# Linux是怎样工作的

TURING
图灵程序设计丛书

实验×图解
直击Linux核心工作原理

“现代计算机操作系统”图解趣味版

[日] 武内觉 / 著　　曹栩 / 译

How Linux Works

本书适合

1 菜鸟程序员入门进阶　2 中级程序员查漏补缺
3 高手程序员 / 教师讲解操作系统基础知识

中国工信出版集团　人民邮电出版社 POSTS & TELECOM PRESS

# 版权信息

书名：Linux是怎样工作的

作者：[日] 武内觉

译者：曹栩

ISBN：978-7-115-58161-7

# 版权声明

*[TAMESHITE RIKAI] Linux NO SHIKUMI : JIKKEN TO ZUKAI DE MANABU OS TO HARDWARE NO KISOCHISHIKI*

by Satoru Takeuchi



Original Japanese edition published by Gijutsu-Hyoron Co., Ltd., Tokyo

This Simplified Chinese language edition published by arrangement with

Gijutsu-Hyoron Co., Ltd., Tokyo in care of Tuttle-Mori Agency, Inc., Tokyo

本书中文简体字版由 Gijutsu-Hyoron Co., Ltd. 授权人民邮电出版社有限公司独家出版。未经出版者书面许可，不得以任何方式或途径复制或传播本书内容。

# 推荐序

我很久以前就认识本书作者武内先生了，回想起来也一同工作十年左右了。

武内先生非常擅长教学。在本职工作之外，他每年还会受邀对大企业中软件开发相关职位的新员工进行操作系统运行原理方面的培训，非常能干。无论是从新员工对培训的满意程度来说，还是从对培训内容的理解程度来说，他的培训都受到了很高的评价，因此在公司内也有口皆碑。另外，在 IPA[1] 的安全知识集训营等中，他所做的操作系统方面的介绍也受到了学生的欢迎。

> [1]即 Information-technology Promotion Agency（信息技术促进机构），是日本的一个独立行政法人，旨在推出解决社会问题和促进产业发展的方针，同时强化信息安全对策，培育优秀的 IT 人才等。——编者注

我也有过培训新人的经历，深知教操作系统有多难。因为不得不从硬件知识开始讲起，所以每一个知识点都需要讲很长时间。而理解这些知识点需要拥有基本的编程知识，这难免会让初学者打起退堂鼓。

武内先生的教学方式非常独特。他经常使用丰富的图表和能够验证所讲知识的实验数据，来简明且直观地解释各个知识点。比如，在讲解对注重性能的程序来说必不可少的高速缓存时，他会以“图解”的方式说明高速缓存的工作原理，并用图表展示内存与高速缓存的速度差距。通过这种方式，他大大提高了新人编写程序的质量。

看到武内先生基于他丰富的教学经验，把他关于操作系统运行原理的见解总结到一本书中，我感到无比兴奋。正如书名所示，这本书的主题是 Linux，因此如果你想了解 Linux 系统是怎样工作的，或者想尝试自制操作系统，或者想优化程序性能，那么这本书一定会对你有所帮助。

Linux 内核黑客、Ruby 语言贡献人
小崎资广
2018 年 1 月 30 日

# 前言

本书的目标是，通过实际动手操作，一边验证结果，一边讲解构成计算机系统的操作系统（以下简称 OS）和硬件设备的运行原理。本书介绍的 OS 是 Linux，目标读者是应用程序开发人员、系统设计师、运维管理人员和技术支持人员等。只要知道 Linux 的基本命令，就能阅读本书。

由于现代计算机系统的层次化和功能细分化，用户一般不会直接意识到 OS 和硬件设备。层次化通常用如图 0-1 所示的漂亮模型来解释。负责任意一层的人，只需了解比他负责的那一层更深的一层就可以了。比如，运维管理人员只需了解应用程序的外部构成就行，应用程序开发人员只需知道如何运用库就行。


图 0-1　计算机系统的层次（漂亮的模型）

但是，现实中的系统其实是像图 0-2 那样的。不管哪一层，都与其他层有着复杂的关联。如果只了解其中一部分，就会遇到很多自己解决不了的问题。而且现实情况是，人们通常不得不在实际工作中花费大量时间去学习覆盖这么多层的知识。

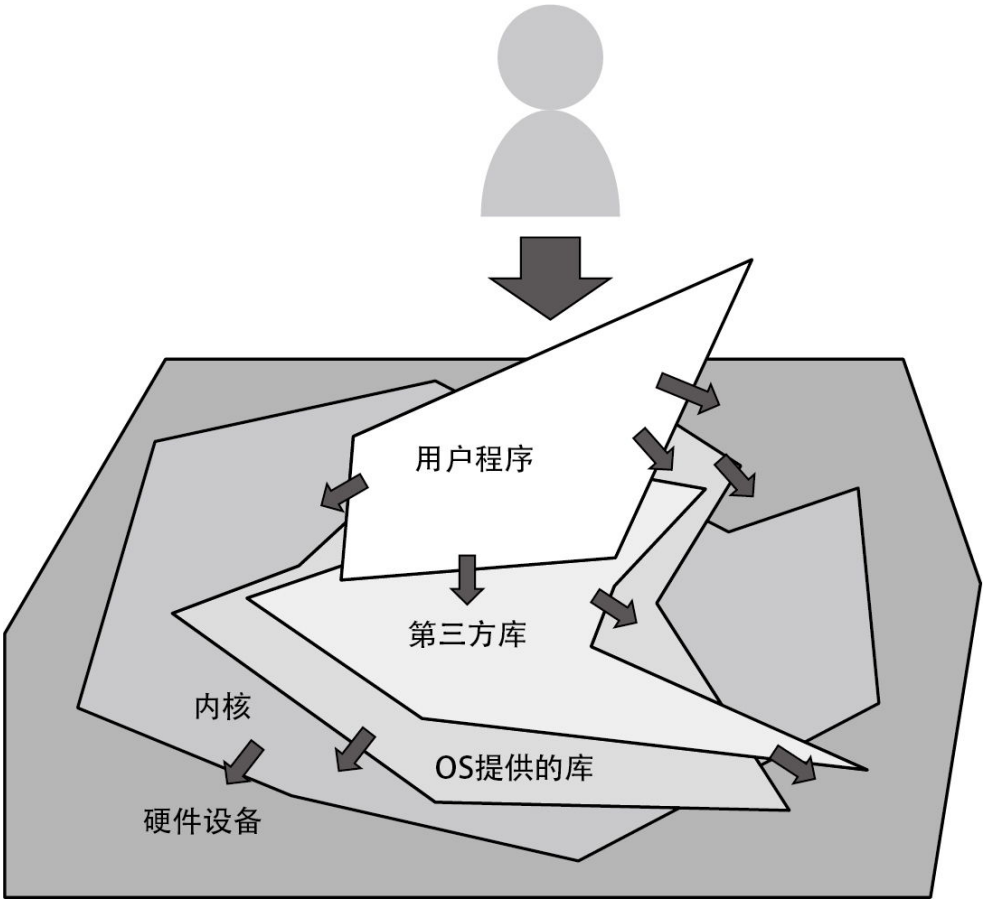


图 0-2　计算机系统的层次（现实情况）

笔者编写本书的目的就是解决这一问题。

通过阅读本书，读者可以在以下方面得到提升[1]。

> [1]虽说如此，但本书的目标读者并非想成为 OS 或硬件方面的专家的人。本书的内容只是笔者根据自己的判断精选的、至少需要了解的 OS 和硬件设备的相关知识。

- **根据硬件的特性更好地开发软件**
- **明白应该依照什么指标来设计系统**
- **冷静地处理 OS 或硬件设备相关的问题**

◆　◆　◆

需要说明的是，由于网络相关的信息庞杂，如果将其写入本书，本书的主旨将变得模糊不清，所以本书不介绍网络相关的内容。

本书提供了大量实验程序，读者可以通过亲自动手运行程序，来确认系统的行为。建议大家一定要亲自动手运行，这是因为与“只看书”相比，“边看边尝试”的学习效果要好得多。实验程序可以从本书的支持页面[2]）下载。此外，关于函数的含义，书中也会稍作解释。由于源代码的开源许可证是 GPL v2，所以大家可以随意使用或更改。对于不想运行程序的读者，本书也展示了笔者的计算机上的运行结果，只要理解了相应的内容，就完全没问题。

[2]请至“随书下载”处下载本书实验程序。——编者注

本书中的实验程序是使用 C 语言和 Python 编写的，另外也有少量 Bash 脚本。这里顺便补充一下使用 C 语言的理由。与现今流行的 Go 或 Python 等编程语言相比，C 语言只有比较原始的功能，因此其生产力较低。但是，拥有比较原始的功能就意味着，可以通过它看到 OS 和硬件设备原本的样子，这一点与本书目标一致，因此本书选择使用 C 语言来编写实验程序。

在本书中，实验程序的运行环境是 Ubuntu 16.04/x86_64。不过，由于实验程序并不依赖 Linux 发行版，所以即使 Ubuntu 的版本不同，或者发行版不同，程序也应该可以正常运行。此外，请尽量使用搭建在实体机而非虚拟机上的系统，因为在虚拟机中，部分实验程序的运行结果会和本书不一样。

在运行实验程序或收集其他统计信息时，需要以下软件包。

- binutils
- build-essential
- sysstat

这些软件包可以通过以下命令来安装。

```
$ sudo apt install binutils build-essential sysstat
```

本书中的数据是在如下配置的计算机上得到的。

- CPU：Ryzen 1800X（超线程关闭）
- RAM：Kingston KVR24N17S8/8×4（32 GB）
- SSD：Crucial CT275X200（256 GB）

- HDD：SEAGATE ST3000DM001（3 TB）
- Ubuntu 16.04/x86_64
- Linux 内核：4.10.0-40-generic