An Information Integration Strategy based on REML

Yuan Zhao, Lijun Zhou, Jie Zhang

Department of Basic Experiment, Naval Aeronautical and Astronautical University, Yantai, 264001, China

email: ascendtop@126.com

Keywords: heterogeneous information; REML Wrapper; ISS;

Abstract. The paper introduces a strategy "Integration of heterogeneous information in the Distributed System based on REML". In the strategy, we wrap different formats, distributed heterogeneous information into the unified REML information, and make it to be re-utilized by Web Service which is a method of communication between some electronic devices or applications over the web, so that Much many heterogeneous information can be accessed by users without developing a new system. In this paper, firstly propose the whole framework and the REML encapsulation protocol, then present the process of wrapping and parsing of the REML information. At last, we discuss the benefits of using this scheme.

Introduction

As known, information system is complex, and its behavior has strong purpose. For a long time, has been used to differentiate itself according to the function, information system construction is mapping the structure. The different departments of the information system using different development tools, development technology, hardware platform, operation system and database management system. But these self-governed and different standard "chimney" system emerge more and more disadvantage in information sharing and mutual work[1][2]. Based on this backgroud, the paper puts forward the solving scheme to wrap heterogeneous information from different data sources and form unified REML information to publish based on Web Service, so as to provide an ideal information integration platform.

Information Encapsulation Protocol—REML

In our strategy the process of converting the source information into REML information is seen as wrapping. Wrapper can be seen as the bridge and the link to connect heterogeneous data and information integration system [3]. It can be said that all the various heterogeneous information will be wrapped into unified REML information by the wrapper, so the development of wrapper is important. The REML unified file format as follows in Table 1.

Table 1. The REML unified file format

```
<? xml version="1.0" encoding="gb2312" >
<root>
<element>
<model>Z9</model>
<service time>2005.1</service time>
<service area>Jiangsu</service area>
<service effect>well</service effect>
<product name> The steering gear </product name>
<batch number>P10</batch number>
</elment>……</root>
```


Among them,<root>is the root element of the REML, <element> is the sub-element of <root>, the numbers of <element> can be expanded, can be one or more, and the name of the child elements are not fixed.

Web Service

Web Service is a method of communication between two electronic devices over the web. The W3C defines Web Service as "a software system designed to support interoperable machine to machine interaction over a network. It has an interface described in a machine-processable format(specifically Web Services Description Language, known by the acronym WSDL). Other systems interact with the Web service in a manner prescribed by its description using SOAP messages, typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards"[4].

There are many ways that a requester might engage and use a Web service. In general, Web Service has three roles, as illustrated in Figure 1:

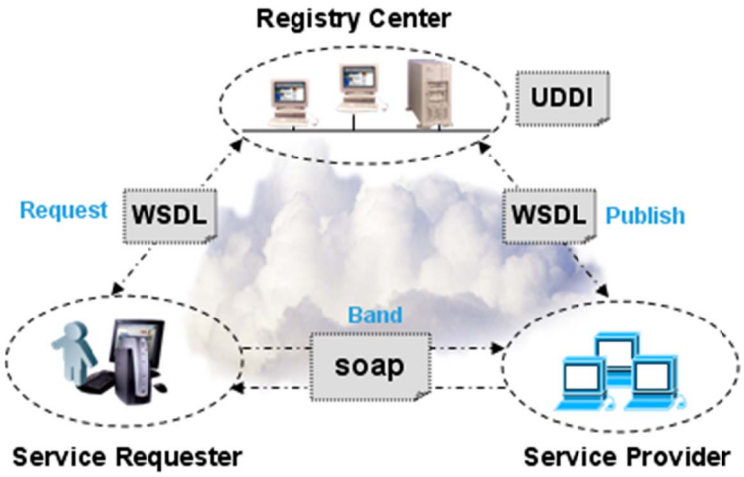

Fig.1. The Architecture of Web Service

Wrapping Strategy

Overall Framework

Information storage formats of each department may be heterogeneous in network,For instance, document information of department 1 is semi-structured Xml information, document information of department 2 is structured information stored in the SQL Server database, the reports of department 3 is Excel structured information, department 4's pictures, sensor datas, etc are unstructured information. The paper put forward an unified middle information format, which can wrap the heterogeneous data objects into unified REML(Root-Element-Markup-Language), and encapsulated into IIS(Information Shared Service) which will be registered and published in the Information Register Center. As shown in Figure 2, it is the overall framework for intergration of heterogeneous information. We can see that, ISS encapsulate all kinds of informations and hide the complexity inside of the heterogeneous information, then REML shield the heterogeneous of the source information.

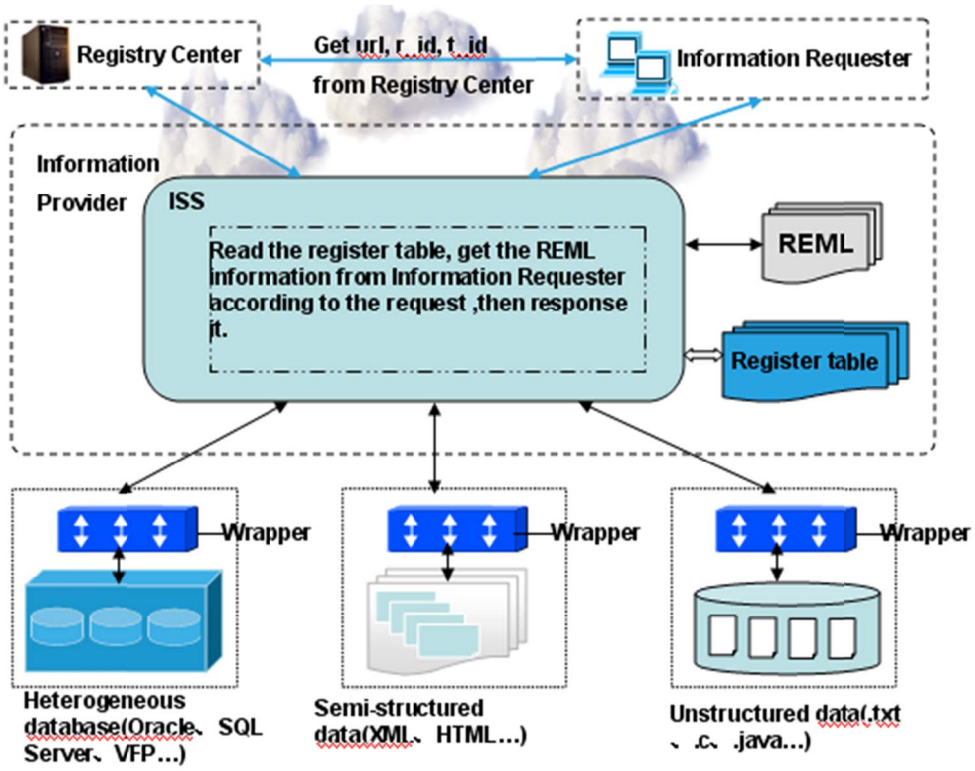

Fig.2. The overall framework of information intergration based on REML

REML Wrapping

The wrapping process to be realized with Java language, in order to meet the system's requirements of portability and good cross-platform performance. The development platform use the Eclipse tool (Eclipse is an open source, extensible application development platform[5]) and access database with JDBC(Java DataBase Connectivity).JDBC is the Java's interface standard to database, which defines the general API(Application Program Interface)supporting the stand SQL function. JDBC API defines several Java classes, allow Java programmer send SQL commands and deal with the results[6][7]. Through the driver manager of program, JDBC API can use different drivers links different database systems. In the scheme, JDOM is used to generate REML information automatically. Noting that, the file "jdom.jar" should be copied to the directory "JDK\lib" in using the JDOM Parser and the environment variables should also be set. The file "jdom.jar" can be download in the jdom's official website. In the process of wrapping, several important methods are as follows:

```
Class.forName("sun.jdbc.odbc.JdbcOdbcDriver")
Connection con=DriverManager.getConnection("jdbc:odbc:studentdb")
//load the drive, and build connection to the data table.
Statement st=con.createStatement( )
ResultSet rs=st.executeQuery(strSQL)
//build the ResultSet after query.
Element rootElement=new Element("root")
Element element=new Element("element")
rootElement.addContent(element)
//add root element <root> and its sub-element <element>.
XMLOutputter
remlOut=new XMLOutputter("",true,"gb2312")
//output to the REML document, the coding is gb2312.
```

Advantage of the Scheme

Openness

The system uses the Web Service technology to construct, provides an effective method to integrate the existing applications of the network[6], so that the system has good adaptability and openness.

Loosely Coupled

As a service oriented computing model, Web Services technology is loosely coupled and platform independent, and it simplifies the application integration between different organizations. In the platform of information integration, the client and server exchange information with the middle fomat of REML, thus make their information exchange system more adaptable and reasonable. Therefore we call the system is loosely coupled.

Reusability

Using Web Service to publish the ISS, can realize the effective information sharing, and also can be the information sources to the other application systems. In addition, some upper application can use the ISS to do the second development.

Conclusions

In the strategy "An Information Integration Strategy based on REML", we wrap different formats, distributed heterogeneous information into unified REML information, and make it to be re-utilized by Web Service. Much many heterogeneous information can be accessed by users without developing a new system, then we can say that achieving a certain degree intergration of information in the network, it can be used to the better development and application afterward.

Acknowledgements

This work is supported by the Foundation of Naval Aeronautical and Astronautical University (HYJC201227), and the authors would like to thank Atiao Yong who is working on doctoral degree in South China Normal University for her useful suggestions on this paper.

References

[1] Zhiwei XU, Baiming Feng, Wei Li. Grid Computing Technology[M]. Bei Jing: Publishing House of Electronics Industry, 2004:25-58.

[2] TUECKE S, FOSTER I, KESSELMAN C, etal Open grid services infrastructure (OGSI)(draft)[EB/OL].(2003-02-01).http://www.gridforum.org/ogsiwg/drafts/draft/draft-ggf-ogsi-gridservice-230_2003-02-17.pdf.

[3] Gray, N. A. B. (2005)."Performance of Java Middleware - Java RMI, JAXRPC, and CORBA". University of Wollongong. pp.31–39. Retrieved January 11, 2011. "The results presented in this paper show that the nature of response data has a greater impact on relative performance than has been allowed for in most previous studies".

[4] "Web Services Glossary". W3C. February 11, 2004. Retrieved 2011-04-22.

[5] Wendy L. Value creation in web services: An integrative model[J]. The Journal of Strategic Information Systems, 2006,15(2),153-174.

[6] Guanyu Li, Jun Liu, Jun Zhang. Researches and Implementation on the Distributed Heterogeneous Data Integration System[J]. Application Research of Computers, 2004,34(3):96-98.