

市场研究部

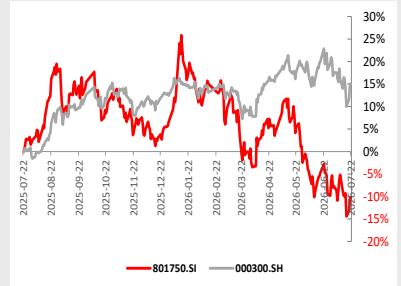
2026 年 7 月 22 日

WAIC 2026 超节点产业观察

看好

市场表现截至

2026.7.21



数据来源：Wind，国新证券整理

相关研究

分析师：钟哲元
登记编码：S1490523030001
邮箱：zhongzheyuan@crsec.com.cn

事件

2026 世界人工智能大会（WAIC）于 7 月 17 日至 20 日在上海举行，华为、中科曙光、阿里巴巴等头部厂商集中落地全系列超节点与超集群方案。

核心观点

2026 年世界人工智能大会清晰印证了国产算力产业的核心转向：竞争主轴已从单芯片制程追赶全面切换至超节点架构主导的系统级协同竞争。超节点通过专用高速互联协议与定制化硬件架构，将数十至数千颗计算芯片整合为统一编址、低时延、高带宽的协同计算系统，直击传统集群的“通信墙”与“内存墙”痛点。

本届 WAIC 上，华为 Atlas 950 SuperPoD 首次真机亮相，1024 卡版本可提供 1 EFLOPS FP8 算力、256 TB 全局统一内存、3 μ s 超低往返时延，满配系统由 128 个计算柜和 32 个互联柜组成，搭载 8192 颗昇腾 950 DT 芯片，总算力达 8 EFLOPS，计划 2026 年 Q4 上市。中科曙光落成全国产十万卡超集群“曙光 8000”，单机柜功率突破 900kW、PUE 降至 1.04。阿里巴巴磐久 AL128 单柜部署 128 颗芯片，实现 Pb/s 级带宽与百纳秒级延迟。新华三 UniPoD S80000 兼容多芯片生态，最高可扩展至 16384 卡。中兴通讯发布 OEX 正交无背板 Matrix 超节点，联合多家国产 GPU 厂商打造开放生态。风冷超节点方案同步面世，覆盖轻量化推理场景。

在芯片层面，国产算力走“架构与封装创新替代制程追赶”路线，通过 Chiplet 多芯粒、2.5D/3D 堆叠、混合键合等技术突破内存墙约束，芯片级先进封装与系统级超节点互联形成两级互补。运营层面，算电协同机制推动行业核心指标从 PUE 转向 TPW，液冷成为高端集群标配。

投资线索

超节点产业链投资逻辑已从主题催化进入订单与业绩兑现阶段，国产算力替代已由政策驱动转向市场驱动。高速互联是确定性最高的增量环节，连接器、光模块量价齐升逻辑明确。先进封装是长期成长空间最广阔的核心赛道，Chiplet 封测、2.5D/3D 堆叠、混合键合相关设备材料及高端封装基板的国产替代空间广阔。整机与系统集成环节，绑定头部芯片生态、具备大规模集群交付能力的厂商将持续受益。上游高端半导体材料、检测设备、高速 PCB 等领域同样蕴含替代机会。建议沿技术壁垒与客户绑定深度两条主线优选标的。

风险提示

1、技术发展不及预期；2、市场竞争加剧；3、地缘政治影响。

证券研究报告

目录

一、算力竞争从单卡性能转向系统整体效率	3
二、超节点方案密集亮相	3
三、先进封装突破制程约束，算电协同优化运营效率	4
四、投资建议	4
五、风险提示	5

一、算力竞争从单卡性能转向系统整体效率

过去十余年，全球算力产业基本沿着摩尔定律路径演进，依靠制程工艺迭代提升单芯片晶体管密度，进而拉动峰值算力增长。但随着大模型参数规模向十万亿级迈进，叠加智能体应用带来的长上下文、高并发推理需求爆发，传统松耦合服务器集群的结构性短板已全面显现。跨节点通信时延、离散内存寻址开销、调度损耗等问题，正在大量吞噬单卡性能提升带来的实际收益。AI 集群中约 80% 的能耗用于数据搬运而非有效计算，单卡制程迭代的边际收益正持续收窄。

超节点架构正是为解决这一核心矛盾而生。其本质是通过专用高速互联协议与定制化硬件架构，将数十至数千颗计算芯片在逻辑上整合为统一编址、低时延、高带宽的协同计算系统，在物理形态上由多台设备组成，在运行逻辑上则等效于一台超级计算机。该架构直击“通信墙”与“内存墙”两大行业痛点，通过统一内存编址、超低时延互联与高密度部署，最大限度降低数据搬运损耗，将芯片的理论峰值算力转化为可实际交付的训练与推理吞吐。本届大会全行业的集体跟进表明，算力竞争的胜负手已从单颗芯片的跑分参数转向整套系统的组织效率与有效产出，这是一次产业范式的本质切换，而非单一产品的迭代升级。

二、超节点方案密集亮相

本届 WAIC 上，华为 Atlas 950 SuperPoD 首次以真机形态公开亮相，是当前国产超节点路线的标杆产品。现场展出的 1024 卡版本可提供 1 EFLOPS FP8 算力，搭载 256 TB 全局统一内存编址空间，依托自研灵衢（Unified Bus）互联协议实现 3 μ s 的超低往返时延。灵衢互联协议采用机柜内电气正交背板与跨机柜光模块相结合的全互联设计，实现 TB 级 NPU 互联带宽。统一内存编址打破了物理节点的内存边界，任意计算单元可直接寻址全系统内存，将远端内存访问的开销压缩至接近本地内存的水平。按照完整产品规划，Atlas 950 满配系统由 128 个计算柜和 32 个互联柜组成，搭载 8192 颗昇腾 950 DT 芯片，总算力可达 8 EFLOPS FP8，总内存容量 1152 TB，全系统互联带宽高达 16.3 PB/s，训推吞吐较前代产品实现十余倍跃升。该方案计划于 2026 年 Q4 正式上市。对比海外主流方案，Atlas 950 支持跨机柜的全域统一内存池化，目前已与头部大模型厂商完成全栈适配，训练到推理的全流程均基于昇腾算力与自研计算架构运行，实现了硬件到软件的国产化闭环。

除华为昇腾外，其他厂商的超节点方案也在 WAIC 2026 集中亮相，呈现出“液冷高密、正交架构、开放生态”三大技术取向。中科曙光落成中国首个全国产十万卡 AI 超集群“曙光 8000（登峰）”，依托曙光数创 C8000 V3.0 兆瓦级相变浸没液冷基础设施整体解决方案，实现单机柜功率突破 900 kW、PUE 降至 1.04，将国内智算规模从万卡时代推入十万卡全新阶段。阿里巴巴磐久 AL128 超节点服务器单柜部署 128 颗芯片，通过正交无缆化架构和自研 ALink System 互连协议，实

现 Pb/s 级带宽与百纳秒级延迟，同等 AI 算力下推理性能较传统架构提升 50%，旨在最大化 GPU 集群的协同计算效率，适配公有云弹性部署需求。紫光股份（新华三）UniPoD S80000 兼容多芯片生态，覆盖从 32 卡到 1024 卡全系列，最高可扩展至 16384 卡，通过软硬协同优化预期可将大模型训练性能提升 70%、推理性能提升 3 倍。中兴通讯发布 OEX 正交无背板 Matrix 超节点，以计算托盘与交换托盘垂直交叉连接实现“零线缆、无背板”，信号传输路径缩短、互联成本大幅下降，并联合壁仞、沐曦、燧原、天数智芯、曦智科技等打造多元 GPU 开放生态。此外，风冷超节点方案同步面世，无需改造现有机房即可实现算力升级，覆盖轻量化推理场景，进一步拓宽了超节点的落地边界。

三、先进封装突破制程约束，算电协同优化运营效率

在芯片技术层面，受外部供应链与制程约束，国产算力普遍走“架构与封装创新替代制程追赶”的技术路线，这与超节点的系统级优化恰好形成两级互补。传统 2D 芯片架构中，数据在逻辑单元与显存之间的搬运时延、功耗占比极高，是制约算力效率提升的核心瓶颈，即行业常说的“内存墙”。先进封装通过 Chiplet 多芯粒、2.5D 中介层、3D 垂直堆叠、混合键合等技术手段，将存储单元与逻辑芯片、多个计算芯粒近距离甚至垂直键合，大幅缩短数据传输路径，提升带宽密度与能效比，本质是以封装架构创新换取等效制程提升。混合键合技术可将互连间距从数十微米级压缩至亚微米级，带来数量级的互连密度提升，是突破内存墙的核心技术方向；Chiplet 多芯粒架构则可通过成熟制程的芯粒拼接，实现高端制程芯片的等效算力，同时提升良率、降低制造成本。当前国产算力芯片已普遍沿这一路线推进，昇腾系列芯片采用 Chiplet 多芯粒架构与 2.5D 扇出封装，通过多颗计算芯粒合封实现算力倍增，同时配套自研高带宽内存方案，在成熟制程基础上实现了访存带宽的持续提升。整体来看，芯片级的先进封装创新与系统级的超节点互联形成了清晰的互补逻辑：前者解决芯片内部的数据搬运损耗，后者解决集群之间的通信损耗，两级优化共同压低系统总时延，提升有效算力产出。

在运营体系层面，超节点高密度部署带动散热方案同步升级，液冷成为高端集群的标配配套。更核心的变化是算电协同机制的落地——通过空间布局优化、分时电价调度、能源与算力协同调度，可显著提升单位电力的有效 Token 产出。行业核心经营指标正在从传统的 PUE 转向 TPW（Token per Watt），即每瓦电能产出的有效 Token 数量，将能源成本与业务收益直接绑定，让超节点的规模化部署具备了更清晰的商业闭环。算电协同带来的节能设备、智能调度软件、液冷方案等环节，同样蕴含着明确的投资机会。

四、投资建议

超节点是本届 WAIC 产业最重要的主线之一，投资逻辑已由主题催化切至订单与业绩兑现阶段。IDC《2025 年度中国云端 AI 加速器市场报告》显示，2025 年中国市场 AI 加速卡总交付量 400 万片，其中国产厂商交付 165 万片，市场份额从 2024 年的约 30% 跃升至 41%；据 IDC 数据，2026 年 Q1 中国 AI 芯片市场中国产 AI 芯片份额突破 55%，同比提升 21 个百分点。国产算力替代已由政策驱动转向市场驱动，为产业链业绩兑现提供了总量基石。

高速互联是确定性最高的增量环节。超节点通过高速互联协议与专用交换芯片，将数十至数百颗算力芯片在逻辑上整合为低延迟、高带宽的协同计算系统，柜内全互联、柜间光互联成为两大硬件刚需。这一架构使交换芯片与 GPU 配比从传统 1:20 跃升至 1:2 甚至 1:4，单机柜连接器与高速线缆用量、单台价值量提升 10 倍以上，800G 及以上速率光模块量价齐升逻辑明确。具备技术壁垒且与头部客户深度绑定的供应商将最先兑现订单弹性。

先进封装是长期成长空间最广阔的核心赛道。"韬定律"指出衡量算力进步的核心标尺不再是制程节点而是完成任务的总时间，先进封装正是压缩芯片内部数据搬运时间、突破内存墙约束的核心手段。在制程迭代边际收益下降、外部制程受限的背景下，先进封装已从配套环节上升为算力竞争的核心瓶颈，Chiplet 封测、2.5D/3D 堆叠、混合键合相关的设备、材料与封测厂商，以及高端封装基板、高带宽存储封装配套材料，将直接受益于国产算力芯片的迭代放量，国产替代空间广阔。

整机与系统集成是需求落地最直接的受益环节。绑定头部芯片生态、具备大规模集群交付与调优能力的服务器与系统集成厂商，将在系统级竞争中占据更大优势，行业份额持续向头部集中，2026 年中报逐步验证 AI 服务器整机、网络交换机等环节业绩高速增长，超节点批量交付将进一步强化景气度。上述技术升级向上游传导，高端半导体材料、检测设备、高速 PCB 等环节中具备技术壁垒的国产厂商，也将持续获得替代空间。

五、风险提示

- 1、技术发展不及预期；
- 2、市场竞争加剧；
- 3、地缘政治影响。

投资评级定义

公司评级		行业评级	
强烈推荐	预期未来 6 个月内股价相对市场基准指数升幅在 15%以上	看好	预期未来 6 个月内行业指数优于市场指数 5%以上
推荐	预期未来 6 个月内股价相对市场基准指数升幅在 5%到 15%	中性	预期未来 6 个月内行业指数相对市场指数持平
中性	预期未来 6 个月内股价相对市场基准指数变动在-5%到 5%内	看淡	预期未来 6 个月内行业指数弱于市场指数 5%以上
卖出	预期未来 6 个月内股价相对市场基准指数跌幅在 15%以上		

免责声明

钟哲元，在此声明，本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。

本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿等。在本人所知情的范围内，本人所在机构、本人以及本人的利害关系人与本报告所评价或推荐的证券不存在任何利害关系。国新证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格，以下简称本公司）已在知晓范围内按照相关法律规定履行披露义务。本公司的资产管理和证券自营部门以及其他投资业务部门可能独立做出与本报告中的意见和建议不一致的投资决策。本报告仅提供给本公司客户有偿使用。

本公司不会因接收人收到本报告而视其为客户。本公司会授权相关媒体刊登研究报告，但相关媒体客户并不视为本公司客户。本报告版权归本公司所有。未获得本公司书面授权，任何人不得对本报告进行任何形式的发布、复制、传播，不得以任何形式侵害该报告版权及所有相关权利。

本报告中的信息、建议等均仅供本公司客户参考之用，不构成所述证券买卖的出价或征价。本报告并未考虑到客户的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时可就研究报告相关问题咨询本公司的投资顾问。本公司市场研究部及其分析师认为本报告所载资料来源可靠，但本公司对这些信息的准确性和完整性均不作任何保证，也不承担任何投资者因使用本报告而产生的任何责任。本公司及其关联方可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务，敬请投资者注意可能存在的利益冲突及由此造成的对本报告客观性的影响。

国新证券股份有限公司市场研究部

地址：北京市朝阳区朝阳门北大街 18 号中国人保寿险大厦 11 层（100020）

传真：010-85556155 网址：www.crsec.com.cn