

G F G Y C B S

Internet

浏览器指南



罗兰 编著

92

国防工业出版社

TP393.0/2
LL/1

Internet 浏览器指南

罗 兰 编著



国防工业出版社

· 北京 ·

040987

图书在版编目(CIP)数据

Internet 浏览器指南/罗兰编著. —北京:国防工业出版社,1997.9

ISBN 7-118-01757-4

I I ... I. 罗... II. 全球网络:互连网络,Internet-检索系统-指南 N. TP393.4-62

中国版本图书馆 CIP 数据核字(97)第 10878 号

JS263/26

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号)

(邮政编码 100044)

涿中印刷厂印刷

新华书店经售

*

开本 787×1092 1/16 印张 14 323 千字

1997 年 9 月第 1 版 1997 年 9 月北京第 1 次印刷

印数:1—5000 册 · 定价:19.00 元

(本书如有印装错误,我社负责调换)

前 言

90年代计算机领域最热门的技术是网络技术,而网络技术最激动人心的应用是 Internet。Internet 无可争议地是曾经建立的网络中最著名、发展最快的网络,更准确地说,它是一个网络的网络,通过它人们可以通达全世界各个角落。

在当今这个信息高度发展的社会,如何在众多的信息中获取有用的精华,如何正确地利用这些信息,这是每一个现代人所关心的事情,而 Internet 正是步入信息高速公路的捷径,这也是 Internet 成为现在最热门话题的原因。

但是,Internet 上的信息中,既有璀璨的明珠,也有腐臭的垃圾,而且,如果不熟悉其检索服务系统,你在网上看到的只是大量的一时不知所云的各种文档。为了帮助读者最大限度的获取 Internet 上的有用信息,我们编写了这本专门介绍现在国内外十分流行的 Internet 信息检索软件——WWW 浏览软件的读物,从最基本的软硬件设置到较高级的软件功能扩展,都有详细的介绍,力图使没有接触过 Internet 的用户,在阅读本书之后,也可成为 Internet 的高手。

本书包括四部分,第一部分对 WWW 作了由浅入深地介绍,第二、三、四部分分别介绍了三个 WWW 浏览软件:Netscape Navigator、Microsoft Internet Explorer 和 Internet In a Box。

目 录

第一部分 五彩纷呈的 WWW 世界

第 1 章 什么是 WWW	1
1.1 WWW 的定义	1
1.2 WWW 的历史	2
第 2 章 如何接入 WWW	3
2.1 如何进入 Internet	3
2.2 选择一个合适的浏览器	4
2.3 第一次进入 WWW	4
2.4 深入了解 WWW	5
第 3 章 WWW 专家必读	7
3.1 有关 WWW 的闲聊	7
3.2 HTTP	8
3.3 CGI	10
3.4 Java	13

第二部分 Netscape Navigator

第 4 章 Netscape 主窗口	15
4.1 总览	15
4.2 File 菜单	22
4.3 Edit 菜单	25
4.4 View 菜单	26
4.5 Go 菜单	29
4.6 Bookmarks 菜单	30
4.7 Options 菜单	30
4.8 Directory 菜单	33
4.9 Window 菜单	34
4.10 Help 菜单	35
4.11 弹出式菜单	36
第 5 章 预置	39
5.1 General Preferences	39
5.2 Mail and News Preferences	47
5.3 Network Preferences	53
5.4 Security Preferences	57

第 6 章	电子邮件,新闻组和书签	59
6.1	Internet 工具	59
6.2	Message Composition 窗口	59
6.3	Mail 窗口	62
6.4	News 窗口	72
6.5	书签列表	78
第 7 章	疑问和解答	84

第三部分 Microsoft Internet Explorer

第 8 章	Microsoft Internet Explorer 概览	106
8.1	什么是 Microsoft Internet Explorer	106
8.2	使用 Internet Explorer	107
8.3	Internet Explorer 的特色	107
第 9 章	如何使用 Internet Explorer	109
9.1	畅游 Internet 的广阔天地	109
9.2	使用页中信息	110
9.3	编页	111
9.4	更改 Internet Explorer 外观	113
9.5	提高查看页时性能	114
9.6	阅读 Internet 新闻组	115
9.7	查看某一页的源文件	116
9.8	打开保存在计算机上的文件	116
9.9	在打开特定类型的文件时启动程序	116
9.10	更改历史记录中的页数	117
9.11	设置 Internet 安全警告的级别	117
9.12	使用其他程序阅读 Internet 新闻组	117
9.13	配置计算机以便用代理程序访问 Internet	117
9.14	保存连接的页和图片	118
9.15	使用电子邮件发送 Internet 上的快捷方式	118

第四部分 SPRY Internet In a Box

第 10 章	SPRY Dialer	119
10.1	何谓 SPRY 拨号器?	119
10.2	如何使用 SPRY 拨号器?	119
10.3	SPRY 拨号器参考大全	121
第 11 章	SPRY Mosaic	143
11.1	启动 SPRY Mosaic	143
11.2	利用 Mosaic 进行浏览	144
11.3	使用热表(Hotlist)	148
11.4	打开 Mosaic 文档	154

11.5	在 Mosaic 中使用 FTP, Gopher, News, 和 Mail	156
11.6	设置 Viewers(观察器)	160
11.7	存储 Mosaic 文档	164
11.8	使用 SPRY Mail 邮寄 Mosaic URLs 文件	166
11.9	在 Mosaic 中打印	166
11.10	在 Mosaic 中查找信息和拷贝信息	167
11.11	设置 Mosaic 选项	167
11.12	查看 Mosaic 文档源码(HTML 码)	173
11.13	退出 Mosaic	174
11.14	Mosaic 安全选项命令	174
11.15	Mosaic 命令行选项	182
第 12 章	SPRY 电子邮件	184
12.1	SPRY 电子邮件简介	184
12.2	使用 SPRY 电子邮件	186
12.3	SPRY 电子邮件参考大全	205

第一部分 五彩缤纷的 WWW 世界

第 1 章 什么是 WWW

1.1 WWW 的定义

WWW(World Wide Web),是一张附着在 Internet 上覆盖全球的信息蛛网,镶嵌着无数以超文本形式存在的信息,其中有璀璨的明珠,也有腐臭的垃圾。有人叫它全球网,也有人叫它万维网,或者就叫它 Web。

WWW 是当前 Internet 上最受欢迎、最为流行、最新的信息检索服务系统。它把 Internet 上现有资源统统连接起来,使用户能在 Internet 上已经建立了 WWW 服务器的所有站点(Site)提供超文本超媒体资源文档。这是因为,WWW 能把各种类型的信息(静止图像、文本声音和音像)天衣无缝地集成起来。WWW 不仅提供了图形界面式的快速信息查找,还可以通过同样的图形用户界面(GUI)与 Internet 上其他服务器对接。

由于 WWW 为全世界的人们提供查找和共享信息的手段,所以也可以把它看作是世界上各种组织机构、科研机关、大专院校、公司厂商甚至个人热衷于研究开发共享的信息集合。它基于 Internet 的查询、信息分布和管理系统,是人们进行交互的多媒体通信动态格式。它的正式提法是:“一种广域超媒体信息检索原始规约,目的是访问巨量的文档”。WWW 已经实现的部分是,给计算机网络上的用户提供一种兼容的手段,以简单的方式去访问各种媒体。它是第一个真正的全球性超媒体网络,改变了人们观察和创建信息的方法。当前,世界上掀起了研究、开放、使用 WWW 的巨大热潮。

WWW 诞生于 Internet 之中,后来成为 Internet 的一部分,而今天,WWW 几乎成了 Internet 的代名词。通过它,加入其中的每个人能够在瞬间抵达世界的各个角落。只要将一根电话线插入你的 PC(它可能是你随身携带的笔记本电脑加上一部移动式电话),谁能相信,此时全球的信息就在你的指尖!

WWW 并不实际存在于世界的任何地方,事实上,WWW 的使用者每天都赋予它新的含义,因而 WWW 将不断扮演新的角色,其中之一就是作为全球最大的超文本出版机构,这是一件具有互联页面和图形的庞大盛装。

Internet 社会的公民们(包括机构和个人),把他们需要公之于众的各类信息以主页(Home Page)的方式嵌入 WWW,主页中除了文本还包括图形、声音和其他媒体形式;至于内容,从招聘广告到电子版圣经,可以说无所不有。

主页是在 Web 上出版的主要形式。主页是一些 HTML 文本(HTML 指 Hyper Text Markup Language,即超文本置标语言),这些分布于全球的超文本以超连接方式集成一

体,阅读者从一个主页跳跃到另一个主页时,也许他不会意识到:他已从华盛顿大学的一台服务器迁移到位于澳大利亚的某一台服务器。简言之,在 Web 上的出版无非是用 HTML 进行一些格式化操作形成主页;阅读其中的一份出版物只是在 Web 浏览器中用鼠标指点相应文档,剩下的工作则由 Web 浏览器包揽了。

1.2 WWW 的历史

WWW 的历史非常短,至今不过五六年时间。

起初,欧洲高能物理实验所 CERN 的一些科学家想寻求一种办法,让分散在世界各地的研究小组能够很快地共享最新的研究信息。由于 CERN 的研究活动有很多都是在彼此相距很远的异地进行的,要想进行一次简单的学术活动(例如查询异地的文件或图片)通常需要同时运行很多不同的软件才能实现,例如需要运行 Telnet、FTP 以及图像阅读软件等。科学家们需要的是一种使用简便、具有统一的界面、能够很快获得信息的系统。

为此,1989 年 3 月,CERN 的科学家 Tim Berners-Lee 首先提出了 WWW 这一新概念,并将它作为高能物理学界科学家传输新想法、新成果的工具。到 1990 年末,第一个 WWW 软件在 NeXT 计算机上实现了。该软件能让用户在 Internet 上查询、传输超文本文档,并具有编辑超文本文档的功能。1992 年,CERN 公开发表了 WWW。

1993 年,Illinois 大学的国家超级计算应用中心(National Center Supercomputing Applications)发布了第一个容易使用的 Web 研究软件包 Mosaic 后,WWW 就迅猛发展起来。到了 1993 年末,针对各种不同类型的计算机系统(例如 X-Windows、PC/Windows、Apple Macintosh)的浏览程序(Browser)相继开发出来了。1994 年夏天,WWW 已成为访问 Internet 资源最流行的手段。现在 WWW 服务器正以每年 2000% 的速度增长。我国于 1994 年春已正式建立了 Internet 上的 WWW 服务器,主要集中在北京。

到今天,许多公司、大学和各种各样的个人用户已经意识到:使用相对简单的超文本置标语言(HTML)命令在 Web 上发布信息是多么容易!每天都有成百上千的主页出现在联机上,这些主页包括从国际贸易法到来自俄勒冈州的儿童指画法作品,甚至各种型号的轰炸机背景资料。虽然深度有时不尽人意,但内容之广泛确实令人吃惊。

目前全球已有数百万活跃的 Web 用户,2000 个商务网点及无数的教育和政府网点。一个 Internet 网点如果没有 Web 服务,现在简直是一件不可思议的事情。当你从一个服务器到另一个服务器、从一个主页到另一个主页作超连接跳跃时,足不出户就可以作环球旅行,所到之处很容易使你感到惊奇甚至欣喜若狂。而经过一番努力达到某些网点后,却没有发现多少有价值的东西,有些内容非常肤浅,有些信息令人置疑,这会令你感到失望。

为什么 WWW 如此流行呢?这是因为对于过去一直难于掌握的 Internet 资源,WWW 提供了一种非常容易使用的界面,吸引了越来越多的用户去创建和开发难以计数的各种应用程序,使网络超文本和超媒体的巨大潜力迅速被挖掘出来。

第 2 章 如何接入 WWW

2.1 如何进入 Internet

我们说过 WWW 是附着在 Internet 上的蛛网,要进入 WWW 世界就得通过 Internet 大门。

2.1.1 连接 Internet 的三种方法

- 将电脑连接到一个局域网,这个局域网的服务器就是 Internet 的一个宿主机。
- 利用 SLIP/PPP 通过电话拨号方式进入一个 Internet 宿主机。
- 通过电话拨号方式进入一个提供 Internet 服务的联机服务系统。

前两种方法是直接连接,因为你在自己的电脑上运行 TCP/IP(Internet 使用的网络协议);第三种方法是间接连接:联机服务系统直接连接 Internet,而你的电脑利用终端仿真软件去访问由联机服务为宿主机提供的服务。

直接和间接连接之间的差别比你的想象具有更深远的含义。你所采用的连接类型将决定你在 Internet 的工作类型。尽管联机服务已增加了一些帮助你与 Internet 交互的功能,但大多数人还是只使用了传递电子邮件业务。Delphi 和 America Online 是两个例外,这两者提供了对 Internet 最好的访问电子邮件目录、新闻组(Newsgroup)以及 Gopher 服务,当然,你只能使用字符方式。

要进入五彩纷呈的 WWW 世界,你必须采用直接连接方式。

如果你能连接到一个与 Internet 连接的网络,可以向你的网络管理员进行咨询。对大多数人而言,也许我们对第二种方式最感兴趣:如果我拥有一台 PC 机和一条电话线的话。

2.1.2 用电话拨号方式(SLIP/PPP)进入 Internet

现在,我们还需要一个调制解调器(MODEM)。

我们能够从以下三个方案中选择其一配置调制解调器:

- 通过电脑的串行口连接一个调制解调器。
- 内置于电脑的调制解调器扩展卡。
- 如果你有一台笔记本电脑,也可以选择具有调制解调功能的 PCMCIA 卡。

第一种和第三种方案更容易掌握,你可以根据自己的情况选择;第二种方案非常便宜,如果你是一个电脑行家,并且热衷“中断”、“I/O 地址”之类的调整,倒也无妨一试。

对调制解调器而言,通信速率是人们经常提及的指标。一般而言,设置调制解调器的速率取决于你所使用的电话线的质量,线路干扰越小,通信速率就可以设置得越高。通常我们使用 9600 波特或 14400 波特的通信速率。

当然,出售调制解调器的厂家应当为你做好以上的调试,他们也许会拨一个传真的号码并发送一个传真,以向你证明它确实能够工作。

下面我们需要找一个 Internet 服务提供商,你可以到电信部门咨询有关事项,申请开户并办理有关手续。然后我们就可以加入 CHINANET 了(我们是通过它访问 Internet 的),然后使用你所选择的 Internet 连接软件拨号(如果你在北京,你可以拨 6011180 或 163),在一阵令人心动的载波歌唱之后,你将敲开 Internet 之门。

接着,我们将看到要求注册的信息,输入帐号和密码之后(你的服务提供商为你开户后将告诉你这一切,不过你不会告诉别人的,对吧?),如果没有问题,屏幕上会有一段服务提供商的公告信息,要仔细阅读,尤其是第一次进网。一般,公告信息会告诉你如何访问邮件服务器、新闻服务器等等,还会告诉你如何启动 SLIP/PPP 会话。为了进入 WWW 世界,这时我会敲入“PPP”启动并进入 PPP 会话,念完这一段“芝麻”咒语之后,WWW 大门终于敞开了。

2.2 选择一个合适的浏览器

Web 浏览器?对,它是你阅读 Web 出版物的工具,发掘 Web 中的宝藏需要一个称为浏览器的特殊软件工具。

Internet 就像一个遍布世界的信息高速公路网,WWW 站点就像公路旁边的加油站。WWW 浏览器就是奔驰在高速公路上的快车,而你则是快车的驾驶者。快车瞬间就可以把你从地球的一端带到另一端,速度几乎是无限的,但它绝对安全,虽然你偶尔也会碰到警察(指那些网络管理员),只要你不是有意挑衅,一般他们是不理会你的种种冒失的。

浏览器知道如何解释和显示在 Internet 上找到的超文本文档。这是由于超文本文档包含着一些借用标题、章节本身等构造文本的命令,从而允许浏览器格式化每一种文本类型,以获得最佳的文本显示效果。例如在 PC 机上可以采用浏览程序 Mosaic for Windows。此外,浏览器还可以访问超媒体,只要在 PC 机上装有声卡和驱动软件,就能听见包含在 WWW 超媒体里的声音片断。

一般浏览器还可以自动调用其他应用程序,以显示特殊类型的文档。例如,如果 WWW 文档包括对 Word for Windows 格式书写的文档访问时,浏览器会自动调用 Word for Windows。有的浏览器还具备访问 Internet 上其他服务器的功能,如匿名 FTP,GO-PHER,WAIS 以及 USENET NEWS 等。

在本书中,我们将向您介绍三种流行的浏览器:Netscape Navigator,Spray Mosaic 以及 Windows 95 中的 Internet Explorer。您可以根据自己的偏好选择一种使用。

2.3 第一次进入 WWW

无论你使用哪一种浏览器,当你按我们前面的方法进入 WWW 世界之后,浏览器一般会把您带到浏览器制造商的家门。例如 Netscape Navigator 会领您进入 Netscape,Internet Explorer 把您带到 Microsoft,不管这是否出于制造商的私心,这些浏览器总要先回一趟自己的出生地。

当然,你一定会去自己想去的方,最好你能知道那儿的地址,就像拜访朋友前一定要知道他的门牌号码(在 WWW 世界里,我们叫它 URL 地址,下一节我们会告诉你什么是 URL)。

如果你知道这个地址,把它用键盘告诉浏览器,然后浏览器会带你造访那里。顺便安慰你一句,你不必告诉它该走哪条路线,这比在真正的高速公路上开车要省心多了。

如果不知道地址该怎么办哪? 这本书后面会告诉你如何去查到这些地址。

如果你想了解关于某个专题的全部资料,最好去访问一些特殊的 WWW 站点,后面我们会介绍一些这样的站点。

可能你会觉得这些方法还不算很方便,那有没有更好的方法呢? 科学家正在开发一种 Internet 上的知识机器人 KNOWBOT。它是一种查找 WWW 资源的计算机程序,只要你确切地告诉它想要找什么,它就可以自动到 Internet 上游历,去找到满足要求的信息和资源。例如对于商业购物,KNOWBOT 可以作出商品的价格比较,它首先在 Internet 上找到提供指定商品的很多家供应商,积累该商品的详细信息,检查价格,然后返回报告,协助买主作出购物决定。

2.4 深入了解 WWW

这里,我们将进一步介绍有关 WWW 的重要概念,如果你急于想了解某种浏览器的使用方法,可以暂时跳过这部分内容。不过,要想成为信息高速公路上的出色车手,这些知识还是必须知道的。

2.4.1 客户机和服务器

在讨论 WWW 时经常听到两个术语:客户机和服务器。WWW 的客户机是指在 Internet 的一个站点请求 WWW 文档的用户计算机;WWW 服务器是指在 Internet 上保存 WWW 信息的计算机,它利用超文本传输协议 HTTP(hyper text transport protocol)允许用户在客户机上发出请求,访问超文本和超媒体信息。

客户机软件是一种程序,用户可以用它查找 WWW 文档。服务器软件则是管理 Internet 上某一宿主机上一组特定文档的软件。

2.4.2 超文本和超媒体

WWW 主要依赖超文本作为它与用户进行交互的手段。超文本专指基于计算机的文档。用户在阅读这种文档时,从其中一个地点移向另一个地点,从其中一个文档移向另一个文档,都是按跳跃方式(或联想方式)进行的。用户不是从头到尾顺章逐节的传统方式去获取信息,而是可以在文档里随机地跳来跳去。这是由于超文本中包含着可用作连接的(LINK)的词或短语(一般用下划线或不同颜色标明)或图标,用户只需用鼠标在其上轻轻一点,就能立即跳至与当前正在阅读的文档相关的新地点和新文档。

超媒体是超文本的自然扩展,是超文本与多媒体的简单组合。在超媒体里的连接不只是连接到文本文档,还可以连接到其他形式的媒体,如图形图像、声音、影视动画等。这样,超媒体就把死板的文本文档变成了活生生的文档,比音响设备、电视机更加生动有趣,更

具有吸引力。

2.4.3 统一资源定位器 URL(Uniform Resource Locator)

设计 WWW 的目的之一是寻求访问 Internet 各种流行文档的标准方法。为此,开发了一种工具叫做统一资源定位器 URL。URL 完整地描述了 Internet 超文本文档的地址。这个地址可以是本地磁盘,也可以是 Internet 上的站点。地址的访问可以是相对的,也可以是绝对的。在相对方式下,假定主机名和路径名就是当前正在使用的名字,只需指出子命令名和文件名即可。绝对方式应包括完整的主机名、路径名和文件名。URL 不仅可以描述 WWW 文档的地址,还可以描述其他服务器(匿名 FTP, GOPHER、WAIS、TELNET、USENET NEWS)的地址。典型的 URL 地址格式如下: `http://www.net.edu.cn/inet/index.html`。其中,“http”代表用于检索文档的协议,即超文本传输协议,它是专门为 WWW 开发的一组规约说明。“//”表明其后紧跟的是 Internet 有效主机名。第一个“/”之后紧跟的是用户要查找文档的 UNIX 风格路径名和文件名。上面 URL 地址的含义是:利用 HTTP 协议,在 Internet 宿主机 `www.net.edu.cn` 上的 `inet` 目录下查找文档 `index.html`。顺便说明一句,这台主机是清华大学的 WWW 站点。

2.4.4 超文本传输协议

设计 WWW 的另外一个目的是能方便地访问 Internet 上的文档,而不管这些文档是在什么地方。决定使用超文本作为 WWW 的标准格式后,科学家立即制定了快速查找这些文档的超文本文档的协议 HTTP,这是一个非常简单的协议。它利用了这样一个事实:该协议所检索的文档包含了用户可以进一步查找的连接。在查找 WWW 文档时,没有必要知道 HTTP 的具体内容。但是如果有兴趣的话,可以在下述地址找到技术说明副本:

`http://info.cern.ch/hypertext/www/Protocols/HTTP/HTTP2.html`

2.4.5 主页(Home Page)

当启动 Internet 上某个 URL 地址上的文档时,首先要显示的文档叫做主页。使用 WWW 的每个用户都可以建立自己的主页。在该文档中,可以加入表征用户特点的图形或图像,列出最常用的一些连接。对于经常使用相同资源的集体或单位,如一家公司、研究所或大学,都可以设计自己的主页。

2.4.6 超文本标记语言 HTML

查阅 WWW 文档时,在屏幕上显示的都是非常漂亮的格式化文本。之所以能做到这一点,就是因为书写这些文档时,利用了超文本标记语言 HTML。HTML 是一组非常简单的命令。这组命令描述了 WWW 文档是如何构造的。它只是定义文档的各个部分,并未真正地进行文档格式化,文档格式化是由浏览器完成的。

第3章 WWW 专家必读

如果你还不打算涉及更为深入和广泛的技术问题,可以跳过本章,以后有兴趣时再读。

3.1 有关 WWW 的闲聊

除了 CERN 的 Tim Berners-Lee 等人的巨大贡献之外,对 WWW 的迅速崛起起了关键作用的是 Internet 已扩展成了全球性的通信网络,它为连接世界各地的计算机资源的应用程序提供了一个巨大的舞台。除此之外,还有两类软件的开发亦功不可没。首先是交互式超文本、超媒体软件的开发,使那些不是计算机专业的人们也能够编写多媒体软件,创建不同类型文档和媒体间的连接。其次是一大批软件界面程序的开发。为用户方便地巡视 WWW 上的各种文档及其上的连接提供了帮助。Internet 发展到现在,在功能上处于一种无政府状态。因为大批的网络都成功地连到它上面去了,但却没有指定也不可能指定谁是 Internet 的拥有者,也未指定谁是它的监督机构。WWW 的情况也是一样,世界上没有一个单独的个人、公司或机构能够声称他拥有 WWW。这是由于 WWW 是拥有成千上万用户的分布式系统,千百万的文档和多媒体作者在为 WWW 提供信息。这些信息提供者包括博士、大中院校的学生、市场营销人员、律师、音乐家等各阶层的人们。这么多人在给这个电子仓库贡献信息,很快就使这个仓库囊括了浩瀚的信息海水,其范围遍及日常购物、娱乐、教育、高技术、天文地理、金融业务等。特别是多媒体技术进入这个巨大的网络之后,更普及、更容易访问的知识数据库开始形成,这就产生了一个问题:如何有效地巡视这个巨大的数据库呢?答案只有一个:开发软件界面工具,即浏览程序。

第一批开发出来的浏览程序都是基于文本的,用户只要敲入适当的键盘命令就可以跟踪 WWW 上的连接。但它却不能让用户访问 WWW 的最精彩部分——多媒体。NCSA 的科学家们在美国国家科学基金会的资助下,于 1991 年初编写了一个名为 Mosaic 的浏览程序,它已经成了 WWW 图形界面的事实上的标准。由于它是免费使用的,在 1993 年 9 月公布之后的半年时间里,就有超过 200 万的人从 NCSA 主机上拷贝下了 Mosaic。前面提及,世界上没有任何单独的组织机构和个人能够拥有 Internet 和 WWW。但却有一个 WWW 组织 W3O 在管理和控制 Internet 和 WWW。W3O 决定如何进一步开发这个系统,使之满足新用户的要求。它要保证 WWW 根本目的的实现:提供这样一个系统,任何人可以利用任何类型的计算机去访问 WWW 上的资源。

在 W3O 创建之前,CERN 是唯一在定义 WWW 的进程中起了重大作用的组织。现在,W3O 除了 CERN 之外,还有一个基本成员单位,即美国麻省理工学院的计算机科学实验室,它把几个研究小组的工作跟 WWW 的开发紧密地联系在一起。这些小组有媒体实验室和人工智能实验室等。虽然,目前 W3O 以这两个单位为基地,但是,它打算和其他

研究所和站点磋商,请他们提出优秀的计算机中心和当地的联系点。在欧洲 W3O 的目标是在每一个国家都建立一个联系点。W3O 最初的任务是为 WWW 开发新的国际信息数据通道和建立检索标准,使人们更容易找到想找到的信息。由于 WWW 现在更像一个巨大的电子信息数据库,那么分类的标准方法要使 WWW 上的信息的可访问性得到改善。根据现有标准提出的新标准要给浏览程序的开发人员提供共同的结构,以保证各公司开发的软件互相兼容。关于 WWW 的全球商业应用,W3O 将建议有关“资金的电子安全保密传输”的研究项目。该项目的经费来自美国政府、欧洲联盟和各跨国公司。尽管目前的 W3O 还不是制定标准的组织,但它正和研究开发人员共同合作,密切注视正在出现的热点技术。目前,它所关注的技术领域包括:名字和地址的语法语义——通用标识符、网络传输协议 HTTP、超文本和超媒体数据格式 HTML、压缩和安全保密方法、记帐填表和合法文件包传输协议、增强 WWW 功能的协议以及其他高速网络技术的应用等。

另一方面 WWW 也不是完美无缺的。首先多媒体信息——图片、声音和影视需要大量的数字空间开销。如果使用传输速率低于 14400bps 的调制解调器,把多媒体信息从宿主机传输到客户机,所用的等待时间是难以忍受的。另外,WWW 上的信息在定期更新前可能是过时的。

不过,对于能够访问全人类知识库的巨大乐趣来说,这些问题是微不足道的。可以毫不夸张地说,全世界各行各业的人们都正在参与或正要参与 WWW 的开发和为 WWW 提供信息资源。

下面是几个主要的 WWW 开发者及其 URL 地址:

- NCSA 的 URL 地址: <http://www.ncsa.uiuc.edu>
- CERN <http://info.cern.ch/hypertext/www/TheProject.html>

以后的几节我们将涉及目前关于 WWW 的极为时髦的话题。

3.2 HTTP

HTTP 是为分布式超媒体系统设计的一个协议。它是无状态、面向对象的协议。HTTP 一般用于名字服务器和对象管理。由于 HTTP 1.0 能够满足 WWW 系统客户与服务器通信的需要,从而成为 WWW 发布信息的主要协议。

3.2.1 HTTP 的主要特点

1. 客户、服务器模式

HTTP 支持客户(浏览器)与服务器间的通信,相互传送数据。一个服务器可以为分布在世界各地的许多客户服务。HTTP 定义的事务处理由四部分组成:

- 客户与服务器建立连接。
- 客户向服务器提出请求。
- 如果请求被接收,则服务器送回应答,在应答中包括状态码和所要的文件。
- 客户与服务器断开连接。

2. 简单

HTTP 既简单又能有效处理大量的请求。在客户与服务器连接后,客户需要传送的

信息只是请求方法和路径。HTTP 规范说明了定义的几种请求的方法,实际上常用的只是其中的三种 GET、HEAD 和 POST。每种方法规定客户与服务器联系的类型不同。正是因为 HTTP 简单,使得 HTTP 服务器程序规模小而简单。这样做的直接效果是时间开销很小,通信速度很快。

3. 灵活

HTTP 允许传输任意类型的数据对象。Content-Type 标识正在传输的数据类型。如果把信息看成藏在“罐头”中的食品,Content-Type 就是贴在罐头上的标签。

4. 无连接

HTTP 是一个无连接协议,每次连接只处理一个请求。客户与服务器连接后提交一个请求,客户接到应答后立即断开连接。使用这种无连接协议,在没有请求提出时,服务器不会在那里闲等着,更不会在完成一个请求后还把着原来的请求不放。使用无连接协议就好像写信,一旦写好信发出去就没事了。对方回信有了新信息,再写另一封信。而保持连接协议更像打电话,双方轮流说许多话后才挂断。对话期间电话线一直被占用。对于无连接协议而言,服务器一方实现起来比较容易。

5. 无状态

HTTP 是无状态协议,这既是优点也是缺点。一方面,由于没有状态,协议对事务处理没有记忆能力。如果后续事务处理需要前面事务处理的有关信息,那么这些信息必须在协议外保存。缺少状态意味着所需要的前面信息必须重现,势必导致每次连接都要传送较多信息。另一方面,也正是由于缺少状态使得 HTTP 累赘少,运行速度高,服务器响应速度快。

6. 元信息

HTTP 1.0 对所有事务处理都加头,即在主要信息前加一块信息,我们称之为元信息,即信息的信息。它使服务器能够提供正在传输数据的有关信息,例如,传送对象是哪种类型,是用哪种语言书写的等等。人们还可以利用元信息进行有条件的请求,或者报告一次事务是否成功等。

HTTP 中有两种提供元信息的方法:一是当服务器回答客户请求时,服务器把元信息作为回答的一部分;二是客户一方随着请求把元信息一同送给服务器,帮助服务器满足客户的请求。当然这些信息块是可选的。

从功能上分,HTTP 支持四类元信息:一般信息头、请求头、应答头和实体头。

3.2.2 客户机与服务器的信息交换

在服务器运行时一直在端口 80 倾听,等待连接的出现。打开一个连接就像拿起电话拨某人的电话号码一样。从技术上讲,客户打开一个套接字(Socket),并把它约束在一个端口上。套接字是一个能进行网络输入输出的特殊文件类型。从客户的观点讲,打开一个套接字就是建立一个虚拟文件。做完这些事就是打开一次连接。当向文件上写完数据后,就是把数据经由网络向外传输。

打开连接后,客户把请求数据行送到服务器驻留的端口上,完成“提出请求”动作。HTTP 1.0 版的请求由数行构成,其中第一行是请求,它包括方法、URL 和协议版本号,其余行依次是一般头、请求头和实体头。由于这时已经与 HTTP 服务器建立了连接,故在

第一行 URL 中不再包括协议名、主机名和端口号。

请求中的方法描述的是在指定资源上应该执行的动作。HTTP 定义了包括 GRT、HEAD 和 POST 在内的 7 种方法。

1. GET

GET 是取回由 URL 指定的资源。如果对象是文件,则 GET 取的是文件内容;如果对象是程序或描述,则 GET 取的是该程序执行的结果,或该描述的输出;若对象是对数据库的查询,则 GET 取的是这次查询的结果。

2. HEAD

要求查询对象的元信息而非对象本身,例如对象大小或最后更改时间。这类请求执行很快。

3. POST

从客户向服务器传送数据,要求服务器和 CGI 程序作进一步处理时会用到 POST。主要用于发送 HTML 的 FORM 内容,让 CGI 程序处理。这时 FORM 的内容和 URL 编码随请求一起送出。在请求头中还要给出 Content-Type 和 Content-Length。服务器根据内容的类型和长度信息动态分配空间,并取到 FORM 的内容。

请求头是要告诉服务器怎样解释本次请求。他们主要有:用户可以接受的数据类型、压缩方法和语言等。服务器应答是对客户请求的回答。

客户和服务器任何一方关闭连接时,不管相应的事务处理成功与否,完成与否,都必须关闭连接。

3.3 CGI

环球网 WWW 上的绝大多数信息都是静止的。信息的内容只是在网点管理人员更动后才能发生变化。这种静态页面最大的问题是缺少交互性,能取到的信息完全由信息提供者决定,而用户完全是被动的。目前不少 WWW 网点已经能够让信息访问达到半自动化水平。用户可以将 WWW 看成交互性媒体,利用在线购物、访问数据库等方式去主动寻找自己感兴趣的信息。这种交互性提高了用户主动参与的积极性。WWW 的交互功能正是由公共网关界面 CGI(Common Gateway Interface)实现的。遵循 CGI 标准编写的服务器侧的可执行程序(CGI 程序)让 HTML 文件在客户机和服务器之间有了更多的交互,使信息网关、反馈机制、访问数据库、订货和查询等一系列灵活复杂的操作得以实现。

3.3.1 CGI 的基本概念

早期很多著名的服务器都以自己独特的方式,支持服务器侧的可执行程序,用来帮助完成客户机的请求。为某个服务器写的程序要在其他服务器上使用时,就必须做较大的修改,因为每个服务器与可执行程序之间传送信息的方式均不相同。为此就形成了一个公共网关界面 CGI,使得为一个服务器写的程序能够在任何服务器上运行。通过 CGI,服务器可以向 CGI 程序发送信息,CGI 程序也可以向服务器程序回送信息。CGI 程序可以用 Csh、Perl、C、C++、FORTRAN 和数据库语言等任何能够生成可执行程序的语言编写。