

·V：RTP协议的版本号，占2位，当前协议版本号为2。

   ·P：填充标志，占1位，如果P=1，则在该报文的尾部将填充一个或多个额外的八位组，它们不是有效载荷的一部分。

    ·X：扩展标志，占1位，如果X=1，则在RTP报头后跟有一个扩展报头。

    ·CC：CSRC计数器，占 4位，指示CSRC 标识符的个数。

    ·M: 标记，占1位，不同的有效载荷有不同的含义，对于视频，标记一帧的结束；对于音频，标记会话的开始。

    ·PT: 有效载荷类型，占7位，用于说明RTP报文中有效载荷的类型，如GSM音频、JPEM图像等。

    ·序列号：占16位，用于标识发送者所发送的RTP报文的序列号，每发送一个报文，序列号增1。接收者通过序列号来检测报文丢失情况，重新排序报文，恢复数据。

    ·时戳 (Timestamp)：占32位，时戳反映了该RTP报文的第一个八位组的采样时刻。接收者使用时戳来计算延迟和延迟抖动，并进行同步控制。

    ·同步信源(SSRC)标识符：占32位，用于标识同步信源。该标识符是随机选择的，参加同一视频会议的两个同步信源不能有相同的SSRC。

    ·特约信源(CSRC)标识符：每个CSRC标识符占32位，可以有0～15个。每个CSRC标识了包含在该RTP报文有效载荷中的所有特约信源。

    这里的同步信源是指产生媒体流的信源，它通过RTP报头中的一个32位数字SSRC标识符来标识，而不依赖于网络地址，接收者将根据SSRC标识符来区分不同的信源，进行RTP报文的分组。特约信源是指当混合器接收到一个或多个同步信源的RTP报文后，经过混合处理产生一个新的组合RTP报文，并把混合器作为组合RTP报文的 SSRC，而将原来所有的SSRC都作为CSRC传送给接收者，使接收者知道组成组合报文的各个SSRC。

    在发送端，上层应用程序以分组形式将编码后的媒体数据传给RTP通信模块，作为RTP报文的有效载荷，RTP通信模块将根据上层应用提供的参数在有效载荷前添加RTP报头，形成 RTP报文，通过Socket接口选择UDP协议发送出去。

    在接收端，RTP通信模块通过Socket接口接收到RTP报文后，将RTP报头分离出来作相应处理，再将RTP报文的有效载荷作为数据分组传递给上层应用。

    考虑到在Internet这种复杂的环境中举行视频会议，RTP定义了两种中间系统：混合器(Mixer)和转换器(Translator)。

    在Internet上举行视频会议时，可能有少数参加者通过低速链路与使用高速网络的多数参加者相连接。为了不强制所有会议参加者都使用低带宽和低质量的数据编码，RTP允许在低带宽区域附近使用混合器作为RTP级中继器。混合器从一个或多个信源接收RTP 报文，对到达的数据报文进行重新同步和重新组合，这些重组的数据流被混合成一个数据流，将数据编码转化为在低带宽上可用的类型，并通过低速链路向低带宽区域转发。为了对多个输入信源进行统一的同步，混合器在多个媒体流之间进行定时调整，产生它自己的定时同步，因此所有从混合器输出的报文都把混合器作为同步信源。为了保证接收者能够正确识别混合器处理前的原始报文发送者，混合器在RTP报头中设置了CSRC标识符队列，以标识那些产生混和报文的原始同步信源。

    在Internet环境中，一些会议的参加者可能被隔离在应用级防火墙的外面，这些参加者被禁止直接使用 IP组播地址进行访问，虽然他们可能是通过高速链路连接的。在这些情况下，RTP允许使用转换器作为RTP级中继器。在防火墙两端分别安装一个转换器，防火墙之外的转换器过滤所有接收到的组播报文，并通过一条安全的连接传送给防火墙之内的转换器，内部转换器将这些组播报文再转发送给内部网络中的组播组成员。