

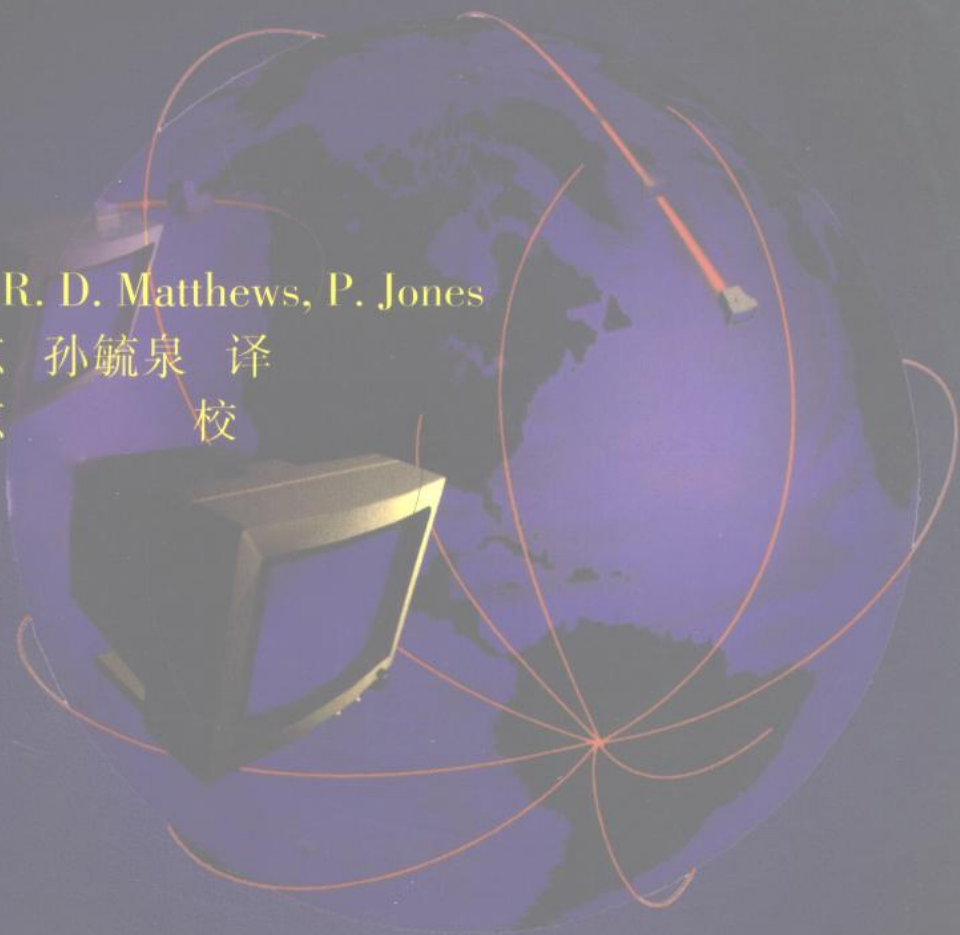
VENTANA

北京科海培训中心

【美】J. Magid, R. D. Matthews, P. Jones

李谦 廖卫东 孙毓泉 译

廖卫东 校



Web

服务器技术指南

3-62

TD/1

机械工业出版社

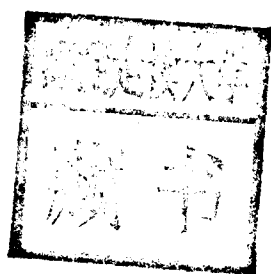
VENTANA

北京科海培训中心

Web 服务器技术指南

丁马结集
[美] J. Magid R. D. Matthews P. Jones 著

李 谦 廖卫东 孙毓泉 译
廖卫东 审校



机械工业出版社

039798

著作权合同登记号:01-96-0533

内 容 提 要

这是一本面向 WWW 服务器的设计和维护人员的技术手册。全书以通俗的语言介绍如何构建一个具有专业性的 WWW 服务。内容包括:WWW 服务器运行机制,HTTP 协议,HTML 语言,WWW 服务器软件安装与配置,WWW 页面内容的设计、装入、布局及润色,如何在 WWW 页面加入图象及声音,使之成为声音、图文并茂的画面,如何创建高级表格并编写 CGI 描述文件,以及 Netscape 与 Hot Java 等 WWW 浏览器的使用。本书附录还提供了 Linux 及 X-Windows 的安装与配置。

本书面向网络服务器的编程与维护人员。

图书在版编目(CIP)数据

Web 服务器技术指南/(美) 马吉德(Magid, J.)等著;李谦,廖卫东,译. —北京:机械工业出版社,1996.9

ISBN 7-111-05224-2

I. W... II. ①马... ②李... ③... III. 计算机网络-控制器, Web-技术手册 IV. TP332.3-62

中国版本图书馆 CIP 数据核字(96)第 16675 号

出版人:马九荣 (北京市百万庄南街 1 号 邮政编码 100037)

责任编辑:科培 责任校对:成昊

门头沟胶印厂印刷·新华书店北京发行所发行

1996 年 9 月第 1 版·1996 年 9 月第 1 次印刷

787mm×1092mm 1/16·23.875 印张·580 千字

0 001-5 000 册

定价:38.00 元

作者简介

Jonathan Magid: 自从位于 Chapel Hill 的北卡罗来纳大学于 1992 年启动其 SunSITE 项目以来, Jonathan Magid 一直是该项目的一名系统管理员和编程人员。他目前正在攻读美国历史学士学位, 同时致力于诗歌方面的一个研究项目。Jonathan 曾经撰写了数篇 Internet 和 UNIX 方面的文章, 并开发了多个用于信息管理方面的实用程序。

R. Douglas Matthews: 他目前是北卡罗来纳大学的一名本科学生, 主修人类学与宗教。Matthews 自从 1994 年 1 月参与 SunSITE 项目以来, 一直致力于 WAIS、Gopher 和 WWW 方面的开发工作。他参加了一家 Web 信息服务公司, 即 Netit.com 公司的设计工作, 同时还负责过 HTML 写作和设计方面的培训工作。

Paul Jones: 他是北卡罗来纳大学信息与图书管理学院和新闻与大众交流学院的讲师。自从北卡罗来纳大学启动其 SunSITE 项目以来, Jones 一直就是此项目的一名管理人员。

Jones: 他获得北卡罗来纳大学计算机科学学士学位, 并已获得 Warren-Wilson 大学诗歌方面的 MFA (类似于硕士学位)。他的文章和诗歌曾经发表在 50 多种文学期刊中, 包括《Southern Review》,《Poetry》,《Southern Humanities Review》和《Ironwood》。Jones 的获奖丛书《What the Welsh and Chinese Have In Common》在首次发行之后又有了 HTML 版本, 该书可以在 <http://sunsite.unc.edu/pjones/poetry/> 找到。

致 谢

首先,我们全体作者都必须向 David McConville 先生致以最诚挚的感谢。David 撰写了第 8 章“真正的多媒体”和第 13 章“未来发展方向”关于虚拟现实模型语言与远程会议部分。事实上,他应该算是本书的一名合作者,同时也是我们大家的亲密朋友。此外,我们还要感谢 Chris Colomb 先生为本书第 3 章“建立服务器”所作的重要贡献。Chris 的知识和经验对充实该章内容起了非常大的作用。

这里我们还要感谢一位 SunSITE 项目的元老 Paul Jones 先生,尽管他也是本书的作者之一。Jones 和他的 SunSITE 项目给我们提供了增长经验和学识的机会。甚至可以说,没有他的努力,本书根本就不可能问世。

这里我想感谢 Walt Bruce 和 Dykki Settle 两位,因为正是他们在得知我们有一些 Web 服务器方面的经验之后,建议我们编写本书的。同时还必须感谢编辑 Jessica Ryan 先生,他的严谨而又细致的工作是本书能够编写到如此完整的主要因素。

我还要感谢这些年来在 SunSite 项目工作过的所有各位同仁,事实上这应该算我们大家的书。尤其要感谢 Simon Spero 先生为本书提出的建议和所提供的源程序。此外,我还非常感谢那些免费软件的编写人员,他们是 Internet 得以蓬勃发展的主要动力之一。感谢 Matt Welsh 为本书附录提供的 Linux 指南。虽然 Slackware 是整个 Linux 和 GNU 界共同努力的结果,但我们还是要特别感谢 Patrick Volkerding 所做的汇总工作。

最后,非常感谢无论我花多少时间忙于工作,都一直给予我支持的我的家人和朋友。当然要特别感谢我心爱的女儿 Merry,她是我的天使!

Jonathan Magid

感谢我的父亲!我觉得无论我现在和将来如何努力都不足以报答他。他一直毫无保留地给我关心、教导和帮助。另外,我还要感谢那些一直给我支持、鼓励、鞭策和帮助,更重要的是能够容忍我的最亲密的朋友 Jason, Mark, Radha, 尤其是 Mary Beth。

R. Douglas Matthews

我能够对本书有所贡献,无疑得归功于自项目启动以来一直或曾经为 UNC SunSITE 工作过的各位同仁,包括 Tom Snee, Darlene Fladager, Dykki Settle, Eric (Max) Leach, Terry Mancour, Beth (Icky) Lyons, Robert (Pyro) Ingram, Kevin Gamiel, Gwenn Conner, Nassib Nassir, Doug Norton, Luke Duncan, Nash Foster, Marcus Cox, Cheryl Friedman, Qin Fang, Glenn Lewis, Ahmet Taylor, Kelly Boley, Jim Fullton, 尤其要归功于 Jim Fullton 和本书的其他几位作者和合作人员,即 Jonathan Magid, Doug Matthews, David McConville 和 Chris Colomb 等。另外,这里还要特别感谢 Sally Greene 和 Tucker Jones 对我生活和情感上的帮助与支持!

Paul Jones

序 言

世界信息网 (World Wide Web) 已经成为最热门的新兴信息技术. 这个由超链接化的多媒体资源组成的巨大网络已遍及全球, 它将信息和娱乐带给越来越多的用户. 拥有这本书, 你能成为这个用户集体中的一员.

现在, 人们已把 Internet 当作自己的印字机一样使用的日常工具. 原版书中配套的光盘 (CD-ROM) 包括了在 UNIX 系统上建立一个 Web 服务器 (Web Server) 所需的全部软件. 此套软件同时也可适用于 Linux 系统, Linux 系统是一个多任务的与 UNIX 兼容的操作系统.

这本关于 Web (世界信息网的简称) 服务器的书将教会你如何使用这些工具, 以便在 Web 上建立一个安全而专业化的形象, 所涉及的内容包括:

- Web 服务器是如何工作的
- 如何将清晰的图形直接加入你的 Web 页面 (Web pages)
- 如何创建可单击的图象
- 如何创建交互表格, 编写 CGI 脚本 (script) 程序
- 如何在你的信息页面嵌入声音和动画

这不是一本普通的 Internet 书籍

两年以前, Internet 并不像当时流行的多媒体一样受人关注. 既没有书或者文章去揭示其神秘细节, 也没有精心制作的所谓“网络王国”之类的电视节目诱惑我们. 只有从事研究的大学和高技术公司才有特权使用 Internet. 对于个人或者小的团体来说, 联网进入 Internet 几乎是不可能的.

今天, 书店里摆满了各种关于 Internet 的书籍. 仅仅美国就有 250 多个 Internet 服务提供者 (ISP), 他们提供大量关于 Internet 的材料, 其质量各异, 从精心编辑的使用手册到单纯的网络 FAQ 文档和文本文件.

然而, 这些资源有一个共同之处, 那就是想把大量的信息提供给你, 你只不过是一个消费者, 而没告诉你如何添加信息. 这些指导手册表明似乎只有网络专家们才能去运行服务器, 让你觉得 Internet 不过是一个让耐心的读者查阅的图书馆.

Internet 远不止如此. 它是一个用于信息通信与分布的平台. 每个人都有一些想说或想呈现给别人的东西. 这些东西或许具有个人特征, 例如某个假期的纪念照片; 或许特别重要, 例如商业活动的目录. 但相互交流的动机是人的本性的一部分. 本书的目的就在于告诉你如何将此动机变成一个 Internet 服务器: 你在世界信息网中的家.

硬件和软件配置需求

本书假定你有一台计算机, 运行 UNIX 系统并且能直接联入 Internet. 当然, 本书提供

的工具和技术,也可用于在企业内联网络或从别处租用的服务器空间上构建 WWW 文档。但如果你需要运行自己的 Internet 服务器,你必须有:

- 一台配有 32MB 内存和 1Gb 硬盘的 UNIX 系统工作站;或者一台 486/33 微机,至少 16MB 内存和 500MB 硬盘。
- 一个 SVGA 显示器(如果你想使用图象的话)。
- 一块声卡(如果需要使用声音的话。请查阅附录 C,验证声卡是否与 Linux 兼容)。
- 一个光盘驱动器(请查阅附录 C,验证是否与 Linux 兼容)。
- 与 Internet 全天候的直接连接。

寻找一个 Internet 服务提供者的过程很复杂,这超出了本书的范围。我们建议你与本地的网络服务站联系。你可以使用你的 WWW 客户程序来查阅提供 Internet 连接的公司清单。地址是 <http://www.yahoo.com/Business/Corporations/Internet.Access.Providers/>。

获得一个 Internet 连接后,需要有一台运行 UNIX 系统的计算机。如果没有一台 UNIX 工作站,也可以使用 486 或 586 微机运行 Linux 系统,它是一个与 UNIX 兼容的操作系统。要知道如何安装和配置 Slackware(Linux 的流行版本),请参阅附录 D“Linux 的安装”。

本书假定你已经了解了 UNIX 系统的基本内容,虽然你不必是 UNIX 专家。如果不熟悉 UNIX,不妨先读一本 UNIX 的入门书,例如由 Charlie Russel 和 Sharon Crawford 所著、并由 Ventana 公司出版的《Voodoo UNIX》,以及由 Matt Welsh 和 Lar Kaufman 所著的《Running Linux》,后者对于 Linux 用户来说是一本很好的入门读物。

本书内容

第 1 章,“什么是 Web”,简要介绍了超文本和它在世界信息网中的使用。本章将解释 Web 服务器、浏览器(browser)和统一资源定位器(URL;Uniform Resource Locator)之间的关系。本章还列出了一些新闻组和地址清单,它们有助于你紧跟 Web 的变化。

第 2 章,“基本组成部分”,简要描述了出入 Web 的各种技术。这章讨论了 Internet 上共同的文件格式,并揭示了信息网如何巧妙地使用“通用 Internet 邮件扩充服务”(MIME: Multipurpose Internet Mail Extensions)类型,最后介绍了超文本标识语言 HTML(Hyper-Text Markup Language)和 tkHTML,一个 WYSIWYG HTML 编辑器。

第 3 章,“建立服务器”,比较了最流行的 Web 服务器软件,并且介绍了安装和设置流行的 NCSA 服务器 1.4 版的过程。对于那些等不及自己建立服务器的用户,本章还简要介绍了如何租用别人的服务器。

第 4 章,“系统安全性”,对安全性这个重要议题进行了阐述。你将学会如何采取一些安全措施,以减少作为电子尖端领域开拓者所带来的风险。

第 5 章,“在 Web 中加入文档”,介绍如何把文档从各种流行的字处理器和桌面出版系统转换到 HTML。

第 6 章,“检查你的工作”,介绍如何使用多种浏览器和 HTML 检验工具,以保证你的文档在每种浏览器中都能很好地显示。

第 7 章,“Web 上的图象”,介绍了如何有效地使用 Web 上的图形。其内容包括常用的

图形格式和如何提高插入图象的质量。并且详尽地解释了不少高级的论题,如透明性、交错(interlacing)和可单击的图形映象。

第 8 章,“真正的多媒体:加入声音和动画”,介绍如何在 WWW 页面加入动画和声音。本章不但描述了各种标准,而且帮助你开始建立自己的多媒体文件以及如何链接它们。

第 9 章“搜索与索引”,告诉你如何使用 freeWAIS 建立 WWW 站点和其他文档集的全文本索引。

第 10 章,“简单表格”揭示了交互式表格的秘密。本章详尽解释了如何用 HTML 编写一个表格,以及如何使用 gform(一个通用表格处理器)来处理表格数据。本章还将讲述如何用表格检索 WAIS 数据库。

第 11 章,“CGI:面向编程人员的高级表格”,介绍 Web 的高级特性。本章还讲述如何使用 CGI 脚本处理表格,在空白页上生成 HTML,以及建立与其他 Internet 服务的连接。

第 12 章,“整合:使你的页面融进 WWW 虚拟区”,介绍怎样才能把用户吸引到你的服务器来并且使他们成为回头客。

第 13 章,“未来发展方向”,解释了那些正在塑造世界信息网的未来的最新技术。其内容包括 HTML 3.0 版的新特性,虚拟现实模型语言 VRML(Virtual Reality Modeling Language),以及实时会议、HTTP-NG、电子商业和其他新兴的 Internet 技术。

附录 A,“原书配套的联机资源库”,介绍了这个 Web 工具和它最有价值的特性——软件库和软件档案。

附录 B,“原书配套的 CD-ROM”,介绍了原版书所附光盘 CD-ROM 的内容:所有为建立用于 Internet 上交互式电子出版平台所需的软件。该光盘不仅包括了所有在书中详细讨论了的与 Web 相关的软件,而且包括 Linux 的 Slackware 版。Linux 是用于 386 或更高档 IBM PC 兼容机的多任务操作系统,它与 UNIX 兼容。

附录 C,“Linux 硬件兼容性指南”列出了 Linux 支持的大多数硬件,该附录有助于你安装任何必要的驱动程序。

附录 D,“Linux 的安装”,介绍如何从 CD-ROM 安装 Slackware Linux 产品。

附录 E,“Linux XFree86 指南”,介绍如何获得安装和配置用于 Linux 系统的 X Windows System(X11R6)的 XFree86 版的 3.11 版。

附录 F,“FTP 和 Gopher”,讲述如何配置 FTP 服务器以允许匿名登录,以及为何我们建议严格限制使用 Gopher。

最后,参考文献中列出了一些联机的和已印刷出版的参考资料。

命令行约定

本书不是一本关于理论方面的书,相反,它将教给你许多使你成为 Web 专家的技巧。因此,本书的许多例子可以在命令行输入。下面便是一个例子:

```
$for i in *.html
>do
>echo $i
>grep -i '<TITLE>' $i
>done
```


本例的第 1 行是以美元符开头。你无需自己输入美元符。这是 Bourne Shell 命令行解释器/bin/sh 的缺省命令行提示符。类似地,你也无需输入>符号。当 shell 程序希望从键盘接收输入时会自动提示。

当一个命令行太长,一行又放不下时,你可插入一个\符来续行:

```
#ls -l *.html|awk '{size+= $5};\n>END {print size}'
```

一旦你写完此行,shell 就会执行该命令。

注意,上例中有一个#符号作为提示符,这表明命令必须以 root 或超级用户身份输入。

本书所有的例子都假定你使用的是 Bourne shell,或者与它向后兼容的 shell,如 GNU 的 bash (Bourne Again Shell)或者 KornShell。本书之所以使用 Bourne shell 作为例子,是因为它有清晰的语法,并且每个 UNIX 机器都支持这种 shell。

如果你使用不兼容的 shell,例如 C shell,那么就得改为 Bourne 或者改变一下例子。本书的每个小例子都不复杂,因此不会太难。例如,第 1 个例子写成 C shell 中的样子是:

```
% foreach i( *.html)\n? echo $i\n? grep -i' <TITLE>' $i\n? end
```

如你所见,它们并不难。其实本书中的例子很少有比这个更难的。

保护与服务

在你开始使用 Internet 发布信息后,也许就不愿再使用传统的媒体。因为好处是明显的:

- 你可以访问到成千上万的人,不管他们是在你家的同一条街上还是在世界的另一端。
- 你可以立即进行更新和更正,可以不断发展你提供的材料,作一些小的更改无需重新印刷的开销。
- 由于大多数人使用计算机制作拷贝,故把这些文档加入 Web 是很简便的。
- 在 Internet 上提供信息是双向的。信息不是像传统媒体那样单向流动,而是在客户机与服务器两者之间相互传递。这种交互性为作者和读者的对话提供了基础。

公司可利用这种交互性实实在在地评估他们广告的效果。一些机构和政府可利用它来参阅由具有类似兴趣和关注对象的人们组成的虚拟团体。Internet 不仅仅是一个通向更多观众的渠道,也还提供了交换意见、寻找新朋友和合作伙伴的机会。

当然,要想认识使用 Web 服务器的益处,你必须采取果断行动并开始提供服务。本书第 1 章介绍超文本的概念以及它在世界信息网中的使用方式。如果你是 Web 老手,那么可跳到第 2 章,该章解释了使 Web 运转的基本技术。

说明:

欲购原版书所附光盘者请与机械工业出版社或北京科海培训中心联系。

联系地址:北京百万庄南街1号机械工业出版社发行部

邮政编码:100037

联系地址:北京海淀路82号北京科海培训中心

邮政编码:100080

目 录

第 I 部分 预备知识

第1章 什么是 Web	1
1.1 客户机与服务器	1
1.1.1 信息服务器	3
1.1.2 其他必须牢记的事情	4
1.2 WWW 的起源	4
1.2.1 超文本	5
1.3 Tim Berners-Lee, CERN 和 WWW	6
1.4 WWW 的发行	7
1.5 NCSA 的 Mosaic	7
1.6 Netscape 和浏览器爆炸	9
1.7 服务器	10
1.8 W3 集团	10
1.9 Internet 工程特遣组 (IETF; Internet Engineering Task Force)	11
1.10 紧跟潮流	11
1.11 小结	14
第2章 基本组成部分	15
2.1 超文本传输协议	15
2.1.1 HTTP 是如何工作的	15
2.2 Web 上的文件	16
2.2.1 文件格式	16
2.2.2 MIME 类型	18
2.2.3 压缩	18
2.3 HTML 的组成	22
2.3.1 文档定位	22
2.3.2 建立文档	23
2.3.3 超链接	27
2.3.4 图象	28
2.3.5 Netscape 的扩充标记	29
2.4 tkHTML: 一个 HTML 编辑器	30
2.5 小结	35
第3章 建立服务器	36
3.1 服务器之间的比较	36
3.1.1 CERN 服务器	37
3.1.2 Plexus	38

3.1.3	WN 服务器	38
3.1.4	发展中的服务器	39
3.1.5	phttpd	40
3.2	NCSA 服务器	40
3.2.1	编译 NCSA 服务器	41
3.2.2	配置服务器	42
3.2.3	服务器配置文件	43
3.2.4	服务器资源映象 (Server Resource Map)	46
3.2.5	服务器访问文件 (Server Access File)	49
3.2.6	建立你的 Web 空间	49
3.2.7	启动服务器	51
3.2.8	访问和身份验证	53
3.2.9	各种服务器的任务	56
3.3	租用空间	58
3.4	小结	59
第4章	系统安全性	60
4.1	普通的 UNIX 安全性	60
4.1.1	帐户 (Accounts)	61
4.1.2	文件许可和拥有权限	65
4.1.3	安全性日志和系统状态	71
4.2	Internet 安全性	75
4.2.1	检查提供的服务	75
4.2.2	NIS 和 NFS	81
4.2.3	高级的 Internet 安全性选择	83
4.3	有人入侵时的对策	84
4.3.1	找到入侵者	84
4.3.2	对策	86
4.3.3	联系人	88
4.4	小结	89
 第 II 部分 增加内容 		
第5章	在 Web 中加入文档	90
5.1	从文本转换到 HTML	90
5.2	从 RTF 转换到 HTML	95
5.2.1	图形	95
5.2.2	命令行选项	96
5.2.3	定义新样式	97
5.3	字处理器转换程序	100
5.3.1	基于 Word for Windows 的 Internet Assistant	100
5.3.2	WordPerfect	101
5.3.3	qt2www.pl	103

5.3.4 其他字处理器	103
5.4 小结	104
第6章 检查你的工作	105
6.1 HTML 样式	105
6.1.1 平台独立性和 HTML	105
6.1.2 根据样式进行设计	106
6.1.3 绝对和相对 URL	109
6.2 有效性验证程序	110
6.2.1 内容和样式检查器	110
6.2.2 用 Anchor 检验程序检验链接	120
6.2.3 用 MOMspider 漫游 WWW 站点	121
6.3 小结	125
第7章 Web 上的图象	127
7.1 链接和内置之间的比较	127
7.2 在 WWW 页面放置内置图象	128
7.2.1 常用 HTML	129
7.2.2 Netscape 扩展	129
7.2.3 GIF	133
7.2.4 JPEG	133
7.2.5 对 GIF 和 JPEG 的进一步思考	134
7.3 从内置图象中得到最佳效果	134
7.3.1 操纵色图	135
7.3.2 隔行 GIF	138
7.3.3 缩放图象	139
7.3.4 透明 GIF	141
7.3.5 综合考虑	143
7.4 常用内置图象	145
7.4.1 漂亮的小点	145
7.4.2 颜色栏	146
7.4.3 导航按钮与图标	147
7.5 可单击的映象图	149
7.6 小结	151
第8章 真正的多媒体:加入声音和动画	152
8.1 Internet 中的数字化声音	152
8.1.1 声音文件格式	155
8.1.2 MIDI 和 MOD	158
8.1.3 语音软件	160
8.2 Internet 上的数字化视频	163
8.2.1 动画文件格式	164
8.2.2 视频软件	166
8.2.3 操作多媒体文件	168
8.3 小结	170

第Ⅲ部分 加入交互性

第9章 搜索与索引	171
9.1 WAIS 综述	172
9.2 WAIS 索引程序	174
9.2.1 建立一个索引	174
9.2.2 导出一个数据库	175
9.2.3 其他的 waisindex 选项	176
9.2.4 waisindex 过滤程序	178
9.2.5 增量索引	182
9.2.6 停止词	183
9.2.7 同义词表	183
9.2.8 扩展索引程序	184
9.3 WAIS 服务器	188
9.3.1 独立服务器	188
9.3.2 使用 Inetd	189
9.3.3 服务器安全性	190
9.4 检查你的 Web 服务器	191
9.5 小结	192
第10章 简单表格	194
10.1 表格的基本结构	194
10.1.1 Submit 和 Reset	197
10.1.2 <INPUT> 标记	198
10.1.3 <SELECT> 标记	201
10.1.4 <TEXTAREA> 标记	202
10.2 通用表格处理器 gform	203
10.2.1 使用 gform	204
10.2.2 可发送的文本	204
10.2.3 关键字	205
10.2.4 综合应用 HTML 中的交互元素	206
10.3 使用网关表格	208
10.3.1 什么是 fwais	210
10.3.2 fwais 的功能	211
10.4 小结	215
第11章 CGI:面向编程人员的高级表格	216
11.1 用 CGI 编程	216
11.2 <ISINDEX> 标记	217
11.3 处理 HTML 表格	221
11.3.1 表格结构与 CGI	221
11.3.2 URL 编码	223
11.3.3 POST 方法	224

11.3.4 GET 方法	227
11.4 动态生成 Web 页面	234
11.5 CGI 参考信息	240
11.5.1 CGI 环境变量	240
11.5.2 CGI 标题	243
11.6 小结	243

第 IV 部分 最终的考虑

第12章 整合:使你的页面融进 WWW 虚拟社区	245
12.1 地址与第一印象	245
12.1.1 站点一致性与通信量控制:设计及内容方面的考虑	249
12.2 选择并关注你的观众	251
12.3 小结	258
第13章 未来发展方向	259
13.1 HTML v3.0	259
13.1.1 定义文档	260
13.1.2 设计文档	261
13.1.3 润色文档:新标记和改进的标记	264
13.2 HTTP-NG	279
13.3 安全性与商业应用	280
13.3.1 安全协议	281
13.3.2 其他安全措施	284
13.4 虚拟现实模型语言	288
13.4.1 HotJava	293
13.5 远程会议技术与高级多媒体	295
13.5.1 MBONE	295
13.6 小结	299
附录 A 本书配套的联机资源库	301
附录 B 本书配套的 CD-ROM	302
附录 C Linux 硬件兼容性指南	305
附录 D Linux 的安装	322
附录 E Linux XFree86 指南	340
附录 F FTP 和 Gopher	355

第 I 部分 预备知识

第 1 章 什么是 Web

世界信息网 WWW(World Wide Web)是 Internet 上那些支持 WWW 协议和超文本传输协议 HTTP(Hyper Text Transport Protocol)的客户机与服务器的集合(现在已成为巨大的集合)。此时此刻,更多地世界各地的客户机与服务器正在进入 Web。每个联机服务提供者都在提供或已宣称不久即将提供 WWW 访问途径。本章将简要介绍一下建立世界信息网所用的概念:客户机、服务器及其协议。在本章中,我们将了解到世界信息网的历史与发展,以及如何才能紧跟上这一潮流。

1.1 客户机与服务器

客户机、服务器和协议虽然都是很简单的概念,但描述它们比理解它们更难。图 1-1 是一种简单的表达方式——两个椭圆由双箭头相连。

客户机是一个需要某些东西的程序,而服务器则是提供某些东西的程序。一个客户机可向许多不同的服务器请求。一个服务器也可以向多个不同的客户机提供服务。通常情况下,一个客户机启动与某个服务器的对话。服务器通常是等待客户机请求的一个自动程序。客户机通常是作为某个用户的请求或类似于用户的某个程序提出的请求而运作的。协议是客户机请求服务器和服务器如何应答请求的各种方法的定义。在本书中,一个 WWW 客户机又称作浏览器。



图 1-1 一个简单的客户机服务器模型图

通常的世界信息网上的客户机包括: Mosaic, Netscape 和 Lynx。通常的服务器来自于 CERN, NCSA 和 Netscape。如果你已使用过 Web,那么就一定有对客户机、服务器和协议概念的理解。我们先来看一下 Web 中客户机与服务器的具体任务。

在 Web 中,客户机的任务是:

- 帮助你制作一个请求(通常在单击某个链接点时启动)。
- 将你的请求发送给某个服务器。
- 通过对直接图象适当解码,呈交 HTML 文档和传递各种文件给相应的“观察器”

(Viewer),把请求所得的结果报告给你。

一个“观察器”是一个可被 WWW 客户机调用而呈现特定类型文件的程序。例如,特技表演文件不能由你的 WWW 客户机来呈现,而应该由特技表演“观察器”呈现。当一个声音文件被你的 WWW 客户机查阅并下载时,它只能用某些程序(例如 Macintosh 上的 Sound-Machine)来“观察”。

通常 WWW 客户机不仅限于向 Web 服务器发出请求(如 Gopher,FTP,news 和 mail),还可向其他服务器发出请求。

如图 1-2 所示,一个 Web 服务器的任务是:

- 接收请求。
- 请求的合法性检查,包括安全性屏蔽。
- 针对请求获取并制作数据,包括使用 CGI 脚本和程序、为文件设置适当的 MIME 类型来对数据进行前期处理和后期处理。
- 把信息发送给提出请求的客户机。

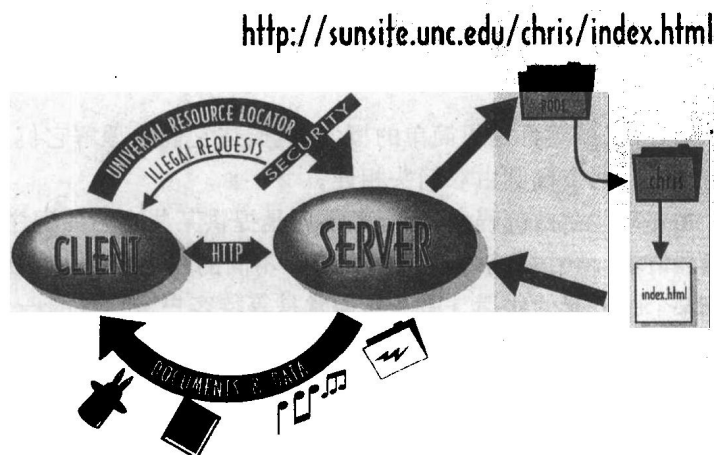


图 1-2 服务器的任务

Web 拥有一个被称为“无状态”的协议。这是因为,服务器在发送给客户机应答信息后便遗忘了此次交互。在“有状态”的协议中,客户机与服务器要记住许多关于彼此和它们的各种请求与应答信息。

Web 是一个易于实现的协议,因为无状态的协议是很轻松的(没有多少必需的核心代码和资源)。此种协议的另一吸引人的特性是,你可以很轻快地从一个服务器转向另一个服务器(在客户机端)或者从一个客户机转向另一个客户机(在服务器端),而无须过多的清理或跟踪。这种快速转移的能力对于超文本而言是非常理想的。(本章后面部分有一节专门介绍超文本的定义)。

然而,无状态环境下的协议处理是非常复杂的,仍然存在许多有待解决的技术问题。因此,Web 上的许多编程机制都试图给无状态协议加上状态。

Internet 和伴随它产生的一切是一个分布极为广泛的网络。它们支持标准的或者至少是有互操作性的协议,允许这种互操作性跨越商业领域和国界。也就是说,Internet 或 TCP/