

# 火山引擎视频云：坚持基础技术创新，打造极致用户体验

原创

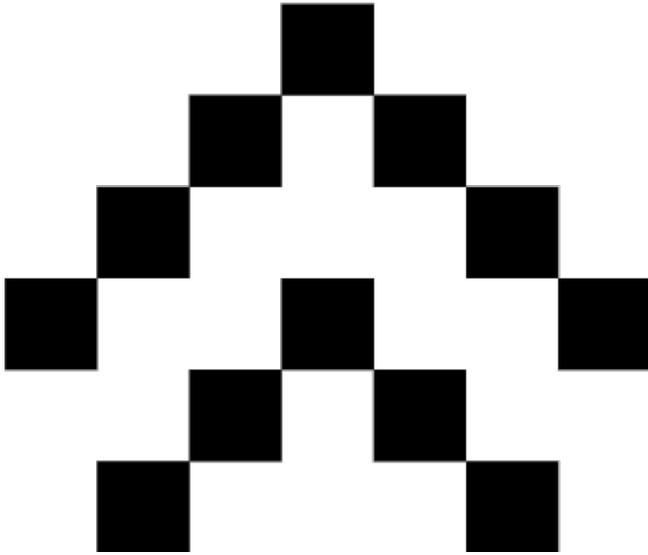
[LiveVideoStack](#) 于 2021-09-22 08:00:00 发布 1103 收藏 1

文章标签：[算法](#) [大数据](#) [编程语言](#) [人工智能](#) [java](#)

版权声明：本文为博主原创文章，遵循 [CC 4.0 BY-SA](#) 版权协议，转载请附上原文出处链接和本声明。

本文链接：<https://blog.csdn.net/vn9PLgZvnPs1522s82g/article/details/120426634>

版权



点击上方“LiveVideoStack”关注我们

从toC到toB，作为抖音背后的视频技术中台，火山引擎视频云经历了哪些转变和挑战？对于国际标准的制定，它是如何实现从跟随者到参与者的跨越？视频盗版问题泛滥，火山引擎视频云又有哪些应对措施？面对日益丰富的直播场景，火山的音视频引擎如何有效支持直播业务不断增加的场景变化？未来还将有哪些直播新玩法？

何予

技术访谈

#005#

LiveVideoStack近期采访了火山引擎视频云架构技术总监何予，和他一起讨论了火山引擎视频云从toC到toB的转变，实现H.266所用到新技术、如何评估视频画质、RTC技术突破以及应对视频盗版等话题。



何予：2006 年于清华大学电子系获得学士学位，2012 年在中国科学院研究生院获得博士学位，目前担任火山引擎视频云架构技术总监，在多媒体领域有丰富的算法、工程架构和产品业务经验。

**LiveVideoStack:** 何老师，您好。请您介绍一下您在火山引擎视频云所负责的工作，以及您目前所关注的技术方向。

何予：我是火山引擎视频云架构技术总监，我们部门的主要业务目标和战略重心是，如何将海量富媒体内容以最低的成本、最优的画质、最低的延时、最安全可靠的方式传递给受众，因此我们所关注的技术方向包括多媒体内容从生产到播放的全链路，包括采集、编码、处理、转码、传输、播放等环节中所涉及到的算法和工程技术，以及大规模高并发高可靠的分发架构。

**LiveVideoStack:** 前段时间，火山引擎召开了品牌发布会。发布会上，火山引擎总经理谭待提到要把字节跳动最好的技术开放出去。这是一个从toC到toB的转变，那么在这种转变的过程中，火山引擎视频云在技术上会面临哪些挑战？

何予：

- 1、从对内中台角色向toB技术服务角色的转化带来的挑战：内部中台要求更专、更精，在技术方案上也可以和业务层有一定的渗透。但toB技术服务需要更通用，和业务层次划分得更明显。因此我们对技术架构进行原子化拆分，将API都设计得更简单通用。
- 2、客户考核标准带来的变化：内部业务要求结果导向，在考核标准和业务目标上有高度共享，比如大家都要从用户体验、单位成本的角度去做优化。但toB因为跨公司，和客户的真正目标之间有隔层，客户更期望云服务公司利用标准方案进行PK。所以我们一方面输出经过业务打磨的最佳实践方案，另一方面也需要向行业标准方案看齐。
- 3、更复杂的业务环境：尤其是客户端方向，内部业务环境的统一，而客户却使用不同组件、不同开发环境，所以我们需要做更多适配、将方案做得更灵活，给出更全面的demo。
- 4、团队人员思维的改变：对内更要求深挖，toB更要求服务性。

**LiveVideoStack:** 去年7月，H.266/VVC标准发布。在保持视频清晰度的情况下，H.266/VVC大约可以节省50%的码率。火山引擎在实现H.266的过程中应用了哪些新技术？对于H.266之后新一代视频压缩标准，您有什么样的期待？

何予：标准被誉为高科技产业王冠上的宝石。谁在标准中拥有更多专利，在产业中就拥有更大的话语权与主动权。正因为如此，各大科技巨头都高度重视标准制定，将相关专利技术视为核心资产。字节跳动火山引擎团队作为后起之秀，坚持基础技术创新，参与国际标准制定，实现了从标准跟随者到制订者的跨越。从多家第三方的报道来看（<https://www.lexology.com/library/detail.aspx?g=cb7164b0-9b22-46f6-a40f-c9b549ab3684>），在国际标准H.266/VVC标准贡献者排行榜上，字节跳动排名第三。同时，H.266/VVC标准贡献者排行榜前五的企业中，字节跳动作为一家年轻的互联网企业，是唯一一家没有参与上一代视频编解码标准开发的公司。但在新一代视频编解码标准制定过程中，字节跳动发起的100多项技术提案获得H.266/VVC标准采纳，形成一系列原创性技术，为公司积累了数百项视频相关的知识产权。公司多名成员出任了标准文本主编及编委、专家小组Chair、核心实验负责人、专题讨论组 Chair等职务，为标准研究工作做出积极贡献。

除了标准化工作的重要贡献之外，火山引擎团队在H.266/VVC商业化落地方向也做出了积极的贡献。早在2019年6月，团队已经完成满足点播场景应用的自研编码器BVC第一个版本，该版本与x265编码器相比，在相同计算资源下，针对大量的1080p高清视频，BVC编码器可以将平均码率下降33%。之后，BVC经过不断迭代，性能又有大幅提升。除了自研编码器，自研解码器也已经配套完成，在高端手机上可以做到实时流畅播放高清和超高清视频。相信BVC的出现和不断迭代必将会给用户带来更好的视频体验。

H.266标准之后，视频领域的专家仍在继续对视频编码技术进行探索。两个方向的探索值得注意，一个是基于新兴深度学习的视频压缩（也包括基于深度学习和传统混合视频编码框架的结合），另外一个是基于传统混合视频编码框架技术的继续挖掘。尽管目前都还只是刚刚开始，我们已经看到不少突破性的进展，比如火山引擎团队提出的一种基于深度学习的自适应滤波器DAM算法已经可以获取15%以上的性能增益；而基于混合视频编码框架的多项技术提升合在一起，可以带来超过13%以上的性能增益。国际标准化组织JVET已经确立了新的参考软件平台ECM，火山引擎的团队也担任了专家小组Co-chair、软件Co-chair、核心实验负责人等重要角色。相信在不久的将来，通过各位专家的持续努力，我们会看到越来越多的新技术的涌现。当再次看到30%左右的压缩性能提高的时候，咱们就可以准备聆听下一代视频编码标准正式启动的钟声了。

**LiveVideoStack:** 随着视频交互逐渐成为一种交流方式，人们对视频画质的要求也越来越高，火山引擎视频云在画质评估方面有哪些经验？

何予：目前常用的画质评估算法（如：SSIM/VMAF等全参考指标），仅适用于转码等业务，且在UGC视频场景表现相对乏力，无法满足实际业务的需求。因此，为了更加准确地评测视频主观画质，依托于抖音、西瓜等视频数据，火山引擎通过海量的众包人工标注精准收集主观标注分数，构建大规模UGC场景质量评估数据集，并基于业界领先的轻量化深度学习算法，开发了**VQScore画质评估算法体系**。VQScore包含无参/全参考清晰度指标，以及噪声、色彩等多维度具有可解释性的细指标。VQScore画质评估体系有几大特点：无参清晰度指标可以用于**监控视频全链路**的画质情况，包含用户投稿侧、前处理增强、服务端转码和播放端等多个节点，能够有效地通过不同维度分析对比展现大盘的画质变化差异，为A/B实验等提供可靠的画质参考信息；针对**UGC视频细分赛道**，在点播、直播和视频会议等不同应用场景设计了不同的画质评估算法；通过结合播放设备的屏幕信息，实现**跨平台适配**，这种device-adaptive VQScore能够准确估计播放端视频播放实际用户体验，补齐端到端链路的最后一环，为客户端播放策略提供更有效的参考信息。最后，画质评估的结论与画质增强、去噪、转码等策略相结合，使用更少的计算资源获取更加明显的画质收益。

VQScore画质评估体系除了在火山引擎得以大规模应用和验证之外，在多媒体领域世界顶级学术会议ICME 2021的“压缩UGC视频质量评估”比赛中，拿下了该比赛“无参考视频质量评价（NR-VQA）MOS赛道”的第一名，也得到了更加广泛的认可和关注度。

**LiveVideoStack**: 过去几年，实时音视频为直播、社交、在线教育等行业带来了巨大增长。火山引擎视频云在RTC技术的应用上取得了哪些进展？对于AI与RTC技术的结合，火山目前做了哪些布局？

何予：

一、火山引擎RTC取得的技术进展包括：

- 1、自研音频编码算法**nico**，更低的编解码复杂度，更强的弱网抗性，可做到50%丢包无感知恢复，最高可对抗80%丢包。
- 2、精细化设备性能自适应，让更多的低端机可以使用实时音视频能力，让高端机可以享受更好的视听体验。
- 3、与业务一起打磨出来的各场景最佳实践，针对互娱、教育、游戏、会议四大行业的十多个主要场景，沉淀出了一系列可复制的实践方案。

基于这些技术进展，目前火山引擎RTC的月均通话分钟数突破**100亿分钟**。

二、RTC与AI结合的布局

火山引擎有非常丰富且优秀的视频与音频的AI产品，而RTC和视频云则相当于是音视频的基础建设，因此在火山引擎，RTC与AI自然而然产生了很多结合点，包括在端上编码前处理和解码后处理，以及在云端解码处理等环节，都已经产生了很多结合AI的应用，比如视频的美颜、特效、手势识别，音频的美声、变声、ASR等。未来RTC与AI的结合应该更自然且紧密，为行业带来更多的新玩法。

**LiveVideoStack**: 现在有这样一种情况：短视频创作者辛辛苦苦地创作出来的视频很容易被他人盗取，发布到其他平台上，获取流量。这对于原创作者来说，就很不公平。所以在应对盗版方面，火山引擎视频云有没有哪些举措？

何予：火山引擎视频云主要在获取视频地址/视频播放两个阶段分别针对隐私/敏感/版权内容的访问提供安全保障解决方案。

- 1、视频地址访问鉴权：通过基于IAM AK/SK 的业务级授权以及基于用户ticket的用户级授权的两层权限控制，实现不同粒度的访问鉴权。
- 2、视频播放：针对视频播放，火山引擎视频云主要从链接访问控制以及DRM提供内容保护解决方案。

链接访问控制：

•隐私地址防扩散：防止URL通过简单copy 传播方式传播，需要通过cookie/header/passport 传递authtoken 进行签名验证之后才能播放。

•源站地址防扩散：隐私地址访问时（请求下载视频时）进行鉴权，通过业务二次鉴权以及源站鉴权等逻辑对访问请求进行管控。

### DRM解决方案：

•内容防扩散：通过点播服务端签发播放凭证PlayAuthToken以及DRM Auth Token，客户端点播SDK获取凭证密钥后，获取点播真实播放地址，并对播放内容进行解密。

**LiveVideoStack:** 从疫情开始，直播行业像是进入快车道，发展势头迅猛，直播的各种场景也在不断丰富中，比如KTV歌房、直播答题、互动游戏、连麦、电商、拍卖、企业直播等。火山的音视频引擎如何有效支持直播业务的不断增加的场景变化？未来还将有哪些直播新玩法？

何予：确实直播的热度在持续提升，疫情更是起到了加速的作用。过往火山音视频引擎支持抖音尝试了很多玩法，在这背后我们从几个方面来重点支持：

1、**底层组件化设计**：把音质、画质、特效、连麦等等能力抽象成底层组件来解耦，保障每个组件可以敏捷迭代，并让专门的团队去做到极致。所以表面上是音视频引擎，其实背后有很多团队在强有力的支持。

2、**靠近业务的解决方案**：直播客户端专门设立解决方案团队，会按照场景进行组件融合，降低上层业务接入门槛，只要调用几个简单的接口即可实现KTV、一起看等玩法。让业务更专注在上侧逻辑，直播解决方案去优化各种音视频能力组合的细节，做到整体最优解。

3、**数据驱动策略**：依赖字节的数据中台能力，把每一种场景，用户画像做到数据标签化，并通过AVB实验平台做大量实验，持续根据反馈取得最佳的直播体验。

相信以上火山音视频引擎的端到端方案，也很快会给到大家，从而加速行业的直播能力升级。

从直播业务演进来看，目前直播业务进入了多样化生态拓展的2.0阶段，直播除了秀场、电商、游戏之外，会探索更多的可能性，比如生活类、赛事、买房、买车等等，我们内部也在探索全新的直播技术升级，比如6dof自由视角、光场视频等可提高内容交互性的前瞻技术，相信很快也会陆续使用起来，敬请期待。

**LiveVideoStack:** 音视频行业的蓬勃发展也吸引了很多其他技术领域的程序员想要转行，对于这些有编程经验，但是对音视频技术不太熟悉的人，您有什么建议吗？

何予：对于有计算机理论和工程基础的同学来说，多媒体的门槛并不高，有很多开源工程、博客、书籍都是很好的入门手段，可以根据自身的兴趣和偏好来选择最合适的学习路径。

这里给大家提供一个我们自己总结提炼出的音视频技术进阶路线，希望能够对大家有所帮助：**音视频 - IT技术知识开源图谱**<https://codechina.gitcode.host/developer-roadmap/av/intro/>。

另外在掘金社区上有一个我们部门的技术专栏，我们的同学会不定期地在上面分享一些技术文章，有兴趣的同学也可以加关注和了解：

<https://juejin.cn/team/6943035964193243150/posts>。

同时，针对开发者，我们也有“火山引擎开发者社区”的公众号，大家也可以在上面浏览到相关的技术文章，了解最新技术活动。

我想更重要的是，加入到一个从事多媒体应用的工程技术团队，活学活用，在实践中更深入地理解技术是如何服务于应用场景的，并在迭代中提升自己的认知和技能。

**LiveVideoStack:** 本次大会火山引擎专场上将会有哪些精彩内容？您可以提前透露一下吗？

何予：本次专场上，火山引擎将会揭秘我们是如何基于自身的业务场景，打磨产品能力，支撑业务的发展；在有限的码率下，设计高效的视频编码压缩方案，为用户提供最佳的视频画质体验；然后聚焦讨论通过性能体验优化促进业务增长的实践；最后，以抖音等业务为例，介绍RTC技术如何追求极致的体验以及直播丰富的玩法。



2021  
北京

新技术  
TECHNOLOGY  
NEW 新机会  
OPPORTUNITY

# 专题 #

# 字节跳动背后的 音视频技术揭秘

2021.10.29 14:00-18:00

出品人

**何予** 火山引擎 视频云架构技术总监

14:00~14:45

《视频编码标准、优化和画质评估》

- 刘明 火山引擎 资深研究员
- 罗亚 火山引擎 资深研究员

15:00~15:45

《抖音背后的RTC是如何追求极致的》

- 朱利安 火山引擎 实时音视频 (RTC) 产品负责人

16:00~16:45

《从体验出发，构建以增长为目标的视频服务体系》

- 浩铭 火山引擎 点播技术研发负责人

17:00~17:45

《直播新玩法背后的音视频技术演进》

- 徐鸿 火山引擎 直播客户端SDK研发负责人



LiveVideoStack  
— 音视频技术社区 —

火山引擎



扫描二维码，了解更多详细信息

2021年10月29-30日  
北京丽亭华苑酒店

扫描图中二维码或点击阅读原文

了解大会更多信息



2021  
北京

NEW 新技术 新机会  
TECHNOLOGY OPPORTUNITY



刘明

火山引擎  
资深研究员



罗亚

火山引擎  
资深研究员



朱利安

火山引擎  
实时音视频(RTC)产品负责人



浩铭

火山引擎  
点播技术研发负责人



徐鸿

火山引擎  
直播客户端SDK研发负责人

2021年10月29-30日  
北京丽亭华苑酒店



了解更多演讲内容  
与大会日程相关信息

喜欢我们的内容就点个“在看”吧！

