

基于 IIS 实现 WWW 与数据库互连的高级应用

魏星 丁伟

(东南大学计算机科学与工程系 南京 210096)

xwei@njnet.edu.cn

【摘要】基于 IIS 实现 WWW 与数据库互连的技术目前已经有了相当广泛的应用，本文在此基础上作了一些着重于高级应用方面的补充说明，并根据华东北地区网络中心 Intranet 的建设经验，总结了实际应用中遇到的问题及几点改进建议，提出了与 Perl 语言配合动态维护源程序的方法。

【关键词】IIS WWW Intranet

一. 引言

在 CERNET 96 上汤立群等同志的论文《SQL 服务器与 WWW 服务器的互连技术》中讲述了一种基于 IIS (Microsoft 的 Internet Information Server) 的 WWW 与数据库的互连方法，另一篇余江等同志的《建立在 Microsoft Windows NT Server 平台上的 Internet 服务》与 CERNET 97 上尹春华等同志的论文《利用 IDC 实现 WWW 与数据库的连接》均对此方法作了详细描述，华东北地区网络中心从 96 年起开始开发基于 IIS 的 Intranet，至今已建立了一套功能较为完整的基于内部网的办公系统，在开发实现的过程中遇到很多问题，由此总结出基于 IIS 实现 WWW 与数据库互连方法的补充说明以及缺点和改进建议，可作为以此方法实现一些高级应用的参考。

二. 系统提供的一些高级功能

在 IIS 上 WWW 与数据库互连的工作流程及具体方法可参看前述三篇文章，在此只在它们的基础上作一些补充说明，主要是系统的高级功能，也有前述文章中未讲到的一些遗漏内容：

1. ODBC 的高级功能

IDC 文件是由一系列“属性：值”对构成的，主要属性除前述文章中所述外，还有：

Expies: 在刷新一个输出缓存前等待的时间

MaxFieldSize: IDC 每个域返回的最大空间，默认为 8196 字节

另外，ODBC 还有一些高级功能，选择这些属性可用于开发 ODBC 驱动程序的功能，格式为：

ODBCOptions: Option Name=Value[,Option Name=Value]

比如，想要在执行 SQL 语句 10 秒后中止，且用户可以跟踪 ODBC 的功能调用，可以采用下面的形式：

ODBCOptions:SQL_QUERY_TIMEOUT=10,SQL_OPT_TRACE=1,SQL_OPT_TRACEFILE=c:\Sql.log

这些高级功能见附 1 (参考文献 4)。

2. 由 HTML 向 IDC 文件传递参数

由 IDC 文件向 HTX 文件传送参数的应用在前述文章中已有详细讲述，形式为<%XXX%>，XXX 为 IDC 由数据库表中选出的列名；在 IDC 文件中，也可以使用来自 HTML 的参数，这是很普通也是很常见的一种应用，形式为%XXX%，如：

在 HTML 中，有

```
<SELECT MULTIPLE NAME="region">
<OPTION VALUE="Western">
<OPTION VALUE="Eastern">
<OPTION VALUE="Northern">
<OPTION VALUE="Southern">
</SELECT>
```

而在 IDC 文件中可以写成：

SQLStatement:SELECT name,region FROM customer WHERE region IN ('%region%')

如果用户选择了"Northern","Western","Eastern"，则 SQL 语句变成：

```
SELECT name,region FROM customer
WHERE region IN ('Northern', 'Western', 'Eastern')
```

3. HTX 文件中的条件判断

HTX 文件中条件选择语句的格式为：

```
<%if condition%>
HTML text
[<%else%>
HTML text]
<%endif%>
```

可以嵌套。其中条件部分可写成“值 1 操作符 值 2”，系统提供四种操作符：

- ! EQ: 等于
- ! LT: 小于
- ! GT: 大于
- ! CONTAINS: 包含

4. HTX 文件中可使用的变量

在 HTX 的条件判断中还可以使用 IIS 和 IDC 定义的全局变量或从对应的 IDC 文件中来的变量，当然如果需要，这些变量也可在条件语句以外的地方使用。

a) IDC 的全局变量

可参看 IDC 文件中的“属性：值”对，例如：MaxRecords、CurrentRecord：均为 IDC 的全局变量，MaxRecords 表示 IDC 方法能得到的最多记录数，CurrentRecord 表示目前记录在<%begindetail%>和<%enddetail%>之间的位置。

b) 从对应的 IDC 文件中来的变量

此处的变量，并非指 IDC 文件从数据库中选出的表列名，该种应用方法已在第 2 点讲过，这里的变量指的是.idc 文件由 HTML 文件得到变量，这些变量也可被.htx 使用，例：如果.idc 中有变量%name%，则在.htx 文件中可以在变量前加上“idc.”，形如：

The Value of the name parameter is:<idc.name%>

c) IIS 提供的全局变量

使用 IIS 提供的全局变量了解目前服务器和浏览器的信息，在使用这些变量时，必须注意：（1）变量前加上 HTTP_；（2）所有字符大写；（3）把“-”转换成“_”。下面是一个例子，若用户以 root 登录，则执行某些操作，若是除 root 外的其它用户，则只给出拒绝信息：

```
<if HTTP_REMOTE_USER EQ 'root'>
Welcome .....
```

```
<%else%>
Sorry,<%HTTP_REMOTE_USER%> hasn't right to do it!
<%endif%>
```

这些变量见附 2（参考文献 4）。

三．系统功能的局限

基于 IIS 实现 WWW 与数据库的互连由于有很多优点（可参看前述三篇文章），已得到了较为广泛的应用，但在使用过程中，根据用户不断提出的新的要求，也发现了许多不便之处：

1. HTX 文件的处理能力不够完备

HTX 作为一种扩展的 HTML 模板文件，已经通过关键字提供了一些很重要的功能，对于一般的应用都足够了，但对于一些较复杂的应用仍显不足。

首先，<%begindetail%>和<%enddetail%>实际相当于一个大的循环调用，逐行处理 IDC 文件由数据库表中取出的信息（实际是 Httpodbc.dll 读取.idc 文件后执行数据库操作，并将执行结果存入.htx 文件中，具体步骤可参看前述论文，在本文中为便于叙述和理解，省去详细步骤，认为是.idc 文件取出数据并存入.htx 文件），每次处理只能针对一行一次到底，若要使用上一行处理的信息则是不可能的；

其次，HTX 只提供了基本的条件逻辑，象循环、分情等语句均没有提供，虽然这可以通过繁复的编程来实现，但显然增加了程序员的负担。

2. IDC 与 HTX 的接口一次调用只能传送一次数据

IDC 文件由数据库表中读出数据，然后写入 HTX 文件模板生成 HTML 文件，这个过程只支持一次数据传送，IDC 文件只能由数据库中一次性读取数据，并一次性写入 HTX 文件处理，若 IDC 要执行两个 SELECT 语句，比如读取两个表或由一个表中读取两次数据，则只有第一个能成功。若将这两个 SELECT 语句合成一个，则读取两个表的情况结果为读出的两个表的数据的笛卡尔乘积。

四．解决方案

上述问题在相当程度上限制了基于 IIS 实现 WWW 与数据库互连方法的使用。在不改变系统基础和数据库结构的前提下，为了满足用户的要求，我们在实践过程中，总结出几点折衷的解决方法：

1. 建临时表

a)HTX 文件的<%begindetail%>和<%enddetail%>之间每次只能处理数据库表中的一行，而且一次到底，若要使用以前处理过的行中的信息或非数据库中的其它信息（如当前时间，或判断标志信息），在 IDC 文件执行时建临时表是较好的方法。临时表的标志是表名的首字符为“#”，建表方式使用普通的 SQL 建表语句，如：

SQLStatement:

+create table #表名(表项名 1 类型, 表项名 2 类型...)

将需要的信息写入临时表，使用时读出，一个 Session 过后，系统会自动将临时表删除。

b)由于 IDC 与 HTX 的接口每次调用只能传送一次数据，若要查询两个以上的表中的数据，只能写在一个 SELECT 语句中，如

SELECT * FROM table1,table2

如前所述，这样得出的结果将是 table1 与 table2 的笛卡尔乘积，

若 table1 中的内容为 a11 a12

	a21	a22		
table2 中的内容为	b11	b12		
	b21	b22		

则上面的 SQL 语句执行结果为

a11	a12	b11	b12
a11	a12	b21	b22
a21	a22	b11	b12
a21	a22	b21	b22

这种反复重复的信息填入 HTX 模板后仍会按照原来的方式重复，前面讲的建临时表后读取信息如果是两个以上的表也会遇到同样的问题。对于这类问题，只能采取类似枚举法的方法解决，且只允许一个表的行数任意，其它表的行数必须是已知或可估范围的。具体方法如下：按照其它表的行数建临时表，建表方式多样，最终目的是使得只有一个表的行数任意，其它表选出的只有一行。如上例，设表 2 为行数可枚举的表，则可建 #table2 和 #table3，表列与 table2 相同，存储的数据分别为 b11 b12 和 b21 b22，“SELECT * from table1,#table2,#table3”的结果为：

a11	a12	b11	b12	b21	b22
a21	a22	b11	b12	b21	b22

虽然仍有重复信息，但已清楚多了，而且这种可枚举的表常用来存储判断信息，即使重复也不需显示。

2. 枚举法

对于某些内容在相当长的一段时间内相对稳定的信息，可以不必入库，直接以枚举法写入程序。例如，在华东北地区网络中心的 Intranet 中有一项值班记录，记录每周的值班情况，要求当班人员登录时与其它用户作不同处理，这就要判断帐号，而显示时又希望以中文姓名而非帐号显示，这就要求有帐号与中文姓名的对应表，若把此表放在数据库中，由于 <%begindetail%>和<%enddetail%>每次只处理一行数据，对应表不能全部读出，而使用临时表在用户多时则很烦琐，考虑到用户的变更情况较少，将此表用枚举法做在程序中。

这种方法思路简单，易于实现，但编程繁复，工作量大，且一旦数据有所变动，维护极为不便。

3. 与 perl 语言的配合

前述的枚举法在维护时工作量很大，且都是机械性工作，为此，我们采用了与 perl 语言配合，以 perl 程序来修改需维护程序的方法。

Perl(Practical Extraction and Report Language)叫做文字分析报告语言，由 Larry Wall 建立。Perl 语言中包含了 C、C++、shell script、sed、awk 这几个语言的语法。它的原始目的就是用来取代 UNIX 原来 sed/awk 与 script 的组合，用来汇集信息，产生报表的一个工具程序语言。且随着版本的改进，功能已超乎原来设计的想象。

我们采用 Perl 是因为它对字符串与数据剖析方面有很强的处理能力，尤其是利用关联性数组来作 CGI 程序剖析输入数据串，可以说是一个功能强大的语言。在 Perl 5.0 版本中还增加面向对象的用法；增加对字符串的处理能力，可以把整个文件当成一个字符串来处理，不受任何数据的大小限制而只受存储器的大小限制；能够处理 DBM 数据库格式的数据，也能够处理二进制的数等；Perl 还可以在 UNIX、Windows NT、Windows 95、DOS、Linux、FreeBSD、OS2、Macintosh 等操作系统中使用。Perl 语言易学习、易使用，写出来的 CGI 应用程序也是简洁有力。因此，目前在国外一些功能强大的 CGI 应用程序几乎都是用 Perl 语言充当主角。在 IIS 上，我们的实践经验表明，以 perl 写的 CGI 程序比 C 等语言更加稳

定。

在上例中，以 perl 写了增加和删除值班成员两组 CGI 程序，去修改 HTX 文件中帐号和中文姓名的对应表。以删除为例，其工作步骤如下：

- a) 由指定的.htx 文件中以特定标志找到该对应表；
- b) 读取表并以下拉选择框的方式显示；
- c) 根据用户的选择，删除所有需修改的.htx 文件中的对应项目。

以 IDC 文件与 HTX 文件实现 WWW 与数据库互连在复杂应用上有很多不足之处，以临时表和枚举法只能解决少量一部分，而采用枚举法与 perl 语言配合去修改源程序则是一种很新的思路，且能收到相当好的效果，极大减轻了程序员的负担。

五. 结论

基于 IIS 的 ODBC 可容易地实现 WWW 与数据库的互连，但这只能用于较简单的应用，对于复杂应用仍有很多问题，在不改变系统基本框架结构的要求下，本文提出了几点解决方法，在华东北地区网络中心的 Intranet 中取得了很好的效果。

六. 参考文献

- 【1】汤立群等，SQL 服务器与 WWW 服务器的互连技术，中国教育和科研计算机网的研究与发展，第一卷，1996
- 【2】余江等，建立在 Microsoft Windows NT Server 平台上的 Internet 服务，同上
- 【3】尹春华等，利用 IDC 实现 WWW 与数据库的连接，中国教育和科研计算机网的研究与发展，第二卷，1997
- 【4】Microsoft Internet Information Server Installation & Planning Guide, 1995

Supper techniques of integrating WWW and Database on IIS

Wei Xing Ding Wei

(Computer Science and Engineering Department, Southeast University, Nanjing, 210096)

【Abstract】It becomes more and more popular of integrating WWW and database on IIS for information services. According to the experience of constructing the Intranet in CERNET eastern China (north) regional network center, the paper summarizes some supplements of advanced technique in this field firstly, then the problems being met during the process are arisen and the ways to cover them are given out at the same time which include a method of maintaining source code dynamically with Perl language etc.

【Keywords】IIS WWW Intranet

附 1

SQL_ACCESS_MODE	0=读写 1=只读	默认值为 0，表示对 ODBC 驱动种的访问方式
SQL_LOGIN_TIMEOUT	整数	表示登录的时间限制，单位为秒，如果置为 0，则用户可以任意登录
SQL_OPT_TRACE	0=关	当开关打开时，每次对 ODBC 功能的调用都会被写入一个文件
SQL_OPT_TRACEFILE	文件名	当 SQL_OPT_TRACE=1 时用于记录的文件名，默认为 sql.log
SQL_PACKET_SIZE	整数	在 DBMS 和 Web 服务器之间进行数据传输的报文

		大小，以字节为单位
SQL_TRANSLATE_DLL	文件名	一个 DLL 的文件名，文件中包含 SQLDriverToData Source 和 SQLDataSourceToDriver 功能，用于特殊特定的目的，如字符集转换
SQL_TRANSLATE_OPTION	整数	用于控制翻译功能的值
SQL_TXN_ISOLATION	整数	设定事务独立级别 1=读（），2=读（），4=多重读，8=可连续，16=版本
SQL_MAX_LENGTH	整数	从一个字符流或二进制流中返回的最大数据量，用于减少网络流量
SQL_MAX_ROWS	整数	一个 SELECT 语句返回的最多行数，如为 0，则返回所有行
SQL_NOSCAN	整数	当置为 0（默认值）时，检查是否有 escape clause，否则把 SQL 语句直接输出到 ODBC 驱动程序 0=检测并转换，1=不检测，不转换
SQL_QUERY_TIMEOUT	整数	一个 SQL 语句的超时限制，单位为秒，超时后会中止查询 0=无时限

附 2

变量	解释
ALL_HTTP	HTTP_ACCEPT:*/*,q=0.300,audio/x-aiff, audio/basic, image/jpeg, image/gif, text/plain, text/html HTTP_USER_AGENT: Microsoft Internet Explorer/0.1(Win32) HTTP_REFERER: http://webserver/samples/dbsamp/dbsamp3.htm HTTP_CONTENT_TYPE: application/x-www-form-urlencoded HTTP_CONTENT_LENGTH: 10
AUTH_TYPE	认证方法，如果用户名被服务器认证过，则为 Basic
CONTENT_LENGTH	用户送来的字符数
CONTENT_TYPE	为请求为 POST 时使用
GATEWAY_INTERFACE	CGI 的版本号
HTTP_ACCEPT	HTTP 的头，它等于把”accept:”域的值连接起来，中间用”,”分开
LOGON_USER	用户用的 Windows NT 的登录名
PATH_INFO	与 CGI 中的 PATH_INFO 含义相同，表示目录信息
PATH_TRANSLATED	与 PATH_INFO 相同，但把虚拟路径转换为时间路径
QUERY_STRING	在用户送来的请求字符串中”?”后的部分
REMOTE_ADDR	用户端的主机 IP 地址
REMOTE_HOST	用户端的主机名
REMOTE_USER	用户端的用户名
REQUEST_METHOD	服务器收到 HTTP 请求类型，如 GET，POST 等
SCRIPT_NAME	被执行的解释程序的文件名
SERVER_NAME	服务器的主机名或 IP 地址
SERVER_PORT	收到请求的 TCP/IP 口
SERVER_PORT_SECURE	值为 1 表示请求是在加密的口上
SERVER_PROTOCOL	用户请求所使用的 HTTP 协议的版本号，一般为 HTTP/1.0
SERVER_SOFTWARE	Web 服务器软件的名称和版本号
URL	用户送来的请求中的 URL