

实时音视频

产品介绍

文档版本 01
发布日期 2025-01-21



版权所有 © 华为云计算技术有限公司 2025。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为云计算技术有限公司

地址：贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

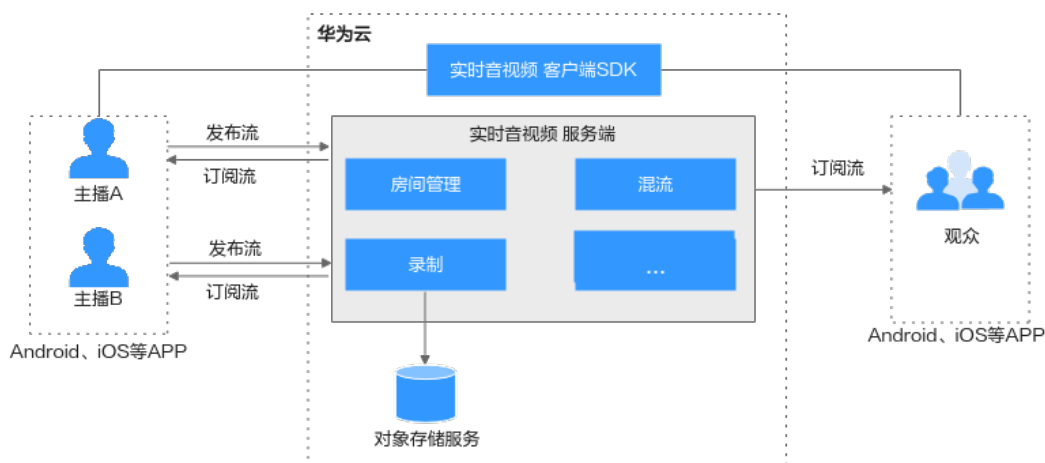
网址：<https://www.huaweicloud.com/>

目录

1 什么是实时音视频?	1
2 产品优势.....	3
3 应用场景.....	4
4 功能特性.....	7
5 使用限制.....	9
6 常用概念.....	10

1 什么是实时音视频？

华为云实时音视频服务（SparkRTC）凭借在视频业务领域长期技术积累，快速为行业提供高并发、低延迟、高清流畅、安全可靠的全场景、全互动、全实时的音视频服务，适用于在线教育、云会议、社交文娱等场景。



- 实时音视频 客户端SDK：提供Web等平台的SDK，方便用户快速集成开发音视频通话、互动直播等功能。
- 实时音视频 服务端：接受连麦请求、指示连麦用户加入到实时音视频流的房间中，支持将主播端发布的流进行混流、录制等实时音视频处理。

访问方式

华为云实时音视频服务提供了Web化的服务管理平台，即管理控制台和基于HTTPS请求的API（Application programming interface）管理方式。您可以通过以下方式访问实时音视频服务：

- 控制台方式
用户可直接登录管理控制台访问实时音视频服务。
 - 如果用户已注册账户，可直接登录管理控制台，从主页选择“视频 > 实时音视频”。
 - 如果未注册，请[单击此处](#)注册华为账号并开通华为云。
- API方式

如果用户需要将华为云实时音视频服务集成到第三方系统，用于二次开发，请使用API方式访问华为云实时音视频服务，具体操作请参见《[服务端API参考](#)》。

2 产品优势

全球实时音视频网络

提供覆盖全球的高质量、大规模的实时音视频网络。自研高效调度算法，具有全网调度能力。丰富的节点资源储备，保证端到端平均时延小于200ms。

低卡顿

基于华为30年音视频编解码能力和优异弱网对抗能力80%丢包下音频通话流畅，50%丢包下视频通话流畅。

高品质音频

稳健的3A算法、智能降噪、回声消除和智能啸叫抑制，48Khz采样，提供一流音质体验。

高品质视频

1080P视频分辨率、H265、感知编码技术，码率降低30%~40%。SCC屏幕共享编码，图文色彩高保真。

稳定可靠

全年SLA服务可用性高达99.99%。全链路安全端到端加密，全球最佳安全和隐私保护。

3 应用场景

在线教育

丰富的一对一教学、小班课、大班课等多种解决方案，轻松构建高品质互动教学体验。高清、低时延、高并发直播，易集成，可快速部署到在线教育视频平台。



文娱互动

零距离互动新玩法，助力业务创新、流量变现。适用于网红直播、企业直播、娱乐直播、游戏直播等场景，提供功能强大的实时媒体处理能力，实现各种直播场景玩法，为客户和合作伙伴打造端到端一站式“音视频互动”解决方案。



电商直播

安全可靠的实时音视频服务，提升互动性和电商客户的转化率，有效助力电商的直播业务发展。



视频会议

视频会议互动，远程协作办公，跨平台全球互通，协同无压力。适用于低时延视频会议场景，提供稳定流畅的低成本高清视频会议服务，具有超强抗丢包、低时延能力。



4 功能特性

华为云SparkRTC提供了视频通话、连麦互动等功能，如表4-1所示。

表 4-1 功能列表

功能	说明	常见场景
视频通话	即1对1或多人视频通话，支持720P及以上高清画质。单个房间最多支持2000人同时在线，最高支持500人互动。	1对1视频通话、视频会议、在线问诊、多人视频聊天、视频客服、视频双录、在线理赔等。
语音通话	即1对1或多人语音通话，单个房间最多支持2000人同时在线，最高支持500人同时发言。	适用于1对1语音通话、多人语音通话、语音聊天、语音会议、语音客服等。
视频互动直播	支持视频连麦、跨直播间PK。	超低延时直播、超级大课堂、主播PK、远程培训、大型会议等。
语音互动直播	支持语音连麦、跨直播间PK。	语音低延时直播、语音直播连麦、K歌、FM电台等。
高音质	支持48kHz采样的高音质。	语音通话、视频通话、互动直播、高音质FM、音乐教学课、K歌房、在线课堂等。
高画质	支持720P、1080P的高清画质视频。	视频通话、互动直播、在线课堂等。
角色切换	支持房间内角色切换，观众自由、平滑上下麦，切换过程无需等待。	互动直播、在线课堂等。
单用户加入多房间	支持单个主播加入多个房间互动PK。	秀场直播、PK连麦、超级大班课等。
屏幕分享	支持将本地电脑桌面、窗口、画面区域分享给他人。	互动课堂、视频会议、远程协助等。

功能	说明	常见场景
3A处理	行业领先的3A算法，在降噪等场景下提供更好的声音质量。3A即AEC（回声消除）、ANS（自动噪声抑制）、AGC（自动增益控制）。	所有语音场景。
音量大小回调	提供音量大小的数值，方便显示成波形动画或提示。	语音通话、视频通话、语聊房、FM电台、K歌房、人声检测等。
跨房连麦	支持主播跨房间连麦互动，观众可自由通过角色切换进行上下麦。	在线教育-子母课、连麦PK。
互动连麦	支持连麦互动，观众可自由通过角色切换进行上下麦，切换过程平滑。	在线教育-小班课、互动直播
云端录制	提供云端录制功能，支持单流/混流的模式将录制下来的文件存储到OBS或者点播服务中。	存储、录制审核、视频双录等。
文件混音	支持将本地或在线的音频和用户声音混音，发送给房间内其他用户。	在线教育、语聊房、在线合唱等。
旁路直播	将房间内的音视频流混流转码后推送RTMP的视频流给直播服务。	互动直播、大型会议等。
网络检测	支持通过网络检测接口判断或预测用户当前的网络状况是否良好。	-
自定义视频数据	支持自定义的视频源和渲染器，使用非摄像头的视频源，例如视频文件、外接设备、第三方定制数据源等。	自定义美颜、定制数据源、多设备管理、视频识别、图像处理等。
自定义音频数据	支持自己采集音频回调，开发者可以对原始数据进行处理，进行自定义操作，例如外接非标设备、音频文件等。	非标设备接入、自定义音频效果、语音处理、语音识别等。
平台兼容	支持Web终端平台，系统要求请参见 使用限制 。	-

5 使用限制

在使用实时音视频前，您需要了解本服务的使用限制。

表 5-1 使用限制

限制项	说明
应用创建	若首次使用SparkRTC，则需 提交工单 联系华为云技术客服进行咨询。
直播间人数	一个直播间最多支持500人同时通话。
Web SDK	<ul style="list-style-type: none">支持Chrome、Firefox、Safari和Opera等主流浏览器，详见Web浏览器适配。不支持用于小程序集成开发。

6 常用概念

应用标识

应用标识或应用ID，简称appID，是SparkRTC用来区分不同应用的唯一标识，在SparkRTC控制台中创建应用时自动生成。

房间

房间是一个音视频空间，同一房间内的用户可以互相接收对方的实时音视频数据。

- SparkRTC使用房间这个虚拟的概念，用于用户之间的相互隔离。
- 只有在同一个房间里的用户才可以相互接收音视频。
- 房间ID的命名规则：可自定义，支持64个字符的大小写字母、数字、下划线（_）中划线（-）随机组合。

用户 ID

简称userId，是指SparkRTC用来区分应用中用户的唯一标识。

- 用户标识是用户在SparkRTC中的一个表示，开发者按照规范定义UserID。
- 命名规则：可自定义，支持64个字符的大小写字母、数字、下划线（_）中划线（-）随机组合。

用户角色

是指用户在房间内的不同角色类型，不同角色类型有不同的权限模型。主要有如下三种角色类型：

- 主播（publisher）：只发流不收流主播型角色。SparkRTC预留的角色类型。
- 互动观众（joiner）：既能发流也能收流的互动型角色。
- 普通观众（player）：只收流的观看型角色。

SparkRTC示例Demo中的角色切换、上台/下台、上麦/下麦主要指在joiner/player两个角色之间进行切换。

主流

主要指摄像头采集的视频流，SparkRTC对同一视频源支持最多同时编码和收发四挡数据类型（720P/360P/180P/90P）。

辅流

也称共享流，指共享屏幕或窗口的视频流。默认1080P，开启辅流流畅优先指切换为720P。

混流

是指将发布的多路音视频流混合转码成一个单流。

大小流模式

也称新模式，指编码和收发两路视频流的模式（默认开启大流720P，小流360P选择性开启）。

订阅

订阅是指房间中的用户接收房间内远端用户音视频流的操作。

自采集、自渲染

SparkRTC默认使用系统音视频设备采集和渲染数据，也支持自定义采集和渲染，常见场景如录屏播放，第三方美颜特效，云游戏等。

丢包

丢包是指实时音视频数据传输过程中发生的数据包丢失的现象。

抖动

实时音视频通信中，连续传输的数据包之间的延时不一致称为抖动。

卡顿

卡顿是实时音视频传输过程中，因网络条件、设备性能受限等原因，引起的音频或视频播放断续、不流畅、甚至定格等现象。

单流录制

是指对房间中每个用户的音视频流进行录制，每个用户都有其对应的音视频录制文件。

合流录制

是指将房间中多个用户的音视频混合录制为一个音视频文件。