

网上多媒体档案系统的开发

戚小滨, 赵文耘, 王俊峰

(复旦大学计算机系, 上海200433)

摘 要: 讨论了如何开发网上多媒体档案系统, 比较了当前流行的几种Web技术, 并给出了一个实际的例子。

关键词: 浏览器; 超文本传输协议; 超文本标识语言

Development of Net Multi-Media-Archives System

Qi Xiaobin, Zhao Wenyun, Wang Junfeng

(Computer Science Department of Fudan University, Shanghai 200433)

【Abstract】 This article is about how to develop Multi-Media-Archives system on the Internet, we compared some of the most popular Web technologies and then described how we developed the system as an example.

【Key words】 Browser; HTTP; HTML

随着Internet的普及, WWW已经深入到了生活的各个领域。WWW最初是作为一种信息发布的方法而出现的, 它仅被用来作为企业宣传、产品广告以及学术交流的一种手段。随着时代的发展, 这种单纯的静态的信息发布越来越不能满足人们的需要了。究其原因, 主要是这种静态网页缺乏交互性, 不能根据用户的需求提供动态的服务。

动态页面的出现极大地改变了WWW。今天, 人们可以通过网页进行查询、购物、数据汇总等工作。在这个过程中, WWW的技术也有了突飞猛进的发展。

1 静态页面向动态页面的过渡

WWW的工作原理是基于Client/Server结构的。客户端是Browser, 服务器端是Web Server。Browser通过Internet向Web Server发出请求, Web Server响应请求并将结果以HTML文本格式返回给Browser。Browser接到文本后将其显示给用户。静态页面的工作原理如图1。

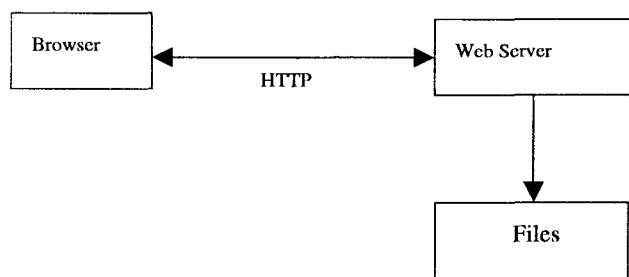


图1 静态页面工作原理

从图1中可以看出, Browser发出调用页面的请求, Web Server找到该页面后将其返回给Browser。在这种情况下, 页面是事先制作好的, 以文件的方式存放在服务器上, Web Server只是简单地查找所需的页面并将其返回, 而Browser只是根据文本的格式将其显示出来。显然, 这种静态的页面的能力是十分有限的。为了提高交互性、动态性, 人们使用了许多机制, 使静态页面逐步过渡为动态页面。下面将简单介绍这些机制。

WWW的工作方式决定了这些扩展的方向--对Browser

端的扩展和对Web Server端的扩展。

(1) 对Browser端的扩展

· 扩展HTML文本格式 在HTML中包含更多的显示格式, 如动态GIF(增加动画效果)、Frameset机制(在一个窗口内显示多个HTML文本)以及增加闪烁、彩色等显示功能。

· 在HTML文本中增加脚本语言(Script) 脚本语言是一种简单的编程语言, 嵌在文本中, 由Browser解释执行。目前流行的有JavaScript和VBScript。在这些语言中, 可以定义变量、响应用户的事件(键盘、鼠标), 这使得页面有了较大的交互性。但由于脚本语言较为简单, 且是在客户端由Browser解释执行, 因而功能有限, 也无法完成与Server端的数据库的连接。

· Java Applet 它是运行于客户端的由Browser解释执行的程序。由于Java是真正的面向对象的编程语言, 因此具有强大的功能。但由于Java语言本身并不十分成熟, 且Java的数据库接口(JDBC)是建立在ODBC之上的(现在已经有了独立的JDBC接口, 但Browser支持并不理想), 因此, 它与数据库的连接效率并不理想。

· Plug-in(插件)和(Browser端的)ActiveX控件 Plug-in的思想最初是由Netscape Communications Corporation提出的, Plug-in是动态代码模块, 是将第三方软件集成进Browser的API。它可以解释Browser自己不能处理的文件(如声音、影像等), 通过不断地增加插件, 可以几乎无限制扩展Browser处理各种文件的能力。ActiveX控件是Microsoft的基于COM模型的控件, 在Browser端, 其作用与Plug-in很相似。Plug-in和(Browser端的)ActiveX控件主要用于解释多媒体文件, 增加网页上的一些动态效果, 并没有真正解决与数据库的连接。

· DHTML(动态HTML) 它将HTML文本的所有元素都视为对象, 允许用脚本语言对其进行控制, 可以动态地改变其内容。并且, 有些DHTML标准(如Microsoft的IE4.0以上)还支持Data Binding, 可以将页面的元素直接Binding到数据集, 比如, 将一个Table的内容直接Binding到服务器的数据库中的某个表或视图, 则该Table将显示表或视图的内容。这种方法可以与数据库连接, 但功能很有限。具体实现是使用ActiveX和DCOM实现的, 此处不赘述。

作者简介: 戚小滨(1972~), 男, 研究生, 主研方向: CASE工具和软件工程

收稿日期: 1999-10-14

因此,只是从Browser端入手并不能完成真正意义上的动态网页,真正的动态网页必须与数据库相连,能够从数据库中查找数据,并对数据库做插入、修改、删除等工作(就像传统的Client/Server结构那样)。数据库通常放在Server端,因此,应对Server端进行扩展。

(2) 对Server端的扩展

· CGI(Common Gateway Interface)应用程序 CGI是为HTTP服务器和服务端间的网关程序通信而规定的标准。服务器和网关程序的工作关系如图2所示。

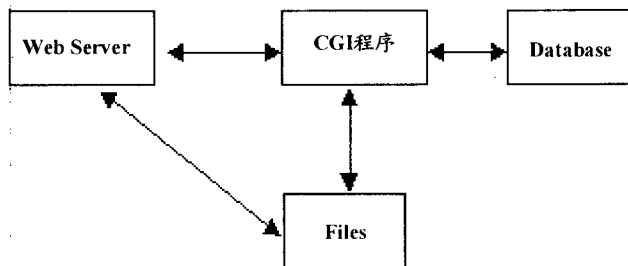


图2 服务器和网关程序的工作关系

网关程序有不可避免的缺点:网关程序是运行在与服务器不同的另一个进程中的,因而效率不高,特别是当访问频繁时,系统中会产生大量进程,从而使效率大大下降。

· ISAPI和NSAPI 针对CGI方式的弱点,ISAPI和NSAPI应运而生。ISAPI和NSAPI程序并非独立的程序,而是被编译成动态链接库(DLL),与Web Server运行于同一进程空间,因而具有较高的效率。其不足之处在于安全性,由于和Web Server运行于同一进程空间,若编制不当,可能造成较严重的后果。

· IDC(Internet Database Connector) 以Microsoft的IIS(Internet Information Server)为例。这种方式下,用户通过编写.idc文件和.htx文件来访问数据库。其中,.idc文件定义了SQL语句和数据源等,IIS通过调用httpodbc.dll读取.idc中的参数,并通过ODBC访问数据库。.htx文件定义了返回给Browser的文件骨架(HTML文本),IIS将访问结果填入到.htx文件的适当位置从而形成标准的HTML文本,再将其返回给Browser。这种方法的缺点在于,访问数据库的部分和返回部分的文本是分开编写、存放的,开发时不方便。

· ASP(Active Server Pages) ASP是Microsoft推出的一种文件格式,它将脚本语言、控件(ActiveX Server Components)有机结合起来,具有强大的功能。其中,控件可以用来连接数据库,进行文件操作,而脚本语言(VBScript或JScript)运行于Server端,可以控制控件,根据用户的不同请求产生不同的数据库动作,返回不同的页面(页面中可以包含Browser端的脚本程序)。由于控件可以由第三方提供,也可以由开发者自己编写,因此,这种方法能极大地扩展Web Server的能力。而且,对控件的调用和对返回的HTML文本的编写是在同一个文件中的,这使得开发十分方便。Microsoft的IIS提供了最常用的控件,可以满足一般的对数据库的访问和文件访问等功能。

· 其它方案 一些厂商提供了自己的Web Server,通过各种方式连接到数据库(如:运行在Server端的Java和JDBC)。

2 档案系统的现状及多媒体档案的一些问题

对一些档案馆的现有系统进行了调查,发现现有档案系

统存在以下特点:

· 档次不齐 有些档案馆已经实现了计算机管理,比较好的采用了Client/Server结构,有些是在单机下采用的Foxpro系统,有些还停留在手工阶段。

· 信息共享不足 因为没有采用Internet技术,资料不能上网,利用率较低。

· 没有很好地利用多媒体信息 当前许多档案馆都有大量的多媒体信息,如图片、录音、录像。但由于保存方面的原因,不能轻易借出(因为一些比较珍贵的图片等资料容易损坏),其利用率非常低,并且保管不便。

针对以上特点,我们认为,网上多媒体档案系统是大势所趋,只有充分利用WWW技术,实现信息共享,才能有效地发挥档案的作用。

3 本项目的需求及完成方法

复旦大学档案馆拥有建校以来的大量档案资料,包括各种文件、照片、录音、影像等,该馆已经实现了内部的计算机管理,拥有一个比较完备的计算机辅助管理系统,可以对档案进行分类管理,并且已经将所有档案都录入到该系统中,采用多媒体方式进行了存储。档案的管理采用"项目-案卷-文件"3级系统,分类清晰。

(1)需求

针对日益增长的用户需求,该馆准备将档案在网上开放,允许注册的用户从网上进行查询、浏览,查询时要验证用户身份,并且要根据用户的身份来确定可以访问的文件范围,最后,要统计每个用户对系统的访问次数,以及所访问的文件。

(2)设计及开发

针对以上需求,决定采用以下方案进行开发:

1) Server端采用Microsoft的IIS作为服务器,并用ASP进行开发,与数据库的接口采用ODBC。

2) 档案的编排信息存放在数据库中(SQL Server),具体的多媒体文件存放在硬盘或光盘中。

3) 对用户指定其可以访问的密级,用户只能访问低于该密级的文件。

4) 充分利用Browser功能,适当地在客户端和服务端分配任务,让客户端做一些"力所能及"的事。在返回的文本中,包含了脚本语言,可以增加动态性,对用户的事件进行响应,对用户的输入作合法性检查,这就减少了服务器的压力,减少了网络开销,提高了效率。对多媒体的播放(声音和影像)是通过Plug-in和ActiveX完成的。

4 结论及今后的工作

档案的价值在于共享和利用,网络技术的发展为这种共享和利用提供了有利的条件。目前,关于如何开发网络应用的研究还处于初级阶段,特别是网上多媒体档案的研究和应用还有很多空白,今后将采用构件和构架的方法对这种应用建立模型并进一步研究。

参考文献

- 1 Microsoft MSDN Library.1999-07
- 2 Jamsa K.Internet编程.北京:电子工业出版社,1996-05
- 3 杨武杰.Java编程与实例.北京:电子工业出版社,1997-06