

**技术白皮书**

OPEN-AVD SDK

开放的**音视频实时通讯 + 视频直播**能力

技术平台

V5.1.1

杭州叁体网络科技有限公司

2020/06



扫一扫关注叁体

3tee.cn

[1. 背景 3](#_Toc143513295)

[2. 产品线 4](#_Toc143513296)

[2.1产品线构成 4](#_Toc143513297)

[2.2 Open-AVD SDK 4](#_Toc143513298)

[2.2.1 概述 4](#_Toc143513299)

[2.2.2 服务器端接口 5](#_Toc143513300)

[2.2.3 客户端接口 5](#_Toc143513301)

[2.2.4 SDK分层 6](#_Toc143513302)

[2.2.5 客户产品架构 6](#_Toc143513303)

[2.2.5 设备与平台 7](#_Toc143513304)

[2.2.6 服务器类型 9](#_Toc143513305)

[2.3佳会 9](#_Toc143513306)

[3. 技术特点与优势 10](#_Toc143513307)

[3.1 叁体SDK，连接一切 10](#_Toc143513308)

[3.1.1 支持的软件操作系统 10](#_Toc143513309)

[3.1.2 支持的硬件终端 12](#_Toc143513310)

[3.1.3 信创 12](#_Toc143513311)

[3.1.4 移动终端 13](#_Toc143513312)

[3.1.5 AR/VR 13](#_Toc143513313)

[3.1.6跨平台开发 13](#_Toc143513314)

[3.1.7 电话技术集成 14](#_Toc143513315)

[3.2 领先音视频技术 14](#_Toc143513316)

[3.3. 领先通讯技术 15](#_Toc143513317)

[3.3.1 P2P与SFU/MCU 15](#_Toc143513318)

[3.3.2多核CPU利用 17](#_Toc143513319)

[3.3.3 分布式服务器集群 18](#_Toc143513320)

[3.3.5 直播 18](#_Toc143513321)

[3.3.6 微信支持 19](#_Toc143513322)

[3.4 其它 20](#_Toc143513323)

[3.4.1 视频硬件编解码 20](#_Toc143513324)

[3.4.2 视频混合编解码 20](#_Toc143513325)

[3.4.3 视频动态码流 20](#_Toc143513326)

[3.4.4 分层视频编码技术 SVC 20](#_Toc143513327)

[3.4.5 音视频服务器端处理技术 20](#_Toc143513328)

[3.4.6 H.323 21](#_Toc143513329)

[3.4.7 SIP 21](#_Toc143513330)

[3.4.8 RTMP 21](#_Toc143513331)

[3.4.9 RTSP 21](#_Toc143513332)

[3.4.10 第三方音视频流倒入与倒出 21](#_Toc143513333)

[3.5 重要技术特性与指标 21](#_Toc143513334)

[4 附录 24](#_Toc143513335)

[4.1 24](#_Toc143513336)

[4.2 术语表 24](#_Toc143513337)

# 1. 背景

杭州叁体网络科技有限公司（3tee.cn）成立于2015年，是一家专业的视频通讯SDK（Software Development Kit-软件开发工具包）开发与运营的软件技术公司。

叁体是国内最早专注于提供应用内视频通讯SDK的软件公司，为第三方软硬件开发商与系统集成商提供覆盖全软件平台、全硬件终端、全通讯协议支持及全场景的技术服务。目前客户覆盖金融、医疗、安防、教育、云服务等众多领域，客户包括工商银行总行、浙江大华股份、杭州创业慧康、中国电信等业界有影响力的公司，他们已选择使用叁体SDK作为产品视频通讯功能的引擎级部件，以此为基础完成自己的行业解决方案。

公司核心成员来自Cisco、WebEx等国内外知名网络视频会议及通讯系统服务公司，在音视频通讯领域的平均工龄超过16年，具有强大的技术实力、丰富的行业经验、深厚的市场背景。

在多年相关经验基础上，结合HTML5/WebRTC等前沿技术，叁体精心打造出了核心产品叁体·视频SDK（Open-AVD SDK），提供了多方音视频实时互动、屏幕分享、白板批注、视频批注、视频直播、视频录制、透明通道及文字聊天等功能。

叁体SDK，连接一切：兼容各种软件系统平台（如Android/iOS、Windows/MacOS/Linux、H5/Webrtc浏览器及微信/QQ浏览器、嵌入式Linux、微信小程序等）；支持所有主流硬件设备（如手机、平板、桌面电脑、电话机、H.323终端、Sip终端、IP摄像头、智能电视、VR眼镜及行业专用设备等等）；更是国内首家（甚至可能是唯一）支持国产计算机系统（麒麟Linux、龙芯/飞腾芯片）的视频通讯SDK。此外，叁体SDK适应公网、专网、内网及混合云等各种网络部署及运行环境。

## 

# 2. 产品线

## 2.1产品线构成

叁体的主要产品线如下图所示：



作为一家专业视频SDK技术公司，叁体的核心产品是Open-AVD SDK，英文全名是Open Audio, Video, Data，意指开放的音频、视频与数据能力，中文花名是“阿梵达”，而“佳会”（Good Meeting）则是一款全功能的会议产品，“钛镁云”（Tele Medicine Cloud）则是参体为远程医疗行业开发的音视频技术平台。

“佳会”和“钛镁云”基于Open-AVD SDK开发，其主要用途是演示Open-AVD SDK的能力，其代码满足一定商务条件下开源，方便第三方团队快速学习，在Open-AVD SDK基础上快速实现自己会议+直播功能。

## 2.2 Open-AVD SDK

### 2.2.1 概述

Open-AVD SDK提供人与人实时沟通协作过程中需要用到的所有基本能力，涵盖了网络会议系统、直播系统及IM系统等三大类终端产品需要的音视频通讯能力。

Open-AVD SDK由业界资深工程师精心打造，稳定可靠，第三方团队拿来就能用，不必自己去造“轮子”，从而降低了第三方团队的技术风险，减少了项目的开发投入，尤其是能大幅缩短第三方团队开发具有多方音视频+数据协作能力的App/Web应用的时间。

Open-AVD SDK可用于几乎所有行业，很多业务场景中需要用到人与人实时沟通与协作的能力，而类似QQ，微信或会议系统这种通用沟通工具又不能直接使用或不能满足功能，这种情况下，Open-AVD SDK就是您最好的选择。市场调研表明，Open-AVD SDK在金融、医疗、教育、能源、交通等各个领域，都有巨大的市场需求。

### 2.2.2 服务器端接口

* 主要功能

帐户管理，用户认证，单点登陆

房间（会议）安排，修改，启动

系统日志

统计、状态报告

* 接口形式

接口以REST API方式提供，相关数据为JSON格式

提供SDK和佳会产品两种接口

* 接口语言

服务器端接口支持如下语言：

* JavaScript
* PHP
* Python
* Java
* C++
* Ruby

### 2.2.3 客户端接口

* 主要功能

客户端接口主要支持以下功能：

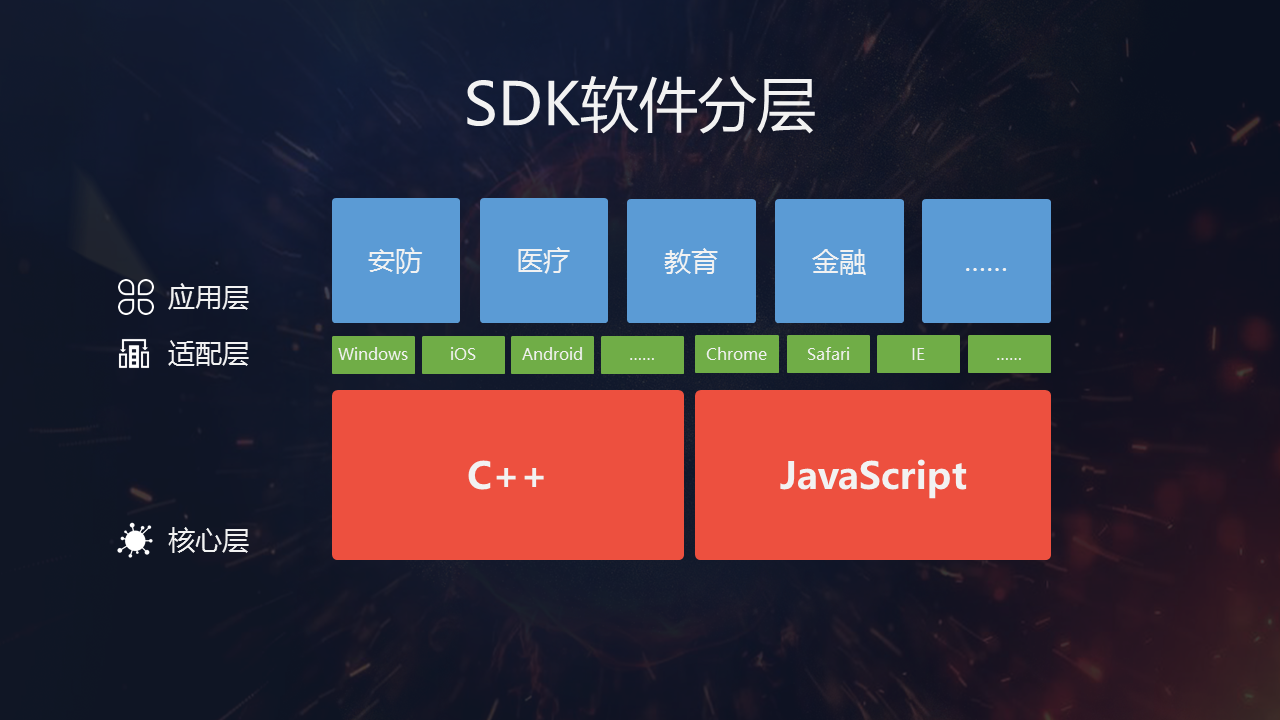
* 文字聊天（单聊，多聊）
* 音频（两方，多方）
* 视频（多方）
* 桌面共享
* 白板
* 批注
* 录制
* 直播（单向音视频，2-10秒延时）
* 各种控制功能，如禁音，踢人等等
* 接口语言

支持主流的编程环境及语言，如：

* JavaScript ( for Web browser )
* C++ （for Windows, Linux）
* Objective C （for iOS, Mac）
* Java（for Android）
* C （for 嵌入式Linux或Windows）
* C#（for Windows）

### 2.2.4 SDK分层

如下图所示，SDK本身分成两层，核心层+适配层，在SDK之上，是客户开发的应用层。

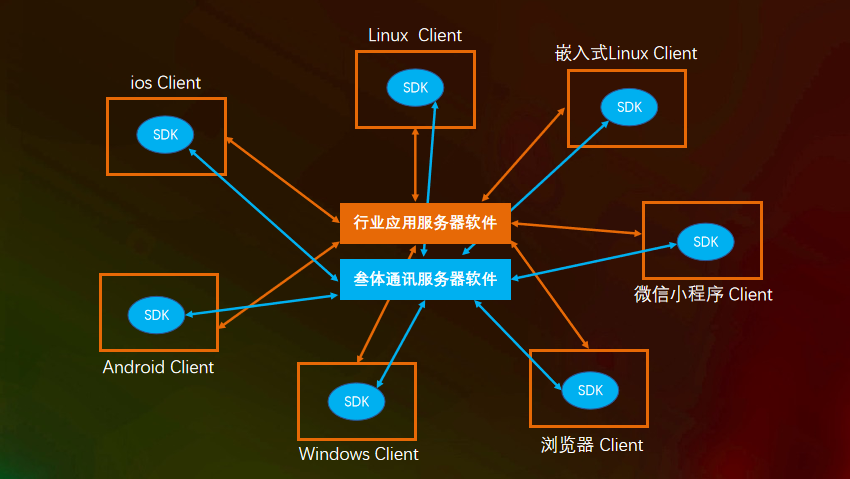


所有和音视频通讯相关的核心代码，都封装在核心层的绿色方块代表的C++库和JavaScript库中，在适配层是非常薄（因而工作量小）的一层代码，做了少量开发，以在需要的地方对不同的操作系统或浏览器做适配。

客户开发团队开发的应用层代码则根据产品需要，调用不同平台的适配层接口完成相应的功能，完全不用理会功能细节及一些技术细节，如线程模型等等。

### 2.2.5 客户产品架构

使用了Open-AVD SDK的第三方业务系统，其基本架构如下图所示：



如上图所示，一个使用了SDK的第三方行业应用，主要由两部分组成，分别是SDK相关功能组件（红色）及行业应用自身特定功能组件（白色）。

SDK本身，也是由两部分组成，一是嵌入在各种业务应用客户端中的SDK代码，二是SDK通讯服务器。SDK代码与服务器代码通过网络（局域网、互联网、移动互联网或专网）连接，交换音视频及其它数据，从而完成相关功能。

行业应用的开发人员在需要使用音视频通讯能力的时候，简单地调用SDK中提供的各种接口即可，不必再去搞清楚音视频通讯过程需要用到的非常复杂的技术细节，如音视频采集、压缩、传输及渲染等等。

行业应用自身与音视频通讯无关的功能，则完全由行业应用开发工程师根据需要自己开发，相关数据传输、存储也完全由行业应用业务服务器自己来处理，与SDK没有关系。

需要说明的是，SDK通讯服务器中安装了所谓的USB Key（俗称硬件加密狗）来控制整个系统能使用的最大音视频并发数，从而控制Box/私有云模式中软件license的数量。除了USB Key这种授权模式，还支持“公网服务授权”及“文件授权”这2种模式。

### 2.2.5 设备与平台

Open-AVD SDK支持现在流行的所有硬件设备及软件平台，可以运行在各种移动终端、桌面系统乃至机顶盒和智能电视上。

2014年10月，在HTML5/WebRTC标准刚刚通过，Open-AVD SDK就在国内第一个宣布对它的全面兼容和支持，使得叁体成为业界第一家支持Html5/WebRTC的公司，能在浏览器中直接启动多方音视频通讯及数据协同功能，免除了用户在浏览器中下载客户端的各种麻烦，开创了极致的用户体验。

以下为支持的主要硬件设备与软件平台：

表1、硬件设备支持表

|  |  |  |
| --- | --- | --- |
| **支持的设备** | **支持方式** | **接口语言** |
| iPhone/iPad | 原生iOS本地客户端 | Objective -C |
| Android Phone/Android Pad | 原生Android本地客户端 | Java |
| WebRTC兼容浏览器，如Chrome/Firefox | JavaScript |
| Windows PC | 原生Windows本地客户端 | C/C++ |
| WebRTC兼容浏览器 | JavaScript |
| Mac PC | WebRTC兼容浏览器，如Chrome/Firefox | JavaScript |
| Linux PC | 原生本地客户端 | JavaScript |
| 自主可控OS  如麒麟OS,UOS等等 | 原生本地客户端 |  |
| 机顶盒 | Android本地客户端 | Java  目前主流的机顶盒都基于Android开发 |
| 智能硬件，如机器人 | 原生Android本地客户端 | Java |
| 各种嵌入式设备 | 原生嵌入式Linux | C/C++ |

表2、软件平台（操作系统等）支持表

|  |  |
| --- | --- |
| **支持平台** | **备注** |
| iOS | 手机、平板 |
| Android | 手机、平板  智能硬件、机器人  机顶盒、智能电视 |
| Windows | Windows桌面  专用工控机 |
| Mac | Mac电脑 |
| Linux | 桌面电脑 |
| 自主可控OS | 自主可控电脑系统 |
| 嵌入式Linux | 一些硬件设备，如摄像头 |
| WebRTC兼容浏览器 | 谷歌Chrome  FireFox  苹果Safari  360极速浏览器  ...... |
| 微信内嵌浏览器 | webrtc技术 |
| 微信小程序 | 微信小程序特有的基于rtmp协议的视频技术 |

### 2.2.6 服务器类型

根据可靠性及复杂度不同，Open-AVD SDK的通讯服务器有不同的类型，现简要描述如下：

表3、服务器类型描述

|  |  |  |
| --- | --- | --- |
| **类型** | **描述** | **特点** |
| 单机 | 服务器软件只有单个实例 | 可靠性低，如果出现服务器软件故联，音视频通讯服务将不可用  实现简单，运维简单 |
| 集群 | 在同一个数据中心中部署很多个物理服务器运行多份SDK通讯服务器软件实例，根据房间内并发人数大小，系统选择多个物理服务器一起工作来提供服务  覆盖了双机热备的所有能力 | 高可靠性  高延展性，支持数千人在同一个房间中 |
| 分布式集群 | 在不同地方的数据中心分别部署服务器集群，所有这些集群中的服务器可以为单个房间的客户端提供音视频服务  覆盖了集群的所有能力 | 高可靠性，可实现异地灾备  客户端可就近智能选择数据中心连接，可实现跨境的大规模实时多方音视频通讯 |

注：叁体服务器支持Linux及Windows操作系统软件。

## 2.3佳会

“佳会”，是基于Open-AVD SDK中间件打造的一个软件视频会议产品，提供所有会议系统的标准功能，如文字聊天、多方实时音视频、旁路直播、屏幕共享、音视频录制等功能。

“佳会”支持所有的软件平台，尤其是支持HTML5/WebRTC兼容浏览器，能直接从浏览器里启动、加入会议，不用下载任何插件。

“佳会”基于Open-AVD SDK开发，可开放源代码给授权合作伙伴，主要目的是演示Open-AVD SDK的能力，同时为第三方团队使用Open-AVD SDK的API/SDK提供样例代码。

# 

# 3. 技术特点与优势

叁体团队核心成员来自于WebEx和Cisco这种做网络视频会议和通讯系统的公司，在音视频及通讯产品的技术方面有多年积累，在音视频处理、音视频通讯、系统稳定性、安全性、大并发、高可用等方面，有丰富经验。

同时，作为一个专业的音视频软件技术公司，叁体有能力和机会紧跟技术发展的最新趋势，兼容业界近几年发展出来的最新的成熟技术，打造出全球领先的技术平台，引领技术潮流。

现对叁体的主要产品SDK的技术特点与优势分别描述如下：

## 3.1 叁体SDK，连接一切

作为业内最专业的视频SDK公司，叁体视频SDK支持市场上所有带音视频能力的设备，无论设备使用什么样的软件操作系统（浏览器），无论设备运行在什么样的网络环境中，也无论设备采用了什么样的音视频技术。

### 3.1.1 支持的软件操作系统

如下图所示，叁体SDK支持了所有种类的软件操作系统及浏览器环境：



尤值得一提的是叁体对H5/Webrtc、自主可控OS及嵌入式Linux的支持。

* H5/WebRTC

2014年10月28日，在历经8年多的艰苦努力后，万维网联盟W3C终于宣布HTML5标准获得通过，这标志着新一代Web应用标准化的努力获得了决定性的进展。

HTML5中最引人入胜的特性就是WebRTC（Web Real　Time Communication　–　Web实时通讯），它第一次使用一种标准的方法，在WebRTC兼容浏览器中原生引入了音视频和双向通讯能力，无需下载任何附加插件或本地Client，就能在浏览器里实现音视频聊天、会议等各种功能，开创了极致的用户体验。

* 自主可控OS

2017年，通过与一个军方项目的合作机会，叁体支持了麒麟OS+飞腾芯片。快4年以来，叁体对自主可控OS及芯片持续跟进和投入，目前已经完成了对主流自主可控OS和芯片的全方位支持，如麒麟OS及UOS，还有飞腾、龙芯、海思及鲲鹏等国产芯片。



2020/05/27，叁体产品通过麒麟OS兼容性认证

* 嵌入式Linux OS

因为低功耗、方案成本低等重要特点，市场上存在大量硬件设备公司都选用嵌入式Linux作为操作系统。叁体SDK深度支持嵌入式Linux，并适配了众多芯片，使得各种专用硬件视频设备可以通过叁体的视频技术平台与其它各种软件终端直接连接（比如与微信小程序互通），极大地丰富了专用视频硬件终端的使用场景。

### 3.1.2 支持的硬件终端

下图为叁体SDK支持的各种种类的硬件设备与终端。



由于叁体SDK的广泛适应性，使得第三方应用能轻松与各种硬件终端集成，保护了已有硬件设备上的投资。

### 3.1.3 信创

作为一家本土的音视频技术公司，叁体对于国产技术的支持一直高度重视，早在2017年，叁体就支持了银河麒麟操作系统，开了业界先河，并在2021年获得麒麟软件及统信OS等主流信创操作系统厂家的兼容性认证证书。

近年来，随着信创技术的广泛应用，叁体紧跟市场需求，把信创支持从操作系统领域扩展到了数据库等技术领域，最大限度满足了各行业对信创的多元化的需求。

以下为叁体支持各种信创产品的列表（持续更新中）

|  |  |  |  |
| --- | --- | --- | --- |
| 信创操作系统 | 芯片 | 支持 | 备注 |
| 麒麟 | 海光x86 | 是 |  |
| 银河麒麟 | 海光x86 | 是 |  |
| 鲲鹏ARM | 是 |  |
| 飞腾服务器ARM | 是 |  |
| 飞腾桌面系统 | 是 |  |
| 飞腾腾锐D2000 | 是 |  |
| 统信 | 飞腾ARM | 是 |  |
| 鲲鹏ARM | 是 |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 信创数据库 | 版本 | 支持 | 备注 |
| 人大金仓 | V8 | 是 |  |
| 达梦 | dm8 | 是 |  |
|  |  |  |  |

### 3.1.4 移动终端

在最终用户为个人消费者的移动应用市场中，各种技术都有大量使用者，这就要求视频技术厂商能提供支持各种技术的sdk，从而帮助应用开发厂商完美适配元样化的市场需求。

叁体在这个方向做了大量努力，提供了技术上可行的几乎所有移动平台音视频技术方案，总结为下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 移动平台 | 场景 | 技术 | 支持 | 备注 |
| Andorid | 原生技术 | Java开发 | 是 |  |
|  | QQ/360/Chrome浏览器 | Web H5 | 是 |  |
|  | 微信内置浏览器 | Web H5 | 是 |  |
|  | 微信小程序 | Web H5/小程序SDK | 是 |  |
|  | 原生集成Webview | Web H5 | 是 |  |
|  | 钉钉/支付宝/淘宝 小程序 |  | 否 |  |
|  | 钉钉/支付宝/淘宝 浏览器 |  | 否 |  |
| Ios | 原生技术 | Object C | 是 |  |
|  | Safari 浏览器 | Web H5 | 是 |  |
|  | 微信内置浏览器 | Web H5 | 是 |  |
|  | 微信小程序 | Web H5/小程序SDK | 是 |  |
|  | 原生集成Webview | Web H5 | 是 |  |
|  | 钉钉/支付宝/淘宝 小程序 |  | 否 |  |
|  | 钉钉/支付宝/淘宝 浏览器 |  | 否 |  |

### 3.1.5 AR/VR

叁体紧跟技术发展，结合客户实际需求，从2018年就开始跟踪VR及AR相关技术，在这个方向已经有相当积累。

目前已经支持了RealWear、Rokid、DreamGlass等多款AR眼镜，并兼容了Pico、创维、大朋等众多VR眼镜，支持合作伙伴落地实施了远程巡检、远程维修、远程医疗等众多场景化应用，受到业界伙伴的好评。

在历时8年的潜心准备后，苹果公司于2023/06/06发布了它在XR领域的第一款产品Vision Pro，革命性地重新定义了人与XR设备的交互模式，开创了全新的可能性。叁体将继续跟进这个领域的潮流，与合作伙伴一道，开发出XR+Video的场景应用，引领市场需求。

### 3.1.6跨平台开发

为适应业界跨平台开发技术的持续需求，叁体对各种跨平台开发技术进行了适配，协助合作伙伴提高了开发效率，降低开发成本。

跨平台开发技术列表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 类别 | 技术 | 描述 | 支持 | 备注 |
| 移动端 | Uni-app | 移动终端跨平台开发 | 是 | 编程语言为JS |
|  |  |  |  |  |
| 桌面端 | Electron | 桌面跨平台开发支持Windows/Mac/Linux | 是 | 编程语言为JS |
|  |  |  |  |  |

### 3.1.7 电话技术集成

**叁体视频平台通过SIP网关与电话实现语音的互联互通，通过VOLTE协议实现音视频的互联互通，拓展了音视频的使用范围和业务模式**

## 3.2 领先音视频技术

针对互联网尤其是移动互联网的特点，叁体设计和使用了专门的技术方案并做了专门的优化与调校，使得系统即使在网络情况不佳的情况下，仍能保持音视频的清晰、流畅与稳定。

尤其是，在唇音同步、音频回声消除与降噪、视频动态码流、视频混合编解码、视频硬件编解码、视频多流等方面，叁体远远领先于市场。

现把叁体Open-AVD SDK音视频特点总结如下：

**表5、【Open-AVD SDK音频特点】**

|  |  |  |
| --- | --- | --- |
| **项目** | **指标、值** | **说明** |
| 编解码 | iSAC,iLBC, opus  Speex,AAC | iSAC，iLBC专门为互联网环境设计，在3G/4G/5G环境下表现优异，质量远优于传统技术  可扩展支持Speex，AAC等编解码 |
| 回声消除 | 优秀 | 直接使用一般的笔记本外放话筒也都不会出现回声或啸叫等异常 |
| 语音动态码流 | 支持 | 支持语音动态码流，根据网络的情况对码流进行自适应调节，因而可以在不同的网络环境中保持出色的语音质量，即使在2G网络环境中也能通话 |
| 语音处理 | 优秀 | 具有出色的降噪、自动增益控制技术，保持语音清晰响亮 |
| 网络处理 | 优秀 | 具有先进的抖动缓冲和丢包补偿技术，保持语音稳定、低延时 |
| 音质 | 电话音质 | 使用了最优秀的语音处理与语音编解码技术 |
| 混音 | 支持 |  |

**表6、【Open-AVD SDK视频特点】**

|  |  |  |
| --- | --- | --- |
| **项目** | **指标、值** | **说明** |
| 编解码 | VP8/VP9  H.264/H.265 | VP8/VP9专门为互联网环境设计，在3G/4G/5G环境下表现优异  支持H.264/H.265，并可支持硬件编解码  可在单个会议类场景中支持VP8/VP9与H.264/H.265混合编码 |
| CPU消耗 | 低 |  |
| 码流 | 低 | 根据网络情况智能调整分辨率与帧率等核心指标，从而动态调整码流 |
| 视频规格 | 全面 | 支持从176\*144到4K的各种规格视频 |
| 图像质量 | 优秀 | 出色的图像处理技术，保持图像清晰稳定 |
| 流畅度 | 优秀 | 具有先进的网络处理技术，保持图像的流畅度 |
| 视频SVC | 支持 | 编码时可将视频编码成高、中、低路不同分辨率的视频放在同一路视频流中，发送到服务器，服务器根据接收端的情况选择发送不同“层”的码流。比如，手机终端接收低分辨率码流，笔记本接收中等分辨率码流，大屏幕终端接收高分辨率码流。 |
| 支持路数 | 理论上不限 | 根据带宽和CPU情况确定，建议PC端不超过25路，移动端不超过8路 |

## 3.3. 领先通讯技术

### 3.3.1 P2P与SFU/MCU

从架构上讲，我们有两种模式来搭建整个音视频通讯系统，分别是P2P(Peer-to-Peer)和SFU（Selective Forwarding Unit）及MCU（Multipoint Control Unit）。



如上图所示，P2P最大特点是，所有音视频的信息直接在通讯的两个客户端之间传递，无需经过服务器的中转，在参与通讯的客户端较多的时候，每个客户端都需要和其它所有客户端建立通讯链路。



SFU

另一种同样被广泛使用的架构是SFU/MCU，其最大特点是，两个通讯的客户端之间并不直接相连，所有的音视频信息传递都通过一个处于中心位置的服务器（集群）来转发。两种架构都有广泛的应用，分别有各自的优势与局限，现简单对比如下：

表7、 P2P vs. SFU/MCU

|  |  |  |  |
| --- | --- | --- | --- |
|  | **P2P** | **SFU** | **MCU** |
| 功能特点 | 所有音视频数据直接在客户端之间传递 | 所有音视频数据都需要经过服务器，但只做“转发” | 所有音视频数据都需要经过服务器，可在服务器上做特殊的功能“处理”，如合流、转码等等 |
| 优点 | 大量小规模会议时，节省带宽 | 通过服务器集群能支持很大规模的会议，如单个房间5000人  可支持一些中心化的功能，如服务器端录制 | 同SFU  合流转码等处理功能 |
| 缺点 | 一个房间只能支持3-5人  对客户端网络质量要求高 | 大量小规模房间时，服务器带宽很浪费 | 同SFU  服务器端需要投入GPU等硬设施，增加成本 |
| 典型产品 | Skype，叁体 | WebEx，叁体 | 叁体 |

MCU与SFU的最大区别是MCU在服务器端有很强的运算能力，除了具备SFU所有功能外，能在服务器端对音视频数据进行特殊“处理”，如多路合屏、转码，由于运算能力要求高，需要借助于GPU等专用硬件设备。MCU在软视频领域的成熟，使得类似宝利通这样的硬件视频会议技术在这个领域不再有优势，从而为软视频技术大规模使用扫除了障碍。

依托在网络视频会议及通讯领域的多年经验积累，叁体的工程师革命性地设计和实现了同时支持P2P、SFU及MCU三种模式的技术架构，使得系统可根据人员数量做智能动态切换成P2P、SFU/MCU模式，既节约了服务器带宽，又能支持超大并发应用场景，鱼与熊掌兼得。

另外，叁体的MCU技术有GPU与CPU两种两种实现方案，现简单比较如下：

|  |  |  |
| --- | --- | --- |
|  | GPU | CPU |
| 适用场景 | 大规模视频处理  高频度 | 很小规模（比如几路）  偶尔使用 |
| 成本 | 专用设备，相对高 | 直接使用服务器CPU，低廉 |

### 3.3.2多核CPU利用

随着近年来CPU制造技术的突飞猛进，16核以上的CPU越来越多，而能否充分使用所有内核的能力，对于大规模PaaS平台的运营十分关键。

和一些传统的音视频厂商不同，叁体汲取了多年音视频及通讯系统设计中的经验教训，在成立之初就把充分利用多核CPU能力设定为最重要的技术目标之一，在服务器技术选择、架构设计和技术实现方面，都做了精心处理，使得系统在运行的时候，能充分利用所有CPU内核的能力，避免了计算资源的浪费，从而在节能降耗及IDC运营成本方面，获得明显的优势。

### 3.3.3 分布式服务器集群

由于历史原因，国内互联网的网络质量存在着地区差异，不同互联网ICP之间，不同地域之间，都存在一定的联通性问题，同时，跨境的互联网访问，也存在巨大的限制，这对于运营一个全国性乃至全球性的音视频+数据实时协同PaaS平台，是一个非常大的挑战。

叁体对于这个问题的解决方案是，设计开发了一个分布式的服务器集群架构（如下图所示），在全国乃至全球多地部署服务器集群。“本地”的client根据一定智能算法，选择最佳（通常是离自己最近的）服务器集群连接，而服务器集群之间则由相关ICP通过特殊的手段来保障其连接，这样，所有client都能获得很好的网络连通性从而保证整个协同沟通的质量。

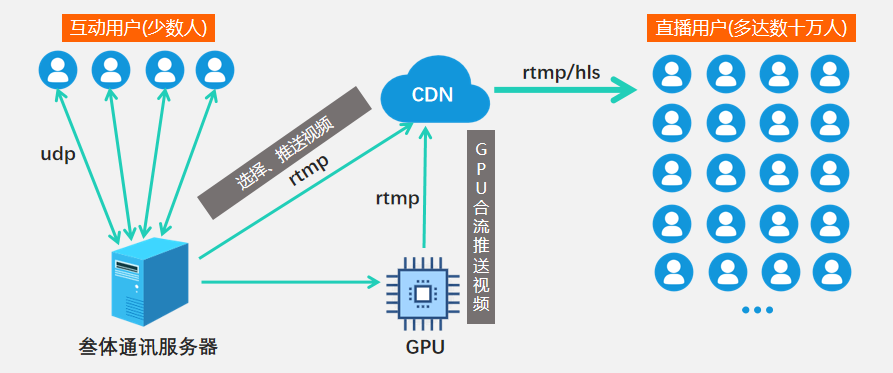
展望未来，通过和ICP的合作，叁体将开发和部署一套类似于CDN（Content Delivery Network）的分布式系统，为所有合作伙伴提供良好的实时音视频和数据共享能力服务，我们叫它RCDN（Realtime Content Delivery Network）。



### 3.3.5 直播

Open-AVD SDK除了能支持实时互动音视频，还能选择某些客户端的音视频，通过第三方成熟的CDN网络推送给其它各种客户端，可支持实时互动音视频与直播功能的混合使用场景，还可将多路视频合流后推送。

其基本架构如下图所示：



注1：需要向第三方CDN供应商租用CDN网络，在向第三方CDN推流前，可以对视频进行合流处理，从而大量节约CDN带宽，降低运营成本。

注2：对于内网环境，叁体提供rtmp服务器软件以替代CDN的功能。

直播的基本特性如下：

|  |  |  |
| --- | --- | --- |
| **特性** | **指标** | **备注** |
| 并发 | 数十万 |  |
| 延时 | 2-10秒 | RTMP：2-5秒  HLS：10秒 |
| 客户端平台 | 微信客户端  Android/ios本地客户端  PC/浏览器/Flash |  |
| 分辨率 | 高至1920\*1080 |  |
| 与实时互动混合 | 支持 |  |

### 3.3.6 微信支持

随着微信的广泛普及，在微信中支持音视频能力变得特别重要。叁体提供了2种技术支持在微信中使用“互动”视频能力，以下简单描述比较：

|  |  |  |
| --- | --- | --- |
|  | 微信内嵌浏览器 | 微信小程序视频接口 |
| 描述 | 微信内嵌浏览器已经可兼容webrtc技术 | 微信小程序开放了自己的视频API |
| 优点 | 直接使用webrtc相关技术  低延时100-300ms | 同时支持ios及android平台 |
| 缺点 | 目前只支持android | 基于rtmp技术，延时500-1000ms |

## 3.4 其它

### 3.4.1 视频硬件编解码

按目前的技术水平，在采用软件编解码技术的情况下，一般的移动终端能只能支持320\*240规格的视频，640\*480只能勉强支持，1280\*720规格以上的高清视频基本上不可能。

为解决这个问题，叁体的视频引擎针对市场上的主流H.264视频处理芯片进行了优化，覆盖了90%以上移动终端，从而为移动终端上显示高清视频提供了可能性，目前该特性支持Android 4.2和iOS7以上系统。

同时，对于一些比较少见的芯片或深度定制过的Android系统，叁体的视频引擎提供了插件接口，第三方可开发专门插件来实现H.264硬编解码。

### 3.4.2 视频混合编解码

目前有些webrtc兼容浏览器只支持VP8，在和只支持H.264的硬件终端共处一个会议场景的时候，会存在互通问题。

为解决这个问题，叁体创造性地开发了混合编解码技术，即在可能的情况下，在视频采集端同时使用两种编码技术分别编码，发出两种不同的视频数据流，由服务器根据连上来的客户端的能力，选择不同的视频数据发出。

### 3.4.3 视频动态码流

互联网尤其是移动互联网的网络情况比较复杂，容易出现各种波动，造成视频卡顿等问题。叁体的系统在出现网络波动的时候，会动态调整视频的分辨率及帧率等核心指标，在网络情况不好的时候降低码流，从而在牺牲一定图像清晰度的情况下，保障视频的流畅度。

### 3.4.4 分层视频编码技术 SVC

受限于CPU的能力和网络带宽和成本，可以预见在相当长时间内，并不是所有移动终端都会或能处理1280\*720规格以上的视频。

叁体SDK支持SVC（Scalable Video Coding - 分层视频编码技术），在视频采集端能在一个视频流中包含不同分辨率及帧率的多层视频数据，视频到达服务器后，服务器根据一定的策略（如终端的CPU能力、网络的状况）智能选择合适的视频发给相应的客户端，从而提高视频的终端适应能力。

### 3.4.5 音视频服务器端处理技术

叁体SDK的通讯服务器软件，能利GPU的运算能力，对经由服务器传输的音视频流做以下处理

* 合流（合屏）

将多路音视频合并成一路，以节省服务器带宽

* 转码

对视频流进行解码，再编码成新的视频，以适应某些特殊需求，如把H.265转成VP9等等。

### 3.4.6 H.323

叁体提供对H.323协议的支持，可方便集成H.323兼容的硬件设备，从而保证客户的大量投资不会被浪费。

### 3.4.7 SIP

叁体提供对H.323协议的支持，可方便集成H.323兼容的软硬件设备，如各种sip软电话及硬电话。

通过使用媒体网关，叁体SDK还能与电话系统PSTN集成。

对于SIP的变种GB/T28181，叁体SDK也提供了兼容，从而可以方便接入安防行业的相关技术平台。

### 3.4.8 RTMP

叁体提供对RTMP协议的支持，可方便集成市场上已经存在的大量支持RTMP协议的音视频系统，并与CDN系统对接，提供互动直播功能。

### 3.4.9 RTSP

叁体提供对RTSP协议的支持，可方便集成市场上已经存在的大量支持RTSP协议的音视频设备和系统，如海康和大华等监控厂商生产的IPC（网络摄像头）。

### 3.4.10 第三方音视频流倒入与倒出

针对一些特殊设备（如智能眼镜），叁体提供了第三方音视频倒入与倒出功能，从而满足一些特殊设备的特殊音视频处理与传输需求。

## 3.5 重要技术特性与指标

表9、重要技术特性

|  |  |  |  |
| --- | --- | --- | --- |
| **类别** | **项** | **值** | **备注** |
| 音视频 | 唇音同步时差 | <100毫秒 | 唇音完全同步，完美用户体验 |
| 音频优先 | 支持 | 在网络不好情况下优先保障声音的质量 |
| 视频 | 视频路数 | 1-16 | 取决于视频规格、客户端运算能力及网络环境，建议移动端不超过3路，桌面端不超过8路 |
| 视频规格 | 176\*144  320\*240  352\*288  640\*480  720\*576  1280\*720  1920\*1080  4K |  |
| 视频编解码 | VP8/VP9  H.264/H.265 | 业内首家支持H.265 |
| 硬件编解码 | 支持 | H.264硬编解码 |
| 动态分辨率调整 | 支持 | 可根据网络情况和CPU能力，动态调整视频发送端的分辨率，从而保证在网络情况不佳或CPU性能不好的时候，保持视频的流畅度 |
| 多层编码SVC | 支持 | 可根据不同的客户端，选择不同规格的视频发送 |
| 视频合屏 | 支持 | 可在服务器端做视频合屏，节约服务器（数据中心）带宽 |
| 视频转码 | 支持 | 可在服务器端做视频合屏，适配不同能力的客户端设备，如机顶盒与高清H.323终端共处一个会议室中 |
| 音频 | 音频路数 | 无限制 |  |
| 音频编解码 | iSac  iLBC  AAC  G711/G729 |  |
| 回声消除 | 支持 | 优秀，笔记本外放无回声 |
| 噪声消除 | 支持 | 优秀，电话音质 |
| 自动增益控制 | 支持 |  |
| 服务器 | P2P与MCU | 支持 |  |
| MCU级联 | 支持 |  |
| 负载均衡 | 支持 |  |
| 水平延展 | 支持 |  |
| 分布式集群 | 支持 | 客户端可就近连接最快最稳定的服务器集群，保证极致的用户体验，在跨境协同沟通应用如远程会诊中，意义重大 |
| 安全 | 传输安全协议 | DTLS |  |
| 数据加密算法 | AES |  |
| 网络 | 防火墙穿越 | 支持 | 自动穿透所有防火墙及代理服务器 |
| 接入方式 | 有线网络  Wifi  3G/4G/5G |  |
| 端口开放要求 | 80/443 |  |
| 抗丢包率 | 20% | 在20%网络丢包情况下，仍能保持音视频质量 |
| 其它 | 支持平台 | 所有浏览器  iOS/Android  Windows/Mac/Linux  微信  嵌入式Linux  国产OS | HTML5/WebRTC兼容浏览器，不用下载，其它浏览器可下载插件及本地客户端来支持  支持微信/QQ扫码加入会议 |
| 免下载 | 支持 | 在h5/webtc兼容浏览器里，不必下载任何客户端及插件就能使用音视频能力 |
| 国产计算机系统 | 支持 | 支持国产麒麟Linux操作系统，飞腾/龙芯/等等 |
| 支持设备 | PC/Mac桌面系统  iPhone/iPad  Android设备  机顶盒，智能电视  各种嵌入式Liunx设备 |  |
| H.323 | 支持 | 与传统硬件会议终端对接 |
| SIP | 支持 | 与传统视频会议对接、支持电话及软件电话对接 |
| RTMP | 支持 | 可与CDN网络集成，完成直播，并可合屏直播 |
| RTSP | 支持 | 可据此与安防摄像头对接 |
| 电话 | 支持 | 可从会议室里拨电话 |
| 微信小程序互动 | 支持 | 可以在微信小程序实现音视频互动 |
| 音视频导入导出 | 支持 |  |

# 

# 4 附录

## 4.1

## 4.2 术语表

表10、术语表

|  |  |  |
| --- | --- | --- |
| **术语** | **术语解释** | **备注** |
| SDK | 英文Software Development Kit的缩写，其含义是软件开发工具包，通常是一个软件公司提供给其它开发公司的工具集合，其它公司可以在SDK基础上做开发 | SDK的典型例子有科大讯飞的语音识别SDK，也有叁体的视频通讯SDK |
| 音频优先 | 实时音视频通讯过程中，会有大量声音与视频数据需要传输，而视频的数据量远大于音频，而从沟通效果来说，保证音频数据的传输远比视频重要  所以，在网络状况不良（比如带宽受限）的时候，需要采取特殊的技术，可以对视频数据进行丢弃处理，以优先保证声音数据的传输，这就是所谓音频优先 | 音频优先会带来辰音同步问题需要解决 |
| 唇音同步 | 由于实时音视频传输过程中，声音与视频数据是由不同的通道来传输的，需要采取专门的技术来保持声音与视频之间的同步性，从效果上来说，需要保持说话时的嘴唇动作和声音的一致性 | 叁体的唇音同步时差小于100毫秒，远远优于市场上一些做网络会议的公司，是赢得类似大华这种大客户的关键 |
| 回声消除 | 声音在播放过程中，由于场地环境等因素，扬声器传出的声音，会从唛克风传回系统，从声音听起来有类似山谷里回声现象，影响收听效果，严重的时候，会形成“啸叫”，完全听不清楚  回声消除就是采取各种手段降低乃至避免回声的技术，是声音处理的关键技术，直接关乎音频通讯的质量 | 叁体的回声消除技术业内遥遥领先，笔记本外放，完全无回声 |
| iLBC | Internet Low Bitrate Codec，互联网低码率编解码，是一种开源的窄带宽话音编解码技术，由WebRTC原生支持的音频技术，专门为互联网环境设计，配合相关音频引擎，在较差的网络条件下仍能保持较好的音质 |  |
| iSAC | Internet Speech Audio Codec，互联网话音编解码，是一种开源音频编码技术，由WebRTC原生支持，专为互联网环境设计 |  |
| 降噪 | 声音在传输处理过程中，会有各种电子噪音混入，从而影声音质量  降噪就是采取各种技术手段，过滤噪音 | 叁体的降噪技术业内遥遥领先，音质和电话音质不相上下 |
| H.264/H.265 | H.264是一种主流的视频编码技术，在安防等行业大量使用  H.265是脱胎于H.264的一种新的视频编码技术，预期2016年开始会逐渐获得工业界应用 | 作为叁体强大技术能力的一个标志，叁体同时支持这两个系列的编码技术，并能在同一个通讯场景（如多方会议）中，混合使用这两种编码 |
| VP8/VP9 | VP8相对于H.264，是一种相对新的视频编码技术，专门针对互联网通讯环境进行了优化，在互联网行业应用广泛，最典型的应用是在html5的音视频规范WebRTC中  VP9是脱胎于VP8的新编码技术，逐渐开始流行，并获是了主流芯片厂商的支持 |
| 硬件编解码 | 视频编码与解码，其中一种方式，是使用手机或PC的CPU来做大量运算，这个过程叫软编解码  如果CPU本身不够强劲，运算过程不够快，在处理比较高清的视频（如720P或1080P）就会造成各种视频问题，如卡断、延时等等  解决这个问题的一种办法是使用手机或PC上已经自带的硬件芯片来做编解码，从而避免上述各种问题 | 叁体Open-AVD SDK针对主流芯片提供了硬编解码能力，从而保证大多数Android、iOS移动终端能处理高清视频 |
| 多流 | 指摄像头在采集视频数据的时候，采集不同规格（分辨率、帧率）的多份视频数据传输到通讯服务器，服务器根据连接上来的不同客户端的能力（如是否是运算能力比较弱的移动终端）选择不同的视频进行传输，从而保证参与视频通讯的所有客户端都能看到流畅的视频 | 叁体Open-AVD SDK支持多流，极大提升了用户体验 |
| 动态码流 | 指在视频信号处理传输过程中，如果网络出现波动导致传输不畅的时候，采取降低视频分辨率及帧率等手段降低传输码流，从而使得在网络状况不好的时候，仍能够保持视频的流畅（质量会有所下降），待网络情况变好的时候，系统自动恢复到更高的分辨率及帧率 | 叁体有非常优异的动态码流技术 |
| P2P | Peer to Peer，意指点对点通讯，这是音视频通讯的典型架构，参与通讯的client直接通过网络传输音视频数据，不必通过服务器中转  这种通讯架构的好处是，大量小规模场景（如远程会诊），可以节约大量服务器（数据中心）网络带宽 | Skype是典型的P2P架构，WebEx是典型的MCU架构  叁体同时支持这两种架构 |
| MCU | Multi-Control Unit，多点控制单元，是硬件视频会议中的一个设备，用于在硬件视频会议终端间传输数据并提供控制功能  目前这个词的使用已经超出了它的原始含义，在软视频会议系统中，主要指用于转发音视频数据并提供控制功能的媒体通讯服务器  采用了MCU的音视频通讯架构，是和P2P相对应的一种通讯架构，在业界也大量使用，其典型优点是能支持大量并发用户，如超过5000人的会议 |
| 分布式服务器集群 | 指服务器集群分布在物理位置不同的城市和国家，使得client端可以采取就近远程最佳服务器来连接，从而保证所有客户端最好的用户体验，如很低的延时、优质的音视频质量等等 |  |
| HLS | Http Live Streaming，Http活动流，一种Apple公司提出的流媒体技术，和RTMP技术一样，是目前市场上主流的直播服务支撑技术 | 叁体Open-AVD SDK已经完成基于HLS技术的直播功能，可在微信里直接看音视频直播 |
| RTMP | Real Time Messaging Protocol，实时消息传输协议，是一种设计用来进行实时数据通信的网络协议，主要用来在Flash/AIR平台和支持RTMP协议的流媒体/交互服务器之间进行音视频和数据通信  目前市场上很多设备和系统，如大华海康的摄像头，支持RTMP协议 | 叁体Open-AVD SDK可以和支持RTMP协议的设备/系统集成 |
| RTSP | Real Time Streaming Protocol，实时流传输协议，该协议定义了一对多应用程序如何有效地通过IP网络传送多媒体数据  目前市场上很多设备和系统，如教育领域里的录播系统，支持RTSP协议 | 叁体Open-AVD SDK可以和支持RTSP协议的设备/系统集成 |
| SIP | Session Initial Protocol，会话初始协议，在一个广泛应用的信令控制协议  支持SIP的各种通讯client，如sip电话，软电话终端可以互相通讯 | 叁体Open-AVD SDK可以和支持SIP协议的设备/系统集成 |
| H.323 | 国际电信联盟ITU-T于1996年公布了H.323协议，它是局域网和企业网使用的多媒体通信标准  H.323是Internet上端与端之间进行实时声音和视频会议的规程和协议，包括系统和构件的描述、呼叫模型的描述、呼叫信令过程、控制报文、服用、语音编码解码器、视像解码器及数据协议  H.323已经有20年历史，围绕这个技术协议系统，业界已经生产了大量的硬件设备和系统。 | 叁体Open-AVD SDK可以和支持H.323协议的设备/系统集成 |
| Html5 | HTML是Hyper-text Markup Language，也就是超媒体标记语言，是Web应用的通用性基础语言  HTML5是该语言的第5代规范，于2014/10/28获是标准化组织的批准 |  |
| WebRTC | Web Real Time Communication，是HTML5标准重要组成部分，专注于解决在web浏览器中提供原生的音视频通讯能力的问题 |