



Integration of task abortion and security requirements in GA-based meta-heuristics for independent batch grid scheduling

Joanna Kołodziej^{a,*}, Fatos Xhafa^b

^a University of Bielsko-Biala, Department of Mathematics and Computer Science, ul. Willowa 2, 43-309 Bielsko-Biala, Poland

^b Department of Languages & Informatic Systems, Technical University of Catalonia, Campus Nord, Ed. Omega C/Jordi Girona 1-3, 08034 Barcelona, Spain

ARTICLE INFO

Keywords:

Scheduling
Computational grid
Genetic algorithm
Security
Resource reliability
Grid simulation

ABSTRACT

Scheduling and resource allocation in large scale distributed environments, such as Computational Grids (CGs), arise new requirements and challenges not considered in traditional distributed computing environments. Among these new requirements, task abortion and security become needful criteria for Grid schedulers. The former arises due to the dynamics of the Grid systems, in which resources are expected to enter and leave the system in an unpredictable way. The latter requirement appears crucial in Grid systems mainly due to a multi-domain nature of CGs. The main aim of this paper is to develop a scheduling model that enables the aggregation of task abortion and security requirements as additional, together with makespan and flowtime, scheduling criteria into a cumulative objective function. We demonstrate the high effectiveness of genetic-based schedulers in finding near-optimal solutions for multi-objective scheduling problem, where all criteria (objectives) are simultaneously optimized. The proposed meta-heuristics are experimentally evaluated in static and dynamic Grid scenarios by using a Grid simulator. The obtained results show the fast reduction of the values of basic scheduler performance metrics, especially in the dynamic case, that confirms the usefulness of the proposed approach in real-life scenarios.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Computational Grid (CG) is usually presented as an alternative to supercomputing for solving large-scale optimization problems within a reasonable time. It utilizes a large amount of computational resources virtually joined through Internet into a large global network composed of many local clusters. The main problem in using CG infrastructure is an efficient mapping of tasks and applications submitted by independent users to Grid resources under assumption that all Grid entities may belong to different administrative domains. Indeed, matching the computational needs of applications within resource availability in the system is crucial to achieve the efficiency and scalability of the mapping methodologies under conditions of heterogeneity, large scale and dynamics of the CGs.

Despite many traditional scheduling approaches to super-computing, cluster computing and LANs, the design of efficient schedulers in Grid environment remains still a challenging task. In fact, scheduling problem in Grid systems may be interpreted as a family of problems. Depending on the restrictions imposed by the application needs, the complexity of the problem can be defined by the number of various objectives (single vs. multi-objective), the type of the environment (static vs. dynamic), the processing mode (immediate vs. batch), task interrelations (independency vs. dependency), etc. In this paper we consider the bi-objective scheduling problem in the Computational Grid, referred to as the *Independent Job*

* Corresponding author. Tel.: +48 33 8279233; fax: +48 33 8279233.

E-mail addresses: jkolodziej@ath.bielsko.pl (J. Kołodziej), fatos@lsi.upc.edu (F. Xhafa).

Scheduling, in which two main metrics of the scheduling effectiveness, namely makespan and flowtime, are simultaneously optimized. We assume that tasks are processed in a batch mode and there are no dependencies among tasks so that they can be independently computed on Grid resources.

Independent Batch Job Scheduling is one of the most useful versions of scheduling problems in Grid systems. There are many realistic scenarios in which the need of independent job scheduling arises. Firstly, this problem is suitable to Grid systems because of the nature of Grid users, who submit jobs or monolithic applications in an independent manner to the system. Secondly, Grid systems are most useful for massive parallel processing, in which large amounts of data are processed independently. In such a case, a batch mode is the appropriate approach for mapping the tasks onto Grid resources, because of periodical runs of Grid-enabled applications and a transfer, replication and processing of a large amount of data. In the batch mode tasks and applications are sampled into *batches* and scheduled as a group. Real life examples of batch scheduling include processing of large log data files of online systems (e.g. banking systems, virtual campuses, health systems, etc.), processing of large data sets from scientific experimental simulations (e.g. High Energy Physics, Parameter sweep applications, etc.), data mining in bio-informatics applications, etc. Even under the independent nature of jobs and the batch processing, the problem is computationally hard to solve. Moreover, it becomes more challenging due to the high degree of heterogeneity of resources, the large scale and the dynamics of Grid systems.

Recently a great interest of researchers in Grid Computing domain has been focused on the secure scheduling, which aims to achieve an efficient assignment of tasks to the reasonable trustful machines. Security in Grid systems usually deals with higher system level comprising the authentication and authorization mechanisms. While these standard mechanisms are important, they may result insufficient to ensure adequate security levels. However, there are several scenarios and cases in which the security requirements should be addressed also at lower levels of the Grid system, at scheduling and resource allocations levels of the middleware stack. Firstly, Grid systems are dynamic in nature. Therefore security requirements should be checked at system run time because some machines may be available after the authorization/authentication is done. Secondly, in scheduling some meta-task or bags-of-tasks applications could span a certain number of other smaller tasks, whose security requirements should be checked together with their availability for planning to the resources. Finally, it is also commonplace in scheduling that some tasks produce data which may be accessed by other tasks, as in parameter sweep or Monte-Carlo simulation applications. Again, binding the security requirements should be done at scheduling time. In current approaches, security and resource reliability are addressed separately, while integrating desired security levels into scheduling is very important for Grid-enabled applications, mainly because of the complex nature of Grid systems.

We propose in this paper a hierarchical model of Grid, which is aware of trust and task abortion. We adapt a general concept of the Meta-broker, who is in our approach responsible for checking the security condition and the resource availability. The main contribution of this paper is developing of a scheduling model that enables the integration of task abortion and security requirements with classical scheduling performance measures. It should be noted however that the aim is not to wave authorization, authentication and other system security levels, rather, to provide an integrated approach for security requirements to be checked at scheduling phase.

Because of the complexity of the problem, heuristic and meta-heuristic approaches are the most feasible methods of scheduling in Grids due to their ability to deliver high quality solutions in reasonable computing time. We define eight variants of genetic-based batch schedulers, which can be easily implemented in a dynamic Grid environment. Firstly, we modified the Expected Time to Compute (ETC) model [1] of the independent task scheduling by integration of the scheduling security and resource reliability requirements defined by the users as additional criteria. We specified the expected times of the executions of the tasks on the machines available in the system using the machine failure and task abortion probabilities. The scheduling problem is defined as a multi-objective global minimization task with makespan and flowtime as the main scheduler's performance measures.

We analyze empirically the performance of the GA-based schedulers in the dynamic Grid environment by using a Grid simulator [2]. We considered two scenarios, where the security and resource reliability requirements are respected (secure scenario) and ignored (risky scenario). We also assumed that the number of tasks and hosts can be kept constant (static mode) or the tasks and machines can dynamically change their availability (dynamic mode). In both cases, the Grid scenarios ranged from small to very large size instances in order to see the scalability of GA-based schedulers.

The remainder of this paper is structured as follows. We define the Independent Job Scheduling problem in Section 3. The specification of the schedulers general framework and a short description of genetic operators are presented in Section 4. Section 5 presents a brief characteristics of the secure variant of HyperSim-G simulator and the experimental evaluation of the proposed genetic schedulers. The paper ends in Section 6 with some final remarks.

2. Related work

2.1. Population-based metaheuristics in Grid scheduling

The stochastic and heuristic approaches to Grid scheduling are widely addressed in the research. Monte Carlo methods, Simulated Annealing (SA), Tabu Search (TS), Genetic Algorithms (GA), among others, are presented as effective schedulers in a complex large-scale Grid environment. GAs are well known for their robustness and have been applied successfully to solve scheduling problems in a variety of fields. Zomaya and Teh [3] used GAs in dynamic load balancing problems. Braun et al. [4] compare the efficiency of a simple GA-based scheduler and Min–Min, Min–Max, Minimum Completion Time (MTC)

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات