

Java Network Programming

第二版
涵盖 Java 2, JDK 1.2 & 1.3



JavaTM 网络编程

O'REILLY[®]
中国电力出版社

Elliott Rusty Harold 著

刘东华 王巍 唐刚 译

Java™ 网络编程

第二版

Elliote Rusty Harold 著

刘东华 王巍 唐刚 译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

O'Reilly & Associates, Inc. 授权中国电力出版社出版

中国电力出版社

图书在版编目 (CIP) 数据

Java™ 网络编程 / (美) 哈罗德 (Harold, E.) 编著; 刘东华等译. - 北京: 中国电力出版社, 2001

书名原文: Java™ Network Programming, Second Edition

ISBN 7-5083-0712-7

I .J... II .①哈... ②刘... III .JAVA 语言 - 程序设计 IV .TP312

中国版本图书馆 CIP 数据核字 (2001) 第 051164 号

北京市版权局著作权合同登记

图字: 01-2001-2664 号

©2000 by O'Reilly & Associates, Inc.

Simplified Chinese Edition, jointly published by O'Reilly & Associates, Inc. and China Electric Power Press, 2001. Authorized translation of the English edition, 2000 O'Reilly & Associates, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly & Associates, Inc. 出版 2000。

简体中文版由中国电力出版社出版 2001。英文原版的翻译得到 O'Reilly & Associates, Inc. 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly & Associates, Inc. 的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式重制。

书 名 / Java™ 网络编程 (第二版)

书 号 / ISBN 7-5083-0712-7

责任编辑 / 刘江

封面设计 / Edie Freedman, 张健

出版发行 / 中国电力出版社 (www.infopower.com.cn)

地 址 / 北京三里河路 6 号 (邮政编码 100044)

经 销 / 全国新华书店

印 刷 / 北京市地矿印刷厂

开 本 / 787 毫米 × 1092 毫米 16 开本 46.5 印张 691 千字

版 次 / 2001 年 8 月第一版 2001 年 8 月第一次印刷

印 数 / 0001-5000 册

定 价 / 79.00 元 (册)

12-08P/02

O'Reilly & Associates 公司介绍

为了满足读者对网络 and 软件技术知识的迫切需求,世界著名计算机图书出版机构 O'Reilly & Associates 公司授权中国电力出版社,翻译出版一批该公司久负盛名的英文经典技术专著。

O'Reilly & Associates 公司是世界上在 UNIX、X、Internet 和其他开放系统图书领域具有领导地位的出版公司,同时是联机出版的先锋。

从最畅销的《The Whole Internet Use's Guide & Catalog》(被纽约公共图书馆评为二十世纪最重要的 50 本书之一)到 GNN(最早的 Internet 门户和商业网站),再到 WebSite(第一个桌面 PC 的 Web 服务器软件),O'Reilly & Associates 一直处于 Internet 发展的最前沿。

许多书店的反馈表明,O'Reilly & Associates 是最稳定的计算机图书出版商——每一本书都一版再版。与大多数计算机图书出版商相比,O'Reilly & Associates 公司具有深厚的计算机专业背景,这使得 O'Reilly & Associates 形成了一个非常不同于其他出版商的出版方针。O'Reilly & Associates 所有的编辑人员以前都是程序员,或者是顶尖级的技术专家。O'Reilly & Associates 还有许多固定的作者群体——他们本身是相关领域的技术专家、咨询专家,而现在编写著作,O'Reilly & Associates 依靠他们及时地推出图书。因为 O'Reilly & Associates 紧密地与计算机业界联系着,所以 O'Reilly & Associates 知道市场上真正需要什么图书。

目录

前言	1
第一章 为什么 Java 要网络化	17
网络程序能做些什么	18
有关 Java 的更多信息	36
第二章 网络的基本概念	37
网络	37
网络的层	39
IP、TCP 和 UDP	45
Internet	48
客户机 / 服务器模型	54
Internet 标准	56
第三章 Web 的基本概念	67
URI	68
HTML、SGML 和 XML	74
HTTP	77

i

MIME	82
CGI	88
applet 和安全性	91
第四章 Java I/O	96
输出流	97
输入流	101
过滤器流	105
Reader 和 Writer	122
第五章 线程	138
运行线程	140
从线程返回信息	144
同步	156
死锁	163
线程的时序安排	164
线程库	179
第六章 查找 Internet 地址	185
DNS、IP 地址以及相关内容	185
InetAddress 类	187
一些有用的程序	201
第七章 用 URL 检索数据	213
URL 类	213
URLEncoder 和 URLDecoder 类	238
通过 GET 实现 CGI 和 Servlet 的通信	243
访问密码保护的网址	248

第八章 Swing 中的 HTML	256
组件中的 HTML	257
JEditorPane	259
解析 HTML	268
第九章 java.applet.Applet 的网络方法	287
使用 java.applet.Applet 下载数据	287
ImageObserver 接口	296
MediaTracker 类	301
java.applet.AppletContext 的网络方法	318
第十章 客户端套接字	322
套接字基础	322
用 Telnet 研究协议	323
Socket 类	325
套接字异常	349
实例	350
第十一章 服务器套接字	369
ServerSocket 类	370
一些有用的服务器	384
第十二章 安全套接字	407
安全通信	408
创建安全客户机套接字	411
SSLSocket 类的方法	415
创建安全服务器套接字	422
SSLServerSocket 类的方法	426

第十三章 UDP 数据报和套接字	429
UDP 协议	429
DatagramPacket 类	431
DatagramSocket 类	441
一些有用的应用程序	453
第十四章 组播套接字	467
什么是组播套接字	468
组播套接字的使用	478
两个简单的例子	484
第十五章 URLConnection 类	490
打开 URLConnection	491
从服务器读数据	493
读首部	494
配置连接	503
配置客户请求 MIME 首部	513
向服务器写数据	515
内容处理器	521
对象方法	523
URLConnections 的安全考虑	523
推测 MIME 类型	524
HttpURLConnection	528
JarURLConnection	544
第十六章 协议处理器	548
什么是协议处理器	548
URLStreamHandler 类	553
写协议处理器	561

更多协议处理器的例子和技术	566
URLStreamHandlerFactory 接口	575
第十七章 内容处理器	579
什么是内容处理器	580
ContentHandler 类	584
ContentHandlerFactory 接口	596
图像格式 image/x-fits 的内容处理器	599
第十八章 RMI	611
什么是 RMI?	611
实现	617
运行时加载类	626
Java.rmi 包	633
java.rmi.registry 包	643
java.rmi.server 包	645
第十九章 JavaMail API	651
什么是 JavaMail API	652
发送 Email	655
接收 Email	666
口令验证	671
地址	675
URLName 类	680
Message 类	683
Part 接口	696
多 part 消息和附件文件	707
MIME 消息	712
文件夹	714

前言

在过去五年中，Java的发展一直非同寻常。它在网络中迅速崛起并引起了广泛的注意，但令人感到惊讶的是，对于许多人来说，Java 仍然非常神秘。这真是不应该。事实上，正如本书将要展示的那样，使用Java实现网络编程相当简单。以往有过在 Unix，Windows 或 Macintosh 环境下进行网络编程经验的读者，将会高兴和惊讶地看到，在Java中编写同样功能的程序是如何更加的容易、简单。Java之所以具有这样的特性，是因为它的内核 API 对于大部分的网络特性都包含了设计良好的接口。实际上，大部分应用层网络软件使用Java编写比用C或者C++要更容易一些。本书将尽力指导你如何利用Java的网络类库优势快速、简便地编写程序，以完成多种常用的网络功能。其内容主要包括：

- 在 Web 上浏览页面
- 解析和翻译 HTML
- 使用 SMTP 协议发送邮件
- 使用 POP 和 IMAP 协议接收邮件
- 编写多线程服务器
- 在浏览器中安装新协议和内容管理器
- 加密通信、身份验证和保证消息的完整性
- 设计网络服务的 GUI 客户程序

- 向 CGI 程序传递数据
- 使用 DNS 查找主机
- 使用匿名 FTP 下载文件
- 底层网络通信的连接套接字 (Socket)
- 用 RMI (Remote Method Invocation, 远程方法调用) 建立跨越多系统的分布式应用程序

Java是第一个提供了如此强大的跨平台网络类库的语言,这些类库能够完成处理各种各样的任务。本书展现了类库的强大能力以及它的精巧之处,目的是使你能够开始使用Java作为编写重要网络程序的平台。为了达到这个目的,本书提供了网络基本原理方面的通用背景并详细讨论了用于编写网络程序的Java技巧。你将学习如何编写游戏、协作、软件更新、文件传输等这些通过 Internet 共享数据的 Java 例程和应用程序;同时,对 HTTP, CGI, TCP/IP 以及其他支持 Internet 和 Web 的协议的了解也不再浮于表面。读完本书,你将会拥有创建充分利用 Internet 优势的新一代软件的工具。

关于第二版

在本书第一版的第一章中,我曾大讲特讲Java可以实现的各种动态的、分布式的网络应用程序;在第二版撰写过程中,很高兴能看到我假设的所有应用程序在实际中都已经实现。程序员利用Java查询数据库服务器,监控Web页面,控制压缩,管理多玩家游戏和其他更多的功能,所有这一切都是通过使用Java的功能访问Internet来实现的。从整体上讲,Java特别是Java网络编程已远远超越了广告宣传的阶段,而迈入了实际工作应用的领域。虽然不是所有的网络软件都依靠Java语言编写,但并不缺乏利用Java编写程序的尝试,使用纯Java改造现有的一些基于C语言的基础软件的工作,正在努力地进行着。Java完全代替C语言实现网络编程在近期是不太可能的;但是,许多人愿意使用Java编写的Web浏览器、Web服务器和其他更多的应用程序这一起码的事实表明,Java自1996年出现以来已经得到了极大发展。

本书的出版也经历了一段相当长的时间。第二版的改写几乎是从零开始,其中有五章内容是全新的,有的章节涉及了自第一版出版以来Java出现的新功能和新的API(第八章、第十二章和第十九章);有些章节是我从教授这些知识时学生们遇到的难

点中获得的宝贵经验（第四章和第五章）。另外，本书删去了关于 Java Servlet API 的一章内容，因为该主题需要一整本书来讲述，实际上 Jason Hunter 已经撰写了《Java Servlet Programming》一书（O'Reilly & Associates, Inc., 1998）。

但是，比新增的内容和删除的内容都重要得多的是在我们对保留的章节所做的改进。其中对第一版最明显的改进，是所有的例程都用 Java 1.1 的 I/O API 重新编写了，这样你使用 Java 1.1 或更高版本编译例程时，不会再有出现警告消息框中断编译过程的现象；同时，改进虽不明显但却更为重要的是，所有的例程都按 Java 命名习惯和设计规范，采用清晰的面向对象的设计方式重新编写了。和大多数人一样，我 1996 年撰写第一版时，也是费尽心思考虑你会使用 Java 做些什么、以及如何实现等等这些细节问题，这恐怕是 Sun 公司没有想到的。尽管原先的例子是正确的，但是从现在看来在许多其他方面，它们显得相当业余。自撰写第一版以来，我已学习了 Java 和面向对象编程的许多知识，我想自己增长的经验已经在第二版中有所表现。仅举一例，我不再使用简单的基于框架就可以满足要求的应用程序，我希望新的例程不仅仅是如何编写网络程序的模范，同时也能成为在普遍意义上如何编写 Java 代码的模范。

当然，本书的正文也进行了整理。实际上，我用了与撰写第一版同样的时间修订了第二版。如前所述，第二版中有五章内容是全新的，但是有十四章的内容大面积地改写并扩充了最新的知识，同时我力求使它们显得简洁、清晰并更加高效。坦白地说，即使不考虑对例程的所有改进，第二版也要强于第一版。我希望你将会发现该版本中对 Java 网络编程的教程和参考，比第一版更加强大、生命期更长、准确度更高，并且更加有趣。

本书结构

本书开始的三章内容概述网络和网络程序如何工作。第一章“为什么 Java 要网络化？”，是对 Java 网络编程以及各种可能的应用的简要介绍；所有的读者都会在本章中找到感兴趣的内容。它探究了一些独特的程序，这些程序在网络与 Java 结合时会变得切实可行。第二章“网络的基本概念”和第三章“Web 的基本概念”详细解释了关于 Internet 和 Web 如何工作方面，程序员需要了解一些什么知识。第二章描述了构成 Internet 基础的协议，例如 TCP/IP 和 UDP/IP；第三章介绍了构成 Web

基础的标准，如 HTTP，HTML 以及 CGI。如果你已经在其他平台上以其他语言编写过很多网络程序，可以略过这两章的内容。

接下来的两章讲述了 Java 的两个组成部分：I/O 和线程，这两个内容在几乎所有的网络程序中都起着关键的作用，但是它们又经常被误解和误用。第四章“Java I/O”探究了 Java 处理输入和输出的独特方式，要知道，理解 Java 在一般情况下如何处理 I/O，是理解它在特殊情况下如何处理网络 I/O 的先决条件；第五章“线程”探讨了多线程和同步，并特别强调了如何使用它们处理异步 I/O 及网络服务器。具有丰富的 Java 编程经验的程序员可以略读或跳过这两章。但是，第六章“查找 Internet 地址”对任何人都属于必读内容。它介绍了 Java 程序如何通过 InetAddress 类与 DNS (Domain Name System, 域名系统) 进行交互，这个类对于所有网络程序都是必不可少的。当你学完该章的内容，就可以按照自己的兴趣和需要翻读本书的其他内容。

第七章“用 URL 寻取数据”探讨了 Java 的 URL 类，这个类是一个强大的抽象类，用于从各种网络服务器上下载信息和文件。URL 类使用户可以直接与网络服务器连接并从服务器上下载文件与文档，而不必了解服务器所采用的协议的细节。它允许用户在连接一个 FTP 服务器时所使用的代码，与访问 HTTP 服务器或在本地硬盘上读取文件的代码相同。

用户一旦从服务器上获取了 HTML 文件，就想对其进行一定的处理。解释和编译 HTML 是网络程序员所面对的最困难的挑战之一。事实上，Mozilla 项目在这个问题上已经奋斗了两年多。第八章“Swing 中的 HTML”介绍了用于解析和翻译 HTML 文档的一些鲜为人知的类，这些类把原来用户承担的负担转移到平台上。

第九章“java.applet.Applet 的网络方法”分析了 Applet，它是每个 Java 程序员学习第一个类的网络编程方法。你将学习如何从服务器上装载图像文件及音频文件并跟踪它们的过程。如果不使用没有文档说明的那些类，这是在 Java 1.2 和更早的版本中处理音频文件的唯一方法。

第十章到第十四章讨论了在 Java 中用于网络访问的底层 Socket 类。第十章“客户端套接字”特别介绍了 Java 套接字 API 及 Socket 类，探究了如何编写与各种包括 whois、Finger 及 HTTP 在内的 TCP 服务器进行交互的网络客户程序。第十一章“服务器套接字”向你展示了在 Java 中如何使用 ServerSocket 类为 Java 中各种类型的协议编写服务器。第十二章“安全套接字”讲述了如何使用 SSL (Secure Sockets Layer, 安全套接字层) 和 JSSE (Java Secure Sockets Extension, Java 安全套接字扩展)

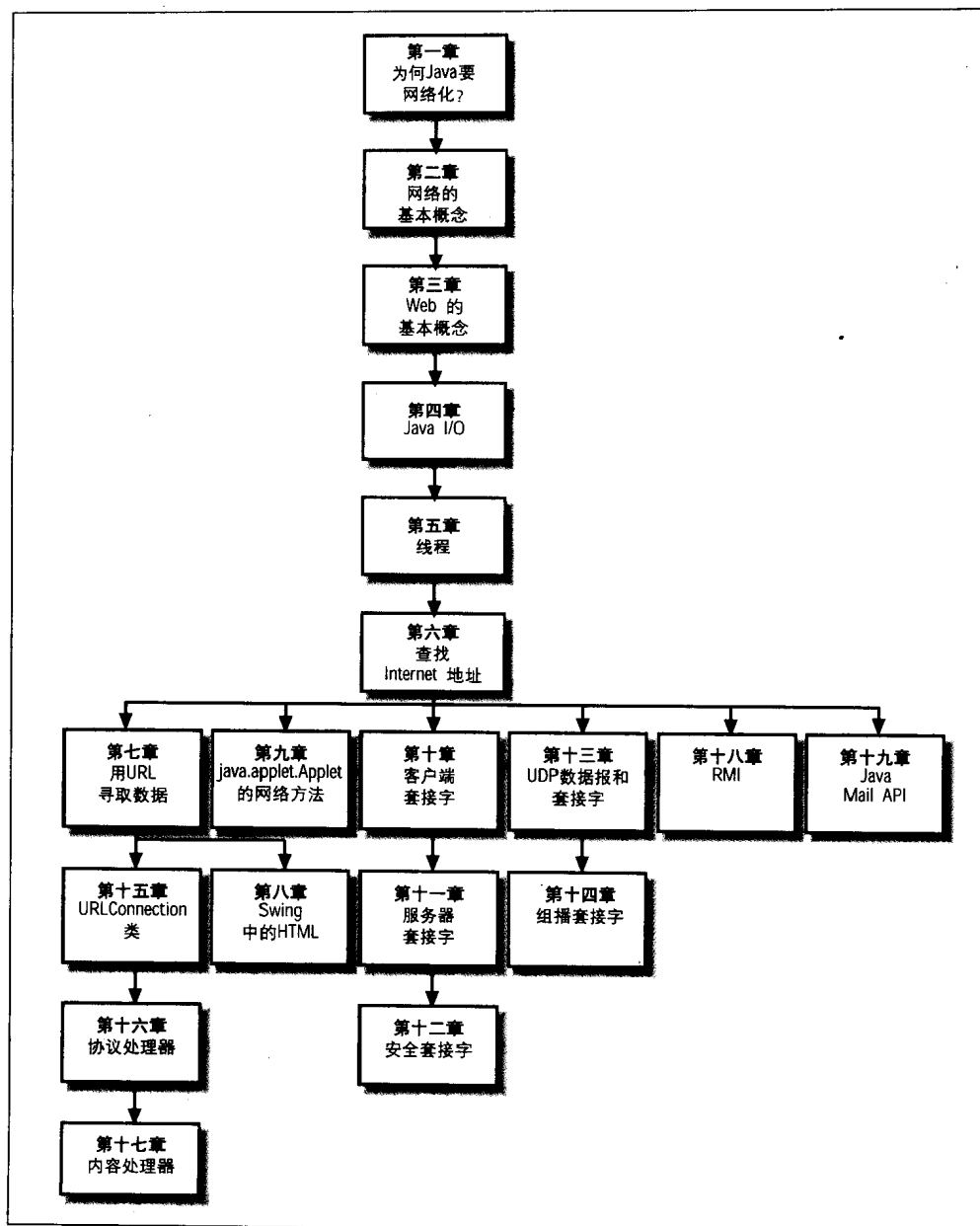


图 0-1 各章阅读的前后顺序

保护用户的客户/服务通信。第十三章“UDP 数据报和套接字”介绍了 UDP (User Datagram Protocol, 用户数据报协议) 以及与其相关的用于有效和可靠通信的类:

DatagramPacket 和 DatagramSocket。最后，第十四章“组播套接字”向你展示了如何使用UDP与多个主机在同一时刻进行通信。所有从Java访问Internet的全部其他类都依赖于在这五章中描述的类。

第十五章到第十七章更为深入地描述了支持URL类的基础结构的相关内容。这两章介绍了Java的一些独特的协议以及内容管理器和概念，这些工具使编写能够自动地理解新的协议和媒体类型的动态扩展软件成为可能。第十五章“URLConnection类”描述了作为第七章中URL类的引擎的URLConnection类，展示了如何通过它的公开（public）API利用这个类。第十六章“协议处理器”也集中讲述URLConnection类，但方向不同，它向你介绍了如何将此类转为子类，从而为新协议和URL创建管理器。最后，第十七章“内容处理器”介绍了Java中一些支持新媒体类型的即将废弃的机制。

第十八与第十九章介绍了网络编程的两种独特的高层API: RMI和JavaMail API。第十八章“RMI”描述了这一编写Java分布式应用程序的强大机制，该种类型的程序能在同一时刻跨越多种不同类型的系统执行，而使用直接访问方式进行通信的称为非分布式的通信。第十九章“JavaMail API”的目的是使用户熟悉这个Java标准扩展，该扩展是访问SMTP、POP、IMAP和其他邮件服务器的底层Socket的另一选择。这两种API都为分布式应用程序提供了比底层协议更轻巧的替代方案。

适用对象

本书假设你已基本熟悉Java语言和编程环境，另外在大体上了解了面向对象编程。本书并不打算成为一本基础语言指南，你应当十分熟悉Java语言的语法，编写过简单的应用程序和applet，同时还要了解AWT。当你遇到要求对网络编程有更深层次理解的主题（例如，线程和流）时，我会对其内容进行概述，至少是简要的论述。

你还应当是Internet的熟练用户，知道如何FTP文件和访问Web站点，知道URL是什么并了解如何找到一个URL。尽管你不需要成为一个高级Web设计者，但应该了解如何编写一个简单的HTML，并能够发布包含Java applet在内的主页。

但是，本书不需要你具有网络编程的经验。你会发现对于网络概念和网络应用程序的发展都有完整的介绍。也不需要你知道数以千计的网络名词缩写（TCP、UDP、

SMTP……), 在本书中将只学习你必须知道的缩写名词。你当然可以把本书作为用类似 Socket 接口进行网络编程的概要介绍, 然后接着学习 Windows 套接字体系 (WSA, Windows Socket Architecture), 并考虑如何用 C++ 编写网络应用程序。但是, Java 的目的是让你用简单的方法写出极其复杂的应用程序。

Java 的版本

自 Java 1.0 版本出现以来, 它的网络类库比其内核 API 的其他部分改变的要少得多。与 AWT 或 I/O 比较, Java 网络类库几乎没有改变而仅是增加了一些内容。当然, 所有的网络程序大量地使用了 I/O 类, 并频繁地利用了 GUI。本书假设你和你的用户至少使用 Java 1.1 版本 (这个假设或许在 2001 年会变得很有把握)。通常, 我使用 Java 1.1 版本的特性, 如 reader、writer 和新的事件模型时, 不再多加解释。

Java 2 版本有了更多的扩展。虽然我使用 Java 2 编写本书中几乎所有的例程, 而且 Java 2 在 Windows 和 Solaris 中的使用也已一年有余, 但是对于 Mac 系统, Java 2 没有执行或开发的环境。尽管 Java 2 已经逐步发展到大多数的 Unix 平台以及 Linux 平台, 但几乎可以确定 Apple 或 Sun 都不会将 Java 2 移植到 MacOS 9.x 或更早的版本上, 因此现在安装了低版本的 100% 的 Mac 机被完全排斥在未来开发工具的门外了 (Java 2 或许会在 2001 年某个时间出现的 MacOS X 上)。这对自称为“一次编写, 随处执行”的一种语言并不是一件好事。此外, Microsoft 的 Java 虚拟机仅支持 Java 1.1 版本, 并且在可预知的未来似乎不愿在这方面改进了 (因为各种官司缠身)。最后, 目前几乎所有浏览器, 包括 Internet Explorer 5.5 及其更早的版本以及 Netscape Navigator 4.7 及其更早的版本仅支持 Java 1.1。applet 的开发者受到其用户能力的极大限制, 而不得不使用 Java 1.1, 这样 Java 2 看起来至少在近期会局限于 Windows 和 Unix 上的应用程序。因此, 当我由于有用或方便而使用了 Java 2 中独有的功能时 (例如, InputStremReader 的 ASCII 编码和 *keytool* 编程), 我会小心地指出。在任何可以安全使用的 Java 1.1 版本的地方, 对它们都有注释。当一种特殊方法或类属于 Java 1.2 或更高版本时, 会在其声明之后加入注释, 形式如下:

```
public void setTimeToLive(int ttl) throws IOException // Java 1.2
```

Java 2 版本的众多使人感到有点云山雾罩。在本书完成时, 当前版本是 Java 2 SDK, Standard Edition v1.2.2 (J2SE, Java 2 标准版), 至少那时是这样命名的。Sun 似乎咨询了市场顾问, 把名字改了, 以前的成品简单地称为 JDK。Sun 还发布了 Java

2 平台的 Enterprise Edition (J2EE, Java 2 企业版) 和 Micro Edition (J2ME, Java 2 微型版)。J2EE 是 J2SE 的超集, 增加了一些特性, 如 JNDI (Java Naming and Directory Interface, Java 命名和目录接口), 并提供了分布式应用程序的高层 API: JavaMail API。这些增加的 API 中的一些接口同样也作为 J2SE 的扩展来使用, 并且在本书中也将它们作为扩展接口对待。J2ME 是 J2SE 的子集, 其应用对象主要是蜂窝电话、顶置盒及其他存储器、CPU 和显示设备。尽管它令人惊奇地保留了本书中讨论的几乎所有的基本网络类和 I/O 类, 但放弃了许多程序员已经学习的与 Java 相关的 GUI API, 最后, 当本书完成约一半时, Sun 发布了 J2SE 的 Java 2 JDK v1.3 试用版, 它在网络 API 方面加入了一些内容, 但对大部分已有的 API 没有进行改动。在接下来的几个月里, Sun 发布了更多的 JDK 1.3 试用版。本书采用了最新版本, 并且所有的代码都使用 JDK 1.3 最终发行版进行了测试 (译注 1)。

坦白地说, 最令人苦恼的事不是重写各个不同版本必须要求改动的部分, 而是考虑如何在文章中表示它们。我不想在每次要指出哪里哪里是 Java 最新版的一些新特性时, 直接用 “Java 2 SDK, Standard Edition v1.3”, 或 “Java 2 1.3”。因此, 采用下列约定:

- *Java 1.0* 指实现了 Sun 公司 JDK 1.0.2 所定义的 Java API 的所有 Java 版本。
- *Java 1.1* 指实现了 Sun 公司 JDK 1.1.x 任何版本所定义的 Java API 的所有 Java 版本, 这里包括第三方的产品, 如 Macintosh Runtime for Java (MRJ) 2.0, 2.1 和 2.2。
- *Java 1.2* 指实现了 Sun 公司 JDK 1.2.x 标准版中所定义的 Java API 的所有 Java 版本, 不包括将作为标准版扩展讨论的 J2EE 的增加部分。它们一般出现在 Javax 包中, 而不是 Java 包。
- *Java 1.3* 指实现了 Sun 公司 JDK 1.3 标准版中所定义的 Java API 的所有 Java 版本。

总而言之, 本书涵盖了使用 Java 2 进行网络编程的方方面面, 这些网络程序与使用 Java 1.1 编写的大体相同。如果可以获得更多的信息, 我将在站点 <http://metalab.>

译注 1: 本书中文版付梓前, 最新的正式版本是 Java™ 2 SDK, Standard Edition, v1.4 Beta (2001 年 7 月 17 日发布)。