

Web 技术基础

裴有福 编著

万水Web技术精品丛书



 中国水利水电出版社
www.waterpub.com.cn

万水 Web 技术精品丛书

主编：裴有福

Web 技术基础

裴有福 编著



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书以较小的篇幅覆盖了 Web 技术的主要方面。其内容主要包括 Web 技术的概括性介绍、Web 浏览器和服务器的 Web 页面的制作技术,而且也介绍了 Java 和 JavaScript 在 Web 页面中的应用,甚至还包括 CGI 的基础内容。

图书在版编目(CIP)数据

Web 技术基础/裴有福 编著.-北京:中国水利水电出版社,1998
(万水 Web 技术精品丛书)
ISBN 7-80124-742-6

I. W… II. 裴… III. 万维网-基本知识, IV. TP393.4

中国版本图书馆 CIP 数据核字(98)第 18526 号

书 名	Web 技术基础
作 者	裴有福 编著
出版、发行	中国水利水电出版社(北京市三里河路 6 号 100044) 网址: www.waterpub.com.cn E-mail: sale@waterpub.com.cn 电话: (010)63202266(总机)、68331835(发行部)
经 售	全国各地新华书店
排 版	北京门槛创作室 WORD 照排部
印 刷	北京市牛山世兴印刷厂
规 格	787×1092 毫米 16 开本 11.75 印张 282 千字
版 次	1998 年 8 月第一版 1998 年 8 月北京第一次印刷
印 数	0001—5000 册
定 价	19.00 元

凡购买我社图书,如有缺页、倒页、脱页者,本社发行部负责调换

版权所有·翻版必究

前 言

不管您愿意不愿意、承认不承认, Internet 正在迅速改变着我们的生活习惯和生活质量, 并且必将越来越深刻地影响人类社会的方方面面。作为 Internet 的重要内容, World Wide Web (简称 Web, 缩写为 WWW, 中译为万维网) 已被公认为是 20 世纪的重大科技突破。它以一种简单方便、易于操作的方式链接着世界范围内各种类型的信息和各种层次的人们, 并将为人们构筑更加美好的生活和工作。在这种形势下, 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。

《万水 Web 技术精品丛书》正是在这种形势下应运而生的。本套丛书的作者长期关注和研究 Web 技术, 目睹着它的迅猛发展, 也体会到它的意义重大。但同时也注意到许多 Web 使用者, 特别是一些初学者, 在广袤的 Web 空间里不知所措, 找不到重点, 分不清主次, 常常事倍功半。有很多读者, 没有条件或不能经常连入 Internet, 不能及时学习 Web 技术, 反被市面上的一些图书弄得懵懵懂懂。因此, 这套丛书力图系统全面地介绍 Web 技术的各个方面, 努力使读者能够在短期内迅速掌握 Web 技术。

Web 技术可谓博大精深。为了帮助初学者迅速进入 Web 世界, 《Web 技术基础》试图以较小的篇幅覆盖 Web 技术的主要方面, 内容不仅包括 Web 技术的概括性介绍、Web 浏览器和服务器的制作技术, 而且也介绍了 Java 和 JavaScript 在 Web 页面中的应用, 甚至还包括 CGI 的基础内容。本书力图使初学者能够“一书在手, Web 全有”。较高水平的读者请选择本书的姊妹书《Web 技术大全》, 这是目前不可多得的 Web 技术全书式著述, 内容涵盖 Web 技术的方方面面。

衷心希望读者可以从本书中获得收益, 并在学习和掌握 Web 技术的过程中享受到无穷的乐趣! 本书作者努力做得更好, 但由于水平有限、时间紧迫, 不足与错误在所难免。希望读者和同行专家不吝指教, 以便在未来的版本中予以修改, 使本书真正成为 Web 技术的众望所归之作。



1998.5

目 录

前 言	
第一章 绪 论	1
1.1 Web 技术概述	1
1.2 Web 相关的协议与语言	4
1.3 Web 浏览器和服务端	9
第二章 Web 页面基础	13
2.1 HTML 基础	13
2.2 基本 HTML 元素	18
2.2.1 页面总体结构	18
2.2.2 文档字体设置	19
2.2.3 页面版式设置	21
2.2.4 页面修饰	30
2.2.5 超链接	33
2.3 页面高级技术	35
2.3.1 表格 (Table)	35
2.3.2 表单 (Form)	48
2.3.3 窗框 (Frame)	55
2.4 页面多媒体技术	60
2.4.1 图像 (Image)	60
2.4.2 音频和视频	65
2.4.3 滚动文本 (Marquee)	67
第三章 Java 在 Web 页面中的应用	70
3.1 在 Web 页面中嵌入 Java 小应用	70
3.2 Java 语言基础	72
3.2.1 程序结构	72
3.2.2 基本规定	72
3.2.3 类型	75
3.2.4 类	77
3.2.5 界面	89
3.2.6 包	90
3.2.7 表达式	91
3.2.8 语句	94
3.3 Java 小程序实例剖析	100
3.3.1 Java 动画	100
3.3.2 Java 汉字处理	107

3.3.3	下拉菜单	111
3.3.4	在 Java 小程序中发送电子邮件	113
3.3.5	在网页上发布统计曲线	116
第四章	JavaScript 在 Web 页面中的应用	124
4.1	在 Web 页面中嵌入脚本程序	124
4.2	JavaScript 实例剖析	125
4.2.1	自动更新文档修改日期	125
4.2.2	链接、状态栏与弹出窗口	125
4.2.3	日期和时间函数	126
4.2.4	随机数的产生	127
4.2.5	自行制作窗口	128
4.2.6	在窗框 (Frame) 中的应用	129
4.2.7	浏览器历史记录跳转	130
4.2.8	单击一次鼠标读入两个文件	130
4.2.9	本地校验用户输入内容	131
4.2.10	浏览器和 JavaScript 版本的测试	132
4.2.11	浏览器状态栏文字的控制	134
第五章	CGI 基础	139
5.1	CGI 的基本概念	139
5.2	CGI 编程语言	140
5.3	程序间的数据通信方式	142
5.4	CGI 的主要用途	144
5.5	CGI 程序设计基础	146
5.5.1	Hello, Web! 程序	147
5.5.2	CGI 输出	148
5.5.3	安装和运行 CGI 程序	148
5.5.4	CGI 输入	150
5.5.5	一个简单的 CGI 程序	153
附录 A	HTML 元素和属性索引	155
附录 B	浏览器支持 HTML 的比较	163
附录 C	HTML 字符集	171
C.1	ISO Latin-1 字符集	171
C.2	基本英文字符集	173
附录 D	Netscape 定义的 HTML 颜色	180

第一章 绪 论

作为未来信息社会基础结构的原型，Internet 正在影响人类社会的各个角落，其中的 Web 技术更是前所未有地超过任何现有的技术，迅速成为 20 世纪最伟大的科技成果之一。本章作为全书的绪论，对 Web 技术作概述性介绍，主要内容包括 Web 技术概述、Web 的协议与语言以及 Web 浏览器与服务器等。

1.1 Web 技术概述

今天，只要有一台与 Internet 相连的计算机，不管它是 PC、Macintosh 还是 Unix 工作站，也不管它是通过什么方式连入 Internet 的，任何人都可以访问处于 Internet 上任何位置的 Web 站点（图 1-1）。Web 正以前所未有的魅力实现人类“世界大同”的梦想，因为它通过技术方式实现了不受限制的全球访问。

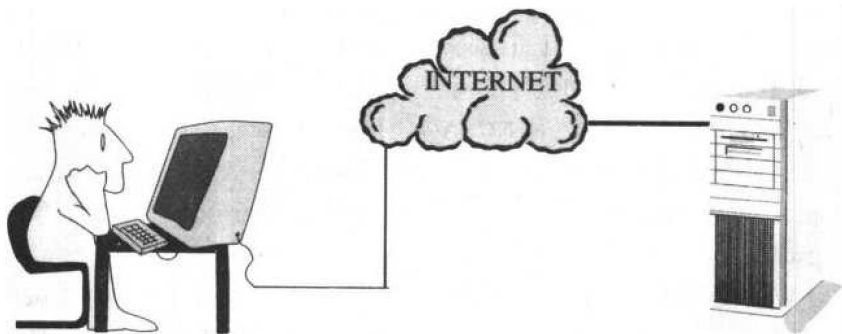


图 1-1 任何人用任何机器都可以访问 Web

那么，究竟什么是 Web 呢？

Web，全称为 World Wide Web，缩写为 WWW，中译为“万维网”。但什么是 Web，目前尚无公认的准确定义。

简单地说，Web 是一种体系结构，通过它可以访问遍布于 Internet 主机上的链接文档。这一说法可以分解为以下几层意思：

Web 是 Internet 提供的一种服务。尽管这几年 Web 的迅猛发展使得有人甚至误认为 Web 就是 Internet，但事实上，Web 是基于 Internet、采用 Internet 协议的一种体系结构，因而它可以访问 Internet 的每一个角落。

Web 是存储在全世界 Internet 计算机中、数量巨大的文档的集合。或者可以通俗地说，Web 是世界上最大的电子信息仓库。

Web 上海量的信息是由彼此关联的文档组成的，这些文档称为主页（Home Page）或页面（Page），它是一种超文本（Hypertext）信息，而使其连接在一起的是超链接（Hyperlink）。由于超文本的特性，用户可以看到文本、图形、视频、音频等多媒体信息，这些媒体称为超媒体（Hypermedia）。

Web 的内容保存在 Web 站点（Web 服务器）中，用户可通过浏览器（Browser）访问

Web 站点。因此, Web 是一种基于客户机/服务器 (Client/Server, 简称 C/S) 的体系结构, 由于它的广泛使用, 现在这种体系结构可以称作浏览器/服务器 (Browser/Server) 结构。也就是说, Web 实际上是一种全球性通信系统, 该系统通过 Internet 使计算机相互传送基于超媒体的数据信息。

Web 以一些简单的方式 (例如点击鼠标即可去往超链接、浏览器通常是标准的应用程序、页面基本上与平台无关等) 联接全球范围的超媒体信息。因此, 它易于使用和普及, 基于 Web 开发各种应用易于实现跨平台、开发成本也较低, 而且基于 Web 的应用常常几乎不需要培训用户。



背景与历史

Web 最早是由欧洲核能研究中心 (CERN, the European center of nuclear research) 提出和开发成功的。CERN 在多个国家建有加速器, 科研人员也分散在各国, 而研究项目经常要延续较长时间。为了便于联系和沟通, CERN 的物理学家 Tim Berners-Lee 于 1989 年 3 月提出了链接文档的构想, 并于 18 个月后实现了第一个基于文本的 Web 原型。1991 年 12 月, Web 在超文本会议 (Hypertext '91) 上进行了公开演示。

经过几年的发展, NCSA (National Supercomputing Application, 美国国家超级计算应用中心) 于 1993 年开发成功了轰动世界的第一个图形界面的浏览器 Mosaic。此后, Mosaic 的主要作者之一 Marc Andreessen 离开 NCSA 并创办了 Netscape 通信公司, 以开发 Web 客户程序、服务器程序和其他 Web 软件为主要目标。Netscape 获得了巨大的成功, 其著名浏览器软件 Netscape Navigator 推出后, 它便成为软件霸主 Microsoft 的主要竞争对手之一, Web 也随着 Netscape 的成功获得了超速发展。

本质上讲, Web 是基于客户机/服务器的一种体系结构 (图 1-2)。PC 等微型计算机的性能/价格比比工作站、小型机和大型机高得多, 因此, 人们将数据等大量信息存储在共享的服务器上, 用价格低廉的个人计算机 (客户机) 与之相连, 以便享用服务器的高性能。在这种客户机/服务器模型中, 典型的通信过程为: 客户机向服务器发送请求, 要求执行某项任务, 而服务器则执行此项任务, 并向客户机返回响应。很容易理解, 在网络上通常客户机的数量比服务器的数量要多得多。

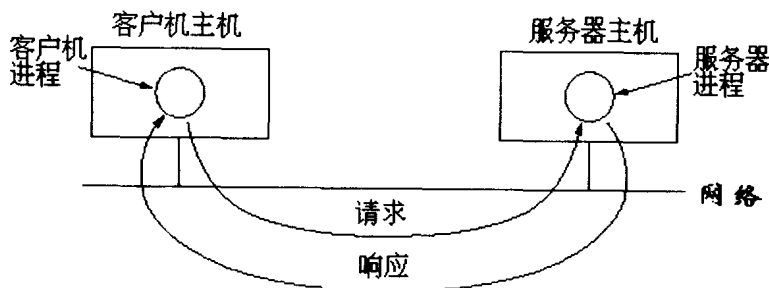


图 1-2 客户机/服务器模型

在客户机/服务器体系结构中, 通常很容易将客户机和服务器理解为两端的计算机。但事实上, “客户机”和“服务器”概念上更多的是指软件, 是指两台机器上相应的应用程序, 或者说是图 1-2 中的“客户机进程”和“服务器进程”。

对于 Web 系统, 客户机上所运行的浏览器程序基本上是标准化的, 所以人们建立客户机/服务器系统的主要任务就落到了服务器端。Web 的体系结构因而可以称作浏览器/服务器结构。

在 Web 环境中, Web 服务器向浏览器提供服务的过程大致可以归纳为以下几个步骤 (图 1-3)。

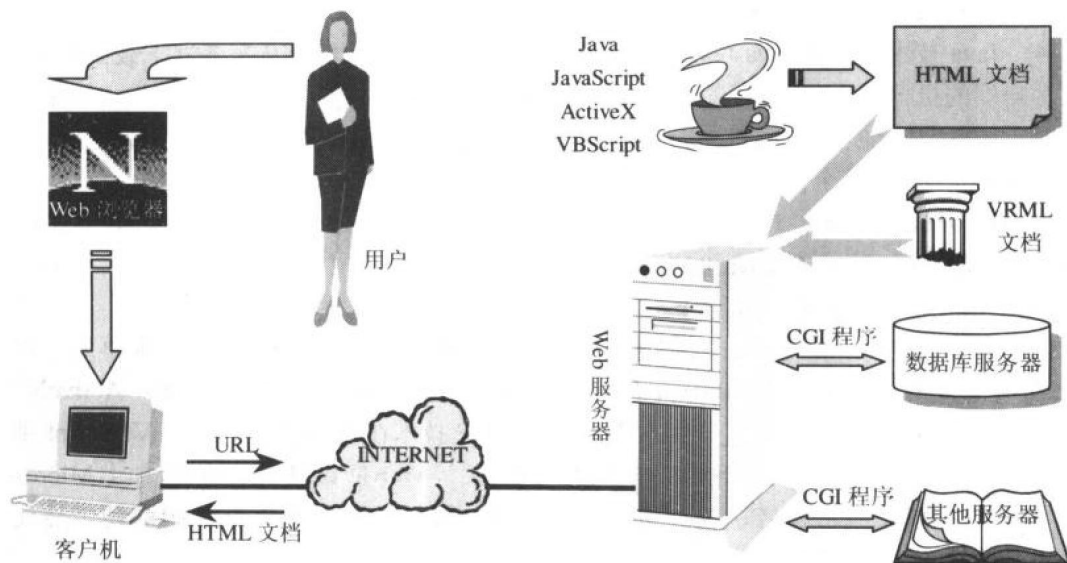


图 1-3 浏览器/服务器的工作过程及 Web 的基本技术

(1) 用户打开计算机 (客户机), 启动浏览器程序 (Netscape Navigator、Microsoft Internet Explore、Sun HotJava、NCSA Mosaic 等), 并在浏览器中指定一个 URL (Uniform Resource Locator, 统一资源定位符), 浏览器便向该 URL 所指向的 Web 服务器发出请求。

(2) Web 服务器 (也称为 HTTP 服务器) 接到浏览器的请求后, 把 URL 转换成页面所在服务器上的文件路径名。

(3) 如果 URL 指向的是普通的 HTML 文档, Web 服务器直接将它送给浏览器。HTML 文档中可能包含有 Java、JavaScript、ActiveX、VBScript 等编写的小应用程序 (applet), 服务器也将其随 HTML 一道传送到浏览器, 在浏览器所在的机器上执行。

(4) 如果 HTML 文档中嵌有 CGI (Common Gateway Interface, 公共网关界面) 程序, Web 服务器就运行 CGI 程序, 并将结果传送到浏览器。Web 服务器运行 CGI 程序时还可能调用数据库服务器和其他服务器。

(5) URL 也可以指向 VRML 文档。只要浏览器中配置有 VRML 插件, 或者客户机主机上已安装 VRML 浏览器, 就可以接收 Web 服务器发送的 VRML 文档。

上述过程涉及到 Web 的许多技术。由于 Web 正处在日新月异的高速发展之中, 它所覆盖的技术领域和层次深度也在不断改变, 所以本书只能讨论主要技术的基础内容。作者认为, 现阶段 Web 的基本技术包括以下几个方面:

(1) 服务器技术。主要指 Web 服务器建构有关的基本技术, 包括服务器策略与结构的设计、服务器软硬件的选择及其他有关服务器建构的问题。

(2) HTML 技术。HTML 是 Hypertext Markup Language (超文本标记语言) 的缩写,

它是构成 Web 页面 (Page) 的主要工具。HTML 是 Web 的基础, 它的超文本思想简单地说就是纯文本加标记 (Tag), 所以它的重要内容就是要熟悉和掌握这些标记。

(3) CGI 技术。CGI (Common Gateway Interface, 公共网关接口) 是 Web 动态页面的基础, 也是 Web 与其他应用交流的基础。例如, 为了向用户提供查询服务, Web 页面就需要通过 CGI 与数据库服务器进行沟通。虽然 CGI 的一些功能已经可以由 Java、JavaScript 等代替, 但 CGI 仍然是不可缺少的。

(4) Java 技术。利用 Java 可以创建小应用程序 (applet), 使之从服务器下载到浏览器, 并在客户端的机器上运行。Java applet 可提供动画、音频和音乐等多媒体服务, 并能产生原本只有 CGI 才能实现的功能。因此, Web 不仅已经成为 Web 技术的重要组成部分, 它对整个 Internet 甚至整个信息业界的影响也是前所未有的、震撼性的。

(5) JavaScript 技术。JavaScript 是由 Netscape 开发的一种解释语言, 它可以直接插入到 HTML 文档中。它具有 Java 的许多特性, 但比 Java 更为简单有效, 而且不需要编译, 因而也是 Web 技术的重要组成部分。

(6) ActiveX 技术和 VBScript 技术。软件巨人 Microsoft 曾在 Internet 方面落后一步, 但它利用其操作系统巨大的市场占有率迅速提出了自己的整套解决方案。ActiveX 是 Microsoft Internet 战略的中心环节, 它本质上是 OLE 技术在 Web 上的应用。VBScript 则是 ActiveX 的脚本功能。由于 ActiveX 很难具有平台无关等 Web 应该具有的特性, 加上 Microsoft 在技术上的垄断作风, 作者认为 Microsoft 在这一战场上很难获得真正的成功。然而, Microsoft 的力量是不可忽视的。

(7) VRML 技术。VRML 是 Virtual Reality Modeling Language (虚拟现实造型语言) 的缩写。Web 已经由静态步入动态, 并正在逐渐由二维走向三维, 将用户带入五彩缤纷的拟实世界。VRML 是目前创建三维对象最重要的工具。它也是一种基于文本的语言, 并可运行于任何平台。

1.2 Web 相关的协议与语言

Web 是信息发现与超文本技术的结合, 它的信息访问是双向的: 一方面是 Web 浏览, 用户可通过浏览器在 Internet 上寻找所需的信息; 另一方面是 Web 展示, 用户可在自己的 Web 服务器上建立和存放按一定结构组织的超文本文件, 供网络其他客户浏览, 传播知识、交流信息和进行广告宣传。

超文本 (HyperText) 是把一些信息根据需要链接起来的信息管理技术。人们可以通过超文本中的链指针打开另外一个相关的文本, 用鼠标点击文本中的高亮度或带下划线的文字, 即可打开相链的文本, 获得所需的信息。目前已有许多软件提供的联机帮助信息或联机手册都是基于超文本的。

超文本由若干互连的文本块组成, 这些信息块可以是若干屏、窗口、文件或更小的块信息, 这样的信息单元称为节点 (node), 如图 1-4 中的 A、B、C、D。不管节点有多大, 每个节点都有若干指针指向其他节点或从其他节点指向该节点的指针, 这些指针称为链接 (link)。超文本链接亦称为锚点 (anchor), 或称锚标。图 1-4 为三个文本、四个节点、五条链接组成的超文本结构。

文本 1

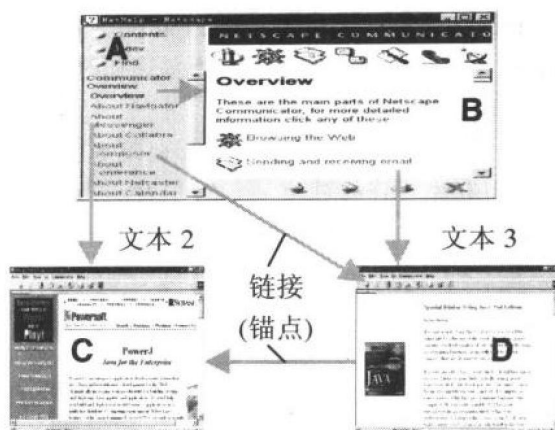


图 1-4 超文本及相关概念

构成超文本的信息网络的每个节点都包含一个特定主题的信息，节点的大小视主题而定。按照节点的功能主要可分为表现型节点、组织型节点和推理型节点三类。

表现型节点是一般性超文本节点，主要有：由文本段组成的文本节点；图形图像节点；由声音或合成语音构成的声音节点；用于执行一个过程的按钮节点。

组织型节点是组织节点的节点，主要有：目录节点，它以条目形式给出信息的索引指针，指向索引节点；索引节点，由索引项组成，索引项用指针指向目的节点或相关索引项。

推理型节点只有智能型的超文本中才有，它主要包括对象节点和规则节点。

超文本的“超”体现在它不仅能包含文本，而且可以包含图像、音频、视频等多媒体信息，即将文本的概念扩充为超文本。因此，超文本也称超媒体（Hypermedia）。

超文本的书写格式有专门的规定，这就是超文本标记语言 HTML（HyperText Markup Language）。HTML 的版本变化非常快，已由 1.0 发展到当前通用的 3.2，最近还发布了最新的 4.0，因而尚未成为国际标准。

在 Internet 上传输超文本文件必须遵循超文本传输协议 HTTP（HyperText Transfer Protocol）。超文本传输协议是 Web 的基本协议。

HTML 将在后文详细讨论，以下简单介绍 HTTP 协议。

超文本传输协议 HTTP 是专门为 Web 设计的一种网络协议，Web 浏览器和服务端用 HTTP 来传输 Web 文档。

HTTP 正在不断改进中，Web 上正在使用一些版本，又有一些新的版本处于开发之中。目前主要使用 HTTP 1.0 和 1.1。

在 TCP/IP 体系结构中，HTTP 属于应用层协议，位于 TCP/IP 协议的顶层。因此，它在设计和使用中要以 TCP/IP 协议族中的其他协议为基础。例如，它要通过 DNS 进行域名与 IP 地址的转换，要建立 TCP 连接才能进行文档传输。

显然，HTTP 也是客户/服务器结构。这里，客户是浏览器（Browser），服务器是 Web 服务器。浏览 Web 时，浏览器通过 HTTP 协议与 Web 服务器交换信息。每当在 Web 上从一个资源转到另一个资源时，浏览器用 HTTP 访问 Web 服务器，其中就包括想要获取的资源信息。

浏览器和服务器通过 HTTP 交换 Web 文档时, 实际可以交换不同的文档类型。这些文档类型的格式由多用途 Internet 邮件扩展 MIME (Multipurpose Internet Mail Extensions) 定义。它早先是一种 Internet 邮件标准, 但后来逐渐成为描述通过 Internet 传输多媒体数据的技术标准。

每一个 Web 主机都有一个服务器进程来监听 TCP 端口 80, 以便同前来建立连接的客户 (通常是浏览器程序) 取得联系。连接建立后, 客户发送一个请求 (简单或完全), 服务器返回一个响应, 然后就释放连接。图 1-5 说明了 HTTP 的工作过程。

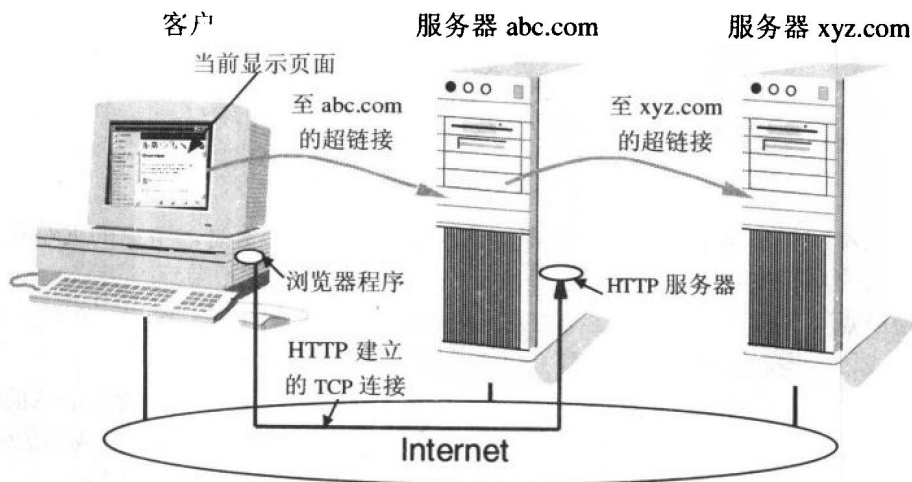


图 1-5 HTTP 的工作过程

为说明简单起见, 假设用户点击一个指向 <http://www.w3.org/hypertext/WWW/TheProject.html> 的超链接或图标, 想要获得该页面。那么, 在用户点击之后到页面显示之前, 浏览器与服务器之间的工作过程大致是:

- (1) 浏览器确定页面的 URL。
- (2) 浏览器向 DNS 查询 www.w3.org 的 IP 地址。
- (3) DNS 返回 18.23.0.23。
- (4) 浏览器向 18.23.0.23 的端口 80 请求建立 TCP 连接。
- (5) 浏览器发送 `GET /hypertext/WWW/TheProject.html` 命令。
- (6) www.w3.org 服务器发送文件 `TheProject.html`。
- (7) TCP 连接被释放。
- (8) 浏览器显示 `TheProject.html` 中的全部文本。
- (9) 浏览器获取和显示 `TheProject.html` 中的全部图像。

许多浏览器都在屏幕底端的状态栏中显示当前执行到哪一步。这样, 当 DNS 没有响应、服务器没有响应或页面传输时网络拥塞时, 用户都能够立即了解, 并据此决定是否继续浏览。

值得注意的是, 对于页面中的每一幅图像 (包括图标、图形、照片等), 浏览器都必须与相关服务器建立一次 TCP 连接, 获取图像然后释放连接。无疑, 当一个页面中包括大量图像而且图像都处于同一服务器时, 对每一幅图像都进行建立、使用和释放连接 “三部

曲”，显然是效率较低的。

除建立与释放连接外，HTTP 事务处理的主要内容是客户端的请求与服务器端的响应。请求分为简单请求（Simple request）和完全请求（Full request），响应也分成简单响应（Simple response）和完全响应（Full response）。

HTTP 是面向对象的协议，因此它将命令看作是方法（Method）。HTTP 方法描述指定资源上执行的动作，指定资源上的方法可以动态变化。HTTP 定义了 GET、HEAD、POST、PUT、DELETE、LINK 和 UNLINK 七种常用方法（表 1-1）。

表 1-1 HTTP 的常用请求方法

方 法	说 明
GET	请求读取一个 Web 页面
HEAD	请求读取一个 Web 页面的头标
PUT	请求存储一个 Web 页面
POST	附加到命名资源（如 Web 页面）中
DELETE	删除 Web 页面
LINK	链接两个已有资源
UNLINK	取消两个资源之间的已有链接

HTTP 请求由数行构成，其中第一行是请求行（Request Line），它包括方法、URL 和协议版本号（各项之间用空格隔开，最后加回车换行符）。请求行后包含若干行 HTTP 头标，依次是一般头标、请求头标和实体头标。一个空行后是实体，即服务器响应对象本身，其长度由实体头标决定，默认值为 0。在简单请求中，请求行的格式中没有版本号。

客户请求后，服务器对请求作出响应。同请求一样，HTTP 响应一般也包括数行：第一行为状态行（Status Line），由 HTTP 版本号、状态码和原因短语（Reason phrase）组成；接下来各行是头标信息（元信息），依次为一般头标、响应头标和实体头标；一个空行后是实体，即服务器响应对象本身，其长度由实体头标决定，默认值为 0。

完全响应至少包括一个状态行。也存在简单响应，但不建议使用。

HTTP 的状态码为三位数字的编码，其中第一位定义了响应码的分类：

- 1xx：提供信息。目前没有明确定义，留作将来使用；
- 2xx：成功。动作被接受、理解或访问；
- 3xx：更改。为完成请求所要求采取的动作；
- 4xx：客户端错。请求中有语法错或请求不可能被执行；
- 5xx：服务器错。服务器不能执行看来有效的请求。

表 1-2 列出了 HTTP 的状态码和相应的原因短语。其中，原因短语是协议推荐文本，服务器通常可用意义相同的文本代替这些短语（但不能影响 HTTP）。

表 1-2 HTTP 响应状态码和原因短语

响应状态码	原因短语
200	OK
201	POST command successful
202	Request accepted
203	GET or HEAD request fulfilled
204	Request fulfilled, but no content to return
300	Resource found at multiple locations
301	Resource moved permanently
302	Resource moved temporarily
304	Resource has not been modified
400	Bad request from client
401	Unauthorized request
402	Payment required for request
403	Resource access forbidden
404	Resource not found
405	Method not allowed for resource
500	Internal server error
501	Method not implemented
502	Bad gateway or server overloaded
503	Service unavailable / gateway timeout
504	Secondary gateway/server timeout

随着 Internet 的迅速发展和商品化, Web 的安全性正在日益受到关注。因此, 人们也在考虑和提出安全的 Web 协议。安全超文本传输协议 S-HTTP (Secure HyperText Transfer Protocol) 就是其中的一种解决方案。

S-HTTP 技术是在电子商贸活动中发展起来的。网络上的商业活动是安全性最敏感的地方, 为了推广电子商贸活动, 美国成立了商业 Internet 联盟, 并组建了 Commerce Net 组织, 推进 Internet 上的电子商贸活动。Internet 上开展商贸活动所必须解决的首要难题就是安全性问题。因此, 商业 Internet 联盟投资 Terisa 公司进行电子商贸安全的研究, S-HTTP 就是 Terisa 在 HTTP 基础上扩充并增加安全功能的结果, 它能够维护同 Web 服务器之间的通信安全。

S-HTTP 是 HTTP 的超集, 可以采用多种方式封装信息。它的封装包括加密、签名和基于消息确认码 MAC (Message Authentication Code) 的认证。同时, 一个消息可以被反复封装加密。S-HTTP 还定义了头标信息来进行密钥传输、认证传输和相似的管理功能。

SHTTP 可以支持多种加密协议, 还为程序员提供了灵活的编程环境。S-HTTP 并不依赖于特定的密钥证明系统, 它目前支持 RSA、带内和带外以及 Kerberos 密钥交换。密钥证明可以由信息包提供, 也可以采用其他方式得到。

Netscape 公司设计的安全套接口层 SSL (Secure Socket Layer) 则是作为传输通信协议 TCP/IP 上的安全协议实现的。SSL 和 S-HTTP 这两种协议互有长处, 有融合成为统一的协议产品的趋势。

SSL 协议的目标是提供两个应用间通信的保密和可靠性。SSL 由两层组成：低层是 SSL 记录层，用于封装不同的上层协议。其中一个被封装的协议即 SSL 握手协议，它可以让服务和客户在传输应用数据之前，协商加密算法和加密键。SSL 独立于应用协议，因而上层应用可能叠加在 SSL 协议上。

SSL 的安全协议主要集中在它的握手协议上。每一个希望提供服务的厂家需要向 CA (Certification Authority) 申请证明书 (Certificate)。在 CA 检验申请合法后，就在申请上加一个数字签名，作为证明书返回给申请厂家。在 WWW 导航器中，有 CA 用于签名的密钥，可以用来检验服务器是否合法。

从体系结构上看，SSL 位于 TCP/IP 之上、Internet 应用层协议之下。因此，它 S-HTTP 并不排斥，所以可以在服务器和客户端同时实现支持。但最终两者必将融合一个统一的协议。

1.3 Web 浏览器和服务

Web 浏览器是一种访问 Web 服务器的客户端工具软件，使用它可以在非常友好的界面下方便地进入 Internet，获取信息。

Web 浏览器的最基本目的在于让用户在自己的计算机上检索、查询、采掘、获取 Web 上的各种资源。由于 Internet 正处在日新月异的飞速发展阶段，Web 每天都被使用它的人们赋予新的涵义，使得浏览器的功能也在不断的扩充和更新。目前市场上已有几十种浏览器，功能有强有弱，归纳起来，浏览器应具备以下几种基本功能：

(1) 检索查询功能：浏览器读入 HTML 文档，解释 HTML 所描述的图表、声音、动画、表格，以及进一步的链接信息，利用 HTTP 协议，可在任意 Web 服务器上畅游。

(2) 文件服务功能：能在下载文档时实时查阅该文档，并可利用 HTTP 去跟踪感兴趣的链接。当感到正在下载的文档不需要时，可以随时中止下载过程。可对正在查阅的文档随时保存、打印、前后浏览等等。

(3) 热表管理：浏览器应能够自动记住用户刚刚访问过的 Web 地址，称为“热表”。当用户想要回到刚才曾访问过的某一 Web 中，用户可以从热表中快速地切换。

(4) 主页制作：一些浏览器可以浏览 Web 页面，而且也带有 HTML 编辑工具，使用户能够制作自己的 Web 页面。

(5) 其他 Internet 服务：浏览器正在努力提供越来越多的 Internet 服务，如 FTP、Gopher、WAIS、Telnet、Usenet 及有 email 等。

浏览器的功能可以通过配置得以扩充。对于浏览器不能处理的超媒体信息，它检查配置文件，查看如何处理接收到的信息。通常，在配置文件中会给定外部查看器 (external viewer) 或帮助程序 (helper application) 的程序名，浏览器自动调用这些程序处理自己不能处理的信息。如果某类文件没有指定外部查看器或帮助程序，浏览器就询问用户对此文件的处理方法。通过这种方法还可以让浏览器调用其他 Internet 服务程序 (例如 TELNET)，从而扩展浏览器的功能。

目前市面上流传着数量非常多的浏览器，它们大多为免费或共享软件，可以从 Internet 上方便地获取 (如果是商业软件，则通常可以下载其测试版或试用版)。最著名的浏览器包括：

● NCSA Mosaic

早期的浏览器只能浏览 ASCII 文本文档。第一个基于图形且得到广泛使用的浏览器是 Mosaic, 该浏览器由美国国家超级计算应用中心 NCSA (National Center for Supercomputing Applications) 的软件开发部设计编写。Mosaic 是一种多用途的 Web 工具, 能够支持信息的定位、合并与恢复。

Mosaic 能够支持许多操作系统平台, 如 X-Window、MacOS 和 Windows 等。它也能够支持多种协议, 并允许用户用不同的方式显示文档。

Mosaic 不仅直接导致了后来极为著名的 Netscape 的产生, 而且也对其他浏览器产品产生了深远的影响。



背景与历史

在 Web 基本协议提出不久的 1993 年, Web 的发展引起了设在美国伊利诺伊州立大学的 NCSA 的 Marc Andreessen 和他的同事们的兴趣。他们迅速研制出一种软件, 能够解释 Web 中的 HTML 文档, 并将 HTML 文档包容的信息以统一的方式显示出来。这就是 Mosaic。这是第一个真正的 Web 浏览器, 最早运行在 Sun 工作站 X-Window 图形环境上。后来, NCSA 又推出了 Mosaic 的 Macintosh 及 Windows 版本。

● Netscape Navigator

Netscape Navigator (常简称为 Netscape) 几乎已经成为浏览器的同义词。Netscape 包括了 Internet 应用的所有工具, 除 Web 浏览和制作外, 还可用于电子邮件、网络新闻、网络视频会议等。它支持各种多媒体对象, 包括多窗页面 (frame)、插件 (plug-in)、JavaScript 和 Java 小应用程序。它还能管理嵌入了带有音频、视频信息的三维对象的 Web 文档。

在 Netscape 的新版本 (4.0 以后) 中, 还包括支持最新的推送 (Push) 技术的 Netcaster 组件, 以及信息与日程管理工具等。Netscape 正在向操作系统的方向进军。



背景与历史

1994 年 4 月, Mosaic 的主要设计者 Marc Andreessen 和 SGI 公司的主要创始人之一 James H. Clark 共同创办了 Netscape 通信公司, Marc Andreessen 任技术副总裁。在他的率领下, Netscape 通信公司在 Mosaic 的基础上进行了大量改进, 开发出了一种新型的浏览器——Netscape Navigator。Navigator 提供了一个与 Mosaic 相似但比它更实用的图形界面。它的各方面性能均有了很大的增强, 是第一个优先快速显示文本和图形的浏览器, 也是第一个在收到整页文档前就允许查看页首的浏览器。

● Microsoft Internet Explore

Microsoft 曾经对 Internet 不够重视。不过, 它起步虽晚, 但有雄厚资金作后盾, 很快就追了上来。Internet Explore 是 Microsoft Internet 战略的重要产品。

目前, Microsoft Internet Explore 已经成为功能强大的主要浏览器之一。它支持 Java、HTML 及其扩展、VRML、JavaScript 和 ActiveX 等。

● Sun HotJava

Sun Microsystem 公司为了展示他们发明的 Java 语言, 专门用 Java 开发了一种浏览器, 称为 HotJava。HotJava 不仅继承了在 Mosaic、Netscape 等浏览器中实现的 Web 浏览技术, 而且扩充了把静态文档变成动态可执行代码的动态特性。HotJava 显示了 Java 语言环境的

强大功能,同时还为在 Internet 上发布 Java 程序提供了一个理想的平台。

● WorldView

目前 Web 站点中显示的大部分页面是基于文本和二维图像的,但用虚拟现实造型语言 VRML (Virtual Reality Modeling Language) 定义三维对象的页面也越来越多。

以前许多浏览器都不支持 VRML,要浏览 VRML 对象必须借助专门的浏览器。WorldView 就是一个较为著名的 VRML 浏览器。

现在只要适当配置 Netscape 和 Microsoft Internet Explore 等浏览器,就可以直接浏览 VRML 对象。

其他的浏览器也很多,其中较为知名的包括: Spry 公司开发的 AIR Mosaic, Spyglass 公司的 Enhanced NCSA Mosaic, NaviSoft 公司的 GNN Works, BookLink Technologies 公司的 Internet Works, Quarterdeck 公司的 Quarterdeck Mosaic, IBM 内置于 OS/2 Warp 内的 Web Explorer 以及 Macintosh 上使用的 MacWeb 等。

为 Web 客户提供服务的是 Web 服务器。尽管服务器也指提供服务的主机,但 Web 服务器主要指提供服务的软件。

Web 服务器软件市场的变化速度非常快,只要有一种新产品出台,功能相似甚至功能更强的产品立即应声而出。各厂商也各有妙计,不断更新其产品版本。目前 Web 服务器软件市场已近饱和,常见的不下几十种,支持各种硬件和操作系统平台,而且除商业软件外,也存在大量自由(免费)、共享类型的服务器。

以下是目前市场上常见、用户常用的一些 Web 服务器软件:

(1) NCSA Server (httpd)。NCSA 可能是当前最流行的 Web 服务器软件。它不仅性能良好,易于操作,而且是免费的。它主要的缺点是缺乏对 Web 加密技术的支持。

(2) CERN Server (httpd)。它是最早的 Web 服务器软件,支持 Unix 和 VMS 平台。CERN 也是免费的,而且有一些很有趣的特性,不过有时它的处理方式比较奇怪。

(3) Netscape Communications Server。它是一个很有特色的、高性能的服务器软件,由 Netscape 公司推出。目前商业上使用相当广泛。

(4) Netscape Enterprise Server。它与 Netscape Communications Server 几乎是一样的,不过它支持 SSL,因而也更昂贵。

(5) Apache Server。它是从 NCSA 派生出来的,主要区别在于 Apache 增加了对 SSL 的支持。目前已能够支持多种平台,是免费服务器软件中较被看好的一种。

(6) WebSite。这是由 O'Reilly and Associates 推出的面向 Windows 平台 (NT 和 95/98) 的一个很流行的商业产品,功能强大,但操作不太直观。

(7) MacHTTP。这是由 BIAP Systems 推出的 MacOS 平台的产品。它使用很方便,但略显得有些低档。它是一个商业产品。

总的看来,目前仍然是 Unix 平台的服务器在 Web 上应用得最广,最流行的分别来自于 CERN (欧洲粒子物理实验室)、NCSA (美国国家超级计算应用中心) 和 Netscape 公司。它们都可以提供服务器侧可执行程序、可点地图和不同等级的访问控制。

Netscape 两种服务器的优势在于响应速度快、效率高和容易使用,特别是在负载很大的情况下,仍然能够具有很高的响应速度和效率。但 Netscape 的服务器是非公开的,售价很高。不过如果用作 Web 商业服务,并需要处理大量访问,它还是很值得的。