

面向算力网络的算力建模与 度量技术研究

Research on Modeling and Measurement Technology for Computing Power Network

王施霁¹,张岩¹,李传宝²,崔童³,曹畅¹,唐雄燕¹(1. 中国联通研究院,北京 100080;2. 烽火通信科技股份有限公司,湖北 武汉 430073;3. 中国联合网络通信集团有限公司,北京 100033)

Wang Shiji¹,Zhang Yan¹,Li Chuanbao²,Cui Tong³,Cao Chang¹,Tang Xiongyan¹(1. China Unicom Research Institute, Beijing 100080, China; 2. FiberHome Telecommunication Technologies Co., Ltd., Wuhan 430073, China; 3. China United Network Communications Group Co., Ltd., Beijing 100033, China)

摘要:

算力度量与建模技术作为算力网络发展的重要基础,通过统一建模算力需求和资源,结合网络性能指标,形成算网能力模板,为算力路由、管理和计费提供统一度量标准。深入探讨了该技术的发展趋势与挑战,分析了算力网络发展中的算力度量需求。在此基础上对算力网络各层级进行抽象建模,设计了一种基于任务基本处理单元的算力度量和管理方法,并提出了一套完善的算力度量与建模体系,为算力网络的进一步发展提供了有力支撑和新的思路。

关键词:

算力网络;度量;建模;管理

doi:10.12045/j.issn.1007-3043.2024.06.001

文章编号:1007-3043(2024)06-0001-06

中图分类号:TN919

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

As an important foundation for the development of computing power network, computing power measurement and modeling technology form computing network capability templates by unifying the requirements of computing power modeling and resources and combining network performance indicators, which provides a unified measurement standard for computing power routing, management, and billing. It delves into the development trends and technical challenges of this technology, analyzes the computing power measurement requirements in the development of computing power networks. Based on this, it abstracts and models each level of the computing power network, designs a computing power measurement and management method based on the basic processing unit of tasks, and proposes a comprehensive computing power measurement and modeling system, providing strong support and new ideas for the further development of computing power network.

Keywords:

Computing power network; Measurement; Modeling; Management

引用格式:王施霁,张岩,李传宝,等. 面向算力网络的算力建模与度量技术研究[J]. 邮电设计技术,2024(6):1-6.

1 算力建模与度量技术的研究背景

1.1 算力度量与建模技术的研究背景

随着全球互联网技术的飞速进步,数字化和智能化浪潮席卷而来,成为推动社会进步的重要引擎。算力网络作为工业互联网、云计算、大数据等新兴技术发展的重要支撑,是数字经济发展的关键资源,并迎来前所未有的发展机遇。面向未来日益多样化的业

务需求,传统的数据中心和服务器集群已无法满足新业务模式对算力灵活性和高效性的需求,算力资源的提供方逐渐转变为遍布云、边、端的泛在算力,通过网络连接实现高效共享。算力网络作为连接这些资源的桥梁,扮演着全网算力资源平台的角色,整合了云、边、端的算力资源,为用户提供透明、一致的算力服务,满足其多样化、复杂化的计算需求。算力网络必须具备高效、灵活的管理和调度能力,实现对网络中算力资源的统一管控。通过精准评估用户计算任务的算力需求,将需求与管理的算力资源匹配,确保任

收稿日期:2024-05-06

务被准确调度至最合适的资源上,从而为用户提供稳定、高效的算力服务^[1-5]。

实现算力资源的高效管理和按需分配,首要任务是对算力进行准确统一的描述。这要求根据算力资源的特点和参数对算力进行抽象建模,从而进行统一的算力度量。根据中国通信学会的《算力网络前沿报告(2020年)》^[6],泛在化、多元化算力资源的统一度量与建模是算力调度过程的基础,为算力网络资源的灵活匹配与调度提供保障。然而,传统的度量标准已无法满足算力网络的需求,因为算力网络不仅融合了算力设施和网络设施,还涉及计算能力、存储、内存等多个维度的指标。传统的算力度量方法通常是基于静态的硬件特性和指标进行评估,缺乏对算力节点动态性和异构性的考虑。因此,在面对复杂的算力网络环境和动态的业务需求时,传统的算力度量方法可能无法提供准确的度量结果。因此,需要进一步研究和探索新的算力度量方法,以适应算力网络的发展和应用需求^[7-10]。

本文探讨了算力度量和建模的技术与方法,旨在解决算力网络中算力资源管理以及资源分配问题。本文对算力供给侧、需求侧和算力服务执行侧进行了深入分析,明确了在算力服务执行过程中的算力度量需求;构建了涵盖算力服务层、算力抽象层、算力资源管理层、算力资源层的4层算力服务模型;根据算力资源的特性和计算任务的需求,设计了一套具体、可操作的衡量指标;通过将复杂的计算任务拆分为若干基本任务单元,设计基于任务单元的算力度量和管理机制,并提出了一套完善的算力度量和建模技术体系。这一体系为算力网络的度量和管理的提供了解决方案,为算网融合的推进提供了有价值的参考。

1.2 算力度量与建模技术的研究现状

随着技术的不断进步,多个研究机构、产业联盟和标准组织正致力于从不同维度对算力度量进行深入研究,旨在形成一套标准规范,用以量化异构算力资源并满足多样化的业务需求。Frank H. McMahon提出了一个创新性的“浮点运算次数”表示方法,用于衡量每秒可以进行多少次浮点运算,他的研究报告深入探讨了这一概念^[11]。目前,被广泛接受的算力评估方式是根据特定的算法和计算类型,将算力细分为逻辑运算能力、并行计算能力和神经网络计算能力,文献[4]初步建立了统一的算力资源度量和标识体系。郭亮等人在对数据中心进行算力评估的研究中,将算力

细分为通用计算能力、高性能计算能力、存储能力和网络能力四大核心要素^[12]。杜宗鹏等学者则提出了四面三级算力度量技术体系,从节点的计算、通信、内存和存储4个方面出发,并按照3级指标进行建模和描述,包括异构硬件算力、节点服务能力度量和业务支撑能力度量^[13]。李建飞等研究者则以业务体验为导向,提出了联合服务能力的算力建模方法^[14]。国内运营商在2021年的研究报告中也将算力节点的资源划分为计算、通信、内存和存储4个维度,并进行细致的度量评估^[15]。然而,尽管研究工作在不断深入,目前尚未形成统一与标准的算力度量单位^[16-17]。

2 算力网络服务建模技术体系

为了准确描述算力网络的节点服务能力,本文提出了一个两面四层的算力服务建模技术体系的框架,两面包括业务面和管控面,业务面保障各类业务的算力服务需求,管控面侧重算力资源管理和调整。四层包括算力设施层模型、算力服务层模型、算力抽象层和算力业务层模型。算力设施层模型全面度量了硬件设施、基础软件系统及基础算力的性能,涵盖了计算能力、通信能力、内存和存储能力等多个维度。算力服务层模型则聚焦于网络操作系统、算力抽象层支撑层和业务运行环境层,为算力服务的提供与创新提供支撑。算力业务层模型则通过业务拆分和任务基本处理单元管理,满足复杂业务的计算需求,两面四层的算力服务建模的能力和指标如图1所示。

2.1 算力设施层模型

算力设施层模型是一个综合性的框架,涵盖了硬件设施、基础软件系统以及它们所支撑的基础算力。该模型需从多个维度对设施层的性能进行度量和建模,以确保能准确评估和优化整个算力系统的性能。

在计算能力建模方面,模型综合考虑了通用算力、机器学习与神经网络算力、并行推理与训练能力,以及区块链与矿机的哈希计算能力。同时,还关注了音频、视频、图像编解码等个性化算力的度量。这些指标共同被用来对节点或计算集群计算能力进行全面评估。

通信能力建模则聚焦于网络带宽、时延和协议支持。网络带宽决定了节点之间的数据传输速度,而时延则反映了数据在网络中的传输效率。协议支持则根据不同应用场景的需求,提供了相应的通信协议支持,以确保数据的高效传输和处理。

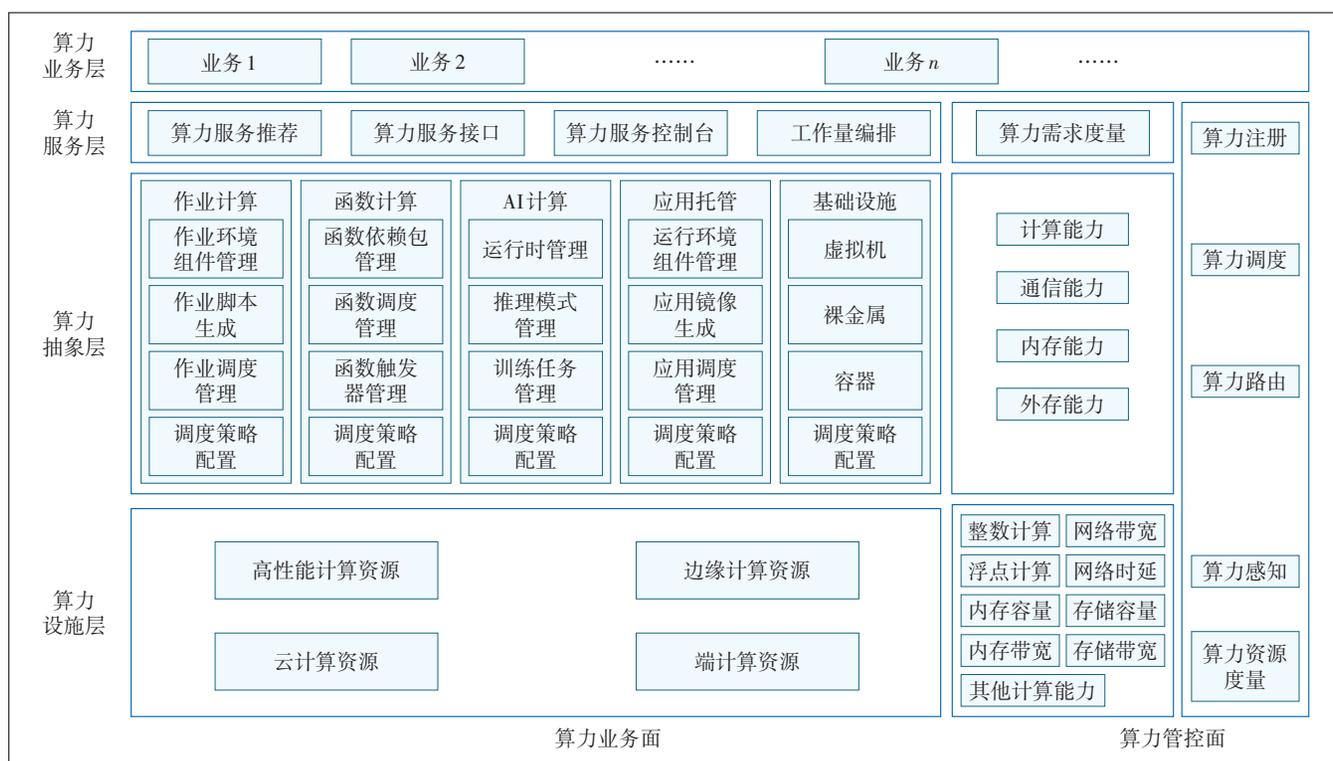


图1 两面四层算力服务模型

内存和存储能力建模也是算力设施层模型的重要组成部分。内存能力从内存容量和内存带宽2个方面进行建模,以反映系统中内存子系统的性能。存储能力则从存储容量、存储带宽和IOPS 3个方面进行建模,以全面评估存储设备的性能和数据传输能力。

2.2 算力服务层和抽象层模型

算力服务层主要由网络操作系统、算力服务支撑层和业务运行环境支撑层构成。

a) 网络操作系统是分布式网络应用的基础操作平台。

b) 算力服务支撑层有分布式存储、分布式计算、容器、虚拟机等。

c) 业务运行环境支撑层是算力服务提供方入网的算力服务提供的运行支撑软件环境,比如BaaS平台、FaaS平台或技术演进产生的其他创新平台。

算力抽象层负责将底层的算力资源(如计算设备、存储设备和网络设备等)进行抽象和封装,以提供统一、灵活的算力服务给上层应用和业务。

2.3 算力业务层模型

算力业务层负责承载多样化的综合算力业务,本文提出了一种基于最小业务单元的建模方式,需将业务逐级拆分为更细粒度的计算任务,并映射至不同计

算节点执行。比如综合业务 k 被划分为 L 个子业务,子业务又被划分为 m 个技术领域任务,技术领域任务根据计算特性被划分为 n 个计算任务,计算任务根据节点算力资源可被拆分为 w 个任务基本处理单元(Task Basic Processing Unit, TBPU)。任务基本处理单元作为最小任务处理单元,由资源配置和代码包组成,具备内存、外存、通信带宽和完成时间等需求配置,以软件模块形式独立运行,支持重复调用和资源释放。TBPU工作于各种运行环境,具备自我统计和执行能力,为实现算力交易与服务提供基础支撑。通过业务逐级拆分和TBPU管理,可有效提升算力资源利用效率,满足复杂业务的分布式计算需求。

表1所示为任务基本处理单元的配置和度量方式描述。

3 基于最小任务单元的算力建模与度量

3.1 算力资源度量功能架构

本文提出了一种基于最小任务单元的算力度量方法,通过监控和采集算力区域内所有计算节点的TBPU信息,综合评估内存、外存、通信带宽和执行时间等关键指标,从而精准度量算力资源并支撑评估与算力交易,其基本原理如图2所示。

表1 任务基本处理单元的配置和度量方式

TBPU 配置	配置子项	度量单位	简要描述
TBPU_Name	-	-	TBPU 命名, 在同一个综合业务里, 名称唯一
TBPU_Run-time	-	ms/Long	TBPU 一个实例完成最大允许时间/长期
TBPU_Mem	TBPU_Mem_Size	MB	包含容量和读写带宽要求
	TBPU_Mem_Spd	MB/s	包含容量和读写带宽要求
TBPU_Stor	TBPU_Stor_Size	MB	包含存储容量和访问带宽和 IOPS 要求
	TBPU_Stor_Spd	MB/s	-
	TBPU_Stor_Iops	-	-
TBPU_Com	TBPU_Com_Spd	MB/s	包含节点通信端口带宽要求
TBPU_Func	-	-	完成 TBPU 需要功能包
TBPU_Handle	-	-	TBPU 入口功能
TBPU_init	-	-	TBPU 冷启动的初始化代码入口
TBPU_init_timeout	-	-	TBPU 启动初始化代码最大执行时间

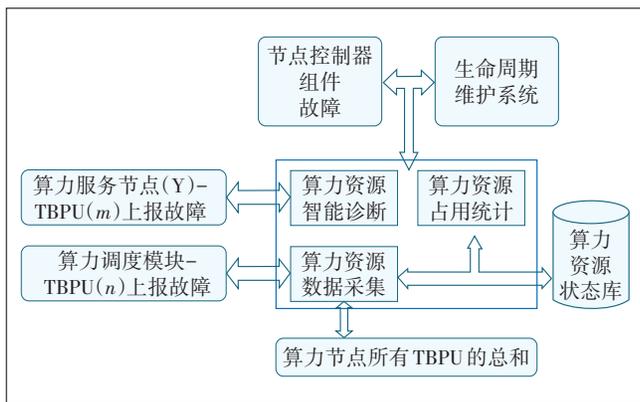


图2 算力资源管度量功能架构

a) 算力资源数据采集。统计区域内算力节点各 TBPU 算力资源(计算、内存、外存、通信)的总和, 并参考控制器上报的节点 CPU 资源及占用率、内存资源及占用率、外存资源及占用率、通信带宽等, 更新各节点的算力资源占用并给节点加上一个标识, 该标识为布尔变量, 有可用和不可用 2 种赋值。

b) 算力资源状态库。采用内存数据库, 以节点为单位统计区域内算力资源状态。

c) 算力资源智能诊断。实时处理算力节点异常上报和调度模块异常上报; 定期查询算力节点维修状况, 更新算力资源状态库, 并通知算力调度模块; 依据诊断结果通知算力资源状态刷新打标签。

d) 算力调度。算力调度模块有自己维护的算力

资源访问列表, 调度反馈失败即堵塞相应的算力供应节点并上报, 直到算力感知服务主动联系重新启用该算力节点。

e) 算力资源占用统计。对某个算力区域内所有计算节点的 TBPU 进行监控和信息采集, 将每个综合业务所有 TBPU 所占用的内存、外存、端口通信带宽和执行时间进行统计求和, 并以此为依据获取算力资源度量值, 支撑算力评估与算力交易。

3.2 算力服务资源管理与度量服务系统

算力服务资源管理与度量服务软件框架的实现如图 3 所示, 主要包括算力调度监控、算力统计与异常监控、算力资源度量服务端。算力资源度量服务端由算力资源智能诊断与数据更新、算力资源统计度量、节点健康数据库、算力资源数据采集等模块构建而成, 各模块功能如下。

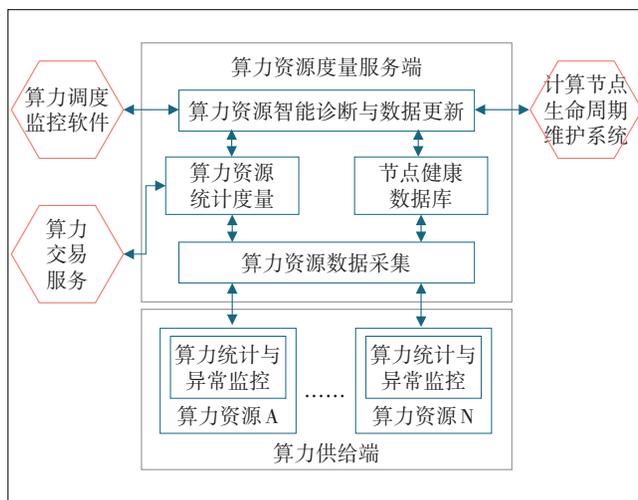


图3 算力度量与管理的软件框架

a) 算力调度监控。运行于算力调度服务系统, 上报算力节点算力调度出现的异常触发。

b) 算力统计与异常监控。运行于算力供给端的主机系统, 上报算力节点所有 TBPU 资源占用状况和 TBPU 个体的异常触发; 算力供给端自带的带外监控系统, 上报节点的算力资源相关组件的健康信息, 主要包括 CPU/内存/物理及虚拟硬盘/网卡/RAID 卡/电源/风扇等信息查询、系统资源门限及超限告警、整机功耗、CPU/内存占用率等。

c) 算力资源度量服务端。包含各类区域算力资源管理软件, 实现区域内算力资源管理与度量, 内嵌组件包括算力资源数据采集、算力资源统计度量模块、节点健康数据库、算力资源智能诊断与数据更新。

3.3 基于TBPU的算力建模

算力资源度量系统对某个算力区域内所有计算节点的TBPU进行信息采集与统计,将每个节点每个TBPU所占用的内存、外存、端口通信带宽和执行时间进行统计求和,并以此为依据获取算力资源度量值,实现算力估价与支付。

这里设某个算力区域内有 M 个节点,节点上有 N 个TBPU,某个TBPU执行时间是 $TBPU_Runtime(n)$,TBPU的内存占用是 $TBPU_Mem_Size(n)$,内存访问速率是 $TBPU_Mem_Spd(n)$,外存占用是 $TBPU_Stor_Size(n)$,外存访问速率为 $TBPU_Stor_Spd(n)$,外存IO访问频次为 $TBPU_Stor_Iops(n)$,TBPU访问占用带宽为 $TBPU_Com_Spd(n)$,可以统计计算获得该区域所有节点对应的算力度量值。

内存容量占用为:

$$MSize = \sum_{m=1}^M Node(m) \sum_{n=1}^N TBPU_Mem_Size(n) \quad (1)$$

内存平均访问速率为:

$$MSpd = \frac{1}{M} \sum_{m=1}^M Node(m) \frac{1}{N} \sum_{n=1}^N TBPU_Mem_Spd(n) \quad (2)$$

外存容量占用为:

$$SSize = \sum_{m=1}^M Node(m) \sum_{n=1}^N TBPU_Stor_Size(n) \quad (3)$$

外存平均访问速率为:

$$SSpd = \frac{1}{M} \sum_{m=1}^M Node(m) \frac{1}{N} \sum_{n=1}^N TBPU_Stor_Spd(n) \quad (4)$$

外存平均IO访问频次为:

$$SIops = \frac{1}{M} \sum_{m=1}^M Node(m) \frac{1}{N} \sum_{n=1}^N TBPU_Stor_Iops(n) \quad (5)$$

任务平均占用通信带宽为:

$$CSpd = \frac{1}{M} \sum_{m=1}^M Node(m) \frac{1}{N} \sum_{n=1}^N TBPU_Com_Spd(n) \quad (6)$$

任务平均执行时间为:

$$T = 1000000 \frac{1}{M} \sum_{m=1}^M Node(m) \frac{1}{N} \sum_{n=1}^N Runtime(n) \quad (7)$$

获取以上算力度量值后,就可以较容易实现算力的估值。选取某次业务在标准节点设备上运行的算力度量值和算力参考价作为标准参考值,实际某次业务的算力交易价Count的计算如下:

$$Count = \frac{MSize}{MSize0} \times K1 + \frac{MSpd}{MSpd0} \times K2 + \frac{SSize}{SSize0} \times K3 + \frac{SSpd}{SSpd0} \times K4 + \frac{SIops}{SIops0} \times K5 + \frac{CSpd}{CSpd0} \times K6 + \frac{T0}{T} \times K7 \quad (8)$$

其中, $MSize0$ 、 $MSpd0$ 、 $SSize0$ 、 $SSpd0$ 、 $SIops0$ 、 $CSpd0$ 、 $T0$ 是业务算力标准参考值, $K1, K2, \dots, K7$ 是其对应的分解评估参考价格。

4 基于TBPU度量方式的算力服务实现方案

基于上述度量方法,本文构建了一种基于TBPU度量方式的算力服务实现方案(见图4)。

a) 算力需求端发起业务申请,先完成业务登记审核。

b) 算力服务编排系统将业务拆解为多个基本处理单元 $a(n)$ 并排序后,将 $a(n)$ 交付调度。

c) 调度系统处理,对 $a(n)$ 业务类型匹配、算力资源位置匹配等,交付算力路由。

d) 算力路由,即网络及节点设备承载传输与路由识别任务,将 $a(n)$ 传送到选定的节点,并建立起临时网络通道。

(a) 如果此时算力供给节点算力资源充足,则TBPU冷启动成功。

(b) 如果此时算力供给侧算力资源不足,则TBPU冷启动失败,进入算力资源度量系统的异常处理流程。

e) $a(n)$ 执行完成并反馈结果,释放节点的计算、内存、外存、网络资源。

f) 如果 $a(n+1)$ 存在,继续跳转到编排系统循环处理。

g) 直到所有基本处理单元处理完成,跳转到编排系统汇总结果并反馈。

h) 一次业务处理结束。

5 结束语

算力建模和度量技术研究是算力网络的发展基础,目前有关算力网络的建模和度量还处于发展初期阶段。本文针对算力网络的算力建模与度量技术进行了深入的研究与探讨,旨在为算力网络的进一步发展提供有力支撑和新的思路。通过对算力网络各层级进行抽象建模,构建了一个两面四层的算力服务建模技术体系的框架,设计了一种基于任务基本处理单

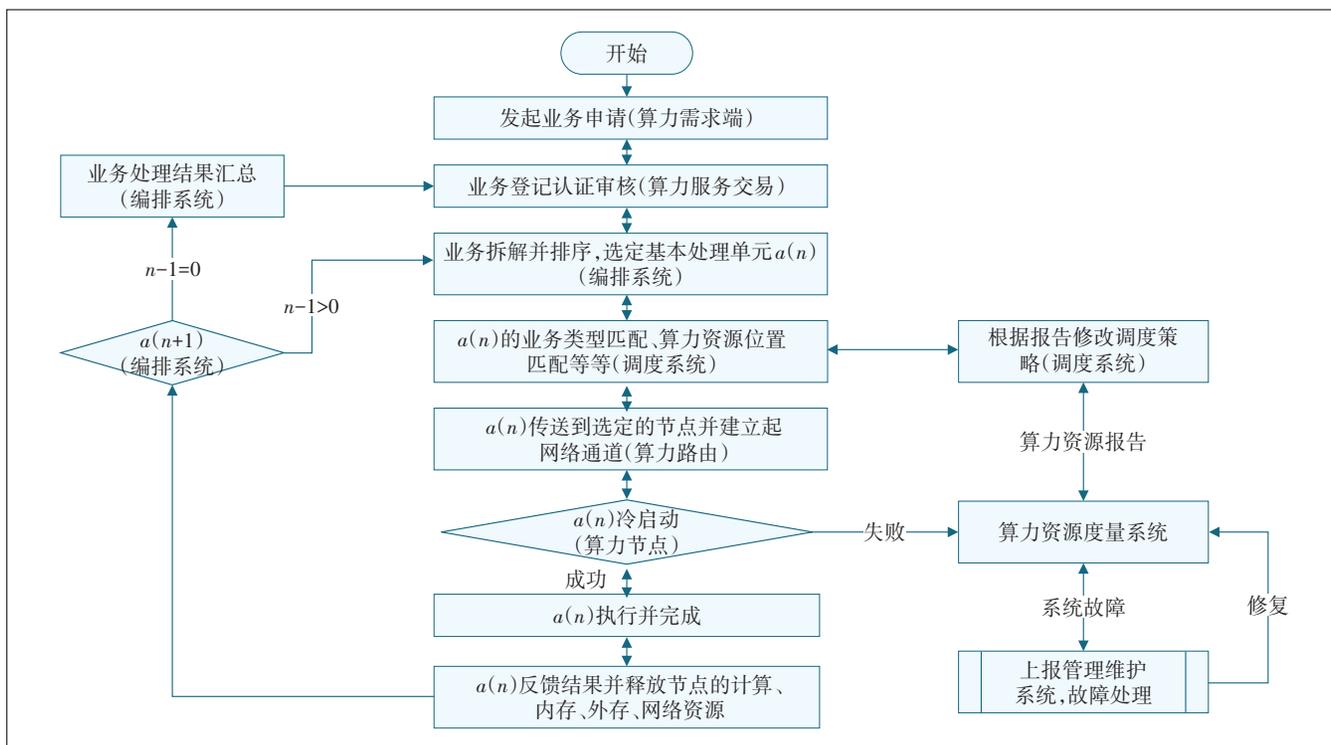


图4 算力服务实现流程

元的算力度量和管理方法,并提出了一套完善的算力度量与建模体系,为算力网络的节点资源提供了统一描述。

参考文献:

[1] 曹畅,唐雄燕. 算力网络关键技术及发展挑战分析[J]. 信息通信技术与政策, 2021, 47(3): 6-11.

[2] 唐雄燕,张帅,曹畅. 夯实云网融合,迈向算网一体[J]. 中兴通讯技术, 2021, 27(3): 42-46.

[3] 雷波,刘增义,王旭亮,等. 基于云、网、边融合的边缘计算新方案: 算力网络[J]. 电信科学, 2019, 35(9): 44-51.

[4] 段晓东,姚惠娟,付月霞,等. 面向算网一体化演进的算力网络技术[J]. 电信科学, 2021, 37(10): 76-85.

[5] 吕廷杰,刘峰. 数字经济背景下的算力网络研究[J]. 北京交通大学学报(社会科学版), 2021, 20(1): 11-18.

[6] 中国通信学会. 算力网络前沿报告[EB/OL]. [2024-03-22]. <https://www.ccsa.org.cn/>.

[7] 乔楚. 算力度量与算网资源调度思路分析[J]. 通信技术, 2022, 55(9): 1165-1170.

[8] 何涛,杨振东,曹畅,等. 算力网络发展中的若干关键技术问题分析[J]. 电信科学, 2022, 38(6): 62-70.

[9] 曹畅,张帅,刘莹,等. 云网向算网演进中的若干关键技术问题[J]. 电信科学, 2021, 37(10): 93-101.

[10] 姚惠娟,陆璐,段晓东. 算力感知网络架构与关键技术[J]. 中兴通讯技术, 2021, 27(3): 7-11.

[11] MCMAHON F H. The livermore fortran kernels: a computer test of

the numerical performance range [R/OL]. [2024-03-22]. <https://www.osti.gov/biblio/6574702>.

[12] 郭亮,吴美希,王峰,等. 数据中心算力评估:现状与机遇[J]. 信息通信技术与政策, 2021, 47(2): 79-86.

[13] 杜宗鹏,李志强,陆璐. 算力网络四面三级算力度量技术体系[J]. 中兴通讯技术, 2023, 29(4): 8-13.

[14] 李建飞,曹畅,李奥,等. 算力网络中面向业务体验的算力建模[J]. 中兴通讯技术, 2020, 26(5): 34-38+52.

[15] 中国通信标准化协会. 面向业务体验的算力需求量化与建模研究[EB/OL]. [2024-03-22]. <https://www.ccsa.org.cn/>.

[16] 李一男,唐琴琴,彭开来,等. 以服务为中心的算力网络度量与建模研究[J]. 信息通信技术与政策, 2023, 49(5): 21-29.

[17] Zhang Y, Cao C, Tang X, et al. Programmable Service System Based on SIDaaS in Computing Power Network [C]//2022 5th International Conference on Hot Information-Centric Networking (HotICN), Guangzhou, China, 2022:67-71.

作者简介:

王施霖,研究员,硕士,主要从事算力网络、未来互联网架构等研究工作;张岩,高级工程师,博士,主要从事算力网络、云网融合/云计算、未来网络体系架构等研究工作;李传宝,高级工程师,主要从事云计算/智算光电混合组网、算力网络、多模态智慧网络等技术研究工作;崔童,高级工程师,硕士,主要从事云计算、AI计算相关方面工作;曹畅,高级工程师,博士,主要从事算力网络、IPv6+网络新技术、未来网络体系架构等研究工作;唐雄燕,中国联通研究院副院长,首席科学家,“新世纪百千万人才工程”国家级人选,主要专业领域为宽带通信、光纤传输、互联网/物联网、SDN/NFV与新一代网络等。