

■ DeepSeek 热浪

DeepSeek 带火算力市场

◎ 记者 李兴彩

算力卡需求增加，此前被冷落的 H20 芯片成为香饽饽；客户尝试大模型服务的意愿增强，智算中心空置率正在下降；算力租赁价格开始上涨，运营商大力推广搭载华为昇腾芯片的 DeepSeek 大模型一体机……

随着 DeepSeek 大模型被广泛适配和接入，算力硬件及租赁市场迎来翻转，需求增长快速显现。多位业内人士表示，DeepSeek 显著降低了全社会数智化的门槛和成本，基于各种应用场景的需求会出现爆发式增长，这种端侧 AI 增长又反向促进云侧 AI 发展，进而形成云、端 AI 算力均快速增长的态势。

算力全链条涨价

在千行百业争相接入 DeepSeek 后，算力市场正在发生显著变化。

“算力市场很火爆，我们在全力推广搭载华为昇腾芯片的 DeepSeek 大模型一体机，全力推销算力服务。”某运营商人士介绍，蛇年开工以来，DeepSeek 大模型带动客户尝试算力服务的意愿大增，现在运营业务全部都融合了语言大模型技术和服务。

事实上，从硬件到算力租赁市场，再到算力需求结构，都在 DeepSeek 影响下呈现出新现象。

“算力卡需求开始增加，市场上之前不太受欢迎、闲置的 H20 芯片（英伟达专供中国市场的产品）都已经陆续售出，英伟达 4090 芯片（显卡）的需求也有所增加。”优刻得相关负责人在接受上海证券报记者采访时介绍，在算力租赁方面，智算中心的空置情况正在得到改善，随着闲置资源的减少，算力租赁的价格已经略有上涨，但涨幅还不太大。

H20 芯片是英伟达在 2023 年底针对中国市场推出的一款人工智能芯片，也是目前中国买家能在合法渠道上买到的用于大模型训练的最先进的 GPU 芯片。但鉴于其性能较低，此前被市场视为“鸡肋”产品，并不受欢迎。

“推理和渲染及各类 AI 场景领域需求显著增长，像立昂技术这样提供灵活算力的公司就比较受益。”立昂技术常务副总裁王子璇在接受记者采访时补充道，具体来看，当前整体算力租赁数量是增长的，但不同品牌硬件接到的需求也不尽相同。另外，在智算中心租赁价格方面，环京、华南等这些适合做推理的热算力，更能感受到价格上涨。

王子璇介绍，从立昂技术来看，公司原来客户更多集中在电竞、游戏等渲染需求上，现在大的 B 端客户、G 端客户明显增加，来寻求资源的平台级客户也开始变多。“立昂技术提供一个为期 3 天的免费测试期，让客户充分感受 DeepSeek 在业务上的转化效果，最近来咨询的客户明显增加。”王子璇表示。

DeepSeek 还带动私有云市场和大模型本地化部署显著增长。

“我看到周围的企业家朋友都在更换自己的 IT 设备，明显提升了预算。”泰凌微董事长王维航表示，DeepSeek 灵活的模型蒸馏机制，降低了对端侧硬件计算资源和存储的要求，完全支持部署在企业本地的服务器或私有云环境中，这对于那些重视数据保密的企业来说非常具有吸引力，也可以更好地提升其业务效率。

云侧与端侧相互推动

“短期的云侧训练需求下降，端侧推理需求大

幅上升；但长期来看，云侧、端侧的需求相互推动，均将大幅提升。”上述优刻得相关负责人强调。

谈及 DeepSeek 对云计算市场的影响，王维航表示，DeepSeek 大模型大幅降低企业和社会的数字化、智能化门槛，基于各种应用场景的需求会出现爆发式增长，端侧 AI 又对云侧产生促进作用，在这个互相促进的双向过程中，云计算市场快速增长。云侧 AI 芯片和端侧 AI 芯片都具有巨大的市场空间。

Canalys 最新报告显示，2024 年全球云支出同比增长 20%，从 2023 年的 2677 亿美元增至 2024 年的 3213 亿美元。AI 模型的快速扩展成为这一增长的核心驱动力，加速了云计算的普及。Canalys 预测，2025 年全球云基础设施服务支出将再增长 19%。

在快速增长的云计算市场中，推理算力需求将表现得更加亮眼。

在 AI 技术的实际应用落地中，推理的需求规模往往是训练需求的 5 倍至 10 倍。随着 AI 应用场景不断开拓，推理场景的算力需求进入快速增长期。IDC 认为，推理在未来几年会占有大部分市场，与训练共存；云端推理占算力的比重将逐步提升，预计到 2026 年推理占 62.2%，训练占 37.8%。

多位业内人士在接受采访时强调，DeepSeek 带动千行百业加速智能化，这种“长尾”需求，有望打破半导体产业的周期性，带动产业持续增长，给中国半导体产业带来新的发展机遇。

“在 DeepSeek 带动下，国产算力芯片的价格也有所上浮。”王子璇介绍，当前国产芯片中，华为昇腾价格最高，几乎是供不应求；沐曦等公司的算力芯片价格也有所上涨。

记者了解到，当前适配 DeepSeek 蒸馏模型最好的国产 GPU 芯片是华为昇腾 920B。另外，在国资智算中心适配国产 GPU 芯片后，更多的非国资智算中心也在积极测试、配置海光信息、寒武纪、沐曦、壁仞、燧原科技等国产 GPU 芯片。

“DeepSeek 带动的长远、持续需求，有望打破半导体产业的周期性，产业持续增长也将吸引更多的人力和财力资源进入行业。”王维航表示，在

DeepSeek 的过程中，相关软件生态的发展得到推动，有助于解决国内生态不完善的问题。

燧原科技创始人、董事长赵立东接受记者采访时表示，在迈向通用人工智能（AGI）的道路上，算法模型与硬件架构会不断创新优化，二者相辅相成。DeepSeek 的成功是 AI 系统软硬件深度协同的结果，也指明了未来 AI 系统的发展方向。只有模型算法与算力芯片更紧密结合，才能在技术与商业两方面实现闭环，从而为大模型商业化部署提供更具市场竞争力的解决方案。

国产 AI 芯片公司紧抓机遇

2 月以来，燧原科技完成了对 DeepSeek 全量模型的高效适配，DeepSeek 全量模型已在其庆阳、无锡、成都等智算中心完成了数万卡的快速部署，将为客户及合作伙伴提供高性能计算资源；摩尔线程也宣布将开放自主设计的夸娥（KUAE）GPU 智算集群，全面支持 DeepSeek 系列模型的分布式部署。

除了国产 AI 芯片算力集群的部署，天数智芯、摩尔线程等公司均推出了 DeepSeek 一体机。据天数智芯方面介绍，其一体机搭载 DeepSeek 各种参数的模型，采用软硬件协同的设计模式，与国内 OEM 厂商、服务器、操作系统等生态伙伴共建，实现多种配置、开箱即用，推动 DeepSeek 在各应用场景落地。

对于 AI 芯片的需求，赵立东认为，当前，训练集群依然是 AI 大模型发展的核心要素，同时随着模型普及和应用拓展，推理市场将被充分激活，推理算力需求将快速放大。

端侧 AI 机遇下，中国半导体产业有望通过场景驱动、算法优化与定制化架构创新实现突破，但需在基础研究、开源生态和产业协同上持续投入。“随着 AI 发展，所有的电子产品都值得重新再设计一遍，这对半导体产业是巨大推动力。”

“国内 AI 芯片企业要积极抓住 DeepSeek 带来的发展机遇，加速推进在工业、医疗、教育等各个垂直领域的落地应用。”赵立东说。

“随着 DeepSeek 带来的应用热潮，市场算力需求大增。天数智芯产品凭借算力性能高、通用性强、开箱即用、性价比高等优势，也将获得更多应用机会。”天数智芯有关负责人说，公司客户群体广泛，在政府、运营商、金融、教育、医疗等领域均有广泛应用。

近日，AI 基础设施公司无问芯穹与七家国产芯片厂商达成合作，打通 DeepSeek-R1、V3 在壁仞科技、海光信息、华为昇腾、摩尔线程、沐曦、燧原科技、天数智芯等七个硬件平台的多芯片适配优化，开发者可以通过无问芯穹 Infini-AI 异构云平台，一键获取 DeepSeek 系列模型与多元异构国产算力服务。

季昕华告诉记者，优刻得已同多家国产芯片厂商完成了对 DeepSeek 满血版和蒸馏版开源模型的适配工作。从测试结果来看，现在可以用起来了。后续，优刻得也会继续同国产芯片公司合作，在“能用”的基础上进行优化，达到“好用”；通过使用量的扩大降低成本，实现“便宜用”，提升使用体验。

上市公司与 AR 眼镜厂商双向奔赴

◎ 记者 柴刘斌

近期，Rokid 创始人祝铭明在一场大会上凭借 AR 眼镜的提词功能，完成了一场脱稿演讲，AR 眼镜进入更广泛的大众视野，完成一次破圈。

资本与产业端同样活跃。近日，海信视像、蓝思科技陆续宣布与 AR 眼镜厂商达成深度合作，另有多家上市公司透露其在 AR 领域的最新布局。

有券商分析师在接受上海证券报记者采访时表示，头部企业加大研发投入，产业链紧密合作，资本不断涌入，都显示出各方对 AI/AR 眼镜发展机遇的一致看好，AI/AR 眼镜产业的商业化有望加速提质。

上市公司纷纷布局 AR 赛道

2 月 25 日，显示产业龙头海信视像与全球 AR 领军企业 XREAL 宣布达成深度战略合作。据了解，双方联合研发的首款 AR 高端观影产品将于今年下半年发布，AI 技术的深度赋能成为关键支点。

双方表示，将在 AR/AI 眼镜领域开展技术协同、生态共建与全球市场拓展。此次合作标志着显示技术与空间计算的跨界融合进入新阶段，有望加速消费级 AR 设备的渗透率提升，重构全球 XR 产业竞争格局。

2 月 22 日，Rokid 官微表示，公司与全球领先的消费电子制造商蓝思科技达成深度战略合作，共同推动 AI 眼镜产业链的技术革新与商业突破。此次合作不仅标志着 AI 眼镜从单点技术创新迈向全链条生态构建，也为 2025 年全球 AI 眼镜出货量的爆发式增长提供坚实支撑。

当前，在 AR 行业发展浪潮中，雷鸟创新、XREAL、Rokid 和影目科技被称为“AR 四小龙”。它们凭借入局早的先发优势，在竞争中脱颖而出，其产品占据国内 AR 市场大部分份额。根据 CINNO Research 数据，雷鸟创新、XREAL、Rokid 分别占据中国消费级 AR 出货量占比前三名。

在技术上，四家公司均有各自的技术特色和研发实力。如雷鸟创新背靠 TCL 显示技术，XREAL 在海外市场技术认可度较高，Rokid 拥有自研操作系统，影目科技独立开发了全球领先的轻量级 SLAM 算法等。

除了与国内 AR 眼镜厂商官宣牵手，还有一些上市公司在互动易平台中透露与国内外厂商的合作情况。例如，三安光电在 2 月 25 日表示，公司 Micro LED 产品与国内外 AI/AR 眼镜领域终端厂商紧密合作，产品已送样验证，部分客户验证结束后进入方案升级优化阶段。

记者注意到，上市公司布局 AR 产业的方式呈现多样化，除了直接与 AR 眼镜厂商合作，也不乏与产业链上下游的同行联手，整合资源，共探发展机遇。

例如，2 月 21 日，天科合达与慕德微纳签署投资合同，共同出资成立合资公司。双方将在 AR 衍射光波导镜片技术研发与市场推广方面展开深度合作，共同推动 AR 行业的技术创新与应用落地。

锁定行业风口落子布局

“AR 眼镜是继电视、手机后的第三代计算平台入口，我们正将 50 年显示技术积淀转化为近眼显示领域的关键参数标准和创新的消费级产品。”海信视像总裁李伟的一番话，道出了公司布局 AR 领域的底层逻辑。

记者梳理上述合作案例发现，技术与资源互补、拓展市场与业务、应对行业趋势与风险成为上市公司寻求深度战略合作的主要原因。

AR 眼镜厂商选择合作对象十分重视产业共识。雷鸟创新创始人兼 CEO 李宏伟接受采访时表示，公司在挑选供应链合作对象时，除了考量能力、技术等方面外，还十分看重双方是否具备共识，即愿意在 AI+AR 眼镜上做真正面向用户的深度投入。

“AI+AR 行业目前仍处于较早期的阶段，与手机等成熟品类相比，市场上并无成熟的解决方案，需要我们与供应链进行深度合作和定制。所以，双方是否能达成这一共识，同时具备相关能力非常关键。”李宏伟强调说。

上市公司纷纷携手 AR 眼镜厂商或行业公司，透露出何种产业趋势？

一位券商分析师对记者表示，近期市场动态及相关企业的动作均表明，市场对 2025 年 AI 眼镜产业趋势高度认同。头部企业加大研发投入、产业链紧密合作、资本不断涌入，显示出各方对 AI/AR 眼镜发展机遇的看好，产业的商业化进程有望加速提质。

郭展凯制图