

Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)

Performance Evaluation of Cloud Services with Profit Optimization

M. Jaiganesh^{a,*}, B. Ramadoss^b, A. Vincent Antony Kumar^a and S. Mercy^c^aDepartment of Information Technology, PSNA College of Engg. and Tech, Dindigul, Tamilnadu, India^bDepartment of Computer Applications, National Institute of Technology, Thiruchirappalli, Tamilnadu, India^cDepartment of Information Science & Engineering, Bangalore Institute of Technology, Bangalore, Karnataka, India

Abstract

Cloud computing is a promising computing paradigm which allows distribution of services from a pool of resources. The services are required by the clients through on-demand via pay and use method. The greatest utilization of resources and maximum profits with scheduling is the main goal of the cloud service providers. The major issue in cloud computing is scheduling of services with improved global throughput and job scheduling. Since, cloud computing is a service based one, the performance evaluation is an important criteria to be dealt with. In this paper, we propose a Priority based Queuing model to evaluate the services leased by the cloud service providers. We consider general service time and response time for arriving requests and the waiting requests are stored in the queue. The services are considered to be SaaS (Software as a Service), PaaS (Platform as a Service) and IaaS (Infrastructure as a Service). We construct this using Queuing model with markovian arrival rate, general service rate, 'm' number of servers, priority queue discipline and a buffer of size 'r'. The advantage of the proposed analytical model is within the time span, the cloud service provider schedules his services to result in maximum profit.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of organizing committee of the Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)

Keywords: Performance evaluation; Cloud computing; Profit optimization; Queuing model.

1. Introduction

Cloud computing is an evolving paradigm to access assortment of data pool via internet using connective devices such as Personal Digital Assistant (PDA), workstation and mobile^{1,2}. It is a utility based computing, which has the capability to deliver services over the internet. It provides on-demand access without the need of any human intervention. Cloud computing provides a basic three key level services called Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). SaaS, which provides the cloud users to access the software applications and hosted services over internet; PaaS, which includes the access to hardware and software computing platforms like operating systems and virtualized servers over internet; and IaaS, which allows the user to access almost all equipments like hardware, storage, servers and networking components. The Cloud Service Provider (CSP) is able to provide the above services in an efficient manner⁹. CSP expects the services leased to provide high profit. The profit provided by the above three cloud services varies greatly. To find the cloud service profit, it is essential to analyze

*Corresponding author. Tel.: +91-9894025367.

E-mail address: ^ajaidevlingam@gmail.com

the performance of cloud computing. Many methods have been delivered to evaluate these performances; especially queuing model is used to evaluate the cloud services in an efficient manner³.

The web applications were modeled as queues and the virtual machines were modeled as service centers. The queue implementation in the above case was done to dynamically create and remove virtual machines and evaluate the scaling up and down of the cloud. There is no VM (Virtual Machine) live migration involved in this model which makes it much simpler than some existing models¹⁰.

Service performance in cloud computing was evaluated using queuing theory. To deliver QoS guaranteed services in cloud computing environment, the relationship among the maximal number of customers, the minimal service resources and the highest level of services is necessary. The above results obtained using queuing theory is useful in the design of this new computing paradigm.

Performance evaluation was done for the recovery on both processing nodes and communication links. Poisson arrivals of users' service requests is considered, and the inter arrival times of service requests follow general probability distribution. The proposed cloud performance evaluation models and methods yielded results which are realistic, and thus are of practical value for related decision-makings in cloud computing.

Performance evaluation of cloud computing center with the arrival of the services in the FCFS manner was done using queuing theory. The mean number of tasks in the system, probability of immediate service, waiting time was calculated to evaluate the services provided to users.

In CSP domain, Queuing model can also be used to achieve profit decisions in cloud services based on the resources needed to execute a particular service. Both the cloud service provider and the clients should be aware of the Quality of Service (QoS) factors in order to evaluate the performance of the cloud farm⁴. The CSP expects a high profit environment in providing the service to the users. But the cloud architecture is not an easy environment for performance evaluation. Since it has

- Large numbers of servers are to support where the conventional queuing model supports only a few numbers of servers.
- The service times are more commonly considered as exponential distribution for convenience in traditional queuing models. The cloud environment needs to be dealt with general service times.
- The loads are not always uniform here because of the dynamic nature of the cloud. The cloud is expected to provide QoS in spite of widely varying loads^{7, 12}.
- Profit yield is another overhead associated with each of the cloud service. QoS should also be evaluated in the perspective of each priority service to attain the profitable environment¹³.

As a solution, we have modeled the cloud as a $M/G/m/m+r$ priority model queuing system. The performance comparison has been done between $M/G/m/m+r$ with First Come First Serve (FCFS) queuing discipline [base], here using FCFS the profit attained is minimum compared with the priority queuing discipline. In priority queuing, we include weighted fair queuing to achieve high profit without affecting the performance. The performance results show that the priority based queuing model is giving comparatively better performance in terms of the profit and QoS compared to FCFS.

The paper is organized as follows. Section 2 gives the existing work in performance evaluation of the cloud and modeling it as $M/G/m/m+r$. Section 3 explains the priority based queuing model in detail. In Section 4, the model is evaluated and the performance comparisons are made. The decisions, findings are summarized and discussions related to future work are done in Section 5.

2. Related Work

Cloud computing has earned a lot of attraction towards it but there is not much work regarding the performance evaluation and profitable cloud environment. The cloud center had been modeled as a classic open network to find the distribution of the response time. But the inter arrival and inter service times were considered as exponential⁵. The cloud center had been modeled as $M/M/m/m+r$ queuing system. The inter arrival times were considered to be exponentially distributed and the model had a buffer of size $m+r$. The response time was split in to waiting, service and execution periods here and they are considered to be independent which seems to be unacceptable⁶.

Download English Version:

<https://daneshyari.com/en/article/487431>

Download Persian Version:

<https://daneshyari.com/article/487431>

[Daneshyari.com](https://daneshyari.com)