

利用 Web 服务和 XML 实现数据库同步

刘福顺, 唐宁九

(四川大学 计算机系, 四川 成都 610065)

摘要: XML(可扩展标记语言)以其自描述性和便于传输的优势迅速成为网络上交换数据的存储格式。Web 服务的应用使得 Web 应用程序可以透明地访问服务, 请求参数和响应结果是被序列化进入 SOAP 消息中的。结合 Web 服务和 XML 技术的各自特点, 提出了利用 Web 服务和 XML 实现数据库同步的策略。

关键词: Web 服务; XML; 数据库同步

中图法分类号: TP311.131

文献标识码: A

文章编号: 1001-3695(2005)02-0051-02

Using Web Service and XML to Achieve Database Synchronization

LIU Fu-shun, TANG Ning-jiu

(Dept. of Computer Science, Sichuan University, Chengdu Sichuan 610065, China)

Abstract: XML (eXtensible Markup Language) becomes the store format of data exchanged on the Web for its self-describe and easy to be transmitted. The use of Web service allows Web application to invoke services transparently, with the request parameters and response results to be serialized into SOAP message. Combining features of Web service and XML technology, this paper proposes a strategy of using Web service and XML to achieve database synchronization.

Key words: Web Service; XML; Database Synchronization

随着 Internet 的迅速发展, 大量的基于 C/S 的传统应用系统都转型到了基于 B/S 的三层系统, 这就是所谓的客户端瘦身运动。特别是以产品销售为主要业务活动的商业性质的公司, 都完全转变了运营模式, 分别搭建起了各自的电子商务平台, 图 1 是这样的商务电子平台的典型模型。该模型通常是由两台服务器构成, 即本地服务器是面向公司工作人员的管理服务器; 远程服务器是面向公众用户的公共服务器。通常两台服务器的物理位置是不同的, 本地服务器通常位于公司的局域网, 而远程服务器通常是由服务提供商进行托管。两台数据库服务器使得公司能在与远程服务器断开的状态下仍然能正常运营, 然而, 这也同时带来了如何迅速、高效、准确地将本地数据库的更新操作同步到远程数据库的问题。

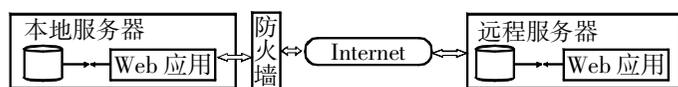


图 1 电子商务典型模型

1 背景知识

1.1 XML

XML 是 SGML 的一个简化标准版本。XML 不是一种编程语言, 而是一种数据的描述语言, 适合于存储半结构化的数据。下面是一个典型的 XML 文档实例。

```

<?xml version="1.0"?>
<books>
  <book>
    <bookname>
      XML: The Complete Reference
    </bookname>
    <author>Heather Williamson</author>
  </book>
</books>
  
```

```

</book>
</books>
  
```

一个 XML 文档通常是由 XML 声明、元素、属性和文本构成。其中文本是真正要存储的数据, 而元素和属性通常是用来描述它所包含的文本数据的性质的。XML 的可扩展性即表现在元素和属性是可以根据需要任意指定的(这完全不同于 HTML 中规定好了的标记)。正是 XML 的这种可扩展性使得 XML 文档是自描述的, 既便于人的阅读, 又便于机器的处理。因此, XML 已成为异构环境和网络环境中数据交换的载体。

由于 XML 本身是一个文本文件, 相对于关系数据库而言具有体积小、便于传输的优势。任何文本传输协议均可快速准确地在网上上传 XML 文档。同时, XML 的文本本质使其独立于特定的平台, 在异构环境下能起到数据的中间存储格式的角色。XML 应用的优势不仅仅取决于其自身的体积小、自描述和便于传输的特性, 大量围绕 XML 的相关技术的完善也大大加强了 XML 的应用范围。例如, XSL 可以将 XML 文档转换成其他格式的数据形式, 如将 XML 转换成 HTML; DOM 和 SAX 分别为操作 XML 文档提供了 API 函数; 高级的软件开发平台, 如 VS.NET, 专门提供了 System XML 的命名空间, 封装了底层的 API 函数, 使得基于 XML 的开发变得容易。本文提出的解决策略的实现就是利用了 VS.NET 平台的 System XML 命名空间提供的操作。

1.2 Web 服务

Web 服务的出现给 B/S 三层模型注入了新的活力。Web 服务的基本思想是: 服务器将对象发布为 Web 服务, 客户端可向该服务发出请求, Web 服务响应该请求, 并在服务器端执行相应的方法再将结果返回客户端调用者。该过程对于客户端调用者来说是透明的。请求的发出和响应的回送是由 Web 服

务的底层 SOAP(简单对象访问协议)支持的。

客户端代理会将请求及参数序列化为 SOAP 消息,并通过 HTTP 发送到 Web 服务所在的服务器。服务器从接收到的 SOAP 消息中获得服务方法名及参数值,并执行该方法,之后将结果再序列化到 SOAP 消息并通过 HTTP 回送到客户端调用者。客户端代理从 SOAP 消息中取出结果并上传给应用程序。可以看出 Web 服务的作用在于将客户端的请求参数带到服务器端执行,并将结果返回到客户端,而该过程对客户端是透明的,不需要客户端应用程序的介入。

1.3 数据库同步的方法

(1) SQL Server 2000 中的复制技术

SQL Server 2000 中的复制技术是基于发布者/订阅者模型的。基本思想是将同步的两个或多个数据库服务器中的一个作为发布者,其他数据库服务器作为订阅者。订阅者可以向发布者进行订阅,订阅成功后,订阅者就可以定期检查发布者的数据库同自己的订阅数据库是否一致,如果不一致则自动进行同步。该技术的优势在于同步过程的自动完成;缺点在于该同步过程的本质是 C/S 模式的,是两台或多台 SQL Server 数据库服务器之间的通信,适合局域网范围内的应用,难以穿透防火墙,不适合 Internet 中 Web 站点数据库服务器同位于公司内部本地数据库之间的数据库同步。

(2) ADO 或 ADO.NET 技术

可以直接通过 ADO 或 ADO.NET 技术远程连接 Web 站点数据库服务器,并将本地数据库的更新操作作用于远程数据库。该方法的优势在于操作直接;不足在于数据在本地数据库与远程数据库之间的往返受到网络因素的制约,同时也会消耗远程数据库服务器的资源(需要保持对数据库连接的长时间打开状态),并使网络流量加大。

鉴于上述两种方法的不足之处,我们提出了利用 Web 服务和 XML 技术的数据库同步的策略,该策略有效地解决了上述问题。

2 利用 Web 服务和 XML 技术实现数据库同步

2.1 总体模型

图 2 给出了利用 Web 服务和 XML 技术实现数据库同步策略的总体模型。

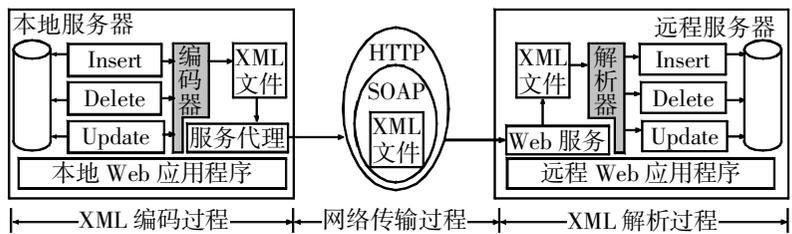


图 2 总体模型

整个模型由三部分组成:本地 XML 编码过程、网络传输过程、远程 XML 解析过程。由于引入了 Web 服务,本地 XML 编码过程的输出——XML 文件的内容(文本字符串)是作为 Web 服务的请求参数由 SOAP 消息带到远程服务器,并作为远程 XML 解析过程的输入。

2.2 本地 XML 编码过程

本地 XML 编码过程的目的是将对本地数据库的更新操作

(Insert, Delete, Update)及所涉及的记录值序列化到本地的一个 XML 文件中,该 XML 文件从性质上讲类似于日志文件,它完全记录了增(Insert)、删(Delete)、改(Update)各项操作。由于查询(Select)不会造成本地同远程数据库的不同步,因此,不需要将查询操作序列化到 XML 文件中。

为了能将影响数据库的三种操作序列化到 XML 文件中,应首先定义 XML 文件的结构。下面是我们给出的 XML 文件的 DTD(文档类型定义):

```
<! DOCTYPE COMMANDS [
<! ELEMENT COMMANDS ( COMMAND ) >
<! ELEMENT      COMMAND
( COMMANDNAME, TABLENAME, VALUE* , CON DITION* ) >
<! ELEMEN  VALUE ( COLUMNNAME, COLUMNVALUE) >
<! ELEMENT  COMMANDNAME ( #PCDATA) >
<! ELEMENT TABLENAME ( #PCDATA) >
<! ELEMENT COLUMNNAME ( #PCDATA) >
<! ELEMENT COLUMNVALUE ( #PCDATA) >
<! ELEMENT CONDITION ( #PCDATA) >
] >
```

有了 XML 文件的 DTD,我们就可以将 Insert, Delete 和 Update 三种操作序列化到 XML 文件了。图 3 分别给出三种操作的通用序列化方式,从中可以看到左侧的 SQL 语句和右侧的 XML 节点是等价的。由 SQL 语句向 XML 节点的转换是由下面提到的 XML 编码器实现的,而从 XML 节点向 SQL 语句的转换是由 XML 解析器完成的,也就是说转换是可逆的。

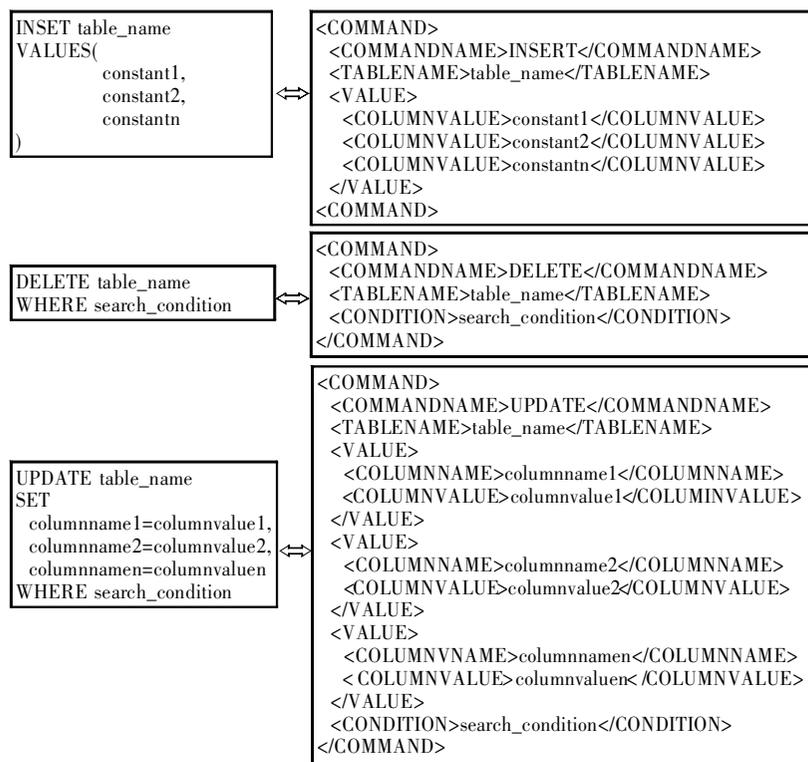


图 3 转换示意图

定义了三种操作的序列化方式后,就可以给出编码器了。我们采用 C# 实现了该编码器(第 3 节将提到该编码器并不完善)。该编码器的作用是接收一个 SQL 语句字符串,依据上面给出的转换方式编码为 XML 节点并写入本地的 XML 文件中。

当需要对数据库进行同步的时候,就会调用远程服务器提供的 Web 服务。该 Web 服务接收的参数就是本地的这个 XML 文件的内容。Web 服务将该 XML 文件的内容作为一个字符串带到远程服务器,由 Web 服务实现在远程服务器上进行处理,处理完成后回复一个信号给本地调用者,调用者接收到该信号后知道同步已经完成,并将本地 XML 文件内容清除,准备写入新的内容。

2.3 远程 XML 解析过程

远程 XML 解析过程就是 Web 服务的具体实(下转第 56 页)