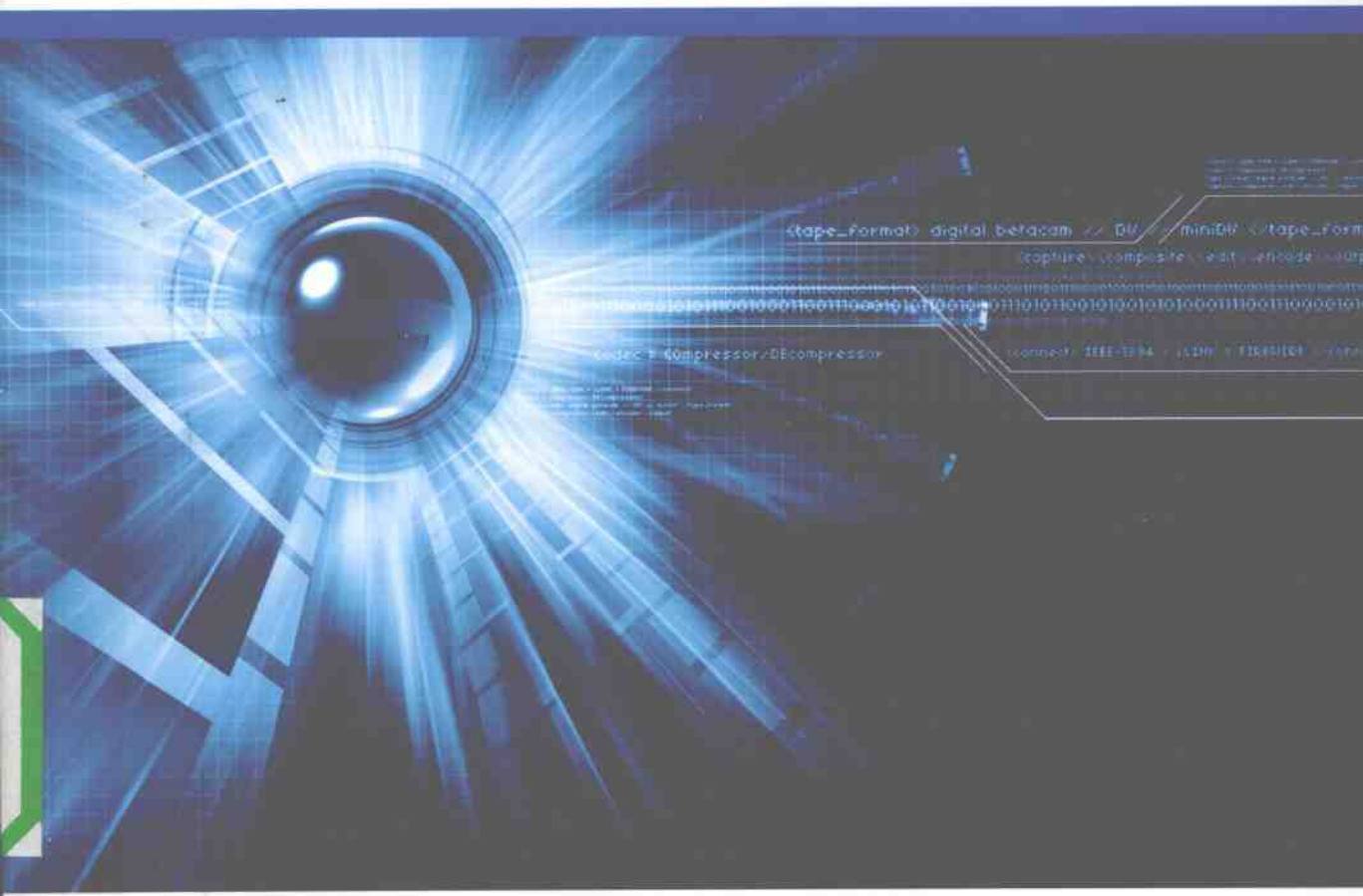
国家信息安全培训丛书

Web系统安全和渗透性测试基础

中国信息安全测评中心　编著



航空工业出版社

国家信息安全培训丛书

Web 系统安全和渗透性测试基础

中国信息安全测评中心　编著

航空工业出版社
北京

内 容 提 要

本书内容从浅入深，依次逐步展开。本书共分两部分：第一部分是Web系统安全基础，主要介绍了Web系统的基础和Web系统安全的基础；第二部分是Web系统渗透性测试基础，主要讲述了Web渗透测试的步骤、Web应用渗透性测试的框架以及如何撰写Web渗透测试报告。另外，书中附录部分介绍了一些常用的Web系统安全渗透性测试工具。

本书是中国信息安全测评中心注册信息安全专业人员（CISP）和注册信息安全员（CISM）的正式教材，可作为高等院校信息安全类专业学生教材，亦可作为信息安全培训教材和IT信息安全从业人员的参考书籍。

图书在版编目（CIP）数据

Web系统安全和渗透性测试基础/中国信息安全测评中心编著. —北京：航空工业出版社，2009.6
（国家信息安全培训丛书）
ISBN 978-7-80243-341-0

Ⅰ. W… Ⅱ. 中… Ⅲ. 计算机网络—安全技术 Ⅳ. TP393.08

中国版本图书馆CIP数据核字（2009）第089264号

Web系统安全和渗透性测试基础
Web Xitong Anquang he Shentouxing Ceshi Jichu

航空工业出版社出版发行
（北京市安定门外小关东里14号 100029）
发行部电话：010-64815615 010-64978486
北京地质印刷厂印刷 全国各地新华书店经售
2009年6月第1版 2009年6月第1次印刷
开本：787×1092 1/16 印张：12.25 字数：306千字
印数：1—5000 定价：35.00元

国家信息安全培训丛书

序

世界正经历一场伟大的信息革命，信息成为一种重要的战略资源。它改变着人们的生活方式和工作方式，形成新的社会形态。

随着我国社会信息化进程的不断发展，计算机网络及信息系统在政府机构、企事业单位及社会团体的工作中发挥着越来越重要的作用。然而，信息化水平的提高在带来巨大发展机遇的同时也带来了严峻的挑战。由于信息系统是一个复杂巨系统，它存在着脆弱性，信息安全问题不断暴露。信息安全关系到国家的经济安全、政治安全、军事安全和文化安全。信息安全已经成为维护国家安全和社会稳定的一个重要因素。

当前，社会对信息安全专业人员的需求逐年增加。发展信息安全技术与产业，关键是人才。培养信息安全领域的专业人才，已成为当务之急。高素质的信息安全人才队伍是保障国家重点基础网络和重要系统安全的基石，是制定信息安全发展战略规划与政策并建设国家信息安全保障体系的骨干力量，是发展我国信息安全产业的排头兵。

目前我国的信息安全教育工作仍相对滞后，信息安全人才十分匮乏，社会需求与人才供给间还存在着很大差距。如何培养信息安全的专业人才，是我国目前面临的重要问题。

《国家信息安全培训丛书》力图涵盖信息安全知识体系的方方面面，蕴含了信息安全保障体系的各个组成部分，是一套很好的信息安全专业人员培训丛书。相信这套丛书的出版，将有利于信息安全专业人员的培养。

何德全

2009 年 5 月

《Web系统安全和渗透性测试基础》编委会

目　录

第一部分　Web 系统安全基础

第二部分 Web 系统渗透性测试基础

第一部分

Web 系统安全基础

本部分包含以下章节：

第 1 章　Web 系统基础

第 2 章　Web 系统安全基础

第一部分

Web 系统安全基础

第 1 章　Web 系统基础

1.1　Web 概述

WWW（World Wide Web）简称 Web，中文译名为万维网，环球信息网等。Web 是由欧洲核物理研究中心（ERN）研制，其目的是为全球范围的科学家利用互联网进行通信、信息交流和信息查询提供便利。

Web 是建立在客户机／服务器模型之上的。Web 是以超文本标注语言（Hypertext Markup Language，HTML）与超文本传输协议（Hypertext Transfer Protocol，HTTP）为基础，能够提供面向互联网服务的、一致的用户界面信息浏览系统。其中 Web 服务器采用超文本链路来链接信息页，这些信息页既可放置在同一主机上，也可放置在不同地理位置的主机上；本链路由统一资源定位器（URL）维持，Web 客户端软件（即 Web 浏览器）负责信息显示与向服务器发送请求。

互联网采用超文本和超媒体的信息组织方式，将信息的链接扩展到整个互联网上。目前，用户利用 Web 不仅能访问到 Web 服务器的信息，而且可以访问到 FTP、Telnet 等网络服务。因此，它已经成为互联网上应用最广和最有前途的访问工具，并在商业范围内发挥着越来越重要的作用。

Web 客户程序在互联网上被称为 Web 浏览器（Browser），它是用来浏览互联网上 Web 主页的软件。目前，最流行的浏览器软件主要有 IE 和 Firefox 等。

Web 页面与 Web 浏览器如图 1-1 所示。

Web 浏览器提供界面友好的信息查询接口，用户只需提出查询要求（输入请求的网址），如何完成查询则由 Web 系统自动完成。因此 Web 为用户带来的是世界范围的超文本服务。用户只要操纵鼠标键盘，就可以通过互联网从全世界调来所需的文本、图像、声音等信息。Web 使得内部非常复杂的互联网使用起来异常简单。

Web 浏览器不仅为用户打开了寻找互联网上内容丰富、形式多样的主页信息资源的便捷途径，而且提供了 Usenet 新闻组、电子邮件与 FTP 协议等功能强大的通信手段。

万维网联盟（World Wide Web Consortium，W3C），又称 W3C 理事会。W3C 是 Web 领域最有影响力的组织，建立了各种 Web 标准。W3C 于 1994 年 10 月在麻省理工学院计算机科学实验室成立。建立者是万维网的发明者蒂姆·伯纳斯·李。

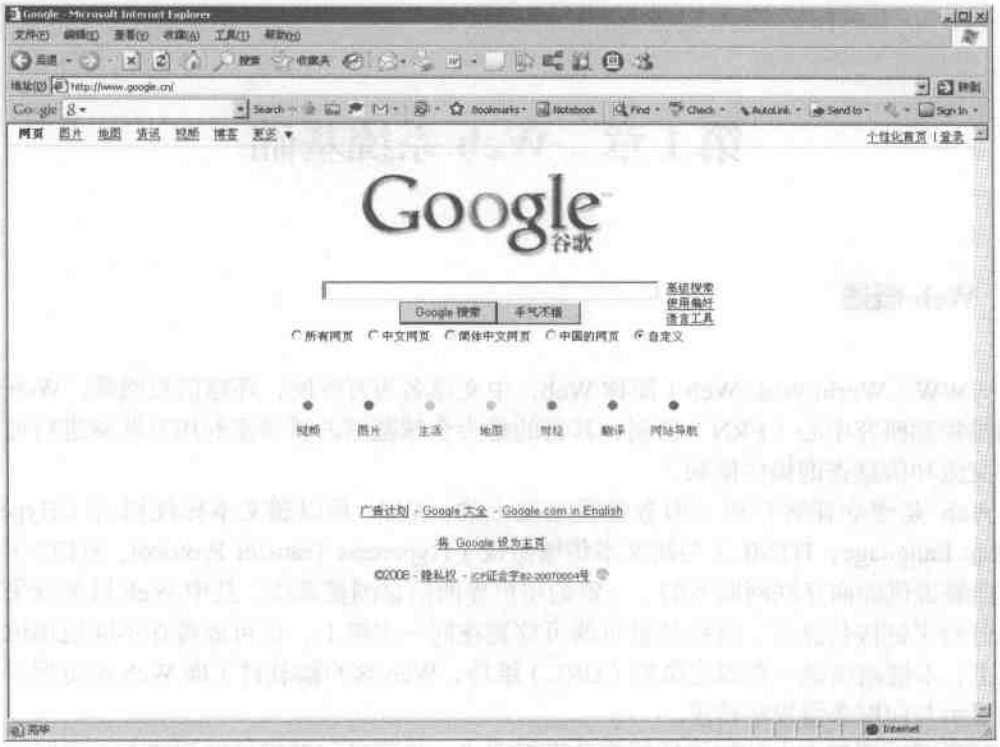


图 1-1　Web 页面与 Web 浏览器

1.1.1　URL

统一资源定位符 (Uniform Resource Locator, URL) 是用于完整地描述互联网上网页和其他资源地址的一种标志方法。

互联网上的每一个网页都具有一个唯一的名称标志，通常称之为 URL 地址，这种地址可以是本地磁盘，也可以是局域网上的某一台计算机，更多的是互联网上的站点。简单地说，URL 就是 Web 地址，俗称“网址”。

URL 方案集，包含如何访问互联网上的资源的明确指令。URL 是统一的，因为它们采用相同的基本语法，无论寻址哪种特定类型的资源（网页、新闻组）或描述通过哪种机制获取该资源。

对于互联网服务器上的目标文件，可以使用“统一资源定位符（URL）”地址（该地址以“http://”开始）。Web 服务器使用“超文本传输协议（HTTP）”，一种“幕后”的互联网信息传输协议。例如，http://www.itsec.gov.cn/ 是中国信息安全测评中心网站的 URL 地址。

URL 的一般格式为（带[]的部分为可选项）：protocol://hostname[:port]/path/[;parameters][?query]#fragment

例如，http://www.imailtone.com:80/WebApplication1/WebForm1.aspx?name=tom&age=20#resume

1.1.2　超文本和超媒体

①超文本（hypertext）

一种全局性的信息结构，它将文档中的不同部分通过关键字建立链接，使信息得以用交互方式搜索。它是超级文本的简称。

②超媒体（hypermedia）

超媒体是超文本（hypertext）和多媒体在信息浏览环境下的结合，它是超级媒体的简称。用户不仅能从一个文本跳到另一个文本，而且可以激活一段声音、显示一个图形，甚至可以播放一段动画。

互联网采用超文本和超媒体的信息组织方式，将信息的链接扩展到整个互联网上。Web 就是一种超文本信息系统，Web 的一个主要的概念就是超文本链接，它使得文本不再像一本书一样是固定的，线性的。而是可以从一个位置跳到另外的位置。用户可以从中获取更多的信息，可以转到别的主题上。想要了解某一个主题的内容，只要在这个主题上点一下，就可以跳转到包含这一主题的文档上。正是这种多链接性才把它称为 Web。

1.2　Web 系统的结构和组成

1.2.1　Web 系统基本架构

Web 从总体上看，可以分为两部分，即服务器和客户端。两者通过协议进行通信，传递 Web 上的各种信息。

从系统角度来看 Web 系统中的服务器和客户端，可以分成三个层面：应用层、支撑技术和中间件以及位于底层的计算资源。各部分相互之间关系如图 1-2 所示。

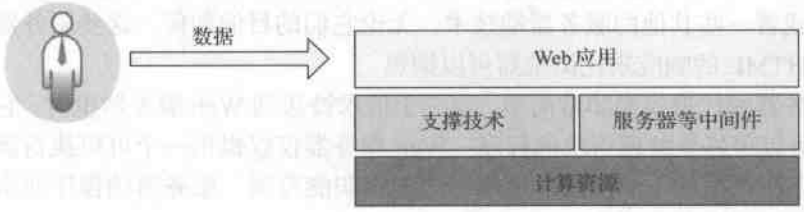


图 1-2　Web 系统基本架构

1.2.2　Web 工作原理

当用户想进入互联网上一个网页或者其他网络资源的时候，通常首先要在浏览器上键入想访问网页的统一资源定位符（URL），或者通过超链接方式链接到指定网页或网络资源。这之后的工作首先是 URL 的服务器名部分，被分布于全球的因特网数据库解析，并根据解析结果决定进入哪一个 IP 地址。接下来的步骤是向在那个 IP 地址工作的服务器发送一个 HTTP 请求。在通常情况下，HTML 文本、图片和构成该网页的一切其他文件很快会被逐一请求并发送回用户。

Web 浏览器接下来的工作是把 HTML、CSS 和其他接收到的文件所描述的内容，加上图像、链接和其他所必需的资源显示给用户，这些构成了所看到的“网页”。

1.2.3 Web 服务器

Web 服务器也称为 WWW（World Wide Web）服务器，主要功能是提供网上信息浏览服务。主要组成包括：

①应用层使用 HTTP 协议；

②HTML 文档格式；

③浏览器统一资源定位符（URL）。

Web 采用的是客户/服务器结构，其作用是整理和储存各种 Web 资源，并响应客户端软件的请求，把客户所需的资源传送到 Windows、Windows NT、UNIX 或 Linux 等平台上。

通俗来讲，Web 服务器传送页面使浏览器可以浏览，然而应用程序服务器提供的是客户端应用程序可以调用的方法。确切一点可以说，Web 服务器专门处理 HTTP 请求，应用程序服务器是通过很多协议来为应用程序提供业务逻辑。服务器主要的工作流程如下：

①接受请求；

②请求的合法性检查，包括安全性屏蔽；

③针对请求获取并制作数据，包括 Java 脚本和程序、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；

④把信息发送给提出请求的客户机。

Web 服务器可以解析 HTTP 协议。当 Web 服务器接收到一个 HTTP 请求时，会返回一个 HTTP 响应，例如，送回一个 HTML 页面。为了处理一个请求，Web 服务器可以响应一个静态页面或图片进行页面跳转，或者把动态响应的内容委托给一些其他的程序例如，CGI 脚本，JSP（JavaServer Pages）脚本，Servlets，ASP（Active Server Pages）脚本，服务器端 JavaScript，或者一些其他的服务器端技术。无论它们的目的如何，这些服务器端的程序通常产生一个 HTML 的响应来让浏览器可以浏览。

Web 服务器的代理模型非常简单。当一个请求被送到 Web 服务器里时，它只单纯地把请求传递给可以很好地处理请求的程序。Web 服务器仅仅提供一个可以执行服务器端程序和返回（程序所产生的）响应的环境而不会超出职能范围。服务器端程序通常具有事务处理、数据库连接和消息等功能。

虽然 Web 服务器不支持事务处理或数据库连接，但它可以配置各种策略来实现容错性和可扩展性，如负载平衡、缓存等。

在 UNIX 和 Linux 平台下使用最广泛的免费 HTTP 服务器是 W3C、NCSA 和 APACHE 服务器，而 Windows 平台 NT/2000/2003 使用 IIS 的 Web 服务器。在选择使用 Web 服务器时应考虑的本身特性因素有：性能、安全性、日志和统计、虚拟主机、代理服务器、缓冲服务和集成应用程序等，下面介绍几种常用的 Web 服务器。

（1）Apache

Apache 仍然是世界上用的最多的 Web 服务器，市场占有率达 60%左右。它源于 NCSA

httpd服务器，当NCSA WWW服务器项目停止后，那些使用NCSA WWW服务器的人们开始交换用于此服务器的补丁，这也是Apache名称的由来（pache——补丁）。世界上很多著名的网站都是Apache的产物，它的成功之处主要在于它的源代码开放、有一支开放的开发队伍、支持跨平台的应用（可以运行在几乎所有的UNIX、Windows、Linux系统平台上）以及它的可移植性等方面。

（2）微软IIS

微软的Web服务器产品是Internet Information Server（IIS），IIS是允许在公共Intranet或互联网上发布信息的Web服务器。IIS是目前最流行的Web服务器产品之一，很多著名的网站都是建立在IIS的平台上。IIS提供了一个图形界面的管理工具，称为互联网服务管理器，可用于监视配置和控制互联网服务。

IIS是一种Web服务组件，其中包括Web服务器、FTP服务器、NNTP服务器和SMTP服务器，分别用于网页浏览、文件传输、新闻服务和邮件发送等方面，它使得在网络（包括互联网和局域网）上发布信息成了一件很容易的事。它提供Intranet Server API（ISAPI）作为扩展Web服务器功能的编程接口；同时，它还提供一个互联网数据库连接器，可以实现对数据库的查询和更新。

（3）Tomcat

Tomcat是一个开放源代码、运行Servlet和JSP Web应用软件的基于Java的Web应用软件容器。Tomcat Server是根据Servlet和JSP规范执行的，因此可以说Tomcat Server也实行了Apache-Jakarta规范且比绝大多数商业应用软件服务器要好。

Tomcat是Java Servlet 2.2和JavaServer Pages 1.1技术的标准实现，是基于Apache许可证下开发的自由软件。Tomcat是完全重写的Servlet API 2.2和JSP 1.1兼容的Servlet/JSP容器。Tomcat使用了JServ的一些代码，特别是Apache服务适配器。随着Catalina Servlet引擎的出现，Tomcat第4版号的性能得到提升，使得它成为一个值得考虑的Servlet/JSP容器，因此目前许多Web服务器都是采用Tomcat。

（4）IBM WebSphere

WebSphere Application Server是一种功能完善、开放的Web应用程序服务器，是IBM电子商务计划的核心部分，它是基于Java的应用环境，用于建立、部署和管理互联网及Intranet Web应用程序。这一整套产品进行了扩展，以适应Web应用程序服务器的需要，范围从简单到高级直到企业级。

WebSphere针对以Web为中心的开发人员，他们都是在基本HTTP服务器和CGI编程技术上成长起来的。IBM将提供WebSphere产品系列，通过提供综合资源、可重复使用的组件、功能强大并易于使用的工具以及支持HTTP和IIOP通信环境，来帮助这些用户从简单的Web应用程序转移到电子商务世界。

（5）BEA WebLogic

BEA WebLogic服务器是一种多功能、基于标准的Web应用服务器，为企业构建自己的应用提供了坚实的基础。各种应用开发、部署关键性的任务，无论是集成各种系统和数据库，还是提交服务、跨互联网协作，起始点都是BEA WebLogic服务器。由于它具有全面的功能、对开放标准的遵从性、多层架构、支持基于组件的开发等优点，基于互联网的

企业都选择它来开发、部署最佳的应用。

BEA WebLogic 服务器在使应用服务器成为企业应用架构的基础方面继续处于领先地位。BEA WebLogic 服务器为构建集成化的企业级应用提供了稳固的基础，它们以互联网的容量和速度，在连网的企业之间共享信息、提交服务，实现协作自动化。

1.2.4 Web 浏览器

Web 浏览器是指可以显示 Web 服务器或者文件系统的 HTML 文件内容，并让用户与这些文件交互的一种软件。网页浏览器主要通过 HTTP 协议与网页服务器交互并获取网页，这些网页由 URL 指定，文件格式通常为 HTML，并由 MIME 在 HTTP 协议中指明。一个网页中可以包括多个文档，每个文档都是分别从服务器获取的。大部分的浏览器本身支持除了 HTML 之外的广泛的格式，如 JPEG、PNG、GIF 等图像格式，并且能够扩展支持众多的插件（plug-ins）。另外，许多浏览器还支持其他的 URL 类型及其相应的协议，如 FTP、Gopher、HTTPS（HTTP 协议的加密版本）。HTTP 内容类型和 URL 协议规范允许网页设计者在网页中嵌入图像、动画、视频、声频、流媒体等。

在 Web 中，客户机的任务是：

①帮助用户制作一个请求（通常在单击某个链接点时启动）；

②将用户的请求发送给某个服务器；

③通过对直接图像适当解码，呈交 HTML 文档和传递各种文件给相应的“查看器”，把请求所得的结果报告给用户。

1.2.5 Web 技术概览

这里通过划分不同的层次和领域，将 Web 领域的主要技术分类列出（见表 1-1）。

表 1-1 Web 技术分类

文档呈现语言	HTML、XHTML、XML、XForms、DHTML
样式格式描述语言	CSS、XSL
动态网页技术	CGI、ASP、ASP.NET、ColdFusion、JSP、PHP、Ruby on Rails
客户端交互技术	ActiveX、Java Applet、Flash、Flex、AJAX、XMLHTTP AIR- Silverlight、JavaFX
客户端脚本语言	JavaScript、JScript、VBScript、ECMAScript、ActionScript
标志定位语言	URL、URI、XPath、URL 重写
文档纲要语言	DTD、XML Schema

1.3 Web 系统及相关技术介绍

1.3.1 HTTP

超文本传输协议（Hypertext Transfer Protocol，HTTP）是互联网上应用最为广泛的一种网络传输协议。所有的 Web 文件都必须遵守这个标准。设计 HTTP 最初的目的是为了提

供一种发布和接收HTML页面的方法。

HTTP的发展是万维网协会和互联网工作小组合作的结果，在一系列的RFC发布中确定了最终版本，其中最著名的是RFC 2616。在RFC 2616中定义了HTTP1.1这个今天普遍使用的版本。以下的介绍围绕HTTP1.1版本进行。

1.3.1.1 HTTP协议是什么

当我们浏览一个网站的时候，只要在浏览器的地址栏里输入网站的地址就可以了，例如，www.baidu.com，但是在浏览器的地址栏里面出现的却是：http://www.baidu.com ，为什么会多出一个“http”？

我们在浏览器的地址栏里输入的网站地址叫URL。就像每家每户都有一个门牌地址一样，每个网页也都有一个互联网地址。当在浏览器的地址栏中输入一个URL或是单击一个超级链接时，URL就确定了要浏览的地址。浏览器通过超文本传输协议（HTTP），将Web服务器上站点的网页代码提取出来，并翻译成网页。因此，在我们认识HTTP之前，有必要先弄清楚URL的组成，例如，http://www.baidu.com/china/index.htm。它的含义如下：

①http://：代表超文本传输协议，通知baidu.com服务器显示Web页，通常不用输入；

②www：代表一个Web（万维网）服务器；

③baidu.com/：这是装有网页的服务器的域名，或站点服务器的名称；

④china/：为该服务器上的子目录，就像我们的文件夹；

⑤index.htm：是文件夹中的一个HTML文件（网页）。

我们知道，互联网的基本协议是TCP/IP协议，然而在TCP/IP模型最上层的是应用层，它包含所有的高层协议。高层协议有：文件传输协议FTP、电子邮件传输协议SMTP、域名系统服务DNS和HTTP协议等。

HTTP协议是用于从Web服务器传输超文本到本地浏览器的传送协议。它可以使浏览器更加高效，使网络传输减少。它不仅保证计算机正确快速地传输超文本文档，还确定传输文档中的哪一部分，以及哪部分内容首先显示（如文本先于图形）等。这就是为什么在浏览器中看到的网页地址都是以http://开头的原因。

HTTP是一个用于在客户端和服务器间请求和应答的协议。一个HTTP的客户端，例如一个Web浏览器，通过建立一个到远程主机特殊端口（默认端口为80）的连接，初始化一个请求。一个HTTP服务器通过监听特殊端口等待客户端发送一个请求序列，就像“GET / HTTP/1.1”（用来请求网页服务器的默认页面），有选择地接收像E-mail一样的MIME消息，此消息中包含了大量用来描述请求各个方面的信息头序列，响应一个选择的保留数据主体。接收到一个请求序列后（如果要的话，还有消息），服务器会发回一个应答消息，诸如“200 OK”，同时发回一个它自己的消息，此消息的主体可能是被请求的文件、错误消息或者其他的一些信息。

HTTP不同于其他基于TCP的协议，诸如FTP。在HTTP中，一旦一个特殊的请求（或者请求的相关序列）完成，连接通常被中断。这个设计使得对于当前页面有规则连接到另一台服务器页面的万维网来说，HTTP是完美的。当持久连接的缺乏成为保持用户状态的