

# 巨人迟暮 英特尔是否还有“第二春”

本报记者 秦泉 北京报道

截止到9月25日收盘,“66岁”的英特尔市值为1007亿美元,不及“65岁”的AMD一半的市值,仅为“31岁”英伟达市值的3%。

近几年在半导体领域,英特尔在制程工艺上疲惫尽显,尽管其仍风光不再

如果说,英特尔与AMD目前的竞争仍然可以用“平分秋色”来形容,那么与“小弟”英伟达的竞争则输得彻彻底底。

过去数十载,英特尔的霸主地位让世人难以望其项背,偶有挑战者也几乎都铩羽而归。

英特尔与AMD同样脱胎于仙童半导体,后被称为“集成电路之父”的罗伯特·诺伊斯和提出了“摩尔定律”的戈登·摩尔在1968年成立了英特尔。翌年,被戈登·摩尔拒之门外,在仙童半导体曾经负责销售的杰里·桑德斯成立了AMD。二者不约而同地都选择了CPU这一赛道,长达数十年的博弈也就此拉开了帷幕。

虽在这场博弈中,AMD偶有高光时刻,但大部分时间均处在英特尔的压制之下。数据显示,AMD的全球份额在2016年时已不足

## IDM2.0步履蹒跚

研究机构TrendForce集邦咨询数据显示,到2024年第二季度,英特尔并未进入全球晶圆代工厂营收前10名。

面对困局,英特尔不是没有考虑过转变。“英特尔重振业务的列车即将离开站台,我希望你们都已经上车。”2022年,帕特·基辛格上任英特尔CEO一年后在美国旧金山举办的投资者大会上表示。

帕特·基辛格所驾驶的这辆“列车”第一站就是IDM2.0。作为IDM(集成设备制造商)模式的典型代表,英特尔在芯片领域以自给自足闻名,从设计到制造,再到封测,已经形成了闭环。但IDM模式对研发能力、资金实力和技术水平都有很高的要求,于是以台积电为代表的Foundry(晶圆代工)模式兴起,再加上专注设计的Fabless(无工厂芯片供应商模式)企业,芯片产业开始由IDM模式转向Foundry+Fabless合作模式。

以“巨头”自居,但出现了亏损百亿美元、拟裁员万余人、AI“掉队”等情况,被一些业界人士认为已渐入“中老年”期的英特尔正在沦为他人的收购对象。近日,据彭博社报道,美国私募巨头阿波罗全球管理公司提出向英特尔投资数十亿美元。除这项尚不明朗的投资计划外,英特尔

10%。尤其是在更强调性能的服务器市场,英特尔巅峰时全球市场占有率高达99%,AMD则不足1%。如今,这种优势正在被AMD蚕食。根据CPU市场追踪机构Mercury Research的数据,2024年第二季度,AMD服务器CPU市场份额达到24.1%,环比增长0.5%,同比增长5.6%。相应地,英特尔服务器CPU市场份额下降至75.9%。

尽管市场份额依然悬殊,但金额差距很小。第二季度,英特尔从服务器CPU销售中获得了30亿美元的销售收入,而AMD获得了28亿美元。

如果说,英特尔与AMD目前的竞争仍然可以用“平分秋色”来形

## IDM2.0步履蹒跚

研究机构TrendForce集邦咨询数据显示,到2024年第二季度,英特尔并未进入全球晶圆代工厂营收前10名。

英特尔IDM2.0中的重头戏就是加快建设各地的晶圆代工厂。为此,英特尔公布了在欧洲、美国等地的建厂计划,总投资额超600亿美元(约4200亿元人民币)。

为此,英特尔专门成立了一个新的独立业务部门“英特尔制造服务部”(Intel Foundry Services, IFS),英特尔未来计划成为美国、欧洲客户的主要晶圆代工、封装服务供应商之一,IFS将为客户提供晶圆代工及封装服务。值得注意的是,其不限于自家的X86,还提供ARM、RISC-V等多种IP组合的代工服务。

不仅如此,2023年英特尔还试图以54亿美元收购高塔半导体公司(Tower Semiconductor)。看似每则消息都足够重磅,但最

近期还传出更多资本动作。路透社的消息称,高通近期已与英特尔接洽收购事宜,潜在交易将面临反垄断审查并需要剥离资产,此次收购将成为自2018年博通试图以1420亿美元收购高通之后科技行业最大的收购尝试。《中国经营报》记者就上述传闻向英特尔方面求证,对方

表示不予置评。记者从多位业内人士了解到,英特尔目前的局面是CPU与GPU在市场竞争中逐渐势弱的缩影。过去几年里,市场拥抱AI成为压垮英特尔的最后一根稻草。而在这种趋势愈加明显的情况下,英特尔仍然“固执”地选择晶圆代工这样的重资产项目,令其积重难返。

在两年AI的风口下,英伟达成为最大赢家,凭借着在AI处理领域超过八成的市场份额,英伟达是股价累计上涨了超过550%。英伟达的市值也一度突破了3万亿美元大关。

独立研究公司CFRA Research的资深行业分析师安杰洛·齐诺(Angelo Zino)表示:“过去两到三年里,市场向AI的转变成了压垮英特尔的最后一根稻草。他们就是没有适当的能力来制造AI芯片。”

## IDM2.0步履蹒跚

研究机构TrendForce集邦咨询数据显示,到2024年第二季度,英特尔并未进入全球晶圆代工厂营收前10名。

效果却不尽如人意。收购高塔半导体公司最终未获监管部门的批准。各地的建厂计划也出现推迟的情况。

研究机构TrendForce集邦咨询数据显示,到2024年第二季度,英特尔并未进入全球晶圆代工厂营收前10名。在过去一年里,英特尔曾经在2023年第三季度短暂上榜,市场份额为1%。

近期的德意志银行会议上,帕特·基辛格表示:“代工业务比预想的更具挑战性,其他芯片公司似乎愿意继续与亚洲制造商合作,而不是将他们的产品送到我们在美国的工厂。我想说,这令人惊讶,也令人失望。”

值得注意的是,由于英特尔代工业务还在持续投入期,仍处于亏损之中,第二季度代工业务运营亏损28亿美元。

表示不予置评。记者从多位业内人士了解到,英特尔目前的局面是CPU与GPU在市场竞争中逐渐势弱的缩影。过去几年里,市场拥抱AI成为压垮英特尔的最后一根稻草。而在这种趋势愈加明显的情况下,英特尔仍然“固执”地选择晶圆代工这样的重资产项目,令其积重难返。

在英特尔2024财年第二财季财报公布后,这种颓势一览无余。在第二财季,英特尔的营收为128.3亿美元,同比下降1%;净利润亏损16.1亿美元,远超过市场预期的亏损5.4亿美元,不仅如此,英特尔方面作出预测,下一季度亏损或将进一步扩大。

与财报一起公布的还有英特尔的裁员计划,英特尔CEO帕特·基辛格在给员工发出的备忘录中宣布了裁员计划:公司将实施一系列“重大成本削减措施”,旨在2025年前节约100亿美元成本。而该计划的核心内容正是裁减约15000个工作岗位,占公司员工总数的比例约15%。

## IDM2.0步履蹒跚

研究机构TrendForce集邦咨询数据显示,到2024年第二季度,英特尔并未进入全球晶圆代工厂营收前10名。

英特尔首席财务官David Zinsner表示:“Foundry的运营亏损为28亿美元,环比恶化。预计第三季度的运营亏损将继续以大致相同的速度持续,超过85%的晶圆量仍来自EUV前节点,具有不具竞争力的成本结构和电源、性能和面积缺陷,反映在基于市场的定价中。我们在爱尔兰工厂的Intel 4和Intel 3设施的持续扩展,以及支持我们领先技术快速发展的研发和启动成本的上升,也将对盈利能力产生影响。”

不过,英特尔的代工产业并非一无是处,起码为英特尔带来了巨额的补助。今年3月,美国商务部宣布,将通过《芯片与科学法案》给予英特尔85亿美元的补助以及110亿美元的贷款。

## 错失AI

从市场角度来看,目前AI对PC销售的实质性推动作用相对有限。

重金押注晶圆代工并不能在短期内取得收益。据帕特·基辛格所言,该业务直到2030年才能实现盈利。作为曾经全球最大的芯片制造商,在晶圆代工过多的投入,也让英特尔错失了AI的风口。

反观曾经的“小弟”,Chat-GPT的爆火和AI(人工智能)时代的到来,让英伟达和它的掌舵人黄仁勋成为最大赢家。数据中心GPU领域,2023年英伟达出货量达到376万片,占据全球近98%的市场份额。

当地时间8月29日,英伟达发布的最新财季的财报显示,截至6月28日的2025财年第二财季,公司实现营收300.40亿美元,同比上涨122%,净利润同比增长168%至166亿美元。

值得注意的是,由戈登·摩尔提出的摩尔定律也被英伟达CEO黄仁勋认定:“摩尔定律死了,它结束了,结束得很彻底。”

摩尔定律的核心内容为处理器性能大约每两年翻一倍,同时价格下降为之前的一半。而英伟达直接将GPU架构的更新频率从两年一次更新加速到一年一次。

与英伟达All in人工智能芯片不同,英特尔选择在自己擅长的PC领域发力AI。帕特·基辛格直言:“我们相信,AI PC是未来数年PC市场的关键转折点,我们的每一款产品都会集成AI。”

不过,从市场角度来看,目前AI对PC销售的实质性推动作用相对有限,众多用户更多的是出于对PC本身的需求进行付费,而非直接为AI技术买单。

帕特·基辛格也坦言:“我们对AI PC类别的定义和推动投资将在短期内给利润率带来压力。我们的收入增长未达预期,

且尚未充分受益于AI的强劲趋势。我们的成本高昂,利润率低下,亟须采取更加果断的措施加以解决,尤其是在考虑到2024年下半年及未来前景比预期更为艰难的情况下。”

不过,在内外交困之际,英特尔频频利好。据悉,美国私募巨头阿波罗全球管理公司已向英特尔公司提出数十亿美元的投资计划。

与此同时,英特尔的代工业务也迎来进展,将为亚马逊代工定制款AI芯片,双方就产品和晶圆达成数年的合作框架。英特尔和亚马逊云服务部门AWS将共同投资于人工智能计算的定制半导体,这项工作将依赖于英特尔的18A工艺,这是一种先进的芯片制造技术。

不过,美国银行分析师Vivek Arya在一份报告中表示:“英特尔已经为AWS提供CPU很长时间了,所以定制化并不是什么新鲜事,而人工智能架构(网络)在18A上的胜利可能只会从2026年开始发挥作用。”

除此之外,近日,英特尔也正式推出了全新的Xeon6处理器和Gaudi3AI加速器,这两款产品旨在增强公司在人工智能(AI)和高性能计算(HPC)领域的竞争力。

Evercore分析师马克认为:“这些更新是积极的,但还不足以让我改变对英特尔发展轨迹的看法。”

伯恩斯坦研究公司的分析师斯泰西·拉斯贡表示:“英特尔的未来取决于下一代芯片制造技术的成败,该技术预计将于明年投产。英特尔希望至少在技术上超越竞争对手,重回技术领先地位有助于提高利润率,增强客户信心。”

## 不做“期货”,节后开测

# 字节跳动推出两款视频生成大模型

本报记者 李静 深圳报道

自今年年初OpenAI的Sora惊艳亮相之后,视频生成大模型就备受业界关注。虽然Sora至今尚未对外公测,但在中国国内却因此掀起了一波不小的视频生成大模型的竞赛,一些大厂和创业公司纷纷加入其中。

9月24日,字节跳动旗下火山引擎在深圳举办AI创新巡展,同时发布了豆包视频生成PixelDance和豆包视频生成Seaweed两款大模型,目前已经面向企业市场开启邀测。

早在此次火山引擎发布两款豆包视频生成大模型之前,字节跳动旗下的文生视频产品Dreamina(隶属于视频剪辑产品“剪映”)在今年5月改名为“即梦”,改名的同时产品做了一次更新,普通用户即可使用AI文生视频工具生成3秒的短视频,VIP用户可以生成6秒的短视频。

火山引擎总裁谭待在接受《中国经营报》等媒体采访时表示,剪映对视频的理解对于豆包视频生成大模型有很大帮助。与此同时,即梦的内测版已经在使用最新的豆包视频生成模型——Seaweed。

同时,谭待还表示:“豆包视频生成大模型不是‘期货’,已经在内测API和应用,(国庆)节后会放出更多公开API。”

值得关注的是,此次火山引擎不只发布了两款视频生成大模型,还带来了多款垂类大模型,包括新发布音乐模型、同声传译模型,升级通用语言模型、文生图模型、语音模型。“垂类大模型产品的发布,对于大模型在B端市场的落地具有重要意义。”北京大数据协会理事、北京融信数联科技有限公司CTO张广志对记者说道。

## 破解多主体互动和一致性难题

此前国内已经发布了多款视频生成大模型。从效果来看,目前对外公测或者开放使用的视频生成大模型在运动平滑度和成像质量方面表现出色,但整体来看,大多数只能完成简单指令,并且仍存在一些稳定性问题。例如,某些模型在生成过程中,生成质量不够稳定。

豆包视频生成大模型在活动现场展示的视频生成效果则有令人惊艳之处。效果显示,豆包视频生成模型已经能够实现自然连贯的多拍动作与多主体复杂交互。有创作者在抢鲜体验豆包视频生成模型时发现,其生成的视频不仅能够遵循复杂指令,让不同人物完成多个动作指令的互动,人物样貌、服装细节甚至头饰在不同运镜下也保持一致,接近实拍效果。

从技术实现路径来看,OpenAI推出的文生视频大模型Sora,其背后采用的是融合架构DiT

## 多款垂类大模型亮相

此次,豆包大模型不仅新增视频生成模型,还发布了豆包音乐模型和同声传译模型,全面覆盖语言、语音、图像、视频等全模态,可以满足不同行业和领域的业务场景需求。

今年5月,字节跳动正式发布豆包大模型时,定价只有0.0008元/千tokens,低于行业99%,带头把大模型的价格打到了“厘时代”,掀起了国内大模型的降价潮。

张广志对记者指出,价格战直接降低了B端企业使用大模型的门槛,使得更多企业能够负担

(Diffusion Transformer)和原生多模态技术路线。

据介绍,豆包视频生成模型也是基于DiT架构。不过,豆包视频生成模型通过高效的DiT融合计算单元,让视频在大动态与运镜中自由切换,拥有变焦、环绕、平摇、缩放、目标跟随等多镜头语言能力。全新设计的扩散模型训练方法攻克了多镜头切换的一致性难题,在镜头切换的同时保持主体、风格、氛围的一致性。

“豆包视频生成大模型能攻克指令遵循、运镜(多镜头下主体一致性)等难题,背后有技术上的突破和全栈能力的优势,还有抖音、剪映对视频理解的优势。”谭待说道。

张广志对记者表示,国外的视频生成大模型可能在算法优化、模型训练、生成效果等方面更加成熟和稳定。然而,国内企业也在不断努力追赶和超越。通过加大研发投入、优化算法模型、提

## 多款垂类大模型亮相

得起大模型的费用,从而促进了AI大模型在B端市场的普及。

“以前大模型的调用成本是大规模应用创新的阻碍,经历了降价潮后,大模型的价格已经不再是阻碍。”谭待说道。

据火山引擎方面披露,截至9月,豆包语言模型的日均tokens使用量超过1.3万亿,相比5月首次发布时增加了10倍,多模态数据处理量也分别达到每天5000万张图片 and 85万小时语音。

谭待认为,随着企业大规模应用,大模型支持更大的并

升数据处理能力等措施,使得国内企业在视频生成大模型领域也取得了显著成果。

据字节跳动方面介绍,经过剪映、即梦AI等业务场景打磨和持续迭代,豆包视频生成模型具备专业级光影布局和色彩调和,画面视觉极具美感和真实感。深度优化的Transformer结构,则大幅提升了豆包视频生成的泛化能力,支持3D动画、2D动画、国画、黑白、厚涂等多种风格,适配电影、电视、电脑、手机等各种设备的比例,不仅适用于电商营销、动画教育、城市文旅、微剧本等企业场景,也能为专业创作者和艺术家们提供创作辅助。

目前,新款豆包视频生成模型正在即梦AI内测版小范围测试,未来将逐步开放给所有用户。

谭待表示:“视频生成有很多难关亟待突破。豆包的两款模型会持续演进,在解决关键问题上探索更多可能性,加速拓展AI视

## 多款垂类大模型亮相

发流量正在成为行业发展的关键能力。据介绍,业内多家大模型目前最高仅支持300K甚至100K的TPM(每分钟token数),难以承载企业生产环境流量。例如,某科研机构的文献翻译场景,TPM峰值为360K,某汽车智能座舱的TPM峰值为420K,某AI教育公司的TPM峰值更是达到630K。

“接下来要做的是在价格基础上,提高大模型质量和性能。质量就是让模型能力更强、更多元;性能方面,就是解锁更多场



新款豆包视频生成模型正在即梦AI内测版小范围测试,未来将逐步开放给所有用户。 本报资料室/图

频的创作空间和应用落地。”

不过,目前豆包生成视频大模型的定价还没确定。谭待透露,视频模型和语言模型应用场景不同,

## 多款垂类大模型亮相

景。”谭待说道。

豆包大模型此次推出视频生成模型、音乐模型和同声传译模型,深意也是通过垂类大模型打开不同行业不同的场景需求。

“目前,大模型市场正在从通用大模型向垂类大模型发展。”张广志对记者说道,这主要是因为通用大模型虽然具有广泛的应用场景,但在解决特定领域的专业问题时往往存在局限性。相比之下,垂类大模型能够针对特定领域进行深度优化和定制,提供更加精准和高效的解

定价逻辑也不同,要考虑新体验、老体验、迁移成本等问题,最终能否广泛应用取决于是否能比以前提升更多的生产力ROI。

## 多款垂类大模型亮相

决方案。“需要强调的是,通用底座和垂直应用,分别属于不同的生态位,有各自差异化的技术迭代和商业化路径。”

谭待说道:“解锁大模型的新场景,其实需要的能力更强。所谓新场景,比如从聊天到群聊的场景,从解小学数学题到解高中数学题的场景。解决新场景的背后成本是关键,成本高无法起量,成本降低也不一定起量,更重要的是需要提升能力。我们认为在大模型使用量增长的同时解锁新场景、新能力更有价值。”