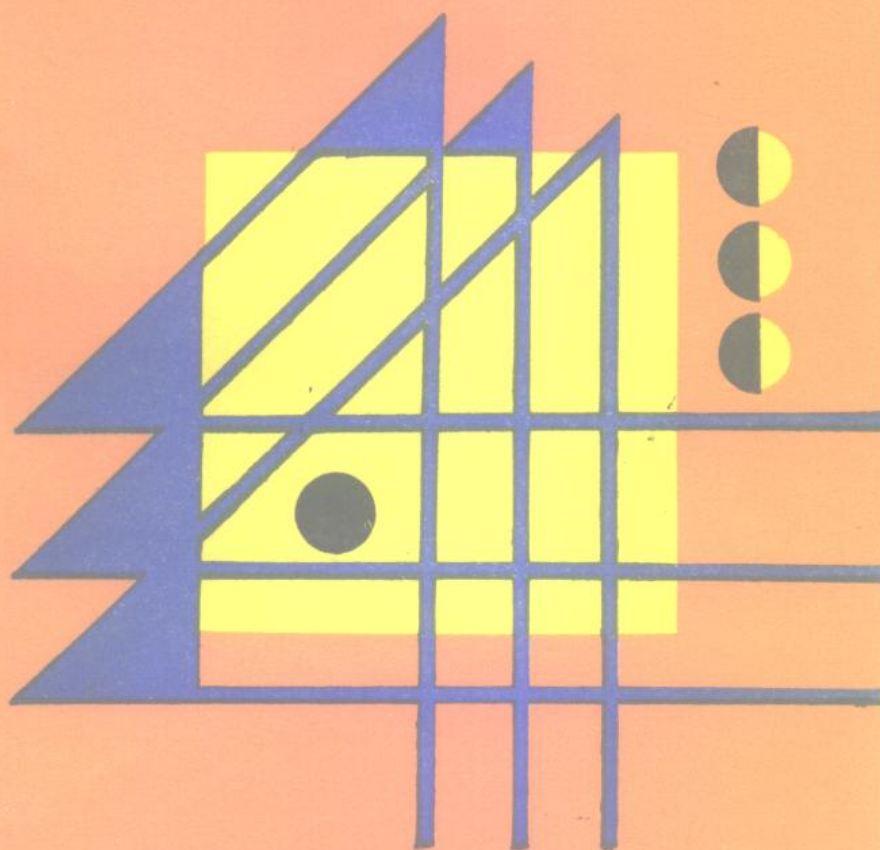


UNIX 丛书

UNIX 系统 工具与应用

孙玉方 主编

孙玉方 吴 健 周保刚 编著



清华大学出版社
广西科学技术出版社

UNIX 系统工具与应用

孙玉方 主编
孙玉方 吴 健 周保刚 编著

清华大学出版社
广西科学技术出版社

内 容 提 要

本书介绍 UNIX 系统的文件处理、程序开发、语言开发、调试等工具。内容包括：查看文件与分划文件；过滤程序与模式匹配；流编辑程序 sed；数据加工和检索工具 awk；C 语言检查工具 lint；系统构造工具 make；源码控制系统 SCCS；造型分析工具 prof 和 lprof；词法分析程序自动生成工具 lex；语法分析程序生成程序 yacc；宏处理程序 m4；目标调试工具 adb；源程序调试工具 sdb；计算器工具 dc 和 bc、文件归档工具 ar、tar 和 cpio 等。

(京) 新登字 158 号

(桂) 新登字 06 号

UNIX 系统工具与应用

孙玉方 吴 健 周保刚 编著

☆

清华大学出版社出版

北京 清华园

广西科学技术出版社出版

南宁市河堤路 14 号

煤炭工业出版社印刷厂印刷

新华书店总店科技发行所发行

☆

开本：787×1092 1/16 印张：22 字数：517 千字

1992 年 12 月第 1 版 1992 年 12 月第 1 次印刷

印数：00001—10000

ISBN 7-302-01049-8/TP·386

定价：12.60 元

引 论

从软件角度来说,UNIX 操作系统是开放系统的杰出代表。从其在 AT&T 贝尔实验室诞生以来,经过二十多年的发展已经成为当今世界上的主流操作系统,并被以此作为操作系统国际标准的基础。

学习并掌握 UNIX 系统的使用已成为研究和开发 UNIX 应用系统的必要手段。尽管学习和掌握 UNIX 系统的使用可以分不同的层次,从不同的角度出发。但行之有效的方法是:第一步学习 UNIX 的基本命令的使用和主要的编辑程序;第二步学习 shell 程序设计方法;第三步学习 UNIX 的一些开发和编程工具。作为软件开发人员可能还要学习 C 编程语言,或者更进一步地学习 UNIX 上的库函数和系统调用的编程方法。

在我们编著的这套丛书中,《实用 UNIX 系统》侧重向用户介绍 UNIX 的基本使用方法,包括基本命令和主要编辑工具的使用方法。《UNIX shell 程序设计》侧重向用户介绍 shell 编程方法。

本书是学习 UNIX 系统使用的第三步,它向使用和开发者提供文件处理、程序开发、语言开发、调试及其它一些有用的工具。

UNIX 系统的特点之一,就是它为软件开发提供了一个环境,配备了许多强有力的工具,为文件处理、程序开发以及调试等工作提供了方便。普通用户和开发者如能掌握这些工具,那么对他进行软件开发以及日常文件处理会有极大好处。

本书共五篇,分别描述五个方面的工具,介绍它们的特点及使用方法。

第一篇,文件处理工具,侧重于对文件内容的处理。其中:

第一章,简单的文件处理工具,介绍了十来个用于文件操作的工具,如查看文件头尾若干行来确定文件之内容;一屏一屏地翻看文件;多栏文件的“剪贴”;按模式分划文件;查看目标文件;多个文件的“联合”等。

第二章,过滤程序与模式匹配工具,包括文件内容的排序;查找两个文件之间的公共行;比较正文文件和二进制文件并作相应处理;去除文件中的重复内容以及若干种模式匹配工具。

第三章,流编辑程序 sed,重点介绍一种与行编辑程序 ed 有关又有区别的编辑程序,它是一种非交互的编辑程序,对于大文件,编辑命令序列太复杂的任务来说,它特别有用。

第四章,awk——数据加工和检索工具,它是一种完成与数据加工和信息检索有关任务的编程语言。它可以把存放在文件中的调查结果制成表格,打印汇总这些结果的各种报告,生成表格信件,把用于一个应用软件包的数据文件重新格式化后使之用于另一个应用软件包等等。

第二篇,程序开发工具,侧重于对程序开发工具的介绍。其中:

第五章,C语言检查工具,解释如何使用C语言程序检查程序 lint。尽管C编译程序能迅速而有效地编译C语言源程序文件,但不能执行许多程序要求的复杂的类型和出错检查。而 lint 检查C源文件,对于程序中的错误或误用提出警告,这些错误或误用均可能引起编译时错或运行时错。

第六章,程序自动维护工具,重点介绍如何使用 make 自动创建中大型程序的简便方法。解释了如何建立每个工程项目的描述文件,如何调用 make 创建程序和更新文件。

第七章,源代码控制系统,重点介绍一组用于对源码版本进行控制和管理的命令,这组命令统称为 SCCS。用 SCCS 可以建立、维护和控制称作 SCCS 文件的特殊文件。解释如何建立 SCCS 文件,如何更新放在 SCCS 文件中的文件及如何维护已创建了的 SCCS 文件。

第八章,造型工具,介绍两个主要的造型工具 prof 和 lprof。它们都用来分析程序的动态行为。具体说来,prof 报告程序的各部分在运行时所耗的时间总量占全部运行时间中的百分比,报告每个函数的调用次数及调用所需的平均执行时间。lprof 记录并报告逐行执行情况。

第三篇,语言开发工具,侧重介绍语言编译及类似软件开发过程中的若干开发工具。其中:

第九章,词法分析程序的生成程序,解释对字符输入流进行词法分析的软件的自动生成工具 lex。lex 接收对字符串匹配进行高层次的、面向问题的描述,生成可以识别正则表达式的C语言程序。重点介绍如何对问题进行描述。

第十章,编译程序的编译程序,解释语法分析程序的生成程序 yacc。yacc 接受一个语法规格说明,然后生成一个语法分析例程的C或 FORTRAN 源码,进而编译这个语法分析例程,然后和一个调用它以便对输入进行语法分析的程序组合在一起。

lex 和 yacc 组合在一起,可以完成编译程序主要部分的自动生成,大大提高了开发效率。

第十一章,m4——宏处理程序,解释了可用于C语言和汇编语言(及其它语言)预处理的通用的宏处理程序 m4。

第四篇,调试工具,侧重于介绍 UNIX 系统上的两种调试工具。其中:

第十二章,目标程序调试工具,介绍主要用于对目标程序进行调试的工具 adb,包括如何启动 adb,如何显示程序的指令和数据,如何运行、设置断点及单步执行一个程序及如何修补程序文件和内存。

第十三章,符号调试程序,介绍主要基于源程序的调试工具 sdb。它允许用户使用源程序中出现的符号名或行号来交互式地引用变量及程序语句。sdb 可以用来控制C语言程序的执行:允许执行用户的程序、设置断点、显示或设定变量值后继续执行程序,跟踪程序的执行情况等等。

第十四章,其它调试方法,介绍利用预处理程序调试和利用 ctrace 调试两种方法。

第五篇,其它工具,侧重介绍其它一些有用的工具。其中:

第十五章,计算器工具,介绍了两种计算器:dc 是一种交互式台式计算器,用来完成任意精度的整数运算,它处理的是逆波兰表达式。bc,它作为 dc 的预处理程序,处理用户所能理解的数学表达式。

第十六章,文件归档工具,介绍了档案建立和维护工具 `ar`,文件后援、存档及转入转出的工具 `tar` 和 `cpio`。

本书分篇对 UNIX 中的一些主要工具作了介绍,每篇有一重点,读者可以根据需要选读其中的一部分。这些工具都是以命令形式提供给用户的,在阅读本书时,请参考相应的“参考手册”,以求得掌握相应命令的全面性和完整性。实践出真知,只有自己用这些工具去编制和开发有关的程序或软件包,才能深刻理解并真正掌握这些工具的精髓。

本书的结构是:每篇前对本篇的主要几方面内容作一概述。每章一开始对本章内容作一说明,然后逐步展开。每章后有习题,可供读者作练习。

在全书正文后有一份参考文献,列出了与本书有关的一些参考资料。最后有一个附录,对 `awk` 语言的语法作了小结。

目 录

第一篇 文件处理工具.....	1
第一章 简单的文件处理工具.....	1
1.1 head 和 tail	1
1.1.1 head	1
1.1.2 tail	2
1.2 more 和 pg	3
1.2.1 more	3
1.2.2 pg	8
1.3 cut 和 paste	11
1.3.1 cut	11
1.3.2 paste	14
1.4 split 和 csplit	15
1.4.1 split	15
1.4.2 csplit	16
1.5 od	18
1.6 join	20
1.7 习题.....	22
第二章 过滤程序与模式匹配工具	23
2.1 sort	23
2.2 comm	26
2.3 diff、diff3 与 cmp	28
2.3.1 diff	28
2.3.2 diff3	33
2.3.3 cmp	35
2.4 uniq	36
2.5 grep	38
2.5.1 grep	39
2.5.2 egrep	41
2.5.3 fgrep	42
2.6 习题.....	43

第三章 流编辑程序 sed	45
3.1 起步	45
3.2 地址	46
3.3 功能	47
3.3.1 面向整行的功能	47
3.3.2 替换功能	48
3.3.3 输入输出功能	49
3.3.4 多输入行功能	49
3.3.5 保留功能与取回功能	50
3.3.6 控制流功能	50
3.3.7 杂项功能	51
3.4 sed 的用法示例	51
3.4.1 普通命令	51
3.4.2 从文件中取 sed 命令	52
3.4.3 附加或插入新的行	53
3.4.4 把输出写至文件	54
3.4.5 在 sed 中使用正则表达式	55
3.5 习题	56
 第四章 awk——数据加工和检索工具	 57
4.1 引言	57
4.2 基础	57
4.2.1 awk 程序的结构	57
4.2.2 用法	58
4.2.3 字段	59
4.2.4 显示	59
4.2.5 格式化后的显示	60
4.2.6 简单模式	61
4.2.7 简单动作	62
4.2.7.1 内部变量	62
4.2.7.2 用户定义的变量	62
4.2.7.3 函数	62
4.2.8 一组单行程序	63
4.2.9 出错信息	63
4.3 模式	64
4.3.1 BEGIN 和 END	64
4.3.2 关系表达式	65
4.3.3 正则表达式	66

4.3.4	模式组合	68
4.3.5	模式范围	68
4.4	动作	69
4.4.1	内部变量	69
4.4.2	算术运算	70
4.4.3	串和串函数	71
4.4.4	字段变量	74
4.4.5	数还是串?	75
4.4.6	控制流语句	76
4.4.7	数组	77
4.4.8	用户定义的函数	79
4.4.9	某些词法约定	80
4.5	输出	80
4.5.1	print 语句	80
4.5.2	输出分隔符	80
4.5.3	printf 语句	81
4.5.4	向文件输出	82
4.5.5	向管道输出	82
4.6	输入	83
4.6.1	文件和管道	83
4.6.2	输入分隔符	83
4.6.3	多行记录	84
4.6.4	getline 函数	84
4.6.5	命令行实参	86
4.7	与其它命令及 shell 一起使用 awk	86
4.7.1	system 函数	87
4.7.2	与 shell 合作	87
4.8	实例应用	88
4.8.1	生成报告	88
4.8.2	其它例子	90
4.8.2.1	词频	90
4.8.2.2	总计	90
4.8.3	随机选择	91
4.8.4	shell 设施	91
4.8.5	表格信件生成	91
4.9	习题	92

第五章 C 程序检查工具	93
5.1 概述	94
5.2 引用 lint 程序	95
5.3 检查语法错误	95
5.4 检查未使用的变量和函数	97
5.5 检查局部变量	100
5.6 检查不可到达的语句	101
5.7 检查类型说明、引用与转换上的一致性	102
5.8 检查函数定义与使用的一致性	104
5.9 检查程序的可移植性	109
5.10 检查奇怪的结构	110
5.11 检查陈旧的 C 文法	112
5.12 检查指针对齐	113
5.13 检查表达式的求值顺序	114
5.14 检查与库函数的一致性	114
5.15 建立 lint 程序库	115
5.16 综合范例	117
5.17 lint 命令选项和伪指令一览	121
5.18 习题	123
 第六章 程序自动维护工具	 124
6.1 概述	124
6.2 依赖关系	124
6.3 建立 Makefile 文件	126
6.4 使用 make	127
6.4.1 使用没有实参的 make	128
6.4.2 make 指定的目标	129
6.4.3 使用指定的 makefile 文件	129
6.4.4 make 的选项	130
6.4.5 内部伪目标名	131
6.5 make 的宏	132
6.5.1 使用宏	132
6.5.2 make 的内部宏	134
6.5.3 使用 shell 环境变量	135
6.6 内部规则	136
6.6.1 使用内部规则	136
6.6.2 修改内部规则	145
6.6.3 自定义新的内部规则	146

6.7 使用与维护档案库	148
6.8 使用 make 的高级课题	151
6.8.1 make 与 shell	151
6.8.2 出错处理	153
6.8.3 make 的嵌套使用	155
6.9 综合范例	156
6.10 习题	158
第七章 源码控制系统	159
7.1 概述	159
7.2 基础知识	160
7.2.1 文件和目录	160
7.2.2 delta 和 SID	161
7.2.3 SCCS 工作文件	162
7.2.4 文件管理员	163
7.3 建立和使用 s-文件	163
7.3.1 建立 s-文件	163
7.3.2 为读而复原文件	164
7.3.3 编辑和修改 SCCS 文件	165
7.3.4 使用多个版本	166
7.3.5 比较版本之间的差异	170
7.3.6 简化版本结构	170
7.3.7 显示 s-文件	171
7.3.8 显示 delta 版本的信息	174
7.3.9 包含与排除 delta	175
7.4 使用标识关键字	176
7.5 使用 s-文件标志	179
7.6 修改 s-文件信息	180
7.6.1 加入注释	181
7.6.2 修改注释	181
7.7 保护和修复 s-文件	182
7.7.1 保护 s-文件	182
7.7.2 检查 s-文件	183
7.7.3 修复 SCCS 文件	183
7.8 同时编辑一个 s-文件	184
7.9 SCCS 的求助功能	185
7.10 make 与 SCCS	186
7.11 习题	187

第八章 直方图工具	188
8.1 概述	188
8.2 怎样使用直方图工具	189
8.3 使用 prof 命令	190
8.4 用环境变量控制程序的运行	192
8.4.1 PROFOPTS 环境变量	192
8.4.2 使用 PROFOPTS	192
8.5 使用 lprof 命令	194
8.5.1 搜索源文件	195
8.5.2 lprof 的输出	195
8.5.3 突出不执行的行	196
8.5.4 总结报告	198
8.5.5 合并统计数据	198
8.6 对库函数进行直方图处理	199
8.7 改善测试的覆盖面	200
8.7.1 检查未列入文档的选项	201
8.7.2 检查永远不被调用的函数	201
8.7.3 检查难于形成的出错条件	202
8.8 综合范例	202
8.9 习题	208
第三篇 语言开发工具	209
第九章 词法分析程序的生成程序	210
9.1 概述	210
9.2 引用 lex	211
9.3 正则表达式	212
9.4 lex 的源	214
9.5 处理二义性规则	217
9.6 lex 动作	218
9.7 lex 应用	221
9.8 习题	225
第十章 编译程序的编译程序	227
10.1 概述	227
10.2 基本规范说明	229
10.2.1 动作	231
10.2.2 词法分析	233

10.3	文法分析程序的工作过程	235
10.4	二义性和冲突	238
10.5	优先级	241
10.6	准备规范说明	243
10.6.1	输入风格	243
10.6.2	左递归	244
10.6.3	词法上的注记	244
10.6.4	处理保留字	245
10.7	yacc 环境	245
10.8	两个例子	246
10.8.1	一个简单的例子	246
10.8.2	一个高级的例子	249
10.9	习题	257
第十一章	m4——宏处理程序	258
11.1	概述	258
11.2	引用 m4	259
11.3	m4 宏处理	260
11.3.1	定义宏功能	260
11.3.2	加引号	261
11.3.3	自变量	262
11.3.4	内部算术运算功能	264
11.3.5	文件管理	265
11.3.6	输出转换	265
11.3.7	系统命令	265
11.3.8	条件	265
11.3.9	字符串操作	266
11.3.10	打印	267
11.4	内部功能一览表	267
11.5	习题	268
第四篇	调试工具	269
第十二章	目标程序调试工具	269
12.1	概述	269
12.2	启动和停止 adb	270
12.3	显示指令和数据	271
12.3.1	形成地址	271
12.3.2	形成表达式	272

12.3.3	选择数据格式	273
12.3.4	使用 (?) 和 (/) 命令文件	273
12.4	调试程序执行	275
12.4.1	调试命令汇总	275
12.4.2	一个例子: 跟踪多个函数	276
12.5	使用 adb 内存映像	280
12.5.1	显示内存映像	280
12.5.2	改变内存映像和建立新的映像项	281
12.6	其它特点	281
12.6.1	建立 adb 命令文件	281
12.6.2	设置默认输入格式	281
12.6.3	使用 UNIX 命令	282
12.6.4	计算数值	282
12.7	修补二进制文件	282
12.8	习题	283
第十三章	符号调试程序	284
13.1	概述	284
13.2	引用 sdb	285
13.3	交互式命令	287
13.3.1	显示堆栈的跟踪数据	287
13.3.2	处理变量的命令	288
13.3.3	源文件显示及处理	291
13.3.4	调试一个活动的进程	292
13.4	sdb 的其它特性	298
13.5	习题	298
第十四章	其它调试方法	299
14.1	概述	299
14.2	利用 C 预处理程序调试	299
14.3	ctrace	301
14.4	习题	304
第五篇	其它工具	305
第十五章	计算器工具	305
15.1	概述	305
15.2	dc: 一个交互式的台式计算器	305
15.2.1	启动和退出	305

15.2.2	dc 栈操作命令	306
15.2.3	详细描述	307
15.3	bc: 一个任意精度的台式计算器	308
15.3.1	启动和退出	308
15.3.2	整数的简单运算	309
15.3.3	特殊内部量	310
15.3.4	函数	312
15.4	习题	314
第十六章	文件归档工具	315
16.1	建立和维护档案文件	315
16.1.1	向档案文件中增加文件	315
16.1.2	列出档案中的文件名	316
16.1.3	从档案文件中删除文件	317
16.1.4	显示或抽取档案中的文件	317
16.1.5	在 C 编译程序中使用档案库	317
16.1.6	ar 的选项	318
16.2	文件存档	319
16.2.1	建立一个新的档案	319
16.2.2	建立多卷档案	320
16.2.3	列出档案中的文件名	320
16.2.4	从档案中抽取文件	321
16.2.5	向档案中增加文件	322
16.2.6	tar 的选项	322
16.3	文件档案的输入和输出	324
16.3.1	建立 cpio 文件	324
16.3.2	从 cpio 档案中抽取文件	325
16.3.3	在目录间复制文件	326
16.3.4	cpio 的选项	326
16.4	习题	327
参考文献		328
附录 A awk 语法小结		329
A.1	命令行	329
A.2	模式	329
A.3	控制流语句	329
A.4	输入—输出	329
A.5	函数	330

A.6	串函数	330
A.7	数值函数	330
A.8	算符（优先级、递增）	330
A.9	正则表达式（优先级递增）	331
A.10	内部变量	331
A.11	限制	332
A.12	初始化、比较和类型强制	332

第一篇 文件处理工具

UNIX 系统提供一组用于检测文件内容的工具。这些工具具有某些共性，并且有许多不同的选项。它们可以读若干输入文件并把结果写至标准输出上。这些实用程序有足够的的能力应付小型数据库和文件的数据处理之需要。但是，它们只提供顺序访问，对于搜索操作来说，可能较慢。它们也不直接更改输入文件的内容，而只是把结果写至标准输出上。下面几章详细介绍这些工具。

第一章 简单的文件处理工具

本章向读者介绍一些基本的文件处理工具，这些工具在平时对 UNIX 系统的文件和目录操作时不一定常常用到，但是在对文件进行格式处理，搜索特定文件模式或文件某些部分等情况下却特别有用，所以我们在介绍这些工具时给出了具体而又详细的实用例子，以便读者在今后具体处理文件时可以直接加以利用。

1.1 head 和 tail

有时我们遇到一些大文件，但并不打算处理其中的内容，而只是想通过显示一些特定行来了解其大致内容或只是判定一下它是不是自己想找的文件，这时可以用 head 和 tail。head 是显示一个文件开头的若干行，tail 是显示一个文件尾部的若干行。

1.1.1 head

head 的格式为：

```
head [-count] [file...]
```

它显示指定的每个文件 *file* 的前 *count* 行。文件 *file* 可以有若干个。若未指定任何文件 *file*，则 head 从标准输入（一般即为键盘）上读取内容。若未指定 *count*，则默认为 10，即显示前 10 行。下面是一些例子。先假设有一个文件 *file2*，可用 cat 显示其全部内容：