

LINUX DEVICE DRIVERS

第三版

# LINUX 设备驱动程序



O'REILLY®  
中国电力出版社

JONATHAN CORBET, ALESSANDRO RUBINI  
& GREG KROAH-HARTMAN 著  
魏永明 耿岳 钟书毅 译

# LINUX

# 设备驱动程序

---

第三版

*Jonathan Corbet,*  
*Alessandro Rubini &*  
*Greg Kroah-Hartman* 著  
魏永明 耿岳 钟书毅 译

O'REILLY®

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

O'Reilly Media, 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

Linux 设备驱动程序: 第 3 版 / (美) 科波特 (Corbet, J.) 等著; 魏永明, 耿岳, 钟书毅译. - 北京: 中国电力出版社, 2005.11

书名原文: Linux Device Drivers, Third Edition

ISBN 7-5083-3863-4

I. L... II. ①科 ... ②魏 ... ③耿 ... ④钟 ... III. Linux 操作系统 IV. TP316.89

中国版本图书馆 CIP 数据核字 (2005) 第 116363 号

北京市版权局著作权合同登记

图字: 01-2005-5621 号

©2005 by O'Reilly Media, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and China Electric Power Press, 2005. Authorized translation of the English edition, 2005 O'Reilly Media, 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 Media, Inc. 出版 2005。

简体中文版由中国电力出版社出版 2005。英文原版的翻译得到 O'Reilly Media, 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 Media, 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 / Linux 设备驱动程序 (第三版)

书 号 / ISBN 7-5083-3863-4

责任编辑 / 陈维宁 牛贵华

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 × 980 毫米 16 开本 36.75 印张 597 千字

版 次 / 2006 年 1 月第一版 2006 年 1 月第一次印刷

印 数 / 0001-4000 册

定 价 / 69.00 元 (册)

# O'Reilly Media, Inc. 介绍

为了满足读者对网络和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 Media, Inc. 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 Media, Inc. 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 User'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 Media, Inc. 一直处于 Internet 发展的最前沿。

许多书店的反馈表明，O'Reilly Media, Inc. 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 Media, Inc. 具有深厚的计算机专业背景，这使得 O'Reilly Media, Inc. 形成了一个非常不同于其他出版商的出版方针。O'Reilly Media, Inc. 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 Media, Inc. 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 Media, Inc. 依靠他们及时地推出图书。因为 O'Reilly Media, Inc. 紧密地与计算机业界联系着，所以 O'Reilly Media, Inc. 知道市场上真正需要什么图书。

# 目录

前言 .....	1
<b>第一章 设备驱动程序简介 .....</b>	<b>9</b>
设备驱动程序的作用 .....	10
内核功能划分 .....	12
设备和模块的分类 .....	14
安全问题 .....	15
版本编号 .....	17
许可证条款 .....	18
加入内核开发社团 .....	19
本书概要 .....	19
<b>第二章 构造和运行模块 .....</b>	<b>21</b>
设置测试系统 .....	21
Hello World 模块 .....	22
核心模块与应用程序的对比 .....	24
编译和装载 .....	28
内核符号表 .....	33
预备知识 .....	35
初始化和关闭 .....	36

模块参数 .....	40
在用户空间编写驱动程序 .....	42
快速参考 .....	44
<b>第三章 字符设备驱动程序 .....</b>	<b>46</b>
scull 的设计 .....	46
主设备号和次设备号 .....	47
一些重要的数据结构 .....	53
字符设备的注册 .....	59
open 和 release .....	62
scull 的内存使用 .....	64
read 和 write .....	67
试试新设备 .....	74
快速参考 .....	74
<b>第四章 调试技术 .....</b>	<b>76</b>
内核中的调试支持 .....	76
通过打印调试 .....	78
通过查询调试 .....	85
通过监视调试 .....	94
调试系统故障 .....	96
调试器和相关工具 .....	102
<b>第五章 并发和竞态 .....</b>	<b>109</b>
scull 的缺陷 .....	109
并发及其管理 .....	110
信号量和互斥体 .....	111
completion .....	116
自旋锁 .....	118
锁陷阱 .....	123

---

除了锁之外的办法 .....	125
快速参考 .....	132
<b>第六章 高级字符驱动程序操作 .....</b>	<b>137</b>
ioctl .....	137
阻塞型 I/O .....	149
poll 和 select .....	163
异步通知 .....	168
定位设备 .....	172
设备文件的访问控制 .....	173
快速参考 .....	179
<b>第七章 时间、延迟及延缓操作 .....</b>	<b>183</b>
度量时间差 .....	183
获取当前时间 .....	188
延迟执行 .....	190
内核定时器 .....	196
tasklet .....	202
工作队列 .....	204
快速参考 .....	208
<b>第八章 分配内存 .....</b>	<b>213</b>
kmalloc 函数的内幕 .....	213
后备高速缓存 .....	217
get_free_page 和相关函数 .....	221
vmalloc 及其辅助函数 .....	225
per-CPU 变量 .....	228
获取大的缓冲区 .....	230
快速参考 .....	231

<b>第九章 与硬件通信</b> .....	<b>235</b>
I/O 端口和 I/O 内存 .....	235
使用 I/O 端口 .....	239
I/O 端口示例 .....	245
使用 I/O 内存 .....	248
快速参考 .....	254
<b>第十章 中断处理</b> .....	<b>258</b>
准备并口 .....	259
安装中断处理例程 .....	259
实现中断处理例程 .....	269
上半部和下半部 .....	274
中断共享 .....	278
中断驱动的 I/O .....	281
快速参考 .....	285
<b>第十一章 内核的数据类型</b> .....	<b>287</b>
使用标准 C 语言类型 .....	287
为数据项分配确定的空间大小 .....	289
接口特定的类型 .....	289
其他有关移植性的问题 .....	291
链表 .....	294
快速参考 .....	298
<b>第十二章 PCI 驱动程序</b> .....	<b>300</b>
PCI 接口 .....	300
ISA 回顾 .....	317
PC/104 和 PC/104+ .....	319
其他的 PC 总线 .....	319
SBus .....	320



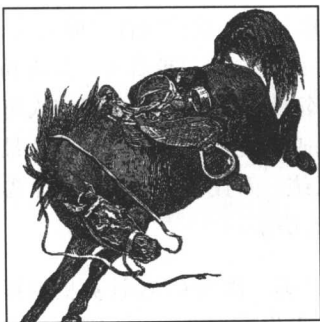
---

NuBus .....	321
外部总线 .....	321
快速参考 .....	322
<b>第十三章 USB 驱动程序 .....</b>	<b>324</b>
USB 设备基础 .....	326
USB 和 Sysfs .....	329
USB urb .....	331
编写 USB 驱动程序 .....	342
不使用 urb 的 USB 传输 .....	352
快速参考 .....	356
<b>第十四章 Linux 设备模型 .....</b>	<b>359</b>
kobject、kset 和子系统 .....	361
低层 sysfs 操作 .....	368
热插拔事件的产生 .....	372
总线、设备和驱动程序 .....	374
类 .....	384
各环节的整合 .....	388
热插拔 .....	394
处理固件 .....	401
快速索引 .....	403
<b>第十五章 内存映射和 DMA .....</b>	<b>408</b>
Linux 的内存管理 .....	408
mmap 设备操作 .....	418
执行直接 I/O 访问 .....	429
直接内存访问 .....	435
快速参考 .....	453

<b>第十六章 块设备驱动程序 .....</b>	<b>458</b>
注册 .....	459
块设备操作 .....	464
请求处理 .....	468
其他一些细节 .....	484
快速参考 .....	487
<b>第十七章 网络驱动程序 .....</b>	<b>491</b>
snll 设计 .....	492
连接到内核 .....	495
net_device 结构细节 .....	499
打开和关闭 .....	508
数据包传输 .....	510
数据包的接收 .....	514
中断处理例程 .....	516
不使用接收中断 .....	518
链路状态的改变 .....	521
套接字缓冲区 .....	521
MAC 地址解析 .....	525
定制 ioctl 命令 .....	527
统计信息 .....	528
组播 .....	529
其他知识点详解 .....	533
快速参考 .....	534
<b>第十八章 TTY 驱动程序 .....</b>	<b>538</b>
小型 TTY 驱动程序 .....	540
tty_driver 函数指针 .....	545
TTY 线路设置 .....	551
ioctls .....	555

---

proc 和 sysfs 对 TTY 设备的处理 .....	557
tty_driver 结构详解 .....	558
tty_operations 结构详解 .....	560
tty_struct 结构详解 .....	562
快速参考 .....	564
<b>参考书目 .....</b>	<b>567</b>



# 前言

顾名思义，本书是讲述如何编写 Linux 设备驱动程序的。面对层出不穷的新硬件产品，必须有人不断编写新的驱动程序以便让这些设备能够在 Linux 下正常工作，从这个意义上讲，讲述驱动程序的编写本身就是一件非常有意义的工作。但本书也涉及到 Linux 内核的工作原理，同时还讲述如何根据自己的需要和兴趣来定制 Linux 内核。Linux 是一个开放的系统，我们希望借助本书使它能够更好地开放，从而能够吸引更多的开发人员。

本书是《Linux 设备驱动程序》的第三版。自本书第一版发行以来，内核已经发生了巨大变化，我们必须努力让本书跟上内核的发展步伐。在这一版本中，我们尽可能完整地描述了 2.6.10 内核。这次，我们决定略去针对先前内核版本的向后兼容性描述，这是因为从 2.4 以来内核发生的改变实在太大了，而针对 2.4 内核的接口描述在本书第二版（可免费获得）中有很好的阐述。

这一版本包括了一些 2.6 内核相关的新内容。关于锁和并发性的内容得到了进一步充实，而且单独成章。我们还详细描述了 2.6 内核中新引入的 Linux 设备模型。我们用新的章节来描述 USB 总线和串行驱动程序子系统；同时，讲述 PCI 的那一章也得到了加强。本书其余部分类似之前的版本，但几乎每一章都彻底更新过了。

我们希望读者能够从本书的学习中获得乐趣，就像我们自己对编写本书的过程中获得乐趣一样。

## Jon 的介绍

这个版本出版的时候，恰好我在 Linux 界工作了 12 年，更惊人的是恰好是我在计算机领域工作的第 25 个年头。1980 年时，计算机领域就已经是个快速发展的领域，然而此后又加速不少。让本书保持更新状态面临越来越大的挑战；Linux 内核黑客在不停地增强他们的代码，但很少有耐心去关心文档是否跟得上步伐。

在市场上Linux保持着成功,但更重要的是Linux赢得了全球开发人员的关注。很明显, Linux的成功证明了其优秀的技术质量以及自由软件的大量好处。但在我看来,其成功的真正关键在于如下事实:Linux将快乐重新带回到计算机领域。利用Linux,任何人可以了解系统并以任何可能的方式贡献自己的代码,当然,代码在技术上的优势是其中最有价值的。Linux不仅为我们提供了一个顶级质量的操作系统,而且也为我们提供了参与到其未来开发过程的机会,我们完全可以从中得到无尽的快乐。

在计算机领域的25年中,我曾经有过许多有意思的经历,从第一次为Cray计算机编程(用Fortran语言在纸带上打孔),到亲历迷你计算机和Unix工作站的变迁,一直到当前微处理器占支配地位的时代。我还没有看到哪个领域可以让人如此着迷并因此开心快乐,也从未有过像现在这样能够完全控制我们的工具及其发展的时候。很明显, Linux和自由软件是这些变化背后的驱动力。

我希望本书能将这种快乐和机会带给新的Linux开发人员。不管你的兴趣在内核级还是用户空间,我都希望本书是一本有用而且有趣的指南,它能够帮助读者发现内核是如何和硬件一同工作的。我希望本书能帮助和启发读者利用自己的编辑器让我们共享的、自由的操作系统更加美好!Linux已经走过了很长的路,然而此刻也正是起点,观察并参与其中将为你带来更大的乐趣。

## Alessandro 的介绍

我一直喜欢玩电脑,就因为通过电脑我可以控制外部的硬件。我曾为Apple II和ZX Spectrum系统焊接我自己的设备,之后,有了大学中学到的Unix和自由软件专业知识,通过在新的386系统上安装了GNU/Linux并再次玩起了自己的电路板,我逃离了DOS陷阱。

那时Linux社区还非常小,也没有太多的文档来描述如何编写驱动程序,于是我开始为《Linux Journal》撰稿。这就是事情的开端:当我发现我自己不喜欢撰写论文后,我离开了大学,并和O'Reilly签订了本书第一版的编写合同。

这是1996年的事情,已经过去好多年了。

现在,计算机世界已经大不相同了:自由软件已经成为一种可行方案,不论在技术上还是在政治上,然而在这两个领域仍然有许多工作要做。我希望本书能够促进如下两个目标的实现:传播技术知识并提高对传播知识必要性的认同。这也是本书第一版被大众广泛接受以来第二版的两位作者在编辑和出版商的支持下转向自由许可证的原因。我坚信这是正确的知识传播途径,并且有利于和认同这种观点的其他人合作。

在嵌入式领域中所发生的一切令我兴奋，我希望本书能够为Linux的应用推波助澜；然而，在今天这个时代，思想的变化尤其迅速，为第四版作计划的时间已经到来，我们也正在寻求第四位作者的帮助。

## Greg 的介绍

从为了编写一个真实的Linux驱动程序而拿起《Linux设备驱动程序》第一版到现在，已经过很长一段时间。本书第一版帮助我理解了Linux操作系统的内部细节，而在此之前，我使用该操作系统有很长的时间，但几乎没有时间来研究内核细节。有了第一版中获得的知识，加上阅读内核中其他程序员的代码，我的第一个充满缺陷、非SMP安全的驱动程序被内核社区接受，并加入到了内核代码的主分支中。尽管在五分钟之后我就收到了我的第一个缺陷报告，但自此我就被希望尽我所能使Linux操作系统成为最好的欲望吸引住了。

我非常荣幸能够为本书贡献一些东西。我希望本书能够帮助其他人掌握与内核相关的一些细节，你会发现驱动程序的开发并不像想像的那么可怕或吓人，当然，我也希望本书能够鼓励其他人加入或者帮助这个大的集体，以便让这个操作系统可以在每一个计算机平台上运行，并支持每一种可获得的设备。开发过程充满着乐趣，加入社区非常值得，因为每个人都能从努力付出中获得好处。

现在，让我们一起寻找这个版本的缺陷，修改API以便让它们更好地工作，或者更加简单而便于每个人理解，或者增加新的特性。来吧，参与其中我们也将得到他人的帮助。

## 本书的读者对象

本书对那些希望编写计算机设备驱动程序的人员、或者那些要解决Linux机器内部问题的程序员来讲，将是非常有帮助的。请注意，“Linux机器”是一个比“运行Linux的PC”更为宽泛的概念，因为Linux现在能够支持许多不同的硬件平台，而内核编程不再绑定到某个特定的平台。我们希望本书能够成为那些想成为内核黑客但却不知如何下手的人们的良好起点。

在技术方面，本书为理解内核内幕以及理解一些Linux开发者所做出的设计决策支了一招。尽管本书的主要目的是告诉读者如何编写设备驱动程序，但同时也给出了内核实现方面的概览。

尽管真正的黑客能够从正式的内核源代码中找到所有必要的信息，但通常来讲，编写好

的书籍能够更好地帮助读者提高编程技巧。读者将要看到的文字来自对内核源代码的仔细分析，我们希望我们所付出的努力是值得的。

Linux 发烧友可从本书找到深入内核代码的足够精神食粮；通过本书的学习，将有能力加入到为某个新功能或性能增强不停工作的开发小组当中。本书并没有涵盖 Linux 内核的全部，但是作为 Linux 设备驱动程序开发人员，你需要的是了解如何与许多的内核子系统一起工作。因此，本书对内核编程作了一个一般性的介绍。Linux 仍然在不断改进和发展，因此新程序员始终有机会加入到这一 Linux 的开发大军中。

另一方面，如果你只是为了给自己的设备编写一个驱动程序，而不想过多了解内核的内幕信息，本书内容则足够模块化以满足你的需求。如果你不想深入到细节当中，则可以跳过大部分的技术章节，而直接阅读可由设备驱动程序使用的、能够和内核的其他部分无缝结合的标准 API。

## 内容的组织

本书内容由简到难，并划分为两大部分。第一部分（第一章到第十一章）首先讲述了如何编写内核模块，然后讲述了编写功能完备的字符设备驱动程序所涉及的各个编程主题。每一章讲述一个特定问题，并在每章结尾包含一个“快速小结”，该“快速小结”可在实际开发中作为参考使用。

在本书第一部分中，内容从面向软件的概念过渡到硬件相关的概念。这种组织方法意味着你能够尽可能不在机器中插入任何外部硬件而测试示例代码。每章都包含有源代码，并给出了能够在任意一台 Linux 计算机上运行的示例驱动程序。但是，在第十章和第十一章中，我们需要读者在并口上连接一些电线，以便测试硬件处理代码，当然，这一要求对任何人来讲都是可以做到的。

本书的第二部分（第十二章到第十八章）讲述了块设备驱动程序和网络接口，并深入讨论了一些更加高级的内容，比如虚拟内存子系统和 PCI、USB 总线等。许多驱动程序作者可能不需要这些内容，但我们鼓励你阅读这些章节。尽管对某个特定的项目来说你并不需要了解这些知识，但第二部分的许多内容和了解 Linux 内核的工作原理一样重要。

## 背景信息

为了更好地利用本书，我们希望读者熟悉 C 语言编程。因为我们经常会提到 Unix 系统调用、命令和管道，因此也需要读者拥有 Unix 的使用经验。

在硬件级，不需要读者有任何预先的经验就可以理解本书内容，当然，一些一般性的概念是必须清楚的。本书内容并不基于某个特定的 PC 硬件，我们在提到某个特定的硬件时会提供给读者所有必要的信息。

构造内核需要一些自由软件工具，而且经常要求使用这些工具的特定版本。太老的工具可能缺少一些必要的特性，而太新的工具又可能会偶尔生成不能正常工作的内核。通常而言，当前流行的 Linux 发行版所提供的工具能够很好地工作。不同的内核版本对工具的版本需求不同，这时，你可以参考内核源代码树中的 *Documentation/Changes* 文件。

## 在线版本和条款

本书作者已经选择本书在 Creative Commons “Attribution-ShareAlike” 许可证版本 2.0 的保护下免费获得：

<http://www.oreilly.com/catalog/linuxdrive3>

## 排版约定

下面是本书中用到的一些排版约定：

### 斜体 (*Italic*)

用于文件和目录的名称、程序和命令的名称、命令行选项、URL 以及新的术语

### 等宽字体 (`Constant Width`)

用于在示例中显示文件内容或者命令的输出，还用于正文中出现的 C 代码或者其他字符串

### 等宽斜体 (`Constant Width Italic`)

用于可变选项、关键词或者需要用户用实际值替换的文字

### 等宽黑体 (`Constant Width Bold`)

用于示例中需要用户照原文键入的命令或者其他文字

## 代码示例的使用

本书用于帮助读者完成自己的任务。通常来讲，读者可以在自己的程序和文档中使用本书中的代码。示例代码采用 BSD/GPL 双许可证发布。



我们希望（但并不要求）你在代码中增加归属信息。归属信息通常包含标题、作者、出版商以及 ISBN。比如：“Linux Device Drivers, Third Edition, by Jonathan Corbet, Alessandro Rubini, and Greg Kroah-Hartman. Copyright 2005 O'Reilly Media, Inc., 0-596-00590-3.”

## 意见和建议

请将有关本书的评论和问题发送给出版商，联系方法如下：

美国：

O'Reilly Media, Inc.  
1005 Gravenstein Highway North  
Sebastopol, CA 95472

中国：

100080 北京市海淀区知春路 49 号希格玛公寓 B 座 809 室  
奥莱理软件（北京）有限公司

我们还为本书建立了一个网页，其中列出了勘误、示例等内容。该网页地址如下：

<http://www.oreilly.com/catalog/linuxdrive3>


如果你希望对本书进行评论，或者遇到有关本书的技术问题，可发电子邮件到：

[info@mail.oreilly.com.cn](mailto:info@mail.oreilly.com.cn)  
[bookquestions@oreilly.com](mailto:bookquestions@oreilly.com)

有关 O'Reilly 的更多信息，包括图书、会议、资源中心以及 O'Reilly Network，可访问我们的 Web 站点：

<http://www.oreilly.com>  
<http://www.oreilly.com.cn>

## Safari Enabled

 如果你在自己喜欢的技术书籍封面上看到 Safari® Enabled 标识，则说明本书可通过 O'Reilly Network Safari Bookshelf 在线获得。

Safari 提供了比电子书更好的一种方案。它是一种虚拟图书馆，能让读者轻松搜索大量的顶级技术书籍、剪切和粘贴代码示例、下载图书章节并在需要最精确和最新的信息时快速找到答案。请点击 <http://safari.oreilly.com> 免费尝试。