

【理论与探索】

基于 XML 和 Delphi 工具的异构数据共享*

董殿伟¹, 杨 军¹, 高 辉²

(1. 军械工程学院 装备指挥与管理系维修工程实验中心, 石家庄 050003;

2. 黑龙江科技学院 计算机与信息工程学院, 哈尔滨 150027)

摘要:分析了 XML 在解决异构数据库集成问题中的优势,并在此基础上提出了基于 XML 的异构数据库集成方案,实现了分布式异构数据库的透明访问和联合查询.对于架构中的主要环节给出了具体实现方法,并着重研究和验证了架构中 XML 与数据库模式映射、数据透明访问、联合查询处理等关键问题.

关键词:数据转换;转换接口;关系数据库;XML

中图分类号:TP311

文献标识码:A

文章编号:1006-0707(2008)04-0117-03

数据库应用发展到今天,已有相当数量的企业和科研机构积累了大量的、以不同形式存储的、依赖于不同数据库管理系统的数据库.如何共享这些数据信息,是企业进一步发展所需解决的问题;而如何合理地使用这些异构数据库,也成为亟待解决的问题.异构数据库系统是相关的多个数据库系统的集合,用以实现数据的共享和透明访问.异构数据库系统的异构性主要体现在以下几个方面:

1) 计算机体系结构的异构:运行的硬件环境可以是大型机、小型机、工作站、微机等.

2) 基础操作系统的异构:基于的操作系统可以是 Linux, Unix, Windows 等.

3) 数据库管理系统(DBMS)的异构:可以是使用相同数据模型不同厂商的数据库产品,如 ORACLE, DB2, SQL SERVER 等,它们之间数据的存储模式不同,所使用的数据库处理语言也有区别,也可以使用不同数据模型的数据库,如关系模型、层次模型、网状模型甚至面向对象模型等^[3].

本研究讨论的异构数据库属于第3种类型,即在不同 DBMS 的 2 个异构数据库系统间(Oracle 数据库和 SQL Server 数据库)进行信息交换.

1 XML 简介

XML(extensible markup language)是 1998 年由 W3C 正式发布的,是 SGML 的一个子集.XML 主要用于描述数据文档中数据的组织和安排的结构,只为数据的结构化提供了一套规则.它的关注点不是数据在浏览器中如何布局和显

示,而是数据内容的组织与结构.

XML 具有以下一些特点^[1]:① 具有可扩展性.XML 是一种元标记语言,可以定义其它语言.② 结构性强,易于处理.XML 对格式的定义非常严格,具有良好层次结构.③ 纯文本,与平台无关.XML 将数据保存在标准的文本文件中,可以使用任何文本编辑工具来创建和编辑 XML 文档,无须任何更改可直接移植到其它平台上.④ 数据存储与数据显式的分离.XML 文档定义的仅仅是数据存储的结构和关系,对于数据的显示没有任何的约束和限制,这也是 XML 最值得称道的地方.⑤ 很强的链接能力,可以定义双向链接、多目标链接、扩展链接和 2 个文档间的链接.

2 异构数据转换方案

XML 与数据库之间的转换包括数据库到 XML 文档的转换和 XML 文档到数据库之间的转换,数据库表达的信息是完全结构化的信息,但无法表达一些层次性质的数据,而 XML 文档表达的是半结构化的信息,特别适合表达层次性的数据,在灵活性上更好,正是数据库和 XML 文档存在结构化和半结构化的区别,主要是非层次和层次型的区别导致了两者之间进行数据转化的困难.

本研究的解决方法是建立数据库和 XML 数据之间的映射关系.这种映射关系是双向的^[2].

从数据库数据到 XML 数据的映射时,XML 文档的根节点为数据库名称,根节点的子节点为表名,代表表中的一条记录,表的各子节点代表每条记录的各个字段.图 1 为在

* 收稿日期:2008-04-16

作者简介:董殿伟(1978—),男,辽宁葫芦岛人,硕士研究生,主要从事装备综合保障数据环境研究.

数据库 database 里的表 student 的结构图.

student : 表	
字段名称	数据类型
ID	数字
name	文本
age	数字

图 1 表 student 结构

则对应到 XML 文档如下:

```
<? xml version = "1.0" encoding = "gb2312"?>
<database>
  <student>
    <id>1</id>
    <name>刘小明</name>
    <age>19</age>
  </student>
</database>
```

从 XML 数据到数据库数据映射时,为 XML 根节点的子节点建立一个表,表名为根节点子节点名.为每一个根节点子节点的属性创建一个字段.则以上的 XML 文档对应到数据库表如表 1 所示.

表 1 数据库表

id	name	age
1	刘小明	19

3 异构数据交换系统的框架

根据上述思想,基于 XML 的数据集成模型的框架结构,如图 2 所示.

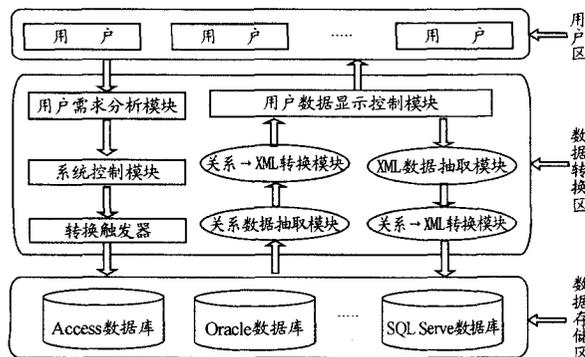


图 2 交换系统框架

用户需求分析模块:主要负责接收用户的需求,进行分析,判断用户是要进行数据存储,还是要进行 XML 与关系数据库的转换,将结果传递到系统控制模块以及全局控

制变量.

系统控制模块:判断系统闲还是忙,如果闲,则提交用户需求,并将系统标志为忙;如果忙,则通知用户需求分析模块,使其等待.

转换触发器:接收用户需求,控制系统是进行数据存储还是进行 XML 与关系数据的转换,并将结果传递给关系数据抽取模块或 XML 数据抽取模块.

关系数据抽取模块:采用 SQL 语言提取异构关系数据库中用户需要的数据,要满足用户的所有需求,将结果传递到关系→XML 转换模块.

XML 数据抽取模块:将用户数据显示控制模块中显示给用户的数据整理成文本形式,将结果传递到 XML→关系转换模块.

关系→XML 转换模块:按照事先规定好的数据映射关系,将关系数据抽取模块提取的结果转换 XML 格式,并存储为文本文件,并传输到用户数据显示控制模块.

XML→关系转换模块:按照事先规定好的数据映射关系,将文本文件转换成熟悉的关系数据库的数据存储方式,并将结果存储到数据库中.

用户数据显示控制模块:将关系数据抽取模块所得关系数据或 XML→关系转换模块所得数据显示给用户.

4 基于编程工具 Delphi 的实例实现

假设数据库 A 为 Access 数据库,内有表 student (ID, name, age),表 stu (ID, score); B 为 Oracle 数据库,内有表 student (S.ID, address); C 为 SQL Server 数据库.现用户需要显示某一学号学生的姓名、年龄、成绩、家庭住址,并将结果存入到数据库 C.

为了实现以上需求,采用开发语言 Delphi.首先在中新建一个工程,放置 3 个 ADOConnection 组件,用于和三个数据库的连接;一个 ADOTable 组件,用于向数据库 C 中存入学生的信息;2 个 ADOQuery 组件,用于从数据库 A 和 B 中查询所需学生信息;一个文本框,用来输入所查学生的学号;3 个标准按钮,分别用来实现查询、保存和退出 3 个功能;一个 DBGrid,用来显示用户所需数据.界面如图 3 所示.

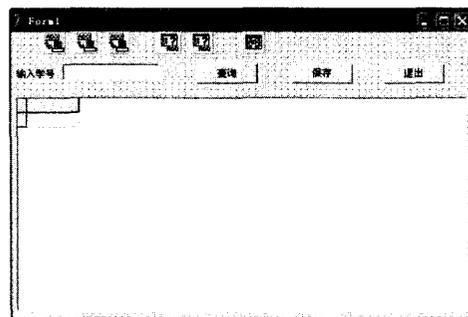


图 3 开发界面

数据库 A 的连接代码为:
ADOConnection1.Provider = 'Microsoft.Jet.OLEDB.4.0';

```
ADOConnection1.LoginPrompt:=false; ADOConnection1.ConnectionString:=extractfilepath(application.ExeName)+'A.mdb';
```

```
ADOConnection1.Connected:=true;
```

数据库 B 与数据库 C 的连接代码与 A 类似.

```
ADOQuery1 与 ADOConnection1 语句为: ADOConnection1.Connection:=ADOConnection1;
```

查询数据库 A 中表 student 中 name 和 age 的 ADOQuery1 代码为

```
With ADOQuery1 do
```

```
Begin
```

```
Close;
```

```
Sql.clear;
```

```
Sql.add('select name,age from studene where id=:a') parameters.ParamByName('a').Value:=edit1.text;
```

```
Open;
```

```
End;
```

ADOQuery2 的连接语句与查询语句和 ADOQuery1 相似.

关系→XML 转换模块的代码:首先定义一个字符型变量 str,再通过程序获得数据库名、表名以及字段名,

```
str:='<? xml Version="1.0"encoding="gb2312"?>
```

```
<A>
```

```
<student>
```

```
<id>1</id>
```

```
<name>刘小明</name>
```

```
<age>19</age>
```

```
<address>北京</address>
```

```
</student>
```

```
</A>'
```

大家注意到,在转换过程中加上了从数据库 B 中取得的家庭住址值 address.

XML→关系转换模块与上一模块类似,是一个反过程,通过判断去掉<? xml Version="1.0"encoding="gb2312"?>以

及标志字符(<)等,取得表名、字段名以及各字段的值赋予给 ADOTable1,最后存入到数据库 C.

由于篇幅有限,各模块代码就不一一详述了.

运行结果如图 4:

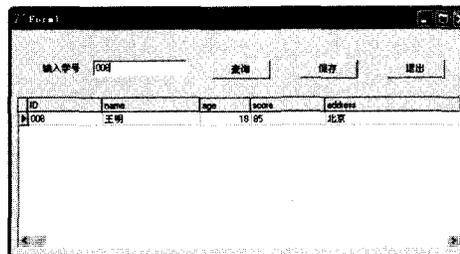


图 4 系统运行结果

5 结束语

本研究是基于 XML 的工具 Delphi 实现的数据交换中间件系统.由于 XML 技术的特性使得它在异构数据库间实现信息的高效传输、交换、共享并保证数据的完整性上发挥着越来越重要的作用.在应用系统中利用上述接口模块进行 XML 数据与关系数据库中表格记录的相互转换可以减少编码的重复.同时由于屏蔽了低层数据库的差异和转换的过程,具有一定的应用价值.

参考文献:

- [1] Peter G Aitken. 微软 XML 技术指南[M]. 北京:中国电力出版社,2003.
- [2] 张家耕,谢晓竹. XML 网络编程技术[M]. 北京:国防工业出版社,2002.
- [3] 梁鹰,罗伟其. 异构数据库的数据转换在大型信息系统中的实现[J]. 计算机工程与应用,2000,36(9):103-105.

(上接第 50 页)

参考文献:

- [1] GJB770A—97[S].
- [2] 傅若农,常永福. 气相色谱和热分析技术[M]. 北京:

国防工业出版社,1989.

- [3] 周起槐. 火药物理化学性能[M]. 北京:国防工业出版社,1983.
- [4] 王泽山,韩盘铭,张续柱. 火药实验[M]. 北京:中国科技出版社,1992.