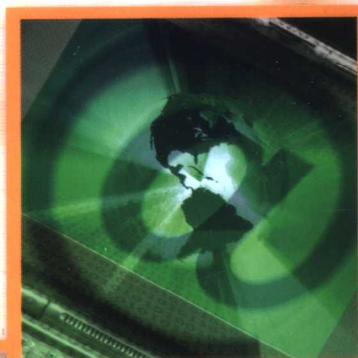


图解 **网络** 技术丛书

# web

# 技术

—— HTTP 到服务器端



[日] 小泉 修 著  
王 浩 译



科学出版社

[www.sciencep.com](http://www.sciencep.com)

图解网络技术丛书

# Web技术

—— HTTP到服务器端

[日] 小泉修 著  
王浩 译



科学出版社

北京

图字: 01-2004-3378 号

## 内 容 简 介

本书是“图解网络系列”之一。本书从互联网的机构和 Web 的功能入手,主要介绍互联网协议和 HTTP,服务器/客户端的 Web 技术,构筑 Web 所使用的各种最新技术,以及 Web 服务器和 Web 应用程序服务器的构筑,Web 服务器的安全对策等。

本书内容详实,实用性强,结合丰富的图表,使枯燥乏味的学习变得更加轻松有趣。

本书可供正从事网站开发、设计和管理工作的初学者参考,也可供 Web 技术爱好者阅读。

### 图书在版编目(CIP)数据

---

Web 技术/(日)小泉修著;王浩译. —北京:科学出版社,2004  
(图解网络技术丛书)

ISBN 7-03-013606-3

I. W… II. ①小… ②王… III. 主页制作-程序设计-图解  
IV. TP393.092-64

中国版本图书馆 CIP 数据核字(2004)第 064618 号

---

责任编辑:杨 凯 崔炳哲 / 责任制作:魏 谨

责任印制:刘士平 / 封面设计:李 祥

**科 学 出 版 社** 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

**新 蕾 印 刷 厂** 印刷

**北京东方科龙图文有限公司** 制作

<http://www.okbook.com.cn>

科学出版社发行 各地新华书店经销

\*

2004 年 9 月第 一 版 开本: B5(720×1000)

2004 年 9 月第一次印刷 印张: 18

印数: 1—5 000 字数: 224 000

**定 价: 27.00 元**

(如有印装质量问题,我社负责调换〈新欣〉)

---

---

# 前 言

---

现在,Web 相关的技术人员大量匮乏。

在企业内外,把掌握 Web 相关知识的技术人员称作 Web 管理员,Web 管理员的培养无法赶上 Web 技术的急速发展及普及的速度。

最初,Web 管理员的主要业务内容只限于整理企业应该发送的信息,使用称为 HTML 的语言制作主页,将用于发送信息的计算机上登录到 Web 服务这样的范畴。

但是现在,企业内局域网及各种互联网服务器构建的知识,提高 Web 站点服务器所必需的各种程度的设计和制作手法,利用称为 XML 的新型语言进行企业内部信息的结构化,再加上与之相对应的通信顺序的控制,面向对象的理解等等,对 Web 管理员提出了更高要求。

因此,重新培养一个 Web 管理员需要很长时间,这将无法满足日益增长的对 Web 管理员的需求。

在日本,具有面向企业的 Web 站点,在同行业企业中进行电子商务和信息交换的企业还不足 10%,在今后的几年里,其普及率估计将要超过 30%。即使是现在,对 Web 管理员的需求还处于不足状态。今后,预计将扩大数倍。即到目前为止,有必要以上面的速度培养 Web 管理员。

但是,由于关于 Web 的历史较短,技术进步速度过快,目前还没有确立起有效的教育体系。

本书目的是探索培养 Web 管理员的最短途径。以 Web 技术爱好者,Web 管理员,以及有必要培养 Web 管理员的企业及其担当者,甚至于今后打算参与到 Web 商业活动的企业的管理者为对象,介绍有关 Web 的最新知识。

关于 Web 的各类技术书籍已经出版了很多,但是在学习特殊技术之前,有必要对 Web 技术有个全面的了解。这是因为为了构筑较好的 Web 站点,只靠一部分特殊技术是难以实现的,这需要较之更高的技术。

如果拿到本书的朋友们正从事 Web 相关工作,或者以 Web 管理员为目标的话,则可以说一定程度上将未来掌握在手中。

今后 Web 技术将与商业活动具有密切关系,并持续发展下去。而且,那时,大家所掌握的技术一定能成为众多企业所不可或缺的部分。

让我立刻加入到 Web 世界的旅行吧!

这数年间,雨后春笋般普及并发展的 Web 将是个怎样的世界!

著者

# 目 录

## 第 1 章 互联网的机构和 Web 的机能

1.1 Web 的发展和检索功能 .....	2
1.1.1 何谓 Web .....	2
1.1.2 Web 的起源 .....	3
1.1.3 Web 浏览器的发展和普及 .....	5
1.1.4 Web 浏览器的功能 .....	7
1.2 实现 Web 的计算机网络 .....	10
1.2.1 计算机网络的优点 .....	10
1.2.2 LAN 和客户机/服务器系统 .....	11
1.2.3 基于 WAN 的 LAN 之间的连接 .....	14
1.3 在互联网上实现通信 .....	15
1.3.1 互联网的概念 .....	15
1.3.2 互联网上的标准协议 .....	16
1.3.3 IP 地址的必要性 .....	17
1.3.4 端口号的必要性 .....	18
1.4 互联网的信赖性 .....	20
1.4.1 数据传送时的故障 .....	20
1.4.2 互联网的高度信赖性 .....	21
1.4.3 数据通信中存在的问题 .....	22
1.4.4 数据包通信的优点 .....	23

1.4.5	数据包通信的信赖性的确保	24
1.5	根据 Web 对信息的访问	25
1.5.1	互联网是 Web 服务器的集合体	25
1.5.2	通过 URL 对资源定位	26
1.5.3	基于 HTTP 的客户机/服务器 之间的交换	29
1.6	Web 服务器站点的动态处理机能	33
1.6.1	服务器和服务器软件	33
1.6.2	Web 服务器所实现的动态处理	35
1.6.3	基于 CGI 的动态处理功能	36
1.6.4	Web 映射的模式	38
1.6.5	SSI 的功能	41
1.6.6	基于 Web 浏览器的动态处理功能	41

## 第 2 章 互联网协议和 HTTP

2.1	所谓的协议	48
2.1.1	协议的必要性	48
2.1.2	协议的标准化的	50
2.2	OSI 参考模型的层次构造	52
2.2.1	基于 OSI 参考模型的标准化的	52
2.2.2	协议层的概念	53
2.2.3	由 7 层构成的 OSI 参考模型	55
2.3	世界标准的 TCP/IP 协议	56
2.3.1	TCP/IP 的特征	56
2.3.2	OSI 参考模型和 TCP/IP 协议群的比较	57
2.3.3	TCP/IP 协议群的内容	58

2.4 TCP/IP 协议群各层的功能 .....	61
2.4.1 应用层的功能 .....	62
2.4.2 传输层的协议——TCP .....	63
2.4.3 TCP 程序段格式的内容 .....	64
2.4.4 代码位的功能 .....	65
2.4.5 网络层的协议——IP .....	67
2.4.6 网络接口层的功能 .....	72
2.5 HTTP 的功能和运行 .....	74
2.5.1 HTTP .....	74
2.5.2 HTTP 和下位层的功能 .....	75
2.5.3 基于 HTTP/1.0 的 Web 信息的交换 .....	76
2.5.4 基于 HTTP 的请求消息 .....	79
2.5.5 基于 HTTP 的应答消息 .....	80
2.5.6 HTTP 消息 .....	81

## 第 3 章 服务器/客户端的 Web 技术

3.1 标记语言的发展 .....	84
3.1.1 “超文本”和“标记”的意义 .....	84
3.1.2 从 SGML 到 HTML .....	85
3.1.3 从 HTML1.0 到 HTML4.01 .....	87
3.2 HTML 的基本构造 .....	89
3.2.1 HTML 文件的特征 .....	89
3.2.2 HTTP 上的标签的功能 .....	91
3.2.3 HTML 记述的例子 .....	93
3.2.4 基于框架的访问环境的改善 .....	94
3.2.5 通过表单收集用户信息 .....	96



3.2.6	通过可点图链接信息 .....	96
3.3	从 HTML 到 DHTML .....	99
3.3.1	在 DHTML 上的扩展功能 .....	99
3.3.2	DHTML 的功能概要 .....	100
3.3.3	DHTML 所支持的技术 .....	101
3.3.4	脚 本 .....	102
3.3.5	能在 Web 浏览器上运行的脚本 .....	105
3.3.6	脚本的记述方法 .....	106
3.3.7	通过脚本实现的控制 .....	107
3.3.8	样式表的便利性 .....	110
3.3.9	样式表的种类 .....	111
3.3.10	通过 DOM 的实现 .....	112
3.4	在 Web 服务器上运行的程序 .....	113
3.4.1	CGI 被作为必要的原因 .....	113
3.4.2	通过 CGI 实现的功能 .....	116
3.4.3	CGI 环境构筑的顺序 .....	118
3.4.4	CGI 制作语言的种类 .....	120
3.5	Cookie 的功能和用途 .....	122
3.5.1	Cookie 是什么 .....	122
3.5.2	基于 Cookie 的交换 .....	123
3.5.3	Cookie 的用途 .....	126

## 第 4 章 Web 使用的各种技术

4.1	Java 的运行环境和开发环境 .....	130
4.1.1	Web 上的 Java 的重要性 .....	130
4.1.2	Java 的开发目的和互联网 .....	131
4.1.3	Java 的特征 .....	133

4.1.4	Java 程序的制作和执行 .....	134
4.1.5	Java 的运行机制 .....	136
4.1.6	Java 环境的种类 .....	137
4.1.7	Java 程序的种类 .....	138
4.2	XML 的功能和文本构成 .....	141
4.2.1	B to B 的 HTML 的界限 .....	141
4.2.2	XML 的登场 .....	143
4.2.3	XML 中的标签功能 .....	144
4.2.4	XML 文件的构成 .....	145
4.2.5	XML 的特征 .....	147
4.2.6	XML 用样式页记述语言——XSL ...	148
4.2.7	解析器和 DOM 功能 .....	151
4.2.8	被对象化了的 XML 信息的优点 ...	152
4.3	服务器端环境 .....	155
4.3.1	通过 CGI 给 Web 服务器的负荷 .....	155
4.3.2	新的服务器端脚本环境 .....	156
4.3.3	ASP 的功能 .....	157
4.3.4	JSP 的功能 .....	159
4.3.5	servlet 的功能 .....	160
4.4	在 Web 构筑中的面向对象的必要性 .....	163
4.4.1	软件开发方法的发展 .....	163
4.4.2	所谓的面向对象 .....	165
4.4.3	属性和方法的概念 .....	166
4.4.4	通过封装隐藏属性 .....	167
4.4.5	类的层次和继承的概念 .....	170
4.5	实现分布式对象的技术 .....	172
4.5.1	分布式对象的概念 .....	173
4.5.2	分布式对象的优点 .....	173
4.5.3	基于 CORBA 的分布式对象 .....	174

## 第 5 章 Web 服务器和 Web 应用程序服务器的构筑

5.1	Web 服务器站点的构成和发展 .....	180
5.1.1	基本的 Web 服务器功能 .....	180
5.1.2	共享 CGI 程序的服务器构成 .....	181
5.1.3	服务器端环境的变化 .....	182
5.1.4	Web 应用程序服务器的登场 .....	183
5.2	构筑 Web 服务器所必须的不间断连接环境 ...	184
5.2.1	Web 服务器所必须的不间断 连接环境 .....	184
5.2.2	适合 Web 服务器的不间断连接服务 ...	185
5.2.3	面向个人用户的不间断连接环境 ...	188
5.3	Web 服务器的构筑环境 .....	190
5.3.1	服务器机器的选定——PC 服务器 ...	190
5.3.2	OS 的选定 .....	192
5.3.3	互联网站点里必需的服务器群 .....	194
5.3.4	互联网服务器群的理论构成 .....	195
5.3.5	互联网服务器群的物理构成 .....	196
5.4	Web 服务器软件的功能 .....	198
5.4.1	Web 服务器软件和其功能 .....	198
5.4.2	Apache 的特征 .....	199
5.4.3	IIS 的特征 .....	201
5.4.4	个人 OS 用的 Web 服务器软件 .....	204
5.5	Web 应用程序服务器的作用 .....	205
5.5.1	围绕 Web 应用程序服务器的构成 ...	206
5.5.2	所谓的三层 Web 客户端/服务器 系统 .....	209

5.5.3	负荷分散装置的功能 .....	210
5.5.4	Windows 环境中的 Web 应用程序 服务器系统的构筑例子 .....	212
5.5.5	.NET 在 Web 上的发展 .....	215
5.5.6	Java 环境中的 Web 应用程序 服务器系统的构筑 .....	217
5.6	支持 Web-EDI 的 XML .....	221
5.6.1	EDI 的正式普及 .....	222
5.6.2	从 EDI 到 Web-EDI .....	223
5.6.3	Web-EDI 上有效的 XML 工具 .....	224
5.6.4	通过 XML 来实现 Web-EDI .....	226

## 第 6 章 Web 服务器的安全对策

6.1	置身于危险中的 Web 服务器 .....	230
6.1.1	危险用户的存在 .....	230
6.1.2	互联网安全概要 .....	232
6.1.3	对 Web 服务器的扫描 .....	233
6.1.4	非法侵入和超级管理员权限的取得 .....	234
6.1.5	密码攻击法 .....	237
6.1.6	欺骗法 .....	238
6.1.7	拒绝服务攻击法 .....	239
6.2	Web 服务器的安全对策 .....	242
6.2.1	安全网络的构成 .....	243
6.2.2	防火墙的功能 .....	244
6.2.3	防火墙的代理功能 .....	246
6.2.4	通过日志工具管理安全 .....	248
6.2.5	VPN 的结构 .....	250

---

6.3	验证和加密技术 .....	252
6.3.1	网络交易中的问题和对策 .....	252
6.3.2	通过加密防止窃听 .....	253
6.3.3	通过电子签名防止删改 .....	254
6.3.4	电子证书和其功能 .....	255
6.3.5	具有超过 SSL 安全性的 SET .....	258
6.4	企业规模的安全对策 .....	261
6.4.1	全公司的意识提升和安全 .....	261
6.4.2	系统管理者的组织构成和配置 .....	263
6.4.3	OS、服务程序的运用管理 .....	264
6.4.4	企业级的病毒对策 .....	265

# 第1章

## 互联网的机构和Web的机能

据统计，日本互联网用户的数量，算上用移动电话上网的人，已经有2000万。现在，不只是商业活动，就连个人信息检索或买东西，作为通信手段，平常都在利用互联网。

实际上，互联网已有30年历史。但是，这期间，并不是谁都可以利用网络。当初，只是在一部分研究者间使用。随着管理互联网的机构渐渐转移至民间，1989年，Web登场，互联网终于迎来了大转机。

本章将介绍有关互联网和Web的基础知识。所谓的互联网及Web是什么样的东西呢？首先在了解这些知识的基础上，再进行技术性的话题吧！

## 1.1 Web 的发展和检索功能

不用说,本书主要内容当然是 Web 技术。要想了解详细内容,有必要对 Web 的发展历史,Web 服务和 Web 检索的获取等基本情况充分了解。

Web 的真实面目到底是什么样的呢?本节将为您揭开 Web 的神秘面纱。

### 1.1.1 何谓 Web

Web 一词最近被频繁使用,其在字典中的意思为蜘蛛网。关于原意,有的字典解释为棉纺布料。无论是蜘蛛网,还是棉纺布料,都是由多条线相互交叉构成的(参见图 1.1)。

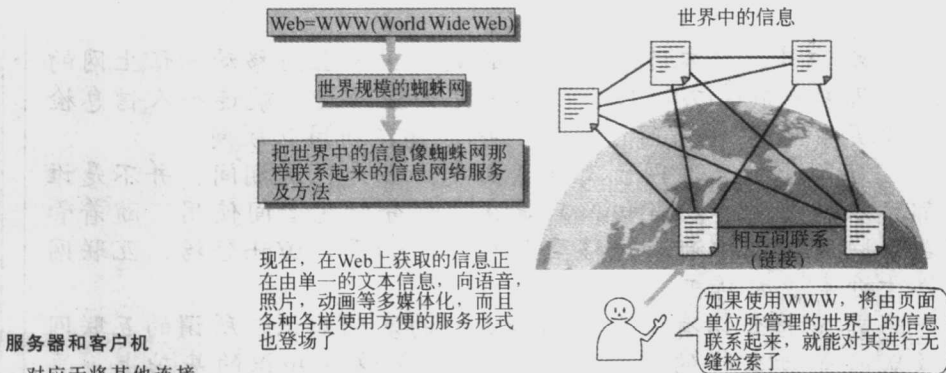


图 1.1 Web 的意义

实际上,Web 是 WWW(World Wide Web)的简称,即 Web 等同于 WWW,它就好像是连接世界的蜘蛛网。总之,可以把连接的网络看成是布满世界的蜘蛛网。

WWW 是互联网提供信息的一个手段。通过这种手段,能够实现以页面为单位管理庞大的信息及其之间的联系,并对其进行无缝检索。

#### 服务器和客户机

对应于将其其他连接在网络上的计算机的请求,进行特化处理的计算机称为服务器(server)。接受互联网上由服务器提供的各种各样服务的计算机称为客户机(client)。互联网上的用户实际操作的计算机终端通常被称为客户端。

所谓 Web 就是世界中的信息关联成蜘蛛网状的信息网络组合,也可以说是在这个网络上所提供的服务。

### 1.1.2 Web 的起源

和 Web 相同意思的词是 WWW,它是 1989 年,汇集于 CERN(Conseil Europeen Pour La Recherche Nucleaire, 欧洲原子核研究机构)的各国科学家间(计算机间)以获取、交流文本信息为目的开发的。

汇集于 CERN 的各国科学家携带各自的计算机,这些计算机之间通过网络相互连接在一起,某些国家的科学家为了得到其他科学家提供的信息,必须要告诉访问装有这些信息的计算机的方法,以及访问后的操作方法等。

但是,如果对于每个计算机都必须牢记其步骤,就不能专心于原来的研究。当时,虽然有可以公开的信息,但是没有将其共享的方法。

因此,一位叫 Tim. Berners-Lee 的科学家,想出了在连接到网络的计算机间,能简单的共同使用文本信息的组合,这就是 WWW 的起源。

在 WWW 方面,将确定了共同使用的信息用标记语言作好,保存于各自的计算机中。当将这些格式的信息作为文本表示出来的程序或其他计算机有了要求时,能够给计算机发送格式化信息,并且可以在计算机之间进行交流了。

还有,这里所说的标记语言,必须要对于其他的文本有关联。具体内容就是,只要用鼠标点击显示和文本有关联的部分,相对应的文本立刻就能显示,利用 WWW 就可以将分散的多个信息作为一个大型信息访问(参见图 1.2)。

像这样,WWW 一方面具有非常优异的性能,一方面构造比较简单,因此在研究者间备受欢迎。但是,在当时存在一个问题,即开发 WWW 的科学家使用的是美国 NeXT 公司称为 STEP 的 OS,因此为了利用 WWW 共同使用信息,前提条件必须使用这个操作系统的计算机,但 STEP 在当时不太普及。

#### OS: Operating System

是基本的软件。是为了使计算机高效运行的程序群,进行应用程序软件的执行管理和周边机器的控制。Windows95/98/Me/XP、WindowsNT/2000、Linux、FreeBSD、MacOSX 等都属于 OS。



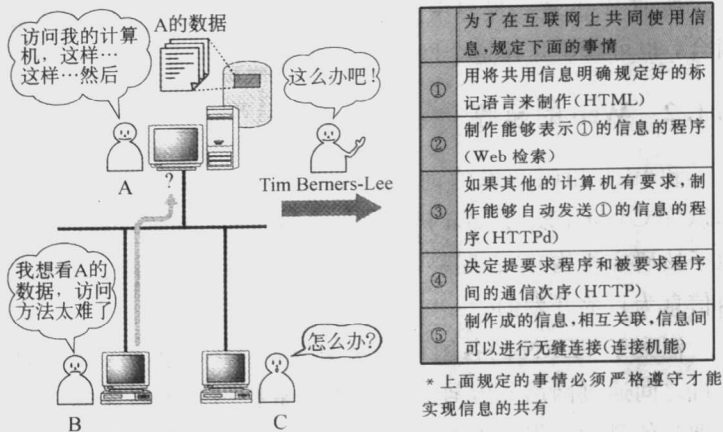


图 1.2 Web 的起源

因此,在众多研究人员认识到 WWW 的优异之处的同时,也有难以利用的一面。在这种背景下,研究了 WWW 的学者决定将 WWW 的规范通过互联网公开。这一决定为互联网的快速普及做出了巨大贡献。

公开的 WWW 规范瞬间扩展至全世界,开始被移植到各种各样的操作系统。在 CERN 以信息共享为目的而名声大振的 WWW,汇集了全世界的信息发展起来。

在先前的说明中,作为实现 WWW 所必需的部分,明确的标记语言、将格式化信息作为文件显示的程序、根据要求发送格式信息的程序,以及计算机间的交换顺序都有其固定的名称。

标记语言是 HTML(HyperText Markup Language),将格式信息作为文件显示的程序是 Web 浏览器(Web browser),根据要求提供格式信息的程序是 Web 服务软件(Web Server Software)或者是 HTTPd(HyperText Transfer Protocol daemon),计算机之间交换顺序称为 HTTP(HyperText Transfer Protocol)。

关于这些,以后会详细说明(参见图 1.3)。