

题目编号：XH-202623

面向教科研多场景的异构算力池智能调度关键技术研究比赛方案

一、发榜单位

上海沅熹科技有限公司

二、题目名称

面向教科研多场景的异构算力池智能调度关键技术研究

三、题目介绍

1. 课题背景

高校的算力资源（异构 GPU 集群、CPU 节点）需支撑多类核心场景，但不同场景算力需求差异大、时间分布不均，传统调度采用“先来先服务”模式，容易导致算力闲置与资源争抢并存——科研批量任务占用大量算力导致教学实训卡顿，实时智能服务因算力不足响应延迟，有限算力池的利用效率仅达 30%-40%，难以适配多样化、动态化的算力需求。

2. 课题需求

本课题主要是针对高校四类高频算力应用，根据其固有规律差异，通过优化调度算法实现“算力需求-资源供给”的精准匹配，解决传统调度“分配粗放、响应滞后、利用率低”的核心痛点，让有限算力池支撑多场景最优运行。

核心应用场景固有规律及算力调度需求如下：

（1）科研训练场景（如大模型微调、工程仿真、基因数据分析等）

➤ 算力需求：单任务算力消耗大（需占用 1-8 卡 GPU）、持续时间长（几小时-几天）。

➤ 时间特性：非实时性、可错峰（多集中在夜间、周末）、优先级动态（国家级项目>一般项目）。

➤ 输出要求：结果准确性优先，允许合理延迟，支持断点续算。

（2）教学实训场景（如 AI 课程实验、虚拟仿真教学、学生毕业设计算力支撑等）

➤ 算力需求：单任务算力消耗中等（1-2 卡 GPU）、批量并发数十-数百学生同时使用）。

➤ 时间特性：强时段集中（上课时间、实验课时段）、实时性中等（响应延迟 $\leq 30s$ ）。

➤ 输出要求：稳定性优先，需保证同一批次任务公平分配算力，避免卡顿。

（3）校园智能服务场景（如学术诚信监测、智能问答助手、多模态文献解析等）

➤ 算力需求：单任务算力消耗小（单卡即可）、请求高频且波动大（峰值集中在工作日白天）。

➤ 时间特性：实时性高（响应延迟 $\leq 5s$ ）、7x24 小时不间断运行。

➤ 输出要求：低延迟优先，算力占用弹性可变，支持动态扩容缩容。

（4）跨学科数据处理场景（如多模态科研数据清洗、遥感图像解析、实验数据批量处理等）

➤ 算力需求：单任务算力消耗稳定（中等规模）、离线批量处理（无实时要求）；

➤ 时间特性：可错峰调度（优先占用闲置算力）、任务可拆分（支持分片并行）；

➤ 输出要求：吞吐量优先，不与实时任务争抢算力资源。

3. 具体应用：

适用于高校科学计算中心、教务处、科研院等部门，覆盖科研训练、教学实训、校园智能服务、跨学科数据处理等场景，可推广至科研院所、中小企业异构算力集群管理。

四、参赛对象

青年科技人才赛道：在高等院校、科研院所、企业等各类创新主体中具有较高科研热情和较强科研能力的青年科技工作者可通过青年科技人才赛道申报作品参赛。参赛人员年龄在 40 周岁以下，即 1986 年 6 月 1 日（含）以后出生。

高校教师在指导参赛的同时不得以参赛人员身份参加同一选题比赛。发榜单位及同发榜单位有相关隶属关系单位的青年不得参加本单位选题比赛。

参赛对象可以团队或个人形式参赛，每个团队不超过

10 人，每件作品可由不超过 3 名指导教师进行指导。可以跨专业、跨学校、跨单位、跨地域组队，但同一团队所有成员均应符合本赛道相关年龄、身份要求。每件作品只可由 1 所高等院校、科研院所或企业等作为参赛主体提交申报。

五、答题要求

参赛者需提交作品报告、作品运行效果视频及其他佐证材料（若有）等，并可接受发榜单位的线上答辩。

六、作品评选标准

1. 场景覆盖与需求契合度（20 分）

（1）覆盖科研训练、教学实训、校园智能服务、跨学科数据处理四类场景，体现各场景固有规律（优先级动态、时段集中、7x24 实时、可拆分错峰等）。

（2）对教学实训、校园智能服务给出 SLA 保障策略：教学交互响应/任务启动延迟 $\leq 30s$ ；校园智能服务响应延迟 $\leq 5s$ 。

2. 调度效果与量化指标（40 分）

（1）算力利用率：在提供的混合负载或自建可复现负载下，相比“先来先服务/基线调度”显著提升；目标平均利用率 $\geq 80\%$ 或相对提升 ≥ 2 倍，并给出统计周期与计算口径。

（2）任务效率：科研训练任务周转时间缩短 $\geq 30\%$ ；

跨学科离线任务吞吐提升（完成时间缩短/处理量提升）。

（3）稳定性与体验：教学实训卡顿率降低 $\geq 80\%$ 或等价指标（掉帧/排队抖动）显著改善；高并发下无明显抖动。

（4）公平性与隔离：教学实训同批次任务公平分配；实时服务 SLA 违约率低，批处理不得挤占实时资源。

3. 算法设计与创新性（15 分）

（1）是否结合场景规律进行差异化调度：优先级动态分层、可抢占/可中断、断点续算、弹性扩缩、任务拆分并行、错峰填谷等。

（2）是否引入需求预测、队列建模、约束优化/学习优化等方法，并给出可解释的决策依据。

4. 系统实现完整度与可落地性（15 分）

（1）可佐证的材料（如技术文档、演示视频、截图等），可在异构 CPU+GPU 资源上运行。

（2）可扩展与容错：支持节点扩容、故障恢复、资源回收；满足 7x24 运行场景。

5. 异构适配与自主可控（10 分）

（1）具备跨架构/多厂商 GPU 与 CPU 统一抽象与资源计量能力，支持国产与进口 GPU 混部调度或迁移策略，并给出验证结果。

（2）兼容主流 AI 框架/容器生态，说明驱动版本与依

赖管理方案。

七、作品提交时间

2026 年 5 月至 9 月上旬，各参赛团队选择榜单中的题目开展研发攻关。

2026 年 8 月 31 日 23:59 前，各参赛团队通过邮件方式完成作品提交，具体要求本方案第八点第（二）款，并严格遵照发榜单位明确的提交规范执行。

2026 年 10 月中旬，由发榜单位完成初审，确定入围终审擂台赛的晋级作品和团队。

2026 年 10 月中下旬至终审擂台赛前，中国青基会平安励志计划将发挥资源优势，组织专家团队，为晋级作品和团队提供帮助和指导，帮助各晋级团队完善作品，冲刺攻关参加终审擂台赛，角逐“擂主”。

2026 年 11 月，组织终审擂台赛，角逐“擂主”。

八、参赛报名及作品提交方式

（一）报名方式

（1）参赛选手登录“挑战杯”官网 www.tiaozhanbei.net，在“揭榜挂帅”擂台赛报名入口注册账号，登录大赛申报系统在线填写报名信息。报名信息提交后，下载打印系统生成的报名表。

（2）申报人在报名表对应位置加盖所在单位公章。

（3）将盖章版报名表扫描件上传至报名系统，等待系统

审核。请参赛选手注意查看审核状态，如审核不通过，需重新提交。

（4）系统开放报名时间为 2026 年 5 月 30 日—6 月 30 日，逾期后系统将自动关闭报名功能。

（二）作品提交方式

请已在官网报名成功的团队，于 8 月 31 日前将盖章的参赛申报表 pdf、作品所有相关材料发送至发榜单位邮箱 fengxiaoya@inspur.com。邮件主题与压缩包命名规则：所在单位-姓名-作品名-联系电话。提交具体作品时，务必一并提交 1 份报名系统中审核通过的参赛报名表（所有信息与系统中填报信息保持严格一致）。以上材料无需在“挑战杯”官网提交。

九、赛事保障

1. 上海云熹配备专门指导人员，组织线上技术宣讲与答疑，讲解四类场景规律、SLA 约束、评测方法与提交规范。

2. 建立专属技术支持通道（企业技术指导团队），提供固定“每周线上 office hour”与邮件/群答疑，48 小时内响应并给出解决建议。

3. 开放可申请的异构算力测试环境（CPU 节点+GPU 节点，多租户隔离）；如受条件限制，提供容器化仿真环境与合成负载生成器，确保参赛队伍可离线复现与对比。

4. 提供脱敏后的典型工作负载画像/Trace（或任务模板），覆盖科研训练、教学实训、校园智能服务、跨学科数据处理四

类负载比例与时间分布，便于验证“闲置与争抢并存”的痛点与优化成效。

十、设奖情况及奖励措施

1. 设奖情况

本题为青年科技人才赛道，拟评“擂主”1个，特等奖、一等奖、二等奖、三等奖各5个，最终授奖数量由发榜单位视作品申报数量和质量情况报组委会同意后动态调整。

2. 奖励措施

本题共设5级奖项：“擂主”、特等奖、一等奖、二等奖、三等奖，奖金金额分别为：10万元/个、2万元/个、1万元/个、0.5万元/个、0.2万元/个，以上均为提供给获奖团队的税后奖励。

3. 奖金发放方式

比赛结束后，单位比赛专班工作人员与获奖团队取得联系，填写奖金申请表，待获奖团队提供银行卡详细信息后1个季度内，将奖金一次性发放至获奖团队提供的银行卡中。

十一、比赛专班联系方式

1. 专家指导团队

顾问专家：冯老师，联系电话：19967502505

顾问专家：李老师，联系电话：18798610427

负责比赛期间技术指导保障。

2. 赛事服务团队

联络专员：李老师，联系电话：18917558774

联络专员：欧阳老师，联系电话：18121358997

负责比赛期间组织服务及后期相关赛务协调联络。

3. 联系时间

比赛期间工作日（9:00-17:00）

附：发榜单位简介

上海沅熹是国有企业浪潮集团下属子公司，注册资金 1.5 亿元，公司汇聚了全球顶尖的数据库、大数据、人工智能领域人才，以分布式多模数据库为基础核心产品，面向工业物联网、数字能源、车联网、智慧产业等各大行业领域，提供稳定自主、安全、高性能的创新软件服务与行业解决方案。

公司拥有分布式计算、多模架构、就地计算、超速分析、云边端协同、原生 AI、数据库自治、主动式时序引擎等多项核心技术；产品体系囊括自主研发分布式多模数据库、数据服务平台及配套的生态工具集。上海沅熹公司的主打产品 **KaiwuDB** 作为市面上首款面向 **AIoT** 的分布式多模数据库，在国际权威数据库性能测试榜单 **BenchANT** 获得 **TSBS** 场景性能世界第一。