International Journal of Engineering & Technology, 7 (4.6) (2018) 174-176

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Multi-Objective Optimization for scientific workflow task scheduling in IaaS Cloud

Arunkumar Panneerselvam[1]*, Bhuvaneswari Subbaraman[2]

[1]*Research Scholar, Department of Computer Science, Pondicherry University, Karaikal campus, India*
[2]*Registrar, Central University of Tamilnadu, Thiruvarur, India*
**Corresponding author E-mail:arunkumar.pselvam@gmail.com*

Abstract

The use of scientific applications on cloud networks increases day by day generating volumes of data and consuming large computational power. These scientific applications find its importance in the field of astronomy, geology, genetics and bio-technology etc. Complex and mission critical scientific applications can be modeled as scientific workflows and can be executed in cloud. The tasks of the scientific applications are generally data intensive and compute intensive. Traditional computer networks are not suitable for handling scientific applications and hence ubiquitous distributed networks like cloud are prominent in hosting scientific applications. The cloud hosted scientific applications and the cloud network need to satisfy many objectives to the interest of its users. This paper explores the multi-objective optimization applications in scientific workflow task scheduling in IaaS cloud and the related algorithms employed.

Keywords: *Cloud; Scientific workflows; Multi-Objective; Resource Provisioning; Task Scheduling.*

1. Introduction

This computer era is drenched with data from business, scientific and personal applications. Complex and mission critical scientific applications requires large and cost effective computing infrastructure to host and execute. Many Scientific applications can be modeled as scientific workflows and can be executed in the distributed environment like cloud. The Cloud forms a promising platform for scientific workflow execution because of its scalable and elastic nature. Also cloud networks are offered on rental basis which frees the user from infrastructure and capital cost for executing applications. Modern approaches models scientific applications in terms of scientific workflows which can be submitted to cloud environment for execution. Scientific workflows contain set of tasks for which cloud resources are provisioned to enable effective execution of the workflows submitted. Several workflow management systems (SWMS) are now integrated with commercial cloud providers such as Amazon (AWS) and Microsoft (Azure) etc.

The big challenge for the cloud providers is to provision cloud resources to the scientific workflows satisfying multiple constraints specified by the Service Level Agreement (SLAs). The cloud providers try to maximize their profit by optimizing the virtual resource usage. The cloud users try to execute their applications in a minimal time and cost. The time and cost are inversely proportional in cloud scientific workflow task execution because to execute tasks in less time the cloud users need to rent high computing virtual resources which is costlier. Hence cloud resource scheduling for scientific workflow task execution can be seen as Multi-Objective Problem. This paper explores the Multi-Objective optimization techniques for scientific workflow task scheduling in IaaS cloud. The task scheduling problem is NP-Hard problem and hence the traditional rule based algorithms are not effective in providing an acceptable solution to the users. This paper explores the literature related to Multi-objective optimization in scientific workflow scheduling in cloud in section 2. The section 3 deals with the Multi-objective optimization techniques and applications. Section 4 explains the important algorithms employed in Multi-objective optimization. Experiments and results are discussed in Section 5. The section 6 contains conclusion and future work.

2. Related work

The works related to Multi-objective scientific workflow scheduling in cloud are limited. We have reviewed selected papers relevant to Multi-objective workflow scheduling in IaaS cloud. J.J Durillo and Radu Prodan proposed Multi-objective Heterogeneous Earliest Finish Time (MOHEFT) which gives better trade off solutions for Makespan and financial cost and the results are compared with well known SPEA2* Algorithms [1]. Phyo Thandar, Courtney Powell et. al. proposed Multi-objective and level wise scientific workflow optimization in IaaS Public cloud environment which concentrates on minimization of Makespan, Virtual machine (VM) deployment cost and VM failure [2]. The results are compared with standard SPEA2 and NSGAII algorithms. Miao Zhang. Huiqi Li et. al. proposed an adaptive penalty function based multi-objective evolutionary approach for scheduling workflows taking the objective to minimize total execution cost and degree of imbalance[3]. The performance is compared with MOPSO, NSGAII and SPEA2 the well known multi-objective algorithms. Zhaomeng Zhu, Gongxuan Zhang et. al used Evolutionary Multi-objective workflow scheduling which uses genetic operators to find trade-off solutions for cost and time[4]. The results are compared with SPEA2*, NSPSO, MOHEFT, MODE ε-Fuzzy PSO. Ji Liu, Ester Pacitti et. al. proposed multisite cloud multi-objective scientific workflow scheduling using ActGreedy algorithm. Their work concentrates on improving trade-off solutions for execution time and monetary cost [5]. Heyang Xu, Bo Yang et. al proposed multi-objective workflow optimization for workflow scheduling considering fault recovery. The author uses min-min based time and cost tradeoffs [6]. The results are compared with LOSS and IC-PCP algorithms.

3. Multi-Objective Application

The scientific workflows execution on cloud from the cloud user perspective requires many objectives to be satisfied. The aim of the mulit-objective optimization is to find the trade-off solutions from multiple non-dominated solutions. The non-dominated solutions that are optimal are called as Pareto optimal solutions. The Pareto optimal solution means that any improvement in one objective further will degrade the other solutions. In scientific workflow executions many objectives like minimizing Makespan, minimizing the execution cost, minimizing failure rate, minimizing the tardiness, maximizing the throughput are some of the Multi-objectives which will be the target of the cloud user.

Mathematically the problem can be represented as,

Minimize $y = f(x) = [f_{makespan}(x), f_{cost}(x), f_{frate}(x)]$ (1)

Subject to $c(x) = [c_{makespan}(x), c_{cost}(x)]$

Where $x \in X$, $y \in Y$

As in (1) the cloud user here intends to minimize the Makespan, cost and failure rate of the virtual machines subject to the constraints specified. A Pareto optimal solution here needs to be obtained which is optimized for the three objectives.

The following Fig. 1 represents the pareto front of the popular benchmark problem DTLZ 7 for three objective functions [7] [8].

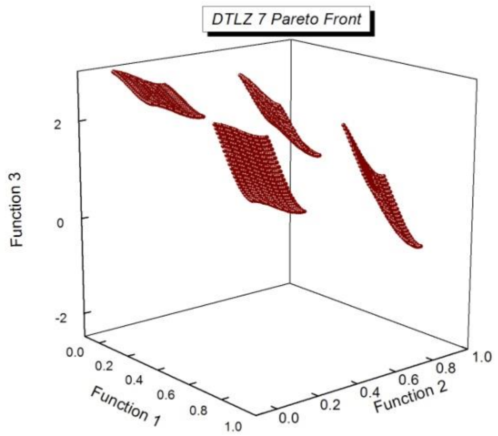

Fig. 1: The Pareto front for DTLZ7 test problem

4. Multi-Objective Optimization Algorithms

As described by Deelman et. al the workflow optimization can be done at composition phase or at instantiation phase or at mapping phase[9]. The most efficient optimization is optimizing the mapping phase of the workflow. The mapping refers to mapping the tasks to the virtual resources available on the cloud. Three popular and well established Multi-Objective optimization algorithms are NSGAII, SPEA2 and MOPSO. All the three algorithms are Evolutionary Meta-Heuristic algorithms. Originally, Non Dominated Sorting Genetic Algorithm (NSGA) was proposed by K. Deb et. al which uses genetic operators such as selection, mutation, cross over and calculates pareto dominance and crowding distance. SPEA2 (Strength Pareto Evolutionary Algorithm) was proposed by E. Zitzler et. al which uses strength pareto and fitness value parameters and is an extension of the genetic algorithm. The Multi-Objective Particle Swarm Optimization (MOPSO) was given by Kennedy et. al which uses local and global best and the fitness values for finding optimal pareto solutions.

5. Experiments

In this section we present our simple experiment results by using MOEA framework tool [9]. The test problem in Fig. 2 is a benchmark function given by E. Zitzler, K. Deb et.al for evaluating Multi-Objective algorithms.

$$\text{Minimize} = \begin{cases} f_1(\boldsymbol{x}) = x_1 \\ f_2(\boldsymbol{x}) = g(\boldsymbol{x})\, h(f_1(\boldsymbol{x}), g(\boldsymbol{x})) \\ g(\boldsymbol{x}) = 1 + \frac{9}{29}\sum_{i=2}^{30} x_i \\ h(f_1(\boldsymbol{x}), g(\boldsymbol{x})) = 1 - \sqrt{\frac{f_1(\boldsymbol{x})}{g(\boldsymbol{x})}} \end{cases}$$

Fig. 2: The ZDT1 test problem

The parameters seed value is set as 10 and maximum NFE is set as 2000 in MOEA framework tool for evaluating the approximation set, elapsed time and hypervolume of the three well known algorithms NSGAII, SPEA2 and SMPSO and results were obtained.

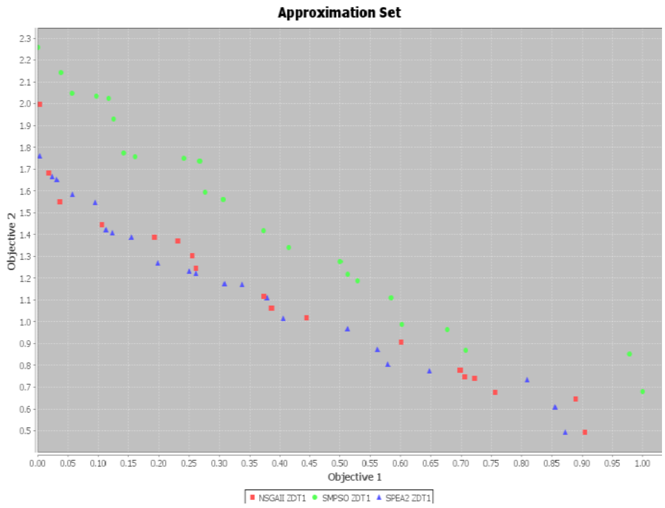

Fig. 3: Approximation set comparison

As in Fig. 3 it is clear that SMPSO gives better approximation set followed by NSGA2 and SPEA2

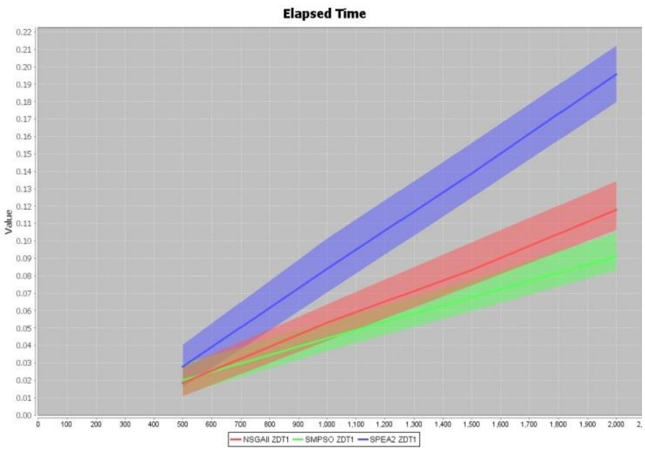

Fig. 4: Elapsed time comparison

As given in Fig. 4 SMPSO is taking less time to execute the ZDT1 problem compared to NGSAII and SPEA2.

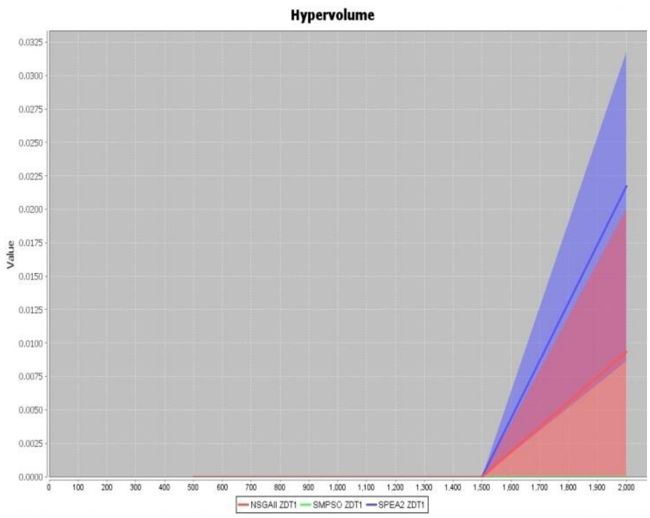

Fig. 5: Hypervolume comparison

The Hypervolume which measures the quality of the non dominated solutions is an important indicator in the Multi-objective optimization. As in Fig. 5 it is evident that SMPSO is having better hypervolume indication followed by NSGAII and SPEA2.

6. Conclusion and Future work

For Multi-Objective scientific workflow optimization in cloud it is better to use the popular Meta-Heuristic algorithms to find optimal Pareto solutions. By our simple experiments we conclude swarm based algorithms give better solutions compared to evolutionary and genetic algorithms. Hybrid algorithms can also be explored much on this research area. However, recently Hyper-Heuritics techniques are becoming popular. Hyper-Heuristics techniques are problem and independent and more general. In our further work we will explore the hyper-heuritics algorithms which employ the strength of the currently available mulit-objective algorithms for scientific workflow scheduling.

Acknowledgement

The corresponding author would like to thank Pondicherry University for providing UGC Non-Net fellowship for carrying out his research work.

References

[1] Juan J. Durillo, Radu Prodan,"Multi-Objective workflow shceduling in Amazon EC2", *Cluster Computing*, Vol.17, No.2, (2014), pp.169-189, available online: https://doi.org/10.1007/s10586-013-0325-0

[2] Phyo Thandar Thant, Courtney Powell, Martin Schlueter, and Masaharu Munetomo, "Multiobjective Level-Wise Scientific Workflow Optimization in IaaS Public Cloud Environment," *Scientific Programming*, Vol. 2017, Article ID 5342727, 17 pages, 2017. https://doi.org/10.1155/2017/5342727.

[3] Miao Zhang, Huiqi Li, Li Liu and Rajkumar Buyya, "An adaptive multi-objective evolutionary algorithm for constrained workflow scheduling in Clouds," *Distributed and Parallel Databases*, Vol. 36 (2017), pp. 339-368.

[4] Zhaomeng Zhu, Gongxuan Zhang, Miqing Li, Xiaohui Liu, "Evolutionary Multi-Objective Workflow Scheduling in Cloud." *IEEE Transactions on Parallel and Distributed Systems* Vol. 27 (2016), pp. 1344-1357.

[5] Ji Liu, Esther Pacitti, Patrick Valduriez, Daniel De Oliveira, Marta Mattoso, "Multi-Objective Scheduling of Scientific Workflows in Multisite Clouds", *Future Generation Computer Systems*, Elsevier, Vol. 2016, 63, pp.76-95, available online: https://doi.org/10.1016/j.future.2016.04.014.

[6] Heyang Xu, Bo Yang, Weiwei Qi and Emmanuel Ahene, "A Multi-objective Optimization Approach to Workflow Scheduling in Clouds Considering Fault Recovery," *KSII Transactions on Internet and Information Systems*,Vol. 10, no. 3, pp. 976-995, 2016. doi: 10.3837/tiis.2016.03.002

[7] K. Deb, L. Thiele, M. Laumanns and E. Zitzler, "Scalable multi-objective optimization test problems," *Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on*, Honolulu, HI, 2002, pp. 825-830.doi: 10.1109/CEC.2002.1007032

[8] Web reference: https://pythonhosted.org/inspyred/reference.html

[9] Web reference: http://www.moeaframework.org/