



(12) 发明专利申请

(10) 申请公布号 CN 118202702 A

(43) 申请公布日 2024.06.14

(21) 申请号 202280073315.1

(74) 专利代理机构 中国贸促会专利商标事务所

(22) 申请日 2022.11.07

有限公司 11038

(66) 本国优先权数据

专利代理人 周磊

202111339237.9 2021.11.12 CN

(51) Int.Cl.

(85) PCT国际申请进入国家阶段日

H04W 28/24 (2006.01)

2024.04.30

(86) PCT国际申请的申请数据

PCT/CN2022/130328 2022.11.07

(87) PCT国际申请的公布数据

W02023/083138 ZH 2023.05.19

(71) 申请人 索尼集团公司

地址 日本

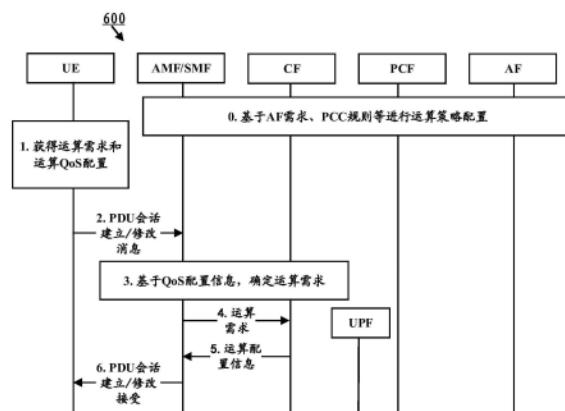
(72) 发明人 李岗涛 孙晨

(54) 发明名称

用于通信系统的电子设备、方法和存储介质

(57) 摘要

本公开内容涉及用于通信系统的电子设备、方法和存储介质。描述了配置和协调通信服务的服务质量的各种实施例。在实施例中，用于网络节点的电子设备包括处理电路，该处理电路被配置为接收来自终端设备的第一请求消息，其中第一请求消息对应于PDU会话建立请求或PDU会话修改请求消息，并且第一请求消息至少包括通信服务的运算QoS配置信息；以及向运算功能CF提供该通信服务的运算需求以实例化相应算力资源。



(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

2023 年 5 月 19 日 (19.05.2023)



WIPO | PCT



(10) 国际公布号

WO 2023/083138 A1

(51) 国际专利分类号:

H04W 28/24 (2009.01)

(72) 发明人; 及

(71) 申请人 (仅对MG): 李岚涛 (LI, Lantao) [CN/CN];
中国北京市朝阳区新源南路1号平安国际金融中心商业栋3层301室, Beijing 100027 (CN)。

(21) 国际申请号:

PCT/CN2022/130328

(22) 国际申请日: 2022 年 11 月 7 日 (07.11.2022)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

202111339237.9 2021年11月12日 (12.11.2021) CN

(71) 申请人: 索尼集团有限公司 (SONY GROUP CORPORATION) [JP/JP]; 日本东京都港区港南1-7-1, Tokyo 〒108-0075 (JP)。

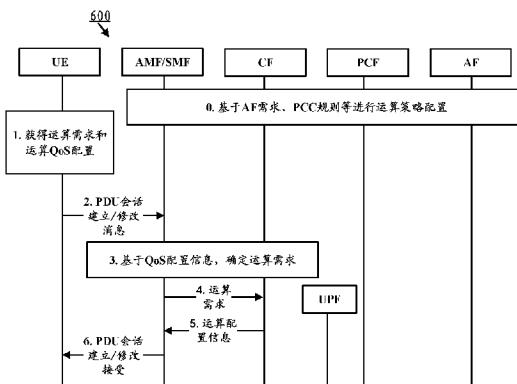
(72) 发明人: 孙晨 (SUN, Chen); 中国北京市朝阳区新源南路1号平安国际金融中心商业栋3层301室, Beijing 100027 (CN)。

(74) 代理人: 中国贸促会专利商标事务所有限公司 (CCPIT PATENT AND TRADEMARK LAW OFFICE); 中国北京市复兴门内大街158号远洋大厦F10层, Beijing 100031 (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,

(54) Title: ELECTRONIC DEVICE AND METHOD FOR COMMUNICATION SYSTEM AND STORAGE MEDIUM

(54) 发明名称: 用于通信系统的电子设备、方法和存储介质



- 0 Configure a computing policy on the basis of an AF requirement, a PCC rule, and the like
- 1 Obtain a computing requirement and a computing QoS configuration
- 2 PDU session establishment/modification message
- 3 Determine the computing requirement on the basis of QoS configuration information
- 4 Computing requirement
- 5 Computing configuration information
- 6 PDU session establishment/modification accepted

图 6

(57) **Abstract:** The present disclosure relates to an electronic device and method for a communication system and a storage medium. Described are embodiments in which the Quality of Service (QoS) of a communication service is configured and coordinated. In the embodiments, an electronic device for a network node comprises a processing circuit, the processing circuit being configured to receive a first request message from a terminal device, wherein the first request message corresponds to a PDU session establishment request or PDU session modification request message, and the first request message at least comprises computing QoS configuration information of a communication service; and provide a computing requirement of the communication service to a computing function (CF) to instantiate a corresponding computing resource.

(57) **摘要:** 本公开内容涉及用于通信系统的电子设备、方法和存储介质。描述了配置和协调通信服务的服务质量的各种实施例。在实施例中, 用于网络节点的电子设备包括处理电路, 该处理电路被配置为接收来自终端设备的第一请求消息, 其中第一请求消息对应于PDU会话建立请求或PDU会话修改请求消息, 并且第一请求消息至少包括通信服务的运算QoS配置信息; 以及向运算功能CF提供该通信服务的运算需求以实例化相应算力资源。



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, 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, 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, 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))。

用于通信系统的电子设备、方法和存储介质

技术领域

本公开一般地涉及通信设备和通信方法，包括用于配置和协调通信服务（例如通信
5 运算融合服务）的服务质量（QoS）的技术。

背景技术

在无线通信系统中，服务质量框架是保证通信服务的端到端服务性能的重要机制。
在第五代（5G）移动通信系统中，提出了基于 QoS 流（QoS flow）的 5G QoS 模型。5G
10 QoS 模型以 QoS 流为粒度，基于相应 QoS 配置信息（Profile）在通信传输性能方面为包
括通信运算融合服务在内的各种通信服务提供适当的服务质量。

在当前网络系统中存在多种多样的算力资源，以便支持在一定程度上涉及运算的各
种服务（例如通信运算融合服务）。一般而言，算力资源可以部署在边缘云、数据中心等
算力资源平台上，可以与无线网络设备或网络功能一起部署，甚至可以在终端设备算力
15 充足（例如存在空闲算力）的情况下由终端设备提供算力资源。多种多样的算力资源可
以满足多种服务场景需求。

针对各种通信服务，合理利用算力资源和通信传输资源以便提供适当且满意的端到
端服务质量是所期望的。

20 发明内容

本公开的第一方面涉及一种用于网络节点的电子设备。该电子设备包括处理电路，
该处理电路被配置为接收来自终端设备的第一请求消息，其中第一请求消息对应于 PDU
会话建立请求或 PDU 会话修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS
配置信息，该运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者；以及
25 向运算功能 CF 提供该通信服务的运算需求以实例化相应算力资源，其中该运算需求是在
该通信服务的运算 QoS 配置信息符合该终端设备的策略的情况下生成的。在一些实施例

中，该网络节点可以被配置为实现接入和移动性管理功能 AMF 和/或会话管理功能 SMF。

本公开的第二方面涉及一种用于网络节点的电子设备，该网络节点被配置为实现运算功能 CF。该电子设备包括处理电路，该处理电路被配置为接收来自会话管理功能 SMF 的通信服务的运算需求；以及向该 SMF 提供经实例化的算力资源的信息。

5 本公开的第三方面涉及一种用于终端设备的电子设备，包括处理电路。该处理电路被配置为向网络发送第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS 配置信息，该运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者。

10 本公开的第四方面涉及一种用于网络节点的电子设备，该网络节点被配置为实现应用功能 AF。该电子设备包括处理电路，该处理电路被配置为向策略控制功能 PCF 发送 AF 需求，其中该 AF 需求基于通信服务的订阅信息或业务场景中的至少一者，并且该 AF 需求包括该通信服务的运算 QoS 配置信息。

15 本公开的第五方面涉及一种用于网络节点的电子设备，该网络节点被配置为实现策略控制功能 PCF。该电子设备包括处理电路，该处理电路被配置为从应用功能 AF 接收 AF 需求，其中该 AF 需求包括通信服务的运算 QoS 配置信息；基于该 AF 需求，生成 PCC 规则；以及向会话管理功能 SMF 提供该 PCC 规则。

本公开的第六方面涉及各种用于通信的方法，该方法包括由例如上述各种电子设备所执行的操作或操作的任何组合。

20 本公开的第七方面涉及一种计算机可读存储介质，其上存储有可执行指令，该可执行指令在由一个或多个处理器执行时，实现根据本公开的各种实施例的方法的操作。

本公开的第八方面涉及一种计算机程序产品，该计算机程序产品包括指令，该指令在由计算机执行时使得实现根据本公开的各种实施例的方法。

25 提供上述概述是为了总结一些示例性的实施例，以提供对本文所描述的主题的各方面基本理解。因此，上述特征仅仅是示例并且不应该被解释为以任何方式缩小本文所描述的主题的范围或精神。本文所描述的主题的其它特征、方面和优点将从以下结合附图描述的具体实施方式而变得明晰。

附图说明

当结合附图考虑实施例的以下具体描述时，可以获得对本公开内容更好的理解。在各附图中使用了相同或相似的附图标记来表示相同或者相似的部件。各附图连同下面的具体描述一起包含在本说明书中并形成说明书的一部分，用来例示说明本公开的实施例和解释本公开的原理和优点。其中：

5 图 1 示出了根据本公开实施例的通信系统的示例框图。

图 2 示出了根据本公开实施例的通信系统的示例结构。

图 3 示出了根据本公开实施例的 5G NR (New Radio) QoS 的示例架构。

图 4 示出了可以实现根据本公开实施例的网络节点的示例电子设备。

10 图 5 示出了可以实现根据本公开实施例的终端设备的示例电子设备。

图 6 示出了根据本公开实施例的用于配置通信服务的 QoS 的示例处理流程。

图 7 示出了根据本公开实施例的用于配置和协调通信服务的运算和通信 QoS 参数的示例处理流程。

图 8 示出了根据本公开实施例的用于实例化算力资源的示例处理流程。

15 图 9 至图 12B 示出了根据本公开实施例的用于通信的示例方法。

图 13 示出了根据本公开实施例的可实现为终端设备或网络节点的计算机的示例框图。

图 14 示出了划分运算需求的示例。

图 15 至图 16 示出了根据本公开实施例的用于通信的示例信令流程。

20 虽然在本公开内容中所描述的实施例可能易于有各种修改和另选形式，但是其具体实施例在附图中作为例子示出并且在本文中被详细描述。但是，应该理解的是，附图以及对其的详细描述不是要将实施例限定到所公开的特定形式，而是相反，目的是要涵盖属于权利要求的精神和范围内的所有修改、等同和另选方案。

25 具体实施方式

以下描述根据本公开的设备和方法等各方面的代表性应用。这些例子的描述仅是为了增加上下文并帮助理解所描述的实施例。因此，对本领域技术人员而言清楚的是，以下所描述的实施例可以在没有具体细节当中的一些或全部的情况下被实施。在其它情况下，众所周知的过程步骤没有详细描述，以避免不必要的模糊所描述的实施例。其它应
5 用也是可能的，本公开的方案并不限制于这些示例。

一般而言，本文使用的所有术语将根据其在相关技术领域中的普通含义来解释，除非在使用上下文中清楚地给出不同的含义和/或暗示。除非明确有其它说明，对元素、装置、组件、单元和操作等的引用旨在开放地解释为元素、装置、组件、单元和操作中的至少一个实例。本文公开的任何方法的操作不必以所公开的精确顺序执行，除非将操作
10 明确或隐含地描述为在另一操作之后或之前。本文公开的任何实施例的任何特征可以应用于任何适当的其它实施例。同样地，任何实施例的任何优点都可以适用于任何其它实施例，反之亦然。实施例的其它目的、特征和优点将从以下描述中变得明晰。

通信系统示例

图 1 示出了根据本公开实施例的通信系统的示例框图。应理解，图 1 仅示出通信系
15 统的多种类型和可能布置中的一种；本公开的特征可根据需要在各种系统中的任一者中实现。

如图 1 所示，通信系统 100 包括基站 120A、120B 以及终端 110A、110B 至 110N。基站和终端可以被配置为通过 Uu 接口进行上下行链路通信。基站 120A、120B 可以被配置为与网络 130（例如，蜂窝服务提供方的核心网、诸如公共交换电话网（PSTN）的电信网
20 网络和/或互联网）进行通信。因此，基站 120A、120B 可以便于终端 110A 至 110N 之间和/或终端 110A 至 110N 与网络 130 之间的通信。进一步地，终端设备 110A 至 110N 可以通过 PC5 接口在有效通信范围内进行直连链路通信。

在图 1 中，基站 120A、120B 的覆盖区域可以被称为小区。根据一种或多种蜂窝通信技术进行操作的基站可以在广阔的地理区域上向终端 110A 至 110N 提供连续或近似连
25 续的通信信号覆盖。

如图 1 所示，通信系统 100 包括云 150 和移动边缘计算节点（Mobile Edge Computing，MEC）140。云 150 可以通过与网络 130 的连接为终端设备提供服务，诸如

IaaS、PaaS 和 SaaS。在云 150 和 MEC 140 中，可以部署算力资源，从而为满足通信服务（例如通信运算融合服务）的运算需求提供支持。

在本公开中，基站可以是 5G NR 基站，例如 gNB 和 ng-eNB。gNB 可以提供与终端设备终接的 NR 用户平面和控制平面协议；ng-eNB 是为了与 4G LTE 通信系统兼容而定义的 5 节点，其可以是 LTE 无线接入网的演进型节点 B (eNB) 的升级，提供与 UE 终接的演进通用陆地无线接入 (E-UTRA) 用户平面和控制平面协议。此外，基站的示例可以包括但不限于以下：GSM 系统中的基站收发信机 (BTS) 和基站控制器 (BSC) 中的至少一者；WCDMA 系统中的无线电网络控制器 (RNC) 和 Node B 中的至少一者；WLAN、WiMAX 系统中的接入点 (AP)；以及将要或正在开发的通信系统中对应的网络节点。本文中基站的部分 10 功能也可以实现为在 D2D、M2M 以及 V2X 通信场景下对通信具有控制功能的实体，或者实现为在认知无线电通信场景下起频谱协调作用的实体。

在本公开中，终端设备可以具有其通常含义的全部广度，例如终端设备可以为移动站 (Mobile Station, MS)、用户设备 (User Equipment, UE) 等。终端设备可以实现为诸如移动电话、手持式设备、媒体播放器、计算机、膝上型电脑、平板电脑、车载单元 15 或车辆或者几乎任何类型的无线设备。在一些情况下，终端设备可以使用多种无线通信技术进行通信。例如，终端设备可以被配置为使用 GSM、UMTS、CDMA2000、WiMAX、LTE、LTE-A、WLAN、NR、蓝牙等中的一者或者进行通信。以下将更多地结合 UE 描述本公开的实施例，然而应理解这些实施例对于任何类型的终端设备均适用。

图 2 示出了根据本公开实施例的通信系统的示例结构。作为示例，系统 200 被示出 20 为具有 3GPP 5G 核心网络 (5GC) 功能。网络功能可被实现为专用硬件上的分立网络元件，被实现为在专用硬件上运行的软件实例，或被实现为在适当平台（例如，专用硬件或云基础结构）上实例化的虚拟化功能。应理解，根据本公开的各种处理、功能和特性可以适用于除 5G 之外的其它（包括已经和将要研究的）核心网络。以下以系统 200 为例描述各种网络功能 (Network Function, NF)。

UPF 可充当 RAT 内和 RAT 间移动性的锚点、与 DN 互连的外部 PDU 会话点，以及支持 25 多宿主 PDU 会话的分支点。UPF 还可执行分组路由和转发，执行分组检查，执行策略规则的用户平面部分，合法拦截分组，执行流量使用情况报告，对用户平面执行 QoS 处理（例如，分组滤波、门控、UL/DL 速率执行），执行上行链路流量验证（例如，SDF 到 QoS

流映射)，上行链路和下行链路中的传输级别分组标记以及执行下行链路分组缓冲和下行链路数据通知触发。UPF 可包括用于支持将流量流路由到数据网络的上行链路分类器。DN 可表示各种网络运营商服务、互联网访问或第三方服务。UPF 可经由 SMF 和 UPF 之间的 N4 参考点与 SMF 进行交互。

5 AUSF 可存储用于 UE 的认证的数据并处理与认证相关的功能。AUSF 可有利于针对各种访问类型的公共认证框架。AUSF 可经由 AMF 和 AUSF 之间的 N12 参考点与 AMF 通信；并且可经由 UDM 和 AUSF 之间的 N13 参考点与 UDM 通信。另外，AUSF 可呈现出基于 Nausf 服务的接口。

10 AMF 可负责注册管理（例如，负责注册 UE 等）、连接管理、可达性管理、移动性管理和对 AMF 相关事件的合法拦截，并且访问认证和授权。AMF 可以是 AMF 和 SMF 之间的 N11 参考点的终止点。AMF 可以是 RAN CP 接口的终止点，其可包括或为(R)AN 和 AMF 之间的 N2 参考点；并且 AMF 可以是 NAS (N1)信令的终止点，并且执行 NAS 加密和完整性保护。

15 AMF 还可通过 N3IWF 接口支持与 UE 的 NAS 信令。N3IWF 可用于提供对不可信实体的访问。N3IWF 可以是控制平面的(R)AN 和 AMF 之间的 N2 接口的终止点，并且可以是用户平面的(R)AN 和 UPF 之间的 N3 参考点的终止点。因此，AMF 可处理来自 SMF 和 AMF 的用于 PDU 会话和 QoS 的 N2 信令，封装/解封分组以用于 IPSec 和 N3 隧道，将 N3 用户平面分组标记在上行链路中，并且执行对应于 N3 分组标记的 QoS，这考虑到与通过 N2 接收到的此类标记相关联的 QoS 需求。N3IWF 还可经由 UE 和 AMF 之间的 N1 参考点在 UE 和 AMF 20 之间中继上行链路和下行链路控制平面 NAS 信令，并且在 UE 和 UPF 之间中继上行链路和下行链路用户平面分组。N3IWF 还提供用于利用 UE 建立 IPsec 隧道的机制。AMF 可呈现出基于 Namf 服务的接口。

25 SMF 可负责 SM（例如，会话建立、修改和释放，包括 UPF 和 AN 节点之间的隧道维护）；UE IP 地址分配和管理（包括任选授权）；UP 功能的选择和控制；配置 UPF 的交通转向以将流量路由至正确的目的地；终止朝向策略控制功能的接口；策略执行和 QoS 的控制部分；合法拦截（对于 SM 事件和与 LI 系统的接口）；终止 NAS 消息的 SM 部分；下行链路数据通知；发起经由 AMF 通过 N2 发送到 AN 的 AN 特定 SM 信息；以及确定会话的 SSC 模式。SM 可指 PDU 会话的管理，并且 PDU 会话或“会话”可指提供或实现由数据网

络名称（DNN）识别的 UE 和数据网络（DN）之间的 PDU 交换的 PDU 连接性服务。PDU 会话可以使用在 UE 和 SMF 之间通过 N1 参考点交换的 NAS SM 信令在 UE 请求时建立，在 UE 和 5GC 请求时修改，并且在 UE 和 5GC 请求时释放。在从应用服务器请求时，5GC 可触发 UE 中的特定应用程序。响应于接收到触发消息，UE 可将触发消息（或触发消息的相关部分/信息）传递到 UE 中的一个或多个识别的应用程序。UE 中的识别的应用程序可建立到特定 DNN 的 PDU 会话。SMF 可检查 UE 请求是否符合与 UE 相关联的用户订阅信息。就这一点而言，SMF 可检索和/或请求以从 UDM 接收关于 SMF 级别订阅数据的更新通知。

SMF 可包括以下漫游功能：处理本地执行以应用 QoS SLA (VPLMN)；计费数据采集和计费接口 (VPLMN)；合法拦截（对于 SM 事件和与 LI 系统的接口，在 VPLMN 中）；以及支持与外部 DN 的交互，以传输用于通过外部 DN 进行 PDU 会话授权/认证的信令。在漫游场景中，两个 SMF 之间的 N16 参考点可包括在系统 700 中，该系统可位于受访网络中的另一个 SMF 与家庭网络中的 SMF 之间。另外，SMF 可呈现出基于 Nsmf 服务的接口。

NEF 可提供用于安全地暴露由 3GPP 网络功能为第三方、内部暴露/再暴露、应用功能（例如，AF）、边缘计算或雾计算系统等提供的服务和能力的装置。在此类实施方案中，NEF 可对 AF 进行认证、授权和/或限制。NEF 还可转换与 AF 交换的信息以及与内部网络功能交换的信息。例如，NEF 可在 AF 服务标识符和内部 5GC 信息之间转换。NEF 还可基于其它网络功能的暴露能力从其它网络功能（NF）接收信息。该信息可作为结构化数据存储在 NEF 处，或使用标准化接口存储在数据存储 NF 处。然后，存储的信息可由 NEF 重新暴露于其它 NF 和 AF，并且/或者用于其它目的诸如分析。另外，NEF 可呈现出基于 Nnef 服务的接口。

NRF 可支持服务发现功能，从 NF 实例接收 NF 发现请求，并且向 NF 实例提供发现的 NF 实例的信息。NRF 还维护可用的 NF 实例及其支持的服务的信息。另外，NRF 可呈现出基于 Nnrf 服务的接口。

PCF 可提供用于控制平面功能以执行它们的策略规则，并且还可支持用于管理网络行为的统一策略框架。PCF 还可实现 FE 以访问与 UDM 的 UDR 中的策略决策相关的订阅信息。PCF 可经由 PCF 和 AMF 之间的 N15 参考点与 AMF 通信，这可包括受访网络中的 PCF 和在漫游场景情况下的 AMF。PCF 可经由 PCF 和 AF 之间的 N5 参考点与 AF 通信；并且经由 PCF 和 SMF 之间的 N7 参考点与 SMF 通信。系统 200 和/或 CN 还可包括（家庭网络中的）

PCF 和受访网络中的 PCF 之间的 N24 参考点。另外，PCF 可呈现出基于 Npcf 服务的接口。

UDM 可处理与订阅相关的信息以支持网络实体对通信会话的处理，并且可存储 UE 的订阅数据。例如，可经由 UDM 和 AMF 之间的 N8 参考点在 UDM 和 AMF 之间传送订阅数据。

UDM 可包括两部分：应用程序 FE 和 UDR。UDR 可存储 UDM 和 PCF 的订阅数据和策略数据，

5 和/或 NEF 的用于暴露的结构化数据以及应用数据（包括用于应用检测的 PFD、多个 UE 的应用请求信息）。基于 Nudr 服务的接口可由 UDR 呈现出以允许 UDM、PCF 和 NEF 访问存储的数据的特定集，以及读取、更新（例如，添加、修改）、删除和订阅 UDR 中的相关数据更改的通知。UDM 可包括 UDM-FE，其负责处理凭据、位置管理、订阅管理等。在不同的事务中，若干不同的前端可为同一用户服务。UDM-FE 访问存储在 UDR 中的订阅信息，并且执行认证凭证处理、用户识别处理、访问授权、注册/移动性管理和订阅管理。UDR 可经由 UDM 和 SMF 之间的 N10 参考点与 SMF 进行交互。UDM 还可支持 SMS 管理，其中 SMS-FE 实现先前所讨论的类似应用逻辑。另外，UDM 可呈现出基于 Nudm 服务的接口。

AF 可提供应用程序对流量路由的影响，提供对 NCE 的访问，并且与策略框架进行交互以进行策略控制。NCE 可以是允许 5GC 和 AF 经由 NEF 彼此提供信息的机制，其可用于边缘计算具体实施。在此类具体实施中，网络运营商和第三方服务可被托管在附件的 UE 接入点附近，以通过减小的端到端延迟和传输网络上的负载来实现有效的服务递送。对于边缘计算具体实施，5GC 可选择 UE 附近的 UPF 并且经由 N6 接口执行从 UPF 到 DN 的流量转向。这可基于 UE 订阅数据、UE 位置和 AF 所提供的信息。这样，AF 可影响 UPF（重新）选择和流量路由。基于运营商部署，当 AF 被认为是可信实体时，网络运营商可允许

20 AF 与相关 NF 直接进行交互。另外，AF 可呈现出基于 Naf 服务的接口。

NSSF 可选择为 UE 服务的一组网络切片实例。如果需要，NSSF 还可确定允许的 NSSAI 和到订阅的 S-NSSAI 的映射。NSSF 还可基于合适的配置并且可能通过查询 NRF 来确定用

于为 UE 服务的 AMF 集，或候选 AMF 的列表。UE 的一组网络切片实例的选择可由 AMF 触发，其中 UE 通过与 NSSF 进行交互而注册，这可导致 AMF 发生改变。NSSF 可经由 AMF 和

25 NSSF 之间的 N22 参考点与 AMF 进行交互；并且可经由 N31 参考点与受访网络中的另一 NSSF 通信。另外，NSSF 可呈现出基于 Nnssf 服务的接口。

根据本公开的实施例，系统 200 可以具有运算功能 CF（Computing/Computation Function）。运算功能可包括运算资源管理、运算资源部署/分配、预算接口适配等子功

能。根据不同上下文，运算功能也可以称为运算感知功能、运算管理功能等。CF 可以用于根据来自终端设备或 AF 的运算需求来实例化算力资源，或者基于算力资源的部署情况与其它功能实体协调算力资源。CF 可呈现出基于 Ncf 服务的接口。

图 3 示出了根据本公开实施例的 5G NR QoS 的示例架构。5G QoS 模型基于 QoS 流，
5 并且支持需要保证流比特率的 QoS 流（GBR QoS 流）和不需要保证流比特率的 QoS 流（非
GBR QoS 流）。因此，在 NAS 级别，QoS 流是 PDU 会话中 QoS 区分的最小粒度。在 PDU 会
话中，通过在 NG-U 上的封装报头中所携带的 QoS 流 ID（QFI）识别 QoS 流。

如图 3 所示，在 NG-RAN 的 QoS 架构（例如用于连接到 5GC 的 NR 以及连接到 5GC 的
E-UTRA）中，对于每个 UE，5GC 可以建立一个或多个 PDU 会话。与 PDU 会话一起，NG-RAN
10 可以建立至少一个数据无线电承载（DRB），并且随后可以为该 PDU 会话的 QoS 流配置附
加的 DRB（何时配置可取决于 NG-RAN）。NG-RAN 可以将属于不同 PDU 会话的分组映射到不
同的 DRB。UE 以及 5GC 中的 NAS 级别数据包过滤器可以将 UL 和 DL 数据包与 QoS 流相关
联，UE 以及 NG-RAN 中的 AS 级别映射规则将 UL 和 DL QoS 流与 DRB 相关联。

15 NG-RAN 和 5GC 通过将数据包映射到适当的 QoS 流和 DRB 来确保服务质量（例如可靠
性和目标延迟）。因此，存在 IP 流与 QoS 流（NAS）以及 QoS 流与 DRB（AS）的两步映射。

在 NAS 级别，QoS 流可以由 QoS 配置信息（Profile）和 QoS 规则（Rule）表征。可
以由 5GC 将 QoS 配置提供给 NG-RAN，可以由 5GC 将 QoS 规则提供给 UE。

NG-RAN 可以使用 QoS 配置信息确定在无线电接口上的处理，QoS 规则可以向 UE 指
示 UL 用户平面业务与 QoS 流之间的映射。根据其配置，QoS 流可以是 GBR 或非 GBR 的。
20 QoS 流的 QoS 配置信息可以包含诸如以下的 QoS 参数（参见例如 3GPP TS 23.501）。

对于每个 QoS 流：

- 5G QoS 标识符（5QI）；
- 分配和保留优先权（ARP）。

仅在 GBR QoS 流的情况下：

- 对于 UL 和 DL，保证流比特率（GFBR）；
- 对于 UL 和 DL，最大流比特率（MFBR）；

- 对于 UL 和 DL，最大丢包率；
- 延迟关键资源类型；
- 通知控制。

仅在非 GBR QoS 流的情况下：

- 5
 - 反射 QoS 属性 (RQA): RQA 在包括时表示在 QoS 流上承载的一些（不一定全部）业务在 NAS 受到反射服务质量 (RQoS);
 - 附加 QoS 流信息。

5QI 与 QoS 特性相关联，可以为设置每个 QoS 流的节点特定参数提供指引。标准化或预先配置的 5G QoS 特性是从 5QI 值导出的，并且不通过显示信令进行通知。相应地，

- 10 可以将需要信令进行通知的 QoS 特性作为 QoS 配置信息的一部分。QoS 流的 QoS 特性可以包含诸如以下内容（参见例如 3GPP TS 23.501）。

- 优先级；
- 数据包延迟预算；
- 误报率；
- 平均窗口；
- 最大数据突发量。

- 在 AS 级别，DRB 可以定义无线电接口 (Uu) 上的数据包处理。DRB 可以通过相同的数据包转发处理为数据包提供服务。NG-RAN 可以基于 QFI 和相关的 QoS 配置信息（例如包括 QoS 参数和 QoS 特性）进行 QoS 流到 DRB 的映射。可以为需要不同数据包转发处理的 QoS 流建立单独的 DRB，或者可以在同一 DRB 中复用属于同一 PDU 会话的多个 QoS 流。

在上行链路中，可以通过映射规则来控制 QoS 流到 DRB 的映射。映射规则可以通过两种不同方式通知，即反射映射以及显式配置（例如通过 RRC 信令）。无论映射规则是通过哪种方式通知的，UE 始终应用最近更新的映射规则。

- 以上 5G QoS 配置信息以及相应的参数、特性等均与传输方面有关，从而提供适当
25 的端到端通信或传输性能。根据本公开的实施例，与上述通信 QoS 相对地定义了运算 QoS

配置信息以及相应的参数、特性，旨在提供一种机制来为通信服务提供适当的运算性能。这对于诸如 AR/VR、V2X 等的通信运算融合服务更加有利。以下示出运算 QoS 参数的一些示例。

- 算力需求 (Computation Power Requirements)，包括例如算力平台的运算能力普适值/平均值（例如单位为 flops/tflops）、算力平台的运算能力最高值（例如单位为 flops/tflops）、哈希率 (H/s)、编解码能力（例如帧率、分辨率、编解码格式（如 H.264/H.265）等；
- 运算优先级 (Computation Priority/Precedence)，指示该运算服务的处理优先级；
- 运算特性 (Computation Characteristics)，包括例如运算所需的运算框架（例如 CUDA）、硬件倾向或需求（例如 CPU、GPU、NPU、FPGA 等）、运算平台的系统配置（例如 win、linux 等）、运算平台的网络通信能力（例如带宽）、运算粒度（例如一次运算需要传输三个数据包的数据量才能进行）；
- 服务标识符 (Service Identifier)，标记服务类型或指示应用场景；
- 运算部署方式 (Computation Deployment)，包括例如容器部署、虚拟机部署、通用服务器（例如 AF 开放 API 支撑运算）等；
- 运算缓存需求 (Computation Buffer Requirements)，包括例如缓冲器大小、缓冲读写速度等、分布式缓存系统类型（例如 redis、MemCache、SSDB）或者硬件规格 (SRAM) 等；
- 算力切片的切片编号 (Computation S-NSSAI)，用于指示核心网络为 UE 选择的对应切片资源。

在一些实施例中，UE 在发送 PDU 会话建立或修改请求消息时，可以以 QoS 流为粒度携带每个 QoS 流的标准化运算 QoS 参数或特性，以发送至核心网络功能实体。根据一些实施例，可以基于通信服务（例如通信运算融合服务）的运算 QoS 参数实例化算力资源，从而为通信服务的运算性能提供保障。可以基于通信服务的通信 QoS 参数配置通信传输资源，从而为通信服务的通信或传输性能提供保障。根据一些实施例，可以在通信服务的运算 QoS 参数和通信 QoS 参数之间协调，从而在当前可用的算力和通信资源下为通信服务实现尽量优异的整体服务质量。

示例电子设备

图 4 示出了可以实现根据本公开实施例的网络节点的示例电子设备。电子设备 400 可以包括各种单元以实现根据本公开的用于控制和协调通信服务 QoS 的各实施例。在图 4 的示例中，电子设备 400 包括控制单元 402 和收发单元 404。控制单元 402 可以被配置为 5 控制或执行与通信服务 QoS 的配置和协调相关的操作，收发单元 404 可以被配置为控制或执行与信令或消息收发相关的操作。以下结合网络节点或网络功能描述的各种操作可以由电子设备 400 的单元 402 至 404 或者其它可能的单元实现。

在一些实施例中，电子设备 400 可以以芯片级来实现，或者也可以通过包括其它外部部件（例如有线或无线链路）而以设备级来实现。电子设备 400 可以作为整机而工作 10 为通信设备，例如诸如 AMF、SMF、CF 的网络节点。

图 5 示出了可以实现根据本公开实施例的终端设备的示例电子设备。电子设备 500 可以包括各种单元以实现根据本公开的用于控制和协调通信服务 QoS 的各实施例。在图 5 的示例中，电子设备 500 包括控制单元 502 和收发单元 504。控制单元 502 可以被配置为 15 控制或执行与通信服务 QoS 的配置和协调相关的操作，收发单元 504 可以被配置为控制或执行与信令或消息收发相关的操作。以下结合终端设备描述的各种操作可以由电子设备 500 的单元 502 至 504 或者其它可能的单元实现。

在一些实施例中，电子设备 500 可以以芯片级来实现，或者也可以通过包括其它外部部件（例如无线电链路、天线等）而以设备级来实现。电子设备 500 可以作为整机而工作为通信设备，诸如 UE、车载单元或配置有通信能力的车辆。

20 应该注意的是，上述各个单元仅是根据其所实现的具体功能划分的逻辑模块，而不是用于限制具体的实现方式，例如可以以软件、硬件或者软硬件结合的方式来实现。在实际实现时，上述各个单元可被实现为独立的物理实体，或者也可由单个实体（例如，处理器（CPU 或 DSP 等）、集成电路等）来实现。其中，处理电路可以指在计算系统中执行功能的数字电路系统、模拟电路系统或混合信号（模拟和数字的组合）电路系统的各种实现。处理电路可以包括例如诸如集成电路（IC）、专用集成电路（ASIC）这样的电路、单独处理器核心的部分或电路、整个处理器核心、单独的处理器、诸如现场可编程门阵列（FPGA）的可编程硬件设备、和/或包括多个处理器的系统。

基于运算 QoS 参数的运算性能保障

图 6 示出了根据本公开实施例的用于配置通信服务的 QoS 的示例处理流程。如图 6 所示，在操作 0 处，在核心网络的各 NF 之间进行运算策略配置。运算策略配置可以基于 AF 需求和 PCC 规则等。具体地，在一些实施例中，AF 可以向 PCF 发送 AF 需求。AF 需求可以是基于通信服务或者用户订阅信息或业务场景中的至少一者生成的。AF 需求可以至少包括通信服务的运算 QoS 配置信息。附加地，AF 需求还包括通信服务的通信 QoS 配置信息。

相应地，PCF 可以接收来自 AF 的 AF 需求。PCF 进一步基于该 AF 需求生成 PCC 规则，并向 SMF 提供该 PCC 规则。PCC 规则可以至少包括通信服务的运算 QoS 配置信息。附加地，PCC 规则还包括通信服务的通信 QoS 配置信息。

相应地，AMF/SMF 可以接收来自 PCF 的 PCC 规则。该 PCC 规则可以包括通信服务的运算 QoS 配置信息或者附加的通信 QoS 配置信息。进一步，AMF/SMF 可以基于 PCC 规则生成各 UE 的策略。AMF/SMF 可以将该策略应用于由相应 UE 发起的 PDU 会话建立流程和 PDU 会话修改流程以及应用于相应的 QoS 流。

如图 6 所示，在操作 2 处，UE 可以向核心网络（例如 AMF/SMF）发送第一请求消息。例如，第一请求消息可以对应于 PDU 会话建立请求或 PDU 会话修改请求消息。一般地，第一请求消息可以由 AMF 接收，AMF 转而将会话相关的信息转发至 SMF。第一请求消息可以至少包括通信服务的运算 QoS 配置信息，运算 QoS 配置信息例如包括运算 QoS 参数或运算 QoS 特性中的至少一者。附加地，第一请求消息可以包括通信服务的通信 QoS 配置信息，通信 QoS 配置信息例如包括通信 QoS 参数或通信 QoS 特性中的至少一者。一般而言，运算 QoS 配置信息和通信 QoS 配置信息以 QoS 流为粒度，携带每个 QoS 流对应的标准化 QoS 参数或特性。例如，运算 QoS 参数或特性可以包括以下中的至少一者，即算力需求、运算优先级、运算特性、服务标识符、运算部署方式、运算缓存需求或者算力切片的切片编号。

应理解，通信服务的运算 QoS 配置信息和通信 QoS 配置信息可以是 UE 基于来自高层（例如应用层、服务层）的服务特性确定的。例如，在操作 1 处，UE 可以通过在 NAS 级别解析数据包（例如特定应用的数据包首包）或者 AT（attention）命令而获得通信服务的运算需求。进一步，UE 可以基于运算策略，将运算需求映射为运算 QoS 配置信息

(例如包括标准化的运算 QoS 参数和/或特性)。类似地，UE 可以获得通信服务的通信需求，并基于传输策略将通信需求映射为通信 QoS 配置信息。

如图 6 所示，在操作 4 处，AMF/SMF 可以向 CF 提供通信服务的运算需求以实例化相应算力资源。运算需求可以是在通信服务的运算 QoS 配置信息符合终端设备的策略的情况下生成的。例如，在操作 3 处，一经获得由 AMF 转发的通信服务的运算 QoS 配置信息，
5 SMF 可以验证其是否符合终端设备的策略。在运算 QoS 配置信息符合终端设备的策略的情况下，SMF 可以基于运算 QoS 配置信息确定通信服务的运算需求。

在操作 3 处，附加地，SMF 可以验证通信服务的通信 QoS 配置信息是否符合终端设备的策略。在通信 QoS 配置信息符合终端设备的策略的情况下，SMF 可以基于通信 QoS 配
10 置信息确定通信服务的通信需求。进一步，SMF 可以在运算和通信 QoS 参数或特性之间进行协调，从而以互补或相互匹配的方式确定并满足通信服务的运算需求和通信需求，如后文详细描述的。更进一步，AMF/SMF 可以与 CF 以及其它网络功能实体协商或协调运算
15 和通信 QoS 参数或特性，确实需要建立/修改的 QoS 流中的初始运算和通信 QoS 参数（例如 QoS 规则及其所包含的数据包过滤器等）。

如图 6 所示，在操作 5 处，CF 可以向 SMF 提供经实例化的算力资源的信息。在一些实施例中，一经接收到来自 SMF 的通信服务的运算需求，CF 可以基于运算需求和可用算力资源情况确定要实例化的算力资源。例如，算力资源可以来自以下中的至少一者，即核心网算力资源实体、基站或基站侧模块、UPF、UE 或第三方算力资源平台。根据一些实施例，经实例化的算力资源的信息可以包括算力资源的配置信息，或者还包括访问算力
20 资源的接口信息。

如图 6 所示，在操作 6 处，一经接收到来自 CF 的经实例化的算力资源的信息，AMF/SMF 可以向 UE 提供第一响应消息。第一响应消息可以对应于 PDU 会话建立接受或 PDU 会话修改接受消息。之后，可以在 UE 与 UPF 之间建立数据面连接。

通过处理流程 600，针对所建立/修改的 QoS 流，可以基于运算 QoS 参数或特性分配并实例化算力资源，从而为通信服务提供适当的运算性能。针对所建立/修改的 QoS 流，
25 基于通信 QoS 参数或特性建立数据面，从而为通信服务提供适当的传输性能。

运算和通信 QoS 参数的协调

在一些实施例中，基于经协调的运算 QoS 参数或特性，可以确定通信服务的运算需求。在一些实施例中，基于经协调的通信 QoS 参数或特性，确定通信服务的通信需求。进一步，AMF/SMF 可以基于运算 QoS 配置信息和通信 QoS 配置信息两者，协调用于运算和通信的 QoS 参数或特性。通过在运算和通信 QoS 参数或特性之间进行协调，可以使得通信服务取得相互匹配的运算和通信性能，或者可以使得通信服务的包括运算和通信两方面的整体性能更优。以下描述协调用于运算和通信的 QoS 参数或特性的一些示例。

在一些实施例中，可以以互补的方式协调用于运算和通信的 QoS 参数或特性。例如，可以相对于通信服务的运算时延确定传输时延，或者相对于通信服务的传输时延确定运算时延，使得运算时延与传输时延之和低于或等于阈值水平。对于 QoS 流所承载的通信服务，用户所感知的整体时延可以由运算时延和传输时延两部分组成。容易理解，只要将运算时延与传输时延之和保持为低于阈值水平，就时延而言便可以提供符合预期的体验质量 (QoE)。这使得通信系统可以在保证服务时延性能方面具有一定灵活性。

例如，在传输资源所能提供的时延性能相对不理想的情况下（即传输时延较大），可以通过配置算力资源使运算时延较小从而与较大的传输时延互补。作为示例，配置算力资源可以包括增加算力、提高运算任务的优先级等）。相反的情况同样成立。又例如，在算力资源所能提供的时延性能相对不理想的情况下（即运算时延较大），可以通过配置传输资源使传输时延较小从而与较大的运算时延互补。相反的情况同样成立。

应理解，AMF/SMF 可以从 RAN、UPF 和 CF 等收集网络状态、无线电资源情况以及 QoS 性能数据，从而为确定传输时延参数和运算时延参数提供基础。例如，传输时延参数可以包括数据包时延预算 (PDB)。数据包时延预算可以对应端到端传输时延，或者可以由以下时延之一构成：上行空口时延、下行空口时延、核心网传输时延、外网/DN 传输时延。运算时延参数可以理解为一个被处理的数据实体从被完整接收到处理完成形成输出结果所需的时间。具体地，运算时延参数可以包括在高速缓存、PIPE、FIFO 等存储器或计算机优先级系统上的等待时间，以及实际运算处理过程中所花费的时间。

在一些实施例中，可以以相互匹配或一致的方式协调用于运算和通信的 QoS 参数或特性。例如，基于所述通信服务的通信和运算处理的相对优先程度，将运算处理能力和通信带宽中的一者确定为与这两者中的另一者相匹配。RAN 在空口上所能够提供的带宽（例如对应每个 QoS 流的最高数据速率 UL/DL GFBR 和 MFBR）需要与算力资源为该 QoS 流

所提供的单位时间数据处理量（即数据处理能力）相匹配；反之亦然。例如，数据处理能力可以包括保障的数据处理能力（Guaranteed Data Processing Capability）和最大数据处理能力（Maximum Data Processing Capability），分别与 QoS 流数据传输速率的 GFBR 和 MFBR 相对应或匹配。

5 根据一些实施例，在通信传输的优先级高于运算优先级的情况下，可以分配通信和算力资源使得通信的单位时间数据吞吐量大于等于运算的单位时间数据吞吐量，以保证可以及时对所有经运算处理的数据进行传输。例如，对于诸如 AR/VR、游戏画面渲染、车联网融合感知、车辆路线决策、车队高速行进、行人碰撞预警等通信运算融合服务，可以将服务设置为流畅优先/实时同步/高帧率优先，则运算处理的优先级可以低于通信传输的优先级。相应地，可以将通信带宽设置为优于运算的单位时间数据处理量（即在运算资源受限时，降低了运算处理能力需求），以满足实时性的需求。
10

根据一些实施例，在运算处理的优先级高于通信传输的优先级的情况下，可以分配通信和算力资源使得运算的单位时间数据吞吐量大于等于通信的单位时间数据吞吐量，以保证可以及时对所有传输的数据进行运算处理。例如，在对处理结果信息浓度/画质/15 细节要求比较高（例如电影、远程操纵遥控、购物直播、传感器数据捕捉）的情况下，可以将通信运算融合服务设置为画质优先，则通信的优先级可以低于运算的优先级。相应地，运算的单位时间内数据处理量需要优于通信带宽，以满足对于画面质量/处理结果信息浓度的需求。

在运算处理的优先级与通信传输的优先级相当或一致的情况下，可以分配通信和算20 力资源使得运算的单位时间数据吞吐量与通信的单位时间数据吞吐量大致相等。

在一些实施例中，可以通过优先级比较标识符（Priority Comparison Identifier）来指示运算处理和通信传输中的哪一者是需要优先保障的。该标识符例如可以在 PDU 会话建立或修改请求消息中携带。通过上述优先级的比较，可以在运算和通信资源中的任25 一者受限的情况下，将受限资源所对应的（运算或通信）优先级设置为高，从而优先保障较高优先级的（运算或通信）需求。相应地，非受限的资源会存在冗余。另外，在实际部署中，单个实例化的算力资源可能需要支撑一个 PDU 会话中具备运算需求的多个 QoS 流。相应地，该算力资源实例所对应的网络带宽需要与需要支撑的所有 QoS 流的数据带宽（即运算的单位时间数据吞吐量）相匹配。在一些实施例中，可以将运算粒度作为 QoS

保障粒度 (QoS guarantee granularity) 对通信运算融合服务的运算和通信的实时 QoS 性能进行监测。通信运算融合服务的 QoS 保障粒度应当针对需要运算处理的完整数据量 (例如从应用角度是完整的需要处理的数据单元 (data unit))，而不是针对空口传输的单个数据包。因此，在空口传输的单个数据包无法携带实例化运算资源一次需处理的完整数据量时，RAN 需要以实例化运算资源的运算粒度来监测实时 QoS 性能，并对应调整空口资源的分配以满足运算粒度下的 QoS 需求。

在一个实施例中，多个数据包可以在报头中携带同一序列号，以标识这多个数据包属于同一数据处理序列/批次。该序列号可以在数据包过滤器中被识别。相应地，RAN (例如基站) 可以被配置为通过数据包报头中的序列号将多个数据包识别为属于同一数据处理序列，并且监测这多个数据包的实际通信 QoS 性能。

在一个实施例中，可以由 UE 或 AF 提供固定的运算粒度，例如将运算粒度指定为 N 个数据包或 X ms 的时间窗口等。相应地，RAN (例如基站) 可以被配置为将特定数量的数据包或者特定时间窗口内的多个数据包识别为属于同一数据处理序列，并且监测这多个数据包的实际通信 QoS 性能。图 7 示出了根据本公开实施例的用于配置和协调通信服务的运算和通信 QoS 参数的示例处理流程。通过处理流程 700，CF、RAN 和 UPF 可以向 AMF/SMF 提供运算或通信资源分配的参考信息。

在 UE 所请求的会话的运算 QoS 配置信息符合 UE 的策略的情况下，在操作 1a 处，AMF/SMF 可以向 CF 发送运算请求。运算请求可以包括可协商的 QoS 参数，例如运算时延、数据处理能力/算力、QoS 保障的粒度 (或运算粒度)、部署方式等参数；运算需求可以包括不可协商的 QoS 参数，例如计算框架需求、应用场景等。在操作 1b 处，基于可调用的算力资源，CF 可以向 AMF/SMF 反馈能够满足的参数，例如当前能够保障的最优运算 QoS 参数，或者所请求的运算 QoS 参数。CF 也可以反馈一系列当前所能够提供的一系列运算实例的参数集合。基于该反馈信息，AMF/SMF 可以控制或调整运算 QoS 参数或需求，或者可以在运算和通信 QoS 参数或需求之间进行协调，如以上所描述的。之后，AMF/SMF 可以基于调整后的运算 QoS 参数或需求形成通信服务的运算需求，并且进一步请求 CF 实例化相应的算力资源。

在一些实施例中，单个通信服务或 QoS 流的运算需求可能需要多个实例化的算力资源来满足。例如，在分布式处理场景下，可以将并行的运算需求 (例如图像处理需求)

划分为多个子运算需求，并在时间同步条件下针对该运算需求进行分布式运算。又例如，在异构化运算需求场景下，可以将总体运算需求（例如基于图像处理进行数据融合）划分为针对不同运算架构/运算实例的子运算需求（例如图像处理需求和数据融合需求）。根据一些实施例，基于多个子运算需求之间是否具有依赖关系，划分得到的子运算需求
5 可以对应串行、并行或者混合的实例化算力资源。可以参考图 14 中的示例理解运算需求的划分。相应地，CF 可以基于所管理的算力资源以及所能协调的第三方算力资源为子运算需求分配/预分配算力资源。在算力资源分配中，还可以根据可实例化的运算资源的地理位置以及 UE 的地理位置，为 UE 通信服务分配距离（例如欧氏距离）较为接近或者网络传输时延较小的算力资源实例。

10 附加地，在操作 2a 处，AMF/SMF 可以向 UPF 发送网络状态报告请求，其中携带有 PDU 会话的应用标识符以及 UE 地址/ID（例如 SUPI（Subscription Permanent Identifier），SUCI（Subscriber Concealed Identifier），GUTI（Global Unique Temporary Identity）等），以获得外网/DN 的 QoS 数据。在操作 2b 处，UPF 可以向 AMF/SMF 提供网络状态报告。例如，UPF 可以提供针对该 UE 现存的 PDU 会话的外网 QoS 数
15 据（即该 UE 存在其它的类似 PDU 会话，且该 PDU 会话与即将建立/修改的会话属于相同的应用服务）；或者当前存在其它 UE 的相同应用服务的数据流，UPF 可以在不暴露其它 UE 隐私的情况下提供其外网 QoS 数据。例如，UPF 可以针对各个应用服务形成统计表，以分别记录不同应用服务的外网数据流 QoS。

附加地，在操作 3a 处，AMF/SMF 可以向 RAN 发送无线电资源报告请求，以获得无线
20 空口 QoS 数据。在操作 3b 处，RAN 基站可以向 AMF/SMF 提供无线电资源报告，其中包括 UE 的无线空口 QoS 数据。如果该 UE 已有建立好的空口数据链路，并且 RAN 存在对 QoS 流粒度/DRB 数据面的统计数据，则可以容易地获得要报告的无线空口 QoS 数据。另选地，即使当前 UE 没有数据面通信，RAN 可以根据机器学习/人工智能/模式匹配/定位技术等，找到与该 UE 存在类似空口通信环境的数据面通信，并提供对应的 QoS 分析数据。基于无
25 线空口 QoS 数据，AMF/SMF 可以确定可满足或新增的 QoS 参数。

可替换的 QoS 参数以及配置切换

对于通信运算融合服务而言，运算和通信 QoS 参数可能分别存在调整空间。因此，通过配置和协调 QoS 参数，可以获得以优先级排列或标识的运算或通信 QoS 参数表。该

QoS 参数表可以包括可替换的多组运算 QoS 参数或通信 QoS 参数。以运算 QoS 参数为例，每一组运算 QoS 参数可以对应一种可实例化的算力资源配置。可以由 AF 向 PCF 提供可替换的（运算或通信）QoS 需求，继而由 PCF 生成相应的 PCC 规则并指示给 SMF，并由 SMF 生成以优先级排列的可替换的 QoS 配置。可以将可替换的 QoS 配置理解为分级的 QoS 规则，这些分级的 QoS 规则与应用层可接受的不同 QoS 等级相对应。这样，在算力资源或空口资源变化时可以动态地调整为 UE 分配的算力资源或空口资源，从而提供不同等级的 QoS 保障。

应理解，可以从可替换的多组 QoS 参数中选择运算 QoS 配置信息和通信 QoS 配置信息中的至少一者。可替换的多组 QoS 参数以优先级排列或标识多组完整的 QoS 参数，可以基于可利用或分配的通信资源和/或算力资源选择其中一组 QoS 参数。以下描述可以触发切换运算 QoS 配置的示例情况。

在一些实施例中，在当前实例化的算力资源可能无法继续满足运算 QoS 需求或者存在可供实例化的算力资源以提供更好的运算 QoS 性能的情况下，可以例如由 SMF 调整以提高运算 QoS 的优先级。该调整可以包括调整运算 QoS 参数中的部分参数或全部参数。相反，在需要减少使用算力资源的情况下，可以例如由 SMF 调整以降低运算 QoS 的优先级。

根据一些实施例，运算 QoS 参数的调整可能触发通信 QoS 参数的调整（以使得两者相互匹配或互补）。例如，该调整可以包括通信时延的互补调整、带宽的匹配调整以及 QoS 保障粒度的调整，等等。

在一些实施例中，在当前分配的传输资源可能无法继续满足通信 QoS 需求或者存在可分配的传输资源以提供更好的通信 QoS 性能的情况下，可以例如由 SMF 调整以提高通信 QoS 的优先级。该调整可以包括调整通信 QoS 参数中的部分参数或全部参数。相反，在需要减少使用传输资源的情况下，可以例如由 SMF 调整以降低通信 QoS 的优先级。

根据一些实施例，类似地，通信 QoS 参数的调整可能触发运算 QoS 参数的调整（以使得两者相互匹配或互补）。例如，该调整可以包括运算时延的互补调整、运算处理能力的匹配调整，等等。

当然，运算 QoS 和通信 QoS 参数可以分别调整，而不必定存在互相影响的关系。除

了由 SMF 生成以优先级排列的可替换的 QoS 配置外，在一些实施例中，还可以通过 PDU 会话修改流程或 AF Influence 流程来实现运算和通信 QoS 配置的切换。

算力资源部署方式以及实例化

在本公开的实施例中，算力资源可以具有不同的部署方式。算力资源可以来自以下 5 中的至少一者，即核心网算力资源实体、基站或基站侧模块、UPF、UE 或第三方算力资源平台。不同的算力资源部署方式可以对应不同的数据面结构。

在一些实施例中，算力资源可以被实现为单个核心网络功能实体，该功能实体可以称为运算实例 CI (Computing Instance)。相应地，在 RAN 与 CI 之间可以设置有 Tunnel 接口 Ni，并且在 UPF 与 CI 之间可以设置有 Tunnel 接口 Nii。图 8 示出了根据本公开实 10 施例的用于实例化算力资源的示例处理流程。

如图 8 所示，在操作 1a 处，AMF/SMF 可以（例如通过算力资源或运算实例请求）将 RAN 的 Ni Tunnel 信息以及 UPF 的 Nii Tunnel 信息指示给 CF，以及进一步指示给 CI。一经完成算力资源实例化，CF 可以为所获得的 CI 分配 Ni Tunnel 信息和 Nii Tunnel 信息以接收上下行链路数据。在操作 1b 处，CF 将为 CI 分配的 Ni Tunnel 信息和 Nii Tunnel 15 信息指示给 AMF/SMF。通过操作 2a 和 2b，SMF 可以将该 Nii Tunnel 信息指示给对应的 UPF，从而完成网络侧数据面通道的建立。

对应于多种业务场景，对于通信服务可以有以下中的一者或者成立。在一些实施例中，来自终端设备的上行链路数据需要处理，数据处理的结果需要传输至 DN/外网。在一些实施例中，来自 DN/外网的下行链路数据需要处理，数据处理的结果需要传输至终端 20 设备。在一些实施例中，来自终端设备的数据需要处理，数据处理的结果需要返回至该终端设备。相应地，为了区分上述不同的数据处理需求，可以存在以下方案。例如，分别建立相应 QoS 流以服务不同的数据处理需求，即使这些数据处理需求属于同一服务请求或同一 PDU 会话。通过不同 QoS 流的配置，可以设置相应的数据处理及输出结果转发规则。又例如，在单个 QoS 流上混合多种数据处理需求。相应地，对应不同数据处理需求的数据包报头可以携带相应的特定标识符以指示数据处理及输出结果转发规则（例如，25 是否处理，输出结果传输至 UE 还是 DN）。

在一些实施例中，算力资源还可以部署在基站侧，例如基站本身或所集成的运算模

块（例如具有运算功能的服务器）等。可以由基站控制算力资源的部署，这种框架有利于计算 QoS 与通信 QoS 之间的协调，这至少是因为空口资源的控制单元与运算资源的控制单元耦合得更加紧密。在这些实施例中，不需要 Tunnel 信息的分配和传递操作。可以采用更细的 QoS 保障粒度（例如相比于运算粒度），例如可以基于每个数据包实际空口传输时延，确定该数据包所传递的数据的运算时延需求。

在一些实施例中，算力资源可以部署在 UPF 侧，例如作为 UPP 的运算功能特性。相应地，可以将 CF 的功能并入 UPF。

在一些实施例中，算力资源可以作为第三方运算资源，而不必受到核心网络的全面控制。第三方运算资源可以仅与核心网络进行运算 QoS 和通信 QoS 之间的协调。相应地，
10 数据面架构不需要任何改动，只需要通过 CF 与第三方运算资源进行 QoS 协调即可。

根据一些实施例，可以将运算通信融合服务高度重合甚至完全一致的多个终端设备形成群组。例如，对于 V2X 场景中的编组、路径规划、感知融合等服务，或者 AR/VR/涉及多媒体的多玩家互动游戏、直播等服务，多个终端设备需要进行的数据运算和通信需求可能高度重合。例如，这多个终端设备可能均需要对同一画面进行渲染，均需要对同
15 一交通场景进行路径规划/环境感知等。相应地，同一群组的多个终端设备在 PDU 会话建立/修改流程中（或者在注册流程中），可以通过携带相同的组 ID（例如应用层的组 ID、层 2 组 ID）或者该群组终端设备共享的运算服务标识符（Computation Service Identifier），辅助 AMF/SMF 以及 CF 在分配运算资源时将群组的共同运算服务部署到相同的实例化的算力资源。

20 示例方法

图 9 示出了根据本公开实施例的用于通信的示例方法。该方法可以由电子设备 400 或相应网络功能（例如 AMF/SMF）执行。如图 9 所示，该方法 900 可以包括接收来自终端设备的第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS 配置信息，该运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者（框 902）。该方法还可以包括向 CF 提供通信服务的运算需求以实例化相应算力资源，其中该运算需求是在通信服务的运算 QoS 配置信息符合终端设备的策略的情况下生成的（框 904）。可以参考上文关于相应网络功能的描述来理解该方法的进一步细节。

在一个实施例中，方法 900 可以包括接收来自 PCF 的 PCC 规则，其中 PCC 规则包含通信服务的运算 QoS 配置信息，其中终端设备的策略是基于 PCC 规则生成的。

在一个实施例中，运算 QoS 配置信息包括以下中的一者或者：算力需求、运算优先级、运算特性、服务标识符、运算部署方式、运算缓存需求或者算力切片的切片编号。

5 在一个实施例中，第一请求消息还包括通信服务的通信 QoS 配置信息，并且方法 900 可以包括基于运算 QoS 配置信息和通信 QoS 配置信息两者，协调用于运算和通信的 QoS 参数或特性；以及基于经协调的运算 QoS 参数或特性，确定所述通信服务的运算需求，和/或基于经协调的通信 QoS 参数或特性，确定所述通信服务的通信需求。

10 在一个实施例中，协调用于运算和通信的 QoS 参数或特性包括以下中的至少一者：相对于通信服务的运算时延和传输时延中的一者确定运算时延和传输时延中的另一者，使得运算时延与传输时延之和低于或等于阈值水平；基于通信服务的通信和运算处理的相对优先程度，将运算处理能力和通信带宽中的一者确定为与这两者中的另一者相匹配；基于运算粒度对通信服务的运算和通信的实时 QoS 性能进行监测。

15 在一个实施例中，运算 QoS 配置信息和通信 QoS 配置信息中的至少一者是从可替换的多组 QoS 参数中选择的，其中可替换的多组 QoS 参数以优先级排列或标识多组完整的 QoS 参数，并且基于可利用或分配的通信资源和/或算力资源选择其中一组 QoS 参数。

在一个实施例中，方法 900 可以包括接收来自 CF 的经实例化的算力资源的信息；以及向 UE 提供第一响应消息，其中第一响应消息对应于 PDU 会话建立接受或 PDU 会话修改接受消息。

20 在一个实施例中，对于通信服务有以下中的一者或者成立：来自终端设备的上行链路数据需要处理，数据处理的结果需要传输至 DN/外网；来自 DN/外网的下行链路数据需要处理，数据处理的结果需要传输至终端设备；来自终端设备的数据需要处理，数据处理的结果需要返回至终端设备。

图 15 示出了根据本公开实施例的用于通信的示例信令流程。该信令流程 1500 可以在终端设备（例如 UE）和核心网络之间执行。如图 14 所示，在操作 1 处，UE 可以向核心网络（例如由 AMF、SMF、CF 等网络功能构成）发送运算资源注册请求消息，其中包括该 UE 所能提供的运算资源的信息，包括例如运算资源特性或参数、运算资源可获得时间

段 (computation resources available period)、运算资源合约 (例如收费策略)、空口路径 (D2D 或 Uu 接口) 等信息元素。一经接收到运算资源注册请求消息，核心网络可以确认其中的信息元素是否符合该 UE 的策略。如果符合，核心网络 (例如 AMF、SMF、CF) 可以记录该 UE 的运算资源信息。该运算资源信息可以存储在 CF 本地，或者存储在 5 UDM。接着，在操作 2 处，核心网络可用向 UE 发送运算资源注册接受消息。在一些实施例中，上述操作可以通过在注册流程 (registration procedure)、服务请求流程 (service request procedure) 等的信令中携带上述信息元素来实现，或者可以通过专用信令流程实现。通过运算资源注册流程，UE 的运算资源可以成为可供核心网络管理的算力资源。

10 图 16 示出了根据本公开实施例的用于通信的示例信令流程。该信令流程 1600 可以在终端设备 (例如 UE) 和核心网络之间执行。如图 15 所示，在需要实例化算力资源的情况下，CF 可以基于运算需求搜索匹配的运算资源。在操作 1 处，在确认特定 UE 可提供的运算资源能够匹配该运算需求 (包括运算特性参数、运算资源可获得时间段、空口路径等信息元素) 的情况下，核心网络可以向 UE 发送运算实例部署请求消息以将 UE 唤起。 15 一经接收到运算实例部署消息，在操作 2 处，在要部署的算力资源与先前向核心网络注册的资源相符或者资源可用的情况下，UE 向核心网络返回运算实例部署确认 (ACK) 消息。在一些实施例中，上述操作可以通过在寻呼流程、网络触发的服务请求流程等的信令中携带上述信息元素来实现，或者可以通过专用信令流程实现。通过运算资源部署流程，核心网络可以调用特定 UE 空闲的运算资源为其它终端设备或用户的通信服务提供算力支撑。 20

图 10 示出了根据本公开实施例的用于通信的示例方法。该方法可以由电子设备 400 或者相应网络功能 (例如 CF) 执行。如图 10 所示，该方法 1000 可以包括接收来自 SMF 的通信服务的运算需求 (框 1002)。该方法还可以包括向 SMF 提供经实例化的算力资源的信息 (框 1004)。可以参考上文关于相应网络功能的描述来理解该方法的进一步细节。

25 在一个实施例中，算力资源是基于通信服务的运算需求实例化的，并且算力资源来自以下中的至少一者：核心网算力资源实体；基站或基站侧模块、UPF、UE 或第三方算力资源平台。

图 11 示出了根据本公开实施例的用于通信的示例方法。该方法可以由电子设备 500

或者任何终端设备执行。如图 11 所示，该方法 1100 可以包括向网络发送第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS 配置信息，运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者（框 1102）。可选地，该方法还可以包括接收来自网络的响应消息，
5 例如 PDU 会话建立响应或 PDU 会话修改响应消息（框 1104）。可以参考上文关于终端设备的描述来理解该方法的进一步细节。

在一个实施例中，该方法 1100 还可以包括获得通信服务的运算需求，以及基于运算策略将运算需求映射为运算 QoS 参数和/或特性。

图 12A 示出了根据本公开实施例的用于通信的示例方法。该方法可以由电子设备
10 400 或者相应网络功能（例如 PCF 和 AF）执行。如图 12A 所示，该方法 1200 可以包括由 AF 向 PCF 发送 AF 需求，其中 AF 需求基于通信服务的订阅信息或业务场景中的至少一者，并且 AF 需求包括通信服务的运算 QoS 配置信息；相应地，由 PCF 接收来从 AF 接收 AF 需求（框 1202）。该方法还可以包括由 PCF 基于 AF 需求生成 PCC 规则（框 1204）。该方法还可以包括由 PCF 向 SMF 提供 PCC 规则（框 1206）。可以参考上文关于相应网络功能的描
15 述来理解该方法的进一步细节。

图 12B 示出了根据本公开实施例的用于通信的示例方法。该方法可以由电子设备
400 或者相应网络功能（例如 RAN 或其基站）执行。如图 12B 所示，该方法 1250 可以包括通过数据包报头中的序列号将多个数据包识别为属于同一数据处理序列，并且监测这多个数据包的实际通信 QoS 性能（框 1252）。另选地，该方法可以包括将特定数量的数据
20 包或者特定时间窗口内的多个数据包识别为属于同一数据处理序列，并且监测这多个数据包的实际通信 QoS 性能（框 1254）。可以参考上文关于相应网络功能的描述来理解该方法的进一步细节。

以上分别描述了根据本公开实施例的各示例性电子设备和方法。应当理解，这些电子设备的操作或功能可以相互组合，从而实现比所描述的更多或更少的操作或功能。各
25 方法的操作步骤也可以以任何适当的顺序相互组合，从而类似地实现比所描述的更多或更少的操作。

应当理解，根据本公开实施例的机器可读存储介质或程序产品中的机器可执行指令可以被配置为执行与上述设备和方法实施例相应的操作。当参考上述设备和方法实施例

时，机器可读存储介质或程序产品的实施例对于本领域技术人员而言是明晰的，因此不再重复描述。用于承载或包括上述机器可执行指令的机器可读存储介质和程序产品也落在本公开的范围内。这样的存储介质可以包括但不限于软盘、光盘、磁光盘、存储卡、存储棒等等。另外，应当理解，上述系列处理和设备也可以通过软件和/或固件实现。

5 另外，应当理解，上述系列处理和设备也可以通过软件和/或固件实现。在通过软件和/或固件实现的情况下，从存储介质或网络向具有专用硬件结构的计算机，例如图 13 所示的通用计算机 1300 安装构成该软件的程序，该计算机在安装有各种程序时，能够执行各种功能等等。图 13 示出了根据本公开实施例的可实现为终端设备或网络节点的计算机的示例框图。

10 在图 13 中，中央处理单元 (CPU) 1301 根据只读存储器 (ROM) 1302 中存储的程序或从存储部分 1308 加载到随机存取存储器 (RAM) 1303 的程序执行各种处理。在 RAM 1303 中，也根据需要存储当 CPU 1301 执行各种处理等时所需的数据。

CPU 1301、ROM 1302 和 RAM 1303 经由总线 1304 彼此连接。输入/输出接口 1305 也连接到总线 1304。

15 下述部件连接到输入/输出接口 1305：输入部分 1306，包括键盘、鼠标等；输出部分 1307，包括显示器，比如阴极射线管 (CRT)、液晶显示器 (LCD) 等，和扬声器等；存储部分 1308，包括硬盘等；和通信部分 1309，包括网络接口卡比如 LAN 卡、调制解调器等。通信部分 1309 经由网络比如因特网执行通信处理。

20 根据需要，驱动器 1310 也连接到输入/输出接口 1305。可拆卸介质 1311 比如磁盘、光盘、磁光盘、半导体存储器等等根据需要被安装在驱动器 1310 上，使得从中读出的计算机程序根据需要被安装到存储部分 1308 中。

在通过软件实现上述系列处理的情况下，从网络比如因特网或存储介质比如可拆卸介质 1311 安装构成软件的程序。

本领域技术人员应当理解，这种存储介质不局限于图 13 所示的其中存储有程序、
25 与设备相分离地分发以向用户提供程序的可拆卸介质 1311。可拆卸介质 1311 的例子包含磁盘（包含软盘（注册商标））、光盘（包含光盘只读存储器 (CD-ROM) 和数字通用盘 (DVD)）、磁光盘（包含迷你盘 (MD)（注册商标））和半导体存储器。或者，存储介质可以是 ROM 1302、存储部分 1308 中包含的硬盘等等，其中存有程序，并且与包含它们的设

备一起被分发给用户。

应理解，可以通过以下示例实施方式来实现本公开的技术方案。

1、一种用于网络节点的电子设备，所述电子设备包括处理电路，所述处理电路被配置为：

5 接收来自终端设备的第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS 配置信息，所述运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者；以及

向运算功能 CF 提供所述通信服务的运算需求以实例化相应算力资源，其中所述运算需求是在所述通信服务的运算 QoS 配置信息符合所述终端设备的策略的情况下生成的。

10 2、根据条款 1 所述的电子设备，其中所述处理电路还被配置为：

接收来自策略控制功能 PCF 的 PCC 规则，其中所述 PCC 规则包含所述通信服务的运算 QoS 配置信息，

其中，所述终端设备的策略是基于所述 PCC 规则生成的。

3、根据条款 1 所述的电子设备，其中所述运算 QoS 配置信息包括以下中的一者或多者：

算力需求、运算优先级、运算特性、服务标识符、运算部署方式、运算缓存需求或者算力切片的切片编号。

4、根据条款 1 所述的电子设备，其中第一请求消息还包括通信服务的通信 QoS 配置信息，并且所述处理电路还被配置为：

20 基于运算 QoS 配置信息和通信 QoS 配置信息两者，协调用于运算和通信的 QoS 参数或特性；以及

基于经协调的运算 QoS 参数或特性，确定所述通信服务的运算需求，和/或基于经协调的通信 QoS 参数或特性，确定所述通信服务的通信需求。

5、根据条款 4 所述的电子设备，其中协调用于运算和通信的 QoS 参数或特性包括
25 以下中的至少一者：

相对于所述通信服务的运算时延和传输时延中的一者确定运算时延和传输时延中的另一者，使得运算时延与传输时延之和低于或等于阈值水平；

基于所述通信服务的通信和运算处理的相对优先程度，将运算处理能力和通信带宽中的一者确定为与这两者中的另一者相匹配；

5 基于运算粒度对所述通信服务的运算和通信的实时 QoS 性能进行监测。

6、根据条款 4 所述的电子设备，其中运算 QoS 配置信息和通信 QoS 配置信息中的至少一者是从可替换的多组 QoS 参数中选择的，

其中，所述可替换的多组 QoS 参数以优先级排列或标识多组完整的 QoS 参数，并且基于可利用或分配的通信资源和/或算力资源选择其中一组 QoS 参数。

10 7、根据条款 1 所述的电子设备，其中所述处理电路还被配置为：

接收来自 CF 的经实例化的算力资源的信息；以及

向 UE 提供第一响应消息，其中第一响应消息对应于 PDU 会话建立接受或 PDU 会话修改接受消息。

8、根据条款 1 所述的电子设备，其中对于所述通信服务，以下中的一者或者成立：

来自所述终端设备的上行链路数据需要处理，数据处理的结果需要传输至 DN/外网；

来自 DN/外网的下行链路数据需要处理，数据处理的结果需要传输至所述终端设备；

来自所述终端设备的数据需要处理，数据处理的结果需要返回至所述终端设备。

9、根据条款 1 所述的电子设备，其中所述处理电路还被配置为：接收来自终端设备的第二请求消息，第二请求消息包括所述终端设备要注册的算力资源的信息；以及向所述终端设备发送第二响应消息以指示接受所述终端设备的算力资源注册。

10、根据条款 9 所述的电子设备，其中所述处理电路还被配置为：向所述终端设备发送第三请求消息，第三请求消息向所述终端设备指示待实例化的算力资源的信息；以及接收来自所述终端设备的第三响应消息，第三响应消息包括所述终端设备对实例化算力资源的确认。

11、一种用于网络节点的电子设备，其中所述网络节点被配置为实现运算功能 CF，所述电子设备包括处理电路，所述处理电路被配置为：

接收来自会话管理功能 SMF 的通信服务的运算需求；以及

向所述 SMF 提供经实例化的算力资源的信息。

5 12、根据条款 11 所述的电子设备，其中所述算力资源是基于所述通信服务的运算需求实例化的，并且所述算力资源来自以下中的至少一者：

核心网算力资源实体；

基站或基站侧模块；

用户平面功能 UPF；

10 UE；或

第三方算力资源平台。

13、一种用于终端设备的电子设备，包括处理电路，所述处理电路被配置为：

向网络发送第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS 配置信息，所述运算 QoS 15 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者。

14、根据条款 13 所述的电子设备，其中所述处理电路还被配置为：

获得所述通信服务的运算需求；以及

基于运算策略，将所述运算需求映射为运算 QoS 参数和/或特性。

15、根据条款 14 所述的电子设备，其中所述处理电路还被配置为：向网络发送第二请求消息，第二请求消息包括所述终端设备要注册的算力资源的信息；以及接收来自 20 网络的第二响应消息，第二响应消息指示算力资源注册被接受。

16、根据条款 15 所述的电子设备，其中所述处理电路还被配置为：接收来自网络的第三请求消息，第三请求消息指示待实例化的算力资源的信息；以及向所述网络发送第三响应消息，第三响应消息包括所述终端设备对实例化算力资源的确认。

25 17、一种用于网络节点的电子设备，其中所述网络节点被配置为实现应用功能 AF，

所述电子设备包括处理电路，所述处理电路被配置为：

向策略控制功能 PCF 发送 AF 需求，其中所述 AF 需求基于通信服务的订阅信息或业务场景中的至少一者，并且所述 AF 需求包括所述通信服务的运算 QoS 配置信息。

18、一种用于网络节点的电子设备，其中所述网络节点被配置为实现策略控制功能
5 PCF，所述电子设备包括处理电路，所述处理电路被配置为：

从应用功能 AF 接收 AF 需求，其中所述 AF 需求包括通信服务的运算 QoS 配置信息；

基于所述 AF 需求，生成 PCC 规则；以及

向会话管理功能 SMF 提供所述 PCC 规则。

19、一种用于无线电接入网络的电子设备，包括处理电路，所述处理电路被配置为：
10 通过数据包报头中的序列号将多个数据包识别为属于同一数据处理序列，并且监测所述多个数据包的实际通信 QoS 性能；或者

将特定数量的数据包或者特定时间窗口内的多个数据包识别为属于同一数据处理序列，并且监测所述多个数据包的实际通信 QoS 性能。

20、一种无线通信方法，所述方法用于会话管理功能 SMF 并且包括：
15 接收来自终端设备的第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息包括通信服务的至少运算 QoS 配置信息，所述运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者；以及

向运算功能 CF 提供所述通信服务的运算需求以实例化相应算力资源，其中所述运算需求是在所述通信服务的运算 QoS 配置信息符合所述终端设备的策略的情况下生成的。

20 21、根据条款 20 所述的方法，还包括：

接收来自策略控制功能 PCF 的 PCC 规则，其中所述 PCC 规则包含所述通信服务的运算 QoS 配置信息，

其中，所述终端设备的策略是基于所述 PCC 规则生成的。

22、根据条款 20 所述的方法，其中第一请求消息还包括通信服务的通信 QoS 配置
25 信息，并且所述方法还包括：

基于运算 QoS 配置信息和通信 QoS 配置信息两者，协调用于运算和通信的 QoS 参数或特性；以及

基于经协调的运算 QoS 参数或特性，确定所述通信服务的运算需求，和/或基于经协调的通信 QoS 参数或特性，确定所述通信服务的通信需求。

5 23、一种计算机可读存储介质，其上存储有可执行指令，所述可执行指令在由一个或多个处理器执行时，实现根据条款 20 至 22 中任一项所述的方法的操作。

24、一种计算机程序产品，所述计算机程序产品包括指令，所述指令在由计算机执行时使使得实现根据条款 20 至 22 中任一项所述的方法。

以上参照附图描述了本公开的示例性实施例，但是本公开当然不限于以上示例。本
10 领域技术人员可在所附权利要求的范围内得到各种变更和修改，并且应理解这些变更和修改自然将落入本公开的技术范围内。

例如，在以上实施例中包括在一个单元中的多个功能可以由分开的装置来实现。替
选地，在以上实施例中由多个单元实现的多个功能可分别由分开的装置来实现。另外，
以上功能之一可由多个单元来实现。无需说，这样的配置包括在本公开的技术范围内。

15 在该说明书中，流程图中所描述的步骤不仅包括以所述顺序按时间序列执行的处理，而且包括并行地或单独地而不是必须按时间序列执行的处理。此外，甚至在按时间序列处理的步骤中，无需说，也可以适当地改变该顺序。

虽然已经详细说明了本公开及其优点，但是应当理解在不脱离由所附的权利要求所
限定的本公开的精神和范围的情况下可以进行各种改变、替代和变换。而且，本公开实
20 施例的术语“包括”、“包含”或者其任何其它变体意在涵盖非排他性的包含，从而使
得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明
确列出的其它要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在
没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要
素的过程、方法、物品或者设备中还存在另外的相同要素。

权 利 要 求

1、一种用于网络节点的电子设备，所述电子设备包括处理电路，所述处理电路被配置为：

接收来自终端设备的第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS 配置信息，所述运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者；以及

向运算功能 CF 提供所述通信服务的运算需求以实例化相应算力资源，其中所述运算需求是在所述通信服务的运算 QoS 配置信息符合所述终端设备的策略的情况下生成的。

2、根据权利要求 1 所述的电子设备，其中所述处理电路还被配置为：

接收来自策略控制功能 PCF 的 PCC 规则，其中所述 PCC 规则包含所述通信服务的运算 QoS 配置信息，

其中，所述终端设备的策略是基于所述 PCC 规则生成的。

3、根据权利要求 1 所述的电子设备，其中所述运算 QoS 配置信息包括以下中的一者或者多者：

算力需求、运算优先级、运算特性、服务标识符、运算部署方式、运算缓存需求或者算力切片的切片编号。

4、根据权利要求 1 所述的电子设备，其中第一请求消息还包括通信服务的通信 QoS 配置信息，并且所述处理电路还被配置为：

基于运算 QoS 配置信息和通信 QoS 配置信息两者，协调用于运算和通信的 QoS 参数或特性；以及

基于经协调的运算 QoS 参数或特性，确定所述通信服务的运算需求，和/或基于经协调的通信 QoS 参数或特性，确定所述通信服务的通信需求。

5、根据权利要求 4 所述的电子设备，其中协调用于运算和通信的 QoS 参数或特性包括以下中的至少一者：

相对于所述通信服务的运算时延和传输时延中的一者确定运算时延和传输时延中的另一者，使得运算时延与传输时延之和低于或等于阈值水平；

基于所述通信服务的通信和运算处理的相对优先程度，将运算处理能力和通信带宽中的一者确定为与这两者中的另一者相匹配；

基于运算粒度对所述通信服务的运算和通信的实时 QoS 性能进行监测。

6、根据权利要求 4 所述的电子设备，其中运算 QoS 配置信息和通信 QoS 配置信息中的至少一者是从可替换的多组 QoS 参数中选择的，

其中，所述可替换的多组 QoS 参数以优先级排列或标识多组完整的 QoS 参数，并且基于可利用或分配的通信资源和/或算力资源选择其中一组 QoS 参数。

7、根据权利要求 1 所述的电子设备，其中所述处理电路还被配置为：

接收来自 CF 的经实例化的算力资源的信息；以及

向 UE 提供第一响应消息，其中第一响应消息对应于 PDU 会话建立接受或 PDU 会话修改接受消息。

8、根据权利要求 1 所述的电子设备，其中对于所述通信服务，以下中的一者或两者成立：

来自所述终端设备的上行链路数据需要处理，数据处理的结果需要传输至 DN/外网；

来自 DN/外网的下行链路数据需要处理，数据处理的结果需要传输至所述终端设备；

来自所述终端设备的数据需要处理，数据处理的结果需要返回至所述终端设备。

9、根据权利要求 1 所述的电子设备，其中所述处理电路还被配置为：接收来自终端设备的第二请求消息，第二请求消息包括所述终端设备要注册的算力资源的信息；以及向所述终端设备发送第二响应消息以指示接受所述终端设备的算力资源注册。

10、根据权利要求 9 所述的电子设备，其中所述处理电路还被配置为：向所述终端设备发送第三请求消息，第三请求消息向所述终端设备指示待实例化的算力资源的信息；以及接收来自所述终端设备的第三响应消息，第三响应消息包括所述终端设备对实例化算力资源的确认。

11、一种用于网络节点的电子设备，其中所述网络节点被配置为实现运算功能 CF，所述电子设备包括处理电路，所述处理电路被配置为：

接收来自会话管理功能 SMF 的通信服务的运算需求；以及
向所述 SMF 提供经实例化的算力资源的信息。

12、根据权利要求 11 所述的电子设备，其中所述算力资源是基于所述通信服务的
运算需求实例化的，并且所述算力资源来自以下中的至少一者：

核心网算力资源实体；

基站或基站侧模块；

用户平面功能 UPF；

UE；或

第三方算力资源平台。

13、一种用于终端设备的电子设备，包括处理电路，所述处理电路被配置为：

向网络发送第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话
修改请求消息，并且第一请求消息至少包括通信服务的运算 QoS 配置信息，所述运算 QoS
配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者。

14、根据权利要求 13 所述的电子设备，其中所述处理电路还被配置为：

获得所述通信服务的运算需求；以及

基于运算策略，将所述运算需求映射为运算 QoS 参数和/或特性。

15、根据权利要求 14 所述的电子设备，其中所述处理电路还被配置为：向网络发
送第二请求消息，第二请求消息包括所述终端设备要注册的算力资源的信息；以及接收
来自网络的第二响应消息，第二响应消息指示算力资源注册被接受。

16、根据权利要求 15 所述的电子设备，其中所述处理电路还被配置为：接收来自
网络的第三请求消息，第三请求消息指示待实例化的算力资源的信息；以及向所述网络
发送第三响应消息，第三响应消息包括所述终端设备对实例化算力资源的确认。

17、一种用于网络节点的电子设备，其中所述网络节点被配置为实现应用功能 AF，
所述电子设备包括处理电路，所述处理电路被配置为：

向策略控制功能 PCF 发送 AF 需求，其中所述 AF 需求基于通信服务的订阅信息或业

务场景中的至少一者，并且所述 AF 需求包括所述通信服务的运算 QoS 配置信息。

18、一种用于网络节点的电子设备，其中所述网络节点被配置为实现策略控制功能 PCF，所述电子设备包括处理电路，所述处理电路被配置为：

从应用功能 AF 接收 AF 需求，其中所述 AF 需求包括通信服务的运算 QoS 配置信息；
基于所述 AF 需求，生成 PCC 规则；以及
向会话管理功能 SMF 提供所述 PCC 规则。

19、一种用于无线电接入网络的电子设备，包括处理电路，所述处理电路被配置为：
通过数据包报头中的序列号将多个数据包识别为属于同一数据处理序列，并且监测所述多个数据包的实际通信 QoS 性能；或者

将特定数量的数据包或者特定时间窗口内的多个数据包识别为属于同一数据处理序列，并且监测所述多个数据包的实际通信 QoS 性能。

20、一种无线通信方法，所述方法用于会话管理功能 SMF 并且包括：

接收来自终端设备的第一请求消息，其中第一请求消息对应于 PDU 会话建立请求或 PDU 会话修改请求消息，并且第一请求消息包括通信服务的至少运算 QoS 配置信息，所述运算 QoS 配置信息包括运算 QoS 参数或运算 QoS 特性中的至少一者；以及

向运算功能 CF 提供所述通信服务的运算需求以实例化相应算力资源，其中所述运算需求是在所述通信服务的运算 QoS 配置信息符合所述终端设备的策略的情况下生成的。

21、根据权利要求 20 所述的方法，还包括：

接收来自策略控制功能 PCF 的 PCC 规则，其中所述 PCC 规则包含所述通信服务的运算 QoS 配置信息，

其中，所述终端设备的策略是基于所述 PCC 规则生成的。

22、根据权利要求 20 所述的方法，其中第一请求消息还包括通信服务的通信 QoS 配置信息，并且所述方法还包括：

基于运算 QoS 配置信息和通信 QoS 配置信息两者，协调用于运算和通信的 QoS 参数或特性；以及

基于经协调的运算 QoS 参数或特性，确定所述通信服务的运算需求，和/或基于经协调的通信 QoS 参数或特性，确定所述通信服务的通信需求。

23、一种计算机可读存储介质，其上存储有可执行指令，所述可执行指令在由一个或多个处理器执行时，实现根据权利要求 20 至 22 中任一项所述的方法的操作。

24、一种计算机程序产品，所述计算机程序产品包括指令，所述指令在由计算机执行时使得实现根据权利要求 20 至 22 中任一项所述的方法。

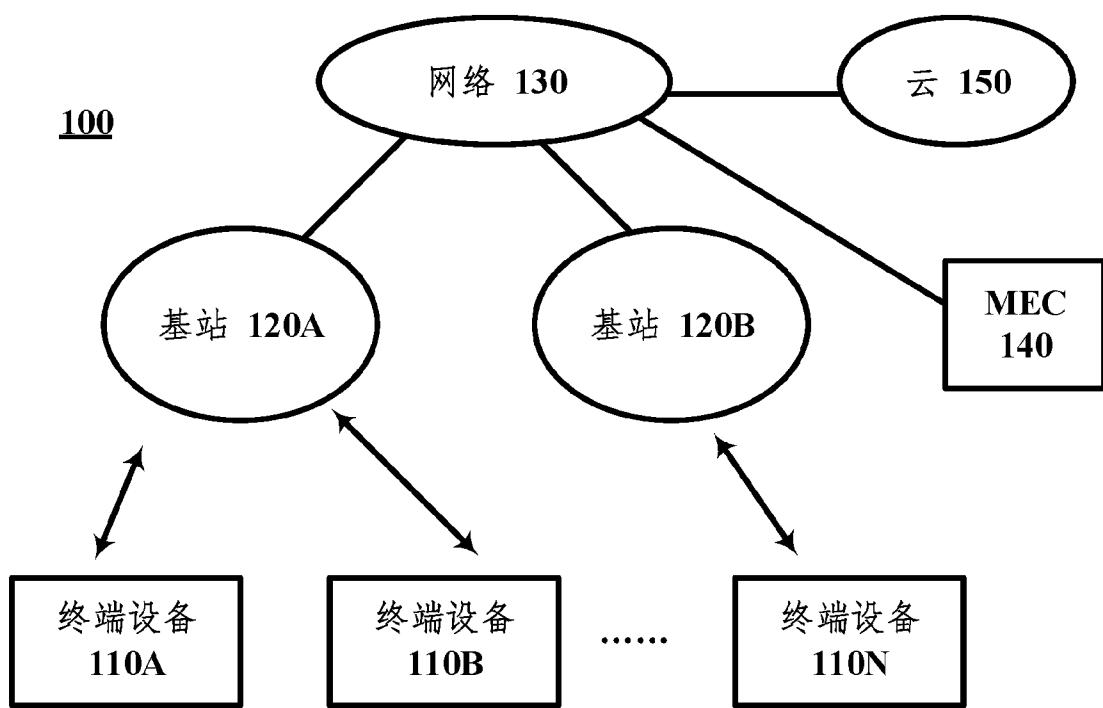


图 1

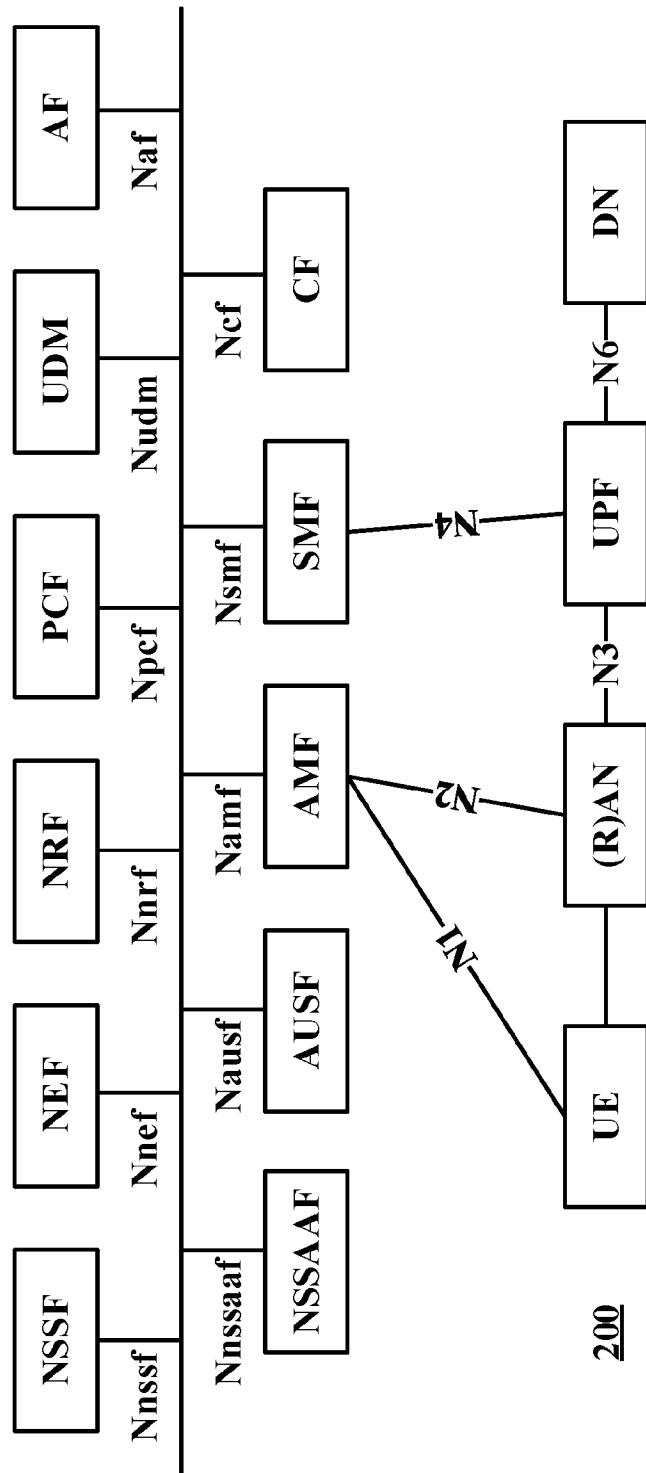


图 2

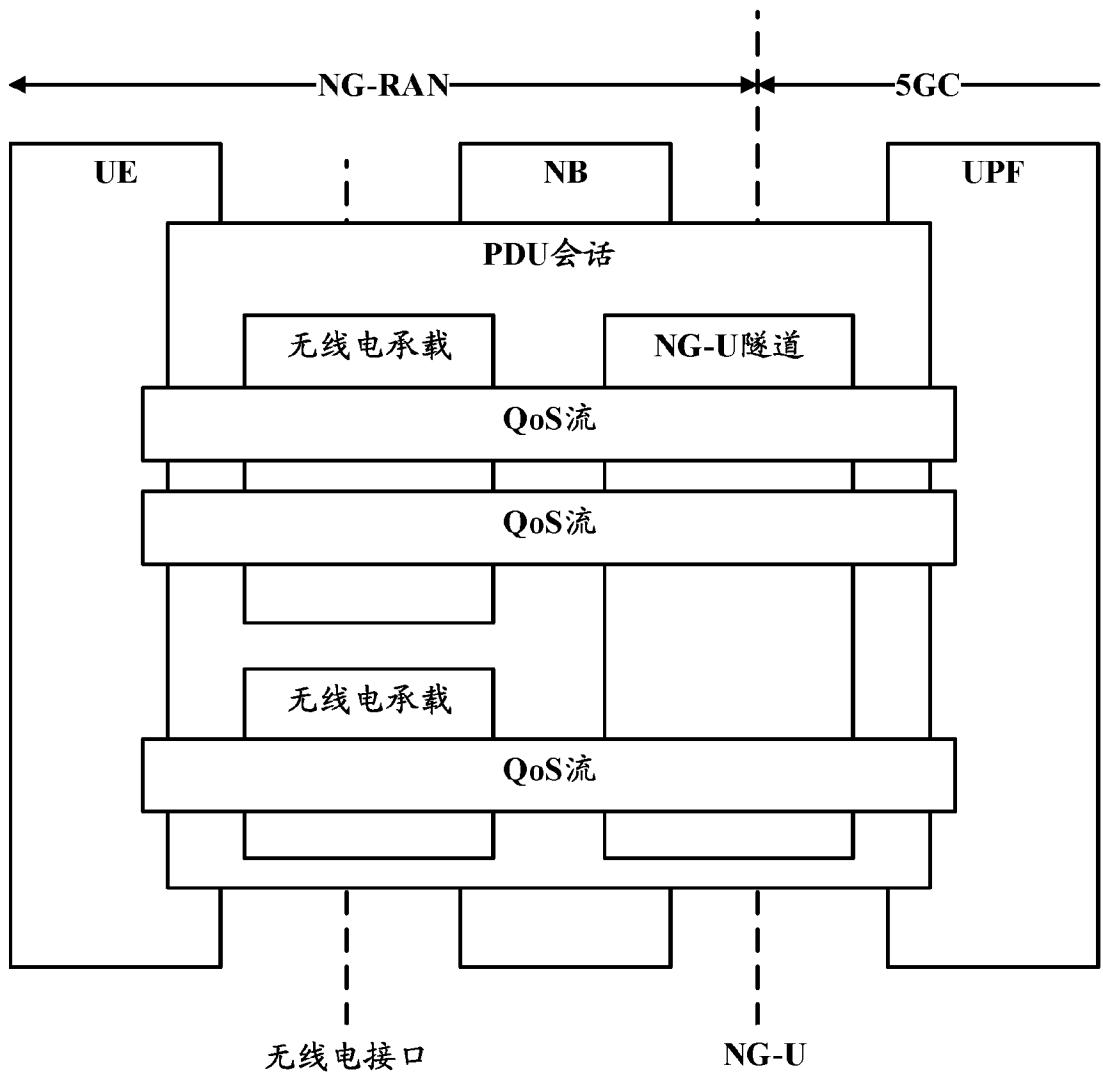


图 3

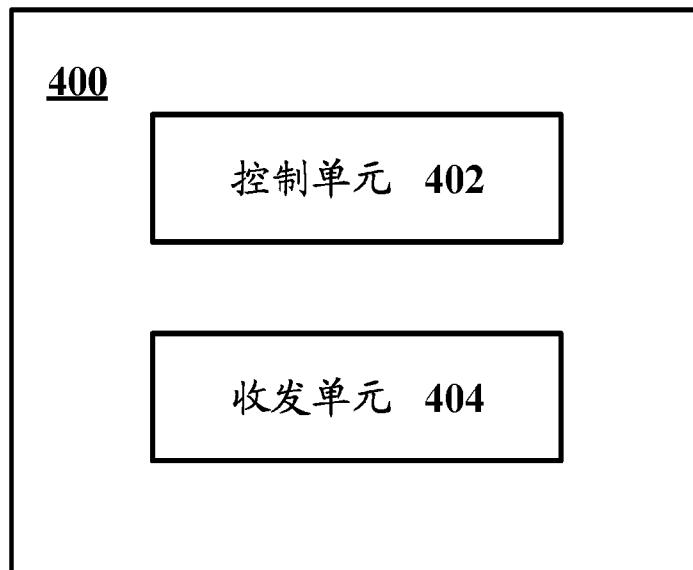


图 4

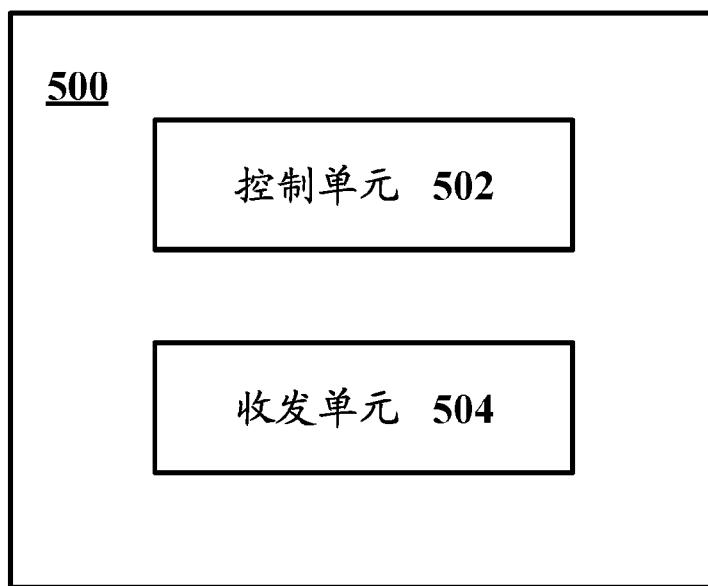


图 5

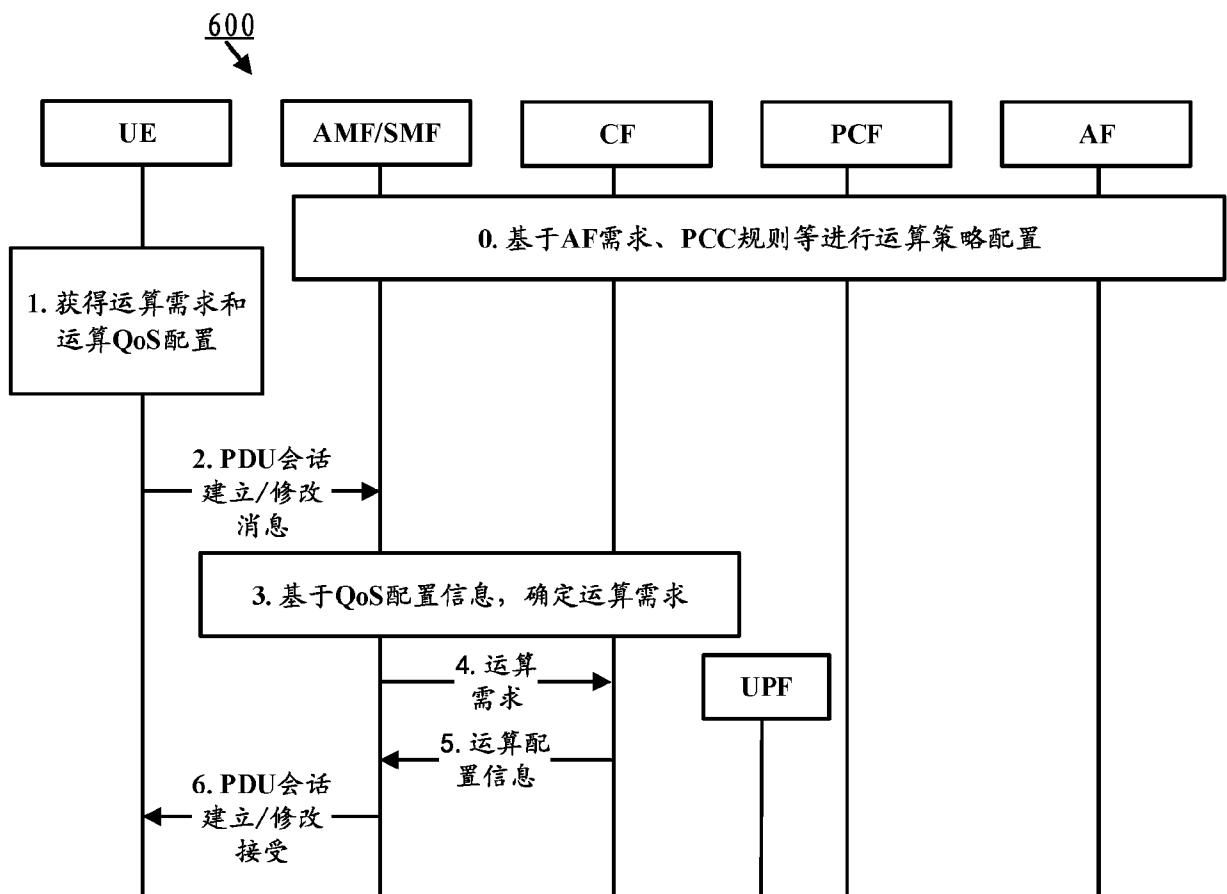


图 6

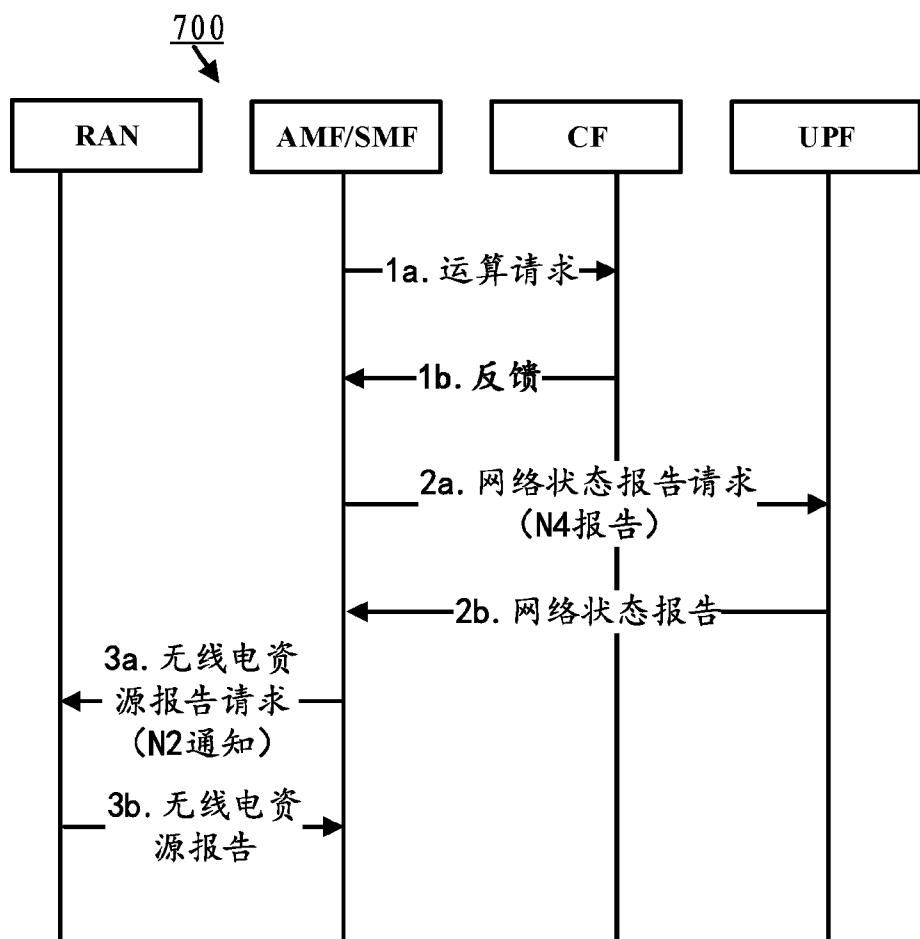


图 7

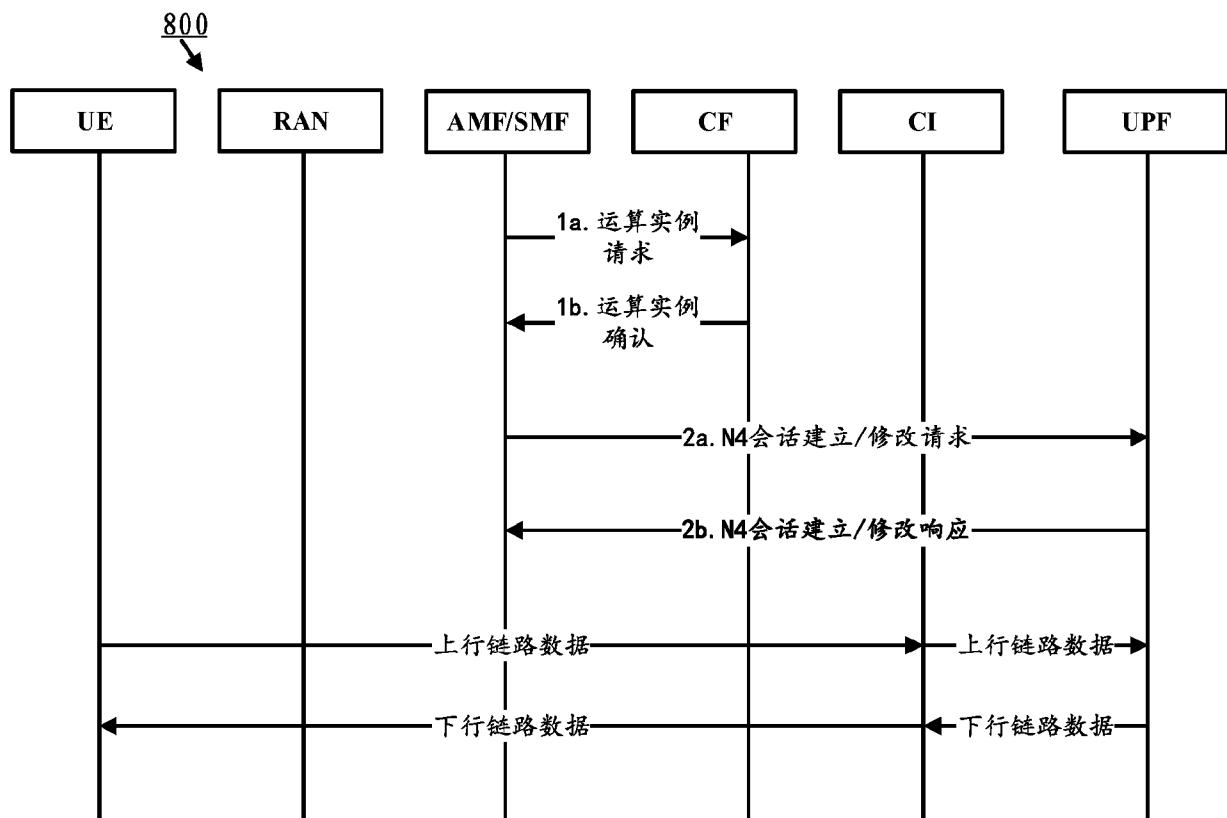


图 8

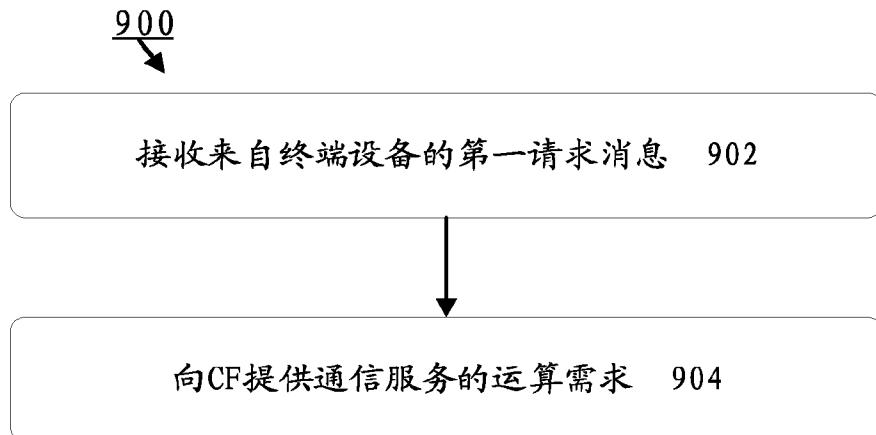


图 9

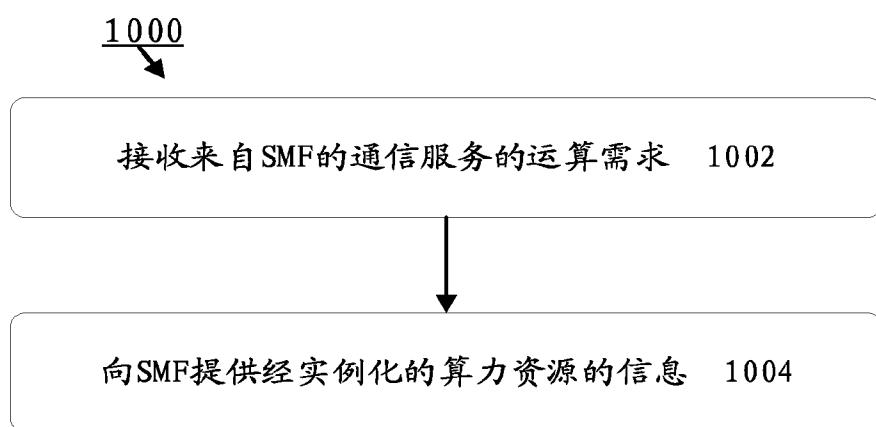


图 10

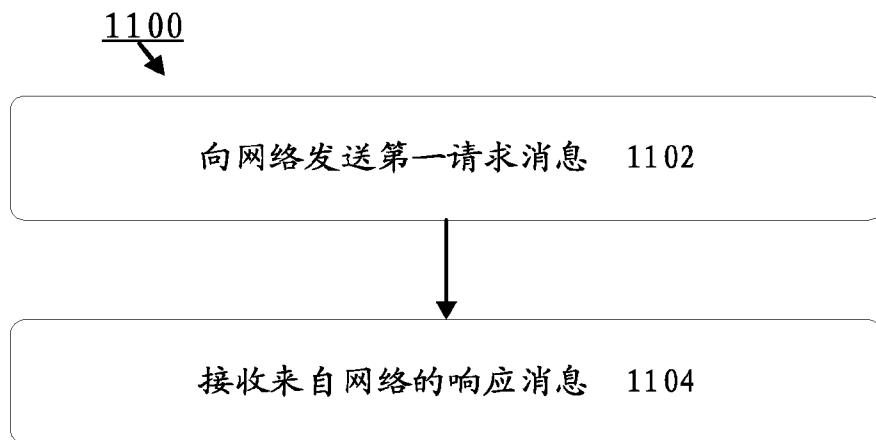
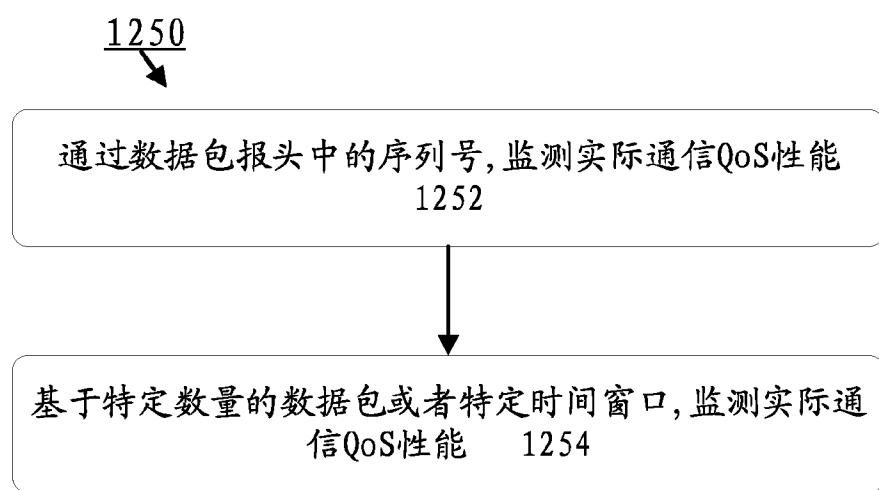
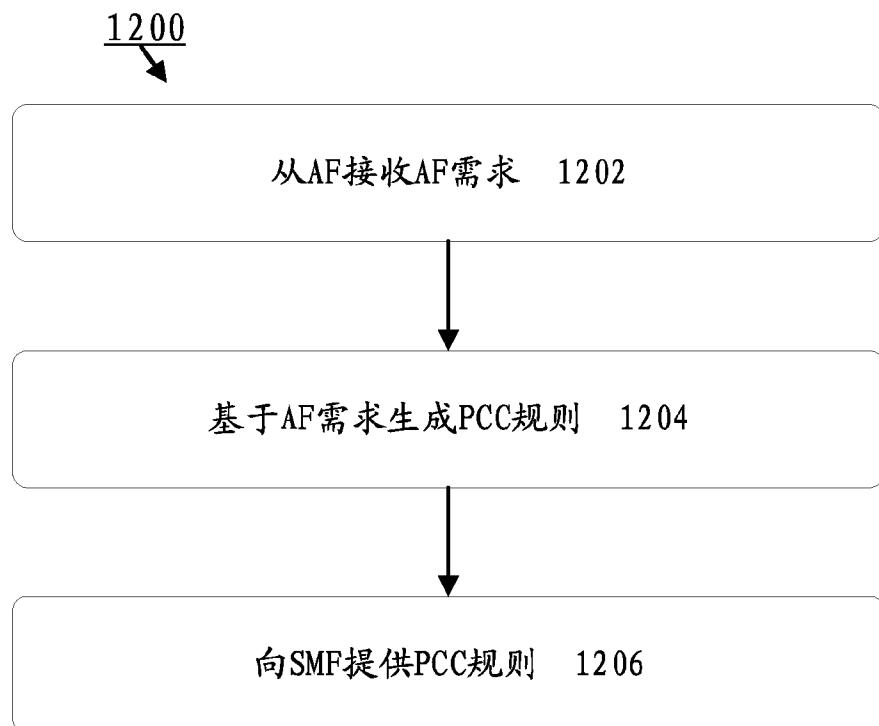


图 11



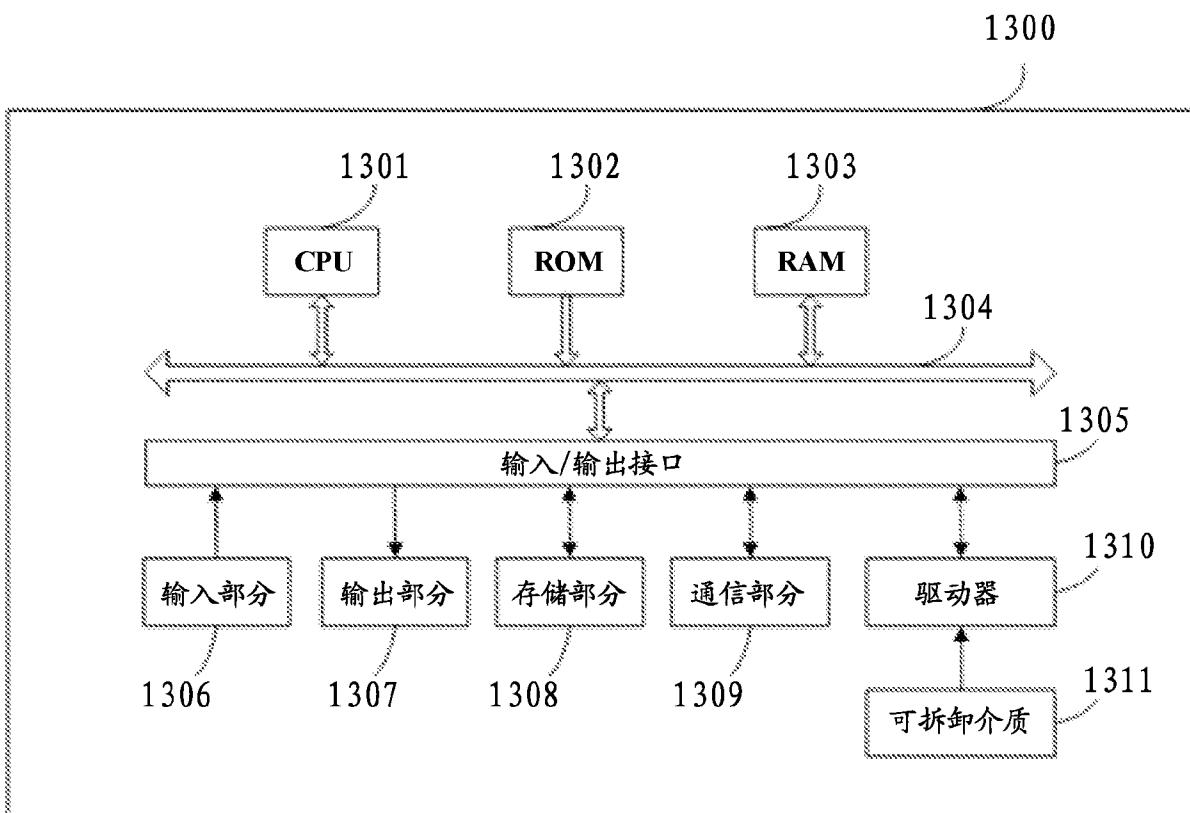


图 13

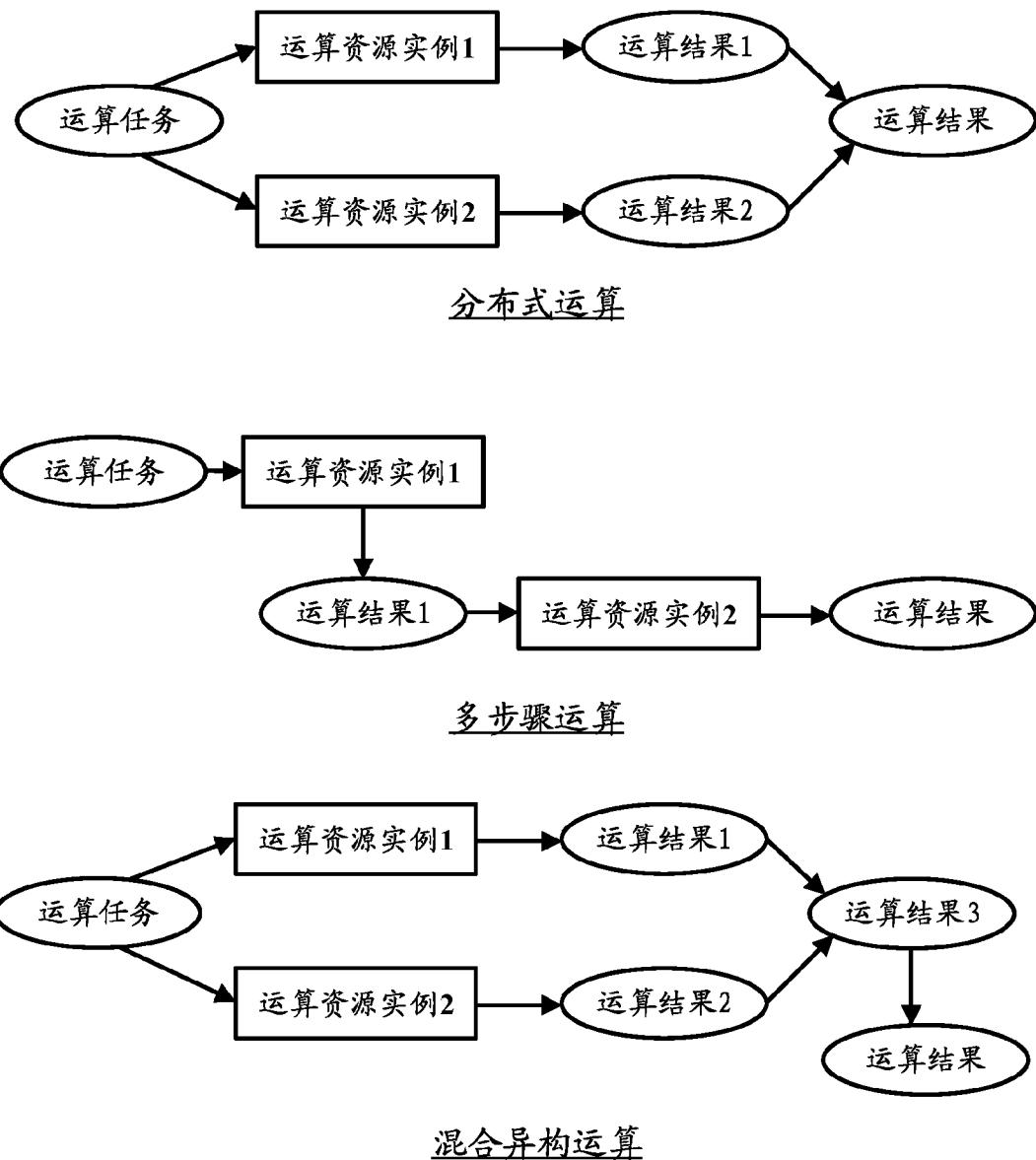


图 14

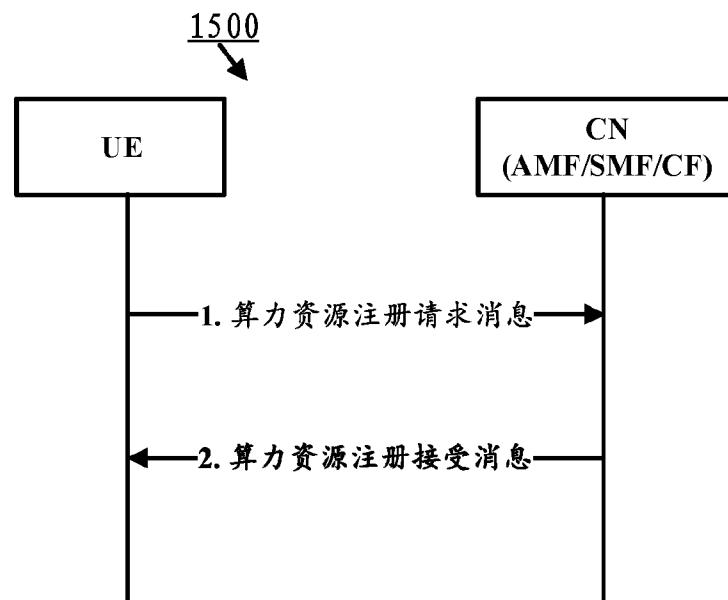


图 15

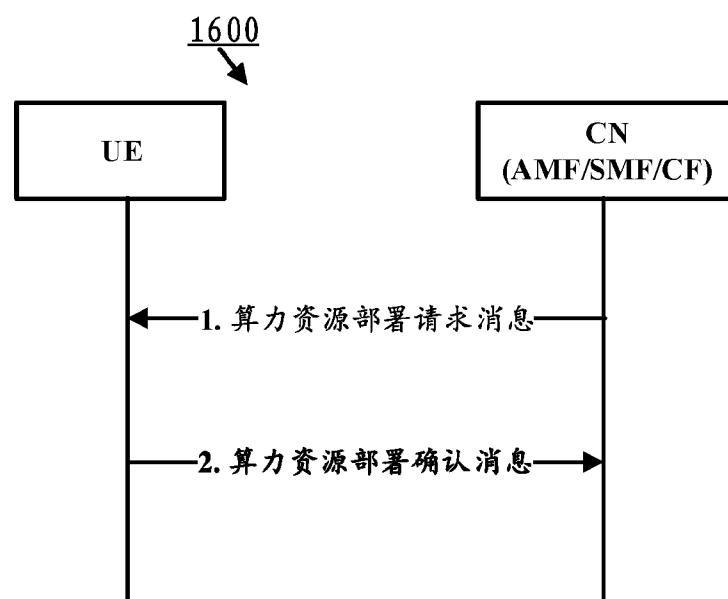


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/130328

A. CLASSIFICATION OF SUBJECT MATTER

H04W 28/24(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)

H04W28/-

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)

CNTXT; WPABSC; ENTXTC; CNKI: 会话, 计算, 算力, 运算, 服务质量, 需求, 能力, 优先级, 切片, 边缘, 性能, 数据包, 分组, 相同, 同一; WPABS; ENTXT; DWPI; 3GPP; IEEE: PDU, session, comput+, calculat+, capacity, ability, power, requirement, QoS, SMF, MEC, CF, PCF, PCC, 5G

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 109644377 B (OPPO GUANGDONG MOBILE TELECOMMUNICATIONS CO., LTD.) 12 January 2021 (2021-01-12) description, paragraphs [0086]-[0087] and [0130]	19
X	CN 112465359 A (CHINA UNITED NETWORK COMMUNICATIONS GROUP CO., LTD.) 09 March 2021 (2021-03-09) description, paragraphs [0043]-[0068] and [0089]	11-12
Y	CN 112465359 A (CHINA UNITED NETWORK COMMUNICATIONS GROUP CO., LTD.) 09 March 2021 (2021-03-09) description, paragraphs [0043]-[0068] and [0089]	1-10, 13-18, 20-24
Y	KR 20200115155 A (SAMSUNG ELECTRONICS CO., LTD.) 07 October 2020 (2020-10-07) description, paragraphs [0046]-[0070] and [0136]	1-10, 13-18, 20-24
A	CN 111771355 A (HUAWEI TECHNOLOGIES CO., LTD.) 13 October 2020 (2020-10-13) entire document	1-24
A	CN 111684774 A (NOKIA SOLUTIONS AND NETWORKS OY) 18 September 2020 (2020-09-18) entire document	1-24

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

16 January 2023

Date of mailing of the international search report

28 January 2023

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

Authorized officer

Facsimile No. **(86-10)62019451**

Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2022/130328

Patent document cited in search report				Publication date (day/month/year)		Patent family member(s)		Publication date (day/month/year)	
CN	109644377	B	12 January 2021	CN	109644377	A	16 April 2019		
				WO	2018170837	A1	27 September 2018		
CN	112465359	A	09 March 2021	None					
KR	20200115155	A	07 October 2020	WO	2020197288	A1	01 October 2020		
				CN	113841432	A	24 December 2021		
				EP	3934291	A1	05 January 2022		
				US	2022191100	A1	16 June 2022		
CN	111771355	A	13 October 2020	WO	2019158218	A1	22 August 2019		
				IN	202037036970	A	02 October 2020		
				US	2020383006	A1	03 December 2020		
				EP	3753202	A1	23 December 2020		
				CN	111771355	B	19 November 2021		
				CN	114222264	A	22 March 2022		
CN	111684774	A	18 September 2020	WO	2019126931	A1	04 July 2019		
				EP	3732846	A1	04 November 2020		
				US	2021076250	A1	11 March 2021		
				US	11134410	B2	28 September 2021		
				CN	111684774	B	19 August 2022		

国际检索报告

国际申请号

PCT/CN2022/130328

A. 主题的分类

H04W 28/24 (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)

H04W28/-

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))

CNTXT;WPABSC;ENTXT;CNKI: 会话, 计算, 算力, 运算, 服务质量, 需求, 能力, 优先级, 切片, 边缘, 性能, 数据包, 分组, 相同, 同一 WPABS;ENTXT;DWPI;3GPP;IEEE: PDU, session, comput+, calculat+, capacity, ability, power, requirement, QoS, SMF, MEC, CF, PCF, PCC, 5G

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 109644377 B (OPPO广东移动通信有限公司) 2021年1月12日 (2021 - 01 - 12) 说明书[0086]-[0087], [0130]段	19
X	CN 112465359 A (中国联合网络通信集团有限公司) 2021年3月9日 (2021 - 03 - 09) 说明书[0043]-[0068], [0089]段	11-12
Y	CN 112465359 A (中国联合网络通信集团有限公司) 2021年3月9日 (2021 - 03 - 09) 说明书[0043]-[0068], [0089]段	1-10, 13-18, 20-24
Y	KR 20200115155 A (SAMSUNG ELECTRONICS CO. LTD.) 2020年10月7日 (2020 - 10 - 07) 说明书[0046]-[0070], [0136]段	1-10, 13-18, 20-24
A	CN 111771355 A (华为技术有限公司) 2020年10月13日 (2020 - 10 - 13) 全文	1-24
A	CN 111684774 A (诺基亚通信公司) 2020年9月18日 (2020 - 09 - 18) 全文	1-24

 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 2023年1月16日	国际检索报告邮寄日期 2023年1月28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	受权官员 解淑瑄 电话号码 (86-28) 62969258

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/130328

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	109644377	B	2021年1月12日	CN	109644377	A	2019年4月16日
				WO	2018170837	A1	2018年9月27日
CN 112465359 A 2021年3月9日			无				
KR	20200115155	A	2020年10月7日	WO	2020197288	A1	2020年10月1日
				CN	113841432	A	2021年12月24日
				EP	3934291	A1	2022年1月5日
				US	2022191100	A1	2022年6月16日
CN 111771355 A 2020年10月13日			WO	2019158218	A1	2019年8月22日	
			IN	202037036970	A	2020年10月2日	
			US	2020383006	A1	2020年12月3日	
			EP	3753202	A1	2020年12月23日	
			CN	111771355	B	2021年11月19日	
			CN	114222264	A	2022年3月22日	
CN 111684774 A 2020年9月18日			WO	2019126931	A1	2019年7月4日	
			EP	3732846	A1	2020年11月4日	
			US	2021076250	A1	2021年3月11日	
			US	11134410	B2	2021年9月28日	
			CN	111684774	B	2022年8月19日	