Chapter IV
Effective Resource Allocation and Job Scheduling Mechanisms for Load Sharing in a Computational Grid

Kuo-Chan Huang
National Taichung University, Taiwan

Po-Chi Shih
National Tsing Hua University, Taiwan

Yeh-Ching Chung
National Tsing Hua University, Taiwan

ABSTRACT

Most current grid environments are established through collaboration among a group of participating sites which volunteer to provide free computing resources. Therefore, feasible load sharing policies that benefit all sites are an important incentive for attracting computing sites to join and stay in a grid environment. Moreover, a grid environment is usually heterogeneous in nature at least for different computing speeds at different participating sites. This chapter explores the feasibility and effectiveness of load sharing activities in a heterogeneous computational grid. Several issues are discussed including site selection policies as well as feasible load sharing mechanisms. Promising policies are evaluated in a series of simulations based on workloads derived from real traces. The results show that grid computing is capable of significantly improving the overall system performance in terms of average turnaround time for user jobs.
INTRODUCTION

Grid computing (Foster and Kesselman, 1999) has recently become a promising trend in high performance computing. Many universities, research institutes, and commercial companies have been devoted to the development of related technologies and applications (Globus Alliance, 2007; Open Grid Forum, 2007; Platform, 2007; IBM Grid Computing, 2007; and Sun Grid Engine, 2007). Among various grid architectures and applications is the computational grid which aims to integrate computing resources located at different places and managed by different authorities to achieve load sharing and better resource utilization.

This article explores computing scenarios related to high throughput computing and on-demand computing discussed in (Berman et al., 2003), which are concerned with improving overall system performance on a computational grid through appropriate workload management approaches. A computational grid has to provide strong incentive for participating sites to join and stay in it. Participating sites are concerned with the performance improvement brought by the computational grid for the jobs of their own local user communities. Feasible and effective load sharing is key to fulfilling such a concern.

BACKGROUND

Without grid computing users can only run jobs on their local site. The owners or administrators of different sites are interested in the potential benefit of participating in a computational grid and whether such participation will result in better service for their local users by improving the job turnaround time. Therefore, it is important to ensure that grid computing can bring performance improvement and that the improvement is achieved in the sense that all participating sites benefit from the collaboration. In the other words, no participating sites’ average turnaround time of their local jobs would increase after joining the computational grid.

Heterogeneity is another important issue in a computational grid. Many previous work (Bucur and Epema, 2003; Ernemann et al., 2002; Zhang et al., 2006) have shown significant performance improvement for multi-site homogeneous grid environments. However, in the real world a grid usually consists of heterogeneous sites that differ in configuration and computing speed. Heterogeneity puts a challenge on designing effective load sharing methods. Methods developed for homogeneous grids have to be improved or even redesigned to make them effective in a heterogeneous environment. This article addresses the potential benefit of sharing jobs among different sites in a speed-heterogeneous computational grid environment. Related issues are discussed, including job scheduling for feasible load sharing and site selection for processor allocation.

Job scheduling for parallel computers has been a subject of research for a long time. As for grid computing, previous work discussed several strategies for a grid scheduler. One approach is the modification of traditional list scheduling strategies for usage on a grid (Hamscher et al., 2000; Ernemann et al., 2002). Some economic based methods are also being discussed (Buyya et al., 2002; 2003; Zhu et al., 2005; Ernemann et al., 2002). In this article we explore non-economic scheduling and allocation policies for a speed-heterogeneous grid environment.

England and Weissman in (England and Weissman, 2004) analyzed the costs and benefits of load sharing of parallel jobs in the computational grid. Experiments were performed for both homogeneous and heterogeneous grids. However, in their work simulations of a heterogeneous grid only captured the differences in the numbers of processing nodes and workload characteristics. The computing speeds of nodes on different sites are assumed to be identical. In this article we deal with load sharing issues regarding heterogeneous
Related Content

An Architecture to Implement the Internet-of-Things using the Prometheus Methodology

Rhotrix Multiplication on Two-Dimensional Process Grid Topologies
[www.irma-international.org/article/rhotrix-multiplication-two-dimensional-process/62995/](http://www.irma-international.org/article/rhotrix-multiplication-two-dimensional-process/62995/)

Application Performance on the Tri-Lab Linux Capacity Cluster -TLCC
[www.irma-international.org/chapter/application-performance-tri-lab-linux/64446/](http://www.irma-international.org/chapter/application-performance-tri-lab-linux/64446/)

Resource Management in Real Time Distributed System with Security Constraints: A Review
[www.irma-international.org/chapter/resource-management-real-time-distributed/72256/](http://www.irma-international.org/chapter/resource-management-real-time-distributed/72256/)

Security in Transnational Interoperable PPDR Communications: Threats, Requirements and Architecture Solution
[www.irma-international.org/article/security-in-transnational-interoperable-ppdr-communications/168576/](http://www.irma-international.org/article/security-in-transnational-interoperable-ppdr-communications/168576/)