International Journal of Innovative
Computing, Information and Control
Volume **6**, Number **12**, December **2010**

 ISSN 1349-4198
pp. 5451–5462

PERFORMANCE ANALYSIS AND DISCUSSION ON A HEURISTIC APPROACH FOR SCHEDULING MULTIPROCESSOR TASKS IN A GRID COMPUTING ENVIRONMENT

Jiann-Fu Lin

Department of Management Information System
Takming University of Science and Technology
No. 56, Sec. 1, HuanShan Rd., NeiHu 11451, Taiwan
alfu@takming.edu.tw

Received August 2009; revised January 2010

Abstract. *The problem considered in this paper is motivated by recent studies in scheduling tasks in a grid computing environment. A grid infrastructure is a virtual organization that integrates a large number of distributed resources and high performance computing capabilities into a super service, which can be viewed as an agreement to share resources among independent organizations. Users may dispatch their tasks to the grid computing environment and use remote computing resources instead of computing locally. Hence, the issues of resource management and tasks scheduling are essential in a grid computing environment. As the problem of nonpreemptively scheduling independent multiprocessor tasks in a grid computing environment is NP-hard, we propose a heuristic algorithm for such a problem. We evaluate the performance of the heuristic algorithm not only by mathematical analysis, but also implementing the heuristic algorithm and exploring phenomena from the experimental results. The derived performance bound of the heuristic algorithm falls between* $(2-1/P)$ *and* $(3-2/P)$*, where* P *is the total number of processors in a grid computing environment. To explain the performance analysis of the heuristic algorithm, we implement the heuristic algorithm and discuss the observed phenomena from the experimental results.*

Keywords: Scheduling, Multiprocessor task, Performance bound, Grid computing environment

1. **Introduction.** A grid computing environment [1,8] is an Internet-based infrastructure that connects geographically distributed and heterogeneous resources into a super service, which allows the use of computing systems belonging to different organizations as a single system. Thus, a grid computing environment can be viewed as an agreement to share resources among independent machines, which could support the execution of many scientific applications [1,8,17,24,28]. In a grid computing environment, users could submit their tasks from any machine and a scheduler assigns the tasks to the machines in the environment. Instead of processing tasks locally, the scheduler might dispatch tasks to remote machines [11,12]. To achieve the potential of a grid computing environment, an efficient scheduling framework within a grid computing environment is necessary. With the increasing importance of scheduling tasks in a grid computing environment, several studies [4,11-16,18-23,25,26,29] have discussed the problem of scheduling tasks in such an environment and evaluated their performance either by experimental results or by mathematical analysis.

In 2004, Martino and Mililotti [19] developed a simulation grid computing environment to evaluate the usefulness of genetic algorithms for scheduling independent tasks in several distributed, parallel machines. In contrast with their previous work on up to 24 tasks