

【本文献信息】胡敏,刘六程,刘鹏.基于 WebRTC 的视频会议系统的设计与实现[J].电视技术 2013 37(1).

# 基于 WebRTC 的视频会议系统的设计与实现

胡 敏,刘六程,刘 鹏

(重庆邮电大学 软件技术中心,重庆 400065)

**【摘 要】**视频会议能够给人们直观、真实的交流效果,提高人们的工作效率,受到了越来越多的企业和研究机构的重视。介绍了视频会议的发展现状;分析和研究了 WebRTC 技术、SIP 协议及现有的 SIP 视频会议系统的框架和模型、Asterisk 系统结构、Zaptel、FFmpeg 等,并对其进行改进,实现了一个普通网络环境下的软件视频会议系统,该系统具有扩展性好、价格低廉、稳定可靠、带宽利用率高等特点。

**【关键词】**WebRTC; SIP; Asterisk; Zaptel; FFmpeg; 软件视频会议系统

**【中图分类号】**TN915.43

**【文献标志码】**A

## Design and Implementation of Video Conference System Based on WebRTC

HU Min, LIU Liucheng, LIU Peng

(Center of Software Technology, Chongqing University of Post and Telecommunication, Chongqing 400065, China)

**【Abstract】**Video conference which can give people an intuitive, real exchange effects and improve people's work efficiency, is subject to the attention of a growing number of companies and research institutions. The current development of video conference is introduced, and the WebRTC, SIP framework and model of SIP video conference systems, Asterisk system, Zaptel, FFmpeg are analyzed and studied. It also provides a way of implementing the software video conference system in the common network environment. The system is scalable, low cost, reliable and high bandwidth utilization.

**【Key words】**WebRTC; SIP; Asterisk; Zaptel; FFmpeg; software video conference system

随着人类经济、政治及文化活动的不断发展,人与人之间的信息交流愈来愈频繁,人们渴望得到直观、真实的交流效果。视频会议系统可以让人们远距离进行实时信息交流,开展协同工作,提供直观、真实的音视频等多种交流方式。此外利用多媒体技术的支持,视频会议系统可以帮助用户对各种信息进行处理、加工、反馈,有利于问题的解决<sup>[1-2]</sup>。视频会议能提高人们的工作效率,降低远距离会议的费用。

目前,视频会议系统主要有两种:基于 H.323 视频会议系统和基于 SIP 的视频会议系统。基于 H.323 视频会议系统目前已经比较成熟,大多数采用基于专线的 MCU 组网方式,有较好的硬件环境支持,其核心音视频处理技术为私有技术。其缺点是技术的更新不及时、价格昂贵、不能满足个性化的定制,而且其体系结构复杂,使中小企业和普通用户望而止步。

简单灵活的 SIP 协议不仅可以实现 H.323 协议的相关功能,而且扩展性强,可以由纯软件来实现会议功能,在普通的因特网(网络环境不稳定,带宽有限)上就可以实现视频会议的应用。Google 于 2011 年将 WebRTC 实时音视频技术开源,为更多研究和开发视频会议系统的人员降低了成本,节约了开发时间。

计算机与通信技术的高速发展,以及网络宽带的不断

改善,视频会议将向着更廉价、更优质、更便捷的方向发展,以满足广大用户的需求。硬件视频会议系统将逐步被软件视频会议系统所取代。因此,研究 WebRTC 和 SIP 协议等技术,对研究和开发基于 IP 网络的纯软件视频会议系统具有重要的价值和意义。

## 1 WebRTC 分析

### 1.1 WebRTC 的背景

在 2010 年,Google 以大约 6 820 万美元收购了 VoIP 软件开发商 Global IP Solutions 公司,并因此获得了该公司拥有的 WebRTC(Web Real Time Communication)技术,并在 2011 年 6 月对 WebRTC 实时通信项目进行了开源。

### 1.2 WebRTC 的结构及分析

WebRTC 是基于 Web 浏览器的实时音频和视频的通信技术。其结构如图 1 所示。该技术主要包括 3 部分:音频模块(Voice Engine)、视频模块(Video Engine)和传输模块(Transport)。

#### 1) WebRTC 的音频模块

音频部分包含采集、编解码(iLBC/iSAC)、加密、声音处理、声音输出、音量控制、音视频同步、网络传输与媒体流控制(RTP/RTCP)等技术。iLBC 编解码器是窄带编解码器,采用 8 kHz 的采样频率,支持比特率为 15.20 kbit/s

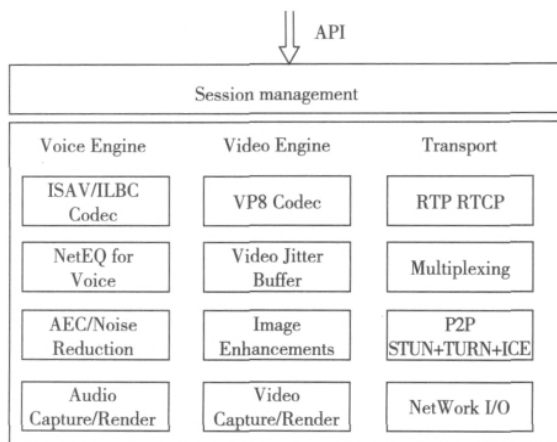


图1 WebRTC 结构

和 13.33 kbit/s 的音频流; iSAC 编解码器是宽带编解码器, 采用 16 kHz 或 32 kHz 的采样频率, 支持 12 ~ 52 kbit/s 自适应和可变的音频流范围。NetEQ 技术是动态的抖动缓存区和错误隐藏算法, 降低网络抖动和包丢失对音频效果的影响, 尽可能保持一个低延迟率语音效果。降噪和回音处理技术( AEC) [3], 能够提供高质量的语音效果。

### 2) WebRTC 的视频模块

视频部分包含采集、编解码( VP8)、加密、媒体文件、图像处理与显示、网络传输与媒体流控制等技术。VP8 视频编解码技术, 可以满足实时视频通信的低延迟特性要求。视频抖动缓存器( Video Jitter Buffer) 和图像增强技术, 减小网络抖动和包丢失对视频效果的影响, 降低视频噪声, 增强视频质量。

### 3) 传输模块

实时音视频的传输与控制, 采用 RTP[4] 与 RTCP[5-6] 技术, 稳定、可靠。此外 WebRTC 还包含了 STUN, ICE, TURN, RTP-over-TCP 的关键 NAT 和防火墙穿越技术。WebRTC 支持 SIP, XMPP/Jingle 等信令协议, 还支持跨平台: Windows, Linux, Mac, Android 等。功能的各部分可扩展, 可根据实际环境进行改进。

WebRTC 实时音视频通信技术能够满足带宽有限的普通网络环境下用户进行实时音视频交流的需求, 对研究视频会议的音视频处理等方面具有重要意义。

## 2 视频会议系统框架设计

### 2.1 SIP 协议

IETF 工作组于 1999 年 9 月推出了基于 IP 网络的会话信令协议 SIP( Session Initiation Protocol) 的第一个版本 RFC2543[7]; 2000 年前后, 3GPP 把 SIP 作为第三代移动通信系统多媒体的控制协议; 2002 年 IETF 公布了 SIP 的第

二版 RFC3261[8]; 2006 年 2 月推出 RFC4353[9], 随后 IETF 发布了多个关于 SIP 的 RFC 文档。

SIP 用于发起、修改和终止 IP 网络上的多媒体会话, 文献[8]中介绍了 SIP 协议主要有 6 类消息, 并对其进行了说明, 它们分别是 INVITE, BYE, OPTIONS, ACK, REGISTER 和 CANCEL 消息。SIP 具有以下 5 个特点: 1) SIP 消息基于文本, 词法和语法简单、可读性好; 2) 协议简单、扩展性好; 3) 定位用户、支持用户移动性; 4) SIP 网络采用 IP 网络常用的 C/S( 客户端/服务器结构); 5) 会话建立的时延较短。文献[9]中对 H.323 和 SIP 作了详细的分析和比较。

随着网络技术的快速发展, 传统固网、移动网、宽带互联网以及有线电视网等网络的融合, 同时终端朝着智能化的方向发展。在这样的网络环境中, SIP 已经被广泛地使用和支持。

### 2.2 SIP 会议系统模型

SIP 会议系统可分为紧耦合和松散耦合两种模型[10]。紧耦合会议中使用一个中心节点来实现对信令的集中控制, 而松散耦合会议中没有集中的 SIP 信令服务器。其中紧耦合模式包括集中混合模式、系统混合模式和信令集中、媒体流分布模式( 如图 2, A, B, C, D, E 为终端, 实线为 SIP 流, 虚线为媒体流)。

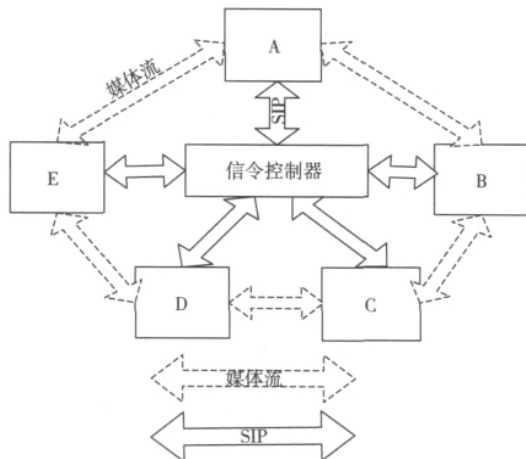


图2 信令集中、媒体流分布模式

在图 2 中, 会议的控制由一个 SIP 信令控制器来完成, 并维护所有终端间的信令连接, 完成对终端用户的控制; 媒体流则由各个与会终端分别进行处理, 媒体流的分发可以采用组播或单播的方式。考虑了媒体流分发过程中会占用大量带宽资源以及实际网络环境中部分节点或路由支持组播等因素。对于支持组播的网络, 采用组播的方式, 最大可能地节省带宽, 提高带宽的利用率, 便于会议规模的扩大。该模型媒体流的传输过程中没有混合编码

转换,处理方便、延迟小。

### 2.3 会议系统框架设计

RFC4353<sup>[11]</sup>描述了一个集中式 SIP 会议系统原始框架。其中定义了一些逻辑实体和应用场景,介绍了如何使用 SIP 协议和 SIP 的扩展来实现一个视频会议系统。其基本功能包括会议控制和管理、音频/视频处理、通信服务(白板、文件共享)等功能模块。文献[12]对该框架做了分析和改进。

本系统采用集中控制式,其视频会议框架结构如图3所示。该模型主要包括 Participant(与会者)、Focus(会议控制中心)、Mixer(媒体混合服务器)、CNS(Conference Notification Server 会议通知服务器)等功能模块。其中,Focus、CNS 和 Mixer 构成会议服务器,完成视频会议的核心功能。文献[13]中给出了该模型中部分模块的功能的介绍。

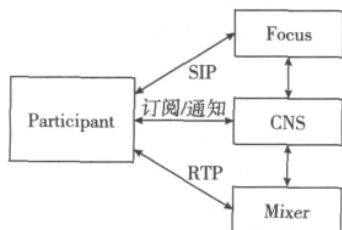


图3 视频会议系统框架

## 3 会议系统的实现

### 3.1 会议控制中心的实现

Focus 主要控制会议的建立或退出、与会者加入或退出等。Focus 的功能是在开源软件 Asterisk 的基础上设计和修改的。Asterisk 具有一个 SIP 协议栈和 RTP 协议栈,实现了一个具有基本功能的语音会议系统 Meetme,其结构如图4所示。修改后,如图5采用 TCP 传输 SIP,增强会议控制的可靠性,增加 RTCP 协议实现对媒体流的实时控制,及时反馈带宽资源的使用情况,当出现拥塞时,视频发送端根据反馈参数调节视频的帧数,保证会议的连续性和稳定性。



图4 Meetme 系统结构

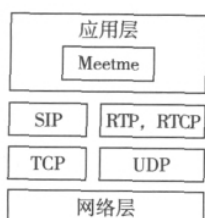


图5 会议系统结构

### 3.2 客户端的实现

与会者通过客户端完成注册后,可以使用音视频交

流、白板、文件传输与共享等通信服务功能。对于音频和视频的采集、编解码、图像和声音同步、显示、传输等分别采用 WebRTC 中的 Voice Engine、Video Engine、Transport 技术来完成,其结构如图6所示。

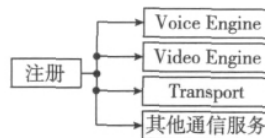


图6 客户端结构

### 3.3 媒体混合服务器的实现

针对音频的混合和分发,采用开源软件 Zaptel 来实现。Zaptel 为音频混合提供了许多方便的接口,可以完成声音的混合和分发等功能。针对视频的混合和分发,采用开源软件 FFmpeg 来实现,并对其进行了改进,首先根据客户端 IP 判断客户端是否在同一个网段内,如客户端 A(192.168.1.100)、B(192.168.1.101)、C(201.202.1.100)。A 与 B 在同一网段内则采用组播方式传输媒体流,这样在一定程度上降低了服务器负载,提高了系统性能。

### 3.4 CNS 服务器的实现

CNS 主要通过定义通知机制来通知与会者相应的会议事件和会议状态<sup>[14]</sup>。与会者通过发送 SUBSCRIBE 请求向 CNS 预定相关的会议事件和状态服务。在会议中,一旦预定的事件发生,CNS 将通过 NOTIFY 方法向与会者通告预定的相关会议事件和当前会议状态。例如,与会者可以预定会议在席信息,当有与会者加入或退出会议时,CNS 将向预定会议的在席信息的与会者发送会议成员实时变更情况。

## 4 结束语

本文通过对目前的 SIP 视频会议体系结构的分析和研究,利用 WebRTC 技术来处理视频会议的音频和视频,采用 SIP 作为视频会议的核心控制信令。并对 Asterisk、Zaptel、FFmpeg 等开源软件进行了研究和改进,设计并实现了一个普通网络环境下的视频会议系统,该系统具有价格低廉、使用方便、带宽利用率高等特点,为研究软件视频会议提供了一个参考。在下一步工作中,将对此系统进行优化,提高系统的安全性,兼容非 SIP 会话,丰富和完善会议的功能。

### 参考文献:

- [1] 仲元昌,林永,蔡增增,等. 视频会议系统关键技术及应用研究[J]. 电视技术,2010,34(6):93-96.

(下转第150页)

## 参考文献:

- [1] 仲元昌 林永 蔡增增 等. 视频会议系统关键技术及应用研究[J]. 电视技术 2010 34(6):93-96.
- [2] 杨高波 韩龙. 视频会议系统综述[J]. 电视技术 2004 28(9):60-63.
- [3] RFC2543 SIP: Session Initiation Protocol[S]. 1999.
- [4] RFC3261 SIP: Session Initiation Protocol[S]. 2002.
- [5] RFC4353 A framework for conference with the session initiation protocol (SIP) [S]. 2006.
- [6] 苏云涛. 基于 SIP 的视频会议系统的设计与实现[D]. 北京: 北京邮电大学 2008.
- [7] 郭敦望. SIP 视频会议服务器的设计与实现[D]. 北京: 北京邮电大学 2010.
- [8] 华中科技大学. 基于分布式 SIP 构架的动态负载均衡的方法: 中国, 201110171741.2[P]. 2011-09-28.
- [9] 烽火通信科技股份有限公司. 会话初始化协议消息分发系统、装置和方法: 中国 201110171741.2[P]. 2011-09-28.

## 作者简介:

何 娇(1988—) 女 彝族 硕士生 主研嵌入式系统及应用方向;  
陈盛云(1956—) 副教授 硕士生导师 主研嵌入式系统及应用方向。

责任编辑: 薛 京

收稿日期: 2012-07-27

## (上接第 132 页)

- [2] RIFE D ,BOORSTYN R. Single tone parameter estimation from discrete-time observations [J]. IEEE Trans. Information Theory ,1974 ,20( 5) : 591-598.
- [3] KAY S. A fast and accurate single frequency estimator[J]. IEEE Trans. Acoustics ,Speech and Signal Processing ,1989 37( 12) :1987-1990.
- [5] MENGALI U ,MORELLI M. Data-aided frequency estimation for burst digital transmission [J]. IEEE Trans. Communications ,1997 45( 1) : 23-25.
- [6] 彭华 李静 葛临东. 一种非判决辅助前向结构载波频差估计方法[J]. 电子学报 2001 29( 7) :984-986.
- [7] 帅涛 任前义 刘静 等. 一种适用于大频偏和低信噪比条件的频率估计器研究[J]. 宇航学报 2010 31( 12) :2471-2475.
- [8] 王旭东 刘渝 邓振森. 基于修正 Rife 算法的正弦波频率估计及 FP-GA 实现[J]. 系统工程与电子技术 2008 30( 4) :621-624.
- [9] LUISE M ,REGGINNAINI R. Carrier frequency recovery in all-digital modems for burst mode transmissions [J]. IEEE Trans. Coications , 1995 43( 2) :1169-1178.

## 作者简介:

周新力(1964—) 博士 教授 主研信息与通信;  
龚岳洲(1987—) 硕士 工程师 主研短波通信;  
孙小东(1974—) 博士 副教授 主研通信与导航;  
孟庆萍(1984—) 女 博士 工程师 主研信号处理。

责任编辑: 薛 京

收稿日期: 2012-06-13

## (上接第 143 页)

- [2] 杨高波 韩龙. 视频会议系统综述[J]. 电视技术 2004 28(9):60-63.
- [3] 潘晓东. 声学回声消除技术(AEC)在会议扩声系统的应用[J]. 电声技术 2012 36(2):2-14.
- [4] IETF. RFC3550 RTP: a transport protocol for real-time applications[S]. 2003.
- [5] IETF. RFC3611 RTP control protocol extended reports (RTCP XR) [S]. 2010.
- [6] RFC3605 Real time control protocol (RTCP) attribute in session description protocol (SDP) [S]. 2010.
- [7] RFC2543 SIP: Session initiation protocol[S]. 1999.
- [8] RFC3261 SIP: Session initiation protocol[S]. 2002.
- [9] 张智江 张云勇 刘韵洁. SIP 协议及其应用[M]. 北京: 电子工业出版社 2005.
- [10] 侯移门. TAIDEN 无纸化多媒体会议系统[J]. 电声技术 2011 35(10):8-11.
- [11] RFC4353 A framework for conference with the session initiation protocol (SIP) [S]. 2006.
- [12] 王森. SIP 多方视频会议应用服务器设计与实现[D]. 沈阳: 沈阳计算技术研究所 2008.
- [13] 苏云涛. 基于 SIP 的视频会议系统的设计与实现[D]. 北京: 北京邮电大学 2008.
- [14] RFC 3265 Session initiation protocol (SIP)-specific event notification [S]. 2002.

## 作者简介:

胡 敏(1971—) 女 硕士 副教授 硕士生导师 主研通信网体系与协议、无线通信;  
刘六程(1986—) 硕士生 主研通信网体系与协议、无线通信;  
刘 鹏(1988—) 硕士生 主研通信网体系与协议、无线通信。

责任编辑: 薛 京

收稿日期: 2012-07-04