

ICS 35.020

CCS L 70

团 体 标 准

T/AI 116.6—2026

信息技术 数字视网膜系统 第 6 部分：端边云协同

Information technology—Digital retina systems—
Part 6: End-edge cloud collaboration

2026-05-29 发布

2025-05-29 实施

中关村视听产业技术创新联盟 发布

T/AI 176.6-2026



版权保护文件

版权所有归属于该标准的发布机构，除非有其他规定，否则未经许可，此发行物及其章节不得以其他形式或任何手段进行复制、再版或使用，包括电子版，影印件，或发布在互联网及内部网络等。使用许可可于发布机构获取。

目 次

前言	III
引言	IV
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 协同架构	2
5.1 基础协同架构	2
5.2 级联协同架构	3
6 系统协同框架	4
7 协同处理模型	5
7.1 参考模型	5
7.2 功能模型	6
7.3 协同处理过程	7
8 协同功能要求	7
8.1 总体要求	7
8.2 基本功能要求	9
8.3 应用能力要求	11
9 通信流程和接口	13
9.1 通用要求	13
9.2 设备资源查询接口	14
9.3 设备资源上报接口	19
9.4 设备算法信息查询接口	21
9.5 任务请求发送接口	22
9.6 任务调度接口	24
9.7 任务控制接口	26
9.8 任务结果发送接口	28
9.9 任务状态查询接口	30
9.10 任务状态上报接口	32
附录 A（资料性） 任务信息格式	34

附录 B（资料性） 端边云协同工作交互过程.....	36
B.1 基础协同模式.....	36
B.2 级联协同模式.....	38
参考文献.....	43

T/AI 116.6-2026

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/AI 116《信息技术 数字视网膜系统》的第6部分。T/AI 116已经发布了以下部分：

- 第1部分：系统结构和通信协议；
- 第2部分：算法模型仓库；
- 第3部分：端子系统；
- 第4部分：边子系统；
- 第5部分：云子系统；
- 第6部分：端边云协同；
- 第9部分：存储系统；
- 第11部分：安全与隐私保护。

本文件由新一代人工智能产业技术创新战略联盟AI标准工作组提出。

本文件由中关村视听产业技术创新联盟归口。

本文件起草单位：鹏城实验室、哈尔滨工业大学（深圳）、中国科学院计算技术研究所、浙江大华技术股份有限公司、中兴通讯股份有限公司、海信集团控股股份有限公司、北京大学、天翼视联科技股份有限公司、杭州海康威视数字技术股份有限公司、浙江视觉智能创新中心有限公司、中国科学院工业人工智能研究所、清华大学。

本文件主要起草人：王耀威、纪雯、白鑫贝、郑清芳、周运红、陈鹏、孔维生、敬太洋、刘海军、山其本、李潘、袁锦宇、温春江、杨哲铭、刘常昱、杨晓玲、贾霞、任文奇、金平、罗伟节、兰艳、李英建、熊雪菲、陈轲、栾晓旭、王智、郑栋宇、高文。

引 言

数字视网膜系统是面向海量视频数据高效处理而提出的新型端边云协同计算架构，采用“特征实时汇聚、视频按需调取、模型在线更新”的新应用范式，能够更好地支撑智慧安防、智能交通、智能制造、自动驾驶等领域的视频大数据分析处理应用。

数字视网膜系统通过端、边、云子系统之间实时交互的控制流和视频、特征、模型等数据流，构建高效的系统协同计算体系，完成高价值信息实时汇聚和分析挖掘，实现端、边、云深度协同。面向城市级应用中不同场景、不同规模的任务，数字视网膜系统端、边、云协同建立了统一的协同处理框架，能够基于统一的资源管理和调度机制，根据一定策略将任务分解、编排并动态调度至最优的端、边、云计算设备，从而提高视觉数据处理效率和系统资源利用率。

数字视网膜系统端、边、云协同能够支持分布式训练或推理、数据协同分析、资源优化调度等功能，从而满足城市级视频大数据实时感知、分析与挖掘需求。

T/AI 116《信息技术 数字视网膜系统》拟由十二个部分构成：

- 第1部分：系统结构和通信协议。目的在于确立数字视网膜系统的参考架构、功能要求、通信流程和数据接口等内容。
- 第2部分：算法模型仓库。目的在于确立算法模型的封装、管理、调度等功能与接口要求，实现对不同算力单元、不同软件框架、不同算法模型、不同计算目标等的可变支持。
- 第3部分：端子系统。目的在于确立数字视网膜端子系统的基本结构和技术要求。
- 第4部分：边子系统。目的在于确立数字视网膜边子系统的逻辑架构、技术要求和数据接口。
- 第5部分：云子系统。目的在于确立数字视网膜云子系统的参考架构、技术要求、数据接口和服务能力。
- 第6部分：端边云协同。目的在于确立数字视网膜端子系统、边子系统和云子系统之间协同工作的内容、机制和接口，为端、边、云子系统协同工作的实现提供参考准则。
- 第7部分：测试规范。目的在于确立数字视网膜系统中的算法模型仓库、端子系统、边子系统、云子系统、端边云协同等部分的测试内容和测试方法。
- 第8部分：系统总体度量及评价体系。目的在于确立数字视网膜系统在建设、验收和使用过程中的系统总体评价准则。
- 第9部分：存储系统。目的在于确立数字视网膜系统中存储系统设计与部署的基本要求。
- 第10部分：应用指南。目的在于确立数字视网膜系统在实际部署实施时的基本要求，为数字视网膜系统的典型行业应用提供参考方案。
- 第11部分：安全与隐私保护。目的在于确立数字视网膜系统在安全、隐私保护等方面的基本要求。
- 第12部分：嵌入表示的技术要求。目的在于确立数字视网膜系统中嵌入表示的总体框架以及功能、接口等方面的技术要求，为嵌入表示的生成和应用提供参考实现。

本文件的发布机构提请注意，声明符合本文件时，可能涉及到8.2.2、8.2.3、9.2、9.3、9.6与《一种视觉智能系统资源信息调度方法、装置、终端及介质》（专利号：202310290470.5）；7、8、9与《任务协同调度方法、装置、设备以及介质》（专利号：202410310666.0）；7.2、7.3、8.1、8.2、9.2与《软件定义的视频处理系统和处理方法》（专利号：201811234417.9）相关的专利的使用。

本文件的发布机构对于该专利的真实性、有效性和范围无任何立场。

该专利持有人已向本文件的发布机构承诺，他愿意同任何申请人在合理且无歧视的条款和条件下，就专利授权许可进行谈判。该专利持有人的声明已在本文件的发布机构备案。相关信息可以通过以下联

系方式获得：

联系人：黄铁军（新一代人工智能产业技术创新战略联盟秘书长）

通讯地址：北京大学理科2号楼2641室

邮政编码：100871

电子邮件：tjhuang@pku.edu.cn

网址：<http://www.aitisa.org.cn>

请注意除上述专利外，本文件的某些内容仍可能涉及专利。本文件的发布机构不承担识别专利的责任。

T/AI 116.6-2026

T/AI 176.6-2026

信息技术 数字视网膜系统 第6部分：端边云协同

1 范围

本文件确立了数字视网膜系统端边云协同架构、系统协同框架和协同处理模型，规定了协同功能要求、通信流程、接口等内容。

本文件适用于数字视网膜系统端、边、云协同的设计、开发与实现，以及设备与之相关的模块研发、生产。其他信息系统可参考采用。

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。

T/AI 116.1—2021 信息技术 数字视网膜系统 第1部分：系统结构和通信协议

3 术语和定义

T/AI 116.1—2021界定的以及下列术语和定义适用于本文件。

3.1

业务 business

与应用场景密切相关的、能够满足具体行业应用需求的事务或服务。

3.2

任务 task

为实现具体业务从技术层面进行解析得到的程序或工作。

[来源：GB/T 37684—2019, 3.3, 有修改]

3.3

子任务 subtask

根据任务内在逻辑关系拆分形成的较小任务单元。

[来源：GB/T 37684—2019, 3.4]

3.4

任务迁移 task migration

将某一个或多个设备中存在的任务转移到其他设备处理的操作。

3.5

协同处理 collaborative processing

数字视网膜系统中多个子系统互相配合，旨在共同完成某任务或提供某服务的一种信息处理过程。

[来源：GB/T 37684—2019, 3.1, 有修改]

3.6

任务发起方 task initiator

在协同工作中，负责生成任务并触发后续任务处理流程的模块或子系统。

3.7

协同调度方 collaborative scheduler

在协同工作中，承担任务规划、任务调度、协同管控等功能的模块或子系统。

3.8

协同执行方 collaborative executor

在协同工作中，承担任务执行、状态返回等功能的模块或子系统。

4 缩略语

下列缩略语适用于本文件。

CPU：中央处理器（Central Processing Unit）

FPS：帧率（Frames Per Second）

HTTP：超文本传输协议（Hypertext Transfer Protocol）

ID：身份标识（Identification）

JSON：对象标记语言（JavaScript Object Notation）

TFLOPS：每秒万亿浮点运算次数（Tera Floating-point Operations Per Second）

TOPS：每秒万亿操作次数（Tera Operations Per Second）

URI：统一资源标识符（Uniform Resource Identifier）

URL：统一资源定位符（Uniform Resource Locator）

5 协同架构

5.1 基础协同架构

数字视网膜系统主要由端子系统、边子系统和云子系统组成，根据参与协同工作的子系统类型，系统存在端-端、端-边、边-云、端-云、边-边、云-云等多种协同工作模式。基于对不同协同工作模式的统一抽象，图 1 给出了数字视网膜系统的基础协同架构，该架构从协同工作角色分工的角度定义了子系统协同工作关系的逻辑模型，而非子系统之间的网络拓扑结构。

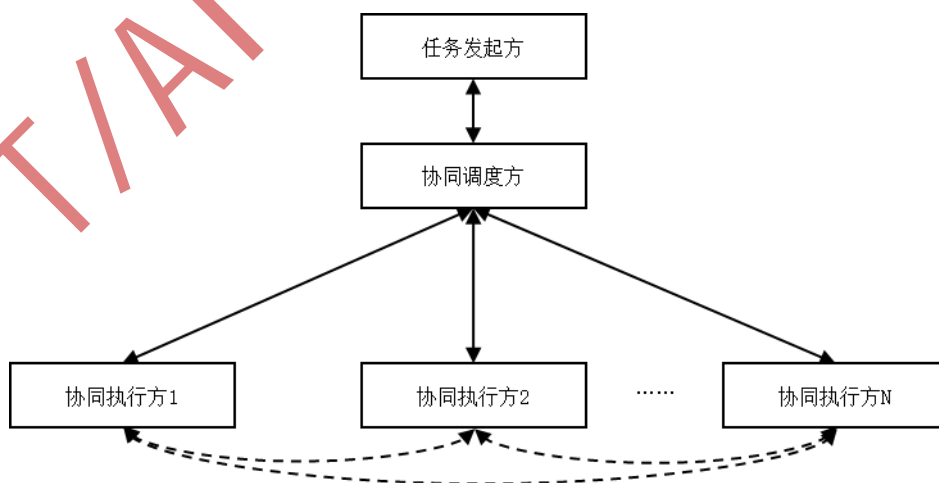


图 1 数字视网膜系统基础协同架构

图 1 描述了系统协同工作过程中，不同角色之间的协同交互关系。协同工作角色包括任务发起方、协同调度方和协同执行方，它们之间以传输网络为基础，实现控制信息和数据的传输与交换。任务发起

方与协同调度方之间进行任务请求、任务结果数据等的交互，协同调度方与协同执行方之间存在控制指令、数据、任务信息等的交互，不同协同执行方之间可存在数据交互。协同调度方接收任务发起方的任务请求，通过对协同执行方的设备、数据、资源、任务等进行协同调度与控制，最终实现任务协同处理。

任务发起方、协同调度方和协同执行方可由不同的子系统承担，也可由同一个子系统承担。相应地，子系统的角色宜根据实际应用需求、业务特点等因素确定，不同类型的子系统可承担相同的角色，同种类型的子系统可承担不同的角色。例如，在端-边-云系统形态中，端子系统作为任务发起方，边子系统作为协同执行方，云子系统作为协同调度方；或者，端子系统和边子系统均作为协同执行方，云子系统同时作为任务发起方和协同调度方。

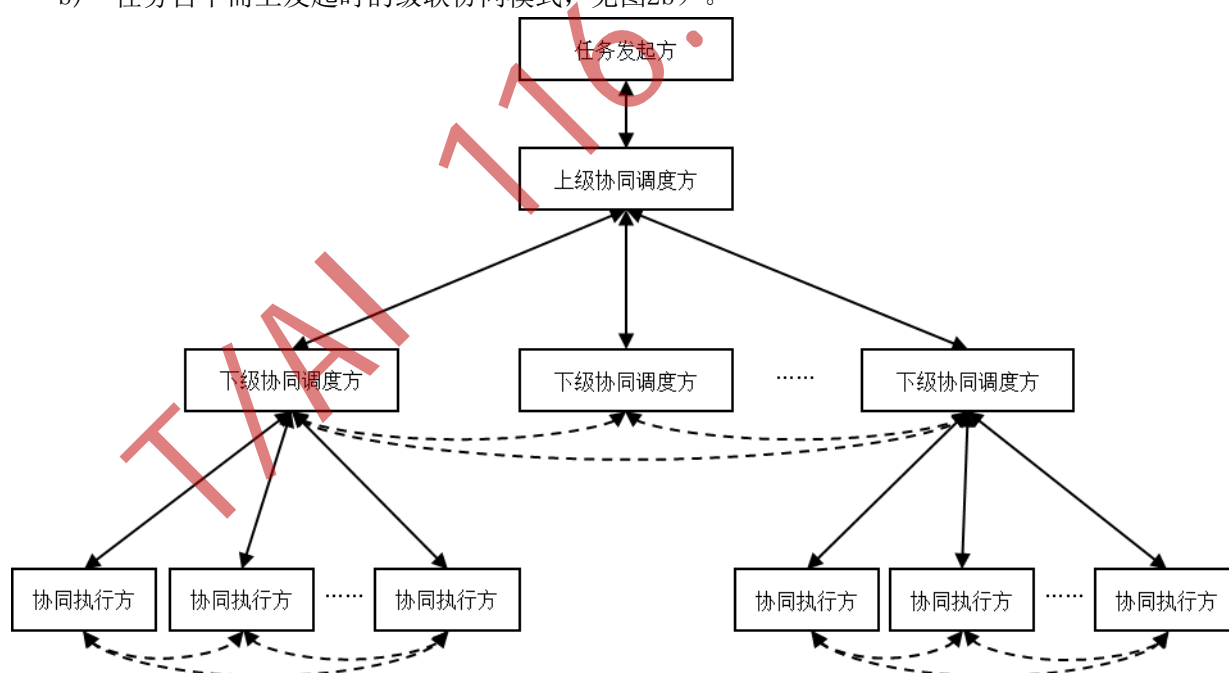
协同调度方的确定主要分为两类情况：对于同种类型子系统之间的协同（如端-端协同、边-边协同等），协同调度方可根据一定的策略协商确定，此外，协同调度方可根据业务内容的变化动态调整；对于不同类型子系统之间的协同，协同调度方通常由具备局部或全局协调管理能力和数据汇聚与处理能力的子系统承担，例如，在端-边协同中，协同调度方为边子系统，在边-云协同中，协同调度方为云子系统，在端-云协同中，协同调度方为云子系统。

5.2 级联协同架构

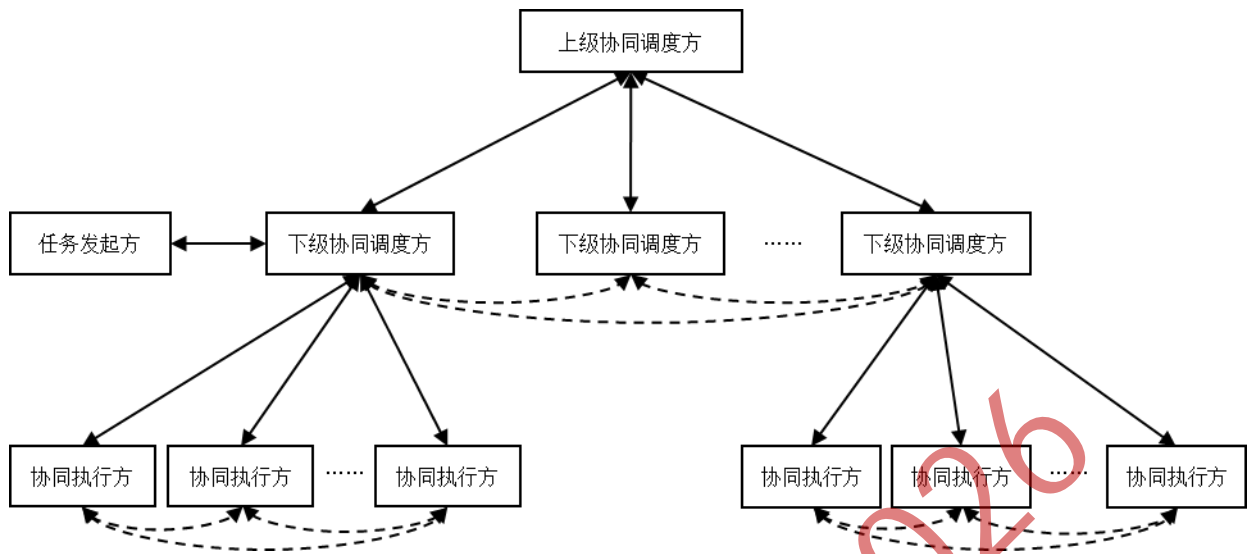
在数字视网膜系统的端-边-云整体架构下，实际业务中可能存在逐级协同调度的情况，基于基础协同架构，通过级联方式可构建数字视网膜系统级联协同架构。

根据任务发起方与不同层级协同调度方的交互关系，级联协同架构主要包括两种情况：

- a) 任务自上而下发起时的级联协同模式，见图2a)；
- b) 任务自下而上发起时的级联协同模式，见图2b)。



a) 自上而下任务发起方式



b) 自下而上任务发起方式

图 2 数字视网膜系统级联协同架构

图 2 给出了不同任务发起方式下的数字视网膜系统级联协同架构，该架构是在协同角色分工视角下采用一般化表示的多层结构，主要由任务发起方、下级协同调度方、上级协同调度方和协同执行方组成。对于自上而下的任务发起方式，任务发起方向上级协同调度方发送任务请求，然后由上级协同调度方向下逐级进行级联调度处理，对于自下而上的任务发起方式，任务发起方向下级协同调度方发送任务请求，任务经过上报和审核后由相关调度方进行级联调度处理。协同执行方、下级协同调度方与上级协同调度方之间通过网络逐级通信，通过各类数据和控制信息的交互，实现任务协同处理。

对于数字视网膜系统的大规模部署应用，此架构支持扩展参与协同工作的节点数量和地理范围，不同层级的协同调度方负责完成不同粒度的任务协同调度，通过级联协同，实现面向不同规模的端、边、云协同应用。

6 系统协同框架

数字视网膜系统协同框架见图 3。图中给出了数字视网膜系统端、边、云协同涉及的核心要素以及它们之间的关系。

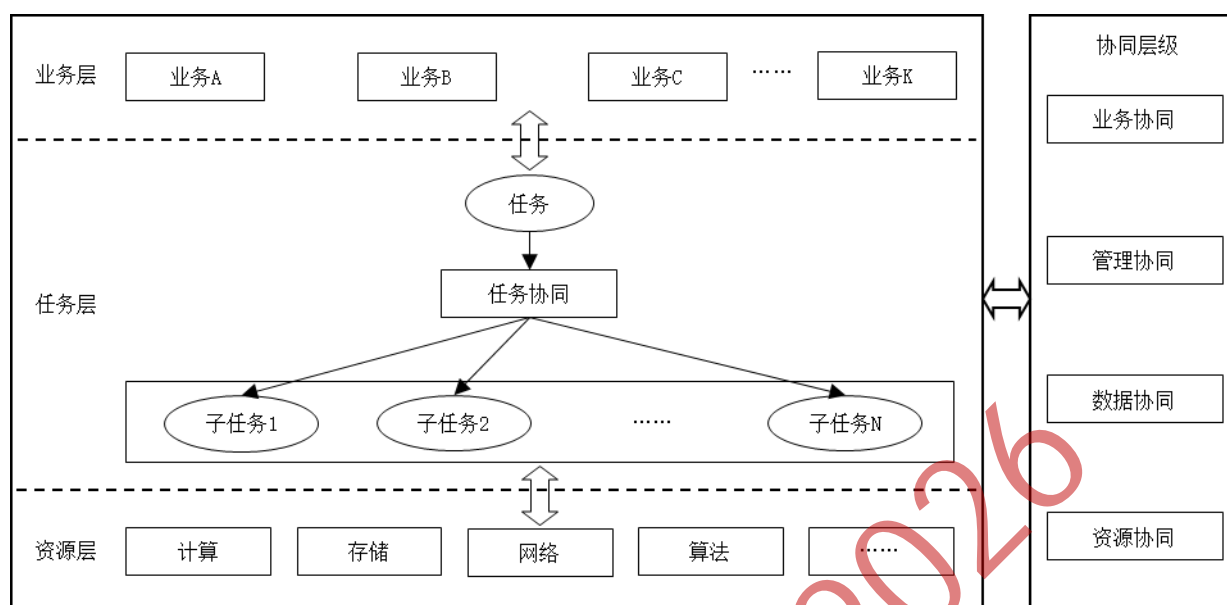


图3 数字视网膜系统协同框架

数字视网膜系统协同框架主要包括业务层、任务层和资源层。业务层面面向应用需求和业务逻辑进行抽象化的业务描述，任务层根据执行逻辑实现任务协同的规划与调度，资源层则提供可供协同调度的基础资源支撑。端、边、云之间协同工作的核心是以顶层业务为驱动，通过一定的任务协同机制，系统对内生成可执行的任务和子任务，选择和分配相适应的算法、计算、存储、网络等资源，通过视频流、特征流、模型流、控制流等的传输与协作，从而实现具体业务，满足实时处理、快速响应、数据联合分析等业务需求。

端、边、云协同工作过程依赖于四个层级的协同能力支撑，具体如下：

- 资源协同：对数字视网膜系统内端、边、云子系统的计算、存储、网络、算法、模型等资源进行统一监控、管理和调度，优化资源分配。
- 数据协同：通过视频、特征、结果等数据在端、边、云子系统之间的有序处理，实现多源数据协同分析、融合、挖掘等的过程。
- 管理协同：统一监控和管理端、边、云协同涉及的软硬件信息和协同工作情况，包括设备信息管理、算法模型运行环境管理、任务信息管理等内容。
- 业务协同：针对具体应用，面向用户称之为业务，面向系统内部实现则对应一项或多项任务。根据业务需求内容，将对应的任务进行解析、分解和编排，选择适当的任务调度策略，确定任务协同调度方案，控制任务执行过程，最终在一个或多个子系统上完成业务。

总体来说，端、边、云之间的协同工作是以资源优化利用和数据实时处理为目标，通过子系统的基本功能和四个层面的基本协同能力，高效支撑实际业务的实现。

7 协同处理模型

7.1 参考模型

数字视网膜系统任务协同处理的参考模型见图4。

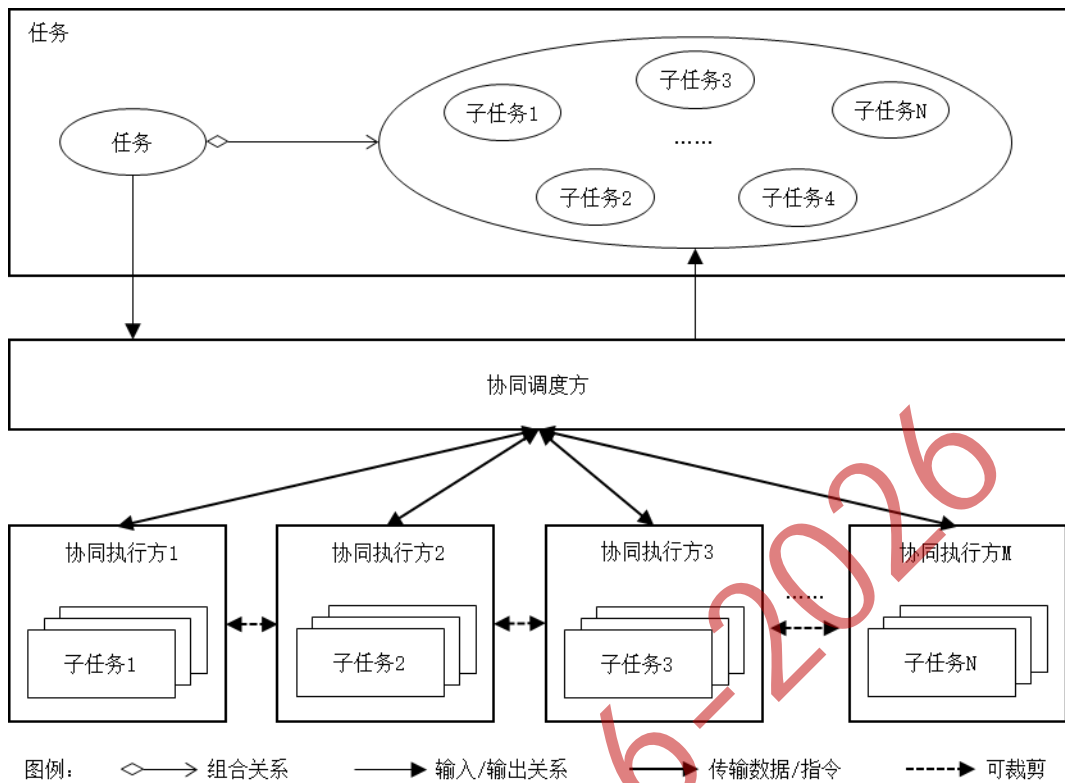


图4 任务协同处理参考模型

按照端边云协同的角色视图和系统协同框架，图4给出了任务、子任务和协同调度方、协同执行方之间的关系。

对于一项任务，可以由协同调度方分解为一项或多项子任务，并将子任务编排起来形成一定的逻辑关系，确定各子任务执行的时序、触发条件、输入输出关系等，此外，可在任务执行过程中进行动态编排，某些情况下，后续子任务的内容和逻辑关系取决于前一个子任务的输出结果。

在端边云协同应用中，协同调度方和协同执行方相互配合完成任务协同处理。协同调度方承担任务导入、任务编排、任务协同调度与管控等功能，协同执行方负责任务的执行，协同调度方、协同执行方之间能够进行各类数据和控制指令的交互，各协同执行方之间也能够通过数据交互完成子任务协同处理，将最终处理结果发送至协同调度方。

任务分解得到的子任务可在一至多个协同执行方上按照一定逻辑关系来协同执行，例如，多个协同执行方串行处理不同子任务，或多个协同执行方并行处理同一类型的子任务、它们的处理结果输入至另一子任务进一步处理输出任务的最终结果。对于动态编排产生的子任务，需对执行该子任务的协同执行方进行动态分配和控制。

7.2 功能模型

协同处理功能模型见图5。图中给出了任务协同处理涉及的协同调度方和协调执行方的功能模块和交互关系。

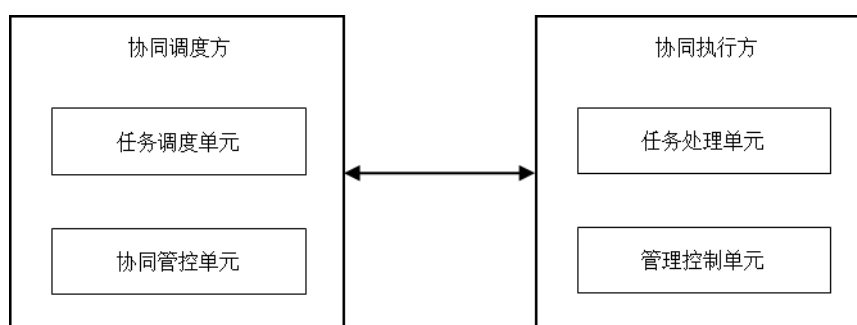


图5 协同处理功能模型

协同调度方包括任务调度单元和协同管控单元，协同执行方包括任务处理单元和管理控制单元。

协同调度方的模块功能描述如下：

任务调度单元：对任务进行分解和编排，根据一定策略确定任务的执行者，即某一个或多个协同执行方，将任务派发至相应的协同执行方，并控制任务执行状态，实现任务的协同调度与运行。具体包括任务分解与编排、选择任务调度策略、选择执行任务的协同执行方、选择算法模型、确定节点联动控制方案等功能。

协同管控单元：监测协同执行方的资源情况和软硬件运行状态，管理协同工作过程中的设备、任务和数据，与协同执行方进行数据与指令交互，实现协同业务流程的组织与控制。

协同执行方的模块功能描述如下：

任务处理单元：利用本地的算力、内存等资源和设备软硬件能力，完成任务的执行。

管理控制单元：对本地资源情况、任务情况和设备状态等信息进行监测，负责与协同调度方交互，上传上述信息和任务执行结果，接收并执行控制指令，完成任务控制、任务管理、设备控制、算法模型更新等功能。

端、边、云子系统协同工作时，各子系统依据承担的角色，应具备相应模块的功能。

7.3 协同处理过程

协同处理过程是一系列对数据的处理转化过程和流向控制过程。基于协同处理功能模型的协同处理过程描述如下：

- a) 任务导入。分析业务需求，获取任务的描述信息。
- b) 任务协同规划，确定任务协同调度方案。基于任务的描述信息对任务进行拆分编排，得到任务逻辑拓扑关系和子任务集合，根据子任务类型、可用算法以及协同执行方的时空信息、硬件平台、软件环境、资源情况等信息，逐步确定执行每个子任务的协同执行方，并生成子任务调度信息。其中，可根据任务约束对任务协同调度方案进行迭代优化。
- c) 任务分配。将子任务调度信息发送至对应的协同执行方上。
- d) 协同处理。各协同执行方按照任务逻辑拓扑关系中描述的处理逻辑和执行条件进行子任务的处理和交互，交互内容例如视频、特征、中间结果等数据，最终完成整个任务。协同处理过程中还需监测和控制任务及其子任务的状态，对发生的异常情况进行处理。

8 协同功能要求

8.1 总体要求

8.1.1 任务描述

任务描述是开展任务协同工作的前提，为任务编排、调度和管理提供支撑，应包括下列信息。

- a) 任务类型描述，支持描述的任务类型包括但不限于：
 - 1) 视频抽帧任务；
 - 2) 视频/图像预处理任务，例如降噪、去雾等；
 - 3) 视频/图像编码、解码、转码任务；
 - 4) 特征编码、解码任务；
 - 5) 模型压缩任务；
 - 6) 传统算法分析处理任务；
 - 7) AI模型训练任务；
 - 8) AI模型推理任务；
 - 9) 特征检索任务；
 - 10) 数据合并、关联、融合等数据分析或聚合任务。
- b) 任务约束信息描述，例如：
 - 1) 任务优先级；
 - 2) 时延要求；
 - 3) 执行时段；
 - 4) 空间范围；
 - 5) 性能要求；
 - 6) 依赖关系。
- c) 环境需求描述，包括任务执行所需的软硬件环境信息等。
- d) 资源需求描述，例如：
 - 1) 任务所需的编/解码资源；
 - 2) 任务消耗的计算资源；
 - 3) 任务数据占用的存储资源；
 - 4) 任务执行消耗的缓存资源；
 - 5) 任务派发及结果上报各自所占用的带宽资源；
 - 6) 任务使用的算法资源。
- e) 任务输入信息，例如，数据源信息、算法分析参数等。
- f) 任务创建信息、状态信息等内容。

8.1.2 基本要求

任务协同处理是指将一项具体任务划分为具有一定逻辑拓扑关系的若干子任务，根据端、边、云资源情况，将子任务派发到合适的设备上执行，并按照任务编排的结果控制子任务执行时序和数据流向，最终完成任务的过程。

任务协同处理涉及任务并发与同步、任务间通信等内容，符合下列要求：

- a) 应支持对多种类型的子任务进行协同处理，例如：视频抽帧任务、视频/图像预处理任务、视频/特征编解码任务、AI推理计算任务、特征检索任务、数据分析/聚合任务等；
- b) 应支持多个任务并发处理能力；
- c) 应支持分布式同步机制，例如，基于一致性协议（如Raft、Paxos）构建的etcd、ZooKeeper等分布式锁；
- d) 应支持任务发起方、协同调度方和协同执行方之间的时间同步和定期校正；
- e) 宜支持任务发起方、协同调度方、协同执行方与外部时钟源之间的时间同步和定期校正；

- f) 应支持通过标记和检查时间信息、任务标识以及任务依赖关系等信息，实现数据按预定的时序和逻辑进行处理；
- g) 任务协同处理中使用的通信机制应具备跨平台兼容性；
- h) 应具备对任务协同处理过程中出现的网络不稳定、资源冲突、任务中断等异常情况的处理能力。

8.1.3 任务协同机制

为协同实施一项任务，协同调度方需提供一定的任务协同机制支撑，符合下列要求：

- a) 应支持以手动或自动的方式进行任务分解与编排；
- b) 应支持按照多种逻辑关系对任务进行分解与编排，例如，串行、并行、条件执行等；
- c) 应支持多种任务调度策略，例如，任务优先级策略、轮询、事件触发等；
- d) 应支持多种资源调度策略，例如，数据就近处理策略、负载均衡策略等；
- e) 宜支持自定义任务调度策略；
- f) 宜支持动态任务编排。

8.1.4 任务调度管理

协同调度方负责对任务进行总体调度、监测和控制，符合下列要求：

- a) 应支持统一记录和维护任务逻辑拓扑关系；
- b) 应支持选择任务所需的算法和模型；
- c) 应支持根据任务执行需求确定执行任务的协同执行方，例如，资源需求、运行环境需求等；
- d) 应支持任务分发、启动、停止、删除等功能；
- e) 应支持监测和管理任务相关的信息，包括任务调度信息、任务状态信息等；
- f) 应支持获取任务执行结果，用于存储、转发或可视化；
- g) 应支持任务迁移，即当负责执行某个任务的协同执行方因发生故障、容量调整、资源受限等情况导致任务无法正常执行时，将任务重新调度到其他具有执行能力的协同执行方上；
- h) 应支持监测和处理任务协同处理过程中的异常情况。

8.2 基本功能要求

8.2.1 任务发起方

任务发起方的基本协同功能要求如下：

- a) 应支持向协同调度方发起任务请求；
- b) 应支持接收任务执行状态和结果数据。

8.2.2 协同调度方

协同调度方的基本协同功能要求如下。

- a) 资源协同
 - 1) 应支持获取协同执行方的设备资源信息；
 - 2) 应支持对所有协同执行方的设备资源进行统一管理、分配和回收；
 - 3) 应支持获取协同执行方的算法资源信息；
 - 4) 宜支持通过池化方式管理和分配异构算力资源；
 - 5) 宜支持在多个协同执行方之间共享数据、服务等资源。
- b) 数据协同
 - 1) 应支持确定和配置任务协同处理过程中的数据流向关系；

- 2) 应支持接收协同执行方输出的视频、特征、算法结果等任务结果数据;
 - 3) 宜支持转发协同执行方输出的任务结果数据。
- c) 管理协同
- 1) 应支持(批量)添加、删除、启用、停用协同执行方;
 - 2) 应支持监测和管理协同执行方的数量和设备信息,例如,在线状态、工作状态、时空信息、异常信息等;
 - 3) 应支持对算法模型运行环境的管理、隔离和配置;
 - 4) 应支持数据或控制指令的收发控制;
 - 5) 宜支持主动感知周边设备的能力;
 - 6) 宜支持动态协商选择协同调度方。
- d) 业务协同
- 1) 应支持任务的导入和创建;
 - 2) 应支持在多个协同执行方中批量部署、更新任务;
 - 3) 应支持配置任务优先级策略,例如,静态优先级、动态优先级等;
 - 4) 宜支持配置任务的触发方式,例如,可将任务配置为定时启动并设定启动时间;
 - 5) 宜支持动态调整并发任务的数量;
 - 6) 应支持生成和下发对协同执行方的控制指令,实现多个协同执行方之间的联动;
 - 7) 应支持调度多个协同执行方按照任务逻辑关系协同完成一项业务;
 - 8) 宜支持调度同一协同执行方中的多个子任务协同完成一项业务;
 - 9) 应支持保存和备份任务信息,任务信息格式见附录A;
 - 10) 应支持对调度失败或执行失败的任务进行重新调度,重试策略可包括设定最大重试次数、重试间隔等;
 - 11) 宜支持对执行失败且无法重试的任务进行回滚,以保持任务状态的一致性;
 - 12) 宜支持设置检查点,当任务执行过程发生故障时从检查点恢复,避免重新执行整个任务。

8.2.3 协同执行方

协同执行方的基本协同功能要求如下。

- a) 资源协同
- 1) 应支持监控自身的设备资源信息,包括总量资源和可用资源;
 - 2) 应支持设备资源信息的查询和上报;
 - 3) 应支持获取和发送自身的算法资源信息;
 - 4) 宜支持通过池化方式管理和分配异构算力资源。
- b) 数据协同
- 1) 应支持将采集或处理得到的数据发送至其他协同执行方,包括视频、图片、特征、中间结果等数据;
 - 2) 应支持至少一种以下数据协同处理功能:对其他协同执行方采集或处理得到的视频、图片、特征或中间结果等数据进行推理计算、二次处理等操作;对其他协同执行方输出的特征数据进行汇聚和检索;对不同算法和模型输出的特征数据进行转换处理以获得兼容特征;对存在一定联系的视频、特征、算法结果等数据进行关联;对其他协同执行方输出的算法结果进行校验、汇总、排序等数据聚合处理;对来自多个协同执行方的数据进行联合分析;或对多模态数据进行关联、融合或挖掘;
 - 3) 应支持将任务结果上报至协同调度方;
 - 4) 应支持对发送失败的数据进行缓存并重新发送。

- c) 管理协同
 - 1) 应支持向协同调度方的登录和注销等操作；
 - 2) 应支持向协同调度方输出设备信息，例如，工作状态、时空信息、异常信息等；
 - 3) 应支持监测任务执行所占用的资源情况；
 - 4) 应支持上报自身承担任务的状态信息；
 - 5) 当发生软件故障时应支持自动重启或切换到备用实例；
 - 6) 宜支持动态协商选择协同调度方。
- d) 业务协同
 - 1) 应支持接收和执行协同调度方下发的任务；
 - 2) 应支持任务可执行性检查；
 - 3) 应支持根据协同调度方的指令对自身承担的任务进行启动、停止等控制操作；
 - 4) 应支持按照协同调度方确定的任务处理逻辑与其他协同执行方协同完成一项业务；
 - 5) 宜支持通过自身的多个子任务协同完成一项业务；
 - 6) 宜支持保存自身的任务信息。

8.3 应用能力要求

8.3.1 概述

数字视网膜系统主要由端子系统、边子系统、云子系统三部分组成。实际中，面向不同的应用场景和业务类型，往往需要多个相同或不同类型的子系统之间协同配合。根据参与协同工作的子系统类型，端、边、云协同的工作模式可分为端-边协同、端-端协同、边-云协同、边-边协同、端-云协同、云-云协同和端-边-云协同。在基本协同功能的支撑下，每种协同工作模式能够支持一定的协同应用能力。

8.3.2 端-端协同

数字视网膜系统宜具备端-端协同能力，宜符合下列要求：

- a) 端子系统具有主动感知周边资源的能力；
- b) 支持多个端子系统协商确定协同调度方，且协同调度方可动态调整；
- c) 多个端子系统之间能够共享视频、特征等数据；
- d) 支持任务在多个端子系统之间进行调度和迁移；
- e) 支持多个端子系统形成自组织网络动态协作完成任务。

8.3.3 端-边协同

数字视网膜系统应具备端-边协同能力，具体要求如下：

- a) 应具备支持以下一种或多种协同方式的能力：
 - 1) 端侧采集数据、边缘侧完成数据处理与分析；
 - 2) 端侧采集数据并完成特征提取、边缘侧完成后续分析、识别等功能；
- b) 应支持端、边之间通过串行处理的方式协同完成分析任务，端侧执行算力要求低或实时性要求高的前置计算子任务，边缘侧执行计算复杂度高的后续计算子任务；
- c) 应支持多个端子系统与边子系统协同完成跨终端数据关联分析或区域优化调控任务，例如，多个端子系统并行执行目标跟踪子任务，边子系统执行数据融合分析子任务实现目标连续追踪；
- d) 宜支持基于模型分割的端-边协同推理，根据资源情况在端侧进行一部分模型的推理计算，然后在边缘侧实现剩余部分模型的推理计算；

- e) 宜支持多个端子系统在边子系统的协同控制下实现联动,完成复杂场景任务,例如,当一个端子系统发现目标时,能够调整拍摄角度、焦距等参数,完成对目标信息的精细提取,同时对目标进行持续跟踪,当目标即将超出该端子系统的监控范围时,由边子系统控制邻近端子系统进行联动布防,识别并锁定目标动向;
- f) 宜支持在边子系统的管理调度下将一个端子系统上的计算任务迁移到另一个端子系统上执行。

8.3.4 端-云协同

数字视网膜系统应具备端-云协同能力,具体要求如下:

- a) 应具备支持以下一种或多种协同方式的能力:
 - 1) 端侧采集数据、云侧完成数据处理与分析;
 - 2) 端侧采集数据并完成特征提取/智能化分析,云侧完成数据检索、协同分析或数据挖掘;
- b) 应支持云侧训练-端侧推理的协同方式;
- c) 宜支持端-云协同推理;
- d) 宜支持云侧训练-端侧增量训练-端侧推理的协同方式;
- e) 宜支持云侧模型和端侧模型的协同进化,基于云侧模型进行压缩转换后生成适配端侧部署应用的模型,基于端侧反馈的数据或模型辅助云侧模型更新优化;
- f) 宜支持通过端、云协同实现强化学习、增量学习等机器学习范式。

8.3.5 边-边协同

数字视网膜系统宜具备边-边协同能力,宜符合下列要求:

- a) 支持边-边协同推理,根据资源情况将模型切分并分配到多个边子系统上,每个边子系统执行一部分模型的推理计算工作;
- b) 支持边-边协同训练,每个边子系统承担模型的部分训练任务,通过整合中间结果(如参数聚合、梯度交换等)完成整体的模型训练;
- c) 支持边-边联邦学习,以保障数据安全和保护隐私信息;
- d) 支持任务在多个边子系统之间进行调度和迁移。

8.3.6 边-云协同

数字视网膜系统应具备边-云协同能力,具体要求如下:

- a) 应支持以容器化的方式在边、云之间灵活地部署任务;
- b) 应支持云侧训练、边侧推理的协同方式;
- c) 应支持边侧处理、云侧分析的协同方式,边子系统通过特征提取、智能化分析、数据处理等算法输出与云侧业务关联的特征、分析结果等关键数据,不同边子系统的数在云侧汇聚,云子系统对大数据进行分析和价值挖掘,实现跨区域、跨系统的多维时空数据融合及协同分析;
- d) 应支持数据边侧存储、云侧按需调取,当云、边网络发生抖动或者故障时,能够通过数据缓存、自动续传、消息校验等手段,保障数据传输的正确性和完整性;
- e) 应支持边-云协同训练,例如,增量学习、联邦学习等;
- f) 宜支持边-云协同推理,例如,深度学习模型经切分后分别部署到边、云上进行协同推理;
- g) 宜支持边、云子系统算法模型仓库之间进行算法模型的双向传输;
- h) 宜支持云侧模型和边侧模型的协同进化;
- i) 宜支持综合利用多个云子系统的业务能力生成边侧应用。

8.3.7 云-云协同

数字视网膜系统宜具备云-云协同能力，宜符合下列要求：

- a) 支持云-云协同训练，能根据数据、模型的规模或复杂度调整训练策略，优化训练效率；
- b) 支持针对海量任务的云-云协同推理，能根据负载情况完成多个云子系统之间的任务迁移；
- c) 支持云-云跨域业务协同，实现不同地域相似业务的数据共享，为用户提供统一的服务。

8.3.8 端-边-云协同

数字视网膜系统应具备端-边-云协同能力，应至少符合以下要求中的一项：

- a) 支持端、边、云协同完成数据处理分析，端侧实现数据采集，边缘侧实现智能化分析，云侧实现数据分析与挖掘；
- b) 支持端、边、云协同的训练与推理，端侧负责数据采集、预处理和特征提取，边缘侧负责完成具体任务（例如识别、分割等），云侧负责模型训练并将算法和模型下发到边缘侧和端侧。

9 通信流程和接口

9.1 通用要求

在端、边、云协同工作过程中，协同调度方应支持获取协同执行方的设备状态信息、设备资源信息和设备算法信息，对任务进行调度与控制，管理任务信息，以及获取任务执行结果、任务运行状态等功能。端、边、云协同接口的通信协议要求应符合T/AI 116.1—2021中9.1的规定，接口功能对照关系见表1。对于基础协同架构和级联协同架构，端、边、云协同工作时的完整交互过程见附录B。

表1 端、边、云协同接口功能对照关系

序号	功能描述	通信协议要求	接口对应章节
1	查询设备资源信息	应符合T/AI 116.1—2021中9.1.3的规定	9.2
2	上报设备资源信息	应符合T/AI 116.1—2021中9.1.2的规定	9.3
3	查询协同执行方已有的算法信息	应符合T/AI 116.1—2021中9.1.3的规定	9.4
4	发送任务请求	应符合T/AI 116.1—2021中9.1.2的规定	9.5
5	将任务调度到相应的协同执行方上	应符合T/AI 116.1—2021中9.1.3的规定	9.6
6	对任务的启动、停止、运行参数等进行控制	应符合T/AI 116.1—2021中9.1.3的规定	9.7
7	发送任务执行结果	应符合T/AI 116.1—2021中9.1.3的规定	9.8
8	查询任务运行状态	应符合T/AI 116.1—2021中9.1.3的规定	9.9
9	上报任务运行状态	应符合T/AI 116.1—2021中9.1.2的规定	9.10

9.2 设备资源查询接口

9.2.1 通信流程

设备资源查询接口分为设备总量资源查询接口和设备可用资源查询接口。当查询设备总量资源时，查询指令接收方回复系统信息、通用资源总量信息、计算加速资源总量信息和视频处理资源总量信息等资源信息。当查询设备可用资源时，查询指令接收方回复实时的可用通用资源、可用计算加速资源和可用视频处理资源等信息。

设备资源查询接口的通信流程见图6。查询发送方为协同调度方，查询接收方为协同执行方。该接口支持协同调度方单独或批量查询协同执行方的设备资源信息，实际中依据子系统之间的通信连接关系而定，例如，云子系统向边子系统发送设备资源查询指令时，可仅查询边子系统的设备资源信息，或者同时查询边子系统和接入该边子系统的端子系统的设备资源信息，对于后者，边子系统需根据查询条件将设备资源查询指令分发给端子系统，并将若干端子系统的设备资源信息组合后返回给云子系统。

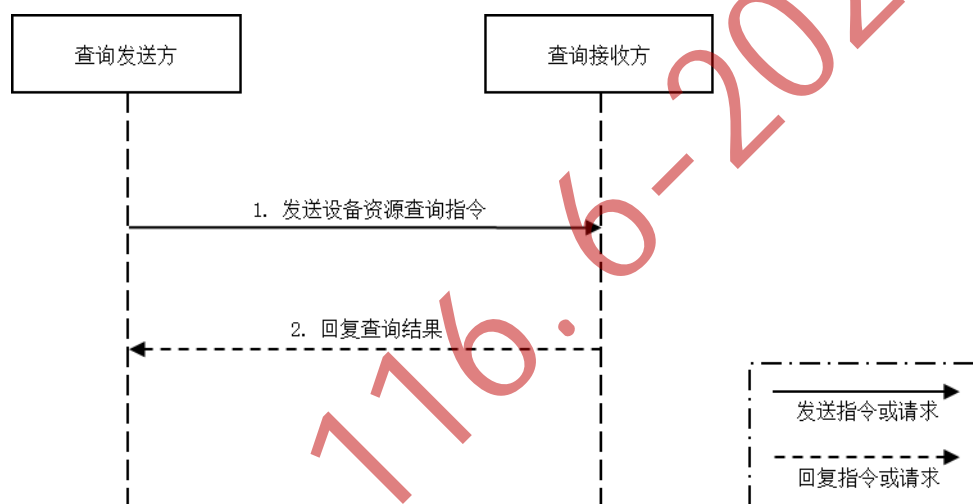


图6 设备资源查询接口通信流程

9.2.2 接口

9.2.2.1 设备总量资源查询接口

设备总量资源查询接口见表2。

表2 设备总量资源查询接口

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“total_res_query”
指令序号	CommandIdx	int	可选	4 字节	
设备 ID	DeviceID	string	可选	20 字节	接收指令的设备的标识
指令数据	MessageBody	JSON 数据包	必选		以下查询条件应至少填写一项
查询设备列表	+DeviceList	JSON array	可选		支持单独查询和批量查询
查询设备 ID	++DeviceID	string	必选	20 字节	以设备 ID 作为条件进行查询
是否逐级查询	++HierarchicalQuery	string	必选	1 字节	0: 不逐级查询, 1: 逐级查询

表 2 (续)

名称	标识符	类型	选择状态	长度	备注
查询经度最大值	+LongitudeMax	string	可选	小于 16 字节	以经度范围作为条件进行查询, 经度格式为 n10,6; 精确到小数点后 6 位
查询经度最小值	+LongitudeMin	string	可选	小于 16 字节	以经度范围作为条件进行查询, 经度格式为 n10,6; 精确到小数点后 6 位
查询纬度最大值	+LatitudeMax	string	可选	小于 16 字节	以纬度范围作为条件进行查询, 纬度格式为 n10,6; 精确到小数点后 6 位
查询纬度最小值	+LatitudeMin	string	可选	小于 16 字节	以纬度范围作为条件进行查询, 纬度格式为 n10,6; 精确到小数点后 6 位
产品类型	+ProductType	string	可选	小于或等于 64 字节	以产品类型作为条件进行查询
计算平台标识	+Chip	string	可选	小于或等于 64 字节	以计算加速平台类型或型号作为条件进行查询
驱动版本	+DriverVersion	string	可选	小于或等于 64 字节	以计算加速平台驱动版本作为条件进行查询
扩展查询条件	+ExtendCond	JSON object	可选		厂商根据需要可自定义扩展其他查询条件

注1: “+”表示JSON数据包的嵌套关系, 每增加一个“+”表示更深一层。

注2: 对于可选字段, 实际中通过检查该字段是否存在来判断接口实现中是否有携带该参数。

设备总量资源查询接口返回值见表3。

表 3 设备总量资源查询接口返回值

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“total_res_query”
指令序号	CommandIdx	int	可选	4 字节	下发的指令序号
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败
资源信息列表	TotalResourceList	JSON 数据包	必选		<TotalResource>对象数组, 详细信息见表 4

设备总量资源查询接口返回消息体<TotalResource>参数见表4。

表 4 设备总量资源查询接口返回消息体<TotalResource>参数

名称	标识符	类型	选择状态	长度	备注
设备 ID	DeviceID	string	必选	20 字节	
设备类型	DeviceType	string	必选	小于 32 字节	0: 终端设备 1: 边缘设备 2: 云设备
产品类型	ProductType	string	可选	小于或等于 64 字节	
返回查询信息	ResourceInfo	JSON object	必选		
信息有效性	+IsValid	string	必选	1 字节	Y: 有效; N: 无效
系统运行环境信息	+SystemEnv	JSON object	可选		
运行环境标识	++EnvID	string	必选	小于或等于 24 字节	
运行环境版本	++EnvVersion	string	必选	小于或等于 64 字节	
通用资源信息	+CommonResource	JSON object	可选		见表 5
计算资源信息	+ComputeResource	JSON object	可选		见表 6
视频处理能力信息	+VideoProcessResource	JSON array	可选		支持的视频格式及编解码能力信息 <VideoProcessResourceObject> 对象数组, 详细信息见表 7
功耗大小	+PowerConsumption	string	可选	小于或等于 64 字节	
扩展资源	+ExtendInfo	JSON object	可选		厂商根据需要自定义返回其他资源信息

表 5 通用资源信息<CommonResource>属性

名称	标识符	类型	选择状态	长度	备注
总通用内存大小	TotalMemory	long	必选	8 字节	单位 MB
总存储大小	TotalStorage	long	必选	8 字节	单位 MB
总网络带宽	TotalBandwidth	int	必选	4 字节	单位 Mb/s

表 6 计算资源信息<ComputeResource>属性

名称	标识符	类型	选择状态	长度	备注
计算平台标识	Chip	string	必选	小于或等于 64 字节	
驱动版本	DriverVersion	string	可选	小于或等于 64 字节	
总算力大小	TotalPower	JSON array	必选		
计算精度	+Precision	string	必选		例如, int4、int8、float16、float32 等
算力值	+ComputingPower	string	必选		例如, 32 TOPS, 16 TFLOPS
总智能内存大小	TotalComputeMemory	long	可选	8 字节	单位 MB

表 7 视频处理能力信息<VideoProcessResourceObject>属性

名称	标识符	类型	选择状态	长度	备注
视频编码格式	FormatName	string	必选	小于 10 字节	支持的视频格式名称，如 H.264、H.265、AVS2、AVS3 等 对于支持的视频格式，其解码能力和编码能力应至少填写一项
解码能力	Decoding	JSON object	可选		
最大解码分辨率	+MaxResolution	string	必选	小于或等于 32 字节	支持的最大的解码分辨率，如 1920*1080
最大解码能力	+MaxPerf	float	必选	4 字节	使用统一基准量化后的解码能力，例如，以 1080P@1FPS 的解码能力为基准
编码能力	Encoding	JSON object	可选		
最大编码分辨率	+MaxResolution	string	必选	小于或等于 32 字节	支持的最大的编码分辨率，如 1920*1080
最大编码能力	+MaxPerf	float	必选	4 字节	使用统一基准量化后的编码能力，例如，以 1080P@1FPS 的编码能力为基准

9.2.2.2 设备可用资源查询接口

设备可用资源查询接口见表8。

表 8 设备可用资源查询接口

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“available_res_query”
指令序号	CommandIdx	int	可选	4 字节	
设备 ID	DeviceID	string	可选	20 字节	接收指令的设备的标识
指令数据	MessageBody	JSON 数据包	必选		以下查询条件应至少填写一项
查询设备列表	+DeviceList	JSON array	可选		支持单独查询和批量查询
查询设备 ID	++DeviceID	string	必选	20 字节	以设备 ID 作为条件进行查询
是否逐级查询	++HierarchicalQuery	string	必选	1 字节	0: 不逐级查询, 1: 逐级查询
查询经度最大值	+LongitudeMax	string	可选	小于 16 字节	以经度范围作为条件进行查询, 经度格式为 n10,6; 精确到小数点后 6 位
查询经度最小值	+LongitudeMin	string	可选	小于 16 字节	以经度范围作为条件进行查询, 经度格式为 n10,6; 精确到小数点后 6 位

表 8 (续)

名称	标识符	类型	选择状态	长度	备注
查询纬度最大值	+LatitudeMax	string	可选	小于 16 字节	以纬度范围作为条件进行查询, 纬度格式为 n10,6; 精确到小数点后 6 位
查询纬度最小值	+LatitudeMin	string	可选	小于 16 字节	以纬度范围作为条件进行查询, 纬度格式为 n10,6; 精确到小数点后 6 位
产品类型	+ProductType	string	可选	小于或等于 64 字节	以产品类型作为条件进行查询
计算平台标识	+Chip	string	可选	小于或等于 64 字节	以计算加速平台类型或型号作为条件进行查询
驱动版本	+DriverVersion	string	可选	小于或等于 64 字节	以计算加速平台驱动版本作为条件进行查询
扩展查询条件	+ExtendCond	JSON object	可选		厂商根据需要可自定义扩展其他查询条件

设备可用资源查询接口返回值见表9。

表 9 设备可用资源查询接口返回值

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“available_res_query”
指令序号	CommandIdx	int	可选	4 字节	下发的指令序号
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败
资源信息列表	AvailableResourceList	JSON 数据包	必选		<AvailableResource>对象数组, 详细信息见表 10

设备可用资源查询接口返回消息体<AvailableResource>参数见表10。

表 10 设备可用资源查询接口返回消息体<AvailableResource>参数

名称	标识符	类型	选择状态	长度	备注
设备 ID	DeviceID	string	必选	20 字节	
返回查询信息	ResourceInfo	JSON object	必选		
信息有效性	+IsValid	string	必选	1 字节	Y: 有效; N: 无效
可用通用资源信息	+AvailCommonResource	JSON object	可选		
可用通用内存大小	++AvailableMemory	long	可选	8 字节	单位 MB
可用存储大小	++AvailableStorage	long	可选	8 字节	单位 MB
可用网络带宽	++AvailableBandwidth	int	可选	4 字节	单位 Mb/s
可用 CPU 百分比	++AvailableCPU	int	可选	4 字节	[0, 100]

表 10 (续)

名称	标识符	类型	选择状态	长度	备注
可用计算资源信息	+AvailComputeResource	JSON object	可选		
可用算力大小	++AvailablePower	JSON array	可选		
计算精度	+++Precision	string	必选		例如, int4、int8、float16、float32 等
算力值	+++ComputingPower	string	必选		例如, 32 TOPS, 16 TFLOPS
可用智能内存大小	++AvailableComputeMemory	long	可选	8 字节	单位 MB
可用视频处理能力信息	+AvailVideoProcessResource	JSON array	可选		支持的视频格式及可用编解码能力信息
视频编码格式	++FormatName	string	必选	小于 10 字节	支持的视频格式名称, 如 H.264、H.265、AVS2、AVS3 等 对于支持的视频格式, 其可用解码能力和可用编码能力应至少填写一项
可用解码能力	++AvailDecodePerf	float	可选	4 字节	使用统一基准量化后的解码能力, 例如, 以 1080P@1FPS 的解码能力为基准
可用编码能力	++AvailEncodePerf	float	可选	4 字节	使用统一基准量化后的编码能力, 例如, 以 1080P@1FPS 的编码能力为基准
扩展资源	+ExtendInfo	JSON object	可选		厂商根据需要自定义返回其他资源信息

9.3 设备资源上报接口

9.3.1 通信流程

设备资源上报接口分为设备总量资源上报接口和设备可用资源上报接口。设备资源上报方主动向接收方上报实时可用通用资源、可用算力资源和可用视频处理资源等信息。设备资源上报接口的通信流程见图7。上报方为协同执行方, 接收方为协同调度方。

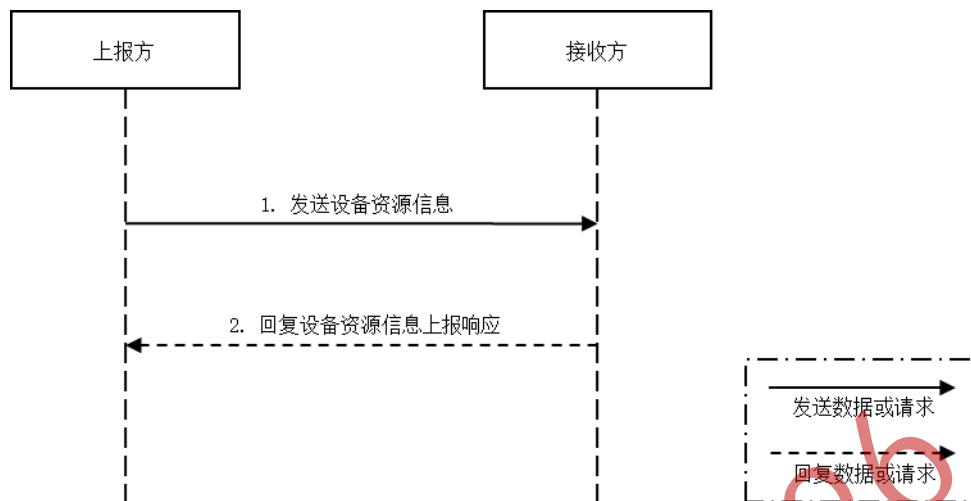


图 7 设备资源上报接口通信流程

9.3.2 接口

9.3.2.1 设备总量资源上报接口

设备总量资源上报接口见表11，接口消息体参数说明见表4。

表 11 设备总量资源上报接口

HTTP 方法	URI	消息体
PATCH	/System/ResourceInfo/TotalResource	<TotalResource>

设备总量资源上报接口返回值说明见表12。

表 12 设备总量资源上报接口返回值

名称	标识符	类型	选择状态	长度	备注
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败

9.3.2.2 设备可用资源上报接口

设备可用资源上报接口见表13，接口消息体参数说明见表10。

表 13 设备可用资源上报接口

HTTP 方法	URI	消息体
PATCH	/System/ResourceInfo/AvailableResource	<AvailableResource>

设备可用资源上报接口返回值说明见表14。

表 14 设备可用资源上报接口返回值

名称	标识符	类型	选择状态	长度	备注
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败

9.4 设备算法信息查询接口

9.4.1 通信流程

数字视网膜系统中,协同调度方应支持对协同执行方已具备的算法进行查询,从而掌握其算法能力,为调度提供参考。

设备算法信息查询接口的通信流程见图8。查询发送方为协同调度方,查询接收方为协同执行方。

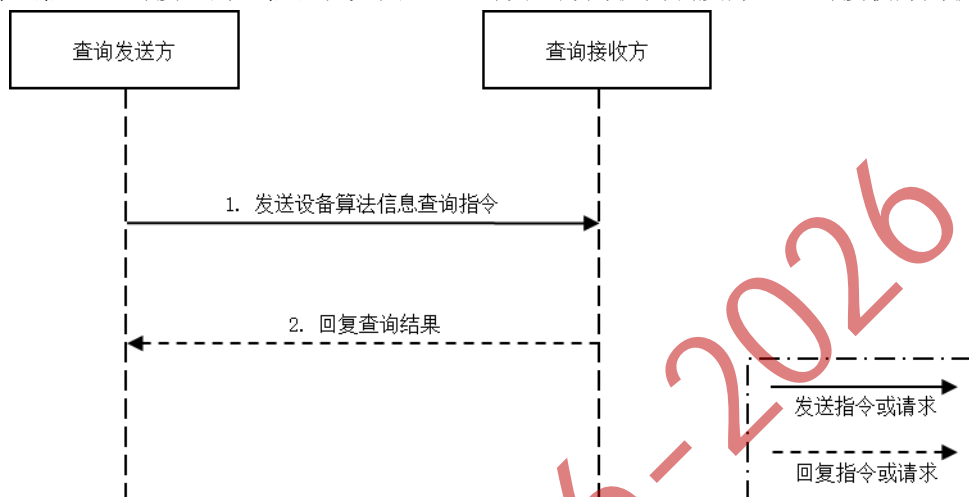


图8 设备算法信息查询接口通信流程

9.4.2 接口

设备算法信息查询接口见表15。

表15 设备算法信息查询接口

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于20字节	值为“algorithm_res_query”
指令序号	CommandIdx	int	可选	4字节	
设备ID	DeviceID	string	可选	20字节	接收指令的设备的标识
指令数据	MessageBody	JSON数据包	必选		
查询设备列表	+DeviceList	JSON array	必选		支持单独查询和批量查询
查询设备ID	++DeviceID	string	必选	20字节	以设备ID作为条件进行查询
是否逐级查询	++HierarchicalQuery	string	必选	1字节	0: 不逐级查询, 1: 逐级查询

设备算法信息查询接口返回值见表16。

表16 设备算法信息查询接口返回值

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于20字节	值为“algorithm_res_query”
指令序号	CommandIdx	int	可选	4字节	下发的指令序号
状态码	StatusCode	string	必选	1字节	返回码 0: 成功, 1: 失败
设备算法信息列表	AlgResourceList	JSON array	必选		<AlgResource>对象数组, 见表17

设备算法信息查询接口返回消息体<AlgResource>参数见表17。

表 17 设备算法信息查询接口返回消息体<AlgResource>参数

名称	标识符	类型	选择状态	长度	备注
设备 ID	DeviceID	string	必选	20 字节	
算法信息列表	AlgInfoList	JSON array	必选		
算法 ID	+AlgID	int	必选	4 字节	
算法类型	+AlgType	string	必选	小于 128 字节	
算法版本	+AlgVersion	string	必选	小于 20 字节	
算法文件名称	+AlgFileName	string	可选	小于 64 字节	
算法状态	+AlgStatus	int	必选	4 字节	0: 使用中 1: 未使用 2: 已失效 3 以上为扩展状态
算法下载时间	+AlgDownloadTime	datetime	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”
算法关联的模型信息	+ModelList	JSON array	可选		<ModelInfo>对象数组，详细 信息见表 18

设备算法信息查询接口返回消息体<ModelInfo>参数见表18。

表 18 设备算法信息查询接口返回消息体<ModelInfo>参数

名称	标识符	类型	选择状态	长度	备注
模型 ID	ModelID	int	必选	4 字节	
模型版本	ModelVersion	string	必选	小于 20 字节	
模型文件名	ModelFileName	string	可选	小于 128 字节	
模型文件大小	ModelSize	string	可选	小于 32 字节	
模型更新时间	ModelUpdateTime	datetime	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”

9.5 任务请求发送接口

9.5.1 通信流程

任务请求发送接口用于发送任务发起方的任务需求，以进行后续调度和执行。

任务请求发送接口的通信流程见图9。发送方为任务发起方，接收方为协同调度方。

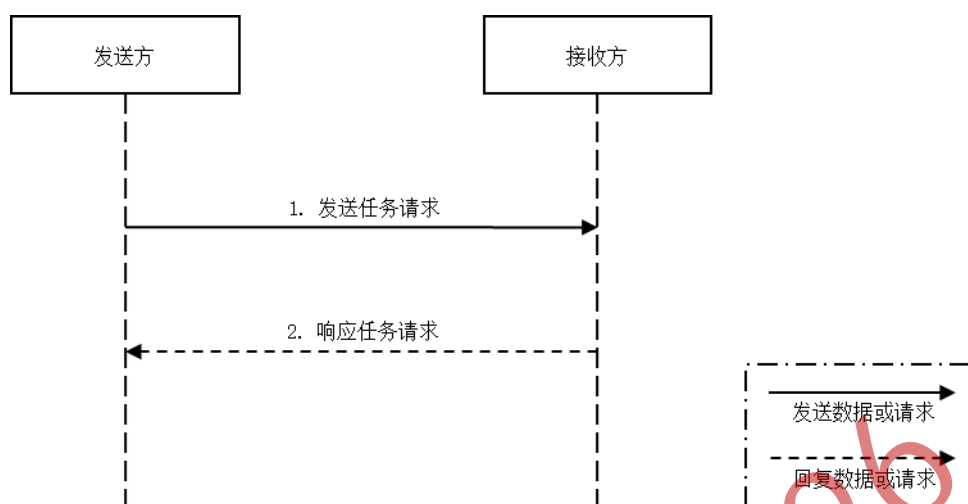


图 9 任务请求发送接口通信流程

9.5.2 接口

任务请求发送接口见表19。

表 19 任务请求发送接口

HTTP 方法	URI	消息体
POST	/System/TaskRequest/	<TaskBasicInfo>对象，具体参数见表 20

任务请求发送接口消息体<TaskBasicInfo>参数见表20。

表 20 任务请求发送接口消息体<TaskBasicInfo>参数

名称	标识符	类型	选择状态	长度	备注
任务类型	TaskType	int	可选	4 字节	
任务名称	TaskName	string	必选	小于或等于 128 字节	
任务描述	Desc	string	可选	小于或等于 512 字节	
任务优先级	Priority	int	可选	4 字节	取值为 1-10，数值越大，优先级越高
任务发起信息	InitiationInfo	JSON object	必选		
设备 ID	+DeviceID	string	必选	20 字节	发起任务请求的设备 ID
任务申请人	+Applicant	string	可选	小于或等于 128 字节	申请人
任务发起时间	+InitiateTime	datetime	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”
任务更新时间	+UpdateTime	datetime	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”
自定义信息	+CustomInfo	JSON object	可选		关于任务发起的自定义参数
任务约束	Constraints	JSON object	可选		例如，执行任务所要求的时间范围、空间范围以及时延、精度等性能指标方面的约束条件

任务请求发送接口返回值见表21。

表 21 任务请求发送接口返回值

名称	标识符	类型	选择状态	长度	备注
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败
任务 ID	TaskID	string	可选	小于或等于 128 字节	当接受任务请求时返回该参数
消息描述	Message	string	可选	小于或等于 512 字节	任务请求处理结果的详细描述, 含请求失败原因

9.6 任务调度接口

9.6.1 通信流程

任务调度接口用于将任务或其子任务发送到满足执行要求的协同执行方上, 并指定任务需处理的数据来源、使用的算法、算法分析参数等内容。

任务调度接口的通信流程见图10。指令发送方为协同调度方, 指令接收方为协同执行方。

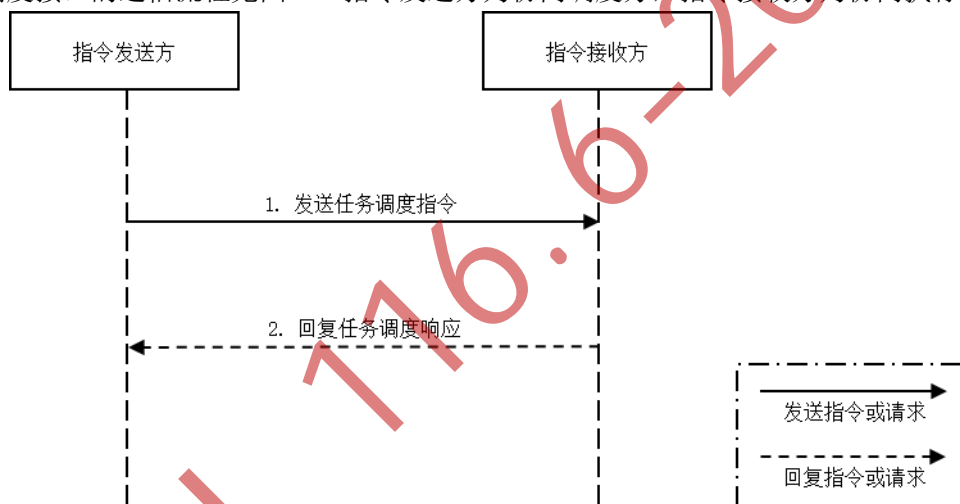


图 10 任务调度接口通信流程

9.6.2 接口

任务调度接口见表22。

表 22 任务调度接口

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_dispatch”
指令序号	CommandIdx	int	可选	4 字节	
设备 ID	DeviceID	string	可选	20 字节	接收调度指令的设备 ID
指令数据	MessageBody	JSON 数据包	必选		支持同时派发某一设备上执行的多个任务
任务执行设备 ID	+TaskExecDeviceID	string	必选	20 字节	具体执行任务的设备 ID
任务调度信息列表	+TaskDispatchList	JSON array	必选		<TaskDispatchInfo> 对象数组, 详细信息见表 23

任务调度接口消息体<TaskDispatchInfo>参数见表23。

表 23 任务调度接口消息体<TaskDispatchInfo>参数

名称	标识符	类型	选择状态	长度	备注
任务 ID	TaskID	string	必选	小于或等于 128 字节	可以是一项独立任务, 或者任务分解得到的某个子任务
任务名称	TaskName	string	可选	小于或等于 128 字节	
任务描述	Desc	string	可选	小于或等于 512 字节	
任务类型	TaskType	int	可选	4 字节	
任务优先级	Priority	int	可选	4 字节	取值为 1-10, 数值越大, 优先级越高
父任务 ID	ParentTaskID	string	可选	小于或等于 128 字节	父任务标识, 当任务被分解时, 该参数为必填项
依赖任务信息列表	Dependencies	JSON array	可选		所有依赖任务的信息列表
依赖任务 ID	+TaskID	string	必选	小于或等于 128 字节	
依赖任务执行设备 ID	+DeviceID	string	可选	20 字节	
依赖任务的扩展信息	+ExtendedInfo	JSON object	可选		用于扩展定义依赖任务的信息, 例如, 任务类型、任务输出信息等
任务性质	TaskAttr	int	必选	4 字节	0: 常驻任务 1: 指定时段任务 2 以上支持扩展其他情况
任务执行条件	ExecutionCond	JSON object	可选		见表 24
任务算法信息	AlgModelInfo	JSON object	必选		见表 25
任务输入信息	InputInfo	JSON object	必选		任务执行所需的输入信息
数据流获取地址	+StreamURL	string	可选	小于或等于 256 字节	任务获取输入数据流的地址
算法分析参数	+AlgArgs	JSON object	可选		应和算法包的算法技术信息描述中的算法分析参数定义对应
任务处理帧率	+TaskFrameRate	int	可选	4 字节	
任务输出的目的设备信息列表	DstDeviceList	JSON array	必选		
设备 ID	+DeviceID	string	可选	20 字节	任务结果上报的目的设备 ID
输出推送 URL	+OutputURL	string	可选	小于或等于 256 字节	任务结果的输出地址

任务执行条件<ExecutionCond>属性见表24。

表 24 任务执行条件<ExecutionCond>属性

名称	标识符	类型	选择状态	长度	备注
时间范围	TimeRange	JSON array	可选		
任务开始时间	+StartTime	string	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”
任务结束时间	+EndTime	string	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”
执行条件的扩展信息	ExtendedInfo	JSON object	可选		用于扩展定义执行任务的其他条件

任务算法信息<AlgModelInfo>属性见表25。

表 25 任务算法信息<AlgModelInfo>属性

名称	标识符	类型	选择状态	长度	备注
算法 ID	AlgID	int	必选	4 字节	算法 ID
算法描述	AlgDesc	string	可选	小于 1024 字节	
算法版本	AlgVersion	string	必选	小于 20 字节	算法版本
算法文件地址	AlgFileURL	string	可选	小于 256 字节	
算法关联的模型信息	ModelList	JSON array	可选		
模型 ID	+ModelID	int	必选	4 字节	
模型描述	+ModelDesc	string	可选	小于 1024 字节	
模型版本	+ModelVersion	string	必选	小于 20 字节	
模型 URL	+ModelURL	string	可选	小于 256 字节	

任务调度接口返回值见表26。

表 26 任务调度接口返回值

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_dispatch”
指令序号	CommandIdx	int	可选	4 字节	下发的指令序号
状态码	StatusCode	string	必选	1 字节	返回码 0：成功，1：失败
调度结果详情	Message	string	可选	小于或等于 512 字节	

9.7 任务控制接口

9.7.1 通信流程

任务控制接口用于控制任务的启动、暂停、停止等操作，也可对任务的算法分析参数进行更新。任务控制接口的通信流程见图11。指令发送方为协同调度方，指令接收方为协同执行方。

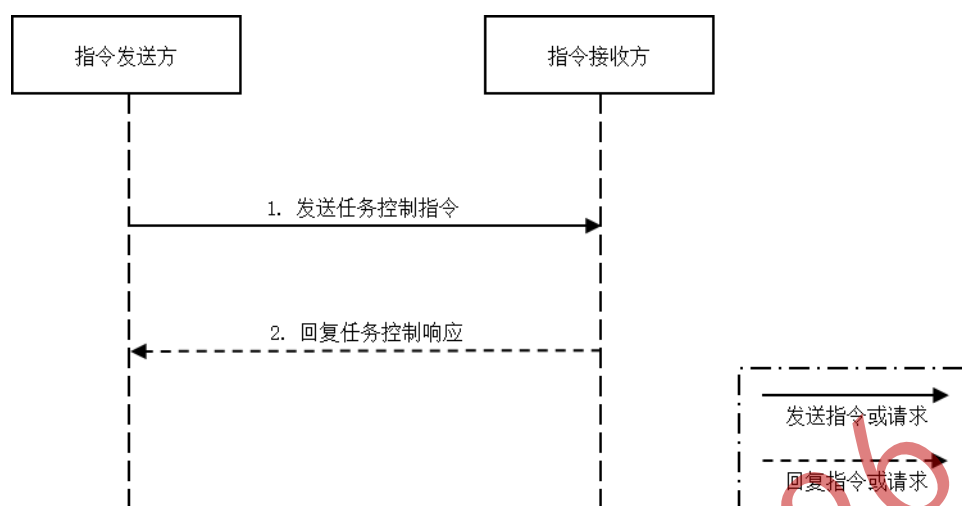


图 11 任务控制接口通信流程

9.7.2 接口

任务控制接口见表27。

表 27 任务控制接口

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_ops_control”
指令序号	CommandIdx	int	可选	4 字节	
指令数据	MessageBody	JSON 数据包	必选		
任务控制信息列表	+OpsControlList	JSON array	必选		<OpsControl>对象数组，详细信息见表 28

任务控制接口消息体<OpsControl>参数见表28。

表 28 任务控制接口消息体<OpsControl>参数

名称	标识符	类型	选择状态	长度	备注
设备 ID	DeviceID	string	可选	20 字节	
任务 ID	TaskID	string	必选	小于或等于 128 字节	
控制操作	Control	int	必选	4 字节	0: 启动任务 1: 停止任务 2: 暂停任务 3: 继续运行任务 4: 调整算法分析参数
算法分析参数	AlgArgs	JSON object	可选		应和算法包的算法技术信息描述中的算法分析参数定义对应

任务控制接口返回值见表29。

表 29 任务控制接口返回值

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_ops_control”
指令序号	CommandIdx	int	可选	4 字节	下发的指令序号
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败

9.8 任务结果发送接口

9.8.1 通信流程

任务结果发送接口用于将任务执行产生的输出数据发送给接收方，以对该数据进行汇聚或后续处理分析。

任务结果发送接口的通信流程见图12。该接口适用于协同调度方与协同执行方之间传输任务结果以及协同执行方之间传输任务结果。

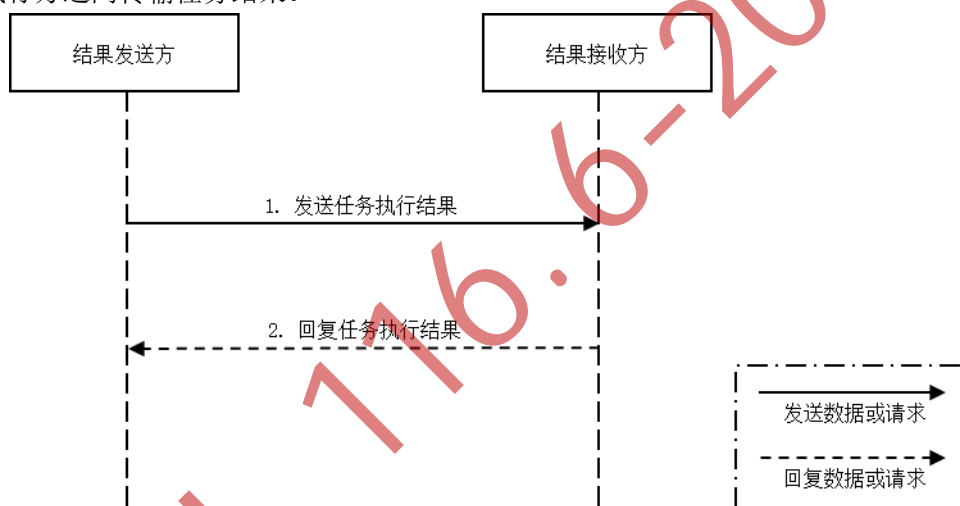


图 12 任务结果发送接口通信流程

9.8.2 接口

任务结果发送接口见表30。

表 30 任务结果发送接口

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_result”
指令序号	CommandIdx	int	可选	4 字节	
设备 ID	DeviceID	string	可选	20 字节	接收指令的设备的标识
指令数据	MessageBody	JSON 数据包	必选		消息体参数见表 31

任务结果发送接口消息体参数见表31。

表 31 任务结果发送接口消息体参数

名称	标识符	类型	选择状态	长度	备注
任务 ID	TaskID	string	必选	小于或等于 128 字节	子任务标识
父任务 ID	ParentTaskID	string	可选	小于或等于 128 字节	父任务标识
设备 ID	DeviceID	string	必选	20 字节	
目的设备信息列表	DstDeviceList	JSON array	必选		
目的设备 ID	+DeviceID	string	必选	20 字节	
算法 ID	AlgID	int	必选	4 字节	算法 ID
算法版本	AlgVersion	string	必选	小于 20 字节	
算法关联的模型信息	ModelList	JSON array	可选		
模型 ID	+ModelID	int	必选	4 字节	
模型版本	+ModelVersion	string	必选	小于 20 字节	
经度	Longitude	string	可选	小于 16 字节	经度格式为 n10,6; 精确到小数点后 6 位
纬度	Latitude	string	可选	小于 16 字节	纬度格式为 n10,6; 精确到小数点后 6 位
时间戳	Ts	long	必选	8 字节	
任务结果生成时间	CreateTime	datetime	可选	小于 24 字节	按照格式: “YYYY-MM-DD hh:mm:ss.ss”
任务执行结果数据	TaskResultData	JSON 数据包	必选		可选择性输出以下类型的数据
图片数据	+ImageResult	JSON object	可选		
非结构化图片数据	++ImageData	base64Binary	可选		图片二进制数据
图片基本信息	++ImageInfo	JSON object	可选		应符合 T/AI 116.1—2021 中 9.2.6.1 的规定
图片特征信息	++ImageFeatureInfo	JSON object	可选		应符合 T/AI 116.1—2021 中 9.2.6.3 的规定
视频数据	+VideoResult	JSON object	可选		
非结构化视频片段数据	++VideoSliceData	base64Binary	可选		视频片段二进制数据
视频片段基本信息	++VideoSliceInfo	JSON object	可选		应符合 T/AI 116.1—2021 中 9.2.6.2 的规定
视频特征信息	++VideoFeatureInfo	JSON object	可选		应符合 T/AI 116.1—2021 中 9.2.6.3 的规定
分析结果数据	+AnalysisResult	JSON array	可选		
特征信息	++FeatureInfo	JSON object	可选		应符合 T/AI 116.1—2021 中 9.2.6.3 的规定

表 31 (续)

名称	标识符	类型	选择状态	长度	备注
算法结果类型	++Type	int	可选	4 字节	0: 人脸 1: 人体 2: 车辆 3: 物品 4: 场景; 类型支持扩展
算法结果数据	++ResultData	JSON object	可选		应符合 T/AI 116.1— 2021 中 9.2.6.4 的规定

任务结果发送接口返回值见表32。

表 32 任务结果发送接口返回值

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_result”
指令序号	CommandIdx	int	可选	4 字节	下发的指令序号
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败

9.9 任务状态查询接口

9.9.1 通信流程

任务状态包括任务运行状态、进度等信息，任务状态查询接口用于通过查询的方式向任务执行者获取任务状态信息。

任务状态查询接口的通信流程见图13。查询发送方为协同调度方，查询接收方为协同执行方。

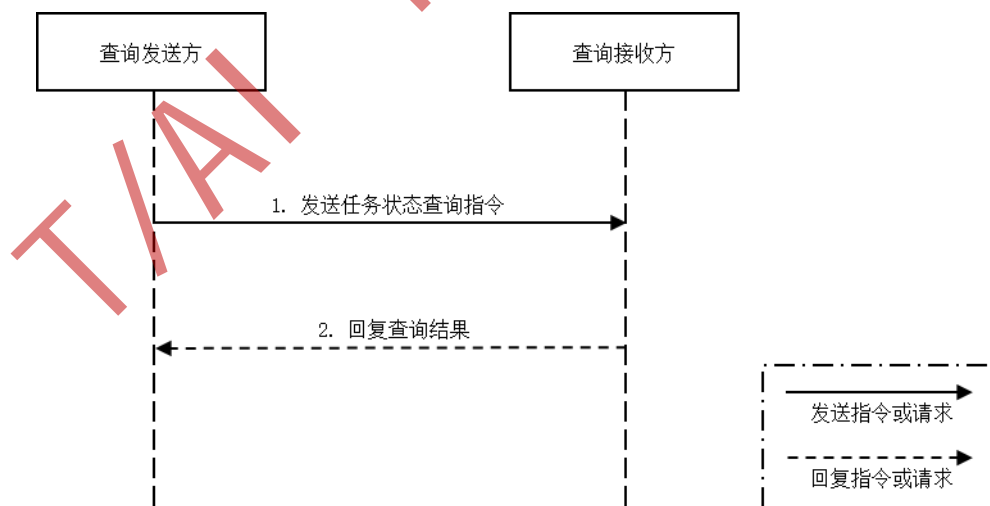


图 13 任务状态查询接口通信流程

9.9.2 接口

任务状态查询接口见表33。

表 33 任务状态查询接口

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_status_query”
指令序号	CommandIdx	int	可选	4 字节	
设备 ID	DeviceID	string	可选	20 字节	接收指令的设备的标识
指令数据	MessageBody	JSON 数据包	必选		以下查询条件应至少填写一项
查询设备列表	+DeviceList	JSON array	可选		支持单独查询和批量查询
查询设备 ID	++DeviceID	string	必选	20 字节	以设备 ID 作为条件进行查询
是否逐级查询	++HierarchicalQuery	string	必选	1 字节	0: 不逐级查询, 1: 逐级查询
任务 ID 列表	++TaskIDList	JSON array	可选		支持一个至多个任务的查询

任务状态查询接口返回值见表34。

表 34 任务状态查询接口返回值

名称	标识符	类型	选择状态	长度	备注
指令类型	CommandType	string	必选	小于 20 字节	值为“task_status_query”
指令序号	CommandIdx	int	可选	4 字节	下发的指令序号
状态码	StatusCode	string	必选	1 字节	返回码 0: 成功, 1: 失败
任务状态信息集合	TaskStatusSet	JSON array	必选		<DevTaskStatus>对象数组, 详细信息见表 35

任务状态查询接口返回消息体<DevTaskStatus>参数见表35。

表 35 任务状态查询接口返回消息体<DevTaskStatus>参数

名称	标识符	类型	选择状态	长度	备注
设备 ID	DeviceID	string	必选	20 字节	
任务状态信息列表	TaskStatusList	JSON array			<TaskStatus>对象数组, 详细信息见表 36

任务状态查询接口返回消息体<TaskStatus>参数见表36。

表 36 任务状态查询接口返回消息体<TaskStatus>参数

名称	标识符	类型	选择状态	长度	备注
任务 ID	TaskID	string	必选	小于或等于 128 字节	
父任务 ID	ParentTaskID	string	可选	小于或等于 128 字节	
任务名称	TaskName	string	可选	小于或等于 128 字节	
任务类型	TaskType	int	可选	4 字节	

表 36 (续)

名称	标识符	类型	选择状态	长度	备注
任务状态	ExecutionStatus	int	必选	4 字节	0: 已完成 1: 执行中 2: 暂停中 3: 排队中 4: 执行异常
执行进度	Progress	int	可选	4 字节	以百分比形式表示, 取值为 0-100%
资源占用情况	ResourcesOccupied	JSON object	可选		任务执行实际占用的资源情况, 例如, CPU、内存、编解码资源等
异常代码	ExceptionCode	int	可选	4 字节	10: 环境不支持 20: 资源不足 30: 前置任务异常 40: 执行性能不足 50: 执行过程出错 60: 数据源错误 支持扩展其他情况
异常描述	ExceptionDesc	string	可选	小于或等于 512 字节	

9.10 任务状态上报接口

9.10.1 通信流程

任务状态上报接口用于任务执行者主动上报任务状态信息。

任务状态上报接口的通信流程见图14。上报方为协同执行方，接收方为协同调度方。

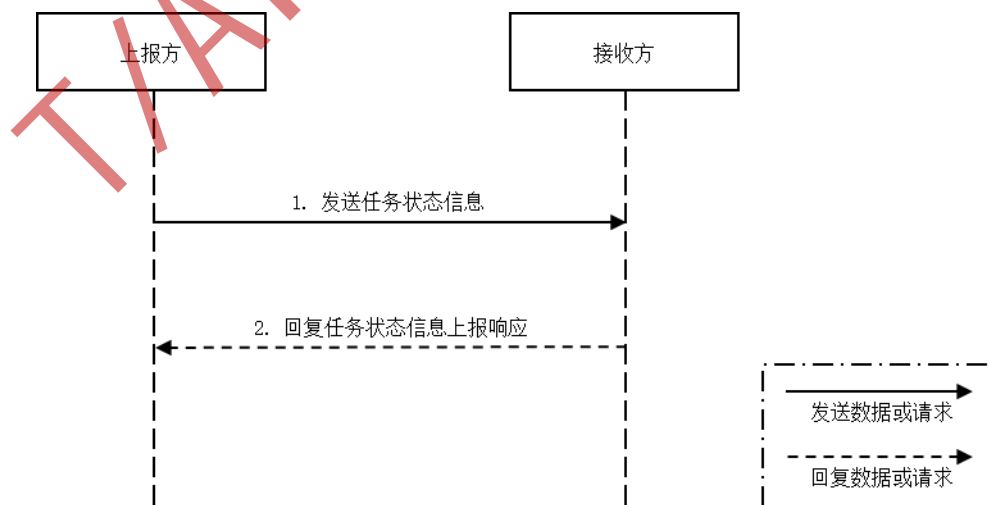


图 14 任务状态上报接口通信流程

9.10.2 接口

任务状态上报接口见表37。

表 37 任务状态上报接口

HTTP 方法	URI	消息体
PATCH	/System/TaskStatus/	<DevTaskStatus>对象或数组，具体参数见表 35

任务状态上报接口返回值见表38。

表 38 任务状态上报接口返回值

名称	标识符	类型	选择状态	长度	备注
状态码	StatusCode	string	必选	1 字节	返回码 0：成功，1：失败

T/AI 116.6-2026

附录 A
(资料性)
任务信息格式

在任务调度与管理过程中,需对任务及其子任务信息进行记录、显示等,任务信息数据结构见表A.1。

表 A.1 任务信息数据结构<TaskInfo>

名称	标识符	类型	选择状态	长度	备注
任务 ID	TaskID	string	必选	小于或等于 128 字节	
任务名称	TaskName	string	必选	小于或等于 128 字节	
任务描述	Desc	string	可选	小于或等于 512 字节	
任务类型	TaskType	int	可选	4 字节	
任务优先级	Priority	int	可选	4 字节	取值为 1-10, 数值越大, 优先级越高
任务创建信息	CreateInfo	JSON object	可选		
任务申请人	+Applicant	string	可选	小于或等于 128 字节	申请人
任务创建时间	+CreateTime	datetime	可选	小于 24 字节	按照格式:“YYYY-MM-DD hh:mm:ss.ss”
任务更新时间	+UpdateTime	datetime	可选	小于 24 字节	按照格式:“YYYY-MM-DD hh:mm:ss.ss”
任务约束	Constraints	JSON object	可选		例如, 执行任务所要求的时间范围、空间范围以及时延、精度等性能指标方面的约束条件
任务状态	ExecutionStatus	int	必选	4 字节	见表 36 中的相关规定
设备列表	DeviceList	JSON array	必选		协同完成该任务的所有设备 ID 列表
设备 ID	+DeviceID	string	必选	20 字节	
子任务信息列表	SubtaskInfoList	JSON array	可选		<SubtaskInfo>对象数组, 详细信息见表 A.2
任务逻辑拓扑关系	TaskTopology	string	可选	小于或等于 256 字节	任务逻辑拓扑关系载体标识

子任务信息数据结构见表A.2。

表 A.2 子任务信息数据结构<SubtaskInfo>

名称	标识符	类型	选择状态	长度	备注
任务执行者信息	TaskExecutor	JSON object	必选		
设备 ID	+DeviceID	string	必选	20 字节	
设备名称	+DeviceName	string	可选	小于 40 字节	
设备类型	+DeviceType	string	可选	小于 32 字节	

表 A.2 (续)

名称	标识符	类型	选择状态	长度	备注
任务调度信息	TaskDispatchInfo	JSON object	必选		具体参数见表 23
任务创建时间	CreateTime	string	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”
任务更新时间	UpdateTime	string	可选	小于 24 字节	按照格式：“YYYY-MM-DD hh:mm:ss.ss”
任务状态信息	TaskStatusInfo	JSON object	必选		
任务状态	+ExecutionStatus	int	必选	4 字节	见表 36 中的相关规定
执行进度	+Progress	int	可选	4 字节	以百分比形式表示，取值为 0-100%
异常代码	+ExceptionCode	int	可选	4 字节	见表 36 中的相关规定
异常描述	+ExceptionDesc	string	可选	小于或等于 512 字节	

附 录 B
(资料性)
端边云协同工作交互过程

B.1 基础协同模式

B.1.1 主要特点

基础协同模式是指逻辑结构可抽象为基础协同架构的数字视网膜系统所对应的协同工作范式，主要特点包括：

- a) 任务可由两层逻辑节点协同完成；
- b) 最小功能单元为数据预处理/后处理、编解码、推理、数据处理、分析与融合、设备控制等类型的任务。

B.1.2 交互流程

基础协同模式下的协同工作交互流程见图B.1。图B.1中概括地给出了基础协同模式下协同交互包含的主要阶段、总体流程和通信要素，图中矩形框内的部分执行过程称之为片段，标识符用于表示相应交互过程执行的条件和方式，其中，Opt表示可选执行项，Loop表示可重复执行的操作。图中的关键点说明如下：

- a) 任务发起方和协同调度方、协同调度方和各协同执行方之间建立网络连接；
- b) 协同执行方可以包括一个或多个；
- c) 协同工作过程包括任务发起、任务创建、资源管理、调度规划和任务协同调度与管理等阶段；
- d) 协同调度方同时具备协同执行方上的子任务下发以及完整任务及其子任务的控制和管理功能；
- e) 协同执行方具备子任务执行和本地子任务管理功能；
- f) 通过协同调度方的调度规划，一项任务可分解为一个或多个子任务，分配给一个或多个协同执行方来执行；
- g) 资源状态查询表示协同调度方需获取参与协同工作的执行方资源情况，可以通过查询的方式获取资源情况，也可以通过资源上报的方式获取实时更新的资源情况；
- h) 任务状态查询可根据一定的策略重复发生，具体时间和间隔不限；
- i) 任务执行过程中，可根据任务中包含的条件或者任务调度规划生成的任务依赖关系对任务进行控制，具体时间和间隔不限。

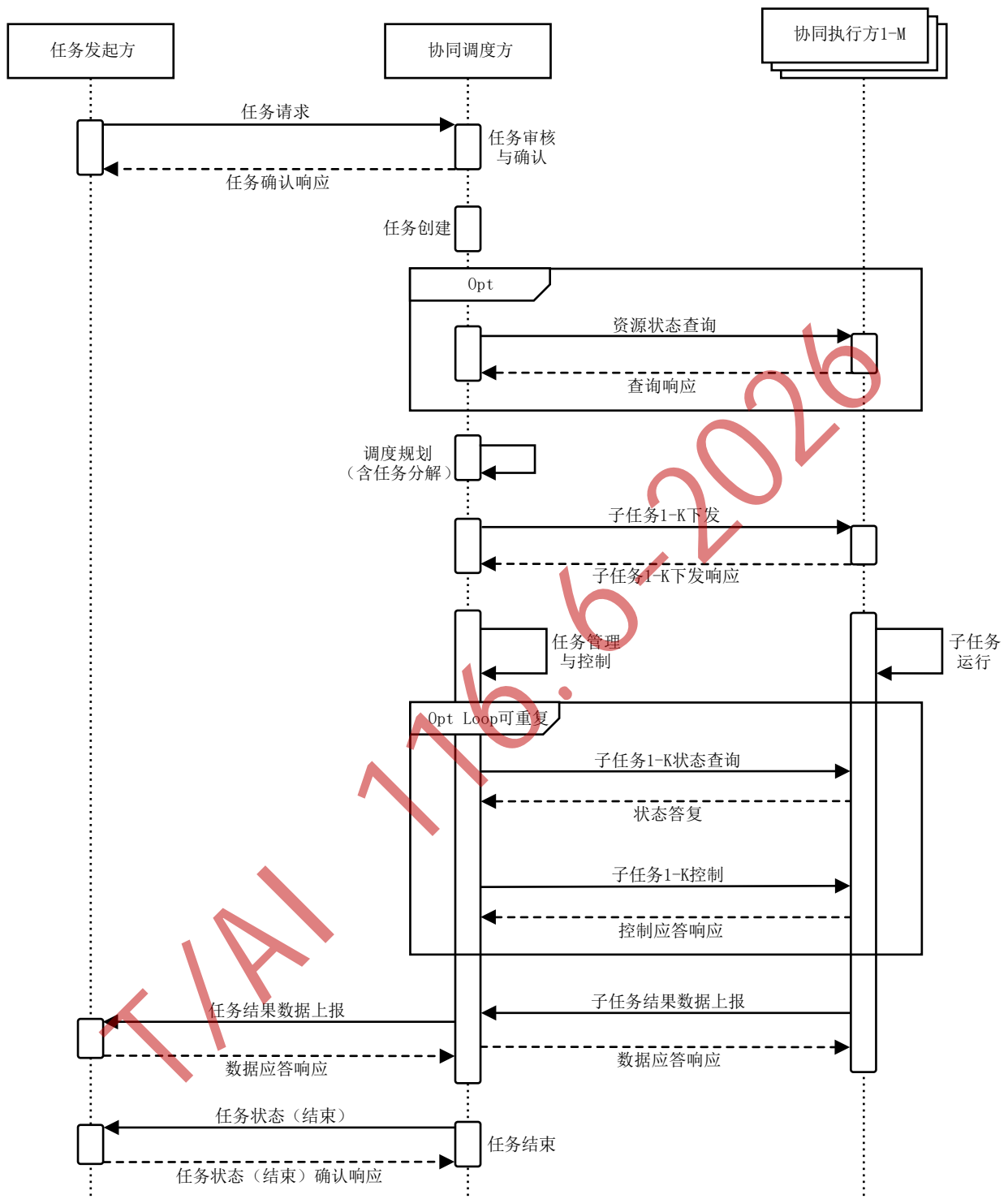


图 B.1 基础协同模式下的协同工作交互流程

B.1.3 协同工作过程

端边云协同处理任务时的具体工作过程描述如下：

- a) 任务发起：任务发起方生成一个任务请求（关于任务需求的描述信息，例如任务类型、输入数据源、任务优先级等），并将任务请求发送给协同调度方；

- b) 任务审核与确认：协同调度方收到任务请求后，验证任务有效性，审核内容例如任务发起方的身份、数据访问权限等，并检查当前系统是否能够处理该任务；
 - 1) 若任务有效且系统具备处理该任务的能力，则协同调度方接受该任务并将其加入调度队列，并发送确认消息给任务发起方；
 - 2) 若任务无效或系统不具备处理该任务的能力，则协同调度方发送错误消息给任务发起方，返回任务无法执行的通知及失败原因。
- c) 任务调度：协同调度方对任务进行分解编排，得到任务逻辑拓扑关系和子任务集合，根据任务优先级、资源需求、算法工作流以及系统平台信息、负载情况等信息，确定每个子任务的算法和模型、依赖关系以及执行每个子任务的协同执行方；
- d) 任务分配：协同调度方将子任务调度信息发送至对应的协同执行方上；
- e) 任务执行：协同执行方接收到任务后，进行任务可执行性检查；
 - 1) 若协同执行方无法执行该任务，将调度失败原因反馈给协同调度方，协同调度方可重新调度；
 - 2) 若协同执行方能够执行该任务，则按照任务条件启动执行任务，执行过程中与协同调度方或其他协同执行方进行数据交互与同步，交互内容例如视频、特征、中间结果等数据，将任务执行状态反馈给协同调度方进行统一监控，同时，对于常驻任务，协同调度方将整体任务执行结果发送给任务发起方。
- f) 错误处理与恢复：若任务执行过程中发生错误（如资源不足、数据不可用、执行超时等），协同执行方立即向协同调度方报告错误信息，包括错误类型和详细信息，协同调度方根据错误类型进行相应的错误处理，例如重新调度任务、通知任务发起方等，通知协同执行方执行恢复操作，如重试、回滚或恢复中断任务等，任务恢复正常执行后，协同调度方更新和保存任务状态，协同执行方继续反馈执行进度和结果数据等内容；
- g) 任务结束与终止：任务成功执行完毕后，协同调度方将任务完成信息发送给任务发起方，任务发起方向协同调度方回复确认收到的信息，标记任务结束状态或发起后续处理操作；
- h) 资源释放与数据清理：协同执行方释放执行任务所占用的内存、算力、存储等资源，清理临时数据和缓存，并向协同调度方上报当前资源状态，协同调度方记录并更新系统资源状态。

B.2 级联协同模式

B.2.1 主要特点

级联协同模式是指逻辑结构可抽象为级联协同架构的数字视网膜系统所对应的协同工作范式，主要特点包括：

- a) 任务可由两层以上的逻辑节点协同完成；
- b) 最小功能单元为数据预处理/后处理、编解码、推理、数据处理、分析与融合、设备控制等类型的任务；
- c) 任务由上级协同调度方或下级协同调度方导入；
- d) 上级协同调度方承担任务整体调度控制功能；
- e) 下级协同调度方具有局部任务调度控制功能。

B.2.2 交互流程

对于级联协同模式下的自上而下发起任务和自下而上发起任务两种情况，协同工作交互流程分别见图B.2和图B.3。图B.2和图B.3中概括地给出了级联协同模式下协同交互包含的主要阶段、总体流程和通信要素，图中标识符的含义见附录B.1.2。图中的关键点说明如下：

- a) 任务发起方和上级协同调度方或下级协同调度方之间建立网络连接；
- b) 上级协同调度方、下级协同调度方和协同执行方之间依次建立网络连接；
- c) 下级协同调度方和协同执行方均可包括一个或多个；
- d) 协同工作过程包括任务发起、任务创建、资源管理、调度规划和任务协同调度与管理等阶段；
- e) 上级协同调度方同时具备调度规划、子任务下发以及完整任务及其子任务的控制和管理功能；
- f) 下级协同调度方具备局部任务管理与控制、子任务相关指令或数据的转发或下发功能；
- g) 下级协同调度方可进行调度规划，以支持对上级协同调度方下发的子任务作进一步分解与调度，或对来自于任务发起方的任务请求进行局部调度；
- h) 协同执行方具备子任务执行和本地子任务管理功能；
- i) 资源状态查询表示上级/下级协同调度方需掌握参与协同工作的执行方资源情况，可以通过查询的方式获取资源情况，可针对部分节点进行查询或级联查询，也可以通过资源上报的方式获取实时更新的资源情况；
- j) 任务状态查询可根据一定的策略重复发生，具体时间和间隔不限；
- k) 任务执行过程中，可根据任务中包含的条件或者任务调度规划生成的任务依赖关系对任务进行控制，具体时间和间隔不限。

T/AI 116.6-2026

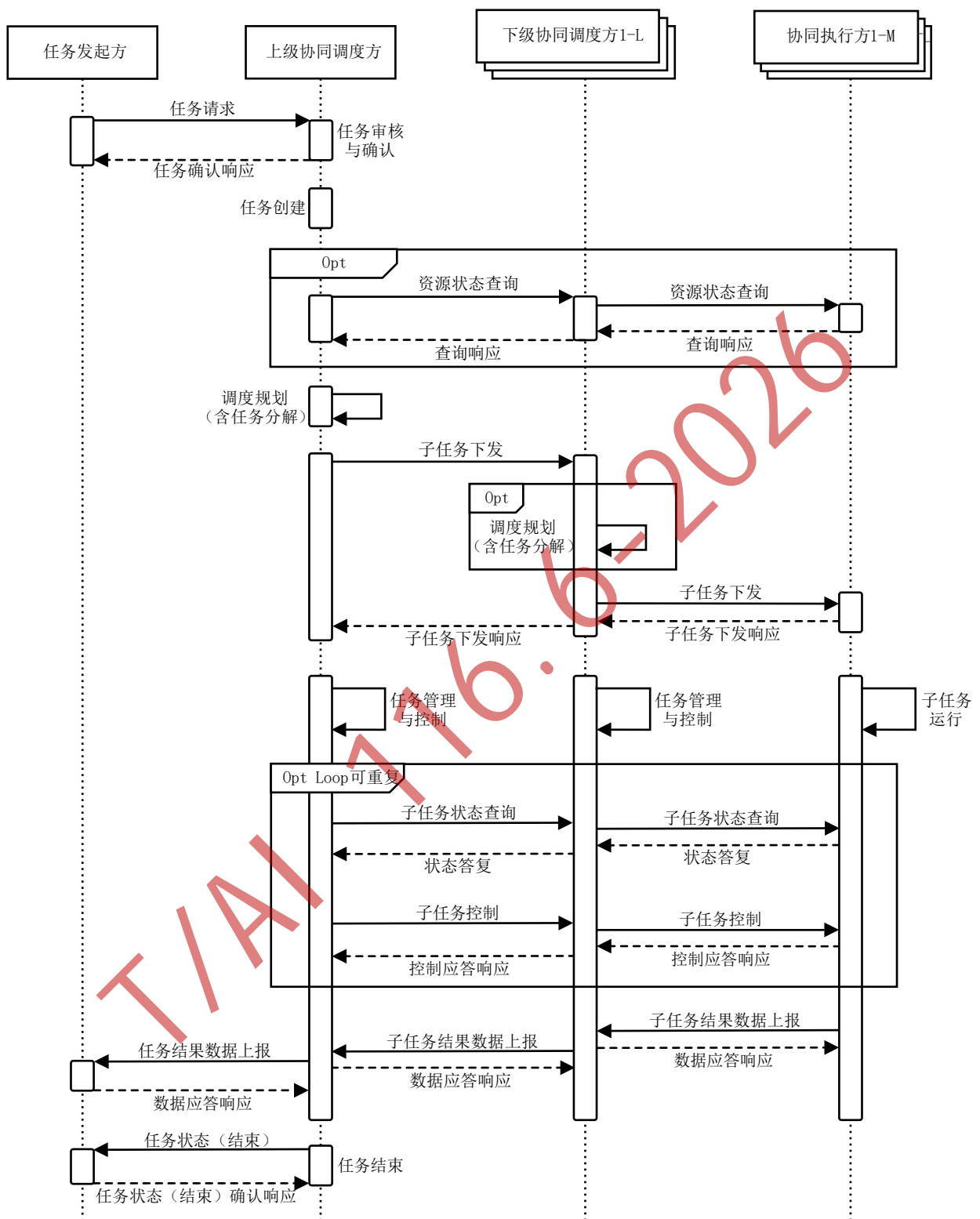


图 B.2 级联协同模式下自上而下发起任务时的协同工作交互流程

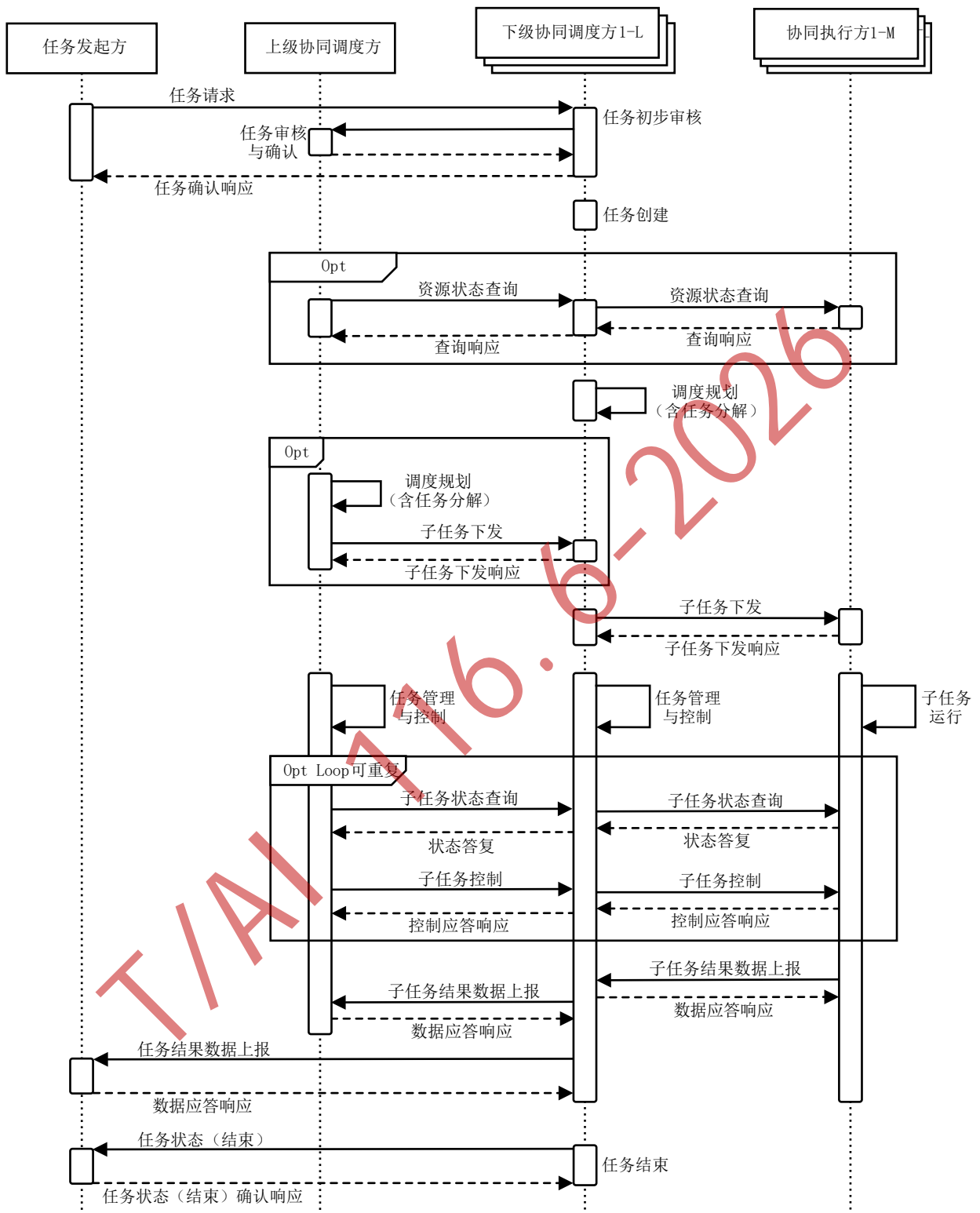


图 B.3 级联协同模式下自下而上发起任务时的协同工作交互流程

B.2.3 协同工作过程

a) 自上而下发起任务

自上而下发起任务时的协同工作过程与基础协同模式基本一致。特别之处说明如下：

- 1) 任务发起方将任务请求发送至上级协同调度方；
- 2) 上级协同调度方对任务进行审核，包括身份、权限和数据等方面的验证，并将系统是否接受该任务的响应反馈给任务发起方；
- 3) 任务调度可由上级协同调度方独立完成，也可由下级协同调度方对分解后的任务做进一步调度；
- 4) 上级协同调度方负责任务全生命周期的统一管理，并将任务结果数据和任务状态反馈给任务发起方。

b) 自下而上发起任务

自下而上发起任务时的协同工作过程与基础协同模式基本一致。特别之处说明如下：

- 1) 任务发起方将任务请求发送至下级协同调度方；
- 2) 下级协同调度方对任务进行初步审核，例如，检查任务请求的格式和参数、验证任务发起方的身份信息和权限等，若初步审核通过，将任务请求转发给上级协同调度方，若初步审核不通过，则将错误信息反馈给任务发起方；
- 3) 上级协同调度方对任务请求进行全面审核，例如，审核任务发起方的身份，任务类型、资源使用是否符合任务发起方的权限范围，任务执行是否符合系统的安全和合规要求等，若审核通过，上级协同调度方确定任务优先级、任务调度策略等全局信息，由自身或下级协同调度方进行任务调度规划，若审核不通过，则将错误信息反馈给下级协同调度方；
- 4) 下级协同调度方可对审核通过的任务请求进行局部范围内的调度，以对依赖本地资源和紧急的任务做出快速响应，如果下级协同调度方发现局部资源不足以完成任务，可请求上级协同调度方进行协助调度，上级协同调度方可依据系统全局资源、策略等因素进行调度决策；
- 5) 任务执行方将执行结果和状态信息反馈给下级协同调度方，并由下级协同调度方发送至上级协同调度方进行审核确认，最终的任务结果数据和任务状态反馈至任务发起方；
- 6) 根据业务情况，任务结果数据可返回至任务发起方，也可发送至指定位置。

参 考 文 献

- [1] GB/T 30269.401—2015 信息技术 传感器网络 第401部分：协同信息处理：支撑协同信息处理的服务及接口
- [2] GB/T 36620—2018 面向智慧城市的物联网技术应用指南
- [3] GB/T 37684—2019 物联网 协同信息处理参考模型
-

T/AI 116.6-2026