



工程与应用

算网融合背景下的异构算力调度系统和应用

龚豪

(中国电信股份有限公司上海分公司, 上海 200120)

摘要: 针对我国近年算网融合及算网基础设施发展的背景, 阐述了异构算力调度的系统定位和功能架构, 提炼了系统实现的4项关键组件, 包括屏蔽异构算力差异的算力调度适配器、基于调度算法执行两级调度的作业调度控制器、提供算力感知和算网路由功能的算网资源编排模块、实现算力交易一体化服务的算力服务3中心模块, 针对调度效果多维约束给出了一套多目标调度算法模型。结合“东数西算”、人工智能(artificial intelligence, AI)训推协同、超算互联、视频直播等场景探讨了异构算力调度的重要作用, 结合工程实践给出了实际部署案例, 并对异构算力调度在算网融合演进中的创新和应用前景进行了总结。

关键词: 算网融合; 异构算力调度; 跨域调度; 算网协同

中图分类号: TP319

文献标志码: A

doi: 10.11959/j.issn.1000-0801.DXKX260013

Heterogeneous computing power scheduling system and applications under the background of computing-network convergence

Gong Hao

China Telecom Co., Ltd., Shanghai Branch, Shanghai 200120, China

Abstract: In view of the development of computing-network convergence and computing-network infrastructure in China in recent years, the system positioning and functional architecture of heterogeneous computing power scheduling were expounded, and four key components for system implementation were extracted, including computing power scheduling adapters that shielded the differences of heterogeneous computing power, a job scheduling controller that performed two-level scheduling based on scheduling algorithms, a computing-network resource orchestration module that provided computing awareness and computing-network routing functions, and three center modules for computing power services that realized integrated computing power transaction services. A multi-objective scheduling algorithm model was proposed for the multi-dimensional constraints of scheduling effects. Combined with scenarios such as the East Data West Computing project, artificial intelligence training-inference collaboration, supercomputing interconnection, and live video streaming, the important role of heterogeneous computing power scheduling was discussed. A prac-

收稿日期: 2025-11-19; 修回日期: 2026-04-13

通信作者: 龚豪, gonghao.sh@chinatelecom.cn



tical deployment case was given combined with engineering practice, and the innovation and application prospects of heterogeneous computing power scheduling in the evolution of computing-network convergence were summarized.

Key words: computing-network convergence, heterogeneous computing power scheduling, cross-domain scheduling, computing-network collaboration

0 引言

在数字经济重塑全球竞争格局的背景下,我国通过算力基础设施规划和“人工智能+”战略形成“双轮驱动”,为数字技术与实体经济深度融合奠定基础。2022年,我国正式启动“东数西算”工程,已在京津冀、长三角、粤港澳大湾区、成渝、贵州、内蒙古、甘肃、宁夏形成8大国家算力枢纽,并持续统筹优化算力布局,促进东中西部算力协同,推动算电一体化融合。与此同时,“人工智能+”战略已经成为我国重要国策,人工智能的高速发展带来算力需求的激增,大模型驱动的智能能力和应用成为千行百业经济增长的新引擎。在此战略背景下,算力基础设施和网络基础设施如何有效地融合和协同成为重要命题。作为实现算网融合的重要技术载体之一,异构算力调度能有效提供多元算力资源协同和全局调配的能力,是实现“东数西算”和人工智能普惠落地的重要技术保证。

当前国内各云服务商对于异构算力资源的调度基本仍以静态开通为主,很少能实现根据业务需求自动感知算力和网络特性,动态进行资源编排、发放和部署。本文基于工程实践提炼出在算力网络^[1]中部署的异构算力调度系统架构和功能方案,通过部署实现动态资源调度的核心组件,包括多维算力调度适配器、包含算力感知和算网路由功能的算网资源编排模块、提供算力调度算法和策略并执行调度过程的智能作业调度控制器等,为实现涵盖通算、智算、超算、算网等多元异构联合调度提供了体系框架指引和实践参考。

1 算网融合需求背景

1.1 “东数西算”加速推动算网融合

“东数西算”工程作为我国算力基建核心,自2022年启动以来成效显著,已在全国范围建成八大枢纽、十大集群,截至2025年3月,八大枢纽节点算力总规模达215.5 EFLOPS,智算规模占枢纽节点算力总规模的80.8%,算力枢纽间20 ms时延圈基本实现,为算力高效互联提供基础支撑。2025年5月,工业和信息化部发布《算力互联互通行动计划》,标志着我国算力基建从大型枢纽“单点建设”进入“网络互联”需求阶段,该计划也为“东数西算”工程后续的算网协同发展提供了行动指引,计划提出“两步走”目标:2026年建立建成各级互联平台并开展试点,2028年实现全国公共算力标准化互联,形成算力互联网^[2]。

“东数西算”工程深化发展的下一阶段是实现算网融合,以网络 and 算力的融合发展为基础,实现算网智能协同和全局优化,核心目标是通过计算、存储及网络资源的统一编排管控,综合用户需求和资源情况,高效精准地将需求调度至最佳资源节点。我国算网融合主导力量是通信运营商。中国电信、中国移动、中国联通纷纷提出云网融合、算网融合相关战略,全国范围内建设低时延骨干网、城域光网并与自建算力中心深度绑定,构建“算-网-边”一体化资源编排平台,支持算力池互联互通。在“东数西算”战略推动下,运营商的网络覆盖和数据中心布局直接决定算力资源的集聚程度与调度效率。

1.2 异构算力调度是实现算网融合的重要实体之一

伴随着我国“东数西算”工程的逐步深入及行业各界对算力的投入和储备,我国算力总体规模发展可观,已经成为世界第二大算力大国。但总体而言,我国算力发展仍呈现如下痛点:(1)算力碎片化严重、广泛存在云边端各规模算力;(2)东西部算力发展不均衡,算力利用率高低不均;(3)囿于各算力芯片架构的非标性,异构算力无法高效协同;(4)单智算集群满足大模型发展有瓶颈,多智算中心分布式训练、业务应用调用多个垂类模型的场景在增加,迫切需要实现多智算集群逻辑一体高速互联;(5)算网深度融合仍需较多技术突破,分布式智算中心无损网络技术^[3]、跨域远程直接存储器访问(remote direct memory access, RDMA)技术、算力中心对网络性能及状况感知技术、跨域动态调度和协同机制、算网安全技术等均有待进一步突破。

发展异构算力调度是解决以上部分痛点的重要路径之一。当前网络中存在多种算力资源且分布广泛,包括以中央处理器(central processing unit, CPU)为主的通用算力、以图形处理器(graphics processing unit, GPU)和人工智能(artificial intelligence, AI)芯片为主的智能算力、超级计算机集群提供的超算算力、量子计算为代表的新型算力等。通常情况异构算力指由指令集、核心架构、内存模型、计算单元互联协议等方面存在差异的算力芯片组成的算力集群。异构算力调度技术是能够屏蔽异构算力底层差异,通过统一的资源建模、状态感知进行资源协同,将各类计算任务智能地分配至最优算力集群的技术;为保证任务执行的带宽、时延、抖动等网络指标满足应用需求,对网络资源的协同感知和网络路由的调度也是异构算力调度体系需要实现的重要环节。

异构算力调度本质上是应对多算力类型应用场景中的融合问题,核心目标为实现跨云商、跨架构、跨网络的多云接入、资源管理和算网协同

调度需求,实现专线入算、算间互联、算力资源的分配和调度、整合一点计费 and 一站式算网管理能力,支持对外提供统一服务。

在理论研究方面,异构算力调度技术包含从底层硬件资源抽象到智能决策算法再到上层应用编排的复杂技术栈,关键技术包括异构算力感知与度量技术、跨域协同调度技术、全局元数据技术、分层分布式调度架构技术、多目标优化调度算法、算网编排智能化技术、数据溯源及算网安全技术等。国内外多家研究机构给出了多项与此相关的理论技术模型^[4-8]。

2 异构算力调度系统定位

尽管异构算力调度具备了较多模型理论进展,但在实际应用中仍面临严峻挑战。例如,算力的统一度量和感知现网实践不足,异构算力存在硬件体系、指令集标准、编程模型和软件栈等显著差异,其统一度量和标准化受到各算力集群天然兼容度差、接口开放较封闭的限制,进展多侧重在理论阶段;资源动态适应性不足问题存在,随着系统节点规模和异构性增加,调度策略和算法难以适应并行任务扩展、任务间干扰和数据依赖的复杂性;在超大规模集群调度场景中还须强化与网络拓扑、RDMA无损传输、跨域带宽等资源联动优化,才能真正实现异构算力间的高效协同等。这些都成为算网调度发展的核心挑战。

面对上述挑战,基于异构算力调度技术理论基础,在工程实践领域通常借助在网络中部署异构算力调度系统来实现对算力及网络资源的一体化管理调度。异构算力调度系统伴随运营商跨域骨干算力网络的发展,可基于算力动态感知和作业智能分发调度实现跨云商、跨架构的算力调度功能,服务于“东数西算”国家战略。

异构算力调度系统定位示意图如图1所示,算力调度系统南向通过算力互联网络连接各算力资源池,北向通过服务门户向行业客户以及生态

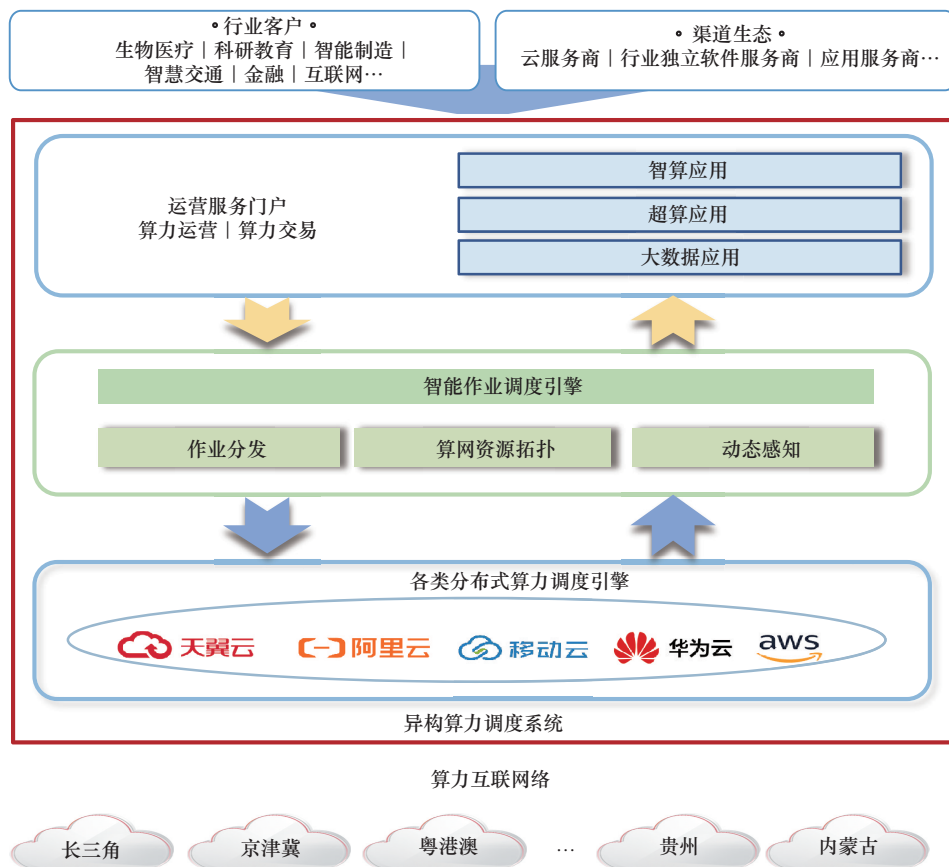


图1 异构算力调度系统定位示意图

合作伙伴等提供调度管理功能服务。算力调度系统可根据管理规模和区域范围大小，在网络中集约部署或分层部署。

借助异构算力调度系统可实现的算力互联互通场景功能包括以下内容。

(1) 异构算力管理

系统通过对接云服务商或算力服务商的应用程序接口（application program interface, API），通过适配层实现第三方算力的统一纳管，屏蔽底层架构差异，实现各类应用在异构云池及算力池间的灵活部署。客户可以通过系统一键订购、开通、纳管及出账，并可实现和云服务商、算力服务商的结算。

(2) 算间互联服务

通过算间网络承载，满足应用系统在多个异

构算力资源间的互访与数据传递需求，实现算间数据调度。

(3) 多云接入

通过采用通信运营商各类专线产品，连接专线传输网络平面路由和调度管理网络平面路由，打通客户专线端到达云内网络接口的逻辑通道，形成以专线方式接入多个云商资源池客户内部网络的能力。

(4) 企业分支互联

用户通过云专网或采用软件定义广域网（software-defined wide area network, SD-WAN）组网上云服务，实现驻地或分支机构接入企业云侧内部系统，全过程由异构算力调度系统网络管理平面自动分发配置并跟踪交付^[9]。

3 我国在异构算力调度领域的实践现状

近年来，在数字中国和算力发展的政策牵引下，各地政府、运营商、研究机构、算网服务企业等在异构跨域算力调度领域进行了诸多实践。近3年我国各地发布的部分重要算力调度平台/调度中心见表1，可以间接地看出东西部均在算力和算网调度领域有投入但不均衡，东部地区的算力调度和算网协同实践活跃于西部。

从表1可以看出，已建的算力调度平台或调度中心大多落在全国一体化大数据中心枢纽节点，聚焦区域内资源统筹管理、调度及实现算力交易。后续算力调度性能监测将成为投建重点，监测范围除算力性能外，更为迫切的是实现全网网络性能监测。当前各地普遍缺乏统一的入算、算间、算内网络的容量、性能指标评估体系，并未通过统一的平台监测到全网算力网络时延、光网络端口利用率、IP网络带宽利用率、光传送网（optical transport network，OTN）覆盖率等指标，无法实现各地算力网络的实时总览。算网协同调度将成为各地在算力调度领域发力的重要目标之一。

4 异构算力调度系统功能架构和关键组件

4.1 异构算力调度系统功能架构

异构算力调度系统基于各类型资源池采集的算力和集群信息，通过智能调度策略实现符合业务的资源池感知及作业智能分发，形成核心调度能力。系统向下基于资源池算力调度适配器，接入通算、超算、智算资源池，基于专线网络实现资源池间按需连接。系统向上为行业客户及生态伙伴提供统一的公共服务门户及相关应用，为企业客户上云入云管云提供智能化和自动化服务。

系统设计采用分层架构支撑各类算力调度需求，异构算力调度系统功能架构如图2所示，主要分为统一门户层、算力服务层、作业调度层、调度适配层、支撑管理层。统一门户层为行业客户及生态伙伴提供统一的公共服务门户及相关功能，同时提供统一的运营运维操作门户。算力服务层主要由应用中心、计费中心以及资源中心组成，向客户提供应用服务及模板，结合数据、客户自定义程序等形成可以被计量的用户作业任务，提供统一视图的资源管理，同时满足算力调

表1 近3年我国各地发布的部分重要算力调度平台/调度中心

序号	发布时间	发布区域	算力调度要素平台
1	2023年2月20日	上海市	上海市人工智能公共算力服务平台正式揭牌，依托上海超算中心建设并运营，支持对公共算力和各种商业算力进行调度
2	2023年2月24日	宁夏回族自治区	宁夏回族自治区银川市正式上线“东数西算”一体化算力服务平台
3	2023年4月25日	粤港澳大湾区	粤港澳大湾区算力调度平台正式发布，由粤港澳大湾区大数据中心设立，探索算力调度和服务实践。平台已被纳入算力交易市场体系核心
4	2023年6月5日	北京市	中国信通院联合中国电信共同发布我国首个实现多元异构算力调度的全国性平台——全国一体化算力算网调度平台
5	2024年3月4日	北京市	中国电信推出天翼云4.0算力分发网络平台“息壤”，在算网资源标准化、算网编排、算网运营3个方面实现技术创新突破，与主流服务商并网
6	2024年5月22日	安徽省	安徽省算力统筹调度平台正式发布，是全国首个通算、智算、超算、量算“四算合一”统筹调度平台
7	2024年5月29日	贵州省	全国一体化算力网络国家（贵州）枢纽节点调度平台新版本上线
8	2024年6月26日	江苏省	江苏省重大项目“东数西算”长三角算力调度中心在苏州吴江启用
9	2025年7月28日	上海市	由上海仪电牵头建设的市级算力调度平台，在世界人工智能大会发布2.0版本，该平台涵盖算力超市、算力纳管、算力融合3大功能模块，纳管算力规模已近20 EFLOPS
10	2025年11月13日	江苏省	江苏省一体化算力调度监测平台正式上线，采用“1+13”省市两级协同运营体系，已监测全省282个数据中心

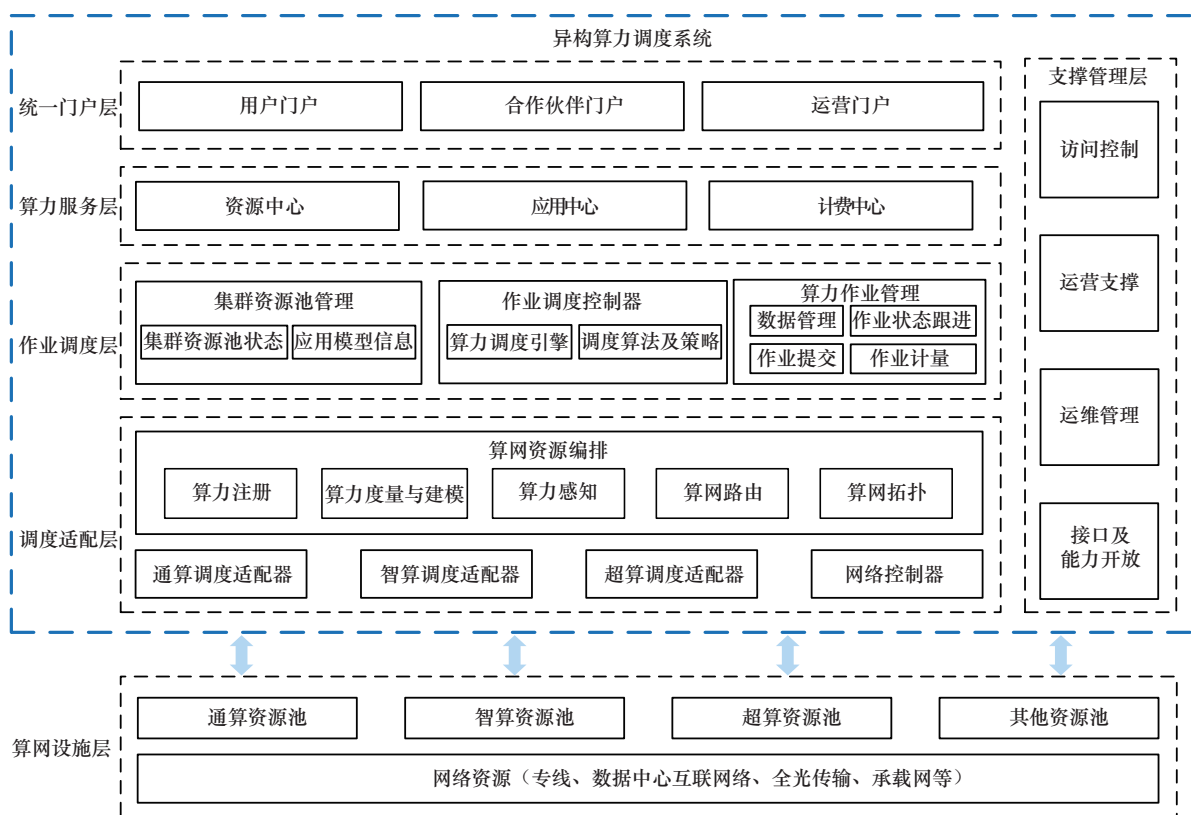


图2 异构算力调度系统功能架构

度和交易所需的计费、定价、账单管理等运营功能。作业调度层通过动态感知、动态更新算网资源拓扑图，基于算力作业调度技术，实现对用户作业任务的路由分发。调度适配层通过通算、智算、超算调度适配器对接异构算力资源，实现向异构资源池接收作业、推送资源及作业状态等，通过网络控制器实现网络连接调度和网络配置下发，并提供算力资源编排和开通能力。支撑管理层提供访问控制、运营支撑、运维管理、接口及能力开放等功能。

架构底层是提供算网资源的算网设施层，涵盖系统纳管和接入范围内的各通算、智算、超算及其他算力资源池，系统纳管和接入范围内的各类专线和网络能力。通常意义上，算网设施层不归入异构算力调度系统本身范围内。

在架构中，实现一体化交易和服务能力的资源中心、应用中心、计费中心，提供算力调度引擎及

算法库的作业调度控制器，实现跨域资源感知、算网路由、多方协同编排的算网资源编排模块，实现标准化适配异构算力的调度适配器等4项关键组件是其区别于传统资源管理平台的关键所在。

4.2 异构算力调度系统关键组件

4.2.1 承担全局资源管理、统一服务和统一计费的算力服务3中心

算力服务层位于异构算力调度系统上层，是连接底层复杂、异构的算网资源与最终用户、应用程序之间的关键桥梁。在算力网络的全局调度体系中，该层不再仅仅管理单一资源池的服务目录，而是承担着汇聚全局资源信息、提供标准化服务接口、支撑跨域计费与交易的核心职能。

模块组成包括资源中心、应用中心和计费中心。资源中心主要是形成和维护覆盖算力网络中所有已注册资源的全局、统一视图。资源中心接收来自调度适配层上报的、关于不同地域、不同

提供商、不同技术架构的算力节点的标准化信息,技术架构包括CPU、GPU、神经网络处理器(neural processing unit, NPU)、高性能计算(high performance computing, HPC)等,信息包括资源类型、规模、属性、位置、状态等。在汇聚异构信息的基础上形成全局算力资源目录,可供查询和上层调度决策使用,是调度引擎获取全网可用资源信息的唯一信息源。在此基础上,资源中心提供逻辑资源池管理功能,可按地域、服务等级、所有权划分为公共或专属资源池,按算力类型分为通算、智算或超算资源池等,形成租户维度的全局资源分类管理和访问控制。

应用中心主要实现基于应用模板按不同流程进行应用构建和部署。应用中心提供标准化服务类型定义,如“AI模型训练服务”“AI模型推理服务”“通用计算实例服务”“超算作业服务”等,用户按服务需求模板提出需求。接收到用户提交的任务后,应用中心进行初步解析和验证,然后将其传递给作业调度层进行全局资源匹配与调度。应用中心还可嵌入各类数据库及中间件、智算框架、超算框架等,以方便用户基于应用模板快速构建应用。

计费中心主要是为算力交易提供统一的异构算力计量标准、灵活的多维度定价模型、支持跨域的订单和账单管理,最终目的是提供清晰、准确、可信的计量计费来实现资源共享和交易。

4.2.2 提供算力调度算法和策略并执行调度操作的作业调度控制器

当前国内各云服务商对于自身算力资源和网络资源的调度基本上仍以静态开通为主,即根据客户或应用需求,在算力资源池进行指定资源的开通、部署和所需网络的配置下发、调测,还未能做到根据业务场景特性,动态地进行资源的自动编排、发放和部署。要做到动态调配,必须具备智能化的作业调度控制层。作业调度层是调度系统的核心决策与执行管理中枢,承接来自算力服务层

的用户作业请求,并基于全局算力资源感知能力,运用智能调度算法,决定将作业分配到算力网络中的哪些资源池执行,并负责管理作业在跨地域、跨主体、异构环境中的执行生命周期。

汇集算力调度引擎和算力调度算法的作业调度控制器是作业调度层最核心的功能。算力调度引擎分为两级调度:一级调度为作业调度,由系统根据用户的作业部署需求,根据调度算法和策略的调度结果分发至目标算力资源池集群,用户作业任务将在目标算力资源池集群队列中排队;二级调度为资源调度,系统通过调度适配层的各调度适配器和目标算力资源池集群形成交互,最终完成作业在资源上的调度和部署^[10]。算力调度引擎工作的底座是算力调度算法及策略库,支持采用动态配置方式实施,可组合使用多种调度策略,按照配置优先级匹配到合适的目标算力节点。

根据业界实践,调度算法和策略最常用的包括成本优先、时延优先、负载优先、能耗优先等。成本优先策略主要根据算力价格进行调度,低价格优先确保作业以更低成本执行,高价格优先则确保作业在更佳的算力集群上快速执行。时延优先策略按照算力响应时延确定算力访问路由,优先调度网络距离近的计算节点。负载优先策略按照算力资源池资源空闲度进行调度;在满足资源需求总量的资源池中优先调度当前算力资源冗余高的集群。针对单一作业在物理机层面的异构计算节点并行计算,湖南大学研究团队提出了一种时延约束下最小资源投入的配置算法^[11]。在能耗和时延联合调度方面,华南理工大学研究团队针对以最小化时延和能量消耗为目标的联合计算卸载和资源分配问题,提出了一种协同任务拆分和资源分配优化算法^[12]。在负载均衡方面,山东科技大学研究团队针对移动边缘计算网络中用户的各种运动模式,提出一种依据固有的移动性规律和数据在连续时间段内随机到达并最大化网络系统数据处理能力的算法^[13]。东南大学研究



团队综合考虑响应时延和执行成本等因素,提出了一种为具有紧迫截止期的任务确定任务优先级并决定对应的容器完成任务部署的算法^[14]。

基于工程实践和算法简化高效原则,本文提出了一种面向异构算力调度多维约束的多目标优化模型。该模型综合考虑算力成本、网络时延、资源负载均衡、能耗4个核心维度,在多作业调度场景中实现调度决策的量化优化。

1) 问题建模与目标函数

设算力网络中包含 m 个异构算力节点,记为 $N=\{n_1, n_2, \dots, n_m\}$,每个节点 n_i 具有算力容量 C_i (以FLOPS度量)、当前负载 L_i (以FLOPS度量)、单位时间算力价格 p_i (元/s)及地理位置坐标 loc_i 。对于某作业请求 J ,设算力需求为 R_j (以FLOPS度量),数据量为 D_j ,时延敏感度为 λ_j ($\lambda_j \in [0, 1]$,值越大表示时延越敏感)。

定义二元决策变量 $X_{ij} \in \{0, 1\}$,当作业 J 调度至节点 n_i 时, $X_{ij}=1$, $X_{ij}=0$ 。

2) 建立多目标优化模型

$$\text{Minimize } F(J) = \omega_1 \times \text{Cost}(J) + \omega_2 \times \text{Latency}(J) + \omega_3 \times \text{Loadbalance}(J) + \omega_4 \times \text{Energy}(J) \quad (1)$$

其中, $\omega_1, \omega_2, \omega_3, \omega_4$ 为各目标权重系数,满足 $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 1$,可取定一组默认权重系数,应用有指定偏好时可根据需求动态调整。

各子目标函数定义如下,成本目标:

$$\text{Cost}(J) = \sum_{i=1}^m X_{ij} \times p_i \times T_{ij} \quad (2)$$

其中, T_{ij} 为作业 J 在节点 n_i 上的预计执行时间,按照 $T_{ij} = R_j / (C_i - L_i)$ 计算。

时延目标:

$$\text{Latency}(J) = \sum_{i=1}^m X_{ij} \times [T(\text{loc}_u, \text{loc}_i) + D_j / B_{ij}] \quad (3)$$

其中, loc_u 为用户位置, $T(\text{loc}_u, \text{loc}_i)$ 为用户位置到算力节点 n_i 的传输时延, B_{ij} 为作业 J 在 n_i 上的可用带宽。

负载均衡目标:

$$\text{Loadbalance}(J) = \sigma^2(L_1/C_1, L_2/C_2, \dots, L_m/C_m) \quad (4)$$

即各算力节点负载率方差,用于衡量系统负载均衡程度, σ^2 代表方差运算。

能耗目标:

$$\text{Energy}(J) = \sum_{i=1}^m X_{ij} \times e_i \times T_{ij} \quad (5)$$

其中, e_i 为节点 n_i 的单位时间能耗系数。

3) 约束条件

资源约束: $X_{ij} \times R_j \leq C_i - L_i, \forall i \in [1, m]$,确保调度后节点负载不超过其容量。

唯一性约束: $\sum_{i=1}^m X_{ij} = 1$,每个作业只能调度至一个算力节点。

时延约束: $\text{Latency}(J) \leq T_{\max} \times (1 - \lambda_j)$, $T_{\max}$ 为最大可接受时延,可根据业务时延敏感度动态调整。

4) 算法实践表现

上述多目标优化算法采用改进NSGA-II^[15]进行求解,工程实践表明,在完成大量作业调度情形下,该算法在成本优化、时延优化及算力节点资源利用率等关键性能指标上均优于只用随机调度、最小成本优先和最短时延调度等算法。当然,上述算法模型忽略了作业并行效率损失、作业干扰、数据传输损耗等因素,在具体应用时可作为基础理论模型进行参考和修正。

4.2.3 提供多元算力纳管对接的调度适配器

算力资源池作为支撑各类计算任务的核心基础设施,其类型涵盖通用计算、智能计算与高性能计算等多种形态。然而,各类资源池在接口开放程度、协议规范及功能支持方面存在较大异质性,导致系统集成与资源统一调度的复杂度增高。为此,调度适配器作为中间层组件,目的是通过标准化接口设计与扩展机制,实现对不同资源池以及网络资源的快速接入与统一管理。针对不同资源,调度适配层分为通算调度适配器、智算调度适配器和超算调度适配器等算力调度适配

器及网络控制器。

算力调度适配器通过 3 类核心接口，即算力节点信息查询接口、集群状态信息查询接口与应用模型信息查询接口，定期采集底层资源状态数据，并对上层调度系统提供统一的资源视图。该机制不仅支撑作业调度层的决策，还可通过门户界面向用户提供多维度的资源状态展示。

通算调度适配器采用服务端-代理端架构实现与各类通用算力资源池的对接。其中，服务端负责接收并分发算力作业，代理端负责资源信息采集、作业接收、状态跟踪与执行反馈。为支持主流云服务商的资源接入，适配器可基于通用算力代理端接口规范及云商弹性容器实例（elastic container instance, ECI）接口^[16]，进一步封装验证鉴权、容器组管理、实例监控、镜像缓存等关键功能。基于该设计，系统可实现多云容器调度能力，支持应用在多个云资源池之间动态调度、部署与跨云互访。

智算调度适配器在功能逻辑上与通算调度适配器相似，主要实现与智能计算平台（如，华为云 ModelArts、阿里云 PAI 等）的深度集成。该模块一方面接收来自算力调度引擎下发的作业数据，包括训练集、模型与代码等，并附带作业编号等信息；另一方面，负责采集智算集群的资源状态、作业执行进度、日志与性能指标。由于不同智算平台在接口规范与调用方式上存在差异，系统可采用嵌入软件开发工具包（software development kit, SDK）或直接调用开放 API 的方式实现对接。

高性能计算领域已存在多种成熟的集群管理与作业调度系统，如，简易 Linux 资源管理工具（simple Linux utility for resource management, SLURM）、负载共享设施（load sharing facility, LSF）、太阳网格引擎（sun grid engine, SGE）、便携式批处理系统（portable batch system, PBS）等，它们广泛用于超算环境中作业的调度与管

理。超算调度适配器通过实现对上述各类调度工具的兼容接入，或将其直接集成至系统中，以复用其高效的资源分配与作业管理能力。

网络控制器以软件定义网络（software defined network, SDN）控制器为代表，主要负责实时获取纳管范围内的网络设备配置信息、全局网络拓扑和链路状态信息，作为算网资源编排的网络信息输入；并可按需创建或调整网络设备配置、网络连接及路由策略等。

4.2.4 实现跨域资源感知、算网路由、多方协同编排的算网资源编排模块

算网资源编排模块作为异构算力调度系统的核心中枢，构建于算力调度适配器与网络控制器之上，承担跨域异构算力与网络资源的全局统筹、动态管控及协同优化职责。算网资源编排模块共分为算力注册、算力度量与建模、算力感知、算力路由和算网拓扑几项子功能，实现“跨域资源感知—算网协同路由—多方联动编排”的流程闭环。算力注册为新算力资源的接入提供标准化接口；算力度量与建模提供标准化的原始数据，支撑资源中心进行全局算力的建模和度量；算力感知通过感知协议实时了解网络中的算力动态信息；算力路由将算力路由决策转化为对具体计算和网络资源的配置指令，并下发至算力调度适配器与网络控制器实现配置落地；算网拓扑整合算力和网络信息，提供全局性的可视化、可叠展的资源拓扑功能。算网资源编排模块是推动算力调度从静态开通走向动态调度的关键网元。

5 异构算力调度的应用场景

聚焦异构算力调度在各行业场景中的应用，主要有如下几类典型场景。

（1）“东数西算”类场景

“东数西算”类场景主要响应国家战略，通过全局算力调度均衡东西部算力生态，推动区域经济协调发展，如，东数西训、东数西存、东数



西备、东数西渲等。因为东西部之间传输距离一般长达几千公里，西部部署的应用或系统大部分为对访问时延、实时响应要求不高的，以不影响用户感知。近两年，各运营商牵头在东西部跨上千公里海量数据无损传输、存算分离拉远大模型训练等方面进行试点和探索，上海人工智能实验室也已推出 DeepLink 超大规模跨域异构混训技术方案，旨在通过跨地域智算中心互联实现大模型的高效训练。

(2) AI 模型训推协同等场景

基于异构算力调度平台的资源协同能力，并依靠公共算力、行业算力以及城市间算力互联网的覆盖优势，可实现“算力产品化供给+广域协同调度”的云网一体新特色，支撑“一云训多边推”、“A地训B地推”等训推协同场景。该场景伴随行业数字化转型及智慧城市应用的发展日益增多。以工业云网大模型应用^[17]为例，工业云网大模型云边端协同架构示意如图3所示。一方面，异构算力调度平台可依据任务特性匹配“端-边-云”异构算力，具体调度逻辑如下：端侧工业相机、采集器等终端仅分配“数据采集+传输”任务，通过万兆接入网络，如50 G无源光网络（passive optical network, PON），将原始图像数据无压缩/浅压缩传输至本地存储设备或边缘计算

节点，避免端侧复杂计算；本地设备/边缘计算节点完成实时性要求较高的模型推理、大模型蒸馏及本地模型场景适配与更新，并生成可使用的各类模型组件；云端算力则承载计算密集型任务，部署大模型训练或云侧推理/统计分析应用等。另一方面，平台可同时感知算力状态与网络状态，在异常情况发生时实现联合优化动态调度：当网络临时拥塞时可端边算力分配动态调整为“端侧浅压缩+边缘解压缩”来平衡传输效率与计算开销；当边缘算力负载过高时，可将部分非实时任务，如模型组件生成等，卸载至云端，同时启动网络服务质量（quality of service, QoS）保障实时推理任务的带宽与时延；云端训练与边缘推理任务开启网络切片隔离，避免模型同步流量抢占工业实时业务的网络资源，保障系统稳定性。

(3) 高速率弹性无损智算网络等场景

当前运营商通过现场试验已基本在现网引入了端到端400 GE大带宽、RDMA广域无损、智能调度等核心技术，搭建出400 G弹性无损智算网络架构，400 G弹性无损智算网络架构示意如图4所示，实现超百公里TB级数据分钟达、远距拉远训练算效达90%以上。弹性无损智算网络通过部署精准流控（subscriber priority-based flow control, SPFC）、自适应流重路由（adaptive

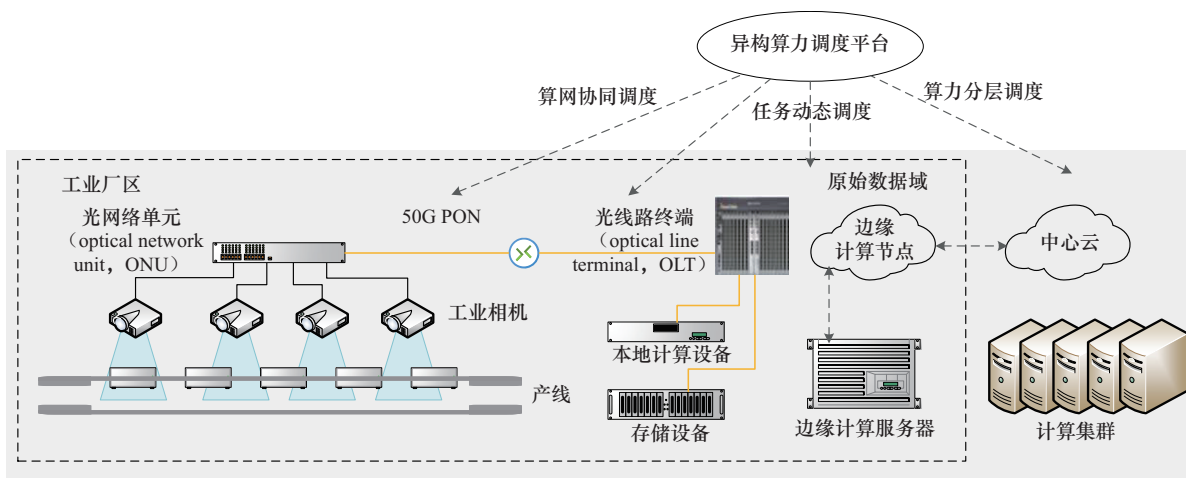


图3 工业云网大模型云边端协同架构示意

flow repathing, AFR) 等技术, 解决了远距广域网络的流量拥塞、无法感知流大小, 导致大象流、老鼠流哈希不均的问题。这些技术的实现将有助于为企业提供跨域协同训练、存算分离、弹性入算等服务。在以上网络技术突破的同时, 调度系统可以灵活调度多地智算中心, 满足“物理分散、逻辑统一”的多地协同高并发训练需求^[2]; 还可以通过调用云端算力实现企业本地样本数据边训边传的存算分离拉远服务, 减少低效的线下数据传递, 提高模型训练效率; 借助算网专线接入广域无损智算网络, 运营商可

通过调度系统按需调用智算、通算、超算等多种异构算力资源, 满足企业一跳入多算^[18]的需求。

(4) 超级计算等科研场景

科学计算、基因测序、金融和交易数据分析、航空模拟建模等领域对算力需求密集, 通过异构算力调度技术实现对超算资源的统一纳管、对计算和传输资源的协同调度, 提升科学大数据的算网协同计算效率。2023 年, 科技部已启动国家超算互联网部署工作, 构建一体化超算算力网络和调度服务平台, 目的是将我国

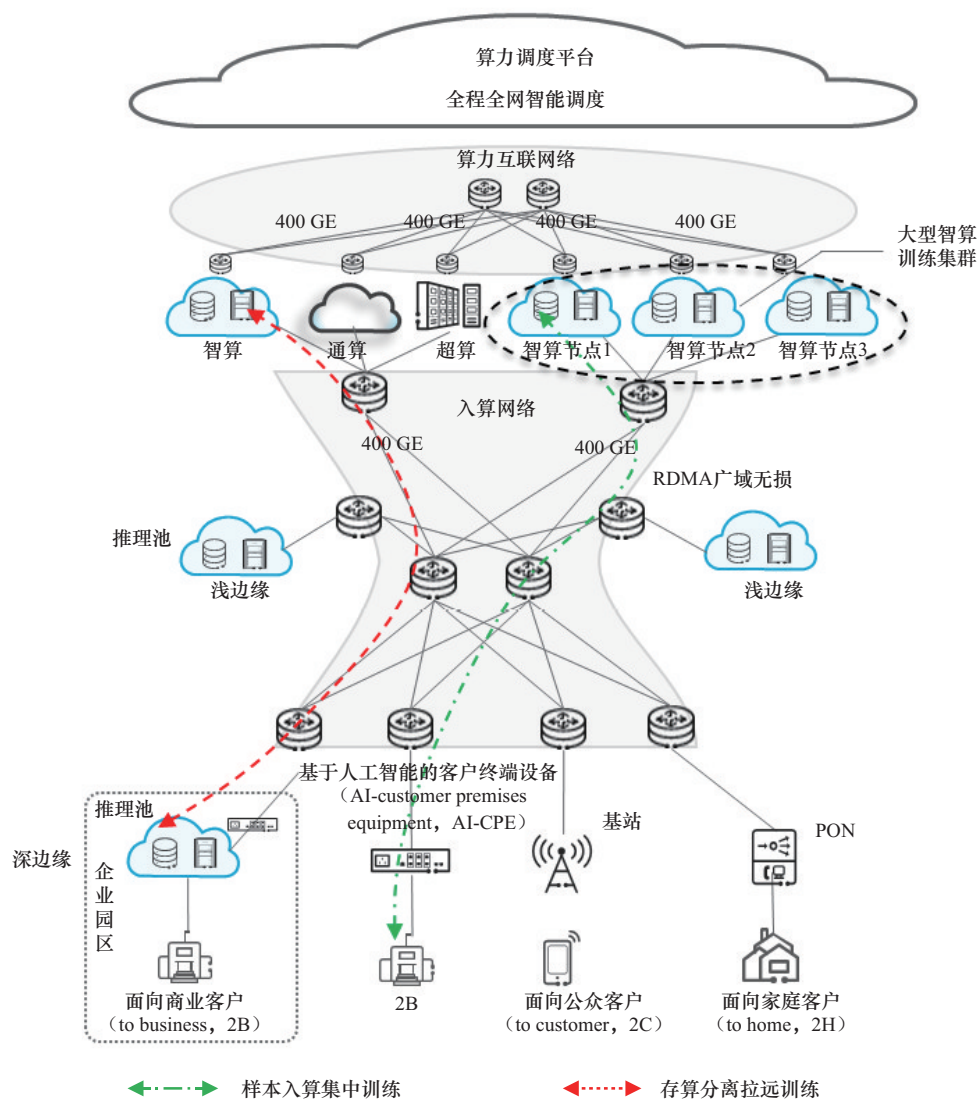


图4 400 G弹性无损智算网络架构示意



各超算中心连接起来, 促进超算算力一体化运营。

(5) 视频直播、视联网、车联网等场景

视频直播、视联网、车联网类场景中普遍存在终端广泛化、AI化, 接入带宽、端边协同要求高的特征。针对视频直播接入带宽需求大、流量波动性变化大的特点, 异构算力调度平台通过全局资源视图和预测算法, 可提前部署调度合适的边缘计算节点, 实现算力资源的弹性伸缩, 以应对流量洪峰。针对视联网场景, 异构算力调度平台可精准感知每个终端或边缘节点的实时任务负载, 如人脸识别、车辆违章分析或突发事件的智能检测, 并综合任务的紧急程度和复杂度进行智能决策, 将实时性要求极高的分析任务调度到最近的边缘节点进行处理, 实现毫秒至微秒级响应。

6 异构算力调度系统工程部署案例

本节以某运营商在枢纽节点部署的某全国异

构算力一体化调度平台为例说明系统的实现方案以供参考, 系统建设者也可考虑自身需求及需纳管资源池情况采用不同的部署方案。

该调度平台支持纳管各云服务商、算力运营商数据中心的智能算力, 包括国产加速芯片、英伟达等, 支持对接公有云算力池以及私有化智算池, 可调度算力达到EFLOPS级别。调度平台具备算力注册能力, 主要通过与云管进行API对接及部署容器两种模式完成资源纳管。算力提供商可提供API在调度平台上触发用户同步, 与调度平台账号关联, 也可采用部署容器的方式, 即在裸算力资源上部署容器组件进行算力注册接入调度平台, 二次封装成销售品, 供通算、智算、超算等算力场景使用。某枢纽节点全国一体化算力调度平台部署示意图如图5所示。

某枢纽节点全国一体化算力调度平台部署资源见表2。平台采用区域内集约部署方案, 部署于运营商集约云资源池, 采用虚拟机全云化部署

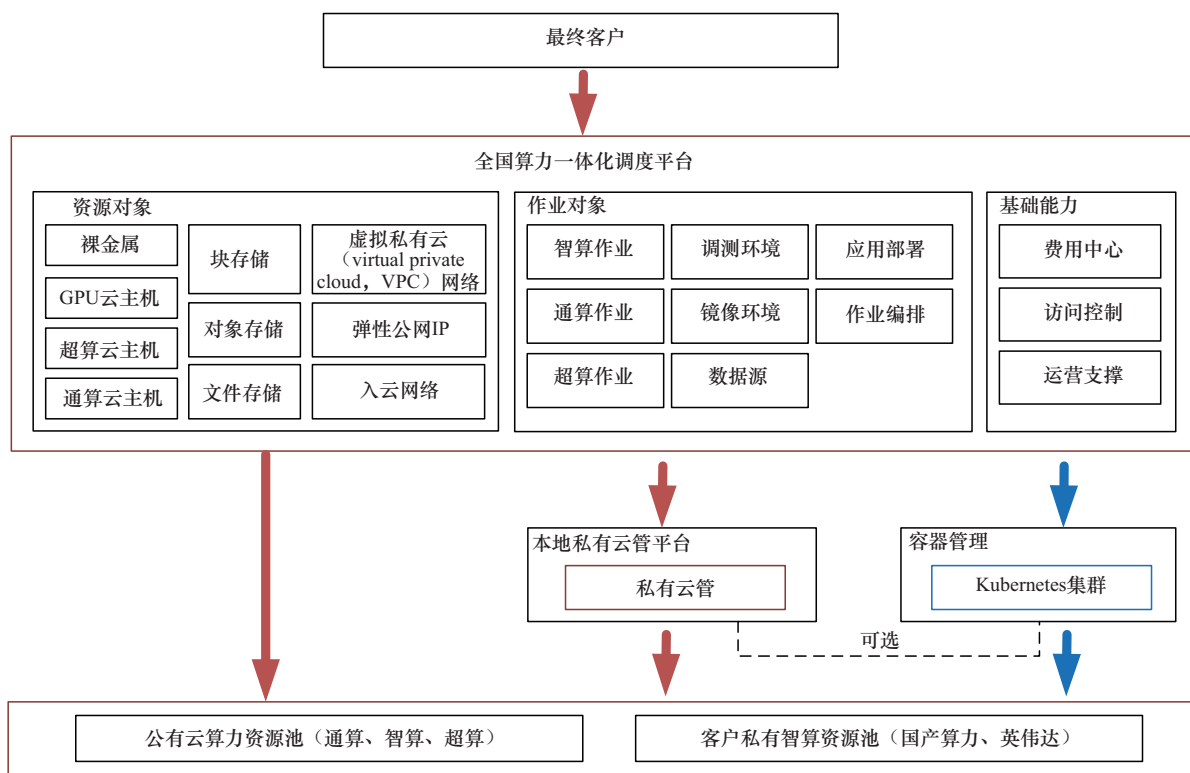


图5 某枢纽节点全国一体化算力调度平台部署示意图

表2 某枢纽节点全国一体化算力调度平台部署资源

用途	配置	数量
门户WEB服务器	至少8vCPU/32 GB内存/200 GB硬盘	2台
接口服务器	至少8vCPU/32 GB内存/200 GB硬盘	至少2台，视纳管算力池数量及规模而定
业务应用服务器	至少8vCPU/32 GB内存/200 GB硬盘	至少4台
数据库服务	至少8vCPU/32 GB内存/1 TB硬盘	2台
中间件服务器（消息、缓存、队列中间件等）	至少8vCPU/32 GB内存/200 GB硬盘	至少8台
共享存储	可采用块存储、网络附接存储（network attached storage, NAS）或多级存储，容量按照需存储数据规模而定	至少1套

方式，无须建设独立的硬件平台，支持规模集群组网和平滑扩容。

自2023年平台正式上线以来，整合了全国数十家服务商的异构算力资源，兼容8大国产AI芯片与国际主流芯片，打破公有云与私有化智算池的算力孤岛，实现资源集约化管理，平台上线前后效果对比如下。

（1）算力利用率提升

平台上线前，各算力资源池独立运营，缺乏统一调度，整体算力利用率仅为30%左右。平台上线后，通过全局资源视图和智能调度算法，实现了跨池资源的动态调配，整体算力利用率提升至60%以上，提升幅度达30%。

（2）调度效率提升

伴随400G OTN、全光交叉（optical cross-connect, OXC）、基于IPv6的分段路由（segment routing over IPv6, SRv6）、广域RDMA等算间网络技术在运营商骨干网络投入应用，平台实现了全国异构算力毫秒级动态调度，关键调度指标如下：枢纽集群内端到端调度时延控制在10 ms内，跨省资源调度时延控制在30 ms内；任务部署响应时间缩短至分钟级，而传统人工开通方式需数小时；日均调度作业成功率达到99%，失败作业主要是资源不足或用户配置错误导致。

（3）成本节约

平台通过灵活的算力服务计费中心功能，一

方面上线AI算力券功能，另一方面支持分钟级精度分卡时计费，支持企业多方式接入算力资源，直接降低企业算力获取成本。

（4）接入效率提升

平台采用API对接与容器部署两种轻量化接入模式，显著地降低了算力提供商的接入门槛。借助标准化的多维算力调度适配器和自动化配置工具，接入周期从传统的2~3周缩短至2~3天，效率提升约90%。

（5）网络协同增强

平台通过网络控制器适配和算网资源编排功能，优化算网联合调度效果。通过算网路由功能，确保传输时延低于应用时延阈值，同时可引导非时延敏感业务向网络带宽利用率低的节点迁移，提升网络带宽利用率。

7 结束语

“东数西算”、人工智能等国家战略推动着算网融合技术体系的发展，而异构算力调度技术是其中一项重要组成部分。随着算网融合在各行各业的深度推进，异构算力调度能够助力解决多种复杂场景中的跨域、跨架构的算网资源和作业协同问题。异构算力调度架构中的重要技术创新包括面向通算、智算、超算等多维度算力的调度适配器，提供算力感知和算网路由功能的算网资源编排，实现算力调度算法和策略并执行两级调



度的作业调度控制器等。异构算力调度契合了AI时代下互联网应用对算力高效互联和智能协同的核心诉求,对于提高我国异构算力综合利用率、推动产业数字化迈向新阶段具有重要意义。

参考文献:

- [1] YD/T 4255-2023. 算力网络总体技术要求[S].
YD/T 4255-2023. General technical requirements for computing power network[S].
- [2] 工业和信息化部. 算力互联互通行动计划: 工信部信管(2025)119号[R]. 2025.
Ministry of Industry and Information Technology. Action plan for interoperability of computing power: MIIT Letter [2025] No. 119[R]. 2025.
- [3] 中国电信股份有限公司研究院. 《分布式智算中心无损网络技术白皮书》[R]. 2024.
Research Institute of China Telecom Co., Ltd. White paper on lossless network technology for distributed intelligent computing centers[R]. 2024.
- [4] 唐卓, 蒋冰婷, 张嘉鹏, 等. 算力网络调度基础理论与关键技术现状及展望[J]. 中国科学基金, 2025, 39(2): 240-249.
Tang Z, Jiang B T, Zhang J P, et al. The foundation theory and key technologies of computing first network scheduling: present situation and prospects[J]. Bulletin of National Natural Science Foundation of China, 2025, 39(2): 240-249.
- [5] YD/T 6064-2024. 算力网络算网编排管理技术要求[S].
YD/T 6064-2024. Technical requirements of orchestration management of computing and network convergence[S].
- [6] YD/T 6044-2024. 算力网络算力度量与算力建模技术要求[S].
YD/T 6044-2024. Technical requirements of computing measurement and computing modeling in computing and network convergence[S].
- [7] 王蕴, 林霄, 楼芝兰, 等. 面向边缘光算力网络的上行链路资源协同调度算法[J]. 光通信技术, 2024, 48(3): 45-51.
Wang Y, Lin X, Lou Z L, et al. Uplink resource coordinated scheduling algorithm for edge-oriented optical computing power networks[J]. Optical Communication Technology, 2024, 48(3): 45-51.
- [8] Yi C Y, Cai J, Su Z. A multi-user mobile computation offloading and transmission scheduling mechanism for delay-sensitive applications[J]. IEEE Transactions on Mobile Computing, 2020, 19(1): 29-43.
- [9] 归律, 王悦, 张钟琴. 东数西算中多云管理技术应用浅析[J]. 信息通信技术与政策, 2022, 48(11): 84-92.
Gui L, Wang Y, Zhang Z Q. The application of multicloud management technology in east-data-west-computing project[J]. Information and Communications Technology and Policy, 2022, 48(11): 84-92.
- [10] Afifi H, Pochaba S, Boltres A, et al. Machine learning with computer networks: techniques, datasets, and models[J]. IEEE Access, 2024, 12: 54673-54720.
- [11] 叶宁, 付康, 胡少文, 等. 基于MPI的异构算力资源融合调度平台[J]. 计算机与现代化, 2025(12): 38-45.
Ye N, Fu K, Hu S W, et al. MPI-based heterogeneous computing resource integration and scheduling platform[J]. Computer and Modernization, 2025(12): 38-45.
- [12] Long Y X, Wang Z Y, Lan S Z, et al. Energy-latency tradeoff for task offloading and resource allocation in vehicular edge computing[J]. Computer Networks, 2025, 258: 111026.
- [13] Pang S C, Cheng Y Z, He X, et al. Multi-user motion state task offloading strategy for load balancing in mobile edge computing networks[J]. Ad Hoc Networks, 2025, 169: 103759.
- [14] Mahesar A R, Li X P, Sajani D K. Enhancing task scheduling and QoS optimization in mobile edge computing via microservice-oriented container selection[J]. Computing, 2025, 107(2): 60.
- [15] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.
- [16] 陆威然, 鲁腾飞, 赵恪振, 等. 基于博弈论-VIKOR的Kubernetes调度策略研究[J]. 计算机时代, 2023(8): 52-56.
Lu W R, Lu T F, Zhao K Z, et al. Kubernetes scheduling strategy based on game theory-VIKOR[J]. Computer Era, 2023(8): 52-56.
- [17] 中国电信股份有限公司上海分公司, 华为技术有限公司. 50G PON全光万兆发展白皮书[R]. 2023.
China Telecom Corp., Ltd. Shanghai Branch, Huawei Technologies Co., Ltd. White paper on all-optical 10G development based on 50G PON[R]. 2023.
- [18] 马智勇, 张坚平, 严森垒. 基于50G-PON的全光万兆示范与应用[J]. 电信科学, 2025, 41(Z1): 274-278.
Ma Z Y, Zhang J P, Yan S L. Demonstration and application of all-optical 10G access based on 50G-PON[J]. Telecommunications Science, 2025, 41(Z1): 274-278.

[作者简介]



龚豪 (1982-), 男, 中国电信股份有限公司上海分公司副总经理, 主要研究方向为人工智能、云计算、算力网络技术及产品创新。