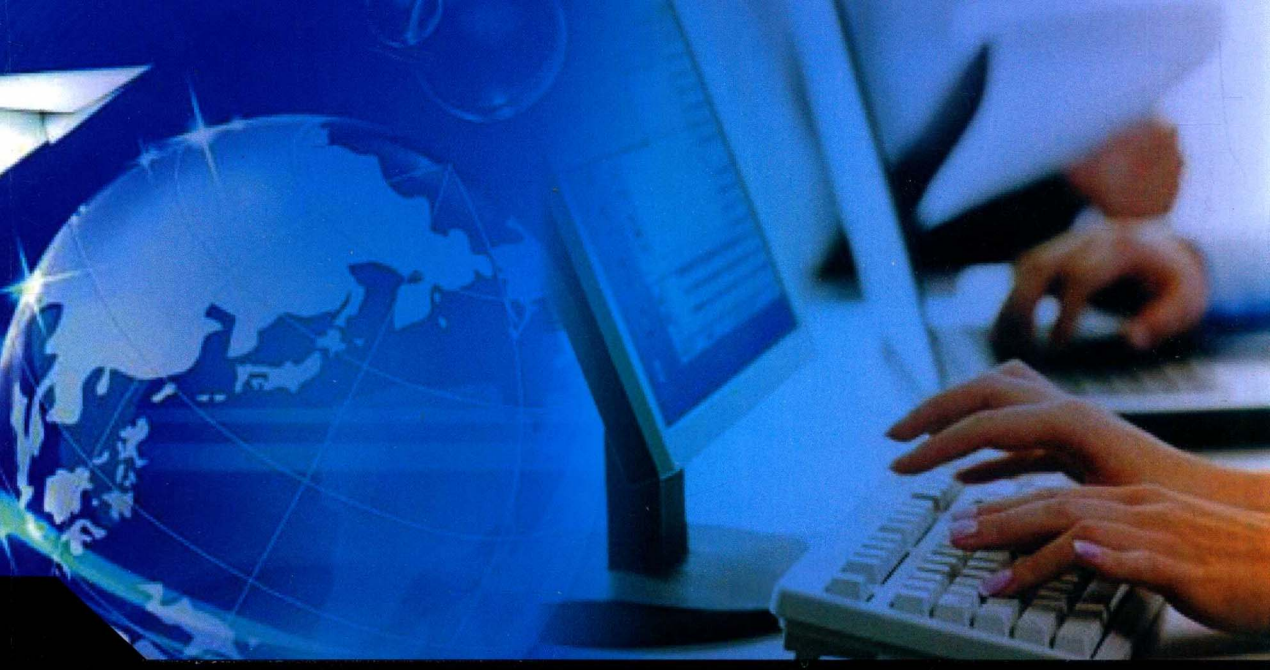


高等学校“十三五”规划教材

网络故障处理

李海龙　罗　眉　李文豪◎主编

WANGLUO GUZHANG CHULI

西北工业大学出版社

WANGLUO GUZHANG CHULI

网络故障处理

主编　李海龙　罗　眉　李文豪

西北工业大学出版社

西安

【内容简介】 本书将计算机网络的理论和工程实践相结合，有针对性地介绍网络故障处理所需的网络工程基础知识、网络故障处理的方法和步骤、网络硬件系统基础、交换机和路由器等骨干设备的软件系统配置、综合布线工程、网络集成与管理维护等内容。以故障的处理为最终目标，重点突出纯粹的“网络工程故障”，不涉及高层协议的软件应用。本书理论联系实际，让读者在学习网络故障处理原理的同时熟悉故障排除的工程实践，在掌握故障处理实践的同时理解故障产生的原理。

本书既可作为高等学校“网络故障处理”相关课程的教材，也可作为相关专业工程技术人员查阅的一本工具书。

图书在版编目（CIP）数据

网络故障处理/李海龙，罗眉，李文豪主编. —西安：西北工业大学出版社，2018.1

ISBN 978-7-5612-5680-0

Ⅰ.①网… Ⅱ.①李… ②罗… ③李… Ⅲ.①计算机网络—故障诊断 ②计算机网络—故障修复 Ⅳ.①TP393.07

中国版本图书馆 CIP 数据核字（2017）第 302879 号

策划编辑：蒋民昌
责任编辑：蒋民昌

出版发行：西北工业大学出版社
通信地址：西安市友谊西路 127 号　　邮编：710072
电　　话：(029)88493844　88491757
网　　址：www.nwpup.com
印 刷 者：兴平市博闻印务有限公司
开　　本：787 mm×1 092 mm　　1/16
印　　张：11.75
字　　数：284 千字
版　　次：2018 年 1 月第 1 版　　2018 年 1 月第 1 次印刷
定　　价：35.00 元

前　言

随着因特网新应用的不断开发，网络中的数据传输也越来越复杂。数据、语音、视频等多种应用对带宽的要求不断增加，一些新技术和老技术的兼容问题也不断出现，网络中的不安全因素越来越难以确定。这些都导致了网络环境的复杂化，意味着网络的连通性和性能发生故障的可能性变大，而且引发故障的原因也越发难以确定。同时，由于人们越来越多的依赖网络处理日常的工作和事务，一旦网络故障不能及时修复，其损失可能很大甚至是灾难性的。

能够正确地维护网络尽量不出现故障，并确保出现故障之后能够迅速、准确地定位并排除故障，对网络维护人员和网络管理人员来说是个挑战，这不但要求对网络协议和技术有着深入的理解，更重要的是要建立一个系统化的故障处理思想并合理应用于实际中，以便将一个复杂的问题隔离、分解或缩减排错范围，从而及时修复网络故障。

本书从网络工程理论和网络故障处理理论基础入手，介绍网络硬件及其故障、网络集成故障与故障测试、网络骨干设备的软件系统管理，最后从网络故障管理的角度介绍网络管理维护常用命令和网络故障分析处理的综合运用。

全书内容共分为 5 章：

第 1 章从计算机网络的软硬件基本概念开始，介绍一般意义上的计算机网络、网络的互连、局域网等基础知识。以“以太网”和“无线局域网”为例，重点讨论了在现实的局域网中，如何实现物理连接、媒体接入控制，以及广播、冲突的管理等计算机网络基础知识。在此基础上对网络故障处理的基本概念、网络故障类别、处理方法和故障的解决步骤等基础知识进行了介绍。

第 2 章介绍传输媒体、网卡、交换机、路由器等网络硬件及其故障的基础知识。

第 3 章介绍一些基本的规划设计和升级的概念、方法，并针对结构化综合布线和施工易出故障及故障测试等细节进行重点讨论。

第 4 章以 Cisco 的设备为例介绍网络骨干设备交换机和路由器的软件配置基础，培养读者依据配置文档进行故障分析与处理，排除故障的技能。

第 5 章介绍计算机操作系统和网络设备操作系统中常见的网络管理命令，并按照网络故障处理的解决步骤为主线，介绍这些命令如何在网络故障分析处理中进行综合的运用。

本书由火箭军工程大学李海龙负责全书策划和统稿，罗眉负责案例编写，李文豪负责校对和图表绘制。编写过程中，单位许多同事为本书提供了许多资料和宝贵的建议。在此表示衷心的感谢！

但愿本书能为读者学习网络应用知识提供有益的帮助。不妥之处，敬请指正。

编　者

2017 年 1 月

目　录

第 1 章　网络故障处理基础知识

本章从计算机网络的软硬件基本概念开始，介绍一般意义上的计算机网络、网络的互连、局域网等基础知识。并以“以太网”和“无线局域网”为例，重点讨论了在现实的局域网中，如何实现物理连接、媒体接入控制，以及广播、冲突的管理等计算机网络基础知识。在此基础上对网络故障处理的基本概念、网络故障类别、处理方法和故障的解决步骤等基础知识进行了介绍。

1.1　计算机网络及 TCP/IP 基本概念

因特网是一个庞大而复杂的计算机通信系统。普通的企业用户，并不需要关心在这个庞大的系统中，如何实现远距离的传输，以及其他网络的实现细节。若企业并不接入因特网或者其他互联网，用户只需要关心在自己的园区网络内部，如何通过传输媒体来有效地组织资源共享和数据传输。即便企业接入了因特网，也只需要关心如何屏蔽自己的园区网络与其他网络的不同，并通过怎样的方法去实现与其他网络的互相连接。

1.1.1　计算机网络概述

提到“计算机网络”这个词，恐怕浮现在人们脑海里的场景可能主要是色彩斑斓的网页、海量的信息和快捷的搜索、个性的微博、方便的电邮、P2P 的资源下载、诙谐的 BBS 回贴，或者温馨的校友录，也可能是视频聊天、IP 电话和视频点播等多媒体应用，也可能是“维客”“威客”“播客”“博客”这些时髦的网络新贵。但是，无论这些应用是多么的令人兴奋，多么的富有创造，它们仍然只是 Internet 的一些应用而已。而且，需要明确一下，Internet 也只是一种特定的计算机网络。

当然，对于普通用户而言，如果要在时下的计算机网络中找到一个区别于 Internet 的网络，的确也是一件困难的事。那些所谓的“专用网”，也使用着和 Internet 一样的组成和软硬件配置。所以描述计算机网络的概况，不妨从 Internet 的特征说起。

这个首字母大写的 Internet 比较通用的称呼叫做公共因特网，本书也统一采用这种叫法。首字母小写的 internet 也表示一种计算机网络，叫做互联网，许多专用网即属于互联网。“inter -”这个词缀的含义是相互之间，internet 就是网络和网络的互联（注：本书将 internet 翻译成专有名词“互联网”，是指网络功能的互相联合，但讨论网络（设备或硬件）互相连接的行为时，使用了“互连”一词。“互联”和“互连”两者所指不同。事实上，因特网就是最大的、遍布全球的互联网。本节通过一个粗线条的概述，阐述其硬件和软件构成。

1. 计算机网络的硬件

用过 Google Earth 的读者一定很熟悉，如果将地图放大到一定程度，就会看到建筑、公路和河流的细节。如果将地图缩小到一定程度，就会看到国家和和国家“拼接”的界线。同样的

道理，如果将镜头深入到计算机网络的细节，就会看到一般意义上计算机网络的硬件构成，包括连接设备、用户设备和传输媒体。但如果站在宏观的角度去观察包括因特网在内的任何互联网，就会看到如图 1.1 所示的互联网结构：通过路由器（本书第 2 章 2.4 节和第 4 章 4.3 节会进一步介绍）将各种不同的网络连接在一起。这些网络的规模、技术各有区别，但它们都平等地互连在一起，构成了一个更大的“网络的网络”。图 1.1 用不同的网络云来表示了这些网络的异构。用户使用计算机或其他智能设备，利用各种接入手段，接入到其中的一个网络。

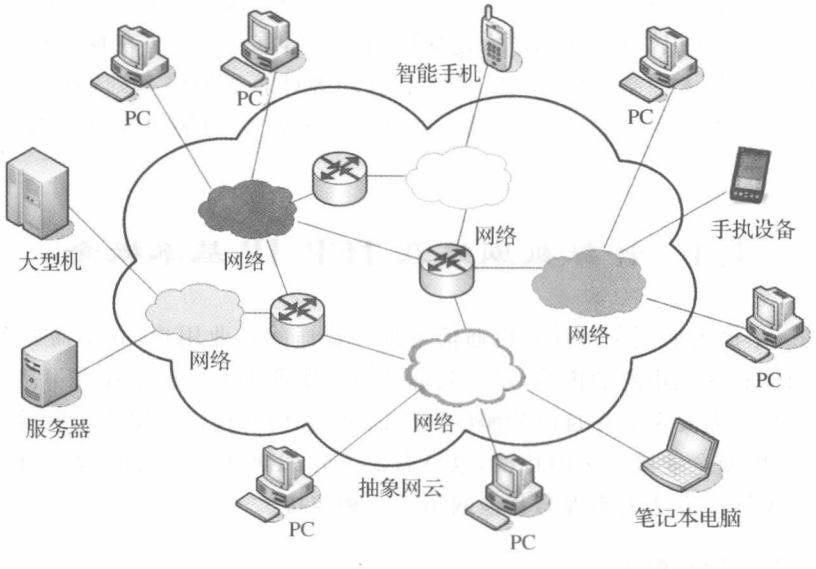


图 1.1　互联网的结构

在因特网中，与因特网相连的计算机通常被称为端系统（end system），它们在图 1.1 中位于边缘。因特网的端系统包括了计算机、便携机、PDA（personal digital assistance）、智能手机，甚至一些智能的家用设备。端系统也被称为主机（host），主机有时又被进一步划分为两类：客户机（Client）和服务器（Server）。这两个概念其实原本是两个软件的概念：客户机和服务器都是指通信中所涉及的两个应用进程。因特网利用“客户-服务器方式”来描述进程之间服务和被服务的关系。客户是服务请求方，服务器是服务提供方。但是由于客户程序经常运行于桌面 PC、便携机和 PDA 等主机上，常被称为客户机。服务器程序常运行于一些可持续工作、功能强大的主机上，用于发布 Web 页面、流媒体、转发电子邮件等，这些主机就经常被称为服务器。所以，客户机和服务器便约定俗成的成了硬件概念。

那些提供用户接入的网络被称为 ISP（Internet Service Providers）。不同的 ISP 提供了各种不同类型的网络接入，包括拨号调制器接入、以 XDSL 为典型的住宅宽带接入、高速局域网接入和无线接入。许多文献中常给这些将端系统接到其边缘路由器的物理链路起了一个名字叫做接入网（Access Network）。

依照计算机系统之间互连距离和网络分布地域范围，这些网络经常被划分为局域网 LAN（Local Area Network）、城域网 MAN（Metropolitan Area Network）和广域网 WAN（Wide Area Network）。但是随着技术的发展，这种划分的界限开始模糊。起初，当网络的作用距离不同时，由于信道的不同，网络采取了不同的技术来实现对数据的传输。局域网由于作用距离

较小，用户数量和传输出错率都比较小。所以一些涉及传输可靠性的技术并不需要和远距离传输一样复杂。而远距离传输则不同，通常需要借鉴传统的电信技术手段来实现。远程的两个局域网需要互连时，就通过与其他跨度比较大的远距离传输网络相连接来实现，如图 1.2 所示。在图中，用一个笼统的“网云”来表示这样一个远距离的传输。这个网云，可能是某种特定的广域网，也可能跨接了好几种广域网。

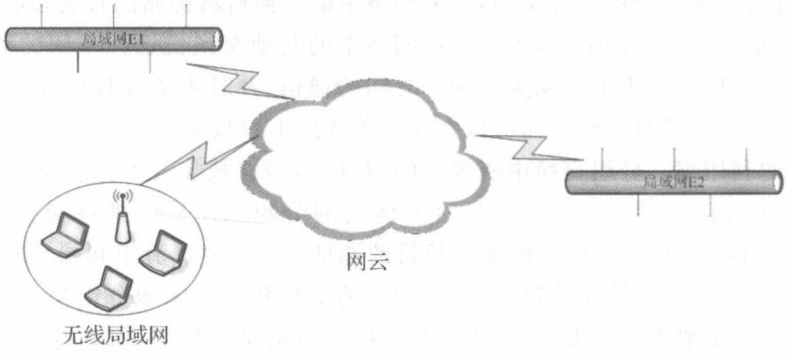


图 1.2　远程局域网的连接

接下来将注意力集中到图 1.1 和图 1.2 中的抽象网云。网云事实上就是由路由器和各种网络所组成的一个网状的网络，正是它互连了端系统和“端局域网”。通常，将它称为网络核心。其连接在拓扑上是四通八达的，不可能也不需要在每一对发送方和接收方之间都铺设专用的传输线路，这就需要在多个站点相同方向上传输到一定的距离后，根据目的地址进行分支（转接）选择，再通向不同的站点。还有信号在传输介质上的衰减和所受到干扰，也需有一个中继节点来完成整形放大。完成这些任务都需要有交换操作。根据所传输信号内容的不同要求，需要相应的交换技术，交换技术的发展与通信和计算机网络技术的应用紧密联系。按照交换技术发展的顺序讲述，目前使用的交换技术有电路交换（circuit switching）、报文交换（message switching）、分组交换（packet switching）、信元交换（cell switching）。构建计算机网络的网络核心主要使用分组交换和信元交换。

电路交换过程类似于打电话，当用户需发送数据时，主叫方通过呼叫，由交换网完成被叫就与它建立一条物理连接数据通路，在通话过程中一直独占该连接线路。通话结束，拆除连接时，由通信双方中任一方完成。在电路交换网络中，沿着端系统通信的路径，为端系统之间通信所提供的资源（缓存、链路传统速率）在通信会话期间将会被预留。它的特点是适合发送一次性大批量的信息。由于建立连接时间长，传递短报文时，效率较低。并且对通信双方在信息传输速率、编码格式、通信协议等方面完全兼容，这就限制了不同速率、不同编码格式、不同通信协议的双方用户进行通信。

分组交换是把电路交换和早期电报通信中所使用的报文交换的优点结合起来产生的一种交换技术。从概念上看，一个分组数据通信系统的硬件组成包括终端用户、分组交换网。其中，终端用户可以是计算机或一般 I/O 设备，它们具有一定的数据处理和发送、接收数据的能力，通常称为数据终端设备 DTE（Data Terminal Equipment）。分组交换网由若干个分组交换机 PSE（Packet Switching Equipment，节点交换机）和连接这些节点的通信链路组成。与 DTE 对应的是数据电路终接设备 DCE（Data Circuit - terminating Equipment）。DCE 指的是

DTE－DTE 远程通信传输线路的终接设备；在物理上，如果传输线路是模拟通道，DCE 就是 MODEM；如果是数字通道，DCE 就是多路复用器或数字通道接口设备。它们提供信号变换、适配和编码功能，和 DTE 同属于用户设施。但是在功能结构上，DCE 属于网络部分，是分组交换机的延伸。

分组交换采用“存储-转发”技术。这种技术最早出现在报文交换。在报文交换中，当源站发送报文时，将目的地址添加在报文中，然后网络中的交换机将源站的报文接收后暂时存储在存储器中，再根据提供的目的地址，不断通过网络中的其他交换机选择空闲的路径转发，最后送到目的地址。这样就解决了不同类型用户之间的通信，并且不需要像电路交换那样在传输过程中长时间建立一条物理通路，而可以在同一条线路上以报文为单位进行多路复用，所以大大提高了线路的利用率。分组交换中所采用的“存储-转发”技术并不像报文交换那样以报文为单位进行交换，而是将报文划分成有固定格式的分组（Packet）进行交换、传输，一般为 1kbit～$n$kbit，每个分组按一定格式附加源与目的地址，分组编号、分组起始、结束标志、差错校验等信息，以分组形式在网络中传输。当源 DTE 将分组传送至本地分组交换机后，本地分组交换机收到每个分组要求的转发信息，不管是否接通目的地址设备，都先存储起来，然后检查目的地址，在分组交换机保存的路由表中找到该目的地址规定的发送通路，分组交换机即按允许的最大发送速率转发该分组。同样，每个中转分组交换机均按此方式存储、转发每个分组，直到将分组送到目的地分组交换机，再由该分组交换机送达目的 DTE。

按上述方式传送的是分组交换中的数据报方式。一般适用于较短的单个分组的报文。其优点是传输可靠性高、传输延时小，由于分组交换机的存储器容量减小，所以提高了经济性。缺点是每个分组附加的控制信息多，增加了传输信息的长度和处理时间，增大了额外开销。

分组交换的另一种方式叫虚电路方式，它与数据报方式的区别主要是在信息交换之前，由源 DTE 向本地分组交换机发送一个特定呼叫请求的分组，其中含有目的 DTE 的地址及逻辑信道识别符，并由分组交换机 PSE 中转转发。若呼叫被目的 DTE 接受，则相应的响应“呼叫接受”予以应答，网络即发出一个“呼叫连通”给源 DTE，此时呼叫建立，在两台 DTE 之间建立一条称作虚电路的逻辑通路，信息就能在这条虚电路上传输，直到数据交换结束，虚电路被拆除，相应的逻辑信道识别符被释放。所以虚电路方式在每次通信时都有虚电路建立、数据传输和拆除三个阶段，类似于电路交换方式，但在网络中的传输是分组交换方式。这种方式对信息传输频率高、每次传输量小的用户不太适用，但由于每个分组头只需标出虚电路标识符和序号，所以分组头开销小，适用长报文传送。虚电路又可分为永久虚电路 PVC（Permanent Virtual Circuit）和交换式虚电路 SVC（Switch Virtual Circuit）。PVC 由网络提供者配置，一旦完成，这种虚电路即长期存在。SVC 则需要由两个远程端用户通过相应的控制协议来建立，在完成数据传输后被拆除。

信元交换技术是一种快速分组交换技术，它结合了电路交换技术延迟小和分组交换技术灵活的优点。信元是固定长度的分组，ATM（Asynchronous Transfer Mode，异步传输模式）采用信元交换技术，其信元长度为 53 字节。由于信元的长度更小，交换所需的时延更少。

2. 计算机网络的软件

计算机网络的软件构成主要包括有网络操作系统软件、网络通信协议、网络工具软件、网络应用软件等。

网络操作系统软件：负责管理和调度计算机网络上的所有硬件和软件资源，使各个部分能

够协调一致的工作。常用的网络操作系统有 Windows、Netware、Unix、Linux 等。

网络通信协议：计算机网络中的数据交换必须遵守事先约定好的规则。这些规则明确规定了所交换的数据格式以及有关的同步问题(同步含有时序的意思)。为进行网络中的数据交换而建立的规则、标准或约定即网络协议(network protocol)，简称为协议。网络协议包括以下三个要素：①语法：数据与控制信息的结构或格式；②语义：需要发出何种控制信息，完成何种动作以及做出何种响应；③同步：事件实现顺序的详细说明。常用的网络通信协议有 TCP/IP 簇、SPX/IPX、NetBEUI 协议等。

网络工具软件：用来扩充网络操作系统功能的软件；如网络浏览器、网络下载软件、网络数据库管理系统等。

网络应用软件：基于计算机网络应用而开发出来的用户软件。如民航售票系统、远程物流管理软件、订单管理软件、酒店管理软件等。

通常提到计算机网络的协议，总是和体系结构的概念分不开。计算机网络的体系结构(architecture)是计算机网络的各层及其协议的集合。这里说的“层”是一种在计算机网络中所使用的方法。通过分层将庞大而复杂的问题，转化为若干较小的局部问题。这些较小的局部问题就比较容易研究和处理。每相邻层间有一接口，下层通过接口向上层提供某种服务，完成特定功能，同时还对上层屏蔽实现该功能的具体过程，使上层可以只简单地使用下层提供的服务而不必关心其具体的实现细节；上层又在其下层提供的服务基础上，向更高层提供更高级的服务。于是，通过接口，各层协议之间能高效地相互作用，协同解决整个通信问题。这种化整为零的思想对计算机网络的研究起到了很大的促进作用。计算机网络大都按层次结构模型去组织计算机网络协议。例如，IBM 公司的系统网络体系结构 SNA。影响最大、功能最全、发展前景最好的网络层次模型，是国际标准化组织(ISO)所建议的“开放系统互连(OSI)”基本参考模型。它由物理层、数据链路层、网络层、运输层、会话层、表示层和应用层等 7 层组成。各层的一些典型服务、标准和协议如下：

7. 应用层(Application Layer)：Http，DNS，Telnet，SMTP，FTP；
6. 表示层(Presentation Layer)：ASCII，EBCDIC，QuickTime，MPEG，GIF，JPG，TIFF；
5. 会话层(Session Layer)：ZIP，NFS，SQL；
4. 运输层(Transport Layer)：TCP，SPX，UDP，NBP，OSI transport protocol；
3. 网络层(Network Layer)：IP，IPX，BGP，OSPF；
2. 链路层(Datalink Layer)：HDLC，PPP；
1. 物理层(physical Layer)：RS232，RS449。

通俗地理解，OSI 模型将一系列复杂的计算机通信问题分解为 7 类，分别进行研究。而且，这些分工的特点是“越往上离接受应用服务的用户越近，越往下离机器越近。”

在计算机网络的分层模型中还有一个很重要的概念“封装”(encapsulation)，封装是在数据前加上报头或者将数据包在首尾里面的过程。封装在 OSI 参考模型的每层上都会出现。来自每层的完整的数据包将插入到下一个层的数据字段中，并且加入另外一个报头。在偶然情况下，层会将一个数据信元(包括前一层的报头)分开为多个部分，更小的数据信元，并且每个更小的数据信元用较低协议层的新报头进行封装。这个过程帮助控制数据流，因为不同的网络允许通过的最大传输单元 MTU(Maximum Transmission Unit)不尽相同。当接收到数据时，接收节点上的对应层在把数据传送到下一个层之前，重新装配数据字段。随着数据逐渐

在目的地的模型上向上移动，逐渐将分段拼装到一起。图 1.3 显示了数据在各层之间传递时进行封装和拆封的这一过程。

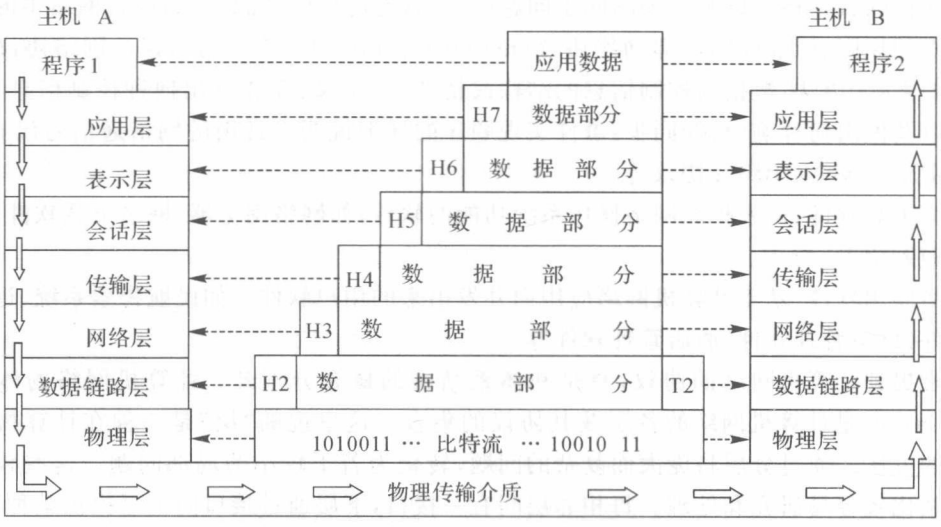


图 1.3 数据在各层之间的传递过程

需要注意的是法律上的国际标准 OSI 并没有得到市场的认可。而非国际标准 TCP/IP 却获得了最广泛的应用。TCP/IP 常被称为事实上的国际标准。TCP/IP 事实上并没有严格的层次体系结构，它们只是在因特网中广泛使用的一系列协议。在设计之初并不具有像 OSI 模型那样强的模型指导作用。所以通常将之称为 TCP/IP 协议簇而不是体系结构。如果用分层的思想去描述 TCP/IP 协议簇，会发现它的层次只有 4 层：高层应用、传输层、网际层、网络接口层。严格意义上的层只有两层：在传输层对高层应用提供可靠的(通过 TCP 协议)和不可靠的(通过 UDP 协议)数据传输服务(这里提到的可靠服务，下文再展开讨论)；在网际层通过 IP 及其相关协议来屏蔽下层各种网络的不同，实现网络的互连。为了便于理解并能够和实际网络中的现状接轨，一些学者提出了一种五层协议的网络体系结构。所谓五层协议的网络体系结构是为便于学习计算机网络原理而采用的综合了 OSI 七层模型和 TCP/IP 的四层模型而得到的五层模型。五层协议的体系结构如图 1.4 所示。

在这种 5 层协议的参考模型中，各层的主要功能如下：

(1)应用层。应用层确定进程之间通信的性质以满足用户的需要。应用层不仅要提供应用进程所需要的信息交换和远地操作，而且还要作为互相作用的应用进程的用户代理(user agent)，来完成一些为进行语义上有意义的信息交换所必须的功能。

图 1.4 五层协议的参考模型

(2)运输层。任务是负责主机中两个进程间的通信。因特网的运输层可使用两种不同的协议。即面向连接的传输控制协议 TCP 和无连接的用户数据报协议 UDP。

(3)网络层。网络层负责为分组选择合适的路由，使源主机运输层所传下来的分组能够交付到目的主机。

(4)数据链路层。数据链路层的任务是将在网络层交下来的数据报组装成帧(frame),在两个相邻结点间的链路上实现帧的无差错传输。

(5)物理层。物理层的任务就是透明地传输比特流。“透明地传送比特流”指实际电路传送后比特流没有发生变化。物理层要考虑用多大的电压代表“1”或“0”,以及当发送端发出比特“1”时,接收端如何识别出这是“1”而不是“0”。物理层还要确定连接电缆的插头应当有多少根针脚以及各个针脚如何连接。

这里有一个非常值得读者总结一下的内容,就是各层的数据格式和所使用的地址或者类似地址作用的标识:

5. 应用层:数据形式就是各种应用报文(message),使用域名(domain name)来表示网站和主机的名字,与 IP 地址等效使用;

4. 运输层:数据格式为报文段(Segments),利用端口来标识高层的应用进程;

3. 网络层:数据格式为分组或数据报(Packets/datagrams),因特网中利用每个主机唯一的合法 IP 地址来找到主机所在网络。注意:分组有时也被译为包;

2. 数据链路层:数据的格式为帧(Frames),使用硬件地址来标识每台主机,并利用主机 IP 地址与硬件地址(也叫物理地址)的映射关系来找到主机;

1. 物理层:比特流(bits)。

通常,在每一层提供的服务中,不丢失、不重复、无差错的传输称之为可靠服务。为了实现可靠服务,通常都会采用面向连接、确认、序号、计时器、流量控制以及拥塞控制等机制来实现。而实现这些机制就需要付出硬件、软件方面的代价。传统的电话网络就设计成一种非常可靠的网络。用户使用非常廉价的电话机就能够享受到清晰的通话质量。电信网负责保证可靠通信的一切措施,因此电信网的节点交换机复杂而昂贵。但这种网络的脆弱性也是显而易见的,一旦电信网的关键节点遇到摧毁,整个通信系统就会瘫痪。

因特网当初的设计思想则不同:网络尽量简单,而智能尽可能放在网络以外的用户端。在计算机网络中,用户所使用的端系统是装载了协议栈的计算机。可靠通信由用户终端中的软件(即 TCP)来保证。所以在层次结构的参考模型中,4 层以上的功能都在网络之外的端系统中。技术的进步使得网络出错的概率越来越小,因而让主机负责端到端的可靠性不但不会给主机增加负担,反而能够使更多的应用在这种简单的网络上运行,大大简化了网络层的结构。

3. 带宽与时延

带宽和时延是计算机网络中两个重要的性能指标,也是最基本的概念。

带宽(bandwidth)本来是通信中的术语,指的是信号所具有的频带宽度,单位是赫兹(Hz)。在计算机网络中,借用这个名词来表示数字信道所能传送的“最高数据率”,单位是比特每秒 b/s (bit/s)。更常用的带宽单位是千比每秒 kb/s(10^3 b/s)、兆比每秒 Mb/s(10^6 b/s)、吉比每秒 Gb/s(10^9 b/s)、太比每秒 Tb/s(10^{12} b/s)。注意在表示带宽时,k、M、G、T 是指 10 的幂。而通常在表示存储容量(字节 B)时所用的 K = 2^{10} = 1024,M = 2^{20}, G = 2^{30}, T = 2^{40}。

时延(delay)是指一个报文或分组从一个网络的一端传送到另一端所需的时间。时延包括三部分:总时延=传播时延+发送时延+处理时延。

其中发送时延是节点在发送数据时使数据块从节点进入到传输媒体所需要的时间。

$$\text{发送时延}=\frac{\text{数据块长度}}{\text{信道带宽(数据在信道上的传输速率)}}$$

传播时延是电磁波在信道中需要传播一定的距离而花费的时间。

$$\text{传播时延}=\frac{\text{信道长度}}{\text{电磁波在信道上的传播速率}}$$

处理时延是数据在交换结点为存储转发而进行一些必要的处理所花费的时间。处理时延的长短往往取决于网络中当时的通信量。

在总时延中，究竟是哪一种时延占主导地位，必须具体分析。通常所说的高速链路，指的是数据的发送速率快而不是比特在链路上的传播速率快。电磁波在特定媒体上传播速率是恒定的。

1.1.2 网络互连原理与 IP 网

1. 网络互连问题

制定体系结构的目的就是为了规范计算机网络的发展，但是实际上，不论是广域网还是局域网都存在着大量的异构网络。各层运行着各种不同的协议。Andrew S. Tanenbaum 教授将网络的这些不同总结为 12 点。

(1)网络所提供的服务不同：有面向连接的报务，也有不连接的服务；
(2)协议不同：比如 IP、IPX、SNA、ATM、MPLS、AppleiTalk 等；
(3)编址方式不同：局域网所采用的地址通常都是平面的，而广域网所采用的地址通常是层次的；
(4)多播和广播的支持：有的网络支持，有的不支持；
(5)分组大小：每个网络都有自己的 MTU 限制；
(6)服务质量 QoS(Quality of Service)：不支持或者采用不同的种类；
(7)错误处理：可能会是可靠的、有序的，以及无序的递交；
(8)流量控制：滑动窗口、速率控制，或者其他控制手段，也可能干脆无控制；
(9)拥塞控制：漏桶、令牌桶、RED、抑制分组等；
(10)安全性：隐私规则、加密等；
(11)一些参数：不同的超时值、流规范等；
(12)记费方式：按连接时间、按分组、按字节，或者根本不记费。

要实现这些异构网络的连接，首先需要的就是互连的设备。通常来说，这些设备工作于不同的层次。而严格意义上的互连只发生在网络层。也就是说只有见到碰到网络层的设备，才认为是两个网络的互连。网络层以下的设备只是实现网络内部的拓展或者信号的中继。网络层以上的设备则是为了实现高层协议的转换。

这些工作于不同层次的设备总结如下：

高层：网关(gateway)，用来实现协议转换。例如传输层网关可以转换 TCP 连接和 SNA 连接，而应用层网关可以翻译消息的语义；
网络层：路由器(router)，用来实现转发分组和路由选择等网络互连任务；
数据链路层：网桥(bridge)、交换机(switch)，实现局域网的拓展；
物理层：中继器(repeater)、集线器(hub)完成信号的放大转发。

注意：由于历史的原因，许多有关 TCP/IP 的文献将网络层使用的路由器称为网关。尤其是微软的 Windows 操作系统中，配置 TCP/IP 的属性时，需要配置的网关，实际上指的就是对外互连的路由器地址。

2. 网络互连的协议

因特网的网际层，使用 IP 协议来屏蔽这些异构网络下层通信技术的不同。互连起来的各种物理网络的异构性本来是客观存在的，但是利用 IP 协议就可以使这些性能各异的网络让端用户看起来好像是一个统一的网络。通常，使用 IP 协议的互连网络常简称为 IP 网。对于端系统而言，看不见互连的各具体的网络异构细节，就好像在一个网络上通信一样。

IP 协议提供无连接的数据报传输机制。IP 协议是点到点的，核心问题是寻径。它向上层提供统一的 IP 数据报，使得各种物理帧的差异对上层协议不复存在。

TCP/IP 体系中与 IP 协议配套使用的还有三个协议：地址解析协议 ARP（Address Resolution Protocol）、逆地址解析协议 RARP（Reverse Address Resolution Protocol）、Internet 控制报文协议 ICMP(Internet Control Message Protocol)。

图 1.5 表示了这三个协议和 IP 协议的关系。在这一层中，ARP 和 RARP 画在最下面，因为 IP 经常要使用着两个协议。ICMP 画在这一层的上部，因为它要使用 IP 协议。

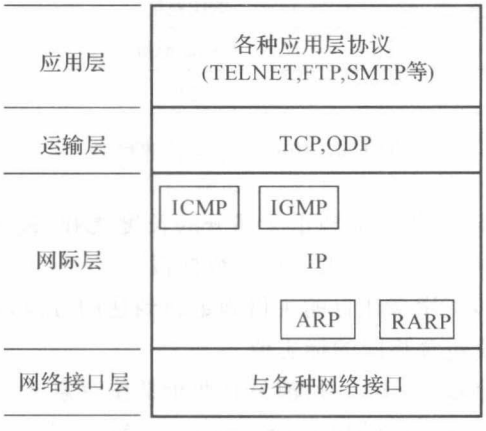


图 1.5　IP 及其配套协议

IP 是 TCP/IP 协议族中最为核心的协议。所有的 TCP、UDP、ICMP 及 IGMP 数据都以 IP 数据报格式传输。IP 数据报以一个头部开始，后跟数据区。一个数据报的数据长度不固定，数据报的大小取决于发送数据的应用。大小可变的数据报使得 IP 可以适应各种应用。但是，采用较大的数据报可以获得更高的效率。目前，已经有 2 种 IP 版本成为标准，它们分别是 IPv4 和 IPv6，后者是前者的升级。现在网络正在使用的是 IPv4。

3. IPv4 的报文

IPv4 报文首部包括一个 20 字节的固定部分和一个可变长度的可选部分，如图 1.6 所示。

(1)版本(version)。说明数据报属于哪一个协议版本，以便可以在运行不同版本协议的机器之间进行版本转换。IPv4 和 IPv6 即在此标示，当该域值为 4 时，表示 IPv4。

(2)首部长度(IHL)。说明包头的长度(单位：4 字节)，最小为 5，最大为 15。故头部最长为 60 字节，即可选部分最大为 40 字节。该域值变化 1，表示包头长度变化 32 个字节。此外，

对于有些可选项，例如记录分组已经走过路由的源路由选项，40 字节就显得太短了。

(3)服务类型(type of service)。允许主机告诉子网它需要什么类型的服务，可能是可靠程度和传输速率的各种组合。例如，对数字话音要求快速传递；而对文件传输无差错比快速更重要。该域中，左起 3 位为优先级(precedence)字段，从 0(正常)到 7(网络控制分组)。后跟 3 个标识(flag)位分别表示延迟、吞吐量和可靠性，它们允许主机指明在以上三项指标中它最关心什么。最后两位没有定义。理论上，这些字段允许路由器在吞吐量大而时延长的卫星链路和吞吐量小而时延短的租用线路之间进行选择。实际上，目前的路由器都不支持服务类型字段。

图 1.6 IPv4 的报文首部格式

(4)总长度(total length)。指头部和正文部分的长度之和，最大为 65535 字节。(目前允许这一上限，但将来的千兆位网络将要求更长的数据报。)

(5)标识(identification)。用来让目的主机确定新到达的分段(fragment)属于哪一个数据报。同一数据报的所有分段包含相同的标志值。

(6)标志(flag)。标志字段占 3bit。目前只有两个低位有意义。

标志字段中的最低位记为 MF(More Fragment)。MF＝1 即表示后面还有分段的数据报。MF＝0 表示这已是若干数据报段中的最后一个。

标志字段中间的一位记为 DF(Don't Fragment)。只有当 DF＝0 时才允许分段。

(7)片偏移量(fragment offset)。告知本分段在当前数据报的位置。除了最后一个分段以外，一个数据报的所有分组必须是 8 字节的倍数，即 8 字节为一个基本分段单位。该域有 13 位，所以每个数据报最多有 8192 个分段，数据报长度最大可达到 65536 字节，比总长度域的最大值大 1 个字节。

(8)生存期 TTL(time to live)。是用来限制分组寿命的计数器，最长生存期为 255 秒。该域在每条链路上都必须递减。若在某个路由器中排了长时间的队，则要以倍数递减。实际上，它只计算链路上的时间。当该域减为 0 时，就将这一分组丢弃，并向源主机发送告警分组。

(9)协议(protocol)。告诉网络层把收到的数据报送给哪一个传输层进程，可能是 TCP，也可能是 UDP 或其他。协议编号在整个 Internet 中是全局唯一的，定义参考 RFC 1700。

(10)首部校验和(header checksum)。只验证 IP 分组头。每条链路中该域都必须重新计算，因为至少有一个域(生存期域)的值是一直在变化的。前面提到过了，每个路由器都会把收

到数据报的 TTL 值减 1。

(11)源地址(source address)和目的地址(destination address)。指明发送数据报的源和目的地址。

(12)选项(options)。用来提供一种余地,使协议的后来版本可以包含原有设计中没有的信息,也可以使试验者能尝试他的新想法。选项域的长度是可变的,每个选项都以一个字节表明内容。某些选项还跟有一个字节的选项长度字段,其后是一个或多个数据字节。选项域以 4 个字节的倍数来安排。目前定义了 5 种选项。

4. IP 编址

IP 报文所使用的地址是 IP 地址。所谓 IP 地址就是给每一个连接在 Internet 上的主机分配一个唯一的 32bit 地址。在 IPv4 中 IP 地址由 4 个字节组成,被表示成用“.”隔开的 4 组 10 进制数,每个数最大为 255。这种表示方法被称为点分十进制表示法,即将每个字节值用十进制数表示。例如 IP 地址 11001000.01100100. 01100100.00000001 的点分十进制表示为 200.100.100.1。IP 地址被分成两部分,按层次结构组成:第一部分是网络号,第二部分是主机号。分组从一个路由器传到另一个路由器就是一跳(hop),它就是这样经过若干跳,最后到达目的网络。在那里,路由器将它送到目的主机。

IP 地址的最初编码是分类的。尽管目前的用法是无类的,这里,还是按照分类的编址方式来讨论目前的地址结构,而且这种结构偶尔还在用。在分类编址方式中,有 5 类地址,如图 1.7 所示,分别是 A 类到 E 类:

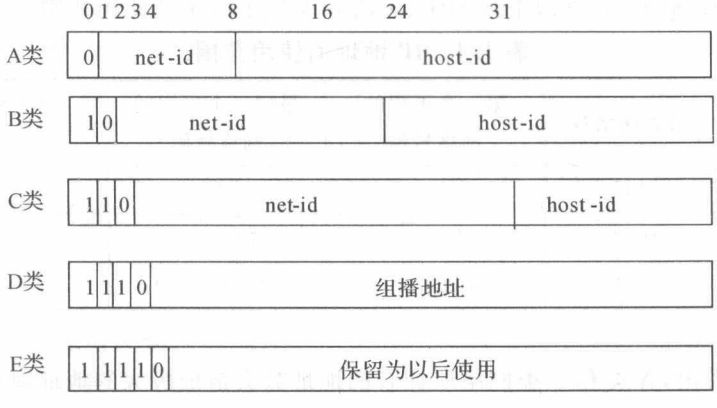
A 类地址,有 8 位网络号,网络号的开头 1 位是 0,后面是 24 位主机地址。因此一个 A 类地址的网络可以有 $2^{24}-2$ 个主机。之所以要减 2 是因为:主机地址部分为 0 时,代表了该主机所在的网络号,主机地址部分全部是 1 是代表广播地址。这两个值都不能代表单个主机地址,因此,要总主机个数上要减去 2。

B 类地址,有 16 位网络号,网络号的开始 2 位是 10,后面是 16 位主机号,因此一个 B 类网络可有 $2^{16}-2$ 个主机。

C 类地址,有 24 位网络号,网络号的开始 3 位是 110,后面是 8 位主机号,因此一个 C 类网络可有 256−2=254 个主机。

D 类地址,地址的开始 4 位是 1110,这是组播地址,将在后面讨论。

E 类地址,以 1111 开头,这类地址保留为以后用。



net-id—网络号码, host-id—主机号码

图 1.7　IP 地址的 5 种类型

在IP地址的使用中，有一些具有特殊意义的保留地址，总结如下：

(1)0.0.0.0。严格说来，0.0.0.0已经不是一个真正意义上的IP地址了。它表示的是这样一个集合：所有不清楚的主机和目的网络。这里的“不清楚”是指在本机的路由表里没有特定条目指明如何到达。对本机来说，它就是一个“收容所”，所有不认识的“三无”人员，一律送进去。如果你在网络设置中设置了默认网关，那么Windows系统会自动产生一个目的地址为0.0.0.0的默认路由。

(2)255.255.255.255。限制广播地址。对本机来说，这个地址指本网段内(同一广播域)的所有主机。如果翻译成人类的语言，应该是这样：“这个房间里的所有人都注意了!”这个地址不能被路由器转发。

(3)127.0.0.1。127.0.0.1～127.255.255.254范围内的地址称为环回地址。最常用于测试的环回地址是127.0.0.1。表示“自己”。在Windows系统中，这个地址有一个别名“Localhost”。寻址这样一个地址，是不能把它发到网络接口的。除非出错，否则在传输介质上永远不应该出现目的地址为“127.0.0.1”的数据包。

(4)224.0.0.1。组播地址，注意它和广播的区别。从224.0.0.0到239.255.255.255都是这样的地址。224.0.0.1特指所有主机，224.0.0.2特指所有路由器。这样的地址多用于一些特定的程序以及多媒体程序。如果你的主机开启了IRDP(Internet路由发现协议，使用组播功能)功能，那么你的主机路由表中应该有这样一条路由。

(5)169.254.x.x。如果主机使用了DHCP功能自动获得一个IP地址，那么当你的DHCP服务器发生故障，或响应时间太长而超出了一个系统规定的时间，Windows系统会为你分配这样一个地址。如果你发现你的主机IP地址是一个诸如此类的地址，很不幸，十有八九是你的网络不能正常运行了。

(6)10.x.x.x、172.16.x.x～172.31.x.x、192.168.x.x。私有地址，这些地址被大量用于企业内部网络中。一些宽带路由器，也往往使用192.168.1.1作为默认地址。私有网络由于不与外部互连，因而可能使用随意的IP地址。保留这样的地址供其使用是为了避免以后接入公网时引起地址混乱。使用私有地址的私有网络在接入Internet时，要使用地址翻译(NAT)，将私有地址翻译成公用合法地址。在Internet上，这类地址是不能出现的。

对一台网络上的主机来说，它可以正常接收的合法目的网络地址有三种：本机的IP地址、广播地址以及组播地址。

对于分类的IP地址，总结以上的约定，可得到表1.1所示的使用范围。

表1.1　IP地址的使用范围

网络类别	最大网络数	第一个可用的网络号码	最后一个可用的网络号码	每个网络中的最大主机数
A	126	1	126	16777214
B	16382	128.1	191.254	65534
C	2097150	192.0.1	223.255.254	254

从表中可以看出，A类和B类网络所拥有的地址数会造成极大的地址浪费。而实际上32位的IP地址，其地址空间本身已经出现耗尽的危机。目前，有两种方法来解决这一危机(当