

全栈算力之巅，物理AI新纪元

投资要点

- 推荐逻辑:** 1) 随着AI算力从芯片级竞争转向系统集群级竞争，英伟达通过创新产品架构、强化互联技术以及整合全栈软硬件生态，有效拓宽了公司的护城河；2) 多模态、AI Agent、主权AI等持续提振算力需求，基于Blackwell平台训练的大模型的后续推出或对算力形成新的催化；3) 公司业绩维持高速增长态势，未来三年净利润年复合增速58%，估值显著低于同业平均水平。
- AI算力竞争转向系统集群级，架构演进和创新支撑公司产品力持续领先。** AI算力已从芯片层级的竞争转向系统集群级的竞争，英伟达最新量产的Blackwell和下一代Rubin平台在架构上进行了演进和诸多创新，或重塑全球算力格局。Blackwell平台是首个整合了CPU、GPU、NVLink、NICK（特定技术组件）、以及网络交换机等多个元素的完整系统。Blackwell NVL72相较于上一代8卡服务器，性能提升了9倍，带宽增加了18倍，每秒浮点运算次数提升了45倍，而功率仅增加10倍。Blackwell已量产，广泛应用于各大科技公司和AI企业。2026年基于Blackwell平台训练的大模型将推出，或对算力形成新的催化。
- 下一代Rubin平台开启算力新纪元。** Rubin平台计划于2026年量产，其集成的Rubin GPU、Vera CPU及NVLink 6等核心组件，在性能指标上实现跨代飞跃。其FP4推理性能相较于Blackwell提升5倍，HBM4内存带宽提升2.8倍。通过引入分离式推理架构和Rubin CPX优化方案，针对Prefill和Decode阶段的硬件进行定制，极大地降低了大模型推理的每token成本。此外，公司通过获得Groq的SRAM架构技术授权，积极布局低延迟推理市场，旨在重塑后训练时代的AI计算价值链。
- 物理AI的“ChatGPT时刻”即将到来，或重塑公司AI底座价值。** 未来AI将具备理解因果、模拟物理运动及自主决策的能力。英伟达通过一系列开源模型，加速AI从数字领域向物理世界渗透。英伟达正构建一个连接数字孪生与真实世界的闭环，这一布局将确保其在具身智能时代，依然能掌握全球算力话语权。
- 盈利预测与评级。** 公司是全球AI芯片龙头企业，预计未来三年Non-GAAP净利润年复合增速为58%。结合业绩增速，给予公司2027财年30xPE(Non-GAAP)，目标价249.54美元，对应市值6万亿美元，给予“买入”评级。
- 风险提示:** AI芯片需求或不及预期、下游客户Capex力度或不及预期、竞争加剧等风险。

| 指标/年度 | FY2024A | FY2025A | FY2026E | FY2027E | FY2028E |
|--------------------|----------|-----------|-----------|-----------|-----------|
| 营业收入(百万元美元) | 60922.00 | 130497.00 | 213192.96 | 342625.44 | 489585.74 |
| 增长率 | 125.85% | 114.20% | 63.37% | 60.71% | 42.89% |
| GAAP净利润(百万元美元) | 29760.00 | 72880.00 | 111073.77 | 182988.94 | 265077.32 |
| 增长率 | 581.32% | 144.89% | 52.41% | 64.75% | 44.86% |
| Non-GAAP净利润(百万元美元) | 32312.00 | 74266.00 | 116707.00 | 202128.00 | 293526.00 |
| 增长率 | 286.23% | 129.84% | 57.15% | 73.19% | 45.22% |
| 每股收益EPS | 1.22 | 3.00 | 4.57 | 7.53 | 10.91 |
| 净资产收益率 | 69.24% | 91.87% | 79.10% | 78.90% | 78.44% |
| PE | 152.89 | 62.43 | 40.96 | 24.86 | 17.16 |

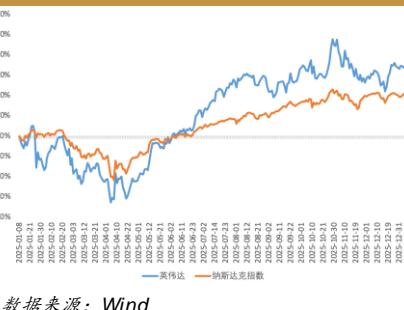
数据来源：公司公告，西南证券

西南证券研究院

分析师：王湘杰
执业证号：S1250521120002
电话：0755-26671517
邮箱：wxj@swsc.com.cn

分析师：杨镇宇
执业证号：S1250517090003
电话：023-67563924
邮箱：yzyu@swsc.com.cn

相对指数表现



数据来源：Wind

基础数据

52周区间(美元) 94.29-207.03
3个月平均成交量(百万) 185.79
流通股数(亿) 243
市值(亿美元) 45954

相关研究

- 英伟达(NVDA.O)：业绩增长超预期，Blackwell四季度出货(2024-11-26)
- 英伟达(NVDA.O)：增长势头强劲，B系列提前出货超预期(2024-05-29)

目 录

| | |
|---|----|
| 1 智能计算的先驱与领导者，软硬一体全栈平台化战略 | 1 |
| 2 AI 算力全栈布局，下游 Capex 支撑高增长 | 3 |
| 2.1 下游资本支出持续攀升，数据中心业务高速增长 | 3 |
| 2.2 AI 算力产品全栈布局，新平台重塑 AI 芯片格局 | 5 |
| 3 物理 AI 的“ChatGPT 时刻”即将到来，或重塑公司 AI 底座价值 | 14 |
| 4 盈利预测与投资建议 | 16 |

图 目 录

| | |
|---|----|
| 图 1: 英伟达重大发展历程 | 1 |
| 图 2: 英伟达硬件产品线——提供底层物理算力支持 | 2 |
| 图 3: 英伟达软件产品线——降低开发门槛, 形成生态黏性 | 2 |
| 图 4: 英伟达营收 (亿美元) 及增速 | 2 |
| 图 5: 英伟达 Non-GAAP 净利润 (亿美元) 及利润率 | 2 |
| 图 6: 英伟达各业务收入 (亿美元) | 3 |
| 图 7: 英伟达各业务收入占比 | 3 |
| 图 8: 北美四大互联网公司季度 Capex (亿美元) | 4 |
| 图 9: 英伟达数据中心业务年度表现 (亿美元) | 4 |
| 图 10: 英伟达数据中心业务季度表现 (亿美元) | 4 |
| 图 11: NVIDIA 加速计算平台 | 5 |
| 图 12: NVIDIA 加速系统与传统系统的对比 | 5 |
| 图 13: 英伟达 AI GPU 产品线规划 | 6 |
| 图 14: 英伟达 Blackwell 平台产品组合 | 7 |
| 图 15: 英伟达 AI GPU 的能耗随产品代际提升而快速降低 | 7 |
| 图 16: 英伟达 Blackwell 机架方案 | 8 |
| 图 17: 英伟达 Spectrum-X | 8 |
| 图 18: 英伟达 Rubin GPU 性能 | 8 |
| 图 19: 英伟达 Vera CPU 性能 | 8 |
| 图 20: 英伟达 BlueField-4 和 ConnectX-9 性能 | 9 |
| 图 21: 英伟达 Vera Rubin NVL72 性能 | 9 |
| 图 22: 英伟达 Vera Rubin DGX SuperPOD | 9 |
| 图 23: 英伟达 Rubin 平台性能相比 Blackwell 大幅提升 | 9 |
| 图 24: 英伟达 Rubin Ultra 平台方案 | 10 |
| 图 25: 英伟达分离式推理架构 | 11 |
| 图 26: 英伟达 Vera Rubin CPX 机架方案 | 12 |
| 图 27: 英伟达 Vera Rubin CPX 性能 | 12 |
| 图 28: 英伟达与 Groq 签署非独家推理技术许可协议 | 13 |
| 图 29: Groq LPU 介绍 | 13 |
| 图 30: Groq 的 LPU: 一种创新的片上 AI 内存架构 | 13 |
| 图 31: 英伟达开源模型生态 | 14 |
| 图 32: 英伟达全栈物理 AI 平台 | 14 |
| 图 33: 英伟达物理 AI 开源世界模型——Cosmos | 15 |
| 图 34: 英伟达开源 VLA 推理模型——Alpamayo | 15 |
| 图 35: 英伟达在 CES 2026 上的机器人演示 | 15 |
| 图 36: 英伟达物理 AI “开源全家桶” | 15 |
| 图 37: AI 不仅仅是大模型 | 16 |
| 图 38: 英伟达 PE/PB-BAND | 17 |

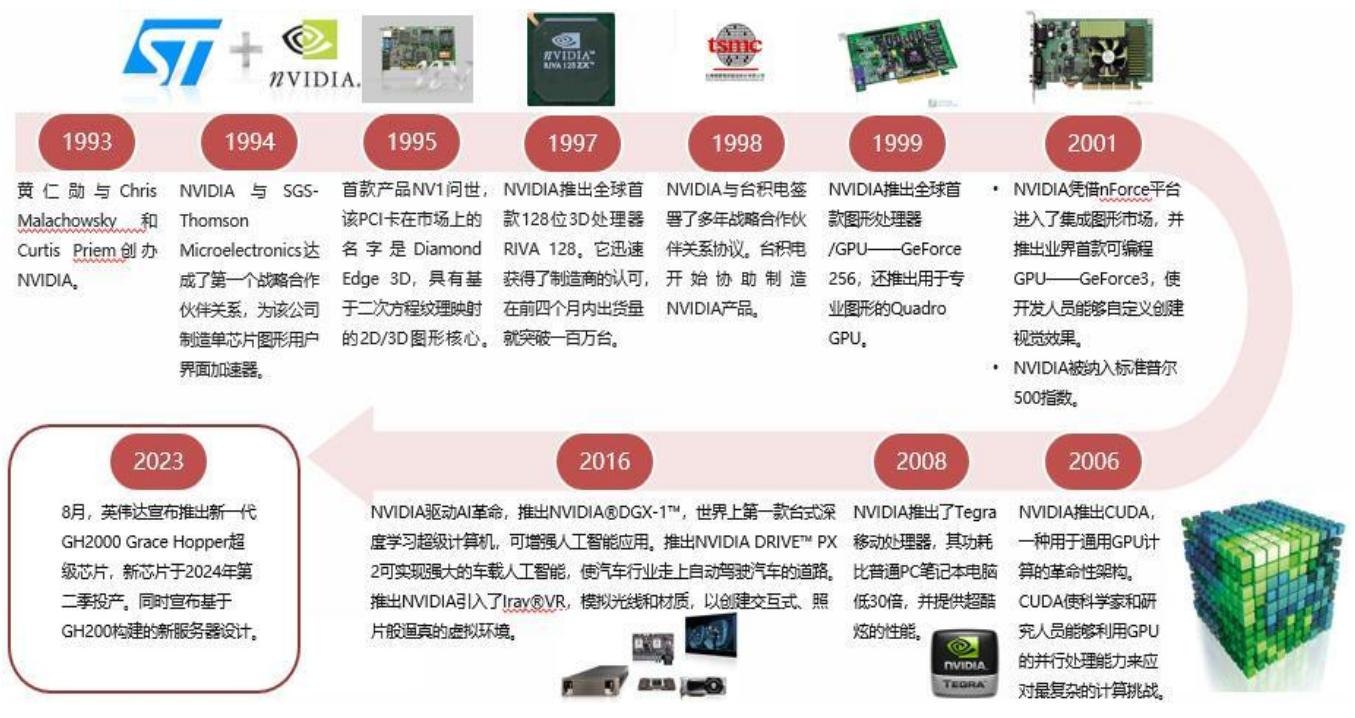
表 目 录

| | |
|-----------------------|----|
| 表 1：英伟达重要机构股东持股 | 3 |
| 表 2：分业务收入 | 16 |
| 表 3：可比公司估值情况 | 17 |
| 附：财务报表 | 18 |

1 智能计算的先驱与领导者，软硬一体全栈平台化战略

英伟达是全球 AI GPU 和消费级显卡的龙头企业。公司早期凭借 CUDA 库的构建，显著降低了 GPU 的编程门槛，开创了 GPU 向通用计算市场拓展的先驱，引领了 GPGPU 的进化之路。英伟达借助硬软件业务形成的 AI、Omniverse、Drive Hyperion 等平台，不断壮大其计算生态，进一步夯实公司的业务护城河，AI 数据中心、智能驾驶等业务将打开公司成长空间。

图 1：英伟达重大发展历程



数据来源：英伟达中国公众号，西南证券整理

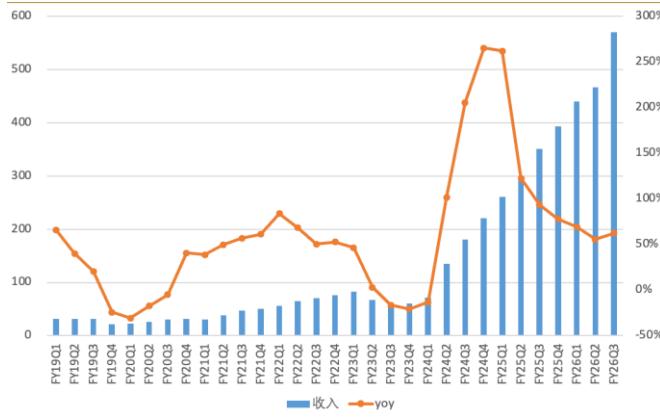
英伟达的产品线包含硬件和软件两部分。硬件产品主要包含数据中心、网络、游戏和创作、嵌入式系统、笔电和专业工作站等板块；软件产品主要有应用框架、应用和工具、游戏和创作、基础架构和云服务等板块。这种软硬一体的协同效应，不仅壮大了公司的计算生态，更通过持续的算法优化与硬件迭代，筑起了极高的准入门槛。AI 数据中心、智能驾驶、以及即将到来的物理 AI，正不断打开公司的长期成长空间。

图 2：英伟达硬件产品线——提供底层物理算力支持

| 数据中心 | | |
|-------------------------------------|--------------------------------------|---------------------------------------|
| 概览 | DGX 平台 | Grace CPU |
| 借助 AI 和加速计算实现数据中心现代化 | 用于模型开发和部署的企业 AI 工厂 | 将数据转化为智能的数据中心架构 |
| HGX 平台 | IGX 平台 | MGX 平台 |
| 专为 AI 和 HPC 打造的超级计算机 | 适用于边缘 AI 的高级功能安全与信息安全 | 使用模块化服务器进行加速计算 |
| OVX 系统 | | |
| 适用于高性能 AI 的可扩展数据中心基础设施 | | |
| 网络 | | |
| 概览 | DPU 和 SuperNIC | 以太网 |
| 适用于现代工作负载的加速网络 | 软件定义的硬件加速器，适用于网络、存储和安全 | 以太网在各种应用中的性能、可用性和易用性 |
| InfiniBand | 网络软件 | 网络加速 |
| 适用于超级计算机、AI 和云数据中心的高性能网络 | 用于优化性能和可扩展性的网络软件 | 适用于现代 GPU 加速数据中心的 IO 系统 |
| 游戏和创作 | | |
| GeForce | GeForce 显卡 | 游戏笔记本电脑 |
| 探索显卡、游戏解决方案、AI 技术等 | RTX 显卡带来颠覆性的 AI 功能 | 经 Max-Q 技术优化的轻薄耐用 RTX 笔记本电脑 |
| G-SYNC 显示器 | DLSS | Reflex |
| 借助 NVIDIA G-SYNC 显示器，畅享流畅、无撕裂的游戏体验 | 可提升帧率并增强画面的神经网络渲染技术 | 出色的响应速度，可加快反应速度并提高精准精度 |
| RTX AI PC | NVIDIA Studio | NVIDIA App |
| 用于游戏、创作、生产力和开发的 AI PC | 专为创作者打造的高性能笔记本电脑和台式机 | 优化游戏、直播和 AI 创新的创意 |
| NVIDIA Broadcast 应用 | | |
| AI 增强的语音和视频，适用于更高级的流式传输、视频和通话 | | |
| 嵌入式系统 | | |
| Jetson | DRIVE AGX | Clara AGX |
| 适用于自主机器和嵌入式应用的领先平台 | 强大的车载计算能力，适用于 AI 驱动的智能汽车系统 | 适用于创新型医疗设备和成像的 AI 计算 |
| 笔记本电脑 | | |
| GeForce 笔记本电脑 | Studio 设计本 | NVIDIA RTX PRO 笔记本电脑 |
| 搭载 RTX GPU 的笔记本电脑，专为游戏玩家和创作者打造 | 专为创作者打造的高性能笔记本电脑 | 随时随地加速专业 AI 和视觉计算 |
| 专业工作站 | | |
| 概览 | DGX Spark | NVIDIA RTX PRO AI 工作站 |
| 加速专业 AI、图形、渲染和计算工作负载 | 桌面上的 Grace Blackwell AI 超级计算机 | 加速 AI 工作流的创新和生产力 |
| 桌面端 NVIDIA RTX PRO | NVIDIA RTX PRO 笔记本电脑 | |
| 强大的 AI、图形、渲染和计算工作负载 | 随时随地加速专业 AI 和视觉计算 | |

数据来源：公司官网，西南证券整理

AI 推动 GPU 的需求大幅提升，英伟达近三年业绩呈现快速增长态势。随着生成式 AI 爆发，推动 GPU 需求大幅增长，公司 FY2023 至 FY2025 财年营收从 269.7 亿美元高速增长至 1305 亿美元，CAGR 高达 120%；期间 Non-GAAP 净利润从 83.7 亿美元高速增长至 742.7 亿美元，CAGR 高达 198%，净利率从 31% 大幅提升至 56.9%。

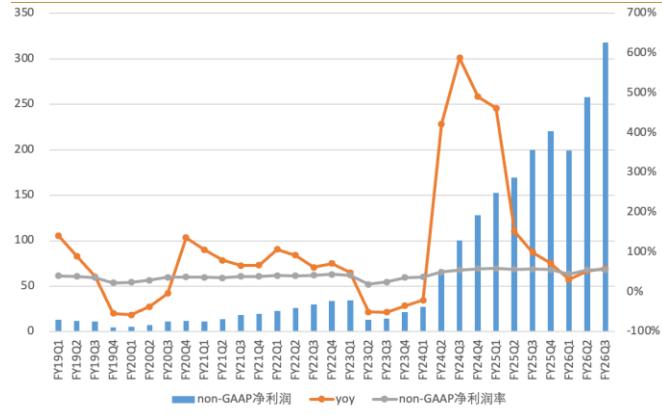
图 4：英伟达营收（亿美元）及增速


数据来源：公司公告，西南证券整理

图 3：英伟达软件产品线——降低开发门槛，形成生态黏性

| | | |
|---|--|---|
| 代理式 AI - Agent Intelligence Toolkit | AI Blueprint | AI 推理 - Dynamo |
| AI 推理微服务 - NIM | AI 微服务 - CUDA-X | 汽车 - DRIVE |
| 虚拟形象 - Tokkio | 云 AI 视频流 - Maxine | 网络安全 - Morpheus |
| 数据科学 - Apache Spark | 数据科学 - RAPIDS | 决策优化 - cuOpt |
| 生成式 AI - NeMo | 医疗健康 - Clara | 工业 AI - Omniverse |
| 智能视频分析 - Metropolis | NVIDIA Mission Control | NVIDIA AI Enterprise 平台 |
| 物理 AI - Cosmos | NVIDIA Runai | 机器人 - Isaac |
| 电信行业 - Aerial | | 查看所有软件 |

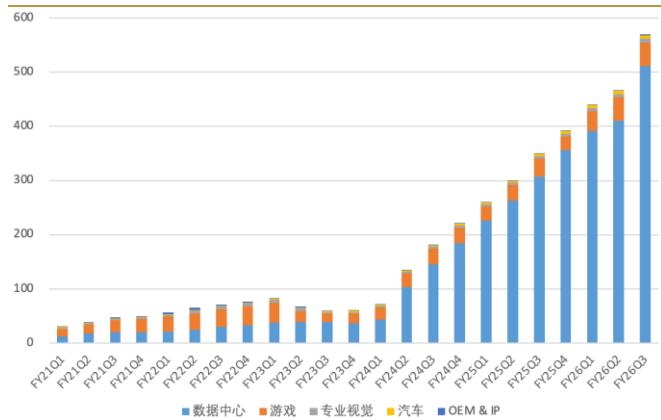
数据来源：公司官网，西南证券整理

图 5：英伟达 Non-GAAP 净利润（亿美元）及利润率


数据来源：公司公告，西南证券整理

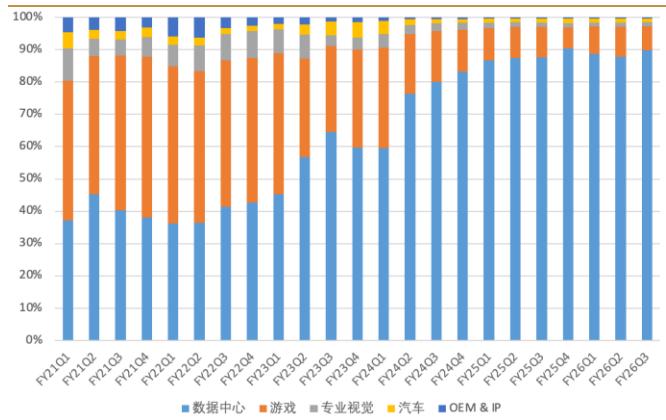
英伟达业务结构可分为数据中心、游戏、专业视觉、汽车、OEM&IP 五大板块。其中，数据中心和游戏为公司主要收入来源。FY26Q3 数据中心业务营收 512.2 亿美元，同比增长 66.4%，环比增长 24.6%，收入占比近 90%，成为英伟达收入和利润增长的最主要贡献。对生成式 AI 和大语言模型的需求不断增长，众多 CSP 厂商、互联网大厂、算力租赁企业等竞相部署 AI 芯片，带动了公司的数据中心业务大幅增长。

图 6：英伟达各业务收入（亿美元）



数据来源：公司公告，西南证券整理

图 7：英伟达各业务收入占比



数据来源：公司公告，西南证券整理

截止 2025 年 9 月 30 日，英伟达前五大机构持股合计 67 亿股，占比 27.6%。其中，Vanguard(先锋领航集团)、BlackRock(贝莱德集团)、FMR LLC、State Street、Geode Capital 五家机构持股占比分别为 9.15%、7.94%、4.04%、4.03% 和 2.41%。

表 1：英伟达重要机构股东持股

| 重要股东 | 持股数量（亿股） | 持股比例 |
|---------------|----------|-------|
| Vanguard | 22.24 | 9.15% |
| BlackRock | 19.30 | 7.94% |
| FMR LLC | 9.83 | 4.04% |
| State Street | 9.80 | 4.03% |
| Geode Capital | 5.85 | 2.41% |

数据来源：Wind，西南证券整理

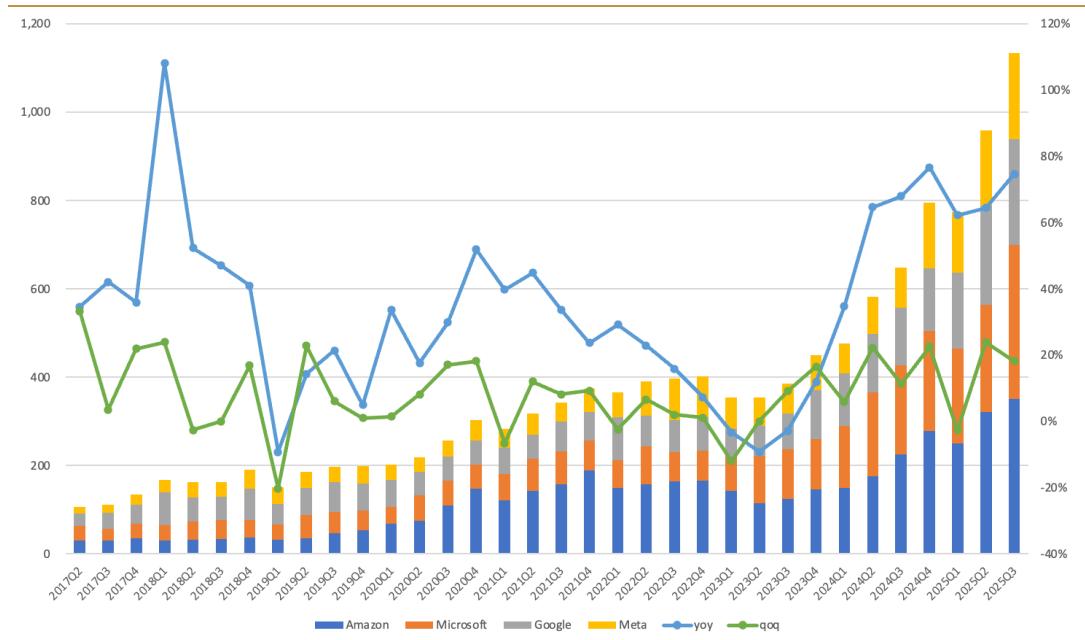
2 AI 算力全栈布局，下游 Capex 支撑高增长

2.1 下游资本支出持续攀升，数据中心业务高速增长

随着大模型的算法复杂度和训练/推理体量不断增长，算力基础设施的供给水平将直接影响应用落地及其迭代的进程。反应各大互联网厂商对算力基础设施投入最直接的资本支出 (Capex) 指标近三年来也在不断攀升。2023 年一季度至今，北美四大互联网公司 (亚马逊、微软、谷歌、Meta) 的 Capex 总额基本实现每季度均环比增长的态势，2025 年三季度四家公司的 Capex 总额达到 1133.2 亿美元，同比和环比分别增长 74.7%、18.2%，相比 23Q1 的体量增长了 2.2 倍。

资本支出的结构方面，微软率先在本轮 AI 周期中加大投资力度，从 23Q1 的 78 亿美元增长至 25Q3 的 349 亿美元，增幅高达 3.5 倍；谷歌 23Q3 开始显著提升 Capex 力度，截至 25Q3 单季度支出近 240 亿美元；亚马逊 24Q1 开始扩张 Capex，从 24Q1 的 149 亿美元提升至 25Q3 的 351 亿美元，目前其 Capex 体量与微软基本持平；Meta 从 24Q2 开始显著上调 Capex，25Q3 其体量为 193.7 亿美元，同比增长 1.1 倍。

图 8：北美四大互联网公司季度 Capex (亿美元)

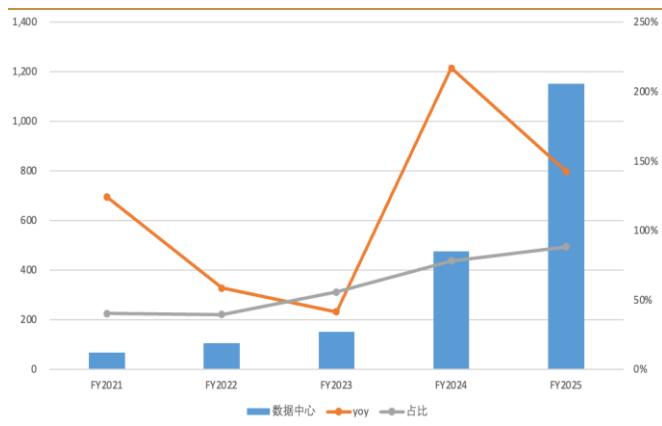


数据来源：各公司公告，西南证券整理

下游巨大的资本支出规模是英伟达数据中心业务高速成长的核心驱动力。公司数据中心业务近五年收入 CAGR 高达 103.7%，FY2026Q3 财季数据中心收入 512.2 亿美元，同比增长 66.4%，占公司整体收入比重 89.8%。

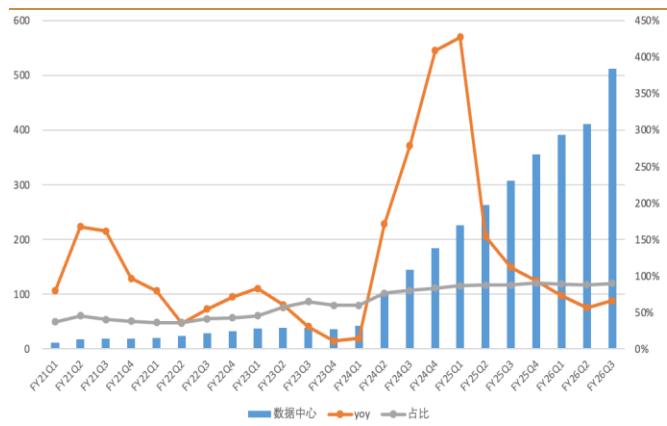
自 FY2024Q2 财季起，随着 AI GPU 的大量出货，数据中心业务迎来拐点。收入单季度平均增速 184%，FY26Q3 体量几乎是 FY24Q2 的 5 倍；高盈利的 AI GPU 推动公司整体毛利率的显著提升，从 FY24Q1 的 64.6% 提升至 FY26Q3 的 73.4%，期间毛利率最高时为 78.4%。

图 9：英伟达数据中心业务年度表现 (亿美元)



数据来源：公司公告，西南证券整理

图 10：英伟达数据中心业务季度表现 (亿美元)



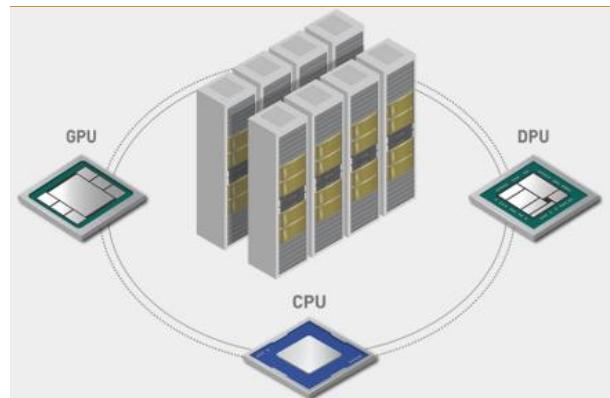
数据来源：公司公告，西南证券整理

2.2 AI 算力产品全栈布局，新平台重塑 AI 芯片格局

英伟达的数据中心解决方案整体提供了 **NVIDIA 加速计算统一平台**，从 AI、数据分析，到高性能计算 (HPC)，再到渲染，数据中心都是攻克某些重要挑战的关键。端到端的 NVIDIA 加速计算平台对硬件和软件进行了集成，可为企业构建强大而安全的基础设施蓝图，支持在现代化工作负载中实施从开发到部署的操作。

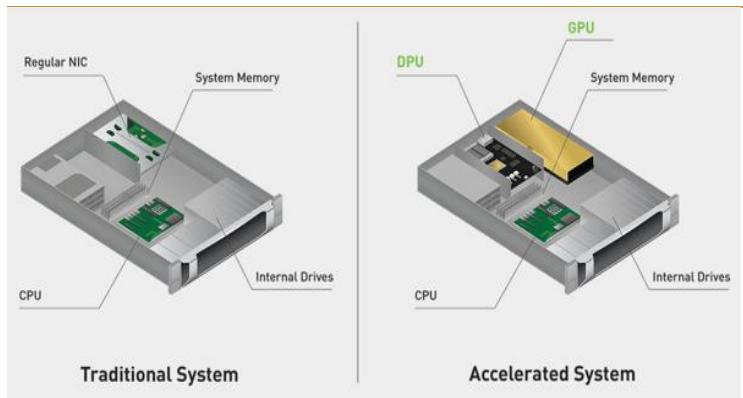
借助基于 **GPU、DPU 和 CPU 架构** 构建的 **NVIDIA 加速计算平台**，英伟达或重塑 AI 时代的数据中心。NVIDIA 加速计算平台可提供支持从用于改善业务预测的数据分析、到自动驾驶汽车使用的 AI、再到用于医疗诊断的高级可视化等多种应用的基础架构。采用 NVIDIA 加速系统的每个服务器和工作站都将配备计算加速器，为当今使用的现代应用（包括 AI、可视化和自主机器等）提供支持。其中许多系统还会配备 DPU，可加速对云原生和云计算框架至关重要的网络、存储和安全服务。

图 11：NVIDIA 加速计算平台



数据来源：英伟达官网，西南证券整理

图 12：NVIDIA 加速系统与传统系统的对比



数据来源：英伟达官网，西南证券整理

2.2.1 步入 Blackwell 时代的系统级飞跃

数据中心 **GPU** 方面，公司产品线三年内经历了 **Ampere、Hopper、Blackwell** 架构的更新迭代。英伟达的计算平台架构迭代时间由 2 年缩短至 1 年，意味着新技术和新工艺能更快地应用到产品中，这种快速迭代策略可以确保英伟达在技术竞争中始终保持领先优势。公司于 2024 年推出 Blackwell GPU，配备 NVLINK 5 Switch (1800GB/s) 和 CX8 SuperNic，支持 Spectrum-X800 以太网交换机和 Quantum-X800 交换机；2025 年推出 Blackwell Ultra GPU，以及 Spectrum Ultra X800 以太网交换机。根据公司的产品路线图，2026 年公司将推出 Rubin 平台，而 2027 年预计推出 Rubin Ultra 平台。

图 13：英伟达 AI GPU 产品线规划



数据来源：英伟达官网，西南证券整理

公司最新量产的 Blackwell 平台的设计并非对单一 GPU 组件的堆砌，而是整合了 CPU、GPU、NVLink、NIC（特定技术组件）、以及网络交换机等多个元素的完整系统。

Blackwell 平台采用 2 颗 B200 GPU Die，通过 10TB/s 的高速链接，结合世界上最先进的 SerDes（高性能接口或连接技术）将它们紧密连接在一起。2 颗 B200 GPU 放置在一个计算节点上，并通过 1 颗 Grace CPU 进行高效协调。Grace CPU 不仅适用于训练场景，还在推理和生成过程中发挥关键作用，如快速检查点和重启；此外，它还能存储上下文，让 AI 系统拥有记忆，并能理解用户对话的上下文，这对于增强交互的连续性和流畅性至关重要。

NVLink 是一种高带宽、低延迟的 GPU 互连技术，最新的 NVLink 5 代和 6 代分别提供 1800GB/s 和 3600GB/s 的带宽。这些技术允许多个 GPU 高效协同工作，显著提高 AI 计算性能。通过 NVLink Spine 技术，NVIDIA 实现了不同 GPU 之间的高效连接，每秒数据传输能力达到 10TB，这使得整个系统性能提升令人瞩目，每秒处理能力达到了惊人的水平。

英伟达算力产品性能的提升不仅体现在计算速度上，还极大地降低了能耗。例如，训练 GPT-4 的能耗从 1000 GWh 减少到 3 GWh，代表了巨大的技术进步。

图 14：英伟达 Blackwell 平台产品组合



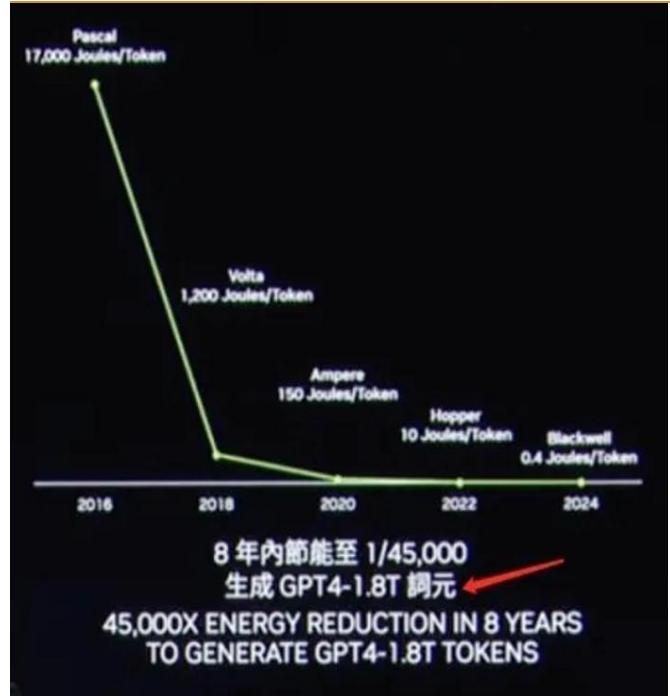
数据来源：英伟达官网，西南证券整理

AI 大模型需要整个 GPU 机架的协同工作。英伟达新的 DGX 系统能容纳数十万亿参数的大模型。Blackwell 系统 (NVL72) 共有 9 个计算节点，共计 72 个 GPU，构成一个庞大的计算集群。这些 GPU 通过全新的 NVLink 技术紧密相连，形成一个无缝的计算网络。NVL72 相较于上一代的 8 个 GPU，性能提升了 9 倍，带宽增加了 18 倍，AI FLOPS (每秒浮点运算次数) 提升了 45 倍，而功率仅增加了 10 倍。

为了应对深度学习和 AI 工厂的需求，英伟达开发了 Spectrum-X 以太网架构。通过 RDMA、拥塞控制、自适应路由和噪声隔离等技术，显著提升以太网性能，使其能够满足大规模 GPU 数据中心需求。Spectrum-X 专为 AI 设计，与传统以太网相比，显著提高生成式 AI 的网络性能。其核心在于端到端的通信能力，确保网络接口卡和交换机高效协作。

2025 年英伟达发布了 Spectrum-X800 Ultra，可支持 10 万卡集群，支持高达 512 路径 Radix。这种扩展能力使得计算集群规模大幅度增加，处理更大规模的数据和更复杂的计算任务。预计 2026 年英伟达将发布 Spectrum-X1600，可支持百万卡集群。这将是革命性的进步，使得 AI 能处理前所未有的大规模计算任务。百万卡集群不仅意味着更强的计算能力，还意味着更高的容错性和灵活性，能更好地适应各种复杂计算需求。此外，NVLink 有望升级至 6.0 版本，支持 3.6TB/s 的互联速率。高速互联技术将极大提高数据传输效率，减少数据传输过程中的延迟和瓶颈。对于大规模计算集群而言，高效的数据传输是保证计算性能的关键，NVLink 6.0 的推出将显著提升整个计算集群的性能和效率。

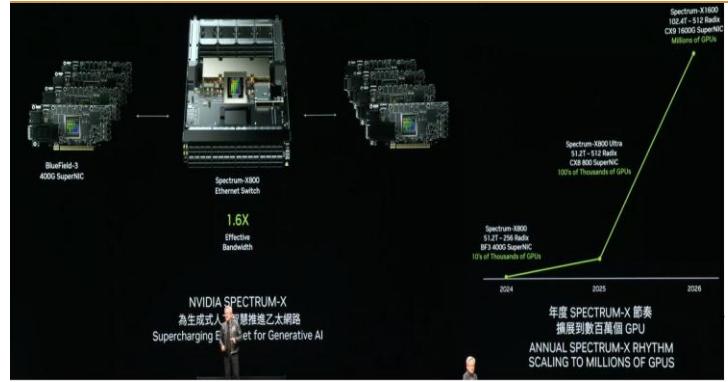
图 15：英伟达 AI GPU 的能耗随产品代际提升而快速降低



数据来源：英伟达官网，西南证券整理

图 16：英伟达 Blackwell 机架方案


数据来源：英伟达官网，西南证券整理

图 17：英伟达 Spectrum-X


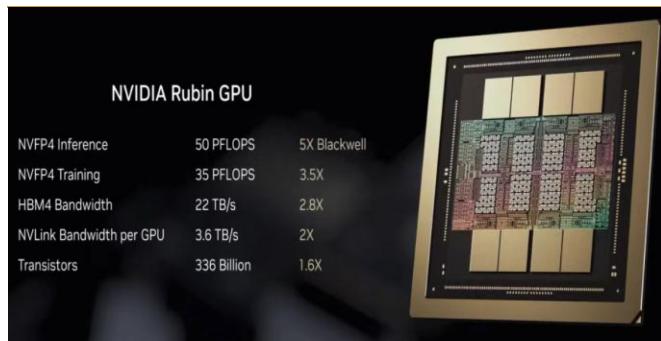
数据来源：英伟达官网，西南证券整理

2.2.2 Rubin 平台开启算力新纪元

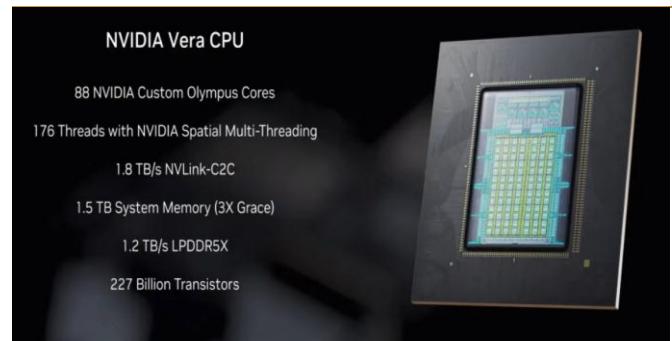
Rubin 是英伟达下一代架构平台。英伟达将于 2026 年量产 Rubin 平台。Rubin 平台整合了 Rubin GPU (2 颗 R200 GPU Die, 8 个 HBM4), Vera CPU, NVLink 6 交换机芯片, ConnectX-9 网卡、BlueField-4、Spectrum-6 等。

Rubin GPU 在关键指标上实现跨代跃升：FP4 推理性能提升至 50 PFLOPS，是 Blackwell 的 5 倍；FP4 训练性能提升至 35 PFLOPS，是 Blackwell 的 3.5 倍；HBM4 内存带宽提升至 22 TB/s，是 Blackwell 的 2.8 倍；单 GPU 的 NVLink 互连带宽翻倍至 3.6 TB/s。

Vera CPU 是专为数据移动和 Agentic 处理设计的核心组件，专注于 AI 推理场景中的数据调度和多步骤推理逻辑处理。Vera CPU 包含 2270 亿个晶体管，具有 88 个基于 ARM 架构自研的 Olympus 内核和 176 个线程，配备 1.5TB 系统内存（是上一代 Grace CPU 的 3 倍），LPDDR5X 带宽为 1.2TB/s，通过 1.8TB/s NVLink-C2C 接口实现与 Rubin GPU 之间的一致性内存连接。

图 18：英伟达 Rubin GPU 性能


数据来源：英伟达官网，西南证券整理

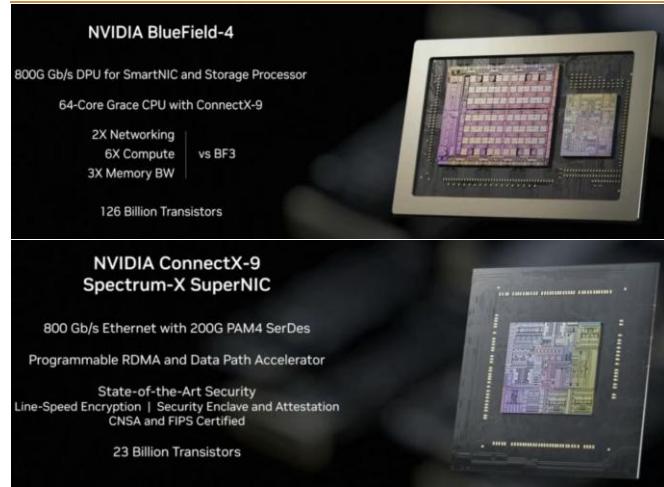
图 19：英伟达 Vera CPU 性能


数据来源：英伟达官网，西南证券整理

在 Rubin 架构中，NVLink 6 负责 GPU 框内协同计算，BlueField-4 负责上下文与数据调度，而 ConnectX-9 则承担系统对外的高速网络连接。它确保 Rubin 系统能够与其他机架、数据中心和云平台高效通信，是大规模训练和推理任务顺利运行的前提条件。NVLink 6 通过 3.6 TB/s 的带宽和网络内计算能力，让 Rubin 架构中的 72 个 GPU 能协同工作，这是实现降低推理成本的关键基础设施。

Vera Rubin NVL72 是业界首个支持机架规模机密计算的平台，包含 72 个 GPU，搭配 20TB HBM4（每个 GPU 的显存容量 288GB），将提供 3.6 EFLOPS 的 FP4 推理算力、以及 2.5 ExaFLOPS 的 FP4 训练算力。NVLink 将吞吐量翻倍，总速度为 260TB/s，机架之间有一个新 ConnectX-9 链路，速度为 28.8TB/s，是 B300 和 ConnectX-8 带宽的两倍。

图 20：英伟达 BlueField-4 和 ConnectX-9 性能



数据来源：英伟达官网，西南证券整理

图 21：英伟达 Vera Rubin NVL72 性能

| NVIDIA Vera Rubin NVL72 | | |
|-------------------------|--------------|--------------|
| NVFP4 Inference | 3.6 EFLOPS | 5X Blackwell |
| NVFP4 Training | 2.5 EFLOPS | 3.5X |
| LPDDR5X Capacity | 54 TB | 3X |
| HBM Capacity | 20.7 TB | 1.5X |
| HBM4 Bandwidth | 1.6 PB/s | 2.8X |
| Scale-Up Bandwidth | 260 TB/s | 2X |
| Transistors | 220 Trillion | 1.7X |

数据来源：英伟达官网，西南证券整理

英伟达在 CES 2026 上宣布推出基于 Rubin 架构的新一代 DGX SuperPOD，将 Rubin 从单机架扩展到整个数据中心的完整方案。新一代 DGX SuperPOD 由 8 个 Vera Rubin NVL72 机架组成，每个机架 72 个 GPU，总共 576 个 GPU；机架内部通过 NVLink 6 将 72 块 Rubin GPU 连接在一起，能够在一个机架内完成大规模推理和训练任务。

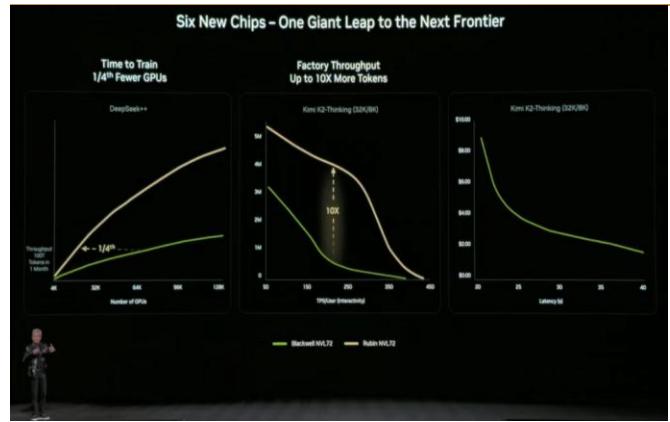
相比 Blackwell 平台，Rubin 可将推理阶段的 token 成本最高降低 10 倍，并将训练混合专家模型 MoE 所需的 GPU 数量减少至原来的 1/4。模型可以运行百万 token 的长下文，企业级 AI 应用也可以部署。

图 22：英伟达 Vera Rubin DGX SuperPOD



数据来源：英伟达官网，西南证券整理

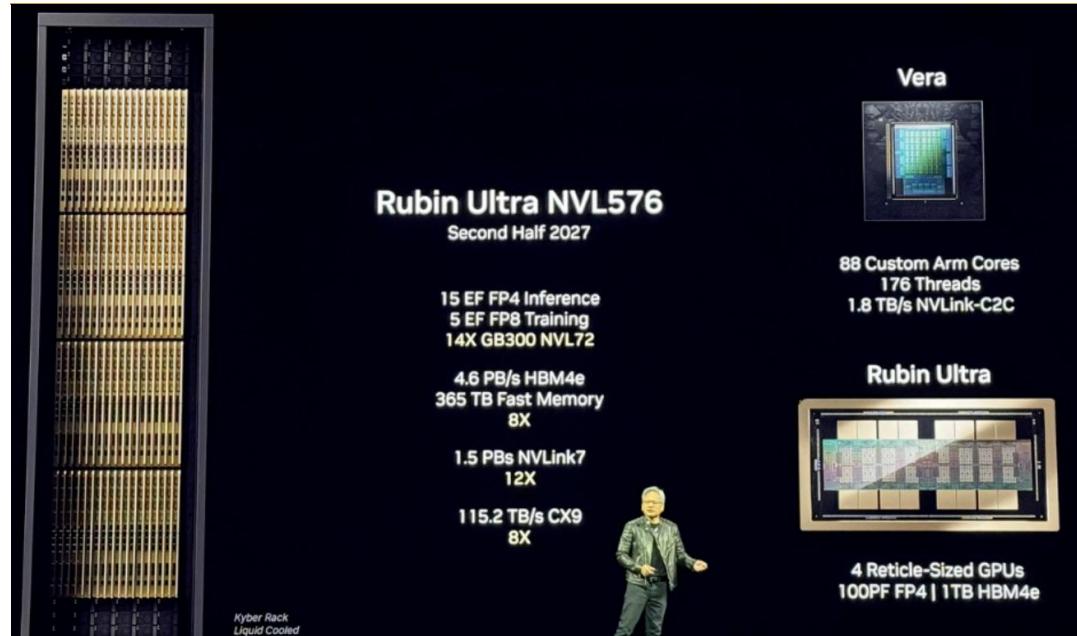
图 23：英伟达 Rubin 平台性能相比 Blackwell 大幅提升



数据来源：英伟达官网，西南证券整理

英伟达计划于 2027 年推出更强大的 Rubin Ultra 平台。其采用 4 个 GPU Die，配备 16 个 HBM4e，每封装总容量达 1TB，而 Vera CPU 则将沿用与 Rubin 相同的型号。完整的“Kyber 机架”为 Rubin Ultra NVL576，将通过 144 个封装部署 576 个 GPU；机架 HBM4e 总容量达 144TB，总带宽为 4.6PB/s。这意味着单个机架的封装数量是当前 Blackwell Ultra 的两倍，GPU 数量则达到四倍。Rubin Ultra NVL576 提供 15 ExaFLOPS FP4 推理算力、以及 5 ExaFLOPS FP8 训练算力。机架或采用旋转 90 度的计算托盘设计，密度较当前机架显著提升。此水冷塔式机架的功耗将达到 600 千瓦。

图 24：英伟达 Rubin Ultra 平台方案

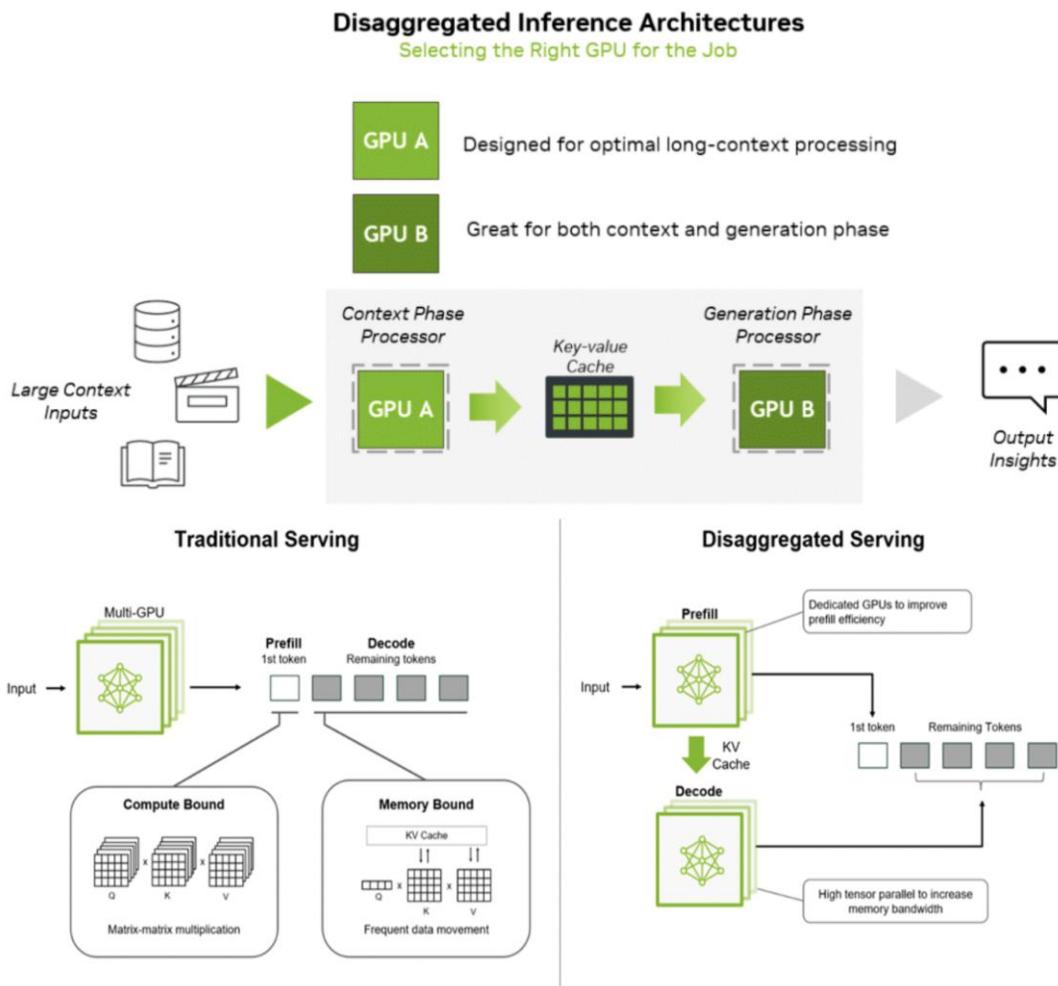


数据来源：英伟达官网，西南证券整理

英伟达于 2025 年三季度发布了 Rubin CPX GPU，这是一款专为推理的 Prefill 阶段优化的解决方案，单芯片的 Rubin CPX GPU 在设计上更强调计算性能而非内存带宽。通过硬件针对推理的 prefill 和 decode 两个不同的阶段进行设计，分离式服务可发挥其全部潜力。

推理过程中的 Prefill 阶段往往大量消耗计算性能，而只少量使用内存带宽，因此在拥有大量昂贵且带宽很高的 HBM 的芯片上运行 Prefill 会造成资源浪费。由于 Prefill 本质上会低效利用内存带宽资源，因此减少浪费的一种方法是降低内存的容量和成本。英伟达的 Rubin CPX GPU 在设计上对内存带宽进行了瘦身简化、在计算能力上相对更加厚实，并采用了容量更小、成本更便宜的 GDDR7 内存。

Rubin CPX GPU 提供了 20 PFLOPS 的 FP4 稠密算力，内存带宽为 2TB/s；配备 128GB 的 GDDR7 内存，与 VR200 相比容量更小成本更便宜；其 TDP 为 800W，远低于 VR200。

图 25：英伟达分离式推理架构


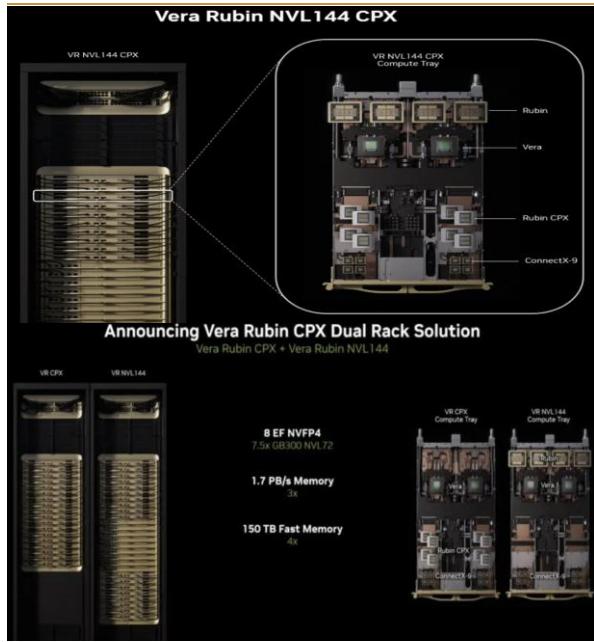
数据来源：英伟达官网，西南证券整理

Rubin CPX GPU 推出后，英伟达 VR200 机架的版本或有三类：

- VR200 NVL144: 18 个计算托盘 (Compute Trays) 中包含 72 个 GPU，每个计算托盘中有 4 个 R200 GPU。
- VR200 NVL144 CPX: 在 18 个计算托盘中，除 144 个 Rubin CPX GPU 封装外还有 72 个逻辑 GPU 封装，每个计算托盘包含 4 个 R200 GPU 封装和 8 个 Rubin CPX GPU 封装。
- Vera Rubin CPX 双机架：两个独立机架——1 个 VR200 NVL144 机架加上一个 VR CPX 机架，总共包含 18 个计算托盘中的 144 个 Rubin CPX GPU，每个计算托盘中有 8 个 Rubin CPX GPU。

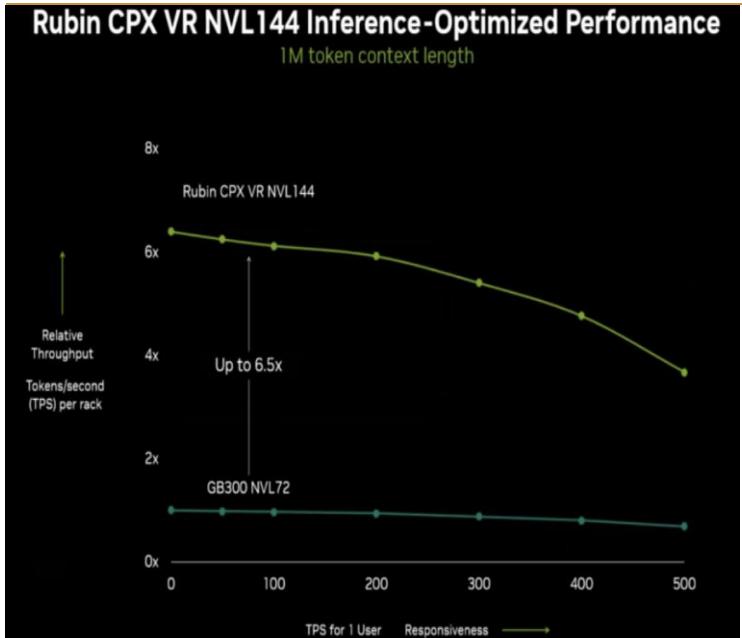
英伟达透露，搭载 Rubin CPX GPU 的 VR200 机架在处理大型上下文窗口时的性能，比 GB300 NVL72 高出最多 6.5 倍。

图 26：英伟达 Vera Rubin CPX 机架方案



数据来源：英伟达官网，西南证券整理

图 27：英伟达 Vera Rubin CPX 性能



数据来源：英伟达官网，西南证券整理

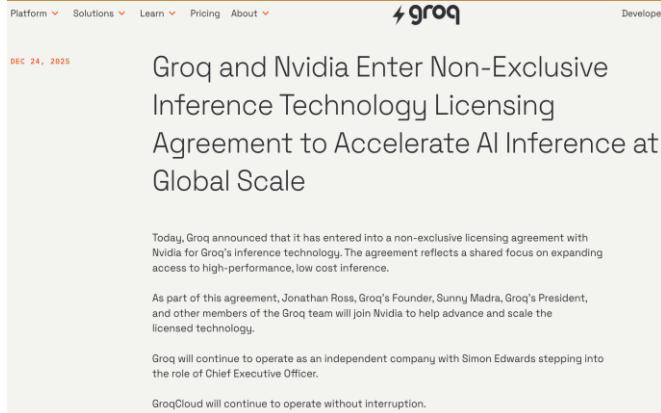
2.2.3 携手 Groq 重塑 AI 推理格局

2025年12月24日，英伟达宣布以200亿美元购买Groq资产。英伟达获得Groq推理技术的非排他性授权，双方共同致力于拓展高性能、低成本推理技术的应用。Groq创始人Jonathan Ross及部分核心团队成员将加入英伟达，以推动和扩展该授权技术；而Groq将继续作为独立公司运营，GroqCloud云服务不受影响。

Groq的核心技术是一种专为AI推理设计的语言处理单元（LPU），其架构与主流的AI GPU存在根本性差异。

- **SRAM-only 内存设计：**Groq芯片摒弃了外部DRAM或HBM，将所有模型权重、键值缓存（KV Cache）和激活值都存储在片上的SRAM中。这消除了“内存墙”瓶颈，是其实现超低延迟的关键。
- **确定性VLIW架构：**采用完全确定性的超长指令字（VLIW）架构，没有缓冲区。这种设计由编译器精确调度执行，从而实现可预测的性能。
- **大规模芯片互连：**由于单芯片SRAM容量有限（当前一代产品为230MB），运行Mixtral 8x7b等实用模型需要通过网络将576颗芯片连接在一起，进行模型分片。
- **芯片规格：**当前一代芯片采用格罗方德14nm工艺，裸片尺寸约为725mm²。

图 28：英伟达与 Groq 签署非独家推理技术许可协议



Platform Solutions Learn Pricing About

DEC 24, 2025

Groq and Nvidia Enter Non-Exclusive Inference Technology Licensing Agreement to Accelerate AI Inference at Global Scale

Today, Groq announced that it has entered into a non-exclusive licensing agreement with Nvidia for Groq's inference technology. The agreement reflects a shared focus on expanding access to high-performance, low cost inference.

As part of this agreement, Jonathan Ross, Groq's Founder, Sunny Madra, Groq's President, and other members of the Groq team will join Nvidia to help advance and scale the licensed technology.

Groq will continue to operate as an independent company with Simon Edwards stepping into the role of Chief Executive Officer.

GroqCloud will continue to operate without interruption.

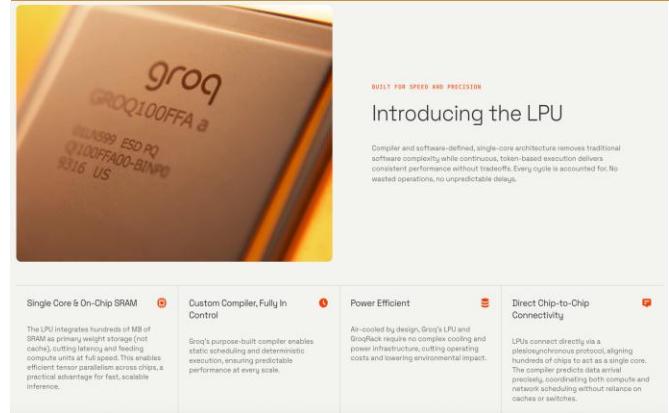
数据来源：Groq 官网，西南证券整理

传统 AI GPU 架构紧密耦合 HBM 显存，HBM 依靠大容量和高带宽的优势，在大规模并行计算和模型训练中占据内存的主导地位。随着模型规模持续扩大，以 HBM 为核心的 GPU 架构在推理应用中逐渐暴露延迟、功耗、成本等瓶颈，且受 HBM 产能和 CoWoS 封装供给的影响，HBM 和 AI GPU 的供应也受限制。

Groq LPU 采用创新的片上内存架构，将权重和 KV 缓存从外部 HBM 转移到 AI 加速器上的 SRAM 中，SRAM 的处理速度是传统内存的 10 倍，可实现极低的延迟、更快的执行和更低的功耗。这一设计专门为推理优化，消除了对外部内存的过度依赖，也绕过了 HBM 和 CoWoS 封装的供给瓶颈。

图 29：Groq LPU 介绍

图 29：Groq LPU 介绍



BUILT FOR SPEED AND PRECISION

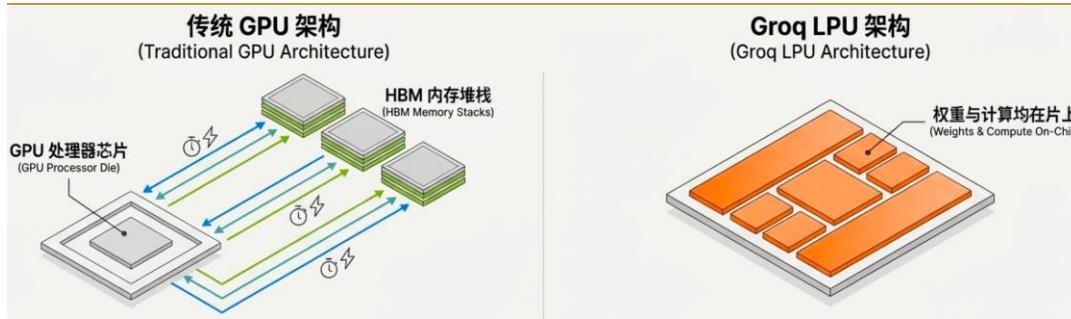
Introducing the LPU

Compiler and software-defined, single-core architecture removes traditional software complexity while continuous, token-based execution delivers consistent performance without tradeoffs. Every cycle is accounted for. No wasted operations, no unpredictable delays.

| | | | |
|---|---|---|---|
| Single Core & On-Chip SRAM | Custom Compiler, Fully In Control | Power Efficient | Direct Chip-to-Chip Connectivity |
| This LPU integrates hundreds of MB of SRAM as primary weight storage (not cache), cutting latency and feeding compute units at full speed. This enables efficient parallelism across chips, a practical advantage for fast, scalable inference. | Groq's purpose-built compiler enables static scheduling and deterministic execution, ensuring predictable performance at every scale. | An-soled by design, Groq LPU and GroqPaks require no complex cooling and power infrastructure, cutting operating costs and lowering environmental impact. | LPUx connect directly via a plesiochronous protocol, aligning hundreds of chips to act as a single core. The compiler predicts data arrival times to enable both compute and network scheduling without reliance on caches or switches. |

数据来源：Groq 官网，西南证券整理

图 30：Groq 的 LPU：一种创新的片上 AI 内存架构



数据来源：西南证券整理

英伟达获得 Groq 的非独家推理技术许可协议，旨在构建一条差异化、专注于极致低延迟和低每 token 成本的推理产品线。这补全了其在 GPGPU 之外的市场版图，以应对那些对延迟和成本极为敏感的应用场景。

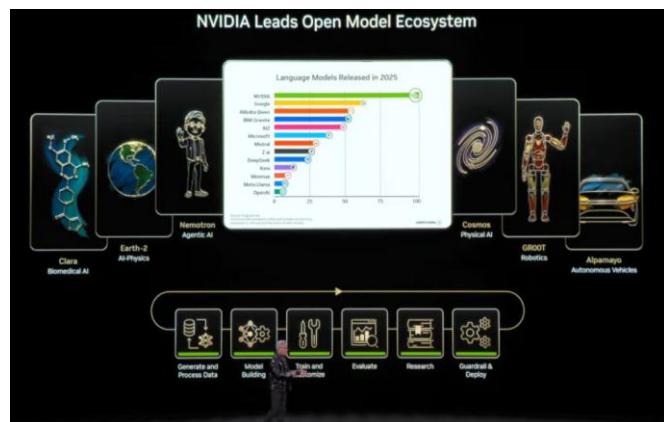
AI 硬件的未来或走向训练与推理的持久架构分叉：训练环节，预计仍将由与 HBM 紧密耦合的 GPU 架构主导，训练系统为处理海量数据集和实现最大并行计算能力而优化，对内存容量和带宽有极高的要求；推理环节或将出现分化，高端、长上下文推理或仍需 HBM，由于 SRAM 在解码阶段有巨大优势，类似 Groq 的基于 SRAM 架构的芯片或适用于延迟敏感、小批量、低成本的部分推理场景。这一分叉或重塑 AI 计算的价值链。

3 物理 AI 的“ChatGPT 时刻”即将到来，或重塑公司 AI 底座价值

英伟达 CEO 黄仁勋在 CES 2026 上表示，物理 AI 的“ChatGPT 时刻”即将到来。物理 AI 的核心任务不仅是处理数字信息，更是要让 AI 理解物理世界的运行逻辑，解决真实世界中的因果推理问题。

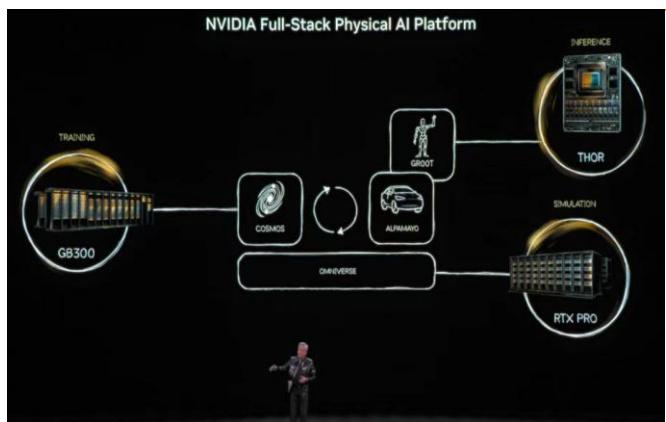
为了支撑这一愿景，英伟达在 CES 2026 上宣布扩展其开源模型生态，新增和更新了一系列模型、数据集、代码库和工具。该生态覆盖六大领域：生物医学 AI (Clara)、AI 物理模拟 (Earth-2)、Agentic AI (Nemotron)、物理 AI (Cosmos)、机器人 (GROOT) 和自动驾驶 (Alpamayo)。其中，英伟达物理 AI 模型的更新包括了：用于理解和生成物理世界视频的 Cosmos，自动驾驶视觉-语言-行动模型 Alpamayo，机器人通用基础模型 Isaac GROOT。

图 31：英伟达开源模型生态



数据来源：英伟达官网，西南证券整理

图 32：英伟达全栈物理 AI 平台



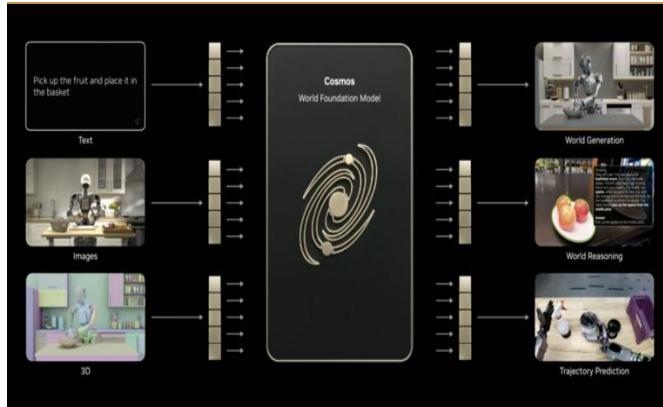
数据来源：英伟达官网，西南证券整理

物理 AI 面对的挑战很多，包括物理世界复杂多变，采集真实数据又慢又贵等。而合成数据是其中一条出路，英伟达相应推出了 Cosmos。Cosmos 是一个开源的物理 AI 世界基础模型，目前已经用海量视频、真实驾驶与机器人数据，以及 3D 模拟做过预训练。Cosmos 能理解世界是怎么运行的，可以把语言、图像、3D 和动作联系起来，能实现不少物理 AI 技能，比如生成内容、做推理、预测轨迹等。Cosmos 可以通过语言、图像、3D 动作的关联，生成逼真的具备一定的物理规律的视频。这种模拟世界的能力对于训练机器人至关重要，因为它能生成罕见的边缘场景数据，从而避免了在真实物理世界中高昂且危险的数据采集过程。

Alpamayo 是一个面向自动驾驶领域的开源工具链，也是首个开源的视觉-语言-行动 (VLA) 推理模型。与之前仅开源代码不同，英伟达这次开源了从数据到部署的完整开发资源。Alpamayo 最大的突破在于它是推理型自动驾驶模型。传统自动驾驶系统基于“感知-规划-控制”的流水线架构，看到红灯就刹车，看到行人就减速，遵循预设规则，如果遇到预设规则之外的突发状况，系统往往反应僵化。而 Alpamayo 引入了推理能力，通过理解复杂场景中的因果关系，预测其他交通参与者的意图，甚至能处理需要多步思考的决策，比如在十字路口，它不只是识别出“前方有车”，还能通过观察做出“前方车辆可能要左转，应该等它先过”的推理。这种多步思考的能力让自动驾驶从“按规则行驶”升级到“像人一样按逻辑思考”，让车辆表现得更像经验丰富的人类驾驶员。

英伟达 DRIVE 系统将正式进入量产阶段，搭载该系统的全新梅赛德斯-奔驰 CLA 计划于 2026 年在美国上路。这款车将搭载 L2+ 级自动驾驶系统，采用“端到端 AI 模型+传统流水线”的混合架构。

图 33：英伟达物理 AI 开源世界模型——Cosmos



数据来源：英伟达官网，西南证券整理

图 34：英伟达开源 VLA 推理模型——Alpamayo



数据来源：英伟达官网，西南证券整理

机器人方面，英伟达正在利用其算力优势重塑机器人产业。包括 Boston Dynamics、Franka Robotics、LEM Surgical、LG Electronics、Neura Robotics 和 XRlabs 等在内的全球机器人领军企业，正在基于英伟达 Isaac 平台和 GR00T 基础模型开发覆盖工业机器人、手术机器人、人形机器人、消费级机器人等多个领域的产品。

为了解决机器人开发中各自为战的痛点，英伟达发布了 Isaac Lab-Arena 和 NVIDIA OSMO。Isaac Lab-Arena 是一套开源框架，它连接了主流基准测试，确保机器人在进厂前已经在虚拟世界里经过了千锤百炼。NVIDIA OSMO 则保证无论是在工作站还是混合云，都能统一调度数据生成、模型训练和测试，大幅缩短开发周期。

此外，全新的 Jetson T4000 模组直接将 Blackwell 架构带到了边缘端：其 FP4 算力达 1200 TFLOPS，是上一代的 4 倍；功耗为 70W，为能源受限的自主设备量身定做。

Isaac 平台和 GR00T 基础模型通过整合多模态感知与复杂任务规划，使人形机器人的大规模部署成为可能。随着物理 AI 基础模型的不断成熟，英伟达有望在未来将其在云端算力市场的绝对垄断地位无缝转移到具身智能机器人的“终端脑”市场。

图 35：英伟达在 CES 2026 上的机器人演示



数据来源：英伟达官网，西南证券整理

图 36：英伟达物理 AI “开源全家桶”



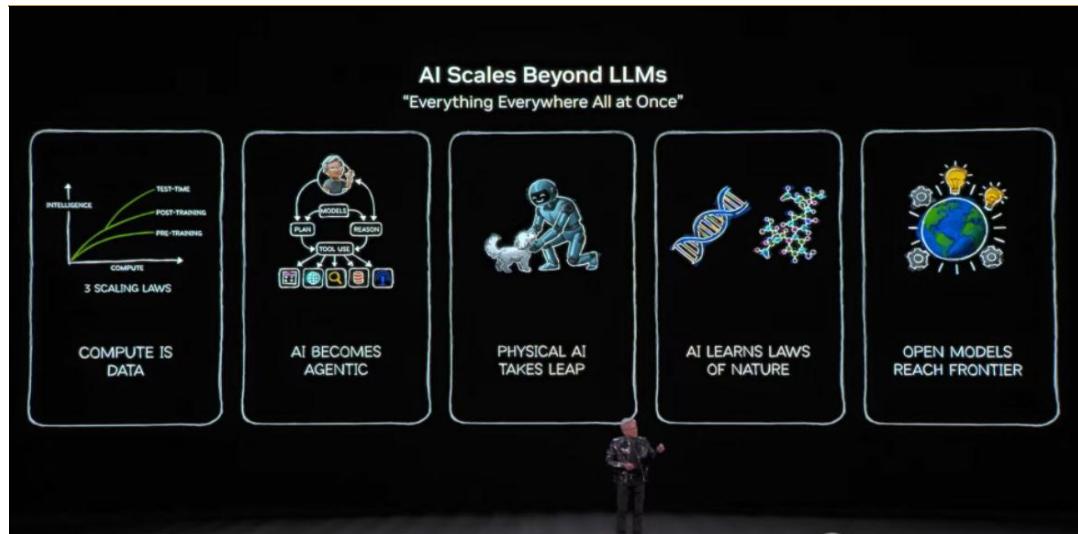
数据来源：英伟达官网，西南证券整理

英伟达不仅是一家芯片设计公司，更是一家定义了 AI 时代生产力法则的系统集成商。通过对算力资本开支趋势的精准把握，英伟达在下游巨头的资本支出竞争实现了自身业绩的高速增长。

从技术维度看，从 Blackwell 到 Rubin 的进阶，不仅是参数的堆叠，更是对内存墙、通信延迟及能耗比等行业极限的持续挑战。通过引入分离式推理架构和对 Groq 技术的整合，英伟达展现了其在多变的市场需求面前极高的进化速度。而物理 AI 的布局，则为其在生成式 AI 热潮之后，预先开辟了具身智能的新增长极。

在可预见的未来，英伟达的护城河将随着 CUDA 生态的深化和百万卡集群架构的落地而变得更加宽广。物理世界的 AI 智能进化不停止，全球算力工业对英伟达的依赖或将持续，其作为“AI 工业革命底座”的价值也将随之不断被重塑。

图 37：AI 不仅仅是大模型



数据来源：英伟达官网，西南证券整理

4 盈利预测与投资建议

关键假设：

假设 1：大模型训练和推理对 AI 加速卡需求持续提升，Blackwell 和 Rubin 系列量产，数据中心业务未来三年复合增速为 58% 左右。

假设 2：先进制程产能供给紧张和元器件涨价带动游戏显卡涨价，显卡 ASP 提升，游戏业务未来三年复合增速达 31%。

假设 3：智驾渗透率持续提升，智能汽车业务未来三年复合增速达到 28% 左右。

基于以上假设，我们预测公司 2026-2028 财年分业务收入如下表：

表 2：分业务收入

| 单位：百万美元 | | FY2025A | FY2026E | FY2027E | FY2028E |
|---------|----|------------|------------|------------|------------|
| 数据中心 | 收入 | 115,186.00 | 191,208.76 | 315,494.45 | 457,466.96 |
| | 增速 | 142.37% | 66.00% | 65.00% | 45.00% |

| 单位：百万美元 | | FY2025A | FY2026E | FY2027E | FY2028E |
|---------|----|------------|------------|------------|------------|
| 游戏 | 收入 | 11,350.00 | 17,025.00 | 21,281.25 | 25,537.50 |
| | 增速 | 8.64% | 50.00% | 25.00% | 20.00% |
| 专业视觉 | 收入 | 1,878.00 | 2,159.70 | 2,375.67 | 2,541.97 |
| | 增速 | 20.93% | 15.00% | 10.00% | 7.00% |
| 汽车 | 收入 | 1,694.00 | 2,371.60 | 3,011.93 | 3,554.08 |
| | 增速 | 55.27% | 40.00% | 27.00% | 18.00% |
| 其他 | 收入 | 389.00 | 427.90 | 462.13 | 485.24 |
| | 增速 | 27.12% | 10.00% | 8.00% | 5.00% |
| 合计 | 收入 | 130,497.00 | 213,192.96 | 342,625.44 | 489,585.74 |
| | 增速 | 114.20% | 63.37% | 60.71% | 42.89% |

数据来源: Wind, 西南证券

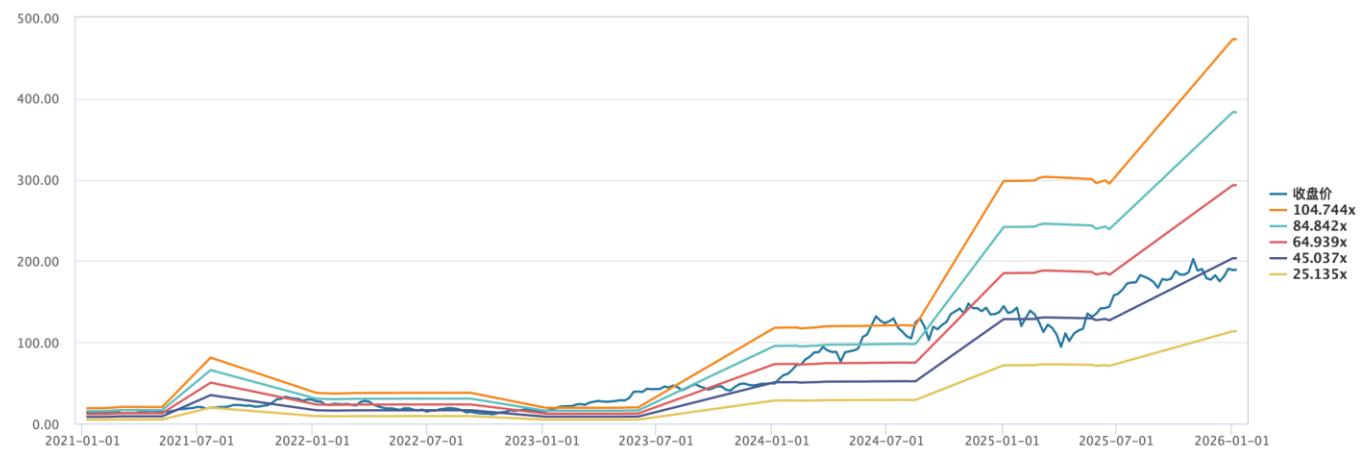
我们选取了半导体产业链的 3 家可比公司, 2025-2026 年平均估值为 158 倍 PE、51 倍 PE。公司是全球 AI 芯片龙头企业, 大模型训练与推理带动 AI 芯片需求爆发, 预计公司未来三年 Non-GAAP 净利润年复合增速为 58%, 当前股价下公司 2027 财年 (对应自然年 2026 年 2 月~2027 年 1 月) 对应估值为 23 倍 PE (Non-GAAP), 相比同业平均水平, 公司显著低估。结合业绩增速, 给予公司 2027 财年 30 倍 PE (Non-GAAP), 目标价 249.54 美元, 对应市值 6 万亿美元, 给予“买入”评级。

表 3: 可比公司估值情况

| 股票代码 | 证券简称 | 股价 (美元) | 市值 (亿美元) | Non-GAAP PE | | | |
|--------|------|---------|-------------|-------------|--------|-------|-------|
| | | | | 2024A | 2025E | 2026E | 2027E |
| AMD.O | AMD | 210.02 | 3419 | 123.80 | 78.04 | 44.88 | 29.65 |
| AVGO.O | 博通 | 343.50 | 16286 | 68.62 | 48.29 | 31.18 | 21.93 |
| INTC.O | 英特尔 | 42.63 | 2033 | -4.67 | 347.01 | 76.68 | 46.30 |
| 平均值 | | | | 96.21 | 157.78 | 50.91 | 32.63 |
| NVDA.O | 英伟达 | 189.11 | 45954 | 61.88 | 39.38 | 22.74 | 15.66 |

数据来源: Wind, 西南证券整理。注: 平均值剔除负值影响

图 38: 英伟达 PE/PB-BAND



数据来源: Wind, 西南证券整理

附：财务报表

| 资产负债表 (百万美元) | FY2025A | FY2026E | FY2027E | FY2028E | 利润表 (百万美元) | FY2025A | FY2026E | FY2027E | FY2028E |
|--------------|-----------|-----------|-----------|-----------|--------------|-----------|-----------|-----------|-----------|
| 货币资金 | 43210.00 | 70979.42 | 125747.51 | 192199.23 | 销售收入 | 130497.00 | 213192.96 | 342625.44 | 489585.74 |
| 应收账款 | 23065.00 | 37681.29 | 60558.14 | 86532.99 | 销售成本 | 32639.00 | 61372.58 | 91967.77 | 124954.54 |
| 预付款项 | 3771.00 | 7730.25 | 12423.39 | 17752.09 | 销售和管理费用 | 3491.00 | 5329.82 | 8565.64 | 12239.64 |
| 存货 | 10080.00 | 25221.61 | 37794.97 | 51351.18 | 研发费用 | 12914.00 | 15989.47 | 27410.04 | 41614.79 |
| 其他流动资产 | 0.00 | 0.00 | 0.00 | 0.00 | 营业利润 | 82992.00 | 126725.92 | 208887.19 | 302736.95 |
| 流动资产总计 | 80126.00 | 141612.57 | 236524.02 | 347835.48 | 其他非经营损益 | 1034.00 | 407.67 | 559.56 | 667.07 |
| 长期股权投资 | 0.00 | 0.00 | 0.00 | 0.00 | 税前利润 | 84026.00 | 127133.59 | 209446.74 | 303404.02 |
| 固定资产 | 6283.00 | 9515.33 | 12807.66 | 16003.32 | 所得税 | 11146.00 | 16059.81 | 26457.80 | 38326.71 |
| 无形资产 | 807.00 | 839.17 | 879.67 | 920.17 | 净利润 | 72880.00 | 111073.77 | 182988.94 | 265077.32 |
| 其他非流动资产 | 24385.00 | 30188.00 | 37266.00 | 43800.00 | 少数股东损益 | 0.00 | 0.00 | 0.00 | 0.00 |
| 非流动资产合计 | 31475.00 | 40542.50 | 50953.33 | 60723.49 | 归属于母公司股东的净利润 | 72880.00 | 111073.77 | 182988.94 | 265077.32 |
| 资产总计 | 111601.00 | 182155.06 | 287477.35 | 408558.97 | EBITDA | 84351.00 | 129279.71 | 212430.10 | 307321.91 |
| 应付账款 | 6310.00 | 8407.20 | 12598.32 | 17117.06 | NOPLAT | 70648.31 | 110901.62 | 182557.89 | 264366.99 |
| 其他流动负债 | 11737.00 | 20177.29 | 30235.98 | 41080.94 | EPS(美元) | 3.00 | 4.57 | 7.53 | 10.91 |
| 流动负债合计 | 18047.00 | 28584.49 | 42834.30 | 58198.00 | Non-GAAP 净利润 | 74266.00 | 116707.00 | 202128.00 | 293526.00 |
| 长期债务 | 8463.00 | 8463.00 | 8463.00 | 8463.00 | | | | | |
| 其他非流动负债 | 5764.00 | 4690.00 | 4268.00 | 3955.00 | | | | | |
| 非流动负债合计 | 14227.00 | 13153.00 | 12731.00 | 12418.00 | | | | | |
| 负债合计 | 32274.00 | 41737.49 | 55565.30 | 70616.00 | | | | | |
| 归属母公司权益 | 79327.00 | 140417.58 | 231912.04 | 337942.97 | | | | | |
| 少数股东权益 | 0.00 | 0.00 | 0.00 | 0.00 | | | | | |
| 股东权益合计 | 79327.00 | 140417.58 | 231912.04 | 337942.97 | | | | | |
| 负债和股东权益合计 | 111601.00 | 182155.06 | 287477.35 | 408558.97 | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| 现金流量表 (百万美元) | FY2025A | FY2026E | FY2027E | FY2028E | 财务分析指标 | FY2025A | FY2026E | FY2027E | FY2028E |
| 税后经营利润 | 71983.16 | 110717.60 | 182500.07 | 264494.51 | 成长能力 | | | | |
| 折旧与摊销 | 1864.00 | 1935.50 | 2917.17 | 4063.84 | 销售收入增长率 | 114.20% | 63.37% | 60.71% | 42.89% |
| 财务费用 | -1539.00 | 210.62 | 66.18 | -145.95 | EBIT 增长率 | 148.39% | 54.38% | 64.52% | 44.74% |
| 其他经营资金 | -8219.16 | -23179.66 | -25893.54 | -29496.04 | EBITDA 增长率 | 142.97% | 53.26% | 64.32% | 44.67% |
| 经营性现金净流量 | 64089.00 | 89684.07 | 159589.88 | 238916.35 | 税后利润增长率 | 144.89% | 52.41% | 64.75% | 44.86% |
| 投资性现金净流量 | -20421.00 | -11720.83 | -13261.13 | -13564.19 | | | | | |
| 筹资性现金净流量 | -42359.00 | -50193.82 | -91560.65 | -158900.4 | | | | | |
| 现金流量净额 | 1309.00 | 27769.42 | 54768.10 | 66451.72 | | | | | |

数据来源：公司公告，西南证券

分析师承诺

本报告署名分析师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，报告所采用的数据均来自合法合规渠道，分析逻辑基于分析师的职业理解，通过合理判断得出结论，独立、客观地出具本报告。分析师承诺不曾因，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获取任何形式的补偿。

投资评级说明

报告中投资建议所涉及的评级分为公司评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 个月内的相对市场表现，即：以报告发布日后 6 个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。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；香港市场以恒生指数为基准；美国市场以纳斯达克综合指数或标普 500 指数为基准。

买入：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅在 20% 以上

持有：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于 10% 与 20% 之间

公司评级 中性：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于 -10% 与 10% 之间

回避：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于 -20% 与 -10% 之间

卖出：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅在 -20% 以下

强于大市：未来 6 个月内，行业整体回报高于同期相关证券市场代表性指数 5% 以上

行业评级 跟随大市：未来 6 个月内，行业整体回报介于同期相关证券市场代表性指数 -5% 与 5% 之间

弱于大市：未来 6 个月内，行业整体回报低于同期相关证券市场代表性指数 -5% 以下

重要声明

西南证券股份有限公司（以下简称“本公司”）具有中国证券监督管理委员会核准的证券投资咨询业务资格。

本公司与作者在自身所知情范围内，与本报告中所评价或推荐的证券不存在法律法规要求披露或采取限制、静默措施的利益冲突。

《证券期货投资者适当性管理办法》于 2017 年 7 月 1 日起正式实施，本报告仅供本公司签约客户使用，若您并非本公司签约客户，为控制投资风险，请取消接收、订阅或使用本报告中的任何信息。本公司也不会因接收人收到、阅读或关注自媒体推送本报告中的内容而视其为客户。本公司或关联机构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或争取提供投资银行或财务顾问服务。

本报告中的信息均来源于公开资料，本公司对这些信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌，过往表现不应作为日后的表现依据。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告，本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本报告仅供参考之用，不构成出售或购买证券或其他投资标的要约或邀请。在任何情况下，本报告中的信息和意见均不构成对任何个人的投资建议。投资者应结合自己的投资目标和财务状况自行判断是否采用本报告所载内容和信息并自行承担风险，本公司及雇员对投资者使用本报告及其内容而造成的一切后果不承担任何法律责任。

本报告及附录版权为西南证券所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。如引用须注明出处为“西南证券”，且不得对本报告及附录进行有悖原意的引用、删节和修改。未经授权刊载或者转发本报告及附录的，本公司将保留向其追究法律责任的权利。

西南证券研究院

上海

地址：上海市浦东新区陆家嘴 21 世纪大厦 10 楼

邮编：200120

北京

地址：北京市西城区金融大街 35 号国际企业大厦 A 座 8 楼

邮编：100033

深圳

地址：深圳市福田区益田路 6001 号太平金融大厦 22 楼

邮编：518038

重庆

地址：重庆市江北区金沙门路 32 号西南证券总部大楼 21 楼

邮编：400025

西南证券机构销售团队

| 区域 | 姓名 | 职务 | 座机 | 手机 | 邮箱 |
|----|-----|-------------------------|--------------|-------------|----------------------|
| 上海 | 蒋诗烽 | 院长助理、研究销售部经理、 上海销售主管 | 021-68415309 | 18621310081 | jsf@swsc.com.cn |
| | 崔露文 | 销售岗 | 15642960315 | 15642960315 | clw@swsc.com.cn |
| | 李煜 | 销售岗 | 18801732511 | 18801732511 | yliyu@swsc.com.cn |
| | 汪艺 | 销售岗 | 13127920536 | 13127920536 | wyf@swsc.com.cn |
| | 戴剑箫 | 销售岗 | 13524484975 | 13524484975 | daijx@swsc.com.cn |
| | 李嘉隆 | 销售岗 | 15800507223 | 15800507223 | ljlong@swsc.com.cn |
| | 欧若诗 | 销售岗 | 18223769969 | 18223769969 | ors@swsc.com.cn |
| | 蒋宇洁 | 销售岗 | 15905851569 | 15905851569 | jyj@swsc.com.c |
| | 贾文婷 | 销售岗 | 13621609568 | 13621609568 | jiawent@swsc.com.cn |
| | 张嘉诚 | 销售岗 | 18656199319 | 18656199319 | zhangjc@swsc.com.cn |
| 北京 | 毛玮琳 | 销售岗 | 18721786793 | 18721786793 | mw1@swsc.com.cn |
| | 李杨 | 北京销售主管 | 18601139362 | 18601139362 | yfly@swsc.com.cn |
| | 张岚 | 销售岗 | 18601241803 | 18601241803 | zhanglan@swsc.com.cn |
| | 姚航 | 销售岗 | 15652026677 | 15652026677 | yhang@swsc.com.cn |
| | 杨薇 | 销售岗 | 15652285702 | 15652285702 | yangwei@swsc.com.cn |
| | 王宇飞 | 销售岗 | 18500981866 | 18500981866 | wangyuf@swsc.com |
| | 王一菲 | 销售岗 | 18040060359 | 18040060359 | wyf@swsc.com.cn |

| | | | | | |
|----|-----|-----|-------------|-------------|--------------------|
| 广深 | 张鑫 | 销售岗 | 15981953220 | 15981953220 | zhxin@swsc.com.cn |
| | 马冰竹 | 销售岗 | 13126590325 | 13126590325 | mbz@swsc.com.cn |
| | 龚之涵 | 销售岗 | 15808001926 | 15808001926 | gongzh@swsc.com.cn |
| | 唐茜露 | 销售岗 | 18680348593 | 18680348593 | txl@swsc.com.cn |
| | 文柳茜 | 销售岗 | 13750028702 | 13750028702 | wlq@swsc.com.cn |
| | 林哲睿 | 销售岗 | 15602268757 | 15602268757 | lzs@swsc.com.cn |