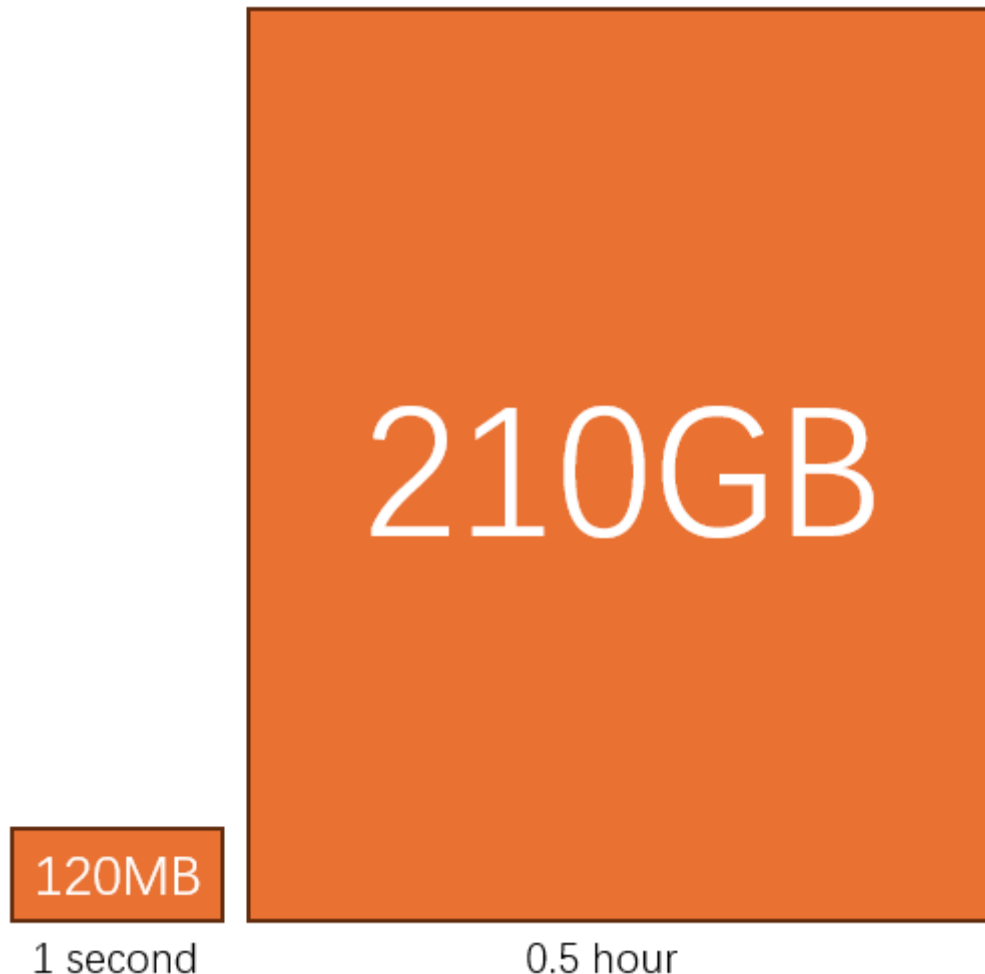


视频的编解码

视频是需要图像一张接着一张的快速切换而成的，那么也就意味着假设帧率为30FPS，每帧1080p的图像即有 $1920 \times 1080 = 2073600$ 像素，如果按照RGB565格式2Byte代表一个像素，就需要

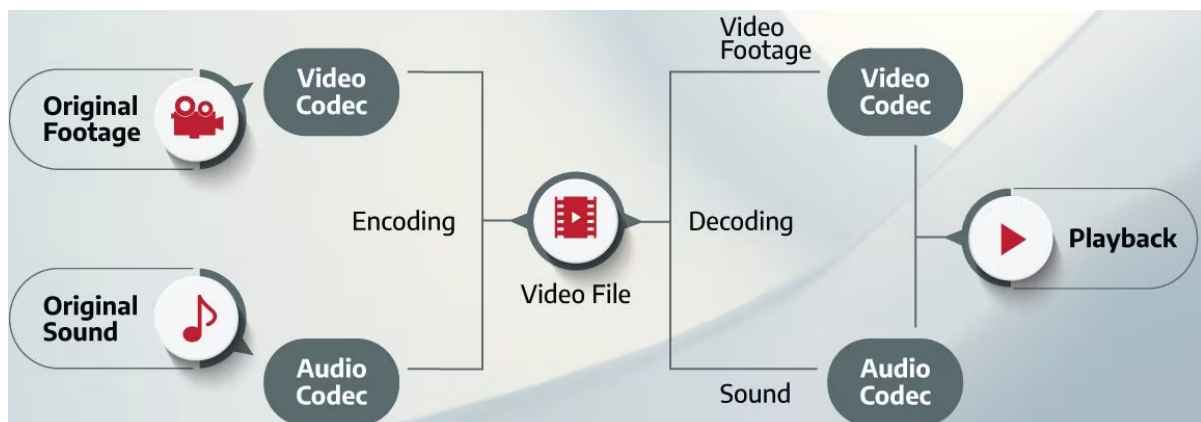
$2 \times 2073600 = 4147200$ (即4MB)，那么一秒钟30帧就需要120MB的空间，如果要播放一个30分钟的视频就需要 $120 \times 60 \times 30 = 210\text{GB}$ 的空间才能播放，那么这种方式在网上进行实时播放视频是不可行。



因为目前我们的网络的传输速率以 bit/s 计算，那么想要达成上面所要求的实时观看的效果，就需要 $120\text{MB} \times 8 = 960\text{Mb/s}$ 即960Mb每秒的带宽，这种接近千兆以太网的速率无法覆盖大部分场景。

为了将视频信号进行压缩，以减小数据量，视频压缩技术应运而生，这种技术又称为：“视频编码”。

视频经过压缩后传输到目标端侧后需要进行视频的解压缩，这种解压缩技术又称为：“视频解码”。



视频压缩编码是一种技术，用于减少视频文件的大小，同时尽可能保持其原始质量。这个过程涉及到多种算法和编解码器，它们利用视频数据中的冗余性和人眼对图像信息的感知特点来实现压缩。视频压缩编码的目标是降低传输成本，提高存储和传输效率，同时保证视频播放时的质量。

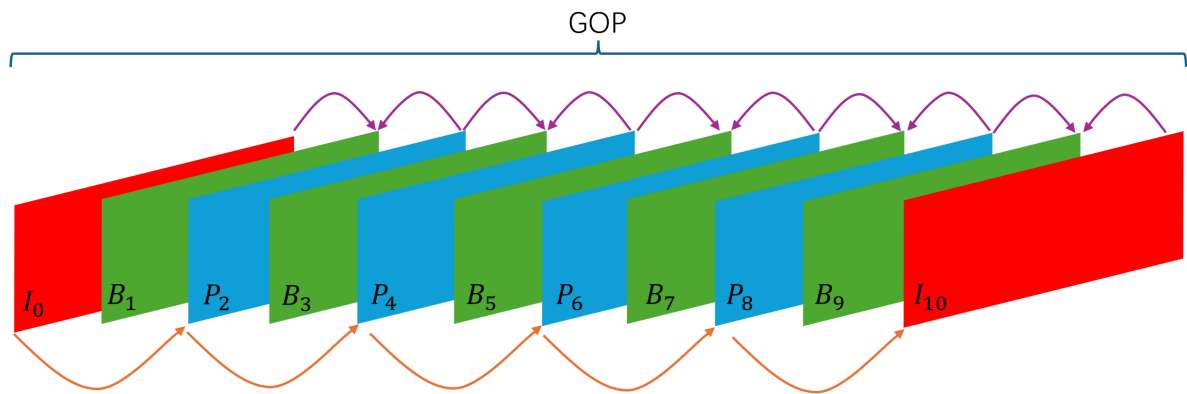
视频压缩编码的关键组成部分包括：

- 编解码器 (Codec)：一组算法，用于编码（压缩）和解码（解压缩）视频数据。
- 比特率 (Bitrate)：表示视频每秒传输的数据量，通常以kbps（千比特每秒）或mbps（兆比特每秒）计算。
- 分辨率 (Resolution)：视频的尺寸，通常以像素为单位，如1920x1080。
- 帧率 (Frame rate)：视频每秒显示的帧数。
- 常见的视频编解码器包括H.264/AVC、H.265/HEVC、H.266/VVC等

一般将YUV格式的视频转换为编码格式的过程称为编码。

在视频编码中会对图像进行分组，分组后的图像称为GOP(Group of Pictures)。其中每组图像就有三种图像帧，分别是：I帧、P帧和B帧。它们各自有不同的作用和特点：

- **I帧** (Intra-coded Frame)：这是关键帧，也称为帧内编码帧。I帧包含了完整的画面信息，不依赖于其他帧就能够独立解码。它通常用于视频序列的开始或者场景变换时，以提供一个清晰的参考点。I帧的压缩率相对较低，因为它需要存储更多的信息。
- **P帧** (Predictive Frame)：这是前向预测编码帧。P帧记录的是与前一个I帧或P帧之间的差异。在解码时，需要用前一个I帧或P帧的信息加上P帧本身的差异信息来重建完整的画面。P帧的压缩率比I帧高，因为它只需要存储与前一帧的变化信息。
- **B帧** (Bi-directional Predictive Frame)：这是双向预测编码帧。B帧记录的是与前后帧的差异。在解码时，需要用前一个I帧或P帧和后一个P帧的信息结合B帧本身的差异信息来重建完整的画面。B帧的压缩率最高，因为它利用了前后帧的信息来最大限度地减少存储的数据量。



视频解码 (Video Decoding) 则是将经过压缩编码的视频数据重新转换为可视化的视频信号的过程。解码器根据压缩编码标准对视频数据进行解码，还原出原始的视频图像，并将其显示在屏幕上或输出到其他设备上。视频解码通常由硬件或软件解码器完成。

视频解码的流程通常包括以下几个步骤：

1. **数据获取**：视频数据通常以文件或网络流的形式传输到解码器中。
2. **解析数据**：解码器首先需要解析视频数据的格式，确定视频的分辨率、帧率、编码方式等参数。

3. **解码压缩数据**：根据视频数据的编码方式，解码器使用相应的解码算法将压缩的视频数据转换成原始的像素数据。这个过程涉及到解码器根据编码标准的规则，解析和还原出视频数据中的各种信息，如运动矢量、残差信息等。
4. **图像重建**：解码器将解码后的像素数据组装成完整的视频图像。这包括解码出每一帧的图像，根据运动矢量和残差信息进行图像预测和补偿，然后将这些信息组合起来以生成完整的图像。
5. **图像显示或输出**：解码器将重建的视频图像传递给显示设备或其他输出设备，如显示器、电视机、或其他视频处理器，以供用户观看或进一步处理。

