

概述

ZEGO 语音通话（AudioRoom）是一款实时的语音通话产品，能够为开发者提供便捷接入、流畅高音质、多平台互通、低延迟、高并发的音频直播服务，可以实现一对多，多对多的实时音频互动，语音聊天，直播K房等场景。

1 主要功能和场景

类别	主要功能	功能描述	典型对标场景
基础功能	直播推拉流	支持RTMP、FLV、HLS和RTP等协议，全球节点实现超低100ms延迟，同时提供多CDN推流，自适应流控等功能。	语音直播
本地音频录制	支持通过 SDK 获取到录制的音频数据，可另外进行本地存储或者云端存储	语音记录保留	
语音直播回放保留	音效变声	支持音调，双声道设置和回声处理，为开发者提供变声，混响，场音，立体声和环绕音等功能。	虚拟立体声
立体环绕音			
语音直播	常用功能	媒体次要信息	支持在主播端和观众端传送音频流数据的同时，同步一些其他信息。
K歌合唱	歌词同步	媒体播放器	支持通过该功能实现在直播过程中，在推流中混入一个音频文件，实现更多直播的应用场景。
直播背景音乐	远程业务办理中插入语音播放	自定义音频前处理	支持对 SDK 采集的原始音频数据进行自定义处理，实现变声等功能。
聊天变声	语速调节	自定义音频采集与渲染	提供了自定义外部音频采集与渲染功能，支持开发者更加灵活的实现采集/渲染相关的自定义需求。
语音直播	多路混流	支持将多路音频流混合成单流。	直播连麦
直播PK			

多主播互动

KTV 合唱

在线教育 ||| 音频加解密 | 支持加解密推拉流的音频，保障音频的安全性。|

远程业务记录安全加密

私人通讯安全加密 || 音频引擎 | 回声消除 | 引擎中自带回声消除功能，在双

讲情况下，各方语音清晰，双工体验良好，支持舒适，适度和激进三种回声消

除模式。| 语音直播 ||| 噪音抑制 | 引擎中自带噪音抑制功能，它结合心理声

学模型，能够提高20dB+的信噪比，同时不损伤语音音质。| 语音直播 |||

自动增益 | 引擎中自带自动增益功能，能够自动调节麦克风音量，适应远近拾

音，在嘈杂环境下带来良好的体验。| 语音直播 ||| 多种编码 | 支持多种硬

编、多种软编，以适应多种应用场景和网络环境的编码要求。| 在语音直播 ||

| 多种解码 | 支持多种硬解、多种软解，以适应多种应用场景和网络环境的解

码要求。| 语音直播 ||| GPU处理 | 数据全部都走GPU显存，性能好，稳定

性好，延迟低。| 语音直播 || 网络传输 | 网络传输Qos | 实现冗余数据(FEC)

和重传策略(ARQ)的动态平衡，一方面保障宽带的充分利用，另一方面不至于

抢带宽造成链路自身拥塞。| 全场景 ||| 网络自适应 | 提供端到端的全链条优

化算法，能根据当前网络情况预测网速并自动进行流控，支持弱网自适应推流

和拉流。| 全场景 ||| 抖动缓冲 | 在弱网环境下，利用良好的抖动缓冲策略找

到延时与流畅的最佳契合点。| 全场景 |

2 特性指标

特性	实时语音指标
SDK 包体积	iOS: 10.8M Android: 11.4M Windows: 7.55M

多人音频	支持 32 路实时音频互动
音频质量	音频采样率： 16k~48k 支持单、双声道 回声消除：支持 10 路同时讲话
海量并发	全球 200+BGP 节点，提供稳定的跨国内容分发网络

3 平台兼容

语音通话支持 **iOS**、**Android**、**Windows**，并支持平台间互通，具体的兼容性要求见下表。

平台	支持版本	支持架构
iOS	7.0+	arm64 armv7 模拟器（x86_64）
Android	4.0.3+	armeabi-v7a arm64-v8a x86 模拟器（x86）

Windows	win7+	x86 x64
---------	-------	------------