

音视频会议实现方法、音视频会议系统及相关装置

本申请要求于 2021 年 07 月 30 日提交中国专利局、申请号 202110875339.6、申请名称为“音视频会议实现方法、音视频会议系统及相关装置”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**

本申请涉及计算机领域，特别是涉及音视频会议实现方法、音视频会议系统及相关装置。

背景技术

10 随着网络技术、通信技术和流媒体技术的迅速发展，以及人们工作、学习的流动性的增加，企业及个人对视频通讯的需求也越来越多，音视频会议系统应运而生。

目前的音视频会议多采用基于选择性转发单元（Selective Forwarding Unit, SFU）的会议架构，由一个服务器和多个终端组成，服务器收到会议房间中某个终端共享的音视频流（即媒体数据流）后，就直接将该音视频流转发给该会议房间内的其他终端。

15 然而，这种会议架构一旦进入一个会议房间的人数过多，将会大大增加服务器的数据分发压力，因此在这种会议架构下一个会议的参会人数有限，并且能够同时开麦或开视频的人数有限，普通观众要举手发言需要重定向，体验上不够顺滑。

发明内容

为了解决上述技术问题，本申请提供了一种音视频会议实现方法、音视频会议系统及相关装置，大大降低了媒体源设备例如第二接口机数据分发压力。因此，可以支持更多参
20 会人数，甚至支持百万人参会，同时由于媒体数据流传输时的数据分发压力大大降低，所有参会人员能够同时开麦或开视频，无需重定向，体验顺滑。

本申请实施例公开了如下技术方案：

第一方面，本申请实施例提供一种音视频会议实现方法，所述方法包括：

25 第一设备组中的回源设备接收所述第一设备组中第一接口机的拉流申请，所述拉流申请中包括媒体数据流的第一标识信息，所述第一设备组为多个设备组中的任一设备组，所述多个设备组中每个设备组中存在回源设备，每个设备组中包括多个接口机，所述多个接口机中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入会议房间；所述第一接口机是所述第一设备组中的任一接口机；所述第一设备组中的回源设备将所述
30 第一标识信息对应的所述媒体数据流转发至所述第一接口机上，以使所述第一接口机将所述媒体数据流转发至对应的会议接入终端；所述媒体数据流是所述第一设备组中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第二设备组为所述多个设备组中的任一设备组。

第二方面，本申请实施例提供一种音视频会议实现装置，所述装置包括接收单元和转发单元：

35 所述接收单元，用于接收第一设备组中第一接口机的拉流申请，所述拉流申请中包括媒体数据流的第一标识信息，所述第一设备组为多个设备组中的任一设备组，所述多个设

-2-

备组中每个设备组中存在回源设备，每个设备组中包括多个接口机，所述多个接口机中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入会议房间；所述第一接口机是所述第一设备组中的任一接口机；

5 所述转发单元，用于将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，以使所述第一接口机将所述媒体数据流转发至对应的会议接入终端；所述媒体数据流是所述第一设备组中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第二设备组为所述多个设备组中的任一设备组。

10 第三方面，本申请实施例提供一种音视频会议系统，所述系统包括数据传输网络和房间管理子系统：

所述数据传输网络用于传输会议房间中产生的媒体数据流，所述数据传输网络包括多个设备组，每个设备组中包括多个接口机；

所述多个接口机中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入所述会议房间；

15 所述多个设备组中每个设备组中存在回源设备，第一设备组中的回源设备用于接收所述第一设备组中第一接口机的拉流申请，所述拉流申请中包括媒体数据流的第一标识信息，将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，所述媒体数据流是所述第一设备组中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第一设备组和所述第二设备组
20 分别为所述多个设备组中的任一设备组；

所述第一接口机是所述第一设备组中的任一接口机，所述第一接口机用于将所述媒体数据流转发至对应的会议接入终端；

所述房间管理子系统包括所述多个设备组和中心服务器，每个设备组中包括组内服务器，所述组内服务器用于辅助所述中心服务器管理所述会议房间。

25 第四方面，本申请实施例提供一种用于音视频会议实现的电子设备，所述电子设备包括处理器以及存储器：

所述存储器用于存储程序代码，并将所述程序代码传输给所述处理器；

所述处理器用于根据所述程序代码中的指令执行第一方面所述的方法。

30 第五方面，本申请实施例提供一种计算机可读存储介质，所述计算机可读存储介质用于存储程序代码，当程序代码在电子设备上运行时，使得电子设备执行第一方面所述的方法。

第六方面，本申请实施例提供一种计算机程序产品，当所述计算机程序产品被执行时，使得电子设备执行第一方面所述的方法。

35 由上述技术方案可以看出，本申请涉及多个设备组，每个设备组中包括多个接口机，多个接口机中每个接口机用于连接会议接入终端，以便会议接入终端通过接口机接入会议房间。与相关技术中多个会议接入终端通过一个服务器接入会议房间相比，多个接口机与相关技术中服务器作用类似，而将会议接入终端分散到不同的接口机，通过多个接口机分

担负载，便于接入更多的会议接入终端，支持大规模会议。多个设备组中每个设备组中存
在回源设备，通过第二接口机接入会议房间产生媒体数据流，每个设备组例如第一设备组
中的回源设备在接收到第一设备组中第一接口机的拉流申请后，由于拉流申请中包括媒体
数据流的第一标识信息，故可以将第一标识信息对应的媒体数据流转发至第一接口机上，
5 以便第一接口机用于将媒体数据流转发至对应的会议接入终端。这样，在每个接口机都需
要拉取第二接口机的媒体数据流时，仅需要第二接口机与各个设备组中的回源设备进行交
互，进而由回源设备为媒体源设备例如第二接口机分担数据分发压力，与其他接口机进行
交互，无需第二接口机与每个接口机进行交互，大大降低了设备例如第二接口机的数据分
发压力。因此，该音视频会议实现方法可以支持更多参会人数，甚至支持百万人参会，同
10 时由于媒体数据流传输时的数据分发压力大大降低，所有参会人员（即用户）能够同时开
麦或开视频，无需重定向，体验顺滑。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术
描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请
15 的一些实施例，对于本领域普通技术成员来讲，在不付出创造性劳动性的前提下，还可以
根据这些附图获得其他的附图。

- 图 1 为相关技术提供了一种基于 SFU 的会议架构图；
- 图 2 为相关技术提供的另一种基于 SFU 的会议架构图；
- 图 3 为相关技术提供了一种音视频会议系统的结构图；
- 20 图 4a 为本申请实施例提供了一种音视频会议实现方法的流程图；
- 图 4b 为本申请实施例提供了一种数据传输网络的结构图；
- 图 5 为本申请实施例提供了一种基于本申请实施例提供的音视频会议系统进行视频会
议的界面图；
- 图 6 为本申请实施例提供了一种房间管理子系统的结构图；
- 25 图 7 为本申请实施例提供了一种基于混音引擎的音频处理架构图；
- 图 8 为本申请实施例提供了一种与第三方会议设备对接的架构图；
- 图 9 为本申请实施例提供的又一种音频处理架构图；
- 图 10 为本申请实施例提供了一种音视频会议实现装置的结构图；
- 图 11 为本申请实施例提供了一种终端的结构图；
- 30 图 12 为本申请实施例提供了一种服务器的结构图。

具体实施方式

下面结合附图，对本申请的实施例进行描述。
目前的音视频会议多采用基于 SFU 的会议架构，参见图 1 所示，图中黑色节点为服务
器，灰色节点为终端，多个终端与服务器连接，用户 A、用户 B、用户 C 分别通过对应的
35 终端接入会议房间参与会议。

然而，这种会议架构一旦进入一个会议房间的人数过多，将会大大增加服务器的数据
分发压力，因此在这种会议架构下一个会议的参会人数有限，并且能够同时开麦或开视频
的人数有限，普通观众要举手发言需要重定向，体验上不够顺滑。

另外，这种方式参会人员不能同时开视频或开麦，为此，在图 1 所示的 SFU 架构基础上增加扩散代理（如图 2 中虚线框中包括的节点所示），把不发言的用户放到扩散代理，举手发言时再重定向到通话节点（如图 2 中实线框中包括的节点所示），利用扩散代理的分发优势扩大会议参与人数。这种方式却需要重定向，体验不够顺滑。

- 5 为了解决上述技术问题，本申请实施例提供一种音视频会议实现方法，该方法在每个接口机都需要拉取第二接口机的媒体数据流时，仅需要第二接口机与各个设备组中的回源设备进行交互，进而由回源设备为第二接口机分担数据分发压力，与其他接口机进行交互，无需第二接口机与每个接口机进行交互，大大降低了媒体源设备例如第二接口机的数据分发压力。因此，该音视频会议实现方法可以支持更多参会人员，甚至支持百万人参会，同时由于媒体数据流传输时的数据分发压力大大降低，所有参会人员能够同时开麦或开视频，
- 10 无需重定向，体验顺滑。

需要说明的是，本申请实施例提供的音视频会议实现方法可以应用于各种音视频会议场景，尤其是用于单会议百万人参会的超大规模视音频会议。

- 15 接下来将对本申请实施例提供的音视频会议系统架构进行介绍。参会人员基于音视频会议系统进入会议房间参与会议时，参会人员主要包括需要开启音视频的嘉宾，以及普通观众，嘉宾开启音视频时会产生媒体数据流，并将媒体数据流传输至普通观众对应的会议接入终端，以及其他嘉宾对应的会议接入终端。其中，会议接入终端为接入会议房间所使用的终端，会议接入终端可以是用户终端。因此，音视频会议主要在于参会人员开启音视频例如开启视频或开麦时媒体数据流的传输，当媒体数据流的传输不受会议房间中参会人员人数限制时，音视频会议系统将可以支持大规模会议，因此本申请实施例重点在于数据传输网络架构的改进。
- 20 另外，因为每个会议都要对应一个会议房间用来保存房间信息、参会列表和维护用户状态（开启或关闭音视频即上下麦、上下视频、静音、进退房等）。在超大规模会议的情况下，所有人都不断开启或关闭视频或进退房，会产生大量体现状态变更的通知消息；超

- 25 大音视频会议的参会列表的拉取和同步也会对房间管理子系统产生很大压力。因此本申请实施例对房间管理子系统的改进也会使得音视频会议系统可以支持超大规模音视频会议。

因此，在本实施例中音视频会议系统架构可以参见图 3 所示，主要包括数据传输网络 301 和房间管理子系统 302，数据传输网络 301 用于传输会议房间中产生的媒体数据流，房间管理子系统 302 用于对会议房间进行管理。

- 30 本申请实施例首先从数据传输网络 301 传输媒体数据流的角度对音视频会议实现方法进行介绍。参见图 4a，所述方法包括：

S401、第一设备组中的回源设备接收所述第一设备组中第一接口机的拉流申请。

- 数据传输网络 301 可以参见图 4b 所示，包括多个设备组 401，例如图 4b 中每个虚线框所示。每个设备组 401 中包括多个接口机，每个接口机例如图 4b 中设备组 401 中的圆圈所表示的节点。多个接口机中每个接口机用于连接会议接入终端，以便会议接入终端通过
- 35 接口机接入会议房间，使得会议接入终端对应的参会人员可以进入会议房间参与会议。其

—5—

中，设备组可以用 SET 表示，即多个接口机按照 SET 划分，接口机可以是服务器。多个设备组 401 中每个设备组 401 中存在回源设备。

其中，第一设备组可以为多个设备组中的任一设备组，第一接口机可以是第一设备组中的任一接口机。

- 5 当某个设备组（例如第一设备组）中的某个接口机（例如第一接口机）需要拉取媒体数据流时，第一接口机可以向第一设备组中的回源设备发送拉流申请。拉流申请中可以包括媒体数据流的第一标识信息，第一标识信息标识了其订阅了的媒体数据的媒体数据流，从而便于回源设备知晓需要拉取哪个媒体数据流。其中，第一接口机可以是订阅了媒体数据的用户对应的会议接入终端所连接的接口机。也就是说，第一接口机上的用户订阅了某个接口机的媒体数据，则触发第一接口机向第一设备组中的回源设备发送拉流申请。
- 10

S402、第一设备组中的回源设备将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，以使所述第一接口机将所述媒体数据流转发至对应的会议接入终端。

- 15 第一设备组中的回源设备接收到拉流申请后，将第一标识信息对应的媒体数据流转发至第一接口机上，再由第一接口机将媒体数据流转发至对应的会议接入终端。其中，产生媒体数据流的会议接入终端通过第二接口机接入会议房间，第二设备组为多个设备组中的任一设备组。

- 20 第二接口机可以为开启了音视频的用户对应的会议接入终端所连接的接口机。例如，当某个用户例如第二用户在其会议接入终端上开启了视频，第二用户对应的会议接入终端通过第二接口机接入会议，第二接口机位于第二设备组中。若参与会议的任一用户例如第一用户希望观看第二用户的视频（例如订阅了第二用户的视频）时，第一用户对应的会议接入终端所连接的第一接口机可以找到其所在设备组例如第一设备组中的回源设备，并向第一设备组中的回源设备发送拉流申请，以便回源设备根据拉流申请将从第二接口机获取的媒体数据流转发至第一接口机上，第一接口机将媒体数据流转发至对应的会议接入终端。其中，本申请实施例涉及的用户可以是参会人员。

- 25 需要说明的是，回源设备上的媒体数据流是从第二接口机上拉取的，拉取的媒体数据流可以保存在回源设备本地。也就是说，一般情况下，当回源设备第一次接收到拉取媒体数据流的拉流申请时，由于本地尚未保存该媒体数据流，故需要从第二接口机上拉取；若之前已经有其他接口机申请过，本地便保存了该媒体数据流，则无需再从第二接口机上拉取。

- 30 因此，在本实施例中，可以根据第一设备组中的回源设备是否在本本地查找到媒体数据流而采用不同的媒体数据流拉取方式。

- 35 在一种可能的实现方式中，第一设备组中的回源设备将第一标识信息对应的媒体数据流转发至第一接口机上的具体方式可以是第一设备组中的回源设备接收到第一接口机发送的拉流申请，根据第一标识信息查找对应的第二接口机的媒体数据流，若未查找到第二接口机的媒体数据流，向第二接口机转发该拉流申请，接收第二接口机根据拉流申请中的第一标识信息转发的媒体数据流，并将媒体数据流转发至第一接口机。

可以理解的是，若拉流申请是在第一用户订阅了第二用户的视频时发送的，则第二接口机接收到拉流申请后，可以在第二接口机的本地记录该回源设备订阅了第二接口机，然后在由第二接口机将媒体数据流发送给回源设备。回源设备也会将接收到的媒体数据流保存至本地，以便后续其他接口机请求该媒体数据流时，回源设备无需再与第二接口机进行交互，可以直接将媒体数据流返回给第一接口机。

例如如图 4b 所示，其中节点 A 可以作为第二接口机，第二接口机为嘉宾所在接口机，即第二接口机上的会议接入终端所对应的用户可以开启视频，上行视频数据流，此时视频数据流作为媒体数据流。节点 B 可以作为第一接口机，节点 D 可以作为其所在设备组例如第一设备组的回源设备，节点 B 上的用户订阅了节点 A 上行的视频数据流，节点 B 会先在第一设备组内找到节点 D，向节点 D 发出拉流申请。节点 D 收到节点 B 的拉流申请后，先在本地查找有没有节点 A 的媒体数据流，如果没有，则向节点 A 发起拉流申请。节点 A 收到节点 D 的拉流申请后，在本地记录节点 D 订阅节点 A，然后将视频数据流转发给节点 D，节点 D 收到视频数据流转发给节点 B，节点 B 最终将视频数据流发送至对应的会议接入终端。

第一设备组中的回源设备根据第一标识信息查找第二接口机的媒体数据流时，若查找到第二接口机的媒体数据流，直接将查找到的媒体数据流转发至第一接口机。通过这种方式可以提高媒体数据流的拉取效率，进一步降低媒体源设备例如第二接口机的数据分发压力。

例如如图 4b 所示，其中节点 A 可以作为第二接口机，第二接口机为嘉宾所在接口机，即第二接口机上的会议接入终端所对应的用户可以开启视频，上行视频数据流，此时视频数据流作为媒体数据流。节点 C 可以作为第一接口机，节点 D 可以作为其所在设备组例如第一设备组的回源设备，节点 C 上的用户订阅了节点 A 上行的视频数据流，节点 C 会先在第一设备组内找到节点 D，向节点 D 发出拉流申请。注意同一设备组中的节点 B 和节点 C 通过一致性哈希或其他方法都找到同一个节点 D，向节点 D 发出拉流申请。节点 D 查找到本地已经有这条视频数据流了（之前节点 B 申请过），节点 D 直接将查找到的视频数据流向节点 C 转发，节点 C 收到来自节点 D 的转发的视频数据流，最终将视频数据流转发给对应的会议接入终端。

可以理解的是，第一设备组中的回源设备是需要申请媒体数据流的接口机例如第一接口机按照预设规则确定得到的，该预设规则使得同一媒体数据流通过同一回源设备路由到其所在设备组中的接口机上。例如上述同一设备组中的节点 B 和节点 C 通过一致性哈希或其他方法都找到同一个节点 D。从而便于在同一设备组中对同一媒体数据流进行统一分发管理。

需要说明的是，第一设备组与第二设备组可以位于同一局域网（即在内网中），也可以位于不同局域网（即在外网中，包括国内网络和国外网络的外网）。若第一设备组与第二设备组位于同一局域网，则第一设备组中的回源设备直接向第二接口机转发拉流申请，第二接口机直接向回源设备发送媒体数据流。若第一设备组与第二设备组位于不同局域网，数据传输网络 301 中还包括中转代理设备 402，例如如图 4b 中 402 所示的设备。则第一设备

组中的回源设备需要通过中转代理设备 402 向第二接口机转发拉流申请，相应的，第二接口机通过中转代理设备 402 将媒体数据流转发至第一设备组中的回源设备。

例如图 4b 所示，其中节点 A 可以作为第二接口机，第二接口机为嘉宾所在接口机，即第二接口机上的会议接入终端所对应的用户可以开启视频，上行视频数据流，此时视频数据流作为媒体数据流。节点 G 可以作为第一接口机，节点 F 可以作为其所在设备组例如第一设备组的回源设备，此时节点 G 所在第一设备组与节点 A 所在第二设备组不在同一局域网，即节点 G 为边缘节点。节点 G 上的用户订阅了节点 A 上行的视频数据流（该视频数据流作为媒体数据流），获取视频数据流的路径会相对比较长，节点 G 会先在第一设备组内找到节点 F，向节点 F 发出拉流申请。节点 F 收到节点 G 的拉流申请后，先在本地查找有没有节点 A 的视频数据流，如果没有，则通过中转代理设备 402（例如图 4b 中节点 E）向节点 A 发起拉流申请。节点 A 收到节点 F 的拉流申请，在本地记录节点 F 订阅节点 A，然后通过节点 E 将视频数据流转发给节点 F，节点 F 再将视频数据流转发给节点 G，节点 G 最终将视频数据流转发给对应的会议接入终端。

图 4b 中节点 H 拉取节点 A 的视频数据流的过程与节点 C 类似，只是中间也要通过节点 E，此处不再赘述。在这一过程中，中转代理设备 402 例如节点 E 的查找，既可以是预先配置的，也可以做动态分配，这里不做限制。在结构上，本申请实施例将用于接入音视频会议的接口机按设备组（SET）进行划分，在某个设备组中的接口机需要拉取媒体数据流时，可以通过在所在设备组中确定回源设备，进而通过回源设备向产生媒体数据流的接口机（即媒体源设备）发起拉流申请。这种设计大大减少了媒体源设备的数据分发压力，极大提升会议房间中参会人员的人数。举个例子，假设每个 SET 有 100 台接口机，整个音视频会议系统有 100 个 SET，每台 8 核 16G 的接口机能够支持 100 路媒体数据流转发，如果一个会议只有一个嘉宾其余都是观众，那么这个会议的最大规模就是 $100(\text{台}) \times 100(\text{SET}) \times 100(\text{分发}) = 100 \text{ 万人}$ 。如果嘉宾不止一人，基于就近接入原则，会分配不同 SET 的不同接口机，上行转发的负载也会分散开，支持百万人的大会很容易。

由上述技术方案可以看出，本申请涉及多个设备组，每个设备组中包括多个接口机，多个接口机中每个接口机用于连接会议接入终端，以便会议接入终端通过接口机接入会议房间。与相关技术中多个会议接入终端通过一个服务器接入会议房间相比，多个接口机与相关技术中服务器作用类似，而将会议接入终端分散到不同的接口机，通过多个接口机分担负载，便于接入更多的会议接入终端，支持大规模会议。多个设备组中每个设备组中存在回源设备，通过第二接口机接入会议房间产生媒体数据流，每个设备组例如第一设备组中的回源设备在接收到第一设备组中第一接口机的拉流申请后，由于拉流申请中包括媒体数据流的第一标识信息，故可以将第一标识信息对应的媒体数据流转发至第一接口机上，以便第一接口机用于将媒体数据流转发至对应的会议接入终端。这样，在每个接口机都需要拉取第二接口机的媒体数据流时，仅需要第二接口机与各个设备组中的回源设备进行交互，进而由回源设备为媒体源设备例如第二接口机分担数据分发压力，与其他接口机进行交互，无需第二接口机与每个接口机进行交互，大大降低了设备例如第二接口机的数据分发压力。因此，该音视频会议实现方法可以支持更多参会人数，甚至支持百万人参会，同

时由于媒体数据流传输时的数据分发压力大大降低，所有参会人员（即用户）能够同时开麦或开视频，无需重定向，体验顺滑。

与基于多点控制单元（Multipoint Control Unit, MCU）的会议架构相比，运营维护简单，不会受性能和硬件价格限制，大大提升实际承载人数，便于实现超大规模会议。

- 5 基于本申请实施例提供的音视频会议实现方法及音视频会议系统进行视频会议的界面图可以参见图 5 所示，在该界面中可以显示开启视频的参会人员的图像、参会列表、相关功能键的图标（例如麦克风图标、摄像头图标、管理成员图标、离开会议图标等等），开麦的参会人员可以显示麦克风的图标，未开麦的参会人员可以在麦克风的图标上划一条斜线。在本申请实施例提供的音视频会议系统下，实现了超大规模会议，例如进入会议的人数上限可以是 50000，即管理成员 50000。

10 接着，本申请实施例从房间管理子系统 302 管理会议房间的角度对音视频会议实现方法进行介绍。

- 房间管理子系统 302 可以参见图 6 所示，包括上述多个设备组 601 和中心服务器（RoomSvc）602，这里的设备组 601 同样是包括前述多个接口机，与图 4b 所示的设备组类似，与此同时，每个设备组 601 中还包括组内服务器（RoomSvcInSet），组内服务器用于辅助中心服务器 602 管理会议房间。其中，组内服务器如图 6 中每个设备组 601 中的圆环所示，例如图 6 中节点 B、节点 C、节点 D、节点 E、节点 F、节点 G；中心服务器 602 例如图 6 中节点 A。

- 20 需要说明的是，在本实施例中同一个设备组中的组内服务器与回源设备可以是同一个设备，也可以是不同设备，本实施例对此不做限定。

管理会议房间主要包括管理会议房间的房间信息、状态变更的通知消息、参会列表的拉取等等。其中，房间信息例如包括会议房间的名称、创建时间、参会人员、会议时间等等。

- 25 本申请实施例通过将房间管理系统分为中心服务器和组内服务器，中心服务器只有一个，组内服务器和接口机一样按 SET 部署，从而使得通知消息的同步和参会列表的拉取都可以通过组内服务器来缓解中心服务器的管理压力，便于实现超大规模会议。

- 30 在设备组内部，对于同一个会议房间，所有接口机都会通过一定的算法注册相同的 RoomSvcInSet，并通过心跳维持一个长连接，然后才能在会议房间中进行媒体数据流的拉取。当会议房间内有通知消息需要进行同步（如用户的状态变更：进入会议房间参加会议、退出会议房间、开关麦、开关视频等）、或者想要拉取参会列表时，都会通过这个 RoomSvcInSet 来完成。

- 35 首先介绍进入会议房间参见会议的注册流程。当任一设备组例如第三设备组中的任一接口机例如第三接口机请求进入会议房间时，第三设备组中的组内服务器接收第三接口机发送的注册请求，注册请求中可以包括房间标识，用于指示需要进入哪个会议房间。若第三设备组中的组内服务器在本地未查找到该会议房间的信息，向中心服务器发送该注册请求。中心服务器将第三设备组中组内服务器的第二标识信息记录到第一会议注册列表，并向第三设备组中的组内服务器返回注册成功信息。第三设备组中的组内服务器将第三接口

机的第三标识信息记录到第二会议注册列表，并向第三接口机返回注册成功信息，从而完成第三接口机的注册。其中，第二标识信息和第三标识信息都是一种标识信息，是对应设备的唯一标识，标识信息例如可以是地址，也就是说，第二标识信息可以是组内服务器的地址，第三标识信息可以是第三接口机的地址。其中，第三设备组为多个设备组中任一设备组，第三接口机为第三设备组中的任一接口机，第三设备组可以是前述实施例中的第一设备组或第二设备组，第三接口机可以是前述实施例中的第一接口机或第二接口机。

例如如图6所示，其中节点a可以作为第三接口机，第三接口机为希望注册会议的接口机，即第三接口机上的会议接入终端所对应的用户希望进入会议房间参加会议。节点A可以作为中心服务器，节点B可以作为第三接口机所在设备组例如第三设备组的组内服务器。其中，节点B可以是接点a根据一定规则（一致性哈希或者分配调度）确定的。节点a向节点B发送注册请求，节点B收到节点a的注册请求后，向调度系统（Scheduler）查找该会议对应的RoomSvc，即节点A。节点B向节点A发送注册请求，节点A收到节点B发出的注册请求，将节点B的地址记录到该会议的注册列表（节点A上保存的第一会议注册列表）中，给B返回注册成功信息。节点B收到节点A返回注册成功信息后，将节点a的地址记录到该会议的注册列表（节点B上保存的第二会议注册列表），向节点a返回注册成功信息。

可以理解的是，在本实施例中，在注册成功后，组内服务器与中心服务器、组内服务器与注册成功的接口机例如第三接口机可以保持长连接，方便后续二者之间的交互。例如节点B还可以向节点A定时发送心跳，与节点A保持一个长连接通道，节点a收到节点B返回注册成功信息后，定时发送心跳，与节点B保持长连接通道。

若第三设备组中的组内服务器在根据注册请求在本地查找会议房间的信息时，若在本地查找到会议房间的信息，可以将第三接口机的第三标识信息记录到第二会议注册列表，并向第三接口机返回注册成功信息。

例如如图6所示，其中节点b可以作为第三接口机，第三接口机为希望注册会议的接口机，即第三接口机上的会议接入终端所对应的用户希望进入会议房间参加会议。节点A可以作为中心服务器，节点B可以作为第三接口机所在设备组例如第三设备组的组内服务器。其中，节点B可以是接点b根据一定规则（一致性哈希或者分配调度）确定的。节点b向节点B发送注册请求，节点B收到节点b的注册请求，发现本地已经有该会议房间的信息，将节点b的地址记录到该会议的注册列表（节点B上保存的第二会议注册列表）中，并向节点b返回注册成功信息。节点b收到节点B返回注册成功信息后，定时发送心跳，与节点B保持长连接通道。

同理，位于边缘节点（即设备组与中心服务器不在同一局域网，设备组中的接口机与中心服务器不在同一局域网）上的用户注册相同会议房间时，注册流程上只需要经过中转代理设备例如如图6中节点O、节点P、节点Q进行中转，其他流程类似。

当接入会议房间的会议接入终端发生状态变更时，例如用户开启或关闭音视频时，需要将开启或关闭音视频的这种状态变更通知给会议房间中的其他用户。以发生状态变更的会议接入终端位于第三设备组中的第三接口机上为例，此时，第三设备组中的组内服务器

接收第三接口机发送的通知消息，并将该通知消息发送至中心服务器。中心服务器将该通知消息发送至每个设备组的组内服务器上，每个设备组的组内服务器将该通知消息发送至注册成功的接口机。

- 继续以上述图 6 为例，其中节点 a 可以作为第三接口机，节点 a 为嘉宾所在接口机，节点 A 可以作为中心服务器，节点 B 可以作为第三接口机所在设备组例如第三设备组的组内服务器。若节点 a 上的嘉宾开了视频，即发生状态变更，节点 a 将指示状态变更的通知消息 (msg-a-openvideo) 发送至节点 B，节点 B 将通知消息发送给节点 A。节点 A 遍历本地注册列表 (节点 A 上保存的第一会议注册列表)，将通知消息发送给所有组内服务器，例如图 6 中的节点 B、节点 C、节点 D、节点 E、节点 F、节点 G。节点 B、节点 C、节点 D、节点 E、节点 F、节点 G 收到通知消息，遍历本地注册列表 (各个作为组内服务器上保存的第二会议注册列表)，将通知消息发送给所有接口机，例如图 6 中的节点 a、节点 b、节点 c、节点 d、节点 e、节点 f、节点 g、节点 o、节点 p、节点 q、节点 r 和节点 s。

- 音视频会议中维护参会列表，参会列表中包括所有参会人员，本申请实施例提供的音视频会议系统实现的超大规模音视频会议中，参会人员数量巨大，故可以采用分页拉取的方式来更新本地参会列表。当某个设备组例如第三设备组中的某个接口机例如第三接口机请求拉取参会列表时，第三设备组中的组内服务器接收第三接口机的列表拉取请求，第三设备组中的组内服务器在本地查找参会列表，若第三设备组中的组内服务器在本地未查找该参会列表，向中心服务器发送该列表拉取请求。第三设备组中的组内服务器接收中心服务器发送的参会列表，并将该参会列表发送至第三接口机。

- 例如图 6 所示，若节点 a 作为第三接口机，节点 a 上的用户向节点 B 请求参会列表，节点 B 在本地缓存中查找，如果没有，就向节点 A 发起列表拉取请求。节点 A 收到列表拉取请求后，将对应的参会列表返回给节点 B，节点 B 将参会列表缓存在本地后，并将参会列表返回给节点 a。其中，参会列表可以是全部参会列表，也可以是分页拉取的某页参会列表例如 Page1。

- 第三设备组中的组内服务器在本地查找参会列表时，若第三设备组中的组内服务器在本地查找到该参会列表，且该参会列表未过期，将该参会列表返回给第三接口机。

- 继续上个例子，若图 6 所示的节点 b 作为第三接口机，如果在节点 b 上之前已经存在其他节点也请求上述参会列表 Page1，此时节点 B 上已经存在该参会列表，故无须再向节点 A 请求参会列表。故节点 a 向节点 B 发送列表拉取请求，节点 B 在本地缓存中查找，发现有 Page1，并且没过期，节点 B 将 Page1 的直接返回给节点 b。

- 接着来看一下容灾恢复的流程：由于 RoomSvc 的所有信息都来自 RoomSvcInSet 和接口机，所以 RoomSvc 当机后，可以很方便通过 RoomSvcInSet 自动重建。同理 RoomSvcInSet 的数据都来自接口机，当 RoomSvcInSet 当机后，可以通过本设备组 (SET) 内的接口机自动重建。

- 本申请实施例提供的音视频会议实现方法，通过组内服务器来缓解中心服务器的管理压力，便于实现超大规模会议。

需要说明的是，在相关技术提供的基于 SFU 的音视频会议，在传输会议房间中产生的媒体数据流例如音频数据流时，需要基于能量值进行音频选路。而这种方式在长距离通话中产生的延迟会非常影响体验，当参会人数比较多的时候，浪费了大量带宽并增大了会议接入终端网络性能损耗。并且对会议中播放动式语音应答(Interactive Voice Response, IVR)

5 和精确录制有要求的场景，在 SFU 的转构下不太容易做精确控制（比如，开始和结束的判断，靠信令还是靠媒体数据流等等）。此外，语音录制的情况下，一些要求严格的行业是不能接受在录制过程中有多录或是少录几个字的情况的，如果没有一个集中式的音频处理服务，就容易出现录了不该录的内容或者该录的没录下来，这些情况对于较为苛刻的用户是绝对不可以接受的。

10 为此，参见图 3 所示，本申请实施例提供的音视频会议系统中还包括混音引擎 303，混音引擎 303 位于媒体传输网络 301 和公共交换电话网络（Public Switched Telephone Network, PSTN）305 之间，作为二者沟通的桥梁，极大提升超大规模房间下的开会体验。其中，混音引擎 303 通过协议例如会话初始协议（Session Initiation Protocol, SIP）或实时传输协议（Real-time Transport Protocol, RTP）连接公共交换电话网络 305。

15 其中，混音引擎 303 包括混音器（Mixer）和选路器（Selector），每个设备组具有对应的混音引擎。当媒体数据流为音频数据流时，混音引擎中的选路器将获取的多路音频数据流向位于同一设备组的混音器发送，多路音频数据流分别来自不同设备组中的第四接口机。混音器从多路音频数据流中选择目标音频数据流，并向产生目标音频数据流的会议接入终端连接的目标接口机发送选中通知。目标接口机向混音器发送目标音频数据流，混音器对目标音频数据流进行混音，并将混音后的音频数据流转发至第四接口机对应的选路器，第四接口机包括目标接口机。第四接口机对应的选路器将混音后的音频数据流发送至第四接口机，第四接口机向其所在设备组中的其他接口机转发混音后的音频数据流。

20 Mixer 和 Selector 也是按 SET 部署的。为了尽可能降低延迟，一个 SET 的 Mixer 和 Selector 部署在同一个可用区里（可用区内部传输延时可以控制在 2ms）。与此同时，同一个会议的音频数据流，会集中在一个 SET 中处理，并且最终会集中到同一个 Mixer 执行混音。

25 假设图 7 代表了一个真实的会议，其中有 3 个参会人员开了麦，这三个参会人员所属接口机为节点 a、节点 f 和节点 g，其他接口机上都是观众。为了简化示例图，假定每个会议选 2 路音频数据流（通常真正的会议选 4-6 路音频数据流），此时音频数据流的处理流程如下：

30 节点 a、节点 f、节点 g 向图 6 中的调度系统请求选路器（为了尽可能减小延迟，调度系统会对同会议房间的音频数据流分配同一个可用区的选路器），分别得到图 7 中节点 B、节点 C、节点 D 所表示的选路器。节点 a、节点 f、节点 g 将音频数据流转发给分别对应的节点 B、节点 C、节点 D。节点 B、节点 C、节点 D 将缓存上行音频数据流，将能量值上报给混音器。混音器对能量值进行排序，并从中选出节点 f、节点 g 的两路音频数据流作为目标音频数据流，然后告知节点 C 和节点 D，节点 C 和节点 D 作为目标接口机。节点 C 和节点 D 收到选中通知后，把相应的目标音频数据流转发给混音器（这两步延迟在 5ms 以

内), 混音器对收到的目标音频数据流进行混音, 然后把混音后的音频数据流转发给节点 B、节点 C、节点 D, 节点 B、节点 C、节点 D 将混音后的音频数据流发送给节点 a、节点 f、节点 g。节点 a、节点 f、节点 g 再进行 SET 内部转发。

- 5 注意, 这里给发送给嘉宾的流和给普通观众的流是不同的, 给嘉宾的混音后的音频数据流要把自己说话的那一路排除掉。所以被选中的接口机会收到两路音频数据流, 一路转发给发言的那个嘉宾, 另一路转发给 SET 内的观众。

- 10 混音引擎 303 是实时的媒体处理系统, 因为经过混音引擎 303 处理过的音频数据流需要实时转发给会议房间内的所有参会人员。而很多情况下还有可能需要异步流处理系统, 即这个系统拿到音频数据流之后不需要再发回给会议房间里的参会人员, 而是另作他用(例如录制、鉴黄、直播推流等), 所以通常不需要那么实时。基于此, 本申请实施例提供的音视频会议系统参见图 3 所示, 还包括旁路媒体处理系统 304, 旁路媒体处理系统 304 用于通过机器人终端拉取会议房间中产生的媒体数据流, 并对媒体数据流进行处理。其中, 机器人终端参见图 3 所示, 包括直播流机器人(LiveStream Robot)、录制机器人(Record Robot)、会议连接机器人(MRA Robot)。

- 15 在本实施例中, 采用机器人(Robot)终端进入会议房间拉流的方式来实现旁路媒体处理功能。以录制为例, 用户业务后台通过应用程序接口(Application Programming Interface, API)方式启动一个录制任务, 录制任务系统实例化一个 Record Robot 模拟参会人员进入会议。Record Robot 将会议中的媒体数据流拉到本地, 混流转码后进行录制。会议结束后, Record Robot 将录制的文件上传到指定的存储(也可以定时上传部分分片)。

- 20 需要说明的是, 鉴黄、直播推流的方案与录制类似, 区别仅在于业务逻辑。与第三方会议设备对接时, 由于要进行双向通信, 所以与录制方案稍有不同:

- 在与第三方会议设备对接时, 参见图 8 所示, 用户业务后台启动一个 MRA(会议连接器) Robot, MRA Robot 进入**会议(此时其他参会人员可以看到这个 MRA Robot), MRA Robot 同时通过连接协议(例如 SIP 或 H323 等协议)与第三方会议设备连接。MRA Robot 25 把**会议中的所有媒体数据流拉到本地、混流之后, 通过连接协议转发给第三方会议设备。从第三方会议设备发来的媒体数据流, MRA Robot 将其转换成**会议的私有协议, 通过媒体传输网络转发给会议房间中其他参会人员。其中, 图 8 中 801 为移动会议终端, 802 为**会议房间智能会议室, 用于接入**会议。

- 需要说明的是, 如果音视频会议系统不引入混音引擎, 还可以使用图 9 所示的架构来实现数据传输网络 301 与公共交换电话网络 305 的对接。即通过选路器与旁路媒体处理系统中涉及到的机器人共同实现混音引擎的功能。其中, 选路器基于能量上报的能量及对应的选路决策进行选路。

- 30 基于前述提供的音视频会议实现方法, 本申请实施例还提供一种音视频会议实现装置, 参见图 10 所示, 所述装置 1000 包括接收单元 1001 和转发单元 1002:

- 35 所述接收单元 1001, 用于接收第一设备组中第一接口机的拉流申请, 所述拉流申请中包括媒体数据流的第一标识信息, 所述第一设备组为多个设备组中的任一设备组, 所述多个设备组中每个设备组中存在回源设备, 每个设备组中包括多个接口机, 所述多个接口机

中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入会议房间；
所述第一接口机是所述第一设备组中的任一接口机；

- 5 所述转发单元 1002，用于将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，以使所述第一接口机将所述媒体数据流转发至对应的会议接入终端；所述媒体数据流是所述第一设备组中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第二设备组为所述多个设备组中的任一设备组。

在一种可能的实现方式中，所述转发单元 1002，用于：

- 10 根据所述第一标识信息查找对应的所述第二接口机的媒体数据流；
若未查找到所述第二接口机的媒体数据流，向所述第二接口机转发所述拉流申请；
所述接收单元 1001，还用于接收所述第二接口机根据所述拉流申请中的第一标识信息转发的所述媒体数据流；

所述转发单元 1002，还用于将所述媒体数据流转发至所述第一接口机。

在一种可能的实现方式中，所述转发单元 1002，还用于：

- 15 若查找到对应的所述第二接口机的媒体数据流，将所述媒体数据流转发至所述第一接口机。

在一种可能的实现方式中，若所述第一设备组与所述第二设备组位于不同局域网，所述转发单元 1002，用于：

- 20 通过中转代理设备向所述第二接口机转发所述拉流申请；
通过所述中转代理设备接收所述第二接口机转发的所述媒体数据流。

在一种可能的实现方式中，所述第一设备组中的回源设备是所述第一接口机按照预设规则确定得到的，所述预设规则使得同一媒体数据流通过同一回源设备路由到其所在设备组中的接口机上。

在一种可能的实现方式中，所述装置还包括发送单元、记录单元和返回单元：

- 25 当第三设备组中的第三接口机请求进入所述会议房间时，所述接收单元 1001，还用于接收所述第三接口机发送的注册请求，所述第三设备组为所述多个设备组中的任一设备组，所述第三接口机为所述第三设备组中的任一接口机；

- 30 所述发送单元，用于若在本本地未查找到所述会议房间的信息，向中心服务器发送所述注册请求，以使所述中心服务器将所述第三设备组中组内服务器的第二标识信息记录到第一会议注册列表，并向所述第三设备组中的组内服务器返回注册成功信息；

所述记录单元，用于将所述第三接口机的第三标识信息记录到第二会议注册列表；

所述返回单元，用于向所述第三接口机返回所述注册成功信息。

在一种可能的实现方式中，若在本本地查找到所述会议房间的信息，所述记录单元，还用于将所述第三接口机的第三标识信息记录到所述第二会议注册列表；

- 35 所述返回单元，还用于向所述第三接口机返回所述注册成功信息。

在一种可能的实现方式中，当第三设备组中的第三接口机连接的会议接入终端发生状态变更时，所述接收单元 1001，还用于接收所述第三接口机发送的通知消息；

所述发送单元，用于将所述通知消息发送至所述中心服务器，以使所述中心服务器将所述通知消息发送至每个设备组的组内服务器上，每个设备组的组内服务器用于将所述通知消息发送至注册成功的接口机。

5 在一种可能的实现方式中，当第三设备组中的第三接口机请求拉取参会列表时，所述接收单元 1001，还用于接收所述第三接口机的列表拉取请求；

所述发送单元，用于若在本本地未查找到所述参会列表，向所述中心服务器发送所述列表拉取请求；

所述接收单元 1001，还用于接收所述中心服务器发送的所述参会列表；

所述发送单元，用于将所述参会列表发送至所述第三接口机。

10 在一种可能的实现方式中，所述发送单元，还用于若在本本地查找到所述参会列表，且所述参会列表未过期，将所述参会列表返回给所述第三接口机。

在一种可能的实现方式中，所述装置还包括选择单元和混音单元：

所述发送单元，还用于当媒体数据流包括音频数据流时，通过混音引擎中的选路器将获取的多路音频数据流向位于同一设备组的混音器发送，所述多路音频数据流分别来自不同设备组中的第四接口机；

所述选择单元，用于通过混音器从所述多路音频数据流中选择目标音频数据流；

所述发送单元，用于向目标接口机发送选中通知，所述目标接口机为产生所述目标音频数据流的会议接入终端连接的接口机；

20 所述接收单元 1001，还用于通过所述混音器接收所述目标接口机发送的目标音频数据流；

所述混音单元，用于对所述目标音频数据流进行混音；

所述转发单元 1002，还用于将混音后的音频数据流转发至所述第四接口机对应的选路器；所述第四接口机包括所述目标接口机；

25 所述发送单元，还用于通过所述第四接口机对应的选路器将混音后的音频数据流发送至所述第四接口机，以使所述第四接口机向其所在设备组中的其他接口机转发所述混音后的音频数据流。

在一种可能的实现方式中，所述装置还包括处理单元：

所述处理单元，用于通过机器人终端拉取所述会议房间中产生的媒体数据流，并对所述媒体数据流进行处理。

30 本申请实施例还提供了一种用于音视频会议实现的电子设备，该电子设备可以是终端，以终端为智能手机为例：

35 图 11 示出的是与本申请实施例提供的终端相关的智能手机的部分结构的框图。参考图 11，智能手机包括：射频（英文全称：Radio Frequency，英文缩写：RF）电路 1110、存储器 1120、输入单元 1130、显示单元 1140、传感器 1150、音频电路 1160、无线保真（英文全称：wireless fidelity，英文缩写：WiFi）模块 1170、处理器 1180、以及电源 1190 等部件。输入单元 1130 可包括触控面板 1131 以及其他输入设备 1132，显示单元 1140 可包括显示面板 1141，音频电路 1160 可以包括扬声器 1161 和传声器 1162。本领域技术人员可以理解，

图 11 中示出的智能手机结构并不构成对智能手机的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。

存储器 1120 可用于存储软件程序以及模块，处理器 1180 通过运行存储在存储器 1120 的软件程序以及模块，从而执行智能手机的各种功能应用以及数据处理。存储器 1120 可主要

5 5 主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）等；存储数据区可存储根据智能手机的使用所创建的数据（比如音频数据、电话本等）等。此外，存储器 1120 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

10 10 处理器 1180 是智能手机的控制中心，利用各种接口和线路连接整个智能手机的各个部分，通过运行或执行存储在存储器 1120 内的软件程序和/或模块，以及调用存储在存储器 1120 内的数据，执行智能手机的各种功能和处理数据，从而对智能手机进行整体监控。可选的，处理器 1180 可包括一个或多个处理单元；优选的，处理器 1180 可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制

15 15 解调处理器主要处理无线通信。可以理解的是，上述调制解调处理器也可以不集成到处理器 1180 中。

在本实施例中，前述实施例中由终端所执行的步骤可以基于图 2 所示的结构实现。

该电子设备还可以包括服务器，本申请实施例还提供服务器，请参见图 12 所示，图 12 为本申请实施例提供的服务器 1200 的结构图，服务器 1200 可因配置或性能不同而产生

20 20 比较大的差异，可以包括一个或一个以上中央处理器（Central Processing Units，简称 CPU）1222（例如，一个或一个以上处理器）和存储器 1232，一个或一个以上存储应用程序 1242 或数据 1244 的存储介质 1230（例如一个或一个以上海量存储设备）。其中，存储器 1232 和存储介质 1230 可以是短暂存储或持久存储。存储在存储介质 1230 的程序可以包括一个或一个以上模块（图示没标出），每个模块可以包括对服务器中的一系列指令操作。更进

25 25 一步地，中央处理器 1222 可以设置为与存储介质 1230 通信，在服务器 1200 上执行存储介质 1230 中的一系列指令操作。

服务器 1200 还可以包括一个或一个以上电源 1226，一个或一个以上有线或无线网络接口 1250，一个或一个以上输入输出接口 1258，和/或，一个或一个以上操作系统 1241，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 等等。

30 30 在本实施例中，所述服务器 1200 中的中央处理器 1222 可以执行以下步骤：

接收所述第一设备组中第一接口机的拉流申请，所述拉流申请中包括媒体数据流的第一标识信息，所述第一设备组为多个设备组中的任一设备组，所述多个设备组中每个设备组中存在回源设备，每个设备组中包括多个接口机，所述多个接口机中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入会议房间；所述第一接口机是所

35 35 述第一设备组中的任一接口机；

将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，以使所述第一接口机将所述媒体数据流转发至对应的会议接入终端；所述媒体数据流是所述第一设备组

中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第二设备组为所述多个设备组中的任一设备组。

根据本申请的一个方面，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用于存储程序代码，当程序代码在电子设备上运行时，使得电子设备执行前述各个实施例所述的音视频会议实现方法。

根据本申请的一个方面，提供了一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。电子设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，使得该电子设备执行上述实施例各种可选实现方式中提供的方法。

本申请的说明书及上述附图中的术语“第一”、“第二”、“第三”、“第四”等（如果存在）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数据在适当情况下可以互换，以便这里描述的本申请的实施例例如能够以除了在这里图示或描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排除他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，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步骤或单元。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组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。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元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存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或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）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory，简称 ROM）、随机存取存储器（Random Access Memory，简称 RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

— 17 —

以上所述，以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术成员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

权 利 要 求

1、一种音视频会议实现方法，所述方法包括：

5 第一设备组中的回源设备接收所述第一设备组中第一接口机的拉流申请，所述拉流申请中包括媒体数据流的第一标识信息，所述第一设备组为多个设备组中的任一设备组，所述多个设备组中每个设备组中存在回源设备，每个设备组中包括多个接口机，所述多个接口机中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入会议房间；所述第一接口机是所述第一设备组中的任一接口机；

10 所述第一设备组中的回源设备将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，以使所述第一接口机将所述媒体数据流转发至对应的会议接入终端；所述媒体数据流是所述第一设备组中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第二设备组为所述多个设备组中的任一设备组。

2、根据权利要求1所述的方法，所述第一设备组中的回源设备将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，包括：

15 根据所述第一标识信息查找对应的所述第二接口机的媒体数据流；

若未查找到所述第二接口机的媒体数据流，向所述第二接口机转发所述拉流申请；

接收所述第二接口机根据所述拉流申请中的第一标识信息转发的所述媒体数据流，并将所述媒体数据流转发至所述第一接口机。

3、根据权利要求2所述的方法，所述方法还包括：

20 若查找到对应的所述第二接口机的媒体数据流，将所述媒体数据流转发至所述第一接口机。

4、根据权利要求2所述的方法，若所述第一设备组与所述第二设备组位于不同局域网，所述向所述第二接口机转发所述拉流申请，包括：

通过中转代理设备向所述第二接口机转发所述拉流申请；

25 所述接收所述第二接口机根据所述拉流申请中的第一标识信息转发的所述媒体数据流，包括：

通过所述中转代理设备接收所述第二接口机转发的所述媒体数据流。

30 5、根据权利要求1-4任一项所述的方法，所述第一设备组中的回源设备是所述第一接口机按照预设规则确定得到的，所述预设规则使得同一媒体数据流通过同一回源设备路由到其所在设备组中的接口机上。

6、根据权利要求1-4任一项所述的方法，所述方法还包括：

当第三设备组中的第三接口机请求进入所述会议房间时，所述第三设备组中的组内服务器接收所述第三接口机发送的注册请求，所述第三设备组为所述多个设备组中的任一设备组，所述第三接口机为所述第三设备组中的任一接口机；

35 若所述第三设备组中的组内服务器在本地未查找到所述会议房间的信息，向中心服务器发送所述注册请求，以使所述中心服务器将所述第三设备组中组内服务器的第二标识信息记录到第一会议注册列表，并向所述第三设备组中的组内服务器返回注册成功信息；

所述第三设备组中的组内服务器将所述第三接口机的第三标识信息记录到第二会议注册列表，并向所述第三接口机返回所述注册成功信息。

7、根据权利要求6所述的方法，所述方法还包括：

5 若所述第三设备组中的组内服务器在本地查找到所述会议房间的信息，将所述第三接口机的第三标识信息记录到所述第二会议注册列表，并向所述第三接口机返回所述注册成功信息。

8、根据权利要求6所述的方法，所述方法还包括：

当第三设备组中的第三接口机连接的会议接入终端发生状态变更时，所述第三设备组中的组内服务器接收所述第三接口机发送的通知消息；

10 所述第三设备组中的组内服务器将所述通知消息发送至所述中心服务器，以使所述中心服务器将所述通知消息发送至每个设备组的组内服务器上，每个设备组的组内服务器用于将所述通知消息发送至注册成功的接口机。

9、根据权利要求6所述的方法，所述方法还包括：

15 当第三设备组中的第三接口机请求拉取参会列表时，所述第三设备组中的组内服务器接收所述第三接口机的列表拉取请求；

若所述第三设备组中的组内服务器在本地未查找到所述参会列表，向所述中心服务器发送所述列表拉取请求；

所述第三设备组中的组内服务器接收所述中心服务器发送的所述参会列表，并将所述参会列表发送至所述第三接口机。

20 10、根据权利要求6所述的方法，所述方法还包括：

若所述第三设备组中的组内服务器在本地查找到所述参会列表，且所述参会列表未过期，将所述参会列表返回给所述第三接口机。

11、根据权利要求1-4所述的方法，所述方法还包括：

25 当媒体数据流包括音频数据流时，通过混音引擎中的选路器将获取的多路音频数据流向位于同一设备组的混音器发送，所述多路音频数据流分别来自不同设备组中的第四接口机；

通过混音器从所述多路音频数据流中选择目标音频数据流，并向目标接口机发送选中通知，所述目标接口机为产生所述目标音频数据流的会议接入终端连接的接口机；

30 通过所述混音器接收所述目标接口机发送的目标音频数据流，并对所述目标音频数据流进行混音；

将混音后的音频数据流转发至所述第四接口机对应的选路器；所述第四接口机包括所述目标接口机；

通过所述第四接口机对应的选路器将混音后的音频数据流发送至所述第四接口机，以使所述第四接口机向其所在设备组中的其他接口机转发所述混音后的音频数据流。

35 12、根据权利要求1-4所述的方法，所述方法还包括：

通过机器人终端拉取所述会议房间中产生的媒体数据流，并对所述媒体数据流进行处理。

13、一种音视频会议实现装置，所述装置包括接收单元和转发单元：

所述接收单元，用于接收第一设备组中第一接口机的拉流申请，所述拉流申请中包括媒体数据流的第一标识信息，所述第一设备组为多个设备组中的任一设备组，所述多个设备组中每个设备组中存在回源设备，每个设备组中包括多个接口机，所述多个接口机中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入会议房间；所述第一接口机是所述第一设备组中的任一接口机；

所述转发单元，用于将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，以使所述第一接口机将所述媒体数据流转发至对应的会议接入终端；所述媒体数据流是所述第一设备组中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第二设备组为所述多个设备组中的任一设备组。

14、一种音视频会议系统，所述系统包括数据传输网络和房间管理子系统：

所述数据传输网络用于传输会议房间中产生的媒体数据流，所述数据传输网络包括多个设备组，每个设备组中包括多个接口机；

所述多个接口机中每个接口机用于连接会议接入终端，以使所述会议接入终端通过接口机接入所述会议房间；

所述多个设备组中每个设备组中存在回源设备，第一设备组中的回源设备用于接收所述第一设备组中第一接口机的拉流申请，所述拉流申请中包括媒体数据流的第一标识信息，将所述第一标识信息对应的所述媒体数据流转发至所述第一接口机上，所述媒体数据流是所述第一设备组中的回源设备从第二设备组中第二接口机获取的，产生所述媒体数据流的会议接入终端通过所述第二接口机接入所述会议房间，所述第一设备组和所述第二设备组分别为所述多个设备组中的任一设备组；

所述第一接口机是所述第一设备组中的任一接口机，所述第一接口机用于将所述媒体数据流转发至对应的会议接入终端；

所述房间管理子系统包括所述多个设备组和中心服务器，每个设备组中包括组内服务器，所述组内服务器用于辅助所述中心服务器管理所述会议房间。

15、一种用于音视频会议实现的电子设备，所述电子设备包括处理器以及存储器：

所述存储器用于存储程序代码，并将所述程序代码传输给所述处理器；

所述处理器用于根据所述程序代码中的指令执行权利要求 1-12 任一项所述的方法。

16、一种计算机可读存储介质，所述计算机可读存储介质用于存储程序代码，当程序代码在电子设备上运行时，使得电子设备执行权利要求 1 至 12 中任一项所述的方法。

17、一种计算机程序产品，当所述计算机程序产品被执行时，使得电子设备执行如权利要求 1-12 任一项所述的方法。