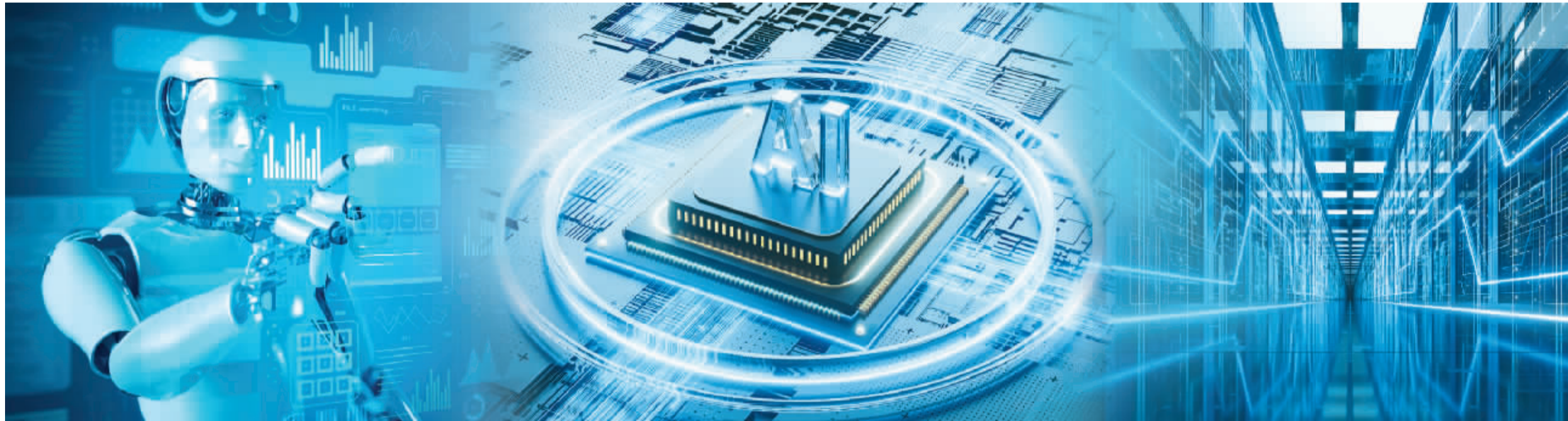


# 抢抓“AI+”机遇 中国企业深度参与产业链



视觉中国图片

3月19日，GTC 2024（英伟达技术大会）在美国开幕。英伟达正式推出了基于更强大的Blackwell架构的B200芯片及其算力集群，并与比亚迪、联想集团等中国企业宣布在硬件支持、平台开发等领域的新合作。

业内人士表示，中国人工智能应用市场潜力巨大，英伟达更加主动地对中国市场进行开放，这对全球人工智能产业链供应链建设起到巨大促进作用。中国企业要抓住“人工智能+”机遇，打造不同行业领域的高水平大模型应用产品，同时进一步加大人工智能芯片研发，提升人工智能产业链韧性。

● 本报记者 彭思雨

## AI应用有望爆发

英伟达CEO黄仁勋表示，人工智能大模型参数量正在呈指数级增长，为了支持人工智能大模型和下游应用蓬勃发展，必须提供支持更快推理速度的GPU（图形处理器）。Blackwell GPU的发布意味着，近八年来人工智能算力增长了一千倍。

赛迪产业研究院院长赵刚对中国证券报记者表示，Blackwell芯片是英伟达GPU的最新升级版，推理性能比上一代最先进的H100芯片提升了30倍，能够让用户在拥有高达万亿参数的模型上实现AI训练和实时语言大模型（LLM）推理，成本和能耗却较上一代产品降低多达25倍，能够全面支撑科学计算、工程分析、数字孪生、自动驾驶等多种大模型应用场景，成为支撑人工智能大模型发展的全球最先进的引擎。

赛迪顾问人工智能与大数据研究中心常务副总经理邹德宝对记者表示：“AI大模型需求推动下，当前全球算力市场供应短缺，英伟达GPU的升级，能够为AI大模型技术及应用提供硬件架构的有力支持，甚至为传统人工智能通过通用人工智能发展创造算力条件，支持更大规模大模型预训练。”

业内人士普遍认为，英伟达试图让投资者相信，一个由生成式AI驱动的价值100万亿美元的新行业开始了。英伟达展现了推理算力的强大能力，为大模型走向行业应用奠定算力基础，刺激AI大模型应用在2024年爆发式增长。

## 中国市场潜力巨大

黄仁勋在大会上表示，比亚迪将采用英伟达下一代智能汽车芯片Thor，还将使用英伟达基础设施进行自动驾驶模型训练。联想集团与英伟达宣布合作推出全新混合人工智能解决方案，为企业和云提供商获得AI关键的加速计算能力。

“得益于实时计算、能源效率及部署便利性的持续改进，全新的AI应用场景正迅速涌现。通过与英伟达的紧密合作，联想集团已经在效率、性能和成本方面取得显著进步，这不仅加快了人工智能在各个行业中的应用部署，还将帮助企业洞察数据，助力提升零售体验、改造城市面貌、推动智能制造进入新台阶。”联想集团执行副总裁、基础设施方案业务集团（ISG）总裁Kirk Skaugen表示。

此次大会期间，英伟达将3月19日设置为“中国人工智能日”（China AI Day），阿里云、腾讯云、百度、蚂蚁集团企业嘉宾参与演讲。

“我明显感受到英伟达对中国市场开放的主动性更强了，这令人出乎意料。”邹德宝表示，过去英伟达更多地聚焦海外市场，中国下游终端厂商往往要主动和英伟达谈芯片或技术支持。英伟达更加主动地对中国市场进行开放，这对全球人工智能产业链供应链建设起到巨大促进作用。

邹德宝认为，无论是以比亚迪为首的智能网联汽车领域，还是以联想集团为代表的面向消费端和企业端客户的人工智能应用领域，都证明了中国市场拥有非常多元化的人工智能应用场景，行业应用客户群体庞大。因此，对英

伟达来说，中国市场的重要性不言而喻。

在邹德宝看来，我国人工智能行业正在朝着“人工智能+”的方向发展。算力中心、数据要素等环节建设的最终目的是实现人工智能技术成果的转化，通过人工智能产业化带动国内产业结构实现新旧动能转换，助推新质生产力建设。

当前，国内人工智能市场正稳步扩张，生成式人工智能（AIGC）应用成为核心驱动力。根据非凡产研发布的《2024年2月全球AIGC行业月报》统计，国内规模以上AI产品数量已经从1月的123个增加到2月的130个。

分析师认为，这一稳步增长不仅反映了中国人工智能应用市场的扩张，也预示着产品更趋多样化。这一增长背后，是中国在人工智能技术研发和应用推广方面的持续投入，以及消费者对于智能解决方案日益增长的兴趣。

多个产品用户使用活跃度高。上述月报显示，2024年2月，百度文心一言以1001万的访问量位居第一，显示出广泛的用户基础和较高的市场认知度；月之暗面以297万的访问量和105.80%的环比增长在市场上脱颖而出。

## 打造核心竞争力

当前，人工智能产业已成为全球技术比拼高地，也焕发着经济发展的勃勃生机。中国企业如何深度参与全球人工智能产业链？

邹德宝表示，人工智能产业链主要包括硬件、软件和服务，中国人工智能产业发展初期更多聚焦在软件和服务层面，硬件仅占大约不到两成的比例。而目前，国内人工智能产业链出现软件和服务终端产品市场占比不断缩小、硬件市场规模不断扩大的走势。目前国产算力已经在汽车、通信、传感、智慧交通等领域，特别是非即时性的应用场景中表现出较好的应用效果，国产算力市场产业链的完备性、企业研发能力以及人才梯队建设能力都得到了显著提升。

工信部数据显示，2022年我国算力核心产业规模达到1.8万亿元，算力正加速向政务、工业、交通、医疗等各行业各领域渗透。截至2023年6月底，全国在用数据中心机架总规模超过760万标准机架，算力总规模达到197EFLOPS（EFLOPS是指每秒百亿亿次浮点运算次数），位居全球第二。算力总规模近5年年均增速近30%，存力总规模超过1080EB。

“长远来看，未来人工智能技术发展可能会形成‘小数据、大运算’的模式，即大模型将具备自我推理能力，不再需要大规模学习，从而对算力要求减小。如果国产算力能力在未来10年持续提升，将弥补当前人工智能产业链薄弱环节，提升产业链整体韧性。”邹德宝说。

赵刚认为，大模型应用创新将是我国人工智能大模型企业的核心竞争力。我国大模型企业要抓住“人工智能+”机遇，打造不同行业领域的高水平大模型应用产品。

“每个人听懂的细节各不相同，但对于AI的热情是相同的。我们的确处于AI加速发展的时代，中国企业也在思考浪潮下的发展方向。”出门问问副总工程师李维在GTC大会现场发出这样的感慨。

华西证券认为，英伟达作为全球算力龙头，正在开启信息科技行业新一轮成长周期。我国产业端积极响应，大模型已经产生相关落地应用，智能算力建设正处于持续提速阶段。

## 英伟达发布人形机器人通用基础模型 商业化落地料加快

● 本报记者 杨洁 见习记者 陈子轩

机器人尤其是人形机器人是3月19日英伟达2024 GTC大会议程的压轴环节。英伟达CEO黄仁勋在发布会上推出人形机器人通用基础模型Project GR00T，以及一款基于NVIDIA Thor系统级芯片（SoC）的新型人形机器人计算机Jetson Thor。此外，英伟达还对其机器人平台NVIDIA Isaac进行了升级。

“开发人形机器人通用基础模型是当今AI领域中最令人兴奋的课题之一。世界各地的机器人技术领导者正在汇集各种赋能技术，致力于在人形机器人通用领域实现突破。”黄仁勋激动地表示。

### 人形机器人进展不断

黄仁勋认为：“下一代机器人可能是人形机器人，因为有了更多的模仿训练数据可以提供给我们构造相似的机器人。人形机器人可以更有效地部署在人类设计的工作站、制造和物流环节中。”

英伟达GTC大会通过视频展示了多台人形机器人完成各种高难度动作指令，无论是打果汁、做家务还是跳舞、打鼓，机器人都应对自如，随后黄仁勋和身后屏幕这一列由英伟达支持的机器人共同向会场观众致意。

黄仁勋现场发布了人形机器人通用基础模型Project GR00T，这一模型将帮助机

器人理解人类语言、模仿人类动作，以使其适应真实世界并有效交互。在运算芯片方面，为了帮助企业成功运营Project GR00T，英伟达还发布了专门用于人形机器人的芯片Jetson Thor。Jetson Thor单芯片搭载Blackwell GPU架构，是一个新的计算平台，能够执行复杂的任务，并与人和机器安全自然地进行交互，具有针对性、功率和尺寸进行优化的模块化架构。

此外，黄仁勋称，正在为多家人形机器人公司开发一个综合AI平台，如1X Technologies、Agility Robotics、Aptronik、波士顿动力公司、Figure AI、傅利叶智能、Sanctuary AI、宇树科技和小鹏鹏行等。

将人工智能大模型与机器人结合开发具身智能，为人形机器人发展按下加速键。OpenAI近期已经联合两大人形机器人初创企业1X和Figure进行具身智能研发，今年3月，Figure公司发布了被OpenAI大模型加持的人形机器人的日常工作演示视频，在各大社交媒体刷屏。

英伟达在人形机器人赛道布局已久。2023年2月英伟达参与了Figure的B轮融资。今年2月，英伟达正式成立通用智能体研究实验室GEAR，正式进入具身智能领域。长城证券认为，人形机器人的本质是“AI+传感器+执行器”的结合，AI算法和运算芯片是当前人形机器人商业化落地的主要瓶颈。英伟达人局或可构建机器人专属的算法算力平台，对整个产业实行AI赋能，驱动更多厂商尝试将大模型用于训练具身智能产

品，加快商业化落地。

### 迎来产业化关键期

本次GTC大会，英伟达还对机器人市场的Isaac机器人平台也进行了重大更新。黄仁勋着重讲解了适用于自主移动机器人的Isaac Robotics平台，该平台的主要落地产品包括英伟达和九号公司联合开发的自主移动机器人平台Nova Carter AMR，以及开发者套件Nova Orin，这些产品大幅简化自主移动机器人二次开发流程，提高开发效率，推动机器人产业进程。

随着大规模设备更新的推进，机器人尤其是人形机器人的关注度居高不下。业内人士认为，人形机器人有望成为继计算机、智能手机、新能源汽车后的又一颠覆性产品，将带动广阔蓝海市场。当前，人形机器人正处于从实验室迈向产业化的关键节点。

广发证券表示，在海内外市场格局方面，人形机器人产业大致可以分为五方面势力：以Tesla为代表的海外车企，以波士顿动力为代表的海外机器人企业，以小鹏为代表的国内车企，以优必选、宇树为代表的国内机器人企业，以及小米、科大讯飞为代表的国内互联网企业。其认为，人形机器人产业化趋势之潮已至，多方势力分别在软件（算法、人工智能技术等）及硬件（自研电机、整体结构、运控能力等）等方面具备先发优势，同时部分厂商已具备商业化能力，有望于今年四季度交付产品。

## 液冷时代有望加速来临

● 本报记者 杨洁

在3月19日英伟达2024 GTC大会上，英伟达CEO黄仁勋发布了更高性能的GPU芯片——基于Blackwell架构的B200以及超级芯片GB200。同时，黄仁勋还公布了由Grace Blackwell超级芯片GB200提供支持的服务器系统。由于功耗太高，液冷的散热方式成为系统标配。正如黄仁勋此前曾提到的，沉浸式液冷技术将是未来方向，将带动整机散热市场迎来全面革新。

英伟达的选择在业内具有风向标意义。多家机构表示，这次发布会或将成为液冷发展的重要推动力。中信证券表示，在数据中心算力提升和能耗降低推动下，液冷技术将加速替代风冷技术，成为新的技术发展方向。在液冷技术中，冷板式液冷技术在性能与成本之间形成平衡，是现阶段最佳的散热方案，建议关注液冷系统及其核心零部件的投资机会。

### 液冷散热大势所趋

据黄仁勋介绍，GB200架构包含两个GPU、一个CPU，通过高速连接组合在一起，这将带来性能的极大提升。以大语言模型LLM推理训练为例，相较于英伟达H100 Tensor Core GPU，GB200超级芯片可实现30倍的性能提升。

英伟达还推出了超级计算机机柜DGX GB200 NVL72，拥有18个GB200节点机架，每个节点搭配2个GB200 GPU。黄仁勋表示，一个DGX GB200 NVL72机柜可以训练27万亿参数的模型。但由于功耗过大，所以也采用了液冷的方式，同时重量高达3000磅（约1361公斤）。

大会上，鸿海集团子公司鸿佰科技（Ingrasys）发布一款使用英伟达GB200芯片的NVL72液冷服务器，该服务器集成了72个英伟达Blackwell GPU和36个英伟达Grace CPU。鸿海的竞争对手和硕也发布了基于GB200 NVL36的液冷解决方案，用于处理计算密集型工作负载。

业内人士认为，算力器件的密度和功耗持续增长，对传统的风冷散热模式带来挑战，能耗成本不断抬升以及政策对PUE指标监管趋严，共同倒逼产业对液冷需求升级，冷板式液冷或率先放量，沉浸式液冷为长期方向。

华西证券研报表示，AI大模型的热火引爆了各行业对算力基础设施的需求。更高密度、更大规模的数据中心建设，给数据中心带来了高功耗的挑战，液冷散热的渗透速度也在加快。主流技术方面，现在市场上主流的液冷解决方案是冷板散热；产业生态方面，从基础设施到AI设备的全栈交付模式已经基本成为行业共识；客户群体方面，互联网和电信这两个数据中心最大的行业客户已经对液冷表示肯定的态度。

### 国内厂商积极布局

据了解，我国液冷服务器普及率不足10%，随着AI和算力需求不断增长，数据中心单机柜功率密度将呈增长趋势，液冷服务器的占比有望进一步提升。根据赛迪顾问数据，单机柜20kW一般认为是风冷可解的散热极限，预计2025年新建数据中心的平均单机柜功率将达到25kW，数据中心液冷散热方式将逐渐取代风冷。

液冷散热目前的技术路线主要是冷板式液冷和沉浸式液冷两种，选择哪种路线，

与单机柜功率有较大关系。一般认为在只考虑初始投资成本时，单机柜功率不超过10kW时，采用风冷相对经济；单机柜功率在10kW-100kW时，采用冷板式液冷较为合适；当单机柜功率超过100kW的时候，采用沉浸式液冷更经济。

曙光数创在接受投资者调研时介绍，随着产业不断成熟，当前冷板式液冷的初始投资成本已经接近甚至低于风冷，而考虑后期运维成本的情况下，沉浸式液冷的CTO（总体拥有成本）在4.5年左右后与风冷持平。

冷板式液冷市场目前竞争激烈。曙光数创是新一代数据中心冷却技术领先的整体解决方案和全生命周期服务供应商。其表示，过去一年，液冷市场处于产业初期阶段，各家厂商以抢占市场份额为目标，尤其是冷板式液冷市场竞争较为激烈，导致公司整体毛利率与上年同期相比有所下降。赛迪顾问《2023中国液冷应用市场研究报告》显示，2021年至2023年上半年，曙光数创以平均58.8%的市场份额，位列中国液冷数据中心基础设施市场部署规模第一。

英维克近期在投资者互动平台介绍，公司拥有全链条液冷技术以及超过900MW液冷交付经验，能够为算力设备以及算力设备的机房部署提供适配各种场景的高效温控解决方案。近日，英维克作为第一中标候选人，中标了中国电信的弹性DC舱（2024-2025年）集中采购项目，该项目合计采购弹性DC舱3200套，其中包括数据中心液冷层DC舱323套。国信证券分析认为，运营商积极推进液冷应用，相关招标有望陆续启动和落地，国内液冷市场前景广阔。