

**Broadview**  
www.broadview.com.cn

THE EXPERT'S VOICE® IN OPEN SOURCE

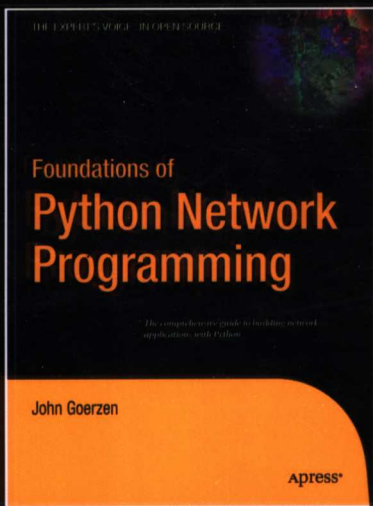
**Apress®**

Foundations of Python Network Programming

# Python 网络

# 编程基础

使用Python构建网络程序的综合指南



[美] John Goerzen 著  
莫 迟 等译



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
<http://www.phei.com.cn>

# Python 网络编程基础

---

Foundations of Python Network Programming

[美] John Goerzen 著

莫 迟 等译

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

## 内 容 简 介

本书全面介绍了使用 Python 语言进行网络编程的基础知识, 主要内容包括网络基础知识、高级网络操作、Web Services、解析 HTML 和 XHTML、XML、E-mail 服务、FTP、使用 Python 操作数据库、SSL、几种服务器端框架(包括 Socket 服务器、SimpleXMLRPCServer、CGI 和 mod\_python), 以及多任务处理(包括 forking、线程和异步通信)等。本书实用性强, 共提供了大约 175 个实例, 6 600 行以上的代码, 是帮助读者全面而快速地学习 Python 语言、编写网络程序的最佳实践。

本书可以作为各层次 Python、Web 和网络程序开发人员的参考书, 在实际工作中使用书中的技术, 效果更佳。

1-59059-371-5 Foundations of Python Network Programming by John Goerzen.

Original English language edition published by Apress L. P., 2560 Ninth Street, Suite 219, Berkeley, CA 94710 USA. Copyright © 2006 by Apress L. P. Simplified Chinese-language edition copyright © 2007 by Publishing House of Electronics Industry. All rights reserved.

本书简体中文专有翻译出版权由 Apress L. P. 公司授予电子工业出版社, 未经许可, 不得以任何方式复制或抄袭本书的任何部分。

版权贸易合同登记号 图字: 01-2007-0689

### 图书在版编目(CIP)数据

Python 网络编程基础 / (美)高森(Goerzen,J.)著; 莫迟等译. —北京: 电子工业出版社, 2007.6

书名原文: Foundations of Python Network Programming

ISBN 978-7-121-04495-3

I. P… II. ①高…②莫… III. 软件工具—程序设计 IV. TP311.56

中国版本图书馆 CIP 数据核字(2007)第 074160 号

责任编辑: 周 筠 杨绣国

印 刷: 北京市天竺颖华印刷厂

装 订: 三河市金马印装有限公司

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: 34 字数: 700 千字

印 次: 2007 年 6 月第 1 次印刷

定 价: 68.00 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 zhts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

谨以此书献给我的妻子 Terah：感谢你在我写书期间，能够容忍院子里面高高的杂草，而且每天工作到很晚以及牺牲周末的时间。感谢你用鼓励、爱和巧克力来支持我的每一步。

# 联系博文视点

---

您可以通过如下方式与本书的出版方取得联系。

读者信箱: [shenguang@broadview.com.cn](mailto:shenguang@broadview.com.cn)

投稿邮箱: [broadvieweditor@gmail.com](mailto:broadvieweditor@gmail.com)

北京博文视点资讯有限公司（武汉分部）

湖北省 武汉市 洪山区 吴家湾 邮科院路特1号 湖北信息产业科技大厦 1402 室

邮政编码: 430074

电话: (027)87690813 传真: (027)87690813 转 817

若您希望参加博文视点的有奖读者调查, 或对写作和翻译感兴趣, 欢迎您访问:

<http://bv.csdn.net>

关于本书的勘误、资源下载及博文视点的最新书讯, 欢迎您访问博文视点官方博客:

<http://blog.csdn.net/bvbook>

# 译 序

---

大概是因为工作的关系，每次去书店的时候，都要看看关于 Python 的书。很遗憾，有的书店根本就没有，而有的书店虽有几本，却和其他一些不好分类的书放在一起。而这仅有的几本基本上也都是介绍 Python 基础的。对比国内 Python 的冷清，浏览国外的招聘网站时，可以看到需要大批的 Python 开发人员。另一方面，当今世界网络无处不在，在掌握了 Python 基础之后，很多人都会想更进一步，而这本书正是一本系统且全面介绍使用 Python 进行网络编程的书。

本书的编排非常清晰，几乎涵盖了网络编程的所有方面。从传统的 FTP、E-mail 到较新的 XML、Web Service，以及到当前流行的多线程和异步通信，本书都有详细的介绍。本书还给出了大量直接（或稍许修改后）可以使用的例子。如果在学习和工作中灵活应用这些例子，一定会得到事半功倍的效果。

个人感觉学习一门计算机语言的最有效的方法就是真正使用它，本书也不例外。所以我建议读者通读完本书后，记住书中大致的内容，便于在实际编程中迅速找到书中相关的部分，得到帮助。

一本书的翻译出版不是一个人的事情，这里首先要感谢原书作者 John Goerzen，没有他就不会有这本书。随着翻译的进行，越来越觉得他是一位 Python 和软件工程方面的大师。然后要感谢的是彭俊先生，他总是仔细且耐心地和我探讨翻译中不合适的地方。互相探讨本身也是一个提高自己的过程。还要感谢周筠女士、杨绣国女士、杨福川先生等电子工业出版社的工作人员，是他们的严谨、专业使得这本书顺利出版。还有 limodou 先生，在我学习 Python 的过程中给予了很多帮助。最后还要感谢我的家人以及我的妻子 Kitty、我的儿子非非，没有他们的鼓励和支持，这本书不会这么快和中国的读者见面。

由于我自身的水平有限以及时间的仓促，翻译中难免有错误和遗漏，有些专业和习惯用语直

接使用了原文，而有些句子的翻译可能无法完全符合汉语的习惯，还请读者以及原书作者原谅。读者能够从本书中得到帮助，将是我开心的事情。

关于本书中的代码注释，编辑建议我也翻译成中文，以方便读者理解。经过再三考虑，我还是决定保持原样。原因如下：

每个例子，书中都有详细的讲解。注释只是起到一个参考的作用；

学习一种语言，阅读别人的程序是一种非常好的方法，我们力争为读者展现作者原汁原味的代码（包括注释）。请读者不要小看注释，作为一个软件从业人员，我一直鼓励周围的开发人员多写注释，而且要多用英文写注释。计算机语言基本上都是英文的，练习用英文写注释一方面可以提高自己的英文水平，另一方面还可以渐渐和世界接轨，将来很有可能外国同事阅读您的代码；

Python 本身是一种跨平台的语言。如果用中文写注释，将来移植的时候也许会出现问题；而当读者按照书中的英文代码和中文注释来试验这些例子的时候，可能会由于运行环境的不同而产生错误。

莫 迟  
2007 年 3 月

# 关于作者

## About the Author

John Goerzen 从 1996 年开始就是 Debian GNU/Linux 操作系统开发组的成员，最近 15 年他一直都在从事软件开发方面的工作。他对操作系统、程序语言和网络都很感兴趣，并且在工作中也开发并用到以上这些方面的各种软件。目前，他在一个中等规模的制造公司中担当程序员和 Unix 管理员，并且已经使用 Python 开发了很多用于公司数据系统的接口。



作为 Debian 工作组成员，John 为系统维护了很多不同的程序。从 1998 年开始，他就致力于启动或向新的体系结构扩展 Debian，并在 Alpha、PowerPC、AMD64 和 NetBSD i386 上取得了很大的成就。

2003 年，John 被选进了 Software in the Public Interest (SPI) 的董事会。这个公司负责管理 Debian 在法律和财务上的事务。同年晚些时候，他被任命为 SPI 的副会长。

除作为 Debian 工作组成员之外，John 还用了大量的时间写软件。在他众所周知的作品中，有两个是用 Python 写的网络激活程序。OfflineIMAP 是一个双向的邮件同步程序，而 PyGopherd 是一个多协议的 Web 和 Gopher 服务器。这两个软件都用 Python 独有的特性来支持模块体系结构，并且提供了很好的灵活性。

John 还写了很多书，包括那本 800 页的《Linux Programming Bible》。他还经常为杂志撰写文章，同时还是很多书的技术编辑。他还建立了 Air Capital Linux 用户组，并经常在组里演讲。



# 关于技术审校

---

## About the Technical Reviewer

Magnus Lie Hetland 是挪威科技大学算法方面的副教授，他从 1997 年开始使用 Python。他是流行的网络杂志 “Instant Hacking” 和 “Instant Python” 的作者。他的出版物除了一些科学著作外还有《Practical Python》（Apress, 2004）。



# 致谢

---

## Acknowledgments

写这本书受益于很多人的知识、经验、贡献和鼓励。没有这些人，这本书恐怕很难完成，所以我尤其想利用这个机会感谢如下这些人：

- **Magnus Lie Hetland**，技术审校。我总是想具备 Magnus 那样的 Python 水平，他观察事情非常敏锐，而且总是能够用很简单的方式来解释事情。我非常感谢他的是，他发现了比我预计要多得多的问题。
- **Ellie Fountain**，产品编辑。在我努力找一台可以运行例子文件的机器时，Ellie 非常地有耐心。
- **Beth Christmas**，项目经理。每当我不知道下一步该做什么的时候，Beth 总是会通过 E-mail 及时提醒我。在时间进度上，她总是做得非常好。我经常奇怪为什么她总能安排得那么好，尽管有时我要落后于计划。
- **Jason Gilmore**，责任编辑。Jason 关于每一章的评论使本书的内容更加丰富。书写到一半的时候，我回头看了一下，想到 Jason 一定会让我以另外的方式来写。是他的鼓励使我度过了那些有意外发生的混乱日子。
- **Mark Nigara**，文字编辑。经过 Mark 的校订，拼写没有问题了，段落更清晰了，而且也更方便阅读了。看到 Mark 的校订后，我经常惊讶于他能对书稿提出如此多的修改意见。

还有很多人为了这本书用到的技术做出了贡献。我想着重感谢一下他们：

- **致 Guido van Rossum**：感谢您为我们发明了 Python 这门通用的语言，用 Python 编写的程序非常容易读懂。

- 致 Richard M. Stallman: 感谢您让我们领略到了协作和共享技术的价值。

最后, 还要感谢 Vint Cerf, Robert Kahn 和 Jonathan Postel 等互联网先驱。是你们让我们能够利用网络来传送信息。

# 简介

---

## Introduction

25年前，世界和现在是完全不同的。很少人能有机会和大洋彼岸的人谈话。寄一封信需要好几周的时间。收听外国的短波新闻广播除了需要专门的技术和耐心外，还要非常好的气候条件。

而今天，收到来自韩国的 E-mail、查看加利福尼亚的天气，以及阅读德国当天的头条新闻对我们来说都是很平常的事情，而且都可以在五分钟之内完成。压缩文件在互联网上的传输使得我们可以做很多事情，从管理投资账户到观看远房亲戚的照片。

尽管 Internet 已经有 25 年的历史了，但它还是处在幼年阶段。作为一种新技术，它还在逐渐成长。

我写这本书的原因是因为 Internet 是那么地让人兴奋。在过去的几年中，我们看到了整个行业的增长，而这些是以前没有的。同时互联网也是产生大量发明家的地方。

而且这也是我希望您从本书获得的，我希望这本书能成为您的实验手册——您为了使 Internet 更美好而进行发明创造的指南。

### 本书的结构

本书共分 6 个部分。第 1 部分介绍了 Internet 是如何工作的。您将学习 Internet 通信的基础。通过例子您还将学到编写您自己的程序要用到的一些基本工具。

第 2 部分主要讲基于 Web 服务方面的知识。第 3 部分主要讲和 E-mail 服务有关的知识。在 Web 或 E-mail 通信中使用了很多新的技术。在这两部分内容中，介绍了如何使用这些新技术来编写程序。

第 4 部分介绍了其他一些技术，例如：数据库和文件传输。它们经常运行在网络程序的后台。您的用户也许永远都不知道您使用了这些技术，但是它们的确是您工具箱中重要的工具。

第 5 部分介绍了如何编写服务器程序，这些应用程序可以回应请求并给出信息。您也许永远不用写一个服务器程序，但是如果您正为一个新的协议打草稿的话，您将需要这些技术。

最后，第 6 部分向您展示了如何同时做多件事情。有些网络程序员也许永远都不会用到这些技术，但是对于一些服务器的设计者来说，没有这些技术是万万不行的。

## 假定

对于本书的读者，我假定您已经具有一定的 Python 知识。如果您没有，我建议您学习 Apress 出版的、由 Magnus Lie Hetland 编著的《Practical Python》。

我还假定您的机器上安装了 Python 2.3 或更高的版本，并且有网络编程的部分。如果您的机器上没有安装 Python，请检查您的操作系统以及它的安装包，因为有的操作系统本身就带有 Python。否则您可以从 [www.python.org](http://www.python.org) 下载 Python 的源文件和安装程序。

最后，我还假定您的机器是和互联网相连的（尽管这个连接或许不是您自己配置的）。此外，阅读本书，您不需具备任何互联网协议的知识。

## 例子

本书包含很多例子，您可以从 Apress 的站点 [www.apress.com](http://www.apress.com) 下载。我建议您试着运行所有的例子，其中有些例子完全可以作为您自己程序的起点，而且绝大多数的例子都是完整的、可以直接运行的程序。您可以自己查看运行的结果。

我还会经常向您演示一些程序交互运行的结果。当看到这样的例子的时候，粗体字表示需要通过键盘输入的内容。当看到计算机命令的时候，“\$”表示操作系统的命令提示符。

网络程序对操作系统的网络支持情况是非常依赖的。操作系统提供了和其他计算机通信的所有基础。这里有一些针对不同操作系统的注释。

**技巧：**如果您在运行例子的时候遇到任何问题，请翻阅本章中的这些注释。有些例子并不是能在所有的操作系统上运行的。

## Linux、FreeBSD、Mac OS X、Solaris和其他UNIX操作系统

所有的例子都假设您已经在系统路径上安装了 Python 解释器，并且是以 Python 命名的。还假设您的 `/usr/bin/env` 程序是有效的。每一个可运行的 Python 程序的第一行都是以 `#!` 开始的，

这可以告诉操作系统去哪里找 Python 解释器来运行您的程序。如果所有的例子都不能运行，那就说明 Python 没有安装或者系统找不到它。可以试着把#!后面的代码换成您机器上 Python 解释器的绝对路径。

同时，所有要运行的程序必须具有可执行的权限。如果当您运行某个程序的时候，系统提示有关权限的问题，请运行诸如 `chmod 0755 scriptname.py` 的命令来使该文件具有可执行的权限。

如果是其他任何错误，您则可以手动运行 Python 解释器，即运行 `python scriptname.py` 来调用这个程序，而不是运行 `./scriptname.py`。

在本书中，如果我提到 Linux/UNIX 平台，我的意思就是任何类似 UNIX 的操作系统，也包括 Mac OS X。任何早于 OS X 的 Mac 操作系统，都不在本书的讨论范围之内。

## Windows操作系统

Windows 操作系统并不支持 Linux/UNIX 平台中用于程序第一行的#!方法。所以在 Windows 操作系统中不能用 `./scriptname.py` 这种方式来执行程序。只要您看到类似 `./scriptname.py` 这样的指令，请用 `python scriptname.py` 来替换。

还有一些 Linux/UNIX 平台特有的功能是 Windows 操作系统不具备的（或者说，有些情况下 Python 不能提供在 Linux/UNIX 平台上提供的功能）。明显的例子是第 3 章中介绍的 `inetd` 类型的服务器（`inetd-style server`）、第 20 章中介绍的 `forking`，还有第 22 章中介绍的部分关于异步通信（`asynchronous communication`）的内容。然而在这些案例中，我们介绍的技术仅仅是多种解决方法中的一种。您完全可以通过另外一种不同的方法来解决同样的问题。在这些章节的注释中，您将看到这些替换的方法。

Python 有很多不同的 Windows 版本。标准的版本来自 [www.python.org](http://www.python.org)，但它也并不能完成所有的任务，而且有的版本还有一些与网络编程有关的错误（Bug）存在。另外一个常用的版本是 ActivePython，可以从 [www.activestate.com](http://www.activestate.com) 得到。无论使用哪种版本，如果遇到一些意料之外的问题，您可以试着换用其他的版本。

部分具有 Linux/UNIX 背景的读者可能想试一下 Cygwin，您可以从 [www.cygwin.com](http://www.cygwin.com) 得到。它为 Windows 操作系统提供了一个类似 Linux 的环境，这种情况下，需要从 [www.python.org](http://www.python.org) 下载 Python 源文件，经过编译后才能使用。

## 其他有用的资源

您也许会遇到一些特殊方面的问题，这里有一些网络上的资源或许能帮上忙：

- 正式的 Internet 协议标准，也叫 RFC（Request for Comments）文档，您可以在 [www.rfc-editor.org](http://www.rfc-editor.org) 和 [www.faqs.org](http://www.faqs.org) 上找到。
- 对于 Python 方面的文档，在 [www.python.org/doc](http://www.python.org/doc) 上的《Python 模块参考》（Python module reference）是非常有用的。
- 当您遇到问题的时候，*comp.lang.python* 新闻组是一个很好的提出问题和寻找答案的地方。详细情况，请访问 [www.faqs.org/faqs/python-faq/python-newsgroup-faq/](http://www.faqs.org/faqs/python-faq/python-newsgroup-faq/)。该网页上还介绍了如果没有新闻组软件（Usenet news access），如何用 E-mail 来参与讨论。
- 您的操作系统开发文档也会提供一些底层的网络操作和网络配置的信息。

## 反馈

我很乐意收到您的来信。欢迎您提出对本书的意见和建议。我的 E-mail 地址是 [jgoerzen+pynet@complete.org](mailto:jgoerzen+pynet@complete.org)。尽管我会阅读所有的 E-mail，但是请原谅我没有时间一一回复，所以请千万不要因为没有收到我的回信而生气。

# 目录一览

---

## Contents at a Glance

关于作者 .....	I
关于技术审校 .....	III
致谢 .....	V
简介 .....	VII
<b>第 1 部分 底层网络 .....</b>	<b>1</b>
第 1 章 客户/服务器网络介绍 .....	3
第 2 章 网络客户端 .....	19
第 3 章 网络服务器 .....	35
第 4 章 域名系统 .....	65
第 5 章 高级网络操作 .....	87
<b>第 2 部分 Web Service .....</b>	<b>111</b>
第 6 章 Web 客户端访问 .....	113
第 7 章 解析 HTML 和 XHTML .....	127
第 8 章 XML 和 XML-RPC .....	145
<b>第 3 部分 E-mail 服务 .....</b>	<b>167</b>



第 9 章 E-mail 的编写和编码.....	169
第 10 章 简单邮件传输协议 (SMTP) .....	197
第 11 章 POP .....	211
第 12 章 IMAP .....	223
<b>第 4 部分 多用途的客户端协议.....</b>	<b>273</b>
第 13 章 FTP .....	275
第 14 章 数据库客户端.....	295
第 15 章 SSL .....	321
<b>第 5 部分 服务器端框架.....</b>	<b>339</b>
第 16 章 SocketServer.....	341
第 17 章 SimpleXMLRPCServer.....	355
第 18 章 CGI.....	369
第 19 章 mod_python .....	393
<b>第 6 部分 多任务处理.....</b>	<b>417</b>
第 20 章 forking .....	419
第 21 章 线程.....	443
第 22 章 异步通信.....	469
索引.....	491