

T/BAX

团体标准

T/BAX XXXX.4—2023

视频图像感知智能应用适配技术要求
第4部分：中心算力适配技术要求

Technical requirements for intelligent application adaptation of video and image
sensing—Part 4 Central computing power technical requirements

(征求意见稿)

XXXX - XX - XX 发布

XXXX - XX - XX 实施

北京安全防范行业协会 发布

目 次

| | |
|--------------------------------|-----|
| 前 言 | III |
| 引 言 | IV |
| 1 范围 | 1 |
| 2 规范性引用文件 | 1 |
| 3 术语、定义和缩略语 | 1 |
| 3.1 术语和定义 | 1 |
| 3.2 缩略语 | 1 |
| 4 总体框架 | 1 |
| 5 中心解析设备 | 2 |
| 5.1 功能要求 | 2 |
| 5.2 接口要求 | 2 |
| 6 中心计算平台 | 2 |
| 6.1 功能要求 | 2 |
| 6.2 接口要求 | 3 |
| 7 中心解析平台 | 3 |
| 7.1 功能要求 | 3 |
| 7.2 接口要求 | 4 |
| 8 输入与输出要求 | 5 |
| 8.1 视图输入 | 5 |
| 8.2 视图输出 | 5 |
| 9 接口安全和协议要求 | 5 |
| 9.1 接口安全 | 5 |
| 9.2 接口协议 | 5 |
| 10 硬件资源接入中心算力要求 | 5 |
| 10.1 算力监控注册要求 | 5 |
| 10.2 算力资源采集注册要求 | 5 |
| 10.3 节点调度标签要求 | 5 |
| 10.4 算力资源信息要求 | 5 |
| 10.5 算力虚拟化配置 | 5 |
| 附 录 A （规范性） 集群算力管理接口 | 6 |
| 附 录 B （规范性） 中心计算平台算法引擎部署 | 17 |
| 附 录 C （规范性） 资源监控接口 | 30 |
| 附 录 D （规范性） 算力管理接口 | 35 |
| 附 录 E （规范性） 算法管理接口 | 42 |
| 附 录 F （规范性） 服务管理接口 | 55 |

| | |
|-------------------------|----|
| 附录 G （规范性） 任务管理接口 | 62 |
| 参考文献 | 86 |

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是T/BAX XXXX-2023《视频图像感知智能应用适配技术要求》的第4部分。

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。

本文件由北京安全防范产品行业协会提出并归口。

本文件起草单位：北京市公安局、北京市经济和信息化局、视频图像信息智能分析与共享应用技术国家工程实验室、北控三兴信息技术有限公司、北京旷视科技有限公司、北京百度网讯科技有限公司、富盛科技股份有限公司、北京太初元芯集成电路有限公司、北京瑞莱智慧科技有限公司。

本文件主要起草人：陶山、步飞、崔云红、赵惠芳、王建勇、欧阳晓智、马忠义、胥洋、张伟、苏志伟、钟永强、王洪磊、张浩天。

引 言

视频图像感知智能应用是国家数字化进程和智慧城市建设中的重要内容，在公共安全管理、社会治理、社区管理等应用场景中，发挥关键作用。

T/BAX XXXX-202X《视频图像感知智能应用适配技术要求》旨在通过建立统一的适配技术规范，指导视频图像解析算法与设备的适配，解决算法与设备的绑定问题，实现算法随业务需求持续快速更新迭代，并通过云边端算力动态调度，逐步实现算法快速部署。

T/BAX XXXX-202X拟由五部分构成。

- 第1部分：总体技术要求。目的在于确定视频图像感知智能化应用适配的总体结构，以及适配需要遵循的总体要求和具体要求等技术要求。
- 第2部分：算法适配技术要求。目的在于规范视频图像感知智能化应用中算法适配的基本要求、功能要求、算法授权要求、接口要求等技术要求。
- 第3部分：感知前端和边缘解析设备适配技术要求。目的在于规范视频图像感知智能化应用中感知前端和边缘解析设备的基本要求、功能要求、接口要求等技术要求。
- 第4部分：中心算力适配技术要求。目的在于规范视频图像感知智能化应用中中心算力的总体框架、功能要求、接口要求等技术要求。
- 第5部分：算法算力服务管理平台技术要求。目的在于规范视频图像感知智能化应用中算法接入、算法管理、算力接入、算力管理、视频图像接入、数据共享等技术要求。

视频图像感知智能应用适配技术要求

第4部分：中心算力适配技术要求

1 范围

本文件规定了视频图像感知智能应用中心算力适配的总体框架、中心解析设备、中心计算平台、中心解析平台、输入与输出要求、接口安全和协议要求、硬件资源接入中心算力要求。

本文件适用于视频图像感知智能应用适配中心算力的规划设计、研发、检测和验收。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GA/T 1399.1—2017 公安视频图像分析系统 第1部分：通用技术要求

GA/T 1400.1—2017 公安视频图像信息应用系统 第1部分：通用技术要求

GB/T 28181—2022 公共安全视频监控联网系统信息传输、交换、控制技术要求

T/BAX XXXX.1—202X 视频图像感知智能应用适配技术要求 第1部分：总体技术要求

T/BAX XXXX.3—202X 视频图像感知智能应用适配技术要求 第3部分：感知前端和边缘解析设备适配技术要求

T/BAX XXXX.5—202X 视频图像感知智能应用适配技术要求 第5部分：算法算力服务管理平台技术要求

IETF RFC 6749 OAuth2.0授权框架（The OAuth2.0 Authorization Framework）

3 术语、定义和缩略语

3.1 术语和定义

GA/T 1399.1—2017、GA/T 1400.1—2017、T/BAX XXXX.1—202X界定的以及下列术语和定义适用于本文件。

3.2 缩略语

下列缩略语适用于本文件。

CAD：中心解析设备（Central Analysis Device）

CCP：中心计算平台（Central Computing Platform）

CAP：中心解析平台（Central Analysis Platform）

4 总体框架

中心算力适配总体结构图见图1，包括最上层的算法算力服务管理平台、中心算力的三种模式，以及下层的操作系统、支撑软件栈、硬件及芯片等。

中心算力三种模式包括中心解析设备、中心计算平台、中心解析平台。中心算力向下接入各异构算力设备厂家的硬件算力资源，向上与算法算力服务管理平台对接，实现算力资源的监控和调度。

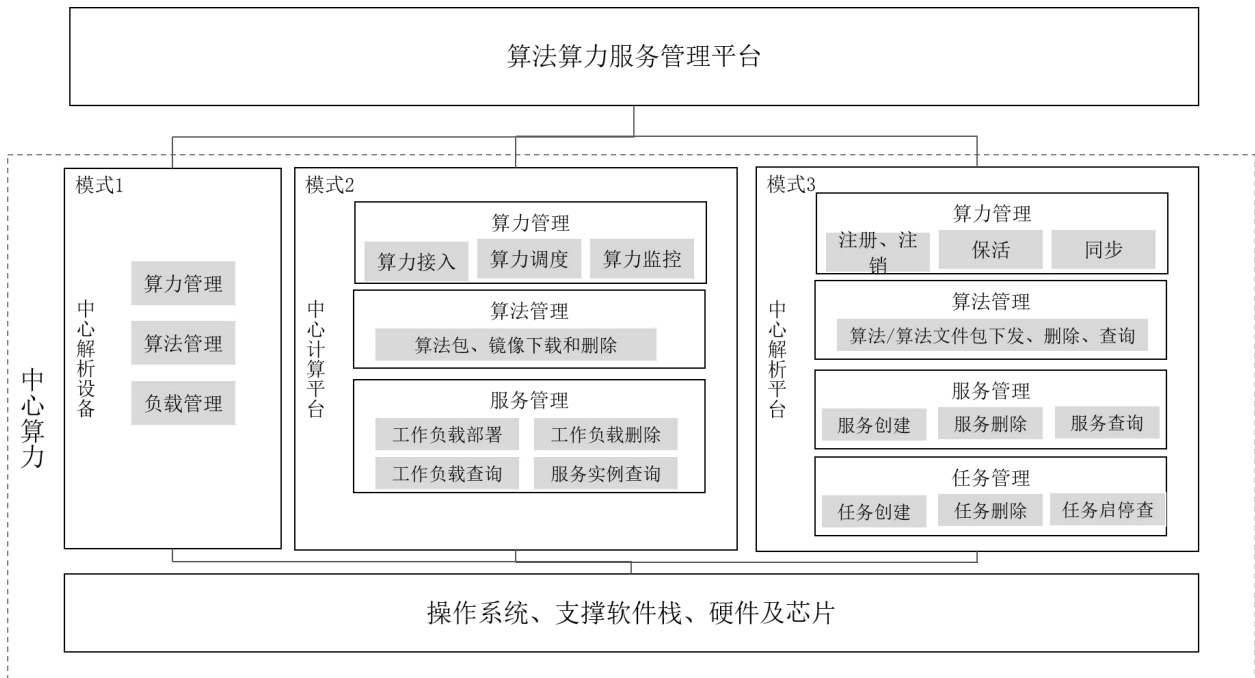


图1 中心算力适配框架

5 中心解析设备

5.1 功能要求

中心解析设备应支持视频图像信息采集、算法加载运行、算力资源的管理，以供算法算力服务管理平台的统一调度管理，同时遵循T/BAX XXXX. 3-202X（《视频图像感知智能应用适配技术要求 第3部分：感知前端和边缘解析设备适配技术要求》）中第5章的规定。

5.2 接口要求

中心解析设备在满足T/BAX XXXX. 3-202X（《视频图像感知智能应用适配技术要求 第3部分：感知前端和边缘解析设备适配技术要求》）中接口要求基础上，提供机器算力接入接口和机器模式下算法包部署接口，接口规范见附录H和附录I。

6 中心计算平台

6.1 功能要求

6.1.1 算力管理

中心计算平台应支持对集群算力的接入、调度和监控

- 算力接入。应支持各类集群算力的接入，可支持算法算力服务管理平台对集群和节点的创建、删除、查询等。
- 算力调度。中心计算平台提供调度算力资源的相关能力，保障上层调度服务按算法所需资源对各中心计算平台发起调度请求，完成算法部署。
- 算力监控。中心算力负责对集群、节点和工作负载关联的算力资源进行监控，监控指标应包括机器信息、卡数量、CPU使用率、节点内存使用量、磁盘使用量、网卡流量、运行状态、异构加速卡使用率和内存使用量信息。

6.1.2 算法管理

中心计算平台支持算法算力服务管理平台对算法引擎的添加，并对中心计算平台的算法引擎进行删除和查询。

中心计算平台应支持创建镜像仓库，能够将算法引擎上传至镜像仓库和删除镜像。

6.1.3 服务管理

中心计算平台支持算法算力服务管理平台对工作负载的部署、删除和查询。

6.2 接口要求

6.2.1 算力管理

为支持算力管理功能，中心计算平台应满足以下接口要求：

- d) 创建集群接口，应符合附录 A.2 的规定；
- e) 删除集群接口，应符合附录 A.3 的规定；
- f) 查询集群列表接口，应符合附录 A.4 的规定；
- g) 查询集群详情接口，应符合附录 A.5 的规定；
- h) 查询集群证书信息接口，应符合附录 A.6 的规定；
- i) 创建节点接口，应符合附录 A.7 的规定；
- j) 删除节点接口，应符合附录 A.8 的规定；
- k) 查询节点列表接口，应符合附录 A.9 的规定；
- l) 获取集群资源监控，应符合附录 C.2 的规定；
- m) 获取节点资源监控，应符合附录 C.3 的规定。

6.2.2 算法管理

算法服务启动时，算力平台应支持从算法算力服务管理平台拉取算法引擎。

6.2.3 负载管理

为提供负载管理功能，中心计算平台应满足以下接口要求：

- a) 部署算法工作负载接口，应符合附录 B.2 的规定；
- b) 删除算法工作负载接口，应符合附录 B.3 的规定；
- c) 算法工作负载扩缩容接口，应符合附录 B.4 的规定；
- d) 查询算法工作负载列表接口，应符合附录 B.5 的规定；
- e) 查询算法工作负载详情接口，应符合附录 B.6 的规定；
- f) 查询算法服务实例列表接口，应符合附录 B.7 的规定；
- g) 查询算法服务实例详情接口，应符合附录 B.8 的规定。

7 中心解析平台

7.1 功能要求

7.1.1 算力管理

中心计算平台应支持以下算力管理能力：

- a) 支持将自身信息注册到算法算力服务管理平台，提供集群算力信息、资源占用情况、物理状态的获取接口。
- b) 支持动态获取集群算力整体情况。

7.1.2 算法管理

中心计算平台应支持以下算法管理能力：

- a) 支持算法算力服务管理平台对算法封装文件、算法包或算法引擎管理以及下发，同时提供查询中心计算平台所拥有的算法封装文件与算法的信息，以及对应的删除的管理能力；
- b) 支持按需下发或删除算法封装文件，下发或删除等行为不应引起中心计算平台的重启或者重置行为，也不会影响与算法无关的其他基础功能（如集群信息获取等）。

7.1.3 服务管理

中心计算平台应支持以下服务管理能力：

- a) 支持算法算力服务管理平台对中心计算平台上的算法服务管理操作，包括算法服务查询、创建、删除等；
- b) 支持按需创建算法服务和算法服务隔离，即相关算法服务的创建、删除等行为，不应引起中心计算平台的重启或者重置行为，不会影响与该算法服务无关的其他基础功能（比如集群监控运维等），也不会影响与该算法服务无关的其他算法服务。

7.1.4 任务管理

中心解析平台应支持以下任务管理能力：

- a) 支持算法算力服务管理平台对中心解析平台相关视图解析任务的管理操作，包括解析任务的查询、创建、删除、暂停、启动等能力。
- b) 支持动态创建解析任务和解析任务隔离的能力，即相关算法解析任务的创建、删除等行为，不应引起中心解析平台的重启或者重置行为，不会影响与该算法解析任务无关的其他基础功能，也不会影响与该算法解析任务无关的其他算法解析任务。

7.2 接口要求

7.2.1 算力管理

为提供算力管理功能，中心解析平台应满足以下接口要求：

- a) 中心解析平台注册接口，应符合附录 D.2 的规定；
- b) 中心解析平台注销接口，应符合附录 D.3 的规定；
- c) 中心解析平台保活心跳接口，应符合附录 D.4 的规定；
- d) 中心解析平台算力同步接口，应符合附录 D.5 的规定。

7.2.2 算法管理

为提供算法管理功能，中心解析平台应满足以下接口要求：

- a) 算法封装文件下发接口，应符合附录 E.2 的规定；
- b) 算法封装文件删除接口，应符合附录 E.3 的规定；
- c) 算法封装文件列表接口，应符合附录 E.4 的规定；
- d) 算法封装文件详情接口，应符合附录 E.5 的规定；
- e) 算法下发接口，应符合附录 E.6 的规定；
- f) 算法删除接口，应符合附录 E.7 的规定；
- g) 算法列表接口，应符合附录 E.8 的规定；
- h) 算法详情接口，应符合附录 E.9 的规定。

7.2.3 服务管理

为提供服务管理功能，中心解析平台应满足以下接口要求：

- a) 算法服务创建接口，应符合附录 F.2 的规定；
- b) 算法服务删除接口，应符合附录 F.3 的规定；
- c) 算法服务查询列表接口，应符合附录 F.4 的规定；
- d) 算法服务查询详情接口，应符合附录 F.5 的规定。

7.2.4 任务管理

为提供任务管理功能，中心解析平台应满足以下接口要求：

- a) 任务创建接口，应符合附录 G.2 的规定；
- b) 任务删除接口，应符合附录 G.3 的规定；
- c) 任务启动接口，应符合附录 G.4 的规定；
- d) 任务暂停接口，应符合附录 G.5 的规定；
- e) 任务查询列表接口，应符合附录 G.6 的规定；
- f) 任务查询详情接口，应符合附录 G.7 的规定。

8 输入与输出要求

8.1 视图输入

中心解析平台应提供符合GB/T 28181-2022中视频流接入要求。
中心解析平台应提供符合GA/T 1400.1-2017中图片流接入要求。

8.2 视图输出

中心解析平台应提供符合GA/T 1399.2-2017中视图输出要求。

9 接口安全和协议要求

9.1 接口安全

中心算力接口应具备安全机制，包括但不限于通过OAuth2获取令牌的方式以及密钥更新的方式。
注1：采用OAuth2可参照rfc6749标准协议。

9.2 接口协议

中心算力接口应采用统一的https协议标准。

10 硬件资源接入中心算力要求

10.1 算力监控注册要求

硬件设备厂商接入中心解析平台时，应支持算力监控注册，包括以下信息：

- a) AI加速卡镜像文件和配置项；
- b) 实例和卡级别AI利用率、显存用量、显存总量、显存利用率；
- c) 卡级别视频、图片解析缓存利用率；
- d) 支持卡-实例关系映射。

10.2 算力资源采集注册要求

硬件设备应提供AI加速卡设备插件，支持对设备资源的访问。

10.3 节点调度标签要求

硬件设备应算力节点上增加具有加速卡类别的选择标签。

10.4 算力资源信息要求

硬件设备宜提供AI加速缓存算力资源、视频解析缓存算力资源、图片解析缓存算力资源。

10.5 算力虚拟化配置

硬件设备宜支持算力虚拟化配置，包括软隔离或硬隔离、时间切片或空间切片、虚拟显存配额。

附录 A
(规范性)
集群算力管理接口

A.1 基本要求

本附录共规定了8个集群算力管理接口，主要包括对集群的创建、删除、查询，以及对集群节点的创建、删除和查询等，见表A.1。其中，算法算力服务管理平台（以下简称：算法算力平台）与算力集群之间的接口调度流程见图A.1，各个接口属性信息应符合表A.2-A.9的规定，具体应符合章节A.2-A.9的规定。

表 A.1 集群算力管理接口

| 序号 | URI | 方法 | 功能说明 |
|----|-------------------------------------|--------|----------|
| 1 | /CCP/clusters | POST | 创建集群 |
| 2 | /CCP/clusters/{clusterCode} | DELETE | 删除集群 |
| 3 | /CCP/clusters | GET | 查询集群列表 |
| 4 | /CCP/clusters/{clusterCode} | GET | 查询集群详情 |
| 5 | /CCP/clusters/{clusterCode}/license | GET | 查询集群证书信息 |
| 6 | /CCP/cluster/nodes | POST | 创建节点 |
| 7 | /CCP/cluster/nodes | DELETE | 删除节点 |
| 8 | /CCP/cluster/nodes | GET | 查询节点列表 |

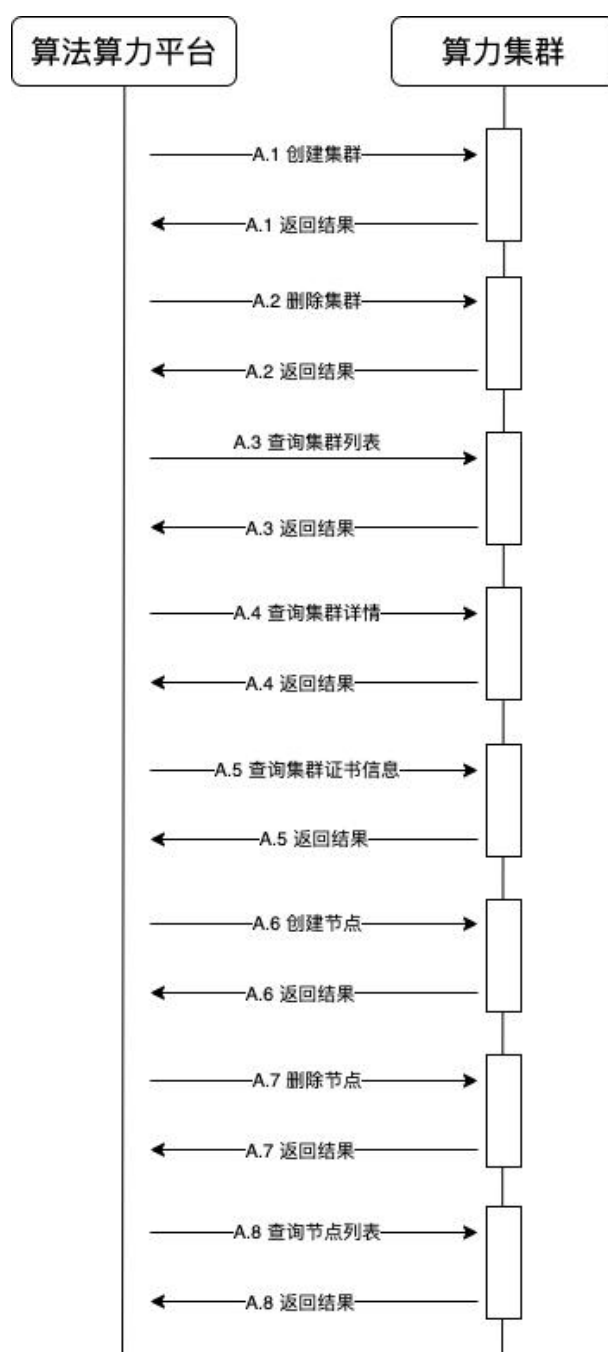


图 A.1 集群算力接口调用时序图

表 A.2 创建集群

| | | | |
|------|---------------|----------------------|-----------------------|
| URI | /CCP/clusters | | |
| 功能 | 创建集群 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | ClusterCreateRequest | ClusterCreateResponse |
| 注释 | 创建集群 | | |

表 A.3 删除集群

| | | | |
|-----|-----------------------------|-------|-------|
| URI | /CCP/clusters/{clusterCode} | | |
| 功能 | 删除集群 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |

| | | | |
|--------|------|---|---------------|
| DELETE | 无 | 无 | EmptyResponse |
| 注释 | 删除集群 | | |

表 A.4 查询集群列表

| | | | |
|-----|---------------|-------|---------------------|
| URI | /CCP/clusters | | |
| 功能 | 查询集群列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | 无 | ClusterListResponse |
| 注释 | 查询集群列表 | | |

表 A.5 查询集群详情

| | | | |
|-----|-----------------------------|-------|-----------------------|
| URI | /CCP/clusters/{clusterCode} | | |
| 功能 | 查询集群详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | 无 | ClusterDetailResponse |
| 注释 | 查询集群详情 | | |

表 A.6 查询集群证书信息

| | | | |
|-----|-------------------------------------|-----------------------|------------------------|
| URI | /CCP/clusters/{clusterCode}/license | | |
| 功能 | 查询集群证书信息 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | ClusterLicenseRequest | ClusterLicenseResponse |
| 注释 | 查询集群证书信息，externalAddress是集群调用地址 | | |

表 A.7 创建节点

| | | | |
|------|--------------------|-------------------|---------------|
| URI | /CCP/cluster/nodes | | |
| 功能 | 创建节点 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | NodeCreateRequest | EmptyResponse |
| 注释 | 创建节点 | | |

表 A.8 删除节点

| | | | |
|--------|--------------------|-------------------|---------------|
| URI | /CCP/cluster/nodes | | |
| 功能 | 删除节点 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 无 | NodeDeleteRequest | EmptyResponse |
| 注释 | 删除节点 | | |

表 A.9 查询节点列表

| | | | |
|-----|--------------------|-----------------|------------------|
| URI | /CCP/cluster/nodes | | |
| 功能 | 查询节点列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | NodeListRequest | NodeListResponse |
| 注释 | 查询节点列表 | | |

A.2 创建集群

POST /CCP/clusters

入参：

表 A.10 ClusterCreateRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|----------|-------------|----|-------|------|
| 1 | 区域 id | regionId | String | | 0 | |
| 2 | 集群名称 | name | String | | R | |
| 3 | 集群 | spec | ClusterSpec | | R | 集群配置 |

表 A.11 ClusterSpec 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|---------------|-------------|----------------|----|-------|-------------------------------------|
| 1 | kubernetes 版本 | version | String | | R | 可选 [1.18.x, 1.20.x, 1.22.x, 1.24.x] |
| 2 | 节点网络 | hostNetwork | String | | 0 | 节点网络, 例如 192.168.0.0/20 |
| 3 | 节点信息 | nodes | List<NodeSpec> | | R | |

表 A.12 NodeSpec 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|--------|--------------|--------------|----|-------|----|
| 1 | 节点资源描述 | nodeResource | ResourceSpec | | R | |
| 2 | 节点个数 | count | Integer | | R | |

表 A.13 ResourceSpec 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|---------------|---------------|----|-------|----------|
| 1 | cpu | cpu | Integer | | R | 单位核数 |
| 2 | 内存 | mem | Integer | | R | 单位 GB |
| 3 | 根目录磁盘 | rootDiskSize | Integer | | R | 单位 GB |
| 4 | 数据磁盘 | localDiskSize | Integer | | R | 单位 GB |
| 5 | 加速卡 | xpuResources | XpuResource[] | | 0 | 加速卡类型及个数 |

表 A.14 XpuResource 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|---------|---------|---------|----|-------|---|
| 1 | xpuType | xpuType | String | | R | xpu 型号, 参考 T/BAX XXXX.5-202X 附录 I 加速卡类型 |
| 2 | num | num | Integer | | R | 个数 |

响应参数:

表 A.15 ClusterCreateResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------|---------|----|-------|-----------------------|
| 1 | 结果状态码 | code | String | | R | 结果状态码 (0: 成功, 其他: 失败) |
| 2 | 结果描述 | message | String | | R | 结果描述 |
| 3 | 数据 | data | Cluster | | 0 | 算法包信息 |
| 4 | 请求 id | requestId | String | | 0 | 请求 ID |

表 A.16 Cluster 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|------|-------------|--------|----|-------|---|
| 1 | 集群编码 | clusterCode | String | | R | |
| 2 | 集群名称 | name | String | | R | |
| 3 | 集群状态 | status | String | | R | 集群状态, 可选 [creating, running, create_failed, deleting, deleted, delete_failed] |

请求示例:

```
{
  "regionId": "cn-bj-d",
  "name": "test-open-clusters",
  "spec": {
    "version": "1.18.9",
    "hostNetwork": "192.168.0.0/20"
    "nodes": [{
      "nodeResource": {
        "cpu": 4,
        "mem": 8,
        "rootDiskSize": 40,
        "localDiskSize": 0,
        "xpuResources": {
          "xpuType": "NVIDIA_T4",
          "num": 4
        }
      },
      "count": 2
    }]
  }
}
```

响应示例

```
{
  "data": [{
    "name": "test-open-clusters",
    "clusterCode": "cluster-1",
    "status": "creating"
  }],
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}
```

A.3 删除集群

DELETE /CCP/clusters/{clusterCode}

表 A.17 入参属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|--------|----|-------|----|
| 1 | 区域 id | regionId | String | | 0 | |
| 2 | 集群编码 | clusterCode | String | | R | |

响应参数:

表 A.18 EmptyResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------|--------|----|-------|-------------------|
| 1 | 结果状态码 | code | String | | R | 结果状态码（0：成功，其他：失败） |
| 2 | 结果描述 | message | String | | R | 结果描述 |
| 3 | 请求 id | requestId | String | | 0 | 请求 ID |

请求示例：

```
{
  "regionId": "cn-bj-d",
  "clusterCode": "cluster-1"
}
```

响应示例

```
{
  "code": "0"
  "message": "success"
  "requestId": "xxxx-xxxx-xxxx"
}
```

A.4 查询集群列表

GET /CCP/clusters

表 A.19 入参属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|----------|--------|----|-------|----|
| 1 | 区域 id | regionId | String | | 0 | |

出参：

表 A.20 ClusterListResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|------|---------|---------------|----|-------|----|
| 1 | 状态码 | code | String | | R | |
| 2 | 结果描述 | message | String | | R | |
| 3 | 集群列表 | data | List<Cluster> | | R | |

表 A.21 Cluster 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|--------|----------|-------------|----|-------|----|
| 1 | 集群基本信息 | metadata | Metadata | | R | |
| 2 | 集群配置 | spec | ClusterSpec | | R | |

表 A.22 Metadata 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|------|-------------|--------|----|-------|----|
| 1 | 集群编码 | clusterCode | String | | R | |
| 2 | 集群名称 | name | String | | R | |

请求示例：

```
{
  "regionId": "cn-bj-d",
}
```

响应示例

```
{
  "data": [{
```



```

    "metadata": {
      "name": "test-open-clusters",
      "clusterCode": "cluster-1"
    },
    "spec": {
      "version": "1.18.9",
      "hostNetwork": "192.168.0.0/20"
    },
    "nodes": [{
      "nodeResource": {
        "cpu": 4,
        "mem": 8,
        "rootDiskSize": 40,
        "localDiskSize": 0,
        "xpuResources": {
          "xpuType": "NVIDIA_T4",
          "num": 4
        }
      },
      "count": 2
    }
  ],
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}

```

A.5 查询集群详情

GET /CCP/clusters/{clusterCode}

表 A.23 入参属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|--------|----|-------|----|
| 1 | 区域 id | regionId | String | | 0 | |
| 2 | 集群编码 | clusterCode | String | | R | |

返回值:

表 A.24 ClusterDetailResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|------|---------|---------|----|-------|----|
| 1 | 状态码 | code | String | | R | |
| 2 | 结果描述 | message | String | | R | |
| 3 | 集群信息 | data | Cluster | | R | |

表 A.25 Cluster 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|--------|----------|-------------|----|-------|----|
| 1 | 集群基本信息 | metadata | Metadata | | R | |
| 2 | 集群配置 | spec | ClusterSpec | | R | |
| 3 | 集群状态 | phase | String | | R | |

请求示例:

```

{
  "regionId": "cn-bj-d",
  "clusterCode": "1"
}

```

响应示例:

```

{
  "data": {
    "metadata": {
      "name": "test-open-clusters",
      "clusterCode": "cluster-1"
    },
    "spec": {
      "version": "1.18.9",
      "hostNetwork": "192.168.0.0/20"
    }
  },
  "nodes": [
    {
      "nodeResource": {
        "cpu": 4,
        "mem": 8,
        "rootDiskSize": 40,
        "localDiskSize": 0,
        "xpuResources": {
          "xpuType": "NVIDIA_T4",
          "num": 4
        }
      },
      "count": 2
    }
  ],
  "phase": "Running",
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}

```

A.6 查询集群证书信息

GET /CCP/clusters/{clusterCode}/license

入参:

表 A.26 ClusterLicenseRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|--------|----|-------|----|
| 1 | 集群编码 | clusterCode | String | | R | |
| 2 | 区域 id | regionId | String | | 0 | |

返回值:

表 A.27 ClusterLicenseResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|------|---------|----------|----|-------|----|
| 1 | 状态码 | code | String | | R | |
| 2 | 结果描述 | message | String | | R | |
| 3 | 证书信息 | data | CertData | | R | |

表 A.28 CertData 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|---------|-----------------|--------|----|-------|---------------------|
| 1 | 外网请求地址 | externalAddress | String | | R | 例如 192.168.1.1:6443 |
| 2 | 证书授权数据 | caCertData | String | | R | 用于访问集群 api-server |
| 3 | 客户端证书 | clientCertData | String | | R | 用于访问集群 api-server |
| 4 | 客户端私钥数据 | clientKeyData | String | | R | 用于访问集群 api-server |
| 5 | 内部请求地址 | innerAddress | String | | 0 | |

请求示例:

```

{
  "regionId": "cn-bj-d",

```

```

    "clusterCode": "cluster-1"
  }

```

响应示例:

```

{
  "data": {
    "innerAddress": "https://127.0.0.1:6443",
    "externalAddress": "https://10.234.199.19:6443",
    "caCertData": "xxxx-xxxx-xxxx",
    "clientCertData": "xxxx-xxxx-xxxx",
    "clientKeyData": "xxxx-xxxx-xxxx"
  },
  "code": "0"
  "message": "success"
  "requestId": "xxxx-xxxx-xxxx"
}

```

A.7 创建节点

POST /CCP/cluster/nodes

入参

表 A.29 NodeCreateRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|----------|----|-------|----|
| 1 | 区域 id | regionId | String | | 0 | |
| 2 | 集群编码 | clusterCode | String | | R | |
| 3 | 机器数 | count | Integer | | R | |
| 5 | 机器配置 | spec | NodeSpec | | R | |

表 A.30 返回值属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|----------|---------|--------------|----|-------|----|
| 1 | 状态码 | code | String | | R | |
| 2 | 结果描述 | message | String | | R | |
| 3 | 节点 id 列表 | nodeId | List<String> | | | |

请求示例:

```

{
  "regionId": "cn-bj-d",
  "clusterCode": "cluster-1",
  "count": 2,
  "spec": {
    "nodeResource": {
      "cpu": 4,
      "mem": 8,
      "rootDiskSize": 40,
      "localDiskSize": 0,
      "xpuResources": {
        "xpuType": "NVIDIA_T4",
        "num": 4
      }
    }
  }
}

```

响应示例:

```

{
  "code": "0"
  "message": "success",

```

```

    "nodeId":["xxx", "xxx"]
}

```

A.8 删除节点

DELETE /CCP/cluster/nodes

入参

表 A.31 NodeDeleteRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|--------------|----|-------|----|
| 1 | 区域 id | regionId | String | | 0 | |
| 2 | 集群编码 | clusterCode | String | | R | |
| 3 | 节点列表 | nodeId | List<String> | | R | |

响应参数: EmptyResponse

请求示例:

```

{
    "regionId": "cn-bj-d",
    "clusterCode": "cluster-1",
    "nodeId": ["xxx", "xxx"]
}

```

响应示例:

```

{
    "code": "0"
    "message": "success"
}

```

A.9 查询节点列表

GET /CCP/cluster/nodes

入参

表 A.32 NodeListRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|--------|----|-------|----|
| 1 | 集群编码 | clusterCode | String | | R | |
| 2 | 区域 id | regionId | String | | 0 | |

返回值:

表 A.33 NodeListResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|------|---------|--------|----|-------|----|
| 1 | 状态码 | code | String | | R | |
| 2 | 结果描述 | message | String | | R | |
| 3 | 节点数据 | data | Node | | R | |

表 A.34 Node 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|---------|----|-------|------------|
| 1 | 节点 id | nodeId | String | | R | |
| 2 | 内网 ip | interallp | String | | R | |
| 3 | 外网 ip | externallp | String | | R | |
| 4 | 机器类型 | machineType | String | | R | [GPU, CPU] |
| 5 | cpu | cpu | Integer | | R | 单位: 核数 |
| 6 | 内存 | memoy | Integer | | R | 单位: GB |
| 7 | 硬盘 | disk | Integer | | R | 单位: GB |

| | | | | | | |
|---|-----------|--------------|---------------|--|---|------------------------------------|
| 9 | AI 卡类型及数量 | xpuResources | XpuResource[] | | R | 参考 T/BAX XXXX.5-202X 附录 I 加速卡类型 |
|---|-----------|--------------|---------------|--|---|------------------------------------|

请求示例：

```
{
  "regionId": "cn-bj-d",
  "clusterCode": "cluster-1",
}
```

响应示例：

```
{
  "code": "0"
  "message": "success"
  "data": {
    "nodeId": "1",
    "interallp": "0.0.0.0",
    "externalIp": "0.0.0.0",
    "machineType": "GPU",
    "cpu": 4,
    "memoy": 8,
    "disk": 40,
    "xpuResources": {[
      "num": 4,
      "aiCardType": "NVIDIA_P4"
    ]}
  }
}
```

附录 B
(规范性)
中心计算平台算法引擎部署

B.1 基本要求

本附录共规定了7个中心计算平台算法引擎部署相关接口，主要包括算法引擎的部署、删除、扩缩和查询，见表B.1。其中，算法算力平台与算力集群之间的算法引擎部署接口调度流程见图B.1，各个接口属性信息应符合表B.2-B.8的规定，具体应符合章节B.2-B.8的规定。

表 B.1 算法引擎部署

| 序号 | URI | 方法 | 功能说明 |
|----|---|--------|------------|
| 1 | /CCP/algoService/workloads | POST | 部署算法工作负载 |
| 2 | /CCP/algoService/workloads/{workloadName} | DELETE | 删除算法工作负载 |
| 3 | /CCP/algoService/workloads/{workloadName} | PUT | 扩缩容算法工作负载 |
| 4 | /CCP/algoService/workloads | GET | 查询算法工作负载列表 |
| 5 | /CCP/algoService/workloads/{workloadName} | GET | 查询算法工作负载详情 |
| 6 | /CCP/algoService/pods | GET | 查询算法服务实例列表 |
| 7 | /CCP/algoService/pods/{podName} | GET | 查询算法服务实例详情 |

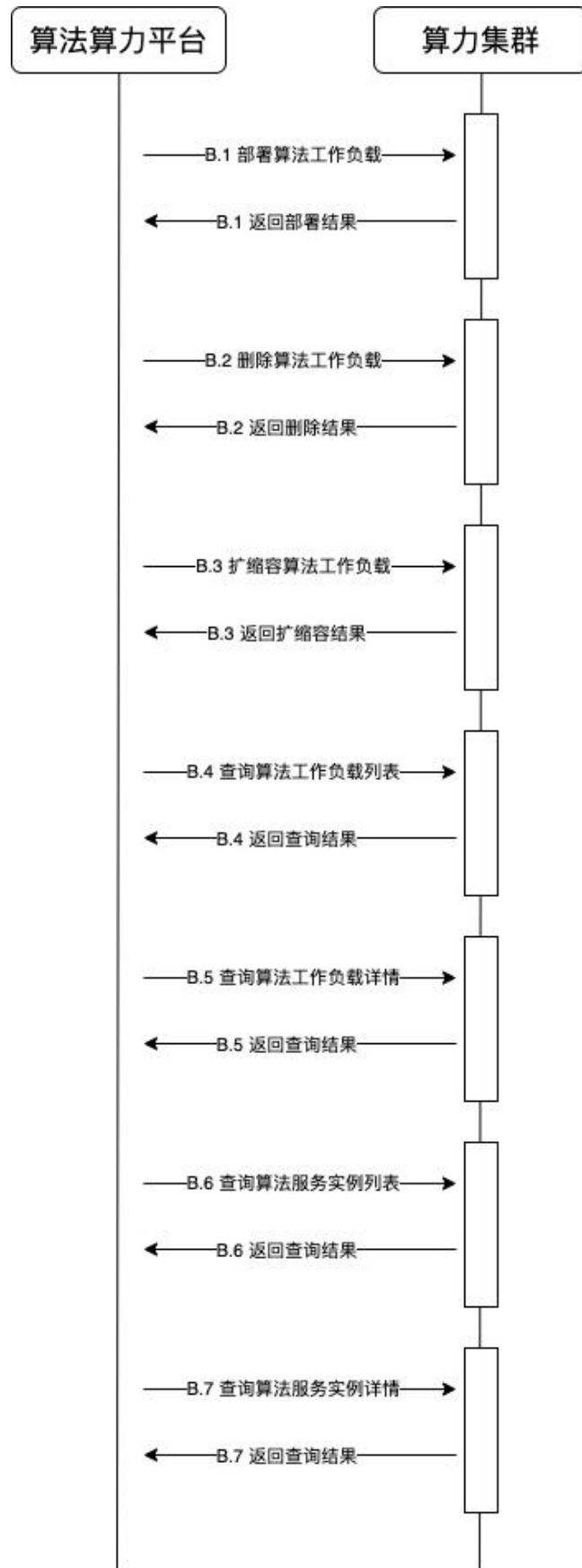


图 B.1 算法引擎部署接口调用时序图

表 B.2 部署算法工作负载

| | | | |
|------|----------------------------|-----------------------|------------------------|
| URI | /CCP/algoservice/workloads | | |
| 功能 | 部署算法工作负载 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | WorkloadCreateRequest | WorkloadCreateResponse |
| 注释 | 算力平台实现的部署算法工作负载接口 | | |

表 B.3 删除算法工作负载

| | | | |
|--------|---|-----------------------|---------------|
| URI | /CCP/algoservice/workloads/{workloadName} | | |
| 功能 | 删除算法工作负载 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 无 | WorkloadDeleteRequest | EmptyResponse |
| 注释 | 算力平台实现的部署算法工作负载接口 | | |

表 B.4 算法工作负载扩缩容

| | | | |
|-----|---|----------------------|---------------|
| URI | /CCP/algoservice/workloads/{workloadName} | | |
| 功能 | 算法工作负载扩缩容 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| PUT | 无 | WorkloadScaleRequest | EmptyResponse |
| 注释 | 算力平台实现的部署算法工作负载接口 | | |

表 B.5 查询算法工作负载列表

| | | | |
|-----|----------------------------|---------------------|----------------------|
| URI | /CCP/algoservice/workloads | | |
| 功能 | 查询算法工作负载列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | WorkloadListRequest | WorkloadListResponse |
| 注释 | 算力平台实现的部署算法工作负载接口 | | |

表 B.6 查询算法工作负载详情

| | | | |
|-----|---|-----------------------|------------------------|
| URI | /CCP/algoservice/workloads/{workloadName} | | |
| 功能 | 查询算法工作负载详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | WorkloadDetailRequest | WorkloadDetailResponse |
| 注释 | 算力平台实现的部署算法工作负载接口 | | |

表 B.7 查询算法服务实例列表

| | | | |
|-----|-----------------------|------------------------|-------------------------|
| URI | /CCP/algoservice/pods | | |
| 功能 | 查询算法服务实例列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | AlgoServiceListRequest | AlgoServiceListResponse |
| 注释 | 算力平台实现的部署算法工作负载接口 | | |

表 B.8 查询算法服务实例详情

| | | | |
|-----|---------------------------------|--------------------------|--------------------------|
| URI | /CCP/algoservice/pods/{podName} | | |
| 功能 | 查询算法服务实例详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | AlgoServiceDetailRequest | AlgoServiceDetailRequest |
| 注释 | 算力平台实现的部署算法工作负载接口 | | |

B.2 部署算法工作负载

将算法引擎部署到中心算力上

URI: POST /CCP/algoservice/workloads

请求参数:

表 B.9 WorkloadCreateRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-----------|---------------|---------------------|----|-------|--|
| 1 | 集群编码 | clusterCode | String | | R | 集群 id |
| 2 | 命名空间 | nameSpace | String | | R | 命名空间 |
| 3 | 工作负载名称 | name | String | | R | 工作负载名称 |
| 4 | 标签组 | labels | Map<String, String> | | R | podTemplate 的标签组 |
| 5 | 标签选择组 | labelSelector | Map<String, String> | | R | 标签选择组 |
| 6 | 副本数 | replicas | Integer | | R | 副本数 |
| 7 | 容器名称 | containerName | String | | R | 容器名称 |
| 8 | 镜像地址 | image | String | | R | 镜像地址 |
| 9 | cpu | cpu | Integer | | R | 容器进程所需 cpu, 单位逻辑核 |
| 10 | AI 加速卡 | xpuType | String | | R | 容器进程所需的 xpu 类型, 参考 T/BAX XXXX. 5-202X 附录 I |
| 11 | AI 加速卡数量 | xpu | Float | | R | 容器进程所需的 xpu 数量, 注: 支持虚拟化的加速卡可传小数(保留一位)。 |
| 12 | memory | memory | Integer | | R | 容器进程所需的 memory, 单位 Gi |
| 13 | 容器启动命令 | command | String | | 0 | 容器启动命令, 未提供则使用容器镜像的 ENTRYPOINT |
| 14 | 容器启动命令的参数 | args | String | | 0 | 如果未提供则使用容器镜像的 CMD |

响应参数:

表 B.10 WorkloadCreateResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|------|--------|----|-------|----------------------|
| 1 | 结果状态码 | code | String | | R | 结果状态码(0: 成功, 其他: 失败) |

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------|----------------|----|-------|---------|
| 2 | 结果描述 | message | String | | R | 结果描述 |
| 3 | 数据 | data | WorkloadResult | | R | 创建的工作负载 |
| 4 | 请求 id | requestId | String | | 0 | 请求 ID |

表 B.11 WorkloadResult 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|--------|--------------|--------|----|-------|--------|
| 1 | 工作负载名称 | workloadName | String | | R | 工作负载名称 |
| 2 | 命名空间 | namespace | String | | R | 命名空间 |

请求示例:

```
{
  "clusterCode": "xxxx",
  "nameSpace": "default",
  "name": "xxxx",
  "labels": {
    "app": "test"
  },
  "labelSelector": {
    "app": "test"
  },
  "replicas": 1,
  "containerName": "test-container",
  "image": "xxxx",
  "xpuType": "NVIDIA_T4",
  "cpu": 3,
  "memory": 5,
  "xpu": 1,
  "replicas": 3
}
```

响应示例

```
{
  "data": {
    "workloadName": "human_detect-deployment-dxxds",
    "namespace": "default"
  },
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}
```

B.3 删除算法工作负载

路径: DELELE /CCP/algoservice/workloads/{workloadName}

请求参数

表 B.12 WorkloadDeleteRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|----|-----|----|----|-------|----|
|----|----|-----|----|----|-------|----|

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|--------|--------------|--------|----|-------|--------|
| 1 | 命名空间 | namespace | String | | R | 命名空间 |
| 2 | 工作负载名称 | workloadName | String | | R | 工作负载名称 |
| 3 | 集群编码 | clusterCode | String | | R | 集群 id |

响应参数：

表 B.13 EmptyResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------|--------|----|-------|-------------------|
| 1 | 结果状态码 | code | String | | R | 结果状态码（0：成功，其他：失败） |
| 2 | 结果描述 | message | String | | R | 结果描述 |
| 3 | 请求 id | requestId | String | | 0 | 请求 ID |

请求示例

```
{
  "clusterCode": "xxxx",
  "namespace": "default"
}
```

响应示例

```
{
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}
```

B.4 扩缩容算法工作负载

路径：PUT /CCP/algoService/workloads/{workloadName}

请求参数

表 B.14 WorkloadScaleRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|--------|--------------|---------|----|-------|--------|
| 1 | 命名空间 | namespace | String | | R | 命名空间 |
| 2 | 工作负载名称 | workloadName | String | | R | 工作负载名称 |
| 3 | 副本数 | replicas | Integer | | R | 期望的副本数 |
| 4 | 集群编码 | clusterCode | String | | R | 集群 id |

响应参数：

表 B.15 EmptyResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------|--------|----|-------|----------------------|
| 1 | 结果状态码 | code | String | | R | 结果状态码(0: 成功, 其他: 失败) |
| 2 | 结果描述 | message | String | | R | 结果描述 |
| 3 | 请求 id | requestId | String | | 0 | 请求 ID |

请求示例

```
{
  "clusterCode": "xxxx",
  "namespace": "default",
  "replicas": 3
}
```

响应示例

```
{
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}
```

B.5 查询算法工作负载列表

路径: GET /CCP/algoservice/workloads

请求参数

表 B.16 WorkloadListRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|--------|--------------|----------------|----|-------|------------------|
| 1 | 集群编码 | clusterCode | String | | R | 集群 id |
| 2 | 命名空间 | namespace | String | | 0 | 命名空间, 默认 default |
| 3 | 工作负载名称 | workloadName | String | | 0 | 工作负载名称 |
| 4 | 分页参数 | pageRequest | ApiPageRequest | | 0 | 分页参数 |

表 B.17 ApiPageRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|------|------------|------------|----|-------|---------------|
| 1 | 页号 | pageNumber | Integer | | R | 页号 [大小: 1~] |
| 2 | 每页大小 | pageSize | Integer | | R | 每页大小 [大小: 1~] |
| 3 | 排序项 | orderList | ApiOrder[] | | 0 | 排序项 |

表 B.18 ApiOrder 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|------|-----------|--------|----|-------|------|
| 1 | 排序属性 | property | String | | 0 | 排序属性 |
| 2 | 排序方向 | direction | String | | 0 | 排序方向 |

响应参数

表 B.19 WorkloadListResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-------|-----------|----------------------|----|-------|----|
| 1 | 状态码 | code | String | | R | |
| 2 | 描述 | message | String | | R | |
| 3 | 请求 id | requestId | String | | 0 | |
| 4 | 数据 | data | WorkloadPage Data | | 0 | |

表 B.20 WorkloadPageData 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|--------|-------------|--------------------|----|-------|----|
| 1 | 工作负载列表 | pageContent | WorkloadInfo [] | | R | |

表 B.21 WorkloadInfo 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|--------|--------------|---------|----|-------|--------|
| 1 | 命名空间 | namespace | String | | R | 命名空间 |
| 2 | 工作负载名称 | workloadName | String | | R | 工作负载名称 |
| 3 | 副本个数 | replicas | Integer | | R | 副本个数 |
| 4 | 创建时间 | createdDate | String | | R | 创建时间 |

请求示例

```
{
  "clusterCode": "xxxx",
  "namespace": "default",
  "workloadName": "xxxxx",
  "pageRequest": {
    "pageNumber": 1,
    "pageSize": 20,
    "orderList": [
      {
        "property": "id",
        "direction": "desc"
      }
    ]
  }
}
```

响应示例

```
{
  "data": {
    "pageRequest": {
      "pageNumber": 1,
      "pageSize": 20,
      "orderList": [
        {
          "property": "id",
```

```

        "direction": "desc"
      }
    ]
  },
  "pageContent": [
    {
      "namespace": 1,
      "workloadName": "",
      "replicas": 1,
      "createdDate": ""
    }
  ],
  "total": 1
},
"code": "0",
"message": "success",
"requestId": "xxxx-xxxx-xxxx"
}

```

B.6 查询算法工作负载详情

路径: GET /CCP/algoservice/workloads/{workloadName}

请求参数

表 B.22 WorkloadDetailRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|--------|--------------|--------|----|-------|--------|
| 1 | 集群编码 | clusterCode | String | | R | 集群 id |
| 1 | 命名空间 | namespace | String | | R | 命名空间 |
| 2 | 工作负载名称 | workloadName | String | | R | 工作负载名称 |

响应参数

表 B.23 WorkloadDetailResponse

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-------|-----------|--------------|----|-------|----|
| 1 | 数据 | data | WorkloadInfo | | R | |
| 2 | 状态码 | code | String | | R | |
| 3 | 状态消息 | message | String | | R | |
| 4 | 请求 ID | requestId | String | | 0 | |

请求示例

```

{
  "clusterCode": "cluster-1"
  "namespace": "default"
}

```

响应示例

```

{
  "data": {
    "namespace": "default",
    "workloadName": "",
    "xpuType": "",

```

```

    "xpu": 1.0,
    "cpuType": "",
    "cpu": 1,
    "replicas": 1,
    "xpuMemory": 1,
    "memory": 1,
    "createdDate": "",
    "status": "",
    "image": "xxxxx"
  },
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}

```

B.7 查询算法服务实例列表

路径: GET /CCP/algoservice/pods

请求参数

表 B.24 AlgoServiceListRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|--------|--------------|--------|----|-------|--------|
| 1 | 集群编码 | clusterCode | String | | R | 集群 id |
| 2 | 命名空间 | namespace | String | | R | 命名空间 |
| 3 | 工作负载名称 | workloadName | String | | R | 工作负载名称 |

响应参数

表 B.25 AlgoServiceListResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-------|-----------|----------------------|----|-------|----|
| 1 | 数据 | data | ClusterPodListResult | | R | |
| 2 | 状态码 | code | String | | R | |
| 3 | 状态消息 | message | String | | R | |
| 4 | 请求 ID | requestId | String | | 0 | |

表 B.26 ClusterPodListResult 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|------|------|---------|----|-------|----|
| 1 | 服务实例 | pods | PodV0[] | | R | |

表 B.27 PodV0 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|--------|------|--------|----|-------|---|
| 1 | pod 名称 | name | String | | R | pod name (eg: vehicle-tracer-deployment-588c78fd4d-599wx) |

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|----------|----------------|---------|----|-------|---|
| 2 | 节点名称 | nodeName | String | | R | node name (eg: bjdd-igu-bce08.bjdd.baidu.com) |
| 3 | 节点 ip | hostIP | String | | R | host IP (eg: 10.234.161.158) |
| 4 | pod ip | podIP | String | | R | pod IP (eg: 10.234.161.158) |
| 5 | 端口 | port | Integer | | R | pod target port (eg: 8750) |
| 6 | pod 状态 | status | String | | R | pod status reusing Kubernetes Pod Lifecycle |
| 7 | pod 状态描述 | statusDesc | String | | R | pod status description |
| 8 | 容器镜像 | image | String | | R | container image (eg: iregistry.baidu-Integer.com/ist/bvs-algo:20210903_1630645199179) |
| 9 | pod 重启次数 | restarts | Integer | | R | pod restart times |
| 10 | pod 运行时间 | age | Integer | | R | age of pod (seconds) |
| 11 | pod 就绪状态 | ready | Boolean | | R | pod ready flag |
| 12 | 工作负载名称 | deploymentName | String | | R | deployment name (eg: vehicle-tracer-deployment) |
| 13 | pod 标签组 | labels | Label[] | | R | pod labels (eg: app: vehicle-tracer) |

表 B.28 Label 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-----------|-------|--------|----|-------|----------------------------------|
| 1 | 标签的 key | name | String | | R | label name (eg: app) |
| 2 | 标签的 value | value | String | | R | label value (eg: vehicle-tracer) |

请求示例

```
{
  "clusterCode": "xxxx",
  "namespace": "default",
  "workloadName": "xxxx"
}
```

响应示例

```
{
  "data": {
    "pods": [
      {
        "name": "",
        "nodeName": "",
        "hostIP": ""
      }
    ]
  }
}
```



```

    "podIP": "",
    "port": 1,
    "status": "",
    "statusDesc": "",
    "image": "",
    "restarts": 1,
    "age": 1,
    "ready": true,
    "deploymentName": "",
    "labels": [
      {
        "name": "",
        "value": ""
      }
    ]
  }
]
},
"code": "0",
"message": "success",
"requestId": "xxxx-xxxx-xxxx"
}

```

B.8 查询算法服务实例详情

路径: GET /CCP/algoservice/pods/{podName}

请求参数

表 B.29 AlgoServiceDetailRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选 | 描述 |
|----|--------|--------------|--------|----|----|--------|
| 1 | 集群编码 | clusterCode | String | | R | 集群 id |
| 2 | 命名空间 | namespace | String | | R | 命名空间 |
| 3 | 工作负载名称 | workloadName | String | | R | 工作负载名称 |
| 4 | pod 名称 | podName | String | | R | pod 名称 |

响应参数

表 B.30 AlgoServiceDetailRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选 | 描述 |
|----|--------|-----------|--------|----|----|----|
| 1 | pod 详情 | data | PodVO | | R | |
| 2 | 状态码 | code | String | | R | |
| 3 | 状态消息 | message | String | | R | |
| 4 | 请求 ID | requestId | String | | 0 | |

请求示例

```

{
  "clusterCode": "xxxx",
  "namespace": "default",
  "workloadName": "xxxx"
}

```

}

响应示例

```
{
  "data": {
    "name": "",
    "namespace": "",
    "nodeName": "",
    "hostIP": "",
    "podIP": "",
    "port": 1,
    "status": "",
    "statusDesc": "",
    "image": "",
    "restarts": 1,
    "age": 1,
    "ready": true,
    "deploymentName": "",
    "labels": [
      {
        "name": "",
        "value": ""
      }
    ]
  },
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}
```

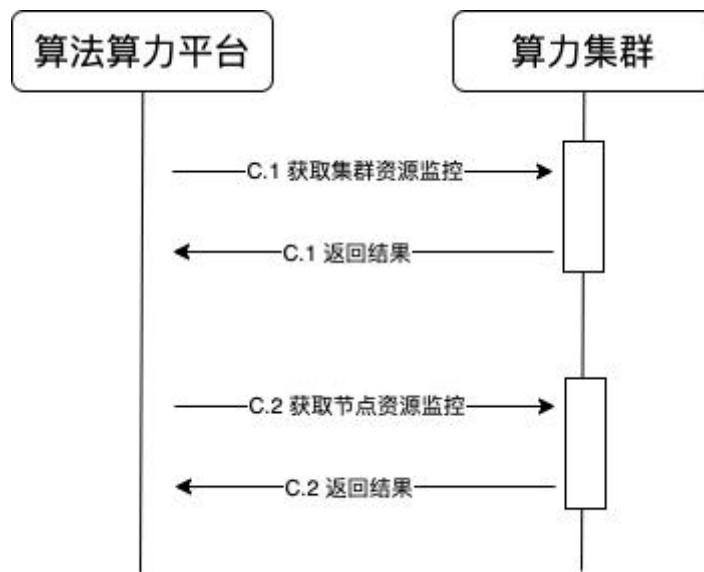
附录 C
(规范性)
资源监控接口

C.1 基本要求

本附录共规定了2个资源监控接口，主要包括集群资源和节点资源的监控，见表C.1。其中，算法算力平台与算力集群之间的资源监控接口调度流程见图C.1，各个接口属性信息应符合表C.2-C.3的规定，具体应符合章节C.2-C.3的规定。

表 C.1 资源监控接口

| 序号 | URI | 方法 | 功能说明 |
|----|-------------------------------------|-----|----------|
| 1 | /CCP/cluster/monitors/{clusterCode} | GET | 获取集群资源监控 |
| 2 | /CCP/node/monitors/{nodeId} | GET | 获取节点资源监控 |



图C.1 资源监控调用接口流程

表 C.2 获取集群资源监控

| | | | |
|-----|-------------------------------------|-------|------------------------|
| URI | /CCP/cluster/monitors/{clusterCode} | | |
| 功能 | 获取集群资源监控 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | 无 | ClusterMonitorResponse |
| 注释 | 获取集群资源监控 | | |

表 C.3 获取节点资源监控

| | | | |
|-----|-----------------------------|--------------------|---------------------|
| URI | /CCP/node/monitors/{nodeId} | | |
| 功能 | 获取节点资源监控 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | NodeMonitorRequest | NodeMonitorResponse |
| 注释 | 获取节点资源监控 | | |

C.2 获取集群资源监控

GET /CCP/cluster/monitors/{clusterCode}

表 C.4 请求参数

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|----|-----|----|----|-------|----|
|----|----|-----|----|----|-------|----|

| | | | | | | |
|---|------|-------------|--------|--|---|-------|
| 1 | 集群编码 | clusterCode | String | | R | 集群 ID |
|---|------|-------------|--------|--|---|-------|

响应参数:

表 C.5 ClusterMonitorResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-------|-----------|--------------------|----|-------|-------|
| 1 | 数据 | data | ClusterMonitorInfo | | R | |
| 2 | 状态码 | code | String | | R | 状态码 |
| 3 | 状态消息 | message | String | | R | 状态消息 |
| 4 | 请求 ID | requestId | String | | 0 | 请求 ID |

表 C.6 ClusterMonitorInfo 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-----------------|--------------------|-------------------------|----|-------|--|
| 1 | ID | id | Integer | | R | ID |
| 2 | 集群名称 | clusterName | String | | 0 | 集群名称 |
| 3 | 总 cpu 资源 | cpu | Integer | | R | 总 cpu 资源, 单位核 |
| 4 | cpu 使用率 | cpuUtilization | Double | | R | cpu 使用率 |
| 5 | cpu 使用率趋势 | cpuUtilizationList | List<Map<Long, Double>> | | 0 | cpu 使用率趋势, [(timestamp, cpuUtilization)] |
| 6 | 总内存资源 | memory | Integer | | R | 总内存资源, 单位 GB |
| 7 | 已使用内存 | memoryUsed | Double | | R | 已使用内存, 单位 GB |
| 8 | 内存使用趋势 | memoryUsedList | List<Map<Long, Double>> | | 0 | 内存使用趋势, [(timestamp, memoryUsed)] |
| 9 | 总硬盘资源 | disk | Integer | | R | 总硬盘资源, 单位 GB |
| 10 | 已使用磁盘 | diskUsed | Double | | R | 已使用磁盘, 单位 GB |
| 11 | 磁盘使用趋势 | diskUsedList | List<Map<Long, Double>> | | 0 | 磁盘使用趋势, [(timestamp, diskUsed)] |
| 12 | xpu 加速卡使用信息(删除) | xpuMonitors | XpuMonitorInfo[] | | 0 | xpu 加速卡使用信息(删除) |

表 C.7 XpuMonitorInfo 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-----------|---------------|---------|----|-------|------------------|
| 1 | xpu 总显存 | xpuMemory | Integer | | 0 | xpu 总显存, 单位 GB |
| 2 | xpu 已使用显存 | xpuMemoryUsed | Double | | 0 | xpu 已使用显存, 单位 GB |

| | | | | | | |
|---|------------|--------------------|-------------------------|--|---|--|
| 3 | xpu 显存使用趋势 | xpuMemoryUsedList | List<Map<Long, Double>> | | 0 | xpu 显存使用趋势, [(timestamp, xpuMemoryUsed)] |
| 4 | xpu 利用率 | xpuUtilization | double | | R | xpu 利用率 |
| 5 | xpu 使用趋势 | xpuUtilizationList | List<Map<Long, Double>> | | 0 | xpu 使用趋势, [(timestamp, xpuUtilization)] |
| 6 | xpu 型号 | xpuType | String | | R | xpu 型号, 参考 T/BAX XXXX. 5-202X 附录 I 加速卡 |

请求示例

```
{
  "clusterCode": "cluster-1"
}
```

响应示例

```
{
  "data": {
    "id": 1,
    "clusterName": "",
    "cpu": 64,
    "cpuUtilization": 0.233444,
    "cpuUtilizationList": [{1686470662199, 0.34345}, {1686433662199, 0.56345}],
    "memory": 128,
    "memoryUsed": 45.3,
    "memoryUsedList": [{1686470662199, 23}, {1686433662199, 45.3}],
    "disk": 1024,
    "diskUsed": 45.0,
    "diskUsedList": [{1686470662199, 23}, {1686433662199, 45.3}],
    "xpuMonitors": [
      "xpuMemory": 16,
      "xpuMemoryUsed": 5.0,
      "xpuMemoryUsedList": [{1686470662199, 3}, {1686433662199, 5.3}],
      "xpuUtilization": 0.233444,
      "xpuUtilizationList": [{1686470662199, 0.34345}, {1686433662199, 0.56345}],
      "xpuType": "NVIDIA_T4"
    ]
  },
  "code": "0",
  "message": "success",
  "requestId": "xxxx-xxxx-xxxx"
}
```

C.3 获取节点资源监控

GET /CCP/node/monitors/{nodeId}

请求参数:

表 C.8 NodeMonitorRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-------------|--------|----|-------|-------|
| 1 | 集群编码 | clusterCode | String | | R | 集群 ID |
| 2 | 节点 ID | nodeId | String | | R | 节点 ID |

响应参数：

表 C.9 NodeMonitorResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-------|-----------|-----------------|----|-------|-------|
| 1 | 数据 | data | NodeMonitorInfo | | R | |
| 2 | 状态码 | code | String | | R | 状态码 |
| 3 | 状态消息 | message | String | | R | 状态消息 |
| 4 | 请求 ID | requestId | String | | 0 | 请求 ID |

表 C.10 NodeMonitorInfo 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 描述 |
|----|-------------|--------------------|-------------------------|----|-------|---|
| 1 | id | id | Integer | | R | ID |
| 2 | ip | ip | String | | R | ip |
| 3 | cpu 资源 | cpu | Integer | | R | 总 cpu 资源，单位核 |
| 4 | cpu 使用率 | cpuUtilization | Double | | R | cpu 使用率 |
| 5 | cpu 使用率趋势 | cpuUtilizationList | List<Map<Long, Double>> | | 0 | cpu 使用率趋势， [(timestamp, cpuUtilization)] |
| 6 | 总内存资源 | memory | Integer | | R | 总内存资源，单位 GB |
| 7 | 已使用内存 | memoryUsed | Double | | R | 已使用内存，单位 GB |
| 8 | 内存使用趋势 | memoryUsedList | List<Map<Long, Double>> | | 0 | 内存使用趋势， [(timestamp, memoryUsed)] |
| 9 | 总硬盘资源 | disk | Integer | | R | 总硬盘资源，单位 GB |
| 10 | 磁盘使用趋势 | diskUsed | Double | | R | 已使用磁盘，单位 GB |
| 11 | 磁盘使用趋势 | diskUsedList | List<Map<Long, Double>> | | 0 | 磁盘使用趋势， [(timestamp, diskUsed)] |
| 12 | xpu 加速卡使用信息 | xpuMonitors | XpuMonitorInfo[] | | 0 | xpu 加速卡使用信息(删除) |

请求示例

```
{
  "clusterCode": "xxx",
  "nodeId": "13"
}
```

响应示例

```
{
  "data": {
    "id": 1,
    "ip": "127.0.0.1",
  }
}
```

```

        "cpu": 64,
        "cpuUtilization": 0.233444,
        "cpuUtilizationList": [{1686470662199, 0.34345}, {1686433662199, 0.56345}],
        "memory": 128,
        "memoryUsed": 45.0,
        "memoryUsedList": [{1686470662199, 23}, {1686433662199, 45.3}],
        "disk": 1024,
        "diskUsed": 45.0,
        "diskUsedList": [{1686470662199, 23}, {1686433662199, 45.3}],
        "xpuMonitors": [
            "xpuMemory": 16,
            "xpuMemoryUsed": 5.0,
            "xpuMemoryUsedList": [{1686470662199, 3}, {1686433662199, 5.3}],
            "xpuUtilization": 0.233444,
            "xpuUtilizationList": [{1686470662199, 0.34345}, {1686433662199, 0.56345}],

            "xpuType": "NVIDIA_T4"
        ]
    },
    "code": "0",
    "message": "success",
    "requestId": "xxxx-xxxx-xxxx"
}

```

附录 D (规范性) 算力管理接口

D.1 基本要求

本附录共规定了4个算力管理接口，主要包括算力向算法算力平台注册、注销、保活和状态查询接口，见表D.1。其中，各个接口属性信息应符合表D.2-D.5的规定，具体应符合章节D.2-D.5的规定，算法算力平台与算力集群之间的算力接口调度流程见图D.1。

表 D.1 算力管理接口

| 序号 | URI | 方法 | 功能说明 |
|----|-----------------|------|---------------|
| 1 | /ACP/Register | POST | 向算法算力平台进行注册 |
| 2 | /ACP/UnRegister | POST | 向算法算力平台进行注销 |
| 3 | /ACP/Keepalive | POST | 向算法算力平台进行心跳保活 |
| 4 | /CAP/Status | GET | 查询中心算力平台算力状态 |

表 D.2 中心解析平台注册

| | | | |
|------|---------------|-----------------|---------------|
| URI | /ACP/Register | | |
| 功能 | 中心解析平台注册 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | RegisterRequest | EmptyResponse |
| 注释 | 向算法算力平台进行注册 | | |

表 D.3 中心解析平台注销

| | | | |
|------|-----------------|-------------------|---------------|
| URI | /ACP/UnRegister | | |
| 功能 | 注销算力设备 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | UnRegisterRequest | EmptyResponse |
| 注释 | 向算法算力平台进行注销 | | |

表 D.4 中心解析平台保活心跳

| | | | |
|------|----------------|------------------|---------------|
| URI | /ACP/Keepalive | | |
| 功能 | 保活算力设备 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | KeepaliveRequest | EmptyResponse |
| 注释 | 向算法算力平台进行心跳保活 | | |

表 D.5 中心解析平台算力同步

| | | | |
|-----|--------------|-------|-------------------|
| URI | /CAP/Status | | |
| 功能 | 查询算力设备状态 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | 无 | CAPStatusResponse |
| 注释 | 查询中心算力平台算力状态 | | |

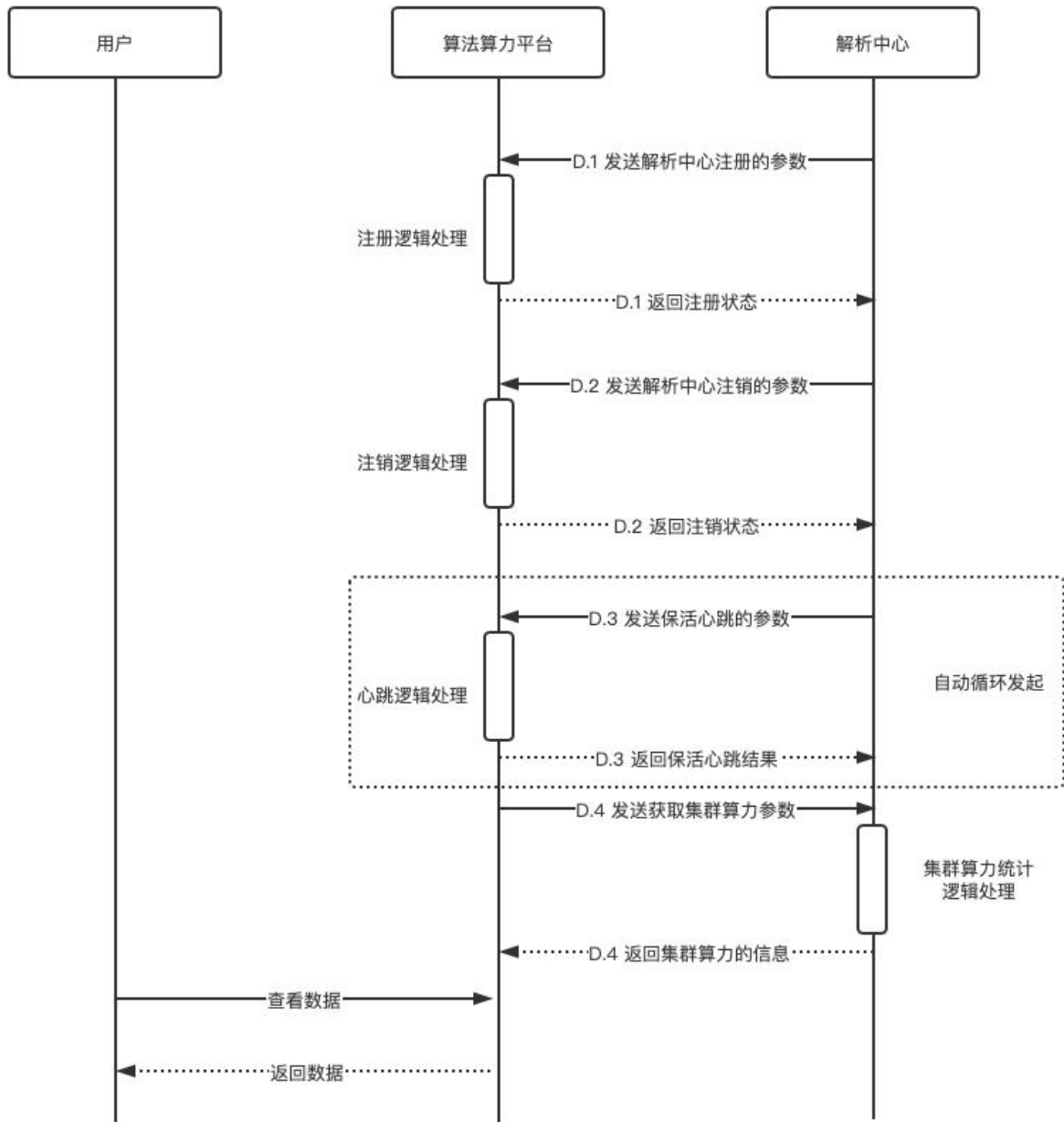


图 D.1 算力管理接口交互接口

D.2 中心解析平台注册

- a) POST /ACP/Register
- b) 请求参数

表 D.6 RegisterRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|----------------------------------|
| 1 | id | R | String | 设备 id |
| 2 | type | R | String | 类型[相机:AI_CAMERA 智能解析盒: AI_BOX |

| | | | | |
|---|----------|---|---------|-----------------|
| | | | | 云(集群):AI_CLOUD] |
| 3 | ip | R | String | 设备 API 服务 ip |
| 4 | port | R | Integer | 设备 API 服务 port |
| 5 | username | 0 | String | API 服务用户名 |
| 6 | password | 0 | String | API 服务密码 |

c) 响应参数

表 D.7 EmptyResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|---------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | message | R | String | 请求结果描述 |
| 3 | data | R | Boolean | 参见示例 |

d) 请求体示例

```
{
  "id": "300000000601",
  "type": "AI_CLOUD",
  "ip": "127.0.0.1",
  "port": 8123,
  "username": "username",
  "password": "password"
}
```

e) 响应体示例:

```
{
  "data": {},
  "code": 0,
  "message": "success"
}
```

D.3 中心解析平台注销

a) POST /ACP/UnRegister

b) 请求参数

表 D.8 UnRegisterRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|-------|
| 1 | id | R | String | 设备 id |

c) 响应参数

表 D.9 EmptyResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|---------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | message | R | String | 请求结果描述 |
| 3 | data | R | Boolean | 参见示例 |

d) 请求体示例

```
{
  "id": "300000000601",
}
```

e) 响应体示例:

```
{
  "data": {},
  "code": 0,
  "message": "success"
}
```

D.4 中心解析平台保活心跳

a) POST /ACP/Keepalive

b) 请求参数

表 D.10 KeepaliveRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|---------|----|---------|-----------|
| 1 | id | R | String | 设备 id |
| 2 | expires | R | Integer | 超时时间, 单位秒 |

c) 响应参数

表 D.11 EmptyResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|---------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | message | R | String | 请求结果描述 |
| 3 | data | R | Object | 参见示例 |

d) 请求体示例

```
{
  "id": "300000000601",
}
```

e) 响应体示例:

```
{
  "data": {},
}
```

```

"code": 0,
"message": "success"
}

```

D.5 中心解析平台算力同步

- a) GET /CAP/Status
b) 请求参数
 无
c) 响应参数

表 D.12 CAPStatusResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|---------------|----|----------|--------------|
| 1 | code | R | String | 请求结果状态码 |
| 2 | message | R | String | 请求结果描述 |
| 3 | data | R | String | 参见示例 |
| 4 | └basicInfos | R | Object[] | 基本信息 |
| 5 | └ip | R | String | IP |
| 6 | └hostName | R | String | HOSTNAME |
| 7 | └cpu | R | Integer | cpu 数 |
| 8 | └gpu | R | Integer | gpu 数 |
| 9 | └gpuType | R | String | gpu 型号 |
| 10 | └memory | R | Double | 内存 |
| 11 | └disk | R | Double | 磁盘空间 |
| 12 | └os | R | String | 操作系统 |
| 13 | └opsInfos | R | Object[] | 运维信息(实际使用情况) |
| 14 | └ip | R | String | IP |
| 15 | └hostName | R | String | HOSTNAME |
| 16 | └cpuUse | R | Double | cpu 使用率 |
| 17 | └gpuUse | R | Double | gpu 使用率 |
| 18 | └memoryUse | R | Double | 内存使用率 |
| 19 | └diskUse | R | Double | 磁盘使用率 |
| 20 | └gpuUsedetail | R | Object[] | 明细 |
| 21 | └algoId | R | String | 算法 Id |
| 22 | └gpuQuota | R | Double | 分配的 gpu 资源 |

- d) 请求体示例
 无
e) 响应体示例:

```

{
  "code": 0,
  "message": "success",
  "data": {
    "basicInfos": [{
      "ip": "10.172.198.208",
      "hostName": "BQF996500332",
      "cpu": 16,
      "gpu": 4,
      "gpuType": "NVidia A2",
      "memory": 256,
      "disk": 14,
      "os": "ubuntu 18.0.4"
    }, {
      "ip": "10.172.198.232",
      "hostName": "BQF996500105",
      "cpu": 32,
      "gpu": 2,
      "gpuType": "NVidia A2",
      "memory": 128,
      "disk": 8,
      "os": "ubuntu 18.0.4"
    }],
    "opsInfos": [{
      "ip": "10.172.198.208",
      "hostName": "BQF996500332",
      "cpuUse": 6,
      "gpuUse": 1,
      "memoryUse": 55.8,
      "diskUse": 2.765,
      "gpuUsedetail": [{
        "algoId": "ASD2LJSAFOWE",
        "gpuQuota": 1
      }
    ]
  }
}

```

```
    }, {  
        "algoId": "ABC8394JDK",  
        "gpuQuota": 0.5  
    }]  
}, {  
    "ip": "10.172.198.232",  
    "hostName": "BQF996500105",  
    "cpuUse": 8,  
    "gpuUse": 1,  
    "memoryUse": 80,  
    "diskUse": 3.545,  
    "gpuUsedetail": [{  
        "algoId": "ASD2LJSAFOWE",  
        "gpuQuota": 1  
    }, {  
        "algoId": "ABC8394JDK",  
        "gpuQuota": 0.5  
    }]  
}]  
}
```

附录 E
(规范性)
算法管理接口

E.1 基本要求

本附录共规定了8个算法管理接口，主要包括算法封装文件和算法的下发、删除、查询等，见表E.1。其中，各个接口属性信息应符合表E.2-E.9的规定，具体应符合章节E.2-E.9的规定，算法算力平台与用户和解析平台之间算法接口调度流程见图E.1。

表 E.1 算法管理接口

| 序号 | URI | 方法 | 功能说明 |
|----|------------------------------|--------|----------|
| 1 | /CAP/AlgoFilePackages | POST | 算法封装文件下发 |
| 2 | /CAP/AlgoFilePackages/{id} | DELETE | 算法封装文件删除 |
| 3 | /CAP/AlgoFilePackages:search | POST | 算法封装文件列表 |
| 4 | /CAP/AlgoFilePackages/{id} | GET | 算法封装文件详情 |
| 5 | /CAP/AlgoPackages | POST | 算法下发 |
| 6 | /CAP/AlgoPackages/{id} | DELTE | 算法删除 |
| 7 | /CAP/AlgoPackages:search | POST | 算法列表 |
| 8 | /CAP/AlgoPackages/{id} | GET | 算法详情 |

表 E.2 算法封装文件下发

| | | | |
|------|-----------------------|----------------------------|-----------------------------|
| URI | /CAP/AlgoFilePackages | | |
| 功能 | 算法封装文件下发 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoFilePackagePushRequest | AlgoFilePackagePushResponse |
| 注释 | 算法封装文件下发 | | |

表 E.3 算法封装文件删除

| | | | |
|--------|----------------------------|-------|-------------------------------|
| URI | /CAP/AlgoFilePackages/{id} | | |
| 功能 | 算法封装文件删除 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 算法封装文件ID | 无 | AlgoFilePackageDeleteResponse |
| 注释 | 算法封装文件删除 | | |

表 E.4 算法封装文件列表

| | | | |
|------|------------------------------|------------------------------|-------------------------------|
| URI | /CAP/AlgoFilePackages:search | | |
| 功能 | 查询算法封装文件列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoFilePackageSearchRequest | AlgoFilePackageSearchResponse |
| 注释 | 查询算法封装文件列表 | | |

表 E.5 算法封装文件详情

| | | | |
|-----|----------------------------|-------|-------------------------------|
| URI | /CAP/AlgoFilePackages/{id} | | |
| 功能 | 查询算法封装文件详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 算法封装文件ID | 无 | AlgoFilePackageDetailResponse |
| 注释 | 查询算法封装文件详情 | | |

表 E.6 算法下发

| | | | |
|-----|-------------------|--|--|
| URI | /CAP/AlgoPackages | | |
| 功能 | 算法下发 | | |

| | | | |
|------|-------|------------------------|-------------------------|
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoPackagePushRequest | AlgoPackagePushResponse |
| 注释 | 算法下发 | | |

表 E.7 算法删除

| | | | |
|--------|------------------------|-------|---------------------------|
| URI | /CAP/AlgoPackages/{id} | | |
| 功能 | 算法删除 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 算法ID | 无 | AlgoPackageDeleteResponse |
| 注释 | 算法删除 | | |

表 E.8 算法列表

| | | | |
|------|--------------------------|--------------------------|---------------------------|
| URI | /CAP/AlgoPackages:search | | |
| 功能 | 查询算法列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoPackageSearchRequest | AlgoPackageSearchResponse |
| 注释 | 查询算法列表 | | |

表 E.9 算法详情

| | | | |
|-----|------------------------|-------|---------------------------|
| URI | /CAP/AlgoPackages/{id} | | |
| 功能 | 查询算法详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 算法ID | 无 | AlgoPackageDetailResponse |
| 注释 | 查询算法详情 | | |

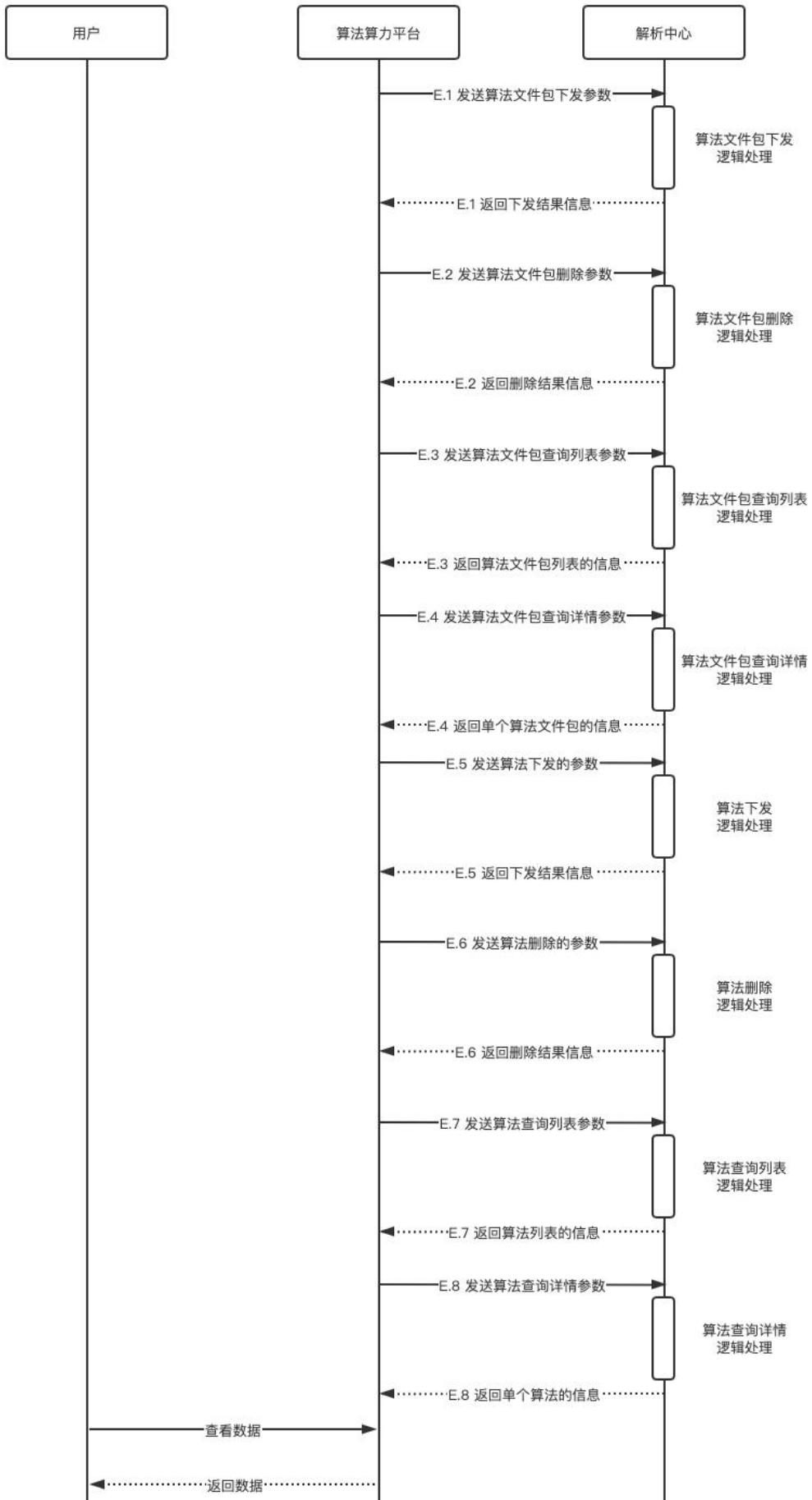


图 E.1 算法管理接口交互

E.2 算法封装文件下发

- a) POST /CAP/AlgoFilePackages
b) 请求参数

表 E.10 AlgoFilePackagePushRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|-----------------|
| 1 | id | R | String | 算法封装文件 id |
| 2 | md5 | R | String | 算法封装文件 md5 |
| 3 | url | R | String | 算法封装文件 http url |

- c) 响应参数

表 E.11 AlgoFilePackagePushResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | |

- d) 请求体示例

```
{
  "id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab",
  "md5": "7a90a32d-8ae8-4ffa-b000-07ef71058bab",
  "url": "http://127.0.0.1/algoFilePackage",
}
```

- e) 响应体示例:

```
{
  "data": {},
  "code": 0,
  "msg": "success"
}
```

E.3 算法封装文件删除

- a) DELETE /CAP/AlgoFilePackages/{id}
b) 请求参数

表 E.12 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|-----------|
| 1 | id | R | String | 算法封装文件 id |

- c) 响应参数

表 E.13 AlgoFilePackageDeleteResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | |

d) 请求体示例

```
{
  "id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab"
}
```

e) 响应体示例:

```
{
  "data": {
    "id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab"
  },
  "code": 0,
  "msg": "success"
}
```

E.4 算法封装文件列表

a) POST /CAP/AlgoFilePackages:search

b) 请求参数

表 E.14 AlgoFilePackageSearchRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------------|----|---------|----------|
| 1 | pageNumber | R | Integer | 页码 |
| 2 | pageSize | R | Integer | size |
| 3 | name | 0 | String | 算法封装文件名称 |

c) 响应参数

表 E.15 AlgoFilePackageSearchResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|-----------|----|---------|-----------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | |
| 4 | └ id | R | String | 算法封装文件 id |
| 5 | └ name | R | String | 算法封装文件名称 |
| 6 | └ version | R | String | 版本 |

| | | | | |
|---|------------------|---|---------|----|
| 7 | L author | R | String | 作者 |
| 8 | L description | R | String | 描述 |
| 9 | L status | R | Integer | 状态 |

d) 请求体示例

```
{
  "pageNumber":0,
  "pageSize":20,
  "name":"xxxx"
}
```

e) 响应体示例:

```
{
  "data": {
    "pageNumber":1,
    "pageSize":20,
    "total":101,
    "records": [
      {
        "id":"",
        "name":"",
        "version":"",
        "author":"",
        "description":"",
        "status":1,
      }
    ]
  },
  "code": 0,
  "msg": "success"
}
```

E.5 算法封装文件详情

a) GET /CAP/AlgoFilePackages/{id}

b) 请求参数

表 E.16 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|----|----|
|----|------|----|----|----|

| | | | | |
|---|----|---|--------|-----------|
| 1 | id | R | String | 算法封装文件 id |
|---|----|---|--------|-----------|

c) 响应参数

表 E.17 AlgoFilePackageDetailResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|--------------------|----|----------|-----------|
| 1 | code | R | Integer | 结果状态码 |
| 2 | msg | R | String | 结果描述 |
| 3 | data | R | Object | |
| 4 | └ id | R | String | 算法封装文件 id |
| 5 | └ name | R | String | 算法封装文件名称 |
| 6 | └ version | R | String | 版本 |
| 7 | └ author | R | String | 作者 |
| 8 | └ description | R | String | 描述 |
| 9 | └ status | R | Integer | 状态 |
| 10 | └ tags | | String[] | |
| 11 | └ env | | Object[] | 环境变量 |
| 12 | └ name | | String | key |
| 13 | └ value | | String | value |
| 14 | └ bin | | Object | |
| 15 | └ type | | | 类型 |
| 16 | └ remoteUri | | | 远程地址 |
| 17 | └ localUri | | | 本地地址 |
| 18 | └ algoRuntime | | | 算法运行时 |
| 19 | └ cmd | | Object | |
| 20 | └ env | | Object[] | |
| 21 | └ name | | | |
| 22 | └ value | | | |
| 23 | └ gpu | | Object[] | |
| 24 | └ quota | | | |
| 25 | └ limit | | | |
| 26 | └ vendor | | | |
| 27 | └ accelerator | | | |
| 28 | └ capacities | | | |
| 29 | └ streamVideoQuota | | | |

| | | | | |
|----|--------------------|--|--|--|
| 30 | └ | | | |
| | streamCaptureQuota | | | |
| 31 | └ analyzeQuota | | | |
| 32 | └ resolution | | | |

d) 请求体示例

```
{
  "id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab",
}
```

e) 响应体示例:

```
{
  "data": {
    "id": "xxx",
    "name": "xxxx",
    "version": "3.4.0",
    "author": "r-face",
    "description": "xxx",
    "env": [
      {
        "name": "_FILELOG_PATH",
        "value": "/data/limbo-log"
      },
      {
        "name": "USE_AIPAAS_PROTO",
        "value": "1"
      }
    ],
    "tags": ["xxx"],
    "bin": {
      "type": "image",
      "remote_uri": "",
      "local_uri": ""
    },
    "algo_runtime": {
      "cmd": "",
      "env": [
```

```

    {
      "name": "_FILELOG_PATH",
      "value": "/data/limbo-log"
    },
    {
      "name": "USE_AIPAAS_PROTO",
      "value": "1"
    }
  ],
  "gpu": [
    {
      "quota": 0.5,
      "limit": 0,
      "vendor": "NVidia",
      "accelerator": "2070",
      "capacities": {
        "stream_video_quota": 6,
        "stream_capture_quota": 0,
        "analyze_quota": 0
      },
      "resolution": "PIXELS_200W",
    }
  ]
},
"code": 0,
"msg": "success"
}

```

E.6 算法下发

- a) POST /CAP/AlgoPackages
- b) 请求参数

表 E.18 AlgoPackagePushRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|----|----|
|----|------|----|----|----|

| | | | | |
|---|-----|---|--------|--|
| 1 | id | R | String | |
| 2 | md5 | R | String | |
| 3 | url | R | String | |

c) 响应参数

表 E.19 AlgoPackagePushResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | |

e) 请求体示例

```
{
  "id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab",
  "md5": "",
  "url": "http://127.0.0.1/algoPackages"
}
```

e) 响应体示例:

```
{
  "data": {},
  "code": 0,
  "msg": "success"
}
```

E.7 算法删除

a) DELETE /CAP/AlgoPackages/{id}

b) 请求参数

表 E.20 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|--------|
| 1 | id | R | String | 算法 id |
| 2 | url | R | String | 算法 url |

c) 响应参数

表 E.21 AlgoPackageDeleteResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | |

d) 请求体示例

```
{
  "id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab"
}
```

e) 响应体示例:

```
{
  "data": {},
  "code": 0,
  "msg": "success"
}
```

E.8 算法列表

a) POST /CAP/AlgoPackages:search

b) 请求参数

表 E.22 AlgoPackageSearchRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------------|----|---------|------|
| 1 | pageNumber | R | Integer | 页码 |
| 2 | pageSize | R | Integer | 分页大小 |

c) 响应参数

表 E.23 AlgoPackageSearchRespons 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|----------------------------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | |
| 4 | └ id | | String | |
| 5 | └ url | | String | |
| 6 | └ relatedAlgoFilePackageid | | String | |
| 7 | └ status | | Integer | |

d) 请求体示例

```
{
  "pageNumber": 1,
  "pageSize": 20
}
```

e) 响应体示例:

```
{
  "data": {
```

```

    "pageNumber":1,
    "pageSize":20,
    "total":101,
    "records": [
      {
        "id": "",
        "relatedAlgoPackageid": "",
        "url": "",
        "status": ""
      }
    ]
  },
  "code":0,
  "msg":""
}

```

E.9 算法详情

- a) GET /CAP/AlgoPackages/{id}
 b) 请求参数

表 E.24 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|-------|
| 1 | id | R | String | 算法 id |

- c) 响应参数

表 E.25 AlgoPackageDetailResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|----------------------------|----|---------|----------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | 返回结果 |
| 4 | └ id | R | String | 算法 id |
| 5 | └ url | R | String | 地址 |
| 6 | └ relatedAlgoFilePackageId | R | String | 关联算法封装文件 |
| 7 | └ status | R | Integer | 状态 |

- d) 请求体示例

```

{

```

```
"id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab"
}
```

e) 响应体示例:

```
{
  "data": {
    "id": "7a90a32d-8ae8-4ffa-b000-07ef71058bab",
    "url": "http://127.0.0.1/",
    "relatedAlgoFilePackageId": "",
    "status": 0,
  },
  "code": 0,
  "msg": "success"
}
```

附录 F (规范性) 服务管理接口

F.1 算法服务创建

本附录共规定了4个服务管理接口，主要包括算法服务的创建、删除、查询等，见表F.1。其中，各个接口属性信息应符合表F.2-F.5的规定，具体应符合章节F.2-F.5的规定，算法算力平台与用户和解析中心之间服务管理接口调度流程见图F.1。

表 F.1 服务管理接口

| 序号 | URI | 方法 | 功能说明 |
|----|--------------------------|--------|----------|
| 1 | /CAP/AlgoServices | POST | 算法服务创建 |
| 2 | /CAP/AlgoServices/{id} | DELETE | 算法服务删除 |
| 3 | /CAP/AlgoServices:search | POST | 算法服务查询列表 |
| 4 | /CAP/AlgoServices/{id} | GET | 算法服务查询详情 |

表 F.2 算法服务创建

| | | | |
|------|-------------------|--------------------------|---------------------------|
| URI | /CAP/AlgoServices | | |
| 功能 | 算法服务创建 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoServiceCreateRequest | AlgoServiceCreateResponse |
| 注释 | 算法服务创建 | | |

表 F.3 算法服务删除

| | | | |
|--------|------------------------|-------|---------------------------|
| URI | /CAP/AlgoServices/{id} | | |
| 功能 | 算法服务删除 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 算法服务ID | 无 | AlgoServiceDeleteResponse |
| 注释 | 算法服务删除 | | |

表 F.4 算法服务查询列表

| | | | |
|------|--------------------------|--------------------------|---------------------------|
| URI | /CAP/AlgoServices:search | | |
| 功能 | 算法服务查询列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoServiceSearchRequest | AlgoServiceSearchResponse |
| 注释 | 算法服务查询列表 | | |

表 F.5 算法服务查询详情

| | | | |
|-----|------------------------|-------|---------------------------|
| URI | /CAP/AlgoServices/{id} | | |
| 功能 | 算法服务查询详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 算法服务ID | 无 | AlgoServiceDetailResponse |
| 注释 | 算法服务查询详情 | | |

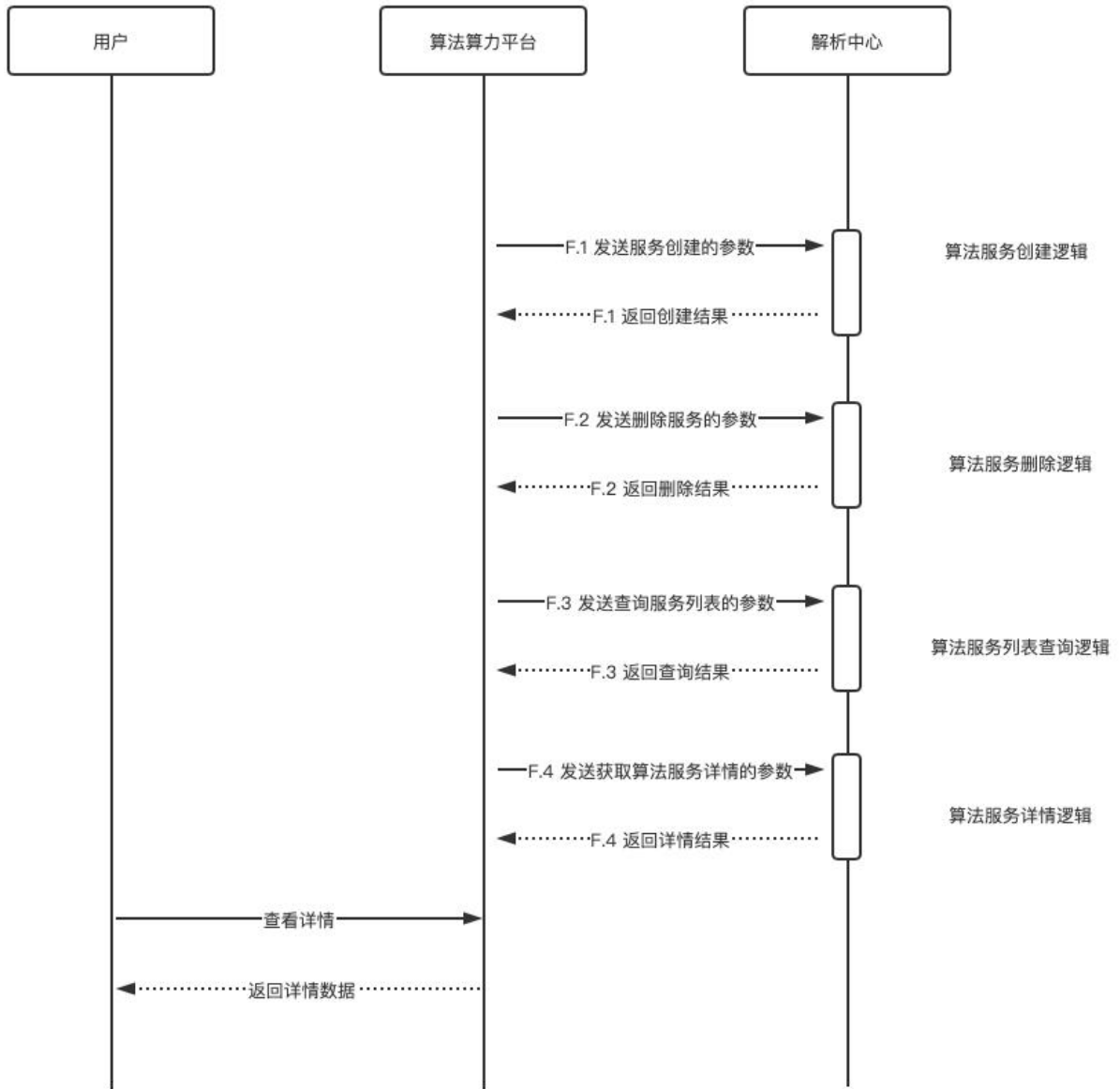


图 F.1 服务管理接口交互

F.2 算法服务创建

- a) POST /CAP/AlgoServices
- b) 请求参数

表 F.6 AlgoServiceCreateRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|----------|----|---------|----------|
| 1 | envs | R | String | 算法服务变量注入 |
| 2 | replicas | R | Integer | 副本数 |
| 3 | gpuType | R | String | gpu 类型 |
| 4 | algoId | R | String | 算法 id |

- c) 响应参数

表 F.7 AlgoServiceCreateResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | String | 服务 id |

d) 请求体示例

```
{
  "envs": "threadCount=8, timeOut=2000, k1=v1, k2=v2",
  "replicas": 2,
  "gpuType": "NVidia A2",
  "algoId": "MEGCITY:3843938347"
}
```

e) 响应体示例:

```
{
  "data": {},
  "code": 0,
  "msg": "success"
}
```

F.3 算法服务删除

a) DELETE /CAP/AlgoServices/{id}

b) 请求参数

表 F.8 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|-------|
| 1 | id | R | String | 服务 id |

c) 响应参数

表 F.9 AlgoServiceDeleteResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Boolean | 参见示例 |

d) 请求体示例

无

e) 响应体示例:

```
{
    "code": 0,
    "msg": "success",
    "data": {}
}
```

F.4 算法服务查询列表

a) POST /CAP/AlgoServices:search

b) 请求参数

表 F.10 AlgoServiceSearchRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------------|----|---------|------|
| 1 | pageSize | 0 | Integer | 分页长度 |
| 2 | pageNumber | 0 | Integer | 页码 |

c) 响应参数

表 F.11 AlgoServiceSearchResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|-----------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | 参见示例 |
| 4 | └id | R | String | 服务 id |
| 5 | └algoId | R | String | 算法 id |
| 6 | └replicas | R | Integer | 副本数 |
| 7 | └gpuType | R | String | gpu 类型 |
| 8 | └envs | R | String | 注入的变量 |

d) 请求体示例

```
{
    "pageNumber": 1,
    "pageSize": 20,
}
```

e) 响应体示例:

```
{
    "code": 0,
    "msg": "success",
    "data": {
```

```

    "pageNumber": 1,
    "pageSize": 20,
    "total": 101,
    "records": {
      ["envs":
"threadCount=8, timeOut=2000, k1=v1, k2=v2",
      "replicas": 2,
      "gpuType": "Nvidia A2",
      "algoId": "MEGCITY:3843938347",
      "id": "28273"
    }
  ]
}

```

F.5 算法服务查询详情

- a) GET /CAP/AlgoServices/{id}
- b) 请求参数

表 F.12 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|---------|
| 1 | id | R | String | 算法服务 id |

- c) 响应参数

表 F.13 AlgoServiceDetailResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|-----------|----|---------|---|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | 参见示例 |
| 4 | └id | R | String | 服务 id |
| 5 | └algoId | R | String | 算法 id |
| 6 | └replicas | R | Integer | 副本数 |
| 7 | └gpuType | R | String | gpu 类型 |
| 8 | └envs | R | String | 注入的变量 |
| 9 | └status | R | String | 服务状态 [starting: 启动中 runing: 运行中 error: 异常 |

| | | | | |
|----|-----------------|---|----------|---------------|
| | | | |] |
| 10 | └gpuQuota | R | Double | 单算法需要的 gpu 资源 |
| 11 | └replicasDetail | R | Object[] | 副本明细 |
| 12 | └nodeHost | R | String | 节点 hostName |
| 13 | └nodeIpv4 | R | String | 节点 IP |

d) 请求体示例
无

e) 响应体示例:

```
{
  "code": 0,
  "msg": "success",
  "data": {
    "envs":
"threadCount=8, timeOut=2000, k1=v1, k2=v2",
    "replicas": 2,
    "gpuType": "NVidia A2",
    "algoId": "MEGCITY:3843938347",
    "status": "runing",
    "id": "28273",
    "gpuQuota": 0.25,
    "replicasDetail": [{
      "node":
"jy-test-kssdr1000h31802110029. ops. megvii-inc. com",
      "nodeIpv4":
"jy-test-kssdr1000h31802110029. ops. megvii-inc. com"
    }, {
      "node":
"jy-test-kssdr1000h31802110030. ops. megvii-inc. com",
      "nodeIpv4":
"jy-test-kssdr1000h31802110030. ops. megvii-inc. com"
    }
  ]
}
```


附录 G
(规范性)
任务管理接口

G.1 基本要求

本附录共规定了6个任务管理接口，主要包括任务的创建、删除、启动、暂停、查询等，见表G.1。其中，各个接口属性信息应符合表G.2-G.7的规定，具体应符合章节G.2-G.7的规定，算法算力平台与用户和解析中心之间任务管理接口调度流程见图G.1。

表 G.1 任务管理接口

| 序号 | URI | 方法 | 功能说明 |
|----|----------------------------|--------|--------|
| 1 | /CAP/AlgoTasks | POST | 任务创建 |
| 2 | /CAP/AlgoTasks/{id} | DELETE | 任务删除 |
| 3 | /CAP/AlgoTasks/{id}:resume | POST | 任务启动 |
| 4 | /CAP/AlgoTasks/{id}:pause | POST | 任务暂停 |
| 5 | /CAP/AlgoTasks:search | POST | 任务查询列表 |
| 6 | /CAP/AlgoTasks/{id} | GET | 任务查询详情 |

表 G.2 任务创建

| | | | |
|------|----------------|-----------------------|------------------------|
| URI | /CAP/AlgoTasks | | |
| 功能 | 任务创建 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoTaskCreateRequest | AlgoTaskCreateResponse |
| 注释 | 任务创建 | | |

表 G.3 任务删除

| | | | |
|--------|---------------------|-------|------------------------|
| URI | /CAP/AlgoTasks/{id} | | |
| 功能 | 任务删除 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 任务ID | 无 | AlgoTaskDeleteResponse |
| 注释 | 任务删除 | | |

表 G.4 任务启动

| | | | |
|------|----------------------------|-------|------------------------|
| URI | /CAP/AlgoTasks/{id}:resume | | |
| 功能 | 任务启动 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 任务ID | 无 | AlgoTaskResumeResponse |
| 注释 | 任务启动 | | |

表 G.5 任务暂停

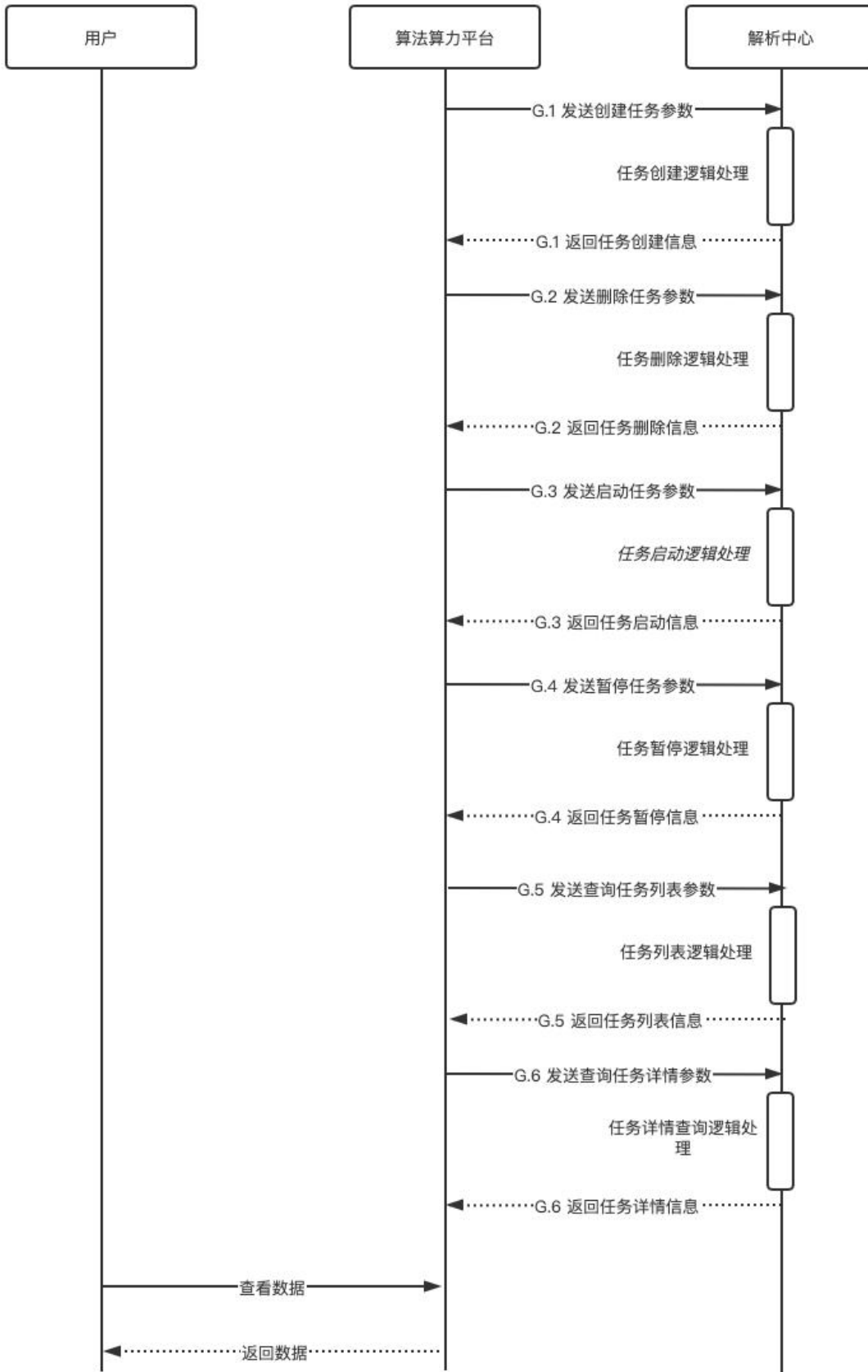
| | | | |
|--------|---------------------------|-------|-----------------------|
| URI | /CAP/AlgoTasks/{id}:pause | | |
| 功能 | 任务暂停 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 任务ID | 无 | AlgoTaskPauseResponse |
| 注释 | 任务暂停 | | |

表 G.6 任务查询列表

| | | | |
|------|-----------------------|-----------------------|------------------------|
| URI | /CAP/AlgoTasks:search | | |
| 功能 | 任务查询列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoTaskSearchRequest | AlgoTaskSearchResponse |
| 注释 | 任务查询列表 | | |

表 G.7 任务查询详情

| | | | |
|-----|---------------------|-------|------------------------|
| URI | /CAP/AlgoTasks/{id} | | |
| 功能 | 任务查询详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 任务ID | 无 | AlgoTaskDetailResponse |
| 注释 | 任务查询详情 | | |



图G.1 任务管理接口交互

G.2 任务创建

- a) POST /CAP/AlgoTasks
b) 请求参数

表 G.8 AlgoTaskCreateRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|----------------------|----|---------------------|---|
| 1 | algoServiceId | R | String | 算法服务 id |
| 2 | restartPolicy | R | String | 重启策略[不重启:NEVER 重启: ALWAYS] |
| 3 | taskInfo | R | Object | 任务信息 |
| 4 | └ labels | R | Object | 任务标签 |
| 5 | └ algoCabinName | R | String | 算法名 |
| 6 | └ version | R | String | 算法版本 |
| 7 | └ rulesParams | R | Object[] | 算法规则参数 |
| 8 | └ areas | R | Object[] | 区域信息集合 |
| 9 | └ areaid | R | Integer | 区域 id |
| 10 | └ areaType | R | String | 区域类型[多边形:POLYGON 单向单绊线 :SINGLE_LINE 双向单绊线 :SINGLE_LINE_BOTH 单向多绊线 :MULTI_LINE 双向多绊线 :MULTI_LINE_BOTH] |
| 11 | └ points | R | Object[] | 点位信息 |
| 12 | └ x | R | Double | 横坐标 |
| 13 | └ y | R | Double | 纵坐标 |
| 14 | └ destinations | R | Object | 解析结果存放 |
| 15 | └ HTTPDest | R | Object | http 方式 |
| 16 | └ url | R | String | 上报 url |
| 17 | └ eventType | R | String | 事件类型 |
| 18 | └ extendParams | 0 | Object | 扩展参数 |
| 19 | └ labels | 0 | Map<String, String> | 业务标签信息 |
| 20 | └ masks | 0 | Object[] | 排除区域 |
| 21 | └ streamOptions | R | Object | 流信息 |
| 22 | └ videoStreamOptions | R | Object | 视频流信息 |

| | | | | |
|----|--------------|---|--------|-----|
| | ns | | | |
| 23 | └ resolution | R | String | 协议 |
| 24 | └ url | R | String | 流地址 |

c) 响应参数

表 G.9 AlgoTaskCreateResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|-------------------------|----|----------|--|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Object | 参见示例 |
| 4 | └ createdStreams | R | Object[] | 创建的解析任务信息 |
| 5 | └ algoServiceId | 0 | String | 算法服务 id |
| 6 | └ taskInfo | 0 | Object | 解析任务信息 |
| 7 | └ id | 0 | String | 解析任务 id |
| 8 | └ status | 0 | String | 任务状态[等待:PENDING 运行:RUNNING 暂停:PAUSED 完成:FINISHED 错误:ERROR] |
| 9 | └ streamOptions | 0 | Object | 流信息 |
| 10 | └ VideoStreamOptions | 0 | Object | 视频流信息 |
| 11 | └ url | 0 | String | 视频流 url |
| 12 | updateStreams | 0 | Object[] | 更新的任务信息 |
| 13 | └ createdStreams | R | Object[] | 创建的解析任务信息 |
| 14 | └ algoServiceId | 0 | String | 算法服务 id |
| 15 | └ taskInfo | 0 | Object | 解析任务信息 |
| 16 | └ id | 0 | String | 解析任务 id |
| 17 | └ status | 0 | String | 任务状态[等待:PENDING 运行:RUNNING 暂停:PAUSED 完成:FINISHED 错误:ERROR] |

| | | | | |
|----|-----------------------|---|--------|---------|
| | | | |] |
| 18 | └─ streamOptions | 0 | Object | 流信息 |
| 19 | └─ VideoStreamOptions | 0 | Object | 视频流信息 |
| 20 | └─ url | 0 | String | 视频流 url |

d) 请求体示例

```

{
  "algoServiceId": "megvii/cloud/megcity-a2_2",
  "restartPolicy": "ALWAYS",
  "taskInfo": {
    "labels": {
      "algoCabinName": "megcity-a2",
      "version": "2.0.1"
    },
    "rulesParams": [{
      "areas": [{
        "areaid": 0,
        "areaType": "POLYGON",
        "points": [{
          "x": 0.0,
          "y": 1.0
        }, {
          "x": 1.0,
          "y": 1.0
        }, {
          "x": 1.0,
          "y": 0.0
        }, {
          "x": 0.0,
          "y": 0.0
        }
      ]
    }],
    "destinations": {
      "eventDest": {
        "mqDest": {
          "mqType": "KAFKA",
          "topic": "PAAS_CAPTURE_ALL",
          "url": "10.172.198.207:23310"
        },
        "imageDest": {
          "cropImageDest":
"aipaas-crop-image",
          "fullImageDest": "aipaas-full-image"
        }
      },
      "eventType": "ILLEGAL_ADV",
      "extendParams": {
        "sensitivity": "低",
        "durationThreshold": 1,
        "repeatAlarmIntervalThreshold": 720
      },
      "labels": {

```



```

        },
        "masks": []
    }],
    "streamOptions": {
        "videoStreamOptions": {
            "resolution": "PIXELS_200W",
            "url":
"iot://10.172.198.207/?deviceid=101234200200000046&mode=video"
        }
    }
}
    
```

e) 响应体示例:

```

{
  "data": {
    "createdStreams": [
      "algoServiceId": "",
      "taskInfo": {
        "id": "1321adad1312312",
        "status": "PENDING",
        "streamOptions": {
          "videoStreamOptions": {
            "url":
"iot://10.172.198.207/?deviceid=101234200200000046&mode=video"
          }
        }
      }
    ],
    "updatedStreams": [],
    "code": 0,
    "msg": "success"
  }
}
    
```

G.3 任务删除

- a) DELETE /CAP/AlgoTasks/{id}
- b) 请求参数

表 G.10 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|---------|
| 1 | id | R | String | 解析任务 id |

c) 响应参数

表 G.11 AlgoTaskDeleteResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|----|----|
|----|------|----|----|----|

| | | | | |
|---|------|---|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Boolean | 参见示例 |

d) 请求体示例

无

e) 响应体示例:

```
{
  "data": {},
  "code": 0,
  "msg": "success"
}
```

G.4 任务启动

a) POST /CAP/AlgoTasks/{id}:resume

b) 请求参数

表 G.12 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|---------|
| 1 | id | R | String | 解析任务 id |

c) 响应参数

表 G.13 AlgoTaskResumeResponse

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Boolean | 参见示例 |

d) 请求体示例

无

e) 响应体示例:

```
{
  "data": true,
  "code": 0,
  "msg": "success"
}
```

G.5 任务暂停

a) POST /CAP/AlgoTasks/{id}:pause

b) 请求参数

表 G.14 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|----|----|
|----|------|----|----|----|

| | | | | |
|---|----|---|--------|---------|
| 1 | id | R | String | 解析任务 id |
|---|----|---|--------|---------|

c) 响应参数

表 G. 15 AlgoTaskPauseResponse

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | msg | R | String | 请求结果描述 |
| 3 | data | R | Boolean | 参见示例 |

d) 请求体示例

{}
}

e) 响应体示例:

```

{
    "data": {},
    "code": 0,
    "msg": "success"
}

```

G. 6 任务查询列表

a) POST /CAP/AlgoTasks:search

b) 请求参数

表 G. 16 AlgoTaskSearchRequest 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|---------------|----|---------|---------|
| 1 | id | 0 | String | 解析任务 id |
| 2 | algoServiceId | 0 | String | 算法服务 id |
| 3 | eventType | 0 | String | 事件类型 |
| 4 | pageSize | R | Integer | size |
| 5 | pageNumber | R | Integer | 页码 |

c) 响应参数

表 G. 17 AlgoTaskSearchResponse

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|--------------|----|---------|---------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | mgs | R | String | 请求结果描述 |
| 3 | data | R | Object | 参加示例 |
| 4 | └ total | R | Integer | 10 |
| 5 | └ pageNumber | R | Integer | 1 |
| 6 | └ pageSize | R | Integer | 10 |

| | | | | |
|----|-----------------|---|---------------------|--|
| 7 | └ records | R | Object[] | 列表 |
| 8 | └ algoServiceId | R | String | 算法服务 id |
| 9 | └ restartPolicy | R | String | 重启策略[不重启: NEVER 重启: ALWAYS] |
| 10 | └ taskInfo | R | Object | 任务信息 |
| 11 | └ id | R | String | 任务 id |
| 12 | └ labels | R | Map<String, String> | 业务标签 |
| 13 | └ algoCabinName | R | String | 算法名 |
| 14 | └ version | R | String | 算法版本 |
| 15 | └ status | R | String | 状态 |

d) 请求体示例

```
{
    "pageNumber": 1,
    "pageSize": 20,
}
```

e) 响应体示例:

```
{
    "code": 0,
    "msg": "success",
    "data": {
        "pageNumber": 1,
        "pageSize": 20,
        "total": 101,
        "records": {
            "algoServiceId": "megvii/cloud/megcity-a2_2",
            "restartPolicy": "ALWAYS",
            "taskInfo": {
                "id": "12313123",
                "labels": {
                    "algoCabinName": "megcity-a2",
                    "version": "2.0.1"
                }
            }
        }
    }
}
```

```

        },
        "status": "PENDING"
    }
]
}
    
```

G.7 任务查询详情

- a) GET /CAP/AlgoTasks/{id}
- b) 请求参数

表 G.18 请求参数属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|------|----|--------|---------|
| 1 | id | R | String | 解析任务 id |

- c) 响应参数

表 G.19 AlgoTaskDetailResponse 属性

| 序号 | 参数名称 | 必填 | 类型 | 说明 |
|----|-----------------|----|---------------------|--------------------------------|
| 1 | code | R | Integer | 请求结果状态码 |
| 2 | mgs | R | String | 请求结果描述 |
| 3 | data | R | Object | 参加示例 |
| 4 | └ algoServiceId | R | String | 算法服务 id |
| 5 | └ restartPolicy | R | String | 重启策略 [不重启: NEVER 重启: ALWAYS] |
| 6 | └ taskInfo | R | Object | 任务信息 |
| 7 | └ id | R | String | 任务 id |
| 8 | └ labels | R | Map<String, String> | 业务标签 |
| 9 | └ algoCabinName | R | String | 算法名 |
| 10 | └ version | R | String | 算法版本 |
| 11 | └ status | R | String | 状态 |

- d) 请求体示例
无
- e) 响应体示例:

```
{  
  "data": {  
    "algoServiceId": "megvii/cloud/megcity-a2_2",  
    "restartPolicy": "ALWAYS",  
    "taskInfo": {  
      "id": "12313123",  
      "labels": {  
        "algoCabinName": "megcity-a2",  
        "version": "2.0.1"}  
      },  
      "status": "PENDING"  
    },  
    "code": 0,  
    "msg": "success"  
  }  
}
```

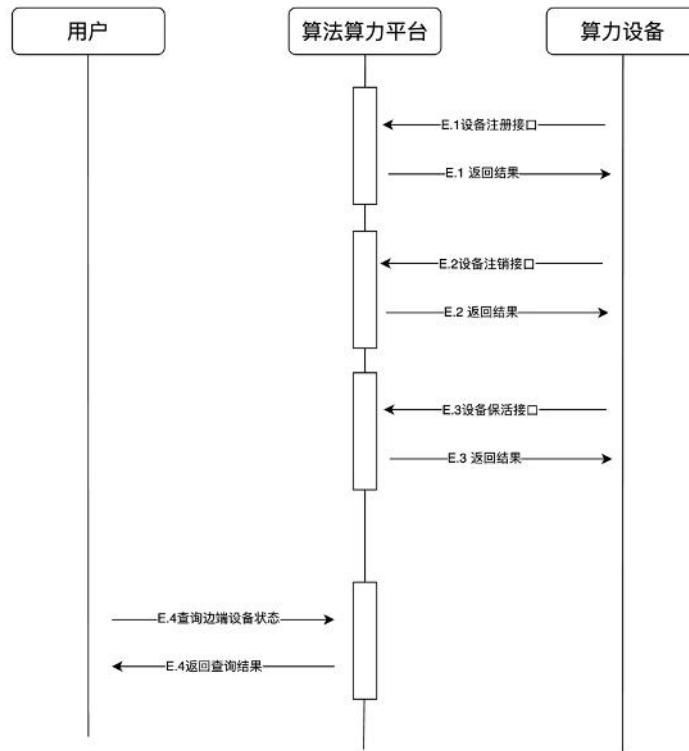
附录 H
(规范性)
机器算力接入接口

H.1 基本要求

本附录共规定了4个机器算力接入接口，主要包括机器的注册、注销、保活和查询等，见表H.1。其中，各个接口属性信息应符合表H.2-H.5的规定，具体应符合章节H.2-H.5的规定，算法算力平台与用户和算力设备之间接口调度流程见图H.1。

表 H.1 机器算力接入接口

| 序号 | URI | 方法 | 功能说明 |
|----|----------------------------------|------|-----------|
| 1 | /CAD/machine/register | POST | 注册机器 |
| 2 | /CAD/machine/unregister | POST | 注销机器 |
| 3 | /CAD/machine/keepalive | POST | 保活机器状态 |
| 4 | /CAD/machine/status/{deviceCode} | GET | 查询机器状态及资源 |



图H.1 机器算力接口调用流程

表 H.2 注册机器

| | | | |
|------|-----------------------|-----------------|---------------|
| URI | /CAD/machine/register | | |
| 功能 | 注册机器 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | RegisterRequest | EmptyResponse |
| 注释 | 注册机器 | | |

表 H.3 注销机器

| | |
|-----|-------------------------|
| URI | /CAD/machine/unregister |
| 功能 | 注销机器 |

| | | | |
|------|-------|-------------------|---------------|
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | UnRegisterRequest | EmptyResponse |
| 注释 | 注销机器 | | |

表 H.4 保活机器状态

| | | | |
|------|--|------------------|---------------|
| URI | /CAD/machine/keepalive | | |
| 功能 | 保活机器状态 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | KeepaliveRequest | EmptyResponse |
| 注释 | 机器算力需实现机器保活接口，定期向平台发送保活信息，超时时间内没收到保活信息会认为机器丢失。 | | |

表 H.5 查询机器状态及资源

| | | | |
|-----|----------------------------------|-------|----------------------|
| URI | /CAD/machine/status/{deviceCode} | | |
| 功能 | 查询机器状态及资源 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| GET | 无 | | DeviceStatusResponse |
| 注释 | 查询机器状态及资源 | | |

H.2 注册机器

- a) POST /CAD/machine/register
- b) 请求参数

表 H.6 RegisterRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|----------|----------|---------|----|-------|------------------------|
| 1 | code | code | String | | R | 编码 code |
| 2 | ip | ip | String | | R | ip |
| 3 | port | port | Integer | | R | port |
| 4 | username | username | String | | 0 | API 服务用户名 |
| 5 | password | password | String | | 0 | 宜采用国密 SM2 加密算法对密码加密后传输 |

- c) 响应参数
EmptyResponse
- d) 请求体示例

```
{
  "code": "machine_code",
  "ip": "127.0.0.1",
  "port": 8123,
  "username": "username",
  "password": "password"
}
```

响应体示例:


```
{
  "data": null,
  "code": "0",
  "message": "success"
}
```

H.3 注销机器

- a) POST /CAD/machine/unregister
- b) 请求参数

表 H.7 UnRegisterRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|----------|----------|--------|----|-------|---------|
| 1 | code | code | String | | R | 编码 code |
| 3 | username | username | String | | R | 用户名 |
| 4 | password | password | String | | R | 密码 |

- c) 响应参数
EmptyResponse

- d) 请求体示例

```
{
  "deviceCode": "device_code",
  "username": "username",
  "password": "password"
}
```

响应体示例:

```
{
  "data": null
  "code": "0",
  "message": "success"
}
```

H.4 保活机器状态

- a) POST /CAD/machine/keepalive
- b) 请求参数

表 H.8 KeepaliveRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|---------|---------|---------|----|-------|-----------|
| 1 | code | code | String | | R | 编码 code |
| 2 | expires | expires | Integer | | R | 超时时间, 单位秒 |

- c) 响应参数

EmptyResponse

d) 请求体示例

```
{
  "code": "machine_code",
  "expires":60
}
```

e) 响应体示例

```
{
  "data": null
  "code": "0",
  "message": "success"
}
```

H.5 查询机器状态及资源

a) GET /CAD/machine/status/{code}

b) 响应参数

表 H.9 DeviceStatusResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|---------|---------|------------|----|-------|-----------------------|
| 1 | code | code | String | | R | 结果状态码 (0: 成功, 其他: 失败) |
| 2 | message | message | String | | R | 请求结果描述 |
| 3 | data | data | DeviceInfo | | 0 | 边端设备状态信息 |

表 H.10 DeviceInfo 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|----------------|----------------|---------|----|-------|----------------------|
| 2 | code | deviceCode | String | | R | code |
| 3 | isOnline | isOnline | Boolean | | R | 是 0 在线 |
| 4 | cpuUtilization | cpuUtilization | String | | 0 | cpu 利用率, 0.11 表示 11% |
| 5 | memUsage | memUsage | String | | 0 | 内存占用, 单位 GB |
| 6 | memFree | memFree | String | | 0 | 内存剩余, 单位 GB |
| 7 | diskUsage | diskUsage | String | | 0 | 磁盘占用, 单位 GB |
| 8 | diskFree | diskFree | String | | 0 | 磁盘剩余, 单位 GB |

c) 请求体示例

```
{
  "code": "machine_code"
}
```

d) 响应体示例:

```
{
  "data": {
    "code": "machine_code",
    "isOnline": true,
    "cpuUtilization": "0.11",
    "memUsage": "24.395",
    "memFree": " 39.605 "
    "diskUsage": "128.1",
    "diskFree": "127.9"
  },
  "code": "0",
  "message": "success"
}
```

附录 I
(规范性)
机器模式算法包部署接口

I.1 下发算法封装文件

本附录共规定了6个机器模式算法包部署接口，主要包括算法封装文件、算法包等的下发、删除、安装卸载等，见表I.1。其中，各个接口属性信息应符合表I.2-I.7的规定，具体应符合章节H.2-H.7的规定，算法算力平台与中心机器之间接口调度流程见图I.1。

表 I.1 机器模式算法包部署接口

| 序号 | URI | 方法 | 功能说明 |
|----|---|--------|----------|
| 1 | /CAD/machine/algoFilePackages | POST | 下发算法封装文件 |
| 2 | /CAD/machine/algoFilePackages/{algoFilePackageId} | DELETE | 删除算法封装文件 |
| 3 | /CAD/machine/algoPackages | POST | 安装算法包 |
| 4 | /CAD/machine/algoPackages/{algoPackageCode} | DELETE | 卸载算法包 |
| 5 | /CAD/machine/algoPackages | GET | 查询算法包列表 |
| 6 | /CAD/machine/algoPackages/{algoPackageCode} | GET | 查询算法包详情 |

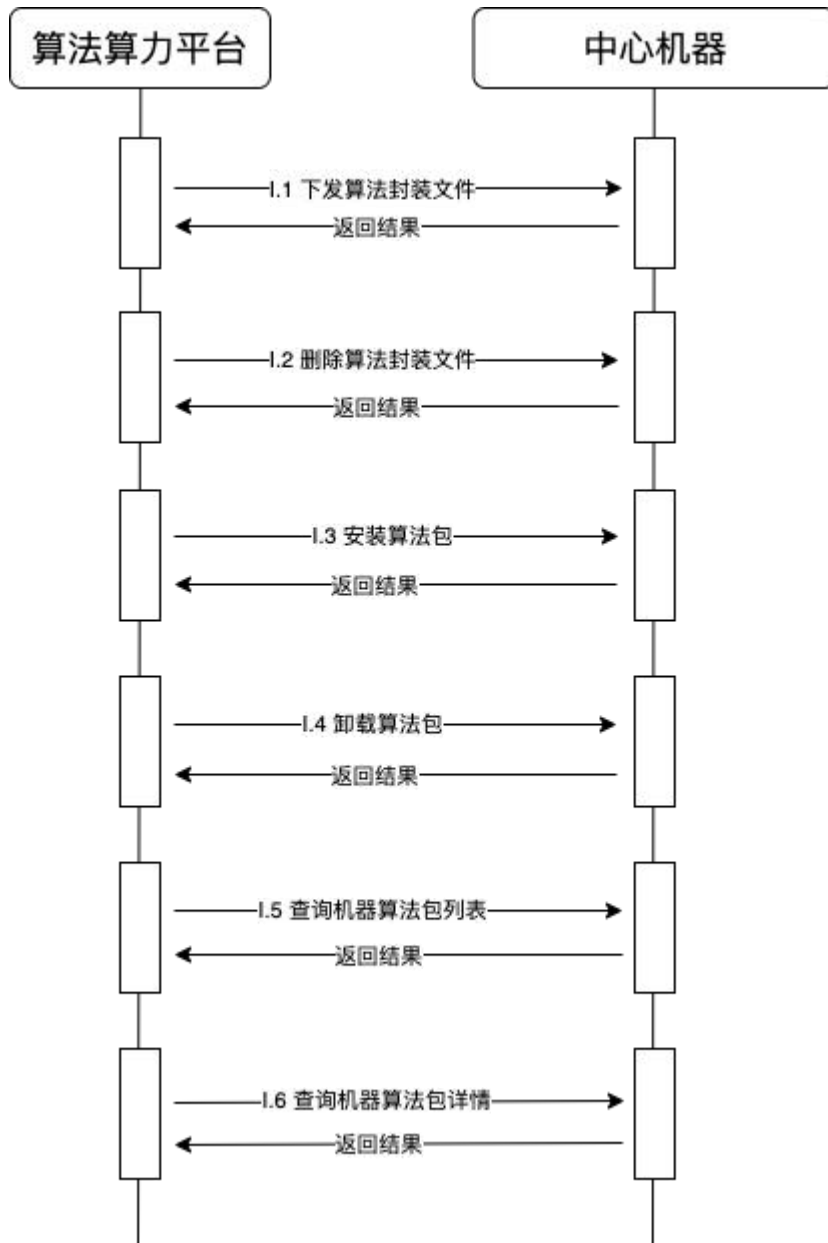


图 I.1 算法包调用接口流程

表 I.2 下发算法封装文件

| | | | |
|------|-------------------------------|------------------------|---------------|
| URI | /CAD/machine/algoFilePackages | | |
| 功能 | 下发算法封装文件 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoFilePackageRequest | EmptyResponse |
| 注释 | 将算法封装文件下发到机器 | | |

表 I.3 删除算法封装文件

| | | | |
|--------|---|-------|---------------|
| URI | /CAD/machine/algoFilePackages/{algoFilePackageId} | | |
| 功能 | 删除算法封装文件包 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 无 | 无 | EmptyResponse |
| 注释 | 从机器删除算法包 | | |

表 1.4 安装算法包

| | | | |
|------|---|---------------------------|---------------|
| URI | /CAD/machine/algoPackages | | |
| 功能 | 安装算法包 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoPackageInstallRequest | EmptyResponse |
| 注释 | 机器算法包安装 算法包通过T/BAX XXXX. 5-202X中A.1 算法封装文件导入，算法封装文件中应包含算法包编码 algoPackageCode，算法包编码由用户自定义，需要唯一，建议格式：provider-算法英文-version，由大小写字母、下划线、横线-、英文.组成，长度小于128字符，例如 BAIDU-HUMAN_ATTR_ALERT-1.0.0。 | | |

表 1.5 卸载算法包

| | | | |
|--------|---|-------|---------------|
| URI | /CAD/machine/algoPackages/{algoPackageCode} | | |
| 功能 | 卸载算法包 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| DELETE | 无 | 无 | EmptyResponse |
| 注释 | 机器算法包卸载 | | |

表 1.6 查询算法包列表

| | | | |
|------|---------------------------|------------------------|-------------------------|
| URI | /CAD/machine/algoPackages | | |
| 功能 | 查询算法包列表 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | AlgoPackageListRequest | AlgoPackageListResponse |
| 注释 | 查询算法包列表 | | |

表 1.7 查询算法包详情

| | | | |
|------|---|-------|---------------------------|
| URI | /CAD/machine/algoPackages/{algoPackageCode} | | |
| 功能 | 查询算法包详情 | | |
| 方法 | 查询字符串 | 请求消息体 | 响应消息体 |
| POST | 无 | 无 | AlgoPackageDetailResponse |
| 注释 | 查询算法包详情 | | |

1.2 下发算法封装文件

- a) URI: POST /CAD/machine/algoFilePackages
- b) 请求参数:

表 1.8 AlgoFilePackageRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-----------|-------------------|--------|----|-------|--------------------------|
| 1 | 算法封装文件 ID | algoFilePackageId | Long | | R | 算法封装文件 ID |
| 3 | 鉴权地址 | authAddress | String | | 0 | 鉴权地址，默认给中心地址 IP:AuthPort |
| 4 | 鉴权文件地址 | licenseFile | String | | 0 | 鉴权文件地址 |

- c) 响应参数:

EmptyResponse

- d) 请求示例

```
{
  "algoFilePackageId":1,
```

```

    "authAddress": "127.0.0.1:8080",
    "licenseFile": "/auth_license.txt"
}

```

e) 响应示例

```

{
  "code": "0",
  "message": "success"
}

```

I.3 删除算法封装文件

a) URI: DELETE /CAD/machine/algoFilePackages/{algoFilePackageId}

b) 请求参数:

表 I.9 请求参数属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-----------|-------------------|------|----|-------|-----------|
| 1 | 算法封装文件 ID | algoFilePackageId | Long | | R | 算法封装文件 ID |

c) 响应参数:

EmptyResponse

d) 请求示例:

DELETE /CAD/machine/algoFilePackages/1

e) 响应示例:

```

{
  "code": "0",
  "message": "success"
}

```

I.4 安装算法包

a) URI: POST /CAD/machine/algoPackages

b) 请求参数:

表 I.10 AlgoPackageInstallRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------------|--------|----|-------|----------------------|
| 1 | 算法包编码 | algoPackageCode | String | | R | 算法包编码,用户在算法封装文件包中自定义 |
| 3 | 运行参数 | args | String | | 0 | 运行参数 |

c) 响应参数:

EmptyResponse

d) 请求示例:

```

{
  "algoPackageCode": "BAIDU-HUMAN_ATTR_ALERT-1.0.0",
  "args": "xpuType=NVIDIA_T4"
}

```

e) 响应示例:

```

{
  "code": "0",
  "message": "success"
}

```

I.5 卸载算法包

- a) URI: DELETE /CAD/machine/algopackages/{algoPackageCode}
 b) 请求参数

表 I.11 请求参数属性

| 序号 | 名称 | 标识符 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------------|----|-------|-------|
| 1 | 算法包编码 | algoPackageCode | | R | 算法包编码 |

- c) 响应参数:

EmptyResponse

- d) 请求示例:

DELETE /CAD/machine/algopackages/BAIDU-HUMAN_ATTR_ALERT-1.0.0

- e) 响应示例:

```
{
  "code": "0",
  "message": "success"
}
```

1.6 查询机器算法包列表

- a) URI: GET /CAD/machine/algopackages
 b) 请求参数:

表 I.12 AlgoPackageListRequest 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|---------|--------|--------|----|-------|--|
| 2 | 算法包运行状态 | status | String | | 0 | 算法包运行状态(RUNNING: 运行中(安装即运行), STOPPED(停止)), 不传表示查全部状态的包 |

- c) 响应参数:

表 I.13 AlgoPackageListResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------|----------------------|----|-------|----------------------|
| 1 | 结果状态码 | code | String | | R | 结果状态码(0: 成功, 其他: 失败) |
| 2 | 结果描述 | message | String | | R | 结果描述 |
| 3 | 数据 | data | MachineAlgoPackage[] | | R | 算法包列表 |
| 4 | 请求 id | requestId | String | | 0 | 请求 ID |

表 I.14 MachineAlgoPackage 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------------|--------|----|-------|-------|
| 1 | 算法包名称 | algoPackageName | String | | R | 算法包名称 |

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|---------|--------------------|--------|----|-------|--|
| 2 | 算法包类型 | algoPackageType | String | | R | 算法包类型, BIN: 二进制算法包, ENGINE: 算法引擎 |
| 3 | 算法包版本号 | algoPackageVersion | String | | R | 算法包版本号 |
| 4 | 算法包编码 | algoPackageCode | String | | R | 算法包编码 |
| 5 | 算法包运行状态 | status | String | | 0 | 算法包运行状态 (RUNNING: 运行中, STOPPED (停止)) |
| 6 | 创建时间 | createdDate | String | | 0 | 创建时间 yyyy-MM-dd HH:mm:ss.sss, 例: 2019-06-01 23:15:22.123 |

d) 请求示例:

```
{
  "status": "RUNNING"
}
```

e) 响应示例:

```
{
  "code": "0",
  "message": "success",
  "data": [
    {
      "algoPackageCode": "BAIDU-HUMAN_ATTR_ALERT-1.0.0",
      "algoPackageName": "人体结构化",
      "algoPackageVersion": "1.0.0",
      "algoPackageType": "BIN",
      "status": "RUNNING",
      "createdDate": "2023-04-24 23:15:22"
    }
  ]
}
```

I.7 查询机器算法包详情

a) URI: GET /CAD/machine/algorithmPackages/{algoPackageCode}

b) 请求参数

表 I.15 请求参数属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------------|--------|----|-------|-------|
| 1 | 算法包编码 | algoPackageCode | String | | R | 算法包编码 |

c) 响应参数:

表 I.16 AlgoPackageDetailResponse 属性

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|----|-----|----|----|-------|----|
|----|----|-----|----|----|-------|----|

| 序号 | 名称 | 标识符 | 类型 | 长度 | 必选/可选 | 说明 |
|----|-------|-----------|-------------------|----|-------|-----------------------|
| 1 | 结果状态码 | code | String | | R | 结果状态码 (0: 成功, 其他: 失败) |
| 2 | 结果描述 | message | String | | R | 结果描述 |
| 3 | 数据 | data | DeviceAlgoPackage | | R | 算法包信息 |
| 4 | 请求 id | requestId | String | | 0 | 请求 ID |

d) 请求示例:

GET /CAD/algoPackage/devices/device-1/BAIDU-HUMAN_ATTR_ALERT-1.0.0

e) 响应示例:

```
{
  "code": "0",
  "message": "success",
  "data": {
    "algoPackageCode": "BAIDU-HUMAN_ATTR_ALERT-1.0.0",
    "algoPackageName": "人体属性事件",
    "algoPackageVersion": "1.0.0",
    "algoPackageType": "BIN",
    "status": "RUNNING",
    "createdDate": "2023-04-24 23:15:22"
  }
}
```

参 考 文 献
