

附件 2

算力强基揭榜行动任务榜单

一、计算

（一）云边端算网协同管理系统

揭榜任务：面向云边端多层级算力环境，研发算网协同应用管理系统，设计面向不同应用软件架构的管理机制，支持对不同架构应用软件的统一管理；研发应用软件在算网协同中的自动化构建部署能力，支持应用软件的自动构建和分发部署；研究算网协同应用系统的一体化观测能力，降低运维复杂度，提高复杂应用软件运行的稳定性和可靠性。

预期目标：到 2026 年，研制应用软件管理系统，支持对传统应用软件、云原生应用软件、AI 应用软件、大数据应用软件等不少于 5 种应用软件的全生命周期管理。研究基于算网协同的分布式构建和部署技术，支持上述应用软件的自动分发和跨算力节点部署，实现零人工介入。研发算网应用一体化观测功能，具备白盒化动态分析及智能故障根因定位能力。在不少于 3 个行业完成试点验证。

（二）支持超大规模参数模型的训推一体化异构智算平台

揭榜任务：面向人工智能大模型训练和推理对计算资源的需求，研发支持超大规模参数模型的训练、推理

一体化智算平台，包括资源调度策略、训推加速套件等，并可支持多种硬件架构，屏蔽底层硬件差异，提升超大规模模型在训练、推理过程中稳定性、资源利用率和运行效率。

预期目标：到 2026 年，研发一套支持万亿参数模型的超大规模训推一体化智算平台，万卡环境下稳定训练时间不低于 30 天，有效训练时长不低于 95%，训练效率较当前主流水平提升不低于 30%，推理效率提升不低于 50%。支持主流深度学习框架，兼容多种硬件架构，并提供统一的编程接口和开发环境，实现不低于 10 个行业用户的落地验证。

（三）异构算力跨域任务编排系统

揭榜任务：针对跨域异构算力协同，研发跨域异构算力管理系统，实现跨域异构算力的管理和应用。研发针对多样性算力的规范化开放互联功能，支持对不同类型的异构算力模型统一抽象封装；研发跨域异构算力的管理功能，支持对跨域异构算力的统一管理和协同；研究跨域多主体算力的安全认证和控制方法，保障跨域协同安全。

预期目标：到 2026 年，研发不少于 6 种跨域协同调度算法，支持数据处理、函数计算、机器学习等不少于 3 个场景的计算任务部署，完成不少于 5 个跨域算力中心的统一管理。研发跨域多主体算力的安全认证方法，支持云边端等不同层级算力协同的安全要求。在不少于 2

个行业完成试点验证。

（四）训推算力一体机

揭榜任务：面向人工智能训练、推理场景，研发基于基础设施即服务（IaaS）和平台即服务（PaaS）的高性能训推一体化解决方案，覆盖对大模型开发训练和部署推理的全流程，包括数据准备、模型训练、模型评测和模型部署。同时，支持大模型加密、攻击防御等能力，解决针对大模型数据泄露、指令攻击等安全问题和风险。

预期目标：到 2026 年，研发支持至少 3 种指令集芯片的训推一体机，针对至少 5 个行业开展人工智能训推一体机应用，为用户提供多元化训推一体化服务，并在至少 10 种不同的场景进行人工智能训推一体机落地。

（五）大规模异构算力集群推理加速技术

揭榜任务：研发存储、网络、计算的协同优化技术，通过模型加速、调度加速等方法实现大规模异构算力集群在大模型推理方面的加速，从而支持更大的模型、更长的上下文、更高的性能及更低的能耗，促进算力芯片在大模型推理方面的更好应用。

预期目标：到 2026 年，实现集群有效吞吐量 5 倍以上提升，实际应用场景中可处理的请求数提升 1 倍以上，首字延迟性能提升 1 倍以上，芯片利用率提升 50% 以上。通过优化算力中心计算、存储、网络的配比以及拓扑结构和系统调度策略，实现千卡以上异构集群在推理加速领域的突破。

二、存储

（六）磁光电融合存储系统

揭榜任务：针对单一存储介质难以满足多样化数据存储需求的现状，依托磁、光、电存储在性能、寿命、功耗等方面的差异化特性，将磁、光、电存储技术进行融合，研发磁光电融合存储系统，构建基于固态硬盘（SSD）、机械硬盘（HDD）和光存储的多级存储架构。根据业务特征，将数据保存在不同级别的存储设备中，实现海量数据的集中、统一存储管理，支撑算力中心高效、低碳、安全持续发展。

预期目标：到 2026 年，研发磁、光、电融合存储系统，支持适配分布式文件、分布式块和分布式对象等至少 3 种存储类型，系统可以根据数据的访问时间、访问频率、文件属性等自定义分级策略，根据业务负载动态调整迁移。系统可通过介质安全、系统安全、软件安全等夯实底层安全能力，通过防勒索、加密算法、远程监控、光存储预警检测等增强数据安全能力。打造磁光电融合存储应用示范，完成至少 20 个业务系统应用，实现至少 4 个东部地区数据流动至西部磁光电存储系统，且数据存储量不少于 10PB。

（七）存储调度管理及应用技术

揭榜任务：针对海量数据存储和算力孤岛问题，研发跨域多算的存力调度、存网编排和存算网一体化系统，实现数据的智能冷热分级、应用的跨域无感访问等能力，

有效降低成本、提高性能和支撑业务。系统具备资源规划、策略调整能力，可优化和调整全网数据存储布局，实现对不断变化的需求的适应。

预期目标：到 2026 年，研制具备高效、可扩展性的存储系统，基于智能算法，对数据进行分析和调度，实现应用无感访问和智能流动。研究存力调度策略，使数据召回率控制在 30% 以下；研究基于潮汐网络调度算法，实现网络带宽利用率提升 50% 以上，达到存网一体的目标。集成存储、计算和网络的能力，支持存算网一体化调度，在算力中心资源池落地应用。

三、网络

（八）高性能数据处理器（DPU）

揭榜任务：开展基于芯粒（Chiplet）和第五代精简指令集（RISC-V）技术的软硬件一体 DPU 芯片技术研究，支持算力中心、智算中心、超算中心场景所需的超高带宽和超低时延，突破 Chiplet 异构芯片封装技术、高速 Serdes 通信、大规模无损网络拥塞算法、硬件密码算法、高性能虚拟化、硬件可编程等技术，实现基于 ARM、X86、RISC-V 等异构核心的 DPU 应用，提升算力中心基础设施处理能力和数据传输能效比。

预期目标：到 2026 年，完成超高性能 DPU 芯片研发工作，吞吐能力达到 400Gbps，单向流量时延不高于 30us，支持与国内外主流 CPU、GPU 芯片平台的适配，支持主流操作系统兼容，支持数据报文硬件处理逻辑可编程。

（九）基于 RoCE 的智算网络

揭榜任务：面向 RoCE 网络开展设备及管控系统研发，通过提高设备带宽、优化负载均衡算法、强化网络流量规划及运维能力等方式，提升 RoCE 网络的吞吐量和时延性能。研制新一代智能化管控工具，引入 AI 大模型能力，简化 RoCE 网络的部署和配置工作，实现全局、多维度的可视化运维。在网络波动、业务变更、故障等情况下，网络参数自动调整，流量快速切换，从而达到提升网络效率和降低运维成本的目标。

预期目标：到 2026 年，实现新型 RoCE 网络整体方案的商用部署，网络性能提升 10%以上。通过智能化管控及运维工具，网络部署难度大幅降低，运维效率提升 50%以上，可支撑更大规模部署和应用。

（十）光交换智算网络技术与验证

揭榜任务：面向智算集群低功耗、高带宽、低延迟技术需求，开展智算集群光交换组网关键技术与验证，重点突破智算集群光交换组网、路由协议适配等关键技术。针对智算集群的功能、性能、可靠性和扩展性等要求，研究光拓扑映射、光电混合路由、多路径负载均衡等技术。

预期目标：到 2026 年，实现支持智算集群的易操作、高可靠、可平滑过渡升级的光网络，支持人工智能等关键业务承载；光交换设备单端口速率支持 100GE/400GE/800GE，交换容量弹性可扩展，可支持不少于 3 种异构算力资源互联，在不少于 2 个智算集群完成验证，并完成不少

于 3 种智算业务承载验证。

（十一）面向分布式智算中心的网络关键技术与验证

揭榜任务：针对智算集群从集中式向分布式部署探索的趋势，攻关算力中心间网络技术，研发面向智算中心间的高可靠传输设备，构建智算中心间超大容量、超低时延、超高可靠光电协同网络，实现智算中心高速、可靠互联。

预期目标：到 2026 年，突破智算中心间超大容量、超高可靠网络传输关键技术，研制面向智算中心间网络的传输设备，单波速率不低于 1.6Tbps，设备时延不超过 30us，支撑分布式智算中心间业务的高可靠传输。

四、应用

（十二）智算中心跨域互联应用

揭榜任务：优化人工智能算力基础设施布局，构建跨地域互补、协同算力调度的超大规模人工智能算力服务能力。加强与人工智能芯片厂商的兼容适配，构筑大规模高性能异构算力池，提供面向大模型训推场景深度优化的弹性调度、弹性容错、高资源利用率的人工智能算力服务。

预期目标：到 2026 年，形成覆盖 5 个以上全国重点算力枢纽节点的人工智能算力中心，支持跨地域、跨云的算力需求感知和动态调度，完成 3 款以上算力芯片适配，聚焦大模型训练和推理场景，构建大规模、高性能、弹性调度、高容错的训推一体算力资源池，具备分钟级断点续训能力，支持万卡级别并行训练。

（十三）算力电力协同应用

揭榜任务：研发基于算力调度技术与能源大模型的多云异构算电协同管理平台，构建基于数据驱动的算力集群用电负荷特性模型、基于计算任务的时空转移特性的能源大模型，推动算力预测与调度技术在智算中心应用落地，提升整体资源利用率，基于新能源、新型储能系统开展算力负荷与电力系统的协同优化，实现精准、动态、实时的能源调度与交易，实现算力与电力等能源的深度协同。

预期目标：到 2026 年，实现智算场景下能源与算力全链路的数据穿透及流程整合，构建“算”随“电”动的直接控制及间接引导机制，实现算力需求预测精准度达到 70%、集群有效负载率提升 25%以上，智算中心整体集群资源利用率提高 10%。结合算力集群用电数据、时间周期、气象数据、工作负载等多种因素，实现“电”随“算”用的能源效率优化与算效提升，实现基础设施用能决策精准度 85%以上，响应时效性达到提前 15 分钟响应级别，智算中心整体算力能效水平提升 30%，算力中心用电成本降低 5%以上。

（十四）大规模通信业务场景中的算力应用

揭榜任务：围绕网络功能虚拟化（NFV）系统架构，针对 NFV 中网络性能、资源利用和灵活展性等方面的挑战，研发面向 NFV 架构的高性能虚拟化、智能化网络管理和资源编排算法等技术和系统，突破虚拟化层与硬件加速器（如 FPGA、DPU、GPU）之间的协同能力。

预期目标：到 2026 年，NFV 算力平台系统中实现对虚拟化网络功能的智能调度，支持异构集群部署、动态扩展，资源动态分配，虚拟化资源利用率提升 20%以上；支持 GPU、FPGA 等硬件加速器的虚拟化调度，加速网络处理性能至 Tbps 以上；支持智能化网络虚拟化功能管理，提升 NFV 系统的自动化运维能力和管理效能，故障修复时间缩减不低于 30%。

五、绿色低碳

（十五）绿色算力技术研究及应用

揭榜任务：围绕算力的绿色节能技术突破，面向算力中的任务调度特性、能源使用模式、负载均衡要求等关键要素，研发适应于绿色计算的动态资源调度算法、能耗优化管理系统，以及面向多场景的协同节能机制，突破节能算法的智能化程度，提升算力网络中多节点的能源利用效率。

预期目标：到 2026 年，能耗管理系统实现对算力中心和网络节点的实时监控与节能调度，通过计算节点支持动态调频、动态电压调节，单节点平均能耗降低 30%以上，满足 AI 推理等应用需求。

（十六）企业绿色计算碳感知平台

揭榜任务：建立企业算力中心碳排放度量体系，能够实时、精准地统计企业各个算力中心碳排放，并能将碳排放量分摊到不同的业务部门、应用场景和工作负载，实现精细化的碳排放的管理。同时，基于碳排放的数据，实现

碳感知调度能力，通过在保证业务体验和连续性的情况下将工作负载调度到更加低碳的算力中心，进一步降低碳排放。

预期目标：到 2026 年，围绕千万核级别跨域的算力中心，构建企业级绿色计算碳感知平台，形成一套行业通用的、可精确度量不同类型工作负载碳排放的技术方法和指标体系，通过生态共建形成绿色度量衡标准体系。构建碳感知调度能力，达到算力中心可再生能源比例 30% 的目标。

（十七）冷板式液冷原生整机柜服务器

揭榜任务：面向新一代液冷算力中心，研发冷板式液冷整机柜，包括液冷服务器节点、无源液冷门等，突破高密算力、多样性算力的散热技术及架构要求，实现支持供电总线、网络互联总线、液冷管路可盲插运维的液冷设备，具备液冷机柜及液冷服务器等多级漏液检测能力，有效降低业务中断范围与损失。

预期目标：到 2026 年，液冷整机柜实现 100% 液冷散热，制冷 PUE 低于 1.15。整机柜服务器内部实现全盲插设计，管理模块可实现整机柜功耗管理、漏液检测、资产管理等功能；通用算力单柜功率不低于 20kW，智能算力单机柜功率不低于 30kW，实现不少于 500 台液冷节点的规模落地应用。

（十八）算力中心节能调优平台

揭榜任务：研制高精确度、高仿真效率、多场景覆盖

的算力中心 PUE 仿真平台，突破物理机理模型构建、仿真引擎集群、模型自动生成等关键技术，实现对算力中心不同运行状态下细分时间颗粒度 PUE 的快速、精准评估。研发基于大数据分析技术的算力中心制冷系统 AI 节能优化系统，通过自动化数据治理、自动推理等关键技术，准确匹配制冷需求，在满足可靠性要求条件下实现算力中心制冷系统整体动态实时优化，优化算力中心 PUE。

预期目标：到 2026 年，支持液冷、水冷等至少 2 类典型制冷场景进行能效优化，支持制冷系统和配电系统联合仿真，系统可输出不同负载及运行工况条件下的 PUE 运行曲线、系统设备运行模拟工况等参数，PUE 仿真精度达到 97% 以上。基于能效优化平台，支持 AI 自动推理，小时级策略自动下发，实现对算力中心能耗的可视、可管、可控。通过 AI 能效优化，实现算力中心 PUE 降低 5% 以上，通过算力中心基础设施与 IT 联动节能，实现总能耗降低 5% 以上，在 5 个以上算力中心落地应用。

（十九）新型制冷系统

揭榜任务：研发人工智能节能系统，针对算力中心基础设施的运行调控和环境监测。提出全新自适应算法，突破原有常见算法的局限性，提升数据的分析和处理效果，搭建基于专家经验的人工智能算法数据库，提升包括能耗管理、能源调度、安全监测、故障诊断、辅助运维等功能的节能性、可靠性、经济性。

预期目标：到 2026 年，在满足制冷要求的基础上，

提高冷却系统的可靠性和自适应性，提高能源使用效率、水资源使用效率和运维效率，其中节电率提升 10%以上。支持冷却系统数据采集、标注、治理、存储，具备系统运行异常告警、告警收敛、自动诊断、远程通信、自动控制等功能，支持冷却系统智能化调优、智能化控制的核心能力，并开展不少于 5 个实际业务场景所提供的 AI 节能调优案例。

六、安全可靠

（二十）算力中心智能运维机器人

揭榜任务：研发算力中心智能运维机器人以及智能机器人管理平台，基于云边端三层架构，实现智能机器人在多层、多房间楼宇机房内的设备设施识别、多模态环境感知、精准空间定位、智能人机协同、多任务联合调度等方面的技术与算法优化。支撑机器人在算力中心设施运维和 IT 运营等典型场景的应用，提升巡检质量，促进算力中心运维、运营的降本增效。

预期目标：到 2026 年，实现大型算力中心内智能机器人的多机房、多楼层协同应用部署；机器人巡检任务成功率不低于 95%，设备识别准确率达到 97%，环境巡检召回率不低于 90%，保障算力中心巡检业务持续运行。实现云边端协同调度，支持不同场景下的自主作业，提高任务并发执行效率，促进稳定、安全、可靠、可控的算力中心智能运维体系建设。

（二十一）云边端一体化智能监测平台

揭榜任务：开发高性能云边端一体化系统，研发以智能化终端或机器人为硬件载体、以多算法模型融合和平台工具为软件载体的软硬结合的集中监测管理与运维巡检方案。突破多层级自动化运维、多维度诊断、多平台覆盖、多模型量化等关键技术。构建综合运维健康度数字化评估体系与模型，实现算力设施从规划、设计、建设、部署、运行、维护的全生命周期数字化管理。

预期目标：到 2026 年，建立大规模集群的智能化运维能力，设备实现跨平台及系统稳定性风险和安全风险识别能力，综合视频识别技术等，结构化告警收敛推送，准确率超过 98%。算力设施全生命周期数字化联动，平台自动化流程推进，实现云端直控覆盖超 10 栋算力中心，落地数字化算力中心健康度评估，智能化终端或机器人的自驱动巡检，视频流识别与告警的联动，系统的智能化运维问答，并保障业务服务级别协议（SLA）达标率 99%以上。

七、其他

（二十二）其他算力领域的特色化技术、产品、服务和平台等，应具有技术先进性，技术成熟度较高，产业化前景较好。