Grid Portal System Based on GPIR

Fang Juan[1], Geoffrey Fox[2], Marlon Pierce[2]

[1] *College of Computer Science, Beijing University of Technology,*
100022 Beijing, China
[2] *Community Grids Lab, Indiana University*
Bloomington, IN 47404, USA
fangjuan@bjut.edu.cn ; {gcf, mpierce}@indiana.edu

Abstract

Grid portal is the bridge between Grid and user. In this paper a Grid portal system is set up based on GPIR of GridPort. The Grid portal system provides an efficient means to user for utilizing grid resources, and provides advantaged condition for achieving effective information of database. The user can customize the application by portlet framework. In this paper we also review multiple choices for portal technology and explain why we choose GPIR.

1. Introduction

A Grid portal is an application server that can provide a secure online environment for gathering information about grid services and resources. There are many projects about Grid portal, such as GridPort, GPDK, Legion Portal, GridSphere, User-friendly Web GUI and advanced grid services are provided. Grid portal can provide a service uniform platform for user, and integrate the services. The user can select different services according to the demand, the Grid portal can conceal the details of grid resources furthest for users, and then the users can utilize and harmonize the resources adequately.

2. The Basic Framework of Grid Portal and GPIR

The basic model of Grid portal[1] can be classified into three layers from the point view of service. The first layer is user layer, it can provide the interface for user. User layer is responsible for displaying the portal content, it can be web browser, or other desktop tools. The second layer is service layer, including authentication, job management service, information service, file service, security service. The third layer is resource layer, including remote compute, data and application resources. Service layer commits the Grid I/O operation to relevant grid resource, and accept the executed result from grid resources.

We use the GridPort as the basic framework. The current version of GridPort is 4.0, developed by TACC(Texas Advanced Computing Center). Grid Portal Toolkit(GridPort) is a toolkit, it can be used to develop the portal which based on web, and it can create the application on the infrastructure of the distributed and grid computing. The GPIR(Grid Portal Information Repository) and CFT(Comprehensive File Transfer Service) web services have been extracted from GridPort and can now be installed separately. GPIR provides a place to store data about your grid that is readily accessible to a portal application. GPIR uses the relational database to store and retrieve the grid data with Web service interface.

3. The Grid Portal Architecture Based on GPIR

We designed a Grid portal architecture based on GPIR, it has three layers[2]. The bottom layer is Information Providers, called resource layer; the middle layer is database, the top layer is client layer. The resource layer provides the resources, it is on the Grid environment. The resources of bottom layer will provide the number of resources, status, and relational information of processing capability to portal and application, the data of grid can be provided by the service to user. The data will be stored to database. In the architecture, resource layer includes the data from the grid resource monitoring system, MDS and so on, the data from the systems will be stored into GPIR, and then displayed to the clients via Grid portal. Data is extracted from GPIR using the Query web service is a query interface to GPIR. It allows you to query GPIR

by a single resource or by a virtual JDBC, it is used to connect the database to User Interface. Almost all the operating system support Java, JDBC is also provided by Java. Fig 1 shows the architecture.
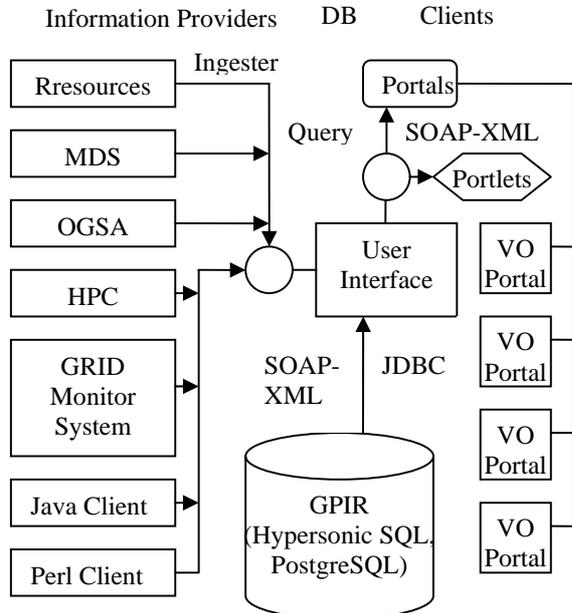
Information Providers   DB   Clients

Figure 1. Grid portal architecture based on GPIR

There are two databases used by GPIR, one is HypersonicSQL, the other is PostgreSQL. Hypersonic SQL is written entirely in Java, so it will run on all the Java platforms; it is compact, weighing only 155 K, since it is an open source project, there is no per CPU license. To summarize, Hypersonic SQL is well adapted for standalone desktop applications written in Java. It is also useful for testing and prototyping. PostgreSQL is originally a database management system that developed at the University of California at Berkeley. The concept of object-relational database is even earlier than many commercial databases. It supports the SQL92/SQL3 language; it also has the integrality of transaction, type of expansibility and so on. The user can access all the services provided by Grid portal by Web Browser or other application. When the request of user gets to Grid portal, it can schedule special portlet to respond the request.

4. Comparison

The GridPort system has several excellent features[3]: the software is written in Perl/CGI, it is portable and runs on most webservers; it supports single login between multiple portals; it is flexible and adaptable; it is easy to support and modify; and portals built with it may run across multiple sites and organizations. The NPACI GridPort Toolkit is a simple, robust, and flexible system. The GridPort software can be extended to support the web-services architecture that is being developed for commercial purposes and implemented in commercial technologies such as JXTA, SOAP, and WSDL. GridPort-based portals require no software downloads or configuration changes on the client side, and run on common web browsers.

Compare it with GridPort, GridSpeed anticipates that the reduced cost and time associated with creating and deploying these portals. GridSpeed should make portal development easier and more appealing to scientists which to migrate their applications to the Grid. Gateway[4] requires that a Gateway server run on each HPC system. GridPort avoids this complexity by using the web server to process and build Globus commands and then communicate directly with the Globus processes on the HPC systems. The Grid Portal Development Kit (GPDK) facilitates the development of Grid portals and provides several key reusable components for accessing various Grid services. It's a Java Servlet implementation with many similar functions to those of GridPort.

5. Conclusion

Grid portal system provides a visualized view and the ability of accessing grid resources for users. The design of this paper is based on the architecture of GPIR, so it avoids the risk of recreating architecture, and it also has much more effective improve. Future work on the Grid portal system includes enhancing capabilities of existing Grid portal, and extending its function. Another area of future development is to make the VO Portal more detailed. We also can implement the grid resource management and schedule on Grid portal.

6. References

[1]Lifish Chen. "Introduction to Grid Portal", http://www.chinagrid.net/grid/paperppt/USTC, 2004
[2] http://gridport.net/services/gpir/arch.html, 2006
[3] Thomas, M. P., Mock, S., Dahan, M., Mueller, K., Sutton, D., "The GridPort Toolkit: a   System for Building Grid Portals", Proceedings of the Tenth IEEE International Symposium on High Performance Distributed Computing, August, 2001.
[4] Akarsu, G. Fox, T. Haupt, A. Kalinichenko, K. Kim, P. Sheethalnath, and C. Youn. "Using Gateway System to provide a desktop access to high performance computational resources", Proceedings of the Eight IEEE International Symposium on High Performance Distributed Computing, August, 1999.