

An Example of Application-Specific Portal in Web-based Supercomputing Environment*

Hong Wu¹, Xue-bin Chi^{1,2}, Hai-li Xiao¹, Tie-jun Zhang³, Sun-gen Deng¹, Bo Chen^{1,2}, Hong-hai Zhang¹

¹Supercomputing Center of Computer Network Information Center, Chinese Academy of Sciences

²The Software Institute, Chinese Academy of Sciences

³BaoTou Water Supply Company, The Inner Mongolia Autonomous Region

Email: {wh, chi, haili}@jupiter.cnc.ac.cn, ztj1@163.com, {dsg, chb, zhh23w}@jupiter.cnc.ac.cn

Abstract

As the Grid computing takes more and more important role in distributed computing, the need for Web-based portal that provide a uniform interface to the heterogeneous HPCs is increasing. A convenient interface, called Scgrid portal is developed, which is based on Globus Toolkits (GT), Java CoG, and Grid Portal Development Kits (GPDK). The architecture and design of this portal are introduced in detail in this paper, and finally some future work about our Computational Portal building is set forth.

1 Introduction

Grids provide the infrastructure to harness the heterogeneous resources in the same or across different domain(s), which give us a good method to utilize the scattered resources in dynamic, distributed “virtual organizations” (VOs). Although more and more scientific and commercial attention has been put since the concept of “Grid” proposed in the 1990s, and especially in these two years, great interest has been taken by a lot of famous IT companies, such as IBM, Microsoft, Sun etc., there is not a clear definition of “Grid”, but the essence of its definitions can be captured into the three point checklist[1]: a) coordinates resources that are not subject to centralized control ..., b) ... using standard, open, general-purpose protocols and interfaces ..., c) ... to deliver nontrivial qualities of service.

The “Grid” has solved the need of utilizing the scattered resources and increased the computing capacity, there are still a variety of reasons for the lack of use of high-performance environment to solve computational problems, such as the sophisticated and inadequate computing environment is the other factor that blocks the use of supercomputers. So a feasible Problem Solving Environment (PSE) that aims at solving the problem of the computational environment is needed to proposed, the CoG Kits[2], a bridge between commodity distributed computing and high-performance grids is developed.

There are 3 high performance computers in our center. They are SGI Power Challenge/16 processors, Hitachi SR2201/32 processors, and Dawning-2000/164 processors, respectively. The first one, SGI Power Challenge with 16 processors, was installed in 1996. The second one, Hitachi SR2201, was installed in 1998 with 32 processors. The third, Dawning 2000 with 164 processors, which is made by National Center for Intelligent Computing System, was installed in 2000. They have provided 6 million CPU hours for the scientific computing research in China till now.

Scgrid portal is an application-specific portal which provide computing service via Internet in our

¹ This work is supported by The Special Funds for Major State Basic Research Projects of China (G1999032805), State Hi-Tech Research and Development Program of China (2001AA111043)

center. What we want to develop at last is a high level web-based interface which aims at high performance computing, shields all different operations of HPCs and provides 1-2 specific application services.

2 Related Work

Although there are a lot of middleware to build the Grid, the most famous two are Globus[3] and Legion[4], and now the Globus is becoming a representative and an accepted standard. Besides of the Globus Toolkits, which is developed by Argonne National Laboratory's Mathematics and Computer Science Division and the University of Southern California's Information Sciences Institute since 1996, a number of efforts and projects have been explored in developing more convenient environment via Internet-based computing, such as the CoG, which can be used to develop applications that tie advanced network-accessible resources into the commodity desktop, Myproxy[5], which is an online credential repository for the grid, the GridPort[6], GPDK[7], and Gridsphere[8], which is Globus Toolkits 3 and Web Services specific), which allow construction of application-specific portals by providing generic interfaces/libraries for accessing Grid resources using Globus. GridBank and G-monitor[9], which provide Grid accounting and monitor application execution on distributed systems respectively, PBSWeb[10] provide web portal for PBS[11].

3 Architecture and Design

3.1 Architecture

Scgrid portal resides on a web server, which sits between the HPCs in our center and Web browser of end users, the end users use the Web browser to access the provided computing services, the Grid nodes is managed by Globus Toolkits, and in order to bridge the incompatibility between Web and Grid security, a Myproxy Server is also needed. The architecture is depicted in Figure 1.

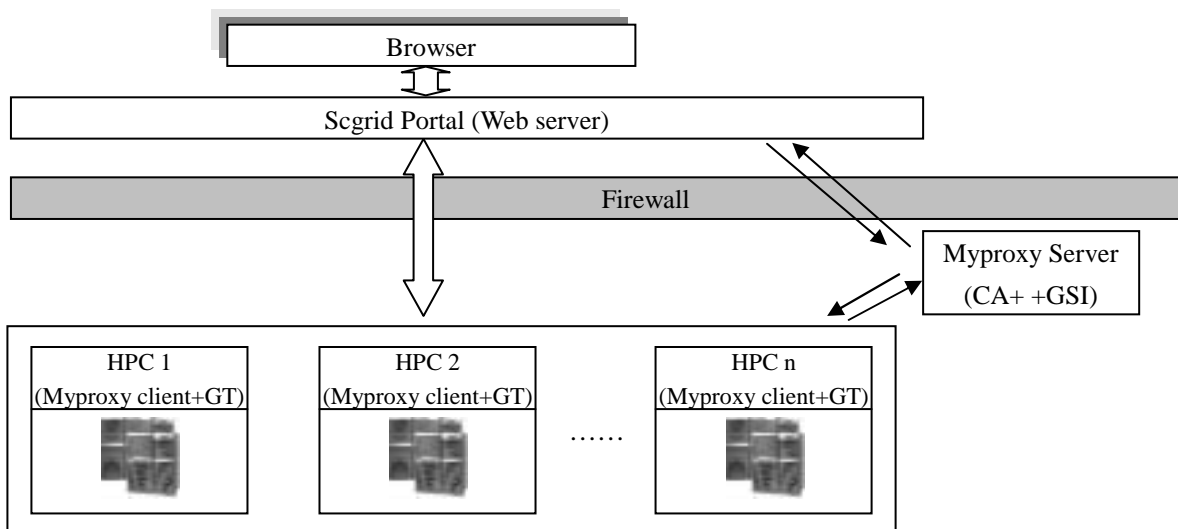


Fig. 1 Scgrid Portal Architecture

3.2 Design

Users we referred here is not only limited to the end users, but also the administrator in our Grid domain (the situation of across Grid domain will not to be considered temporarily), so the function of Scgrid portal can be categorized into two main modules, one is for ordinary end users, the other is for administrator of our Grid domain, which are called module of user computing service and module of

administrator service respectively.

The function of module of user computing service includes: file service/resource service/job service/accounting service, the former three have been finished, and we should add the visualization interface and 1-2 interface of specific application in the future. The function of module of administrator service includes: user manage/log manage/accounting manage/other daily operation, the former two had been finished, but the user manage just limited inner Grid domain and based on the assumption that there is already an real accounts on each HPCs to access resources. The disadvantage of this realization is obvious, such as unsupport in uniform manage, unscalability, difficulty in accounting, etc., but a real user account in each of HPCs is one of the fundamental demand of Globus Toolkits, so it still needs us to explore how to find a feasible method not only considering the compatibility of portal infrastructure based on GT/Myproxy/Java CoG but also is an ideal user manage model which is adaptive inner- and across- Grid domain. How to solve the problem of global user mapping inner and across the Grid domain and put this into practice of portal development is the next step of our project.

To get a better understanding of the Scgrid portal, we can walk through a work flow as below:

1. The user fills in a registration form and submits it to our web portal.
2. Some prerequisite login information, such as username and password, will be sent to him according to the email address he provided in 7 workdays.
3. Then, he logs in our portal in IE or Netscape with his username and password.
4. After logs in and passes the authentication and authorization successfully, he could submit jobs to a desired HPC, examine result of submitted jobs, check status of job queues, view information of resources.

4 Implementation

One of the main aims of developing Scgrid portal is to provide a web-based interface which is easy to use, there has been a variety of efforts and projects to explore the computational Grid portal, such as the Java CoG, GridPort, GPDK, which mentioned before, they are all robust and flexible middleware and can be used as rapid prototype tools in building portal. The platform we based on is GT 2.0 (although the GT 3.0 was released on 30, June, 2003, but which have full compatibility with Java CoG is GT2.0, as shown in Fig. 2), So the fundamental middleware is GT2.0. The web server environment is: JDK+Tomcat+Apahce+GPDK, Myproxy server 0.4.4. The programming and scripting language we used for was Java and Javascript. And the demo version of our portal will be released at the end of 2003.

Java CoG Kit Version	Globus Toolkit Services									
	Job submission (GRAM) (1)					File Transfer (GridFTP) (2)				
	2.0	2.2.2	2.2.3	2.2.4	2.4.0	2.0	2.2.2	2.2.3	2.2.4	2.4.0
0.9.13	pass	N/A	N/A	N/A	N/A	pass	fail	fail	fail	fail
0.9.13 + jftp	pass	N/A	N/A	N/A	N/A	fail (3)	pass	pass	pass	pass
1.1a	pass	pass	pass	pass	pass	fail (3)	pass	pass	pass	pass

Fig.2 Results of the test of Java CoG with Globus Toolkit Services

5 Conclusion and Future Work

Operation of computing has experienced the console to telnet, till now distributed computing, and the Grid is the research hotspot of distributed computing since middle period of last century. Scgrid is a web-based supercomputing portal, which provide uniform interface to access HPCs in the supercomputing center of Chinese Academy of China, and now a lot of fundamental function of telnet method had been implemented. And our future work will focus in the following:

1. interoperating computing:

Change the vendor-MPI on our HPCs now into mpich, and configuration of Nimrod/G[10] and GT, There is a lot of work should be do to let our scgrid portal to be a real computational Grid.

2. Uniform user management

User management is a very critical problem in Grid computing environments, a new model have been put forward by us[11]. How to solve the contradiction between global user mapping and mechanism of Myproxy and Java CoG in user management, which still needs us to explore.

3. Strong compatibility

The basis of our portal is GT2.0 which released in Feb, 2002, with application and application communication via WWW become more and more used, Web services become more and more prevail and the release of the GT3.0 based on Web service as scheduled, how to realize the smooth transfer of our portal between GT2 and GT3 is also a problem.

Availability

Website of our center: <http://www.sccas.cn>

Website of our portal: <http://www.scgrid.cn>

References

- [1] I. Foster, What is the Grid? A Three Point Checklist, <http://www.globus.org>
- [2] G. Laszewski, I. Foster, J. Gawor, CoG Kits: A Bridge between Commodity Distributed Computing and High-Performance Grids, Proceedings of the ACM java Grande Conference, 2000
- [3] The Globus Projects: <http://www.globus.org>
- [4] The Legion Projects: <http://legion.virginia.edu/>
- [5] J. Novotny, S. Tuecke, Y. Welch, An Online Credential Repository for the Grid: Myproxy, Proc. 10th IEEE Symp. On High Performance Distributed Computing 2001
- [6] J. Boisseau, M. Dahan, S. Mock, K. Mueller, D. Sutton and M. Thomas, The Gridport Toolkit: A System for Building Grid Portals, Proc. of the 10th IEEE Intl. Symp. on High Perf. Dist. Comp 2001
- [7] J. Novotny, The Grid Portal Development Kit, Concurrency: Pract. Exper. 2000; 00:1-7
- [8] The Gridsphere Project, <http://www.gridsphere.org>
- [9] The Gridbus Project, <http://www.gridbus.org>
- [10] PBSWeb, <http://www.cs.ualberta.ca/~pinchak/PBSWeb/>
- [11] The OpenPBS Project, <http://www.openpbs.org>
- [12] The Nimrod project, <http://www.csse.monash.edu.au/~davida/nimrod.html>
- [13] Bo Chen, Xue-bin Chi, Hong Wu, A Model for User Management in Grid Computing Environments (To be accepted)