

多人即时通讯方法、系统、装置及电子设备

本公开要求于 2020 年 01 月 16 日提交的申请号为 202010048012.7、发明名称为“多人即时通讯方法、系统、装置及电子设备”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。

技术领域

本公开涉及即时通讯领域，具体涉及一种多人即时通讯方法、系统、装置及电子设备。

背景技术

随着信息技术的快速发展，多人即时通讯越来越被广泛地进行应用。而随着通讯规模的扩大、同一时间在同一即时通讯房间内参与的用户数量的增加，现有技术暴露了以下问题：在多人即时通讯中，由于接收方的条件参差不齐，而现有技术中的多人技术通讯对于接收方的兼容性较低，导致不能为部分接收方提供足够良好的服务。

发明内容

本公开的一个目的在于提出一种多人即时通讯方法、系统、装置及电子设备，能够提高多人即时通讯对接收方的兼容性。

根据本公开实施例的一方面，公开了一种多人即时通讯方法，所述方法由服务器执行，包括：

获取第一数目个的第一多媒体数据流；

将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目；

将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

根据本公开实施例的一方面，公开了一种多人即时通讯系统，所述多人即时通讯系统包括调度节点、至少两个候选数据适配节点，其中：

所述调度节点配置为从所述至少两个候选数据适配节点中选取一候选数据适配节点，作为目标数据适配节点；

所述目标数据适配节点配置为获取第一数目个的第一多媒体数据流，将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目，将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

根据本公开实施例的一方面，公开了一种多人即时通讯装置，包括：

获取模块，配置为获取第一数目个的第一多媒体数据流；

融合模块，配置为将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目；

发送模块，配置为将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

根据本公开实施例的一方面，公开了一种多人即时通讯电子设备，包括：存储器，存储有计算机可读指令；处理器，读取存储器存储的计算机可读指令，以执行以上多人即时通讯方法。

根据本公开实施例的一方面，公开了一种计算机可读存储介质，该计算机可读存储介质上存储有计算机可读指令，该计算机可读指令被处理器执行，以实现以上多人即时通讯方法。

本公开实施例中，在多人即时通讯中，获取到各发送方分别上传的第一多媒体数据流后，对获取到的第一多媒体数据流进行融合，得到数目更小的第二多媒体数据流，再将第二多媒体数据流发送给接收方，从而降低了接收方在接收数据时对带宽的要求，提高了对接收方的兼容性。

本公开的其他特性和优点将通过下面的详细描述变得显然，或部分地通过本公开的实践而习得。

应当理解的是，以上的一般描述和后文的细节描述仅是示例性的，并不能限制本公开。

附图说明

通过参考附图详细描述其示例实施例，本公开的上述和其它目标、特征及优点将变得更加显而易见。

图1示出了根据本公开一个实施例的多人即时通讯的基本体系构架组成。

图2示出了根据本公开一个实施例的多人即时通讯的体系构架组成。

图3示出了根据本公开一个实施例的多人即时通讯方法的流程图。

图4示出了根据本公开一个实施例的多人即时通讯系统的组成。

图5A示出了根据本公开一个实施例的区块链系统的组成。

图5B示出了根据本公开一个实施例的区块链系统中的区块的结构。

图5C示出了根据本公开一个实施例的一个新区块产生的过程。

图6示出了根据本公开一个实施例的多人即时通讯系统的组成。

图7示出了根据本公开一个实施例的多人即时通讯系统的组成。

图8示出了根据本公开一个实施例的多人即时通讯系统的组成。

图9示出了根据本公开一个实施例的多人即时通讯系统在直播场景中的初始处理过程。

图10示出了根据本公开一个实施例的多人即时通讯系统在直播场景中申请混流的处理过程。

图11示出了根据本公开一个实施例的多人即时通讯系统在直播场景中申请观看的处理过程。

图12示出了根据本公开一个实施例的多人即时通讯系统在直播场景中上行

数据的处理过程。

图 13 示出了根据本公开一个实施例的多人即时通讯系统在直播场景中下行数据的处理过程。

图 14 示出了根据本公开一个实施例的多人即时通讯装置的框图。

图 15 示出了根据本公开一个实施例的多人即时通讯电子设备的硬件图。

具体实施方式

现在将参考附图更全面地描述示例实施方式。然而，示例实施方式能够以多种形式实施，且不应被理解为限于在此阐述的范例；相反，提供这些示例实施方式使得本公开的描述将更加全面和完整，并将示例实施方式的构思全面地传达给本领域的技术人员。附图仅为本公开的示意性图解，并非一定是按比例绘制。图中相同的附图标记表示相同或类似的部分，因而将省略对它们的重复描述。

此外，所描述的特征、结构或特性可以以任何合适的方式结合在一个或更多示例实施方式中。在下面的描述中，提供许多具体细节从而给出对本公开的示例实施方式的充分理解。然而，本领域技术人员将意识到，可以实践本公开的技术方案而省略所述特定细节中的一个或更多，或者可以采用其它的方法、组元、步骤等。在其它情况下，不详细示出或描述公知结构、方法、实现或者操作以避免喧宾夺主而使得本公开的各方面变得模糊。

附图中所示的一些方框图是功能实体，不一定必须与物理或逻辑上独立的实体相对应。可以采用软件形式来实现这些功能实体，或在一个或多个硬件模块或集成电路中实现这些功能实体，或在不同网络和/或处理器装置和/或微控制器装置中实现这些功能实体。

首先对本公开实施例涉及的概念进行简要解释。

第一多媒体数据流指的是尚未经过服务器处理的多媒体数据流。

第二多媒体数据流指的是经过服务器融合的、待发送给接收方的多媒体数据流。

混音处理指的是将多路音频进行混合的处理过程，例如：将多路音频混合为一路音频。经过混音处理的多路音频被接收方播放后，能够被接收方同时播放出这多路音频。

混屏处理指的是将多路视频进行混合的处理过程，例如：将多路视频混合为一路视频。经过混屏处理的多路视频被接收方播放后，能够在接收方的屏幕上同时呈现出这多路视频，例如：将 4 路视频进行混合为 1 路视频，这 1 路视频被接收方播放后，能够在接收方的屏幕上以四宫格的布局（具体的布局可以由接收方进行定制，例如：接收方通过预设有关视频显示模板的参数）同时呈现出这 4 路视频。

下面参考图 1、图 2 对本公开实施例的体系构架进行描述。

图 1 示出了本公开一实施例的基本体系构架组成：服务器 10、第一用户端

20. 各个第一用户端 20 之间通过服务器 10 进行多媒体数据流的发送与接收, 从而实现多人即时通讯。其中, 每个第一用户端 20 都可以作为发送方, 将自己所采集、生成的第一多媒体数据流发送给服务器 10; 每个第一用户端 20 都可以作为接收方, 从服务器 10 处接收由多个第一用户端 20 所采集、生成的第一多媒体数据流融合而成的第二多媒体数据流。具体的, 服务器 10 可以接收每一第一用户端 20 发送的第一多媒体数据流; 将接收到的各第一多媒体数据流进行融合, 得到数目更小、占用宽带更少的第二多媒体数据流; 再将得到的第二多媒体数据流发送给接收方。通过这种方法, 降低了对接收方的带宽要求, 提高了多人即时通讯中对接收方的兼容性。

例如: 在多人视频会议中, 5 个参会者都使用多媒体采集设备(例如: 麦克风、摄像头) 将自己的多媒体数据流(例如: 能够呈现出自己声音的音频数据流、能够呈现出自己影像的视频数据流) 上传至服务器——服务器会接收到 5 路多媒体数据流。服务器将这 5 路多媒体数据流融合为 1 路多媒体数据流, 再将这 1 路多媒体数据流发送给每个参会者。从而每个参会者都可以实时呈现出所有 5 个参会者分别想展示的多媒体(例如: 多媒体数据流对应的视频、多媒体数据流对应的音频), 而只需要接收这 1 路多媒体数据流。

需要说明的是, 该实施例只是示例性的说明, 不应对本公开的功能和使用范围造成限制。

图 2 示出了本公开一实施例的体系构架组成: 服务器 10、第一用户端 20、第二用户端 30。各个第一用户端 20 之间通过服务器 10 进行多媒体数据流的发送与接收, 各第二用户端 30 可以通过服务器 10 进行多媒体数据流的接收, 从而实现多人即时通讯。其中, 每个第一用户端 20 都可以作为发送方, 将自己所采集、生成的第一多媒体数据流发送给服务器 10; 每个第一用户端 20 与每个第二用户端 30 都可以作为接收方, 从服务器 10 处接收由多个第一用户端 20 所采集、生成的第一多媒体数据流融合而成的第二多媒体数据流。具体的, 服务器 10 可以接收每一第一用户端 20 发送的第一多媒体数据流; 将接收到的各第一多媒体数据流进行融合, 得到数目更小、占用宽带更少的第二多媒体数据流; 再将得到的第二多媒体数据流发送给接收方。通过这种方法, 降低了对接收方的带宽要求, 提高了多人即时通讯中对接收方的兼容性。

例如: 在多人直播的场景中, 5 个主播先后加入同一线上直播房间进行直播, 该线上直播房间内有 1000 个观众在观看该直播。在直播过程中, 这 5 个主播均通过音视频采集设备(麦克风、摄像头) 将自己的音视频数据流上传至服务器——服务器会接收到 5 路音视频数据流。服务器将这 5 路音视频数据流融合为 1 路音视频数据流, 再将这 1 路音视频数据流发送给每个主播、每个观众。从而这 5 个主播、这 1000 个观众都可以观看到所有 5 个主播的视频、听到所有 5 个主播的音频, 而只需要接收并播放这 1 路音视频数据流。

需要说明的是, 该实施例只是示例性的说明, 不应对本公开的功能和使用范围造成限制。

下面对本公开实施例所提出的多人即时通讯方法的具体实施过程进行详细描述。

如图 3 所示，一种多人即时通讯方法，包括：

步骤 S410、获取第一数目个的第一多媒体数据流；

步骤 S420、将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目；

步骤 S430、将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

本公开实施例中，在多人即时通讯中，获取到各发送方分别上传的第一多媒体数据流后，对获取到的第一多媒体数据流进行融合，得到数目更小的第二多媒体数据流，再将第二多媒体数据流发送给接收方。从而降低了接收方在接收数据时对带宽的要求，提高了对接收方的兼容性。

在步骤 S410 中，获取第一数目个的第一多媒体数据流。

本公开实施例中，在多人即时通讯中，每有一个发送方，服务器就会接收到对应的至少一个第一多媒体数据流。例如：在多人视频会议中，每有一个参会者，服务器就会接收到该参会者发送的至少一路多媒体数据流。

在一实施例中，获取第一数目个的第一多媒体数据流，包括：在与该第一数目个的发送方分别建立连接后，获取该第一数目个的发送方分别通过对应的连接所上传的对应的第一多媒体数据流。

该实施例中，每个发送方向服务器发送对应的一个第一多媒体数据流。具体的，服务器首先与多人即时通讯中的第一数目个的发送方分别建立连接。连接建立完成后，对应的发送方可以实时将其采集、生成的第一多媒体数据流上传至服务器。

例如：在多人视频会议中，5 个参会者首先分别与服务器建立连接，以进行数据的发送与接收。在会议过程中，每个参会者都会使用多媒体采集设备实时地将自己的多媒体数据流通过对应的连接上传至服务器——服务器获取到 5 路多媒体数据流，从而后续经由服务器对这 5 路多媒体数据流的处理，使得自己能够被其他参会者看到、听到。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在步骤 S420 中，将该第一数目个的该第一多媒体数据流融合为第二数目个的第二多媒体数据流，该第二数目小于该第一数目。

在步骤 S430 中，将该第二数目个的该第二多媒体数据流发送给接收方，以使得该接收方呈现对应的多媒体。

本公开实施例中，服务器获取到第一数目个的第一多媒体数据流后，将该第一数目个的该第一多媒体数据流融合为数目更小的第二多媒体数据流，即，第二数目个的第二多媒体数据流。由于第二数目小于第一数目，从而降低了对接收方的带宽要求，提高了对接收方的兼容性。

在一实施例中，将该第一数目个的该第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

从该第一多媒体数据流中分离出对应的第一音频数据流；

对分离出的各该第一音频数据流进行混音处理，获取该第二数目个的该第二多媒体数据流。

该实施例中，获取到的第一多媒体数据流可以只包括音频数据，也可以既包括音频数据又包括视频数据。与此同时，要发送给接收方的第二多媒体数据流只包括音频数据，相应的，接收方接收到该第二多媒体数据流之后，只能播放出对应的音频。

具体的，根据具体应用场景，若服务器事先被配置为只能发送音频数据，或者服务器判断该接收方只能接收音频数据，则，服务器获取到第一数目个的第一多媒体数据流后——从每个第一多媒体数据流中分离出对应的第一音频数据流，进而对分离出的各第一音频数据流进行混音处理，从而得到待发送给该接收方的、只包括音频数据的、第二数目个的第二多媒体数据流。

例如：在 A 直播平台上，5 个主播同时在线直播房间进行直播。与此同时，A 直播平台与 B 电台平台进行合作，进行引流——A 直播平台上的直播内容也能够同步地由 B 电台平台进行电台播放。

在该直播过程中，每个主播将自己的多媒体数据流（既包括音频数据，又包括能够呈现自己影像的视频数据）上传至 A 直播平台的 A 服务器的同时，也会将自己的多媒体数据流同步上传至 B 电台平台的 B 服务器。其中，由于 B 电台平台自身的业务需求，B 服务器事先被配置为只能发送音频数据。

B 服务器获取到对应的 5 路多媒体数据流后——从每个多媒体数据流中分离出对应的音频数据流，进而将这些分离出的音频数据流进行混音处理，得到待发送给接收方的、只包括音频数据的、1 路多媒体数据流。

又例如：在 A 直播平台上，5 个主播同时在线直播房间进行直播。在该直播过程中，每个主播将自己的多媒体数据流（既包括音频数据，又包括能够呈现自己影像的视频数据）上传至 A 直播平台的 A 服务器。

小明作为正在该线上直播房间中的一观众，出于自身原因暂时不便观看视频。因此，小明调整该线上直播房间的观看线路——选择“纯音频线路”。A 服务器接收到小明将观看线路调整为“纯音频线路”的信号后，即确定小明只能接收音频数据。

从而针对小明作为接收方的情况，A 服务器获取到对应的 5 路多媒体数据流后——从每个多媒体数据流中分离出对应的音频数据流，进而将这些分离出的音频数据流进行混音处理，得到待发送给接收方的、只包括音频数据的、1 路多媒体数据流。

该实施例的优点在于，通过将分离出的各第一音频数据流进行混音，进而获取更小数目的第二多媒体数据流，降低了对接收方带宽的要求。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在一实施例中，在对分离出的各第一音频数据进行混音处理之前，该方法还包括：若一该第一多媒体数据流来源于该接收方，将从该第一多媒体数据流中分离出的第一音频数据流进行剔除。

也就是说，对于目标多媒体数据流，若目标多媒体数据流来源于接收方，则将从目标多媒体数据流中分离出的第一音频数据流进行剔除；该目标多媒体数据流是第一数目个的第一多媒体数据流中的任意一个。

该实施例中，在对分离出的各第一音频数据进行混音处理之前，会针对每一接收方分别进行如下判断：获取到的各第一多媒体数据流中，是否有一第一多媒体数据流来源于该接收方。若有，则将从该第一多媒体数据流中分离的第一音频数据流进行剔除。从而使得该接收方接收到的第二多媒体数据流中不包括其自身的音频数据。

例如：5个主播同时同一直播房间进行直播。这5个主播中，每个主播在作为多媒体数据流的发送方的同时，也会作为多媒体数据流的接收方——接收来源于其他主播的多媒体数据流，以看到其他主播的视频、听到其他主播的声音。服务器获取到的5路多媒体数据流中，包括该主播的音频数据。显然，若该主播再从服务器接收自己的音频数据流、播放出自己的音频，会对该主播造成干扰。

因此，服务器获取到这5路多媒体数据流，并分离出各音频数据流后——针对特定主播，将从来源于该主播的多媒体数据流中分离出的音频数据流进行剔除，只对从其他4个主播的多媒体数据流中分离出的4路音频数据流进行混音，得到待发送给接收方的、只包括其他4个主播的音频数据的、1路多媒体数据流。从而使得该主播不会被自己的音频造成干扰。

该实施例的优点在于，若有一第一多媒体数据流来源于该接收方，则将从该第一多媒体数据流中分离的第一音频数据流进行剔除，从而使得接收方不会被自己的音频干扰。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在一实施例中，将该第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

从该第一多媒体数据流中分离出对应的第一视频数据流；

对分离出的各该第一视频数据流进行混屏处理，获取该第二数目个的第二多媒体数据流。

该实施例中，获取到的第一多媒体数据流可以只包括视频数据，也可以既包括音频数据又包括视频数据。与此同时，要发送给接收方的第二多媒体数据流只包括视频数据，相应的，接收方接收到该第二多媒体数据流之后，只能播放出对应的视频。

具体的，根据具体应用场景，若服务器事先被配置为只能发送视频数据，或者服务器判断该接收方只能接收视频数据，则，服务器获取到第一数目个的第一多媒体数据流后——从每个第一多媒体数据流中分离出对应的第一视频数

据流，进而对分配出的各第一视频数据流进行混屏处理，从而得到待发送给该接收方的、只包括视频数据的、第二数目个的第二多媒体数据流。

例如：在 A 直播平台上，5 个主播同时在线直播房间进行直播。在该直播过程中，每个主播将自己的多媒体数据流（既包括能够呈现自己声音的音频数据，又包括能够呈现自己影像的视频数据）上传至 A 直播平台的 A 服务器。

小明作为正在该线上直播房间中的一观众，出于直播内容的原因（例如：该直播内容为这 5 个主播协同进行绘画创作）、或者自身原因（例如：节省数据流量），只需要观看视频。因此，小明调整该线上直播房间的观看线路——选择“纯音频线路”。A 服务器接收到小明将观看线路调整为“纯视频线路”的信号后，即确定小明只能接收视频数据。

从而针对小明作为接收方的情况，A 服务器获取到对应的 5 路多媒体数据流后——从每个多媒体数据流中分离出对应的视频数据流，进而将这些分离出的视频数据流进行混屏处理，得到待发送给接收方的、只包括视频数据的、1 路多媒体数据流。从而小明的终端（例如：手机、电脑）接收到这 1 路多媒体数据流后，能够在终端屏幕上的 5 个区域（例如：屏幕右上角预定大小的方框区域、屏幕右下角预定大小的方框区域、屏幕左上角预定大小的方框区域、屏幕左下角预定大小的方框区域、屏幕中间预定大小的方框区域）分别显示出对应主播的视频，供小明观看。

该实施例的优点在于，通过将分离出的各第一视频数据流进行混屏，进而获取更小数目的第二多媒体数据流，降低了对接收方带宽的要求。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在一实施例中，将该第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

从该第一多媒体数据流中分离出对应的第一音频数据流；

从该第一多媒体数据流中分离出对应的第一视频数据流；

对分离出的各第一音频数据流进行混音处理；

对分离出的各第一视频数据流进行混屏处理；

基于混音处理的各第一音频数据流与混屏处理的各第一视频数据流，获取该第二数目个的第二多媒体数据流。

该实施例中，获取到的第一多媒体数据流既包括音频数据又包括视频数据。与此同时，要发送给接收方的第二多媒体数据流也是既包括音频数据又包括视频数据，相应的，接收方接收到该第二多媒体数据流之后，能够播放出对应的音频与视频。

具体的，服务器获取到第一数目个的第一多媒体数据流后——从每个第一多媒体数据流中分离出对应的第一音频数据流，进而对分配出的各第一音频数据流进行混音处理；从每个第一多媒体数据流中分离出对应的第一视频数据流，进而对分离出的各第一视频数据流进行混屏处理；进而基于分离出的各第一音

频数据流与分离出的各第一视频数据流，从而得到待发送给该接收方的、既包括音频数据又包括视频数据的、第二数目个的第二多媒体数据流（例如：将分离出的各第一音频数据流与分离出的各第一视频数据流整合进一个数据包中，从而得到待发送给该接收方的、既包括音频数据又包括视频数据的、1个第二多媒体数据流；或者，将分离出的各第一音频数据流整合进一个数据包中，将分离出的各第一视频数据流整合进另一个数据包中，从而得到待发送给该接收方的、既包括音频数据又包括视频数据的、2个第二多媒体数据流）。

例如：在A直播平台上，5个主播同时在线直播房间进行直播。在该直播过程中，每个主播将自己的多媒体数据流（既包括能够呈现自己声音的音频数据，又包括能够呈现自己影像的视频数据）上传至A直播平台的A服务器。

A服务器获取到对应的5路多媒体数据流后——从每个多媒体数据流中分离出对应的音频数据流，进而将这些分离出的音频数据流进行混音处理；从每个多媒体数据流中分离出对应的视频数据流，进而将这些分离出的视频数据流进行混屏处理；将混音处理的这些音频数据流与混屏处理的这些视频数据流整合进一个数据包中，得到待发送给接收方的、既包括音频数据又包括视频数据的、1路多媒体数据流。

从而小明的终端接收到这1路多媒体数据流后，能够在终端屏幕上的5个区域分别显示出对应主播的视频；并播放出这5个主播的音频。

该实施例的优点在于，通过将分离出的各第一音频数据流进行混音，将分离出的各第一视频数据流进行混屏，进而获取更小数目的第二多媒体数据流，降低了对接收方带宽的要求。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在一实施例中，将该第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

基于属性配置信息，将该第一数目个的第一多媒体数据流转换为匹配该属性配置信息的该第一数目个的第一多媒体数据流；

将匹配该属性配置信息的该第一数目个的第一多媒体数据流，融合为匹配该属性配置信息的第二数目个的第二多媒体数据流；

其中，在将该第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流之前，获取该接收方针对多媒体播放预设的该属性配置信息。

其中，上述属性配置信息是用于指示对应的多媒体数据流的呈现属性的信息，以多媒体数据流是视频数据流为例，该呈现属性包括视频分辨率、视频码率、视频显示模版等。

该实施例中，接收方可以针对多媒体播放，预设相应的属性配置信息（例如：视频分辨率、视频码率、视频显示模板）。服务器在将第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流之前，还会获取接收方对应的该属性配置信息。

从而，服务器在将第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流时——基于该属性配置信息，对第一数目个的第一多媒体数据流进行属性转换，转换为匹配该属性配置信息的第一数目个的第一多媒体数据流；再将匹配该属性配置信息的第一数目个的第一多媒体数据流，融合为匹配该属性配置信息的第二数目个的第二多媒体数据流。进而服务器可以将匹配该属性配置信息的第二数目个的第二多媒体数据流发送给接收方，使得接收方能够呈现出匹配该属性配置信息的多媒体。

例如：在 A 直播平台上，5 个主播同时在同一直播房间进行直播。在该直播过程中，每个主播将自己的多媒体数据流上传至 A 直播平台的 A 服务器。

小明作为正在该线上直播房间中的一观众，由于自身原因（例如：网速慢、手机能显示的最大分辨率低），小明预设针对视频播放配置“视频分辨率：480P”，从而 A 服务器预先获取到小明预设的该属性配置信息——“视频分辨率：480P”。

从而针对小明作为接收方的情况，A 服务器获取到对应的 5 路多媒体数据流后——将这 5 路多媒体数据流转换为视频分辨率为 480P 的、5 路多媒体数据流；进而将这视频分辨率为 480P 的、5 路多媒体数据流融合为待发送给小明的、视频分辨率为 480P 的、1 路多媒体数据流。从而小明的终端在接收到这 1 路多媒体数据流后，能够播放出视频分辨率为 480P 的多媒体。

该实施例的优点在于，通过获取接收方预设的属性配置信息，进而对多媒体数据流进行相应处理，避免了接收方对接收到的第二多媒体数据流进行处理的过程，降低了接收方的处理负担。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在一实施例中，在将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流之后，包括：基于属性配置信息，将该第二数目个的该第二多媒体数据流转换为匹配该属性配置信息的该第二数目个的该第二多媒体数据流；其中，在将该第二数目个的该第二多媒体数据流发送给接收方之前，获取该接收方针对多媒体播放的属性配置信息。

该实施例中，接收方可以针对多媒体播放，预设相应的属性配置信息（例如：视频分辨率、视频码率、视频显示模板）。服务器在将第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流之前，还会获取接收方对应的该属性配置信息。

从而，服务器在将第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流之后——基于该属性配置新，对第二数目个的第二多媒体数据流进行属性转换，转换为匹配该属性配置信息的第二数目个的第二多媒体数据流。进而服务器可以将匹配该属性配置信息的第二数目个的第二多媒体数据流发送给接收方，使得接收方能够呈现出匹配该属性配置信息的多媒体。

例如：在 A 直播平台上，5 个主播同时在同一直播房间进行直播。在该直播过程中，每个主播将自己的多媒体数据流上传至 A 直播平台的 A 服务器。

小明作为正在该线上直播房间中的一观众，由于自身原因（例如：网速慢、手机能显示的最大分辨率低），小明预设针对视频播放配置“视频分辨率：480P”，

从而A服务器预先获取到小明预设的该属性配置信息——“视频分辨率:480P”。

从而针对小明作为接收方的情况，A服务器将获取到对应的5路多媒体数据流融合为待发送给小明的、1路多媒体数据流后——将这1路多媒体数据流转换为待发送给小明的、视频分辨率为480P的、1路多媒体数据流。从而小明的终端在接收到这1路多媒体数据流后，能够播放出视频分辨率为480P的多媒体。

该实施例的优点在于，通过获取接收方预设的属性配置信息，进而对多媒体数据流进行相应处理，避免了接收方对接收到的第二多媒体数据流进行处理的过程，降低了接收方的处理负担。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

本公开实施例还提出了一种多人即时通讯系统。

如图4所示，本公开一实施例中的多人即时通讯系统包括：调度节点101、至少两个候选数据适配节点102。其中，该调度节点101配置为从该至少两个候选数据适配节点102中选取一候选数据适配节点，作为目标数据适配节点。

该目标数据适配节点配置为获取第一数目个的第一多媒体数据流，将该第一数目个的该第一多媒体数据流融合为第二数目个的第二多媒体数据流，该第二数目小于该第一数目，将该第二数目个的该第二多媒体数据流发送给接收方，以使得该接收方呈现对应的多媒体。

本公开实施例所提出的该多人即时通讯系统是建立在分布式系统的基础上得到的一种MCU（Multipoint Control Unit，多点控制单元）系统——用于进行多人即时通讯中数据处理（获取第一数目个的第一多媒体数据流；将该第一数目个的该第一多媒体数据流融合为第二数目个的第二多媒体数据流，该第二数目小于该第一数目；将该第二数目个的该第二多媒体数据流发送给接收方，以使得该接收方呈现对应的多媒体）的目标数据适配节点是从分布设置的各候选数据适配节点中选取的。其中，该多人即时通讯系统中的每一节点可以分别设置在位于不同物理位置的机器上。例如：调度节点设置在位于深圳的一台服务器上，其中一个候选数据适配节点设置在位于广州的一台服务器上，另一个候选数据适配节点设置在位于杭州的一台服务器上。

该多人即时通讯系统通过融合第一数目个第一多媒体数据流，将更小数目的第二数目个第二多媒体数据流发送给接收方，降低了接收方对带宽的要求，提高了对接收方的兼容性；同时，该多人即时通讯系统通过建立在分布式系统之上，多人即时通讯系统的各部分处理过程可以灵活分配给各个分布式节点，提高了多人即时通讯系统的处理过程的灵活性；而且，该多人即时通讯系统通过建立在分布式系统之上，提供了强大的平行拓展能力，能够轻松支持十万级、百万级用户在同一通讯房间内得到多人即时通讯的服务。

在一实施例中，该调度节点配置为：

定期获取该各候选数据适配节点分别对应的工作状态信息；

基于该工作状态信息，分别获取该各候选数据适配节点对应的状态分数；

基于该状态分数，从该各候选数据适配节点中选取目标数据适配节点。

该实施例中，调度节点会定期获取每一候选数据适配节点对应的工作状态信息（例如：该候选数据适配节点的负载信息）。当需要选取目标数据适配节点时（例如：当多个主播要在同一个线上直播房间中进行直播时，调度节点要选取目标数据适配节点，从而由该目标数据适配节点进行第一数目个的第一多媒体数据流的获取、将第一数目个的第一多媒体数据流融合为第二数目个的第二多媒体数据流、将第二数目个的第二多媒体数据流发送给接收方），调度节点即可以基于候选数据适配节点的工作状态信息，衡量候选数据适配节点工作状态的优劣——具体体现为状态分数。进而基于状态分数，从各候选数据适配节点中选取目标数据适配节点。

其中，基于工作状态信息，获取候选数据适配节点的状态分数——可以通过对工作状态信息中各项工作属性的表现分别进行衡量，得到每一工作属性对应的属性分数；进而将各项属性分数相加，或者根据预先为各项工作属性分配的权重对各项属性分数进行加权，从而得到对应的状态分数。

其中，基于状态分数从各候选数据适配节点中选取目标数据适配节点——可以将状态分数最大的候选数据适配节点选取为目标数据适配节点；也可以从按照从高到低的顺序状态分数排在前三位的候选数据适配节点中，随机选取一个作为目标数据适配节点。

该实施例的优点在于，基于状态分数选取目标数据适配节点，能够灵活地保证选取出的目标数据适配节点的处理效率。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在一实施例中，调度节点通过区块链系统获取各候选数据适配节点分别对应的工作状态信息。例如，该实施例中，各候选数据适配节点均分别作为区块链系统中的一区块链节点，基于预设的共识算法定期对各候选数据适配节点的工作状态信息进行共识、打包为一个数据区块、将该数据区块进行上链（将该数据区块链入区块链）。从而调度节点可以通过查询区块链中最新的数据区块，获取到各候选数据视频分别对应的工作状态信息。

其中，区块链系统是指用于进行节点与节点之间数据共享的系统。参见图 5A 所示的区块链系统 50，该区块链系统中可以包括多个节点 501，多个节点 501 可以是指区块链系统中各个客户端。每个节点 501 在进行正常工作可以接收到输入信息，并基于接收到的输入信息维护该区块链系统内的共享数据。为了保证区块链系统内的信息互通，区块链系统中的每个节点之间可以存在信息连接，节点之间可以通过上述信息连接进行信息传输。例如，当区块链系统中的任意节点接收到输入信息时，区块链系统中的其他节点便根据共识算法获取该输入信息，将该输入信息作为共享数据中的数据进行存储，使得区块链系统中全部节点上存储的数据均一致。

对于区块链系统中的每个节点，均具有与其对应的节点标识，而且区块链系统中的每个节点均可以存储有区块链系统中其他节点的节点标识，以便后续根据其他节点的节点标识，将生成的区块广播至区块链系统中的其他节点。每

个节点中可维护一个如下表所示的节点标识列表，将节点名称和节点标识对应存储至该节点标识列表中。其中，节点标识可为 IP（Internet Protocol，网络之间互联的协议）地址以及其他任一种能够用于标识该节点的信息，表 1 中仅以 IP 地址为例进行说明。

节点名称	节点标识
节点 1	117.114.151.174
节点 2	117.116.189.145
...	...
节点 N	119.123.789.258

表 1. 节点标识列表

区块链系统中的每个节点均存储一条相同的区块链。区块链由多个区块组成，参见图 5B，区块链由多个区块组成，创始块中包括区块头和区块主体，区块头中存储有输入信息特征值、版本号、时间戳和难度值，区块主体中存储有输入信息；创始块的下一区块以创始块为父区块，下一区块中同样包括区块头和区块主体，区块头中存储有当前区块的输入信息特征值、父区块的区块头特征值、版本号、时间戳和难度值，并以此类推，使得区块链中每个区块中存储的区块数据均与父区块中存储的区块数据存在关联，保证了区块中输入信息的安全性。

在生成区块链中的各个区块时，参见图 5C，区块链所在的节点在接收到输入信息时，对输入信息进行校验，完成校验后，将输入信息存储至内存池中，并更新其用于记录输入信息的哈希树；之后，将更新时间戳更新为接收到输入信息的时间，并尝试不同的随机数，多次进行特征值计算，使得计算得到的特征值可以满足下述公式：

$$SHA256(SHA256(version + prev_hash + merkle_root + ntime + nbits + x)) < TARGET$$

其中，SHA256 为计算特征值所用的特征值算法；version（版本号）为区块链中相关区块协议的版本信息；prev_hash 为当前区块的父区块的区块头特征值；merkle_root 为输入信息的特征值；ntime 为更新时间戳的更新时间；nbits 为当前难度，在一段时间内为定值，并在超出固定时间段后再次进行确定；x 为随机数；TARGET 为特征值阈值，该特征值阈值可以根据 nbits 确定得到。

这样，当计算得到满足上述公式的随机数时，便可将信息对应存储，生成区块头和区块主体，得到当前区块。随后，区块链所在节点根据区块链系统中其他节点的节点标识，将新生成的区块分别发送给其所在的区块链系统中的其他节点，由其他节点对新生成的区块进行校验，并在完成校验后将新生成的区块添加至其存储的区块链中。

该实施例的优点在于，通过区块链系统管理各候选数据适配节点的工作状态信息，提高了工作状态信息管理的透明度与安全性，同时也使得调度节点能够更便捷地获取各候选数据适配节点的工作状态信息。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

如图6所示,在一实施例中,该系统还包括至少两个候选音频处理节点103,该调度节点101配置为:从该至少两个候选音频处理节点103中选取一候选音频处理节点103,作为目标音频处理节点。

候选音频处理节点指的是待选取的、可以进行音频处理的节点;目标音频处理节点指的是确定选取的、用以进行音频处理的节点。

该实施例中,多人即时通讯系统中除了包括调度节点101、至少两个候选数据适配节点102外,还包括至少两个候选音频处理节点103。其中,候选音频处理阶段103配置为候选进行音频处理;调度节点101配置为从该至少两个候选音频处理节点103中选取一候选音频处理阶段,作为目标音频处理节点。调度节点101、至少两个候选数据适配节点102与至少两个候选音频处理节点103可以合并看作服务器10。

可以理解的,对于目标音频处理节点的选取及相应优点,与对于目标数据适配节点的选取及相应优点同理,故在此不再赘述。需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

在一实施例中,该目标数据适配节点配置为从该第一多媒体数据流中分离出对应的第一音频数据流。

该目标音频处理节点配置为对分离出的各该第一音频数据流进行混音处理,将混音处理的各所述第一音频数据流发送给所述目标数据适配节点,以使得该目标数据适配节点获取该第二数目个的第二多媒体数据流。

该实施例中,包括了至少两个候选音频处理节点的多人即时通讯系统,其中选取出的目标数据适配节点配置为从第一多媒体数据流中分离出对应的第一音频数据流,进而由选取出的目标音频处理节点对分离出的各第一音频数据流进行混音处理,将混音处理的各第一音频数据流发送给目标数据适配节点,从而使得目标数据适配节点获取第二数目个的第二多媒体数据流。

该实施例的优点在于,通过将分离出的各第一音频数据流进行混音,进而获取更小数目的第二多媒体数据流,降低了对接收方带宽的要求。

需要说明的是,该实施例只是示例性的说明,不应对本公开的功能和使用范围造成限制。

在一实施例中,该目标音频处理节点配置为在对分离出的各所述第一音频数据流进行混音处理之前,若一该第一多媒体数据流来源于该接收方,将从该第一多媒体数据流中分离出的第一音频数据流进行剔除。

该实施例中,包括了至少两个候选音频处理节点的多人即时通讯系统,其中选取出的目标音频处理节点除了对分离出的各第一音频数据流进行混音处理,将混音处理的各第一音频数据流发送给目标数据适配节点之外,还会在对分离出的各所述第一音频数据流进行混音处理之前,将从来源于接收方的第一多媒体数据流中分离出的第一音频数据流进行剔除,从而使得该接收方接收到的第二多媒体数据流中不包括其自身的音频数据。

该实施例的优点在于,若有一第一多媒体数据流来源于该接收方,则将从该第一多媒体数据流中分离的第一音频数据流进行剔除,从而使得接收方不会被自己的音频干扰。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

如图7所示，在一实施例中，该系统还包括至少两个候选视频处理节点104，该调度节点101配置为：从该至少两个候选视频处理节点104中选取一候选视频处理节点104，作为目标视频处理节点。

候选视频处理节点指的是待选取的、可以进行视频处理的节点；目标视频处理节点指的是确定选取的、用以进行视频处理的节点。

该实施例中，多人即时通讯系统中除了包括调度节点101、至少两个候选数据适配节点102外，还包括至少两个候选视频处理节点103。其中，候选视频处理阶段103配置为候选进行视频处理；调度节点101配置为从该至少两个候选视频处理节点103中选取一候选视频处理节点，作为目标视频处理节点。调度节点101、至少两个候选数据适配节点102与至少两个候选视频处理节点103可以合并看作服务器10。

可以理解的，对于目标视频处理节点的选取及相应优点，与对于目标数据适配节点的选取及相应优点同理，故在此不再赘述。需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

在一实施例中，该目标数据适配节点配置为从该第一多媒体数据流中分离出对应的第一视频数据流；

该目标视频处理节点配置为对分离出的各该第一视频数据流进行混屏处理，将混屏处理的各所述第一视频数据流发送给所述目标数据适配节点，以使得该目标数据适配节点获取该第二数目个的第二多媒体数据流。

该实施例中，包括了至少两个候选视频处理节点的多人即时通讯系统，其中选取出的目标数据适配节点配置为从第一多媒体数据流中分离出对应的第一视频数据流，进而由选取出的目标视频处理节点对分离出的各第一视频数据流进行混屏处理，将混屏处理的各第一视频数据流发送给目标数据适配节点，从而使得目标数据适配节点获取第二数目个的第二多媒体数据流。

该实施例的优点在于，通过将分离出的各第一视频数据流进行混屏，进而获取更小数目的第二多媒体数据流，降低了对接收方带宽的要求。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

如图8所示，在一实施例中，该系统还包括至少两个候选音频处理节点103，至少两个候选视频处理节点104，该调度节点101配置为：

从该至少两个候选音频处理节点103中选取一候选音频处理节点103，作为目标音频处理节点；

从该至少两个候选视频处理节点104中选取一候选视频处理节点104，作为目标视频处理节点。

该实施例中，当需要进行音频处理时（例如：目标数据适配节点分离出各第一音频数据流，需要对各第一音频数据流进行诸如混音、音频格式转换之类

的处理时), 调度节点从各候选音频处理节点中选取目标音频处理节点; 当需要进行视频处理时 (例如: 目标数据适配节点分离出各第一视频数据流, 需要对各第一视频数据流进行诸如混屏、视频码率转换之类的处理时), 调度节点从各候选视频处理节点中选取目标视频处理节点。

需要说明的是, 该实施例只是示例性的说明, 不应对本公开的功能和使用范围造成限制。

在一实施例中, 该目标数据适配节点配置为:

从该第一多媒体数据流中分离出对应的第一音频数据流;

从该第一多媒体数据流中分离出对应的第一视频数据流;

该目标音频处理节点配置为对分离出的各第一音频数据流进行混音处理, 将混音处理的各所述第一音频数据流发送给所述目标数据适配节点, 以使得该目标数据适配节点获取该第二数目个的第二多媒体数据流;

该目标视频处理节点配置为对分离出的各第一视频数据流进行混屏处理, 将混屏处理的各所述第一视频数据流发送给所述目标数据适配节点, 以使得该目标数据适配节点获取该第二数目个的第二多媒体数据流;

该目标数据适配节点配置为基于混音处理的各第一音频数据流与混屏处理的各第一视频数据流, 获取该第二数目个的第二多媒体数据流。

该实施例中, 目标数据适配节点负责从第一多媒体数据流中分离出对应的第一音频数据流与对应的第一视频数据流。之后, 目标音频处理节点负责对分离出的各第一音频数据流进行混音处理, 将混音处理的各所述第一音频数据流发送给所述目标数据适配节点; 目标视频处理节点负责对分离出的各第一视频数据流进行混音处理, 将混音处理的各所述第一视频数据流发送给所述目标数据适配节点。再由目标数据适配节点负责基于混音处理的各第一音频数据流与混屏处理的各第一视频数据流, 获取第二数目个的第二多媒体数据流。

该实施例的优点在于, 通过将分离出的各第一音频数据流进行混音, 将分离出的各第一视频数据流进行混屏, 进而获取更小数目的第二多媒体数据流, 降低了对接收方带宽的要求。

需要说明的是, 该实施例只是示例性的说明, 不应对本公开的功能和使用范围造成限制。

下面参考图 9 到图 13 对本公开一实施例在直播场景中的完整实施过程进行详细描述。

图 9 示出了本公开一实施例在直播场景中的初始处理过程: 在已建立的直播房间中, 存在两个模块——流控机与接口机。其中, 流控机用于直播进行的过程中, 对该直播房间内的音视频参数实时调节; 接口机用于与外部系统 (例如: 服务器 10) 交换数据, 以及与流控机之间进行信令的传输。

该实施例中, 服务器 10 中的调度节点会定期获取各候选数据适配节点的负载信息、位置信息; 定期获取各候选音频处理节点的负载信息、位置信息; 定期获取各候选视频处理节点的负载信息、位置信息。以根据获取到的负载信息、位置信息, 于后续流程中进行目标数据适配节点的选取、目标音频处理节点的

选取、目标视频处理节点的选取。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

图 10 示出了本公开一实施例在直播场景中申请混流的处理过程：其中，混流指的是当多个主播加入直播房间后，将多个主播的多媒体数据流上传至服务器 10 的过程；相应的，申请混流即是在多个主播加入直播房间，要开启直播时触发的。

例如，该实施例中，第 1 步：流控机向调度节点申请混流。

第 2 步：调度节点根据全局的负载信息、位置信息，选取出处理申请混流过程的目标数据适配节点后，向目标数据适配节点发送端口分配请求。

第 3 步：目标数据适配节点在本地建立会话，分配空闲端口。并将分配的端口发送给调度节点。

第 4 步：调度节点将目标数据适配节点的 IP、目标数据适配节点的端口发送给流控机。

第 5 步：流控机将目标数据适配节点的 IP、目标数据适配节点的端口与接口机进行同步。

第 6 步：接口机根据接收到的目标数据适配节点的 IP、目标数据适配节点的端口，与目标数据适配节点建立连接。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

图 11 示出了本公开一实施例在直播场景中申请观看的处理过程：其中，主播可以触发申请观看，观众也可以触发申请观看。

例如，该实施例中，第 1 步：流控机向调度节点发送申请观看请求，并附带属性配置信息——例如音频编码格式、声道数、音频码率、视频编码格式、视频分辨率、视频码率、视频显示模板。

第 2 步：调度节点根据全局的负载信息、位置信息，选取出处理申请观看过程的目标数据适配节点、目标音频处理节点、目标视频处理节点，并向目标数据适配节点、目标音频处理节点、目标视频处理节点均发送端口分配请求。

第 3 步：目标数据适配节点、目标音频处理节点、目标视频处理节点均在对应的本地建立会话（目标音频处理节点、目标视频处理节点在对应的本地建立会话时，保存属性配置信息），分配对应的空闲端口。并将分配的端口发送给调度节点。如果申请观看的用户在该目标数据适配节点上已经有申请混流的会话，则此时目标数据适配节点可以复用该申请混流的会话；并且如果申请观看的用户是混流和观看都需要的（例如：该直播房间中的主播），则针对该用户，可以将申请混流请求与申请观看请求合并为同一个请求。

第 4 步：调度节点将目标数据适配节点的 IP、目标数据适配节点的端口发送给流控机。

第 5 步：流控机将目标数据适配节点的 IP、目标数据适配节点的端口与接口机进行同步。

第 6 步：接口机根据接收到的目标数据适配节点的 IP、目标数据适配节点

的端口，与目标数据适配节点建立连接。

第 7 步：调度节点将目标音频处理节点的 IP、目标音频处理节点的端口、目标视频处理节点的 IP、目标视频处理节点的端口发送给目标数据适配节点。

第 8 步：目标数据适配节点根据接收到的目标音频处理节点的 IP、目标音频处理节点的端口、目标视频处理节点的 IP、目标视频处理节点的端口，与目标音频处理节点建立连接，与目标视频处理节点建立连接。

需要说明的是，该实施例中，第 4、5、6 步可以与第 7、8 步并行执行。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

图 12 示出了本公开一实施例在直播场景中上行数据的处理过程：该实施例中，有三个用户：主播 A、主播 B、观众 C。其中，只有主播 A 与主播 B 能够触发上行数据的处理过程。

主播 A 和主播 B 在直播过程中，会将对应的多媒体数据流上传。例如，主播 A 将其音视频数据流发送给目标数据适配节点；主播 B 将其音视频数据流发送给目标数据适配节点。

目标数据适配节点接收到主播 A 与主播 B 发送的多媒体数据流后，针对主播 A——分离出主播 A 的音频、主播 A 的视频；针对主播 B——分离出主播 B 的音频、主播 B 的视频。进而将主播 A 的音频、主播 B 的音频发送给目标音频处理节点进行处理（目标音频处理节点根据属性配置信息，对音频进行相应处理）；将主播 A 的视频、主播 B 的视频发送给目标视频处理节点进行处理（目标视频处理节点根据属性配置信息，对视频进行相应处理）。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

图 13 示出了本公开一实施例在直播场景中下行数据的处理过程：该实施例中，有三个用户：主播 A、主播 B、观众 C。其中，任一用户均能够触发下行数据的处理过程。

在直播过程中，任一用户申请观看通过后，均会触发相应的下行数据的处理过程。

对于主播 A，只需要听主播 B 的音频、看主播 A 和主播 B 共同的视频（视频的显示布局可以定制）。目标音频处理节点对音频进行选路（剔除掉 A 的音频）以及根据属性配置信息进行相应处理，将主播 B 的音频发送给目标数据适配节点，目标数据适配节点按照相应的数据传输协议将主播 B 的音频传输给接口机，再由接口机转发至主播 A。目标视频处理节点将主播 A 的视频与主播 B 的视频进行混屏以及根据属性配置信息进行相应处理后，发送给目标数据适配节点，目标数据适配节点将视频帧切分为小于网络的最大传输单元大小的帧并按照相应的数据传输协议进行封装，传输给接口机，再由接口机转发至主播 A。

对于主播 B，只需要听主播 A 的音频、看主播 A 和主播 B 共同的视频（视频的显示布局可以定制）。目标音频处理节点对音频进行选路（剔除掉主播 B 的音频）以及根据属性配置信息进行相应处理，将主播 A 的音频发送给目标数据适配节点，目标数据适配节点按照相应的数据传输协议将主播 A 的音频传输给

接口机，再由接口机转发至主播 A。目标视频处理节点将主播 A 的视频与主播 B 的视频进行混屏以及根据属性配置信息进行相应处理后，发送给目标数据适配节点，目标数据适配节点将视频帧切分为小于网络的最大传输单元大小的帧并按照相应的数据传输协议进行封装，传输给接口机，再由接口机转发至主播 B。

对于观众 C，需要听主播 A 和主播 B 的音频、看看主播 A 和主播 B 共同的视频（视频的显示布局可以定制）。目标音频处理节点将主播 A 的音频和主播 B 的音频混成一路以及根据属性配置信息进行相应处理后发送给目标数据适配节点，目标数据适配节点再发给接口机，再有接口机转发给观众 C。目标适配处理节点将主播 A 的视频和主播 B 的适配进行混屏以及根据属性配置信息进行相应处理后，发送给目标数据适配节点，目标数据适配节点将视频帧切分为小于网络的最大传输单元大小的帧并按照相应的数据传输协议进行封装，传输给接口机，再由接口机转发至观众 C。

需要说明的是，该实施例只是示例性的说明，不应对本公开的功能和使用范围造成限制。

根据本公开一实施例，如图 14 所示，还提供了一种多人即时通讯装置，所述装置包括：

获取模块 610，配置为获取第一数目个的第一多媒体数据流；

融合模块 620，配置为将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目；

发送模块 630，配置为将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

在本公开的一示例性实施例中，所述融合模块 620 配置为：

基于属性配置信息，将所述第一数目个的第一多媒体数据流转换为匹配所述属性配置信息的所述第一数目个的第一多媒体数据流；

将匹配所述属性配置信息的所述第一数目个的第一多媒体数据流，融合为匹配所述属性配置信息的第二数目个的第二多媒体数据流，其中，在将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流之前，获取所述接收方针对多媒体播放预设的所述属性配置信息。

在本公开的一示例性实施例中，所述融合模块 620 配置为：

从所述第一多媒体数据流中分离出对应的第一音频数据流；

对分离出的各所述第一音频数据流进行混音处理，获取所述第二数目个的所述第二多媒体数据流。

在本公开的一示例性实施例中，所述融合模块 620 配置为：若一所述第一多媒体数据流来源于所述接收方，将从所述第一多媒体数据流中分离出的第一音频数据流进行剔除。

也就是说，融合模块 620 配置为：对于目标多媒体数据流，若目标多媒体数据流来源于接收方，则将从目标多媒体数据流中分离出的第一音频数据流进行剔除；该目标多媒体数据流是第一数目个的第一多媒体数据流中的任意一个。

在本公开的一示例性实施例中，所述融合模块 620 配置为：

从所述第一多媒体数据流中分离出对应的第一视频数据流；

对分离出的各所述第一视频数据流进行混屏处理，获取所述第二数目个的所述第二多媒体数据流。

在本公开的一示例性实施例中，所述融合模块 620 配置为：

从所述第一多媒体数据流中分离出对应的第一音频数据流；

从所述第一多媒体数据流中分离出对应的第一视频数据流；

对分离出的各第一音频数据流进行混音处理；

对分离出的各第一视频数据流进行混屏处理；

基于混音处理的各第一音频数据流与混屏处理的各第一视频数据流，获取所述第二数目个的所述第二多媒体数据流。

下面参考图 15 来描述根据本公开实施例的多人即时通讯电子设备 70。图 15 显示的多人即时通讯电子设备 70 仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

如图 15 所示，多人即时通讯电子设备 70 以通用计算设备的形式表现。多人即时通讯电子设备 70 的组件可以包括但不限于：上述至少一个处理单元 710、上述至少一个存储单元 720、连接不同系统组件（包括存储单元 720 和处理单元 710）的总线 730。

其中，所述存储单元存储有程序代码，所述程序代码可以被所述处理单元 710 执行，使得所述处理单元 710 执行本说明书上述示例性方法的描述部分中描述的根据本公开各种示例性实施方式的步骤。例如，所述处理单元 710 可以执行如图 3 中所示的各个步骤。

存储单元 720 可以包括易失性存储单元形式的可读介质，例如随机存取存储单元（RAM）7201 和/或高速缓存存储单元 7202，还可以进一步包括只读存储单元（ROM）7203。

存储单元 720 还可以包括具有一组（至少一个）程序模块 7205 的程序/实用工具 7204，这样的程序模块 7205 包括但不限于：操作系统、一个或者多个应用程序、其它程序模块以及程序数据，这些示例中的每一个或某种组合中可能包括网络环境的实现。

总线 730 可以为表示几类总线结构中的一种或多种，包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。

多人即时通讯电子设备 70 也可以与一个或多个外部设备 800（例如键盘、指向设备、蓝牙设备等）通信，还可与一个或者多个使得用户能与该多人即时通讯电子设备 70 交互的设备通信，和/或与使得该多人即时通讯电子设备 70 能与一个或多个其它计算设备进行通信的任何设备（例如路由器、调制解调器等）通信。这种通信可以通过输入/输出（I/O）接口 750 进行。输入/输出（I/O）接口 750 与显示单元 740 相连。并且，多人即时通讯电子设备 70 还可以通过网络适配器 760 与一个或者多个网络（例如局域网（LAN），广域网（WAN）和/或公共网络，例如因特网）通信。如图所示，网络适配器 760 通过总线 730 与

多人即时通讯电子设备 70 的其它模块通信。应当明白，尽管图中未示出，可以结合多人即时通讯电子设备 70 使用其它硬件和/或软件模块，包括但不限于：微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID 系统、磁带驱动器以及数据备份存储系统等。

通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本公开实施方式的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是 CD-ROM，U 盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、终端装置、或者网络设备等）执行根据本公开实施方式的方法。

在本公开的示例性实施例中，还提供了一种计算机可读存储介质，其上存储有计算机可读指令，当所述计算机可读指令被计算机的处理器执行时，使计算机执行上述方法实施例部分描述的方法。

根据本公开的一个实施例，还提供了一种用于实现上述方法实施例中的方法的程序产品，其可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码，并可以在终端设备，例如个人电脑上运行。然而，本公开的程序产品不限于此，在本文件中，可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

所述程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。可读存储介质的更具体的例子（非穷举的列表）包括：具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器（RGM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。

计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了可读程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质，该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于无线、有线、光缆、RF 等等，或者上述的任意合适的组合。

可以以一种或多种程序设计语言的任意组合来编写用于执行本公开操作的程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如 JAVA、C++ 等，还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中，远程计算设备可以通过任意种类的网络，包括局域网（KGN）或广域网（WGN），

连接到用户计算设备，或者，可以连接到外部计算设备（例如利用因特网服务提供商来通过因特网连接）。

应当注意，尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元，但是这种划分并非强制性的。实际上，根据本公开的实施方式，上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之，上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。

此外，尽管在附图中以特定顺序描述了本公开中方法的各个步骤，但是，这并非要求或者暗示必须按照该特定顺序来执行这些步骤，或是必须执行全部所示的步骤才能实现期望的结果。附加的或备选的，可以省略某些步骤，将多个步骤合并为一个步骤执行，以及/或者将一个步骤分解为多个步骤执行等。

通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本公开实施方式的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是 CD-ROM，U 盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、移动终端、或者网络设备等）执行根据本公开实施方式的方法。

本领域技术人员在考虑说明书及实践这里公开的发明后，将容易想到本公开的其它实施方案。本申请旨在涵盖本公开的任何变型、用途或者适应性变化，这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的，本公开的真正范围和精神由所附的权利要求指出。

权利要求书

1、一种多人即时通讯方法，其特征在于，所述方法由服务器执行，所述方法包括：

获取第一数目个的第一多媒体数据流；

将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目；

将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

2、根据权利要求1所述的方法，其特征在于，将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

基于属性配置信息，将所述第一数目个的第一多媒体数据流转换为匹配所述属性配置信息的所述第一数目个的第一多媒体数据流；所述属性配置信息是用于指示对应的多媒体数据流的呈现属性的信息；

将匹配所述属性配置信息的所述第一数目个的第一多媒体数据流，融合为匹配所述属性配置信息的第二数目个的第二多媒体数据流。

3、根据权利要求2所述的方法，其特征在于，所述在将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流之前，还包括：

获取所述接收方针对多媒体播放预设的所述属性配置信息。

4、根据权利要求1所述的方法，其特征在于，将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

从所述第一多媒体数据流中分离出对应的第一音频数据流；

对分离出的各所述第一音频数据流进行混音处理，获取所述第二数目个的所述第二多媒体数据流。

5、根据权利要求4所述的方法，其特征在于，在对所述分离出的各第一音频数据进行混音处理之前，所述方法还包括：

若目标多媒体数据流来源于所述接收方，将从目标多媒体数据流中分离出的第一音频数据流进行剔除；所述目标多媒体数据流是所述第一数目个的所述第一多媒体数据流中的任意一个。

6、根据权利要求1所述的方法，其特征在于，将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

从所述第一多媒体数据流中分离出对应的第一视频数据流；

对分离出的各所述第一视频数据流进行混屏处理，获取所述第二数目个的所述第二多媒体数据流。

7、根据权利要求1所述的方法，其特征在于，将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，包括：

从所述第一多媒体数据流中分离出对应的第一音频数据流；

从所述第一多媒体数据流中分离出对应的第一视频数据流；

对分离出的各第一音频数据流进行混音处理；

对分离出的各第一视频数据流进行混屏处理；

基于混音处理的各第一音频数据流与混屏处理的各第一视频数据流，获取所述第二数目个的所述第二多媒体数据流。

8、一种多人即时通讯系统，其特征在于，所述多人即时通讯系统包括调度节点、至少两个候选数据适配节点，其中：

所述调度节点配置为从所述至少两个候选数据适配节点中选取一候选数据适配节点，作为目标数据适配节点；

所述目标数据适配节点配置为获取第一数目个的第一多媒体数据流，将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目，将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

9、根据权利要求8所述的系统，其特征在于，所述调度节点配置为：

定期获取所述各候选数据适配节点分别对应的工作状态信息；

基于所述工作状态信息，分别获取所述各候选数据适配节点对应的状态分数；

基于所述状态分数，从所述各候选数据适配节点中选取目标数据适配节点。

10、根据权利要求8所述的系统，其特征在于，所述系统还包括至少两个候选音频处理节点，所述调度节点配置为：从所述至少两个候选音频处理节点中选取一候选音频处理节点，作为目标音频处理节点。

11、根据权利要求10所述的系统，其特征在于，所述目标数据适配节点配置为从所述第一多媒体数据流中分离出对应的第一音频数据流；

所述目标音频处理节点配置为对分离出的各所述第一音频数据流进行混音处理，将混音处理的各所述第一音频数据流发送给所述目标数据适配节点，以使得所述目标数据适配节点获取所述第二数目个的所述第二多媒体数据流。

12、根据权利要求11所述的系统，其特征在于，所述目标音频处理节点配置为在对分离出的各所述第一音频数据流进行混音处理之前，若目标多媒体数据流来源于所述接收方，将从所述目标多媒体数据流中分离出的第一音频数据流进行剔除；所述目标多媒体数据流是所述第一数目个的所述第一多媒体数据流中的任意一个。

13、根据权利要求8所述的系统，其特征在于，所述系统还包括至少两个候选视频处理节点，所述调度节点配置为：从所述至少两个候选视频处理节点中选取一候选视频处理节点，作为目标视频处理节点。

14、根据权利要求13所述的系统，其特征在于，所述目标数据适配节点配置为从所述第一多媒体数据流中分离出对应的第一视频数据流；

所述目标视频处理节点配置为对分离出的各所述第一视频数据流进行混屏处理，将混屏处理的各所述第一视频数据流发送给所述目标数据适配节点，以使得所述目标数据适配节点获取所述第二数目个的所述第二多媒体数据流。

15、一种多人即时通讯装置，其特征在于，所述装置包括：

获取模块，配置为获取第一数目个的第一多媒体数据流；

融合模块，配置为将所述第一数目个的所述第一多媒体数据流融合为第二数目个的第二多媒体数据流，所述第二数目小于所述第一数目；

发送模块，配置为将所述第二数目个的所述第二多媒体数据流发送给接收方，以使得所述接收方呈现对应的多媒体。

16、一种多人即时通讯电子设备，其特征在于，包括：

存储器，存储有计算机可读指令；

处理器，读取存储器存储的计算机可读指令，以执行权利要求 1-6 中的任一个所述的方法。

17、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行，以实现如权利要求 1-6 中的任一个所述的方法。