Virtual Machine Scheduling using Improved Time Shared Policy in Cloud Computing

Amanjot Kaur and Jyotsna Sengupta

***Abstract*—Cloud Computing has emerged as a popular computing model with increasing number of users and organizations shifting towards cloud based services. To deal with large number of users and tasks efficient scheduling of Virtual Machines has become an important concern in cloud computing environment. Scheduling policies in cloud computing are characterized into two approaches namely Time Shared and Space Shared. The same scheduling policies are applicable at VM level and host level. But the existing scheduling policies does not provide efficient allocation with respect to certain parameters especially response time. In this research work an algorithm is purposed for efficient allocation of processing elements to virtual machines. The proposed algorithm works by effectively utilizing the available RAM in the host according to processing elements. It minimizes the free RAM according to the number of processing elements. The evaluation of these algorithms is done in CloudSim 3.0.3 simulator and simulation results show that proposed algorithm performs better in terms of response time.**

***Index Terms*-- Cloud Computing, CloudSim, Scheduling, Virtual Machines**

I. INTRODUCTION

Computing based on the internet using shared resources is called cloud computing. It enables users to access applications that actually reside at a location other than user's own computer. The word cloud is actually a metaphor for the internet hence it is also known as internet based computing. It also refers to the fact the all the underlying infrastructure is hidden as if obscured by a cloud therefore it's known as cloud computing. It is based on pay as you use model, which means users have to pay according to their usage only hence it is cost effective. We often use it daily for example when we are using many social networking sites, gmail etc. we are actually using cloud computing services. Cloud computing has overcome many issues like latency, transaction control and many more. For example in the past, many of us worried about losing our documents, photos and files if something bad happened to our computers, like a virus or a hardware malfunction. But today all our data would still safely reside on the web, accessible from anywhere in the world because of cloud computing. According to NIST [1] definition of cloud computing it is defined as "Cloud Computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models."

The services provided by Cloud Computing are broadly classified into three service models which are Software as a Service (SAAS), Infrastructure as a Service (IAAS) and Platform as a Service (PAAS). On the basis of deployment clouds are classified as Public Cloud, Private Cloud, Community Cloud and Hybrid Cloud.

Among all the benefits provided by Cloud Computing there are also some issues faced by Cloud computing. Scheduling of Virtual Machines is one such issue. An efficient scheduling algorithm should satisfy user's QoS constraints such as response time, cost, energy efficiency etc.

In this paper two different Virtual Machine Scheduling Policies namely Time Shared and Space Shared are compared and an improved Time Shared Scheduling Algorithm is proposed to reduce the response time of cloudlet (task) execution.

The remainder of the paper is organized as: Section 2 gives a glimpse about related work, Section 3 describes about Scheduling Architecture of Cloud Computing, Section 4 describes about Virtual Machine Scheduling Policies, Section 5 describes about proposed algorithm, Section 6 describes about simulation scenario, Section 7 describes results and discussion and Section 8 finally concludes the paper.

II. RELATED WORK

Virtual machine scheduling refers to the way in which tasks are assigned to VMs. Broadly VM scheduling is classified as QoS aware, energy aware, resource aware, cost aware and work load aware scheduling.[2] There are different algorithms depending upon the scheduling requirements. Regarding the same issue the various research papers that have evaluated VM placement policies are summarized as follows:-

A novel compromised-time-cost (CTC) scheduling algorithm was proposed for instance-intensive cost-constrained cloud workflows. Simulation results

showed that CTC algorithm is efficient in terms of cost. They concluded that in future more cloud workflow scheduling algorithms for various scenarios will be developed. [3]

In this literature [4] analysis is done on scheduling issues in grid, cloud computing environments. By investigating various scheduling issues they concluded that an efficient scheduling algorithm should focus on the QoS requirements such as response time, cost, security, energy and load balancing features. It should also have optimal resource allocation policy. They also concluded that considering all the parameters in a single scheduling framework is not a feasible solution as it increases design complexity.

Authors in [5] have done performance evaluation of Time Shared and Space Shared Scheduling Policies using Cloud Analyst Simulator. The main focus of their research is to prove that the choice of VM Scheduling Policy in Cloud computing model significantly improves performance under resource and service demand variations. They concluded that response time is less if Space Shared as VM Scheduling policy is implemented in comparison to Time Shared VM Scheduling Policy

Authors in [6] have evaluated virtual machine placement policies in geographically distributed datacenter using CloudSim toolkit. In addition to that they have also evaluated Time Shared and Space Shared virtual machine placement policies. They compared these policies with different parameters such as total energy consumption and total delay. However they have not taken response time into consideration. They concluded that in future experiments will be performed to show the impact of different parameters and applications on virtual machine placement decisions.

III. SCHEDULING IN CLOUD COMPUTING

The main objective of scheduling is to increase the utilization of resources and to reduce the processing time of the tasks. The scheduler should assign VMs in such a way such that quality of service is improved. An efficient scheduling algorithm has less response time or delay and must allow more number of tasks to be submitted by the user. In scheduling architecture user submits the set of cloudlets (which are basically the tasks) to the broker, and broker in return assigns these cloudlets to some virtual machines running on the host. Each datacenter has number of hosts and each host can be further virtualized into a number of virtual machines. On these VMs tasks (i.e. cloudlets) are assigned according to the scheduling policy adopted by the broker.

IV. VM SCHEDULING POLICIES

A. Time Shared Policy

In Time Shared scheduling policy the resources are being shared among the cloudlets. Each cloudlet gets the resources for execution for a certain period of time. After that period of time the resources are taken away from that cloudlet and are assigned to another cloudlet. The concept of Round Robin (RR) algorithm is used in this policy.

B. Space Shared Policy

In Space Shared the resources are not shared among the cloudlets. A cloudlet owns the Virtual Machine until it finishes execution. It works on the basis of first come first served; therefore waiting time is high. It works in the same way as that of First Come First Served (FCFS) algorithm.

V. PROPOSED ALGORITHM

The algorithm is purposed for efficient allocation of processing elements to virtual machines. The proposed algorithm works by effectively utilizing the available RAM in the host according to processing elements. It minimizes the free RAM according to the number of processing elements.

Proposed Algorithm

1. The first step is the initialization of common variables. In this step we set the number of users which directly correlates with the broker count.
2. The next step is the creation of data center; this in turn will lead to the creation of host along with their characteristics. The characteristics of host can be size of processing element, bandwidth utilization, RAM etc.
3. After datacenter is created next step is the creation of datacenter broker instance. The creation of broker is very important as it is responsible for communication between datacenters and the submission of cloudlets.
4. Next step is the creation of virtual machine instances. Virtual machines have same characteristics as that of host.
5. Virtual machines then submitted to the broker.
6. After virtual machines are submitted the next step is to specify the cloudlets which are basically the tasks to be submitted. Cloudlets are also created with certain parameters like MIPS requirement, bandwidth requirement etc.
7. Obtain the current utilization of RAM.
8. If free RAM is less than processing element in use then allocate processing element for virtual machine otherwise again go back to previous step.

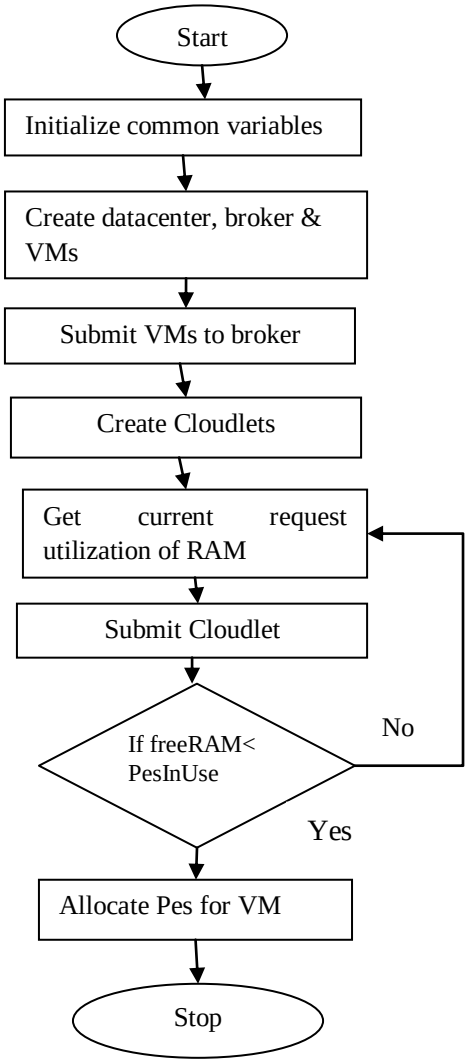

Fig 5.1 Flowchart of Proposed Algorithm

VI. SIMULATION SCENARIO

Table VI. Host Configuration

Parameter	Value
VM Size	10000
VM Memory	512 MB
VM Bandwidth	1000
Standard Architecture	x86
Data Center OS	Linux
Data Center VMM	Xen
Data Center Memory per Machine	2048 MB
Data Center Storage per Machine	1000000
Data Center Available Bandwidth per machine	10000
Data Center VM policy	Time Shared, Space Shared
Server Broker Policy	Time Shared, Space Shared

VII. RESULTS AND DISCUSSION

This section explains the simulation results of Time Shared, Space Shared and Improved Time Shared (proposed) virtual machine scheduling algorithms in cloud computing. These algorithms are compared using different parameters by graphs in MATLAB. The simulation results are obtained on different performance parameters. These parameters include response time, delay, energy consumption and cost.

A. Comparison of Response Time (in ms)

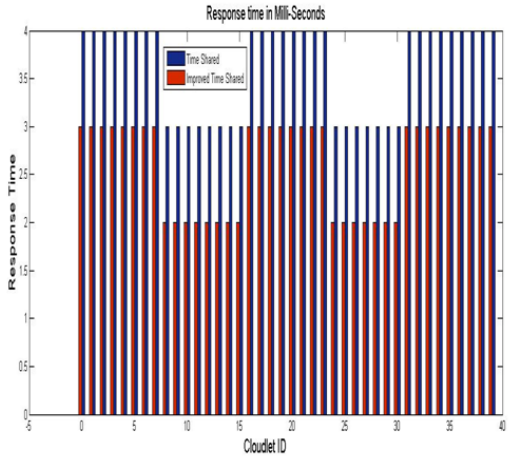

Fig 7.1 Response Time Comparison

Fig. 7.1 illustrates the comparison of response time. Response time is the time taken to complete the execution of a task (cloudlet). As observed from the above figure it is less for proposed algorithm which indicates that it is more efficient in terms of response time than the existing one.

B. Delay Comparison (in ms)

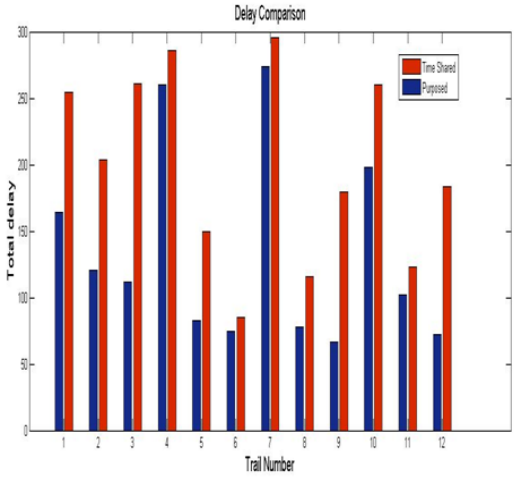

Fig 7.2 Delay Comparison

Fig. 7.2 illustrates the comparison of delay between space shared and proposed virtual machine scheduling algorithm. Simulation results indicate that delay is less in proposed algorithm.

C. Energy Consumption Comparison (in kWh)

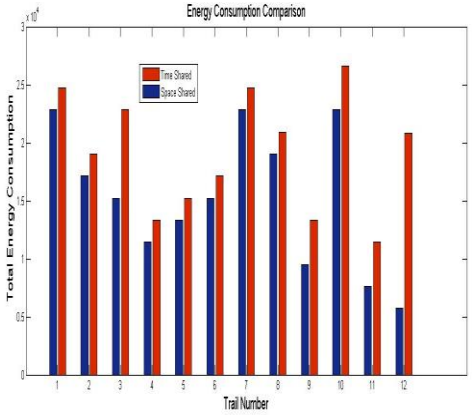

Fig 7.3 Energy Consumption Comparison

Fig. 7.3 illustrates the comparison of total energy consumption between space shared and time shared virtual machine scheduling policies. The energy consumption values are computed in kWh(kilo-watt-perhour). Simulation results indicate that energy consumption is less in space shared approach.

D. Cost Comparison

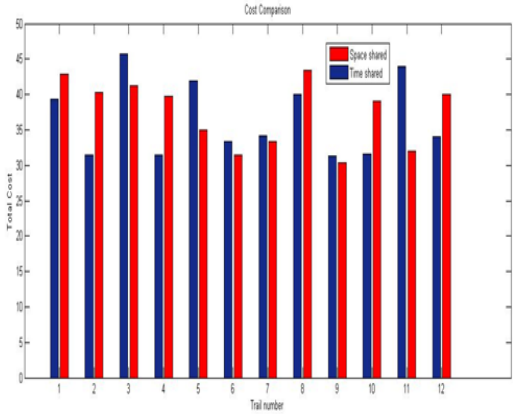

Fig 7.4 Cost Comparison

Fig. 7.4 illustrates the comparison of cost between two approaches. It is sometimes less in time shared and in some cases less in space shared, this is because cost changes as the cloudlet length varies.

VIII. CONCLUSION AND FUTURE SCOPE

Virtual machine scheduling has become a prime concern of present day. The efficiency of scheduling algorithms depends upon several factors such as response time, cost, energy efficiency etc. In this research work the results of Space Shared and Time Shared approach for both Cloudlet level and VM level allocation policy are compared. A new algorithm for efficient allocation of processing elements to virtual machines is purposed. The purposed algorithm can be considered as an improved form of time shared policy. It minimizes the free RAM available in the host according to the processing elements. Simulation results show that proposed algorithm performs better in terms of response time and total delay.

As a future work, a wide set of experiments can be performed with different parameters showing the impact of different types of applications and their effect on the placement decisions (i.e. time shared and space shared).

REFERENCES

[1] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.

[2] Liu, L., & Qiu, Z. (2016, October). A survey on virtual machine scheduling in cloud computing. In Computer and Communications (ICCC), 2016 2nd IEEE International Conference on (pp. 2717-2721). IEEE.

[3] Liu, K., Jin, H., Chen, J., Liu, X., Yuan, D., & Yang, Y. (2010). A compromised-time-cost scheduling algorithm in swindew-c for instance-intensive cost-constrained workflows on a cloud computing platform. The International Journal of High Performance Computing Applications, 24(4), 445-456.

[4] Shenai, S. (2012). Survey on scheduling issues in cloud computing. *Procedia Engineering*, *38*, 2881-2888.

[5] Khurana, S., & Marwah, K. (2013, July). Performance evaluation of Virtual Machine (VM) scheduling policies in Cloud computing (spaceshared & timeshared). In Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on (pp. 1-5). IEEE.

[6] Benali, R., Teyeb, H., Balma, A., Tata, S., & Hadj-Alouane, N. B. (2016, June). Evaluation of traffic-aware VM placement policies in distributed Cloud using CloudSim. In Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2016 IEEE 25th International Conference on (pp. 95-100). IEEE.

[7] Buyya, R., Ranjan, R., & Calheiros, R. N. (2009, June). Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities. In High Performance Computing & Simulation, 2009. HPCS'09. International Conference on (pp. 1-11). IEEE.

Amanjot Kaur, currently pursuing MTech in the Department of Computer Science Punjabi University Patiala, Punjab

Dr. Jyotsna Sengupta, currently working as Professor in the Department of Computer Science, Punjabi University Patiala, Punjab, India. She has over 30 years of experience in teaching field and has published research papers in various National and International Journals. Her research areas are Computer Networks, Mobile Networks, Cloud Computing, Parallel and Distributed Systems