

Web 编程丛书

CGI 编程指南

黄光奇 等编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
URL: <http://www.phei.com.cn>

Web 编程丛书

CGI 编程指南

黄光奇 等编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书首先介绍 CGI 基本知识,包括 CGI 简介、编程初步、运行准备,然后通过两个实例“用户意见簿(UserAdvice)”和“网络考试(NetworkExam)”来讲解 CGI 编程应用,最后讲解了 CGI 编程的高级话题:数据库访问、搜索引擎、SSI、调试 CGI 程序和提高 CGI 程序的安全性。

本书内容深入浅出,叙述简洁,代码详尽,适合所有 Web 开发人员和管理人员参考。

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,翻版必究。

图书在版编目(CIP)数据

CGI 编程指南/黄光奇编著. - 北京:电子工业出版社,1999.1
(Web 编程丛书)
ISBN 7-5053-5192-3

I. C… II. 黄… III. 因特网-接口, CGI-程序设计 IV. TP393

中国版本图书馆 CIP 数据核字(98)第 32021 号

丛 书 名: Web 编程丛书

书 名: CGI 编程指南

编 著 者: 黄光奇 等

责任编辑: 秦 梅

特约编辑: 郝 刚

排版制作: 电子工业出版社计算机排版室

印 刷 者: 北京兴华印刷厂

装 订 者: 三河市双峰装订厂

出版发行: 电子工业出版社 URL: <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×1092 1/16 印张: 19 字数: 486.4 千字

版 次: 1999 年 6 月第 1 版 1999 年 6 月第 1 次印刷

书 号: ISBN 7-5053-5192-3
TP·2580

定 价: 32.00 元

凡购买电子工业出版社的图书,如有缺页、倒页、脱页、所附磁盘或光盘有问题者,请向购买书店调换。

若书店售缺,请与本社发行部联系调换。电话 68279077

《Web 编程丛书》序

目前,许多公司和机构都在建立可供对外访问的 Internet 站点以及仅供公司内部使用的 Intranet。而且,这种势头正在迅猛增长。这样,对于建立 Web 站点所需要的各种软件技术,包括程序设计语言、开发和设计工具,则是各类 Web 站点开发和设计人员急切了解的热门主题。本套丛书正是基于这种市场需求,组织有关专业技术人员编写而成。

本套丛书大致分为两类:程序设计语言和开发工具。程序设计语言包括了 HTML 4.0、JavaScript 和 CGI; Web 站点开发和设计工具包括 FrontPage 98、Visual InterDev 1.0、IntraBuilder 1.51 和简易 Intranet。下面分别介绍各个主题,读者可从中把握各种技术、各书的主要内容以及各书之间的相互关系,从而决定自己的选择。

《HTML 编程指南》

HTML(超文本标志语言)是 Web 页面描述语言,不管是在 Web 上发布信息,或者编写可供用户交互的程序,都需通过 HTML 语言来实现。

本书详细介绍了 HTML 4.0 规范中的每一个元素及其属性,并给出了相应的编程实例。

由于 HTML 是 Intranet 的基础,本套丛书中的其他主题不可避免要涉及到 HTML 的运用。我们在策划本套丛书时,为了保持丛书的一致性,各主题没有或仅用了非常有限的篇幅介绍 HTML。但这并不是说,读者不需要具备 HTML 的知识。所以,读者在阅读其他主题的过程中遇到有关 HTML 的内容时,请参考本书。

《JavaScript 编程指南》

JavaScript 是 Netscape 在 Java 的基础上开发出的一种脚本编程语言,这种语言功能比 HTML 强大,但比 HTML 复杂,由 JavaScript 编写的代码可直接嵌入到 HTML 文档中,由浏览器解释执行。

本书主要介绍了 JavaScript 的语法规则、对象应用和服务端编程等内容,并在讲解基本概念和原理时,辅以代码实例以及实用程序编程。

在阅读本书之前,读者应具备相当的 HTML 知识;否则请参考《HTML 编程指南》;同时,本书是《IntraBuilder 编程指南》的基础,因为 IntraBuilder 把 JavaScript 当做开发语言。

《CGI 编程指南》

CGI 是 Web 服务器提供交互式信息服务的标准接口。通过 CGI,Web 服务器就能够执行应用程序,并将它们的输出(如文字、图形、声音等)传送给一个 Web 浏览器。

本书首先介绍 CGI 基本知识,包括 CGI 简介、编程初步、运行准备,然后通过两个实用的例子“用户意见簿”和“网络考试”来讲解 CGI 编程应用,最后讲解了 CGI 编程的高级话题:数据库访问、搜索引擎、SSI、调试 CGI 程序和提高 CGI 程序的安全性。

由于 CGI、Visual InterDev 和 IntraBuilder 都是 Web 服务器编程的方式,请读者在阅读这三本书时注意相互参考,选择一种最适合当前项目的开发形式。

《FrontPage 98 编程指南》

FrontPage 98 是 Microsoft 公司的 Web 站点创建和管理的第三个版本,也是目前主页制作市场上占有率最高的产品。Microsoft FrontPage 98 包括 FrontPage 管理器和 FrontPage 编辑器。使用 FrontPage 管理器可用来创建、设计、查看和管理整个 Web 站点;使用 FrontPage 编辑器可创建和编辑任意复杂的 Web 页面,甚至可以不需要知道 HTML 的知识。本书以创建和管理 Web 页面的各种应用为线索,全面展示了 FrontPage 管理器和编辑器的各项功能和操作方法。

在使用《HTML 编程指南》时可以参考本书,因为 FrontPage 98 提供了一个 HTML 编辑和预览非常简便的环境;在使用《Visual InterDev 编程指南》时也可参考本书,因为 Visual InterDev 自带一个 FrontPage Editor for InterDev 版本,在《Visual InterDev 编程指南》中只用了非常有限的篇幅介绍 FrontPage,要深究 FrontPage 的应用,需参考本书。

《Visual InterDev 编程指南》

与 FrontPage 98 不同,Visual InterDev 是 Microsoft 公司的站点开发工具,主要用于 Web 客户端和服务端各种交互性功能的实现。

本书主要介绍了 Visual InterDev 的基本操作、使用 FrontPage Editor 来开发页面、用 VBScript 开发客户端和服务端脚本,以及如何进行 Web 应用程序与数据库集成等内容。

本书以 HTML 和主页制作为基础,这样,如果没有这些方面基础的读者,请注意参考《HTML 编程指南》和《FrontPage 98 编程指南》。

《IntraBuilder 编程指南》

IntraBuilder 是 Borland 公司推出的站点建立、Web 与数据库集成的解决方案,其采用 JavaScript 为开发语言,可支持多种数据库,包括 Paradox 及各类支持 ODBC 的数据库系统。

本书紧密结合数据库的开发和应用,详细介绍 IntraBuilder 的使用方法,主要内容有生成数据库表格、构造窗体、设计报表、数据库查询、安全管理、JavaScript 面向对象编程等。另外,在本书的最后,还给出了综合性的应用示例。

IntraBuilder 以 HTML 和 JavaScript 为基础,如果没有这些方面基础的读者,请注意参考《HTML 编程指南》和《JavaScript 编程指南》。

《简易 Intranet 开发指南》

采用 Office 97 和 FrontPage 98 可以创建一个简单实用的 Intranet,这种方法特别适合于那些包含几台到几十台电脑、由非计算机专业人员组成的办公室进行企业网组建。

本书专为那些想快速建立单位或工作组内部网络的用户编写。全书以通俗的语言、丰富的内容和耐心细致的操作指导,着重介绍了如何使用 Office 97 和 FrontPage 98 来快速构建 Web 服务器和 Intranet。

电子工业出版社

1999 年 1 月

前 言

通用网关接口(Common Gateway Interface, CGI)是 Web 服务器提供交互式信息服务的标准接口。通过 CGI, Web 服务器就能够执行应用程序,并将它们的输出,如文字、图形、声音等传送给一个 Web 浏览器。可以说,有了 CGI, WWW 才变得更为实用,界面更为友好,信息服务更为多种多样。

本书分为三部分。第一部分介绍了 CGI 的基本概念,包括 CGI 的工作原理以及 CGI 编程语言的选择等,然后阐述了有关 CGI 编程的初步知识,包括 CGI 环境变量、CGI 程序结构和 HTML 简介等。第二部分通过两个完整的 CGI 程序范例“用户意见簿(UserAdvice)”和“网络考试(NetworkExam)”,讲解了 CGI 程序设计的详细过程和编程思路。第三部分讲解了 CGI 编程的高级话题:通过 CGI 访问数据库,给 Web 站点建立索引,通过 CGI 使用 WAIS,使用服务器端包含(SSI),调试 CGI 程序和提高 CGI 程序的安全性。

本书的读者可以是想学习 CGI 的新手,因为书中提供了许多有关 CGI 的基本概念和编程基础知识,也可以是一名 CGI 编程人员,因为书中提供了许多非常有用的程度代码,可以非常容易地加以修改利用,并进一步提高编程水平。总之,无论你的基础如何,都可以从本书中得到帮助,很快编写出你自己的 CGI 程序。

本书是集体创作的结晶,除了封面上署名的主要编著者外,李子木、凌云翔、万俊伟、史湘宁等都参加了本书的编著工作。

目 录

第一部分 CGI 编程基础

第 1 章 通用网关接口 (CGI) 简介	(3)
1.1 CGI 与 WWW	(3)
1.2 CGI 的工作原理	(5)
1.3 CGI 程序语言	(6)
1.3.1 解释型语言	(7)
1.3.2 编译型语言	(7)
1.3.3 混合型语言	(8)
1.3.4 总结	(8)
1.4 常见 CGI 问题	(9)
第 2 章 CGI 编程初步	(10)
2.1 POST 和 GET	(11)
2.1.1 URL 编码	(11)
2.1.2 POST 方法	(12)
2.1.3 GET 方法	(13)
2.2 CGI 环境变量	(14)
2.2.1 环境变量概述	(14)
2.2.2 环境变量分类	(16)
2.3 CGI 格式化输出	(18)
2.3.1 CGI 标题	(18)
2.3.2 MIME 类型	(19)
2.3.3 Location	(21)
2.3.4 HTTP 状态码	(22)
2.4 CGI 程序结构	(24)
2.4.1 初始化	(25)
2.4.2 处理	(25)
2.4.3 终止	(26)
2.5 CGI 程序的可移植性	(26)
2.5.1 平台独立性	(26)
2.5.2 服务器独立性	(28)
2.6 几点忠告	(29)
第 3 章 CGI 程序运行准备	(31)
3.1 配置服务器	(31)
3.1.1 配置 NCSA 服务器	(32)
3.1.2 配置 Website 服务器	(34)
3.1.3 执行 CGI 程序	(35)

3.2	HTML 简介	(35)
3.3	HTML 文档结构定义	(37)
3.3.1	<HTML> ... </HTML>	(37)
3.3.2	<HEAD> ... </HEAD>	(37)
3.3.3	<BODY> ... </BODY>	(38)
3.4	HTML 表单	(39)
3.4.1	<FORM> ... </FORM> 标识对	(40)
3.4.2	INPUT 标识	(41)
3.4.3	文本输入框	(43)
3.4.4	SUBMIT 和 RESET	(44)
3.4.5	RADIO 控件	(45)
3.4.6	CHECKBOX 控件	(47)
3.4.7	<TEXTAREA> ... </TEXTAREA> 标识对	(48)
3.4.8	<SELECT> ... </SELECT> 标识对	(50)
3.5	HTML 表单实例	(53)

第二部分 CGI 编程实例

第 4 章	用户意见簿	(59)
4.1	UserAdvice 的基本框架	(59)
4.1.1	UserAdvice 的功能	(59)
4.1.2	UserAdvice.html	(60)
4.1.3	UserAdvice.pl	(60)
4.1.4	测试 UserAdvice	(64)
4.2	进一步完善 UserAdvice	(66)
4.2.1	用户直接访问	(66)
4.2.2	添加计数器	(67)
4.3	使 UserAdvice 更健壮	(68)
4.4	UserAdvice 的完整代码	(69)
4.5	UserAdvice 的响应结果	(74)
4.5.1	用户直接访问	(74)
4.5.2	用户输入	(75)
4.5.3	文件错误	(75)
4.6	用 C 改写 UserAdvice	(76)
第 5 章	网络考试	(87)
5.1	NetworkExam 的基本框架	(87)
5.1.1	NetworkExam 的功能	(87)
5.1.2	NetworkExam.html	(88)
5.1.3	NetworkExam.pl	(89)
5.1.4	测试 NetworkExam	(95)
5.2	进一步完善 NetworkExam	(97)
5.2.1	修改前台 HTML 文档	(97)
5.2.2	修改后台 CGI 程序	(101)

5.3	NetworkExam 的完整代码	(103)
5.4	NetworkExam 的响应结果	(110)
5.4.1	考生直接访问	(110)
5.4.2	考生答题	(110)
5.4.3	文件错误	(111)
5.5	用 cgi-lib.pl 改写 NetworkExam	(112)
5.6	用 C 改写 NetworkExam	(119)
5.7	用 cgihtml 改写 NetworkExam	(129)

第三部分 CGI 高级话题

第 6 章	Web 数据库应用	(141)
6.1	文本数据库	(141)
6.1.1	文本数据库的格式	(141)
6.1.2	客户界面	(141)
6.1.3	student.cgi 程序	(142)
6.2	Web 索引	(145)
6.2.1	给 HTML 表单加关键字	(145)
6.2.2	WAIS 索引	(148)
6.3	建立 Web/RDBMS 应用	(153)
6.3.1	为何要建立 Web 数据库应用	(153)
6.3.2	建立 Web 数据库应用面临的问题	(153)
6.3.3	简单的 Web 数据库应用示例	(154)
6.3.4	特殊 Web 数据库应用	(160)
6.3.5	Web/Database 工具介绍	(166)
第 7 章	建立 Web 站点上的搜索引擎	(169)
7.1	信息收集程序	(169)
7.1.1	机器人概念	(169)
7.1.2	机器人 Harvest 介绍	(169)
7.1.3	服务器上 robot.txt 文件的配置	(171)
7.2	信息组织程序	(173)
7.2.1	ICE 的建索引程序	(173)
7.2.2	SWISH 索引系统	(173)
7.3	信息查询	(175)
7.3.1	搜索引擎的信息查询方式	(175)
7.3.2	将搜索函数加入你的站点	(177)
7.3.3	基于索引的查询	(181)
7.3.4	无索引的查询	(184)
7.4	搜索引擎一览	(190)
7.4.1	基于小型站点的简单搜索引擎	(190)
7.4.2	商业站点搜索引擎	(191)
7.5	在你的页面中加入搜索引擎	(193)
第 8 章	SSI	(197)

8.1	SSI 概念	(197)
8.1.1	概念	(197)
8.1.2	SSI 的功能	(197)
8.1.3	SSI 与其他动态 HTML 技术的区别	(198)
8.2	SSI 语法	(198)
8.2.1	HTML 注释语法	(198)
8.2.2	SSI 语法	(198)
8.3	常用 SSI 命令	(199)
8.4	配置 SSI	(201)
8.5	SSI 编程示例	(202)
8.5.1	访问计数器	(202)
8.5.2	随机输出产生器	(210)
8.5.3	页面首次访问时间	(214)
第 9 章	调试 CGI 程序	(225)
9.1	避免 CGI 出错	(225)
9.1.1	使用完全路径名	(226)
9.1.2	CGI 标题出错	(226)
9.1.3	使用正确的 MIME 类型	(226)
9.1.4	检查输入信息的类型	(226)
9.1.5	注意编程语言的特点	(227)
9.1.6	服务器的问题	(227)
9.2	与 Web 服务器交互调试	(227)
9.2.1	服务器返回的错误	(227)
9.2.2	附加语句	(229)
9.3	输入信息的类型	(233)
9.4	浏览器与服务器之间的通信	(234)
9.5	调试器	(237)
9.6	小结	(238)
第 10 章	提高 CGI 程序的安全性	(239)
10.1	配置安全的 Web 服务器	(239)
10.1.1	配置 Web 服务器的原则问题	(239)
10.1.2	保证服务器平台自身的安全性	(240)
10.1.3	把 Web 服务器目录和文档目录分开	(241)
10.1.4	chroot 环境	(241)
10.2	CGI 程序的安全性原则	(241)
10.2.1	原则问题	(242)
10.2.2	示例: CGI 语言带来的安全问题	(242)
10.3	外部函数派生的安全问题	(244)
10.4	CGI 目录管理	(249)
10.5	一个隐蔽的安全漏洞	(251)
10.6	小结	(252)

附 录

附录 A	cgi-lib.pl 简介	(255)
A.1	cgi-lib.pl 的使用	(255)
A.2	cgi-lib.pl 子程序	(256)
A.3	cgi-lib.pl 源代码	(258)
附录 B	cgihtml 简介	(269)
B.1	cgihtml 的组成	(269)
B.2	cgihtml 的使用	(269)
B.3	cgihtml 头文件	(270)
B.4	cgihtml 子程序	(272)
B.5	cgihtml 源代码	(275)

第一部分 CGI 编程基础

第 1 章 通用网关接口(CGI)简介

通用网关接口(Common Gateway Interface, CGI)是一个 Web 服务器(Server)主机提供信息服务的标准接口,通过提供这样一个标准接口,Web 服务器能够执行应用程序并将它们的输出,如文字、图形、声音等传送给一个 Web 浏览器。一般来说,CGI 标准接口的功能就是在 HTML 文档与服务器应用程序之间传递信息。

CGI 这个标准接口主要解决以下几个问题:

1. Web 服务器和 CGI 程序之间如何传送信息?
2. Web 服务器向 CGI 程序传送什么信息?
3. Web 服务器向 CGI 程序传送信息的格式是什么?
4. CGI 程序向 Web 服务器返回什么信息?
5. CGI 程序向 Web 服务器返回信息的格式是什么?

如果没有 CGI,一个 Web 服务器将只能提供静态文本文件或连接到其他服务器。可以说,有了 CGI,全球互联网(World Wide Web, WWW)才变得更为实用,界面更为友好,信息服务更为多种多样。

CGI 的主要用途是使用户能够编写用于与服务器相交互的程序,并借助 CGI 可编写程序用于处理如下几种典型工作:

1. 动态创建新的 WWW 页面。
2. 处理 HTML 表单输入。
3. 在 Web 和其他网络之间建立相互沟通的渠道。

在这一章里,主要介绍 CGI 的基本知识。具体包括:

- CGI 与 WWW。
- CGI 的工作原理。
- CGI 程序语言。
- 常见 CGI 问题。

1.1 CGI 与 WWW

自从 Internet 兴起以来,计算机领域出现了从所未有的飞速发展,各种 Internet 应用也应运而生,其中,全球互联网(WWW)的出现是 Internet 发展史上一件划时代的大事。

WWW 是欧洲量子实验室(the European Laboratory of Particle Physics)在 1990 年提出构建的一个分布式多媒体网络信息系统。它采用开放式客户机/服务器结构(Client/Server),分成服务器结构(Server)及客户机结构(Client)两个部分。其中服务器结构规定服务器方的信息传输设定,包括信息传输格式及服务器本身的基本开放结构等;客户机结构规定客户机方的信息接收设定,包括信息接收格式以及适当的信息接收工具,如:浏览器(Browser)。

从硬件的角度来看,一个完整的 WWW 系统至少包含一个 WWW 服务器、一个作为物质基础的计算机网络以及多个 WWW 客户机,只要系统能力允许,它几乎可以被无限扩充。

从软件的角度来看,一个完整的 WWW 系统包含 WWW 服务器软件、Internet 网络协议和 WWW 浏览器。任何一台与 Internet 互联的计算机,只要配备了 WWW 服务器软件并提供 WWW 资源,就可以成为一个 WWW 服务器,本地或异地的 WWW 用户就可以使用 WWW 浏览器来访问它。

简单地说,访问 WWW 可以分为以下六个步骤:

1. 你(客户机)的浏览器分析所提供的统一资源定位器(URL)的第一部分,并试图连接到相应的服务器。
2. 你的浏览器提供 URL 的剩余部分给服务器。
3. 服务器将 URL 翻译成一个路径和文件名。
4. 服务器将所需文件返回给浏览器。
5. 服务器断开连接。
6. 浏览器显示文件。

图 1.1 显示了 Netscape 公司的 Netscape Navigator 浏览器。

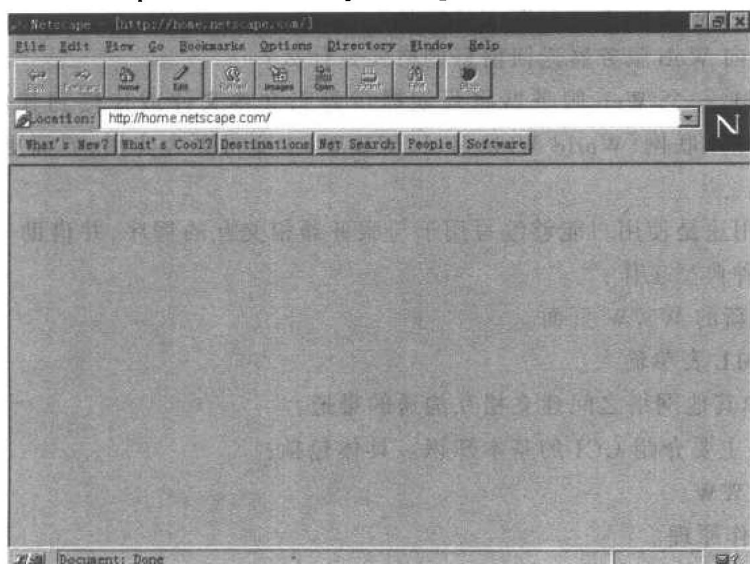


图 1.1 Netscape Navigator

不论是 WWW 服务器软件还是 WWW 浏览器必然都要牵涉到一个重要概念:WWW 中的主页(Hompage)。主页中提供的信息包含有:文字(Text)、图形(Graphic)、声音(Sound)以及影像(Video)。目前,在 WWW 中,已提供了一个大家都遵循的多媒体主页写作技术标准,称为超文本标记语言(Hypertext Markup Language, HTML)。要想在 WWW 提供多媒体信息,就可以利用 HTML 标记语言来编写多媒体主页以表达所要呈现的信息。WWW 服务器用来存放这些编写好的多媒体主页,而 WWW 浏览器用来阅读这些多媒体主页文件,以达到信息交流、信息共享的目的。目前,市场上有许多 HTML 编辑工具,可以帮助用户大大简化 WWW 资源的开发工作。另外,WWW 还有一个非常重要的特性:超链接,超链接使全球的 WWW 主页链接成一个网状结构,使得页面之间的互访变得非常容易。

通用网关接口(CGI)是一个 Web 服务器(Server)主机提供信息服务的标准接口。CGI 接口是为了提供在超文本 HTML 的文件编写时,可以结合其他外部的程序语言,让使用者能通过浏览器将使用者的数据输入到文件里,然后经 HTTPd 或 CERN 服务器主机处理后,使其他

用户能够进行访问。CGI 与 WWW 几乎同时出现,曾经是能提供交互性的唯一方法,目前,几乎所有的 WWW 服务器软件都支持 CGI。

可以预知,在未来信息综合交流下,WWW 势必占有极其重要的地位,而 CGI 作为信息服务的标准接口必然也将起到举足轻重的作用。

1.2 CGI 的工作原理

一个 CGI 程序并不是什么特别神秘的东西,其中也并没有什么魔法可言。实际上,使用一些最常见的编程语言如 C 或 Perl 等来编写 CGI 程序是一件简单不过的事情。

第一个 CGI 程序是用 UNIX 的 shell 语言(sh)和 Perl 语言编写而成的,它们都是解释型语言,就象 DOS 的 BAT 文件,当然,它们比 BAT 文件更为有效。其他 CGI 程序语言如 C 语言则是编译型语言。由此可见,一个 CGI 程序既可以使用解释型语言又可以使用编译型语言来编程实现。具体将在以后阐述。

那么,编写 CGI 程序的关键问题是什么呢?

当客户机端发出一个执行 CGI 程序的请求给服务器以后,服务器便根据 CGI 程序的类别,决定某种数据传递方式,即服务器与 CGI 程序的沟通方式,这就是编写 CGI 程序的关键所在。一般情况下,服务器与 CGI 程序之间都是通过标准输入输出(Standard Input and Output)进行数据传递的。

从操作系统的角度来说,标准输入(STDIN)与标准输出(STDOUT)是两个预先定义好的文件流句柄。每个进程在创建时,都同时创建了这两个文件流句柄,而且,它们总是打开的。当你对一个程序的输入进行重定向时,实际上你是对标准输入进行重定向;当你对一个程序的输出进行重定向时,实际上你是对标准输出进行重定向。例如,当你进行一次列目录(DIR)操作,并将其输出结果重定向到一个分屏显示程序(MORE),你实际上是对 DIR 操作的标准输出进行了重定向,并对 MORE 操作的标准输入进行了重定向。

对服务器与 CGI 程序而言就是:服务器的标准输出与 CGI 程序的标准输入相连接,服务器标准输入与 CGI 程序的标准输出相连接。如此,就构成了一个完整的数据循环通路。

从 CGI 程序的角度来看,标准输入是从服务器来的数据的通路,标准输出则是它所输出的数据的通路。这样,CGI 程序就无须担忧可能的输入输出重定向。从服务器的角度来看,服务器使用它的标准输出向 CGI 程序发送数据与结果,同时,服务器希望 CGI 程序通过标准输出向它发送有关信息。

简单地说,通过 CGI 访问全球互联网 WWW 可以分为以下十个步骤:

1. 你(客户机)的浏览器分析所提供的 URL 的第一部分,并试图连接到相应的服务器。
2. 你的浏览器提供 URL 的剩余部分给服务器。
3. 服务器将 URL 翻译成路径和文件名。
4. 服务器将 URL 指向一个应用程序而不是一个静态文件。
5. 服务器为应用程序执行做准备。
6. 应用程序执行,读取标准输入和有关环境变量。
7. 应用程序准备进行标准输出。
8. 应用程序发送标准输出并结束。
9. 服务器断开连接。

10. 浏览器显示应用程序的输出结果。

这种通过标准输入输出来实现服务器与 CGI 程序之间的数据传递的标准在基于文本的 UNIX 系统中尤其适用。

对于 Windows 与 Windows NT 操作系统而言,还可以使用另外一种方法来实现服务器与 CGI 程序之间的数据传递,这就是通过 profile 文件(如 INI 文件)。目前最常见的 Windows 服务器中,支持 Windows CGI 程序的有 Website, Netscape Enterprise Server, Microsoft IIS 等几种。而且,这种方法正在逐步成为一个标准,称为 CGI-WIN。支持这种标准的服务器将它的输出结果写入一个 INI 文件中,而不是输出到标准输出中。CGI 程序通过打开这个 INI 文件,就可以从这个 INI 文件中读取和处理数据,以实现其与服务器之间的数据通信。

另外,在服务器执行 CGI 程序之前,服务器必须从一些环境变量中找到与该 CGI 程序相关的一些执行信息,如:是谁在请求执行 CGI 程序?需要注意的是,因为每个服务器可能同时执行多个 CGI 程序,所以每个 CGI 程序在执行时都应该有自己独有的环境变量集。事实上,每个 CGI 程序的环境变量集在该 CGI 程序开始执行时建立,CGI 程序的执行一终止,其对应的环境变量集就失效了。关于环境变量的更多信息,我们将在后面的章节中更进一步去学习、探讨。

在 WWW 页面中,主要通过 HTML 表单或者超链接的方式来执行 CGI 程序。当服务器看到被请求的程序位于 CGI 目录时,就知道它是一个 CGI 程序,从而去执行该程序,并将执行结果返回给浏览器。浏览器接收到执行结果以后,将其在新的页面中显示出来。由此可知,CGI 工作目录对于 CGI 程序非常重要。

一般地,CGI 工作目录是一个特定的目录,所有编译好的 CGI 程序都要集中放在这个 CGI 工作目录中。对于 UNIX 系统而言,最常见的 CGI 工作目录名是 CGI-BIN,事实上,最早的支持 CGI 程序的服务器就是使用 CGI-BIN 目录。对于 Windows 与 Windows NT 操作系统(注意,这里不包括 NT 4.0),以 Webstar 或者 Website 作 WWW 服务器,最常见的 CGI 工作目录名则是 CGI-WIN。当然,目前的 WWW 服务器一般都允许你为 CGI 工作目录设置特定的目录名。

CGI 的跨平台性能极佳,几乎可以在任何操作系统上实现,如 DOS、UNIX、Windows、OS/2、Macintosh 等。实现 CGI 的编程语言也有许多种选择,目前最流行的 CGI 编程语言有四种:C, Perl, UNIX Shell 和 Visual Basic。正因如此,CGI 提出以后,非常受欢迎,而且被广泛地应用。

但是,CGI 应用程序一般都是一个独立的可执行程序,和 WWW 服务器各自占据着不同的进程,而且一般一个 CGI 应用程序只能处理一个用户请求。这样,每有一个用户请求,都会激活一个 CGI 进程,当用户请求数量非常多时,将会大量挤占系统的资源,如内存、CPU 等,而造成效能低下。

1.3 CGI 程序语言

首先需要明确的是:到底什么程序语言才能称为 CGI 程序语言呢?事实上,任何一个程序语言所设计的应用程序,如果可以在服务器上执行,并利用服务器与 CGI 程序的标准接口来传送有关信息,则这种程序语言就可以称之为“CGI 程序语言”。因此,非常自然地,有很多程序语言可供选择,而且,现在看来,这些选择中所存在的余地越来越大。但是,必