Volume 6, Issue 7, July 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Review Paper on Efficient Usage of MIPS to Allocate Virtual Machine to Cloudlets

Kuldip Kaur*, Er. Khushdeep Kaur
Department of CSE, I.K Gujral PTU,
Punjab, India

Abstract: ***Cloud could be a term used as a trope for the wide space networks (like internet) or any such giant networked atmosphere. It came partially from the cloud-like image wont to represent the complexities of the networks within the schematic diagrams. In this paper, we have proposed a new load balancing algorithm with efficient usage of virtual machines in a cloud environment by computing the utilization power and capacity of Each virtual machine. We will be implementing the above proposed work in cloud sim simulation tool.***

Keywords: Cloud computing, Datacenter, Virtual machine, load balancing, cloudlets.

I. INTRODUCTION

With traditional desktop computing, we run copies of software programs on our own computer. the documents we create are stored on our own pc. although documents can be accessed from other computers on the network, they can't be accessed by computers outside the network. this is pc-centric. with cloud computing, the software programs one use are not run from one's personal computer, but are rather stored on servers accessed via the Internet. If a computer crashes, the software is still available for others to use. Same goes for the documents one create; they are stored on a collection of servers accessed via the Internet. Anyone with permission can not only access the documents, but can also edit and collaborate on those documents in real time. Unlike traditional computing, this cloud computing model isn't PC-centric, it's document-centric. Cloud Computing may be a term accustomed describe each a platform and kind of application. As a platform it provides, configures and reconfigures servers, whereas the servers is physical machines or virtual machines.

II. RELATED WORK

Al-Rayis et al. [1] explains that basically, load balancers can be deployed based on three different architectures. The centralized load balancing architecture which includes a central load balancer to make the decision for the entire system regarding which cloud resource should take what workload and based on which algorithm(s). In the hierarchical load balancing architecture, a main load balancer (parent) receives all job requests, and then it spreads them to other connected load balancers (children) where each load balancer in the tree may use a different algorithm. The special characteristics of cloud environments that result from the complexities of cloud computing virtual infrastructure require advanced load balancing solutions that are capable of dynamically adapting the cloud platform while providing continuous service and performance guarantees. These difficulties have created different views on which load balancing architecture would better suites cloud computing.

Bhoi et al. [2] discussed that in enhanced Max-Min Task Scheduling Algorithm in cloud computing helps in supplying a high performance computing based on protocols which allowed shared computation and storage over long distances. It depends upon expected execution time instead of completion time. Max-Min algorithm assign task with maximum execution time to resource produces minimum completion time while Enhanced Max-min assign task with average execution time to resource produces minimum execution time.

Bhadani et al. [3] proposed a Central Load Balancing Policy for Virtual Machines (CLBVM) that balances the load evenly in a distributed virtual machine/cloud computing environment.

Bendiab et al. [4] introduced the Map Reduced based Entity Resolution load balancing technique in networking which is based on large datasets. In this technique, two main tasks are done: Map task and Reduce task which the author has described. For mapping task, the PART method is executed where the request entity is partitioned into parts. And then COMP method is used to compare the parts and finally similar entities are grouped by GROUP method and by using Reduce task. Map task reads the entities in parallel and process them, so that overloading of the task is reduced.

Birattari et al. [5] proposed troubleshoot of load balance in Cloud computing using Stochastic Hill Climbing. Buzato et al. [6] proposed Bee Life algorithm which was used for scheduling in Cloud computing. Bee Life algorithm is inspired by the behavior and reproduction of bee to find food source. The algorithm evaluated the performance of the resources and it has the aim to reduce time and complexity of work.

Babu et al. [7] proposed a Honey Bee Behavior inspired Load Balancing [HBB-LB] technique which helps to achieve even load balancing across virtual machine to maximize throughput. It considers the priority of task waiting in queue for execution in virtual machines. After that work load on VM calculated decides whether the system is overloaded, under loaded or balanced. And based on this VMs are grouped. New according to load on VM the task is scheduled on VMs.

Task which is removed earlier. To find the correct low loaded VM for current task, tasks which are removed earlier from over loaded VM are helpful. Forager bee is used as a Scout bee in the next steps.

Dorigo et al. [8] has proposed a load balancing technique called colony of cooperating agents in ants based on soft computing for solving the optimization problem. This technique solves the problem with high probability. It is a simple loop moving in direction of increasing value which is uphill. And this make minor change in to original assignment according to some criteria

Deldari et al. [9] proposed a novel load balancing algorithm called VectorDot in intelligent ants. It handles the hierarchical complexity of the datacenter and multidimensionality of resource loads across servers, network switches, and storage in an agile data center that has integrated server and storage virtualization technologies.

Desai et al. [10] discusses about the emerging technology i.e. a new standard of large scale distributed computing and parallel computing. It provides shared resources, information or other resources as per clients' requirements at specific times. For better management of available good load balancing techniques are required. And through better load balancing in cloud, performance is increased and user gets better services. So in this author has discussed many different load balancing techniques used to solve the issue in cloud computing environment.

Elzeki et al. [11] discussed in Improved Max-Min Algorithm in Cloud Computing that focuses on the cloud computing which further deals with the allocation of the tasks to the resources while observing different parameters like waiting time, Average waiting time, Turn Around time, processing cost. So, an algorithm named as Max-Min in improved manner from load balancing has been shown to overcome such kinds of problems. The algorithm calculates the expected completion time of the submitted tasks on each resource. Then the task with the overall maximum expected execution time is assigned to a resource that has the minimum overall completion time.

Fahringer et al. [12] introduced a static load balancing technique called Ant Colony Optimization. In this technique, an ant starts the movement as the request is initiated. This technique uses the Ants behavior to collect information of cloud node to assign task to the particular node. In this technique, once the request is initiated, the ant and the pheromone starts the forward movement in the pathway from the "head" node. The ant moves in forward direction from an overloaded node looking for next node to check whether it is an overloaded node or not. Now if ant find under loaded node still it move in forward direction in the path. And if it finds the overloaded node then it starts the backward movement to the last under loaded node it found previously. In the algorithm if ant found the target node, ant will commit suicide so that it will prevent unnecessary backward movement.

Fang et al. [13] discussed a two-level task scheduling mechanism based on load balancing to meet dynamic requirements of users and obtain high resource utilization. It achieves load balancing by first mapping tasks to virtual machines and then virtual machines to host resources thereby improving the task response time, resource utilization and overall performance of the cloud computing environment.

Gellerb et al. [14] introduced a static well-known load balancing technique called Round Robin, in which all processes are divided amid all available processors. The allocation order of processes is maintained locally which is independent of the allocation from the remote processor. In this technique, the request is sent to the node having least number of connections, and because of this at some point of time, some node may be heavily loaded and other remain idle. This problem is solved by CLBDM (Central Load Balancing Decision Model).

III. PROBLEM FORMULATION

- The currently proposed work will work effectively and efficiently only in the homogeneous environment where all the machines are of same capacity. But as we know that in the cloud computing model, the configurations of the machines will be different from each other as the different users will have different requirements.
- The tasks of the user are allocated on the basis of availability of virtual machine. There is no checking of capacity of the virtual machine before allocating the request. There can be a scenario where a machine with high configuration is sitting idle and we have assigned the task to the low configuration machine. This may lead to overutilization and under-utilization of resources.
- There is no checking of the requirement of the user whether user wants to use the machine of high configuration or low configuration.

IV. OBJECTIVES

The primary purpose of the cloud system is that its client can utilize the resources to have economic benefits. A resource allocation management process is required to avoid underutilization or overutilization of the resources which may affect the services of the cloud.

- To study and evaluate the performance of existing load balancing algorithm.
- To take the heterogeneous Virtual machines of different capacities inside different hosts of multiple datacenters.
- To take into account the MIPS of VMs along with the Utilization Power and capacity of each VM.
- To create the cloudlets of different requirements and services.
- To allocate the virtual machines to the cloudlets by matching their specifications.

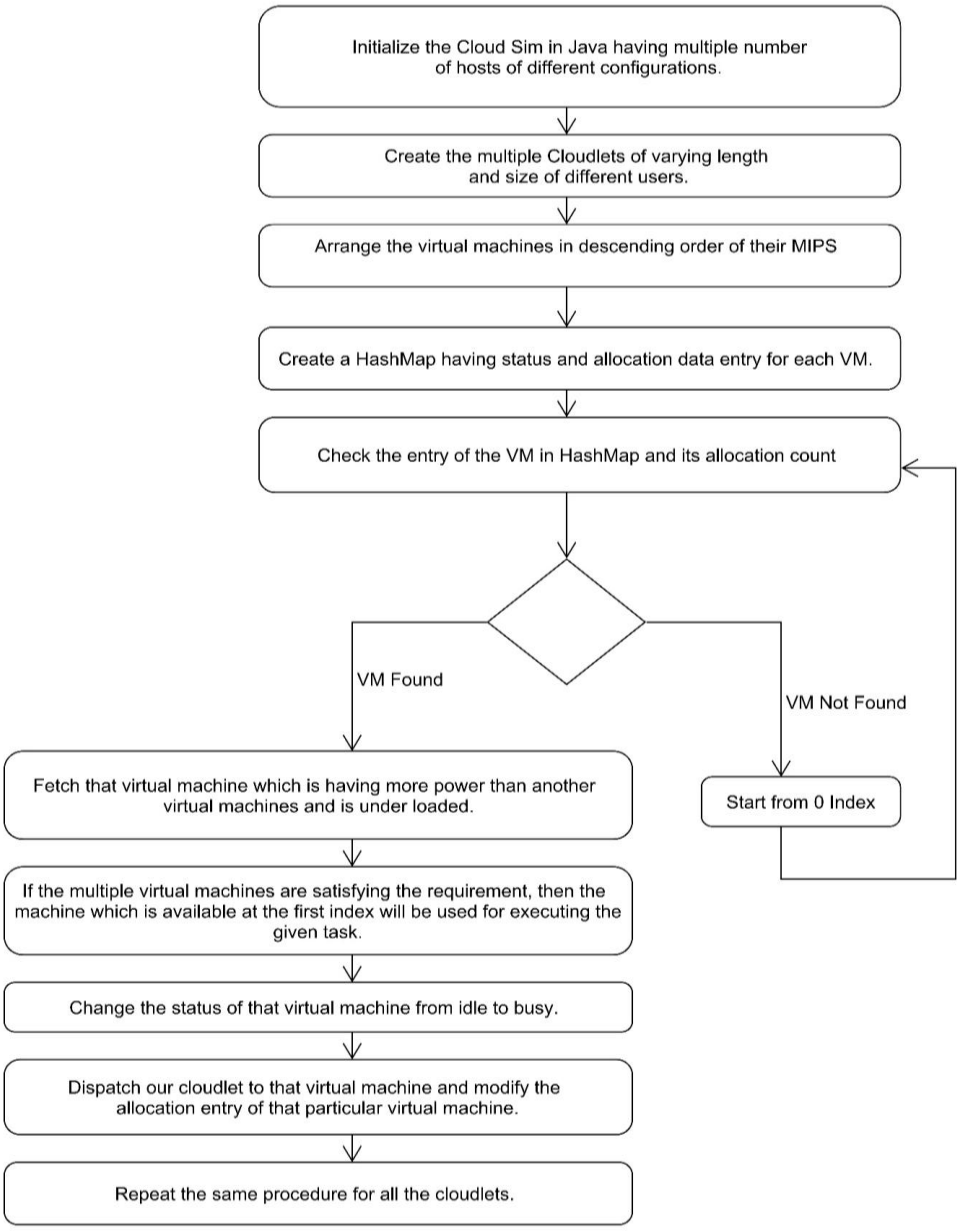
V. PROPOSED WORK

Methodology/Planning of Work

Cloud services provide computing on demand in real time. Number of users accessing cloud environment are always more than that were using it on previous day. Cloud has application areas for developing applications, providing and

managing infrastructure, patching applications. Users and their requests for accessing cloud infrastructure are highly dynamic and loading servers running in data center. We need efficient strategy to balance load on these servers so that the servers don't get crash and they can persist long. Precisely Objective is to achieve accuracy, performance of servers and the cloud environment can be maintained.

1. Initialize the Cloud Sim in Java by creating different data centers having multiple number of hosts of different configurations.
2. Each Host in the datacenter will have the different numbers of Virtual machines of different capacities.
3. Then we will create the Cloudlets of varying length and size of different users.
4. Arrange the virtual machines in descending order of their MIPS (Milliion Instructions per Second). The virtual machine having more MIPIS, is more powerful.
5. The HashMap will contain the list of all the Virtual Machines will their status and allocation count. The status will specify whether the virtual machine is busy or idle. The allocation entry in the table will define the total number of tasks submitted to this particular virtual machine.
6. The list containing the Virtual machines and Cloudlets will be given to the Data Center Broker (DCB).
7. Arrange the cloudlets in descending order of their instruction size.
8. DCB will fetch the first cloudlet from the list and will check the requirements of that virtual machine.
9. Fetch that virtual machine which is having more power than another virtual machines and is under loaded.
10. If the multiple virtual machines are satisfying the requirement, then the machine which is available at the first index will be used for executing the given task.
11. Change the status of that virtual machine from idle to busy.
12. Dispatch our cloudlet to that virtual machine and we will modify the rating of that particular virtual machine.
13. Repeat the same procedure for all the cloudlets.



Flow Chart of Proposed Methodology

VI. IMPLEMENTATION TOOLS

Implementation Language: Java

Java is a general-purpose computer programming language that is concurrent, class-based, object-oriented, and specifically designed to have as few implementation dependencies as possible. It is intended to let application developers "write once, run anywhere" (WORA), meaning that code that runs on one platform does not need to be recompiled to run on another. Java applications are typically compiled to bytecode [16] that can run on any Java virtual machine (JVM) regardless of computer architecture.

Cloud Sim

CloudSim is an extensible simulation toolkit that enables modeling and simulation of Cloud computing systems and application provisioning environments. The CloudSim toolkit supports both system and behavior modeling of Cloud system components such as data centers, virtual machines (VMs) and resource provisioning policies. It implements generic application provisioning techniques that can be extended with ease and limited effort.

VII. CONCLUSION

Cloud computing involves the Sharing of resources, software, information via internet with an objective to reduced capital and operational cost, better performance in terms of response time and data processing time. To maintain the system stability and to accommodate future modification in the system , there are various technical challenges that needs to be addressed like load balancing. In this paper, we have proposed a new load balancing mechanism by computing the utilization power and capacity of each virtual machine before assigning the task to the virtual machine. We will be implementing the proposed methodology in the cloud sim simulator by taking various types of workloads. It wil increase the overall efficiency of the cloud environment.

REFERENCES

[1] Al-Rayis, Ektemal, “Performance Analysis of load balancing architectures in cloud computing”, IEEE, European Modeling Symposium, pp. 520-524, ISBN 978-1-4799-2577-3, November 2013.

[2] Bhoi Upend ra Bhoi, “Enhanced max-min Task scheduling Algorithm in cloud computing”, International Journal of Application or Innovation in Engineering & Management (IJAIEM), Vol. 2, Issue 4, ISSN 2319 - 4847, April 2013.

[3] Bhadani A. and Chaudhary S., "Performance evaluation of web servers using central load balancing policy over virtual machines on cloud", COMPUTE '10, Proceedings of the Third Annual ACM Bangalore Conference, Article no. 16, ISBN 978-1-4503-0001-8, 2010.

[4] Bendiab A.,Randles M., Lamb D., "A comparative study into distributed load balancing algorithms for cloud computing", IEEE, 24th International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 551-556, April 2010.

[5] M. Birattari and T. Stutzle, "Ant Colony Optimization-Artificial Ants as a Computational Intelligence Technique", IEEE, Computational Intelligence Magazine, Vol. 1, Issue 4, pp. 28-39, November 2006.

[6] Buzato L.E, Nakai A.M., Madeira E, 5th Latin-American Symposium on Dependable Computing, pp. 156-165, 2011.

[7] Babu, "Honey bee behavior inspired load balancing of tasks in cloud computing environments", Elsevier, Applied Soft Computing Journal, ISSN : 1568-4946, 2013.

[8] M. Dorigo, and T. Stutzle, "Ant Colony Optimization-Artificial Ants as a Computational Intelligence Technique", IEEE, Computational Intelligence Magazine, Vol. 1, Issue 4, pp. 28-39, November 2006.

[9] H. Deldari and M. Salehi , "Grid Load Balancing using an Echo System of Intelligent Ants", Proceedings of the 24th IASTED International Conference on Parallel and Distributed Computing and Networks, pp. 47-52, 2006.

[10] Desai, “A survey of various load balancing techniques and challenges in cloud computing”, International Journals of scientific and technology research, Vol. 2, Issue11, ISSN 2277-8616, November 2013.

