

白强
丁蔚斌
李斌

著

高级实用

WWW技术



中国科学技术大学出版社

高级实用 WWW 技术

白強 丁蔚 李斌 著

中国科学技术大学出版社

1999 · 合肥

内 容 提 要

本书是为那些具有一定计算机和网络知识, 并想学习如何建立一流的 Web 站点的人而编写的。本书并不涵盖所有 WWW 技术细节, 而是在阐述一些理论知识的基础上着重介绍实用的方法和技巧, 并尽可能多地介绍该领域的最新进展。

本书提供建立 WWW 服务器和 Web 站点所需要的各种实用知识, 包括通用 WWW 图像处理技术、HTML 编程、CGI 编程、面向 Web 的 Java 技术、Javascript、微软的 ASP 以及层叠风格页 CSS。本书根据白强先生多年的工作经验编写而成, 书中的例子都经过实际运行检验, 所介绍的方法全部可以即学即用。

本书适用于从事网络工程的技术人员和用户, 也可供高等院校信息和计算机专业的高年级学生和研究生阅读参考。

图书在版编目(CIP)数据

高级实用WWW技术 / 白强等著. —合肥: 中国科学技术大学出版社, 1999.3

ISBN 7-312-00905-0

I. 高... II. 白... III. 万维网-HTTP IV. TP393.4

中国版本图书馆 CIP 数据(98)第 4082 号

中国科学技术大学出版社出版发行

(安徽省合肥市金寨路 96 号, 邮编: 230026)

合肥晓星印刷厂印刷

全国新华书店经销

开本: 787 × 1092 / 6 印张: 15 字数: 360 千

1999 年 3 月第 1 版 1999 年 3 月第 1 次印刷

印数: 1—5000 册

ISBN 7-312-00905-0 / TP · 178 定价: 20.00 元

前 言

正如书名“高级实用WWW技术”所提示，“高级”与“实用”是本书的特色。本书不是为初学者编写的入门指导，不是罗列全部HTML标记的WWW参考手册，也不是提供一连串干巴巴CGI例程的服务器开发指南。本书是为Web站点管理员、高级Web开发人员、从事实际工作的Web技术人员而写的。本书假设读者已具有操作系统和编程语言的知识，了解Internet和WWW的基本概念。我们的目标是让读者掌握从头开始建立第一流Web站点的实用技术和技巧，同时，也介绍相应的Web理论。本书重点讨论如何建立和维护Web服务器，如何由简到繁创建主页，HTML编程，CGI编程，面向Web的Java和Javascript技术，以及微软的ASP技术和层叠风格页CSS。

本书采用如下方式展开：假设有一家新开业的公司，需要创建和维护一个Web站点。该公司的网络技术人员包括系统管理员、Web管理员以及若干普通开发人员，那么这家公司该怎么做呢？本书一步一步展示该公司是如何做的。首先，建立Web服务器；接着，配置Web站点并创建主页；然后，为各种应用和服务添加CGI程序；最后，使用Java小程序、Javascript、ASP、CSS等高级功能使站点达到一流的水准。本书类似这家公司的Web工作日记。

本书不打算包括有关HTML编程的所有方面，因为现在市场上有不少HTML参考书和HTML编辑软件。本书也不打算教授如何编写Perl或C程序，假设读者已经具备了Perl和C编程的知识。本书也不想成为Java编程的教科书，关于Java本身就要用几本书去讲解。本书包括尽可能多的HTML、CGI和Java编程的实用知识以及精心准备的一批演示功能很强的例子。我们不保证本书含有关于WWW的所有知识，但是，我们保证本书将包含这方面足够多的知识和技巧，掌握了这些知识与技巧，读者就能创建有较高水平和品味、足以和大多数现有Web站点相媲美的站点。

本书由9章组成。第一章WWW简介，第二章建立完整的WWW环境，第三章WWW图像，第四章使用HTML创建主页，第五章公网网关接口CGI，第六章Java，第七章Javascript，第八章微软ASP技术，第九章层叠风格页。书中讨论的内容都来自作者近几年的实际工作经验，并尽可能多地包括WWW领域的最新进展。

郝炳焜、谭立湘同志参加了一些章节的编写，白方周教授对本书的出版做了大量的组织、协调工作，还有许多同志为本书的出版付出了辛勤的劳动，作者向他(她)们表示衷心的感谢。

由于作者水平所限，书中欠妥乃至错误之处在所难免，请读者不吝指正。

作者

1998.10

目 录

前言I

1	WWW 简介	1
1.1	概述	1
1.1.1	Internet	1
1.1.2	超文本和超媒体	1
1.1.3	超文本传输协议 HTTP	2
1.2	WWW 的过去、现在和未来	2
1.2.1	WWW 的诞生	3
1.2.2	Mosaic	3
1.2.3	Netscape	4
1.2.4	Java 语言	5
1.2.5	基于 Web 的数据库	5
1.2.6	WWW 和电子商务	6
1.2.7	插件	6
1.2.8	WWW 图形功能	6
1.2.9	WWW 发展趋势	7
1.3	WWW 的概念和术语	7
2	建立完整的 WWW 环境	11
2.1	概述	11
2.1.1	确定人员	11
2.1.2	确定系统	11
2.1.3	选择 Web 服务器软件	12
2.1.4	获得 Internet 域名	13
2.2	安装 NCSA 服务器	13
2.2.1	概述	13
2.2.2	安装	15
2.2.3	服务器配置文件 httpd.conf	16
2.2.4	资源配置文件 srm.conf	20
2.2.5	访问配置文件 access.conf	25
2.2.6	ABC 公司的配置	30
2.3	安装网景服务器	30
2.4	安全套节字层协议 SSL	31

II

2.4.1	SSL 与网景 Commerce server.....	31
2.4.2	申请证书.....	32
2.4.3	身份认证.....	32
2.4.4	加密.....	32
2.4.5	数据完整性.....	32
2.4.6	证书代理及其要求.....	33
2.4.7	服务器安全系统的安装.....	33
2.5	Windows NT 与 Internet 信息服务器 IIS.....	33
2.5.1	安装 IIS.....	34
2.5.2	文件的安装.....	34
2.5.3	远程管理.....	35
2.6	后记.....	35
3	WWW 图像.....	36
3.1	概述.....	36
3.1.1	生成图像的工具.....	36
3.1.2	术语.....	36
3.2	GIF 的高级功能.....	38
3.2.1	透明 GIF 和交织 GIF.....	38
3.2.2	交织 GIF 的进一步说明.....	38
3.2.3	GIF 动画.....	39
3.3	后记.....	40
4	使用 HTML 创建主页.....	41
4.1	HTML 简介.....	41
4.2	HTML 文件的结构.....	42
4.2.1	概述.....	42
4.2.2	头部段.....	43
4.2.3	正文段.....	44
4.2.4	语义格式化元素.....	47
4.2.5	物理格式化元素.....	48
4.2.6	字符引用与实体引用.....	50
4.2.7	图像.....	55
4.2.8	命名链接.....	56
4.3	演示.....	57
4.3.1	创建第一个 HTML 文件.....	57
4.3.2	列表.....	58
4.3.3	背景.....	62
4.4	高级功能与新增特性.....	64

4.4.1	字体元素.....	64
4.4.2	新增格式化元素.....	65
4.4.3	图像映射.....	65
4.4.4	表格.....	66
4.4.5	分框.....	72
5	公共网关接口 CGI.....	76
5.1	概 述.....	76
5.1.1	浏览器与 CGI 程序交互作用的示例.....	77
5.1.2	数据表单的数据传输.....	77
5.1.3	虚拟文件.....	78
5.1.4	URL 编码.....	79
5.2	示例程序.....	80
5.2.1	Hello World 程序	80
5.2.2	简单回显程序 Bounce	80
5.2.3	HTML 乘法程序	81
5.2.4	包装程序 Swords.....	82
5.2.5	完全交互式程序 Hangman	83
5.2.6	注释与数据表单采集程序.....	85
5.2.7	数据表单和 CGI 程序配合的例子.....	87
5.2.8	处理电子邮件的 CGI 程序.....	87
6	Java	91
6.1	概述.....	91
6.2	Java 的特点.....	92
6.3	理解 Java 的要点.....	95
6.4	Java 程序的例子.....	98
7	JavaScript.....	133
7.1	概述.....	133
7.2	Javascript 基础.....	135
7.2.1	事件.....	136
7.2.2	函数.....	136
7.2.3	Javascript 的层次结构.....	139
7.2.4	地址对象.....	143
7.2.5	分框.....	144
7.2.6	导航栏.....	146
7.2.7	窗口.....	150
7.2.8	自动创建文件.....	153
7.2.9	状态栏.....	157

7.2.10	定时器.....	158
7.2.11	移动文本.....	158
7.2.12	日期对象.....	160
7.2.13	数组对象.....	162
7.2.14	数学对象.....	163
7.3	JavaScript 高级功能	164
7.3.1	验证数据表单输入.....	164
7.3.2	检查特定字符.....	165
7.3.3	传送数据表单输入.....	166
7.3.4	设置镜头.....	167
7.3.5	Web 页面的图像.....	168
7.3.6	层次.....	171
8	微软的 ASP 技术	180
8.1	概述.....	180
8.1.1	ASP 的特点	180
8.1.2	ASP 与 OLE.....	182
8.1.3	ASP 注册表	182
8.1.4	创建一个简单的 ASP 页面	182
8.1.5	ASP 与数据表单	183
8.2	ASP 的内置对象.....	184
8.2.1	应用对象.....	184
8.2.2	会话对象.....	185
8.2.3	请求对象.....	189
8.2.4	响应对象.....	196
8.2.5	服务器对象.....	200
8.2.6	服务器构件.....	202
8.2.7	global.asa 文件	206
9	层叠风格页.....	208
9.1	概述.....	208
9.2	基本语法.....	208
9.2.1	被选项.....	208
9.2.2	其它构件.....	209
9.2.3	伪类和伪元素.....	210
9.2.4	层叠次序.....	211
9.3	CSS1 特性.....	212
9.3.1	特性定义的语法.....	212
9.3.2	字体特性.....	212

9.3.3 颜色和背景特性..... 214

9.3.4 文本特性..... 217

9.3.5 边框特性..... 220

9.3.6 分类特性..... 227

9.3.7 单位..... 229

参考文献.....231

1 WWW 简介

1.1 概 述

半个世纪以来，人们一直梦想着能有一种通用的信息数据库，不仅全世界的人可以读取信息，而且，通过信息与信息间的联结，用户能很快检索到重要的数据。在本世纪 60 年代，人们对这种思想进行了深入的研究，形成了新概念“docuverse”，就是可以任凭人类思想遨游的文档世界。它会使人类信息交流的各个方面发生重大的变革。但是 1990 年，WWW 技术的出现才使这一梦想成为现实。

WWW(World Wide Web，即万维网，亦简称 Web)的权威定义是：建立在超文本传输协议 HTTP 之上的广域分布式超媒体信息检索系统，用于实现大量文档的通存通取。WWW 在计算机网络上，用简单明了的风格，为用户提供访问各种媒体信息的统一方法。通过浏览器软件，如网景 Navigator，WWW 改变了人们浏览和创建信息的方法，创建了第一个真正的全球超媒体网络。

1.1.1 Internet

Internet 是建立在 TCP/IP 协议上的世界范围的计算机网络。Internet 字面上的意思是网间网，它由全球各地成千上万个大小网络互联而成。每天，它要连接一百多个国家和地区的大约三千多万个用户。WWW 在 Internet 上使用。但是它们不是一回事，WWW 指的是一个信息体，一个知识的抽象空间，而 Internet 指的是全球计算机网络的物理连接，即大量连接着的电缆、光缆和计算机。

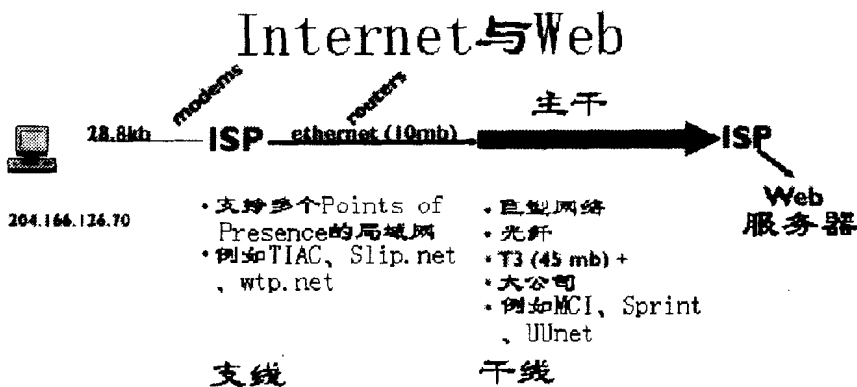


图 1.1 Internet 与 Web

Internet 由美国的军事网络 ARPANET 演化而来，可追溯到 60 年代末。70 年代初，由美国国防部开发的 TCP/IP 协议向公众开放。1982 年，美国国防部决定 ARPANET 上的所有系统

必须转换到TCP/IP协议。在以后的几年里，互联网协议IP被广泛地安装。直到九十年代，才出现商业性服务。随着WWW的出现，Internet开始成为许多普通人生活的一部分。

1.1.2 超文本和超媒体

WWW的运行依赖于超文本，以它作为与用户交互的手段。超文本基本上与普通文本相同——可存储、可读取、可检索、可编辑。但是，它们有一个很重要的区别：超文本包含与其他文档的链接(Link)，普通文本却没有。

例如，你在超文本中点击一个链接，即突出显示的单词或短语，如“hypertext”，会有一个或多个关于“超文本”的文件出现在你面前，或者介绍“超文本”的历史，或者描述Web管理员对“超文本”的定义。这些文件又能链接到其他文件，连续选择超文本将带你进入自由自在的信息之旅。通过这种方式，超文本链接可创建复杂的、包罗万象的Web链接网络。

超媒体也是超文本，所不同的是：超媒体文件不仅包含对其他文本的链接，而且包含对其他媒体形式的链接，如语音、图像、电影等。图像也能用于链接声音或文本。

1.1.3 超文本传输协议 HTTP

HTTP(Hypertext Transfer Protocol)是为WWW而创建的网络协议，用于创建分布式协作超媒体信息系统。用Tim Berners-Lee自己的话说，HTTP是通用、无状态、面向对象的协议，易用易实现。

利用HTTP，可以浏览信息库，从Internet获取海量的信息资源，并可从任何资源查询相关的资源。

HTTP的功能包括查询、检索、前端更新、文档注释。HTTP允许设计和配置一个可扩展的方法集。HTTP还有通过客户/服务器协商提供动态数据表示的功能。这使WWW信息系统的建设与新信息表示方式的开发及配置无关。信息表示方式可作为从服务器到客户(或反之)传送数据过程的一部分。

HTTP建立在统一资源标识符URI的基础上，URI是Internet工程任务工作组(IETF)精心规划的方案。有两类URI：统一资源名URN和统一资源定位符URL。

在Internet上，网络通讯是通过TCP/IP连接进行的。偶尔会出现这个协议被安装在另一个网络协议之上的情况。在这种情况下，HTTP对象被映射到外来传输机制或网络上。这是一个简单、直接的过程，通常不为WWW用户注意。

HTTP协议是无状态的，即客户和服务器都不存储对方的状态信息。服务器不必存储为之提供服务的客户的信息，客户也不必跟踪服务器的事件状态。对于超媒体系统，响应时间要足够短，这需要快速、无状态的协议。无状态使得HTTP可以支持最简单的客户/服务器应用程序，并有助于在WWW世界中获得广泛的访问支持和平台支持。

许多网络协议建立在“事务”的概念之上，这包括以下要素：

- 客户和服务器之间建立连接，以进行通信。
- 从客户发到服务器的、对某些特定的服务、资源或其他已知的通信主题的请求。
- 服务器对客户请求的响应，向客户提供所请求的资源，或解释不能传送的原因。

- 连接的中止，也就是请求/响应通信的结束。

连接

这是指建立从客户端到 Web 服务器的 TCP/IP 连接。当在 UNIX 服务器上使用 TCP/IP 协议时，端口 80 是众所周知的 WWW 通讯端口。

请求

请求由客户端以请求报文的格式初始化，发往特定的服务器。该请求包括将应用于被请求的网络对象的方法、该对象的标识符、所用 HTTP 协议的版本，以及包含关于传送报文的信息的请求头。

响应

由服务器发出的响应，返回到请求服务的客户端。它由三部分组成：响应状态行、响应头和响应数据。

响应状态行包括：服务器所用的 HTTP 版本的标识符，描述服务器执行客户请求的结果的状态码，关于服务器响应的注释。

响应头一般包含关于被请求对象的信息。此外，还可能包含 MIME(多用途互联网邮件扩展)声明，如 Content-Type，用于传送响应。作为一种确认，响应头通常包含与客户端提供的请求头相同的内容。

响应数据紧跟在响应头之后，提供附加信息。数据一般用 MIME 消息格式编码。数据是否恰当完全取决于服务器返回的状态码。应用于数据的 MIME 声明 Content-Type 可以是任何客户指明可以接受的类型。两个常用的 Content-Type 类型是：text/plain(普通的 ASCII)和 text/html(普通的 HTML 文档)。根据 HTTP 的定义，可以假设任何 WWW 客户都能处理这两种数据。

关闭

TCP/IP 连接的终止，可由服务器或客户端任一方发出。

1.2 WWW 的过去、现在和未来

1.2.1 WWW 的诞生

Web 开始于 1989 年 3 月，由当时在欧洲核物理研究中心(CERN)工作的 Tim Berners-Lee 首先提出，目的是为在该组织内部传播研究成果和思想提供有效的方法。有效地交流是 CERN 多年来的目标，因为它的成员分布于多个不同的国家。1990 年，在瑞士日内瓦 CERN 的一个地下室，WWW 诞生了。

最初的草案，勾画了一个在 CERN 的成员之间利用网络化的超文本传送文件、进行通讯的简单系统，没有考虑传送语音、图像或视频的能力。Tim Bernes-lee 的思想来源于 TCP/IP、MIME、SGML、客户/服务器体系开放系统构想。他综合 UNIX 协议集中几个较通用的部分，设计了简单有效的 WWW。作为结果，形成了两个既成事实的标准——HTTP 和 HTML，它们为 WWW 的巨大成功奠定了基础。

1990 年底，第一个 WWW 软件被移植到 NEXT 机。它不仅能在 Internet 上浏览和传送超文本文档给其他人，还可以在屏幕上编辑超文本。开发人员把该软件演示给 CERN 的专家，并在 1991 年超文本会议上做了演示。

1991 年夏，Tim Berners-Lee 在 Internet 上公布了该 WWW 软件，此后，该软件以一种出乎开发人员预料的速度在全球范围内蔓延开来。Tim Berners-Lee 创建了 WWW 协会，此协会由 MIT 计算机科学实验室、法国计算机科学和自动化研究所和日本 Keio 大学联合运作，为 WWW 开发公共的标准，Tim Berners-Lee 一直是该协会的会长。至今已有 150 多个组织成为该协会的成员。Tim Berners-Lee 始终坚信 WWW 应是开放的，不应被垄断的商业组织利用，他的这一信念促进了 WWW 的飞速发展。

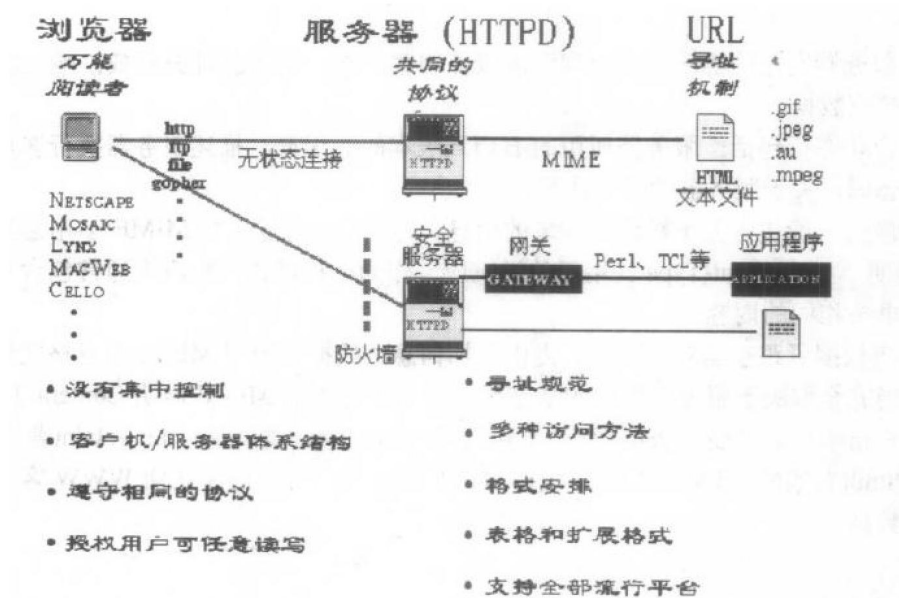


图 1.2 Web 的工作原理

以下是 WWW 发展过程中的一些统计数字:

1993 年 6 月，130 个 WWW 站点，1.5% 是 com 网点，每个服务器 13000 个主机。

1993 年 12 月，623 个 WWW 站点，4.6% 是 com 网点，每个服务器 3475 个主机。

1994 年 6 月，2738 个 WWW 站点，13.5% 是 com 网点，每个服务器 1095 个主机。

1994 年 12 月，10022 个 WWW 站点，18.3% 是 com 网点，每个服务器 451 个主机。

1995 年 6 月，23500 个 WWW 站点，31.3% 是 com 网点，每个服务器 270 个主机。

1996 年 1 月，100000 个 WWW 站点，50% 是 com 网点，每个服务器 94 个主机。

1999 年 1 月，有 5012287 个 WWW 站点，60% 是 com 站点。

1993 年，在 Internet 上有 130 万台主机，1994 年有 220 万台，1995 年有 490 万台，1996 年 1 月有 950 万台。

1.2.2 Mosaic

CERN 提出 WWW 草案后的几个月，位于伊利诺斯州 Champaign 市的美国国家超级计

算应用中心(NCSA)启动了一个项目,开发访问 WWW 的界面。NCSA 的一个目的是向科研人员提供容易获取的非商业 Web 软件,它的另一个目标是开发可以得到商业利益的新技术。1993 年初, NCSA 软件设计组开发成功了访问 WWW 的通用多平台界面,称其为 Mosaic。

1993 年上半年, NCSA 在 Internet 发布了 Mosaic 浏览器的第一个版本,吸引了越来越多的普通应用软件开发人员和 Internet 用户。

1.2.3 网景公司

NCSA Mosaic 的成功导致了对 WWW 商用软件和全球网络服务的迫切需求。网景通信公司(Netscape Communications)的创立顺应了这个潮流。

Jim Clark 和 Mark Anderson 于 1994 年 4 月创立了网景通信公司,公司位于硅谷的心脏——维纽山,到 1995 年就有了超过 75 人的雇员队伍。网景公司的目标是成为 Internet 上信息交换软件和商务软件的主要供应商。

Jim Clark 是网景公司的主席。在创建网景公司之前,Clark 是硅图公司(SGI)的主席,这个他于 1982 年创立的公司年收入达 15 亿美元,是《幸福》杂志评出的美国五百家发展最快的公司之一。在创立 SGI 之前,Clark 是斯坦福大学的副教授,在那里他和研究生开发了 SGI 最初依赖的技术。Clark 于 1994 年 2 月辞去 SGI 主席一职,和开发 Mosaic 的年轻程序员们开始了新的冒险。Clark 拥有犹它大学的计算机科学博士学位。

Mark Anderson 是网景公司技术副总裁。1992 年秋,当他还是伊利诺斯大学的学生时,就已成为 NCSA 的成员,是 NCSA Mosaic 浏览器的主要设计和编程人员之一。1993 年初,他和 NCSA 的同事开发了一个友好易用的 Internet 导航工具。在网景公司,Mark 负责确立和监督公司的技术方向。他于 1993 年获得了伊利诺斯大学的计算机科学学士学位。

除 Anderson 外,网景公司的核心技术组还包括 Mosaic 7 个开发者中的其他 5 个。

网景公司提供全面的开放软件,致力于开发电子商务软件和安全信息交换软件,这些软件应用于 Internet 和基于 TCP/IP 的局域网。网景的所有产品都与其他基于 HTTP 的客户/服务器软件完全兼容,它们向那些想在全球性网络上提供或获得信息服务的公司和个人提供安全的通讯、先进的性能和简化的操作。

直到现在,网景公司仍然是 WWW 浏览器市场的领头羊,占有超过 70% 的市场份额。尽管微软公司通过发布 Internet Explorer 浏览器向网景发起了有力的冲击,但其自身并没有特别令人感兴趣的标准。Internet 深深根植于 Unix 用户,WWW 服务更适合于 Unix 系统的事实,挫败了微软占据浏览器市场头把交椅的企图。

1.2.4 Java 语言

从 1995 年第二季度开始,Java 逐渐变成 Internet 与 WWW 开发技术中最具吸引力的部分。Java 是为 Internet 编程而设计的语言,由太阳微系统公司(SUN)发布。它是一种面向对象、分布式、解释性、安全、与结构无关、可移植、高性能、多线程的动态编程语言。Java,尤其是 Java Applet(小应用程序),把 WWW 带入了全新的境界。Applet 是小型的 Java 程序,可以嵌入另一个应用程序之中运行,绝大多数情况下嵌入 Web 浏览器(如网景的

Navigator 2.0 或更高版本)。Java 是与结构无关的、十分理想的 Internet 编程语言, 因为 Java 程序能在连接到 Internet 的所有计算机上运行, 不论它在什么平台上, 使用什么操作系统。过去一直嘲笑 WWW 开发技术(通常是用 HTML 编程)的开发人员现在遇到了真正的挑战。随着 Java 新版本的发布, 可以预见在 WWW 开发中将出现更加实用的 Java 语言。

随着 Java 热, 网景开发了脚本语言 JavaScript, 它比 Java 简单, 介于真正的编程语言和脚本语言之间。它可将 WWW 站点带人更高的水平, 前提是用户使用网景的 Navigator 浏览器。

1.2.5 基于 Web 的数据库

将 Web 作为数据库的前端, 可以吸引大量用户访问数据库。可以说 WWW 生来就是和数据库绑在一起的。当前许多应用程序都用 WWW 作前端, CGI 程序作桥梁, 去访问关系数据库。嵌入式 SQL 语句能更加直接地实现以上目标, 同时大大拓宽了可作为 Web 应用程序的软件的范围。面向对象数据库和 Web 的结合, 使 Web 服务器自身成为一个巨大的数据库。

1.2.6 WWW 与电子商务

人们讨论电子商务已经很长时间了, 但由于种种制约, 至今仍未取得重大进展。随着 WWW 的快速发展, 通过 WWW 从事电子商务成为 WWW 的一大特色。尽管 WWW 的设计目的为信息交换, 它还是给电子商务带来了巨大的可能性。WWW 容易访问, 操作简便, 一天二十四小时都开通, 不必支付额外费用。上网用户只要拥有 WWW 浏览器, 就不必购买额外的商务软件。因为 Internet 不像专线那样安全, 安全就成为主要问题。但随着先进的安全算法的实现, 这将不再是用户的主要顾虑。总之, WWW 开辟了一个商机无限的虚拟世界, 电子商务也使 WWW 超越了研究和娱乐的层次。

1.2.7 插 件

插件(Plug in)的思想由网景首先提出。插件的意思是: 当使用 Web 浏览器漫游 WWW 世界时, 假如浏览器遇到某种特定格式的文件, 浏览器将自动搜索用户本地系统, 寻找支持此文件格式的软件, 一旦它找到了这个软件, 它就将启动该软件, 把该格式的文件正确地呈现给用户。譬如, Real Audio 是一个插件软件, 用于把音频格式文件通过 WWW 实时提供给用户。“插件”的确扩大了 WWW 内容的范围, 特别在音频和视频方面, “插件”至今仍是最好的解决方案。随着网景 Navigator 4.0 的发布, “插件”将成为浏览器的一部分, 这将消除用户在本地系统安装“插件”软件的麻烦。

1.2.8 WWW 图形功能

人们喜欢 WWW 的一个重要原因是它有强大的图形功能。通过 Web 浏览器, 可以浏览各种精美的图片和令人陶醉的设计图案。这方面的趋势是动画、三维图像和 VRML(虚拟现实操纵语言)。

随着越来越多的图形进入 WWW 世界, 网络传输速度(带宽)成为越来越严重的问题。已进行了许多解决带宽问题的研究, 但至今还没有较好的解决方法。

1.2.9 WWW 的发展趋势

除了以上几个发展方向, 人们正投入越来越多的精力开发 WWW 计算机和 WWW 电视。WWW 计算机是简化的主要作为 WWW 浏览器的计算机, 拥有普通 PC 机的部分功能。WWW 电视是把电视和 WWW 结合起来的令人神往的构想, 将使 Internet 和有线电视一样受大众喜爱。

可以预见, WWW 将进入一个缓慢发展的时期。目前人们基本是把各种已有的技术引进 WWW 世界, 一旦这个时期结束, 人们将努力为 WWW 寻找新的特色。

1.3 WWW 的概念和术语

和进入数学世界一样, 进入 WWW 世界, 你将遇到一些新概念。下面不给出严格的定义, 以通俗的语言解释这些概念, 使读者能基本理解。

协议

协议是正式的通信行为规则。当计算机之间进行通信时, 必须定义一组规则来管理通信过程, 这组规则叫做协议。

IP 地址

一组用句点“.”分开的四个数字, 这些数字都介于 0 到 255 之间, 用于唯一确定一个 Internet 地址。IP 地址包含网络部分和主机部分。

TCP/IP

IP(Internet Protocol)是将 Internet 上的计算机连接在一起的协议。TCP(Transmission Control Protocol)是传输控制协议, 位于 IP 的上层。TCP/IP 协议组用于操作系统和机器结构无关的开放协议创建异构网络。任何人都可以获得 TCP/IP 协议, 开发满足这些协议的产品。TCP/IP 还在发展和完善之中。

套节字和端口

在 TCP 中, 完整的源地址或目的地址叫做套字节(socket)。主机内的用户或进程标识叫做端口。应用程序进程(网络服务)由 16 位的端口号标识。IP 地址和端口号的组合叫做套节字。套节字在整个 Internet 上唯一地标识一个网络进程。一对套节字(一个对应接收主机, 一个对应发送主机)定义一个面向连接协议(如 TCP)的网络连接。

公共网关接口 CGI

为进行 WWW 的动态交互, 需要一个在 HTTP 服务器的指示下运行外部程序的接口。CGI 提供了这个功能, 它是 WWW 和其他资源(如 SQL 数据库)之间的通道。CGI 应用程序能为 SQL 查询提供一种机制, 从 WWW 服务器传送数据到发出查询请求的客户端。该程序通过 WWW 执行, 不能直接浏览。

CGI 定义了 Web 服务器运行外部应用程序的接口。可以把 CGI 视为一种扩展 WWW 服务器功能的机制，通过服务器上的 CGI 程序，可以扩充你的站点的功能，使之与众不同。

超文本标记语言 HTML

HTML 是标准通用标记语言 SGML 的一个应用。SGML(ISO-8879)在 1986 年被批准为国际标准。SGML 提供了一种表示文本和超文本结构的方法，它也是对超文本进行编码以使它们能相互交换的一种方法。SGML 还是描述文档标记系统的形式分析语言。

WWW 服务器

为了让世界上其他地方的人能通过 Web 访问某个局域计算机系统，即 Internet 的一个节点，这个计算机系统必须运行 WWW 服务器软件。服务器可以是通用服务器，如 `http://www.stat.purdue.edu`，它支持所有本地用户的主页；或者是个人服务器，如 `http://b.stat.purdue.edu:1280`，是笔者白强用过的个人服务器。

WWW 服务器的名字总是以 `http://` 开头。通用服务器的缺省端口号为 80，通常可以省略。个人服务器总是以一个大于 1250 的数字结尾。前几年流行的服务器软件是 `httpd`，由 NCSA 提供，现在是 1.4 版。

主页

主页(homepage)是在互联网上传递信息的手段。它可能属于某个用户，包含用户自己的图片、研究项目、简历、兴趣等。它也可能属于某个组织，如某大学的主页。假定某人访问笔者白强的服务器 `http://b.stat.purdue.edu:1280`，他实际上是访问白强的主页 `homepage.html`，这是第一级主页，也可以命名为 `index.html`、`welcome.html` 等。

WWW 浏览器

浏览器是访问 WWW 的工具，如网景的 Navigator、微软的 Internet Explorer、Mosaic 等。网景的 Navigator、微软的 Internet Explorer 是其中的佼佼者。浏览器连接到 WWW 服务器，接收信息，把它们转换为用户能理解的形式。大部分浏览器支持图像。



图 1.3 浏览器与 Web 服务器的交互作用

通常，为了使浏览器工作，需要直接的互联网连接。没有直接连接，安装浏览器将遇到一些困难，但仍然可以通过特殊的浏览器访问 Web，当然速度很慢而且性能有限。对拨