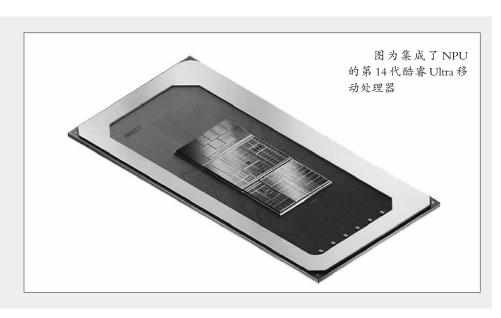
AI PC加速NPU集成应用

本报记者 沈丛

2024年被视为AI PC元年,根据市场研究机构群智咨询的预测,2024年全球AI PC整机出货量将达到约1300万台。而作为AI PC的算力中枢,集成了神经处理单元的计算处理器(NPU)也将在2024年大规模应用落地。英特尔、AMD等第三方处理器供应商及苹果等PC厂商均表示,将在2024年推出搭载NPU的电脑处理器。与此同时,如何在电脑处理器中发挥出NPU的最大功效,也成为了业内热议的话题。



NPU助力 PC 更好运行 AI 应用

据了解,NPU可以根据网络运算的特点,通过软件或硬件编程,实现多种特定用途的计算功能。相比传统 CPU 和 GPU,NPU能够以更高的效率、更低的功耗执行神经网络任务。

相比FPGA,NPU更具成本优势。业内专家告诉《中国电子报》记者,FPGA虽然同样在灵活性和可编程性方面具有优势,但主流FPGA芯片的制程节点通常在14nm~45nm之间,而NPU的制程节点往往在10nm以下,更加适合小型的PC和手机等消费电子设备。

在软件方面,NPU与主流软件实现了更好的兼容运行。"NPU能够通过标准化的接口,如OpenCL、CUDA、OpenVX等实现与AI框架交互;还更容易集成到不同的软件环境中。因此,NPU通常可以与主流的AI框架和软件栈更好地兼容,开发者也能更容易

利用这些硬件加速器。"业内专家表示。

2017年,华为最先将NPU处理器集成到手机CPU中,显著提升了手机CPU单位时间计算的数据量和单位功耗下的AI算力,此举也让业内人士看到了NPU应用于级端设备的潜力

随着AI应用越来越广泛,PC用户对于AI能力的需求也在持续增加,行业不少品牌开始尝试将NPU集成到PC处理器中。此举可以让用户不必依赖云端的计算资源,在PC端侧进行AI模型的训练和推理;NPU加持下,PC还支持用户根据个人所需进行模型的训练和优化,符合当前软件和服务的个性化、定制化发展趋势。

从目前来看,各家处理器供应商对内置 NPU 的电脑处理器表示出极大的兴趣,纷纷发布相关新产品。英特尔在近期发布了内置 NPU 的第 14 代酷睿 Ultra 移动处理器。英特尔官方表示,2024年将有 230 多款机型搭载酷睿 Ultra。苹果将在 2024年发布搭载 M3 处理器的 MacBook 笔记本电脑,据

透露,其M3处理器的NPU性能相较于M1提升了60%。AMD宣布,其首次集成独立NPU的锐龙8040处理器将于2024年年初正式发布,AMD表示,加入NPU后,锐龙8040处理器的大语言模型性能提升了40%。

业内人士普遍认为,未来"CPU+NPU+GPU"的组合将成为AIPC的算力基座。CPU主要用于控制和协调其他处理器的工作、GPU主要用于大规模并行计算、NPU则专注于深度学习和神经网络的计算。三种处理器协同工作可以充分发挥各自的优势,提高AI计算的效率和能效比。

消费级高端PC将率先应用NPU

2024年,NPU作为AI计算的核心部件 将会在PC上得到广泛应用,消费级高端PC 用户或将先一步领略其产品魅力。

业内专家认为,在PC处理器中集成NPU需要一定的技术和资金投入,这也会导致PC成本上升。因此,NPU将首先应用在

消费级高端PC产品线。

"目前搭载 NPU 的处理器主要集中在消费级 PC上。因为消费级 PC 用户的需求和使用场景相对灵活,对产品性能、功能和使用体验的重视程度超过了对可靠性和使用寿命的关注。因此,PC厂商通常更愿意在其高端产品上应用新技术。而企业级 PC对可靠性和使用寿命的要求更加严苛,企业用户更倾向于采用已经经过验证的技术,以确保系统的稳定性。因此,NPU 在企业级PC的应用会更晚一些。"业内专家向《中国电子报》记者表示。

以英特尔最新发布的两款 AI PC 处理器为例,英特尔针对企业级 PC 用户推出的第五代至强可扩展处理器中,既没有采用最新制程,也没有集成 NPU,而是继续沿用CPU 处理器,并依旧采用较为成熟的 Intel 7制程。与之形成鲜明对比的是,英特尔推出的消费级酷睿 Ultra 处理器,不仅采用了最新的 Intel 4制程,还首次在处理器中加入了NPU。

在PC上用好NPU不容易

NPU 的使用难点在于如何能释放出有效的算力。

英特尔公司中国区技术部总经理高宇表示,想要NPU充分发挥性能,并非"堆积"这么简单,而是要了解NPU的理论算力和实际应用算力,在此基础上优化现有的编程模型,使其更好地适应NPU的并行处理架构。这需要开发者和工程师深入了解NPU的工作原理和特性。除了编程模型优化外,还需要对硬件设备的性能进行优化,例如散热、功耗管理等

为了进一步实现NPU在AIPC中的普及并释放出有效算力,需要建设针对NPU的生态。"为了推动NPU在PC领域的发展,需要建立一个包括硬件供应商、软件开发者、PC应用厂商在内的完善生态系统,健康活跃的生态系统,有助于NPU技术更好地发展。"业内专家说道。

汽车进入"比芯"阶段

本报记者 姬晓婷

2023年12月28日,在小米集团创始人、董事长兼CEO雷军宣布造车的第1003天,小米汽车的真容终于揭晓:这款名为小米SU7的纯电轿车,分为SU7和SU7Max两个版本。其中,SU7Max搭载了双电机、101kWh三元锂电池、HUD、电吸门和激光雷达,零百加速(指汽车从静止到时速100km/h所需的时间)仅为2.78秒。SU7的汽车芯片也首次曝光:SU7Max的自动驾驶系统采用了2颗NVIDIADRIVEOrin,智能座舱采用了高通骁龙8295。

雷军在发布会上介绍说,SU7 Max 采用了2颗 NVIDIA DRIVE Orin,能够实现508TOPS的算力支 撑。这是什么水平?

据了解,NVIDIA DRIVE Orin在2022年3月正式销售,单芯 片算力为254TOPS,可谓当前市场上算力最强的汽车自动驾驶集成芯片。 NVIDIA DRIVE Orin 自上市以来,颇受自动驾驶汽车厂商的青睐,被大量汽车品牌应用在最新款车型。记者粗略统计:截至目前,搭载NVIDIADRIVE Orin的汽车包括极氪、蔚来等品牌在内的多种新车型。

当前,2023年上市车辆的自动驾驶级别多集中在L2+,所需算力在10TOPS左右。汽车自动驾驶功能的竞争,更多集中在软件和算法层面。小米及其他汽车厂商为自动驾驶配置的500+TOPS算力,相比于当前L2+自动驾驶级别来说,显然是冗余的,专业人士猜测这一举动多是为自动驾驶升级提供空间。

除了自动驾驶系统,智能座舱也 是当前车企比拼的"红海战场":多屏 联动、舱内游戏、端侧 AI······支持这 些功能,同样需要一颗强大的智能座 舱芯片。小米的新车选择了高通骁龙 8295,该芯片采用5nm工艺,是目前 汽车座舱平台中工艺最先进的产品。

在2023年11月刚刚举办的第二十一届广州国际汽车博览会上,多款中外汽车品牌的热门车型均采用骁龙8295作为智能座舱平台芯片。

从性能来看,骁龙8295的优势体现在核心数、CPU算力、GPU浮点运算速度等方面。据了解,骁龙8295搭载了高通第六代Kryo CPU,采用四颗超大核+四个大核的架构,CPU算力可以达到230K DMIPS;骁龙8295采用了第六代高通Adreno GPU,可以实现每秒2.9万亿次单精度浮点运算(2.9TFLOPS)或每秒5.8万亿次半精度浮点运算(5.8TFLOPS)。甚至有业内专家认为,骁龙8295的性能已经超过苹果M1。

正因为强大的 GPU以及显示处理单元 DPU, 骁龙 8295 最多可支持32 块屏幕显示, 或者最多6 块 4K 屏幕显示——接近5000 万像素的显示

算力。此次雷军宣称,SU7可以实现 "五屏联动",实现视频播放在不同屏幕之间的自由流转,其中,中控生态 屏具备16.1英寸和3K的分辨率。

在此次发布会上, 雷军对小米汽车座舱功能的描述, 更多地着墨在小米生态的联动上, 例如: 点击车身按钮, 可以实现与手机的一键连接; 手机上的导航可以滑动投屏在汽车挡风玻璃上; 小米平板挂上车, 就成为车机的一部分……

对于小米这样一家主攻手机、家电、智能家居等产品的企业来说,完整的生态是其区别于传统车企的一大优势。国家新能源汽车技术创新中心主任、总经理、中国汽车芯片产业创新战略联盟秘书长原诚寅在接受《中国电子报》记者采访时表示:"(小米)这些企业造车时,可以很好地把原有的技术迁移到汽车上来。其服务意识也更强,对于新技术的引入和迭代创新能力也更有优势。"

(上接第1版)

中国电子报评出2023年电子信息产业十件大事



大模型引领生成式人工智能发展浪潮

2023年,受OpenAI ChatGPT的强力驱动,大模型发展迈向新阶段,在全球范围内掀起了一股生成式人工智能发展浪潮。百度文心一言、阿里巴巴通义千问、腾讯混元大模型、华为盘古大模型等国内企业纷纷推出自己的人工智能大模型。赛迪顾问数据显示,中国有至少130家公司研究大模型产品,其中100亿级参数规模以上的大模型超过10个,10亿级参数规模以上的大模型已近80个,大模型数量位居世界第一梯队。生成式人工智能应用加速融入制造业、零售业、电信行业和医疗健康等垂直行业,催生出一系列新的商业模式和商业场景,深刻改变着人类社会的生产生活方式。

8

鸿蒙欧拉开源操作系统渐成主流

2023年,在移动操作系统方面,搭载开源鸿蒙操作系统的设备数量超过7亿台,220多万名开发者投入到鸿蒙生态的开发,425款软硬件产品通过兼容性测试,覆盖金融、能源、工业、交通、医疗、教育等关键产业。在服务器操作系统方面,开源欧拉操作系统累计部署610万套,已经实现对Arm、x86、RISC-V等全球主流通用计算架构及英伟达、昇腾等主要AI处理器的支持,中国市场占有率达36.8%,为150余个国家和地区提供服务。经过多年积累,鸿蒙、欧拉等开源操作系统版本快速迭代,能力迅速增强,市场占有率不断提升,构筑起中国



基础软件之"魂"。

我国高端手机产业链供应链建设取得突破

8月,采用新型麒麟9000S芯片的华为新款手机Mate 60 Pro开始销售。麒麟9000S芯片采用了先进工艺,实现了我国手机高端芯片自主制造的重大突破。华为Mate 60 Pro可支持在无地面网络拨打和接听卫星电话,是全球首款支持卫星通话的大众智能手机。该手机同时将AI算法与多模低辐射天线技术相融合,能够在高铁、地铁、电梯、车库等弱信号场景下实现更稳定的网络连接。华为Mate 60 Pro"现象级"热销,标志着我国高端手机产业链供应链建设取得突破。



中国显示产业迎来投资大年

2023年,中国显示产业再一次迎来投资大年。2月,京东方总投资290亿元的第6代新型半导体显示器件生产线项目在北京开工。4月,总投资22亿元的重庆迈特光电光掩膜版项目正式开工建设。10月,莱宝高科宣布拟投资90亿元在湖州建设徽腔电子纸显示器件(MED)项目。11月,东旭集团总投资155亿元的超薄柔性玻璃、LED显示模组和芯片项目落地衢州。12月,杉杉股份投资约60亿元在扬州建设高端显示用偏光片生产线项目。2023年,我国显示产业瞄准IT等高端显示领域,OLED、Micro LED等前沿技术,玻璃、偏光片等上游领域持续扩充产能,全年新建产线总规模超过千亿元,我国显示产业高质量发展再上新台阶。

(上接第1版)陈晓红建议,一是聚焦标志性产业链,贯通上下游关键环节,稳链强链、优化布局,加强产业统筹协调,推动产业基础高级化、产业链现代化;二是梳理重点产业链,精准绘制产业链画像,科学支撑产业链诊断,找准薄弱环节,补齐短板和弱项,确保不"掉链子";三是提高产供分析与预警能力,结合产业政策、产业要素、供应商信息、国内外环境等多维度关联数据,应用人工智能算法模型,提供产供风险分析与预警,有效提升产业链供应链的自主性和韧性。

下一步,我们要着力补短板,在 关系安全发展的领域,实施"一链一 策",强化产业链上下游协同攻关,化 点成珠、串珠成链,确保产业链供应 链稳定畅通;着力锻长板,聚焦光伏、 新能源汽车、5G等优势产业链,锻造 一批"杀手锏"技术,提升产业质量。

推动重点产业链高质量发展,提 升科技创新能力,要加强重点领域关 键核心技术攻关。

创新是新型工业化的根本动力。2023年,工业和信息化部坚持把产业发展基点放在创新上,加强重点领域关键核心技术攻关,产业基础再造工程和重大技术装备攻关工程稳步推进。工业母机、关键软件等重点领域创新实现新突破,C919大型客机投入商业运营,国产首艘大型邮轮"爱达·魔都号"正式命名交付,全球单机容量最大的16兆瓦海上风电机组并网发电,国产ECMO完成注册取证,智能6行采棉机实现量产,

神舟十六号、十七号顺利升空。 "当前,我国的新型工业化推进 和工业体系发展取得了积极成效,上下游的产业链布局进一步协同,存在的主要问题是未能掌握全部核心和源头技术。"中国科学院院士马於光在接受《中国电子报》记者采访时表示,"实现新型工业化,需要对重点产业的核心问题进行分析,并集中力量解决。只有解决了核心问题,产业才能更好地发展。" 马於光建议,一是要选好发力方向,既要面向当前急需解决的问题,又要着眼于长远布局,实现核心技术的突破和积累;二是要集中力量办大事,一方面强化新型举国体制带来的优势,另一方面避免过度投入导致的重复和内卷。

下一步,我们要深入实施创新驱动发展战略,以科技创新推动产业创新,催生新产业、新模式、新动能,发展新质生产力。系统布局重点领域关键核心技术攻关,加快重大项目实施,发挥新型举国体制优势,一体化推进技术攻关、迭代应用、生态培育,推进重大战略性技术和产品攻关突破。

推动重点产业链高质量发展,提 升科技创新能力,要强化企业的科技 创新主体地位。

企业是推进新型工业化的主体力量,企业强工业才能强。2023年,工业和信息化部进一步强化企业科技创新主体地位,整合行业科技力量,完善产业创新生态,打造以大企业为主力军、专精特新中小企业为生力军、大中小企业融通发展、"政产学研用金"深度融合的发展生态。目前,已累计培育制造业单项冠军1187家、专精特新"小巨人"企业1.2万家、专精特新中小企业10.3万家,

一大批优质企业在强链补链中发挥 重要作用。

"制造业的'链主'是在市场竞争中自然形成的,行业'领头羊'的产值、创新能力、研发能力、市场认可度最为关键,所以链主的创新才能带动全产业链的创新。"中国工程院院士谭建荣强调要发挥"链主"的关键作用,"比如主机厂可以推出一代又一代的新产品,配件厂就要根据主机厂创新的产品不断地更换新产品,不断迭代工艺。"

陈晓红也认为,应注重培优塑强 "链主",充分发挥头雁作用,引领关 联企业高质量发展。

下一步,我们要培育更多具有国际竞争力的世界一流企业,壮大专精特新企业群体,让各类企业在新型工业化进程中的创新活力充分涌流、创业激情竞相迸发。要进一步发挥企业作为出题人、答题人和阅卷人的作用,鼓励更多龙头企业发布产业技术创新和配套需求,中小企业"揭榜"攻关,共同攻克产业技术难题,着力构建以企业为主体、市场为导向、产学研用高效协同深度融合的创新体系。

推动重点产业链高质量发展,提 升科技创新能力,要推进园区提质增 效,加快打造世界领先的产业科技创 新高地。

国家高新区是我国重要创新策源地、高成长企业和高端产业集聚的重要载体。2022年,国家高新区生产总值达到17.3万亿元,创造了全国14.3%的GDP,贡献了全国13.6%的税收。2023年1—9月,园区生产总值达到12.33万亿元,同比名义增长

7.11%,新注册企业57.63万家,占全国新注册企业的22%,同比增长18.59%,为实体经济高质量发展持续注入澎湃动力。

工业和信息化部规划司副司长 吴家喜表示,从基地平台看,国家高 新区聚集了近80%的全国重点实验 室、70%的国家制造业创新中心、 78%的国家技术创新中心;从研发 投入看,国家高新区企业研发经费 投入超万亿元,占全国企业研发经 费投入近一半;从成果产出看,国家 高新区企业拥有的发明专利占全国 接近一半,智能机器人、卫星导航等 一批引领性原创成果在高新区加速 产业化。

作为第一个国家高新区,中关村科技园拥有国家高新技术企业1.7万家、上市企业498家。据北京市科学技术委员会、中关村科技园区管理委员会党组成员、副主任张宇蕾介绍,2023年1—10月,中关村规模以上企业总收入6.7万亿元,实现工业总产值9169.6亿元,占北京市47.8%。中关村将按照工业和信息化部统一部署,加快谋划实施重点任务和项目,强化科技创新与产业创新深度融合,努力在推进新型工业化、打造新质生产力方面奋力走在前列。

下一步,我们要发挥科技创新的 引领作用,推动更多国家战略科技力 量和重大创新基地平台向国家高新 区布局,巩固提升特色优势产业,培 育壮大新兴产业,前瞻布局未来产 业,开辟更多新领域新赛道,加快打 造世界领先的科技园区和产业科技 创新高地,进一步支持国家高新区加 快形成新质生产力。