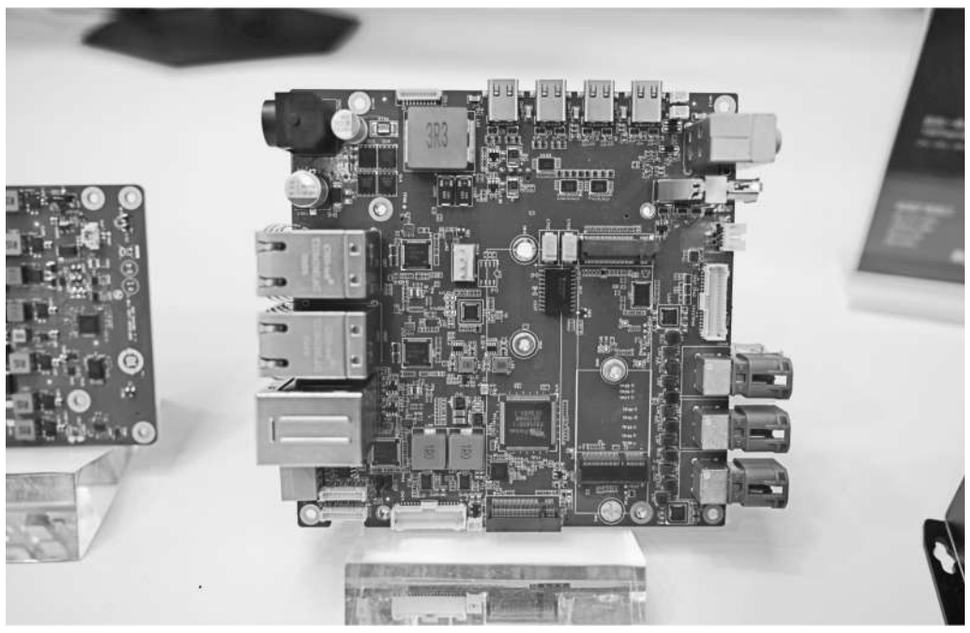
# 打破AI参数比拼 智能体重构云边端算力分配

证券时报记者 刘灿邦

近期,谷歌、OpenAI等科技巨头推出新的智能体,腾讯也将企业微信与WorkBuddy进一步打通,智能体加速进入企业办公场景。与此同时,智能手机、PC(个人电脑)等终端开始运行更大参数的模型,端侧AI正从本地问答向本地任务处理演进。

这些变化的核心均指向智能体,而智能体的发展也在推动AI基础设施重新分工:云侧从以GPU(图形处理器)为核心转向GPU、CPU(中央处理器)并行;同时,云边端之间的算力分配也在发生变化,端侧任务处理正照进现实。

不过,基础设施的重新分工并不会一蹴而就。对英特尔而言,CPU的重要性上升,不意味着产业价值分配必然向CPU倾斜;对高通、联发科等厂商而言,模型迭代周期已缩短至数月,芯片如何跟上模型变化、持续满足端侧部署需求也是现实挑战。



图为2026英特尔技术创新与产业生态大会期间展示的一款嵌入式开发板。 刘灿邦/供图

## 1 GPU与CPU配比收敛

数据显示,全球已有近80%的企业开始在内部尝试和导入智能体。在中国,活跃的企业智能体预计到2031年将超过3.5亿个。这意味着,无论是全球还是国内,智能体都呈现加速前进的状态。

在英特尔数据中心集团副总裁、中国区总经理陈葆立看来,一个智能体从思考到执行,牵动的是模型推理、任务编排、数据搬运与记忆管理的完整链路。例如,用户的电脑上可能没有专门的修图软件,但如果用户下达了修图指令,智能体就会想尽办法到网络上寻找各种资源来完成这个任务。

“过去AI止于生成答案,由GPU主导,如今AI充满了主动性,CPU的角色变得举足轻重,智能体AI对基础设施进行了重新分工。”陈葆立认为,数据中心的架构必须随智能体AI而变,全新的智算中心应以CPU为核心,CPU与GPU、IPU(基础设施单元)、存储协同运作,而不是单纯比拼某一环节的参数。

不久前,英特尔推出了至强6+处理器,最高可支持超过1000个智能体的稳定部署,同时,在满足200毫秒启动时延的SLA(服务级别协议)约束条件下,单颗至强6+最多可支持350个沙箱并发创建,以此回应了数据中心客户对真实环境下性能的诉求。

基于上述背景,英特尔判断,智能体AI时代,数据中心CPU与GPU的配比将逐步收敛至1:1,而在过去,智算中心里一块CPU需要搭配4—8块GPU。

对此,英特尔的技术专家向记者表示,想象这样一种情景,人的大脑每秒可以想到一个字,而将这个字输入到电脑只需要一毫秒。现在的情况是,原本每秒钟只能想出一个字的大脑正被一个每秒钟想出1000甚至10000个字的AI所替代,原本绰绰有余的键盘输入速度可能就不够用了。

在该技术专家看来,人的大脑好比模型本身,过去几年进步非常快,但负责敲字的键盘,也就是承接和处理这些输入任务的基础设施却跟不上了。“从这个比喻出发,GPU算力和CPU算力其实是相辅相成的,未来两者都会持续增长,但CPU算力的需求缺口会更明显。”

从更基础的原理来看,传统大模型是在思考后通过语言输出,并不会直接执行任务,随着智能体执行的任务和涉及的场景越来越多,“主智能体—子智能体—技能—沙箱”这套任务执行链路的增长速度非常快,而这个执行过程一定是放在CPU上。

## 2 端侧模型能力飙升

智能体的任务类型不同,算力需求也存在差异。复杂推理、跨系统检索等任务适合由云端提供支持;强调即时响应、个人数据调用和本地交互的任务则有机会更多地在终端完成。

随着智能体进入手机、PC等设备,AI基础设施的竞争也从数据中心延伸至终端芯片、内存、存储、操作系统和应用生态。其中,SoC(系统级芯片)需从单纯追求峰值算力转向解决智能体持续运行时的计算、功耗、调度等问题。

9月底,高通在发布第六代骁龙8超级至尊版芯片时披露,基于该芯片,高通携手阶跃、无量火与江波龙率先将阶跃StepEdge-Omni30B-MoE(混合专家)模型完整部署于端侧,打造具备全链路智能体AI能力的端侧智能体助手。

高通公司中国区董事长孟樸介绍,三四年前,高通演示的端侧文生图模型参数只有3B(Billion,十亿),去年还是支持7B—8B的模型,而现在,一些互联网领先企业已经在关注高通何时能在终端上支持100B模型。

他认为,在端侧,这些模型怎样快速地下沉到智能硬件、端侧的智能硬件在能力上怎样向上匹配是一个相向而行的过程,两边的速度都会影响整个端侧模型的使用。

值得注意的是,云侧与端侧的协同还离不开网络能力的支撑。例如,中国联通就围绕AIDN(AI定义网络)架构推进通信网与算力网的协同,对跨设备、跨系统运行的智能体而言,网络不再只是传输指令和结果的通道,还关系到任务能否及时访问远端模型、数据和工具。算力部署在哪里、任务在哪一端执行,最终都需要与网络连接能力相匹配。

“从落地节奏来看,要真正实现各类智能体在不同终端之间协同,并持续服务个人用户,可能要到2027年或2028年左右,才会开始逐步落地。”孟樸判断,端侧AI本身是一个新事物,整个产业还处于探索阶段,各种智能体在终端侧的应用,同样处在非常早期的阶段;另一方面,不同AI终端厂商的发展各有侧重,很少有企业能够覆盖所有通用终端,因此,要真正实现跨厂商、跨设备的协作,仍然需要一段时间。

## 3 技术愿景需跨越落地关

“高通具备的优势之一是在技术研发上注重前瞻性,我们强调要提前看到产业十年后可能遇到的瓶颈,并尽早去克服;对高通而言,风险和机会是并存的,关键在于快速落地、执行到位。”孟樸直言,高通比较早地意识到,AI要在端侧落地,需要产业链各方通力合作。

他举了一个例子,中国市场上的大模型数量众多、类型多样,每一个大模型都需要与终端进行适配。过去,智能手机完成适配后,可能9个月甚至12个月都无需再调整;而现在,大模型每3个月,甚至更短时间内就会迭代一次。“这部分工作量显著增加,对高通而言,非常具有挑战性。”

随着端侧应用场景不断丰富,相关市场的参与者越来越多,竞争也日趋活跃。上个月,联发科发布了天玑9600M芯片,支持端侧AI成为其卖点之一。与此同时,端侧AI的关注度也延伸到PC市场,围绕AI PC的布局和探索持续升温,整体市场竞争更加多元。

正如英特尔方面介绍,在端侧,依托CPU、GPU与NPU(神经处理单元)协同的XPU异构算力架构,英特尔加速智能体应用开发,推动传统PC向车载AIBox、AINAS等全新产品形态延伸。同时,英特尔毫不掩饰地提出,PC是办公智能体的最佳载体,要驱动PC重回生产力核心位置。

与此同时,英特尔针对CPU重回C位的叙事也并非没有挑战。当市场看到AMD与英伟达等厂商提出类似观点时,硬币往往展现出了它的另一面。

一方面,同属x86架构的AMD正在蚕食英特尔的基本盘,数据显示,2026年第二季度AMD在核心服务器CPU市场出货份额已达46.4%。另一方面,英伟达也不会坐视CPU这块蛋糕而不顾,通过供应整机柜系统,英伟达直接打包CPU、GPU与网络,CPU自然而然地跟随GPU一起进入智算中心。

对于英特尔而言,挑战还来自其供应能力。去年年底到现在,成熟的CPU平均售价提升了10%—30%。据介绍,英特尔将随市场需求和供应链状况做出一些反应,价格提升仍会回归供需基本面。

人工智能 AI瞭望台


## 交通银行“人生有戏”深圳站启幕 邀请银发客户共赏粤剧《帝女花》

10月25日下午,交通银行“沃德·颐享家园·人生有戏(深圳站)”粤剧《帝女花》鉴赏专场在深圳大剧院温情上演,邀请银发客户共赏经典。这是交通银行深圳分行本月发布“沃德·颐享家园”养老金融品牌后,面向长者客户的首场大型文化专场活动,也是该行深耕养老金融、助力老年友好城市建设的具体实践。

“落花满天蔽月光,借一杯附荐凤台上。”唐涤生的传世之作《帝女花》,经任剑辉、白雪仙演绎,成为粤港澳几代人共同的文化记忆。本场由深圳市粤剧团倾力演绎,蒋文端、黎骏声、李嘉宜、彭庆华四位中国戏剧“梅花奖”得主同台,香港粤剧名家李秋元助阵,粤港艺术家联袂献演。不少长者跟着熟悉的曲调轻声和唱。一位定居深圳的香港长者说,在家门口看到原汁原味的《帝女花》,一句唱腔就唤起了几十年前的记忆。

“人生有戏”是交通银行“沃德·颐享家园”为长者客户打造的文化品牌。交通银行深圳分行相关负责人表示,养老金融不只是管好一笔养老金,更要回应长者对美好生活的期待;把长者请进剧场,是希望金融服务从一次业务办理,变成贯穿全生命周期的温暖陪伴。围绕“沃德·颐享家园”,交通银行深圳分行正加快构建覆盖养老金金融、养老服务金融、养老产业金融的服务体系:资金端为社保、年金、个人养老金提供专业配置,服务端依托“交银长者学堂”常态化开设金融反诈、健康科普、文化文娱公益课堂,并持续推进网点适老化改造。品牌坚持“准公益”导向,既讲金融专业,也讲健康管理、长者教育、文化娱乐等非金融权益,坚决守护长者“钱袋子”。

深圳的区位,让这套服务有了特别的用武之地。罗湖口岸相通、文脉相融,是不少港人北上的“第一站”。针对跨境客群,交通银行深圳分行在口岸周边网点配备粤语服务专员,围绕账户开立、跨境汇款、社保绑定提供合规、便捷的一站式服务。本场活动由交通银行深圳分行主办、罗湖区文化广电旅游体育局支持。罗湖素有深圳“戏窝子”之称,政银携手把高品质公共文化服务送到长者身边,也让“看一场戏”带动“游一座城”,为地方文旅消费添一把火。

戏中演绎家国深情,戏外传递金融温度。交通银行深圳分行将持续厚植养老金融根基,丰富“沃德·颐享家园”服务内涵,以专业与温度托起长者“稳稳的幸福”,让更多长者的晚年生活有戏、有味、有陪伴。 (CIS)


<<上接A1版

# 从百亿市值到千亿阵营 沪电股份的PCB进阶路

创业故事联合调研组

值得关注的是,深交所持续陪伴企业成长。陈梅芳介绍:“交易所相关负责人来企业调研时也会关注我们的扩产进度。”深交所还通过“一对一”沟通方式,对信披、公司治理、股东回报、股权激励等规则问题,为企业答疑解惑。对长周期的PCB产业而言,这种稳定的制度环境,为后续的产能扩张和研发迭代提供了重要保障。

第二阶段是“升级期”(2010—2019年)。借着昆山“退二进三”搬迁的契机,沪电股份青淞厂区按照高多层数据通信板的要求设计,新建了20层以上高多层板、HDI(高密度互连板)等高阶产能的产线。这次战略转身,让沪电股份在2019年5G基站建设高峰期成为主要受益者,当年沪电股份的归母净利润约为12.06亿元,同比增长约111.41%。

眼下,沪电股份已步入第三阶段。这一次的关键词是“AI放量”。AI正在把PCB从低端制造推向高端精密制造。从多层板、HDI,到高速材料、封装基板,AI算力军备竞赛背后,承载着一整条PCB产业链的价值重估。

为何沪电股份能抓住这轮AI机遇?“不是我们踩准AI风口,而是本来就在PCB这一赛道上。”陈梅芳坦言,“公司早年服务通信客户,积累了生产高性能、高信赖的PCB板的工艺能力和产品经验。”

灼识咨询数据称,按2025年收入计,沪电股份位居全球PCB制造商前列,并在数据通信PCB、数据中心PCB、22层及以上PCB等细分市场占据领先位置。

从昆山起步,到深交所上市完成技术升级,再到站稳全球AI算力基础设施的“骨架”位置,沪电股份三十余年的三段式跃迁,既是一家电子制造企业的进化史,也是中国PCB产业突围的缩影。

## 从5G到AI:技术积累夯实算力底座

PCB层数越高、纵横比越大,工艺窗口就越窄,良率控制也越难,这对制造端的技术能力和管理能力提出了更高要求。陈梅芳随手比画着PCB的层数,“一般PCB板约3毫米厚,沪电股份做的厚产品可到14.6毫米,我们能做100层以上,技术难度非常高!”

而这恰恰是沪电股份多年来服务通信客户、应对高标准要求所积累的本事。陈梅芳回忆:“早期还得从国外请工程师、买设备,去做通信板,下游客户对产品指标、交付与验证流程要求严苛。做一块板、测试运行、判断稳定性,至少要几个月甚至一年时间,才决定能否通过认证。”沪电股份的工艺标准、品质要求不断拔高,这样才接得住AI服务器、数据交换机的订单。

PCB行业还有“连带赔偿”的规矩。陈梅芳向记者解释:“PCB板子运行出错,上面贴装的芯片、元器件要一起赔。”

沪电股份能够接住AI带来的机遇,离不开5G时代的技术积累。5G基站电路板对信号完整性、散热性和可靠性要求极高,这些技术能力和工艺经验,恰好可以平移到AI服务器领域。

从上游看,尽管北美服务器厂商在尝试简化电路板设计,但AI服务器机柜越来越大,需要将更多功能集成到一块大型基板上减少连接点,这对电路板的要求反而更高。大型基板需承载所有芯片,如同把鸡蛋放在同一个篮子里,且篮子必须足够结实。

眼下,更多PCB厂家涌入AI高阶板领域,外界不免担心同质化与价格战。“入场者多、通关者少,高多层板的认证很难,这本身就是门槛。”陈梅芳介绍,尽管AI爆火,沪电股份仍保持客户结构的多元性,除了AI服务器、数据交换机等通信业务订单,汽车板、工业控制板的订单也持续承接。

一家来自长三角的数据通信设备商企业市场负责人对记者介绍:“25年前,第一次跟沪电股份合作背板产品,如今,合作产品已经拓展到了高速网络交换机、路由器等领域。”目前,沪电股份与5家主要终端客户的合作时间在20年以上。

AI时代,能做高阶电路板的公司,业绩跟着GPU厂商,一路腾飞,也愈发重视投资者回报。公司10月16日公告称,控股股东提议,针对2026会计年度实施的现金分红总额不低于22亿元。

当全球算力竞赛进入硬件决胜阶段,PCB产业链的战略地位日益凸显。高盛认为,随着AI基础设施的快速扩张,产能持续扩大且利用率保持在较高水平,对AI领域的PCB市场持乐观态度。

目前,深交所已聚集约30家PCB上市企业,总市值约1.37万亿元,形成一个颇具规模的产业板块。其中,沪电股份与深南电路、鹏鼎控股、东山精密、胜宏科技等头部公司,均加入千亿元市值阵营。这些企业虽同处PCB高景气赛道,但各自依托自身技术积淀,既有技术、产品路线的差异化布局,也在数据通信、汽车电子等市场直面同位竞争,多元并进,共同构筑起中国PCB产业的领先格局。

沪电股份董秘李明贵是跟随吴陈夫妇创业的老员工,他透露:“公司也在加快资本开支以满足客户群体需求,2026年上半年资本开支超过35亿元,近两年宣布的投资规划超过前30多年的投资总额。”

## 从中国到全球:产能网络织就交付弹性

全球化始终根植于沪电股份的基因之中。创业之初,订单便来自全球各地;如今,外销占比已稳定在80%以上。2022年,公司设立泰国生产基地,随后上下游配套企业也陆续跟进布局。陈梅芳坦言,此举意在满足海外客户对本地化供应和供应链多元化的需求,泰国生产基地在2026年已迈入规模化运营期。

李明贵介绍,算力订单要效率,昆山的配套设施齐全,水电稳定、熟练老师傅都在岗,更适合做高端板;海外客户要的是产地全球化布局,保持供应链稳定。因此昆山总部继续“啃”最高等级的产品,待工艺成熟后再交给湖北黄石、泰国生产基地。

当前,泰国生产基地已实现30层以内相关PCB产品批量量产,从产能爬坡期逐步迈入规模化运营期,2026年前三季度实现扭亏为盈。

去年至今,外资机构不止一次组团前往泰国生产基地调研。全球化对沪电股份而言,不仅仅是产品外销,更在生产效率与供应链约束之间,搭建起一套可灵活调配的产能网络。

当前,PCB产业正处于由AI驱动的强劲上行周期,产能供给维持高位、有效产能受限,东亚供应链尤其是中国的PCB制造商受益最大。

“公司目前的研发聚焦于高速、高密度数据通信等下游应用领域。”沪电股份技术负责人介绍,就数据通信而言,重点研发高速材料、信号与电源完整性、复杂高密度结构,多层压合PCB的产品可靠性。

中证鹏元研究发展部资深研究员翁欣表示:“PCB在散热、隔离等综合功能上具有不可替代性,推动行业迈入长周期强劲发展阶段。随着全球核心算力产品的大规模批量出货,行业景气度有望迎来拉升,产能利用率攀升与订单饱满,将在业绩端得到直接验证。”

面对外界对PCB行业的乐观预期,陈梅芳的态度更为笃定。她表示:“我们只做PCB一件事,相信未来20年,没有其他国家能超越中国的PCB。”

创业故事



印制电路板被称为“电子产品之母”,沪电股份的故事就从这块绿色“底座”开始。

图虫创意/供图