实时性　实用性　可靠性

TCP/IP Analysis and Troubleshooting Toolkit

TCP/IP 分析与故障诊断

(美) Kevin Burns　著
战晓苏　张　敏　译

清华大学出版社

Kevin Burns

TCP/IP Analysis and Troubleshooting Toolkit

EISBN：0-471-42975-9

北京市版权局著作权合同登记号　图字：01-2003-7899

图书在版编目(CIP)数据

TCP/IP 分析与故障诊断/(美)本斯(Burns,K.)著；战晓苏等译. —北京：清华大学出版社，2005.1

书名原文：TCP/IP Analysis and Troubleshooting Toolkit

ISBN 7-302-08984-1

Ⅰ.T…　Ⅱ.①本…②战…　Ⅲ.计算机网络—通信协议—故障诊断　Ⅳ.TN915.04

中国版本图书馆 CIP 数据核字(2004)第 066373 号

出 版 者：清华大学出版社　　**地　　址**：北京清华大学学研大厦
http://www.tup.com.cn　**邮　　编**：100084
社 总 机：010-62770175　**客户服务**：010-62776969
组稿编辑：曹　康
文稿编辑：王　黎
封面设计：康　博
版式设计：康　博
印 刷 者：北京季蜂印刷有限公司
装 订 者：三河市新茂装订有限公司
发 行 者：新华书店总店北京发行所
开　　本：185×260　**印张**：18.75　**字数**：480 千字
版　　次：2005 年 1 月第 1 版　2005 年 1 月第1 次印刷
书　　号：ISBN 7-302-08984-1/TP・6354
印　　数：1～4000
定　　价：38.00 元

前　言

为什么写本书

网络工程师每天都面对巨大的挑战。服务器可能崩溃，广域网链路可能达到饱和，或者由于某些未知的原因应用程序的性能变得像爬虫蠕动般缓慢，这些都导致了网络工程师和应用程序开发人员毫无根据地相互指责。当问题发生时，如果没有适当的工具和培训，网络工程师通常不得不问：为什么用户得不到 DHCP 地址？为什么用户不能登录到服务器？甚至还有更令人厌烦的问题：为什么网速这么慢？在所有这些混乱中，上层管理人员也经常要问：为什么这些问题没有得到解决？大部分的大型网络基础结构在其配置中都有综合故障检修工具，但是他们通常在错误的工作中选用了错误的工具。如何才能最好地使用配置中的工具和自己的网络知识来快速、果断地解决网络基础结构中的问题？本书内容就是对这个问题的解答。

作者写这本书就是为了前面提到的网络工程师。我们十分尊敬他们，他们是实干家，是他们使网络正常运行，而当网络不能运行时又是他们最先得到通知。在我们支持台式计算机、服务器及大型复杂网络基础结构的十多年经验中，我们得到一个结论：最好的网络工程师就是那些能真正解决疑难故障的人。

那些善于解决故障的人通常都是顽强的和有求知欲的。这两种品质可以使这些人整晚不睡觉去尝试着解决一些故障。他们知道答案就在某个地方等待着他们去揭示，而他们也一直顽强地钻研，直到他们找到答案为止。那些真正有求知欲的人很可能读了许多关于 TCP/IP 协议的好书，其中包括 W.Richard Stevens 的 *TCP/IP Illustrated*(Addison-Wesley，1994 年 1 月出版)和 Douglas Comer 的 *Internetworking with TCP/IP*(Prentice Hall 2000 年 1 月出版)这两本书。迄今为止，这些书仍是了解 TCP/IP 的标志性书籍，但是这些书主要集中于理论而缺乏实际的例子(当然，我们仍然建议每一个分析员自己的书架要有这两本书)。我们找出协议中最重要的概念并将它们应用于网络分析员在 TCP/IP 网络中遇到的最常见的故障中，以此来填补以上两本书与实际应用之间的差距。为了满足读者进一步的求知欲望，对于那些对复杂细节及协议内部操作有兴趣的读者，我们还提供了附录 A，附录 A 中有对此进行进一步详述的站点。

本书的目的就是为读者提供成功维护现实网络中的协议所必需的信息。由于 TCP/IP 是现在最通用的协议，因此就很容易决定整本书都集中于 TCP/IP 的分析和 TCP/IP 故障诊断方法。作者试图介绍一些可帮助您掌握理解并诊断与 TCP/IP 协议相关问题的知识，而不是一本有很多复杂和通俗细节的书。

您很快就将认识到这本书中的很多例子都是针对 Cisco 或 Microsoft 公司产品的。因为这两个供应商的产品是目前使用最普遍的，所以作者选用了和它们系统相关的例子。这些例子决不是 Cisco 或 Microsoft 公司所特有的，在绝大多数的示例中，可以将书中的例子应用到任何一个供应商的硬件或软件上。特定的例子则应用于某个标注的供应商。顺着这条线，您会注意到在

例子中提到和使用了许多分析工具。工具的类型不是最重要的，只要它能提供所必须或所描述的功能就行。

对技术的理解是最重要的，同时这也是这本书着重介绍的内容。

本书读者对象

虽然本书提供了网络分析技术和 TCP/IP 协议的简介，但此书并不是针对初学者的。读者对 OSI 模型需要有一个基本的理解，对运行 TCP/IP 的服务器操作系统也要有一定的管理经验。

很多已经熟悉协议的高级读者将能够从每章的示例研究中得到更多的收获。本书将帮助您成为一个更加出色的网络分析员。如果您是一位网络管理员并渴望对客户机和服务器之间的通信有更多的了解，那么本书可以作为一个好的起点。如果您已经熟悉路由器和交换机的配置，那么本书将教会您配置命令中的技巧，教会您打破常规的思考方式。

本书介绍各种技巧以及如何在配置中以最佳方式使用工具来保持网络的平稳运行。

本书的组织结构

本书分成三个部分

- **第 I 部分：网络分析基础。**将回答“为什么要进行协议分析”以及“应该使用什么工具”等问题。该部分解释了捕获和操作跟踪文件的过程。同时对 OSI 模型和网络通信的基本概念进行了复习，复习这些知识有助于您学习后面的章节。
- **第 II 部分：核心协议。**建立了理解这些核心协议的基础，TCP/IP 正是基于这些核心协议而构建的。正是这些协议为应用层的其他所有协议提供了支持。
- **第III部分： TCP/IP 相关协议。**通过揭示实现标准协议和供应商无关协议的内部工作原理来扩展对它们的了解和研究。该部分详细分析了域名系统(Domain Name System，DNS)、超文本传输协议(Hypertext Transport Protocol，HTTP)以及文件传输协议(File Transport Protocol，FTP)等应用程序，并且深入研究了 Microsoft 公司的 TCP/IP 实现，其中包括了服务器消息块协议(Server Message Block Protocol)。

每章都补充了很多来自实际网络中的例子和示例研究。这些例子和示例研究用来说明如何应用所讨论的知识和技术。

工具

本书使用了几个不同的分析工具来阐明故障诊断的例子。然而理解这些例子不必掌握这些工具，查看合作中网站的跟踪文件时需要用到这些工具。网站会指导您下载 Ethereal 协议分析仪的免费版本，该软件用于查看跟踪文件。

合作网站

本书的合作网站(可以通过在浏览器中输入以下的地址访问：www.wiley.com/compbooks/burns)包括了诸如RFC(Request for Comment，请求注解)和IETF (Internet Engineering Task Force，Internet 工程任务组)制定的标准协议，还包括了与本书讨论的协议相关的其他资源。该网站还包括了本书中大部分例子的在线视频和示例研究中的跟踪文件，您可以自己加载和检查这些内容。最后，它还提供了网络分析员的工具包中所必需的几种共享软件和免费软件。如果需要了解该网站的更多信息，请参见附录A。

目　　录

第 I 部分　网络分析基础

第 1 章　协议分析导论 …… 1
1.1　网络通信简史 …… 1
1.2　OSI 解决方案 …… 2
1.2.1　定义层 …… 3
1.2.2　各层的协议分析 …… 4
1.2.3　总结 …… 16
1.3　TCP/IP 的历史 …… 16
1.4　本章小结 …… 18

第 2 章　分析工具与分析方法 …… 19
2.1　网络管理工具回顾 …… 19
2.1.1　按功能对网络管理工具分类 …… 19
2.1.2　按功能的执行方式对网络管理工具分类 …… 21
2.2　解决问题的工具—— 协议分析仪 …… 22
2.2.1　为什么要进行协议分析 …… 23
2.2.2　协议分析仪的功能 …… 24
2.2.3　配置并使用分析仪 …… 30
2.2.4　分析提示 …… 42
2.3　本章小结 …… 45

第II部分　核 心 协 议

第 3 章　网际协议内部运行机制 …… 46
3.1　回顾第 2 层通信 …… 46
3.1.1　多路复用 …… 46
3.1.2　差错控制 …… 47
3.1.3　寻址 …… 47
3.1.4　示例研究：NetBEUI 通信 …… 48
3.1.5　第 2 层通信网络的局限 …… 51
3.2　网络层协议 …… 52

3.3 网际协议寻址 …… 53
3.3.1 IP 寻址 …… 54
3.3.2 保留寻址 …… 57
3.3.3 类别寻址 …… 58
3.3.4 无类寻址 …… 59
3.4 IP 通信 …… 61
3.4.1 地址解析协议(ARP) …… 62
3.4.2 IP 路由选择 …… 70
3.5 IP 报文格式 …… 80
3.5.1 版本 …… 80
3.5.2 报文头长 …… 80
3.5.3 服务类型 …… 80
3.5.4 数据报长 …… 81
3.5.5 段 ID …… 81
3.5.6 分段标志 …… 82
3.5.7 段偏移 …… 82
3.5.8 生存期 …… 82
3.5.9 协议 …… 82
3.5.10 报头校验和 …… 83
3.5.11 源端 IP 地址 …… 83
3.5.12 目的端 IP 地址 …… 83
3.5.13 选项 …… 83
3.5.14 数据 …… 83
3.5.15 示例研究：TTL 过期 …… 83
3.5.16 示例研究：本地路由选择回顾 …… 85
3.6 IPv6 简介 …… 87
3.6.1 IPv6 报文头 …… 88
3.6.2 IPv6 地址格式 …… 89
3.6.3 IPv6 的其他变化 …… 90
3.7 本章小结 …… 90

第 4 章 网际控制报文协议 …… 91
4.1 网络的可靠性 …… 91
4.1.1 面向连接的网络与无连接网络 …… 92
4.1.2 反馈 …… 92
4.2 对 ICMP 的研究 …… 92
4.2.1 ICMP 报文头 …… 93
4.2.2 ICMP 类型与代码 …… 93

4.2.3 ICMP 报文细节 ····· 95
4.2.4 使用 ICMP 进行网络诊断 ····· 106
4.3 本章小结 ····· 108

第 5 章 用户数据报协议 ····· 109
5.1 传输层回顾 ····· 109
5.2 UDP 报文头 ····· 110
5.2.1 源端口 ····· 110
5.2.2 目的端口 ····· 110
5.2.3 UDP 长 ····· 111
5.2.4 UDP 校验和 ····· 111
5.2.5 数据 ····· 112
5.3 UDP 通信过程 ····· 112
5.4 UDP 通信示例研究 ····· 117
5.4.1 名字解析服务 ····· 117
5.4.2 路由选择信息协议 ····· 118
5.4.3 简单网络管理协议 ····· 120
5.4.4 UDP 与防火墙 ····· 120
5.4.5 路由跟踪警告 ····· 123
5.5 本章小结 ····· 124

第 6 章 传输控制协议 ····· 125
6.1 TCP 简介 ····· 125
6.1.1 可靠传输协议的要求 ····· 125
6.1.2 TCP 报文头 ····· 128
6.1.3 TCP 实现 ····· 130
6.2 TCP 连接管理 ····· 131
6.2.1 TCP 连接打开 ····· 131
6.2.2 TCP 连接关闭 ····· 136
6.2.3 半关闭 ····· 137
6.2.4 TCP 重置 ····· 137
6.3 TCP 通信量管理 ····· 141
6.3.1 数据排序和确认 ····· 141
6.3.2 TCP 重传 ····· 143
6.3.3 延迟确认 ····· 145
6.3.4 Push 标志 ····· 146
6.3.5 TCP 滑动窗口 ····· 147
6.3.6 慢启动和避免拥塞 ····· 150
6.3.7 Nagle 算法 ····· 151

6.3.8 数据保护……152
6.3.9 TCP 专家诊断……154
6.4 TCP 应用程序分析……155
6.4.1 TCP 与吞吐量……155
6.4.2 TCP 的高性能扩展……159
6.5 本章小结……161

第III部分 TCP/IP 相关协议

第 7 章 上层协议……162
7.1 上层协议简介……162
7.1.1 分析上层协议……163
7.1.2 本章目标……165
7.2 域名系统……166
7.2.1 DNS 数据库……168
7.2.2 DNS 消息格式……169
7.2.3 使用 NSLookup……171
7.2.4 名字服务器……172
7.2.5 资源记录……176
7.2.6 分析 DNS……180
7.3 文件传输协议……184
7.3.1 FTP 命令和响应……184
7.3.2 示例研究：主动传输故障……187
7.3.3 示例研究：被动传输故障……189
7.3.4 示例研究：通过防火墙的 FTP 故障……189
7.3.5 示例研究：回顾 FTP 传输故障……192
7.4 超文本传输协议……193
7.4.1 HTTP 请求……193
7.4.2 HTTP 响应……195
7.4.3 HTTP 报文头和消息……198
7.4.4 高速缓存控制报文头……201
7.4.5 HTTP 代理……202
7.4.6 测量代理延迟……203
7.4.7 分析高级 Web 体系结构……204
7.4.8 示例研究：Web 站点故障……205
7.5 简单邮件传输协议……206
7.6 本章小结……208

第 8 章　与 Microsoft 公司相关的协议 …… 209
8.1　动态主机配置协议 …… 209
8.1.1　DHCP 报文头 …… 209
8.1.2　DHCP 处理 …… 211
8.1.3　DHCP 消息 …… 214
8.1.4　DHCP 选项 …… 215
8.1.5　DHCP 租用 …… 217
8.2　TCP/IP 上的 NetBIOS …… 218
8.2.1　NetBIOS 名字 …… 219
8.2.2　NetBIOS 服务 …… 221
8.2.3　NetBIOS 操作 …… 225
8.3　服务器消息块(SMB) …… 232
8.3.1　SMB 报文头 …… 233
8.3.2　SMB 命令 …… 235
8.3.3　SMB 响应 …… 237
8.3.4　SMB 操作分析 …… 239
8.3.5　进程间通信 …… 252
8.3.6　Microsoft 公司应用程序 …… 257
8.4　本章小结 …… 264

附录 A　Web 站点上的内容 …… 265
A.1　系统要求 …… 265
A.2　Web 站点上的内容 …… 265
A.2.1　标准和 RFC …… 265
A.2.2　作者提供的资料 …… 266
A.2.3　应用程序 …… 266
A.3　使用 Flash 视频示例 …… 266
A.4　故障诊断 …… 266

附录 B　SMB 状态代码 …… 267

第Ⅰ部分　网络分析基础

第1章　协议分析导论

什么是协议分析？协议就是用来调整计算机之间数据传输的标准过程。协议分析是检验这些标准过程的进程。我们利用协议分析仪(protocol analyzer)这一专门工具来进行分析。协议分析仪对网络中传输的比特流进行译码，然后以协议的结构化格式来显示这些比特。使用协议分析技术来了解在网络中出现的过程是本书的重点。根据十多年来网络分析与网络实现的经验，作者认识到要想了解供应商的硬件平台是如何工作的，如路由器或交换机，就必须了解使硬件实现操作功能的协议的工作方式。如果没有协议，路由器、交换机、Hub(集线器)和网关等设备将毫无用处，有了协议才有网络。路由器和其他设备实现协议。了解了协议就可以更好地了解网络内部的运行机制。

1.1　网络通信简史

多年以来，对复杂处理的需求成了计算机系统发展的动力。早期，超级计算机的发展满足了这一需求。超级计算机用来高速地为单一的应用程序服务，因此与手工计算相比，节省了宝贵的时间。

超级计算机主要为单个应用程序服务，因此不能完全满足计算系统能支持多用户的业务需求。那些设计给多人使用的应用程序要求具有多路输入/输出系统，而超级计算机不具备这种功能。由于每个用户从整个处理系统中分配一个很小的时间片，因此这种系统称为分配时间系统。最早出现的这种系统称为大型机。虽然不如超级计算机速度快，但是大型机可以满足多个用户同时运行多个应用程序的业务需求。这个特点使大型机可以更有效地满足多业务需求。

大型机的出现促使了集中式计算的诞生。集中式计算可以在被严密控制的结合性系统中提供全方位的网络通信系统。这种系统(如 IBM 公司的 S/390)在大型的中央处理系统中提供了通信路径、应用程序和存储系统。客户工作站仅仅是一个让用户和在中央处理单元中运行的应用程序进行交互的文本屏幕。

集中式计算之后紧接着出现了分布式计算。分布式计算的特征是业务过程分布在各个单独的计算机系统上。在 20 世纪 80 年代后期和 90 年代早期，集中式计算体系结构中使用的哑终端屏幕开始被计算机工作站所取代。这种工作站具有自己的处理能力和存储器，更重要的是它具有了可以离开大型机独立运行应用程序的能力。早期的分布式系统仅仅是对调制解调器或专用

租用线路的单一供应商解决方案(购自单个厂商)的扩充。因为供应商控制了系统的各个方面，所以对供应商来说开发将集中式系统发展成分布式系统所必须的通信功能是很容易的。这种系统称为“封闭的”系统，因为它们只能和同一供应商的其他系统进行互操作。Apple 公司和 Novell 公司是最早开发出分布式网络系统(虽然还是专用的)的公司。

分布式处理很复杂，它需要在系统之间寻址、进行错误控制和同步协调。不幸的是，各供应商设计的满足这些要求的通信体系结构不能相互兼容。大部分的封闭式专用系统得到了发展，其中最著名的是 IBM 公司的 SNA(System Network Architecture，系统网络结构)和数字设备公司的 DECNet。在这种情况下，其他公司，如 Novell 公司和 Apple 公司，也都跟随了这一潮流。为了能改变这种“封闭式系统”的局面，就需要一个允许在不同供应商的系统之间进行互操作的架构。

1.2 OSI 解决方案

由 ISO (International Organization for Standardization，国际标准化组织)制定的开放系统互连(Open System Interconnection，OSI)是一种用来促进不同供应商间系统互操作的解决方案。OSI 定义了一套支持分布式处理的通信体系结构。OSI 模型描述了允许不同系统通过网络成功通信的功能，它使用了分层的方法，通信功能被分解到 7 个不同的层中。从 OSI 模型的底层开始，这 7 层分别是：

- 第 1 层：物理层
- 第 2 层：数据链路层
- 第 3 层：网络层
- 第 4 层：传输层
- 第 5 层：会话层
- 第 6 层：表示层
- 第 7 层：应用层

每一层都为它的上层提供服务，同时又依赖于它的下层所提供的服务。该模型同时提供了层的抽象，即上层不必知道下层的操作细节，而是只要具有使用下层提供的服务的能力就可以了。之所以建立这个模型，是为了在理想状态下任何网络层协议的传输操作都可以忽略其所处的物理介质。网络层协议有 IP(Internet Protocol，网际协议)、IPX(Internet Packet Exchange，网际报文交换)或者 X.25。这一概念适用于所有的层，在后面的章节中，可以看到一些应用程序协议如何在不同的网络协议上同样运行(有时甚至是在不同的供应商间——Microsoft、IBM 以及 Banyan 公司的服务器操作系统使用的服务器消息块(Server Message Block，SMB)就是最好的例子)。大部分的通信协议可以很好地映射到 OSI 模型。

注意：

实际上，OSI 不仅包含了模型而且还包含一套复杂的协议。虽然这些协议现在很少使用，但是它们的最初目的是提供一套协议让所有的供应商都在它们的系统中采用，允许在不同供应商的系统之间进行互操作。尽管这个模型保留了下来，但不幸的是，这套协议没有保留下来。

1.2.1　定义层

因为几乎所有的协议都是基于OSI模型的，所以彻底了解这个模型如何工作是非常重要的。而为了了解协议，必须首先了解该模型的框架。下面将具体地解释这 7 层，在图 1-1 中给出了每一层中的一些协议示例。

OSI 模型	协议示例
应用层	SMB，HTTP，FTP，SMTP，NCP，TELNET
表示层	JPG，GIF，MPEG，ASN.1，SMB Negotiation
会话层	NetBIOS，TCP 3-way handshake
传输层	TCP，SPX
网络层	IP，IPX，DDP
数据链路层	Ethernet，Token Ring，FDDI，Frame Relay，HDLC
物理层	X.21，RS-232，DS1，DS3

图 1-1　OSI 模型

1. 第 1 层：物理层

简单地说，物理层负责将二进制数据转换为信号，并且在通信介质中传输。物理层还负责比特传输的功能和过程。例如在 B8ZS T1 线路中的零置换功能和 RS-232 握手过程。物理层本身只关心通过网络在两个设备间发送比特流。

2. 第 2 层：数据链路层

第 2 层是数据链路层，负责处理在设备间协调帧所必须的功能和过程。在数据链路层中，“0”和“1”按照逻辑顺序被划分成一些定义好了开始和结尾的帧。与物理层不同的是，数据链路层具有一定的智能。例如，该层中一个最普通的协议—— 以太网(Ethernet)协议就有一个检测算法用来控制冲突检测、损坏的帧，以及进行地址识别。上面的层不仅需要数据链路层提供无差错线路，还需要它可以检测可能发生的差错。被损坏的数据不能传给上面的层。

3. 第 3 层：网络层

第 3 层提供的是一种端到端的通信。数据链路层的职责在第 2 层设备处就结束了，网络层负责在第 2 层多路线路上将数据从源端传到目的端。应用程序使用第 3 层的协议而不必知道第 2 层网络以下的细节。第 3 层网络中使用的网际协议与第 2 层中使用的以太网、令牌环、帧中继以及异步传输模式(Asynchronous Transfer Mode，ATM)等技术有很大的差别。第 3 层的协议包括 IP、IPX 以及 AppleTalk 的 DDP(Datagram Delivery Protocol，数据报传送协议)。虽然网络层负责将数据从源端传送到目的端，但是它不保证传送过程。

4. 第 4 层：传输层

网络是不可靠的。在以太网中，冲突导致的数据丢失、拥塞会造成交换机丢失报文，甚至由于超负荷链路，网络本身也会造成数据丢失(互联网每天都会有这种异常发生)。传输层中的协议要重新传输丢失的数据、在端系统间监视流控制并经常要给应用程序数据额外添加一个防止出错的层。网络层在两个终端间传输数据，而传输层则保证数据可以到达目的地。

5. 第5层：会话层

会话层是处在传输协议和应用程序之间的层，它对端系统之间的通信进行进一步控制。如果有一个应用层的协议具有此功能，就不需要会话层协议了。本章后面看到的NetBIOS就是会话层协议的一个典型例子。有时，会话层本身并不表现为一个协议，而更像是一个让协议继续发挥其作用的执行过程。即使有时在某一层有一个协议，但该协议的过程却可以在另一层中发挥作用。在以后的章节中出现这种特例的地方会进行说明。

6. 第6层：表示层

表示层是另外一个有时表现不明显的层。表示层的功能是确保应用层协议所使用的数据格式在端系统间是一致的。该层中的协议有ASCⅡ、JPG和ASN.1等。就像在会话层中一样，表示层也有一些在其他层中起作用的协议。

7. 第7层：应用层

很多人经常把第7层和那些在服务器或工作站上使用的应用程序混淆。应用层协议并不是用户应用程序，而是允许那些应用程序在网络上运行的一些协议。用户使用Internet Explorer浏览因特网时就利用一个名为HTTP的应用层协议。Microsoft Word的用户将文件存储到网络服务器上就使用了SMB (Server Message Block，服务器消息块)协议。对用户来说，一个网络驱动器仅仅表现为“G：\”，但在这后面要有很多功能强大的应用层协议才能使“G：\”表示为远程服务器上的一个存储单元。应用层上的其他协议有FTP和Telnet。

1.2.2　各层的协议分析

本小节介绍OSI模型的协议分析方法，解释了每一层所做的工作，更重要的是解释了为什么要这样做。至于每一层如何实现其功能就留给协议设计人员吧。我们将在第3章到第6章将讨论TCP/IP是如何执行其功能的。很多高级用户可能会注意到在下面的内容中有一些含糊的或很普通的报文描述。这是为了对层的功能性的描述提供一个一般性的蓝图，具体的细节将在本书的随后部分提到。

1. 第1层：物理层

就像在本章前面提到的那样，物理层本身只关心通信信号如何在介质中传输。介质可以精确地定义为用来传送通信信号的路径。只要可以在其中成功地传输信号，路径可以是铜、水、空气甚至是带刺的铁丝中的任何一种。介质传送通信信号。在无线网络中，信号以无线电频率(Radio Frequency，RF)电波形式在空气中传播。在10BaseT以太网中，信号以电压形式传送。在光纤分布式数据接口(Fiber Distributed Data Interface，FDDI)网络中使用玻璃作为介质，信号以光脉冲形式在玻璃光纤电缆中传输。在不同的技术中使用特定类型的介质是有很多原因的。理论上可以使用任何一种介质来传输信号，但不幸的是，信号的表示形式限制了所能使用的介质类型。

模拟信号方式

通信信号有两种传输方式。第一种是模拟方式，用来传输数值随时间变化而变化的信号。

声音就是模拟信号最典型的例子。声音作为模拟信号，在一个周期中是以每秒或每赫兹来度量的。人声音的变化范围是 100Hz～1500Hz。在电话网络发展的早期，使用模拟信号进行长距离通信很难有较好的传输质量，因为模拟信号被放大后无法将噪音和声音信号区分开来。在模拟声音信号被放大的同时，噪音也被放大了。将模拟的声音信号转换成数字信号是解决这个问题的一种方法。

数字信号方式

与模拟信号不同，数字信号只有离散值：1 或者 0。在早期，数字电话工程师找到了一种方法可以将模拟信号调制到数字载波上，就是脉冲编码调制(Pulse Code Modulation，PCM)。PCM 可以用一个二进制数字来表示模拟信号的即时频率。现在只需要重复 0 或者 1 即可，而不要去判断放大器应该放大哪个信号。使用这种方法极大地提高了长距离通信的质量。当需要通过网络链接来传输计算机数据时，使用数字信号比较简单。因为计算机已经使用了 0 和 1 来表示数据，所以 0 和 1 可以很容易地在数字化网络中传输。

“0”和“1”表示了数字信号方式的全部。在 10BaseT 以太网中，数据用电压表示：“1”用 - 2.05V～0V 的电压范围表示，“0”用 0V～ - 2.05V 的电压范围表示。在光纤网络中，“1”用光脉冲来表示，而“0”则用无光脉冲来表示。处理过程并不像说的那么简单，但概念是基本相同的。不同的数字信号处理方法都可以在介质中产生“1”和“0”。现在由于只有两种信号需要区别，放大器就可以更加容易地将数字的“1”和“0”从背景噪音中提取出来。具有了将信号从噪音中分离的能力后，就可以更容易地在远距离上建立一个网络来传送用计算机处理的二进制数据。

注意：

数字信号方式有很多种。数字信号方式的效率和它的比特表示方法是决定在具体技术中使用何种类型数字信号方式的一个重要因素。例如，在 10-Mb 以太网中使用的是 Manchester 编码方式(数字信号方式的一种类型)，但是对于 100-Mb 快速以太网，由于在当时(20 世纪 80 年代后期)可利用的电缆无法支持这么高的带宽，因此 Manchester 编码方式因效率太低而不能使用。取而代之的是，快速以太网使用了 NRZI(Non Return to Zero Inverted)编码，有时还采用 MLT-3 编码(Multi-Level Three)。其他的数据链路技术使用不同的数字信号方式。令牌环使用的有差异的 Manchester 编码方式，而 T1 线路使用 AMI 或 B8ZS 编码方式。

2. 第 2 层：数据链路层

“1”和“0”是如何变成在网络中传输的 IP 报文的呢？对网络接口卡(Network Interface Card，NIC)来说，要把比特输入线路中，首先要有访问介质的方法。这个方法称为介质访问方法。所有用来在共享网络中使用的数据链路协议都具有一种介质访问方法。介质访问方法的功能之一就是让目的站识别哪一个比特是 MAC(Media Access Control，介质访问控制)帧的第一个比特。一旦帧的第一个比特找到了，NIC 就可以开始将“1”和“0”分组装入到一个 DLC(Data Link Control，数据链路控制)帧中。正如有不同的数字信号方式一样，也有不同类型的 DLC 帧。在以太网中，IP 协议是通过 Ethernet II 帧来传送的。在令牌环中，IP 协议则通过 Token_Ring_SNAP 帧来传送。

注意：

由于本书的目的是介绍如何更好地分析 TCP/IP 网络，因此不会详细介绍很多帧类型。想对多种帧类型有更多的了解，请参见由 Ulysses Black 著的“Data Link Protocol”一书(Prentice Hall Professional 1993 年出版)。

然而，了解第 2 层中构成帧方式的基本细节十分重要。每一个 DLC 帧都由以下 5 个基本部分构成：

- 介质访问部分
- 寻址
- 服务访问点
- 上层数据
- 帧保护

这 5 个基本部分如图 1-2 所示，我们将在随后的部分详细加以介绍。

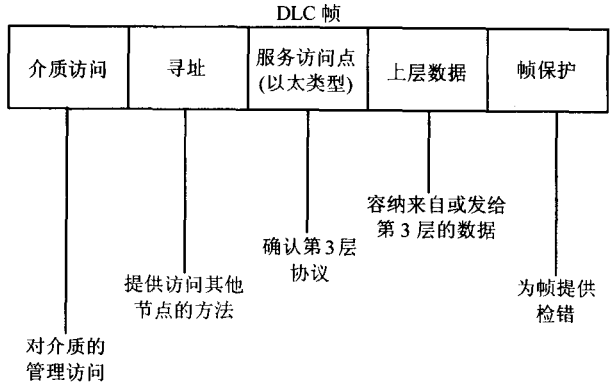

图 1-2　数据链路控制帧

介质访问部分

帧的介质访问部分由特定的比特模式和 NIC 驱动软件使用的保留比特组成。介质访问，顾名思义就是 NIC 必须访问介质。NIC 不可能总是按照自己的意愿发送信号，有时介质会被网络的其他节点使用，这就产生了共享网络(Shared Network)。在共享网络中，任一时刻只有一个节点可以通过线路向外传输数据。一个共享网络在物理结构上可能是由许多线路和 Hub 构成，但在逻辑结构上它只相当于一条线路。在某一时刻只有一个站点可以通过那条线路传输数据。考虑以下的例子：

- 以太网使用一个名为 CMSA/CD 的冲突回退算法。使用这个算法，站点会侦听传输介质是否为空闲，如果空闲则传输数据。如果它发现另外一个站点正在传输数据(即所谓的冲突)，则所有的站点都回退，等待一个随机时间再开始侦听传输介质，直到有一个站点可以成功地访问介质。
- 令牌环和 FDDI 使用了名为令牌模式的访问方案，这个方案是凭借一个小的令牌帧在逻辑回路中流通。当令牌到达相应的站点时，站点将令牌标记为“忙”并把数据附在令牌

上，以便在环路中传输数据。

这两种方法都有各自的优点和缺点，但概念在本质上是相同的。每个数据链路层协议都必须提供访问介质的方法。

MAC 寻址

通信是在网络上的节点间进行的，每个节点必须有一个惟一的标识符。这个标识符称为数据链路控制地址，即 DLC 地址，它也称为 MAC 地址(MAC 是 Media Access Control 的缩写)。本书会交替使用这两个术语。MAC 地址是由数据链路控制终端提供的，如 NIC。(这些地址也称作预烧硬件地址(burned-in-address)，这是因为它们是永久性地写入了只读存储器中(Read-Only Memory，ROM)。制作 ROM 芯片的过程实际包括了将一些很小的保险丝烧制到芯片中，以用来表示“1”或“0”，因此取名为预烧硬件地址。MAC 地址是一个 6 字节的十六进制数，可以惟一地标识一个节点的接口。要记住：MAC 地址标识的并不是节点，而只是节点上的一个接口。节点可以是工作站、服务器、路由器、网桥甚至是无线网络上的访问点，任何一个这种节点在网络上都有多重的 NIC 卡(即终端)。例如，一个路由器就有多个接口。另一方面，一个服务器可能有两个链路，一个连向产生式 LAN，另一个连向后备 LAN。

MAC 地址有 3 种类型，表 1-1 列出了这 3 种类型。

表 1-1　MAC 地址的 3 种类型

类　型	举　例
单播	00-00-0C-45-A9-D5
多播	01-23-7D-34-1E-9A
广播	FF-FF-FF-FF-FF-FF

- 单播：由单个终端处理。
- 多播：由多个终端处理。
- 广播：由所有终端处理。

第 1 种类型，单播地址由 6 个字节(十六进制的)构成完整的地址。第 2 种类型，多播地址也是由 6 个字节构成。第 3 种类型，广播地址同样是 6 个字节，但每个字节的数值都是 FF。

当半双工 NIC 卡不传输数据时，就在线路上侦听包含有自己地址的 MAC 帧。例如，在以太网上，一个节点能以自己的位模式收到另一个站点的传输和同步信息。当它识别出帧的第 1 位后，就会查看前 48 位，以决定是否将帧从线路中复制下来并发送给上面的层。为什么是 48 位呢？因为一个字节有 8 位，所以 48 位相当于 6 个字节，这就是 MAC 地址的长度。

注意：

以太网本身就是一个半双工协议。在以太网刚建立的时候，只有共享式 Hub 而没有交换机。当把一个以太网卡直接连接到转换端口上时，该段上仅有两个站点：计算机上的以太网卡和转换端口。通过取消冲突检测并允许所有的 NIC 及交换机任意地传输数据，连接就变了全双工。全双工在一个端到端的以太网段上的两端真正地禁用冲突检测。

网络上的节点必须能够将数据帧传输给单个站点、所选择的多个站点或所有站点。传输给