

存算一体芯片迭代进入冲刺期

长江日报记者杨佳峰

6月20日,华中科技大学信息存储系统教育部重点实验室里,电流嗡嗡声不绝于耳。“喻家山二号”芯片研发进入关键期。

2024年11月底,全球存储容量最大的MRAM(磁性随机存储器)存算一体芯片“喻家山一号”芯片在此诞生。这款由华科大副校长、信息存储系统教育部重点实验室主任冯丹教授带领团队研发的芯片,突破了统治计算机架构70余年的冯·诺依曼技术瓶颈,将“存”与“算”融为一体,通过原位计算大幅降低数据移动能耗,算力高达每秒能做30.1万亿次AI运算,能效比超过高端边缘AI芯片。

冯丹表示,“喻家山二号”芯片将在“喻家山一号”技术基础上迭代升级,进一步提升算力与能效。

“小个子,大能耐”

武汉光电国家研究中心展厅,两块银灰色“喻家山一号”芯片呈现在长江日报记者眼前,单芯片面积不足2厘米见方。

“小个子,能耐大。”实验室童薇教授拿起其中一块,语气里带着自豪,“存储容量64Mb,算力达30.10TOPS,能效140.54TOPS/W。”

这组数字意味着什么?冯丹解释,在传统芯片上,数据必须像搬运工一样,从存储区长途跋涉到计算区,这个过程消耗了90%以上的能量和时间。而“喻家山一号”能让数据“原地办公”,直接在存储单元完成计算。这不仅让它每秒能完成30.1万亿次AI运算,更让每瓦功耗能驱动140.54万亿次运算。简而言之,它不仅算得快,而且还节能。

AI算力的“马力”越大,数值越大,跑大模型、图像识别、自动驾驶等AI任务越快。2024年11月底,冯丹教授拿到团队刚研发完成的“喻家山一号”封装芯片时,迫不及待放到板卡上测试其是否具备设计预期的超凡“马力”。

“当时我们研发了两款芯片,一款基于MRAM介质,另一款基于ReRAM(阻变存储器)介质,没想到两款都能点亮,非常顺利。”童薇介绍,首先确认是否能正常上电,随后测试读写等基本功能,包括存储容量是否达标、计算能力是否正常等。

测试过程中,ReRAM芯片比较顺利,能正常工作,但MRAM芯片连续一个月测试没有任何反应。

“示波器测量不到任何信号,输入激励后也没有输出,虽然能上电,但无法激活。”童薇回忆,那段时间实验室气氛压抑到了极点。冯丹带领团队一遍遍排查,最终发现是芯片的一个引脚定义弄反了,“那一刻,悬着的心才重重落下”。

随着ChatGPT和DeepSeek相继上线,2025年,冯丹决定做一个大胆的尝试:将1.5B(15亿)参数规模的模型部分计算任务,卸载到“喻家山一号”上运行。

“成功了!”测试结果验证了芯片具备支撑大模型推理的系统级能力。冯丹解释:“就像CPU需要编译器将高级语言转化为机器码一样,我们的存算一体芯片也需要一套专用编译工具链,将目标操作转换为芯片可理解的指令。”

“完全符合设计预期!”冯丹带领团队完成了存算一体芯片的设计到软件工具链的全套工作。但是团队却异常平静,第二天的实验室依然安静如昨,因为他们前面还有更大的目标。

“要做就做最好的”

1987年,17岁的冯丹考入华中工学院(现华中科技大学)计算机专业,她首次系统学习了冯·诺依曼架构。那时的她未曾预料到,未来自己会成为这一经典理论的突破者。

2006年,已在校任教十载的冯丹,投身国家973计划“下一代互联网信息存储的组织模式和核心技术研究”。在探索“多层次、可扩展的存储对象”模式的过程中,她认识到冯·诺依曼架构的局限性,脑海中萌生出存算一体的创新构想。

“存算一体的核心,是把计算功能迁移到存储端,让数据在存储位置就地运算,彻底消除搬运开销。”冯丹向记者阐释道。这一构想早在2009年看到曙光——惠普实验室发现忆阻器效应,为存算一体的物理实现奠定了理论基础。存算一体旨在最大限度降低乃至消除数

据传输的能耗。其实质是让数据与计算单元尽可能贴近,理想状态下,数据可在存储位置直接运算,无需任何移动。

2014年,他们研发的“主动对象海量存储系统及关键技术”获得国家技术发明奖二等奖。这项技术让存储系统具备了一定的处理能力,被视为存算一体的早期雏形。

“距离实现存算一体芯片的目标近在咫尺,我们必须突破这一关键瓶颈。”凝视着实验室陈列的主动对象存储卡,冯丹为团队定下了新的攻关方向。

研发之路布满荆棘。当时,存算一体芯片主要有两条技术路线:SRAM(静态随机存储器)和MRAM(磁性随机存储器)。国内同行多选择SRAM路线,因为更易适配现有工艺。但SRAM断电数据即失,且存储密度受限,难以承载大模型时代对海量数据的需求。

“要做就做最好的!”冯丹毅然选择了更难走的路——MRAM非易失性存储技术。这种技术兼具断电数据保存、读写速度快、功耗低、寿命长、抗干扰强等优势,更适配存算一体芯片的应用需求,是未来高端存储芯片的核心方向。

从材料特性研究到单元结构设计,再到阵列集成与算法优化,这支四五十人的团队埋头苦干了10年。“很多人的黑发熬成了白发。”童薇感慨道。正是这份坚持,让“喻家山一号”在算力与能效上达到全球领先水平,一举打破国外对高端存储技术的长期垄断。

“要打破生态壁垒”

“喻家山一号”问世后,冯丹团队马不停蹄地推进产业化进程,他们对接产业链上下游企业,开展应用测试与市场推广,致力于让顶尖技术走出实验室,赋能千行百业。

“存算一体技术短期内必将大规模应用。”童薇描述了应用前景——该芯片采用异构计算架构,并非要取代CPU或GPU,而是将基础CPU改造为更适应智能处理的计算单元。通过在存算一体芯片外围集成更多逻辑模块,可将其打造为功能全面的通用芯片。

目前,团队正与多家行业领军企业合作,推进芯片的小批量试产与场景化应用,着力解决工艺适配、成本优化及批量生产等关键问题,加速实现商业化落地。“AI发展的主要瓶颈是能耗,而存算一体技术凭借其高能效比,必将在各层级应用中大放异彩。”冯丹表示。

但在落地过程中,一道无形的“生态壁垒”横亘在前。长期以来,全球集成电路技术标准多由国外主导,我国产业发展受制于人。

“技术创新离不开标准引领。”冯丹深知,要实现存算一体规模化发展,必须建立自主可控的标准体系。去年,她牵头组织高校、科研院所及龙头企业,启动存算一体技术国家标准的制定工作。项目组系统梳理了技术架构、性能指标、测试方法等核心要素,形成了完整科学的国标草案。

目前,团队已正式提交国家标准申报材料。冯丹表示,这将彻底改变国内产业“无标可依”的局面,破除技术生态壁垒,构建自主可控的存算一体产业生态。

“喻家山一号”芯片问世仅一年,“喻家山二号”芯片的研发已进入冲刺阶段。冯丹透露,“喻家山二号”芯片将在现有技术基础上,进一步提升算力与能效,优化存储容量与集成度,拓展更多应用场景,攻克更高精度的原位计算、多场景适配等核心技术,推动存算一体技术向更高水平、更宽领域迈进。



华科大信息存储系统教育部重点实验室存算一体芯片研发团队,前排右四为冯丹教授,前排右五为童薇教授。(资料照片)

“喻家山一号”存算一体芯片。

存算芯片板卡。

“喻家山一号”流片。

对话

记者:存算一体到底是什么?怎么去理解?

冯丹:最开始的核心概念是“原位计算”,也就是在数据存储的位置直接进行计算,数据不需要移动。比如说,传统冯·诺依曼架构中,CPU是一个中心处理器,所有操作数(比如做加减乘除时需要的数值)都要从其他地方读取到CPU的寄存器里才能运算。这就像是处理单元定在那不动,数据得一个一个运过来才能处理。

计算本身的能耗其实很低,可能只占整个计算过程能耗的十分之一,甚至更低。大量的能量都消耗在数据从存储位置到处理器的搬运上,而且搬运还会有延迟。比如做加法运算,其逻辑运算的能耗可能在皮焦级别,而数据搬运的能耗要高得多。这就是大家常说的“冯·诺依曼瓶颈”。

存算一体解决的核心问题就是尽可能降低甚至消除数据移动的开销,让数据和计算单元尽可能靠近,极致状态就是数据在存储位置直接运算,完全不需要移动。

记者:存算一体为何如此重要?

冯丹:眼下,我国全面实施“人工智能+”行动,数据存储事关国家战略。未来,不仅需要更强的算力,更需要与之匹配的存力。存算一体技术相比传统的存储分离架构,能效比提升一个数量级以上,是突破现有算力瓶颈,实现绿色低碳发展的颠覆性技术路径。

记者:“喻家山一号”存算一体芯片性能如何?它是如何工作的?

冯丹:我们研发的存算一体芯片,是目前全球存储容量最大的MRAM存算一体芯片。当我们使用一台电脑时,系统会将存储芯片里的数据搬运到计算芯片中。七八成的电量耗在了搬运途中。存算一体芯片的厉害之处是数据存在哪儿,就在哪儿算,不用搬来搬去,能节省大量能耗。耗电量是同规格计算芯片的千分之一。

记者:存算一体芯片能做哪些事情?

冯丹:存算一体芯片不仅可以应用于手机端,也可支持图像识别、声纹识别、指纹识别等场景。其核心优势是能耗更低:不仅能延长待机时间(充一次电可用两三天甚至更久),还能减少设备发热,提升使用体验。

具体场景包括所有采用RISC-V处理器的设备(机顶盒、智能电视、智能手机、电脑)、物联网设备(家庭机器人、摄像头),以及带智能标签的家电——它能提供更强大的智能算力。就拿家用摄像头来说,用上存算一体芯片后,在本地完成数据处理,不需要把实时画面传到云端处理,既安全又高效。

记者:存算一体芯片在国际上研发现状如何?

冯丹:目前,存算一体芯片的研发,国际和国内几乎齐头并进,台积电、三星等在产业化落地上领先。联发科去年也推出过存算一体芯片,采用的是SRAM(静态随机存储),跟我们MRAM路径不一样:SRAM更容易适配现有工艺,而MRAM需要专门的后道生产线实现,优势是存储密度更高,算力也能更高,更适合大模型时代的需求。此外,武汉也有相关企业在研发存算一体芯片,产业集聚效应正在形成。

遗失声明 湖北聚鑫茂建筑劳务有限公司遗失财务章一枚,特此声明作废。

第六季“泰康一夏”送健康福利

与家长一起守护孩子的眼睛与口腔健康

暑期来临,正是青少年关注身心健康、提升生活品质的黄金时期。武汉大学泰康临床学院、泰康同济(武汉)医院(以下简称泰康同济)正式启动第六季“泰康一夏”暑期公益活动,为广大读者朋友送出健康福利。本次活动中,大型三甲医院口腔科与眼科携手发力,针对青少年牙齿矫正、近视矫正等热门需求,推出专业诊疗服务与惠民福利,为广大青少年的健康成长保驾护航。

暑期是牙齿矫正黄金时间 但选择机构时务必谨慎

暑期历来是青少年牙齿矫正的高峰期。对于许多家长和孩子来说,利用假期时间完成牙齿矫正,既不影响学业,又有充足的恢复时间。泰康同济口腔科以三甲医院的技术保障与贴心服务,为广大青少年的正畸之路保驾护航。

14岁的小诗(化名)曾因龅牙、深覆合问题在外院开始正畸治疗,然而一年半过去,牙齿前突和咬合问题不仅没有解决,原本就医的口腔机构还突然倒闭,托槽粘接不规范、咬合错乱等问题让家长忧心忡忡。

来到泰康同济口腔科后,专家团队重新评估病情,为她制定了全新的治疗方案:拆除不符合规范的托槽重新粘接,去除不合理的咬合垫高材料,先解决深覆合问题,再在上颌前牙区植入两颗支抗钉内收拢牙。经过一年的系统治疗,小诗不仅拥有了整齐的牙齿,脸型也得到了明显改善,终于绽放出自信灿烂的笑容。

口腔科主任雷敬提醒,暑期确实是牙齿矫正的黄金时间,但矫正周期长、技术要求高,家长选择机构时务必谨慎。医生的专业水平、机构的稳定性都直接影响矫正效果。泰康同济作为大型三甲医院,既有过硬的技术保障,又有优质便捷的医疗服务,是正畸

治疗的可靠选择。

近视手术并非“想做就能做” 术前检查是关键环节

对于即将步入大学的青年来说,摆脱眼镜束缚、拥有清晰视力是一个美好心愿。暑期时间充裕、术后恢复便利,正是进行近视矫正手术的最佳时机。泰康同济眼科屈光专科以先进的设备与精湛的技术,帮助广大青年顺利摘镜。

18岁的小卓(化名)梦想报考的大学与专业对视力有严格限制,然而250度近视加150度散光成为他梦想实现道路上的障碍。为此,他来到泰康同济寻求摘镜方案。

眼科副主任医师成琼首先为小卓安排了全套术前专项检查,结果显示他的角膜厚度、眼底状况、眼轴长度等各项指标均达到手术标准。综合考虑小卓的情况,成琼为他

选择了全飞秒SMILE手术——无需制作角膜瓣,仅通过2至4微米微小切口完成透镜取出,角膜完整性更好,术后恢复快、稳定性高,特别适合运动爱好者和特殊志愿报考考生。手术全程微创无痛,术后小卓成功摆脱眼镜,视力恢复到理想水平,向着理想大学迈进了一大步。

眼科主任陈中山特别提醒:近视手术并非“想做就能做”,术前检查是关键环节;一定要选择正规医院和经验丰富的医生;术后遵医嘱用药、定期复查、合理用眼,才能确保视力稳定恢复。

多科室联袂打造暑期健康盛宴 推出一站式优质诊疗服务

作为泰康同济一年一度的品牌公益活动,第六季“泰康一夏”暑期公益活动汇聚了口腔科、眼科、皮肤科、整形美容外科等20

余个优势科室,针对不同人群的暑期健康需求,打造一站式健康福利平台。

活动期间,口腔科为青少年正畸患者提供专属优惠方案,无论是传统金属牙套矫正还是美观舒适的隐形矫正,都能享受活动专属福利;眼科屈光专科推出近视手术专项优惠,为高中毕业生、大学生群体量身定制摘镜方案。此外,各参与科室还将提供专家号源优先预约、免费健康咨询等多项便民服务,让市民在暑期就能享受到优质、实惠的医疗服务。

泰康同济始终坚持以患者为中心,用专业的医疗技术和温暖的人文关怀,守护每一位市民的健康。第六季“泰康一夏”暑期公益活动,将以实实在在的福利与优质的诊疗服务,陪伴大家度过一个健康、美好的夏天。

欲了解活动详情,可关注“泰康同济(武汉)医院”官方微信公众号,回复“暑期”获取最新活动资讯。(文/张剑 王慧文)