

LINUX

Linux 平台下的 Web 编程

农孙博 范正刚 编著



人民邮电出版社
www.pptph.com.cn

Linux 平台下的 Web 编程

农孙博 范正刚 编著

人民邮电出版社

图书在版编目 (CIP) 数据

Linux 平台下的 Web 编程/农孙博, 范正刚编著. —北京: 人民邮电出版社, 2000.9

ISBN 7-115-08715-6

I.1.... II.(1)农...(2)范... III.操作系统 (软件), Linux...程序设计 IV.TP316.89

中国版本图书馆 CIP 数据核字 (2000) 第 12232 号

内 容 提 要

本书主要讲述了 Linux 平台下的 Web 开发技术。主要内容包括: Linux 和 Web 的基本概念; Linux 平台下 Web 服务器的安装和配置; Linux 支持的多种编程语言, 如 Shell、Perl、C/C++、PHP3 等; Linux 平台下 Web 数据库的开发等。最后一章还给出了一个综合实例: 图书销售系统的实现。

本书主要读者对象为设计制作 Web 页面和创建 Web 站点的 Web 技术人员。

JS287/13

Linux 平台下的 Web 编程

- ◆ 编 著 农孙博 范正刚
责任编辑 梁 凝
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@puph.com.cn
网址 <http://www.puph.com.cn>
北京汉魂图文设计有限公司制作
北京鸿佳印刷厂印刷
新华书店总店北京发行所经销
- ◆ 开本: 787 × 1092 1/16
印张: 21.75
字数: 528 千字 2000 年 10 月第 1 版
印数: 1 - 6 000 册 2000 年 10 月北京第 1 次印刷

ISBN 7-115-08715-6/TP·1771

定价: 32.00 元

前 言

据统计，目前 Internet 上已有几十万个 Web 站点，至于访问 Web 的用户则更是多得难以计数。由于 Web 拥有大量的用户群，因此建设 Web 站点对于一个机构或个人的发展具有十分重要的作用。Web 问世之后的最初几年，建立 Web 站点的大多数单位还是教育、科研机构，而近几年建立 Web 站点的机构已扩展到了金融、商业、新闻出版、娱乐、体育等各行各业。在这种形式下，学习和掌握 Web 技术，就成为越来越多的人的迫切要求和强烈愿望。

Linux 是一种全新概念的操作系统。它派生自 UNIX，但与“父辈”不同的是，在 Linux 的设计中简化了传统 UNIX 操作系统的部分功能，因而 Linux 具有对硬件配置的要求低，占用内存小，便于使用等特点，与流行的 MS-Windows 操作系统相比，Linux 的运行更加稳定、可靠，效率也更高。Linux 的应用和发展非常迅速，在短短几年中它的用户数骤增，并已经在操作系统的市场上占有了一席之地，而且展示出了自己的活力，有着很好的发展前景。

本书选择了很有活力的 Linux 作为开发平台，结合 Linux 本身的特点来介绍 Web 的开发技术。它所针对的读者不是普通的 Web 用户，而是针对设计制作 Web 页面和创建 Web 站点的 Web 技术人员。本书覆盖面广、技术新颖、实例丰富，是一本很好的 Web 技术参考书。衷心希望读者能从本书获得收益，并在学习和掌握 Web 技术的过程中享受无穷的乐趣！

本书由农孙博、范正刚执笔编写，吴集光、农孙雄参加了资料收集与整理方面的工作，并提出了许多忠恳的建议。高波负责全书的策划和统稿，并给予作者多方面的支持和协助，在此深表感谢。

由于作者水平所限，加之时间仓促，书中谬误和不当之处在所难免，恳请读者和有关专家批评指正。

作者

2000 年 5 月于北京

目 录

第一章 Linux 的基础知识	1
1.1 Linux 简介	1
1.1.1 Linux 的历史	1
1.1.2 Linux 系统	1
1.1.3 Linux 的主要特点	2
1.2 Linux 的主要版本	3
1.2.1 RedHat	3
1.2.2 Slackware Linux	4
1.2.3 Debian GNU/Linux	4
1.3 Linux 常用的应用程序与开发工具简介	4
1.3.1 应用程序(Applications)	4
1.3.2 开发(Development)工具包	7
1.4 Linux 平台下 Web 编程的特点	12
1.5 小结	13
第二章 Web 概述.....	14
2.1 万维网 (World Wide Web).....	14
2.1.1 Web 的由来及现状	14
2.1.2 Web 的特点	15
2.2 客户/服务器模式	15
2.3 URL	16
2.4 超文本传输协议(HTTP)	16
2.5 TCP/IP 协议	17
2.5.1 TCP 协议	18
2.5.2 IP 协议	18
2.6 HTML	18
2.7 Web 浏览器	18
2.8 小结	19
第三章 Web 服务器的安装和配置	20
3.1 Apache 服务器简介	20
3.2 Apache 服务器的安装	21
3.2.1 如何获得 Apache Server	21
3.2.2 编译源程序	21
3.2.3 Apache 服务器的目录结构	22

3.2.4 Apache 服务器的启动方法和主要选项	22
3.3 Apache 服务器的配置	23
3.3.1 httpd.conf	24
3.3.2 access.conf	36
3.3.3 srm.conf	38
3.4 小结	38
第四章 公共网关接口 (CGI)	39
4.1 通用网关接口 (CGI)	39
4.1.1 CGI 的含义	39
4.1.2 使用 CGI 的理由	39
4.1.3 CGI 的工作过程	40
4.2 程序间的通信方式	40
4.2.1 从客户机到服务器的通信方式	40
4.2.2 从服务器到 CGI 程序的通信方式	41
4.2.3 CGI 程序返回数据的方式	41
4.3 环境变量	41
4.3.1 服务器专有环境变量	41
4.3.2 请求专有环境变量	42
4.3.3 客户专有环境变量	43
4.4 数据传送方法	43
4.4.1 传送方法	43
4.4.2 GET 方法	44
4.4.3 POST 方法	44
4.4.4 实现方法	44
4.5 CGI 脚本使用的数据类型	44
4.5.1 MIME 的含义	44
4.5.2 MIME 数据类型方案简介	44
4.5.3 常用 MIME 数据类型	45
4.6 Linux 上 CGI 编程语言简介	47
4.6.1 Perl	47
4.6.2 C/C++	48
4.6.3 Shell Script	48
4.6.4 PHP3	48
4.6.5 TCL/TK	48
4.7 小结	48
第五章 表单的实现	49
5.1 表单	49
5.2 INPUT	50

5.2.1 TEXT 输入域	50
5.2.2 PASSWORD 输入域	50
5.2.3 RADIO	50
5.2.4 CHECKBOX 输入域	51
5.2.5 FILE	51
5.2.6 SUBMIT 和 RESET	51
5.2.7 例子	51
5.3 SELECT 和 OPTION	53
5.4 TEXTAREA	54
5.5 小结	56
第六章 Shell 语言简介	57
6.1 Shell 概述	57
6.1.1 Shell 基础	57
6.1.2 Shell 的种类	57
6.1.3 Shell 的基本功能	58
6.1.4 Shell 的特点	58
6.2 Bash 的基础知识	58
6.2.1 Bash 的特点	59
6.2.2 Bash 基本语法结构	59
6.2.3 Bash 的环境变量	65
6.3 小结	66
第七章 用 Shell 编写 CGI 程序	67
7.1 使用 Shell 编写 CGI	67
7.1.1 Shell CGI 的优点	67
7.1.2 CGI 对 Shell 的要求	67
7.1.3 Shell CGI 的使用范围	67
7.2 Shell 如何完成 CGI 的功能	67
7.3 一个简单的例子	68
7.4 显示环境变量	69
7.5 表单的处理	70
7.5.1 思路	71
7.5.2 调查表的表单	71
7.5.3 表单的处理	73
7.6 计数器	76
7.6.1 思路	76
7.6.2 计数器的实现	76
7.7 Shell CGI 的缺点	77
7.8 小结	78

第八章 Perl Script 简介	79
8.1 Perl 语言简介	79
8.1.1 Perl 的主要用途	79
8.1.2 一个简单的例子	80
8.1.3 程序的运行	80
8.1.4 程序的调试方法	80
8.2 标量	81
8.2.1 标量的概念	81
8.2.2 标量值	81
8.2.3 标量变量	83
8.2.4 标量变量操作符	83
8.2.5 操作符优先级	87
8.3 数组变量	88
8.3.1 数组变量的赋值	88
8.3.2 数组变量元素访问	89
8.3.3 数组变量运算符	89
8.4 控制结构	90
8.4.1 if/unless	90
8.4.2 while/until	91
8.4.3 for 语句	92
8.4.4 foreach 语句	92
8.4.5 die()	92
8.4.6 last 语句退出循环	92
8.5 关联数组	93
8.5.1 关联数组的变量	93
8.5.2 关联数组操作符	93
8.6 基本 I/O 处理	94
8.6.1 从 STDIN 输入	94
8.6.2 从 <> 操作符输入	94
8.6.3 向 STDOUT 输出	95
8.7 子程序	95
8.7.1 子程序的调用	95
8.7.2 子程序返回值	96
8.8 内部函数	96
8.9 文件处理	99
8.9.1 打开文件	99
8.9.2 读文件	99
8.9.3 写文件	100
8.9.4 关闭文件	100

8.9.5 检测文件状态	100
8.10 Perl 的进程管理	100
8.11 包	101
8.11.1 包的基本概念	101
8.11.2 作用范围	101
8.11.3 符号表	101
8.12 模块	102
8.12.1 模块的定位	102
8.12.2 模块引用	103
8.13 小结	103
第九章 用 Perl 编写 CGI 程序	104
9.1 为什么使用 Perl	104
9.2 处理表单信息	104
9.2.1 用 Perl 进行表单的处理	105
9.2.2 表单处理实例	106
9.3 简单例子——访问计数器	112
9.3.1 思路	112
9.3.2 计数器的实现	112
9.4 来宾留言板的实现	114
9.4.1 思路	114
9.4.2 留言板的表单	114
9.4.3 表单的处理	116
9.5 聊天室	119
9.5.1 聊天室的执行过程	119
9.5.2 连接的方法	119
9.5.3 交谈服务器的实现	120
9.5.4 用户信息处理	121
9.5.5 一个简单的聊天室	122
9.6 cgi-lib.pl	128
9.7 CGI.pm	129
9.7.1 如何使用 CGI.pm	129
9.7.2 引用 CGI.pm 的方法说明	129
9.7.3 CGI.pm 方法的使用	130
9.7.4 常用的 CGI.pm 方法	133
9.7.5 CGI.pm 中的 HTML 标记	135
9.7.6 CGI.pm 中的表单处理	136
9.8 小结	140
第十章 C/C++ 的 CGI 开发技术	141

10.1 C/C++ 支持 CGI 的原因	141
10.1.1 C/C++ CGI 的优点	141
10.1.2 C/C++ 的公共库	141
10.2 用 C/C++ 编写 CGI 程序	142
10.2.1 CGI 的输入输出	142
10.2.2 使用 C/C++ 的 API	142
10.2.3 简单例子: hello world!	142
10.3 表单的处理方法	143
10.3.1 数据的解码	144
10.3.2 数据的输出	147
10.4 用 C 语言进行表单处理	147
10.5 电子信箱	152
10.5.1 思路	152
10.5.2 表单内容的实现	152
10.5.3 发信服务的实现	153
10.6 小结	156
第十一章 PHP3	157
11.1 PHP3 概述	157
11.1.1 PHP3 的历史	157
11.1.2 PHP3 的特点	157
11.2 PHP3 的安装和配置	158
11.2.1 PHP3 的安装	158
11.2.2 PHP3 的配置	159
11.2.3 PHP 脚本的执行	159
11.3 PHP3 的语法和结构	160
11.3.1 PHP3 的定界符和分隔符	160
11.3.2 变量	160
11.3.3 类	161
11.3.4 变量作用域	162
11.3.5 类型转换	162
11.3.6 运算符	163
11.4 控制结构	164
11.4.1 while	164
11.4.2 if-elseif-else	164
11.4.3 FOR	165
11.4.4 SWITCH	165
11.5 自定义函数	165
11.5.1 函数的定义	165
11.5.2 返回值	166

11.5.3 参数	166
11.5.4 include 和 require	166
11.6 常用库函数	167
11.6.1 字符串操作函数	167
11.6.2 正则表达式函数	168
11.6.3 目录遍历函数	168
11.6.4 文件操作函数	168
11.6.5 时间函数	169
11.6.6 图像函数	170
11.6.7 数据库函数	171
11.7 PHP 的安全问题	174
11.8 小结	175
第十二章 动态创建图像	176
12.1 GD 库	176
12.1.1 GD 库的安装	176
12.1.2 简单例子	176
12.1.3 常用函数说明	178
12.1.4 C 语言 API 使用示例	180
12.1.5 GD.pm 使用示例	182
12.1.6 PHP3 API 使用示例	185
12.2 Ghostscript	189
12.2.1 计数器的图形化例子	189
12.2.2 Ghostscript 解释器和 Postscript 代码	193
12.3 小结	199
第十三章 Sever Side Includes (SSI)	200
13.1 SSI 概述	200
13.2 SSI 工作原理	200
13.3 为 SSI 配置 Apache 服务器	200
13.4 SSI 常用命令	201
13.4.1 #include 命令	201
13.4.2 #echo 命令	202
13.4.3 #fsize	203
13.4.4 #filemtime	203
13.4.5 #config	203
13.4.6 #exec	204
13.5 安全性	205
13.6 小结	205

第十四章 安全问题	206
14.1 Web 站点安全问题	206
14.2 Web 站点的保护对象	207
14.2.1 保护信息和资源	207
14.2.2 保护用户	207
14.2.3 保护重要的对象	207
14.3 Web 站点的安全隐患	208
14.3.1 Web 服务器安全	208
14.3.2 客户与服务器间的交易安全	208
14.3.3 错误与疏漏	208
14.3.4 危险的方式	209
14.3.5 不合理数据	209
14.3.6 内部伤害	210
14.4 如何保护 Web 站点	210
14.4.1 语言的选择	210
14.4.2 不要相信表单数据	211
14.4.3 拒绝不合理要求的表单数据	212
14.4.4 处理路径数据	213
14.4.5 校验数据长度	214
14.4.6 处理文件名	214
14.4.7 验证数据满足字段规则	215
14.4.8 处理外部进程	216
14.4.9 CGI 脚本权限	217
14.4.10 本地文件安全	217
14.4.11 使用他人 CGI 脚本时的注意事项	217
14.5 安全性评定	217
14.6 小结	218
第十五章 Web 数据库及其工具	219
15.1 数据库的基本知识	219
15.1.1 关系数据库	219
15.1.2 SQL(Structured Query Language)结构化查询语言	220
15.1.3 SQL 的优点	220
15.2 Web 数据库的选择	221
15.2.1 Linux 下常用的关系数据库系统简介	221
15.2.2 数据库方案分析	223
15.3 MySQL	224
15.3.1 MySQL 概述	224
15.3.2 MySQL 的主要特点	224

15.3.3 获取 MySQL	225
15.3.4 MySQL 的安装	225
15.3.5 MySQL 的启动方法	226
15.3.6 MySQL 的系统管理	227
15.3.7 MySQL 的使用	229
15.3.8 MySQL 的客户端工具和 API 接口函数	230
15.3.9 MySQL 的稳定性	231
15.4 PostgreSQL 数据库	232
15.4.1 PostgreSQL 数据库简介	232
15.4.2 PostgreSQL 的特点	232
15.4.3 PostgreSQL 的安装和配置	233
15.4.4 PostgreSQL 的启动	234
15.4.5 PostgreSQL 的使用	235
15.4.6 Postgres 的客户端工具和接口简介	235
15.4.7 PostgresSQL 的安全机制	237
15.5 Web 数据库的性能分析	238
15.6 小结	238
第十六章 使用 DBM 数据库	239
16.1 DBM 数据库的不同版本	239
16.1.1 DBM	239
16.1.2 NDBM	239
16.1.3 Berkeley Db	239
16.1.4 GDBM	240
16.2 DBM 数据库的使用	240
16.3 DBM 数据库的信息检索	243
16.4 DBM 数据库的信息查询	246
16.5 小结	250
第十七章 Web 数据库的开发	251
17.1 方案的选择	251
17.2 一个简单数据库查询的工作过程	252
17.3 MySQL 和 PHP3 的组合	252
17.3.1 准备工作	252
17.3.2 连接数据库	253
17.3.3 数据库操作	254
17.3.4 数据操作	254
17.3.5 一个典型的例子	254
17.3.6 综合实例: phpMyAdmin 分析	257
17.4 PostgreSQL 和 Perl 组合	267

17.4.1 准备工作	268
17.4.2 连接数据库	269
17.4.3 数据库操作	269
17.4.4 获取查询结果	269
17.4.5 实现查询的例子	270
17.4.6 关于数据安全的讨论	271
17.5 DBI	272
17.5.1 DBI 的标准方法	273
17.5.2 连接数据库驱动	274
17.5.3 执行 SQL 语句	274
17.5.4 结果处理	275
17.5.5 典型例子	276
17.6 各种组合的比较	277
17.6.1 C 语言	277
17.6.2 Perl 语言	277
17.6.3 PHP3	278
17.7 小结	278
第十八章 综合实例：图书销售系统的实现	279
18.1 应用程序涉及的知识点	279
18.1.1 数据库结构	279
18.1.2 DBI 接口	280
18.1.3 DBM 数据库	281
18.1.4 书库管理程序	281
18.1.5 图书订购程序	281
18.2 系统的主页	281
18.2.1 公共例程	282
18.2.2 SSI(Server-side include)的应用	283
18.2.3 密码保护	285
18.3 书库管理	287
18.3.1 书库管理的主页面	288
18.3.2 对书库的各表进行维护	289
18.4 图书订购	298
18.4.1 基本思路	298
18.4.2 客户 ID 的生成	299
18.4.3 客户 ID 的传递	300
18.4.4 给客户发送 HTML 页面	301
18.4.5 图书查询	302
18.4.6 把图书放入购物车	308
18.4.7 订单的管理	312

18.4.8 获取客户信息	319
18.4.9 结帐离开	325
18.5 小结	329

第一章 Linux 的基础知识

目前世界上有十多种计算机操作系统,但是,这其中,有两大类是最流行的,一个是 UNIX 类,另一个是 Windows 类。UNIX 系统占据的主要领域是大型机和工作站等高性能计算机,而 Windows 的主要领域是 PC 机。但是,由于 Windows 软件比较昂贵,对于一些计算机爱好者来讲,他们急需一种廉价的,而且性能稳定的操作系统。Linux 的设计者 Linus Torvalds 就是本着这种思想开发了这一套系统。Linux 继承了 UNIX 系统的各种优点,包括安全性、稳定性、可扩展性及强大的网络功能等,又对 UNIX 系统的某些内容进行了改进,比如增加了更加丰富的工具,如 GNU 的编译程序、文本处理工具等。同时, Linux 更容易与硬件融合,各种类型的 PC 机都能够很顺利地安装和使用 Linux。

在本章中,首先介绍 Linux 的基础知识。本章主要目的是让读者了解 Linux 的根源和发展过程,了解在 Linux 下进行软件开发的基本特点。

1.1 Linux 简介

1.1.1 Linux 的历史

1991 年 8 月,芬兰的一个学生在 comp.os.minix 新闻组贴上了以下这段话:

“你好,所有使用 minix 的人 — 我正在为 386(486)AT 做一个免费的操作系统(只是为了爱好,不会像 GNU 那样很大很专业)”。这名学生就是 Linus Torvalds,而他所说的“免费操作系统”就变成我们今天知道的 Linux。

作为一个具备所有特性的类似 POSIX 的操作系统, Linux 并非仅由 Linus 一人开发,而是由全世界几百个程序员共同开发的。其发展过程吸引了全球的玩家以及部分商业组织的参与,这样的软件发展模式也只有透过 Internet 才有可能。有趣的是这个巨大的全世界的开发活动不是很协调。当然, Linus 为内核定了调子,但 Linux 并不只是内核。当时 Linus 的目标是改进原有系统的一些缺点,加入一部分有活力的新东西。

由于 Linux 属于 GNU 软件,你可不花钱就得到 Linux。你只要有时间并愿意下载,你就可以免费得到大多数软件。任何人都可以拷贝、传播、发展这个系统。但是,使用者只可以拷贝该产品,不可以牵涉到商业行为。同时,发布软件的人也有必要将程序的源代码一起交给别人。

1.1.2 Linux 系统

Linux 是一套操作系统,就其 Linux 字眼而言指的是 Linux 的 kernel (系统核心服务程序)。对一般 Linux 使用者或管理者来说,我们所谈的 Linux,并不是这一小区块,而是指一个可以具体提供某些解决方案的操作环境,用计算机术语来说,就是包含有 Linux kernel、

utilities (系统工具) 以及应用程序的一个整体。

1.1.3 Linux 的主要特点

1. 自由软件

Linux 的最大特点就是它的 Free 性。这种思想源于 GNU。所谓的自由软件就是发放的程序包含程序的所有源代码。这种思想是非常进步的, 在那些非 Free 的软件中, 是不会包含程序的源代码的, 这是由于商业公司的利益导致的。而 Linux 彻底摒弃了这种做法, 它保留版权, 但是允许任何用户获得程序的源代码, 而且这些程序可以在 Internet 上完全免费获取。即使是通过发放的 CDROM 获取, 花费的价格也只是复制 CDROM 的成本钱, 远远低于其他操作系统的价格。正是这种 Free 又免费的特点, 使 Linux 系统迅速成长起来, 逐渐成为越来越受欢迎的操作系统。

2. 多任务系统

Linux 可以同时执行多个程序、多个进程, 而且可以用相应的指令安排各种进程间的优先级, 达到合理的分配效果。同时, Linux 下的多任务分配相当不错, 各程序可以在键入命令以后及时地响应。

3. 多用户系统

采用 Linux 操作系统的同一台机器可以供多人同时使用。它不但可以在 Telnet 下面支持多用户, 也可以同时响应许多用户的不同请求。

4. 必须在 386 protected mode (386 保护模式) 下运行

要求系统硬件至少是 Intel 系列 80386 以上级别, 如果需要在 80286 以下的微机上使用, 就必须使用专用的 UNIX 系统, 比如 Minix。

5. 采用保护内存的方式执行程序

Linux 系统采用保护内存方式执行程序, 所以在个别的程序失控的时候, 不会引起整个系统的崩溃而造成死机, 或者是造成系统对外无响应能力。同样, 还可以使用远程登录的方式进入系统, 清除掉已经失掉的进程。

6. 使用动态链接库

使用动态链接的思想, 在磁盘上是读取程序中实际用到的部分。这样可以大大地提高系统的响应速度, 增强系统实际性能。

7. 共享内存

多个程序可以同时使用同一个内存块。最初, 由几个程序共用一段内存块, 但当某个程序尝试写入这个内存块时, 该 page 就被拷贝到新的地方, 以后该程序的 page 就指向新的 page。这样就可以提高速度并减少内存的使用。所以, 共享内存提高了内存的使用效率, 用同样的内存做了更多的工作。

8. 使用虚拟内存

可以使用虚拟内存, 但需在硬盘上划出一块区域作为交换用的交换分区 (Swap partition)。

9. 符合 POSIX 标准

Linux 符合 POSIX (Portable Operating System Interface), 即标准操作系统界面定义标准。源代码与 UNIX SYSTEM V、BSD 和 SVR4 完全兼容, 可以大大地减少软件移植的工作量。这也是其迅速流行的一个主要原因。