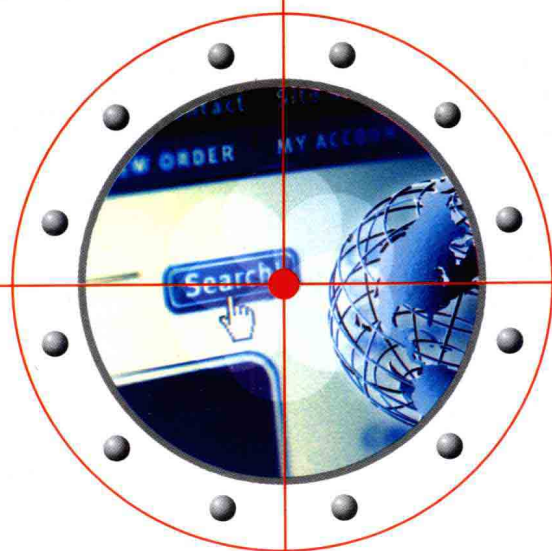


21世纪高职高专“十一五”规划教材

主编 盛 宜 吴晓葵



网站建设 与维护

西北工业大学出版社

21 世纪高职高专“十一五”规划教材

网站建设与维护

主 编	盛 宜	吴晓葵	
副主编	高振江		
编 者	盛 宜	吴晓葵	高振江
	侯维刚	赵 琳	

西北工业大学出版社

【内容简介】 本书以 Windows Server 2003 为平台,以 IIS 6.0 为基础,系统、全面地介绍了网站系统开发、建设、维护、管理的相关技术和方法,最后通过实例分析,对以上知识进行综合运用。全书共分 9 章,主要从网站建设的基础知识、网站建设的方法、网站的维护和管理技术三个方面系统地介绍网站建设、网站维护、页面设计、数据库应用、网络安全的管理、网站宣传和网站推广等知识。

本书可作为高等学校、高职高专院校、成人高校等计算机及相关专业“网站建设与维护”课程的教材,也可供从事网站开发和维护工作的工程技术人员使用。

5

图书在版编目 (CIP) 数据

网站建设与维护/盛宜,吴晓葵主编. —西安:西北工业大学出版社,2008.12

21 世纪高职高专“十一五”规划教材

ISBN 978-7-5612-2501-1

I. 网… II. ①盛…②吴… III. ①网站—开发—高等学校:技术学校—教材②网站—维护—高等学校:技术学校—教材 IV. TP393.092

中国版本图书馆 CIP 数据核字 (2008)第 197923 号

出版发行:西北工业大学出版社

通信地址:西安市友谊西路 127 号 邮编:710072

电 话:(029)88493844 88491757

网 址:www.nwpup.com

印 刷 者:陕西宝石兰印务有限责任公司

开 本:787 mm×1 092 mm 1/16

印 张:18

字 数:437 千字

版 次:2008 年 12 月第 1 版 2008 年 12 月第 1 次印刷

定 价:28.00 元

前 言

计算机网络已经深入到生活的各个角落,对人们的生活产生了深远的影响。而计算机网络正是由一个一个网站构成的,因此网站的建设与维护成为目前计算机网络工作者的一项常规和繁重的工作,市场上对于这类人才的需求量也与日俱增。不管是企业、学校,还是其他的社会部门,都需要建立网站来对自己的业务、工作进行宣传和推广,所以掌握一定的网站建设与维护知识是十分必要的。

网站的建设与维护是一项系统工程,要求用户既要掌握网络基础知识,又要能从管理角度考虑其实用性和易于管理的措施,还要从美观和友好的角度考虑其界面的设计方法。本书从网站建设的基础知识、网站建设的方法、网站的维护和管理技术三个方面系统地介绍网站建设、网站维护、页面设计、数据库应用、网络安全的管理、网站宣传和网站推广等知识。

本书由多年从事网站建设与维护教学并有一定实际工作经验的教师精心编写,以 Windows Server 2003 为平台,以 IIS 6.0 为基础,兼而介绍 Apache,PHP 等网站建设的相关技术和方法。通过技术讲解,首先解决其中的技术性问题,然后对站点建设的方法进行详细讲解,包括 Web 服务器、FTP 服务器、DNS 服务器、流媒体服务器和电子邮件服务器等服务器的建设方法。书中使用较为简单的 Access 2000 为数据库平台,对网络数据库的建设方法和建设技巧进行讲解。

本书侧重于培养读者在具备一定基础知识的前提下进行独立思考和独立完成工作的能力,可作为高等学校、高职高专院校、成人高校等计算机及相关专业“网站建设与维护”课程的教材,也可供从事网站开发和维护工作的工程技术人员使用。

本书由盛宜、吴晓葵主编,高振江担任副主编,侯维刚、赵琳参编。其中第 1,2 章由盛宜编写,第 3 章由赵琳编写,第 4 章由吴晓葵编写,第 5,6,7 章由高振江编写,第 8,9 章及附录部分由侯维刚编写。最后由高振江统稿,盛宜、吴晓葵审稿并定稿。

本书在编写的过程中,参阅了大量的文献资料、教材、论文等,在此恕不一一罗列致谢,谨向各位作者一并表示衷心的感谢。

由于网站建设的技术不断进步,发展速度非常快,加之作者水平有限,书中疏漏之处在所难免,欢迎广大读者批评指正。

编 者

2008 年 11 月

目 录

第 1 章 概述	1	1.9.2 Browser/Server 模式	23
1.1 Internet 概述	1	1.9.3 Browser Application/Server 模式	24
1.1.1 Internet 的概念与历史	1	1.9.4 Web 的工作原理	25
1.1.2 Internet 的组成	2	第 2 章 网页编程技术	26
1.1.3 Internet 的特点与结构	3	2.1 HTML 基础	26
1.2 Intranet 技术	4	2.1.1 HTML 语言概述	26
1.2.1 Intranet 的概念与组成	4	2.1.2 HTML 标记	28
1.2.2 Intranet 网络结构	4	2.2 ASP 介绍	47
1.3 Web 的基本知识	5	2.2.1 ASP 概述	47
1.3.1 Web 的起源	5	2.2.2 ASP 语法	48
1.3.2 超文本与超链接	6	2.3 VBScript 脚本介绍	49
1.3.3 URL	6	2.3.1 VBScript 概述	49
1.3.4 HTML 和 XML	7	2.3.2 VBScript 基础知识	51
1.4 TCP/IP 协议	7	2.3.3 VBScript 常用函数	68
1.5 IP 地址	10	2.4 JavaScript 脚本介绍	74
1.5.1 IP 地址分类	10	2.4.1 JavaScript 基础	74
1.5.2 IP 地址相关说明	11	2.4.2 JavaScript 事件	79
1.5.3 IPv4 地址编码存在的问题	11	2.5 PHP 脚本介绍	82
1.6 子网掩码	11	2.5.1 PHP 脚本基础知识	82
1.6.1 子网掩码的计算	12	2.5.2 服务器脚本比较	82
1.6.2 子网掩码的应用	12	第 3 章 数据库访问技术	86
1.7 域名及 DNS 服务器	13	3.1 Access 数据库介绍	86
1.7.1 域名结构	13	3.1.1 信息与数据	86
1.7.2 域名服务器(DNS)	14	3.1.2 数据库	86
1.8 网络系统的组成	14	3.1.3 数据库管理系统(DBMS)	87
1.8.1 网络概述	14	3.1.4 数据库系统(DBS)	87
1.8.2 网络的基本组成	18	3.1.5 关系	88
1.9 Web 系统的组成与模式	23		
1.9.1 Web 的组成结构	23		

3.1.6 数据库访问技术	88	4.4.2 网站结构设计	127
3.1.7 Access 2000 简介	89	4.4.3 网站形象设计	128
3.1.8 Access 的界面及对象	90	4.4.4 网页设计规则	128
3.2 Access 数据库常用操作	91	4.4.5 网站技术规范	129
3.2.1 创建数据库	91	4.5 网站系统设计	130
3.2.2 创建表	95	4.5.1 网络服务器	131
3.2.3 创建和编辑表之间的关系	101	4.5.2 网络操作系统	133
3.2.4 编辑表	103	4.5.3 服务器配置与集群	136
3.2.5 创建查询	104	4.6 网站的宣传与推广	138
3.3 数据库管理	107	第 5 章 建立 IIS 信息服务	141
3.3.1 Access 数据库的格式转换	107	5.1 Windows Server 2003 概述	141
3.3.2 备份、还原数据库	108	5.2 IIS 6.0 介绍	142
3.3.3 Access 数据库的压缩与修复	108	5.2.1 IIS 6.0 的新特性	142
3.3.4 数据库的导入与导出	108	5.2.2 IIS 6.0 的安装	143
3.4 数据库加密	112	5.2.3 IIS 的目录、用户和服务	143
第 4 章 网站项目管理及网站规划设计	114	5.3 FTP 服务的建立	144
4.1 网站设计方法	114	5.3.1 使用 IIS 6.0 建立 FTP 服务	144
4.1.1 网站设计流程	114	5.3.2 使用 Ser-U 建立 FTP 服务	154
4.1.2 网站设计原则	116	5.4 Web 站点的建立	169
4.1.3 网站设计方法	118	5.4.1 WWW 概述	169
4.1.4 网站设计技术	119	5.4.2 建立 Web 站点	169
4.2 网站需求分析	120	5.5 DNS 服务器的建立	174
4.2.1 网站规划要求	120	第 6 章 邮件和流媒体服务器的组建 ...	179
4.2.2 网站规划分析	121	6.1 电子邮件的基本知识	179
4.3 网站规划设计	122	6.1.1 电子邮件	179
4.3.1 网站规划内容	122	6.1.2 电子邮件的工作过程	179
4.3.2 网站可行性分析	123	6.1.3 电子邮件的地址	180
4.3.3 网站规划方案编写	124	6.1.4 SMTP 协议	180
4.4 网站内容设计	126	6.1.5 中继代理	183
4.4.1 网站主题定位	126		

6.1.6 邮件服务器软件.....	184	7.4.1 防火墙的定义.....	219
6.2 组建 Webmail 邮件系统	185	7.4.2 防火墙的结构.....	221
6.2.1 IIS 6.0 邮件系统	185	7.4.3 防火墙的缺点.....	223
6.2.2 电子邮件安全.....	191	7.4.4 防火墙的发展.....	224
6.3 流媒体的基本知识.....	192	7.5 网络病毒的防范技术.....	225
6.3.1 流媒体的概念.....	192	7.5.1 病毒的分类.....	226
6.3.2 流式传输技术.....	193	7.5.2 病毒的防范技术.....	227
6.3.3 流媒体传输协议.....	194	第 8 章 Web 站点管理与维护	229
6.3.4 流媒体的播放方式.....	195	8.1 网站的性能与缩放性.....	229
6.4 组建流媒体服务器.....	196	8.1.1 网站性能与缩放性概述	229
6.4.1 安装 Windows Media		Services	229
Services	196	8.1.2 网站性能与缩放性需求	231
6.4.2 安装 Windows Media		Encode	231
Encode	198	8.1.3 网站性能与缩放性测试	232
6.5 使用 Windows Media Encode			232
编辑流媒体文件.....	201	8.1.4 网站性能与缩放性测试方法	232
第 7 章 网站安全	205		232
7.1 网站安全概述.....	205	8.1.5 网站性能和缩放性的改进	233
7.1.1 网站的主要安全问题.....	205		233
7.1.2 不安全的网络环境.....	206	8.1.6 性能缩放性与成本.....	234
7.2 Web 服务器安全设置	207	8.2 Windows Server 2003 性能的监视	234
7.2.1 系统服务包和安全补丁			234
.....	207	8.2.1 性能管理.....	235
7.2.2 限制用户权限.....	208	8.2.2 性能监视器的使用.....	237
7.2.3 加固文件系统.....	210	8.3 事件查看器(日志).....	241
7.2.4 删除或禁用不必要的组件和		8.3.1 日志.....	241
服务.....	211	8.3.2 日志的查看.....	243
7.2.5 日志和审核.....	214	8.3.3 日志的审核.....	246
7.3 数据加密技术.....	215	8.3.4 日志的维护.....	247
7.3.1 密码学概述.....	215	第 9 章 网站建设方案实例	249
7.3.2 密码系统的安全性.....	216	9.1 建立网站的必要性和可行性	249
7.3.3 秘密密钥和公开密钥密码			249
系统.....	217	9.2 网站定位及网站设计需求.....	250
7.4 防火墙技术及应用.....	219		

9.3 网站结构方案.....	250	9.5.5 网站功能模块说明.....	258
9.4 网站建设的流程.....	254	9.5.6 合生创展集团网站建设 投资预算.....	261
9.5 网站架设的步骤.....	255	附录	264
9.5.1 树立企业形象.....	255	附录 A HTML 常用标记一览.....	264
9.5.2 网上产品展览方案.....	256	附录 B ASP 常用函数	266
9.5.3 信息反馈系统设置及交流	257	附录 C ASP 错误代码注解	276
9.5.4 在线商务推广.....	257		

第 1 章 概 述

1.1 Internet 概述

1.1.1 Internet 的概念与历史

Internet 是由 ARPAnet 的发展而逐步形成的,是一个使用 TCP/IP 协议连接各个国家、各个地区、各个机构的计算机网络的数据通信网。它是由各种不同类型和规模的主机或网络组成的一个特大型网络,是成千上万条信息资源的汇总。所有的资源分布在世界各地的数十万台计算机上,涉及政治、经济、文化、科学、娱乐等社会的各个领域。

ARPAnet 于 1969 年建成,采用分组交换技术。1974 年,TCP/IP 协议诞生。1980 年,ARPAnet 上的计算机开始采用新的 TCP/IP 协议进行通信。1983 年,ARPAnet 上的协议完全过渡到 TCP/IP。

由于 ARPAnet 起初用在军事上,为保障国防安全,美国在 1983 年将原先的 ARPAnet 分为两个网络,一个是 Milnet,即网络的机密部分,仅供美国国防部使用;另一个是剩余的 ARPAnet,仅供与政府签约合作的科研和教育机构使用。1985 年,美国国家科学基金会(NSF)筹建了 6 个超级计算机中心及国家教育科研网,1986 年形成了用于支持科研和教育的全国规模的计算机网络 NSFNET,并面向社会开放,实现超级计算机中心的资源共享。NSFNET 也采用 TCP/IP 协议,并连接 ARPAnet。NSFNET 实际上取代了 ARPAnet,成为 Internet 的主干网。从此,Internet 开始迅速发展起来。

20 世纪 90 年代初,商业机构开始进入 Internet。1994 年,NSF 宣布不再给 NSFNET 的运行和维护以经费支持,而由其他一些公司运行维护。至此,Internet 彻底完成商业化,形成今天最为著名的遍及世界各地的互联网。目前,Internet 已进入了全新的时期,并朝着商业化、全民化和全球化的趋势发展。我国 Internet 的发展大致经历了以下三个阶段。

第一阶段:1986 年,北京市计算机应用技术研究所开始与国际联网,建立了中国学术网 CANET(Chinese Academic NETwork)。1987 年 9 月,CANET 建成中国第一个国际互联网电子邮件节点,并于 9 月 20 日发出了中国第一封电子邮件,标志着我国 Internet 时代的到来。

第二阶段:1993 年 12 月,以高速电缆和路由器实现将北京中科院、北京大学、清华大学的校园网组成的主干网互连成中关村地区教育与科研示范网(NCFC)。1994 年 4 月,NCFC 开通了连入 Internet 的 64 Kb/s 国际专线,实现了与 Internet 的全功能连接。1994 年 5 月,建立了中国国家顶级域名(CN)服务器,并设在国内。

第三阶段:1995 年 1 月,NCFC 开始向社会提供 Internet 接入服务。从 1995 年开始,我国 Internet 应用进入商业应用领域并得到了突飞猛进的发展。目前,我国有六大 Internet 国际

出口,国内有 4 个与 Internet 连接的网络,分别是中国科学院的中国科学技术网(CSTNET)、教育部的中国教育和科研计算机网(CERNET)、信息产业部的中国公用计算机互联网(CHINANET)、信息产业部所属的吉通公司负责建设的中国金桥信息网(CHINAGBN)。其中,CHINANET 国际线路容量所占比重最大。目前,我国互联网发展状况如表 1.1 所示。

表 1.1 中国互联网发展状况统计

截至时间	CN 域名个数/万个	网站数量/万个	国际出口带宽/(Mb/s)	IPv4 地址数 /万个	网民规模/亿人
2004 年 6 月	38	62.7	53 941	4 942	
2004 年 12 月	43	66.9	74 429	5 995	
2005 年 6 月	62	67.8	82 617	683	1.03
2005 年 12 月	110	69.4	136 106	7 439	1.11
2006 年 6 月	119	78.8	214 175	8 479	1.23
2006 年 12 月	180	84.3	256 696	9 802	1.37
2007 年 6 月	615	131.2	312 346	11 825	1.62
2007 年 12 月	900	150.4	368 927	13 527	2.1
2008 年 6 月	1 190	191.9	493 729	15 814	2.53

1.1.2 Internet 的组成

Internet 是使用 TCP/IP 协议连接各个国家、各个地区、各个机构的计算机网络的数据通信网。Internet 互连了全世界的各种网络和计算机,在组成上可归纳为以下几个部分。

1. 通信线路

通信线路将 Internet 中的计算机、路由器等设备连接起来,是 Internet 的基础设施,如光缆、铜缆、卫星和无线等。通信线路带宽越宽,传输速率越高,传输能力就越强。

2. 路由器

路由器是 Internet 中最重要的设备,实现 Internet 中各种异构网络间的互连,并提供最佳路径选择、负载平衡和拥塞控制等功能。

3. 终端设备

接入 Internet 的终端设备可以是普通 PC 机、服务器、巨型机等,它是 Internet 中不可缺少的设备。终端设备分为服务器和客户机两大类,服务器是 Internet 服务和信息资源的提供者,有 WWW(World Wide Web)服务器、电子邮件(E-mail)服务器、文件传输(FTP,File Transfer Protocol)服务器等,为用户提供 Internet 上的信息服务。客户机是 Internet 服务和信息资源的使用者。

4. 计算机网络

计算机网络是指连接在 Internet 上,分布于世界各地的各种计算机网络,这些网络可以是采用不同的局域网或广域网技术实现的异构网络,所有网络统一使用 TCP/IP 协议互连在一起,实现资源共享。

Internet 的“网际网”逻辑结构如图 1.1 所示。

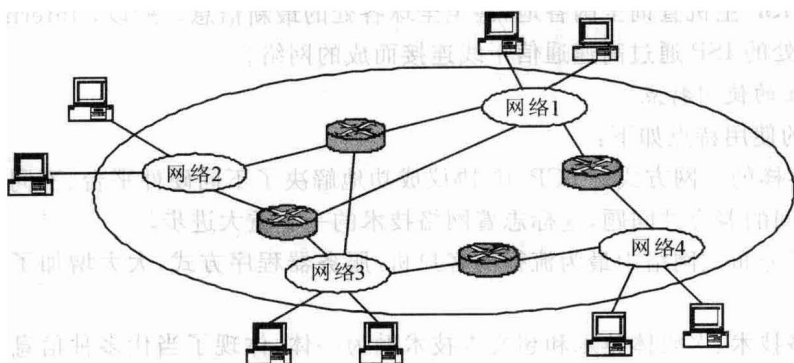


图 1.1 Internet 的“网际网”逻辑结构

1.1.3 Internet 的特点与结构

1. Internet 的结构特点

Internet 的结构特点如下：

(1) Internet 是无中心网,再生能力很强。在 Internet 上,并没有一个中心,一个局部的破坏不会影响整个系统的运行。因此,Internet 特别能适应各种特殊环境,如突发战争、自然灾害等情况。

(2) Internet 已实现无线上网以及移动通信、多媒体通信等多种服务。它所提供的电子邮件(E-mail)、文件传输(FTP)、万维网——全球浏览以及多媒体、移动通信等服务,是以往任何一种通信方式都无法比拟的。因此,Internet 不可避免地会对传统的通信方式(如电报、信函、电话等)带来冲击。

(3) Internet 一般分为外部网和内部网。从安全保密的角度来看,Internet 的安全主要指内部网(Intranet)的安全,因此其安全保密系统要靠内部网的安全保密技术来实现,在内部网与外部网的连接处设置防火墙,以确保内部网不受外部黑客的侵袭。

Internet 覆盖整个地球,凡是采用 TCP/IP 协议并且能够与 Internet 的任何一台主机进行通信的计算机,都可以看成是 Internet 的一部分。到目前为止,Internet 联系着 180 多个国家与地区、40 000 多个子网、500 多万台主机,已经成为世界上最大的计算机广域网络。Internet 采用了目前分布式网络最为流行的客户机/服务器工作方式,在很大程度上增加了网络信息服务的灵活性。

Internet 不仅仅局限于某种特定的计算机或操作系统。不同类型的主机与各种操作系统都可以在 TCP/IP 协议上互连。例如,让 PC 机与大型机互相连接起来就像两台 PC 机相连一样方便。Internet 同时是一个巨大的数据分配网,各台主机都可以识别和理解来自对方的数据和命令。

严格地说,用户不是把自己的计算机直接连接到 Internet 上,而是连接到一个网络上,这个网络又通过网络干线与其他网络相连。网络干线之间可通过路由器互连,使得某个网上的计算机能够和其他任一相连网中的任何一台计算机进行数据和信息交换。如在家中通过拨号上网,只是将家中的计算机连了一台 ISP 主机,而 ISP 主机又通过高速通信干线与我国各城

市、世界各国各地区许许多多的主机相连,因此,虽然我们仅仅连接了一台 ISP 主机,但我们可以通过这台 ISP 主机查询全国各地,甚至全球各地的最新信息。所以,Internet 也可以说是分布在全球各处的 ISP 通过高速通信干线连接而成的网络。

2. Internet 的使用特点

Internet 的使用特点如下:

(1)灵活多样的人网方式。TCP/IP 协议成功地解决了不同硬件平台、不同网络产品和不同操作系统之间的兼容性问题,这标志着网络技术的一个重大进步。

(2)采用了分布式网络中最为流行的客户机/服务器程序方式,大大增加了网络信息服务的灵活性。

(3)把网络技术、多媒体技术和超文本技术融为一体,体现了当代多种信息技术互相融合的发展趋势。

(4)方便易行,有多种上网方式。例如,在任何地方仅需通过一根电话线,普通电脑即可与 Internet 相连。

(5)丰富的信息服务功能和友好的用户接口,操作简单,无须掌握更多的专业计算机知识也可以方便地上网浏览、收发电子邮件、获得极为丰富的信息资源,而且许多是免费的。目前 Internet 已成为服务全世界的通用信息网络。

1.2 Intranet 技术

1.2.1 Intranet 的概念与组成

Intranet 是指应用 Internet 技术组建的企业内部专用网络。它是基于 TCP/IP 协议,使用 WWW 工具,采用防止外界侵入的安全措施,为企业内部服务并具有连接 Internet 功能的企业内部网络。Intranet 网络的基本组成如图 1.2 所示。

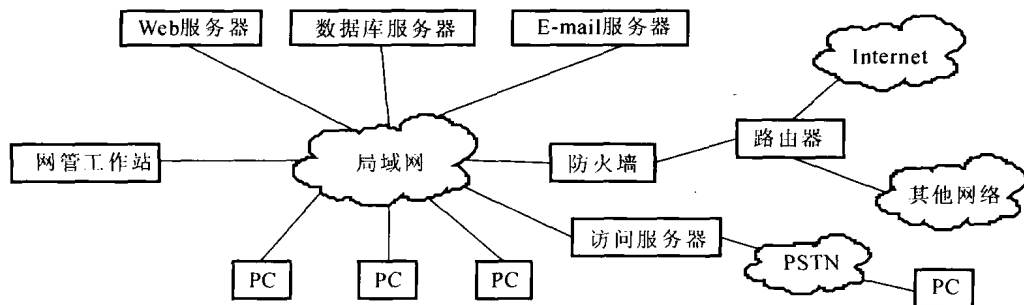


图 1.2 Intranet 网络的基本组成

1.2.2 Intranet 网络结构

Intranet 网络结构采用浏览器/服务器模式,如图 1.3 所示。Intranet 客户端不再与服务器直接相连,而是与 Web 服务器相连,Web 服务器再与数据库服务器相连。用户请求先送到

Web 服务器,再由 Web 服务器通过公用网关接口(CGI)送到数据库服务器,Web 服务器负责将处理结果格式化为 HTML 格式,最后返回客户端的浏览器。



图 1.3 Intranet 网络结构

Intranet 的服务器端是一组 Web 服务器,用以存放 Intranet 上共享的 HTML 标准格式信息以及应用;客户端则为配置浏览器的工作站,用户通过浏览器以 HTTP 协议提出存取请求,Web 服务器则将结果回送给用户。其中,Intranet 通常包含有多个 Web 服务器,一个大型的企业集团的 Intranet 通常会有多达数百个 Web 服务器和工作站,这些服务器有的与机构组织的全局信息及应用有关,有的仅与某一个客户端有关,这些分布组织不仅有利于降低系统的复杂度,也便于开发和维护管理。

此外,考虑到安全性能,可以使用防火墙将 Intranet 与 Internet 隔离开。这样,既可提供对公共 Internet 的访问,又可防止机构内部机密的泄露。

必须指出,Intranet 并不等于局域网,Intranet 可以是局域网、城域网甚至是广域网的形式。

1.3 Web 的基本知识

Web 的全称是 World Wide Web,缩写为 WWW。它是由遍及全球的信息资源组成的系统。这些信息资源包含的内容不仅可以是文本,还可以是图像、表格、音频与视频文件。这种 Web 系统统称为“万维网”,是 Internet 上一种崭新的、活跃的信息服务形式。

1.3.1 Web 的起源

Web 起源于欧洲粒子物理实验室(European Laboratory for Particle Physics,其法文名称的缩写为 CERN)。该机构位于瑞士日内瓦附近,由欧洲共同体国家联合资助,专门从事复杂物理学、工程学和信处理工程学的研究,是世界高能物理研究的精英汇集的场所。

由于从事高能物理研究的科学家分布于世界各地,因此及时地传递思想、共享研究成果比较困难。1989 年,CERN 的 Tim Berners-Lee 首先提出 Web 的发展计划,以便为科学家们提供一种有效的通信手段。初期的 Web 传递信息时采用传统的文本方式,并且仅仅局限于某些学科与地理范围。那时,人们还无法认识到 Web 对现代生活的巨大影响。

在规划 Web 的发展前景时,专家们明智地预测到即将到来的多媒体时代,因而开始在 Web 内运用多媒体技术。漂亮的图片、多样的字体、动听的音乐、视频剪辑、可单击的按钮、超链接等都成为 Web 的一部分,为人们接受及使用 Web 奠定了基础。

现在,Web 已成为存储在 Web 服务器上的大量信息的代名词。服务器通过电缆联系在一起,人们使用 Netscape 的 Navigator, Microsoft 的 Internet Explorer 等 Web 浏览器,就可以访问存储在这些服务器上的网页文件。

Web 上具有相似性质、共同内容的一组信息资源就是一个 Web 站点,网页文件和文件夹是站点的基本组成要素。用于介绍站点内容,引导人们访问站点的网页又称为主页(homepage),它是联系站点内容的纽带。

1.3.2 超文本与超链接

为了理解 Web,首先需要弄清超文本(hypertext)与超链接(hyperlink)的概念。

超文本文件是与其他数据有关联(link)的数据。假设用户正在读“树”这个条目,在文章的最末有一句话:“相关信息参见‘植物’”。从“树”到“植物”这个条目,这就是一个关联。当然,这只是一个简单的例子,万维网基于一个个复杂得多的超文本文件,在文件的任何地方都可能有关联,而不仅是在文件的末尾。

超链接的出现改变了人们按顺序阅读的习惯。单击组成超链接的文本、图片之后,链接目标就会出现在浏览器窗口中;而单击返回主页的超链接时,则将从当前网页直接返回站点主页。使用超链接的 Web 又称为超媒体(hypermedia)。

超链接不仅可以指向其他网页,还可以指向网页内的文本、按钮与图片,从而建立内部链接。内部链接指向的目标又称为书签。单击内部链接时,设置为书签的内容将出现在浏览器窗口中。如果将书签也设置为内部链接,那么用户还可以直接从书签处跳转到网页内的任意位置。超链接的出现,极大地方便了用户对网页内容的访问,轻轻一点鼠标,人们就可以在浏览器窗口中欣赏音乐、发布消息、观看视频剪辑、查询信息资源。

万维网服务的基础是 Web 页面,每个服务站点都包括若干相互关联的页面。超链接指向的资源,可以是另一个 Web 页面、另一个文件或另一个 Web 站点。另一个站点的页面又可指向其他站点,使全球范围的 WWW 服务器连成一体。这就是所谓的超文本和超链接技术,用户只要用鼠标在 Web 页面上单击,就可获得全球范围的多媒体信息服务。

1.3.3 URL

超链接帮助人们访问网页的内容,网页通过统一资源定位器(URL, Uniform Resource Locator)定位。URL 用于表示网络资源的地址,它的功能相当于日常使用的通信地址,也有人将 URL 称为网址。例如,在 Internet Explorer 6.0 窗口的地址栏内输入 URL“http://www.ccert.edu.cn/advisories/index.htm”并回车之后,浏览器就会在 Web 上搜索该网址,并把 CERNET 网络应急响应组的主页传送到用户的面前。

URL 可以说是由四部分内容组成的,这四部分内容分别代表网络协议、域地址、文件夹与网页文件的名称。它的基本形式是“How://Who/Where/What”。How 表示网络协议(protocol),它是网络资源发送、接收时要共同遵守的规定。Web 专用的协议是超文本传输协议(HTTP),该协议与 Web 站点(域)类型之间使用冒号和斜杠分隔。由于 WWW 已经成为代表 Web 站点类型的标志,因此,凡是以“http://www”开头的 URL 都可以初步认定是 Web 站点的网址。如果未安装 TCP/IP 网络协议或者安装的协议遭到破坏,那么,用户可通过 Windows 98/2000 的控制面板来安装 TCP/IP 网络协议。Who 表示存储网络资源的 Web 服务器的名称,又称为域地址(domain address),它是由站点名和域名(domain)组成的。例如,在 URL“http://www.ccert.edu.cn”中,www 代表站点类型,ccert 代表站点名,edu.cn 代表域名。Where 和 What 则用于确定网络资源在 Web 服务器上的位置。前者一般是指保存网络

资源的文件夹,后者是指用户访问的网页文件的名称。网页名称需要使用扩展名,否则浏览器可能会认为它是文件夹。

1.3.4 HTML 和 XML

超文本标记语言(HTML,Hyper Text Markup Language)是编制 Web 页面文档的语言。HTML 是标准通用型标记语言(SGML,Standard Generalized Markup Language)的一种。SGML 是定义结构化文本类型和标记这些文本类型的标记语言系统。HTML 的标记符定义了文档结构、字形字体、版面布局、超链接等超文本文档结构,使 Web 浏览器能够阅读和重新格式化任何 Web 页面。HTML 的最新版本是 HTML 4.0,它支持 object 标记和层叠风格表单(CSS)。通过支持 ISO 10646 字符集,它支持任何语言所需的标记,为插入对象和脚本提供了方便,并在文件中支持格式项。

HTML 仅关注 Web 浏览器的页面安排,缺乏对结构化数据的表示能力,采用单一固定语法,缺乏可扩展性,因此,难以满足用于显示除文本之外的其他需求,如数据处理、刷新排列、存储、转发、交换、加密及签名等,同时也难以满足协同工作、信息重用等对大量动态数据使用统一数据格式进行表示和管理的需要。

可扩展标记语言(XML,Extensible Markup Language)是一种元语言,即用来描述其他语言的语言。它可以让用户定义自己的标记语言,从而在 XML 文件中描述并封装数据。XML 实际上是 SGML 的一个优化子集。SGML 是 ISO 在 1986 年推出的用来创建标记语言的语言标准,它提供一种将数据内容与显示分离开来的数据表示方法,使得数据独立于机器平台和处理程序。SGML 在大型文档管理、出版业等领域得到了非常好的应用,但它的复杂性使其难以直接应用到 Internet 上。基于 SGML 的 HTML 语言主要考虑数据的外观表示,在可扩展性、交互性和语义方面则先天不足,缺乏对结构化数据的表示能力,鉴于 SGML 和 HTML 各自的优点与缺陷,W3C 协会(W3C,World Wide Web Consortium)于 1998 年正式公布了 XML 规范。XML 在数据交换、存储和描述等方面都具有自己的优点,它保留了 SGML 80% 的功能,但只有 SGML 约 20% 的复杂度,使之可以直接用于网上传输。XML 突出了数据的语义与元素结构描述能力,从而可以解决协同工作、信息重用中面临的许多问题。

2001 年 5 月,W3C 又推出了一个替代 DTD(Data Type Definition)的建议标准,称为模式(Scheme)。Scheme 比 DTD 更具灵活性,功能也更强大,例如,支持多种数据类型、支持命名空间(Namespace)等。它将逐渐代替 DTD,并进一步推动 XML 的广泛应用。

简而言之,XML 和 HTML 的本质区别是:HTML 标记是用来描述怎样在屏幕上显示信息,解决人一机交互的问题,而 XML 用来描述信息是什么,解决机器之间的交互问题。

1.4 TCP/IP 协议

TCP/IP 协议全称为“传输控制协议和互联网协议”,是目前 Internet 中最常用的一种网络协议,也是国际互联网的基础。TCP/IP 协议的目的是使不同厂家生产的计算机能在各种网络环境下运行。

TCP/IP 是一种可选择路由的通信协议。它具有很强的灵活性,可支持任意规模的网络,几乎可连接所有的服务器和工作站。但其灵活性也给它的使用带来了一些不便,例如它的设

置和管理比 IPX/SPX 兼容协议、NetBEUI 协议都要困难和复杂一些。

TCP/IP 协议是一个协议集,如图 1.4 所示。其中最重要的协议一个是传输控制协议(TCP, Transmission Control Protocol),另一个是网际互联协议(IP, Internet Protocol),统称为 TCP/IP 协议。该协议是 Web 通信的基础,Web 的通信协议(如 HTTP, FTP)需要 IP 和 TCP 的支持。

TCP/IP 协议具有如下 4 个特点:

- (1)开放的协议标准,可以免费使用,并且独立于特定的计算机硬件与操作系统;
- (2)独立于特定的网络硬件,可以运行在局域网、广域网中,更适用于网络互连;
- (3)统一的网络地址分配方案,使得网络中的每台主机在网中都具有唯一的地址;
- (4)标准化的高层协议(如 HTTP, FTP, SMTP, DNS 等),可以提供多种可靠的用户服务。

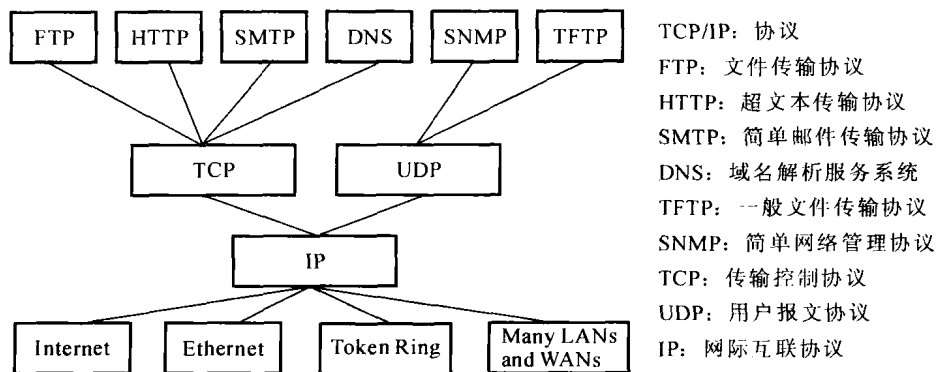


图 1.4 TCP/IP 协议集示意图

在 TCP/IP 协议中, TCP 协议和 IP 协议各有分工。TCP 协议是 IP 协议的高层协议, TCP 在 IP 之上提供一个可靠的、面向连接的协议。TCP 协议能保证数据包的传输,以及正确的传输顺序,并且它可以确认包头和包内数据的准确性。如果在传输期间出现丢包或错包的情况, TCP 负责重新传输出错的包,这样的可靠性使得 TCP/IP 协议在会话式传输中得到充分应用。

IP 协议为 TCP/IP 协议集中的其他所有协议提供“包传输”功能, IP 协议为计算机上的数据提供一个最有效的无连接传输系统。也就是说, IP 包不能保证到达目的地, 接收方也不能保证按顺序收到 IP 包, 它仅能确认 IP 包头的完整性。确认包是否最终到达目的地, 还要依靠 TCP 协议, 因为 TCP 协议是面向连接服务的。

OSI 参考模型是国际标准化组织提出的一个网络互连模型, 该模型将网络体系结构划分为 7 层, 分别为物理层、数据链路层、网络层、传输层、会话层、表示层和应用层。图 1.5 是对 OSI 和 TCP/IP 模型进行的对比。

TCP/IP 体系结构的网络接口层(Network Interface)、网络互连层(即 IP 层, 包括 ICMP, Any RARP 等协议)、传输层(TCP, UDP)、应用层(高层协议), 分别对应 OSI 参考模型的物理层与数据链路层、网络层、传输层、会话层、表示层和应用层。

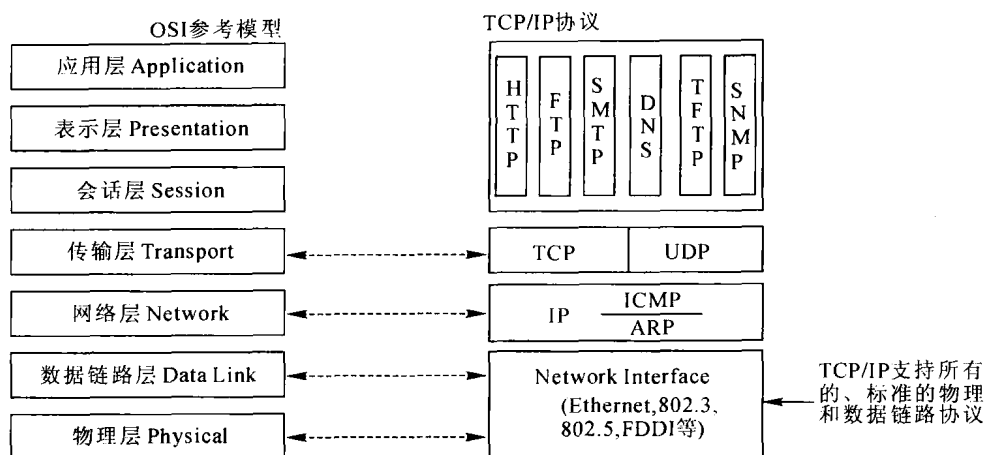


图 1.5 OSI 与 TCP/IP 的对比

TCP/IP 体系结构分为 4 个层次,各层的功能如下:

(1)网络接口层。该层是整个体系结构的基础部分,负责接收 IP 层的 IP 数据包,通过网络向外发送;或接收处理从网络上来的物理帧,抽出 IP 数据包,向 IP 层发送。该层是主机与网络的实际连接层,与邮政系统相比,网络接口层中的比特流传输相当于信件的运送。

(2)网络互连层。该层是整个体系结构的核心部分,负责处理互联网中计算机之间的通信,向传输层提供统一的数据包。它的主要功能是处理来自传输层的分组发送请求,处理接收的数据包,处理互连的路径。

网络互连层 IP 协议提供了不可靠、无连接的数据传输服务,数据包从一个主机经过多个路由器到达目的主机。如果路由器不能正确地传输数据包,或者检测到异常现象影响数据包的正确传输,路由器就要通知原主机或路由器采取相应的措施。ICMP(Internet Control Messages Protocol)为 IP 协议提供了差错控制、网络拥塞控制和路由控制等功能。

网络互连层的 ARP(Address Resolution Protocol)提供地址转换服务,查找与给定 IP 地址对应的主机的网络物理地址(网卡的 MAC 地址)。与 ARP 功能相反的是 RARP(Reverse ARP),RARP 主要解决物理地址到 IP 地址的转换。

ARP 采用广播消息的方法来获取网上 IP 地址对应的 MAC 地址。对于使用低层介质访问机制的 IP 地址来说,ARP 是非常通用的。当一台主机要发送报文时,首先通过 ARP 获取 MAC 地址,并把结果存储在 ARP 缓存的 IP 地址和 MAC 地址表中,下次该站需要发送报文时,就不用再发送 ARP 请求,只要在 ARP 缓存中查找就可以了。与 ARP 类似,RARP 也采用广播消息的方法来决定与 MAC 地址相对应的 IP 地址。RARP 对于网络上的无盘客户机来说显得尤为重要,因为无盘客户机在系统引导时根本无法知道它自己的 IP 地址。

(3)传输层。该层是整个体系结构的控制部分,负责应用进程之间端到端的通信。传输层定义了两种协议:传输控制协议 TCP 与用户数据报协议 UDP。TCP 是一种可靠的面向连接的协议,允许从一台机器发出的字节流无差错地发往互联网上的其他机器。TCP 将应用协议的字节流分成报文段,并将报文段传输给网络层。在接收端,TCP 接收进程将收到的报文段再组装成应用协议字节流。TCP 还可用于处理流量控制,以避免快速发送方向低速接收方发