

大数据战略与双碳战略双轮驱动 加快大数据存储技术创新和产业变革——《“十四五”大数据产业发展规划》解读

北京邮电大学数据科学中心主任 刘军

2021—2025 年是我国国民经济和社会发展第十四个五年规划时期,也是“两个一百年”奋斗目标的历史交汇期,对我国国民经济和社会发展具有重要的意义。在这承前启后的重要时期,各部门都以习近平新时代中国特色社会主义思想为指导,推出“十四五”规划,明确“十四五”期间的发展方向 and 路径。工信部印发了《“十四五”大数据产业发展规划》,围绕“十四五”时期大数据产业发展的总体要求、主要任务和保障措施作出全面部署。

随着我国信息经济发展,数据作为新型生产要素已经成为重要的战略资源,数据所蕴藏的价值逐渐释放。《“十四五”大数据产业发展规划》明确了充分激发数据要素价值潜能,夯实产业发展基础,构建稳定高效产业链的发展方向,同时提出重点提升数据生成、采集、存储、加工、分析、安全与隐私保护等通用技术水平。“十四五”期间,要加大数据通用技术创新力度,推动数据处理各环节的技术发展,赋能大数据产业,支撑大数据产业蓬勃发展。在数据处理各环节中,数据采集能力的提升、数据生成速度的加快、数据加工分析的多样化都需要强大的数据存储能力做支撑,因此数据存储需求在“十四五”期间将呈现加速增长的态势。只有保证海量数据存储的高效安全,才能加强数据资源整合,提高全要素生产率,支撑我国社会经济数字化建设。在数据存储技术上实现突破,是落实《“十四五”大数据产业发展规划》的关键之一。

数据存储技术创新的驱动力不仅仅来自大数据产业自身发展的需求,节能环保也是数据存储必须达到的目标。以往数据中心所导致的高能耗与高成本等问题很大程度来源

- 数据作为新型生产要素已经成为重要的战略资源。
- 优化存储密度和提升性能是数据存储系统技术创新的主要方向之一。
- 应培养基础软硬件的底层支撑能力,用“揭榜挂帅”等方式激励创新。

于数据存储所消耗的能源。为实现我国双碳目标,数据存储技术也开启了一场浩浩荡荡的绿色革命。国家发展改革委印发了《“十四五”循环经济发展规划》,为“十四五”时期我国循环经济发展提供了指引,对加快促进我国发展方式绿色转型、实现资源高效利用和循环利用、推动碳达峰碳中和具有重要意义。因此,如何兼顾双碳目标和大数据战略,推动数据存储技术变革,成为当前所面临的重要挑战。

优化存储密度和提升性能同样是数据存储系统技术创新的主要方向。《“十四五”大数据产业发展规划》推动数据集约化整合和高效化利用,释放数据红利。集约化整合意味着同样规模的数据要发掘更大的价值。这要求数据必须得到妥善的存储,以便于持续或者反复利用,不断与多源数据进行关联,充分发挥不同类型数据的短期和长期价值。大量的数据需要长期保留,从而数据存储密度成为技术优化的方向。同样,高效化利用意味着大数据计算的效率要进一步提高,对于存储数据的检索和读取速度,将成为计算效率的决定因素之一。因此,提升

数据存储在空间利用率和读取效率,可以放大大数据效益。

可扩展性也是数据存储技术创新的一个基线要求。数据正以指数级的速度增长,存储系统必须跟上数据增长的速度,需要不断扩展容量。但是,可扩展性并不仅仅是简单地增加存储设备,要考虑到扩充设备对整体系统可靠性的影响,解决“冷”、“热”数据的分级存储、处理多种类型不同大小的文件、查询检索的最优化、元数据管理等很多伴生的问题。只有实现高效、灵活的数据存储可扩展,才能加快《“十四五”大数据产业发展规划》所提到的“大数据容量大、类型多、速度快、精度高、价值高等特性优势转化”,推动大数据产业高质量发展。

综上所述,为了更好地实现《“十四五”大数据产业发展规划》,数据存储领域以下工作将成为重点:

一是明确数据存储技术的发展目标,通过多元化的技术路线形成数据存储体系的顶层设计。充分认识各种数据存储技术的优势和缺点,基于当前数据规模和数据增速预期,针对不同类型数据的访问频次和读写操

作特征,选择块存储、文件存储、对象存储技术路线,同时综合考虑规模、成本、能耗、寿命、可靠性、可扩展性等多方面因素,积极探索在保证读取速率的前提下最大程度降低能源消耗的软硬件技术和产品。在架构上,除数据中心外,边缘数据中心、高性能计算中心、智能计算中心等正成为数据存储的重要发力者,构建分层分类的数据存储体系能够提升数据存储安全性、完备性和持久性,为数字经济持续稳定发展提供有力支撑。

二是推动数据存储技术的自主创新。培养基础软硬件的底层支撑能力,用“揭榜挂帅”等方式激励创新,突破瓶颈,实现数据存储核心技术的自主可控。鼓励政府产业基金、创业基金及社会资本积极参与创新项目,培养创新型人才,鼓励有能力的企业参与算法、基础设备及软件的研发,推动自主开源框架、组件和工具的研发,促进前沿领域技术的借鉴和融合,推动大数据技术与新一代信息技术集成,对大数据产业进行深层次的升级改造,夯实产业发展基础,构建稳定高效产业链。

三是形成配套的指标体系,验证和评估

大算力芯片量产之战

(上接第1版)领先的车企开始统一部署车端、云端、边端和终端协同计算能力,以此建立高效的智能驾驶运行和迭代体系。

杨宇欣给出了一组数据:2014—2016 年特斯拉 ModelS 的算力为 0.256TOPS,2017 年蔚来 ES8 的算力是 2.5TOPS,2019 年特斯拉 Model3 算力为 144TOPS,2021 年智己 L71070TOPS,2022 年 蔚来 ET7 是 1016 TOPS。这组数据进一步印证了这样一个事实:智能驾驶每前进一小步,后面都需要算力前进一大步。

但目前自动驾驶芯片尚不能满足车企快速增长的算力需求,最近还出现了整车厂主动约请自动驾驶芯片公司,“甲方”与“乙方”之间的姿态有了微妙变化。

从预测来看,2022—2023 年,算力超过 100TOPS 的芯片将量产装车;2025 年,500TOPS 的芯片将量产装车;而 2030 年,超 1000TOPS 算力的芯片才普及。

芯驰科技自动驾驶负责人陶圣给出了相同的判断:“我们认为,2023 年将是 L3 自动驾驶芯片的量产年代,2025 年则将是 L4 规模研发的时代。”

所以不断推高自动驾驶芯片算力,就成为汽车芯片公司追逐的珠穆朗玛峰。

大算力车芯之难

既然自动驾驶大算力芯片是“车芯”竞赛的珠穆朗玛峰,那在设计定义芯片的时候把参数定得高一点,不就能跑在时间前面,跑在对手前面了吗?

其实没那么容易,因为有一系列的挑战横在汽车自动驾驶大算力芯片的面前。

王平曾透露,大算力智能驾驶芯片面临四大挑战。一是芯片系统架构的挑战。200TOPS 以上的芯片对于访存能力的要求非常高,需要支持更高的带宽,这带来系统架构设计复杂度的大幅度提升。二是通用 AI 软件栈的挑战。智能驾驶的算法目前还处于不断演变的过程,激光点云算法和多传感器融合算法也还在快速迭代中,所以不断变化的算法需求需要借由 OTA 以通用的硬件架构和软件栈来支持算法不断升级。三是大尺寸芯片工程的挑战。大算力芯片的尺寸非常大,在封装、电源和热管理、成本控制、良率等方面都存在严峻的挑战。四是先进工艺平台的挑战。大算力芯片需要 7nm 甚至 5nm 先进工艺,只有先进工艺才能做到更高集成度、更低能耗。

“汽车芯片之所以难做,除了算力的竞赛,还需要处理门类众多的数据,感知各种各样的系统,涉及激光雷达、超声波雷达、毫米波雷达、摄像头等各种各样的运算,芯片的参数非常复杂,除了车本身

功能安全问题之外,它还要把整个系统里面的路径规划、决策、控制这些算法都放在芯片上面,所以对 CPU、传感器融合以及路径规划算法等都有较高要求,同时还必须考虑降低功耗和易用性等。”单记章说。

车规级汽车芯片和消费电子芯片有很大不同,芯驰科技创始人兼 CEO 仇雨菁此前曾表示,与消费级芯片注重性能和算力且快速迭代不同,汽车芯片生命周期长,事关人的生命安全,其安全性、可靠性、耐用性更为重要。

杨欣宇认为,自动驾驶芯片目前有两个方面的难点需要突破:第一方面是芯片核心的算力需要突破。这不是指 TOPS 算力,而是芯片的 CPU 算力,TOPS 算力是用于深度神经网络加速的。如何把强大的性能带到在终端领域对性能需求最高的汽车上,是难点。第二方面是如何在实现如此复杂功能和强大性能的前提之下达到车规标准。“芯片定义的时候一定要考虑到 5~10 年之后的功能需求,因为芯片可能会 5 年之后上车,供货周期可能要 5~10 年。”杨欣宇说。

正因为难,才有机会,因为智能汽车时代才刚刚开始,无论是汽车芯片企业,还是计算芯片企业,大家都在同一起跑线上,“就像 PC 与互联网时代诞生了英特尔、移动互联网时代诞生了高通一样,在智能汽车时代有机会同样有机会诞生新的芯片巨头。”仇雨菁认为。

国际巨头抢量产头把交椅

正因智能汽车对于计算的诉求如此高歌猛进,正因无论传统汽车芯片还是计算芯片企业都处于同一起跑线上,所以越来越多的芯片巨头加入了自动驾驶芯片市场。进入 2022 年,自动驾驶大算力芯片市场的关键词是“量产上车”。

中国汽车工程学会发布的《2022 年度中国汽车技术趋势报告》预测,2022 年中国汽车市场的重要趋势之一是大算力芯片量产装车。中国汽车工程学会副秘书长侯福深曾在 2021 年中国汽车工程学会的年会上表示,当前,自主车规级芯片已形成面向 ADAS/智能座舱等功能的批量应用,大算力车规级计算芯片(单芯片算力大于 100TOPS)正在开展测试试验。作为高度自动驾驶汽车“大脑”的核心部件,100TOPS(处理器运算能力单位)以上车规级计算芯片将在 2022 年实现量产并装车。

英特尔旗下的 Mobileye 算是“老牌”高级辅助驾驶芯片公司,目前其辅助驾驶芯片出货量已经超过 1 亿片,在不久前举行的

2022CES 上,Mobileye 宣布推出专为自动驾驶打造的 EyeQ Ultra 系统集成芯片, EyeQUltra 的算力为 176TOPS,预计将于 2023 年年底供货,2025 年全面实现车规级量产。

与此同时,美国的另一家汽车芯片公司安霸也在 CES 上宣布推出 AI 域控制器芯片 CV3 系列汽车专用 SoC,采用了 5nm 制程,16 个 Arm Cortex-A78AE CPU 内核,单芯片 AI 算力达到 500 eTOPS(eTOPS 中的“e”指的是“Equivalent”,意思是等效算力),预计今年上半年将为开发者提供评估样片。

不管是 Mobileye 规划的 176TOPS,还是安霸强调调算法的等效算力 500 eTOPS,他们描绘得再好,距离量产面市都尚需时日。而比 Mobileye、安霸起得更早的英伟达和高通,经过几年卧薪尝胆,其大算力汽车芯片都将在今年进入量产。

高通通过给汽车提供调制解调器芯片、智能座舱芯片等多年历练,两年前推出了 Snapdragon Ride 自动驾驶解决方案,可为 L1/L2、L2+/L3、L4 等不同的自动驾驶系统提供从 10TOPS 到超过 700TOPS 的算力。目前,面向 L2+ 到 L4 级别的基于 Snapdrag-onRide 的 SoC 和加速器芯片已出样,预计在 2022 年开始搭载于长城和通用汽车的量产车型中。

英伟达历时 4 年投入数十亿美元打造的自动驾驶芯片 Orin SoC 也将于今年量产。Orin 系列芯片可实现 254TOPS 的运算性能,能够为自动驾驶功能、置信视图、数字集群、车载信息娱乐以及乘客与 AI 的交互提供算力支持。搭载 Orin 的车型可通过 OTA 软件更新实现自动驾驶能力在汽车全生命周期内的持续升级。据英伟达透露,蔚来、智己、沃尔沃等多个品牌采用的 Orin 芯片都将在今年量产上车。

是骡子是马,量产来看看。谁都可以描绘蓝图,但是从蓝图到流片到上市再到量产,中间会不会跳票,会不会延迟,都面临诸多不确定性。

今年英伟达和高通的大算力芯片(100TOPS 以上)将量产上车,这意味着大算力芯片竞赛的第一回合,至少在时间抢先性上这两家已经胜出。

之后的赛程,有后来居上者吗?事实上,除了第一回合,英伟达还为其后续发展布下了“护城河”。2021 年的 GTC 大会上,英伟达公布了单颗算力达到 1000TOPS 的下一代自动驾驶 SoC Atlan 芯片规划,比上一代 Orin 芯片算力提升近 4 倍,将于 2023 年向开发者提供样品,2025 年量产装车。这样看来,无论是眼下,还是未来,英伟达都处于暂无对手的位置上。

国内车芯争相角逐大算力

在国际市场上,高通、英伟达、Mobileye 咬得很死。与此同时,在中国市场,包括华为、黑芝麻智能、地平线、寒武纪、芯驰科技等在内的几家汽车芯片企业也同样步步紧逼,自动驾驶大算力芯片市场火药味越来越浓。因为未来两到三年将是自动驾驶芯片量产上车的关键节点,而去年和今年则会是大算力自动驾驶芯片拿到项目定点的关键时刻。

2021 年是中国汽车大算力芯片的“曝光年”。在 2021 年上海车展上,黑芝麻智能发布“华山二号”A1000Pro 芯片,其中 DynamicAINN 神经网络处理器的算力达到 106TOPS(INT8)或 196TOPS(INT4)。2021 年 7 月,地平线发布了全场景整车智能中央计算芯片——地平线征程 5,单颗芯片最大算力为 128TOPS。在 2021 世界人工智能大会上,寒武纪 CEO 陈天石披露了正在研发中的行歌车载智能芯片的数据:超 200TOPS AI 性能、7nm 制程、车规级。华为因为众所周知的原因,其没有刻意提及芯片能力,在去年上海车展上,华为展示了赋能智能汽车的全栈能力,推出的多款自动驾驶计算平台 MDC 可以搭配多种传感器,适用于更多车型。

华为没有刻意提及单颗汽车芯片参数,寒武纪的大算力芯片并没有给出上市的时间表,芯驰科技的自动驾驶大算力芯片到今年第四季度发布,目前看只有黑芝麻智能与地平线的大算力芯片有望今年量产上车。基于这样的市场状态,让黑芝麻智能与地平线竞争的隔空口水战此起彼伏。

在黑芝麻看来,地平线是一家算法芯片公司,其强调的是算法。此前杨欣宇曾表示,决定自动驾驶能力其实是综合能力的体现,这既需要算法、操作系统,也需要芯片,但真正从技术源头来讲,推动自动驾驶发展的一定是芯片。“所有电子行业的发展都是从硬件先开始的,因为芯片决定了整个自动驾驶性能和功能的边界。如果硬件上不能支持的东西,软件是怎么也实现不了的,这是技术规律。”杨宇欣说,“目前能够给自动驾驶提供大算力芯片的只有英伟达、高通,而黑芝麻在全球第一阵营。”

研究 AI 算法出身的地平线创始人兼 CEO 余凯则认为,算力也并不代表汽车智能芯片的“真实性能”,芯片计算效率也同样需要关注。正如对于汽车来说,马力不如百公里加速时间更真实地反映整车动力的性能,算力并不反映汽车智能芯片的实际性能,而每秒准确识别帧率 MAPS 才是更

数据存储技术与产品的效果。坚持标准先行,明确数据存储系统的核心指标和基线要求,积极开展标准编制工作,并推动标准落地应用。围绕数据存储技术和产品的功能性、行业应用、节能环保等需求,聚焦低功耗、高性能、可扩展等重点方向,加快关键标准研制。建立标准验证和试点示范流程,培育标准符合性评测服务能力,推动以数据存储标准为重要组成部分的大数据技术标准体系的落地实施。

四是对符合标准的、技术先进的数据存储方案和产品进行示范性应用推广。《“十四五”大数据产业发展规划》明确了适度超前统筹推进新型基础设施;全国一体化大数据中心体系统筹推进,高技术、高算力、高效能、高安全数据中心加快布局建设。适度超前的部署数据中心,必然需要建设数据存储能力,因此,要将经过验证符合标准的、技术先进的数据存储方案和产品率先应用到新型数据中心中。在大数据行业应用方面,工业大数据平台、金融大数据平台等已释放出卓著能效,应积极在数据容量大、类型多、大数据技术应用基础好的消费领域与工业领域推广创新型数据存储技术,发挥示范效应,拉动大数据技术服务更多的行业。

面对世界百年未有之大变局和新一轮科技革命和产业变革深入发展的机遇,我国适时出台的《“十四五”大数据产业发展规划》,开启了大数据产业创新发展新赛道,聚力数据要素多重价值挖掘,抢占了大数据产业发展制高点。各项支撑大数据产业的通用技术应基于面向未来的动向和战略态势研究,统筹发展,系统布局,加强技术、产品和服务协同,全面发力,为打造数字经济发展新优势、建设数字中国奠定坚实的基础。

真实的性能指标。“在算法驱动 AI 芯片设计的软硬结合趋势下,新一代汽车智能芯片的领导者,将会是世界级 AI 算法公司。”余凯说。

关于算法、数据和软件,黑芝麻智能并不示弱。黑芝麻智能应用工程副总裁邓堃表示,黑芝麻智能目前已经打造了丰富的算法模型,搭建了资深的技术团队和专业的研发体系,推出的山海人工智能开发平台拥有 50 多种 AI 参考模型库转换用例,降低客户的算法开发门槛;能够实现 QAT 和训练后量化的综合优化,保障算法模型精度;支持动态异构多核任务分配,同时还支持客户自定义算子开发,完善的工具链开发包及应用支持,能够助力客户快速移植模型和部署落地的一体化流程。数据方面,黑芝麻智能能够向客户提供面向数据安全的数据闭环体系,确保数据安全可控。软件方面,目前,黑芝麻智能已经能够为客户提供从开发到量产的软硬件全解耦解决方案。

黑芝麻智能与地平线明里暗地较劲很正常,谁都希望成为大算力芯片量产上车的“一哥”。据业内人士透露,地平线的征程 5 将在今年首先搭载在哪吒汽车上,此外上汽集团、岚图汽车等也是意向客户,而黑芝麻的首款大算力芯片的量产上车也将是国内自主品牌,但未透露客户名字。

自动驾驶大算力芯片竞争的大幕刚刚拉开,究竟是“算法派”更有优势还是强调硬核能力的“芯片派”更能赢得用户,现在市场和用户都还没有做出选择,因为大家都还没有量产。目前汽车大算力芯片玩家们是“八仙过海”各有各的招数,事实上,英伟达做大算力汽车芯片也是在四年前,高通是在两年前,“老牌”辅助驾驶芯片厂商 Mobileye 的大算力芯片也要明年才出来。这样看起来,无论是成立 5 年的黑芝麻,还是成立 6 年的地平线、寒武纪,或是成立 3 年的芯驰科技,都是前后脚踏到了机会的时间窗口上,目前看地平线与黑芝麻将最先跑完第一棒。

未来究竟谁能够做得“更快、更高、更强”,谁的路径更优,每一家公司有每一家的基因,每一家有每一家的优势。但有一点可以肯定,打造高级别自动驾驶大算力芯片属于跨界融合,既要懂汽车又要懂计算,无论是传统的计算芯片公司,还是传统的汽车芯片公司,都需要补齐短板,所以现在大家彼此并不是对方的“敌人”,真正的敌人是自己——是否能够做出满足用户需求的大算力芯片,技术上是否过硬,是否能够赢得生态和用户的选择,“朋友圈”是不是够给力,才是硬道理。在这场与智能汽车产业同步迭代的大演进中,我们期望中国的创新企业当中能够成长出“小巨人”。