International Journal of Distributed Systems and Technologies
Volume 9 • Issue 1 • January-March 2018

Performance Aware Planning Algorithms for Cloud Environments

Jyoti Thaman, Department of Computer Science and Engineering, Maharishi Markandeshwar University, Mullana, Ambala, India

Kamal Kumar, University of Petroleum and Energy Studies, Dehradun, India

ABSTRACT

For the last decade, cloud computing has been spreading its application base from the small enterprises to the large, from the domestic user to the professional, from buyers to sellers and from research to implementation. Subscribers submit their jobs or workflows for executions on clouds. Workflow scheduling is a very important aspect in cloud computing and it imitates industrial operations, constraints and dependencies. Several approaches such as Greedy, Heuristic, Meta-heuristic and Hybrid have been tried to reschedule workflows. This article proposes Modified HEFT (MHEFT) and Cluster Based Modified HEFT (C-MHEFT). MHEFT modifies the mapping of ranked tasks to the VMs. C-MHEFT is the cluster based extension of MHEFT. The simulations were performed in WorkflowSim and were compared with existing benchmarks in planning algorithms like HEFT and DHEFT. The proposed schemes will help industries, enterprises to model and sequence the Industrial process which will be faster and efficient.

KEYWORDS

Dynamic Scheduling, HEFT, Performance, Task Types, WorkflowSim

INTRODUCTION

Cloud computing is driving the computing everywhere. A lot of services, applications, infrastructure and managerial tasks have been mapped on to the cloud platform, both Public and Private. Ever increasing impact of cloud computing requires a lot of research. It should focus to improve the performance of clouds. This would prevent drowning of clouds and establish cloud as a long-term reality. However, increasing pace of cloud platforms has exposed several issues related to QoS parameters of cloud platforms.

Scheduling of tasks, resource management, network management, privacy and security etc. should be addressed so that QoS parameters agree with clients and meet to the best possible levels. Scheduling of tasks is one of the most important challenges as client. Many scheduling and planning algorithms have been proposed for addressing this issue. HEFT, DHEFT, Min-Min, Max-Min, Round-Robin, SJF etc. are some well-known proposals. This paper is a proposal of the recently researched a Modified HEFT algorithm (MHEFT) and Cluster Based Modified HEFT (C-MHEFT). This MHEFT algorithm is using a novel task-VM mapping based on principle of reducing the difference between task requirements and VM's expected performance for given set of tasks. As set of tasks are changed, VM's performance gets tuned as per task's characteristics. This adds to the concept of

DOI: 10.4018/IJDST.2018010101

dynamic self-configuration of VM's performance when physical characteristics were fixed for whole length of experimentations. CMHEFT is the extension of MHEFT which modifies the HEFT by using a clustering approach. In this paper, workflows have been considered as representing different task types. This proposal has been organized into several sections. Introduction section introduces the concepts and expectations from the proposal. Related Work presents the existing related work and identifies the gaps. The Preliminary section presents a briefing of preliminaries in the form of HEFT planning policy. The fourth section presents proposed Scheduling Approach. The fifth presents Modified HEFT (MHEFT) planning algorithm and Cluster Based Modified HEFT (C-MHEFT). The sixth section presents the simulation and analysis of the scheme. Finally comes the Conclusion of the research and future work.

RELATED WORK

Recent development and marketing of cloud based solutions have opened up scheduling of task as a novel and interesting research domain. Several remarkable scheduling algorithms including planning algorithms have been proposed for task scheduling in literature. This section reviews few remarkable and citable contributions in research of scheduling of independent and dependents tasks.

Authors Chen and Deelman (2012) have proposed the extension of CloudSim for inclusion of distributed jobs which are called workflows. The management and scheduling aspects has been analyzed in a review of workflows application (Liu et al., 2014). A useful conclusion was drawn from the review that single cloud model is inefficient with respect to increasing scale and complexity of workflows. Another proposal (Zhou & Garg, 2015) evaluated the performance of various existing workflow scheduling schemes for different workflows. This work has endorsed that dynamism may be important parameter to keep in knowledge while evaluating any workflow scheduling proposal. Some study (Thaman & Singh, 2016a) has proposed a Robust HEFT (RHEFT) planning algorithm by considering a novel VM-Task mapping algorithm. Authors Thaman and Singh (2016b) have presented a simulated performance of various dynamic scheduling schemes available in WorkflowSim. Authors (Chopra & Singh, 2013) have proposed hybrid of HEFT with a novel concept of sub-deadline to achieve deadline constraints. Almost all the tasks were scheduled within their deadline. A novel mechanism called Partial Deadline Constraint (PDC) was proposed which not only achieves a better mapping but also exploits the underlying concept of elasticity (Arabnejad & Bubendorfer, 2015). Authors Thaman and Singh (2016c) have proposed a novel taxonomy for classification of task scheduling algorithms based on the goal oriented and constraint oriented approach. This paper presents a highly useful current perspective of scheduling approaches. Another research (Singh et al., 2014) has presented a comparison of various scheduling algorithms on distinct aspects like suitability, feasibility and adaptability to cloud platform. This paper also proposed a hybrid scheduling approach which can bring better returns to cloud service providers and ensure better quality of services. A Genetic Algorithm based load balancing approach was presented (Dasgupta et al., 2013). Authors have discussed and proved that dynamic nature of cloud environment might be the best addressed through adaptive scheduling schemes. The proposal claimed that un-balanced state is inevitable whereas system might have started well in balanced state. Another Genetic Algorithm based heuristic was proposed (Yu & Buyya, 2006) where two objectives, namely budget and deadlines were addressed. Optimizations were achieved by either minimizing execution times while at the same time meeting the budget and deadline constraints posed by the subscriber. Evaluation was done using both balanced and unbalanced workflows and results show this newly proposed approach was better in handling complex workflow structures. Authors (Sheng & Li, 2016) have proposed a new concept named Template Based Genetic Algorithm (TBGA) where users specify QoS constraints. Based on specified values, TBGA computes the maximal number of tasks that could be scheduled on each machine. These computed results are known as template. These templates are used to segregate the tasks into multiple subsets which are to be allocated to one processor each. This allocation was made considering