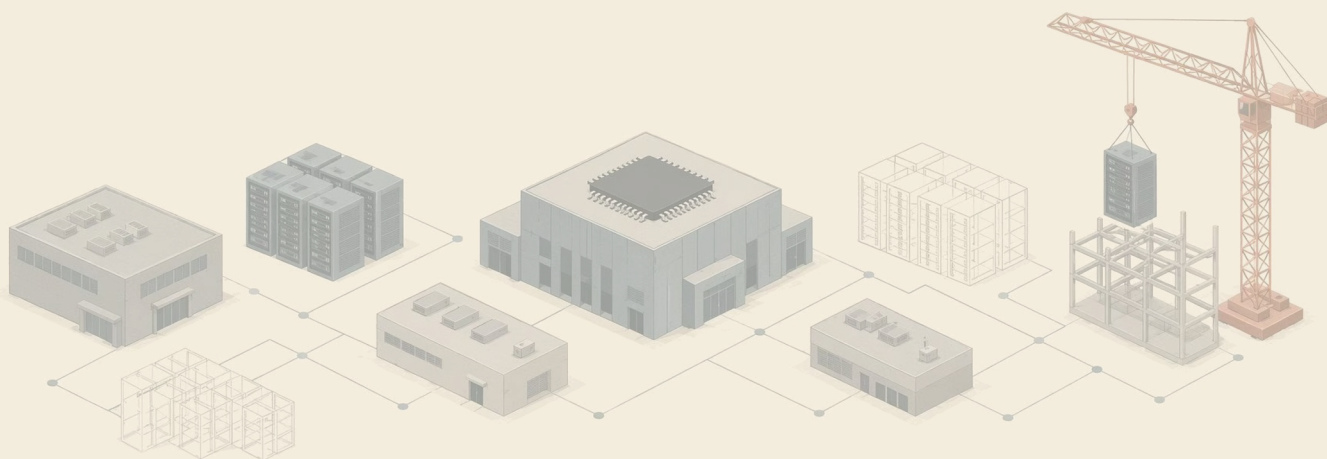


产业观察系列

2026

# 中国算力网投资发展报告

---



和 坤 经 济 笔 记

2026年9月

## 摘要

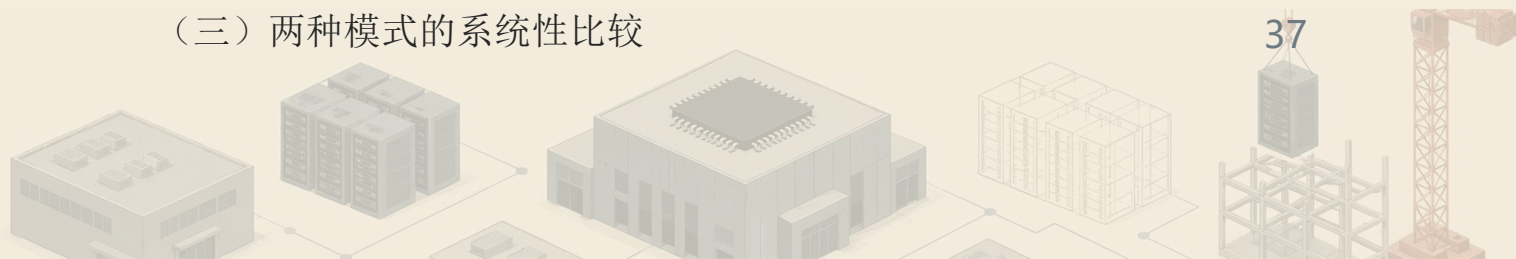
算力是人工智能时代的核心生产要素，算力网是支撑智能经济的基础性、战略性设施。2026 年以来，算力网建设在我国被提升到前所未有的战略高度：“算电协同”首次写入政府工作报告并纳入“十五五”规划纲要；4 月中央政治局会议将算力网纳入国家“六张网”统筹规划建设；9 月 11 日国务院常务会议专题研究算力网建设工作，要求坚持合理布局、规范有序，进一步完善算力基础设施。工信部《信息通信行业发展“十五五”规划》提出到 2030 年智能算力规模达到 9800 EFLOPS、较 2025 年基期增长约 5.2 倍的目标。“十五五”时期算力网建设将新增直接投资约 4 万亿元，成为扩内需、稳增长、育动能的重要抓手。

与此同时，我国算力投资与算力网建设仍面临一系列突出矛盾。供需两侧存在区域、结构、时序三重错配，全国智算中心平均算力使用率仅约 30%，一边一卡难求、一边算力空转并存。电力与绿色约束趋紧，绿电直连、隔墙售电等机制尚未打通。商业模式尚未闭环，算力租赁经历价格战后盈利分化加剧，重资产、长周期、高折旧特征与需求不确定性之间的矛盾凸显。投融资体系与算力资产属性之间存在系统性不匹配，资产标准化和可流转程度偏低，社会资本参与不足，债务资金期限与资产回收周期缺口明显。需求侧兑现进度慢于算力投资节奏，中国大模型厂商商业化体量仅为海外龙头十分之一，应用层收入规模尚不足以支撑算力投资的持续回收。算力与模型出海面临地缘政治、合规壁垒与生态短板的多重约束。

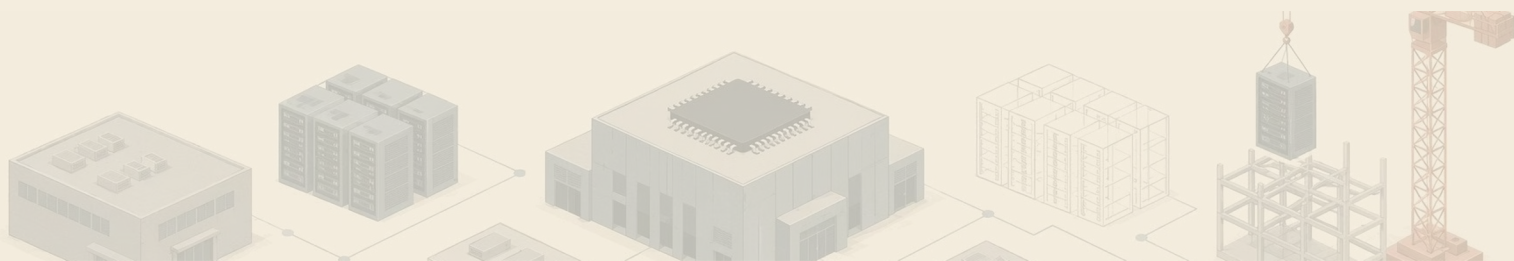
本报告认为，“十五五”时期算力网建设的核心命题，是如何在保

持投资强度的同时提升投资效能。报告在系统梳理政策脉络与发展态势的基础上，从投资主体格局、商业闭环、中美模式比较、需求侧验证和出海路径五个维度展开分析，并从提升投资精准度、建设全国统一算力大市场、创新投融资机制、深化算电协同、保障芯片供应链安全、培育需求侧生态六个方面提出对策建议，供决策参考。

|                                |           |
|--------------------------------|-----------|
| <b>一、我国算力网建设战略的演进与发展动力</b>     | <b>1</b>  |
| （一）政策脉络：从“东数西算”到“六张网”的历史跃升     | 2         |
| （二）布局成型：“8+10+3”空间格局与一体化监测体系落地 | 5         |
| （三）需求爆发：词元经济重构算力需求的量级与结构       | 7         |
| <b>二、算力投资与算力网建设的瓶颈剖析</b>       | <b>9</b>  |
| （一）算力闲置与低效运行问题突出               | 10        |
| （二）供需错配：区域、结构、时序三重矛盾交织         | 11        |
| （三）调度梗阻：跨区域算力的调度、定价和结算规则尚未统一   | 13        |
| （四）芯片瓶颈：国产替代加速，但训练端与生态链仍是短板    | 14        |
| （五）电力与绿色约束：算电协同前景广阔，但落地机制尚未打通  | 17        |
| <b>三、算力设施商业闭环的运营与资本分化</b>      | <b>20</b> |
| （一）算力设施运营的七类商业模式               | 21        |
| （二）运营分化：头部闭环跑通与中小出清加速          | 22        |
| （三）资本分化：REITs扶强不救弱与期限结构错配      | 24        |
| <b>四、中美算力投资模式的比较分析</b>         | <b>27</b> |
| （一）美国模式：私人资本主导下的债务驱动扩张         | 28        |
| （二）中国模式：科技企业主导下的多元主体协同扩张       | 30        |
| （三）两种模式的系统性比较                  | 37        |



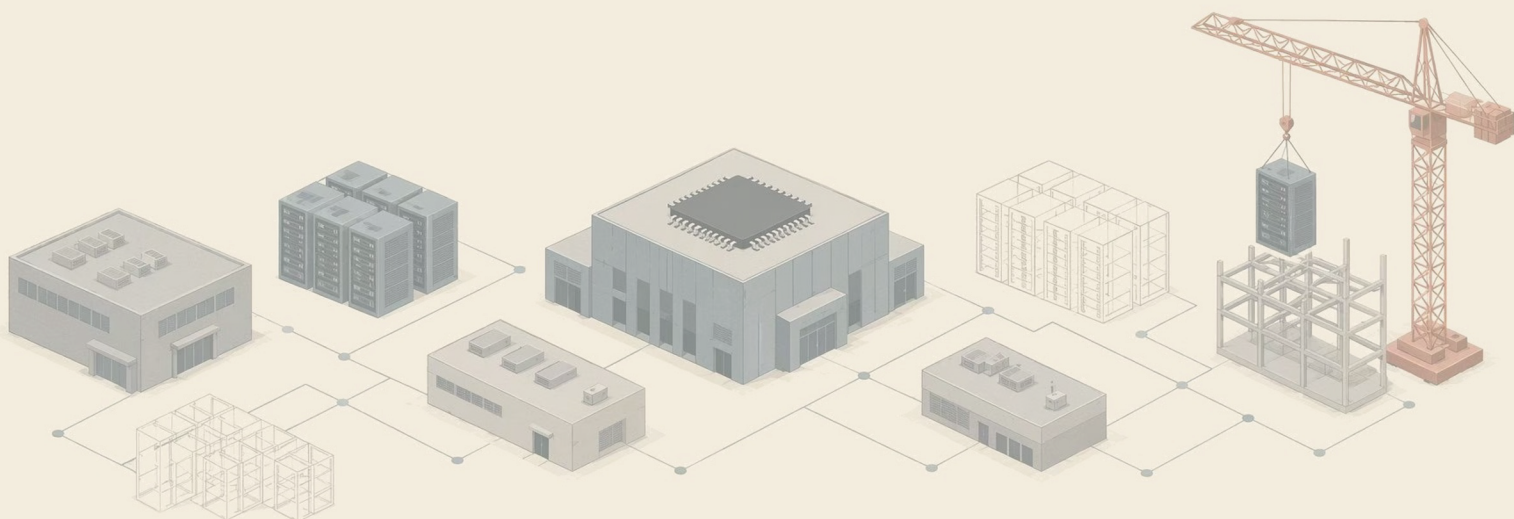
|                              |    |
|------------------------------|----|
| ■ 五、需求侧验证：AI应用盈利前景与算力投资的安全边际 | 41 |
| （一）产业链利润分布与算力投资的回报逻辑         | 42 |
| （二）中美应用层商业化差距：需求侧成熟度的结构性差异   | 42 |
| （三）需求侧验证对算力投资节奏的含义           | 45 |
| ■ 六、算力与模型出海态势、障碍与路径          | 48 |
| （一）出海态势：从产品出海到模型与算力协同出海      | 49 |
| （二）出海面临的多重障碍                 | 50 |
| （三）算力与模型出海路径选择               | 52 |
| ■ 七、对策建议                     | 55 |
| （一）以用定建，提升投资精准度              | 56 |
| （二）完善算力交易与流通机制，促进算力资源高效配置    | 57 |
| （三）创新投融资机制，走通“投融管退”全周期       | 58 |
| （四）深化算电协同，把绿电优势转化为算力成本优势     | 60 |
| （五）保障芯片供应链安全，夯实算力投资的硬件基础     | 62 |
| （六）培育需求侧生态，以应用繁荣锚定算力投资       | 63 |
| ■ 结语                         | 65 |



# 01

## SECTION 01

# 我国算力网建设战略的演进与发展动力



（一）政策脉络：从“东数西算”到“六张网”的历史跃升

“十四五”以来，我国算力基础设施的顶层布局经历了一个由点到网、由网到体系的渐进过程。2022 年 2 月，“东数西算”工程全面启动，在京津冀、长三角、粤港澳大湾区、成渝、内蒙古、贵州、甘肃、宁夏布局 8 大国家算力枢纽节点和 10 个国家数据中心集群，首次从国家层面统筹算力资源的空间配置。2023 年 10 月，工信部等六部门印发《算力基础设施高质量发展行动计划》；同年 12 月，国家发改委、国家数据局等五部门联合印发《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》，提出到 2025 年底综合算力基础设施体系初步成型，国家枢纽节点新建数据中心绿电占比超过 80%，1 毫秒、5 毫秒、20 毫秒三级算力时延圈在示范区域初步实现，并在国家层面明确“创新算力电力协同机制”。

2026 年是算力网战略地位发生质变的一年，标志性事件密集落地，形成了一条清晰的政策升级链条，如下表所示。

表 1 2026 年国家有关算力投资政策规划

| 时间         | 部门/活动             | 核心内容                                                                               |
|------------|-------------------|------------------------------------------------------------------------------------|
| 2026 年 3 月 | 全国两会              | “算电协同”首次写入政府工作报告，明确实施超大规模智算集群、算电协同等新基建工程，加强全国一体化算力监测调度，支持公共云发展；国家发改委首次系统披露“六张网”概念。 |
| 2026 年 4 月 | 中央政治局会议（4 月 28 日） | 提出“加强水网、新型电网、算力网、新一代通信网、城市地下管网、物流网等规划建设”，算力网正式纳入国家“六张网”统筹体系。                       |
| 2026 年     | 四部门联合行            | 国家发改委、国家能源局、工信部、国家                                                                 |

| 时间         | 部门/活动                           | 核心内容                                                                                                  |
|------------|---------------------------------|-------------------------------------------------------------------------------------------------------|
| 5 月        | 行动方案                            | 数据局联合印发《关于促进人工智能与能源双向赋能的行动方案》，部署算电协同等十方面任务。                                                           |
| 2026 年 7 月 | 中央政治局会议（7 月 30 日）               | 再次强调扎实推进“六张网”规划建设。                                                                                    |
| 2026 年 8 月 | 国家发改委、国家能源局                     | 发布《新型电力系统建设“十五五”规划》，首次系统性将算电协同纳入新型电力系统建设核心任务，提出“统筹能源资源配置与算力设施建设，协同规划布局算力、电力项目，做到以电强算、以算促电”。           |
| 2026 年 9 月 | 工信部发布《信息通信行业发展“十五五”规划》（9 月 7 日） | 提出到 2030 年智能算力规模达 9800 EFL OPS，有序部署万卡、十万卡及以上智算集群，适时启动 6G 商用，推进城市及热点区域网络向“双万兆”演进，信息基础设施五年累计投资 3.8 万亿元。 |
| 2026 年 9 月 | 国务院常务会议（9 月 11 日）               | 专题研究算力网建设工作，提出“坚持合理布局、规范有序”“推动算电协同、算网融合，加强算力资源监测调度和骨干光纤网络建设，强化算电设施规划衔接，加快绿电直连、源网荷储等项目落地”。             |

资料来源：各部门公告，和坤经济笔记整理。

尤其值得注意的是，“十五五”规划纲要已将算力设施建设的总方针明确为“统筹布局、有序建设算力设施，推进算力资源规模化、集约化、绿色化、普惠化发展”，并要求“推动绿色电力与算力协同布局”。“规模化、集约化、绿色化、普惠化”四化并举，标志着算力网建设从规模扩张阶段正式转入高质量发展阶段。

从投资体量看，“六张网”已成为“十五五”时期扩大有效投资的最大载体之一。据国家发改委披露，“十五五”时期国家水网、新型电网、

城市地下管网投资规模分别超过 6 万亿元、5 万亿元和约 5 万亿元；算力网方面，国家发改委引述有关机构测算称，将新增直接投资 4 万亿元。<sup>1</sup> 财通证券测算“六张网”总投资规模约 25 万亿元，较“十四五”增长近四成，其中算力网投资增幅达 50.9%，增速在六张网中位居前列；2026 年全年“六张网”及相关重点领域投资规模将超过 7 万亿元。<sup>2</sup> 在固定资产投资整体承压、房地产投资深度调整的背景下，算力网作为既能稳投资扩内需、又能发展新动能与培育新质生产力的新型基础设施，其战略价值已远超工程本身。

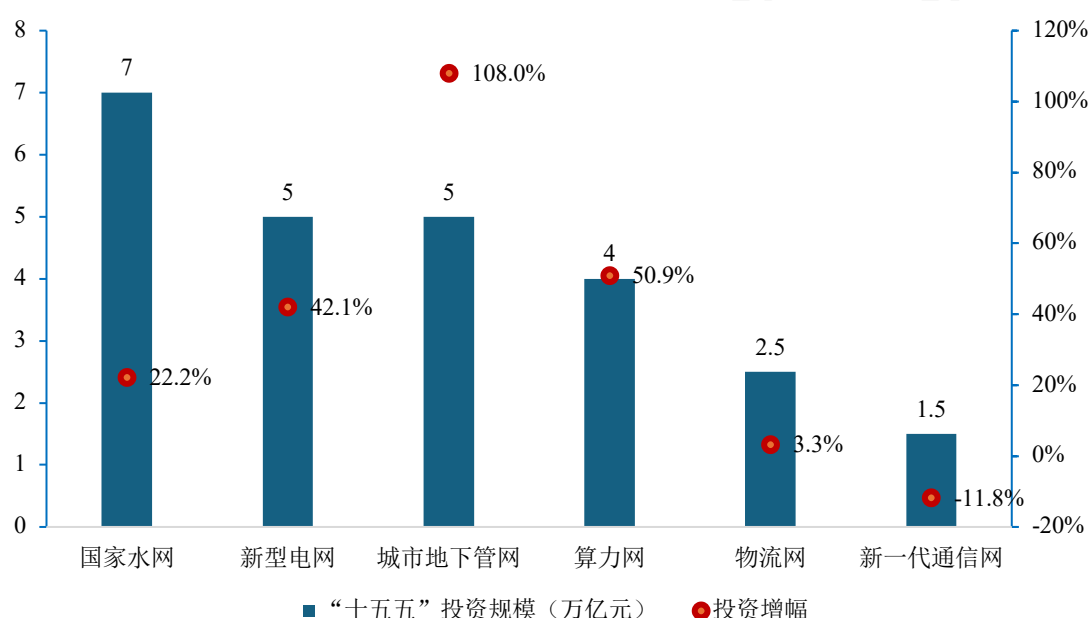


图 1 “十五五”时期“六张网”投资规模及同比增速测算

资料来源：新型电网、城市地下管网投资规模为国家发改委披露数据，其余投资规模及增速采用财通证券测算数据，和坤经济笔记整理。对通信网建设投资的测算口径，与工信部《信息通信行业发展“十五五”规划》中信息基础设施累计投资 3.8 万亿元口径不同，后者涵盖算力设施、5G/5G-A、万兆光网、卫星互联

<sup>1</sup> 新华网：《国家发改委介绍算力网、物流网等情况——“六张网”建设有新进展》，2026 年 6 月 19 日，<https://www2.xinhuanet.com/fortune/20260619/dd7cb6a2923b4466aadcfaabd15e51152/c.html>。

<sup>2</sup> 财新网：《“六网协同”提速，多元化投融资模式如何落地？》，2026 年 8 月 27 日，<https://www.caixin.com/2026-08-27/102478356.html>。

网等信息通信全行业基础设施投资，不可直接比较。

不同于水网、管网以财政投入为主，算力网以企业投资为主，市场化属性最强，对社会资本的撬动效应和商业模式创新的要求也最高。

## （二）布局成型：“8+10+3”空间格局与一体化监测体系落地

经过多年建设，我国已形成以 8 大国家算力枢纽、10 个国家算力集群、3 个算电协同发展区域为重点的“8+10+3”算力空间布局，全国算力基础设施总量规模快速增长、算电协同深入推进、算网融合加速发展。<sup>3</sup> 截至 2026 年 6 月，全国算力设施整体上架率达 71.4%；建成超 70 条连通算力枢纽、重点区域的算力传输通道。<sup>4</sup> 当前已批复 17 个国家算力互联互通区域节点。<sup>5</sup> 在 9 月 11 日至 13 日举行的 2026 中国算力大会上，我国宣布实现算力一体化统筹监测，标志着全国算力“一本账”、调度“一盘棋”迈出关键一步。<sup>6</sup>

从基础规模看，根据国家数据局报告，截至 2025 年底，全国在用算力设施机架数超过 1373 万标准机架，建成万卡智算集群 42 个，智能算力规模达到 1590 EFLOPS（FP16），位居全球第二；八大枢纽地区各类新增算力占全国新增算力的 80%以上，算力向国家枢纽集聚的态势已经确立；全国超大型算力设施平均电能利用效率（PUE）降至 1.34。<sup>7</sup> 到 2026 年 7 月底，全国智算总规模达到 2450 EFLOPS，

<sup>3</sup> 新华社，2026 年 9 月 12 日，[https://sti.xizang.gov.cn/xxgk/fdzdgk/zcjd/202609/t20260914\\_558454.html](https://sti.xizang.gov.cn/xxgk/fdzdgk/zcjd/202609/t20260914_558454.html)。

<sup>4</sup> 人民日报：《截至 6 月底我国智能算力规模同比增长 177%》，2026 年 7 月 26 日，[https://paper.people.com.cn/rmrb/pc/content/202607/26/content\\_30170977.html](https://paper.people.com.cn/rmrb/pc/content/202607/26/content_30170977.html)。

<sup>5</sup> 工业和信息化部办公厅关于公布国家算力互联互通区域节点建设名单的通知，2026 年 8 月 20 日，[https://wap.miit.gov.cn/jgsj/xgj/wjfb/art/2026/art\\_30f3cbfada28440c8ec9807f160c689f.html](https://wap.miit.gov.cn/jgsj/xgj/wjfb/art/2026/art_30f3cbfada28440c8ec9807f160c689f.html)。

<sup>6</sup> 经济参考报：《全国一体化算力监测落地 算力网建设提速》，2026 年 9 月 14 日，<http://jjckb.xinhua.net.com/20260914/de40e118a58c445b9510a887f0551e5e/c.html>。

<sup>7</sup> 国家数据局《数字中国发展报告（2025 年）》，2026 年 5 月，<https://www.nda.gov.cn/sjj/ywps/sjzg/0608/ff808081-9e81e847-019e-a7020a06-00a9.pdf>。

国家级监测调度平台已接入智能算力 1450 EFLOPS，约占全国总量的 60%。<sup>8</sup> 按照工信部规划，到 2030 年智能算力规模将达 9800 EFLOPS，以 2025 年 1590 EFLOPS 为基期约增至 6.2 倍，新建大型和超大型算力设施运行 PUE 须低于 1.2。<sup>9</sup>

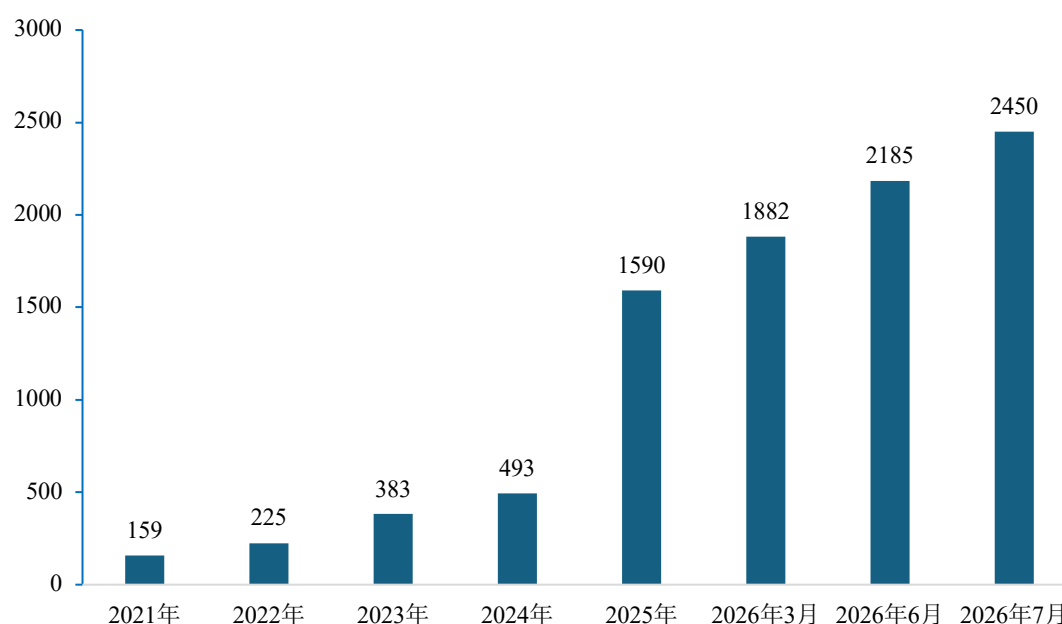


图 2 2021 年以来我国智算规模增长态势（EFLOPS）

资料来源：国家数据局、工信部《信息通信行业发展“十五五”规划》及公开报道，和坤经济笔记整理。1EFLOPS 等于 1000PFLOPS。

如图 2 所示，2025 年底全国智能算力规模为 1590 EFLOPS，同比增长 225%，2026 年以来继续呈现加速增长态势，7 月相比 6 月即已增加 265 EFLOPS，2026 年底智能算力规模极有可能达到 4000 EFLOPS。不过，要达到 2030 年 9800 EFLOPS 的目标，意味着未来四年多仍需扩大大约 3 倍，年均复合增速约 40%，算力投资的强度和持续性要求极高。

<sup>8</sup> 人民日报：《国家数据局：全国 60% 的智能算力纳入监测范围》，2026 年 8 月 29 日，<https://www.peopleapp.com/column/30053045113-500007673137>。

<sup>9</sup> 工信部：信息通信行业发展“十五五”规划，<https://www.gov.cn/zhengce/zhengceku/202609/P020260907785145440077.pdf>。

### （三）需求爆发：词元经济重构算力需求的量级与结构

支撑这一轮算力投资周期的根本动力，是人工智能从模型竞赛转向智能体（Agent）规模落地所带来的需求革命。中国电信研究院《智能体时代 AI 基础设施发展研究报告（2026）》显示，预计 2026 年我国词元（Token）年消耗量将达 10 亿亿，到 2030 年将超过 3500 亿亿，年复合增速接近 12 倍；未来 2 至 3 年我国算力需求平均每年将增长近 10 倍；从需求结构看，推理算力需求将远超训练算力，预计 2029 年我国推理算力市场占比将达到 80%；2026 年我国头部科技型企业人工智能资本支出接近 6000 亿元，占全社会技术投资总额的 10% 以上。<sup>10</sup> 从实际运行看，中国日均 AI 词元调用量 2024 年初约 1000 亿，2025 年底飙升至 100 万亿<sup>11</sup>，2026 年 3 月超过 140 万亿<sup>12</sup>；据央视财经报道，2026 年 6 月已突破 500 万亿，增长超过 5000 倍<sup>13</sup>。

需要清醒认识到，词元经济的爆发不仅带来算力需求量的增长，更重塑了算力需求的内在性质。一是**推理主导**，需求从周期性、集中式的训练任务转向持续性、并发式的推理运营，对算力的时延、稳定性、成本敏感度提出全新要求；二是**智能体放大**，单个智能体任务的词元消耗可达普通对话的数十倍乃至上百倍，部分智能体编程（Agentic Coding）任务的输入输出词元比可达 154:1，单位用户的算力消耗差距被急剧拉大<sup>14</sup>；三是**成本约束硬化**，当词元成为可计价的

<sup>10</sup> 上海证券报：《中国电信研究院报告显示：未来 2-3 年我国算力需求年均增长近 10 倍》，2026 年 9 月 14 日，<https://www.cnstock.com/commonDetail/789884>。

<sup>11</sup> 国家数据局：《全国数据资源调查报告（2025 年）》，<https://www.nda.gov.cn/sjj/swdt/xwfb/0429/ff808081-9b5e8626-019d-d86b2f54-1154.pdf>。

<sup>12</sup> 人民日报：《“词元经济”崛起，调用量两年增千倍》，2026 年 3 月 27 日，[https://www.nda.gov.cn/sjj/swdt/mtsy/0327/20260327221103432511765\\_pc.html](https://www.nda.gov.cn/sjj/swdt/mtsy/0327/20260327221103432511765_pc.html)。

<sup>13</sup> 央视网：《需求井喷，创造空间！国产算力产业链从“产品验证”走向“批量交付”》，2026 年 8 月 27 日，<https://news.cctv.cn/2026/08/27/ARTIGSQLolId7AguticEGNP8260827.shtml#1>。

<sup>14</sup> 钛媒体：《告别“堆卡”：2026，中国算力开启“Token 价值产出”大考》，2026 年 8 月 19 日，<https://www.tmtpost.com/8108840.html>。

商品，每一瓦电、每一张卡、每一秒折旧的有效产出都成为竞争胜负手，有效算力（业务实际可稳定产出的算力）取代标称算力（硬件理论峰值算力）成为产业考核的核心口径。

中泰证券的研究进一步将词元经济概括为“一场关于性价比的全产业链优化运动”，每元词元创造的任务价值等于任务价值量除以词元消耗量与单价的乘积，供给端按量计价与下游按结果、按订阅付费之间的错配，正在催生套餐化、路由聚合、效果打包的词元运营层；而在杰文斯悖论<sup>15</sup>作用下，单位词元成本的下降非但没有压缩总开支，反而解锁更多应用场景，把总消耗推向更高量级。<sup>16</sup> 这一系列变化，既为算力网建设提供了确定性需求支撑，也对既往“重建设、轻运营”的算力投资模式构成了根本性挑战。

---

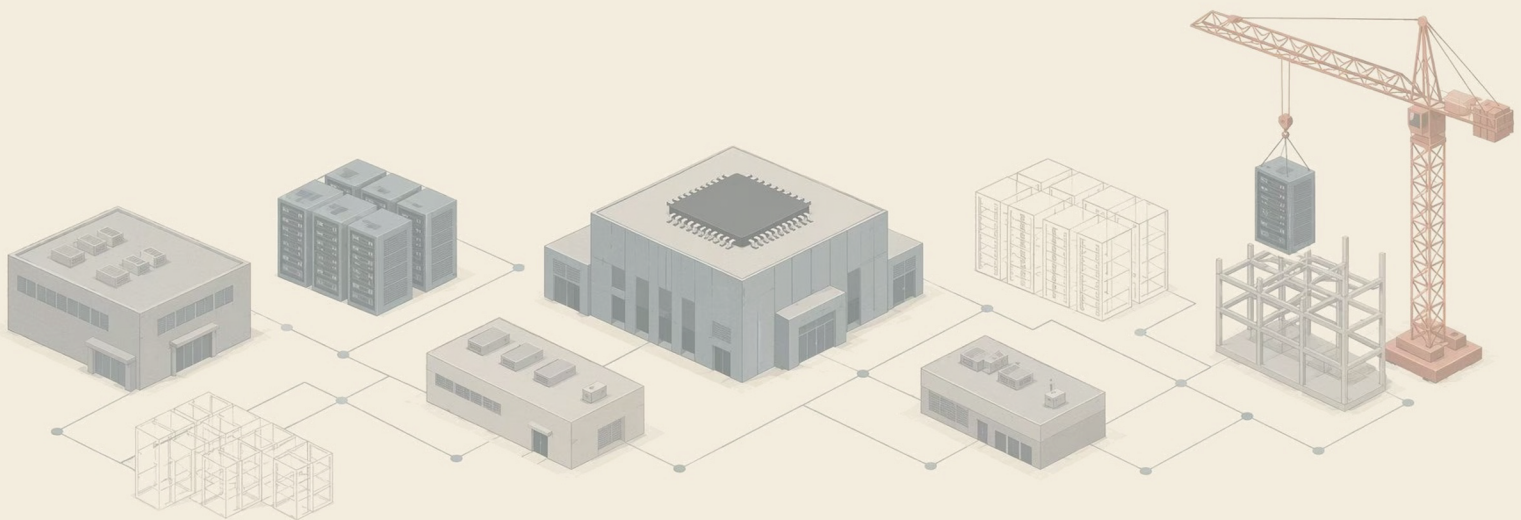
<sup>15</sup> 杰文斯悖论（Jevons Paradox），由英国经济学家杰文斯于 1865 年提出，指技术进步提升资源使用效率、降低单位成本后，反而因需求扩张导致该资源总消耗量不降反增。

<sup>16</sup> 苏仪、刘一哲：《在“卡瓦格博之心”持续瞭望顶峰——解析 Token 经济》，中泰证券研究所，2026 年 6 月 12 日。

# 02

## SECTION 02

### 算力投资与算力网建设的瓶颈剖析



## （一）算力闲置与低效运行问题突出

设施闲置与低效运行的问题早已引起多方关注。早在 2021 年底，国家发改委在“东数西算”枢纽节点批复中即设定 65% 的上架率基准线；工信部逐季监测发布全国上架率，多家智库和研究机构持续跟踪智算中心算力使用率，有关部门和地方也多次开展专项调研。

中国信通院数据显示，2025 年国内智算的总算力已经到了 38 亿左右卡时，同期用算量在 14 亿卡时左右，用算占比只有 36.8%。<sup>17</sup> 浪潮人工智能研究院测算，2024 年全国智算中心平均算力使用率仅约 30%。<sup>18</sup> 部分早期建设项目以及大量中小城市的智算中心长期处于低负载运行状态。与之形成鲜明对照的是，头部互联网企业受益于自有产品和用户基数，算力中心利用率普遍较高甚至满负荷运行，阿里云负责人公开表示服务器内“几乎没有一张卡是闲置的”<sup>19</sup>。

利用率分化呈现清晰的层级结构。头部云厂商专属智算集群依托自有大模型业务和稳定政企订单，GPU 利用率稳定在 75%-98%；省会城市中型公共智算中心以零散租赁为主，平均利用率 25%-40%；三四线城市近年新建的千卡级小型集群运营最差，GPU 利用率集中在 10%-22%<sup>20</sup>，项目普遍无法覆盖年度刚性成本。

<sup>17</sup> 中国 IDC 圈：《你有万卡集群，但你有算力资产么？》，2026 年 1 月 14 日，<https://news.idcquan.com/news/206246.shtml>。

<sup>18</sup> 科技日报：《利用率仅三成，“沉睡”的算力如何激活》，2025 年 7 月 17 日，[https://www.stdaily.com/web/gdxw/2025-07/17/content\\_371093.html](https://www.stdaily.com/web/gdxw/2025-07/17/content_371093.html)。

<sup>19</sup> 澎湃新闻：《吴泳铭：阿里服务器内没一张卡是空的，希望实现芯片全栈自研》，[https://m.thepaper.cn/newsDetail\\_forward\\_33174268](https://m.thepaper.cn/newsDetail_forward_33174268)。

<sup>20</sup> 亿欧智库：《2026 年中国 AI 芯片发展趋势报告》，2026 年 7 月。

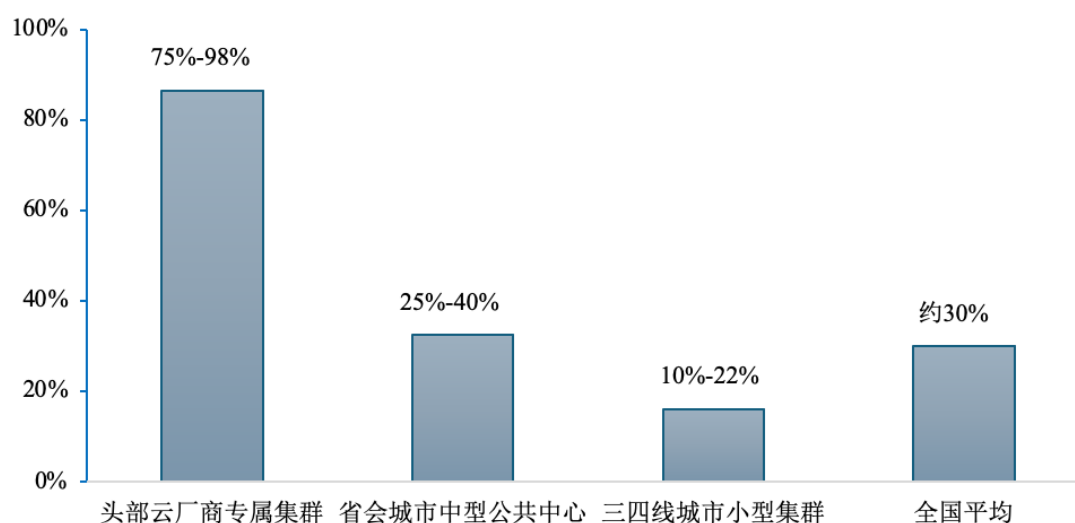


图3 我国智算中心 GPU 利用率分层分化

资料来源：亿欧智库、浪潮研究院，和坤经济笔记整理。

## （二）供需错配：区域、结构、时序三重矛盾交织

算力供需呈现显著的区域性空间错配。京津冀、长三角、粤港澳大湾区的算力需求占全国 65% 以上，且 AI 企业、高端制造业的智算需求时延要求严苛，而西部算力枢纽因跨区调度未贯通，难以以富余通用算力弥补东部智算缺口。<sup>21</sup> “东数西算”工程启动以来，八大枢纽节点的物理骨架虽已基本形成，但实地调研显示行业落差现象明显，比如，内蒙古和林格尔集群部分数据中心机柜上架率不足四成，建好机房却缺少业务需求；而距北京 200 公里的张家口集群算力持续供不应求，土地、电力指标却日益收紧。<sup>22</sup>

核心城市与周边算力承载区负载率落差悬殊。研究显示，北京数据中心峰值负载率接近 85%，而张家口、廊坊等环京算力枢纽节点同

<sup>21</sup> 半月谈：《算力供给：短缺与闲置并存》，2026 年 6 月 10 日，[http://www.banyuetan.org/jrt/detail/20260610/1000200033134991781073256356852352\\_1.html](http://www.banyuetan.org/jrt/detail/20260610/1000200033134991781073256356852352_1.html)。

<sup>22</sup> 证券日报：《算力网要“建”更要“连”》，2026 年 8 月，[https://www.szzg.gov.cn/2026/xwzx/szkx/202608/t20260811\\_5357172.htm](https://www.szzg.gov.cn/2026/xwzx/szkx/202608/t20260811_5357172.htm)。

期负载率不足 40%，两者落差折射出算力动态调度机制不成熟、供需对接渠道不畅的现实困境；与之相对，廊坊燕郊光子 1 号金融算力中心凭借与北京仅 1 毫秒的数据传输时延和约 30% 的电力成本节约，清晰展现了环京节点承接首都核心业务的现实价值。<sup>23</sup> 东西部之间的供需落差，本质上反映了“算力跟着资源走”与“需求跟着产业走”两种逻辑的张力——西部有绿电、土地和气候优势，但东部有数据、场景和用户，时延敏感型业务难以西迁，导致“西部建得多、用得少，东部用得少、建不下”。

**算力供给结构与需求演进方向存在深层错位。**财联社 2025 年 9 月调研显示，中小企业苦苦寻求高端算力，却一卡难求，而部分已建成智算中心实际利用率不足 40%；通用算力阶段性过剩与高端智能算力持续短缺并存，形成一边不足、一边闲置的典型结构性失衡。<sup>24</sup> 随着大模型训练向推理迁移、向国产芯片迁移，早期按照英伟达训练场景设计的部分智算中心，在软硬件结构上难以适配新需求。

**技术迭代周期与基础设施建设周期存在时序冲突。**算力技术约 18 至 24 个月迭代一代，而数据中心从规划、建设到投运往往需要两三年，机电设备和建筑的折旧周期更长。部分算力中心按照数年前的市场需求和技术条件规划设计，建成之时技术性能和服务能力已经落后。有芯片制造商反映，“项目等几年都没人用，等用的时候，发现已经是 3 年前建的”。<sup>25</sup> 这种“建成即落后”的风险，在推理架构快速演进、国产超节点密集迭代的背景下被进一步放大。

错配反复出现，很重要的原因是政府主导项目背后的投资逻辑偏

<sup>23</sup> 孙琪、赵静：《京津冀算力产业布局与协同发展》，《中国外资》2026 年第 10 期，第 35—37 页。

<sup>24</sup> 财联社：《高端算力“饥渴难耐”VS 智算中心“吃不饱”，破局之路已在脚下》，2026 年 9 月 26 日，<https://www.cls.cn/detail/2156687>。

<sup>25</sup> 同前引：半月谈：《算力供给：短缺与闲置并存》。

差。以往一些地方把智算中心当作看得见、可考核的政绩工程，沿用先把房子盖好、再等企业入驻的基建思路，以土地、电价优惠和财政补贴吸引项目落地，本地却没有相应的产业需求。

脱离需求建起来的设施，很快陷入闲置。前述《半月谈》调研发现，中部某省规划建设 3 个智算中心，总规模 1200 PFLOPS，当地数字经济企业不足百家，实际算力需求不足 200 PFLOPS；东部某市两个区县分别建设生物医药和工业互联网智算中心，实际入驻企业均不足 10 家。受访专家将这类现象归结为分散建设和同质化竞争，同类中心功能重叠，客户被分流，资源难以整合。

### （三）调度梗阻：跨区域算力的调度、定价和结算规则尚未统一

算力网的核心价值在“连”而不在“建”。2025 年 5 月，工信部印发《算力互联互通行动计划》，提出到 2026 年建立较为完备的算力互联互通标准、标识和规则体系，到 2028 年基本实现全国公共算力标准化互联。截至 2026 年 1 月 15 日，“1+M+N”国家算力互联互通节点体系已接入 31 个省区市的 155 家企业、578 个资源池、316 EFLOPS 智算资源、72 万张 GPU 加速卡，部署完成 1032 万条算力标识，月均调度近 300 次。<sup>26</sup>

但对照“像用电一样用算”的目标，梗阻依然明显。

其一，“路网”成本高企。以庆阳集群为例，已建成算力规模突破 18.55 万 P（FP16 口径，约合 185.5 EFLOPS），但东西部跨域网络专

<sup>26</sup> 人民邮电报：《找算难、用算贵？构建“1+M+N”算力互联互通节点体系 推动算力产业高质量发展》，2026 年 3 月，[https://www.cnii.com.cn/rmydb/202603/t20260303\\_722850.html#1](https://www.cnii.com.cn/rmydb/202603/t20260303_722850.html#1)。

线成本居高不下，受访企业负责人直言“算力产能已经落地，但数据传输‘路费’过高，综合成本亟待压降”。<sup>27</sup>

其二，“路权”尚未打通。根据工信部《2025 年通信业统计公报》，2025 年底三大运营商对外提供服务数据中心机架数量 93.8 万个<sup>28</sup>，占全国在用机架总量不足 10%，其余多数机架由第三方 IDC 服务商和自建自用主体构成。异构芯片、异构云平台之间缺乏统一的资源描述和调度接口，跨主体、跨架构、跨地域的任务级调度仍处试点阶段<sup>29</sup>，月均 300 次的调度量与全国算力体量相比微不足道。

其三，“计价”规则缺失。不同厂商算力规格、性能指标、服务等级各不相同，采购方难以有效比价，市场亟需标准化度量单位；现行“1+M+N”体系侧重技术互联和调度管理，尚未覆盖价格形成、交易撮合与价格指数等市场化定价机制。<sup>30</sup>

#### （四）芯片瓶颈：国产替代加速，但训练端与生态链仍是短板

外部管制倒逼下，国产算力芯片实现了从能用到可用的跨越。IDC 数据显示，2025 年中国 AI 加速卡整体出货约 400 万张（通常 1 张加速卡对应 1 颗 GPU），本土厂商合计出货 165 万张、份额达 41%，其中华为昇腾出货 81.2 万张、份额约 20%，如下图所示。

<sup>27</sup> 同前引：证券日报《算力网要“建”更要“连”》。

<sup>28</sup> 工信部：《2025 年通信业统计公报》，2026 年 1 月 28 日，[https://www.miit.gov.cn/gxsj/tjfx/txy/art/2026/art\\_5c99d65350f7452f999e8efcb1ee2d6a.html#1](https://www.miit.gov.cn/gxsj/tjfx/txy/art/2026/art_5c99d65350f7452f999e8efcb1ee2d6a.html#1)。

<sup>29</sup> 中国信通院：《算力互联网体系架构研究报告》，2026 年 1 月，[https://www.thepaper.cn/newsDetail\\_forward\\_32481946](https://www.thepaper.cn/newsDetail_forward_32481946)。

<sup>30</sup> 中国新闻网：《全国政协委员张云泉：让算力价格“看得见”》，2026 年 3 月 7 日，<https://www.chinanews.com/gn/2026/03-07/10582911.shtml>。

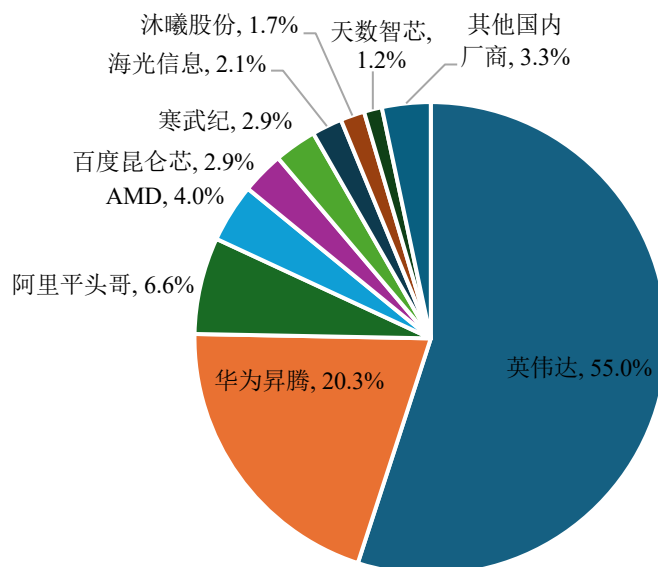


图4 2025年中国AI加速卡市场份额（出货量）

资料来源：IDC，和坤经济笔记整理。

**2026年国产AI芯片将在国内市场占据主导地位。**根据2026年8月TrendForce发布的报告，2026年中国高端AI芯片市场中，国产解决方案的份额将接近90%，出货量将同比增长超过83%，英伟达、AMD的份额将降至一成左右。2025年底投行伯恩斯坦的报告预测，英伟达在中国AI半导体市场的份额可能从2025年的约40%下降到2026年的8%左右，华为可能超过50%。

**国产芯片在推理场景已具备真实商业竞争力。**华为昇腾950PR推理卡单卡算力达H20的2.87倍<sup>31</sup>，DeepSeek V4已在技术报告中将昇腾与英伟达GPU并列为验证平台<sup>32</sup>，寒武纪思元690芯片的性能目标对标英伟达H100的80%<sup>33</sup>，海光深算三号对标A100八成以上<sup>34</sup>。

31 观察者网：《华为推出新款AI加速卡：国内唯一支持FP4，算力是H20近三倍》，2026年3月21日，<https://m.163.com/dy/article/KOII0L29051481US.html>。

32 光锥智能：《中国AI算力的突围，昇腾生态的“破”与“立”》，2026年5月1日，[https://m.thepaper.cn/newsDetail\\_forward\\_33082842](https://m.thepaper.cn/newsDetail_forward_33082842)。

33 国海证券2025年10月《寒武纪2025年季报点评》。

34 大湾区经济网：《海光信息推出全新AI算力架构 国产高端芯片“换道超车”新路径》，2026年8月28日，<https://4g.stockstar.com/detail/IG2026082800030689>。

整体看，国产芯片在推理场景已具真实竞争力，但制程与绝对算力仍落后英伟达先进产品，生态兼容与集群互联成为补差距的主要路径。

国产芯片短板主要体现在以下三个方面：

一是训练端的单卡算力差距依旧显著。华为昇腾 950PR 对 H20 的 2.87 倍优势，本质上是针对推理 Prefill 阶段和推荐场景的定向优化，并非通用训练能力的反超。从单芯片绝对算力看，950PR 的 FP8 算力为 1 PFLOPS，而英伟达 B200 的 FP4 算力高达 9 PFLOPS<sup>35</sup>，即便考虑精度差异，两者在单卡算力上仍存在显著差距。据 Semi Analysis 测算，单颗国产 AI 芯片的性能约为英伟达 Blackwell 架构的三分之一。<sup>36</sup> 尽管国产替代加速，但在万亿参数以上的超大规模预训练领域，对英伟达高端芯片（如 H200）仍存在阶段性需求。

二是 HBM 供应链的对外依赖。HBM 是 AI 加速卡中最昂贵的单一组件之一，其供给约束直接决定了国产芯片在单卡容量和带宽上的上限，进而制约了训练场景中大规模参数模型对显存的需求。美国对华 HBM 出口禁令与芯片禁令并行，H200 虽获对华出口许可，但其配套的 HBM3e 仍然不能出口。当前国产 HBM 与国际巨头相比，存在工艺、TSV 硅通孔、先进堆叠封装、良率、客户认证周期等综合差距，机构普遍预测落后三到五年。

三是软件生态差距尚难以跨越。根据近期一项对华为昇腾、摩尔线程、寒武纪、沐曦、海光五家厂商软件栈的系统性研究明确指出，国产 AI 芯片软件生态已从“基础可用”迈向“特定场景好用”，但在工

<sup>35</sup> 半导体产业纵横：《160%、155%、75%，国产 AI 芯片集体向上》，2026 年 5 月 8 日，<https://www.tmtpost.com/7979392.html>。

<sup>36</sup> 中国经济网：《知名半导体研究机构：华为云 CloudMatrix 384 可满足超大规模 AI 计算需求》，2026 年 4 月 19 日，<http://finance.people.com.cn/n1/2025/0419/c1004-40463652.html>。

具链完备性、生态成熟度和开发者基础方面仍与国际主流存在差距。<sup>37</sup> 从厂商之间看，昇腾依托 CANN、海光基于 DTK、寒武纪使用 Neuware，不同芯片之间缺乏统一的资源描述与调度接口，迁移成本极高。硬件差距可以通过制程进步和系统设计逐步缩小，但软件生态既依赖单一厂商的开发者规模与模型库积累，又受制于跨厂商的生态割裂，是国产芯片面临的最深层、最难速成的短板。

## （五）电力与绿色约束：算电协同前景广阔，但落地机制尚未打通

算力用电正以远超全社会用电增速的水平扩张，电力保障和绿电供给能力已成为算力网建设的硬约束。中国信通院数据显示，2025 年我国算力中心用电量达 1960 亿千瓦时，同比增长 18.1%，增速是全社会用电增速的 3.6 倍；预计到 2030 年用电量为 5000—7000 亿千瓦时，占全社会用电量约 3.7%—5.3%。<sup>38</sup> 中电联预计，2030 年全国算力用电量达 8000 亿千瓦时、占全社会用电量约 6%。<sup>39</sup>

算电协同的政策框架已初步成型。2026 年政府工作报告将其列为新基建工程，“十五五”规划纲要明确“推动绿色电力与算力协同布局”，《新型能源体系建设“十五五”规划》进一步提出“以电强算、以算促电”；国家枢纽节点新建数据中心绿电占比须超过 80%，绿电直连从试点转向规模化推广。截至 2026 年 4 月，全国已有 99 个绿电直

<sup>37</sup> 陈果等，《“可用”到“好用”：国产 AI 芯片软件生态的架构体系与产业实践》，《计算》2026 年第 2 期，第 77-88 页，<https://cccf.hrbueu.edu.cn/en/article/doi/10.11991/cccf.202607012#1>。

<sup>38</sup> 中国能源报：《算电网 协同赋能——多维度构建算电协同新生态》，2026 年 5 月 25 日，[https://paper.people.com.cn/zgnyb/pc/content/202605/25/content\\_30159648.html](https://paper.people.com.cn/zgnyb/pc/content/202605/25/content_30159648.html)。

<sup>39</sup> 21 世纪经济报道：《AI 赢得了芯片，输不起一次停电》，2026 年 9 月 28 日，<https://www.cnenergynews.cn/article/4TOIKqrASKJ>。

连项目完成审批，新能源总装机规模 3405 万千瓦。<sup>40</sup> 据国盛证券研报预测，全年预计落地超 200 个、总规模超 15 吉瓦。<sup>41</sup>

但落地层面仍有三道坎。

**一是成本坎。**绿电直连项目需自建专用线路、配套储能，初始投入大，在现有输配电价体系下经济性尚不稳定。根据光大证券报告测算，绿电直连模式相较于传统购电模式的综合度电成本节省约 0.028 元，降幅约 4.0%，自发自用比例是影响项目经济性的敏感因素，比例越高，成本节约越明显。<sup>42</sup> 高可靠负荷情形下绿电直连的接入容量费用高企，算力负荷与新能源开发规模错配，风光与算力设施在建设周期、资产回报上不同步，“谁受益、谁负担”原则下的输配成本折算与责任分担框架尚未建立<sup>43</sup>，业内受访企业普遍反映“落地难在成本高”<sup>44</sup>。

**二是机制坎。**绿电直连与隔墙售电是两种不同的就近消纳模式，各自面临不同的制度障碍。绿电直连通过自建专用线路实现电源与用户的点对点直供，物理上不经过公共电网，通过专线物理直连规避了输配电价定价难题，仍需全额缴纳系统成本，且政策要求年自发自用电量占总可用发电量的比例不低于 60%，占总用电量的比例不低于 30%（2030 年前不低于 35%），余电上网比例不超过 20%，对项目的负荷匹配能力提出了较高要求。

隔墙售电则依托公共配电网，允许分布式电源向邻近用户直接售

<sup>40</sup> 国家能源局：《面向“十五五”的绿电要素重塑 多主体直连的政策演进与机制破局》，2026 年 5 月 2 5 日，<https://nnenergy.in-en.com/html/newenergy-2452802.shtml>。

<sup>41</sup> 国盛证券：《电力行业周报：绿电直连 3259 万千瓦装机完成审批 算电协同进一步推进》，2026 年 3 月 1 日。

<sup>42</sup> 光大证券：《绿电直连深度报告：破解资源消纳与成本矛盾，政策持续完善赋能绿电直连长效发展》，2026 年 6 月 24 日。

<sup>43</sup> 邢庆科、许成东：《算电协同破局：从绿电直连到绿色算力出口》，《新经济导刊》2026 年第 6 期，第 14—21 页。

<sup>44</sup> 澎湃新闻：《数据中心绿电直连落地难在成本高，需从新型电力系统构建出发》，2026 年 4 月 16 日，[https://m.thepaper.cn/newsDetail\\_forward\\_32984153](https://m.thepaper.cn/newsDetail_forward_32984153)。

电，属于“电网代输”模式。企业自建风光并向多方售电涉及“供电专营”边界，根据《电力法》“一营业区一供电机构”的规定，多主体售电面临法律层面的限制；电网因免交叉补贴和基金附加而出现收入倒挂，0.15 元/度的输配电价可能降至不足 0.05 元/度；分布式电源的间歇性导致系统调频备用成本增加，渗透率超过 15% 时系统成本显著上升，责任界定、调度关系、备用费用等规则不清。电网企业对供电安全的考量使模式推广谨慎，配套政策有待进一步完善。<sup>45</sup>

**三是技术坎。**数据中心负荷要求 99.999% 以上供电可靠性，而风光出力波动性与长时储能技术成熟度之间仍有差距<sup>46</sup>，现阶段绿电直连自发自用比例和稳定性难以同时满足 AI 训练的刚性负荷要求<sup>47</sup>。此外，东部负荷中心面临电力缺口与能耗指标双重收紧，江苏、浙江、广东等地电力缺口最大，跨区域绿电调配需求迫切，绿电“产得出、送不到、用不上”的空间割裂依然突出<sup>48</sup>。

<sup>45</sup> 《财经》杂志：《对话腾讯翟永平：AI 浪潮下，数据中心如何算电协同》，2026 年 1 月 15 日，<https://m.caijing.com.cn/article/202601/391833>。

<sup>46</sup> IEEE 论文，《Energy Supplies for Future Data Centers》，2024 年 6 月 26 日，<https://ieeexplore.ieee.org/ielx8/2/10574404/10574518.pdf>。

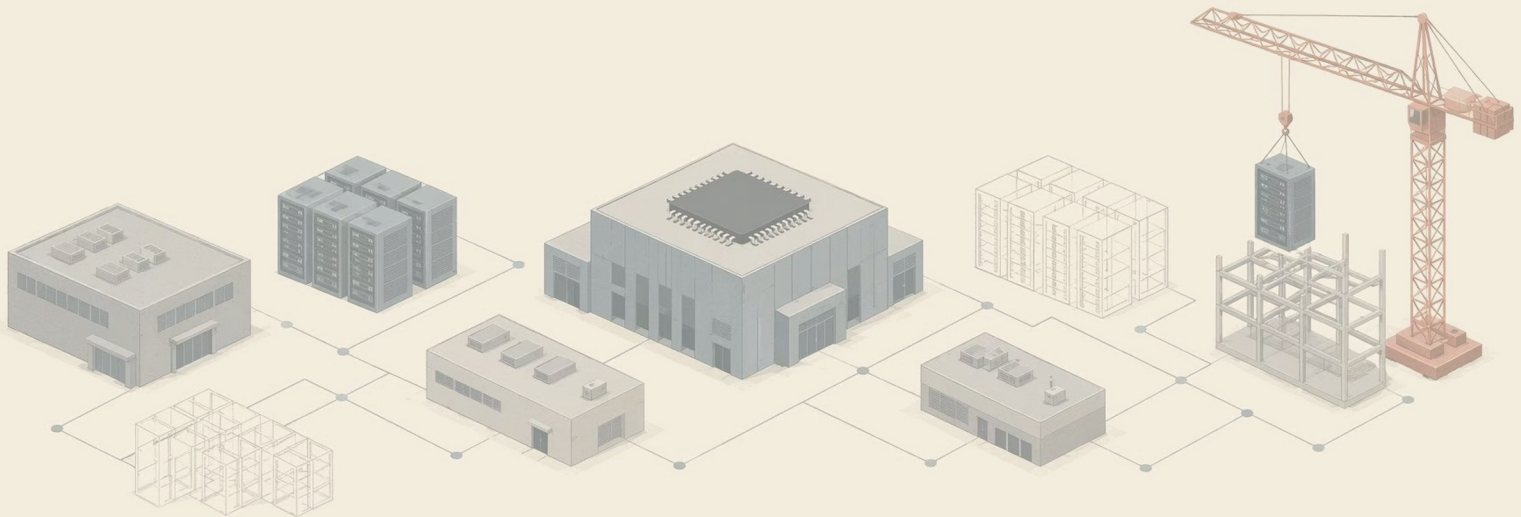
<sup>47</sup> 同前引：《财经》杂志：《对话腾讯翟永平：AI 浪潮下，数据中心如何算电协同》。

<sup>48</sup> 证券日报：《算力降本还需迈过三道坎》，2026 年 9 月 16 日，[http://epaper.zqrb.cn/html/2026-09/16/content\\_1274563.htm](http://epaper.zqrb.cn/html/2026-09/16/content_1274563.htm)。

# 03

## SECTION 03

### 算力设施商业闭环的运营与资本分化



（一）算力设施运营的七类商业模式

当前我国算力设施运营大体形成七类商业模式，其盈利逻辑、壁垒与风险各不相同，如下表所示。

表 2 算力设施运营商业模式

| 模式类型         |                | 代表主体                        | 盈利逻辑                                                                 | 核心壁垒                               | 主要风险                     |
|--------------|----------------|-----------------------------|----------------------------------------------------------------------|------------------------------------|--------------------------|
| 传统 IDC（数字地产） | 批发型数据中心        | 润泽科技（IDC 业务）、万国数据、数据港、秦淮数据等 | 向头部互联网/云厂商整栋批发机柜（不拥有服务器，客户自备硬件），长约（5-20 年）锁定现金流                      | 区位选址、能耗指标、大客户绑定                    | 客户集中度高，议价能力弱             |
|              | 零售型数据中心        | 世纪互联（城市型 IDC）、光环新网（零售 IDC）  | 按机柜/U 向中小型互联网公司、金融、政企客户零售，客户自备硬件，单价高但需自行承担空置风险                       | 一线城市区位稀缺性、获客能力、网络质量与运维服务           | 上架率波动大、价格战、客户退租导致机柜变成负资产 |
| 算力租赁         | 批发型算力租赁        | 利通电子、协创数据、润泽科技（AIDC 业务）     | 自购 GPU 服务器，向头部客户长协出租（合同期 3-5 年），合同订单收入覆盖设备投入；部分转向“基础费用保底+Token 用量分成” | 高端 GPU 采购渠道与配额（英伟达 NPN 认证资质）、大客户绑定 | 客户违约/合同到期降价、芯片断供、客户集中度极高 |
|              | 零售型算力租赁（“二房东”） | 中小算力服务商                     | 向上游长租或融资租赁获取算力，按小时/天碎片化零售给中小客户，赚取批零差价与调度运营效率收                        | 长尾客户获客能力、算力调度与运维效率、平台易用性（API/      | 上架率/出租率波动大、大厂降价降维打击、设备贬值 |

| 模式类型          |  | 代表主体                       | 盈利逻辑                                                                         | 核心壁垒                           | 主要风险                           |
|---------------|--|----------------------------|------------------------------------------------------------------------------|--------------------------------|--------------------------------|
|               |  |                            | 益                                                                            | 计费系统)                          |                                |
| 云厂商一体化服务      |  | 阿里云、腾讯云、三大运营商云             | 算力+平台+生态捆绑, IaaS/PaaS 分层变现, MaaS 平台抽成                                        | 全栈技术、用户基数、生态                   | 资本开支巨大, 回报周期长                  |
| Token 工厂/MaaS |  | 大模型厂商（智谱、Kimi）、商汤大装置、软通智算等 | 自研模型推理能力直接变现, 按 Token 使用量计费, 逐步向按任务结果计价延伸; 部分厂商（如 Kimi）探索向托管云平台收取 Token 收入分成 | 模型能力的持续迭代、推理工程优化、单位 Token 成本控制 | 模型迭代风险, 单位成本竞争                 |
| 算力调度与交易平台     |  | 中国电信“息壤”、中国联通“星罗”          | 为政府/企业提供跨域异构算力纳管与调度服务, 收取并网费/服务费                                             | 统一资源标准、算网互联能力、平台公信力            | 跨厂商接口不统一导致纳管成本高、双边市场冷启动难、交易量不足 |

资料来源：网络公开信息，和坤经济笔记整理。

（二）运营分化：头部闭环跑通与中小出清加速

从市场格局看，批发型与云厂商模式凭借客户与生态优势保持较高上架率和盈利水平。根据科智咨询报告，截至 2025 年底，全国在运营 IDC 机柜 IT 总容量达 19933 兆瓦，2021—2025 年复合增长率 20.1%；2025 年定制化 IDC 机房上架率达 71.4%，明显高于非定制化机房的 55.7%——前者在规划阶段即与头部客户需求深度绑定，后者

面向多租户和中小客户、AI 适配能力不足，去化压力突出；行业头部集中趋势延续，TOP10 服务商市场份额已超 55%。<sup>49</sup>

这种分化在算力租赁领域同样显著。以润泽科技和利通电子为代表的批发型算力租赁服务商，凭借自建基础设施或核心渠道资源，实现了高盈利和快速扩张。以润泽科技为例，其 2025 年 AIDC 业务收入 25.10 亿元、同比增长 73%，毛利率达 48.50%，并依托公募 REITs 打通了“投融管退”闭环；利通电子则通过与腾讯的深度绑定，算力业务收入占比快速提升至 36.3%，毛利率近 50%。

与之相对，以中低端通用算力和散租业务为主的中小算力服务商正经历剧烈出清，困境具有系统性。一是价格战击穿成本线，2025 年成都地区 A100 服务器租赁单价从 8 元/小时降至 5.8 元/小时，而单台服务器含托管、电力、运维的小时综合成本约 6.5 元，每租出一小时亏损 0.7 元。二是 API 模式替代散租需求，2026 年以来 DeepSeek 等模型厂商以极低价 API 直接服务终端客户，中小客户转向调用 API 而非租卡跑推理，低端推理散租需求趋势性萎缩。三是设备迭代加速贬值，算力芯片 18 至 24 个月更新一代，而智算中心投资回收期普遍需要 5 至 8 年，部分中低端项目回本周期从 2 年拉长至 4 至 5 年，超过设备经济寿命，形成尚未回本即已过时的困局。

受此影响，2026 年以来多家持有中低端算力的公司选择出售资产或终止项目，首都在线拟出售服务器回笼资金，亿田智能将刚购入的智算机作价 1.56 亿元转让，广脉科技在合同解约后出售算力资产回血。中小租赁企业客源持续萎缩，毛利率跌破 20%。<sup>50</sup>

<sup>49</sup> 科智咨询：《2026 年中国第三方算力中心服务商发展研究报告》。

<sup>50</sup> 财闻：《算力租金雪崩、机柜大面积空置相关概念下挫 利通电子、盈峰环境双双跌停》，2026 年 6 月 26 日，[https://m.sohu.com/a/1041969204\\_122014422](https://m.sohu.com/a/1041969204_122014422)。

但零售型算力租赁并非全面消亡。拥有高端卡源和优质客户的二房东反而在 Token 需求爆发中实现量价齐升，H100 一年期租赁合同价格从 2025 年 10 月的 1.70 美元/卡时低点反弹至 2026 年 3 月的 2.35 美元，涨幅近 40%<sup>51</sup>。市场资源正加速向具备卡源优势和客户绑定能力的企业集中。

盈利分化的决定性变量不是有没有卡，而是四项能力：稳定的低价电力供给、通畅可靠的网络、智能调度平台与 7×24 小时运维、以及绑定头部客户的确定性需求。总体而言，算力设施的商业模式正在从卖资源向卖能力、卖服务、卖生态跃迁，能够完成这一跃迁的主体才能在价格战中存活。

### （三）资本分化：REITs 扶强不救弱与期限结构错配

算力设施重资产、长周期、高折旧、技术迭代快的属性，决定了其融资必须走股债结合、长短搭配的路子。当前我国算力投融资体系已初步形成多层次格局<sup>52</sup>。

股权端，国家级和地方算力产业基金、市场化 VC/PE、上市公司 IPO 及定增共同供给资本金。地方层面，甘肃庆阳设立数字经济产业基金 50 亿元，投资方向明确包括算力集群建设、算力调度运营、AI 场景应用等前沿领域。<sup>53</sup> 上市公司股权融资同步放量，协创数据拟定增募资不超 80 亿元加码智算中心建设，利通电子拟定增募资不超 50 亿元投向宁夏中卫智算中心项目。

<sup>51</sup> 财联社：《算力太紧俏！英伟达四年前发布的 H100 租赁费用近半年飙升近 40%》，2026 年 4 月 3 日，<https://www.cls.cn/detail/2334039>。

<sup>52</sup> 中国金融四十人论坛：《以用为本、协同为要，扎实推进“算力网”建设》，2026 年 9 月 8 日，<https://www.cf40.org.cn/article/1/289341>。

<sup>53</sup> 甘肃日报：《庆阳智算规模再攀新高》，2026 年 8 月 24 日，<https://gansu.gansudaily.com.cn/system/2026/08/24/031421759.shtml>。

债权端，根据中国信通院报告，当前算力中心融资渠道以政府专项债、银行信贷等传统路径为主<sup>54</sup>。项目贷款是当前最主要融资方式之一，企业主导的智算中心项目中银行贷款占比普遍达 70%左右；设备融资租赁因能匹配设备生命周期而广泛应用，已有金租公司落地行业首单 SPV 数据中心设备租赁项目<sup>55</sup>，部分算力采购项目中融资租赁占比高达 80%以上<sup>56</sup>；绿色信贷、“算力贷”“词元贷”等创新产品开始密集出现，中国银行已成为首家参与中国算力平台生态建设的金融机构<sup>57</sup>。

资产证券化端实现标志性突破。2025 年 8 月润泽科技、万国数据两单数据中心公募 REITs 登陆沪深交易所，合计认购资金超 4700 亿元，标志着数据中心正式进入基础设施 REITs 版图，投融管退全周期通道初步走通。<sup>58</sup> 截至 2026 年 7 月，国家发改委累计向证监会推送各类基建 REITs 项目 133 单，其中数据中心、智算中心类算力项目合计 4 单，除已上市两单首发产品外，还有江天数据首发项目和润泽科技存量扩募项目在途，算力 REITs 梯队正在形成。<sup>59</sup>

但结构性短板同样明显，集中体现为三个错配。

一是 REITs 门槛与现实资产的错配。公募 REITs 要求权属清晰、现金流稳定、上架率高，只有头部优质资产能够发行，最需要盘活的

<sup>54</sup> 中国信通院：《算力中心创新融资研究报告（2025 年）》，<http://221.179.172.81/images/20251020/8761760929099239.pdf>。

<sup>55</sup> 新华财经：《交银金租落地行业首单 SPV 数据中心设备租赁项目，赋能算力基础设施建设》，2026 年 7 月 13 日，[https://m.cnfin.com/wx/share?url=/m.cnfin.com/cmjj-lb/zixun/20260713/4439574\\_1.html](https://m.cnfin.com/wx/share?url=/m.cnfin.com/cmjj-lb/zixun/20260713/4439574_1.html)。

<sup>56</sup> 央广网：《蜂助手公告全资子公司签署超 46 亿元算力大单》，2026 年 8 月 5 日，[https://finance.cnr.cn/ycbd/20260805/t20260805\\_527749241.shtml](https://finance.cnr.cn/ycbd/20260805/t20260805_527749241.shtml)。

<sup>57</sup> 中国银行：《中国银行成为首家参与中国算力平台生态建设的金融机构》，2026 年 5 月 9 日，[http://wap.bank-of-china.com/bif/bi1/202605/t20260509\\_25666106.html](http://wap.bank-of-china.com/bif/bi1/202605/t20260509_25666106.html)。

<sup>58</sup> 经济参考报：《全国首批数据中心 REITs 登陆沪深交易所》，2025 年 8 月 8 日，<http://jjckb.xinhuanet.com/20250808/58df35deadc64212becb5241b4f84d92/c.html>。

<sup>59</sup> 睿思网：《老钢厂变身印钞机？A 股第三单 IDC REIT 来了》，2026 年 7 月 10 日，<https://wap.zhennguannews.cn/html/zgh/412037.html>。

大量低利用率中小智算中心反而进不了通道，资产证券化扶强不救弱。这意味着行业最需要资金盘活的存量资产，恰恰是最难获得资本市场支持的资产。

**二是债务资金期限与资产回收期的错配。**银行贷款期限多为 5 至 8 年，而智算中心投资回收期普遍在 10 年以上，两者存在明显期限缺口；融资租赁利率普遍高于项目实际收益率，在价格战环境下极易倒挂。算力设施需要在折旧周期内完成投资回收，但债务偿还节奏快于资产产生现金流的节奏，形成结构性资金缺口。

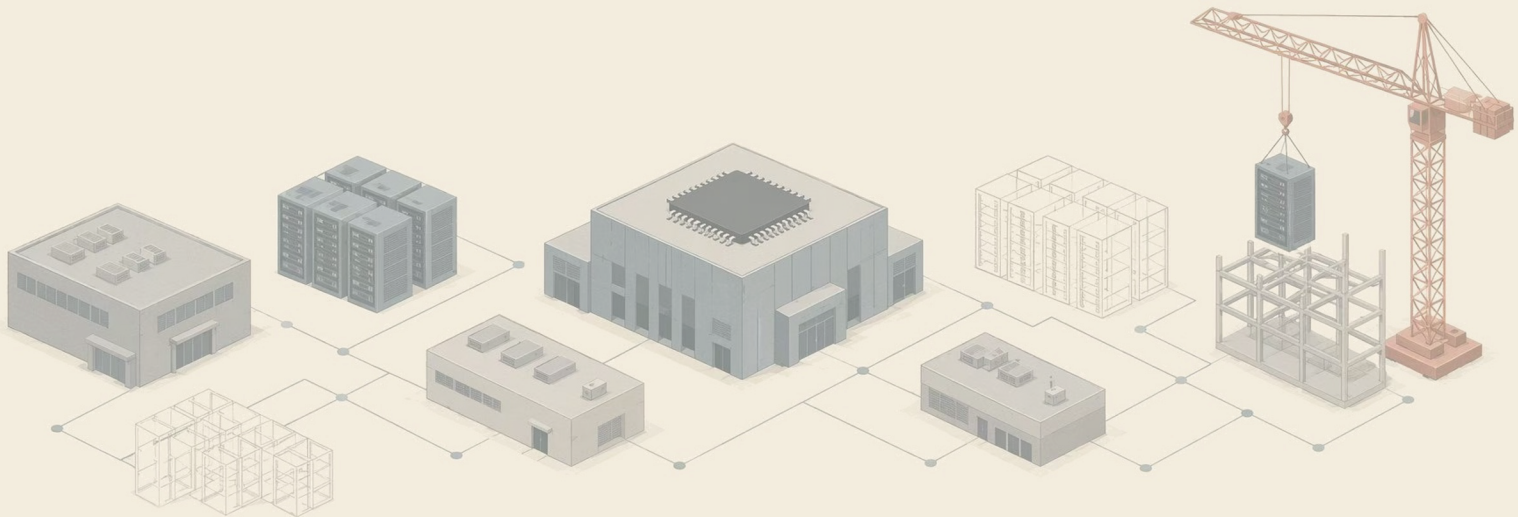
**三是社会资本风险偏好与资产属性的错配。**算力中心单项目投资 20 亿至 100 亿元，投资回收期长，上架率和电价是生死线，中小投资者和专业机构难以判断技术迭代风险。这一重资产、长周期、高门槛的属性，使其天然不适合缺乏长期资金和风险承受能力的中小民间资本。由此形成政府资金的替代性介入，地方政府平台公司以招商和政绩为导向举债建设智算中心，既推高了自身债务，也在局部挤压了更有效率的民间资本空间。

这三个错配共同指向一个核心矛盾：算力资产的现金流特征与现有金融工具的供给特征之间存在系统性不匹配。商业闭环的形成，不仅需要运营层面跑通盈利模型，还需要资本层面找到与资产属性相匹配的融资结构。REITs 破冰提供了方向，但从首批两单到覆盖行业多数资产，从批发型数据中心到智算中心，从首发到扩募，制度供给仍有较大缺口。

# 04

## SECTION 04

### 中美算力投资模式的比较分析



## （一）美国模式：私人资本主导下的债务驱动扩张

美国算力投资呈现市场主导、科技巨头主导的格局，投资规模加速扩大，且主要金融机构对投资规模的预测持续上调。从资本开支看，微软、谷歌、亚马逊、Meta 四家科技巨头 2026 年第二季度单季资本开支合计达 1712 亿美元，同比增长约 79%；2026 年全年四大巨头资本开支指引合计约 7250 亿美元<sup>60</sup>。摩根士丹利 2026 年 9 月最新预测，亚马逊、Alphabet、微软与 Meta 的合并资本支出将于 2027 年达到约 1.47 万亿美元，同比增长约 60%。<sup>61</sup> 高盛 2026 年 9 月预测，美国五大超大规模云服务商 2027 年 AI 基础设施支出将增长 54%至 1.2 万亿美元。<sup>62</sup> 美国银行数据显示，截至 2026 年 8 月，全球 AI 相关融资规模已达 3440 亿美元，远超 2025 年全年的 1360 亿美元。<sup>63</sup>

当资本开支超过自由现金流，美国科技巨头构建了一套多层次的外部融资体系，本质上是用金融工程手段将巨额基建开支移出资产负债表，但担保、租赁等隐性风险并未消失。其金融工程手段如下：

**第一层，债券融资放量。**2025 年美国超大规模云厂商为 AI 基础设施发行新债约 1210 亿美元；2026 年以来发行节奏进一步加快，仅亚马逊、字母表（谷歌）、Meta、甲骨文四家截至 7 月上旬即发行约 1940 亿美元债券，高盛预计计入微软后五家全年发债规模可能达 2500 亿美元、2027 年升至 4000 亿美元<sup>64</sup>；云厂商债券占美国非金融投资

<sup>60</sup> 财联社：《7250 亿美元！美股 AI 四巨头挥舞支票本 分析师：供应链面临“全方位瓶颈”》，2026 年 4 月 30 日，<https://www.cls.cn/detail/2360617>。

<sup>61</sup> 网址：<https://www.edgen.tech/zh-tw/news/post/hyperscaler-capex-to-peak-at-15-trillion-in-2027-morgan-stanley-says#1>。

<sup>62</sup> 彭博社：<https://www.bloomberg.com/news/articles/2026-09-25/goldman-sees-hyperscaler-ai-capex-rising-50-to-1-2-trillion>。

<sup>63</sup> Odaily 星球日报，2026 年 8 月 7 日，<https://followin.io/zh-Hans/feed/26839994>。

<sup>64</sup> 界面新闻：《AI 大厂开始借钱烧算力》，2026 年 7 月 30 日，<https://finance.sina.com.cn/jjxw/2026-07-30/doc-inikqcsml1316041.shtml>。

级公司债供应的比例从 2025 年的约 2% 升至 2026 年的 19%<sup>65</sup>。瑞银数据显示，2026 年全球超大规模云厂商资本开支将几乎耗尽全部经营现金流，远高于过去十年约 40% 的平均水平。<sup>66</sup> 资金价格也在上升：上述公司 2—4 年期债券利差中位数已从 2025 年的 30 个基点升至 40 个基点，20 年以上长期债券利差中位数从 108.5 个基点升至 118 个基点，超额认购倍数则从 2 月接近 5 倍降至 7 月不足 2 倍<sup>67</sup>。

**第二层，SPV 表外融资。**以 Meta 为代表，通过设立特殊目的实体（SPV），由私募资本（如 Blue Owl、贝莱德）持有多数股权并承担大部分债务融资，Meta 仅保留少数股权，将一次性资本开支转化为长期租金和担保义务。Meta 已用此结构完成两笔合计约 410 亿美元的数据中心交易；xAI 也以类似方式为 GPU 集群融资。

**第三层，GPU 金融化与私募信贷。**以 CoreWeave 为代表，GPU 抵押贷款利率从 2023 年的约 15% 降至投资级水平，GPU 从折旧设备演变为可抵押、可证券化的金融资产，数据中心证券化规模快速扩张。2026 年以来趋势加速，推理芯片首次成为大额贷款抵押品<sup>68</sup>；英伟达联合多家大型资管机构筹划算力融资平台，拟长期动员超 5000 亿美元第三方资本，并为项目提供部分残值支持<sup>69</sup>。芯片厂商正从单纯卖方转向算力金融生态的关键组织者。

**第四层，表外承诺与循环交易。**资本闭环之外还有更大规模的隐性敞口。超大规模云厂商已承诺的未开始执行租赁费用合计约 1.09 万

<sup>65</sup> 财联社：《摩根士丹利：AI 投资进入融资时代 科技巨头表外承诺已超 3 万亿美元》，2026 年 8 月 26 日，<https://www.cls.cn/detail/2465220>。

<sup>66</sup> 财联社：《摩根士丹利：2026 年全球 AI 相关债券发行规模将接近 5700 亿美元》，2026 年 6 月 10 日，<https://www.chinastarmarket.cn/detail/2396460>。

<sup>67</sup> 同前引：界面新闻：《AI 大厂开始借钱烧算力》。

<sup>68</sup> TechCrunch，2026 年 6 月 20 日，<https://techcrunch.com/2026/07/17/why-the-first-gpu-financiers-are-turning-to-inference-chips-in-a-400-million-deal/>。

<sup>69</sup> 光明网：《芯片巨头，官宣万亿元大计划》，2026 年 8 月 11 日，[https://m.gmw.cn/toutiao/2026-08/11/content\\_1304546628.htm](https://m.gmw.cn/toutiao/2026-08/11/content_1304546628.htm)。

亿美元，其中微软约 3290 亿、Meta 约 2790 亿、甲骨文约 2610 亿美元。<sup>70</sup> 若叠加采购承诺以及英伟达、博通等芯片厂商提供的担保和信用支持，相关表外承诺及担保总额已超过 3.1 万亿美元。<sup>71</sup> 与此同时，承诺关系在链条内部相互交织。OpenAI 对甲骨文的算力采购承诺约 3000 亿美元，对 CoreWeave 超 220 亿美元，对亚马逊 380 亿美元后扩展至约 1380 亿美元。英伟达曾计划对 OpenAI 进行千亿美元级投资，但该框架已于 2026 年 1 月搁置，实际注资规模为 300 亿美元。各方收入与估值在链条内自我强化，一旦需求或融资端出现松动，风险也将沿同一链条传导。

## （二）中国模式：科技企业主导下的多元主体协同扩张

与美国私人科技巨头主导的模式不同，在 2025 年及以前，我国算力设施投资主要是政府引导与产业资本双轮驱动，三大运营商、地方国有企业及平台公司承担了公共智算中心和 IDC 基建的主要建设任务，而大型科技企业则在算力设备采购与 AI 技术研发领域占据绝对主导。

根据中金公司报告，2025 年中国 AI 基础层投资规模约 1400 亿美元，其中龙头企业出资 650 亿美元、三大运营商 114 亿美元、政府引导基金 634 亿美元。这一统计覆盖了龙头企业和运营商的投资，但未纳入地方国有企业和平台公司、第三方 IDC/AIDC 服务商的算力设施投入，且政府基金 634 亿美元主要投向芯片研发与半导体制造等重资产领域，并非全部用于算力基础设施。具体来说，数据中心与配套

<sup>70</sup> 半导体产业纵横：《1.09 万亿美元，AI 租赁还在疯狂烧钱》，2026 年 8 月 6 日，[https://www.thepaper.cn/newsDetail\\_forward\\_33721537](https://www.thepaper.cn/newsDetail_forward_33721537)。

<sup>71</sup> 财联社：《摩根士丹利：AI 投资进入融资时代 科技巨头表外承诺已超 3 万亿美元》，2026 年 8 月 26 日，<https://www.chinastarmarket.cn/detail/2465220#1>。

基建等投资规模为 704 亿美元，其中私人部门投资 590 亿美元、三大运营商 114 亿美元。<sup>72</sup> 因此，该数据可用于观察中美 AI 投资的宏观结构差异，但大幅低估了国内的算力基建规模，不宜直接作为中国算力设施投资规模的完整口径。

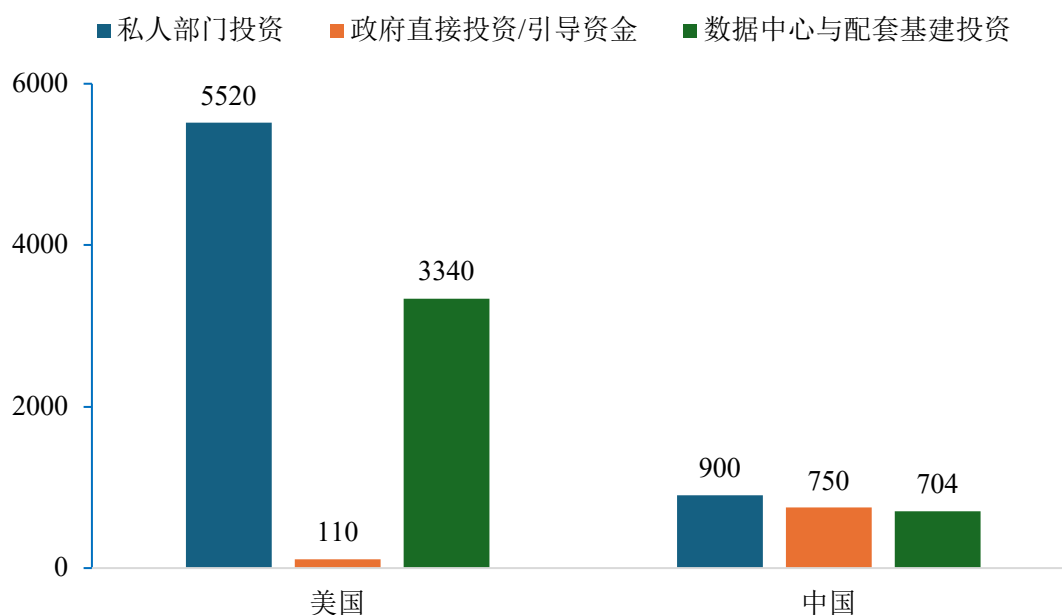


图 5 2025 年中美 AI 投资结构对比（亿美元）

资料来源：中金公司，和坤经济笔记整理。

2026 年起，大型科技企业跃升为国内算力投资主导力量，运营商、地方国资投入退居其次，第三方服务商则构成重要补充。从投资主体看，算力投资已形成“一超多强”格局：

一是大型科技企业，占据绝对主导地位。根据穆迪的估算，中国头部科技企业（预计主要指字节跳动、阿里巴巴、腾讯）AI 相关资本支出将从 2025 年的约 650 亿美元，增长至 2026 年的约 1400 亿美元，

<sup>72</sup> 中金公司：《中美 AI 投资比较研究》，2026 年 1 月，<https://www.163.com/dy/article/KJMVCQ130519QIKK.html>。

预计 2027 年将达到 1650 亿美元。<sup>73</sup> 如包含其他大型互联网平台（百度、快手、京东、美团、拼多多等）和主要大模型厂商（如 DeepSeek、智谱、MiniMax、Kimi），2026 年度合计算力投入约 6800-10400 亿元，如下表所示。科技企业已不再是与政府/国有部门并列的“一极”，而是算力投资增量的主要贡献者。

表 3 2026 年中国主要科技企业算力投资规模

| 企业类型   |      | 算力投资         | 说明/依据                                                                                           |
|--------|------|--------------|-------------------------------------------------------------------------------------------------|
| 头部科技企业 | 字节跳动 | 1900-4700 亿元 | 资本开支官方确认超 2000 亿元，同时彭博社报道至多 700 亿美元，算力投资占比预计 95%以上；同时预计算力租赁 180-250 亿元                          |
|        | 阿里巴巴 | 1700-2080 亿元 | 高盛预测 2027 财年（2026 年 4 月-2027 年 3 月）资本开支 2090 亿元；算力投资预计占资本开支 80%-90%；租赁算力支出预计 100-200 亿元（属于运营支出） |
|        | 腾讯   | 2000-2200 亿元 | 高盛预测 2026 年资本开支 2100 亿元；算力投资占比预计 95%以上；算力租赁预计 100-200 亿元                                        |
| 互联网平台  | 百度   | 330-360 亿元   | 花旗预测 2026 年资本开支 320 亿元；算力投资占比预计 95%以上；算力租赁预计 30-60 亿元                                           |
|        | 快手   | 270-300 亿元   | 2026 年集团资本开支 260 亿元；算力投资占比预计 95%以上；算力租赁预计 20-50 亿元                                              |
|        | 京东   | 约 90 亿元      | 已披露的在建项目合计投资额（武汉、廊坊二期、大同灵丘）                                                                     |
|        | 美团   | 120-130 亿元   | MarketScreener 聚合分析师预测 2026                                                                     |

<sup>73</sup> 网址：<https://technode.global/2026/08/28/chinese-hyperscalers-ramp-up-ai-spending-but-trail-us-rivals-moodys/>。

| 企业类型      |          | 算力投资            | 说明/依据                                                                                                                                                    |
|-----------|----------|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| 大 模 型 厂 商 |          |                 | 年资本开支 128 亿元；算力投资占比预计 95%以上                                                                                                                              |
|           | 拼多多      | 数十亿元            | 主要是算力租赁服务费                                                                                                                                               |
|           | DeepSeek | 190-200 亿元      | IT 之家报道 The Information 援引消息称 2026 年 1-7 月在 AI 基础设施方面投入约 110 亿元，对应年化在 200 亿元左右                                                                           |
|           | 智谱       | 80-86 亿元        | 主要采用算力租赁模式，研发开支中 70%-80%为算力服务费，2026 年全年研发开支预计 45-50 亿元；中金预测 2026 年收入 60.7 亿元，按 24.6%毛利率、推理算力占销售成本 90%计算推理侧算力服务费约 36 亿元                                   |
|           | Kimi     | 45-70 亿元        | 租赁阿里 2 万卡集群（按 H100 计算）年租金 30-50 亿元，研发侧算力消耗推算 15-20 亿                                                                                                     |
|           | MiniMax  | 60-75 亿元        | 算力投入以云服务采购为主、国产集群自建为辅；根据招商证券数据，其研发开支中算力占比（70%-80%）与营业成本中推理算力占比（90%以上）；花旗预测 2026 财年总收入 5.77 亿美元，按毛利率 17%-20%可测算推理侧算力约 28 亿—30 亿元；训练算力预计 30-40 亿元；另有自建集群投入 |
| 合计        |          | 约 6800-10400 亿元 |                                                                                                                                                          |

资料来源：高盛、花旗、南华早报、彭博社、The Information、MarketScreener、招商证券、各公司公告，和坤经济笔记整理。

二是电信运营商，投资结构向算力倾斜。2026 年中国移动、中国电信、中国联通资本开支分别计划为 1366 亿元、730 亿元和约 500 亿

元，在总开支收缩 8%—9.5%的背景下，算力投资逆势增长。三大运营商算力投资合计超 800 亿元，占各自资本开支比重均超过 35%。<sup>74</sup> 从上半年执行情况看，中国电信算力网投资同比增长 97%，占比提升 24.6 个百分点；中国移动上半年算力网资本开支 156 亿元、同比增加 65 亿元；中国联通算力投资占比提升至 37%，投资额同比提升超过 80%。<sup>75</sup> 三大运营商资本开支结构集体向算力倾斜，但绝对规模显著低于头部科技企业。

表 4 2026 年三大运营商规划算力投资

| 企业   | 算力投资     | 说明/依据      |
|------|----------|------------|
| 中国移动 | 378 亿元   | 同比增长 62.4% |
| 中国电信 | 255 亿元   | 同比增长 26%   |
| 中国联通 | 超 175 亿元 | 占资本开支超 35% |

资料来源：各公司公告，和坤经济笔记整理。

三是地方国资或平台公司，承担区域算力兜底角色。主要通过政府专项债、平台举债等融资方式建设公共智算中心。根据中国信通院统计，截至 2024 年 7 月，全国由地方国资公司主导投资建设的智算中心已有 55 个，大部分首批规划建设算力峰值在 100—300P（FP16）之间。<sup>76</sup> 根据科智咨询统计，截至 2024 年 8 月，地方政府主导建设的智算中心项目数量占比 34%，是数量上最大的建设力量，但按算力规模统计占比仅为 14.2%。<sup>77</sup> 地方国资平台在项目数量上仍占较高比

<sup>74</sup> 上海证券报：《三大运营商 2025 年业绩保持稳健 2026 年投资主攻算力赛道》，2026 年 3 月 27 日，<https://www.cnstock.com/commonDetail/657732>。

<sup>75</sup> 中国邮电报：《三大运营商上半年交出“韧性答卷”》，2026 年 8 月 25 日，[https://www.cnii.com.cn/rmydb/202608/t20260825\\_756409.html](https://www.cnii.com.cn/rmydb/202608/t20260825_756409.html)。

<sup>76</sup> 中国信通院：《智算基础设施发展研究报告（2024 年）》，2024 年 9 月，<http://221.179.172.81/images/20241105/51051730819245496.pdf>。

<sup>77</sup> 科智咨询：《2024 年中国智算中心产业发展白皮书》，2024 年 12 月。

例，但单体投资规模和算力密度远低于科技企业主导的项目。

2025 年以来，浙江、辽宁、湖北、四川、山东、北京、海南等省市及宜昌、十堰、湖州、青岛、雄安、宁德等地市密集设立了省级或市级数据集团与算力投资平台，注册资本多在数亿元至数十亿元之间。这些平台的职能高度趋同，主要是承担公共数据授权运营、建设算力调度平台、推进全省算力统筹。部分县级市乃至区级层面也出现了专门的算力建设运营主体。密集的平台化布局，直接动因是地方对算力需求爆发的判断和增加算力投资的意愿。但与此同时，受新增智算中心立项审批趋严、利用率红线约束和隐性债务管控加强等因素影响，预计地方国资在算力投资总量中的权重持续下降。

四是专业化第三方服务商，投资向头部企业集中。2026 年主要算力/IDC 服务上市公司已披露的资本开支合计超过 600 亿元，但集中在头部企业中。万国数据 2026 年全年资本开支指引约 100 亿元，世纪互联资本支出指引 100 亿—120 亿元，协创数据 2026 年 2 月新增 110 亿元服务器采购合同。2026 半年报的固定资产及在建工程变化也体现了算力投资向少数头部企业集中的特征。

表 5 2026 年国内第三方服务商算力投资规模测算

| 企业   | 资本开支       | 说明/依据                                                                                                           |
|------|------------|-----------------------------------------------------------------------------------------------------------------|
| 协创数据 | 230-290 亿元 | 2026 半年报固定资产及在建工程相比 2025 年末已增加 114.6 亿元，主要为算力服务设备投入；2026 年 2 月公告新增不超过 110 亿元服务器采购合同；定增募资 80 亿元（其中 46.86 亿投智算中心） |
| 万国数据 | 100 亿元     | 2026 年 8 月上调资本支出指引至 100 亿元，2026 年 5 月宣布未来三年投入 300—500 亿元在国内 AI 基础设施                                             |

| 企业   | 资本开支         | 说明/依据                                                                                                     |
|------|--------------|-----------------------------------------------------------------------------------------------------------|
| 世纪互联 | 100-120 亿元   | 2026 年资本支出指引，计划交付 450—500MW                                                                               |
| 润泽科技 | 100 亿元（权益融资） | 2026 年通过 REIT 扩募、可转债合计权益融资 100 亿元；2026 上半年在建工程 99.2 亿元                                                    |
| 宝信软件 | 15-20 亿元     | 2026 上半年固定资产增加 8 亿元，在建工程 17.62 亿元                                                                         |
| 光环新网 | 7-10 亿元      | 2026 上半年固定资产新增 3.6 亿元，在建工程 11.5 亿元；和林格尔智算中心拟投 12.35 亿元、呼和浩特算力基地拟投 22.95 亿元（均处前期审批阶段）                      |
| 奥飞数据 | 5-10 亿元      | 2026 上半年固定资产新增 0.6 亿元，在建工程 13.2 亿元                                                                        |
| 证通电子 | 2-5 亿元       | 2026 上半年在建工程 8.5 亿元                                                                                       |
| 首都在线 | 2.6 亿元       | 2026 上半年固定资产新增 2.6 亿元，在建工程 0.1 亿元；2026 年 4 月拟定增募资不超 15.87 亿元，投向京北云计算算力中心（二期，7.87 亿元）、智算中心建设（5.00 亿元）但尚未批复 |
| 云赛智联 | 约 2 亿元       | 2026 上半年固定资产新增 1.7 亿元，在建工程 0.7 亿元                                                                         |
| 利通电子 | <1 亿元        | 2026 上半年固定资产及在建工程合计减少 2.1 亿元（折旧增加所致），在建工程仅 0.2 亿元，预计年度算力直接投入较少                                            |
| 数据港  | <1 亿元        | 2026 上半年固定资产及在建工程合计减少 2.8 亿元，在建工程仅 0.2 亿元，预计年度算力直接投入较少                                                    |
| 网宿科技 | <1 亿元        | 2026 上半年在建工程 0.8 亿元                                                                                       |
| 浙大网新 | 预计无投资        | 无算力设施在建工程                                                                                                 |
| 浙数文化 | 预计无投资        | 无算力设施在建工程                                                                                                 |

| 企业   | 资本开支  | 说明/依据     |
|------|-------|-----------|
| 二六三  | 预计无投资 | 无算力设施在建工程 |
| 科华数据 | 预计无投资 | 无算力设施在建工程 |
| 铜牛信息 | 预计无投资 | 无算力设施在建工程 |

资料来源：各公司公告，和坤经济笔记整理。

从结构变化看，2025 年至 2026 年中国算力投资主体格局发生了实质性位移：科技企业的资本开支增速远超其他主体，在投资总额中的占比显著提升；运营商算力投资规模快速增长，但在总投资中的份额被科技企业的增量稀释；地方国资平台在项目数量上仍占较高比例，但单体投资规模有限；第三方服务商的头部企业与科技企业形成产业链绑定，成为算力基础设施的重要补充供给方。这种“科技企业领跑、多元主体协同”的格局，与美国私人科技巨头主导但缺乏政府统筹协调的模式形成对照——中国的优势不在于政府替代市场进行大规模算力投资，而在于政府通过规划引导、绿电供给和标准制定，为科技企业的大规模市场化投资创造制度条件。

（三）两种模式的系统性比较

中美两种模式的差异根植于各自的制度逻辑，很难简单以优劣论之。具体比较如下表所示。

表 6 中美算力设施投资模式比较

| 维度   | 美国模式                       | 中国模式                               |
|------|----------------------------|------------------------------------|
| 主导力量 | 私人科技巨头（微软、谷歌、Meta、亚马逊、甲骨文） | 科技企业主导投入+地方国资平台+央企（运营商）压舱+第三方服务商补充 |
| 资金来源 | 经营现金流→债券→SPV               | 企业自有资金+银行贷款+融                      |

| 维度   | 美国模式                                        | 中国模式                                                                                               |
|------|---------------------------------------------|----------------------------------------------------------------------------------------------------|
| 源    | 表外融资→私募信贷→GPU抵押证券化                          | 资租赁+产业基金+上市公司可转债+REITs起步                                                                           |
| 投资结构 | 2025年3340亿美元，2026年约7800亿美元（高盛5大厂商口径），均为私人投资 | 2025年：头部科技企业590亿美元、三大运营商114亿美元、地方国资及平台预计数千亿元<br>2026年：主要科技企业约6800-10400亿元，将超过三大运营商（800亿元）与地方国资平台之和 |
| 投资节奏 | 资本开支约为经营现金流1.2倍，依赖债务扩张                      | 科技企业资本开支翻倍式扩张；运营商结构倾斜；地方平台新增立项趋谨慎                                                                  |
| 布局逻辑 | 需求导向、点对点布局超大型AI工厂（500MW—1GW级）               | 规划导向、“8+10+3”全国统筹，兼顾区域均衡                                                                           |
| 能源约束 | 电力瓶颈突出，多州暂停并网审批，倒逼巨头直购核电、自建电源               | 电力供给充裕但绿电消纳与跨区调配机制待完善                                                                              |
| 风险形态 | 债务杠杆、表外承诺、循环交易、估值泡沫                         | 低效重复建设、利用率不足、地方隐性债务、财政补贴依赖                                                                         |
| 优势   | 资本效率高、技术迭代快、商业闭环清晰                          | 企业投资意愿强、应用场景丰富、绿电与制造优势；政府统筹规划与动员能力提供支撑                                                             |
| 软肋   | 回报不及预期即引发债务与估值双杀                            | 部分地方平台与中小项目财务回报敏感度低，易局部“建多用少”                                                                      |

资料来源：中金公司、高盛、穆迪、瑞银、财联社等，和坤经济笔记整理。

美国模式以私人科技巨头的商业回报为驱动，资本运作能力极强。债券融资、SPV表外融资、GPU抵押证券化等工具使算力建设摆脱了对单一主体资产负债表的依赖，资本效率和技术迭代速度都很突出。

中国模式则是在政府引导下以科技企业的大规模市场化投资为主体，地方国资平台作为规模可观的第二梯队，运营商承担算力兜底职能，第三方服务商构成重要的市场化补充。这种模式动员能力强，战略定力足，能够通过统筹规划和政策引导，在算力设施布局上实现超前推进。但资本运作工具尚不成熟，资产标准化和可流转程度偏低，地方投资项目对财务回报的敏感度也相对不足，容易滋生低效重复建设。

从美国的实践看，有三条经验值得中国在完善算力投融资体系时参考。其一，资本运作能力本身就是基础设施竞争力。SPV 项目融资、REITs、资产证券化等工具将重资产转化为可交易、可定价的金融资产，使算力建设摆脱对单一主体资产负债表的依赖。我国首批数据中心 REITs 的成功发行已初步印证这一路径的可行性，但覆盖范围和产品层次仍有较大拓展空间。其二，长期租赁与确定性合同是融资的基石。美国 AI 数据中心融资均建立在投资级租户长期租约之上，本质是把算力需求信用转化为项目融资信用。我国大量智算中心恰恰缺少这种确定性合同，融资自然困难。其三，资产流动性决定资本进入意愿。GPU 从设备变为抵押品、数据中心从房产变为证券，极大拓宽了资金供给面。我国在算力资产标准化、可估值、可流转方面刚刚起步，制度供给仍有明显缺口。

**美国模式的风险集中于债务链条。**相关表外承诺及担保总额已超过 3.1 万亿美元，相当比例的公司债基于对 OpenAI 等少数公司的需求预期发行。表外融资安排只是改变了债务的会计呈现，SPV、经营租赁、残值担保等或有负债一旦触发回表条件，将即时冲击资产负债结构。循环交易则使收入与投资在链条内部相互放大，估值与真实需求脱钩，风险沿同一链条传导。对我国而言，这些潜在风险的含义是：

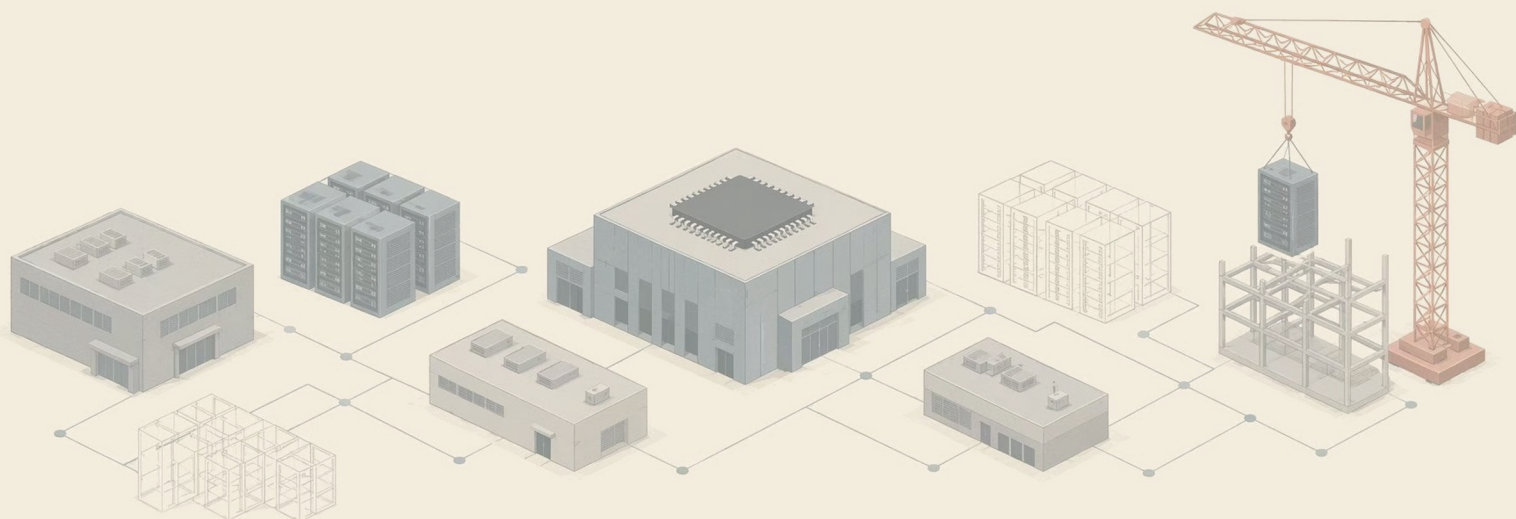
算力投资必须锚定真实需求与现金流，警惕以债务堆出来的繁荣，防止地方政府和企业算力领域重走高杠杆、过剩、出清的老路。

和坤经济笔记

# 05

## SECTION 05

### 需求侧验证：AI应用盈利前景与算力投资的 安全边际



## （一）产业链利润分布与算力投资的回报逻辑

算力投资能否持续，最终取决于下游应用能否产生足够收入来消化算力供给。如果应用层长期无法形成自我造血能力，算力设施的利用率红线就难以守住，投资回收更无从谈起。

从全球 AI 产业链看，当前利润分布呈“哑铃型”失衡。芯片环节和云基础设施环节攫取了绝大部分利润；模型层除 Anthropic 于 2026 年二季度实现调整后营业利润首次转正（约 5.59 亿美元，单季营收超 115 亿美元）外，其余国内外大模型公司普遍亏损。OpenAI 2025 年账面营收 130.7 亿美元、经营亏损约 209 亿美元，预计 2026 年仍亏损约 140 亿美元，最早 2030 年现金流转正。应用层则处于融资驱动向利润驱动过渡的前夜，仅 Midjourney 等极少数订阅制且推理成本可控的公司跑通盈利。

这一分布对算力投资的含义很明确。算力作为“卖铲人”，在产业扩张期是确定性的受益者，但其长期回报终究取决于模型与应用层能否赚钱、何时形成自我造血能力。其中的传导逻辑是：算力的单位成本持续下降，应用层用算力的门槛随之降低，更多场景变得经济可行，应用层的收入规模提高。只有应用层赚到钱，算力投资才有持续回收的基础。

## （二）中美应用层商业化差距：需求侧成熟度的结构性差异

中美大模型厂商的商业化进度存在代际级差距。美国方面，截至 2026 年 7 月底，Anthropic 年度经常性收入（ARR）已突破 650 亿美元，较 2025 年底的 90 亿美元增长逾 7 倍；OpenAI 的 ARR 也在 9 月接近 700 亿美元，约为 2025 年底的 3.5 倍。中国方面，根据企业公

开信息及相关媒体披露，2026 年 8 月智谱、Kimi、DeepSeek、MiniMax 的 ARR 分别达到 16 亿美元、超 10 亿美元、超 10 亿美元和 8 亿美元。国内视频生成模型中，字节跳动 Seedance 在 2026 年 6 月 ARR 达到 20 亿美元，同期快手可灵达到 5 亿美元。据 Rhodium Group 2026 年 9 月根据 3 月至 8 月披露的数据测算，中国 7 家主要 AI 开发商模型业务 ARR 合计约 107 亿美元，约为 OpenAI 与 Anthropic 合计规模的十分之一。<sup>78</sup>整体看，国内厂商仍处于 ARR 高速爬坡期，但绝对体量与海外龙头存在约一个数量级的整体差距。

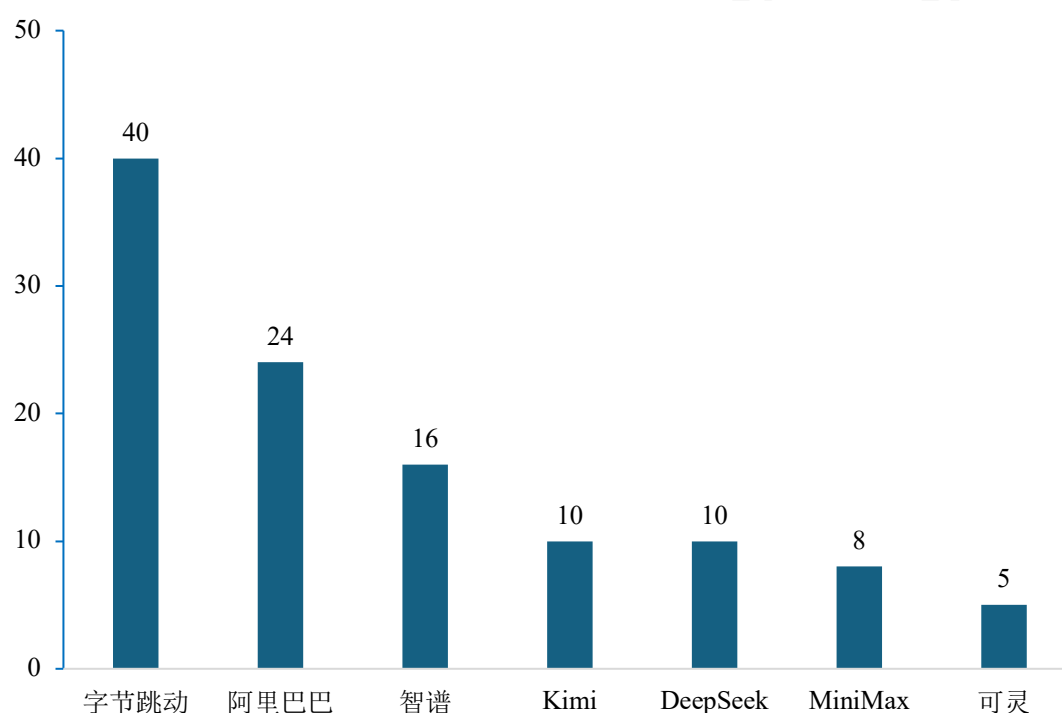


图 6 中美大模型厂商 ARR 对比（亿美元）

数据来源：字节跳动（2026 年 7 月）、阿里巴巴（2026 年 8 月）数据来自 Rhodium Group，Kimi（2026 年 8 月）、智谱（2026 年 8 月）数据分别来自彭博社、财新网；DeepSeek（2026 年 8 月）数据来自 36 氪（援引 The Information）；MiniMax（2026 年 8 月）数据为官方披露；可灵（2026 年 3 月）为官方披露数

<sup>78</sup> <https://www.digitaltoday.co.kr/cn/view/105569/chinas-ai-open-weights-strategy-revenue-10-percent-of-openai-anthropic>。

据。和坤经济笔记整理。

差距的成因是多重的。从市场结构看，美国企业级软件付费生态成熟，Anthropic 企业级 API 与服务占比约八成、年支出超百万美元的企业客户突破 1000 家（2026 年 3 月）、超 10 万家企业通过 AWS Bedrock 运行 Claude，Ramp 数据显示 34.4% 的美国企业为 Anthropic 付费、已反超 OpenAI。<sup>79</sup>

相比之下，中国企业客户付费意愿与客单价显著偏低，表现为三个层面的结构性差距。

其一，企业 IT 预算中软件与订阅服务的占比远低于美国。根据长江商学院 2026 初对 2000 家规上工业企业的调研，在有 AI 支出的企业中，资金平均有 57.2% 投向了 AI 硬件，38.8% 用于开发/定制内部 AI 系统，而用于 AI 订阅、服务与培训的比例仅占 3.5%。<sup>80</sup>

其二，中小微企业价格敏感度构成付费转化的直接瓶颈。赛迪智库中小企业研究所与网商银行联合发布的《AI 时代小微经营者观察》报告显示，2026 年一季度中国小微经营者净利润率均值仅为 4.4%，58.5% 曾因价格原因放弃购买或停止续费 AI 工具，44% 倾向于使用免费或低价产品。<sup>81</sup>

其三，单位调用价格与客单价存在量级差异。即使经历 2026 年的集体涨价，大部分中国前沿模型完成单项任务的价格仍只有约 0.04 至 0.50 美元，而 Claude 顶级模型的同类调用价格在 2 至 4 美元之间，GPT 最昂贵的一批模型约为 1 至 2 美元。这一价差直接体现在整体

<sup>79</sup> 第一财经：《不到一年增长超七倍，Anthropic 年化收入突破 650 亿美元》，2026 年 8 月 18 日，<https://www.stcn.com/article/detail/4083354.html>。

<sup>80</sup> 《财经》杂志：《渗透率 10%，一季度实体产业 AI 应用数据观察》，<https://m.caijing.com.cn/s/202605/5159374>。

<sup>81</sup> 新华网：《AI 时代小微经营者观察》：近六成小微因成本弃用 AI 工具，生态共助成破局关键，2026 年 6 月 25 日，<https://imgs.xinhuanet.com/tech/20260625/f450798a75c74507bdd8ede2ec47d204/c.html>。

毛利率的差距上：根据 Rhodium Group 的估算，DeepSeek 毛利率约 45%，智谱 AI 和 MiniMax 分别仅为 26%和 18%，而 Anthropic 约为 65%。<sup>82</sup>

中美应用层商业化差距对算力投资的含义，首先体现在需求侧成熟度差异上。中国企业客户付费意愿与客单价显著偏低，同等规模算力投资面临需求不足的风险。但中国开源模型的系统性成本优势——训练成本约为海外头部企业的十分之一，推理定价仅为 10%–20%，头部厂商 GPU 利用率超 70%——可在相当程度上对冲这一劣势，且头部厂商 ARR 快速爬坡，投资收益将逐步显现。因此，同等规模的算力投资在中国并不必然面临更低的效率。真正拉长回报周期、推高利用率风险的，是地方补贴驱动、与真实需求脱节的中低端算力。

### （三）需求侧验证对算力投资节奏的含义

综合判断，国内 AI 应用环节的盈利拐点将沿“三步走”渐次到来：第一步（2026—2027 年），模型层头部厂商通过提价与推理工程优化实现毛利转正，编码（Coding）等核心赛道率先跑通；第二步（2027—2028 年），随着推理成本继续下降与企业级渗透率提升，垂直行业应用（金融、政务、制造、医疗）形成规模付费；第三步（2028 年以后），智能体经济的价值分配机制成熟，应用层利润池真正打开。对算力投资而言，这意味着“十五五”前期算力需求有模型竞赛与推理爆发的双重支撑，确定性较高；但中期必须跟踪应用层收入兑现进度，一旦应用盈利持续不及预期，算力投资的高速扩张将失去需求锚点。

鉴于算力投资体量大、周期长、不可逆性强，上述判断仍有必要

<sup>82</sup> BT 财经：《107 亿对 1000 亿：中国 AI 的收入账本》，2026 年 9 月 21 日，<https://business.timescn.com/articles/626914.html?in=wanfang#1>。

放到不同需求情景下做敏感性检验。

表 7 “十五五”算力投资不同情景敏感性分析

| 情景   | 核心假设                                                      | 需求侧表现                                                 | 对 4 万亿元投资的含义                                          | 政策含义                            |
|------|-----------------------------------------------------------|-------------------------------------------------------|-------------------------------------------------------|---------------------------------|
| 乐观情景 | Agent 经济超预期爆发，应用层 2027 年规模盈利；国产芯片窗口期守住并扩大                 | Token 消耗增速超中国电信研究院预测上沿；推理占比提前至 2028 年达 80%；高端算力持续供不应求 | 4 万亿元投资偏紧，枢纽节点与高端智算需适度超前；投资重心向推理算力、东部高端集群倾斜           | 加快项目审批与要素供给，防止“算力缺口”制约 AI 产业竞争力 |
| 基准情景 | 应用拐点按“三步走”渐次到来；Token 需求按预测兑现（2026 年 10 亿亿、2030 年 3500 亿亿） | 算力需求未来 2—3 年年增近 10 倍后回落至 40%—50%；利用率分化延续但中枢抬升         | 4 万亿元投资规模总体适度，但结构上约两到三成中小项目需要通过整合、转型或退出消化；存量盘活与增量优化并重 | 坚持“以用定建”、利用率红线与动态清退，投资节奏前高后稳    |
| 悲观情景 | 模型能力提升放缓、应用付费不及预期；海外管制变化冲击国产芯片商业化                         | Token 价格通缩快于量增，算力租金再下台阶；中小智算中心利用率长期低于 20%             | 约三分之一规划项目面临推迟或取消；地方政府与租赁企业债务风险显性化                     | 预留投资弹性空间，建立需求监测与项目退出机制，严防债务风险传染 |

资料来源：基于中国电信研究院、中国信通院等公开预测的情景推演。

综合以上分析，应用层盈利前景对“十五五”算力网 4 万亿元新增

投资有三点判断。

**第一，需求侧的方向性支撑是扎实的。**即使在国内应用层商业化进度偏慢的情况下，Token 消耗量的爆发式增长和推理算力需求的持续扩大，仍然为算力投资提供了确定性的需求基础。风险主要集中在低端、同质、无运营能力的供给。

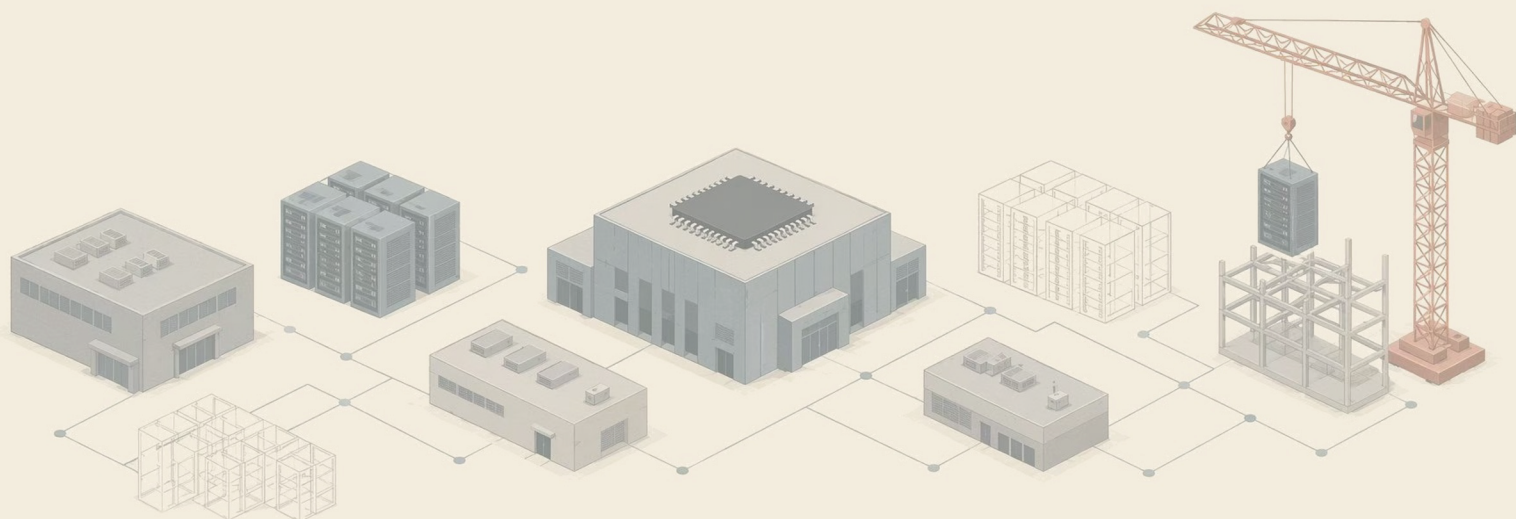
**第二，投资安全的关键变量是节奏与结构。**应用层从烧钱换规模切换到量价齐升需要时间，算力投资的节奏应与应用层收入兑现的进度动态匹配。建议将 Token 消耗量、头部模型厂商收入、算力租金价格、上架率四类高频指标纳入算力投资景气度监测体系，作为发改等部门在项目审批和要素供给上把握节奏的参考依据。当指标显示需求侧兑现明显滞后时，可通过放缓审批节奏、收紧能耗指标和电价优惠等方式，引导企业自主调整投资节奏，避免形成新的低效算力供给。

**第三，培育需求侧生态与扩建设施必须同步甚至优先。**悲观情景的触发点主要在于应用层收入兑现，这意味着如果应用层商业化进度持续低于预期，算力设施的利用率红线将面临更大压力。政策重心应从单纯的扩建设施转向供需两端协同发力，通过普惠算力、政府应用牵引、Token 经济商业模式创新等方式，加快应用层形成自我造血能力。

# 06

## SECTION 06

### 算力与模型出海态势、障碍与路径



## （一）出海态势：从产品出海到模型与算力协同出海

2026 年被业界视为国产大模型出海元年，出海形态已从应用产品延伸至模型能力与算力基础设施三个层次。

**模型出海势头最为强劲。**OpenRouter 数据显示，中国 AI 大模型周调用量已连续 22 周超过美国、居全球首位；2026 年 9 月 21 日至 27 日当周，平台总调用量达 146 万亿 Token，其中中国上榜模型达 62.22 万亿 Token，约为美国模型（14.2 万亿）的 4.4 倍；在按付费金额排名的榜单中，Kimi K3 高居第二、智谱 GLM 与 DeepSeek 进入前六。<sup>83</sup> 需要注意的是，OpenRouter 的 Token 调用量仅占全球总量的 2% 左右，平台上显示的中国大模型服务提供方，其数据中心实体多部署在海外，海外开发者实际使用的是部署在海外云平台上的大模型服务；这意味着这些调用并未回流国内数据中心，Token 出海在物理意义上还未实现，本质仍是模型出海。<sup>84</sup>

**海外商业化兑现同步提速。**MiniMax 2026 年上半年海外收入占总收入比重超六成；Kimi 海外付费用户与 API 收入均增长 400%；AI 内容制作企业昆仑万维 2026 年上半年海外收入 52.03 亿元、占营收 97%，同比增长 51.21%；快手可灵 2026 年第一季度约 75% 收入来自海外；DeepSeek 甚至登顶 2026 年 5 月美国企业软件采购趋势榜，Coinbase 等美国企业直接付费采购。<sup>85</sup>

**算力设施出海稳步铺开。**中国信通院数据显示，2025 年东南亚重

<sup>83</sup> 每日经济新闻：《中国 AI 大模型调用量连续二十二周领跑》，2026 年 9 月 28 日，<https://m.nbd.com.cn/articles/2026-09-28/4592714.html>。

<sup>84</sup> 人民邮电报：《“Token 出海”热潮：伪命题还是真机遇？》，2026 年 3 月 31 日，[https://www.cnii.com.cn/rmydb/202603/t20260331\\_728001.html](https://www.cnii.com.cn/rmydb/202603/t20260331_728001.html)。

<sup>85</sup> 财联社：《国产大模型出海“元年”谁在海外市场淘金？》，2026 年 9 月 1 日，<https://www.cls.cn/detail/2470443>。

点区域算力中心市场规模达 87 亿美元，2026 年将突破 100 亿美元；阿里云、腾讯云、万国数据、秦淮数据、润泽科技等已在东南亚规划建设多座数据中心。<sup>86</sup> 根据广发证券报告，字节跳动、万国数据、腾讯云、润泽科技在东南亚已宣布大型数据中心投资规划，至 2030 年总投资额达到 174 亿美元。<sup>87</sup>

## （二）出海面临的多重障碍

一是欧美市场地缘政治与安全审查。美国构建覆盖创新链、供应链、数据链的“三链协同”管制体系，对 GPU 和 HBM 实施出口管制、限制模型权重出口与数据跨境，并依托盟友体系推动“联盟式围堵”。<sup>88</sup> 欧盟《通用数据保护条例》（GDPR）的数据跨境规则使中国云厂商从境内直接服务欧盟企业近乎不可行，在欧盟境内建设数据中心、实现数据本地化处理成为唯一实际路径；欧盟《网络安全法案》修订草案（CSA 2.0）若生效，还将从供应链安全角度进一步限制其进入欧盟关键领域云服务市场。

二是新兴市场监管呈现碎片化与高不确定性。东南亚、南亚、拉美等新兴经济体在数据本地化、跨境传输、外资准入和网络安全审查等方面标准各异、调整频繁。印度、印尼等国对核心数据跨境设置刚性本地化约束，合规要求趋严。<sup>89</sup> 罚款力度上，巴西、南非等国以境内营收为基数设定上限，整体低于欧盟全球营收 4% 的标准，但拉美

<sup>86</sup> 福建省信息中心（转引中国信通院《东南亚区域算力中心服务商分析报告（2026 年）》），2026 年 6 月，[https://xxzx.fujian.gov.cn/jjxx/xxhdt/202606/t20260603\\_7155599.htm](https://xxzx.fujian.gov.cn/jjxx/xxhdt/202606/t20260603_7155599.htm)。

<sup>87</sup> 广发证券：《算力网专题：AIDC 模块化大势所趋，北美向高渗透率迈进、国内和东南亚加速》，2026 年 8 月 17 日。

<sup>88</sup> 申青青、孙成昊：《“硅幕”已至：美国对华人工智能出口管制与中国应对》，《国际经济合作》2026 年第 1 期，第 1—13 页。

<sup>89</sup> 经济形势报告网：《中国算力出海的战略机遇与风险挑战》，2026 年 4 月 3 日，<http://www.china-cer.com.cn/juece/2026040331537.html>。

部分国家（如智利）已开始向全球营收口径靠拢。外资持股方面，中东主要经济体总体开放，沙特、阿联酋自贸区允许数据中心领域外资全资持有；中亚国家则多以外资自有资金比例作为门槛。

**三是海外数据中心建设面临当地电力与水资源基础设施的客观约束。**电力并网是首要瓶颈，专用变电站与输电线路建设通常需要 18 至 36 个月，一座电厂从规划到投产往往需要 4 至 5 年，远跟不上算力项目落地速度。当地电网接入审批和电力基础设施建设周期，是项目落地的硬约束。水资源方面，马来西亚柔佛州因水资源压力已停止批准基础设施较简易的数据中心，新加坡曾出台暂停审批政策不再受理新建项目。即便芯片供应和合规审批问题得到解决，海外数据中心的建设周期和成本仍面临较大不确定性。

**四是芯片来源面临两难困境。**中国企业在东南亚等地建设数据中心，若选择英伟达芯片，不仅面临出口管制下的许可证审批障碍，还需应对穿透式合规审查。美国商务部工业与安全局（BIS）2026 年 5 月发布的指导明确，先进计算芯片出口至总部位于中国等受限地区的实体，即使接收实体位于境外，也需要出口许可证，中资企业在泰国、新加坡等地租用合规机房训练大模型获批可能性极低。与此同时，美国商务部正在起草新规，拟将管制范围从实体芯片延伸到远程算力调用，直接封堵中国企业通过第三国数据中心远程租用英伟达 GPU 的路径。若转向国产 AI 芯片，虽可规避上述管制风险，但面临生态适配的现实障碍。全球 90% 以上 AI 大模型基于英伟达 CUDA 生态开发，国产芯片在推理侧的应用正在扩大，但在训练端和海外部署场景中，生态迁移成本仍构成实质性门槛。

**五是“境内处理、结果回传”的物理交付模式尚未规模化实现。**所

谓 Token 出海，其理想化叙事是将海外用户的推理请求传回国内数据中心，由国内电力和算力完成计算后回传结果，实现“电力不出境、算力不出国、价值流向全球”。但将海外用户的请求先传输至国内计算再返回结果，涉及跨国数据传输，不仅难以满足大模型推理的时延要求，数据跨境流动的合规性也为这一模式划定了红线。<sup>90</sup> 无论是中国对数据出境的管理规定，还是用户所在国对数据出境的严格限制，都使这一方式面临较高的法律风险与操作成本。<sup>91</sup> 目前真正实现“境内处理、结果回传”的案例，主要依托汕头“来数加工”等特殊试点框架，在物理隔离的算力集群中完成推理后向境外返回结果<sup>92</sup>，尚未形成可规模化复制的通用路径。

**六是前沿 AI 模型分级出口管制带来的政策不确定性。** 2026 年 7 月，商务部牵头与阿里巴巴、字节跳动、智谱等主要 AI 企业召开多场会议，讨论对前沿 AI 模型实施分级出口管制的可能性，核心措施包括限制训练数据出境、限制外国用户下载模型权重，以及按技术能力分级审查（最敏感的前沿模型可能被限制在中国境内使用）。目前规则草案尚未公布，筛选标准和例外情况均不明确，仍处于讨论阶段。这一动向对海外客户使用中国模型构成直接的不确定性。即便管制最终以较温和的形式落地，规则边界的不清晰本身就会抑制海外企业的长期投资决策，尤其是那些依赖本地部署和二次开发的企业。

### （三）算力与模型出海路径选择

<sup>90</sup> 人民邮电报：《“Token 出海”热潮：伪命题还是真机遇？》，2026 年 3 月 31 日，[https://www.cnii.com.cn/rmydb/202603/t20260331\\_728001.html](https://www.cnii.com.cn/rmydb/202603/t20260331_728001.html)。

<sup>91</sup> 贸企通：《中国 AI 产业出海的重点法律问题研究》，2026 年 9 月 7 日，<https://www.eccpit.com/news/Y21zcG86MzE3NjU#1>。

<sup>92</sup> 汕头日报：《统筹推进“人工智能+”行动 打造智能经济“试验场”》，2026 年 7 月 27 日，<https://strb.strtv.cn/content/202607/27/c197108.htm>。

中国算力与模型出海宜沿三条路径推进。

**其一，市场分层，差异化拓展重点区域。**以东南亚、中东、拉美等新兴市场为主攻方向，发挥地理邻近、时延优势和性价比优势，推动中国模型与算力服务优先落地。这些市场本土 AI 能力薄弱，与中国的地缘政治摩擦相对可控，适合作为出海的首选区域。在欧美发达市场，则两条腿走路：一方面以云服务商为先导推进数据中心和智算中心的本地化布局，比如阿里巴巴计划未来 12 个月内在土耳其、芬兰、荷兰建立云区域，并扩大在马来西亚、德国、阿联酋、法国的数据中心布局，全球数据中心容量目标提升至 20 吉瓦；另一方面以 API 性价比渗透开发者工作流，避免与头部厂商正面争夺受监管敏感的大型政企订单。

**其二，模式升级，从模型输出走向主权 AI 共建。**出海的重心应从单纯追求调用量转向帮助新兴市场国家建设本国可控的 AI 基础设施。2025 年底，智谱 GLM 开源模型在马来西亚上线国家级 MaaS 平台，同步启用主权 AI 人才培养实验室，针对当地多语言环境与本地文化优化，保持以主权为核心的数据安全架构，成为第一家在海外实现国家级主权大模型实质性落地的中国公司。<sup>93</sup> 这一实践表明，输出国产芯片、开源模型与本地化部署的整体方案，既能规避数据主权敏感，又能形成长期绑定的生态扎根，是中国方案区别于美国闭源体系的差异化竞争力。

**其三，链条协同，推动全产业链抱团出海。**鼓励芯片、服务器、数据中心、云服务、模型、应用企业组成联合体参与海外项目，把东南亚区域算力中心建设与国内算力网的调度标准、绿电模式、运维体

<sup>93</sup> 智谱微信公众号，2025 年 12 月 31 日，<https://mp.weixin.qq.com/s/3TGqTsfBlfOh6QcDx-MtQ>。

系整体输出，使中国算力网的技术标准与治理规则随产能一同国际化。与此同时，依托绿电直连形成的低成本绿色算力，建设绿色算力出海试点，建立绿色算力认证、绿电溯源与碳足迹核算体系，推动绿色算力从“接入可行”迈向“可出口、可结算、可持续”。<sup>94</sup>

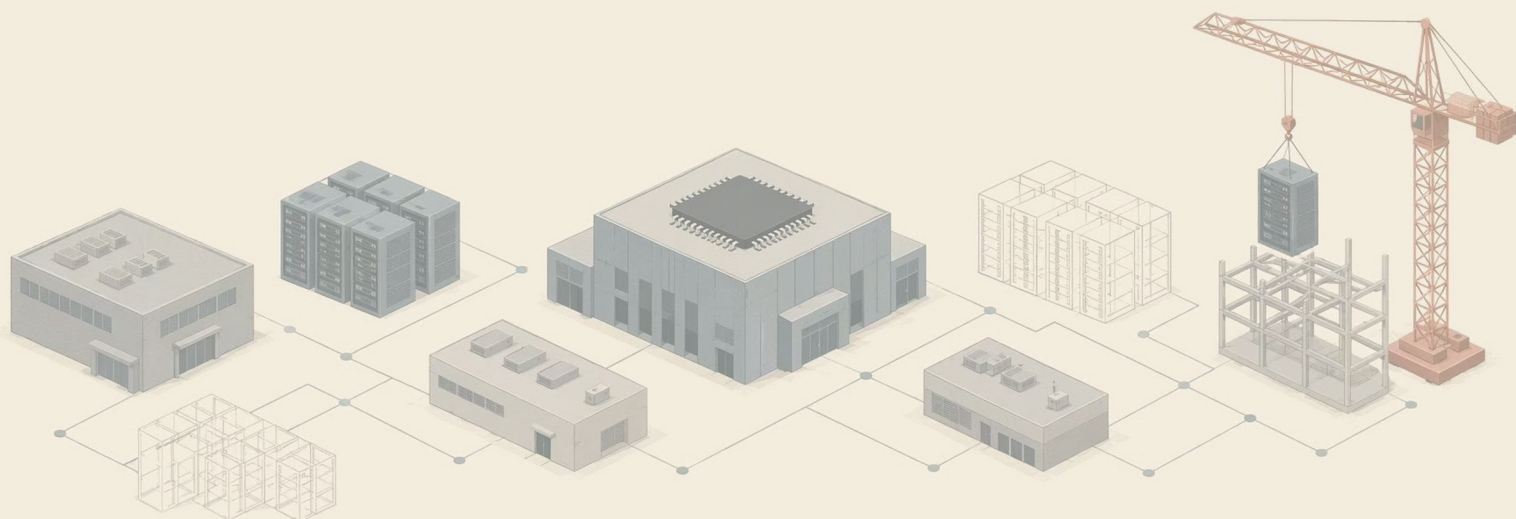
**路径落地需要政策配套。**一是升级出海合规服务。依托国家层面海外综合服务平台，增设算力出海专属服务模块，整合重点市场数据合规、外资准入、税务筹划等信息，为企业提供跨境合规一站式服务。二是完善金融与外交配套。将海外数据中心和智算中心纳入政策性金融机构重点支持范围，发挥国家开发银行、中国银行等机构对海外算力项目的融资支持作用；用好反外国不当域外管辖条例等现有救济工具，完善跨部门协调响应机制，为企业应对歧视性审查提供法律依据和救济渠道。

<sup>94</sup> 同前引：邢庆科、许成东：《算电协同破局：从绿电直连到绿色算力出口》。

# 07

## SECTION 07

### 对策建议



“十五五”时期是我国算力网从规模扩张转向质量效能的关键窗口期。把4万亿元新增投资用好，关键在于提升增量质量、盘活存量资源，让投资跟着产出走。基于这一判断，提出以下六个方面的对策建议。

## （一）以用定建，提升投资精准度

针对同质化建设与利用率偏低问题，建议将“合理布局、规范有序”的国常会要求具体化为可操作的约束机制。

**一是建立利用率承诺制。**六部门2023年行动计划已明确，整体上架率低于50%的地区规划新建项目应加强论证。上海2025年建设导则执行上架率承诺制，要求前期项目两年内IT设备上架率不低于75%。建议把这类约束推广到全国新建智算项目，投运两年上架率仍低于50%的，暂停后续扩建审批。对绑定头部客户的批发型项目，可在规划阶段要求落实意向客户；对面向中小企业的零售型项目，则应降低一次性建设规模，分期投产、以销定建，避免一次性建成后的空置风险。这类约束要见效，前提是地方考核从算力规模转向利用效率，否则承诺容易变成纸面数字。

**二是完善监测基础，让实际负载可核查。**工信部自2025年下半年启动算力态势感知监测试点，并于2026年1月正式发文全面铺开，截至2026年9月，31个省区市已完成平台对接，全国一体化算力统筹监测体系基本建成，上架率监测已基本覆盖。但穿透到上架后实际负载的监测仍处试点阶段，安徽芜湖在国家数据局指导下率先完成算力资源自动化监测试点，实现算力规模、存储规模、资源负载和能耗等核心指标的实时采集与上报，经验尚待推广。建议将监测数据与扩建审批、能耗指标和财政支持挂钩，并接受第三方核验。

三是实行分层分级布局，并与利用率约束联动。明确枢纽节点主攻训练与大规模推理、区域节点主攻行业推理、边缘节点主攻低时延场景的分工。东部严控通用型数据中心扩张，重点发展高端智算与推理算力；西部以绿电优势承接训练与冷数据业务。但分层布局不能成为继续铺摊子的借口，对未达到利用率红线的区域，即使符合功能定位，也应暂缓新建审批。

## （二）完善算力交易与流通机制，促进算力资源高效配置

算力网的价值最终要通过市场机制实现。当前跨区域算力的调度、定价和结算规则尚未统一，富余算力难以通过市场化渠道流向有需求的地区和行业。缺少统一的交易出口，也在一定程度上助长了各地自建自营、重复建设。算力不像公路，建好后不会自动有车来，必须依靠市场机制把分散、异构的算力转化为可购买的标准服务。

一是加快设立全国性算力交易平台。可参照全国政协委员张云泉的建议，依托国家算力互联网服务平台设立独立于供需双方的算力交易所或交易板块，提供集中挂牌、公开竞价、撮合交易服务，先行覆盖智能算力现货交易，适时推出远期合约，建立全国统一的算力价格指数，让算力价格看得见、比得了。<sup>95</sup>

二是统一度量与计价标准。加快制定以有效算力（标称算力×效能系数×满载系数）为核心的计量计费国家标准，推动 Token 计量、卡時計价、服务质量分级的标准化，解决异构算力不可比、不可换的问题。

三是完善并网与调度规则。明确各类算力资源并网的技术规范、

<sup>95</sup>九三学社中央：《全国政协委员张云泉：建议设立全国性算力交易所》，<https://www.93.gov.cn/m/site/content.do?id=783634>。

结算规则与违约责任，将月均调度量、跨域调度占比纳入枢纽节点考核，推动“1+M+N”体系从技术调度网络升级为调度、交易、结算一体化的市场基础设施。

**四是降低跨域传输成本。**建立国家枢纽节点间专项网络传输计费机制，推广弹性带宽、任务式服务、数据快递等模式，组织运营商对算力干线实行公益性定价，把东西部数据传输的路费降下来，这是西部算力盘活的前提条件。

**五是培育专业资源整合平台，以市场化方式提升存量利用率。**趋动科技等企业通过软件将物理 GPU 虚拟化为资源池，将不同用户的算力需求错峰复用到同一批物理卡上，帮助客户将 GPU 利用率从不足 15%提升数倍，再按提升的效益抽成收费。<sup>96</sup> 这类不新增机柜、靠技术提升存量周转率的模式，比单纯新建机房更能快速释放存量价值。对沉淀资产，支持通过算力租赁平台和跨区域调度盘活，允许头部运营商以市场化方式整合低效产能，打通退出通道。

### （三）创新投融资机制，走通“投融管退”全周期

当前社会资本面临“不敢进、不会进”的困境，债务风险也在局部累积。建议从资产证券化、债务工具、绿色金融和风险管理四个方面，构建与算力资产属性匹配的金融支持体系。

**一是扩大数据中心 REITs 发行并分层设计产品。**在首批两单试点基础上，将发行范围从批发型数据中心扩展至现金流稳定的智算中心，鼓励首发与扩募滚动盘活。同步发展持有型不动产 ABS、私募

<sup>96</sup> 上海证券报：《趋动科技王鲲：让每一块 GPU 发挥极致价值》，2026 年 5 月 6 日，<https://www.cnstock.com/commonDetail/710725>。

REITs 等中间层次，为中等质量资产提供过渡性退出通道。<sup>97</sup> REITs 的核心价值不只是融资，更在于通过公开发行人形成对资产质量和运营效率的市场化约束，反向推动项目前期合规与运营质量提升。

**二是发展长周期、低成本的债务工具。**用好超长期特别国债、政策性开发性金融工具对国家级算力枢纽重大项目的支持，推广算力贷、绿色信贷，引导贷款期限向 10 年以上延展以匹配资产回收周期。对算力设备融资租赁实行差异化监管，要求充分披露设备贬值假设，避免金融机构在抵押率评估中低估技术迭代风险。

**三是健全绿色金融支持体系。**中国人民银行等三部门联合印发的《绿色金融支持项目目录（2025 年版）》，已将算力基础设施纳入绿色金融支持范围。发改委等四部门联合印发的《关于促进人工智能与能源双向赋能的行动方案》进一步明确，鼓励金融机构对符合上述目录要求的算力基础设施项目提供资金支持。在信贷产品层面，兴业银行已落地全国首批绿电直连零碳算力产业园绿色贷款，将贷款利率与清洁能源电力使用量、碳减排成效双向挂钩<sup>98</sup>；浦发银行落地全国首单数据中心绿色算力指数可持续发展挂钩贷款<sup>99</sup>。建议在此基础上，将 PUE、可再生能源电量占比等指标纳入授信评价的常规流程，用好用足碳减排支持工具，使绿电使用和能效表现真正转化为可计量、可核证的融资增信。

**四是探索算力期货、算力保险等风险管理工具。**支持金融机构开发算力价格指数衍生品、设备贬值保险、需求违约保险，帮助运营商对冲技术迭代与需求波动风险。算力作为新兴资产类别，其价格波动

<sup>97</sup> 同前引：中国信通院：《算力中心创新融资研究报告（2025 年）》。

<sup>98</sup> 新华财经：《绿色金融创新赋能！兴业银行北京分行助力零碳算力产业升级》，2026 年 9 月 21 日，[https://www.cnfin.com/cmjj-lb/detail/20260921/4472930\\_1.html](https://www.cnfin.com/cmjj-lb/detail/20260921/4472930_1.html)。

<sup>99</sup> 浦发银行 2024 年年度报告。

和残值不确定性远高于传统基础设施，风险对冲工具的缺位是当前社会资本不敢进入的重要原因之一。建议选取算力交易量较大、价格形成机制相对成熟的枢纽节点先行试点，逐步积累定价经验和风险管理数据，再向全国推广。

#### （四）深化算电协同，把绿电优势转化为算力成本优势

算电协同是我国区别于美国的重要优势。美国电力瓶颈突出，多州暂停并网审批，科技巨头不得不直购核电、自建电源。我国电力供给总体充裕，但绿电消纳与跨区调配机制尚未打通，绿电“产得出、送不到、用不上”的问题依然存在。把绿电优势转化为算力成本优势，需要从以下五个方面推进机制突破。

**一是完善绿电直连配套规则。**国家发改委、国家能源局已明确绿电直连项目实行按容（需）量缴纳输配电费，下网电量不再缴纳系统备用费和输配环节的电量电费。下一步应加快明确过网费、备用费、系统运行费的收费细则，建立直连绿电的物理溯源与绿证核发衔接机制，稳定项目投资预期。支持“源网荷储算一体化”项目以聚合主体身份参与电力市场，通过市场化方式回收调峰和备用成本。<sup>100</sup>

**二是稳妥推进多用户绿电直连试点。**2026年5月，国家发改委、国家能源局发布通知，将绿电直连从单用户扩展至多用户，明确优先支持算力设施。建议在算力枢纽节点选择若干园区先行试点，重点验证多用户模式下专线投资分摊、负荷消纳匹配和调度运行的实际效果。多用户模式通过一对多分摊专线投资和运维成本，有利于降低单一用户的进入门槛，但试点中需重点解决多用户之间的负荷曲线差异和费

<sup>100</sup> 同前引：邢庆科、许成东：《算电协同破局：从绿电直连到绿色算力出口》。

用分摊机制问题。

**三是探索隔墙售电在算力场景中的适用边界。**目前隔墙售电推进较为缓慢，面临过网费标准不明确、系统成本分摊机制缺位、与《电力法》“一营业区一供电机构”规定的法律冲突等制度障碍，全国仅少数省份有项目落地。在算力枢纽节点，建议将隔墙售电定位于绿电直连的补充方案，适用于中小规模、分散布点的算力设施就近消纳分布式绿电。可依托零碳园区等载体开展探索，重点厘清过网费标准和结算规则，明确电网企业的安全责任边界和备用费用分摊方式，成熟后再研究是否具备推广条件。

**四是推动算力负荷参与电力调节。**算力负荷具有柔性可调潜力，训练任务可以按实时电价跨区域调度，推理任务则更适合就近消纳绿电。国家能源局等四部门已明确提出推动建立算力与电力互动机制，鼓励算力设施作为负荷侧灵活可调节资源参与电网运行，提升电力系统调节能力。2026 年 5 月，中国联通韶关数据中心等已接入粤能投虚拟电厂运营平台，参与电力现货市场交易与结算，实现了国内首次“算随电动”落地。<sup>101</sup> 建议推广“算随电动”（按实时电价跨区域调度训练任务）与“电随算用”（绿电基地配套算力）两种模式，加快算电协同调度系统建设，使数据中心从单纯的用电大户转变为新型电力系统的柔性调节资源，通过需求响应收益反哺算力成本<sup>102</sup>。参与电力调节的前提是计量和结算体系能够准确区分可调节负荷与刚性负荷，上海已在数据中心专项快速削峰响应中实现 99.3% 的响应准确率，验证了

<sup>101</sup> 证券日报：《算电协同推动虚拟电厂迈入“黄金发展期”》，2026 年 5 月 19 日，[https://www.cpn.cn.cn/news/baogao2023/202605/t20260519\\_1888640.html](https://www.cpn.cn.cn/news/baogao2023/202605/t20260519_1888640.html)。

<sup>102</sup> 通信信息报社：《“电老虎”一口吞下 3600 亿度电 电与算系统耦合打通绿电算力闭环》，2026 年 7 月 3 日，<https://www.hubpd.com/detail/index.html?contentId=8358680908404929286>。

分时计量和响应结算机制的技术可行性<sup>103</sup>。建议在试点园区先行建立分时计量和响应结算机制，为算力负荷规模化参与电力调节积累经验。

**五是健全跨区绿电调配机制。**东部负荷中心面临电力缺口与能耗指标双重收紧，跨区域绿电调配需求迫切。建议加快“西电东送”通道与绿电交易扩容，建立东部算力用户与西部绿电基地的长期购电协议撮合机制。跨区调配的关键在于建立稳定的长期合同关系，而非依赖短期交易，这样才能为西部绿电基地和东部算力用户提供可预期的投资信号。

## （五）保障芯片供应链安全，夯实算力投资的硬件基础

国产芯片的供应能力和生态成熟度，直接决定算力设施的投资成本和运营效率。当前国产芯片在推理场景已具备商业竞争力，但训练端单卡算力仍有差距，HBM 供应链对外依赖尚未解除，软件生态的迁移成本仍然较高。建议从以下三个方面降低算力投资的芯片供应风险。

**一是加快 HBM 国产化验证和先进封装产能建设。**HBM 是 AI 加速卡中价值量最高的单一组件，其供应链对外依赖尚未解除。美国对华 HBM 出口管制持续收紧，国产 AI 芯片所需的 HBM 仍主要依赖进口。建议支持国内存储企业加快 HBM 工程验证和客户认证，同步推进先进封装产能建设，发展 Chiplet、3D 堆叠等绕开制程限制的技术路线。

**二是推动国产芯片与主流模型框架的适配。**软件生态的迁移成本是当前国产芯片在训练端和海外部署场景中面临的主要障碍。建议发

<sup>103</sup> 上海证券报：《达卯科技：“算电协同 2.0”平台响应准确率高达 99.3%》，2026 年 7 月 31 日，<http://www.cnstock.com/commonDetail/753793>。

挥信创采购和政务、金融、能源等场景的牵引作用，以应用反哺生态。在政府引导基金和国资投资决策中，可将适配模型数、开发者规模作为投后评估的参考维度；在信创采购中，仍以安全合规和场景适配度为核心标准，避免将生态规模作为准入门槛。同时，支持建设国产芯片与模型框架的公共适配平台，降低中小芯片企业和模型开发者的适配成本。武汉国产加速芯片软件适配中心的实践表明，平台化适配可将成本降至企业单独适配的五分之一<sup>104</sup>，这一模式的推广具有显著的杠杆效应。

**三是建立算力设施芯片选型分级指引。**明确推理场景优先采用自主可控方案、训练场景混合部署的策略，避免“一刀切”排斥国产或“一刀切”强制国产两种极端。分级指引的定位是整合已有体系，而非另建标准：安全等级维度直接对接国家信息安全测评中心发布的 AI 训练推理芯片安全可靠测评结果，设施层级维度参照城市、园区、边缘的分类布局，场景维度按模型参数量和部署位置交叉划分。这一指引以“技术选型参考”而非“产业政策目录”的形式发布，嵌入信创采购和算力券的评审流程，为算力设施的投资决策提供清晰的芯片选型依据。

## （六）培育需求侧生态，以应用繁荣锚定算力投资

算力投资的长期合理性取决于应用层收入兑现，必须坚持供需两端协同发力。

**一是降低用算门槛，释放长尾需求。**面向中小企业和科研机构发放算力券，通过中国算力平台提供普惠价算力，降低 AI 应用试错成本。当前算力券的发放和使用多集中在头部企业，中小企业和科研机

<sup>104</sup> 长江日报：《三大科技巨头携手，国产加速芯片软件适配中心落地武汉》，2026 年 6 月 12 日，[http://news.cjn.cn/whpd/ttxw\\_19945/202606/t5304457.htm](http://news.cjn.cn/whpd/ttxw_19945/202606/t5304457.htm)。

构实际受益有限，建议简化申领流程、降低使用门槛，把一卡难求的长尾需求转化为有效市场。

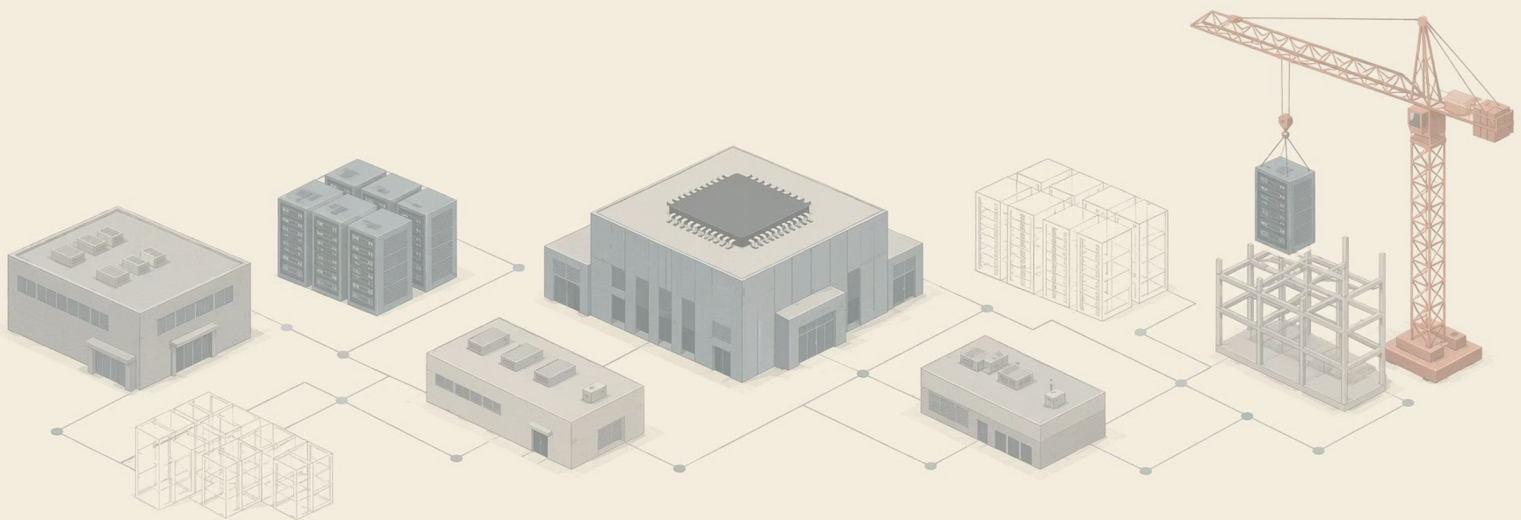
**二是以政府应用牵引市场。**政务、医疗、教育、制造等领域开放 AI 应用场景，推行“人工智能+”采购，优先采购基于自主可控算力的模型服务，形成应用、收入、再投资的良性循环。政府场景的价值不仅在于订单本身，更在于通过真实业务反馈帮助模型和芯片企业打磨产品、积累案例，为后续市场化推广奠定基础。

**三是支持商业模式创新。**鼓励模型厂商探索分层定价、按效付费、行业订阅等模式，支持套餐化、路由聚合、效果打包等 Token 运营层业态发展，引导行业从价格竞争转向价值竞争。对 AI 应用企业实施首购首用等支持，帮助其跨越付费转化阶段。

# 08

## SECTION 08

### 结语



展望“十五五”，算力网建设大概率呈现三个趋势性走向。一是投资逻辑从规模竞赛转向效能竞赛，随着全国一体化算力监测调度体系完善和利用率约束强化，算力投资将更加集中于枢纽节点、高端智算和推理算力，低效供给将有序退出。二是价值链条从卖算力转向卖Token，算力的计价单位、交易形态和商业模式都将围绕Token经济重构，绿电、算力、Token贯通的价值主轴将打开价值外溢的国际空间。三是竞争焦点从单点能力转向体系能力，芯片、集群、网络、调度、电力、模型的全栈协同能力，将取代单一要素优势成为国家和企业竞争力的决定变量。

把握这三条主线，关键在于坚持“以用为本”。“十五五”时期算力网建设新增直接投资约4万亿元，是扩大有效投资、培育新质生产力的重要抓手，投资方向明确，力度也需要保持。决定这笔投资最终成效的，不是设施规模的大小，而是算力能否被真实需求有效消化，能否转化为可持续的产业竞争力。以真实需求牵引投资，以统一市场配置资源，以机制创新盘活存量，以应用繁荣锚定未来，算力网建设才能在保持投资强度的同时，走出一条高质量、高效益的发展路径。

算力网建设需要政府与市场协同发力。市场化企业是投资主力，国有平台承担基础性和兜底性投资功能，政府的职责则在于把制度和服务做到位。统一的市场规则、透明的运行数据、可预期的收益环境，与投资支持同样重要。唯有如此，才能把算力网真正建成像电网一样高效运转、普惠易用的国家级基础设施，为人工智能时代的国家竞争力提供坚实支撑。

# 和 坤 经 济 笔 记

HEKUN ECONOMIC NOTES · RESEARCH REPORT

## 研 究 合 作 联 系

电 话 159 8580 5268

邮 箱 linhekun@hotmail.com

## 版 权 声 明

本研究文件系依据公开可查信息、数据统计分析及专业研究方法独立编纂而成，仅用于行业探讨、学术交流及决策参考之目的，不构成任何商业指引、投资建议、法律意见、税务筹划或其他专业领域的具体操作依据。本报告为免费提供，未经作者事先授权，任何单位或个人不得擅自转载、复制、修改、节选本报告的全部或部分内容。版权归属厦门美知经济咨询有限公司。