

内容管理系统中动态网页静态转化的三种方法

Automatic Generation of Static Page from Dynamic Webpage in the Content Management System

王 晖

WANG Hui

(广西瀚特信息产业股份有限公司, 广西桂林 541004)

(Guangxi Hunter Information Industry Co., Ltd., Guilin, Guangxi, 541004, China)

摘要: 提出网站内容管理系统中动态网页自动生成静态页面的三种方法, 阐述了每种方法的控制流程、关键技术及核心代码。

关键词: 管理系统 动态网页 静态网页

中图分类号: TP393.092 **文献标识码:** A **文章编号:** 1002-7378(2012)01-0041-02

Abstract: In the content management system, 3 methods that can automatically generate static page from dynamic webpage are proposed. The control process, key technology and core code of each method are illustrated.

Key words: management system, dynamic page, static page

自动生成静态页面技术是提高网站性能的必要手段, 但是对于有大量内容并且频繁更新的网站, 如行业门户网站和政府门户网站, 我们无法全部手动挨个实现每个页面的静态化, 于是就迫切需要一个能自动生成静态页面的内容管理系统。如新浪、搜狐等门户站点的新闻频道, 甚至其它频道, 每条新闻都是通过该系统来管理和实现的。自动生成静态页面一直是大型网站内容管理系统的重要功能之一。它能够有效提高网站访问速度、安全性、稳定性, 有利于搜索引擎对网站内容的收录, 从而使得网站的流量也能得到进一步的提高。由于采用 asp.net 的动态页面^[1] 都需要读取数据库程序才能显示内容, 如果访问量较大, 就会产生大量的数据库访问请求, 也产生了大量的数据库读写次数, 导致占用了大量的服务器资源, 最终影响了网站访问的速度。而采用全静态的网站避免了数据库的读写操作, 减少了访客获取内容的环节, 有效减轻了服务器在大并发、大流量下的压力。从安全角度分析, 由于黑客从前

台页面上无法了解网站采用程序的参数和数据库的地址, 就很难对网站进行有效的攻击。从稳定性角度分析, 由于程序或者数据库出现问题会经常导致网站无法访问, 而静态页面就从根本上杜绝了此类现象, 避免了因为网站程序出现问题而导致用户无法访问, 严重影响网站的信用度。从网站优化来分析, 搜索引擎更喜欢静态的网页。由于静态网页便于抓取, 搜索引擎 SEO 排名更容易提高。因此动态网页自动生成静态页面对网站内容的管理具有重大的现实意义, 而且能给用户带来较好的经济效益。本文提出了三种动态网页转换成静态页面的方法, 指出各自的优缺点并阐述其核心代码。

1 直接将动态网页生成静态页面

该方法(图 1)主要用于对原有的动态站点进行静态化页面的生成, 也就是内容管理系统在更新内容的同时自动生成静态页面, 技术实现比较简单, 对原来站点改动很小, 但是生成静态网页的性能低、速度慢。实现的关键技术是采用重载动态页面的 Render() 函数^[2], 当系统请求生成静态页面时将代码写到静态页面中, 接着将原来的动态链接修改为静态链接。

收稿日期: 2011-12-03

作者简介: 王 晖(1980-), 男, 工程师, 主要从事网站开发技术研究。

直接将动态网页生成静态页面的核心代码如下:

```
//采用重载动态页面的 Render() 函数将代码
写入页面
base.Render(sw);
//创建 sw 对象用于读取文件
System.IO.StreamWriter sw;
//将需要生成的文件写入 sw
sw = new System.IO.StreamWriter(Server.
MapPath("CreatedFileName"), false, System.
Text.Encoding.UTF8);
```

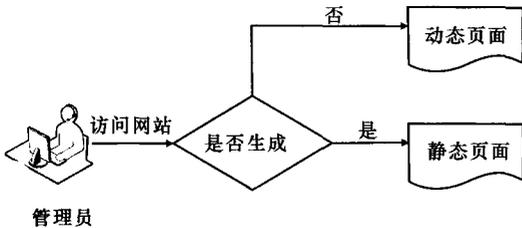


图1 直接将动态网页生成静态页面的控制流程

2 通过自定义网站模板方式生成静态页面

通过自定义网站模板的方式生成静态页面(图2)效率高、速度快,但是需要在网站建设之前就确定采用此方法生成,对于已建好的动态站点改版时需要很大的工作量。该方法的关键技术必须先把网页的公共部分做成一个网站模板,而动态变化的部分采用特殊字符(如标签)代替。在生成静态页面的时候,通过程序读取网站模板,接着再去网站数据库中找到需要的数据替换网站模板中的特殊字符,然后将生成真正的网页保存在网站的目录下。

自定义网站模板的方式生成静态页面的核心代码如下:

```
//通过模版文件地址将内容赋给对象 sr
System.IO.StreamReader sr = new System.
IO.StreamReader(templatefile, System.Text.
Encoding.UTF8);
//将生成的文件写入指定的目录
System.IO.StreamWriter sw = new System.
IO.StreamWriter(Server.MapPath("/Article/" +
dr["ArticleId"].ToString() + ".html"), false,
System.Text.Encoding.UTF8);
//定义变量 templete 用于替换标签
string templete = sr.ReadToEnd(); sr.Close
(); sr.Dispose();
//替换 Title 文章标题的标签
```

```
templete = templete.Replace("<@Title@>",
dr["Title"].ToString());
//替换 Content 文章内容的标签
templete = templete.Replace("<@Content
@>",dr["Content"].ToString());
//替换 Editor 文章编辑人的标签
templete = templete.Replace("<@Editor
@>",dr["Editor"].ToString());
//替换 Hits 文章点击数的标签
templete = templete.Replace("<@Hits@>",
dr["Hits"].ToString());
//将标签写入生成的文件
sw.Write(templete); sw.Flush(); sw.Close
();sw.Dispose();}
```

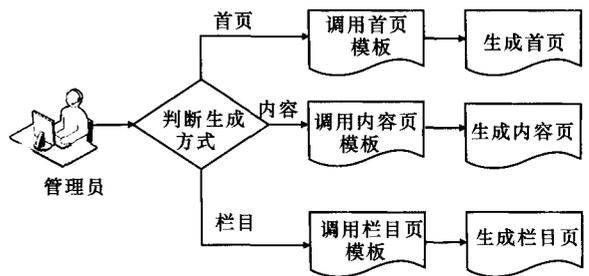


图2 通过自定义模板生成静态页面的控制流程

3 通过爬虫生成静态页面

通过爬虫生成静态页面的方式(图3)生成的效率高,无需人工管理,访客访问页面后自动生成,但是由于需要在之前就设置爬虫的条件,如果设置不当会造成错误。关键技术是根据动态页面生成静态页面方法,逐个页面的访问才能生成静态页面,如果我们需要批量生成,那就需要用采用 http 访问的方式^[3],并运用生成管理器的爬虫程序去爬遍每一个页面来达到自动生成静态页面的目的。

爬虫生成静态页面的核心代码如下:

```
//定义需要生成的文章 Url 地址
System.Collections.Generic.List<string> urls
= new System.Collections.Generic.List<string>
();
//判断表格的总行数
if (ds != null && ds.Tables.Count > 0
&& ds.Tables[0].Rows.Count > 0)
//循环遍历所有行
{for (int i = 0; i < ds.Tables[0].Rows.
Count; i++)
```

(下转第 46 页)

没有做好详细勘察,对项目的技术要求认识不够充分,盲目地签订了施工合同,结果在施工过程中遇到了困难却无力解决,只好自己选择了退出,但是项目的工期就延误了,进度也赶不上,而且给后面接替的施工队伍带来了压力。所以 IT 项目经理在这个时候需要调整施工计划,加强对施工队伍的选择和管理,在控制好质量的前提下,用“赶工”的方法加快施工进度,赶上原来的进度计划。

3 结束语

有几种软件可以帮助进行项目进度管理。常用的项目管理软件比如微软 Project 2000 一类的软件,是专为实施项目管理任务而设计的。可以用它来绘制网络图,确定项目的关键路径、创建甘特图以及报告、浏览和筛选特定项目时间管理信息。项目管理软件还可以计算所有活动的自由和总浮动时间或时差^[1]。项目管理软件有利于项目经理管理项目

进程,但是在使用项目管理软件还应注意理解隐藏在创建网络图、确定关键路径或设置进度基准背后的概念^[1],以避免误用项目管理软件,或者过分依赖于示范文件或模板。

俗话说“以不变应万变”,只要项目经理在 IT 项目实施前做好充分的准备,做好工作计划,当出现突发情况时,处乱不惊,相应地对项目进行调整,合理有效地安排好各项工作,就能把住 IT 项目进度的脉搏,顺利完成 IT 项目。

参考文献:

- [1] (美)凯西·施瓦尔贝. IT 项目管理[M]. 邓世忠,译. 北京:机械工业出版社,2004.
- [2] 柳纯录. 系统集成项目管理工程师教程[M]. 北京:清华大学出版社,2009.

(责任编辑:邓大玉)

(上接第 42 页)

```

//定义爬虫去爬遍所有页面
urls.Add("http://" + ServerName + "/Article/Show.aspx? ArticleId=" + ds.Tables[0].Rows[i].ToString());}
//批量生成静态页面
return CreateHtml(urls); }

```

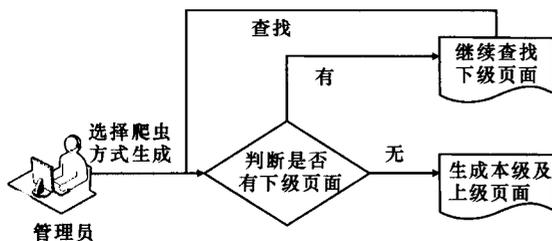


图 3 通过爬虫方式生成静态页面的控制流程

4 结束语

自动生成静态页面一直是大型网站内容管理系统的重要功能之一。随着互联网行业的迅速崛起,传统的动态网站内容数据库信息不断扩大,对服务器的资源消耗也越来越严重,造成访问者打开网站

的速度特别慢;并且原有的传统动态网站不容易获得很好的搜索引擎排名。本文提出的三种动态网页自动生成静态页面的方法各有优缺点,可以用于中、小型网站来提升内容管理系统的访问速度和安全性,可以使得用户的网站更加容易被最终用户找到,从而节约了广告方面的投入。采用自动生成静态页面技术也会给访客带来了更好的访问体验,从而便于网站获取最佳的知名度和经济效益。

参考文献:

- [1] 丁士锋,蔡平. ASP.NET 项目开发指南[M]. 北京:清华大学出版社,2010.
- [2] 王杰瑞,孙更新,宾晟. ASP.NET3.5 从入门到精通[M]. 北京:科学出版社,2009.
- [3] (美)内格尔,(美)埃夫琴. C# 高级编程[M]. 第 7 版. 北京:清华大学出版社,2010.

(责任编辑:尹 闯)