International Journal For Technological Research In Engineering
Volume 3, Issue 7, March-2016 *ISSN (Online): 2347 - 4718*

SURVEY ON VM SCHEDULING ALGORITHMS FOR CLOUD DATA CENTERS

Arti Patel[1], Shrikant Patel[2]
[1]Master of Engineering, [2]Department of Computer Engineering, [1,2]Alpha College of Engineering,

Abstract: ***Now a day's Cloud Computing is rapidly increase technology, moving of more and more applications on cloud and demand of clients for more services and better results, load balancing in Cloud has become a very interesting and important research area. The primary aim of Cloud Computing is to provide efficient access to remote and geographically distributed resources with the help of Virtualization in Infrastructure as a Service (IaaS). In cloud computing, multiple computers are allowed to run as virtual machines in a single physical computer is called virtual machine scheduling. For virtual machine scheduling Large amount of requests are scheduled by cloud providers, so cloud provider require an efficient scheduling algorithm for VM scheduling. In this paper, we has been discusses different type of algorithm s for VM scheduling.***

I. INTRODUCTION

Cloud Computing is a computing technology where computing is moved away from personal computers to a cloud of computers. Cloud computing enables users to access computing resources on demand without the burden of owning, managing, and maintaining the resources. Cloud provided different modeled architecture to work on with like SaaS, IaaS, and PaaS. It is on demand network access and internet based development. Cloud computing provides services like storage, server application and network components. Hence, cloud computing is a technology that uses internet to deliver its services.[1] Now a day's application ranging from simple web site to large e-commerce application is being deployed to cloud infrastructure. These applications respond client request by processing data within desirable amount of time. The application, especially e-commerce application, often needs to communicate client request frequently, so overall performance of application would be affected. Therefore, Virtual machine scheduling required maintaining application performance. As per the define requirements from the customer side computational resources are shared among users according to the Pay-Per-Use model. The allocation of resource (i.e. VMs, Applications etc) and proper scheduling has significant impact on the performance of the system. An efficient Scheduling, Provisioning, Load Balancing, Security aware infrastructure is needed to manage the access to the different resources. There are different types of resource scheduling in Cloud that are based on different parameters like cost, performance, resource utilization, time, priority, and physical distances, through put, bandwidth, and resource availability. In this paper we have concentrate on the time effective, Energy efficient and security aware scheduling and provisioning algorithms. If we need commercial cloud then Cost effective and Energy efficient cloud environment is needed. If some organization need secure data transmission and sharing of secure remote application then we have to think about secure scheduling and provisioning that do not affect the neighbor resources and prevent side-channel attack and DOS attack of one VM to the other VMs and PMs.

II. VM SCHEDULING

The role of Virtual Machine's (VMs) has important issue because, though virtualization, it make cloud computing infrastructure to be scalable. Virtualization also has the ability to run multiple operating system on a single physical system and share the underlying hardware resources. Therefore developing on optimal scheduling of virtual machines is an important issue.

Scheduling is a balancing scenario in which processes or tasks are scheduled as per the given requirements of the users. In Cloud Computing VM scheduling algorithms are used to schedule the incoming VM requests to the Physical Machines (PM) of the particular Data Center (DC) as per the requirement fulfilled with the requested resources (i.e. RAM, Memory, Bandwidth etc).

In general scheduling algorithm works in three levels as per given below [2]

1. For the set of VMs find the appropriate Physical Machine.
2. Determine the proper provisioning scheme for the VMs.
3. Scheduling the tasks on the VMs.

Basic idea of cloud sim and cloud analyst

CLOUDSIM

CloudSim is a toolkit (library) for simulation of Cloud computing scenarios. It provides basic classes for describing data centers, virtual machines, applications, users, computational resources, and policies for management of diverse parts of the system (e.g., scheduling and provisioning). It provides features for modeling and creating the data center. By using CloudSim, developers can focus on particular system design issues of the infrastructure of cloud system, without getting into the lower level details related to Cloud-based infrastructures and services. CloudSim uses SimJava as the simulation engine that supports various functions of cloud system entities (services, host, data center, broker, VMs) such as queuing and processing of events & communication between components.[3] Therefore, CloudSim is not a ready-to-use solution were you set parameters and collect results for use in your project. Being a

library, CloudSim requires that you write a Java program using its components to compose the desired scenario. Nevertheless, CloudSim can be used to build such a ready-to-use solution. for example: CloudSim is a simulator, so it doesn't run any actual software technology. Simulation can be defined as "running a model of a software in a model of hardware".

CLOUD ANALYST

Analyst is another cloud simulator that is completely GUI-based and also supports the evaluation of social network tools according to the geographic distribution of users and data centers. Cloud Analyst is very user friendly graphical user interface (GUI) simulation which provides high degree of configurability and flexibility.

Cloud Analyst also performs different types of experiments with frequently connectivity with Java for extension.

The output obtained from the simulating kit is more appropriate and understandable as it is in the form of graphical representation which can be examined by the developer and researchers in more precise form. Due this feature, it can quickly highlight the problem with the performance and accuracy of simulation results

Types of scheduling algorithm

Basicaly VM scheduling algorithm are divided as static and dynamic algorithm. For example First-Come-First-Severed (FCFS) is a good example of static VM scheduling algorithm. Genetic Algorithm is a good example of dynamic VM scheduling algorithm.[2]

III. PROPOSED WORK AND METHODOLOGY

In this paper proposed algorithm utilizes the characteristics of generalized priority algorithm. Many algorithms have been made on priority algorithm. These scheduling algorithms provide minimum execution time and best utilization of resources. In generalized priority algorithm the tasks are initially prioritized according to their size such that one having largest size has highest rank. Generalized priority scheduling algorithms takes more execution time than first come first serve and priority scheduling algorithms. But it is performing better than priority scheduling algorithm in terms of total energy consumption. Therefore we need an algorithm which

reduces the execution time as well as the energy. However, in this work, allocation function is a new hybrid virtual machine scheduling scheme for public cloud. The execution of all the tasks with minimum execution time is being addressed in this work.

Steps Involved implementing the proposed scheduling algorithm

1. Start: It is the first step of execution that initializes the process, loads all files and supported libraries in memory. After configuring system with client it transfers control to the next step that is either generate system or generate task.
2. Generate System: After initialing all the content and libraries, system is ready to perform their working. It requires the nodes to generate network for execution.
3. Generate Task: It is used to generate tasks for execution and get results from generated network of nodes with the help of proposed scheduling algorithm that schedule and generate results with different parameters.
4. Properties: For generation of tasks and systems some properties are used to define their behavior during execution. These properties are like RAM, Processor etc. These are used to behave node as a system in network for execution task as real environment.
5. Allocate task: This step define the number of tasks that accomplish in the process of scheduling. It manages them and makes a queue with their groups and allocates them.
6. Find threshold: This step is used to find a threshold and create group accordingly. Threshold is a limit that is used to create groups and divide the queue into two parts, which makes the system reliable and executes the task with less time and minimum energy consumption.
7. Limit: : It means some time nodes are busy or their limit to execute task is over and new task occur for execution than proposed system get all the possible routes and broadcast them to network for refresh their working capacity to optimize their working.
8. List Value: Now list value is used to check that the entire step that is performed before execution are successfully executed and system ready to allocate task on network of nodes that are attached with scheduler.
9. Broadcast: If some time system limit exceeds or a task is not executed. In that case system setup a virtual environment which creates a VM for their execution.
10. VM: System allocates jobs to the VM and calculates results.

IV. RESULTS AND DISCUSSION

It is necessary to understand the need of use of proposed technique. In this paper we present different scheduling algorithms,

A. First Come First Serve Scheduling Algorithm (FCFS)

In this algorithm we have randomly selected 3 tasks T0=1, T1=10 and T2=10. As we know FCFS execute the process which come first in queue. So according to this criteria first T0 task will be execute, then T1 and T2. The total time consumption by FCFS algorithm is 21ms (1+10+10).

Its main disadvantage is that the shortest tasks which are at the back of the queue have to wait for the long task at the front to be finished.

B. Priority Scheduling Algorithm

In priority algorithm, priority is defined according to the

user's demand. In this different parameters are defined like size, memory and time. In this algorithm, the tasks are initially prioritized according to their priority number.
For example: Consider 3 virtual machines represented by their priority number as Vm = {{1, 1}, {2, 10}, {3, 10}}. Here Vm1 will get first preference because of its highest priority number, second preference is given to Vm2 and then Vm3 get last preference. After observation , it is observed that total time consumed by the priority algorithm is equal to the first come first serve scheduling algorithm and its energy is more than first come first serve algorithm.

C. Genetic Algorithm
Genetic algorithms are stochastic search algorithms based on the mechanism of natural selection strategy. It starts with a set of initial solution, called initial population, and will generate new solution using genetic operators. The genetic algorithm approach computes the impact in advance, that it will have on the system after the new VM resource is deployed in the system, by utilizing historical data and current state of the system. It then picks up the solution, which will have the least effect on the system [10].
The advantage of this technique is it can handle a large searching space, applicable to complex objective function and can avoid trapping by local optimum solution.

D. Match-Making Algorithm
The MMA (match-making algorithm), algorithm first filter out the nodes or hosts those do not meet the VM requirements and do not have enough resources (Like CPU, Memory, Processors etc) to place and run the VM. Rank will be given to nodes as per the gathered information by the monitoring drivers. If any variable comes in monitoring then it will be included in to rank expression [4]. The result of rank expression is given to the cloud scheduler and monitoring driver makes decision for VMs placement and reconfiguration. The goal of this algorithm is to prioritize resources those are most suitable for the VM. Those resources with a higher rank are used first to allocate VMs.

E. Gang scheduling Algorithm
The use of gang scheduling algorithm in cloud computing responsible for selection of best suitable resources for task execution, by taking some static and dynamic parameters and restrictions of VM into the considerations.[5] Gang scheduling is a scheduling algorithm for parallel system that scheduled related VM to run simultaneously on different machines. Gang Scheduling is an efficient job scheduling algorithm for time sharing, already applied in parallel and distributed systems. Gang scheduling can be effectively applied in a Cloud Computing environment both performance-wise and cost-wise. Gang scheduling is a special case of job scheduling that allows the scheduling of such virtual Machines. Gang scheduling is a special case of scheduling parallel jobs in which tasks of jobs need to communicate very frequently. Gang scheduling involves high overhead since network status must be saved and then be restored when switching between tasks.
Moschakis et. al. [8] gives improved version of gang scheduling and performance and Cost evaluation of Gang Scheduling. The study takes into consideration both performance and cost while integrating mechanisms for job migration and handling of job starvation. The number of Virtual Machines (VMs) available at any moment is dynamic and scales according to the demands of the jobs being serviced. Results highlight that this scheduling strategy can be effectively deployed on Clouds. The number of VM's available at any moment is dynamically scales according to the demand of the job being served. For this purpose they applied "shortest Queue First (SQF) algorithm which dispatches the tasks to VMs with the shortest queue.
Gang scheduling can be effectively applied in cloud computing environment both performance wise and cost-wise. Response Time Rj, of a job j is the time interval between the arrival and the departure of the job. Its average is defined as [8]:
Response Time Rj = (Rj / n)
(Where n is the total number of jobs.)

F. Round Robin Scheduling Algorithm
We have randomly selected 3 tasks T0=1, T1=10 and T2=10. As we know each job in a queue has the same execution time and it will be executed in turn. If a job can't be completed during its turn, it will be stored back to the queue waiting for the next turn. We have created the execution time of job from 2 to 6 time slices. The advantage of round robin algorithm is that each job will be executed in turn and it doesn't have to wait for the previous one to get completed. But if the load is found to be heavy, round robin will take more time to complete all the jobs. The main advantage of this algorithm is that it utilizes all the resources in a balanced order. An equal number of VMs are allocated to all the nodes which ensure fairness. The drawback of round robin is that, the largest job takes enough time for completion. So overall performance of job will be degreed.

G. Content-Based Virtual Machine Scheduling Algorithm
The content based VM scheduling algorithms were designed with the goal of lowering the amount of data transferred between racks in the data center when virtual machines disk image are being copied to the host node [5]. The algorithm returns the selected node and the VM on that node with the highest similar content. When deploying a VM, we search for potential hosts that have VMs that are similar in content to the VM being scheduled. Then, we select the host that has the VM with the highest number of disk blocks that are identical to ones in the VM being scheduled. Once we have chosen that host node, we calculate the difference between the new VM and the VMs residing at the host; then, we transfer only the difference to the destination host. Finally, at the destination host, we can reconstruct the new VM from the difference that was transferred and the contents of local VMs. Content based VM scheduling algorithm that can significantly reduced the network traffic associated with transfer of VM's from storage racks to host racks in cloud data center.

H. Efficient Resource Utilization Algorithm

In this paper [6] explains use of the massive pool of resources in terms of pay-as-you use policy. On demand the resources are delivers by the cloud through the use of network resources under different conditions. Based on their usage the effective utilization of resources the users will be charged. His named his proposed algorithm as "Effective Resource Utilization Algorithm (ERUA)" is based on 3-tier cloud architecture (Consumer, Service Provider and the Resource Provider) which benefits both the user (QoS) and the service provider (Cost) through effective schedule reallocation based on utilization ratio leading to better resource utilization. Performance analysis made with the existing scheduling techniques shows that efficient resource utilization algorithm gives out a more optimized schedule and enhances the efficiency rate [6]. The service provider hires resources from the resource provider and creates Virtual Machine (VM) instances dynamically to serve consumers.

I. Deadline-Aware Algorithm

In this paper [7] explained a novel approach to optimize job deadlines when run in virtual machines by developing a deadline-aware algorithm that responds to job execution delays in real time and dynamically optimizes jobs to meet their deadline obligations. The algorithm intelligently schedules the jobs and learns over time about the missed deadlines under various conditions and tries to predict whether job would be meeting its deadline di, and if not then take appropriate measures to improve it chances in meeting deadline di. This implies: virtual exec. time = (duration * overhead) + duration The main goal of the Deadline aware scheduler is to guarantee a start service time for a request. It does that by imposing a deadline on all I/O operations to prevent starvation of requests, reducing job deadline miss rate and increasing job throughput rate. Deadline-aware algorithm has flexible use and high utilization of the datacenter resources

V. CONCLUSION AND FUTURE WORK

Existing scheduling algorithms give improved performance and resource utilization. Each algorithm is based on set of policies to control the order of work. Good scheduling algorithm must be energy efficient, cost effective and minimize the power requirement. FCFS, round robin and priority scheduling algorithms contained these parameters, but not efficient for all types of scenarios. We use the characteristics of generalized priority algorithm. Many algorithms have been made on priority algorithm. In this paper, we surveyed multiple algorithms and techniques for virtual machine management for Cloud Computing. We discussed the challenges that must be addressed to provide the most suitable and efficient VM scheduling algorithms. We also discussed the advantages and disadvantages of these algorithms. Our research focuses on efficient use of time consumption algorithm to VM management at different data centre's As our future work, we are planning to improve existing VM management algorithms to make it more suitable for multimedia services and application where long term connection between client and datacenter is applicable. Also to schedule Vm's such as to make Cloud environments more efficient in terms of storage utilization.

REFERENCES

[1] Harmeet Kaur, Rama Krishna Challa, "A New Hybrid Virtual Machine Scheduling Scheme for Public Cloud", in 2015 Fifth International Conference on Advanced Computing & Communication Technologies

[2] Karan D. Prajapati, Prof. Pushpak Raval, Miren Karamta, Dr. M. B. Potdar, "Comparison of Virtual Machine Scheduling Algorithms in Cloud Computing" in International Journal of Computer Applications

[3] Sumit Khurana , Khyati Marwah "PERFORMANCE EVALUATION OF VIRTUAL MACHINE (VM) SCHEDULING POLICIES IN CLOUD COMPUTING", in IEEE - 31661

[4] GWDG eScience Group, "Virtual Machine Allocation in Current Cloud Computing Middleware", 2012.

[5] Dr. Chenna Reddy , "An Efficient Profit-based Job Scheduling Strategy for Service Providers in Cloud Computing Systems" in International Journal of Application or Innovation in Engineering & Management (IJAIEM) , Volume 2, Issue 1, January 2013.

[6] Ramkumar N, Nivethitha , "Efficient Resource Utilization Algorithm (ERUA) for Service Request Scheduling in Cloud" in International Journal of Engineering and Technology (IJET) , Vol 5 No 2 Apr-May 2013.

[7] K. Parrott et. al. ," Deadline Aware Virtual Machine Scheduler for Grid and Cloud Computing" in 2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops

[8] I. Moschakis, H. Karatza, "Performance and Cost evaluation of Gang Scheduling in a Cloud Computing System with Job Migrations and Starvation Handling" in Department of Informatics Aristotle University of Thessaloniki, Greece , IEE 2013

[9] Marc E. Frincu, St´ephane Genaud and Julien Gossa, "Comparing Provisioning and Scheduling Strategies for Workflows on Clouds", 2013.

