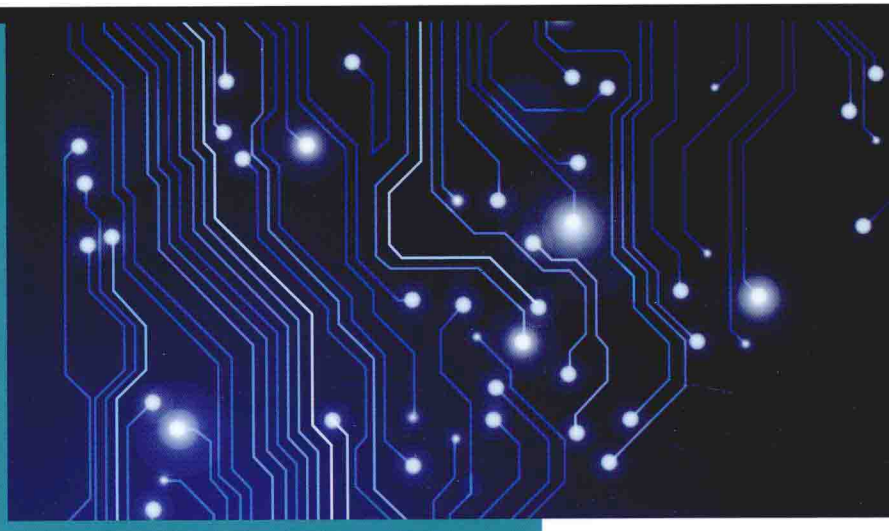
思科系列丛书



网络故障分析——路由篇

（上册）

李涤非◎编著

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

思科系列丛书

网络故障分析——路由篇

（上册）

李涤非　编著

電子工業出版社
Publishing House of Electronics Industry
北京 • BEIJING

内 容 简 介

本书阐述了计算机网络中与路由相关的故障诊断方法，并通过案例讲解了如何应用理论知识分析网络故障产生的原因，重点在于分析过程，旨在为读者提供一种易于理解和掌握的网络故障分析方法，以达到有效排除网络故障的目的。

全书共 6 章，分上、下两册出版。上册内容包括：网络基础和故障排除方法、直连路由和静态路由的故障分析、RIP 协议的故障分析。下册内容包括：EIGRP 协议的故障分析、链路状态路由协议（OSPF）的故障分析、与路由协议相关的安全技术。本书为上册。

本书适合网络工程师、管理员和自学网络技术的读者阅读，既可作为思科网络技术学院的教辅用书，也可作为相关院校师生的教学参考读物。

图书在版编目（CIP）数据

网络故障分析. 路由篇. 上册 / 李涤非编著. —北京：电子工业出版社，2015.1
（思科系列丛书）

ISBN 978-7-121-24713-2

Ⅰ. ①网… Ⅱ. ①李… Ⅲ. ①计算机网络—故障诊断②互联网络—路由器—故障诊断 Ⅳ. ①TP393.07 ②TN915.05

中国版本图书馆 CIP 数据核字（2014）第 260376 号

策划编辑：宋 梅
责任编辑：宋 梅
印 刷：北京中新伟业印刷有限公司
装 订：北京中新伟业印刷有限公司
出版发行：电子工业出版社
北京市海淀区万寿路 173 信箱 邮编 100036
开 本：787×980 1/16 印张：16 字数：358 千字
版 次：2015 年 1 月第 1 版
印 次：2015 年 1 月第 1 次印刷
印 数：3 000 册 定价：49.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

前　言

互联网是一个全球性的网络，它的主体是终端设备（包括服务器、个人计算机以及目前各种手持 PC 设备），它们直接向用户提供各种资源。而互联网的另一个组成部分是各种网络设备，路由器又是其中的重要部分。路由器使用路由协议在网络设备之间传递网络可达信息，这些信息保证了用户终端之间能够通过互联网相互通信，因此路由器在互联网中承担了极其重要的角色。随着互联网规模的增长，网络技术人员维护网络的任务更加繁重，各种疑难问题也大大增加。由于路由设备在互联网中处于核心地位，因此，如果能够快速、有效地解决路由故障问题，对于提高网络的可用性和可靠性都具有非常重要的意义。

本书阐述了计算机网络中与路由相关的故障分析方法，这种方法区别于其他同类书籍所采用的程序框图式的叙述方式，而是根据数据分组在网络中传输的路径、所经历的处理过程来确定故障发生的位置，然后分析可能发生的故障原因。本书通过设计好的案例讲解如何应用理论知识分析和解决网络故障问题，重点在于讲解分析过程，旨在为读者提供一种易于理解和掌握的网络故障分析方法。

全书共 6 章，分上、下两册出版。本书为上册，以下是各章的简要内容。

第 1 章，介绍 OSI 参考模型、TCP/IP 协议栈，以及其中主要协议的工作机制、IP 地址分类和子网划分方法。这些内容主要为后面章节提供预备知识，同时也作为后续章节的理论参考。在本章的最后还介绍了业界通用的网络故障排除模型，作为后面章节的理论指导。

第 2 章，首先介绍路由表的工作机制及路由表的构建过程，并以此为理论基础讲解直连路由如何生成以及静态路由的不同配置方法。最后以路由器构建路由表项的原理和过程为线索，重点阐述直连路由和静态路由的故障分析方法。

第 3 章，介绍距离矢量路由协议 RIP，首先介绍 RIP 协议的工作原理并重点阐述路由环路产生的原因和防止环路的各项措施。因为 RIP 是本书中第一个介绍的路由协议，所以在这一章中还详细介绍了“度量”和“管理距离”等重要的基本概念，它们是正确理解路由协议如何选路的基础。本章详细讲述了 RIPv1 和 RIPv2 的主要区别以及它们的基本配置方法，重点介绍了利用 RIP 自动传播默认路由的方法、产生自动汇总的条件以及 RIPv2 的手工汇总配置方法，详细介绍了负载均衡的两种工作方式以及它们与路由器转发方式之间的关系。在本章的最后，按照 RIP 协议的工作过程，分别阐述了路由发布、路由安装方面的故障分析和故障排除方法。

笔者在编写过程中注重对基本理论和协议原理的讲解，以循序渐进的方式介绍网络中遇到的各种问题和解决方法，启发读者对这些问题进行深入思考，希望读者能够在真正理解的基础上掌握所学的内容。这种写作方式不同于目前图书市场上的其他同类书籍。笔者多年在思科网络技术学院担任培训教师，以笔者对目前学校和一些培训机构授课方式的了解，近年来，对网络知识的讲授方法逐渐演变为强调操作的重要性，在授课过程中以学生能否在设备上把预期的

实验结果做出来为依据来判断学生是否掌握了学习内容。长此以往，导致学生过多地注重对各种命令的使用而不知不觉地忽略了对协议原理和实现机制的理解和掌握，以为熟练操作就是掌握所学知识的标志。这个问题反映到网络故障排除上就表现为遇到问题后只注意对各项配置命令的检查，或检查命令是否输入有误、或检查是否漏掉配置命令。如果确认配置命令无误，则面对故障便不知所措。他们通常将“故障排除”称为“排错”，意思是说排除故障的过程就是查找错误。殊不知在实际网络环境中，大部分网络故障都不是因为输入了错误的配置命令引起的。例如，目前互联网中不同网络设备接口的 MTU 尺寸不尽相同，同时网络中大量存在的过滤设备很可能造成 MTU 探测过程的失败，而很多主机的应用程序在发送 IP 分组时不允许在中途进行分片，从而造成数据分组传输失败。类似这种故障并非由配置错误造成，仅仅熟练掌握配置命令而对网络原理、数据分组经过网络设备时的处理过程等协议的实现细节不了解，不会运用相关网络协议的理论知识去分析遇到的问题，是无法排除这类网络故障的。

笔者在本书中尝试从介绍网络原理入手，希望读者能够在掌握基本理论和网络协议的工作原理与实现方法的基础上，应用本书中介绍的故障排除方法，对遇到的网络故障问题进行分析、判断，最终找到故障原因并解决问题。这也是笔者在每一章的前面花大量篇幅详细介绍相关协议的实现机制、工作过程和配置方法的原因。希望读者在阅读本书时不要忽略这部分内容而直接跳到故障分析部分。因为前者是内在的基础，后者是外在的应用，忽略前者而重视后者便是舍本逐末。全书上册的第 1 章是其他章节的预备知识和理论基础，对于上面列举的 MTU 问题，在本书的第 1 章中有详尽的论述。如果读者在故障排除过程中遇到对网络原理、协议的运行机制等不清楚或不理解的地方，请返回到本书的理论部分仔细阅读。

本书面向初学者或刚入职的工程技术人员，因此不涉及一些复杂、难深的网络技术。同时在讲解故障排除方法时尽量避免采用单纯罗列所有可能的故障原因，然后直接列举解决方法的叙述方式。本书的重点在于介绍故障分析方法和过程，使读者最终获得运用已掌握的理论知识分析和解决未知问题的能力。

本书由李涤非编写并统稿，参加编写工作的还有董燕、李国鼐、邢学东、王炬和吴晓明。感谢思科公司刘亢经理对本书部分章节提出的意见和建议，更要感谢电子工业出版社的宋梅老师，没有宋梅老师的耐心帮助和鼓励，本书将无法完成。

由于作者水平有限，不足之处在所难免，请读者给予批评指正。

邮件地址：lidf2014@163.com

李涤非

2014 年 12 月于北京

目　录

第1章 >>>

网络基础和故障排除方法

本章要点

- OSI 参考模型介绍
- 了解 OSI 参考模型各层的功能
- TCP/IP 协议栈
- 地址解析协议
- 网络故障排除方法
- 本章小结

本章介绍有关理解计算机网络所必备的一些概念，OSI 参考模型的体系结构、分层及各层的功能与相互关系，重点阐述 TCP/IP 协议栈的主要协议内容、IP 地址的结构、分类和子网划分方法。这些内容是后续章节的理论基础。本书的核心部分虽然是网络故障排除，但了解网络原理尤其是网络协议的内部实现机制，是有效排除网络故障的必要前提。建议读者仔细阅读本章内容，理解所述网络协议的设计思想、工作方式和原理。本章最后介绍网络故障排除的通用方法，它是资深网络工程师们多年工作经验的结晶，在这里作为一个框架提供给读者，具体的分析、诊断方法将会在后面的章节中详述。

1.1 OSI 参考模型介绍

1.1.1 为什么需要 OSI 参考模型

在回答这个问题之前，先简单回顾一下 OSI 参考模型出现之前的计算机网络发展历程：1974 年，著名的 IBM 公司提出了世界上最早的计算机网络体系结构 SNA（System Network Architecture），它的主要目的是为了实现 IBM 本公司设备之间的互连。随后 DEC 公司（20 世纪 80 年代初 DEC 在计算机行业排名仅次于 IBM）于 1975 年提出了自己的网络体系结构 DNA（DIGITAL Network Architecture），图 1-1 是两种网络结构的示意图。由于相互之间缺乏沟通，这些不同厂商自己提出的网络体系结构之间存在差异，互不相容。因此，将不同厂商设备通过网络互连存在很大的困难，必须在不同厂商设备之间做一些翻译和转换的工作。这种专用的体系结构实际上体现了一种封闭性。

图 1-1 IBM 与 DEC 公司的计算机网络体系结构

随着计算机网络规模与数量的急剧增长，这种不同厂商设备之间的不兼容性，严重阻碍了计算机网络的健康发展。各厂商也意识到各自的体系结构之间缺乏兼容性所造成的问题，于是开始想办法解决这个难题以促进网络的进一步发展。这时，在制定国际标准方面具有权威性的国际组织——国际标准化组织（International Organization for Standardization，ISO）着手制定统一的标准：OSI 参考模型（Open System Interconnection Reference Model），如图 1-2 所示。该参考模型的研究和起草工作起始于 20 世纪 70 年代末，于 1984 年正式发布。这个标准的第一个词之所以称为“Open”（意为“开放”），是相对于上面所提到的厂商私有网络体系结构的封闭性而言的。这是一个对所有厂商和机构都开放的标准，只要遵守这个标准，就可以和其他任何同样遵守该标准的网络相互通信。

第7层	应用层	Application
第6层	表示层	Presentation
第5层	会话层	Session
第4层	传输层	Transport
第3层	网络层	Network
第2层	数据链接层	Data Link
第1层	物理层	Physical

图 1-2　OSI 参考模型

读者不难看出图 1-1 所示的两个私有体系结构与图 1-2 的 OSI 标准模型之间的相似性。实际上，在制定 OSI 参考模型的过程中，工作组的成员研究了当时已存在的一些解决方案，包括 IBM 公司的 SNA 和 ARPANET（Internet 的前身）等网络体系结构，在它们的基础上提出的 OSI 参考模型。这个标准的参考模型提出后，已有的网络体系结构都与之建立对应关系。例如，SNA 最初的模型只定义了 6 层，并没有定义物理层，这部分功能由其他标准实现，而图 1-1 中 SNA 之所以加入物理层是为了与 OSI 参考模型对应。

需要强调的是：OSI 参考模型只是一个理论框架，它定义了信息要通过网络传递所要完成的各项任务，但并不规定具体如何去实现。即它只定义了需要做什么，并没有规定如何做。虽然 OSI 参考模型的实际应用意义不是很大，但它对于理解计算机网络内部运作的机制很有帮助，也为我们学习网络知识提供了一个很好的参考。这也是几乎每一本计算机网络教科书都以 OSI 参考模型为主体框架描述相关网络协议与标准的原因。本书在一开始介绍 OSI 参考模型同样是为了便于在后续章节中清楚地描述和定位网络故障。

说明：协议（Protocol）是指网络中通信实体之间就交换信息等问题所做的某种约定或制定的相应规则。

1.1.2　OSI 的层次结构

图 1-2 清楚地展示了 OSI 参考模型的层次结构，那么为什么要分层呢？在回答这个问题之前，首先了解一下计算机之间互连的目的和意义。在人们应用计算机的早期，由于价格非常昂贵，计算机的数量很少，设备之间互连的主要目的是为了资源共享以节约成本，典型的例子是通过网络共享打印机、文件服务器等。资源共享是计算机网络产生的动因。随着互联网的飞速发展，目前，除资源共享之外，计算机网络已成为用户信息交流和协同工作的平台。

1. 在网络中传输数据信息需要完成的任务

无论是简单的资源共享还是复杂的协同工作，网络中最基本的操作就是将数据从源可靠地

传递到目的地，这看似简单的一句话，背后却隐藏着复杂的处理过程。我们以大家都熟悉的发送电子邮件的过程为例，了解一下在网络中传输数据信息需要完成哪些任务。

① 发送方在写好邮件后交给网络以便通过它传送到接收方。

② 与现实生活中邮寄信件相似，要有统一的信件格式才能保证双方相互理解，因此要预先定义好标准的格式。

③ 网络上的设备所能接收的信息长度是有限的，因此，需要对邮件的内容进行分割，并且规定好分割的最大长度。

④ 要送达到接收方必须正确地标明目的方地址，同时，为了在发生意外情况下能够收到退信，还要写明发送方的源地址。当然，这种地址要定义成网络设备（类似于邮递员）能够理解的格式。

⑤ 网络设备要能够找到一条可达的路径将邮件送达目的地。

⑥ 要确定数据信息以何种形式在媒介（有线或无线）上传送。

2. 在网络中传输数据信息可能出现的问题

以上这些任务如果都完成了，理论上，就能够将邮件送达到接收方。但谁也不能保证网络中不发生意外，因此，需要考虑一下可能出现的下列问题。

① 数据损坏：硬件故障或环境的影响可能造成数据内容丢失或被破坏。

② 数据之间过长的延迟：数据传递可能要经过一个冗长的路径，在这个过程中难免会出现延迟，对于延迟的容忍时间应该如何规定？

③ 数据丢失与重传：如果延迟时间超出了容忍范围就认为发生了数据丢失，那么发方就需要重传。这时，重传的数据与收方已接收到的数据之间很可能出现顺序颠倒的现象，因为重传的是早先已发送的数据，而到达的时间却较晚，在这种情况下如何排序？

对于上述故障，网络协议首先要能够检测，继而必须能够纠正。

通过上面对发送电子邮件所需完成的各项任务以及在网络传输中可能出现的故障的分析，不难看出在网络中将数据从源可靠地传递到目的地是一件非常复杂的工作。人们在处理复杂的事务时往往将其分解成多个子任务，对于整个任务来说每个子任务相对简单。完成所有的子任务后，整个任务也即告完成，这种分而治之的方法极大地简化了问题。OSI 参考模型分层的目的就是为了简化问题而将任务分解。如图 1-2 所示，它分为 7 层，将数据在网络中可靠传输这一复杂的任务分解成 7 个子任务。上述发送邮件的过程中所要做的全部工作都要在这 7 个子任务中完成，才能保证邮件可靠地送达。

需要强调的是，各层之间是相互关联的，下层（层号较小）为上层（层号较大）提供服务，类似于工业化生产中的流水线作业。主机发送数据时由上层逐层传递给下层，接收数据时则相反，从下层逐层传递到上层。分层的方法除了能够简化问题外，还有如下优点。

① 每一层对应一个功能模块，各模块之间相互独立，即相邻层之间定义出标准界面，而本层内部实现的功能对其他层是不可见的，当某一层需要修改其功能模块时不影响其他层。这

样做的好处是可以并行开发、维护不同层的功能模块，提高了工作效率。

② 在进行网络故障分析时，分层的方法还能够帮助我们分解、简化问题，定位故障点。

3. 常用计算机网络分类

在具体介绍 OSI 参考模型各层功能之前，先介绍一种常用的计算机网络分类方法，按照网络的规模和覆盖的地域范围分为如下几类。

① 局域网（Local Area Network，LAN）：一般在几千米的范围以内，通常在一座建筑物或一个园区（Campus）内。办公室的计算机网是最典型的局域网。

② 城域网（Metropolitan Area Network，MAN）：城域网的覆盖范围比局域网更广，通常覆盖一个城市，从几十千米到 100 千米不等。城域网是由一个城市范围内的局域网互连而成的。

③ 广域网（Wide Area Network，WAN）：广域网所覆盖的范围比城域网更广，地理范围可从几百千米到几千千米。跨国公司在不同国家和地区的办公局域网可互连起来构成一个规模更大的广域网。大家熟知的互联网就是典型的广域网。

1.2　了解 OSI 参考模型各层的功能

OSI 参考模型的 7 个层次按照由上至下的顺序分别为应用层、表示层、会话层、传输层、网络层、数据链路层和物理层，如图 1-2 所示。上面的应用层最靠近用户，下面的物理层最接近网络介质。从本书的侧重点出发，将这 7 层进一步划分为两个部分：上部层次和下部层次。上部层次为用户的应用程序提供网络服务，而下部层次主要负责数据在网络中的传输工作。由于本书的焦点在于网络故障分析，因此，将重点关注下部各层的功能。

注意：这里所说的上部层次和下部层次是根据本书的需要划分的，并不是标准的分类方式。

1.2.1　OSI 上部层次

OSI 参考模型的上部层次（第 5、6、7 层）靠近用户的应用程序，应用程序通过它们与网络进行交互。下面分别简要介绍各层功能（按照发送数据的顺序）。

1. 第 7 层——应用层

应用层对用户的应用软件提供接口以便它们能够使用网络服务。在上面提到的发送邮件的例子中，写好邮件后交给网络传输的工作就是首先由应用层接手的。在这个例子中，应用层为用户的邮件应用程序提供服务，准备将邮件通过网络传递到目的地。应用层中除了提供大家熟悉的邮件服务外，还提供文件传输（FTP），远程登录（Telnet）和万维网（WWW/HTTP）等服务。

2. 第 6 层——表示层

在发送邮件的例子中，为了通信双方相互能够理解而预先定义标准格式的工作就是表示层要完成的主要任务，将不同类型的表达格式转换为标准格式的工作也称为“翻译”。除了定义标准的信息格式以外，表示层还包括数据的加密与解密、压缩与恢复等任务。

> 说明：表示层中的加密与压缩任务是可选的，并不一定对所有的用户都需要。此外，数据的加密与解密处理原则上也可以在其他层中实现，不一定必须在表示层中完成。例如，网络通信中常用的 IPSec 协议就是在网络层实现加 / 解密操作的。

3. 第 5 层——会话层

应用层接收到的用户信息，经过表示层转换成标准格式，交给会话层以对话的方式完成双方的信息交换。会话层的任务是在两个节点间建立和维护会话连接。例如，服务器验证用户登录的过程就是会话层的一个典型实例。

1.2.2 OSI 下部层次

OSI 参考模型的下部层次（第 1、2、3、4 层）负责数据传输工作，网络工程师所重点关注的就是这些部分。下面分别介绍各层功能（按照发送数据的顺序）。

1. 第 4 层——传输层

传输层负责端到端节点之间的数据传输和控制功能。首先解释一下什么是叫端到端：仍然以发送邮件为例，假设某个在北京的用户 A 给在广州的用户 B 发送一封邮件，北京到广州相距上千千米，邮件在传输过程中还需要经过很多其他的网络设备，在这条路径上的所有节点设备都要对邮件信息做某种处理，而传输层定义的任务仅在发送端 A 和接收端 B 上完成，中间的设备不需要执行传输层任务，由于 A 和 B 是整个路径上的两个端点，因此我们强调传输层负责端到端之间的数据传输和控制任务。

在发送端，为了方便网络传输，传输层将上面三层传来的信息分割成段，然后传递给下面的网络层；在接收端，传输层接收到这些分段信息后，将其重新组合起来传递给上面的会话层。用户会有多种不同的应用程序，因此传输层还要能够区分不同的上层应用。为信息分段以及区分不同的上层应用是传输层最基本的任务。

传输层协议分为两种类型：可靠的和不可靠的。对于可靠协议（如 TCP/IP 协议栈中的 TCP 协议）除了完成上述基本任务外，还要保证信息能够可靠传输。它首先要检查信息是否完好到达，如果到达还可能需要为它们排序、重组；如果没有到达或出现差错就通知发送方重新发送。同时可靠的传输层协议还具有流量控制的机制。对于不可靠协议（如 TCP/IP 协议栈中的 UDP 协议）除了可选的差错检查功能外，不对数据做可靠性检查也不提供流量控制的机制。因此，

可靠的协议实现起来相对复杂一些，而不可靠协议实现比较简单。

注意：这里所说的可靠性并不能保证信息在传递过程中不发生问题，而是通过在接收端检查信息是否完好到达，如果有问题再通知发送方重传来实现的。

既然传输层协议分为可靠的和不可靠的两种，那么上层协议如何在两者之间选择呢？如果选择了不可靠的传输层协议，一旦数据损坏或丢失将如何补救？这些问题将在 1.3.2 节中介绍 TCP 和 UDP 协议时再详细讨论。

在了解了传输层定义的任务之后，简要说明一下 1.1.2 节那个发送邮件的例子中，属于传输层的任务有：信息的分割，数据损坏、丢失后的重传以及重新排序等。从这些任务的内容看，邮件服务选用的是可靠的传输层协议。

综合我们讨论过的上述 4 层所定义的任务，读者应该能够体会到一项较复杂的任务是如何被逐层分解的。当然，仅分解到传输层整个信息传输的任务还没有完成，接下来讨论下三层分别完成哪些任务。

2．第 3 层——网络层

在讨论传输层任务时我们曾提到过，信息在传输过程中还需要经过很多网络设备，并且传输的距离越长可选的路径就越多，其中的一些设备要负责选择一条合适的路径将信息送达目的地。这些负责选路的设备叫作路由器，网络层的一项重要任务——选路，就是由它们完成的。更进一步地说，选路要有依据，根据地址才能选路，因此网络层还要定义地址格式。网络层定义的地址称为网络层地址（也叫逻辑地址），在 1.3.4 节中我们将具体介绍目前最常用的网络层地址——IP 地址。

除了选路和定义网络层地址以外，网络层的任务还包括不同类型网络之间的互连和拥塞控制。不同类型的网络之间协议各不相同，当它们互连时网络层负责协议间的转换工作。网络拥塞是指网络设备的某个接口收到过多的数据，超出了它的处理能力而发生的延迟增加甚至丢失数据的现象。这时网络层负责调度资源，让重要的数据优先通过，缓解拥塞所造成的负面影响。这就与公路上发生交通堵塞时，警察指挥疏导的情况相类似。

说明：对于网络拥塞的解决办法不仅限于网络层。例如，利用传输层 TCP 协议也可以实现拥塞避免。该技术内容超出本书范围，有兴趣的读者可参考相关资料。

3．第 2 层——数据链路层

网络层选好路径之后，下一步的任务是将数据朝目的方向送出，具体到网络设备就是将数据从设备的某个接口发送出去。数据在通往目的地传输过程中需要经过很多网络设备，我们所说的链路就是指这些网络设备之间的连接通路。总体来说，数据链路层的任务就是要保证数据在网络设备之间的链路上正确传递，相对于传输层负责管理源端与目的端之间的端到端通信，

数据链路层负责相邻网络设备之间的通信。可以将网络设备理解为中间经过的信息点，因此也叫作点到点通信。端到端通信是建立在点到点通信的基础上的，它由多个点到点通信信道组成，如图 1-3 所示。端到端通信是传输层的概念；而点到点通信是数据链路层的概念。读者应当熟悉这些术语，以便阅读技术文档时能准确理解其含义。

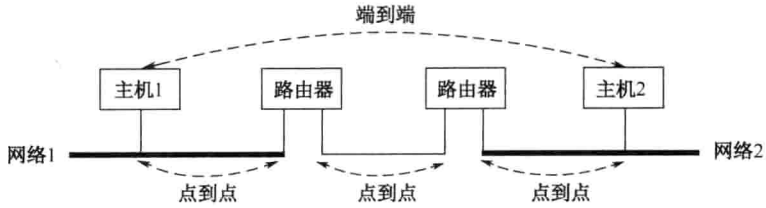


图 1-3　点到点通信与端到端通信

根据 IEEE 802 标准，数据链路层又分两个子层：媒介访问控制（Media Access Control，MAC）子层和逻辑链路控制（Media Access Control，LLC）子层，如图 1-4 所示。分层即任务的分解，数据链路层又分为子层即对该层任务的进一步分解。LLC 的主要任务是差错校验和流量控制。MAC 的主要任务是将数据组成帧、定义数据链路层地址（也叫作物理地址）以及控制对网络媒介的使用权，这两个子层的任务合起来就是数据链路层所要完成的所有任务。

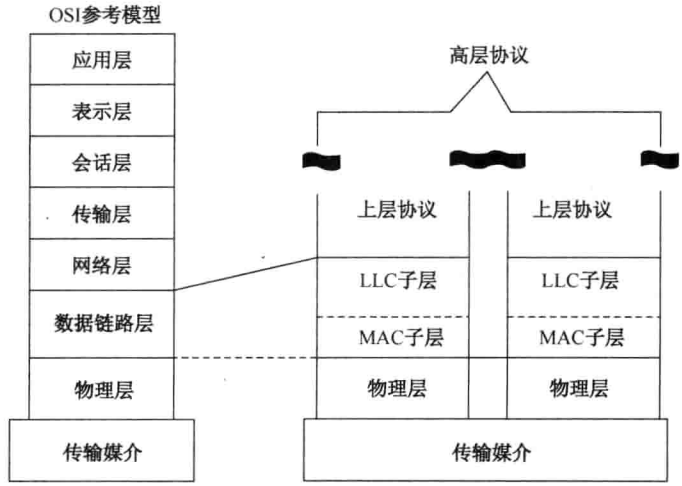


图 1-4　数据链路层子层

说明：IEEE 802 有一系列的标准，它们定义 OSI 最下面两层——物理层和数据链路层的功能。如前所述，OSI 只确定了哪一层需要做什么，并没有规定如何去实现。IEEE 802 标准规定了如何实现物理层和数据链路层的功能。

IEEE 802 标准定义 LLC 子层（对应 IEEE 802.2）的主要目的是为了对它的上层——网络

层屏蔽不同 MAC 子层的差异，如图 1-5 所示，IEEE 802.3 和 IEEE 802.5 分别对应以太网和令牌环网协议类型，它们与 LLC 子层有不同的接口。LLC 对它的上层即网络层的接口是统一的，这样网络层就不需要了解 LLC 下面到底是何种网络类型。遗憾的是 IEEE 这一良苦用心在现实网络环境中并没有得到认可，目前在网络中占统治地位的低层协议——以太网协议在封装用户数据时并没有采用划分子层的方式，而是使用传统的 DIX 2.0 以太网标准，虽然在这个标准中没有定义 LLC 子层，但它同样要完成 LLC 所规定的任务，因为它要涵盖整个数据链路层的功能。

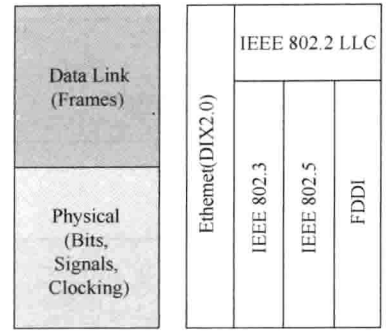


图 1-5　LLC 子层屏蔽下层差异

说明：网络中的以太网数据帧有两种格式：其中，用于传送用户数据的帧格式为 DIX 2.0 标准（它没有划分子层），例如，发送电子邮件、网页浏览等；而用于网络设备之间交换控制信息的帧格式大多为 IEEE 802.2/802.3（它划分了两个子层），例如，生成树协议即采用这种帧格式，如图 1-6 所示。

DIX 的由来：以太网是一个开放式的、多厂商参与的局域网标准。施乐（Xerox）公司发明了以太网以后，另外两家公司（Digital Equipment Corporation 和 Intel）也加入进来。1980 年，Digital Equipment Corporation、Intel 与 Xerox 三家公司宣布了 10 Mbps 以太网标准，该标准的名称由这三家公司的英文首写字母组合而成，即 DIX 以太网标准（DIX 标准的最高版本为 2.0）。随后以太网协议交给 IEEE 管理，成为了一个开放性的协议。开放性使得以太网成为目前使用最为广泛的局域网技术。

由于目前网络中绝大多数的以太网帧格式都使用 DIX 2.0 标准，所以我们以该标准为例介绍数据链路层定义的任务。

（1）组帧（封装）

当网络层请求发送数据时，数据链路层将网络层传来的数据组成一定的格式准备发送，这一过程称为组帧，以太网的帧格式如图 1-6 所示。下面介绍一下帧的组成。

① 前同步信号（Preamble）：左侧第一个字段为 8 字节的前同步信号，它的任务是保证网卡的接口在重要数据字段到来之前与之同步，这一过程可以简单理解为通知接口做好接收数据的准备工作，它本身的内容没有实质性的意义。

② 目的地址（Destination Address）：紧接着前同步信号的是 48 bit 目的地址字段，该字段对应帧的目的地接口地址。以太网是广播型网络，一个站点发送的数据帧会被网络中多个站点收到，接收站点读取帧中的目的地址与自己的接口地址进行比较，如果相同则继续接收并处理后面的内容；如果不同就可以忽略帧的其余内容（将其丢弃）。

7	1	6	6	2	46-1500	4
Preamble	Start of Frame Delimiter	Desination Address	Source Address	Type	Date	Frame Check Sequence

7	1	6	6	2	46-1500	4
Preamble	Start of Frame Delimiter	Desination Address	Source Address	Length	IEEE 802.2 Header and Data	Frame Check Sequence

图 1-6　DIX 2.0（上）和 IEEE 802.2/802.3 帧格式（下）

③ 源地址（Source Address）：目的地址后面的是 48 bit 的源地址字段，以太网站点在它传送的每个帧中用自己的接口地址作为源地址。接收站点可根据源地址对发送站点进行应答。

④ 类型（Type）与长度（Length）字段：该字段 DIX 与 IEEE 802.3 标准定义有所不同。

在 DIX 以太网标准中，源地址字段后面是 2 字节的类型字段（参见图 1-6 上部）。它包含一个标识符，用来说明帧的数据字段中携带的上层协议（网络层协议）的类型，例加，标识符 0x0800（0x 表示后面的数字为十六进制数）指示上层协议为 IP。类型字段的意义重大，如果没有它标识上层协议类型，以太网协议将无法支持多种网络层协议。

IEEE 802.3 标准在最初确立时这部分对应的是 2 字节的长度字段（参见图 1-6 下部），而上层协议类型由 IEEE 802.2（对应 LLC 子层）中定义。长度字段中的数值指明了在它后面的数据字段中数据的长度，以字节为单位。目前的 IEEE 802.3 标准中对此做了修改，这个字段被称为长度 / 类型字段。由于以太网帧中数据字段的最大长度定义为 1 500 字节（也称为以太网的最大传输单元 MTU），协议规定若该字段中的数值小于或等于 1 500，则作为长度字段使用（这时其后包含 IEEE 802.2 协议内容），若大于或等于 1 536 则作为类型字段使用（因为根据 IEEE 802.3 标准不可能有这么长的以太网帧）。这时，IEEE 802.3 帧的组成结构与 DIX 标准定义的完全相同。因此，当前的 IEEE 802.3 标准与 DIX 完全兼容。

⑤ 数据字段（Data field）：这个字段的内容与前面的类型与长度字段相关，当前面的字段代表类型时，数据字段的内容是网络层传递下来的数据，规范规定的长度范围是 46～1 500 字节；当前面的字段代表长度时（IEEE 802.3 最初的标准），数据字段包含 LLC 协议（IEEE 802.2）及网络层传递下来的数据这两部分内容，规范规定的长度范围也是 46～1 500 字节。

⑥ 帧检查序列字段（Frame Check Sequence field，FCS）：这是以太网帧的最后一个字段。字段的长度为 4 字节，其中的数值用来检查帧的完整性。这个值是通过循环冗余校验（Cyclic Redundancy Check，CRC）的算法计算出来的。发送站点首先计算出帧的 CRC 数值放入 FCS 字段，接收站点收到数据帧后重新计算该值并与 FCS 字段做比较，以判断帧是否有误。

在这里逐字段地解释每一部分的内容有两点考虑，一是使读者了解以太网帧的组成结构及各字段的功能，以便在故障分析过程中能够读懂并理解数据帧的内容；二是通过分析具体的以