

云视频会议关键技术浅析

Analysis of Key Technologies of Cloud Video Conference

陈浩然, 胡悦(中国联通网络技术研究院,北京 100048)

Chen Haoran, Hu Yue(China Unicom Network Technology Research Institute, Beijing 100048, China)

摘要:

描述了实现云视频会议的3种架构类型,描绘了不同架构类型的网路拓扑图与系统分层图,并分析其各自特点与优劣势,同时详细阐述了云视频会议中媒体处理的关键技术,具体包括媒体分层编码技术、媒体数据交换同步技术、媒体网络适应性技术与媒体安全技术。最后对云视频会议技术发展趋势进行了展望。

关键词:

云视频会议;架构设计;分层编码;网络适应性;媒体安全

doi: 10.12045/j.issn.1007-3043.2019.05.002

中图分类号:TN915

文献标识码:A

文章编号:1007-3043(2019)05-0005-06

Abstract:

It describes three types of architectures for cloud video conferencing, and describes their network topology diagrams and system hierarchy diagrams, and analyzes their characteristics, advantages and disadvantages. Then, the key technologies of media processing in cloud video conferencing are elaborated in detail, including scalable media coding technology, media data exchange synchronization technology, media network adaptability technology and media security technology. Finally, it prospects development trend of cloud video conferencing technology.

Keywords:

Cloud video conference; Architecture design; Scalable codec; Network adaptability; Media security

引用格式:陈浩然,胡悦.云视频会议关键技术浅析[J].邮电设计技术,2019(5):5-10.

0 前言

近些年,随着各行各业对成本费用的控制以及对经营效益预期的不断提高,视频会议以其便捷、省时、省钱三大优势,打破了地域限制,提高了沟通效率,节省了沟通成本,越来越受政企客户青睐,成为日常办公的刚性需求。与此同时,主流的视频会议业务实现方式也发生了变革,服务侧从传统驻地式硬件设备向云架构平台迁移,并通过SaaS模式提供视频会议服务,终端侧也以手机端、Pad端与PC端的APP形式为主,进一步降低了视频会议的投资成本。这种云架构的视频会议系统以先进的音视频技术、丰富的数据业务、广泛的终端类型支持、轻资产的投资方式、计费灵

活的商业模式,逐渐替代传统视频会议系统,并通过技术的快速迭代满足企业用户日益增长的音视频实时通话与数据协同业务的需求。

1 云视频会议架构类型

架构设计是云视频会议系统中最核心的技术。目前云视频会议架构在实现上有MCU资源池架构、IMS架构、云计算虚拟机架构3种类型,下面对其进行详细说明与对比。

1.1 MCU资源池架构

1.1.1 架构描述

MCU资源池架构是把分散在各地的专用MCU设备、流媒体转发服务器设备以及其他高性能服务器,通过资源抽象的方式虚拟成资源池,实现对全网资源统一的分配、管理、调度、监控。资源池对外统一提供

收稿日期:2019-03-27

会议端口资源,并可根据网络条件有选择地实现MCU间以服务器或板卡为单位的资源负载均衡和本地/异地容灾备份,视频业务的使用者只需要通过网络就近注册接入到所辖的资源池即可享受到服务。网络拓扑结构如图1所示,系统架构图如图2所示。

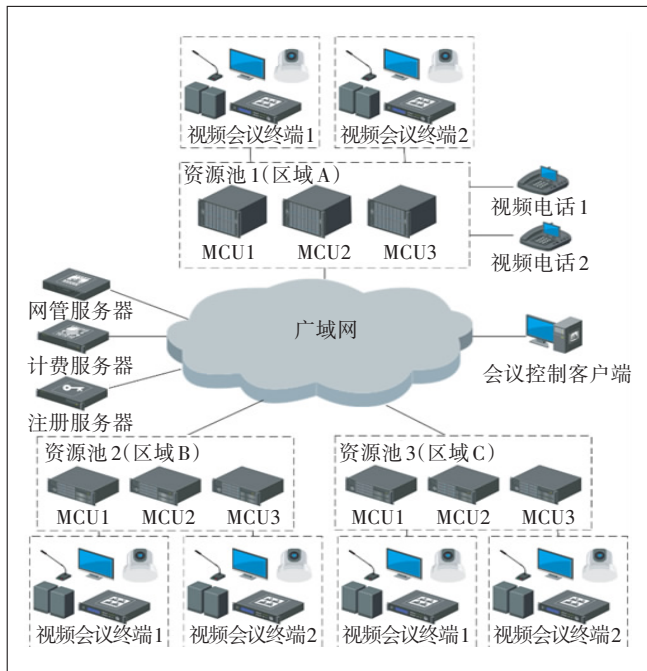


图1 MCU资源池的网络拓扑结构图

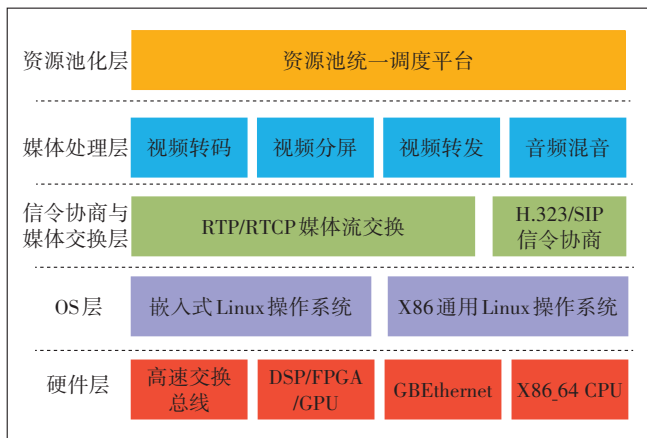


图2 MCU资源池的系统架构图

由图2可见,从硬件层到媒体处理层其实是一台MCU服务器或MCU业务板卡的内部架构,而最上层的资源池统一调度平台真正负责资源的抽象与汇聚,MCU资源池主要是通过这一调度平台实现对多台MCU资源使用上的智能判断,并实现负载均衡、就近接入、资源灾备等功能。资源池统一调度平台具体实现的功能包括MCU资源获取、MCU资源池构建、会议

资源调配、资源状态通知。

1.1.2 架构特点

该架构多数为传统视频会议厂商的云视频会议解决方案,技术成熟,有着较多的成功案例。整个系统网元除MCU资源池外,还包括注册服务器(H.323的GateKeeper或SIP的Register Sever),客户端通常是视频会议终端或支持H.323/SIP协议的软终端。

1.1.3 架构的优劣势

该架构具有如下2个方面的优势。

a) 视频编解码、音频混音由ASIC或DSP硬件实现,设备有MTBF、MTTR指标,系统较稳定。

b) 对于国际标准H.323、SIP支持较好,不同厂商间设备兼容性较好。

该架构具有如下3个方面的劣势。

a) 受限于H.323、SIP协议本身,视频画面合成由中心侧实现,需要中心侧二次编解码,终端侧多画面切换时较不灵活。

b) 资源灾备是以单台服务器或板卡为单位的,无法进行更细粒度的资源灾备。

c) 业务不灵活,如新增视频监控等视频融合类型业务,还需要新增其他硬件网元设备。

1.2 IMS架构

1.2.1 架构描述

IMS架构是指遵循IMS体系中信令传输、媒体协商、媒体控制的标准,通过IMS体系定义的会议服务器、MRFC、MRFP等核心网元,实现视频会议的业务功能。由于以IMS为基础的架构定义了完整的电信业务与多媒体服务,既支持语音、视频、数据等多种形式的会议,又支持多种类型终端接入会议,还可与IMS共用运营平台,充分体现运营商通信网络和业务的优势。网络拓扑结构如图3所示,系统架构图如图4所示。

由图4可见,通信能力层与业务能力层共同实现了典型的基本通信业务,业务逻辑层对所提供的能力进行组织,形成包括音视频会议、群组消息、数据业务等完整的业务服务,并通过业务开放层的SDK接口与行业应用进行融合。

1.2.2 架构特点

由于IMS体系结构的设计初衷是成为下一代网络的核心,体系结构包括了电信网的接入层、控制层、承载层和业务应用层各个部分,故基于IMS架构的视频会议在与其他基础通信业务的融合上有着天然优势,

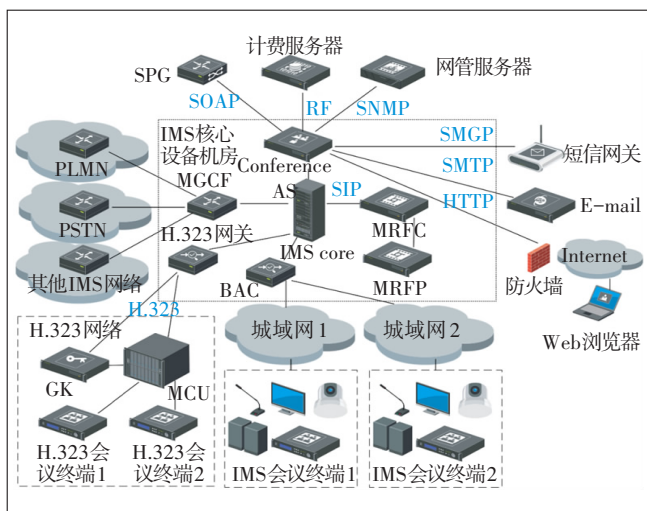


图3 IMS架构的网络拓扑结构图

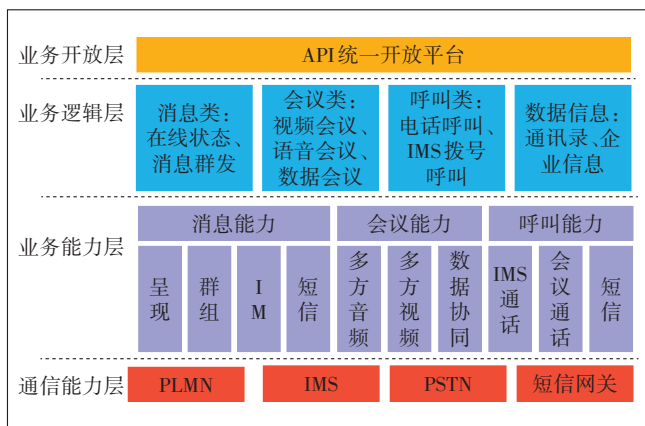


图4 IMS架构的系统架构图

因此该架构通常立足于融合通信而非仅仅会议业务本身。其次该架构中涉及到的业务应用更为广泛,涵盖了数据协同业务,并可以以中间件技术通过SDK提供能力开放接口,与行业应用进行融合。主要网元包括 Conf AS、MRFC、MRFP,而其他网元如 HSS、CSCF、CCF、SPG 可与IMS核心网设备共用。

1.2.3 架构的优劣势

该架构具有如下4个方面的优势。

a) 支持各种类型电信终端,除专用IMS视频会议终端外还可接入电话、手机、IP话机、各类软终端以及传统H.323/SIP视频会议终端。

b) 能够便捷地与基础电信业务相融合,如通话、短信、传真等业务。

c) 可为行业应用提供能力开放接口,如用户管理、通讯录同步、状态呈现、消息推送、传真发送等。

d) IMS提供电信级的QoS保证。在会话建立时,

可以按需为策略控制功能和媒体预留等机制进行网络资源分配,优先满足使用者获得多媒体通信服务的需求。

该架构的劣势为:IMS整个协议体系较复杂,虽然IMS信令基于SIP进行扩展,但要比传统SIP视频会议交互信息更多,交互流程更繁琐。

1.3 基于云计算虚拟机架构

1.3.1 架构描述

基于云计算虚拟机架构的云视频会议系统是符合NIST对云计算服务形式定义真正的云架构,该架构利用云计算中的虚拟化技术,将视频系统业务逻辑、信令交互与媒体流处理等过程在云端实现,并支持公有云、私有云或混合云方式部署,一方面以SaaS形式为用户提供流畅、低延时、高并发的实时音视频通信服务,另一方面以PaaS形式为用户提供SDK/API调用的能力,为用户自有系统集成和个性化定制开发提供灵活的服务。网络拓扑结构如图5所示,系统架构图如图6所示。

由图6可见,IaaS层提供基础架构与基础服务,PaaS层提供软件研发平台的服务,其中PaaS层也存在业务的负载均摊,与IaaS层虚拟机管理节点对工作节点的负载均摊具有不同维度。SaaS层在PaaS层之上实现了能力开放,为产品多元化与产品定制化提供手段,SaaS层业务除基本的视频会议与数据协同业务之外,还支持视频监控、流媒体点播/直播、即时通信以及

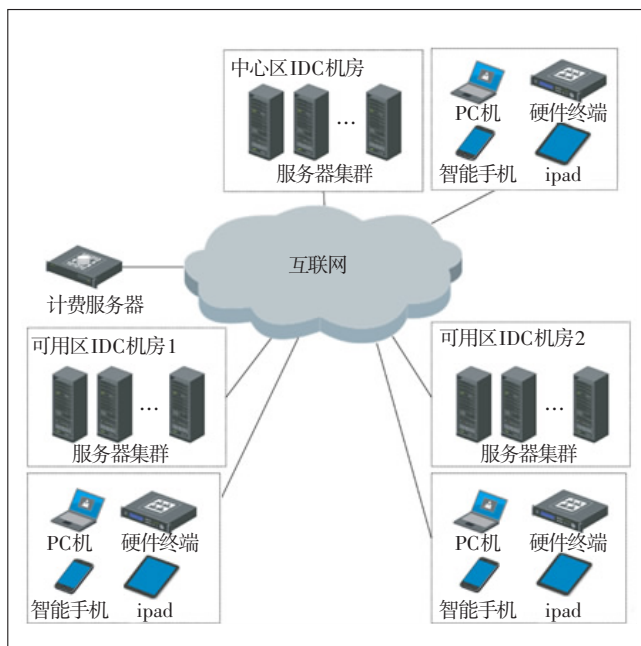


图5 云计算虚拟机架构的网络拓扑结构图

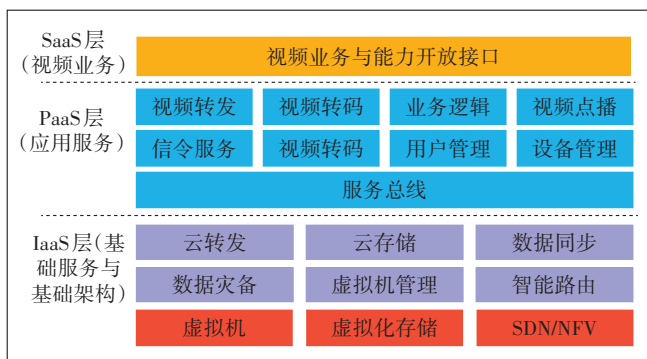


图6 云计算虚拟机架构的系统架构图

其他视频融合业务。

1.3.2 架构特点

该架构是对视频会议业务进行软件化、服务化、微服务化的功能性重组,并依托虚拟机、容器来对计算存储网络资源进行分配、管理、调度、监控。由于构建于IP互联网之上,通常并不关心IaaS所处的地理位置。部署于多个IDC机房的云视频会议系统分为中心区和可用区,中心区是需要集中统一操作功能的软件模块,包含管理监控各可用区服务状态、资源使用状况、许可资源,同步各可用区间数据,跨可用区间的会议召集时协调各可用区之间的会议智能级联,生成媒体在应用层面智能路由的策略。而可用区是可分布式部署的业务、信令、媒体功能的软件模块,提供会议管理、业务Portal提供、多方会议媒体交换等功能。考虑到稳定性和扩展性,系统设计会将中心区、可用区内部采用多台服务器集群式部署,进而支持异地灾备与分布式部署就近接入。

1.3.3 架构的优势劣势

该架构具有如下4个方面的优势。

a) 功能易扩展,纯软件的平台环境能够支持更丰富的数据业务,并可细粒度将业务划分为原子功能进行功能的柔性重组。

b) 性能易扩容,通过服务器的堆叠及软件的再部署,可以响应类似互联网业务指数型爆发式增长的速度,实现弹性扩容。

c) 负载均衡与资源灾备以虚拟机或容器为单位,相比以服务器或板卡为单位,资源使用效率高、资源恢复速度快。

d) 多画面合成多数在终端进行,故在中心侧不进行二次编解码,减少了交互时延,画面窗口的灵活变化使用户体验更佳。

该架构具有如下2个方面的劣势。

a) 与基础电信业务互通(如电话入会、移动终端入会(非APP)、VoLTE终端入会)略显不足。

b) 信令为私有信令,故不同厂商间兼容较差,与传统视频会议设备兼容需要网关服务单元,网关服务单元也是软件实现的一项服务,由于需要二次编解码来完成媒体格式的转换,计算资源占用较大。

1.4 行业应用情况

中国电信商用的“摩云”云视频会议系统采用第1种架构,中国移动商用的“云视讯”视频会议系统采用第2种架构,中国联通网研院试商用的“云视联”视频会议系统采用第3种架构。第3种架构是纯粹的互联网化架构,考虑到未来业务提供的灵活性、业务使用的便捷性、业务增长的爆发性、前瞻技术的吸纳速度等特征,第3种架构将成为实现云视频会议系统最主流的架构。

2 云视频会议媒体处理的关键技术

云视频会议除体系架构外,最核心的技术即是音视频媒体处理技术,包含媒体的编解码、传输策略、网络适应性、安全性等多个方面的内容,下面对关键技术进行阐释说明。

2.1 媒体分层编码技术

云视频会议系统构建于IP网络之上,不同于带宽质量有保障的专线,使用IP网络需要更高的压缩率、更低的失真率、更鲁棒的网络变化适应性的编码技术来保障实时音视频交互服务,进而提供更佳的主观体验。当前一项主流技术是针对不同信宿端网络传输带宽以及解码能力,在信源端提供灵活的时域、空域、质量分级编码来最大限度适应网络动态特性和信宿端的处理能力,从而为用户提供尽可能高的音视频质量,该技术被称为分层编码(Scalable Codec)。

分层编码对不同能力(包括网络收发能力和处理能力)的节点提供不同QoS的媒体服务。通过分层编码技术,信源被分成0~n层,0层为基础层,其余层为增强层,每个信宿端都可以选择下载其中的0~k层($k \leq n$),如果只下载0层能够实现基本播放,下载层数越高播放的品质越好,实现了QoS分级的媒体服务。H.264标准中SVC、H.265标准中的SHVC视频编码以及某些厂商私有的SAC音频编码都是基于分层编码来实现的。

2.2 媒体数据交换与同步技术

云视频会议系统通过媒体数据的交换实现不同

终端间的互听互视,媒体数据交换主要有“推”和“拉”2种机制。“推”就是节点主动向另一个节点发送数据,通常情况下节点之间有一种父子关系,父节点依据这种关系将数据主动发送给子节点。这种机制优点在于无须发送数据的状态信息,无须存储转发,所以负载较低、延迟较低,而缺点在于父节点失败后子节点需要重新选择父节点,需要维护复杂的树结构且对动态变化的应对能力较弱。视频监控、视频直播以及构建于组播的业务场景一般用“推”的机制。“拉”是以存储转发为基础,节点首先获得对方节点所拥有的数据状态信息,而后发起请求,后者再根据请求发送数据。这种机制的优点在于可以随意同时从多个节点获取数据,节点间也完全对等,缺点在于需要在每个节点进行必要的缓冲,传统视频会议的MCU、云视频会议的MRU、视频点播一般用“拉”的机制。

云视频会议系统中音视频是通过不同的RTP会话通道进行传输,同一个接收端会接收来自云端的多组媒体流,并在接收端进行视频合成与音频混音,而在传输过程中,由于不同类型的媒体流的网络时延不同会造成播放时间偏差,故需要媒体同步技术来恢复。同步机制具体方法是对每个类型的媒体数据流单元统一添加时间码,在发送时将按时间顺序分成单元,在同一时间轴上给每个单元都打上一个统一时间标记,在各个媒体到达信宿端相同时间戳的媒体单元同时进行解码呈现,借此用来保证接收端以正确的时间收到媒体数据,实现了不同媒体类型之间的媒体同步。

2.3 媒体网络适应性技术

云视频会议系统需要解决网络数据包丢失、时延抖动过大、带宽时变等恶劣网络或弱网络下的各种问题,解决上述问题的技术组合称为媒体网络适应性技术,除前文所提的分层编码之外,还包括自动请求应答(ARQ——Automatic Request reQuest)、自适应码率控制(ABC——Adaptive Bitrate Control)、丢失包错误隐藏(PLC——Packet Loss Concealment)、前向纠错(FEC——Forward Error Correction)等技术。

ARQ技术是通过重传关键数据包来纠错的信道保护算法,有3种常见策略:停等ARQ、退回N步的ARQ和选择重传的ARQ,目前主流丢包重传算法大多都是效率较高的选择重传的ARQ,即接收端通过序列号检查,对未收到序列号的数据包向发送端重新申请。ABC技术具体方法是发送端依据接收端周期性

的网络状态报文对网络带宽进行感知,自动调整码率来适应网络环境的变化,即在网络好的情况下,适当提高码率来提高语音视频的质量和降低延迟,在网络差的情况下,适当降低码率,适当牺牲音画质量来保障语音视频通话的可用性和流畅性。FEC技术是通过增加冗余数据对丢失的数据包进行恢复的信道编码算法。具体来说,由发送端对原始数据进行FEC编码,生成冗余奇偶校验数据包,原始数据和冗余数据包合并称作FEC数据块。接收端接收到FEC数据块后,通过冗余数据包和原始数据包来恢复出丢失或者出错的数据包。比较成熟的FEC算法有RS、LDPC、Raptor和Tornado算法。PLC技术应用于实时语音通话的场景,通过前一个语音数据包和后一个语音数据包的相关性来“推测出”当前丢失的语音数据包,从而“隐藏”了信道传输所造成的错误,减少语音数据包丢失对语音通话质量的影响。错误隐藏PLC算法在接收端进行,不需要发送端参与。

一般情况下,根据网络感知后的往返时延(RTT)和丢包率(PLR——Packet Lost Rate)来智能决定采用哪种策略或哪些策略的组合(见图7)。

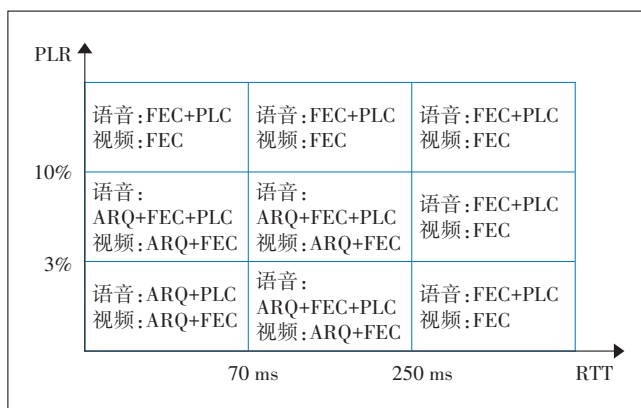


图7 不同RTT和PLR下的丢包策略组合

2.4 媒体安全技术

云视频会议的媒体安全也是系统设计中的重中之重。早期的媒体安全是依赖国际标准通信协议的安全服务协议,如H.323协议族中的H.235协议、SIP协议族中的HTTPS摘要认证与S/MINE的安全协议,用来解决通信安全面临的威胁,如注册劫持、服务器伪装、消息篡改、拆卸会话以及拒绝服务。随着互联网上多媒体应用的繁荣,更多媒体应用的安全依赖互联网媒体的安全协议,如信令上使用安全传输层协议(TLS)、媒体传输上使用数据包传输层安全性协议

(DTLS)、安全实时传输协议(SRTP)、组合Z的实时传输协议(ZRTP)、安全实时消息传输协议(RTMPs)等,用来解决传输安全、编码安全面临的威胁,如会话劫持、媒体流截获。近2年随着国密算法的成熟以及信息安全技术国家标准的完善,媒体安全开始更多地关注于自主可控的国密标准或技术,逐步构建在SM2、SM3、SM4、GB35114等加密算法或安全体系之上,并开始考虑在数据、编码、传输、应用等角度全方位地实现安全设计。

3 云视频会议与新技术的融合

3.1 云视频会议+微服务容器技术

云视频会议可以结合微服务容器技术,利用其快速迭代、敏捷交付、弹性部署的特点,迅速满足业务扩张过程中对系统架构的新要求以及行业定制化过程中对功能的新需求。使用微服务技术,围绕着视频业务领域来创建可独立开发、管理的组件与应用,让产品迭代与交付变得更简单;而使用容器技术,将视频会议轻量级业务进行隔离部署,通过容器集群调度技术,可实现视频业务的高并发处理、业务动态编排管理,让弹性部署成为可能,进而更迅速地响应用户的爆发式增长。

3.2 云视频会议+AI技术

云视频会议可以结合主流的AI技术,为用户提供更富科技感的操作体验。使用语音识别技术,可以实现语音助手协助下的会议预约与会议控制,并可自动将语音转换成文本,生成会议记录;使用图像识别技术,可以实现自动点名、人物身份标注、会场背景更换、AR效果叠加;使用自然语言处理,可以实现会议摘要生成、会议主题词云生成以及关联内容检索。AI技术对云视频会议的赋能,优化了会议流程,简化了文书记录操作,方便了会议历史回放的检索,让用户更关注于沟通本身信息的传递。

3.3 云视频会议+采集呈现技术

云视频会议可以结合多视点视频采集、麦克风阵列、环绕立体声、全息成像、超高清大视场显示覆盖等手段以及VR/AR等设备,营造更为逼真的远端环境,全方位、全场景提升用户对媒体信息的感受能力,使用户达到对媒体影像级的临场感、沉浸感和真实感。使用VR技术,既可将全景计算机虚拟世界进行呈现,又可将全维视角进行呈现,达到虚实交融的呈现效果;使用AR技术,可实现视频、音频、图像、数据等多

维信息的汇聚融合,将可视化的信息数据叠加到现实空间,丰富用户的信息获取渠道。超高清高保真的视音频采集呈现技术连同VR/AR设备将进一步增进视频通信服务框架内现实世界的远端、近端与虚拟世界的互动。

3.4 云视频会议+5G

云视频会议可以结合5G面向视频业务的解决方案,为移动客户端用户提供更优质的音视频体验。使用eMBB切片所提供的差异化服务,能够满足移动端4K超高清、低时延的视频传输需求。使用边缘计算技术,能够将云计算框架体系进行系统性延伸,视频分发、视频转码服务可下沉至边缘设备中,构建区域级小规模云视频业务,减少了业务数据路由,降低了云端计算的负载,增强了网络的韧性。5G技术能够从根本上为带宽消耗时延敏感的视频交互业务提供强有力的传输保障。

4 结束语

云架构的视频会议满足了当前视频通信系统大规模服务的需求,降低了子系统间、组件间的耦合性,实现了功能可扩展、性能易扩容、数据易灾备、服务高可用的能力,满足业务内容的多样性、产品迅速迭代的要求。

随着基础技术的迭代式更替与应用技术的颠覆式创新,云视频会议媒体处理的关键技术也将不断向前演进。同时,云视频会议通过与新的Web架构技术、采集显示技术、AI技术、通信技术的融合,将实现自身的能力重构,释放更大的产业化潜力。

参考文献:

- [1] 薛森,胡悦,符刚.云视频会议趋势及探讨[J].邮电设计技术,2017(7):4-8.
- [2] 史俊潇,贺琛,卢晓帆.IMS技术在4G多媒体视频会议中的应用研究[J].电力信息通信,2017(10):120-125.
- [3] 管磊.P2P技术揭秘[M].北京:清华大学出版社,2011:362-363.

作者简介:

陈浩然,毕业于电信科学技术研究院,高级工程师,硕士,主要从事视频通信业务、能力开放研究工作;胡悦,毕业于西安电子科技大学,工程师,硕士,主要从事核心网新技术研究、企业云通信研究工作。

