On Construction of a Multi-Grid Resource Selection Strategy on Grids

Chao-Tung Yang, Department of Computer Science, Tunghai University, Taichung, Taiwan

Wen-Jen Hu, Department of Computer Science, Tunghai University, Taichung, Taiwan

Kuan-Chou Lai, Department of Computer Science and Information Engineering, National Taichung University, Taichung, Taiwan

ABSTRACT

Grid computing is now in widespread use, integrating geographical computing resources across multiple virtual organizations to achieve high performance computing. A single grid does not often provide a vast resource because virtual organizations have inadequate computing resource restrictions for management on an organizational scale. This paper presents a new grid architecture named Multi-Grid, which integrates multiple computational grids from different virtual organizations. This study builds a resource broker on multiple grid environments, integrating a number of single grids from different virtual organizations without the limit of organizations. The purpose of the multiple-grid resource is to avoid wasting resources. In addition, this study proposes a Multi-Grid Resource Selection Strategy (MRGSS) for the resource broker to better allocate resources before submitting jobs, to avoid network congestion that consequently causes a decrease in performance.

Keywords: *Computing, Grid Computing, MGRSS, Multi-Grid, Resource Broker*

1. INTRODUCTION

Grid technology plays a major role in managing large-scale problems, by integrating distributed resources to provide users with a supercomputer with a capacity for data sharing and computation (Tang & Zhang, 2005; Ian & Carl, 1997; Ian & Carl, 1999). Participating sites may be physically distributed, heterogeneous, and governed by different administrative domains. Numerous related studies and projects have proposed solutions for large-scale scientific problems, such as earthquake simulation, atmosphere and ocean simulation, high energy and nuclear physics, bioinformatics, and medical image processing. A number of other proposed grid projects are Globus, Condor, LEGION, Grid PP, EGEE, P-Grid, DutchGrid, ESnet, and Grid Bus (http://www.globus.org/; http://www.cs.wisc.edu/condor/; http://www.cs.virginia.edu/~legion/; http://www.gridpp.ac.uk/; EGEE, n.d.; http://www.p-grid.org/; http://www.dutchgrid.nl/; http://www.es.net/; http://www.gridbus.org/).

DOI: 10.4018/ijghpc.2014010103

Though grid computing comprises a conceptual framework, the Globus Toolkit® (GT) enables grid computing. The GT, an open source project developed by the Globus Alliance® for building grids, not only provides users with an implementation of the necessary services of a middleware to build grid infrastructures, but is also capable of implementing immense applications on grid infrastructures. However, the GT lacks some features, such as a queuing system, a friendly interface, and a proper resource brokering manager to process, accept, and reject jobs or workloads submitted by users. In addition, a monitoring mechanism adept at monitoring the status of jobs is also expected.

In a grid environment, applications share grid resources to improve performance. The target function is usually dependent on several parameters, for example, scheduling strategies, machine configurations and topology, workloads in a grid, the degree of data replication, and others. This work examines how these parameters may affect performance, choosing an application's overall response time as an object function and dynamically scheduling independent tasks. The study defines the job, scheduler, and performance model of a grid site, and conducts experiments on the Tiger Grid platform. The study uses the Ganglia (Massie, Chun, & Culler, 2004; Sacerdoti, Katz, Massie, & Culler, 2003; http://ganglia.sourceforge.net/) and NWS (Wolski, Spring, & Hayes, 1999; http://nws.cs.ucsb.edu/ewiki/) tools to monitor resource status and network-related information (Yang, Shih, & Chen, 2006), respectively. Understanding the influence of each parameter is not only crucial for an application to achieve a good performance, but also helps develop effective schedule heuristics and design high-quality grids.

In Taiwan, a number of research institutes are devoted to building grid platforms for academic research, such as Tiger Grid (http://gamma2.hpc.csie.thu.edu.tw/ganglia/) and Medical Grid (http://eta1.hpc.csie.thu.edu.tw/ganglia/), which integrate available computing resources in several universities and high schools. However, various virtual organizations do not possess adequate computing resources for restriction on an organizational scale. For this reason, this study constructs a multi-grid resource broker (http://gamma2.hpc.csie.thu.edu.tw/) to integrate the single grid environment into a multiple grid environment, for more effective use of resources.

This paper proposes a multi-grid architecture to resolve the abovementioned problem. When a new grid joins the multi-grid environment via the proposed system, the user can utilize the entire multi-grid system to receive more computing resources. This approach is a cost-effective method for virtual organizations to possess more computing resources.

We also provide a user-friendly web portal that integrates with the resource broker (Krauter, Buyya, & Maheswaran, 2002; Chung & Chang, 2009; Yang, Shih, & Li, 2005), the cross-grid service, the monitoring service, and the multi-grid resource selection strategy, called the "Multi-Grid Resource Broker." The main function of the resource broker is to match the available resources to the user's requirements. The resource broker helps users select suitable resources according to user preferences and job characteristics. In the architecture, we chose the Globus Toolkit to be the grid middleware. Though Globus provides a monitoring tool called MDS, it is incapable of providing a rich set of all the requisite information. Thus, the study required another monitoring tool, "Ganglia", (Massie, Chun, & Culler, 2004; Sacerdoti, Katz, Massie, & Culler, 2003; http://ganglia.sourceforge.net/) for the multi-grid system. Ganglia is a scalable open-source distributed system for monitoring the status of nodes in wide-area cluster-based systems.

The study also proposes a multi-grid resource selection strategy called "MGRSS". The strategy helps users select machines according to their performance, to shorten the execution time of programs. Furthermore, the study required adjusting each grid, thereby providing the user with a quota of resources. Experimental results show that MGRSS performs better than other strategies.

Related Content

Recommender Systems in Digital Libraries Using Artificial Intelligence and Machine Learning: A Proposal to Create Automated Links Between Different Articles Dealing With Similar Topics

Namik Delilovic (2021). *Handbook of Research on Methodologies and Applications of Supercomputing (pp. 162-178).*

www.irma-international.org/chapter/recommender-systems-in-digital-libraries-using-artificial-intelligence-and-machine-learning/273402