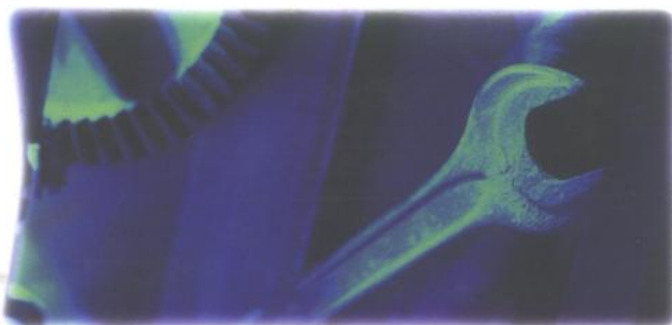


Linux

Web编程

Open Source Linux
Web Programming



本书附带的
CD-ROM上含有
开放源代码工具
和所有示例的源
代码!

利用开放源代码工具创建先进的
Linux Web应用程序

利用面向对象的编程增强可扩展性

利用XML建立开放、可扩展的数据结构

[美] Christopher A. Jones, Drew Batchelor 著

张立新 张 华 高寿福 等译

魏永明 审校



电子工业出版社

Publishing House Of Electronics Industry

URL: <http://www.phei.com.cn>

Linux Web 编程

Open Source Linux Web Programming

[美] Christopher A. Jones 著
Drew Batchelor

张立新 张 华 高寿福 等译
魏永明 审校

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是围绕 Web 编程展开的,分三个部分,共 12 章。内容包括:介绍一种广泛应用于 Internet 应用程序和分布式系统的 Web 体系结构,读者对这种体系结构的理解是开发强壮的、可升级的 Internet 应用程序的关键;介绍在 Linux 中如何使用 Apache 应用服务器和 Perl 语言来完成 Web 编程工作;介绍内容对象的概念,以便读者能够很容易地将显示逻辑和显示内容区分开来;介绍了如何布置 Internet 应用程序,从而实现应用程序的伸缩性。本书最精彩的部分在于对使用新兴技术来完成 Web 开发工作的讨论,其中包括如何使用 XML,如何使用 Java 程序,以及如何使用 GPL 工具。本书在讨论上述内容时,向读者提供了大量使用 Perl 语言、Java 语言、HTML 和 XML 开发的实例,并由浅入深地讲解了每一个概念和实例。因此,本书所面向的读者范围十分广阔,无论是刚刚了解 Linux 的爱好者还是网站的资深管理员,在阅读完本书之后,都会有一种获益匪浅的感觉。



Open Source Linux Web Programming by Christopher A. Jones, Drew Batchelor

Copyright ©2000 by Publishing House of Electronics Industry. Original English language edition copyright

©1999 by IDG Books Worldwide, Inc. All rights reserved including the right of reproduction in whole or in part in any form. This edition published by arrangement with the original publisher, IDG Books Worldwide, Inc., Foster City, California, USA.

本书中文简体专有翻译出版权由美国 IDG Books Worldwide, Inc. 公司授予电子工业出版社及其所属今日电子杂志社。未经许可,不得以任何手段和形式复制或抄袭本书内容。该专有出版权受法律保护,侵权必究。

图书在版编目(CIP)数据

Linux Web 编程/(美)克里斯托弗(Christopher)等著;张立新等译.-北京:电子工业出版社,2000.4

ISBN 7-5053-5918-5

I.L… II.①克…②张… 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(1999) 第 20503 号

书 名:Linux Web 编程

著 者:[美]Christopher A. Jones, Drew Batchelor

译 者:张立新 张 华 高寿福 等

审 校 者:魏永明

责任编辑:嘉 益

特约编辑:赤 耳

印 刷 者:北京天竺颖华印刷厂

出版发行:电子工业出版社 URL:<http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销:各地新华书店经销

开 本:787×1092 1/16 印张: 30.75 字数: 738 千字

版 次:2000 年 4 月第 1 版 2000 年 4 月第 1 次印刷

书 号:ISBN 7-5053-5918-5 著作权合同登记号: 图字:01-1999-3325
TP·3085

定 价:59.00 元(含光盘一张)

凡购买电子工业出版社的图书,如有缺页、倒页、脱页者,本社发行部负责调换。联系电话:68214070

译者序

Internet 已经成为当今 IT 行业最引人瞩目的焦点。Internet 发展到今天,已不仅仅只是一种工具,它已经轰轰烈烈地走入了我们的生活,并开始影响我们每一个人的生活和工作方式。

Linux 操作系统作为自由软件行业的一朵奇葩,令从事软件行业的人士感到异常振奋。从 1992 年开始,Internet 上无数的程序员为 Linux 的发展注入了自己的心血和智慧。如今, Linux 已经成为 32 位的跨平台、多任务、高效而稳定的操作系统。

在 Linux 上的 Internet 开发工作,结合了软件行业中 Internet 和 Linux 这两大焦点的内容,越来越得到软件行业人士的重视,并取得了巨大的发展。尽管有许多操作系统可以用来完成 Internet 开发工作,但 Linux 由于其在网络应用上的卓越表现以及低成本的特点,在 Internet 开发方面有非常优越的竞争力。早在 1996 年中期的一次调查中,统计数字就表明采用 Linux 的网络主机数量比采用 Windows NT 的两倍还多,在 UNIX 系统中名列第二。现在的情况要比 1996 的情况更加乐观,由于 Linux 操作系统的飞速发展,由于大型公司(如 Red Hat)对 Linux 市场的商业化运作,采用 Linux 作为网络主机操作系统的数量要比上面的数字还要多得多。

面对这种软件行业的发展趋势,本书论述了如何在 Linux 平台上进行 Internet 开发工作。其中包括:在交互式 Internet 应用开发中使用到的分布式 Web 体系结构、Apache 软件服务器和 Perl 语言在 Linux 中的使用、内容对象的概念、软件的伸缩性和可供用户选择的实现方法等问题。并且本书将重点放在向读者提供解决实际问题的方法,即提供给读者使用 Linux 开放系统或其他开放资源的工具来解决 Internet 开发问题的方法。本书在讨论上述问题的同时,向读者提供了大量使用 Perl 语言、Java 语言、HTML 和 XML 开发的实例,并由浅入深地讲解了每一个概念和实例,做到了通俗易懂且实用性强。相信在本书的指导下,读者的 Web 编程能力和经验将会得到巨大的提高。

本书由张立新、张华、高寿福、游华云、赵广涛等组织翻译,全书由魏永明审校。参与本书翻译工作的还有包桂秋、王邦江、李铁民、王忠华、孙昆鹏、孙登峰、唐晓强、胡波、白杰文、李峰厚、朱玉、李健峰、陈涛生、谢盈、崔铁军等。由于译者水平有限,错误之处在所难免,还望读者不吝指正。

译者
2000 年 1 月

关于作者

Christopher A. Jones 是西雅图地区的一名 Internet 开发顾问,是《UNIX Shell Objects》(IDG Books Worldwide 出版)一书的作者。他的专长在于接口开发,其客户包括 Boeing、AT&T、Microsoft 以及 Planet 7 Technologies 等。休假期间,Chris 喜欢与他的妻子 Barb 及儿子 Miles 在太平洋西北享受欣赏奇异的动植物世界。

Drew Batchelor 是 Onstage Media Inc. 的股东,该公司是座落于宾夕法尼亚州西切斯特的一家 Internet 咨询公司。他在公司担任商业 Web 站点的设计及开发人员。在空闲时间,他喜欢写作及录制音乐,当然,更喜欢的还是玩他的计算机。

目 录

前言	(1)
第一部分 当今的 Internet 应用程序	(5)
第一章 分布式系统和网络互联技术	(7)
多层体系结构	(8)
什么是层	(8)
客户机-服务器模式的僵硬性	(9)
大规模的分布式系统	(9)
Internet 应用程序	(11)
了解 Web 用户的能力	(12)
真正理解 HTML	(12)
用户脚本	(12)
Java 小应用程序	(15)
ActiveX 控件	(15)
XML	(16)
建立定制 Web 客户	(16)
使用应用程序服务器	(17)
理解通用网关接口	(17)
在应用程序服务器中运行软件	(18)
使用应用程序服务器的资源	(19)
了解商务对象层	(20)
定义商务规则	(20)
Business 对象的优点	(20)
Business 对象在分布式系统中的地位	(21)
访问数据层和后端服务	(23)
对数据访问的理解	(23)
事务处理	(23)

服务器数据的作用	(25)
了解 XML 的功能	(25)
在软件之间调节数据	(26)
小结	(26)
第二章 介绍 Apache HTTP 服务器	(27)
编译和安装 Apache	(28)
为 CGI 配置 Apache	(33)
编辑 httpd.conf	(33)
设置 Document Root	(34)
配置目录选项	(35)
允许服务器端包含(Server-Side Include)	(36)
在 Document Root 中使用 CGI	(37)
将 CGI 作为文件类型来使用	(38)
使用服务器端包含	(39)
第一次运行 Apache	(40)
在系统启动时运行 Apache	(42)
对 Apache 的剖析	(43)
理解 HTTP 在 Web 开发中的作用	(44)
请求一个 HTML 页	(44)
接收 HTML 页	(46)
向服务器发送 Web 表单	(46)
向服务器发送一个 URL 请求	(48)
发送和接收 cookie	(49)
处理 CGI 脚本	(52)
创建一个规范的 CGI 脚本	(53)
理解 CGI 环境	(53)
创建脚本	(54)
对核心 CGI 变量的详细介绍	(56)
小结	(58)
第三章 在 Linux 中用 Perl 实现 CGI	(59)
为什么要使用 Perl	(59)
Perl 的基本知识	(60)
运行 Perl 脚本	(60)

变量	(61)
一个例子:日志文件.....	(75)
模式匹配和正则表达式	(81)
在日志分析程序中使用模式匹配	(84)
引用	(91)
在日志分析程序中使用引用	(94)
CGI 的另一面:用户交互	(101)
包、对象和模块	(104)
下载并使用 CGI 模块	(107)
完成日志分析程序.....	(112)
小结.....	(120)
第四章 将内容、逻辑与对象分开	(121)
将 HTML 和 CGI 分离开	(121)
改变文本文件需要一个程序设计员	(122)
重新设计需要重新编码	(122)
克服常见的 Web 站点问题	(123)
使用服务器端包含	(123)
将错综复杂的站点结构变得更加简单	(124)
基于上下文内容动态更新内容	(124)
内容创建对象.....	(124)
自确定的定位(导航)控件	(125)
自动化的页眉和页脚	(125)
连接内容	(125)
处理文件系统结构.....	(126)
编写对象.....	(127)
创建 HeaderFooter Perl 对象	(127)
理解 HeaderFooter 方法	(128)
为 HeaderFooter.pm 编写代码	(130)
创建 Navigator Perl 对象	(138)
理解 Navigator 方法	(138)
编写 Navigator 代码	(139)
创建一个 Navigator/HeaderFooter 实例	(147)
设置 Document Root	(147)
编写代码	(149)

配置 Display 对象	(150)
写入到 HTTP 流	(152)
评价脚本	(155)
创建 Content Perl 对象	(156)
理解 Content 方法	(156)
编写 Content 对象代码	(157)
创建一个 Template Perl 对象	(162)
创建更多的灵活性	(163)
揭示 Template 方法	(163)
编写 Template 代码	(164)
创建 Producer Perl 对象	(172)
理解和扩展 Producer 方法	(172)
编写 Producer 对象代码	(173)
创建一个 Producer/Template 实例	(175)
编写代码	(175)
创建一个在线编目系统	(181)
向用户致意	(181)
介绍产品分目	(182)
显示一个单独的产品	(183)
查明附带的 CD-ROM 中的文件	(183)
创建一个分目引擎	(183)
找出解决方案	(184)
开发代码	(184)
扩展这个系统	(198)
小结	(198)
 第五章 用 XML 创建结构优良的对象	(199)
表述信息	(199)
XML 的鼻祖	(200)
从 GML 到 HTML	(200)
HTML 所存在的问题	(201)
HTML 和样式表	(201)
从 HTML 转向 XML	(204)
序言	(205)
文档类型定义	(207)

一般的实体	(207)
记号	(209)
元素	(210)
属性	(214)
参数型实体	(218)
INCLUDE 和 IGNORE	(219)
CDATA 标记段	(220)
XSL 和 Xlink:新兴的标准	(220)
小结.....	(222)

第六章 用 CGI 应用程序服务器处理 XML	(223)
安装解析器模块.....	(223)
关于解析器.....	(224)
对解析器进行初始化.....	(225)
把标记翻译成实际的动作.....	(227)
本章的项目:设计一个把 XML 转换成 HTML 的翻译器	(230)
使用处理程序.....	(230)
处理程序与 Subs 的对比	(234)
XML 文件	(234)
Memo.xml	(235)
Style.xml	(237)
开始编写转换脚本.....	(239)
initStyles 方法	(240)
style_start_handler 方法	(241)
style_end_handler 方法	(242)
style_char_handler 方法	(243)
处理源文件.....	(244)
source_start_handler 方法	(245)
source_end_handler 和 source_char 处理程序	(246)
创建一个脚本.....	(247)
处理属性.....	(248)
修改样式处理程序.....	(251)
修改源处理程序.....	(253)
访问属性值.....	(255)
proc_handler 和 sub_att_value 方法	(257)

连接 IDREF 和 ID 属性	(259)
Attlist 和 Element 处理程序	(261)
实现文本置换	(252)
实现 Final 处理程序	(264)
小结	(273)
第七章 创建基于客户的 Web 内容管理器	(275)
基于 XML 的留言板	(275)
MESSAGE_1.xml	(276)
style_MESSAGE.xml	(278)
应用程序用到的其他文件	(282)
把 XML 转换成 HTML 表单	(283)
枚举输入	(286)
添加选择框	(289)
修改 XMLStyle 模块	(290)
from_directory 方法	(292)
from_file 方法	(294)
from_att_type 方法	(296)
必需元素内容	(297)
传送文档的结构	(299)
edit.cgi 脚本	(305)
form2xml.cgi 脚本	(307)
对参数进行排序	(308)
删除不必要的元素	(310)
生成 XML	(312)
Admin 程序	(316)
admin.cgi 脚本	(316)
查看器程序	(320)
小结	(324)
第二部分 超越今天的 Internet 模型	(325)
第八章 用 Linux 提供 XML 服务	(327)
查看 Internet 应用程序	(327)
超越客户/服务器时代	(327)
引入 Internet	(328)

通向电子商务的条条道路	(328)
需求分析	(329)
获得一种解决方法	(329)
完全使用另外一种方法	(331)
查看目前企业版 XML 模型	(332)
XSL 转换	(333)
跨应用程序通讯	(333)
了解无状态问题	(333)
单向与固定状态	(334)
HTTP/CGI 模型的缺点	(334)
通向真正交互性的道路	(334)
XML 应用程序服务器	(335)
序列化和共享 XML	(335)
理解进程	(337)
较大的图片	(338)
安装 XML 应用程序服务器	(338)
安装软件	(339)
测试 XAS 设置	(340)
小结	(341)
 第九章 在 Linux 上开发 Java 小应用程序	(343)
Java 是怎样工作的	(343)
Java 开发软件包	(344)
安装 Blackdown JDK	(345)
检查 JDK	(346)
进行安装测试	(346)
HelloEarl 探秘	(347)
从应用程序到小应用程序	(349)
接口	(352)
把所学到的全部用起来:Life 游戏	(356)
关于这个游戏	(356)
Life 小应用程序概述	(356)
Life.java	(357)
init 方法:获取 HTML 参数	(358)
异常处理	(359)

initImage 方法	(362)
initAWTComponents 方法	(363)
实现 Life 逻辑的方法	(365)
Java 线程	(368)
myCanvas 类	(371)
小结	(374)
 第十章 创建 Java/XML 网络应用程序	(377)
应用程序概述	(377)
使用 XML 的应用程序服务器	(378)
创建 XAS 客户程序	(378)
XASReceiver 接口	(380)
Aelfred XML 解析器	(381)
下载和安装 Aelfred	(381)
XMLParser 类	(382)
XmlHandler 接口和 HandlerBase 类	(383)
开始小应用程序的编写	(383)
NetApp.java	(383)
NetApp 类的 init 方法	(386)
initAWTComponet 方法	(388)
XASReciever 方法	(391)
Hub 接口方法	(392)
myHandler 内部类	(393)
创建组件对象	(397)
Chat 组件	(398)
Chat 构造函数	(400)
initAWTComponets	(400)
actionPerformed 方法	(402)
refreshChatterList XASReceiver 接口方法	(402)
XML Handler 方法	(404)
Whiteboard 组件	(404)
Whiteboard 的构造函数	(406)
generateGraphics 和 doClear 方法	(406)
MouseEvent 处理过程	(408)
sendPointVector 方法	(410)

XASReveiver 方法	(411)
XML 处理程序方法	(411)
drawRemote 方法	(412)
paint 方法	(413)
启动网络应用程序	(415)
小结	(417)
第十一章 采用错误处理保护网站	(419)
处理应用程序的错误	(419)
CGI::Carp API 简介	(420)
用 Apache 处理错误	(420)
理解 ErrorDocument 的指令	(421)
实现自定义错误页	(421)
错误处理体系结构	(422)
分析 Judy Jumpsuit 事例	(422)
环境分析	(423)
设计一个错误处理方案	(424)
从消息中提取错误代码	(425)
增加错误情况	(425)
编写错误代码	(425)
一个简单的 Perl 模块	(426)
使模块与信息分离	(427)
实现错误消息	(427)
编写双重错误处理程序	(428)
使用方法	(429)
编写对象代码	(430)
写一个 HTML 错误脚本	(438)
扩展商品目录网站	(445)
创建顾客表单	(446)
获取顾客数据	(448)
伸出你的网络触角	(453)
为什么捕获特殊字符	(453)
小结	(454)

第十二章 部署 Linux Web 应用程序	(455)
这不是你家的 Web 站点	(455)
采取正确的第一步	(456)
采取正确的第二步	(457)
测试和调试	(457)
检查 HTTP 报头	(457)
理解 cookie 的使用	(460)
性能监控	(463)
可用性和浏览器的兼容性	(464)
规划资源消耗	(465)
规划 Intranet 应用程序	(465)
规划商务应用程序	(466)
选定服务器的配置	(469)
采用多个服务器分配负载	(469)
负载均衡工具软件	(469)
负载均衡的优点	(471)
对负载均衡方案缺陷的理解	(471)
了解伸缩性问题	(473)
小结	(473)
 附录 CD-ROM 的内容	(475)
每一章的示例程序	(475)
XML 应用程序服务器	(475)
Microstar AElfired 解析器	(476)
Perl 模块	(476)
开放源代码的 Kaffe Java 虚拟机	(476)
PostgreSQL 6.5	(477)

前言

Internet 已经逐渐成为当今分布式系统的主要表现形式。从提供超文本开始,Internet 业已成为信息时代最充满活力的主要方面。

在本书中,我们提出了当今先进的 Internet 体系结构,并规划了未来的应用软件以及在解决新问题时所采取的技术路线。本书可以说是我们工作成果的总结。

在本书的结尾,我们努力超越读者在其他 Web 编程书籍中可以找到的内容,并努力使读者接触到真正的信息,这种信息是 Internet 开发的基础。本书详细讲述了在交互式应用开发中用到的分布式结构、软件的伸缩性和可供用户选择的实现方法,并且本书的重点在于提供解决问题的方法,提供给读者使用 Linux 开放系统或其他开放资源的工具来解决开发问题的方法。

本书所面向的读者

如果读者对在 Linux 平台上使用开放源代码的工具进行 Internet 开发这种前沿技术感兴趣,那么阅读本书就会受益。虽然本书重点讨论编程,但本书包括了从伸缩性到配置的所有内容,这些内容对于专业 IT 人员和中高级程序员都是有吸引力的。

本书的内容

本书首先介绍了一种广泛应用于 Internet 应用软件和分布式系统的复杂的 Web 体系结构。对这种体系结构的理解是开发强壮的、可升级的应用软件的关键。当今的 Internet 系统已经从获得一个主页或执行一个 CGI 脚本这种简单的应用水平得到了长足的发展,它一般包括多重的计算机和逻辑,从而使进程和计算机的界线交织在一起。本书就是以这种思想为指导来写作的。

本书介绍了在 Linux 中如何使用 Apache,并对 Perl 语言的特点进行了简单的回顾,以使 Web 开发者能够从一个开发语言转入到 Perl 语言中。

在第4章,将建立内容对象的概念,以便于阐述显示逻辑和显示内容的区别。假设我们建立了一个 Web 站点,而在这个站点中,我们想要显示的内容和显示逻辑是密切相关的,这对于我们来说简直就是一个恶梦。本书对于内容对象的讨论,可以使读者看到如何将程序员的逻辑从 HTML 内容中清晰地分离出来。

本书更精彩的部分在于讨论如何使用新兴技术,例如使用 XML 将读者的 Internet 应用软件的功能推广到新的应用领域。通过本书对 XML 的简要介绍,读者对可扩展标记语言 XML(eXtensible Markup Language)的重要特点有所了解,这种可扩展标记语言可能和读者的 Web 开发工作有关。本书通过剖析 XML::Parser Perl 模块的使用,向读者阐述了如何将动态结构和 XML 结构文件连接起来。如果将上述思想扩展到一个自支持的管理系统中,用户就可以通过简单的浏览器方式对 XML 文件进行以 Web 为基础的编辑。

众所周知,如果正确使用 Java 小应用程序,这些程序可以为 Web 主页提供优异的功能。因此,我们对在 Linux 上开发 Java 小应用程序进行了全面的分析。我们创建了一个动态的 Life 游戏,来阐明 Java 的特点,其中包括 Java 的多线程和事件模型。

本书通过阐述一种全功能的、以 Java 为基础的聊天室和留言板(whiteboard)实时通讯应用软件,扩展了 Internet 应用程序的范围。这个实例是用 XML 应用服务器开发的,因此在本书中也介绍了一种 GPL 工具。通过使用 Java、XML Parsers 和实时网络连接来生成用户应用软件,可以使读者作到以最少的工作量得到高品质的软件。书中所举的聊天室和留言板应用软件的例子阐明了一种优异的、以顾客为基础的一类,这种类通过使用 XML 应用服务器所提供的服务来创建应用程序。

在本书的结尾,对如何配置 Linux 应用程序进行了切实的分析,其中包括对负载均衡的硬件和可伸缩关系的讨论。用户在编写代码时,需要不断地访问这些关系,因此这些关系的存在避免了用户在开发程序阶段为阐明这些关系而反复编码。

本书约定

本书假定读者对 Linux 系统和程序编制具有中等熟练的水平,因此本书没有将 Linux 系统和编程的基本知识包括在内,而是将重点放在应用程序和普通 Web 编程技巧的扩展上,即利用开放源代码的 Linux 资源和本书所提供的增强的体系结构来编制应用程序和提高普通 Web 编程的技巧。

纵观本书,有很多图标贯穿于其中。这些图标表示的含意为:对读者的警告、需要读者注意等。书中所提到的软件都存放于随书附带的 CD-ROM 中。