Chapter 111
Solving Job Scheduling Problem in Computational Grid Systems Using a Hybrid Algorithm

Tarun Kumar Ghosh
Haldia Institute of Technology, India

Sanjoy Das
Kalyani University, India

ABSTRACT

Grid computing is a high performance distributed computing system that consists of different types of resources such as computing, storage, and communication. The main function of the job scheduling problem is to schedule the resource-intensive user jobs to available grid resources efficiently to achieve high system throughput and to satisfy user requirements. The job scheduling problem has become more challenging with the ever-increasing size of grid systems. The optimal job scheduling is an NP-complete problem which can easily be solved by using meta-heuristic techniques. This chapter presents a hybrid algorithm for job scheduling using genetic algorithm (GA) and cuckoo search algorithm (CSA) for efficiently allocating jobs to resources in a grid system so that makespan, flowtime, and job failure rate are minimized. This proposed algorithm combines the advantages of both GA and CSA. The results have been compared with standard GA, CSA, and ant colony optimization (ACO) to show the importance of the proposed algorithm.

1. INTRODUCTION

A computational Grid aims to aggregate the power of heterogeneous, geographically distributed, multiple-domain-spanning computational resources to provide high performance or high-throughput computing. It also provides dependable, consistent, pervasive, and inexpensive access to computational resources existing on the network. Job scheduling is one of the major challenges in computational Grids to efficiently exploit the capabilities of dynamic, autonomous, heterogeneous and distributed resources for

DOI: 10.4018/978-1-7998-5339-8.ch111

execution of different types of jobs. Also, the size of a Grid has been ever increasing. Therefore, the job scheduling in such complex systems is a challenging problem.

The Grid scheduler, often known as Grid resource broker, acts in three phases: resource discovery phase, resource allocation phase and job execution phase. The resource discovery phase involves identifying the available resources from the resource pool, whereas the resource allocation phase involves selection of suitable resources and allocating the selected resources to the jobs. The third phase is executing the jobs at resource locations. The Grid scheduler searches the fittest resource for a job so that certain criteria (like minimization of makespan or execution time, best utilization of the resources, and maximization of user satisfaction) are met.

In this chapter, a hybrid algorithm is proposed which combines the advantages of Genetic Algorithm (GA) and Cuckoo Search Algorithm (CSA) to minimize three key performance issues, viz. makespan, flowtime and job failure rate, of a computational Grid system. Genetic Algorithm (GA) is one of the widely used evolutionary heuristic algorithms for the constrained optimization problems, but the disadvantage of the algorithm is that it can easily be trapped in local minima. In order to avoid such local minima problem, the Cuckoo Search Algorithm (CSA) can be used to perform the local search more efficiently.

This chapter is organized as follows. Section 2 briefly outlines the relevant past works done on job scheduling in computational Grid environment. In Section 3, the framework of Grid job scheduling problem has been defined. Sections 4 - 6 highlight GA, CSA and ACO methods. Section 7 describes our proposed hybrid technique for scheduling jobs in computational Grid systems. Section 8 exhibits the results obtained in this study. Finally, Section 9 concludes the chapter.

2. RELATED WORKS

Due to various complex characteristics of resources and jobs, the job scheduling in Grid is a NP-complete problem. Meta-heuristic methods have proven to be efficient in solving such problems. Various meta-heuristic algorithms have been designed to schedule the jobs in computational Grid (Thilagavathi et al., 2012). These sorts of approaches make realistic assumptions based on a priori knowledge of the concerning environment and of the system load characteristics. The most commonly used meta-heuristic algorithms are Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Simulated Annealing (SA), Ant Colony Optimization (ACO) and Cuckoo Search Algorithm (CSA). In general, meta-heuristic approaches manage to obtain much better performance, but take a longer execution time (Bianco et al., 2015).

The Genetic algorithm (GA) is a meta-heuristic algorithm that imitates the principle of genetic process in living organisms. GA mimics the evolutionary process by applying selection, crossover, and mutation to generate solution from the search space. The GA is a very popular algorithm to solve various types of combinatorial optimization problems. GAs for the Grid scheduling problems have been studied by Abraham et al. (2000); Kolodziej et al. (2012); Braun et al. (2001); Zomaya and The (2001); Di Martino and Mililotti (2004); Moghaddam et al. (2012); Page and Naughton (2005); Gao et al. (2005); Xhafa et al. (2008); Aggarwal et al. (2005). Prakash and Vidyarthi (2015) have proposed a new technique to maximize the availability of resources for job scheduling in computational Grid using GA. Enhanced Genetic-based scheduling for Grid computing is proposed in (Kolodziej and Xhafa, 2011).

The PSO algorithm is a population-based optimization technique that tries to find the optimal solution using a population of particles. Each position of a particle in the search space corresponds to a potential