

下 篇

莱昂氏UNIX源代码分析

本书是UNIX操作系统版本6源代码的姐妹篇。它对UNIX操作系统的源代码进行了分析。UNIX软件系统是由贝尔实验室的肯·汤姆森和丹尼斯·里奇编写的。

本文档包含了贝尔系统专有的信息。仅限于贝尔系统中工作与此相关的授权雇员使用。不允许向非授权雇员泄露本书的内容或为其作复制。在贝尔系统之外，只限于向 UNIX分时操作系统版本6许可证的持有者配售此书。使用、复制本文档受到 Western Electric 公司发出的这种许可权上所指明的限制。

前 言

本书力图详细解释一个操作系统的内核，该操作系统在几年内将成为最重要的系统之一。这个操作系统就是UNIX分时系统，它在数字设备公司的PDP11计算机系统的较大型机上运行，由贝尔实验室的肯·汤姆森和丹尼斯·里奇设计并实现。《ACM通信》(Communication of ACM)1974年7月号上宣布其问世。

在对UNIX稍加研究后，我们就发现UNIX可作为学生们深入学习的正式课程，其理由是：

- UNIX在我们已具有的系统上运行。
- 该系统非常紧凑，源代码和有关资料都可方便取用。
- 它提供非常广泛的可用功能。
- 它是一个非常优良的操作系统，事实上它开辟了许多操作系统的新领地。

在UNIX分时系统的魅力和优越性中的一个重要点是其源代码的紧凑性。当只提供少量外部设备时，该系统常驻内存的内核源代码长度大约只有9000行。通常认为一个人能够理解和维护的程序长度约为10000代码行。大多数操作系统或者其长度超过这种限制1~2个数量级，或者只向用户提供非常有限的功能，也就是说或者除极少数非常专注、能投入大量时间的学生外大多数学生都无法了解其详细情况，或者这种系统是专用的，在技术方面没有进行学习的实际价值。

教授操作系统课程大致有三种方法：第一种是“一般原理”法 (general principle)，这种方法详细阐述基本工作原理，并引用若干个现存操作系统作为示例，但一般学生对这些系统都缺少直接经验。这种方法是COSINE委员会所倡导的，但是根据我们的经验，大多数学生的基础和经验都与这种方法不相适应。

第二种是“积木块”法 (building block)，这种方法使学生能够设计并实现一个小规模或“玩具式”的操作系统，毫无疑问，如果组织得当，这会是非常有价值的经验。但是，这种方法通常偏重于操作系统设计的某些方面，例如进程同步，因此难于使学生体会实际操作系统的复杂性以及有关的技术难度。

第三种是“实例学习”法 (case study)，这是68教程对于系统程序设计课程所推荐的方法，68教程是ACM计算机科学教程委员会 (ACM Curriculum Committee on Computer Science) 编写的一个报告，发表于《ACM通信》1968年3月号上。

十年前，这种方法所倡导的将该课程的大部分内容 and 时间安排到一个单一系统的学习上是不现实的，其原因是，那时为一定数量的学生提供这种了解和适当系统所需的开销太高了。

十年后，经济状况发生了很大变化，如果课程学习所需的只是小型计算机，那么费用就不再是主要的障碍。这种方法的主要优点——对一个现存系统可以作详细分析，现在能够变成现实。

我们的意见是：使学生有机会学习一个实际工作的操作系统是极为有益的。还有，主修计算机科学专业的学生在学习生涯中至少应该读并理解一道相当规模的程序，这会使他们受益匪浅。

1976年，我们在新南威尔士大学的两门操作系统课程中采用UNIX作为实例进行教学。本书中对UNIX所作的各种分析最初的目的就是为了帮助选修这些课程的学生。

这两门课程的每一门都安排为一个学期。对选修这些课程学生的要求是，他们已学习过PDP11系统结构和汇编语言，并且在他们先修课程的练习中已经得到机会使用过UNIX操作系统。

一般而言，学生们发现使用 UNIX作为实例进行教学的新课程需要花费很多时间，但与使用 COSINE委员会推荐的“一般原理”法进行教学的课程比较却更为满意。

需要提及的是 UNIX作者提供的有关文献资料。我们把所复制并用于校园内的 UNIX文献资料装订成两册。

关于该资料的第一个想法是，整个资料能放在学生的书包中，而且以清晰、简洁、不墨守成规的风格撰写的这些文献资料基本上可敷需要。

在得到了足够经验后，我们的第二个想法是，这些文献在用作参考资料时是相当全面和完整的。但是，还留有相当余地可以编写一些引导性的资料，两本分析的编写目的就是为了填补其中一部分空间。

实际的UNIX操作系统源代码的名称是“UNIX Operating System Source Code”，它在1976年7月第一次印刷。这是对我们1975年12月收到的UNIX 版本6代码进行特殊编辑处理后构成的。

在1976年，本分析的原始版本首先以复写方式散发，这一年的后期，nroff文本格式化程序开发成功，借此机会我们对原稿进行了修改和更正，并用 nroff对其进行了格式化处理，最后得到了现在的形式。

关于源代码的编排顺序，必须尽早作出决定。我们当时的意图是为想学习 UNIX整个系统的学生提供一种合理的逻辑顺序。现在回想起来，当时的处理仍遗留很多可改进之处，我们将在以后的版本中考虑到这些并作修改。

我们希望本书对学习UNIX分时系统的学生是有益的。虽然本书的编写目的不是用作参考文献，但是某些读者可能为此目的而使用本书。本书的索引部分可在某种程度上满足此种要求。

因为本书涉及的 UNIX归属于 Western Electric 公司并受该公司的管理，所以只有具有 UNIX分时的许可证的读者才可使用本书，受此限制本书不能通过其他渠道出版。

非常欢迎读者指出本书的错误，进行批评和提出改进的建议。

致谢

本书的编写得到很多同事和学生的鼓励和支持，包括：David Carrington、Doug Crompton、Ian Hayes、David Horsfall、Peter Ivanov、Ian Johnstone、Chris Maltby、Dave Milway、John O'Brien和格里格·罗斯。

原稿打印主要是由 Pat Mackie和Mary Powter完成的，而用 nroff对原稿进行排版则得到 Adele Green的许多帮助。

David Millis和新南威尔士大学的出版部对本书的出版提供了重要支持，Ian Johnstone和澳大利亚研究生管理院为本书最后稿的形成提供了所需的设施。

在编写本书的全过程中，我的妻子玛丽安妮给予了我无限的支持，她校对了全书。

最后要感谢的是肯·汤姆森和丹尼斯·里奇，没有他们开创的 UNIX，就不会有本书。

对上面提及的诸位，我对他们表示诚挚的感谢！

也要提及 nroff 的重要作用，没有 nroff 就不会有本书现今的格式。但是 nroff 也有一些令人难以捉摸之处，对此很难进行处理。确实，nroff 本身必须提供一片沃土，使将来的程序文档处理人员或软件有用武之地。

约翰·莱昂

1977年5月

第1章 绪 论

UNIX是在PDP11计算机上运行的一个分时系统的名字，它由贝尔实验室的肯·汤姆森和丹尼斯·里奇编写。《ACM通信》的1974年7月号刊载了他们的论文，对该系统进行了说明。

事实说明UNIX功能丰富、高效和可靠，到1976年底为止，它已安装在150多台机器上并得到应用。

直至UNIX的版本6，编写UNIX共投入的人力约为10人年，与其他系统相比较，这是非常小的一个数字。（例如，到1968年，OS/360已投入5000人年，而另一个IBM操作系统TSS/360则投入了1000多人年。）

UNIX的强大能力主要来自于它的简练和直接，UNIX由两位在同一办公室工作的人员编写，他们用了一种称作为C的程序设计语言，还受到PDP11非常有限的存储容量和计算能力的限制。

不仅是UNIX的能力令人注目，其可理解性也是大多数其他系统难以比拟的。为了较深入地理解UNIX，必须掌握的资料数量并不太多。与之相比较，任一个人想要完全理解OS/360及其后继者都是不太可能的，因为它们过于复杂。在对任一主流操作系统作较大修改前，一个人需要先花费数月时间对该系统进行学习。

自然，有些系统比UNIX更易于理解，但是可以断言这些系统会简单得多，而且它们企图达到的目标也非常有限。如果关心的是向用户提供的功能类型和数量，那么UNIX就位于其中。事实上，UNIX提供了很多某些非常著名的主流操作系统所缺少的功能。

1.1 UNIX操作系统

《莱昂氏UNIX源代码分析》的目的是详细地说明我们从UNIX分时系统中选出的部分，这部分代码在UNIX运行时常驻内存，我们将其称作为“UNIX操作系统”。这部分代码提供下列功能：

- 系统初启。
- 进程管理。
- 系统调用。
- 中断处理。
- 输入/输出操作。
- 文件管理。

1.2 公用程序

UNIX的其余部分(它比内核大得多!)由一组剪裁适当的程序组成，它们以“用户程序”方

式运行，常被称为“公用程序”。

在此标题下，有一组与操作系统关系特别紧密的程序，例如：

“shell”（命令语言解释程序）

“/etc/init”（终端配置控制程序）

以及一组文件系统管理程序，例如：

check	du	rmdir
chmod	mkdir	sync
clri	mkfs	umount
df	mount	update

应当指出的是：由上述程序执行的很多功能在其他计算机系统中是由操作系统提供的，而且在操作系统中占了相当大的比重。

上述程序的功能性说明及使用方法包含在《UNIX程序员手册》（UNIX Programmer's Manual）的UPM的第一部分（公共使用的程序）和第三部分（仅由系统管理员使用的程序）中。

1.3 其他文档

本书中的注释频繁地引用《UNIX程序员手册》（UPM），偶而也引用《UNIX文档》（UNIX Documents）手册，而对《UNIX操作系统源代码》（UNIX Operating System Source Code），则几乎时时处处都要引用。

所有这些文档都关系到对UNIX系统的完整理解。另外，为了学习汇编语言例程，需要参考DEC出版的《PDP11处理机手册》（PDP11 Processor Handbook）。

1.4 UNIX程序员手册

《UNIX程序员手册》（UPM）分成八个主要部分，在它们之前还有一个目录和一个KWIC(Key Word In Context,上下文中的关键词)索引。KWIC是极其有用的，但有时也使人烦恼，某些被索引到的材料并不存在，而某些现存的材料却没有索引到。

在该手册的每一部分中，材料按主题名的字母顺序编排。在主题名后按惯例附加了它在部分的编号，某些主题会出现在几个部分中，例如，“CHDIR()”和“CHDIR()”。

第一部分(Section)包含了各种命令，它们或者是由“shell”命令解释程序所识别的，或者是标准用户公用程序的名字。

第二部分(Section)包含“系统调用”(system calls)。系统调用是用户程序可以调用以获得操作系统服务的操作系统例程。学习UNIX后，我们会对大多数系统调用相当熟悉。

第三部分(Section)包含“子例程”(subroutines)。这些子例程是可由用户程序调用的库例程。对一般程序员而言，第二和第三部分的区分并不很重要。第三部分的大多数与操作系统并无关系。

第四部分(Section)说明“特殊文件”(special files)，在UNIX中将外部设备称为“特殊文件”。某些特殊文件是相关的，有些则关系不大。这取决于你所在的位置。

第五部分(Section)说明“文件格式和约定”(File Formats and Conventions)。大量与理解UNIX紧密相关的信息都可在这一部分中找到。

第六部分(Section)和第七部分(Section)说明“用户维护的”程序和子例程。UNIX的爱好者不会忽略这两个部分,但是这两部分和操纵系统并没有密切的联系。

第八部分(Section)说明“系统维护”(软件部分,不包括硬件部分)。如果你对如何管理UNIX特别有兴趣,那么这一部分就包含了大量你所需的信息。

1.5 UNIX文档

本文档手册是相关程度不等的文集,其主要内容包括:

- 构造和安装UNIX。实际上属于UPM的第八部分(这部分内容是有关的)。
- UNIX分时系统。是原先发表在《ACM通信》上论文的修改版。这篇文章极其重要,应当每月至少重读一次。
- UNIX入门。如果你在UNIX方面缺少经验,那么这是非常有用的。
- 除非你是C语言等方面的专家,那么“C”语言导引、编辑程序以及“C”语言和汇编语言的参考手册都是非常有用的。
- UNIX I/O系统。提供了UNIX系统很多特征和功能的综述。
- UNIX摘要。提供了一张检查表,该检查表在回答“一个操作系统是做什么的?”此类问题是有用的。

1.6 UNIX操作系统源代码

这是对Bell实验室所提供的UNIX系统代码经编辑处理后形成的一个版本。

该代码假定的硬件“模型”系统包括:

PDP11/40处理机。

RK05磁盘驱动器。

LP11行式打印机。

PC11纸带阅读/穿孔机。

KL11终端接口。

对Bell实验室提供的源代码所作的主要更改是:

- 各文件的编排顺序已经更改。
 - 在许多文件中材料的先后顺序已经更改。
 - 代码在文件之间作了极少量的迁移(现在回顾起来当时)
 - 大约5%的代码行以各种方式缩短成少于66个字符(删去)
 - 增加了一些由下划线组成的行,特别在程序过程的结
 - 每个文件的长度都调整为50行的整数倍,原先文件的
 - 每行的行首加了一个4位数标识该行的编号,以便于:
- 在本书的开始部分有若干摘要性内容,包括:

- UNIX操作系统过程分类索引:其中各过程都带了行号
- UNIX操作系统文件及过程:显示各文件及文件中包
- UNIX操作系统定义的符号列表:各定义符号都带它

- UNIX操作系统源代码交叉引用列表：给出了每个符号引用处的行号。其中省略了“ C ”的保留字和若干通用符号，例如“ p ”、“ u ”。

1.7 源代码中各部分

源代码分成五部分，每个部分致力于系统的一个主要方面。

这种划分的意图是使每个部分尽可能自包含，无需掌握后继部分就能作为一个单元学习每一部分，这一目的已经基本达到。各部分的主要内容是：

第一部分：处理系统初启和进程管理。它也包含了所有汇编语言例程。

第二部分：处理中断、陷入(trap)、系统调用和信号(signal),信号又称为软件中断(software interrupt)。

第三部分：主要处理程序换进/换出(swapping)、基本磁盘操作，以及面向块的输入/输出。它也处理块缓冲池的资源管理和使用。

第四部分：处理文件和文件系统，包括文件创建、维护、使用和删除。

第五部分：涉及“字符特殊文件”(“character special files”)。UNIX将低速外部设备称为字符特殊文件，这类设备共用面向字符的缓冲池。

各部分的内容将在第4章中更加详细地列出。

1.8 源代码文件

上述五个源代码部分的每一个都由若干源代码文件组成。每个文件的名称都带一个标识

的文件。

语言语句的文件。

语句的文件，但该文件并不单独进行编译，而需包含在其“ .h ”文件包含了全局说明。

源代码中的注释进行补充，它对理解 UNIX操作系统并不时可以不阅读本书，而且鼓励你尽可能这样做。

到难点而不能前进，那么你就可以把这本“分析”当作拐你完全独立地阅读每一个文件或过程，那么在开始阶段可阅读他人编写的程序是一种应当学习并加以实践的艺术，

注释

很高的水平。很多部分在你开始阅读时似乎很复杂也不清

晰，但在进一步加以分析和回顾后，你会发现它们的思路很完美，而且可能是唯一的编写方法。

由于这一原因，所以在本“分析”中偶而对程序设计风格所做的评论几乎都是针对一些稍稍偏离通常的高水平之处而言的。

造成这些小问题的原因是什么呢？部分原因是原先的代码加了补丁。有些我们找到的“坏”代码事实上往往与采用的巧妙方法有关连，而这种情况在开始阅读时并不能清楚地意识到。当然也确实具有可进一步斟酌之处。但是作为一个整体，你会发现 UNIX的作者创造了一个强有力的、有机结合成一个整体的、功能非常丰富的程序，对此应当赞美并加以仿效。

