ChinaGrid: Making Grid Computing a Reality*

Hai Jin

Cluster and Grid Computing Lab,
Huazhong University of Science and Technology, 430074, Wuhan, China
hjin@hust.edu.cn

Abstract. Grid computing presents a new trend to distributed computing and Internet applications, which can construct a virtual single image of heterogeneous resources, provide uniform application interface and integrate widespread computational resources into super, ubiquitous and transparent aggregation. ChinaGrid project, founded by Ministry of Education of China, is an attempt to achieve above goals by exploring the various resources on existing and well developed internet infrastructure, CERNET (*China Education and Research Network*). In this paper, I will introduce the general picture of ChinaGrid project, its vision and mission. The design of ChinaGrid support platform, called CGSP, is also discussed briefly. To illustrate the reality of ChinaGrid project, five different grid computing applications and its application supporting platform are discussed in detail. The purpose of this paper is to introduce this great grid project to the world completely for the first time.

1 Introduction

Grid computing presents a new trend to distributed computing and Internet applications, which can construct a virtual single image of heterogeneous resources, provide uniform application interface and integrate widespread computational resources into super, ubiquitous and transparent aggregation. According to [10, 11], grid computing is a "resource sharing and coordinated problem solving in dynamic, multi-institutional virtual organizations". The purpose of grid computing is to eliminate the resource islands in the application level, and to make computing and services ubiquitous.

The prerequisites for grid computing lie in three aspects: network infrastructure; wide area distribution of computational resources; and continuous increasing requirement for resource sharing. Nearly all the existing grid computing projects are based on existing network infrastructure, such as UK e-Science Programme [20], Information Power Grid (IPG) [12], and TeraGrid [18]. In TeraGrid, the five key grid computing sites are interconnected via 30 or 40 Tb/s fast network connections. The ChinaGrid project, which will be discussed in detail in this paper, is also based on a long running network infrastructure *China Education and Research Network* (CERNET) [1].

* This paper is supported by ChinaGrid project of Ministry of Education of China and National 863 Hi-Tech R&D Research Program under grant 2004AA104280.

Z. Chen et al. (Eds.): ICADL 2004, LNCS 3334, pp. 13–24, 2004.

There are five major grid computing projects in China [21], they are: China education and research grid, called ChinaGrid [2]; China National Grid [8]; China Spatial Information Grid; China Science Grid; and Shanghai City Information Grid. In this paper, I will talk about the detail of ChinaGrid project and the applications running on top of ChinaGrid.

The following of this paper is organized as follows: in section 2, the history, vision and mission of ChinaGrid project is introduced. In section 3, the supporting grid computing platform for ChinaGrid project, called CGSP, is presented briefly. In section 4, five ongoing ChinaGrid computing application platforms and pilot applications running on top of these application platforms are discussed in detail. Section 5 ends this paper with conclusions.

2 ChinaGrid Project: Vision and Mission

In 2002, China *Ministry of Education* (MoE) launched the largest grid computing project in China, called **ChinaGrid** project, aiming to provide the nationwide grid computing platform and services for research and education purpose among 100 key universities in China. The vision for ChinaGrid project is to deploy the largest, most advanced and most practical grid computing project in China or even around the world.

The underlying infrastructure for ChinaGrid project is the CERNET, which began to run from 1994, covering 800 more universities, colleges and institutes in China. Currently, it is second largest nationwide networks in China. The bandwidth of CERNET backbone is 2.5Gbps, connected by 7 cities, called local network center. The bandwidth of CERNET local backbone is 155Mbps.

The ChinaGrid project is a long term project with three different stages. The first stage period is from 2002 to 2005, covering 12 top universities in China. They are: Huazhong University of Science and Technology (HUST), Tsinghua University (THU), Peking University (PKU), Beihang University (BUAA), South China University of Technology (SCUT), Shanghai Jiao Tong University (SJTU), Southeast University (SEU), Xi'an Jiaotong University (XJTU), National University of Defense Technology (NUDT), Northeastern University (NEU), Shandong University (SDU), and Sun Yat-Sen University (ZSU). The site view of ChinaGrid project is shown in Fig.1.

The focus for the first stage of ChinaGrid project is on platform and applications on computation grid (e-science). These applications are varied in all scientific disciplines, from life science to computational physics. The second stage of ChinaGrid project will be from 2005 to 2007, covering 20 to 30 key universities in China. The focus will extend from computational grid applications to information service grid (e-info), including applications for distance learning grid, digital Olympic grid, etc. The third stage will from 2007 to 2010, extending the coverage of ChinaGrid project to all the 100 key universities. The focus of third stage grid application will be even more diverse, include instrument sharing (e-instrument).

3 Design Philosophy of ChinaGrid Support Platform (CGSP)

The underlying common grid computing platform for ChinaGrid project is called *ChinaGrid Supporting Platform* (CGSP), supporting all above three different stages grid applications, they are e-science, e-info, and e-instrument. CGSP integrates all kinds of resources in education and research environments, makes the heterogeneous and dynamic nature of resource transparent to the users, and provides high performance, high reliable, secure, convenient and transparent grid service for the scientific computing and engineering research. CGSP provides both ChinaGrid service portal, and a set of development environment for deploying various grid applications.

Fig. 1. Site Distribution for the Initial ChinaGrid Project

The detail software building blocks for CGSP are shown in Fig. 2. The current version, CGSP 1.0, is based on the core of Globus Toolkit 3.9.1, and is WSRF [19] and OGSA [13] compatible. According to the roadmap of CGSP development, CGSP 1.0 is to be released in October 2004. There are five 5 building blocks in CGSP 1.0. They are:

1. Grid portal: grid portal is the entrance for the end user to use grid services. By using grid portal, users can submit their jobs, monitor the running of jobs, manage and transfer data, inquiry the grid resource information. Grid portal also provides other facilities such as user management and accounting of grid resource usage.
2. Grid development toolkits: they provide toolkit to pack the resource to grid services, the deployment and management toolkit for grid, and programming model to deploy complex grid application in grid environment.
3. Information service: it is responsible for the management of various resources within grid environment, provides a global resource view and grid information services, and updates grid resource information in real time manner. The main

purpose is to provide real time information of various grid resources for end users and other modules in grid environment.

4. Grid management: it provides basic support for various jobs in grid environment. It consists four parts:
   - Service container: it provides a grid service installation, deployment, running, and monitoring environment on each node in grid environment. It also provides necessary support to monitor the real time resources status of each grid node.
   - Data manager: it is responsible for the management of various storage resources and data files in grid environment. It provides a global file view, so that users can access various data files transparent.
   - Job manager: based on information services and data management, it provides support for job management, scheduling, and monitoring for end users' computational task, so that data and resources can be accessed transparently within grid and cooperative working among distributed resources.
   - Domain manager: ChinaGrid is organized in domain. A domain refers to a independent grid system to provide services to the others. A domain can be a specialized grid, or a regional grid. The main function of domain manager is responsible for user management, logging, accounting within domain and interacting with other domains. It makes the domain administrator easily manage the users, services, and resources within domain, and interactive policies among domains.
5. Grid security: it provides user authentication, resources and services authorization, encrypted transmission, and the mapping between users to resources authorization.

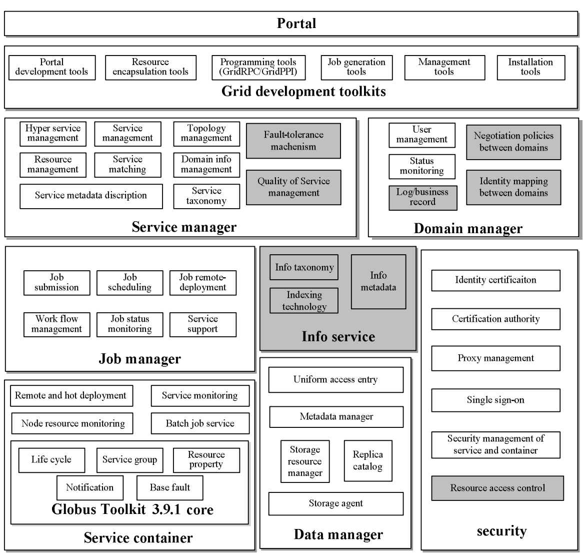

Fig. 2. The Software Building Block of CGSP System

4 Grid Computing Application Platforms

Besides the development of CGSP, another major task for ChinaGrid project is the deployment of grid computing application platform and some pilot grid applications, this makes ChinaGrid very unique to other existing grid projects. In the first stage of ChinaGrid project, five ongoing main grid computing application platforms are under development. They are: bioinformatics grid [3], image grid [6], computational fluid dynamics grid [4], course online grid [5], and massive data processing grid [7]. The first three grids are actually computational oriented. The course online grid is very unique grid, which is the first step to information service grid. The last one, massive data processing grid, is the data grid in nature [9]. Next, I will discuss these grid application platforms in detail.

4.1 Bioinformatics Grid

Bioinformatics merges biology, computer science and information technology into one discipline. The ultimate goal of the field is to enable the discovery of new biological insights. It mainly focuses on grasping biological meaning of plentiful biological data. Bioinformatics uses database, data processing method and software to get results through mass computation. Each bioinformatics research institute holds their own computing facilities, software resources and storage devices. They setup their own research platform independently. Although most of these resources are connected with the Internet, they are only used by their owners, and the utilization rate is low. At the same time, many researchers in Chinese universities have no such research resources to use.

The bioinformatics grid integrates heterogeneous large-scale computing and storage facilities within ChinaGrid to provide bioinformatics supercomputing services for bioinformatics researchers through the Web interface, which is very user-friendly. There are two steps and only two steps are required for users to use bioinformatics grid through this portal. First, input the computing requests according to the submission form from web page. Second, get the computing results from bioinformatics grid. Where the requests are submitted, how the computing requests is allocated, and the monitoring and management for computing tasks are all completed by the bioinformatics grid itself.

When bioinformatics grid server receives the computational requests from the client, it locates a suitable node in the grid to perform the mathematical computation according to the users' requirement and task allocation rule, or integrates a virtual supercomputer to perform the larger computational requests from users.

There are three key modules in bioinformatics grid. They are:

1. Management of Heterogeneous Resources

In bioinformatics grid, resource management consists of both resource abstraction and resource organization. In order to integrate the diverse resources in a uniform and seamless way, all hardware and software resources are packaged into certain services. A set of attributes is abstracted to describe the services (such as OS information, computation power, software category, software name). Objectives of OMR (Organization and Management of Resources) are:

- Integrated heterogeneous computer;
- Make bioinformatics computing tools grid-enable;
- Basic communication mechanisms;

  These mechanisms must permit the efficient implementation of a wide range of communication methods, including message passing, remote procedure call, and multicast.

- Process creation

  This component is used to initiate computation on a computing grid node once it has been located and allocated. This task includes setting up executables, creating an execution environment, starting an executable, passing arguments, integrating the new process into the rest of the computation, and managing termination and process shutdown.

2. Monitoring

The bioinformatics grid provides uniform protocol and interface specification for grid resources monitoring. Monitoring is used to identify system bottleneck. The system can be adjusted and optimized by using the information gathered by monitoring. Objectives of Grid Resources Monitoring (GRM) in bioinformatics grid are:

- Maintaining the latest copy of crucial data in the root node;
- Task scheduling;
- Monitoring bioinformatics grid;
- Keeping log of operations, errors and statistics.

3. Integration of Heterogeneous Bioinformatics Computing Tools

At present, many bioinformatics computing tools have been designed for different application purposes. All of these tools are installed into the grid hardware environment and are used through the Web interface easily and widely. In bioinformatics grid, objectives of Integration of heterogeneous Bioinformatics Computing Tools (IBCT) are:

- Classification and installation of shared bioinformatics computing and research tools;
- Integration of bioinformatics database;
- Pipeline running control of computing flow;
- Graphic user interface;
- Billing system.

4.2 Image Grid

The image grid is a grid image-processing platform based on ChinaGrid infrastructure, which is an integrated environment hiding inner heterogeneous resources and dynamic information. It not only realizes cooperative characteristic, but also affords a secure and transparent image-processing environment with correlative services [23].

The remarkable characteristics of the image-processing platform are: application development, runtime environment, and remote visualization.

The image processing grid application development and runtime environment is based on web service and grid service technology, in which image processing toolkits or software are deployed as web services or grid services components. These services are usually deployed redundantly for all kinds of reasons such as fault-tolerance, availability, and performance. The concept of virtual services, which are the abstract of several physical services, is introduced to implement common functions and interfaces. Therefore during the development of image processing grid application, users just care about the virtual services of image applications, which then convert into different physical services running on ChinaGrid through service-scheduling system.

Various image-processing functions are afforded and extended in the image grid environment with much higher performance and QoS. Currently, the following applications are integrated into image grid as pilot applications: reconstruction of digital virtual human body, image processing of remote sensing, and medical image diagnoses.

The Virtual Human Project (VHP) is the creation of complete, anatomically detailed, three-dimensional representations of the normal male and female human bodies. The long-term goal of VHP is to describe perfectly the conformations and functions of gene, protein, cell and organ so as to achieve the whole accurate simulation of body information in the end.

The basic realization of digital virtual human is to identify and classify inner viscera, rebuild the edge profile data of outer body and three-dimensional profile data of inner viscera, then construct high-precision digital human grid, and use these data to get the visualization for a human body. In addition to the massive amount of body slices (a suit of raw data requires 20GB storage space), the dada amount will increase greatly for mesh construction. Only a computer resource grid can afford its processing. Before we analyze and dispose these data, some optimized resource scheduling algorithm must be bring forward, such as high-precise algorithm for contour extraction and high-precise parallel algorithm for reconstructing three-dimensional grids of digital virtual human.

Remote sensing can be defined as the collection of data about an object from a distance. Geographers use the technique of remote sensing to monitor or measure phenomena found in the Earth's lithosphere, biosphere, hydrosphere, and atmosphere. Remote sensing imagery has many applications in mapping land-use and cover, agriculture, soils mapping, forestry, city planning, archaeological investigations, military observation, and geomorphological surveying, among other uses.

Medical imaging processes are dedicated to the storage, retrieval, distribution and presentation of images, which are handled from various modalities, such as ultrasonography, radiography, magnetic resonance imaging, positron emission tomography, and computed tomography. It consists of a central server which stores a database containing the images. Doctor and his patients can get more reliable and precise medical images (such as X-ray slice, CT and MR) with more quick, exact and simple mode, so as to increase the rate of inchoate morbid detection and diagnoses and the living opportunity.

4.3 Computational Fluid Dynamics Grid

Computational Fluid Dynamics (CFD) grid provides the infrastructure to access different CFD applications across different physical domains and security firewalls. By defining a standard CFD workflow and general interfaces for exchanging mesh data, the platform facilitates interoperating between CFD applications. An application development packages (ADP) is developed and an application development specifications (ADS) is defined shipping with the platform to make migrating CFD applications feasible. The major function of ADP is to schedule jobs, recover from fault and discover service effectively.

The CFD platform includes two levels: high level for Service Oriented Application (SOA) and low level for parallel computing. The physical architecture of the CFD platform is shown in Fig.3.

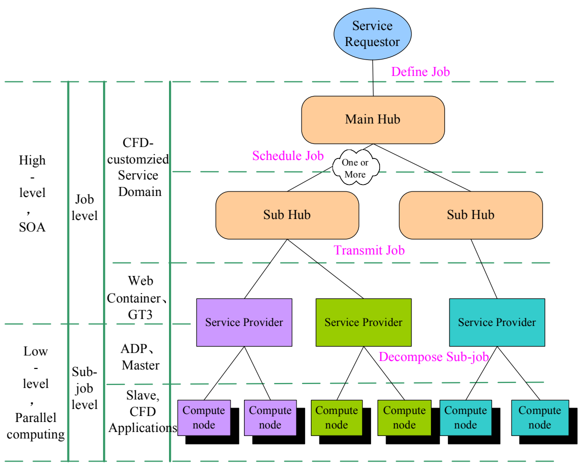

Fig. 3. Physical Architecture of CFD Grid Platform

In the high level, all services are registered in the Main hub or Sub hubs. The function of the Main hub is to provide job definition. Service requestor visits the Main hub for the service interface definition and finds the invocation entrance. Main hub is also responsible for the task scheduling and resources allocation. It collects all the service information from Sub hubs and then find the best resource solution for a given service requestor. The Sub hubs also have those functions which Main hub has, but within a smaller area. Sub hubs delivery the sub tasks (jobs) to the specific service provider, and then collect the results. All the services are provided as web service. By this way (hierarchy architecture and general service interface), the SOA has a good scalability.

In the low level for parallel computing layer, all kinds of parallel environments could be employed. In this system, MPICH2 is used as the low layer parallel environment. The service provider decomposes tasks that distributed by Sub hubs and sends them to computing nodes via MPI mechanism. This feature enable us utilize the distributed computing nodes as a whole to provide scalable and stable services.

4.4 Course Online Grid

Course Online Grid (*realcourse* in short) is a video stream service supported by a collection of physical servers distributed all over China [22]. Since its birth in April of 2003, *realcourse* has been in non-stop operation for more than one and half year, in spite of some of the servers up-and-down, in-and-out from time to time. As of this writing, *realcourse* is composed of 20 servers spanned over 10 major cities distributed on CERNET with 1500 hours quality course videos, which was grown up from initially 4 servers within Peking University campus. As a result we normally observe 20-100 on line users at any time. In turn, these user activities help us to bring the service to a level of robustness.

The 1500 hours of videos are from different universities within ChinaGrid. They are uploaded to different servers with a permanent backup copy created in a server other than original one, and flow from server to server on demand. If one server is down for some reason, its duty is automatically taken up by some other server(s). And once it is up again, a maintenance session will occur automatically to bring it up to a consistent state, in case some transactions occurred during its down time (for example, some more videos were uploaded to the system).

Realcourse is a successful application of distributed computing technologies in a geographically wide area. Different from some traditional distributed fault-tolerant services, *realcourse* emphasizes on giving clients access to the service - with reasonable response times - for as much of the time as possible, even if some results do not conform to sequential consistency. *Realcourse* owns some distinguished characters:

Realcourse aims at non-critical application of video on demand and file storage in which temporal inconsistency is acceptable only if consistency of whole system will finally reached. By eliminating the "update" operation, *realcourse* greatly reduce the consistency requirement. The length of exchanged message among servers is kept constantly small, no heavy communication traffic is observed with the growth of server numbers.

The servers in *realcourse* are not equal to each other. Only two servers are chosen to keep two permanent physical copies of files. No effort is made to keep track of those temporal copies of files but the existence of these copies greatly improve the performance of downloading operation in a way of wide area network cache. In the new version of consistency-check procedure a pervasive search is started to find those "lost" copies in case that both two permanent copies of file are corrupted.

The loop topology of all servers makes it possible for each server to run a consistency-check procedure independently without any overhead. Actually, a more sophisticated consistency-check procedure is in development at the time of this publish. Knowledge from area of “autonomous computing” does much help.

By exploiting the reliable communication provided by Asynchronous Messaging Middleware, *realcourse* hides the failures of network, which is not rare by our observation in a wide area network from servers. The consistency of servers is eventually kept in a “delayed” fashion.

4.5 Massive Data Processing (MDP) Grid

MDPGrid includes three data intensive grid applications: High energy Physics Computing, *Alpha Magnetic Spectrometer* (AMS) Experiment, and University Digital Museums.

High energy physics computing based on MDPGrid, is a solution for the processing and analyzing of the massive data, which are generated by CERN's ATLAS experiment, Sino-Italy cooperated cosmic rays experiment at YangBaJing, Tibet, and BEPC/BES project, at Shandong University.

AMS experiment project is the large-scale physics experiment on *International Space Station* (ISS), with the main purpose to look for the universe composed of antimatter, search for the source of darkmatter, and measure the source of cosmic ray. Space Detector Type Two (AMS-02) will send the space shuttle in 2005 onto ISS to carry on the experiment for three to five years. The data collected by AMS-02 will be finally stored, indexed and analyzed in *Science Operation Centers* (SOC). The Data Processing Center in Southeast University will be able to directly receive the data sent by NASA and carry on data analysis independently. Currently, SOC in SEU has put up a grid platform SEUGrid for the Monte-Carlo simulation computing of AMS-02 MC.

University Digital Museum Grid is to integrate the enormous dispersed resources of various digital museums, to share the resources effectively and eliminate the information island, to filter and classify the collection information, and to provide appropriate information service to users according to their knowledge levels and motivation through unified grid portal. For the time being, the university digital museums involved include the Digital Museum of Aeronautics and Astronautics (BUAA) [15], the Archaeological Digital Museum (SDU) [14], the Geoscience Digital Museum (NJU) [16], and the Mineralogical Digital Museum (KUST) [17]. The services and resources in each university digital museum compose one site. The collection provider service gives access to digital museum’s collections as raw data.

5 Conclusion

Although ChinaGrid has just begun for two years, it has gained much attention from all aspects. Just within one year, the coverage of ChinaGrid project has already been to 20 universities in China. Besides, some provinces and cities are trying to deploy their regional grids. CGSP, which is the potential grid computing middleware candidate to use, makes these regional grids as the extension of ChinaGrid system. All these together with above grid computing application platforms and pilot applications make ChinaGrid very unique from other existing grid computing projects. We are very confident that ChinaGrid will be the largest, the most advanced, and the most practical grid in the near future.

Acknowledgement

Special thanks to following people providing materials for this paper of each ChinaGrid application platform: Professor Weiming Zheng from Tsinghua University, Professor Xiaoming Li from Beijing University, Dr. Xiaowu Chen from Beihang University, Professor Minglu Li from Shanghai Jiao Tong University, and Prof. Qin Zhang from Huazhong University of Science and Technology. I also would like to take this opportunity to thank all the experts in ChinaGrid project for their devoting time to this project, and making grid computing a reality in China.

References

1. China Education and Research Network, http://www.edu.cn/
2. ChinaGrid, http://www.chinagrid.edu.cn
3. ChinaGrid Bioinformatics Grid, http://166.111.68.168/bioinfo/tools/index.jsp
4. ChinaGrid Computational Fluid Dynamics (CFD) Grid, http://grid.sjtu.edu.cn:7080/grid/
5. ChinaGrid Course Online Grid, http://realcourse.grids.cn
6. ChinaGrid Image Processing Grid, http://grid.hust.edu.cn/ImageGrid/
7. ChinaGrid Mass Data Processing Grid, http://athena.vrlab.buaa.edu.cn/gcc/
8. China National Grid, http://www.863.org.cn/863_105/applygide/applygide2/information2_app/200206270024.html
9. Data Grid Project WP1, "Definition of Architecture, Technical Plan and Evaluation Criteria for Scheduling, Resource Management, Security and Job Description", *Datagrid document DataGrid-01-D1.2-0112-0-3*, 14/09/2001.
10. I. Foster and C. Kesselman, "Globus: A Metacomputing Infrastructure Toolkit", *International Journal of Supercomputer Applications*, Vol.11, No.2, pp.115-128, 1997.
11. I. Foster, C. Kesselman, and S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations", *International Journal of High Performance Computing Applications*, 15 (3), 200-222, 2001.
12. W. E. Johnston, D. Gannon, and B. Nitzberg, "Grids as Production Computing Environments: The Engineering Aspects of NASA's Information Power Grid", *Proceedings of 8th IEEE Symposium on High Performance Distributed Computing*, 1999.
13. Open Grid Services Architecture, http://www.ggf.org/Public_Comment_Docs/Documents/draft-ggf-ogsa-specv1.pdf.
14. The Archaeological Digital Museum, http://museum.sdu.edu.cn/index/index.asp
15. The Digital Museum of Aeronautics and Astronautics, http://digitalmuseum.buaa.edu.cn/
16. The Geoscience Digital Museum, http://202.119.49.29/museum/default.htm

17. The Mineralogical Digital Museum, http://www.kmust.edu.cn/dm/index.htm
18. The TeraGrid Project, http://www.teragrid.org/
19. The Web Services Resource Framework, http://www.globus.org/wsrf/.
20. UK e-Science Programme, http://www.rcuk.ac.uk/escience/
21. G. Yang, H. Jin, M. Li, N. Xiao, W. Li, Z. Wu, Y. Wu, and F. Tang, "Grid Computing in China", *Journal of Grid Computing*, accepted.
22. J. Zhang and X. Li, "The Model, Architecture and Mechanism Behind *Realcourse*", Technical Report, Beijing University, 2004.
23. R. Zheng, H. Jin, Q. Zhang, Y. Li, and J. Chen, "IPGE: Image Processing Grid Environment Using Components and Workflow Techniques", *Grid and Cooperative Computing – Lecture Notes in Computer Sciences*, Vol.3251, 2004.