

## <<Linux内核设计与实现>>

### 图书基本信息

书名：<<Linux内核设计与实现>>

13位ISBN编号：9787111338291

10位ISBN编号：7111338294

出版时间：2011-4-30

出版时间：机械工业出版社华章公司

作者：Robert Love

页数：335

译者：陈莉君,康华

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## &lt;&lt;Linux内核设计与实现&gt;&gt;

## 前言

随着Linux内核和Linux应用程序越来越成熟，越来越多的系统软件工程师涉足Linux开发和维护领域。他们中有些人纯粹是出于个人爱好，有些人是为Linux公司工作，有些人是为硬件厂商做开发，还有一些是为内部项目工作的。

但是所有人都必须直面一个问题：内核的学习曲线变得越来越长，也越来越陡峭。

系统规模不断扩大，复杂程度不断提高。

虽然现在的内核开发者对内核的掌握越发炉火纯青，但新手却无法跟上内核发展的步伐，长此以往将出现青黄不接的断层。

我认为这种新老鸿沟已经成为内核质量的一个隐患，而且问题将继续恶化。

所以那些真正关心内核的人已经开始致力于扩大内核开发群体。

解决上述问题的一个方法是尽量保证代码简洁：接口定义合理，代码风格一致，“一次做一件事，做到完美”等。

这也就是LinusTorvalds倡导的解决办法。

我提倡的解决办法是对代码慷慨地加上注释，即能够让读者立刻了解代码开发者意图的文字（识别意图和实现之间差异的工作称为调试。

如果意图不明确显然调试就难以进行）。

可是，即使有注解，也没办法清楚地展现内核的各个主要子系统的全景，说明它们到底要做什么。

那么，这些开发者又该从何下手呢？

由文字材料来说明这些在起步阶段就该理解的材料，其实是最合适的。

RobertLove的贡献就在于此，有经验的开发者可以通过本书全面了解内核子系统提供的服务，同时还可以了解这些服务是怎么实现的。

对不少人来说，这些知识就已经足够了：那些好奇的人，那些应用程序开发者，那些想对内核的设计品头论足一番的人，都有足够的谈资了。

但是学习本书同样可以作为那些有抱负的内核开发者更上一层楼的契机，可以帮他们更改内核代码以达到预定的目标。

我建议有抱负的开发者能够亲身实践：理解内核某部分的捷径就是对它做些修改，这样能为开发者揭示仅仅通过看内核代码无法看到的深层机理。

严肃认真的内核开发者应该加入开发邮件列表，不断和其他开发者交流。

这是内核开发者相互切磋和并肩前进的最好方法。

而Robert在书中对内核生活中至关重要的文化和技巧都做了精彩介绍。

请学习和欣赏Robert的书吧。

想必你也希望能精益求精，继续探索，成为内核开发社区中的一员，那么首先你要清楚的是：社区欢迎你。

我们评价和衡量一个人是根据他所作的贡献，当你投身于Linux时，你要明白：虽然你仅仅贡献了一小份力，但马上就会有数千万或上亿人受益。

这是我们的欢乐之源，也是我们的责任之本。

## <<Linux内核设计与实现>>

### 内容概要

《Linux内核设计与实现（原书第3版）》详细描述了Linux内核的设计与实现。内核代码的编写者、开发者以及程序开发人员都可以通过阅读本书受益，他们可以更好理解操作系统原理，并将其应用在自己的编码中以提高效率和生产率。

本书详细描述了Linux内核的主要子系统和特点，包括Linux内核的设计、实现和接口。从理论到实践涵盖了Linux内核的方方面面，可以满足读者的各种兴趣和需求。

作者Robert Love是一位Linux内核核心开发人员，他分享了在开发Linux 2.6内核过程中颇具价值的知识和经验。

本书的主题包括进程管理、进程调度、时间管理和定时器、系统调用接口、内存寻址、内存管理和页缓存、VFS、内核同步、移植性相关的问题以及调试技术。

同时本书也涵盖了Linux

2.6内核中颇具特色的内容，包括CFS调度程序、抢占式内核、块I/O层以及I/O调度程序。

《Linux内核设计与实现（原书第3版）》新增内容包括：

- 增加一章专门描述内核数据结构
- 详细描述中断处理程序和下半部机制
- 扩充虚拟内存和内存分配的内容
- 调试Linux内核的技巧
- 内核同步和锁机制的深度描述
- 提交内核补丁以及参与Linux内核社区的建设性建议

## <<Linux内核设计与实现>>

### 作者简介

Robert

Love是一位资深的开源社区达人，很早就开始使用Linux。

目前他是Google公司高级软件工程师，是开发Android移动平台内核的团队成员；他曾在Novell公司任职Linux桌面系统的首席架构师；他之前也曾是MontaVista和Ximain公司的内核开发工程师。

他参与的内核项目包括抢占式内核、进程调度器、内核事件层、通知机制、VM改进，以及设备驱动程序。

他是《Linux journal》杂志的编辑。

另外他还著有《Linux System Programming》和《Linux in a Nutshell》。

陈莉君，西安邮电学院教授，十多年来一直致力于推动Linux在中国的发展，多年从事Linux内核的教学和研究，并积极跟踪Linux内核的发展动向，对Linux内核版本的不断演化有着深刻的理解。

著译作品有《Linux操作系统原理与应用》、《Linux操作系统内核分析》、《深入分析Linux内核源代码》、《深入理解Linux内核》和《Linux内核编程》等。

## &lt;&lt;Linux内核设计与实现&gt;&gt;

## 书籍目录

译者序

序言

前言

作者简介

第1章 Linux内核简介1

1.1 Unix的历史1

1.2 追寻Linus足迹: Linux简介2

1.3 操作系统和内核简介3

1.4 Linux内核和传统Unix内核的比较5

1.5 Linux内核版本7

1.6 Linux内核开发者社区8

1.7 小结8

第2章 从内核出发10

2.1 获取内核源码10

2.1.1 使用Git10

2.1.1 安装内核源代码10

2.1.3 使用补丁11

2.2 内核源码树11

2.3 编译内核12

2.3.1 配置内核12

2.3.2 减少编译的垃圾信息14

2.3.3 衍生多个编译作业 14

2.3.4 安装新内核14

2.4 内核开发的特点15

2.4.1 无libc库抑或无标准头文件15

2.4.2 GNU C16

2.4.3 没有内存保护机制18

2.4.4 不要轻易在内核中使用浮点数18

2.4.5 容积小而固定的栈18

2.4.6 同步和并发18

2.4.7 可移植性的重要性19

2.5 小结19

第3章 进程管理20

3.1 进程20

3.2 进程描述符及任务结构 21

3.2.1 分配进程描述符22

3.2.2 进程描述符的存放23

3.2.3 进程状态23

3.2.4 设置当前进程状态25

3.2.5 进程上下文25

3.2.6 进程家族树25

3.3 进程创建26

3.3.1 写时拷贝27

3.3.2 fork()27

3.3.3 vfork()28

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 3.4 线程在Linux中的实现28
  - 3.4.1 创建线程29
  - 3.4.2 内核线程30
- 3.5 进程终结31
  - 3.5.1 删除进程描述符32
  - 3.5.2 孤儿进程造成的进退维谷32
- 3.6 小结34
- 第4章 进程调度35
  - 4.1 多任务35
  - 4.2 Linux 的进程调度36
  - 4.3 策略36
    - 4.3.1 I/O消耗型和处理器消耗型的进程36
    - 4.3.2 进程优先级37
    - 4.3.3 时间片38
    - 4.3.4 调度策略的活动38
  - 4.4 Linux调度算法39
    - 4.4.1 调度器类39
    - 4.4.2 Unix 系统中的进程调度40
    - 4.4.3 公平调度41
  - 4.5 Linux调度的实现42
    - 4.5.1 时间记账42
    - 4.5.2 进程选择44
    - 4.5.3 调度器入口48
    - 4.5.4 睡眠和唤醒49
  - 4.6 抢占和上下文切换51
    - 4.6.1 用户抢占53
    - 4.6.2 内核抢占53
  - 4.7 实时调度策略54
  - 4.8 与调度相关的系统调用54
    - 4.8.1 与调度策略和优先级相关的系统调用55
    - 4.8.2 与处理器绑定有关的系统调用55
    - 4.8.3 放弃处理器时间56
  - 4.9 小结56
- 第5章 系统调用57
  - 5.1 与内核通信57
  - 5.2 API、POSIX和C库57
  - 5.3 系统调用58
    - 5.3.1 系统调用号59
    - 5.3.2 系统调用的性能59
  - 5.4 系统调用处理程序60
    - 5.4.1 指定恰当的系统调用60
    - 5.4.2 参数传递60
  - 5.5 系统调用的实现61
    - 5.5.1 实现系统调用61
    - 5.5.2 参数验证62
  - 5.6 系统调用上下文64
    - 5.6.1 绑定一个系统调用的最后步骤65

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 5.6.2 从用户空间访问系统调用67
- 5.6.3 为什么不通过系统调用的方式实现68
- 5.7 小结68
- 第6章 内核数据结构69
  - 6.1 链表69
    - 6.1.1 单向链表和双向链表69
    - 6.1.2 环形链表70
    - 6.1.3 沿链表移动71
    - 6.1.4 Linux 内核中的实现71
    - 6.1.5 操作链表73
    - 6.1.6 遍历链表75
  - 6.2 队列78
    - 6.2.1 kfifo79
    - 6.2.2 创建队列79
    - 6.2.3 推入队列数据79
    - 6.2.4 摘取队列数据80
    - 6.2.5 获取队列长度80
    - 6.2.6 重置和撤销队列80
    - 6.2.7 队列使用举例 81
  - 6.3 映射 81
    - 6.3.1 初始化一个idr82
    - 6.3.2 分配一个新的UID82
    - 6.3.3 查找UID83
    - 6.3.4 删除UID84
    - 6.3.5 撤销idr84
  - 6.4 二叉树84
    - 6.4.1 二叉搜索树84
    - 6.4.2 自平衡二叉搜索树 85
  - 6.5 数据结构以及选择 87
  - 6.6 算法复杂度88
    - 6.6.1 算法88
    - 6.6.2 大o 符号88
    - 6.6.3 大 符号89
    - 6.6.4 时间复杂度89
  - 6.7 小结 90
- 第7章 中断和中断处理91
  - 7.1 中断91
  - 7.2 中断处理程序92
  - 7.3 上半部与下半部的对比93
  - 7.4 注册中断处理程序93
    - 7.4.1 中断处理程序标志94
    - 7.4.2 一个中断例子95
    - 7.4.3 释放中断处理程序95
  - 7.5 编写中断处理程序96
    - 7.5.1 共享的中断处理程序97
    - 7.5.2 中断处理程序实例97
  - 7.6 中断上下文99

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 7.7 中断处理机制的实现100
- 7.8 /proc/interrupts102
- 7.9 中断控制103
  - 7.9.1 禁止和激活中断103
  - 7.9.2 禁止指定中断线105
  - 7.9.3 中断系统的状态105
- 7.10 小结106
- 第8章 下半部和推后执行的工作107
  - 8.1 下半部107
    - 8.1.1 为什么要用下半部108
    - 8.1.2 下半部的环境108
  - 8.2 软中断110
    - 8.2.1 软中断的实现111
    - 8.2.2 使用软中断113
  - 8.3 tasklet114
    - 8.3.1 tasklet的实现114
    - 8.3.2 使用tasklet116
    - 8.3.3 老的BH机制119
  - 8.4 工作队列120
    - 8.4.1 工作队列的实现121
    - 8.4.2 使用工作队列124
    - 8.4.3 老的任务队列机制126
  - 8.5 下半部机制的选择127
  - 8.6 在下半部之间加锁128
  - 8.7 禁止下半部128
  - 8.8 小结129
- 第9章 内核同步介绍131
  - 9.1 临界区和竞争条件131
    - 9.1.1 为什么我们需要保护132
    - 9.1.2 单个变量133
  - 9.2 加锁134
    - 9.2.1 造成并发执行的原因135
    - 9.2.2 了解要保护些什么136
  - 9.3 死锁137
  - 9.4 争用和扩展性138
  - 9.5 小结140
- 第10章 内核同步方法141
  - 10.1 原子操作141
    - 10.1.1 原子整数操作142
    - 10.1.2 64位原子操作144
    - 10.1.3 原子位操作145
  - 10.2 自旋锁147
    - 10.2.1 自旋锁方法148
    - 10.2.2 其他针对自旋锁的操作149
    - 10.2.3 自旋锁和下半部150
  - 10.3 读-写自旋锁150
  - 10.4 信号量152

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 10.4.1 计数信号量和二值信号量153
- 10.4.2 创建和初始化信号量154
- 10.4.3 使用信号量154
- 10.5 读-写信号量155
- 10.6 互斥体156
  - 10.6.1 信号量和互斥体158
  - 10.6.2 自旋锁和互斥体158
- 10.7 完成变量158
- 10.8 BLK：大内核锁159
- 10.9 顺序锁160
- 10.10 禁止抢占161
- 10.11 顺序和屏障162
- 10.12 小结165
- 第11章 定时器和时间管理166
  - 11.1 内核中的时间概念166
  - 11.2 节拍率：HZ167
    - 11.2.1 理想的HZ值168
    - 11.2.2 高HZ的优势169
    - 11.2.3 高HZ的劣势169
  - 11.3 jiffies170
    - 11.3.1 jiffies的内部表示171
    - 11.3.2 jiffies的回绕172
    - 11.3.3 用户空间和HZ173
  - 11.4 硬时钟和定时器174
    - 11.4.1 实时时钟174
    - 11.4.2 系统定时器174
  - 11.5 时钟中断处理程序174
  - 11.6 实际时间176
  - 11.7 定时器178
    - 11.7.1 使用定时器178
    - 11.7.2 定时器竞争条件180
    - 11.7.3 实现定时器180
  - 11.8 延迟执行181
    - 11.8.1 忙等待181
    - 11.8.2 短延迟182
    - 11.8.3 schedule\_timeout()183
  - 11.9 小结185
- 第12章 内存管理186
  - 12.1 页186
  - 12.2 区187
  - 12.3 获得页189
    - 12.3.1 获得填充为0的页190
    - 12.3.2 释放页191
  - 12.4 kmalloc()191
    - 12.4.1 gfp\_mask标志192
    - 12.4.2 kfree()195
  - 12.5 vmalloc()196

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 12.6 slab层197
  - 12.6.1 slab层的设计198
  - 12.6.2 slab分配器的接口200
- 12.7 在栈上的静态分配203
  - 12.7.1 单页内核栈203
  - 12.7.2 在栈上光明正大地工作203
- 12.8 高端内存的映射204
  - 12.8.1 永久映射204
  - 12.8.2 临时映射204
- 12.9 每个CPU的分配205
- 12.10 新的每个CPU接口206
  - 12.10.1 编译时的每个CPU数据206
  - 12.10.2 运行时的每个CPU数据207
- 12.11 使用每个CPU数据的原因208
- 12.12 分配函数的选择209
- 12.13 小结209
- 第13章 虚拟文件系统210
  - 13.1 通用文件系统接口210
  - 13.2 文件系统抽象层211
  - 13.3 Unix文件系统212
  - 13.4 VFS 对象及其数据结构213
  - 13.5 超级块对象214
  - 13.6 超级块操作215
  - 13.7 索引节点对象217
  - 13.8 索引节点操作219
  - 13.9 目录项对象222
    - 13.9.1 目录项状态222
    - 13.9.2 目录项缓存223
  - 13.10 目录项操作224
  - 13.11 文件对象225
  - 13.12 文件操作226
  - 13.13 和文件系统相关的数据结构230
  - 13.14 和进程相关的数据结构232
  - 13.15 小结233
- 第14章 块I/O层234
  - 14.1 剖析一个块设备234
  - 14.2 缓冲区和缓冲区头235
  - 14.3 bio结构体237
    - 14.3.1 I/O向量238
    - 14.3.2 新老方法对比239
  - 14.4 请求队列240
  - 14.5 I/O调度程序240
    - 14.5.1 I/O调度程序的工作241
    - 14.5.2 Linus 电梯241
    - 14.5.3 最终期限I/O调度程序242
    - 14.5.4 预测I/O调度程序244
    - 14.5.5 完全公正的排队I/O调度程序244

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 14.5.6 空操作的I/O调度程序245
- 14.5.7 I/O调度程序的选择245
- 14.6 小结246
- 第15章 进程地址空间247
  - 15.1 地址空间247
  - 15.2 内存描述符248
    - 15.2.1 分配内存描述符249
    - 15.2.2 撤销内存描述符250
    - 15.2.3 mm\_struct 与内核线程250
  - 15.3 虚拟内存区域251
    - 15.3.1 VMA标志251
    - 15.3.2 VMA 操作253
    - 15.3.3 内存区域的树型结构和内存区域的链表结构254
    - 15.3.4 实际使用中的内存区域254
  - 15.4 操作内存区域255
    - 15.4.1 find\_vma()256
    - 15.4.2 find\_vma\_prev()257
    - 15.4.3 find\_vma\_intersection()257
  - 15.5 mmap()和do\_mmap(): 创建地址区间258
  - 15.6 mummap()和do\_mummap(): 删除地址区间259
  - 15.7 页表260
  - 15.8 小结261
- 第16章 页高速缓存和页回写262
  - 16.1 缓存手段262
    - 16.1.1 写缓存262
    - 16.1.2 缓存回收263
  - 16.2 Linux 页高速缓存264
    - 16.2.1 address\_space对象264
    - 16.2.2 address\_space 操作266
    - 16.2.3 基树267
    - 16.2.4 以前的页散列表268
  - 16.3 缓冲区高速缓存268
  - 16.4 flusher线程268
    - 16.4.1 膝上型计算机模式270
    - 16.4.2 历史上的bdflush、kupdated 和pdflush270
    - 16.4.3 避免拥塞的方法: 使用多线程271
  - 16.5 小结271
- 第17章 设备与模块273
  - 17.1 设备类型273
  - 17.2 模块274
    - 17.2.1 Hello, World274
    - 17.2.2 构建模块275
    - 17.2.3 安装模块277
    - 17.2.4 产生模块依赖性277
    - 17.2.5 载入模块278
    - 17.2.6 管理配置选项279
    - 17.2.7 模块参数280

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 17.2.8 导出符号表282
- 17.3 设备模型283
  - 17.3.1 kobject283
  - 17.3.2 ktype284
  - 17.3.3 kset285
  - 17.3.4 kobject、ktype和kset的相互关系285
  - 17.3.5 管理和操作kobject286
  - 17.3.6 引用计数287
- 17.4 sysfs288
  - 17.4.1 sysfs中添加和删除kobject 290
  - 17.4.2 向sysfs中添加文件291
  - 17.4.3 内核事件层293
- 17.5 小结294
- 第18章 调试295
  - 18.1 准备开始295
  - 18.2 内核中的bug296
  - 18.3 通过打印来调试296
    - 18.3.1 健壮性296
    - 18.3.2 日志等级297
    - 18.3.3 记录缓冲区298
    - 18.3.4 syslogd和klogd298
    - 18.3.5 从printf()到printk()的转换298
  - 18.4 oops298
    - 18.4.1 ksymoops300
    - 18.4.2 kallsyms300
  - 18.5 内核调试配置选项301
  - 18.6 引发bug并打印信息301
  - 18.7 神奇的系统请求键302
  - 18.8 内核调试器的传奇303
    - 18.8.1 gdb303
    - 18.8.2 kgdb304
  - 18.9 探测系统304
    - 18.9.1 用UID作为选择条件304
    - 18.9.2 使用条件变量305
    - 18.9.3 使用统计量305
    - 18.9.4 重复频率限制305
  - 18.10 用二分查找法找出引发罪恶的变更306
  - 18.11 使用Git进行二分搜索307
  - 18.12 当所有的努力都失败时：社区308
  - 18.13 小结308
- 第19章 可移植性309
  - 19.1 可移植操作系统309
  - 19.2 Linux移植史310
  - 19.3 字长和数据类型311
    - 19.3.1 不透明类型313
    - 19.3.2 指定数据类型314
    - 19.3.3 长度明确的类型314

## &lt;&lt;Linux内核设计与实现&gt;&gt;

- 19.3.4 char型的符号问题315
- 19.4 数据对齐315
  - 19.4.1 避免对齐引发的问题316
  - 19.4.2 非标准类型的对齐316
  - 19.4.3 结构体填补316
- 19.5 字节顺序318
- 19.6 时间319
- 19.7 页长度320
- 19.8 处理器排序320
- 19.9 SMP、内核抢占、高端内存321
- 19.10 小结321
- 第20章 补丁、开发和社区322
  - 20.1 社区322
  - 20.2 Linux编码风格322
    - 20.2.1 缩进323
    - 20.2.2 switch 语句323
    - 20.2.3 空格324
    - 20.2.4 花括号325
    - 20.2.5 每行代码的长度326
    - 20.2.6 命名规范326
    - 20.2.7 函数326
    - 20.2.8 注释326
    - 20.2.9 typedef327
    - 20.2.10 多用现成的东西328
    - 20.2.11 在源码中减少使用ifdef328
    - 20.2.12 结构初始化328
    - 20.2.13 代码的事后修正329
  - 20.3 管理系统329
  - 20.4 提交错误报告329
  - 20.5 补丁330
    - 20.5.1 创建补丁330
    - 20.5.2 用Git创建补丁331
    - 20.5.3 提交补丁331
  - 20.6 小结332
- 参考资料333

## &lt;&lt;Linux内核设计与实现&gt;&gt;

## 章节摘录

不知不觉涉足Linux内核已经十多个年头了，与其他有志（兴趣）于此的朋友一样，我们也经历了学习—实用—追踪—再学习的过程。

也就是说，我们也是从漫无边际到茫然无措，再到初窥门径，转而觉得心有戚戚焉这一路走下来的。其中甘苦，犹然在心。

Linux最为人称道的莫过于它的自由精神，所有源代码唾手可得。

侯捷先生云：“源码在前，了无秘密。

”是的，但是我们在面对它的时候，为什么却总是因为这种规模和层面所造就的陡峭学习曲线陷入困顿呢？

很多朋友就此倒下，纵然Linux世界繁花似锦，纵然内核天空无边广阔，但是，眼前的迷雾重重，心中的阴霾又怎能被阳光驱散呢？

纵有雄心壮志，拔剑四顾心茫然，脚下路在何方？

Linux内核入门是不容易，它之所以难学，在于庞大的规模和复杂的层面。

规模一大，就不易现出本来面目，浑然一体，自然不容易找到着手之处；层面一多，就会让人眼花缭乱，盘根错节，怎能让人提纲挈领？

“如果有这样一本书，既能提纲挈领，为我理顺思绪，指引方向，同时又能照顾小节，阐述细微，帮助我们更好更快地理解STL源码，那该有多好。

”孟岩先生如此说，虽然针对的是C++，但道出的是研习源码的人们共同的心声。

然而，Linux源码研究的方法却不大相同。

这还是由于规模和层面决定的。

比如说，在语言学习中，我们可以采取小步快跑的方法，通过一个个小程序和小尝试，就可以取得渐进的成果，就能从新技术中有所收获。

而掌握Linux呢？

如果没有对整体的把握，即使你对某个局部的算法、技术或是代码再熟悉，也无法将其融入实用。

其实，像内核这样的大规模的软件，正是编程技术施展身手的舞台（当然，目前的内核虽然包含了一些面向对象思想，但还不能让C++一展身手）。

那么，我们能不能做点什么，让Linux的内核学习过程更符合程序员的习惯呢？

RobertLove回答了这个问题。

RobertLove是一个狂热的内核爱好者，所以他的想法自然贴近程序员。

是的，我们注定要在对所有核心的子系统有了全面认识之后，才能开始自己的实践，但却完全可以舍弃细枝末节，将行李压缩到最小，自然可以轻装快走，迅速进入动手阶段。

因此，相对于DanielP.Bovet和MarcoCesati的内核巨著《UnderstandingtheLinuxKernel》，它少了五分细节；相对于实践经典《LinuxDeviceDrivers》，多了五分说理。

可以说，本书填补了Linux内核理论和实践之间的鸿沟，真可谓“一桥飞架南北，天堑变通途”。

就我们的经验，内核初学者（不是编程初学者）可以从本书着手，对内核各个核心子系统有个整体把握，包括它们提供什么样的服务，为什么要提供这样的服务，又是怎样实现的。

而且，本书还包含了Linux内核开发者在开发时需要用到的很多信息，包括调试技术、编程风格、注意事项等。

在消化本书的基础上，如果你侧重于了解内核，可以进一步研究《UnderstandingtheLinuxKernel》和源代码本身；如果你侧重于实际编程，可以研读《LinuxDeviceDrivers》，直接开始动手工作；如果你想有一个轻松的内核学习和实践环节，请访问我们的网站[www.kerneltravel.net](http://www.kerneltravel.net)。

Linux内核是一艘永不停息的轮船，它将驶向何方我们并不知晓，但在这些变化的背后，总有一些原理是恒定不变的，总有一些变化是我们想知晓的，比如调度程序的大幅度改进，内核性能的不不断提升，本书第3版虽然针对的是较新的2.6.34Linux内核版本，但在旧版本上积累的知识 and 经验依然有效，而新增内容将使读者在应对变化了的内核代码时更加从容。

感谢牛涛和武特，他们在第2版和第3版差异的校对中花费了大量精力。

## &lt;&lt;Linux内核设计与实现&gt;&gt;

感谢素不相识的网友ChengRenquan，他主动承担了其中一章的修订。

还要感谢苏锦绣、黄伟、王泽宇、赵格娟、刘周平、周永飞、曹江峰、陈白虎和孟阿龙，他们参与了后期的校对和查错补漏。

最后，特别感谢我的合作者康华，从十年前一块分析Linux内核代码到今天，他对技术孜孜不倦的追求不但在业界赢得声誉，也使我们在翻译过程中所遇到的技术难点和晦涩句子被一一迎刃而解。

感谢合作者张波，他流畅有趣的文笔让本书少了份枯燥，多了份趣味。

陈莉君2010年10月前言在我刚开始有把自己的内核开发经验集结成册，撰写一本书的念头时，我其实也觉得有点头绪繁多，不知道该从何下手。

我实在不想落入传统内核书籍的窠臼，照猫画虎地再写这么一本。

不错，前人著述备矣，但我终究是要写出点儿与众不同的东西来，我的书该如何定位，说实话，这确实让人颇费思量。

后来，灵感终于浮现出来，我意识到自己可以从一个全新的视角看待这个主题。

开发内核是我的工作，同时也是我的嗜好，内核就是我的挚爱。

这些年来，我不断搜集与内核有关的奇闻轶事，不断积攒关键的开发诀窍，依靠这些日积月累的材料，我可以写一本关于开发内该做什么—更重要的是—不该做什么的书籍。

从本质上说，这本书仍旧是描述Linux内核是如何设计和实现的，但是写法却另辟蹊径，所提供的信息更倾向于实用。

通过《Linux内核设计与实现（原书第3版）》，你就可以做一些内核开发的工作了—并且是使用正确的方法去做。

我是一个注重实效的人，因此，这是一本实践的书，它应当有趣、易读且有用。

我希望读者可以从这本书中领略到更多Linux内核的精妙之处（写出来的和没写出来的），也希望读者敢于从阅读本书和读内核代码开始跨越到开始尝试开发可用、可靠且清晰的内核代码。

当然如果你仅仅是兴致所至，读书自娱，那也希望你能从中找到乐趣。

从第1版到现在，又过了一段时间，我们再次回到本书，修补遗憾。

本版比第1版和第2版内容更丰富：修订、补充并增加了新的内容和章节，使其更加完善。

本版融合了第2版以来内核的各种变化。

更值得一提的是，Linux内核联盟做出决定，近期内不进行2.7版内核的开发，于是，内核开发者打算继续开发并稳定2.6版。

这个决定意味深长，而本书从中的最大受益就是在2.6版上可以稳定相当长时间。

随着内核的成熟，内核“快照”才有机会能维持得更久远一些。

本书可作为内核开发的规范文档，既认识内核的过去，也着眼于内核的未来。

使用这本书开发Linux内核不需要天赋异秉，不需要有什么魔法，连Unix开发者普遍长着的络腮胡子都不一定要有。

内核虽然有一些有趣并且独特的规则和要求，但是它和其他大型软件项目相比，并没有太大差别。

像所有的大型软件开发一样，要学的东西确实不少，但是不同之处在于数量上的积累，而非本质上的区别。

认真阅读源码非常有必要，Linux系统代码的开放性其实是弥足珍贵的，不要无动于衷地将它搁置一边，浪费了大好资源。

实际上就是读了代码还远远不够呢，你应该钻研并尝试着动手改动一些代码。

寻找一个bug然后去，改进你的硬件设备的驱动程序。

增加新功能，即使看起来微不足道，寻找痛痒之处并解决。

只有动手写代码才能真正融会贯通。

内核版本《Linux内核设计与实现（原书第3版）》基于Linux2.6内核系列。

它并不涵盖早期的版本，当然也有一些例外。

比如，我们会讨论2.4系列内核中的一些子系统是如何实现的，这是因为简单的实现有助于传授知识。

特别说明的是，本书介绍的是最新的Linux2.6.34内核版本。

尽管内核总在不断更新，任何努力也难以捕获这样一只永不停息的猛兽，但是本书力图适合于新旧内

## &lt;&lt;Linux内核设计与实现&gt;&gt;

核的开发者和用户。

虽然本书讨论的是2.6.34内核，但我也确保了它同样适用于2.6.32内核。

后一个版本往往被各个Linux发行版本奉为“企业版”内核，所以我们可以各种产品线上见到其身影。

该版本确实已经开发了数年（类似的“长线”版本还有2.6.9、2.6.18和2.6.27等）。

读者范围《Linux内核设计与实现（原书第3版）》是写给那些有志于理解Linux内核的软件开发者的。

本书并不逐行逐字地注解内核源代码，也不是指导开发驱动程序或是内核API的参考手册（如果存在标准的内核API的话）。

本书的初衷是提供足够多的关于Linux内核设计和实现的信息，希望读过本书的程序员能够拥有较为完备的知识，可以真正开始开发内核代码。

无论开发内核是为了兴趣还是为了赚钱，我都希望能够带领读者快速走进Linux内核世界。

本书不但介绍了理论而且也讨论了具体应用，可以满足不同读者的需要。

全书紧紧围绕着理论联系实践，并非一味强调理论或是实践。

无论你研究Linux内核的动机是什么，我都希望这本书都能将内核的设计和实现分析清楚，起到抛砖引玉的作用。

因此，本书覆盖了从核心内核系统的应用到内核设计与实现等各方面的内容。

我认为这点很重要，值得花工夫讨论。

例如，第8章讨论的是所谓的下半部机制。

本章分别讨论了内核下半部机制的设计和实现（核心内核开发者或者学者会感兴趣），随即便介绍了如何使用内核提供的接口实现你自己的下半部（这对设备驱动开发者可能很有用处）。

其实，我认为上述两部分内容是相得益彰的，虽然核心内核开发者主要关注的问题是内核内部如何工作，但是也应该清楚如何使用接口；同样，如果设备驱动开发者了解了接口背后的实现机制，自然也会受益匪浅。

这好比学习某些库的API函数与研究该库的具体实现。

初看，好像应用程序开发者仅仅需要理解API——我们被灌输的思想是，应该像看待黑盒子一样看待接口。

另外，库的开发者也只关心库的设计与实现，但是我认为双方都应该花时间相互学习。

能深刻了解操作系统本质的应用程序开发者无疑可以更好地利用它。

同样，库开发者也决不应该脱离基于此库的应用程序，埋头开发。

因此，我既讨论了内核子系统的设计，也讨论了它的用法，希望本书能对核心开发者和应用开发者都有用。

我假设读者已经掌握了C语言，而且对Linux比较熟悉。

如果读者还具有与操作系统设计相关的经验和其他计算机科学的概念就更好了。

当然，我也会尽可能多地解释这些概念，但如果你仍然不能理解这些知识的话，请看本书最后参考资料中给出的一些关于操作系统设计方面的经典书籍。

本书很适合在大学中作为介绍操作系统的辅助教材，与介绍操作系统理论的书相搭配。

对于大学高年级课程或者研究生课程来说，可直接使用本书作为教材。

第3版致谢与其他作者一样，我决非是一个人躲在山洞里孤苦地写出这本书来的（那也是一件美差，因为有熊相伴）。

我能最终完成本书原稿是与无数建议和关怀分不开的。

仅仅一页纸无法容纳我的感激，但我还是要衷心地感谢所有给予我鼓励、给予我知识和给予我灵感的朋友和同事。

首先我要对为《Linux内核设计与实现（原书第3版）》付出辛勤劳作的所有编辑表示感谢，尤其要感谢我的组稿编辑Mark Taber，他为这一版的出版从头到尾倾注了许多心血。

还要特别感谢业务开发编辑Michael Thurston和项目组织编辑Tonya Simpsom。

本版的技术编辑Robert P. J. Day也是我需要倍加感谢的人，他独到的洞察力和准确的校对使书稿质量大大提高。

## <<Linux内核设计与实现>>

尽管他的工作称得上完美，但如果书中仍留有错误，责任由我承担。

需要十分感谢的还有AdamBelay、ZackBrown、MartinPool以及ChrisRivera，他们对第1版和第2版所做的一切努力我依然记忆犹新。

许多内核开发者为我提供了大力支持，回答了许多问题，还有那些撰写代码的人——本书正是由于有了这些代码，才有了存在的意义。

他们是：AndreaArcangeli、AlanCox、GregKroah-Hartman、DaveMiller、PatrickMochel、AndrewMorton、NickPiggin和LinusTorvalds。

我要大力感谢我的同事们。

他们的创造力和智慧无与伦比，能与他们一起工作其乐无穷。

因为篇幅原因，请谅解我不能列出所有人的名字。

但不得不提的是：AlanBlount、JayCrim、ChrisDanis、ChrisDiBona、EricFlatt、MikeLockwood、SanMehat、BrianRogan、BrianSwetland、JonTrowbridge和SteveVinter，感谢他们给予我的支持、知识和友谊！

还有许多值得尊敬和热爱的人，他们是：PaulAmici、MikeyBabbitt、KeithBarbag、JacobBerkman、NatFriedman、DustinHall、JoyceHawkins、Migueldelcaza、JimmyKrehl、DorisLove、LindaLove、BretteLuck、RandyOttowd、SalRibaud和他了不起的妈妈ChrisRivera、CarolynRodon、JoeyShaw、SarahStewart、JeremyVanDoren和他的家人、LuisVilla、SteveWeisberg和他的家人以及HelenWhisnant等。

。

## <<Linux内核设计与实现>>

### 媒体关注与评论

能够把linux内核的内容在300多页内叙述一遍，本身就是一件高难度的事情。但《Linux内核设计与实现》确实做到了。

《Linux内核设计与实现》很少涉及具体实现，而是把握思想，讲解算法，读者可以学习到linux内核的知识，而不用纠缠于具体细节。

—豆瓣读者Googol，《Linux内核设计与实现》很适合系统学习了OS理论之后直接看代码详解又觉得暂且还不够功力的读者，它可以带你由理论学习阶段逐渐过渡到实践阶段。

这本书对Linux内核内容的范围和深度把握得恰到好处。

——豆瓣读者 纳兰经若

## <<Linux内核设计与实现>>

### 编辑推荐

《Linux内核设计与实现(原书第3版)》编辑推荐：畅销图书新版《Linux内核设计与实现》第3版翻译版、影印版同步上市，《Linux内核设计与实现(原书第3版)》详细描述Linux内核的主要子系统和特点，《Linux内核设计与实现(原书第3版)》涵盖Linux内核从理论到实践的方方面面。

## <<Linux内核设计与实现>>

### 版权说明

本站所提供下载的PDF图书仅提供预览和简介, 请支持正版图书。

更多资源请访问:<http://www.tushu007.com>