

基于校园 LAN 的 IP 电话设计与实现^①

Design and implementation of Ethernet based IP phone system

李志浩 唐红 (重庆邮电学院 400065)

摘要:介绍了基于校园网络的 IP 电话系统的设计和实现方法。系统利用校园局域网这一高效的通讯介质,结合当前先进的 CTI(计算机电信集成)技术,针对目前 Ethernet 的特点进行性能优化,实现了局域网与 PSTN 电话网间的拨号及语音通信。

关键词:IP 电话 CTI Ethernet 时延

1 引言

本文介绍的局域网 IP 电话系统,利用目前先进成熟的数字语音通讯技术而开发实现的。它不仅可以节省可观的电信网络建设费用和使用费用,同时还能充分利用已有的网络资源。

布式设计,实现局域网内终端用户与电话用户的实时语音通信。它由 IP 电话网关服务器和客户终端组成,工作在 100BASE-T 局域网。整个系统的拓扑结构如图(1)所示。

2.1 硬件构成

服务器端兼具了网关和网守的功能,它通过 PSTN 接口卡建立到 PSTN 网络的连接,标准以太网卡建立到 IP 网络的连接,IP 语音资源卡则在 IP 网与 PSTN 网之间建立起通信桥梁。IP 电话网关服务器是分组网络的一个端点,用于在分组网络和电路交换网两侧完成呼叫的建立和清除,并进行音频数据编码格式的编译。总的来说,网关就是在分组网络端点和电路交换网络端点之间进行透明的映射。局域网客户端可通过统一的 IP 电话界面,

经由 IP 电话网关服务器拨叫 PSTN 的用户。

2.2 软件构成(图 2)

(1) 服务器端软件模块

PSTN 接口卡模块:负责实现与 PSTN 外线电话的实时接续和正常的语音通信;

IP 资源卡模块:实现 RTP 数据包的解包,DSP 的轮询,语音数据的编/解码处理,压缩数据的 RTP 打包处理,通过 RTP/UDP/IP 方式完成 RTP 包的发送和接收;

地址映射模块:通过数据库查询,实现局域网内用

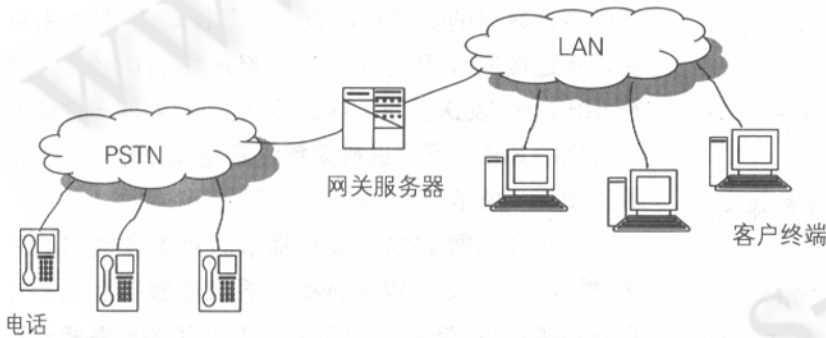


图 1 系统拓扑结构图

目前的校园网络普遍采用的是 100BASE Ethernet 网络,该网络的设计初衷不是面向语音、视频等实时业务的,考虑到分组丢失、延迟等对实时通讯的影响,本系统在 ITU-T 的 H.323 协议基础上,针对网络的实际情况和实际应用中的需求,采取了一些初步的解决措施,取得了较好的效果。

2 系统设计与实现

局域网 IP 电话系统采用客户/服务器模式进行分

① 基金项目:教育部春晖计划和重庆市科技攻关项目(编号:7220-13-20)资助

户分机号与对应 IP 地址的解析。

(2) 客户端软件模块

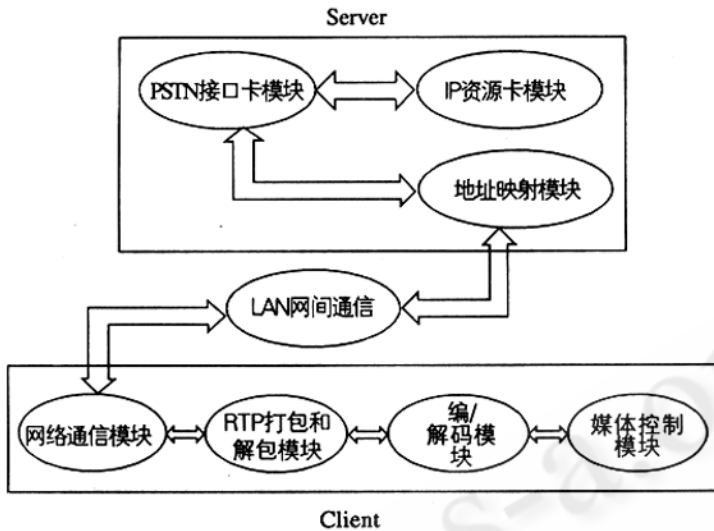


图 2 系统软件结构模块图

网络通信模块:通过 RTP/UDP/IP 协议完成语音数据包的发送和接收;

RTP 打包和解包模块:将压缩语音数据进行 RTP 封装以及对接收到的 RTP 数据包进行解包处理;

编/解码模块:通过 G. 723.1 语音压缩算法,完成语音编/解码功能;

媒体控制模块:实现对语音数据流的采集和播放。

2.3 系统功能

PSTN 用户可作为主叫方发起呼叫。当 PSTN 电话用户拨打网关服务器对应的接入号码后,网关系统将在两声振铃后自动接听并向 PSTN 线路上播放语音菜单的提示音。用户根据语音向导的指引进入二次拨号环节,即输入分机号,网关系统将接收到的分机号通过地址映射模块映射到对应的局域网 IP 地址,然后对其发起呼叫,局域网用户发现有呼叫到来后可选择是否接听。一旦通话建立,网关系统将 PSTN 线路上的 PCM 数据压缩后发送到局域网终端,客户端将接收到的数据解压还原为原始语音播放出来;由局域网终端传来的语音数据在网关处解压缩后,经 PSTN 接口卡播放到 PSTN 线路上。通话双方依靠网关达到异种网络语音通讯的目的。

局域网终端用户作为 0 主叫方,只需在 IP 电话客户端界面输入外线电话号码,网关服务器将用户拨叫的 PSTN 外线号码通过接口卡送到 PSTN 交换机,从而

对外线用户发起呼叫,当电话接通后,接下来的过程与外线用户拨打局域网用户的情况是一致的。无论是 PSTN 用户还是局域网内的终端用户,一旦话音通讯通道建立成功,任何一方都可通过挂机操作结束此次通话。

2.4 客户端的实现及关键技术

客户端的主要工作是:对话音的实时采集、处理、播放,同时完成数据在网络上的发送和接收。对于语音的采集、播放,系统采用 Windows MDK 的底层音频编程函数实现。当应用程序不断向设备驱动程序提供音频数据时,设备驱动程序控制音频设备在后台完成录音和放音的具体操作,通过回调机制,又可以检测到何时用完一个数据块,并及时传送下一个数据块,从而保证了声音的连续性。有了这种单机上的实时采集、回放功能后,接下来的工作就是对采集到的音频数据进行处理。为了进行高压缩比的传送,系统采用 ITU-T 的 G. 723.1 算法进行语音数据的编码和解码处理。最后,实现在网络间语音数据的发送和接收。为了保证实时性,系统采用无连接的 UDP 协议并对已经压缩的语音数据进行 RTP 封装,发送至服务器端,同时也要对服务器端发送过来的 RTP 数据包进行解包和解压缩处理。其工作原理如图 3 所示。

(1) 音频数据的采集和播放。对语音数据的采集、播放,系统采用 Windows 底层音频服务实现。其主要功能包括:查询音频设备、打开和关闭设备驱动程序、分配和准备音频数据块、管理音频数据块等。使用低级音频函数之所以能够对各个声音数据块进行操作,要归功于 Windows 的消息映射机制。在声音采集与回放之前,首先要检查音频设备的能力,定义音频数据的结构来设置相应的音频参数。在使用底层音频函数对声音进行采集、回放时,声音存放在内存数据块中,当采集缓冲区中数据满(发出 MM_WIM_DATA 消息)时,以及回放缓冲区中数据为空(发出 MM_WOM_DONE 消息)时,将执行相应的消息映射函数来进行处理。

在录音和放音的程序处理中,有一个问题尤为重要:就是分配的内存缓冲区的大小和数量。缓冲区的大小与语音的连续性和延时有直接关系。缓冲区越大,则语音的连续性越好,但延时越大;反之,缓冲区越

小,则语音的延时越小,但连续性越差;所以要取一个折中值。经试验表明,大约每个内存块录音的时间长度为 0.1 秒比较合适。分配缓冲区的数量则与内存块的大小和对每个内存块中数据的处理时长有关。

一般应保证所有分配的内存块录音时间长度总和约为 0.5 秒左右比较恰当。

(2) 音频数据的编/解码处理。系统采用 ITU - T G. 723.1 算法进行音频数据的编解码处理,独立开发的 G. 723.1 算法模块以动态链接库的形式提供。主要导出四个函数:音频数据的编码初始化函数 G723_1_Init_Encode,解码初始化函数 G723_1_Init_Decode,编码函数 G723_1_Encode 和解码函数 G723_1_Decode。前两个函数用于对音频数据进行编码和解码处理时变量的定义,内存块的分配等初始化工作;后两个函数分别实现了 G. 723.1 算法对音频数据的编码和解码处理。

本系统中还在 G. 723.1 算法中使用了静音抑制技术,即检测通话过程中的安静时段,并在这些静音期间,只发送极少量的静音语音包。通过这种方式进一步节省网络带宽,有利于其它业务在现有网络上传输。

(3) 实时传输协议与音频数据封装。如何提高实时性,确保通信的 QoS 是 VoIP 中的关键技术要求。系统采用了 IETF 提出的实时传输协议 (Real-time Transport Protocol) 对语音压缩数据包进行封装传输。RTP 协议是 IP 网络中针对实时业务的一种传输协议,一般架构在 UDP 协议之上。另外,它也是一个数据封装协议,实时业务数据封装于 RTP 包的数据域中。RTP 提供序列号 (Sequence Number) 以恢复数据包的顺序,实现丢包检测,为实时传输提供网络拥塞等信息;提供时间戳 (Timestamp) 用于媒体同步,使接收端按正确的速率回放数据;提供同步源标识 (Source Identification) 使接收端可获得发送端的相关信息。

在此要注意封装效率与全程时延的关系。在 RTP 包中打进的压缩数据越多,封装效率越高,线路带宽利

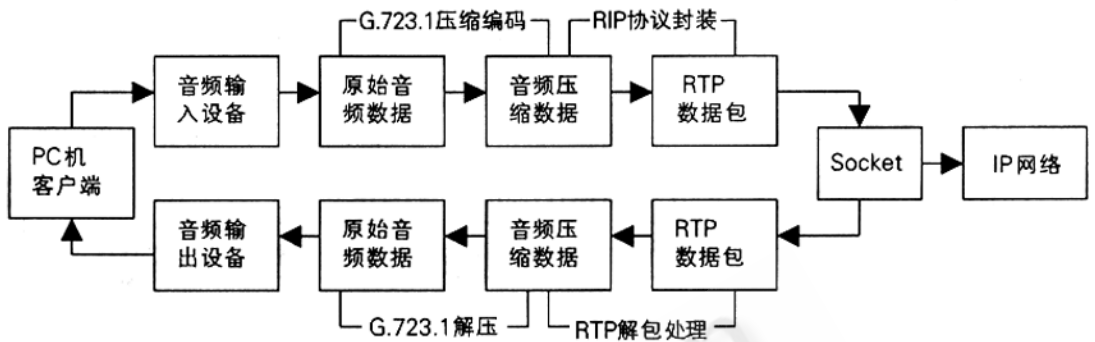


图 3 局域网终端工作原理图

用率也越高;与此同时,由于 RTP 包所携带的语音数据量增加,单个 IP 数据包的传输时延加大,导致全网时延增加,给听者最直接感觉就是“对方反应迟钝”。一般在单向时延(包括编码时延、打包时延、处理时延、缓冲排队时延等)中分配给编码的时延为 30ms 左右。因此必须考虑在 RTP 包中打进适当的压缩语音数据以满足该时延要求。

2.5 服务器端的实现及关键技术

服务器端的工作原理是这样的:系统首先启动监听线程,目的是实时监听客户端对外线电话发起的呼叫,监听线程将接收到的客户端发送过来的外线电话号码送入 PSTN 接口卡进行外线拨号,当检测到外线电话摘机信号时,就发送数据包通知客户端用户外线已经拨通,可以开始通话了。在对客户端监听的同时,系统还会启动一个外线电话的呼叫监听线程,以实时地监听是否有外线电话拨叫网内终端用户。当外线用户拨通网关接入号码后,PSTN 接口卡会播放录音提示用户通过按键选择要呼叫的网内用户,然后通过查询数据库得到该用户对应的 IP 地址,继而向该用户发出是否接听的询问,若选择接听,就可以开始通话了。只要外线电话不挂机或客户端不停止通话,服务器就要持续的对语音数据流进行处理:所获取的来自外线电话的 PCM 语音流会自动进入 IP 资源卡的 DSP 中进行编码处理,然后进行 RTP 协议封装,再发送至客户端;同时实时接收客户端发送来的 RTP 数据包,对其进行解包后,自动送入 DSP 中解码还原成原始的 PCM 语音数据。通过 PSTN 接口卡和 IP 资源卡连通的时隙通道,传送到外线电话,这样就实现了双向的实时语音通信。其工作原理如图 4 所示。

2.5.1 PSTN 接口卡的接口处理过程

- (1) 对语音卡的初始化,获得通道数和通道类型;
- (2) 设置定时器监听外线通道是否有振铃音以及内线通道是否有摘机信号;

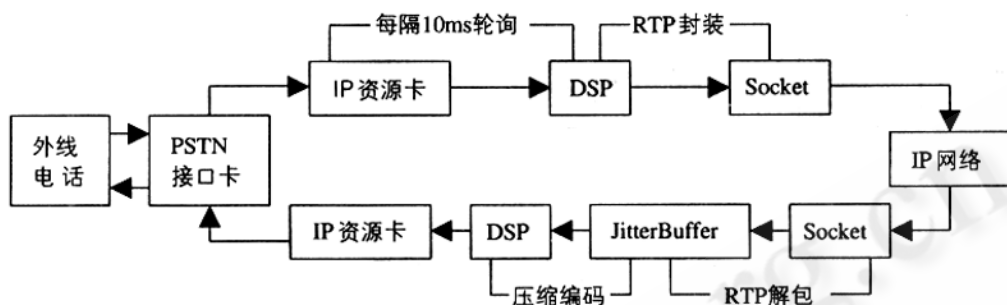


图 4 IP 电话网关工作原理图

- (3) 外线电话主叫和被叫的情况分别在两个线程内处理,两者的通话过程是基本相同的,唯一的区别是在后者时不需要对所有的通道进行检测,由特定的语音通道对呼叫进行处理;

- (4) 主线程结束时,必须释放接口卡,释放所有的内存资源。

2.5.2 IP 资源卡的接口处理过程

- (1) 对资源卡进行资源管理:包括对资源卡的初始化,获取资源卡的通道数,对通道的初始化,设置 DSP 的各项参数(语音编解码格式、回声使能、静音抑制等);

- (2) 为资源卡分配时隙,连接语音通道和 IP 通道的时隙,保证两卡的同步工作;

- (3) 采用多个线程实现接收来自客户终端的 RTP 数据包,进行解包、解码处理,获取外线电话语音流进行编码处理,继而进行 RTP 封装并发送至客户端;

- (4) 由于 RTP 数据包的到达不是匀速的,为了消除由此造成的语音抖动现象,必要的做法是在接收端设计一个平滑缓冲区,用于在播放之前先缓存一段时间的语音包以平滑突发期到达的语音包,在此要注意播放时延和包丢失率之间的平衡;

- (5) 通话结束后,要断开两卡相连的通道和时隙,关闭所有的逻辑通道,释放 IP 资源卡,释放所占用的内存资源。

2.5.3 地址映射

必须事先建立一个局域网客户端分机号及其对应 IP 地址的数据库,当有外线电话呼入时,服务器获得由按键输入的分机号,通过查询数据库获取局域网用户的

IP 地址,然后由 PSTN 接口卡模块负责对其发起呼叫。

2.6 系统测试

系统在实际的校园网环境中,分别选取不同的时间段进行性能测试,以单向时延作为主要的测试指标,

分别得到以下数据:编码时延:30ms;发送端录音时延:<30ms;数据压缩解压时延:<20ms;RTP 打包和解包时延:<20ms;数据接收端排队缓冲时延:<90ms;IP 资源卡通过 DSP 处理外线电话语音数据时延:<10ms;校园网中一般网络传输时延:<

40ms。即总的单向延迟不超过 240ms,根据欧洲电信标准化委员会(ETSI)提出的 IP 电话 QoS 标准,已经完全能够满足普通用户对于实时通讯的需求。

3 结束语

系统结合了当前 CTI 领域的先进技术,按照 H.323、TCP/IP 等协议规范实现了 PSTN 电话用户与局域网终端用户的实时语音通信。该系统具有较高的呼叫建立成功率,达到了良好的通话质量。在提高语音分组传输效率,保证语音清晰度,减小时延等方面做了有益的探索,值得在校园网或企业 Intranet 中加以推广。今后进一步的工作是考虑在开放的网络环境下需要增加安全模块的设计,同时还要继续改善服务质量以及加强网络管理的功能。

参考文献

- 1 廉正琨,IP 网络电话技术[M],人民邮电出版社,2000。
- 2 ITU-T Recommendation H.323 Packet based multimedia Communication Systems[S],1998。
- 3 姜玉萍,基于电子政务的交互式语音应答系统的开发与设计[D],重庆邮电学院,2004。
- 4 李宁溪,张峡,基于 LAN 的语音通讯软件设计[J],计算机工程与应用,2003,39(4):137-138。
- 5 Schulzrinne H, Casner S, Frederick R, Jacobson V. RTP: A Transport Protocol for Real-Time Application, RFC 1889[S]. 1996。