

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024年6月20日 (20.06.2024)



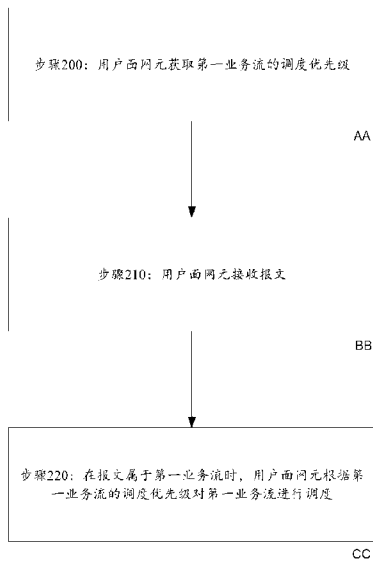
(10) 国际公布号
WO 2024/125232 A1

- (51) 国际专利分类号:
H04W 72/566 (2023.01)
- (21) 国际申请号: PCT/CN2023/132949
- (22) 国际申请日: 2023年11月21日 (21.11.2023)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202211604610.3 2022年12月13日 (13.12.2022) CN
- (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: 胡翔 (HU, Xiang); 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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

(54) Title: SERVICE FLOW SCHEDULING METHOD AND APPARATUS

(54) 发明名称: 一种业务流调度方法及装置



- AA Step 200, A user plane network element acquires a scheduling priority level of a first service flow
- BB Step 210, The user plane network element receives a packet
- CC Step 220, When the packet belongs to the first service flow, the user plane network element schedules the first service flow according to the scheduling priority level of the first service flow

(57) Abstract: A service flow scheduling method and apparatus. The method comprises: a user plane network element acquires a scheduling priority level of a first service flow, wherein the scheduling priority level of the first service flow is the scheduling priority level of the first service flow in a plurality of service flows, and the first service flow is one of the plurality of service flows; the user plane network element receives a packet; and when the packet belongs to the first service flow, the user plane network element schedules the first service flow according to the scheduling priority level of the first service flow. By using the design, the scheduling priority levels of different service flows in a plurality of service flows can be distinguished, and scheduling is performed according to the scheduling priority levels of different service flows.

(57) 摘要: 一种业务流调度方法及装置, 该方法包括: 用户面网元获取第一业务流的调度优先级, 第一业务流的调度优先级为第一业务流在多个业务流中的调度优先级; 第一业务流为多个业务流中的一个, 用户面网元接收报文。在报文属于第一业务流时, 用户面网元根据第一业务流的调度优先级对第一业务流进行调度。采用上述设计, 可以实现对多个业务流中的不同业务流的调度优先级进行区分并根据不同业务流的调度优先级进行调度。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

一种业务流调度方法及装置

相关申请的交叉引用

本申请要求在2022年12月13日提交中华人民共和国知识产权局、申请号为202211604610.3、发明名称为“一种业务流调度方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请实施例涉及无线通信领域，尤其涉及一种业务流调度方法及装置。

背景技术

10 在第5代移动通信技术（the 5th generation mobile communication technology, 5G）系统中，服务质量流（QoS Flow）可以承载很多内容，其中，每个QoS流可以承载多种业务。

当前，描述QoS流在终端设备和用户面网元之间的报文转发处理的参数可以包括以下参数：资源类型、优先级（priority level）、分组时延预算（包括核心网分组时延预算）、包错误率、平均窗口、最大数据突发卷等。

15 其中，上述参数中的优先级仅用于不同QoS流之间的优先级调度，而由于每个QoS流可以承载多种业务，不同业务对应的业务流也可能具有不同的优先级需求，但当前在QoS流内部的不同业务对应的业务流无法实现差异化调度。

此外，对于多个业务流共用一个固定带宽的其他可能场景也存在相同的问题。

20 发明内容

本申请提供一种业务流调度方法及装置，用以解决在多个业务流共用一个固定带宽时无法实现对多个业务流进行差异化调度的问题。

第一方面，本申请提供一种业务流调度方法，该方法包括：

25 用户面网元获取第一业务流的调度优先级，所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级；所述第一业务流为所述多个业务流中的一个；所述用户面网元接收报文，在所述报文属于所述第一业务流时，所述用户面网元根据所述第一业务流的调度优先级对所述第一业务流进行调度。

采用上述设计，用户面网元可以获得业务流的调度优先级，进而能够对不同的调度优先级的业务流进行调度，实现对多个业务流进行差异化调度。

30 在一种可能的设计中，在用户面网元获取第一业务流的调度优先级时，所述用户面网元从会话管理网元接收第一规则，所述第一规则用于指示所述第一业务流的调度优先级。

在一种可能的设计中，所述第一规则包括第一预定义规则的名称，所述第一预定义规则包括所述第一业务流的调度优先级，所述第一预定义规则为所述用户面网元存储的多个预定义规则中的一个。

35 采用上述设计，用户面网元可以提前获得多个预定义规则，并在获得第一预定义规则的名称时，激活第一预定义规则。

在一种可能的设计中，所述用户面网元还获取所述第一业务流的带宽参数，所述第一业务流的带宽参数包括所述第一业务流的保障带宽，和/或所述第一业务流的最大带宽；或者，所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率，和/或所述第一业务流的最大带宽比率；其中，所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率，所述第一业务流的最大带宽比率为所述

40 为所述第一业务流所能占用的最大带宽与总带宽的比率，所述总带宽是为所述多个业务流分配的共用带宽。

采用上述设计，可以实现用户面网元在调度业务流时保证业务流的保障带宽，并基于业务流的调度优先级保证业务流的最大带宽。

在一种可能的设计中，所述第一预定义规则还包括所述第一业务的带宽参数。

在一种可能的设计中, 在所述多个业务流属于同一个 QoS 流时, 所述第一规则包括 QoS 执行规则的标识, 所述 QoS 执行规则包括所述第一业务流的调度优先级。

采用上述设计, 可以通过扩展 QoS 执行规则将第一业务流的调度优先级加入 QoS 执行规则。

在一种可能的设计中, 所述 QoS 执行规则还包括所述第一业务流的带宽参数。

5 在一种可能的设计中, 在所述多个业务流属于同一个 QoS 流时, 所述第一规则为包检测规则。

在一种可能的设计中, 在所述用户面网元根据所述第一业务流的调度优先级对所述第一业务流进行调度时, 在所述多个业务流分别对应的需求带宽之和大于所述总带宽时, 若在所述多个业务流中存在第二业务流, 所述第二业务流的保障带宽大于所述第二业务流的需求带宽, 则为所述第二业务流分配的带宽等于所述第二业务流的需求带宽; 根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽, 按照调度优先级由高到低的顺序为所述多个业务流中除所述第二业务流之外的业务流分配差值带宽, 所述差值带宽为所述第二业务流的保障带宽与所述第二业务流的需求带宽的差值。

采用上述设计, 可以实现对于保障带宽大于需求带宽的业务流时, 保证该业务流的需求带宽, 基于其他业务流的调度优先级和需求带宽, 为这些业务流分配差值带宽, 进而提升总带宽的利用效率。

15 在一种可能的设计中, 在所述用户面网元根据所述第一业务流的调度优先级对所述第一业务流进行调度时, 在所述多个业务流分别对应的需求带宽之和大于所述总带宽时, 满足所述多个业务流分别对应的保障带宽比率, 所述多个业务流分别对应的保障带宽比率之和小于或等于 1; 若所述总带宽有剩余, 根据所述多个业务流分别对应的调度优先级, 按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽。

采用上述设计, 在各个业务流的需求带宽均大于相应的保障带宽时, 优先保证各个业务流的保障带宽, 基于各个业务流的调度优先级和需求带宽, 为各个业务流分配剩余带宽。因此, 可以实现保证各个业务流的保障带宽, 且优先满足高调度优先级的业务流的需求带宽。

在一种可能的设计中, 在按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽时, 按照所述顺序满足所述多个业务流中部分或全部业务流的最大带宽比率。

采用上述设计, 可以基于各个业务流的调度优先级、需求带宽和最大带宽比率, 为各个业务流分配剩余带宽, 进而可以实现优先满足高调度优先级的业务流的最大带宽比率。

25 在一种可能的设计中, 所述多个业务流分别对应的实际带宽之和等于所述总带宽。

采用上述设计, 可以提升总带宽的利用效率。

在一种可能的设计中, 所述报文中为下行报文; 所述用户面网元在所述下行报文中添加所述第一业务流的调度优先级, 获得第一待发送下行报文; 所述用户面网元向接入网设备发送所述第一待发送下行报
30 文。

采用上述设计可以实现将业务流的调度优先级通过下行报文传输至接入网设备。

在一种可能的设计中, 所述第一待发送下行报文包括通用无线分组业务隧道协议报文头, 所述通用无线分组业务隧道协议报文头包括所述第一业务流的调度优先级。

35 在一种可能的设计中, 所述报文中为下行报文; 所述用户面网元在所述下行报文中添加反射 QoS 指示和所述 QoS 流的 QoS 流标识, 获得第二待发送下行报文。所述用户面网元向接入网设备发送所述第二待发送下行报文。

采用上述设计可以实现将反射 QoS 指示和所述 QoS 流的 QoS 流标识通过下行报文传输至接入网设备, 以使接入网设备将反射 QoS 指示和所述 QoS 流的 QoS 流标识传输至终端设备。

40 在一种可能的设计中, 所述第二待发送下行报文包括通用无线分组业务隧道协议报文头, 所述通用无线分组业务隧道协议报文头包括所述反射 QoS 指示和所述 QoS 流的 QoS 流标识。

在一种可能的设计中, 所述多个业务流为多个终端设备分别对应的业务流, 所述多个终端设备属于同一个终端设备组。

45 在一种可能的设计中, 所述多个终端设备的等级不同, 所述第一业务流的调度优先级是根据第一终端设备的等级确定的, 所述第一业务流为所述第一终端设备签约的业务对应的业务流, 所述第一终端设备为所述多个终端设备中的一个。

在一种可能的设计中, 所述多个业务流属于同一个 QoS 流。

第二方面, 本申请提供一种业务流调度方法, 该方法包括: 会话管理网元获取第一业务流的调度优先级, 所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级; 所述第一业务流

为所述多个业务流中的一个；所述会话管理网元向用户面网元发送第一规则，所述第一规则用于指示所述第一业务流的调度优先级。

采用上述设计，会话管理网元可以获得业务流的调度优先级，并通过第一规则向用户面网元指示业务流的调度优先级，以使用户面网元可以根据业务流的调度优先级，进而对不同的调度优先级的业务流进行调度，实现对多个业务流进行差异化调度。

在一种可能的设计中，在会话管理网元获取第一业务流的调度优先级时，所述会话管理网元从策略控制网元接收策略与计费控制规则，所述策略与计费控制规则用于指示所述第一业务流的调度优先级；所述会话管理网元根据所述策略与计费控制规则确定所述第一规则。

采用上述设计，会话管理网元可以从策略控制网元接收策略与计费控制规则，进而根据策略与计费控制规则确定第一规则。

在一种可能的设计中，所述策略与计费控制规则包括第一预定义规则的名称，所述第一规则包括所述第一预定义规则的名称；所述第一预定义规则包括所述第一业务流的调度优先级，所述第一预定义规则为所述会话管理网元和所述用户面网元存储的多个预定义规则中的一个。

采用上述设计，会话管理网元可以提前获得多个预定义规则，并在获得第一预定义规则的名称时，激活第一预定义规则，以及通过第一规则携带第一预定义规则的名称。

在一种可能的设计中，所述策略与计费控制规则包括所述第一业务流的调度优先级；所述第一规则包括 QoS 执行规则的标识，所述 QoS 执行规则包括所述第一业务流的调度优先级。

采用上述设计，可以通过扩展 QoS 执行规则将第一业务流的调度优先级加入 QoS 执行规则。

在一种可能的设计中，所述会话管理网元根据所述策略与计费控制规则确定 QoS 规则，所述 QoS 规则包括所述第一业务流的调度优先级；所述会话管理网元向所述终端设备发送所述 QoS 规则。

在一种可能的设计中，所述第一预定义规则还包括反射 QoS 指示。

在一种可能的设计中，所述策略与计费控制规则还包括反射 QoS 指示；所述第一规则还包括所述反射 QoS 指示。

在一种可能的设计中，所述会话管理网元还获取所述第一业务流的带宽参数，所述第一业务流的带宽参数包括所述第一业务流的保障带宽，和/或所述第一业务流的最大带宽；或者，所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率，和/或所述第一业务流的最大带宽比率；其中，所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率，所述第一业务流的最大带宽比率为所述第一业务流所能占用的最大带宽与总带宽的比率，所述总带宽是为所述多个业务流分配的共用带宽。

在一种可能的设计中，所述 QoS 规则还包括所述第一业务流的带宽参数。

在一种可能的设计中，所述多个业务流为多个终端设备分别对应的业务流，所述多个终端设备属于同一个终端设备组。

在一种可能的设计中，所述多个终端设备的等级不同，所述第一业务流的调度优先级是根据第一终端设备的等级确定的，所述第一业务流为所述第一终端设备签约的业务对应的业务流，所述第一终端设备为所述多个终端设备中的一个。

在一种可能的设计中，所述多个业务流属于同一个服务质量 QoS 流。

第三方面，本申请提供一种业务流调度方法，该方法包括：

接入网设备接收第一业务流中的报文，所述报文包括第一业务流的调度优先级，所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级；所述第一业务流为所述多个业务流中的一个；所述接入网设备根据所述第一业务流的调度优先级对所述第一业务流进行调度。

采用上述设计，接入网设备可以根据接收到的报文中的业务流的调度优先级，对该业务流进行调度，进而能够对不同的调度优先级的业务流进行调度，实现对多个业务流进行差异化调度。

在一种可能的设计中，所述报文还包括所述第一业务流的带宽参数，所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率，和/或所述第一业务流的最大带宽比率；

其中，所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率，所述第一业务流的最大带宽比率为所述第一业务流所能占用的最大带宽与总带宽的比率，所述总带宽是为所述多个业务流分配的共用带宽。

在一种可能的设计中，在所述接入网设备根据所述第一业务流的调度优先级对所述第一业务流进行

调度时,在所述多个业务流分别对应的需求带宽之和大于所述总带宽时,若在所述多个业务流中存在第二业务流,所述第二业务流的保障带宽大于所述第二业务流的需求带宽,则为所述第二业务流分配的带宽等于所述第二业务流的需求带宽;根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽,按照调度优先级由高到低的顺序为所述多个业务流中除所述第二业务流之外的业务流分配差值带宽,所述差值带宽为所述第二业务流的保障带宽与所述第二业务流的需求带宽的差值。

采用上述设计,可以实现对于保障带宽大于需求带宽的业务流时,保证该业务流的需求带宽,基于其他业务流的调度优先级和需求带宽,为这些业务流分配差值带宽,进而提升总带宽的利用效率。

在一种可能的设计中,在所述接入网设备根据所述第一业务流的调度优先级对所述第一业务流进行调度时,在所述多个业务流分别对应的需求带宽之和大于所述总带宽时,满足所述多个业务流分别对应的保障带宽比率,所述多个业务流分别对应的保障带宽比率之和小于或等于 1;若所述总带宽有剩余,根据所述多个业务流分别对应的调度优先级,按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽。

采用上述设计,在各个业务流的需求带宽均大于相应的保障带宽时,优先保证各个业务流的保障带宽,基于各个业务流的调度优先级和需求带宽,为各个业务流分配剩余带宽。因此,可以实现保证各个业务流的保障带宽,且优先满足高调度优先级的业务流的需求带宽。

在一种可能的设计中,在按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽时,按照所述顺序满足所述多个业务流中部分或全部业务流的最大带宽比率。

采用上述设计,可以基于各个业务流的调度优先级、需求带宽和最大带宽比率,为各个业务流分配剩余带宽,进而可以实现优先满足高调度优先级的业务流的最大带宽比率。

在一种可能的设计中,所述多个业务流分别对应的实际带宽之和等于所述总带宽。

采用上述设计,可以提升总带宽的利用效率。

在一种可能的设计中,所报文包括通用无线分组业务隧道协议报文头,所述通用无线分组业务隧道协议报文头包括所述第一业务流的调度优先级。

在一种可能的设计中,所述多个业务流为多个终端设备分别对应的业务流,所述多个终端设备属于同一个终端设备组。

在一种可能的设计中,所述多个终端设备的等级不同,所述第一业务流的调度优先级是根据第一终端设备的等级确定的,所述第一业务流为所述第一终端设备签约的业务对应的业务流,所述第一终端设备为所述多个终端设备中的一个。

在一种可能的设计中,所述多个业务流属于同一个 QoS 流。

第四方面,本申请提供一种业务流调度方法,该方法包括:

终端设备获取 QoS 规则,所述 QoS 规则包括第一业务流的调度优先级,所述第一业务流为所述终端设备签约的业务,所述第一业务流的调度优先级为所述第一业务流在 QoS 流中的调度优先级;所述终端设备生成上行报文;在上行报文为属于所述第一业务流时,所述终端设备根据所述第一业务流的调度优先级对所述第一业务流进行调度。

采用上述设计,终端设备可以根据 QoS 规则中的业务流的调度优先级,对该业务流中的上行报文进行调度。

在一种可能的设计中,在终端设备获取 QoS 规则时,所述终端设备从会话管理网元接收所述 QoS 规则。

在一种可能的设计中,在终端设备获取 QoS 规则时,所述终端设备接收下行报文,所述下行报文包括所述第一业务流的调度优先级、反射 QoS 指示和所述 QoS 流的 QoS 流标识;所述终端设备根据所述第一业务流的五元组、所述第一业务流的调度优先级、所述 QoS 流的 QoS 流标识和所述反射 QoS 指示生成所述 QoS 规则。

采用上述设计,终端设备可以根据下行报文中的一业务流的调度优先级、反射 QoS 指示等信息生成 QoS 规则。

在一种可能的设计中,所述 QoS 流规则还包括所述第一业务流的带宽参数,所述第一业务流的带宽参数包括所述第一业务流的保障带宽,和/或所述第一业务流的最大带宽;或者,所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率,和/或所述第一业务流的最大带宽比率;其中,所述第一业务流的保障带宽比率为所述第一业务流所占用的最低带宽与所述 QoS 流对应的总带宽的比率,所

述第一业务流的最大带宽比率为所述第一业务流所占用的最大带宽与所述QoS流对应的总带宽的比率。

在一种可能的设计中,在所述终端设备根据所述第一业务流的调度优先级对所述第一业务流进行调度时,在所述多个业务流分别对应的需求带宽之和大于所述总带宽时,若在所述多个业务流中存在第二业务流,所述第二业务流的保障带宽大于所述第二业务流的需求带宽,则为所述第二业务流分配的带宽等于所述第二业务流的需求带宽;根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽,按照调度优先级由高到低的顺序为所述多个业务流中除所述第二业务流之外的业务流分配差值带宽,所述差值带宽为所述第二业务流的保障带宽与所述第二业务流的需求带宽的差值。

采用上述设计,可以实现对于保障带宽大于需求带宽的业务流时,保证该业务流的需求带宽,基于其他业务流的调度优先级和需求带宽,为这些业务流分配差值带宽,进而提升总带宽的利用效率。

在一种可能的设计中,在所述终端设备根据所述第一业务流的调度优先级对所述第一业务流进行调度时,在所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率时,若所述QoS流承载的至少一个业务分别对应的需求带宽之和大于所述QoS流对应的总带宽时,根据所述QoS流承载的至少一个业务中每个业务的调度优先级,按照调度优先级由高到低的顺序依次满足每个业务的保障带宽比率,所述至少一个业务包括所述第一业务流,所述至少一个业务分别对应的保障带宽比率之和小于或等于1。

采用上述设计,在各个业务流的需求带宽均大于相应的保障带宽时,优先保证各个业务流的保障带宽,基于各个业务流的调度优先级和需求带宽,为各个业务流分配剩余带宽。因此,可以实现保证各个业务流的保障带宽,且优先满足高调度优先级的业务流的需求带宽。

在一种可能的设计中,在按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽时,按照所述顺序满足所述多个业务流中部分或全部业务流的最大带宽比率。

采用上述设计,可以基于各个业务流的调度优先级、需求带宽和最大带宽比率,为各个业务流分配剩余带宽,进而可以实现优先满足高调度优先级的业务流的最大带宽比率。

在一种可能的设计中,所述多个业务流分别对应的实际带宽之和等于所述总带宽。

采用上述设计,可以提升总带宽的利用效率。

在一种可能的设计中,所述终端设备在所述上行报文中添加所述第一业务流的调度优先级,获得待发送上行报文;所述终端设备向接入网设备发送所述待发送上行报文。

采用上述设计,终端设备可以在上行报文中添加业务流的调度优先级。

在一种可能的设计中,所述待发送上行报文包括通用无线分组业务隧道协议报文头,所述通用无线分组业务隧道协议报文头包括所述第一业务流的调度优先级。

第五方面,本申请实施例提供了一种通信装置,包括用于执行以上任一方面中各个步骤的单元。

第六方面,本申请实施例提供了一种通信设备,包括至少一个处理元件和至少一个存储元件,其中该至少一个存储元件用于存储程序和数据,该至少一个处理元件用于读取并执行存储元件存储的程序和数据,以使得本申请以上任一方面提供的方法被实现。

第七方面,本申请实施例还提供了一种计算机程序,当所述计算机程序在计算机上运行时,使得所述计算机执行上述任一方面提供的方法。

第八方面,本申请实施例还提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有计算机程序,当所述计算机程序被计算机执行时,使得所述计算机执行上述任一方面提供的方法。

第九方面,本申请实施例还提供了一种芯片,所述芯片用于读取存储器中存储的计算机程序,执行上述任一方面提供的方法。

第十方面,本申请实施例还提供了一种芯片系统,该芯片系统包括处理器,用于支持计算机装置实现上述任一方面提供的方法。在一种可能的设计中,所述芯片系统还包括存储器,所述存储器用于保存该计算机装置必要的程序和数据。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。

第十一方面,提供一种通信系统,包括第三方面所述的通信装置和第四方面所述的通信装置。

本申请在上述各方面提供的实现的基础上,还可以进行进一步组合以提供更多实现。

45 附图说明

图1为本申请应用的移动通信系统的架构示意图;

图2为本申请实施例提供了一种业务流调度方法的概述流程图;

图3为本申请实施例提供了一种在5G通信系统中业务流调度的流程图;

图 4 为本申请实施例提供的另一种在 5G 通信系统中业务流调度的流程图；

图 5 为本申请实施例提供的又一种在 5G 通信系统中业务流调度的流程图；

图 6 为本申请实施例提供的再一种在 5G 通信系统中业务流调度的流程图；

图 7 为本申请实施例提供的一种通信装置的结构示意图；

5 图 8 为本申请实施例提供的另一种通信装置的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述。显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”以及相应术语标号等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换，这仅仅是描述本申请的实施例中对相关属性的对象在描述时所采用的区分方式。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，以便包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它单元。

10 在本申请的描述中，除非另有说明，“/”表示或的意思，例如，A/B 可以表示 A 或 B；本申请中的“和/或”仅仅是一种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 这三种情况。另外，在本申请的描述中，“至少一项”是指一项或者多项，“多项”是指两项或两项以上。“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-b,b-c,或 a-b-c，其中 a,b,c 可以是单个，也可以是多个。

15 为了便于理解本申请的实施例的技术方案，下面首先简要介绍相关的现有技术。

本申请实施例提供的技术方案可以应用于各种通信系统。例如：可以适用于 4G 系统或 5G 系统，也可以适用于其它面向未来的新系统等。本申请实施例对此不作具体限定。此外，术语“系统”可以和“网络”相互替换。

25 1. 5G 系统

如图 1 所示为第三代合作伙伴计划 (3rd generation partnership project, 3GPP) 标准制定的 5G 的通信系统的架构图。该通信系统包括终端设备 (例如，用户设备 (user equipment, UE))，无线接入网 (radio access network, RAN)，核心网 (Core network, CN)。数据网络 (data network, DN) 在逻辑上，核心网的网元它们可以分为用户面和控制面两部分，控制面负责移动网络的管理，用户面负责业务数据的传输。

30 其中，终端设备是移动用户与网络交互的入口，能够提供基本的计算能力，存储能力，向用户显示业务窗口，接收用户操作输入。下一代终端设备 (NextGen UE) 可以采用新空口技术，与 RAN 建立信号连接，数据连接，从而传输控制信号和业务数据到移动网络。终端设备可以包括各种具有无线通信功能的手持设备、车载设备、可穿戴设备、计算设备或连接到无线调制解调器的其它处理设备，以及各种形式的终端，移动台 (mobile station, MS)，终端 (terminal)，软终端等等，例如水表、电表、传感器等。

RAN：部署在靠近终端设备的位置，为特定区域的授权用户提供入网功能，并能够根据用户的级别，业务的需求等确定不同质量的传输隧道来传输用户数据。RAN 能够管理自身的资源，合理利用，按需为终端设备提供接入服务，并负责把控制信号和用户数据在终端设备和核心网之间转发。

40 核心网：负责维护移动网络的签约数据，管理移动网络的网元，为终端设备提供会话管理，移动性管理，策略管理，安全认证等功能。在终端设备附着的时候，为终端设备提供入网认证；在终端设备有业务请求时，为终端设备分配网络资源；在终端设备移动的时候，为终端设备更新网络资源；在终端设备空闲的时候，为终端设备提供快恢复机制；在终端设备去附着的时候，为终端设备释放网络资源；在终端设备有业务数据时，为终端设备提供数据路由功能，如转发上行数据到数据网络；或者从数据网络接收终端设备的下行数据，转发到 RAN，从而由 RAN 发送给终端设备。

45 数据网络 (data network, DN)：为用户提供业务服务的数据网络，一般客户端位于终端设备，服务端位于数据网络。数据网络可以是私有网络，如局域网，也可以是不受运营商管控的外部网络，例如互联网 (Internet)，还可以是运营商共同部署的专有网络，例如提供 IP 多媒体网络子系统 (IP multimedia

core network subsystem, IMS) 服务的网络。

5G 网络架构对下一代核心网设备的网络架构进行了重新调整, 通常控制面会采用服务化接口对外提供相关功能。控制面与用户面之间通过 N4 接口进行消息交互, 实现控制面到用户面的用户策略下发和用户面到控制面的事件上报处理。按照目前的网络架构, 策略控制网元负责用户签约的用户策略定义、
5 下发和更新功能, 会话管理网元需要负责管理用户面网元的选择, 策略下发, 事件上报、用户面网元的心跳检查、用户面网元的负荷上报等非会话级的管理功能。用户面网元负责基于会话管理网元为终端设备建立的会话上下文及其策略, 对用户的数据报文进行业务感知、规则和策略匹配、执行计费与控制策略等。

其中, 核心网用户面包括用户面功能 (user plane function, UPF); 核心网控制面包括接入和移动性
10 管理功能 (access and mobility management function, AMF), 会话管理功能 (session management function, SMF), 网络开放功能 (network exposure function, NEF), 网络功能仓储功能 (NF repository function, NRF), 统一数据管理 (unified data management, UDM), 策略控制功能 (policy control function, PCF), 应用功能 (application function, AF), 鉴服务器功能 (authentication server function, AUSF), 网络切片选择功能 (network slice selection function, NSSF)。

15 核心网控制面采用服务化架构, 控制面网元之间的交互采用服务调用的方式, 来替换传统架构中的点对点通信方式。在服务化架构中, 控制面网元会向其他控制面网元开放服务, 供其他控制面网元调用; 在点对点通信中, 控制面网元之间通信接口会存储一套特定的消息, 只能由接口两端的控制面网元在通信时使用。

以下对核心网中的部分功能实体的功能进行简单介绍:

20 1、会话管理网元: 主要用于会话管理、终端设备的 IP 地址分配和管理、选择可管理用户设备平面功能、策略控制、或收费功能接口的终结点以及下行数据通知等。在 5G 通信中, 会话管理网元可以是 SMF 网元, 在未来通信如 6G 通信中, 会话管理功能网元仍可以是 SMF 网元, 或者有其它名称, 本申请对此不作限定。Nsmf 是 SMF 提供的基于服务的接口, SMF 可以通过 Nsmf 与其他的网络功能通信。

25 2、接入管理网元: 主要用于移动性管理和接入管理等, 例如可以是 4G 通信网络中的移动性管理实体 (mobility management entity, MME) 功能或者 5G 网络中的 AMF 网元。在未来通信如 6G 通信中, 接入管理网元仍可以是 AMF 网元, 或者有其它名称, 本申请对此不作限定。Namf 是 AMF 提供的基于服务的接口, AMF 可以通过 Namf 与其他的网络功能通信。

30 3、网络开放网元: 用于安全地向外部开放由 3GPP 网络功能提供的业务和能力等。在 5G 通信中, 网络开放网元可以是 NEF 网元, 在未来通信如 6G 通信中, 网络开放功能网元仍可以是 NEF 网元, 或者有其它名称, 本申请对此不作限定。其中 Nnef 是 NEF 提供的基于服务的接口, NEF 可以通过 Nnef 与其他的网络功能通信。

35 4、网络存储网元: 用于提供服务注册、发现和授权, 并维护可用的网络功能 (network function, NF) 实例信息, 可以实现网络功能和服务的按需配置以及 NF 之间的互连。在 5G 通信中, 网络存储网元可以是 NRF 网元, 在未来通信如 6G 通信中, 网络存储功能网元仍可以是 NRF 网元, 或者有其它名称, 本申请对此不作限定。Nnrf 是 NRF 提供的基于服务的接口, NRF 可以通过 Nnrf 与其他的网络功能通信。

40 5、策略控制网元: 用于指导网络行为的统一策略框架, 为控制平面功能网元 (例如 AMF, SMF 等) 提供策略规则信息等。在 5G 通信中, 策略控制网元可以是 PCF 网元, 在未来通信如 6G 通信中, 策略控制网元仍可以是 PCF 网元, 或者有其它名称, 本申请对此不作限定。其中 Npcf 是 PCF 提供的基于服务的接口, PCF 可以通过 Npcf 与其他的网络功能通信。

6、数据管理网元: 用于处理用户标识、签约、接入鉴权、注册、或移动性管理等。在 5G 通信中, 数据管理网元可以是 UDM 网元, 在未来通信如 6G 通信中, 数据管理网元仍可以是 UDM 网元, 或者有其它名称, 本申请对此不作限定。其中 Nudm 是 UDM 提供的基于服务的接口, UDM 可以通过 Nudm 与其他的网络功能通信。

45 7、应用网元: 用于进行应用影响的数据路由, 接入网络开放功能, 或与策略框架交互进行策略控制等。在 5G 通信中, 应用网元可以是 AF 网元, 在未来通信如 6G 通信中, 应用网元仍可以是 AF 网元, 或者有其它名称, 本申请对此不作限定。Naf 是 AF 提供的基于服务的接口, AF 可以通过 Naf 与其他的网络功能通信。

8、用户面网元：用于分组路由和转发、或用户面数据的服务质量（quality of service, QoS）处理等。在 5G 通信中，用户面网元可以是用户面功能（user plane function, UPF）网元，在未来通信如 6G 通信中，用户面网元仍可以是 UPF 网元，或者有其它名称，本申请对此不作限定。

9、认证服务网元：主要用于用户鉴权等。在 5G 通信中，认证服务网元可以是 AUSF 网元，在未来通信如 6G 通信中，认证服务网元仍可以是 AUSF 网元，或者有其它名称，本申请对此不作限定。Nausf 是 AUSF 提供的基于服务的接口，AUSF 可以通过 Nausf 与其他的网络功能通信。

10、网络切片选择功能网元：用于为终端设备选择网络切片，在 5G 通信中，网络切片选择功能网元可以是 NSSF 网元，在未来通信如 6G 通信中，网络切片选择功能网元仍可以是 NSSF 网元，或者有其它名称，本申请对此不作限定。

可以理解的是，核心网还可以包括其他网元，本申请对此不作限定。

2. QoS

每一个 QoS 流由服务质量流标识（QoS flow ID, QFI）唯一识别，QFI 在每个 PDU 会话中是唯一的。

5G QoS 分为保证速率比特流（GBR Qos flow）和非保证速率比特流（Non-GBR Qos flow）两种类型，具体可以包括但不限于以下参数：5G QoS 标识（5G QoS identifier, 5QI）（对应 QoS 级别），分配保留优先级（allocation and retention priority, ARP）。

其中，保证速率比特流的专有参数可以包括但不限于以下参数：保证比特流速率、最大比特流速率、通知控制（notification control）、最大可容忍丢包率（上下行语音媒体）（maximum packet loss rate-UL/DL（voice media））等。非保证速率比特流的专有参数可以包括但不限于单个 PDU 会话总最大比特率（session-AMBR）、UE 总最大比特率（UE-AMBR）、反射型 QoS 属性（reflective QoS attribute, RQA）等。

在 5G QoS 流映射方式下，在下行数据传输过程中，UPF 根据服务数据流（service data flows, SDF）模板将数据包映射到 QoS 流，并在 N3 隧道头标记 QFI，接入网设备（access network, AN）根据 QFI 将数据包映射到数据无线承载（data radio bearer, DRB）上，并传输至 UE。在上行数据传输过程中，UE 的非接入层（non-access stratum, NAS）层根据 QoS 规则将数据包映射到 QoS 流，接入层（access stratum, AS）层负责 QoS 流到 DRB 的映射。其中，QoS 规则包含 QoS 流对应的 QFI，一个包过滤集合以及一个优先级值。AN 根据 DRB 上接收到的数据包的 QFI，在 N3 隧道头标记 QFI。UPF 接收 AN 发送的数据包，并执行验证。

在 QoS 流建立机制中，支持信令控制 QoS 机制和反射 QoS 机制（仅用于非保证速率比特流类型的 QoS 流）。

本申请实施例提供的方法可以应用于多个业务流共用一个固定带宽的场景，用于实现对多个业务流进行差异化调度。

在第一种可能的应用场景中，多个业务流为多个终端设备分别对应的业务流，多个终端设备属于同一个终端设备组。

示例性地，终端设备组包括 UE1、UE2 和 UE3，其中，UE1 签约业务 A，UE2 签约业务 B，UE3 签约业务 C，则多个业务流包括业务 A 的业务流，业务 B 的业务流，业务 C 的业务流。

此外，终端设备组中的终端设备也可以签约相同的业务，本申请对此不作限定。例如，终端设备组包括 UE1、UE2 和 UE3，其中，UE1 签约业务 A，UE2 签约业务 A，UE3 签约业务 C，则多个业务流包括与 UE1 对应的业务 A 的业务流，与 UE2 对应的业务 A 的业务流，业务 C 的业务流。

示例性地，终端设备组包括 UE1、UE2 和 UE3，其中，UE1 签约业务 A，UE2 签约业务 B，UE3 签约业务 C 和业务 D，则多个业务流包括业务 A 的业务流，业务 B 的业务流，业务 C 的业务流和业务 D 的业务流。

在第二种可能的应用场景中，多个业务流属于同一个 QoS 流，此时多个业务流为一个终端设备签约的多个业务分别对应的业务流。例如，UE1 签约业务 A 和业务 B，业务 A 的业务流与业务 B 的业务流属于同一个 QoS 流。

可以理解的是，上述两种可能的应用场景仅为举例，不作为本申请的限定。本申请中所涉及的多个业务流均是指共用一个固定带宽的多个业务流。

以下对本申请涉及的几个技术概念进行说明：

1、业务的调度优先级

针对多个业务流中的任意一个业务流，该业务流的调度优先级为该业务流在多个业务流中的调度优先级，即该业务流的调度优先级用于确定该业务流在多个业务流中的调度先后顺序。

5 示例性地，第一业务流的调度优先级为第一业务流在多个业务流中的调度优先级，第一业务流为多个业务流中的一个。

作为一种可能的设计，在第一种可能的应用场景中，多个终端设备的等级可以相同或不同。当多个终端设备的等级不同时，第一业务流的调度优先级可以根据第一终端设备的等级确定的，其中，第一业务流为第一终端设备签约的业务对应的业务流，第一终端设备为多个终端设备中的一个。

10 当第一终端设备的等级高于第二终端设备的等级时，第一业务流的调度优先级高于第二业务流的调度优先级。其中，第二终端设备为多个终端设备中的一个，第二业务流为第二终端设备签约的业务对应的业务流。

15 示例性地，终端设备组包括 UE1、UE2 和 UE3，其中，UE1 签约业务 A，UE2 签约业务 B，UE3 签约业务 C。若 UE1 的等级为等级 1、UE2 的等级为等级 1，UE3 的等级为等级 2，其中，等级 1 高于等级 2，则业务 A 的业务流的调度优先级与业务 B 的业务流的调度优先级相同，业务 A 的业务流的调度优先级高于业务 C 的业务流的调度优先级。

20 示例性地，终端设备组包括 UE1、UE2 和 UE3，其中，UE1 签约业务 A，UE2 签约业务 B，UE3 签约业务 C 和业务 D。若 UE1 为金牌用户、UE2 为银牌用户，UE3 为铜牌用户，其中，金牌用户的等级高于银牌用户的等级，银牌用户的等级高于铜牌用户的等级，则业务 A 的业务流的调度优先级高于业务 B 的业务流的调度优先级，业务 B 的业务流的调度优先级高于业务 C 的业务流的调度优先级，业务 C 的业务流的调度优先级可以与业务 D 的业务流的调度优先级相同。

作为一种可能的设计，在第二种可能的应用场景中，每个业务流的调度优先级可以根据终端设备的签约信息确定。

可以理解的是，业务流的调度优先级还可以采用其他方式确定，本申请对此不作限定。

25 2、业务流的带宽参数

以第一业务流的带宽参数为例，第一业务流的带宽参数包括第一业务流的保障带宽，和/或所述第一业务流的最大带宽，或者第一业务流的带宽参数可以包括第一业务流的保障带宽比率，和/或第一业务流的最大带宽比率。

30 其中，第一业务流的保障带宽比率为第一业务流的保障带宽与总带宽的比率，其中，保障带宽又可称为预留带宽。在第一业务流的实际带宽低于保障带宽时，第一业务流将不受到管控处理，但不代表第一业务流的实际带宽一定要高于保障带宽。

35 第一业务流的最大带宽比率为第一业务流所能占用的最大带宽与总带宽的比率。其中，第一业务流的最大带宽比率又可称为第一业务流的最大突发带宽比率，最大带宽也是指该业务在 QoS 策略中运行占用的最大带宽比例，在第一业务流的实际带宽超过最大带宽时，第一业务流超过最大带宽限制的流量将会被管控的。

需要说明的是，这里的总带宽是指为多个业务流分配的共用带宽，即一个固定带宽。总带宽又可称为管道带宽。第一业务流的保障带宽可以由总带宽和第一业务流的保障带宽比率相乘获得，第一业务流的最大带宽可以由总带宽和第一业务流的最大带宽比率相乘获得。

40 基于此，本申请提供一种业务流调度方法，如图 2 所述，该方法包括：

步骤 200：用户面网元获取第一业务流的调度优先级，第一业务流的调度优先级为第一业务流在多个业务流中的调度优先级，第一业务流为多个业务流中的一个。

其中，多个业务流以及第一业务流的调度优先级的相关描述可以参考前文中的相关内容，此处不再赘述。

45 在一种可能的实现方式中，用户面网元可以提前配置多个业务流的调度优先级，或者用户面网元可以通过从其他网元（例如，会话管理网元等）获取多个业务流的调度优先级。下述仅以用户面网元获取第一业务流的调度优先级为例进行说明。

此外，用户面网元还可以获取第一业务流的带宽参数，其中，用户面网元获取第一业务流的带宽参

数的具体方式与用户面网元获取第一业务流的调度优先级的具体方式类似。其中，第一业务流的带宽参数包括的具体内容可以参考前文中的相关内容，此处不再赘述。

示例性地，用户面网元可以从会话管理网元接收第一规则，第一规则用于指示第一业务流的调度优先级。可选地，第一规则还用于指示第一业务流的带宽参数。

5 其中，第一规则可以为报文检测规则（packet detection rule, PDR）。

示例性地，会话管理网元从策略控制网元接收策略与计费控制（policy and charging control, PCC）规则，策略与计费控制规则用于指示第一业务流的调度优先级。会话管理网元可以根据策略与计费控制规则确定第一规则，并向用户面网元发送第一规则。

10 在一示例中，策略与计费控制规则包括第一预定义规则的名称。会话管理网元可以根据策略与计费控制规则确定第一规则，第一规则包括第一预定义规则的名称。

其中，第一预定义规则包括第一业务流的调度优先级，第一预定义规则为会话管理网元和用户面网元存储的多个预定义规则中的一个。可选地，第一预定义规则还包括第一业务流的带宽参数。具体可以参考下述如图4和图5所示的实施例。

15 需要说明的是，会话管理网元保存的第一预定义规则的一部分内容（以下简称第一部分内容），用户面网元保存的第一预定义规则的另一部分内容（以下简称第二部分内容），两部分内容共同构成第一预定义规则，第一部分内容和第二部分内容可以存在重叠的内容，同时也存在不重叠的内容。

此外，在第二种可能的应用场景中，第一预定义规则还可以包括反射 QoS 指示（reflective QoS indication, RQI）。若第一预定义规则不包括反射 QoS 指示，会话管理网元还可以根据策略与计费控制规则确定 QoS 规则，QoS 规则包括第一业务流的调度优先级。会话管理网元向终端设备发送 QoS 规则。
20 可选地，若第一预定义规则还包括第一业务流的带宽参数，则 QoS 规则还包括第一业务流的带宽参数。

在另一示例中，在第二种可能的应用场景中，会话管理网元可以根据策略与计费控制规则确定第一规则，第一规则包括 QoS 执行规则（QoS enforcement rule, QER）的标识，QoS 执行规则包括第一业务流的调度优先级，也即会话管理网元可以将第一业务流的调度优先级扩展到第一规则引用的 QoS 执行规则中。可选地，若策略与计费控制规则还包括第一业务流的带宽参数，则 QoS 执行规则还可以包括第一业务流的带宽参数。具体可以参考下述如图3所示的实施例。
25

此外，策略与计费控制规则还可以包括反射 QoS 指示，则第一规则还可以包括反射 QoS 指示。若策略与计费控制规则不包括反射 QoS 指示，会话管理网元还可以根据策略与计费控制规则确定 QoS 规则，QoS 规则包括第一业务流的调度优先级。会话管理网元向终端设备发送 QoS 规则。可选地，若策略与计费控制规则还包括第一业务流的带宽参数，则 QoS 规则还可以包括第一业务流的带宽参数。
30

步骤 210：用户面网元接收报文。

示例性地，这里的报文可以为上行报文或下行报文，本申请对此不作限定。

示例性地，若报文为下行报文，用户面网元可以在下行报文中添加第一业务流的调度优先级，获得第一待发送下行报文，用户面网元向接入网设备发送第一待发送下行报文。

35 其中，第一待发送下行报文包括通用无线分组业务隧道协议（GPRS tunneling protocol, GTP）报文头，通用无线分组业务隧道协议报文头包括第一业务流的调度优先级。

在一种可能的设计中，在第二种可能的应用场景中，用户面网元除了可以在下行报文中添加第一业务流的调度优先级，用户面网元还可以在下行报文中添加反射 QoS 指示和 QoS 流的 QoS 流标识，获得第二待发送下行报文，用户面网元向接入网设备发送第二待发送下行报文。其中，反射 QoS 指示和 QoS 流的 QoS 流标识也可以由通用无线分组业务隧道协议报文头携带。

40 步骤 220：在报文属于第一业务流时，用户面网元根据第一业务流的调度优先级对第一业务流进行调度。

示例性地，用户面网元可以根据报文检测规则中的报文检测信息（packet detection information, PDI）确定检测到的报文属于第一业务流。

45 示例性地，用户面网元在根据第一业务流的调度优先级对第一业务流进行调度时，可以采用但不限于如下规则：

首先，用户面网元可以判断多个业务流分别对应的需求带宽之和是否大于总带宽，若多个业务流分别对应的需求带宽之和小于或等于总带宽，则用户面网元可以保证多个业务流分别对应的需求带宽。

若多个业务流分别对应的需求带宽之和大于总带宽，即总带宽不够用，用户面网元可以采用如下规

则对多个业务流进行调度：

规则 (1)：

用户面网元根据多个业务流的调度优先级由高至低的顺序，对多个业务流进行调度，使得多个业务流分别对应的实际带宽之和等于总带宽，也即，多个业务流分别对应的实际带宽之和尽量用满总带宽，又可描述为多个业务流分别对应的实际带宽之和接近总带宽。

规则 (2)：

若在多个业务流中存在第二业务流，第二业务流的保障带宽大于第二业务流的需求带宽，则为第二业务流分配的带宽等于第二业务流的需求带宽，进一步地，根据多个业务流分别对应的调度优先级和多个业务流分别对应的需求带宽，按照调度优先级由高到低的顺序为多个业务流中除第二业务流之外的业务流分配差值带宽，差值带宽为第二业务流的保障带宽与第二业务流的需求带宽的差值。

采用上述规则，可以实现对于保障带宽大于需求带宽的业务流时，保证该业务流的需求带宽，基于其他业务流的调度优先级和需求带宽，为这些业务流分配差值带宽，进而提升总带宽的利用效率。

规则 (3)：

用户面网元优先满足多个业务流分别对应的保障带宽比率，其中，多个业务流分别对应的保障带宽比率之和小于或等于 1。在满足多个业务流分别对应的保障带宽比率之后，根据多个业务流分别对应的调度优先级和多个业务流分别对应的需求带宽，按照调度优先级由高到低的顺序为多个业务流中的部分或全部业务流分配剩余带宽，以实现每个业务流的实际带宽与总带宽的比值大于或等于该业务流的保障带宽比率。

在一种可能的实现方式中，在按照调度优先级由高到低的顺序为多个业务流中的部分或全部业务流分配剩余带宽时，按照该顺序满足多个业务流中部分或全部业务流的需求带宽，以实现调度优先级较高的业务流的需求带宽被优先满足。

在一种可能的实现方式中，在按照调度优先级由高到低的顺序为多个业务流中的部分或全部业务流分配剩余带宽时，按照该顺序满足多个业务流中部分或全部业务流的最大带宽比率，以实现调度优先级较高的业务流的最大带宽比率被优先满足。

示例性地，在按照调度优先级由高到低的顺序为多个业务流中的部分或全部业务流分配剩余带宽时，若剩余带宽小于调度优先级最高的业务流的需求带宽与调度优先级最高的业务流的保障带宽之差，则将剩余带宽全部分配给调度优先级最高的业务流。

若剩余带宽大于调度优先级最高的业务流的需求带宽与调度优先级最高的业务流的保障带宽之差，按照该顺序满足多个业务流中部分或全部业务流的需求带宽。

若剩余带宽大于调度优先级最高的业务流的需求带宽与调度优先级最高的业务流的保障带宽之差，且剩余带宽小于或等于调度优先级最高的业务流的最大带宽与调度优先级最高的业务流的保障带宽之差，则将剩余带宽全部分配给调度优先级最高的业务流。

若剩余带宽大于调度优先级最高的业务流的最大带宽与调度优先级最高的业务流的保障带宽之差，则按照该顺序满足多个业务流中部分或全部业务流的最大带宽比率。

例如，假设总带宽为 10G，多个业务流包括业务流 A 和业务流 B，业务流 A 的调度优先级低于业务流 B 的调度优先级。业务流 A 的带宽参数包括业务流 A 的保障带宽比率为 30%，业务流 A 的最大带宽比率为 50%，业务流 B 的带宽参数包括业务流 B 的保障带宽比率为 40%，业务流 B 的最大带宽比率为 60%。若业务流 A 的需求带宽为 5G，业务流 B 的需求带宽为 6G，则业务流 A 的需求带宽与业务流 B 的需求带宽之和为 13G，大于 10G，优先满足业务流 A 的保障带宽比率和业务流 B 的保障带宽比率，则业务流 A 分得带宽为 3G，业务流 B 分得带宽为 4G，此时总带宽剩余 3G 未分配，可以优先分配给业务流 B，由于业务流 B 的需求带宽为 6G，最终，业务流 A 的实际带宽为 4G，业务流 B 的实际带宽为 6G。此时，可以实现优先满足业务流 B 的需求带宽。

又例如，假设总带宽为 10G，多个业务流包括业务流 A、业务流 B 和业务流 C，按照调度优先级由高到低的顺序依次为业务流 B，业务流 C 和业务流 A。业务流 A 的带宽参数包括业务流 A 的保障带宽比率为 30%，业务流 A 的最大带宽比率为 50%，业务流 B 的带宽参数包括业务流 B 的保障带宽比率为 40%，业务流 B 的最大带宽比率为 60%。业务流 C 的带宽参数包括业务流 C 的保障带宽比率为 10%，业务流 C 的最大带宽比率为 30%。若业务流 A 的需求带宽为 5G，业务流 B 的需求带宽 7G，业务流 C 的需求带宽 3G，则业务流 A 的需求带宽、业务流 B 的需求带宽以及业务流 C 的需求带宽之和为 14G，

大于 10G, 优先满足业务流 A 的保障带宽比率、业务流 B 的保障带宽比率和业务流 C 的保障带宽比率, 则业务流 A 分得带宽为 3G, 业务流 B 分得带宽为 4G, 业务流 C 分得带宽为 1G, 此时总带宽剩余 2G 未分配, 可以优先分配给业务流 B, 最终, 业务流 A 的实际带宽为 3G, 业务流 B 的实际带宽为 6G, 业务流 C 的实际带宽为 1G。

5 需要说明的是, 基于各个业务流的需求带宽, 保障带宽和最大带宽以及总带宽, 可以设计多种规则, 上述规则和举例不作为本申请实施例的限定。

可以理解的是, 上述图 2 所示的实施例还可以应用于第 4 代移动通信技术 (the 4th generation mobile communication technology, 4G) 通信系统, 具体的, 用户面网元可以替换为数据网关控制面 (PGW user plane, PGW-U), QoS 流可以替换为专用承载, 会话管理网元可以替换为数据网关控制面 (PGW control plane, PGW-C), 策略控制网元可以替换为策略与计费规则功能单元 (policy and charging rules function, PCRF)。

以下结合具体实施例对图 2 所示实施例进行说明:

15 如图 3 所示为在 5G 通信系统中业务流调度的具体流程之一。图 3 所示实施例应用于第二种可能的应用场景。其中, 策略控制网元为 PCF, 会话管理网元为 SMF, 用户面网元为 UPF, 终端设备为 UE, 第一规则为 PDR。该流程可以包括以下步骤:

S301: UE 向 AMF 发送 PDU 会话建立请求 (PDU Session Establishment Request) 消息。

20 示例性地, UE 在 5G 网络中激活, 发起 PDU 会话建立流程。其中, PDU 会话建立请求消息包括 UE 的标识, 例如, UE 的标识可以为国际移动用户识别码 (international mobile subscriber identification number, IMSI), 移动台国际用户识别码 (mobile subscriber international ISDN number, MSISDN) 等。本申请对此不作限定。

S302: AMF 向 SMF 发送 PDU 会话创建会话管理上下文请求 (Nsmf_PDUSession_Creat-SMContext Request) 消息。

其中, PDU 会话创建会话管理上下文请求消息包括 UE 的标识。

S303: SMF 向 PCF 发送请求消息, 该请求消息用于请求 PCC 规则。

25 示例性地, SMF 通过 N7 接口向 PCF 发送会话管理策略控制创建请求 (Npcf_SMPolicy-Control_Create Request) 消息。其中, 会话管理策略控制创建请求消息包括 UE 的标识。

S304: PCF 向 SMF 发送 PCC 规则。

30 示例性地, PCF 根据 UE 的标识查询 UE 的签约规则, UE 的签约规则包括第一业务流的调度优先级。可选的, UE 的签约规则还可以包括第一业务流的带宽参数, 第一业务流的带宽参数包括第一业务流的保障带宽比率, 和/或第一业务流的最大带宽比率。

其中, UE 的签约规则又可称为 UE 的签约信息。可以理解的是, UE 的签约规则包括多个业务流的调度优先级。多个业务流包括第一业务流, 这里仅以第一业务流为例进行说明。对于不同的业务流, PCF 可以向 SMF 发送的不同的 PCC 规则。

35 示例性地, PCF 通过 N7 接口向 SMF 发送会话管理策略控制创建响应 (Npcf_SMPolicy-Control_Create Response) 消息, 该消息包括 PCC 规则。其中, PCC 规则包括第一业务流的调度优先级。可选的, PCC 规则还包括第一业务流的带宽参数。

S305: SMF 将该 PCC 规则映射为 PDR。

40 示例性地, PDR 包括 QER ID, QER 包括第一业务流的调度优先级, 可选的, QER 还包括第一业务流的带宽参数。

此外, 在一种可能的实现方式中, QER 中的第一业务流的带宽参数, 还可以替换为第一业务流的保障带宽, 和/或第一业务流的最大带宽。也即在第一业务流的带宽参数包括第一业务流的保障带宽比率时, 根据 QoS 流对应的总带宽和第一业务流的保障带宽比率计算得到第一业务流的保障带宽。在第一业务流的带宽参数包括第一业务流的最大带宽比率时, 根据 QoS 流对应的总带宽和第一业务流的最大带宽比率计算得到第一业务流的最大带宽。

S306: SMF 向 UPF 发送 N4 会话建立请求 (N4 Session Establishment/Modification Request) 消息, 该消息包括上述 S305 中确定的 PDR。

S307: UPF 向 SMF 发送 N4 会话建立响应 (N4 Session Establishment/Modification Response) 消息。

UPF 可以安装接收到的 PDR，用于报文的匹配，以及确定第一业务流的 QFI 和第一业务流调度优先级。

此外，PDR 中还包括 PDI，PDI 包括第一业务流的应用标识。UPF 可以根据第一业务流的应用标识作为业务感知的条件，记录在该 PDR 的上下文中，启动对第一业务流的识别。

5 可以理解的是，S306 中的 N4 会话建立请求消息，还可以替换为 N4 会话修改请求（N4 Session Modification Request）消息，相应的，S307 中的 N4 会话建立响应消息，还可以替换为 N4 会话修改响应（N4 Session Establishment/Modification Response）消息。

10 S308：SMF 向 AMF 发送 N1N2 消息传输消息（Namf_Communication_N1N2MessageTransfer），该消息包括 N1 会话管理容器（SM container），N1 会话管理容器包括 PDU 会话建立接受（PDU Session Establishment Accept）消息。PDU 会话建立接受消息包括 QoS 规则。

示例性地，SMF 根据 PCC 规则和第一业务流的五元组信息生成 QoS 规则，QoS 规则包括第一业务流的调度优先级。可选的，QoS 规则还包括第一业务流的带宽参数。五元组信息包括源地址、目标地址、源端口、目标端口、协议号。

S309：AMF 通过 RAN 向 UE 发送 PDU 会话建立接受消息。

15 UE 可以安装接收到的 QoS 规则，用于报文的匹配，确定第一业务流的 QFI 和第一业务流调度优先级。

可以理解的是，以上流程中还可以替换为对于会话更新或修改流程。例如，当 UE 签约新业务时，可以通过会话更新流程或会话修改流程使得 UPF 和 UE 获得新业务的业务流的调度优先级。

S310：UPF 确定检测到属于第一业务流的下行报文。

20 示例性地，PDR 中还包括 PDI，PDI 可以包括 IP 地址、端口（Port）等 L3 层信息和 L4 层信息，也可以是应用标识(application ID)或统一资源定位系统（uniform resource locator;URL）等 L7 层信息。UPF 可以根据 PDI 判断接收到的报文是否属于第一业务流。

其中，应用标识可以映射到 UPF 上预定义的多个应用过滤条件，比如应用名称、IP 地址、端口等信息用于后续匹配流程。

25 例如，当 PDI 包括第一业务流的应用标识时，UPF 可以根据 PDR 判断下行报文中的应用标识是否与第一业务流的应用标识匹配，若匹配，则确定检测到属于第一业务流的下行报文。

S311：UPF 根据第一业务流的调度优先级对下行报文进行调度。

30 示例性地，UPF 根据 PDR 中包括的 QER 的标识，执行该 QER 所定义的 QoS 策略。UPF 基于第一业务流所在的 QoS 流包括的至少一个业务流和 QER 包括的第一业务流的调度优先级对下行报文进行调度。可选地，若 QER 还包括第一业务流的带宽参数，则 UPF 还基于 QER 包括的第一业务流的带宽参数对下行报文进行调度。

具体调度方式可以参考上述步骤 220 中的相关描述。

S312：UPF 向 RAN 发送下行报文。

35 示例性地，UPF 在下行报文中添加第一业务流的调度优先级，将包括第一业务流的调度优先级的下行报文发送至 RAN。例如，UPF 在下行报文的 GTP 头中添加第一业务流的调度优先级。此外，GTP 头还可以包括 QFI。此外，可选地，UPF 还可以在下行报文中添加第一业务流的带宽参数。

S313：RAN 根据第一业务流的调度优先级对接收到的下行报文进行调度。

40 示例性地，RAN 可以根据 GTP 头中的 QFI 和第一业务流的调度优先级，确定 QFI 所指示的 QoS 流包括的至少一个业务流，进而根据 QFI 所指示的 QoS 流包括的至少一个业务流和第一业务流的调度优先级对下行报文进行调度。此外，可选地，若下行报文中还包括第一业务流的带宽参数，则 RAN 根据第一业务流的带宽参数对下行报文进行调度。

具体调度方式可以参考上述步骤 220 中的相关描述。

S314：RAN 向 UE 发送下行报文。

示例性地，RAN 也可将第一业务流的调度优先级通过空口扩展携带到 UE。

45 示例性地，RAN 将接收到的来自于核心网（即 UPF）的报文进行解封装，获取原始报文内容和 GTP 报文头携带的 QFI 和第一业务流的调度优先级，然后再通过空口的封装格式将原始报文内容和第一业务流的调度优先级重新封装，并将重新封装后的报文发送至 UE。

S315：UE 生成属于第一业务流的上行报文，根据第一业务流的调度优先级对该上行报文进行调度。

示例性地，UE 生成上行报文，确定该上行报文属于第一业务流，以及第一业务流所在的 QoS 流，进而匹配到相应的 QoS 规则，根据 QoS 规则确定第一业务流的调度优先级，根据该 QoS 流包括的至少一个业务流和第一业务流的调度优先级对该上行报文进行调度。此外，可选地，若 QoS 规则还包括第一业务流的带宽参数，则 UE 还根据第一业务流的带宽参数对下行报文进行调度。

5 具体调度方式可以参考上述步骤 220 中的相关描述。

S316:UE 向 RAN 发送上行报文。

示例性地，UE 可以在上行报文中添加第一业务流的调度优先级，将包括第一业务流的调度优先级的上行报文发送至 RAN。示例性地，UE 可以将报文内容和第一业务流的调度优先级通过空口的封装格式进行封装，并将封装后的报文发送至 RAN。此外，UE 还可以在上行报文中添加 QFI，可选地，UE 还可以在上行报文中添加第一业务流的带宽参数。

10

S317: RAN 根据第一业务流的调度优先级对接收到的上行报文进行调度。

示例性地，RAN 可以根据接收到的上行报文中的 QFI 和第一业务流的调度优先级，确定 QFI 所指示的 QoS 流，以及该 QoS 流包括的至少一个业务流，进而根据该 QoS 流包括的至少一个业务流和第一业务流的调度优先级对接收到的上行报文进行调度。此外，可选地，若上行报文中还包括第一业务流的带宽参数，则 RAN 还根据第一业务流的带宽参数对上行报文进行调度。

15

具体调度方式可以参考上述步骤 220 中的相关描述。

S318: RAN 向 UPF 发送上行报文。

示例性地，RAN 可以在上行报文中添加第一业务流的调度优先级，将包括第一业务流的调度优先级的上行报文发送至 UPF。例如，包括第一业务流的调度优先级的上行报文包括 GTP 头，GTP 头包括第一业务流的调度优先级。此外，GTP 头还可以包括 QFI，可选地，RAN 可以在上行报文中添加第一业务流的带宽参数。

20

S319: UPF 根据第一业务流的调度优先级对上行报文进行调度。

示例性地，UPF 确定接收到的报文中的 QFI 匹配之前接收到的 PDR，进一步地，UPF 根据 PDR 中包括的 QER 的标识，执行该 QER 所定义的 QoS 策略。UPF 基于第一业务流所在的 QoS 流包括的至少一个业务流和 QER 包括的第一业务流的调度优先级对上行报文进行调度。可选地，若 QER 还包括第一业务流的带宽参数，则 UPF 还基于第一业务流的带宽参数对上行报文进行调度。

25

具体调度方式可以参考上述步骤 220 中的相关描述。

采用上述设计，可以实现对一个 QoS 流中的不同业务流的调度优先级进行区分并根据不同业务流的调度优先级进行调度，同时还可以保障调度优先级低的业务流的保障带宽以及调度优先级高的业务流的最大带宽。

30

如图 4 所示为在 5G 通信系统中业务流调度的具体流程之二。其中，图 4 所示实施例应用于第二种可能的应用场景。策略控制网元为 PCF，会话管理网元为 SMF，用户面网元为 UPF，终端设备为 UE，第一规则为 PDR。该流程可以包括以下步骤：

35

S401: UE 向 AMF 发送 PDU 会话建立请求消息。

S402: AMF 向 SMF 发送 PDU 会话创建会话管理上下文请求消息。

S403: SMF 向 PCF 发送请求消息，该请求消息用于请求 PCC 规则。

其中，S401 至 S403 可以参考上述图 3 所示实施例中的 S301 至 S303 的描述，不再赘述。

S404: PCF 向 SMF 发送 PCC 规则，PCC 规则包括第一预定义规则的名称。

40

S405: SMF 根据第一预定义规则的名称生成 PDR。

其中，PDR 包括第一预定义规则的名称。

S406: SMF 向 UPF 发送 N4 会话建立请求消息，该消息包括 PDR。

S407: UPF 根据第一预定义规则的名称确定第一预定义规则。

45

第一预定义规则包括第一业务流的调度优先级，可选的，第一预定义规则还包括第一业务流的带宽参数。

UPF 可以安装第一预定义规则，用于报文的匹配，以及确定第一业务流的 QFI 和第一业务流调度优先级。

此外，PDR 中还包括 PDI，PDI 包括第一业务流的应用标识。UPF 可以根据第一业务流的应用标识

作为业务感知的条件，记录在该 PDR 的上下文中，启动对第一业务流的识别。

S408: UPF 向 SMF 发送 N4 会话建立响应消息。

此外，S406 中的 N4 会话建立请求消息，还可以替换为 N4 会话修改请求消息，相应的，S408 中的 N4 会话建立响应消息，还可以替换为 N4 会话修改响应消息。

5 S409: SMF 向 AMF 发送 N1N2 消息传输消息，该消息包括 N1 会话管理容器，N1 会话管理容器包括 PDU 会话建立接受消息。PDU 会话建立接受消息包括 QoS 规则。

示例性地，SMF 根据第一预定义规则和第一业务流的五元组信息生成 QoS 规则，QoS 规则包括第一业务流的调度优先级。可选的，QoS 规则还包括第一业务流的带宽参数。

S410: AMF 通过 RAN 向 UE 发送 PDU 会话建立接受消息。

10 UE 可以安装接收到的 QoS 规则，用于报文的匹配，确定第一业务流的 QFI 和第一业务流调度优先级。

可以理解的是，以上流程中还可以替换为对于会话更新或修改流程。

S411: UPF 确定检测到属于第一业务流的下行报文。

S412: UPF 根据第一业务流的调度优先级对下行报文进行调度。

15 示例性地，UPF 基于第一业务流所在的 QoS 流包括的至少一个业务流和第一预定义规则包括的第一业务流的调度优先级对下行报文进行调度。可选地，若第一预定义规则还包括第一业务流的带宽参数，则 UPF 还基于第一业务流的带宽参数对下行报文进行调度。

具体调度方式可以参考上述步骤 220 中的相关描述。

S413: UPF 向 RAN 发送下行报文。

20 S414: RAN 根据第一业务流的调度优先级对接收到的下行报文进行调度。

S415: RAN 向 UE 发送下行报文。

S416: UE 生成属于第一业务流的上行报文，根据第一业务流的调度优先级对该上行报文进行调度。

S417: UE 向 RAN 发送上行报文。

S418: RAN 根据第一业务流的调度优先级对接收到的上行报文进行调度。

25 S419: RAN 向 UPF 发送上行报文。

其中，S413 至 S419 可以参考上述图 3 所示实施例中的 S312 至 S318 的描述，不再赘述。

S420: UPF 根据第一业务流的调度优先级对上行报文进行调度。

30 示例性地，UPF 确定接收到的报文中的 QFI 匹配之前接收到的 PDR，进一步地，UPF 基于第一业务流所在的 QoS 流包括的至少一个业务流和第一预定义规则包括的第一业务流的调度优先级对上行报文进行调度。可选地，若第一预定义规则还包括第一业务流的带宽参数，则 UPF 还基于第一业务流的带宽参数对上行报文进行调度。

具体调度方式可以参考上述步骤 220 中的相关描述。

35 采用上述设计，可以实现对一个 QoS 流中的不同业务流的调度优先级进行区分并根据不同业务流的调度优先级进行调度，同时还可以保障调度优先级低的业务流的保障带宽以及调度优先级高的业务流的最大突发带宽。

如图 5 所示为在 5G 通信系统中业务流调度的具体流程之三。其中，图 5 所示实施例应用于第二种可能的应用场景。策略控制网元为 PCF，会话管理网元为 SMF，用户面网元为 UPF，终端设备为 UE，第一规则为 PDR。

40 下述流程仅以 PCC 规则包括 RQI 为例进行说明，此外，在一种可能的实现方式中，PCC 规则还可包括预定义规则的名称，该预定义规则包括 RQI。

该流程可以包括以下步骤：

S501: UE 向 AMF 发送 PDU 会话建立请求消息。

S502: AMF 向 SMF 发送 PDU 会话创建会话管理上下文请求消息。

45 S503: SMF 向 PCF 发送请求消息，该请求消息用于请求 PCC 规则。

其中，S501 至 S503 可以参考上述图 3 所示实施例中的 S301 至 S303 的描述，不再赘述。

S504: PCF 向 SMF 发送 PCC 规则。

示例性地，PCF 通过 N7 接口向 SMF 发送会话管理策略控制创建响应

(Npcf_SMPolicy-Control_Create Response) 消息, 该消息包括 PCC 规则。其中, PCC 规则包括第一业务流的调度优先级和 RQI。可选的, PCC 规则还包括第一业务流的带宽参数。

S505: SMF 将该 PCC 规则映射为 PDR。

5 示例性地, PDR 包括 QER ID, QER 包括第一业务流的调度优先级和 RQI, 可选的, QER 还包括第一业务流的带宽参数。

S506: SMF 向 UPF 发送 N4 会话建立请求消息, 该消息包括上述 S505 中确定的 PDR。

S507: UPF 向 SMF 发送 N4 会话建立响应消息。

UPF 可以安装接收到的 PDR, 用于报文的匹配, 以及确定第一业务流的 QFI 和第一业务流调度优先级。

10 此外, PDR 中还包括 PDI, PDI 包括第一业务流的应用标识。UPF 可以根据第一业务流的应用标识作为业务感知的条件, 记录在该 PDR 的上下文中, 启动对第一业务流的识别。

此外, S506 中的 N4 会话建立请求消息, 还可以替换为 N4 会话修改请求消息, 相应的, S507 中的 N4 会话建立响应消息, 还可以替换为 N4 会话修改响应消息。

S508: UPF 确定检测到属于第一业务流的下行报文。

15 S509: UPF 根据第一业务流的调度优先级调度对下行报文进行调度。

S510: UPF 向 RAN 发送下行报文。

20 示例性地, UPF 在下行报文中添加第一业务流的调度优先级和 RQI, 将包括第一业务流的调度优先级和 RQI 的下行报文发送至 RAN。例如, UPF 在下行报文的 GTP 头中添加第一业务流的调度优先级和 RQI。此外, GTP 头还可以包括 QFI。此外, 可选地, UPF 还可以在下行报文中添加第一业务流的带宽参数。

S511: RAN 根据第一业务流的调度优先级对接收到的下行报文进行调度。

25 示例性地, RAN 可以根据 GTP 头中的 QFI 和第一业务流的调度优先级, 确定 QFI 所指示的 QoS 流包括的至少一个业务流, 进而根据该 QoS 流包括的至少一个业务流和第一业务流的调度优先级对下行报文进行调度。此外, 可选地, 若下行报文中还包括第一业务流的带宽参数, 则 RAN 还根据第一业务流的带宽参数对下行报文进行调度。

具体调度方式可以参考上述步骤 220 中的相关描述。

S512: RAN 向 UE 发送下行报文。

示例性地, RAN 也可将第一业务流的调度优先级通过空口扩展携带到 UE。

S513: UE 基于下行报文携带的 RQI 和其他信息生成 QoS 规则。

30 示例性地, UE 基于下行报文中携带的 QFI、第一业务流的调度优先级、RQI 以及第一业务流五元组 (包括源 IP+端口、目标 IP+端口、L4 协议 (Protocol)) 生成上行方向的 QoS 规则。

此外, 可选地, 若下行报文还包括第一业务流的带宽参数, 则 UE 生成的 QoS 规则还包括第一业务流的带宽参数。

S514: UE 生成属于第一业务流的上行报文, 根据第一业务流的调度优先级对该上行报文进行调度。

35 S515: UE 向 RAN 发送上行报文。

S516: RAN 根据第一业务流的调度优先级对接收到的上行报文进行调度。

S517: RAN 向 UPF 发送上行报文。

S518: UPF 根据第一业务流的调度优先级对上行报文进行调度。

其中, S514 至 S518 可以参考上述图 3 所示实施例中的 S315 至 S319 的描述, 不再赘述。

40 采用上述设计, 可以实现对一个 QoS 流中的不同业务流的调度优先级进行区分并根据不同业务流的调度优先级进行调度, 同时还可以保障调度优先级低的业务流的保障带宽以及调度优先级高的业务流的最大带宽。此外, 采用上述设计, 还可以实现减少从 SMF 到 UE 之间传输 QoS 规则所需的信令消耗, 快速实现 QoS 规则的生成。

45 如图 6 所示为在 5G 通信系统中业务流调度的具体流程之四。其中, 图 6 所示实施例应用于第一种可能的应用场景。策略控制网元为 PCF, 会话管理网元为 SMF, 用户面网元为 UPF, 终端设备为 UE, 第一规则为 PDR。该流程可以包括以下步骤:

S601: UE 向 AMF 发送 PDU 会话建立请求消息。

S602: AMF 向 SMF 发送 PDU 会话创建会话管理上下文请求消息。

S603: SMF 向 PCF 发送请求消息, 该请求消息用于请求 PCC 规则。

其中, S601 至 S603 可以参考上述图 3 所示实施例中的 S301 至 S303 的描述, 不再赘述。

S604: PCF 向 SMF 发送 PCC 规则, PCC 规则包括第一预定义规则的名称。

5 S605: SMF 根据第一预定义规则的名称生成 PDR。

其中, PDR 包括第一预定义规则的名称。

S606: SMF 向 UPF 发送 N4 会话建立请求消息, 该消息包括 PDR。

S607: UPF 根据第一预定义规则的名称确定第一预定义规则。

10 第一预定义规则包括第一业务流的调度优先级, 可选的, 第一预定义规则还包括第一业务流的带宽参数。

UPF 可以安装第一预定义规则。示例性地, 第一预定义规则也可称为组级带宽管控策略。其中, 该组级带宽管控策略为基于某个条件构成的终端设备组的业务整体的带宽控制策略, 该终端设备组中每个终端设备对应的至少一个业务流所构成的多个业务流共用一个固定带宽, 也即终端设备组对应的多个业务流共用一个固定带宽。

15 示例性地, 终端设备组可以基于签约信息, 接入点 (access point name, APN)、接入网类型、漫游属性、位置信息等确定。

例如, 通过 APN: xxx 接入的终端设备构成一个终端设备组, 在该终端设备组中的终端设备均访问业务 A 时, 该终端设备组中每个终端设备对应的业务 A 的业务流所构成的多个业务流允许访问的带宽为 5Gpbs, 即多个业务流共用带宽的带宽为 5Gpbs。

20 S608: UPF 向 SMF 发送 N4 会话建立响应消息。

此外, S606 中的 N4 会话建立请求消息, 还可以替换为 N4 会话修改请求消息, 相应的, S608 中的 N4 会话建立响应消息, 还可以替换为 N4 会话修改响应消息。

S609: UPF 确定检测到属于第一业务流的下行报文。

S610: UPF 根据第一业务流的调度优先级对下行报文进行调度。

25 示例性地, UPF 基于终端设备组对应的多个业务流和第一预定义规则包括的第一业务流的调度优先级对下行报文进行调度。可选地, 若第一预定义规则还包括第一业务流的带宽参数, 则 UPF 还基于第一业务流的带宽参数对下行报文进行调度。

具体调度方式可以参考上述步骤 220 中的相关描述。

S611: UPF 向 RAN 发送下行报文。

30 示例性地, UPF 在下行报文中添加第一业务流的调度优先级, 将包括第一业务流的调度优先级的下行报文发送至 RAN。例如, UPF 在下行报文的 GTP 头中添加第一业务流的调度优先级。此外, 可选地, UPF 还可以在下行报文中添加第一业务流的带宽参数。

S612: RAN 根据第一业务流的调度优先级对接收到的下行报文进行调度。

35 示例性地, RAN 可以根据终端设备组对应的多个业务流和第一业务流的调度优先级对下行报文进行调度。此外, 可选地, 若下行报文中还包括第一业务流的带宽参数, 则 RAN 还根据第一业务流的带宽参数对下行报文进行调度。

具体调度方式可以参考上述步骤 220 中的相关描述。

S613: RAN 向 UE 发送下行报文。

示例性地, RAN 也可将第一业务流的调度优先级通过空口扩展携带到 UE。

40 示例性地, RAN 将接收到的来自于核心网 (即 UPF) 的报文进行解封装, 获取原始报文内容和 GTP 报文头携带的第一业务流的调度优先级, 然后再通过空口的封装格式将原始报文内容和第一业务流的调度优先级重新封装, 并将重新封装后的报文发送至 UE。

45 采用上述设计, 可以实现对一个终端设备组对应的多个业务流中的不同业务流的调度优先级进行区分并根据不同业务流的调度优先级进行调度, 同时还可以保障调度优先级低的业务流的保障带宽以及调度优先级高的业务流的最大带宽。

图 7 示出了本申请实施例中所涉及的一种通信装置的可能的示例性框图, 该装置 700 包括: 收发模块 720 和处理模块 710, 收发模块 720 可以包括接收单元和发送单元。处理模块 710 用于对装置 700 的

动作进行控制管理。收发模块 720 用于支持装置 700 与其他网络实体的通信。可选地，装置 700 还可以包括存储单元，所述存储单元用于存储装置 700 的程序代码和数据。

可选地，所述装置 700 中各个模块可以是通过软件来实现。

5 可选地，处理模块 710 可以是处理器或控制器，例如可以是通用中央处理器(central processing unit, CPU)，通用处理器，数字信号处理(digital signal processing, DSP)，专用集成电路(application specific integrated circuits, ASIC)，现场可编程门阵列(field programmable gate array, FPGA)或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本申请实施例公开内容所描述的各种示例性的逻辑方框，模块和电路。所述处理器也可以是实现计算功能的组合，例如包含一个或多个微处理器组合，DSP 和微处理器的组合等等。收发模块 720 可以是通信接口、收发器或收发电路等，其中，该通信接口是统称，在具体实现中，该通信接口可以包括多个接口，存储单元可以是存储器。

当装置 700 为用户面网元或用户面网元中的芯片时，装置 700 中的处理模块 710 可以支持装置 700 执行上文中各方法示例中用户面网元的动作。

收发模块 720 可以支持装置 700 与会话管理网元进行通信。

15 收发模块 720，用于获取第一业务流的调度优先级，所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级；所述第一业务流为所述多个业务流中的一个；接收报文；

处理模块 710，用于在所述报文属于所述第一业务流时，根据所述第一业务流的调度优先级对所述第一业务流进行调度。

20 在一种可能的设计中，收发模块 720，用于在获取第一业务流的调度优先级时，从会话管理网元接收第一规则，所述第一规则用于指示所述第一业务流的调度优先级。

在一种可能的设计中，所述第一规则包括第一预定义规则的名称，所述第一预定义规则包括所述第一业务流的调度优先级，所述第一预定义规则为所述用户面网元存储的多个预定义规则中的一个。

在一种可能的设计中，在所述多个业务流属于同一个 QoS 流时，所述第一规则包括 QoS 执行规则的标识，所述 QoS 执行规则包括所述第一业务流的调度优先级。

25 在一种可能的设计中，收发模块 720，用于还获取所述第一业务流的带宽参数，所述第一业务流的带宽参数包括所述第一业务流的保障带宽，和/或所述第一业务流的最大带宽；或者，所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率，和/或所述第一业务流的最大带宽比率；其中，所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率，所述第一业务流的最大带宽比率为所述第一业务流所能占用的最大带宽与总带宽的比率，所述总带宽是为所述多个业务流分配的共用带宽。

30 在一种可能的设计中，处理模块 710，用于在根据所述第一业务流的调度优先级对所述第一业务流进行调度时，在所述多个业务流分别对应的需求带宽之和大于所述总带宽时，若在所述多个业务流中存在第二业务流，所述第二业务流的保障带宽大于所述第二业务流的需求带宽，则为所述第二业务流分配的带宽等于所述第二业务流的需求带宽；根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽，按照调度优先级由高到低的顺序为所述多个业务流中除所述第二业务流之外的业务流分配差值带宽，所述差值带宽为所述第二业务流的保障带宽与所述第二业务流的需求带宽的差值。

35 在一种可能的设计中，处理模块 710，用于在根据所述第一业务流的调度优先级对所述第一业务流进行调度时，在所述多个业务流分别对应的需求带宽之和大于所述总带宽时，若所述多个业务流分别对应的需求带宽均大于相应的保障带宽，则满足所述多个业务流分别对应的保障带宽比率，所述多个业务流分别对应的保障带宽比率之和小于或等于 1；在满足所述多个业务流分别对应的保障带宽比率之后，根据所述多个业务流分别对应的调度优先级，按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽。

40 在一种可能的设计中，处理模块 710，用于在按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽时，按照所述顺序满足所述多个业务流中部分或全部业务流的最大带宽比率。

45 在一种可能的设计中，所述多个业务流分别对应的实际带宽之和等于所述总带宽。

在一种可能的设计中，所述报文为下行报文；处理模块 710，用于在所述下行报文中添加所述第一业务流的调度优先级，获得第一待发送下行报文；

收发模块 720, 用于向接入网设备发送所述第一待发送下行报文。

在一种可能的设计中, 所述第一待发送下行报文包括通用无线分组业务隧道协议报文头, 所述通用无线分组业务隧道协议报文头包括所述第一业务流的调度优先级。

5 在一种可能的设计中, 所述多个业务流为多个终端设备分别对应的业务流, 所述多个终端设备属于同一个终端设备组。

在一种可能的设计中, 所述多个终端设备的等级不同, 所述第一业务流的调度优先级是根据第一终端设备的等级确定的, 所述第一业务流为所述第一终端设备签约的业务对应的业务流, 所述第一终端设备为所述多个终端设备中的一个。

在一种可能的设计中, 所述多个业务流属于同一个 QoS 流。

10 应理解, 根据本申请实施例的装置 700 可对应于前述方法实施例中用户面网元, 并且装置 700 中的各个模块的操作和/或功能分别为了实现前述方法实施例中用户面网元的方法的相应步骤, 因此也可以实现前述方法实施例中的有益效果, 为了简洁, 这里不作赘述。

当装置 700 为会话管理网元或会话管理网元中的芯片时, 装置 700 中的处理模块 710 可以支持装置 700 执行上文中各方法示例中会话管理网元的动作。

15 收发模块 720 可以支持装置 700 与用户面网元进行通信。

处理模块 710 调用收发模块 720, 执行: 获取第一业务流的调度优先级, 所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级; 所述第一业务流为所述多个业务流中的一个; 向用户面网元发送第一规则, 所述第一规则用于指示所述第一业务流的调度优先级。

20 在一种可能的设计中, 收发模块 720, 用于在获取第一业务流的调度优先级时, 从策略控制网元接收策略与计费控制规则, 所述策略与计费控制规则用于指示所述第一业务流的调度优先级; 处理模块 710, 用于根据所述策略与计费控制规则确定所述第一规则。

在一种可能的设计中, 所述策略与计费控制规则包括第一预定义规则的名称, 所述第一规则包括所述第一预定义规则的名称;

25 所述第一预定义规则包括所述第一业务流的调度优先级, 所述第一预定义规则为所述会话管理网元和所述用户面网元存储的多个预定义规则中的一个。

在一种可能的设计中, 在所述多个业务流属于同一个 QoS 流时, 所述第一规则包括 QoS 执行规则的标识, 所述 QoS 执行规则包括所述第一业务流的调度优先级。

在一种可能的设计中, 处理模块 710, 用于在所述多个业务流属于同一个 QoS 流时, 根据所述策略与计费控制规则确定 QoS 规则, 所述 QoS 规则包括所述第一业务流的调度优先级;

30 收发模块 720, 用于向所述终端设备发送所述 QoS 规则。

在一种可能的设计中, 收发模块 720, 用于还获取所述第一业务流的带宽参数, 所述第一业务流的带宽参数包括所述第一业务流的保障带宽, 和/或所述第一业务流的最大带宽; 所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率, 和/或所述第一业务流的最大带宽比率; 其中, 所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率, 所述第一业务流的最大带宽比率为
35 所述第一业务流所能占用的最大带宽与总带宽的比率, 所述总带宽是为所述多个业务流分配的共用带宽。

在一种可能的设计中, 所述多个业务流为多个终端设备分别对应的业务流, 所述多个终端设备属于同一个终端设备组。

40 在一种可能的设计中, 所述多个终端设备的等级不同, 所述第一业务流的调度优先级是根据第一终端设备的等级确定的, 所述第一业务流为所述第一终端设备签约的业务对应的业务流, 所述第一终端设备为所述多个终端设备中的一个。

在一种可能的设计中, 所述多个业务流属于同一个 QoS 流。

应理解, 根据本申请实施例的装置 700 可对应于前述方法实施例中会话管理网元, 并且装置 700 中的各个模块的操作和/或功能分别为了实现前述方法实施例中会话管理网元的方法的相应步骤, 因此也可以实现前述方法实施例中的有益效果, 为了简洁, 这里不作赘述。

45 当装置 700 为接入网设备或接入网设备中的芯片时, 装置 700 中的处理模块 710 可以支持装置 700 执行上文中各方法示例中接入网设备的动作。

收发模块 720 可以支持装置 700 与用户面网元进行通信。

收发模块 720, 用于接收第一业务流中的报文, 所述报文包括第一业务流的调度优先级, 所述第一

业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级;所述第一业务流为所述多个业务流中的一个;

处理模块 710, 用于根据所述第一业务流的调度优先级对所述第一业务流进行调度。

5 在一种可能的设计中, 所述报文还包括所述第一业务流的带宽参数, 所述第一业务流的带宽参数包括所述第一业务流的保障带宽, 和/或所述第一业务流的最大带宽; 或者, 所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率, 和/或所述第一业务流的最大带宽比率; 其中, 所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率, 所述第一业务流的最大带宽比率为所述第一业务流所能占用的最大带宽与总带宽的比率, 所述总带宽是为所述多个业务流分配的共用带宽。

10 在一种可能的设计中, 处理模块 710, 用于在根据所述第一业务流的调度优先级对所述第一业务流进行调度时, 在所述多个业务流分别对应的需求带宽之和大于所述总带宽时, 若在所述多个业务流中存在第二业务流, 所述第二业务流的保障带宽大于所述第二业务流的需求带宽, 则为所述第二业务流分配的带宽等于所述第二业务流的需求带宽; 根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽, 按照调度优先级由高到低的顺序为所述多个业务流中除所述第二业务流之外的业务流分配差值带宽, 所述差值带宽为所述第二业务流的保障带宽与所述第二业务流的需求带宽的差值。

15 在一种可能的设计中, 处理模块 710, 用于在根据所述第一业务流的调度优先级对所述第一业务流进行调度时, 在所述多个业务流分别对应的需求带宽之和大于所述总带宽时, 若所述多个业务流分别对应的需求带宽均大于相应的保障带宽, 则满足所述多个业务流分别对应的保障带宽比率, 所述多个业务流分别对应的保障带宽比率之和小于或等于 1; 在满足所述多个业务流分别对应的保障带宽比率之后, 根据所述多个业务流分别对应的调度优先级, 按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽。

20 在一种可能的设计中, 处理模块 710, 用于在按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽时, 按照所述顺序满足所述多个业务流中部分或全部业务流的最大带宽比率。

在一种可能的设计中, 所述多个业务流分别对应的实际带宽之和等于所述总带宽。

25 在一种可能的设计中, 所报文包括通用无线分组业务隧道协议报文头, 所述通用无线分组业务隧道协议报文头包括所述第一业务流的调度优先级。

在一种可能的设计中, 所述多个业务流为多个终端设备分别对应的业务流, 所述多个终端设备属于同一个终端设备组。

30 在一种可能的设计中, 所述多个终端设备的等级不同, 所述第一业务流的调度优先级是根据第一终端设备的等级确定的, 所述第一业务流为所述第一终端设备签约的业务对应的业务流, 所述第一终端设备为所述多个终端设备中的一个。

在一种可能的设计中, 所述多个业务流属于同一个 QoS 流。

35 应理解, 根据本申请实施例的装置 700 可对应于前述方法实施例中接入网设备, 并且装置 700 中的各个模块的操作和/或功能分别为了实现前述方法实施例中接入网设备的方法的相应步骤, 因此也可以实现前述方法实施例中的有益效果, 为了简洁, 这里不作赘述。

图 8 示出了根据本申请实施例的通信装置 800 的示意性结构图。如图 8 所示, 所述装置 800 包括: 处理器 801。

当装置 800 为用户面网元或用户面网元中的芯片时, 一种可能的实现方式中, 当所述处理器 801 用于调用接口执行以下动作:

40 获取第一业务流的调度优先级, 所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级; 所述第一业务流为所述多个业务流中的一个; 接收报文; 在所述报文属于所述第一业务流时, 根据所述第一业务流的调度优先级对所述第一业务流进行调度。

应理解, 所述装置 800 还可用于执行前文实施例中用户面网元侧的其他步骤和/或操作, 为了简洁, 这里不作赘述。

45 当装置 800 为会话管理网元或会话管理网元中的芯片时, 一种可能的实现方式中, 当所述处理器 801 用于调用接口执行以下动作:

获取第一业务流的调度优先级, 所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级; 所述第一业务流为所述多个业务流中的一个; 向用户面网元发送第一规则, 所述第一规则

用于指示所述第一业务流的调度优先级。

应理解, 所述装置 800 还可用于执行前文实施例中会话管理网元侧的其他步骤和/或操作, 为了简洁, 这里不作赘述。

5 当装置 800 为接入网设备或接入网设备中的芯片时, 一种可能的实现方式中, 当所述处理器 801 用于调用接口执行以下动作: 接收第一业务流中的报文, 所述报文包括第一业务流的调度优先级, 所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级; 所述第一业务流为所述多个业务流中的一个; 根据所述第一业务流的调度优先级对所述第一业务流进行调度。

应理解, 所述装置 800 还可用于执行前文实施例中接入网设备侧的其他步骤和/或操作, 为了简洁, 这里不作赘述。

10 应理解, 所述处理器 801 可以调用接口执行上述收发动作, 其中, 调用的接口可以是逻辑接口或物理接口, 对此不作限定。可选地, 物理接口可以通过收发器实现。可选地, 所述装置 800 还包括收发器 803。

可选地, 所述装置 800 还包括存储器 802, 存储器 802 中可以存储上述方法实施例中的程序代码, 以便于处理器 801 调用。

15 具体地, 若所述装置 800 包括处理器 801、存储器 802 和收发器 803, 则处理器 801、存储器 802 和收发器 803 之间通过内部连接通路互相通信, 传递控制和/或数据信号。在一个可能的设计中, 处理器 801、存储器 802 和收发器 803 可以通过芯片实现, 处理器 801、存储器 802 和收发器 803 可以在同一个芯片中实现, 也可能分别在不同的芯片实现, 或者其中任意两个功能组合在一个芯片中实现。该存储器 802 可以存储程序代码, 处理器 801 调用存储器 802 存储的程序代码, 以实现装置 800 的相应功能。

20 上述本申请实施例揭示的方法可以应用于处理器中, 或者由处理器实现。处理器可能是一种集成电路芯片, 具有信号的处理能力。在实现过程中, 上述方法实施例的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器、数字信号处理器 (digital signal processor, DSP)、专用集成电路 (application specific integrated circuit, ASIC)、现成可编程门阵列 (field programmable gate array, FPGA) 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, 还可以是系统芯片 (system on chip, SoC), 还可以是中央处理器 (central processor unit, CPU), 还可以是网络处理器 (network processor, NP), 还可以是数字信号处理电路 (digital signal processor, DSP), 还可以是微控制器 (micro controller unit, MCU), 还可以是可编程控制器 (programmable logic device, PLD) 或其他集成芯片。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。

25 通用处理器可以是微处理器或者该处理器也可以是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译码处理器执行完成, 或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器, 闪存、只读存储器, 可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器, 处理器读取存储器中的信息, 结合其硬件完成上述方法的步骤。

30 可以理解, 本申请实施例中的存储器可以是易失性存储器或非易失性存储器, 或可包括易失性和非易失性存储器两者。其中, 非易失性存储器可以是只读存储器 (read-only memory, ROM)、可编程只读存储器 (programmable ROM, PROM)、可擦除可编程只读存储器 (erasable PROM, EPROM)、电可擦除可编程只读存储器 (electrically EPROM, EEPROM) 或闪存。易失性存储器可以是随机存取存储器 (random access memory, RAM), 其用作外部高速缓存。通过示例性但不是限制性说明, 许多形式的 RAM 可用, 例如静态随机存取存储器 (static RAM, SRAM)、动态随机存取存储器 (dynamic RAM, DRAM)、同步动态随机存取存储器 (synchronous DRAM, SDRAM)、双倍数据速率同步动态随机存取存储器 (double data rate SDRAM, DDR SDRAM)、增强型同步动态随机存取存储器 (enhanced SDRAM, ESDRAM)、同步连接动态随机存取存储器 (synchlink DRAM, SLDRAM) 和直接内存总线随机存取存储器 (direct rambus RAM, DR RAM)。应注意, 本文描述的系统和方法的存储器旨在包括但不限于这些和任意其它适合类型的存储器。

35 40 45

应理解, 在本申请实施例中, 编号“第一”、“第二”... 仅仅为了区分不同的对象, 比如为了区分不同的参数信息或者消息, 并不对本申请实施例的范围构成限制, 本申请实施例并不限于此。

还应理解, 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

5 程的执行顺序应以其功能和内在逻辑确定。上述各个过程涉及的各种数字编号或序号仅为描述方便进行的区分，而不应对本申请实施例的实施过程构成任何限定。

还应理解，本文中术语“和/或”，仅仅是一种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 这三种情况。另外，本文中
5 字符“/”，一般表示前后关联对象是一种“或”的关系。

本申请中出现的类似于“项目包括如下中的一项或多项：A，B，以及 C”表述的含义，如无特别说明，通常是指该项目可以为如下中任一个：A；B；C；A 和 B；A 和 C；B 和 C；A，B 和 C；A 和 A；
10 A，A 和 A；A，A 和 B；A，A 和 C，A，B 和 B；A，C 和 C；B 和 B，B，B 和 B，B，B 和 C，C 和 C；C，C 和 C，以及其他 A，B 和 C 的组合。以上是以 A，B 和 C 共 3 个元素进行举例来说明该项目的
15 的可选用条目，当表达为“项目包括如下中至少一种：A，B，……，以及 X”时，即表达中具有更多元素时，那么该项目可以适用的条目也可以按照前述规则获得。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，
20 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。

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
25 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。

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
30 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。

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，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
35 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 ROM、随机存取存储器 RAM、磁碟或者光盘等各种可以存储程序代码的介质。

权利要求

1.一种业务流调度方法，其特征在于，该方法包括：

用户面网元获取第一业务流的调度优先级，所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级；所述第一业务流为所述多个业务流中的一个；

5 所述用户面网元接收报文；

在所述报文属于所述第一业务流时，所述用户面网元根据所述第一业务流的调度优先级对所述第一业务流进行调度。

2.如权利要求1所述的方法，其特征在于，用户面网元获取第一业务流的调度优先级，包括：

10 所述用户面网元从会话管理网元接收第一规则，所述第一规则用于指示所述第一业务流的调度优先级。

3.如权利要求2所述的方法，其特征在于，所述第一规则包括第一预定义规则的名称，所述第一预定义规则包括所述第一业务流的调度优先级，所述第一预定义规则为所述用户面网元存储的多个预定义规则中的一个。

4.如权利要求3所述的方法，其特征在于，在所述多个业务流属于同一个QoS流时，所述第一规则包括QoS执行规则的标识，所述QoS执行规则包括所述第一业务流的调度优先级。

5.如权利要求1-4任一项所述的方法，其特征在于，还包括：

所述用户面网元还获取所述第一业务流的带宽参数；

20 所述第一业务流的带宽参数包括所述第一业务流的保障带宽，和/或所述第一业务流的最大带宽；或者，所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率，和/或所述第一业务流的最大带宽比率；其中，所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率，所述第一业务流的最大带宽比率为所述第一业务流所能占用的最大带宽与总带宽的比率，所述总带宽是为所述多个业务流分配的共用带宽。

6.如权利要求5所述的方法，其特征在于，所述用户面网元根据所述第一业务流的调度优先级对所述第一业务流进行调度，包括：

25 在所述多个业务流分别对应的需求带宽之和大于所述总带宽时，若在所述多个业务流中存在第二业务流，所述第二业务流的保障带宽大于所述第二业务流的需求带宽，则为所述第二业务流分配的带宽等于所述第二业务流的需求带宽；

30 根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽，按照调度优先级由高到低的顺序为所述多个业务流中除所述第二业务流之外的业务流分配差值带宽，所述差值带宽为所述第二业务流的保障带宽与所述第二业务流的需求带宽的差值。

7.如权利要求5所述的方法，其特征在于，所述用户面网元根据所述第一业务流的调度优先级对所述第一业务流进行调度，包括：

35 在所述多个业务流分别对应的需求带宽之和大于所述总带宽时，若所述多个业务流分别对应的需求带宽均大于相应的保障带宽，则满足所述多个业务流分别对应的保障带宽比率，所述多个业务流分别对应的保障带宽比率之和小于或等于1；

在满足所述多个业务流分别对应的保障带宽比率之后，根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽，按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽。

40 8.如权利要求7所述的方法，其特征在于，按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽，包括：

按照所述顺序满足所述多个业务流中部分或全部业务流的最大带宽比率。

9.如权利要求6-8任一项所述的方法，其特征在于，所述多个业务流分别对应的实际带宽之和等于所述总带宽。

10.如权利要求1-9任一项所述的方法，其特征在于，所述报文为下行报文；

45 所述方法还包括：

所述用户面网元在所述下行报文中添加所述第一业务流的调度优先级，获得第一待发送下行报文；所述用户面网元向接入网设备发送所述第一待发送下行报文。

11.如权利要求 10 所述的方法,其特征在于,所述第一待发送下行报文包括通用无线分组业务隧道协议报文头,所述通用无线分组业务隧道协议报文头包括所述第一业务流的调度优先级。

12.如权利要求 1-3、5-11 任一项所述的方法,其特征在于,所述多个业务流为多个终端设备分别对应的业务流,所述多个终端设备属于同一个终端设备组。

5 13.如权利要求 12 所述的方法,其特征在于,所述多个终端设备的等级不同,所述第一业务流的调度优先级是根据第一终端设备的等级确定的,所述第一业务流为所述第一终端设备签约的业务对应的业务流,所述第一终端设备为所述多个终端设备中的一个。

14.如权利要求 1-11 任一项所述的方法,其特征在于,所述多个业务流属于同一个服务质量 QoS 流。

15.一种业务流调度方法,其特征在于,该方法包括:

10 会话管理网元获取第一业务流的调度优先级,所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级;所述第一业务流为所述多个业务流中的一个;

所述会话管理网元向用户面网元发送第一规则,所述第一规则用于指示所述第一业务流的调度优先级。

16.如权利要求 15 所述的方法,其特征在于,会话管理网元获取第一业务流的调度优先级,包括:

15 所述会话管理网元从策略控制网元接收策略与计费控制规则,所述策略与计费控制规则用于指示所述第一业务流的调度优先级;

所述会话管理网元根据所述策略与计费控制规则确定所述第一规则。

17.如权利要求 16 所述的方法,其特征在于,所述策略与计费控制规则包括第一预定义规则的名称,所述第一规则包括所述第一预定义规则的名称;

20 所述第一预定义规则包括所述第一业务流的调度优先级,所述第一预定义规则为所述会话管理网元和所述用户面网元存储的多个预定义规则中的一个。

18.如权利要求 16 所述的方法,其特征在于,在所述多个业务流属于同一个 QoS 流时,所述第一规则包括 QoS 执行规则的标识,所述 QoS 执行规则包括所述第一业务流的调度优先级。

19.如权利要求 16-18 任一项所述的方法,其特征在于,还包括:

25 在所述多个业务流属于同一个 QoS 流时,所述会话管理网元根据所述策略与计费控制规则确定 QoS 规则,所述 QoS 规则包括所述第一业务流的调度优先级;

所述会话管理网元向所述终端设备发送所述 QoS 规则。

20.如权利要求 15-19 任一项所述的方法,其特征在于,还包括:

所述会话管理网元还获取所述第一业务流的带宽参数;

30 所述第一业务流的带宽参数包括所述第一业务流的保障带宽,和/或所述第一业务流的最大带宽;或者,所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率,和/或所述第一业务流的最大带宽比率;其中,所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率,所述第一业务流的最大带宽比率为所述第一业务流所能占用的最大带宽与总带宽的比率,所述总带宽是为所述多个业务流分配的共用带宽。

21.一种业务流调度方法,其特征在于,该方法包括:

35 接入网设备接收第一业务流中的报文,所述报文包括第一业务流的调度优先级,所述第一业务流的调度优先级为所述第一业务流在多个业务流中的调度优先级;所述第一业务流为所述多个业务流中的一个;

所述接入网设备根据所述第一业务流的调度优先级对所述第一业务流进行调度。

40 22.如权利要求 21 所述的方法,其特征在于,所述报文还包括所述第一业务流的带宽参数;

所述带宽参数包括所述第一业务流的保障带宽,和/或所述第一业务流的最大带宽;或者,所述第一业务流的带宽参数包括所述第一业务流的保障带宽比率,和/或所述第一业务流的最大带宽比率;其中,所述第一业务流的保障带宽比率为所述第一业务流的保障带宽与总带宽的比率,所述第一业务流的最大带宽比率为所述第一业务流所能占用的最大带宽与总带宽的比率,所述总带宽是为所述多个业务流分配的共用带宽。

45 23.如权利要求 22 所述的方法,其特征在于,所述接入网设备根据所述第一业务流的调度优先级对所述第一业务流进行调度,包括:

在所述多个业务流分别对应的需求带宽之和大于所述总带宽时,若在所述多个业务流中存在第二业

务流, 所述第二业务流的保障带宽大于所述第二业务流的需求带宽, 则为所述第二业务流分配的带宽等于所述第二业务流的需求带宽;

5 根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽, 按照调度优先级由高到低的顺序为所述多个业务流中除所述第二业务流之外的业务流分配差值带宽, 所述差值带宽为所述第二业务流的保障带宽与所述第二业务流的需求带宽的差值。

24. 如权利要求 22 所述的方法, 其特征在于, 所述接入网设备根据所述第一业务流的调度优先级对所述第一业务流进行调度, 包括:

10 在所述多个业务流分别对应的需求带宽之和大于所述总带宽时, 若所述多个业务流分别对应的需求带宽均大于相应的保障带宽, 则满足所述多个业务流分别对应的保障带宽比率, 所述多个业务流分别对应的保障带宽比率之和小于或等于 1;

在满足所述多个业务流分别对应的保障带宽比率之后, 根据所述多个业务流分别对应的调度优先级和所述多个业务流分别对应的需求带宽, 按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽。

15 25. 如权利要求 24 所述的方法, 其特征在于, 按照调度优先级由高到低的顺序为所述多个业务流中的部分或全部业务流分配剩余带宽, 包括:

按照所述顺序满足所述多个业务流中部分或全部业务流的最大带宽比率。

26. 如权利要求 23-25 任一项所述的方法, 其特征在于, 所述多个业务流分别对应的实际带宽之和等于所述总带宽。

20 27. 如权利要求 21-26 任一项所述的方法, 其特征在于, 所报文包括通用无线分组业务隧道协议报文头, 所述通用无线分组业务隧道协议报文头包括所述第一业务流的调度优先级。

28. 一种通信装置, 其特征在于, 所述通信装置用于执行如权利要求 1-14 任一项所述的方法, 或权利要求 15-20 任一项所述的方法, 或权利要求 21-27 任一项所述的方法。

25 29. 一种芯片, 其特征在于, 所述芯片与存储器耦合, 所述芯片读取所述存储器中存储的计算机程序, 执行权利要求 1-14 任一项所述的方法, 或权利要求 15-20 任一项所述的方法, 或权利要求 21-27 任一项所述的方法。

30. 一种计算机可读存储介质, 其特征在于, 所述存储介质中存储有计算机程序或指令, 当所述计算机程序或指令被通信装置执行时, 实现如权利要求 1-14 任一项所述的方法, 或权利要求 15-20 任一项所述的方法, 或权利要求 21-27 任一项所述的方法。

30 31. 一种包含指令的计算机程序产品, 其特征在于, 当所述指令在计算机上运行时, 使得计算机执行如权利要求 1-14 中任一项所述的方法, 或权利要求 15-20 中任一项所述的方法, 或权利要求 21-27 中任一项所述的方法。

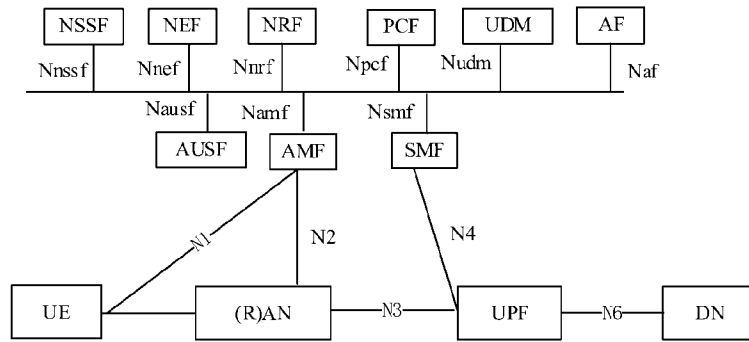


图 1

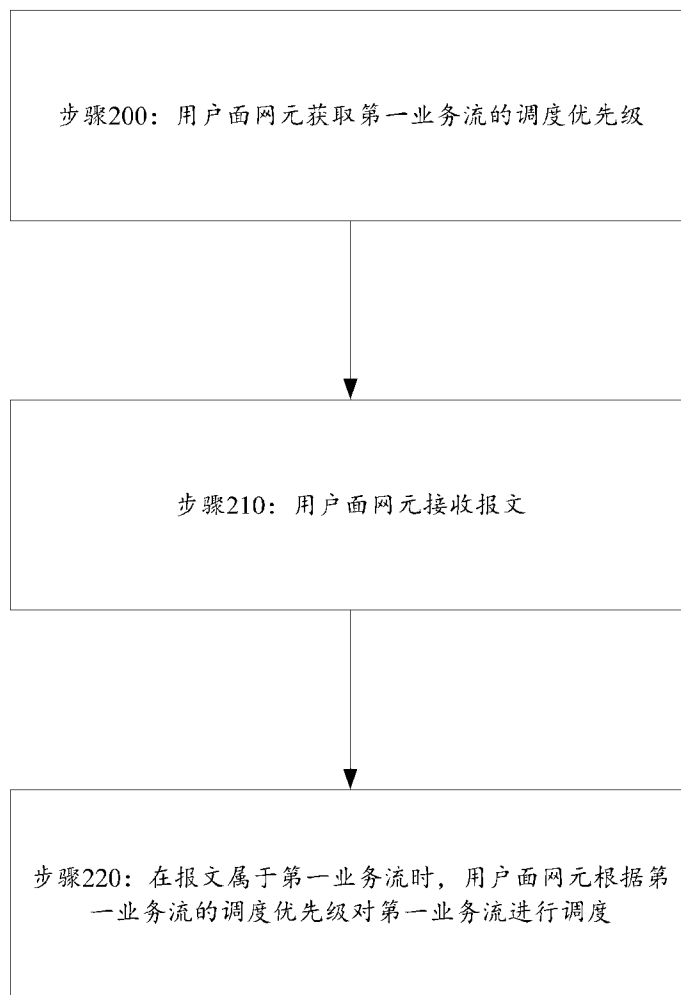


图 2

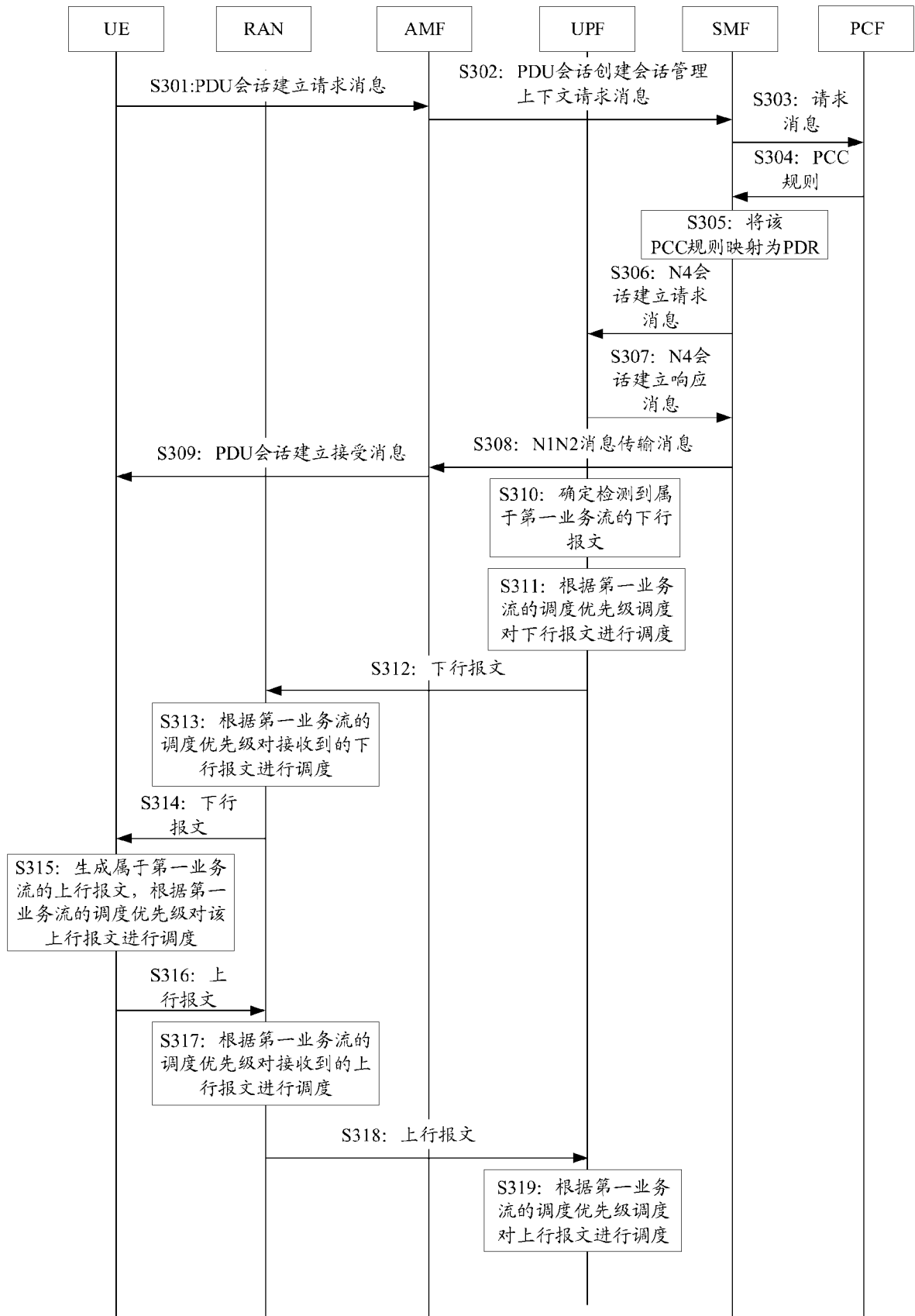


图 3

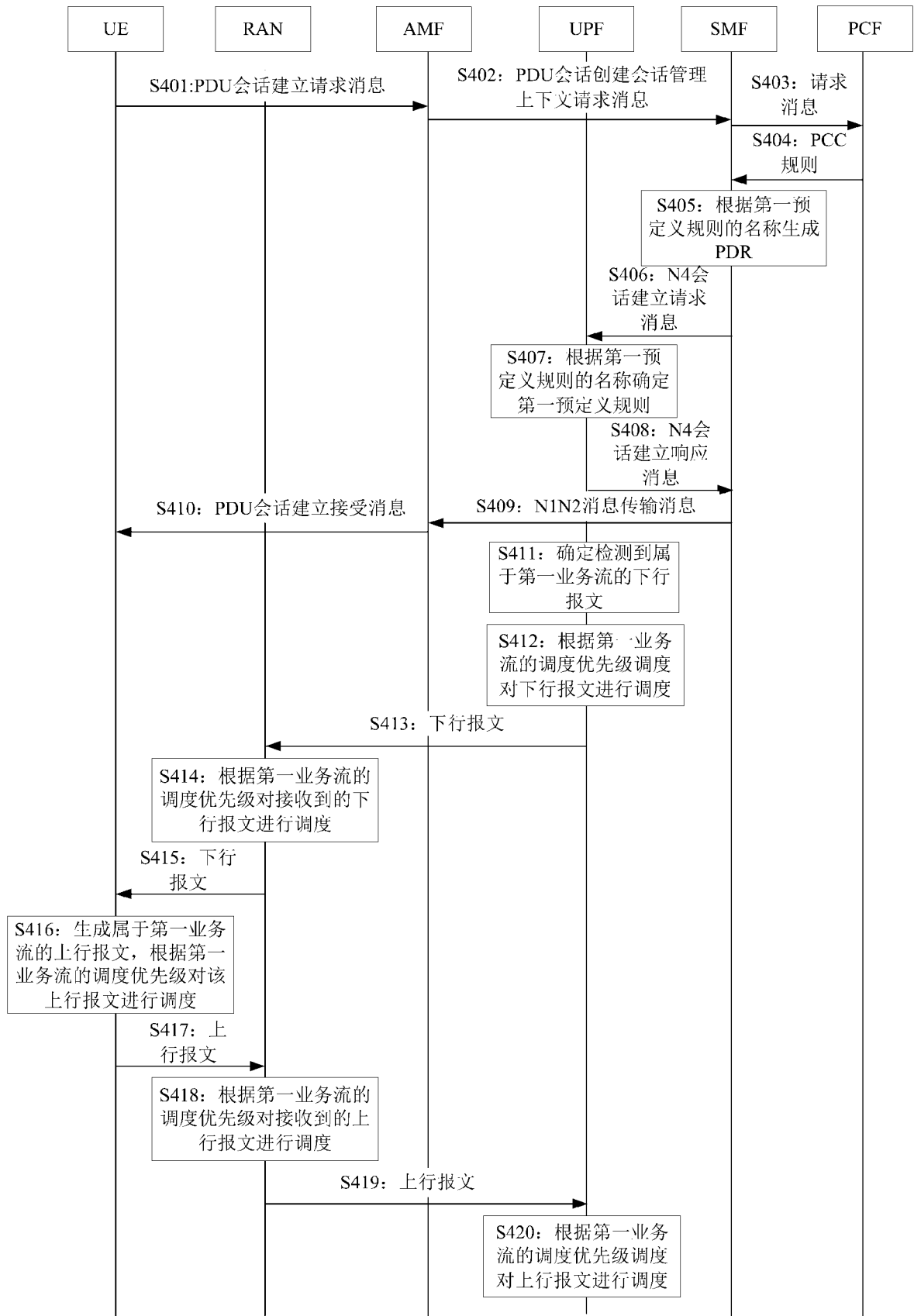


图 4

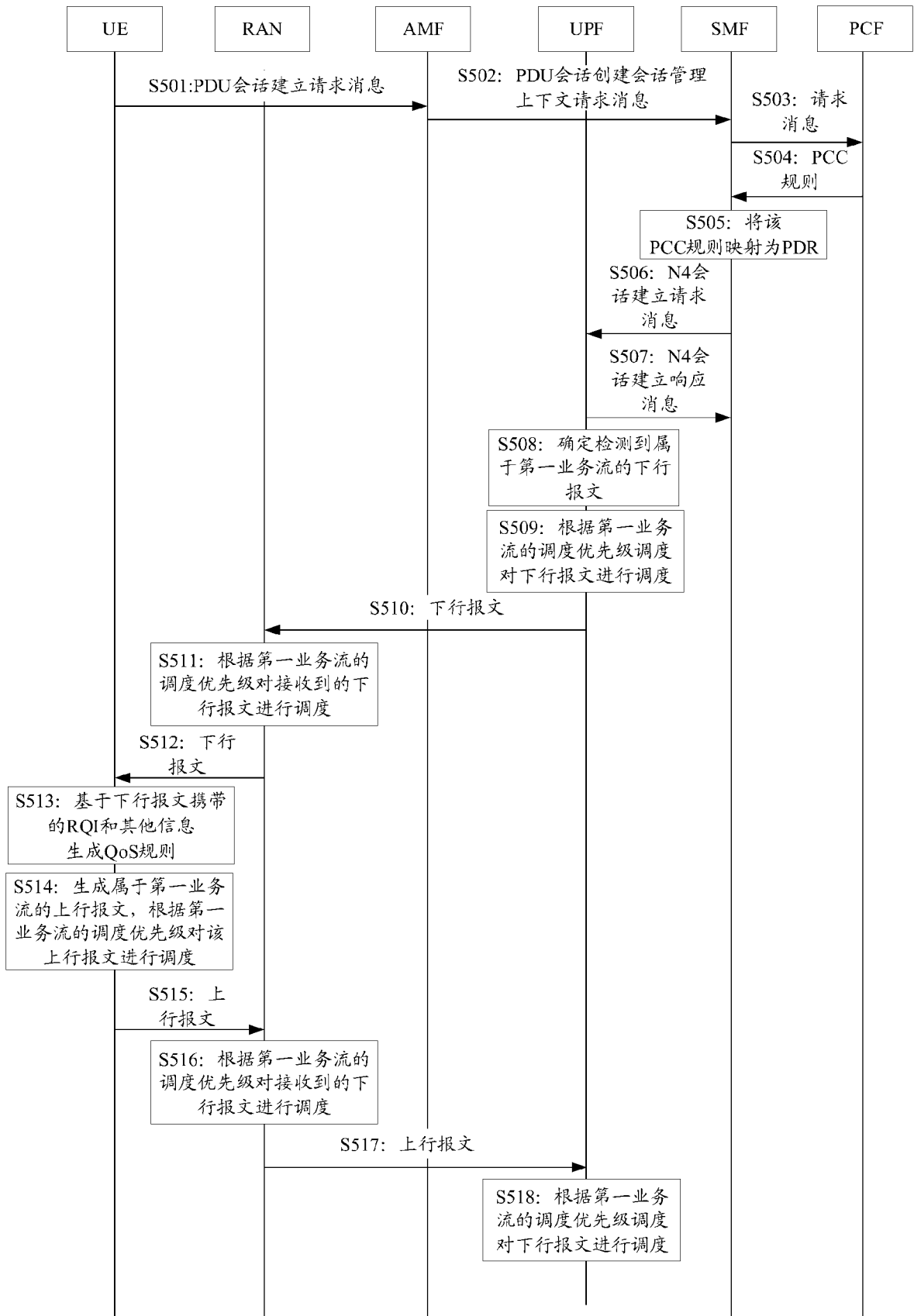


图 5

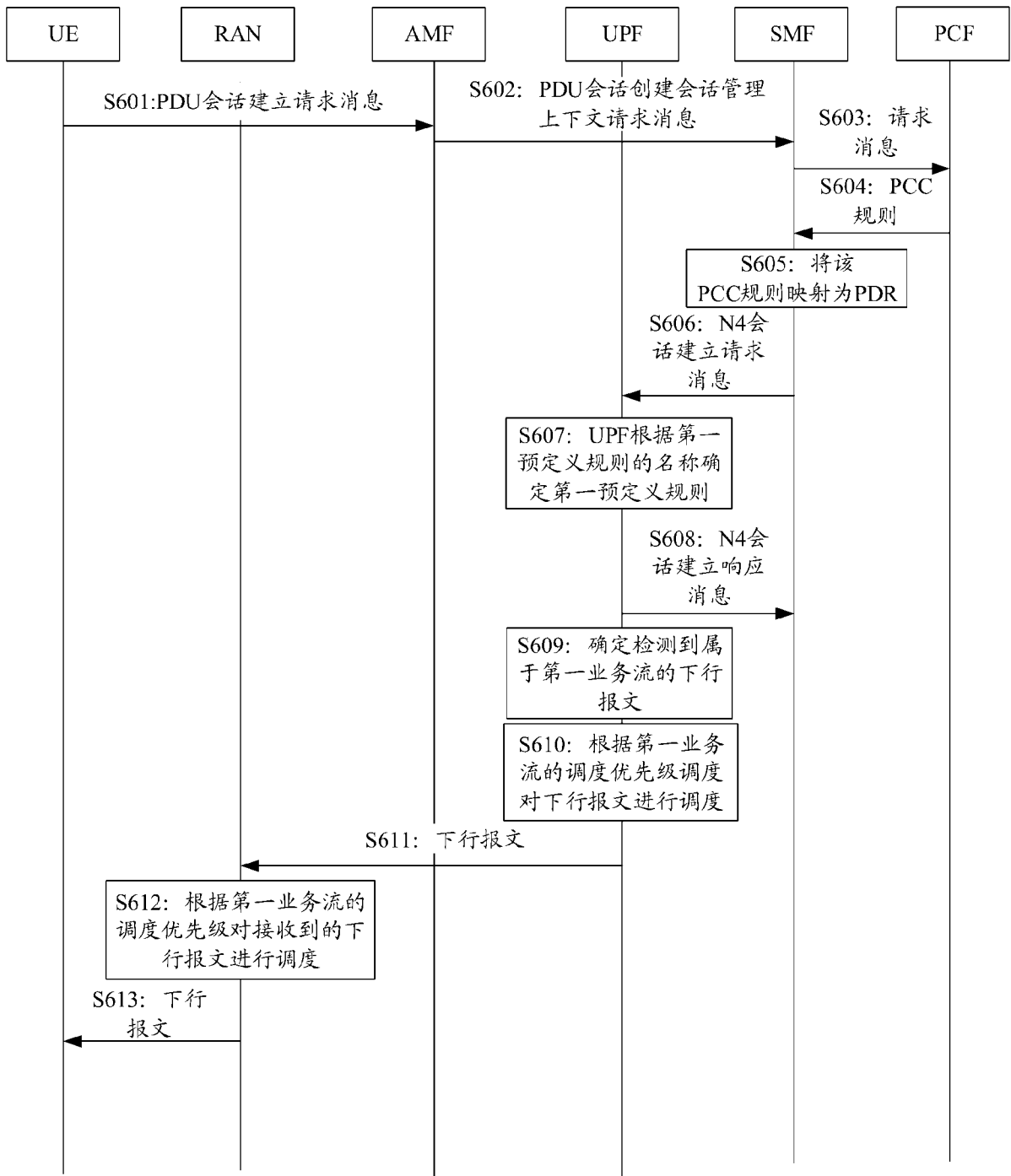


图 6

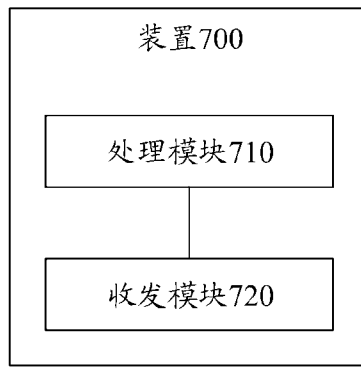


图7

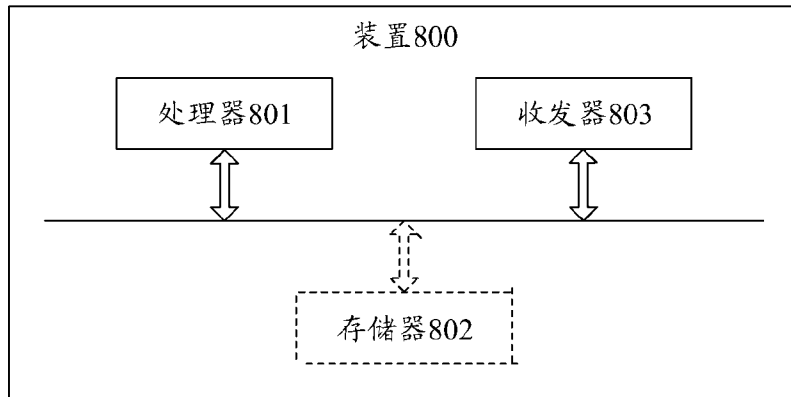


图8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/132949

A. CLASSIFICATION OF SUBJECT MATTER		
H04W 72/566(2023.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
IPC: H04L, H04W		
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, ENTXTC, DWPI, 3GPP: 业务流, 调度, 用户面网元, 优先级, 报文, 不同; traffic flow, scheduling, UPF, priority, packet, different		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 02 June 2022 (2022-06-02) description, paragraphs 13-91	1, 5, 10-14, 28-31
Y	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 02 June 2022 (2022-06-02) description, paragraphs 13-91	2-4, 15-22, 27
A	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 02 June 2022 (2022-06-02) description, paragraphs 13-91	6-9, 23-26
Y	ERICSSON et al. "Clarifications for 5QI priority level" 3GPP TSG-SA2 Meeting #129 S2-1811305, 12 October 2018 (2018-10-12), section 5.7	2-4, 15-22, 27
A	CN 101778431 A (ZTE CORP.) 14 July 2010 (2010-07-14) entire document	1-31
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
31 January 2024		06 February 2024
Name and mailing address of the ISA/CN		Authorized officer
China National Intellectual Property Administration (ISA/CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/132949

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2022174544	A1	02 June 2022	US	2023007537	A1	05 January 2023
CN	101778431	A	14 July 2010	None			
WO	2020192438	A1	01 October 2020	CN	111757511	A	09 October 2020

<p>A. 主题的分类</p> <p>H04W 72/566(2023.01)i</p> <p>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</p>																							
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>IPC: H04L, H04W</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNXTX, ENXTXC, DWPI, 3GPP: 业务流, 调度, 用户面网元, 优先级, 报文, 不同; traffic flow, scheduling, UPP, priority, packet, different</p>																							
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段</td> <td>1, 5, 10-14, 28-31</td> </tr> <tr> <td>Y</td> <td>US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段</td> <td>2-4, 15-22, 27</td> </tr> <tr> <td>A</td> <td>US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段</td> <td>6-9, 23-26</td> </tr> <tr> <td>Y</td> <td>ERICSSON等. "Clarifications for 5QI priority level" 3GPP TSG-SA2 Meeting #129 S2-1811305, 2018年10月12日 (2018 - 10 - 12), 第5.7节</td> <td>2-4, 15-22, 27</td> </tr> <tr> <td>A</td> <td>CN 101778431 A (中兴通讯股份有限公司) 2010年7月14日 (2010 - 07 - 14) 全文</td> <td>1-31</td> </tr> <tr> <td>A</td> <td>WO 2020192438 A1 (华为技术有限公司) 2020年10月1日 (2020 - 10 - 01) 全文</td> <td>1-31</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 引用文件的具体类型: "A" 认为不特别相关的表示了现有技术一般状态的文件 "D" 申请人在国际申请中引证的文件 "E" 在国际申请日的当天或之后公布的在先申请或专利 "L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) "O" 涉及口头公开、使用、展览或其他方式公开的文件 "P" 公布日先于国际申请日但迟于所要求的优先权日的文件 "T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 "X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 "Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 "&" 同族专利的文件</p>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段	1, 5, 10-14, 28-31	Y	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段	2-4, 15-22, 27	A	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段	6-9, 23-26	Y	ERICSSON等. "Clarifications for 5QI priority level" 3GPP TSG-SA2 Meeting #129 S2-1811305, 2018年10月12日 (2018 - 10 - 12), 第5.7节	2-4, 15-22, 27	A	CN 101778431 A (中兴通讯股份有限公司) 2010年7月14日 (2010 - 07 - 14) 全文	1-31	A	WO 2020192438 A1 (华为技术有限公司) 2020年10月1日 (2020 - 10 - 01) 全文	1-31
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
X	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段	1, 5, 10-14, 28-31																					
Y	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段	2-4, 15-22, 27																					
A	US 2022174544 A1 (VERIZON PATENT AND LICENSING INC.) 2022年6月2日 (2022 - 06 - 02) 说明书第13-91段	6-9, 23-26																					
Y	ERICSSON等. "Clarifications for 5QI priority level" 3GPP TSG-SA2 Meeting #129 S2-1811305, 2018年10月12日 (2018 - 10 - 12), 第5.7节	2-4, 15-22, 27																					
A	CN 101778431 A (中兴通讯股份有限公司) 2010年7月14日 (2010 - 07 - 14) 全文	1-31																					
A	WO 2020192438 A1 (华为技术有限公司) 2020年10月1日 (2020 - 10 - 01) 全文	1-31																					
<p>国际检索实际完成的日期</p> <p>2024年1月31日</p>	<p>国际检索报告邮寄日期</p> <p>2024年2月6日</p>																						
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088</p>	<p>授权官员</p> <p>汤广强</p> <p>电话号码 (+86) 010-53961734</p>																						

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/132949

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)
US	2022174544	A1	2022年6月2日	US	2023007537	A1	2023年1月5日
CN	101778431	A	2010年7月14日	无			
WO	2020192438	A1	2020年10月1日	CN	111757511	A	2020年10月9日