

微软宣布史上最大自研芯片押注



明年起,微软将以吉瓦级规模部署新一代自研AI芯片Maia 300,主要用于AI推理任务。同时除了微软之外,其他超大规模云计算企业,包括亚马逊、谷歌和Meta,以及OpenAI和Anthropic等也已经推出或正在开发自己的定制芯片。为什么这些竞争关系错综复杂的公司,会不约而同地走向同一条路?答案或许在于它们共同面对的算力困境。而传统的芯片设计巨头如英伟达、AMD,需要重新审视这些昔日的客户正变成潜在的竞争对手。

定档2027年

“今天我们宣布迄今为止在自研芯片上最大的一笔押注。”当地时间10月12日,微软董事长兼首席执行官萨提亚·纳德拉(Satya Nadella)在社交媒体上公布了微软自研芯片的最新计划。微软表示Azure Maia 200现已投入生产,并且发布最新芯片Azure Maia 300。同时,微软公布新的部署计划:从2027年开始将把Azure Maia系列扩展至数吉瓦级规模部署,使Maia成为微软产品和服务的主要推理系统。

从微软公布的进度看,Maia 200目前已投入使用,服务于包括Microsoft Copilot在

内的微软生产工作负载,并为MAI团队开发下一代模型提供支持,GitHub Copilot和Microsoft Foundry也将于年底前投入使用。在这些工作负载中,Maia 200运行着OpenAI模型、微软自研的MAI模型以及领先的开源模型系列。

Maia 300则专为当今涌现的新型AI工作负载而设计,包括长上下文推理、延迟敏感的对话体验以及持续的多步骤智能体任务。微软预计,相比领先的基于GPU的系统,Maia 300每兆瓦可多输出40%的Token(词元)。受这一消息影响,微软股价当天上涨1.14%至541.15美元,市值重回4万亿美元。

近年来,微软加快了自研芯片的进程。2023年11月,微软发布了Maia 100,该芯片

采用台积电5纳米工艺,旨在为OpenAI模型、Bing、GitHub Copilot和ChatGPT等人工智能工作负载运行基于云的训练和推理。

两年后,微软在今年1月宣布推出第二代人工智能芯片Maia 200。当时微软表示Maia 200是公司迄今为止部署的最高效推理系统,其每美元性能比微软当时部署的最新一代硬件提升了30%。

在去年11月的一次访谈中,萨提亚·纳德拉提及了自研芯片的逻辑。他表示,微软计划在自己的MAI模型和芯片之间建立一个闭环,赋予公司自主研发芯片的权利,即根据自己的需求设计微架构,然后不断更新自己的模型。如今距离微软发布Maia 200不到9个月,微软再次发布了Maia 300。

天使投资人、人工智能专家郭涛对此表示,随着AI推理需求持续增长,云厂商正通过自研芯片提高能效、控制成本。若Maia 300达到预期表现,微软有望降低对英伟达GPU的依赖,并改善AI服务的运营经济性。

集体突围

除了微软,为降低对于英伟达的依赖,亚马逊、谷歌、OpenAI等巨头都在加速自研芯片。相比之下,亚马逊和谷歌比微软更早启动自研芯片项目。通过针对自有数据中心的需求定制芯片,微软有望更高效地运行推理任务,并减少采购英伟达GPU所带来的部分成本。

其中OpenAI的动作最为激进。其CEO山姆·阿尔特曼被曝正在全球范围内,为一项代号“Tigris”的芯片项目筹集巨额资金,目标是建设能支撑超级智能的专属芯片制造网络。与此同时,OpenAI已与博通等芯片设计巨头合作,开发定制化的AI推理芯片,并频繁与台积电接触。

不过,不同于谷歌的张量处理单元(TPU)以及亚马逊的Inferentia和Trainium芯片,微软的AI加速器并没有按小时出租给其他公司使用。咨询公司Moor Insights & Strategy的CEO Patrick Moorhead指出,还没有微软以外的企业测试过这些芯片。

此外,Meta近日宣布第三代自研芯片MTIA 450(代号Arke)将于明年上半年开始在数据中心投入使用。第四代产品MTIA 500(代号Astrid)的设计工作将完成,预计2027年底进入数据中心,部署规模将更为广泛。Meta工程副总裁Yee Jiun Song表示,与英伟达目前出货的产品相比,自研芯片在运行AI模型时效率更高,“正是因为我们自己承担了大量工程工作”。

另据外媒披露,Anthropic已与三星电

子展开深度谈判,核心议题是开发一款专为其大模型设计的AI推理芯片。这并非一次简单的商业采购,而是一场涵盖架构定义、制造工艺与高带宽内存集成的全面技术合作。消息一出,业界震动,因为这意味着大模型公司对算力的焦虑,已从软件层面的“提示词优化”“模型蒸馏”,延伸到了物理层面的晶体管电路设计。

深层博弈

大模型公司集体造芯,也正引发芯片制造环节的权力重构。当最纯粹的模型公司也开始亲自下场定义芯片,AI产业的垂直整合已然不可逆转。

比如,Meta与博通合作进行芯片设计,与台积电合作负责制造,整体方向是降低对英伟达行业主导地位的依赖。而Anthropic与三星的牵手,是大模型竞争从参数规模、多模态能力向算力主权延伸的里程碑。

不过在郭涛看来,客户自研芯片并不意味着英伟达业务必然流失。“设计一款能与英伟达在软件生态上匹敌的芯片,投入动辄数亿美元,周期长达两到三年,且充满失败风险。其CEO黄仁勋不久前公开表示,即便竞争对手的芯片免费,也未必比英伟达的综合成本更低。”

数据显示,英伟达第二季度营收翻倍至960亿美元,显示整体AI算力需求足以容纳多方玩家。目前,微软等竞争对手主要聚焦高利润的推理领域,而英伟达在训练领域仍占据绝对主导。

此外,供应链制约依然存在。郭涛指出,开发自有芯片的公司仍依赖英伟达所倚重的有限制造产能、先进封装技术和高带宽内存。这意味着,即便减少了对英伟达产品的依赖,也无法完全消除供应链的限制。

北京商报记者 赵天舒

聚焦 Focus

高房贷利率重创美国楼市

随着政策利率和债券市场收益率变化,居民住房融资成本也受到影响。美国住房金融机构房地美(Freddie Mac)近日公布的数据显示,美国30年期固定房贷平均利率升至7.4%,高于前一周的7.28%,也明显高于一年前同期的6.3%,达到2023年11月以来的最高水平。美媒认为,房贷利率的上升削弱潜在购房者的购买力,让更多潜在购房者选择观望,房地产市场的低迷状态还将持续。

对正在寻找住房的家庭而言,利率上升并非一个抽象的宏观数字,而是每月还款额和可负担房价的直接变化。许多业主在2020年至2021年间锁定了3%至4%左右的低利率贷款。如果此时卖掉原有住房再买新房,就意味着要以7%以上的利率重新借款。即使换房需求仍然存在,不少业主也会因为融资成本大幅增加而选择继续持有现有住房。

光大证券在美国房地产专题研究中也指出,美国住房市场面临供需两弱的局面:高房价与高房贷利率压制购房需求,而存量房贷低利率形成的“锁仓效应”又限制二手房供给。即使美联储降息,房贷利率也未必同步大幅下降,住房市场因此难以迅速恢复活跃。

不过,购房者并非完全没有调整空间。随着固定房贷利率升高,部分借款人开始考虑初始利率较低的浮动利率贷款。美国存量住房抵押贷款往往采用固定利率,美国联邦住房金融局(FHFA)2024年的研究显示,其近5000万笔存量住房抵押贷款,几乎全部采用固定利率。而美国抵押贷款银行家协会(MBA)最新数据显示,截至10月2日当周,浮动利率贷款占其统计

的房贷新申请活动超10%,表明部分购房者正在尝试通过不同贷款产品降低初始还款压力。

当前美国房地产市场也并非所有板块都受到同等程度的冲击。洛杉矶地产经纪人表示,相较于二手房交易,新房市场表现出一定韧性。一些大型开发商会通过补贴贷款利率、承担部分交易费用等方式降低买家的初始购房成本。例如,开发商可以通过贷款利率补贴,帮助符合条件的购房者获得“6字头”的融资利率,从而提高项目吸引力。这也使部分经纪人在当前环境下更愿意向客户推荐新房。

贷款申请数据进一步反映了利率上升对市场的影响。MBA于10月7日公布的周度调查显示,截至10月2日当周,美国抵押贷款申请总量环比下降4.2%;其中,再融资申请指数环比下降8%,同比下降56%,购房申请指数环比下降2%,同比下降15%。在房贷利率升至近三年高位后,再融资的吸引力进一步减弱,购房者也开始推迟决策。

东方金诚研究发展部高级副总监白雪指出,无论是澳洲四大行集体上调房贷利率,还是美国30年期房贷利率逼近7.5%创三年新高,本质上都是全球高利率周期下,货币政策向居民端传导的典型体现,印证了当前发达经济体抗通胀优先的政策取向。

值得注意的是,近期美国长期国债收益率上行,正是房贷利率走高的重要背景之一。10月上旬,10年期美国国债收益率一度升至5.29%,达到2002年以来高位。市场对能源价格上涨、通胀压力以及美国财政赤字和政府债务规模的担忧,均可能影响长期利率定价。

中金公司研究部宏观分析师于文

博指出,美债收益率对房地产市场利率具有较强的影响。往期测算显示,10年期美债收益率变动贡献了30年房贷利率变动的近95%,是决定30年房贷利率波动的主要因素。

他预计,美债收益率上升冲击对抵押贷款的负面影响更为持久,并随着时间推移逐渐放大。

那么,四季度美债有哪些关键变量?白雪认为,核心观察变量集中在货币政策、财政供给、地缘能源和经济韧性四个方面,这几个方面将共同决定长端收益率的高位区间与拐点。

其中,当前市场已基本计入年内再加息的预期,若四季度核心通胀尤其是服务通胀黏性超预期,叠加能源价格的二次传导效应,将推动美联储再度加息,进一步强化“更高更久”的政策预期,从而推升长端收益率。

多家机构分析还共同提示,判断美国房贷利率的后续走势,不能只看美联储是否加息,还需要观察通胀、财政融资需求以及长期债券市场的供需变化。

从目前的预测看,高利率环境短期内可能难以明显缓解。MBA最新预测显示,2026年第四季度30年期房贷平均利率将达到7.3%,并在2027年继续维持在7%以上,2027年美国住房抵押贷款发放总额预计将下降5%。

这意味着,即使购房需求在部分地区逐步恢复,高融资成本仍可能限制交易规模。对美国住房市场而言,真正的转折点不仅在于房贷利率是否从高位回落,也在于回落幅度能否足以改善家庭支付能力,以及房源供给能否跟上需求变化。

北京商报综合报道

· 图片新闻 ·

冰箱升温至50度 法国超市遭黑客攻击



12日,法国拉尼永一家超市内,所有生鲜和冷冻食品区的冰箱均已关闭。视觉中国/图

近日,法国超市遭遇了前所未有的网络攻击,黑客在夜间将数百家门店冰箱与冰柜的温度调高至50摄氏度,导致大量食品变质,部分冰柜甚至起火,损失严重。

当地时间10月12日,巴黎检察官办公室表示,一场“大规模”攻击正影响法国多家大型食品零售商的联网冷藏设备。目前,法国国家网络安全局(ANSSI)已展开调查。

通常情况下,超市冷冻柜会在夜间进行短暂除霜,温度略有升高。黑客却远程指令这些设备持续升温并长时间保持高温。

家乐福表示,其旗下80家超市受到波及;在法国经营约490家门店的蒙普里则称,约5%的门店(约24家)受到影响,其他主要食品零售商也未能幸免。据悉,此次攻击针对的是丹麦工程公司丹佛斯提供的软件,该软件使零售商能够远程管理冷藏设备。

零售商英特玛榭的母公司火枪手集团表示,公司旗下约150家门店的冷藏设备在夜间被加热至50摄氏度,并持续了六个小时。公司发言人称,所有受影响食品均已丢弃,不会对顾客健

康构成风险。另有约100家门店已紧急切断网络连接,以防类似情况再次发生。

目前尚无全行业损失的即时估算数据,但英特玛榭首席执行官蒂埃里·科蒂亚尔表示,这场“史无前例”的袭击将导致“全国范围内单日损失高达数千万欧元”。据各连锁店提供的数据,约有400家店铺受到影响。

ANSSI工作人员表示,法国正与欧洲伙伴协调应对,因其他国家也可能受到波及;当局同时正在评估是否有其他法国公司面临同样风险。法国私营网络安全公司塞科亚的高级网络威胁情报分析师费利克斯·艾梅表示,西班牙、葡萄牙、波兰、德国和比利时也受到了影响。

一位消息人士透露,ANSSI在丹麦丹佛斯公司提供的软件中发现了一个漏洞。而丹佛斯指出,事故中受影响的制冷设备均为较旧型号,其安全软件无法更新以排除已识别的风险,并补充称,公司“长期以来一直建议客户保持受支持系统的固件更新,并将旧版系统迁移至当前一代的解决方案”。

北京商报综合报道