

基于 HTML5 的视频通信云服务应用技术研究

张志明¹,柯卫²

(1.中国电信股份有限公司上海研究院 上海 200122;2.中国电信股份有限公司上海分公司 上海 200120)

摘要

HTML5 技术的快速发展大大提升了 Web 应用的用户体验,逐渐呈现出取代本地应用客户端的趋势。本文介绍了主流视频通信技术,并在 HTML5 技术与电信 IMS 现有架构的基础上,提出了一种基于 Web 技术的视频通信云服务应用解决方案,并阐述了其应用优势。

关键词 视频通信;HTML5;WebRTC;IMS

文献标识码 A doi: 10.3969/j.issn.1000-0801.2012.10.006

1 背景

HTML5 是近 10 年来 Web 标准最巨大的飞跃。和以前的版本不同,HTML5 并非仅仅用来表示 Web 内容,它的使命是将 Web 带入一个成熟的应用平台,在这个平台上,音视频、图像、动画及一些交互特性都被标准化。尽管完善 HTML5 标准还有很长的路要走,但 HTML5 的出现正在改变互联网,特别是基于 HTML5 的 Web App 将为互联网应用带来跨越式发展。

HTML5 技术出现以前,Native App(NA)占领了移动终端应用的主体市场,NA 是指使用移动平台开发语言(iOS 为 objective-c,Android 为 Java)和移动平台 SDK 来开发的软件包,可以充分发挥系统特性(调用系统服务、内存管理等),操控设备硬件(如照相机、蓝牙、传感器),提供最佳的用户体验。然而各平台的开发语言、SDK 和开发环境不同,导致学习成本高、多平台的开发移植比较复杂、开发周期长,多设备的调试和适配较麻烦,基于 HTML5 技术的 Web App 因此而出现并得到飞速发展。

如图 1 所示,Web App 是指使用 Web 标准技术(HTML、CSS、JavaScript)开发的针对移动设备优化的网站,其完全使用 Web 标准,学习和开发成本低,可兼容多种平台(实际是兼容浏览器),跨平台和设备的适配成本低;开发周期快,无需安装也可使用。

随着 HTML5 标准的不断完善和发展,基于其关键特性,目前已经出现了实现视频通信类 Web App 的标准,即 WebRTC(Web real time communication),使视频通信在移动设备中得到广泛应用成为可能,并构建了一种全新的 IP 视频通信体系架构。此外,基于 WebRTC 的 Web App 功能可以与 IMS(IP multimedia subsystem)相结合,只要尝试 IP 视频电话到核心网的连通,便可为用户提供极为方便的实时通信功能。

2 现有视频通信技术

2.1 IP 视频通信

IP 视频通信主要采用两种技术:一种是采用压缩编码及基于 TCP/IP 的分组交换技术;另一种是利用具有统计



图1 Web App 介绍

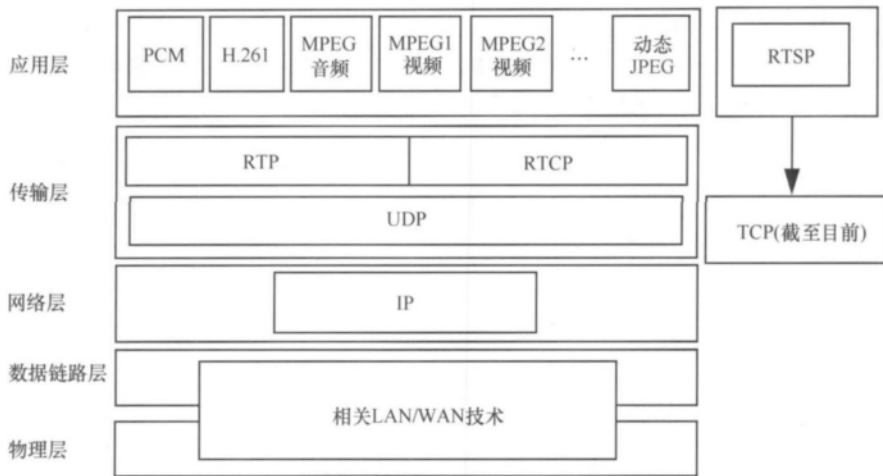


图2 IP 视频通信技术的主要协议架构

复用功能的IP数据网。以基于TCP/IP的分组交换技术为例,IP视频信息以数据分组的形式在数据网中传输,首先由信源编码器对连续视频信号进行分割压缩、采用,在传送前,将压缩后得到的数据封装到IP数据分组中,实现在IP交换网中的传输,网络上实际传送的码流是经过封装后的码流。封装形式为在数据分组前加上IP分组头、UDP分组头和RTP分组头,封装的效率取决于一个RTP分组中打多少数量的数据分组。RTP分组头所打的视频分组越多,封装效率就越高,单位流量也就越小,但会加大全网时延。图2为IP视频通信技术的主要协议架构。

目前,IP视频通信技术基本采用客户端(client)方式及Web页面访问方式。

(1)IP视频通信客户端方式

目前该方式以其视频编解码质量好、通信稳定的优势一直占据视频通信类的主流地位。IP视频通信客户端与

服务器的互通如图3所示,客户端是指与服务器相对应,为客户提供本地服务的程序。该方式的通信基本采用私有协议,由于客户端自身受到跨操作系统、跨终端的限制,用户体验不太理想,比如苹果的Facetime仅限于苹果设备之间的视频通信,其他客户端存在无法跨屏视频、体验效果

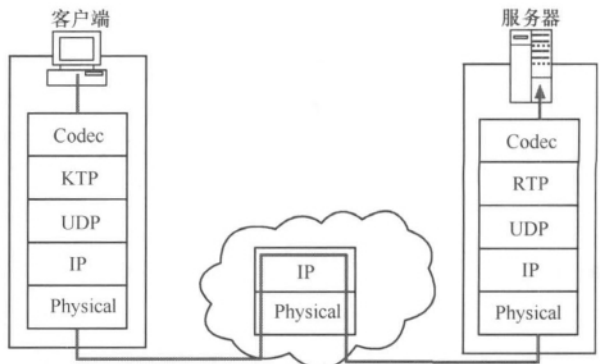


图3 IP 视频通信客户端与服务器的互通

较差等问题。

(2) Web 页面访问方式

目前该方式主要采用 flash 插件、Ajax 异步通信交互等方式完成,需要对现有浏览器安装插件,可在页面上完成视频通信。但由于受到插件能力的限制,插件需要考虑不同浏览器、版本,甚至 OS,且视频编解码能力较差,没有开放的标准,使得 Web 开发者开发和部署与平台无关的 Web 视频通信应用较为困难。

新一代视频通信 Web App 可以比较好地解决以上两种视频通信方式出现的问题,基于 WebRTC 标准的视频通信应用,通过 video 标签支持无插件的高清视频格式,通过 Web Socket 特性实现全双工通信,比之前的异步通信效率更高。基于 WebRTC 的视频通信应用具有安全性好、互通方便、视频质量好等优点,所有浏览器可以原生支持 RTC 已经成为大势所趋,可以成功避免客户端受到跨平台、跨设备的困扰,以实现 Web 高速创新的步伐。

2.2 IMS 视频通信

随着用户需求量的不断增大,单纯的 IP 视频通信只能满足 IP 电话与固话及运营商核心网的互通及访问需求,要解决 IP 电话与核心网络的互通问题,必须借助 IMS。IMS 是一种全新的多媒体业务形式,能够满足终端客户更新颖、更多样化的多媒体业务需求,是 RCS (rich communication system) 解决方案中的重要一环。

IMS 协议架构如图 4 所示。最底层为承载层,用于提供 IMS SIP 会话的接入和传输功能,承载网必须是基于分组交换的,主要的承载层设备有 SGSN (业务支撑节点)、GGSN 以及 MGW (媒体网关),MGW 是负责媒体流在 IMS 域和 CS 域互通的功能实体。基于 IP 技术,所有 IMS 用户

信令可以很好地传送到信令控制层 (即中间层),所有 IP 多媒体业务的信令控制都在这一层完成,最后通过 SIP 信令到达应用层。

在 IETF、3GPP2 和 OMA 中,作为 IMS 客户端的基本协议簇,SIP 用于两个或者多个 IP 节点间会话的建立、维护和拆除;SDP (会话描述协议)、RTP/RTCP (实时传送协议/实时控制协议)都是实现 IMS 客户端最重要的基础协议。其中,SDP 作为一种应用层协议,用来描述媒体会话能力、媒体格式、媒体流地址和端口等信息;RTP 是用于端到端传递实时数据的协议,RTCP 用于实时监控数据服务质量。

IMS 作为基于 SIP 的通用平台,可以同时支持固定和移动的多种接入方式,很好地实现固定网与移动网的融合,如果要将 IMS 与 IP 视频通信应用进行互通,则在服务器端也需采用 SIP。

3 WebRTC——基于 HTML5 的视频通信技术

3.1 WebRTC 现状

WebRTC 简单来说就是使用 Web 浏览器打语音电话、视频电话或者开视频会议。已经有部分浏览器实现了上述功能,但需要安装额外的插件 (如在 Chrome 上安装 Google Voice 插件、Firefox 上安装 Skype 插件),而且不同的浏览器之间不能互通。为了将通信功能内置在浏览器中,同时实现各个浏览器之间、浏览器与现有通信系统之间的互通,主流浏览器厂商 (Google、Mozilla 等)联合主要的通信设备厂商 (Ericsson、Cisco 等)在 IETF 和 W3C 中发起并成立了相应工作组进行标准化工作,即对 WebRTC 标准进行制定。

基于 WebRTC 的 Web App 相对于传统客户端模式的视频通信应用,有着免安装免更新、跨平台跨终端的技术优势,且运营商可为没有安装客户端的被叫用户提供 Web 页面方式,使其可以与安装了客户端的用户进行通信,扩大业务范围,还可与 IMS 互通,实现跨设备的固移融合。WebRTC 标准中主要用到了 HTML5 中的两个特性:Web Socket 和 video 标签,Web Socket 主要负责全双工视频通信,而 video 标签主要负责视频通信中的视频编解码工作。

3.2 WebRTC 特性

3.2.1 Web Socket

在 HTML5 技术标准中,Web Socket 是一个非常重要

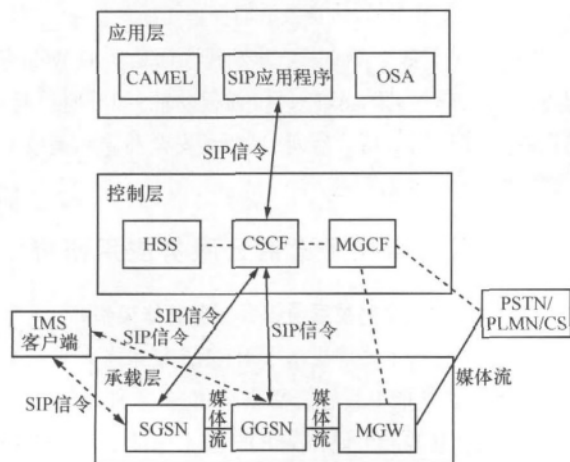


图 4 IMS 协议架构

的新特性,其定义了一个全双工通信信道,即浏览器可以利用 Web Socket 协议和主机进行双向通信,比 XML HTTP Request 更加强大、高效并能有效减少流量。

正常情况下,浏览器访问 Web 页面时会向服务器发送一个 HTTP 请求,Web 服务器识别请求后返回响应。大多数情况下,一些实时性要求比较高的数据(如视频流、股票价格、余票查询等)的实时 Web 应用是围绕轮询和服务器推送技术展开的,这些方式提供实时数据时都会涉及 HTTP 请求和响应报头,包含大量额外的不必要的报头数据。建立 Web Socket 通信时,客户端和服务端在初始握手时就将 HTTP 升级为 Web Socket 协议,一旦连接成功,就可以以全双工模式在客户端和服务端之间传送 Web Socket 消息,在同一时间、任何方向,都可发送基于文本的消息,所以 Web Socket 的高效和节流成为开发者们津津乐道的新特性之一。

浏览器对 Web Socket 的支持程度见表 1。可以看出,目前主流浏览器均可以很好地支持该特性。

表 1 主流浏览器对 Web Socket 的支持程度

浏览器	支持情况
Internet Explorer	从版本 10 开始支持
Chrome	从版本 4 开始支持
Firefox	从版本 4 开始支持
Opera	从版本 10 开始支持
Safari	从版本 5 开始支持

3.2.2 video 标签

现今大多数音视频是通过插件(如 flash)来显示的,然而并非所有浏览器都拥有同样的插件。HTML5 规定了一种通过 video/audio 元素来包含视频/音频的标准方法。video/audio 元素有两个关键概念:Container 和 Codec。

- Container: 音频和视频文件实际上只是一个 Container 文件,如视频文件,可以包含音频轨道、视频轨道和其他元素。目前,主流视频 Container 支持 audio video interleave、flash video、MPEG4、Matroska、Ogg 等视频格式。
- Codec: 视频和音频的编解码是一种算法,但有些编解码算法是有专利的,加上其他因素的限制,使得 HTML5 不能指定编解码算法,只能由开发人员根据不同浏览器的情况,针对不同浏览器的环境对媒体文件进行重编码。主流的编解码器有 AAC、MPEG3、Ogg vorbits、H.264、VP 8、Ogg theora。

当前,video 元素支持 3 种视频格式: 带有 theora 视频编码和 vorbis 音频编码的 Ogg 文件; 带有 H.264 视频编码和 AAC 音频编码的 MPEG4 文件; 带有 VP8 视频编码和 vorbis 音频编码的 WebM 文件。而 audio 元素支持 3 种音频格式:Ogg vorbits、MP3、WAV。

目前,很多浏览器已经实现了对 video 和 audio 元素的支持,见表 2。

表 2 浏览器对 video 标签的支持情况

浏览器	支持情况
Internet Explorer	IE9 支持
Chrome	从 3.0 版本开始支持
Firefox	从 3.5 版本开始支持
Opera	从 10.5 版本开始支持
Safari	从 3.2 版本开始支持

video 标签能成为最受欢迎的 HTML5 新特性之一,主要是因为其可以完全抛弃传统基于私有协议的 flash 插件或其他插件,而使得视频通信协议标准完全公开化,可实现跨平台、跨终端。

3.3 WebRTC 技术

基于 WebRTC 标准的视频通信应用主要包括数据传输和连接管理两个方面。

- 数据传输: 包括 TCP、UDP 及安全连接的技术手段、RTP 及数据加密以及拥塞管理、带宽估计、数据发送时机、NAT/FW 穿越等, 传输的数据格式包括 Codec 规范以及数据或者文档共享的格式。
- 连接管理: 包括一些重要的连接协议(SIP、Jingle/XMPP)、呼叫信令等建立连接、协商能力、改变拆除连接等,即通常所说的呼叫信令;还包括会议成员控制、屏幕布局、语音激活的屏幕切换等。

因此,对于基于 WebRTC 开发视频通信类的 Web 应用,在浏览器的 RTC API 设计、通信功能(如呼叫控制、NAT 穿越)在应用与浏览器间的分布、安全和隐私是技术上要解决的几个核心问题。

4 电信运营商视频通信云服务应用研究

4.1 基于 WebRTC 的视频通信云服务总体架构

WebRTC 技术主要解决基于 HTML5 技术实现可跨平台、跨终端的高性能视频通信 Web 应用,是一种基于 IP 网络的云应用。从本质上说,WebRTC 技术标准是对 HTML5 关键性能的组合封装,比如其使用了 HTML5 的 Web

Socket、video、audio、DeviceAPI 等特性。在浏览器侧,该视频通信应用需要设计 RTC 的 API,而在服务器端,则需要实现一些关键的云服务能力,除了基本的 WebRTC 功能外,与 IMS 互通的网关、视频转码服务等都是需要考虑的因素。基于 HTML5 的云服务应用总体架构如图 5 所示。

4.2 WebRTC 服务器的主要功能

基于 IP 视频通信的基本 WebRTC 功能协议架构如图 6 所示。该架构说明了用户使用不同的浏览器实现视频通信的原理,是 WebRTC 视频通信应用的标准流程。在客户侧,浏览器基于 W3C 的 WebRTC 功能通过 Web Socket/ HTTP 实现与服务器端的互联互通,而服务器端通过 SIP/XMPP

相连接,最终两个浏览器可基于 RTP/RTCP/UDP/IP 等协议实现用户 A 和用户 B 端到端的 IP 视频流通信。

4.3 IMS 转换网关的主要功能

如何解决 IP 视频通信与 IMS 互通的问题是电信运营商关注的重点之一,其中一个方案是引入媒体网关,如图 7 所示。

实现基本的视频通信需要服务器端具备 WebRTC 功能,而与 IMS 互通需要增加一个互通网关,所以可以将互通网关合并至云服务器,作为云端服务器的功能之一,如图 8 所示。云服务平台兼具了 WebRTC 和与 IMS 互通的能力。用户使用终端 1(手机或 PC 等可接入互联网设备),浏览器支持 WebRTC 引擎,可通过 Web Socket 协议与运营商

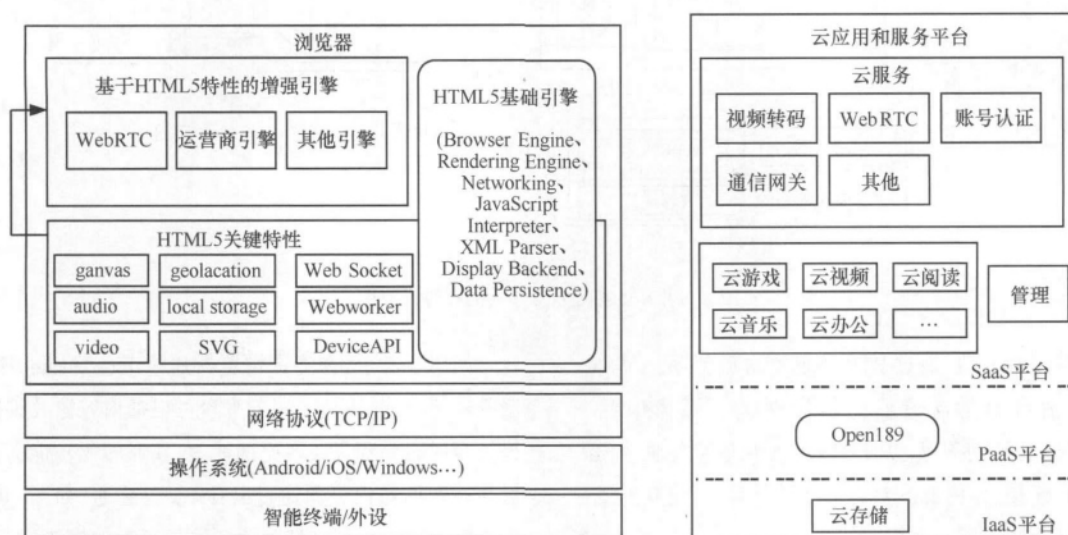


图 5 基于 HTML5 的云服务应用总体架构

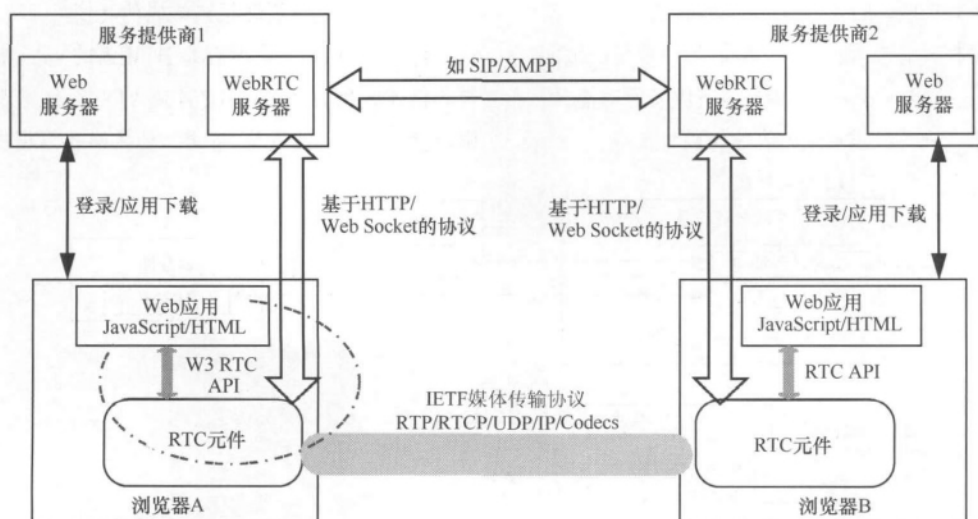


图 6 基于 IP 视频通信的 WebRTC 功能协议架构

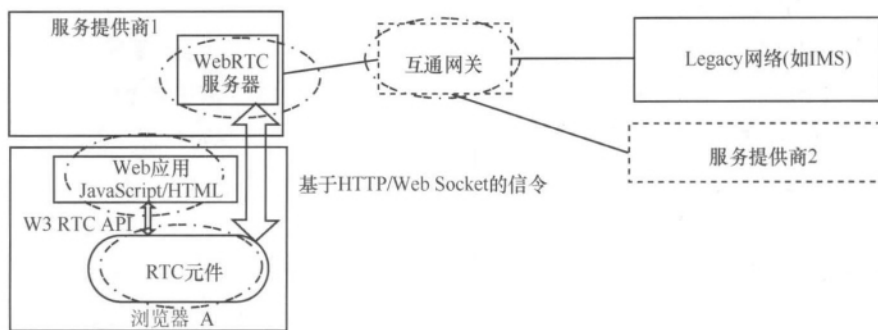


图7 引入媒体网关的WebRTC协议架构

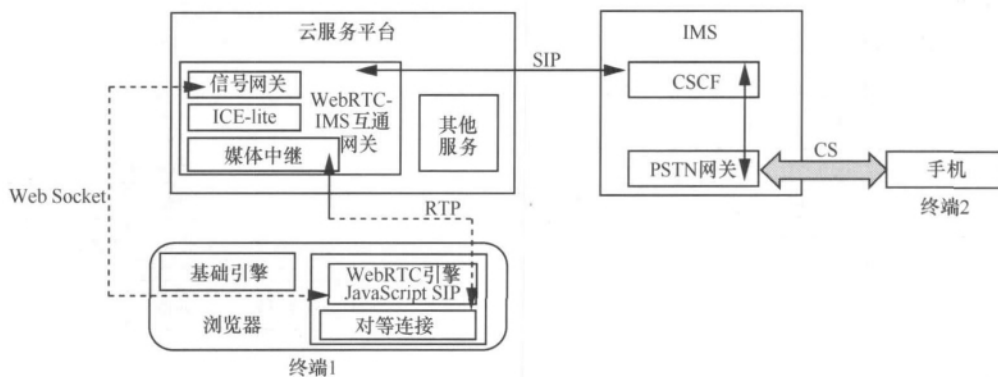


图8 引入互通网关的WebRTC协议架构

云服务平台进行互通,通过 RTP 实现实时流传输,云服务平台侧包括与浏览器通信的信号网关、媒体中继、WebRTC-IMS 互通网关等。WebRTC-IMS 互通网关通过 SIP 与 IMS 进行连接,再由 IMS 对终端 2 进行寻址,建立连接,便可完成浏览器到核心网终端的视频通信。

4.4 视频云转码的主要功能

根据 HTML5 中的 video 标签特性,WebRTC 标准中浏览器支持的视频格式有限,而且不尽相同,所以要该问题,云服务平台需要增加另一个功能——视频转码功能,即通过云转码引擎实现对各类视频内容的转码与支持。

如图 9 所示,在终端浏览器通过 Web Socket/HTTP 与服务器建立连接后,云服务平台可根据内容源及编码格式提供实时转码和离线转码的能力。实时转码功能需要服务器与视频网站内容源通过 RSTP 实时互通,如图 10 所示,在线将视频编码格式转换为浏览器支持的格式;而离线转码功能可根据云服务平台的云存储功能预留视频内容源,在离线情况下进行视频格式编码的转换。

转码原则如下:将所有非 HTML5 的 video 标签支持的视频格式转化为 theora、H.264 或 VP8 格式,音频编码格式则转换为 vorbis 或 AAC 格式,具体格式要根据云转码能

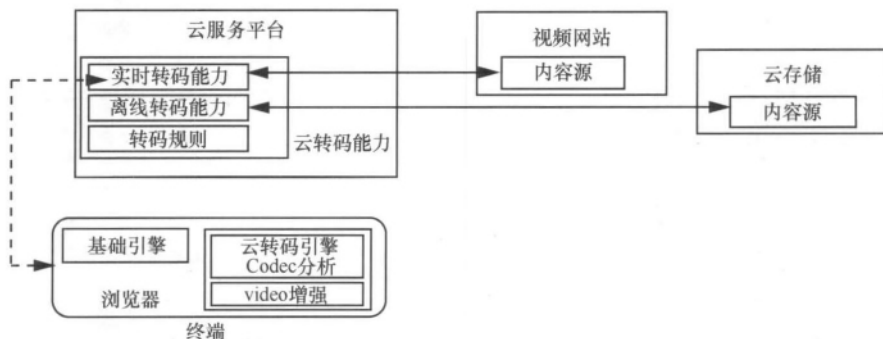


图9 WebRTC 的转码云服务平台

力对浏览器属性进行判断以后确定。

云服务平台具备了云转码引擎后,可看作对 HTML5 的 video 标签特性的扩展支持,而云服务平台转码能力的关键在于做好平台与视频在线和离线内容源的对接。

5 结束语

本文通过介绍主流的 IP 视频通信技术,结合 HTML5 的新特性及技术优势,详细分析了 WebRTC 标准的现状及价值,并在结合 WebRTC 技术标准与 IMS 现有架构的基础上,结合电信运营商的实际情况,提出了基于 WebRTC 视频通信的云服务应用解决方案。该方案整合了 HTML5 的技术优势及 IMS 的通信效率,并对视频格式支持进行了优化。相信在不久的将来,基于 WebRTC 的视频通信应用作为 RCS 的一个重要方面,将以更轻松简便的方式为终端用户提供更加丰富的通信服务,实现在不同国家、不同运营商网络间的视频通信自由沟通。

参考文献

- 1 Keith J. HTML5 for web designers. Jeffrey Zledman, New York, 2010
- 2 Network Working Group Internet-Draft. RTCWeb Network

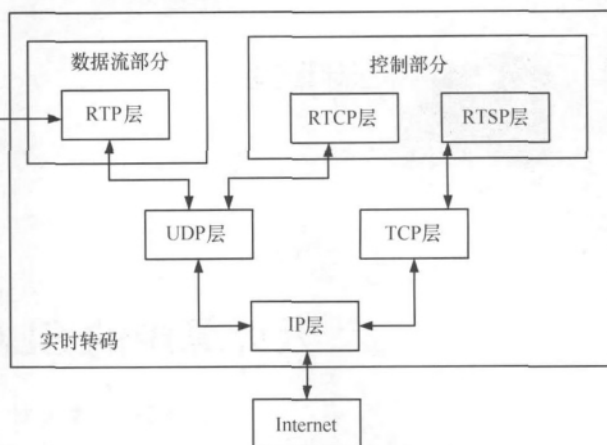


图 10 RSTP 音视频转码

Address Translation draft-cbran-rtcWeb-nat-01, 2012

- 3 RTCWeb Internet-Draft. Security Considerations for RTCWeb draft-ietf-rtcWeb-security-00, 2012
- 4 付航. IMS 体系研究与网络架构演进. 移动通信, 2007(Z1)
- 5 陈奇峰, 刘锦高. 关于 IMS 中基于 SIP 的应用分析. 中国科技信息, 2007(11)
- 6 林鸿, 陈自力, 王松. IMS 客户端技术标准及软件特性分析. 电信科学, 2007(1)
- 7 刘晓梅, 王彦永, 魏立峰等. 基于 IE 浏览器 H.264 视频播放插件的实现. 计算机系统应用, 2009(3)
- 8 彭涛, 杨炼. 基于移动浏览器的 HTML5 核心技术的研究及其应用. 广东通信技术, 2012(4)
- 9 刘华星, 杨庚. 下一代 Web 开发标准研究. 计算机技术与发展, 2011, 21(8)

Research on Techniques of Cloud Service Applications in HTML5-Based Video Communication

Zhang Zhiming¹, Ke Wei²

(1. Shanghai Research Institute of China Telecom Co., Ltd., Shanghai 200122, China;

2. Shanghai Branch of China Telecom Co., Ltd., Shanghai 200120, China)

Abstract The development of HTML5 technique increases user experience of web apps greatly, and tends to become the substitutions of native apps. This paper introduces main techniques of video communication, and proposes a solution for web techniques-based cloud service application in video communication by combining HTML5 and the existing IMS architecture. Then advantages of the solution are illustrated.

Key words video communication, HTML5, WebRTC, IMS

(收稿日期:2012-09-10)