

LINUX DEVICE DRIVERS

第二版
涵盖 Linux 2.4

LINUX

设备驱动程序



O'REILLY®
中国电力出版社

ALESSANDRO RUBINI & JONATHAN CORBET 著
魏永明 骆刚 姜君 译

LINUX

设备驱动程序

第二版

ALESSANDRO RUBINI & JONATHAN CORBET 著

魏永明 骆刚 姜君 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

05
10
02

图书在版编目 (CIP) 数据

Linux 设备驱动程序 (第二版) / (美) 鲁比尼 (Rubini, A.) 等著; 魏永明等译.
—北京: 中国电力出版社, 2002.10

书名原文: Linux Device Drivers, Second Edition
ISBN 7-5083-1294-5

I. L... II. ①鲁... ②魏... III. LINUX 操作系统—驱动程序 IV. TP316.89
中国版本图书馆 CIP 数据核字 (2002) 第 069178 号

北京市版权局著作权合同登记
图字: 01-2002-4326 号

© 2001 by O'Reilly & Associates, Inc.

Simplified Chinese Edition, jointly published by O'Reilly & Associates, Inc. and China Electric Power Press, 2002. Authorized translation of the English edition, 2001 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 2001。

简体中文版由中国电力出版社出版 2002。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复制。

书 名 / Linux 设备驱动程序 (第二版)
书 号 / ISBN 7-5083-1294-5
责任编辑 / 夏平
封面设计 / Edie Freedman, 张健
出版发行 / 中国电力出版社
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 36.5 印张 537 千字
版 次 / 2002 年 11 月第一版 2002 年 11 月第一次印刷
印 数 / 0001-5000 册
定 价 / 69.00 元 (册)

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 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 & 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。



译者序

由 Linus Torvalds 领导开发的 Linux 操作系统把我们带入了一个自由开放的世界。而在庞大的 Linux 内核源代码中,大约超过 85% 是各种驱动程序代码,其重要性可想而知。本书就是一本讲解如何在 Linux 系统中编写驱动程序,如何对各种新的计算机外设提供支持的著作。

本书由浅入深地探讨了驱动程序开发所涉及到的各方面问题,依次讲述了字符设备驱动程序、块设备驱动程序、网络设备驱动程序,以及常见外设硬件的工作特性和内核中的相关机制。同时提供了大量的程序实例。可以预见,本书严谨的组织、全面细致的讲解、丰富的示例程序,将使它和第一版一样成为 Linux 爱好者和开发人员的必备读物。

学习本书,不仅要理解原理,更重要的是动手实践。毕竟驱动程序是和设备紧密相关的,尝试运行示例驱动程序,并随着学习的深入逐渐扩展和完善这些驱动程序,将是学习本书的最好方法。知识只有被灵活应用时才能算作真正地被掌握了。

本书的第一版由 Alessandro Rubini 撰写,主要讲解了 Linux 2.0 内核系列中的驱动程序。由于 Linux 发展得很快,那本书已不能满足读者的需要了。于是 Alessandro Rubini 和 Jonathan Corbet 合著了第二版,主要讲述 Linux 内核 2.4 版本,系统地阐述了各种驱动程序编写的方法,同时也向后兼容了 2.0 及 2.2 版本的相关内容。

参与本书翻译的人员主要有魏永明、骆刚、姜君、刘新华、余义军、陈雷等。全书由魏永明统稿。本书翻译过程中,译者深切体会到原作者对知识和技术细节的执着追求和精益求精的精神。希望读者不仅能从本书中获得编写驱动程序的技巧,而且也能和我们一样,体会到计算机技术的精妙之处。

最后,感谢本书责任编辑夏平的细致工作,还要感谢技术审校人员和一些网友提出的宝贵修正意见。但由于我们自身水平所限,以及中英文本身的差异,难免会出现差错。恳请读者提出宝贵意见,我们将不胜感谢!

译者

2002 年 4 月

目录

前言	1
第一章 设备驱动程序简介	11
设备驱动程序的作用	12
内核功能划分	14
设备和模块分类	16
安全问题	18
版本号	20
许可证条款	21
加入内核开发社团	23
本书概要	23
第二章 构建和运行模块	25
核心模块与应用程序的对比	27
编译和装载	32
内核符号表	38
初始化和关闭	40

使用资源	46
自动和手动配置	53
在用户空间编写驱动程序	56
向后兼容性	58
快速参考	62
第三章 字符设备驱动程序	65
scull 的设计	65
主设备号和次设备号	67
文件操作	75
file 结构	79
open 和 release	81
scull 的内存使用	86
竞态简介	88
read 和 write	91
试试新设备	98
设备文件系统	98
向后兼容性	105
快速参考	108
第四章 调试技术	111
通过输出调试	111
通过查询调试	117
通过监视调试	123
调试系统故障	125
调试程序和相关工具	136
第五章 增强的字符驱动程序操作	145
ioctl	146
阻塞型 I/O	158

poll 和 select	171
异步通知	176
定位设备	180
设备文件的访问控制	182
向后兼容性	189
快速参考	195
第六章 时间流	199
内核中的时间间隔	199
获取当前时间	202
延迟执行	204
任务队列	207
内核定时器	219
向后兼容性	223
快速参考	224
第七章 获取内存	227
kmalloc 函数的内幕	227
后备高速缓存	231
get_free_page 和相关函数	234
vmalloc 与相关函数	237
引导时的内存分配	240
向后兼容性	243
快速参考	244
第八章 硬件管理	246
I/O 端口和 I/O 内存	246
使用 I/O 端口	249
使用数字 I/O 端口	255
使用 I/O 内存	259

向后兼容性	268
快速参考	269
第九章 中断处理	272
中断的整体控制	273
准备并口	274
安装中断处理程序	274
实现中断处理程序	286
tasklet 和底半部处理	290
中断共享	296
中断驱动的 I/O	300
竞态	301
向后兼容性	311
快速参考	312
第十章 合理使用数据类型	316
使用标准 C 语言类型	317
为数据项分配确定的空间大小	318
接口特定的类型	319
其他有关移植性的问题	320
链表	324
快速参考	327
第十一章 kmod 和高级模块化	329
按需加载模块	329
模块间通信	336
模块中的版本控制	339
向后兼容性	343
快速参考	344

第十二章 装载块设备驱动程序	346
注册驱动程序	347
头文件 blk.h	354
请求处理简介	356
请求处理详解	361
挂装和卸装是如何工作的	375
ioctl 方法	376
可移动设备	379
可分区设备	382
中断驱动的块设备驱动程序	389
向后兼容性	391
快速参考	393
第十三章 mmap 和 DMA	397
Linux 的内存管理	397
mmap 设备操作	410
kiobuf 接口	424
直接内存访问和总线主控	429
向后兼容性	448
快速参考	451
第十四章 网络驱动程序	455
snll 的设计	457
连接到内核	460
net_device 结构的细节	465
打开和关闭	474
数据包传输	476
数据包的接收	480
中断处理程序	482
链路状态的改变	483

套接字缓冲区	484
MAC 地址解析	487
定制 ioctl 命令	490
统计信息	491
组播	492
向后兼容性	496
快速参考	499
第十五章 外设总线综述	502
PCI 接口	502
ISA 回顾	526
PC/104 和 PC/104+	529
其他的 PC 总线	529
SBus	531
NuBus	532
外部总线	532
向后兼容性	535
快速参考	536
第十六章 内核源代码的物理布局	539
引导内核	540
引导之前	543
init 进程	545
kernel 目录	546
fs 目录	546
mm 目录	547
net 目录	549
ipc 和 lib	550

include 和 arch	551
drivers 目录	551
参考文献	561
词汇表	563

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，同时还将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，我们希望借助本书它能够更加开放，从而能够吸引更多的开发人员。

自本书第一版问世以来，Linux 的变化非常巨大。现在的 Linux 能够在更多的处理器上运行，并且支持更加广泛的硬件，许多内部的编程接口也相应发生了重大变化，因此，我们决定编写本书的第二版。本版以 Linux 2.4 版本的内核为主，讲述了新内核提供的所有新特色，同时，仍然兼顾了早期的内核版本。

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

Alessandro 的介绍

作为一个喜欢 DIY 的电子工程师，我一直乐于使用计算机来控制一些外部的硬件设备。从小时候使用父亲的 Apple IIe 计算机起，我就开始寻找另外一个平台，以便能够将我自制的电路板连接到它上面，并能够编写自己的驱动程序。不幸的是，不管是从硬件级别还是从软件级别来看，20 世纪 80 年代 PC 的功能都不是非常强大：

PC的内部设计比起Apple II来简直是差远了,而且可供利用的文档也远远不能令人满意。但在Linux出现之后,我决定尝试利用这个新的操作系统,为此,我购买了一个昂贵的386主板,但没有购买任何受到所有权保护的软件。

那时,我在大学里使用Unix系统,这个设计精巧的系统令我激动不已,尤其在有了由GNU项目提供给用户使用的、更加精巧的工具之后,这个系统更加令我着迷了。对我来讲,在我自己的PC主板上使用Linux内核,一直是最为难忘的经历,我不仅可以编写自己的设备驱动程序,而且还有机会再次拿起电烙铁。我不停地告诉别人:“我长大之后,一定要成为一名黑客”,而GNU/Linux则是实现这一梦想的最佳平台。可是,我不知道我是否已经长大。

随着Linux的成熟,越来越多的人开始乐于为自制的电子设备或者商用设备编写驱动程序。就像Linus Torvalds所说的那样,“我们又回到了能掌握自己的命运、从而自己编写设备驱动程序的时代”。

回顾1996年,我经常为那些从别人那里借来的,或者别人给我的,或者是我自己制作的硬件设备编写自己的设备驱动程序,并且乐此不疲。那时,我已经为Michael Johnson所著的《Kernel Hacker's Guide》撰写了一些内容,并开始为由Michael创办和负责的《Linux Journal》编写内核相关的文章。在Michael的引荐下,我认识了在O'Reilly工作的Andy Oram,他希望我能就设备驱动程序编写一本书,我非常高兴地接受了这一工作,有很长一段时间我一直忙于编写这本书。

到1999年,我发现,我已经没有足够的精力来独自完成本书的更新工作了:我的家庭添丁进口,而且我更多的时间要花在编写GPL'd软件的工作上。除此之外,内核变得更大,而且可以支持比以往更多的平台,而API也变得更加复杂和成熟。这时,Jonathan开始帮助我更新本书。他拥有足够的技巧、能力和热情,不断催促我跟上落下很多的进度。他凭着自己良好的技能和热情,已经成为推进这个进程的最得力助手,这些却是我无法企及的。我非常高兴能够和他共事,不管在技术上还是在私人方面。

Jon 的介绍

我从1994年初开始接触Linux,那时,我正在说服自己的老板为我购买一台Fintronic Systems公司生产的笔记本电脑。作为20世纪80年代初期(从那时起我就在和源代码打交道)的一名Unix用户,我立即被Linux所吸引。甚至在1994年,Linux就已

经成为一个非常实用的系统，而且也是我所遇到的第一个真正自由的系统。那时，我几乎完全丧失了对专有系统的兴趣。

但我并没有一个完整的、想为Linux编写什么著作的计划。当O'Reilly和我讨论有关帮助编写本书第二版的时候，我刚刚从我工作了18年的公司辞职，并成立了一个Linux咨询公司。为了吸引别人的注意力，我们建立了一个Linux新闻站点，即Linux Weekly News (<http://lwn.net>)，该站点的内容主要集中于内核开发。随着Linux的大众化，该Web站点也变得非常知名，而我们的咨询业务却最终被人遗忘。

然而，我的第一兴趣却始终还是系统编程。早些时候，我“修正”最初的BSD Unix系统当中的分页代码（这是一个可怕的黑客工作），或者在VAX/VMS系统上编写磁带驱动器的驱动程序（这些源代码是可获得的，如果你不在意这些由汇编和Bliss语言编写的代码的话）。随着时间的推移，我又为Alliant、Ardent和Sun等系统编写驱动程序。后来，我开始利用Linux开发实时雷达数据收集系统。这个时候，也就是编写本书的时候，我正致力于修正Linux软盘驱动程序中的I/O请求队列锁的实现。

我为能参与本书的编写而感到高兴。首先，通过本书的编写，我能够更加深入地研究内核代码，同时能够帮助别人达到同样的目的。Linux是个实用的系统，同时也是一个能带给人乐趣的系统，而围绕内核工作则是所有事情当中最有趣的（至少对那些兴趣乖僻的黑客是如此）。和Alessandro一起工作也令人高兴，我必须感谢他对我如此信任，让我修改他撰写的优秀文字，也感谢他在我出现错误或者不能赶上进度时的耐心，当然也得感谢那次到Pavia的自行车旅行。编写本书的那些时光的确令人难忘！

本书的读者对象

在技术方面，本书提供了一条便捷的途径，让读者可以理解内核内幕以及一些Linux开发者所做出的设计决策。尽管本书的主要目的是教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，好的书籍能够更好地帮助读者提高编程技巧。本书内容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。

本书对那些希望编写计算机设备驱动程序的人员，或者对那些要解决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，所以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。

Linux的狂热爱好者将从本书中找到深入内核代码的足够精神食粮；通过本书的学习，此类读者将有能力加入到为增加功能或提高性能而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设备驱动程序的开发人员，需要了解如何和许多内核子系统一起工作。因此，本书对内核编程做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Linux的开发大军中。

另一方面，如果你只是为了为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的话，那么本书的模块化结构足以满足你的需求。如果你不想深入到细节当中，则可以简单地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部分。

本书的主要目标是为Linux内核2.4版本编写内核模块。模块(module)是能够在运行时装载到内核的目标代码，它能够为正在运行的内核添加新的功能。我们尽可能地让示例代码也能够在内核的2.2和2.0版本中运行，如果需要做些改动则会特别指出。

内容的组织

本书内容由简到难，并被划分为两大部分。第一部分（第一章到第十章）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。每一章分别讲述某一个特定问题，并在每章结尾包含了一个“符号表”，该符号表可在实际开发中作为参考。

在本书的第一部分中，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，你能够尽可能地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，并给出了能够在任何Linux计算机上运行的示例驱动程序。但是，在

第八章和第九章中，我们需要你在并行端口上连接一些电线，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。

本书的第二部分讲述了块设备驱动程序和网络接口，并深入讨论了一些更高级的内容。许多驱动程序编写者可能不需要这些内容，但我们建议你阅读这些章节。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对了解 Linux 内核的工作原理也很重要。

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 命令和管线，因此，也需要读者拥有 Unix 的使用经验。

在硬件级，即使没有任何经验的读者也可以理解本书的内容，当然，对一些一般性的概念必须清楚。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。

建立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，而太新的工具有时又可能会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要求也不同，你可以参考内核源代码树中的 *Documentation/Changes* 文件。

其他信息的来源

本书提供的大部分信息直接来自内核源代码以及相关文档。我们要特别注意内核源代码树中的 *Documentation* 目录，其中包含有大量有用的信息，比如内核 API 中新增部分的文档（在 *DocBook* 子目录中）。

还有一些有用的书籍包含了更为广泛的内容，这些书籍列在本书的“参考文献”中。

Internet 上有大量可用的信息，下面将列出部分站点。当然，Internet 站点的更新非常快，而印刷书籍却难以及时更新。因此，下面的清单可能已经过时了。