Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 98 (2016) 199 – 204

Procedia Computer Science

The 7th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2016)

# A Budget-Constrained Time and Reliability Optimization BAT Algorithm for Scheduling Workflow Applications in Clouds

Navneet Kaur[a], Sarbjeet Singh[b,*]

[a]*Computer Engineering and Technology, GNDU, Amritsar 143005, India*

[b]*Computer Science and Engineering, UIET, Panjab University, Chandigarh 160014, India*

**Abstract**

Effective scheduling is one of the key concerns while executing workflows in the cloud environment. Workflow scheduling in clouds refers to the mapping of workflow tasks to the cloud resources to optimize some objective function. In this paper, we apply a recently developed meta-heuristic method called the BAT algorithm to solve the multi-objective problem of workflow scheduling in clouds that minimizes the execution time and maximizes the reliability by keeping the budget within user specified limit. Comparison of the results is made with basic, randomized, evolutionary algorithm (BREA) that uses greedy approach to allocate resources to the workflow tasks on the basis of low cost, high reliability and improved execution time machines. It is clear from the experimental results that the BAT algorithm performs better than the basic randomized evolutionary algorithm.


Peer-review under responsibility of the Program Chairs

*Keywords:* Workflow scheduling; BAT algorithm; Cloud computing; Budget-Constrained time and reliability optimization.

## 1. Introduction

Cloud computing is emerging as a popular and widely used platform in almost every field. Ambrust et al.[1] define cloud computing as applications which are delivered over the Internet in the form of services, as well as the

* Sarbjeet Singh. Tel.: +91-9815951674; fax: +91-172-2547986.
*E-mail address:* sarbjeet@pu.ac.in


Peer-review under responsibility of the Program Chairs
doi:10.1016/j.procs.2016.09.032

hardware and software systems in the datacentre that actually provide these services. Buyya et al.[2] expresses cloud computing on the basis of its utility-oriented aspect. Cloud computing also offers several features like pay-per-use, broad network access, virtualization, rapid elasticity and on-demand access. These features make cloud environment highly suitable for handling the scientific workflow applications.

Workflow is a series of related tasks connected by control-flow and data-flow dependencies. To run workflows on cloud, high-level approach is to make a virtual cluster, formed of cloud resources, and then schedule the tasks of the workflow on that cluster. Scheduling in clouds is an NP-Hard problem. Many scheduling methods which include heuristic as well as meta-heuristic techniques have been applied for workflow scheduling in clouds and more are constantly being experimented and researched, still a lot of work needs to be done for making scheduling in clouds more effective. The quality of the output also depends on whether the tasks are scheduled on reliable virtual machines or not because virtual machines having less failure rate are expected to successfully execute tasks allocated to them. However, scheduling a task on more reliable machine incurs more cost. Similarly, time and cost are conflicting parameters that also needs to be balanced and optimized. So, in order to get optimal solution, the trade off among these objectives must be effectively managed.

Unlike most of the scheduling algorithms that assign cloud resources to the input tasks based on single objective function, usually cost or execution time, in this paper, we deal with multi-objective scheduling problem where we link three objectives namely cost, execution time and reliability into a single objective function. The aim of our paper is to experiment a recently developed meta-heuristic approach called the BAT algorithm for the purpose of workflow scheduling in clouds with the objective of maximizing reliability and minimizing execution time ensuring that the cost does not exceed the user specified budget.

## 2. Related Work

This section describes the related work done in the area of multi-objective workflow scheduling problem using meta-heuristic approaches in clouds. Fard et al.[3] proposed heuristic algorithm and a general framework for solving multi-objective problem of workflows. A double strategy approach is used, where for solutions, which are dominant, distance to the constrained vector is maximized and for non- dominant solutions, it is minimized. A case study comprising of four objectives including execution time, cost, reliability and energy consumption is presented. The algorithm is compared with two bi-criteria algorithms and delivered superior results. Another scheduling algorithm based on PSO technique was proposed by Rodriguez et al.[4]. It aims at minimizing overall cost by taking into account the deadline constraint. Comparison of the results is made with the SCS and the IC-PCP algorithms on different scientific workflows of varying sizes. The authors also suggested that in future, the strategies like GA can also be used and resource model can also be extended by considering data transfer costs between different datacentres. Wang et al.[5] proposed an algorithm named Look Ahead Genetic algorithm (LAGA) and RD reputation to calculate the reliability of a resource. Reliability and execution time are the two objectives considered for the optimization. Results are compared with the GA and it shows that the RD reputation effectively improves the reliability of the application and the LAGA provides better quality solutions. A new meta-heuristic technique inspired from cat-swarm behavior was proposed by Bilgaiyan et al.[6]. The technique is based on multi-objective scheduling approach for workflows. It targets at optimizing execution time, cost and CPU idle-time. Performance of the algorithm is evaluated by comparing it with the multi-objective PSO technique. Tan et al.[7] proposed a trust based multi-objective scheduling model which is explained with the help of a case study. A future direction is also given which says that the presented model can also be extended for rescheduling of tasks in response to the changes in the cloud environment. This paper considers trust value, cost and execution time as the objectives. Another multi-objective algorithm named RD-PSO algorithm was proposed by Wu et al.[8]. It considers both execution time and cost for optimization. For evaluating cost, the data transmission costs and the computation costs are taken into account. Different sets of workflows are used for experimentation and comparisons are made with the standard PSO and the BRS algorithm. Results show that the proposed algorithm outperforms both the algorithms. A score based algorithm which takes into account the reliability of the virtual machines was designed by Singh and Singh et al.[9] which considers deadline as a constraint. It aims at minimizing cost. Score value is calculated for the virtual machines and according to this the tasks are allocated to the resources. The results are compared with the same algorithm where the score value is not considered. Cost and time optimization, PSO based technique was proposed by Yassa et al.[10]