

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024 年 9 月 26 日 (26.09.2024)



(10) 国际公布号
WO 2024/192584 A1

(51) 国际专利分类号:
G06F 9/50 (2006.01) **H04W 4/50** (2018.01)

(21) 国际申请号: PCT/CN2023/082299

(22) 国际申请日: 2023 年 3 月 17 日 (17.03.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 北京小米移动软件有限公司 (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) [CN/CN]; 中国北京市海淀区西二旗中路 33 号院 6 号楼 8 层 018 号, Beijing 100085 (CN)。

(72) 发明人: 陈栋 (CHEN, Dong); 中国北京市海淀区西二旗中路 33 号院 6 号楼 8 层 018 号, Beijing 100085 (CN)。何智斌 (HE, Zhibin); 中国北京市海淀区西二旗中路 33 号院 6 号楼 8 层 018 号, Beijing 100085 (CN)。

(74) 代理人: 北京法胜知识产权代理有限公司 (FASHENG INTELLECTUAL PROPERTY

COMPANY, LTD.); 中国北京市海淀区北洼路 45 号 1 号楼 2 层 202 室, Beijing 100142 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,

(54) Title: CALLING METHOD AND APPARATUS FOR COMPUTING POWER SERVICE, AND STORAGE MEDIUM

(54) 发明名称: 算力服务调用方法、装置及存储介质

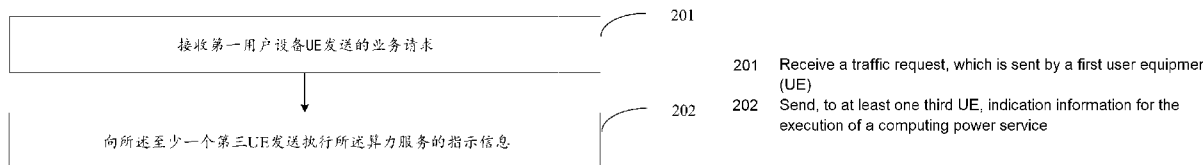


图 2

(57) Abstract: The present disclosure belongs to the technical field of communications. Provided are a calling method and apparatus for a computing power service, and a storage medium. The method comprises: receiving a traffic request, which is sent by a first user equipment (UE), wherein the traffic request is used for requesting the calling of a computing power service; sending, to at least one third UE, indication information for the execution of the computing power service, wherein the at least one third UE is a UE that meets requirements corresponding to the computing power service; the first UE sending the traffic request to a computing service module; the third UE receiving indication information for the execution of the computing power service, which is sent by the computing service module; and a first network element sending computing power registration information of at least one second UE to the computing service module. Therefore, UEs can collaborate with a network to carry out distributed computing, thereby providing a computing power service.

(57) 摘要: 本公开提出一种算力服务调用方法、装置、及存储介质, 属于通信技术领域。该方法包括: 接收第一用户设备 UE 发送的业务请求, 其中, 业务请求用于请求调用算力服务; 向至少一个第三 UE 发送执行算力服务的指示信息, 其中, 至少一个第三 UE 为满足算力服务对应需求的 UE; 第一 UE 向计算服务模块发送业务请求; 第三 UE 接收计算服务模块发送的执行算力服务的指示信息; 第一网元向计算服务模块发送至少一个第二 UE 的算力登记信息, 从而实现 UE 协同网络进行分布式计算, 提供算力服务。

CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

算力服务调用方法、装置及存储介质

技术领域

[0001] 本公开涉及通信技术领域，尤其涉及算力服务调用方法、装置及存储介质。

背景技术

[0002] 在未来网络系统中，用户设备UE有望能够向网络提供计算能力，终端设备将通过基于业务的接口实现灵活、直接的通信，海量终端设备将自身的算力能力登记存储于网络，协同网络进行分布式计算，减轻了网络承载的任务压力，提高了空闲终端设备的资源利用效率。相关技术中对于用户设备如何协同网络提供算力服务尚不明确。

[0003] 发明内容

[0004] 本公开提出的算力服务调用方法、装置及存储介质，用于终端设备协同网络提供算力服务。

[0005] 第一方面，本公开实施例提供一种算力服务调用方法，该方法由网络设备执行，包括：接收第一用户设备UE发送的业务请求，其中，业务请求用于请求调用算力服务；向至少一个第三UE发送执行算力服务的指示信息，其中，所述至少一个第三UE为满足所述算力服务对应需求的UE。

[0006] 第二方面，本公开实施例提供一种算力服务调用方法，该方法由第一UE执行，方法包括：向计算服务模块发送业务请求，其中，业务请求用于请求调用算力服务。

[0007] 第三方面，本公开实施例提供一种算力服务调用方法，该方法由第三UE执行，方法包括：接收计算服务模块发送的执行算力服务的指示信息。

[0008] 第四方面，本公开实施例提供一种算力服务调用方法，方法由第一网元执行，方法包括：向计算服务模块发送至少一个第二UE的算力登记信息，其中，至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE。

- [0009] 第五方面，本公开实施例提供一种计算服务模块，计算服务模块部署于核心网中，包括：
- [0010] 收发单元，用于接收第一用户设备UE发送的业务请求，其中，业务请求用于请求调用算力服务；
- [0011] 收发单元还用于：向至少一个第三UE发送执行算力服务的指示信息，其中，至少一个第三UE为满足算力服务对应需求的UE。
- [0012] 第六方面，本公开实施例提供一种第一用户设备UE，包括：
- [0013] 收发单元，用于向计算服务模块发送业务请求，其中，业务请求用于请求调用算力服务。
- [0014] 第七方面，本公开实施例提供一种第三用户设备UE，包括：
- [0015] 收发单元，用于接收计算服务模块发送的执行算力服务的指示信息。
- [0016] 第八方面，本公开实施例提供一种第一网元，包括：
- [0017] 收发单元，用于向计算服务模块发送至少一个第二UE的算力登记信息，
- [0018] 其中，至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE。
- [0019] 第九方面，本公开实施例提供一种通信装置，装置包括处理器和存储器，其中，存储器中存储有计算机程序，处理器执行存储器中存储的计算机程序，以使装置执行上述第一方面至第四方面任一方面的方法。
- [0020] 第十方面，本公开实施例提供一种通信装置，包括：处理器和接口电路，其中接口电路，用于接收代码指令并传输至处理器；处理器，用于运行代码指令以执行上述第一方面至第四方面任一方面的方法。
- [0021] 第十一方面，本发明实施例提供一种计算机可读存储介质，用于存储有指令，当指令被执行时，使上述第一方面至第四方面的任一方面的方法被实现。
- [0022] 第十二方面，本公开还提供一种包括计算机程序的计算机程序产品，当其在计算机上运行时，使得计算机执行上述第一方面至第四方面的任一方面的方法。
- [0023] 第十三方面，本公开提供一种通信系统，系统包括计算服务模块、第一用户设备UE、第三UE、第一网元，其中，
- [0024] 计算服务模块用于执行第一方面的方法；

- [0025] 第一UE用于执行第二方面的方法；
- [0026] 第三UE用于执行第三方面的方法；
- [0027] 第一网元用于执行第四方面的方法。
- [0028] 综上所述，根据本公开提供的算力服务调用方法，该方法由计算服务模块执行，计算服务模块部署于核心网中，方法包括：接收第一用户设备UE发送的业务请求，其中，业务请求用于请求调用算力服务；向至少一个第三UE发送执行算力服务的指示信息，其中，至少一个第三UE为满足算力服务对应需求的UE。本公开的方案实现了UE协同网络提供算力服务，通过将UE的算力能力登记存储于网络，协同网络进行分布式计算，减轻了网络承载的任务压力，实现了动态、灵活的算力调度和部署，提高了空闲终端设备的资源利用效率。

附图说明

- [0029] 本公开上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：
- [0030] 图1为本公开实施例提供的一种通信系统的架构示意图；
- [0031] 图2为本公开的一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0032] 图3为本公开另一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0033] 图4为本公开再一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0034] 图5为本公开又一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0035] 图6为本公开另一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0036] 图7为本公开再一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0037] 图8为本公开又一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0038] 图9为本公开另一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0039] 图10为本公开再一个实施例所提供的一种算力服务调用方法的流程示意图；
- [0040] 图11为本公开一个实施例提供的一种算力服务调用的交互方法的流程图；
- [0041] 图12为本公开一个实施例所提供的一种算力服务调用的交互方法的示例图；
- [0042] 图13为本公开一个实施例所提供的一种计算服务模块的结构示意图；
- [0043] 图14为本公开一个实施例所提供的一种计算服务模块的结构示意图；
- [0044] 图15为本公开一个实施例所提供的一种第一UE的结构示意图；

- [0045] 图16为本公开一个实施例所提供的一种第三UE的结构示意图；
- [0046] 图17为本公开一个实施例所提供的一种第三UE的结构示意图；
- [0047] 图18为本公开一个实施例所提供的一种第一网元的结构示意图；
- [0048] 图19为本公开一个实施例所提供的一种第一网元的结构示意图；
- [0049] 图20为本公开一个实施例所提供的一种通信装置的结构示意图；
- [0050] 图21为本公开一个实施例所提供的一种通信装置为芯片或芯片系统的结构示意图；
- [0051] 图22为本公开一个实施例所提供的一种通信系统的结构示意图。

具体实施方式

- [0052] 这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本公开实施例相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本公开实施例的一些方面相一致的装置和方法的例子。
- [0053] 在本公开实施例使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本公开实施例。在本公开实施例和所附权利要求书中所使用的单数形式的“一种”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。
- [0054] 应当理解，尽管在本公开实施例可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本公开实施例范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”及“若”可以被解释成为“在……时”或“当……时”或“响应于确定”。
- [0055] 下面详细描述本公开的实施例，实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的要素。下面通过参考附图描述的实施例是示例性的，旨在用于解释本公开，而不能理解为对本公开的限制。

- [0056] 为了便于理解，首先介绍本申请涉及的术语。
- [0057] 1、接入和移动性管理功能 (Access and Mobility management Function, AMF) 网元
- [0058] 支持具有不同移动性管理需求的终端设备。可以执行以下主要任务：非接入层 (NAS) 信令终端；NAS 信令安全；接入层安全控制；用于 3GPP 接入网之间移动性的核心网节点间信令；空闲模式终端设备可达性 (包括寻呼重传的控制和执行)；注册区管理；支持系统内和系统间的移动性；接入认证；访问授权，包括漫游权限检查；移动性管理控制 (订阅和策略)；支持网络切片；会话管理功能 (Session Management Function, SMF) 选择。
- [0059] 2、会话管理功能 (Session Management Function, SMF) 网元
- [0060] 可以与 AMF 一起支持定制的移动性管理方案，如 “仅限移动启动的连接” (Mobile Initiated Connection Only, MICO) 或 RAN 增强功能，如 “RRC Inactive” 状态。SMF 网元可以执行以下主要任务：会话管理；终端设备 IP 地址分配与管理；用户面功能 (User Plane Function, UPF) 的选择与控制；在 UPF 配置流量控制，将流量路由到适当的目的地；策略执行和服务质量 (Quality of Service, QoS) 控制部分；下行数据通知。
- [0061] 3、统一数据管理功能 (Unified Data Management, UDM) 网元
- [0062] UDM，用于用户标识、签约数据、鉴权数据的管理、以及用户的服务网元注册管理。
- [0063] 本公开实施例中涉及到的各种网元/功能，其既可以是一个独立的硬件设备，也可以是在硬件设备内的通过计算机代码实现的功能，本公开实施例中并不对此做出限定。
- [0064] 4、网络开放功能 (Network Exposure Function, NEF) 网元
- [0065] NEF 是 5GC 对外开放和提供标准接口的网络功能。基于 3GPP 网络功能通过 NEF 向其他系统公开功能和事件，既提供开放性，又能保障系统的安全。NEF 的存在，将 5GC 的功能实现标准化的展现，为第三方接入提供便利。
- [0066] 5、应用功能 (Application Function, AF) 网元

- [0067] AF类似于应用服务器，其与其他5G核心网控制面NF交互，并提供业务服务。AF可以针对不同的应用服务而存在，可以由运营商或可信的第三方拥有。
- [0068] 6、鉴权服务功能(Authentication Server Function, AUSF)网元
- [0069] AUSF是5G中支持统一鉴权服务功能的网元，实现3GPP和非3GPP的接入认证，是5G核心网络(5GC)中的网络实体。
- [0070] 7、策略控制功能(Policy Control Function, PCF)网元
- [0071] PCF支持统一的策略框架去管理网络行为，提供策略规则给网络实体去实施执行，访问统一数据仓库(UDR)的订阅信息。
- [0072] 8、统一数据存储(Unified Data Repository, UDR)网元
- [0073] UDR用于UDM存储订阅数据或读取订阅数据以及PCF存储策略数据或者读取策略数据。
- [0074] 9、网络存储功能(Network Repository Function, NRF)网元
- [0075] NRF支持服务发现功能，从NF实例接收NF发现请求，并将发现的NF实例(被发现)的信息提供给NF实例，同时可以维护可用NF实例及其支持的服务的NF配置文件。
- [0076] 10、用户面功能(User Plane Function, UPF)网元。
- [0077] UPF作为移动基础设施(Mobile Infrastructure, 例如：RAN)和数据网络DN(Data Network)之间的互连点，完成UP上GTP-U(GPRS隧道协议)协议的封装和解封装。通常，UPF用于无线接入技术(Radio Access Technologies, RAT)内部、或RAT之间移动性的协议数据单元(Protocol Data Unit, PDU)会话锚点(Session Anchor Point)，包括向gNB(NG-RAN节点)发送一个或多个结束标记数据包(End Marker Packets, EMP)。
- [0078] 为了更好的理解本公开实施例公开的方案，下面首先对本公开实施例适用的通信系统进行描述。
- [0079] 6G(第六代移动通信系统)服务化架构是将网元的功能抽象成多个服务。其中，相对于传统网络架构而言，6G服务化架构中引入了通感融合技术，以为终端设备提供多种服务(如感知服务、算力服务、人工智能(Artificial Intelligence, AI)服务等)。其中，6G服务化架构中包括用于提供服务的网

元，6G服务化架构中的其他网元通过与该提供服务的网元之间的协同与共享，实现多维感知、协作通信、智能计算功能的深度融合、互惠增强，进而使网络具备新型信息流智能交互与处理及广域智能协作的能力。

[0080] 请参见图1，图1为本公开实施例提供的一种6G服务化网络架构的示意图。

图1中包括6G服务化网络架构中的网元以及各个网元之间用于通信的接口。参考图1，虚线框用于表示提供服务的网元，如图1所示，该网元可以实现感知(Sensing)、AI计算(Calculation)等能力，并且还具备存储(Storage)能力。

[0081] 以及，图1所示的无线接入网(Radio Access Network, RAN)可以是网络侧的一种用于发射或接收信号的实体。例如，RAN可以为演进型基站(evolved NodeB, eNB)、发送接收点(transmission reception point, TRP)、NR系统中的下一代基站(next generation NodeB, gNB)、其他未来移动通信系统中的基站或无线保真(wireless fidelity, WiFi)系统中的接入节点等。本公开的实施例对网络设备所采用的具体技术和具体设备形态不做限定。本公开实施例提供的网络设备可以由集中单元(central unit, CU)与分布式单元(distributed unit, DU)组成的，其中，CU也可以称为控制单元(control unit)，采用CU-DU的结构可以将网络设备，例如基站的协议层拆分开，部分协议层的功能放在CU集中控制，剩下部分或全部协议层的功能分布在DU中，由CU集中控制DU。

[0082] 图1所示的用户设备(User Equipment, UE)(或也可称为终端设备)可以是用户侧的一种用于接收或发射信号的实体，如手机。终端设备也可以称为终端(terminal)、移动台(mobile station, MS)、移动终端设备(mobile terminal, MT)等。终端设备可以是具备通信功能的汽车、智能汽车、手机(mobile phone)、穿戴式设备、平板电脑(Pad)、带无线收发功能的电脑、虚拟现实(virtual reality, VR)终端设备、增强现实(augmented reality, AR)终端设备、工业控制(industrial control)中的无线终端设备、无人驾驶(self-driving)中的无线终端设备、远程手术(remote medical surgery)中的无线终端设备、智能电网(smart grid)中的无线终端设备、运输安全(transportation safety)中的无线终端设备、智慧城市(smart city)中的无线终端设备、智慧家庭(smart

home)中的无线终端设备等等。本公开的实施例对UE所采用的具体技术和具体设备形态不做限定。

[0083] 以及,图1中还包括有AMF、SMF、UDM等网元。其中,基于AMF网元和SMF网元的协作调度,可以使得终端设备(即图1中的UE)通过RAN与提供服务的网元建立用户面连接,以使得该提供服务的网元能够向终端设备提供该终端设备所请求调用的服务。关于该部分的具体流程可以参见后续实施例介绍。

[0084] 可以理解的是,本公开实施例描述的通信系统是为了更加清楚的说明本公开实施例的技术方案,并不构成对于本公开实施例提供的技术方案的限定,本领域普通技术人员可知,随着系统架构的演变和新业务场景的出现,本公开实施例提供的技术方案对于类似的技术问题,同样适用。

[0085] 下面参考附图对本公开实施例所提供的算力服务调用方法、装置、设备及存储介质进行详细描述。该方法可以应用于第五代移动通信技术(Fifth Generation, 5G)及其后续通信技术,诸如第五代移动通信技术演进(5G-advanced)、第六代移动通信技术(Sixth Generation, 6G)等,在本公开中不予限制。

[0086] 图2为本公开实施例所提供的一种算力服务调用方法的流程示意图,该方法由计算服务模块执行,计算服务模块部署于核心网中。在本公开中,计算服务模块可以以网元设备的形式部署于核心网中,以用于统筹调度和部署算力,其具体呈现形式在本公开中不予限制。

[0087] 应理解,本公开所提供的算力服务调用方法可以建立在下述前提下:

[0088] 1、UE(在本公开中称为第一UE)与核心网之间已建立用户面连接;

[0089] 2、可以为核心网提供算力支持的多个UE(在本公开中称为第二UE)已经向核心网中的算力存储网元(在本公开中称为第一网元)进行算力登记,经核心网中计算服务模块的部署,全部或部分第二UE(在本公开中称为第三UE)可以接收来自核心网分配的算力任务。

[0090] 如图2所示,该算力服务调用方法可以包括以下步骤:

[0091] 步骤201、接收第一用户设备UE发送的业务请求。

[0092] 在本公开的一些实施例中,业务请求用于请求调用算力服务。

- [0093] 在本公开的一些实施例中，业务请求包括以下至少一项：第一UE的标识；算力服务的标识；算力服务的类型；算力需求；网络带宽。
- [0094] 在本公开的一些实施例中，第一UE可以在需要算力服务时发送业务请求。例如，第一UE需要进行图像处理，则可以触发向核心网请求调用图像处理的算力服务。
- [0095] 在本公开的一些可选实施例中，UE的标识可以是用户永久标识符 (Subscription Permanent Identifier, SUPI)，或其他能够唯一标识UE的标识，在本公开中不予限制。算力服务的标识 (CalService ID) 可以是标记可被调用的算力服务的代号。算力服务的类型 (Cal Service Type) 可以是标记可被调用的算力服务的类型的代号或字段。算力需求为UE对当前待进行的计算任务的算力大小需求。网络带宽可以是下行网络带宽，以便于网络设备向UE返回符合带宽要求的计算结果的数据包。
- [0096] 在本公开中，计算服务模块可以通过用户面连接接收第一UE发送的业务请求。其中，第一UE与核心网之间的用户面连接可以通过如下方式建立的：当SMF网元接收到服务调用请求时，说明终端设备当前请求调用服务，此时，SMF网元可以与第一网元 (提供终端设备请求调用的服务的网元) 建立连接，以便SMF网元后续通过与第一网元交互来使得第一网元和终端设备建立用户面连接，进而使得第一网元能够通过用户面与终端设备进行通信，来为终端设备提供该终端设备所请求调用的服务。应理解，本公开对用户面连接的建立方式不予限制，以其他可能方式建立的用户面连接均落入本公开范围之内。通过用户面连接发送业务请求以进行算力服务调用，无需再通过控制面连接的信令交互流程，能够提升通信效率。
- [0097] 在本公开中，计算服务模块也可以通过其他连接接收第一UE发送的业务请求，对此本公开不予限制。
- [0098] 步骤202、向至少一个第三UE发送执行算力服务的指示信息。
- [0099] 在本公开的一些实施例中，指示信息可以是用于指示第三UE执行算力服务的信息。

- [0100] 在本公开的一些实施例中，指示信息包括执行算力服务所需的数据信息和/或任务需求。
- [0101] 计算服务模块可以将第一UE请求的计算任务整体下发给一个第三UE执行计算服务。
- [0102] 在本公开的一些实施例中，根据业务请求的类型不同，计算服务模块也可将UE请求的该计算业务分割成一个或多个子业务，即，进行分布式计算业务的拆分，基于分割后的子业务向确定的至少一个第三UE发送指示信息，将多个子业务分别分配给一个或多个第三UE进行计算，其中第三UE所能提供的算力大于且相近于子业务所需算力，多个第三UE的总算力能力可以大于或等于第一UE所请求的计算任务的算力需求，以为第一UE提供对该业务的合理算力服务，避免算力资源浪费。
- [0103] 综上所述，根据本公开提供的算力服务调用方法，该方法由计算服务模块执行，计算服务模块部署于核心网中，方法包括：接收第一用户设备UE发送的业务请求，其中，业务请求用于请求调用算力服务；从第一网元获取至少一个第二UE的算力登记信息；基于至少一个第二UE的算力登记信息，从至少一个第二UE中确定执行算力服务的至少一个第三UE；向至少一个第三UE发送执行算力服务的指示信息。本公开的方案实现了UE协同网络提供算力服务，通过将UE的算力能力登记存储于网络，协同网络进行分布式计算，减轻了网络承载的任务压力，实现了动态、灵活的算力调度和部署，提高了空闲终端设备的资源利用效率。
- [0104] 在一些可选实施例中，该方法还包括：从第一网元获取至少一个第二UE的算力登记信息。
- [0105] 在本公开的一些实施例中，该步骤可以在步骤201和步骤202之间执行。第一网元可以为算力存储网元(参见图1中的“storage”网元)，其中能够存储提供算力服务的算力节点的信息，在本公开中，UE可以协同核心网进行算力部署，第一网元中能够存储可以提供算力服务的UE的信息，还可以存储其他算力节点的信息，例如核心网中的算力节点的信息或核心网外的第三方算力节点的信息。

息，对此本公开不予限制。计算服务模块可以通过获取第一网元中的信息，统筹部署第一UE所请求执行的计算任务。

[0106] 在一些可能的实施方式中，第一网元可以是专门用于存储算力节点信息的网元，也可以是核心网中的通用或共享存储网元，用于存储算力节点信息以及其他信息，对此本公开不予限制。

[0107] 在一些可能的实施方式中，第一网元还可以部署于核心网之外，其存储的信息可以通过NEF网元与核心网进行数据交换，对此本公开不予限制。

[0108] 在一些可选实施例，第二UE可以为第一网元存储的、能够为核心网提供协同算力服务的全部或部分UE，其中，UE所能提供的算力大小、算力类型等算力信息登记在第一网元中，计算服务模块可以获取第一网元中存储的全部或部分第二UE的算力登记信息，以从中确定执行算力服务的一个或多个UE。

[0109] 在本公开的一些实施例，算力登记信息可以是UE在核心网(具体地，例如第一网元)中注册或登记的关于UE算力能力的信息。

[0110] 在本公开的一些实施例，算力登记信息至少包括每个第二UE的标识、算力类型、可用算力大小。

[0111] 在本公开的一些实施例，计算服务模块接收算力存储网元发送的可提供算力服务的一个或多个第二UE的算力登记信息，以确定第二UE中是否存在至少一个第三UE满足业务请求中的算力需求。

[0112] 在一些可选实施例，该方法还包括：基于至少一个第二UE的算力登记信息，从至少一个第二UE中确定执行算力服务的至少一个第三UE。

[0113] 在本公开的一些实施例，该步骤可以在步骤201和202之间执行。计算服务模块响应于接收到业务请求，可以对该业务请求进行业务感知，其中，业务感知为根据请求的类型不同，例如，语音处理或图像处理，对该请求所需算力大小、算力类型等需求进行的智能评估。

[0114] 在本公开的一些实施例，计算服务模块基于此业务感知以及登记在第一网元中的各UE的算力信息，从上述至少一个第二UE中确定执行算力服务的至少一个第三UE。

- [0115] 在本公开中计算服务模块可以使用预设算法来确定至少一个第三UE，该算法的类型在本公开中不予限制。
- [0116] 图3为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由计算服务模块执行，如图3所示，该算力服务调用方法可以包括以下步骤：
- [0117] 步骤301、接收第一用户设备UE发送的业务请求。
- [0118] 其中，业务请求用于请求调用算力服务。
- [0119] 在本公开中，步骤301的原理与图2所示实施例中的步骤201相同，可以参照图2相关描述，在此不再赘述。
- [0120] 步骤302、向第一网元发送获取请求。
- [0121] 其中，获取请求用于请求UE的算力登记信息。
- [0122] 在本公开的一些实施例中，获取请求中可以包括执行算力服务所需的算力大小，以请求第一网元查找能够提供该算力大小的UE的算力登记信息；或者，获取请求中可以包括算力类型，以请求第一网元查找能够提供该类型服务的UE的算力登记信息；或者，获取请求中可以包括UE的状态，以请求第一网元查找处于特定状态的UE(例如，处于空闲态的UE)的算力登记信息。
- [0123] 在本公开的一些实施例中，算力存储网元中存储有一个或多个UE(即上述第二UE)的算力登记信息，各UE的算力资源可以相同或不同，即可提供的算力大小、算力类型等相同或不同。
- [0124] 在本公开的一些实施例中，第一网元中还可以存储各个UE的状态，例如空闲态或占用态。已经被调用的UE在第一网元中存储的状态为占用态，在其完成当前算力服务前可以不再被调用；尚未被调用的UE在算力存储网元中存储的状态为空闲态，可为上述业务提供算力服务。
- [0125] 在一些可能的实施方式中，第一网元中还可以存储各个UE的剩余算力。例如，某一UE总算力能力为X，该UE被调用执行算力服务A，执行算力服务A所需算力大小为Y，第一网元可以存储该UE的剩余算力 $X-Y$ ，并当需要执行算力服务B且算力服务B所需算力大小小于 $X-Y$ 时，计算服务模块可以调用该UE执行算力服务B。
- [0126] 步骤303、接收第一网元发送的至少一个第二UE的算力登记信息。

- [0127] 在本公开的一些实施例中，在本公开的一些实施例中，算力登记信息可以是UE在核心网中(具体地，例如第一网元)注册或登记的关于UE算力能力的信息。
- [0128] 在本公开的一些实施例中，算力登记信息至少包括每个第二UE的标识、算力类型、可用算力大小。
- [0129] 在本公开的一些实施例中，计算服务模块接收算力存储网元发送的可提供算力服务的一个或多个第二UE的算力登记信息，以根据业务感知，确定第二UE中是否存在至少一个第三UE满足业务请求中的算力需求。
- [0130] 步骤304、从至少一个第二UE中确定是否存在至少一个第三UE。
- [0131] 在本公开的实施例中，至少一个第三UE的总算力资源满足业务请求中的算力需求。
- [0132] 在本公开的实施例中，计算服务模块可以基于业务请求和至少一个第二UE的算力登记信息，利用资源匹配算法，从第二UE中确定是否存在第三UE。例如，计算服务模块可以通过业务感知提取出业务请求中的相关信息，例如第一UE的标识、算力服务标识、算力服务类型、算力需求等等，确定第一UE请求执行的计算任务需要的算力大小、类型等信息，并通过从第一网元中获取的第二UE的算力登记信息，确定是否存在能够提供算力服务的第三UE，即，确定是否存在满足业务请求中算力需求的至少一个第三UE。
- [0133] 例如，若第一UE请求执行的计算任务的计算量较小，则其需要的算力也较小，此时单独的一个第三UE可以单独完成该算力服务。若第一UE请求执行的计算任务的计算量较大，则其需要的算力也较大，此时单独的一个第三UE可能不能完全完成该算力服务，则需要多个第三UE分布式的计算支持。计算服务模块通过特定算法衡量资源的利用率，从而确定一个或多个第三UE，以提供最优的算力部署决策。
- [0134] 步骤305、确定至少一个第二UE中不存在至少一个第三UE，向第一UE发送反馈信息。
- [0135] 其中，反馈信息用于通知第一UE调用失败。例如，反馈信息中可以携带调用失败字段，该字段表示当前没有可用的算力节点为其提供算力服务，此次调用失败，第一UE可以放弃该次算力服务的调用。

- [0136] 或者，反馈信息可以用于通知第一UE等待满足第一UE所需算力需求的UE。例如，反馈信息中可以携带等待字段，该字段表示当前没有可用的算力节点为其提供算力服务，第一UE可以等待可以为其提供算力服务的其他算力节点。
- [0137] 在本公开的实施例，若计算服务模块确定第二UE中没有能够为第一UE提供算力服务的第三UE，即，第二UE对于经拆分的任一个子业务均不存在满足算力需求的第三UE，则计算服务模块无法为第一UE请求的计算任务调度第三UE，计算服务模块可以向第一UE发送反馈信息，以通知第一UE调用失败，则第一UE可以决定是否放弃当前计算任务，或等待，或更换其他计算任务。
- [0138] 在一些可选实施例，计算服务模块也可以通过反馈信息直接通知第一UE等待，第一UE可以不放弃当前计算任务，而是等待能够为第一UE提供服务的、符合当前计算任务的算力需求的第三UE。
- [0139] 步骤306、确定至少一个第二UE中存在至少一个第三UE，向至少一个第三UE发送执行算力服务的指示信息。
- [0140] 其中，指示信息包括执行算力服务所需的数据信息和/或任务需求。
- [0141] 在本公开的实施例，若计算服务模块确定第二UE中存在至少一个能够为第一UE提供算力服务的第三UE，则计算服务模块将经拆分的子业务分别分配给第三UE，即，分别向第三UE发送指示信息，以指示第三UE执行算力服务。
- [0142] 指示信息中携带有执行该算力服务所需的数据信息和/或任务需求。例如，第一UE请求的算力服务是对100张图片进行图像处理，计算服务模块可以将该任务进行拆分，例如拆分为5个子任务，每个子任务下对20张图像进行处理，并在第二UE中确定执行每个子业务的第三UE，例如5个，计算服务模块可以将20张图像以及进行图像处理的业务需求下发给每个第三UE，以使各个第三UE执行对所接收到的图像的计算处理任务。
- [0143] 计算服务模块确定第二UE中存在多个UE的满足计算任务的需求，称该多个UE为第四UE；则 计算服务模块可以进一步从多个第四UE中选择一个或多个第三UE。其中，选择可以基于一个或多个维度，如利用率和/或计算效率和/或计算能力等。

- [0144] 在本公开的实施例中，可以向能够提供总体算力大小大于并最接近于第一UE请求算力服务所需算力的大小的一个或多个第三UE发送算力服务的指示信息，以避免终端设备资源的浪费。例如，第一UE请求的算力服务是对500张图像进行图像处理，此时第二UE中空闲态的第三UE有两个，一个可以对550张图像进行处理(利用率为 $5/5.5$)，一个可以对600张图像进行处理(利用率为 $5/6$)，此时可以向可以处理550张图像的第三UE发送算力服务的指示信息，以合理利用终端资源。
- [0145] 类似地，在本公开的实施例中，考虑到不同UE的计算效率和/或计算能力等，可以选择计算效率和/或计算能力高的UE作为第三UE。例如，第一UE请求的算力服务是对500张图像进行图像处理，此时第二UE中空闲态的第三UE有两个，一个UE的计算能力为1.5TFlops，另一个UE的计算能力为2TFlops，则选择计算能力为2TFlops的UE作为第三UE。例如，第一UE请求的算力服务是对500张图像进行图像处理，此时第二UE中空闲态的第三UE有两个，一个UE的计算效率为0.1568(单位计算能力/碳排放)，另一个UE的计算效率为0.4325(单位计算能力/碳排放)，则选择计算效率为0.4325的UE作为第三UE。
- [0146] 综上所述，根据本公开提供的算力服务调用方法，该方法由计算服务模块执行，包括：接收用户设备UE通过用户面连接发送的业务请求；向第一网元发送算力登记信息获取请求；接收第一网元发送的至少一个第二UE的算力登记信息；基于业务请求和至少一个第二UE的算力登记信息，利用资源匹配算法，从至少一个第二UE中确定是否存在至少一个第三UE，其中，至少一个第三UE的总算力资源满足业务请求中的算力需求；确定至少一个第二UE中不存在至少一个第三UE，向第一UE发送反馈信息；确定至少一个第二UE中存在至少一个第三UE，向至少一个第三UE发送执行算力服务的指示信息。本公开的方案实现了计算服务模块通过业务感知与算力资源信息相匹配，为业务请求分配所需第三UE，进而使第一UE向所分配的第三UE发送指示信息，以使计算服务模块高效、合理地调用第三UE，实现了灵活、准确的UE算力调度，提高了网络利用率和通信效率。

[0147] 图4为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由计算服务模块执行，基于图3所示的实施例，该算力服务调用方法可以包括以下步骤：

[0148] 步骤401、接收第一用户设备UE发送的业务请求。

[0149] 其中，业务请求用于请求调用算力服务。

[0150] 步骤402、从第一网元获取至少一个第二UE的算力登记信息。

[0151] 步骤403、基于业务请求和至少一个第二UE的算力登记信息，利用资源匹配算法，从至少一个第二UE中确定是否存在至少一个第三UE，其中，至少一个第三UE的总算力资源满足业务请求中的算力需求。

[0152] 在本公开中，步骤401-403的原理与上述图2-图3所示实施例中的相关步骤相同，可以参照上述实施例的相关描述，在此不再赘述。

[0153] 步骤404、确定至少一个第二UE中存在至少一个第三UE，向至少一个第三UE发送执行算力服务的指示信息。

[0154] 在本公开的实施例中，若计算服务模块确定第二UE中存在至少一个能够为第一UE提供算力服务的第三UE，则计算服务模块将经拆分的子业务分别分配给确定的第三UE，即，分别向确定的第三UE发送指示信息，以指示各个第三UE执行算力服务。

[0155] 在一种可选实施例中，步骤404可以包括：向第二网元发送指示信息。

[0156] 其中，指示信息用于指示第二网元向至少一个第三UE发送执行算力服务所需的数据信息和/或任务需求。

[0157] 在本公开的一些实施例中，第二网元可以是AMF网元。具体地，算力服务可能会部署在多个算力节点(例如，在本公开中通过UE协同执行算力服务时，UE可以作为算力节点)，计算服务模块可以向AMF网元发送指示信息，指示信息中还可以包括第三UE的标识，AMF接收到指示信息并将该指示信息发送至第三UE，以实现计算服务模块对第三UE的调用。

[0158] 指示信息中携带有执行该算力服务所需的数据信息和/或任务需求。例如，第一UE请求的算力服务是对100张图片进行图像处理，计算服务模块可以将该任务进行拆分，例如拆分为5个子任务，每个子任务下对20张图像进行处理，并确定

执行每个子业务的第三UE，例如5个，计算服务模块可以将20张图像以及进行图像处理的任务需求下发给每个第三UE，以使各个第三UE执行对所接收到的图像的计算处理任务。

[0159] 步骤405、接收至少一个第三UE反馈的计算结果。

[0160] 其中，计算结果为至少一个第三UE基于指示信息执行算力服务得到的结果。

[0161] 在本公开的实施例，第三UE依据接收到指示信息提供的的数据信息和/或任务需求进行任务处理，计算服务模块可以接收第二UE中确定的一个或多个第三UE完成所分配的算力服务后得到的计算结果。

[0162] 步骤406、对计算结果进行聚合处理，以得到聚合结果。

[0163] 在本公开的一些实施例，每个第三UE的任务分工可以不同，计算服务模块会将接收到的算力任务分成多个步骤，例如：数据清洗、模型训练等，在一种示例中，其中一个第三UE在完成数据清洗后，将该数据清洗后的数据返回至计算服务模块，计算服务模块再将该数据发送至另一个第三UE进行模型训练，待训练完成后将该数据返回至计算服务模块，计算服务模块可以对各第三UE的计算结果进行聚合处理，得到聚合结果。

[0164] 在本公开的一些实施例，各第三UE的任务分工也可以相同，各第三UE可以将计算结果返回计算服务模块进行汇总，完成对各第三UE计算结果的聚合处理，得到聚合结果。在该情形下，相较于各第三UE分工不同的情形，计算服务模块可以对各模块返回的计算结果进行汇总和打包，无需再对各计算结果进行二次处理。

[0165] 步骤407、将聚合结果发送至第一UE。

[0166] 在本公开的一些实施例，计算服务模块可以将聚合结果通过用户面连接反馈给第一UE。第一UE通过用户面与核心网建立连接，当计算服务模块接收到聚合结果，说明计算服务模块完成算力调用服务，通过计算服务模块和终端设备建立的用户面连接，将上述聚合结果发送至第一UE。与通过控制面建立连接不同，通过用户面建立连接，可以实现直接从用户面与各提供算力的第三UE直接建立连接，同时可以实现数据或参数的上传，而控制面默认不能进行数据包发送。

- [0167] 应理解，上述步骤406-407可以是可选步骤。换言之，计算服务模块可以直接将各计算结果发送至第一UE，不进行聚合处理，对此本公开不予限制。
- [0168] 步骤408、接收第一UE发送的确认信息。
- [0169] 在本公开的一些实施例中，第一UE依据接收到的计算服务模块发送的聚合结果，向计算服务模块发送确认信息，计算服务模块接收该确认信息。在一些实施例中，若聚合结果未达到预设要求，第一UE可能再次请求调用算力服务或进行其他方式修改，本公开对此不予限制。
- [0170] 在一些实施例中，该方法还包括：向第一网元发送状态更新请求。
- [0171] 在本公开的实施例中，状态更新请求可以用于请求更新第一网元中所存储的UE的状态。其中，该状态可以是占用态或空闲态。
- [0172] 应理解，该步骤可以是可选步骤，并且本公开不限制其执行顺序。在一种可能的实施方式中，该步骤可以在步骤404之后、步骤405之前执行，即，当计算服务模块将计算任务下发至第三UE时，计算服务模块可以向第一网元发送状态更新请求，将该第三UE的状态由空闲态变更为占用态。在另一种可能的实施方式中，该步骤可以在步骤405之后执行，即，当计算服务模块接收到第三UE反馈的计算结果时，计算服务模块可以向第一网元发送状态更新请求，将该第三UE的状态由占用态变更为空闲态。在本公开中不予限制。
- [0173] 在本公开的一些实施例中，状态更新请求可以用于请求第三UE的剩余算力。其中，第一网元中还可以存储各个UE的剩余算力。例如，某一UE总算力能力为 X ，该UE被调用执行算力服务A，执行算力服务A所需算力大小为 Y ，第一网元可以存储该UE的剩余算力 $X-Y$ ，并当需要执行算力服务B且算力服务B所需算力大小小于 $X-Y$ 时，计算服务模块可以调用该UE执行算力服务B。
- [0174] 在本公开的一些实施例中，计算服务模块完成业务分割及分配，将子业务发送至第三UE，可以通过向第一网元发送状态更新请求来变更被调用的第三UE的算力状态，将被调用的第三UE由空闲态变更为占用态。
- [0175] 在本公开的一些实施例中，计算服务模块接收到第三UE反馈的计算结果，表明第三UE的算力服务结束，向第一网元发送状态更新请求，变更第三UE的算力状态，将已完成算力服务的第三UE由占用态变更为空闲态。

[0176] 综上所述，根据本公开提供的算力服务调用方法，该方法由计算服务模块执行，包括：接收第一用户设备UE发送的业务请求，从第一网元获取至少一个第二UE的算力登记信息，基于业务请求和至少一个第二UE的算力登记信息，利用资源匹配算法，从至少一个第二UE中确定是否存在至少一个第三UE，其中，至少一个第三UE的总算力资源满足业务请求中的算力需求，确定至少一个第二UE中存在至少一个第三UE，向至少一个第三UE发送执行算力服务的指示信息，接收至少一个第三UE反馈的计算结果，对计算结果进行聚合处理，以得到聚合结果，将聚合结果发送至第一UE，接收第一UE发送的确认信息，向第一网元发送状态更新请求，向第二网元发送指示请求，从而实现了算力服务模块合理调用UE侧进行算力服务，有效减轻了网络承载的任务压力，提高了对算力任务的处理速度。

[0177] 图5为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由第一UE执行，如图5所示，该算力服务调用方法可以包括以下步骤：

[0178] 步骤501、向计算服务模块发送业务请求。

[0179] 其中，业务请求用于请求调用算力服务。

[0180] 在本公开的一些实施例中，业务请求包括以下至少一项：第一UE的标识；算力服务的标识；算力服务的类型；算力需求；网络带宽。

[0181] 在本公开的一些可选实施例中，第一UE的标识可以是用户永久标识符(SUBSCRIPTION PERMANENT IDENTIFIER, SUPI)，或其他能够唯一标识UE的标识，在本公开中不予限制。算力服务的标识(Cal Service ID)可以是标记可被调用的算力服务的代号。算力服务的类型(Cal Service Type)可以是标记可被调用的算力服务的类型的代号或字段。算力需求为UE对当前待进行的计算任务的算力大小需求。网络带宽可以是下行网络带宽，以便于网络设备向UE返回符合带宽要求的计算结果的数据包。

[0182] 在本公开的一些实施例中，第一UE与核心网之间的用户面连接可以通过如下方式建立的：当SMF网元接收到服务调用请求时，说明终端设备当前请求调用服务，此时，SMF网元可以与第一网元(提供终端设备请求调用的服务的网元)建立连接，以便SMF网元后续通过与第一网元交互来使得第一网元和终端设备建立

用户面连接，进而使得第一网元能够通过用户面与终端设备进行通信，来为终端设备提供该终端设备所请求调用的服务。应理解，本公开对用户面连接的建立方式不予限制，以其他可能方式建立的用户面连接均落入本公开范围之内。

[0183] 综上所述，根据本公开提供的算力服务调用方法，该方法由第一UE执行，包括：向计算服务模块发送业务请求，其中，业务请求用于请求调用算力服务。从而将用户设备第一UE同计算服务模块连接起来，为第一UE能够成功调用核心网算力服务提供了基础。

[0184] 图6为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由第一UE执行，如图7所示，在执行上述步骤601之后，该算力服务调用方法可以包括以下步骤：

[0185] 步骤601、接收计算服务模块发送的聚合结果。

[0186] 其中，聚合结果为计算服务模块对至少一个第三UE的计算结果进行聚合处理得到的结果。

[0187] 在本公开的一些实施例中，第一UE可以接收计算服务模块发送的聚合结果，以获取本次算力服务调用的结果。

[0188] 在本公开的一些实施例中，第一UE可以接收计算服务模块发送的不进行聚合处理的各计算结果，以获取本次算力服务调用的结果。

[0189] 步骤602、向计算服务模块发送确认信息。

[0190] 在本公开的一些实施例中，第一UE接收到计算服务模块的计算结果，发送确认信息，确认其收到该算力调用服务。

[0191] 综上所述，根据本公开提供的算力服务调用方法，该方法由第一UE执行，包括：接收计算服务模块发送的聚合结果，其中，聚合结果为计算服务模块对至少一个第三UE的计算结果进行聚合处理得到的结果，向计算服务模块发送确认信息，使UE能及时、准确收到算力服务调用结果，并反馈收到的算力服务，以及及时变更第三UE算力状态，提高了空闲终端设备的资源利用效率。

[0192] 在本公开的一些实施例中，第一UE可以接收计算服务模块发送的反馈信息，该反馈信息用于通知第一UE调用失败和/或等待满足算力需求的UE。第一UE响应于该反馈信息，确定是否放弃当前计算任务，或等待，或更换其他计算任务，或

者响应于该反馈信息，等待能够为第一UE提供服务的、符合当前计算任务的算力需求的第三UE。

[0193] 图7为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由第三UE执行。如图8所示，该算力服务调用方法可以包括以下步骤：

[0194] 步骤701、接收计算服务模块发送的执行算力服务的指示信息。

[0195] 在本公开的一些实施例中，指示信息包括执行算力服务所需的数据信息和/或任务需求。

[0196] 在本公开的一些实施例中，第三UE接收计算服务模块发送的执行算力服务的指示信息，通过计算服务模块的部署，各第三UE可以执行相应的计算任务，以提供算力服务。

[0197] 综上所述，根据本公开提供的算力服务调用方法，该方法由第三UE执行，包括：接收计算服务模块发送的执行算力服务的指示信息。从而使第三UE依据指示信息执行所分配的算力任务。

[0198] 图8为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由第三UE执行，如图8所示，在执行上述步骤701之后，该算力服务调用方法可以包括以下步骤：

[0199] 步骤801、基于指示信息，执行算力服务，以得到计算结果。

[0200] 在本公开的一些实施例中，第三UE基于指示信息，执行算力服务，例如数据清洗、模型训练等，得到计算结果。不同的第三UE可以执行不同的算力任务，也可以执行相同的算力任务，其取决于计算服务模块对第一UE所请求的计算任务的拆分和部署。

[0201] 步骤802、向计算服务模块发送计算结果。

[0202] 在本公开的一些实施例中，第三UE将上述计算结果发送至计算服务模块，通过计算服务模块将该结果发送至第一UE。

[0203] 此外，计算服务模块可以将各第三UE的计算结果直接反馈至第一UE，也可以将各第三UE的计算结果进行打包汇总后再反馈至第一UE，还可以将各第三UE的计算结果进行二次聚合处理后在反馈至第一UE，对此本公开不予限制。

- [0204] 综上所述，根据本公开提供的算力服务调用方法，该方法由第三UE执行，包括：基于指示信息，执行算力服务，以得到计算结果，向计算服务模块发送计算结果，从而实现将第三UE的计算结果发送至计算服务模块。
- [0205] 图9为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由第一网元执行，如图9所示，该算力服务调用方法可以包括以下步骤：
- [0206] 步骤901、向计算服务模块发送至少一个第二UE的算力登记信息。
- [0207] 其中，至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE。
- [0208] 在本公开的一些实施例中，第一网元可以为算力存储网元(参见图1中的“storage”网元)，其中能够存储提供算力服务的算力节点的信息，在本公开中，UE可以协同核心网进行算力部署，第一网元中能够存储可以提供算力服务的UE的信息，还可以存储其他算力节点的信息，例如核心网中的算力节点的信息或核心网外的第三方算力节点的信息，对此本公开不予限制。计算服务模块可以通过获取第一网元中的信息，统筹部署第一UE所请求执行的计算任务。
- [0209] 在一些可能的实施方式中，第一网元可以是专门用于存储算力节点信息的网元，也可以是核心网中的通用或共享存储网元，用于存储算力节点信息以及其他信息，对此本公开不予限制。
- [0210] 在一些可能的实施方式中，第一网元还可以部署于核心网之外，其存储的信息可以通过NEF网元与核心网进行数据交换，对此本公开不予限制。
- [0211] 在本公开的一些实施例中，第一网元将存储的至少一个第二UE的算力登记信息发送至计算服务模块，帮助计算服务模块根据业务感知结果和接收到的第二UE的算力资源信息，为该服务请求选择合适的第三UE进行算力服务部署。
- [0212] 综上所述，根据本公开提供的算力服务调用方法，该方法由第一网元执行，包括：向计算服务模块发送至少一个第二UE的算力登记信息，从而实现第一网元向算力服务模块发送第二UE的算力信息，帮助其完成业务分配。
- [0213] 图10为本公开实施例所提供的一种算力服务调用方法的流程示意图，该方法由第一网元执行，如图10所示，该算力服务调用方法可以包括以下步骤：

- [0214] 步骤1001、接收计算服务模块发送的算力登记信息获取请求。
- [0215] 在本公开的一些实施例中，第一网元接收计算服模块发送的算力登记信息获取请求，以确定能提供算力服务的第三UE。
- [0216] 步骤1002、响应于算力登记信息获取请求，将处于空闲态的UE的算力登记信息确定为至少一个第二UE的算力登记信息。
- [0217] 其中，算力登记信息至少包括每个第二UE的标识、算力类型、可用算力大小。
- [0218] 在本公开的一些实例中，第一网元将处于空闲态的UE的算力登记信息确定为至少一个第二UE的算力登记信息，该第二UE所能提供的总算力满足第一UE业务请求所需算力，且可以提供算力大小接近于所需算力需求的第二UE。
- [0219] 步骤1003、接收计算服务模块发送的状态更新请求。
- [0220] 在本公开的一些实施例中，第一网元接收计算服务模块向第三UE下发算力任务后发送的状态更新请求，请求将调用的第三UE由空闲态更新为占用态，表明该第三UE在提供算力服务，此期间可以不被调用服务其他算力业务。
- [0221] 在本公开的一些实施例中，第一网元接收计算服务模块接收第一UE发送的确认信息后发送的状态更新请求，请求将调用的第三UE由占用态更新为空闲态，表明该第三UE结束算力服务，可以为其他算力业务提供算力支持。
- [0222] 步骤1004、响应于状态更新请求，更新至少一个第三UE的算力状态，算力状态包括空闲态、占用态。
- [0223] 在本公开的一些实施例中，第一网元接收算力服务模块发送的状态更新请求，将被调用的第三UE由空闲态变更为占用态或将完成算力服务的第三UE由占用态变更为空闲态，及时变更第三UE状态以提高空闲终端设备的资源利用效率。
- [0224] 综上所述，根据本公开提供的算力服务调用方法，该方法由第一网元执行，包括：接收计算服务模块发送的算力登记信息获取请求，响应于算力登记信息获取请求，将处于空闲态的UE的算力登记信息确定为至少一个第二UE的算力登记信息，接收计算服务模块发送的状态更新请求，响应于状态更新请求，更新至少一个第三UE的算力状态，算力状态包括空闲态、占用态，通过及时变更第三UE状态，提高空闲终端设备的资源利用效率，减轻网络承载的任务压力。

- [0225] 图11为本公开一个实施例提供的一种算力服务调用的交互方法的流程图，该方法可以由通信系统执行，通信系统中包括第一UE、第三UE、计算服务模块、第一网元。如图11所示，该交互方法可以包括以下步骤：
- [0226] 1101、第一UE向计算服务模块发送业务请求。
- [0227] 在本公开的一些实施例中，业务请求用于请求调用算力服务。
- [0228] 1102、第一网元向计算服务模块发送至少一个第二UE的算力登记信息。
- [0229] 在本公开的一些实施例中，至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE。
- [0230] 1103、计算服务模块从至少一个第二UE中确定执行算力服务的至少一个第三UE。
- [0231] 1104、计算服务模块向至少一个第三UE发送执行算力服务的指示信息。
- [0232] 上述步骤1101-1104的原理与上述图2-10所示实施例各个步骤的原理相同，可以参照图2-10相关描述，在此不再赘述。
- [0233] 综上所述，根据本公开提供的算力服务调用方法，通过第一UE、第三UE、计算服务模块以及第一网元之间的交互，第一UE向计算服务模块发送业务请求，其中，业务请求用于请求调用算力服务；第一网元向计算服务模块发送至少一个第二UE的算力登记信息，至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE；计算服务模块从至少一个第二UE中确定执行算力服务的至少一个第三UE；计算服务模块向至少一个第三UE发送执行算力服务的指示信息，从而实现了一种完整的核心网中的算力服务工作流程，以使得终端设备协同网络提供算力服务，减轻了网络承载的任务压力，提高了空闲终端设备的资源利用效率。
- [0234] 作为一种示例，算力服务调用工作流程具体如图12所示。
- [0235] 其中，第一UE向核心网内的计算服务模块发送业务请求，该请求信息包含：UE标识符、算力服务的标识(CalService ID)、算力服务的类型(Cal Service Type)、算力大小需求、网络带宽等，如图示步骤1所示。
- [0236] 其中，计算服务模块收到业务请求信息并进行分析转译，完成业务感知，如图示步骤2所示。

- [0237] 其中，计算服务模块根据业务感知结果，如果需要调用UE侧算力协同网络完成业务，计算服务模块向第一网元Storage网元发送第二UE算力登记信息获取请求，Storage网元返回结果，如图示步骤3所示。
- [0238] 其中，计算服务模块根据收到的第二UE的算力登记信息，进行算力服务部署，如图示步骤4所示。
- [0239] 其中，计算服务模块将算力任务进行分配，将数据、任务需求和分配的第三UE的信息发送至接入与移动性管理功能AMF网元，如图示步骤5所示。
- [0240] 其中，算力服务可能会部署在多个算力节点(例如，在本公开中通过UE协同执行算力服务时，UE可以作为算力节点)，AMF将对应的数据、任务需求下发至不同的第三UE，如图示步骤6所示。
- [0241] 其中，计算服务模块向Storage网元发送更新第三UE状态信息，将分配任务的第三UE调整为占用状态，如图示步骤7所示。
- [0242] 其中，各第三UE将计算结果上传至计算服务模块，如图示步骤8所示。
- [0243] 其中，计算服务模块向Storage网元发送更新第三UE状态信息，将完成任务的第三UE调整为空闲状态，如图示步骤9所示。
- [0244] 其中，计算服务模块收到所有子任务计算结果后，对收到的计算结果进行聚合计算处理，如图示步骤10所示。
- [0245] 其中，计算服务模块通过用户面将最终计算结果返回请求业务的第一UE，如图示步骤11所示。
- [0246] 上述本申请提供的实施例中，分别从网络设备、用户设备的角度对本申请实施例提供的方法进行了介绍。为了实现上述本申请实施例提供的方法中的各功能，网络设备和用户设备可以包括硬件结构、软件模块，以硬件结构、软件模块、或硬件结构加软件模块的形式来实现上述各功能。上述各功能中的某个功能可以以硬件结构、软件模块、或者硬件结构加软件模块的方式来执行。
- [0247] 与上述几种实施例提供的网络设备管理方法相对应，本公开还提供一种算力服务调用装置，由于本公开实施例提供的算力服务调用装置与上述几种实施例提供的算力服务调用方法相对应，因此算力服务调用方法的实施方式也适用于本实施例提供的算力服务调用装置，在本实施例中不再详细描述。

- [0248] 图13为本公开实施例提供的一种计算服务模块1300的结构示意图，计算服务模块部署于核心网中。
- [0249] 如图13所示，该装置1300可以包括收发单元1301，接收第一用户设备UE发送的业务请求，其中，业务请求用于请求调用算力服务；确定单元1302，用于基于至少一个第二UE的算力登记信息，从至少一个第二UE中确定执行算力服务的至少一个第三UE。
- [0250] 在一些实施例中，收发单元1301还用于从第一网元获取至少一个第二UE的算力登记信息。
- [0251] 在一些实施例中，收发单元1301还用于向至少一个第三UE发送执行算力服务的指示信息。
- [0252] 在一些实施例中，确定单元1302还可以用于根据业务请求的类型不同，将第一UE请求的该计算业务分割成一个或多个子业务，即，进行分布式计算业务的拆分，基于分割后的子业务向确定的至少一个第三UE发送指示信息，将多个子业务分别分配给一个或多个第三UE进行计算，其中第三UE所能提供的算力大于且相近于子业务所需算力，多个第三UE的总算力能力可以大于或等于第一UE所请求的计算任务的算力需求，以为第一UE提供对该业务的合理算力服务，避免算力资源浪费。
- [0253] 综上所述，根据本公开实施例提供的计算服务模块，计算服务模块部署于核心网中；接收第一用户设备UE发送的业务请求；从第一网元获取至少一个第二UE的算力登记信息；从至少一个第二UE中确定执行算力服务的至少一个第三UE；向至少一个第三UE发送执行算力服务的指示信息；可以使得第三UE与计算服务模块建立连接，进而使得计算服务模块通过与第三UE通信来为第一UE提供该第一UE所请求调用的算力服务，从而实现UE协同网络提供算力服务。
- [0254] 在本公开的一些实施例中，业务请求包括以下至少一项第一UE的标识；算力服务的标识；算力服务的类型；算力需求；网络带宽。
- [0255] 在本公开的一些实施例中，收发单元1301还用于向第一网元发送算力登记信息获取请求，其中，算力登记信息获取请求用于请求处于空闲态的UE的算力登记信息。

- [0256] 在本公开的一些实施例中，收发单元1301还用于接收第一网元发送的至少一个第二UE的算力登记信息，算力登记信息至少包括每个第二UE的标识、算力类型、可用算力大小。
- [0257] 在本公开的一些实施例中，确定单元1302还用于基于业务请求和至少一个第二UE的算力登记信息，利用资源匹配算法，从至少一个第二UE中确定是否存在至少一个第三UE，其中，至少一个第三UE的总算力资源满足业务请求中的算力需求。
- [0258] 在本公开的一些实施例中，确定单元1302还用于确定至少一个第二UE中不存在至少一个第三UE。
- [0259] 在本公开的一些实施例中，收发单元1301还用于向第一UE发送反馈信息，其中，反馈信息用于通知第一UE调用失败和/或等待满足算力需求的UE。
- [0260] 在本公开的一些实施例中，确定单元1302还用于确定至少一个第二UE中存在至少一个第三UE，其中，指示信息包括执行算力服务所需的数据信息和/或任务需求。
- [0261] 在本公开的一些实施例中，收发单元1301还用于接收至少一个第三UE反馈的计算结果，其中，计算结果为至少一个第三UE基于指示信息执行算力服务得到的结果。
- [0262] 在本公开的一些实施例中，装置还包括聚合单元1303，用于：对计算结果进行聚合处理，得到聚合结果。
- [0263] 在本公开的一些实施例中，收发单元1301还用于将聚合结果发送至第一UE。
- [0264] 在本公开的一些实施例中，收发单元1301还用于接收第一UE发送的确认信息。
- [0265] 在本公开的一些实施例中，收发单元1301还用于向第一网元发送状态更新请求，其中，状态更新请求用于请求第一网元更新至少一个第三UE的算力状态，算力状态包括空闲态、占用态。
- [0266] 在本公开的一些实施例中，收发单元1301还用于向第二网元发送指示信息，其中，指示信息用于指示第二网元向至少一个第三UE发送执行算力服务所需的数据信息和/或任务需求。
- [0267] 图15为本公开实施例提供的一种第一用户设备UE1400的结构示意图。

- [0268] 如图15所示,该装置1400可以包括收发单元1401,用于向计算服务模块发送业务请求,其中,业务请求用于请求调用算力服务。
- [0269] 在本公开的一些实施例中,业务请求包括以下至少一项:第一UE的标识;算力服务的标识;算力服务的类型;算力需求;网络带宽。
- [0270] 综上,根据本公开实施例提供的算力服务调用装置,装置配置于第一UE:第一UE用于向计算服务模块发送业务请求,其中,业务请求用于请求调用算力服务,从而使第一UE与计算服务模块连接,以调用算力服务。
- [0271] 在本公开的一些实施例中,收发单元1401还用于接收计算服务模块发送的反馈信息,其中,反馈信息用于通知UE调用失败和/或等待满足算力需求的UE。
- [0272] 在本公开的一些实施例中,收发单元1401还用于接收计算服务模块发送的聚合结果,其中,聚合结果为计算服务模块对至少一个第三UE的计算结果进行聚合处理得到的结果。
- [0273] 在本公开的一些实施例中,收发单元1401还用于向计算服务模块发送确认信息。
- [0274] 图16为本公开实施例提供的一种第三用户设备UE1500的结构示意图。
- [0275] 如图16所示,该装置1500可以包括收发单元1501,用于接收计算服务模块发送的执行算力服务的指示信息。
- [0276] 在本公开的一些实施例中,指示信息包括执行算力服务所需的数据信息和/或任务需求。
- [0277] 综上,根据本公开实施例提供的算力服务调用装置,装置配置于第三UE:用于接收计算服务模块发送的执行算力服务的指示信息,其中,指示信息包括执行算力服务所需的数据信息和/或任务需求,从而使第三UE依据指示信息执行所分配的算力任务。
- [0278] 在本公开的一些实施例中,装置还包括处理单元1502,处理单元1502用于执行算力服务,得到计算结果。
- [0279] 在本公开的一些实施例中,收发单元1501还用于向计算服务模块发送计算结果。
- [0280] 图18为本公开实施例提供的一种第一网元1600的结构示意图。

- [0281] 如图18所示,该装置1600可以包括收发单元1601,用于向计算服务模块发送至少一个第二UE的算力登记信息,其中,至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE。
- [0282] 综上,根据本公开实施例提供的算力服务调用装置,装置配置于第一网元:用于向计算服务模块发送至少一个第二UE的算力登记信息,其中,至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE,从而确定提供算力服务的第三UE,实现第三UE协同网络提供算力服务。
- [0283] 在本公开的一些实施例中,收发单元1601还用于接收计算服务模块发送的算力登记信息获取请求。
- [0284] 在本公开的一些实施例中,装置还包括处理单元1602,处理单元1602用于将处于空闲态的UE的算力登记信息确定为至少一个第二UE的算力登记信息,其中,算力登记信息至少包括每个第二UE的标识、算力类型、可用算力大小。
- [0285] 在本公开的一些实施例中,收发单元1601还用于接收计算服务模块发送的状态更新请求。
- [0286] 在本公开的一些实施例中,处理单元1602还用于更新至少一个第三UE的算力状态,算力状态包括空闲态、占用态。
- [0287] 请参见图20,图20是本申请实施例提供的一种通信装置1700的结构示意图。通信装置1700可以是网络设备,也可以是终端设备,也可以是支持网络设备实现上述方法的芯片、芯片系统、或处理器等,还可以是支持终端设备实现上述方法的芯片、芯片系统、或处理器等。该装置可用于实现上述方法实施例中描述的方法,具体可以参见上述方法实施例中的说明。
- [0288] 通信装置1700可以包括一个或多个处理器1701。处理器1701可以是通用处理器或者专用处理器等。例如可以是基带处理器或中央处理器。基带处理器可以用于对通信协议以及通信数据进行处理,中央处理器可以用于对通信装置(如,基站、基带芯片,终端设备、终端设备芯片,DU或CU等)进行控制,执行计算机程序,处理计算机程序的数据。
- [0289] 可选的,通信装置1700中还可以包括一个或多个存储器1702,其上可以存有计算机程序1704,处理器1701执行计算机程序1704,以使得通信装置1700执行上

述方法实施例中描述的方法。可选的，存储器1702中还可以存储有数据。通信装置1700和存储器1702可以单独设置，也可以集成在一起。

[0290] 可选的，通信装置1700还可以包括收发器1705、天线1706。收发器1705可以称为收发单元、收发机、或收发电路等，用于实现收发功能。收发器1705可以包括接收器和发送器，接收器可以称为接收机或接收电路等，用于实现接收功能；发送器可以称为发送机或发送电路等，用于实现发送功能。

[0291] 可选的，通信装置1700中还可以包括一个或多个接口电路1707。接口电路1707用于接收代码指令并传输至处理器1701。处理器1701运行代码指令以使通信装置1700执行上述方法实施例中描述的方法。

[0292] 在一种实现方式中，处理器1701中可以包括用于实现接收和发送功能的收发器。例如该收发器可以是收发电路，或者是接口，或者是接口电路。用于实现接收和发送功能的收发电路、接口或接口电路可以是分开的，也可以集成在一起。上述收发电路、接口或接口电路可以用于代码/数据的读写，或者，上述收发电路、接口或接口电路可以用于信号的传输或传递。

[0293] 在一种实现方式中，处理器1701可以存有计算机程序1703，计算机程序1703在处理器1701上运行，可使得通信装置1700执行上述方法实施例中描述的方法。计算机程序1703可能固化在处理器1701中，该种情况下，处理器1701可能由硬件实现。

[0294] 在一种实现方式中，通信装置1700可以包括电路，电路可以实现前述方法实施例中发送或接收或者通信的功能。本申请中描述的处理器和收发器可实现在集成电路(integrated circuit, IC)、模拟IC、射频集成电路RFIC、混合信号IC、专用集成电路(application specific integrated circuit, ASIC)、印刷电路板(printed circuit board, PCB)、电子设备等上。该处理器和收发器也可以用各种IC工艺技术来制造，例如互补金属氧化物半导体(complementary metal oxide semiconductor, CMOS)、N型金属氧化物半导体(nMetal-oxide-semiconductor, NMOS)、P型金属氧化物半导体(positive channel metal oxide semiconductor, PMOS)、双极结型晶体管(bipolar junction transistor, BJT)、双极CMOS(Bi CMOS)、硅锗(SiGe)、砷化镓(GaAs)等。

[0295] 以上实施例描述中的通信装置可以是网络设备或者终端设备，但本申请中描述的通信装置的范围并不限于此，而且通信装置的结构可以不受图20的限制。

通信装置可以是独立的设备或者可以是较大设备的一部分。例如通信装置可以是：

[0296] (1)独立的集成电路IC，或芯片，或，芯片系统或子系统；

[0297] (2)具有一个或多个IC的集合，可选的，该IC集合也可以包括用于存储数据，计算机程序的存储部件；

[0298] (3)ASIC，例如调制解调器(Modem)；

[0299] (4)可嵌入在其他设备内的模块；

[0300] (5)接收机、终端设备、智能终端设备、蜂窝电话、无线设备、手持机、移动单元、车载设备、网络设备、云设备、人工智能设备等等；

[0301] (6)其他等等。

[0302] 对于通信装置可以是芯片或芯片系统的情况，可参见图21所示的芯片的结构示意图。图21所示的芯片包括处理器1801和接口1802。其中，处理器1801的数量可以是一个或多个，接口1802的数量可以是多个。

[0303] 可选的，芯片还包括存储器1803，存储器1803用于存储必要的计算机程序和数据。

[0304] 本领域技术人员还可以了解到本申请实施例列出的各种说明性逻辑块(illustrative logical block)和步骤(step)可以通过电子硬件、电脑软件，或两者的结合进行实现。这样的功能是通过硬件还是软件来实现取决于特定的应用和整个系统的设计要求。本领域技术人员可以对于每种特定的应用，可以使用各种方法实现的功能，但这种实现不应被理解为超出本申请实施例保护的范围。

[0305] 本申请还提供一种可读存储介质，其上存储有指令，该指令被计算机执行时实现上述任一方法实施例的功能。

[0306] 本申请还提供一种计算机程序产品，该计算机程序产品被计算机执行时实现上述任一方法实施例的功能。

[0307] 图22为本公开实施例所提供的一种通信系统的结构示意图，如图22所示，该通信系统包括：计算服务模块用于执行上述图2—图4所示方法；第一UE用于执行上述图5和6所示方法；第三UE用于执行上述图7和8所示方法；第一网元用于执行上述图9和10所示方法。

[0308] 在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。计算机程序产品包括一个或多个计算机程序。在计算机上加载和执行计算机程序时，全部或部分地产生按照本申请实施例的流程或功能。计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。计算机程序可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，计算机程序可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(digital subscriber line, DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。可用介质可以是磁性介质(例如，软盘、硬盘、磁带)、光介质(例如，高密度数字视频光盘(digital video disc, DVD))、或者半导体介质(例如，固态硬盘(solid state disk, SSD))等。

[0309] 本领域普通技术人员可以理解：本申请中涉及的第一、第二等各种数字编号仅为描述方便进行的区分，并不用来限制本申请实施例的范围，也表示先后顺序。

[0310] 本申请中的至少一个还可以描述为一个或多个，多个可以是两个、三个、四个或者更多个，本申请不做限制。在本申请实施例中，对于一种技术特征，通过“第一”、“第二”、“第三”、“A”、“B”、“C”和“D”等区分该种技术特征中的技术特征，该“第一”、“第二”、“第三”、“A”、“B”、“C”和“D”描述的技术特征间无先后顺序或者大小顺序。

- [0311] 本申请中各表所示的对应关系可以被配置，也可以是预定义的。各表中的信息的取值仅仅是举例，可以配置为其他值，本申请并不限定。在配置信息与各参数的对应关系时，并不一定要求必须配置各表中示意出的所有对应关系。例如，本申请中的表格中，某些行示出的对应关系也可以不配置。又例如，可以基于上述表格做适当的变形调整，例如，拆分，合并等等。上述各表中标题示出参数的名称也可以采用通信装置可理解的其他名称，其参数的取值或表示方式也可以通信装置可理解的其他取值或表示方式。上述各表在实现时，也可以采用其他的数据结构，例如可以采用数组、队列、容器、栈、线性表、指针、链表、树、图、结构体、类、堆、散列表或哈希表等。
- [0312] 本申请中的预定义可以理解为定义、预先定义、存储、预存储、预协商、预配置、固化、或预烧制。
- [0313] 本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。
- [0314] 所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。
- [0315] 以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以所述权利要求的保护范围为准。

权利要求书

- [权利要求 1] 一种算力服务调用方法，其特征在于，所述方法由网络设备执行，所述方法包括：
- 接收第一用户设备UE发送的业务请求，其中，所述业务请求用于请求调用算力服务；
- 向至少一个第三UE发送执行所述算力服务的指示信息，其中，所述至少一个第三UE为满足所述算力服务对应需求的UE。
- [权利要求 2] 根据权利要求1所述的方法，其特征在于，所述业务请求包括以下至少一项：
- 所述第一UE的标识；
- 所述算力服务的标识；
- 所述算力服务的类型；
- 算力需求；
- 网络带宽。
- [权利要求 3] 根据权利要求1或2所述的方法，其特征在于，所述方法还包括：
- 从第一网元获取至少一个第二UE的算力登记信息；
- 基于所述至少一个第二UE的算力登记信息，从所述至少一个第二UE中确定执行所述至少一个第三UE。
- [权利要求 4] 根据权利要求3所述的方法，其特征在于，所述从第一网元获取至少一个第二UE的算力登记信息包括：
- 向所述第一网元发送获取请求，其中，所述获取请求用于请求UE的算力登记信息；
- 接收所述第一网元发送的所述至少一个第二UE的算力登记信息，所述算力登记信息至少包括每个第二UE的标识、算力类型或可用算力大小。
- [权利要求 5] 根据权利要求3或4所述的方法，其特征在于，所述至少一个第二UE为空闲态的UE。

- [权利要求 6] 根据权利要求3至5中任一项所述的方法，其特征在于，所述基于所述至少一个第二UE的算力登记信息，从所述至少一个第二UE中确定所述至少一个第三UE包括：
基于所述业务请求和所述至少一个第二UE的算力登记信息，从所述至少一个第二UE中确定所述至少一个第三UE，
其中，所述至少一个第三UE的总算力资源满足所述业务请求中的算力需求。
- [权利要求 7] 根据权利要求1至6中任一项所述的方法，其特征在于，所述方法包括：
其中，所述指示信息包括执行所述算力服务所需的数据信息和/或任务需求。
- [权利要求 8] 根据权利要求1至7中任一项所述的方法，其特征在于，所述方法还包括：
接收所述至少一个第三UE反馈的计算结果，
其中，所述计算结果为所述至少一个第三UE基于所述指示信息执行所述算力服务得到的结果。
- [权利要求 9] 根据权利要求1至8中任一项所述的方法，其特征在于，所述方法还包括：
对所述计算结果进行聚合处理，得到聚合结果；
将所述聚合结果发送至所述第一UE；
接收所述第一UE发送的确认信息。
- [权利要求 10] 根据权利要求1至9中任一项所述的方法，其特征在于，所述方法还包括：
向所述第一网元发送状态更新请求，
其中，所述状态更新请求用于请求所述第一网元更新所述至少一个第三UE的算力状态，所述算力状态包括空闲态、占用态。
- [权利要求 11] 根据权利要求1至9中任一项所述的方法，其特征在于，所述方法还包括：

向所述第一网元发送状态更新请求，
其中，所述状态更新请求用于请求所述至少一个第三UE的剩余算力。

[权利要求 12] 根据权利要求1至11中任一项所述的方法，其特征在于，所述向所述至少一个第三UE发送 执行所述算力服务的指示信息包括：
向第二网元发送所述指示信息，
其中，所述指示信息用于指示向所述至少一个第三UE发送执行所述算力服务所需的数据信息和/或任务需求。

[权利要求 13] 根据权利要求1至6中任一项所述的方法，其特征在于，所述方法还包括：
基于所述业务请求和所述至少一个第二UE的算力登记信息确定所述至少一个第二UE中不存在所述至少一个第三UE；
向所述第一UE发送反馈信息；
其中，所述反馈信息用于通知所述第一UE调用失败和/或等待满足所述算力需求的UE。

[权利要求 14] 一种算力服务调用方法，其特征在于，所述方法由网络设备执行，所述方法包括：
接收第一用户设备UE发送的业务请求，其中，所述业务请求用于请求调用算力服务；
基于所述业务请求和所述至少一个第二UE的算力登记信息，从所述至少一个第二UE中确定所述至少一个第三UE，
确定是否存在至少一个第三UE，其中，所述至少一个第三UE为满足所述算力服务对应需求的UE。

[权利要求 15] 根据权利要求14所述的方法，其特征在于，所述方法还包括：
在确定不存在所述至少一个第三UE的情况下，向所述第一UE发送反馈信息；
其中，所述反馈信息用于通知所述第一UE调用失败和/或等待满足所述算力需求的UE。

- [权利要求 16] 根据权利要求14所述的方法，其特征在于，所述方法还包括：
在确定不存在所述至少一个第三UE的情况下，向所述至少一个第三UE发送执行所述算力服务的指示信息，
其中，所述指示信息包括执行所述算力服务所需的数据信息和/或任务需求。
- [权利要求 17] 根据权利要求14至16中任一项所述的方法，其特征在于，所述确定是否存在至少一个第三UE包括：
获取至少一个第二UE的算力登记信息；
基于所述业务请求和所述至少一个第二UE的算力登记信息，从所述至少一个第二UE中确定是否存在所述至少一个第三UE。
- [权利要求 18] 一种算力服务调用方法，其特征在于，所述方法由第一UE执行，所述方法包括：
向计算服务模块发送业务请求，其中，所述业务请求用于请求调用算力服务。
- [权利要求 19] 根据权利要求18所述的方法，其特征在于，所述业务请求包括以下至少一项：
所述第一UE的标识；
所述算力服务的标识；
所述算力服务的类型；
算力需求；
网络带宽。
- [权利要求 20] 根据权利要求19或20所述的方法，其特征在于，所述方法还包括：
接收所述计算服务模块发送的反馈信息，
其中，所述反馈信息用于通知所述第一UE调用失败和/或等待满足所述算力需求的UE。
- [权利要求 21] 根据权利要求19或20所述的方法，其特征在于，所述方法还包括：

接收所述计算服务模块发送的聚合结果，其中，所述聚合结果为所述计算服务模块对至少一个第三UE的计算结果进行聚合处理得到的结果；

向所述计算服务模块发送确认信息。

[权利要求 22] 一种算力服务调用方法，其特征在于，所述方法由第三UE执行，所述方法包括：

接收计算服务模块发送的执行算力服务的指示信息。

[权利要求 23] 根据权利要求22所述的方法，其特征在于，
所述指示信息包括执行所述算力服务所需的数据信息和/或任务需求。

[权利要求 24] 根据权利要求23所述的方法，其特征在于，所述方法还包括：
基于所述指示信息，执行所述算力服务，以得到计算结果；
向所述计算服务模块发送所述计算结果。

[权利要求 25] 一种算力服务调用方法，其特征在于，所述方法由第一网元执行，所述方法包括：
向计算服务模块发送至少一个第二UE的算力登记信息，
其中，所述至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE。

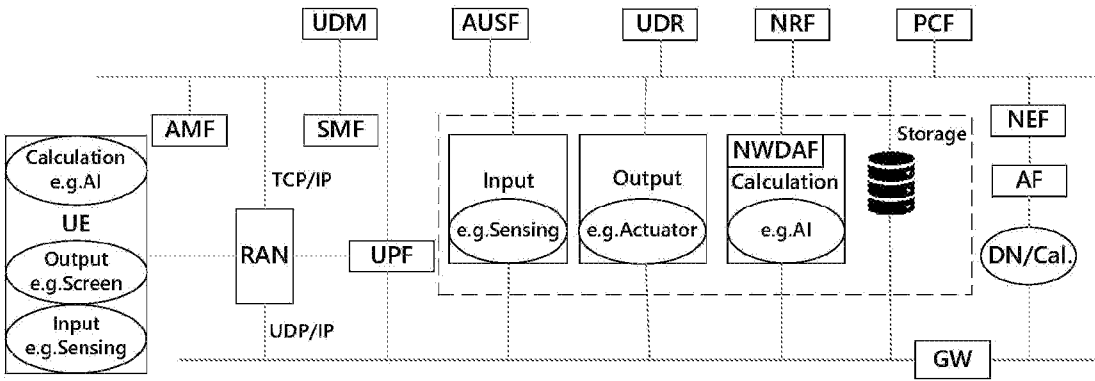
[权利要求 26] 根据权利要求25所述的方法，其特征在于，所述方法还包括：
接收所述计算服务模块发送的获取请求；
响应于所述获取请求，将处于空闲态的UE的算力登记信息确定为所述至少一个第二UE的算力登记信息，
其中，所述算力登记信息至少包括每个第二UE的标识、算力类型、可用算力大小。

[权利要求 27] 根据权利要求25或26所述的方法，其特征在于，所述方法包括：
接收所述计算服务模块发送的状态更新请求；
响应于所述状态更新请求，更新所述至少一个第三UE的算力状态，
所述算力状态包括空闲态、占用态。

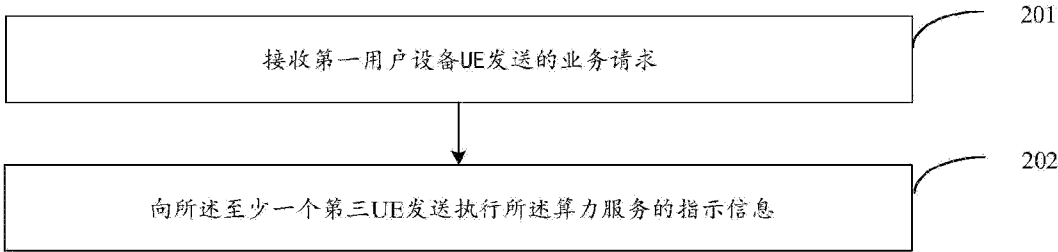
- [权利要求 28] 根据权利要求25或26所述的方法，其特征在于，所述方法包括：
接收所述计算服务模块发送的状态更新请求；
响应于所述状态更新请求，更新所述至少一个第三UE的剩余算力。
- [权利要求 29] 一种计算服务模块，其特征在于，所述计算服务模块部署于核心网中，包括：
收发单元，用于接收第一用户设备UE发送的业务请求，其中，所述业务请求用于请求调用算力服务；
向至少一个第三UE发送执行所述算力服务的指示信息，其中，所述至少一个第三UE为满足所述算力服务对应需求的UE。
- [权利要求 30] 一种第一用户设备UE，其特征在于，包括：
收发单元，用于向计算服务模块发送业务请求，其中，所述业务请求用于请求调用算力服务。
- [权利要求 31] 一种第三用户设备UE，其特征在于，包括：
收发单元，用于接收计算服务模块发送的执行算力服务的指示信息。
- [权利要求 32] 一种第一网元，其特征在于，包括：
收发单元，用于向计算服务模块发送至少一个第二UE的算力登记信息，
其中，所述至少一个第二UE的算力登记信息用于辅助计算服务模块确定执行算力服务的至少一个第三UE。
- [权利要求 33] 一种通信装置，其特征在于，所述装置包括处理器和存储器，其中，所述存储器中存储有计算机程序，所述处理器执行所述存储器中存储的计算机程序，以使所述装置执行：
如权利要求1至13中任一项所述的方法；或者
如权利要求14至17中任一项所述的方法；或者
如权利要求18至21中任一项所述的方法；或者
如权利要求22至24中任一项所述的方法；或者
如权利要求25至28中任一项所述的方法。

- [权利要求 34] 一种通信装置，其特征在于，包括：处理器和接口电路，其中所述接口电路，用于接收代码指令并传输至所述处理器；
所述处理器，用于运行所述代码指令以执行：
如权利要求1至13中任一项所述的方法；或者
如权利要求14至17中任一项所述的方法；或者
如权利要求18至21中任一项所述的方法；或者
如权利要求22至24中任一项所述的方法；或者
如权利要求25至28中任一项所述的方法。
- [权利要求 35] 一种计算机可读存储介质，用于存储有指令，当所述指令被执行时，使以下方法被实现：
如权利要求1至13中任一项所述的方法；或者
如权利要求14至17中任一项所述的方法；或者
如权利要求18至21中任一项所述的方法；或者
如权利要求22至24中任一项所述的方法；或者
如权利要求25至28中任一项所述的方法。
- [权利要求 36] 一种通信系统，其特征在于，所述系统包括计算服务模块、第一用户设备UE、第三UE、第一网元，其中，
所述计算服务模块用于执行如权利要求1至13中任一项所述的方法或14至17中任一项所述的方法；
所述第一UE用于执行如权利要求18至21中任一项所述的方法；
所述第三UE用于执行如权利要求22至24中任一项所述的方法；
所述第一网元用于执行如权利要求25至28中任一项所述的方法。

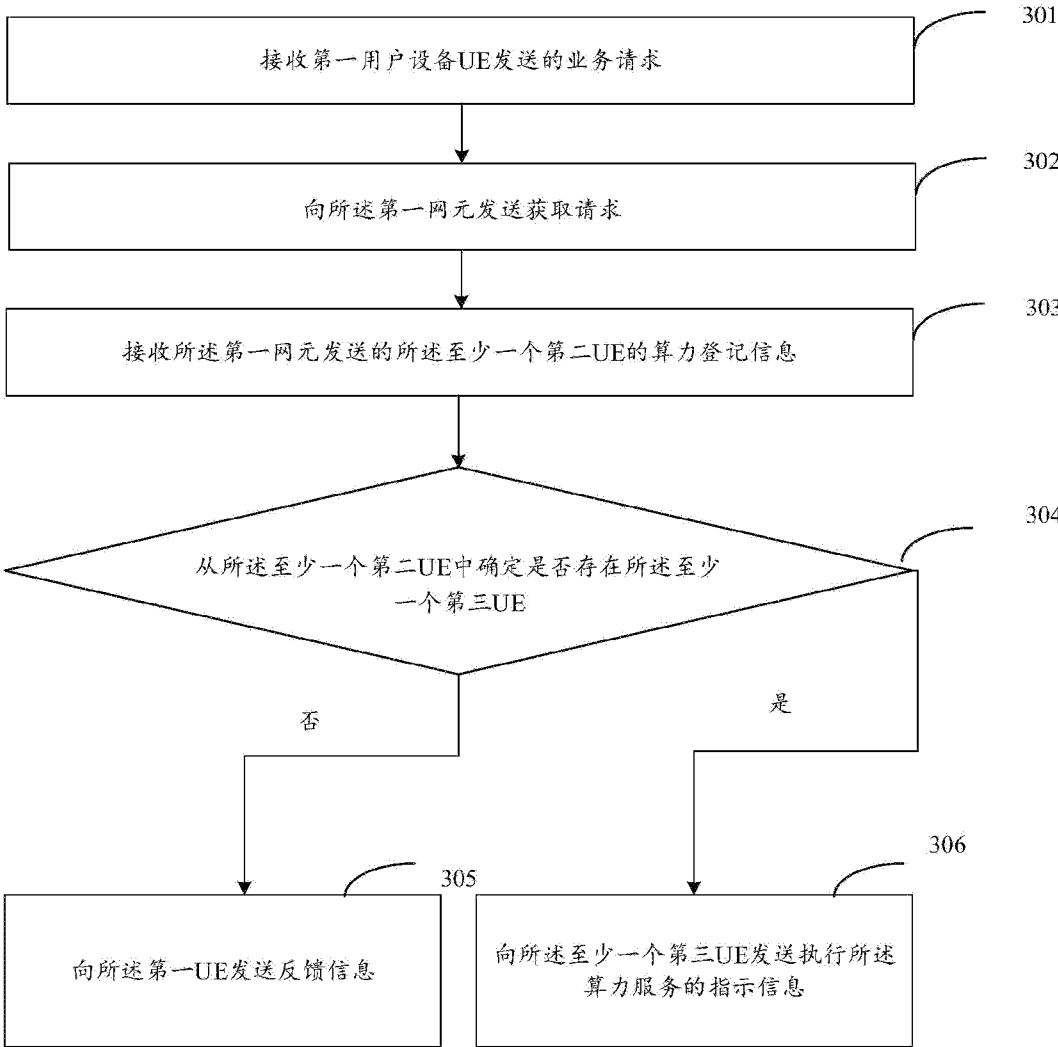
[图 1]



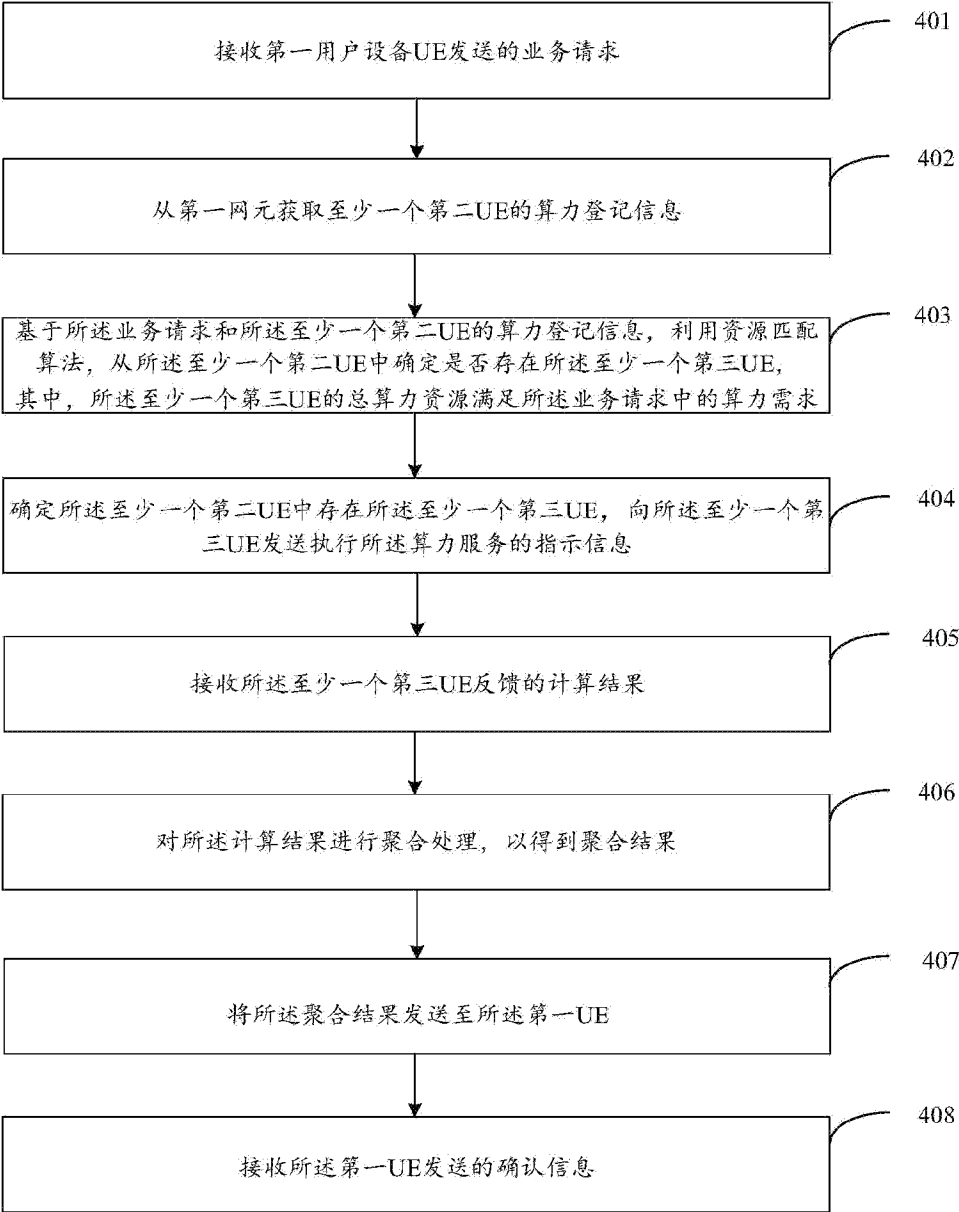
[图 2]



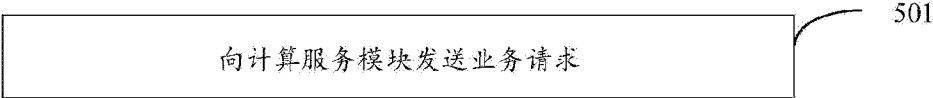
[图 3]



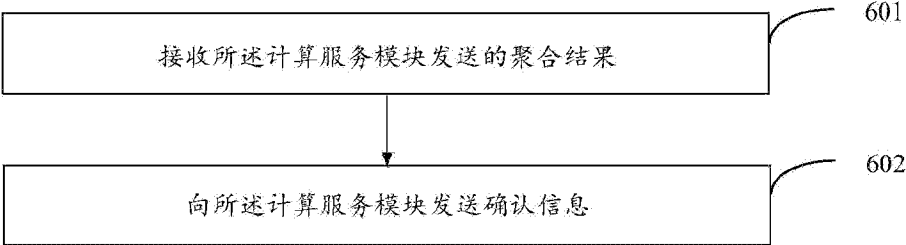
[图 4]



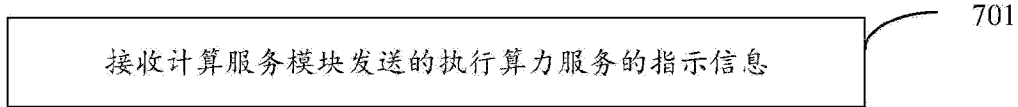
[图 5]



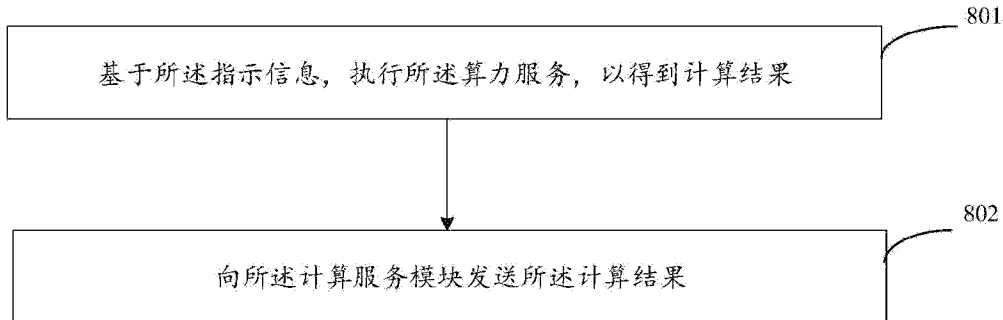
[图 6]



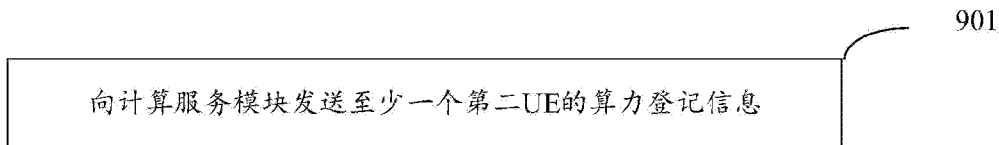
[图 7]



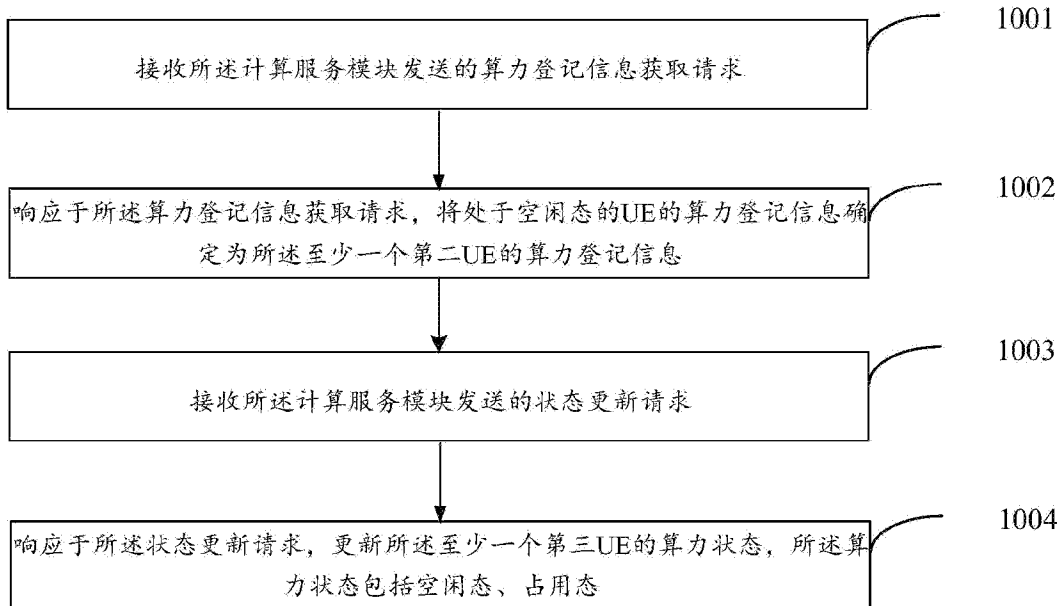
[图 8]



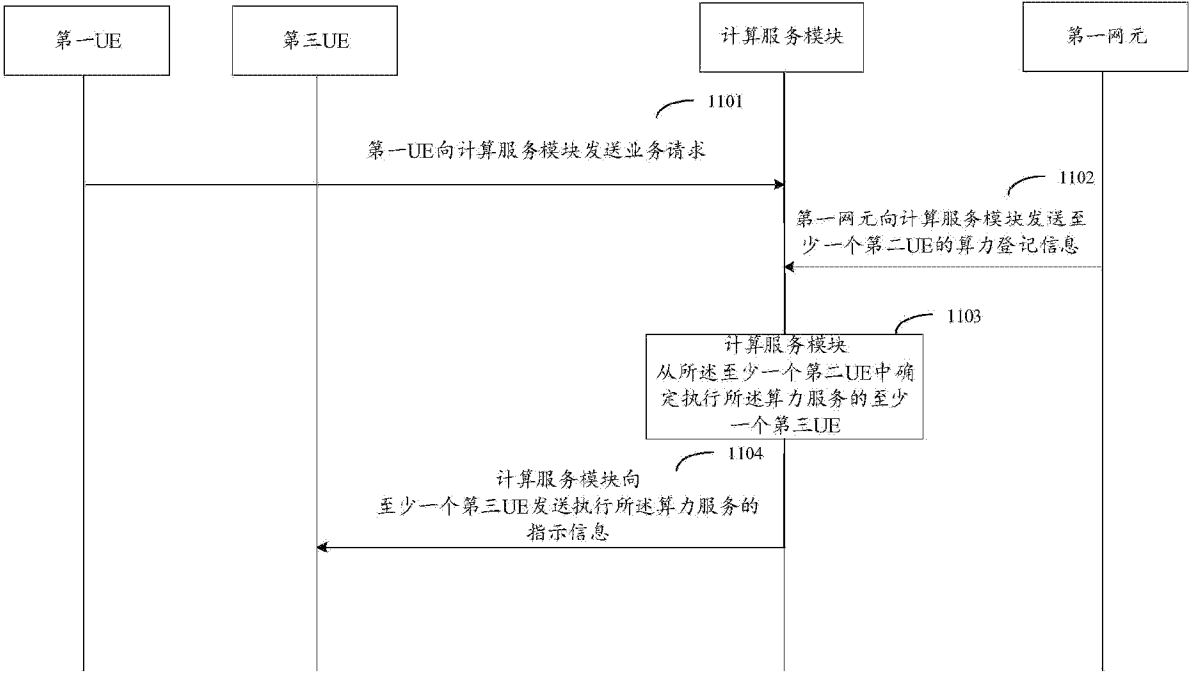
[图 9]



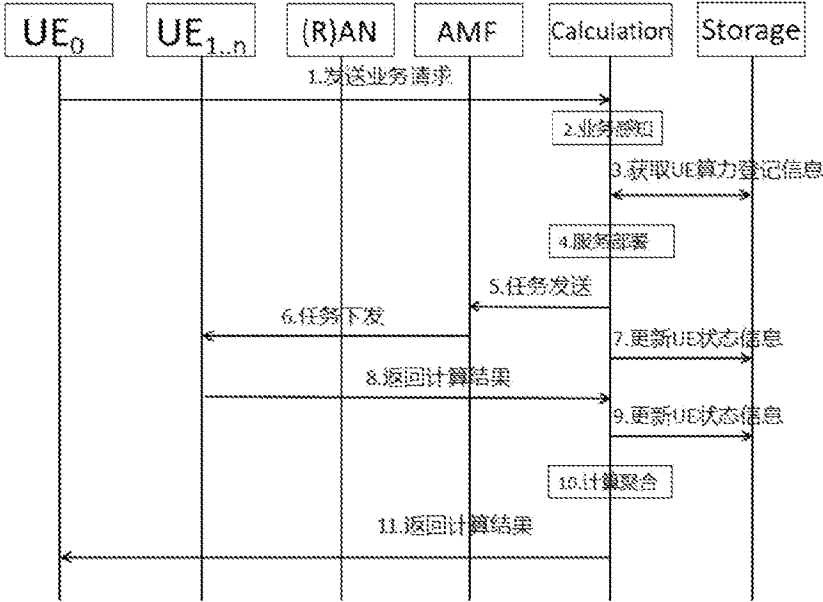
[图 10]



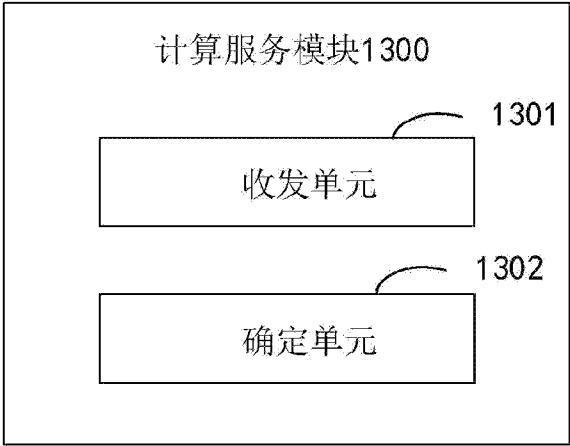
[图 11]



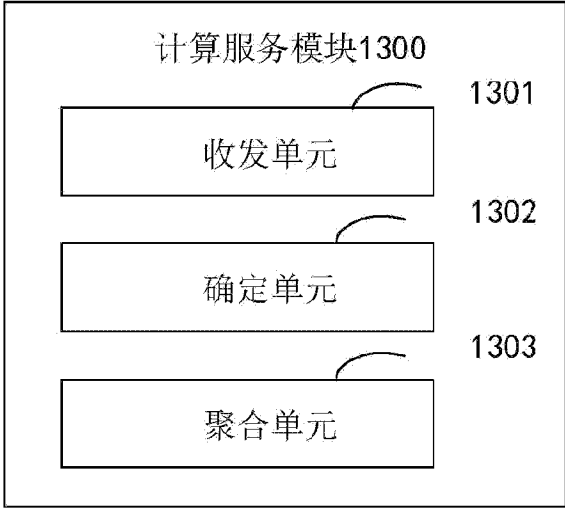
[图 12]



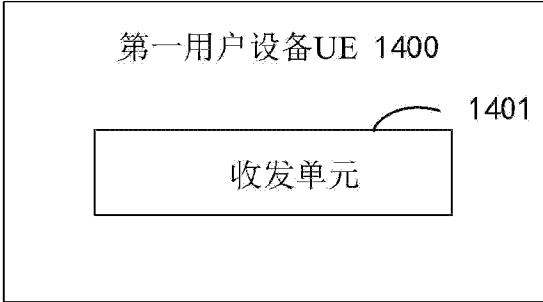
[图 13]



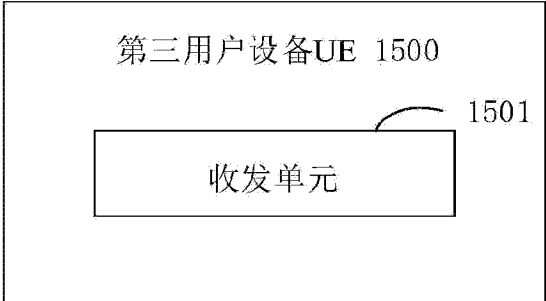
[图 14]



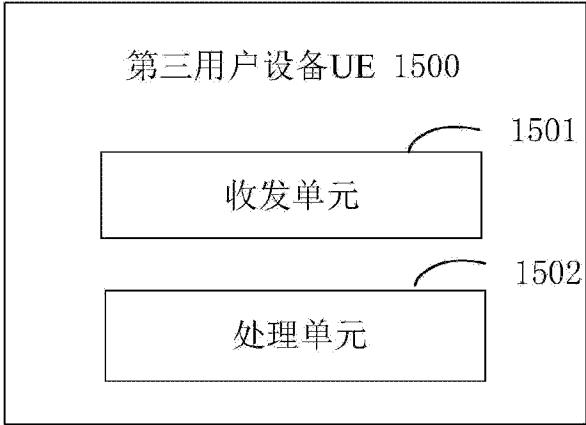
[图 15]



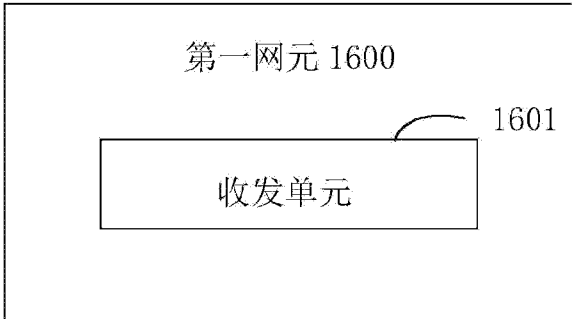
[图 16]



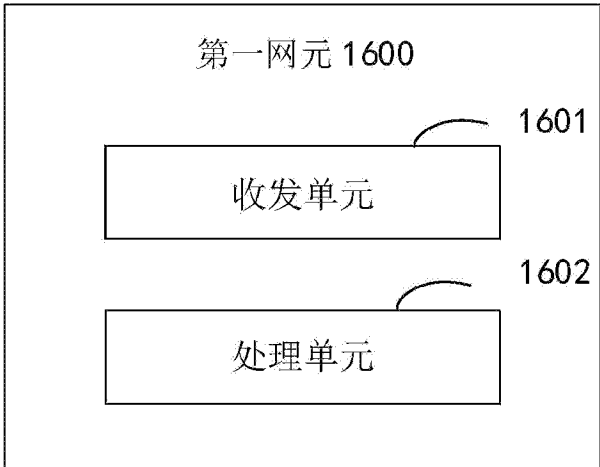
[图 17]



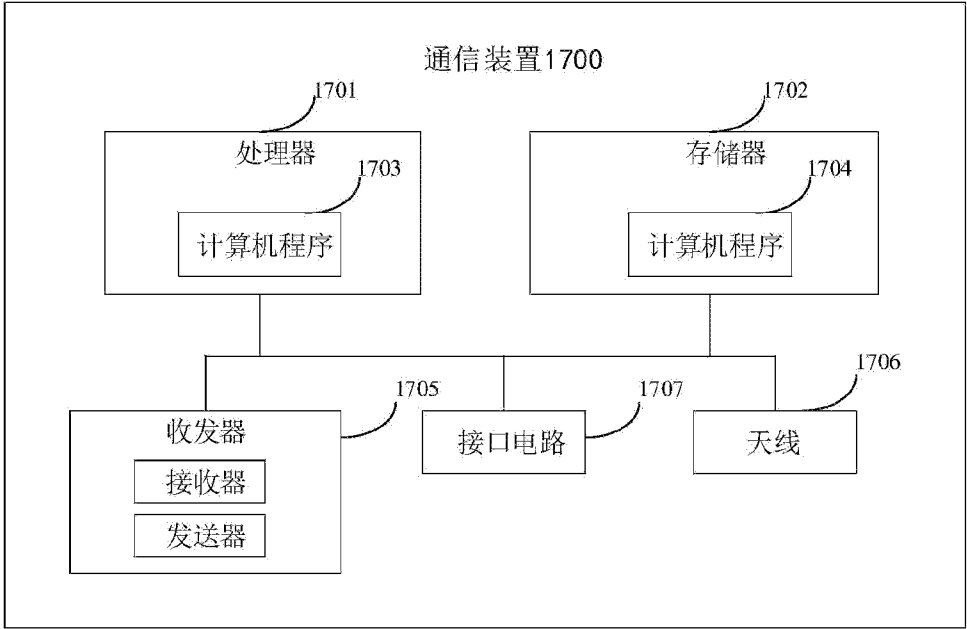
[图 18]



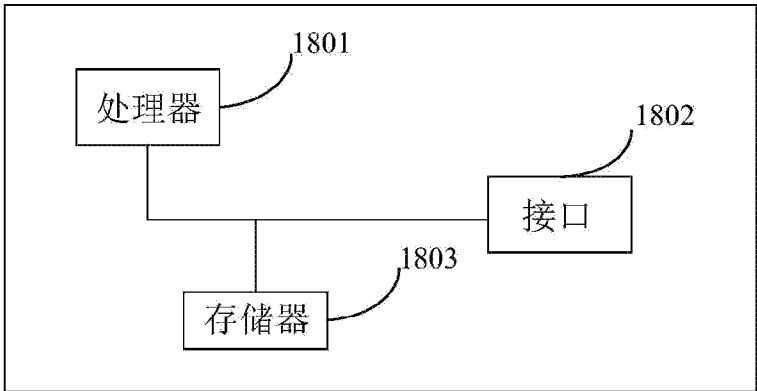
[图 19]



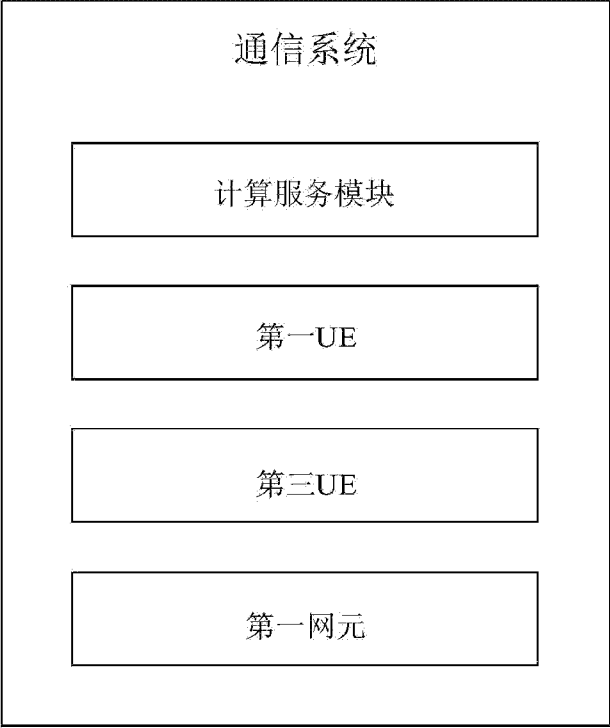
[图 20]



[图 21]



[图 22]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/082299

A. CLASSIFICATION OF SUBJECT MATTER

G06F9/50(2006.01)i; H04W4/50(2018.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC:H04W,H04L,G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNTXT, ENTXT, ENTXTC, WPABSC, WPABS, VEN, 3GPP: 算力, 分配, 调度, comput+ 3d capability, schedul+, distribut+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021208915 A1 (SPREADTRUM SEMICONDUCTOR (NANJING) CO., LTD.) 21 October 2021 (2021-10-21) description, page 8, line 1 to page 22, line 11, and figures 2-9	1-36
X	WO 2021208914 A1 (SPREADTRUM SEMICONDUCTOR (NANJING) CO., LTD.) 21 October 2021 (2021-10-21) description, page 3, line 16 to page 14, line 25, and figures 1-7	1-36
X	WO 2022143744 A1 (VIVO MOBILE COMMUNICATION CO., LTD.) 07 July 2022 (2022-07-07) description, page 12, line 15 to page 30, line 16, and figures 1-15	1, 2, 7, 8, 12-16, 18-20, 22-25, 29-36
A	CN 114756340 A (CHINA UNITED NETWORK COMMUNICATIONS GROUP CO., LTD.) 15 July 2022 (2022-07-15) entire document	1-36
A	CN 112988377 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 18 June 2021 (2021-06-18) entire document	1-36

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

09 November 2023

Date of mailing of the international search report

22 November 2023

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/082299

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2021208915	A1	21 October 2021	JP	2023521222	A	23 May 2023
				US	2023144062	A1	11 May 2023
				KR	20230035516	A	14 March 2023
WO	2021208914	A1	21 October 2021	None			
WO	2022143744	A1	07 July 2022	EP	4261683	A1	18 October 2023
CN	114756340	A	15 July 2022	None			
CN	112988377	A	18 June 2021	None			

A. 主题的分类 G06F9/50(2006.01)i; H04W4/50(2018.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) IPC:H04W,H04L,G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNKI,CNXTXT,ENTXT,ENTXTC,WPABSC,WPABS,VEN,3GPP:算力, 分配, 调度, comput+ 3d capability, sche- dul+, distribut+		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	WO 2021208915 A1 (展讯半导体(南京)有限公司) 2021年10月21日 (2021 - 10 - 21) 说明书第8页第1行至第22页第11行, 图2-9	1-36
X	WO 2021208914 A1 (展讯半导体(南京)有限公司) 2021年10月21日 (2021 - 10 - 21) 说明书第3页第16行至第14页第25行, 图1-7	1-36
X	WO 2022143744 A1 (维沃移动通信有限公司) 2022年7月7日 (2022 - 07 - 07) 说明书第12页第15行至第30页第16行, 图1-15	1, 2, 7, 8, 12-16, 18-20, 22-25, 29-36
A	CN 114756340 A (中国联合网络通信集团有限公司) 2022年7月15日 (2022 - 07 - 15) 全文	1-36
A	CN 112988377 A (腾讯科技(深圳)有限公司) 2021年6月18日 (2021 - 06 - 18) 全文	1-36
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
★ 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “D” 申请人在国际申请中引证的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 2023年11月9日		国际检索报告邮寄日期 2023年11月22日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088		授权官员 张慧 电话号码 (+86) 010-62412032

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/082299

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
WO	2021208915	A1	2021年10月21日	JP	2023521222	A	2023年5月23日
				US	2023144062	A1	2023年5月11日
				KR	20230035516	A	2023年3月14日
WO	2021208914	A1	2021年10月21日	无			
WO	2022143744	A1	2022年7月7日	EP	4261683	A1	2023年10月18日
CN	114756340	A	2022年7月15日	无			
CN	112988377	A	2021年6月18日	无			