



【中国银河证券】DeepSeek 冲击波： 通信算力降本增效，Deepseek 变与不变

通信行业首席分析师：赵良毕

并购家 免费行业报告网

IPOIPO.CN 每日更新 不用注册会员 无广告 永久免费

通信算力降本增效，Deepseek 变与不变

2025 年 02 月 21 日

核心观点

- **DeepSeek 的变：降本增效算法变革，推动算力+应用的共同繁荣：**
DeepSeek 模型通过使用 PTX 编程语言，以及工程能力上的创新变革等，使得其在具有更强的性能的同时，实现更低的成本，或加速推动 AI 应用与硬件的普及落地。将算力分为训练算力及推理算力两部分，我们认为成本更低的推理算力有望推动应用的进一步增长，带动应用端繁荣度提升，AI 能力边际扩张依然需要依赖更强的模型和强大的算力，中长期来看，DeepSeek 在算法和架构上的创新给 AI 发展增加了一条新的道路，有望形成大模型从训练到推理端再到应用端的进一步繁荣共赢局面。

同时，应用端发展有望带动算力产业链加速发展，国产化进程有望加速。DeepSeek 的推出，对推理侧成本带来巨大降低，将显著带来训练测迭代的加速，应用场景落地或将进一步加速，同时反推推理侧模型效率的进一步提升，从而带动通信行业算力相关方向的持续高景气，形成良性循环。同时，更强训练模型的未来需求将带动光模块产业链快速发展，在全球经济形势复杂化趋势下，核心器件光芯片等方向自主可控进程进一步加速，当前国产化 GPU 芯片及相关产业链较同类竞品差距逐步缩小，在具备一定的国产化进程加速基础上，推理端算力的高速发展或将促使国产化器件的普及度进一步加速。
- **DeepSeek 的不变：高端算力资源仍是必须，AIDC 有望实现投入产出的良性循环：**运营商作为央企，基本盘稳健，盈利能力、现金流资产不断改善、资产价值优势凸显，持续增加分红回馈股东，且基础业务具备持续增长性。当前三大运营商云业务持续高增长，在人工智能大背景下，此次 DeepSeek 开源模型推出后，我们认为大模型的重心逐步从预训练转移到了后训练，推理端算力规模有望提升，对于硬件端的要求也有所降低，运营商对算力底座的投资有望与应用端的普及形成良性循环，资本开支的投入逐步转化为规模化提升的回报，协同运营商的通道优势，推动运营商等产业链云业务盈利能力的边际改善。
- **投资建议：**我们认为运营商作为我国最大的流量管道，具备数据优势及接口优势，运营商作为算力行业龙头，利用自身云资源池及硬件的突出能力带动 AI 应用的普及，同时，更强训练模型的未来需求将带动光模块及 AIDC 产业链快速发展，在全球经济形势复杂化趋势下，核心器件光芯片等方向自主可控进程或进一步加速。**建议关注：**运营商中国移动、中国联通、中国电信等；光模块中际旭创、新易盛、天孚通信、光迅科技、华工科技等；光芯片源杰科技、仕佳光子等；AIDC 相关光环新网、数据港等。
- **风险提示：**AI 应用发展不及预期的风险；国际经济形势复杂度进一步提升的风险；AI 硬件发展速度不及预期的风险；AI 产业链上下游短期波动的风险等。

通信行业

推荐 维持评级

分析师

赵良毕

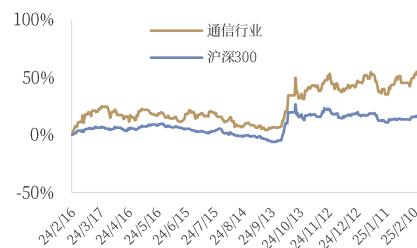
☎：010-8092-7619

✉：zhaoliangbi_yj@chinastock.com.cn

分析师登记编码：S0130522030003

相对沪深 300 表现图

2025-02-21



资料来源：中国银河证券研究院

相关研究

1. 【银河证券】2025 年度策略报告：高成长高景气，科技变革创长牛
2. 【银河证券】科技行业专题报告：大拐点，大机遇，创新变革，拥抱科技新动能
3. 【银河通信】运营商行业 2024 年中报专题：运营商利润增速稳健，数智化转型全球领先
4. 【银河通信】出海专题报告：技术+成本优势驱动，市占率持续提升
5. 【银河通信】中期策略报告：AI 为基算力为石，科技变革浩瀚星辰

目录

Catalog

一、 算法提升带动应用发展，Deepseek 变中取胜.....	4
（一） DeepSeek：从硬件算力竞赛到算法效率革命的拓展.....	4
（二） DeepSeek 引领 AI 成本革命，算法突破促进算力需求正向循环.....	6
（三） deepseek 带来的变与不变：运营商、光模块等板块高景气有望延续.....	8
二、 运营商接入 DeepSeek，第二成长曲线蓬勃发展	9
（一） 国企改革深化，运营商资本开支向智算转型	9
（二） 算力部署持续增加，运营商接入 DeepSeek 降本增效	10
三、 光通信景气度持续，DeepSeek 推动需求高增	13
（一） 高速率光模块需求无虞，硅光加深光模块技术护城河	13
（二） 训练侧光模块持续升级，DeepSeek 提升光模块行业整体景气度	14
（三） 后训练增长有望带动光芯片需求增长，国产化进程加速.....	16
四、 DeepSeek 推动 AIDC 行业加速发展，供需改善.....	18
（一） 我国智能算力崛起，软硬件自主可控能力有望显著提升.....	18
（二） 政策支撑高性能智算发展，DeepSeek 推动 AIDC 上架率及需求双增	20
五、 投资建议：关注运营商、光通信、AIDC 产业链	24
六、 风险提示	26

一、算法提升带动应用发展，Deepseek 变中取胜

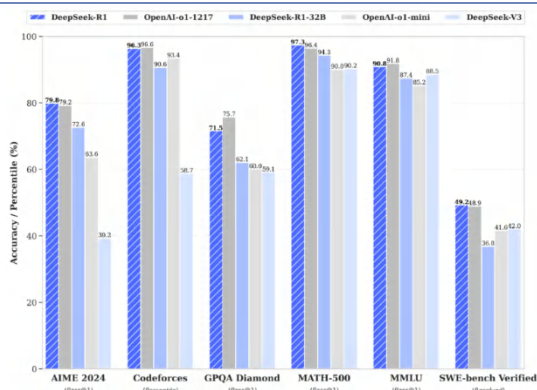
（一）DeepSeek：从硬件算力竞赛到算法效率革命的拓展

大模型在 AI 行业中占据核心地位，是推动技术创新、拓展应用场景及提升行业效率的关键因素。全球范围内的领军企业持续推动大模型性能的提升，随着模型规模的不断扩张，其性能也实现了显著提升。然而，这种规模的扩大也相应地带来了训练和部署成本的急剧增加，成为制约大模型广泛应用的瓶颈。

在机器学习领域，尤其是大型语言模型（LLMs）的应用场景中，模型性能的提升与模型规模、数据集的大小以及计算资源之间存在着紧密的关联，这一关系通常被描述为“规模定律”（Scaling Law）。根据规模定律，模型的性能会随着模型规模的指数级增加而实现线性提升。目前，国际上主流的大模型，诸如 OpenAI 的 GPT 系列、Anthropic 的 Claude 以及谷歌的 Gemini 等，其最新版本规模均已突破千亿参数大关。尽管这些模型在性能上展现出了卓越的表现，但对于众多公司和开发者而言，其高昂的硬件资源使用成本、计算时间等依然构成了巨大的挑战。长期以来，大算力训练一直是基座模型厂商用于融资与构建竞争壁垒的重要手段。

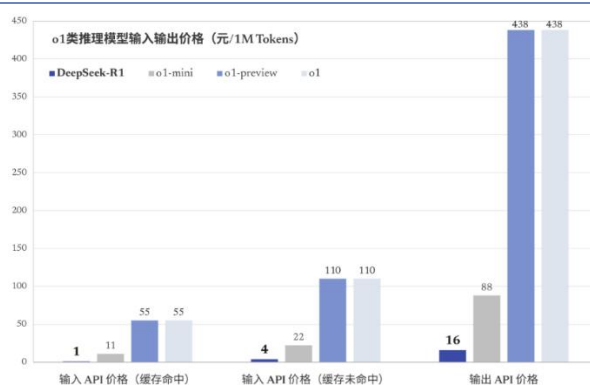
从技术层面来看，GPU 等硬件设施效率的提升以及算法的优化等方式，均有望带动大模型成本的显著下降。在全球 GPU 短缺以及美国限制政策的双重压力下，我国的人工智能公司 DeepSeek 通过算法优化的创新路径，进一步降低了训练成本，为大模型的大规模应用提供了前所未有的可能性。DeepSeek 在 1 月 20 日正式发布了其 R1 模型，并同步开源了模型权重。在第三方的基准测试中，DeepSeek-R1 的表现优于 OpenAI、Meta 和 Anthropic 等美国领先的人工智能公司。在 AIME2024 数学基准测试中，DeepSeek-R1 的成功率高达 79.8%，成功超越了 OpenAI 的 o1 推理模型。在标准化编码测试中，DeepSeek-R1 更是展现出了“专家级”的性能，在 Codeforces 上获得了 2029Elo 的评级，并超越了 96.3% 的人类竞争对手。同时，DeepSeek-R1 真正令人瞩目的地方并不仅在于其卓越的性能，也在于其极低的成本。它打破了硅谷传统的“堆算力、拼资本”的发展路径，仅用 557.6 万美元和 2048 块英伟达 H800 GPU 便完成了性能对标 GPT-4o 的模型训练，成本仅为 OpenAI 同类模型的十分之一，推理成本更是低至每百万 Token 0.14 美元，而 OpenAI 的推理成本则为 7.5 美元。

图1：DeepSeek 性能对齐 OpenAI-o1 正式版



资料来源：DeepSeek 官网，中国银河证券研究院

图2：推理成本低至每百万 Token 0.14 美元



资料来源：DeepSeek 官网，中国银河证券研究院

图3: DeepSeek 蒸馏小模型超越 OpenAI o1-mini

	AIME 2024 pass@1	AIME 2024 cons@64	MATH- 500 pass@1	GPQA Diamond pass@1	LiveCodeBench pass@1	CodeForces rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759.0
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717.0
o1-mini	63.6	80.0	90.0	60.0	53.8	1820.0
QwQ-32B	44.0	60.0	90.6	54.5	41.9	1316.0
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954.0
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189.0
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481.0
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1691.0
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205.0
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1633.0

资料来源: DeepSeek 官网, 中国银河证券研究院

与专有模型不同, DeepSeek-R1 的代码和训练方法均在 MIT 许可下完全开源, 这意味着任何人都可以无限制地获取、使用和修改该模型。全球开发者对 DeepSeek-R1 的贡献代码使其推理效率每小时提升 0.3%, 这一开放性的举措极大地激发了业界的创新活力。DeepSeek-R1 在芯片资源利用、算法复杂性和推理速度上实现了重大突破, 为 AI 行业的发展树立了新的标杆。

DeepSeek 突破的核心在于算法层次和系统软件层次的创新:

1) 首先是算法层次的创新。他们采用了新的 MoE 架构, 使用了共享专家和大量细粒度路由专家的架构。通过将通用知识压缩到共享专家中, 可以减轻路由专家的参数冗余, 提高参数效率; 在保持参数总量不变的前提下, 划分更多的细粒度路由专家, 通过灵活地组合路由专家, 有助于更准确和针对性的进行知识表达。同时, 通过负载均衡的算法设计, 有效地缓解了传统 MoE 模型因负载不均衡带来训练效率低下的问题。

2) 其次在系统软件层次的创新。DeepSeek 采用了大量精细化的系统工程优化。例如, 在并行策略方面, 采用双向流水的并行机制, 通过精细的排布, 挖掘了计算和通信的重叠, 有效的降低了流水并行带来的气泡影响; 在计算方面, 采用 FP8 等混合精度进行计算, 降低计算复杂度; 在通信方面, 采用低精度通信策略以及 token 路由控制等机制有效降低通信开销。

DeepSeek-R1 的成功或许证明, 未来的 AI 竞赛将不再单纯依赖于芯片的纳米级较量, 而是算法效率、生态活力与政策弹性的多维度博弈, AI 行业的发展将呈现出更加多元化和复杂化的变化趋势可期。

图4: DeepSeek-R1 与其他代表性模型在各个维度性能上的对比

Benchmark (Metric)	Claude-3.5- Sonnet-1022	GPT-4o 0513	DeepSeek V3	OpenAI o1-mini	OpenAI o1-1217	DeepSeek R1
Architecture	-	-	MoE	-	-	MoE
# Activated Params	-	-	37B	-	-	37B
# Total Params	-	-	671B	-	-	671B
English	MMLU (Pass@1)	88.3	87.2	88.5	85.2	90.8
	MMLU-Redux (EM)	88.9	88.0	89.1	86.7	92.9
	MMLU-Pro (EM)	78.0	72.6	75.9	80.3	84.0
	DROP (3-shot F1)	88.3	83.7	91.6	83.9	90.2
	IF-Eval (Prompt Strict)	86.5	84.3	86.1	84.8	83.3
	GPQA Diamond (Pass@1)	65.0	49.9	59.1	60.0	75.7
	SimpleQA (Correct)	28.4	38.2	24.9	7.0	47.0
	FRAMES (Acc.)	72.5	80.5	73.3	76.9	82.5
	AlpacaEval2.0 (LC-winrate)	52.0	51.1	70.0	57.8	87.6
	ArenaHard (GPT-4-1106)	85.2	80.4	85.5	92.0	92.3
Code	LiveCodeBench (Pass@1-COT)	38.9	32.9	36.2	53.8	63.4
	Codeforces (Percentile)	20.3	23.6	58.7	93.4	96.6
	Codeforces (Rating)	717	759	1134	1820	2061
	SWE Verified (Resolved)	50.8	38.8	42.0	41.6	48.9
	Aider-Polyglot (Acc.)	45.3	16.0	49.6	32.9	61.7
Math	AIME 2024 (Pass@1)	16.0	9.3	39.2	63.6	79.2
	MATH-500 (Pass@1)	78.3	74.6	90.2	90.0	96.4
	CNMO 2024 (Pass@1)	13.1	10.8	43.2	67.6	78.8
Chinese	CLUEWSC (EM)	85.4	87.9	90.9	89.9	92.8
	C-Eval (EM)	76.7	76.0	86.5	68.9	91.8
	C-SimpleQA (Correct)	55.4	58.7	68.0	40.3	63.7

资料来源: DeepSeek 官网, 中国银河证券研究院

(二) DeepSeek 引领 AI 成本革命, 算法突破促进算力需求正向循环

DeepSeek-R1 以超越美国顶尖模型的性能、更低的研发成本和较弱的芯片性能要求, 引发了美国对其 AI 全球领先地位的担忧, 同时也对科技公司在构建 AI 模型和数据中心上的巨额投入提出了质疑。在 DeepSeek 的冲击下, 国内外大模型厂商紧急上线新模型, 采取降价、免费等措施以证明自身的领先优势。同时, 亚马逊、英伟达、微软等海外 AI 巨头纷纷上线部署用户访问 DeepSeek-R1 模型。2 月 1 日, OpenAI 发布全新推理模型 o3-mini 并首次向免费用户开放。这是 OpenAI 推理模型系列中最新、成本效益最高的模型。在定价方面, o3-mini 每百万 token 的输入(缓存未命中)/输出价格分别为 1.10 美元/4.40 美元, 比完整版 o1 便宜 93%。不过, o3-mini 的性价比或依然不及 DeepSeek。作为对比, DeepSeek 的 API 提供的 R1 模型, 每百万 token 的输入(缓存未命中)/输出价格仅分别为 0.55 美元/2.19 美元。在 o3-mini 推出后, OpenAI CEO 表示, 中国竞争对手 DeepSeek 的崛起削弱了 OpenAI 的技术领先优势, 并就开源与闭源的问题回应称, OpenAI 过去公司曾经开源部分模型, 但主要采用闭源的开发模式, 未来将重新制定开源战略。与此同时, 国产大模型的降价浪潮仍在持续。1 月 30 日, 阿里云发布百炼 qwen-max 系列模型调整通知, qwen-max、qwen-max-2025-01-25、qwen-max-latest 三款模型输入输出价格调整, qwen-max batch 和 cache 同步降价, AI 大模型行业竞争加剧。

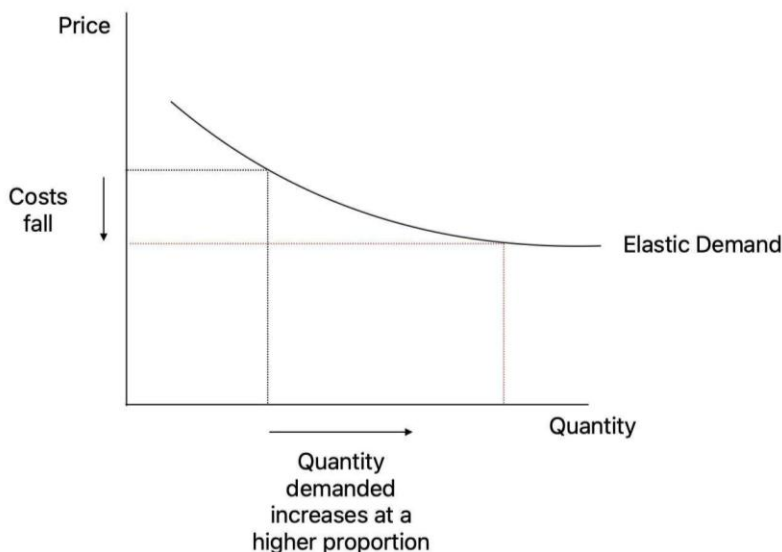
DeepSeek 的成本突破不仅是模型训练的从“硬件驱动”向“算法驱动”的范式拓展, 更为普惠化应用打开了新空间, 反映 AI 技术向实用化、低成本化演进。行业对算力的依赖相较之前发生了“结构性”而非“总量性”变化: DeepSeek 的技术进步短期内或许能够局部缓解算力压力, 但由于算法与算力的“螺旋上升”关系、应用场景的爆发式扩展以及数据增长的不可逆等特性, 我们认为算力资源需求会从预训练端逐渐转移到推理端, DeepSeek 的兴起不会削弱高端芯片需求, 而会促使大模型发展进入“算法进步→模型复杂化→硬件升级”的正向循环。

大模型成本优化与算力需求之间相互成就, 高资源使用效率反而可能会增加算力的总消耗量。DeepSeek 通过降低训练成本, 提高训练效率, 看似减少算力需求, 但同时, 大模型成本缩减意味着降低了企业的训练与推理门槛, 即每单位成本所能提供的训练和推理服务更多了, 算力效率提升有望激活更广泛的用户与应用场景, 从而引发对更大参数以及更复杂的大模型迭代需求。算法优化(如模型压缩、蒸馏)确实能提升单次任务效率, 但 AI 能力的边界扩展(如多模态、复杂推理、通用人工智能)仍依赖更大规模模型和更复杂计算。这可能会对均衡下的算力需求产生正面影响, 整

算力需求不会减少而是更加旺盛，从而形成对硬件需求的新一轮推升，即步入“算法进步→模型复杂化→硬件升级”的正循环。

微软首席执行官引用了“杰文斯悖论”来解释这一现象：Jevons 在《煤炭问题》中发现，随着蒸汽机效率的提升，煤炭消耗量不降反增。其核心观点为：技术进步提高了资源使用效率，效率提高降低了资源使用成本，成本下降刺激了资源需求的增长，需求增长可能超过效率提升带来的节约，最终导致资源总消耗增加。在算力日益成为数字经济“水电煤”的今天，DeepSeek 的技术方向与开源定位，恰恰是算力民主化革命的关键参与者。我们认为，大模型成本优化与算力需求并不是直接的此长彼消关系，而是互相搭台、相互成就，高资源使用效率反而可能增加算力的总消耗量。定价的持续走低有望带来更快的商业化落地，进而会衍生出更多的微调及推理等需求，将逐步盘活全球 AI 应用及算力发展。

图5：“杰文斯悖论”——高资源使用效率反而可能增加总消耗量



资料来源：维基百科，中国银河证券研究院

（三）DeepSeek 带来的变与不变：运营商、光模块等板块高景气有望延续

我们认为 DeepSeek 对通信行业的推动作用主要体现在两方面：

1) **强化国产算力产业链**：为中国 AI 发展带来新机遇，为中美科技竞争增添变数。DeepSeek-R1/V3 支持华为昇腾平台及 MindIE 引擎，自研推理加速引擎使硅基流动与华为云昇腾服务上的模型效果媲美高端 GPU，同时降低成本。这一突破为 AI 生态提供自主多元化方案，助力我国本土芯片厂商商业化落地，促进高效能 AI 的普及。

2) **提升中国 AI 国际影响力**：DeepSeek-R1 的开源实践标志着开源模式对闭源模式的一次重要胜利，这一开源模式对社区的贡献能够迅速转化为整个开源社区的繁荣。中小企业和个人开发者等长尾市场原本因成本限制无法参与的领域（如小规模模型微调、实验性研究）将被激活，形成分散的算力需求增量，产业或将迎来结构性变化，有望为 AI 技术的多元化创新增添更多可能性。同时，AI 加速走向千行百业，智能驾驶、机器人、元宇宙等新兴领域对实时计算和低延迟的高要求，将持续助推高算力需求。DeepSeek 迅速吸引了全球开发者瞩目，曾短时间内即在苹果中国及美国应用商店免费应用下载榜超越 ChatGPT 登顶，彰显了中国 AI 技术以更开放姿态融入全球。未来创新将聚焦于效率、开放性和生产力转化，DeepSeek 或成全球 AI 科技发展转折点。

虽然 DeepSeek 的推出，对降低推理侧成本带来大幅降低，但我们认为推理侧的成本降低，将显著带来训练测迭代的加速，由于推理侧成本的降低，应用场景落地或将进一步加速，推动推理侧模型效率的进一步提升，从而带动通信行业相关方向的持续性高景气。我们维持此前对通信细分子领域运营商、光芯片、光模块的推荐方向，认为 DeepSeek 的推出，运营商作为我国最大的流量管道，具备数据优势及接口优势，AI 应用的普及将持续推进，同时，更强训练模型的未来需求将带动光模块产业链快速发展，在全球经济形势复杂化趋势下，核心器件光芯片等方向自主可控进程或进一步加速。

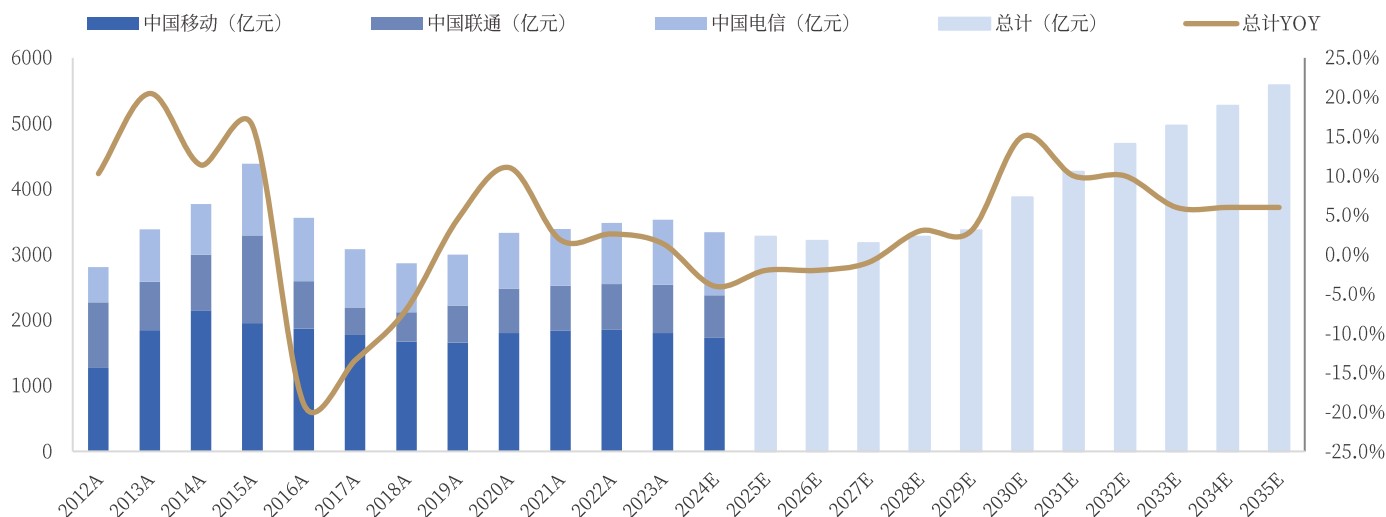
二、运营商接入 DeepSeek，第二成长曲线蓬勃发展

（一）国企改革深化，运营商资本开支向智算转型

国企改革深化，AI 新动能赋能新质生产力。 国资国企是我国新型算力基础设施建设的重要主体，在我国数字经济发展中发挥着“底座”作用。二十届三中全会提出推动国有资本和国有企业做强做优做大，向前瞻性战略性新兴产业集中。立足当下，我国科技发展的大时代已然而至，国央企推动我国经济新旧动能转换有望超预期。展望未来，受益算网基建新变革带来新增量，运营商 AI+ 产业链发挥“链长”作用赋能新质生产力，运营商作为国央企盈利能力逐步提升实现更高质量发展，有望推动我国数字经济 AI 大发展超预期。

运营商资本开支结构向算网倾斜，持续完善 AI 智算能力布局。 中国移动 2024 上半年资本开支为 640 亿元，资本开支占主营业务收入比为 11.71%。公司持续构建技术和规模领先的全国性算力网络，通用算力规模达 8.2 EFLOPS (FP32)，智能算力规模达 19.6 EFLOPS (FP16)。2023 年公司各项资本开支合计约 1,803 亿元，预计 2024 年资本开支约 1,730 亿元，主要用于推进连接感知领先、算力发展提速、能力集约增效和基础拓展布局，以及支撑 CHBN 业务发展等方面。**中国联通** 2024 上半年资本开支为 239 亿元/-13.4%，投资重点由稳基础的联网通信业务转向高增长的算网数智业务。公司 2023 年资本开支为 739 亿元，网络投资显现拐点，着力强化网业协同，突破“算网一体”技术能力，对全国海量算力资源实现统一运营，连接“东数西算”枢纽节点；推动 5G 与边缘云深度融合，实现全国云池毫秒级智能调度；基于大规模数据治理和服务实践，打造超 2 万节点的数据计算引擎，具备万亿级实时处理能力，构建“1+1+M”大模型体系。**中国电信** 2024 上半年资本开支为 472 亿元，其中移动网投资 186 亿元，产业数字化投资 161 亿元。2023 年累计完成资本开支 988 亿元，其中移动网投资 348 亿元，产业数字化投资 355 亿元。公司打造云智、训推一体的算力基础设施，在京津冀、长三角建成国内领先的全液冷万卡池，已承载万亿参数大模型训练任务；在 31 省部署推理池，智能算力达到 21EFLOPS，推动数据中心全面向 AIDC 升级。运营商资本开支实现有效控制，全覆盖智算能力布局不断完善。

图6：运营商资本开支结构化向数通增长，赋能 AI+产业链新发展



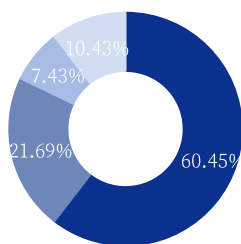
资料来源：中国移动年报、中国电信年报、中国联通年报，中国银河证券研究院

从三大运营商资本开支角度分析，智能算力成为建设重点。 中国移动 2024 年预计资本开支约 1730 亿元，其中连接投入 874 亿元，占比 50.52%/-9.93pct，算力投入 475 亿元/+5.77pct，能力

投入 163 亿元/+1.99pct，基础投资 218 亿元/+2.17pct，其中算力投入增速最快。**中国联通**持续完善云网边一体化布局，建设上海临港、呼和浩特等万卡智算中心，全网智算算力达到 10EFLOPS；布局 20 多个大型算力中心园区，全面覆盖“东数西算”枢纽节点。加快建设“新八纵八横”国家骨干网，在算力中心间部署超高速互联通道，打造一张高通量、高性能、高智能的智能算力智联网络。**中国电信** 2024 年将重点聚焦战新业务，在产业数字化方面的投资占比预计为 38.5%，比 2023 年提升 2.5%，将在云、算力相关投资 180 亿元，全面打造灵活、弹性、绿色的 AIDC，智算能力提升超 10EFLOPS。

图7：2023 年中国移动资本开支占比

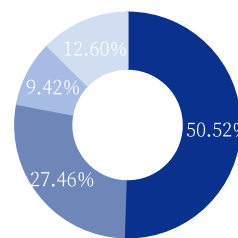
■连接 ■算力 ■能力 ■基础



资料来源：中国移动公告，中国银河证券研究院

图8：2024 年中国移动资本开支占比

■连接 ■算力 ■能力 ■基础



资料来源：中国移动公告，中国银河证券研究院

表1：中国电信及中国联通资本开支结构变化（单位：亿元）

运营商名称	资本开支细分	2020A	2021A	2022A	2023A
中国电信	产业数字化	132	173	272	356
	移动网	482	398	322	348
	宽带网	114	161	187	167
	运营系统和基础设施	120	135	149	118
中国联通	基础设施、传输网及其他	196	200	237	244
	固网宽带及数据	101	103	89	118
	移动网络	379	386	416	377

资料来源：中国电信公告，中国联通公告，中国银河证券研究院

（二）算力部署持续增加，运营商接入 DeepSeek 降本增效

国资委支持中央企业发展人工智能，政策面推动运营商发展智能算力。2024 年 2 月，国务院国资委召开“扎实推动 AI 赋能产业焕新”中央企业人工智能专题推进会，强调加快推动人工智能发展，是国资央企发挥功能使命，抢抓战略机遇，培育新质生产力，推进高质量发展的必然要求。中央企业要主动拥抱人工智能带来的深刻变革，把加快发展新一代人工智能摆在更加突出的位置。2024 年 7 月，国新办举行“推动高质量发展”系列主题新闻发布会，国务院国资委提出未来五年，中央企业预计安排大规模设备更新改造总投资超 3 万亿元，更新部署一批高技术、高效率、高可靠性的先进设备，中央企业要率先行动，推动大规模设备更新重点任务落实落地。2024 年 10 月，国务院国资委《在推进新型工业化中彰显国资央企使命担当》中提到，中央企业是我国新型基础设施建设的重要主体，在我国数字经济发展中发挥着“底座”作用。中央企业智能算力资源是我国算力网络的重要组成部分。下一步将适度超前建设一批智能算力中心，强化人工智能算力供给；同时构

建更加完善的网络安全保障体系，提升数据安全治理能力，筑牢数字经济安全屏障。

表2：电信运营商 2024 算力基础设施建设规划

	中国移动			中国电信			中国联通		
业务规模	2022A	2023A	2024E	2022A	2023A	2024E	2022A	2023A	2024E
云业务收入 (亿元)	503	833	1166	579	972	1361	361	510	714
算力规模&边缘节点	云服务器>71 万台，8.0EFLOPs。边缘计算节点 1000 个。	云服务器>95 万台，18.8EFLOPs。数据中心能力覆盖国家“东数西算”全部枢纽节点	>26EFLOPs	3.8EFLOPs	15.1EFLOPs	>31EFLOPs	170 个城市实现“一市一池”，边缘计算节点达 400 个。	250 个城市实现“一市一池”，边缘计算节点达 600 个。	计划资本开支共计 650 亿元，算网数智投资坚持适度超前，加强布局。
IDC 机架数量	46.8 万架	>50.7 万架，其中境外机架 1.22 万架		51.3 万架	>56 万架		36 万架，数据中心覆盖 23 个省	>40 万架，数据中心覆盖 29 个省	

资料来源：中国移动公告、中国联通公告、中国电信公告，中国银河证券研究院

中国电信：天翼云自研“息壤”智算平台完成国产算力与 DeepSeek-R1/V3 系列大模型的深度适配优化，成为国内首家实现 DeepSeek 模型全栈国产化推理服务落地的运营商机云平台；

中国移动：移动云全面上线 DeepSeek，全版本覆盖、全尺寸适配、全功能畅用。现已全面兼容 DeepSeek 所有主流版本，灵活应对多样化业务场景需求；

中国联通：联通云与 DeepSeek 实现多场景深度对接。联通云已基于“星罗”平台实现多规格 DeepSeek-R1 模型适配，并在全 270 多个骨干云池预部署，接入多种产品场景。

当前三大运营商云业务持续高增长，但在相对较高的基数下增速同比有所放缓，在人工智能大背景下，互联网厂商针对性的高投入以及高研发大幅提升了大模型训练端的发展激烈程度，此次 DeepSeek 开源模型推出后，我们认为大模型的重心逐步从预训练转移到了后训练，推理端算力规模有望提升，对于硬件端的要求也有所降低，运营商对算力底座的投资有望与应用端的普及形成良性循环，资本开支的投入逐步转化为规模化提升的回报，协同运营商的通道优势，形成人工智能应用端的快速发展以及市占率的提升。

表3：电信运营商 AI 大模型部署情况

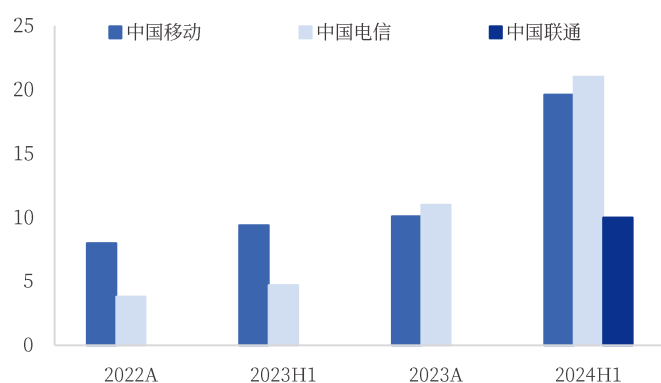
大模型	中国电信	中国移动	中国联通
模型名称	TeleChat	九天	鸿湖
模型类型	大语言模型	行业大模型	多模态大模型
功能特点	使用大量高质量中英文语料进行预训练，采用千万级问答数据进行微；研发了自校准微调技术，提升强化学习效果；利用知识图谱增强大模型的预训练和推理能力	首创“政务大模型-信息场-应用”端到端政务服务体系，具备强大的政务事项理解能力、多维度的信息关联能力	“鸿湖”图文大模型是首个面向运营商增值业务的大模型，可以实现文本生成图像，视频剪辑和图像生成图像等功能
产业链布局	围绕行业应用需求、联合产业链上下注重生态合作伙伴，为各行业提供定制化应用	建设面向通用智能的“人工智能大平台”，为新型人工智能基础设施基于泛在的算网基础设施，实现数据的高质量	面向未来，构建大算力平台底座、行业多场景大模型、高可靠的案例体系

资料来源：中国移动公告、中国联通公告、中国电信公告，中国银河证券研究院

新兴业务收入保持两位数增长，AI 下游应用收入有望进一步加大。根据工信部数据，截止 2024 全年，运营商完成包括云计算、大数据、移动物联网、数据中心等在内的新兴业务收入 4348 亿元，

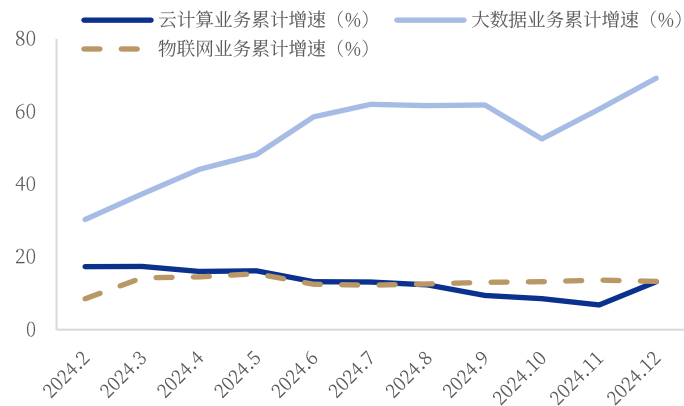
比上年增长 10.6%，在电信业务收入中占比由上年的 21.6% 提升至 25%，拉动电信业务收入增长 2.5pct。其中，云计算、大数据、移动物联网业务收入比上年分别增长 13.2%、69.2% 和 13.3%。伴随“数字中国”AI+建设持续推进、AI+应用方案政策的提出，运营商有望赶上中国数字经济信息发展变革的大时代，市场化程度持续优化，迎来行业景气度的复苏，经营业绩有望更上一层楼。我们认为随着大模型推动 AI 产业趋势进一步加强，国内云厂商对人工智能相关投入有望继续提升，有望带动 AIDC 增速进一步回暖。

图9：运营商智能算力规模增长快速（单位：EFLOPS）



资料来源：中国移动公告、中国联通公告、中国电信公告，中国银河证券研究院

图10：运营商新兴业务分业务收入增速，AI 下游应用有望多点开花



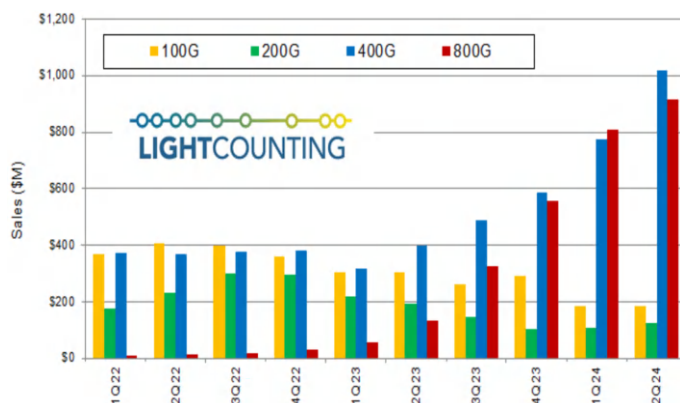
资料来源：Wind，中国银河证券研究院

三、光通信景气度持续，DeepSeek 推动需求高增

（一）高速率光模块需求无虞，硅光加深光模块技术护城河

未来 5 年数通市场的增长驱动力主要来自 400G 以上高速率光模块的需求。全球云计算服务提供商对计算能力和带宽需求的持续增长，以及他们在服务器、交换机和光模块等硬件设备上的资本支出的增加，将推动光模块产品向更高速率的 800G、1.6T 甚至更高端产品的迭代升级。根据我们估算，全球光模块 400G 客户主要集中于亚马逊（约 45%）和谷歌（约 25%）、800G 主要集中于英伟达（约 50%）、谷歌（约 30%）和 Meta（约 20%）等，2025 年 1.6T 光模块的主要需求方预计将是英伟达和谷歌。在 GTC 2024 大会上，英伟达发布了其最新产品 GB200，其服务器与交换机端口速率也实现了翻倍提升，更有望引领 AI 光模块从现有的 800G 向更高性能的 1.6T 升级。此外，英伟达明确了 2026 年将使用 1.6T 网卡，对应 3.2T 光模块需求，明确了光模块升级迭代的节奏。LightCounting 预测，到 2029 年，400G+市场预计将以 28% 以上的复合年增长率（每年约 16 亿美元以上）扩张，达 125 亿美元。其中 800G 和 1.6T 产品的增长尤为强劲，这两个产品共占 400G+ 市场的一半以上。与此同时，200G 以下速率光模块产品的市场规模预计将以每年约 10% 的速度缩减。光模块头部厂商产品的高度可靠性、领先的研发实力及交付能力等优势将进一步凸显，行业集中度有望进一步提高。因此，那些能够与客户同步研发、快速融入客户供应链，并能提前把握客户需求的光模块厂商，将有机会在产品更新换代时抢先获利。

图11：2Q24 400G 和 800G 光模块需求强劲（百万美元）



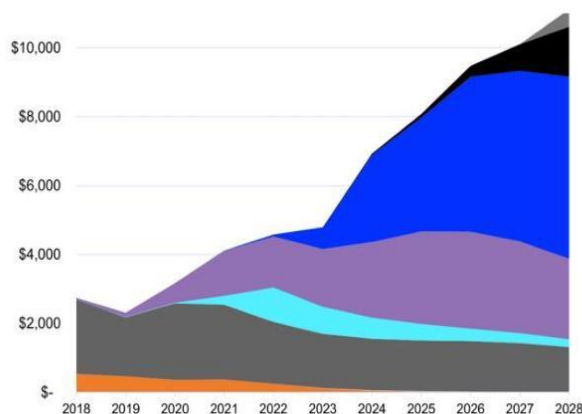
资料来源：LightCounting，中国银河证券研究院

图13：英伟达数据中心芯片产品迭代线路图



资料来源：英伟达，中国银河证券研究院

图12：2018-2028 年全球数通光模块各速率市场空间（百万美元）



资料来源：LightCounting，中际旭创 24 半年报，中国银河证券研究院

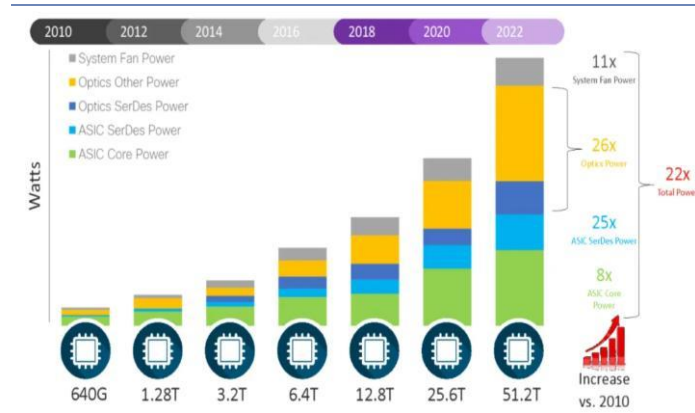
图14：互联速率在过去每四年翻一倍，2023 年开始每两年翻一倍



资料来源：Marvell，中国银河证券研究院

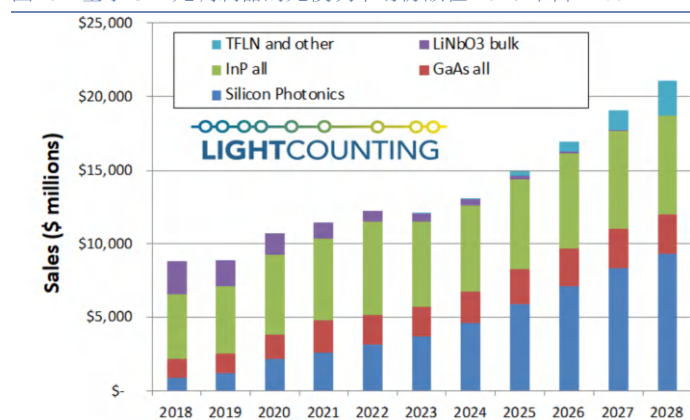
高速光模块的应用导致网络设备功耗大幅增加，硅光等新技术加固护城河。在以 400G 和 800G 光模块为典型配置的 51.2T 和 100T 交换机中，光模块驱动 SerDes 的功耗占比在 40~45%。预计到 2030 年，在 400G+SerDes 和 6.4T 光模块代际时，OSFP 光模块功耗、SerDes 驱动距离将成为很难突破的瓶颈。据统计 2010-2022 年全球数通光模块的整体功耗提升了 26 倍，2024 年 800G 光模块正式放量后该问题更为突出，这种能耗增长对智算中心的运营成本构成了重大压力，降功耗成为光模块技术发展的核心诉求之一。硅光技术利用现有的 CMOS 工艺将光器件与电器件开发和集成到同一个作为光学介质的硅基衬底上，令光电处理深度融合，较传统分立器件更能发扬“光”（高速率、低功耗）与“电”（大规模、高精度）的各自优势。目前由于良率和损耗问题，硅光模块方案的整体优势尚不明显，在功耗、速率、成本、体积四个方面的突破是未来新技术发展的重点方向，也是未来光模块厂商竞争力的体现。根据 LightCounting 的预测，使用基于 SiP 的光模块市场份额将从 2022 年的 24% 增加到 2028 年的 44%，硅光有望凭借硅基产业链的工艺、规模和成本优势迎来产业机遇。

图15：2010-2022 年光模块的整体功耗提升了 26 倍



资料来源：Cisco, Rosenberger, 中国银河证券研究院

图16：基于 SiP 光调制器的光模块市场份额在 2028 年占 44%



资料来源：LightCounting, 中国银河证券研究院

（二）训练侧光模块持续升级，DeepSeek 提升光模块行业整体景气度

LPO 和 CPO 技术在功耗及成本上也各具明显优势，或成未来发展方向之一。LPO（线性驱动）技术通过移除 DSP 降低了光模块的成本和功耗，以 400G 光模块为例，其 7nm DSP 的功耗约 4W，占模块总功耗的一半，而 BOM 成本则占 20-40%，无 DSP 的 LPO 在功耗和成本上更具优势。然而，由于 DSP 的功能不能完全由 TIA 和驱动芯片替代，LPO 可能会增加误码率，进而缩短传输距离。因此 LPO 更适合短距离应用，如数据中心内部服务器与交换机的连接，以及机柜间的连接。而在 CPO（光电共封装）技术中，光学组件被直接封装在交换机芯片旁边，进一步缩短了光信号输入和运算单元之间的电学互连长度，在减少信号损耗问题的同时实现了更低的功耗，还有助于缩小设备体积，使得数据中心的布局更加紧凑。LightCounting 统计，CPO 出货预计将从 800G 和 1.6T 端口开始，并于 2024 至 2025 年开始商用，2026 至 2027 年开始规模上量，CPO 端口在 2027 年 800G 和 1.6T 出货总数中占比预计达约 30%。

表4：传统光模块与 LPO 与 CPO 方案技术优缺点对比

特性	传统光模块	LPO	CPO
功耗	高	较低	低
成本	高	较低	低
时延	高	较低	低
产品成熟度	高	较低	较低

可维护性	好	好	较差
链路性能	好	一般	好
互联性	好	较差	较差

资料来源: Rosenberger, 中国银河证券研究院

在光电子器件方面,随着算力资源的广泛部署及其网络基础设施建设的加速推进,MTP、MPO这类密集连接的典型产品,以其独特的高密度设计显著降低了布线成本,同时增强了系统的可靠性和可维护性,为数据中心的长期发展提供了有力支持,需求展现出快速增长的态势。此外,传输速率的显著提升也驱动了光有源器件光口向多通道方向的快速发展,进而带动了市场对多通道密集连接器件产品的需求增长。在此背景下,研发、制造 MTP、MPO 等高密度光网络关键无源器件的企业将显著受益。比如太辰光是全球最大的密集连接产品制造商之一,其中 MT 插芯及部分无源光器件产品的技术水平在细分行业处于领先地位,公司凭借产品的高性价比优势,有望进一步提升在产业链的市场份额。

目前在数据中心和算力点内部,美国已经完成 400G 光口向 800G 光口的演进,正在向 1.2T、1.6T 推进。我国目前仍然以 400G 光口为主,预计明年 800G 光口成为主流。因此在数据中心、算力点和算力集群之间迫切需要 400G/800G 光传送设备进行承载和传输。比如德科立在长距离光电子器件产品上不断推陈出新,在宽谱放大器、小型化可插拔放大器、高速率长距离相干和非相干光收发模块等领域保持较强的技术优势,有望随数据中心互联互通的建设升级而迎来更加广阔的发展空间。

TSV(硅通孔)技术是硅光芯片封装中的关键技术,其通过在硅片中创建垂直通道实现光芯片与电芯片间的高效电互连,促进了高密度集成和 3D 堆叠,增强了光电混合集成的性能和可靠性,对提升硅光芯片封装技术至关重要。晶方科技作为全球晶圆级芯片尺寸封装服务的主要技术引领者,拥有包括 TSV 在内的多样化先进封装技术,具备 8 英寸、12 英寸晶圆级芯片尺寸封装技术规模量产封装线,有望在提升高端光模块性能方面发挥关键推动作用。

我们认为 DeepSeek 对于光模块行业并非利空而是利好:首先,在推理端算力增加的情况下,低速率光模块出货量具备需求端承载基础,200G、400G 光模块的部署较此前预期有所增长;其次,在需求端强刺激下,训练侧算力模型有望持续迭代,从而反哺应用端日益增长的需求,对高端光模块的刺激有望进一步提升。综合来看,我们认为光模块板块虽然经历了短期的回调,但整体仍旧处于高景气周期内,且对光模块的需求有望更进一步,即中低速率光模块的使用量增加,高速率光模块的需求不变,但迭代周期在应用端强刺激的基础上有望进一步缩减,整体行业规模及景气度有望再上新台阶。

(三) 后训练增长有望带动光芯片需求增长，国产化进程加速

光芯片是光模块核心器件，应用场景较广。光芯片分为激光器芯片及探测器芯片，当前人工智能相关光模块内光芯片主要以 VCSEL 及 EML 芯片为主，高速率光芯片主要以 VCSEL 芯片为主，该款芯片具备线宽窄，功耗低，调制速率高，耦合效率高，传输距离短等特点，主要应用于 500 米内短距离传输，集中在数据中心机柜内布传输，消费电子等领域；EML 芯片则由于其调制频率高，稳定性好，传输距离长等特性，广泛应用于长距离传输，高速率远距离的电信骨干网，城域网和 DCI 等领域，但其具备成本较高的特性，虽然可以部署于短距离传输市场中，但不具备性价比优势。

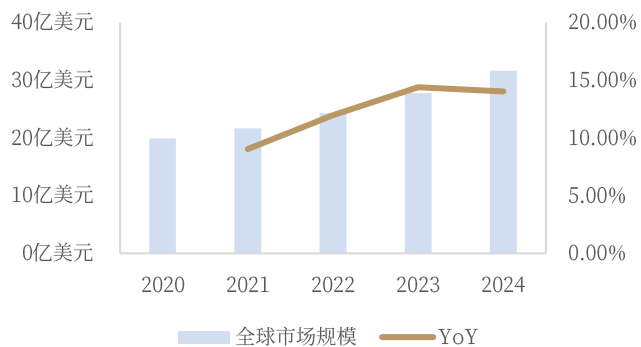
表5：光芯片按功能分类

分类	产品类别	工作波长	产品特性	应用场景
激光器芯片	VCSEL	800-900nm	线宽窄，功耗低，调制速率高，耦合效率高，传输距离短，线性度差	500 米内短距离传输，数据中心机柜内布传输，消费电子等领域
	FP	1310-1550nm	调制速率高，成本低，耦合效率低，先行督察	中速度无线接入短距离市场，部分应用场景逐步被 DFB 激光器芯片取代
	DFB	1270-1610nm	谱线窄，调制速率高，波长稳定，耦合效率低	中长距离传输，如 FTTx 接入网、传输网、无线基站、数据中心内部互联等
	EML	1270-1610nm	调制频率高，稳定性好，传输距离长，成本高	长距离传输，高速率远距离的电信骨干网，城域网和 DCI
探测器芯片	PIN	830-860/1100-1600nm	噪声小，工作电压低，成本低，灵敏度低	中长距离传输
	APO	1270-1610nm	灵敏度高，成本高	长距离单模光纤

资料来源：中商情报网，中国银河证券研究院

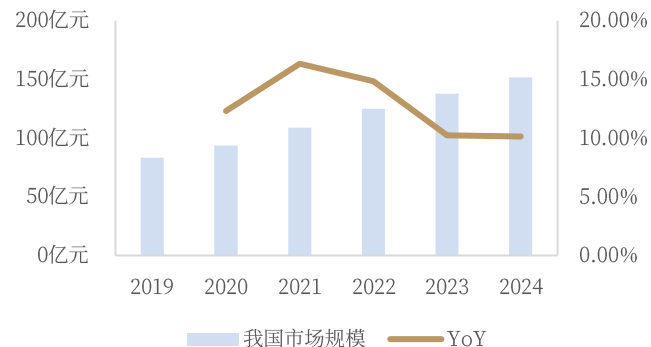
光芯片市场规模持续提升，推理侧算力部署加速及国产化进程加速有望直接带动行业增长。随着云计算、大数据、人工智能等技术的快速发展，对高速、高效、低能耗的数据传输需求日益增长，光芯片的市场需求也随之增加，推动全球光芯片市场规模持续扩大。根据中商产业研究院发布的《2024-2029 年全球及中国光芯片行业发展趋势与投资格局研究报告》显示，2023 年全球光芯片市场规模约 27.8 亿美元，较上年增长 14.4%。中商产业研究院预测，2024 年全球光芯片市场规模将达到 31.7 亿美元。随着国产替代的加速推进，中国光芯片市场规模持续增长，并展现出强劲的发展势头。中商产业研究院发布的《2024-2029 年全球及中国光芯片行业发展趋势与投资格局研究报告》显示，2023 年中国光芯片市场规模约为 137.62 亿元，较上年增长 10.24%。中商产业研究院预测，2024 年中国光芯片市场规模将增长至 151.56 亿元。从国产化率来看，国内相关企业仅在 2.5G 和 10G 光芯片领域实现核心技术的掌握，2.5G 及以下速率光芯片国产化率超过 90%；10G 光芯片国产化率约 60%；25Gbs 及以上的光芯片国产化率低，仅有 4%。预计随着推理侧算力部署的逐步增多，以及后训练算力的规模提升，相对较低速率光芯片市场空间将进一步提升，在我国光芯片相关企业有望受益。

图17：全球光芯片市场规模



资料来源：中商情报网，中国银河证券研究院

图18：我国光芯片市场规模



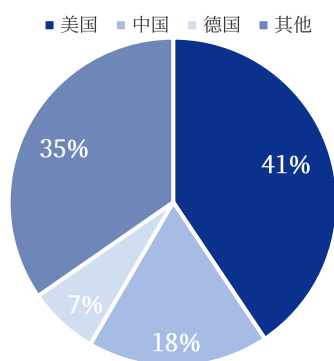
资料来源：中商情报网，中国银河证券研究院

四、DeepSeek 推动 AIDC 行业加速发展，供需改善

（一）我国智能算力崛起，软硬件自主可控能力有望显著提升

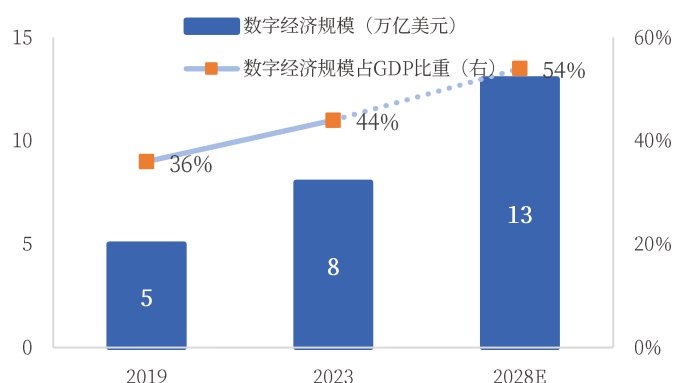
中国数字经济规模稳步增长，核心产业贡献显著。数字经济已成为全球经济发展的强大引擎。在全球版图上，美国、中国和德国以其卓越的数字经济实力，形成三足鼎立的领导格局。受益于大力推动产业数字化转型的国家战略，以及海量数据资源优势，中国数字经济规模增速显著超越美国 and 德国。2019-2023 年，中国数字经济年复合增长率达到约 11.9%，整体规模在 2023 年实现约 56 万亿人民币，同比增长 9.8%，占 GDP 比重约为 44%，其中以云计算、大数据、物联网等为代表的数字技术为主要增长点。

图19：美国、中国和德国在数字经济实力上形成三足鼎立的领导格局



资料来源：Frost&Sullivan，中国银河证券研究院

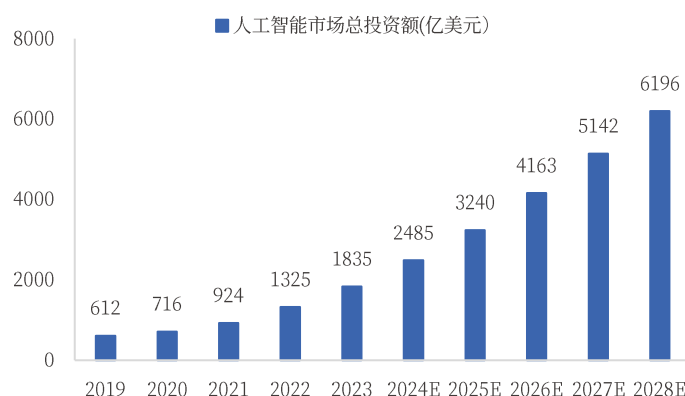
图20：中国数字规模在 2023 年占 GDP 比重约为 44%



资料来源：Frost&Sullivan，中国银河证券研究院

中国人工智能领域投资活跃，硬件领域占总投资的 60%以上。人工智能产业作为数字经济生态中的技术推进器和创新加速器，受到全球企业持续性的技术资金加码。Frost&Sullivan 统计，2023 年全球人工智能市场总投资额达 1835 亿美元，同比增长 38.5%，2023-2028 年全球人工智能市场规模将保持 28% 的年复合增长率，预计到 2028 年总额将超过 6000 亿美元。中国人工智能领域技术投资活跃，投资额在 2023 年突破 193 亿美元，同比增长 48.2%，创下历史新高，投融资主要集中在生活服务，智慧医疗，智能制造和汽车物流领域。2023 年中国人工智能市场总投资额突破 190 亿美元，占全球总量的 10.5%，2019-2023 年的年复合增长率为 43.4%。全球人工智能领域的投资预计将重点聚焦在四个关键业务板块：AI 模型开发企业、AI 基础设施提供商、AI 应用软件开发商，以及企业终端用户。从技术层面进行分析，中国市场的投资主要集中在硬件领域，其占市场总规模的 60% 以上，其次投资主要流向软件领域，投资增速位于技术市场首位。

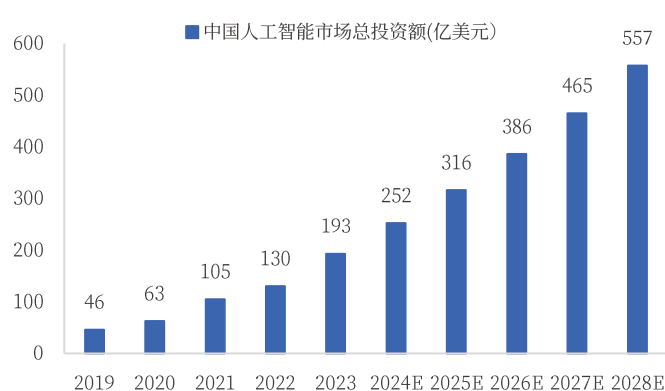
图21: 2023 年全球人工智能市场总投资额达 1835 亿美元



资料来源: Frost&Sullivan, 中国银河证券研究院

注: 总投资规模指企业在包括硬件、软件和服务在内的 AI 市场技术投资总额

图22: 2023 年中国人工智能领域投资额达 193 亿美元

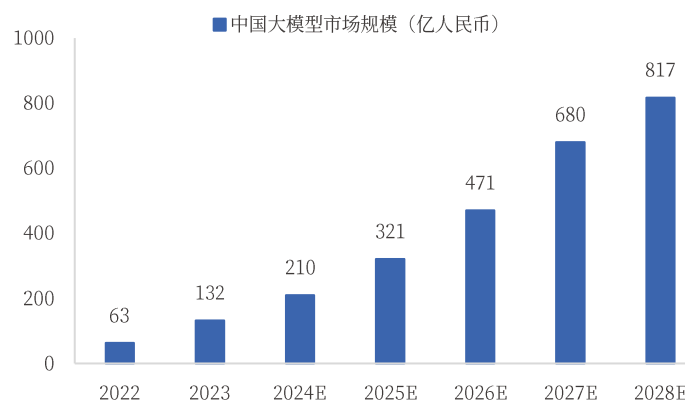


资料来源: Frost&Sullivan, 中国银河证券研究院

自 2020 年起, 全球生成式人工智能应用市场规模经历了显著增长。2022 年被视为生成式 AI 技术的突破性元年, ChatGPT 的推出预示着强人工智能拐点的到来, 激发了新一波的人工智能革新。至 2023 年, 全球该市场规模已膨胀至约 675 亿美元, 实现了 70% 的同比大幅增长, 并预测将于 2028 年跃升至 5160 亿美元, 期间年均复合增长率预计达到约 50.2%。在此市场构成中, 硬件部分占据最大份额, 而基于生成式 AI 的商业服务则展现出最快的增长速度。

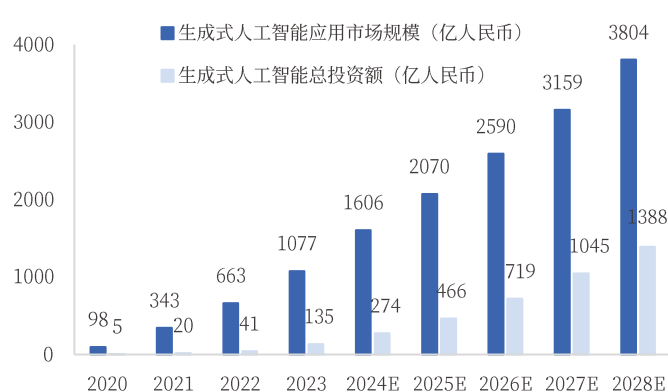
中国生成式 AI 市场得益于上游技术投资的持续增加与下游商业化应用的不断落地, 正展现出强大的增长潜力和市场活力。尽管中国的大模型产业相较于西方企业起步较晚, 但从 2023 年起, 该领域也步入了快速发展阶段。国内众多高科技企业、高等院校、科研机构及创业团队纷纷推出自主研发的大模型, 大模型市场呈现出“百家争鸣”的热烈态势。当前, 领军企业的大模型技术水平已与国际平均水平相持平, 预计至 2028 年, 中国应用端市场规模将突破 3800 亿元人民币。展望未来, 随着大模型技术的不断进步, 其发展趋势将趋向于轻量化小模型、垂直化应用及多功能化拓展。

图23: 预计至 2028 年, 中国大模型市场规模将突破 800 亿元人民币



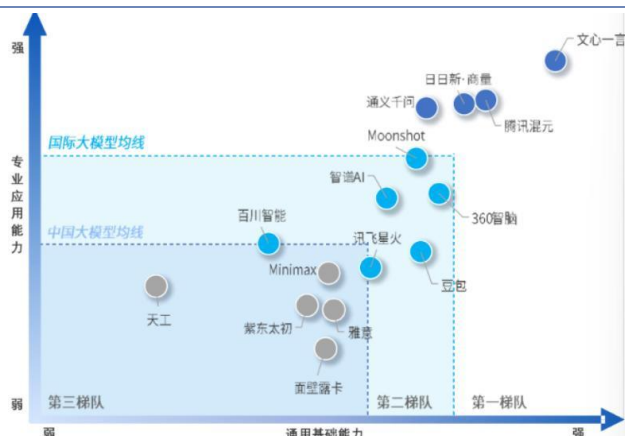
资料来源: Frost&Sullivan, 中国银河证券研究院

图24: 预计至 2028 年, 中国应用端市场规模将突破 3800 亿元人民币



资料来源: IDC, Frost&Sullivan, 中国银河证券研究院

图25：国内领军企业的大模型技术水平已与国际平均水平相持平



资料来源：Frost&Sullivan，头豹研究院，中国银河证券研究院

图26：硬件部分占据最大份额，商业服务则展现出最快的增长速度



资料来源：Frost&Sullivan，中国银河证券研究院

（二）政策支撑高性能智算发展，DeepSeek 推动 AIDC 上架率及需求双增

高性能智算资源是未来大模型产业发展的重要基石。全球头部大模型企业引领和推动技术的革新和落地，商业化应用逐步从文本向图像、音频和视频等领域推进，从而催生了对高性能智能算力的需求，全球主要国家相继出台扶持政策，从资金、基础建设、数据供给、人才、下游应用等多方面谋篇布局。自 2012 年后，驱动 AI 的底层机器学习技术计算量呈指数级增长，根据 OpenAI 论文，深度学习前期，算力翻倍时间为 21.3 个月，深度学习时期，算力翻倍时间为 5.7 个月，大模型时期，AI 训练任务所用的算力每 3.43 个月就会翻倍，远超摩尔定律（晶体管每 18 个月翻一倍）带来的算力提升速度。根据 OpenAI 数据，训练 GPT-3175B 的模型，需要的算力高达 3640PF-days（假如每秒做一千万亿次浮点运算,需要计算 3640 天）。当前大模型的参数量级已达到数千亿参数的量级规模，新增算力中智能算力成为增长新引擎。根据 IDC 与浪潮信息联合发布的《2022-2023 中国人工智能算力发展评估报告》，未来 5 年中国智能算力规模年复合增长率将达到 52.3%，通用算力规模年复合增长率为 18.5%。预计到 2026 年中国智能算力将达到 145EFLOPS，占比将达到 36.7%，渗透率将显著提升，主要用于互联网领域大模型的训练和推理,其占比超过五成。

表6：国家出台一系列政策文件，智算产业被提升至国家战略高度

政策名称	政策主体	政策要点
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	国家发改委等五部门	到 2025 年底，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
算力基础设施高质量发展行动计划	工信部等六部门	到 2025 年算力规模超过 300EFLOPS，智能算力占比达到 35%。推动算力结构多元配置，逐步提升智能算力占比，推动智能算力与通用算力协同，满足不同类型算力业务需求
数字中国建设整体布局规划	国务院	系统优化算力基础设施布局，促进东西部算力高效互补和协同联动，引导通用数据中心超算中心、智能计算中心、边缘数据中心等合理梯次布局
国家发展改革委等部门关于同意京津冀地区启动建设全国一体化算力网络国家枢纽节点的复函	国家发改委等四部门	同意在京津冀地区启动建设全国一体化算力网络国家枢纽节点，发展高密度高能、低碳数据中心集群，通过云网协同、云边协同等优化数据中心供给结构，扩展算力增长空间，实现大规模算力部署
“十四五”国家信息化规划	中央网络安全和信息化委员会	统筹建设面向区块链和人工智能等的算力和算法中心，构建具备周边环境感知能力和反馈回应能力的边缘计算节点提供低时延、高可靠强安全边缘计算服务

“十四五”大数据产业发展规划	工信部	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
----------------	-----	---

资料来源：Frost&Sullivan、中国银河证券研究院

表7：各级政府响应国家号召，政策引导和鼓励覆盖包括基础设施建设、区域协同发展等多个维度

省份/城市	政策名称	政策要点
陕西	陕西省加快推动人工智能产业发展实施方案 (2024-2026 年)	整合省内算力资源，建设省级算力统筹调度平台，实现“算力一网化、统筹一体化、调度一站式”，2026 年建设运营智能算达到 3000P 以上，可统筹的公共智能算力达到西部领先水平
北京	北京市算力基础设施建设实施方案 (2024-2027 年)	改变智算建设“小散”局面，集中建设一批智算单一大集群，到 2025 年，本市智算供给规模达到 45 EFLOPS。到 2027 年，实现智算基础设施软硬件产品全栈自主可控，整体性能达到国内领先水平，具备 100%自主可控智算中心建设能力
上海	上海市智能算力基础设施高质量发展“算力浦江”智算行动实施方案	到 2025 年，新建智算中心国产算力芯片使用占比超过 50%，国产存储使用占比超过 50%，PUE 值将降至 125 以下。到 2025 年，上海市智能算力规模将超过 30 EFLOPS，占比达到总算力的 50%以上，智算中心内先进存储容量占比达到 50%以上
广东	广东省算力基础设施高质量发展行动暨“粤算”行动计划 (2024-2025 年)	到 2025 年，在算力方面，算力规模达到 38 EFLOPS，智能算力占比达到 50%。建成智能计算中心 10 个，基本形成算力规模体量与数字化发展需求相适应、算力供给结构与业务需求相匹配的发展格局
深圳	深圳市算力基础设施高质量发展行动计划 (2024-2025)	到 2025 年，通用算力达到 14 EFLOPS(FP32)，智能算力达到 25 EFLOPS(FP16)，超算算力达到 2 EFLOPS(FP64)，存储量达到 90EB。先进存储容量占比达到 80%以上，重点行业核心数据、重要数据灾备覆盖率达到 100%
山东	山东省数字基础设施建设行动方案 (2024-2025 年)	到 2025 年，全省数据中心在用标准机架总数达到 45 万个，总算力达到 12.5EFLOPS，智能算力占比达到 35%，存力规模达 65EB，先进存储占比达到 35%以上
河南	河南省重大新型基础设施建设提速行动方案 (2023-2025 年)	到 2025 年智算和超算算力规模超过 2000PFLOPS，高性能算力占比超过 30%。持续提升国家超算郑州中心超算能力，建设智算中心和郑州城市算力网调度中心，资源利用率达到 70%
上海	临港新片区加快构建算力产业生态行动方案 (2023-2025 年)	到 2025 年，新片区算力供给形成以智算算力为主、基础算力和超算算力协同的多元算力供给体系，总算力超过 5 EFLOPS(FP32)，AI 算力占比达到 80%，新建数据中心 PUE 控制在 1.25 以内
贵州	面向全国的算力保障基地建设规划	围绕高可靠、高可用目标，从备份中心提升为计算中心、效益中心，重点布局智算基础设施，形成低时延人工智能算力基地、全国低成本中心、高安全中心，到 2024 年通用算力、智算算力、超算算力的总规模达到 5EFLOPS

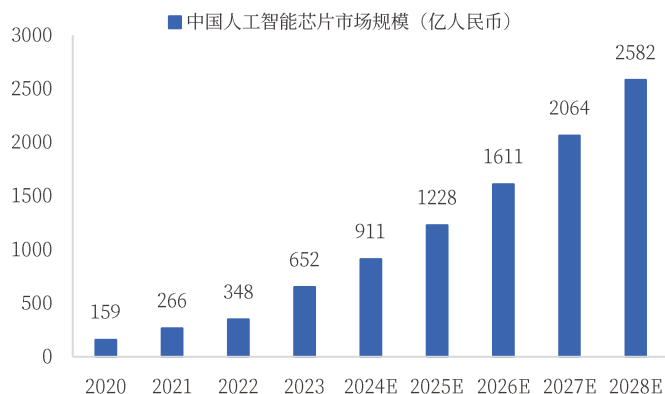
资料来源：Frost&Sullivan、中国银河证券研究院

在智能算力产业的上游，尽管中国已在基础设施和部分核心设备领域实现了国产化替代，但在人工智能芯片领域，现阶段市场主导权仍由海外厂商掌握，这对我国智能算力供应的稳定性构成潜在隐患。作为算力产业的基础构成部分，人工智能芯片对于支撑智能算法及应用所需的计算能力至关重要，其设计与制造水平直接关联到中下游智算资源的有效供给与利用，掌握自主可控的人工智能芯片技术是保障智算产业持续健康发展的基石。鉴于美国对中国半导体行业的监管不断强化，加之全球经济增速放缓，随着芯片产业链结构的持续优化，如芯片设计领域的快速增长及封装环节占比的逐渐降低，预计中国芯片市场将在未来恢复增长，至 2028 年市场规模有望达到 1.5 万亿元人民币，其中人工智能芯片市场规模预计将达 2582 亿元人民币。

人工智能服务器作为数据中心的硬件设备及算力的重要支撑载体，2023 年中国市场规模达到约 692 亿人民币，其中 GPU 加速服务器占比超过 90%，这一显著增长反映了全球人工智能企业

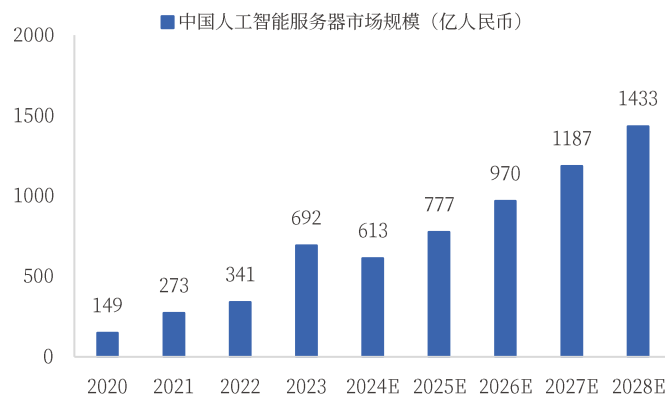
正加速布局 AI 及智算业务资源。在这一市场中,浪潮、新华三等企业在 2023 年合计提供了超过 50% 的人工智能服务器,特别是基于 GPU、ASIC 等加速芯片的 AI 服务器需求增长最为明显。展望未来,随着人工智能应用的日益成熟,市场对人工智能服务器的需求将持续扩大。

图27: 2028 年人工智能芯片市场规模预计将达 2582 亿元人民币



资料来源: Frost&Sullivan, 中国银河证券研究院

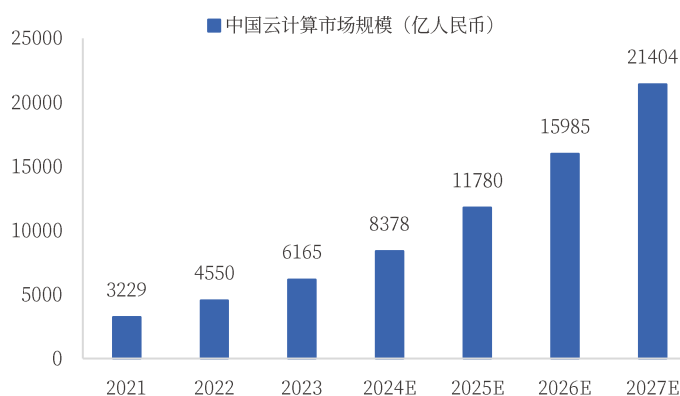
图28: 2023 年中国 AI 服务器市场规模首次达到约 692 亿元人民币



资料来源: Frost&Sullivan, 中国银河证券研究院

在智能算力产业的中游环节, 数据中心为智算资源供给方搭建和运维物理平台, 并为其提供海量的数据存储, 高速的网络传输和实时处理服务, 为配合高性能智算资源的供给, 传统数据中心在架构、性能、可扩展性和安全性等方面都进行了升级改造。2023 年, 我国政府印发《数字中国建设整体布局规划》, 明确提出要夯实数字基础设施和数据资源体系, 预计未来我国数据中心机架数量将持续增长, 在 2025 年达到 1400 万台。三大电信运营商作为国内数据中心建设的主导力量, 依托其在政企市场的深厚渠道优势, 正积极整合其在互联网数据中心和网络资源方面的领先优势, 加速布局数字基础设施资源以满足日益增长的数据处理和存储需求。云服务商是目前智算市场上的主要智算资源供给方, 凭借其自身在云计算方面全面的技术能力和完善的产品矩阵, 为下游客户提供算力资源和算法增值服务。2023 年中国云计算市场规模超 6000 亿人民币, 其中公有云服务占比 75.4%。随着 AI 技术革新和未来大模型的应用落地, 我国云服务市场即将开始新一轮增长, 预计到 2027 年市场规模将达到 21404 亿人民币。

图29: 预计到 2027 年, 中国云计算市场规模将达 21404 亿人民币



资料来源: Frost&Sullivan, 中国信息通信研究院, 中国银河证券研究院

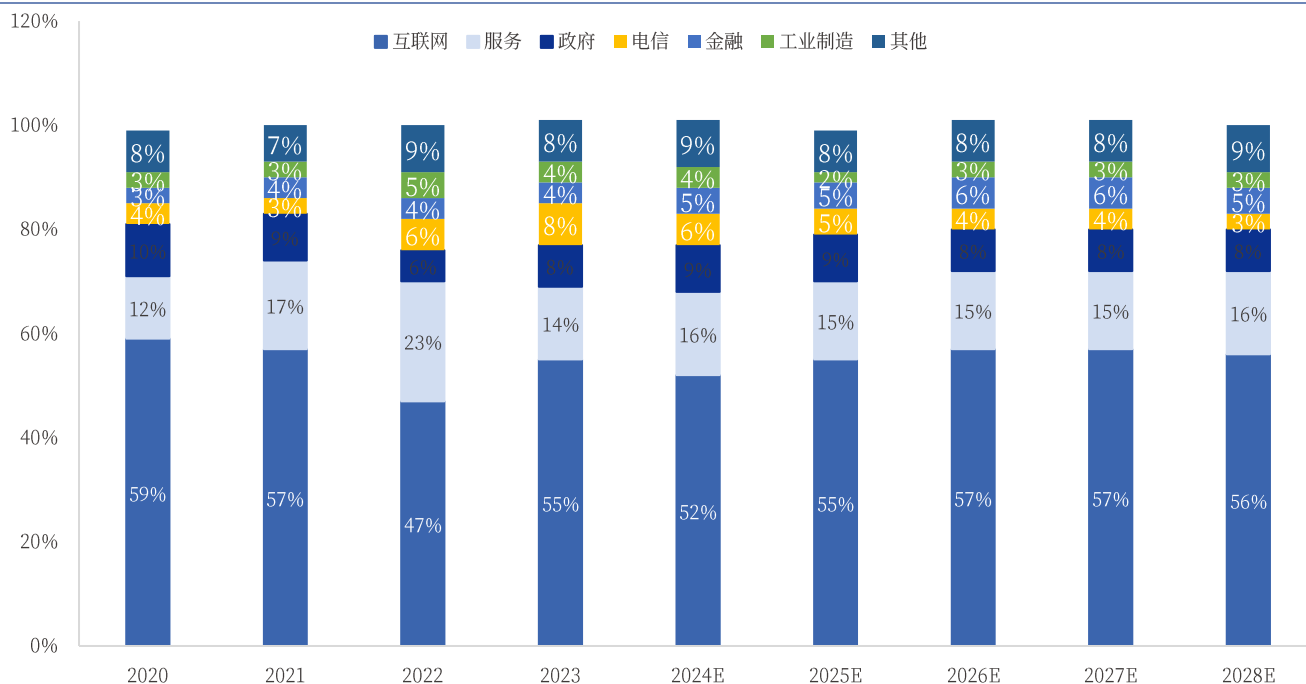
图30: 2023 年公有云服务占云计算总市场规模比重为 75.4%



资料来源: Frost&Sullivan, 中国信息通信研究院, 中国银河证券研究院

在智能算力产业的下游环节, 互联网一直处于大模型训练和推理领域的第一梯队, 占据每年人工智能加速服务器市场的半壁江山, 随着垂直行业大模型的逐步渗透和商业化落地, 金融、能源、交通运输、教育、政府和智能制造等板块对智算资源的需求将逐渐凸显。

图31：中国加速服务器出货量按下游行业拆分



资料来源：Frost&Sullivan, IDC, 中国银河证券研究院

总结来看，中国智算产业的发展受到以下三点关键因素的驱动：1) 政策支持：我国通过发布《十四五大数据产业发展规划》等政策文件，对智能算力产业进行了宏观规划，并启动了“东数西算”工程，以促进智算中心建设和大模型研发，满足市场需求。地方政府也积极出台政策，推进智算基础设施建设，提供普惠算力服务，降低中小企业和个人使用算力的门槛；**2) AI 下游应用需求增长：**人工智能技术在金融、电信、工业制造等多个领域的应用不断深化，推动产业智能化和数字化转型，同时在元宇宙、人形机器人等新兴领域的需求不断涌现，带动智能算力需求激增；**3) 硬件升级需求：**算力技术迭代创新推动了硬件的升级，国内算力基础设施建设蓄势待发。

我们认为 DeepSeek 的部署，有望提升 AIDC 的现有机房上架率，同时刺激新机房的建设需求。我们可以发现 DeepSeek 的推出，对于算力的需求并非缩减而是持续增加，当前算力规模并不足以支撑 DeepSeek 全时间段大规模运行，算力的强需求将带动现有 AIDC 的上架率提升，以及新机房建设需求也成为应用端发展的确定性方向。

五、投资建议：关注运营商、光通信、AIDC 产业链

运营商：低估值高成长，国家政策支持下算力新业务发展有望超预期。

运营商盈利能力、现金流资产不断改善、资产价值优势凸显，持续增加分红回馈股东，相对历史估值和国外水平，通信运营商均处于估值低位。总体来说，运营商业绩持续增长或超预期，5G“收获期”大有可为。当前运营商云业务发展如火如荼，DeepSeek 对于成本端的降低有望协同运营商云业务部署以及运营商的海量数据资产，推动运营商第二曲线的快速增长。

建议关注：中国移动、中国联通、中国电信等。

光通信：技术迭代产品量价齐升，技术壁垒增强竞争格局有望边际改善。

AIGC 引领新一轮科技革命，DeepSeek 对于成本端的降低或将推动应用端的繁荣，继而反哺推理侧模型的快速迭代，推动应用端的进一步发展。光模块 100G/200G→400G→800G→1/6T 迭代速率持续提升，带来产品量价齐升有望延续，带来业绩高增持续可期。同时国内的算力部署有望推动国内光模块产业链景气度提升，带来较强的规模效应。

建议关注：中际旭创、新易盛、天孚通信、光迅科技、华工科技等。

光芯片产业链国产化进程加速，复杂国际经济形势下渗透率有望进一步提升。

DeepSeek 对于成本端及训练精度的降低或将使得推理侧对光芯片的技术需求边际改变，国产光芯片在推理侧算力部署中具备成本优势且技术可靠性较强，产业链渗透率有望跟随推理侧算力部署的增加而有所上升，在复杂国际形势下，国产光芯片在推理侧部署的可靠性及成本优势有望将进一步提升自身在采购链条中的话语权和份额。

建议关注：源杰科技、仕佳光子等。

DeepSeek 推动 AIDC 行业加速发展，供需改善。

DeepSeek 的推出，对于算力的需求并非缩减而是持续增加，当前算力规模并不足以支撑 DeepSeek 全时间段大规模运行，算力的强需求将带动现有 AIDC 的上架率提升，以及新机房建设需求也成为应用端发展的确定性方向。

建议关注：AIDC 相关光环新网、数据港等。

表8：建议关注相关标的盈利预测情况-（截至 2025 年 2 月 21 日）

代码	标的名称	EPS（元）			P/E		
		2024E	2025E	2026E	2024E	2025E	2026E
600941.SH	中国移动	6.47	6.81	7.20	17.70	16.81	15.90
600050.SH	中国联通	0.29	0.32	0.35	23.88	21.62	19.65
601728.SH	中国电信	0.36	0.39	0.41	23.98	22.29	20.84
300308.SZ	中际旭创	4.74	7.89	9.76	25.07	15.05	12.16
300502.SZ	新易盛	3.51	6.29	8.23	34.53	19.25	14.71
300394.SZ	天孚通信	2.58	4.16	5.46	40.35	25.01	19.08
002281.SZ	光迅科技	0.98	1.38	1.74	59.06	41.73	33.16
000988.SZ	华工科技	1.33	1.71	2.17	37.07	28.78	22.70
688498.SH	源杰科技	0.54	1.62	2.46	269.11	89.91	59.33
688313.SH	仕佳光子	0.13	0.26	0.41	169.77	82.84	53.54
603881.SH	数据港	0.25	0.32	0.37	163.05	126.20	108.61
300383.SZ	光环新网	0.30	0.40	0.48	68.68	51.80	43.00

资料来源：Wind 一致预期、中国银河证券研究院

六、风险提示

AI 应用发展不及预期的风险：当前 AI 应用正在快速发展，行业正处于全面拥抱初期，若当前 AI 应用发展不及预期，则有可能降低应用推广速度，从而使得硬件发展及软件发展脚步降低；

国际经济形势复杂度进一步提升的风险：当前国际形势较为复杂，全球各地地缘政治博弈较大，一定程度上影响了经济的进一步发展，若国际经济形势复杂度进一步提升，则有可能增加出口占比较大的公司的不确定性；

AI 硬件发展速度不及预期的风险：虽然 DeepSeek 算法创新，但算法的创新并不能代替硬件的迭代，若 AI 硬件发展不及预期，则人工智能发展可能迎来瓶颈；

AI 产业链上下游短期波动的风险：在 DeepSeek 算法及模型开源的影响下，应用及 AI 硬件需求度均有所提升，短期内可能对时效性要求较高，促使产业扩大生产规模，由于 DeepSeek 带来的产业格局变化，新生的推理侧及应用相关企业较多，需求提升后上下游产能准备可能一定程度上不及预期，造成产业链上下游短期波动风险。

图表目录

图 1: DeepSeek 性能对齐 OpenAI-o1 正式版.....	4
图 2: 推理成本低至每百万 Token 0.14 美元.....	4
图 3: DeepSeek 蒸馏小模型超越 OpenAI o1-mini.....	5
图 4: DeepSeek-R1 与其他代表性模型在各个维度性能上的对比	6
图 5: “杰文斯悖论”——高资源使用效率反而可能增加总消耗量	7
图 6: 运营商资本开支结构化向数通增长, 赋能 AI+产业链新发展	9
图 7: 2023 年中国移动资本开支占比.....	10
图 8: 2024 年中国移动资本开支占比.....	10
图 9: 运营商智能算力规模增长快速 (单位: EFLOPS)	12
图 10: 运营商新兴业务分业务收入增速, AI 下游应用有望多点开花	12
图 11: 2Q24 400G 和 800G 光模块需求强劲 (百万美元)	13
图 12: 2018-2028 年全球数通光模块各速率市场空间 (百万美元)	13
图 13: 英伟达数据中心芯片产品迭代线路图	13
图 14: 互联速率在过去每四年翻一倍, 2023 年开始每两年翻一倍	13
图 15: 2010-2022 年光模块的整体功耗提升了 26 倍	14
图 16: 基于 SiP 光调制器的光模块市场份额在 2028 年占 44%.....	14
图 17: 全球光芯片市场规模.....	17
图 18: 我国光芯片市场规模.....	17
图 19: 美国、中国和德国在数字经济实力上形成三足鼎立的领导格局	18
图 20: 中国数字规模在 2023 年占 GDP 比重约为 44%	18
图 21: 2023 年全球人工智能市场总投资额达 1835 亿美元	19
图 22: 2023 年中国人工智能领域投资额达 193 亿美元	19
图 23: 预计至 2028 年, 中国大模型市场规模将突破 800 亿元人民币	19
图 24: 预计至 2028 年, 中国应用端市场规模将突破 3800 亿元人民币	19
图 25: 国内领军企业的大模型技术水平已与国际平均水平相持平	20
图 26: 硬件部分占据最大份额, 商业服务则展现出最快的增长速度	20
图 27: 2028 年人工智能芯片市场规模预计将达 2582 亿元人民币	22
图 28: 2023 年中国 AI 服务器市场规模首次达到约 692 亿元人民币	22
图 29: 预计到 2027 年, 中国云计算市场规模将达 21404 亿人民币	22
图 30: 2023 年公有云服务占云计算总市场规模比重为 75.4%	22
图 31: 中国加速服务器出货量按下游行业拆分	23

表 1： 中国电信及中国联通资本开支结构变化（单位：亿元） 10

表 2： 电信运营商 2024 算力基础设施建设规划..... 11

表 3： 电信运营商 AI 大模型部署情况 11

表 4： 传统光模块与 LPO 与 CPO 方案技术优缺点对比 14

表 5： 光芯片按功能分类..... 16

表 6： 国家出台一系列政策文件， 智算产业被提升至国家战略高度 20

表 7： 各级政府响应国家号召， 政策引导和鼓励覆盖包括基础设施建设、区域协同发展等多个维度..... 21

表 8： 建议关注相关标的盈利预测情况-（截至 2025 年 2 月 21 日） 25

分析师承诺及简介

本人承诺以勤勉的执业态度，独立、客观地出具本报告，本报告清晰准确地反映本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告的具体推荐或观点直接或间接相关。

赵良毕，通信&中小盘首席分析师，科技组组长。北京邮电大学通信硕士，复合学科背景，2022 年加入中国银河证券。8 年中国移动通信产业研究经验，6 年证券从业经验。曾获得 2018/2019 年（机构投资者 II-财新）通信行业最佳分析师前三名，2020 年获得 Wind（万得）金牌通信分析师前五名，获得 2022 年 Choice（东方财富网）通信行业最佳分析师前三名。

免责声明

本报告由中国银河证券股份有限公司（以下简称银河证券）向其客户提供。银河证券无需因接收人收到本报告而视其为客户。若您并非银河证券客户中的专业投资者，为保证服务质量、控制投资风险、应首先联系银河证券机构销售部门或客户经理，完成投资者适当性匹配，并充分了解该项服务的性质、特点、使用的注意事项以及若不当使用可能带来的风险或损失。

本报告所载的全部内容只提供给客户做参考之用，并不构成对客户的投资咨询建议，并非作为买卖、认购证券或其它金融工具的邀请或保证。客户不应单纯依靠本报告而取代自我独立判断。银河证券认为本报告资料来源是可靠的，所载内容及观点客观公正，但不担保其准确性或完整性。本报告所载内容反映的是银河证券在最初发表本报告日期当日的判断，银河证券可发出其它与本报告所载内容不一致或有不同结论的报告，但银河证券没有义务和责任去及时更新本报告涉及的内容并通知客户。银河证券不对因客户使用本报告而导致的损失负任何责任。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的银河证券网站以外的地址或超级链接，银河证券不对其内容负责。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。

银河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。银河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。

银河证券已具备中国证监会批复的证券投资咨询业务资格。除非另有说明，所有本报告的版权属于银河证券。未经银河证券书面授权许可，任何机构或个人不得以任何形式转发、转载、翻版或传播本报告。特提醒公众投资者慎重使用未经授权刊载或者转发的本公司证券研究报告。

本报告版权归银河证券所有并保留最终解释权。

评级标准

评级标准		评级	说明
评级标准为报告发布日后的 6 到 12 个月行业指数（或公司股价）相对市场表现，其中：A 股市场以沪深 300 指数为基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准，北交所市场以北证 50 指数为基准，香港市场以恒生指数为基准。	行业评级	推荐：	相对基准指数涨幅 10%以上
		中性：	相对基准指数涨幅在-5%~10%之间
		回避：	相对基准指数跌幅 5%以上
	公司评级	推荐：	相对基准指数涨幅 20%以上
		谨慎推荐：	相对基准指数涨幅在 5%~20%之间
		中性：	相对基准指数涨幅在-5%~5%之间
	回避：	相对基准指数跌幅 5%以上	

联系

中国银河证券股份有限公司 研究院

深圳市福田区金田路 3088 号中洲大厦 20 层

上海浦东新区富城路 99 号震旦大厦 31 层

北京市丰台区西营街 8 号院 1 号楼青海金融大厦

公司网址：www.chinastock.com.cn

机构请致电：

深广地区：程曦 0755-83471683 chengxi_yj@chinastock.com.cn

苏一耘 0755-83479312 suyiyun_yj@chinastock.com.cn

上海地区：陆韵如 021-60387901 luyunru_yj@chinastock.com.cn

李洋洋 021-20252671 liyangyang_yj@chinastock.com.cn

北京地区：田薇 010-80927721 tianwei@chinastock.com.cn

褚颖 010-80927755 chuying_yj@chinastock.com.cn