

边缘AI将为MCU带来什么?

本报记者 王信豪

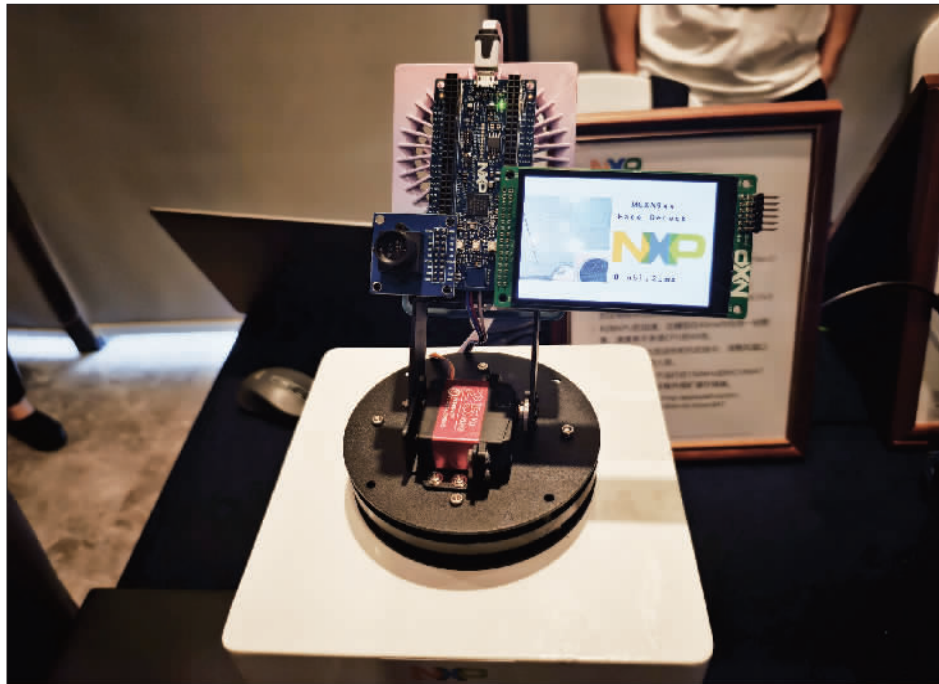
在人工智能和机器学习的推动下,边缘计算的需求正在逐渐增加。相关机构预测,全球被创建、采集和复制的数据将急剧扩张,至2025年达到175ZB(泽字节)。为减轻数据在云端处理的负荷,越来越多的数据也将被置入边缘侧进行运算。作为电子设备的主控制芯片,MCU对于边缘计算设备的数据处理和决策能力提升有着重要作用,将在边缘AI的浪潮中迎来新的发展机遇,同时也面临着更多的挑战。

边缘AI浪潮来袭 拓展MCU市场空间

物联网(IoT)设备、工控、智能家居和智能穿戴等领域的需求,正在带动MCU市场的增长。从长期来看,在保证低功耗、强实时性的前提下,让MCU具备更强的边缘侧计算和智能决策能力成为了下游市场的期许。

在这个过程中,边缘AI成为智能设备发展的重点。边缘AI是在物理世界的设备中部署AI应用程序。之所以称其为“边缘AI”,是因为这种AI计算在靠近数据的位置完成。对比云端运算,边缘AI具备强实时性,数据处理不会因为长途通信而产生延迟,而是能更快响应终端用户的需求,同时也可保证数据的隐私和安全。对比传统边缘计算只能响应预先完成的程序输入这一特性,边缘AI又具备更强的灵活性,从而允许更多样的信号输入(包括文本、语音及多种声光信号等)和针对特定类型任务的智能解决方案。

边缘AI的以上特性,与MCU有着较强的契合度。一方面,MCU具备低功耗、低成本、实时性、开发周期短等特性,适合对成本和功耗敏感的边缘智能设备。另一方面,人工智能算法的融入也能补强MCU,使其兼顾更高性能的数据处理任务。因此,“MCU+边缘AI”正在图像监控、语音识别、健康状况监测等越来越多的领域得到应用。此外,在IoT的基础之上,MCU和边缘AI的结合也将推动AIoT(人工智能物联网)的发



展,使各种设备能够更加智能地互联互通。数据机构Mordor Intelligence预测,边缘人工智能硬件市场在2024至2029年间将以19.85%的年复合增长率增长,2029年将达到75.2亿美元。

“未来的MCU将面向专业化、智能化方向发展。”北京奕斯伟计算技术股份有限公司首席市场官刘帅告诉《中国电子报》记者,“其中智能化体现在两方面,一是强化对AI算法和机器学习模型的支持,使得MCU拥有一定智能决策能力;二是性能提升,高端产品将采用多核设计以提升处理能力,满足高性能的需求。”

集成AI加速器 强化MCU性能

边缘AI为MCU带来了诸多市场机遇,而想要满足智能设备在边缘侧进行人工智能的运算需求,强化MCU的AI性能是重中之重。

“面向边缘AI和端侧AI需求,MCU需要做出以下调整以增强AI计算能力。”兆易创新MCU事业部产品市场总监陈思伟表示,“一是集成AI加速器,如神经网络加速器或者专用的向量处理器,以加速AI推断和训练任务;二是优化能效比,在保持性能的同时降低功耗,延长设备续航时间;三是增强安全性,包括数据加密、安全引导和安全存储,以保护用户数据不受攻击;四是支持多模态感知;五是优化系统集成,提供更多的硬件接口和软件支持,使得开发人员能够更轻松地将AI功能集成到边缘设备中。”

在AI加速器方面,数字信号处理器(DSP)和神经网络处理器(NPU)都成为在MCU中集成的重要加速组件,让MCU能够在边缘运行AI算法。具体而言,DSP更适合信号处理任务,包括音频、视频、通信等,而NPU则更聚焦于高效处理大量的矩阵运算和并行计算任务。

为此,各大厂商积极布局。意法半导体于2023年推出STM32N6,采用Arm Cortex-M55内核,集成ISP和NPU以提供

机器视觉处理能力和AI算法部署。恩智浦推出MCX N系列MCU,具有双核Arm Cortex-M33,并集成了eIQ Neutron NPU,据了解,该NPU可将机器学习推理性能提升约40倍。

作为大多数MCU内核的供应方,Arm也在边缘侧NPU上发力。4月,Arm推出Ethos-U85 NPU,作为一款AI微加速器,其支持Transformer架构和CNN(卷积神经网络),配合Arm v9 Cortex-A CPU可提供4TOPS的端侧算力,助力AI推理。

“许多不同的处理器都可以实现人工智能。AI可以在Arm核中运行,也可以在NPU、DSP当中,不同处理器将会运行不同算法,那么功耗也会不同。”恩智浦执行副总裁兼安全连接边缘业务总经理Rafael Sotomayor表示,“是否在Arm核或者DSP中实现AI功能,这取决于工程团队的技术专长。我们使用NPU来做,是因为(从我们的技术出发)速度更快,而且功耗更小。当然,如果客户对特定机器学习算法有要求,也可以继续使用DSP。”

实现性能功耗成本平衡 构建平台化方案

尽管可以通过集成NPU等AI加速器使MCU支持AI算法,但是高性能往往会带来更高的功耗。因此在提升性能的同时,也要保证MCU在功耗和成本等诸多要素之间达到平衡。这不仅考验MCU厂商的芯片设计能力,也对公司整体成本和功耗优化提出了更高要求。

使用自研NPU成为厂商平衡成本与效能的选择之一。据悉,意法半导体所发布的STM32N6采用了自研的Neural-Art加速器,恩智浦MCX N系列所集成的eIQ Neutron NPU同样为自研。这既有利于降低授权成本,也能保证在自身技术路径下对NPU迭代节奏的合理把控。

同时,内核架构的选项也在增加。当前在MCU市场中,除少部分8位MCU使用CISC架构,Arm架构的Cortex-M系列核由

于功耗表现较好占据主流,而随着RISC-V架构逐渐发展,这一新兴架构也逐渐获得厂商的青睐。据悉,ADI推出的MAX78000/2在集成专用神经网络加速器的基础上,提供了Arm核与RISC-V核两种方案,可在本地以低功耗执行AI处理,最大程度地降低CNN运算的能耗和延迟。

刘帅向记者表示,采用RISC-V架构能够为MCU产品带来更多优势,主要体现在两方面:一方面,灵活可定制,RISC-V作为开源架构,可根据具体应用场景进行深度定制,吸引更多上下游伙伴共同建设生态;另一方面,超低功耗,定制RISC-V架构更加简洁,同类架构功耗最优。

而站在客户的视角,除了需要MCU产品在高性能、低功耗、低成本等多方面达成平衡,也希望MCU厂商能够提供平台化的整套解决方案。

“过去客户的选择是自下而上——先选择芯片,再思考需要何种软件、应用等。现在则是从应用层开始,自上而下到芯片,以获取技术支撑。”Rafael Sotomayor指出,“这些技术非常复杂,不仅涉及人工智能,还涉及信息安全、功能安全、视觉、音频等。因此,帮助客户简化技术的复杂性,成为厂商为客户提供产品和服务的核心价值。”今年4月,恩智浦宣布与英伟达合作,将英伟达TAO工作组件集成在恩智浦eIQ机器学习开发环境中,以便开发者加速开发,并部署经过训练的AI模型。

ADI基于MAX78000/2,提供了开发工具MAX78000EVKIT#,以帮助开发者实现平稳的评估和开发体验。意法半导体同样推出云端开发者平台STM32Cube.AI,支持使用者在云端对已有资源进行配置,进一步降低边缘人工智能技术开发的复杂度。

综合来看,面对边缘AI浪潮所带来的挑战,MCU厂商正在积极探索,并展示出多样化的发展路径。尽管架构及AI加速器等方案的“最优解”目前还未有定数,但是对于MCU而言,在保证低功耗和低成本等基本特性的前提下,不断提升计算能力和安全性来适应越发复杂的边缘计算环境,已是大势所趋。

大力推进现代化产业体系建设 加快发展新质生产力

德州仪器 MSPM0 系列 MCU 的进阶之路

本报记者 姬晓婷

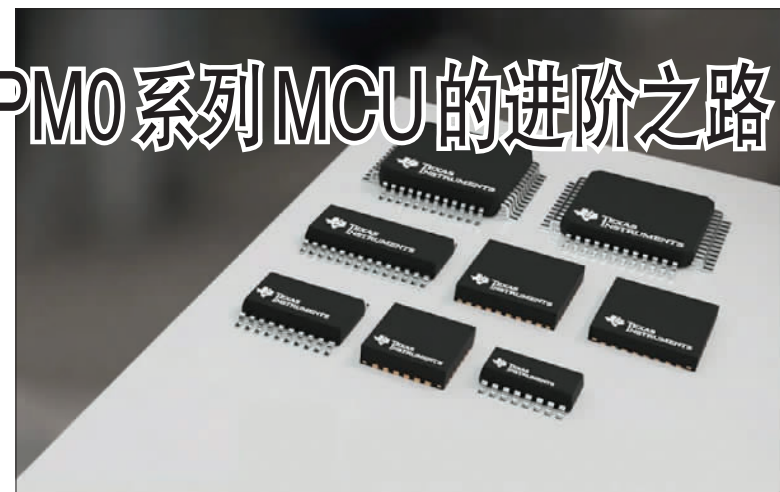
研究机构发布数据称,2023年全球MCU出货量达到320亿片。经历了几十年的迭代与发展,MCU现已在家电、汽车、工业设备等产业链中发挥着不可替代的作用。而在电路和系统设计尺寸越来越小、性能提升的要求越来越迫切的当下,行业工程师仍面临着性能和成本的双重挑战,市面上留存的8位、16位MCU已经很难满足需求,亟须新技术的注入。

德州仪器(TI)的32位MSPM0系列产品就是在这样的背景下推出的,以其丰富的功能和可扩展性,帮助终端产品提高性能、降低功耗、增加集成度和安全性,以适应智能化、物联网和工业控制等诸多应用场景。

庞大产品群和丰富规格

MSPM0 MCU基于Arm Cortex-M0架构设计,在2023年3月发布之时,Arm的M0架构已经推出十余年。“为什么在这个时间节点上推出市场成熟度已经很高的M0系列芯片”,是该系列产品要回答的第一个问题。而今,经过一年多的发展,该系列产品已经通过实践证明了其在通用性、性价比、可拓展性、低功耗、开发友好等方面具备的行业优势,并凭借庞大的产品矩阵,获得来自工业、汽车、消费电子等领域客户的认可。

TI的MSPM0系列MCU在发布之时提供G和L两个子系列:G系列可以实现更快的运算,芯片支持的频率更高;L系列可感应更低的电流,更适合传感器的应用,相应地,其支持的频率更低。该系列产品还提供从8kb到128kb不同规格的闪存容量、从16到64的不同引脚封装选项,为客户提供丰富的可选空间。自MSPM0发布以来,产品线不断扩展,今年,德州仪器更是凭借其深厚的嵌入式技术积累和对8位及16位MCU市场技术挑战的敏



锐洞察,推出了基于创新制造技术和先进工艺节点的MSPM0C110x系列微控制器,该系列产品主打超小型封装,同时保持了极低的成本,在各种空间敏感型的应用场景中展现了独特的优势。

MSPM0共计推出了100余款产品,涵盖不同的封装、引脚数、温度范围及模拟、通信等外设,同时保持同封装产品的引脚兼容。同年12月,德州仪器新发布了符合AEC-Q100标准的MSPM0 MCU,以满足车身电子产品的应用需求。

降低动态功耗与静态功耗

随着物联网、工业、汽车等领域与日俱增的智能化需求,MCU——这类工作时长最久的半导体器件之一,需要不断提升能效表现以应对市场需求。而低功耗恰恰是MSPM0系列产品最突出的优势之一。

MCU的功耗大致分为动态功耗和静态功耗两大类。其中,动态功耗主要是指CPU和大部分外设运行时的电流;静态功耗主要是指CPU和大多数外设器件关闭时模拟电路和数字电路的偏置电流与漏电流。

为了降低功耗,TI选择了“两条腿走路”的方式:在降低动态功耗方面,MSPM0可在完成任务后快速返回低功耗模式。MSPM0提供不同的时钟源、时钟频率和电源域,以实现不同的功耗级别,用户可通过设置产品在不同的功耗模式(运行、睡眠、停止、待机和关断)和三个低功耗模式策略选项(XX0、XX1、

XX2)之间切换,实现更低的动态功耗。在降低静态功耗方面,MSPM0的策略是关闭未使用的模块,并在合适的工作频率和适当的MCU工作模式下使用这些模块。

在上述整体设计思路之下,MSPM0还采用了调整条件代码执行、非阻塞编程、优化代码大小等软件编程策略,并对MCU电源、电阻器、电容器、电源IC等不同部件分别做了降低功耗的调整,从而实现MSPM0系列产品的整体功耗优化。

构建易用的生态

作为一家具有浓厚工程师风格的IC企业,德州仪器在交付产品的同时也非常注重为使用这些MCU的客户打造更舒适的开发体验,帮助下游工程师轻松上手、一键迁移、高效拓展,把更多的时间用在优化和提高自己的系统上。为此,德州仪器为MSPM0客户提供了全栈开发工具。其中包括多种硬件编程工具,三种不同频率的LaunchPad开发套件,优化的软件驱动程序,数百个代码示例的软件支持,以及涵盖汽车、楼宇自动化、功能安全、医疗、电机控制、电源管理等十余个应用场景的库和协议栈。如果遇到开发问题,工程师还可进入MSPM0开发人员专区,查找相应的培训教程。

作为一家老牌半导体公司,德州仪器与全球的生态系统合作伙伴紧密合作,提供一站式的解决方案,为客户从硬件到软件、从开发工具到生产制造的全方位支持。在未来,德州仪器将持续进行MSPM0系列产品技术迭代,为行业伙伴提供源源不断的创新动能。