



本书内容适用于Linux内核2.6以上版本

Linux

操作系统之奥秘

邱世华 著 白涛 审校

全面呈现操作系统新思维

ACPI高级配置电源管理、省电状态切换，以及省电模式的设定等

EFI可扩展固件接口、全新形态的BIOS，明显提升开机性能；

upstart启动机制、Fedora所采用的全新开机流程，兼具向下兼容与弹性

Virtual Machine虚拟机技术再提升：包含Xen、KVM、VMWare等。

Linux



操作系统之奥秘

邱世华 著 白涛 审校



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内容简介

本书是一本关于Linux整体概念的工具书，书中用实际的操作步骤和开机流程来说明操作系统的本质与精髓，希望通过阅读本书，读者能提高对Linux操作系统架构的理解。

本书分三篇共9章内容，主要包括BIOS、引导管理程序、kernel与initrd、进入系统前的Script File、进入系统、系统中的电源管理、系统检查、系统性能、Virtual Machine等内容。

本书适合任何对Linux感兴趣的读者，同时也可作为大中专院校相关专业辅导读本。

本书为精诚资讯股份有限公司授权电子工业出版社于中国大陆（台港澳除外）地区之中文简体版本。本著作物之专有出版权为精诚资讯股份有限公司所有。该专有出版权受法律保护，任何人不得侵害之。

版权贸易合同登记号图字：01-2010-1367

图书在版编目（CIP）数据

Linux操作系统之奥秘 / 邱世华著. —2版. —北京：电子工业出版社，2011.10
ISBN 978-7-121-14233-8

I. ①L… II. ①邱… III. ①Linux操作系统 IV. ①TP316.89

中国版本图书馆CIP数据核字（2011）第153677号

策划编辑：刘 皎

责任编辑：贾 莉

印 刷：北京中新伟业印刷有限公司

装 订：

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：787×1092 印张：21.25 字数：476千字

印 次：2011年10月第1次印刷

印 数：4000册 定价：56.00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zlts@phei.com.cn，盗版侵权举报请发邮件至dbqq@phei.com.cn。

服务热线：（010）88258888。

作者序

没想到这本书可以再版，这都要感谢读者们的“爱护”。一开始笔者只想把所学的东西、所知道的事情做个整理，作为一本参考书，分享给大家，没想到大家对于这本书的内容非常感兴趣，让笔者十分欣慰。在写作过程中，也陆续收到读者的来信（这是我最兴奋的事情），询问书中一些相关的问题，也希望通过这些邮件，在本版中能补充读者们所需要知道的答案与知识。

一开始会希望有再版的机会，主要是书中所提到的“upstart”与“ACPI”两大部分。笔者总是纳闷，为什么遍寻网络、书本都找不到有关“upstart”的文章？但是，这部分在较新版的Linux中，却是一项非常重要的特色，尤其在最近发布的Ubuntu 9.10中，可以看到把upstart的精神发挥得淋漓尽致（该目录已被搬到“/etc/init”下），很多的系统设置都是通过upstart来启用的。

至于“ACPI”，笔者认为这个功能在系统的运行中，是特别值得注意的。不论是关机、重新启动还是休眠等，尤其我们已身处在一个重视环保的年代，似乎更需要了解电源管理的意义。再有就是笔记本电脑（比服务器计算机更注重ACPI的功能）使用率的提高，而且目前的Linux对ACPI的支持度比起以前实在好太多了（难道是与笔者所使用的版本有关？）。不论如何，笔者相信读者会和我一样对电源管理的设置及应用感兴趣的。

随着家中小朋友的年纪越来越大，老婆的脾气也越来越受到“考验”，真的很为难在这样的時候还要让我“安静”地写书，希望这本书可以当做已经过期的生日礼物（好像很久没过生日了……）。另外，也要感谢两个小朋友的阿姨们对我老婆的关心，她们一直都是我老婆可以“发泄”的对象。当然也要感谢这本书的编辑书瑜以及小世，忙到快翻掉还要被我催稿（好像

反过来了)，希望大家可以长长久久地一直合作下去。

当然没有忘记最伟大的读者们，没有你们就没有这本书。大家若对这本书有任何意见，都可以告诉笔者，大家都是Linux User，互相交流是唯一进步最快的方式。当你们询问问题时，就等于是让笔者在Linux领域有进一步钻研的机会，这也是为何很高兴收到读者写信询问的原因，期待你们的问题！

邱世华

Juergen.chiu@gmail.com

2009年11月

使用Linux这么多年，一直以来，都没有时间好好地把研究及经验整理出来，一方面也是因为Linux下的kernel、软件更新速度实在太快了。因此，在规划本书时，便希望尽量不要因套件、版本的不同，而造成内容的差异。

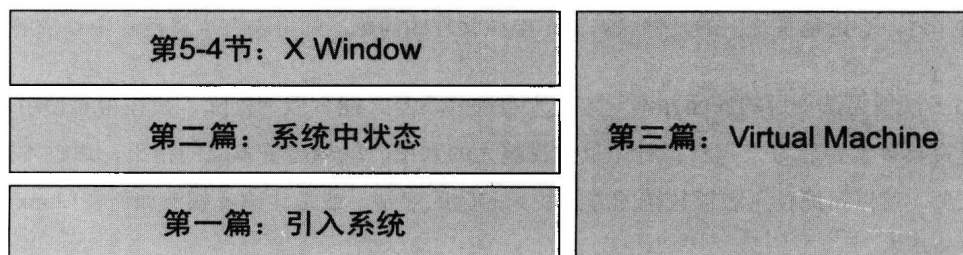
另一方面，也期望将本书规划成非常实用的工具书，不论遇到任何系统的疑难杂症，都可以利用这本书找到真正核心的问题，因为Linux并不是短时间即可学好的操作系统，所需要的入门知识实在太多了！

建议你在阅读本书时，一定要掌握一个重点——Linux与硬件的关系是十分密切的，这也是本书一再强调的。虽然本书是以Fedora为主要的范例说明，但只要是使用Linux当成kernel的操作系统，其内容都适用。本书结构主要分成三大篇：

■第一篇：引入系统

■第二篇：系统中状态

■第三篇：Virtual Machine



本书篇名与系统的关系图

三篇的关系就好比是整个操作系统的架构，从硬件到X Window都绑在一起，彼此关系密不

可分。尤其是第三篇的Virtual Machine，是一个很新、很特别的机制，因为跨越了硬件、软件及操作系统的界线，而成为了一个独立的虚拟操作系统。

第一篇：引入系统

本篇内容主要以Linux基本观念为主，让读者可以通过实例，了解Linux进入操作系统的每一个步骤，以及软硬件之间的关系。以整个启动流程为章节顺序开始介绍，从BIOS一直到登录Linux都涵盖在内。第1章虽然和Linux没有直接关系，但BIOS对Linux启动以及系统的实际操作上有一定程度的影响，因而将其加入到基本观念中。接着，之后的第2章、第3章和第4章都是Linux中不得不知的细节。

kernel及initrd是比较偏重硬件的部分，如果是CPU、内存或硬件的问题，将造成无法开机，比较有可能会在第3章及第4章启动的过程中看到，因为在这两章中的阶段是开始启动kernel，却又尚未进入真正的操作系统。

第二篇：系统中状态

在第5章中，许多启动时遇到的问题都是在此阶段产生的。因为这是进入系统的第一个阶段，所有的环境、区域变量，以及系统的设置都是在此建立的，也就是说，如果问题是由系统本身所造成的，将很有可能会停在这个步骤。

在硬件之上，如何通过Linux内置的指令以及一些外在的工具程序，判断系统的稳定度？本篇的重点在于如何检查系统软硬件的搭配状况，也针对目前很重要的观念——“电源管理”做深入的探讨。不论是系统中的硬件状态或是电源管理的问题，都可以在本篇范围中寻求解答。

很多硬件无法使用的原因并不一定是来自硬件本身，能否顺利排除，往往得靠使用者对软硬件是否具备整合能力来判断。对一个负载越大的系统，就越需要本篇的指引，通过本篇所介绍的概念，可以帮助使用者以较快的方式找到问题的症结，或是让读者知道如何在Linux中对应到硬件的状态。

从系统管理者的角度来看，整个系统的稳定度都操控在手中，若没有将系统与软硬件之间的关系理清楚，很容易会产生非预期的结果。本书中的第一篇及第二篇都是非常重要的部分，

尤其像现在的服务器，很多会因为性能取向而直接牺牲图形接口，这样的导向就更需要第二篇之前的基本观念。

对一般使用者来说，计算机一买来，就应该按照第7章所谈到的系统检查先行检查过，这是个很容易忽略的关键，虽然不是每一项都会影响到启动的进行，但性能的差别是非常大的。

之后再运用第8章所介绍的系统性能观念，并参考其中所介绍到的硬件测试工具或方法，以了解性能所带来的差异。有时候某些硬件并不一定需要工具程序才可以评估好坏，在有限的资源中，也有可以即时知道的方法。

第5.4节：X Window

会把这一节拉出来独立做说明，是因为本节主要是以探讨X Window为主，而与其他相关的章节进行区分。虽然说这也是属于系统中的状态，但重点是Linux中图形界面的基本观念，本节的重点是让读者对Linux如何让使用者具备图形界面能有较清楚的认识。

和以往所看到如何在X Window下使用各式各样的软件不同，本节是以一个全新的角度来看窗口界面架构的。与本书的基本精神一样，我们并不会提到如何操作X Window，取而代之的，是说明X Window的启动流程以及整个X Window的架构，这对于基于笔记本电脑使用X Window的使用者而言，是非常有帮助的，尤其是当遇到启动X Window的问题时。

第三篇：Virtual Machine

另一个独立的章节就是第9章——Virtual Machine，原因在于Virtual Machine于Linux是一个全新的架构，虽不能说是Linux的观念，但往后所有的Linux，都有可能会加入Virtual Machine的功能。因为是全新的功能，本章重点先放在如何操作上，这也是本书唯一介绍操作层面的章节，希望让读者可以以最快的速度熟悉此项最新的功能。

Virtual Machine是未来的趋势，不只是Linux，Windows也已经走向这样的机制，Windows Server 2008其实就已经是Virtual Machine的操作系统。连Windows都已经默认以Virtual Machine为系统架构，身为Linux使用者的你当然一定要知道。

本书所要带给读者的，不是速成找到问题的解决之道，而是希望通过建立对整体Linux操

作系统架构的理解，让所有使用者可以在问题发生时，了解其问题原因所在以及未来防范的方法。相信唯有如此，才可以让一台Linux的主机无后顾之忧，而不需要一直有追随版本的迷思。



Note|如何取得Linux操作系统——Fedora

Fedora 是由Red Hat所主导的免费Linux-based操作系统，读者可从以下网站下载各版本的安装光盘image文件：

■ 官方网站：

<http://fedoraproject.org/>

■ 中国台湾地区相关网站：

<http://free.nchc.org.tw/fedora/linux/releases/>

<ftp://ftp.isu.edu.tw/Linux/Fedora/linux/releases/>

目 录

第一篇 引入系统

第1章 BIOS——启动硬件及加载固件的灵魂.....	2
1.1 何谓BIOS	3
1.2 Power On	10
1.3 POST.....	12
1.4 BIOS信息	14
1.5 BIOS修复	17
1.6 硬盘的零磁道.....	18
1.7 BIOS与操作系统的交互	26
第2章 引导管理程序.....	28
2.1 何谓GRUB	29
2.2 GRUB的设置方式	38
2.3 多重引导的管理.....	44
2.4 安装GRUB	50
2.5 除错方式.....	59
2.6 拯救无法开机的情况.....	61
第3章 kernel & initrd: 内核程序与初始化文件.....	68
3.1 何谓kernel.....	71
3.2 编译kernel.....	74
3.3 何谓initrd.....	91
3.4 开机时常用的kernel参数.....	102

第4章 进入系统前的Script File	107
4.1 认识nash	109
4.2 挂载主要的文件系统	111
4.3 建立设备文件所需的文件系统	117
4.4 转移前的准备	131

第二篇 系统中状态

第5章 进入系统	140
5.1 upstart启动机制	143
5.2 runlevel的介绍	147
5.3 登录Terminal	153
5.4 X Window	180
第6章 系统中的电源管理	207
6.1 何谓ACPI	208
6.2 G State	215
6.3 S State	217
6.4 C/D/P State	223
第7章 系统检查	228
7.1 硬件	229
7.2 ext3与journaling	241
7.3 ext4	246
7.4 固件	249
第8章 系统性能	262
8.1 性能调校	263
8.2 硬件测试	274
总结	287

第三篇 Virtual Machine

第9章 Virtual Machine	289
9.1 虚拟机的原理	293
9.2 XEN	306
9.3 KVM	313

The background is a grayscale abstract composition. It features several wireframe spheres of varying sizes, some of which are partially obscured by other elements. A horizontal band of binary code (0s and 1s) runs across the middle of the image. In the lower foreground, there are two vertical, cylindrical structures that resemble stylized antennas or sensors, each with a circular base and a pointed top. The overall aesthetic is technological and futuristic.

第一篇 引入系统

第 1 章

BIOS——启动硬件及加载固件的灵魂

本章学习重点：

- BIOS在启动过程中的重要性
- 如何通过Port 80找出无法启动的原因
- BIOS POST的意义
- BIOS与操作系统交接的方式
- 认识引导扇区

计算机一开机，电源刚从Power进入到主板时，第一个启动的软件便是主板上的BIOS，BIOS存在的目的包括：启动各组件、检查硬件、分配资源，以及协助加载操作系统。在加载操作系统这部分，很多介绍操作系统的书籍都忽略了操作系统和BIOS关联性的问题，其实BIOS与操作系统的关系就好比人的身体和小脑，密不可分，在操作系统运行时，往往需要请BIOS出面协调许多事情。一般而言，用户在操作系统下会忘记BIOS，主要原因是当启动进入大家所熟悉的，如Windows、Linux、FreeBSD等操作系统之后，就看不到BIOS为操作系统所带来的系统信息及硬件关联性了。此外，参照BIOS信息的程度也因操作系统的不同而不同。

在本章，首先要强调的便是这个小型操作系统——BIOS。

1.1 何谓BIOS

一般提到的BIOS，就是指操作系统与硬件或固件间的一种桥梁，此为固件的一种，早期只有BIOS这套机制可以执行这样的功能，因此统称为BIOS。就目前的现况来看，BIOS可分为下列两大类：

- ◎ Legacy BIOS：一般的家用计算机或是笔记本电脑，都还是以Legacy BIOS为主流，短时间内并不会会有太大的变动。
- ◎ EFI：这是一种较新的软件接口，目前大部分都是在服务器上出现的，并且数量还不是很 多，但之后应该会慢慢地全部转到EFI。

1. Legacy BIOS

Legacy BIOS其实就是一般所说的BIOS，全名为基本输入输出系统（Basic Input Output System, BIOS），一般的教科书或是文件中所提及的操作系统（Operating System, OS），都是像Windows或Linux这种形态的系统软件，而BIOS则是被归类为固件（Firmware）。其实，如果以面向任务来区分，BIOS比较偏向操作系统，理由很简单，当用户接上电源的那一刻，负责和所有硬件沟通并将计算机呈现到用户面前的，便是BIOS，而这不就是操作系统的工作吗？

想要简单认识BIOS，首先需要知道下列名词：

◎ 北桥、南桥

◎ Flash Memory

◎ CMOS

◎ SMBIOS

在此不针对每个名词做深入的解释，重点将放在各组件彼此之间的关系上，这几个组件彼此的关联性如图1-1所示，下文会再一一补充说明。

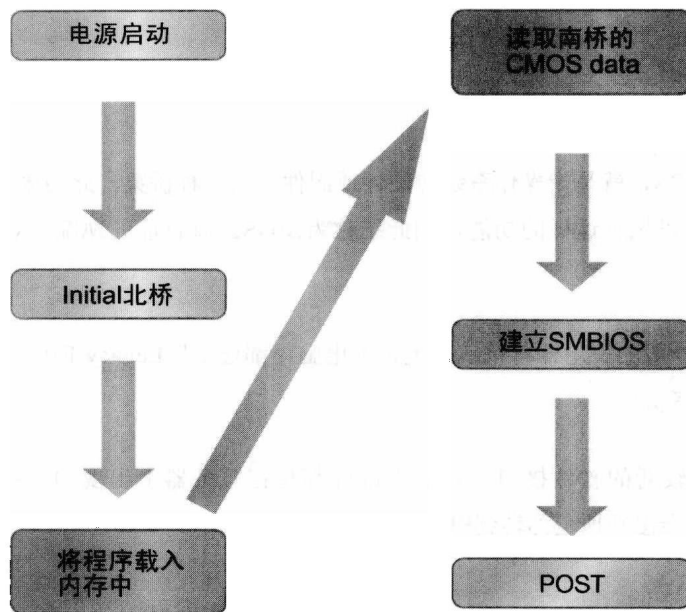


图1-1 BIOS与各组件的关系图

北桥、南桥是主板最重要的芯片组（目前AMD已经不使用这一种架构），一般而言，北桥主要控制着CPU、内存；南桥则是负责PCI、PCI-E、USB、VGA等所有周边相关设备。在南桥里面有一个特殊的区块，是负责存储CMOS的空白区域，是用来让BIOS存储用户设定的地方（在BIOS选项中可以设定的值都是存储在这一个区域中的），而主板上的水银电池，其主要目的是为CMOS保存资料提供电源的。

当BIOS启动时，会先去确认CMOS中的资料是否正确，如果正确，便会将之前用户存储在

CMOS中的资料，加上已存在或找到的硬件信息，整合成为一个表格，写入内存中，也就是所谓的SMBIOS，如果发现错误，则会自动以默认值取代CMOS提供的资料，所以，这一份写入内存的SMBIOS表格，就是用户进入BIOS选项后，可以看到或选择的所有信息。

市面上最常见的BIOS厂商，如Phoenix Technologies (<http://www.phoenix.com/>) 或AMI (American Megatrends Inc, <http://www.ami.com/>)，目前都被很多PC大厂采用，大部分的计算机一开机都可以看到它们的Logo（有些大厂会将Logo换成自己的），因为核心的版权还是归属于BIOS厂商。一般OEM或ODM厂商的做法是先选一家BIOS厂商合作，购买部分源代码的版权进行修改，再配合主板的特性，发放给用户使用。

在计算机研发的初期，其实BIOS的工程很浩大，原因是在市面上销售后，用户在计算机使用上可能会面临到的问题有一大半是来自于BIOS，而测试软件所测试的计算机性能结果，也有部分会因为BIOS中的参数而导致不一样的数据。所以，开发者在品质及效率上都要兼顾，而BIOS涉及的硬件又非常多，因而在制作上相当麻烦。

BIOS的作用主要有以下四点：

- ◎ 自检及进行初始化：BIOS在开机的同时，会首先检测主板上的所有相关芯片，并通知各芯片开始运行。
- ◎ 记录系统设定值：用户可以通过设定BIOS来改变各种不同的设置，如onboard显卡的内存大小等。
- ◎ 中断处理：主板上的任何资源都是由BIOS分配的，包括IRQ，即所有AGP/PCI/PCI-E插槽所需要被赋予的珍贵资源。
- ◎ 加载操作系统：用户手上所有的操作系统，都是由BIOS转交给引导扇区，再由引导扇区转到各分区所启动的。

会在BIOS之前产生的动作就是Power On，也就是按下计算机上电源开关的那一瞬间（我们将在1.2节“Power On”中详细介绍）。在Power On阶段一开始，会进入BIOS的启动流程（如图1-2所示），当用户一按下电源开关，CPU会先启动去寻找BIOS，接着，BIOS会先让本身在flash memory中开始执行，再加上CMOS中用户喜好的设定值，将自己解压缩到计算机的内存中。

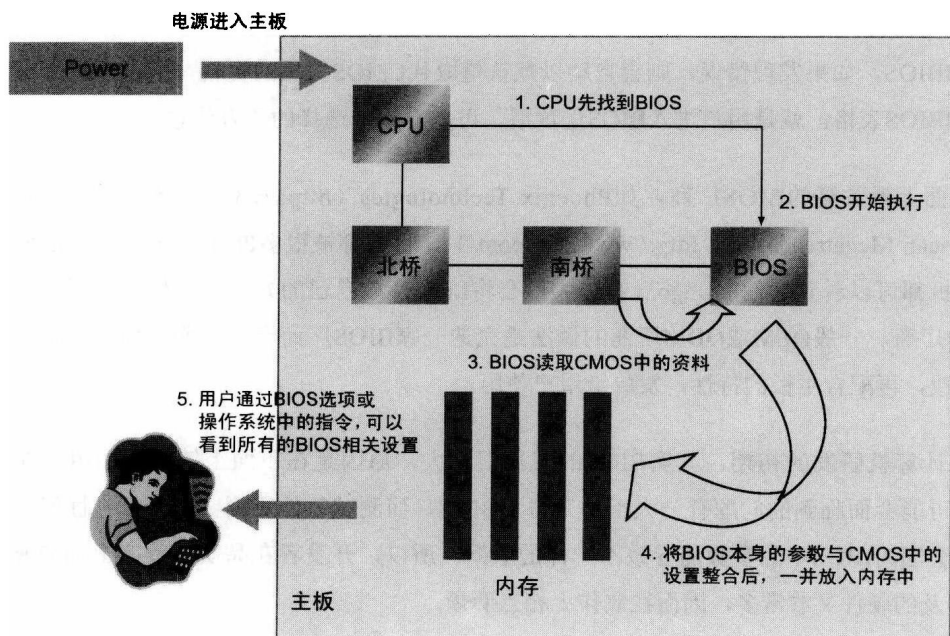


图1-2 BIOS的启动流程

此时，用户若按下【Del】（一般计算机的BIOS设置键），就可以进入BIOS所提供的设置选单，而看到所有的设定值，或是当用户进入操作系统后，操作系统就可以参考内存中的BIOS数据，读取每一个硬件的地址。

既然知道了BIOS在主板上的重要性，就应该知道BIOS对整个启动过程（从按下电源开关到进入操作系统）及系统的运行性能而言，占有极为重要的地位。电源进入CPU后，CPU就会先找到BIOS，BIOS便成为硬件和操作系统间的一个桥梁。在Power On阶段，BIOS该做的动作有哪些，以及对整个操作系统的影响如何呢？接下来，将会做更详细的介绍。

2. EFI

EFI (Extensible Firmware Interface) 的做法最早是由 Intel 提出的，不过，现在已经交由 Unified EFI Forum (www.uefi.org) 管理。后来，EFI 的名称也直接改为 UEFI，这也是为何在以下部分或是图中所看到的都是 UEFI 的原因。UEFI 组织主要是由一些市面上的大厂共同组成的，其中包含原本的 BIOS 公司 (AMI、Phoenix 和 Insyde)、CPU 硬件厂商 (AMD 和 Intel)、操作系统厂商 (Microsoft 和 Apple) 和一些知名的计算机品牌大厂 (IBM、HP、Dell 和