

# TC609

## 全国数据标准化技术委员会技术文件

TC609-6-2025-04

### 全国一体化算力网 智算中心算力池化 技术要求

National integrated computing power network—Technical requirements for pooling  
of computing power resource in artificial intelligence data center

2025-08-29 发布

2025-08-29 实施

全国数据标准化技术委员会 发布



# 目 次

前 言 .....	II
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
4 缩略语 .....	2
5 智算中心算力池化概览 .....	2
5.1 智算中心算力池化与资源层各模块的关系 .....	2
5.2 功能架构 .....	3
6 智算中心算力池化功能要求 .....	4
6.1 异构设备管理 .....	4
6.2 算力资源抽象 .....	4
6.3 资源池化管理 .....	5
6.4 任务执行管理 .....	6
7 智算中心算力池化接口要求 .....	6
参 考 文 献 .....	9

# 前 言

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国数据标准化技术委员会（SAC/TC609）提出并归口。

本文件起草单位：中国移动通信集团有限公司、中国移动通信有限公司研究院、中移（苏州）软件技术有限公司、鹏城实验室、中国科学院计算技术研究所、国家信息中心、中兴通讯股份有限公司、江苏省数据集团有限公司、联通数字科技有限公司、天翼云科技有限公司、中国联合网络通信集团有限公司、曙光信息产业股份有限公司、北京邮电大学、北京智源人工智能研究院、上海人工智能实验室、京东科技信息技术有限公司、视联动力技术股份有限公司、江苏博云科技股份有限公司、江苏未来网络集团有限公司、北京市大数据中心、江西省大数据中心、中国信息通信研究院、国家数据发展研究院、中国电子技术标准化研究院、海光信息技术股份有限公司、青海省海南藏族自治州数据局、北京趋动智能科技有限公司、新华三技术有限公司、浪潮电子信息产业股份有限公司。

# 全国一体化算力网 智算中心算力池化技术要求

## 1 范围

本文件规定了全国一体化算力网资源层中智算中心内的算力池化技术要求，包括异构设备管理、算力资源抽象、资源池化管理、任务执行管理等功能及接口要求。

本文件适用于在全国一体化算力网监测调度平台，算力网资源层的智算中心异构算力资源相关池化能力的规划、建设与实现。

## 2 规范性引用文件

本文件没有规范性引用文件。

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

**算力** `computing power`

图形处理器（GPU）、中央处理器（CPU）等设备执行计算密集型任务的计算能力。

### 3.2

**算力网** `computing power network`

支撑数字经济高质量发展的关键基础设施，可通过网络连接多源异构、海量泛在算力，实现资源高效调度、设施绿色低碳、算力灵活供给、服务智能按需。

### 3.3

**智算中心** `artificial intelligence data center`

智算中心是以人工智能计算需求为核心，集成高性能异构算力（如GPU、NPU、TPU、FPGA等加速卡也称加速芯片）、高速互联网络与全栈AI软件栈的新型基础设施。

### 3.4

**算力池化** `pooling of computing power resources`

通过将异地异构异属的通算、超算、智算等多元算力资源进行抽象化、虚拟化和逻辑化整合，形成可在全国一体化算力网中监测、调度、运营的业务资源集合，满足按需分配、动态调度、一点接入、即取即用的算力服务需求。

### 3.5

**智算中心算力池化** `pooling of computing power resources in artificial intelligence data center`

将智算中心内的异构智能算力（如GPU、NPU、TPU、FPGA等加速卡，也称加速芯片）进行抽象化、虚拟化，形成服务器位置无关、物理架构无关的逻辑智能算力，并面向云资源、云原生等资源封装提供标准化逻辑智能算力规格，支持智能算力任务的灵活、按需、无感使用。

### 3.6

**热迁移** `live migration`

在业务无感知条件下将任务或资源动态转移至其他物理设备的过程。

### 3.7

**逻辑智能算力** `logical artificial intelligent computing power resources`

逻辑智能算力是对高性能异构智能算力（如GPU、NPU、TPU、FPGA等加速卡，也称加速芯片）的抽象，具备可自定义规格的计算内核与显存，本文简称XPU。

## 4 缩略语

下列缩略语适用于本文件。

CPU：中央处理单元（Central Processing Unit）

FPGA：现场可编程门阵列（Field-Programmable Gate Array）

NPU：神经网络处理单元（Neural network Processing Unit）

GPU：图形处理单元（Graphics Processing Unit）

QoS：服务质量（Quality of Service）

SLA：服务等级协议（Service Level Agreement）

TPU：张量处理单元（Tensor Processing Unit）

## 5 智算中心算力池化概览

### 5.1 智算中心算力池化与资源层各模块的关系

《全国一体化算力网 监测调度平台建设指南》中定义了全国一体化算力网监测调度平台总体架构，包括算力网运营层、算力网调度层、算力网资源层以及算力网监测层。其中算力网资源层通过资源并网将异属异构异地算力资源接入算力网，并依托于算力网服务用户业务需求。算力资源并网是通过网络连接实现算力资源的可达、可用，并通过API接口实现算力资源的管理、调度与计量。

智算中心是以集群方式提供智能算力资源的典型方式之一，位于算力网资源层中，智算中心内一般存在一种或多种类型的异构物理智能算力，其中每种物理智能算力指一张由硬件厂商提供的物理加速卡设备。智算中心中的智能算力资源可以全量或部分通过智算中心网络实现同地异构算力资源接入算力网。

如图1所示，智算中心算力池化位于算力网资源层，将智算中心内异构智能算力资源进行抽象化、虚拟化和逻辑化整合，形成逻辑智能算力，提供给上层资源封装，具体包括：

- a) 智算中心算力池化与算力网资源层中的资源封装对接，为云资源封装、云原生封装等资源封装形式提供逻辑智能算力资源描述形式和资源管理方式，并支持为各类资源封装提供逻辑智能算力需求与物理智能算力的映射；
- b) 智算中心算力池化可为使用云资源封装、云原生封装等资源封装形式的智算应用提供所需逻辑智能算力，并将智算应用内的计算任务调度到任意物理智能算力上执行。

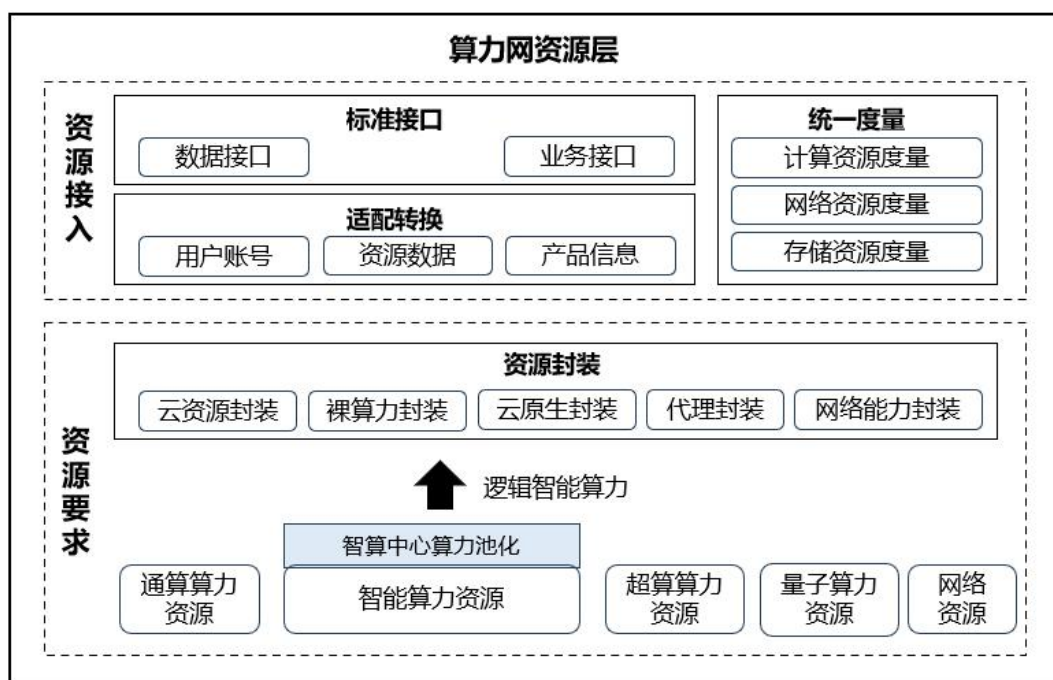


图 1 智算中心算力池化与资源层各模块的关系

## 5.2 功能架构

针对算力网资源层中的各智算中心，算力池化对其提供的异构智能算力进行差异化屏蔽和池化，对上呈现统一智能算力，以便实现上层调度及管理系统对底层异构硬件的一致化灵活调度与管理，以及应用程序在底层异构智算硬件上的灵活跨架构运行与迁移。

如图2所示，智算中心算力池化功能架构主要包括算力资源抽象、资源池化管理、任务执行管理、异构设备管理四个功能模块：

- 算力资源抽象：对异构算力资源进行统一抽象，形成包含设备、显存、计算在内的统一算力资源结构，面向上层调度及业务实现对异构算力资源硬件结构差异的屏蔽，并提供异构算力资源的标准化操作接口；
- 资源池化管理：根据算力资源抽象对异构算力资源统一管理，面向上层调度及业务实现算力资源架构无感的池化能力，支持根据应用任务式资源需求申请、分配并编排算力资源，并提供算力资源抽象与异构算力的相互映射；
- 任务执行管理：负责提供以计算任务的执行调度，支持将智算应用中的多个计算任务拆分，按需将每个计算任务调度到异构算力资源实施计算；
- 异构设备管理：负责发现、注册并记录智算中心每个服务器节点的硬件信息，包括但不限于异构加速卡的型号、数量、拓扑等信息。此外，本模块还对异构算力提供监控、统计、配置、管理等能力。

智算中心池化相关模块的协同逻辑为：异构设备管理首先获取智算中心内物理智能算力信息，由算力资源抽象将物理加速卡抽象为逻辑智能算力，并面向资源池化管理提供统一的资源结构描述，资源池化管理向任务执行管理输出可用物理及逻辑智能算力相关资源信息，并为任务执行管理中计算任务提供所需物理及逻辑智能算力资源的生命周期管理。

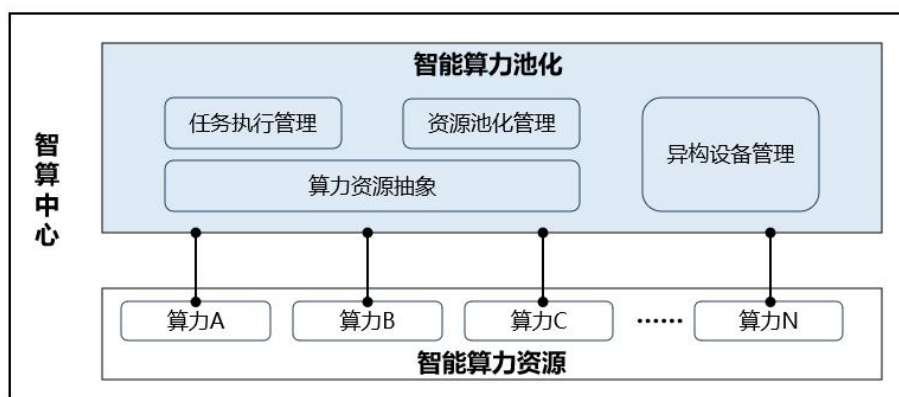


图 2 智算中心算力池化功能架构

## 6 智算中心算力池化功能要求

### 6.1 异构设备管理

异构设备管理对智算中心异构智能算力进行统一管理与监控，具体要求如下：

- 异构资源纳管：应支持纳管异构物理智能算力，资源获取的频率为每半个时一次，提供资源注册、发现、管理等功能，并支持获取设备及异构智能算力基本信息，包括但不限于设备型号、设备供应商、物理加速卡类型、物理加速卡数量、物理加速卡标识、物理加速卡健康状态（运行、离线、故障）、实时负载率、温度、功耗、物理加速卡拓扑等；
- 异构资源监控：应支持对纳管的异构智能算力的状态进行监控，监控频率为每分钟一次，监控内容包括加速卡状态、加速卡利用率、加速卡显存用量等；
- XPU 视角资源管理：应支持将纳管的异构智能算力以逻辑智能算力（即 XPU）形式呈现，具体包括本智算资源池中 XPU 数量、XPU 规格、XPU 资源用量、XPU 负载、XPU 健康状态等；
- 数据呈报：应支持物理智能算力视角以及 XPU 视角的全局监控、告警、日志、数据统计、报表等。

### 6.2 算力资源抽象

#### 6.2.1 设备抽象

设备抽象旨在定义统一的算力设备结构及信息，以面向多样异构设备形成统一的设备形态，具体要求如下：

- 每台算力抽象设备应具备一个主机端和一个或多个设备端，其中主机端指 CPU 及内存，设备端指逻辑智能算力。算力抽象设备可兼容无主机端的异构设备（如 DPU）；
- 设备端逻辑智能算力包含符合计算抽象的逻辑计算单元、符合内存抽象的逻辑内存单元，并支持自定义规格。

#### 6.2.2 计算抽象

计算抽象旨在对异构物理智能算力的计算内核进行统一抽象，具体要求如下：

- 计算内核抽象应具备四级粒度不同的计算单元，具体包括标量计算单元、向量计算单元、张量计算单元、函数计算单元；



- b) 标量计算单元为计算内核抽象中的一级计算单元，应支持一维数据的 add、mul、div、sin、cos、sqrt、log 等基础运算；
- c) 向量计算单元为计算内核抽象中的二级计算单元，应支持二维数据的 add、mul、div、sin、cos、sqrt、log、max、min 等基础运算；
- d) 张量计算单元为计算内核抽象中的三级计算单元，应支持多维（大于等于三维）数据的 add、mul、div、sin、cos、sqrt、log、max、min 等基础运算；
- e) 函数计算单元为计算内核抽象中的四级计算单元，应支持卷积、矩阵乘法、激活函数等预定义复杂算子运算；
- f) 计算内核抽象应支持对 INT4、INT8、BF16、FP16、FP32 等精度的数据计算，并支持根据物理加速卡硬件特性自动适配精度，屏蔽不同加速卡的精度差异；
- g) 计算内核抽象应支持与异构算力资源加速卡中的至少一种计算内核进行分级映射；
- h) 应定义数据依赖关系，如计算单元间的数据传递关系（如“向量单元的输出作为张量单元的输入”），为后续调度提供依据。

### 6.2.3 内存抽象

内存抽象旨在对存储、主机内存及异构物理智能算力的显存进行统一抽象，具体要求如下：

- a) 内存抽象应具备五种存储介质抽象，具体包括存储系统、主机端内存、设备端全局内存、设备端本地内存、设备端私有内存，其中存储系统可选；
- b) 存储系统为设备之外的独立存储系统，应支持缓存数据、与主机端内存数据交换，支持数据缓存一致性机制，当主机端与设备端内存数据同步时，需通过锁机制或原子操作保障数据完整性，避免读写冲突；
- c) 主机端内存应支持缓存数据、从存储中读写数据、与设备端内存数据交换，支持 DMA 等传输方式；
- d) 设备端全局内存应支持数据缓存、与主机端内存数据交换、全部计算内核数据读写访问；
- e) 设备端本地内存应支持数据缓存、一组计算内核共享数据读写访问；
- f) 设备端私有内存应支持数据缓存、一个计算内核数据读写访问。

### 6.3 资源池化管理

资源池化管理支持面向上层调度及业务实现算力资源架构无感的池化资源供给，具体要求如下：

- a) 异构资源调度：应支持以异构算力资源抽象中的资源粒度进行异构资源调度，最小调度单位为一个 XPU，其中 XPU 是一个抽象的加速卡，具备符合计算抽象的逻辑计算单元、符合内存抽象的逻辑内存单元，计算单元数量及内存容量支持自定义；
- b) 异构资源映射：应支持 XPU 与实际物理加速卡资源的映射，支持根据物理计算内核与抽象计算内核、物理内存与抽象内存之间的比例进行映射，一般一个物理加速卡可映射为 1 个或多个 XPU，当物理加速卡资源不足够映射为整数个 XPU 时，按最小资源单元碎片化映射；资源映射时可根据计算单元、内存、算力等多维度映射；
- c) 池化资源管理：应支持应用与物理加速卡资源的解耦，面向应用提供来源于智算中心内一台或多台物理计算节点的同构或异构的智能算力资源；
- d) 服务质量：服务质量应支持资源限制（算力、显存配额）及性能隔离（计算内核、内存隔离），应支持自动 QoS 管理，按分配的 XPU 数量限制上层应用可以使用的智算资源而避免争抢，按需实现算力与显存隔离；
- e) 弹性扩缩：应支持根据负载以 XPU 为单位自动在线 Scale Up/Down 而业务不感知；
- f) 分布式调用：应支持在上层应用无感知的情况下调用 XPU；

- g) 迁移能力：应支持 XPU 上计算任务的热迁移能力，将 XPU 上任务热迁移到本地其它服务器上的剩余资源充足的卡上，任务能够继续运行而上层业务无感知，热迁移过程中资源量仍使用 XPU 最小调度单位；
- h) 资源生命周期管理：应支持管理虚拟机、容器的资源生命周期，包括但不限于实例创建、查询、终止等，具体包括：
  - 实例创建：为应用执行虚拟机或容器实例创建时，执行 XPU 与异构物理资源映射，并根据 XPU 资源需求选择对应物理加速卡资源执行应用实例化创建，并同步执行本地网络、存储、服务、状态、扩缩容相关编排及配置；
  - 查询：支持查询应用虚拟机或容器实例化创建后，所用物理加速卡资源、XPU 资源、网络、存储、状态等信息的查询；
  - 终止：终止应用虚拟机或容器时，执行 XPU 及对应物理加速卡、网络、存储等资源的释放；
  - 资源调度日志：提供资源的调度日志。

#### 6.4 任务执行管理

任务执行管理支持按计算任务粒度实现资源调度，具体要求如下：

- a) 计算任务拆分：计算任务拆分应基于任务依赖关系、算力需求强度（如高算力需求任务单独分组）进行，应支持将应用中的多个计算任务分为一个或多个计算任务组，每个计算任务组中包含一个或多个应在同构加速卡上执行的计算任务，其中，一个计算任务一般指 1 个可在加速卡上执行的核函数；
- b) 计算任务调度：应支持按计算任务组粒度调度并执行计算任务；在任务调度时，如用户指定物理加速卡型号，可定向调度至指定型号的物理加速卡，若用户未指定物理加速卡型号，可随机调度到任意可提供所需资源的物理加速卡；
- c) 调度队列：应支持计算任务调度队列，计算任务可加入队列等待分发与计算执行、按优先级分配，应支持多计算任务优先级及亲和性等策略设置，高优先级任务可抢占低优先级任务资源，抢占时应保存低优先级任务状态；
- d) 任务质量管理：应支持计算任务执行的 SLA 管理；
- e) 任务迁移：应支持在异构加速卡间执行计算任务热迁移，任务迁移过程中应通过快照机制保障数据一致性，迁移中断时支持任务回滚至迁移前状态，迁移后计算任务能够继续运行而业务无感知；
- f) 任务管理：应支持任务的可视化界面展现，包括对当前任务计算状态、在运行期间的 XPU 分配及实际使用情况进行展现，对以上信息支持任务日志记录及任务报表导出；
- g) 任务隔离：应通过命名空间或硬件隔离保证多任务并发执行时的安全边界；
- h) 任务执行监控日志：支持定期推送任务执行监控数据记录及日志。

#### 7 智算中心算力池化接口要求

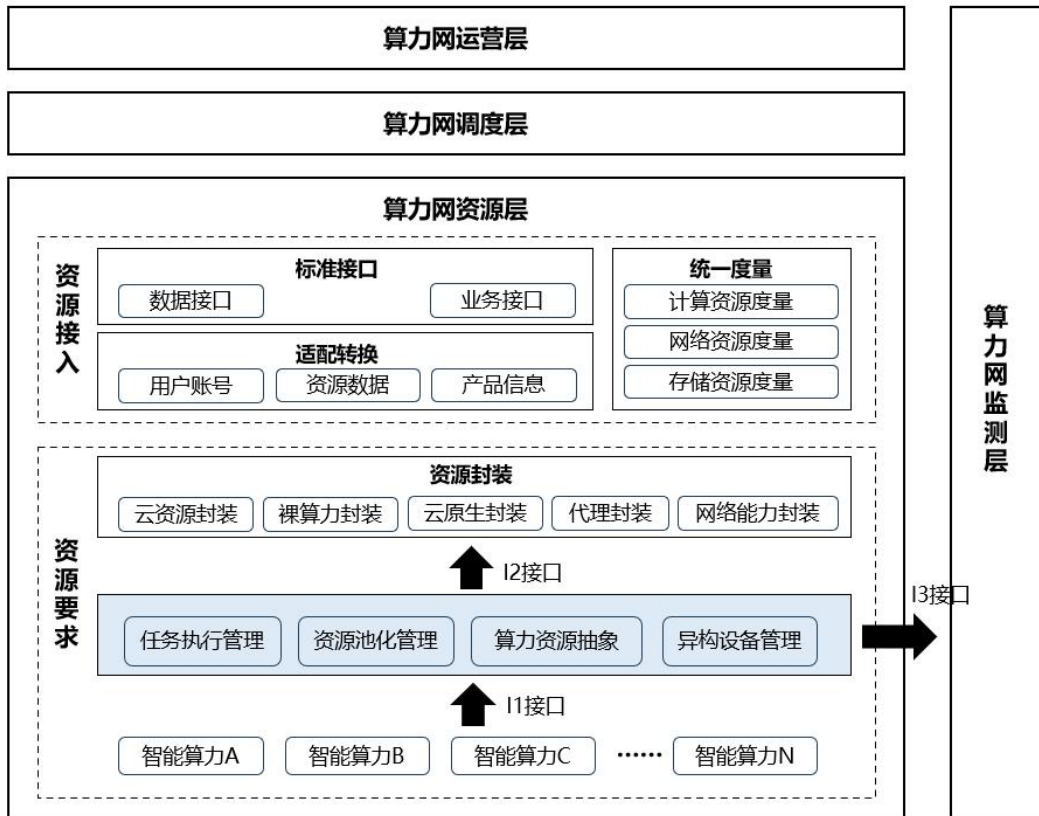


图3 智算中心算力池化接口

智算中心算力池化需通过接口与异构智能算力、资源封装、算力网监测层交互，具体包括异构智能算力池化接入接口（如图3中I1接口所示）、池化算力资源封装接口（如图3中I2接口所示）、池化算力资源监测接口（如图3中I3接口所示），接口详细要求如下：

异构智能算力池化接入（I1接口）：异构算力资源通过此接口与算力资源抽象、异构设备管理对接，以实现设备、计算、内存等资源的上报与映射，以及异构资源个性化操作接口和底层异构算力资源驱动接口的差异化屏蔽。接口具体要求如下：

- 应提供用于查询设备、计算、内存等资源信息的接口，并通过该接口获取底层异构资源信息以及XPU资源信息，资源信息内容包括资源总量、可用/已用资源量、显存总量、可用/已用显存量、功耗等；
- 应支持设备管理系列接口，至少支持对异构加速卡的设备注册、设备获取、设备信息、健康状态获取、设备占用、设备释放；
- 应支持内存管理系列接口，至少支持对异构加速卡的内存空间创建、内存空间释放；
- 应支持计算任务核函数管理系列接口，至少支持计算任务创建、计算任务下发、任务失败后自动重试、计算任务释放、计算任务参数设置、计算任务使用计算核信息获取；
- 应支持上下文管理系列接口，至少支持上下文创建、上下文保持、上下文释放；
- 应支持事件管理接口，至少支持事件等待、事件释放；
- 应支持根据获取的设备信息将异构加速卡与标准规格XPU设备映射，并提供映射比例查询接口和映射比例动态调整接口；
- 应支持接口版本管理、多版本兼容与平滑升级。

池化算力资源封装（I2接口）：池化的逻辑智能算力资源通过此接口与资源封装对接，以实现池化算力资源的差异化屏蔽及统一封装接口调用。接口具体要求如下：

- a) 应支持资源生命周期管理系列接口，至少支持算力应用的实例化、升级、查询、终止、扩缩容、迁移，同时支持资源预冻结，防止资源冲突；
- b) 应支持标准 XPU 规格的资源申请接口，并支持将标准规格 XPU 资源申请与物理加速卡资源映射，支持抢占式资源预留和预留资源超时释放，接口应返回申请的 XPU 与物理加速卡的映射关系，包括物理卡唯一标识、映射比例、资源分配起始时间，支持通过该接口查询映射关系的实时有效性；
- c) 应支持以 XPU 为标准规格的算力资源注册接口，智能算力可通过该接口将其注册为 XPU 资源；
- d) 应支持计算任务队列接口，支持业务按单个任务或任务组的粒度向物理加速卡提交计算任务；
- e) 应支持计算任务资源查询接口，向业务提供标准 XPU 的资源申请建议；
- f) 应支持计算任务组合配置接口，支持用户及调度按应描述多个计算任务之间的组合关系，支持任务组合的拓扑监控；
- g) 应支持计算任务迁移接口，支持将特定计算任务从当前执行加速卡迁移至其他加速卡；
- h) 应支持计算任务状态查询接口，支持查询计算任务当前所在加速卡位置、执行状态等；
- i) 应支持状态监控接口：支持接口调用量、成功率、响应时间的实时统计，通过状态监控接口提供监控数据查询能力；
- j) 应支持接口版本管理、多版本兼容与平滑升级。

池化算力资源监测（I3接口）：智算中心内的智能算力资源通过此接口与算力网监测层对接，以实现智能算力资源及池化的逻辑智能算力信息的上报，具体包括资源总量、可用/已用资源量、显存总量、可用/已用显存量、功耗、资源状态等。

### 参 考 文 献

- [1] 《全国一体化算力网 监测调度平台建设指南》标准草案
  - [2] 《数据领域常用名字解释（第二批）》
-