

市场研究部

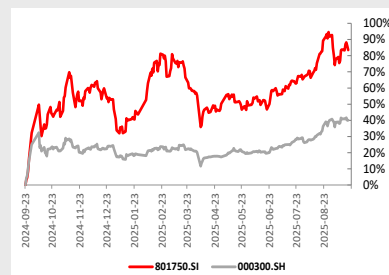
2025 年 9 月 22 日

计算机行业周度：华为昇腾 AI 芯片与超节点技术

看好

市场表现截至

2025.9.19



数据来源：Wind，国新证券整理

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

本周计算机行业指数表现

本周（9.15-9.19）计算机（申万）板块下跌 0.16%，沪深 300 指数下跌 0.44%，计算机板块跑赢沪深 300 指数 0.29 个百分点。和申万其他行业对比，计算机行业涨幅排名位列第 14 位。

本周涨幅前 3 名分别为卡莱特（30.19%）、信息发展（25.71%）、有棵树（25.05%），跌幅前 3 名分别为*ST 东通（-65.05%）、新炬网络（-18.88%）、纬德信息（-15.33%）。

本周关注

华为昇腾 AI 芯片与超节点技术

在 2025 年全联接大会上，华为发布了昇腾 AI 芯片路线图及超节点战略，标志着国产算力技术向高端 AI 市场迈进。该战略不仅涵盖芯片迭代，更涉及互联协议、软件生态与基础设施的全方位创新，旨在构建自主可控的 AI 算力体系，并挑战英伟达主导的 CUDA 生态。

昇腾芯片持续迭代，从 2026 年的 950 系列到 2028 年的 970 系列，算力不断提升。950 系列针对推理与训练场景分别优化，960 及 970 系列更支持自研 HiF4 格式，显著提升推理精度与吞吐。微架构层面将内存访问颗粒度优化至 128 字节，增强离散数据处理能力。

超节点技术是另一突破。Atlas 950 SuperPoD 支持 8192 张卡，通过自研“灵衢”互联协议实现全对等互联，带宽高达 16.3PB/s，时延降至纳秒级，并依托液冷散热有效控制高功耗。该架构使万卡集群可统一调度资源，显著提升训练与推理性能，并已扩展至通用计算领域。

灵衢互联协议采取开放策略，支持第三方硬件开发，其 UBoE 协议在时延与可靠性上优于传统方案。软件层面通过 CANN 开源栈、AscendC 语言及与主流框架深度集成，持续优化开发生态。

华为的进展正值全球 AI 算力竞争加剧，其开放策略带动芯片制造、封装、光通信和温控等产业链发展。然而，生态成熟度、技术迭代与国际竞争仍是未来需应对的挑战。

投资线索

国产 AI 算力产业链的核心环节，包括芯片制造与先进封装、高速互联、液冷温控及超节点相关硬件等领域，将直接受益于头部厂商的技术突破和规模化部署。同时，算力调度与集群运维系统对于实现超节点的高效协同和资源池化也很重要。

风险提示

1、政策落地不及预期；2、技术发展不及预期；3、市场竞争加剧。

证券研究报告

目录

一、本周市场回顾.....	4
二、本周关注.....	4
华为昇腾 AI 芯片与超节点技术	4
1、昇腾芯片架构的创新与迭代.....	5
2、超节点技术的突破与应用扩展.....	5
3、灵衢互联协议的开放生态建设.....	6
4、软件开源与软硬协同优化.....	6
5、产业影响和未来挑战	6
三、投资线索.....	7
四、本周要闻.....	7
1、进一步改善芯片散热：曝英伟达推动上游开发微米级水冷组件 MLCP.....	7
2、编程能力大幅提升，OpenAI 发布 GPT-5-Codex 新模型.....	8
3、2025 腾讯全球数字生态大会：腾讯 AI 能力全面开放 助力产业提效	8
4、阿里自研 AI 芯片现身，部分性能参数比肩英伟达 H20	10
5、华为公布昇腾 AI 芯片三年发展路线图	10
6、英伟达斥资 50 亿美元入股英特尔	12
五、风险提示.....	13

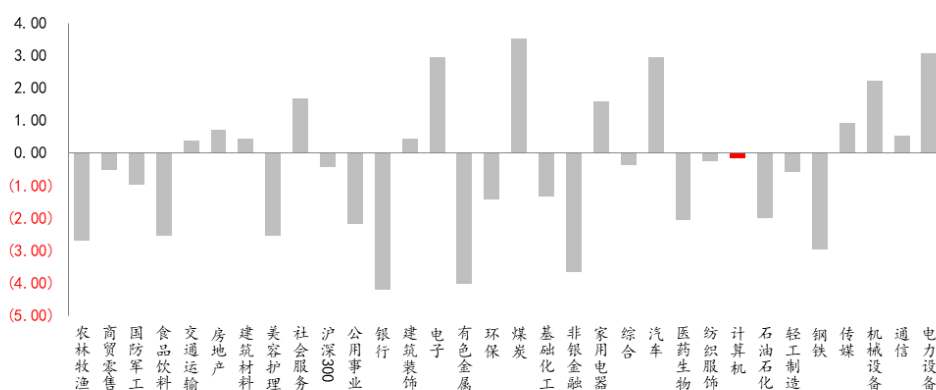
图表目录

图表 1：计算机行业板块相对表现（%）	4
图表 2：计算机行业个股涨跌幅	4

一、本周市场回顾

本周（9.15-9.19）计算机（申万）板块下跌 0.16%，沪深 300 指数下跌 0.44%，计算机板块跑赢沪深 300 指数 0.29 个百分点。和申万其他行业对比，计算机行业涨幅排名位列第 14 位。

图表 1：计算机行业板块相对表现（%）



数据来源：Wind，国新证券整理

目前计算机行业上市公司共 335 家，其中，88 家公司收涨，占比 26.19%。本周涨幅前 3 名分别为卡莱特（30.19%）、信息发展（25.71%）、有棵树（25.05%），跌幅前 3 名分别为*ST 东通（-65.05%）、新炬网络（-18.88%）、纬德信息（-15.33%）。

图表 2：计算机行业个股涨跌幅

板块涨幅前三名				板块跌幅后三名			
证券代码	证券简称	细分行业	周涨跌幅 (%)	证券代码	证券简称	细分行业	周涨跌幅 (%)
301391.SZ	卡莱特	智能硬件	30.19	300379.SZ	*ST 东通	网络安全	-65.05
300469.SZ	信息发展	数据要素	25.71	605398.SH	新炬网络	IT 基础设备	-18.88
300209.SZ	有棵树	交运 IT	25.05	688171.SH	纬德信息	网络安全	-15.33

数据来源：Wind，国新证券整理

二、本周关注

华为昇腾 AI 芯片与超节点技术

华为在 2025 年全联接大会上发布的昇腾 AI 芯片路线图及超节点战略，成为国

算力技术进军高端 AI 市场的重要事件，标志着相关领域实现重大突破。这并非局限于芯片硬件的迭代升级，而是涵盖了互联协议、软件生态及基础设施的全方位创新，其核心目标主要是两大方向：一是挑战英伟达主导的 CUDA 生态，二是推动具备自主可控属性的 AI 算力发展。从技术层面来看，华为正通过自研芯片架构、推行开源策略以及开展系统级优化，构建一套从芯片到集群的全栈式解决方案。

1、昇腾芯片架构的创新与迭代

昇腾 AI 芯片的核心技术演进，主要围绕算力、互联和内存架构三个维度展开升级。根据公布的路线图，从 2026 年将推出的昇腾 950 系列，到 2028 年计划发布的昇腾 970 系列，华为将微架构升级为 SIMD/SIMT，提升了芯片的编程灵活性与计算效率。其中，昇腾 950PR 和 950DT 虽共享同一 Die 设计，但针对不同应用场景进行了差异化优化：950PR 聚焦于推理预填充与推荐算法，采用华为自研的低成本 HBM HiBL 1.0 技术，旨在实现功耗与成本的双重降低；950DT 则侧重训练与解码阶段，配备 HiZQ 2.0 HBM，内存带宽达到 4TB/s，同时支持 FP8/MXFP4 等低精度格式，以此提升数据处理吞吐量。

根据华为公布的路线图，昇腾 AI 芯片将持续进行高性能迭代。后续的昇腾 960 和 970 系列算力分别翻倍至 2PFLOPS(FP8) /4PFLOPS(FP4)和 4PFLOPS(FP8) /8PFLOPS(FP4)。值得注意的是，昇腾 960 支持华为自研的 HiF4 数据格式，能进一步提升推理吞吐，并且比业界 FP4 方案的推理精度更优。这些性能升级不仅依赖于先进的封装工艺，更得益于微架构的创新，比如将内存访问颗粒度从 512 字节降至 128 字节，极大优化了离散数据处理能力，充分体现出华为在芯片设计过程中，对实际工作负载的深度适配与考量。

2、超节点技术的突破与应用扩展

华为在 2025 年全联接大会上发布的超节点技术，是其 AI 基础设施领域的另一项关键创新，旨在解决大规模算力集群在互联效率与能效方面的核心挑战。以 Atlas 950 SuperPoD 为例，该产品支持 8192 张昇腾卡，采用了华为自研的“灵衢 (UnifiedBus)”互联协议，实现了超节点内部的全对等互联。其互联带宽高达 16.3PB/s，显著超越了英伟达 NVL144 等业界方案。该互联技术将卡间时延降至纳秒级别（例如，华为 384 卡超节点的卡间时延为 150-200 纳秒），并且通过覆盖 70% 以上的液冷方案（如 Atlas 950 SuperPoD 即为全液冷数据中心超节点），有效应对了单机柜功耗超 100kW 的散热难题，提升了系统的稳定性。

超节点架构的核心在于其“总线级互联”与“平等协同”机制，使得万卡集群在逻辑上能够作为一台统一的计算机高效运行，实现计算与内存资源的全局调度。这一架构优势直接体现在性能提升上：Atlas 950 超节点的训练吞吐达 4.91M TPS，推理吞吐达 19.6M TPS，相比前代产品有显著提升。此外，华为还将超节点技术扩展至通用计算领域，推出了 TaiShan 950 超节点，该节点支持鲲鹏处理器的多核设

计，并结合 GaussDB 数据库，旨在实现对传统大型机和小型机的替代，展现了该技术在多样化场景下的强大适应能力。

3、灵衢互联协议的开放生态建设

作为支撑超节点技术的核心，互联协议灵衢（UB）的开放策略，是华为构建生态壁垒的重要举措。灵衢 2.0 协议成功攻克长距离高可靠互联的技术难题，通过光路百纳秒级故障检测与多端口聚合技术，将互联距离扩展至 200 米以上，同时保障 TB 级带宽与低时延的稳定性能。与传统 RoCE 协议相比，UBoE 协议在静态时延与可靠性上更具优势，不仅减少了对交换机和光模块的依赖，还降低了组网成本。

华为采取开放灵衢技术规范的策略，允许第三方企业开发兼容硬件，这一举措有助于加速技术在产业内的应用与普及，但能否成功关键取决于技术文档质量、实际用例支持以及问题响应机制的完善程度。灵衢的生态建设还延伸至软件层面，例如与 PyTorch/TensorFlow 等主流框架的深度集成，以及 MindStudio 工具链提供的全链路支持，这些措施有效缩小了与 CUDA 在开发者体验上的差距。不过，要进一步完善生态，仍需持续优化编译器前端与 API 兼容性。

4、软件开源与软硬协同优化

华为将其 AI 生态战略的核心置于软件开源与软硬协同优化之上，通过分层解耦的 CANN 软件栈实现从算子库到驱动层的全栈开源。该软件栈作为连接 AI 框架与昇腾硬件的关键中间件，支持主流框架如 PyTorch 和 TensorFlow，并通过图编译优化、算子融合等技术提升性能。华为提供了 AscendC 编程语言供开发者进行自定义算子开发，以优化性能。

华为与 DeepSeek 的合作典型地体现了其软硬协同策略的成效。双方通过采用 MoE 架构和华为昇腾硬件的动态专家调度等优化技术，有效解决了大模型推理中的负载均衡和卡间通信问题。该合作使得 DeepSeek 大模型在昇腾平台上的单推理吞吐量显著提升，解码时延大幅降低，同时显存占用减少。尽管取得了不少进展，但在处理动态输入形状等复杂场景时，其性能稳定性仍是业界持续挑战的方向之一。华为也在通过不断优化其软件栈来应对这些问题。

为了构建更繁荣的开发者生态，华为已将许多 AI 相关的软硬件根技术开源，例如将 MindSpore 框架开源，并将欧拉操作系统（openEuler）等推向开源社区。华为通过“众智”计划等项目，鼓励产学研合作，并计划发展数百万的 MindSpore 社区开发者。生态的繁荣需要华为提供更完善的基准测试体系、详尽的技术文档以及高效的问题响应机制，以吸引更多开发者参与并推动技术的持续迭代升级。

5、产业影响和未来挑战

从产业影响来看，华为在昇腾 AI 芯片与超节点技术上的突破，恰逢全球 AI 算力投资热潮，中美双方也在“星际之门”、国家基金等相关计划推动下，加速展开算力领域的竞争。国产算力产业链借助华为的开放策略，在芯片制造、HBM 封装、光通信及液冷温控等多个领域获得发展机遇，但核心挑战仍集中在生态成熟度层面。

华为此次发布的昇腾 AI 芯片路线图及超节点战略，标志着国产算力正式迈入国际高端竞争舞台，为投资者带来了结构性机会。凭借自研芯片架构、全光互联技术及开源生态建设，华为在算力性能、能效成本及自主可控性上形成显著优势，未来三年有望推动国产替代进程加速。昇腾芯片的迭代升级（如 950PR/DT、960 及 970 系列在算力、互联带宽和内存架构上的突破）、超节点架构（如 Atlas950 SuperPoD 支持 8192 张卡的全光互联设计）以及灵衢互联协议的开放生态，共同构成了华为 AI 基础设施的核心竞争力，但技术迭代风险、生态成熟度及国际竞争压力仍需谨慎考量。

三、投资线索

建议重点关注国产 AI 算力产业链的核心环节，涵盖芯片制造、HBM 封装、光通信、液冷温控及超节点相关硬件领域。具体来看，芯片与 HBM 环节可关注中芯国际、通富微电、中微公司、北方华创、精智达和华海诚科等；超节点相关的全光互联企业，如德科立、腾景科技、凌云光、光库科技和炬光科技等；正交背板领域可关注深南电路、南亚科技（内存）和博敏电子等；液冷温控供应商方面，申菱环境、飞荣达和英维克具备较大潜力。这些企业直接受益于华为超节点部署与昇腾芯片量产带来的需求增长，部分已进入华为供应链，拥有先发优势。投资者应优先选择技术壁垒高、第二增长曲线清晰的标的，并密切关注 2026 年第一季度昇腾 950PR 量产及超节点商业化进度带来的验证机会。

同时，投资者需高度关注潜在风险因素，包括技术迭代不及预期（如先进工艺量产延迟或生态适配问题）、生态发展缓慢（开发者社区活跃度及与 CUDA 的兼容性差距）、下游需求波动（受 AI 资本开支周期影响）以及国际竞争加剧的影响。

四、本周要闻

1、进一步改善芯片散热：曝英伟达推动上游开发微米级水冷组件 MLCP

IT 之家消息，综合台媒《经济日报》《工商时报》报道，英伟达正推动上游供应商开发一类名为 MLCP（IT 之家注：微通道水冷板）的水冷散热组件，以应对英伟达 AI GPU 芯片随代际更替不断上升的发热。

报道指出，英伟达下代的 "Rubin" GPU 将在单一封装中包含两颗 GPU 芯片，功耗预计将超过 2000W，尽管较大的表面积有利于热量传导但其仍然对冷却系统

提出了超越现有水冷板解热能力的严苛要求。

传统水冷板的微水道尺寸在 0.1mm (100 μ m) 到数毫米水平，而在 MLCP 中通过在芯片或封装上的蚀刻水道尺寸可降低至微米级别，同时均热板、水冷板、IHS 封装顶盖、芯片裸晶也将高度整合，进一步提升散热效率。

据悉 MLCP 单价可达传统水冷板的 3~5 倍且能贡献较高毛利率，但由于其在流体力学、气泡动力学上的复杂性和更高冷液渗漏风险尚需一段时间才能成熟。

2、编程能力大幅提升，OpenAI 发布 GPT-5-Codex 新模型

IT 之家 9 月 16 日消息，OpenAI 今日凌晨宣布推出全新的 GPT-5-Codex，并已在其 AI 编程助手 Codex 中上线。这是一个专门为 Agent 编程而优化的 GPT-5 新版本。

目前，GPT-5-Codex 已通过 Codex 向 ChatGPT Plus、Pro、Business、Edu 与 Enterprise 用户开放，可在终端、IDE、GitHub 或 ChatGPT 中使用。IT 之家从官方获悉，OpenAI 还计划在未来向 API 客户开放该模型。

OpenAI 表示，GPT-5-Codex 的一大改进在于动态计算能力。该模型能够更灵活地分配“思考时间”，在执行代码任务时可能从数秒到长达七小时不等，从而在多项编程基准测试中展现出更优表现。

与此前依靠“路由”机制来预判任务复杂度并分配资源的方式不同，GPT-5-Codex 能在任务进行过程中实时调整所需时间和算力。例如，它可能在开始五分钟后判断需要额外花费一小时，甚至在个别情况下长达七小时。这种机制使其在复杂编程任务上更具优势。

根据 OpenAI 数据，GPT-5-Codex 在 SWE-bench Verified 测试以及大型代码库的重构任务中，均优于标准 GPT-5 模型。

此外，该模型还专门接受了代码审查训练，并由资深软件工程师进行评估。结果显示，GPT-5-Codex 的错误更少，同时能提供更多“高影响力的评论”。

OpenAI Codex 产品负责人 Alexander Embiricos 在简报中指出：“GPT-5-Codex 的动态思考能力带来了显著提升。它不依赖固定路由，而是能在任务中途判断需要增加工作时间，这在复杂问题上尤其有优势。”

3、2025 腾讯全球数字生态大会：腾讯 AI 能力全面开放 助力产业提效

9 月 16 日，2025 腾讯全球数字生态大会在深圳举行，会上公布多项 AI 技术和产品最新进展，并宣布全面开放腾讯 AI 落地能力及优势场景，助力“好用的 AI”在千行百业中加速落地。

“如果说智能化引擎致力于让 AI 从时髦概念，变成切实驱动业务的实际效率，那么全球化引擎则希望企业在大航海时代，赢得新增长。”腾讯高级执行副总裁、云与智慧产业事业群 CEO 汤道生表示，腾讯将从基础设施、技术产品以及服务能力三个方向，升级腾讯云国际化战略，打造“全球化引擎”。腾讯特色产品如超级应用解决方案 Superapps a Service、刷掌服务平台等，受到了海外企业广泛欢迎。同时，通过覆盖全球的本地化技术支持和服务团队，及时响应客户需求，敏捷应对市场变化。

会上，腾讯云宣布了国际化产品的全系列升级方案，腾讯云智能体开发平台 ADP、代码助手 CodeBuddy、企业级全渠道交易解决方案云 Mall、星星海服务器、腾讯云数据库 TDSQL、专有云 TCE、边缘安全加速平台 EdgeOne 等腾讯云优势产品全面适配全球技术生态。

汤道生强调，腾讯将全面开放 AI 落地能力及优势场景，让腾讯 AI 成为各行各业的提效与增长新引擎。

会上，腾讯云发布了多项 AI 技术产品与最新升级：腾讯云智能体开发平台（ADP）面向全球发布 3.0 版。面对产业实战中的智能体需求，ADP 在 3 个月内完成了近 600 项需求的开发，持续迭代 LLM+RAG、Workflow、Multi-Agent 等多种智能体开发框架。

腾讯云还发布了多模态智能数据湖 TCLake，通过多层链路的自研，让企业可以更高效地处理多模态数据，构建统一的数据底座，更好地推动 Agent 应用服务落地。

腾讯混元大模型也再次升级。最新发布的混元 3D 3.0 模型，首创 3D-DiT 分级雕刻模型，建模精度比前代模型提升 3 倍，几何分辨率达到了 1536³，现已集成于混元 3D AI 创作引擎，面向用户免费开放，并同步上线腾讯云 API。

基于多项最新能力，腾讯云在会上正式发布腾讯云智能体战略全景图，全面开放腾讯云的开发生态、应用场景、基础设施、模型能力以及伙伴生态。

会上，腾讯云公布了全球化最新成绩单：截至目前，腾讯云服务的海外客户规模在过去一年同比翻番。90%以上的中国互联网企业、95%以上的头部游戏公司出海都选择了腾讯云，美团 Keeta、泡泡玛特等一系列中国企业，借助腾讯云加速出海进程。

基础设施层面，腾讯计划未来投入 1.5 亿美元，在沙特阿拉伯建设首个中东数据中心；同时，将在大阪新建日本的第三个数据中心，并设立大阪新办公室。截至 2025 年上半年，腾讯云已在雅加达、马尼拉、吉隆坡、新加坡、曼谷、东京、首尔、帕洛阿尔托、法兰克福部署 9 大全球技术支持中心。

作为产业提效创新的新机遇，智能化和国际化不仅是腾讯助力实体产业的重要

抓手，也助力生态伙伴稳步增长。伙伴通过腾讯云的收入增速，远超行业平均水平；SaaS 服务客户数增长了 60%；腾讯云给生态伙伴带来的商机规模同比增长了 330%。本次大会上，腾讯云还联合数十家合作伙伴共同发布了“AI 应用繁荣计划”，发起“AI 共创营”和“AI 百校行”行动，联合伙伴力量借助腾讯云技术能力孵化更多垂直领域 AI 应用，助力培育未来 AI 人才，为产业创新注入新动能。（证券日报）

4、阿里自研 AI 芯片现身，部分性能参数比肩英伟达 H20

市场传闻半个月之久的阿里自研 AI 芯片，9 月 16 日晚出现在央视《新闻联播》。

央视画面披露，阿里旗下平头哥 PPU 芯片在部分重要参数上比肩英伟达的 H20 芯片，并超过 A800 芯片。

平头哥 PPU 集成 HBM2e（第三代高性能内存），和 A800 相同，但落后于 H20 的 HBM3（第四代高性能内存）。但 PPU 的显存容量与 H20 相同，均为 96G。卡间互联带宽上，PPU 为 700GB/s，介于 A800 和 H20 之间。PCIe（外围组件快速互连）接口层面，PPU 也优于 A800，与 H20 等同。功耗上，PPU 与 A800 一致，均为 400W，少于 H20 的 550W。

如果和国产阵营的华为昇腾 910B 芯片对比，央视画面显示，平头哥 PPU 在上述大多数性能参数指标上均处于领先。不过，昇腾系列最新款是 910C 芯片。

有媒体此前 8 月下旬报道，阿里已开发出一种比其旧款芯片更通用的新款芯片，用于 AI 推理任务，以填补英伟达留下的市场空白。

9 月 11 日，硅谷科技媒体 The Information 进一步爆料称，阿里和百度都在使用自研的 AI 芯片训练模型。其中，阿里将自研芯片用于开发较小的 AI 模型；百度则尝试使用其昆仑芯 P800 芯片来训练文心大模型的新版本。但两家公司并没有完全放弃英伟达的芯片。阿里和百度均未对此回应。

“AI+云”已成为阿里内部与电子商务并列的两大增长引擎之一。今年 2 月，阿里方面透露，未来三年将投入至少 530 亿美元，建设云和 AI 的硬件基础设施。

今年 6 月，阿里集团主席蔡崇信、阿里集团 CEO（首席执行官）吴泳铭在股东信中表示，未来十年，最大的增量和变量都是以 AI 为核心的驱动力带来的变革。围绕 AI 这个战略核心，阿里会加大三个领域的投资力度：投入 AI 和云计算基础设施建设，投入 AI 基础模型和 AI 原生应用，投入现有业务的 AI 转型升级。（南方都市报）

5、华为公布昇腾 AI 芯片三年发展路线图

9月18日，在上海世博中心举办的2025华为全联接大会上，华为副董事长、轮值董事长徐直军登台发表演讲，首次对外公布了昇腾AI芯片未来三年的产品迭代路线图，同时明确表示2026年一季度发布的新产品将采用华为自研HBM（高带宽内存）。

根据规划，2026年至2028年期间，华为将分阶段推出四款昇腾系列芯片，具体包括：2026年第一季度推出昇腾950PR，该芯片采用华为自研HBM；2026年第四季度推出昇腾950DT；2027年第四季度推出昇腾960芯片；2028年第四季度推出昇腾970。

与英伟达基于通用集成电路设计的GPU有所不同，华为昇腾芯片属于专用集成电路架构的NPU（Neural Processing Unit，神经网络处理器），专为处理AI神经网络计算任务设计。

从2019年开始，华为已经发布多款昇腾910系列芯片，包括910B、910C多款产品，大会晒出的路线图显示910C为今年第一季度最新发布。该系列是基于华为自研的达芬奇架构，专为云端AI训练和推理使用。

从芯片技术指标来看，昇腾910C算力高达800TFLOPS（以行业衡量AI算力规模的半精度浮点数FP16为标准），支持业内标准的FP32/HF32/FP16/BF16/INT8等数据格式，互联带宽784GB/s，HBM容量为128GB、内存带宽为3.2TB/s。

作为对比，英伟达最新Blackwell B300在同等标准下的算力约为3840TFLOPS（B300在FP4标准下算力为15P，1P等于1024T，即15360TFLOPS，在理想精度无损转换场景下，FP4算力折算成FP16需除以4），配备的是288GBHBM3e，带宽为8TB/s。

2026年将要发布的昇腾950PR/DT微架构将升级为SIMD/SIMT，算力达到1PFLOPS（FP8）/2PFLOPS（FP4），支持FP32/HF32/FP16/BF16/FP8/MXFP8/HiF8/MXFP4/HiF4等数据格式，互联带宽为2TB/s。内存容量及带宽上，昇腾950PR为128GB和1.6TB/s；昇腾950DT为144GB和4TB/s。

徐直军在演讲中指出，由于受美国的制裁，华为不能到台积电去投片，单颗芯片的算力相比英伟达存在差距。但华为“有三十多年联人、联机器”的积累，并在超节点互联技术上强力投资、实现突破，能够做到万卡级的超节点，从而做到世界上算力最强。

借助超节点技术，走“集群规模化”路线是华为一直发力的方向。超节点可以简单理解为凭借高速总线互联技术，将多个CPU、GPU或NPU加速卡等组成的“小计算单元”连接成一整个超大的计算单元，以集群化、规模化弥补单芯片性能不足，实现算力供给。

此前以昇腾 910 芯片为基础，华为已经推出了超节点产品。华为今年 5 月推出的昇腾 384 超节点即将 384 张昇腾 NPU 与 192 张鲲鹏 CPU 连接在一起，并在 7 月举办的世界人工智能大会上线下展出亮相。

在今日的全联接大会上，徐直军还公布了华为最新的超节点产品 Atlas 950 SuperPoD、Atlas 960 SuperPoD 超节点，分别支持 8192 及 15488 张昇腾卡。

按照他在演讲中展示的对比图，Atlas 950 SuperPoD 超节点已超越英伟达将在 2027 年发布的最大规模超节点 NVL576。NVL576 是英伟达计划在 2027 年下半年推出的基于下一代 Rubin 架构的超大规模 AI 算力集群系统，包含 576 张 Rubin Ultra GPU。

按照华为披露的计划，Atlas 950 SuperPoD 预计今年四季度上市，Atlas 960 SuperPoD 计划将在 2027 年四季度上市。

徐直军演讲中多次强调“算力过去是，未来也将继续是人工智能的关键，更是中国人工智能的关键”。而“基于中国可获得的芯片制造工艺”，华为正在努力打造‘超节点+集群’算力解决方案，以满足持续增长的算力需求。（界面新闻）

6、英伟达斥资 50 亿美元入股英特尔

今日，美国两大芯片巨头英伟达与英特尔共同宣布，双方将开展深度合作，联合开发多代定制化数据中心及个人电脑（PC）产品，以加速超大规模、企业级及消费级市场中各类应用与工作负载的运行效率。此次合作不仅涉及技术层面的深度融合，为深化合作信任，英伟达还计划以每股 23.28 美元的价格，向英特尔普通股投资 50 亿美元，进一步巩固双方合作关系。不过，此次股权投资需满足常规成交条件，其中包括获得全球相关监管机构的批准，具体完成时间将根据监管流程推进情况确定。

双方在公告中表示，将以英伟达的 NVLink 技术为核心纽带，实现英伟达与英特尔架构的无缝连接，此举旨在整合英伟达在人工智能（AI）与加速计算领域的技术优势，以及英特尔领先的 CPU 技术与庞大的 x86 生态系统。

据了解，双方此次的合作将聚焦两大核心场景。一是数据中心领域，用定制化 CPU 赋能 AI 基础设施。根据协议，英特尔将为英伟达单独设计并制造 x86 架构 CPU。这款定制化 CPU 后续将被英伟达整合至其 AI 基础设施平台中，通过“英特尔 CPU+英伟达 AI 技术”的组合，提升数据中心在 AI 训练、推理等高强度工作负载下的运行效率，并面向全球市场推出相关解决方案；二是个人计算领域，双方将集成 RTX GPU 的 SoC 重塑 PC 体验。英特尔将打造并向市场推出集成英伟达 RTX GPU 芯片的 x86 系统级芯片（SoC）。这类全新的 x86 RTX SoC 将重点满足高性能 PC 的需求，为游戏、创意设计、专业计算等对 CPU 与 GPU 协同算力要求极高的场景提供支持，填补当前 PC 市场中“顶级 CPU 与 GPU 深度集成”的产品

空白。(中国电子报)

五、风险提示

- 1、政策落地不及预期；
- 2、技术发展不及预期；
- 3、市场竞争加剧。

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。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