

视频点播

快速入门

快速入门

开始使用视频点播

概述

本文档将帮助您快速开始使用视频点播VOD服务。

前提条件

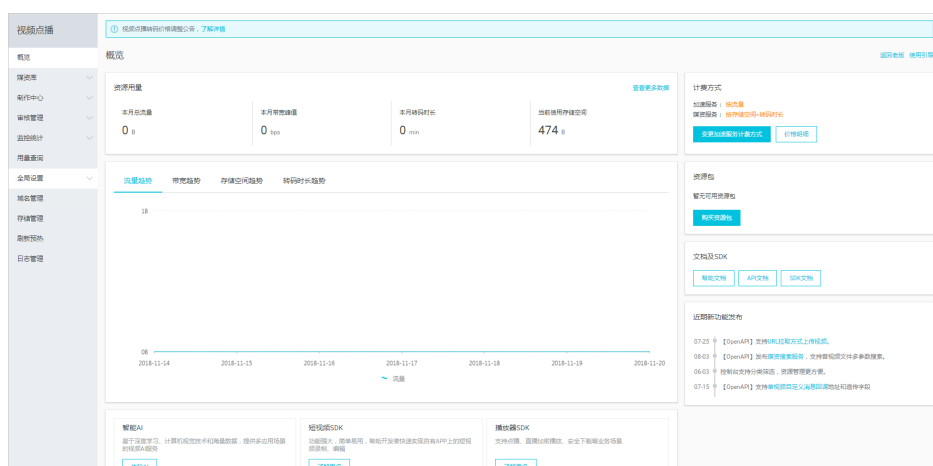
在使用阿里云VOD服务之前，请确保您已经注册了阿里云账号并完成实名认证。如果您还没有创建阿里云账号，系统会在您开通VOD时提示您注册账号。

开通服务

1. 访问阿里云官网，在右上角单击**登录**。
2. 在阿里云官网进入VOD产品详情页，单击**立即开通**。
3. 在VOD服务订购页面，单击**立即开通**，选择加速服务计费方式并确认订单。



4. 开通完成后，即可进入VOD控制台**概览**页面，开始使用点播服务。



基础配置

在您开始使用服务前，需要完成一些基础的配置，以下步骤将指导您进行处理。

管理存储

VOD提供存储服务，会默认帮您分配一个存储bucket，默认区域为华东2(上海)，无需任何配置即可进行上传和媒体资源管理，如您对存储区域有要求可再进行添加，目前VOD服务支持华北2(北京)和华东2(上海)两个服务中心，操作请参考 [存储管理文档](#)

域名管理

视频相对于一般的图文内容，文件体积更大，在传播过程对传输能力的体验要求更高，点播CDN服务可将您的资源缓存至阿里云遍布全球的加速节点上。当终端用户请求访问和获取这些资源时，无需回源，系统将就近调用CDN节点上已经缓存的资源，提高播放的速度，并提供稳定的流畅体验，添加完成加速域名即可开启点播加速服务。

添加域名，进入VOD控制台，添加用于视频分发与加速的自有域名。

注意：请确保该域名已经备案，并拥有使用权。

视频点播

概览

媒资库

制作中心

审核管理

监控统计

用量查询

全局设置

域名管理

存储管理

刷新预热

日志管理

< 添加域名

* 域名

请输入单个域名

暂不支持添加泛域名，如* .test.com * [了解更多](#)

* 源站信息

类型

OSS域名

IP

源站域名

点播服务默认Bucket

OSS作为源站，为您节省更多回源流量费用

* 端口

80端口

443端口

8080端口

* 加速区域

中国大陆

海外加速和国内加速价格有差别，如需开通海外加速请工单进行申请 [了解更多](#)

取消

提交

CNAME解析处理，添加域名后可获得阿里云加速域名CNAME地址。完成CNAME配置后，即可完成域名添加操作。

注意：请根据提示信息，到您的域名解析商处完成CNAME绑定。如果使用万网、新网或DNSPod域名，域名绑定操作参考 CNAME绑定。

业务配置

根据视频业务的特点和不同客户的业务要求，在基础配置外，还需要进行必要的面向业务场景的配置服务，您可以根据实际需要进行配置

设置转码模版

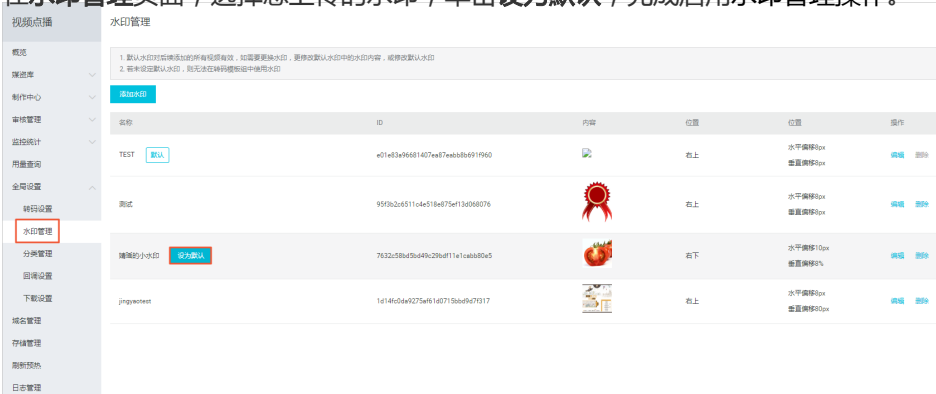
转码即将一个音视频文件转换成另一个或多个不同清晰度等条件音视频文件，以适应不同网络带宽、不同终端设备和不同的用户需求，您可以根据实际业务场景需求，进行转码模版配置

- 选择**全局设置 > 转码设置**，单击**添加转码模板组**。
- 在**视频转码模板组**页面，根据您的业务需求选择封装格式和清晰度。
- 如您不需要对内容进行转码处理，上传视频时直接选择不转码模版即可



添加水印

- 选择**全局设置 > 水印管理**，单击**添加水印**。
- 在**水印**页面，上传一张PNG图片，您可自定义相关参数，单击**保存**。
- 在**水印管理**页面，选择您上传的水印，单击**设为默认**，完成启用水印管理操作。



管理分类

为便于您管理内容，可以对内容设置分类，选择**全局设置 > 分类管理**，在**分类管理**页面进行分类管

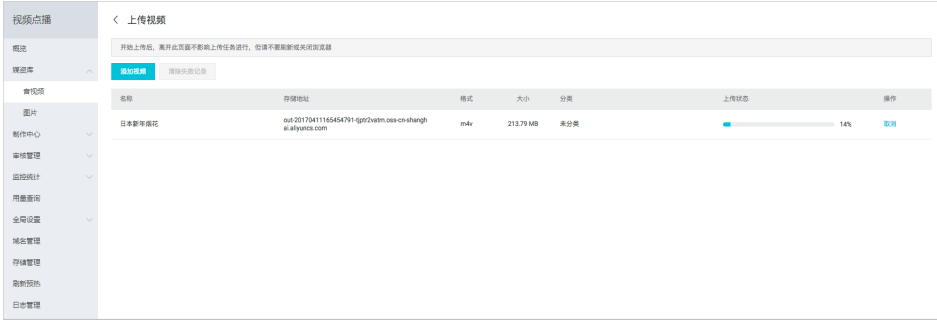


完成上述配置之后，即可在控制台进行视频文件的上传

- 在上传音视频页面，您可以根据需要连续添加多个音视频或选择移除某个音视频，单击开始上传。



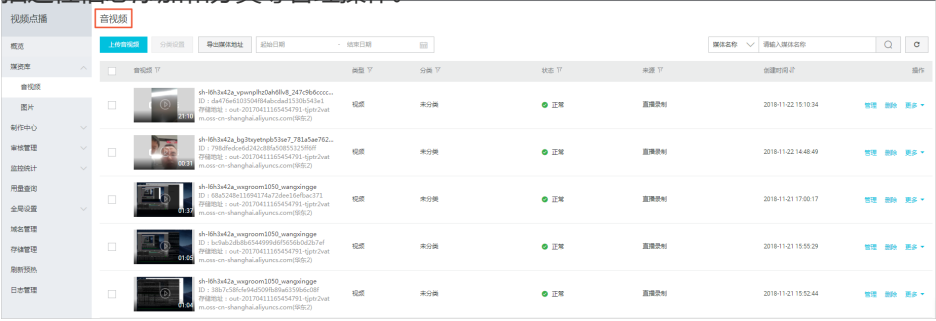
5



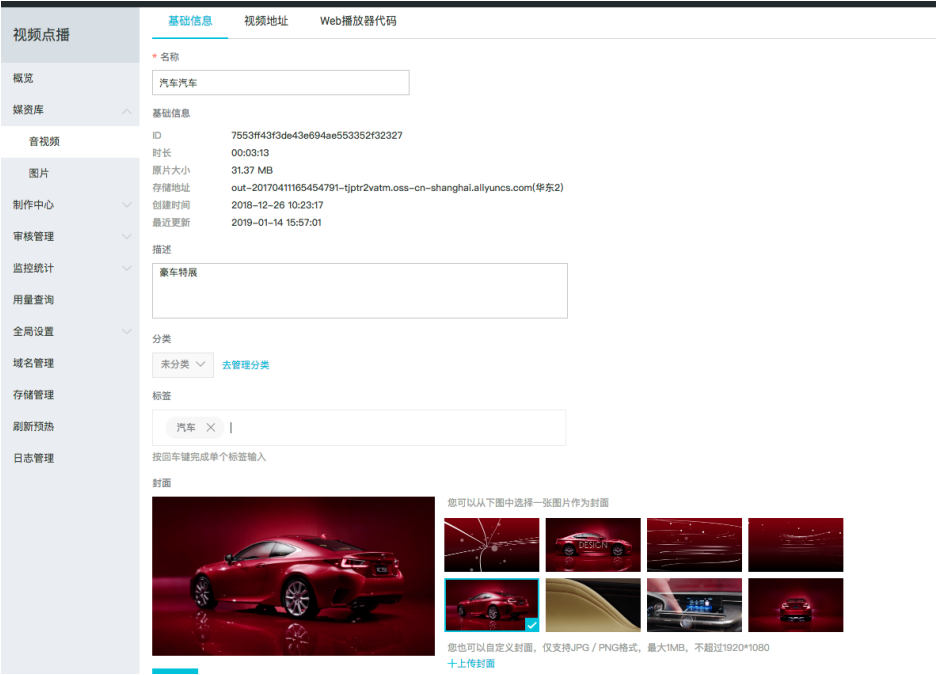
说明：在上传阶段，您可以取消当前音视频的上传操作或重新上传。在上传视频页面，您可以继续上传其它音视频。

管理资源

上传完成后，资源统一在媒资库呈现，如在音视频页面，可查看音视频的处理状态，并对资源的进行描述性信息添加和分类等管理操作。



单击目标音视频对应的管理，查看和编辑视频的详细信息，包括标题、描述、分类、标签等。



转码结束后，单击视频可进行预览，且支持清晰度切换。



使用限制

点播服务对资源使用、API的调用次数等都有默认限制，如要提高限制，可联系售后，或提交工单联系我们，并请详细说明使用场景和预期的上限（如域名数、播放接口访问频次等）。

点播资源使用限制

资源类型	使用限制	限制场景
域名	单用户不超过20个，域名必须已备案完成	控制台添加域名限制
刷新预热缓存	每日刷新URL上限2000，刷新目录上限100，预热URL上限	控制台和调用接口刷新预热限制

	500（预热不支持目录）	
存储	系统bucket单区域不超过1个	控制台添加存储限制
分类	单用户最多支持三级分类，每个分类最多支持100个子分类	控制台创建分类限制
自定义标签	单视频不超过16个	控制台视频管理时限制
工作流	单用户不超过20个	暂无
转码模板组	单用户不超过20个	控制台转码设置时限制
审核、剪辑、截图模板	单用户不超过20个	控制台转码设置时限制
转码模板组关联的模板数	单模板组不超过20个	控制台转码设置时限制
水印	单用户不超过20个	控制台添加水印时限制
模板关联的水印数	单模板不超过5个	控制台转码设置时限制
视频剪辑	只能编辑相同存储地域的媒体文件	剪辑视频时限制
事件回调	单用户大区内不超过1个	控制台设置回调时限制

服务端接口访问限制

点播服务端接口默认有访问的频次限制，如超过访问次数会拒绝访问。

超过阈值会随机拒绝访问，返回的HTTP状态码为400，错误码：Throttling.User，错误信息：Request was denied due to user flow control.

API维度限制

接口分类	频次限制(次/秒)	接口名称
媒体上传	100	CreateUploadVideo、RefreshUploadVideo、CreateUploadImage、CreateUploadAttachedMedia、UploadMediaByURL、RegisterMedia
音视频播放	300	GetPlayInfo、GetVideoPlayAuth
高频查询	100	媒资查询、AI查询等API，如SearchMedia、GetVideoInfo、GetMezzanineInfo 等
配置管理	10	转码模板、截图模板、视频水印、AI模板、域名管理等配置类API，如AddTranscodeTemplateGrou

		p 等
其它接口	30	提交转码/截图/AI处理/媒体审核、修改/删除视频信息、剪辑合成等API，如 DeleteVideo、SubmitTranscodeJobs、ProduceEditingProjectVideo 等

IP维度限制

单IP默认访问限制：300次/秒，请注意是以互联网出口IP为准。

词汇表

- 基础术语：文件格式、封装格式、编码方式 文件格式封装格式编码方式
- 转码基础术语 转码转封装
- 视频编码术语 码率、码流分辨率帧率GOP（关键帧间隔）IDR 帧对齐编码级别比特率码率控制方法
- 音频编码术语 编码格式采样率码率声道、声道数
- 其它术语 UTC时间（ISO 8601标准时间格式）

基础术语：文件格式、封装格式、编码方式

在开始了解和使用视频点播产品之前，先介绍三个基础概念：视频文件格式、封装格式、编码方式。

文件格式

我们知道操作系统中的文件名都有后缀，即扩展名，例如1.doc，2.jpg，3.avi等。设置扩展名的目的是让系统中的应用程序来识别并关联这些文件，让相应的文件由相应的应用程序打开。常见的文件格式如1.avi，2.mpg，3.mp4这些都叫做视频的文件格式，可由电脑上安装的视频播放器关联和打开（播放）。

封装格式

封装格式（Format），也称多媒体容器（Multimedia Container），是将已编码压缩好的视频轨道、音频轨道和元数据（视频基本信息如标题、字幕等）按照一定的格式规范，打包放到一个文件中，形成特定文件格式的视频文件。

封装格式主要分为两大类：面向存储的和面向流媒体的。

- 面向存储的，常见的有AVI、ASF（WMA/WMV）、MP4、MKV、RMVB（RM/RA）等；

- 面向流媒体的，常见的有FLV、TS（需要配合流媒体网络传输协议，如HLS、RTMP等），MP4也支持流媒体方式（配合HTTP等）。

下面以流媒体传输协议的视角重点介绍面向流媒体的封装格式：

- MP4：经典的视频封装格式，移动端（iOS/Android）、PC Web多终端都能良好支持。但MP4的视频文件头太大，结构复杂；如果视频较长（如数小时），则其文件头会过大，影响视频加载速度，故更适合短视频场景。

MP4由一个个的box（以前叫atom）组成，所有的Metadata（媒体描述元数据），包括定义媒体的排列和时间信息的数据都包含在这样的一些结构box中。Metadata对媒体数据（比如视频帧）引用说明，而媒体数据在这些引用文件中的排列关系全部在第一个主文件中的metadata描述，这样就会导致视频时长越大文件头就会越大、加载越慢。

- HLS（HTTP Live Streaming）：苹果公司推出的基于HTTP的流媒体网络传输协议，视频的默认封装格式是TS，除了多个TS分片文件，还定义了用来控制播放的m3u8索引文件（文本文件），可以规避普通MP4长时间缓冲头部数据的问题，比较适合点播场景。移动端(iOS/Android)支持较好，但PC端IE存在兼容性问题依赖播放器的二次开发（建议使用阿里云Web播放器）。
- FLV：Adobe公司推出的标准，在PC端有Flash的强力支持，但在移动终端只有APP实现播放器才能支持（建议使用阿里云播放器），大部分手机端浏览器均不支持，特别是苹果的移动设备都不支持。
- DASH（Dynamic Adaptive Streaming over HTTP）：使用fragmented MP4(fMP4)格式，将MP4视频分割为多个分片，每个分片可以存在不同的编码形式（如分辨率、码率等）；播放器端可自由选择需要播放的视频分片，实现自适应多码率、不同画质内容的无缝切换，提供更好的播放体验。其中MPD文件类似HLS的m3u8文件，国外视频网站如YouTube、Netflix等较多使用DASH。
- HLS+fMP4（HTTP Live Streaming with fragmented MP4）：此处单列，但本质上还是HLS协议。苹果公司于WWDC 2016宣布新的HLS标准支持文件封装格式为fragmented MP4，使用方法与TS分片类似，意味着一次转码可同时打包成DASH和HLS。

HLS（包括HLS+fMP4）和DASH是最常用的自适应流媒体传输技术（Adaptive Video Streaming），推荐使用。

编码方式

视频编码方式（Codec）是指能够对数字视频进行压缩或解压缩（视频解码）的程序或者设备。通常这种压缩属于有损数据压缩。也可以指通过特定的压缩技术，将某个视频格式转换成另一种视频格式。常见的编码方式有：

H.26X系列：由ITU（国际电信联盟）主导，包括H.261、H.262、H.263、H.264、H.265。

- H.261：主要在老的视频会议和视频电话产品中使用。
- H.263：主要用在视频会议、视频电话和网络视频上。
- H.264：H.264/MPEG-4第十部分，或称AVC（Advanced Video Coding，高级视频编码），是一种视频压缩标准，一种被广泛使用的高精度视频的录制、压缩和发布格式。
- H.265：高效率视频编码（High Efficiency Video Coding，简称HEVC）是一种视频压缩

标准，H.264/MPEG-4 AVC的继任者。HEVC不仅提升图像质量，同时也能达到H.264/MPEG-4 AVC两倍的压缩率（等同于同样画面质量下码率减少50%），可支持4K分辨率甚至超高画质电视，最高分辨率可达8192×4320（8K分辨率），这是目前发展的趋势。

MPEG系列：由ISO（国际标准组织机构）下属的MPEG（运动图象专家组）主导，视频编码方面主要有：

- MPEG-1第二部分：主要使用在VCD上，有些在线视频也使用这种格式，该编解码器的质量大致上和原有的VHS录像带相当。
- MPEG-2第二部分：等同于H.262，使用在DVD、SVCD和大多数数字视频广播系统和有线分布系统（Cable Distribution Systems）中。
- MPEG-4第二部分：可以使用在网络传输、广播和媒体存储上，比起MPEG-2和第一版的H.263，它的压缩性能有所提高。
- MPEG-4第十部分：技术上和ITU-T H.264是相同的标准，二者合作，诞生了H.264/AVC标准，ITU-T将其命名为H.264，而ISO/IEC称它为MPEG-4高级视频编码（Advanced Video Coding，AVC）。

AVS（Audio Video coding Standard）：我国自主知识产权的信源编码标准，是《信息技术 先进音视频编码》系列标准的简称，目前已完成两代AVS标准的制定。

- 第一代AVS标准包括国家标准《信息技术 先进音视频编码 第2部分：视频》（简称AVS1）和《信息技术 先进音视频编码 第16部分：广播电视视频》（简称AVS+）。AVS+的压缩效率与国际同类标准H.264/AVC最高档次（High Profile）相当。
- 第二代AVS标准，简称AVS2，首要应用目标是超高清清晰度视频，支持超高分辨率（4K以上）、高动态范围视频的高效压缩。AVS2的压缩效率比上一代标准AVS+和H.264/AVC提高了一倍，超过国际同类型标准HEVC/H.265。

其他系列，如，VP8、VP9（Google 主导），RealVideo（RealNetworks推出）等编码方式，在互联网视频使用较少，此处不再介绍。

选择编码方式要充分考虑播放终端（如移动端APP、Web浏览器等）的兼容性，尽量使用最常见和广泛支持的。阿里云视频点播支持视频编码格式：H.264/AVC（默认）、H.265/HEVC，音频编码格式：MP3（默认）、AAC、VORBIS、FLAC。

转码基础术语

转码

视频转码（Video Transcoding）是指将已经压缩编码的视频码流转换成另一个视频码流，以适应不同的网络带宽、不同的终端处理能力和不同的用户需求。转码本质上是一个先解码、再编码的过程，因此转换前后的码流可能遵循相同的视频编码标准，也可能不遵循相同的视频编码标准。

转封装

转封装指的是将视频或音频的封装格式进行转换，如将AVI的视频转换为MP4，其间并不会进行音视频的编码和解码工作，而是直接将视频和音频压缩码流从一种封装格式文件中获取出来然后打包成另一种封装格式的文件。相比转码，转封装有两大特点：

- 处理速度极快。音视频编解码过程十分复杂，占据了转码的绝大部分时间。转封装不需要进行编码和解码，节约了大量的处理时间。
- 音视频质量无损。没有解码（解压缩）和编码（压缩）过程，所以不会有音视频的压缩损伤。

转封装后的文件与原始文件的分辨率、码率等几乎一致，故播放时也称其为“原画”。

视频编码术语

码率、码流

码率（Bitrate）是指视频文件在单位时间内使用的数据流量，也叫码流或码流率，是视频编码中画面质量控制最重要的部分。量度单位为“比特每秒”（bit/s或bps），常使用Kbps（每秒多少个比特）或Mbps。一般来说同样分辨率下，视频文件的码率越大，压缩比就越小，画面质量就越高。码率越大，说明单位时间内取样率越大，数据流精度就越高，处理出来的文件就越接近原始文件，图像质量越好，画质越清晰，要求播放设备的解码能力也越高。

当然，码率越大，文件体积也越大，其计算公式是文件体积=时间X码率/8。例如，网络上常见的一部60分钟的码率为1Mbps的720P的视频文件，其体积就大概为3600秒×1Mb/8=450MB。

分辨率

分辨率（Resolution）用来描述视频对细节的分辨能力，通常表示为每一个方向上的像素数量，比如1280x720等。分辨率决定了视频画面细节的精细程度。通常情况下，视频的分辨率越高，所包含的像素就越多，画面就越清晰。

分辨率是决定码率的主要因素，不同的分辨率要采用不同的码率。总体而言，视频的分辨率越高，所要求的码率也越大，但并不总是如此，不同分辨率都有合理的码率选择范围。所谓“合理的范围”指的是，如果低于这个范围，视频画面质量会很差；如果高于这个范围，画面提升有限甚至几乎无提升，且浪费网络流量和存储空间。

帧率

帧率（Frame Rate）是单位时间内视频显示帧数的量度单位，也就是每秒钟刷新的图片的帧数，量度单位为“每秒显示帧数”（Frame Per Second，FPS）或“赫兹”。

高的帧率可以得到更流畅、更逼真的画面效果。一般来说25~30fps就可接受，提升至60fps则可以明显提

升交互感和逼真感，但一般来说超过75fps就不容易察觉到有明显的流畅度提升了。如果帧率超过屏幕刷新率只会浪费图形处理的能力，因为显示设备不能以这么快的速度更新，这样超过刷新率的帧率就浪费掉了。在分辨率不变的情况下，帧率越高，则对显卡的处理能力要求越高。

GOP（关键帧间隔）

GOP（Group of Pictures）是一组以 MPEG 编码的影片或视讯串流内部的连续图像，以 I 帧开头，到下一个 I 帧结束。一个 GOP 包含如下图像类型：

- I 帧(Intra Coded Picture)：又称帧内编码帧，为关键帧，是一种自带全部信息的独立帧，无需参考其他图像便可独立进行解码，可以简单理解为一张静态画面。视频序列中的第一个帧始终都是I 帧，每个 GOP 由I 帧开始。
- P 帧(Predictive Coded Picture)：又称帧间预测编码帧，需要参考前面的I帧才能进行编码。表示的是当前帧画面与前一帧（前一帧可能是I帧也可能是P帧）的差别。解码时需要用之前缓存的画面叠加上本帧定义的差别，生成最终画面。与I帧相比，P帧通常占用更少的数据位，但不足是，由于P帧对前面的P和I参考帧有着复杂的依赖性，因此对传输错误非常敏感。
- B 帧(Bidirectionally Predictive Coded Pictures)：又称双向预测编码帧，也就是B帧记录的是本帧与前后帧的差别。也就是说要解码B帧，不仅要取得之前的缓存画面，还要解码之后的画面，通过前后画面的与本帧数据的叠加取得最终的画面。B帧压缩率高，但是对解码性能要求较高。

GOP值表示关键帧的间隔(即两个关键帧之间的帧数)，也就是两个IDR帧之间的距离，一个帧组的最大帧数。一般而言，每一秒视频至少需要使用 1 个关键帧。增加关键帧个数可改善视频质量，但会同时增加带宽和网络负载。GOP值（帧数）除以帧率即为时间间隔，如阿里云视频点播默认的GOP值为250帧，帧率为25fps，则时间间隔为10秒。

GOP值需要控制在合理范围，以平衡视频质量、文件大小（网络带宽）和seek效果（拖动、快进的响应速度）等：

- 加大GOP值有利于减小视频文件大小，但也不宜设置过大，太大则会导致GOP后部帧的画面失真，影响视频质量。
- GOP值也是影响视频seek响应速度的关键因素，seek时播放器需要定位到离指定位置最近的前一个关键帧，如果GOP太大意味着距离指定位置可能越远（需要解码的预测帧就越多）、seek响应的时间（缓冲时间）也越长。
- 由于P、B帧的复杂度大于I帧，GOP值过大，过多的P、B帧会影响编码效率，使编码效率降低。
- 但如果设置过小的GOP值，则需要提高视频的输出码率，以确保画面质量不会降低，故会增加网络带宽。

IDR 帧对齐

IDR帧（Instantaneous Decoding Refresh Picture），即时解码刷新帧，是 I 帧的一种。与普通 I 帧的区别在于，一个 IDR 帧之后的所有帧都不能引用该 IDR 帧之前的帧的内容；相反，对于普通的 I 帧，其后的 P 帧和 B 帧可以引用该普通 I 帧之前的其他 I 帧。在编码和解码中为了方便，将首个I帧和其他I帧区别开，称为IDR，这样就方便控制编码和解码流程。

IDR帧的作用是立刻刷新，使错误不致传播，从IDR帧开始，重新算一个新的序列开始编码。而普通I帧不具有随

机访问的能力，这个功能是由IDR承担。视频播放时，播放器一般都支持随机seek（拖动）到指定位置，而播放器直接选择到指定位置附近的IDR帧进行播放最为便捷，因为可以明确知道该IDR帧之后的所有帧都不会引用其之前的其他I帧，从而避免较为复杂的反向解析。

在对同一个视频进行多码率转码时，如果指定IDR帧对齐（IDR Frame Alignment），则意味着所有输出视频的IDR帧在时间点、帧内容方面都保持精确同步，此时播放器便可实现多码率视频平滑切换，从而不会出现较为明显的切换卡顿。

编码级别

编码档次（Profile）是针对特定应用场景的特定编码功能的集合。H.264规定了三种主要级别：

- Baseline：支持I/P帧，只支持无交错（Progressive）和CAVLC，一般用于低阶或需要额外容错的应用，比如视频通话、手机视频等即时通信领域；
- Main：提供I/P/B帧，支持无交错（Progressive）和交错（Interlaced），同样提供对于CAVLC和CABAC的支持，用于主流消费类电子产品规格如低解码（相对而言）的MP4、便携的视频播放器、PSP和iPod等；
- High：在Main的基础上增加了8x8内部预测、自定义量化、无损视频编码和更多的YUV格式（如4:4:4），用于广播及视频碟片存储（蓝光影片），高清电视的应用。

比特率

比特率（Bit rate）是指每秒传送的比特(bit)数，单位为bps(Bit Per Second)，比特率越高，传送的数据越大。在视频领域，**比特率等同于码率**。比特率表示经过编码（压缩）后的音、视频数据每秒钟需要用多少个比特来表示，而比特就是二进制里面最小的单位，要么是0，要么是1。与码率类似，比特率与音、视频压缩的关系，简单的说就是比特率越高，音、视频的质量就越好，但编码后的文件就越大；如果比特率越低则文件越小。

码率控制方法

码率控制是指视频编码中决定输出码率的过程，常用的码率控制方法有：

VBR（Variable Bitrate）：动态比特率，也就是没有固定的比特率，音视频压缩软件在压缩时根据音视频数据的复杂程度即时确定使用什么比特率，这是以质量为前提兼顾文件大小的方式。

CBR（Constant Bitrate）：固定比特率，指文件从头到尾都是一种位速率。相对于VBR和ABR来讲，它压缩出来的文件体积很大，而且视频质量相对于VBR和ABR不会有明显的提高。

ABR（Average Bitrate）：平均比特率，是VBR的一种插值参数。LAME针对CBR不佳的文件体积比和VBR生成文件大小不定的特点独创了这种编码模式。ABR在指定的文件大小内，以每50帧（30帧约1秒）为一段，低频和不敏感频率使用相对低的流量，高频和大动态表现时使用高流量，可以做为VBR和CBR的一种折衷选择。

ABR在一定的时间范围内达到设定的码率，但是局部码率峰值可以超过设定的码率，平均码率恒定。ABR是VBR的改良版，能确保输出平均码率在合理范围，且在这个范围内，还是动态根据复杂度编码

，也是阿里云默认的编码控制方式。

音频编码术语

编码格式

定义参考视频编码方式。主要包括无损编码和有损编码两种。根据采样定理，音频编码只能“无限接近”自然信号，故而所有的音频编码方式实质上都是有损的。在计算机领域中，通约定能够达到最高保真水平的 PCM 编码为无损编码。互联网中常见的音频编码均为有损编码，常见的编码格式有 MP3、AAC 等。

采样率

采样率（Sample Rate），也称为采样速度或者采样频率，定义了每秒从连续信号中提取并组成离散信号的采样个数，单位为赫兹(HZ)。采样率是指将模拟信号转换成数字信号时的采样频率，也就是单位时间内采样多少点，采样频率越高声音的还原就越真实越自然。

码率

参考视频编码的“码率”介绍。

声道、声道数

声道（Sound Channel）是指声音在录制（或播放）时，在不同空间位置采集（或播放）的相互独立的音频信号。所谓声道数，也就是声音录制时的音源数量或播放时的扬声器数量。

其它术语

UTC时间（ISO 8601标准时间格式）

UTC时间指的是协调世界时，又称世界统一时间、世界标准时间、国际协调时间。由于英文（CUT）和法文（TUC）的缩写不同，作为妥协，简称UTC。协调世界时是以原子时秒长为基础，在时刻上尽量接近于世界时的一种时间计量系统。中国大陆采用ISO 8601-1988的《数据元和交换格式信息交换日期和时间表示法》（GB/T 7408-1994）称之为国际协调时间。

如无特别指定，阿里云视频点播默认返回的时间字段或接口请求的时间参数都为UTC时间，时间格式按照ISO8601标准表示，格式为：YYYY-MM-DDThh:mm:ssZ。例如：2017-01-11T12:00:00Z 表示北京时间2017年1月11日20点0分0秒。注：北京时间与UTC的时差为+8，也就是UTC+8。