A STUDY ON RS IMAGE SHARING BASED ON DATA GRID

WANG Lianbei [a, *]

[a]Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052, China - Wlbzhm@sohu.com

KEY WORDS: Data Grid, Metadata, Globus, Spatial Data Sharing, Geo-spatial Information Service

ABSTRACT:

Based upon the in-depth analysis of the remote sensing image sharing demand and the status of the current research on it, this paper points out the existing Web-based sharing-model did not eliminate "Information Islands" fundamentally. To tackle the problem, the author has been working on the research topic of finding a remote sensing image data sharing model which is based on the remote sensing image information system technology and the Data Grid technology. Hereby the author has provided the Data Grid-based model. Experiments are carried out on the remote sensing image information sharing service based on the Data Grid Technology under the environment of GlobusTookit4.0 Grid Computing Platform. The result of the experiment shows, the remote sensing image data sharing service model is technologically available for the integration of heterogeneous and geographically distributed image storage and management system.

1. INTRODUCTION

The spatial information, especially the RS image presents various location information of the earth surface. As a main part of the scientific data, the sharing of the RS image has an important meaning to achieve high level research results. Along with the implementation of various earth observation plans, there are more and more RS image information in China while utilization ratio is deficient. The main reason is that the data is bounded in the different RS data management systems built by its' owners. Thus, the "Isolated Information Islands" come into being. The solution for the above mentioned problem, which can realize the data sharing by integrating the heterogeneous and geographically distributed RS image storage and management systems, has become the current research focus.

2. RESEARCH STATUS AND ANALYSIS

2.1 Research Status

Driven by the RS image applications requirement and with the development of the Internet technology, the research institutes and organizations around the world have developed various RS image information management and sharing systems. Several representatives are given as follows.

TerraServer: Developed by Microsoft, Digital Device, Aerial Images and USGS. Through the system, users could browse and down-load images via the internet. The technical principle is that it provides the meta-data release and the data browse function via the WWW technology as well as the data down-load service via the HTTP and the FTP technology.

Image Web Server (IWS): Product of Earth Resource Mapping Co., Ltd. It enables users to manage the digital image at GB or TB level on the line and share the image data via the public networks. More than the data browse and down-load based on the WWW technology, IWS provides the online data service realizing the WMS (Web Map Services) criterion based on the WEB Service technology, which enables the real time data service available for the image application systems.

Google Earth, created by Google, combines satellite imagery, maps and the power of Google Search that puts the world's geographic information at your fingertips, giving you access to terabytes of geographical, political and social data. Google Earth is a GIS program that allows users to overlay datasets of their own choosing.

2.2 Research Status Analysis

After the comprehensive analysis of the characteristics and the technical principle of the existing RS image sharing systems based on the internet technologies, we could find that the existing systems can be divided into two main categories.

One kind system realizes data sharing through finding the required spatial data for the users by searching the spatial metadata, then, sharing the data by downloading from the internet or via the disk/CD from the data owners. It is an off-line and manual model. It works in this way. With the support of WWW technology, the data browse is available. The data downloading is realized via the HTTP or the FTP protocols, etc. TerraServer is the examples of this model.

The other kind system provides the RS image information search and browsing services on the internet via the client application (or browser plug-in). It is the Web Map Service mode. The client application acquires the data from the server via the special protocol (Web Map Services protocol etc.). The service on the server is the HTTP service model or the WEB Services model. Google Earth and IWS are the examples of this model.
By analyzing the above materials, we could find two disadvantages exist in the current sharing models and technologies.

* Corresponding author.

(1) The current sharing model does not solve the problem of "Isolated Information Island".

The internet makes the information transmission available via the TCP/IP protocol but it does not ensure the data sharing. The incompatible data access methods are the barriers among the various RS image storage and management systems, thus the RS data resources are isolated from each other for the users. Thus, the more heterogeneous image storage and management systems are developed, the more "Isolated Information Islands" are formed.

Technically speaking, the internet technology does not ensure a favourable data sharing environment for the integration of the heterogeneous and geographically distributed systems since it does not solve the problems such as the resource management, service release and identification, safety, etc.

(2) Storage Facility Sharing is not available.

The storage facilities providing speedy and reliable data access service have a great ration in the cost of various RS image application systems since the RS image data is massive. It is obvious that the off-line sharing mode of browsing and downloading needs more storage facilities in the process of data transmission rather than sharing them, thus it is not an economical model.

Technically speaking, the problem is that the current internet does not provide a favourable way of the data integration service for the data users. The difficulties are how to release the data service information, how to find the appropriate data and data service, how to integrate these services in the application programs and how to control the data accessing activities, etc.

3. THE DATA GRID TECHNOLOGY

A group of technical experts in the computer science is concentrating on the research of universal sharing mode of the distributed resources when the RS experts are trying to solve the above mentioned problems. The universal sharing mode means the realization of cooperation resource sharing. Thus an open platform including a series of protocols, services and API (Application Program Interface) and SDK (Software Development Kit) is developed for the sharing of the distributed resources. The technology is named the Grid Technology. The open platform is called OGSA (Open Grid Services Architecture).

The majority of the researches in the areas of digital earth and RS etc are related to the sharing and accessing of massive data. The accessed data is at the TB level or the PB level. Meanwhile, the data is probably from various management systems which are heterogeneous and geographically distributed. Thus the current data management architecture, methods and technologies are not sufficient for the access of massive data resources which are geographically distributed and of different architectures.

The Data Grid came into being along with the development of the Grid Technology. It is distributed data management architecture. The Data Grid architecture is composed of three layers: the Bottom Resource, the Data Grid Middle Ware and the Data Grid Application. The Bottom Resource includes the data resources, storage facilities, sensors and large scale equipments for scientific researches, etc. The Data Grid Middle Ware is for the resource sharing and coordinated activities related functions as well as the safe, sufficient and favourable Data Grid environment. The Data Grid Application system provides the user-required services which are developed under the environment provided by the Data Grid Middle Ware.

The Data Grid provides the technologies such as: the Safety Management Technology, the Information Service Technology, the Coordinated Access of the Distributed Data Resource, the Speedy Transmission of the Massive Data, etc.

4. THE RS IMAGE DATA SHARING MODEL BASED ON THE DATA GRID TECHNOLOGY

Targeting on the "Isolated Information Islands" in the RS image data sharing, the author tries to find out a RS image data sharing model based on the Data Grid technology which is able to solve the following problems.
(1) The approach how the data providers issue the metadata and data service information of the remote sensing image data resource.
(2) The approach how the data users identify the data and bind the data service according to their requirements.
(3) The approach how the data users get access to the data provided by others.

Then, the following researches are carried out：
(1)The study on the model of remote sensing image information service metadata：Studying the inquiry way of remote sensing image data, defining the metadata which describes the image data contents； Studying and defining the data which is applicable to the Grid Service, and in the meanwhile, it should be easy for the find, bind and integration of the data service.
(2) The study on the metadata sharing service system: Studying the registration mechanism and the update mechanism for the accurate find of the data.
(3) The study on the interface model of the Grid data service：
Based on the study of Grid Services, providing an easy way of data service integration, thus making it convenient for the client application program to integrate the service and get the data from the service.

After the above researches, the author designs the remote sensing image information sharing service model which is based on the Data Grid technology. The model is composed of the following four layers, shown in the Figure 1:

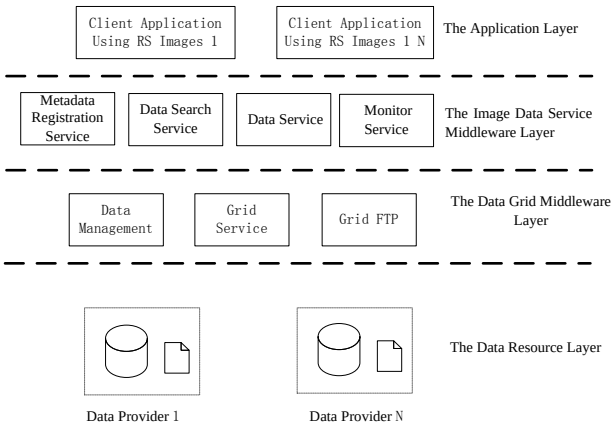

Figure 1. The RS Image Data Sharing Model

The Data Resource Layer: It is composed of various remote sensing image storage and management systems. The systems might be the heterogeneous systems which are geographically distributed. They are the data providers in the model.

The Data Grid Middleware Layer: It is composed of the data grid components. It provides the basic grid computing environment to the image data service components in the upper layer.

The Image Data Service Middleware Layer: It is composed of the metadata registration and update service component, the data inquiry component, the data service component and the monitor component. The metadata registration and update service components are for the data providers to register and update the relevant metadata of the data contents and the data service. The data inquiry components and the data service components are for the client program to find and acquire the data. The monitor component is for the administrator to monitor the status of the sharing system. It is the interface for the data users to find and acquire the data as well as the agent for the data providers to issue their data.

The Application Layer: It is composed of various client programs which use the remote sensing image. It is the data user in the model.

In this model, the remote sensing image information sharing service works like this:
(1) The data providers register and update the metadata on the metadata service system.
(2) The client program searches the data via the metadata service system.
(3) The metadata service system finds out the data provider catering to the requirements and provides the information to the applicant.
(4) The client program binds the data service interface provided by the data provider and gets the data.
(5) The data service interface acquires the data from the data storage system and gives it to the client program.
(6) The monitor service monitor all the activities motioned above and report to administrator.

5. EXPERIMENTS BASED ON THE GLOBUSGRID COMPUTING PLATFORM

Making a feasibility study of the model, the author presents the remote sensing image information sharing service experiment supported by the Data Grid Technology under the environment of GlobusTookit4.0 Grid Computing Platform. The registration service based on the Registry interface criteria and the inquiry service based on the Grid Service technology are designed and developed during the process of the experiment, thus having realized the dynamic bind of the data service and the data distribution. Based on that, A Prototype System is developed.

5.1 Main Works of The Experiments

The main works of the experiment are as follows.
(1) The Metadata Registration Service based on the Registry Interface is developed. It is designed for the data providers to manage the metadata like the registration data description, service binding information and integration information etc.
(2) The Data Inquiry Service based on the Grid Service Technology is built up. It is designed for the data users' client application program to find out the required data service dynamically.
(3) The Local Metadata Registration Service is constructed. The Registration Service Interface on the Metadata Server provides the registration service and the dynamic data updating.
(4) The Local Data Service based on the Grid Service is set up. It is for the client application program integration. It provides the real-time data searching and local data distribution service.
(5) The Client Application Program is developed. It is used to test the data search service and the data services mentioned above.
(6) The Monitor proxy/stub components are developed. They are used to monitor all the interactive actions among the Client Application, the Metadata Registration Service, the Data Search Service and the Local Metadata Registration Service.

5.2 Experimental Environment and deployment of Services

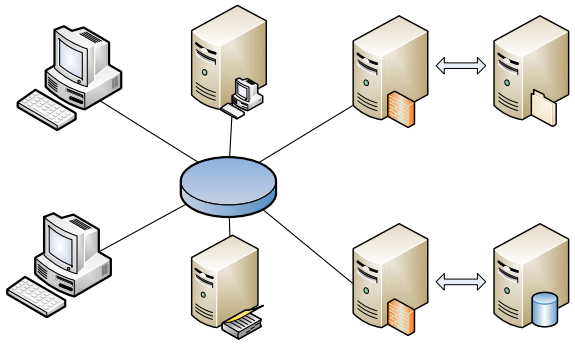

Figure 2. Experimental Environment of Prototype System

As shown in figure 2, two client applications are deployed on the PC A and B; The Metadata Registration Service component and the Data Inquiry Service component are deployed in Metadata Management Server; The Data Service component is deployed in the Local Data Proxy Server; The monitor proxy component is deployed in the Monitor server, which can receive the interactive information given by the monitor stub component and report to the administrator. The monitor stub component is deployed in the servers in which Services are deployed, such as Metadata Management Server and the Local Data Proxy Server. There are two local RS data storage and management systems in the experimental environment. They are heterogeneous and geographically distributed systems.

5.3 Results of Experiments

In the experiment, the client application program can find out the appropriate data provider via the Metadata Management Services deployed on the Metadata Management Server and bind the data service deployed on the Local data proxy server. After that, the data service retrieve data from the local data management system and send to the client application. Thus, the client application can get the data which it need. The Interface of the client application program is shown in figure 3 and figure 4. The report of the monitor component is shown in figure 5.

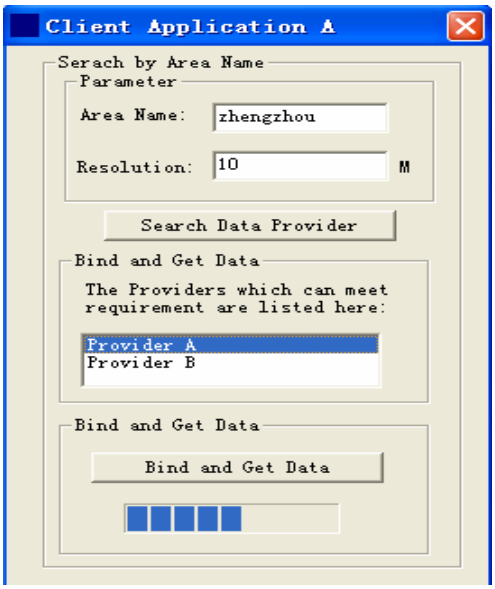

Figure 3. The interface of the client application A

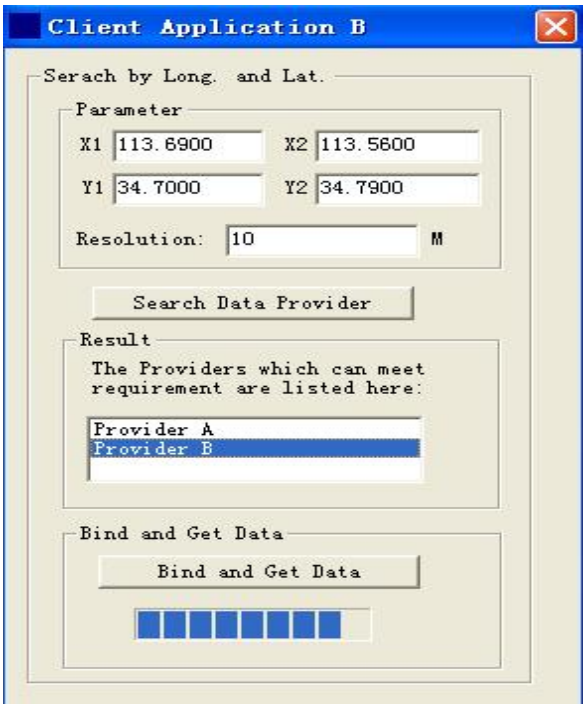

Figure 4. The interface of the client application B

6. CONCLUSIONS

The result of the experiment shows: the remote sensing image data sharing service model shown in this paper is technologically available for the integration of heterogeneous and geographically distributed image storage system. The problem of "Information Island" is solved and the sharing of data and stored resource are realized by the model. The main features of the model are as follows:

(1) The model expanded the remote sensing image information metadata model to meet the requirements of Grid Data Service. It enables the client application program to find, bind and integrate the data service since it provides both the metadata of remote sensing image and the metadata of data service at the same time.

(2) It has designed two services, i.e.: the metadata registration service which is based on the Registry interface criteria and the inquiry service which is based on the Grid Service criteria. The sharing service mechanism, which is based on the Data Grid technology and supports the dynamic update, not only ensures the data providers issue their data and service information at a comparatively low cost but also makes the data users convenient to get more accurate search and integration of the data service.

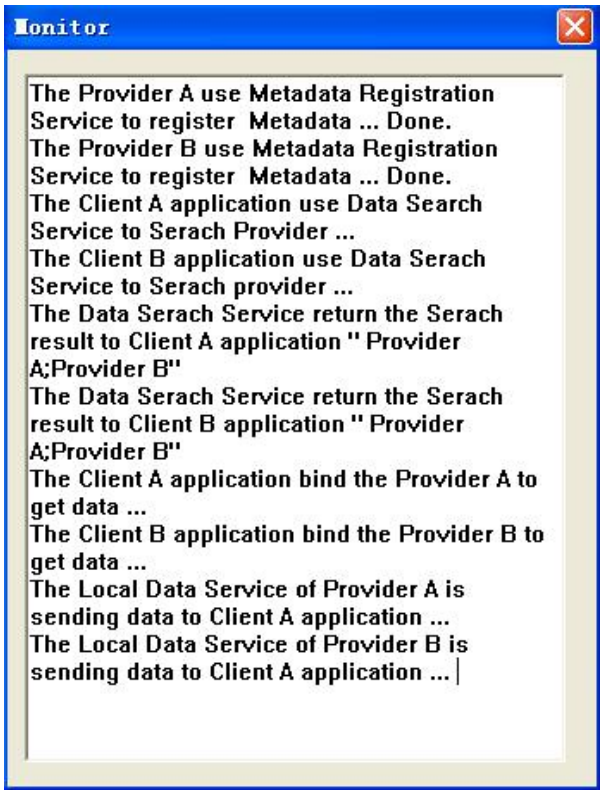

Figure 5. The interactive information reported by the monitor component.

REFERENCES

Allcock, W., 2005b. The Globus Striped GridFTP Framework and Server. *Proceedings of Super Computing,(SC'2005).*

Deelman, E., 2004. Grid-Based Metadata Services. *16th International Conference on Scientific and Statistical Database Management.* http://www.globus.org/alliance/publications/papers/deelman3.pdf(accessed 28 Nov. 2007)

Foster, I., 2002a. Grid services for distributed system integration. *Computer*, 35(6) , pp. 37 -46.

Foster, I., 2005b, Modelling and Managing State in Distributed Systems: The Role of OGSI and WSRF. *Proceedings of the IEEE*, pp. 604-612.

Paton, N. Data Access and Integration Services on the Grid. *UK e-Science Programme Technical Report Series (UKeS-2002-03).* http://www.nesc.ac.uk/technical_papers/dbtf.pdf(accessed 28 Jan. 2008)