

Research of Heterogeneous Data Integration in Ship Design

An Yin^{1, a}

¹ China Ship Development and Design Center, Wuhan 430064, China

^aibm_hust@yahoo.com.cn

Keywords: heterogeneous; integration; XML; SQLToXML; XSL

Abstract. The China Ship Development and Design Center has built several separate database systems in the design center and the working groups. However, the data cannot be shared and exchanged due to no network connection as the secrecy provision limited. The paper presents the data integration method based on the XML to implement the data integration between the heterogeneous databases, which makes the data interactive convenient. The key technologies included the mapping from database to XML documents, the XML documents importing to database, querying and displaying XML documents were analyzed. The SQLToXML and XMLToSQL libraries are used to develop the program. The method has put into test, the result shows the design scheme is rationality and effective.

Introduction

The ship design includes the general project design phase, the construction detail design phase, the cooperation for manufacture phase and the cooperation for sea trial phase. The design center takes charge of the general project and most of the construction detail design. The working groups in the shipyard are responsible for solving the problems in the construction phase and modifying the construction drawings according to the practical situation. The duty of the working groups in the ship testing base is to modify the construction drawings and complete the sea trial documents according to the sea trial results. Some important information is stored in the databases of the working groups and is needed to feed back to the design center. For example, the monthly summary information about the modified drawings and the problem-solving rate of progress is important to the project control department and the officers who are in charge of the ship-building. Meanwhile, the work statistics from working groups is also important to optimize the ship design. The working groups also need to get the necessary information from the design center.

Currently, the information system of the design center includes the Product Data Management (PDM) and the Office Automation (OA). The documents and drawings in the general project design phase and construction detail design phase are managed by the PDM. The PDM also covers the management of the work plan and the modification requisitions. The working groups in the shipyard and the ship testing base use the information system named "Information Integrated Manage System for Cooperation for Ship Manufacture and Testing" to record the working information and manage the work flow [1]. Because of the different development period of these information systems, there are some problems such as the heterogeneous databases, lack of the unified data management and so on. In the other hand, there is no network connection from the design center to the working groups due to the secrecy provision limited. Therefore, the data in the design center and the working groups is isolated from each other.

In this paper, a data integration method is presented for solving the above problem. The method is based on the XML documents to implement the data integration between the heterogeneous databases. The SQLToXML and XMLToSQL libraries provided by the Websphere Application Developer (WSAD) are used to develop the program. The method makes the data interactive between the design center and the working groups convenient and effective.

Data integration method

Problem analysis and solutions. Some data integration methods have been proposed to integrate the data from different data sources and heterogeneous platforms [2-5]. The data pipeline of the PowerBuilder is a traditional way to backup and update multi-databases. The mainstream trend is to integrate heterogeneous data based on SOA and Web service, e.g. constructs a central database to synchronize and exchange the data to the corresponding databases. Considering there is no network connection between the design center and the working groups due to the secrecy provision limited, the above methods are not applicable. Using the XML document as the data carrier to implement the data integration is an effective solution for the special situation. XML is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. XML as the de facto standard for data exchange makes data integration easier and is playing an increasingly important role in the exchange of heterogeneous data [6-7]. The work flow diagram of the data integration method based on XML is shown in Fig. 1. The key technologies will be discussed in the rest of the paper.

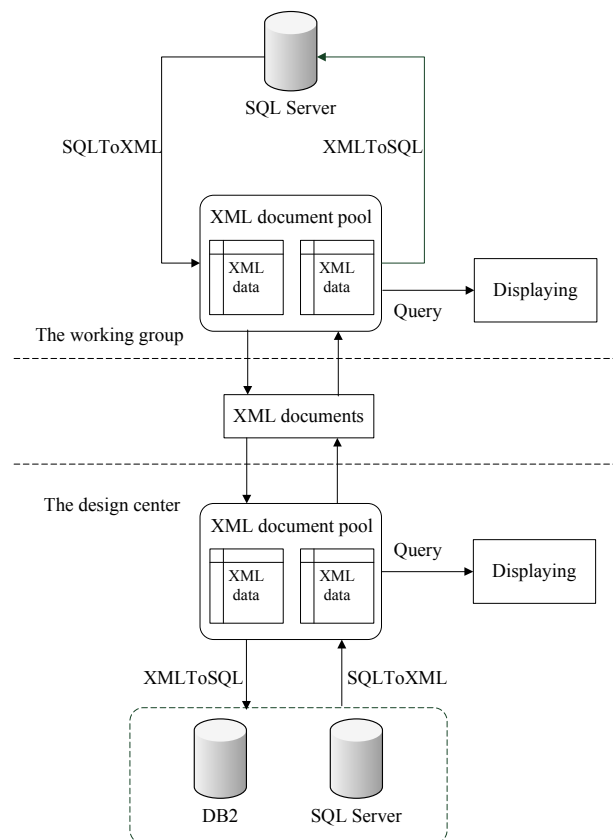


Fig.1 The work flow diagram of the data integration method

Data conversion to XML documents. According to the solution presented above, the first step is to convert the data from the source databases to XML documents. The SQLToXML library provided by WSAD gives a quick and effective way to implement the conversion. The WSAD IDE provides the “SQL to XML guide” to generate the xst file which is used by SQLToXML library to get the necessary information for the conversion such as the account information and the driver of the source database, the SQL statement for getting the data from source database and so on. The next codes show the main content of the xst file which is used to get the monthly modified drawings information from the DealHQDTable table.

```
<SQLGENERATEINFORMATION>
<DATABASEINFORMATION>
// setting up the account information for logging to the source database
<LOGINID>...</LOGINID>
```

```

<PASSWORD><![CDATA[...]]></PASSWORD>
//setting up the driver information
<JDBC_DRIVER>com.sqlserver.jdbc2.jdbc.SQLDriver</JDBC_DRIVER>
<JDBC_SERVER>jdbc:microsoft:sqlserver://127.0.0.1:1433;DatabaseName=PJGZZ
</JDBC_SERVER>
</DATABASE_INFORMATION>
<STATEMENT>
    //setting up the SQL statement for getting data from the source database
    <![CDATA[ SELECT * from DealHQDTable datediff(mm,shijian,getdate())=0]]>
</STATEMENT>
.....
</SQLGENERATE_INFORMATION>

```

We can get the different data from the source database by writing the correlative SQL statement between <STATEMENT> and </STATEMENT> in xst file.

XML documents import to database. The XML documents generated from the source database of the working groups will be brought back to the design center and stored in the XML document pool per month. Similarly, the XML documents generated from the source database of the design center will be stored in the XML document pool in the working group. The naming rule for the XML document obeys “file-name_YYYYMM.xml” (e.g., DealHQD_201207.xml). The technician can get the needed information by querying the monthly or historical work information in the related XML documents. In order to avoid the problem of the XML documents missing or corruption and to improve the data reliability, the XML documents are imported to the database of the design center or the working group. We use the XMLToSQL library to implement the import from XML documents to database. The main procedure is shown in Fig.2.

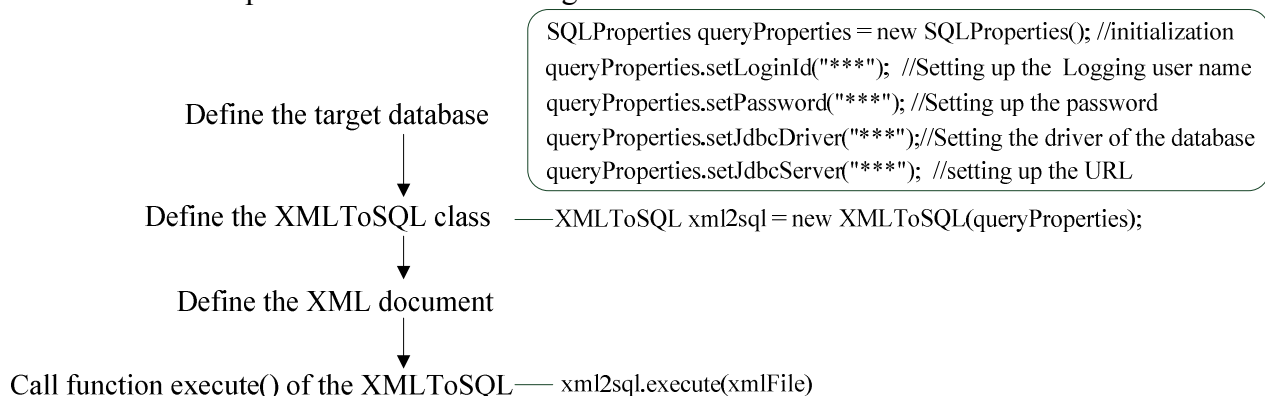


Fig.2 The main procedure of XML documents import to database

XML documents query. The key technologies included XPath and XQuery are used to access and return the needed data from the selected XML documents in the XML document pool. XPath is a language for accessing and referring any part of an XML document. XQuery is an XML-oriented query language rooted in XPath and XML schema. It provides methods to extract data from XML documents [8-9]. According the usage of the XML query language, the main steps are listed as below:

1. Using doc() function to locate the selected XML document.
2. Using XPath to navigate and find the needed data. We used a flexible facility named “XPath expression guide” provided by the WSAD IDE to write and also test the XPath expression. For example, the statement for querying the name of the modified drawings during July of 2012 is shown as:

```
doc("DealHQD_201207.xml ")/SQLRESULT/ModifiedDrawings
```

If the selected historical XML document not exists in the XML document pool, the system will generate the related XML document by SQLToXML first as all historical data from XML document had imported into the database of the design center or the working group. The query flow is shown in Fig.3.

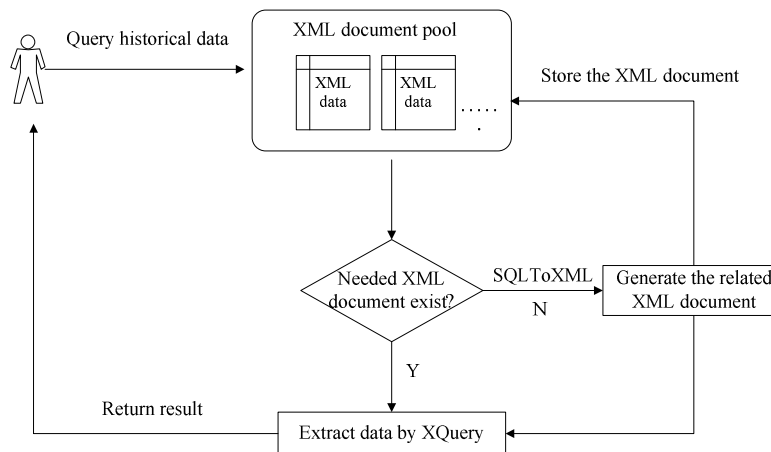


Fig.3 The XML document query flow

XML documents display. XML is designed to carry data, not to display data as XML documents do not contain any display pattern. Displaying XML documents directly will be hard to understand the data information. XSL is an XML-based language used to display XML documents. It consists of three sub-parts: XSLT, XPath and XSL-FO [10-11]. We use XSL to transform and display XML documents. The main steps are shown as follows:

1. Define the stylesheet document that contains a set of template rules and styling features. It is an effective way to gain the code reusability by importing the sharing extern stylesheet document in each XSL.

2. Declare XSL imported in the XML document.

The display interface of the monthly modified drawings information formatted by XSL is shown in Fig.4.

NO.	File Number	Modified Drawing Number	Modified Reason	Result	Analyst	Date
1	D1-162	D100666-009B second edition D100666-009A cancelled	Interference	Processing	Yin An	2012-05-18
		D100747-006A second edition D100747-006 cancelled	Mounting Position Changed			
2	D1-111	D123145-010	Design Optimization	Processing	Yin An	2012-05-11
3	D1-164	D12125-10	Design Optimization	Processing	Yin An	2012-05-09

Fig.4 The display interface formatted by XSL

Conclusions

Considering there is no network connection between the design center and the working group, the traditional data integration methods such as the data pipeline, the heterogeneous data integrated based on SOA and Web service can not be applied. The paper presents the integration method based on XML and discusses the key technologies. The method has put into test, the result shows the design scheme is rationality and effective.

Acknowledgement

This work has been performed as a part of the informationization construction of the China Ship Development and Design Center supported by the Youth Innovation Foundation of the China Ship Development and Design Center under the grant No. YFB11-04-49.

References

- [1] An Yin, Li Gong, Lei Tang: Computer Technology and Development, 2012 (03), p.28-30 (In Chinese)
- [2] Haige Song, Zhibin Zhang: Research of Campus Heterogeneous Database Middleware Based on SOA (Proceedings of the Third International Symposium on Computer Science and Computational Technology 2010)
- [3] M. Arenas, V. Kantere, A. Kementsietsidis, I. Kiringa, R.J. Miller and J. Mylopoulos: The Hyperion Project: From Data Integration to Data Coordination (In SIGMOD Record, Special Issue on Peer-to-Peer Data Management, 2003)
- [4] Web Services Architecture on <http://www.w3.org/TR/ws-Arch>
- [5] XuXin Huang, Ruihua Nie, HuiQiong Luo, HuaiGong Lin, Jiang Ma: Computer Engineering and Design, Vol.30 (14), 2009, p.3338-3340 (In Chinese)
- [6] XML information on <http://en.wikipedia.org/wiki/XML>
- [7] Extensible Markup Language (XML) on <http://www.w3.org/XML>
- [8] Cho SR, Koudas N, Srivastava D: Meta-Data indexing for XPath location steps (Proceedings of the ACM SIGMOD Int'l Conf. on Management of Data (SIGMOD). Chicago 2006)
- [9] XQuery 1.0: An XML Query Language (Second Edition) on <http://www.w3.org/TR/xquery/>
- [10] Extensible Stylesheet Language (XSL) Version 1.1 on <http://www.w3.org/TR/xsl/>
- [11] XSL Transformations (XSLT) on <http://www.w3.org/TR/xslt/>