Accepted Manuscript

An efficient cost optimized scheduling for spot instances in heterogeneous cloud environment

Shridhar G. Domanal, G. Ram Mohana Reddy

PII: S0167-739X(17)30366-7
DOI: https://doi.org/10.1016/j.future.2018.02.003
Reference: FUTURE 3974

To appear in: *Future Generation Computer Systems*

Received date : 8 March 2017
Revised date : 6 November 2017
Accepted date : 3 February 2018

Please cite this article as: S.G. Domanal, G.R.M. Reddy, An efficient cost optimized scheduling for spot instances in heterogeneous cloud environment, *Future Generation Computer Systems* (2018), https://doi.org/10.1016/j.future.2018.02.003

An Efficient Cost Optimized Scheduling for Spot Instances in Heterogeneous Cloud Environment

Shridhar G. Domanal[a], G. Ram Mohana Reddy[a]

[a]*Department of Information Technology, National Institute of Technology Karnataka, Surathkal, Mangalore, 575025, India.*

Abstract

In this paper, we propose a novel efficient and cost optimized scheduling algorithm for a Bag of Tasks (BoT) on Virtual Machines (VMs). Further, in this paper, we use artificial Neural Network to predict the future values of Spot instances and then validate these predicted values with respect to the current (actual) values of Spot instances. On-Demand and Spot are the key instances which are procured by the cloud customers and hence, in this paper, we use these instances for the cost optimization. The key idea of our proposed algorithm is to efficiently utilize the cloud resources (mainly VMs instances, Central Processing Unit (CPU) and Memory) and also to optimize the cost of executing the BoT in the heterogeneous Infrastructure as a Service (IaaS) based cloud environment. Experimental results demonstrate that our proposed scheduling algorithm outperforms state-of-the-art benchmark algorithms (Round Robin, First Come First Serve, Ant Colony Optimization, Genetic Algorithm, etc.) in terms of Quality of Service (QoS) parameters (Reliability, Time and Cost) while executing the BoT in the heterogeneous cloud environment. Since the obtained results are in the form of ordinal, hence we carried out the statistical analysis on both predicted and actual Spot instances using the Spearman's Rho Test.

Keywords: Scheduling Algorithms, Load Balancing, On-Demand and Spot instances, Resource Utilization.

1. Introduction

Efficient VM scheduling and utilization of cloud resources is a challenging problem of distributed computing that comes under NP-Hard/NP-Complete complex class. Cloud computing is the emerging technology in a distributed environment consisting of several data centers, servers, virtual machines, load balancers etc. which are connected intelligently. Further, it deals with many things such as storing and retrieving of documents, sharing of multimedia, lending the related resources on a pay-as-you-go model and much more. Even though there is much advancement in different components of cloud environment with respect to reliability and quick responsiveness, still there is a scope for further improvement in Infrastructure as a Service (IaaS)

Email addresses: shridhar.domanal@gmail.com (Shridhar G. Domanal), profgrmreddy@gmail.com (G. Ram Mohana Reddy)