

来自 2024 中国算力大会的呼声

打造更多“撒手铜”级算力产业应用场景

► 本报记者 李洋

“全国在用算力中心机架总规模超过 830 万标准机架,算力总规模达 246 EFLOPS(每秒百亿亿次浮点运算),位居世界前列。”9 月 27—29 日,以“聚智向新 算领未来”为主题的 2024 中国算力大会在郑州市举行,工业和信息化部总工程师赵志国在大会开幕式上表示。

赵志国强调,目前,城域算力节点间 1ms 时延保障能力已逐渐形成,区域集群到周边主要城市间 5ms 时延保障能力广泛覆盖,国家枢纽节点间 20ms 时延保障能力全面实现;全国算力中心平均电能利用效率(PUE)降至 1.47,创建国家绿色数据中心 246 个;工业、教育、医疗、能源等多个领域算力应用项目超过 1.3 万个,为在更大范围、更高水平上共享数字化发展成果奠定了坚实基础。

“以算筑基”是大势所趋

目前,以数字经济为代表的新经济形态日益成为我国经济体系结构优化、换挡升级的重要引擎,迫切需要以算力设施为代表的新型数字信息基础设施使其畅通经济循环的“大动脉”。数据显示,截至今年 6 月,我国在用算力中心超过 830 万标准机架,算力规模达 246EFLOPS,智算同比增速超过 65%。

我国已建成全球规模最大、技术领先的“双千兆”网络基础设施,骨干网络综合承载能力跃升至新高度,超低损耗纤、灵活全光交叉和智能管控等新技术实现大范围应用,启动省际干线超长距 400Gbps 光传输系统的线网部署,云网、算网协同承载能力不断提升。

中国联通集团公司副总经理郝立谦表示,依托“算网一体”资源禀赋,加速推进自动标识与数据采集(AIDC)建设升级,在数据中心“5+4+31+X”基础上打造“1+N+X”智算集群,覆盖东数西算所有枢纽节点和重点城市,我国正不断创新智能算力供给,打造全面覆盖、适智升级、高效节能、智慧安全的智算基础设施新格局。

中国电信集团公司副总经理栾晓维表示,中国电信打造了业界领先的“息壤”算力平台调度,突破公有云万卡纳管和调度核心技术,研发算力插件实



图片来源:主办方供图



以“聚智向新 算领未来”为主题的 2024 中国算力大会在郑州举行,图为大会现场。

现异构算力接入标准化、研发算力网关,支持社会算力云化纳管安全接入,同时建立算力度量体系,使得异构算力可统一衡量。目前“息壤”已接入 39 家算力合作伙伴,聚合各类智能算力合计达到 22EFLOPS。

中国移动通信集团公司副总经理李慧镭表示,中国移动实现全球规模最大的“IPv6+”骨干承载网、全球首个 400G 全光骨干传送网规模商用,基于云间数据备份场景,完成 5500 公里超长距高吞吐试验,算力路由、全向智感互联、全调度以太网等原创算网技术取得突破,自研算网大脑在全网推广。

“全国在用算力中心业务的 58.3% 已连接国家骨干网,其中大型、超大型算力中心连接比例为 45%,规划在建算力中心预计有 64%连接骨干网。”中国信息通信研究院院长余晓晖说,“接下来,需要加快软件产业协同,打造开放兼容的智能体系。”

“以网强算”是关键路径

随着算力形态从 PC 到云再到算力网络的转变,网络的支撑作用日益凸显,从传送信息到连接算力,进一步升级为感知调动算力,呈现显著的网络效益。依托我国高速、移动、安全、泛在的网络优势,将有效提升数据汇聚、处理、流通、交易效率,助力我国从算力大国向算力强国加速迈进。

去年,国家超算互联网核心节点在郑州市启动建设。该节点立足河南省、面向全国,承担国家超算互联网运营、服务和资源调度等核心枢纽功能,成为一个多维度服务的综合平台。

国家高性能计算机工程技术研究中心主任厉军表示,国家超算互联网的建成,将使分布在全国的超级计算机和智能计算设备形成一个高效运行的网络。目前,该平台已经汇聚 1000 多种应用,覆盖 100 多个行业领域,为 6300 多

种商品和 4300 万行的源代码提供服务,未来目标是接入 500 家服务商和 1 万种商品,达到 20 万用户的规模。

中国工程院院士刘韵洁则认为,确定性网络是中国算力网的一个重要基础,能够实现端到端的时延抖动小于 50 微米、零丢包、完全达到无损传输的标准,中国团队已经在确定性网络的关键技术上实现了突破。“原来的互联网还在发挥作用,但是它发挥的只是一个普通网络的作用,确定性网络是要实现算力网络的高铁、高速公路、航空的功能。”刘韵洁说。

“为智铸算”是迫切需求

当下,随着大模型赋能千行百业,“为智铸算”成为迫切需求。

“2003—2023 年 20 年间智能算力需求增长了百亿倍,是构成目前算力增长最主要的驱动力。”余晓晖表示。

“新的算力对传输的要求越来越高。”刘韵洁表示,大模型、空间计算以及有感知、有思维的智能体都对网络提出了新需求、新标准,特别是前 5 年 GPU 增长 90 倍,而网络带宽仅增长 10 倍,带宽瓶颈问题日益凸显。“无损协议如远程直接数据存取(RDMA)已成为算力网的必需。”

为加快打造统一的行业认证标准,更好评估国产 GPU 芯片在不同场景下的性能表现,加快促进工序匹配和生态成熟,新华三集团已经在上海、杭州建立了国产 GPU 测试适配实验室和大模型、中小模型适配实验室,并与超过 15 家国内 GPU AI 芯片厂商进行深度协同,目前适配超过 50 多款国产 AI 的加速卡,同时积极参与同大模型公司联合的芯片的测评标准与 AI 加速卡适配标准的研究与制定。

“算力用于模型训练是小场景,小应用,而赋能千行百业智能化才是算力真正的用武之地、是大场景。”紫光股份董事长、新华三集团总裁兼 CEO 于英涛说。

于英涛建议,统筹遴选、开放关键场景、核心数据,深化大模型在垂直行业的应用探索,打造出更多的“撒手铜”级应用。“政府主导的算力企业应积极面向中小企业创新创业者提供低成本甚至免费的算力支持,这才有希望打造出一批生成式人工智能(AIGC)时代的独角兽企业。”

超算+AI 演绎“超智融合”发展趋势

► 本报记者 张伟

近年来,超算与人工智能持续融合,应用边界不断拓展,应用场景不断丰富,成为科技创新和产业变革的重要驱动力。在人工智能发展催生海量算力需求背景下,超算与智算技术如何融合发展?如何定义“超智融合”?

9 月 23 日,中国智能计算产业联盟、国家高性能计算机工程技术研究中心等联合在武汉市发布国内首部“超智融合”研究报告《2024 中国算力发展研究报告之超智融合技术路线与趋势》,集中阐述对这些问题所进行的系统性思考与总结。在第二十届全国高性能计算学术年会(CCF HPC China 2024)期间,记者就此采访了多位业界专家。

“超智融合”成必然趋势

该研究报告指出,算力赋能人工智能的关键在于超算、智算基础设施融合演进。作为计算技术演进的里程碑,超智融合标志着从通用计算主导传统科学计算过渡到异构、通用计算融合的新时代。

研究报告强调,随着人工智能技术的快速发展,传统高性能计算(HPC)与人工智能计算(AI)的融合已成为不可逆转的趋势。这种融合被业界称之为超智融合,并且正在重塑计算科学、IT 产业和人类社会发展格局。

一方面,应用侧对算力的需求越来越多元化,超智融合兼具超算的双精度浮点运算处理能力和智算的整数运算或半精度 16 位的浮点运算能力,可满足“AI for Science”等混合计算精度应用场景需求;另一方面,随着 AI 大模型复杂度提高,超智融合可提升 AI 大模型训练、推理计算效率与可信度。

此外,超智融合发展将呈现出 3 个阶段:一是超算

支撑 AI 应用(For AI),利用强大算力提升 AI 性能;二是 AI 改进传统超算(By AI),通过 AI 技术使计算系统更加智能和高效;三是超智实现内生融合(Being AI),使 AI 技术成为计算系统的核心。

中国智能计算产业联盟秘书长安静分析认为,推动超智融合发展有如下几大关键因素:首先,人工智能催生的巨大算力需求,正推动超算性能向适 AI 化升级。目前,千亿参数人工智能大模型的训练已成为名副其实的算力黑洞。其次,应用侧对算力结构转型存在迫切需求,基础算力、智算算力、超算算力等应用的多元化发展催生超智融合。第三,单一计算架构已无法应对日益复杂的计算场景,通过超智融合技术将超算的强大处理能力与智算的算法优化能力融合发展已成为大势所趋。

在中国科学院计算技术研究所研究员张云泉看来,人工智能大模型迅猛发展,展现出新质生产力特质,但目前也遇到了算力瓶颈。他期待“我国在超算领域尽快拥有深厚技术积累,希望超算与人工智能融合发展以有效化解这一挑战。”

“超智融合”依托先进计算平台

研究报告指出,国家超算互联网等先进计算平台是超算、智算基础设施融合演进的重要依托,这些平台通过链接全国的超算、智算基础设施,汇聚全国各个算力中心的异构算力资源,可以让多个超算中心、智算中心之间更加紧密地融合,更高效快速地共同解决大型复杂计算难题。

数字化、智能化趋势下,算力已成为推动经济社会发展的重要驱动力。计算场景纷繁复杂,单一计算架构无法应对,超算与人工智能融合发展将有助于把超

算的强大处理能力与智算算法的优化能力充分结合起来。

自 4 月 11 日国家超算互联网平台上线以来,该平台依托一体化的算力调度、数据传输、生态协作体系,紧密连接供需方,通过市场化运营和服务体系,实现算力资源统筹调度,构建和完善全国一体的算力调度网络和面向超算应用的生态协作网络,打造国家先进算力底座。

目前该平台已连接算力产业上游 270 余家各类应用服务商、下游 100 余家行业软件生态企业,上线 6000 余款算力商品,服务用户超 10 万家,推动算力引擎赋能千行百业。

安静介绍了超智融合在多个行业和领域的创新应用潜力。在金融领域,超智融合发挥着重要作用,尤其是在处理复杂数据和优化算法方面,为金融分析和决策提供了强有力的支撑。在能源矿产勘探领域,超智融合显示出巨大潜力,能够处理大规模的数据集,提高资源勘探的效率和精度。

上海超级计算中心主任李根国指出,以上海气象局为例,在全球气候研究领域,超智融合通过其强大的数据处理能力和高效的问题求解策略,对复杂系统进行精准模拟,为全球气候研究提供了重要支撑。以上海药物所为例,在生物医药研发领域,超智融合的应用加速了科技成果转化和知识边界拓宽,有助于加快药物发现和开发过程。

“随着超智融合技术的不断突破和应用深化,它将成为推动科技创新的关键力量,为各行业的转型升级和数字经济发展提供核心驱动力。”安静强调,超智融合技术的发展需要产学研用各方的共同努力。她呼吁业界加强合作,共同推动超智融合技术的研发和应用,为我国计算产业的创新发展注入新动力。



图片来源:主办方供图

“未来启航 6G 创新发展论坛”现场

6G 发展将走向何方? 大咖们这样说

► 本报记者 罗晓燕

6G 作为未来产业和新一代信息技术的典型代表,是新质生产力的重要组成部分,已成为全球信息产业新一轮科技创新的高地。

近日,由中国移动通信集团有限公司主办,中关村泛联移动通信技术创新应用研究院、中国信息通信科技集团有限公司协办的“未来启航 6G 创新发展论坛”在北京举行。论坛现场,业界大咖围绕 6G 未来发展方向,分享前沿技术观点、研判机遇挑战。

6G 标准化工作 进入实质阶段

“当前,3GPP 正式启动首个 6G 标准项目,标志着 6G 从技术预研阶段正式进入实质性标准化阶段,进入至关重要的蓝图绘制期。”中国移动副总经理高同庆说。

9 月 12 日,由中国移动担任主报告人的 6G 场景用例与需求研究项目,在澳大利亚墨尔本召开的第三代合作伙伴计划(英文缩写 3GPP,是制定全球通信国际标准的重要组织)业务与系统技术规范组全会上获得通过。这是 3GPP 的首个 6G 标准项目,得到全球超过 90 家企业的支持。

高同庆表示,进入到这一阶段之后,6G 发展将呈现出 3 个明显的特征:一是场景需求更加多元,低空经济、空天一体沉浸式通信、智能制造等潜在应用,驱动着 6G 向更高层次、更宽领域发展;二是技术融合加速深入,特别是随着人工智能技术的快速发展和广泛应用,AI+6G 已成为 6G 创新的应有之义;三是创新主体加速跨越,6G 研发更需要突破不同领域和行业的界限才能共同推动 6G 技术的创新和应用。

“标准化是通信技术发展的基础,也是推动产业协同、促进技术创新、保障网络安全、提升国际竞争力的关键。”中国通信标准化协会理事长闻库表示,在 6G 研究中,更应强调标准化的引领作用,以标准推动技术创新,以标准引导产业发展,以标准保障网络安全,发挥中国在全球 6G 标准化中的作用。

在闻库看来,6G 发展不能单打独斗,而是需要产学研用各方的深度合作形成合力。“我们要打破壁垒、搭建平台,鼓励跨领域跨行业的交流与合作,共同攻克 6G 关键技术,共同制定全球统一的 6G 标准,共同构建开放共享的 6G 产业生态。”

未来 3—5 年是 6G 技术 突破关键时期

中国工程院院士、北京邮电大学教授、鹏城实验室宽带通信研究部主任、中关村泛联移动通信技术创新应用研究院院长张平指出,当前 6G 研究已从“远景需求定义”进入“关键技术收敛和与标准化”阶段,未来 3—5 年将是 6G 技术突破和产业培育的关键时期。

张平表示,相对 5G,6G 技术演进面临一些未知挑战:一是通信性能大幅提升,可持续发展成为挑战。6G 的通信效率、时延和连接数等传统通信性能在质量上将大幅提升。二是多要素融合,尤其是通信与 AI 的融合是关键难题,需要通过通信、感知、计算、人工智能、数据、安全等要素的深度融合。三是业务和能力的按需供给,按需的定制

化和个性化服务与刚性的系统设计成为矛盾。四是空天地一体全域覆盖。6G 包含多样化的接入网,如移动蜂窝、卫星通信、无人机通信、可见光通信等,将构建空天地一体化网络,实现真正意义上的全球无缝覆盖。

“6G 的愿景之一就是实现物理世界和数字世界的深度融合,构建万物智联的新世界。”中国工程院外籍院士、英国皇家工程院院士王江舟指出,6G 的重要应用场景在于行业应用,6G 要考虑行业个性化,如扩展现实、智能交互、全息通信、数字孪生等。为了满足这些应用,相对于 5G 来讲,6G 要在诸如频谱效率、峰值速率、高可靠低时延等关键指标上实现量级的提升,“不是 10%、20% 的提升,而是要实现‘倍’的提升。”

王江舟表示,6G 将实现从信息阐述到信息服务的重大转变,实现通信、感知、人工智能、计算等多要素高度融合,成为全覆盖的崭新信息网络。以全息通信为例,6G 会从视觉应用扩展到操控应用,实现多人远程实时交互。“用户虽然相隔几千里,交流却能像‘面对面’一样,增加沉浸式的体验。”

推进 6G 与 AI 深度融合

当前业界达成的一个共识是,6G 要服务于各行各业,其通信性能必须大幅度优于 5G。然而,随着“香农定律”接近理论极限,通信行业面临瓶颈和挑战。

“香农定律”描述了在有限的带宽和一定的信号噪声比下,信道能够达到的最大信息传输速率是有限的。“香农线是一条走向崩溃的道路,对我们思想是一个束缚,要突破必须找到‘拐点’技术。”张平说。

“语义通信是未来通信的‘拐点’技术。”张平认为,AI 应该突破对香农的理解。语义信息是通信与人工智能融合的关键,借助语义信息,通信融合 AI 将实现跨越“最后一公里”的目标,推动无线 AI、具身智能等应用的成熟与普及。

“6G 将是 5G 的增强与扩展,人机物充分互联,业务的种类也更加多样,业务应用需求推动网络向智能化演进。”中国移动研究院院长黄宇红认为,6G 一定是 AI 的最大受益者以及最重要的赋能者之一。

在黄宇红看来,6G 发展存在一个“不可能三角”关系,即有效性、经济性和包容性之间的系统性问题。

对于 6G 如何突破“不可能三角”,黄宇红认为,中国移动提出了以简为约束,以智为手段,以智简网络打造开放底座、多元能力和开放范式。通过端到端服务化设计实现功能和资源的按需编排和组合,以轻量内核应对差异化需求,将 6G 打造成开放的底座,赋能千行百业。

黄宇红进一步表示,中国移动针对 AI 赋能网络性能与复杂度间需平衡、数据有效获取难等问题,提出低开销低复杂度 AI 设计及网络数据自生成技术,以提升网络运行、运维性能和效率;针对网络赋能 AI 面临的多能力编排、最优部署、电信级安全保障等问题,提出高效 AI 任务编排、多维资源协同部署、可信内生安全等技术,为网络自身和第三方用户提供泛在普惠、安全可信的智能服务。