



A Survey on Virtual Machine Scheduling in Cloud Environment

Kiran Kumar Shakya, Prof. Dr. D.Singh Karaulia

Computer Technology & Applications

India

Abstract- *In cloud environment multiple cloud users can demand for a number of cloud services simultaneously. So there must be a provision that all resources are made available to requesting user in efficient manner to satisfy their need. To get the maximum benefit from cloud services, developers must design mechanisms that optimize the use resources. Virtualization [1] technology plays an important role in the cloud because it makes cloud computing infrastructures to be scalable. In the cloud computing there are number of resources with some capacity and multiple users can request for a same resources, so there should be an algorithm that optimally schedules the virtual machine resources. This paper show an analysis of different existing Virtual Machine's scheduling algorithms.*

Key Terms- *Cloud Computing, Distributed Computing, On Demand Resources, Round Robin, Virtualization, VM scheduling.*

I. INTRODUCTION

Now days cloud computing is one of the fast growing technology in the field of computer science and information technology because of online, cheap and pay as use scheme [2, 3, 4]. Main features of the cloud computing is the virtualization. Virtualization is the technology where the all the physical resource of a host are shared by the multiple users. In the virtualization all users have their own virtual machine, which is created according to the user's requirement. All the VMs are independent, so they are not effect to each other. Furthermore, one or multiple virtual machines can be run on a single host, and the utilization of resources is improved effectively. An essential requirement of cloud computing environment is scheduling the current jobs to be executed with the given constraints. The VM scheduler should order the jobs in a way that there should be a balance between quality of services and at the same time maintaining the efficiency and fairness among the jobs. In the cloud computing, the computing resources are provided to the client through the VM via internet. Each user has its own VM and each VM has its own resources such as CPU, memory, bandwidth etc.

So VM scheduling is an important part of cloud computing, which is a mechanism that maps users' VM to the appropriate host to execute, its efficiency will directly affect the performance of the whole cloud computing environment. Grid computing is different from the cloud computing. In grid computing several computers are connected parallel to solve a individual problem or to run a specific application. On the other hand cloud computing refer to the leveraging of multiple resources to deliver a services to the end user.

VM scheduling in the cloud is different from the scheduling in the grid computing, because the resource requirement in the cloud is dynamically change according to the user requirement. Since resources in the cloud environment are heterogeneous in nature and geographically distributed so the demand of the cloud user cannot predict previously, it may change dynamically on runtime. So the resource management and the resource scheduling in such a large-scale distributed environment is a very challenging task. For the cloud environment, numbers of scheduling algorithms have been proposed to enhance the total performance of the system.

II. RELATED WORK

So many scheduling algorithms have been proposed by different researchers that run under cloud computing environment. Mostly of the scheduling algorithms try to achieve two main objectives [5].

- I. User task should be run within the deadline
- II. Algorithm should be maintained the efficiency (load balancing) and fairness for all tasks.

Yang et al. [6] highlighted the issue related to the job scheduling in cloud computing environment. They argued that there is no well-defined job scheduling algorithms are available in the cloud environment that considers the system state in the future.

The existing job scheduling algorithms in cloud do not consider the hardware or software failure and recovery in the cloud. To tackle this issue they proposed a Reinforcement Learning (RL) based algorithm that helps the scheduler in making scheduling decision with fault tolerable while maximizing utilities attained in the long term.

Sindhu et al [7] proposed two algorithms for scheduling tasks into the cloud computing. They considered the computational complexity and the computing capacity of the processing elements to schedule the task. This algorithms work in private cloud environment where the resources are limited. The first algorithm is named Longest Cloudlet Fastest Processing Element (LCFPE), it uses the cloudlets computational complexity to make scheduling decisions. The second

algorithm, Shortest Cloudlet Fastest Processing Element (SCFP), maps the shorter cloudlets to the Processing Elements (PEs) having high computational power. Since it uses the largest system, so it reduces the number of ON physical host.

Paul et al. [8] proposed an algorithm which focuses on efficient resource utilization in cloud computing and gain maximum profit. For this purpose, a credit based scheduling algorithm is proposed which evaluates the entire group of tasks in the task queue and finds the minimal completion time of all tasks. The proposed scheduling method considers the scheduling problem as an assignment problem in mathematics where the cost matrix gives the cost of a task to be assigned to a resource. Problem with this algorithm is that it only considers the probability of a resource to be free soon after executing a task so that it will be available for the next waiting, but processing time of a job is not considered.

Isam Azawi Mohialdeen [9] uses the random algorithm to assign the job to the VM. In this algorithm jobs are randomly assigned to the available VM. So any process can be assigned to any VM. This method has no overhead and very less complexity.

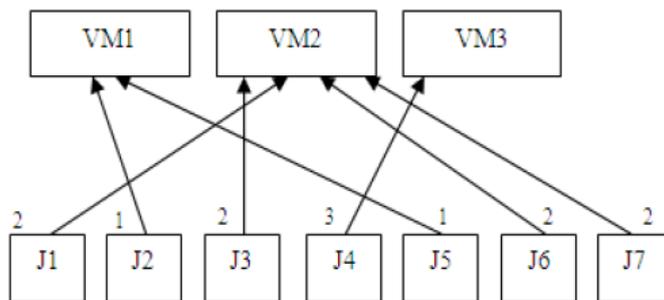


Figure1. The process of random algorithm

Problem with this approach is that it does not consider the size of the process so sometimes it considers the machine which is overloaded.

Subramanian S. et al. [14] proposed a Greedy algorithm for the VM scheduling in cloud. It is a default VM scheduling algorithm used in Eucalyptus. It is very simple and straight forward. In this algorithm first node that it finds with suitable resources for running the VM that is to be allocated. The first node that is identified is allocated the VM. This means that the greedy algorithm exhausts a node before it goes on to the next node.

Advantage:

- 1) The main advantage of the Greedy algorithm is its simplicity.
- 2) Easy to implement and also the allocation of VMs do not require any complex processing.

Drawback:

- 1) The major drawback would be the low utilization of the available resources. This drawback can be overcome by the Round Robin algorithm.

Isam Azawi Mohialdeen [9] uses the Round Robin method to assign the job to the VM. In this approach select processes are assigned to the available VM in round robin manner, where all jobs are treated equally.

Yiqiu Fang et al.[10] proposed a two level task scheduling mechanism based on load balancing in cloud computing. Cloud Computing Architecture consist of three layers, application layer, platform layer and infrastructure layer [11].The application layer is implemented to provide the interaction mechanism between user and service provider . Users can submit tasks and receive the results through the application layer. The infrastructure layer is a set of virtual hardware. Platform layer is a set of software resources with versatility and reusability, which can provide an environment for cloud application to develop, run, manage and monitor. They are using two level scheduling, first level scheduling is from the users' application to the virtual machine, and the second is from the virtual machine to host resources. First level scheduler creates a VM description and the second level scheduler find the appropriate host to the VM according to the task description.

Problem with approach is that they schedule the virtual machine to the host with lightest load each time. The advantage is to avoid overloading for the host holding more resources. Recent studies show that on average an idle server consumes approximately 70% of the power consumed when it is fully utilized. Only task response time and the demand for resources are considered in this paper.

III. FUTURE WORK

Yiqiu Fang et al.[10] select a lightest host to place a VM. So most of the times all host are ON, which will increase the energy consumption and resources wastages. This energy consumption and resources wastages can be reduced by minimizing the number of ON host, which can be achieved by the threshold. So we want to develop a threshold based approach for scheduling a VM. In this approach VM are allocated to the host till the upper level threshold value, when the host utilization near the upper threshold next host in the data center are chose for placing the VM. Because of the threshold number of working host are minimize that will increase the system performance by reducing the power consumption.

IV. CONCLUSION

This paper presents the basic knowledge about the cloud computing as a paradigm in which large numbers of resources are allocated on demand. Cloud computing provides all the computing related services through the internet. For storage there is data as a service, for application there is software as a service, for computing there is platform as a service and infrastructure as a service etc. VM scheduling is an important feature of virtualization. Resource utilization can be increased by proper scheduling of VMs. It can also play an important role to reduce the number of migrations. In this paper we discussed some exiting virtual machine scheduling algorithms along with the work. Cloud computing has numerous advantages. There are some critical issues also which need to resolved with urgency. One of the major issues of cloud computing is user privacy & security, and has found considerable favor in the research community.

REFERENCES

- [1] Peer1 hosting site puts a survey on “Visualized: ring around the world of data center power usage”. From engadget.com ,2011
- [2] Peter Mell, Timothy Grance, "Cloud Computing" by National Institute of Standards and Technology - Computer
- [3] Y. sahu et al., “cloud computing overview with load balancing techniques” International Journal of Computer Applications (0975 – 8887) Volume 65– No.24, March 2013, pp: 40-44.
- [4] Wenke Ji, Jiangbo Ma “A Reference Model of Cloud Operating and Open Source Software Implementation Mapping” in 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, 2009..
- [5] Li et al. “A scheduling algorithm for private clouds”. Proceedings of 12th IEEE International Conference on Cloud and Service Computing, 2011
- [6] Yang, B et al. “An utility based job scheduling algorithm for cloud computing considering reliability factor. Proceedings of IEEE International Conference on Cloud and Service Computing, Dec 2011.
- [7] Sindhu, S. et al. “Efficient task scheduling algorithms for cloud computing environment. Commun. Comput. Inform. Sci., PP.79-83,2011.
- [8] Paul, M. et al. “Task-scheduling in cloud computing using credit based assignment problem”, International journal of Comput. Sci. Eng., pp. 26-30, 2011
- [9] Miyuki sato “creating next generation cloud computing based network services and the contribution of social cloud operation support system (OSS) to society” 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, 2009.
- [10] Yiqiu Fang et al. “A Task Scheduling Algorithm Based on Load Balancing in Cloud Computing”, 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, 2010.
- [11] Armbrust, M.: Above the Clouds: A Berkeley View of Cloud Computing. In: EECS Department, University of California, Berkeley (2009)
- [12] Börje Ohlman, Anders Eriksson “What Networking of Information Can Do for Cloud Computing” 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, 2009.
- [13] Anas BOUA Y AD, Asmae BLILA T, Nour el houda MEJHED, Mohammed EL GHAZI “Cloud computing: security challenges” IEEE Computer Society, 2012
- [14] Subramanian S, Nitish Krishna G, Kiran Kumar M, Sreesh P4 and G R Karpagam, “An Adaptive Algorithm For Dynamic Priority Based Virtual Machine Scheduling In Cloud” IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 6, No 2, November 2012.