&腾讯云I •腾讯云音视频| 中国信通院

超低延时直播

白皮书

腾讯云

中国信息通信研究院云计算与大数据研究所

2022年2月

■前言 1

01 /直播行业发展背景 2

[®秀场游戏场景推动网络直播产业初具规模 5](#bookmark13)

[®在线教育兴起为超低延时直擂带来发展机遇 8](#bookmark16)

[®各行业“直播+”转型加速超低延时直播规模化落地 9](#bookmark19)

02 /直播技术发展襪述 11

[®标准直播延时问题仍成为直播技术发展瓶颈 12](#bookmark25)

[®直播流媒体协议向低时延目标不断演进 13](#bookmark28)

[® WebRTC技术优势促进超低延时直播快速发展 16](#bookmark43)

03 /超低延时直播技术架构及性能优化 19

®超低延时直播技术架桐 20

[-上行接入 20](#bookmark81)

[-网络加速 21](#bookmark105)

-下行接入 22

[®超低延时直播技术优化 23](#bookmark114)

[-延时优化 23](#bookmark117)

・卡顿优化 24

-首帧优化 37

[-媒体处理优化源 28](#bookmark150)

[04/超低延时直播推动多场景应用发展与迭代 29](#bookmark156)

®在线教育 30

®电商帯货 31

®赛事直播 32

®会展直播 33

05 /超低延时直播未来发展趋势 34

・参考文档 36

前言

上世纪末，流媒体直播技术起始于传统广电行业发展。早期的媒体协议基于MPEG-TS(Mpeg Transport Stream) 格式传输于专用网络环境。随着互联网技术的发展，直播技术逐步应用于包括泛互、教育、广电等多个行业。

我国的网络直播起始于20世纪初，早期主要是娱乐秀场和实时游戏赛事场景，逐渐成为一种新兴的网络社交与互 动方式，网络直播平台也成为了一种崭新的社交媒体。网络直播相比广播、电视、互联网视频，以及近几年广泛使用的 微信微博，有着独一无二的媒介特性。随着新政策的引导及资本市场的追捧，直播行业首先在娱乐、游戏直播等行业 经历了一次大规模增长，2020年后，疫情爆发进一步促进了直播行业高速发展，正在全面赋能电商、旅游、文化传播 等众多领域。

直播业务规模的高速增长，直播技术本身也同时得到了长足发展。目前业界所使用的直播传输技术，具有代表性的如： HLS(HTTP Live Streaming)、MPEG-DASH(MPEG Dynamic Adaptive Streaming over HTTP)其延时一般高达 10-30 秒;而基于 HTTP 协议的 FLV(Flash Video)、RTMP(Real Time Messaging Protocol)传输技术，进一步将 延时降低到了 3-5秒，但仍然难以满足业务对低延时的渴求。随着人们对直播过程中强互动、低延时的迫切需要，基 于WebRTC协议的超低延时技术应运而生；然而其昂贵的网络成本、有挑战性的伸缩部署，进一步限制了其在大规 模、超低延时直播领域的应用。

本报告主要介绍了直播行业发展背景和技术演进路线，详细阐述了基于WebRTC的技术升级、传统CDN(Content Distribution Network)网络架构的融合改造，以解决大规模、超低延时直播的分发所面临的成本挑战。通过深度 优化各项QoS指标解决WebRTC技术在直播领域面临的用户体验的挑战，最终使得超低延时直播具备替代传统 直播技术的能力，被应用到更多的业务场景，例如电商、教育、事件直播等垂直领域,促进了众多行业业务的快速增长。

直播行业发展背景

1. -秀场游戏场景推动网络直播产业初具规模
2. -在线教育兴起为超低延时直播带来发展机遇
3. -各行业“直播+”转型加速超低延时直播规模化落地

*直播行业发展背景*

中国网络直播产业起始于2005年，早期PC端直播主要集中在娱乐秀场、游戏赛事行业等。2015年以后，随着移动 直播的兴起，直播产业逐步渗透到教育、电商等众多行业，直播平台、观众数量都呈现井喷式发展。时至今日，网络 直播已经形成一个新的经济体系，成为当前互联网经济的重要组成部分，在网络文化内容供应、技术创新、商业模 式创新的方面发挥了重要作用。

据部分行业机构统计发现，2016-2019年，我国网络直播市场规模不断扩大，2019年为843.4亿元，同比增长 63.4%,占中国网络视听行业总规模的19%o 2020年后，在新冠肺炎疫情的影响下，“全民直播”时代来临，网络直 播市场达到历史新高。2021年，中国网络直播预计将持续高速增长，市场规模有望达到0.3万亿元。随着网络直播渗 透率的不断提升，我国网络直播习惯基本养成，预计至2026年我国网络直播市场规模将超过2万亿元。

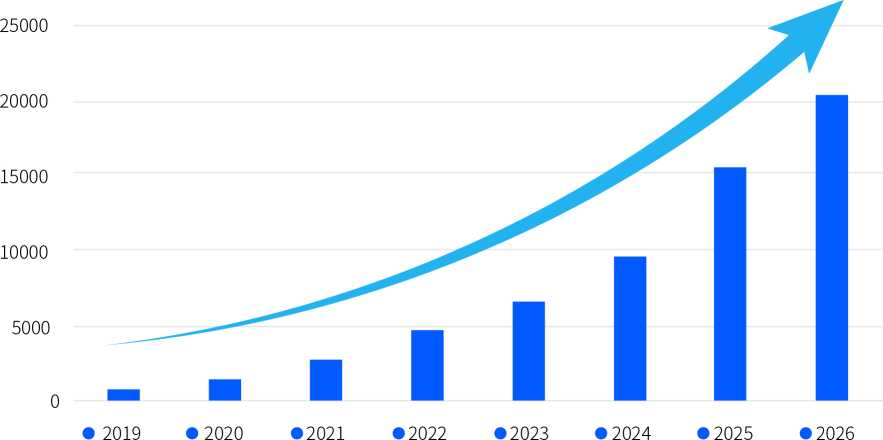
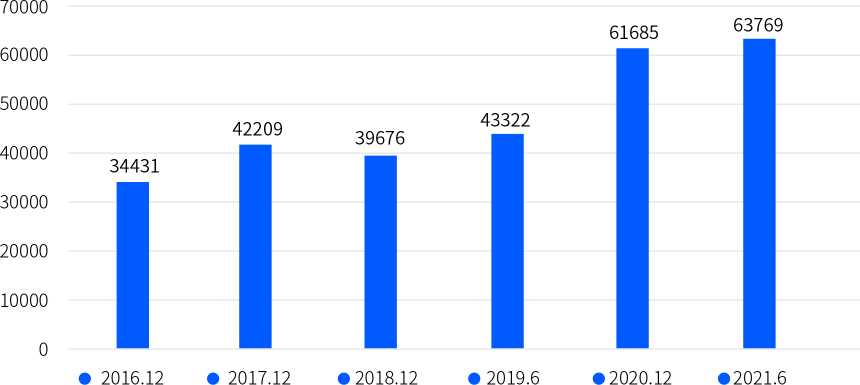


图1： 2019-2026年中国网络直播行业市场规模及预测1

从整体网络直播用户规模来看，自2020年初新冠疫情的影响以来，我国网络直播用户规模也出现了大幅增长。根据中 国互联网络信息中心（CNNIQ2021年2月3日发布的第48次《中国互联网络发展状况统计报告》统计数据显示，截止 2021年6月末，中国网络直播用户规模达到6.37亿人，比2018年12月末增加了2.4亿人。



YR蜒g仏此曄\*|蠟匿

图2： 2016-2021年中国网络直播用户规模2

通过对2021年中国网络直播梯队的主要构成数据分析看，传统的网络直播模型以及用户群体集中在秀场、游戏以及 教育行业。直到疫情的爆发，类似电商直播等新的行业趋势和方向才逐渐明朗。

。*秀场游戏场景推动网络直播产业初具规模*

娱乐秀场直播是最早的网络直播模式，营收方式大多依赖于观众各种形式的打赏。随着网络直播技术以及形态的 不断发展，新的模式不断冲击传统的秀场直播市场，目前秀场直播平台纷纷布局多元化业务，加码社交、强调互动。

整体来看，我国秀场直播行业发展经历了4个阶段：2013年前的萌芽期.2013-2014年的成长期.2015-2017年的爆 发期，2018年至今进入成熟期。

**2005-2012 2013-2014 2015-2017 2018-**至今

萌芽期

成长期 爆发期

成熟期

图3：中国秀场直播发展历程

早期的秀场直播互动主要表现为消息弹幕、点赞、送礼等较为传统的形式。然而，这种二维的互动方式限制了主播 和观众的互动沉浸感，连麦互动的逐渐成为娱乐秀场直播的刚需。

连麦互动最大的技术难题就是解决双向直播带来的延时，超低延时(延时＜ls)直播应运而生。为了满足业务场景，业 界众多实时音视频解决方案和产品快速涌现。实时音视频技术底层通常采用uDP(User Datagram Protocol)协议， 结合jitterbuffer以及播控策略，具备较好低延时特性，多数此类产品的时延控制在500ms内，为直播互动提供了有 力的技术支撑。

游戏直播是早期直播发展较好的另一个领域。游戏直播起始于20世纪初期，由最早期的视频网站和语音平台萌芽，之 后先后经历了增长期、爆发期，目前正处于稳步发展的成熟期。

萌芽期:2013年以前

增长期:2013-2014年

爆发期:2015-2018

成熟期:2019-至今

•电竞塞事网络直播萌芽

•基础设施升级,4G以及光纤入户促进网络提速

• “千播”大战,热梦游戏内容，催生游戏直播爆发

•反垄断监管，头部平台合并终止

•早期视频网站探索直播平台

•资本催生直播平台快速扩张

•港讯控股斗鱼、虎牙两大游戏直播平台

•游戏内容“寒冬”，版号限制影响大

•哗哩哗哩、快手、字节持续加大游戏直播投入

•游戏版号问题缓解,新游戏促进规模进一步增张



持续演进发展

起始于20世纪初期

图4：中国游戏直播行业发展历程游戏直播早期主要是部分游戏平台对大型游戏赛事进行实时直播的视频内容服务，逐步发展为部分主播通过直播平 台对热门游戏内容或游戏实时场景进行解说的形式。为了增加和用户的实时互动，大部分直播平台还会为每个游戏直 播间设置弹幕，方便观众实时评论。

当前，我国游戏直播已处于成熟期，伴随着国家对整个游戏行业、特别是针对青少年的监管体系日趋完善，整个游戏 直播行业进入有序竞争环境，预计未来行业市场规模将进入稳步发展，预计到2022年，游戏直播市场规模将达到530 亿元人民币。

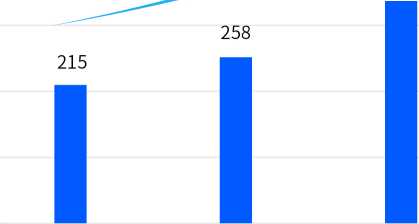
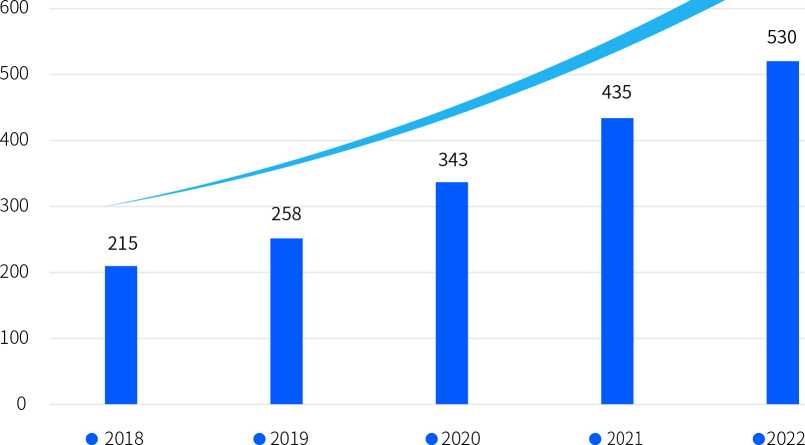


图5：游戏直播行业发展情景预测3

从游戏直播行业的发展趋势看，游戏直播平台不仅仅在用户流量、主播资源这两方面竞争，更在向用户精细化运营、 内容多元化挖掘、直播技术特性创新等更广泛的领域进行全面竞争，以建立独有的平台优势。

2020年底，“全真互联”的概念描绘了一个万物可互联、互动超真实的互联网世界。随着低延时直播、云渲染等新兴技

术的持续发展。“全真互联”逐步在游戏直播领域中落地。

3.数据来源：www.iresearchchina.com

—方面，更多的直播创新玩法应用到游戏直播中。以虚拟人为例，2021年元宇宙概念爆发，虚拟人投资热潮在国内一 级市场方兴未艾。同时，国家政策鼓励虚拟人创新发展，扶持文化与科技产业深度融合。2021年10月广电总局发布《 广播电视和网络视听“十四五”科技发展规划》，指出“要推动虚拟主播、动画手语广泛应用于新闻播报、天气预报、 综艺科教等节目生产，创新节目形态，提高制播效率和智能化水平”，首次明确地鼓励和支持虚拟人等新直播技术的 发展和应用。

另一方面，游戏直播逐渐呈现更沉浸式发展趋势。随着5G等更大传输带宽的新的网络技术发展，大型游戏赛事可以将 4K实时画面信号输出全部都上传至云端，可以通过多角度的VR机位应用实现VR直播。用户可在客户端中观看高清VR 直播，自由选择切换视角，真正实现了足不出户也能亲临现场的沉浸式体验，且在空间感的感受上会更加舒适。

新的游戏直播玩法下，信息接触、人机交互的模式发生更丰富的变化。这是一个从量变到质变的过程，意味着线上线 下的一体化，虚拟世界和真实世界的融合，帮助用户实现更真实的体验。这些强烈的低延时直播技术能力需求，推动 着实时通信到音视频、计算能力等一系列基础技术进行升级优化。

*,在线教育兴起为超低延时直播带来发展机遇*

直播是在线教育行业的重要手段。通常理解，在线教育是基于互联网基础技术能力进行授课和学习的方式，这种教 育方式的优势在于突破空间和时间的限制，可以最大程度降低线下固定的教室、排班表固定时间等问题的影响和制 约，在一定程度上有效降低了人们接受教育的门槛。基于这样的优势，2013年以后，在资本的推动下中国在线教育 行业开始蓬勃发展，无论是1对1还是1对多，还是职业教育等众多形式，都借助直播实现了初步规模化发展，并在 2017年后逐步推向发展高潮，走向成熟。目前在线教育行业已初步形成多样的细分领域，行业呈现精细化发展趋势。

在经济层面上，随着人民生活水平的提高，中国居民消费结构不断迭代，教育文化娱乐支出比重上升明显。据国家 统计局发布的《2020年居民收入和消费支出情况》数据显示，2019年，中国居民人均教育文化娱乐支出为2513元， 同比增长12.9%。随着生活质量的提高，人民对教育的重视程度和投入将不断增加，为在线教育行业的发展提供了 良好的经济基础。

在技术层面上，实时音视频产品也为在线教育的爆发提供了有力的技术支撑。这种直播方式打破时间、空间限制， 增强教、学互动性，提高性价比，受到大量教育企业和消费者的青睐，如学科、职业教育、素质教育等均进入线上直 播教学。

然而，早期的实时音视频技术存在的不足也限制了直播技术在在线教育领域的大规模应用。在基础资源上通常使用 BGP网络以及核心骨干网区域的IDC机房，总体硬件成本较高，限制了低延时直播技术的大规模应用。行业内的实 时音视频底层技术往往采用私有协议，整体封闭性较高，必须集成各家厂商SDK(Software Development Kit)[[1]](#footnote-2),技 术侵入性较高。这些弊端往往会导致一个在线教育平台的产品需要集成多家实时音视频产品的SDK,接入和使用复 杂度较高。

。*各行业“直播+”转型加速超低延时直播规模化落地*

随着直播行业高速发展，直播应用正在全面赋能电商、旅游、文化传播等众多领域。中国网络视听节目服务协会发 布的《2021年中国网络视听发展研究报告》显示，受直播经济增长趋势推动，越来越多的网络文化企业、传统企业 增项经营内容或入局网络直播领域，为产品和服务拓展新的推广和销售渠道，通过“直播+”的形式实现转型升级。

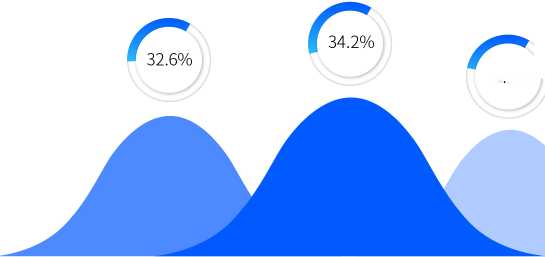
电商直播从2016年淘宝、京东等传统电商探索“直播+电商”新模式开始，经历了萌芽期、成长期、快速发展期三个 阶段，2020年开始，受疫情影响，直播电商GMV已经突破1万亿元。从艾瑞咨询等相关机构数据预计未来几年直播 电商仍将迅猛发展，2025年直播电商市场GMV将接近7万亿元。

■注:1.直播电商用户渗透率二直播电商用 户/同期互联网平均月活用户规模

2.直播电商用户指在统计周期内提交 过一次直播电商购物订单的用户

• 2019 年 • 2020年 • 2021 年 H1

图6：直播电商用户渗透率分析5



29.1%

• 2020年8月

• 2019年8月

■注:1.直播电商重度用户指当月提交订 单，5单的用户

2.直播间互动包括点赞、评论、送礼物 及转发等

• 2021年8月

图7：直播电商重度用户直播间互动比例分析6 从2021年中国直播电商行业报告(以下简称“报告”)中可以看出，2019年到2021年直播电商用户的渗透率逐年大幅 提升。深入分析不难发现，电商直播这种形式不断通过各种各样的直播互动形式，加强主播和观众之间的联系，以抖 音为例，根据抖音流量分配的算法逻辑来看，用户留存率和留存时长、互动率、商品点击率、转化率都是直播间流量分 配的重要指标，商家或主播在直播带货时，可以通过吸引眼球的内容增加用户存留，通过派发福利引导用户在直播间 打字、点赞，提升直播间互动，通过秒杀特价等方式提升用户的商品点击率和转化率，以此来增加抖音对直播间的流 量分配。直播电商用户消费习惯正在形成，改善服务体验，提升用户粘性成为最棘手挑战。

疫情的突然爆发，各行各业不得不将生产活动迁移到线上，类似电商直播这种大规模低延时直播应用场景不断涌现。 实时音视频产品的成本和技术问题被再次放大，整个行业都在积极探索有效的解决方案。2021年初，以腾讯云为主的 —些云厂商，通过对开放的WebRTC技术优化，同时融合边缘计算等全新的资源升级对传统的CDN技术架构进行优化 改造，最终演化出来的超低延时直播技术，填补了传统直播CDN和实时音视频产品服务的不足，既能满足低延时体验 需求，性价比也适中，包括：

®播放器兼容性、开放性较好，能够支持跨平台播放需求

®性价比高，能够支持大规模并发需求

®技术特性能够满足低延时流媒体特性和可扩展性的要求

2021年是超低延时直播的发展元年，国内众多知名电商平台逐步将超低延时直播作为精品秒杀、活动派发红包的核心 利器，随时随地都可以发起秒杀活动，摆脱了传统活动需要复杂导播流程的束缚。在演唱会等事件直播的场景下，主 办方利用多机位多直播流方案结合超低延时直播特性实现了多视角无缝切换，现场沉浸感更强。职业教育大班课、 互动游戏等越来越多的应用场景开始大量使用超低延时直播来优化体验和促进业务增长。从多家电商平台公开的数 据来看，超低延时直播在观众观看时长增长和平台交易额提升都有正向促进作用。另外，在QoS优化上，超低延时直 播在核心指标比如渲染卡顿、首帧耗时等指标都超过了传统直播CDN。

以腾讯云为例，腾讯云超低延时直播产品能将直播延时降低到500ms以内，齐备的云端+终端能力能帮助行业客户轻 松接入超低延时直播，实现毫秒级超低延时直播，极大提升主播PK、电商秒杀、在线教育等场景的直播体验。同时，通 过实时数据监控等增值能力，全面掌握实时业务质量，并且通过ABR多码率自适应、极速高清转码等能力提 升客户QoE(Quality of Experience)。目前，已助力上百家客户，实现线上业务新突破，创造更高的商业价值。

直播技术发展襯述

1. -标准直播延时问题仍成为直播技术发展瓶颈
2. -直播流媒体协议向低延时目标不断演进
3. **-WebRTC**技术优势促进超低延时直播快速发展

直播技朮发展概述

近20年来，流媒体直播技术已经从新鲜事物和试验品逐步成为成为一个蓬勃发展的业务，广泛服务于泛互、教育、 广电等各行各业，业务规模呈现井喷式增长。究其原因，随着全球整体基础网络能力建设以及3G、4G、5G移动互联 网能力升级，流媒体传输已经不再依赖传统媒体或较小的专用流媒体网络，大容量、大带宽的互联网正在逐步取而 代之。从承载流媒体内容的传输协议发展进程来看，早在2000年代中期，流媒体行业过渡到HTTP自适应流媒体7 (HAS)，它使用标准的HTTP服务器和TCP(Transmission Control Protocol)[[2]](#footnote-3) [[3]](#footnote-4) [[4]](#footnote-5) [[5]](#footnote-6)来传输内容，使得基于互联网联通能 力的内容传输网络(CDN)能够利用HTTP网络来传输流媒体内容。

❸*标准直播延时问题仍成为直播技术发展瓶颈*

以直播电商为例的直播场景对低延时提出较高要求，电商场景最初在标准直播上实施低延时优化，但随着延时的降 低，卡顿率大幅上升，原因在于标准直播的传输-播放模型中，上行推流和下行播放均采用固定的缓存(buffer)s完全 可靠的传输，其传输和播放控制完全割裂，没有对变化的网络进行适配；固定的缓存(buffer)一旦减小，弱网下的卡 顿率就大幅上涨；其次可靠传输(基于TCP的RTMP、HTTP with FLV、HLS)协议层面无法区分视频帧的优先级，在 网络抖动或变差时，无法满足延时优先的要求。9

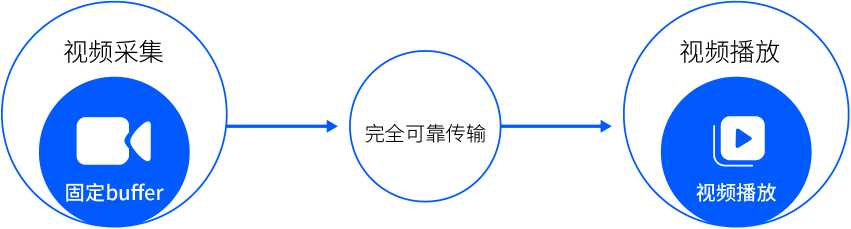
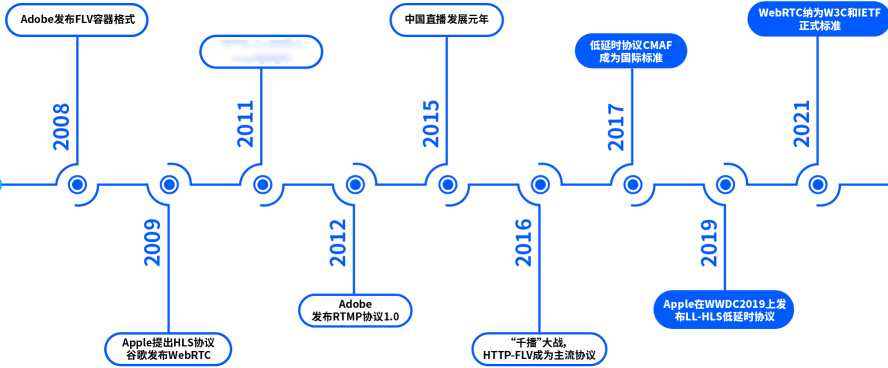


图11:标准直播播放模型

另外，标准直播采用单向完全可靠传输，缺少传输和播放控制的反馈联动，传输不理解流媒体特性，这些都限制了 标准直播的QoS/QoE的优化，比如端云双边加速、Simucast/ABRi。自适应码率策略在标准直播上被限制，当前标 准直播的QoS/QoE的优化已经到达了一个瓶颈。

零 直*播流媒体协议向低延时目标不断演进*

早期的HAS服务标准和应用众多。比如，Adobe公司使用自研的RTMP[[6]](#footnote-7)和HTTP with FLV作为实时流媒体传输 的基础，这种协议为了优化播放体验，一般会在服务侧缓存一个GOP[[7]](#footnote-8) [[8]](#footnote-9)长度的视频帧；与此同时，苹果公司(Apple Inc.)发布他们的HTTP实时流媒体(HLS)格式，这种格式最初建议在播放器中使用10秒的分片时长和保守的3个 分片的起始缓冲区。随着行业技术的发展，这些设置逐渐成为流媒体的标准，被整个行业采用。这种固化的认知，使 直播实际应用的延时保持在2-40秒的范围内。为了降低端到端延时，常规的手段可以通过缩短分片的时间来实现。 然而，这种优化手段不但会降低视觉质量，在片段持续时间只有往返时间(RTT% Round-Trip Time)的几倍场景下, 还会导致CDN的交付基础设施不稳定。面对如何优化延时过大这个难题，整个行业都在积极探索正确方向。随着 WebRTC技术的出现，仿佛见到了曙光，基于WebRTC的低延时直播技术畅想不断涌现，整个直播行业向着更低延 时、更强互动、更高画质的“全真直播”方向发展。



**MPEG-DASH** 技术 成为国际标准

图8：流媒体技术发展历程

流媒体传输技术经过了漫长的发展和迭代演进。如图9所示，近期外部平台行业开发者的调研报告表明，行业开发者

对直播低延时关注度最高，流媒体传输技术最大的趋势是向低延时方向发展。

直播低延迟

成本控制（比如带宽，存储）

41%

33%

32%

全平台支持

数字版权保护

分发**（CDN**分发）

21%

视频质量

21%

质量保证和测试 **17%**

视频互动 **16%**

弱网（限速或掉线）情况**Qos**的保障 **16%**

接入点方案

4%

12%

其他

图9：流媒体技术开发者热点调研14

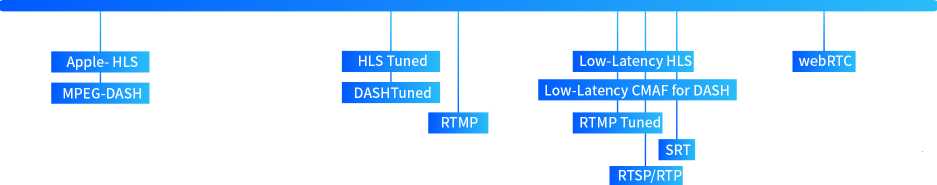
苹果公司（Apple Inc.）在2009年推出HLS,其延迟在10-45秒左右，在2019年又推出了LL-HLS （Low-Latency HLS）,采

用chunk编码传输模式，将延迟降低到3秒左右。目前基于TCP方案的LL-HLS和CMAF（Common Media Application

Format）方案的极限延迟在2-3秒左右，而基于UDP方案的WebRTC可将延迟降低到1秒以内。[[9]](#footnote-10) [[10]](#footnote-11) [[11]](#footnote-12)

|  |  |  |  |
| --- | --- | --- | --- |
| **当前常见的基于 http协议的延迟情况** | **降低延迟** *厂* | **♦**  **、 低延迟** | **・**  **近乎实时** |
| 单向大并发直播事件 线性直播 | 新闻和体育赛事直播 | *Sf）*  **UGC**直播  游戏直播以及体育赛事直播 | 双向网络会议、远程共享、实时设备控制） 例如云台摄像机、无人机） |

45+ I I I I I I 18丨丨S S丨丨丨丨丨05川川川川叫川川lOllllllllllvOl



soconcs

soconds

soconcs

soccnd

soccnd

图10：流媒体技术延迟比较16

如下表所示，对**WebRTC. RTMP**、**LL-HLS. CMAF**进行了对比分析。

■延迟

LL-HLS和CMAF经过优化，延迟有较大的改善，但受限于TCP可靠传输协议及播控模型和chunk大小的限制，其极

限延迟是2秒左右；WebRTC基于UDP传输结合jitterbuffer播控可将延迟降低到1秒以内；

■多平台支持

WebRTC、LL-HLSx CMAF表现相当，均支持多平台；

■流媒体特性

WebRTC表现最好，同时融合网络与编码技术，可理

解音视频类型、帧类型、多流、冗余包和原生包等等；

■ Codec支持

原生WebRTC表现较弱，如视频不支持H.265,音频不

支持AAC (Advanced Audio Coding) 17 等；

■扩展性

WebRTC扩展性较强，可通过扩展头实现自定义的特 性。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **r**  流媒体协议 | **WebRTC** | **RTMP** | **LL-HLS** | 1  **CMAF** |
| 延迟 | <lsec | >2sec | >2.5sec | >2.5sec |
| 多平台 | 高 | 低 | 高 | 高 |
| Codec支持 | 中 | 低 | 高 | 高 |
| 自适应码率 | 高 | 低 | 高 | 高 |
| 流媒体特性 | 高 | 低 | 低 | 低 |
| 扩展性 | 高 | 低 | 低 | 低  *y* |

表1:常见流媒体协议特性对比

。*WebRTC技术优势促进超低延时直播快速发展*

超低延时直播技术摒弃传统直播的传输播控模型，借鉴WebRTC通信模型，将传输和播放控制实时反馈联动，形成反 馈闭环，通过感知网络状态来调整播控缓存策略和传输策略，使传输和播控缓存根据实时网络进行最优匹配，使用 户在变动的网络环境下获取到最优的体验。



图12:超低延时直播播放模型

**WebRTC**的初衷是用于低延迟**P2P(Peer-to-Peer)**通信，在直播场景也面临挑战，具体体现为:

**07**信令流程繁杂，难以满足直播首帧要求。标准 WebRTC 的信令需要经过 SDP(Session Description Protocol)交换,ICE(lnteractive Connectivity Establis hmerit)交互,DTLS(DatagramTransport Layer Security)[[12]](#footnote-13) 握手之后方能传输流媒体数据，导致其首帧耗时严重恶 化。

*03*原生的WebRTC在重传、发包策略上基于P2P 通话和会议，不适合一对多的直播场景。在通话场景，为 保证低延迟，重传较激进，但在高码率、高画质的直播场 景，激进的重传会带来大量的带宽浪费。

*02* 原生WebRTC的音视频编码格式支持受限。标 准WebRTC音频不支持AAC,视频不支持H.265和B 帧，不支持私有数据的透传，不支持可选加密，这些都限 制了在大规模直播场景的应用。

**04**原生WebRTC通过探测带宽，按照带宽容量来调 整编码码率发包，而超大房间的直播，主播编码推流被 多个下行用户复用，这个场景下，单个用户无法反馈给主 播去调整推流编码，因此需要新的解决方案。

针对上述问题，超低延时直播基于**WebRTC**在适配低延迟直播方面进行了大量的优化升级。

**01**信令改造

标准WebRTC的信令交互是一个繁复冗长的过程，不利于直播的快速开播，解决这个问题需要对信令大幅简化。

超低延时直播使用的miniSDP二进制压缩方案(压缩比10%)是一个事实上的标准，它将SDP压缩到一个MTIP之内， 在一个UDP包内完成SDP交互。信令简化压缩后，再通过一定的冗余，使得在50%的丢包下也能保证首帧成功。在此 基础上，进一步提出的0-RTT方案，降低70%的延迟。miniSDP和0-RTT的结合，大幅减少信令耗时、提升信令交互成 功，进而降低首帧耗时和提升开播成功率。

**02**音视频改造

■支持AAC

AAC作为传统直播领域最为常用的音频格式，但WebRTC不支持。超低延时直播基于WebRTC根据RFC6416和 ISO/IEC 14496-3,实现对上述AAC格式的完整支持，以避免WebRTC中Opus音频格式的转码。与此同时，附加其上 的Audio FEC(Forward Error Correction)能够根据网络配置不同的冗余度，使音频能抗50%以上的丢包。

■支持H.265

H.265比H.264有更高的压缩率，在直播领域已经普遍采用，标准WebRTC不支持H.265,为避免H.265到H.264转码带 来的成本和耗时。超低延时直播以RFC7798为基础，实现了对H.265的支持，从而避免转码成本和转码延迟。

■支持B帧

B帧在直播流中广泛存在，其增强了画质，同时大幅减少了码率。标准WebRTC为了减少编解码引入的耗时，不支持B 帧，但超低延时直播在延迟和画质、压缩率、成本之间取得了一个较好的平衡，因此支持B帧。通过客户端SDK配合 CDN,实现B帧的支持。

1. MTU:在计算机网络中，最大传输单元(MTU)是指在单个网络层事务中可以通信的最大协议数据单元(PDU)的大小。MTU与数据链路层可以 传输的最大帧大小有关，但不完全相同，例如以太网帧。

**03**传输改造

・柔性分级传输

标准WebRTC弱网应对策略是通过反馈网络到推流端，使推流端调整码率来适应网络。直播场景，主播和观众是1对 多模型，不能因为某些观众的网络问题去调整主播推流码率。超低延时直播通过服务端和客户端的配合，WebRTC扩 展帧属性和依赖关系，采样柔性分级丢帧的传输策略来渐进式降低码率，以适应弱网情况。

■自适应码率(Simulcast/ABR)

超低延时直播通过扩展RTCP2。作为切流信令，客户端和服务端都具备根据网络来无缝切流的能力，服务端通过渐进 式的超发来探测网络的承载能力，作为切流决策依据，达到快速、精准、无缝切流的目的。

■ P2P分发网络

超低延时直播利用WebRTC原生自带的Peer to Peer的能力，能够将看同一视频流的用户群就近地组织成网络，相互 分享传输，每个客户端节点一边通过RTCP与CDN协商数据，同时与其他客户端节点约定内容共享，在保持低延迟的前 提下依然能够取得不错的效果。

**04**其他改造

・支持私有数据透传以适配业务

众多客户在使用标准直播时,需要通过Meta Data、SEI(Supplemental Enhancement Information)和自定义 NALU(Network Abstraction Layer Unit)等方式传递业务信息；超低延时直播通过RTP?】扩展，支持全链路的私 有数据透传；使得标准直播到超低延时直播的迁移过渡平滑无缝。

■可选加密

标准WebRTC设计应用于音视频通信领域，为保证安全，加密为必选项，而直播部分场景，其内容本身公开，但对性能 和终端消耗敏感，安全性可以适当降低，超低延时直播可根据SDP协商选择开关加密。关闭加密可明显减少前后端开 销,也进一步节省DTLS握手耗时,最终减少首帧耗时。

1. RTCP: RTP控制协议(RTCP)是实时传输协议(RTP)的一个姐妹协议。其基本功能和数据包结构在RFC 3550中定义。RTCP为RTP会话提供 带外的统计和控制信息。它与RTP合作进行多媒体数据的传输和包装，但本身不传输任何媒体数据。
2. RTP:实时传输协议(RTP)是一个在IP网络上传输音频和视频的网络协议。RTP用于涉及流媒体的通信和娱乐系统，如电话、视频电话会议应 用，包括WebRTCs电视服务和基于网络的推送功能。

*超低延时直播 技术架购及性能优化*

1. -超低延时直播技术架构
2. -超低延时直播技术优化

*超低延时直播*

*技术架莉及性能优化*

。超低延时直播技朮架』构

超低延时直播技术架构如下图所示，主要分为上行、加速网络、下行三个部分。



图13：超低延时直播技术架构

上行接入

上行的接入方式多样以针对不同场景，一般而言，分为三类:

■方式

使用WebRTC推流，优点是省去H5播放场景下的Opus/去B帧转码，可做到延迟最优；

■方式二

复用标准直播的上行，采用RTMP/TS OVER SRT22推流，平滑兼容标准直播，无缝迁移；

■方式三

除去RTMP、SRT等常用上行协议以外，一些新兴或者行业相关的上行协议如RIST23、DASH[[13]](#footnote-14) [[14]](#footnote-15) [[15]](#footnote-16)-IF Live Media

Ingest ProtocokGB2828181等也在OTT、安防等领域具备一定的市场。

超低延时直播如果要实现更稳定和更低延迟直播质量，可以在主播端使用**WebRTC**推流。目前常见的三种**WebRTC** 推流接入方案如下：

® OBS WebRTC 插件

OBS是PC端广泛使用的推流软件。提供OBS WebRTC推流插件实现PC端的WebRTC推流是一个 必然的选择。

® 原生WebRTC SDK

谷歌原生WebRTC可在移动Apps端支持推流能力。

® H5 WebRTC 推流 SDK

利用浏览器自身的WebRTC能力，使用H5 WebRTC SDK实现WebRTC推流。

网络加速

超低延时直播加速网络对超低延时直播传输进行了深度优化:

®动态加速 。功能扩展

支持就近接入、智能路由、链路容灾等； 支持H.265、B帧、音频AAC、私有数据透传；

®传输优化

采用深度优化的自研用户态协议栈对传输进行加速：基于网络和码率自适应的Pacing[[16]](#footnote-17).云端双边加速、柔性丢帧、 Simulcast/ABR（低延迟多码率自适应）、O-RTT数据下发结合miniSDP二进制信令压缩、FEC/NACK/RTX融合优 化、首GOP丢帧与重编时间戳等策略深度优化了 WebRTC超低延时直播的体验,QoS远超标准直播。

**® P2P**内容分发

利用WebRTC自带的ICE体系，将用户群就近组网，保证低延迟的体验。

下行提供TAndroid/iOS/Windows/Mac/Linux/H5全终端平台的SDK,采用多层次的接入方案满足用户灵活的接

入需求：

®针对大客户

提供了传输层接入SDK （可通过FFmpeg, ijkplayer 等方式），可最大化减少对原播放器的侵入，复用播 放和业务逻辑，降低接入成本，同时传输层SDK经 过极致精简，客户端APP集成后打包增量不超过 500KB；

。针对中小客户

提供了全功能层接入方式，可最大化接入效率，保障

播放质量。

以腾讯云超低延时直播为例，提供多种接入方式。按端类型区分，可分为Native SDK和H5 SDK接入。同时，Native SDK按功能可分为全功能SDK和传输层SDK,按接入方式分为FFmpeg接入和播放器接入。通过超低延时直播优 质网络链路和SDK的配合优化，超低延时直播的延迟比标准直播下降90%,卡顿下降30%,秒开和标准直播持平。



图14：全平台SDK

*• 超低延时直播技术优化*

超低延时直播从编码、传输、双边加速等多个方面来解决直播所面临的延迟、卡顿、首帧等挑战。

市延时优化

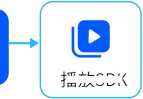
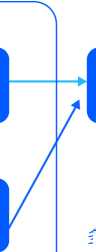
® WebRTC传输改造

WebRTC26具有低延迟相关的流媒体特性、网络反馈与缓存控制机制，双向信令通道等优势，其特性专门用于低延

迟场景的体验优化，使用方便、灵活。超低延时直播首创将WebRTC引入直播领域，作为方案的基础。

超低延时直播对新客户改造上行、下行传输与缓存播放部分，对老客户可仅改造下行传输与缓存播放部分。下行釆

用分级QoS、UDP传输协议、动态jitterbuffer,根据网络自适应调节jitterbuffer及QoS策略，将整体延迟控制在

500ms左右，同时优化起播和卡顿。

多种接入方式以平滑迁移

*丿*

**RTMP**

**推流**

**WebRTC**

**推流**

播放SDK

**WebRTC**

**拉流**

推流SDK

OBS

图15：超低延时直播接入架构

®首**GOP**丢帧、重编时间戳

直播场景受GOP长度的影响，起播时缓存1个GOP长度 的数据，导致起播时必须快进，否则会导致1个GOP （约 2秒）的延迟；但一方面快进追帧倍率大易影响播放体验； 另一方面，追帧的倍率太小，会导致延迟持续时间长，也 会影响体验。

超低延时直播通过首GOP丢帧与时间戳重编，使得播放 侧免去追帧快进，大幅改善播放体验。

视频的处理 当前收帧点

Will

丁〒一I一

回退 收帧驱动发送 被丢弃的视频帧

图16：超低延时直播起播策略优化

26.在众多低延迟视频传输协议中,WebRTC是唯——个被W3C、IETF同时纳为正式标准的协议，其生态已经非常丰富，超过93%的浏览器支持 WebRTCo

、3.2.2卡顿优Q

。自适应**Pacing**降卡顿

平滑发包**（Pacing）**采用**2**级调整，分为**GOP**级别的一级调整和以帧为基础的二级调整:

■周期性（GOP级别）初平滑

以GOP的码率为基准才能做到最大限度的平滑，能 最大限度减少对网络的冲击，但保证不了帧级别的 时效性。

■考虑时效性（帧级别）的二次平滑

初平滑能最大限度平滑发包，若GOP级别平滑超过 单帧最大发送时间，则应该以单帧发送时间为窗口 进行二次平滑，以单帧粒度来保证时效性。

®端云双边加速

超低延时直播根据视频帧的重要性等级**（1**帧＞ 卩帧＞ **B**帧）实现分级重传，可提升重传效率，能抗**30%+**的丢包率。

■视频内容分级

根据重要性将音视频帧分级，分别为：音频＞1帧＞P帧＞B帧（可能有多级）。服务端RTP包扩展帧属性+ seq range,客户端在未收全帧的时候即知道帧属性及依赖关系。

・重传优化

原生WebRTC客户端在特定情况下，会产生大量重复的NACK请求，这会导致重传风暴。重传风暴会产生较大的发 包毛刺，并降低重传效率。

超低延时超低延时直播对此场景进行了优化：首先超低延时直播考虑了网络质量、帧类型优先级对重传进行了优 化：如在首帧画面渲染之前，若同时有I、P、B帧的NACK,则优先响应I帧的NACK,待确认I帧收全后，再响应其他帧 的 NACK。

I 帧 NACK P 帧 NACK B 帧 NACK

--Illi I 而 IIIII 而 III

I I

I

无效重复NACK

图17：超低延时直播选择重传 其次，超低延时直播考虑重传延时和时效性，依据时效性指标过滤重传队列。如下图所示，新的GOP已经到来，前 序GOP仍然有大量过期的重传，超低延时直播根据网络质量选择丢弃过时的NACKo

低优先级NACK 1帧NACK

IIIIIIIIIIIIIIIIIIIH

前**GOP**未收全 新**GOP**已达到

I I

I

过时NACK

图18：超低延时直播依据优先级重传

■网络自适应分级Qos

对于实际带宽容量低于流码率的场景，无论如何平滑，都会导致丢包。此类问题需服务端-客户端SDK配合，对视频 内容进行分级QoS保障。

网络质量根据RTT、丢包率、重传率对网络的影响权重，设置影响因子，对其进行归一化评级如下：27

net\_score = l/exp(kl\*rtt/rtt\_max+ k2\*loss\_rate + k3\*retrans\_rate) \* 100

评估完网络质量，则可以根据不同质量，执行不同的网络分发策略。下表是一个可能的选择:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 网络评级 | 发包及重传策略 | | |  |
| **Net bcore** | **B** 帧**(P3)** | **P** 帧**(P2)** | **1** 帧**(P1)** | 音频**(P0)** |
| 90-100 | 优秀 | 正常发包 正常重传 | 正常发包 正常重传 | 正常发包  重传2次 | 正常发包 正常重传 加 40%fec |
| 80-90 | 良 | 正常发包 不重传B帧 | 正常发包 正常重传 | 正常发包 正常重传 |
| 70-80 | 中 | 丢B帧 | 尾部丢5%P帧 正常重传 |
| 60-0 | 差 | 尾部丢n%P帧 (n负相关于 网络评级)  减少重传带宽 |

表2：网络自适应发包策略

27.采用类sigmod函数对影响因子进行归一化，影响因素的权重不同：RTT权重最小，因一般在数据路由过程中是有buffer存在，RTT抖动正常， 且RTT增加一般先于重传丢包，仅是RTT少量增加和抖动，被认为是正常的；其次是重传率，其表示网络有乱序或丢包，但重传可能恢复；最后是丢 包率，影响权重最大；此处选择重传后的丢包率，反映的是网络丢包修复能力，此值较大，表明很多包重传不成功。

® Simulcast/ABR （自适应码率优化弱网卡顿）

Simulcast是一种WebRTC客户端以不同的分辨率和比特率对同一视频流进行多次编码，并将其发送到路由器，然 后由路由器决定谁接收哪一个视频流。Simulcast与SVC（Scalable Video Coding）密切相关，在SVC中，单个编码的 视频流可以分层，每个参与者只接收他能处理的层。

虽然可以简单地在流媒体到达时转发，但现实世界的网络条件带来了一些挑战。具体来说，不是每个人都一直能有 足够快的网络连接来接收发布的流媒体。为了确保流畅和高质量的传输，我们有几个选择。

■使用一个可扩展的编解码器，如**VP9**或**AV1**

・向参与者发送单独的流，在网络变动时候切换不同的码率或者分辨率，以满足每个用户的实时可用带宽

Simulcast允许服务侧发布同一源轨道的多个版本（一般为高、中、低分辨率，不同的比特率编码版本），并采用不 同的编码方式（即空间层），实时检测用户的带宽和CPU能力，并相应调整媒体流的质量。

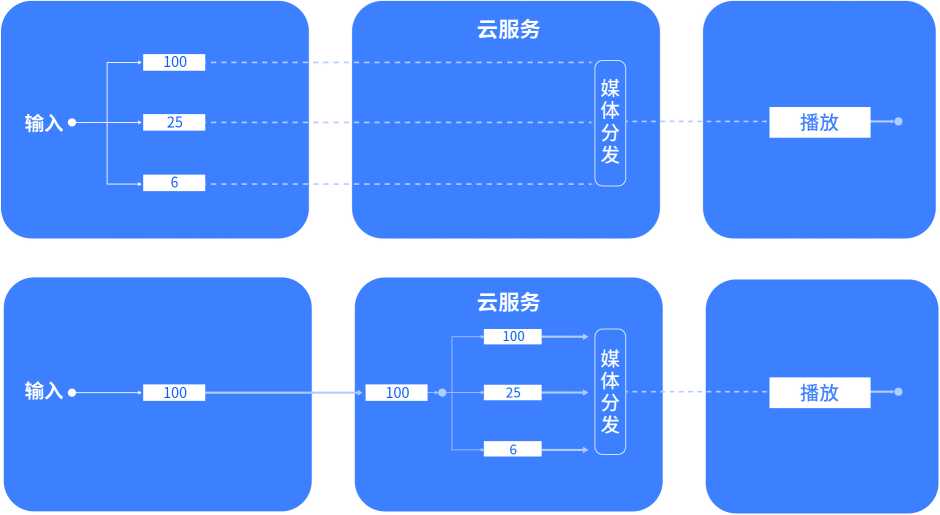
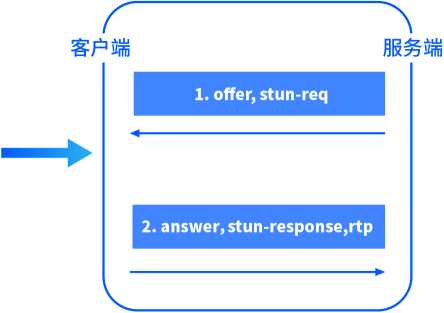


图19：端到端Simulcast与基于服务侧的ABR

J3.2.3*首帧优死［*

。信令简化**-O-RTT**数据下发

原生WebRTC信令交互流程较长（需要6-7个RTT）,通过O-RTT信令方案精简交互流程，可减少60%的信令耗时。



土业握手

**2. sdp-offer**

**3. sdp-answer**

**4. stun-req**

**5. stun-response**

**6. dtls**握手

**7. rtp/rtcp**

图20：信令简化

® 信令压缩与冗余

SDP为文本信息，体积较大，需要几个RTT才能传完。通过信令二进制极致压缩，节省95%的体积，只需一个UDP包即 可装载；同时，结合信令冗余传输，在节省传输耗时的同时,保证信令的可靠性。

®视频传输的初始窗口控制



播放器播放时，需要积累一定的数据才开始起播。如某 播放器积累700ms数据才开始起播，而TCP协议，需要

协商初始窗口并经历慢启动才能达到带宽容量上限，导 致需要多个RTT才增大到带宽上限。

理想状态下，直接使用可用带宽容量来发包，这样首帧 最快，但在首帧发送的时候，丢包率、RTT、带宽容量等 信息未知,则需要200ms以上才能获取到。这使得开播阶 段并未充分使用带宽，并直接影响了起播的速度。

超低延时直播直接跳过TCP的慢启动阶段，以历史带 宽容量作为参考，通过SDK和服务端配合，在起播阶段 加速RTCP反馈，以大窗口快速适配带宽容量，达到降 低首帧耗时的目标。

图21： Fb/与WebRTC起播性能开销对比

324媒体处理优化源

以AV1为代表的新一代编码技术，也为超低延时直播带了更多的可能以及更好的体验。在WebRTC中采用AVI的动因 可以简单归结于：

■更高的压缩效率降低了带宽消耗，提高了视觉质量

・使低带宽网络上的用户能够访问视频流

■与**RTC**场景具备天然协同性,包括对时域可伸缩性**(TemporalScalability)**、屏幕内容编码**(Screen Content Coding, see)**等工具

除去新一代编码器带来的极致的压缩能力，感知编码、细节增强、视频去噪等前处理也随着服务侧算力的提升和算法 本身的进化，逐步在超低延时场景发挥作用。结合AI的能力，在编码前对视频源进行前置处理，使视频源更适配编码 场景。其中，前置处理基于预训练模型，消除视频源附带的毛边细节，使整体边缘更加平滑，更有利于视频压缩。另 外，从人眼主观的角度来处理视频，进而使视频更容易被压缩，同时低码率下也能保持一个良好的观感，结果是更低 的码率、更好的画面质量、更低的分发带宽和更好的观看体验。除去极致压缩，画质重生技术也被用于视频画质的提 升，以视频超分、智能插帧、色彩增强等能力增强源的画质。除去视频处理之外，音频降噪、去混响、响度管理、立体声 等音频相关能力，也为超低延时直播的快速发展带来更好的视听体验。

*超低延时直播*

*推动多场景应用发展与迭代*

1. -在线教育
2. -电商带货
3. -赛事直播
4. -会展直播

*超低延时直播*

推动多场*景应用发展与迭代*

直播的魅力在于即时和多向的互动，尤其涉及体育赛事、电商直播、在线教育等特定场景，延迟会极大地影响用户

观看和互动体验。试想，一场精彩紧张的足球赛的跌宕起伏因为延迟，早已在隔壁此起彼伏的欢呼声失去了悬念， 这是多么令人沮丧的体验。

超低延时直播能够带给用户实时互动体验的核心在于低延迟。基于WebRTC的低延迟技术，在含有“互动”的直播 场景中已经突破到了400毫秒以内的延迟边界。超低延时直播能够还原真实互动场景，让在线教育、电商带货、跨域 联动这类直播场景拥有更低延时更强互动的观影体验。



场景需求

在传统教育直播中，通常分为大班课或小班课，小班课通常使用实时音视频互动产品，进行小范围的音视频

互动，大班课面向于几十人甚至上百万人同时在线的要求。

业务痛点

在面向大房间高并发要求时，实时音视频产品无论是技术架构还是使用成本都无法满足用户的要求，比如 在连麦和下麦切换时常会出现延时较高的情况，影响讲师与学生的互动。并且还会出现当实时音视频流量 成本较高时，很多在线教育平台会放弃部分大班课规划。

接入效果

腾讯课堂作为理财实践线上教学模式的开创者，选择将腾讯云快直播技术融入其直播课堂，用毫秒级的延 时来保障老师和学生之间畅通无阻的沟通，真实还原线下课堂的互动体验，再辅以答题、投票、白板等互动 功能模块，使线上课堂的教学质量和学员完课率都得到大幅度提升。



场景需求

对于电商直播来说，场景的核心便是带货主播与观众之间的互动。早期传统的电商直播还是继承了传统直 播单向讲解产品的枯燥形式，互动往往依赖于简单的文字聊天，观众的响应和反馈延迟，用户问题得不到及 时解答，用户流失率较高，平台成交量和整体直播电商的规模都相对较小。为了解决用户粘性以及付费意愿 等问题，众多平台都在探索更好地加强互动，通过更有效的互动体验，增加用户播放时长和平台成交量。比 如2018年以后，众多平台通过直播在线答题的方式派发现金或优惠券方式增强平台粘性，通过趣味性游戏 的方式和观众增加互动。

业务痛点

为了满足公平性要求，这些互动形式对直播延迟要求较高（Is左右），传统的直播技术和架构问题突出，众 多云厂商都在探索低延迟直播技术，绝部分方案都是基于传统的直播技术进行CDN配置或播控策略优化 来尽可能优化延迟，总体效果不尽如人意。

接入效果

近年来，随着部分云厂商低延迟直播技术能力突破性发展，都在低延迟这个赛道实现了本质突破。以腾讯 云为例，通过腾讯云快直播产品，不需要编排和时间规划，随时随地可以发起红包或福利派发，极大的促进 了观众交互和购买热情，大大提升了平台的交易量。以国内某头部电商平台举例，当前已大规模应用低延迟 直播技术，让主播更快的接收到观众的信息并实时反馈，提升消费者边看边买的体验，提升了用户活跃度和 商品成交率。



场景需求

活动赛事对于直播播放的延时、流畅以及画面清晰度都有比较高的要求。比如游戏赛事、体育赛事等，对赛 事结果的时效性非常敏感，精彩瞬间如果不能实时投放给观众，那么对于平台的体验反馈将大打折扣。

业务痛点

活动赛事中，传统直播架构引发的延迟大问题一直被观众所诟病。早期包括英雄联盟总决赛、英超直播、 世界杯等重大赛事，众多平外都选用了HLS协议用于平台直播分发，整体延迟在10-30S左右。为了解决这种 超大规模并发低延迟问题，各大平台尝试过各种优化手段，比如通过HTTP-FLV替换HLS协议，进一步将延 迟降低到3-5S。然而，由于FLV协议的安全问题，导致此种分发方式在海外不被广大CDN厂商所接受。

接入效果

2019年，腾讯云率先推出了完整产品化能力的超低延时直播产品，腾讯自有游戏直播平台抢先接入快直播 能力，在2021英雄联盟全球总决赛(S11),首先让国内的观众体验了超低延时观看体验，提前为EDG夺 冠而喝彩。企鹅电竞作为一站式赛事解决方案的供应商，采用了腾讯云快直播产品，以最快速度将精彩赛 事传递给用户，让线上观众能快人一步获取赛事精彩瞬间，全量提升用户的沉浸式参与感。



场景需求

公司跨域联动的企业直播活动都对直播场景有着更低延迟的要求。比如某些庆典活动具备多个分会场，需 要釆集多个会场的画面推多路流到云端，或者大型演唱会多视角镜头如何实现多机位的无缝切换。

业务痛点

早期的方案通常使用较为复杂的导播台方案，为了解决同步问题，往往还需要在流里面插入同步依赖信息， 切换画面的复杂度较高和效果较差。极大的限制了跨区联动直播的大面积推广和应用。

接入效果

国内某企业直播服务商，以知识产品与用户服务为核心，利用腾讯云快直播产品为其平台客户实现了多场跨 地域多会场联动的公司年会、企业直播活动，消除了各地观众们之间的地域限制，达到异地实时联动的强 观感现场效果。

超低延时直播 未来发展趋势

超*低延时直播未来发展趋势*

超低延时直播技术架构不断升级，有力支撑直播与各行业深度融合。伴随消费互联网向产业互联网升级的浪潮，信 息传递的形式也发生了翻天覆地的变化一早期的图文信息从贴吧、论坛到微博，视频信息从下载、点播再到直播 ——从演变的过程看，不难发现，整体的互动性不断增强。业务形态对互动性的依赖，促使直播技术不断演进升级， 通过协议优化链路传输延时的同时，结合流媒体本身的技术升级包括对新的编解码协议H.265、AVI、H.266的应用， 完成了整体的技术架构升级，直播服务具备了较强互动性的能力。然而在面对大规模强互动直播的应用时，高昂的 流量成本阻塞了超低延时直播在部分场景的大规模应用。近几年的时间里，通过持续的技术架构的升级以及边缘 计算的应用逐步使得超低延时直播大规模应用成为了现实。

超低延时直播场景持续探索，新业态、新模式、新价值不断涌现。例如在电商直播行业，随时发起的低延时秒杀活 动赋予了卖货更强的互动娱乐性，拉动了用户观看时长以及交易额增长；例如在企业直播行业，多机位镜头结合超 低延时直播技术应用，使得用户摆脱了导播台的束缚，能够无缝切换现场多视角镜头，现场沉浸式体验更佳。类似 的应用场景不断被各行各业发掘，一些依赖于线下或专业设备的特殊应用逐步在超低延时直播技术的助力下迁移 到线上。在不久的将来，超低延时直播的技术应用会释放直播行业更多的商业价值，助力更多产业实现业务变现和 增长。

■•考文档

© Y.-K. Wang,Y. SanchezJ. Schierl,S. Wenger and M. M. Hannuksela, RTP Payload Format for High

Efficiency Video Coding (HEVC), IETF RFC 7798, March 2016;

**<https://www.rfc-editor.org/rfc/rfc7798.txt>.**

® M. SchmidtF. de Bont,S. Doehla,J. Kim, RTP Payload Format for MPEG-4 Audio/Visual Streams, IETF RFC

6416, October 2011;

**<https://www.rfc-editor.org/rfc/rfc6416.txt>**

® WebRTC 1.0: Real-Time Communication Between Browsers,

**<https://www.w3.org/TR/2021/REC-webrtc-20210126/>**

® Bitmovin's Video Developer Report 2021

**https:**〃**f.hu bspotusercontent30.net/hubfs/3411032/Premium%20Content%20PDF%20Files%20 %20whitepaper,%20case%20studv,%20report,%20/Video%20Developer%20Reports/Bitmovin- Video-Developer-Report-2021.pdf**

® DASH-IF Live Media Ingest Protocol,

**<https://dashif-documents.azurewebsites.net/lngest/master/DASH-IF-lngest.html>**

4. SDK：软件开发工具包(SDK)是一个可安装的软件包中的软件开发工具的集合。

17. AAC:高级音频编码(AAC)是一种用于有损数字音频压缩的音频编码标准。AAC被设计为MP3格式的继承者，在相同的比特率下，AAC通常比 MP3编码器取得更高的音质。

1. 1. 数据来源：《2021年中国直播电商行业报告》
   2. 数据来源：《2021年中国直播电商行业报告》

   [↑](#footnote-ref-2)
2. 自适应流媒体：在自适应流媒体中，有多个轨道可供选择，它们以不同的比特率呈现相同的媒体。在播放过程中，使用ABR算法动态地选择所选 的轨道。 [↑](#footnote-ref-3)
3. TCP：传输控制协议(TCP)是互联网协议套件的主要协议之一。它起源于最初的网络实施，它是对互联网协议(IP)的补充。 [↑](#footnote-ref-4)
4. 标准直播基于完全可靠的TCP,首先基于TCP的传输要求接收方确认所有收到的数据包，以便发送方能够重新发送任何丢失的数据，这会带来 确认延时；其次，基于TCP的传输有窗口确认机制，在弱网场景下,形成数据积压，导致延时大幅增加。另外，基于TCP的传输采用的是单边加速，没 有反馈应用层信息的信令通道，不理解音视频信息，其拥塞控制可优化的空间受限。 [↑](#footnote-ref-5)
5. ABR：自适应比特率。ABR算法是一种在播放过程中在若干轨道之间进行选择的算法，其中每个轨道呈现相同的媒体，但比特率不同。 [↑](#footnote-ref-6)
6. RTMP：实时信息传输协议(RTMP)是一个用于在互联网上传输音频、视频和数据的通信协议。最初是由Macromedia开发的专有协议，用于 Flash Player和服务器之间的流媒体,Adobe收购Macromedia后发布了一个不完整的协议规范版本供公众使用。 [↑](#footnote-ref-7)
7. GOP：在视频编码中，一组图片，或称GOP结构，规定了帧内和帧间的排列顺序oGOP是一个编码视频流中连续图片的集合。每个编码视频流 由连续的GOP组成，可见的帧就是从这些GOP中产生的。在压缩视频流中遇到一个新的GOP意味着解码器不需要之前的任何帧来解码下一个帧， 并允许在视频中快速寻找。 [↑](#footnote-ref-8)
8. RTT：往返时间(RTT)是指发送一个信号所需的时间加上收到该信号的确认时间。这个时间延时包括两个通信端点之间路径的传播时间。在计 算机网络中，信号通常是一个数据包。 [↑](#footnote-ref-9)
9. 数据来源：from <https://bitmovin.com/> [↑](#footnote-ref-10)
10. 基于TCP的传输要求接收方确认所有收到的数据包，以便发送方能够重新发送任何丢失的数据。然而，对于实时视频流来说，TCP/IP的通信效 率很低，特别是在长距离上。太多的接收确认充斥着连接，极大地降低了带宽效率。工作流程中每个发送方和接收方之间建立的缓冲区（例如每个路 由器）也会带来巨大的传输延迟。这就是为什么基于TCP的协议（如RTMP.HLS和MPEG-DASH）不适合低延退直播要求。 [↑](#footnote-ref-11)
11. 图片来源：wowza.com [↑](#footnote-ref-12)
12. DTLS：数据报传输层安全(DTLS)是一个通信协议，通过允许它们以一种旨在防止窃听、篡改或信息伪造的方式进行通信，为基于数据报的应 用提供安全保障。 [↑](#footnote-ref-13)
13. SRT: Secure Reliable Transport (SRT)是一个开源的视频传输协议，利用UDP传输协议。 [↑](#footnote-ref-14)
14. RIST:可靠互联网流传输(RIST)是一个开源的、开放规范的传输协议，旨在通过有损网络(包括互联网)以低延迟和高质量可靠地传输视频。它 目前正在视频服务论坛的"RIST活动小组”中进行开发。 [↑](#footnote-ref-15)
15. DASH: HTTP上的动态自适应流。一个行业驱动的自适应流媒体协议。它由ISO/IEC23009定义，可在ISO公开可用标准页面上找到。 [↑](#footnote-ref-16)
16. Pacing：在计算机网络领域.Pacing是一套技术的名称,用于使传输协议产生的数据包传输不那么突兀。 [↑](#footnote-ref-17)