

基于 H.323 的 MCU 多点会议系统^①

许 胜 徐 捷 (国防科学技术大学 电子科学与工程技术学院 湖南 长沙 410073)

摘 要: 简单介绍了三种多点会议模式,并对多点控制器和多点处理器的设计方法进行了分析。对于混音过程中音频数据溢出的问题,简单介绍了几种常用的溢出算法以及它们的性能。最后对影响多点会议 Qos 的因素进行分析,提出了基于终端的 Qos 策略。

关键词: H.323; MCU; 多点处理器; 多点控制器; 溢出算法; Qos

MCU Multipoint Conference System Based on H.323

XU Sheng, XU Jie

(School of Electronic Science and Engineering National University of Defense Technology, Changsha 410073, China)

Abstract: This paper briefly describes three multipoint conferencing modes, and analyzes the multi-point controller and multi-point processor design method. To address the audio data overflow in the mixing process, it introduces the overflow of several commonly used algorithms and their performance. Finally, the factors that affect the MultiSite Qos analysis are proposed based on the terminal Qos policy.

Keywords: H.323; MCU; multipoint processors; multipoint controller; overflow algorithm; Qos

H.323 协议是 ITU-T(国际电信联盟)制订的一套在分组网上提供实时音频、视频和数据通信的标准^[1]。近年来以 H.323 为标准的多点会议系统得到广泛的发展和运用,很多的厂商都对多点会议系统进行深入的研究和开发。多点控制单元 MCU 是 H.323 协议中支持多点会议的终端,也是多点会议中最重要的组成部分,它负责整个会议的控制管理以及媒体数据的处理。因此,MCU 设计是整个多点会议系统的关键。

1 H.323会议系统总体设计

1.1 MCU 组成及功能介绍

多点控制单元(MCU)用以同时举行多个会议,会议个数理论上没有限制,但与实际选用的硬件性能相关,如在 Windows 操作系统下可使用的 SOCKET 端口的数量有关,MCU 接入网络的带宽和主机的处理能力如:CPU 型号、内存等。MCU 主要包括两个部分,多点控制器(MC)和多点处理器(MP)^[2]。MC 提供支持多点会议的管理控制功能,MP 在 MC 管理控制下负责对媒体流进行交换、混音以及数据处理。

MC 对多点会议的控制功能主要体现在两个方面:一方面负责会议控制,包括邀请终端加入会议、接受或拒绝终端与会要求以及将终端踢出会议等;另一方面是对通信的控制,包括终端能力协商、通信模式选择及逻辑通道建立等。主要功能:获取会议(成员)列表、呼叫连接、能力协商、会议记录保存与查询、维护逻辑信道、结束会议等^[3]。

MP 存在于集中或混合多点会议中,接收来自终端的音频、视频和数据流,经过处理后将这些媒体流送回相应的终端区。音频处理功能,指通过选择、混合、切换或组合操作将多路音频信息处理后再发送相应终端;视频处理功能,提供视频切换及混合等功能。

1.2 MCU 会议系统模式

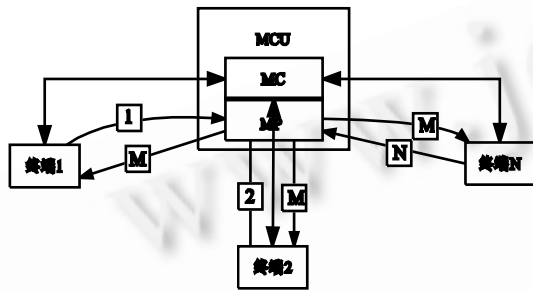
MCU 提供在以太网上三个以上的终端和网关参与多点会议的能力。MCU 至少存在一个 MC,根据选择会议模式不同决定 MP 存在或不存在。如:集中式多点会议的 MCU 是由一个多点控制器 MC 和一个音频的多点处理器 MP、一个视频的多点处理器和一个数据的多点处理器组成;而典型的支持分布式多点会议

① 收稿时间:2009-09-22;收到修改稿时间:2009-11-15

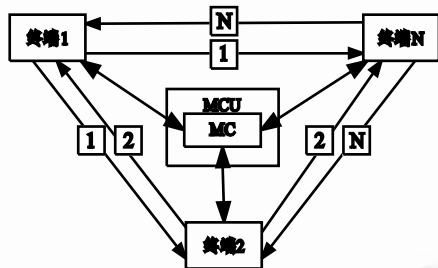
的 MCU 则仅有一个多点控制器 MC，视频、音频和数据的多点处理器则分散在各个终端处。

1.2.1 集中式多点会议系统

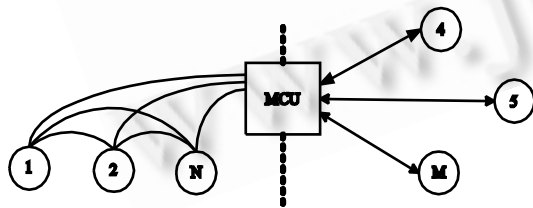
在集中式多点会议系统中如图 1(a)所示，所有终端(包括网关)都通过点对点方式(逻辑上的连接)与 MCU 中的 MC 建立 H.245 控制信道连接，并打开媒体通道，通过音频、视频及数据通道与 MP 进行通信，由 MP 对媒体数据进行处理。MP 可以实现不同音频、视频及数据格式和传输速率之间的转换，允许终端通过不同的通信方式加入会议。每一终端将自己的媒体流(音频、视频、数据、控制)传给 MCU，MCU 集中处理后再选择将相应的媒体流分发给终端。



(a) 集中式多点会议系统模式



(b) 分布式多点会议系统模式



(c) 混合式多点会议系统模式

图 1 多点会议系统模式

集中式会议又分为两种会议模式，主席模式和自由模式。在会议的进行过程中，可由主席动态地指定会议模式为自由模式或主席模式。自由模式下，会议终端可以自由进入或离开会议，各终端接收所有发言

终端的声音混合，接收的视频流有两种形式：一种与会成员较少，每个终端接收其它终端的混合视频流；另一种由 MCU 根据声音大小选择某一终端的视频图像发送给其它终端。主席模式下，会议中只有一个终端在发言，其它终端只接收演讲者的音频流和视频流。终端存在三种角色：会议主席(会议的管理者)、演讲者、听众。当没有演讲者的时候，听众接收会议主席的音、视频流。会议由会议主席发起，由某终端在 MCU 上创建会议终端成为会议主席，它即参与也管理会议。会议主席的权限包括：发起会议、控制会议、结束会议^[4]。普通参会的终端，可以申请作为演讲者，由会议主席批准或结束发言。主席模式下，最重要的是为用户数据成员角色做好类型标示；而自由模式下，由 MCU 在上一发言者完毕后选择下一个发言者。

1.2.2 分布式多点会议系统

在分布式多点会议系统中如图 1(b)所示，所有终端节点都是平等的，各终端仍以点对点方式和 MC 建立 H.245 控制信道连接。MCU 的功能以 MC 和 MP 功能模块的方式存在于系统中终端中，如：MC 可以位于 MCU、网关、关守或某个终端中。会议管理和媒体数据流处理均由终端节点来实现，每个终端将自己的媒体流分发其它同类的端点而不是通过 MCU。终端具有多播能力，可以通过音频、视频通道向其它会议终端发送媒体数据。MC 仍然具有对会议的控制功能。

1.2.3 混合式会议系统

混合式会议系统则是结合了集中式和分布式多点会议的特点如图 1(c)所示，部分数据流传给 MCU，另外一部分直接传给参加会议的终端节点。例如音频可以由集中式 MCU 控制，同时视频却在所有终端间分布式多目广播。同样，会议中一些终端可以以集中方式参与，而另一些终端允许以分布式方式参与。

选择不同的会议系统模式，对 MCU 和终端的要求各不相同，集中式多点会议系统对 MCU 要求高，而分布式会议系统对终端要求很高。在选择采取那种模式，要根据硬件的配置情况，要求网络交换机和路由要支持组播时候才能选择使用分布式会议系统，因此当前 MCU 多数都采用的是集中式的会议系统。

2 MCU多点会议系统研究

2.1 多点控制单元模块(MC)

多点控制单元模块(MC)是整个 MCU 系统的主控

部分,从功能上划分为呼叫控制和会议管理两大模块如图2所示。呼叫控制模块负责对会议成员的呼叫控制,包括呼叫的建立、控制信道的建立、主从确定和能力集交换以及控制媒体信道的建立和释放。会议管理模块负责整个会议的管理,包括会议的建立、会议成员的加入、删除、查找以及会议成员身份确定;控制会议开始与结束以及会议信息的记录与更改等功能[5]。

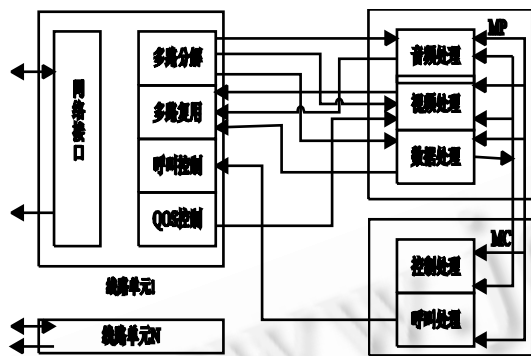


图2 MCU构成图

2.2.1 呼叫控制模块

在呼叫建立过程中,MCU即可充当会议的发起者也可以是被叫一方,不管是主叫还是被叫方在呼叫建立的Setup消息中都要包含以下两点内容:一是选择是新创建一个会议还是参加一个已存在的会议,二是会议的别名(conference Alias)。当终端与MC建立H.245控制信道连接后,进行主从确定和能力集交换,在能力交换过程中选择会议的形式,会议形式的选择要根据端点和MC的能力来确定。对于每个新加入的终端,MC通过信令(terminal NumberAssign)分配一个终端号给新加入端点,并通知其它终端新端点的加入。MC将分别为每个成员建立四个媒体信道:音频的输入输出和视频的输入输出信道。会议成员要求离开会议,MC接收到终端发来的离开的请求后,首先关闭媒体信道,最后关闭H.245控制信道。

2.2.2 会议管理模块

会议管理模块在MC中负责对会议以及会议成员的管理。MCU支持同时举行多个会议,以会议别名标示。每组会议成员的管理由链表类来实现,链表类对象中每个元素唯一对应一个该会议的成员,包括成员的全部信息如:用户名、别名、IP地址及端口号等。在呼叫建立Setup消息中包含呼叫别名(destina-

tionAdres)、呼叫传输信令(destCallSignalAdres)、会议号CID。通过对呼叫链表类的每个元素的添加、删除、查找等操作,实现对会议成员的创建、清除、查找等[5]。

2.2 多点处理器

2.2.1 音频处理

音频处理是多点会议系统中较为重要的功能部分,音频处理最重要的任务就是对音频流的实时混合,混音算法如何处理语音数据是关键。实时混音算法是音频处理的关键部分,尤其是在数据溢出的处理上。

实时混音需要注意几方面问题:一是不同音频流要转换后才能进行混音;二是混音的数据上限溢出问题;三是本地声音抑制;四是声音滤波处理。假设一组会议中有M个终端,即有M个音频输入流,每帧语音流有N个样本(即采样值), $Out[i]$ 表示当前混音输出第i个样本, $In[i,j]$ 为当前第j个音频流第i个样本输入(其中 $0 \leq i \leq N$, $0 \leq j \leq M$),按照一般线性叠加原则如下:

$$Out[i] = \sum_{j=0}^{M-1} In[i,j] \quad (1)$$

由于多个样本按照(1)式直接叠加,结果可能溢出,产生噪音。为了解决混音溢出,常用的有以下几类方法:

一是均值法,对 $Out[i]$ 取平均值作为混音输出:

$$Out[i] = \frac{1}{M} \sum_{j=0}^{M-1} In[i,j] \quad (2)$$

二是箝位法,当有溢出,则溢出的样本用最大和最小临界值代替:

$$Out[i] = \begin{cases} MAX & , \quad Out[i] > MAX \\ Out[i] & , \quad MIX \leq Out[i] \leq MAX \\ MIX & , \quad Out[i] < MIX \end{cases} \quad (3)$$

三是加权法,即利用权重 $w_j[i]$ (第j路语音第i个样本值)将语音样本值缩小:

$$Out[i] = \sum_{j=0}^{M-1} In[i,j] * w_j[i] \quad (4)$$

四是最大值法,当有溢出时,选择语音样本值最大的为输出:

$$Out[i] = MAX \{ In[i][0], In[i][1], \dots, In[i][M-1] \} \quad (5)$$

表 1 溢出算法性能比较

方法	均值法	箱位法	加权法		最大值法
			固定加权	贡献加权	
复杂度	M次加法、 一次乘法	M次加法、 M次比较	M次加法、 M次乘法	M平方次	M次加法、 N次比较
方法	较差	较差	一般	较好	一般
不足	音量变小，	波形失真、引 入噪声，	不适应语音 的实时变化	计算时延 最大	新成员加入 引起声音不 连续

一个会议中终端要求接收的通常是别人的音频流混音，自己的音频数据不需要被混音，一般采用图 3 中所示两种方法。图 3(a)是先由将所有语音数据混音后，在每个终端处将自身音频流减去；图 3(b)是每个终端都带有一个混音器，接收除自身外的其它终端发送来的音频流进行混音。

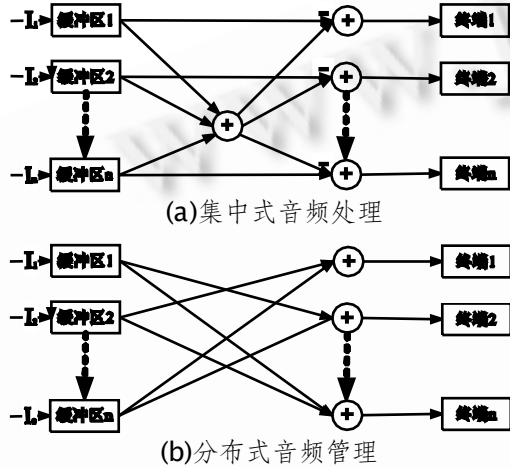


图 3 本地声音抑制

2.2.2 视频处理

在 MCU 的视频处理中，会议处理模块主要执行视频的复合及切换功能。在每组会议中，MCU 为每个终端建立一个缓冲区，既是接收缓冲区也是发送缓冲区。视频复合在多终端参与的会议中，利用发言者检测算法检测将最新发言人加入到视频缓存列表，采用先进先出的方式对视频缓存列表进行更新，选择四路视频经过复合后发送至各终端。视频切换，在会议主席或主持人根据需要，选取需要复合的记录终端进行复合。

由于 H.323 采用的是 RTP/RTCP 网络传输协议进行媒体数据的传输，当网络差丢包率较高的时候，对视频会议影响很大，因此，在多视频会议情况下一方面要采取技术措施保证网络传输的稳定，另一方面

可以通过 MC 对视频大小进行调整减小占用的带宽，来保证视频的连贯性。

2.3 服务质量 QoS 研究

IP 网络的 QoS 问题研究是当前的热点问题，目前常用的 QoS 解决技术，如资源预留(RSVP)、区分服务(DiffServ)、多协议标记交换(MPLS)等，多是基于网络传输上的 QoS 解决方案。而对于多点视频通信来说更多的是采用基于终端的 QoS 技术，终端的 QoS 技术主要包分为：拥塞控制和差错控制。

2.3.1 拥塞控制

由于在实时多媒体业务中，通常采用 UDP 协议传输，因为 TCP 重传造成的时延对于实时业务的影响是不能忍受的^[6]。通常在网络终端处采用基于速率控制的方法，即发送方根据接收方的反馈信息，计算当前网络可承载的最大码率，并以此作为根据调整编码器的输出。

基于速率的控制方法就是利用端对端的 RTCP 反馈控制信息，从中估算出数据包的分组丢失率和分组延迟抖动信息。一般来说数据包分组丢失主要是两方面原因，缓冲区溢出和比特错误。通常终端会对包丢失率 Ploss 设置一个门限 Pthreshold，当丢失率超过 Pthreshold 时候，可以通过减少发送的帧率或减少混合的视频终端数，减少网络带宽的使用量。在选择上有两种保证机制，一是以损失质量为代价保证视频传输的数量；另一是以减少混合视频数量保证传输质量。

2.3.2 差错控制

基于 RTP 的差错控制通常有重传法和前向纠错(forward error correction, FEC)法。由于视频会议对实时要求比较严格，所以侧重于选择前向纠错法(前向纠错编码法)。重传丢失或错误的数数据一方面增加网络和系统的处理的压力，另一方面增加网络延时，对于实时的视频通信是不可容忍的。因此多采用前向纠错编码的方法 FEC，FEC 采用特殊的 RTP FEC 包结构，通过一定的解码和纠错算法，将丢失或错误的包恢复过来。

3 结语

本文主要简单介绍了基于 ITU-T H.323 建议下的多点控制单元 MCU 的设计，对多点控制和多点处理模块的功能和结构进行详细分析。分析了在多点处理模块设计中需要注意的几方面问题，并对常用的混

(下转第 182 页)

(上接第 4 页)

音溢出算法进行了详细的说明。在工程应用多点控制单元设计中要考虑的问题是多方面的,本文给出的解决方案是当前应用较多的设计方案,主要适用于小规模的工作平台下。

参考文献

- 1 H.323v4-2001, Packet-based multimedia communication system, ITU-T Recommendation, 2001. 3 - 12.
- 2 刘瑞芳,孟庆昌,等. H323 视频会议中 MCU 的设计与

实现.北京邮电大学学报, 2003,26(2):77 - 81.

- 3 刘岩.基于 H.323 协议的 IP 视频会议中 MCU 的研究与实现[硕士学位论文].郑州:信息工程大学, 2006.
- 4 赵娜. IP 网络视频会议的 MCU 的设计和实现.现代计算机, 2005,10:41 - 44.
- 5 颜金辉,徐捷.基于 H.323 的 MCU 的分析及设计.计算机与数字工程, 2008,230(12):173 - 177.
- 6 林闯,单志广,任丰原.计算机网络的服务质量.北京:清华大学出版社, 2004.

182 经验交流 Experiences Exchange