

从全球超算 TOP 500 榜单看中国超算发展：

# 中国超算行稳方能致远

■本报记者 赵广立

当地时间 6 月 17 日,新一期(第 53 届)全球高性能计算机 TOP500 榜单作为国际超算大会(ISC19)的重头戏在德国法兰克福发布。昔日的最快超算,来自中国的“神威·太湖之光”和“天河二号”虽然被美国新升级的“顶点”和“山脊”抢去冠军,但中国大陆仍以 219 台超算上榜的总数蝉联份额第一。

这成为了国内各路媒体报道的标题。然而,在当天的 TOP500 超算榜单发布中,劳伦斯伯克利国家实验室超算科学家、超算 TOP500 榜单联合创始人埃里奇·斯特罗迈尔(Erich Strohmaier)在分析过这份榜单的 TOP50 和 TOP100 之后,留下了一个问题,引起了在现场的钱德沛、卢宇彤、张云泉等中国超算专家的关注。

斯特罗迈尔的问题是:“中国是不是超算强国?”

## 中国超算不够“超”?

斯特罗迈尔提出问题的出发点在于,榜单上的中国超算好像不那么“超”——“在全球超算 TOP100 之中,中国仅有 9 台设备上榜”“在 TOP50 的超算设备中,仅有 1 台来自中国的商用系统”。

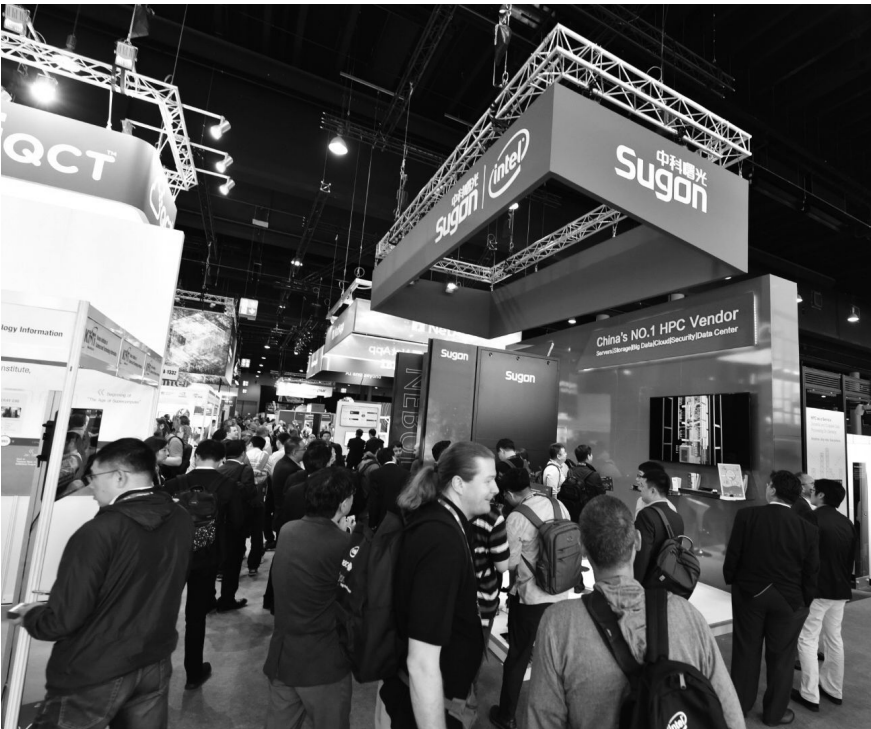
反映在整个榜单上,美国 116 台超算占据了全榜单 38.4% 的计算能力,而中国大陆 219 台超算所占的计算力只有不到 30%。

“在 TOP500 中我们数量最多,但在性能方面落后美国;而在 TOP50 之中,我们无论在性能还是数量上,都远远落后于美国、日本甚至欧洲的多个国家。”在由中国超算厂商中科曙光举办的“超算之夜”活动上,中山大学数据科学与计算机学院院长钱德沛就这一数据评论道:“这反映了中国超算的一个问题:和国外相比我们的应用类型、领域宽度等还有差距,中国超算的任务还很艰巨。”

担任本届 ISC19 大会程序主席的国家超级计算广州中心主任卢宇彤也关注到这一细节。在接受《中国科学报》采访时她表示:“TOP50 的超算设备中仅有 1 台来自中国的商用系统”,就是说在大中型超算应用领域,中国的高性能计算厂商还没有安装一些大规模系统。

“中国生产高性能计算机的能力很强,可以说超过了美国,但最尖端的还要差一些。TOP500 榜单中中国数量最多,但不排除有厂商拿一堆小机器去充数。如果打榜的测试基准发生改变——这也正是组委会在考虑的问题——这些小机器能否上榜还是未知数。”在展会现场,一位超算专业人士告诉记者,在超算的金字塔尖,依次是美国、欧洲、日本,然后才是中国。

“这些国家的整体实力在中国之上?”面对《中国科学报》的提问,这位专



ISC19 展会上中科曙光展台人流如织

家的回答是:“第一梯队中,我们还是弱一些。”

## 硬件、软件都要“赶”

如果把高性能计算机“拆开”看,中国超算的弱势就更加一览无余。

首先从驱动计算最关键的部件核心处理器来看,全球最快的 500 台超算中,有 488 台设备采用的是美国公司生产的芯片,其中芯片霸主英特尔的处理器出现在 478 台设备中,另有 7 台超算系统选用了 IBM Power 系列处理器,3 台系统选择了 AMD 处理器;另外,全球超算 TOP500 名单上共有 133 个系统采用了加速器或协处理器技术,其中的 122 台使用的是美国英伟达不同型号的 GPU 处理器。

另外,在高性能计算机所需的存储、通信等的高端核心芯片的选用上,中国超算还没有多少可匹配的国产化产品的选择。

“中国的‘神威·太湖之光’‘天河二号’等高性能计算机在核心处理器的选择上分别采用了申威、飞腾等国产芯片,总体而言,差距已经在缩小。”卢宇彤对记者说。

中国超算在软件方面的短板更加明显。“中国超算软件包括基础软件和应用软件,对国外商用软件的依赖程度非常大。”中科院软件所一位要求匿名的研究员告诉《中国科学报》,软件上的短板有很多因素,其中“重视不够”是重要原因

之一。

“与美国相比,我国超算设备的采购机制还不完善。”在 ISC19 展会上,中科曙光高性能计算产品事业部首席科学家吉青告诉《中国科学报》,美国的超级计算机经费一般分为三部分:假设硬件设备需要 3000 万美元,那么他们还有接近 3000 万美元支撑所有相关人员在其上开发相关软件和应用;另外还有一笔经费用于支持其他可能会用到的技术研发,“这种资金投入相对来说更为合理。”

记者就这一问题在采访卢宇彤时得到了同样的答复。她对《中国科学报》表示,中国在超算软件和应用开发上需要“与硬件开发相匹配的投入”。

## 从超算到超算能力

正如中国超算在 TOP500 榜单中的表现备受关注一样,在 ISC19 的展会上,中科曙光、联想、浪潮、华为等中国超算厂商也非常吸睛。记者在中科曙光展台注意到,有不少国外友人“组团”前来参观曙光今年在展台上展出的“身高”2.6 米的新一代硅立方高性能计算机,以及它所使用的曙光第四代液冷技术产品——全浸没式相变液冷散热系统。据介绍,利用该技术,系统的电能使用效率(PUE)值可降至 1.04 以下,相当于 40 瓦的功率冷却 1000 瓦的设备,而传统的风冷系统需要 500~1000 瓦。

“以前都是我们组队去看别人,今年我们都没机会离开展台,有许多机构

来看我们的。”吉青对《中国科学报》说,外国友人向她提出的问题涉及到高性能计算机的方方面面,除了液冷技术,还包括能效比、通用性、网络结构、运维管理,等等。

“每回答一个问题,都代表着我们在这个领域的能力,这也是新一代‘硅立方’要展示的内容——中国厂商的超算设计能力和用好超算的能力。”吉青说,高性能计算机不仅是国之重器、超级计算工具,还是实验最先进计算技术的平台。因此,只有高性能计算机做到最好,才能支撑最“上乘”、最“高大上”的应用,才会有更好的人工智能、精准医疗、气象模拟和预报。

以曙光为代表,中国超算厂商正努力做到这一点。比如,曙光在新一代硅立方高性能计算机的架构上选择了可广泛兼容的异构结构,用户可以根据所需选取不同的处理器和加速器;在通信方面支持业界主流的胖树拓扑和 6D-Torus 高速网络技术,力求系统网络性能的高带宽、低延时;曙光还成立一支专业队伍组建“EasyOP 在线运维平台”,在提供超算用户 7×24 小时的 PaaS 级运维的同时,还可根据用户不同应用定制、分配甚至出售超算服务。

“只有把高性能计算机的每一个环节都做好——搭好硬件、做好软件、做好管理和运维,才能有最好的超算。”吉青说,这也是中科曙光发展超算的理念,行稳方能致远。

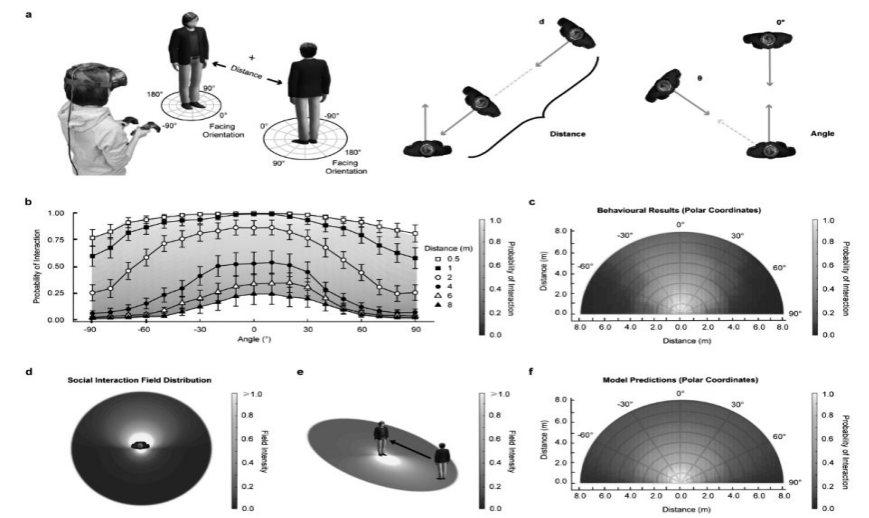
卢宇彤也告诉记者,中国超算的应用市场广阔,仅“天河二号”注册用户就有 3500 多个,平均负载率在 75% 以上。而且,随着超算应用越来越复杂,系统也会变得越来越复杂,超算厂商和应用科学家将逐渐形成一个整体来共同构建超算系统。

“我们有个提法叫做‘全软件栈/系统的协同设计’,它要求系统设施提供方和应用方互相配合,系统去适配应用,应用也要适配系统,通过协同设计,建设用户和系统之间的桥梁。”卢宇彤告诉记者,未来超算面貌可能发生的变化,对于中国超算而言既是挑战也是机遇。

吉青认为,通过协同攻关,中国超算有望实现从超算(Super Computer)到超算技术能力(Super Computing)的跃迁。

当地时间 6 月 19 日,中科院计算技术研究所研究员、国家超算济南中心主任张云泉在随中科院代表团访问斯洛文尼亚共和国卢布尔雅那大学的行程中,带来一个好消息:“欧洲可能不会再坚持用 Arm 处理器研发欧洲 E 级系统,甚至不排除由中国为欧洲建设 E 级超算。”

“如果实现,将是中国超算走向世界的一个里程碑。”张云泉对《中国科学报》表示,这代表着中国超算的研发、设计及产品、服务能力逐步得到国际的认可。



利用虚拟现实技术对人类社交交互行为进行量化测量的过程

图片来源：剃曙光团队

过程中,产生的数据可以作为判断其交互行为的指标。

研究结果提示人际距离、角度更接近的两个个体有较大概率被试判断为一个社会群组。基于这一实验的结果,剃曙光研究团队进一步建立了人类社会交互行为模型,首次量化地测量了人类交互行为的强度。

为了开展此项研究,剃曙光团队历时两年,共开展了 7 个实验,招募被试人员 179 人次,共计 190320 试次。但在审稿过程中,却面临不小挑战。为了回应外界疑问,研究团队在多种复杂、动态场景中对模型进行了验证,结果显示,其对于现实场景中的社会分组均达到了较准确的预测效果。最终,《自然—人类行为》杂志审稿人对于该研究在领域中的意义以及所使用的研究技术予以赞扬。

## 更懂人类 更大价值

在虚拟现实技术的帮助下,社交交互信息的量化测量成为了可能。此外,这一模型还能回答更多问题,为其他领

域带来更多想象力。

剃曙光表示,此前心理学研究无法量化的社会因素,比如权力、亲密度、吸引力等,借由该模型可以设计出量化测量的实验框架。

未来,在临床治疗方面,该实验范式与计算模型也将发挥重要作用,可借此考察自闭症、孤独症以及视力、听觉障碍患者社交功能情况,评估特殊教育或治疗效果。

“这样一项人工智能与心理学的交叉研究,在工业领域也有一定价值。”剃曙光说。他举例,目前,公共场合存在大量服务机器人,该模型的引入将使这些机器人更加人性化。服务机器人能够判断出与人类合适的交流距离;能够识别正在交互的人群,参与群聊或不打扰人类交流从容穿梭。

社交交互行为的发生基于个体,研究也需要回归个体。剃曙光表示,接下来,还需要对交互行为发生的姿态、表情、种族、性别以及区別细化,不断优化模型,达到还原真实人类反应的效果。

相关论文信息:  
DOI:10.1038/s41562-019-0618-2

## 异言堂

最近一段时间,我国部分技术受制于人的局面引起人工智能(AI)领域人士的担忧:架构于开源深度学习平台发展起来的 AI 技术产业,会不会再次陷入“缺芯少魂”的境地?

中国科学院大学一位教授曾在接受采访时表示,国内无论学术圈还是产业界都某程度踏进了“拿来主义”的怪圈:盲目跟随研究热点,缺少对人工智能的基础理论、基础模型、基础算法、基础工具等方面的研究与开发,经得起时间考验的工作较少。相应的,这导致我们在相关的 AI 应用研发过程中对外高度依赖。

以深度学习框架为例。处于硬件层和应用层之间的深度学习框架平台,相当于 AI 时代的操作系统和编译系统,其向上可以作为图像分割、文本分类等各种应用的入口,向下可以用于定义硬件功能和角色,发挥着承上启下的作用。然而,我国开源深度学习平台主要控制在国外巨头手中。在科研领域,谷歌的 TensorFlow、脸书的 PyTorch 等开源平台已经占据垄断地位;在产业界,近 90% 的国内开发者都在使用国外开源平台,其中半数以上使用的是谷歌 TensorFlow。

此外,这些科技巨头还将开源深度学习平台与自身的云计算服务结合,在加强数据运算能力的同时持续增强用户黏性,以培养开发者依托其平台开展 AI 相关开发的习惯。更值得一提的是,国外企业以“开源深度学习平台+AI 芯片”的模式,正构建起智能时代的“Wintel”式生态——如谷歌自研的机器学习芯片 TPU 与 TensorFlow 深度融合,让 AI 计算和开发更加高效的同时,也设置了越来越高的生态门槛。

反观国内,具备完全自主知识产权的开源深度学习平台,目前只有一棵独苗——百度的飞桨(PaddlePaddle)。而且,即便飞桨平台在技术上具备了与国外平台比拼的实力,但由于推出时间较晚,其在市场占有率和社区资料丰富度方面,仍难以与国外平台抗衡。换言之,单凭一家企业之力赶超,有相当大的难度。

目前,学术界和产业界对使用国外平台问题的认识还不足。但从

## 速递

## 复旦大学等将建集成电路产教融合创新平台

**本报讯** 清华大学、北京大学、复旦大学、厦门大学四所高校日前获批建设“国家集成电路产教融合创新平台”,将联手攻克集成电路领域研发难题并培养产研人才。

6 月 12 日,记者从复旦大学获悉,该校项目建设方案正式获批,实施期从 2019 年起至 2021 年,项目建设总经费 4.7 亿元,额度最高,体现了国家对上海及长三角地区集成电路技术创新、产业发展和人才培养的高度重视。

复旦大学国家集成电路产教融合创新平台将重点突出集成电路紧缺人才的培养与工程实践,建成后具备每年为 2000 人次提供集成电路实训手段的能力。此外,该平台将

AI 发展的形势来看,今后 AI 应用仍将主要基于开源深度学习平台开发和运行。开源深度学习平台不可控,意味着我们在智能时代也会面临失去核心技术主导权、规则制定权等巨大风险。

也就是说,在人工智能的新时代,我们可能面临从芯片到操作系统和应用的一整套技术和产业体系再次被国外控制的局面,让信息化时代受制于“Win-tel”的历史重演。

因此,政府决策部门应通过政策引导、设立专项、开展竞赛等方式,鼓励国内芯片、软件算法、硬件等企业、科研机构、开源社区和广大开发者,基于我国自主可控的开源深度学习平台进行应用创新;同时应注意到 AI 技术不仅仅是一项软件技术,AI 芯片等硬件也扮演着智能时代的重要角色。因此,建议飞桨平台应以开放共赢心态扩大与我国 AI 芯片厂商合作,共同突破核心技术,建立起国产 AI 软硬件技术的“统一战线”,更好地为我国 AI 产业服务。

总之,在芯片、操作系统等很多方面,我们确实需要下定决心,攻坚克难,自力更生。因为一旦核心技术受制于人,就有人“釜底抽薪”的威胁。所以,我们要有“宝剑锋从磨砺出,梅花香自苦寒来”的定力,逐渐在 AI 领域形成一批拥有自主知识产权的核心技术和产业,把发展的主动权牢牢掌握在自己手中。只有这样,不管是“狼来了”还是“狼走了”,才能泰然自若、从容应对。

# 量化人类社会交互行为

你问我爱你有多深,数据代表我的心

■本报见习记者 卜叶 记者 黄辛

清早,小明走进教室,发现同学们正在聊天,老师敲黑板试图引起学生们的注意,李雷和韩梅梅明白了老师的意图,停止聊天并坐下。此时,小明也快速走到自己的座位前坐下。

看似顺理成章的交互行为到底是如何发生的呢?小明是如何识别出正在交互的群体的呢?又是基于什么原则感知交互信息,指导自我行为的呢?社会心理学家与认知心理学家研究多年,却始终未能解开谜团。

日前,华东师范大学心理与认知科学学院剃曙光团队研究取得新进展,该研究利用虚拟现实技术结合经典心理物理学方法,首次量化测量了人类之间的交互行为,并在此基础上提出了社交交互场模型以及用于解释人类社会知觉分组的基本算法。相关研究结果,近日发表于《自然—人类行为》杂志。

## 对亲密距离的深深误解

人与人之间吸引力不可名状,却让亲密的人越来越靠近。这一现象在美国人类学家爱德华·霍尔的个人空间理论(Personal Space Theory)中得到印证。个人空间存在 4 种距离,0.45 米以内是亲密距离,常发生在爱情、亲友关系之间。0.45—1.22 米属于个人距离,适合熟悉的人之间,可以亲切地握手、交谈。1.22—3.7 米属于社交距离,常发生在工作场合和公共场所。大于 3.7 米的属于公众距离,比如在大会堂发言、演讲、戏剧表演、电影放映时与观众保持的距离。

自 1966 年该理论提出以来,这一距

离描述着人与人之间的关系,殊不知,这却是一项定性描述,0.45、1.22、3.7 只是对理论的描述,数字的科学性无从验证,人类社会交互描述这一重要问题的探索始终停留在理论层面。

其间,研究人员从未停止量化描述人类社会交互行为的尝试。传统实验中,研究人员通过实验对象对真实人类行为的反应,量化判断交互行为。但实验场景、人类行为发出方的些许差异就会导致实验结果的不确定性。

该论文通讯作者剃曙光介绍,传统实验存在诸多难以操控的社会性刺激,无法精确重复同一个交互行为的实验条件,也难以真实客观还原社交交互场景,因此,基于此种研究方法的结果很难具有普适性。

## 在虚拟世界追踪现实

长达 9 年的心理学学习过程中,剃曙光一度对这一科学难题产生兴趣,但始终未找到解决办法。一个偶然的机会,剃曙光进入了工业领域,虚拟现实技术在该领域的应用像万有引力一样吸引、包围着他。他突然想到,“能不能用虚拟人来模拟交互行为呢?”

2017 年,剃曙光带领研究生团队利用虚拟现实技术,量化地操作虚拟空间中的各类变量,开展认知心理学、社会心理学与人因工效学的研究。

研究团队使用 3D 引擎编程,实现了虚拟人物之间的距离和角度的操控,量化地测量了人类之间的社交交互强度。具体而言,实验对象在使用虚拟现实设备,调控两个虚拟人物的距离、角度的

谨防受制于人的历史在智能时代重演

■赵广立

经济高效、安全可靠、可持续等特点。然而,因缺乏统一的设计和工程建设标准指导,导致 POL 的设计定位、主要建设内容和工程质量存在一定差异,制约了 POL 全光园区产业的推广与使用。

该标准系统性地从规划设计、设备配置、施工、调试与试运行、检测验收四大维度阐述和定义 POL 全光园区设计和施工规范,旨在加强 POL 工程建设项目的设计、施工水平和效率,解决 POL 全光园区产业发展的核心痛点。(田晓雨)