THE RESEARCH OF FAST PROCESSING AND DISTRIBUTION REMOTE SENSING IMAGE BASED ON THE GRID TECHNIQUE

Liang zhong[a,] *, Hongchao ma[a]， Jie sun[a], xun zen[a]

[a]School of Remote Sensing Information Engineering,Wuhan University
Luoyu Street 129, Wuhan city, Hubei Province, P.R.China –
bluewind20000@163.com

Commission WG IV/5

KEY WORDS: Distributed Computing, Image Processing, Networking, Systems, Remote Sensing, Compression

ABSTRACT:

In this paper, a brand new method utilizing the service-oriented software architecture and the rising technology of Grid computing with available software Globus Toolkit 4 is introduced. The traditional centralized remote sensing image processing model under stand-alone computer environment is changed into a model in Grid environment by decomposing the system into different services which enables the users to process the remotely sensed images, perceive quickly and obtain the result image using the service in Grid bases on their given demands. In this way the whole processing from raw data to product generation is completed rapidly.

1. INTRODUCTION

Currently, geo-spatial information has been applied widely to every field and industry that related to “Geo-spatial position”. A paper titled “the opportunities of mapping” published on the famous《Nature》in January, 2004 pointed out that: "from the very beginning of this year, along with the Nanotechnology and Biotechnology， the geo-information technology is accepted by the labour ministry of the US as one of these there most important developing technologies. And the expression of its value in every area will increase the employment opportunities and makes it be diversity."

Remote sensing – one of the Earth Observation technologies- as a critical supporting technique of the GIS is developing rapidly. Newly designed sensors with high multiphase, high optical resolution and high spectral resolution can provide us with plenty of geo-spatial information. But it also brings some problems: how to timely process and distribute these large numbers of data for human comprehension, application and for computer interpretation and generalization analysis, especially transform to requisite information in real-time, there are all difficult tasks ahead of us. The current problem is that, the contradiction between the rapid development of data acquisition with relatively lagging of data processing and distribution has becoming increasingly prominent and sharp. So, the huge image data of the observed objects gets from high multiphase, high optical resolution, high spectral resolution and high overlapping also bring us new challenges. Traditional remote sensing image processing system always completes the production of spatial information in stand-alone, centralized processing mode, which lacks the technique of sharing and distribution information real-time or quasi-real-time. Another reason is that the demand of the processing of raw data in different fields varied widely makes the centralized processing mode cannot meet the requirement of providing the data networked, distributed, allocated according to need, and the requirement to release or share the information resources real-timely or quasi-real-timely. So the image processing, information extraction and distribution or sharing of the mass remote sensing data rapidly, effectively and intelligentized becomes an urgent scientific problem.

The construction of the new generational network, the raising of the Grid technology (I. Foster, Kesselman C,1998) and the emergence of the advanced compression techniques have formed a good theoretical base for the new generational network oriented remote sensing data processing system.

2. THE STRUCTURE AND CHARACTER OF SYSTEM BASED ON GRID

2.1 The overall architecture of the system

“Grid” computing has emerged as an important new field, distinguished from conventional distributed computing by its focus on large-scale resource sharing, innovative applications, and, in some cases, high-performance orientation (I. Foster, C. Kesselman, S. Tuecke., 2001). In this paper, the overall architecture of remote sensing processing system based on Grid and service is composed of a structure which includes the following four layers: as shown in Figure 1

1:.Data layer it can offer various raw products to processing-service unified system upon users’ request for further and deeper processing results.

2: processing layer is formed by various services for remote sensing processing whose package mode are Grid services based on high-performance environment or web service conforms international norm of OGC.

3: Grid layer, the main role it plays is managing both the data and service resources so that to achieve the goal of sharing the

* Corresponding author. - bluewind@163.com; phone 0086-1347 6140 343

resources, providing a unified system operating environment of remote sensing application and coordinating the tasks of remote sensing processing. Due to that the Grid technology develops very quickly, the GT4 that has been widely used in the world is the core of the network environment layer when considering the scalability of the operating environment and compatibility and the conforming to the current trends of technology to interconnect and interwork with other Grid system.

4: Application support layer is a virtual logical layer. Its role is the connectivity between the remote sensing systems facing end-users and various sources of remote sensing processing in Grid. It is formed by middleware supporting Grid, API and tools such as automatic generation tools of remote sensing application, user interfaces, application deployment tools, etc.

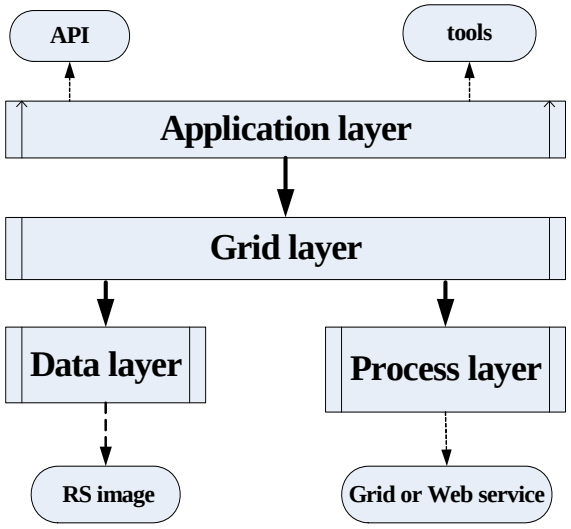

Figure 1. The layered system architecture and its relationship

2.2 The fundamental characteristics of the system

1: Rich and practical processing algorithms system of remote sensing based primarily on the common and stable algorithms formed by summarizing the universal methods and technologies on them and absorbing the forefront research results. The remote sensing platform based on Grid and high-performance computing technology can ensure the running of designed algorithms of parallel processing and high-performance computing capabilities, according to the designing and developing specifications of Grid environments, the algorithms become a sort of standard treatment services providing users good abilities of secondary development and customizing.

2:The system can meet requests of producing and processing raw data, realize the geometric correction, fusion, classification and post-processing of large remote sensing images, it can achieve automatically data manufacturing and provide a good basis for advanced processing and application based on its standards and norms of remote sensing data and general processing algorithms.

3: As the middle layer with processing algorithms and Grid environment offers an efficient capability of data accessing, supports the accessing of local data and their special formats, thus ensures the efficient processing found on Grid.

4: These processes can be flexibly customized depending on the module of resource management and service on Grid platform. For instance, task wizard, multi-tasking schedule, Grid service whose algorithms can be customized, multi-user supported working environment and visually realized work flow creating based on the algorithm service component. All together make it easy to use.

5: The system with a human-machine interactive interface supports data editing operation, quick browsing to high-fidelity data, quickly displaying of 2D/3D data and the overlaying of data with their spatial information.

3. THE DESIGN FEATURES

3.1 Application of the tool GT4

Globus software is designed to enable applications that federate distributed resources, whether computers, storage, data, services, networks, or sensors. Globus Toolkit (GT) development has tracked this trend, which is Grid computing moving from legacy computing-intensive applications to service-oriented computing based on open standards. With GT4 (M. Humphrey, G. Wasson, J. Gawor, et al., 2005, I. Foster, 2005)].It contains many functions such as monitoring and discovering resources, security, management and distribution of data, etc. The way to integrate Grid computing services with remote sensing data processing depending on various tools offered by GT is as following:

And the GT4 core can be deployed in different system platforms like Linux, windows. It can be used in processing service including the creation and destruction of service, the life-cycle of service, the notification and subscription of service, and the registration and release of that. The above function can be accomplished through manipulating, which is from GT4 core, the generated service: the package of mature remote sensing process algorithms.

The GRAM, sites at the bottom of the structure of Globus resources management system， mainly deals with resource request about remote task， resource distribution of remote tasks, management of the activation of remote task, at the same time the availability and situation of the computing resources are dynamically updated and passed to the registration centre. The user in Grid can directly access resource based on a standard interface provided by GRAM for local resource scheduling system, without noticing local resource interface. In the system, the main function is the distribution and management of data resource and processing algorithm service resource of remote sensing.

GridFTP：(Grid File Transfer Protocol)－GridFTP based on the FTP protocol norms, extends them all-round and supports high speed data transmission. GridFTP supports GSI and Kerberos security mechanism, and therefore ensures the data integrity in different levels depending on flexible identification, reliable security and integrity check. It is used in remote sensing data transmission on network and can improve performance and efficiency of data transmission..

GT4 MDS: the registration service does play a crucial role. It can resolve the registering, publishing and searching of resource through networks, thus it is the most important part while designing the architecture based on service-oriented

framework. It plays a vital role for improving the capability of quick and stable transmission of the mass remote sensing data and therefore promotes the performance of the platform.

3.2 Packaged service

Web service technology is a standard mechanism for application releasing through internet and utilizing the software service. Web service can easily exchange and integrate the data between different systems (Cheng Yongxing, Chen Ping, 2003). Its models provide scalable, loosely coupled capability to exchange information under non-specific platforms (Zheng Xiaodong Wang Zhijian Zhou Xiaofeng Fei Yukui Xu Feng ,2004, Zheng Xiaodong Wang Zhijian Zhou Xiaofeng Fei Yukui Xu Feng ,2004). The information exchange depends on a series of technology standards such as XML (Extensible Markup Language), SOAP (Simple Object Access Protocol) and WSDL (Web Services Description Language). Its major advantage is to actualize the interoperability between heterogeneous platforms. Users need no concerning on the structure and functions in the Web service because its communication interface has been predefined and therefore the complex processing function is made to be transparent to the them.

In GT4, it provides some registered default services and makes extensive use of Web Service (Booth D,Haas H, McCabe F et al.,2003), so many extension models in the form of service are in its containers. Therefore, the traditional methods of remote sensing data processing can be used in GT environment if it exists in the form of service. So when considering the algorithm of processing realized under stand-alone version of remote sensing processing platform and that in platform of Grid-oriented one whose main idea is distributed, thus to decompose remotes sensing processing as a series services，each algorithm correspondences to a service and can be packaged and translated. The algorithms included in these services are the integration of that for geometry processing, multi-spectral classification and multi-source integration.

WS Java Core as part of Globus Toolkit V4 is the Java implementation of WSRF. A web service container placing service is needed for using the WS Java Core. It provides a number of API and tools for building web service (WS-Resources) with state. The java class packaged with containers and various scripts can be used to deal with WS-Resources. Only Java WS Core is concerned when developing remote sensing processing based on Grid.

The Grid service is described as five simple steps: as shown in Figure 2

1: Define service interface, and then write WSDL files to describe external exposure operation of remote sensing image processing.

2: Realization service: Implementation service by using java language or JNI technology to transfer the processing algorithms developed via other languages such as c, c++.

3: The Definition and deployment of files: These files of remote sensing processing service can be implemented by writing JNDI and WSDD file.

4: Compile and generate the GAR file.

5: Deploy service to the container.

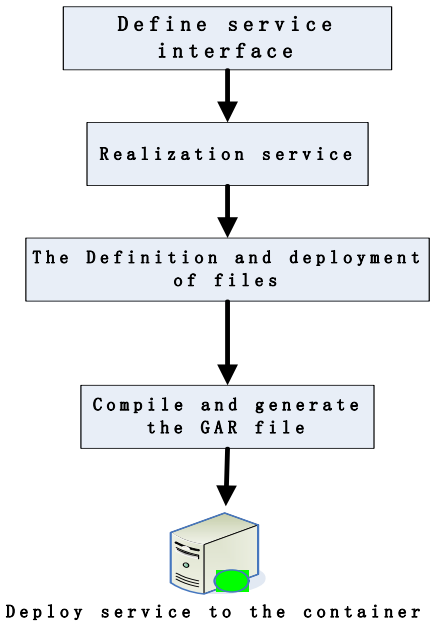

Figure 2. The step of packaged service

3.3 Strategy on fast distribution

Several problems of remote sensing data processing will face in Grid environment are as follows: first of all, the global remote sensing images contain not only different kind of the scales, bands, types and many other dimensions but also different the time, space, spectrum, multiple-resolution and different formats. Different images need different way to process, thus when users need to get detailed information of the earth observation image before processing it. The result of many algorithms of remote sensing processing is related with parameters chosen based on image status, therefore the users use different parameters to treat image for optimal ones. So, to sum up, users need to have a detailed understanding about the image in the whole process. The traditional browser such as Internet Explorer or Netscape Navigator fail to support the display and ramble of large raster image from different sensors and various formats, and the transmission rate of current internet is limited, so we must get some way for the rapid cognition of images information and comparisons between them.

Therefore, in accordance with the above difficulties and purposes of fast transferring data on current internet environment and showing image information in client browsers which are resolved, here by not only JPEG coding technology is used to recode the image but also extracting and interpreting the spatial information from the remote sensing image. A professional module of compressing images and extracting spatial information is devised for rapidly generating the publication information which can be sent to the pre-designated cache database when the user calls this module after the processing of image data. Users can browse quickly the data from the cache database in IE browsers and know images better such as simple measurement in recoding image without downloading the result of large data, thus it is more effective for users to process images.

Image information is released from the cache database through building web service and JSP pages developed by dynamic technologies, so that remote users on the internet will be able to access the compressed image quickly with its spatial encoding information.

The compressed image will be loaded into the same size of image-layer automatically generated by client browsers, and then data from analytical XML file will also be loaded to these pages. Affine transformation coefficient can be gained when the image information contains these defined as padfTransform [5], and it must be pointed out that the image will be resampled to the restricted size before being compressed if the image size is bigger than the maximum size IE browsers supported, and the sampling interval value m_changrate will be sent to the coordinate conversion equation. The flow of real coordinate conversion as is shown in figure 3;

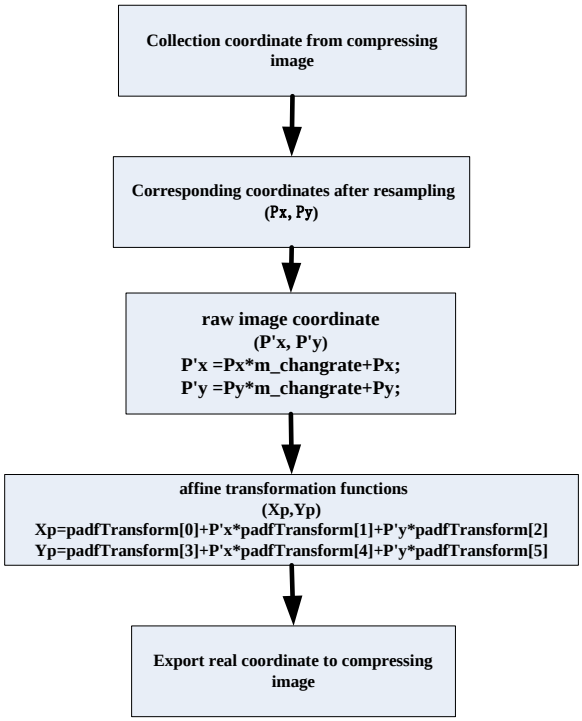

Figure 3. The flow of coordinate conversion

And the real coordinates of points specified by the users in compression images in client browsers: Xp and Yp can be got through this equation. Then it is measuring distances and computing area based on real coordinates that can be realized, therefore the user can understand images in more detail and can compare them.

4. THE WORKFLOW OF PROCESSING PLATFORM

The prototype platform structure, as is shown in Figure4, First of all, the user accesses the index of raw images in registration centre established by GT tools, and then calls the module of processing and extracting to compress the raw image and extract its spatial information, the results are sent to the information cache database designated on internet. The remote user can decide whether to deal with the image after seeing the information of an image, which is queried from the cache database, in IE browser. Appropriate services such as geometry processing, classification or multi-source integration services can be chosen by the registration centre when the user decides to processing the image referred through Gridftp part of GT tools to the computer which is deployed service needed. After completing the processing by module of compression and extraction information, the user can see the data information through the IE browsers, and then decides whether to re-adjust the processing parameters for renewedly processing the image or download the result to the client computer by Gridftp.

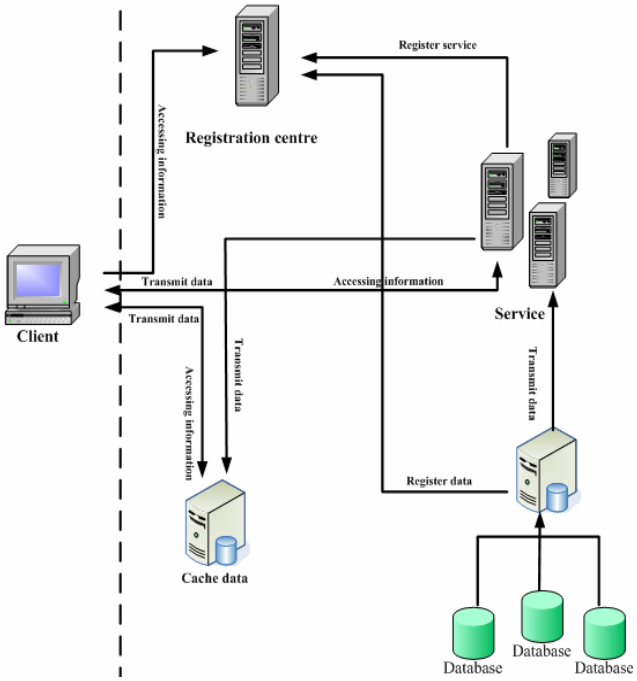

Figure 4. Prototype system structure

5. CONCLUSIONS

This paper introduces the structure of mass remote sensing processing platform based on the Grid technology considering the characteristics of remote sensing data, it studies the way to design the related services and the strategy for rapidly recognizing the remote sensing images in current internet environment, finally, a highly professional platform which promotes greatly the application effectiveness of researches in remote sensing field was created. Therefore the traditional development and designate of various quantitative applications based on single sensor could be changed by the Grid platform. It realizes the resource communication and information complementary; it changes the status that more satellites were launched and more data were received but the current application request is still not meet. The existing sensor potential will be fully exploited; the quantitative remote sensing product with high-precision, high-quality and long time sequence will become possible. Through a large number of remote sensing data continuously become useful information resources, the Global Change Research and Geo-related disciplines will be provided with solid support; the development of Earth System Science will be promoted.

ACKNOWLEDGEMENTS

Support by the ministry of education Foundation of China (A1420060213)

REFERENCE

I. Foster, Kesselman C,1998 The Grid: Blueprint for a New Computing Infrastructure [M] . 1st ed. SanFransisco : Morgan Kaufmann Publishers , pp.259-278

I. Foster, C. Kesselman, S. Tuecke., 2001, the Anatomy of the Grid: Enabling Scalable Virtual Organizations. International J. Supercomputer Applications, 15(3), pp.200-222

M. Humphrey, G. Wasson, J. Gawor, et al., 2005, "State and Events for Web Services: A Comparison of Five WS-Resource Framework and WS-Notification Implementations," 14th International Symposium on High Performance Distributed Computing (HPDC14), pp.3-13.

Foster, 2005, "Globus Toolkit Version 4: Software for Service-Oriented Systems," IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779, 2-13.pp.513-520

Cheng Yongxing, Chen Ping,2003,Web Service in Trust Information Engineering [J] Computer Engineering and Applications,，:39(29), pp 202-204

Xiaohua liu ,2002,Net Web Service Development Guide [M]Beijing：Publishing House of Electronics Industry，

Zheng Xiaodong Wang Zhijian Zhou Xiaofeng Fei Yukui Xu Feng ,2004,Research and Implementation of a Distributed-Computing Model Based on Web Service [J] Computer Engineering and Applications 40(1) pp 144-147

Booth D,Haas H, McCabe F et al.,2003, Web Services Architecture. W3C, Working Draft. http://www.w3.org/TR/2003/WD-ws-arch-20030808/. (accessed 18 Apr. 2008).