

操作系统 之编程观察

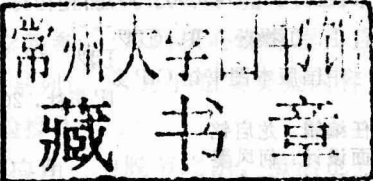
罗秋明 著



清华大学出版社

操作系统 之编程观察

罗秋明 著



清华大学出版社

北京

内 容 简 介

本书弥合了操作系统理论和操作系统编码实现之间的断层。以 Linux 真实操作系统为观察对象,主要利用 proc 文件系统展示出的内核行为数据来讲述操作系统的核心概念。第 1 章简单介绍 Linux (CentOS)和 VirtualBox 虚拟机的安装。第 2 章介绍进程实体并通过编程实践来观察进程创建及其组织关系;对比了进程和线程的资源开销差异;通过编程观察了孤儿进程、僵尸进程和守护进程。第 3 章讨论进程状态转换、CFS 进程调度和实时进程调度;观察 CFS 和实时进程的负载均衡行为。第 4 章学习和观察使用管道、消息队列、共享内存等通信手段;观察了 System V 信号量集和 POSIX 信号量对进程进行同步的活动。第 5 章通过 /proc/PID/maps 观察进程虚存空间,以及程序操作对虚存空间的改变;模拟 CPU 地址部件通过查找多级页表来完成地址转换的过程;观察进程的物理页帧分配、回收等行为。第 6 章对前面 5 章的知识进行一次整合,以 Shell 上执行一个可执行文件的全过程而将虚存空间管理、物理内存管理、用户空间与内核空间、中断与系统调用、进程调度等概念有机地联系起来。第 7 章在 VFS 通过文件模型之下观察文件的基本操作、目录操作、文件系统挂载/卸载等活动行为;通过编程观察页缓存和交换空间变化等。第 8 章针对 EXT2 具体文件系统,通过在一个块设备建立起 EXT2 文件系统,并通过设备的读写操作而完成文件数据和目录内容的读写。通过编程实例,将进程控制、进程线程差异、进程调度、进程通信、内存管理、文件系统和设备管理等关键内容,都以各种编程代码和内核数据的数值形式将操作系统内核行为直观地展现给读者。

本书不仅可以作为计算机及相关专业高年级本科生或研究生学习操作系统课程的教学用书,对希望深入了解 Linux 内核的读者和相关开发人员也非常有参考价值。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

操作系统之编程观察/罗秋明著. —北京:清华大学出版社,2018
ISBN 978-7-302-48973-3

I. ①操… II. ①罗… III. ①操作系统—程序设计 IV. ①TP316

中国版本图书馆 CIP 数据核字(2017)第 293290 号

责任编辑:龙启铭
封面设计:何凤霞
责任校对:李建庄
责任印制:刘祎淼

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>
地 址:北京清华大学学研大厦 A 座 邮 编:100084
社 总 机:010-62770175 邮 购:010-62786544
投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn
质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn
课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市铭诚印务有限公司

经 销:全国新华书店

开 本:185mm×230mm 印 张:20.5 字 数:410千字
版 次:2018年5月第1版 印 次:2018年5月第1次印刷
印 数:1~1500
定 价:59.00元

产品编号:076707-01

序

PREFACE

Preface

序

深

圳大学计算机与软件学院正在进行教学改革,基于明仲教授和王志强教授两位领导的构想,计算机与软件学院参照美国纽约宾汉姆敦大学教学要求,将其课程在深圳大学原样重现。对计算机系统系列课程和操作系统课程都进行了改革,其中操作系统课程缩减了理论授课,增加了实验操作环节。同时深圳大学计算机与软件学院计算机系统课程组正在承担广东省教育厅应用型人才培养项目(计算机系统系列核心课课程),强调丰富的动手实践经验并提高系统能力、系统思维。在上述环境下,为了充实实验内容,保持学生在课程学习中全程充实,我们完成了本书以配合理论教学。

在过去的操作系统教学过程中,作者深感理论教学与实践的脱节之困。虽然也有老师讲授 Linux 工程实践和系统编程,从而拉近了两者的距离,但都还未能与操作系统的核心概念紧密联系,仍似隔靴搔痒,可望而不可即。另外也考虑过将 Linux 内核源码分析或增强作为实践内容,但是由于学习曲线过于陡峭需要花费太多的时间,并不太适合作为课程内容——难以在一个学期课程中结合进来,甚至还可能让学生产生无功而返的挫败感。

根据我们在个人高性能计算机(PHPC)系统研制过程中对研究生培养的经验积累,将 Linux 系统编程的基础知识结合 Linux 的内核行为观察,利用 /proc 文件系统中探测到的内核数据,以及其他各种工具收集的内核数据,直观生动地将进程与内核的交互、内核的行为展示给学生,获得非常好的学习效果。学生对进程行为、内存分配管理、进程间通信和文件系统等各方面的认知,都远比传统的操作系统课程教学效果好。

读者在学习和体验操作系统各种概念的同时,也获得了初步的系统编程实践锻炼,并为进一步阅读 Linux 内核源代码做好了充足的准备。将 proc 文件系统和相关工具加入到

操作系统的学习过程中,相当于有了电路系统课程中的“万用表、示波器和逻辑分析仪”等工具,有了观测工具后,操作系统的教学和实验才算基本成熟了。正因为这些观测工具,使得一些读者在完成全书学习后可能会觉得:“哦,原来这才是操作系统!”经过这样的实践锻炼后,读者不仅可以在后续学习中加快系统编程的学习进度,还可以加快获得分析和修改 Linux 内核代码能力的培养进程。

本书是作者这些年在本科教学和研究生培养过程中积累的经验汇总,相信读者会喜欢。

限于作者的水平和能力,书中会有疏漏,欢迎读者指正,联系邮件: lqm@szu.edu.cn。

编者

2018年4月

新

感

谢深圳大学计算机与软件学院操作系统课程组的各位老师,大家一起完成了操作系统实验课程的改革,特别是张滇和周明洋两位老师在相关实验内容的检查和教学工作中做出了极大的贡献。

还需要感谢 2014 级几位同学在相关的材料整理和实验代码的设计中做出的贡献。其中林润胜同学完成了第 5 章的匿名映射、文件映射、meminfo 和 zoneinfo 解读及相关代码,第 8 章的 EXT2 文件系统中文件内容读取、目录读取及相关代码,协助完成了第 6 章的勘误;张永昌同学完成了 5.1.1 节、7.3.3 节的内容和相应的代码。这两位同学一起提供了 5.3.3 节的部分内容,完成了第 5 章内存管理和第 8 章 EXT2 文件系统的勘误工作。罗文杰同学完成了 4.2.1 节的信号量集相关代码和材料。2017 级研究生汤钊扬同学作为第一位读者,协助验证代码和完成勘误工作。

在上述老师和同学的大力支持下,本书终于完稿并与读者见面,再次对他们表示衷心的感谢!

^① 本书获得深圳市科创委基础研究 JCYJ20150930105133185 项目和 JCYJ20170302153920897 云环境中的异构存储资源分配与性能优化研究的资助。

CONTENTS

目 录

第 1 章 Linux 安装与访问	1
1.1 安装 Linux	1
1.1.1 下载 CentOS 7	1
1.1.2 CentOS 7 安装	3
1.2 虚拟机安装 Linux	9
1.2.1 VirtualBox 安装	9
1.2.2 虚拟机配置	17
1.2.3 虚拟机安装 Linux	19
1.3 ssh 远程终端访问	21
1.4 初次接触 Linux	24
1.4.1 简单操作	24
1.4.2 运行 HelloWorld 程序	29
1.4.3 操作系统接口	32
1.5 阅读注意事项	33
1.6 本章小结	34
第 2 章 进程控制	35
2.1 进程基本概念	35
2.1.1 进程实体	37
2.1.2 进程间组织关系	38

2.1.3	进程控制命令	42
2.2	创建与撤销进程	44
2.2.1	fork()创建子进程	44
2.2.2	孤儿进程和僵尸进程	47
2.2.3	exec 函数族	49
2.2.4	通过 kill()撤销进程	51
2.2.5	创建守护进程	51
2.3	创建 pthread 线程	56
2.3.1	进程与线程	57
2.3.2	创建方法	58
2.4	进程和线程资源开销	60
2.4.1	PCB 开销	60
2.4.2	内存描述符开销	63
2.5	本章小结	64
2.6	练习	65
第 3 章	进程调度	66
3.1	调度与均衡	66
3.1.1	调度与均衡框架	71
3.1.2	全系统的调度统计	72
3.2	进程状态及其转变	75
3.2.1	进程状态	75
3.2.2	状态转换	80
3.2.3	进程的调度统计	83
3.3	进程的调度	88
3.3.1	普通进程的 CFS 调度	88
3.3.2	实时进程调度	93
3.4	进程迁移与负载均衡	103
3.4.1	CFS 进程的负载均衡	104
3.4.2	实时进程的负载均衡	111
3.5	本章小结	114
3.6	练习	114

第 4 章 进程间通信与同步	115
4.1 进程间通信.....	115
4.1.1 管道.....	115
4.1.2 System V IPC.....	120
4.2 进程间同步.....	133
4.2.1 System V IPC 信号量集.....	133
4.2.2 POSIX 信号量.....	145
4.3 本章小结.....	154
4.4 练习.....	154
第 5 章 内存管理	156
5.1 虚存空间管理.....	156
5.1.1 进程映像.....	156
5.1.2 堆区.....	161
5.1.3 文件映射区.....	168
5.1.4 栈区.....	171
5.1.5 访问任意进程的虚存.....	176
5.1.6 虚存使用的物理页帧.....	179
5.2 分页机制与页表.....	181
5.2.1 分页机制.....	182
5.2.2 进程页表.....	186
5.3 物理内存组织管理.....	193
5.3.1 页帧、节点、内存域.....	193
5.3.2 空闲页帧管理——buddy 系统.....	200
5.3.3 物理内存分配与回收.....	204
5.3.4 内存回收.....	215
5.4 本章小结.....	227
5.5 练习.....	228
第 6 章 综合——新进程创建到运行	229
6.1 shell 读入命令.....	229
6.1.1 用户空间与内核空间.....	229
6.1.2 读入命令.....	231

211	6.2	创建进程	237
211	6.2.1	fork()复制进程	237
217	6.2.2	替换进程映像	238
251	6.2.3	开始运行新进程	239
221	6.2.4	进程映像与缺页	240
221	6.3	本章小结	240
231	6.4	练习	241
271			
	第 7 章	VFS 文件系统	242
	7.1	VFS	242
221	7.1.1	VFS 对象	243
221	7.1.2	文件系统类型	245
231	7.2	文件基本操作	246
241	7.2.1	命令行基本操作	246
241	7.2.2	编程接口	248
251	7.3	目录结构	253
251	7.3.1	树形结构	253
251	7.3.2	软/硬链接	254
281	7.3.3	文件系统创建与安装	258
281	7.4	页缓存	264
281	7.4.1	页缓存基本概念	265
281	7.4.2	页缓存动态变化	266
291	7.5	非文件功能	268
292	7.5.1	交换	268
292	7.5.2	设备接口	273
291	7.5.3	proc 文件系统	279
292	7.6	本章小结	280
292	7.7	练习	280
	第 8 章	EXT2 文件系统	281
292	8.1	EXT2 磁盘数据的组织	281
292	8.1.1	整体布局	281
292	8.1.2	超级块	283

8.1.3	块组描述符	285
8.1.4	索引节点	286
8.1.5	目录结构	289
8.2	EXT2 文件系统的创建	291
8.2.1	分配磁盘空间	291
8.2.2	创建环回设备	292
8.2.3	创建 EXT2 文件系统	293
8.2.4	安装文件系统	293
8.3	查看 EXT2 磁盘数据	294
8.3.1	布局信息	294
8.3.2	块组描述符	299
8.3.3	索引节点与文件内容	300
8.3.4	目录结构	304
8.4	本章小结	309
8.5	练习	310
附录	vi 编辑命令	311

本书以 CentOS 7 为例，为读者提供从获取下载到安装配置的全过程。不同发行版之间的章节有所不同，读者可以通过 CentOS 或百度等搜索引擎查找相应的安装方法。

1.1.3 下载 CentOS 7

读者可到 CentOS 官方网站(<http://www.centos.org>)下载如图 1-1 所示。



图 1-1 CentOS 网站

第 1 章

Linux 安装与访问

在接触 Linux 系统之前,需要先获得一个可访问的 Linux 系统。读者可以自己安装配置一个 Linux 系统(真实机器或虚拟机都可以),也可以让 Linux 系统管理员给你开设一个账号。如果读者已经使用过或安装配置过 Linux 系统,可以直接跳过本章内容,从第 2 章开始学习。

1.1 安装 Linux

这里以安装 CentOS 7 为例,为读者展示从映像下载到安装配置的各个步骤。不同的 Linux 发行版细节有所不同,读者可以通过 Google 或百度等搜索引擎查找相应的安装方法。

1.1.1 下载 CentOS 7

首先访问 CentOS 官方网站(<https://www.centos.org/>),如图 1-1 所示。

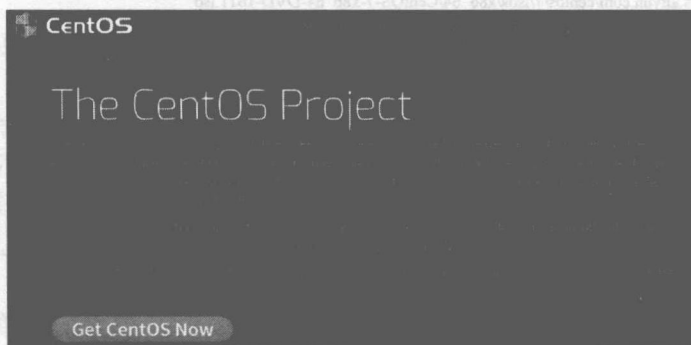


图 1-1 CentOS 官网

然后单击“Get CentOS Now”按钮,进入下载 ISO 镜像文件的选择页面,可以下载 DVD ISO、Everything ISO 和 Minimal ISO,如图 1-2 所示。这三者对操作系统的编程实践而言差别不大。另外网页上也提供了 Torrent 等其他下载方式的选择。



图 1-2 CentOS 的 ISO 镜像选择页面

接着单击“DVD ISO”按钮进入下载链接的选择网页(见图 1-3),从中选择一个速度较快的即可。

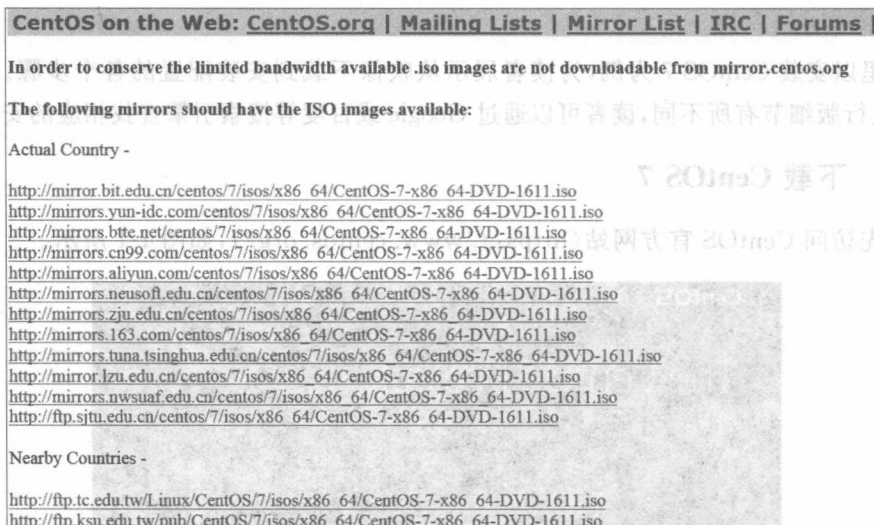


图 1-3 CentOS 7 x86-64 ISO 下载链接

单击其中一个链接将开始下载 CentOS 7 的 ISO 文件(约 4.7GB)。此时浏览器将弹出一个保存对话框(不同浏览器界面不同),如图 1-4 所示。

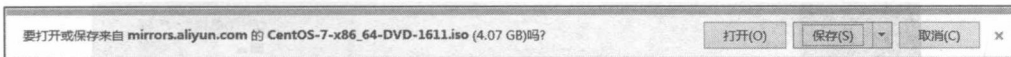


图 1-4 浏览器保存下载文件的对话框

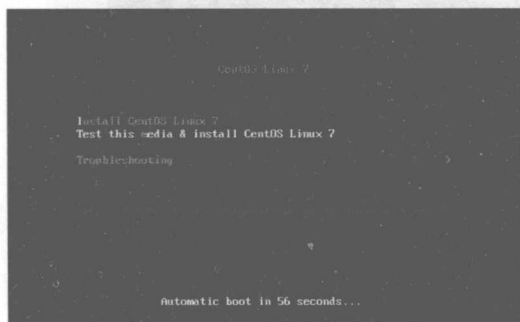
直接单击“保存”或“另存为”按钮将文件保存到本机的某个目录下。然后将 CentOS-7-x86_64-DVD-1611.iso 映像刻录成光盘启动、U 盘启动或通过 PXE 完成安装。

1.1.2 CentOS 7 安装

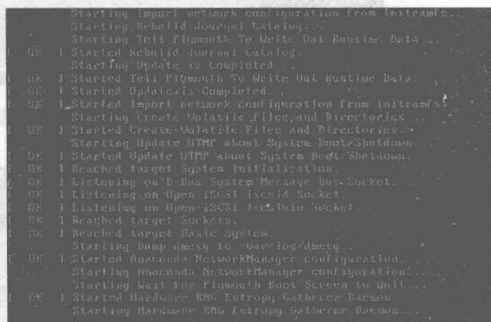
下面介绍 CentOS 7 的安装,图 1-5(a)是字符界面的安装,图 1-5(b)是图形界面的安装,读者可以根据自己的喜好选择安装方法。

1. 简易安装

将刻录好的 CentOS 7 安装启动盘插入光驱,重新启动计算机,将进入安装界面如图 1-5 所示。



(a)



(b)

图 1-5 CentOS 7 安装启动界面

选择第一项“Install CentOS Linux 7”并按 Enter 键,等待片刻将出现如图 1-6 所示的安装界面。

在里面选择“中文”→“简体中文”,然后再单击“继续”按钮进入下一窗口,如图 1-7 所示。

单击图 1-7 所示的“系统”→“安装位置”,进入后直接选择“完成”即可,此时图 1-7 的“开始安装”按钮变成可用状态。如果希望使用图形桌面,则需要单击“软件选择”进行必要的选择(参见图 1-12)。如果不需要图形桌面,则单击“开始安装”按钮进入下一步,如图 1-8 所示。

然后分别单击图 1-8 所示的两个带叹号的图标,在相应的对话框中进行 ROOT 用户密码设置和创建用户的设置,这两项工作完成后,图标上的叹号将消失。窗口的底端有安装进度条(如图 1-9 所示),完成后在进度条下方出现“重启”按钮。



图 1-6 语言选择



图 1-7 安装信息摘要

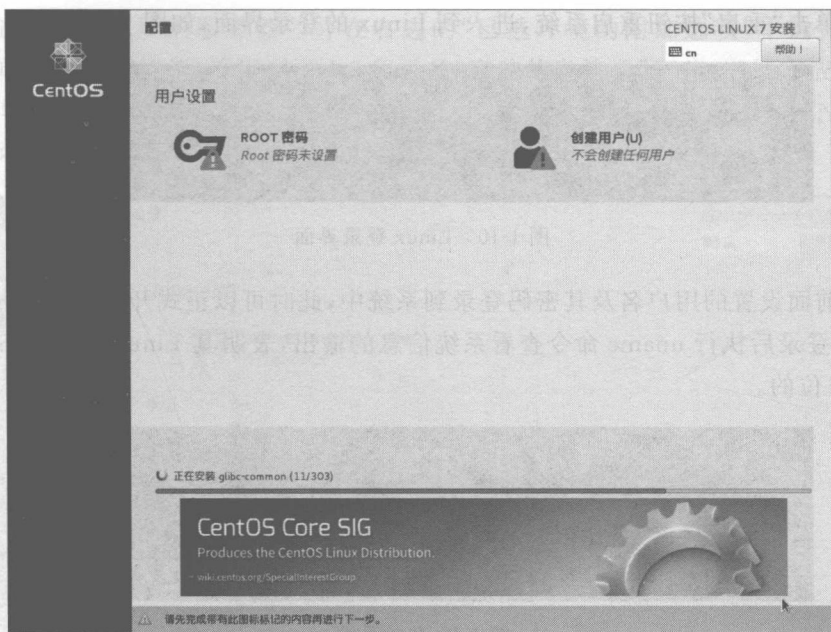


图 1-8 设置 ROOT 用户密码并创建普通用户

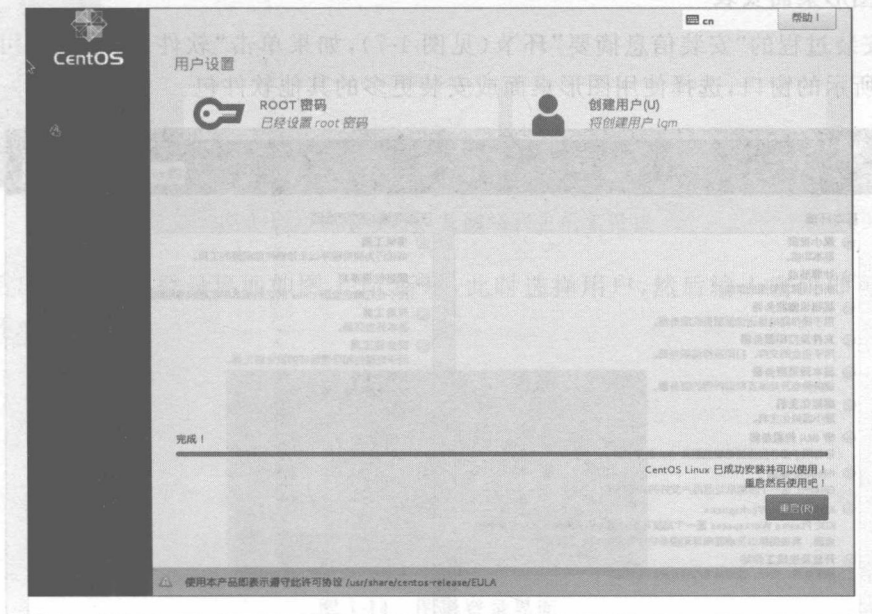


图 1-9 安装完成窗口

然后单击“重启”按钮重启系统,进入到 Linux 的登录界面,如图 1-10 所示。

```
CentOS Linux 7 (Core)
Kernel 3.10.0-514.el7.x86_64 on an x86_64

localhost login: _
```

图 1-10 Linux 登录界面

输入前面设置的用户名及其密码登录到系统中,此时可以正式开始使用 Linux 系统。图 1-11 是登录后执行 `uname` 命令查看系统信息的输出,表明其 Linux 内核是 3.10.0 版本、x86-64 位的。

```
CentOS Linux 7 (Core)
Kernel 3.10.0-514.el7.x86_64 on an x86_64

localhost login: lqm
Password:
lqm@localhost ~]$ uname -a
Linux localhost.localdomain 3.10.0-514.el7.x86_64 #1 SMP Tue Nov 22 16:42:43 UTC 2016 x86_64 x86_64 x86_64 GNU/Linux
lqm@localhost ~]$
```

图 1-11 登录系统

2. 图形桌面安装

在安装过程的“安装信息摘要”环节(见图 1-7),如果单击“软件选择”按钮可以进入图 1-12 所示的窗口,选择使用图形桌面或安装更多的其他软件包。

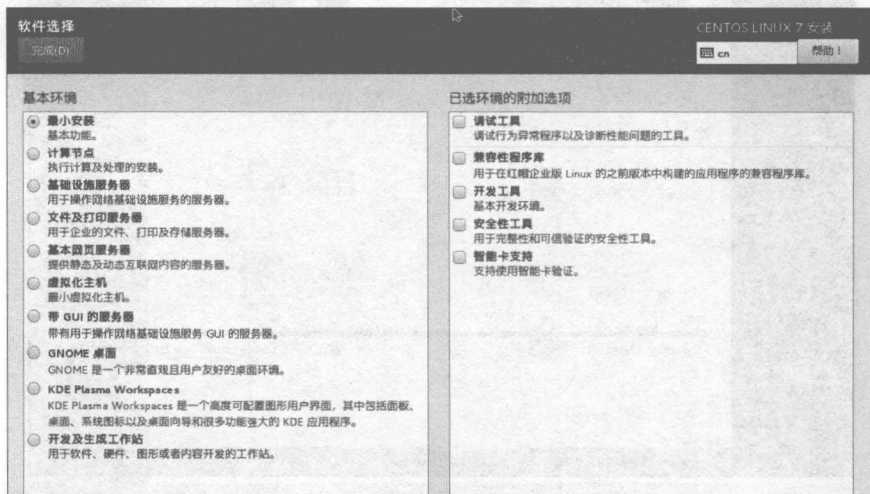


图 1-12 安装软件选择