Journal of Engineering Sciences
Vol 10, Issue 9, Sept /2019
ISSN NO:0377-9254

AN EFFICIENT VIRTUAL MACHINE MANAGEMENT AND SCHEDULING IN CLOUD COMPUTING

[1]G. Prathibha Priyadarshini, [2]S. Somasekhar, [3]H. Ateeq Ahmed
[1]Assistant Professor, [2]Assistant Professor, [3]Assistant Professor
DEPARTMENT OF CSE
Dr. K V SUBBA REDDY INSTITUTE OF TECHNOLOGY, KURNOOL

Abstract - Due to rapid increase in use of Cloud Computing, 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. VM Scheduling is essential for efficient operations in distributed environments. In cloud computing the load balancing concept broadly classify in three stages as Data Centre Selection, V irtual Machine Scheduling and Task scheduling at particular data centre. Many algorithms were suggested to provide efficient mechanisms and algorithms for assigning the client's requests to available Cloud nodes. In this paper, we explained different algorithms and techniques proposed for Virtual Machine Scheduling either at single data centre or multiple data center. Also infers their characteristics to resolve the issue of efficient Virtual Machine Management in Cloud Computing. We discuss and compare these algorithms and techniques in regards of various performance matrices to provide an overview of the latest approaches in the field.

Keywords— Cloud Computing, VM management, scheduling algorithms, data center, scheduling techniques

I. INTRODUCTION

Any solution where data storage and any processing take place without the user being able to pinpoint the specific computer carrying. Cloud computing refers to both thee application delivered as services over the internet and the hardware and system software in the data center that provides those services. Cloud computing provides shared pool of resources on-demand over network on pay per use. Cloud computing insures access to virtualized it resources that data center are presented and are shared by others. It is common to divide cloud computing into three categories:

A. Infrastructure as a service (IaaS)

It provides flexible ways to create use and manage virtual machines. In IaaS model, computing resources such as storage, network, and computation resources are provisioned as services. Consumers are able to deploy and run arbitrary software, which can include operating systems and applications. Consumers do not manage or control the underlying cloud infrastructure but have to control its own virtual infrastructure typically constructed by virtual machines hosted by the IaaS vendor. This thesis work mainly focuses on this model, although it may be generalized to also apply to the other models.

Platform as a service (PaaS)

Focused on providing the higher level capabilities more than just virtual machines required to supports applications. In the PaaS model, cloud providers deliver a computing platform and/or Solution stacks typically including operating system, programming language execution environment, database, and web server [5]. Application developers can develop and run their software on a cloud platform without having to manage or control the underlying hardware and software layers, including network, servers, operating systems, or storage, but maintains the control over the deployed applications and possibly configuration settings for the application-hosting environment.

II. LITERATURE SURVEY

The goal of scheduling algorithms in distributed systems is spreading the load on processors and maximizing their utilization while minimizing the total task execution time Job scheduling, one of the

Journal of Engineering Sciences

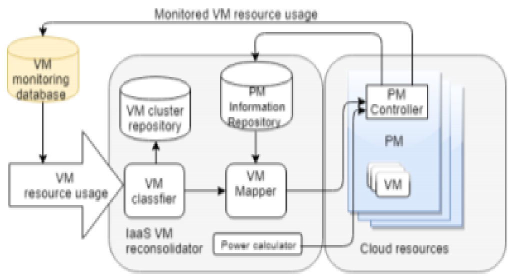
Vol 10, Issue 9, Sept /2019
ISSN NO:0377-9254

most famous optimization problems, plays a key role to improve flexible and reliable systems. The main purpose is to schedule jobs to the adaptable resources in accordance with adaptable time, which involves finding out a proper sequence in which jobs can be executed under transaction logic constraints [2].

Reddy [7] explain 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. 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.

Round Robin is proportionally fair algorithm, or maximum throughput scheduling (throughput). The main advantage of this algorithm is that it utilizes all the resources in a balanced order (resource utilization). The scheduler starts with a node and moves on to the next node, after a VM is assigned to that node. This is repeated until all the nodes have been allocated at least one VM and then the scheduler returns to the first node again. Hence, in this case, the scheduler does not wait for the exhaustion of the resources of a node before moving on to the next (Fault tolerant) [6].

III. PROPOSED SYSTEM MODEL



Our proposed architecture is shown in Fig 2 and its components are explained in this section. The first part is called IaaS VM reconsolidator because the process of consolidating a VM starts again after A VM has been profiled.

VM Classifier: this component is used to classify VMs based on their historical resource usage. It is trained using historical data harvested from VMs. It receives VM resource usage from the VM monitoring database and then classifies it based on CPU usage, memory usage and disk usage. The complete process of clustering is discussed in Section V. After a VM has been classified, the classification results are stored in a VM cluster repository and forwarded to the VM mapper.

VM Mapper: this component receives classification results from the VM classifier and determines the new host for the classified VM. This is our modified form of VM allocation policy, which we refer to as First Fit Increasing Similarity (FFIS). From the host list, we find all hosts, we call them candidate hosts (candidateHostList), which have enough resources to accommodate the classified VM. The candidate hosts are then sorted in order of increasing similarity of VMs in running hosts with classified VM. Similarity Index, I, of a host machine with classified VM is computed as shown in Equation 1. The first host in the sorted candidate host becomes the new host. The complete operation of VM mapper is shown in Algorithm 1.

Algorithm 1: VM Mapper Operation

Input: *hostList, classfiedVm, VmClass, oldHost*

Output: *newHost, classfiedVm, migrationVerdict*

1. for *host* in *hostList* do
2. if *host* has enough resources to accommodate *classfiedVm* then
3. *candidateHostList.add(host)*
4. end if
5. end for
6. *candidateHostList.sort(VmClass, candidateHostList)*
7. *newHost* equal to *candidateHostList.get(0)*
8. *migrationVerdict* equals to '*do nothing*'
9. if *newHost* is not same as *oldHost* then
10. *migrationVerdict* equals '*migrate*'
11. end if
12. return *newHost, classfiedVm, migrationVerdict*

Vol 10, Issue 9, Sept /2019
ISSN NO:0377-9254

VIRTUAL MACHINE CLUSTERING USING KMEANS

In order to group the pool of VMs (520 in number in Materna-Trace-3) k-means clustering algorithm has been used. The basic k-means algorithm is shown in Algorithm 2. Closeness is computed using Euclidian distance As our clustering feature set, we have used the following features for each VM

Algorithm 2: **Basic k-means algorithm**

Input: *Historical VM resource usage (CPU and Memory), Number of the cluster, K*

Output: *Centers of cluster*

1: Select *K* points as initial centroids

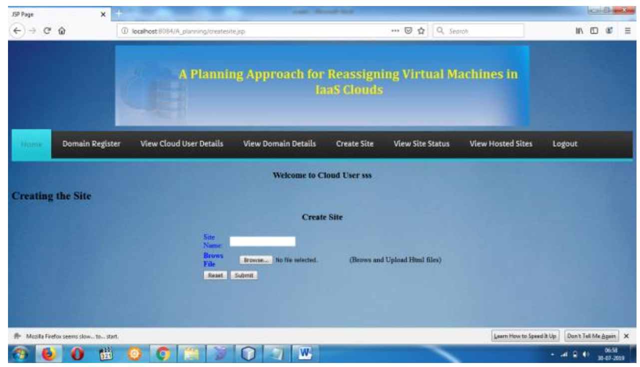
2: **repeat**

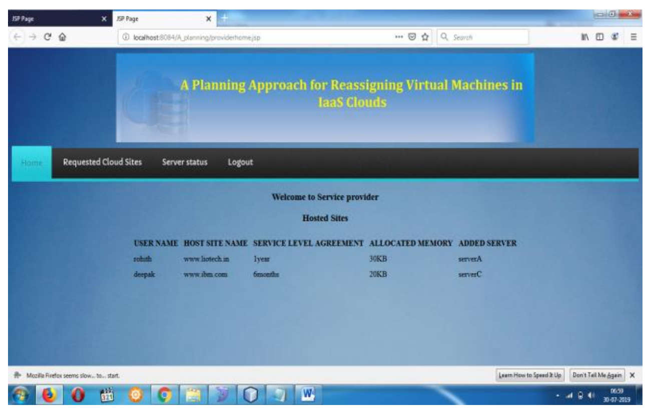
3:From *K* clusters by assigning each to its closest centroid

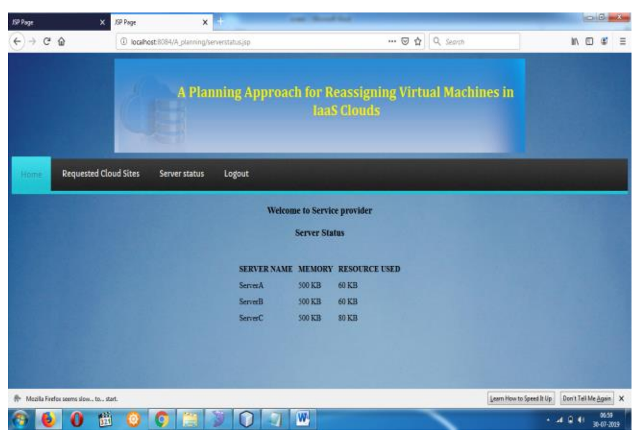
4: Recalculate centroids for each cluster

5: **until** Centroids do not change

IV. IMPLEMENTATION AND RESULTS







V. CONCLUSION

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. Then, we compared the existing algorithms based on the performance matrices we discussed. Our research focuses on efficient use of Renewable Energy Source provisioned algorithm to VM management at different data centre's located at different geolocatable locations and also to consider the parameters like bandwidth available and latency delays before use of such scheduling algorithm. 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]. X. Chen, L. Rupprecht, R. Osman, P. Pietzuch, F. Franciosi and W. Knottenbelt, "CloudScope: Diagnosing and Managing Performance Interference in Multi-tenant Clouds," in 2015 IEEE 23rd International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 2015.

[2]. Industry Outlook, "Industry Outlook Data Center Energy Efficiency," 2014. [Online]. Available: http://www.datacenterjournal.com/industryoutlook-data-center-energy-efficiency/. [Accessed 10 October 2018].

[3]. M. D. Kenga, V. Omwenga and P. Ogao, "Energy Consumption in Cloud Computing Environments," in Pan African Conference on Science, Computing and Telecommunications (PACT) 2017, Nairobi, 2017.

[4] Hadi Salimi , "Advantages, Challenges and Optimizations of Virtual Machine Scheduling in Cloud Computing Environments" in International Journal of Computer Theory and Engineering Vol. 4, No. 2, April 2012.

[5] Pinal Salot , "A Survey Of Various Scheduling Algorithm In Cloud Computing Environment" in M.E, Computer Engineering, Alpha College of Engineering, Gujarat, India , Volume: 2 Issue: 2.

[6] MR.NISHANT, "Pre-Emptable Shortest Job Next Scheduling In Private Cloud Computing" in journal of information, knowledge and research computer engineering, NOV 12 TO OCT 13 | VOLUME – 02, ISSUE – 02.

[7] TARUN GOYAL, "Host Scheduling Algorithm Using Genetic Algorithm In Cloud Computing Environment", International Journal of Research in Engineering & Technology (IJRET) Vol. 1, Issue 1, June 2013.

[8]. F. P. Sareh, R. N. Calheiros, J. Chan, A. V. Dastjerdi and R. Buyya, "Virtual Machine Customization and Task Mapping Architecture for Efficient Allocation of Cloud Data Center Resources," The Computer Journal, 2015.

[9]. B. Adrian and L. Heryawan, "Analysis of Kmeans Algorithm For VM Allocation in Cloud Computing," in 2015 International Conference on Data and Software Engineering (ICoDSE), Yogyakarta, Indonesia, 2015.

[10]. R. Neha and J. Rishabh, "Cloud Computing: Architecture and Concept of Virtualization," International Journal of Science, Technology & Management, vol. 4, no. 1, 2015.