

中国实时音视频行业研究报告

©2022.8 iResearch Inc.

发展背景



网络基础设施升级、音视频传输技术迭代、WebRTC开源等因素，驱动音视频服务时延逐渐降低，使实时音视频（RTC）技术成为炙手可热的研究方向。实时音视频业务在消费互联网领域蓬勃发展，并逐渐向产业互联网领域加速渗透。经历了行业第一轮的红利爆发期，**我国实时音视频行业的场景效能逐渐深化，步入到理性增长阶段。**

市场规模



2021年中国实时音视频（RTC）PaaS市场规模为16亿元，消费互联网领域为实时音视频行业贡献了核心收入来源。受到社交娱乐头部应用的高度渗透及“双减”政策的持续影响，预计未来三年的复合增长率为28.4%，**2024年实时音视频（RTC）PaaS市场规模将达到30亿元。**

竞争格局



实时音视频赛道玩家的主要竞争策略分为两大类：（1）**提供通用的PaaS层能力**，包括RTC PaaS厂商、通信云PaaS厂商、综合型IaaS厂商。其中，**RTC PaaS厂商在技术研发和产品打磨上具备更高的行业专注度，占据市场的主导地位**；（2）**根植于特定行业场景**，主要为垂直行业解决方案商。

核心技术



实时音视频行业存在较高的资源门槛与技术壁垒。主要包括：软件定义的实时音视频传输网络（Real-time Network, RTN），基于UDP的协议层优化，以及弱网传输保障策略。在实际应用中，RTC与CDN技术的融合，衍生出**实时互动直播、超低延时直播**两大技术路径。

趋势洞察



实时音视频标准化工作的推进将引导行业迈向更高的服务质量，推动PaaS层厂商与垂直行业解决方案商之间的良性竞争与协作共赢。同时，实时音视频逐渐向**实时互动领域的延展**，元宇宙相关应用的出现将为消费互联网创造更大的想象空间。此外，实时音视频技术有望成为我国泛娱乐出海企业的差异化竞争机会，**海外业务版图的扩大**也将驱动行业迎来新一轮的业务增长。

概念定义



实时音视频

Real Time Communications (RTC)

基本释义：

在远程条件下，以接近实时/可忽略延迟交换信息的通信方式，包括固定通话、移动通话、音频会议、视频会议、网真等形式。

本篇报告研究的实时音视频（RTC）服务特指：

- **通信方式**：以音频、视频的形式赋能行业；
- **服务模式**：基于云计算技术理念提供，包括实时音视频的底层通用能力，以及垂直行业的解决方案；
- **底层技术**：采用软件定义的实时音视频传输网络（Real-time Network, RTN）和基于UDP的传输协议；
- **时延区间**：本篇报告聚焦于广义的实时音视频，包含时延400ms内的强互动，以及800ms内的超低延时互动。

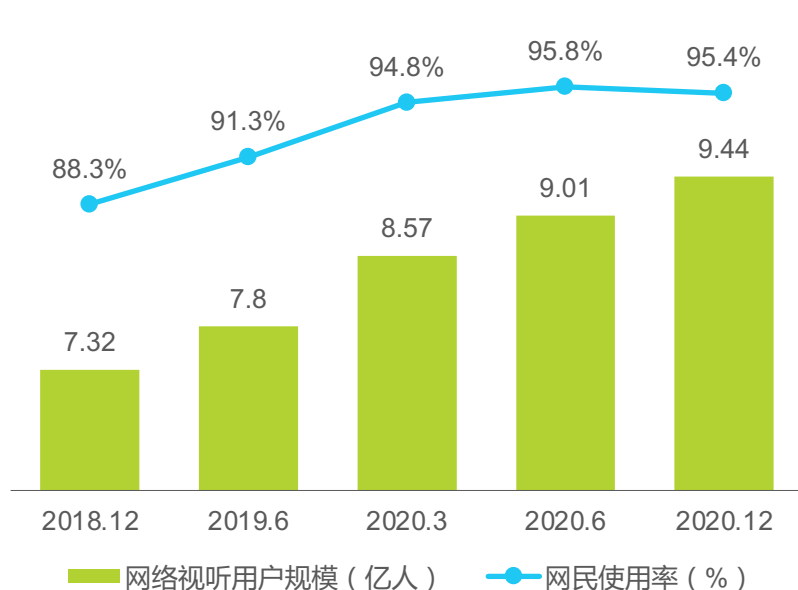
中国实时音视频行业发展分析	1
供给侧：实时音视频行业发展能力洞察	2
需求侧：垂直行业实时音频应用实践分析	3
中国实时音视频行业典型企业案例	4
中国实时音视频行业发展趋势	5

音视频消费习惯的迁移

用户音视频习惯养成，并逐渐向强实时性、强互动性场景延伸

富媒体信息时代，音视频已成为人们获取、发布、交换信息的重要方式。截至2020年底，中国网络视听用户占整体网民规模的比例已高达95.4%。高饱和的渗透率水平，使用户更加关注音视频服务的体验感。得益于底座能力与关键技术的持续突破，音视频服务体系实现了从点播、直播到实时音视频的深化发展。实时性与互动性的长足优化，逐步激活了更多场景下的音视频互动模式，驱动用户的音视频消费习惯向更加还原真实、更加沉浸式的实时音视频服务迁移。同时，伴随疫情的常态化发展，人们远程办理业务的习惯已逐渐养成。实时音视频作为一种通用型能力，撬动了传统行业中众多强实时、强互动场景的数字化升级，也使得实时音视频的消费趋势进一步从消费互联网向产业互联网延伸发展。

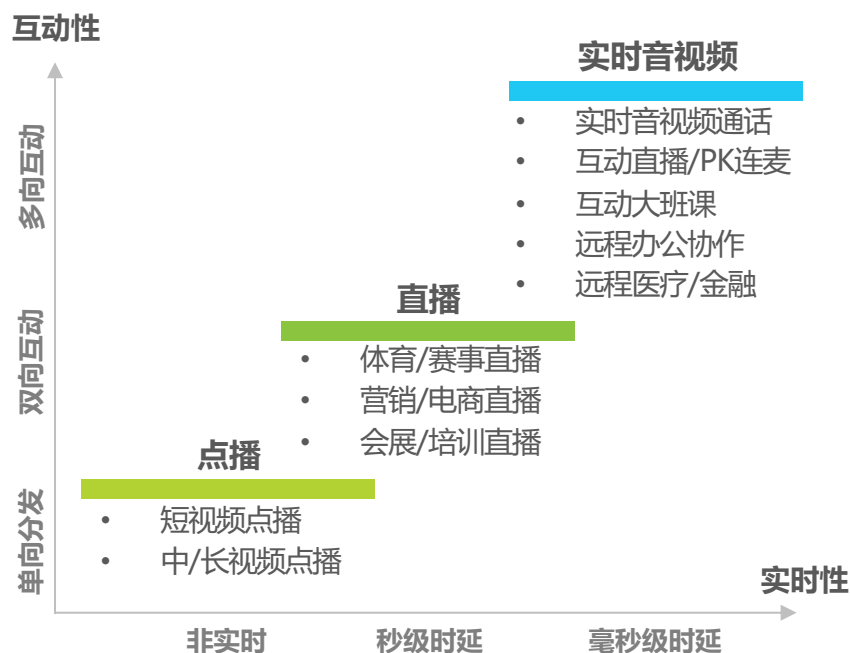
2018-2020年中国网络视听用户规模及使用情况



注释：网络视听用户为综合视频、短视频、网络音频、网络直播用户的并集；网民使用率指网络视听用户占网民规模的比例。

来源：CNNIC；中国网络视听节目服务协会；艾瑞咨询研究院自主研究及绘制。

音视频消费习惯迁移趋势及典型应用场景



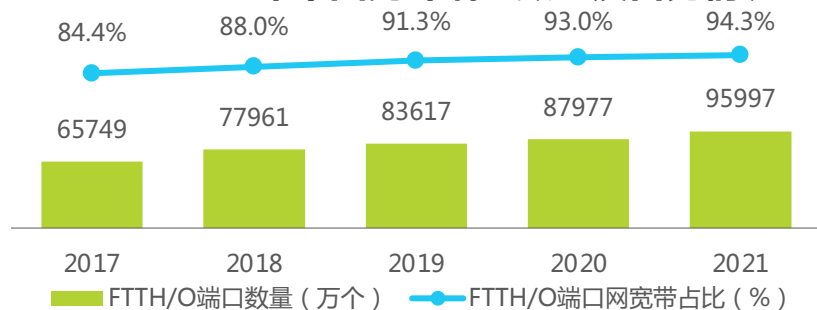
来源：艾瑞咨询研究院自主研究及绘制。

网络基础设施持续升级

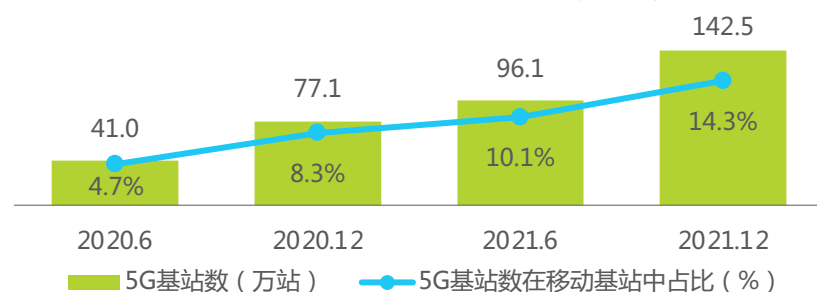
高速泛在的骨干网络，为实时音视频传输提供坚实保障

实时音视频行业的高速发展离不开国家骨干网络的建设。以千兆光网和5G为代表的“双千兆”网络，近年来取得突破性进展。2021年，我国光纤接入（FTTH/O）端口已达到对互联网宽带端口的高度覆盖，光纤接入能力已普遍超过百兆，并向千兆以上速率不断升级；5G网络已覆盖国内所有地级以上城市，5G基站数累计高达142.5万，用户数占全球的89%。网络基础设施的广泛渗透，使更多用户可以受益于终端设备与骨干网络之间的快速连接，享受到无感接入、触手可及的高质量网络服务。同时，“双千兆”网络在带宽能力上的显著提升，可大幅降低音视频信息的延迟和缓冲时间，提高端到端毫秒级传输时延的满足率，为实时音视频的数据传输与落地应用提供良好的底层网络支撑。

2017-2021年中国光纤端口数量及占比情况



2020-2021年中国5G基站数量及占比情况



来源：工信部；中国信通院；艾瑞咨询研究院自主研究及绘制。

国内某省会城市运营商4G/5G下视频通话达标率

	4G网络	5G网络	相对增益
无卡顿样本端到端时延均值	541.81ms	412.60ms	23.8%
200ms时延满足度	0.54%	10.40%	1827.6%
300ms时延满足度	24.49%	66.72%	172.5%
400ms时延满足度	51.89%	78.06%	50.4%
音频卡顿达标率	70.37%	85.05%	20.9%

注释：1、该视频通话测试，测试时段涵盖网络忙闲时（7:00-21:00），采集端帧率为30fps，传输协议是UDP+FEC，编码方式为H.264，编码后分辨率为1280*720，帧率为15fps，编码目标码率为2Mbps；2、音频卡顿率参照声网“体验等级协议XLA”，单位时间（1分钟）内200ms音频卡顿率大于3%时，记为不达标。

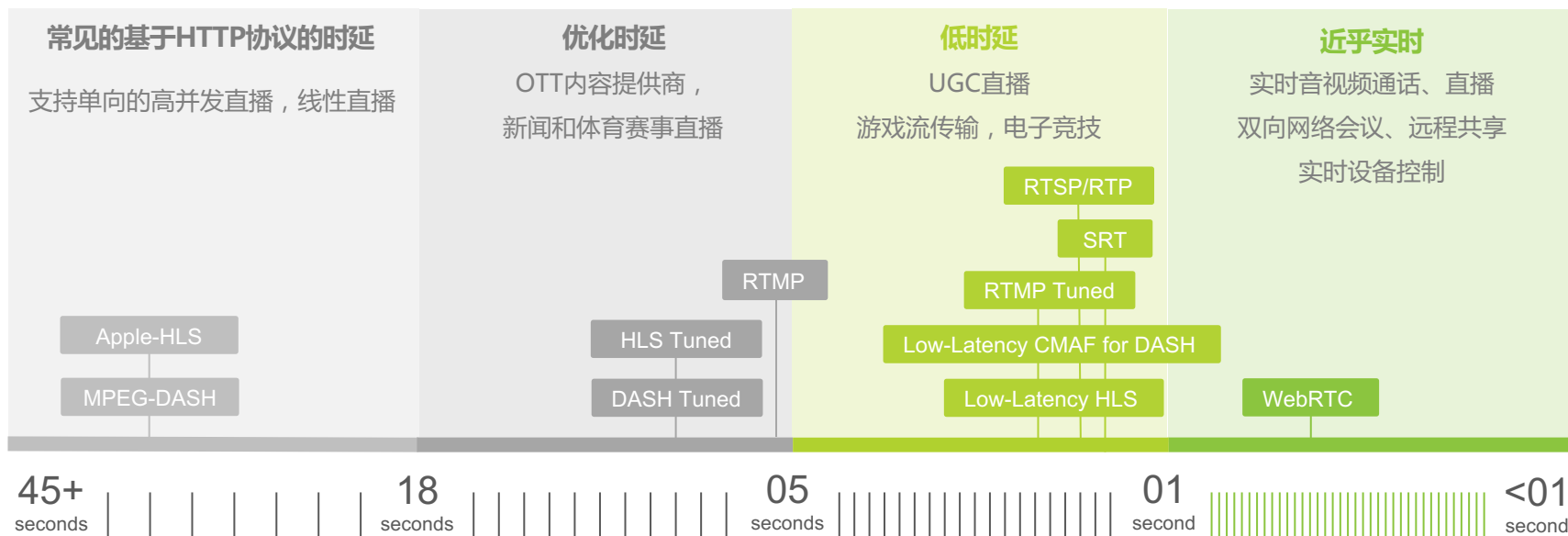
来源：中国信通院；艾瑞咨询研究院自主研究及绘制。

音视频传输技术更新迭代

技术驱动音视频走向实时，流媒体协议迈入毫秒级传输阶段

流媒体技术可将压缩处理后的音频、影像连续性地上传到网站服务器，使文件无需完全下载到本地即可观看。在编解码方面，H.265、国产AVS3等视频解编码技术可以使文件的压缩体积更小、传输速度更快。其中，AVS3作为全球首个面向5G产业应用的音视频信源编码标准，已被成功纳入DVB（数字视频广播组织）标准体系，可以在同等画面质量下比H.265节省40.09%的码率。在协议方面，基于TCP的RTMP、HLS、DASH等协议的优化方案仍存在2-3秒左右的延迟，而基于UDP的WebRTC则突破性地将延迟降低至毫秒级别。与其他协议不同，WebRTC作为流媒体通信框架，覆盖音视频采集、编解码、传输和渲染的全部环节，可以为实时音视频提供全流程的理论依据与技术支持。

流媒体传输协议/框架的适用场景及时延对比



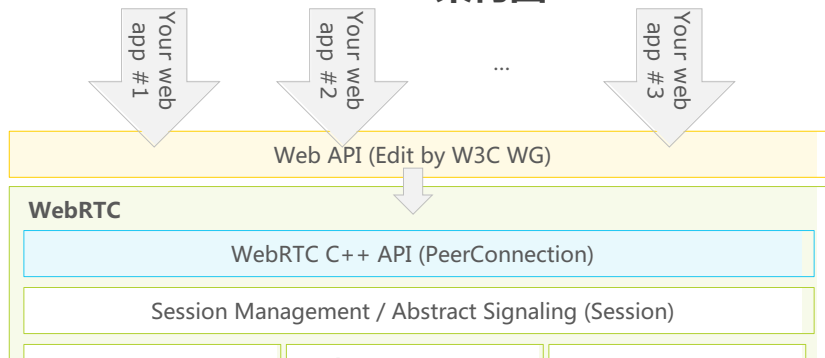
来源：WOWZA官网；艾瑞咨询研究院自主研究及绘制。

商用方案成为行业主流

WebRTC奠定技术框架，在基础上演化出第三方服务商

WebRTC是目前实时音视频领域最流行的开源框架。2010年Google收购GIPS引擎后，将其纳入Chrome体系且开源后，命名为“WebRTC”。WebRTC获得各大浏览器厂商的支持并纳入W3C标准，促进了实时音视频在移动互联网应用中的普及。2021年1月，W3C和IETF两大标准制定组织宣布WebRTC成为官方标准，用户无需下载额外组件或单独的应用程序，便可以支持在网络上的实时音视频通信。尽管WebRTC具有免费开源的特性，但其庞大、繁杂，学习门槛高，又缺乏服务器方案的设计和部署，为基于WebRTC搭建的商用方案留下了发展空间。第三方的RTC PaaS厂商凭借规模效应和技术优势成为开发者的首选，推动实时音视频行业进入发展的快车道。

WebRTC架构图



WebRTC开源服务器存在的主要缺陷

1 终端接入

WebRTC主要面向Web应用，在手机端支持上项目复杂度高，需要兼容不同版本、机型，容易出现回声、摄像头打开失败、屏幕录制失败等各种问题

2 并发支持

面对多地多用户接入的场景，单台服务器带宽有限，需要考虑服务器集群之间的级联，以及多地部署的分布式服务器方案，但WebRTC开源服务器缺乏整体的服务器设计和部署方案

在增加转发层的情况下，WebRTC的传输策略

预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=1_45539

