



[12] 发明专利申请公开说明书

[21] 申请号 200410076930.1

[43] 公开日 2005年3月9日

[11] 公开号 CN 1592276A

[22] 申请日 2004.8.31

[21] 申请号 200410076930.1

[30] 优先权

[32] 2003.9.2 [33] GB [31] 0320540.8

[71] 申请人 西门子公司

地址 联邦德国慕尼黑

[72] 发明人 E·霍尔瓦斯 K·克拉霍菲

G·-D·施帕尔

[74] 专利代理机构 中国专利代理(香港)有限公司

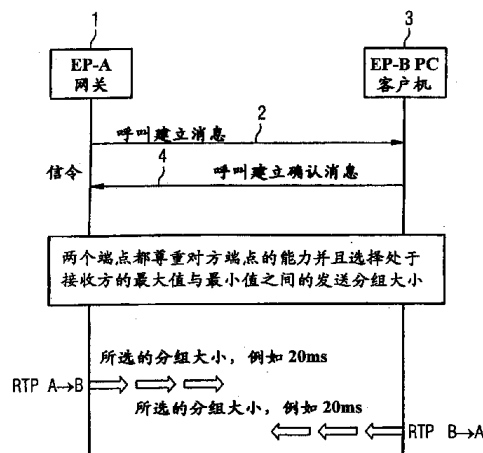
代理人 杨凯 张志醒

权利要求书1页 说明书7页 附图2页

[54] 发明名称 控制网络上的音频通信的方法

[57] 摘要

一种控制网络上音频通信的提供的方法，该网络包括至少两个端点(1、3)，其中至少一个端点作为源，至少一个端点作为目标，所述方法包括：在源上设置所需的最大和最小分组大小；在目标上设置所需的最大和最小分组大小；确定最小发送分组大小为源设置的所需最小值和目标设置的所需最小值中较大的值；在目标中把抖动缓冲器设置为针对所确定的最小发送分组大小的适当大小；以及发送音频分组，其大小大于或等于所确定的最小发送分组大小。



1. 一种控制网络上音频通信的方法，所述网络包括至少两个端点，其中至少一个端点作为源，而至少一个端点作为目标，所述方法包括：在所述源上设置所需的最大和最小分组大小；在所述目标上设置所需的最大和最小分组大小；确定最小发送分组大小为所述源设置的所述所需最小值与所述目标设置的所述所需最小值中较大的值；在所述目标中将抖动缓冲器设置为针对所述确定的最小发送分组大小的适当大小；以及发送大小大于或等于所述确定的最小发送分组大小的音频分组。

2. 如权利要求 1 所述的方法，其特征在于，所述目标通知所述源所述最小分组大小的增加；以及发起所述抖动缓冲器的大小的相应增加。

3. 如权利要求 1 所述的方法，其特征在于，所述源通知所述目标所述最小分组大小的增加；以及所述目标发起所述抖动缓冲器的大小的相应增加。

4. 如权利要求 3 所述的方法，其特征在于，所述源的所述最小分组大小设置在所述目标所设置的最小分组大小与最大分组大小之间。

5. 如上述任一权利要求所述的方法，其特征在于，所述源包括 IP 网络中的终端、服务器和网关其中之一。

6. 如上述任一权利要求所述的方法，其特征在于，所述目标包括 IP 网络中的终端、服务器和网关其中之一。

7. 如上述任一权利要求所述的方法，其特征在于，所述网络采用基于因特网协议的语音传输(VoIP)。

8. 如上述任一权利要求所述的方法，其特征在于，所述音频通信包括话音通信。

控制网络上的音频通信的方法

5 技术领域

本发明涉及控制网络上的音频通信的方法，具体地说，涉及 VoIP 系统。

背景技术

10 基于因特网协议的语音传输(VoIP)的特征在于如下事实：发送方编译连续音频流(PCM：每 125 微秒 1 字节)，然后将此音频流打包成分组，并通过局域网(LAN)接口发送这些分组，但是不限于在 LAN 中使用。根据编码方式，存在对分组大小的不同最小值要求(G.711：1 字节；G.723.1：240 字节；G.729A：80 字节)。为减少所需的带宽，
15 发送方还可发送包含多个此最小长度的分组。

接收方依次接收经过 LAN 接口的分组，并将它们重新转换为连续音频流。音频流的这种转换(以及如果合适的压缩)一般由称为数字信号处理器(DSP)的器件来执行，而在 DSP 与 LAN 接口之间的分组传输由中央处理器(CPU)来执行。

20 CPU 上的负荷不取决于分组类型，而仅取决于这些分组到达的速率。分组大小越小，分组速率越高，因此 CPU 负荷以及带宽要求也越高。另一方面，分组速率越高，嘴到耳的延迟越小，因此用户的通话质量也越好。

25 在具有大量信道的网关的情况下，现在的问题是，为了获得理想的通话质量，终端以高分组速率发送分组，导致网关上的 CPU 负荷较高，可能太高而使 CPU 无法应付。

目前，没有用于控制呼入语音分组的速率的设施，因为在协议中只有一个用于最大分组大小的字段，该最大分组大小对应于最小

分组速率。因此，终端的分组速率取决于折衷速率，一方面不使网关过载，另一方面在两个终端直接连接的情况下使通话质量尽可能地高。

5 发明内容

根据本发明，一种网络包括至少两个端点，其中至少一个端点用作源，而至少一个端点用作目标，控制所述网络上的音频通信的方法包括：在源上设置所需的最大和最小分组大小；在目标上设置所需的最大和最小分组大小；确定最小发送分组大小为源设置的所需最小值与目标设置的所需最小值中较大的值；在目标中将抖动缓冲器设置为针对所确定的最小发送分组大小的适当大小；以及发送大小大于或等于所确定的最小发送分组大小的音频分组。

本发明在语音编码的协商期间在源和目标中都设置最大分组大小和最小分组大小，并将设定结果用信号通知给发送音频通信的源。最大分组大小符合现有标准以确保最小分组速率，此外，确定最小分组大小，以便得到以嘴到耳延迟的形式表示的通话质量与资源需求之间的平衡关系。对于 IP 端点，这意味着达到了可接受的 CPU 负荷，以及在网络中满足带宽要求。用于发送的最小分组大小不得小于目标所设置的用于接收的最小值。

在一个实施例中，目标通知源最小分组大小的增大；以及发起抖动缓冲器的容量的相应增大。

如果 CPU 负荷看来变得过高，则目标在呼叫期间可将此变化通知给源，并在通话中断期间，增大抖动缓冲器的容量。

或者，源通知目标最小分组大小的增大；以及目标发起抖动缓冲器的容量的相应增大。

源也可选择增大最小分组大小，并将此增大通知目标。

源的最小分组大小最好设置在目标所设置的最小分组大小与最大分组大小之间。

源和目标最好包括 IP 网络中的终端、服务器以及网关中的任何一个。

可能存在各种各样的组合，包括终端用作源而网关用作目标；或者两个终端，一个用作源而另一个用作目标，或者它们的其它组合。

通常，网络采用基于因特网协议的语音传输(VoIP)，包括载波 VoIP 系统。

附图说明

下面参照附图说明控制包括至少两个端点的网络上的音频通信的提供的方法的例子，其中：

图 1 说明采用本发明的方法在两个端点之间消息往来的例子；

图 2 说明作为分组大小的函数的主处理器的 CPU 的利用率；以及

图 3 说明对于三种不同编解码器、作为分组大小的函数的每个信道的带宽要求。

具体实施方式

每个端点(EP)具有两个参数。这些参数中的第一个参数是端点可接收分组的最小分组大小，第二个参数是端点可发送分组的最小分组大小。发送分组大小不得小于目标 EP 所设定的最小值，而源 EP 不得超出它自己的性能极限。因此，源 EP 则可选择发送这样的分组，其大小范围是从源 EP 用于发送的最小分组大小与目标 EP 用于接收的最小分组大小中较大的值到目标 EP 的用于接收的最大分组大小。

传统 H.323 系统通过 UDP/IP 传送作为 RTP 流的音频数据。音频流的特性经由作为终端能力交换程序的一部分的 H.245 消息、在为音频流开通逻辑信道的过程中被交换。音频流参数之一是每个分组的最大帧数。该值表示一个分组所传送的最大时间间隔，因此，也

表示分组大小以及音频源发送分组的分组速率。源不得超过此参数所规定的分组大小，但是可以更频繁地发送较小的分组。例如，对于 G.711 流，可能已经协商好 20 毫秒的最大间隔(160 帧/分组)，但是源可以选择每隔 10 毫秒(80 帧/分组)或者甚至更频繁地发送分组。

5 出于 QoS 的原因，希望较高的分组速率(较小的分组大小)，因为在此情况下，音频帧以较低延迟发送，且丢失分组的影响不太严重。可是，较小的净荷大小意味着相对较高的分组开销以及较低的带宽效率。更频繁的分组到达还将增加接收机的处理负荷。一种实现必须考虑效率和 QoS 要求之间的这种权衡。对于大型网关，在高业务
10 负荷期间，分组到达速率可能变得至关重要。如果在这种情况下，发射机通过加大分组大小来降低分组速率，那么网关的性能可得到改善，使网关可以维持较高数量的呼叫，但存在个别呼叫的 QoS 有所下降的风险。

本发明引入了作为两个实体之间 VoIP 信令的一部分的最小分组
15 大小表示。在本发明的一个例子中，将最小分组大小(最小帧数/分组)添加到 H.245 消息中。然后，要求发送侧发送分组的频率不高于此参数所确定的频率。可作为终端能力的一部分并在逻辑信道信令期间规定该参数。另一个特征是分组大小(即分组速率)对业务负荷的动态适应的规定。例如，如果网关中达到临界负荷，则将以 10 毫秒/分
20 组的分组大小开始的呼叫提高到 30 毫秒/分组的大小。对于这种适应，网关需要一种告知发送侧暂时降低其分组速率的装置，类似于 H.245 中现有的流量控制机制。

图 1 中，通过举例说明针对采用本发明的方法的 VoIP 系统中的
净荷来控制接收分组速率。图 1 表示了 VoIP 分组网络中的两个 VoIP
25 端点之间用于控制发送分组大小的典型消息流。端点 A 1(此例中是网关)向端点 B 3(此例中是个人计算机)发送呼叫建立消息 2。端点 B 发回呼叫建立确认消息 4。VoIP 端点可包括网关、IP 电话、PC 或服务
器端点上的软件客户机。可使用各种 VoIP 信令协议，例如标准

H.323[ITU-T 建议 H.323 - 基于分组的多媒体通信系统]、标准 SIP[RFC 3261 SIP - 会话发起协议]或者专有 VoIP 协议。采用 RFC 1889 RTP/RTCP - 实时协议/实时控制协议通过分组网络来发送用户数据。

在图 1 的例子中，端点 A 1 是接收方向上的端点(EP)，对于例如
5 G.711 编解码器，要求具有 20 毫秒的最小分组大小的音频媒体流。
这在呼叫建立消息($G711PSRvMaxA = 60$, $G711PSRvMinA = 20$)中指明，对于 G.711 编解码器，该 EP 在接收方向还要求 60 毫秒的最大
10 分组大小。最小分组大小使 EP-A 能确保在 EP-A 中每单位时间的呼
入实时协议(RTP)语音分组的数量可保持相对较少，例如在 EP-A 中
10 为了性能，这样做可能是必要的。EP-A 可以是必须能同时处理几个
包括音频流的呼叫的网关(每单位时间要处理的分组越多，对 EP 的
性能要求就越苛刻)。

对于本例，EP-B 3 是具有 H.323 软件客户机的 PC。该 PC 具有
15 足够的的能力，并且通常仅需要能一次处理一个呼叫(媒体流)。因此，
EP-B 也提供接收方向上的 60 毫秒的最大分组大小，但是对于 G.711
编解码器提供 10 毫秒的最小分组大小，这是相对较小的分组大小，
因为 EP-B 每单位时间能处理相对较大数量的接收分组，这一事实例
20 如对端到端延迟有正面影响。在呼叫建立 2 和呼叫建立确认 4 消息
期间，端点 1、3 将用于 G.711 编解码器(但也可以是其它所支持的编
解码器，诸如 G.723.1 或 G.729)的相关的最小和最大接收分组大小发
信号通知给对方端点 1、3。本例的呼叫建立确认消息是
25 $G711PSRvMaxB = 60$, $G711PSRvMinB = 10$ 。例如，在 SIP 标准的情
况下，例如通过在 SIP INVITE 或者 200 - OK 消息中的 SDP[RFC 2327
SDP - 会话描述协议]中扩展的方式，发信号通知新参数。在 H.323
的情况下，例如通过在 H.245 标准[ITU-T 建议 H.245 - 多媒体通信的
控制协议]中单元音频能力或一般能力扩展的方式，发信号通知新参
数。在 H.323 和 SIP 标准中，只公开了最大分组大小的细节。在 SIP
标准中，有关于推荐的分组大小的要素，但这不是强制的。也没有

任何标准包含了最小分组大小的详细说明。

本发明的另一个特征在于，如果诸如在呼叫建立时导致定义并发送初始最小和最大分组大小值的 EP 中性能情况、网络中带宽等标准在呼叫期间发生改变，则可能出现包括随后修改的最小或最大值的能力的重复交换。这些修改后的值必须受到对方端点尊重。在 SIP 系统中，这可以例如通过利用扩展的参数的 RFC 3264 SDP 提供/应答模型[RFC 3264 SDP 提供/应答模型]发信号通知。在 H.323 系统中，这可以例如通过 H.245 终端能力设置消息(具有扩展的参数集)的重复交换或者在 H.460.6 标准[ITU-T 建议 H.460.6 - 扩展的快速连接]中通过扩展的快速连接来进行。如果接收方在呼叫期间请求发送方增大分组大小，则为了避免语音中断，在发送增大分组大小的请求之前，也必须在通话暂停时将抖动缓冲器增加此增量。或者，发送方可以通知接收方：分组大小在输出方向上正被增大。此通知随后被用来在通话暂停时将抖动缓冲器增加此值。

另一种选择是，为了在服从资源限制的同时设置最佳语音质量，源 EP1、3 将从它自己的用于发送的最小大小以及由对方 EP 设定的用于接收的最小大小中选出的最大可能值选择作为发送方向上的分组大小，即，从这两个值中选择较大的值。在一个示例中，网关可考虑到通话质量来配置，即，用于 G.711 的分组大小可以被例如从 240 字节减少到 80 字节，在这种情况下，仅仅支持信道的 1/3。此时，仅能控制自己的发送方向。

图 2 和图 3 说明分组速率对资源的影响。图 2 表示了作为分组大小的函数的主处理器的 CPU 的利用率。CPU 能力利用率不取决于所选的编解码器。CPU 工作负荷与分组大小成反比，即，如果分组大小减半，则在两个分组之间的时间减半，分组速率以及 CPU 工作负荷加倍。

在图 3 中，表示了作为分组大小和所选的编解码器的函数的以太网中每个信道的带宽要求。在压缩后实际净荷数据所用的带宽，

在 G.711 的情况下由曲线 5 表示为 64 千比特/秒，在 G.729A 的情况下由曲线 6 表示为 8 千比特/秒，在 G.723.1 的情况下由曲线 7 表示为 6.4 千比特/秒。然而，对于每个分组需要如下协议首标：RTP 12 字节、UDP 8 字节、IP 20 字节、以太网(包括前同步码和帧校验序列，
5 但没有 VLAN 标记)26 字节。该首标组成部分增加了根据分组速率实际所需的带宽。

然而，由于编解码器本身的限制，因此不是所有的分组大小都是可行的。G.729A 仅允许 10 毫秒的倍数，而 G.723.1 允许 30 毫秒的倍数。因此，图 2 中关于 G.729A 和 G.723.1 的曲线 6、7 是理论上
10 计算的，因为事实上只有以方块或三角表示的离散值可以实际出现。

本发明实现了低成本 IP 网关解决方案，而不会丧失在 IP 电话之间呼叫的高通话质量的优点。

15

图 1

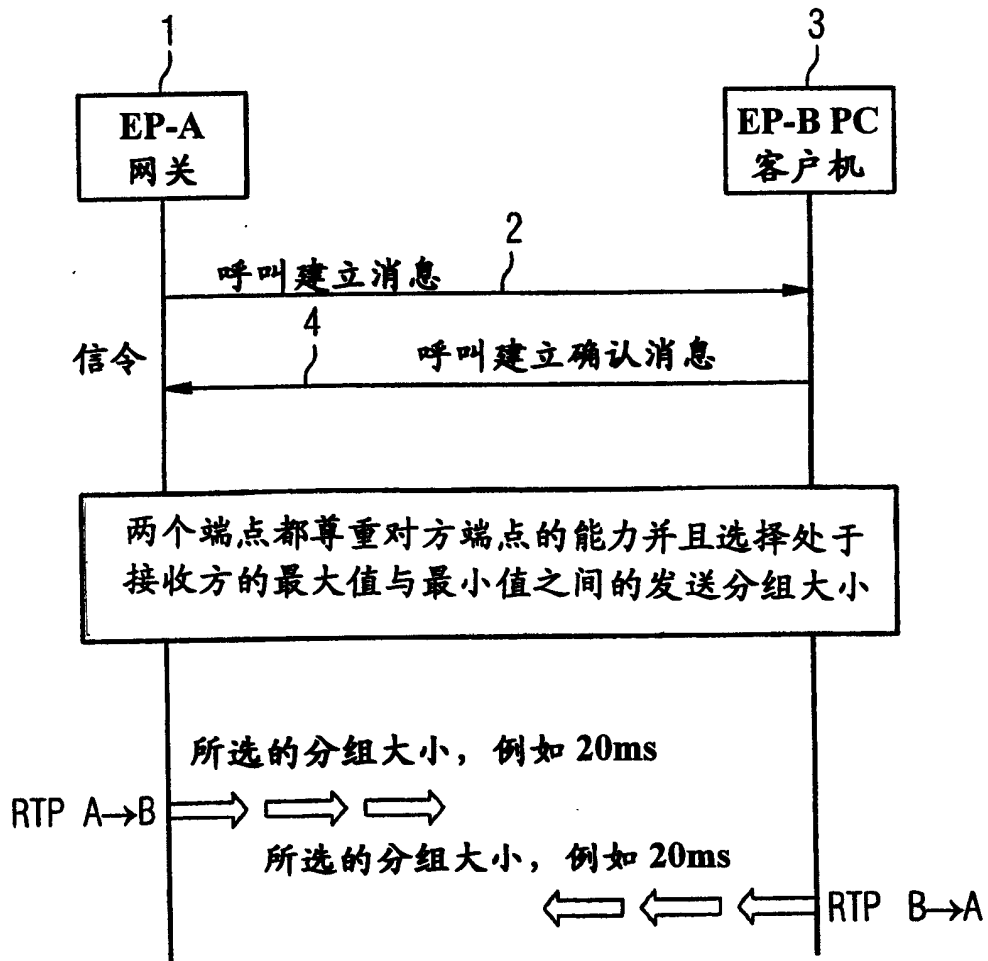


图 2

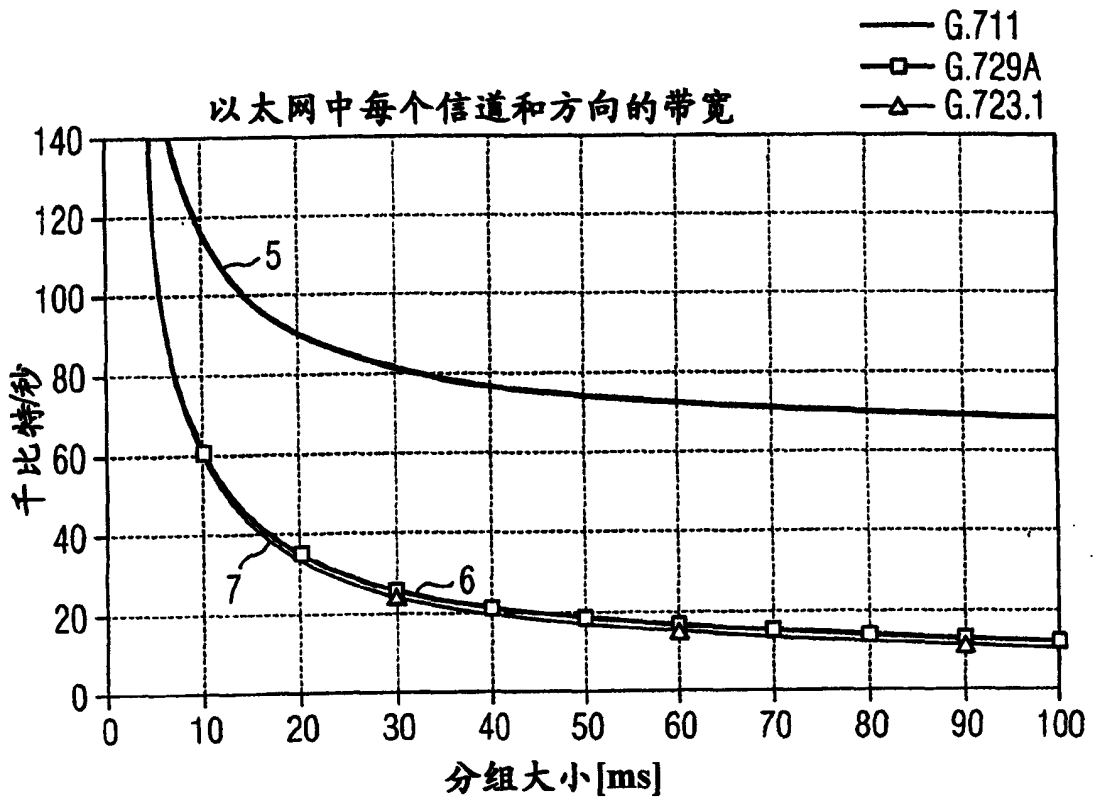
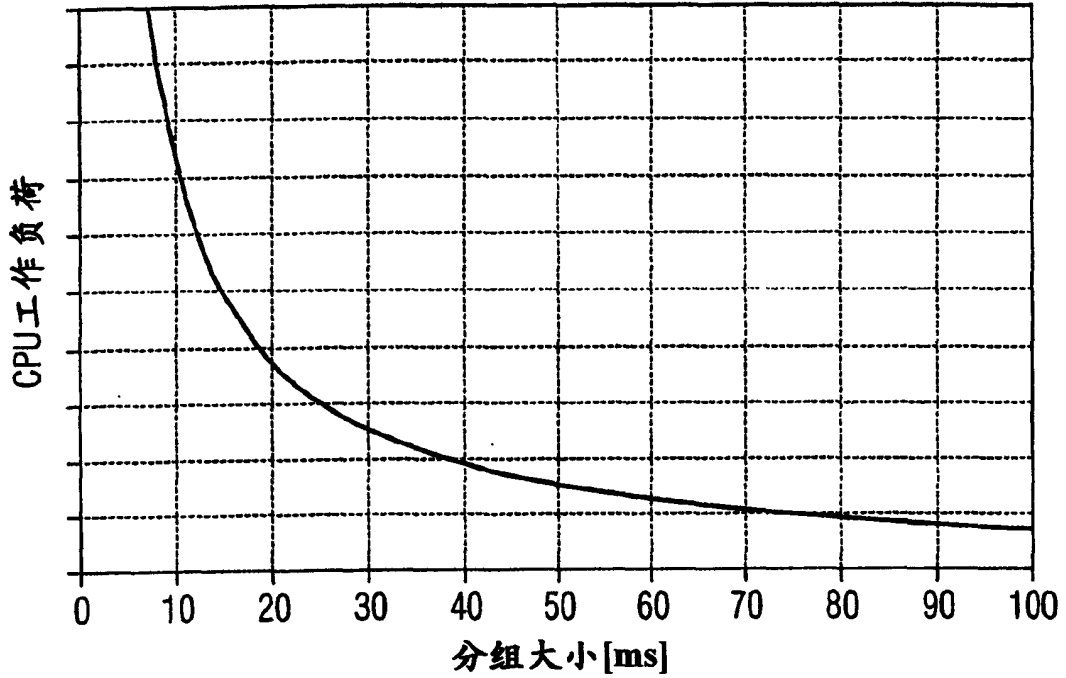


图 3