

芯片厂商携手上下游共建 AI PC 生态



图为 AMD AI PC 创新峰会现场，AMD 宣布与 OEM 厂商持续建设 AI PC 生态

本报记者 王信豪

2024 年第一季度，AI 在云端的内容生成正在逐渐向 PC、手机等终端靠近。英特尔 CEO 帕特·基辛格在年初的 CES 2024 上表示“AI PC 是个人电脑的‘寒武纪’时刻”，言外之意，AI PC 将迎来如寒武纪时的“生命大爆发”；高通 CEO 安蒙在近日召开的中国发展高层论坛中也表示“下一代 AI PC 将成为行业转折点”；在 3 月举办的 AMD AI PC 创新峰会上，AMD CEO 苏姿丰现身会场与合作伙伴共议 AI PC 生态。一边是这些高层对于 AI PC 的前景表示看好，另一边是搭载着 AI 处理器的众多 AI PC 先后面世，AI PC 元年来的势凶猛。

2024 年不仅是 AI PC 元年，也将是 AI PC 的转折之年。相比 2023 年的概念期，2024 年的 AI PC 更加注重端云一体，处理器也从多种芯片和架构的“融合”走向架构与场景的“适合”。与此同时，为 AI PC 供应处理器的芯片企业加强软件生态建设，加速与 OEM 厂商、品牌方、下游应用企业共建生态，推动 AI PC 走向更加成熟的发展阶段。

五部门发布集成电路企业或项目 税收优惠政策

本报讯 近日，国家发展改革委、工业和信息化部、财政部、海关总署和税务总局五部门发布《关于做好 2024 年享受税收优惠政策的集成电路企业或项目、软件企业清单制定工作有关要求的通知》（以下简称《通知》）。《通知》提到，2024 年享受税收优惠政策的集成电路企业或项目、软件企业清单（以下简称“清单”）制定工作，基本沿用 2023 年清单制定程序，享受税收优惠政策的企业条件和项目标准。

《通知》指出，2023 年已列入清单的企业如需享受新一年度税收优惠政策（进口环节增值税分期纳税政策除外），2024 年需重新申报。申请列入清单的企业应于 2024 年 3 月 25 日至 4 月 16 日在信息填报系统中提交申请，并生成纸质文件加盖企业公章，连同必要证明材料（电子版、纸质版）报本省、自治区、直辖市及计划单列市、新疆生产建设兵团发展改革委或工业和信息化主管部门（由地方发展改革委确定接受单位）。经审计的企业会计报告须在提交申请时一并提交。

记者在中国半导体行业协会官微注意到，中国半导体行业协会提醒相关集成电路生产企业或项目、重点设计企业、先进封装测试企业、集成电路产业的关键原材料、零配件（靶材、光刻胶、掩膜版、封装基板、抛光垫、抛光液、8 英寸及以上硅单晶、8 英寸及以上硅片）生产企业可关注《通知》，符合条件企业可于 3 月 25 日至 4 月 16 日期间在

系统完成填报提交，集成电路研发费用加计扣除也需要在 3 月 25 日至 4 月 16 日申报。

此外，《财政部、税务总局、国家发展改革委、工业和信息化部关于提高集成电路和工业母机企业研发费用加计扣除比例的公告》提及的国家鼓励的集成电路生产企业或项目归属企业、重点集成电路设计企业也需要在本次进行系统申报。

中国半导体行业协会请集成电路相关企业或项目在申报时关注以下要点：

一是本次申报涉及的研发费用为 2023 年 1 月 1 日至 2023 年 12 月 31 日期间实际发生的可加计扣除研发费用，用于申报 2023 年所得税税前扣除。

二是所得税优惠与提高研发费用加计扣除优惠可单独申请其中一项，也可同时申请。

三是超过所得税减免优惠期或处于税务亏损状态的生产企业和项目可以单独申请提高研发费用加计扣除优惠。

四是针对重点设计企业，单独申请提高研发费用加计扣除优惠或同时申请所得税优惠及提高研发费用加计扣除优惠，应纳税所得额需按照提高后的研发费用加计扣除金额进行计算，结果应符合申报条件要求。只申请所得税优惠不申请提高研发费用加计扣除优惠的重点设计企业，无上述要求，但应纳税所得额应符合申报条件要求。

（张心怡）

中芯国际发布 2023 年年报 营收 452.5 亿元

本报讯 记者王信豪报道：3 月 28 日，中芯国际发布 2023 年年报。中芯国际董事长刘训峰在致股东信中称，中芯国际 2023 年总营收为 452.5 亿元，调整波动幅度好于行业平均水平，毛利率为 22%，年平均产能利用率为 75%，基本符合年初指引。

2023 年是中芯国际营收在五年内（2019—2023 年）首次下降的一年。年报显示，公司总营收同比下降 8.6%，其中晶圆代工业务营收为 408.8 亿元，同比减少 9.8%，毛利率由 2022 年的 38.3% 降至 2023 年的 21.9%，净利率为 14.1%，同比减少 15.5%。

数据显示，中芯国际 2023 年 8 英寸和 12 英寸晶圆片总产量达 607.4 万片，销售量为 586.7 万片，库存量为 72.4 万片，库存量同比增长 40.1%，主要原因是企业生产备货。

中芯国际表示，过去一年，半导体行业处于周期底部，全球市场需求疲软，行业库存较高，去库存缓慢且同业竞争激烈。受此影响，集团平均产能利用率降低，晶圆销售数量减少，产品组合变动。此外，集团处于高投入期，折旧较 2022 年增加。以上因素都影响了公司 2023 年度的财务表现。

与此同时，中芯国际在研发方面的投

入约 49.9 亿元，研发投入总额占收入比例达 11%，同比增加 1%，申请专利 644 件，授权专利 581 件。据了解，中芯国际 28nm 超低功耗项目平台、40nm 嵌入式存储工艺汽车项目平台、4Xnm NOR Flash 工艺平台等项目在 2023 年已完成研发，进入小批量试产。

中芯国际在报告中指出，2023 年下半年，终端市场的需求呈一定复苏迹象，但整体供应链库存处于高位，终端产品销售状况处于调整阶段，库存消化仍为 2023 年半导体行业主旋律。从长远角度看，全球半导体行业兼具周期性和成长性，短期的供需失衡不会影响行业的长期向好。而随着新一轮智能化科技应用逐渐落地，产业变革趋势也在逐步确立。家居、教育、科研、商业、工业、交通、医疗等领域的智能化应用需要大量逻辑、模拟、射频、光电以及传感器器件，这将为晶圆代工企业注入新一轮的市场增量，产业链各个环节有望逐轮回暖。

展望 2024 年，中芯国际表示，公司仍然面临宏观经济、地缘政治、同业竞争和老产品库存的挑战，预计公司表现“中规中矩”，随着半导体产业链一起摆脱低迷，在客户库存逐步好转、手机与互联需求持续回升的共同作用下实现平稳温和的增长。

总投资 43 亿元 美光西安新厂房奠基

本报讯 记者姬晓婷报道：3 月 27 日，美光西安新厂房奠基仪式在西安高新区举行。美光已在西安布局封装测试工厂，本次奠基的项目就是在原有工厂基础上的扩建。该项目总投资 43 亿元，新建产线将主要用于生产移动 DRAM、NAND、SSD 等存储产品。

在奠基仪式上，美光总裁兼首席执行官桑杰·梅赫罗特拉（Sanjay Mehrotra）表示，在此次投资之前，美光在西安的投资已达到 110 亿元。桑杰表示，西安将与上海、北京和深圳的客户实验室合作，根据中国市场的需求定制产品和解决方案。美光中国区总经理吴明霞在奠基仪式上表示，此次加建的新厂房，将用于制造包括但不限于移动 DRAM、NAND 和 SSD 等产品，这些产品将广泛应用于国内外制造的个人电脑、旗舰智能手机和汽车等产品中，上述产品将在西安进行本土化生产和供货。

美光还将在西安工厂投资多个工程实验室，提升产品可靠性认证、监测、故障分析和测试的效率，从而帮助客户加快产品上市时间。新厂房预计于 2025 年下半

年投产，并根据市场需求逐步增产。新厂房落成后，美光西安工厂的总面积将超过 13.2 万平方米。

美光也在推进收购力成半导体（西安）有限公司（力成西安）的资产，并将向力成西安 1200 名员工提供新的就业机会。新厂房项目还将额外增加 500 个就业岗位，使美光在中国的员工总数增至 4800 余人。

在当天下午举行的媒体见面会上，美光全球运营执行副总裁 Manish Bhatia 在答记者问时表示，此次投资落成是由美光与客户关系的需求推动的。他表示，低功耗 DRAM 需要多种制造工艺来实现产品高容量、高可靠性、低功耗的要求。制造过程也相较此前更加复杂，需要的设备、材料都有不同，这就要求美光需要更多的技术能力。“我们之所以选择西安，既是因为其在工程、质量和运营出口方面的出色表现，也是因为在低功耗 DRAM，包括智能手机、汽车和個人计算领域，我们有许多中国客户，因此可以利用这一优势。”Manish Bhatia 表示。

面向未来，端云一体是更具性价比的解决方案，端侧注重隐私，云端能满足部分大算力处理任务的需要。

手、AI 本地知识库、AI 图像处理以及 AI PC 管理。其中的部分功能通过云端已经可以实现，而更多的应用，则需要照顾本地化、私人化的使用需求。

“面向未来，端云一体是更具性价比的解决方案。端侧注重隐私，云端能满足部分大算力处理任务的需要，二者结合能够带来更好的 AI PC 体验。”在 2024 英特尔商用客户端 AI PC 产品发布会上，Extend Office 相关负责人向记者表示。

不论是 GPU 还是 NPU，单一处理器并非 AI 计算的唯一解，更重要的是芯片算力与 AI 应用需求之间的互相匹配。

“特立独行”，不同于其他芯片厂商的集思路，英伟达更加关注自家 RTX 系列的独显。如同在 CES 2024 上，英伟达发布了 RTX 40 系列 Super 的消费级显卡，其 AI 算力同样十分强劲。

独显与集显，乃至异构集成这种不同路径之间的差异并无优劣之分。具体而言，英伟达 RTX 显卡针对游戏市场，例如实现 3A 游戏中的光追渲染；而搭载如英特尔酷睿 Ultra 处理器的 AI PC，则更加关注轻薄笔记本电脑所在的商业办公领域，两块市场消费者的习惯与体量均有不同。

综合来看，不论是 GPU 还是 NPU，单一处理器并非 AI 计算的唯一解，更重要的是芯片算力与 AI 应用需求之间的互相匹配。

AI PC 的生态建设不能各自为战，需要芯片厂商、ISV 厂商和 OEM 厂商协同运作，推进 AI PC 产业整体长远发展。

高通在 MWC 2024 上推出了预优化模型库高通 AI Hub。“它能够帮助开发者快速地将预优化的 AI 模型集成进应用程序，从而打造更快、更可靠且更具隐私性的用户体验。”高通高级副总裁兼技术规划和边缘解决方案业务总经理马德嘉表示。

AMD 推出的开发工具也将支持多个 AI 模型，并与微软合作，在操作系统层面提供支持。英特尔深入推进其 CEO 帕特·基辛格在去年 10 月提出的“AI PC 加速计划”，并于近日向开发者提供最新版本的开发套件 Open VINO，以优化 AI 模型的特性。

“AI PC 的生态建设不能各自为战，需要芯片厂商、ISV 厂商和 OEM 厂商协同运作，形成协议与标准，更好地推进 AI PC 产业整体长远发展。”谭斌向《中国电子报》记者表示。

管理、设计、办公协同等。相对来讲，AI PC 的处理能力要高于 AI 手机，但又低于云端大模型。”深圳中为智研咨询有限公司研究员谭斌告诉《中国电子报》记者。另一方面，AI PC 相较于云端大模型的内容生成，也能在数据隐私的保护上给到消费者更大的安全感。

3 月 26 日，英特尔中国区技术部总经理高宇分享了关于 AI PC 的六大应用场景，分别为 AI Chatbot、AI PC 助理、AI Office 助

行 AI 运算的前提下，依照功耗和算力要求选择不同的处理器，从而提高效率，并保证其他非 AI 任务的流畅处理；要求异构集成则是因为，并不是所有时候都需要独显，能够进行 AI 计算的集显更具性价比。”英特尔的相关生态伙伴告诉记者。

在 CES 2024 上，AMD 选择取消独显，推出 Ryzen 8040 的 PC 处理器，该处理器集成了 Zen4 架构的 CPU、RDNA3 架构的 GPU 和 XDNA 架构的 NPU。据了解，三种架构负责的主要工作不同，Zen 架构用于通用处理和 AI 推理，RDNA 用于游戏和基于 AI 的内容开发，XDNA 用于低功耗的 AI 推理。

值得一提的是，在云计算领域取得骄人成绩的英伟达在 AI PC 上则显得有些

从广度来看，满足消费者日常工作需求的应用数量正在扩大，从深度来看，AI 应用也在教育、医疗、金融、法律等更加专业的细分领域落地生根。

有道首席科学家段亦涛表示，教育领域的个性化需求很强。通过算法、模型，可针对学生的数据进行深度分析，从而为每个学生量身打造个性化的学习路径。香港中文大学（深圳）和深圳市大数据研究院联合开发的医疗大模型“华佗”，可在终端离线时进行简要问诊等工作。此外，AI 助理也可依据大模型内含的法律专业知识迅速拟写合同、文书等文本文件，进一步提升相关行业工作者的工作效率。

应用的逐渐丰富需要更加成熟的软件开发环境。

从云到端，再到端云一体

如果从外观上看 AI PC 与传统 PC 的区别，微软给出的答案是“是否具有 Copilot”按钮。但是，由于 Copilot 基于 ChatGPT4，这也意味着其大部分现实功能需要在云端完成。对于 AI PC 而言，一个重要的任务就是在云和端两侧做好平衡。

新一阶段的 AI PC 将更加关注终端推理。一方面，这与 AI PC 的应用场景有关。“AI PC 更多是办公场景应用，解决个人办公与工作需求，譬如高效文档作业、编程、

处理器从融合到适合

相比广泛应用于数据中心的 GPU（图形处理器）、NPU（神经单元处理器）被认为是更适合进行边缘侧 AI 运算的处理器。因此，高通、英特尔和 AMD 在 AI PC 处理器的设计上均选择了“CPU+GPU+NPU”的异构模式。

2024 年，这种异构集成的模式被要求“分工更加明确”。

在近期 AMD 和英特尔举办的 AI PC 主题活动中，记者观察到，许多测试者在 PC 上运行 AI 应用时都会查看后台 CPU 和 GPU 的负载情况。从实际情况来看，CPU 的负载率普遍在 10% 左右。

“我们希望 AI PC 在运行时，处理器有更好的分工，同时又保持一定的集成度。希望有所分工，是因为在 GPU 和 NPU 均可执

从概念引出到生态建设

相较于 2023 年年底 AI PC 概念刚刚被提出，现在 AI PC 的最大变化在于其生态建设的进展。

目前看来，芯片企业与下游 OEM 企业已经实现了初步打通，联想、华硕、戴尔等品牌将推出的 AI PC 将会搭载对应的 AI 处理器，下一步的目标便是使 AI PC 产品与 AI 应用实现更深层次的融合，从而构筑“芯片—设备—应用”的完整闭环。

据悉，AMD 及其生态伙伴在 AI PC 上实现了绘画、视频剪辑、音乐创作、辅助编程等功能。搭载英特尔最新酷睿处理器的 AI PC，不仅可以完成日常的图文生成与语义识别，还可以在离线状态下使用 Office 插件对文本和表格进行处理，同时，也实现了知识库在 PC 端的本地部署。