Grid Computing in Hong Kong: Research and Development

Francis C.M. Lau, Roy S.C. Ho, and C.L. Wang
Department of Computer Science and Information Systems,
The University of Hong Kong
Email: {fcmlau,scho,clwang}@csis.hku.hk

Abstract

Grid is a new exciting arena for re-application of many distributed computing techniques from the past as well as brand new ideas because of new requirements imposed upon by a form of computing that was never possible before. We present the Hong Kong Grid which is a major grid platform for local institutions to conduct their R&D projects.

1. Introduction

Over the last decade, computing has become increasingly concerned with collaboration, data sharing, and other new modes of interaction that involve distributed resources. While no one would argue that this paradigmatic shift was mainly due to the phenomenal growth of the Internet, the emergence of grid computing technologies actually play an important role. Grid computing technologies enable sharing and coordinated use of distributed computing resources [4]. Compared to traditional computing platforms, grids feature superior cost-effectiveness and flexibility [3].

In this position paper, we present the Hong Kong Grid (HKGrid), a grid test bed that links up a host of HPC facilities in the local institutions. We also describe the research and development activities to address various challenges in grid computing.

2. The Hong Kong Grid

The HKGrid project [14] is to construct and make available a grid test bed to facilitate the development of grid middleware and applications by local industry and institutions in Hong Kong and their partners in the region. The purpose is to demonstrate the benefits of adopting grid technologies and to showcase any outstanding results of development or application. Table 1 summarizes the current constituents and computing facilities of the HKGrid. The Hong Kong Academic and Research Network (HARNET) [13] is used as the backbone interconnect. The HKGrid currently consists of more than 500 compute machines, offering a theoretical maximum computing power of about 4 Tflop/s. Furthermore, through linking to the Asia-Pacific Grid (ApGrid) [9] and the China National Grid (CNGrid) [10], as well as an Internet2 connection to the Abilene backbone at Chicago, USA [11], the HKGrid plays the role of a gateway for these other bigger grids. These three grid systems in the Asia region collectively present a powerful computing platform which involves 50+ institutions spanning 13 countries.

The HKGrid forms a real grid environment for supporting the R&D projects as well as pilot applications in the participating institutions. In the following sections, we briefly describe some of these projects.

Table 1: Members and Computing Facilities in HKGrid

Institution (Department)	Computing facilities
City University of Hong Kong (CS)	Service gateway
Hong Kong Baptist University (CS)	2-way Xeon SMP x 64; Gigabit Ethernet (ranked #300 in TOP500 [15] in 6/2003)
Hong Kong University of Science and Technology (CS)	4-way SMP cluster
The Hong Kong Institute of High Performance Computing	Service gateway
The Hong Kong Polytechnic University (Computing)	Service gateway
The University of Hong Kong (Computer Centre)	2-way Xeon SMP x 128; Gigabit Ethernet (ranked #240 in TOP500 in 11/2003)
The University of Hong Kong (CSIS)	Pentium 4 x 300; Fast Ethernet (ranked #340 in TOP500 in 6/2003)

Proceedings of the 10th IEEE International Workshop on Future Trends of Distributed Computing Systems (FTDCS’04)
0-7695-2118-5/04 $20.00 © 2004 **IEEE**

3. Projects at HKPU

The Hong Kong Polytechnic University (HKPU) is committed to integrating peer-to-peer (P2P) technologies with grid computing, and supporting fault tolerance in grids.

3.1 Integration of P2P with grid computing

Most of the existing solutions to resource management and task-scheduling in computing grid are based on a traditional client-server model, employing a central administrative server/manager. Since a computing grid provides resources distributed in multiple domains, it should not have a single or central authority for resource management and task scheduling. Existing solutions based on the client-server model use a centralized monotonic mechanism in the scheduling server which is not flexible or scalable. Due to this reason, a P2P-based decentralized approach is proposed, which off-loads the intermediate server by letting the peers in the grid to make the scheduling decision among themselves using their own scheduling policies. A generic architecture for meta-scheduler on peers, called PGS (P2P Grid Scheduler), and the task scheduling framework based on PGS have been developed. Both push and pull modes are used for allocating the tasks to peers with the support of load balancing and fault tolerance.

3.2 Fault tolerance in grid computing

Resource failure in grids is more a norm than an exception. With so many resources in a grid environment, the probability of having some failures is naturally high. Consequently, support for the development of fault-tolerant applications has been identified as one of the major technical challenges in designing production grid systems. In general, a computing grid may contain many heterogeneous subsystems for large-scale, coordinated resource sharing. Each subsystem may, at a lower level, contain many processes performing complex operations which collectively perform a single required function. In such a system, traditional approaches to fault tolerance such as a single rollback recovery scheme for the whole system is inadequate since each subsystem may require different rollback recovery scheme due to the distinct functions and natures of each subsystem. The goal is thus to seamlessly reuse and bridge the existing check-pointing schemes deployed in different subsystems. In this project, a bridging mechanism is proposed which allows subsystems with different rollback-recovery schemes to cooperate with their recovery activities while leaving the underlying recovery protocols unchanged. In addition, a complementary portable recovery scheme is designed which allows a failed process to restart in another subsystem dynamically designated by a grid task scheduler.

4. Projects at HKU

In The University of Hong Kong (HKU), the research effort is focused on the design of an advanced grid platform that aims to address a number of fundamental issues relating to different usage paradigms of grids. In general, the usage paradigms can be classified into two broad categories: service-centric and resource-centric.

In the *service-centric paradigm*, the users access the remote resources through a grid service interface and protocol. This paradigm provides a standard semantic for controlling accesses to remote resources in “virtual organizations” (VO’s) [4]. Nevertheless, existing approaches (such as [5], [6], and [16]) seem to fall short of being able to flexibly *coordinate* the use of distributed resources in a VO, let alone the use of resources across VO’s. The problem could have been due to the lack of (1) VO-based collective operations to coordinate resource usage, and (2) a standard abstraction of VO. Without the support of collective operations that manage and “virtualize” distributed resources, resources could at best be shared in a “point-to-point” fashion, leading to complication for nontrivial forms of resource sharing. Without a standard abstraction of VO, it is nearly impossible to derive universal protocols for inter-VO resource sharing, and hence coordination.

In the *resource-centric paradigm*, resources are offered “as is.” The applications that require certain computing resources are dispatched to where the resources are. Since the users or the grid middleware have certain knowledge of the resources and their present state, most HPC applications fit in this paradigm. Nevertheless, the resource-centric paradigm does introduce two fundamental problems. First, load balancing is difficult to achieve in grid systems than in a traditional distributed computing environment because of the heterogeneity in and dynamic nature of grids. Second, it is difficult to create a *single system image* (SSI) over the distributed computing resources due to their evasive nature. The SSI support, however, is much desired because the users can make use of the distributed resources as if they were in a single, powerful computer.

Apart from the paradigm-specific issues, there are other issues which arise in general grid computing. First, large-scale grid deployments should support not only

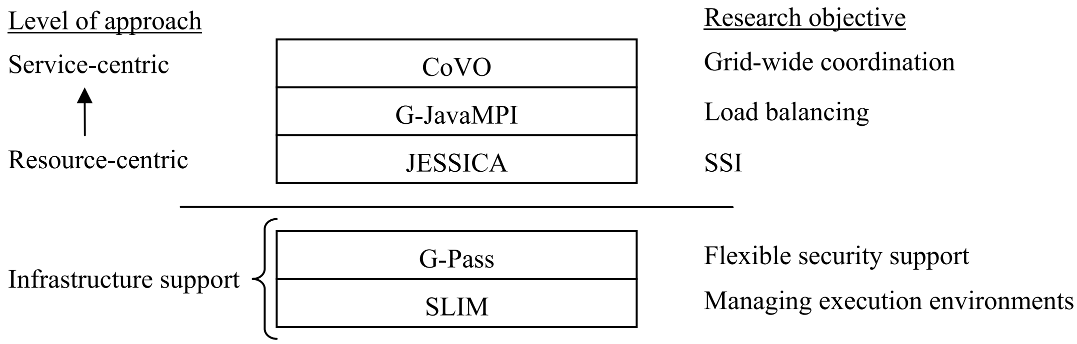

Figure 1: An Advanced Grid Platform

user-to-host authentication (which is the main approach of existing grid middleware), but also grid-wide authentication. This is because if a user's credential is only valid within a single grid point, it would be tedious to enable certain advanced features such as process migration and collective operations since the credential needs to be propagated somehow. Second, based on our participations in the ApGrid, CNGrid, and HKGrid, we realize that managing the execution environments for different collaborators and their applications could be complicated and time-consuming. The induced management hassle might well offset the benefits brought forth by the institutional collaborations. Therefore, it would be highly desirable to support on-demand construction or adjustment of execution environments in order to provide the suitable grid middleware needed by different user applications.

4.1 An advanced grid platform

We have responded to the above research issues by initiating a number of projects, including CoVO, G-JavaMPI, JESSICA, G-Pass, and SLIM. These projects (Figure 1) represent different and yet complementary approaches to providing an advanced grid platform for future applications.

In the **CoVO** project, we aim to design a standard abstraction of VO, and a set of collective, VO-centric operations for managing, virtualizing, and coordinating accesses to distributed resources. They enable flexible and scalable resource sharing among grid entities belonging to the same or different VOs. The research effort will be focused on the following aspects: (1) VO abstraction, (2) VO naming, (3) service virtualization and discovery, (4) a simplified service framework and protocol bindings for grid computing on small devices, and (5) design issues of ad-hoc VO's in highly-dynamic and unreliable computing environments.

G-JavaMPI [1] is a grid-enabled implementation of the Java language bindings of the MPI v1.1 standard. The main goal of G-JavaMPI is to support dynamic balancing of both computational and network load. This is achieved by employing several communication-efficient process redistribution algorithms and a custom mobility support that enables transparent process migration within and across grid points. The decision of Java process migration is made through prediction of future communication patterns of the target applications using bytecode parsing and execution tracing.

JESSICA [7][17] stands for "Java-Enabled Single-System Image Computing Architecture," a middleware that runs on top of the standard UNIX operating system to support parallel execution of Java applications and preemptive thread migration in clusters and grid platforms. With JESSICA, all threads running in the distributed nodes could share all the resources that each thread has created or allocated, and they see the underlying grid platform as a single computing system with multiple processors – i.e., SSI.

G-Pass aims to support dynamic process management and resource access in grids without compromising security. It is a distributed authentication system that enables application mobility on grids, where each process is given a G-Pass credential which is valid within a predefined security context. G-Pass forms a foundation for secure process migration in grids, thus providing the needed support for our G-JavaMPI, JESSICA, and DOS/VM projects.

SLIM [8] is a network service for managing and constructing Linux-based execution environments, and disseminating them to remote computing platforms. It is particularly useful for dynamic and efficient construction of grid points since it facilitates customization of execution environments, therefore reducing the management hassles in dealing with large grid systems. SLIM has been deployed since 2002, and is currently managing the HKU-CSIS grid point [12] and an additional 600+ Linux workstations for undergraduate teaching and research purposes.

4.2 Applications

The HKU-CSIS grid point has been acting as the gateway to the HKGrid for a number of inter-disciplinary projects in HKU that involve computationally-intensive tasks. These projects include (1) modeling of air quality in Hong Kong; (2) bioinformatics computations for approximate string matching, whole genome alignment, and DNA shuffling; (3) speech recognition; and (4) parallel simulation of turbulent flow model.

Besides, the HKU Computer Centre is currently studying the feasibility of adopting grid technologies for consolidating the campus-wide HPC facilities so that the user applications can be deployed seamlessly on any available resources. The goals are to achieve a better utilization of the existing resources, and to realize the concept of *computing-on-demand*. We anticipate that the incremental adoption of grid technologies in HKU would benefit not only the research in computer science, but also other disciplines which traditionally rely on HPC such as the natural science, engineering, statistics and actuarial science, and financial engineering. These applications will create opportunities for us to experiment with and refine our advanced grid platform in production environments. We wish the research output will then be able to enhance the experience of grid users.

5. Projects at the Other Universities

The Centre for E-Transformation Research [2] at the Hong Kong Baptist University (HKBU) is focusing on autonomous grid service composition. The research objectives include distributed resource discovery and management, services matchmaking, and on-demand planning under uncertainty. The research is intended for supporting future applications in the areas of e-business and e-learning.

The Hong Kong University of Science and Technology (HKUST) is addressing the research problems associated with allocation and scheduling of computing resources in computing grids; and the overlay topology optimization problem on grid networks. Furthermore, the university is working with some local secondary schools to develop educational grid systems for the possible dissemination of educational information (e.g., class material, problem sets, etc.) and operational information (e.g., student records).

Acknowledgements

We thank Dr. Jiannong Cao (HKPU), Dr. William Cheung (HKBU), and Prof. Lionel Ni (HKUST) for providing us with information on their projects. The research at HKU is supported in part by the HKU Foundation Seed Grant 28506002 and the HKU Large Equipment Grant 01021001.

References

[1] L. Chen, C.L. Wang, and F.C.M. Lau. A Grid Middleware for Distributed Java Computing with MPI Binding and Process Migration Supports. Journal of Computer Science and Technology (China), Vol. 18, No. 4, July 2003, pp. 505-514.

[2] Centre for E-Transformation Research, Hong Kong Baptist University. http://www.comp.hkbu.edu.hk/~wic-hk

[3] I. Foster, and C. Kesselman (editors). The Grid: Blueprint for a New Computing Infrastructure, 2nd edition. Morgan Kaufmann. 2004.

[4] I. Foster, C. Kesselman, and G. Tsudik. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. International Journal of Supercomputing Applications, 15(3), 2001.

[5] I. Foster, C. Kesselman, J.M. Nick, and S. Tuecke1. The Physiology of the Grid - An Open Grid Services Architecture for Distributed Systems Integration. White Paper, The Globus Project. http://www.globus.org/

[6] A. Grimshaw, A. Ferrari, F. Knabe, and M. Humphrey. Legion: An Operating System for Wide-Area Computing. IEEE Computer, 32:5, May 1999, pp. 29-37.

[7] M.J.M. Ma, C.L. Wang, F.C.M. Lau. JESSICA: Java-Enabled Single-System-Image Computing Architecture. Journal of Parallel and Distributed Computing, Vol. 60, No. 10, October 2000, pp. 1194-1222.

[8] Single Linux Image Management (SLIM). http://www.csis.hku.hk/~cmlee/slim/

[9] The Asia-Pacific Grid (ApGrid). http://www.apgrid.org/

[10] The China National Grid (CNGrid). http://www.chinagridforum.org/

[11] The HARNET-Internet2 Connection. http://www.hku.hk/cc/home/networks/internet2/overview.htm

[12] The HKU-CSIS Grid Point. http://grid.csis.hku.hk/

[13] The Hong Kong Academic and Research Network (HARNET). http://www.jucc.edu.hk/jucc/harnet.html

[14] The Hong Kong Grid Project. http://www.hkgrid.org/

[15] The TOP500 Supercomputer List. http://www.top500.org/

[16] The UNICORE Forum. http://www.unicore.org/

[17] W.Z. Zhu, C.L. Wang, and F.C.M. Lau. A Lightweight Solution for Transparent Java Thread Migration in Just-in-Time Compilers. The 2003 International Conference on Parallel Processing, Taiwan, Oct. 6-10, 2003, pp. 465-472.