



WebRTC 技术研究及其应用

付 斌¹, 杨 鑫¹, 王 松², 林 鸿²

(1. 中国电信股份有限公司北京研究院 北京 100035; 2. 法国电信北京研发中心 北京 100190)

摘 要: WebRTC 技术在浏览器上直接实现语音、视频和文件共享等业务, 极大地提高了使用的便利性。由于其开源和开放, 使得 VoIP 应用的开发更为简单和快速。在对 WebRTC 进行介绍的基础上, 分析了其发展趋势以及对运营商通信业务的影响, 与 IMS 进行了对比, 然后对 WebRTC 与 IMS 结合的 3 种方式进行了讨论。最后, 对目前 WebRTC 的主要研究方向进行了介绍, 展望了未来的发展趋势。

关键词: WebRTC; RTCWeb; IMS; VoIP

doi: 10.3969/j.issn.1000-0801.2013.09.021

Research on WebRTC Technology and Application

Fu Bin¹, Yang Xin¹, Wang Song², Lin Hong²

(1. Beijing Research Institute of China Telecom Co., Ltd., Beijing 100035, China;

2. France Telecom R&D Beijing Centre, Beijing 100190, China)

Abstract: WebRTC technology enables audio, video communication and file sharing services over web browsers without installing add-ons, which improves the convenience of such applications greatly. Its openness and open source policy will simplify and facilitate the development of applications. WebRTC technology was introduced, and then the trend and influence on operator's communication services were analyzed. A comparison and three ways of combination with IMS were depicted as well. At the end, a summary of research topics and overview of future development were given.

Key words: Web real-time communication, RTCWeb, IP multimedia subsystem, voice over IP

1 引言

WebRTC(Web real-time communication)是近年来兴起的一项 OTT(over the top)VoIP(voice over IP)技术, 目标是通过简单的 JavaScript 调用即可在浏览器上实现点对点语音、视频和文件共享等业务。因为其开放、开源且具有极大的使用便利性, 从推出以来就受到了业界特别是互联网和通信行业的广泛关注。在互联网应用爆发式增长的大环境下, 这种实时通信功能已经成为各种互联网应用的基础要素, 也是 WebRTC 受到如此关注的主要原因。对于电信业而言, WebRTC 提供的语音、视频通信服务一定程度上会

与传统的电信业务产生竞争关系, 对于同样基于 VoIP 技术的 IMS(IP multimedia subsystem)会产生一定的替代。另一方面, 人们也在探索把 WebRTC 技术应用于 IMS 中, 如以 Web 方式接入 IMS 服务, 甚至整体上作为运营商 VoIP 业务的一种选项。

WebRTC 包括语音引擎(voice engine)、视频引擎(video engine)、传送(transport)功能和会话管理(session management)功能以及对本地音频、视频、网络的操作功能和向浏览器提供的 API(application programming interface)。目前公开的核心技术包括音视频编解码技术, 如音频的 NetEQ 算法、回声抑制、噪声消除, 视频的抖动处理、图像

增强等。绝大部分技术来源于 Google 向业界开源的一些项目(参阅 <http://www.webrtc.org>)。媒体的传送部分重用 ICE(interactive session establishment)技术,支持对 4 种类型 NAT(network address translation)或防火墙的穿透,实现点对点、点对多点等多种方式媒体通路的建立和传送。

在 Google 等公司的推动下,WebRTC 的标准化工作取得了快速发展。2011 年 5 月,IETF 成立 RTCWeb 工作组负责需求、总体方案和协议的制定,并为 W3C(World Wide Web Consortium)提供输入。与此同时,W3C 成立 WebRTC 工作组负责客户端 API 的定义,几个主要的浏览器(Firefox、Chrome、Opera)也都紧跟着标准化进展的步伐推出了支持 WebRTC 的版本。另外,3GPP 也启动了 WebRTC access to IMS 的工作项目,并在 2013 年 7 月得到了全球主要运营商和设备商的一致支持和积极参与。研究的内容主要是实现 WebRTC 客户端与 IMS 客户端的互操作性,因此需要对 IMS 进行增强。

2 WebRTC 发展趋势及其对通信业务的影响

2.1 WebRTC 技术发展趋势

近年来电信市场通信业务增长缓慢,来自电信市场研究公司 TeleGeography 的一份报告显示,2012 年国际电话通信量仅增长 5%,达到 4 900 亿分钟。而与此同时,互联网通信类应用却保持着爆发式的增长速度。以 Skype 为例,2012 年语音和视频流量增长 44%,达到 1 670 亿分钟,相当于全部电信运营商国际电话通信总量的 1/3,这还不包括其他流行的通信类应用(如 GTalk、微信、Viber、Line、Nimbuzz、KakaoTalk 等),对比如图 1 所示。

WebRTC 技术的出现,必将进一步加剧这一趋势,因为用户在使用这些应用时将不再需要下载安装客户端(或

插件),直接打开支持 WebRTC 的浏览器即可,这将为用户带来极大的便利性。如果这些通信类应用都提供 WebRTC 方式的应用支持,将会大大提高服务的可用性——只要有浏览器的地方,用户就可以使用它们的服务,而且会带来新的使用方式和用户体验。

WebRTC 可以作为一种全新的 VoIP 应用形态,催生其他应用向 VoIP 领域转移和拓展。例如,各种电商、社交类应用,通过对应用平台的增强实现 WebRTC 支持,这样用户之间可以直接在 Web 页面上发起 VoIP 呼叫,把原有的服务能力自然延伸到 VoIP 领域,提供更为丰富和全面的用户体验。

从产业发展的情况看,目前已经有众多著名互联网厂商投入 WebRTC 技术研究和规范制定领域,更多的厂商潜心研究 WebRTC 已经公开的技术和代码,并在此之上开发各式各样的应用,可以预见,将来基于 Web 的通信会广泛出现在各类互联网应用之中,而这必将对运营商的通信业务带来深远的影响。

从产业发展看,以 Skype 为代表的 OTT VoIP 应用对运营商通信业务带来巨大的业务分流和替代已经是大势所趋,而 WebRTC 技术所带来的便利性和新的用户体验也必将极大地促进这一趋势的发展,从而侵蚀运营商的语音业务。与此同时,WebRTC 的兴起也给运营商带来更多的机遇,具体如下:

- WebRTC 带动互联网语音和视频应用的发展,为运营商带来巨大的数据流量,并拉动接入需求的增长;
- WebRTC 用户成为运营商的潜在客户,运营商可以通过与 WebRTC 的互通获得更多的呼入呼出话务量,从而扩大营收规模;
- 借助 WebRTC 技术,可以开发新的业务形式和解决方案,如使用户从任意支持浏览器的设备上接入运营商网络。

2.2 与 IMS 的关系

WebRTC 自推出以来就受到了广泛关注,特别是与 IMS 关系的话题。事实上,两者在服务能力、主要技术等方面有很多共同点,主要体现在如下几个方面:

- 提供基于 IP 的基本语音、视频服务;
- 在控制面有很多技术相通,如呼叫状态处理、终端保活、呼叫连续性等,媒体协商也基本一致(WebRTC 还需要考虑呼叫状态保存);

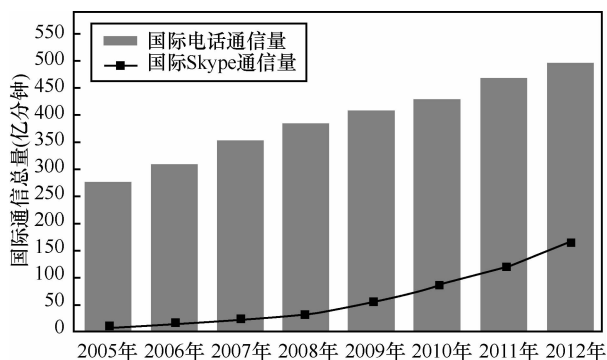


图 1 国际电话通话分钟数与 Skype 的对比(来源:TeleGeography)



- 语音、视频编解码技术以及回声消除、纠错、码率适应算法等方面,可以共用(出于规范定义的原因,IETF RTCWeb 工作组对必须支持的语音、视频编解码和 3GPP 规范对 IMS 的要求存在很大不同);
- NAT 和防火墙穿越方面,采用相同的 ICE 技术(IMS 还有 ALG(application layer gateway)方式的解决方案)。

从产品定位看,WebRTC 技术定位于为双方或者多方基于 Web 浏览器直接进行实时交互式沟通,而 IMS 则为电信用户提供多媒体通信(multimedia telephony)服务,需要遵循原有通信业务和各种管制法规的要求,而 WebRTC 则没有这些约束,这也从本质上决定了两者必然存在的差异,两者对比见表 1。

除此之外,两者还有很多细节上的区别,如语音编码方面,WebRTC 只有 G.711 和 Opus 是必需的,而 IMS 则要求支持几乎所有传统电话网的编码;媒体路由方面,IMS 路由需要经过核心网,而 WebRTC 则可以是完全点对点的。另外,IMS 作为基础电信服务面向普遍的有通信需求的用户,接入方式既包括智能终端,也包括各种类型的终端,如传统电话、手机、IAD(integrated access device)、SIP 话机等,如果 WebRTC 期望把服务延伸到电信用户领域,也需要考虑与 IMS 的互联互通。同样的,IMS 也可以引入 WebRTC 的技术成果,以提高服务的便利性和扩展用户规模。

2.3 WebRTC 技术典型应用

需要指出的是,虽然 WebRTC 的目标是提供基于浏览器的点对点语音、视频等业务,但其使用方式并不能完全局限于浏览器,其应用方式主要包括如下几种。

- 基于浏览器的应用,浏览器必须支持 WebRTC,服务器提供呼叫控制。浏览器与服务器之间以 HTTP 或 Web Socket 交互。
- 在本地应用中使用 WebRTC,即应用客户端集成 WebRTC 能力,与网络侧的服务器以 XMPP、SIP 等协议交互。有很多原有的 VoIP 类应用都是通过这种方式改造使用 WebRTC 的一些技术成果。

3 WebRTC 与 IMS 的结合

经过前面的对比和分析,WebRTC 与 IMS 有着很多区别,但两者并不是对立的,而是有着各自的应用场合,能够相互补充和结合。从 IMS 角度看,与 WebRTC 的结合主要有 3 种方式:利用 WebRTC 接入、与 WebRTC 互通和与 WebRTC 融合。

3.1 WebRTC 接入 IMS

传统的 IMS 客户端包括硬终端和软终端,WebRTC 的出现使得单纯基于浏览器的 IMS 客户端成为可能。WebRTC 浏览器提供了对本地麦克风和摄像头的访问功能,封装了媒体面所需要的能力,包括语音、视频编解码和 NAT 穿越等。

以 WebRTC 方式接入 IMS 的系统架构如图 2 所示,需要部署 WebRTC 接入网关连接到 IMS 网络。WebRTC 接入网关提供 IMS 应用逻辑,包括用户鉴权、呼叫控制等。媒体面需要有媒体服务器提供 NAT 穿越、转码等功能。IMS 用户通过支持 WebRTC 的浏览器连接到接入网关并以 HTTP/Web Socket 交互,首先实现用户鉴权。鉴权通过后,在用户 UI 上显示拨号盘或者通讯录列表供用户发起和接

表 1 IMS 与 WebRTC 的对比

IMS	WebRTC
设计严密、完善,高度可控,还需要考虑跟原有通信网络的演进、互操作以及用户迁移等问题	无统一的顶层设计,高度开放和灵活,没有先决限制条件
作为基础电信业务的解决方案,信令、媒体面完全受控,需要考虑互联互通、监管等要求,需满足计费、结算等固有商业模式要求	作为互联网业务,完全按需设计,不需要满足任何特别的管制要求;商业模式灵活多样,无计费、结算等要求
端到端全面的方案,从系统到接口、协议、流程、标识都严格标准化,体系庞大复杂	完全开放,仅对浏览器 API、会话协商、编解码等进行标准化,其他留给应用开发者决定
主要面向语音、视频类业务	提供数据通道,用于文件共享等远程协作类服务
用户鉴权和可识别,必须是运营商用户,且通常要求具有 tel 或 SIP URI 标识	完全开放,用户鉴权不是必选项,甚至可以是物,如监控类应用
与通信网络紧密耦合,能力开放有限,能力扩展主要通过应用服务器	灵活扩展,便于集成互联网上的各种能力,如语音识别/翻译、语音/图像处理、定位等,也能作为能力轻松地被其他应用集成
有 QoS 保证,可以提供 3GPP 定义的高清语音业务	随着网络条件的改善,在多数情况下 QoS 已经足够好并持续提升,能提供近乎无限提高的高清语音、视频服务
系统复杂度高,对新的需求改动难度大,难以及时响应	改动灵活,便于实现快速、多样化的应用需求

收呼叫。呼叫建立的过程中,通过调用 JavaScript API 建立起媒体面连接,经由媒体服务器转接与 IMS 网络建立媒体流。

这一方案应用于运营商为 IMS 用户提供 Web 方式登录使用 IMS 业务的场景,即用户首先是运营商的用户。带来的好处是显而易见的,用户不需要安装或购买专用的客户端即可使用 IMS 服务。由于是通过浏览器方式访问 WebRTC 接入网关,免去了终端软件升级带来的烦恼。鉴权时所使用的身份信息与通过其他方式登录 IMS 一致。使用业务时如何处理与其他 IMS 终端的关系,可以由运营商的策略来控制。

3.2 IMS 与 WebRTC 互通

IMS 用户和 WebRTC 用户之间相互进行语音、视频通话的场景,即两个用户分别属于 IMS 网络和 WebRTC 网络,由一端发起与另一端的通话。WebRTC 呼叫服务平台与 IMS 之间通过 IMS 与 WebRTC 的互通网关连接,如图 3 所示。

图 3 中 WebRTC 用户通过浏览器接入 WebRTC 服务平台。IMS 用户接入 IMS 网络,没有在图 3 中画出,两个用户分别在各自的服务平台鉴权。进行互通时,WebRTC 平台和 IMS 都需要进行增强,增加呼叫路由规则。当 WebRTC 用户发起呼叫时,WebRTC 平台需要判别该用户为 IMS 用户,路由呼叫到 WebRTC 互通网关,WebRTC 网关向该 IMS 用户标识发起呼叫,然后按照 IMS 的规则进行呼叫处理,反之亦然。需要说明的是,此时 WebRTC 互通网

关与 IMS 网络之间通过 NNI (network to network interface) 交互,交互的网元可以是 I-CSCF 或 SBC (session border controller)。为了 NAT 穿越的需要,WebRTC 用户的媒体流可能需要经过中间节点(如 TURN 服务器)中转,具体与网络部署有关。

3.3 IMS 与 WebRTC 融合

在这种场景下,用户同时归属于 WebRTC 用户和 IMS 用户,用户既可以在 WebRTC 平台发起和接收呼叫,也可以在 IMS 网络发起和接收呼叫,甚至可以在呼叫过程中在两个域之间切换。这种方式对用户而言有极大的自主性,可以根据偏好在不同的时候采用不同的服务。比如,网络条件好时采用 WebRTC 服务,条件变差时切换到 IMS 服务,或者设置在国际漫游时自动切换到 WebRTC 服务,以节省资费。对于这种场景的研究目前比较少,实现起来也比较复杂,需要对 WebRTC 和 IMS 网络都进行更多的增强并引入类似 IMS VCC 的能力,在这里不做详细介绍。

4 主要研究方向

WebRTC 作为一项发展中的技术,仍然有很多问题处于研究之中或有待研究,主要包括以下几个方面。

- 呼叫建立过程中的状态保存问题,包括呼叫的会话状态、媒体描述信息和 ICE 协商状态,浏览器刷新可能导致这些状态信息丢失,目前考虑的方式是封装成对象保存到服务器端,在完成信息保存的同时,又避免了 API 的复杂化。

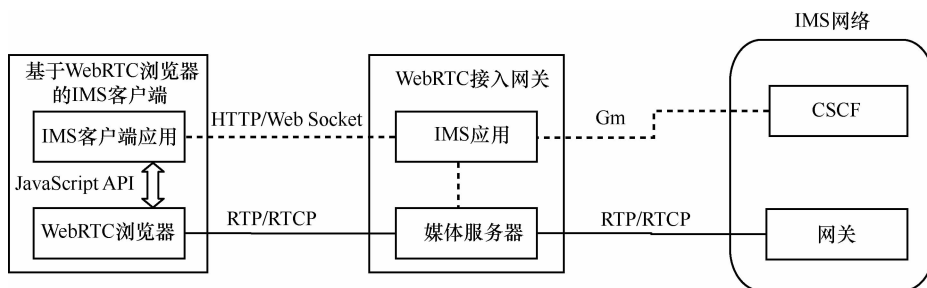


图2 用户以 WebRTC 方式接入 IMS

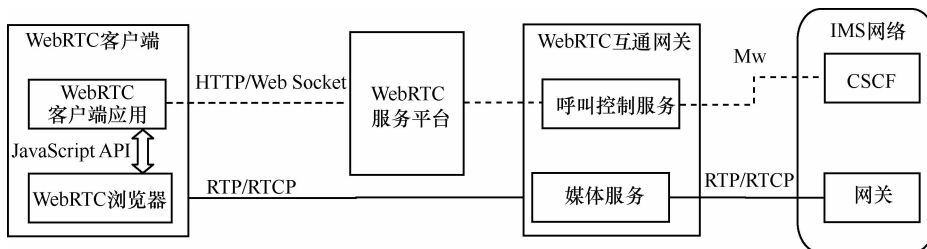


图3 IMS 与 WebRTC 互通



- 呼叫相关的协议和流程定义,包括对呼叫分支的处理,如串振、并振等方式以及 SDP 扩展和交互过程,协议的扩展目前由 IETF RTCWeb 工作组协同 MMUSIC 工作组进行。
- QoS。对于提供 DiffServ 的网络,通过定义不同媒体流(音频、交互式视频、非交互式视频、数据)的优先级提高 QoS。
- WebRTC 的安全。WebRTC 的安全威胁是显而易见的。由于浏览器提供了对本地资源的访问,一个恶意的 Web 站点可能在用户毫无察觉的情况下启动本地摄像头、麦克风等侵犯用户的隐私,或者冒充某个用户的身份发起呼叫而难以识别。另外,屏幕共享时,Web 站点通过 JavaScript 代码可能窃取超出用户共享范围的内容。
- IMS 用户以 WebRTC 方式接入,提供用户的鉴权、计费、紧急服务以及合法监听等业务。
- WebRTC 与 IMS 互通。
- WebRTC 与 PCC(policy and charging control)结合,以提高 QoS 和管控能力。

此外还有一些技术问题等待探索,如 HTTP-only 型 NAT/防火墙的穿越、WebRTC 紧急服务和合法监听的提供等。

5 结束语

在这个一切走向 Web 的时代,WebRTC 以其极大的便利性必将获得长足的发展。从现有的态势看,传统客户端方式的 VoIP 应用都纷纷推出对 WebRTC 的支持。各个主流的浏览器厂商对 WebRTC 的支持更是亦步亦趋。WebRTC 的开源和对通信能力的封装大大降低了 VoIP 应用的开发难度,应用开发者不需要具备专门的 VoIP 经验就可以开发出专业的 WebRTC 应用,这将促进互联网化的 VoIP 应用能力的蓬勃发展。而 WebRTC 使用的便利性使得用户可以随时利用此类应用发起语音、视频呼叫或者文件共享,因此 VoIP 用户规模也会迅速扩大。一方面,这种趋势会对运营商的通信类业务带来极大的影响;另一方面,VoIP 应用的普及和用户群的扩大也有利于运营商扩大服务范围和话务流量,也会带来可观的数据流量。无论是引入 WebRTC 技术到运营商业中,还是与 WebRTC 互通,都会为运营商带来更多的价值。也许未来会形成以运营商通信业务为代表的、受控的 VoIP 生态系统和一个以 WebRTC 为代表的 OTT VoIP 的生态系统共存的局面。前

者服务于有 QoS 保障、高可靠性、提供普遍服务和全球互通的用户群体,后者服务于互联网化、低成本的有通信需求的用户群体,两者之间的互联实现更大规模的互联互通。

参考文献

- 1 WebRTC.<http://www.webrtc.org>
- 2 W3C HTML5. <http://www.w3.org/TR/html5/>
- 3 W3C WebRTC Workgroup. <http://www.w3.org/2011/04/webrtc-charter.html>
- 4 W3C WebRTC API. <http://dev.w3.org/2011/webrtc/editor/webrtc.html>
- 5 IETF RTCWeb Workgroup. <http://tools.ietf.org/wg/rtcweb/charters>
- 6 3GPP TR 23.701. Study on the Support of WebRTC IMS Client Access to IMS, 2010
- 7 IETF RFC5245. ICE (interactive connectivity establishment). <http://tools.ietf.org/html/rfc5245>, 2010
- 8 IETF RFC5389. STUN (session traversal utilities for NAT). <http://tools.ietf.org/html/rfc5389>, 2008
- 9 IETF RFC5766. TURN (traversal using relays around NAT). <http://tools.ietf.org/html/rfc5766>, 2010
- 10 IETF RFC6716. Opus (definition of the opus audio codec). <http://tools.ietf.org/html/rfc6716>, 2013
- 11 IETF RFC3261. SIP (session initiation protocol). <http://tools.ietf.org/html/rfc3261>, 2002
- 12 IETF RFC3264. SDP (session description protocol). <http://tools.ietf.org/html/rfc3264>, 2010
- 13 IETF RFC6455. The Web socket protocol. <http://tools.ietf.org/html/rfc6455>, 2011

[作者简介]

付斌,男,中国电信股份有限公司北京研究院社交通信产品线高级工程师,主要研究方向为下一代通信网络及创新业务产品。

杨鑫,男,博士,中国电信股份有限公司北京研究院社交通信产品线产品线总监,主要研究方向为互联网化的创新通信业务及平台。

王松,男,法国电信北京研发中心融合业务平台实验室主任,主要研究方向为创新业务及服务平台。

林鸿,男,博士,法国电信北京研发中心融合业务平台实验室高级项目经理,主要研究方向为业务平台及创新应用。

(收稿日期:2013-09-01)