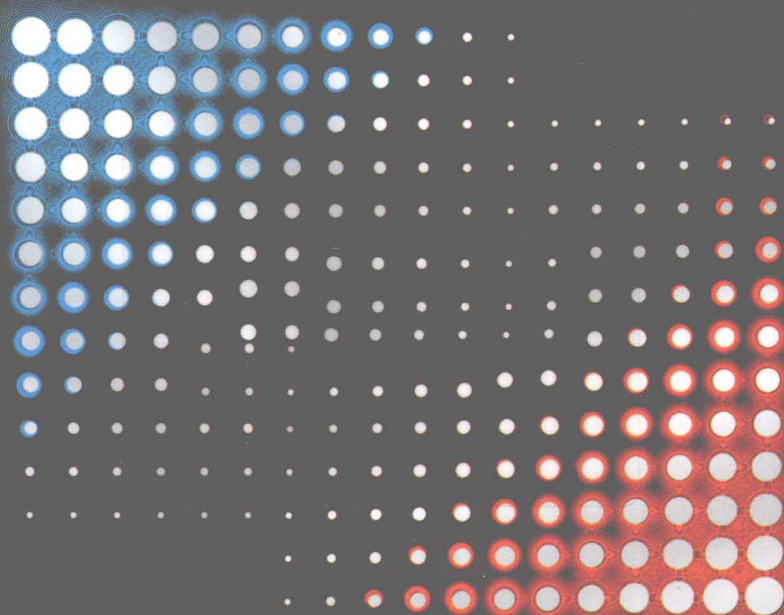


新 编 计 算 机 类 本 科 规 划 教 材

网站组建、管理与维护 (第2版)

杨 威 主 编

高立同 杨陟卓 赵 鑫 孙清亮 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

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

新编计算机类本科规划教材

网站组建、管理与维护

(第2版)

杨 威 主编

高立同 杨陟卓 赵 鑫 孙清亮 编著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书概要介绍了网站组建与管理的概念、技术发展和理论基础,较全面、系统地介绍了网站组建基础工作,Web 服务器配置与管理,Web 主页设计基础,Web 网站安全设置与管理,邮件服务器组建与管理,视频服务器组建与管理,网站服务器高级配置与管理,电子政务网站构建基础,以及 Web 网站运维管理等内容。结合笔者多年从事网络系统集成与管理的实践,为读者提供一些典型案例。

本书内容通俗易懂、结构清晰、实用性强,具有教材和技术文档的双重特征。本书适合高等院校计算机科学与技术、网络工程、软件工程、电子信息工程、电子商务、信息管理与信息系统、教育技术学等专业的学生使用,也适合作为系统集成培训自学教材,亦是网络工程技术人员和管理人员的技术参考文档与必备的工具书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

网站组建、管理与维护 / 杨威主编; 高立同等编著. —2 版. —北京: 电子工业出版社, 2011.5

(新编计算机类本科规划教材)

ISBN 978-7-121-13520-0

I. ①网… II. ①杨… ②高… III. ①网站—高等学校—教材 IV. ①TP393.092

中国版本图书馆 CIP 数据核字(2009)第 084510 号

策划编辑: 冉 哲

责任编辑: 冉 哲

印 刷: 涿州市京南印刷厂

装 订: 涿州市桃园装订有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编: 100036

开 本: 787×1092 1/16 印张: 20 字数: 512 千字

印 次: 2011 年 5 月第 1 次印刷

印 数: 3 000 册 定价: 38.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

再 版 前 言

本书第 1 版自 2005 年出版后,曾多次印刷、出版发行,深受广大读者的喜爱。为适应网站组建、管理与维护的发展,并保持教材内容的先进性及可操作性,又提出了对该书进行修订的要求。本次修订在第 1 版的基础上,去掉一些与网站技术关联不密切或滞后的内容,整合相关章节的内容,增加与网站组建有关的新技术、新方法;从整体上对网站组建、管理与维护的理论、方法和技术,进行了全新组织、充实和完善;尽可能地反映当今网站组建、管理与维护的新思想,以适应读者的需求。

教材修订,充分考虑应用型本科学生的认知特征和学习目标。依据系统观点,从网站组建、管理与维护的视角组织相关内容。全书内容连贯、层次结构分明,方法与技术融合、具有良好的逻辑性。通过“案例学习”和“上机实验”等环节,体现教材内容的实践性与可操作性。将网站组建、管理与维护中较难理解的技术和方法,分散在不同的章节介绍,实现了难度分散的编写目的,便于学生理解与掌握。

教材修订,紧扣“网站组建、管理与维护”主题,突出内容的实用性和整体性。这些内容包括:高校网站服务器总体架构,网站节能机房设计,基于 IIS 7.0 的 Web 服务器配置与管理,主页构图与色彩设计,主页创意设计,Web 网站安全设置与管理,Windows Server 2008 高级安全防火墙应用,基于 ACL 的边界网络安全设置,Postfix 的安装配置和电子邮件安全技术,Webmail 管理系统 ExtMail 的使用,Windows Server 2008 视频服务器组建与管理,Windows Server 2008 的数据备份与恢复,Windows Server 2008 集群设置,Web 网站双机负载均衡设计与配置,安全可信电子政务网站技术路线,网站性能改善措施,使用 Sniffer Pro 诊断网站性能,以及网络运维管理方法等。这些网络实用性技术和相关案例均来自工程实践。读者可以直接应用在网站设计与系统集成项目之中,或稍加修改即可作为实际的网络工程使用。

教材修订,注重网络工程设计与系统集成的新技术、新方法的介绍。例如,IPv6 地址与域名解析,网站节能机房设计,基于 IIS 7.0 的 Web 服务器配置与管理,Windows Server 2008 高级安全防火墙应用,Windows Server 2008 集群设置,Web 网站双机负载均衡设计与配置,安全可信电子政务网站技术路线等。

教材修订,注重精品课程建设,以方便读者学习。采用 Web 网站提供了与本书配套的课程大纲、PPT 课件、学习案例、网络实训、习题参考答案及问题答疑等教学资源。为了降低纸质教材成本,该教材的 Web 网站补充了一些实用技术和学习案例。

总之,本次修订仍坚持“系统观点为纲、实用技术为主、工程实践为线、侧重主流产品”的原则;立足于“看得懂、学得会、用得上”的策略;由浅入深、循序渐进地介绍了网站组建、管理与维护的理论、方法和技术。教材仍保持了选题适当、结构完整、层次清晰、实用性强等特点。

本书由山西师范大学杨威教授主编。本书第 1, 2, 5 由杨威编写;第 4 章由孙清亮

(山西师范大学)编写;第6章由赵鑫讲师(山西师范大学,北京理工大学2010级博士生)编写,第7章由高立同讲师(山西师范大学)编写,第8,10章由杨陟卓(北京理工大学2009级博士生)编写,第3,9章由杨威、高立同合编。全书由杨威统稿、定稿。

本书的出版得益于电子工业出版社的关怀和支持。本书在编写过程中,汲取了许多Web网站建设、管理与维护方面专著和论文的思想,得到了许多老师的帮助。在本书出版之际,对以上给予我们帮助、鼓励、支持的老师,在此一并表示感谢。

由于作者水平有限,错误、疏漏之处在所难免,敬请广大读者不吝赐教。

作 者

目 录

第 1 章 Web 网站概述	1	2.1.3 网站工程项目阶段划分	23
1.1 Web 的基本知识	1	2.2 网站需求分析	24
1.1.1 Web 的起源	1	2.2.1 需求分析思想	25
1.1.2 超文本与超链接	2	2.2.2 项目经理的职责	25
1.1.3 URL	2	2.2.3 需求调查文档记录	26
1.1.4 HTML 和 XML	3	2.2.4 用户调查内容	26
1.1.5 Web 2.0	4	2.2.5 市场调研内容	26
1.2 TCP/IP 基本知识	4	2.2.6 网站需求规格说明与网站功能 描述书	27
1.2.1 网络协议	5	2.3 网站开发的 RSI 方法	28
1.2.2 OSI 参考模型与数据封装	5	2.3.1 网站开发用例分析	28
1.2.3 TCP/IP 体系结构	6	2.3.2 用例分析的 RSI 方法	29
1.3 IPv4 地址与域名	9	2.3.3 网站开发的 RSI 模型	32
1.3.1 IPv4 地址	9	2.4 网站支持平台构建基础	35
1.3.2 子网与子网掩码	10	2.4.1 网站平台组建原则	35
1.3.3 IPv4 域名解析	11	2.4.2 网络服务器选型	36
1.4 IPv6 地址与域名	12	2.4.3 网络操作系统选型	37
1.4.1 IPv6 地址	12	2.4.4 网站拓扑结构设计	38
1.4.2 IPv6 域名解析	14	2.4.5 网站安全管理思想	39
1.5 Web 系统的组成与模式	16	2.5 网站运行环境设计	40
1.5.1 Web 服务器的组成	16	2.5.1 机房布线设计	40
1.5.2 Browser/Server 模式	16	2.5.2 机房供电设计	40
1.5.3 Browser/Application/Server 模式	17	2.5.3 机房制冷设计	42
1.5.4 Web Service 的体系结构	18	2.5.4 绿色节能机房	44
1.5.5 Web 系统的工作原理	19	2.5.5 机房接地保护	45
1.6 Web 系统的安全	20	2.6 网站创意设计	45
1.6.1 Web 安全体系结构	20	习题与思考	47
1.6.2 Web 访问过程的安全	20	第 3 章 Web 服务器配置与管理	48
习题与思考	21	3.1 服务器技术概述	48
第 2 章 网站组建基础工作	22	3.1.1 服务器功能与分类	48
2.1 网站组建概述	22	3.1.2 服务器的 CPU 结构	49
2.1.1 网站项目相关概念	22	3.1.3 对称多路处理技术	51
2.1.2 网站组建层面	23	3.1.4 内存技术	51

3.1.5 磁盘存储接口技术	53	4.2.2 色彩三要素	101
3.1.6 网络存储技术	55	4.2.3 色彩对比与调和	102
3.1.7 服务器虚拟化技术	56	4.3 主页风格设计	103
3.1.8 控制与管理技术	56	4.3.1 资讯主页	103
3.2 Web 服务器选型与配置	57	4.3.2 娱乐主页	105
3.2.1 Web 服务器性能与配置	57	4.3.3 企业主页	106
3.2.2 Windows Server 2008 功能	62	4.3.4 教育主页	107
3.2.3 安装 Windows Server 2008 中文版	63	4.3.5 政府主页	108
3.2.4 配置 Windows Server 2008 服务器	64	4.3.6 商务主页	109
3.3 活动目录的结构及安装配置	66	4.4 Web 网站发布	111
3.3.1 活动目录的逻辑结构及其 相关概念	66	4.4.1 发布前的测试	112
3.3.2 安装设置活动目录域控制器	68	4.4.2 取得相应的权限	112
3.4 Windows Server 2008 组管理	72	4.4.3 网站发布	113
3.4.1 用户组概念与特点	72	4.5 Web 网站测试	114
3.4.2 创建用户组	72	4.5.1 网页测试注意问题	114
3.4.3 使用组时应该注意的问题	74	4.5.2 网页测试项目	114
3.5 配置 DNS 与 DHCP 服务	74	4.5.3 服务器的错误提示信息	115
3.5.1 配置域名解析服务	74	4.6 Web 网站维护	116
3.5.2 配置动态主机分配服务	79	4.6.1 网站维护	116
3.5.3 客户端安装与配置	84	4.6.2 网站更新	116
3.6 Web 和 FTP 站点设置与管理	84	4.6.3 网站升级	117
3.6.1 安装 IIS 7.0	84	习题与思考	117
3.6.2 Web 站点的设置	85	第 5 章 Web 网站安全设置与管理	118
3.6.3 FTP 站点的设置	87	5.1 Web 系统安全弱点	118
3.6.4 多域名与 IP 地址指派	90	5.2 加固操作系统的安全	120
习题与思考	91	5.2.1 系统服务包和安全补丁	121
网络实训	91	5.2.2 系统账户安全配置	121
第 4 章 Web 网站设计与管理基础	94	5.2.3 文件系统安全设置	124
4.1 网站策划	94	5.2.4 安全模板创建与使用	125
4.1.1 网站目标	94	5.2.5 使用安全配置和分析	126
4.1.2 网页设计原则	94	5.2.6 使用安全配置向导	127
4.1.3 网站框架组织	96	5.3 设置 Web 服务器的安全	134
4.1.4 网页造型	97	5.3.1 IIS 的安全机制	134
4.1.5 网页可访问性	98	5.3.2 设置 IP 地址限制	135
4.2 网站配色设计	100	5.3.3 设置用户身份验证	136
4.2.1 色彩体系	100	5.3.4 设置授权规则	137
		5.3.5 设置 SSL 证书验证	137
		5.3.6 设置文件的 NTFS 权限	138
		5.3.7 审核 IIS 日志记录	139

5.4 高级安全 Windows 防火墙	140	7.1.4 流媒体技术的应用	186
5.4.1 高级安全 Windows 防火墙概述	141	7.1.5 流媒体应用系统的组成	187
5.4.2 设置入站规则保护 Web 站点	143	7.1.6 流媒体软件的选择	188
5.5 保护网络边界安全	145	7.2 Media Service 协议及功能	189
5.5.1 防火墙和路由器应用	145	7.2.1 Media Services 协议	189
5.5.2 使用 DMZ 保护内网安全	147	7.2.2 Windows Media Service 2008 新增功能	191
5.6 设置边界路由器的 ACL	147	7.3 组建 Windows Media 服务器	192
5.6.1 ACL 的作用与分类	148	7.3.1 流媒体服务器选型	192
5.6.2 ACL 的配置	148	7.3.2 安装 Media Server	194
5.6.3 ACL 的执行与位置	149	7.4 流媒体文件的制作	198
5.6.4 扩展 ACL 的应用	150	7.4.1 Windows Media 支持的源文件类型	199
习题与思考	152	7.4.2 Windows Media 流文件制作工具	199
网络实训	152	7.4.3 Windows Media 编码器	200
第 6 章 邮件服务器组建与管理	154	7.4.4 文件格式转换	202
6.1 电子邮件的基本知识	154	7.4.5 视频文件的录制	204
6.1.1 SMTP 的基本结构	154	7.5 Media Service 的点播和广播服务	207
6.1.2 电子邮件工作原理	156	7.5.1 组建点播服务	207
6.1.3 电子邮件信头结构及分析	158	7.5.2 组建广播服务	213
6.1.4 如何选择邮件服务器软件	159	7.5.3 现场实况直播	214
6.2 安装和配置 Postfix 邮件服务器	161	7.5.4 制作播放列表	216
6.2.1 安装邮件服务器	161	7.5.5 发布广告	217
6.2.2 配置邮件服务器	166	7.5.6 Media Server 运行监视	219
6.2.3 测试邮件服务器	171	习题与思考	220
6.3 电子邮件的安全	173	网络实训	220
6.3.1 影响邮件安全的因素	173	第 8 章 网站服务器高级配置与管理	222
6.3.2 邮件安全的主要问题	174	8.1 网络存储技术	222
6.3.3 保密电子邮件及特性	175	8.1.1 服务器传统存储技术	222
6.3.4 数字证书与 S/MIME	176	8.1.2 网络附加存储技术	223
6.3.5 邮件系统病毒防御	178	8.1.3 存储区域网络技术	224
6.3.6 Webmail 系统安全	180	8.1.4 NAS 与 SAN 的比较	226
6.3.7 邮件安全的其他技术	180	8.1.5 iSCSI 存储技术	227
习题与思考	181	8.2 数据备份与恢复技术	228
网络实训	182	8.2.1 备份与恢复基本知识	228
第 7 章 视频服务器组建与管理	183	8.2.2 数据备份类型与方法	230
7.1 流媒体的基本知识	183	8.2.3 网络存储备份技术	231
7.1.1 流媒体的概念	183	8.3 服务器备份与恢复管理	233
7.1.2 流媒体技术原理	184		
7.1.3 流媒体播放方式	184		

8.3.1 安装 Windows Server Backup	233	9.4.2 申请和安装服务器证书	272
8.3.2 使用 Windows Server Backup 备份数据	234	9.4.3 设置 Web 服务器的安全通信 ...	279
8.3.3 使用 Windows Server Backup 恢复数据	236	9.4.4 创建网站证书	280
8.4 服务器集群与负载均衡	238	9.4.5 访问网站证书	281
8.4.1 服务器集群与容错技术	238	习题与思考	283
8.4.2 双机集群工作模式与原理	239	网络实训	284
8.4.3 Web 网站双机负载均衡设计	241	课程设计	284
8.4.4 Web 网站负载均衡配置	243	第 10 章 Web 网站运维管理	286
8.5 服务器集群与虚拟化应用	248	10.1 网站性能测试	286
8.5.1 服务器虚拟化节能	248	10.1.1 网站性能及指标	286
8.5.2 服务器选型与配置	248	10.1.2 性能测试类型与方法	288
8.5.3 多服务器整体架构	249	10.1.3 网站可靠性测试	290
8.6 多服务器集中存储的应用	250	10.1.4 网站吞吐率测试	290
8.6.1 集中存储需求分析	250	10.2 网站性能改善	291
8.6.2 IP SAN 产品选型	250	10.2.1 网站性能改善技术措施	291
8.6.3 多服务器集中存储	251	10.2.2 调整和优化服务器内存	294
习题与思考	251	10.2.3 服务器资源优化	295
网络实训	252	10.2.4 建立与完善网络配置文档	296
第 9 章 电子政务网站构建基础	253	10.3 Windows 可靠性和性能监视器	296
9.1 电子政务概述	253	10.3.1 Windows 可靠性和性能监视器 概述	296
9.1.1 电子政务建设背景	253	10.3.2 使用资源视图监视系统活动 ...	297
9.1.2 电子政务网络体系结构	254	10.3.3 使用性能监视器	298
9.2 政务网站设计基础	255	10.3.4 使用可靠性监视器	299
9.2.1 政务网站信息模型	256	10.4 网络嗅探技术与运行监测	300
9.2.2 政务信息系统结构与特点	256	10.4.1 网络嗅探技术	300
9.2.3 政务网站功能结构	257	10.4.2 Sniffer 软件	301
9.2.4 政务处理逻辑结构	258	10.4.3 使用 Sniffer 软件监测网络	301
9.3 电子政务安全技术	259	10.5 使用日志维护网络系统	305
9.3.1 电子政务 PKI 部署	259	10.5.1 基于 Web 的网管日志	305
9.3.2 基于 VPN 的安全通信	261	10.5.2 利用日志排除故障	306
9.3.3 基于 MPLS 的安全通信	262	10.5.3 利用日志分析网络性能	306
9.3.4 电子政务业务互访	265	10.6 网站系统性能与安全评估	307
9.3.5 电子政务安全通信协议	267	10.6.1 评估基本知识	307
9.3.6 电子政务的点对点通信	268	10.6.2 网站健壮性评估	308
9.3.7 网络行为监管与审计	270	10.6.3 网站安全性评估	309
9.4 使用 SSL 与证书建立可信网站	271	习题与思考	310
9.4.1 基于 SSL 的 Web 服务器	271	网络实训	310
		参考文献	312

第1章 Web 网站概述

本章简单介绍了 Web 的起源，超文本和超链接，URL，HTML 和 XML，以及 Web 2.0，重点叙述了 TCP/IP 协议栈，IPv4 地址和子网划分及域名系统，概要介绍了 IPv6 地址表示，域名解析原理，简要分析了 Web 平台组成、工作模式、工作原理，以及 Web 系统的安全问题。通过本章学习，要达到以下目标。

① 理解超文本和超链接、URL，Web 2.0 等概念；了解 IPv6 地址表示，域名解析原理；理解 Web 平台组成，工作模式和工作原理，以及 Web 安全体系结构和 Web 访问过程的安全。

② 熟悉 TCP/IP 体系结构，TCP/IP 协议，ARP 协议，PDU 及数据封装过程，IPv4 地址分类，以及域名系统原理；掌握 IPv4 子网划分与子网掩码设置；了解 Web 网站相关知识，建立 Web 平台体系结构。

1.1 Web 的基本知识

Web 的全称是 World Wide Web，缩写为 WWW。Web 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 系统通称为“万维网”或“全球信息网”，是 Internet 上流行的信息服务形式。

1.1.1 Web 的起源

Web 起源于欧洲粒子物理实验室（英文为 European laboratory for particle physics，其法文名称缩写为 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们聚会的场所。

高能物理研究的科学家分布于世界各地，及时地传递思想、共享研究成果就显得非常必要。1989 年，CERN 的 Tim Berners-Lee 首先提出了 Web 的发展计划，为科学家们提供了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，并且仅局限于某些学科与有限的范围。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，多媒体时代即将到来，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、多彩的视频剪辑，可单击的按钮和超链接等都成为 Web 的一部分，为普通人接受及使用 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。服务器通过电缆和网络联系在一起，Internet 漫游者使用 Netscape 的 Navigator，Microsoft 的 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 Internet 漫游者访问站点的网页又称为主页 (Home Page)，它是联系站点内容的纽带。

1.1.2 超文本与超链接

为了理解 Web，首先需要清楚超文本 (Hypertext) 与超链接 (Hyperlink) 的概念。超文本文件是与其他数据有关联 (Links) 的数据。超文本文件的一个简单例子是大百科全书。假设用户正在读“树”这个条目，在文章的最后有一条参考写到，“相关信息参见‘植物’”。这最后一行就是一个关联，从“树”到“植物”这个条目。当然，这只是一个简单例子，万维网是一个复杂得多的超文本文件，特别是，在文件的任何地方都可能有关联，而不仅是在文件的末尾。

多样化的文字、漂亮的图片、动听的声音、流畅的视频使 Web 具有了多媒体特性，超链接的出现则改变了人们按顺序阅读的传统习惯。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至站点主页。使用超链接的 Web 又称为超媒体 (Hypermedia)。

超链接不仅可指向其他网页，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，用户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个相互关联的页面。这种链接点指向的资源，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，这样可使全球范围的 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。

1.1.3 URL

超链接为用户访问其他网页内容提供了方便。Web 网站与外界是通过统一资源定位符 (Uniform Resource Locator, URL) 进行联系的。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 窗口的地址栏内输入“<http://www.ccert.edu.cn/advisories/index.htm>”之后，浏览器就会在 Web 上搜索该网址，把 CERNET 网络应急响应组的主页（限于篇幅，图略）传送到用户的面前。

URL 是由 4 部分内容组成的，这 4 部分内容分别代表了网络协议、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**”。其中，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反斜杠分隔。WWW 已成为代表 Web 站点类型的标志，凡是以“<http://www>”开头的 URL 都可初步认定是 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 XP/7 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“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，如“/advisories”；后者是指用户访问的网页文件的名称，如“/index.htm”。网页名称需要使用扩展名，否则浏览器可能会认为它是存储资源的文件夹。表 1.1 中列出了 Internet 上常用资源的 URL 的书写格式。

表 1.1 常用网络资源的 URL 格式

Internet 资源	URL 格式
超媒体传输	http://Who/Where/What
超媒体传输（加密）	https://Who/Where/What
FTP(目录)	ftp://Who/Where/
FTP（文件）	ftp://Who/Where/What
流媒体	rtsp:// Who/Where/What; mms:// Who/Where/What
远程登录	telnet://Who

1.1.4 HTML 和 XML

超文本置标语言（Hyper Text Markup Language，HTML）是编制 Web 页面文档的语言。HTML 是标准通用型置标语言（Standard Generalized Markup Language，SGML）的一种。SGML 是定义结构化文本类型和标识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 支持 Object 标记和层叠风格表单（Cascading Style Sheets，CSS）。通过支持 ISO 10646 字符集，支持任何语言所需的标记，为编程者插入对象和脚本提供了方便，并在文件中支持格式项。HTML 最新版本 HTML 5.0 增加了更多样化的 API，提供了嵌入音频、视频、图片的函数、客户端数据存储，以及交互式文档。其他特性包括新的页面元素，如<header>、<section>、<footer>和<figure>。通过制定处理所有 HTML 元素的规则，以及从错误中恢复的规则，HTML 5.0 改进了互操作性，并降低了开发成本。

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，HTML 难以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；难以满足协同工作、信息重用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。

可扩充的置标语言（Extensible Markup Language，XML）是一种元语言（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，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，它提供了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之外。SGML 在大型文档管理、出版业等领域取得了非常好的应用，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，WWW 协会（World Wide Web Consortium，

W3C) 于 1998 年正式公布了 XML 1.0 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 (Schema)。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 标记是定义数据的,用来描述怎样在屏幕上显示信息,解决人机交互的问题,而 XML 是用来描述信息是什么,用来存储数据的,重在数据本身,解决机器之间的交互问题。

1.1.5 Web 2.0

Web 2.0 是相对于 Web 1.0 的新的一类互联网应用的统称。Web 1.0 的主要特点在于用户通过浏览器获取信息。Web 2.0 则注重用户的交互作用,用户既是网站内容的浏览者,也是网站内容的制造者。在模式上由单纯的“读”向“写”及“共同建设”发展,由被动地接收互联网信息向主动创造互联网信息发展。Web 2.0 的特点如下。

(1) 用户参与网站内容制造。Web 2.0 网站的内容通常是用户发布的,Web 2.0 网站为用户提供了更多参与机会。例如,博客 (Blog) 网站和维客 (Wiki, 一种多人协作的写作工具) 就是典型的用户制造内容的工具。Web 2.0 的 Tag 技术 (用户设置标签) 将传统网站中的信息分类工作直接交给用户来完成。

(2) Web 2.0 侧重交互性。不仅用户在发布内容过程中实现与服务器之间交互,而且,也实现了同一网站不同用户之间的交互,以及不同网站之间信息的交互。

(3) 符合 Web 标准的网站设计。Web 标准是目前国际上正在推广的网站标准,通常所说的 Web 标准一般是指网站建设采用基于 XHTML 语言的网站设计语言。实际上,Web 标准并不是某一标准,而是一系列标准的集合。Web 标准中典型的应用模式是“CSS+DIV (层叠样式表+标签)”,摒弃了 HTML 4.0 中的表格定位方式,其优点之一是网站设计代码规范,并且减少了大量代码,减少网络带宽资源浪费,加快了网站访问速度。更重要的一点是,符合 Web 标准的网站在支持用户访问和搜索引擎方面更加友好。

(4) Web 2.0 网站与 Web 1.0 没有绝对的界限。Web 2.0 技术可成为 Web 1.0 网站的工具,一些在 Web 2.0 概念之前诞生的网站本身也具有 Web 2.0 特性。例如,B2B 电子商务网站的免费信息发布和网络社区类网站的内容来源于用户。

(5) Web 2.0 的核心不是技术而在于指导思想。Web 2.0 有一些典型的技术,但技术是为了达到某种目的所采取的手段。Web 2.0 技术本身不是 Web 2.0 网站的核心,重要的在于 Web 2.0 技术体现了具有 Web 2.0 特征的应用模式。因此,与其说 Web 2.0 是互联网技术的创新,不如说是互联网应用指导思想的革命。

1.2 TCP/IP 基本知识

Web 网站是 Internet 的信息资源站点。无论是用户主动访问站点获取信息,还是站点主

动向用户推送信息，Web 站点均需要应用协议（如 HTTP）支撑。应用协议工作需要底层协议 TCP（传输控制协议）和 IP（网际协议）支撑。

1.2.1 网络协议

计算机网络由多个互连的节点组成，节点之间要不断地交换数据和控制信息。要做到有条不紊地交换数据，每个节点必须遵守一些事先约定好的共同规则。为数据交换而制定的规则、约定和标准统称为网络协议（Network Protocol）。

一般，一个网络协议由三个要素构成：语法、语义和时序。语法确定通信双方之间“如何讲”，由逻辑说明构成，还确定通信时采用的数据格式、编码、信号电平及应答结构等。语义确定通信双方之间“讲什么”，由通信过程的说明构成，要对发布请求、执行动作及返回应答予以解释，并确定用于协调和差错处理的控制信息。时序则确定事件的顺序及速度匹配和排序等。

1.2.2 OSI 参考模型与数据封装

IEEE 802 委员会于 1981 年提出了 OSI（开放系统互连，Open System Inter Connection）参考模型。OSI 模型定义了异构计算机（硬件结构、软件指令均不同）连网标准的框架结构，受到计算机和通信行业的极大关注。OSI 模型作为计算机网络体系结构标准，大大推动了计算机网络与通信的发展。

1. IEEE 802 参考模型 OSI

OSI 模型采用三级抽象，即体系结构、服务定义和协议规格说明。体系结构部分定义 OSI 的层次结构、各层间关系及各层可能提供的服务。服务定义部分详细说明了各层所具备的功能。协议规格部分的各种协议精确定义了每一层在通信中发送控制信息及解释信息的过程。提供各种网络服务功能的计算机网络系统是非常复杂的。根据分而治之的原则，OSI 模型将整个通信功能划分为 7 个层次，如图 1.1 所示。

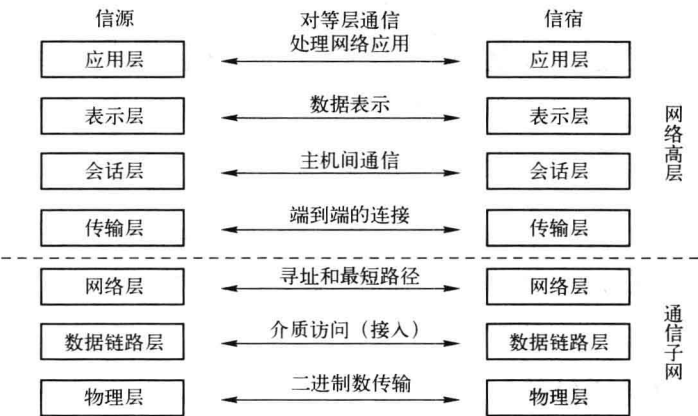


图 1.1 IEEE 802 参考模型 OSI

从总体上看，计算机网络分为“通信子网”和“网络高层”两大层次：通信子网（1～3 层）支持通信接口，提供网络访问；网络高层（4～7 层）支持端端通信，提供网络服务。

无论怎样分层，较低的层次总是为与它紧邻的上层提供服务的。

OSI 参考模型是理论模型。模型的建立有利于将网络通信作业拆解成较小的，也较简单的部分，方便设计制造。将网络元件标准化，使更多的厂商加入开发及技术支持，让各种不同类型的网络硬件与软件彼此互通信息；防止一层中的改变影响到其他各层，便于技术更迅速发展。将网络通信作业拆解成较小的部分，在学习和了解时就更加简单明了。

2. PDU 与数据封装

在 OSI 参考模型中，对等层协议之间交换的信息单元统称为协议数据单元（Protocol Data Unit, PDU）。传输层及以下各层的 PDU 有各自特定的名称：传输层——数据段（Segment），网络层——分组数据报文（Packet），数据链路层——数据帧（Frame），物理层——二进制比特流（Bit）。

一台计算机要发送数据到另一台计算机中，数据首先必须打包，打包的过程称封装。封装就是在数据前面加上特定的协议头部。这如同发送信件一样，信纸装入写有源地址和目的地址的信封中发送，还要写明是“平信”或“挂号信”。

OSI 参考模型中每一层都要依靠下一层提供的服务。为了提供服务，下层把上层的 PDU 作为本层的数据封装，然后加入本层的头部（和尾部）。头部中含有完成数据传输所需的控制信息。这样，数据自上而下递交的过程实际上就是不断封装的过程，到达目的地后自下而上递交的过程就是不断拆封的过程，如图 1.2 所示。

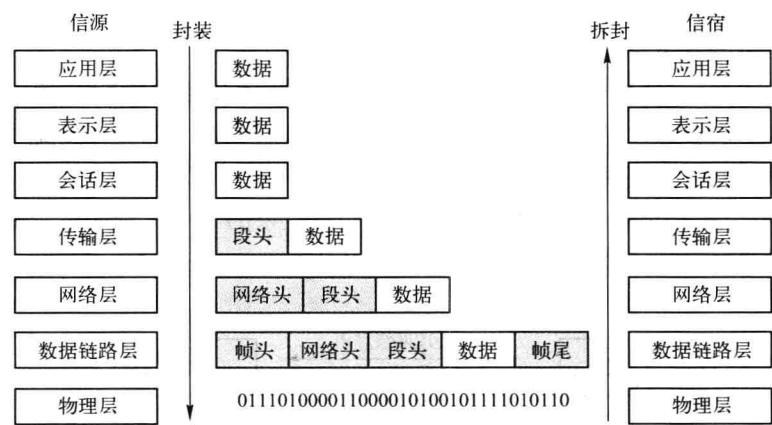


图 1.2 数据多层封装与拆封

由此可知，在物理线路上传输的数据，其外面实际上被包封了多层“信封”。但是，某一层只能识别由对等层封装的“信封”，而对于被封装在“信封”内部的数据，仅仅是拆封后将其提交给上层，本层不做任何处理。

1.2.3 TCP/IP 体系结构

由于种种原因，OSI 模型并没有成为真正应用在工业技术中的网络体系结构。Internet 在全世界的飞速发展使其所遵循的 TCP/IP 参考模型得到了广泛的应用。

1. TCP/IP 协议

TCP/IP 协议是一个协议集，如图 1.3 所示。TCP/IP 协议集中最重要是传输控制协议

(Transmission Control Protocol, TCP) 和网际协议 (Internet Protocol, IP), 通称为 TCP/IP 协议。TCP/IP 协议具有如下 4 个特点。

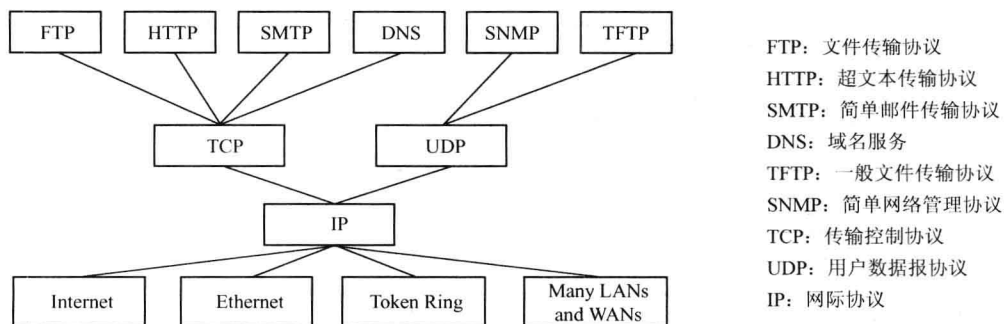


图 1.3 TCP/IP 协议集

- 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 FTP, HTTP, SMTP 等), 如图 1.3 所示, 可以提供多种可靠的信息服务。

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 之上提供了一个可靠的面向连接的协议。TCP 协议能保证数据包传输及正确传输顺序, 并且它可以确认数据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 协议为计算机网络上的数据提供了一个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。最终确认数据包是否到达目的地, 还要依靠 TCP 协议, 因为 TCP 协议是面向连接服务的。

2. TCP/IP 体系结构及功能

TCP/IP 体系结构分为 4 个层次: 网络接口层、IP 层、传输层和应用层。TCP/IP 体系结构与 OSI 参考模型对应关系, 以及 TCP/IP 数据封装与拆封流程, 如图 1.4 所示。TCP/IP 网络接口层 (Network Interface) 对应 OSI 的物理层和数据链路层, TCP/IP 的 IP 层 (包括 ICMP、ARP、RARP 等协议) 对应 OSI 的网络层, TCP/IP 传输层 (TCP、UDP) 对应 OSI 的传输层, TCP/IP 的应用层 (高层协议) 对应 OSI 的会话、表示及应用层。

TCP/IP 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; 或者接收、处理网络上的物理帧, 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, 网络接口层中的比特流传输相当于邮政系统中信件的运送。

(2) 网络互连层。该层是整个体系结构的核心部分, 负责处理互连网络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, 处理接收的数据包和处理互连的路径。

网络互连层 IP 协议提供了无连接 (不可靠) 的数据包传输服务, 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，或者检测到影响数据包的正确传输的异常状况，路由器就要通知信源主机或路由器采取相应的措施。

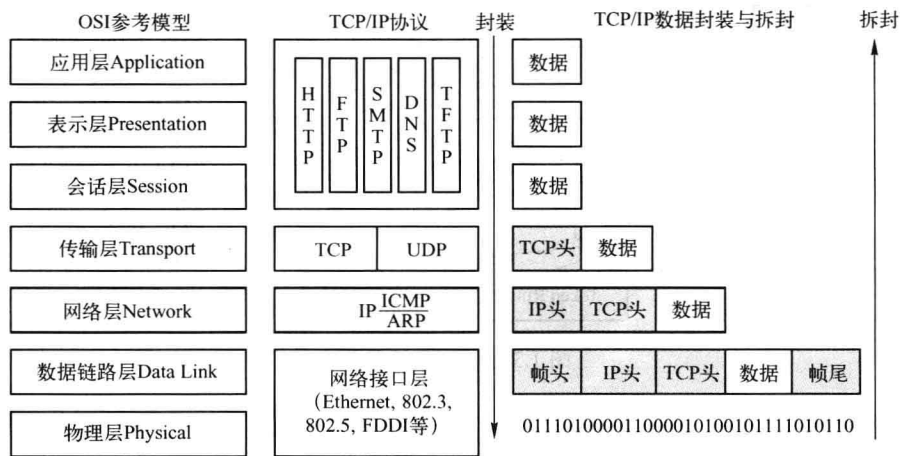


图 1.4 OSI 与 TCP/IP 对比及数据封装与拆封流程

(3) 传输层。该层是整个体系结构的控制部分，负责应用进程之间的端到端通信。传输层定义了两种协议：传输控制协议（Transfer Control Protocol，TCP）与用户数据报协议（User Datagram Protocol，UDP）。TCP 协议是一种可靠的面向连接的协议，允许从一台主机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数据段，并将数据段传输给 IP 层打包。在接收端，IP 层将接收的数据包解开，再由 TCP 层将收到的数据段组装成应用协议字节流。TCP 还可处理流量控制，以避免快速发送方，向低速接收方发送过多数据包而使接收方无法处理。UDP 协议是一种无连接（不可靠）的协议，它与 TCP 协议不同的是它不进行分组顺序的检查和差错控制，而是把这些工作交给上一级应用层完成。

(4) 应用层。该层是整个体系结构的协议部分，它包括了所有的高层协议，并且总是不断有新的协议加入。与 OSI 模型不同的是，在 TCP/IP 模型中没有会话层和表示层。由于在应用中发现，并不是所有的网络服务都需要会话层和表示层的功能，因此，这些功能逐渐被融合到 TCP/IP 协议中应用层的那些特定的网络服务中。应用层是网络操作者的应用接口，就像发信人将信件放进邮筒一样，网络操作者只需在应用程序中按下发送数据按钮，其余的任务都由应用层以下的各层完成。

3. ICMP 与 ARP 的功能

ICMP（Internet Control Messages Protocol）协议封装在 IP 数据包中，通过 IP 协议进行传送。ICMP 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，最常用的是“目标无法到达（Destination unreachable）”和“回声（Echo）”消息。ICMP 工作示意图如图 1.5 所示。

网络互连层的 ARP（Address Resolution Protocol）提供地址转换服务，查找与给定 IP 地址对应主机的物理地址（网卡的 MAC 地址）。与 ARP 功能相反的是 RARP（Reverse ARP），RARP 协议主要将主机物理地址转换为对应的 IP 地址。