第 8 章　互联网上的音频/视频服务

本章首先对互联网提供音频/视频服务进行概述。然后介绍流式音频/视频中的媒体服务器和实时流式协议 RTSP，并以 IP 电话为例介绍交互式音频/视频所使用的一些协议，如实时运输协议 RTP、实时传送控制协议 RTCP、H.323 以及会话发起协议 SIP。接着讨论改进“尽最大努力交付”服务的一些措施，包括怎样使互联网能够提高服务质量，并介绍综合服务 IntServ、资源预留协议 RSVP 和区分服务 DiffServ 的要点。

本章最重要的内容是：

(1) 多媒体信息的特点（如时延和时延抖动，播放时延等）。

(2) 流媒体的概念。

(3) IP 电话使用的几种协议。

(4) 改进“尽最大努力交付”服务的几种方法。

8.1　概述

计算机网络最初是为传送数据设计的。互联网 IP 层提供的“**尽最大努力交付**”服务以及**每一个分组独立交付**的策略，对传送数据信息十分合适。互联网使用的 TCP 协议可以很好地解决 IP 层不能提供可靠交付这一问题。

然而技术的进步使许多用户开始利用互联网传送音频/视频信息。在许多情况下，这种音频/视频常称为**多媒体信息**①。本来电路交换的公用电话网传送话音和多媒体信息早已是成熟的技术。例如视频会议（又称为电视会议）原先是使用电路交换的公用电话网。使用电路交换的好处是：一旦连接建立了（也就是只要拨通了电话），各种信号在电话线路上的**传输质量就有保证**。但使用公用电话网的缺点是**价格太高**。因此要想办法改用互联网。

多媒体信息（包括声音和图像信息）与不包括声音和图像的数据信息有很大的区别，其中最主要的两个特点如下。

第一，**多媒体信息的信息量往往很大。**

含有音频或视频的多媒体信息的信息量一般都很大，下面是简单的说明。

对于电话的声音信息，如采用标准的 PCM 编码（8 kHz 速率采样），而每一个采样脉冲用 8 位编码，则得出的声音信号的数据率就是 64 kbit/s。对于高质量的立体声音乐 CD 信息，虽然它也使用 PCM 编码，但其采样速率为 44.1 kHz，而每一个采样脉冲用 16 位编码，因此这种双声道立体声音乐信号的数据率超过了 1.4 Mbit/s。

---

① 注：多媒体信息和传统数据信息不同，它是指内容上相互关联的文本、图形、图像、声音、动画和活动图像等所形成的复合数据信息。而多媒体业务则应有集成性、交互性和同步性的特点。集成性是指对多媒体信息进行存储、传输、处理、显示的能力，交互性是指人与多媒体业务系统之间的相互控制能力，同步性是指在多媒体业务终端上显示的图像、声音和文字是以同步方式工作的。在本章中，我们经常把音频/视频信息和多媒体信息作为同义词来使用，虽然它们并不严格地等同。