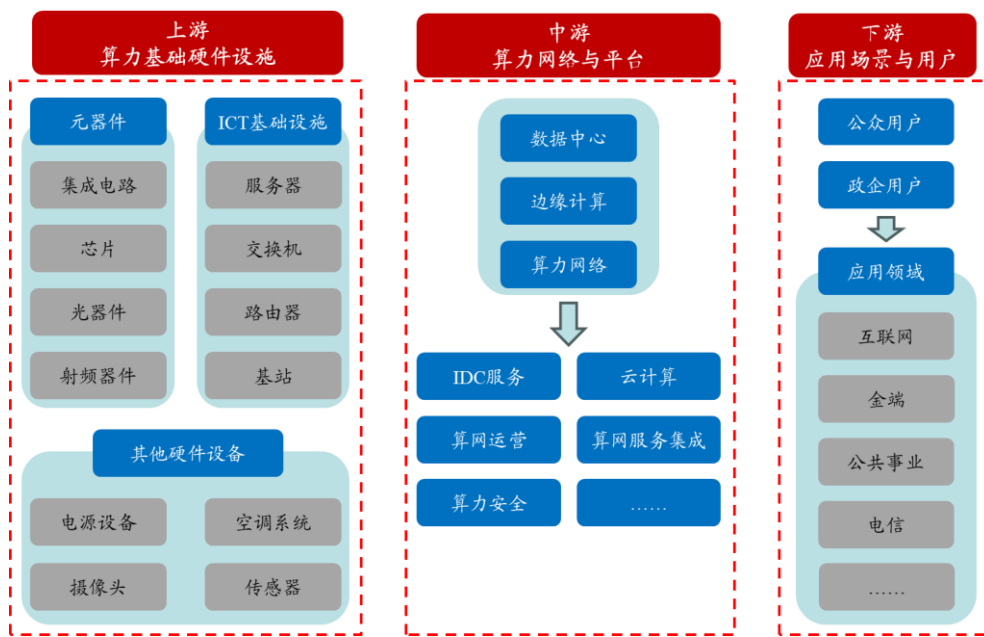
	 尖端超算	 通用超算	 业务超算
典型场景	<ul style="list-style-type: none"> <li>海洋勘探</li> <li>地球物理</li> <li>爆炸模拟</li> <li>灾害预警</li> <li>精尖制造</li> <li>材料研发</li> <li>航空航天</li> <li>气象监测</li> </ul>	<ul style="list-style-type: none"> <li>高校科研</li> <li>石油勘探</li> <li>生物制药</li> <li>工业仿真</li> </ul>	<ul style="list-style-type: none"> <li>芯片仿真</li> <li>汽车机械</li> <li>金融经济</li> <li>多媒体渲染</li> </ul>
核心需求	<ul style="list-style-type: none"> <li>以国家级的技术创新、理论验证和精密测试为主</li> <li>对计算资源和性能有极致的需求，相比通用超算与业务超算，对计算结果的精度和准度有着最高的标准和期望</li> </ul>	<ul style="list-style-type: none"> <li>核心需求仍然以产业级的技术创新和验证为主，但相比尖端超算，对项目时间的敏感性更低，故而对超算服务的经济性有一定需求</li> </ul>	<ul style="list-style-type: none"> <li>业务超算相比尖端超算和通用超算，核心需求以满足其产品业务优化为主，对超算服务的需求相对零散，对极致性能的追求也更加有限</li> </ul>
经济性要求	<ul style="list-style-type: none"> <li>尖端超算项目的经济性要求非常低，满足极致的计算服务是其核心标准</li> <li>由此，尖端超算项目通常由政府主导，或是在政府支持下举办</li> </ul>	<ul style="list-style-type: none"> <li>通用超算的经济性要求中等，会在极致性能与项目成本之间做一定的取舍</li> <li>此外，通用超算的性能要求比尖端超算较低，集群式的云服务在一定程度上能够满足通用超算的要求</li> </ul>	<ul style="list-style-type: none"> <li>企业作为以盈利为目标的商业主体，通常对经济性的要求非常高，在投入、产出比方面有着严格的把控，此外企业对超算服务的付费模式也有着较高的要求，超算云服务凭借灵活的收费机制，对企业客户吸引力更大</li> </ul>

资料来源：沙利文咨询、开源证券研究所

从产业链维度而言，算力产业链上游是算力基础硬件设施，包括元器件、ICT基础设施、其他硬件设备三大类。中游是算力网络与平台，数据中心、边缘计算、算力网络等为用户提供 IDC 服务、云计算服务等服务。产业链下游应用领域主要包括互联网、金融、公共事业等行业。随着 AI 产业快速发展，算力行业从上游芯片到下游应用均持续进展。



图8: AI 产业链上游是算力基础设施, 下游包括互联网、金融、公共事业等领域



资料来源：前瞻产业研究院、开源证券研究所

## 2.1、上游：英伟达为中国研制 B20 芯片，燧原科技等国产企业实现量产

从上游来看，芯片是算力基础设施中技术难度较高的器件，近期海内外芯片产业均有较大突破，为算力行业的发展提供条件。2024 年 7 月 22 日，据路透社报道，英伟达正在为中国市场开发新旗舰 AI 芯片 B20。该芯片基于英伟达 2024 年 3 月推出的“最强 AI 芯片”Blackwell 系列，将与目前美国政府对华芯片出口管制相兼容，B20 芯片计划于 2025 年第二季度开始出货。自 2023 年美国政府加强对中国芯片出口限制以来，英伟达先后为中国市场定制了三款专用芯片：L20 PCIe、L2 PCIe、HGX H20，B20 即将为第四款定制芯片。路透社认为，H20 系列芯片由于性能大幅缩减、定价相对较高，使得 H20 芯片对华交付时开局表现不佳。随着美国政府在 10 月份对半导体出口管制政策的年度审查，英伟达 H20 GPU 或将面临销售禁令。该禁令可能包括对特定产品的直接禁售、对芯片计算能力的进一步降低，或是对内存容量的限制。对此，英伟达 B20 芯片在继承 Blackwell 系列优秀性能的基础上，针对中国市场的特殊需求进行了深度优化，与之前 Hopper 架构相比，Nvidia 声称在原始浮点精度方面快了 2.5~5 倍。

**图9：B20 芯片基于的 Blackwell 性能优于 Hopper 架构**

Blackwell vs Hopper Basic Specifications					
	H100	H200	B100	B200	GB200 NVL72
Price <sup>1</sup>	\$24,000	\$24,000	\$30,000		
Watts Per GPU	700	700	700	1,000	1,200
All-in System Watts Per GPU	1,275	1,275	1,275	1,788	1,667
NVLink Bandwidth (Unidirectional - GB/s)	450	450	900	900	900
Memory Capacity (GB)	80GB	141GB	Up to 192GB <sup>3</sup>	Up to 192GB	192GB
Memory Bandwidth (GB/s)	3,352	4,800	Up to 8000 <sup>3</sup>	Up To 8,000	8,000
Memory Bandwidth Improvement	0%	43%	79%	139%	139%
TF32 TFLOPS <sup>2</sup>	495	495	900	1,100	1,250
TF32 Improvement	0%	0%	77%	127%	153%
FP16/BF16 TFLOPS <sup>2</sup>	989	989	1,750	2,250	2,500
FP16/BF16 Improvement	0%	0%	77%	127%	153%
FP8 / FP6 / Int8 TFLOPS <sup>2</sup>	1,979	1,979	3,500	4,500	5,000
FP8 / FP6 / Int8 Improvement	0%	0%	77%	127%	153%
FP4 TFLOPS <sup>2</sup>	1,979	1,979	7,000	9,000	10,000
FP4 Improvement	0%	0%	254%	355%	405%

1. Full ASP and volume details available in Accelerator Model  
 2. All FLOPS are dense  
 3. B100 memory specs to be finalized - we assume 168GB and 6,000 GB/s

资料来源：semianalysis、半导体产业研究公众号

**在国内芯片市场**，华为、腾讯支持的燧原科技等国产厂商也取得显著进展，国产 GPU 在产品上取得了一定的突破。行业内超过 20 家 GPU 领域的厂商中，已有许多厂商将 GPU 芯片产品投入量产。**国产 GPU 领域发展迅速。**

**表2：国内超 20 家 GPU 领域厂商发展成果显著**

序号	主要厂商	技术背景	应用领域
1	中科曙光	高性能 GPU	高性能计算机、存储设备的研发制造，以及配套系统、软件和其他技术服务
2	凌久电子	GP101 实现了我国通用 3D 显卡	涉及船舶、航空、航天等电子邻域及轨道、能源、电力、半导体制造等高科技行业
3	瀚博半导体	GPU 专家做 DSA 架构设计云端 AI 推理芯片	深度学习、AI 推理、云计算等领域
4	景嘉微	国防科大、自研	应用于人工智能、云计算、金融、教育等领域
5	中船重工	JARIG12 是 2018 年性能最强的国产通用图形处理器	应用于量子通信领域
6	芯原微电子	收购图芯美国，获得了 GPU IP	专注于服务广泛的应用市场，包括数字多媒体、语音及无线通讯市场
7	燧原科技	GPGPU AI 芯片	广泛用于云数据中心、超算中心、泛互联网、传统行业及智慧城市等多个人工智能场景
8	航锦科技	并购的长沙韶光	应用于监测检测、精准农业、测量测绘、共享单车等领域
9	璧仞科技	GPGPU/图形 GPU	聚焦云端通用智能计算，逐步在人工智能训练和推理、图形渲染等多个领域发展
10	沐曦集成电路	高性能 GPU 领域	应用于高性能计算、数据中心及 AI 等方面的计算需求
11	摩尔线程	首颗国产全功能 GPU 研制成功	物理仿真、AI 计算、游戏娱乐、自动驾驶等领域
12	兆芯	上海国资和中国台湾威盛，世界上第三家拥有 X86 授权的微处理器公司	致力于技术创新与兼容主流的发展路线
13	登临科技	GPU+为核心云端 AI 计算平台公司	应用在安防、交通、医疗、金融、能源、电力、智

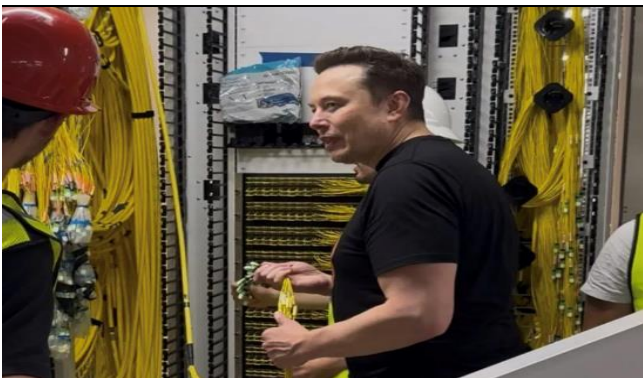
序号	主要厂商	技术背景	应用领域
			能驾驶、互联网等
14	天数智芯	中国第一家通用 GPU 高端芯片及超级算力提供商	自动驾驶、智慧医疗、智慧金融、智慧教育等领域
15	芯动科技	中国一站式 IP 和 GPU 领军企业，分别实现了国产服务器和桌面 GPU 跑分纪录并已大规模量产	5G 数据中心/服务器，云游戏/云桌面/云渲染，以及桌面/笔记本电脑和嵌入式工控机等主流商业应用
16	龙芯	集成 GPU 集成显卡	金融、政务办公、网安、教育通信、医疗等领域
17	芯瞳半导体	党政八大行业、云游戏	信创、能源、交通、金融、人工智能等领域
18	海思半导体	自研 GPU、嵌入式 GPU	覆盖无线网络、固定网络、数字媒体等领域
19	深圳中微电	自主可控	应用于智能家居，边缘计算和人工智能等领域
20	昆仑芯	前身是百度智能芯片及架构部	深耕人工智能芯片领域
21	智绘微	国产自主可控创新型 GPU 芯片设计的细分领域	高性能计算机系统所需的图形渲染、图像处理、虚拟现实、人工智能等通用处理器芯片
22	深流微智能	专注于下一代 GPU 芯片设计研发，拥有世界领先的下一代超级流处理 XST 架构	应用于视觉计算（图形渲染、虚拟现实、元宇宙）、人工智能、高性能计算
23	砺算科技	专注自研架构、自有知识产权，打造对标国际主流的国产 GPU 芯片	现实加速、游戏电竞、数字孪生、元宇宙、VR/AR、超高清视频以及专业设计等多个领域

资料来源：EDA365 电子论坛公众号、开源证券研究所

## 2.2、下游：xAI 在“孟菲斯超级集群”训练，2023 公司 AI 云营收过亿

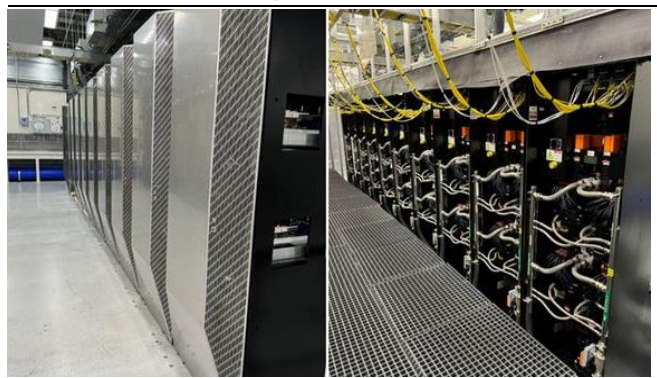
从下游来看，大模型领域也有较大进展。2024 年 7 月，埃隆·马斯克宣布旗下 xAI 已开始“孟菲斯超级集群”上进行训练。该集群拥有 10 万台液冷 H100 GPU，利用单个 RDMA 结构(远程直接内存访问)，所有 GPU 可以高效地共享和传输数据，从而极大地提高了计算效率。此外，H100 GPU 具备第四代 Tensor Cores 和 Transformer Engine，支持 FP8 精度，能够显著提升大型语言模型的训练速度和推断速度。以 H100 芯片为例，H100 可以在大型语言模型上实现高达 9 倍的训练加速和 30 倍的推理速度。同时，H100 GPU 还通过第二代 Multi-Instance GPU (MIG)、内置 NVIDIA 保密计算和 NVLink 交换系统来确保所有工作负载的安全加速，在安全性方面提供保障。

图10：马斯克亲自参与孟菲斯超级集群建设



资料来源：半导体行业观察公众号

图11：马斯克自研 Dojo AI 加速器超级计算机集群



资料来源：半导体行业观察公众号

2023 年，AI 大模型已能够成熟应用于文本、代码领域，在图形、视频、3D、游戏等领域也已接近成熟。据甲子光年预测，2024-2025 年各模态的 AI 将在各赛道全面走向深化成熟阶段，可实现文本撰写水平接近人类平均值，根据文本生成初版应用程序，进行图像等产品设计，生成短视频等。

图12：从文字到视频的各类大模型逐步发展，预计将在 2024 年进一步实现成熟应用

领域类型	大模型成熟难度： 初级尝试 接近成熟 成熟应用						
	2020年之前	2020年	2022年	2023年	2024年E	2025年E	2030年E
文本领域	诈骗垃圾信息识别 翻译 基础问答回应	基础文案撰写 初稿	更长的文本 二稿	垂直领域的文案 撰写实现可精调 (论文等)	终稿，水平接近人类平均值	终稿，水平高于人类平均值	终稿，水平高于专业写手
代码领域	单行代码补足	多行代码生成	更长的代码 更精确的表达	支持更多语种 领域更垂直	根据文本生成 初版应用程序	根据文本生成 初版应用程序	根据文本生成 终版应用程序， 比全职开发者 水平更高
图像领域			艺术 图标 摄影	模仿 (产品 设计、建筑等)	终稿 (海报 设计、产品设计 等)	终稿 (产品 设计、建筑等)	终稿，水平高于 专职艺术家、 设计师等
视频/3D/游戏领域				视频和3D文件的 基础版/初稿	根据文本生成 初版的短视频	根据文本生成 初版的长视频， 并实际应用于 制作环节	AI版Roblox 可依个人梦想 定制的游戏与 电影

资料来源：甲子光年《2024AIGC 视频生成：走向 AI 创生时代》

随着 AI 逐步深化与成熟，AI 模型训练、推理等算力需求也将会随之得到进一步释放。2019 年，GPT-2 的模型参数量为 15 亿，只需要使用 8 张训练卡，约 10 天时间就能完成训练。到了 2022 年，GPT-4 的模型参数量增长到了 1750 亿，需要 8 千张卡，约 50 天左右才能完成训练。当前大模型的参数量正在向万亿、十万亿级增长，训练一个大模型将需要更大的算力规模。

图13：AI 大模型需要大算力，训练时长与模型参数量、训练数据量成正比

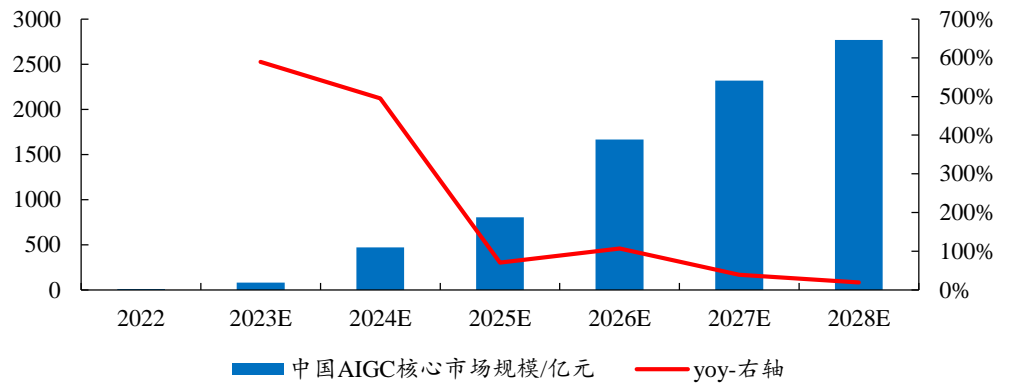
根据业界论文<sup>①</sup>理论推算，端到端训练AI大模型的理论时间为 $E_t = 8 * T * P / (n * X)$ ，其中 $E_t$ 为端到端训练理论时间(s)，T为训练数据的token数量(B Tokens)，P为模型参数量(B)，n为AI硬件卡数(卡)，X为每块卡的有效算力(标\*利用率)。



① (Efficient Large-Scale Language Model Training on GPU Clusters Using Megatron-LM)  
 ② 二次训练，基于预训练模型使用行业数据增量训练(含有标注+无标注)，使模型学得行业领域知识；根据业界相关实践，建议训练数据量>=1000B Tokens  
 ③ 根据业界最佳实践，为保证预训练模型效果，建议数据量>=1000B Tokens，同时根据业界论文理论推算<sup>②</sup>，大模型取得最优训练效果所需的训练数据是其参数的20倍以上(Training Compute-Optimal Large Language Models)

资料来源：华为培训公众号

根据艾媒咨询数据，2022 年中国 AIGC 核心市场规模已达 11.5 亿元，预计 2024 年将增长至 471.7 亿元，AIGC 市场潜力较大。并行科技于 2020 年前后推出 AI 云产品线，面向教育、科研及企业客户提供 GPU 算力服务，重要的大模型客户主要包括智谱华章、中科闻歌、瑞莱智慧、虎博网络、澜舟科技、北京智源人工智能研究院等。2023 年度公司 AI 云实现营业收入 1.1 亿元，同比增长 262.70%，且公司 AI 智算云在手订单情况较好。下游大模型领域的发展有望为公司打开业绩新增量。

**图14：预计 2024 年中国 AIGC 核心市场规模将增长至 471.7 亿元**


数据来源：艾媒咨询、开源证券研究所

### 2.3、中游：“东数西算”等规划颁布，华为云、华南数谷等智算中心投产

超级计算是国家的科技实力的体现，也是国家高精尖技术发展的重要保障。根据《2021-2022 全球算力指数评估报告》显示，国家算力指数与数字经济发展和 GDP 的走势呈现显著的正相关，从 15 个重点国家的算力指数看，算力指数平均每提高 1 个点，数字经济规模和 GDP 将分别增长 3.5%和 1.8%。为此，我国各级政府颁布多项政策，大力支持算力行业发展。2023 年 10 月，工业和信息化部等政府部门颁布《算力基础设施高质量发展行动计划》，提出到 2025 年，全国算力规模应超过 300EFLOPS，智能算力占比应达到 35%。2023 年 12 月，国家发展改革委等部门发布关于深入实施“东数西算”工程加快构建全国一体化算力网的实施意见。2024 年 4 月，江苏省颁布《江苏省算力基础设施发展专项规划》，是全国首个省级算力基础设施发展专项规划。

**表3：国家和各级地方颁布多项政策，支持算力行业发展**

层级	颁布日期	颁布单位	政策名称	主要内容
	2021.11	工业和信 息部	《“十四五”大数 据产业发展规划》	加快构建全国一体化大数据中心体系，推进国家工业互联网大数据中心建设，强化算力统筹智能调度，建设若干国家枢纽节点和大数据中心集群。建设高性能计算集群，合理部署超级计算中心。
国家	2023.10	工业和信 息化部等	《算力基础设施 高质量发展行动 计划》	到 2025 年，算力方面，算力规模超过 300EFLOPS，智能算力占比达到 35%，东西部算力平衡协调发展。运载力方面，国家枢纽节点数据中心集群间基本实现不高于理论时延 1.5 倍的直连网络传输，重点应用场所光传送网（OTN）覆盖率达到 80%，骨干网、城域网全面支持 IPv6，SRv6 等创新技术使用占比达到 40%。存储力方面，存储总量超过 1800EB，先进存储容量占比达到 30%以上，重点行业核心数据、重要数据灾备覆盖率达到 100%。应用赋能方面，打造一批算力新业务、新模式、新业态，工业、金融等领域算力渗透率显著提升，医疗、交通等领域应用实现规模化复制推广，能源、教育等领域应用范围进一步扩大。每个重点领域打造 30 个以上应用标杆。
	2023.12	国家发展 改革委等	关于深入实施“东 数西算”工程加快 构建全国一体化 算力网的实施意 见	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。1ms 时延城市算力网、5ms 时延区域算力网、20ms 时延跨国家枢纽节点算力网在示范区域内初步实现。算力电力双向协同机制初步形成，国家

层级	颁布日期	颁布单位	政策名称	主要内容
				枢纽节点新建数据中心绿电占比超过 80%。用户使用各类算力的易用性明显提高、成本明显降低，国家枢纽节点间网络传输费用大幅降低。算力网关键核心技术基本实现安全可靠，以网络化、普惠化、绿色化为特征的算力网高质量发展格局逐步形成。
	2024.7	国家发展改革委等	《数据中心绿色低碳发展专项行动计划》	到 2025 年底，全国数据中心布局更加合理，整体上架率不低于 60%，平均电能利用效率降至 1.5 以下，可再生能源利用率年均增长 10%，平均单位算力能效和碳效显著提高。到 2030 年底，全国数据中心平均电能利用效率、单位算力能效和碳效达到国际先进水平，可再生能源利用率进一步提升，北方采暖地区新建大型及以上数据中心余热利用率明显提升。
	2023.12	深圳市工业和信息化局	《深圳市算力基础设施高质量发展行动计划（2024-2025）》	到 2025 年，全市数据中心机架规模达 50 万标准机架，算力算效水平显著提高。到 2025 年，通用算力达到 14EFLOPS（FP32），智能算力达到 25EFLOPS（FP16），超算算力达到 2EFLOPS（FP64）。存储总量达到 90EB。先进存储容量占比达到 30% 以上，重点行业核心数据、重要数据灾备覆盖率达到 100%。市内数据中心间时延不高于 1ms，至韶关国家枢纽节点时延不高于 3ms，至贵安国家枢纽节点时延不高于 10ms。到 2025 年，新建数据中心电能利用效率（PUE）降低到 1.25 以下，绿色低碳等级达到 4A 级以上。启动“老旧小散”数据中心升级改造。强化网络、数据、算力设施安全管理和能力建设，构建完善的安全保障体系。
	2024.2	武汉市人民政府办公厅	《武汉市推进算力基础设施及应用产业高质量发展行动方案（2024-2025 年）》	到 2025 年，全市建成布局合理、服务高效、技术先进、绿色安全的算力基础设施，形成算力、存力、运力及应用产业与数字经济高质量发展相适应的发展格局，打造中部算力高地。
地方	2024.4	江苏省通信管理局等	《江苏省算力基础设施发展专项规划》	江苏将全力构建一个“2+N+X”算力网络梯次多元布局，即建成南京、苏州 2 个国家级核心算力枢纽集群，N 个省内城市级数据中心（云计算中心、智算中心）和 X 个边缘计算节点。到 2030 年，全省数据中心机架规模达 120 万标准机架，全省在用总算力超过 50EFLOPS，智能算力占比超过 45%，智算中心突破 20 个，算力算效水平显著提高；存储总量超过 500EB，推进全闪存、硬件高密等先进存储创新技术规模应用，将先进存储容量占比提升至 40% 以上，重点行业核心数据、重要数据灾备覆盖率达到 100%；建成省级算力调度平台，覆盖纳入国家枢纽节点数据中心集群范围的重点项目算力，实现算力资源的灵活调度，算力资源市场化高效调配；算力、算法、数据、应用资源实现集约化和服务化创新发展，多要素融合创新孵化至少 50 项创新应用，打造不少于 80 个算力产业应用推广案例、智算典型示范场景 35 个。
	2024.4	北京市经济和信息化局等	《北京市算力基础设施建设实施方案（2024-2027 年）》	到 2025 年，基本建成智算资源供给集群化、智算设施建设自主化、智算能力赋能精准化、智算中心运营绿色化、智算生态发展体系化的格局，引领京津冀蒙地区建成具有国际影响力的智算产业创新应用高地。到 2027 年，优化京津冀蒙算力供给质量和规模，力争自主可控算力满足大模型训练需求，算力能耗标准达到国内领先水平，建立布局合理、算网协同、技术领先、绿色集约、产业链完备的京津冀蒙算力一体化协同发展格局，融入联网调度、普惠易用、绿色安全的全国一体化算力体系，有力支撑落实国家人工智能产业发展战略部署。
	2024.7	天津市人民政府办公厅	《天津市算力产业发展实施方案（2024—2026 年）》	力争到 2026 年，全市智能算力规模达到 10EFLOPS 以上，全市超级算力应用率达到 70% 以上，全市既有大型及以上数据中心能源利用效率（PUE）值力争降至 1.4 以下，

资料来源：各级政府官网、开源证券研究所

为响应国家号召，国内 AI 算力企业算力建设持续加码，在全国多地兴建智算中心。中国移动落地“4+N+31+X”算力网络布局，统筹规划“N+X”智算中心布局体系，在京津冀、长三角、粤港澳大湾区、成渝、内蒙古、宁夏、贵州、黑龙江、山东、湖北等地打造了首批 12 个智算中心，整体算力规模达 11EFLOPS。2024 年 6 月，华为云华东（芜湖）数据中心正式开服，标志着华为云全国存算网的枢纽节点布局全面完成；安徽芜湖、贵州贵安、内蒙古乌兰察布三大云核心枢纽，将构建“算力一张网”，服务全国算力需求。截至 2024 年 6 月，华为云已覆盖全球 33 个地理区域、93 个可用区，打造了一张性能领先、运行稳定、安全可靠的全球存算网，为 170 多个国家和地区的客户提供服务。截至 2023 年底，我国提供算力服务的在用机架数达到 810 万标准机架，各类算力提供主体超 5000 家，算力总规模位居全球第二。

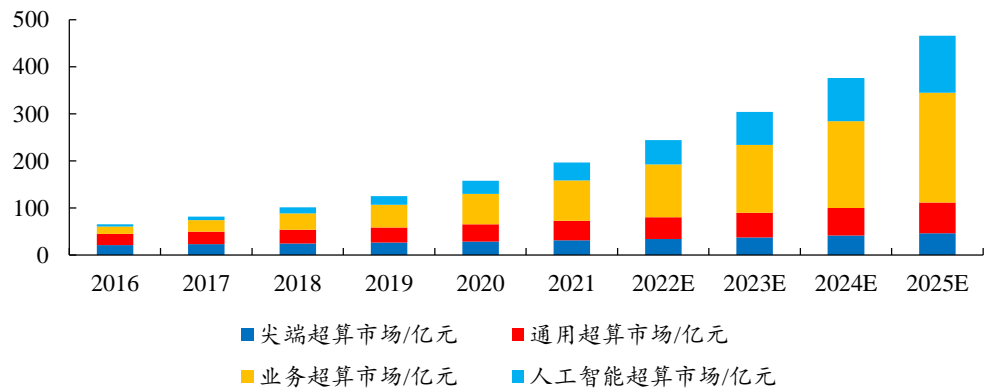
**表4：浪潮集团、阿里等国内 AI 算力企业在全国多地兴建智算中心。**

企业	时间	投资动向
浪潮集团	2022.07	浪潮与济南市政府签约，建设一体化大数据中心，提升济南算力服务供给能力；建设中国算谷科技园，为中国算谷提供科技支撑，推动新一代信息技术产业生态发展；浪潮与京东科技集团合作，通过开发“天枢”服务器，助力京东云建设第一代绿色数据中心；浪潮与济南市能源集团签约，围绕社会治理、智慧灯杆、城市大脑、智慧社区等领域，建设算网融合数据中心，支撑济南市智慧城市运营。
阿里	2022.08	阿里云宣布正式启动张北超级智算中心。该智算中心总建设规模为 12EFLOPS（每秒 1200 亿亿次浮点运算）AI 算力，将超过谷歌的 9EFLOPS 和特斯拉的 1.8EFLOPS，成为全球最大的智算中心，可为 AI 大模型训练、自动驾驶、空间地理等人工智能探索应用提供强大的智能算力服务。
百度	2023.06	百度智算中心项目落地四川雅安，响应国家东数西算政策，雅安百度智算中心作为国内中西部地区首个集智能算力与大模型技术平台一体化的新型人工智能算力基础设施，不仅包含物理数据中心建设能力、智算基础设施-百度 AI 大底座能力，还包括以大模型为代表的 AI 软硬件生态能力的综合性解决方案，可支持智算时代下企业不同业务场景的计算任务和大模型的产业应用需求。
中国联通	2023.11	中国联通长三角（芜湖）智算中心项目开工。该项目总投资 60 亿元，终期规划总建筑面积约 8.77 万平方米，具备等效 1 万架功率 6kW 机柜能力，由 4 栋数据中心、2 栋运维楼、1 栋 110kV 变电站、2 个油机平台及其他配套设施组成。
云赛智联	2024.02	仪电智算中心（松江）二期入选上海市 2024 年重大工程清单，预计将于 2024 年 10 月竣工交付。
博大数据	2024.03	博大数据深圳前海智算中心开园。该项目基础设施投资 20 亿元，总面积超过 5 万平方米，整体规划容量约为 15000 个标准机柜，并支持高功率深度定制，整体按照国家 A 级标准建设，一期可支持算力规模达 40,000P，是深圳市在建规模、投入规模首屈一指的高等级智算中心。
燧原科技	2024.03	燧原智算中心项目签约落地无锡高新区。根据燧原科技与无锡高新区合作构建的新一代人工智能算力平台，这座智算中心采用异构算力架构，其中英伟达高端算力 H800/H100 占比约 80%，燧原国产 AI 算力占比约 20%，总算力不低于 1000P（FP16）的 AI 算力。
鹰硕集团	2024.03	华南数谷智算中心投产。中心由中国电子集团和鹰硕集团共同投资建设，总投资 23.6 亿元，按照国家 A 级机房标准规划建设，总规模约 3 万个标准机柜（7000 个算力机柜），是粤港澳大湾区枢纽节点首批投入运营的智算中心。
中国移动	2024.04	中国移动呼和浩特智算中心正式投产，该智算中心是全球运营商最大单体液冷智算中心，入选“2023 年度央企十大超级工程”，部署约 2 万张 AI 加速卡，AI 芯片国产化率超 85%，智能算力规模高达 6.7EFLOPS（每秒 670 亿亿次浮点运算）。
华为	2024.06	华为云华东（芜湖）数据中心正式开服，成为华为云三大数据中心之一，标志着华为云全国存算网枢纽节点布局全面完成。安徽芜湖、贵州贵安及内蒙古乌兰察布三大核心枢纽是华为最高规格、最大规模、最新技术的数据中心，通过全国上万公里的光纤骨干网，及百 Tb 级带宽互联，华为云算力枢纽可面向全国提供算力。
中贝通信	2024.07	“中贝合肥智算中心”在合肥蜀山经济技术开发区正式上线运行。该项目是安徽首家“低碳 AI 智算中心”，可满足华东、华南区域 AI 企业“大模型”训练和“大数据”推理的各项需求。

资料来源：各公司官网、新华网、各级政府官网、前瞻产业研究院等、开源证券研究所

在国家政策、市场需求等多重因素叠加下，我国超算服务市场持续扩容。据沙利文数据，2021年超算服务市场规模约为196.6亿元，并预计在2025年达到466亿元，2021年-2025年CAGR为24.1%。其中，增长主要由业务超算与人工智能超算带动。2021年业务超算市场规模为85.6亿元，并预计在2025年达到233.2亿元，2021年-2025年CAGR为28.5%；2021年人工智能超算市场规模为38.3亿元，并预计在2025年达到121.1亿元，2021年-2025年CAGR为33.3%。

**图15：2021年超算服务市场规模约为196.6亿元，预计在2025年达到466亿元**



数据来源：沙利文、开源证券研究所

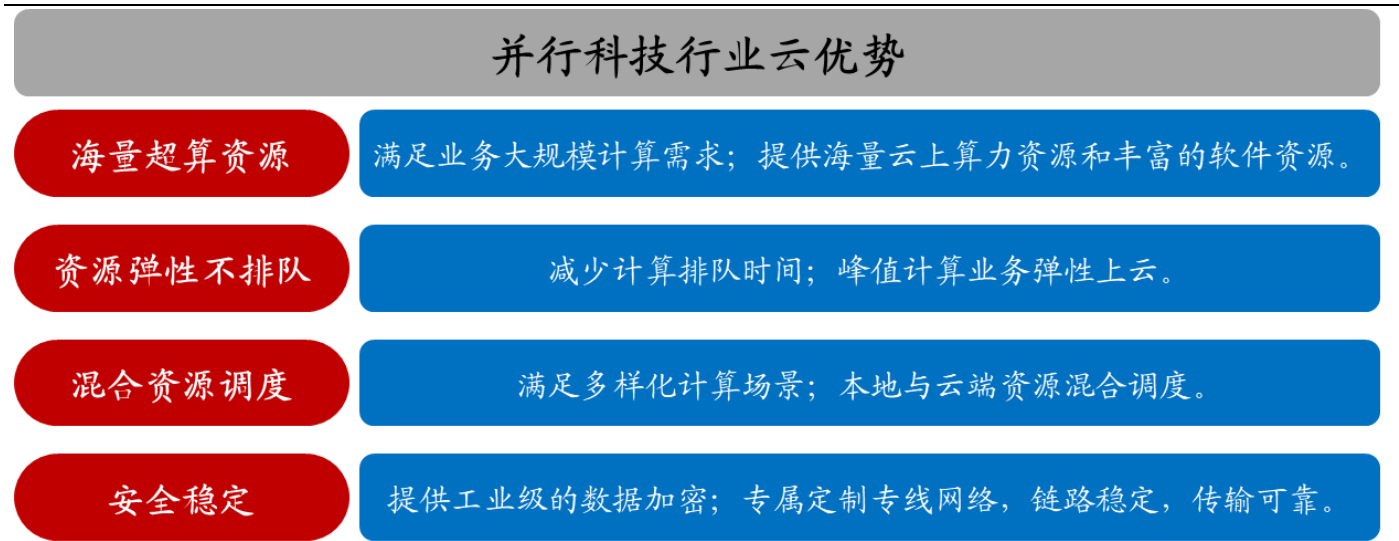
### 3、行业云+AI云双引擎持续驱动，绘制第二增长曲线

#### 3.1、并行行业云近三年CAGR 138%，解决气象海洋等多个行业痛点

并行行业云在并行通用超算云的基础上增加了业务系统对接、提供专线链路等物理层、系统层、应用层的专业化服务等以满足用户需求，提供整体解决方案。既保障用户业务流程快速、高效、动态实现，又能满足企业混合部署、自主可控、安全稳定、易运维的需求。

总体而言，公司的行业云产品具有大量的超算资源，并随着公司“算海计划”的持续推进，公司的算力资源规模将进一步扩大。并行科技的行业云还具有混合资源调度模式，以及良好的资源弹性，有效提升客户应用的运行效率。除此自外，在安全性方面，公司为用户提供工业级的数据加密，保证链路稳定、传输可靠。



**图16：并行科技行业云在超算资源、资源弹性、资源调度程度、安全性方面具有优势**


资料来源：公司官网、开源证券研究所

面向在超算云上构建业务系统的行业客户，并行行业云服务为客户提供硬件资源、专线连接、灵活的计费模式、以及混合部署方式，以满足用于不同行业的不同业务需求。

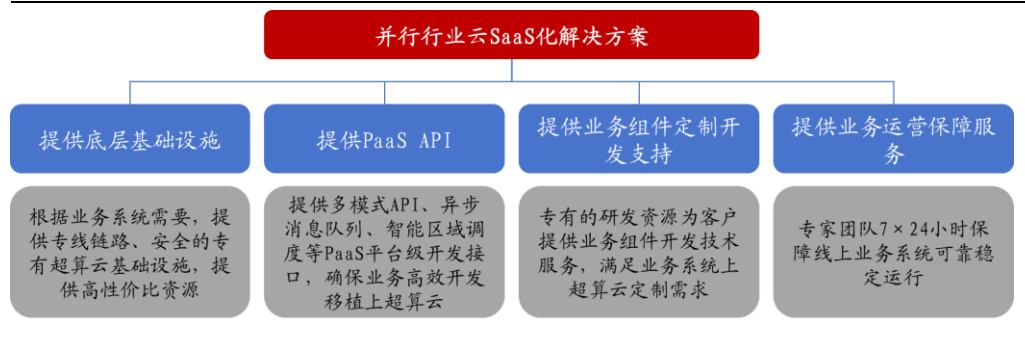
**表5：并行行业云为客户提供硬件资源、专线连接、灵活的计费模式等服务**

资源服务	内容
硬件资源	提供多种硬件配置，采用 Intel Xeon 或 AMD EPYC 的 CPU，配置全互联的 Infiniband 高速网络，满足不同应用软件对硬件的特殊要求等。
专线连接	提供专线、VPN 等多种接入方式，通过多链路互备保证传输稳定可靠，业务高峰期还可弹性扩容，满足临时大量数据传输。
灵活计费	提供包年、包月、按需的混合计费模式，在满足基本的固定计算资源外，还能周期性业务和临时性需求提供短期计算能力，降低整体 TCO。
混合部署	可与自建集群整合，通过 Paracloud 企业版对接本地集群与超算云，实现混合云部署；云端不存结果直接回传本地存储，确保数据安全。

资料来源：公司官网、开源证券研究所

并行行业云积极为用户提供 SaaS 化解决方案，对于有兴趣在云端无缝交付软件的 ISV，帮助用户实现软件上云，使用户能够在云端运行软件，并通过订阅模式产生新的收入。并行行业云可根据业务系统需求，为用户提供专线链路、安全的专有超算云基础设施，提供高性价比资源；还可以提供多模式 API、异步消息队列、智能区域调度等 PaaS 平台级开发接口，快速满足文件管理、数据传输、作业调度和作业监控等业务需求，确保业务高效开发移植上超算云；公司可在业务开发、移植上超算云的过程中，为客户提供研发级支持；公司专有的研发资源亦为客户提供业务组件开发技术服务，满足业务系统上超算云的定制需求。

图17: 行业云 SaaS 化解决方案为用户提供底层基础设施、PaaS API 等服务



资料来源: 公司官网、开源证券研究所

并行行业云可针对不同行业客户痛点定制专门解决方案, 如气象海洋超算云解决方案、生命科学行业云解决方案、高端制造行业云解决方案、企业级 HPC 超算云解决方案、院校级 HPC 超算云解决方案等。行业云为不同行业提供定制化超算中心, 满足行业业务对于超算资源的需求。

表6: 并行行业云针对气象海洋、生命科学等不同行业制定解决方案

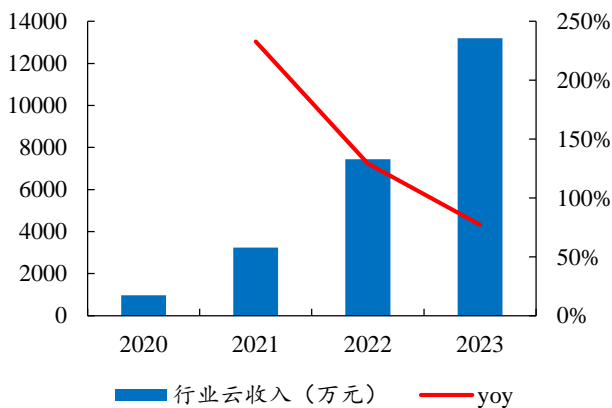
行业	痛点	解决方案
气象海洋行业	数值模式软件运行数据来源渠道多且使用过程中数据流转多; 数值模式软件种类多且种类复杂; 数值模式软件运行过程中对各类资源(尤其是 CPU 资源、网络资源)要求高; 数值模式软件计算结果输出量大; 运行时对时效性要求高等	公司气象海洋超算云解决方案为客户提供面向全球的数据下载服务, 并进行工业级的数据加密, 链路稳定, 传输可靠。气象海洋超算云解决方案提供优质的 CPU 性能、充足的内存容量、快速的网络带宽。同时, 可支持业务前后处理接入, 提供业务全流程的开发和服务支持
生命科学行业	用户需短时间调配海量算力资源, 本地服务器很难满足项目进度要求; 软件环境复杂, 部署和安装调试占用大量时间; 需要多种软件资源以及组件, 自建需要长时间的调研以及大量的资金投入; 多资源配置形态、多业务场景模式、场景间交互以及流程化需求较为普遍等	可以按需调配纯物理/虚拟算力资源, 提供多云、多资源、混合管理平台、作业运行诊断分析、云原生组件等各类产品及服务, 方便客户便捷、高效使用超算云资源, 提高其整体平台架构以及方案规划能力, 帮助客户计算业务迅速上线
高端制造业	高端制造业领域, 无论是汽车、精密仪器还是集成电路, 超算能够加速其产品设计与优化周期, 避免潜在的产品缺陷并提高业务运行整体效率。自建超算中心前期资金投入大, 日常的设备维护和升级改造产生的管理成本高昂, 硬件设备及软件运维难度大, 技术维护门槛较高, 在整体业务链条中经济适配性较差	可提供稳定、全面、配套的一栈式超算服务, 具备成熟的超算服务能力, 能在满足高端制造业对计算性能的高标准要求的同 时, 实现算力资源的高效调度及软硬件系统的灵活搭配以满足客户的精细化需求。公司完善的服务矩阵以及灵活付费模式可为行业用户提供轻资产化的超算服务, 并实现业务运营效率的提升及业务成本的优化
企业级 HPC 超算云	企业一般随业务增长逐年规划建设超算集群, 由于从需求收集到项目落地存在采购周期长且项目实施与交付较为复杂的情形, 规划困难且建设维护成本高; 在业务高峰期时, 企业研发人员排队严重, 而闲时资源利用率较低, 企业难以找到规模平衡点; 超算中心建设初期投入大, 成本高, 企业将面临较大的资金压力	案结合本地共建集群、专线网络、云端超算、调度软件、专业服务及行业经验等优势, 打包输出混合云整体解决方案, 云端资源可大幅缩短交付时间。云端超算可以按月/年固定租赁或按需弹性使用进行交付, 使用灵活, 初期少量投入即可享受充足算力资源, 大幅降低企业用户资金压力。企业级 HPC 超算云解决方案可本地实现前后处理与用户数据保存, 云端只计算不储存数据, 在提升用户体验的同时保证数据安全。另外, 企业级 HPC 超算云解决方案可对接用户认证与研发管理系统, 实现一体化管理与维护
院校级 HPC 超算云	大部分高校的超算中心或网络中心缺少专业的高性能计算技术人员, 课题老师专注科研领域, 对高性能计算机知识较为生疏。超算中心亦缺乏有效管理手段,	基于并行超算云服务, 为高校、科研院所构建云上科研环境, 将用户科学计算任务弹性扩展到云端, 有效解决本地超算中心算力资源不足等问题, 节约了 HPC 集群构建成本, 计算无需排队, 提升科研效率;

行业	痛点	解决方案
	用户体验和资源使用率较低	为老师及科研人员提供海量硬件和丰富软件资源，满足紧急大量的科研计算需求；提供计量计费系统，通过本地或云上统一管理，并为其提供 7×24 小时响应服务，解决课题老师在使用过程中的各种问题

资料来源：公司招股书、开源证券研究所

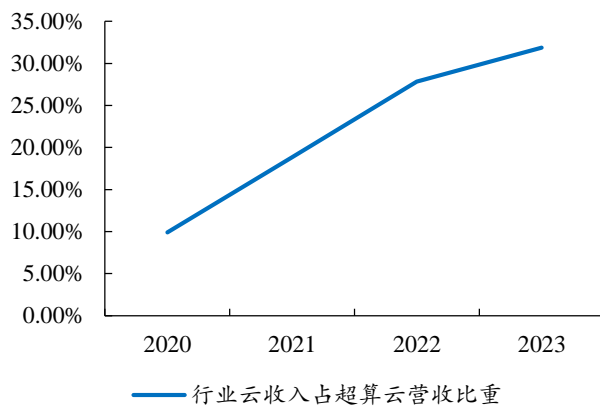
收入维度而言，并行科技行业云业务收入高速增长，2020年-2023年CAGR为138.32%。虽然行业云业务收入放缓，但仍处于迅速增长阶段。2023年，并行行业云业务收入达1.32亿元，同比上升了77.44%。2020-2022年，行业云收入分别为975.3万元、3247.11万元、7442.52万元。2020年来，行业云业务收入占超算云业务收入的比例持续攀升。2020年，行业云业务收入占比仅为9.93%，至2023年，该比例已达到31.89%。

图18：2023年行业云收入为1.32亿元（+77.44%）



数据来源：Wind、开源证券研究所

图19：2023年行业云收入占超算云收入比例为31.89%



数据来源：Wind、开源证券研究所

### 3.2、并行 AI 云近三年 CAGR 253%，服务人工智能、高性能计算领域

并行 AI 云面向高校、科研院所、行业用户等在 AI 深度学习和科学计算等方向的 GPU 算力资源需求，提供高性能 GPU 算力资源及相关 PaaS 服务。并行 AI 云可应用到人工智能及高性能计算等场景，可灵活地满足相关人工智能技术研究在训练和推理阶段对 GPU 算力资源复杂多样需求。并行 AI 超算云裸金属集群为 GPU 高性能计算场景锻造，提供最新架构 GPU 卡，以裸金属形态输出，消除虚拟化性能损耗；提供高速互联、高性能并行存储，支持大规模并行；基于大规模集群架构提供丰富的 GPU 队列资源池，可根据业务需要便捷弹性获取大规模资源，满足大规模计算需求。

图20：并行 AI 云以 GPU 为核心资源池



资料来源：公司招股书、开源证券研究所

从公司 AI 云产品提供的服务角度而言，并行 AI 云具有高性价比、资源丰富、易用高效、服务贴心的特点及优势。AI 云的应用场景通常包括人工智能与高新能计算领域。在人工智能领域，AI 云主机可灵活地满足相关人工智能技术研究，在训练和推理阶段对 GPU 算力复杂多样需求。在高性能计算领域，AI 云的 GPU 集群可以支持高性能计算场景的应用。

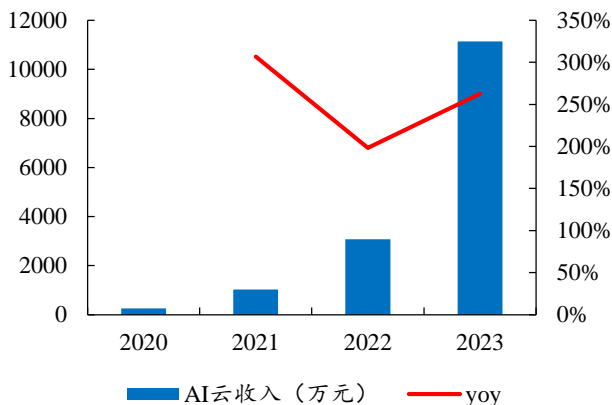
表7：并行 AI 云具有高性价比、资源丰富、易用高效、服务贴心等优点

产能优势	内容
高性价比	建设与运维“0”成本，按需灵活选用海星资源，不操心管，只专注用。
资源丰富	提供云主机平台和裸金属云服务等多种平台，算力资源包括 A100\V100\3090\A30\A10\T4\国产 DCU 等多种型号，支持多机多卡调度，满足训练、推理等不同场景需求。
易用高效	预装 AI 框架、内置数据集，可通过简单直观的界面快速创建 AI 计算实例，裸金属服务器、独占显卡，性能强劲。
服务贴心	专家团队 7×24 小时在线服务，提供开发框架部署、算例编译、调试优化等多元化服务支持，为 AI 客户深度赋能。

资料来源：公司官网、开源证券研究所

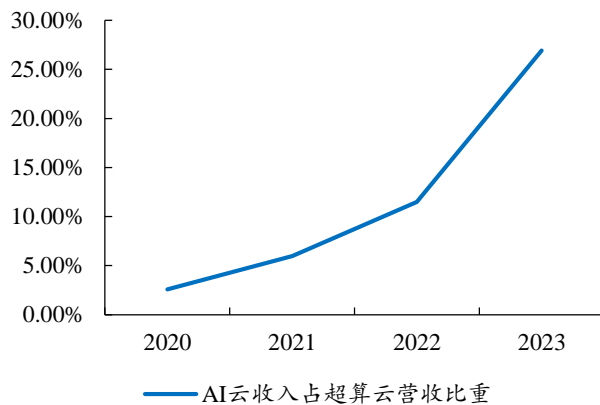
聚焦于并行 AI 云业务收入角度，2023 年 AI 云收入达到 1.11 亿元，相较于 2022 年涨幅高达 262.70%，2020 年至 2023 年，CAGR 为 253.09%。2020-2022 年，公司 AI 云业务收入分别为 253.11 万元、1029.81 万元、3071.95 万元。2020 年来，AI 云业务收入占算云业务收入的比例分别为 2.58%、5.98%、11.50%、26.91%，收入占比持续升高。

图21：2023年AI云业务收入为1.11亿元(+262.70%)



数据来源：Wind、开源证券研究所

图22：2023年AI云业务收入占超算云营收26.91%



数据来源：Wind、开源证券研究所

#### 4、运营效率+算力利用率双提升，2024H1 业绩持续修复已扭亏为盈

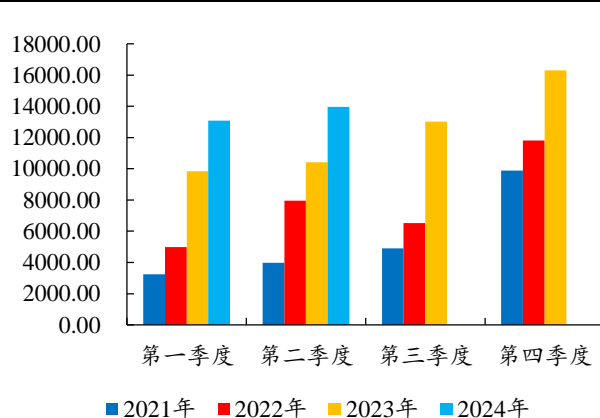
从营收情况看，公司的营业收入呈现一定季节性波动，第一季度收入占比相对偏低，第四季度占全年收入比例较高，各季度收入总体呈上升趋势。但总体看，2021-2023年全年营收分别为2.2亿元、3.1亿元和5.0亿元，CAGR为50%，公司营收增速较快。2021Q3-2024Q2，各季度营收均保持同比正增长，2024Q2公司营收达1.4亿元，同比增长34%。

图23：2021Q3-2024Q2 营收呈现上升趋势



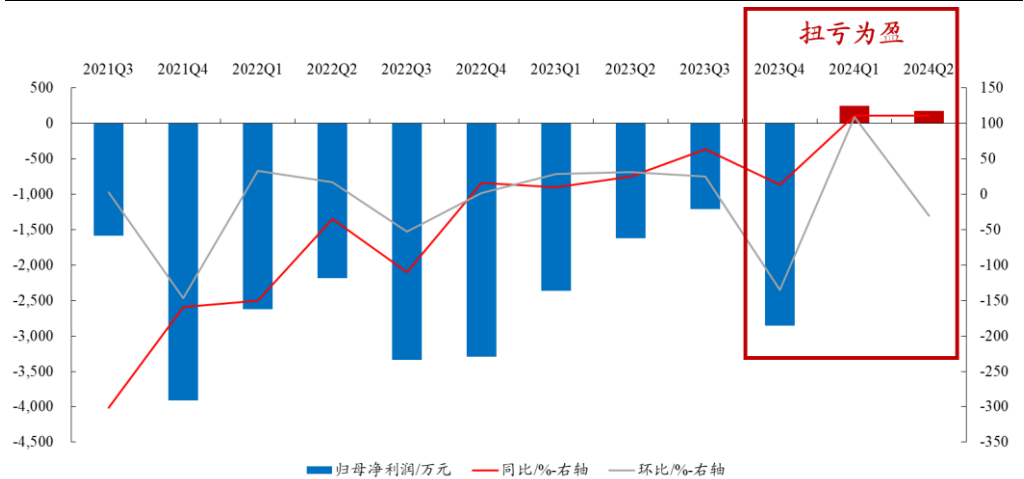
数据来源：Wind、开源证券研究所

图24：2021-2024Q2 营收保持同比增长 单位：万元



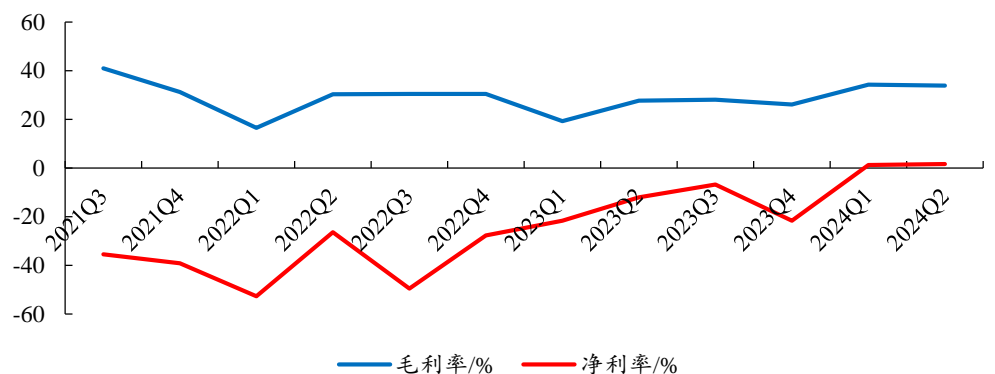
数据来源：Wind、开源证券研究所

从利润情况看，2021Q3-2023Q4，公司归母净利润均为负值，但整体来看亏损规模有所收缩。2024年第一季度，公司实现扭亏为盈，归母净利润达250万元。2024年第二季度业绩持续修复，根据公司2024半年报数据，2024H1归属于上市公司股东的净利润达到423万元。

**图25：2024年Q2归母净利润达173万元，同比上涨111%**


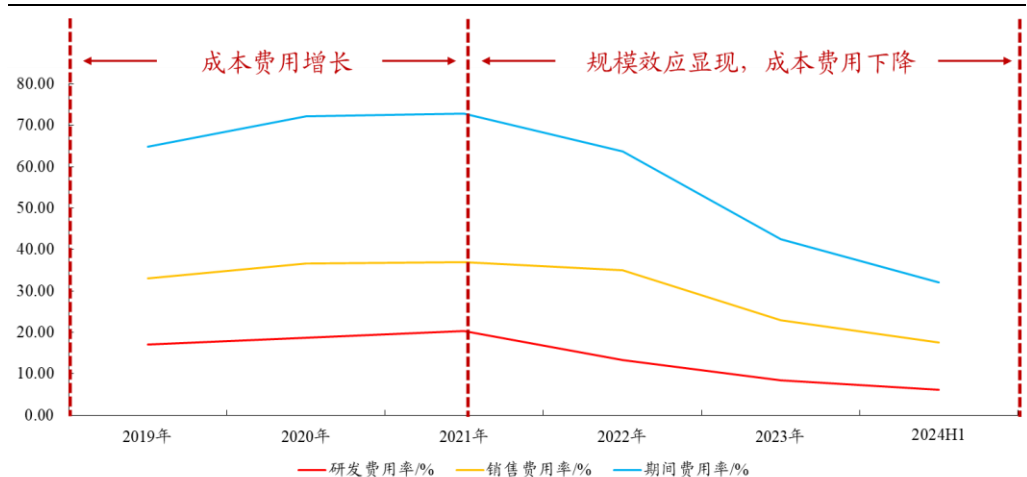
数据来源：Wind、开源证券研究所

从毛利率、净利率来看，公司2020-2023年毛利率分别为42.99%、34.02%、28.26%和25.63%，受固定资产折旧影响，毛利率下滑较为明显，且净利率为负。2024Q1受公司算力服务规模效应逐渐显现、运营效率持续提升、算力资源利用率提升等多重积极因素叠加影响，毛利率回升至34.27%，净利率由负转正。2024Q2毛利率较Q1无明显变化，净利率小幅上升至1.62%。

**图26：2024Q2毛利率为33.86%，Q2净利率在Q1由负转正后小幅上升**


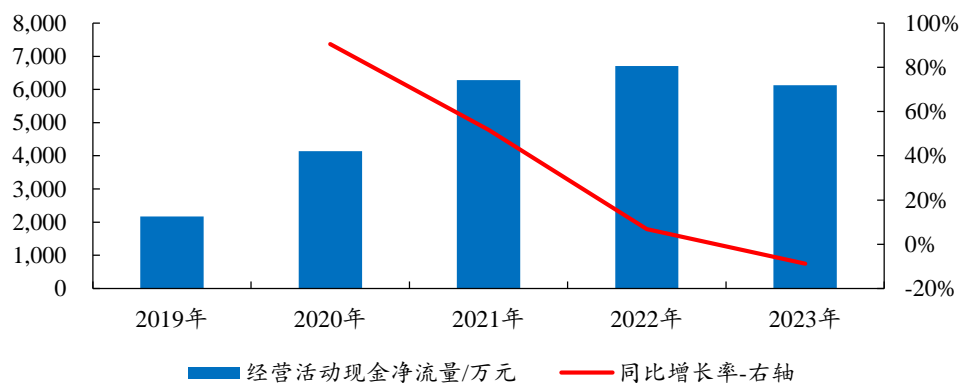
数据来源：Wind、开源证券研究所

从成本费用来看，2021年之前，随着公司业务规模的拓展，公司成本费用逐年提高，2021年达到最高点，期间费用率为72.86%，研发费用率为20.31%，销售费用率为37.02%。2021年之后，公司规模效应逐步显现，成本费用持续降低，2024上半年，期间费用率为32.14%，研发费用率为6.09%，销售费用率为17.62%，三者较2021年缩减一半以上。

**图27：2024年H1期间费用率为32.14%，较2021年缩减一半以上**


数据来源：Wind、开源证券研究所

从经营活动现金流来看，2019-2021年，随着公司规模扩大和业务拓展，经营活动现金净流量逐年提升，CAGR为70%。2021-2023年，公司经营活动现金净流量规模相对稳定，2023年公司经营活动现金净流量为6,125.37万元。

**图28：2019-2021年公司经营活动现金净流量逐年提升，2021-2023年保持稳定**


数据来源：Wind、开源证券研究所

## 5、盈利预测与投资建议

选取**云赛智联**作为AI算力云服务商领域的可比公司，**恒为科技**作为算网服务及第三方AI算力运维领域的可比公司，**奥飞数据**作为为客户提供互联网数据中心软件服务以及相关的解决方案，**深信服**作为AI算力云服务、行业云等领域的可比公司。

并行科技是国内领先的超算云服务和算力运营服务提供商，主要服务包括**并行通用超算云**、**并行行业云**、**并行AI云**等。并行科技合作内蒙古算力基地于2024年5月15日正式上线运营，并向算力建设合作伙伴发起号召：共同筹建面向模型训练的大规模单体集群。通过多年积累，公司具备分布式超算集群、算力资源网络、AI算力调优等领域的领先技术，参与多项国家重点研发计划。近年来，公司业绩持续修复，2021Q3-2024Q2，各季度营收均保持同比正增长，2024Q1公司实现扭亏为盈，2024H1归母净利润达423万元，2024H1毛利率回升至34.06%。我们上调2024-2026

年盈利预测, 预计 2024-2026 年归母净利润为-0.15/0.15/0.54(原:-0.37/0.07/0.45)亿元, 对应 EPS 分别为-0.26/0.25/0.92 元/股, 对应当前股价的 PS 分别为 3.0/2.3/1.8 倍, 可比公司 PS 2024E 均值为 3.6X, 看好公司 AI 算力云及行业云释放增量, 上调“买入”评级。

**表8: 可比公司 PS 2024E 均值为 3.6X**

公司名称	代码	最新收盘价 (元/股)	最新总市值 (亿元)	PS			PE		
				2024E	2025E	2026E	2024E	2025E	2026E
云赛智联	600602.SH	9.46	111.29	1.8	1.6	1.4	55.6	47.3	39.4
恒为科技	603496.SH	20.48	65.58	5.1	3.8	3.0	30.6	22.0	16.7
奥飞数据	300738.SZ	8.79	84.81	4.9	3.8	3.1	41.9	28.4	21.4
深信服	300454.SZ	46.83	196.61	2.4	2.1	1.9	64.2	38.1	29.1
<b>均值</b>			<b>114.57</b>	<b>3.6</b>	<b>2.8</b>	<b>2.3</b>	<b>48.1</b>	<b>33.9</b>	<b>26.6</b>
并行科技	839493.BJ	36.20	21.08	3.0	2.3	1.8	-140.4	143.1	39.2

数据来源: Wind、开源证券研究所 注: 数据截至 2024.8.26; 并行科技和深信服盈利预测来自开源证券研究所, 其他可比公司盈利预测数据来自 Wind 一致预期

## 6、风险提示

行业需求不及预期、新业务开拓受阻、客户合作风险



**附：财务预测摘要**

资产负债表(百万元)	2022A	2023A	2024E	2025E	2026E
<b>流动资产</b>	323	611	497	619	852
现金	154	410	248	325	493
应收票据及应收账款	41	46	57	69	84
其他应收款	2	5	5	7	8
预付账款	14	3	28	14	35
存货	75	87	97	144	168
其他流动资产	37	60	61	61	62
<b>非流动资产</b>	423	523	667	617	473
长期投资	0	0	0	0	0
固定资产	256	313	374	393	327
无形资产	3	7	8	8	8
其他非流动资产	163	203	285	216	138
<b>资产总计</b>	746	1135	1164	1236	1325
<b>流动负债</b>	569	759	789	852	888
短期借款	27	183	183	183	183
应付票据及应付账款	88	88	115	126	160
其他流动负债	453	488	491	542	545
<b>非流动负债</b>	49	29	43	38	38
长期借款	3	0	0	0	0
其他非流动负债	46	29	43	38	38
<b>负债合计</b>	618	788	831	890	926
少数股东权益	6	10	11	10	8
股本	47	58	58	58	58
资本公积	469	753	753	753	753
留存收益	-394	-475	-489	-475	-423
<b>归属母公司股东权益</b>	122	337	322	336	390
<b>负债和股东权益</b>	746	1135	1164	1236	1325

现金流量表(百万元)	2022A	2023A	2024E	2025E	2026E
<b>经营活动现金流</b>	67	61	86	149	161
净利润	-112	-78	-15	14	52
折旧摊销	66	85	82	110	130
财务费用	8	7	9	9	6
投资损失	0	0	0	0	0
营运资金变动	67	13	8	14	-30
其他经营现金流	37	34	1	2	2
<b>投资活动现金流</b>	-131	-212	-226	-60	14
资本支出	132	213	139	76	3
长期投资	0	0	0	0	0
其他投资现金流	1	2	-87	16	17
<b>筹资活动现金流</b>	-47	428	-21	-13	-6
短期借款	-5	156	0	0	0
长期借款	3	-3	0	0	0
普通股增加	0	12	0	0	0
资本公积增加	6	284	0	0	0
其他筹资现金流	-51	-20	-21	-13	-6
<b>现金净增加额</b>	-110	278	-162	76	169

利润表(百万元)	2022A	2023A	2024E	2025E	2026E
<b>营业收入</b>	313	496	701	918	1184
营业成本	224	369	508	664	859
营业税金及附加	1	1	2	3	3
营业费用	109	114	105	119	130
管理费用	41	48	46	47	53
研发费用	42	42	50	64	77
财务费用	8	7	9	9	6
资产减值损失	-0	0	0	0	0
其他收益	2	8	6	5	4
公允价值变动收益	0	0	0	0	0
投资净收益	0	0	0	0	0
资产处置收益	0	1	0	0	0
<b>营业利润</b>	-112	-77	-14	15	58
营业外收入	0	0	0	0	0
营业外支出	0	0	0	0	0
<b>利润总额</b>	-112	-78	-15	15	58
所得税	-1	-0	-0	1	6
<b>净利润</b>	-112	-78	-15	14	52
少数股东损益	3	3	0	-0	-2
<b>归属母公司净利润</b>	-114	-80	-15	15	54
EBITDA	-47	9	71	129	191
EPS(元)	-1.96	-1.38	-0.26	0.25	0.92

主要财务比率	2022A	2023A	2024E	2025E	2026E
<b>成长能力</b>					
营业收入(%)	42.1	58.5	41.3	31.0	29.1
营业利润(%)	-42.4	31.1	81.6	207.5	279.5
归属于母公司净利润(%)	-39.8	29.5	81.4	198.1	264.9
<b>获利能力</b>					
毛利率(%)	28.3	25.6	27.5	27.7	27.5
净利率(%)	-36.5	-16.2	-2.1	1.6	4.5
ROE(%)	-87.2	-22.4	-4.4	4.1	13.1
ROIC(%)	-60.4	-13.8	-2.2	3.5	9.4
<b>偿债能力</b>					
资产负债率(%)	82.8	69.4	71.5	72.0	69.9
净负债比率(%)	-69.5	-56.7	-18.4	-39.8	-76.8
流动比率	0.6	0.8	0.6	0.7	1.0
速动比率	0.3	0.6	0.4	0.5	0.7
<b>营运能力</b>					
总资产周转率	0.4	0.5	0.6	0.8	0.9
应收账款周转率	10.7	11.7	14.0	15.0	16.0
应付账款周转率	4.4	4.8	5.0	5.5	6.0
<b>每股指标(元)</b>					
每股收益(最新摊薄)	-1.96	-1.38	-0.26	0.25	0.92
每股经营现金流(最新摊薄)	1.15	1.05	1.48	2.57	2.76
每股净资产(最新摊薄)	2.09	5.78	5.52	5.77	6.70
<b>估值比率</b>					
P/E	-18.5	-26.2	-140.4	143.1	39.2
P/B	17.3	6.3	6.6	6.3	5.4
EV/EBITDA	-43.0	203.0	29.1	15.3	9.5

数据来源：聚源、开源证券研究所

### 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