

第一章 介绍

40 年代末，一个偏爱雪茄烟的英国人曾引用了一个著名且常常被重复的短语“喋喋不休强于战争”。大约 50 年后，Winston Churchill，如果他能从天国看到当今的世界，将对现今的喋喋不休倍感惊讶。今天，电话与个人电脑已成为我们大多数人每天最主要的办公及生活用具。电话是基于模拟技术的，因为它必须能够支持人类发音所产生的模拟波形。比较而言，个人电脑是基于数字技术的，因为数据是由 0 和 1 这样的二进制字符串进行编码和操作的。

直到最近，将话音与数据传输集成在一起的常规办法主要分为两种：一是利用现有的话音网络满足数据传输的要求；一是构建专为在普通网络上传输语音、数据、视频信号和图像等多种信息而设计的异步传输网络（即 ATM 网）。将数据传输集成到原本用于传输远距离话音呼叫的网络中也只是最近才出现的。在 80 年代进入市场的第一代 T1 多路复用器就提供了这样的功能；通过声音数字化技术的改进以及一系列新型多路复用器的开发使企业通信网络可以利用 T1 和 T3 传输设备在传送路由器至路由器通信的同时还可以传输大量的语音对话外加传真和视频会议。尽管基于话音网络的数据传输仍将继续作为在普通网络基础结构上进行信息传输的重要机制，但它主要适用于那些在不同地点之间话音通信占较大比重的机构。

尽管 ATM 被设想成作为一种统一的技术被开发，以便支持在普通网络基础结构上的语音、数据、视频和图像应用，但和许多其它技术一样其具体实现结果并非所宣传的那样。由于 ATM 技术具有可扩展性，它可使基于局域网的以 25 或 155Mbps 操作的数据以单元流的统一形式和 1.544Mbps 的 T1 速率传送到通信公司的中心站；再通过光缆以高达 2.4Gbps 的速率在各公司的通信中心之间进行传输。这是一个绝妙的主意，但是由于局域网设备所具有的价格低廉的竞争优势，使许多机构推迟了将 ATM 发展到桌面的计划。而 ATM 在通信公司中被广泛认为是对其它传输形式的一种分类机制，范围从 T1 线路上的传统的数字化声音到帧中继以及理论上的构建于 ATM 之上的 TCP/IP 数据网络。因此寻找一种更实际有效和成本低廉的机制来传输声音和数据的网络管理者和局域网管理员转向了第三种选择，这就是本书的焦点：基于数据网络的话音传输。

当我们将注意力转向 IP 话音传输和帧中继话音传输时，也将说明在许多情况下可以利用许多因特网服务供应商（即 ISP）和通信公司使用的 ATM 基础结构的优势。因为 ATM 可以提供真正的服务质量（QoS），它可以保证带宽和延迟，如果可以把 IP 和帧中继映射到 ATM 我们就可以利用 ATM 的优势。因此，在通过数据网传输话音为 ATM 提供选择的同时，也可以对已有的 ATM 基础结构加以利用。

1.1 概 述

基于数据网络的语音传输技术的开发满足了一项最基本的网络管理需求：即以最低的成本获得设备，其安装和操作要尽量不影响正在进行的机构活动。许多读者会认为，对提供基于数据网络的语音传输功能的设备的开发代表了技术上的一大进步。尽管这是通信技术的一个进步，但大多数底层技术在几年之前就已存在了。基本上说，基于数据网络的语音传输可以看作是一个涉及多种硬件和软件产品的有关声音数字化和声音压缩的应用，它使得声音可以在原来只进行数据通信的网络上进行传输。设备开发者主要针对两种网络形式，IP 网络和帧中继网络。本书也将详细介绍这两种网络在基于数据网络的语音传输中的使用情况。本书后面将谈到，有关获得数据网络上的语音传输能力所使用的硬件和软件的基本概念将同样适用于公共网络和专用网络。此书所提供的信息将适用于公共 Internet，专用 Intranet 和公用、专用帧中继网络上的语音传输。现在我们已大致了解了如何以集成的数据和话音网络实现基于数据网的语音传输，并针对底层技术及其优势进行了简单的讨论。让我们开阔视野将注意力转向为什么在新世纪到来之际基于数据网络的语音传输会成为最热门的技术。由于这项技术与其它技术相似，他们在不同的网络方案中的使用都有一定的局限性，我们将涉及到影响话音在数据网络上传输的潜在的联网问题，就已有的和正在研究的有助于克服这些困难的解决方法进行讨论，然后在本章结束时我们将简要介绍一下随后各章的内容。将帮助你决定是直接转到你所关注的信息所在的章节，或者如果你在基于数据网络的语音传输方面是一个新手，就应该按顺序阅读各章节以便获得整体的认识。

1.2 基 本 原 理

现在许多机构的话音和数据网络是分开运行的，由于传统上这两种技术是分离的，设备最初也是为话音和数据网络分别设计的。另外一些机构则通过增加设备使数据可以在原先只支持语音传输的网络上进行传输，即形成基于话音网络的数据传输的基础结构，从而将全部或部分话音和数据网络集成起来。一些机构未能将话音与数据网络集成起来的原因有财政和技术两个方面的因素。在某些情况下，维持原有的分离网络是合理的，这是由于将二者集成起来的花费可能超过潜在的节省预算，或系统的实现需要相当一段时间，结果使得投资回报率很低。至于技术方面，直到最近，设备还不能非常有效地在 IP 和帧中继网络上传输话音，这样就需要考虑其它技术，如时分多路，在不考虑话音和数据各自的带宽需求时，它将对话音和数据传输能力具有永久性影响。

1.2.1 传统的基于话音的数据传输的局限性

图 1-1 说明了通过时分多路技术集成话音和数据的传统方法。分析图 1-1，注意到每个多路复用器的帧都可提供一个话音和数据源之间的带宽静态分配。例如，如果 PBX 设置成可提供 20 路 PCM 解码话音对话，每路为 64Kbps，则每帧将有 1.28Mbps 带宽分配给 PBX ($64\text{Kbps} \times 20$)，其余 256Kbps 分给数据源。由于每帧的时间片划分是固定的，如果

数据源停止传输，PBX 就不能利用它的时间片传输另外的 PCM 数字化语音对话。同样，如果 20 路语音对话中的一路或多路停止传输，多路复用器也不能把更多的带宽分配给数据源。

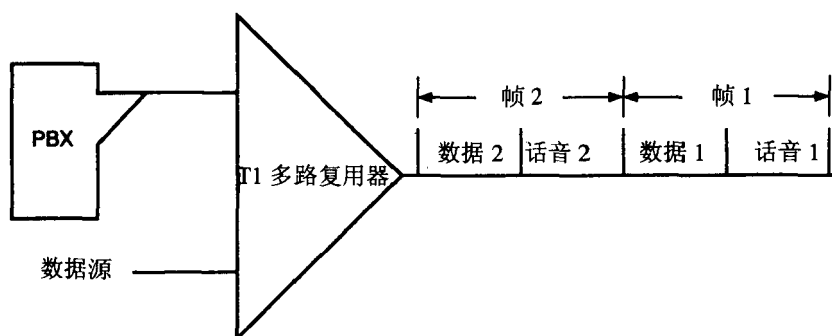


图 1-1 集成语音和数据的传统方法是采用时分多路技术，它为语音和数据分配固定的带宽

另一个与传统的 TDM 技术相关的局限是需要用户通过安装一系列点对点数字租用线路来建设自己的网络。尽管这种类型的网络可以为该机构所专用，但这种专用不是没有代价的。比较而言，使用事实上可以支持无限多个机构的传输需求的报文分组网络的花销更为便宜，因为大家共同分担了构建报文分组网络的传输设施的花销。

专用的专线网络的可靠性是另一个必须考虑的局限性。最普遍的增强可靠性的方式是，为不同路由安装附加线路以保证源和目的地之间的每条线路经过不同的中心站，所以这项附加费用也是相当可观的。一个机构将负担每个网络节点间两条线路的花销以及实现各节点间一对线路的不同路由所需的每月的路由费用。

1.2.2 基于数据网络的语音传输的优点

结合使用报文分组网络进行数字化语音传输的优点是相当多的。以下列举了与基于分组报文共享网络的语音传输相关的 6 个优点。

有效的带宽分配

- 可以使用现代化的语音压缩方法

降低开销

- 可以使用单一接口

报文分组网络的增强的可靠性

- 与共享网络使用相关的经济性

1.2.2.1 带宽分配

IP 和帧中继都是分组报文共享式网络，其带宽只在发生传输时才消耗。因此，这些网络取消了第一代 T1 多路复用器的固定带宽分配。但是这些网络最初是为在低速传输设施，如 56K 和 64K 数字线路，进行数据通信而开发的，对这些网络的访问速率在最近才增加到 T1，对 IP 网增加到 T3，而对帧中继网的 T3 连接在本书写作之时还在开发之中。

1.2.2.2 现代语音压缩技术

由于 PCM 数字化语音对话需要 64Kbps 的传输速率，如果试图将 56K 或 64Kbps 的数字线路上的数字化对话传输到 IP 或帧中继网络是没有意义的，因为这样做将排除同时传输数据的能力。认识到这个问题，设备开发商把新一代语音压缩算法结合到他们的产品之中。今天，你可以考虑使用一些帧中继接入设备（FRAD），它们包含数字化语音模块可将语音对话的操作速率降低到 4 或 8Kbps。这意味着你可在一个 64Kbps 的帧中继连接中传送 8 至 16 路语音对话，或在单一的帧中继连接中混合传送语音和数据。

1.2.2.3 降低开销

语音压缩技术的采用以及使压缩的语音可以在报文分组网络上进行有效流动技术的发展是两个关键，它们使这一技术获得认可，成为在数据网络上进行语音传输的具有实际意义和经济意义的主要机制。与之相似将语音压缩算法和 IP 设备相结合使得语音对话可以使用非常低的带宽在 IP 网上进行传输。

由于使用帧中继网络的大部分开销在于每月访问线路和帧中继网络提供商的接入端口的费用，任何可以使得其它数据源在这个公用连接上进行传输的方法都是很值得的。同样，如果你所在的机构拥有一个因特网连接，或者你将内部 IP 网络构建成为一个 Intranet，则对现有传输设施的能力的充分利用实际意味着降低了传输成本。即，你已经在为诸如路由器端口，通道服务单元和传送设施等设备付费。所以除去语音数字化的开销以外，传送数字化语音不必增加额外开支，从而使每分钟通话费可以达到很低。

通过帧中继和基于 IP 的网络进行语音传输可以提供几个关键性优势。这些优势包括：通过到分组网络的单一连接，可以在多个地点之间收发信息、增强传输的可靠性，并能更经济合理地让多个机构共享分组交换网络。

1.2.2.4 通过单一网络连接进行多种访问

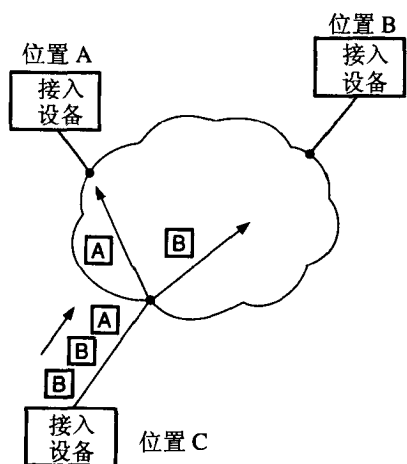


图 1-2 通过单一报文分组网络访问多个连接

一个报文分组网络允许在单一物理线路上传送多个逻辑连接，这样允许一个机构通过单一的网络连接支持对多个地点的传送。图 1-2 说明了这一概念，显示了通过使用报文分组网络接入设备以及从该设备到报文分组网的单一网络连接，将从地点 C 发给地点 A 和 B 的报文进行混合的方式。比较而言，使用基于专线的网络将需要一个机构从地点 C 到地点 A 和地点 B 分别建立专线，这样还会需要用于专线网络的网络设备，如：路由器或多路复用器，具有两个端口，因为与每条专线的连接都需要一个端口来支持。

随后需要加上建立专线网络的开销，因为路由器和多路复用器的每个端口要花费 500-1000 美元。另外，当你的网络需求扩大使你需要考虑大量节点之间的互联时，每个点对点网络连接的成本，以及每一个地点

路由器或多路复用器的大量端口的开销将使计划成为泡影。在这种环境下，你就需要考虑另外一种网络结构——将不同地理区域内的各点先连接到一个集线器，然后再将几个集线器互联起来。尽管这种类型的网络结构比使用专线和路由器、多路复用器等花费要少，但它更适用于支持需要大量话音传输以及相当数量的数据源之间的互联，而对于支持一定容量的基于数据网络的话音传输并不很适用。对后一种情况采用帧中继或 IP 网络传输语音和数据通常更具有经济优势。这带来了大量的联网可能性，如将国内和国际的站点通过报文分组网络连接起来，或使用 FRAD 支持在两点间原来只有少量数据会话的专线上增加话音传输功能，这是由于在某些供货商的产品上增加语音支持只需要很少的一次性费用即可实现。

1.2.2.5 增强的可靠性

报文分组网络的主干结构通常是基于网状拓扑结构而构建的。这种拓扑结构，如图 1-3 所示，通常在网络节点之间提供 2 条或多条路径，这取决于两个网络接入点之间的网络节点的个数，网状结构可在两个网络接入点之间提供众多的可选路径。例如，如果在地点 A 和 C 之间的直接线路连接不起作用，则这两点之间的传输可通过路径 A-B-C 或 A-D-C 继续进行。报文分组网络的这种网状结构主干提供了一种内置的备用路由能力，其成本由网络上大量的个人和机构用户共同分担。相比较而言，对于单独的机构要开发一个网状基础结构以提高网络的可靠性将会需要非常大的投入。

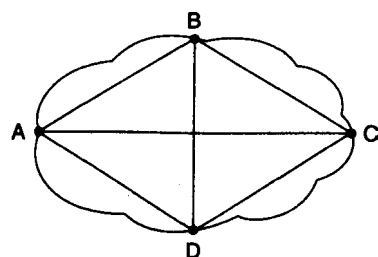


图 1-3 大多数报文分组网络使用网状结构连接各接入点，提供增强的冗余度以改善网络的可靠性

1.2.2.6 使用的经济性

在我们前面讨论报文分组网络的优点时，我们谈到它们可以通过单一网络连接进行多点访问，以及通过重复构建专线网络而形成网状基础结构所带来的可靠性的增强将是非常昂贵的。这两个优点为报文分组网络的使用提供了经济上的局限。但是通过使用不同的帧中继网络和因特网将产生一个更具实质性的经济优势。例如：一些帧中继网络的使用费是根据平均月费用即承诺信息速率（CIR）收取的，CIR 是帧中继网络提供商所保证提供的速率。CIR 可能小于或等于访问线路的实际操作速率，这将在本书后面介绍。这意味着如果你尚未使用全部的 CIR，则可以不必增加传输开支而以常规的数据或数字化语音的形式传输额外的信息，需要的花费只是一次性获取用于在帧中继网络上传送语音所需的设备。对于使用 IP 网，大部分 ISP 对用户机构每月收取的费用是根据连接到该机构的专线的传输速率而确定的。这样在因特网上以常规的数据或数字化语音的形式传输额外的信息将不会影响该机构的通信费用（除去购置用于支持数字化语音传送设备的一次性开支）。如果 IP 网络是一个机构的内部专用网络，在网络上增加话音传输功能将带来很大的节约。这是由于使用现代化语音压缩方法可以使话音包通过许多网络的多余带宽进行传送。再者，这种数字化语音包并不改变该机构传输设施的每月费用。而通常只是需要一次性购置支持对语音进行数字化和打包处理，使之可以在 IP 网络上传

输的设备费用。既然我们已经了解利用数据进行语音传输的优点，让我们把注意力转向一些可能发生的网络问题。理解通过数据进行语音传输的某些潜在缺陷，可以帮助你实施这种网络技术时做出明智的决定，其基本依据是：你所在机构当前和未来的通信需求，现有网络的基础结构，以及对不同形式的报文分组网络的潜在使用情况。

1.3 对潜在问题的考虑

基于数据网络的语音传输包含一定的风险，这与其使用的技术有关。网络管理者和局域网管理员应考虑的问题可以分成 4 个方面：可靠性、可预见性、安全性和标准化。

1.3.1 可靠性

出于本书的目的，我们将可靠性定义为一个包到达目的地的能力。在原来传输数据的报文分组网络上传输语音承担了一定的风险，即能否将语音编码包进行可靠的传送。一些报文分组网络包含一个丢弃机制，即当网络的使用达到一个预设的水平时，所传输的包将被象征性地送到“空中的巨大的存储记录中”（即被丢掉）。尽管采取了很多措施，但性能没有很大改善，这些措施包括在数据包中传送电子邮件消息，附带交互式的问答数据，或重传掉包的（dropped）文件（或包序列）。当传送数字化语音时情况不是这样。语句必须被重组以使其听起来自然，这需要去除由于掉包重发而带来的延迟。因而在数据网络上进行有效语音传输的能力需要一种机制以保证包传输的可靠性。当我们在本书后面介绍和讨论某些类型的设备时，我们将会注意到生产商主要是通过采用限制包的长度，减少它们被报文分组网络丢弃的概率的这种方法来增加语音编码包的传递可靠性的。

另一个与数据网络传送语音有关的可靠性问题是当包含声音的包被丢掉后，接收方所采取的方式。大多数设备只是不做任何事情，从而产生一段静默，如果掉包只是偶尔发生，人耳就不会对此有所察觉。然而，如果传送数字化语音的一些包被丢掉，造成语音的静默间隙，人耳就会察觉。要克服这个问题，一些设备生产商就生成一些噪音，它不像静默间隙那样引人注目。

1.3.2 可预见性

许多人认为可预见性与可靠性是相似的，实际上二者代表了基于数据网络的语音传输的两类不同问题。可靠性指的是包含数字化语音的包到达预定目的地而未被网络丢掉的能力。可预见性指的是在包的传递过程中没有很大的延迟，如果延迟很大就需要重新建立听起来不太舒服的传输对话，从而影响系统的可预见性。可靠性与可预见性是否会成为问题取决于所传输的语音类型。为了理解这一点，我们介绍一下您的机构在数据网络上可能希望传输的两种语音类型。

1.3.2.1 对语音应用的影响

需要在数据网络上进行语音传输的应用分为两大类：实时的和非实时的。一个实时语音应用，例如电话呼叫，对可靠性和可预见性都有要求。比较而言，附在电子邮件中的语

音文件，或将预先录制的声音文件通过邮件从一地发往另一地，对可靠性和可预见性就没有要求。

针对预先录制好的声音信息，就不要求传输数字化语音的过程中尽量减少掉包的可靠性。因为最终的掉包重传可以确保将包传到目的地。在所有声音片段到达并存成一个整体后对方就可以听到了。尽管包到达时带有随机出现的延迟，由于不是实时收听，所以可预见性也不成为问题。因为包在被收听前要去掉在报文分组网络上传输所形成的间隙，组合成一个整体。

1.3.2.2 改善可预见性的技术

有两项基本技术可以增加包到达预定目的地的可预见性。第一项技术中引入服务质量（Quality of Service, QoS）功能，为语音包序列的传输保留网络资源以保证最小延迟。尽管 QoS 功能是 ATM 所固有的，现在已发展成可通过资源保留协议（Resource ReSerVation Protocol, RSVP）将服务质量（QoS）间接用于 IP 网络的机制。但是，RSVP 要求在包经过的所有 IP 网络路径上都安装符合 RSVP 要求的设备以便能够获得从源到目的地的网络资源保留。在一个像因特网这样的公共 IP 网上，几年前相当数量的厂家就已进行了设备的 RSVP 升级，以便为哪怕是很少量的 RSVP 需求提供保证。当然如果你的机构拥有内部 IP 网络，你就有更大的控制权将设备进行升级以支持 RSVP。

在意识到 RSVP 真正成为实用的工具还需要几年时间，而且在帧中继方面还没有相应的方法的情况下，设备开发者通过对语音包的控制来满足对可预见性的要求。他们将语音包分成很小的包，以利于在网络中传输，并制定优先级策略使语音包的优先级高于数据包，以使语音包先于数据包到达网络。另外，他们还开始研究其它技术，这些将在本书的后面详细介绍。尽管这些技术不能保证语音包能够按时到达目的地，但它们大大地提高了这种可能性。

1.3.3 安全性

除非你为 CIA, DIA 或其它神秘机构工作，当你使用公共交换电话网（PSTN）或公司内部话音网时，毫无疑问你的电话被窃听的可能性是很高的。当你使用基于数据网络的话音传输方式进行通话时，尽管被窃听的可能性不大，但当话音包或帧通过公共报文分组网络与公司内部话音网连接时就会带来新的安全性危险——这种危险给你所在机构带来的潜在威胁远大于泄露计划中的行程，或旅行安排甚至针对建设项目的竞标信息等。这是因为从机构内部专用网到公用报文分组网的连接，已将内部专用网对所有连接到公用网的用户开放。

尽管一个机构在使用公用报文分组网时总会有意或无意地泄露一些数字化语音包的内容，但更严重的问题是当网络连接到公用报文分组网时，黑客和解密高手们就会试图进入你的内部网。一般窃贼和骗子需要花费数小时或数天才能干的事，黑客们可以通过个人计算机在几分钟内完成。从安全角度看，有三个问题需要考虑：未被邀请的人禁止访问你的网络，对访问网络的人需进行身份验证，对包进行加密编码使它无法直接阅读。这些问题涉及访问控制，身份验证和对经由公共报文分组网络的传输进行加密，见图 1-4。

1.3.3.1 访问控制

访问控制是根据预先定义的准则对某些传输进行允许或禁止的一种机制，例如禁止传输具有某些源地址的包，或对具有另一些源地址的包则允许传输。这种机制控制公共网络和专用网络之间，两个公共网络之间，或两个专用网络之间的传输。实现访问控制的最常用的方法是在路由器中建立一个访问列表，前提是路由器支持这一特性。

有关访问控制列表的最难理解的问题是：如何在基于 IP 环境的话音安全性与由于处理扩展的列表所带来的延迟之间进行权衡，对列表的处理反过来还会影响对目的地端经重组后的话音的可理解度。由于语音使用的 TCP 和 UDP 端口不是标准化的，这意味着如果你的机构支持多个厂商的产品并为了安全而使用路由器访问控制列表的话，你的列表还要包括一些与 IP 语音相关的语句。由于访问列表是自上而下检查的，所以如果将与语音有关的语句放在列表的最后也会造成几毫秒的额外延迟。

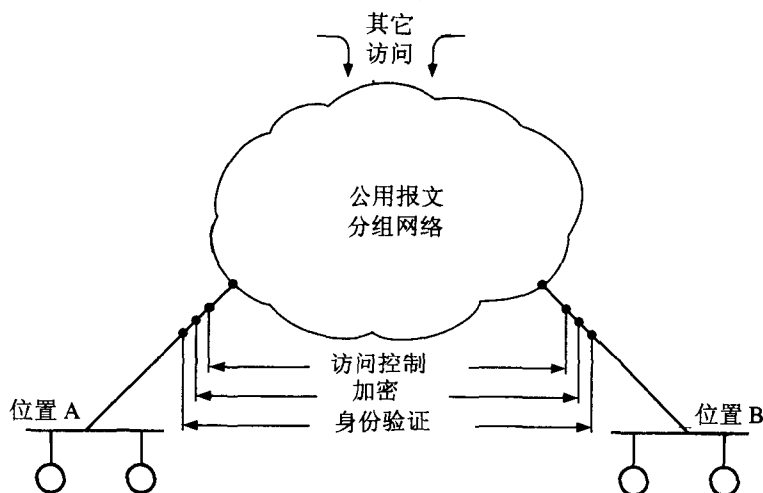


图 1-4 需要考虑的网络安全因素

应用注释 如果使用访问控制列表，将与语音相关的语句移到列表的开始处。

为了争取几毫秒的时间，你应该考虑把与语音相关的语句移到靠近访问控制列表的开始处。由于大多数列表开始处的语句都是禁止来自你的网络的入站包以防地址欺骗，所以较好的方法是将与语音有关的语句紧跟在反地址欺骗语句之后。

分析图 1-4 中网络安全的三个因素，考虑到两个机构的网络通过公用报文分组网进行联系，访问控制位于最里层。因为最常用的访问控制设备、路由器，可作为公用网络和专用网络的分界。第二种广泛使用的访问控制设备是防火墙。除了可通过访问列表提供访问控制功能外，防火墙还带有其它的安全特性，包括身份验证、加密、警告信息的生成以及代理服务等。路由器访问列表和防火墙的使用将在后面介绍。

1.3.3.2 加密

由于加密操作通常是由路由器后面的一个设备完成的，所以在图 1-4 中它被列为中间一层安全因素。但是读者应该知道，在许多厂商的防火墙产品中除提供身份验证外，加密也作为一个可选特性。所以加密和身份验证功能完全可能共存于一个设备之中。同样，也

可以由一个防火墙来提供身份验证，而由一个专用网络设备，如在专用网络中一台计算机上运行的一个程序，执行加密任务。在这种情况下，身份验证应该在加密或解密之前进行，从实用的观点看这样做在接收包时更有意义。因为如果在对一个包解密之后，其发送者无法通过身份验证，该包将被拒绝，这样就会影响设备的工作效率。

现在有许多设备可用来对包进行加密。但是与点对点线路不同，在点对点线路中可以对整个包进行加密和解密而不影响其路由，而对在公用网络上传输的数据流进行加密，需要使用智能化的加密设备。该设备必须能够把信息字段与包的报头和报尾区分开来，并只对信息字段进行加密操作。否则包就无法在网络中传送。

使用加密来隐藏包的内容又会带来延迟，它本身可能并无意义，但是当与其它因素结合起来，在数据网络传输语音时会反过来影响对重组后的语音的可理解性。因为一个黑客必须不仅能截取你的语音包，还要知道语音压缩算法，所以要听到未被授权的对话，不是一件简单的事情。因此，在许多情况下，你需要在加密之前认真权衡其利弊。

应用注释 加密会增加传送数字化语音包的延迟，所以应在绝对必要时采用。

1.3.3.3 身份验证

网络安全的第三个主要因素是验证包的始发者“确是其人”，对这种处理的正式称谓是身份验证。尽管一个私人对话中的声音足以验证一个呼叫的始发者，当你将公司内部网络与公用报文分组网络连接起来后，就可能需要对一些数据访问请求进行身份验证。你首先需要拥有这样的设备，它可以区别以不同的报文类型传输的信息，而且只对试图访问专用网络上某些预定服务的报文包进行身份验证。现在你对通过数据网络传输语音的合理性和其中潜在的问题有了了解，让我们将注意力转向在你所在机构的基本网络基础上增加基于数据网络的语音传送功能时，应考虑的一些关键性原因（除了与技术相关的经济问题）。我们将看几个应用实例，单从应用自身看，在你的网络结构中增加这些功能是很合理的。但在开始之前，先简单讨论一下有关的管理和标准。

1.3.4 管理的考虑

当通过数据网络传送语音可带来许多效率和经济上的优势时，它就成为一种非常值得考虑的技术，读者需要知道一些可能影响其实现的管理问题，国内和国际多方面正在努力避免单一机构在其内部网络上传送语音，并阻止语音服务经销商通过使用网关提供对话服务，从而与通信公司竞争。

1.3.4.1 基本原理

经济问题是国内和国际的管理机构试图阻止或限制各机构在数据网络上传送语音的主要原因。在过去的一个世纪里，全球的通信公司投资数千亿美元开发出使用 64Kbps 时间片传输数字化语音的技术。能够在 2.4Kbps 这样低的速率上传输语音的能力不仅非常具有竞争性，而且使得新组建的通信公司最终有能力将历史悠久的老牌通信公司挤出市场。这就像一个配件制造商能够将另一家公司的成本降低 8 到 16 个点或更多。

1.3.4.2 国内事件

在美国国内，电信公司协会（Association of Telecommunication Carrier），一个代表成千上万小型电话公司的组织，请求联邦通信委员会（Federal Communications Commission, 简称 FCC）阻止在数据网络上建立话音传输，因为如果用网关将各地区的交换机连接起来将减少他们的长途通信收入。

有关规则的另一个令人感兴趣的方面是：各地区的电话公司希望向打给因特网用户的长途电话收取传送费。如果一个朋友或业务伙伴通过 MCIWorldcom, AT&T, Sprint 或另一家长途通信公司拨打一个属于另一个地区电话公司的电话号码，提供当地电话服务的公司将因为协助完成通话而获得适当的补偿。现在，在因特网电话这个精彩的世界中，无法保证地区性电话公司可以获得任何补偿，这令它们感到非常气愤。FCC 提出了新规定，规定因特网通话就代表长途通话，但这一规定对当地电话服务提供者在传送这类通话后能否获得补偿所起的作用还有待进一步观察。

为理解收取传送费的潜在影响，假设当地信道的提供者对每次通话的收费为 5 美分。如果网络经销商当前对因特网通话的收费是每分钟 4.9 美分，它就会因为它们没有利润而停业。但对于使用专线的专用网络而言，前述的本地信道传送费用就不适用了。因此一个公司如果使用专线构成专用网络或使用专线对公用报文分组网络进行访问，它就不应缴纳本地传送费用。

1.3.4.3 国际事件

在国际上，有几个国家已经阻断了对某些网站的访问，原因是它们提供在因特网上拨打长途电话的软件。1997 年 6 月，捷克共和国、匈牙利、冰岛、葡萄牙和 Palau（一个南太平洋的小岛）制定法规，如果用户访问那些提供在因特网上拨打电话的软件的网站将被视为非法。1999 年 9 月，黎巴嫩在官方禁止因特网服务提供商向其用户提供国际长途电话之外，又增加了禁止因特网电话的规定。由于软件使得通话只需每分钟 10 到 30 美分，而那些国家收取的长途通话费为每分钟 2 到 4 美元，前面谈到的做法可以被看作是对提供国际长途通话的某些国有通信公司的一种自我保护。一个受到这种做法影响的厂家要求 FCC 和白宫出面干预。尽管国内和国际针对基于数据网络的语音应用采取了一些做法，但其潜在的影响实质上只是一些猜测。以作者之见，限制性的规定最终将弄巧成拙。尽管基于数据网络的语音技术还处于非常稚嫩的阶段，但它代表了一种相对易于实现的技术，是很难用行政方式来禁止的。国内和国际的通信公司应该考虑将这种技术集成到自己通信基础结构中去，而不是与技术进行抗争。他们应该能够设计出新的收费方案，为低速率数字化话音传输提供相对常规价格更具实质性的折扣，从而鼓励人们使用他们的设施。

1.3.5 标准

标准可以看作是粘合剂，它使不同厂家的产品具有互操作性。在标准领域内，IP 语音仍然是非常不成熟的，许多标准具有明细的缺陷，这就迫使厂家使用其它不太适于操作的标准。由于帧中继论坛在制定“实现规范 II（Implementation Agreement II，简称 IA II，它涵盖了帧中继上话音传输的很多特征）”方面的努力，使得帧中继话音现在较为标准化，

但仍有其它一些领域缺乏标准。例如，帧中继提供商们还没有将他们的服务水平协议（SLA）标准化，如优先级的划定，用于支持优先级的排队方法。

现在，普及的 IP 标准包括实时协议（Real-Time Protocol，简称 RTP），它涉及时间标记和包的顺序；RSVP，涉及对带宽的预定以及 H.323，它是一个保护性标准，原来是为电话会议而开发的，但由于控制带宽的使用，现在许多厂商都把它包含在自己的产品中。本书中我们将详细介绍每个标准。

除前面谈到的标准之外，还有两个进一步的标准，尽管它们对因特网电话的成功至关重要，但对于一个机构实现 IP 话音功能又不是必要的。这两个标准是开放结算协议（Open Settlement Protocol，简称 OSP），它被设计用于处理身份验证、呼叫路由和 ISP 之间的呼叫详细帐单；和 iNOW!，由几个厂商提出以使不同厂家的 IP 电话平台具有互操作性。

1.4 应 用

当讨论潜在的以及已有的数据网络话音应用时，我们进入了一个新的空间。类似于三维图形的出现给工业带来的革命。以下列出了五个基于数据网络的话音应用。我们将简要地讨论一下以便说明这项新技术的潜力。

- 文档会议
 - 访问求助热线
 - 集成的呼叫管理
 - 定单安排
 - 统一消息

1.4.1 文档会议

文档会议使得多个团体可以在各自的计算机屏幕上查看同一个文件，同时还可以进行声音的交流。直到最近，文档会议还是基于使用公用交换电话网进行话音传输，多台计算机上的软件同时使用公司网络或因特网提供实时文档编辑功能。由于具有在数据网络中传输话音的能力，所以使将数据和话音的通信需求结合到普通网络基本结构中变为可能。这不仅使软件的协调更容易，还避免了使用 PSTN，这会带来大量的节省，特别是当文档会议的话音和数据在各国间传输时。

1.4.2 访问求助热线

想象一下，客户购买了你们公司的软件和硬件产品。假如在安装过程中，他们遇到了问题或有疑问，再进一步假设在安装软件和硬件产品的过程中，在客户的多媒体计算机他所喜爱的浏览器上产生一个屏幕显示：“请单击这里与我们的客户服务代表谈话”。单击图标，用户的浏览器就建立了与你们机构中的“产品求助热线”的连接。它支持数字化电话的应用，使客户可以通过语音描述问题和寻求帮助。

前述的求助热线操作代表了一种数据网络上的话音传输应用，从客户的满意程度和经济前景两方面来看都很成功。这种应用可以先通过免费电话接到你们机构的求助热线设施

上。另外由于许多公司不提供国际免费电话访问它们的用户求助热线，因此通过因特网提供话音支持可能成为一个市场策略，使得一个产品比起没有提供此类功能的同类其它厂家的产品更有价值。这样通过数据网络提供到一个机构的求助热线的话音访问，使得其市场和经济方面的优势都有所增强。

1.4.3 集成的呼叫管理

呼叫管理包括入站和出站呼叫的路径选择，使用最适当的技术使呼叫到达其预定目的地。当你增加了在数据网络上传输话音的功能后，你可以将这一能力集成到现有的话音应用和传输方法中以获得集成化的呼叫管理能力。例如，可对 PBX 编程，使其识别开始为 6 的拨号并将其从本地 PBX 经由内部专用 IP 网络，帧中继网络或因特网路由到另一个地点。

另一个可提供令人感兴趣的讨论话题的呼叫管理技术是因特网回呼（callback）。如果你在国外旅行，并阅读了“International Herald Tribune”或浏览了一些外国报纸，你会注意到一些宣称“廉价回呼”，“国际长途节省 50% 或更多”的广告颇为引人注目。尽管你可能想到某些违法的阴谋，但厂家收费比国际通话费率低 50% 或更多的是合法的。它们使用一种叫做回呼的功能，使用户可拨叫它们在美国的设施，在多数情况下并不应答他们的呼叫而只是读出他们的呼叫者 ID 电话号码，然后使用非常低的费率拨回来。在接受回呼时，用户输入一个访问代码和要接通的电话号码，在美国的设备就产生到目的地的第二个电话呼叫。由于同是从美国发出的呼叫，所以比一个来自其它国家的国际长途电话便宜很多，节省的部分就返还给用户以鼓励其继续使用回呼功能。

为了使国际通话更经济实惠，许多回呼操作者现在转向具有话音功能的 FRAD，将它们用于国际线路，在一条 64Kbps 的帧中继线路或 64Kbps 专线上可同时传送多达 16 路电话。事实上，ACT 网络公司的报告称他们的具有话音功能的 FRAD 大量卖给回呼操作者，原因在于他们的设备可以大大降低长途国际通话的开销。

1.4.4 定单安排

尽管尚处于不成熟的阶段，使用因特网订购从书籍到百吉饼等产品已成为可能。可以与预定台直接对话，为通过因特网销售产品增加了新的空间。以咖啡制造商的销售为例，让我们考虑一下现在的产品销售方式以及产品售出的原因。假定你找到一个网页，它显示一幅图画和有关“12 杯咖啡”制造商的介绍。让我们假设你有一个关于它的过滤器的问题。通常你会单击电子邮件的图标，它产生一个表，在“收件人”一栏包含有公司邮件地址。填写完信息后单击“发送”按钮，你就可能去浏览其它网站，因为你知道对你的问询还不能做到马上回答。现在假设你不是单击邮件的图标，而是一个语音图标，并且直接与客户服务代表取得联系，他马上回答了你的问题并说“我可以处理您的定单吗？”这显然是一个获得销售额的更有效的方法。

另一个可能的变化是使用话音对潜在的客户进行因特网定单安排，如果客户不具有相应的话音能力，就需要将一个公司呼叫管理能力集成到 Web 浏览器的活动中。例如代替单击语音图标，潜在客户可能需要提供选项，单击“帮助”按钮，就会让读者在 WWW 网页上输入用户名和电话号码。这个信息被传给生产厂商，结果，呼叫管理系统就会拨通

号码，并将一个客户代表连接到已拨通的连接上。针对本节讨论的两种方法，最终结果是将语音与数据结合起来使用可帮助潜在客户，并有望提高达成销售的可能性。

1.4.5 统一消息

统一消息指的是使用一套系统去管理所有用于传送消息的方法，例如语音邮件、电子邮件、传真和实时语音。将基于数据网络的语音添加到一个统一的消息系统，使雇员可以充分利用所有方式传送信息。这是由于可以学习和使用常用的图形用户界面（GUI），免去用户花费很多时间到不统一的分离系统中去查找和检索信息。另外，通过将数据网上的语音应用集成到统一的信息系统中，系统可以促进语音应用项目的使用，如雇员可使用实时语音或语音邮件，来代替在 PSTN 上的拨号，这将会大大提高工作效率，并避免当无法使用 PSTN 时，不得不留下实时语音邮件的开销。

1.5 技术上的成功

数据网络的语音传输同其它成功的新技术相似都是由于它们提供了一定的客户满意度。通过对电话公司的电路交换和包交换进行比较，我们可以对客户的满意程度进行评价。然后我们将关注一个我称之为“获得成功的金钥匙”的问题——即在构造基于数据网络的语音传输之前必须努力降低的最大延迟。

1.5.1 电路交换与包交换

表 1-3 提供了传统的电话公司使用的电路交换与包交换的比较。在分析了表 1-1 中各项之后，我们注意到电话公司使用的脉冲编码调制（PCM）代表了一种“长途电话品质”的语音数字化方法，它可形成高质量的重组语音。与之相比，大多数基于数据网络设备传输的语音都使用低速率编码器，它可提供近似长途电话品质的重组语音。但是由于报文分组网络会随机出现延迟，语音重组也会受不断变化的延迟的影响。

表 1-1 电路交换与包交换的比较

参数	电路交换	包交换
指定带宽	是	否
服务质量		
语音质量	长途电话品质	非长途电话品质
延迟等待	小	可变的
使用水平	低	高
使用经济性	低	高
呼叫管理功能	很多	很少

尽管在表 1-1 中没有指出，但是需要注意的是：一旦你提供交换电话线路形成了一个连接，就不会有意外的断线。相对而言在网络阻塞时，路由器和帧中继交换机都会掉包。考虑到表 1-1 的最后一项，呼叫管理特性可包括呼叫等待、呼叫转发、呼叫者 ID 和统一收费。尽管大多数电话公司在市场上推出了全系列呼叫管理功能，而大多数基于数据网络

的语音解决方案现在只能提供很少的这类功能。以我之见，包交换语音要想获得成功就需要能提供与传统的电话公司类似的功能和特点。

1.5.2 解决延迟的关键性问题

人类可以忍受大约 250 毫秒的延迟，超过此限度则通话就变得很糟糕以至于不可容忍。当话音经由报文分组网络传输时，必须考虑到固定的和可变的两种延迟。当了解了何时会出现何种延迟后，你才能考虑采取何种办法来保证将端到端的延迟时间限制在 250 毫秒之内；如果你无法提供这种功能，你可以考虑暂时放弃你的语音应用项目，直到你能够保证将延迟控制在合理的范围内时再执行。因此，理解阻塞和延迟就成为实现数据网络上的话音传输的金钥匙。

表 1-2 固定和可变的报文分组网络延迟

原因	延迟 (ms)
固定延迟	
压缩 (语音编码)	20—45
源端的交互处理	10
源端的网络访问	0.25—7
网络延迟	20—100
目的地端网络出口	0.25—7
目的地端的交互处理	10
抖动缓冲区 (可配置)	10
解压缩 (语音)	10
固定延迟	90—199

最坏情况下一个呼叫可容忍 250—80 至 280—199 或 170 至 51ms 的可变延迟。

表 1-2 列出了报文分组网中常见的固定和可变的延迟。分析表 1-2 中各项，交互处理延迟表示路由器的路由处理时间，并未考虑处理访问列表和加密所引起的额外延迟。还有网络延迟和网络出口延迟是与线路的工作速率相关的。例如，当一个相对较小的语音包的访问和出口发生在操作速率为 1.544Mbps 的 T1 线路时可能需要 2.5 毫秒，当访问和出口发生在 56Kbps 的线路时，延迟就要增加到 7 毫秒。

应用注释 理解了由接入和出口线路、不同的语音压缩方式和网络等因素引起的延迟，使得我们可以考虑使用不同的技术来减少整体的延迟。

既然对限制我们成功实行基于数据网络的话音传输技术的制约因素已进行了评估，就让我们谈一谈这一技术采用的具体方法。由于这些方法是本书余下部分的重点，我们就预先介绍一下以后各章的话题，作为本章的结束。

1.6 话题预览

你可以将本节提供的信息独立使用，或与本书索引相结合，将焦点集中到你所感兴趣的话题。你可以按顺序阅读各话题，也可以直接阅读与你的特殊需求相关的资料。

1.6.1 网际协议

在第 2 到第 4 章我们将了解 IP 和帧中继网络，以及它们的特点，还要了解对人类语音的编码技术。有些读者对这几章的内容可能已经很熟悉。如果你符合这种情况，可能就希望跳过或简单浏览一下这些内容。对其它读者，这些章节中介绍的信息将会给他提供一个有关这两种类型的报文分组网络的基础知识，正是采用本书介绍的许多新近开发的硬件和软件产品使基于这两种网络的语音传输既经济又实用。

在第 2 章，我们将讨论 TCP/IP，包括 IP 栈的构成，处于栈中的应用，以及为什么 IP 操作缺乏可预见性。我们还将讨论 RSVP 的作用，以及如何通过它与不同网络的结合最终提供一定的可预见性，来满足基于公共 IP 网络的话音传输的要求，同时还将介绍 RTP 和 H.323 标准的使用。

1.6.2 帧中继

第 3 章将对基本 X.25 的包交换技术进行综述，并介绍帧中继的基本原理，说明快速发展的帧中继产品和服务的市场所具有的优势和弱点。综述之后，我们将焦点集中到帧中继。我们将了解专用网络是如何连接到这个公用网络上的，定义各种帧中继参数，讨论信息通过这种网络的常规流程。包括帧的丢失以及协议栈中的较高层次是如何保证帧的丢失不会转变成实际数据的丢失。

1.6.3 语音基础

介绍基于数据网络的话音传输技术的书，都应包括有关声音的特性、数字化和编码的详细介绍，这也是本书第 4 章的重点。在这章中，我们将分析人类语音的特性，这是讨论对声音进行分析和合成的基础。在讨论不同的声音数字化方法时，还将介绍不同的国际标准，为介绍后面 3 章的内容打下基础。

1.6.4 电话操作

在数据网络上传输语音需要选择支持你的 PBX 的信号传输方式的设备。故此需要了解电话呼叫是如何在用户交换机 PBX 间传送的，是如何通过交换电话网络的，以及建立呼叫需要的不同类型的信号。这些内容见第 5 章。

1.6.5 IP 语音

正如前面所提到的，有两种类型的报文分组网络适合于传送语音。一种是基于 IP 协议的，包括从专用 IP 网络到所有公共 IP 网络之母的因特网。在第 6 章，我们将关注 IP 网上的语音传输，对实现这一功能的不同技术进行分析，包括对不同厂家设备的操作，与 IP 语音相关的经济效果，以及如何将这个技术集成到已有的或建议中的专用网络上，而这些专用网络是通过内部 IP 网络或像因特网这样的公用 IP 网络连接起来的。

1.6.6 帧中继语音

继续进行有关在数据网络上传输语音的实用分析，在本书结尾部分的第 7 章，将讨论基于帧中继网络的话音传输。首先回顾了各种不同的技术，它们既能实现这种功能，又使

网络具有一定的可预见性，然后将关注相关厂家设备的操作，与这种技术相关的经济效果，以及你应考虑的各种网络集成问题。本章将使你了解把话音移到帧中继网络会遇到的实际问题。包括现有设备的使用以及与已有的和计划中的专用网的集成。

1.6.7 管理

一本有关 IP 话音和帧中继话音的书如果不考虑管理工具和技术，将是不完整的。在本书的结尾，我们分析了几个管理问题，评估了不同工具的使用，以便帮助你在进行设备和 / 或线路设施投资之前决定这项技术能否使工作达到可接受的水平。

第二章 IP 及相关协议

任何有关基于 IP 网络的语音传输都需要有一定的网际协议的知识。由于 IP 代表 TCP/IP 族协议的一部分，要将 IP 作为独立实体进行讨论是很困难的。如要真正实现 IP 语音，需要有关设置 IP 地址和子网掩码以及将计算机指向正确的网关和域名服务器等知识。本章将向读者提供配置和操作任何包含 IP 语音的 IP 设备所需要的实际信息。另外有关 TCP/IP 协议栈操作的知识将有助于我们了解某些活动是如何通过 IP 网络反过来对数据流产生影响的，我们还将注意到某些网络操作对延迟的影响。这使我们可以考虑进行某些网络调整使原先无法实现的 IP 语音传输方式成为可能。

本章的重点放在 IP；但由于它代表 TCP/IP 协议族的一部分，我们将先复习一下这一族协议对应于国际标准化组织（International Standard Organization, ISO）的开放系统互联（Open System Interconnection, OSI）参考模型的分层结构。在此之后，我们介绍网络和数据链路层的编址方式。为此我们还要分析以太网和令牌环帧在局域网中的传输方式，以及 TCP/IP 使用 32 位 IP 地址的方法。这些信息将用于讨论在数据链路和网络地址之间的映象以及通过地址解析协议（ARP）实现的翻译机制。我们还将讨论域名服务器的作用和用法，以及网际协议的结构。另外，由于 IP 协议不能自己独立操作，我们还要回到协议栈，看一下传输层和传输控制协议（TCP）、用户数据报协议（UDP）以及其它可以提供 IP 语音传输的有效方式的相关协议。

2.1 TCP/IP 协议族

传输控制协议 / 网际协议（TCP/IP）是为实现预定的任务经过一段相当长的时间发展形成的一系列的协议。TCP/IP 的出现可以追溯到美国国防部高级计划研究 7 项目署（DARPA）开发的一系列通信协议，它们的目的是通过一个公用的网络基础结构，即高级计划研究项目署网（ARPANET），在不同地域的网络之间传输数据。对 ARPANET 的研究奠定了当今因特网发展的基础。首字母大写时，即 Internet，指的是互联在一起的许多网络。

尽管 ARPANET 和因特网的出现早于 ISO 的 OSI 参考模型，但 TCP/IP 协议族与 OSI 参考模型类似都具有分层的网络结构。因此，在讨论 TCP/IP 协议族之前，让我们先快速回顾一下 OSI 参考模型。

2.1.1 OSI 参考模型

OSI 参考模型是将网络功能划分成相关活动的逻辑组，这种逻辑组又称为层。每个网络层代表一系列独立的任务，从编程的角度看，具有一定的自包容性，但又能很容易地与

层号	层名
7	应用层
6	表现层
5	会话层
4	传输层
3	网络层
2	数据链路层
1	物理层

图 2-1 OSI 7 层参考模型