

英伟达重金收编潜在挑战者

当地时间周三(12月24日),被视作英伟达“挑战者”的Groq在官网宣布,与英伟达达成一项“非独家授权协议”。英伟达支付约200亿美元现金,获得Groq的核心AI推理技术知识产权和相关资产,而Groq的云服务业务(Groq Cloud)将继续独立运营。此外,Groq创始人兼CEO Jonathan Ross、总裁Sunny Madra等核心高管及团队将加入英伟达。



非典型性收购

“这不是一次对公司整体的收购。”Jonathan Ross表示。

据悉,Groq是专攻AI推理芯片的明星初创公司,成立于2016年,总部位于美国加利福尼亚,创始人Jonathan Ross曾是谷歌自研AI芯片TPU(张量处理单元)项目的核心研发成员,深度参与了专为AI优化的芯片设计。这个项目后来被用于击败围棋冠军李世石的AlphaGo比赛,也是Google AI服务的关键硬件。2016年,他带领谷歌TPU团队10名核心成员中的7位一同离职,创立了Groq。当时他发现,传统计算架构(如CPU/GPU)无法高效处理现代AI任务,这一认知促使他决定创办一家突破传统限制的公司。

2024年2月,Groq推出了一款全新的AI芯片,声称实现了“地表最强推理”——在

Groq上运行大模型的推理速度较英伟达GPU提高10倍甚至更高。

2025年11月,美国白宫和美国能源部的最新声明显示,24家顶尖的人工智能企业已与美国政府签署协议,加入“创世纪计划”,英伟达和Groq都位列其中。

目前,Groq已与Meta合作为其Llama API提供推理加速;与IBM合作整合其AI推理平台;与沙特阿美签署巨额协议,计划建设大型AI推理数据中心。

英伟达CEO黄仁勋在发给员工的内部邮件中明确了此次交易的战略意图。他表示,计划将Groq的低延迟处理器整合进英伟达的AI工厂架构中,从而扩展平台能力,以服务更广泛的AI推理和实时工作负载。这意味着英伟达正试图在其昂贵的高性能训练芯片之外,补齐在推理芯片领域的高效能短板。

虽然此次与Groq的200亿美元交易并非全资收购,但其资金规模已远超以往,显示出

英伟达在面对日益专业化的芯片需求时,愿意付出高昂代价以消除潜在威胁并整合前沿技术。

值得注意的是,这笔交易的架构与微软、亚马逊和谷歌过去两年采用的模式如出一辙,即通过“许可技术+聘用人才”的方式,在不正式收购公司的情况下通过监管雷达,这被认为是科技巨头争夺顶级AI人才与技术的典型方式,能以绕过复杂反垄断审查的形式,快速获取关键创新。

中国人工智能学会天使投资人、资深人工智能专家郭涛表示,近两年,大型科技公司频繁采用“准收购”,也就是许可技术加聘用人才的合作模式,核心目的是规避日趋严格的反垄断审查,高效获取前沿技术与稀缺人才,无需承担全资收购带来的监管风险与高昂成本。这种模式允许大公司在不正式并购标的企业的前提下,吸纳其创新技术与核心团队,快速补强自身技术实力、消除潜在竞争威胁;同时保

障被投企业的独立性,以更灵活的姿态应对瞬息万变的市场需求与技术迭代节奏。

双方补齐短板

本次交易的核心驱动力在于英伟达对AI推理市场的争夺。

尽管英伟达的GPU及配套软件在AI模型开发、训练方面占据绝对主导地位,但对于运行聊天机器人等实际应用(推理)而言,其现有芯片往往体积过大且成本高昂。市场一直在寻找更廉价、更高效的替代方案,而Groq的技术正是为此而生。

而Groq的核心产品是LPU(语言处理单元),这类芯片主要用于加快大语言模型完成推理相关任务的速度,被外界视为英伟达GPU替代方案之一。Groq声称其芯片在涉及AI应用的特定任务中,数据处理速度优于英伟达。相比之下,英伟达的芯片虽然能在处理多种类型操作时保持灵活性,但在处理速度和延迟方面存在优化空间。

通过此次“非独家授权协议”,英伟达获得了Groq的知识产权,因此获得更具成本效益且速度更快的芯片设计能力,以应对日益增长的AI应用运行需求。

芯片咨询公司SemiAnalysis的首席分析师Dylan Patel指出,虽然Groq的第一代芯片尚未能与英伟达形成有力竞争,但其后续两代产品即将推出。他认为,英伟达可能在Groq的新一代技术中看到了威胁,因此选择出手。

截至10月底,英伟达持有的现金及短期投资规模为606亿美元,高于2023年初的133亿美元。随着现金储备不断增加,英伟达近年来持续加大对芯片初创公司及整个生态系统的投资。

今年9月,英伟达表示计划向OpenAI投资最高1000亿美元。同月,英伟达还表示将向英特尔投资50亿美元。

除此以外,英伟达还投资了AI和能源基础设施企业Crusoe、AI模型开发商Cohere,并在AI云服务商CoreWeave上市之际进一步加大了投资力度。

对于Groq而言,尽管获得了数十亿美元的风险投资,包括Groq在内的挑战者们一直难以打破英伟达对高端AI芯片市场的严密控制。英伟达的芯片在性能上由于其专有的CUDA编程语言生态,使得客户对其产生了极高的黏性。

Groq近期的经营状况也反映了挑战巨头的难度。公司最近将2025年的收入预期下调了约四分之三。Groq发言人曾表示,这是由于计划部署芯片的地区缺乏数据中心容量,导致部分收入预期推迟至明年。Groq曾在7月预计其云业务今年收入将超过4000万美元,总销售额超过5亿美元。

这笔约200亿美元的资金规模,已达到Groq数月前融资时69亿美元估值的约3倍。这可能意味着其独立的硬件挑战者征程告一段落,但其核心技术将在英伟达的生态中获得更广阔的开发平台。

资深产业经济分析师、钉科技创始人丁少将分析称,英伟达此次通过“准收购”模式获得Groq的核心AI推理技术,主要目的在于快速补齐自身在低延迟、高能效推理芯片领域的技术短板,从而巩固其市场主导地位。

对Groq而言,郭涛则认为,该交易带来约200亿美元的充足现金流,极大缓解了公司财务压力,也为现有投资者创造了可观回报。尽管核心团队被纳入英伟达体系, Groq仍保留独立的云服务业务Groq Cloud,且已任命新CEO,持续维持独立运营。此外,借助英伟达的资源优势, Groq的技术有望触达更丰富的应用场景,加速商业化落地进程;同时避免了被全资收购后品牌消失的风险,为未来发展保留了充分的自主性。

北京商报记者 和岳

骗局远离我

毕方:中国传统瑞兽之一。传说上古时期,毕方作为上天的神鸟,毅然衔着火种来到人间,拯救因为没有火种而忍饥挨冻的人类。人类感念其恩德而将其奉为祥瑞。毕方也唤作火鸟,所到之处,播撒希望,去伪存真。

