****

**THE LINUX PROGRAMMING INTERFACE**

**A Linux and UNIX System Programming Handbook**

**MICHAEL KERRISK**

no starch press

San Francisco

*Translated by: Kevin Zhang*

*2011-01-26 – 2011-05-15(TODO)*

***本资料仅供学习所用，请于下载后24小时内删除，否则引起的任何后果均由您自己承担。本书版权归原作者所有，如果您喜欢本书，请购买正版支持作者。***

目录

[前言 7](#_Toc285366210)

[主题 7](#_Toc285366211)

[目标读者 7](#_Toc285366212)

[Linux和UNIX 8](#_Toc285366213)

[使用和组织 8](#_Toc285366214)

[例子程序 9](#_Toc285366215)

[练习 10](#_Toc285366216)

[标准和可移植性 10](#_Toc285366217)

[Linux内核和C库版本 11](#_Toc285366218)

[其它语言使用编程接口 11](#_Toc285366219)

[关于作者 11](#_Toc285366220)

[致谢 11](#_Toc285366221)

[许可 12](#_Toc285366222)

[网站和例子程序源代码 12](#_Toc285366223)

[反馈 12](#_Toc285366224)

[第1章 历史和标准 13](#_Toc285366225)

[第2章 基础概念 14](#_Toc285366226)

[第3章 系统编程概念 15](#_Toc285366227)

[第4章 文件I/O：统一的I/O模型 16](#_Toc285366228)

[第5章 文件I/O：更多细节 17](#_Toc285366229)

[第6章 进程 18](#_Toc285366230)

[第7章 内存分配 19](#_Toc285366231)

[第8章 用户和组 20](#_Toc285366232)

[第9章 进程凭证 21](#_Toc285366233)

[第10章 时间 22](#_Toc285366234)

[第11章 系统限制和选项 23](#_Toc285366235)

[第12章 系统和进程信息 24](#_Toc285366236)

[第13章 文件I/O缓冲 25](#_Toc285366237)

[第14章 文件系统 26](#_Toc285366238)

[第15章 文件属性 27](#_Toc285366239)

[第16章 扩展属性 28](#_Toc285366240)

[第17章 访问控制列表 29](#_Toc285366241)

[第18章 目录和链接 30](#_Toc285366242)

[第19章 监控文件事件 31](#_Toc285366243)

[第20章 信号：基础概念 32](#_Toc285366244)

[第21章 信号：信号处理器 33](#_Toc285366245)

[第22章 信号：高级特性 34](#_Toc285366246)

[第23章 定时器和睡眠 35](#_Toc285366247)

[第24章 进程创建 36](#_Toc285366248)

[第25章 进程结束 37](#_Toc285366249)

[第26章 监控子进程 38](#_Toc285366250)

[第27章 程序执行 39](#_Toc285366251)

[第28章 进程创建和程序执行的更多细节 40](#_Toc285366252)

[第29章 线程：介绍 41](#_Toc285366253)

[第30章 线程：同步 42](#_Toc285366254)

[第31章 线程：线程安全和线程存储 43](#_Toc285366255)

[第32章 线程：线程取消 44](#_Toc285366256)

[第33章 线程：更多细节 45](#_Toc285366257)

[第34章 进程组、会话和任务控制 46](#_Toc285366258)

[第35章 进程优先级和调度 47](#_Toc285366259)

[第36章 进程资源 48](#_Toc285366260)

[第37章 Daemon 49](#_Toc285366261)

[第38章 编写安全的特权程序 50](#_Toc285366262)

[第39章 能力 51](#_Toc285366263)

[第40章 登录会计 52](#_Toc285366264)

[第41章 共享库基础 53](#_Toc285366265)

[第42章 共享库高级特性 54](#_Toc285366266)

[第43章 进程间通信简介 55](#_Toc285366267)

[第44章 管道和FIFO 56](#_Toc285366268)

[第45章 System V IPC介绍 57](#_Toc285366269)

[第46章 System V消息队列 58](#_Toc285366270)

[第47章 System V信号量 59](#_Toc285366271)

[第48章 System V共享内存 60](#_Toc285366272)

[第49章 内存映射 61](#_Toc285366273)

[第50章 虚拟内存操作 62](#_Toc285366274)

[第51章 POSIX IPC介绍 63](#_Toc285366275)

[第52章 POSIX消息队列 64](#_Toc285366276)

[第53章 POSIX信号量 65](#_Toc285366277)

[第54章 POSIX共享内存 66](#_Toc285366278)

[第55章 文件锁 67](#_Toc285366279)

[第56章 Sockets：介绍 68](#_Toc285366280)

[第57章 Sockets：UNIX Domain 69](#_Toc285366281)

[第58章 Sockets：TCP/IP网络基础 70](#_Toc285366282)

[第59章 Sockets：Internet Domain 71](#_Toc285366283)

[第60章 Sockets：服务器设计 72](#_Toc285366284)

[第61章 Sockets：高级主题 73](#_Toc285366285)

[第62章 终端 74](#_Toc285366286)

[第63章 可选I/O模型 75](#_Toc285366287)

[第64章 伪终端 76](#_Toc285366288)

[附录A：跟踪系统调用 77](#_Toc285366289)

[附录B：解析命令行参数 78](#_Toc285366290)

[附录C：转换NULL指针 79](#_Toc285366291)

[附录D：内核配置 80](#_Toc285366292)

[附录E：更多信息来源 81](#_Toc285366293)

[附录F：部分习题解答 82](#_Toc285366294)

[参考书目 83](#_Toc285366295)

[索引 84](#_Toc285366296)

# 前言

## 主题

本书描述Linux编程接口——Linux（UNIX操作系统的一种免费实现）提供的系统调用、库函数、和其它底层接口。这些接口被直接或间接地使用在Linux上运行的每个程序中。它们允许应用程序完成各种任务：如文件I/O、创建删除文件和目录、创建新进程、执行程序、设置定时器、本机进程和线程间通信、通过网络连接的不同机器进程间通信等等。这些底层接口有时候也叫做系统编程接口。

尽管本书关注于Linux，但我也非常注意标准和可移植性问题，清晰地区分了Linux特有的接口、多数UNIX实现共有的特性、以及POSIX和Single UNIX Specification标准定义的特性。因此本书也提供了UNIX/POSIX编程接口的详尽描述，能够适用于编写UNIX系统应用或跨平台应用的程序员。

## 目标读者

本书主要面向以下读者：

* 为Linux、UNIX、或者其它遵循POSIX的系统开发应用的程序员和软件设计师；
* 在Linux、UNIX、或其它操作系统之间移植应用的程序员；
* Linux或UNIX系统编程课程的教师和高年级学生；
* 希望深入理解Linux/UNIX编程接口，以及系统软件是如何实现的系统管理员和“高级用户”。

我假设你拥有一定的编程经验，但不要求系统编程经验。我还假设你了解C编程语言，并且知道如何使用shell和常用的Linux或UNIX命令。如果你是Linux/UNIX的新手，你会发现第2章非常有用，我们以程序员的视角来讲述Linux和UNIX的基础概念。

## Linux和UNIX

本书原本可以纯粹地讲解标准UNIX（也就是POSIX）系统编程，因为UNIX和Linux的大多数特性都是相同的。不过虽然编写可移植程序是很好的目标，理解Linux对标准UNIX编程接口的扩展也是非常重要的。理由之一是Linux非常流行；其二是有时候为了性能、或使用标准UNIX没有的功能，我们不得不使用非标准的扩展（所有UNIX实现都提供类似的非标准扩展）。

因此本书在适用于标准UNIX的程序员时，还提供了Linux特定编程特性的详细描述。这些特性包括：

* epoll，获得文件I/O事件通知的机制；
* inotify，监控文件和目录改变的机制；
* 能力，授予进程一组超级用户能力的机制；
* 扩展属性；
* i-node标志；
* clone()系统调用；
* /proc文件系统
* Linux对文件I/O、信号、定时器、线程、共享库、进程间通信、和socket的特殊实现细节。

## 使用和组织

你至少可以按两种方式来使用本书：

* 作为Linux/UNIX编程接口的介绍手册。你可以从头到尾阅读本书。后续章节建立在之前章节的基础之上，我尽量避免依赖后续章节的情况。
* 作为Linux/UNIX编程接口的索引参考手册。详细的索引和频繁的交叉引用，允许你随机地阅读任何主题。

我把本书分为以下几部分：

1. 背景和概念：UNIX、C和Linux的历史；UNIX标准简介（第1章）；以程序员的视角介绍Linux和UNIX的基本概念（第2章）；Linux和UNIX系统编程的基本概念（第3章）。
2. 系统编程接口的基础特性：文件I/O（第4章和第5章）；进程（第6章）；内存分配（第7章）；用户和组（第8章）；进程凭证（第9章）；定时器（第10章）；系统限制和选项（第11章）；获取系统和进程信息（第12章）。
3. 系统编程接口的高级特性：文件I/O缓冲（第13章）；文件系统（第14章）；文件属性（第15章）；扩展属性（第16章）；访问控制列表（第17章）；目录和链接（第18章）；监控文件事件（第19章）；信号（第20章到第22章）；定时器（第23章）。
4. 进程、程序、和线程：进程创建、进程结束、监控子进程、执行程序（第24章到第28章）；POSIX线程（第29章到第33章）。
5. 进程和程序的高级主题：进程组、会话、任务控制（第34章）；进程优先级和调度（第35章）；进程资源（第36章）；daemon（第37章）；编写安全的特权程序（第38章）；能力（第39章）；登录会计（第40章）；共享库（第41章到第42章）。
6. 进程间通信（IPC）：IPC简介（第43章）；管道和FIFO（第44章）；System V IPC——消息队列、信号量、共享内存（第45章到第48章）；内存映射（第49章）；虚拟内存操作（第50章）；POSIX IPC——消息队列、信号量、共享内存（第51章到第54章）；文件锁（第55章）。
7. Socket和网络编程：IPC和socket网络编程（第56章到第61章）。
8. 高级I/O主题：终端（第62章）；可选I/O模型（第63章）；伪终端（第64章）。

## 例子程序

我用短小但完整的例子程序来阐述多数接口的使用方法，这些例子都被设计为很容易就能从命令行体验，来查看不同的系统调用和库函数如何工作。所以本书包含大量的示例代码——大概15000行C代码和shell会话日志。

尽管阅读和试验例子程序是不错的起点，掌握本书讨论的概念最有效的方法是编写代码，按你的想法修改例子程序，或者编写新程序都可以。

本书的所有源代码都可以在网站上下载。源代码包含许多书中没有的程序。这些程序的目的和细节在注释中都有相关描述。我提供了Makefile编译这些程序，以及一个README文件，给出了例子程序更多的细节信息。

源代码采用GNU Affero 通用公共授权版本3，可以自由分发和修改。源代码中也包含一份该协议的拷贝。

## 练习

多数章节都以一组练习结束，其中一些是要你按不同方式来试验例子程序，另外一些是该章讨论过的概念相关的问题，还有就是要求你来编写代码以巩固你对本书的理解。你可以在附录F找到部分练习的解答。

## 标准和可移植性

贯穿整本书，我都对可移植性问题特别地关注。你会发现很多相关标准的引用，特别是POSIX.1-2001和Single UNIX规范版本3（SUSv3）标准。同时你还将看到这些标准最新修订的细节改变，也就是POSIX.1-2008和SUSv4标准。（由于SUSv3是更大的修订版本，也是本书编写时最广泛有效的UNIX标准，本书讨论的标准大多是SUSv3，并标注出SUSv4不同的地方。除非我明确地提到，你可以假设我们对SUSv3规范的描述也适用于SUSv4）。

对于那些不是标准的特性，我会指出在不同UNIX实现间的差别。我还会突出那些Linux特定的特性，以及Linux与其它UNIX对系统调用和库函数实现上的细小差别。当某个特性我没有明确指出是Linux专有时，你也通常可以假设它在多数或所有UNIX上都有实现。

本书大多数例子程序我都在Solaris、FreeBSD、Mac OS X、Tru64 UNIX、和HP-UX上测试通过（除了那些Linux独有的特性）。为了提高代码在这些系统上的可移植性，本书网站上提供的某些例子程序有一些额外的代码。

## Linux内核和C库版本

本书主要关注Linux 2.6.x系列，这是本书写作时最广泛使用的内核版本。Linux 2.4的某些细节也会提到，我也会指出Linux 2.4和2.6的区别。当Linux 2.6.x系列出现了新特性时（例如2.6.34），我也会特别指出相应的内核版本号。

至于C库，本书则主要关注于GNU C库（glibc）版本2。当然，glibc 2.x系列版本存在差异时，我也会特别指出。

在本书即将印刷时，Linux内核刚刚发布了2.6.35版本，glibc则已经发布2.12版本。本书完全适用于这两个软件版本。Linux内核和glibc将来接口的变化，会在本书的网站上列出。

## 其它语言使用编程接口

尽管例子程序用C语言编写，你也可以在其它编程语言中使用本书讨论的接口——例如编译型语言C++、Pascal、Modula、Ada、FORTRAN、D；解释型语言Perl、Python、Ruby等。（Java则需要采用一种不同的方式JNI）。不同的语言要获取必要的常量定义和函数声明，需要使用不同的技术（C++除外），另外传递函数参数时可能也需要一点额外的工作。此外就没有太大的区别了，核心概念其实都是一样的。因此即使你使用其它的编程语言，你也会发现本书提供的信息是适用的。

## 关于作者

（略）

## 致谢

（略）

## 许可

电子工程学会和开放组织非常友好地许可我引用IEEE Std 1003.1, 2004版本，以及信息技术标准——可移植操作系统接口（POSIX），开放组织基本规范Issue6。完整的标准可以在<http://www.unix.org/version3/online.html>上在线查阅。

## 网站和例子程序源代码

你可以在<http://www.man7.org/tlpi>上找到关于本书更多的信息，包括勘误表和例子程序的源代码。

## 反馈

我非常欢迎代码bug报告、代码改进建议、以及代码可移植性的提高。同样我也欢迎本书的bug报告和改进建议。由于Linux编程接口总是在变化，我也非常高兴能获得关于本书将来版本的改进意见，包括新特性和变化特性。

Michael Timothy Kerrisk

Munich, Germany and Christchurch, New Zealand

August 2010

[mtk@man7.org](mailto:mtk@man7.org)

# 第1章 历史和标准

# 第2章 基础概念

# 第3章 系统编程概念

# 第4章 文件I/O：统一的I/O模型

# 第5章 文件I/O：更多细节

# 第6章 进程

# 第7章 内存分配

# 第8章 用户和组

# 第9章 进程凭证

# 第10章 时间

# 第11章 系统限制和选项

# 第12章 系统和进程信息

# 第13章 文件I/O缓冲

# 第14章 文件系统

# 第15章 文件属性

# 第16章 扩展属性

# 第17章 访问控制列表

# 第18章 目录和链接

# 第19章 监控文件事件

# 第20章 信号：基础概念

# 第21章 信号：信号处理器

# 第22章 信号：高级特性

# 第23章 定时器和睡眠

# 第24章 进程创建

# 第25章 进程结束

# 第26章 监控子进程

# 第27章 程序执行

# 第28章 进程创建和程序执行的更多细节

# 第29章 线程：介绍

# 第30章 线程：同步

# 第31章 线程：线程安全和线程存储

# 第32章 线程：线程取消

# 第33章 线程：更多细节

# 第34章 进程组、会话和任务控制

# 第35章 进程优先级和调度

# 第36章 进程资源

# 第37章 Daemon

# 第38章 编写安全的特权程序

# 第39章 能力

# 第40章 登录会计

# 第41章 共享库基础

# 第42章 共享库高级特性

# 第43章 进程间通信简介

# 第44章 管道和FIFO

# 第45章 System V IPC介绍

# 第46章 System V消息队列

# 第47章 System V信号量

# 第48章 System V共享内存

# 第49章 内存映射

# 第50章 虚拟内存操作

# 第51章 POSIX IPC介绍

# 第52章 POSIX消息队列

# 第53章 POSIX信号量

# 第54章 POSIX共享内存

# 第55章 文件锁

# 第56章 Sockets：介绍

# 第57章 Sockets：UNIX Domain

# 第58章 Sockets：TCP/IP网络基础

# 第59章 Sockets：Internet Domain

# 第60章 Sockets：服务器设计

# 第61章 Sockets：高级主题

# 第62章 终端

# 第63章 可选I/O模型

# 第64章 伪终端

# 附录A：跟踪系统调用

# 附录B：解析命令行参数

# 附录C：转换NULL指针

# 附录D：内核配置

# 附录E：更多信息来源

# 附录F：部分习题解答

# 参考书目

# 索引