

通信行业深度报告

云厂商仍保持算力高需求，算力租赁需求有望边际向上

增持（首次）

投资要点

- **算力租赁与 IDC 在算力生态中互补共存，两者的核心区别可分为资源控制权、服务内容与成本结构三方面：**算力租赁是“购买服务”模式，用户无需自购硬件，即可通过租赁 GPU 集群等计算资源获得算力，所有硬件和技术迭代由服务商负责，按需付费模式也更具灵活性，适合短期弹性需求与技术更新较快的场景；IDC 则属于“自建或租赁基础设施”模式，用户需持有或自行租用机柜、电力、网络等资源，并自行管理服务器、存储设备及应用部署，前期投入较高但更利于长期稳定业务，亦适合需要高度数据控制和安全保障的企业。总体而言，算力租赁强调“算力即服务”的轻量化输出，IDC 则提供“重资产持有+基础设施服务”，两种模式在不同需求场景下发挥优势，形成互补的算力服务体系。
- **云厂商逐步向 DeepSeek 的效率与成本对标，训练仍有高需求。**以 DeepSeek 为代表的顶尖 AI 公司，通过算法-硬件协同设计，如稀疏化训练、混合精度优化及分布式训练框架升级（如 3D 并行策略），云厂商对算力租赁具有明确需求，其核心逻辑在于灵活应对算力需求波动、降低硬件投入成本、优化资源利用率。Agent（智能体）对算力的需求提升主要源于其渗透率、日均调用次数及任务复杂度的三重增长。例如企业端因 AI 应用扩展（如智能制造、金融分析）开始自建模型训练，进一步加剧算力压力——例如 DeepSeek 单日推理调用 2500 万次，需 1814 张 H800 芯片支持。在这种双重驱动趋势下，算力需求从训练端开始向推理端迁移。
- **供给端硬约束下，因稀缺性溢价与成本传导机制，算力租赁仍有涨价空间。**美国对华芯片出口限制持续加码，英伟达 H100/H20 等高端 GPU 对华供应被严格限制。英伟达 H20 禁售后，H800 服务器单价持续逐步涨价，硬件端的成本压力已经传导到租赁端。根据《智能计算中心创新发展指南》，2022 年我国智能算力规模快速增长，达到 268 百亿亿次每秒 (EFLOPS)，超过通用算力规模，预计未来 5 年中国智能算力规模的年复合增长率将达 52.3%。
- **IDC 行业的竞争本质是资源禀赋争夺战，算力租赁的核心在于 GPU 资源的规模和渠道稳定性。**IDC 核心壁垒在于获取一线城市及周边地区的土地资源与能耗指标，并通过绿色节能技术降低运营成本。政策要求数据中心 PUE 逐步降低，且在绿色电力比例要求下，一线城市能耗指标审批趋严，数据中心逐步向其他非一线城市部署。算力租赁行业的竞争力取决于 GPU 囤积规模与获卡渠道稳定性，尤其在英伟达高端芯片如 H100、H20 供应受限背景下，资源垄断性成为关键。国产替代（如华为昇腾 910B）虽加速，但性能仍有差距，短期内无法完全替代英伟达系芯片单次大模型训练需消耗千卡级 GPU 集群。
- **相关标的：**润建股份、大位科技、利通电子、朗科科技、润泽科技、光环新网、铜牛信息、中贝通信、恒润股份、海南华铁、奥飞数据、弘信电子。
- **风险提示：**技术发展不及预期，中美地缘政治风险，行业竞争加剧。

2025 年 04 月 03 日

证券分析师 张良卫

执业证书：S0600516070001

021-60199793

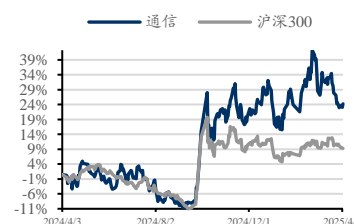
zhanglw@dwzq.com.cn

证券分析师 周高鼎

执业证书：S0600523030003

zhougd@dwzq.com.cn

行业走势



相关研究

《AI 算力跟踪深度：辨析 Scale Out 与 Scale Up——AEC 在光铜互联夹缝中挤出市场的 What、Why、How》

2025-01-07

《光通信跟踪深度：以太网在 AI 算力投资中的 Why、How 与 What》

2024-06-20

内容目录

1. IDC 和算力租赁区别：“地皮”与“楼盘”的关系	4
1.1. IDC 和算力租赁的商业模式区别在于“卡”的所有权.....	4
1.2. 需求视角：IDC 倾向帮运维，算力租赁则是“算力即用”	4
1.3. 供给视角：IDC 赚服务器量增的运维，算力租赁价格弹性大.....	5
2. 模型训练加速优化，算力租赁业务弹性更大	6
2.1. 模型训练优化，Agent 带动应用训练.....	6
2.2. 供给受限，算力租赁仍有涨价空间.....	7
3. 拥有算力资源和大厂订单为核心壁垒	9
4. 相关标的	10
5. 风险提示	11

图表目录

图 1: IDC 商业模式.....4

图 2: 第三方算力租赁商.....4

图 3: 需求端对比.....5

图 4: IDC 运营内部示意图.....5

图 5: 英伟达 H100 2023 年流向.....5

图 6: Generative Agents 运行流程图.....7

图 7: 中国算力产业链.....8

图 8: 中国智能算力发展情况 百亿亿次/秒(EFLOPS).....8

图 9: 中国整体 IDC 市场规模及趋势（单位：亿元）.....8

图 10: IDC 行业第三方服务商竞争梯队.....9

表 1: 腾讯和阿里对于云相关服务的需求点.....6

表 2: 算力租赁与 IDC 相关公司梳理.....10

1. IDC 和算力租赁区别：“地皮”与“楼盘”的关系

1.1. IDC 和算力租赁的商业模式区别在于“卡”的所有权

IDC 仅提供算力运营环境：互联网数据中心（Internet Data Center），简称 IDC，指的是一种拥有完善的设备，包括高速互联网接入带宽、高性能局域网络、可靠的机房环境等、专业化的管理、完善的应用的服务平台，**专业化服务器机房，提供数据中心租赁托管服务。**

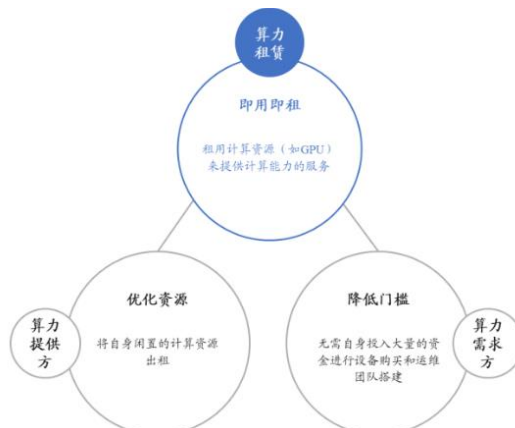
本质为数字地产商，提供物理空间+电力+网络，客户自备服务器或租赁整机柜。

图1：IDC 商业模式



数据来源：CDSN，东吴证券研究所

图2：第三方算力租赁商



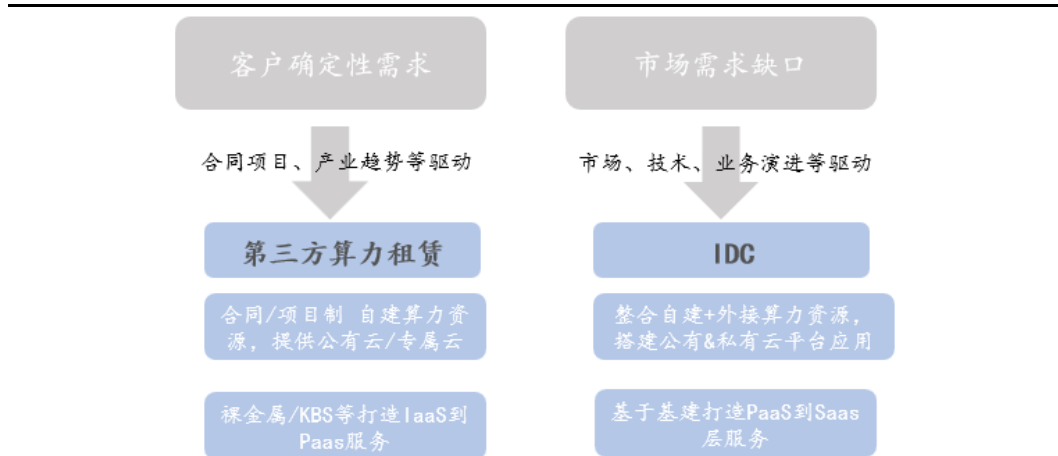
数据来源：未来智库，东吴证券研究所

第三方算力租赁基本是云服务的垂直形态，仅提供虚拟化计算资源（如 GPU/CPU 集群），客户按需租用。对于使用客户来说，算力租赁是一种按需付费的轻资产模式，依赖外部 IDC 部署服务器，聚焦资源调度与客户对接，如 AI 公司训练需求。位于算力产业链上游，是算力中心基础设施集成建设方和运营方，提供算力中心机柜租赁服务、端口服务以及其他增值服务

1.2. 需求视角：IDC 倾向帮运维，算力租赁则是“算力即用”

IDC 的本质是解决企业“硬件安家”需求——客户已拥有服务器、GPU 等硬件设备，但缺乏专业场地与运维能力，需借助 IDC 提供稳定的电力、散热和网络环境。企业以此规避自建数据中心的巨额建设投入（CAPEX）与日常运维复杂度（OPEX），以“拎包入住”模式实现轻量化运营。比如金融机构自建服务器保障数据主权，却无力承担一线城市机房的高昂土地与能耗成本；或视频平台部署定制化存储设备时，通过 IDC 实现异地容灾。

图3：需求端对比



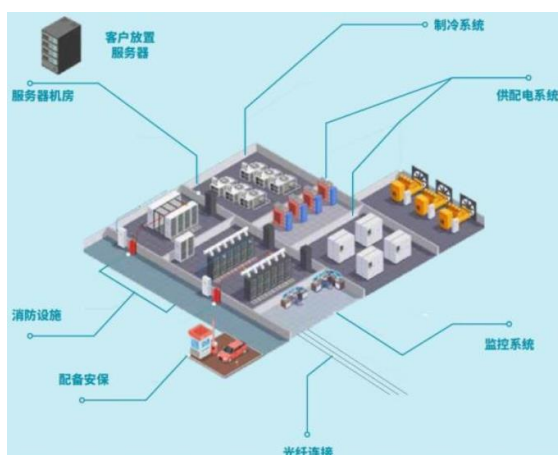
数据来源：东吴证券研究所绘制

算力租赁瞄准“算力即用”需求，即客户无自有硬件或面临短期算力缺口，如 AI 大模型训练、影视渲染，需快速获取即租即用的计算资源，避免硬件采购的周期风险与长期闲置。例如 AI 创业公司为千卡级 GPU 集群租赁 3-6 个月完成模型训练，或影视公司按项目周期租用算力平衡成本与效率。其核心价值在于将算力投入从固定资产（CAPEX）转化为随用随付的可变成本（OPEX），适配业务快速迭代。

1.3. 供给视角：IDC 赚服务器量增的运维，算力租赁价格弹性大

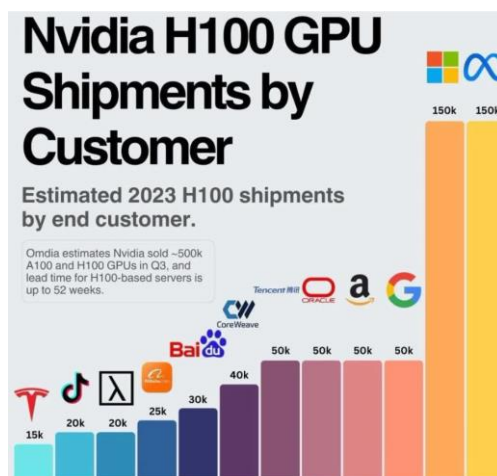
第三方 IDC 的盈利本质是：通过**机柜规模扩张与资源利用率提升**的双重驱动提升收入与利润率。收入增长与服务器/机柜的**物理数量扩张强相关，规模经济效益明显**。头部企业如润泽科技、数据港通过大规模布局一线城市周边机柜（如润泽已部署超 32 万架机柜），利用土地、能耗指标等稀缺资源抢占市场份额。IDC 机柜上架率直接影响毛利率，**上架率提升摊薄成本**。

图4：IDC 运营内部示意图



数据来源：首程控股《数据中心行业投资与价值观察》，东吴证券研究所

图5：英伟达 H100 2023 年流向



数据来源：Omdia Research，东吴证券研究所

算力租赁的核心价值在于**弹性供给与资源稀缺性溢价**。在供需错配前提下，算力租赁商通过**资源垄断性和需求刚性**获得定价权。单次大模型训练需消耗 128 张 H100 连续运行 3 个月，对应成本约 200 万元。英伟达 H100 芯片 2023 年全球出货量约 55 万张，中国仅获得约 5%（2.75 万张），在 GPU 硬约束下，若 100 家 AI 公司同时训练，需 1.28 万张卡——已接近国内供给上限，推高租赁溢价。

2. 模型训练加速优化，算力租赁业务弹性更大

2.1. 模型训练优化，Agent 带动应用训练

云厂商逐步向 DeepSeek 的效率与成本对标，以 DeepSeek 为代表的顶尖 AI 公司，通过**算法-硬件协同设计**，如稀疏化训练、混合精度优化及分布式训练框架升级（如 3D 并行策略），在降低成本基础上，提升了千卡级集群训练效率。云厂商需加速技术对齐，避免客户因成本差转向自建算力。

云厂商对算力租赁具有明确需求，其核心逻辑在于**灵活应对算力需求波动、降低硬件投入成本、优化资源利用率**。同时通过租赁模式规避美国对高端 GPU 芯片（如 A100/H100）的出口管制风险。具体而言，算力租赁使企业无需长期投入自建硬件即可应对 AI 大模型训练、云业务扩容等场景的突发需求，例如腾讯 2025 年算力租赁占比达 25%（约 200 亿元），主要用于应对算力暴增且自建周期长达一年半的挑战。此外，租赁模式将算力开支从资本支出（CAPEX）转为运营支出（OPEX），减轻现金流压力，同时通过第三方合规供应商（如国资背景的数据中心）获取被禁芯片资源，例如阿里云通过租赁获取 A800 卡，腾讯则依赖协创等供应商的海外算力资源。这种“自建+租赁”双轨策略既保障了算力可控，又实现了资源的高效复用与成本优化。

表1：腾讯和阿里对于云相关服务的需求点

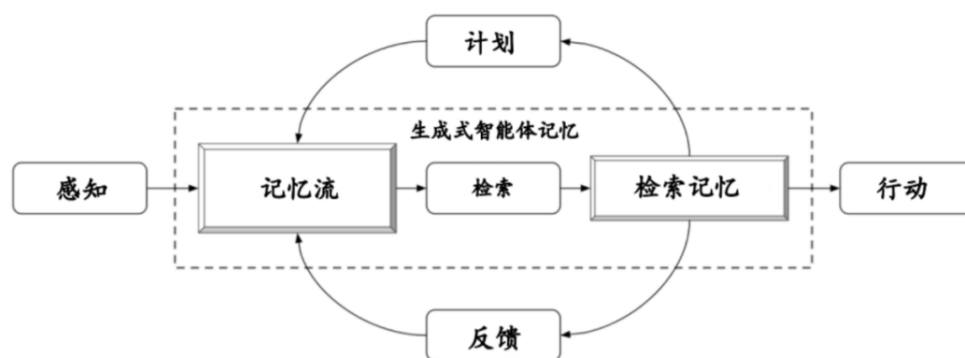
维度	腾讯	阿里
核心场景	AI 推理、全球化业务、短期弹性需求	AI 训练、国产化适配、数据中心基建
合规策略	海外回收卡租赁	东南亚代采+国产芯片生态
成本结构	租赁占比 25%，侧重 H20 推理	租赁占比 70%，自建与租赁协同
技术布局	自研推理芯片（12nm）	开源框架优化+平头哥芯片生态

数据来源：东吴证券研究所整理

腾讯与阿里在算力需求上的核心差异，体现在资源控制策略与技术路线选择的分化，进而导致其 GPU 卡采购方向不同。腾讯选用 H20 和 H100 系列以兼顾合规性与弹性，应对双十一、游戏渲染等突发场景及金融医疗等对本地化数据存储的要求，H20 占比需求更高，适配国产化需求，H100/H200 则通过租赁海外回收卡来兼顾性能与成本控制。相比之下，阿里则以自建为基础、租赁为辅，加速国产化替代，主推 H200 并大规模部署寒武纪 590 与昇腾 910C 等国产卡，以支撑多模态大模型高强度推理需求，同时通过东南亚布局服务出海企业，在规避供应链风险的同时实现算力成本优化。

Agent（智能体）对算力的需求提升主要源于其渗透率、日均调用次数及任务复杂度的三重增长。随着 DeepSeek、Manus 等 AI Agent 的爆发，其处理复杂任务（如数据整理、研究报告生成）时需消耗远超对话场景的 token 量，推动算力需求陡增。同时，Agent 作为流量入口，预计 2030 年前将创造近千万 B200 的算力需求，远超多模态模型。此外，企业端因 AI 应用扩展（如智能制造、金融分析）开始自建模型训练，进一步加剧算力压力——例如 DeepSeek 单日推理调用 2500 万次，需 1814 张 H800 芯片支持，而企业级训练场景对算力冗余和效率的要求更高。这种双重驱动下，算力需求从训练端向推理端迁移，推动全球算力市场规模持续扩张。

图6：Generative Agents 运行流程图



数据来源：infoq，东吴证券研究所

2.2. 供给受限，算力租赁仍有涨价空间

供给端硬约束下，因稀缺性溢价与成本传导机制，算力租赁仍有涨价空间。美国对华芯片出口限制持续加码，英伟达 H100/H20 等高端 GPU 对华供应被严格限制。2024 年中国市场 H100 供应量远远小于需求量，且 H20 芯片被完全禁售，导致国内算力租赁商面临“无卡可用”的困境。华为昇腾 910B 等国产芯片虽加速替代，但性能与生态适配

度（CUDA 生态依赖）仍存在显著差距，短期内难以填补需求缺口。

英伟达 H20 禁售后，H800 服务器单价从 200 万元涨至 300 万元，增加 50%，**硬件端的成本压力已经传导到租赁端**。而**运维方面，算力集群能耗占训练总成本大，液冷等节能技术普及需额外投入，进一步推高租金**。

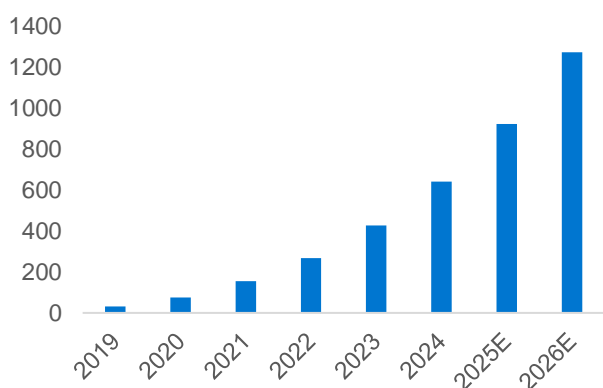
图7：中国算力产业链



数据来源：科智咨询、东吴证券研究所

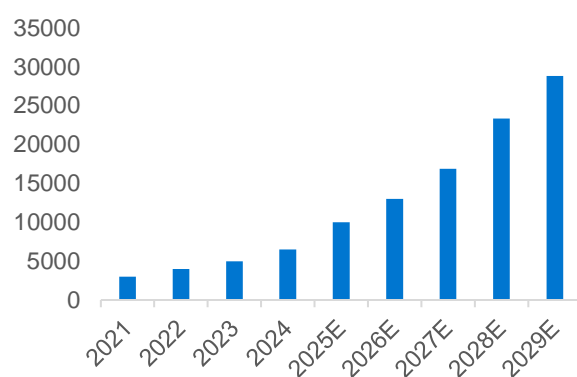
未来 5 年智算将有望高速增长。在智能算力方面,根据《智能计算中心创新发展指南》,2022 年我国智能算力规模快速增长,达到 268 百亿亿次每秒(EFLOPS),超过通用算力规模,预计未来 5 年中国智能算力规模的年复合增长率将达 52.3%。据国际数据公司（IDC）预测，2024 年，中国智能算力规模达 725.3 百亿亿次/秒(EFLOPS)，同比增长 74.1%。中研普华产业研究院报告显示，到 2026 年，国内算力租赁潜在收入市场规模有望达到 2600 亿元，且将以每年 20%以上的速度快速增长。

图8：中国智能算力发展情况 百亿亿次/秒(EFLOPS)



数据来源：《2023 新型算力中心调研报告》，东吴证券研究所

图9：中国整体 IDC 市场规模及趋势（单位：亿元）



数据来源：前瞻产业研究院，东吴证券研究所

3. 拥有算力资源和大厂订单为核心壁垒

IDC 行业的竞争本质是**资源禀赋争夺战**，核心壁垒在于获取一线城市及周边地区的**土地资源与能耗指标**，并通过绿色节能技术降低运营成本。政策要求数据中心 PUE 需低于 1.3，且在绿色电力比例要求下（如绿电使用率需达 80%），一线城市能耗指标审批趋严。

一线城市及周边地区（如北京、上海、广州）因网络延迟低、客户集中度高，成为 IDC 布局的核心区域，但土地资源稀缺且价格高昂。例如，润泽科技在京津冀、长三角布局超 32 万架机柜，依托土地先发优势形成竞争壁垒；西部地区虽土地成本低，但需求密度不足，需依赖“东数西算”政策引导存储类需求西迁。

图10：IDC 行业第三方服务商竞争梯队



数据来源：前瞻研究院、东吴证券研究所

算力租赁行业的竞争力取决于 GPU 囤积规模与获卡渠道稳定性，尤其在英伟达高端芯片如 H100、H20 供应受限背景下，资源垄断性成为关键。国产替代（如华为昇腾 910B）虽加速，但性能差距显著，短期内无法完全替代英伟达系芯片单次大模型训练需消耗千卡级 GPU 集群，如通义千问需万卡级算力，头部租赁商通过大规模囤卡形成供给壁垒。国际关系导致英伟达 H100/H20 对华出口受限，2024 年全球 H100 基本被海外大厂所购买，因此国内**具备海外供应链能力的厂商可绕过限制，形成渠道优势。**

4. 相关标的

表2: 算力租赁与 IDC 相关公司梳理

公司名称（股票代码）	机柜/算力规划	合作伙伴 / 重点布局	其他亮点
润建股份（002929）	- 规划建设总规模 20,000 个机柜- 一期已建成 6,000 个机柜- 单集群算力最高可达 3216P 单精度浮点算力	- 与阿里云合作建设“中国-东盟智算云”，提供 2500P 算力资源池- 通过统一平台实现资源调度	- 总投资 36 亿元- 通过国标 A 级认证- 融合绿色能源技术实现“零碳智算”，目标成为国内领先的智算服务商
大位科技（600589）	- 子公司张北榕泰云计算数据中心规划 12,000 个机柜- 总用电量 130MW，目标提供近万 P 算力	- 获得“东数西算”政策支持	- PUE 值优化至 1.25 以下- 享受低价电价及能耗指标倾斜
利通电子（603629）	- 2024 年战略转型 AI 算力租赁- 对外出租算力超 10,000P（自有 4,000P + 转租 6,000P）	- 与腾讯长三角超算中心、深圳奔峰数据中心深度绑定	- 重点发力 AI 算力租赁市场
朗科科技（300042）	- 机柜/规模未具体说明，主要提供 GPU 租赁服务	- 客户涵盖大模型训练、图像渲染等高算力需求场景	- GPU 型号齐全，灵活满足多场景算力需求
润泽科技（300442）	- 一线城市卫星城布局，现有 32 万架机柜- 2025 年规划超过 40 万架机柜	- 承接字节跳动等头部客户 60% 以上订单- 甘肃庆阳“风光储一体化”智算中心项目，总投资 55 亿元	- 深耕低 PUE 技术- 政策倾斜明显- 机柜资源规模大
光环新网（300383）	- 在北京、上海核心区域自建 5.2 万个机柜	- 与亚马逊云服务合作深度运营	- 机柜签约率超 95%- 位于京沪核心地段，资源优势明显
铜牛信息（300895）	- 北京密云 AIDC 项目：一期 6,800 个机柜（单柜 20kW），对应 10 万 P 算力，PUE<1.2 • 二期建设中，未来可扩容至 300MW（约 15 万 P）	- 在北京、天津等地运营天坛、三元桥等高品质数据中心- 政务、金融等客户占比高	- 机柜出租率稳定- 提供高安全等级服务
中贝通信（603220）	- 2024 年算力规模规划超 8000P	- 服务腾讯长三角 AI 超算中心- 政企客户占比提升至 30%	- 布局智能驾驶与工业质检算力租赁
恒润股份（603985）	- 算力规划未具体说明，逐步提高政企客户占比	- 政企客户为主要目标群体	- 具备国产替代路径- 借助华为昇腾芯片提升自主可控能力
海南华铁（603300）	- 2025 年目标建成 20 万 P 算力池	- 海南国资委提供资金与信用支持- 依托自贸港政策拓展跨境算力业务	- 目标跻身国内算力租赁第一梯队- 融资成本低
奥飞数据（300738）	- 2025 年投资 48 亿元扩建廊坊固安智算中心，新增 25,000 个高功率机柜- 总机柜数目标突破 5 万	- 深度绑定百度、阿里、华为等头部企业-	- 扩建力度大- 行业龙头客户资源丰富

		中标联通智算云项目 (2.16 亿元)	
弘信电子 (300657)	- 截至 2024 年底已建成 2.5 万 P	- 满足客户对英伟达算	- 跻身国内算力第一
	算力集群- 2025 年规划扩展至 5 万 P 以上	力芯片及 AI 服务器 租赁需求	梯队- 掌握优质 GPU 资源
数据来源：东吴证券研究所整理			

5. 风险提示

- 技术发展不及预期。AI 算力基建能否持续主要取决于 AI 大模型等技术的发展，如果 AI 技术发展不及预期，可能会影响算力基建投资速度。
- 中美地缘政治风险。美国可能加大对中国的科技制裁，导致中国无法获取最先进的 AI 芯片。
- 行业竞争加剧。算力基建市场发展迅速，新进参与者众多，可能会导致行业竞争加剧。

免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司（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，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。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。

在法律许可的情况下，东吴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。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，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。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 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），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

买入：预期未来 6 个月个股涨跌幅相对基准在 15%以上；

增持：预期未来 6 个月个股涨跌幅相对基准介于 5%与 15%之间；

中性：预期未来 6 个月个股涨跌幅相对基准介于-5%与 5%之间；

减持：预期未来 6 个月个股涨跌幅相对基准介于-15%与-5%之间；

卖出：预期未来 6 个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

增持：预期未来 6 个月内，行业指数相对强于基准 5%以上；

中性：预期未来 6 个月内，行业指数相对基准-5%与 5%；

减持：预期未来 6 个月内，行业指数相对弱于基准 5%以上。

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应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所
苏州工业园区星阳街 5 号
邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>