

Oracle 签 300 亿美元大单，英伟达算力需求旺盛

2025 年 07 月 09 日

► **事件：**7 月 1 日，据新浪财经报道，Oracle 公开表示已签署多项大型云服务协议，其中一项关键订单预计自 2028 财年起，每年将为公司带来超过 300 亿美元的营收，约占 2025 财年总收入的 52%，相当于公司 OCI 业务收入的近三倍。为满足大体量算力需求，Oracle 已采购约 40 万张英伟达 GB200 高端算力卡，成为全球第二大 NV 算力卡持有者。我们此前持续强调英伟达 GB200 出货节奏改善，而 Oracle 订单超预期，验证了英伟达的 GPU 算力需求旺盛。

► **除了传统云厂商的需求以外，英伟达 GPU 还有三大需求场景。Oracle 的算力供给更偏重租赁，其主要需求由第三方大模型训练、主权 AI 建设和 To B 定制化私有云三大应用场景驱动——**1) 第三方 LLM 训练：以 OpenAI、Anthropic 等为代表的 AI 公司正在推进超大参数量模型的开发训练，对高密度通用算力资源的需求持续攀升；2) 主权 AI：随着数据安全和自主的战略地位提升，欧洲、中东、日本、韩国等国家和地区正在加速推进本土云和主权 AI 基础设施建设，英伟达成为关键合作伙伴；3) To B 市场：大型行业客户出于业务安全、合规和专属定制化需求，持续建设私有云和混合云 AI 基础设施，而英伟达 GPU 具备算力性能好、通用性强的优势，成为相关企业的理想选择。

► **全球 AI 算力需求旺盛，GPU 与 ASIC 均快速增长。我们同时看好英伟达 GPU 和云厂商 ASIC 的快速增长。**1) GPU 阵营以英伟达、Oracle、Coreweave 等厂商为代表，英伟达的 GPU 凭借其高度通用性、成熟生态和出色扩展能力，成为支撑大模型训练、主权 AI 建设以及多行业 AI 基础设施的主流选择；2) ASIC 阵营包括谷歌、亚马逊、Meta 等 CSP 厂商和博通、Marvell 等 ASIC 定制厂商，为大厂的 ASIC 聚焦于 AI 推理、搜索推荐等特定任务的加速和成本优化，凭借自研或合作定制芯片，在降低功耗、提升单位算力性价比方面形成差异化优势。在目前算力需求高增的背景下，GPU 和 ASIC 均有望实现快速增长。

投资建议：我们认为，算力的长期需求，不仅在于云厂商的 Token 数增长，ROI 闭环等，主权 AI、大模型训练加速以及 B 端客户私有云的部署均拉动了算力的强劲需求。**我们继续看好 ASIC+NV 链的投资逻辑**，后续重点关注优质龙头的中报业绩表现，以及产品/客户突破，建议关注：**ASIC：芯原股份；服务器：工业富联、华勤技术；PCB：沪电股份、胜宏科技、广合科技、生益科技、威尔高；光通信：旭创科技、新易盛、东山精密、长芯博创、太辰光；铜连接：瑞可达、沃尔核材；电源：江海股份、禾望电气。**

► **风险提示：**算力需求不及预期的风险；AI 应用落地不及预期的风险；产业链供应受限的风险；全球贸易政策变动的风险。

重点公司盈利预测、估值与评级

代码	简称	股价 (元)	EPS (元)			PE (倍)			评级
			2025E	2026E	2027E	2025E	2026E	2027E	
688521.SH	芯原股份	91.08	-0.18	0.07	0.61	/	1301	149	推荐
601138.SH	工业富联	23.98	1.52	1.90	2.27	16	13	11	推荐
603296.SH	华勤技术	83.17	3.62	4.52	5.45	23	18	15	推荐
002463.SZ	沪电股份	45.80	1.89	2.32	2.71	24	20	17	推荐
002384.SZ	东山精密	40.46	1.57	2.13	2.70	26	19	15	推荐
300548.SZ	长芯博创	61.47	1.38	1.77	2.24	45	35	27	推荐
688800.SH	瑞可达	45.38	1.78	2.32	2.89	25	20	16	推荐
300476.SZ	胜宏科技	144.23	5.40	7.38	8.99	27	20	16	/

资料来源：Wind，民生证券研究院预测；（注：股价为 2025 年 7 月 7 日收盘价；未覆盖公司数据采用 wind 一致预期）

推荐

维持评级



分析师 方竞

执业证书：S0100521120004

邮箱：fangjing@mszq.com

分析师 宋晓东

执业证书：S0100523110001

邮箱：songxiaodong@mszq.com

相关研究

- 1.电子行业点评：HBM 需求强劲，国产替代势在必行-2025/06/30
- 2.电子行业点评：谷歌端侧大模型迭代，泰凌微借势高增乘红利-2025/06/30
- 3.电子行业 2025 年中期投资策略：AI 投资的新范式-2025/06/30
- 4.电子行业动态：影石引领全景相机，AI 眼镜新品升温-2025/06/17
- 5.电子行业深度：智能眼镜：AI 形态领航，探索 AR 新视界-2025/06/11

目录

1 Oracle 大单背后的算力风向：英伟达 GPU 需求超预期	3
1.1 Oracle 签订 300 亿美元算力订单，GPU 需求旺盛	3
1.2 全球 AI 算力格局：GPGPU 与 ASIC 双轨并进.....	3
2 NV GPGPU 需求新动能：三大应用场景驱动增长.....	9
2.1 第三方 LLM 训练需求旺盛.....	9
2.2 主权 AI 算力加速布局	9
2.3 To B 客户与定制化私有云需求	10
3 投资建议	12
4 风险提示	14
插图目录	15
表格目录	15

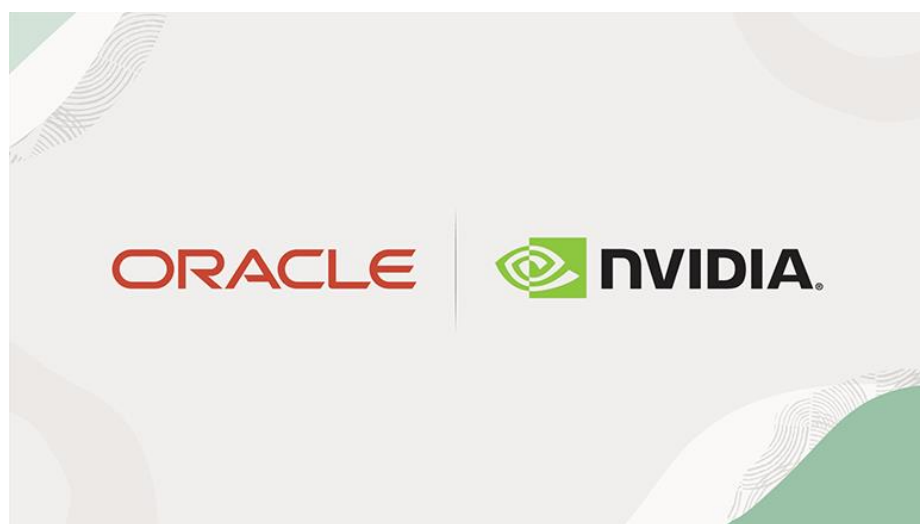
1 Oracle 大单背后的算力风向：英伟达 GPU 需求超预期

1.1 Oracle 签订 300 亿美元算力订单，GPU 需求旺盛

2025 年 7 月 1 日，据新浪财经报道，Oracle 公开表示已签署多项大型云服务协议，其中一项关键订单预计自 2028 财年起，每年将为公司带来超过 300 亿美元的营收。这一消息引发市场广泛关注，因为该订单规模不仅创下 Oracle 历史纪录，也对公司未来收入结构产生深远影响。按照公司披露数据，该项目的年收入贡献将达到 Oracle 2025 财年总收入的约 52%，相当于当前 Oracle 云基础设施业务（OCI）收入的近三倍，充分体现出 AI 大模型及相关云服务市场需求的高速增长。

为了支撑如此大体量的算力需求，Oracle 在硬件资源上的投入同样令人瞩目。公司已采购约 40 万张英伟达 GB200 系列算力卡，这使得 Oracle 成为全球第二大英伟达高端算力卡持有者。这批算力卡为 Oracle 在 AI 基础设施市场上构建了坚实壁垒，确保公司能够在未来数年持续满足第三方大模型训练、主权 AI 和大型企业定制化私有云等多元化需求。与传统云厂商不同，Oracle 的 OCI 业务更偏向算力租赁，通过全面携手英伟达产品，公司在供应链紧张的背景下获得了更优先的资源保障，也进一步强化了其在 AI 云市场的独特竞争力。

图1：Oracle 携手 NVIDIA 助力企业和开发者加速 AI 创新



资料来源：Oracle 官网，民生证券研究院

1.2 全球 AI 算力格局：GPGPU 与 ASIC 双轨并进

在 AI 算力需求持续高速增长背景下，全球算力市场逐步分化为两大主流技

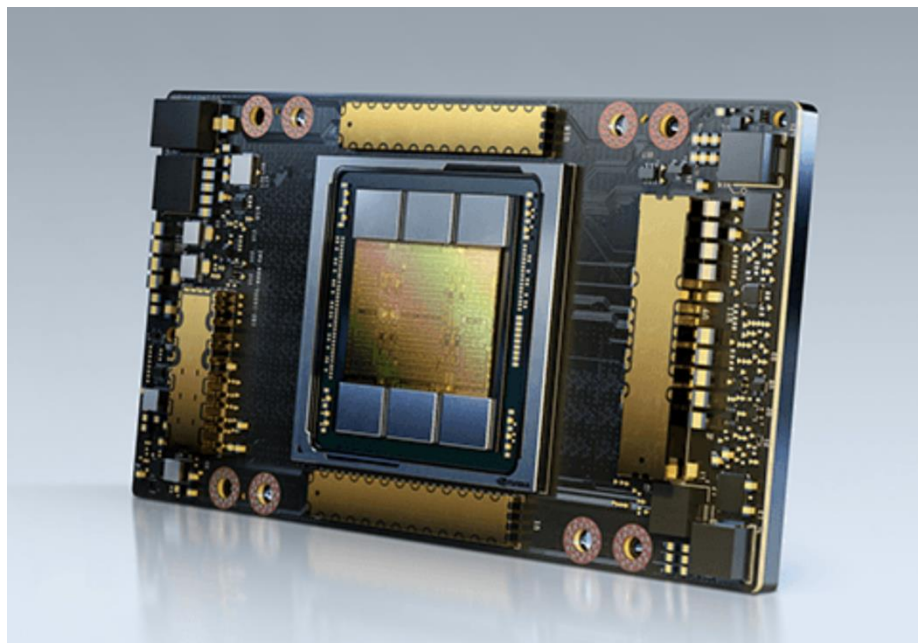
术阵营：以英伟达 GPGPU 为核心的通用计算阵营，以及由 ASIC 驱动的专用计算阵营。两类方案各有侧重，既体现了技术路线和应用场景的不同选择，也共同构建起支撑全球 AI 基础设施发展的算力基石。

1.2.1 NV GPGPU 阵营：适配大模型训练与通用 AI 需求

GPGPU(General-Purpose GPU)指通用计算图形处理器。与传统 CPU 相比，GPGPU 由于其大规模并行架构和高内存带宽，特别适合处理计算密集型任务。在 AI 时代，GPGPU 已成为相当重要的协处理器解决方案，在大模型训练，通用场景，主权 AI 等方面发挥作用。

在大模型训练领域，训练大型语言模型 (LLM) 需要处理海量参数和巨量训练数据，这对计算硬件提出了极高要求，GPGPU 在此场中凭借其高能效比、低延迟等技术特点展现出不可替代的价值。在通用计算场景中，GPGPU 被应用于生命科学，石油勘探，气象预测等领域。在主权 AI 场景中，是否能提供本地化部署、安全合规、定制化的服务成为云基础的重点，高性能的 GPGPU 具有构建该设施的显著优势，更受到大型 to B 客户的青睐。

图2：NVIDIA A100 Tensor 核心 GPU



资料来源：NVIDIA，民生证券研究院

该阵营的主要云厂商是 CoreWeave 和 Oracle。

CoreWeave：公司的核心业务是为 AI 工作负载提供高性能 GPU，并提供 AI 解决方案。该公司拥有 32 个数据中心，运行超过 250000 个 GPU。公司是 NVIDIA 的早期采用者，其在 2019 年起专注于购买企业级 GPU 芯片组，整体围绕 NVIDIA 的芯片调整业务。

CoreWeave 的商业模式是通过从 NVIDIA 租用 GPU 并转租给客户，在此过

程中承担芯片供应商与需求方之间的桥梁角色。作为英伟达首家计算精英云解决方案提供商，CoreWeave 可访问云中的十多种 NVIDIA GPU 型号，涵盖 H100、A100、A40 和 RTX A6000 等。此外，CoreWeave 还部署了全球最大的 NVIDIA HGX H100 GPU 集群，每个 GPU 配备 40GB 的 HBM2 内存。

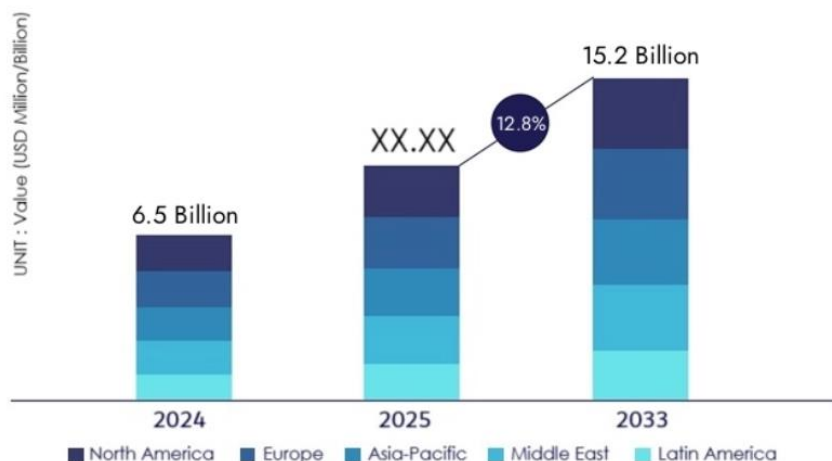
Oracle: 公司因开发首个商用 SQL 关系数据库管理系统 Oracle Database 而闻名，经过近 50 年发展，Oracle 从单一数据库提供商成长为全球领先的企业级软件和云服务供应商，提供包括数据库、中间件、应用软件和云基础设施在内的技术栈。

公司的核心优势在于其企业级数据库和云服务，其旗舰产品 Oracle Database 以高性能、高可用性和强大安全性著称，支持从传统 OLTP 到大数据分析的各类工作负载。公司与 NVIDIA 密切合作，将 NVIDIA 的全栈 AI 计算平台与其高性能、安全性以及部署灵活性的云基础设施 (OCI) 相结合，从而在全球范围内推进 AI 工作。同时，公司是首批与 NVIDIA DGX Cloud Lepton 集成的超大规模企业之一，可帮助开发人员访问高级 GPU 资源。

1.2.2 ASIC 阵营：推理加速与成本优化驱动新机遇

近年来，全球 ASIC 市场空间持续扩张，成为 AI 算力体系重要组成部分。根据 Verified Market Reports 数据，2024 年全球 AI ASIC 市场规模已达到约 65 亿美元，主要需求来自北美、欧洲和亚太地区，涵盖 AI 推理、搜索推荐、广告算法等高效算力场景。在 AI 大模型训练与推理任务日益复杂化的背景下，ASIC 以其高效、低功耗、定制化优势在全球算力市场中的战略地位持续上升。预计从 2025 年至 2033 年，该市场将保持约 12.8% 的年均复合增长率，到 2033 年市场规模有望增至 152 亿美元。

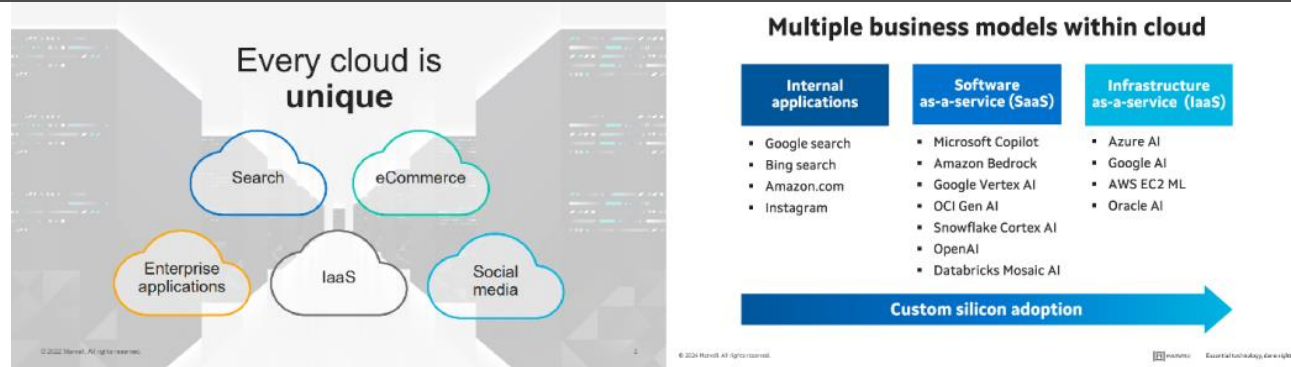
图3：2024-2033 年全球 ASIC 市场空间及增速



资料来源：Verified Market Reports，民生证券研究院

与以 GPGPU 为核心的通用计算阵营不同，ASIC 阵营在 AI 算力市场中主要聚焦于低精度计算场景提升以及算力成本的持续优化。当前，以谷歌、亚马逊、微软等大型云服务商（CSP）为代表的企业，已大规模自研或采购 ASIC 方案，用于满足 AI 推理、大规模搜索推荐等场景的高效算力需求。这些厂商在训练阶段可能仍然依赖 NV GPGPU 提供高通用性算力，但同时也在加速自研 ASIC 芯片进度，争取压低运营成本与单位算力成本、优化功耗比。

图4：ASIC 需要满足不同业务/应用的加速计算需求



资料来源：Marvell，民生证券研究院

表1：云厂商自研 ASIC 加速卡进展

厂商	大类	型号	发布时间	制程 nm	峰值算力 TOPS/TFLOPS			内存信息		互联带宽 GB/s
					INT8/FP8 Dense/Sparse	BF16/FP16 Dense/Sparse	TF32/FP32 Dense/Sparse	类型	容量 GB	
谷歌	训练	TPUv5E	2023	-	394	197	-	HBM2	16	400
		TPUv5P	2023	-	918	459	-	HBM2	95	800
Meta	推理	MTIA v2	2024	5	354/708	177/354	2.76	LPDDR5	128	-
微软	训练	Maia 100	2023	5	1600	800	-	HBM3	64	1200
亚马逊	训练	Trainium2	2023	4	861	431	215	-	96	-
	推理	Graviton4	2023	-	-	-	-	DDR5	-	-
TESLA	训练	D1	2021	7	362	362	22.6	-	32	-

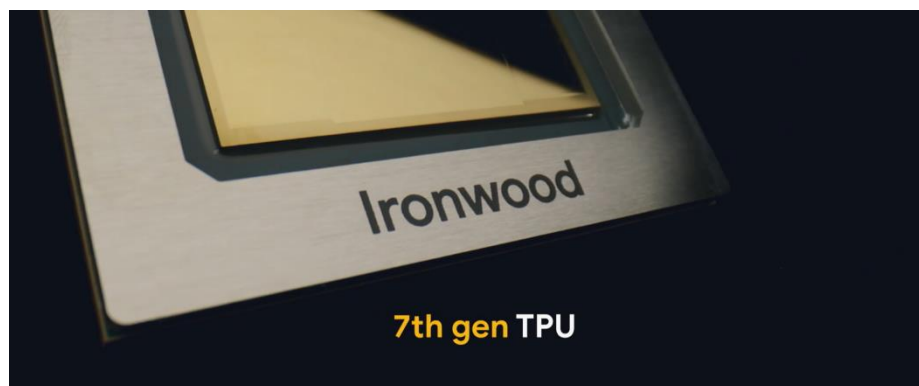
资料来源：各公司官网，Semianalysis，VerifiedMarket 等，民生证券研究院整理

注：未标注的数据为没有在公开渠道披露的信息

在 AI 算力加速需求持续攀升的背景下，定制化 ASIC 方案逐渐成为云厂商和 AI 巨头优化推理性能、控制成本的重要战略方向。谷歌自 2016 年推出 TPU 以来，已迭代至第七代 Ironwood——专为大规模思考、推理 AI 模型设计，也是谷歌迄今为止最节能的 TPU。凭借其强大的计算能力，TPU 在性能、能效比和成本方面为 AI 计算带来了新的拓展，成为谷歌内部 AI 任务以及外部云客户的重要算力选

择。

图5：谷歌第七代 Tensor Processing Unit (TPU) Ironwood



资料来源：Google 官网，民生证券研究院

亚马逊 AWS 则通过自研 Inferentia、Trainium 等芯片，强化其 AI 基础设施能力。Inferentia 可通过芯片之间的超高速连接支持横向扩展分布式推理，帮助客户以最低的成本获得深度学习和生成式人工智能推理所需的高性能。Trainium 是 AWS 专门为人工智能训练和推理打造的人工智能芯片系列，试图打破英伟达芯片在训练领域的垄断局面，尽管目前仍以补充角色为主，但亚马逊在 AI 芯片上的持续投入已逐步构筑差异化竞争力。

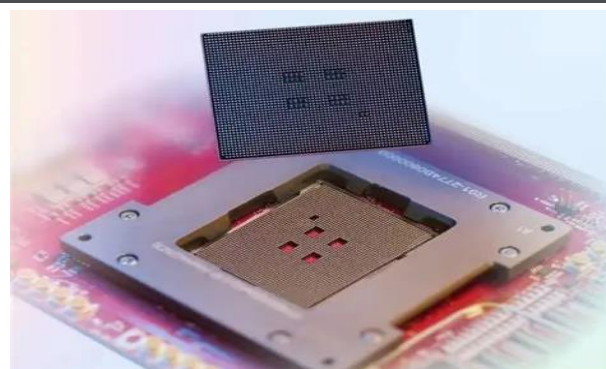
Meta 在 ASIC 领域的布局主要聚焦 AI 推理加速器，其自研的 MTIA (Meta Training and Inference Accelerator) 项目正在测试阶段，未来有望支撑其庞大的推荐系统和生成式 AI 业务。OpenAI 虽然尚未大规模部署自研 ASIC，但该公司首款自主研发的 AI 芯片即将完成流片，并预计在 2026 年实现量产，这也表明 OpenAI 正致力于构建一个涵盖芯片、云服务和 AI 应用的完整技术生态体系。

图6：亚马逊 AWS 发布新一代 AI 芯片 Trainium3



资料来源：ICsmart，民生证券研究院

图7：Meta 下一代 MTIA 芯片模型



资料来源：Meta，民生证券研究院

在 ASIC 产业链的供给侧，博通、Marvell 等主流芯片公司普遍对该市场持积极乐观态度，二者占 ASIC 市场超 60% 的份额。博通一直是谷歌自研 AI 芯片 TPU 的主要制造商，并将在谷歌第七代 TPU 处理器 (TPU v7) 中继续合作；Marvell 与亚马逊、谷歌和微软有多年的 ASIC 合作经验，随着台积电 CoWoS 产能的大幅

提升以及亚马逊扩大产量的承诺, Marvell 对 ASIC 市场保持乐观态度。与此同时, 由于 ASIC 方案对工艺、封装、测试等环节提出更高要求, 也带动了相关产业链的需求增长。

总体来看, 随着推理应用规模的不断扩大, 以及全球云厂商对 AI 运营成本控制需求的提升, ASIC 阵营在 AI 算力市场中的战略地位正加速抬升, 逐步形成多元化格局, 与 GPGPU 共同支撑全球 AI 算力基础设施的发展。

2 NV GPGPU 需求新动能：三大应用场景驱动增长

NVIDIA 的 GPGPU（通用图形处理器）需求持续强劲，主要来自三大方向：第三方训练、主权 AI 建设以及企业级客户的定制化私有云部署。

2.1 第三方 LLM 训练需求旺盛

在第三方训练领域，随着 AI 模型参数量的增长（如 GPT-3 的 1750 亿参数），训练所需的算力也随之上升。**NVIDIA 在 GPGPU 上的技术优势明显。**例如，训练 GPT-3 需要约 $3.14E23$ FLOPS 的计算量，而英伟达 H100 GPU 集群（3,584 块）仅需 11 分钟即可完成训练，其 NVLink 与 InfiniBand 网络能够支持 GPGPU 间高效通信，实现超线性的扩展性能。

随着大模型参数规模向万亿级迈进，云服务商和 AI 公司对英伟达 GPU 的依赖度持续攀升，NVIDIA 最新的 BlackWell 架构是首个专为人工智能从训练到大规模推理的每个阶段而设计的架构，它为 GPT-4 或 Grok-3 等万亿参数模型提供高达 30 倍的推理速度，而且每瓦、每个机架和每美元投资的总成本最低。

在第三方大模型训练领域，人工智能巨头 OpenAI 近期丰富了其算力策略，在 2024 年 6 月与 Oracle 达成了算力租赁方面的合作。通过此次合作，OpenAI 旨在获取更多计算能力，从而能够在 Oracle 的基础设施上使用微软 Azure AI 平台。而 OpenAI 与微软 Azure 的合作最为深入，微软是其核心算力来源。早在 2019 年，微软就向 OpenAI 投资了 10 亿美元，并成为了其主要的云合作伙伴。微软不仅提供资金支持，同时也在 Azure 云平台上为 OpenAI 定制了大规模的超级计算机。该台超级计算机拥有超过 28.5 万个 CPU 核心和 1 万个 GPU，每个 GPU 服务器都配备了每秒 400 千兆比特的网络连接。这种专属且庞大的算力基础设施是 OpenAI 训练其大型语言模型（如 GPT 系列）的关键。

2.2 主权 AI 算力加速布局

主权 AI 指各国政府为保障数据安全、技术自主而构建的本土化 AI 基础设施，数据主权和技术自主成为各国建立主权 AI 的核心驱动力。随着大模型技术的发展，各国寻求更安全更稳定的手段监管和管理本国 AI，构建“主权云”成为各国的重要选择。**NVIDIA 的 GPGPU 芯片提供了卓越的计算能力，这些芯片不仅被用于构建国家级 AI 超级计算机，也支持了各地的 AI 研发中心和数据主权云平台。**

在以沙特阿拉伯为代表的中东地区，各国正在大力投资主权 AI 建设，向实现经济多元化和成为全球 AI 领导者的目标而努力。他们积极采购 NVIDIA 的 GPGPU 芯片，用于建设大规模的国家级 AI 超级计算机和数据中心。2025 年，NVIDIA 与沙特主权财富基金 PIF 新成立的 AI 公司 Humain 达成芯片供应协议，为其提供 1.8 万颗 GB300 芯片，支撑高达 500 兆瓦的数据中心建设，同时协作部署

Omniverse Cloud 数字孪生平台，共同培训开发者数千名。此类投资旨在本地化 AI 数据处理，加速中东地区创新转型。

在欧洲，许多国家正积极建设由 NVIDIA 芯片驱动的超级计算中心，以支持科学研究、工业 AI 应用以及开发符合欧洲价值观的 AI 模型。为加速国家资源开发，NVIDIA 正与法国、英国和德国等各国企业合作。在法国，Mistral AI 正在和 NVIDIA 合作，现阶段的目标为布设 NVIDIA 的 Blackwell 系统支持的云平台，为下一步加速采用 AI 代理应用作铺垫。在英国，NVIDIA 与 Nebius 和 Nscale 合作，力图使全英国能够使用可扩展且数据安全的 AI 基础设施。在德国，NVIDIA 及其合作伙伴正在为欧洲制造商打造全球首个工业 AI 云，让欧洲的行业领导者能够加速各种制造应用。

在东南亚地区中，新加坡计划投资超过 10 亿新元发展 AI 战略，在此战略中，NVIDIA 与新加坡通信服务提供商 Singtel 达成合作，使用 NVIDIA Hopper 架构 GPU 和高性能 NVIDIA AI 参考架构，将 NVIDIA AI 平台引入岛屿国家和其他地区的企业，为东南亚各地构筑能效高的数据中心提供助力。

中国在 AI 发展方面取得了显著成就，并高度重视 AI 的自主可控性。虽然直接获取 NVIDIA 最先进的 GPGPU 芯片受到限制，但 NVIDIA 仍通过符合出口管制的芯片型号（如 H20）或早期型号在中国市场发挥作用。同时，中国本土企业也在大力研发国产 AI 芯片，以实现算力自给自足。

表2：主权 AI 建设地区情况

地区	英伟达贡献
中东地区	-沙特阿拉伯：提供 18000 颗 GB300 芯片用于 500MW 超算中心，部署 Omniverse Cloud 数字孪生平台，培训开发者 -阿联酋：与 G42 旗下的数据中心开发商 Khazna 合作建立数据中心，由 NVIDIA Blackwell 芯片提供支持 -卡塔尔：Ooredoo 推出基于 NVIDIA Hopper GPU 的主权 AI 云服务，提供 NVIDIA AI Enterprise 平台
欧洲地区	-法国：与 Mistral AI 合作部署 Blackwell 系统云平台，支持 AI 代理应用开发。 -英国：与 Nebius、Nscale 合作部署 14,000 个 NVIDIA Blackwell GPU 驱动新的数据中心 -德国：打造首个工业 AI 云，由 NVIDIA DGX B200 系统和搭载 10000 个 NVIDIA Blackwell GPU 的 NVIDIA RTX PRO 服务器提供支持 -意大利：与 Domyn 合作，采用 NVIDIA Grace Blackwell 平台提升主权 AI 能力。
东南亚地区	-新加坡：与 Singtel 合作部署 Hopper 架构 GPU，构建高能效数据中心 -印度尼西亚：为 Sahabat-AI 大语言模型提供由 NVIDIA 加速的主权 AI 云服务 GPU Merdeka by Lintasarta，提供 NVIDIA NeMo 平台 -马来西亚：合作开发 500MW 数据中心，部署 NVIDIA H100 Tensor Core GPU
中国	提供符合出口管制的 H20 芯片（已受限），支持有限市场需求，本土企业加速国产替代

资料来源：NVIDIA，Developing Telecoms，新浪财经等，民生证券研究院整理

2.3 To B 客户与定制化私有云需求

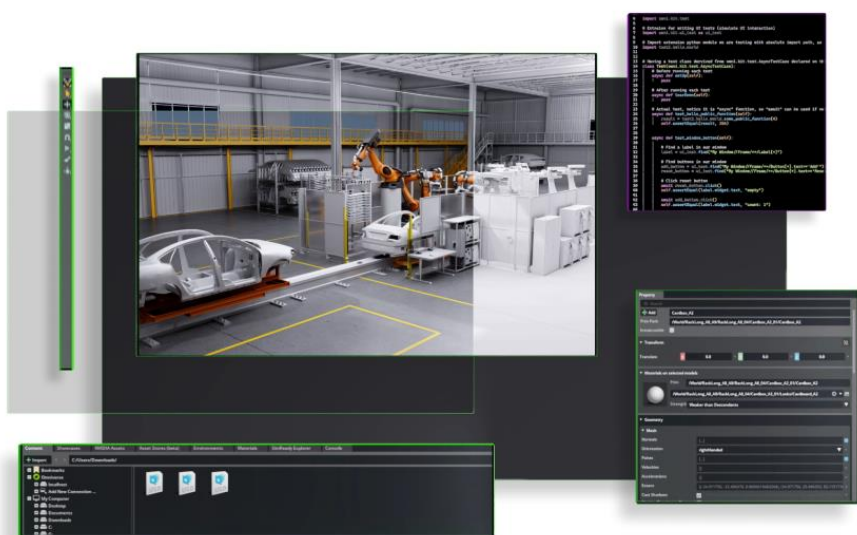
在 To B 市场，企业级客户更倾向于私有化部署和混合云方案。为了跟上行业的发展步伐，组织部署的产品、技术和工具日益多样化，NVIDIA 采用混合云解决方案扩展公有云和私有云的优势。企业可通过混合云动态扩展训练资源，在需求高峰时调用云端 GPU，平时使用本地集群。当需要私有云进行定制化大模型微调时，

NVIDIA 的 SageMaker HyperPod 灵活的基础架构由于其支持大规模模型训练和推理，成为有类似需求的企业理想选择。在定制化解决方案方面，NVIDIA 也具有相当优势。

生物技术巨头安进 (Amgen) 正在与 NVIDIA 合作，利用 NVIDIA BioNeMo 平台进行蛋白质 LLM 的训练。通过租赁云端 GPU 算力，Amgen 的研究团队结合 BioNeMo 平台，借助 DGX Cloud 强大的算力和多节点能力，将蛋白质大模型训练速度提升 3 倍，训后分析速度提升 100 倍，使药物发现和研发流程获得了相当大的效率提高。

宝马集团与 NVIDIA 合作，通过 NVIDIA Omniverse 构建用于优化全球生产的虚拟工厂。该团队使用 Omniverse 的 OpenUSD 框架整合建筑、设备和人体工程学数据，使德布勒森电动汽车工厂在虚拟环境中完成全流程仿真验证，预计节省 30% 的改造成本。

图8：NVIDIA Omniverse 平台



资料来源：NVIDIA 官网，民生证券研究院

3 投资建议

在全球AI产业快速发展、大模型训练和主权算力建设加速推进的背景下，英伟达GPGPU需求高位运行的趋势大概率将持续较长时间。Oracle斩获的300亿美元大单、全球主权云和第三方AI训练中心的扩建，充分验证了通用AI算力的长期成长空间。与此同时，各国对AI自主可控的高度重视，也带动了高性能算力芯片、服务器、通信连接和电源管理等上下游产业链的需求扩张。

我们认为，算力基础设施建设已进入“高景气、多元化”的新阶段。短期看，NV GPGPU订单驱动下相关产业链景气度持续超预期；中长期看，ASIC、GPGPU、服务器、光通信等方向将形成协同增长格局，带来全产业链投资机会。在此背景下，建议投资者重点关注拥有核心技术、产能扩张能力强、在AI算力产业链中具备明显竞争优势的公司。

我们认为，算力的长期需求，不仅在于云厂商的Token数增长，ROI闭环等，主权AI、大模型训练加速以及B端客户私有云的部署均拉动了算力的强劲需求。**我们继续看好ASIC+NV链的投资逻辑**，后续重点关注优质龙头的中报业绩表现，以及产品/客户突破，建议关注：**ASIC：芯原股份；服务器：工业富联、华勤技术；PCB：沪电股份、胜宏科技、广合科技、生益科技、威尔高；光通信：旭创科技、新易盛、东山精密、长芯博创、太辰光；铜连接：瑞可达、沃尔核材；电源：江海股份、禾望电气。**

表3：电子行业重点关注个股

证券代码	证券简称	股价 (元)	EPS			PE			评级
			2025E	2026E	2027E	2025E	2026E	2027E	
688521.SH	芯原股份	91.08	-0.18	0.07	0.61	/	1301	149	推荐
601138.SH	工业富联	23.98	1.52	1.90	2.27	16	13	11	推荐
603296.SH	华勤技术	83.17	3.62	4.52	5.45	23	18	15	推荐
002463.SZ	沪电股份	45.80	1.89	2.32	2.71	24	20	17	推荐
002384.SZ	东山精密	40.46	1.57	2.13	2.70	26	19	15	推荐
300548.SZ	长芯博创	61.47	1.38	1.77	2.24	45	35	27	推荐
688800.SH	瑞可达	45.38	1.78	2.32	2.89	25	20	16	推荐
300476.SZ	胜宏科技	144.23	5.40	7.38	8.99	27	20	16	/

资料来源：Wind，民生证券研究院预测；（注：股价为2025年7月7日收盘价；未覆盖公司数据采用wind一致预期）

工业富联：工业富联作为全球AI服务器OEM核心厂商，在AI服务器市场份额达到四成左右，主要客户涵盖微软、亚马逊、谷歌等核心CSP厂商，有望充分受益AI服务器放量。前期市场担忧GB200出货延后，而目前GB200良率问题已经解决，机柜服务器进入大规模出货阶段，工业富联作为英伟达GB200机柜最重要的供应商之一，在4Q24已经率先完成GB200机柜的量产交付，预计今年GB200将给公司带来较大业绩增量。此外，2025年4月工业富联公告回购5-10

亿元公司股票，回购价格上限为 20 元，截止 6 月底公司累计回购 1.47 亿元，仍有 3.53-8.53 亿元回购额度，预计后续公司仍将继续回购，彰显长期增长信心。

华勤技术：华勤技术为国内 AI 服务器核心制造商，1Q25 数据中心业务实现超 100 亿元营收，占公司营收比重近 3 成，在头部互联网客户实现全栈式发货和高速增长；其他主业方面，1Q25 智能手机和智能穿戴收入均实现了超过翻倍的增长，同时 PC 业务保持稳健的成长趋势；此外公司汽车业务前期定点项目开始落地，进入高速成长期，未来将实现多增长极协同发展。我们预计公司 2H25 业绩环比较为平稳，公司预计 25 年收入及利润均保持 25%同比增长；当前英伟达特供版 B30 芯片已在客户端开启测试，B40 芯片已进入量产冲刺阶段，用于替代之前受到出口限制的 H20 芯片，华勤作为国内 AI 服务器龙头有望显著受益。

4 风险提示

1) 算力需求不及预期的风险：尽管当前全球 AI 热潮带动了高端算力需求的快速增长，但未来市场对 GPGPU 和相关算力的需求仍存在不确定性。大模型训练和推理的技术路径尚在快速演进，算法优化、模型压缩、推理加速等技术进步可能降低对高算力硬件的需求强度；同时，下游客户的投入节奏、资金实力及商业化进展亦可能受宏观经济波动、资本市场环境等因素影响，导致算力采购需求阶段性放缓。

2) AI 应用落地不及预期的风险：目前 AI 大模型和各类行业专属 AI 解决方案的商业化路径和落地速度仍存在较大不确定性。AI 应用的技术成熟度、数据安全、隐私保护、伦理合规等方面尚需克服多重挑战，这可能导致部分行业客户持观望态度，延缓大规模部署；同时，商业模式探索进展缓慢或回报周期过长，也可能影响终端客户对高性能算力的持续采购意愿。

3) 产业链供应受限的风险：当前全球半导体产能紧张、大宗原材料价格波动、关键零部件交付周期延长等因素，均可能对算力设备生产及交付形成制约。此外，产业链中部分高精尖环节集中在少数企业，若相关厂商因产能扩张不及预期、工艺瓶颈或突发事件导致供应中断，也将对高端算力设备的交付能力和市场供应节奏带来不利影响。

4) 全球贸易政策变动的风险：虽然 AI 产业链参与主体遍布全球，市场需求于美洲、欧洲、亚洲等地广泛分布，但国际贸易环境的变化、主要经济体对高技术产品的出口管制、关税政策调整、贸易壁垒增加等因素，仍可能在局部或阶段性影响算力设备的供需格局。

插图目录

图 1: Oracle 携手 NVIDIA 助力企业和开发者加速 AI 创新	3
图 2: NVIDIA A100 Tensor 核心 GPU	4
图 3: 2024-2033 年全球 ASIC 市场空间及增速	5
图 4: ASIC 需要满足不同业务/应用的加速计算需求	6
图 5: 谷歌第七代 Tensor Processing Unit (TPU) Ironwood	7
图 6: 亚马逊 AWS 发布新一代 AI 芯片 Trainium3	7
图 7: Meta 下一代 MTIA 芯片模型	7
图 8: NVIDIA Omniverse 平台	11

表格目录

重点公司盈利预测、估值与评级	1
表 1: 云厂商自研 ASIC 加速卡进展	6
表 2: 主权 AI 建设地区情况	10
表 3: 电子行业重点关注个股	12

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登记为注册分析师，基于认真审慎的工作态度、专业严谨的研究方法与分析逻辑得出研究结论，独立、客观地出具本报告，并对本报告的内容和观点负责。本报告清晰准确地反映了研究人员的研究观点，结论不受任何第三方的授意、影响，研究人员不曾因、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。

评级说明

投资建议评级标准	评级	说明
以报告发布日后的 12 个月内公司股价（或行业指数）相对同期基准指数的涨跌幅为基准。其中：A 股以沪深 300 指数为基准；新三板以三板成指或三板做市指数为基准；港股以恒生指数为基准；美股以纳斯达克综合指数或标普 500 指数为基准。	推荐	相对基准指数涨幅 15%以上
	谨慎推荐	相对基准指数涨幅 5% ~ 15%之间
	中性	相对基准指数涨幅-5% ~ 5%之间
	回避	相对基准指数跌幅 5%以上
公司评级	推荐	相对基准指数涨幅 5%以上
	中性	相对基准指数涨幅-5% ~ 5%之间
	回避	相对基准指数跌幅 5%以上
行业评级	推荐	相对基准指数涨幅 5%以上
	中性	相对基准指数涨幅-5% ~ 5%之间
	回避	相对基准指数跌幅 5%以上

免责声明

民生证券股份有限公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。

本报告仅供本公司境内客户使用。本公司不会因接收人收到本报告而视其为客户。本报告仅为参考之用，并不构成对客户的投资建议，不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告所包含的观点及建议并未考虑获取本报告的机构及个人的具体投资目的、财务状况、特殊状况、目标或需要，客户应当充分考虑自身特定状况，进行独立评估，并应同时考量自身的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见，不应单纯依靠本报告所载的内容而取代自身的独立判断。在任何情况下，本公司不对任何人因使用本报告中的任何内容而导致的任何可能的损失负任何责任。

本报告是基于已公开信息撰写，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、意见及预测仅反映本公司于发布本报告当日的判断，且预测方法及结果存在一定程度局限性。在不同时期，本公司可发出与本报告所刊载的意见、预测不一致的报告，但本公司没有义务和责任及时更新本报告所涉及的内容并通知客户。

在法律允许的情况下，本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易，也可能为这些公司提供或正在争取提供投资银行、财务顾问、咨询服务等相关服务，本公司的员工可能担任本报告所提及的公司的董事。客户应充分考虑可能存在的利益冲突，勿将本报告作为投资决策的唯一参考依据。

若本公司以外的金融机构发送本报告，则由该金融机构独自为此发送行为负责。该机构的客户应联系该机构以交易本报告提及的证券或要求获悉更详细的信息。本报告不构成本公司向发送本报告金融机构之客户提供的投资建议。本公司不会因任何机构或个人从其他机构获得本报告而将其视为本公司客户。

本报告的版权仅归本公司所有，未经书面许可，任何机构或个人不得以任何形式、任何目的进行翻版、转载、发表、篡改或引用。所有在本报告中使用的商标、服务标识及标记，除非另有说明，均为本公司的商标、服务标识及标记。本公司版权所有并保留一切权利。

民生证券研究院：

上海：上海市虹口区杨树浦路 188 号星立方大厦 7 层； 200082

北京：北京市东城区建国门内大街 28 号民生金融中心 A 座 18 层； 100005

深圳：深圳市福田区中心四路 1 号嘉里建设广场 1 座 10 层 01 室； 518048