仅供内部参考，请勿外传

国泰海通证券 GUOTAI HAITONG SECURITIES

从高端计算到 AI 超集群，国产算力龙头加速进阶

中科曙光(603019)
计算机/信息科技

姓名	电话	邮箱	登记编号
杨林(分析师)	021-23183969	yanglin2@gtht.com	S0880525040027
魏宗(分析师)	021-23180000	weizong@gtht.com	S0880525040058

评级： 增持

目标价格(元)： 123.88

当前价格(元):	89.63

本报告导读：

AI 算力需求与政策推动超大规模智算基础设施建设，公司依托“两光”协同、十万卡超集群及全栈产品能力，有望持续受益国产算力基础设施升级。

投资要点：

- **投资建议：维持“增持”评级。**公司是国内高端计算及算力基础设施领域的重要企业，已形成覆盖高端计算机、通用服务器、智能计算、存储、液冷、数据中心、软件平台及算力服务的完整产品体系。我们预计 2026-2028 年公司营业收入分别为 175.74/206.72/244.71 亿元，归母净利润分别为 28.77/35.95/43.89 亿元，EPS 分别为 1.97/2.46/3.00 元。按照 PE 和 PB 估值，最终给予公司 2026 年 63 倍 PE，公司合理市值为 1812.51 亿元，对应目标价 123.88 元。
- **国产算力由单点替代向全栈系统升级，超节点与超大规模集群成为基础设施演进核心方向。**政策持续推动国产算力由芯片、服务器等单点突破向软硬件生态和超大规模智算基础设施建设延伸，国产算力已进入万卡级集群规模化部署阶段。随着加速卡数量增加，系统性能约束由单卡算力扩展至卡间互联、存储带宽、散热及集群调度。
- **“曙光+海光”贯通核心芯片与算力系统，全栈能力由服务器进一步延伸至超节点、互联、存储和液冷。**海光 CPU、DCU 构成国产计算芯片底座，中科曙光围绕处理器构建服务器、超节点、高速网络、存储和集群管理平台，曙光数创提供液冷基础设施。ScaleX640 将单机柜级高带宽通信域扩大至 640 张加速卡，scaleFabric 提供 400Gb/s 端口带宽、低于 1 微秒端侧时延并支持十万卡以上集群扩展，公司产品体系已由单体计算设备向一体化算力基础设施延伸。
- **曙光 8000 验证十万卡级系统工程与运营能力，80 亿元募投继续强化下一代超节点、训推和存储能力。**2026 年 7 月，全国产十万卡 AI 超集群曙光 8000 落成投用并接入国家超算互联网郑州核心节点，上线首周实现满载运行，日均处理作业超过 15 万个、单日峰值超过 50 万个，并已完成近千项超智融合应用适配；公司随后与北京科学智能研究院启动第二套全国产十万卡超智融合算力系统。公司 80 亿元可转债募投拟投入建设先进算力集群、下一代 AI 训推一体机和国产化先进存储，其中先进算力集群项目规划推动 Scale-Up 域由 640 卡向 1024 卡及以上扩展、网络端口由 400G 向 800G/1600G 升级，为公司承接超大规模国产算力基础设施建设提供产品和技术储备。
- **风险提示：**技术迭代及研发进展不及预期；关键部件供应链及国产化适配进度不及预期；产业政策支持力度及国际合作环境变化；超大规模算力集群建设进度及市场需求不及预期。

财务摘要(百万元)	2024A	2025A	2026E	2027E	2028E
营业总收入	13,148	14,964	17,574	20,672	24,471
(+/-)%	-8.4%	13.8%	17.4%	17.6%	18.4%
净利润(归母)	1,911	2,176	2,877	3,595	4,389
(+/-)%	4.1%	13.9%	32.2%	24.9%	22.1%
每股净收益(元)	1.31	1.49	1.97	2.46	3.00
净资产收益率(%)	9.4%	9.8%	11.9%	13.4%	14.7%
市盈率(现价&最新股本摊薄)	66.02	57.98	43.86	35.10	28.75

资料来源： Wind，国泰海通证券研究

请务必阅读正文之后的免责条款部分

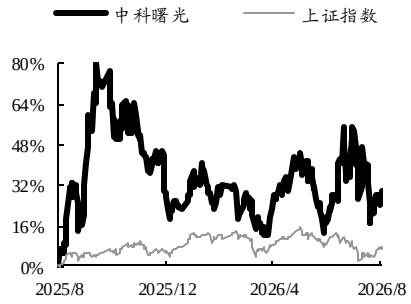
交易数据

52 周内股价区间(元)	69.73-125.50
总市值(百万元)	131,139
总股本/流通 A 股(百万股)	1,463/1,463
流通 B 股/H 股(百万股)	0/0

资产负债表摘要(LF)

股东权益(百万元)	22,556
每股净资产(元)	15.42
市净率(现价)	5.8
净负债率	-10.91%

52周股价走势图



升幅(%)	1M	3M	12M
绝对升幅	-16%	-11%	29%
相对指数	-16%	-3%	23%

相关报告

IT 基础设施领军，所处服务器、存储赛道保持高景气 2025.11.13

仅供内部参考，请勿外传

财务预测表

资产负债表(百万元)	2024A	2025A	2026E	2027E	2028E
货币资金	6,517	6,226	8,852	12,289	16,492
交易性金融资产	0	0	0	0	0
应收账款及票据	2,344	2,847	3,047	3,430	3,940
存货	3,859	3,533	4,218	4,790	5,552
其他流动资产	1,267	2,756	2,634	2,665	2,730
流动资产合计	**13,987**	**15,362**	**18,750**	**23,174**	**28,714**
长期投资	7,559	8,967	8,967	8,967	8,967
固定资产	2,854	3,304	3,420	3,479	3,480
在建工程	749	729	729	729	729
无形资产及商誉	4,159	3,781	3,588	3,376	3,143
其他非流动资产	7,311	8,811	8,834	8,834	8,838
非流动资产合计	**22,631**	**25,592**	**25,537**	**25,384**	**25,157**
总资产	**36,617**	**40,954**	**44,287**	**48,557**	**53,870**
短期借款	22	846	846	846	846
应付账款及票据	3,228	4,609	5,181	6,413	8,010
一年内到期的非流动负债	772	199	163	163	163
其他流动负债	1,564	2,722	3,012	3,504	4,121
流动负债合计	**5,586**	**8,376**	**9,201**	**10,925**	**13,140**
长期借款	98	87	87	87	87
应付债券	0	0	0	0	0
租赁负债	26	2	2	2	2
其他非流动负债	9,580	9,420	9,900	9,900	9,900
非流动负债合计	**9,704**	**9,509**	**9,988**	**9,988**	**9,988**
总负债	**15,291**	**17,885**	**19,190**	**20,914**	**23,128**
实收资本(或股本)	1,463	1,463	1,463	1,463	1,463
其他归母股东权益	18,939	20,770	22,761	25,268	28,330
归属母公司股东权益	**20,402**	**22,233**	**24,224**	**26,731**	**29,793**
少数股东权益	925	836	874	912	949
股东权益合计	**21,327**	**23,069**	**25,098**	**27,644**	**30,742**
总负债及总权益	**36,617**	**40,954**	**44,287**	**48,557**	**53,870**

现金流量表(百万元)	2024A	2025A	2026E	2027E	2028E
经营活动现金流	2,722	1,313	3,302	4,662	5,506
投资活动现金流	-1,320	-1,194	-212	-118	44
筹资活动现金流	-2,257	-233	-457	-1,106	-1,347
汇率变动影响及其他	6	-14	-6	0	0
现金净增加额	-849	-128	2,626	3,437	4,203
折旧与摊销	878	872	1,086	1,163	1,237
营运资本变动	134	-1,317	-335	408	548
资本性支出	-858	-889	-1,007	-1,006	-1,005

利润表(百万元)	2024A	2025A	2026E	2027E	2028E
营业总收入	**13,148**	**14,964**	**17,574**	**20,672**	**24,471**
营业成本	9,314	10,387	12,096	14,094	16,621
税金及附加	56	75	80	96	114
销售费用	773	703	817	951	1,113
管理费用	353	377	439	506	587
研发费用	1,292	1,671	1,757	2,067	2,447
EBIT	**1,772**	**1,805**	**3,317**	**4,115**	**4,994**
其他收益	544	303	534	596	683
公允价值变动收益	0	0	0	0	0
投资收益	563	683	725	888	1,049
财务费用	-53	-68	-43	-69	-103
减值损失	-234	-318	-330	-330	-330
资产处置损益	2	8	5	7	9
营业利润	**2,287**	**2,494**	**3,362**	**4,187**	**5,102**
营业外收支	3	-8	-3	-4	-4
所得税	297	328	444	551	672
净利润	**1,993**	**2,158**	**2,915**	**3,633**	**4,425**
少数股东损益	82	-19	38	38	36
归属母公司净利润	1,911	2,176	2,877	3,595	4,389

主要财务比率	2024A	2025A	2026E	2027E	2028E
ROE(摊薄,%)	9.4%	9.8%	11.9%	13.4%	14.7%
ROA(%)	5.8%	5.6%	6.8%	7.8%	8.6%
ROIC(%)	6.9%	6.5%	11.0%	12.4%	13.6%
销售毛利率(%)	29.2%	30.6%	31.2%	31.8%	32.1%
EBIT Margin(%)	13.5%	12.1%	18.9%	19.9%	20.4%
销售净利率(%)	15.2%	14.4%	16.6%	17.6%	18.1%
资产负债率(%)	41.8%	43.7%	43.3%	43.1%	42.9%
存货周转率(次)	2.6	2.8	3.1	3.1	3.2
应收账款周转率(次)	5.5	6.0	6.2	6.6	6.9
总资产周转周转率(次)	0.4	0.4	0.4	0.4	0.5
净利润现金含量	1.4	0.6	1.1	1.3	1.3
资本支出/收入	6.5%	5.9%	5.7%	4.9%	4.1%
EV/EBITDA	37.82	44.90	26.90	21.78	17.78
P/E(现价&最新股本摊薄)	66.02	57.98	43.86	35.10	28.75
P/B(现价)	6.18	5.68	5.21	4.72	4.24
P/S(现价)	9.60	8.43	7.18	6.10	5.16
EPS-最新股本摊薄(元)	1.31	1.49	1.97	2.46	3.00
DPS-最新股本摊薄(元)	0.27	0.45	0.59	0.74	0.91
股息率(现价,%)	0.3%	0.5%	0.7%	0.9%	1.1%

资料来源：Wind，国泰海通证券研究

仅供内部参考，请勿外传

目录

1. 盈利预测和投资建议 4
1.1. 收入和盈利预测 4
1.2. 估值分析 4
2. 中科曙光：国产高端计算龙头，经营业绩稳健增长 5
2.1. 深耕高端计算：国产算力基础设施龙头 5
2.2. AI 算力需求驱动收入增长，经营质量稳步提升 7
3. 行业趋势：AI 驱动算力基础设施扩容，国产化进程持续推进 9
3.1. Scaling Law 持续演进，训练与推理共同拉动算力需求 9
3.2. CSP 资本开支维持高位，AI 基础设施建设周期延续 11
3.3. AI 算力基础设施市场持续扩容，AI 服务器、超节点与 HPC 协同增长 12
3.4. 算力架构由单机向超节点演进，Scale-Up 互联成为系统升级核心 12
3.5. 算力功率密度持续提升，液冷由可选配置迈向基础设施刚需 14
3.6. 国产算力加快突破，生态适配与系统集成推动规模化落地 15
4. 曙光+海光协同构筑国产算力底座，全栈能力向超大规模算力集群延伸 16
4.1. 曙光+海光贯通芯片与系统，光合生态强化国产算力适配 16
4.2. ScaleX 超节点与 scaleFabric 高速互联推动计算产品向超集群升级 17
4.3. 存储与液冷补齐数据供给与热管理能力，提升单项目价值量 19
4.4. 曙光 8000 十万卡系统落地，验证超大规模集群工程与运营能力 20
4.5. 80 亿元募投推动 Scale-Up 向 1024 卡升级，持续强化集群、训推与存储能力 21
5. 风险提示 22

仅供内部参考，请勿外传

1. 盈利预测和投资建议

1.1. 收入和盈利预测

1）IT 设备业务：公司 IT 设备业务主要包括高端计算机、通用服务器、智能计算设备及存储等产品，高端计算和通用服务器构成业务基本盘，AI 服务器、ScaleX 超节点、万卡集群、高性能存储等产品则打开增量空间。考虑到国产算力基础设施投入持续增长，公司产品由单台服务器向机柜级超节点及集群级系统升级，同时存储、液冷等产品配套范围扩大，我们预计 2026-2028 年 IT 设备业务营收将分别实现 15.00%、16.00%和 18.00%的同比增长，营收额分别达 143.79 亿元、166.80 亿元和 196.82 亿元。考虑到服务器及设备业务市场竞争较为充分，同时高价值量智能计算产品占比提升有望对冲部分价格压力，我们预计 2026-2028 年该业务毛利率维持在 27.00%左右。

2）围绕高端计算机的软件开发、系统集成及技术服务：公司依托高端计算设备、存储、网络、液冷和基础软件能力，为客户提供软件开发、系统集成、云计算、数据中心建设、算力调度及技术服务，业务形态由单一设备销售逐步延伸至“算、存、网、电、冷、管、软”一体化交付。随着智算中心、云计算和数据中心项目持续落地，后续增速预计逐步回归稳健。我们预计 2026-2028 年软件开发、系统集成及技术服务业务收入将分别实现 30.00%、25.00%和 20.00%的同比增长，营收额分别达 31.80 亿元、39.75 亿元和 47.71 亿元。随着软件平台、定制开发、系统优化和运维服务等高附加值业务占比提升，我们预计 2026-2028 年该业务毛利率分别为 50.00%、52.00%和 53.00%。

3）其他业务：来源于主营业务以外的其他业务，占营业收入比例较低。考虑到该业务缺乏明确的大规模扩张驱动，我们预计 2026-2028 年其他业务收入每年均实现 10.00%的同比增长，营收额分别为 0.15 亿元、0.17 亿元和 0.18 亿元；基于历史盈利水平，预计同期毛利率维持在 40.00%左右。

表1：公司收入拆分表（百万元）

		2025	2026E	2027E	2028E
IT 设备业务	营收	12,503.50	14,379.03	16,679.67	19,682.01
	同比	6.81%	15.00%	16.00%	18.00%
	毛利率	27.31%	27.00%	27.00%	27.00%
围绕高端计算机的软件开发、系统集成及技术服务	营收	2,446.42	3,180.35	3,975.44	4,770.53
	同比	75.34%	30.00%	25.00%	20.00%
	毛利率	47.25%	50.00%	52.00%	53.00%
其他业务	营收	13.73	15.10	16.62	18.28
	同比	-70.37%	10.00%	10.00%	10.00%
	毛利率	42.17%	40.00%	40.00%	40.00%
总营收		14,963.64	17,574.48	20,671.73	24,470.82
同比		13.81%	17.45%	17.62%	18.38%
毛利率		30.58%	31.17%	31.82%	32.08%

资料来源：Wind，国泰海通证券研究

最终我们预计 2026-2028 年公司营业收入分别为 175.74/206.72/244.71 亿元，归母净利润分别为 28.77/35.95/43.89 亿元，EPS 分别为 1.97/2.46/3.00 元。

1.2. 估值分析

我们采用 PE 和 PB 两种估值方法，对公司进行估值。

1）PE 估值法

我们选取浪潮信息、软通动力、紫光股份和锐捷网络作为可比公司。其中，浪潮信息与中科曙光均聚焦智算基础设施，产品覆盖通用服务器、AI 服务

仅供内部参考，请勿外传

器、存储及液冷等环节；软通动力通过软通计算机布局服务器、存储、超节点和 AI 基础设施，同时拥有软件开发、数字技术服务和智算中心建设运营能力；紫光股份控股子公司新华三的产品覆盖服务器、网络、存储、云计算和安全等 ICT 基础设施；锐捷网络主要提供交换机、无线产品、云桌面及网络安全产品，并面向互联网和运营商客户提供智算网络、数据中心交换机及云网解决方案，与中科曙光 AI 基础设施、数据中心和网络业务存在一定可比性。

我们预计 2026-2028 年中科曙光营业收入分别为 175.74/206.72/244.71 亿元，归母净利润分别为 28.77/35.95/43.89 亿元，EPS 分别为 1.97/2.46/3.00 元。参考可比公司 2026 年平均估值水平约为 62.46 倍 PE，考虑公司在国产高端计算、液冷、存储和 ScaleX 超节点领域的产品布局及 AI 基础设施成长潜力，我们给予公司 2026 年 63 倍 PE，合理市值为 1812.51 亿元。

表2：可比公司 PE 估值表（表中数据更新于 2026/8/4）

股票代码	股票简称	收盘价（元）	EPS（元）			PE		
			2026E	2027E	2028E	2026E	2027E	2028E
000977.SZ	浪潮信息	74.14	2.74	3.72	4.92	27.06	19.93	15.07
301236.SZ	软通动力	38.36	0.34	0.47	0.61	112.82	81.62	62.89
000938.SZ	紫光股份	34.32	1.02	1.34	1.71	33.65	25.61	20.07
301165.SZ	锐捷网络	107.59	1.41	2.04	2.85	76.30	52.74	37.75
	平均值					**62.46**	**44.97**	**33.94**

资料来源：Choice，国泰海通证券研究（公司盈利预测为 Choice 一致预期。）

2）PB 估值法

我们仍选取浪潮信息、软通动力、紫光股份和锐捷网络作为可比公司。

预计 2026-2028 年中科曙光净资产分别为 242.24/267.31/297.93 亿元。参考可比公司 2026 年平均估值水平为 7.09 倍 PB，综合考虑公司在国产算力基础设施领域的龙头地位、软件及系统服务收入占比提升，以及海光信息股权账面价值与市场价值之间的差异，我们给予公司 2026 年 7.5 倍 PB，对应合理估值为 1816.8 亿元人民币。

表3：可比公司 PB 估值表（表中数据更新于 2026/8/4）

股票代码	股票简称	收盘价（元）	每股净资产（元）			PB		
			2026E	2027E	2028E	2026E	2027E	2028E
000977.SZ	浪潮信息	74.14	17.39	20.88	25.48	4.26	3.55	2.91
301236.SZ	软通动力	38.36	10.96	11.94	12.37	3.50	3.21	3.10
000938.SZ	紫光股份	34.32	6.09	7.05	8.04	5.64	4.87	4.27
301165.SZ	锐捷网络	107.59	7.19	9.02	11.88	14.96	11.93	9.06
	平均值					**7.09**	**5.89**	**4.83**

资料来源：Choice，国泰海通证券研究（公司盈利预测为 Choice 一致预期。）

综合以上估值方法，按照谨慎取低原则，我们认为公司合理估值为 1812.51 亿元人民币，对应目标价为 123.88 元。

2. 中科曙光：国产高端计算龙头，经营业绩稳健增长

2.1. 深耕高端计算：国产算力基础设施龙头

公司技术积淀深厚，已由高性能计算产业化主体成长为国内算力基础设施领域的重要企业。公司技术与产业化基础可追溯至国家“863”计划支持的曙光高性能计算机项目，上市公司前身天津曙光计算机产业有限公司成立于 2006 年，2010 年整体变更为曙光信息产业股份有限公司，并于 2014 年在上海证券交易所上市。公司早期主要从事高性能计算机、通用服务器和存储产品的研发制造，经过多年业务拓展，目前已形成覆盖高端计算机、通用

仅供内部参考，请勿外传

服务器、存储、安全、数据中心、液冷、云计算、智能计算及算力服务的产品与服务体系，业务定位逐步由计算设备供应商向算力基础设施综合服务商延伸。

公司经营模式以自主研发为核心，采用“研发＋生产＋销售＋服务”一体化运营模式。1）研发：坚持核心技术自主可控，持续加大智能计算、国产化软硬件、液冷、算力调度等关键技术投入；2）生产：采用自主生产为主、外协生产为辅的生产模式，保障产品品质与交付效率；3）销售：销售模式分为公司直接销售及通过经销商、集成商销售两种方式，其中以直销模式的销售为主；4）服务：建立全生命周期服务体系，提升客户粘性与综合价值。公司聚焦算力基础设施全产业链，以研发、生产、销售高性能计算机、通用服务器、存储等 IT 设备及围绕 IT 设备提供软件开发、系统集成与技术服务为核心，持续拓展云计算、数据中心、液冷、算力服务等高附加值业务。

图 1：中科曙光 AI 超集群系统

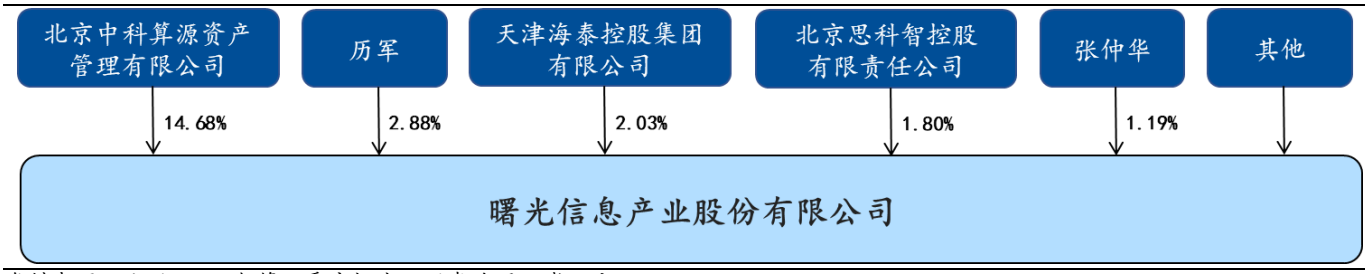


资料来源：科技日报

公司主要股东持股比例较为分散，控制权状态已变为无控股股东、无实际控制人。截至 2026 年一季度末，北京中科算源资产管理有限公司持有公司 2.15 亿股，持股比例为 14.68%，仍为公司第一大股东；公司董事长历军、天津海泰控股集团有限公司和北京思科智控股有限责任公司分别持股 2.88%、2.03%和 1.80%，除中科算源外不存在其他持股 5%以上的股东。中科算源原为公司控股股东，其唯一股东中国科学院计算技术研究所原为公司实际控制人；2026 年 4 月 15 日，公司公告称，鉴于中科算源持股比例降至 14.68%、其余股东持股较为分散，且不存在单一股东能够控制股东会表决结果、决定半数以上董事选任或通过一致行动及表决权委托安排影响公司控制权的情形，公司控制权状态变更为无控股股东、无实际控制人。本次变化不涉及股份及董事会构成变动，公司经营管理团队保持稳定。

仅供内部参考，请勿外传

图 2：公司主要股东持股比例较为分散（数据更新至 2026 年一季报）



资料来源：公司 2026 年第一季度报告，国泰海通证券研究

高管团队长期深耕曙光体系，兼具经营管理、技术研发、市场渠道和财务管理经验。公司现任高级管理人员共 5 名，董事长、总经理历军为经济学博士、教授级高级工程师，自 1996 年加入曙光体系，长期负责公司经营管理；高级副总裁李斌具有博士学历，由工程师、部门经理逐步成长为公司高级管理人员，具备技术研发和组织管理经验；任京暘、关宏明分别长期负责市场渠道、营销管理及云计算业务，翁启南自 2001 年起在公司财务体系任职，现任财务总监、董事会秘书。从任职经历看，除关宏明未披露具体入职时间外，其余 4 名高级管理人员截至 2026 年在曙光体系任职均超过 15 年，管理团队内部培养特征较为突出，对公司产品技术、客户渠道和经营体系具有较深理解。2026 年 5 月，公司完成董事会换届，历军出任董事长并继续兼任总经理，任京暘、李斌、关宏明继续担任高级副总裁，翁启南继续担任财务总监、董事会秘书，管理层整体保持稳定。

表 4：公司部分高级管理人员简历

姓名	职务	主要工作经历
历军	董事长、总经理	1968 年出生，中国国籍，经济学博士，教授级高级工程师。1996 年加入北京曙光天演信息技术有限公司，2001 年起任曙光信息产业（北京）有限公司总裁，2006 年加入天津曙光计算机产业有限公司，即曙光信息产业股份有限公司前身，现任公司董事长、总经理。
翁启南	财务总监、董事会秘书	1969 年出生，中国国籍，大学本科学历，高级会计师。2001 年加入曙光信息产业（北京）有限公司，历任财务部副总经理、财务部总经理，2006 年加入天津曙光计算机产业有限公司，现任公司财务总监、董事会秘书。
任京暘	高级副总裁	1971 年出生，中国国籍，大学本科学历。2001 年加入曙光信息产业（北京）有限公司，历任市场部总经理、渠道业务群组总经理，2006 年加入天津曙光计算机产业有限公司，现任公司高级副总裁。
李斌	高级副总裁	1982 年出生，中国国籍，博士学历。2009 年加入曙光信息产业股份有限公司，历任公司工程师、部门经理、副总裁，现任公司高级副总裁。
关宏明	高级副总裁	1969 年出生，中国国籍，大学本科学历。历任公司营销管理部副总经理、总裁助理、副总裁，曙光云计算集团股份有限公司总裁；2022 年 5 月至 2023 年 5 月任公司董事，现任公司高级副总裁。

资料来源：公司向不特定对象发行可转换公司债券募集说明书，国泰海通证券研究

海光信息与曙光数创构成公司重要产业及财务资产。截至 2025 年末，中科曙光持有海光信息 6.499 亿股、持股比例 27.96%，为其第一大股东；海光信息作为联营企业不纳入公司合并报表，中科曙光按照权益法确认投资损益，2025 年确认权益法投资收益 7.11 亿元，期末长期股权投资账面价值为 70.29 亿元。中科曙光全资持有曙光信息产业（北京）有限公司，后者直接持有曙光数创 62.07%股份，因此曙光数创作为控股子公司纳入中科曙光合并报表，其收入、成本、资产和负债直接进入合并报表，并在归母净利润计算中扣除少数股东损益；2025 年曙光数创实现营业收入 8.82 亿元。

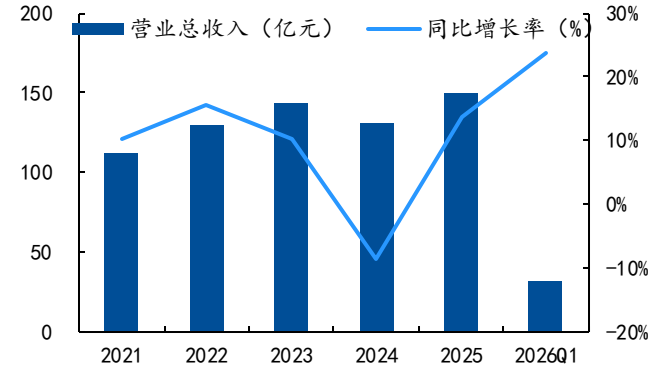
2.2. AI 算力需求驱动收入增长，经营质量稳步提升

营业收入与利润保持增长，扣非净利润增速明显高于收入增速。2025 年公司实现营业收入 149.64 亿元，同比增长 13.81%；实现归母净利润 21.76 亿元，同比增长 13.87%；实现扣非归母净利润 18.38 亿元，同比增长 33.97%。

仅供内部参考，请勿外传

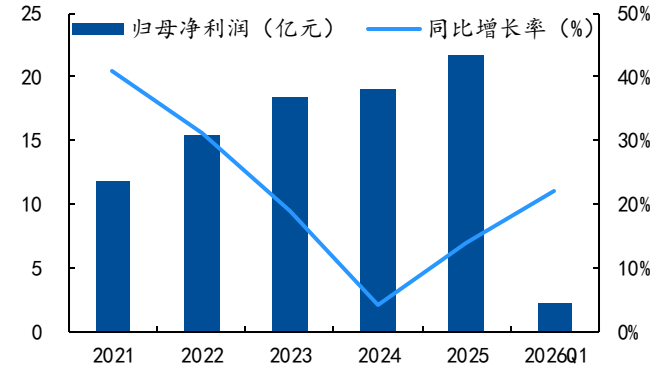
2026Q1 公司实现营业收入 31.99 亿元，同比增长 23.71%；实现归母净利润 2.28 亿元，同比增长 22.19%；实现扣非归母净利润 1.64 亿元，同比增长 53.30%，主营业务盈利改善趋势延续。

图 3：营业收入维持增长趋势



资料来源：Wind，国泰海通证券研究

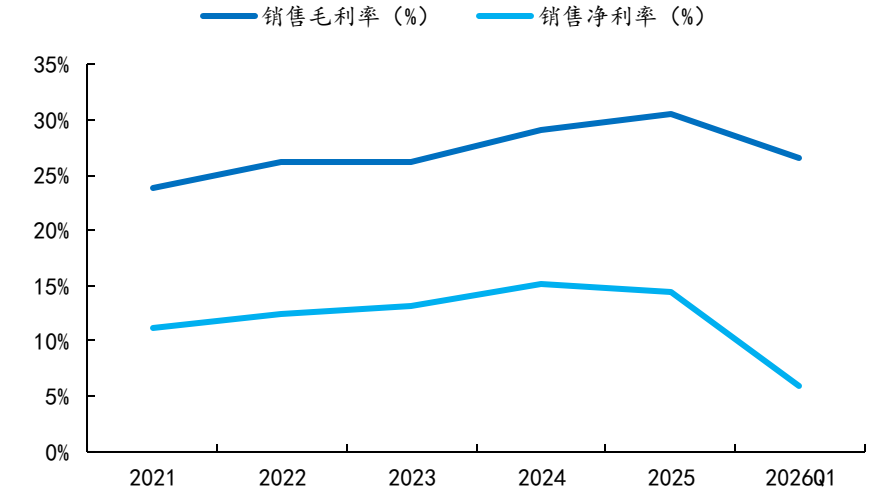
图 4：归母净利润稳健增长



资料来源：Wind，国泰海通证券研究

公司毛利率同比改善，归母净利率总体保持稳定。2025 年公司综合毛利率为 30.58%，同比提升 1.42 个百分点；归母净利率为 14.54%，同比基本持平。2026Q1 公司综合毛利率为 26.56%，同比提升 0.49 个百分点；归母净利率为 7.12%，同比下降 0.09 个百分点，整体保持稳定。公司收入和利润具有一定季节性，季度毛利率与全年毛利率不宜直接比较。

图 5：公司毛利率净利率整体呈现逐年改善趋势



资料来源：Wind，国泰海通证券研究

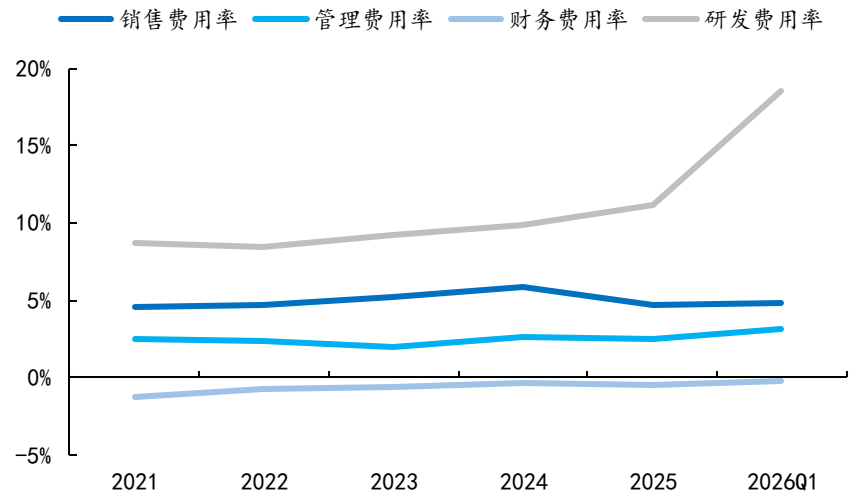
公司维持较高研发强度，研发费用增速明显高于收入增速。2025 年公司研发费用为 16.71 亿元，同比增长 29.33%，研发费用率由 2024 年的 9.83%提升至 11.17%。2026Q1 公司研发费用为 5.92 亿元，同比增长 51.73%，研发费用率由 2025Q1 的 15.09%提升至 18.51%，主要围绕智能算力、国产化软硬件、先进存储、液冷散热及算力调度等方向开展技术攻关。截至 2025 年末，公司拥有研发人员 3,199 人，占员工总数的 53.10%，研发人员占比维持较高水平。

费用管控有所优化，销售费用率和管理费用率同比下降。2025 年公司销售费用为 7.03 亿元，同比下降 9.08%，销售费用率由 5.88%下降至 4.70%；管理费用为 3.77 亿元，同比增长 6.93%，但管理费用率由 2.68%下降至 2.52%；财务费用为-0.68 亿元，财务费用率为-0.45%，主要由于利息收入高于利息支出。2026Q1 公司销售费用率和管理费用率分别为 4.82%和 3.19%，同比分别下降 0.69 和 0.16 个百分点；财务费用率为-0.22%，继续维持净利息收

仅供内部参考，请勿外传

入状态。

图 6：销售费用率和管理费用率同比下降，研发费用率总体呈上升趋势



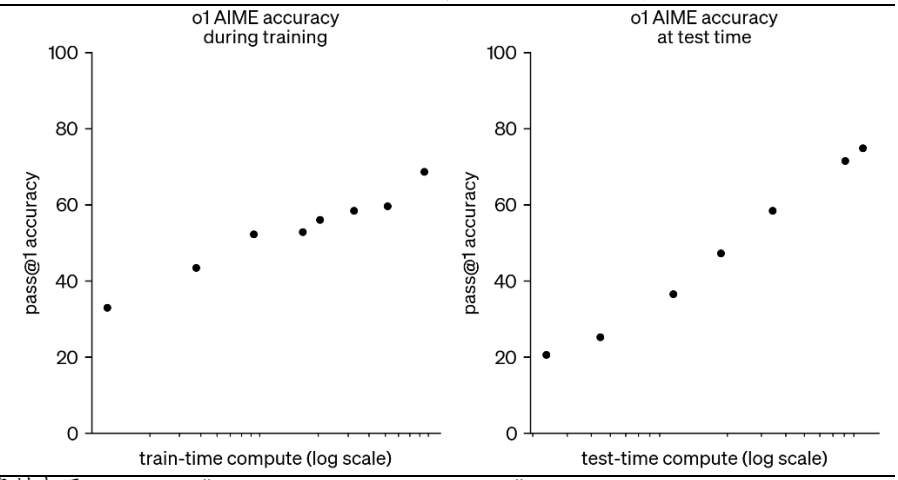
资料来源：Wind，国泰海通证券研究

3. 行业趋势：AI 驱动算力基础设施扩容，国产化进程持续推进

3.1. Scaling Law 持续演进，训练与推理共同拉动算力需求

Scaling Law 仍是推动大模型能力提升的重要技术路径，模型性能随参数规模、数据规模和训练计算量增加呈现可预测改善。OpenAI 在《Scaling Laws for Neural Language Models》中指出，语言模型交叉熵损失与模型规模、数据集规模和训练计算量之间呈幂律关系，部分规律横跨超过 7 个数量级。该研究意味着，在模型架构和训练方法相对稳定的条件下，持续增加参数、数据和计算投入，仍能够以可预测方式降低模型损失。

图 7：模型性能随训练时和测试时计算投入增加而提升



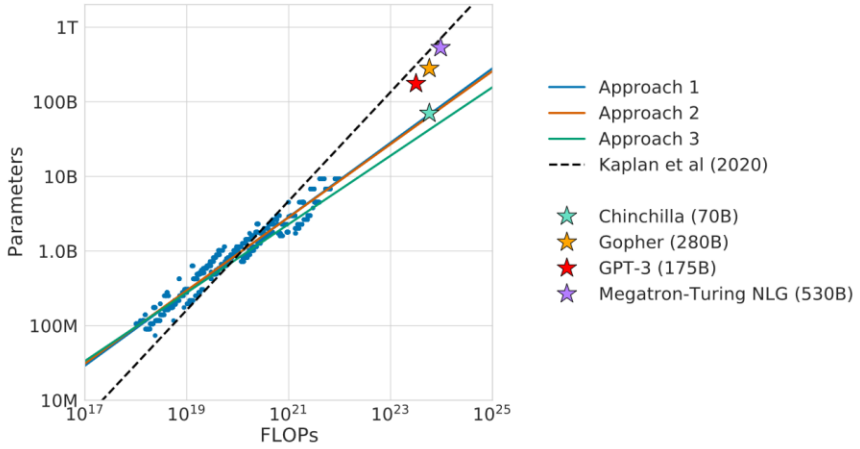
资料来源：OpenAI《Learning to Reason with LLMs》

Scaling Law 已由单纯扩大模型参数量，演进为参数规模、训练数据和计算预算的协同扩展。DeepMind 在 Chinchilla 研究中训练了超过 400 个语言模型，模型参数范围覆盖 7000 万至 160 亿，训练数据量覆盖 50 亿至 5000 亿 Token。研究结果显示，在计算最优条件下，模型参数量与训练 Token 数量应大致同比例增长，即模型参数每扩大一倍，训练 Token 数量也应扩大一倍。由此可见，更大的参数量需要更大规模的训练数据和计算资源配合。按照 Transformer 训练计算量近似与参数量和训练 Token 数量乘积成正比推算，

仅供内部参考，请勿外传

当二者同步扩大一倍时，训练计算量约扩大至原来的 4 倍。

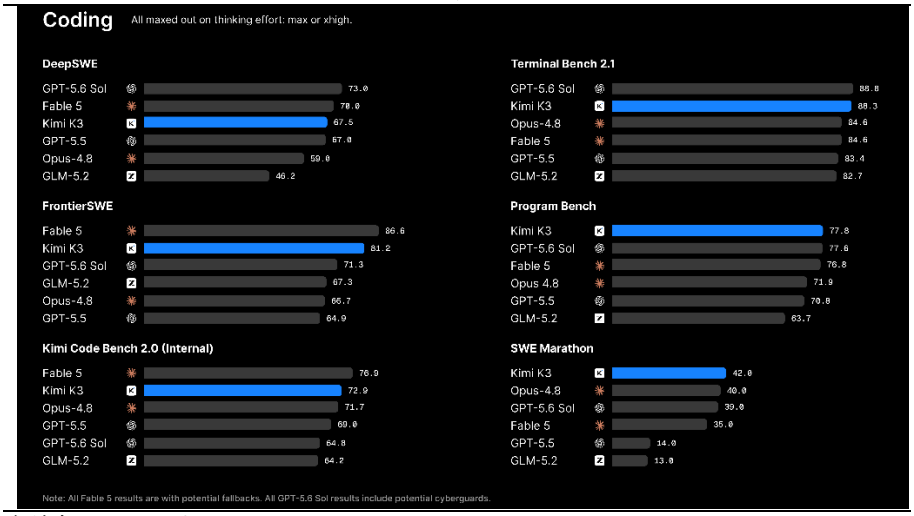
图 8：计算最优 Scaling 要求模型参数与训练数据协同扩展



资料来源：Hoffmann 等《Training Compute-Optimal Large Language Models》

Kimi K3 将开放模型参数规模推升至 2.8 万亿，表明前沿模型仍在向更大规模、更长上下文和更复杂任务扩展。Kimi K3 是首个达到 2.8 万亿参数的开放权重模型，原生支持视觉理解和 1M Token 上下文窗口，并面向长程编程、知识工作和深度推理等场景设计。Kimi K3 采用 KDA 混合线性注意力、Attention Residuals 和 Stable Latent MoE 架构，每次从 896 个专家中激活 16 个，叠加训练方法和数据配方优化后，官方测算其整体 Scaling 效率较 Kimi K2 提高约 2.5 倍。

图 9：Kimi K3 在多项编程基准测试中位居前列

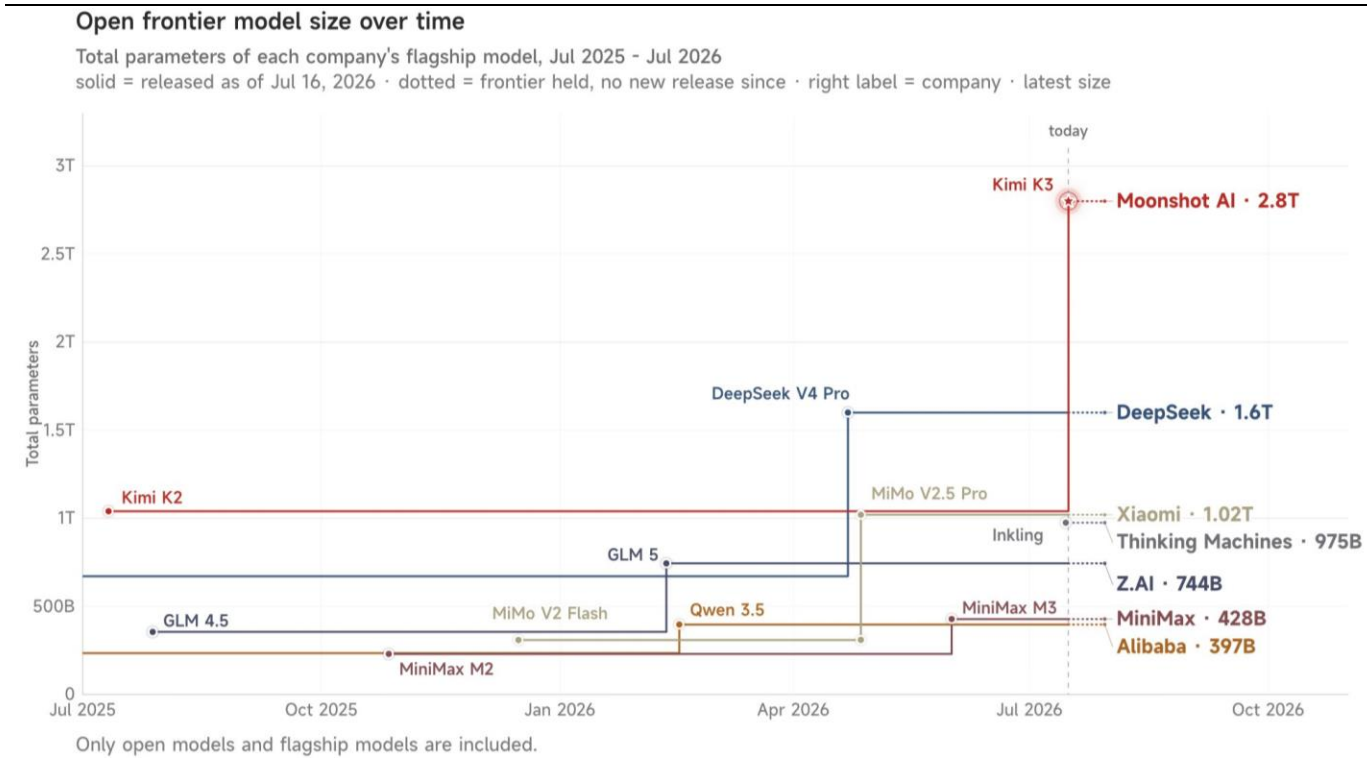


资料来源：月之暗面

模型规模扩大通常会增加单次推理计算量，并抬升显存容量、内存带宽和跨卡通信要求。Kimi K3 属于稀疏 MoE 模型，推理时仅激活 896 个专家中的 16 个，实际推理成本还取决于激活参数规模、上下文长度、输出长度、推理强度和缓存命中率。即便如此，2.8 万亿参数模型仍需存储和调度完整权重，并在专家并行过程中进行跨设备通信。月之暗面据此建议 Kimi K3 部署在拥有 64 张及以上加速卡的超节点配置中，以扩大高带宽通信域并提高推理效率。前沿模型参数量和 MoE 稀疏度继续提升，将同步强化对大容量显存、高带宽互联、超节点、分布式存储和集群调度能力的需求。

仅供内部参考，请勿外传

图 10：Kimi K3 将开放模型参数规模推升至 2.8 万亿



资料来源：月之暗面

模型能力提升正在打开编程、智能体和知识工作等新应用场景，推理需求增速预计快于训练需求。IDC 与浪潮信息联合报告指出，算法效率提升并未抑制算力需求，更多用户和应用场景加入反而推动大模型普及，并带动数据中心、边缘端和终端算力建设。2025 中国移动云智算大会披露，未来 3 年推理算力年复合增速预计达到训练算力的近 4 倍，到 2028 年推理算力规模将超过训练算力规模。模型能力由简单问答向深度推理、代码工程、工具调用和长周期智能体任务扩展后，算力需求驱动因素将由模型训练次数，进一步扩展至活跃用户数、调用频次、输入输出 Token 数量、推理强度及智能体执行步骤。

3.2. CSP 资本开支维持高位，AI 基础设施建设周期延续

海外头部云厂商及互联网平台资本开支维持高位，AI 基础设施仍处于集中扩张阶段。根据各公司最新披露，Microsoft、Amazon、Alphabet 和 Meta 预计 2026 年资本开支合计达到 7200 亿至 7450 亿美元，资金主要投向 AI 芯片、服务器、存储、网络设备、数据中心及配套电力基础设施。四家公司同时维持高额资本投入，表明 AI 算力扩张并非个别企业的短期行为，而是头部科技公司围绕模型训练、推理服务和云计算竞争形成的行业性投资周期。

近半年大型数据中心项目持续开工和投产，CSP 资本开支正在转化为实际算力供给。Microsoft 位于威斯康星州的 Fairwater AI 数据中心已投入运营，后续设施仍在建设；Amazon 宣布投资 120 亿美元建设路易斯安那州数据中心集群；Meta 在印第安纳州建设投资超过 100 亿美元、规划容量达到 1GW 的数据中心园区；Google 则推动数据中心与新增能源设施共址建设。相关项目投资规模由数十亿美元向百亿美元级提升，单体容量向吉瓦级扩展，建设内容由计算设备采购进一步延伸至电力、网络、散热和水务等配套基础设施。

仅供内部参考，请勿外传

图 11：Meta 规划建设 1GW 级 AI 数据中心园区



资料来源：Meta

算力供给持续受限，叠加云业务和 AI 调用增长，验证资本开支具有实际需求支撑。Microsoft 表示算力供给至少在 2026 年内仍将受到限制，Alphabet 亦计划在自有数据中心容量上线前增加第三方算力使用。与此同时，Azure、AWS 和 Google Cloud 业务继续增长，AI 模型调用量和 Token 消耗持续提升。CSP 通常根据预期需求增长、在手合同和现有容量缺口提前启动算力建设；当前持续存在的容量约束进一步验证新增资本开支具有实际需求支撑。

3.3. AI 算力基础设施市场持续扩容，AI 服务器、超节点与 HPC 协同增长

AI 服务器市场规模持续增长，2026 年产业价值预计超过 3800 亿美元。TrendForce 数据显示，2024 年全球 AI 服务器产业价值约 2050 亿美元，预计 2025 年接近 2980 亿美元，按可比口径计算同比增长约 45.4%，占全球服务器产业价值的比例超过 70%。2026 年全球 AI 服务器产业价值预计同比增长 30%以上，占全球服务器产业价值的比例进一步提升至 74%；以 2025 年接近 2980 亿美元的产业价值为基数测算，2026 年全球 AI 服务器产业价值将超过 3870 亿美元。

整柜式超节点进入规模放量阶段，NVIDIA GB/VR 系列已成为当前可直接观察的主要市场口径。NVIDIA 披露，全球供应链每周约生产 1000 柜 GB300 系统；TrendForce 预计，2026 年 GB300 在 NVIDIAGB/VR 整柜系统出货中的占比接近 80%，VR200 将于 2026 年第三季度末开始爬坡。TrendForce 进一步预计，GB/VR 整柜系统占 NVIDIA AI 服务器出货量的比例将接近 78%，北美五大 CSP 将贡献全球 GB/VR 系列服务器需求的 60%以上，表明超节点已由产品验证进入规模化采购阶段，需求主要集中于头部云厂商，产品结构将由 GB300 逐步向 Vera Rubin 切换。

全球 HPC 服务器市场稳步扩张，2028 年预计达到 222 亿美元。全球高性能计算服务器需求持续增长，科学研究、工业仿真、气象预测、能源勘探及 AI 融合计算共同推动 HPC 基础设施扩容。Hyperion Research 数据显示，2023 年全球本地部署 HPC 技术服务器市场收入为 150 亿美元，预计 2028 年达到 222 亿美元，2023-2028 年复合增长率为 8.2%。Hyperion Research 将单套价格不低于 50 万美元的 HPC 系统划分为超级计算机子市场，预计其 2026 年市场规模达到 95 亿美元，约占全球 HPC 服务器市场的一半。

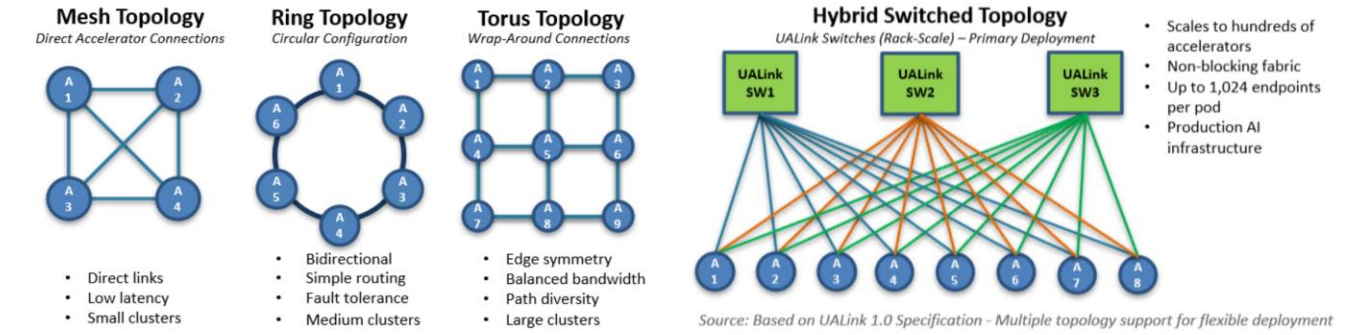
3.4. 算力架构由单机向超节点演进，Scale-Up 互联成为系统升级核心

仅供内部参考，请勿外传

模型规模扩大推动算力架构从芯片堆砌转向系统级协同。中兴通讯《超节点技术白皮书》指出，当模型规模超过单颗芯片的算力与显存容量后，传统分布式计算面临通信开销增加和算力利用率下降等问题。超节点通过高速互联协议和专用交换芯片构建高带宽域，将数十至数百颗加速芯片在逻辑上整合为统一编址、低时延和高带宽的协同计算系统，形成类单机的编程和调度体验。

Scale-Up 互联通过扩大高带宽通信域，将机柜内加速器组织为统一计算单元。UALink 白皮书将 UALink 定义为面向 AI 加速器的 Scale-Up 互联，通过低时延、高带宽连接加速器与交换机，并支持在一个 Scale-Up Pod 内连接最多 1024 个加速器。Scale-Up 由此扩大加速器之间的高带宽通信域，为大模型训练和推理中的高频同步通信提供基础。

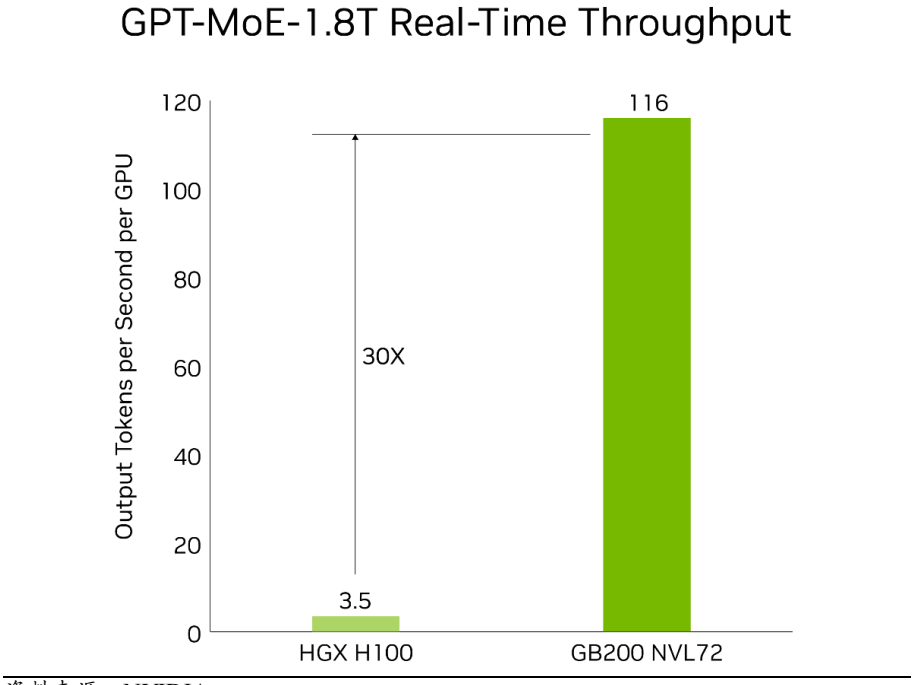
图 12：UALink 支持的互联拓扑



资料来源：UALink Consortium

超节点推动算力竞争由单卡性能转向系统有效算力。NVIDIA GB200 NVL72 将 36 颗 Grace CPU 和 72 颗 Blackwell GPU 连接至同一 NVLink 域，使整套系统能够作为一颗大型 GPU 运行，并提供 130TB/s 的低时延 GPU 通信带宽。NVIDIA 测试显示，GB200 NVL72 在万亿参数模型实时推理和大模型训练中的性能分别达到 H100 平台的 30 倍和 4 倍，体现高带宽互联对并行效率和算力利用率的提升作用。

图 13：GB200 的实时处理能力相比 H100 高出 30 倍



资料来源：NVIDIA

仅供内部参考，请勿外传

超节点提高整机柜集成度，带动交换、连接和液冷系统同步升级。NVIDIA GB200 NVL72 由 18 个计算托盘、9 个 NVLink 交换托盘及连接 GPU 与交换机的铜缆组件构成，计算托盘配置液冷冷板和液冷接口。超节点由此增加对 Scale-Up 交换芯片、交换托盘、高速铜缆、连接器和液冷系统的需求，并推动服务器形态由独立设备向计算、交换和散热一体化的整机柜系统演进。

图 14：GB200 NVL72 通过 NVLink 交换系统构建 72 卡 Scale-Up 域

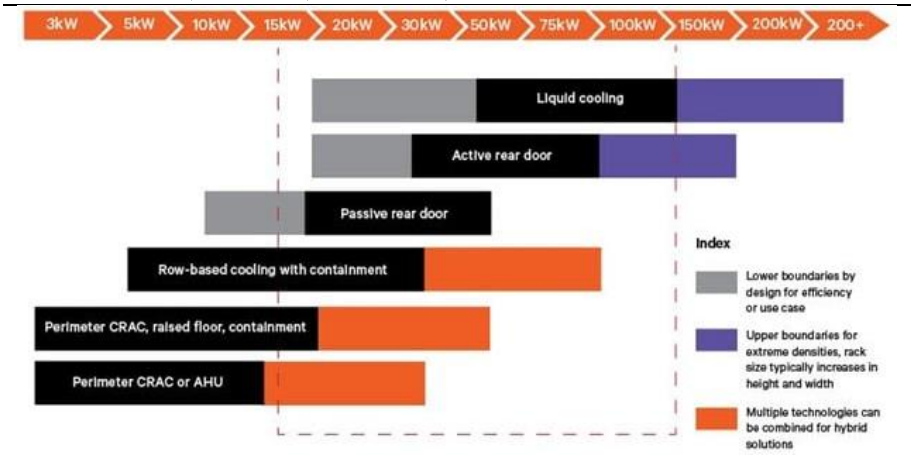


资料来源：NVIDIA

3.5. 算力功率密度持续提升，液冷由可选配置迈向基础设施刚需

机柜功率密度提升推动散热技术由传统风冷逐步向液冷演进。Vertiv《2023 Data Center Liquid Cooling Technology and Market Update》指出，人工智能、机器学习等计算密集型应用正在推动更高密度计算架构发展，芯片厂商也开始推出高功率、仅支持液冷的 CPU 和 GPU。其散热技术适用范围显示，传统周边风冷主要覆盖低功率密度机柜，行级制冷和后门换热器可向中高密度场景延伸，液冷则能够覆盖约 20kW 至 200kW 以上的高密度区间。随着高密度机柜部署增加，数据中心运营商正由局部配置液冷设备转向在每列机柜预留液冷能力，并从整体设施设计角度规划液冷系统。

图 15：机柜功率密度提升推动数据中心散热方式向液冷演进



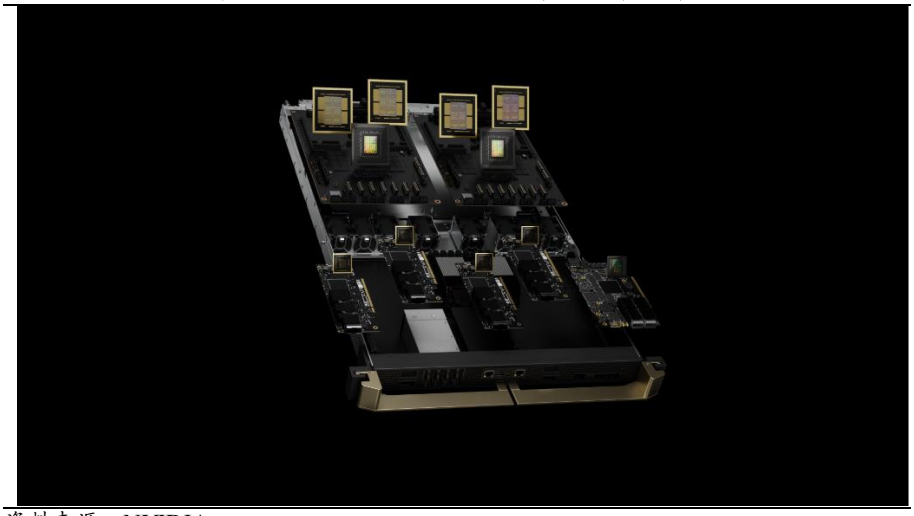
资料来源：Vertiv

海外旗舰 AI 系统已将液冷纳入整机柜标准设计。NVIDIA GB200 NVL72 采用机架级液冷架构，单个计算托盘配置液冷冷板和液冷连接接口，并在一个机柜内集成 18 个计算托盘、72 颗 Blackwell GPU 和 9 个 NVLink 交换托盘。NVIDIA 表示，液冷系统能够支持更高的计算密度和功率效率，液冷已由数

仅供内部参考，请勿外传

据中心的外部配套设备进一步深入服务器和计算托盘内部。

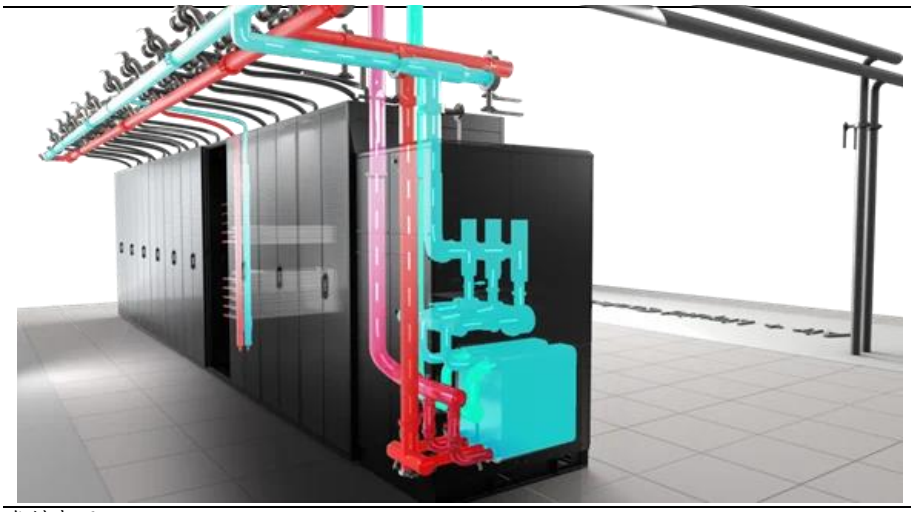
图 16：GB200 计算托盘通过液冷冷板支持高密度算力部署



资料来源：NVIDIA

液冷渗透将带动 CDU、歧管、管路和快速接头等配套基础设施需求。Vertiv 指出，液冷数据中心需要建立独立的二次侧流体网络，由 CDU 控制冷却液温度、流量和洁净度，再通过列级歧管将冷却液分配至多个机柜，并进一步连接服务器内部冷板。液冷系统由此形成“冷源、CDU、歧管、管路、快速接头、冷板”的完整热管理链条，散热系统价值量由传统空调设备向机柜级和芯片级部件扩散。

图 17：液冷二次侧网络连接 CDU、歧管与服务器冷板



资料来源：Vertiv

3.6. 国产算力加快突破，生态适配与系统集成推动规模化落地

政策持续推动国产算力由单点芯片突破向超大规模智算基础设施建设升级。国务院《关于深入实施“人工智能+”行动的意见》提出支持人工智能芯片攻坚创新与使能软件生态培育，加快超大规模智算集群技术突破和工程落地，优化国家智算资源布局；“十五五”规划纲要进一步提出加强高性能高质量智算资源供给、论证建设超大规模智算集群，并推进全国一体化算力网建设和算力设施市场化运营。截至 2026 年 3 月底，全国智能算力总规模达到 188 万 PFLOPS（FP16），国家级监测调度平台已接入 137 万 PFLOPS。国产算力建设由单点芯片和设备研发进一步向软件生态、超大规模集群、算力网络及运营服务体系协同演进。

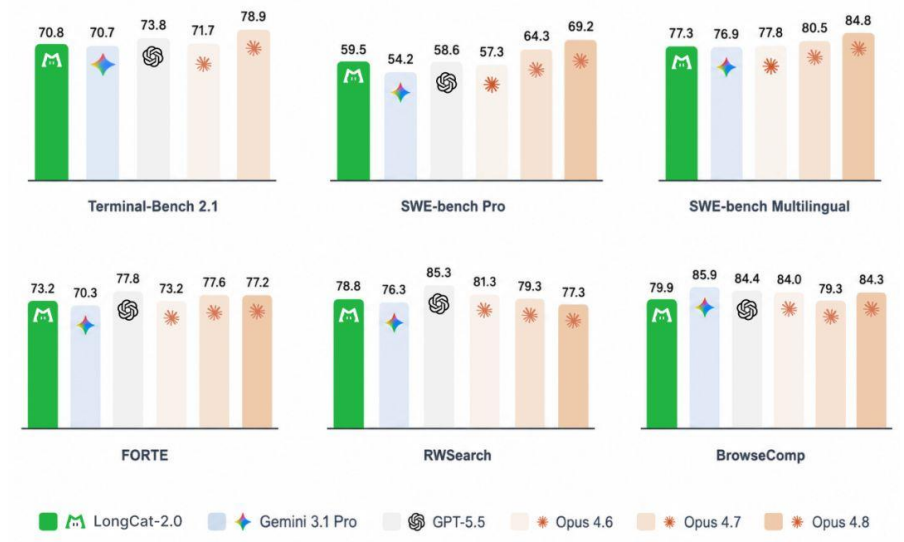
仅供内部参考，请勿外传

国产算力软硬件生态已由单一产品适配向模型、框架和基础软件全栈协同演进。中国信通院《人工智能产业发展研究报告》指出，国内正在推进国产软硬件全栈适配，已实现上千个模型对国产算力的兼容适配，智算生态将加速走向开放协同。主流模型持续适配国产算力，使芯片、服务器、计算框架、编译工具和算子库能够在真实训练与推理任务中协同优化，国产算力的应用范围由测试验证逐步扩大至模型开发、行业部署和规模化服务。

国产算力已由单机和小规模验证迈向万卡级全栈集群部署。2026 年 3 月，深圳建成 14000P 智能算力集群，其中新点亮的 11000P 集群全部采用国产先进芯片，是全国首个万卡级全栈自主可控智算集群，也是全国首个万卡昇腾 910C 超节点智算集群。该项目依托“昇腾+CANN”自主软硬件生态，覆盖芯片、集群、基础软件和开发工具，，验证了国产软硬件体系在万卡级集群中的系统集成和规模部署能力。

LongCat-2.0 进一步验证国产算力支撑万亿参数模型大规模训练与推理的工程能力。2026 年 6 月，美团发布 LongCat-2.0，该模型采用 MoE 架构，总参数量达到 1.6 万亿，并在五万卡国产算力集群上完成从零预训练及后续训练、推理流程，预训练数据规模超过 30 万亿 Token。针对大规模集群运行中的通信异常、节点故障和负载调度等问题，团队通过弹性扩缩卡、自动故障恢复、流水线调度和显存优化等方式，将训练 MFU 提升至原来的 1.5 倍，实现稳态日吞吐超过 1 万亿 Token。五万卡集群承载万亿参数模型的持续训练与真实业务推理，进一步验证国产算力从集群建设向大规模模型生产负载运行延伸的工程能力。

图 18：LongCat-2.0 在编程能力和通用智能体场景表现优异



资料来源：美团

4. 曙光+海光协同构筑国产算力底座，全栈能力向超大规模算力集群延伸

4.1. 曙光+海光贯通芯片与系统，光合生态强化国产算力适配

两光协同构建从核心芯片到算力系统的完整产业链条。海光信息提供国产 CPU 和 DCU，夯实自主计算的芯片底座；中科曙光围绕处理器布局服务器、超节点、高速网络、先进存储、集群管理和基础软件，将芯片能力转化为整机及大规模集群算力；曙光数创进一步提供冷板式、浸没式和相变液冷基础设施。三者贯通核心芯片、计算系统与数据中心基础设施，形成覆盖“芯片、整机、集群、液冷”的国产算力产品体系。

仅供内部参考，请勿外传

中科曙光的系统工程能力充分释放海光芯片在大规模集群中的算力价值。AI 集群的实际性能由处理器性能、卡间通信、存储带宽、软件栈、任务调度、故障管理和散热供电共同决定，公司通过 scaleFabric 高速网络、ParaStor 分布式存储、FlashNexus 全闪存储、资源管理软件和系统集成平台，将分散的计算资源组织为可规模扩展、统一调度和稳定运行的算力系统。公司由服务器设备供应商进一步升级为国产算力集群的系统平台，持续提升在超节点、万卡集群和十万卡超集群中的产品价值量。

光合生态持续扩大国产计算平台的软件适配范围和行业应用边界。围绕海光处理器和中科曙光计算系统，光合生态汇聚操作系统、编译工具、数据库、中间件、模型框架、基础软件及行业应用合作伙伴，推动国产软硬件由单点适配向体系化协同演进。中科曙光依托长期积累的客户资源和系统集成能力，将芯片、基础软件、AI 平台与行业应用组合为完整解决方案，加快国产算力在科研、政务、金融、能源、通信和制造等场景中的部署落地。

图 19：光合组织围绕海光 CPU+DCU 构建全栈软硬件生态



资料来源：光合组织官网

公司进一步向 AI 软件平台和算力运营延伸，强化国产算力由基础设施向模型开发和行业应用的承接能力。公司围绕国产加速卡推出 DeepAI 深算智能引擎 DAP，并通过相关联合测评；2025 年发布 AI 计算开放架构，对“算、存、网、电、冷、管、软”进行一体化设计，并依托曙光 AI 超集群推出 OneScience 科学大模型开发平台，推动国产算力在金融、医疗、科研及公共服务等场景落地。

4.2. ScaleX 超节点与 scaleFabric 高速互联推动计算产品向超集群升级

公司智能计算产品向机柜级超节点和大规模集群系统延伸。公司推出的 scaleX640 通过超高速互联技术将大量高性能计算核心整合在一个机柜内，形成“超级计算单元”，并以此为基础推出 scaleX 万卡超集群系统。公司进一步发布 AI 计算开放架构，在“算、存、网、电、冷、管、软”领域进行一体化紧耦合设计。

仅供内部参考，请勿外传

图 20：曙光 AI 基础设施全景图



资料来源：中科曙光官网

ScaleX640 扩大柜内 Scale-Up 高带宽通信域，推动单机柜计算由节点级服务器向超节点升级。ScaleX640 采用“一拖二”高密架构，单机柜可实现 640 张加速卡高速总线互联，将更多加速卡纳入统一的高带宽、低时延通信域；双 ScaleX640 超节点可进一步组成 1280 卡计算单元，并由液体冷凝换热装置提供最高 1.72MW 散热能力。通过提升单机柜加速卡密度并强化柜内高速互联，ScaleX640 将计算单元由传统节点级 AI 服务器扩展至机柜级 Scale-Up 超节点，为大规模模型训练和推理提供更大的紧耦合计算域。

图 21：中科曙光 ScaleX640 单机柜级 640 卡超节点



资料来源：中关村在线

scaleFabric 承担超节点间 Scale-Out 高速互联，推动算力系统由超节点向万卡、十万卡超集群扩展。在机柜级 Scale-Up 域基础上，公司通过 scaleFabric 原生 RDMA 高速网络连接多个计算单元，实现超集群横向扩展；scaleFabric 可提供 400Gb/s 端口带宽、低于 1 微秒的端侧通信时延，并支持超集群向 10 万卡以上扩展。配套智能调度引擎可管理万级节点、服务十万级用户并支持每秒万级作业调度，结合数字孪生、故障恢复和智能运维能力，公司产品能力由单一计算设备进一步延伸至跨节点高速互联、资源调度和超大规模集群运维。

仅供内部参考，请勿外传

图 22：中科曙光 scaleFabric 国产原生无损 RDMA 高速网络



资料来源：中科曙光官网

4.3. 存储与液冷补齐数据供给与热管理能力，提升单项目价值量

高性能存储与液冷基础设施分别解决大规模算力集群的数据供给和热管理约束。加速卡数量增加后，训练数据读取、模型 Checkpoint 和参数交换对存储吞吐提出更高要求；单机柜功率密度提升则使传统风冷难以满足稳定运行需求。公司在计算产品之外同步布局 ParaStor 分布式存储、FlashNexus 集中式全闪存储及冷板、浸没和相变液冷产品，使计算、数据和基础设施能够在同一集群内协同设计。

存储产品由一般数据保存设备升级为影响集群有效算力的关键组件。ParaStor 主要面向 AI 训练和高性能计算中的海量非结构化数据及并行访问需求，采用横向扩展架构并支持文件、对象、块和 HDFS 等多种协议。FlashNexus 采用端到端 NVMe 架构，主要面向低时延、高并发及高可靠企业级负载。集中式全闪存储在相关测试中可将 GPU 利用率由 30%至 40%提升至 70%至 80%，并将万亿参数模型 Checkpoint 时间由分钟级缩短至秒级，表明数据读取性能能够直接影响计算资源利用率。

图 23：中科曙光存算传紧耦合设计提升计算效率



资料来源：中科曙光官网

液冷产品持续向高功率密度和系统级交付升级。公司已形成冷板式、相变间接式和浸没相变式液冷产品体系，其中 C8000 通过将服务器部件浸没于不导电冷媒中实现全液冷散热，相较传统风冷数据中心节能超过 30%。C8000 产品功率密度持续迭代，早期产品单柜功率为 220kW，根据公司 2025 年半年度报告，当时最新产品单机柜功率密度已突破 750kW。2026 年 4 月，曙光数创发布 C8000 V3.0 相变浸没液冷整机柜及基础设施整体解决方案，最高可支持单机柜功率超过 900kW。公司液冷能力由单一散热设备进一步延伸至供配电、机柜、监控和运维一体化基础设施。

仅供内部参考，请勿外传

图 24：中科曙光 C8000 浸没液冷数据中心基础设施



资料来源：中科曙光官网

4.4. 曙光 8000 十万卡系统落地，验证超大规模集群工程与运营能力

曙光 8000 落地验证公司具备十万卡级超大规模算力系统建设能力。2026 年 7 月，国内首个全国产十万卡 AI 超集群曙光 8000（登峰）正式落成投用，并接入国家超算互联网郑州核心节点，成为国家超算互联网平台上线以来接入的全国最大单体国产 AI 算力资源池，可对外提供超 10 万卡国产 AI 算力。曙光 8000 采用“超智融合”技术路线，支持 FP64 至 INT8 全精度计算，可覆盖科学计算、大模型训练、AI 推理、工业仿真等多类场景；上线首周即实现满载运行，日均处理作业超过 15 万个、单日峰值超过 50 万个，并已完成近千项超智融合应用适配。中科曙光通过计算、网络互联、软件适配和统一调度等系统级能力，将国产计算芯片整合为可规模扩展、统一运营并对外提供服务的超大规模算力基础设施。

第二套十万卡系统启动研制与建设，进一步验证公司超大规模集群能力具备跨项目复用基础。首套曙光 8000 投入使用后，中科曙光与北京科学智能研究院启动第二套全国产十万卡超智融合算力系统，重点面向 AI for Science、大模型训练和高强度科学计算等场景。连续项目落地有助于公司复用计算、互联、存储、资源调度和应用适配等系统能力，并积累不同客户及应用场景下的超大规模集群实施经验。

城市云及曙光智算平台运营经验进一步拓展公司在算力基础设施全生命周期中的服务能力。2009 年投入运营的成都城市云采用服务商投资、建设和运营、政府按需采购服务的模式；2021 年，公司推出公共计算服务平台并由全资子公司曙光智算统一运营，通过连接数据中心算力资源，为客户按需

仅供内部参考，请勿外传

提供计算、应用、数据、调优、运营及运维服务。公司能力不仅覆盖计算设备和集群建设，同时已延伸至资源调度、平台运营及持续运维服务，为承接大型智算基础设施项目提供运营和服务经验。

图25：曙光 8000（登峰）正式落成并接入国家超算互联网



资料来源：中国证券报

4.5. 80 亿元募投推动 Scale-Up 向 1024 卡升级，持续强化集群、训推与存储能力

80 亿元募投聚焦先进算力集群、AI 训推一体机和国产化先进存储三大方向。公司拟投入 35 亿元建设面向人工智能的先进算力集群系统，投入 25 亿元建设下一代高性能 AI 训推一体机，投入 20 亿元建设国产化先进存储系统，分别强化超大规模集群、模型训推部署和 AI 数据基础设施能力。

表5：中科曙光可转债募集资金投向

募投项目	拟投入募集资金（亿元）	占募集资金比例	建设周期	主要建设内容	战略定位
面向人工智能的先进算力集群系统	35.00	43.75%	4 年	研发先进算力超节点、高速互连系统、系统级基础软件栈及异构算力资源管理平台	强化超节点、集群互联与异构算力调度能力，支撑大规模 AI 训练和推理
下一代高性能 AI 训推一体机	25.00	31.25%	3 年	开发 8 卡、16 卡及桌面级液冷工作站，集成国产算力硬件、训推引擎、模型工具链和运维平台	推动公司由 AI 服务器硬件交付向软硬一体化模型部署方案延伸
国产化先进存储系统	20.00	25.00%	3 年	研发国产化全闪存阵列、分布式存储、高速并行文件系统及云原生存储	完善 AI 数据基础设施，提升存储性能、扩展能力及核心部件国产化水平
合计	**80.00**	**100.00%**			

资料来源：中科曙光向不特定对象发行可转换公司债券募集说明书，国泰海通证券研究

先进算力集群项目进一步强化超大规模系统能力。项目拟研发下一代超节点硬件、高速互联系统、系统级基础软件栈和异构算力资源管理运营平台，推动 Scale-Up 域由现有 640 卡向 1024 卡及更大规模扩展，网络端口由 400G 向 800G、1600G 及更高速率升级，并通过 Scale-Out 网络支撑十万卡级线性扩展，为百万卡级集群预留技术空间。

AI 训推一体机项目推动公司由硬件交付向软硬一体化模型部署方案延伸。

仅供内部参考，请勿外传

项目拟开发 8 卡 AI 一体机、16 卡 AI 一体机和桌面级液冷工作站，并配套大模型一站式部署与管理平台、训推加速工具链及一体机服务平台，将国产 CPU、AI 加速卡与模型部署、性能优化和运维工具进行软硬件集成。项目面向金融、医疗、制造、政务和科研等行业的私域模型部署需求，有望降低客户在算力环境搭建、模型部署优化及后续运维等环节的实施复杂度，进一步完善公司从 AI 计算硬件向模型训练、微调和推理部署的产品能力。

图 26：中科曙光 StackCube-K 系列超融合一体机



资料来源：中科曙光官网

国产化先进存储项目进一步提升 AI 集群数据供给能力及核心部件国产化水平。项目拟基于国产 CPU、IO 控制器、网络控制器和国产 PCIe Switch 研发全闪存阵列，并升级分布式存储、高速并行文件存储、智能存储和云原生存储软件。按照募集说明书规划，新一代产品双控制器最大支持 175 块数据盘，控制器内 25 块盘总带宽最高提升 30%，整机工作功率降低 10%以上。项目有望进一步提升公司高性能存储产品的扩展能力、数据吞吐能力和能效水平，并补强大规模 AI 集群的数据基础设施能力。

图 27：中科曙光高端全闪存存储 FlashNexus 9000



资料来源：中科曙光公众号

5. 风险提示

1）技术迭代及研发进展不及预期的风险。计算机及算力基础设施行业技术迭代速度较快，AI 算力、高速互联、液冷散热、存储及基础软件等领域持续演进。若公司未能及时跟踪技术路线变化、丰富技术储备并完成产品迭代，或新产品在性能、可靠性、生态适配及服务能力等方面未能满足客户需求，

仅供内部参考，请勿外传

可能削弱公司现有技术和市场竞争优势，并对经营业绩产生不利影响。

2）关键部件供应链及国产化适配进度不及预期的风险。公司在部分前沿信息技术产品中仍需使用国外先进部件，关键芯片及零部件的供应数量、交付周期和采购价格可能受到国际贸易环境、产业政策及市场供需变化影响。同时，若国产芯片、操作系统及其他基础软硬件的产品性能、规模化供应或适配进度未能满足公司业务发展需要，可能影响公司产品研发、生产交付及成本控制。

3）产业政策支持力度及国际合作环境变化的风险。公司所处的算力基础设施、人工智能和产业数字化行业受国家产业政策、财政投入及相关项目建设节奏影响。若产业政策支持力度、政府及国有企业信息化资本开支低于预期，可能导致下游算力基础设施需求和项目落地速度放缓。此外，若国内外技术标准、贸易规则及产业合作环境发生不利变化，可能限制公司技术应用、产业生态建设及海外市场拓展。

4）超大规模算力集群建设进度及市场需求不及预期的风险。若相关项目建设、交付进度或系统性能未达预期，可能对公司相关产品及系统集成业务产生不利影响。同时，若大模型训练与推理、AI for Science 及行业智能化等下游需求增长不及预期，或客户算力基础设施资本开支放缓，可能导致超大规模算力集群项目数量、建设规模及相关产品需求低于预期，进而对公司经营业绩产生不利影响。

仅供内部参考，请勿外传

本公司是本报告所述海光信息(688041)的做市券商。本报告系本公司分析师根据海光信息(688041)公开信息所做的独立判断。

仅供内部参考，请勿外传

本公司具有中国证监会核准的证券投资咨询业务资格

分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

免责声明

本报告仅供国泰海通证券股份有限公司（以下简称“本公司”）授权客户使用。本公司不会因接收人收到本报告而视其为本公司的当然客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。

本报告的信息来源于已公开的资料，本公司对该等信息的准确性、完整性或可靠性不作任何保证。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、价值及投资收入可升可跌。过往表现不应作为日后的表现依据。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。本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

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。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也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。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

本公司利用信息隔离墙控制内部一个或多个领域、部门或关联机构之间的信息流动。因此，投资者应注意，在法律许可的情况下，本公司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或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的情况下，本公司的员工可能担任本报告所提到的公司的董事。

市场有风险，投资需谨慎。投资者不应将本报告作为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“国泰海通证券研究”，且不得对本报告进行任何有悖原意的引用、删节和修改。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息或进而交易本报告中提及的证券。本报告不构成本公司向该机构之客户提供的投资建议，本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

评级说明

投资建议的比较标准

投资评级分为股票评级和行业评级。

以报告发布后的 12 个月内的市场表现为比较标准，报告发布日后的 12 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数涨跌幅为基准。

	评级	说明
股票投资评级	增持	相对沪深 300 指数涨幅 10%以上
	谨慎增持	相对沪深 300 指数涨幅介于 5%～10%之间
	中性	相对沪深 300 指数涨幅介于-5%～5%
	减持	相对沪深 300 指数下跌 5%以上
	无评级	对于个股未来市场表现与基准指数相比无明确观点
行业投资评级	增持	明显强于沪深 300 指数
	中性	基本与沪深 300 指数持平
	减持	明显弱于沪深 300 指数

国泰海通证券研究所

地址 上海市黄浦区中山南路 888 号 邮编 200011 电话 （021）38676666