

(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
2022 年 4 月 7 日 (07.04.2022)



(10) 国际公布号
WO 2022/067816 A1

(51) 国际专利分类号:
H04W 88/18 (2009.01)

(21) 国际申请号: PCT/CN2020/119747

(22) 国际申请日: 2020 年 9 月 30 日 (30.09.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 孙飞 (SUN, Fei); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 曾青海 (ZENG, Qinghai); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 罗海燕 (LUO, Haiyan); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

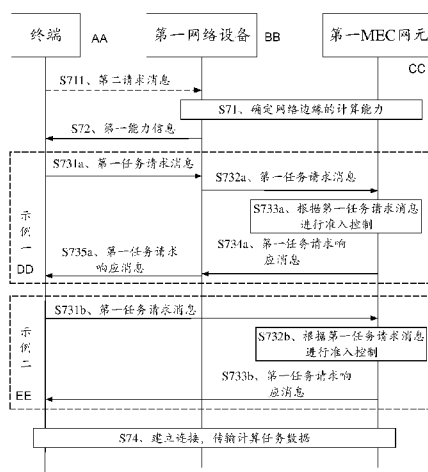
(74) 代理人: 北京同达信恒知识产权代理有限公司 (TDIP & PARTNERS); 中国北京市西城区裕民路 18 号北环中心 A 座 2002, Beijing 100029 (CN)。

(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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, MK, MT, NL, NO, PL, PT,

(54) Title: NETWORK EDGE COMPUTING METHOD AND COMMUNICATION APPARATUS

(54) 发明名称: 一种网络边缘计算方法及通信装置



S71 Determine a computing capability of a network edge
S72 First capability information
S74 Establish a connection, and transmit computing task data
S711 Second request message
S731a, S731b, S732a First task request message
S732b, S733a Perform admission control according to the first task request message
S733b, S734a, S735a First task request response message
AA Terminal
BB First network device
CC First MEC network element
DD Example I
EE Example II

(57) Abstract: Disclosed are a network edge computing method and a communication apparatus, which can ensure the security of internal data of a network side. The method comprises: a first network device determining a first computing capability of a first MEC function, wherein the computing capability of the first MEC function is used for indicating the capability of the first MEC function in providing a computing service; the first network device sending first capability information, wherein the first capability information is used for determining capability parameters of the first computing capability, and the capability parameters comprise one or more of a supported computing service type, hardware parameter and computing load; the first network device receiving a first task request message from a first terminal, wherein the first task request message is used for requesting the first MEC function to execute at least one computing task of the first terminal, and the at least one computing task is determined according to the first capability information; and the first network device sending the first task request message to the first MEC function.

RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, 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))。

(57) 摘要：本申请公开了一种网络边缘计算方法及通信装置，能够保证网络侧内部数据的安全性。该方法包括：第一网络设备确定第一MEC功能的第一计算能力，第一MEC功能的计算能力用于指示第一MEC功能所能提供计算服务的能力；第一网络设备发送第一能力信息，第一能力信息包括用于确定第一计算能力的能力参数，能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种；第一网络设备接收来自第一终端的第一任务请求消息，该第一任务请求消息用于请求第一MEC功能执行第一终端的至少一个计算任务，其中，至少一个计算任务是根据第一能力信息确定的；第一网络设备向第一MEC功能发送第一任务请求消息。

一种网络边缘计算方法及通信装置

技术领域

本申请涉及无线通信技术领域，尤其涉及一种网络边缘计算方法及通信装置。

5 背景技术

多接入边缘计算(multi-access edge computing, MEC)是指通过在靠近设备侧部署 MEC 计算功能，为用户提供更快的服务和达到更好的网络性能。

基于现有的 MEC 架构，即 MEC 计算功能部署在与核心网侧连接的数据网络（data network, DN）中。终端设备与网络侧设备建立连接，通过无线接入网将数据传输给核心网中的用户面功能，用户面功能可将数据传输给 DN，DN 中的 MEC 计算功能处理来自终端的数据，从而分担网络侧的计算负载。

通常，网络侧的计算业务包括网络内部数据，而 DN 中部署的 MEC 计算功能通常是第三方服务商提供的，如果要借助 MEC 执行网络侧的计算业务，就会将网络内部数据提供给第三方，安全性较低。另外，随着终端设备的计算能力的提升，例如互联自动驾驶汽车(connected and autonomous vehicles, CAV)技术使得车载终端具备较强的计算能力，终端设备也可以提供计算服务。基于现有 MEC 架构，MEC 设备执行网络侧的计算任务过程中，计算任务相关数据需要与核心网交互，不能保证网络侧内部数据的安全性。

发明内容

20 本申请提供一种网络边缘计算方法及通信装置，能够保证网络侧内部数据的安全性。

第一方面，本申请实施例提供第一种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为第一网络设备为例进行描述。该方法包括：

25 第一网络设备确定第一 MEC 功能的第一计算能力，第一 MEC 功能的计算能力用于指示第一 MEC 功能所能提供计算服务的能力；

所述第一网络设备发送第一能力信息，所述第一能力信息包括用于确定所述第一计算能力的能力参数，所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种；

30 第一网络设备接收来自第一终端的第一任务请求消息，该第一任务请求消息用于请求所述第一 MEC 功能执行所述第一终端的至少一个计算任务，其中，所述至少一个计算任务是根据所述第一能力信息确定的；

所述第一网络设备向所述第一 MEC 功能发送所述第一任务请求消息。

35 在该方案中，第一 MEC 功能设置于第一网络设备内或者该第一 MEC 功能独立于第一网络设备，即网络侧设置 MEC 功能，使得终端就近得到计算服务，从而提高终端的服务体验。

在可能的实现方式中，第一网络设备确定第一 MEC 功能的计算能力包括如下三种方式，本申请实施例不限制使用如下三种方式的任意一种，较为灵活。

方式一，第一网络设备向第一 MEC 功能发送第一请求消息，该第一请求消息用于请

求获取所述第一 MEC 功能的计算能力;第一网络设备接收来自第一 MEC 功能的第一响应消息,该第一响应消息包括能力参数,所述能力参数用于确定第一 MEC 功能的第一计算能力,所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种。

该方案可认为是第一网络设备向第一 MEC 功能请求查询第一 MEC 功能的第一计算能力。即在需要获取第一计算能力的情况下,发送第一请求消息,可避免不必要的信令开销。

方式二,第一网络设备向所述第一 MEC 功能发送第二请求消息,该第二请求消息用于请求与第一 MEC 功能建立连接;第一网络设备接收来自第一 MEC 功能的第二响应消息,该第二响应消息包括能力参数,所述能力参数用于确定所述第一 MEC 功能的第一计算能力,所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种。

该方案可复用第一网络设备与第一 MEC 功能的建立连接的请求消息,即第一网络设备只要与第一 MEC 功能请求建立连接,第一 MEC 功能就向第一网络设备通知该第一 MEC 功能的第一计算能力,这样不需要针对 MEC 能力交互额外的信令,以尽量节约信令的开销。

方式三,第一网络设备接收来自第一 MEC 功能的第一更新消息,该第一更新消息包括第一 MEC 功能的第一计算能力的相关信息。

该方案中,第一 MEC 功能的计算能力可认为是第一 MEC 功能的一种配置信息。第一更新消息可认为是第一 MEC 功能的配置信息发生变化后,主动向第一网络设备发送的更新消息。由于第一 MEC 功能主动告知第一网络设备,第一 MEC 功能更新后的计算能力,所以可避免第一网络设备发送不必要的第一任务请求消息,以尽量节约信令的开销。

在一种可能的实现方式中。所述方法还包括:

第一网络设备接收来自第一终端的第三请求消息,所述第三请求消息用于请求获取 MEC 功能的计算能力。

在该方案中,第一网络设备可在接收到第三请求消息的情况下,获取第一 MEC 功能的计算能力。即基于第一终端的触发,获取第一 MEC 功能的计算能力,以尽量避免不必要的信令开销。

示例性的,所述第三请求消息为网络边缘能力查询请求消息;或者,所述第三请求消息为按需系统消息。

在一种可能的实现方式中,第一网络设备发送第一能力信息,包括:

第一网络设备广播所述第一能力信息。该方案中,第一网络设备可广播获取的网络侧的计算服务能力,以告知任意终端网络侧的计算服务能力。这样任意终端可根据网络侧的计算能力请求合适的 MEC 功能执行至少一个计算任务。

在一种可能的实现方式中,所述方法还包括:

第一网络设备确定所述第一 MEC 功能的所述第一计算能力不满足所述至少一个计算任务的计算需求,第一网络设备向第二网络设备发送第三任务请求消息,所述第三任务请求消息用于请求所述第二网络设备执行所述至少一个计算任务;

第一网络设备接收来自第二网络设备针对所述第三任务请求消息的响应消息。

在该方案中,如果第一网络设备确定第一 MEC 功能具有的计算能力不满足执行第一终端的至少一个计算任务的计算需求,那么第一网络设备可请求第二网络设备侧的 MEC 功能执行该至少一个计算任务,以尽量保证终端侧的计算任务的顺利执行。

在一种可能的实现方式中,所述方法还包括:

第一网络设备确定第二网络设备侧的第二 MEC 功能的第二计算能力能够满足所述至少一个计算任务的计算需求, 所述第二 MEC 功能设置于所述第二网络设备内或者所述第二 MEC 功能独立于所述第二网络设备, 其中, 所述第二 MEC 功能的第二计算能力是所述第二 MEC 功能通知所述第一 MEC 功能的。

5 在该方案中, 第一网络设备可从第一 MEC 功能处获知第二 MEC 功能的第二计算能力, 从而在第二计算能力满足至少一个计算任务的计算需求的情况下, 向第二网络设备发送第三任务请求消息。这样可避免向不能满足计算需求的网络侧 MEC 功能请求执行至少一个计算任务, 以尽量保证交互更少的信令。

在一种可能的实现方式中, 所述方法还包括:

10 第一网络设备接收来自第一 MEC 功能的指示信息, 所述指示信息用于指示第二终端, 所述第二终端能够提供计算服务;

第一网络设备将来自第一终端的第三任务请求消息转发给所述第二终端, 所述第三任务请求消息用于请求执行所述至少一个计算任务。

15 该方案中, 第二终端能够提供计算服务, 那么第二终端可认为是 MEC 功能实体, 从而执行第一终端的至少一个计算任务。对于第一网络设备来说, 可根据至少一个计算任务的计算需求, 以及第一 MEC 功能或第二终端的计算资源、计算负载等信息, 确定由第一 MEC 功能或第二终端执行该至少一个计算任务。即综合考虑网络的计算资源, 合理分配计算任务, 提高系统容量。

20 第二方面, 本申请实施例提供第一种网络边缘计算方法, 该方法可由第二通信装置执行, 第二通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置, 例如芯片或芯片系统。下面以所述通信设备为终端设备为例进行描述。该方法包括:

25 第一终端接收来自第一网络设备的第一能力信息, 所述第一能力信息用于指示所述第一 MEC 功能的第一计算能力, 所述第一计算能力用于指示所述第一 MEC 功能能够提供计算服务的能力, 所述第一能力信息包括用于确定所述第一计算能力的能力参数, 所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种;

所述第一终端向所述第一网络设备或者所述第一 MEC 功能发送第一任务请求消息, 该第一任务请求消息用于请求执行至少一个计算任务;

所述第一终端接收来自所述第一 MEC 功能的准入信息, 并根据准入信息将所述至少一个计算任务的相关数据发送给所述第一 MEC 功能。

30 在该方案中, 第一 MEC 功能设置于第一网络设备内或者该第一 MEC 功能独立于第一网络设备。第一终端可根据第一 MEC 功能的第一计算能力, 向第一 MEC 功能请求执行至少一个计算任务, 提高终端的服务体验。

在一种可能的实现方式中, 所述方法还包括: 第一终端向第一网络设备发送第二请求消息, 所述第二请求消息用于请求获取 MEC 的计算能力。

35 该方案中, 第一终端有计算需求时, 可主动触发第一网络设备获取第一 MEC 功能的第一计算能力, 从而根据第一计算能力确定是否请求第一 MEC 功能执行至少一个计算任务。这样可保证第一终端选择能够满足至少一个计算任务的计算需要的第一 MEC 功能, 以避免不必要的信令开销。

40 示例性的, 所述第二请求消息为网络边缘能力查询请求消息; 或者, 所述第二请求消息为按需系统消息。

在一种可能的实现方式中，所述方法还包括：

第一终端根据第一能力信息确定驻留小区，所述驻留小区支持所述第一终端需要的计算服务；或者，

第一终端根据第一能力信息从当前驻留小区切换到邻小区，所述当前驻留小区不满足所述第一终端的计算服务需求，所述邻小区满足所述第一终端的计算服务需求。

该方案中，第一终端可根据第一 MEC 功能的第一计算能力优先选择支持该终端所需 MEC 计算服务的小区进行驻留，这样当终端有业务需要处理时，不需要重选到能够支持其计算服务的小区，可减少接入时延。

第三方面，本申请实施例提供第一种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为 MEC 功能为例进行描述。该方法包括：

MEC 功能确定第二终端执行第一网络设备或所述 MEC 功能的至少一个计算任务，所述第二终端能够提供计算服务，所述 MEC 功能设置于所述第一网络设备内或者所述 MEC 功能独立于所述第一网络设备；

MEC 功能向第一网络设备发送会话请求消息，该会话请求消息用于指示第一网络设备建立第一网络设备与所述 MEC 功能之间的第一传输通道，以及建立第一网络设备到所述第二终端的第二传输通道；

MEC 功能通过第一传输通道以及第二传输通道将所述至少一个计算任务的相关数据发送给所述第二终端。

该方案中，第二终端能够提供计算服务，那么第二终端可认为是 MEC 功能实体。对于网络侧设备来说，可确定将网络侧的计算任务由第二终端来执行。例如网络侧设备可根据至少一个计算任务的计算需求，以及第一 MEC 功能或第二终端的计算资源、计算负载等信息，确定由第一 MEC 功能或第二终端执行该至少一个计算任务。即综合考虑网络的计算资源，合理分配计算任务，提高系统容量。

在可能的实现方式中，所述方法还包括：

MEC 功能接收来自第二终端的注册请求消息，该注册请求消息用于请求所述 MEC 功能注册所述第二终端的计算能力信息，所述至少一个计算任务是根据所述计算能力信息确定的。

该方案中，能够提供计算服务的第二终端可将自身的相关信息，例如身份信息、计算能力信息等注册到 MEC 功能。这样 MEC 功能可根据这些注册信息选择合适的第二终端执行至少一个计算任务。

第四方面，本申请实施例提供第二种网络边缘计算方法，该方法可由第二通信装置执行，第二通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为终端设备为例进行描述。该方法包括：

第二终端向 MEC 功能发送注册请求消息，该注册请求消息用于请求所述 MEC 功能注册所述第二终端的计算能力信息，所述计算能力信息用于指示所述第二终端能够提供的计算服务；

第二终端接收第一网络设备发送的第一任务请求消息，所述第一任务请求消息用于请求执行所述第一网络设备或所述 MEC 功能的至少一个任务，所述至少一个计算任务是根据所述计算能力信息确定的。

关于第四方面或第四方面的各种可能的实施方式所带来的技术效果，可参考对于第三方面或第三方面的各种可能的实施方式的技术效果的介绍。

第五方面，本申请实施例提供第一种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为第一终端为例进行描述。该方法包括：

第一终端确定第二终端，所述第二终端能够提供计算服务；

第一终端向所述第二终端发送第一任务请求消息，该第一请求消息用于请求所述第二终端执行所述第一终端的至少一个计算任务；

第一终端接收来自第二终端的准入信息，并根据所述准入信息将所述至少一个计算任务的相关数据发送给所述第二终端。

该方案中，由于第二终端能够提供计算服务，那么第二终端可认为是 MEC 功能实体。对于第一终端来说，也可以请求第二终端执行至少一个计算任务。即第一终端通过直连通信的方式可将所述至少一个计算任务的相关数据发送给所述第二终端，请求第二终端执行至少一个计算任务。

在可能的实现方式中，第一终端确定第二终端，包括：

第一终端广播 MEC 服务信息，所述 MEC 服务信息包括所述至少一个任务的计算需求参数；

第一终端接收来自某个终端针对所述 MEC 服务信息的响应消息，并将该终端确定为所述第二终端。

该方案中，第一终端可广播计算服务需求，满足该计算服务需求的第二终端可响应第一终端。即第一终端接收到某个终端的响应，可认为该终端是第二终端。

在可能的实现方式中，第一终端确定第二终端，包括：

第一终端向第一网络设备发送所述第一任务请求消息；

第一终端接收来自所述第一网络设备的第一响应消息，所述第一响应消息包括所述第二终端的信息。

该方案中，第一终端可通过第一网络设备确定第二终端。

第六方面，本申请实施例提供第二种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为第二终端为例进行描述。该方法包括：

第二终端广播计算能力信息，所述第二终端能够提供计算服务，所述计算能力信息用于指示所述第二终端的第二计算能力；

第二终端接收来自第一终端的第一任务请求消息，所述第一任务请求消息用于请求执行所述第一终端的至少一个计算任务。

在一种可能的实现方式中，所述方法还包括：

所述第二终端将所述计算能力信息发送给 MEC 功能。

关于第六方面或第六方面的各种可能的实施方式所带来的技术效果，可参考对于第五方面或第五方面的各种可能的实施方式的技术效果的介绍。

第七方面，本申请实施例提供第三种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为第二终端为例进行描述。该方法包括：

第一 MEC 功能接收来自第一终端的第一任务请求消息，所述第一任务请求消息用于请求执行所述第一终端的计算任务；

第一 MEC 功能向至少一个边缘计算设备发送第二任务请求消息，每个边缘计算设备对应的所述第二任务请求消息用于请求该边缘计算设备执行所述计算任务中的部分计算子任务，所述计算任务包括的多个计算子任务具有逻辑关系，所述至少一个边缘计算设备包括第二 MEC 和/或第二终端，所述第二终端能够提供计算服务。

该方案中，可将第一终端的计算任务划分为具有逻辑关系的多个计算子任务，并由多个边缘计算设备执行。即多个边缘计算设备可执行基于服务功能链(service function chain, SFC)的多个计算子任务，从而可通过缩小任务调度粒度来提高资源调度的精度，提高系统的计算资源利用效率。

在一种可能的实现方式中，所述第一请求消息包括如下信息的一种或多种：

所述多个计算子任务、所述多个计算子任务之间的逻辑关系，以及多个计算子任务中各个子任务的服务需求。

该方案中，第一终端可根据需求将第一终端的计算任务划分为基于 SFC 的多个计算子任务，并将该多个计算子任务的相关信息告知第一 MEC 功能。这样第一 MEC 功能可根据这些信息选择合适的至少一个边缘计算设备来执行这多个计算子任务，即合理利用各个边缘计算设备的计算资源等，以提高系统容量。

在一种可能的实现方式中，所述方法还包括：

第一 MEC 功能将所述计算任务划分为基于 SFC 的所述多个计算子任务，其中，所述第二任务请求消息包括第一信息和第二信息，所述第一信息用于指示所述多个计算子任务所述第二信息用于指示所述多个计算子任务之间的逻辑关系。

该方案中，第一 MEC 功能可将第一终端的计算任务划分为基于 SFC 的多个计算子任务，并根据这些信息选择合适的至少一个边缘计算设备来执行这多个计算子任务，即合理利用各个边缘计算设备的计算资源等，以提高系统容量。

应理解，至少一个边缘计算设备可包括多个 MEC 功能，这多个 MEC 功能彼此之间可交互自身的相关信息，例如计算能力、计算负载等信息。从而每个 MEC 功能可根据基于 SFC 的多个计算子任务，以及各个 MEC 功能的计算能力以及计算负载等，从多个 MEC 功能中选择合理的 MEC 功能执行这多个计算子任务，以尽量提高计算资源利用率。

示例性的，所述方法还包括：第一 MEC 功能接收来自第二 MEC 功能发送第一信息，所述第一信息用于指示所述第二 MEC 功能的第二计算能力，以及所述第二 MEC 的第二计算负载。该方案即第二 MEC 功能告知第一 MEC 功能，第二 MEC 功能自身的计算能力、计算负载等。

示例性的，所述方法还包括：所述第一 MEC 功能向第二 MEC 功能发送第二信息，所述第二信息用于指示所述第一 MEC 功能的第一计算能力，以及所述第一 MEC 的第二计算负载。该方案即第一 MEC 功能告知第二 MEC 功能，第一 MEC 功能自身的计算能力、计算负载等。

在一种可能的实现方式中，所述方法还包括：

第一 MEC 功能接收来自第二终端的注册请求消息，所述注册请求消息用于请求注册所述第二终端的计算能力，以及所述第二终端的计算负载，所述第二终端能够提供计算服务。

同理，能够提供计算服务的第二终端也可以将自身的计算能力、计算负载告知第一 MEC 功能。这样第一 MEC 功能也可以选择合适的第二终端来执行第一终端的多个计算子任务。

第八方面，本申请实施例提供第一种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为第一终端为例进行描述。该方法包括：

第一终端向第一 MEC 功能发送第一任务请求消息，所述第一任务请求消息用于请求执行所述第一终端的计算任务；

在一种可能的实现方式中，所述第一请求消息包括如下信息的一种或多种：

所述多个计算子任务、所述多个计算子任务之间的逻辑关系，以及多个计算子任务中各个子任务的服务需求。

关于第八方面或第八方面的各种可能的实施方式所带来的技术效果，可参考对于第七方面或第七方面的各种可能的实施方式的技术效果的介绍。

第九方面，本申请实施例提供第一种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为网络设备为例进行描述。该方法包括：

网络设备接收第一信息，所述第一信息用于指示终端建立的协议数据单元（protocol data unit）PDU 会话与计算服务器的映射关系；

所述网络设备根据所述映射关系对来自所述终端的至少一个计算任务的相关数据进行处理。

在一种可能的实现方式中，所述方法包括：

所述网络设备被配置数据包过滤规则 PDR，所述 PDR 用于保证所述至少一个计算任务关联的数据流映射到正确的服务质量 QoS 流。

第十方面，本申请实施例提供第一种网络边缘计算方法，该方法可由第一通信装置执行，第一通信装置可以是通信设备或能够支持通信设备实现该方法所需的功能的通信装置，例如芯片或芯片系统。下面以所述通信设备为终端为例进行描述。该方法包括：

终端确定第一信息，所述第一信息包括所述终端建立的 PDU 会话与远端服务器的对应关系，所述第一信息用于指示 MEC 功能对来自所述远端服务器的终端的至少一个计算任务的相关数据进行处理；

终端向所述 MEC 功能发送第一信息；

在一种可能的实现方式中，所述方法还包括：

所述终端向所述 MEC 功能发送第一任务请求消息，所述第一任务请求消息用于请求所述 MEC 功能执行所述终端的至少一个计算任务，所述第一任务请求消息包括所述远端服务器的相关信。

在一种可能的实现方式中，所述方法还包括：

所述终端接收来自所述 MEC 功能发送第一任务请求响应消息，所述第一任务请求消息包括所述 MEC 功能选择的计算服务器的相关信息。

第十一方面，本申请实施例提供了一种通信装置，该通信装置具有实现上述第一方面至第十方面中任一方面中方法实施例中的行为的功能。所述功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模

块。

在一种可能的实现方式中，该通信装置具有实现上述第一方面方法实施例中的行为的功能，该通信装置包括收发模块和处理模块，其中：

所述处理模块用于确定第一 MEC 功能的第一计算能力，第一 MEC 功能的第一计算能力用于指示第一 MEC 功能所能提供计算服务的能力；

所述收发模块用于发送第一能力信息，所述第一能力信息包括用于确定所述第一计算能力的能力参数，所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种；

所述收发模块还用于接收来自第一终端的第一任务请求消息，该第一任务请求消息用于请求所述第一 MEC 功能执行所述第一终端的至少一个计算任务，其中，所述至少一个计算任务是根据所述第一能力信息确定的；

所述收发模块还用于向第一 MEC 功能发送所述第一任务请求消息。

在一种可能的实现方式中，该通信装置具有实现上述第二方面方法实施例中的行为的功能，该通信装置包括收发模块和处理模块，其中：

所述收发模块用于接收来自第一网络设备的第一能力信息，该第一能力信息用于指示第一 MEC 功能的第一计算能力，该第一计算能力用于指示第一 MEC 功能能够提供的计算服务，该第一 MEC 功能设置于第一网络设备内或者该第一 MEC 功能独立于第一网络设备；

所述收发模块还用于向第一网络设备或者第一 MEC 功能发送所述处理模块生成的第一任务请求消息，该第一任务请求消息用于请求执行至少一个计算任务；

所述收发模块还用于接收准入信息，并根据准入信息将至少一个计算任务的相关数据发送给所述第一 MEC 功能。

在一种可能的实现方式中，该通信装置具有实现上述第三方面方法实施例中的行为的功能，该通信装置包括收发模块和处理模块，其中：

所述处理模块用于确定第二终端执行第一网络设备或所述通信装置的至少一个计算任务，所述第二终端能够提供计算服务，所述通信装置设置于所述第一网络设备内或者所述通信装置独立于所述第一网络设备；

所述收发模块用于向第一网络设备发送会话请求消息，该会话请求消息用于指示第一网络设备建立第一网络设备与所述通信装置之间的第一传输通道，以及建立第一网络设备到所述第二终端的第二传输通道；

所述收发模块用于通过第一传输通道以及第二传输通道将所述至少一个计算任务的相关数据发送给所述第二终端。

在一种可能的实现方式中，该通信装置具有实现上述第四方面方法实施例中的行为的功能，该通信装置包括收发模块和处理模块，其中：

所述收发模块用于向 MEC 功能发送注册请求消息，该注册请求消息用于请求所述 MEC 功能注册所述通信装置的计算能力信息，所述计算能力信息用于指示所述通信装置能够提供的计算服务；

所述收发模块还用于接收第一网络设备发送的第一任务请求消息，所述第一任务请求消息用于请求执行所述第一网络设备或所述 MEC 功能的至少一个任务，所述至少一个计算任务是根据所述计算能力信息确定的。

在一种可能的实现方式中，该通信装置具有实现上述第五方面方法实施例中的行为的

功能, 该通信装置包括收发模块和处理模块, 其中:

所述处理模块用于确定第二终端, 所述第二终端能够提供计算服务;

所述收发模块用于向所述第二终端发送第一任务请求消息, 该第一请求消息用于请求所述第二终端执行所述通信装置的至少一个计算任务;

5 所述收发模块用于接收来自第二终端的准入信息, 并根据所述准入信息将所述至少一个计算任务的相关数据发送给所述第二终端。

在一种可能的实现方式中, 该通信装置具有实现上述第六方面方法实施例中的行为的功能, 该通信装置包括收发模块和处理模块, 其中:

10 所述收发模块用于广播所述处理模块确定的计算能力信息, 所述通信装置能够提供计算服务, 所述计算能力信息用于指示所述通信装置的第二计算能力;

所述收发模块还用于接收来自第一终端的第一任务请求消息, 所述第一任务请求消息用于请求执行所述第一终端的至少一个计算任务。

在一种可能的实现方式中, 该通信装置具有实现上述第七方面方法实施例中的行为的功能, 该通信装置包括收发模块和处理模块, 其中:

15 所述收发模块用于接收来自第一终端的第一任务请求消息, 所述第一任务请求消息用于请求执行所述第一终端的计算任务;

所述收发模块还用于向至少一个边缘计算设备发送所述处理模块生成的第二任务请求消息, 每个边缘计算设备对应的所述第二任务请求消息用于请求该边缘计算设备执行所述计算任务中的部分计算子任务, 所述计算任务包括的多个计算子任务具有逻辑关系, 所述至少一个边缘计算设备包括第二 MEC 和/或第二终端, 所述第二终端能够提供计算服务。

20 在一种可能的实现方式中, 该通信装置具有实现上述第八方面方法实施例中的行为的功能, 该通信装置包括收发模块和处理模块, 其中:

所述收发模块用于向第一 MEC 功能发送第一任务请求消息, 所述第一任务请求消息用于请求执行所述通信装置的计算任务。

25 在一种可能的实现方式中, 该通信装置具有实现上述第九方面方法实施例中的行为的功能, 该通信装置包括收发模块和处理模块, 其中:

所述收发模块用于接收第一信息, 所述第一信息用于指示终端建立的协议数据单元 (protocol data unit) PDU 会话与计算服务器的映射关系;

30 所述处理模块用于根据所述映射关系对来自所述终端的至少一个计算任务的相关数据进行处理。

在一种可能的实现方式中, 该通信装置具有实现上述第十方面方法实施例中的行为的功能, 该通信装置包括收发模块和处理模块, 其中:

35 所述处理模块用于确定第一信息, 所述第一信息包括所述通信装置建立的 PDU 会话与远端服务器的对应关系, 所述第一信息用于指示 MEC 功能对来自所述远端服务器的通信装置的至少一个计算任务的相关数据进行处理;

所述收发模块用于向所述 MEC 功能发送第一信息;

40 第十二方面, 本申请实施例提供一种通信装置, 包括至少一个处理器。该处理器可用于执行计算机程序, 以实现上述第一方面至上述第十方面中任一方面中任一可能实现方式中的方法。可选地, 该通信装置还包括通信接口, 处理器与通信接口耦合, 所述通信接口用于输入和/或输出信息, 所述信息包括计算机程序、指令和数据中的至少一项。可选地,

该通信装置还包括存储器，所述存储器用于存储计算机程序、指令和数据中的至少一项，处理器与存储器耦合。

在一种实现方式中，该通信装置为终端设备或网络设备。当该通信装置为终端设备或网络设备时，所述通信接口可以是收发器，或，输入和/或输出接口。可选地，所述收发器可以为收发电路，也可以通过终端设备或网络设备中的天线、馈线和编解码器等实现。可选地，所述输入和/或输出接口可以为输入和/或输出电路。

在另一种实现方式中，该通信装置为配置于终端设备或网络设备中的芯片或芯片系统。当该通信装置为配置于终端设备或网络设备中的芯片或芯片系统时，所述通信接口可以是输入和/或输出接口、接口电路、输出电路、输入电路、管脚或相关电路等。所述处理器也可以为处理电路或逻辑电路。

第十三方面，本申请实施例提供了一种芯片系统，该芯片系统包括处理器，还可以包括存储器，用于实现第一方面至第十方面中任一方面中的通信装置执行的方法。在一种可能的实现方式中，所述芯片系统还包括存储器，用于保存程序指令和/或数据。该芯片系统可以由芯片构成，也可以包含芯片和其他分立器件。

第十四方面，本申请实施例提供了一种通信系统，所述通信系统包括上述第九方面-上述第十三方面所述的一种或多种通信装置。

第十五方面，本申请提供了一种计算机可读存储介质，该计算机可读存储介质存储有计算机程序，当该计算机程序被运行时，实现上述各方面中由通信装置执行的方法。

第十六方面，提供了一种计算机程序产品，所述计算机程序产品包括：计算机程序代码，当所述计算机程序代码被运行时，使得上述各方面中由通信装置执行的方法被执行。

附图说明

图 1 为 MEC 应用层架构的示意图；

图 2 为现有技术中 MEC 部署的一种架构示意图；

图 3 为本申请实施例提供的 MEC 网络架构的示意图；

图 4 为本申请实施例提供的 MEC 管理功能的网络架构示意图；

图 5 为本申请实施例提供的嵌入式 MEC 管理功能的协议架构；

图 6 为本申请实施例提供的独立 MEC 管理功能的协议架构；

图 7 为本申请实施例提供的第一种网络边缘计算方法的流程图；

图 7a 为本申请实施例提供网络设备确定 MEC 网元的计算能力的一种流程示意图；

图 8 为本申请实施例提供的第二种网络边缘计算方法的流程图；

图 9 为本申请实施例提供的第三种网络边缘计算方法的流程图；

图 10 为本申请实施例提供的第四种网络边缘计算方法的流程图；

图 11 为本申请实施例提供的第五种网络边缘计算方法的流程图；

图 12 为本申请实施例提供的第六种网络边缘计算方法的流程图；

图 13 为本申请实施例提供的第七种网络边缘计算方法的流程图；

图 14 为本申请实施例提供的第八种网络边缘计算方法的流程图；

图 15 为本申请实施例提供的基于 SFC 的 MEC 方法；

图 16 为基于现有 MEC 架构本申请实施例提供的中间渲染信令交互流程图；

图 17 为本申请实施例提供的 MEC 架构的中间渲染信令交互流程图；

图 18 为本申请实施例提供的计算服务器处理数据包的一种流程示意图;

图 19 为本申请实施例提供的计算服务器处理数据包的一种流程示意图;

图 20 为本申请实施例提供的计算服务器处理数据包的一种流程示意图;

图 21 为本申请实施例提供的 D2D 本地渲染的流程示意图;

5 图 22 为本申请实施例提供的通信装置的一种结构示意图;

图 23 为本申请实施例提供的通信装置的一种结构示意图。

具体实施方式

为了使本申请实施例的目的、技术方案和优点更加清楚,下面将结合附图对本申请实
10 施例作进一步地详细描述。

为便于本领域技术人员理解本申请实施例提供的技术方案,首先对本申请实施例中的
部分用语进行解释说明。

1) 网络侧设备,也可称为网络设备,是网络侧中一种用于发射或接收信号的实体,
是通信系统中将终端设备接入到无线网络的设备,一般通过有线链路(例如光纤线缆)连
15 接到核心网,如新一代基站(generation Node B, gNodeB)。网络侧设备可负责接收核心网
的数据并转发给无线回传设备,或者接收无线回传设备的数据并转发给核心网。网络设备
可以是用于与移动设备通信的设备。网络设备可以是无线局域网(wireless local area
networks, WLAN)中的 AP,长期演进(long term evolution, LTE)中的演进型基站(evolutional
Node B, eNB 或 eNodeB),或者也可以包括 5G NR 系统中的下一代节点 B(next generation
20 node B, gNB),或者中继站或接入点,或者车载设备、可穿戴设备以及未来 5G 网络中的
网络设备或者未来演进的公共陆地移动网络(public land mobile network, PLMN)网络中的
网络设备,或 NR 系统中的 gNodeB/gNB 等。下面以网络侧设备是 gNB 为例。

gNB 可以包括天线,基带单元(base band unit, BBU)和射频拉远单元(remote radio
unit, RRU)。其中, BBU 可以通过公共无线接口(common public radio interface, CPRI)
25 或增强的 CPRI(enhance CPRI, eCPRI)等与 RRU 相连, RRU 可以通过馈线与天线相连。
该天线可以为无源天线,其与 RRU 是分离的,之间可以通过电缆连接。或者该天线可以
为有源天线单元(active antenna unit, AAU),即 AAU 的天线单元和 RRU 是集成在一块的。
AAU 实现部分物理层处理功能、射频处理及有源天线的相关功能。

在一些部署中, gNB 可以包括集中式单元(centralized unit, CU)和分离单元(distributed
30 unit, DU)。CU 实现 gNB 的部分功能, DU 实现 gNB 的部分功能,例如, DU 可用于实现
射频信号的收发,射频信号与基带信号的转换,以及部分基带处理。CU 可用于进行基带
处理,对基站进行控制等。在一些实施例中, CU 负责处理非实时协议和服务,实现无线
资源控制(radio resource control, RRC),分组数据汇聚层协议(packet data convergence
35 protocol, PDCP)层的功能。DU 负责处理物理层协议和实时服务,实现无线链路控制(radio
link control, RLC)层、介质接入控制(media access control, MAC)层和物理(physical,
PHY)层的功能。由于 RRC 层的信息最终会变成 PHY 层的信息,或者,由 PHY 层的信
息转变而来,因而,在这种架构下,高层信令,如 RRC 层信令,也可以认为是由 DU 发送
的,或者,由 DU 和 AAU 发送的。可以理解的是,网络设备可以为包括 CU 节点、DU 节
点、AAU 节点中一项或多项的设备。此外,可以将 CU 划分为接入网(radio access network,
40 RAN)中的网络设备,也可以将 CU 划分为核心网(core network, CN)中的网络设备,

本申请对此不做限定。

2) 终端设备, 也可以称为终端设备或者终端, 可以是能够接收网络设备调度和指示的无线终端设备, 终端设备可以是指向用户提供语音和/或数据连通性的设备, 或具有无线连接功能的手持式设备、或连接到无线调制解调器的其他处理设备。

终端设备可以经无线接入网 (如, radio access network, RAN) 与一个或多个核心网或者互联网进行通信, 与 RAN 交换语音和/或数据。该终端设备可以包括用户设备 (user equipment, UE)、无线终端设备、移动终端设备、设备到设备通信 (device-to-device, D2D) 终端设备、V2X 终端设备、机器到机器/机器类通信 (machine-to-machine /machine-type communications, M2M/MTC) 终端设备、物联网 (internet of things, IoT) 终端设备、订户单元 (subscriber unit)、订户站 (subscriber station)、移动站 (mobile station)、远程站 (remote station)、接入点 (access point, AP)、远程终端 (remote terminal)、接入终端 (access terminal)、用户终端 (user terminal)、用户代理 (user agent)、或用户装备 (user device) 等。例如, 终端可以包括移动电话 (或称为“蜂窝”电话), 具有移动终端设备的计算机, 便携式、袖珍式、手持式、计算机内置的移动装置等。又例如, 终端可包括虚拟现实 (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、无人驾驶 (self driving) 中的无线终端、远程手术 (remote medical surgery) 中的无线终端、智能电网 (smart grid) 中的无线终端、运输安全 (transportation safety) 中的无线终端、智慧城市 (smart city) 中的无线终端、智慧家庭 (smart home) 中的无线终端、未来演进的公用陆地移动通信网络 (public land mobile network, PLMN) 中的终端设备、或者车联网 (vehicle to everything, V2X) 中的车辆设备, 客户前置设备 (customer premises equipment, CPE) 等等。再例如, 终端可包括个人通信业务 (personal communication service, PCS) 电话、无绳电话、会话发起协议 (session initiation protocol, SIP) 话机、无线本地环路 (wireless local loop, WLL) 站、个人数字助理 (personal digital assistant, PDA)、等设备。或者终端还可包括受限设备, 例如功耗较低的设备, 或存储能力有限的设备, 或计算能力有限的设备等。例如包括条码、射频识别 (radio frequency identification, RFID)、传感器、全球定位系统 (global positioning system, GPS)、激光扫描器等信息传感设备。

作为示例而非限定, 在本申请的实施例, 该终端设备还可以是可穿戴设备。可穿戴设备也可以称为穿戴式智能设备或智能穿戴式设备等, 是应用穿戴式技术对日常穿戴进行智能化设计、开发出可以穿戴的设备的总称, 如眼镜、手套、手表、服饰及鞋等。可穿戴设备即直接穿在身上, 或是整合到用户的衣服或配件的一种便携式设备。可穿戴设备不仅仅是一种硬件设备, 更是通过软件支持以及数据交互、云端交互来实现强大的功能。广义穿戴式智能设备包括功能全、尺寸大、可不依赖智能手机实现完整或者部分的功能, 例如: 智能手表或智能眼镜等, 以及只专注于某一类应用功能, 需要和其它设备如智能手机配合使用, 如各类进行体征监测的智能手环、智能头盔、智能首饰等。而如上介绍的各种终端设备, 如果位于车辆上 (例如放置在车辆内或安装在车辆内), 都可以认为是车载终端设备, 车载终端设备例如也称为车载单元 (on-board unit, OBU)。

在本申请的实施例, 网络设备和终端可以部署在陆地上, 包括室内或室外、手持或车载; 也可以部署在水面上; 还可以部署在空中的飞机、气球和卫星上。本申请的实施例对网络设备和终端的应用场景不做限定。

3) MEC 可以理解为边缘式的大数据处理平台, 即把传统的数据中心切割成各种小型

数据中心后放置到网络的边缘节点（服务器），以期更靠近用户，为用户提供更快的服务和达到更好的网络性能。换句话说，MEC 是将终端侧的计算负载卸载(offloading)至网络边缘的计算服务器，在网络边缘提供近似云端计算、存储和处理的能力。

如图 1 所示，为 MEC 应用层架构的示意图。终端侧包括一个或多个应用客户端(application client)，以及一个或多个边缘启动客户端(edge enabler client, EEC)。应用客户端可认为是驻留在终端的应用，用于执行客户端功能。网络边缘侧包括边缘数据网络(edge data network, EDN)和边缘配置服务器（edge configuration server, ECS）。EDN 是本地数据网络，包括边缘应用服务器(edge application server, EAS)和边缘启动服务器(edge enabler server, EES)。其中，EES 用于提供 EAS 和 EEC 所需的支撑功能，例如 EES 可提供配置信息，以便与 EAS 交互应用层数据流量；EES 也可以向 EEC 提供 EAS 的相关信息；EES 支持 3GPP TS 23.222 中规定的 API 的调用和开放功能。EEC 用于提供应用客户端所需的支持功能，例如 EEC 用于获取和提供配置信息，以便与 EAS 交互应用层数据流量；EEC 用于发现 EDN 中可用的 EAS。ECS 用于提供 EEC 连接 EES 所需的支持功能，例如 ECS 向 EEC 提供 Edge 配置信息，例如包括 EEC 连接 EES 需要的信息，例如适用于局域数据网(local area data network, LADN)的服务区信息等；又例如还可包括 ECS 与 EES 建立连接的信息，例如 通用资源标识符(uniform resource identifier, URI)。EAS 是部署在 EDN 中的应用服务器，用于执行服务器功能。EAS 对核心网能力的使用方式可能不同，例如 EAS 是核心网信任的实体，则可以直接调用核心网功能 API；或者 EAS 可通过 EES 调用核心网能力；或者，EAS 可通过能力开放功能(例如 SCEF 或 NEF 网元)调用核心网能力。

终端侧通过核心网网络可与边缘数据网络进行交互。例如应用客户端连接到 EAS，可将应用层数据流量（也可以认为是计算任务）传输给 EAS，这样就可利用边缘计算的优势来使用应用程序的服务。应用程序中的服务器功能可能仅作为 EAS 使用。但是，如果应用的服务器功能同时作为边缘应用服务器和驻留在云端的应用服务器可用，则边缘应用服务器和应用服务器的功能可能不同。另外，如果边缘应用服务器和应用服务器的功能不同，应用数据流量也可以不同。

如上介绍了本申请实施例涉及的一些概念，下面介绍本申请实施例的技术特征。

随着应用业务多样化的发展，终端的计算业务呈现多样化、大量化的发展趋势，要求终端支持更高的计算量。例如在线游戏、语音识别、人脸识别、增强现实(augmented reality, AR)以及虚拟现实(virtual reality, VR)等等，要求终端具备更强的计算能力。然而受终端电池容量以及计算能力的限制，终端的计算能力已经难以满足日益增长的业务需求。为此，可基于云计算技术，将终端的部分计算业务卸载到云端，利用云端来执行终端的部分计算业务。但是，由于终端到云端的距离遥远，这就会带来更高的业务响应延迟，且会增加回程链路的负载。

为此提出可将终端的计算任务卸载到无线接入网络(radio access network, RAN)侧，这就需要 RAN 支持更高的计算量，以满足日益增长的业务的需求。例如要求 RAN 支持的计算量能够满足流量定向(Traffic Steering)业务、体验质量(quality of experience, QoE)优化(optimization)业务、大规模多入多出(massive-multi-input multi-output, M-MIMO)优化(optimization)业务等网络内部优化业务的需求。通常来说，网络内部优化业务需要借助人工智能(artificial intelligence, AI)来完成，这会导致 RAN 侧的计算负载激增，传统的 RAN 侧已经难以满足日益增长的业务需求。

继而提出了 MEC 的概念，即由网络边缘的计算服务器执行终端侧的计算任务，在网络边缘提供近似云端计算、存储和处理的能力。MEC 可以理解为边缘式的大数据处理平台，即把传统的数据中心切割成各种小型数据中心后放置到网络的边缘节点（服务器），以期更靠近用户，为用户提供更快的服务和达到更好的网络性能。

如图 2 所示，为 MEC 部署的一种架构示意图。MEC 系统可包括部署在数据网络（data network, DN）中的一个或多个计算服务器，以及 MEC 管理平台。MEC 管理平台可用于管理一个或多个计算服务器处理 DN 中的数据。其中，终端与 RAN 建立连接，通过 RAN 可将数据传输给核心网中的用户面功能（user plane function, UPF）。UPF 可通过 N6 接口将来自终端的数据传输给 DN。即 DN 可通过 N6 接口与核心网中的用户面功能（user plane function, UPF）进行用户面应用层数据的交互，获得来自终端的数据。DN 中的一个或多个计算服务器可处理来自终端的数据，从而分担终端侧的计算负载。

通常，以 RAN 侧的网络内部优化业务为例，相关计算业务会包括网络内部数据，而 DN 中部署的计算服务器通常是第三方服务商提供的，如果要借助 MEC 执行 RAN 侧的计算业务，就会将网络内部数据提供给第三方，安全性较低。另外，随着终端设备的计算能力的提升，终端设备也可以提供计算服务。即终端设备也可以作为网络边缘节点，能够分担 RAN 侧的部分计算任务，以提高系统整体的计算容量。但是图 2 所示的架构并不支持 RAN 侧的计算任务卸载到 UE 侧。

鉴于此，本申请实施例提供了一种网络边缘计算方法以及 MEC 网络架构，该 MEC 网络架构，在网络侧（本文中也称为 RAN 侧）和终端侧（本文中也称为 UE 侧）引入了 MEC 管理功能。基于该 MEC 网络架构可实现将 RAN 侧的计算任务卸载到 UE 侧执行，这样可降低 RAN 侧的计算负载。基于该 MEC 网络架构也可将 UE 侧的计算任务卸载到 RAN 侧执行，以减轻 UE 侧的计算负担。进一步地，由于 RAN 侧和 UE 侧都可以提供 MEC，所以可根据 RAN 侧和 UE 侧实际的计算负载以及 MEC 计算能力灵活调度网络系统的计算资源，从而提高系统计算资源的利用率。且由于 RAN 侧引入 MEC 管理功能，RAN 的内部数据可在 RAN 侧处理，不需要提供给如图 2 所示的第三方 MEC 计算服务器，能够保证 RAN 侧内部网络数据的安全性。需要说明的是，本申请实施例以 MEC 计算任务为例，但是本申请实施例不限制计算任务的类型，也可以是其它类型的计算任务。MEC 计算能力指的是 MEC 网元的计算能力，可用于指示 MEC 网元能够提供的 MEC 计算服务。为了简便，下文中将 MEC 计算能力统称为计算能力。

请参见图 3，为本申请实施例提供的 MEC 网络的架构示意图。该 MEC 网络包括一个或多个终端设备、RAN 侧设备以及核心网设备和云端服务器。RAN 侧设备可包括一个或多个网络设备，图 3 以包括网络设备 1 和网络设备 2 为例。相比于现有技术，图 3 所示的架构，在终端侧设置了 MEC 管理功能和至少一个 MEC 计算功能，在 RAN 侧也设置了 MEC 管理功能和至少一个 MEC 计算功能。应理解，MEC 计算功能是用于提供计算服务的软件、硬件或软件和硬件的组合。本申请实施例中，MEC 管理功能也可以认为是管理 MEC 计算功能的虚拟化资源的 MEC 管理网元。

位于 RAN 侧的 MEC 管理功能可管理 MEC 计算功能，并管理所要执行的计算任务（也可以称为 MEC 任务）。例如 MEC 管理功能可将要执行的计算任务分配给一个或多个 MEC 计算功能，以尽量提供系统计算资源的利用率。RAN 能力开放（RAN exposure）通过标准应用程序接口（application programming interface, API）可向 MEC 计算功能提供无线网络

信息 (radio network information, RNI), 以及 API 发布和 RNI 访问管理。

位于终端侧的 MEC 管理功能可用于发现 MEC 计算功能, 并管理终端侧的计算任务。例如可将终端侧的部分或全部计算任务卸载给 RAN 侧的一个或多个 MEC 计算功能, 也可以将终端侧的部分或全部计算任务分配给终端侧的一个或多个 MEC 计算功能, 或者也可以用于管理来自其它终端的计算任务。

需要说明的是, MEC 管理功能和/或 MEC 计算功能可以设置在 RAN 侧设备内或终端设备内, 也可以是独立于 RAN 侧设备或终端设备。例如 MEC 管理功能和 MEC 计算功能独立于 RAN 侧设备, 或者 MEC 管理功能设置在 RAN 侧设备内部, MEC 计算功能独立于 RAN 侧设备。又例如, MEC 管理功能设置在终端设备内部, MEC 计算功能可理解为是终端设备应用 (也可以理解为终端设备的上层) 提供的服务功能。图 3 以 MEC 管理功能位于终端设备内部, 以及 MEC 管理功能位于网络侧设备内部为例。应理解, MEC 管理功能和 MEC 计算功能独立于 RAN 侧设备, 那么 MEC 管理功能和 MEC 计算功能可以作为一个逻辑实体, 为了便于描述, 在本文中可称为 MEC 网元。

请参见图 4, 为本申请实施例提供的 MEC 管理功能的网络架构示意图。图 4 示出了嵌入式 MEC 管理功能的网络架构 (即 MEC 管理功能设置于设备内), 以及作为独立逻辑实体的 MEC 管理功能的网络架构 (即 MEC 管理功能设置于设备外部)。图 4 以 MEC 管理功能设置在 RAN 侧为例。

如图 4 中的 (a) 所示, MEC 管理功能是嵌入式 MEC 管理功能, 即 MEC 管理功能设置在 RAN 侧设备内。MEC 管理功能作为 RAN 侧设备的一部分, 可管理 RAN 侧设备的 MEC 计算功能, 例如指示 MEC 计算功能执行 RAN 侧或终端侧执行计算任务。应理解, RAN 侧设备还可以包括连接管理功能、移动管理功能和服务质量 (quality of service, QoS) 管理功能。

如图 4 中的 (b) 所示, MEC 管理功能是独立的逻辑实体, 即 MEC 管理功能设置在 RAN 侧设备外部。MEC 管理功能可与一个或多个 RAN 侧设备交互信息, 例如 MEC 管理功能可通过与网络设备 1 或网络设备 2 之间的 H 接口交互信息。终端设备可通过 RAN 侧设备与 MEC 网元交互信息。应理解, 如果网络中存在多个 MEC 网元, 那么多个 MEC 网元彼此可通过 MEC 网元之间的 X 接口交互信息。需要说明的是, 本文中 H 接口指的是网络设备间的接口, 本申请实施例对 H 接口的具体名称不作限制。同理, X 接口指的是 MEC 网元之间的接口, 本申请实施例对 X 接口的具体名称不作限制。

MEC 计算功能执行计算任务, 需要与 RAN 侧设备或者终端设备交互执行计算任务的相关信息, 例如与计算任务相关的数据 (可简称为计算任务数据)。例如 RAN 侧设备与 MEC 计算功能可能会交互计算任务数据。在本申请实施例中, MEC 管理功能设置在 RAN 侧设备, 相较于现有技术中 MEC 管理功能设置在核心网侧来说, RAN 侧设备与 MEC 计算功能的交互是在 RAN 侧内部实现的, 能够保证 RAN 侧设备数据的安全性。同时, RAN 侧设备与 MEC 计算功能的交互是在 RAN 侧内部实现的, 相较于与核心网侧的 MEC 计算功能交互来说, 可降低交互时延。

应理解, RAN 侧的 MEC 管理功能和 RAN 侧设备需要建立通信接口, 并完成承载映射的配置, 以根据配置来进行 MEC 管理功能和 RAN 侧设备之间的数据传输。该通信接口需要支持用户面协议层和控制面协议层。对于终端设备而言, 控制面由 MEC 管理功能与 RAN 侧的 MEC 管理功能实现, 用户面可承载现有的无线信令承载 (signaling radio bearer,

SRB)、数据承载 (data radio bearer) DRB 或者新定义类型的 RB。

请参见图 5 和图 6, 其中图 5 示出了嵌入式 MEC 管理功能的协议架构, 图 6 示出了独立 MEC 管理功能的协议架构。

如图 5 中的 (a) 所示, 终端侧和网络侧之间建立通信接口, 对等的协议层包括物理 (physical, PHY) 层、媒体访问控制 (media access control, MAC) 层、无线链路控制 (radio link control, RLC) 层、分组数据汇聚层协议 (packet data convergence protocol, PDCP) 层和无线资源控制 (radio resource control, RRC) 层。终端侧和网络侧可重用现有 RRC 层协议交互用于实现 MEC 功能的信令。例如可新定义一些 RRC 信令, 用于交互实现 MEC 功能的相关信息 (本文中也可简称为 MEC 功能相关信息); 或者可沿用现有的 RRC 信令, 在现有的 RRC 信令中增加新的信元, 通过新的信元携带用于实现 MEC 功能相关信息。

如图 5 中的 (b) 所示, 与如图 5 中的 (a) 的不同之处在于, 可引入与 RRC 层并列的新协议层, 例如边缘应用管理协议 (edge application control, EAC) 层。即终端侧和网络侧对等的协议层包括 PHY 层、MAC 层、RLC 层、PDCP 层和 RRC 层, 以及与 RRC 层并列的 EAC 层。这种协议架构中, 可以引入新的信令承载 (radio bearer, RB) 类型, 例如 MEC-RB 类型。MEC-RB 类型与 SRB 类似, 可用于传输实现 MEC 功能相关信息。

如图 5 中的 (c) 所示, 与如图 5 中的 (a) 的不同之处在于, 可在 RRC 层上引入新协议层, 例如 EAC 层。即终端侧和网络侧对等的协议层包括 PHY 层、MAC 层、RLC 层、PDCP 层、RRC 层以及 EAC 层。这种协议架构中, EAC 层可用于 MEC 功能的协商, 协商的相关信息可作为容器 (container) 承载在 RRC 消息中, 即 MEC 功能相关信息对网络侧透明。

上述结合图 5 介绍了嵌入式 MEC 管理功能的协议架构, 下面结合图 6 介绍独立 MEC 管理功能的协议架构。

如图 6 中的 (a) 所示, 与图 5 中的 (a) 相同, 终端侧和网络侧之间对等的协议层包括 PHY 层、MAC 层、RLC 层、PDCP 层和 RRC 层。终端侧和网络侧可重用现有 RRC 层协议交互用于实现 MEC 功能的信令。例如可新定义一些 RRC 信令, 用于交互实现 MEC 功能的相关信息; 或者可沿用现有的 RRC 信令, 在现有的 RRC 信令中增加新的信元, 通过新的信元携带用于实现 MEC 功能的相关信息。不同之处在于, 网络侧接收到终端侧的 MEC 功能相关信息, 可将该 MEC 功能信息转发给 MEC 管理功能。相对的, MEC 管理功能也可将 MEC 功能信息发送给网络侧, 经由网络侧发送给终端设备。应理解, 网络侧接收到来自终端侧的 MEC 功能相关信息, 可对 MEC 功能相关信息进一步处理, 例如将 MEC 功能相关信息进行格式转化, 即将适配 RRC 层协议的 MEC 功能相关信息转化为适配于 H 接口的 MEC 功能相关信息, 之后通过 H 接口将 MEC 功能相关信息发送给 MEC 管理功能。同理, 网络侧接收到来自 MEC 管理功能的 MEC 功能相关信息, 可对 MEC 功能相关信息进一步处理, 例如将 MEC 功能相关信息进行格式转化, 即将适配于 H 接口的 MEC 功能相关信息转化为适配 RRC 层协议的 MEC 功能相关信息, 之后通过 RRC 信令将 MEC 功能相关信息发送给终端侧。

如图 6 中的 (b) 所示, 终端侧和网络侧对等的协议层包括 PHY 层、MAC 层、RLC 层、PDCP 层和 RRC 层, 以及与 RRC 层并列的 EAC 层。与如图 5 中的 (b) 的不同之处在于, 网络侧接收到来自终端侧承载于 EAC 信令的 MEC 功能相关信息, 通过 H 接口将 MEC 功能相关信息发送给 MEC 管理功能。相应的, MEC 管理功能通过 H 接口将 MEC 功

能相关信息发送给网络侧,网络侧再通过EAC信令将MEC功能相关信息发送给终端设备。

如图6中的(c)所示,终端侧和网络侧对等的协议层包括PHY层、MAC层、RLC层、PDCP层、RRC层以及EAC层。与如图5中的(c)的不同之处在于,网络侧接收到来自终端侧承载于EAC信令的MEC功能相关信息,通过H接口将MEC功能相关信息发送给MEC管理功能。相应的,MEC管理功能通过H接口将MEC功能相关信息发送给网络侧,网络侧再通过EAC信令将MEC功能相关信息发送给终端侧。

上述图3所示的MEC架构中,在网络侧和终端侧均可引入MEC管理功能,可为网络侧或终端侧提供MEC应用服务。基于该MEC网络架构,网络侧的计算任务可卸载到终端侧执行,以降低网络侧的计算负载。终端侧的计算任务也可以卸载到网络侧执行,以减轻终端侧的计算负担。且可根据网络侧和终端侧实际的计算负载以及计算能力灵活调度网络系统的计算资源,从而提高系统计算资源的利用率。

基于图3所示的MEC架构,本申请实施例还提供了网络边缘计算方法。本申请实施例提供的技术方案可应用于终端与一个网络设备连接的场景,或者也可用于终端与多个网络设备连接场景,或者还可以用于其他的场景中,例如终端与其它终端连接的场景,具体的不作限制。为了便于理解本申请实施例,下面结合具体的应用场景介绍基于图3所示的MEC架构,网络侧或终端侧如何实现计算任务的卸载。这里计算任务主要指的是计算任务,因此下文的描述中以计算任务是计算任务为例,且以MEC管理功能是独立于RAN侧设备的网元为例。

应用场景一、终端与一个网络设备连接,终端侧将计算任务卸载到网络侧。

请参见图7,为本申请实施例提供的第一种网络边缘计算方法的流程图,应用于上述应用场景一。该网络边缘计算方法的具体流程描述如下:

S71、第一网络设备确定第一MEC网元的计算能力,该第一MEC网元的计算能力用于指示第一MEC网元能够提供的MEC计算服务。

第一MEC网元是部署在网络侧的MEC网元,该第一MEC网元包括MEC管理功能和MEC计算功能。MEC计算功能能够提供MEC计算服务,MEC管理功能可用于管理MEC计算功能,例如MEC管理功能通过第一网络设备可以接收来自终端的计算任务请求,可指示MEC计算功能执行终端的部分或全部计算任务。

不同的MEC计算功能能够提供的MEC计算服务可能相同,也可能不同。例如不同MEC计算功能的硬件参数不同,导致不同MEC计算功能所具有的计算能力不同,那么不同MEC计算功能所能提供的MEC计算服务也不同。即使同一MEC计算功能在不同时间段能够提供的MEC计算服务也可能不同。例如同一个MEC计算功能在不同时间段内的计算负载不同,那么在不同时间段该MEC计算功能所具有的计算能力可能不同,该MEC计算功能所能提供的MEC计算服务也不同。在本文中,MEC计算功能的计算能力可通过计算能力参数来表征。MEC计算功能可以是硬件功能,也可以是软件功能,或者是硬件功能和软件功能的结合。对于MEC计算硬件功能,计算能力参数可包括MEC计算功能计算负载、该MEC计算功能支持的服务类型、以及该MEC计算功能的硬件参数等。示例性地,硬件参数可包括如下的一种或多种参数:

1) 吞吐量,即MEC计算功能在某一时间段内能够处理的信息量。

2) 响应时间,即从输入有效到系统产生响应的的时间。

3) 利用率,即给定的时间间隔内系统被实际使用的时间所占的百分百。

4) 处理机字长, 即运算器中一次能够运算二进制数的位数。

5) 总线宽度, 即中央处理单元 (central processing unit, CPU) 中连接运算器与存储器的内部总线二进制位数。

6) 存储器带宽, 即单位时间内从存储器中读出的二进制信息量, 一般用字节数/秒表示。

7) CPU 主频, 即主时钟产生信号的固定频率, 其倒数是 CPU 的时钟周期。

8) CPU 执行时间= CPU 时钟周期数乘以 CPU 时间。

9) 每秒执行多少百万次定点指令 (million instructions per second, MIPS)。

10) 每秒钟浮点操作的次数 (floating-point operations per second, FLOPS)。

对于 MEC 计算软件功能, 计算能力参数也可包括 MEC 计算功能计算负载、该 MEC 计算功能支持的服务类型、软件参数等。示例性的, 软件参数还可以包括如下的一种或多种:

1) 吞吐量, 即 MEC 计算功能在某一时间段内能够处理的信息量。

2) 响应时间, 即从输入有效到系统产生响应的的时间。

3) 最大用户访问量, 指在同一时间段内访问系统的用户数量。

4) 最大任务处理量, 指在同一时间段内系统并行处理的任务数量。

5) 性能计数器, 描述服务器或操作系统性能的一些数据指标。例如, 对 Windows 系统来说, 使用内存数 (memory in usage), 进程时间 (total process time) 等都是常见的计数器。

6) 思考时间, 也被称为“休眠时间”, 从业务的角度来说, 这个时间指的是用户在进行操作时, 每个请求之间的间隔时间。

7) TPS (Transaction per second), 每秒钟系统能够处理的交易或者事务的数量。

8) 点击率, 每秒钟用户向服务器提交的超文本传输协议(hyper text transfer protocol, HTTP)请求数量。

应理解, 终端可请求第一网络设备将终端的计算任务卸载到 MEC 计算功能。如果 MEC 计算功能具有的计算能力不足以满足终端的计算任务的计算需求, 而第一网络设备将终端的计算任务卸载到 MEC 计算功能, 显然会导致 MEC 计算功能执行计算任务的失败。

为此, 第一网络设备可确定网络边缘的计算能力。在该场景中, 以网络边缘节点是第一 MEC 网元为例, 即第一网络设备可确定第一 MEC 网元的计算能力。这样第一网络设备清楚地知道 MEC 计算功能具有的计算能力, 从而合理地分配终端的计算任务, 以保证 MEC 计算功能执行计算任务的成功率。第一网络设备可主动获取第一 MEC 网元的计算能力, 第一 MEC 网元也可主动告知第一网络设备该第一 MEC 网元的计算能力。下面列举几种第一网络设备确定第一 MEC 网元的计算能力的方式。

确定方式一, 请参见图 7a, 为第一网络设备确定第一 MEC 网元的计算能力的一种流程示意图。

S701a、第一网络设备向第一 MEC 网元发送第一请求消息, 该第一请求消息用于请求或指示查询第一 MEC 网元的计算能力。

第一网络设备需要获取第一 MEC 网元的计算能力, 可向第一 MEC 网元发送第一请求消息, 以请求获取第一 MEC 网元的计算能力。本申请实施例对第一请求消息的名称不作限制, 例如第一请求消息也可称为 MEC 能力查询(MEC capability enquiry)请求消息。

S702a、第一 MEC 网元向第一网络设备发送第一请求响应消息，该第一请求响应消息包括第一能力信息，该第一能力信息用于指示第一 MEC 网元的计算能力。

第一 MEC 网元接收到第一请求消息，响应该第一请求消息，可向第一网络设备发送第一请求响应消息。即第一 MEC 网元向第一网络设备发送针对第一请求消息的响应消息。
5 第一请求响应消息携带第一能力信息，该第一能力信息可用于指示第一 MEC 网元的计算能力。示例性的，该第一能力信息可包括上述的一种或多种计算能力参数。第一网络设备获取第一能力信息，可根据第一能力信息确定第一 MEC 网元的计算能力。

需要说明的是，本申请实施例对第一请求响应消息的名称不作限制，例如第一请求响应消息也可称为 MEC 能力应答（MEC capability ACK）消息。

10 确定方式二、请继续参见图 7a，示出了第一网络设备确定第一 MEC 网元的计算能力的另一种流程。

S701b、第一 MEC 网元向第一网络设备发送 H 接口建立请求消息，该 H 接口建立请求消息用于请求与第一 MEC 网元建立 H 接口连接。

15 S702b、第一网络设备向第一 MEC 网元发送 H 接口建立请求响应消息，该 H 接口建立请求响应消息包括第一能力信息，该第一能力信息用于指示第一 MEC 网元的计算能力。

第一 MEC 网元接收第一网络设备的 H 接口建立请求消息，可以将第一能力信息携带在 H 接口建立请求响应消息中发送给第一网络设备。该确定方式，即第一网络设备只要与第一 MEC 网元请求建立连接，第一 MEC 网元就向第一网络设备通知该第一 MEC 网元的计算能力，这样不需要针对 MEC 能力交互额外的信令，以尽量节约信令的开销。

20 上述第一网络设备通过确定方式一和确定方式二可主动获取第一 MEC 网元的计算能力。作为一种可替换的方式，即第三确定方式，第一 MEC 网元可主动告知第一网络设备该第一 MEC 网元的计算能力。

确定方式三，请继续参见图 7a，示出了第一网络设备确定第一 MEC 网元的计算能力的再一种流程。

25 S701c、第一 MEC 网元向第一网络设备发送第一配置更新消息，该第一配置更新消息包括第一能力信息，该第一能力信息用于指示第一 MEC 网元的计算能力。

第一 MEC 网元可以周期性地通知第一网络设备，该第一 MEC 网元的计算能力。例如第一 MEC 网元可周期性地向第一网络设备发送第一配置更新消息，该第一配置更新消息可携带第一能力信息，该第一能力信息可包括上述的一种或多种计算能力参数。或者，第一 MEC 网元的计算能力发生变化后，第一 MEC 网元可将变化后的计算能力告知第一网络设备。例如第一 MEC 网元的计算能力发生变化后，向第一网络设备发送第一配置更新消息，该第一配置更新消息可携带第一能力信息，该第一能力信息可包括上述的一种或多种计算能力参数。第一网络设备获取第一能力信息，可根据第一能力信息确定第一 MEC 网元的计算能力。由于第一 MEC 网元主动告知第一网络设备，第一 MEC 网元更新后的计算
30 能力，所以可避免第一网络设备发送不必要的第一请求消息，以尽量节约信令的开销。

需要说明的是，本申请实施例对第一配置更新消息的名称不作限制。由于第一配置更新消息是用于通知 MEC 的计算能力的，所以第一配置更新消息也可称为 MEC 能力更新消息。

S702c、第一网络设备向第一 MEC 网元发送第一配置更新响应消息。

40 第一网络设备根据第一配置更新消息确定第一 MEC 网元的计算能力之后，可向第一

MEC 网元反馈确认消息，以告知第一 MEC 网元，第一网络设备是否确定第一 MEC 网元的计算能力。例如第一网络设备可向第一 MEC 网元发送第一配置更新响应消息，该第一配置更新响应消息是针对第一配置更新消息的响应消息。如果第一配置更新响应消息指示第一网络设备不确定第一 MEC 网元的计算能力，例如虽然第一 MEC 网元向第一网络设备发送了第一配置更新消息，但是第一配置更新消息中不包括计算能力信息，又或者第一网络设备中保存的第一 MEC 网元的计算能力信息的时效性已过。那么第一 MEC 网元可再次向第一网络设备发送第一配置更新响应消息，以尽量保证第一网络设备确定第一 MEC 网元的计算能力。

需要说明的是，确定方式一中的第一请求消息可以是专用消息，例如新定义的 EAC 消息。或者第一请求消息可复用已有的消息，例如第一请求消息是 RRC 消息。同理，确定方式中的第一配置更新消息也可复用已有的消息，例如第一配置更新消息是 RRC 消息。或者第一配置更新消息也可以是专用消息，例如新定义的 EAC 消息。

应理解，如果终端没有将计算任务卸载的需求，第一网络设备可不必获取第一 MEC 网元的计算能力。

在一些实施例中，第一网络设备可基于终端的需求，获取第一 MEC 网元的计算能力，以尽量节约信令的开销。即 S701a 可在 S71? 1 之后执行。

S711、终端向第一网络设备发送第二请求消息，该第二请求消息用于指示第一网络设备获取网络边缘的计算能力。

终端可向第一网络设备发送第二请求消息，该第二请求消息可用于指示第一网络设备获取网络边缘的计算能力。或者第二请求消息可用于请求获取网络边缘的计算能力。

终端有卸载计算任务的需求时，向第一网络设备发送第二请求消息。而第一网络设备接收到第二请求消息，向第一 MEC 网元发送第一请求消息，可避免向第一 MEC 网元发送不必要的消息，节约信令的开销。需要说明的是，本申请实施例对该第二请求消息的名词不作限制，例如第二请求消息可称为网络边缘能力查询（NW edge capability enquiry）请求消息。第二请求消息可以是新定义的消息，例如 EAC 消息。又例如，第二请求消息可复用已有的消息，例如第二请求消息可以是消息 1（Msg1）或者消息 3（Msg3），或者，终端对网络边缘计算能力的获取方式可采用按需系统消息（On demand SI）的形式发送给第一网络设备。

需要说明的是，S711 不是必须执行的步骤，因此在图 7 中以虚线进行示意。如果 S711 是需要执行的步骤，本申请实施例不限制 S711 与 S71 的先后执行顺序。例如 S71 可在 S711 之前执行，也可以在 S711 之后执行。

第一网络设备确定网络侧的计算能力，可告知终端网络侧的计算能力。这样终端可参考网络侧的计算能力，合理地请求网络侧执行终端的部分计算任务。

S72、第一网络设备发送第一能力信息，该第一能力信息用于指示第一 MEC 网元的计算能力。

应理解，如果终端执行 S711，那么第一网络设备获取第一 MEC 网元的计算能力之后，可向该终端发送第一能力信息。如果不执行 S711，第一网络设备自行获取第一 MEC 网元的计算能力之后，可广播第一能力信息，这样处于网络中的任意终端都可以知道第一 MEC 网元的计算能力，从而按需请求网络侧执行终端的计算任务。

在一些实施例中，第一网络设备可将第一能力信息携带在新定义的 EAC 消息（例如

称为网络边缘能力消息)中发送给终端。

需要说明的是,终端根据第一能力信息除了确定要卸载的计算任务,还可以根据第一能力信息选择要驻留的小区。例如终端优先选择支持该终端所需 MEC 计算服务的小区进行驻留,这样当终端有 MEC 计算服务需要处理时,不需要重选到能够支持其计算服务的小区,可减少接入时延。当然如果终端确定终端当前驻留的小区支持的 MEC 计算服务不满足终端的 MEC 计算服务需求,那么而邻区支持的 MEC 计算服务能够满足终端的 MEC 计算服务需求,且邻区满足小区重选条件,那么终端可重选到邻区。或者终端重选小区时,可在满足重选条件的小区列表中查找能够满足 MEC 计算服务需求的小区进行重选。

终端确定了网络边缘计算能力,可请求网络侧执行终端的计算任务,以减轻终端的计算负担。在可能的实现方式中,终端可向第一网络设备请求网络侧执行终端的计算任务;或者终端也可以向第一 MEC 网元请求执行终端的计算任务。下面分别介绍终端请求网络侧执行终端的计算任务的两种可能的实现方式。

作为一种可能的实现方式,在示例一中,终端向第一网络设备请求网络侧执行终端的计算任务,具体流程可参考如下的 S731a-S735a。应理解,S731a-S735a 所涉及的信令交互可基于图 6 中的(a)所示的协议层架构。

S731a、终端向第一网络设备发送第一任务请求消息,该第一任务请求消息用于请求网络侧执行终端的至少一个计算任务。

终端需要网络侧执行终端的至少一个计算任务,终端可向第一网络设备发送第一任务请求消息。本申请实施例对第一任务请求消息的名词不作限制,例如该第一任务请求消息可称为计算任务请求(MEC Task Req)消息。第一任务请求消息用于请求网络侧执行终端的至少一个计算任务,那么第一任务请求消息可包括与至少一个计算任务的相关参数,以使得第一网络设备根据该相关参数合理地将至少一个计算任务分配给 MEC 计算功能。例如第一任务请求消息可包括 MEC 计算服务类型信息、资源需求信息、任务连续性保证需求信息、任务关键绩效指标(key performance indicator, KPI)需求信息等一种或多种信息。应理解,MEC 计算服务类型信息可用于指示 MEC 计算服务类型,也就是计算任务的类型,例如人工智能类型、微服务类型等。应理解,一个计算业务可由多个微服务组成。资源需求信息用于指示计算至少一个计算任务所需要的资源,例如存储资源,图形处理单元(graphics processing unit, GPU)资源等。任务连续性保证需求信息可用于指示任务连续性的保证需求,即任务中断,是否需要从任务断开处继续执行后续过程,还是需要重新对该任务进行处理。任务 KPI 需求信息可用于指示该任务的性能指标需求,例如时延需求、精度需求等。

S732a、第一网络设备向第一 MEC 网元发送第一任务请求消息,第一 MEC 网元接收该第一任务请求消息。

第一网络设备可将第一任务请求消息转发给第一 MEC 网元,以请求第一 MEC 网元执行终端的至少一个计算任务。

S733a、第一 MEC 网元根据第一任务请求消息进行准入控制。

第一 MEC 网元接收到第一任务请求消息,可根据第一 MEC 网元支持的 MEC 计算服务类型、计算负载等情况对终端的请求进行准入控制。

S734a、第一 MEC 网元向第一网络设备发送第一任务请求响应消息,所述第一任务请求响应消息用于通知所述至少一个计算任务的准入信息。

如果第一 MEC 网元确定能够执行至少一个计算任务,第一 MEC 网元可告知第一网络设备,该第一 MEC 网元确定执行至少一个计算任务。例如第一 MEC 网元可向第一网络设备发送第一任务请求响应消息。应理解,该第一任务请求响应消息是针对第一任务请求消息的响应消息。如果第一 MEC 网元支持至少一个计算任务的执行,那么该第一任务请求响应消息可包括至少一个计算任务的准入信息。

如果第一 MEC 网元不支持至少一个计算任务的执行,那么第一任务请求响应消息可用于指示第一 MEC 网元不支持至少一个计算任务的执行,以告知第一网络设备,第一 MEC 网元不支持至少一个计算任务的执行,以避免第一网络设备将至少一个计算任务的数据发送给第一 MEC 网元,引入不必要的开销。

S735a、第一网络设备向终端发送第一任务请求响应消息。

第一网络设备接收到第一任务请求响应消息,可将该第一任务请求消息转发给终端,即将所述至少一个计算任务的准入信息发送给终端。

作为另一种可能的实现方式,在示例二中,终端向第一 MEC 网元请求执行终端的计算任务,具体流程可参考如下的 S731b-S733b。应理解,S731a-S734a 所涉及的信令交互可基于图 6 中的 (b) 所示的协议层架构或图 6 中的 (c) 所示的协议层架构。

S731b、终端向第一 MEC 网元发送第一任务请求消息,该第一任务请求消息用于请求网络侧执行终端的至少一个计算任务。

S732b、第一 MEC 网元根据第一任务请求消息进行准入控制。

S733b、第一 MEC 网元向第一网络设备发送第一任务请求响应消息。

上述的 S731b 与 S731a 类似,不同之处在于,S731b 基于图 6 中的 (b) 所示的协议层架构或图 6 中的 (c) 所示的协议层架构实现,S731a 基于图 6 中的 (a) 所示的协议层架构实现,即 S731b 中第一任务请求消息可以是新定义的 EAC 消息,S731a 中第一任务请求消息可以是 RRC 消息。应理解,在 S731b 中,第一网络设备收到第一任务请求消息后将该消息转发给第一 MEC 网元,第一网络设备并不解析第一任务请求消息的内容,而是由第一 MEC 网元对第一请求消息进行解析和处理。具体 S731b 的实现可参考 S731a 的描述,这里不再赘述。

同理,S732b 的实现同 S733a 类似,可参考 S733a 的描述,这里不再赘述。S733b 的实现同 S734a 类似,可参考 S734a 的描述,这里不再赘述。

该实现方式中,终端可以与第一 MEC 网元建立控制面和数据面的传输通道,从而使终端可与第一 MEC 网元直接交互计算任务数据,不需要借助第一网络设备的转发,可降低交互时延。

S74、终端与第一 MEC 网元建立连接,并传输计算任务数据。

终端接收到来自第一 MEC 网元的第一任务请求响应消息,终端可请求与第一 MEC 网元建立连接。终端与第一 MEC 网元建立连接后,可相互传输计算任务的相关数据(下文中称为计算任务数据)。

示例性的,终端和第一 MEC 网元可建立控制面的传输通道,即终端可通过控制面的传输通道与第一 MEC 网元交互计算任务数据。例如计算任务数据承载于 SRB;或者计算任务数据承载于针对 MEC 业务新定义的无线承载 RB,例如 MEC-RB;或者计算任务数据可承载于 MEC 接口的应用层控制信息,例如可称为 HAP 消息。终端和第一 MEC 网元可建立数据面的传输通道,即终端可通过数据面的传输通道与第一 MEC 网元交互计算任务

数据。例如计算任务数据承载于 DRB，或者新定义的无线承载，例如 MEC-RB，或者第一网络设备到第一 MEC 网元的 H 接口数据通道。

该应用场景一支持示例一和示例二，即终端可如示例一或示例二与第一 MEC 网元交互执行相关信令，更为灵活。

5 应用场景二、终端与多个网络设备连接，网络侧包括多个 MEC 网元。终端侧可将计算任务卸载到网络侧的一个或多个 MEC 网元。

请参见图 8，为本申请实施例提供的第二种网络边缘计算方法的流程图，应用于上述应用场景二。图 8 以终端和 2 个网络设备连接，且这 2 个网络设备分别与 MEC 网元连接，即图 8 以包括 2 个 MEC 网元（第一 MEC 网元和第二 MEC 网元）为例。那么终端可将计算任务卸载到网络侧的第一 MEC 网元，也可以将计算任务卸载到第二 MEC 网元。该网络边缘计算方法的具体流程描述如下：

S811、第一网络设备确定第一 MEC 网元的计算能力，第一 MEC 网元的计算能力用于指示第一 MEC 网元能够提供的 MEC 计算服务。

15 S812、第二网络设备确定第二 MEC 网元的计算能力，第二 MEC 网元的计算能力用于指示第二 MEC 网元能够提供的 MEC 计算服务。

S811 的实现和 S812 的实现与前述 S71 的实现相同，具体可参考前述 S71 的相关描述，这里不再赘述。另外本申请实施例对 S811 和 S812 的执行顺序不作限制，即 S811 可在 S812 之前执行，S811 也可在 S812 之后执行。

S82a、第一网络设备和第二网络设备交互各自所连接的 MEC 网元的计算能力。

20 应理解，如果终端请求第一 MEC 网元执行计算任务，但是第一 MEC 网元的计算能力不足以满足该计算任务的计算需求，例如第一 MEC 网元的计算负载较重，此时第一网络设备指示第一 MEC 网元执行该计算任务，会导致该计算任务的失败。所以第一网络设备确定第一 MEC 网元无法执行该计算任务，可从网络中确定其它能够执行该计算任务的其它 MEC 网元，例如第二 MEC 网元。

25 示例的，第一网络设备可从其他网络设备，例如第二网络设备确定第二 MEC 网元的计算能力。例如第一网络设备可向第二网络设备发送用于请求获取第二 MEC 网元的计算能力的消息，第二网络设备接收该消息之后，回复第一网络设备，第二 MEC 网元的计算能力。类似的，第二网络设备也可以获取第一 MEC 网元的计算能力。即网络中多个网络设备中两两网络设备间可获知彼此的 MEC 网元的计算能力。这样当某个网络设备确定不满足终端的计算任务需求，可请求其它网络设备执行计算任务，以尽量保证终端的计算任务正常执行。

30 应理解，S82a 中第一网络设备通过与第二网络设备交互，确定第二网络设备侧的 MEC 的计算能力。即多个网络设备可直接交互彼此的计算能力，以确定彼此的计算能力。作为一种可替换的实现方式，位于第一网络设备侧的第一 MEC 网元可与位于第二网络设备的第二 MEC 网元交互彼此的计算能力，这样第一网络设备从第一 MEC 网元可间接获取第二 MEC 网元的计算能力，第二网络设备从第二 MEC 网元也可获取第一 MEC 网元的计算能力。

S82b、第一 MEC 网元和第二 MEC 网元交互各自的计算能力。

40 与 S82a 类似，当第一 MEC 网元需要获取第二 MEC 网元的计算能力，第一 MEC 网元可向第二 MEC 网元发送用于请求获取第二 MEC 网元的计算能力的消息，第二 MEC 网

元接收该消息之后,向第一 MEC 发送请求响应消息,该请求响应消息包括第二 MEC 网元的计算能力相关信息。类似的,第二 MEC 网元也可以获取第一 MEC 网元的计算能力。即网络中多个 MEC 网元中两两 MEC 网元间可获知彼此的计算能力。这种情况下,第一网络设备可请求从第一 MEC 网元获取第二 MEC 网元的计算能力。第二网络设备可请求从第二 MEC 网元获取第一 MEC 网元的计算能力。

应理解,由于第一 MEC 网元和第二 MEC 网元是独立的逻辑实体,所以第一 MEC 网元可与第一网络设备通过 H 接口交互消息或者信息。同理,第二 MEC 网元可与第二网络设备通过 H 接口交互消息或者信息。第一 MEC 网元可与第二 MEC 网元通过 X 接口交互消息或信息。如果第一 MEC 网元和第二 MEC 网元是嵌入式结构,即第一 MEC 网元设置在第一网络设备内,第二 MEC 网元设置在第二网络设备内,那么第一 MEC 网元和第二 MEC 网元可通过 Xn 接口交互消息或信息。

需要说明的是,S82a 和 S82b 的先后执行顺序不作限制,即 S82a 可在 S82b 之前执行,也可以在 S82b 之后执行。且 S82a 和 S82b 不必都执行。作为一种可能的实现方式,在示例三中,终端通过第一网络设备向第二网络设备请求网络侧执行终端的计算任务;本申请实施例执行了 S82a,那么后续可执行 S831a-S837a。作为另一种可能的实现方式,在示例四中,终端通过第一 MEC 网元向第二 MEC 网元请求网络侧执行终端的计算任务;本申请实施例执行了 S82b,那么后续可执行 S831b-S835b。

S831a、终端向第一网络设备发送前述第一任务请求消息,第一网络设备接收该第一任务请求消息。

具体的,S831a 的实现与前述 S73a 的实现相同,可参考前述 S73a 的相关描述,这里不再赘述。

第一网络设备接收到第一任务请求消息,如果第一 MEC 网元的计算能力满足终端的计算任务的计算需求,例如第一 MEC 处于计算过载的情况,那么第一网络设备可将该第一任务请求消息发送给第一 MEC 网元,即执行前述的 S732a。与图 7 所示流程的不同之处在于,如果第一 MEC 网元的计算能力不满足终端的计算任务的计算需求,第一网络设备可请求其它 MEC 网元执行终端的计算任务。即第一网络设备可执行 S832a。

S832a、第一网络设备向第二网络设备发送第一任务请求消息,第二网络设备接收该第一任务请求消息。

由于第一网络设备与其它网络设备可交互各自的计算能力,因此当第一网络设备确定第一 MEC 网元的计算能力不满足终端的计算任务的计算需求,可从其它网络设备中确定能够满足终端的计算任务计算需求的 MEC 网元,例如第二 MEC 网元。或者也可以理解为,第一网络设备可从其它网络设备中确定目标网络设备,例如第二网络设备。应理解,与目标网络设备连接的 MEC 网元能够满足终端的计算任务的计算需求。这种情况下,第一网络设备可将来自终端的第一任务请求消息转发给第二网络设备。

作为一种可替换的实现方式,第一网络设备也可以将第一任务请求消息发送给其它网络设备中的某个网络设备,例如第三网络设备。如果第三网络设备连接的第三 MEC 网元能够满足终端的计算任务的计算需求,那么第三网络设备请求该第三 MEC 网元执行终端的计算任务。如果第三 MEC 网元不满足终端的计算任务的计算需求,那么第三网络设备可从其它网络设备中再次确定目标网络设备,将第一任务请求消息发送给目标网络设备。或者第三网络设备将第一任务请求消息再次转发给其它网络设备中的另一网络设备。

S833a、第二网络设备向第二 MEC 网元发送第一任务请求消息,第二 MEC 网元接收 该第一任务请求消息。

具体的, S833a 的实现与前述 S732a 的实现相同,可参考前述 S732a 的相关描述,这里不再赘述。

5 S834a、第二 MEC 网元根据第一任务请求消息进行准入控制。

具体的, S834a 的实现与前述 S733a 的实现相同,可参考前述 S733a 的相关描述,这里不再赘述。

S835a、第二 MEC 网元向第二网络设备发送第一任务请求响应消息,第二网络设备接收该第一任务请求响应消息。

10 具体的, S834a 的实现与前述 S733a 的实现相同,可参考前述 S733a 的相关描述,这里不再赘述。

S836a、第二网络设备向第一网络设备发送第一任务请求响应消息,第一网络设备接收该第一任务请求响应消息。

15 第二网络设备通过与第一网络设备的接口发送第一任务请求响应消息,即第一任务请求响应消息可以是第二网络设备与第一网络设备之间的接口消息。

S837a、第一网络设备将第一任务请求响应消息发送给终端。

20 具体的, S837a 的实现与前述 S735a 的实现相同,可参考前述 S735a 的相关描述,这里不再赘述。与 S735a 的不同之处在于,第一任务请求响应消息除了包括终端的计算任务的准入信息之外,第一任务请求响应消息还可以包括第二网络设备的相关配置信息。该相关配置信息可包括第二网络设备的空口配置信息,例如随机接入信道(andom access channel, RACH)信息,以供终端可通过切换或直接连接等方式接入第二网络设备。

通过示例三上述步骤,即第一网络设备与第二网络设备交互彼此的计算能力,这样第一网络设备可确定能够执行终端的计算任务的其它 MEC 网元,从而请求第二网络设备侧执行终端的计算任务,以尽量为终端侧提供就近计算服务,提升用户体验。

25 S831b、终端向第一 MEC 网元发送前述第一任务请求消息,第一 MEC 网元接收该第一任务请求消息。

S832b、第一 MEC 网元向第二 MEC 网元发送第一任务请求消息,第二 MEC 网元接收该第一任务请求消息。

S833b、第二 MEC 网元根据第一任务请求消息进行准入控制。

30 S834b、第二 MEC 网元向第一 MEC 网元发送第一任务请求响应消息,第一 MEC 网元接收该第一任务请求响应消息。

S835b、第一 MEC 网元向第一网络设备发送第一任务请求响应消息,第一网络设备接收该第一任务请求响应消息。

35 S836b、第一网络设备向终端发送第一任务请求响应消息,终端接收该第一任务请求响应消息。

具体的, S831b 的实现与前述 S731b 的实现相同,可参考前述 S731b 的相关描述; S833b 的实现与前述 S732b 的实现相同,可参考前述 S732b 的相关描述; S835b 的实现与前述 S733b 的实现相同,可参考前述 S733b 的相关描述; S836b 的实现与前述 S837a 的实现相同,可参考前述 S837a 的相关描述,这里不再赘述。S832b 和 S834b 中第一 MEC 网元和第二 MEC 网元可通过第一 MEC 网元和第二 MEC 网元之间的 X 接口交互第一任务请求消息或第一

40

任务请求响应消息。或者第一 MEC 网元和第二 MEC 网元可通过第一网络设备和第二网络设备之间的 Xn 接口交互第一任务请求消息或第一任务请求响应消息。

需要说明的是, S834b 中第一任务请求响应消息可携带第二网络设备的相关配置信息, S835b 中第一任务请求响应消息可携带第一网络设备的相关配置信息和第二网络设备的相关配置信息。

通过本申请实施例的步骤, 即第一 MEC 网元与第二 MEC 网元交互彼此的计算能力, 这样终端可直接向第一 MEC 网元请求执行计算任务, 不需要借助网络设备, 可节省信令的开销。

S84、终端与第二 MEC 网元建立连接, 并传输计算任务数据。

通过本实施例上述步骤, 终端将计算任务数据发送给第一 MEC 网元, 第一 MEC 网元再通过与第二 MEC 网元之间的 X 接口将计算任务数据发送给第二 MEC 网元, 不需要借助第一网络设备和第二网络设备转发, 这样可减少切换或者双连接带来的信令开销和时延。

如上介绍了基于图 3 所示的 MEC 架构, 网络侧执行终端侧的计算任务的流程。随着终端计算能力的提升, 一些终端具有 MEC 计算功能, 也能够提供计算服务。为了便于描述, 本申请实施例根据终端是否能够提供计算服务可将终端分为两类, 例如第一类终端和第二类终端。第一类终端指的是不能够提供计算服务的终端 (可记为 C-UE), 相对而言, 第二类终端指的是能够提供计算服务的终端 (可记为 H-UE)。第一类终端可认为是需要获取计算服务的终端 (consumer UE, 可记为 C-UE), 第二类终端可认为是用于提供计算服务的终端 (helper UE, 可记为 H-UE) 这种情况下, 第一类终端可请求第二类终端执行第一类终端的计算服务。下面结合附图介绍第二类终端执行第一类终端的计算服务的具体流程。

请参见图 9, 为本申请实施例提供的第三种网络边缘计算方法的流程图。该网络边缘计算方法的具体流程描述如下:

S91、第二终端和第一 MEC 网元交互所述第二终端的计算能力信息。

应理解, 第二终端属于第二类终端类型, 即第二终端可以是网络中任意一个 H-UE。虽然第二终端能够提供计算服务, 但是第一终端并不知道第二终端的存在。应理解, 第一终端属于第一类终端类型。如果第一终端直接请求另一终端执行第一终端的计算任务, 会导致失败。因此, 在本申请实施例中, 第二终端在第一 MEC 网元上注册自身的计算能力信息, 第一 MEC 网元可管理第二终端。

作为一种可实现的方式, 第二终端可向第一 MEC 网元发送第一注册请求消息, 该第一注册请求消息可用于请求注册第二终端的计算能力。第一注册请求消息可携带第二终端的注册信息, 例如第二终端的身份标识信息, 以及第二终端的计算能力信息, 例如包括上述的一种或多种计算能力参数。第一 MEC 网元接收到第一注册请求消息之后, 可存储第二终端的注册信息, 并向第二终端反馈针对第一注册请求消息的响应消息 (可称为第一注册请求响应消息)。

当然, 第二终端的计算能力可能发生变化, 当第二终端的计算能力发生变化, 第二终端可将变化之后的计算能力告知第一 MEC 网元, 以保证第一 MEC 网元存储第二终端的最新计算能力。示例性的, 第二终端的计算能力发生变化, 第二终端可向第一 MEC 网元发生计算能力更新信息, 该计算能力更新信息可携带如上的一种或多种计算能力参数。例如第二终端的计算负载发生变化, 那么计算能力更新信息可携带更新后的计算负载。应理解,

第一 MEC 网元可存储多个第二终端的计算能力。第一 MEC 网元知晓具有计算能力的第二终端，当接收到第一终端的计算任务请求，第一 MEC 网元可确定自身执行第一终端的计算任务，也可以请求第二终端执行第一终端的计算任务。

S92、第一终端向第一网络设备发生第一任务请求消息，第一网络设备接收该第一任务请求消息。

应理解，第一终端属于第一类终端类型，即第一终端可以是网络中任意一个 C-UE。

S93、第一网络设备向第一 MEC 网元转发第一任务请求消息。

S92 的具体实现同 S73a 相同，S93 的具体实现同 S732a 相同，具体可参考前述 S73a 和 S732a 的相关描述，这里不再赘述。

S94、第一 MEC 网元向第一网络设备发送目标终端的相关信息，第一网络设备接收该相关信息。

第一 MEC 网元接收到第一任务请求消息，可根据第一任务请求消息中至少一个计算任务的计算需求，从多个第二类终端中确定满足该计算需要的目标终端。之后第一 MEC 网元将该目标终端的相关信息，例如目标终端的身份标识信息发送给第一网络设备，以告知第一网络设备，该目标终端可执行第一终端的计算任务。第一网络设备可请求目标终端执行第一终端的计算任务。为便于描述下文中以第二类终端为例。

应理解，S93 所涉及的信令交互可基于图 6 中的(b)所示的协议层架构或图 6 中的(c)所示的协议层架构。例如目标第二类终端的相关信息可携带在 H 接口请求消息中，也可以携带在新定义的 EAC 消息。

S95、第一网络设备将第一任务请求消息转发给第二终端。

应理解，在该步骤中，第二终端是多个第二类终端中的一个 H-UE，该 H-UE 是在上述步骤 S94 中第一网络设备获知的目标终端。在本实施例后续步骤中，第二终端对应为该目标终端。

S96、第二终端根据第一任务请求消息进行准入控制。

S96 的具体实现同 S733a 相同，具体可参考前述 S733a 的相关描述，这里不再赘述。

S97、第二终端向第一网络设备发送第一任务请求响应消息，第一网络设备接收该第一任务请求响应消息。

如果第二终端确定能够执行至少一个计算任务，第二终端可告知第一网络设备，第二终端确定执行至少一个计算任务。例如第二终端可向第一网络设备发送第一任务请求响应消息。应理解，第一任务请求响应消息是针对第一任务请求消息的响应消息。如果第二终端支持至少一个计算任务的执行，那么第一任务请求响应消息可包括至少一个计算任务的准入信息。

由此可见，第二终端支持至少一个计算任务的执行，可以使网络将部分的计算任务卸载到第二类终端执行，节省了网络侧资源，也充分利用了终端侧的资源。

下面分别介绍第一终端和第二终端通过不同方式连接时的三种可能的实现方式。

作为一种可能的实现方式，在示例五中，第一网络设备作为中继实现第一终端和第二终端之间的通信，具体流程可参考如下的 S98a-S911a。

S98a、第一网络设备确定作为第一终端和第二终端之间的中继，用于传输第一终端和第二终端之间的计算任务数据。

S99a、第一网络设备向第二终端发送 RRC 重配置消息，该 RRC 重配置消息包括用于

计算任务数据的承载信息，例如 DRB 相关的配置信息，以使第一网络设备和第二终端之间可以交互计算任务数据。

S910a、第一网络设备向第一终端发送第一任务请求响应消息以及 RRC 重配置消息，以使第一网络设备和第一终端之间可以交互计算任务数据。

5 S911a、第一终端与第二终端通过第一网络设备交互计算任务数据。

通过示例五上述步骤，第一网络设备作为第一终端和第二终端之间的中继，将来自第一终端的计算任务数据转发给第二终端，也将来自第二终端的计算任务数据等转发给第一终端。这样，利用网络侧的中继功能，实现第一终端将计算任务数据传输给第二终端。

需要说明的是，图 9 所示的流程也可以适用于多网络设备的中继场景，即第一终端通过多个网络设备的中继，将消息或信息传输给第二终端。与通过一个网络设备中继的不同之处在于，多个网络设备间针对 MEC 业务需要建立 Xn 通道进行数据的转发。

另外，本申请实施例对第二终端的具体实现形式不作限制。根据第二终端的不同，图 9 所示流程也有改变。

作为另一种可能的实现方式，在示例六中，第二终端可以是集成接入回传（integrated access backhaul, IAB）设备，具体流程可参考如下的 S98b-S910b。即如果第二终端是 IAB 设备，那么图 9 所示的网络边缘计算方法包括 S91-S97，以及 S98b-S910b。

S98b、第一网络设备向第一终端发送第一任务请求响应消息以及 RRC 重配置消息。

与 S910a 的不同之处在于，该第一任务请求响应消息包括第二终端的相关信息。由于第二终端为 IAB 设备，那么第一终端与第二终端之间可不需要通过第一网络设备进行通信。第一终端可通过第二终端的相关信息与第二终端建立连接，进而实现计算任务数据的直接交互。

S99b、第一终端与第二终端建立连接。

S910b、第一终端与第二终端交互计算任务数据。

在 IAB 设备设置 MEC 计算功能，通过 S98b-S910b，第一终端可随机接入到第二终端，能够实现将计算任务数据传输给第二终端。应理解，第二终端，即 IAB 设备需要配置配备 CU-UP 协议层，例如 PDCP 层和 SDAP 层。计算任务数据传输给 IAB 设备的 DU，例如传输给 DU 的 PHY 层或 MAC 层或 RLC 层后，再经由 CU-UP 协议层传输给 IAB 设备的应用层进行处理。

通过示例六上述步骤，第二设备可以是 IAB 设备，这样第一设备可随机接入到第二终端，实现与第二终端的连接，进而交互计算任务数据。不需要借助其他设备，例如网络设备，这样可减少第二终端执行第一终端的计算任务的信令开销。

作为又一种可能的实现方式，在示例七中，第二终端和第一终端具有 D2D 功能，即第二终端和第一终端可通过 D2D 的方式交互计算任务数据。这种情况下，具体流程可参考如下的 S98c-S910c。即如果第二终端和第一终端可以通过 D2D 方式交互计算任务数据，那么图 9 所示的网络边缘计算方法包括 S91-S97，以及 S98c-S910c。

S98c、第一网络设备向第一终端发送第一任务请求响应消息以及 RRC 重配置消息。

与 S910a 的不同之处在于，该第一任务请求响应消息包括第二终端的相关信息。由于第二终端为具有 D2D 功能的设备，那么第一终端与第二终端之间可不需要通过第一网络设备进行通信。第一终端可通过第二终端的相关信息与第二终端建立 D2D 连接，进而实现计算任务数据的直接交互。

S99c、第一终端与第二终端建立 D2D 连接。

与前述 S99b 的不同之处在于，第一终端和第二终端建立 D2D 连接。

S910c、第一终端与第二终端交互计算任务数据。

第一终端和第二终端建立 D2D 连接之后，第一终端可通过 D2D 通道将计算任务数据发送给第二终端。

通过本示例七上述步骤，第一终端和第二终端可直接建立 D2D 连接，进而交互计算任务数据。不需要借助其他设备，例如网络设备，这样可减少第二终端执行第一终端的计算任务的信令开销。

上述图 9 示出了第一终端通过第一网络设备将至少一个计算任务卸载到第二终端的流程。即第一终端不知道第二终端，第一终端通过第一网络设备告知能够提供计算服务的第二终端，以请求第二终端执行第一终端的计算任务。应理解，如果第一终端知道能够提供计算服务的第二终端，那么第一终端也可以直接请求第二终端执行第一终端的计算任务，如图 10 所示的流程。

请参见图 10，为本申请实施例提供的第四种网络边缘计算方法的流程图。该网络边缘计算方法的具体流程描述如下：

S101、第二终端广播自身的计算能力。

在本申请实施例中，第二终端可广播用于指示自身计算能力的计算能力信息，这样第一终端接收该计算能力信息之后，可向第二终端请求执行第一终端的计算任务。

应理解，在实际网络中，多个第二类终端中的各个终端分别广播自身的计算能力。在这种情况下，第一终端可以分别接收多个第二类终端的计算能力信息，并可根据计算任务的需求或其他准则，确定向该多个第二类终端中的其中一个请求执行第一终端的计算任务。第二终端可基于预先配置的资源广播计算能力信息；或者，第二终端可基于网络设备配置的资源广播计算能力信息。其中，用于广播计算能力信息的资源可以是 UE 配置的专用资源（UE Specific），也可以不区分 UE，即 Non UE Specific 广播资源。用于广播计算能力信息的资源可以是业务级别的，例如针对 MEC 业务配置的广播资源。

第二终端可周期性地广播计算能力信息。或者第二终端在触发事件的触发下，广播计算能力信息。例如第二终端在计算能力发生变化后，广播计算能力信息。例如第二终端的计算负载低于预设门限，第二类终端广播计算能力信息。

第一终端可周期性地监听第二终端广播的计算能力信息。或者第一终端被触发监听第二终端广播的计算能力信息。例如第一终端需要其它终端执行第一终端的计算任务，第一终端监听第二终端广播的计算能力信息。第一终端根据监听到的计算能力信息，确定能够提供计算服务的第二终端，从而请求第二终端执行第一类终端的至少一个计算任务。

S102、第一终端向第二终端发送第一任务请求消息，第二终端接收该第一任务请求消息。

S103、第二终端根据第一任务请求对第一终端的至少一个计算任务进行准入控制。

S104、第二终端向第一终端发送第一任务请求响应消息，第一终端接收该第一任务请求响应消息。

S105、第一终端与第二终端建立 D2D 连接。

S106、第一终端与第二终端交互计算任务数据。

第一类终端可通过网络设备调度的侧行链路（sidelink，SL）资源与第二类终端进行计

算任务数据的交互。或者第一类终端也可以确定网络设备配置或者预配置的 Sidelink 资源与第二类终端进行计算任务数据的交互。

应理解, S102、S103 以及 S104 可参考前述图 8 或图 9 中第一任务请求消息, 第一任务请求响应消息等的相关介绍, 这里不再赘述。不同之处在于, 图 10 中第一终端和第二终端之间的信令交互基于 D2D 方式。

通过本实施例上述步骤, 第一终端和第二终端之间的交互不需要借助例如网络设备转发信令, 这样可减少第二终端执行第一终端的计算任务的信令开销。

作为图 10 的一种可替换的实现方式, 请参见图 11, 为本申请实施例提供的第五种网络边缘计算方法的流程图。该网络边缘计算方法的具体流程描述如下:

S111、第一终端广播计算服务请求消息。

其中, 第一终端是第一类终端中的任意一个终端。计算服务请求消息指的是执行第一终端的至少一个计算任务所需要的计算服务。第一终端需要第二终端执行第一终端的计算任务时, 可广播计算服务请求消息。例如第一终端可基于预先配置的资源广播计算服务请求消息; 或者第一终端可基于网络设备配置的资源广播计算服务请求消息。其中, 用于广播计算服务请求消息的资源可以是针对配置的专用资源(UE Specific), 也可以不区分 UE, 即 Non UE Sepcific 广播资源。或者用于广播计算服务请求消息的资源可以是业务级别的, 例如针对 MEC 业务配置的广播资源。需要说明的是, 本申请实施例对计算服务请求消息的名称不作限制。

第二终端可周期性地监听第一终端广播的计算服务请求消息。或者第二终端的计算负载低于某一门限, 第二终端监听第一终端广播的计算服务请求消息。当第二终端确定满足第一终端的计算服务需求, 第二终端可响应计算服务请求消息。

S112、第二终端向第一终端发送计算服务响应消息。

其中, 第二终端是第二类终端中的任意一个终端。第二终端确定满足第一终端的计算服务需求, 第二终端可向第一终端发送计算服务响应消息, 该计算服务响应消息可用于指示第二终端能够满足第一类终端的计算服务需求。当然如果第二终端确定不能够满足第一终端的计算服务需求, 那么第二终端可不响应第一终端的计算服务请求消息, 即不向第一终端发送计算服务响应消息。需要说明的是, 本申请实施例对计算服务响应消息的名称不作限制。

S113、第一终端向第二终端发送第一任务请求消息, 第二终端接收该第一任务请求消息。

S114、第二终端根据第一任务请求消息对第一终端的至少一个计算任务进行准入控制。

S115、第二终端向第一终端发送第一任务请求响应消息, 第一终端接收该第一任务请求响应消息。

S116、第一终端与第二终端建立 D2D 连接。

S117、第一终端与第二终端交互计算任务数据。

应理解, S113-S117 的具体实现可参考前述图 10 所示的实施例的相关介绍, 这里不再赘述。

通过本实施例上述步骤, 第一终端有计算任务需求时, 可广播计算服务请求消息, 以获取多个第二终端的最新的计算能力, 从而可选择合理的第二终端执行计算任务, 以节省不必要的第二终端的广播, 减少资源开销。

应理解，第二类终端能够提供计算服务，那么可用于分担网络侧的计算负担。即第二类终端可执行网络侧的计算任务。考虑到通常来说，终端侧的计算能力低于网络侧的计算能力，一个第二类终端的计算能力可能不满足网络侧的计算任务需求，因此，网络侧可将计算任务划分为多个计算子任务，这多个计算子任务彼此之间是独立的，即一个计算子任务的执行，对另一个计算子任务的执行没有影响。网络侧可请求终端侧的多个第二类终端执行这多个计算子任务，这样每个第二类终端可根据自身的计算能力执行合理数量的计算子任务，从而保证业务的及时响应，提高网络的运行效率。

下面结合附图介绍终端侧执行网络侧的计算任务流程。

请参见图 12，为本申请实施例提供的第六种网络边缘计算方法的流程示意图。该方法的具体流程描述如下：

S121、第二终端和第一 MEC 网元交互所述第二终端的计算能力信息。

具体地，第二类终端中的多个终端分别和第一 MEC 网元交互各自的计算能力信息。该 S121 的具体实现同前述的 S91，可参考前述 S91 的相关描述，这里不再赘述。

S122、第一 MEC 网元将需要执行的计算任务划分为多个计算子任务，并将所述多个计算任务分配给多个第二类终端。

应理解，第一 MEC 网元可存储多个第二类终端的计算能力。不同的第二类终端的计算能力有所不同，能够提供的计算服务也有所不同。某个第二类终端能够提供的计算服务可能不满足第一 MEC 网元的计算任务需求，那么由该第二类终端执行第一 MEC 网元的计算任务，显然会导致失败。

为此，本申请实施例中，第一 MEC 网元可将计算任务划分为多个计算子任务，再将这多个计算子任务分配给多个第二类终端执行，这样可保证第一 MEC 网络的计算任务的顺利执行。具体的第一 MEC 网元可根据各个第二类终端的计算负载、通信速率、可用计算资源等信息确定适合执行计算子任务的多个第二类终端，并为每个第二类终端分配合理数量的计算子任务。示例性的，可将涉及隐私数据的计算子任务分配给第二类终端，以尽量保证数据的安全性。

具体地，第一 MEC 网元从获知计算能力的多个第二类终端中确定由哪些第二类终端共同执行计算任务，并可进一步确定由哪个第二类终端执行哪个计算子任务。换言之，在该步骤中，第一 MEC 网元确定哪个目标第二类终端执行哪个目标计算子任务。其中，有多个目标第二类终端和多个目标计算子任务，并且目标第二类终端的数目和目标计算子任务的数目是相同的。

需要说明的是，本申请实施例不限制第一 MEC 网元将计算任务划分为多个计算子任务的方式。例如第一 MEC 网元可通过拆分人工智能/机器学习（Split AI/ML）模型，例如联邦学习（federated learning），将计算任务拆分为多个计算子任务。

S123、第一 MEC 网元向第二终端发送第一任务请求消息，第二终端接收该第一任务请求消息。

应理解，S123 中的第二终端为多个目标第二类终端中的任意一个目标第二类终端。

具体地，第一 MEC 网元分别向在上述步骤 S122 中确定的各个目标第二类终端发送第一任务请求消息，其中第一任务请求消息中包含第一 MEC 网元请求该目标第二类终端执行的目标计算子任务的服务需求。

S124、第二终端分别根据第一任务请求消息对要执行的计算任务进行准入控制。

具体地, 各个目标第二类终端对接收到的目标计算子任务进行准入控制, 以决定是否执行相应的目标计算子任务。

S125、第二终端向第一 MEC 网元发送第一任务请求响应消息。

5 具体地, 各个目标第二类终端进行准入控制后分别向第一 MEC 网元发送第一任务请求响应消息, 以指示该目标第二类终端是否执行相应的目标计算子任务。

S123-S125 的具体实现可参考前述图 7 中 S731b-S732b 的具体实现, 这里不再赘述。

S126、第一 MEC 网元向第一网络设备发送会话建立请求消息, 该会话请求消息用于指示第一网络设备建立第一网络设备到各个目标第二类终端的用户面传输通道。

10 该会话建立请求消息可包括针对分配给第二终端的计算任务的 QoS 流配置信息, 以保证计算业务中不同数据流的 QoS 需求。该 QoS 流配置信息可包括服务质量流标识(QoS flow identity, QFI) 和 QoS 模板(profile), 该配置信息可以由网络侧设备(例如 SMF)生成。

S127、第一网络设备与第二终端交互 RRC 重配置消息。

15 第一网络设备可分别对各个目标第二终端进行 RRC 重配置, 以配置计算任务的无线承载, 以及无线承载与计算任务的映射关系。其中, 计算任务的无线承载可以是 DRB, 也可以是 RB。

S128、第一网络设备向第一 MEC 网元发送会话建立请求响应消息。

第一网络设备建立第一网络设备到各个目标第二终端的用户面传输通道之后, 可向第一 MEC 网元发送会话建立请求响应消息, 以告知第一 MEC 网元, 第一网络设备到各个目标第二终端的用户面传输通道已建立。

20 S129、第一 MEC 网元与第二终端交互计算任务数据。

具体地, 第一 MEC 网元分别与各个目标第二终端交互计算子任务数据。针对第一 MEC 网元与一个目标第二终端交互对应的计算子任务数据的操作, 具体实现可参考前述 S704 的具体实现, 这里不再赘述。

25 通过本申请实施例的步骤, 可以使网络将部分的计算任务卸载到第二类终端执行, 节省了网络侧资源, 也充分利用了终端侧的资源。

应理解, 根据第二终端处于连接态(RRC_connected)、非激活态(RRC_inactive)或空闲态(RRC_idle)的不同, 上述第六种网络边缘计算方法的流程也有差异。图 12 所示的流程可认为是将网络侧的计算任务卸载到处于连接态的第二终端的流程。

30 如果第二终端处于非激活态, 那么将网络侧的计算任务卸载到第二终端的流程如图 13 所示。应理解, 如果第二终端处于非激活态, 第二终端与第一 MEC 网元交互已经交互了第二终端的计算能力信息。即第一 MEC 网元已经知道了第二终端的计算能力。

S131、第一 MEC 网元将需要执行的计算任务划分为多个计算子任务, 并将所述多个计算任务分配给多个第二类终端。

35 S132、第一 MEC 网元向第一网络设备发送第一任务请求消息, 该第一任务请求消息携带多个第二类终端的相关信息。

应理解, 这里的多个第二类终端中的任意一个是第一 MEC 网元确定用于执行计算子任务的第二终端。

S133、第一网络设备向第二终端发起随机接入寻呼消息, 该寻呼消息包括第一任务请求消息。

40 具体地, 第一网络设备分别向上述多个第二类终端发随机接入寻呼消息, 各个寻呼消

息包含对各个第二类终端的计算子任务的请求。

S134、第二终端根据第一任务请求消息对计算子任务进行准入控制。

S135、第二终端向第一网络设备发送第一任务请求响应消息。

5 该第一任务请求响应消息可以是 RRC 恢复请求消息。可认为通过 RRC 恢复请求消息隐式指出接受网络侧的任务请求，否则不会发起 RRC 恢复请求消息。

S136、第一网络设备向第一 MEC 网元转发第一任务请求响应消息。

S137、第一 MEC 网元向第一网络设备发送会话建立请求消息，该会话请求消息用于指示第一网络设备建立第一网络设备到第二终端的用户面传输通道。

S138、第一网络设备与第二终端交互 RRC 重配置消息。

10 第一网络设备可对第二终端进行 RRC 重配置，以配置计算任务的无线承载，以及无线承载与计算任务的映射关系。其中，计算任务的无线承载可以是 DRB，也可以是 RB。

S139、第一网络设备向第一 MEC 网元发送会话建立请求响应消息。

S1391、第一 MEC 网元与第二终端交互计算任务数据。

15 应理解，图 13 与图 12 的不同之处在于，由于第二终端处于非激活态，所以 S132-S136 有所不同。

通过本申请实施例的步骤，可以使网络将部分的计算任务卸载到处于非激活态的第二类终端执行，即网络侧除了将计算任务卸载到处于激活态的第二类终端执行，也可以将计算任务卸载到处于非激活态的第二类终端执行，充分利用了终端侧的资源。

如图 14 所示，示出将网络侧的计算任务卸载到处于空闲态的第二终端的流程。

20 S140、第一 MEC 网元将需要执行的计算任务划分为多个计算子任务。

由于第二终端处于空闲态，所以第一 MEC 网元与第二终端没有交互第二终端的计算能力信息。这种情况下，第一 MEC 网元可事先将计算任务划分为多个计算子任务。之后再请求一个或多个第二终端执行这多个计算子任务，以尽量保证每个第二终端执行合理数量的计算子任务，能够保证整个计算任务的执行成功率。

25 S141、第一 MEC 网元向第一网络设备发送第一任务请求消息，第一网络设备接收该第一任务请求消息。

S142、第一网络设备广播第一任务请求消息。

30 由于第二终端处于空闲态，所以第一网络设备可广播第一任务请求消息，这样第二终端接收到第一任务请求消息，可根据自身的计算负载、计算资源等确定是否满足一个或多个计算子任务。若某个第二终端确定能够满足一个或多个计算子任务的计算服务需求，那么该第二终端可根据第一任务请求消息对一个或多个计算子任务进行准入控制。

S143、第二终端根据第一任务请求消息对一个或多个计算子任务进行准入控制。

S144、第二终端向第一网络设备发送第一任务请求响应消息，第一网络设备接收该第一任务请求消息。

35 与 S135 的不同之处在于，第一任务请求响应消息可以是 RRC 连接建立请求消息。

S145、第一网络设备向第一 MEC 网元发送第一任务请求响应消息，第一 MEC 网元接收该第一任务请求消息。

S146、第一 MEC 网元向第一网络设备发送会话建立请求消息，该会话请求消息用于指示第一网络设备建立第一网络设备到第二终端的用户面传输通道。

40 S146 的具体实现同前述的 S126 的具体实现相同，可参考前述 S126 的相关介绍，这里

不再赘述。

S147、第一网络设备与第二终端交互 RRC 重配置消息。

S147 的具体实现同前述的 S127 的具体实现相同，可参考前述 S127 的相关介绍，这里不再赘述。

5 S148、第一网络设备向第一 MEC 网元发送会话建立请求响应消息。

S148 的具体实现同前述的 S128 的具体实现相同，可参考前述 S128 的相关介绍，这里不再赘述。

S149、第一 MEC 网元与第二终端交互计算任务数据。

该 S149 的具体实现可参考前述 S704 的具体实现，这里不再赘述。

10 第一 MEC 网元与处于空闲态的第二终端没有交互第二终端的计算能力信息，而通过本实施例上述步骤，即第一 MEC 网元可事先将计算任务划分为多个计算子任务，之后再请求一个或多个第二终端执行这多个计算子任务，可尽量保证每个第二终端执行合理数量的计算子任务，能够保证整个计算任务的执行成功率。

15 前述图 12-图 14 所示的实施例中，多个计算子任务彼此之间是独立的，即一个计算子任务的执行，对另一个计算子任务的执行没有影响。换言之，各个计算子任务之间可以认为是并行执行的。然而实际中，多个计算子任务之间也可存在输入输出逻辑关系，例如一个计算子任务的执行需要参考另一个计算子任务的执行结果。这种情况下，可以认为各个计算子任务之间是串行执行的。存在输入输出逻辑关系的计算任务可以理解为是服务功能链（service function chain, SFC）的计算任务。

20 SFC 通常与软件定义网络（software-defined networking, SDN）相结合，SDN 主要提供的服务可包括两类。例如第一类即提供网络基础架构（networking infrastructure 功能，也可理解为提供 SDN 中各个终端设备（例如，虚拟机）的网络连通性功能。又例如第二类即提供网络服务（networking service）功能，例如向 SDN 用户提供网络服务（Networking As A Service）、负载均衡功能、防火墙功能、QoS 功能等等。其中，SFC 是实现 Networking Service 的一种很灵活的方式。云服务运营商提供各种各样的服务功能，由 SDN 管理 SFC，将服务功能加入到网络流量路径中，即能实现网络服务。

25 本申请实施例还提供了一种基于 SFC 的 MEC 机制，可通过缩小任务调度粒度来提高资源调度的精度，提高系统的计算资源利用效率。

30 请参见图 15，为本申请实施例提供的基于 SFC 的 MEC 方法，该方法可应用于如图 3 所示的 MEC 网络架构，即网络侧和终端侧均设置有 MEC 功能为例。图 15 以将第一终端的计算任务通过第二类终端以及多个 MEC 网元（例如第一 MEC 网元和第二 MEC 网元）为例。该方法的具体流程描述如下：

S150、第一 MEC 网元和第二 MEC 网元交互彼此支持的服务功能的类型，以及计算负载。

35 该 S150 的具体实现可参考前述 S82b 的具体介绍，与 S82b 的不同之处在于，S150 中第一 MEC 网元和第二 MEC 网元交互的内容与 S82b 中的不同。在本申请实施例中，第一 MEC 网元可告知第二 MEC 网元，第一 MEC 网元支持的服务功能的类型（也称为服务功能类型），以及该服务功能类型对应的当前计算负载。同理。第二 MEC 网元可告知第一 MEC 网元，第二 MEC 网元支持的服务功能类型，以及该服务功能类型对应的当前计算负载。这样第一 MEC 网元和第二 MEC 网元可根据各自支持的服务功能类型合理分配多个计

40

算子任务。

S151、第二类终端和第一 MEC 网元交互所述第二类终端的计算能力信息。

具体地，第二类终端中的多个终端分别和第一 MEC 网元交互各自的计算能力信息。该 S151 的具体实现可参考前述 S91 的具体介绍，这里不再赘述。需要说明的是，本申请
5 实施例对 S150 和 S151 的执行顺序不作限制。即 S150 可在 S151 之后执行，或者 S151 在 S150 之前执行。

之后，第一终端可请求网络侧执行第一终端的计算任务。在一种可能的实现方式中，第一终端可将计算任务划分为基于 SFC 的多个计算子任务，再请求网络侧执行这多个计算子任务，即执行 S152a-S153a。在另一种可能的实现方式中，第一终端可请求网络侧执行
10 第一终端的计算任务，网络侧将第一终端的计算任务划分为基于 SFC 的多个计算子任务，再执行这多个计算子任务，即执行 S152b-S153b。

S152a、第一终端将计算任务划分为基于 SFC 的多个计算子任务。

第一终端可将需要执行的计算任务划分为基于 SFC 的多个计算子任务。示例性的，第一终端可将计算任务建模为 SFC 模型，从而获得多个计算子任务，其中 SFC 模型的建立
15 方法本申请实施例不作限制。例如第一类终端将计算任务划分为基于 SFC 的 6 个计算子任务这 6 个计算子任务组成线性拓扑结构。假设这 6 个计算子任务的索引为{1,2,3,4,5,6}，组成线性拓扑结构。

S153a、第一终端向第一 MEC 网元发送第二任务请求消息，第一 MEC 网元接收该第二任务请求消息。

第一终端获得基于 SFC 的多个计算子任务之后，可向第一 MEC 网元发送任务请求消息，例如第二任务请求消息，以请求第一 MEC 网元执行这多个计算子任务。其中，第二任务请求消息的实现方式同上述 S731b 中的第一任务请求消息的实现方式。与上述 S731b 中的第一任务请求消息的不同之处在于，第二任务请求消息携带的内容包括的是基于 SFC 的多个计算子任务，以及这多个计算子任务之间的逻辑拓扑关系，以及各个计算子任务的服务需求。以这多个计算子任务是 6 个计算子任务为例，这 6 个计算子任务的索引为{1,2,3,4,5,6}，那么这 6 个计算子任务的逻辑拓扑关系例如可以为 1->2->...->6，即第二个
25 计算子任务是基于第一个计算子任务的执行结果才能执行，第三个计算子任务是基于第二个计算子任务的执行结果才能执行，以此类推。

S152b、第一终端向第一 MEC 网元发送第一任务请求消息，第一 MEC 网元接收该第一任务请求消息。

S152b 的具体实现可参考前述 S731b 的相关介绍，这里不再赘述。

S153b、第一 MEC 网元将计算任务划分为基于 SFC 的多个计算子任务。

第一 MEC 网元接收到第一任务请求消息，可获取第一终端请求执行的计算任务。第一 MEC 网元可根据计算任务的计算需求，以及结合网络中提供计算服务的网元（例如第二 MEC 网元和第二类终端）将计算任务划分为基于 SFC 的多个计算子任务。
35

示例性的，第一 MEC 网元可根据各个提供计算服务的网元的计算能力、各个网元之间的链路状态、时延需求等参数对第一类终端请求执行的计算任务进行划分，获得多个计算子任务。例如第一 MEC 网元将该计算任务划分为 6 个计算子任务，这 6 个计算子任务的索引为{1,2,3,4,5,6}。

S154、第一 MEC 网元确定通过第二 MEC 网元执行多个计算子任务中的第一部分计算
40

子任务, 通过第二类终端执行多个计算子任务中的第二部分计算子任务。需要说明的是, 第一 MEC 网元可以请求一个或多个第二 MEC 网元, 和/或, 一个或多个第二终端执行多个计算子任务。例如第一 MEC 网元可以请求一个或多个第二 MEC 网元执行多个计算子任务中的第一部分计算子任务, 请求一个或多个第二终端执行多个计算子任务中的第二部分计算子任务。又或者, 第一 MEC 网元可以请求一个或多个第二 MEC 网元以及一个或多个第二终端执行多个计算子任务中的第一部分计算子任务, 请求其他一个或多个第二终端执行多个计算子任务中的第二部分计算子任务。

进一步的, 第一 MEC 网元还可根据各个网元的计算能力、各个网元之间的链路状态、时延需求等参数从各个网元中选择合适的网元执行部分计算子任务。例如第一 MEC 网元将该计算任务划分为三部分, 这三部分包括的计算子任务的索引分别为{1,2}{3,4}{5,6}, 这三部分计算子任务分别由第一 MEC 网元、第二 MEC 网元和第二类终端执行。

S155a、第一 MEC 网元向第二 MEC 网元发送第三任务请求消息, 第二 MEC 网元接收该第三任务请求消息。

第一 MEC 网元确定需要某个第二 MEC 网元执行的第一部分计算子任务(例如计算子任务 3 和计算子任务 4)之后, 可向该第二 MEC 网元发送第三任务请求消息。也就是该第二 MEC 网元是 S154a 中第一 MEC 网元确定的一个或多个第二 MEC 网元中的一个。其中, 第三任务请求消息的具体实现可参考前述 S832b 的相关介绍, 这里不再赘述。应理解, 该第三任务请求消息除了携带第一部分计算子任务之外, 还携带第一部分计算子任务之间的逻辑关系, 即 3->4。该第三任务请求消息还可以携带第一部分计算子任务的计算需求。

S156a、第二 MEC 网元根据第三任务请求消息对第一部分计算子任务进行准入控制。

第二 MEC 网元接收到第三任务请求消息可根据第二 MEC 网元的当前计算任务的负载、网络状态等信息对第一部分计算子任务进行准入控制, 具体可参考前述 S834a 的相关介绍, 这里不再赘述。

S157a、第二 MEC 网元向第一 MEC 网元发送第一准入信息, 第一 MEC 网元接收该第一准入信息。

第二 MEC 网元对第一部分计算子任务进行准入控制之后, 可向第一 MEC 网元反馈第一准入信息。例如第二 MEC 网元可向第一 MEC 网元发送第三任务请求响应消息, 即针对第三任务请求消息的响应, 该第三任务请求响应消息可携带第一准入信息。具体可参考前述 S834b 的相关介绍, 这里不再赘述。

S155b、第一 MEC 网元向第二终端发送第四任务请求消息, 第二终端接收该第四任务请求消息。其中, 第二终端是第一 MEC 网元在上述步骤 S154 中确定的一个或多个第二类终端中的任意一个终端。

第一 MEC 网元确定需要第二终端执行的第二部分计算子任务(例如计算子任务 5 和计算子任务 6)之后, 可向第二终端发送第四任务请求消息。其中, 第四任务请求消息的具体实现可参考前述 S123 的相关介绍, 这里不再赘述。应理解, 该第四任务请求消息除了携带第二部分计算子任务之外, 还携带第二部分计算子任务之间的逻辑关系, 即 5->6。该第四任务请求消息还可以携带第二部分计算子任务的计算需求。

S156b、第二终端根据第四任务请求消息对第二部分计算子任务进行准入控制。

第二终端接收到第四任务请求消息可根据第二终端的当前计算任务的负载、网络状态等信息对第二部分计算子任务进行准入控制, 具体可参考前述 S124 的相关介绍, 这里不

再赘述。

S157b、第二终端向第一 MEC 网元发送第二准入信息，第一 MEC 网元接收该第二准入信息。

第二终端对第二部分计算子任务进行准入控制之后，可向第一 MEC 网元反馈第二准入信息。例如第二终端可向第一 MEC 网元发送第四任务请求响应消息，即针对第四任务请求消息的响应，该第四任务请求响应消息可携带第二准入信息。具体可参考前述 S125 的相关介绍，这里不再赘述。

需要说明的是，本申请实施例对 S155a 和 S155b 的执行顺序不作限制，以及对 S156a 和 S156b 的执行顺序不作限制、对 S157a 和 S157b 的执行顺序不作限制。

S158、第一 MEC 网元向第一终端发送第一任务请求响应消息，第一终端接收该第一任务请求响应消息。

第一 MEC 网元接收到第三任务请求响应消息和第四任务请求响应消息，可向第一终端发送第一任务请求响应消息，以告知第一终端，第一终端请求执行的计算任务的准入信息。具体可参考 S733b 的相关介绍，这里不再赘述。

S159、第一终端和第一 MEC 网元、第二 MEC 网元以及第二终端建立通道连接。

应理解，需要建立第一终端的计算任务承载节点之间的用户面隧道以通过该隧道传输计算任务数据。在本申请实施例中，第一终端请求执行的计算任务被划分多个计算子任务，这多个计算子任务由第一 MEC 网元、第二 MEC 网元以及第二终端执行。即计算任务的传输路径为第一终端->第一 MEC 网元->第二 MEC 网元->第二终端->第一终端。这就需要建立第一终端和第一 MEC 网元之间的用户面隧道、第一 MEC 网元和第二 MEC 网元之间的用户面隧道、第二 MEC 网元和第二终端之间的用户面隧道，以及第二终端和第一终端之间的用户面隧道。

例如，第一终端和第一 MEC 网元建立用户面隧道的过程可参考前述 S126-S128 的相关介绍，这里不再赘述。

例如，第一 MEC 网元和第二 MEC 网元建立用户面隧道的过程可通过第一 MEC 网元和第二 MEC 网元之间的 X 接口交互建立用户面隧道的相关信息。例如第一 MEC 网元通过 X 接口向第二 MEC 网元发送会话建立请求消息，以请求建立第一 MEC 网元和第二 MEC 网元之间的用户面隧道。

例如，第二 MEC 网元和第二终端建立用户面隧道的过程可参考前述 S126-S128 的相关介绍，这里不再赘述。

例如，第二终端和第一终端建立用户面隧道的过程可参考前述 S921 或 S931 的相关介绍，这里不再赘述。

第一终端和第一 MEC 网元、第二 MEC 网元以及第二终端之间的通道连接后，各个节点交互计算任务数据。其中，在用户面隧道建立的过程中，第一 MEC 网元可以指示第二 MEC 网元在处理完成后，将数据发送给第二终端进行处理。或者，第二 MEC 网元处理完成后返回给第一 MEC 网元，再由第二 MEC 转发给第二终端进行处理。类似的，第二终端可以直接将最终结果发送给第一终端，或者通过第一 MEC 网元转发给第一终端。

通过本申请实施例提供基于 SFC 的 MEC 机制，可通过缩小任务调度粒度来提高资源调度的精度，从而提高了系统的计算资源利用效率。

针对某些业务，例如 AR 业务或 VR 业务的处理会涉及到渲染的处理。如果本地渲染

需要通过 MEC 功能进行处理,那么需要在网络边缘的网络侧配置包过滤功能,以保证 MEC 功能处理后获得的数据包能够映射到正确的 QoS 流。应理解,针对网络侧向终端侧发送的数据包(也称为下行数据包),UPF 基于该数据包检测规则(packet detection rule, PDR)中的分组数据包过滤器(packet filter)对数据包进行分类,并将其映射到对应的 QoS 流;针对终端侧向网络侧发送的上行数据包,终端侧基于 QoS 规则中的 packet filter 对数据包进行分类。packet filter 主要包括 IP packet filter 和 Ethernet packet filter,具体可参见 3GPP 协议规范 TS23.501 中的相关内容。

作为一种示例,基于图 1 所示的 MEC 架构,终端可通过本地(local)的 UPF(L_UPF)与 EAS 建立的 local PDU(L_PDU)会话,云服务器(cloud server, CS)将数据经由公网发送给 EAS 进行中间处理(本地渲染),EAS 处理之后,将数据由 L_PDU 会话发送给终端。

具体的,请参见图 16,示出了应用于现有 MEC 架构的中间渲染信令交互流程。该流程包括如下步骤:

S161、终端获取 EAS 配置信息。

EAS 配置信息可包括发送端的网际协议(internet protocol, IP)或者端口(port)信息(简称为 IP 地址/port 信息)、发送端的 IP 地址/port 信息,或者也可以包括发送端的 MAC 地址信息、接收端的 MAC 地址信息等。终端可通过计算服务发现过程,例如边缘启动服务器(Edge Enabler Server, EES)/边缘配置服务器(Edge Configuration Server, ECS)获取 EAS 的配置信息。

S162、终端通过本地 UPF 建立与 EAS 的本地 PDU 会话,并获取本地 PDU 会话的 IP 信息。

该 IP 信息可以是内网 IP 地址/port 信息,也可以是公网 IP 地址/port 信息。终端和 EAS 交互的信息依据会话管理功能(session management function, SMF)为 L_UPF 配置的 PDRs。需要说明的是,这里 PDRs 表示一个 PDR 或多个 PDR。例如针对 CS 的 PDRs, packet filter 中的发送端的 IP 地址/port 值为 CS 的 IP 地址/port;针对 EAS 的 PDRs, packet filter 中的发送端的 IP 地址/port 值为 EAS 的 IP 地址/port。应理解, EAS 位于 UPF 之后,收到的数据包源地址为 UE 的公网 IP 地址/port, EAS 与 UPF 基于终端的公网 IP 地址/port 进行交互。例如终端获取 PDU 会话对应的公网 IP 地址/port,并通知 EAS,这样 EAS 可将处理后的数据包发送给正确的目的地址(UPF)。如果终端只告诉 EAS, PDU 会话对应的内网 IP,那么 EAS 在公网内可能找不到目的地址。

需要说明的是,3GPP 网络内部(内网)的终端的 PDU 会话只有内网 IP 地址,没有端口号(port),UPF 负责将 PDU 会话的内网 IP 映射到公网对应的 IP 地址/port。下行数据到达 UPF 后,UPF 根据数据包的目的 IP 地址/port(即公网目的 IP 地址/port)来区分目标 UE 的 PDU 会话。

终端可通过本地 UPF 建立与 EAS 的本地 PDU 会话,从而获取本地 PDU 会话的 IP 信息。或者,终端也可通过单个 PDU 会话与 EAS 和 CS 建立连接,获取本地 PDU 会话的 IP 信息。或者终端分别与 EAS 和 CS 建立本地 PDU 和默认(default)PDU 会话,之后获取本地 PDU 会话的 IP 信息。

S163、终端向 EAS 发送第一对应关系。

该第一对应关系用于指示 L_PDU 会话与 CS 的对应关系。例如可以是 L_PDU 内网 IP 与 CS IP 地址/port 的对应关系,也可以是 L_PDU 公网 IP 地址/port 与 CS IP 地址/port 的对

应关系。EAS 可根据第一对应关系对来自 CS 的数据进行中间处理，例如渲染处理。

S164、EAS 与 CS 交互彼此的 IP 地址/port 信息。

EAS 可基于 TCP/IP 或者 Ethernet 协议，根据第一对应关系在公网与 CS 通信，以获取 EAS 的 IP 地址/port。之后 EAS 可指示 CS 将来自终端的下行数据发送给 EAS。

5 S165、EAS 对接收的下行数据进行处理。

EAS 接收到该下行数据之后，对该下行数据进行处理，并生成新的传输控制协议（transmission control protocol, TCP）TCP/IP 数据包。应理解，TCP 数据包的结构可参考 TCP 协议的定义，IP 数据包的结构可参考 IP 协议的定义。IP 数据包通过 IP 地址信息把数据包发送给指定的终端，而 TCP 数据包通过端口号把数据包分发给正确的应用。

10 EAS 可以修改数据包的发送端的 IP 地址/port 为 CS 的 IP 地址/port，而接收端的 IP 地址/port 为 L_PDU 会话的 IP 地址/port。L_UPF 基于 S162 中 CS 的 PDRs 对发送端发送的数据包过滤。或者，EAS 对数据包不进行处理，即发送端的 IP 地址/port 仍为 EAS 的 IP 地址/port，接收端的 IP 地址/port 为 L_PDU 会话的 IP 地址/port，L_UPF 基于 S162 中 EAS 的 PDRs 对发送端发送的数据包进行过滤。

15 S166、L_UPF 通过 PDRs 将数据包映射至 QoS 流发送给网络设备，网络设备将其映射到对应的 DRB 发送给终端。

本申请实施例中，由于终端向 EAS 发送第一对应关系，所以在多级处理的 MEC 应用场景中，可保证数据包在传输和处理过程中 QoS 映射关系不变，从而保证业务的 QoS 需求。

20 下面介绍本申请实施例基于图 3 所示的 MEC 网络架构，可能的中间渲染信令交互流程。与图 16 的不同之处在于，终端需要通过 MEC 管理功能进行网络侧的计算功能，例如云服务器（cloud service, CS）的发现和计算任务分配过程。之后，终端与计算功能建立 PDU 会话连接，下行数据到达网络设备后，网络设备将下行数据转发至对应的计算服务器进行处理。计算服务器处理下行数据，将处理后的数据反馈给网络设备，网络设备将处理后的数据映射到对应的 DRB 发送给终端。应注意，MEC 管理功能并不参与数据传输和处理过程。

具体请参见图 17，该流程包括如下步骤：

S171、终端向 MEC 网元发送第一任务请求消息，MEC 网元接收该第一任务请求消息。

30 第一任务请求消息可用于请求 MEC 网元执行终端的至少一个计算任务。该第一任务请求消息可包括计算功能的相关信息，例如 CS 的 IP 地址/port 信息。

MEC 网元接收到第一任务请求消息，根据该 MEC 网元上注册的计算服务器的计算能力信息，选择合适的第一计算服务器。之后，MEC 网元将第一任务请求消息发送给第一计算服务器。

35 计算服务器根据第一任务请求消息对终端请求执行的至少一个计算任务进行准入控制。之后计算服务器可将至少一个计算任务的准入信息发送给 MEC 网元。例如计算服务器向 MEC 网元发送第一任务请求响应消息，该第一任务请求响应消息携带至少一个计算任务的准入信息。当然如果计算服务器拒绝执行终端的至少一个计算任务，第一任务请求响应消息可用于指示计算服务器拒绝执行终端的至少一个计算任务。

S172、终端接收 MEC 网元发送的第一任务请求响应消息。

40 MEC 接收到来自计算服务器的第一任务请求响应消息，可将该第一任务请求响应消息

转发给终端。MEC 网元发送的该第一任务请求响应消息可包括至少一个计算任务的准入信息，以及计算服务器的身份标识信息、计算服务器的 IP 地址/port 信息。

S173、终端与计算功能建立 PDU 会话。

终端可将要处理的数据携带在 PDU 会话中发送给网络设备。因此终端在发送数据之前，可与计算功能，例如 CS 建立 PDU 会话。示例性的，终端可接入和移动管理功能(access and mobility management function, AMF)发送面向 CS 的 PDU 会话建立请求消息，请求与 CS 建立 PDU 会话。该 PDU 会话建立请求消息携带计算服务器的信息，例如计算服务器的身份标识信息、计算服务器的 IP 地址/port 信息。

AMF 将 PDU 会话与计算服务器的映射关系告知网络设备，指示网络设备将来自该 PDU 会话数据发送给计算服务器进行处理。其中，PDU 会话与计算服务器的映射关系可以是 PDU 的 IP 与计算服务器的 IP 地址/port 的映射关系，也可以是用于标识隧道的通用分组无线服务(general packet radio service, GPRS)隧道协议(GTP)隧道端点标识(tunnel endpoint identifier, TEID)与计算服务器的 IP 地址/port 的映射关系。

UE 与 CS 建立 PDU 会话过程中, SMF 可为网络设备配置基于计算服务器或 CS 的 PDRs, 如前述 S172 中的介绍。

S174、MEC 网元对来自终端的数据进行处理。

与图 16 所示实施例不同，通过本申请实施例上述步骤，AMF 将 PDU 会话与计算服务器的映射关系告知网络设备。这样基于图 3 所示的 MEC 网络架构中进行多级处理计算任务过程中，可保证数据包在传输和处理过程中 QoS 映射关系不变，从而保证业务的 QoS 需求。

终端与计算功能建立会话之后，可将要处理的数据携带在 PDU 会话中发送给网络设备，网络设备将来自该 PDU 会话数据发送给计算服务器进行处理。针对网络设备与计算服务器之间的隧道情况的不同情况，MEC 网元处理数据的流程也有不同。

第一种情况，网络设备与计算服务器之间建立 QoS 流粒度 GTP-U 隧道。

由于网络设备与计算服务器之间建立 QoS 流粒度 GTP-U 隧道，所以 PDU 会话建立过程中，SMF 会通过 AMF 告知网络设备，该 PDU 会话中包括哪些 QoS 流，每个 QoS 流对应的 QFI，以及对应 QoS profile。网络设备为该 PDU 会话中的每个 QoS 流分别建立从网络设备到计算服务器(或者 MEC 网元)的 GTP-U 隧道。当然，如果只有部分 QoS 流需要进行处理，会话建立请求可指定部分 QoS 流，网络设备为该 PDU 会话中的这部分 QoS 流分别建立从网络设备到计算服务器(或者 MEC 网元)的 GTP-U 隧道。不同 QoS 流对应的网络设备到计算服务器 GTP-U 隧道可通过网络设备 TEID 到计算服务器 RS TEID 进行区分。不同 QoS 流对应的网络设备到计算服务器 GTP-U 隧道指明不同 QoS 流对应的处理信息。

这种情况下，网络设备接收到来自终端的数据后，根据 GTP-U 报头中的 QFI，如果该 QoS 流的数据需要本地处理，则将数据映射到对应的 QoS 流隧道发送给计算服务器，否则将数据映射到对应的空口 DRB 上。

计算服务器从隧道接收数据，记录每个数据包对应的隧道 ID，该 ID 可以是 QFI，也可以是 TEID。计算服务器对数据包处理后，通过接收数据包的隧道，将处理后的数据包发送给网络设备。

请参见图 18，示出了计算服务器处理数据包的一种流程。计算服务器对数据包的处理过程包括但不限于以下 2 种情况。

过程一、计算服务器仅对每个 QoS 流隧道来的数据包单独处理，每个隧道输入的数据包个数与输出的数据包个数可以不同。

具体的，计算服务器从某个 QoS 流隧道传输的数据包，去除并保存数据包的原有报头，并对数据进行处理。给处理后的数据重新添加原有报头。以数据包是 IP 数据包为例，报头可认为包括数据包五元组，例如发送端的 IP 地址/port 信息、接收端的 IP 地址/port 信息、协议类型等参数。报头还可包括其他数据相关报头元素，例如数据总长，序号等。计算服务器处理数据后，为数据添加原有的报头，其中添加的报头的数据包五元组保持不变，其他数据相关报头元素可以根据实际的数据情况对应修改。类似的，如果数据包括是 ethernet 数据包，也是保持 ethernet packet filter 中的对应参数不变，数据相关参数可以根据数据实际处理情况对应修改。

过程二、计算服务器将不同 QoS 流隧道的数据报混合在一起进行处理。这就需要为计算服务器配置数据过滤功能(packet filter)，可以在 PDU 会话建立时，由 SMF 将相关的 packet filter 配置参数(通过 AMF 或网络设备)发送给 MEC 网元，再通过 MEC 网元配置给计算服务器。计算服务器对原始数据处理后，生成新的数据包，该新的数据包可以是 IP 数据包，也可以是 ethernet 数据包。计算服务器发送所生成的数据包，可通过 packet filter 对其进行过滤分类，将数据包映射到不同的 QoS 流，并根据对应的 QFI，将数据包通过对应的 QoS 流隧道发送给网络设备。

网络设备根据保存的 QoS 流隧道与 PDU 会话和 QoS flow 的映射关系，将计算服务器发送的数据映射到对应的空口 DRB 上。如果 PDU 会话中仅部分 QoS 流需要处理，为了保证处理和未处理数据同时到达目的端，处理后的数据包的包延迟预算(packet delay budget, PDB)可减去计算服务器的处理时延和计算服务器到网络设备间的往返数据传输时延。

考虑到处理过的 QoS 流和未处理的 QoS 流可能需要映射到相同的 DRB，这种情况下，则需要考虑根据数据包的剩余 PDB 在缓存上插队，例如优先发送剩余 PDB 小的数据包。或者可在 SDAP 层引入每个 QoS 流的缓存来缓存数据，并根据数据包的生存时间(survival time)向 PDCP 层优先注入 survival time 小的数据包。

需要说明的是，本申请实施例对数据包的报头进行处理，仅针对数据包的外层报头，(outer IP header)，即与传输相关的报头。并不涉及应用层数据，例如数据净荷的报头的处理。这里的外层报头和数据净荷的报头是相对 IP 里面封装 IP(IP Encapsulation within IP)，可简称为 IP-in-IP 来说的。IP-in-IP 描述了如何获取 IP 数据报并将其作为另一个 IP 数据报的载荷。具体可参见移动 IP 封装数据的过程。

第二种情况，网络设备与计算服务器之间建立 PDU 会话粒度 GTP-U 隧道。

由于网络设备与计算服务器之间建立 PDU 会话粒度 GTP-U 隧道，所以 PDU 会话建立过程中，核心网设备例如 SMF、AMF 或者 MEC 网元可指示网络设备该 PDU 会话需要进行中间处理，且该 PDU 会话携带对应的本地处理信息。网络设备建立从网络设备到计算服务器或者 MEC 网元的 PDU 会话粒度 GTP-U 隧道，并记录 N3 隧道(网络设备到 UPF)与上述隧道(网络设备到计算服务器)的对应关系。

网络设备接收到核心网设备发送的数据后根据上述隧道映射关系将数据通过对应隧道发送给计算服务器。计算服务器从隧道接收数据后，对数据进行处理，并将处理后的数据包发送给网络设备。网络设备根据数据包对应的 QFI，将计算服务器返回的数据映射到对应的空口 DRB 上。

请参见图 19, 示出了计算服务器处理数据包的另一种流程。计算服务器从隧道接收数据后, 可通过如下三种处理方式处理接收的数据。

处理方式一、计算服务器对具有相同 QFI 的数据包进行统一处理。

具体的, 计算服务器根据每个数据对应的隧道 ID 和 GTP-U 报头中的 QFI 将数据放入对应的输入队列。应理解, 不同 QFI 对应的输入队列不同。计算服务器对数据处理后, 为处理后的数据添加报头, 其中报头元素与接收到数据的报头信息一致, 具体可参考前述第一种场景种的情况一的相关介绍, 这里不再赘述。

处理方式二, 计算服务器具有数据包过滤功能, 可通过数据包过滤功能将数据包通过对应的 GTP-U 隧道发送给网络设备。具体可参考前述第一种场景种的情况二的相关介绍, 这里不再赘述。

处理方式三、如果计算服务器不具有数据包过滤功能, 可以将处理后的数据包通过对应 PDU 会话粒度隧道发送给网络设备。网络设备对数据包进行过滤将数据包映射到对应的 QoS 流。例如网络设备设置有数据包过滤功能。

第三种情况, 网络设备与计算服务器之间不存在隧道。

网络设备与计算服务器之间不存在隧道, 网络设备在 PDU 会话建立过程知道该 PDU 会话所有 QoS 流或部分 QoS 流需要进行本地处理, 以及本地处理的详细信息。

请参见图 20, 示出了计算服务器处理数据包的再一种流程。网络设备接收到核心网设备发送的数据后, 网络设备可去除原有报头 (例如称为报头 1), 生成新的数据包。该新的数据包的报头例如称为报头 2。举例来说, 该数据包是 IP 数据包, 网络设备未处理该数据包, 该数据包包括的报头 1 包括发送端 (CS, 云服务器) 的 IP 地址/port 信息和接收端 (终端 PDU 会话) 的 IP 地址/port 信息。网络设备处理该数据包之后, 该数据包包括的报头 2 包括发送端 (gNB) 的 IP 地址/port 信息和接收端 (RS, 计算服务器) 的 IP 地址/port 信息。不同之处在于, 报头 2 和报头 1 包括的发送端的 IP 地址/port 信息不同, 接收端的 IP 地址/port 信息不同。网络设备可基于 TCP/IP 或者 Ethernet 协议将处理后的数据包发送给计算服务器, 由计算服务器对该数据包进行处理。计算服务器处理数据包之后, 为该数据包添加报头, 例如报头 3, 其中, 报头 3 包括发端 CS 的 IP 地址/port 信息, 报头 3 包括收端终端的 PDU 会话 IP 地址。之后计算服务器将重新添加报头的数据包发送给网络设备。网络设备根据配置的 PDRs 对数据包进行过滤, 将数据包映射至对应的 QoS 流, 再映射到对应 DRB 发送给终端。

应理解, 第一终端执行的计算任务可在第二终端执行渲染。请参见图 21, 为本申请实施例提供的 D2D 本地渲染的流程示意图。该流程包括如下步骤:

S211、第一终端和第二终端交互计算服务的类型。

该 S211 的具体实现可参考前述图 9、图 10 和图 11 中相关实施例的介绍, 这里不再赘述。

S212、第二终端请求与 CS 建立针对第一终端的 PDU 会话, 并处理接收的数据。

第二终端接收到第一终端请求执行至少一个计算任务之后, 可请求建立与 CS 的 PDU 会话连接, 该 PDU 会话连接可用于中继第一终端的数据。同时, 可指示第二终端处理来自第一终端的数据。核心网设备可为该 PDU 会话分别分配针对第二终端和第一终端的 IP 地址, 以确定数据的发送端和接收端。可选的, 核心网设备可为第二终端配置 QoS 规则。

应理解, 该 QoS 规则可用于数据包到 DRB 映射。

S213、第一终端通过为第二终端配置 QoS 规则。

在本申请实施例中，第一终端和第二终端可通过 D2D 方式传输数据包。具体的，第一终端可通过侧行链路配置信息（sidelink reconfiguration, SL reconfiguration）为第二终端配置 QoS 规则，用于 D2D 侧的数据包到 DRB 映射。

5 需要说明的是，S213 不是必须执行的步骤，因此在图 21 中以虚线进行示意。

S214、第二终端对接收的数据包进行处理。

CS 发送的数据包到达第二终端，第二终端可以接收的数据包进行处理，例如去掉数据包原始的报头，对数据包数据进行处理。之后为处理后的数据包添加新的报头，例如该新的报头包括发送端即 CS 的 IP 地址/port 和接收端即第一终端的 IP 地址。之后第二终端可
10 基于 QoS 规则将 IP 数据包映射至对应的 QoS 流，再通过对应的 SLRB 发送给第一终端。

通过本申请实施例的上述步骤，在终端与终端间在多级处理计算任务的过程中，可保证数据包在传输和处理过程中 QoS 映射关系不变，从而保证业务的 QoS 需求。

上述本申请提供的实施例中，分别从终端和网络设备以及 MEC 网元等之间交互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
15 功能，终端和网络设备可以包括硬件结构和/或软件模块，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以硬件结构、软件模块、还是硬件结构加软件模块的方式来执行，取决于技术方案的特定应用和设计约束条件。

下面结合附图介绍本申请实施例中用来实现上述方法的通信装置。因此，上文中的内容均可以用于后续实施例中，重复的内容不再赘述。

20 图 22 为本申请实施例提供的通信装置 2200 的示意性框图。该通信装置 2200 可以对实现上述各个方法实施例中由终端或基站实现的功能或者步骤。该通信装置可以包括处理模块 2210 和收发模块 2220。可选的，还可以包括存储单元，该存储单元可以用于存储指令（代码或者程序）和/或数据。处理模块 2210 和收发模块 2220 可以与该存储单元耦合，例如，处理模块 2210 可以读取存储单元中的指令（代码或者程序）和/或数据，以实现相
25 应的方法。上述各个单元可以独立设置，也可以部分或者全部集成。

一些可能的实施方式中，通信装置 2200 能够对应实现上述方法实施例中终端的行为和功能。例如通信装置 2200 可以为终端，也可以为应用于终端中的部件（例如芯片或者电路）。收发模块 2220 可以用于执行图 9-图 17 任一所示的实施例中由终端或网络设备或第一 MEC 功能，以及其他设备所执行的全部接收或发送操作。处理模块 2210 用于执行如
30 图 9-图 17 任一所示的实施例中由终端或网络设备或第一 MEC 功能，以及其他设备所执行的除了收发操作之外的全部操作。

作为一种示例，处理模块 2210 用于确定第一 MEC 功能的第一计算能力，该第一 MEC 功能设置于通信装置内或者该第一 MEC 功能独立于通信装置，第一 MEC 功能的计算能力用于指示第一 MEC 功能能够提供的计算服务；

35 收发模块 2220 用于发送第一能力信息，该第一能力信息用于指示所述第一 MEC 功能的第一计算能力；

收发模块 2220 还接收来自第一终端的第一任务请求消息，该第一任务请求消息用于请求所述第一 MEC 功能执行所述第一终端的至少一个计算任务，其中，所述至少一个计算任务是依据所述第一能力信息确定的。

40 在一种可能的实现方式中，收发模块 2220 用于向第一 MEC 功能发送第一请求消息，

该第一请求消息用于请求获取所述第一 MEC 功能的计算能力；收发模块 2220 用于接收来自第一 MEC 功能的第一响应消息，该第一响应消息包括能力参数，所述能力参数用于确定第一 MEC 功能的第一计算能力，所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种。

5 在一种可能的实现方式中，收发模块 2220 用于向所述第一 MEC 功能发送第二请求消息，该第二请求消息用于请求与第一 MEC 功能建立连接；收发模块 2220 用于接收来自第一 MEC 功能的第二响应消息，该第二响应消息包括能力参数，所述能力参数用于确定所述第一 MEC 功能的第一计算能力，所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种。

10 在一种可能的实现方式中，收发模块 2220 用于接收来自第一 MEC 功能的第一更新消息，该第一更新消息包括第一 MEC 功能的第一计算能力的相关信息。

在一种可能的实现方式中，收发模块 2220 还用于接收来自第一终端的第三请求消息，所述第三请求消息用于请求获取 MEC 功能的计算能力。

在一种可能的实现方式中，收发模块 2220，具体用于广播第一能力信息。

15 在一种可能的实现方式中，处理模块 2210 还用于确定第一 MEC 功能的第一计算能力不满足至少一个计算任务的计算需求，所述收发模块 2220 用于向第二网络设备发送第三任务请求消息，第三任务请求消息用于请求第二网络设备执行所述至少一个计算任务；以及收发模块 2220 还用于接收来自第二网络设备针对所述第三任务请求消息的响应消息。

20 在一种可能的实现方式中，处理模块 2210 还用于确定第二网络设备侧的第二 MEC 功能的第二计算能力能够满足该至少一个计算任务的计算需求，该第二 MEC 功能设置于该第二网络设备内或者该第二 MEC 功能独立于该第二网络设备，其中，该第二 MEC 功能的第二计算能力是该第二 MEC 功能通知该第一 MEC 功能的。

25 在一种可能的实现方式中，收发模块 2220 用于接收来自第一 MEC 功能的指示信息，该指示信息用于指示第二终端，该第二终端能够提供计算服务；以及将来自第一终端的第三任务请求消息转发给该第二终端，该第三任务请求消息用于请求执行该至少一个计算任务。

30 作为另一种示例，收发模块 2220 用于接收来自第一网络设备的第一能力信息，该第一能力信息用于指示第一 MEC 功能的第一计算能力，该第一计算能力用于指示第一 MEC 功能能够提供的计算服务，该第一 MEC 功能设置于第一网络设备内或者该第一 MEC 功能独立于第一网络设备；以及向第一网络设备或者第一 MEC 功能发送第一任务请求消息，该第一任务请求消息用于请求执行至少一个计算任务；以及接收准入信息，并根据准入信息将至少一个计算任务的相关数据发送给所述第一 MEC 功能。

在一种可能的实现方式中，收发模块 2220 还用于向第一网络设备发送第二请求消息，所述第二请求消息用于请求获取 MEC 的计算能力。

35 在一种可能的实现方式中，处理模块 2210 用于根据第一能力信息确定驻留小区，所述驻留小区支持所述通信装置需要的计算服务；或者，根据第一能力信息从当前驻留小区切换到邻小区，所述当前驻留小区不满足所述通信装置的计算服务需求，所述邻小区满足所述通信装置的计算服务需求。

40 作为又一种示例，处理模块 2210 用于确定第二终端执行第一网络设备或所述通信装置的至少一个计算任务，所述第二终端能够提供计算服务，所述通信装置设置于所述第一

网络设备内或者所述通信装置独立于所述第一网络设备;

收发模块 2220 用于向第一网络设备发送会话请求消息, 该会话请求消息用于指示第一网络设备建立第一网络设备与所述通信装置之间的第一传输通道, 以及建立第一网络设备到所述第二终端的第二传输通道; 以及通过第一传输通道以及第二传输通道将所述至少一个计算任务的相关数据发送给所述第二终端。

在一种可能的实现方式中, 收发模块 2220 用于接收来自第二终端的注册请求消息, 该注册请求消息用于请求所述通信装置注册所述第二终端的计算能力信息, 所述至少一个计算任务是根据所述计算能力信息确定的。

作为又一种示例, 收发模块 2220 用于向 MEC 功能发送注册请求消息, 该注册请求消息用于请求所述 MEC 功能注册所述通信装置的计算能力信息, 所述计算能力信息用于指示所述通信装置能够提供的计算服务; 以及接收第一网络设备发送的第一任务请求消息, 所述第一任务请求消息用于请求执行所述第一网络设备或所述 MEC 功能的至少一个任务, 所述至少一个计算任务是根据所述计算能力信息确定的。

作为又一种示例, 处理模块 2210 用于确定第二终端, 所述第二终端能够提供计算服务;

收发模块 2220 用于向所述第二终端发送第一任务请求消息, 该第一请求消息用于请求所述第二终端执行所述通信装置的至少一个计算任务; 以及接收来自第二终端的准入信息, 并根据所述准入信息将所述至少一个计算任务的相关数据发送给所述第二终端。

在一种可能的实现方式中, 处理模块 2210 具体用于通过收发模块 2220 广播 MEC 服务信息, 所述 MEC 服务信息包括所述至少一个任务的计算需求参数; 以及通过收发模块 2220 接收来自某个终端针对所述 MEC 服务信息的响应消息, 并将该终端确定为所述第二终端。

在一种可能的实现方式中, 处理模块 2210 具体用于通过收发模块 2220 向第一网络设备发送所述第一任务请求消息; 以及接收来自所述第一网络设备的第一响应消息, 所述第一响应消息包括所述第二终端的信息。

作为又一种示例, 收发模块 2220 用于广播计算能力信息, 所述通信装置能够提供计算服务, 所述计算能力信息用于指示所述通信装置的第二计算能力; 以及接收来自第一终端的第一任务请求消息, 所述第一任务请求消息用于请求执行所述第一终端的至少一个计算任务。

在一种可能的实现方式中, 收发模块 2220 还用于将所述计算能力信息发送给 MEC 功能。

作为又一种示例, 收发模块 2220 用于接收来自第一终端的第一任务请求消息, 所述第一任务请求消息用于请求执行所述第一终端的计算任务; 以及向至少一个边缘计算设备发送第二任务请求消息, 每个边缘计算设备对应的所述第二任务请求消息用于请求该边缘计算设备执行所述计算任务中的部分计算子任务, 所述计算任务包括的多个计算子任务具有逻辑关系, 所述至少一个边缘计算设备包括第二 MEC 和/或第二终端, 所述第二终端能够提供计算服务。

在一种可能的实现方式中, 所述第一请求消息包括如下信息的一种或多种:

所述多个计算子任务、所述多个计算子任务之间的逻辑关系, 以及多个计算子任务中各个子任务的服务需求。

在一种可能的实现方式中，处理模块 2210 用于将所述计算任务划分为基于 SFC 的所述多个计算子任务，其中，所述第二任务请求消息包括第一信息和第二信息，所述第一信息用于指示所述多个计算子任务所述第二信息用于指示所述多个计算子任务之间的逻辑关系。

5 在一种可能的实现方式中，收发模块 2220，用于接收来自第二终端的注册请求消息，所述注册请求消息用于请求注册所述第二终端的计算能力，以及所述第二终端的计算负载，所述第二终端能够提供计算服务。

10 作为又一种示例，收发模块 2220 用于向第一 MEC 功能发送第一任务请求消息，所述第一任务请求消息用于请求执行所述第一终端的计算任务；

在一种可能的实现方式中，所述第一请求消息包括如下信息的一种或多种：

所述多个计算子任务、所述多个计算子任务之间的逻辑关系，以及多个计算子任务中各个子任务的服务需求。

15 作为又一种示例，收发模块 2220 用于接收第一信息，所述第一信息用于指示终端建立的协议数据单元（protocol data unit）PDU 会话与计算服务器的映射关系；

处理模块 2210 用于根据所述映射关系对来自所述终端的至少一个计算任务的相关数据进行处理。

在一种可能的实现方式中，所述网络设备被配置数据包过滤规则 PDR，所述 PDR 用于保证所述至少一个计算任务关联的数据流映射到正确的服务质量 QoS 流。

20 作为又一种示例，处理模块 2210 用于确定第一信息，所述第一信息包括所述通信装置建立的 PDU 会话与远端服务器的对应关系，所述第一信息用于指示 MEC 功能对来自所述远端服务器的通信装置的至少一个计算任务的相关数据进行处理；

收发模块 2220 用于向所述 MEC 功能发送第一信息；

25 在一种可能的实现方式中，收发模块 2220 用于向所述 MEC 功能发送第一任务请求消息，所述第一任务请求消息用于请求所述 MEC 功能执行所述通信装置的至少一个计算任务，所述第一任务请求消息包括所述远端服务器的相关信。

在一种可能的实现方式中，收发模块 2220 用于接收来自所述 MEC 功能发送第一任务请求响应消息，所述第一任务请求消息包括所述 MEC 功能选择的计算服务器的相关信息。

30 应理解，本申请实施例中的处理模块 2210 可以由处理器或处理器相关电路组件实现，收发模块 2220 可以由收发器或收发器相关电路组件或者通信接口实现。

应理解，本申请实施例中的处理模块 2210 可以由处理器或处理器相关电路组件实现，收发模块 2220 可以由收发器或收发器相关电路组件或者通信接口实现。

35 如图 23 所示为本申请实施例提供的通信装置 2300，其中，通信装置 2300 可以是终端，能够实现本申请实施例提供的方法中终端的功能，或者，通信装置 2300 可以是网络设备，能够实现本申请实施例提供的方法中网络设备的功能；通信装置 2300 也可以是能够支持终端实现本申请实施例提供的方法中对应的功能的装置，或者能够支持网络设备实现本申请实施例提供的方法中对应的功能的装置。其中，该通信装置 2300 可以为芯片或芯片系统。本申请实施例中，芯片系统可以由芯片构成，也可以包含芯片和其他分立器件。

40 在硬件实现上，上述收发模块 820 可以为收发器，收发器集成在通信装置 2300 中构成通信接口 2310。

通信装置 2300 包括至少一个处理器 2320, 用于实现或用于支持通信装置 2300 实现本申请实施例提供的方法中网络设备或终端的功能。具体参见方法示例中的详细描述, 此处不做赘述。

通信装置 2300 还可以包括至少一个存储器 2330, 用于存储程序指令和/或数据。存储器 2330 和处理器 2320 耦合。本申请实施例中的耦合是装置、单元或模块之间的间接耦合或通信连接, 可以是电性, 机械或其它的形式, 用于装置、单元或模块之间的信息交互。处理器 2320 可能和存储器 2330 协同操作。处理器 2320 可能执行存储器 2330 中存储的程序指令和/或数据, 以使得通信装置 2300 实现相应的方法。所述至少一个存储器中的至少一个可以包括于处理器中。

通信装置 2300 还可以包括通信接口 2310, 用于通过传输介质和其它设备进行通信, 从而用于通信装置 2300 中的装置可以和其它设备进行通信。示例性地, 当该通信装置为终端时, 该其它设备为网络设备; 或者, 当该通信装置为网络设备时, 该其它设备为终端。处理器 2320 可以利用通信接口 2310 收发数据。通信接口 2310 具体可以是收发器。

本申请实施例中不限定上述通信接口 2310、处理器 2320 以及存储器 2330 之间的具体连接介质。本申请实施例在图 23 中以存储器 2330、处理器 2320 以及通信接口 2310 之间通过总线 2340 连接, 总线在图 23 中以粗线表示, 其它部件之间的连接方式, 仅是进行示意性说明, 并不引以为限。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示, 图 23 中仅用一条粗线表示, 但并不表示仅有一根总线或一种类型的总线。

在本申请实施例中, 处理器 2320 可以是通用处理器、数字信号处理器、专用集成电路、现场可编程门阵列或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件, 可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者任何常规的处理器等。结合本申请实施例所公开的方法的步骤可以直接体现为硬件处理器执行完成, 或者用处理器中的硬件及软件模块组合执行完成。

在本申请实施例中, 存储器 2330 可以是非易失性存储器, 比如硬盘 (hard disk drive, HDD) 或固态硬盘 (solid-state drive, SSD) 等, 还可以是易失性存储器 (volatile memory), 例如随机存取存储器 (random-access memory, RAM)。存储器是能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质, 但不限于此。本申请实施例中的存储器还可以是电路或者其它任意能够实现存储功能的装置, 用于存储程序指令和/或数据。

需要说明的是, 上述实施例中的通信装置可以是终端也可以是电路, 也可以是应用于终端中的芯片或者其他具有上述终端功能的组合器件、部件等。当通信装置是终端时, 收发模块可以是收发器, 可以包括天线和射频电路等, 处理模块可以是处理器, 例如: 中央处理模块 (central processing unit, CPU)。当通信装置是具有上述终端功能的部件时, 收发模块可以是射频单元, 处理模块可以是处理器。当通信装置是芯片或芯片系统时, 收发模块可以是芯片或芯片系统的输入输出接口、处理模块可以是芯片或芯片系统的处理器。

本申请实施例还提供一种通信系统, 具体的, 通信系统包括网络设备和终端, 或者还可以包括更多个网络设备和多个终端。示例性的, 通信系统包括用于实现上述图 4-图 7 任一实施例的相关功能的网络设备和终端。

所述网络设备分别用于实现上述图 9-图 17 任一实施例任一相关设备部分的功能。具体请参考上述方法实施例中的相关描述, 这里不再赘述。

本申请实施例中还提供一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行图 9-图 17 任一实施例中任一相关设备执行的方法。

本申请实施例中还提供一种计算机程序产品，包括指令，当其在计算机上运行时，使得计算机执行图 9-图 17 任一实施例中任一相关设备执行的方法。

5 本申请实施例提供了一种芯片系统，该芯片系统包括处理器，还可以包括存储器，用于实现前述方法中图 9-图 17 任一实施例中任一相关设备的功能；或者用于实现前述方法中图 9-图 17 任一实施例中任一相关设备的功能。该芯片系统可以由芯片构成，也可以包含芯片和其他分立器件。

10 应理解，本申请实施例中的术语“系统”和“网络”可被互换使用。“至少一个”是指一个或者多个，“多个”是指两个或两个以上。“和/或”，描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 的情况，其中 A，B 可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。“以下至少一项(个)”或其类似表达，是指的这些项中的任意组合，包括单项(个)或复数项(个)的任意组合。例如，a，b 或 c 中的至少一项(个)，可以表示：a，b，c，a 15 和 b，a 和 c，b 和 c，或 a、b 和 c，其中 a，b，c 可以是单个，也可以是多个。

以及，除非有相反的说明，本申请实施例提及“第一”、“第二”等序数词是用于对多个对象进行区分，不用于限定多个对象的顺序、时序、优先级或者重要程度。例如，第一信息和第二信息，只是为了区分不同的指示信息，而并不是表示这两种信息的优先级、或者重要程度等的不同。

20 应理解，在本申请的各种实施例中，上述各过程的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

另外，在本申请实施例中，“示例性的”一词用于表示例子或说明。本申请实施例汇总被描述为“示例”的任何实施例或实现方案不应被解释为比其他实施例或实现方案更优选。也就是，使用“示例”一词旨在以具体方式呈现概念。

25 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种说明性逻辑块(illustrative logical block)和步骤(step)，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。

30 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。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组 35 件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

40 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络

单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

5 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（read-only memory, ROM）、随机存取存储器（random access memory, RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

10 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

1、一种网络边缘计算方法，其特征在于，包括：

第一网络设备确定第一多接入网络边缘计算 MEC 功能的第一计算能力，第一 MEC 功能的所述第一计算能力用于指示第一 MEC 功能所能提供计算服务的能力；

5 所述第一网络设备发送第一能力信息，所述第一能力信息包括用于确定所述第一计算能力的能力参数，所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种；

10 第一网络设备接收来自第一终端的第一任务请求消息，所述第一任务请求消息用于请求所述第一 MEC 功能执行所述第一终端的至少一个计算任务，其中，所述至少一个计算任务是根据所述第一能力信息确定的；

所述第一网络设备向所述第一 MEC 功能发送所述第一任务请求消息。

2、如权利要求 1 所述的方法，其特征在于，所述方法还包括：

所述第一网络设备接收来自第一终端的第二请求消息，所述第二请求消息用于请求获取 MEC 功能的计算能力。

15 3、如权利要求 1 或 2 所述的方法，其特征在于，所述第一网络设备发送第一能力信息，包括：

所述第一网络设备广播所述第一能力信息。

4、如权利要求 1-3 任一项所述的方法，其特征在于，所述第一网络设备确定所述第一 MEC 功能的第一计算能力，包括：

20 所述第一网络设备向所述第一 MEC 功能发送第一请求消息，所述第一请求消息用于请求获取所述第一 MEC 功能的所述第一计算能力；

所述第一网络设备接收来自所述第一 MEC 功能的第一响应消息，该第一响应消息包括所述能力参数。

25 5、如权利要求 1-3 任一项所述的方法，其特征在于，所述第一网络设备确定第一 MEC 功能的第一计算能力，包括：

所述第一网络设备向所述第一 MEC 功能发送第二请求消息，所述第二请求消息用于请求所述第一网络设备与所述第一 MEC 功能建立连接；

所述第一网络设备接收来自所述第一 MEC 功能的第二响应消息，所述第二响应消息包括所述能力参数。

30 6、如权利要求 1-3 任一项所述的方法，其特征在于，所述第一网络设备确定第一 MEC 功能的第一计算能力，包括：

所述第一网络设备接收来自所述第一 MEC 功能的第一更新消息，所述第一更新消息包括与所述第一 MEC 功能的所述第一计算能力相关的信息。

7、如权利要求 1-6 任一项所述的方法，其特征在于，所述方法还包括：

35 所述第一网络设备确定所述第一 MEC 功能的所述第一计算能力不满足所述至少一个计算任务的计算需求，所述第一网络设备向第二网络设备发送所述第一任务请求消息；

所述第一网络设备接收来自所述第二网络设备针对所述第一任务请求消息的响应消息。

8、如权利要求 7 所述的方法，其特征在于，所述方法还包括：

所述第一网络设备确定第二网络设备侧的第二 MEC 功能的第二计算能力能够满足所述至少一个计算任务的计算需求,其中,所述第二 MEC 功能的所述第二计算能力是所述第二 MEC 功能通知所述第一 MEC 功能的。

9、如权利要求 1-3 任一项所述的方法,其特征在于,所述方法还包括:

5 所述第一网络设备接收来自所述第一 MEC 功能的指示信息,所述指示信息用于指示第二终端,所述第二终端能够提供计算服务;

所述第一网络设备将来自所述第一终端的所述第二任务请求消息转发给所述第二终端,所述第二任务请求消息用于请求执行所述至少一个计算任务。

10、一种网络边缘计算方法,其特征在于,包括:

10 第一终端接收来自第一网络设备的第一能力信息,所述第一能力信息用于指示所述第一 MEC 功能的第一计算能力,所述第一计算能力用于指示所述第一 MEC 功能能够提供计算服务的能力,所述第一能力信息包括用于确定所述第一计算能力的能力参数,所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种;

15 所述第一终端向所述第一网络设备或者所述第一 MEC 功能发送第一任务请求消息,该第一任务请求消息用于请求执行至少一个计算任务;

所述第一终端接收来自所述第一 MEC 功能的准入信息,并根据准入信息将所述至少一个计算任务的相关数据发送给所述第一 MEC 功能。

11、如权利要求 10 所述的方法,其特征在于,所述方法还包括:

20 所述第一终端向所述第一网络设备发送第二请求消息,所述第二请求消息用于请求获取 MEC 的计算能力。

12、如权利要求 10 或 11 所述的方法,其特征在于,所述方法还包括:

所述第一终端根据所述第一能力信息确定驻留小区,所述驻留小区支持所述第一终端需要的计算服务;或者,

25 所述第一终端根据所述第一能力信息从当前驻留小区切换到邻小区,所述当前驻留小区不满足所述第一终端的计算服务需求,所述邻小区满足所述第一终端的计算服务需求。

13、如权利要求 11 所述的方法,其特征在于,所述第二请求消息为网络边缘能力查询请求消息;或者,所述第二请求消息为按需系统消息。

14、一种通信装置,其特征在于,包括收发模块和处理模块,其中:

30 所述处理模块用于确定第一多接入网络边缘计算 MEC 功能的第一计算能力,第一 MEC 功能的所述第一计算能力用于指示第一 MEC 功能能够提供计算服务的能力;

所述收发模块用于发送第一能力信息,所述第一能力信息包括用于确定所述第一计算能力的能力参数,所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种;

35 所述收发模块还用于接收来自第一终端的第一任务请求消息,所述第一任务请求消息用于请求所述第一 MEC 功能执行所述第一终端的至少一个计算任务,其中,所述至少一个计算任务是根据所述第一能力信息确定的;

所述收发模块还用于向所述第一 MEC 功能发送所述第一任务请求消息。

15、如权利要求 14 所述的通信装置,其特征在于,所述收发模块还用于接收来自第一终端的第二请求消息,所述第二请求消息用于请求获取 MEC 功能的计算能力。

40 16、如权利要求 14 或 15 所述的通信装置,其特征在于,所述收发模块具体用于广播

所述第一能力信息。

17、如权利要求 14-16 任一项所述的通信装置，其特征在于，所述收发模块具体用于：
向所述第一 MEC 功能发送第一请求消息，所述第一请求消息用于请求获取所述第一 MEC 功能的所述第一计算能力；

5 接收来自所述第一 MEC 功能的第一响应消息，该第一响应消息包括所述能力参数。

18、如权利要求 14-16 任一项所述的通信装置，其特征在于，所述收发模块具体用于：
所述通信装置向所述第一 MEC 功能发送第二请求消息，所述第二请求消息用于请求
所述通信装置与所述第一 MEC 功能建立连接；

10 所述通信装置接收来自所述第一 MEC 功能的第二响应消息，所述第二响应消息包括
所述能力参数。

19、如权利要求 14-16 任一项所述的通信装置，其特征在于，所述收发模块具体用于：
所述通信装置接收来自所述第一 MEC 功能的第一更新消息，所述第一更新消息包括
与所述第一 MEC 功能的所述第一计算能力相关的信息。

20、如权利要求 14-19 任一项所述的通信装置，其特征在于，所述收发模块还用于：
15 在所述处理模块确定所述第一 MEC 功能的所述第一计算能力不满足所述至少一个计
算任务的计算需求时，向第二网络设备发送所述第一任务请求消息；

接收来自所述第二网络设备针对所述第一任务请求消息的响应消息。

21、如权利要求 20 所述的通信装置，其特征在于，所述处理模块还用于：

20 确定第二网络设备侧的第二 MEC 功能的第二计算能力能够满足所述至少一个计算任
务的计算需求，其中，所述第二 MEC 功能的所述第二计算能力是所述第二 MEC 功能通知
所述第一 MEC 功能的。

22、如权利要求 14-16 任一项所述的通信装置，其特征在于，所述收发模块还用于：

接收来自所述第一 MEC 功能的指示信息，所述指示信息用于指示第二终端，所述第
二终端能够提供计算服务；

25 将来自所述第一终端的所述第二任务请求消息转发给所述第二终端，所述第二任务请
求消息用于请求执行所述至少一个计算任务。

23、一种通信装置，其特征在于，包括收发模块和处理模块，其中：

所述收发模块用于接收来自第一网络设备的第一能力信息，所述第一能力信息用于指
示所述第一 MEC 功能的第一计算能力，所述第一计算能力用于指示所述第一 MEC 功能所
能提供计算服务的能力，所述第一能力信息包括用于确定所述第一计算能力的能力参数，
30 所述能力参数包括支持的计算服务类型、硬件参数、计算负载中的一种或多种；

所述收发模块用于向所述第一网络设备或者所述第一 MEC 功能发送所述处理模块确
定的第一任务请求消息，该第一任务请求消息用于请求执行至少一个计算任务；

35 所述收发模块用于接收来自所述第一 MEC 功能的准入信息，并根据准入信息将所述
至少一个计算任务的相关数据发送给所述第一 MEC 功能。

24、如权利要求 23 所述的通信装置，其特征在于，所述收发模块还用于：

所述通信装置向所述第一网络设备发送第二请求消息，所述第二请求消息用于请求获
取 MEC 的计算能力。

25、如权利要求 23 或 24 所述的通信装置，其特征在于，所述处理模块还用于：

40 所述通信装置根据所述第一能力信息确定驻留小区，所述驻留小区支持所述通信装置

需要的计算服务；或者，

所述第一终端根据所述第一能力信息从当前驻留小区切换到邻小区，所述当前驻留小区不满足所述第一终端的计算服务需求，所述邻小区满足所述第一终端的计算服务需求。

5 26、如权利要求 24 所述的通信装置，其特征在于，所述第二请求消息为网络边缘能力查询请求消息；或者，所述第二请求消息为按需系统消息。

27、一种通信装置，其特征在于，所述通信装置包括处理器和存储器，所述存储器用于存储计算机程序，所述处理器用于执行存储在所述存储器上的计算机程序，使得所述装置执行如权利要求 1~9 或 10~13 中任一项所述的方法。

10 28、一种通信装置，其特征在于，所述通信装置包括处理器和通信接口，所述通信接口用于输入和/或输出信息，所述处理器用于执行计算机程序，使得所述装置执行如权利要求 1~9 或 10~13 中任一项所述的方法。

29、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有计算机程序，所述计算机程序当被计算机执行时，使所述计算机执行如权利要求 1~9 或 10~13 中任意一项所述的方法。

15

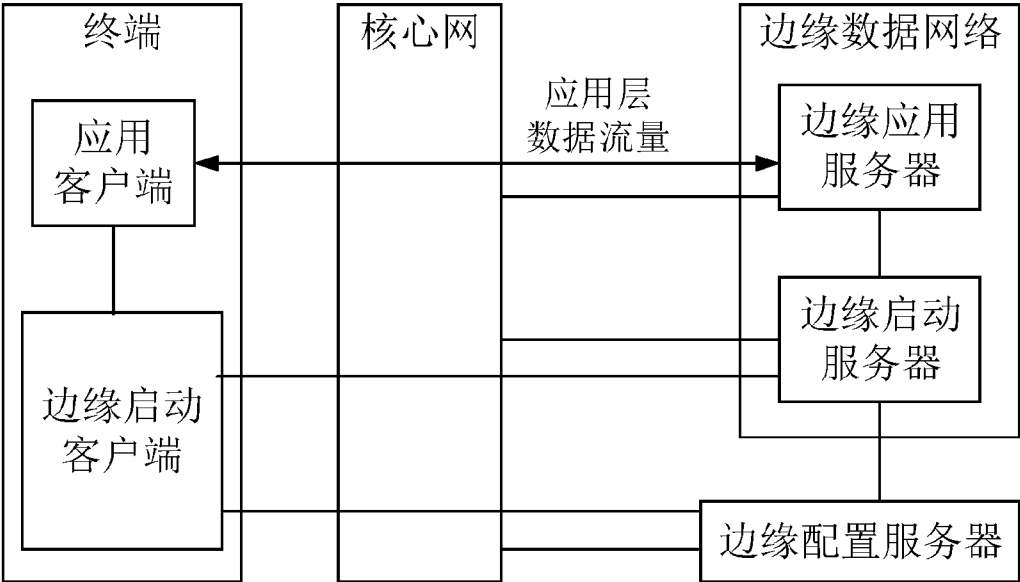


图 1

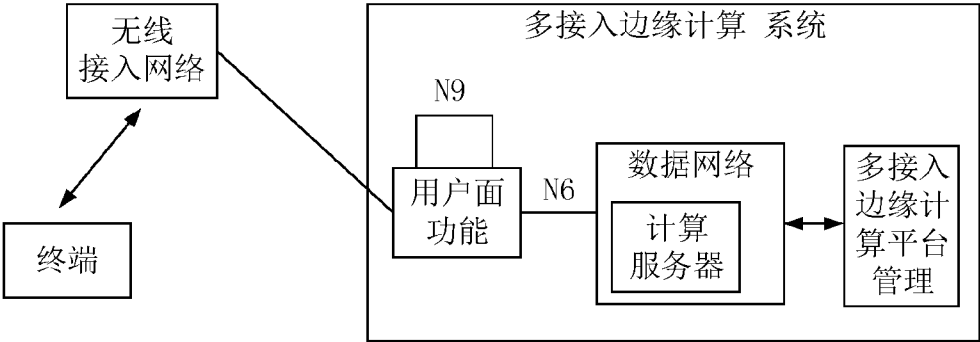


图 2

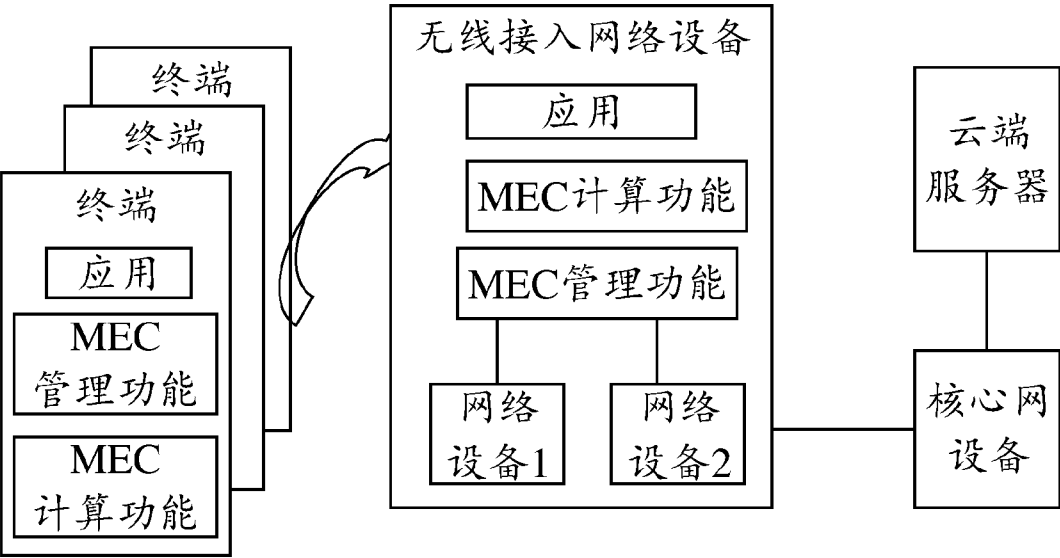


图 3

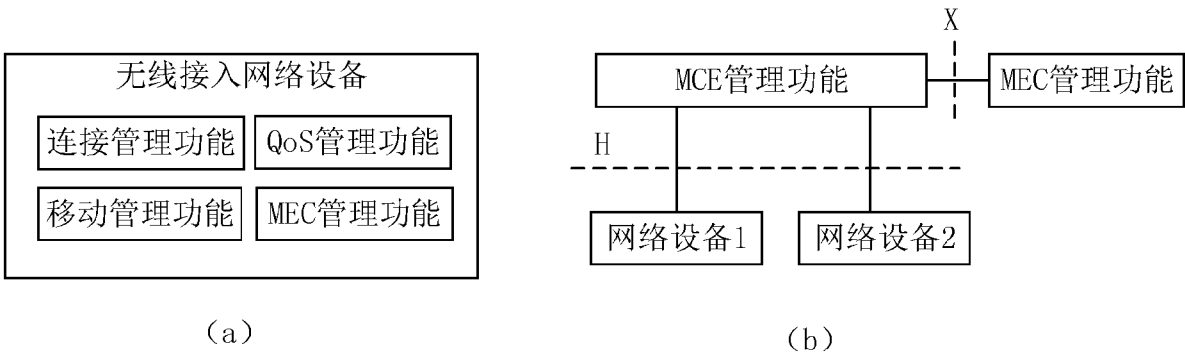


图 4

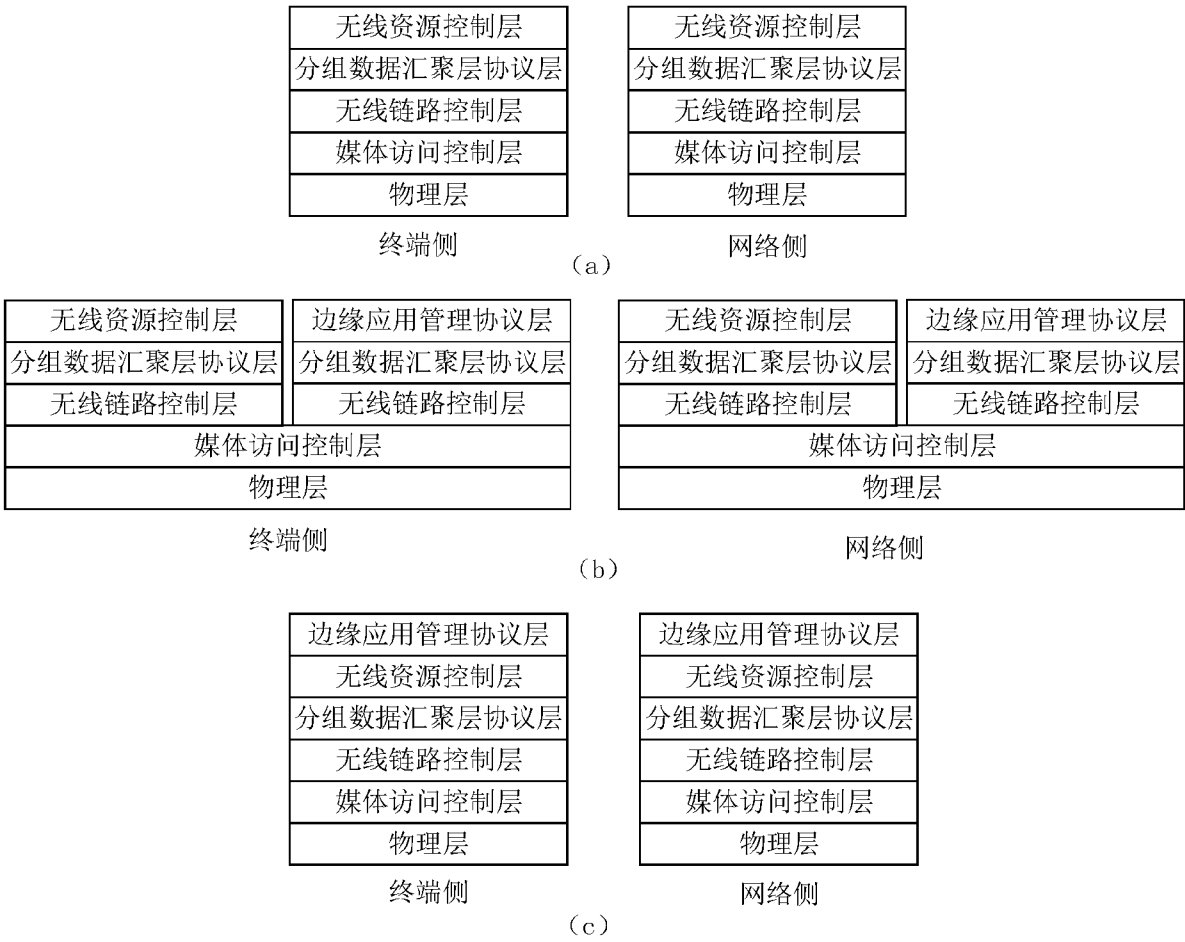


图 5

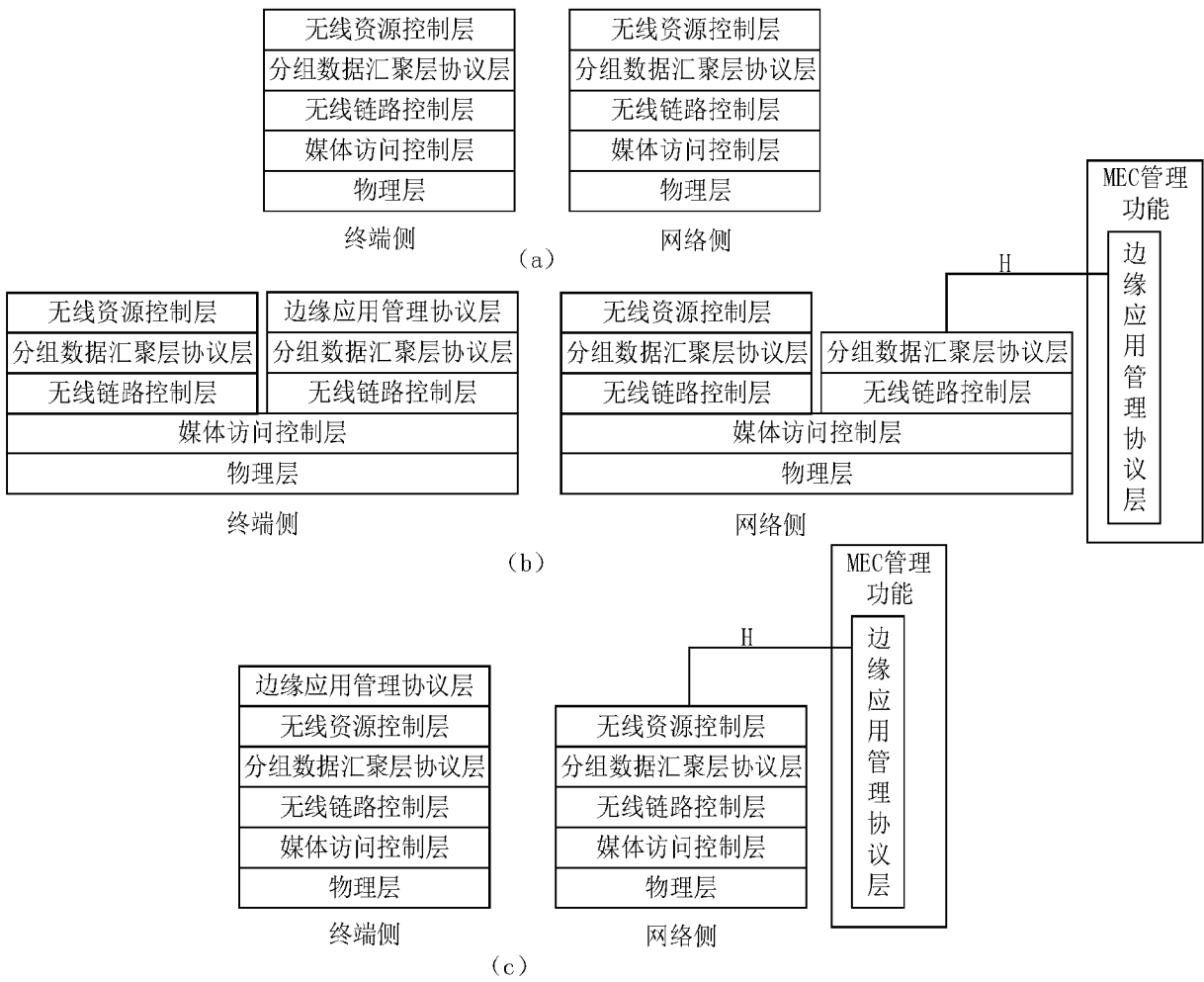


图 6

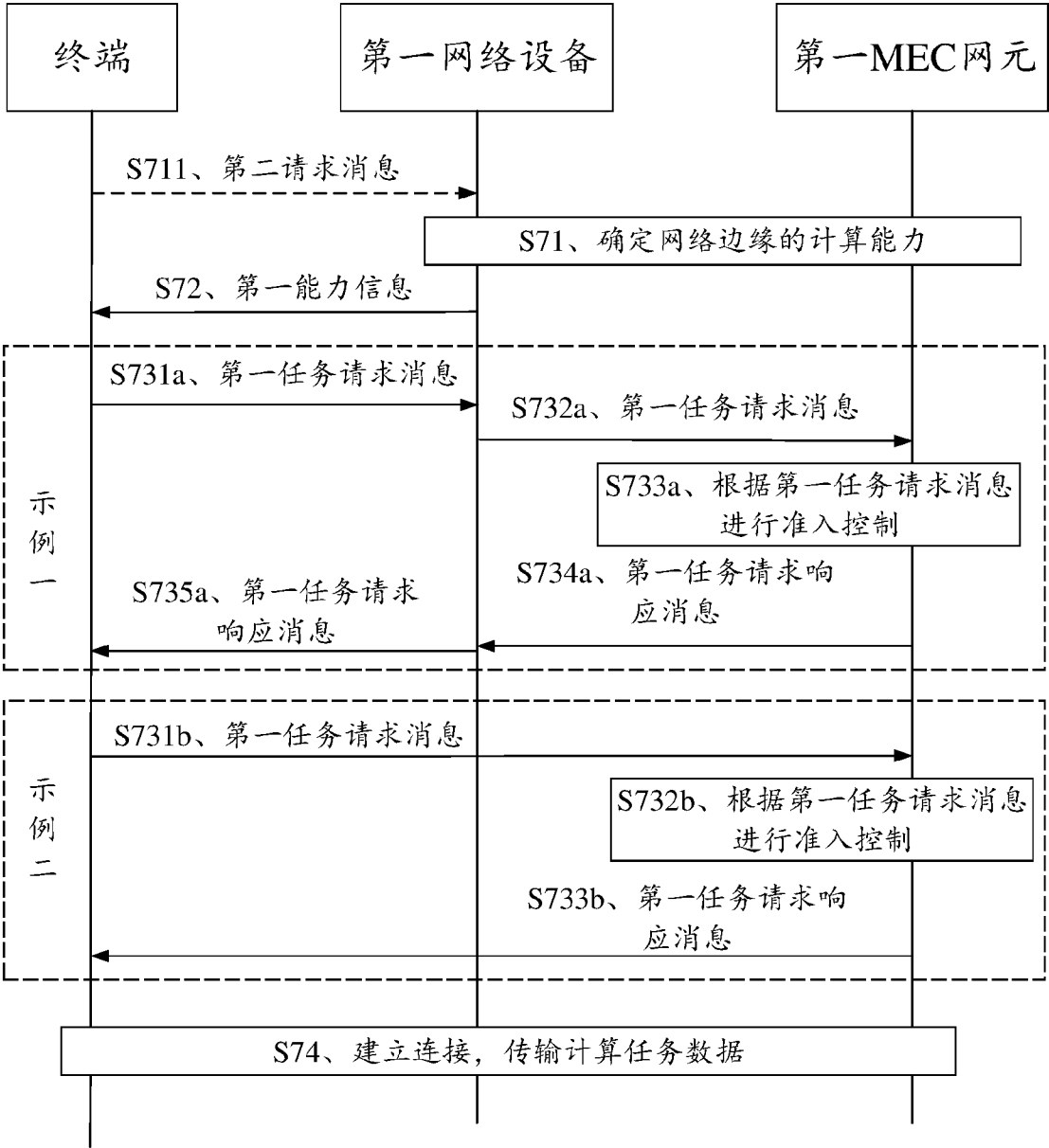


图 7

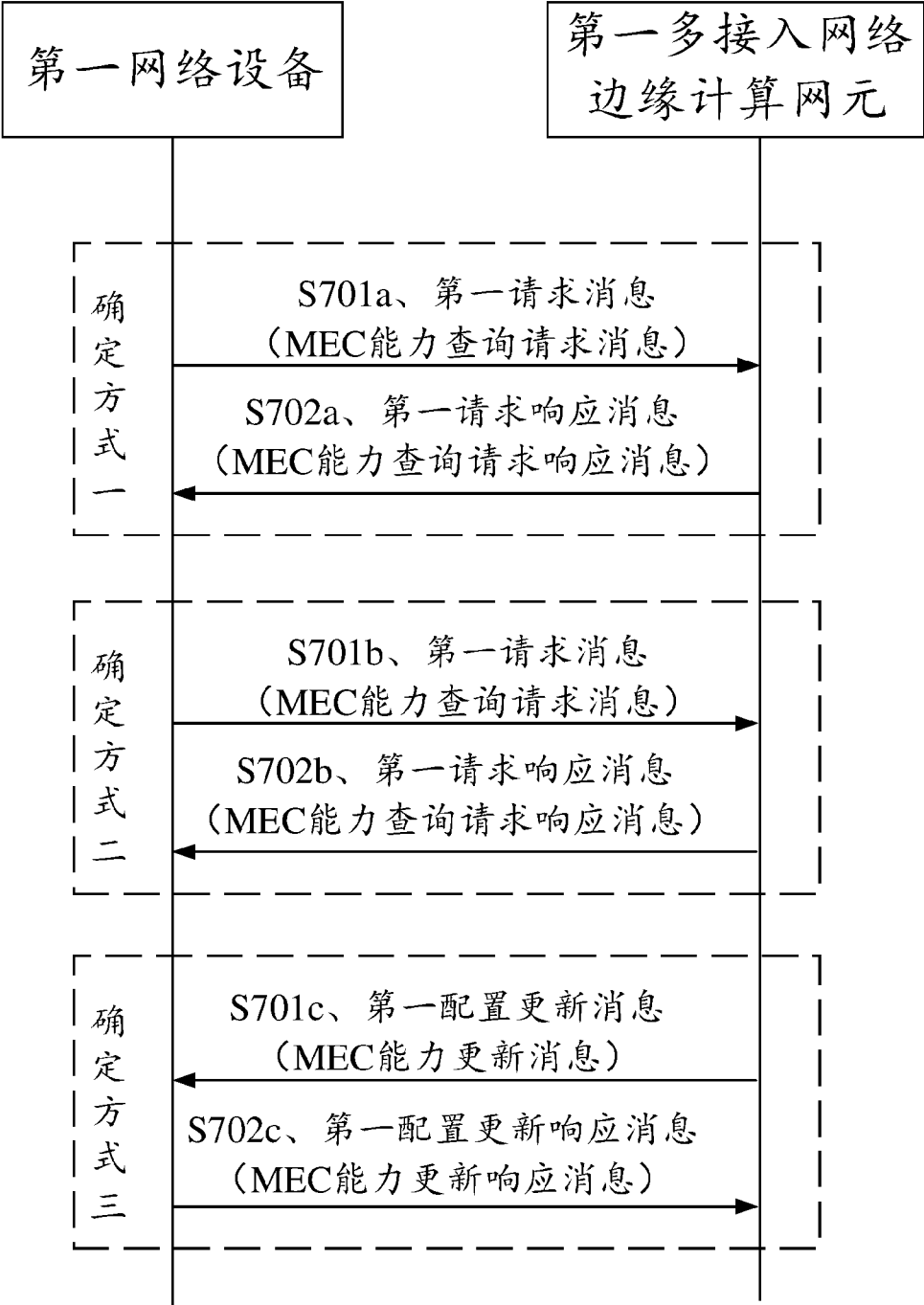


图 7a

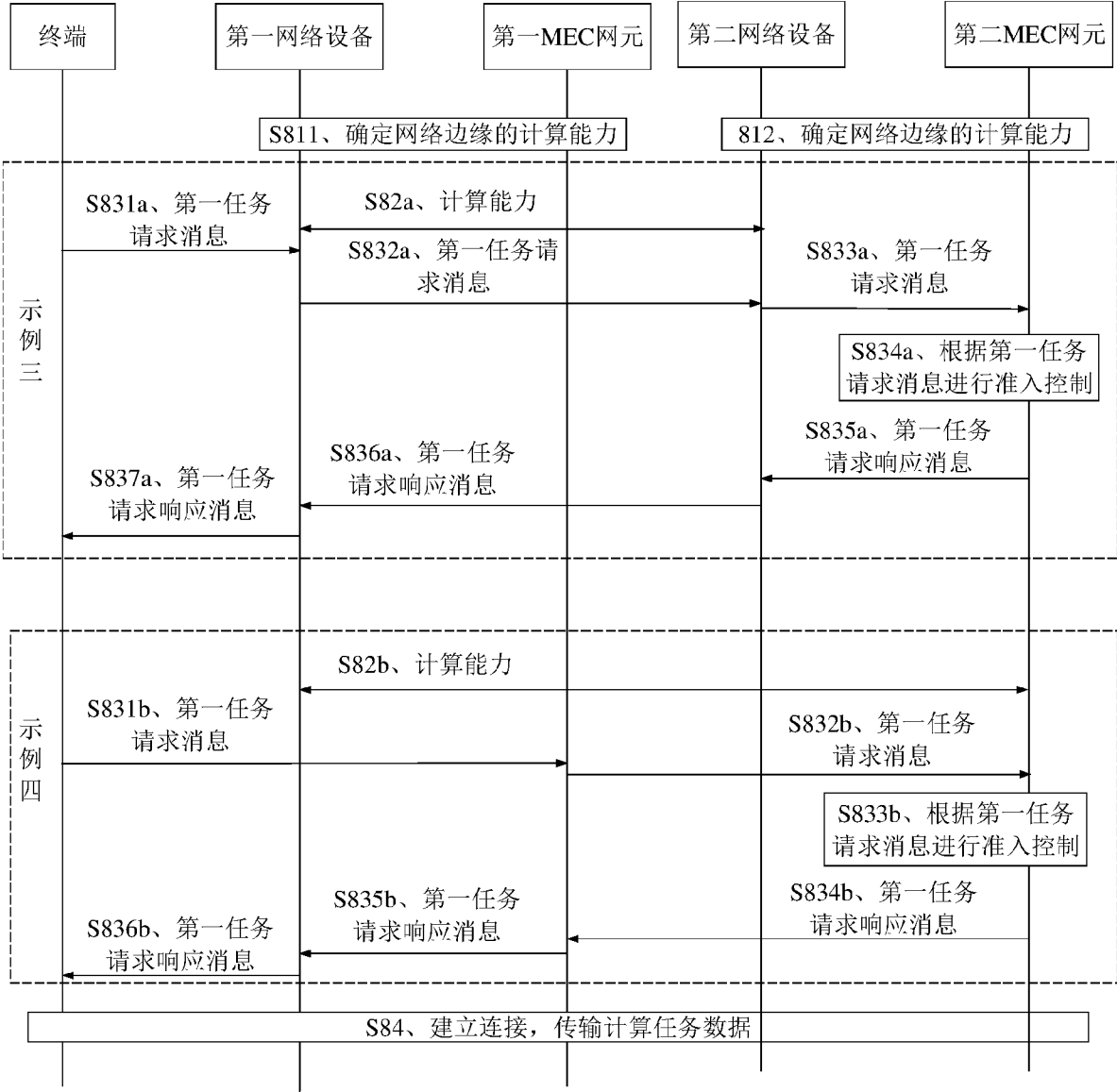


图 8

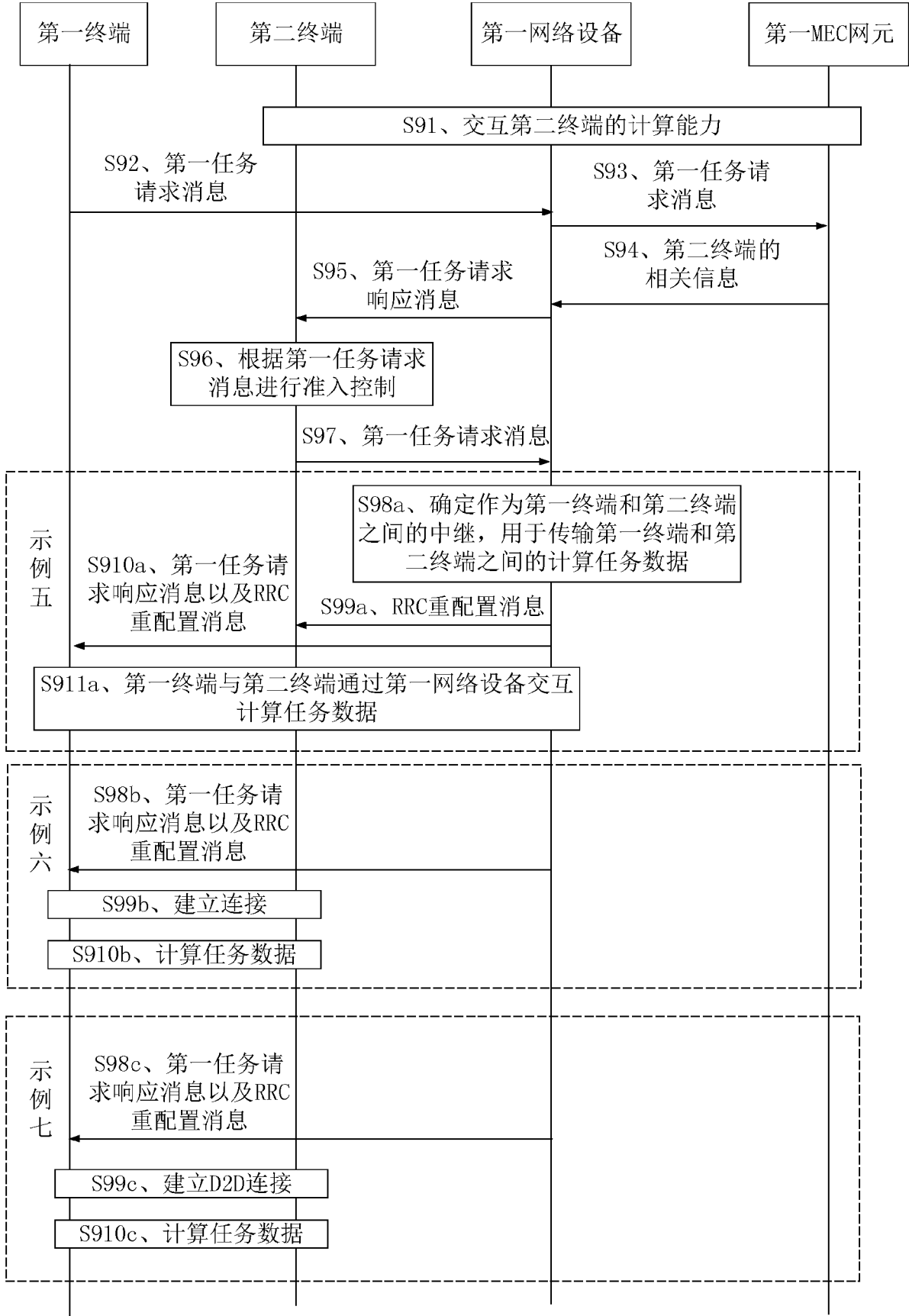


图 9

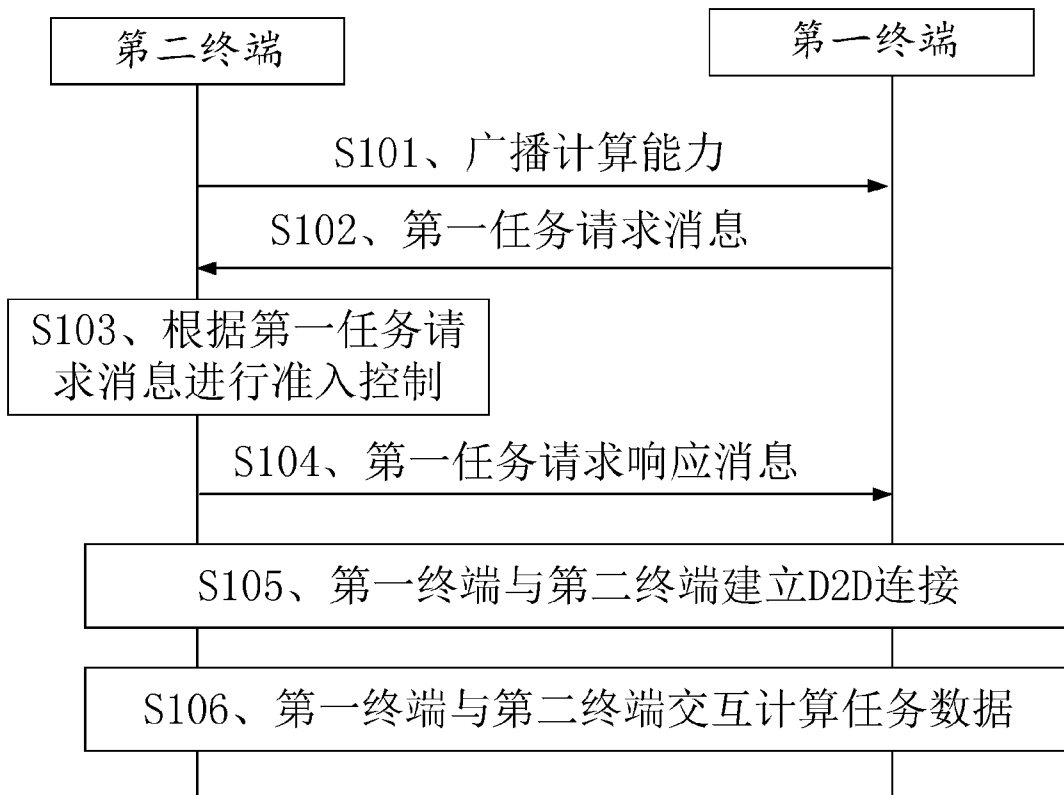


图 10

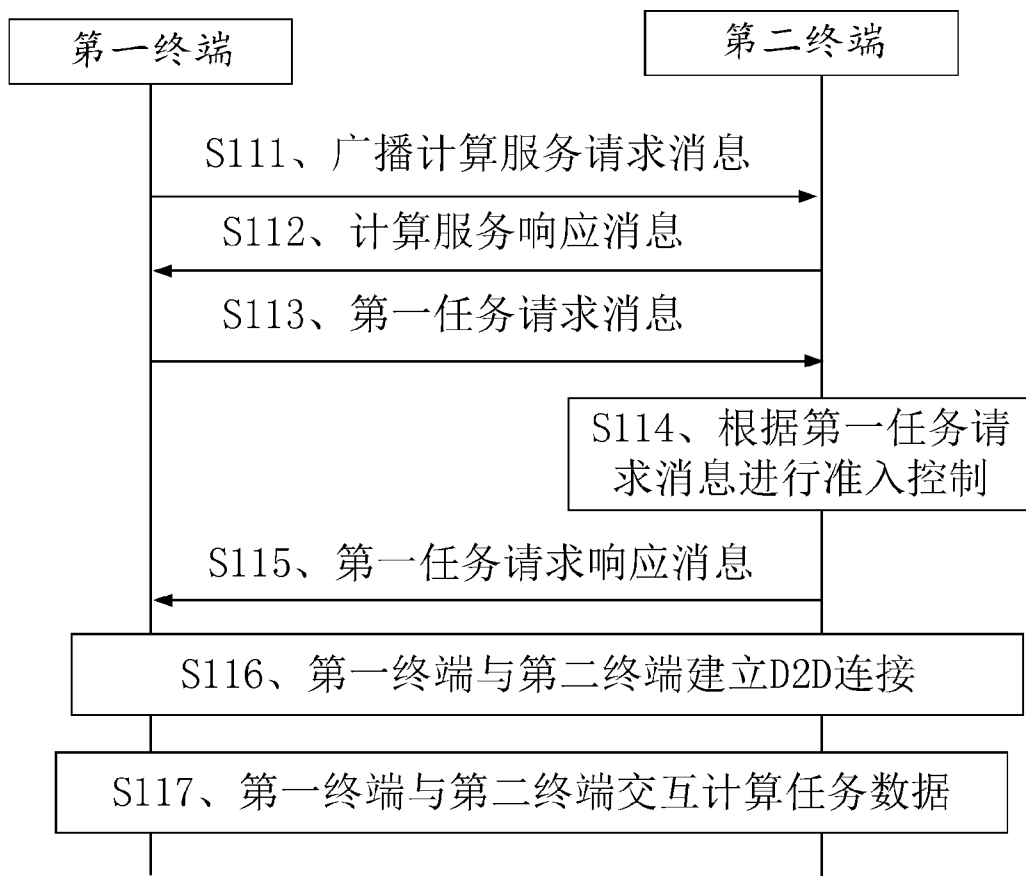


图 11

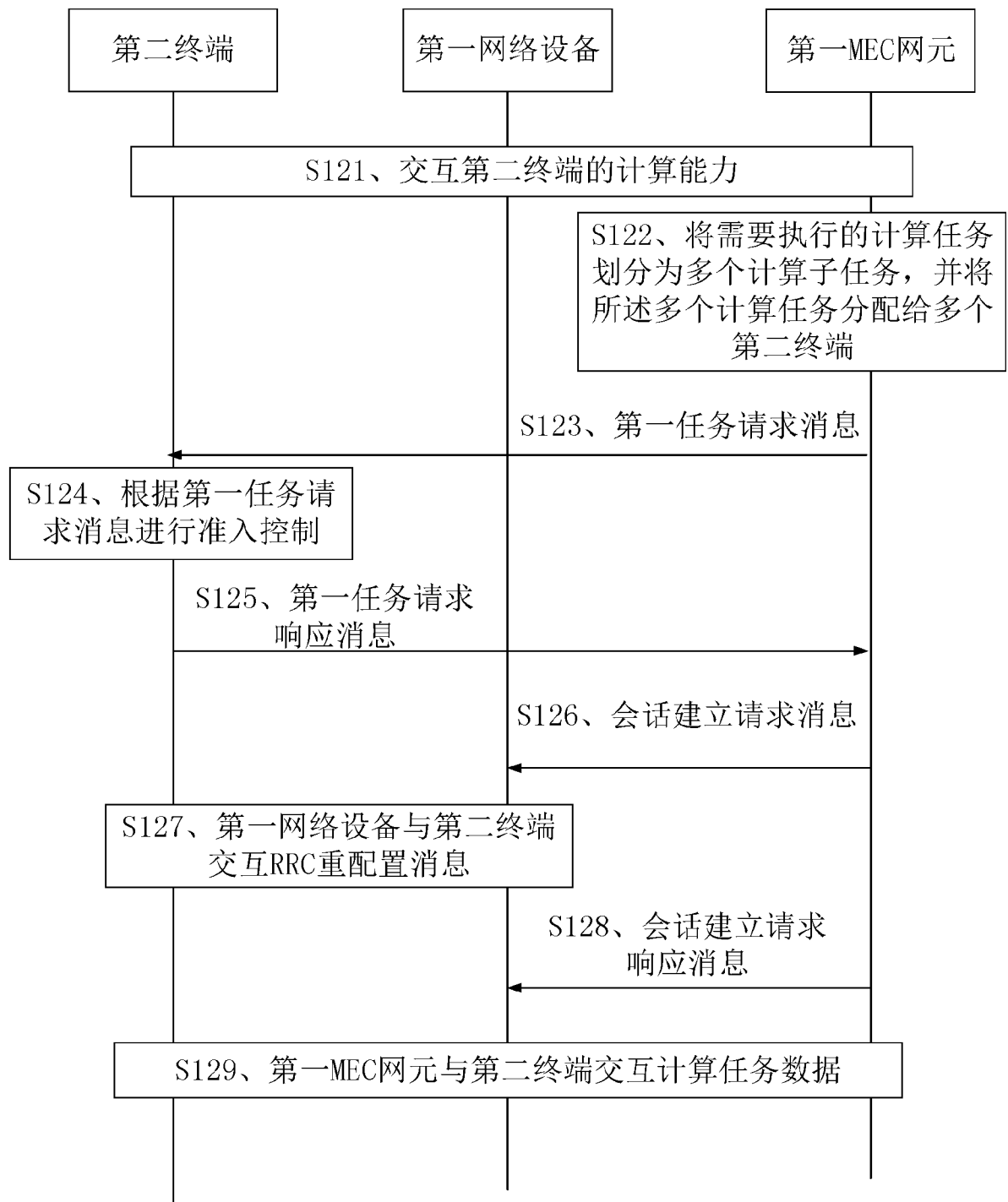


图 12

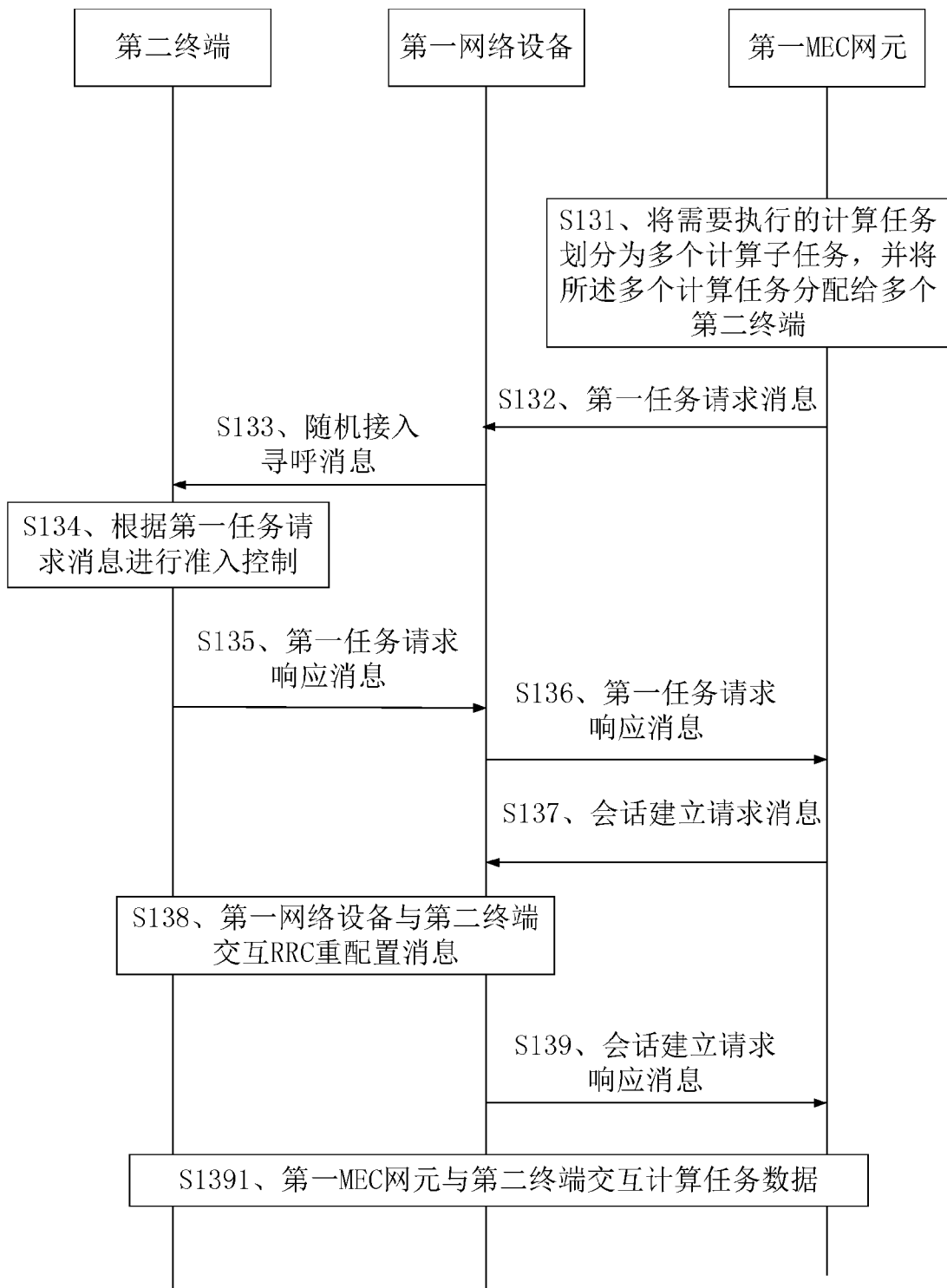


图 13

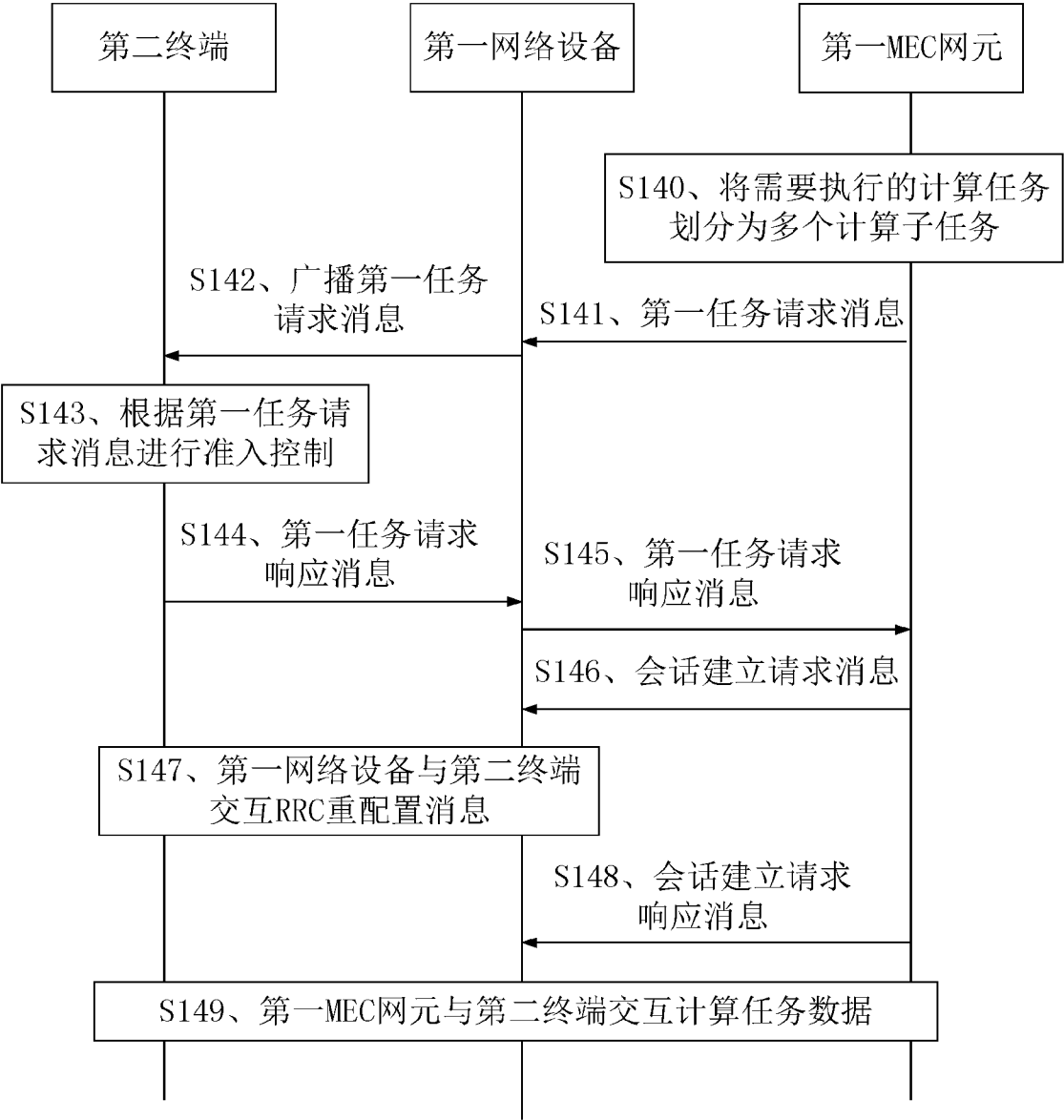


图 14

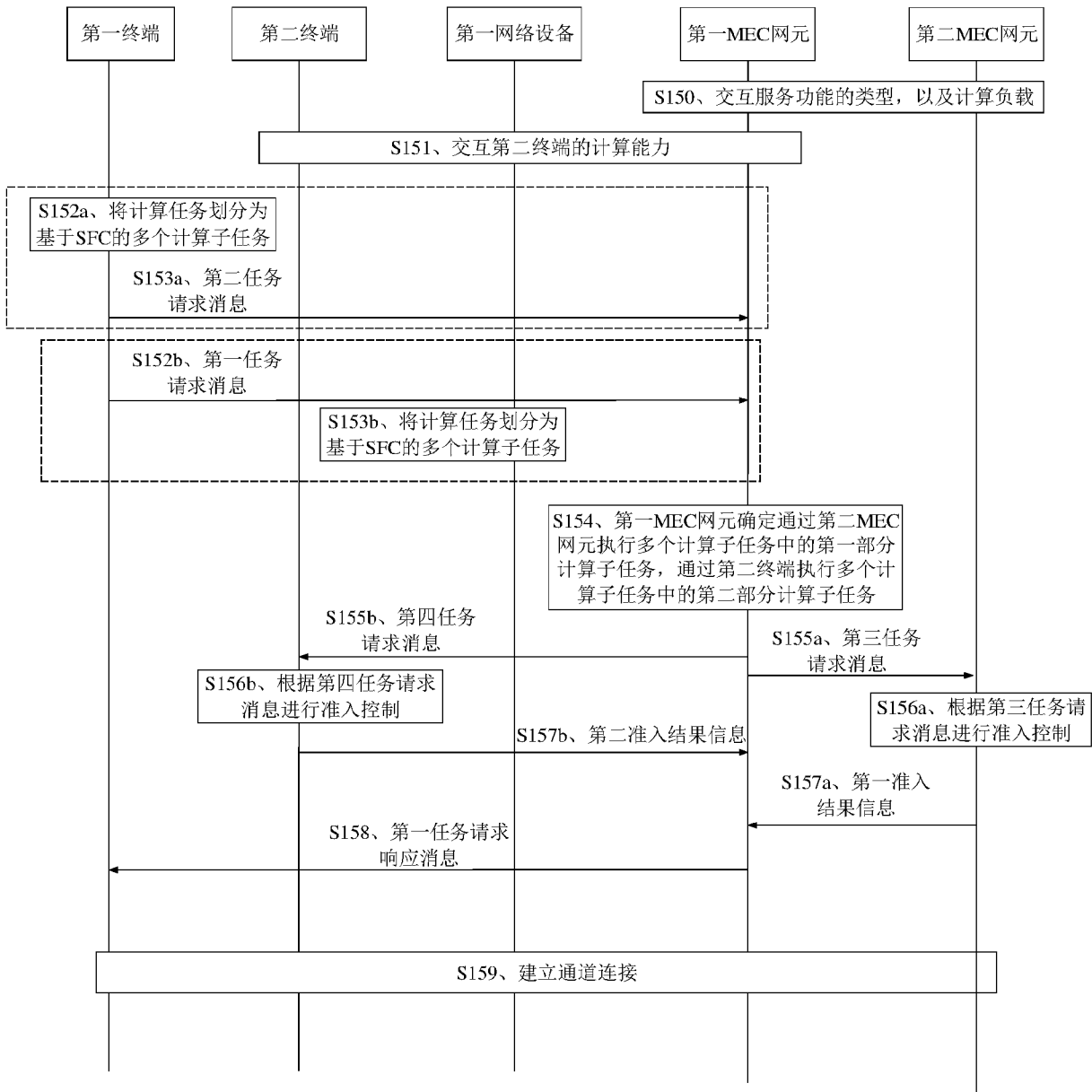


图 15

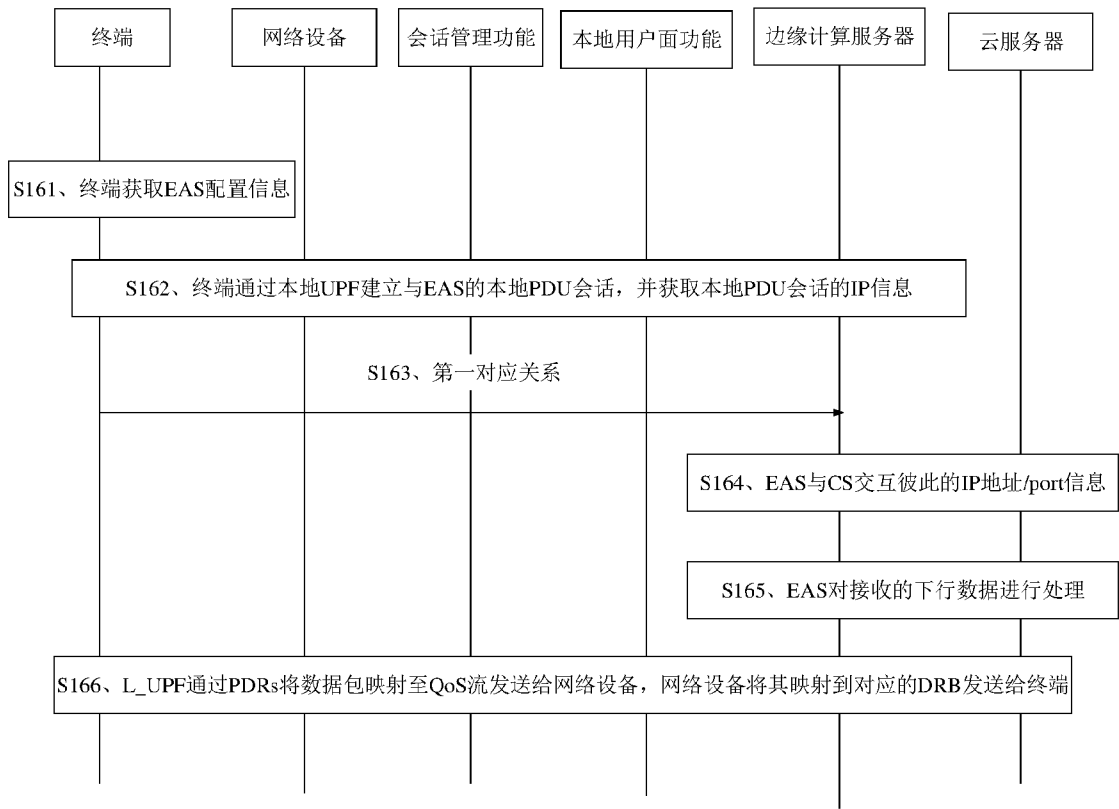


图 16

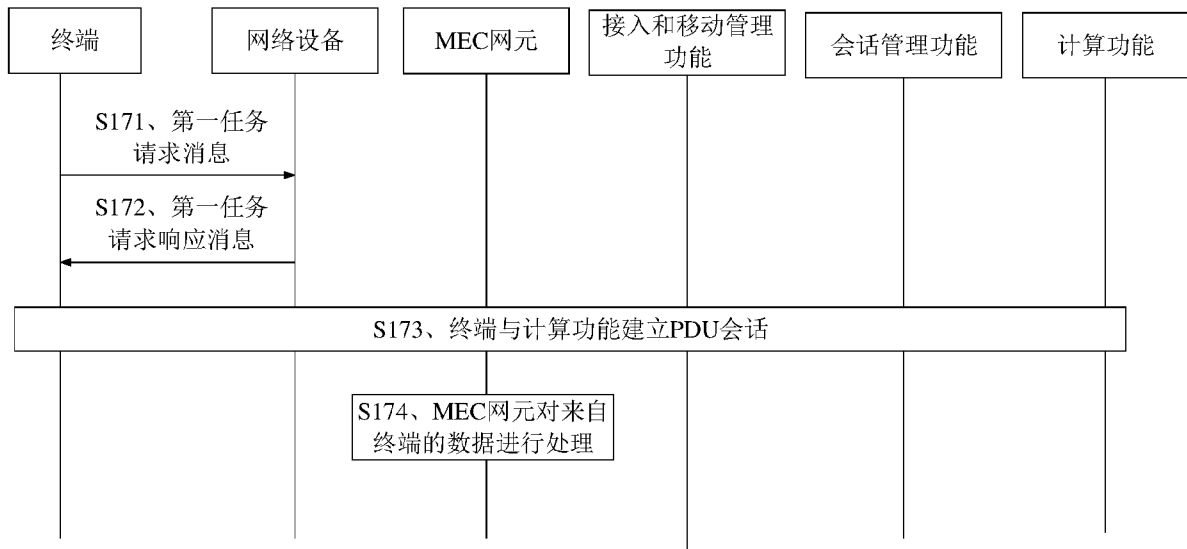


图 17

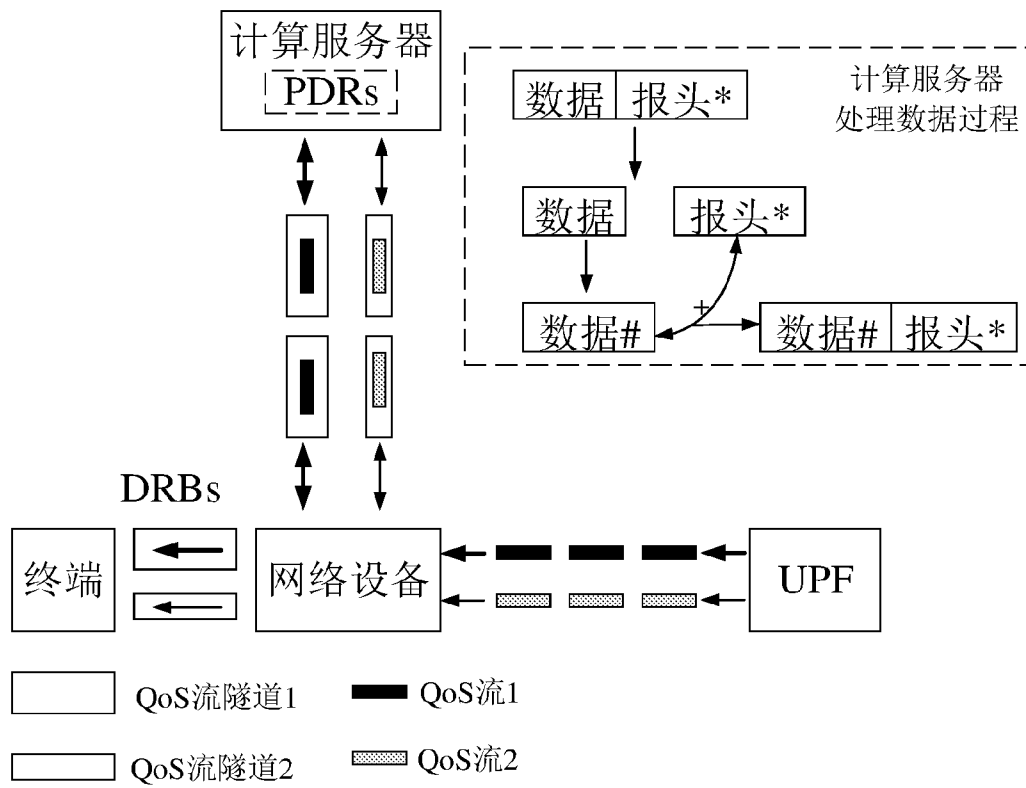


图 18

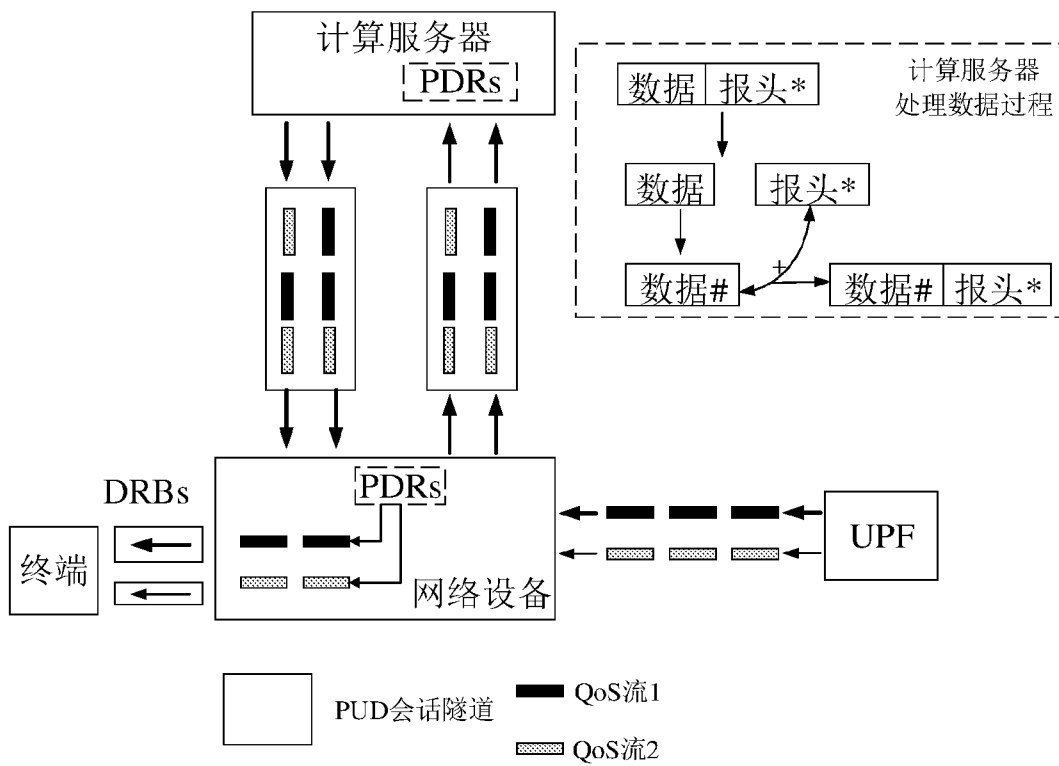


图 19

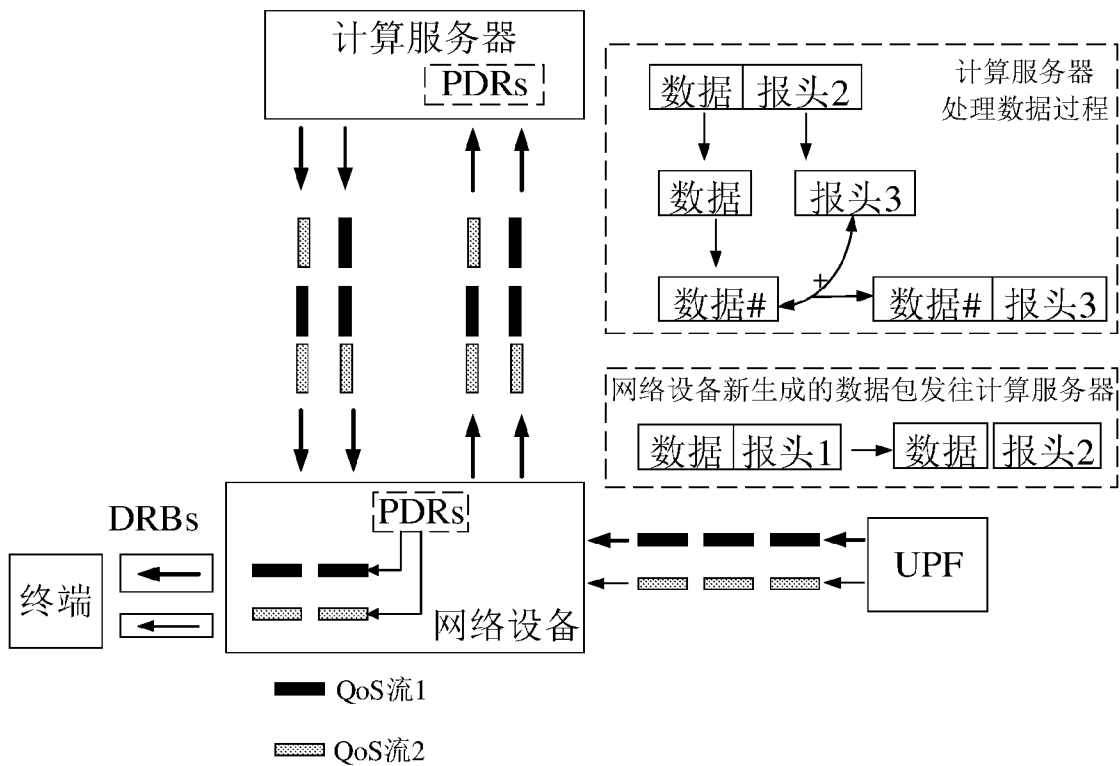


图 20

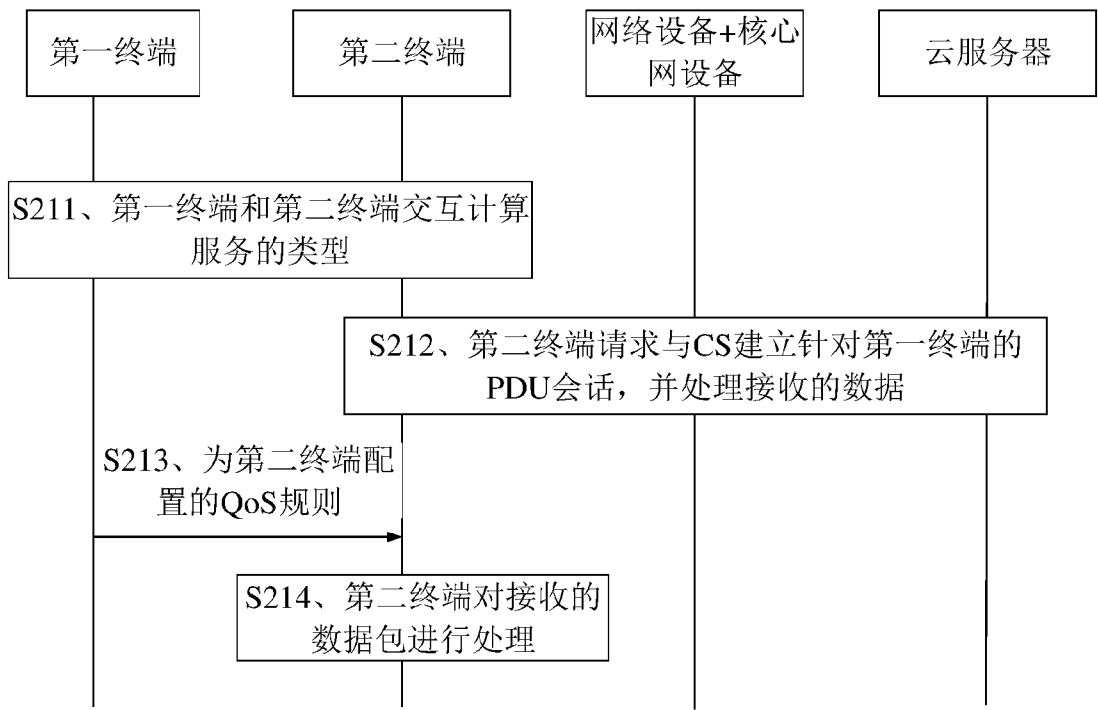


图 21

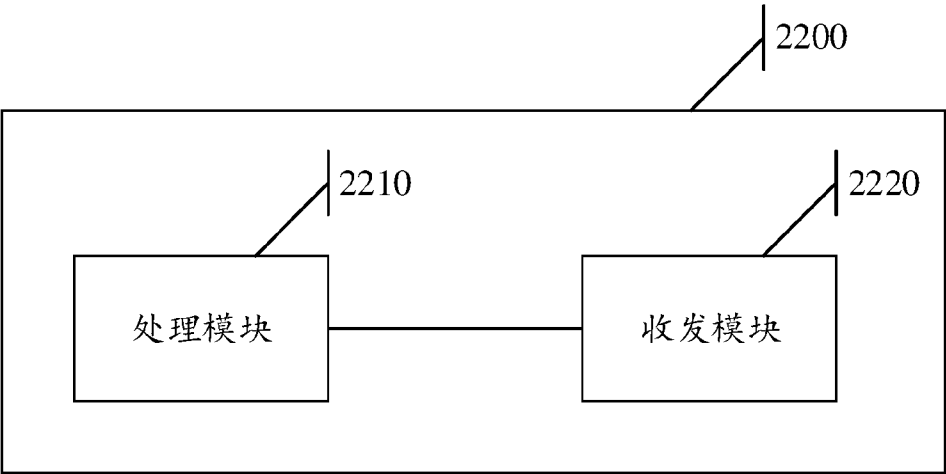


图 22

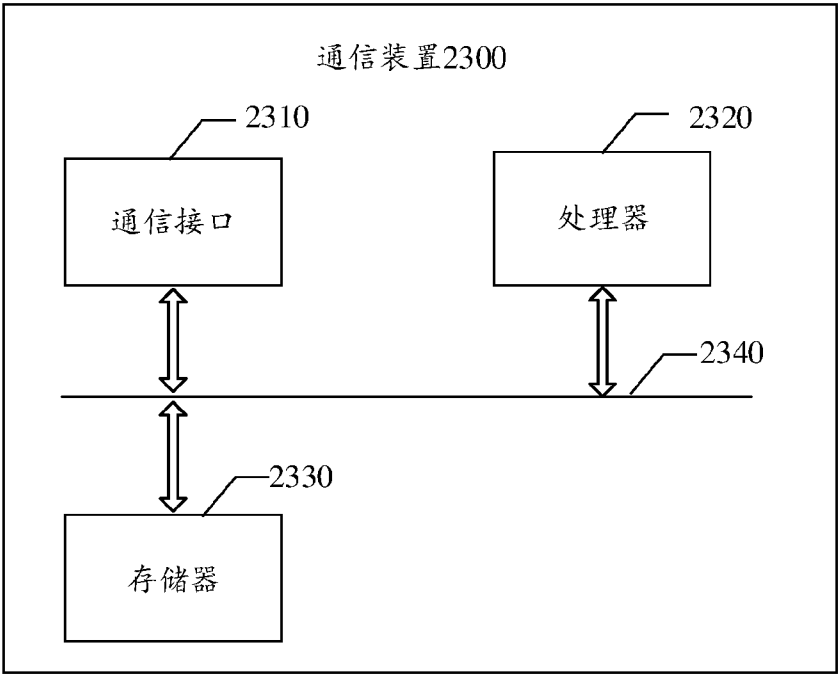


图 23

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/119747

A. CLASSIFICATION OF SUBJECT MATTER

H04W 88/18(2009.01)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)

H04W; H04L; H04B

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)

CNPAT, CNKI, WPI, EPODOC, 3GPP: 多接入边缘计算, 能力, 指示, 参数, 类型, 负载, 任务, 请求, 广播, 满足, 支持, MEC, capability, indication, parameter, type, load, task, request, broadcast, satisfy, support

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 108496341 A (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 04 September 2018 (2018-09-04) description paragraphs [0096]-[0239]	1-29
A	CN 109495929 A (HUAWEI TECHNOLOGIES CO., LTD.) 19 March 2019 (2019-03-19) entire document	1-29
A	CN 109451793 A (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 08 March 2019 (2019-03-08) entire document	1-29
A	US 2020076875 A1 (INTEL IP CORPORATION) 05 March 2020 (2020-03-05) entire document	1-29
A	CHINA MOBILE COM. CORPORATION. "Edge computing and service profile in 5GMSA3" 3GPP TSG-SA4 Meeting #104 S4-190806, 08 July 2019 (2019-07-08), entire document	1-29



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

"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

15 June 2021

Date of mailing of the international search report

24 June 2021

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/119747

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108496341	A	04 September 2018	WO	2019000414	A1	03 January 2019
CN	109495929	A	19 March 2019	US	2020213419	A1	02 July 2020
				AU	2018332696	A1	16 April 2020
				EP	3668174	A1	17 June 2020
				WO	2019052376	A1	21 March 2019
CN	109451793	A	08 March 2019	US	2020137642	A1	30 April 2020
				WO	2019000376	A1	03 January 2019
US	2020076875	A1	05 March 2020	US	2018183855	A1	28 June 2018

国际检索报告

国际申请号

PCT/CN2020/119747

A. 主题的分类

H04W 88/18 (2009.01) 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)

H04W; H04L; H04B

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))

CNPAT, CNKI, WPI, EPDOC, 3GPP: 多接入边缘计算, 能力, 指示, 参数, 类型, 负载, 任务, 请求, 广播, 满足, 支持, MEC, capability, indication, parameter, type, load, task, request, broadcast, satisfy, support

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	CN 108496341 A (北京小米移动软件有限公司) 2018年 9月 4日 (2018 - 09 - 04) 说明书第[0096]-[0239]段	1-29
A	CN 109495929 A (华为技术有限公司) 2019年 3月 19日 (2019 - 03 - 19) 全文	1-29
A	CN 109451793 A (北京小米移动软件有限公司) 2019年 3月 8日 (2019 - 03 - 08) 全文	1-29
A	US 2020076875 A1 (INTEL IP CORPORATION) 2020年 3月 5日 (2020 - 03 - 05) 全文	1-29
A	CHINA MOBILE COM. CORPORATION. "Edge computing and service profile in 5GMSA3" 3GPP TSG-SA4 Meeting #104 S4-190806, 2019年 7月 8日 (2019 - 07 - 08), 全文	1-29

☐ 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

“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

2021年 6月 15日

国际检索报告邮寄日期

2021年 6月 24日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

杨丹

电话号码 86-(10)-53961569

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/119747

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)
CN	108496341	A	2018年 9月 4日	WO	2019000414	A1	2019年 1月 3日
CN	109495929	A	2019年 3月 19日	US	2020213419	A1	2020年 7月 2日
				AU	2018332696	A1	2020年 4月 16日
				EP	3668174	A1	2020年 6月 17日
				WO	2019052376	A1	2019年 3月 21日
CN	109451793	A	2019年 3月 8日	US	2020137642	A1	2020年 4月 30日
				WO	2019000376	A1	2019年 1月 3日
US	2020076875	A1	2020年 3月 5日	US	2018183855	A1	2018年 6月 28日