

- 连续播放。一旦视频开始播放，它应该根据初始记录的时序进行。因此，为了在客户端播放，必须从服务器中及时接收数据；否则，用户经历视频帧停滞（这时客户等待延迟的帧）或帧跳过（这时客户漏掉延迟的帧）。

到目前为止，对流式视频最重要的性能测量是平均吞吐量。为了提供连续的播放，网络为流式应用提供的平均吞吐量必须至少与该流视频本身的比特率一样大。如我们将在7.2节所见，通过使用缓存和预取，即使在吞吐量波动的时候，提供连续播放也是可能的，只要平均吞吐量（在5~10秒区间平均）保持在视频速率之上 [Wang 2008]。

对于许多流式视频应用，预先录制的视频被存储起来，并且从CDN而非从单一的数据中心流式播放。也有许多P2P视频流式应用，其中视频被存储在用户主机（对等方）上，不同视频块从可能分布在全球的不同对等方到达。在得知了因特网流式视频的性能后，我们将在7.2节更加深入地研究流式视频，特别关注客户缓存、预取、对可用带宽的适应性质量和CDN分发。

## 2. 会话式IP语音和视频

在因特网上的实时会话式语音通常称为**因特网电话**（Internet telephony），因为从用户的角度看，它类似于传统的电路交换电话服务。它也常被称为**IP语音**（Voice-over-IP, VoIP）。会话式视频与之类似，除了它包括参与者的语音以及视频外。今天的大多数语音和视频会话式系统允许用户生成具有三个或更多个参与者的会议。会话式语音和视频广泛地应用于今天的因特网中，因特网公司Skype、QQ和Google Talk自称每天都有数亿用户。

在第2章有关应用服务需求的讨论中（图2-4），我们确定了一些轴，服务需求可以根据它们分类。其中的两个轴（即定时考虑和数据丢失容忍度）对会话式语音和视频应用尤其重要。定时考虑是很重要的，因为音频和视频会话式应用是高度**时延敏感**（delay-sensitive）的。对于具有两个或更多个交互讲话者的会话来说，从用户讲话或移动开始到该动作显示在其他端的时延应当小于几百毫秒。对于语音，小于150ms的时延不会被人类听者觉察到，150~400ms的时延能够被接受，当时延超过400ms时，即使不会使对话变得完全无法理解，也会使语音会话变得令人沮丧。

另一个方面，会话式多媒体应用**容忍丢包**（loss-tolerant），即偶尔的丢失只会在音频/视频回放时偶尔出现干扰信号，而且这些丢失经常可以部分或者全部地隐藏。这些时延敏感但容忍丢包的特性明显不同于那些弹性数据应用（如Web浏览、电子邮件、社交网络和远程注册等）的特性。对于这些弹性应用，长时延令人恼火，但并不是特别有害，然而传输数据的完全和完整性是首要的。我们将在7.3节中更加深入地探讨会话式语音和视频，特别关注适应性播放、前向纠错和差错掩盖是如何减缓网络引入的分组丢失和时延的。

## 3. 流式实况音频和视频

这种第三类应用类似于传统的电台广播和电视，只是它通过因特网来传输而已。这些应用允许用户接收从世界上任何角落发出的实况无线电广播和电视传输。今天有数以千计、遍及全球的无线电台和电视台正在因特网上广播内容。

实况是类似于广播的应用，它们经常有很多接收相同音频/视频节目的客户。尽管通过使用4.7节中描述的IP多播技术，能够有效地完成向多个接收方分发实况音频/视频，但是多播分发今天更多的是通过应用层多播（使用P2P网络或CDN）或通过多个独立的单播流来实现。由于使用流式存储多媒体，网络必须为每个实况多媒体流提供大于该视频