

基于XML的岸船数据同步方案设计

魏江涛,黄爱军,周 勇

(中国卫星海上测控部,江苏 江阴 214431)

摘要:随着航天测控领域的系统应用不断深入发展,需要通过分布式部署来进行管理,即对岸站和测量船分别进行部署。为对多个信息管理系统的数据进行同步,通过分析XML和Windows服务的特点,提出了将二者结合起来解决分布式部署应用系统的数据同步的解决方案,有效地利用有限带宽解决了岸船数据同步问题。该设计方案在岸船业务应用系统中经过了实验验证,验证效果良好,达到了预期效果。

关键词:可扩展标记语言;触发器;数据同步;Windows服务;分布式部署

中图分类号:TP311.1

文献标识码:A

文章编号:1673-629X(2014)03-0049-05

doi:10.3969/j.issn.1673-629X.2014.03.013

Design of Data Synchronization Scheme Based on XML

WEI Jiang-tao, HUANG Ai-jun, ZHOU Yong

(Department of Chinese Maritime Satellite Monitoring and Control, Jiangyin 214431, China)

Abstract: With the in-depth development of system application in aerospace measurement and control field, need to manage through the distributed deployment, namely to deploy landed station and vessel respectively. For synchronization of multiple information management system data, through the analysis of the characteristics of XML and Windows Service, put forward a solution to combine the two for solving the data synchronization of distributed deployment application system, efficiently using limited bandwidth to solve the landed ship data synchronization problem. The design is tested in onshore ship business application system, and the test result is good, reaching the desired effect.

Key words: eXtensible Markup Language (XML); trigger; data synchronization; Windows Service; distributed deployment

0 引言

随着海上航天测控技术的不断发展,各种业务应用系统不断出现,如设备管理信息系统、综合训练信息管理系统等。为了管理这些不断深入发展的系统应用,需要通过分布式部署,即对岸站和测量船分别进行部署。

由于测量船经常出海执行任务,所以对这些应用系统来说就存在一个共性问题:岸站应用系统和测量船应用系统之间的数据如何进行同步。同时,由于测量船出海之后通信带宽有限,所以应用系统数据库本身所具有的“出版和订购”的同步方式无法实现岸船业务系统数据库之间的数据同步。

针对以上数据同步问题,文中提出了一种基于XML文件和Windows服务的数据同步方案,有效地利用有限带宽解决了岸船数据同步问题。

1 基本概念

1.1 XML

XML(eXtensible Markup Language)是由多家软件公司共同制定的一种通用语言规范,是一种可扩展性标记语言。XML与其他语言相比,具有明显的优势,XML的特点主要有^[1-3]:

(1)XML具有高度结构化和层次化的数据组织形式。XML允许开发者建立它们自己的、用于存储信息的标记结构。在HTML中数据是没有含义的,用户无法知道某个数据具体代表什么,但是在XML文档中人们可以使用特定的标记为数据定义相关的语义,其数据内容与格式是分离的。

(2)XML格式以文本为基础,建立在Unicode基础上,这使它能更容易创建国际化的文档,具有良好的可读性。所以通过XML文档可以使数据被更多的用户、

更多的设备所读取利用。

(3)XML 具有开放性、与平台无关性、灵活的可扩展性,可用于交换数据。通过 XML,可以在不兼容的系统及数据库之间交换数据,将大大减少交换数据时的复杂性,并且还可以使得这些数据被不同的程序读取。因此用 XML 进行复杂的信息流交换比用其他格式进行信息交换更简单。

基于 XML 的这些优势,XML 目前已成为信息描述的事实标准以及成为不同应用之间数据交换的主要方式,具有非常广阔的应用前景^[4]。

1.2 Windows 服务

Windows 服务即 Windows Service,是 Windows 操作系统中一种长期运行在后台完成特定功能的可执行的应用程序。它们长期在后台运行,没有用户界面,默默无闻,但它们却是支持 Windows 正常运行的幕后英雄。Windows 服务程序没有用户界面,程序员可以编写一系列的方法来完成其特定的服务功能,这样运行在后台就不会受到用户界面对其产生的影响。

Windows 服务程序虽然是可执行的,但是它不像一般的可执行文件通过双击就能开始运行了,它必须有特定的启动方式。这些启动方式包括了自动启动和手动启动两种。对于自动启动的 Windows 服务程序,它们在 Windows 启动或是操作系统重启之后用户登录之前就开始执行了,而对于手动启动的 Windows 服务程序,你可以通过命令行工具或是控制面板中的管理工具来启动它,通过这两种方式启动服务后,该服务就一直在后台运行。

Windows 服务程序的应用范围很广,典型的 Windows 服务程序包含了硬件控制、应用程序监视、系统级应用、诊断、报告、Web 和文件系统服务等功能。

该方案正是基于 Windows 服务的这些优点,在 XML 数据文件的生成和解析过程中使用 Windows 服务。

1.3 触发器

触发器(Trigger)是用户定义在数据库关系表上的一类由事件驱动的特殊存储过程,触发器由触发事件或语句、触发条件、触发动作体三部分组成^[5]。触发条件是一个逻辑表达式,用于判断触发器是否触发。触发动作体是一个 PL/SQL 语句块,是触发器触发后执行的语句块,这个语句块也可以是对已创建存储过程的调用。每当一个特定的数据操作语句在指定数据表上发生时,当产生的逻辑值为真时就会引发触发器的执行,就会执行触发动作体中的 PL/SQL 语句块。

数据库允许创建 INSERT、UPDATE、DELETE 触发器;当 INSERT 触发器被触发时,系统将会建立一个名为 Inserted 的逻辑表,被插入的数据行会被复制到 Inserted 中;当 DELETE 触发器被触发时,系统将会建立一个名为 Deleted 的逻辑表,被删除的数据行会被复制到 Deleted 中;当 UPDATE 触发器被触发时,系统将会分别建立名为 Deleted、Inserted 的逻辑表,更新前的数据行会被复制到 Deleted 中,更新后的数据行将被复制到 Inserted 中^[6]。

该方案在设计时对应用系统需要同步的数据表的 INSERT(插入)、UPDATE(更新)和 DELETE(删除)操作建立触发器。每当用户增加、修改或删除数据表中的数据时,将自动触发对应的触发器,在触发器中编写 SQL 代码,把变动的数据保存到一个专门记录触发事件的数据表中,以供 XML 数据文件生成服务进行处理。

2 岸船应用系统数据同步的设计方案

基于 XML 和 Windows 服务的特点,比较适合解决系统间的数据同步问题^[7],文中提出了一种基于 XML 和 Windows 服务的岸船应用系统数据同步方案(如图 1 所示)。

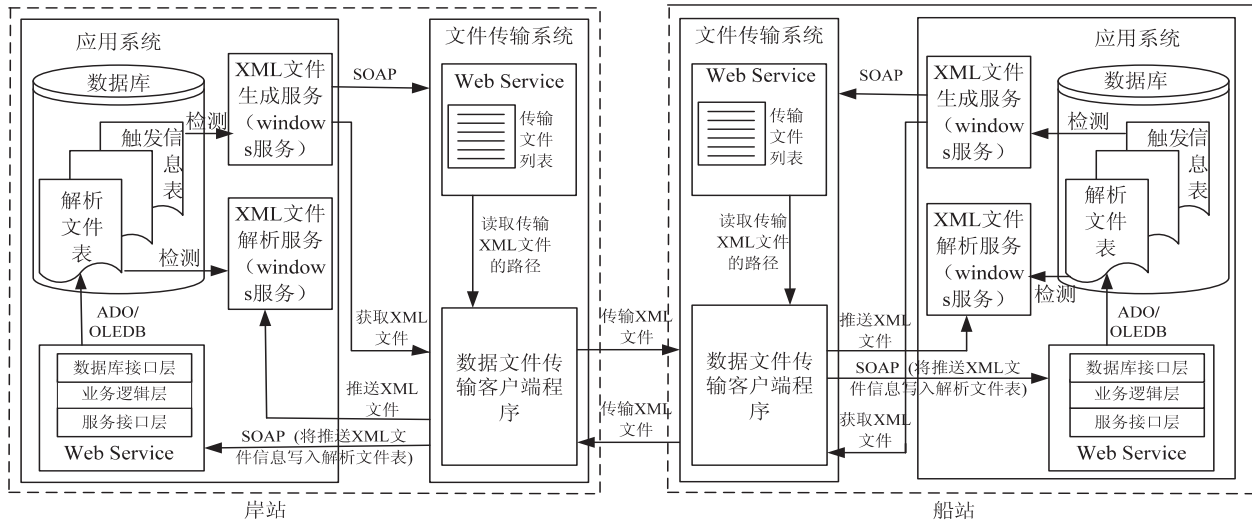


图 1 数据同步设计方案

方案主要由岸站应用系统、XML 文件生成服务(Windows 服务)、XML 文件解析服务(Windows 服务)、文件传输系统、船站应用系统五部分组成。下面详细介绍方案的设计:

2.1 应用系统数据库触发器及数据表的设计

(1) 在应用系统数据库中增加一张触发信息表, 这张触发信息表的主要字段有: 发生信息变更的表名、发生信息变更的主键、操作方法(是指该记录变更时是增加、编辑还是删除)、记录发生变更的时间。当应用系统的数据表记录发生变更以后, 触发事件发生, 将发生变更记录的信息写入这张触发信息表中。供接下来的 XML 文件生成服务检测和生成相应 XML 记录文件。

(2) 在需要数据同步的应用系统数据表中建立触发器。触发器中主要工作是当这张数据表信息发生变更时将该数据表的信息: 表名、主键、操作方法、发生变更的时间写入(1)中建立的触发信息表。

在对数据库表创建触发器时, 某些数据表建立的触发器船站和岸站是不同的, 因为某些信息是单向的, 只是岸站到船站, 或船站到岸站, 目的方站数据库不需要触发。所以在建立触发器时船站和岸站数据库触发器需要分别创建。

(3) 在应用系统的数据库中再增加一张解析文件数据表。解析文件数据表主要是记录文件传输系统推送到应用系统的文件信息(包括文件存放的目的路径、文件名称、文件大小、存放时间以及文件是否被解析的状态标记)。该数据表主要是供 XML 文件解析服务检测和解析使用。

2.2 XML 文件生成服务

XML 文件生成服务是运行在计算机后台的 Windows 服务, 它定时检测应用系统数据库触发信息表, 检测触发信息表是否有新的记录。当应用系统的数据库某个数据库表发生变更时, 触发器触发, 触发信息表记录增加, XML 文件生成服务(Windows 服务)检测到数据变更, 则 XML 文件生成服务开始执行, 完成以下功能:

(1) 将触发信息表中增加的记录信息即变更信息生成 XML 文件。这样每一条数据库原生操作都由 XML 文件生成服务生成一个 XML 记录文件。

(2) XML 文件生成服务调用外部的文件传输系统提供的 Web Service 接口将生成的 XML 文件信息(主要包括文件存放路径、文件大小、文件名称、目的地标识和目的地存放路径)写入文件传输系统的传输文件列表。

对于一般的数据表, XML 文件生成服务做以上两步处理即可, 但对于附件表和走流程的数据表, XML

文件生成服务需要做特殊处理。

①附件表的特殊处理。

当用户在应用系统中提交附件后, 应用系统数据库的附件表增加了一条附件记录, 但同时应用系统也多了一个附件。所以在对附件进行岸船数据同步时不仅要同步数据库记录, 还要同步附件。因此 XML 文件生成服务如果检测到变更信息是附件表时首先将变更信息的数据记录生成 XML 文件, 而后调用文件传输系统提供的 Web Service 接口将生成的 XML 文件信息以及上传附件的文件信息同时写入传输文件列表。这样在附件的同步中不仅将附件数据记录同步到目的方, 同时也将相应的附件同步到目的方。

②流程的处理。

流程是指应用系统的某项功能可能经过好几个过程, 这几个过程由不同的人来完成, 如设备管理信息系统中的设备采购, 首先由基层部门申请采购物品种类和数量, 而后由上级机关进行审核和确认。

由于流程的处理由多个不同的人来共同完成, 所以每发起一个流程在整个应用系统中只能有一个唯一的主键 ID 号, 这样才能保证应用系统流程数据记录的完整性和正确性, 才能保证在分布式部署应用时不发生逻辑错误。

下面以设备管理信息系统中的设备采购流程为例进行说明。由于标识某一个流程唯一的主键 ID 号是动态生成的, 所以针对此情况, 流程的处理过程如下:

首先某个基层单位如船站在应用系统中填写采购设备的基本信息, 填写完毕后, 提交给船站服务器, 提交后船站应用系统并不直接将该数据记录插入到船站服务器的数据库中, 而是直接生成相应的 XML 数据记录文件, 由岸船之间的文件传输系统将 XML 数据记录文件传回岸站主服务器, 由岸站主服务器的 XML 解析服务将 XML 文件解析, 解析后, 将采购信息插入岸站主服务器数据库, 此时这个流程的唯一主键 ID 产生, 岸站主服务器数据库信息插入成功后, 该数据表的触发器触发, 进而由 XML 文件生成服务生成该记录的 XML 文件, 再由岸船之间的文件传输系统将此 XML 文件传输到船站, 船站的 XML 解析服务将 XML 文件解析写入船站的应用系统的数据库, 这样就完成了走审批流程的信息处理, 保证了流程信息的完整性和符合业务逻辑。

2.3 文件同步传输

文件传输系统是外部应用系统, 一般是基于 C/S 结构的应用程序, 文件传输系统主要由数据文件传输客户端程序和文件传输列表组成。数据文件传输客户端程序根据传输文件列表获取要传输的文件, 根据传输列表中的目的地将文件传送到相应的地方。

在岸站与船站的通信中,岸站对船站文件传输不仅有一对一的文件传输,可能也有一对多的文件传输,比如一个文件同时需要传输到三条船站,所以文件传输系统根据传输文件列表中目的地标识来决定文件需要传输到哪个船站或哪些船站^[8-10]。

从图 1 中可看出,岸站文件传输到达目的端后并不是直接到船站应用系统,而是到船站文件传输系统,船站文件传输系统客户端程序根据收到文件的目的地存放路径将文件推送到相应的目录路径下,推送完成后,船站文件传输系统客户端程序调用船站应用系统提供的 Web Service 接口将推送的文件信息(包括文件存放路径、文件名称、文件大小、存放时间以及文件是否被解析的状态标记)写入船站应用系统数据库的解析文件数据表中。

2.4 XML 文件解析服务

文件传输系统完成文件的传输后接着就是 XML 文件的解析。XML 文件解析服务(Windows 服务)定时扫描应用系统的解析文件数据表,当扫描到新记录时,XML 文件解析服务根据解析文件数据表的记录找到相应的 XML 文件,并将 XML 文件内容解析到应用系统数据库中,同时将解析文件数据表中状态改变,标识为已解析。

同样对于 XML 文件解析服务,附件也要做特殊处理:如果 XML 文件解析服务遇到的文件是附件时,则不解析,将附件文件拷贝到相应的附件目录下即可。

由于岸站和船站在文件传输中文件可能不能做到 100% 的可靠传输,中间可能偶尔有丢包的现象,为此,为了确保数据解析的可靠,XML 文件解析服务对于插入和更新操作做了以下处理:

如果 XML 文件中记录的操作方法是插入操作则首先在数据库中查找该条记录是否存在,如果存在则不插入,不存在则插入。如果记录的方法是更新操作则同样在数据库中查找,如果有该条记录就更新,如果没有该条记录就直接进行插入操作。

这样需要同步的数据记录一旦从源站到达目的站,目的站的应用系统就会将数据呈现出来,达到数据的同步更新。

上述描述的就是岸站和船站的数据同步过程,对于岸站到岸站或船站到岸站的数据同步及实现是类似的。

2.5 方案的实现及验证结果

上述设计方案设计完成后在设备管理信息系统中进行验证。选择设备管理信息系统的周运行情况功能进行了实验验证。设备管理信息系统数据库采用的是 SQL Server 2005,在船站的应用系统数据库的周运行情况数据表建立触发器,触发器 SQL 语句如下:

```
Create TRIGGER [ Trigger_ZbglOpinionReturn ] on ZbglOpinionReturn
after insert, update, delete
as
BEGIN
--//触发删除
IF Exists(Select 0 From Deleted) AND NOT Exists(Select
0 From Inserted)
Begin
insert
Tongbu_Table( tablename, ..., operatormethod)
values( 'ZbglOpinionReturn', ..., delete')
End
--//触发插入
Else IF NOT Exists(Select 0 From Deleted) AND Exists(Select
(Select 0 From Inserted)
Begin
insert
Tongbu_Table( tablename, ..., operatormethod)
values( 'ZbglOpinionReturn', ..., 'insert')
End
--//触发更新
Else IF Exists(Select 0 From Deleted) AND Exists(Select
0 From Inserted)
Begin
insert
Tongbu_Table( tablename, ..., operatormethod)
values( 'ZbglOpinionReturn', ..., 'update')
End
END
```

XML 文件生成服务和解析服务这两个 Windows 服务采用 Visual Studio 2008^[11-12] 进行开发,开发完成后将两个服务分别部署在岸站服务器和船站服务器,在船站应用系统中录入周运行情况后经过中间的处理环节,最后同步到岸站服务器的数据库,岸站的设备管理信息系统进而在页面中显示出同步后数据,实验效果良好。

3 设计方案的特点

(1) 岸船间的数据同步可以做到准实时,自动化运行。由于 XML 文件生成服务和 XML 文件解析服务两个 Windows 服务始终在后台运行,文件传输客户端程序也始终在运行,所以对于变更的数据信息记录所生成 XML 文件、传输 XML 文件以及解析 XML 文件可以做到准实时,不需要人工干预。

(2) 由于各个业务系统对于岸站和船站数据同步所面临的问题是共性的,所以上述方案不受业务应用系统的不同的制约,其他的业务应用系统也适用。

(3) 该方案无论是在有线网络还是无线网络情况下即离岸还是靠岸都适用,不受环境变化的影响,不会因为线缆连接的改变而更改设置。

(4) XML 文件生成服务和解析服务不会受到人为误操作而关闭。方案在设计时,对于生成和解析的处理如果设计成客户端程序可能会受到人为因素而误关闭,导致数据无法同步,采用 Windows 服务很好地避免了这一问题。

4 结束语

中国卫星海上测控部经过多年的信息化建设,目前已有多个管理信息系统,文中主要研究了如何将这多个管理信息系统数据进行同步,以及如何利用有限带宽解决岸船数据同步问题。实践表明,通过岸船间的数据同步使得这些信息系统的管理效率大大提高。

由于岸船数据传输的特殊性,测量船出海后,主要靠通信卫星通信,中间传输文件偶尔可能会丢失,所以下一步可以对方案再进行优化,使得数据同步真正做到稳妥可靠。

参考文献:

[1] World wide web consortium,extensible markup language(X-ML) 1.0, 3rd[EB/OL]. 2004. <http://www.w3.org/TR/>

(上接第48页)

4 结束语

目前的大部分文本分类系统中采用统计和机器学习的方法,这类方法在词典语义的水平上来分析文本内容,并且判断其相似度,从而得到文本类别的划分,因此其文本分类的准确率存在瓶颈。文中设计了一种新的用于文本分类的概率分类器,该分类器在产生规则的基础上计算正、负概率值,然后得出分类结果,从而提高文本分类的有效性。

参考文献:

[1] 薛为民,陆玉昌. 文本挖掘技术研究[J]. 北京联合大学学报(自然科学版),2005,19(4):59-63.

[2] Sebastiani F. Machine learning in automated text categorization[J]. ACM computing surveys,2002,34(1):1-47.

[3] Yang Y,Liu X. A re-examination of text categorization methods[C]//Proceedings of 22nd annual international ACM SIGIR conference on research and development in information retrieval. Berkeley:[s. n.],1999:42-49.

[4] 吴晓洲,万里明,韩霄松,等. 基于隐马尔可夫模型的转录因子文本挖掘算法[J]. 吉林大学学报(理学版),2012,50(2):320-322.

[5] 艾海麦提江·阿布来提,吐尔地·托合提,艾斯卡尔·艾

2004/REC-xml-20040204/.

[2] XML:Binary characterization working group public page,world wide web consortium[EB/OL]. 2005. <http://www.w3.org/XML/Binary/>.

[3] 东方人华.XML3.0 技术内幕[M]. 北京:清华大学出版社,2001.

[4] Miklau G,Suciu D. A formal analysis of information disclosure in data exchange[J]. Journal of computer and system services,2007,73(3):507-534.

[5] 徐国华,白凤娥. 数据库中的触发器技术[J]. 太原大学学报,2006,7(1):64-65.

[6] 陈绍钧. 新编 SQL Server 2005 数据库管理入门与提高[M]. 北京:人民邮电出版社,2007.

[7] 刘福顺,唐宁九. 利用 Web 服务和 XML 实现数据库同步[J]. 计算机应用研究,2005,22(2):51-52.

[8] 刘杰,徐立臻. XML 在移动数据库同步服务中的应用研究[J]. 计算机工程,2007,33(13):81-83.

[9] 沈敏,许华虎,季永华,等. 基于 XML 的分布式异构数据库数据同步系统的研究[J]. 计算机工程与应用,2005,41(5):184-186.

[10] 鱼滨,郑娅峰. 基于 XML 的异构系统集成框架的研究[J]. 计算机应用与软件,2005,22(7):14-15.

[11] 刘丽霞. C#范例开发大全[M]. 北京:清华大学出版社,2010.

[12] 李文强. C#程序设计[M]. 北京:清华大学出版社,2010.

木都拉. 基于 Naive Bayes 的维吾尔文文本分类算法及其性能分析[J]. 计算机应用与软件,2012,29(12):27-29.

[6] 张玉峰,何超. 基于领域本体的语义文本挖掘研究[J]. 情报学报,2011,30(8):832-839.

[7] 田丰,桂小林,杨攀,等. 采用类别相似度聚合的关联文本分类方法[J]. 西安交通大学学报,2012,46(12):6-11.

[8] Yang Xiquan,Sun Na. The application of latent semantic indexing and ontology in text classification[J]. International journal of innovative computing, information and control,2009,5(12):1-9.

[9] 卢中宁,张保威. 一种基于改进 TF-IDF 函数的文本分类方法[J]. 河南师范大学学报(自然科学版),2012,40(6):158-160.

[10] Vries A,Mamoulis N,Nes N,et al. Efficient k-NN search on vertically decomposed data[C]//Proceedings of the ACM SIGMOD conference on management of data. [s. l.]:[s. n.],2002:322-333.

[11] 李荣陆,胡运发. 基于密度的 kNN 文本分类器训练样本裁剪方法[J]. 计算机研究与发展,2004,41(4):539-545.

[12] Liaw Y,Leow M,Wu C. Fast exact k nearest neighbors search using an orthogonal search tree[J]. Pattern recognition,2010,43(6):2351-2358.

基于XML的岸船数据同步方案设计

作者：[魏江涛](#)，[黄爱军](#)，[周勇](#)，[WEI Jiang-tao](#)，[HUANG Ai-jun](#)，[ZHOU Yong](#)

作者单位：[中国卫星海上测控部](#)，[江苏 江阴](#)，[214431](#)

刊名：[计算机技术与发展](#)

英文刊名：

Computer Technology and Development

ISTIC

年，卷(期)：

2014(3)

本文链接：http://d.wanfangdata.com.cn/Periodical_wjfz201403013.aspx