IPHV8I109X
International Journal Of Advanced Research and Innovation -Vol.8, Issue .I
ISSN Online: 2319 – 9253
Print: 2319 – 9245

AUTOMATED CONTROL OF MULTIPLE SERVERS STORAGE VIRTUALIZATION

[#1] P. MAHEEDHAR VARMA – Pursuing M.Tech,
[#2] Dr. PENMETSA VAMSI KRISHNA RAJA -Principal,

Dept of CSE, AMALAPURAM INSTITUTE OF MANAGEMENT SCIENCES AND COLLEGE OF ENGINEERING, MUMMIDIVARAM, WEST GODAVARI, A.P., INDIA.

Abstract*— Cloud computing has become the collective model for grouping the different technologies which are collaborated to provide the services on demand. Cloud computing is known as a dynamic service provider using very large scalable and virtualized resources over the Internet. These services include IaaS (Infrastructure as a Service), PaaS (Platform as a Service), SaaS (Software as a Service) and DaaS (Data as a Service). VM allocation allows efficient sharing of virtual machines to available datacenters and these allocation policies help to evaluate and enhance the cloud performance. Different allocation policies are available and they have their own advantages and limitations. In this paper a new dynamic VM allocation policy is introduced that takes VM`s as per user requirement and allocate them in cluster form to the available datacenters. These clusters of VM`s are formed by using K-Means clustering algorithm. So before moving to any datacenters, sets of VM`s are created and then passed to the nearest datacenterId. It allows fast accessing of servers and also efficient utilization of available resources. This reallocation of VM`s improves the performance of CPU, memory and network operations by reducing the load on datacenters. The implementation of proposed algorithm is performed by using CloudSim3.0.1 simulator.

***Keywords*— Cloud Computing, VM allocation, Clustering, Virtualization, CloudSim**

I. INTRODUCTION

Cloud computing is a model for enabling convenient and on demand accessing of the resources. This model of cloud offers a shared pool of resources that are available on customer`s demand and can be accessed at anytime from anywhere. The collaboration of technologies has allowed most talked and popular term to cloud computing in recent times. Most of the organizations have been shifted to the cloud computing and had a great impact on their services and efficiency. Thus the growth of the IT technologies with cloud computing has reached to the customer in an efficient manner. As cloud offers a package of services that enables a whole new way of using IT and can be accessed on demand.

VM allocation is a keyword used in cloud computing for virtual sharing of physical machine among the datacenters. It provides the knowledge of allocated VM to a particular datacenterId. This allocation is based on different policies that make it efficient and easy to understand. These allocation policies can be implemented at virtualization level. The virtualization of cloud elements takes place at the infrastructure layer. Depending on these allocation policies, cloud infrastructure is highly structured and scalable.

II. LITERATURE REVIEW

Mahendiran et al [17] have implemented K-Means clustering algorithm in cloud computing environment. It is obtained that both Data Mining techniques and Cloud Computing helps the business organizations to achieve maximized profit and cut costs in different possible ways. Thus K-Means clustering algorithm, which is one of the very popular and high performance clustering algorithms, is used in cloud. The main aim of this work was to implement and deploy K-Means algorithm in Google Cloud using Google App Engine with Cloud SQL. Implementation of K-Means algorithm is done in java, so Eclipse IDE is chosen for design and development of the application. To deploy the application in Google, Google App Engine Plug-In is used and Google Cloud SQL is chosen for creating Database and tables.

G.Malathy et al [1] proposed the Reservation Cluster approach for performance enhancement in cloud computing. The concept of reservation cluster is to schedule the unscheduled tasks. Unscheduled tasks are sent to the reservation cluster and in this cluster all the tasks are scheduled simultaneously without any iteration. It reduces the amount of computation time and resource usage and allows better performance. Figure 1 shows the Structure of reservation cluster-based cloud computing approach:

Michael Shindler et al [5] proposed fast and accurate k-means clustering for large datasets. As cloud computing deals with large data centres, so huge amount of data need to be accessed simultaneously. Thus it impacts on the performance. In this way, fast and accurate clustering algorithm seems to be a better option to adopt in clouds. Results of this work show that K-Means clustering is much faster than other Divide and Conquer Algorithms.

Jasmin James [14] has proposed efficient VM load balancing algorithm and implemented for an IaaS framework in Simulated cloud computing environment; i.e. „Weighted Active Monitoring Load Balancing Algorithm" using CloudSim tools. This algorithm is proposed for the Data centre to effectively load balance the requests between the available virtual machines

assigning a weight, in order to achieve better performance parameters such as response time and Data processing time.

Sonam Rathore [15] has proposed efficient Allocation algorithm of Virtual Machine in Cloud Computing Environment which uses all the combination of allocation sequence and chooses the allocation sequence on the basis of strength of allocation. Proposed model is implemented in JAVA using NetBeans IDE in CloudSim. The algorithm is tested for different sets of VM instance request and computing nodes. In this work the experimental results show that proposed algorithm can improve resource utilization by efficient VM allocation.

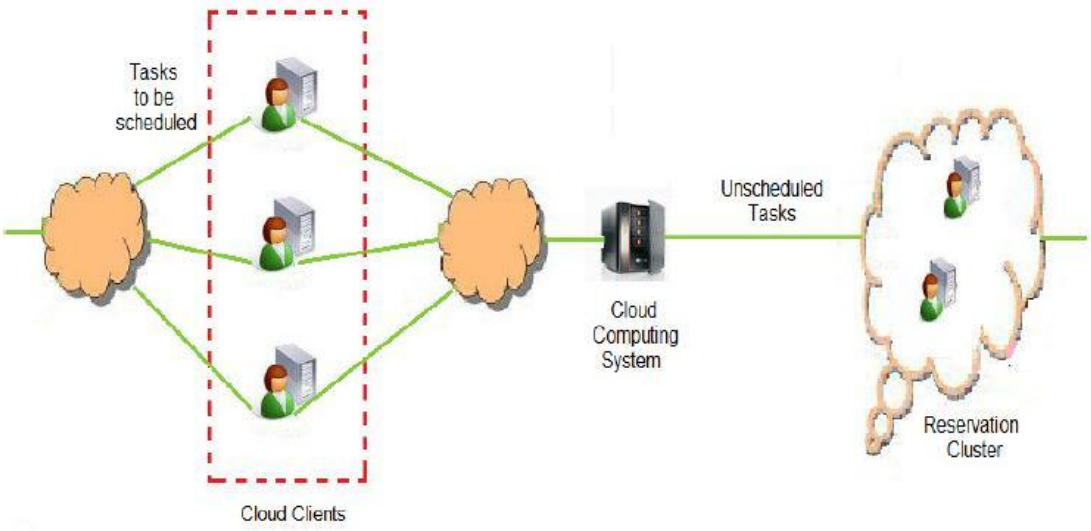

Fig1: Structure of reservation cluster-based cloud computing approach

Soumya Ray et al [16] have performed execution analysis of load balancing algorithms in Cloud Computing environment. This work was aimed to identify qualitative components for simulation in cloud environment and then based on these components, execution analysis of load balancing algorithms such as Round Robin Algorithm, Central Queuing Algorithm, Randomized Algorithm, Token Routing, and Least Connection Mechanism are presented. These algorithms are implemented in CloudSim simulator. Execution analysis of the simulation has shown that change of MIPS effect the response time. It is also observed thorough the study that, load balancing algorithm works on the principle on which situation workload is assigned, during compile time or run time. Depending on the compile time or run time it may be static or dynamic. Static algorithms are more stable than dynamic algorithm and it is easy to predict the behavior of static algorithm also. From the results analysis it is obtained that Dynamic algorithms are really works better in case of distributed environments.

III. MODELLING THE VM ALLOCATION

Cloud computing infrastructure is the massive deployment of virtualization tools and techniques as it has an extra layer i.e. Virtualization layer that acts as a creation, execution, management, and hosting environment for application services. The modelled VMs in the above virtual environment are contextually isolated but still they need to share computing resources- processing cores, system bus etc. Hence, the amount of hardware resources available to each VM is constrained by the total processing power i.e. CPU, the memory and system bandwidth available within the host. The choice of virtual machine, meaning that you can select a configuration of CPU, memory, storage, bandwidth etc. that is optimal for an application (Jasmin James et. al[14]).

Contemporary VM Load Balancers

Virtual machine enables the abstraction of an OS and Application running on it from the hardware. The interior hardware infrastructure services interrelated to the Clouds is modelled in the CloudSim simulator by a Datacenter element for handling service requests. These requests are application elements sandboxed within VMs, which need to be allocated a share of processing power on Datacenter˝s host components. Datacentre object manages the data centre management activities such as VM creation and destruction and does the routing of user requests received from User Bases via the Internet to the VMs. The Data Center Controller uses a VmLoadBalancer to determine which VM should be assigned the next request for processing. The contemporary Vmloadbalancer are Round Robin, throttled and active monitoring load balancing algorithms.

A. **Round Robin Load Balancer (RRLB)**

In this, the datacenterId controller assigns the requests to a list of VMs on a rotating basis. The first request is allocated to a VM- picked randomly from the group and then the Datacentre controller assigns the subsequent requests in a circular order. Once the VM is assigned the request, the VM is moved to the end of the list.

B. **Throttled Load Balancer (TLB)**

The TLB maintains a record of the state of each virtual machine (busy/ideal). If a request arrived concerning the allocation of virtual machine, the TLB sends the ID of ideal virtual machine to the data centre controller and datacenter

controller allocates the ideal virtual machine.

C. **Active Monitoring Load Balancer (AMLB)**

The AMLB maintains information about each VMs and the number of requests currently allocated to which VM. When a request to allocate a new VM arrives, it identifies the least loaded VM. If there are more than one, the first identified is selected. ActiveVmLoadBalancer returns the VM id to the Data Center Controller. The data Center Controller sends the request to the VM identified by that id. DataCenterController notifies the ActiveVmLoadBalancer of the new allocation and cloudlet is sent to it.

IV. PROPOSED DYNAMIC VM ALLOCATION ALGORITHM

Mainly Clustering is the method which includes the grouping of similar type objects into one cluster and a cluster which includes the objects of data set is chosen in order to minimize some measure of dissimilarity. For scheduling the virtual machines, K-Means clustering algorithm is used. K-Means clustering is a clustering method in which the given data set is divided into K number of clusters. K-means clustering is a well known partitioning method. In this objects are classified as belonging to one of K-groups. The results of partitioning method are a set of K clusters, each object of data set belonging to one cluster. In each cluster there may be a centroid or a cluster representative

Dynamic VM allocation using Clustering algorithm: K-Means follows the partitioned or non-hierarchical clustering approach. It involves partitioning the given data set into specific number groups called Clusters. Each cluster is associated with a centre point called centroid. Each point is assigned to a cluster with the closest centroid. Proposed dynamic VM allocation algorithm using clustering is as:

Input: List V of Virtual Machine`s with their location around the globe List D of datacenters

Algorithm:

1. Select K points according to the number of datacenters in D 2. Choose datacenter from D
3. Form K clusters of VM`s from V by assigning closest centroid 4. Recomputed the centroid of each cluster
5. Arrange all the requested VM`s in cluster form 6. Allocate the VM`s to the available Host
7. If all the VM`s are allocated
8. Assign the VM`s cluster to the selected datacenter 9. Endif
10. Repeat [2] until D is empty
11. If all the VM`s are created in the datacenters 12. Send the cloudlets to the created VM`s
13. Endif
14. compute the results

The Initial centroid will be chosen randomly. The centroid is nothing but the mean of the points in the cluster. Euclidean distance is used to measure the closeness. K-Means generates different clusters in different runs.

V. IMPLEMENTATION OF PROPOSED ALGORITHM IN CLOUDSIM

Before perform the implementation, it is required to know the places where modification can be made. In CloudSim different classes are there that support the simulation environment for the cloud computing. So in order to implement our own policy, it is essential to have knowledge about existing allocation policies and the classes that support these allocation strategies. As in the previous sections, we have studied the required classes that are for our purpose. DatacenterBroker class is the place where the VM allocation policies are carried out. Different functions are there in this class that help to process the virtual machines and their assignment to the datacenters.

Since clustering is the new concept in the CloudSim, so some new classes are also created in CloudSim to compute our working policies.

Implementation of CloudSim in Eclipse

CloudSim is a java based simulation tool, so it can be used either with the eclipse IDE or NetBeans IDE. For our work, we select eclipse IDE to implement proposed VM allocation policy. Different versions of eclipse IDE are available to run the CloudSim such as eclipse Indigo, eclipse Juno etc. Our work is implemented on the eclipse Juno. To run the CloudSim in eclipse Juno, we have to download the eclipse IDE and install it. Since eclipse is java base platform, so a java run time environment is needed before installing it. When this installation is completed, latest CloudSim package is extracted and it is imported in the eclipse. CloudSim package contains list of source files, jar files, list of supporting classes and some examples to understand the behaviour of cloud computing simulation. When we import the CloudSim in eclipse, it asks for the path where the extracted CloudSim package is available. This package is then copied into the workspace of eclipse from

where it can be easily configured and run.

Implementing Dynamic VM Allocation Using Clustering Algorithm

K-Means follows the partitioned or non hierarchical clustering approach. It involves partitioning the given data set into specific number groups called Clusters (Navjot Kaur et. al [19]). Each cluster is associated with a center point called centroid. Since our concept is to allocate the virtual machines dynamically, as per user request. So a graphical interface is prepared for selecting the position of virtual machines. From here lists of virtual machines and their location are created that are used for further processing. When these lists are submitted, simulation with CloudSim starts. A broker is created first, that runs the datacenters. Number of datacenters is assumed to be known in advance. On the basis of datacenters, clusters of VM`s are created. When each element is processed through simulation methods. In the CloudSim, DatacenterBroker class is the base class for any VM level event. When createVmsInDatacenter (int datacenterId) function is called, it performs the VM allocation to the respective datacenter.

VI. SIMULATED OUTCOMES

In the CloudSim simulator, we have performed the test in both cases: first with the existing allocation algorithm that is based on selecting host with minimum Pe`s and second with the proposed Dynamic VM allocation algorithm with clustering. For both the cases, we set the same parameters to perform the test such as:

- Number of VM`s: 15
- Number of Datacenters: 3
- Number of Cloudlets: 40
- Number of Hosts: 3

On comparing the simulation results, we have different simulation outcomes for both cases. Results show that proposed algorithm is having better performance than existing one. In the same case, load sharing among the available datacenters is better than the previous algorithm. Simulation results for both the cases are given below:

Test Perform with Existing VM Allocation Algorithm

Existing VM allocation algorithm select the host with minimum Pe`s and then allocate virtual machines to available datacenters. For the simulation test, we take 15 VM`s and 3 datacenters. Output window is shown in the figure 2:

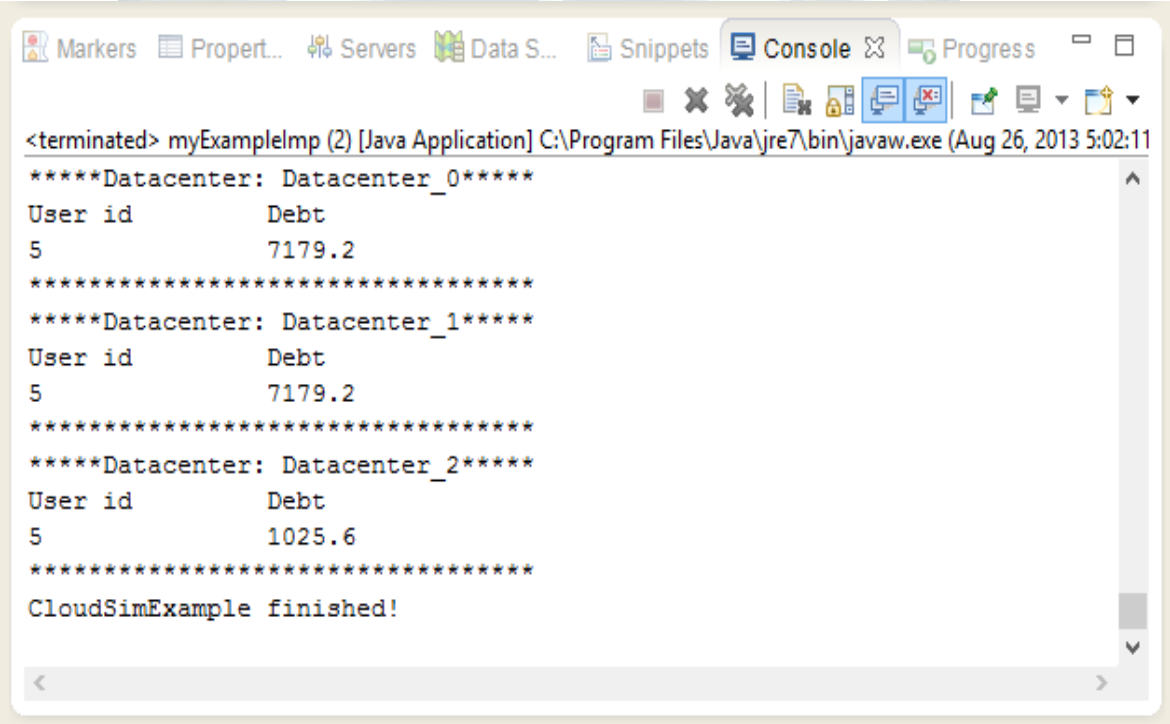

Fig 2: Test Perform with Existing VM Allocation Algorithm

As shown in the figure 2, all the VM`s are allocated the three datacenters. The selection of hosts in the datacenter is based on selecting minimum Pe`s. in Datacenter _0 and Datacenter_1 debt is same i.e. 7179.2, and in Datacenter_2 debt is 1025.6 as given in the figure. From these values, it is obtained that Datacenter_0 and Datacenter_1 are highly loaded and Datacenter_2 is underutilized. This inefficient utilization of datacenters leads to the performance degradation. Here the performance is

IPHV8I109X
International Journal Of Advanced Research and Innovation -Vol.8, Issue .I
ISSN Online: 2319 – 9253
Print: 2319 – 9245

measured in terms of simulation processing time.

Test Perform with Proposed Dynamic VM Allocation Algorithm with Clustering

Same input parameter is given for the proposed algorithm i.e. 15 VM`s and 3 datacenters, but the allocation strategy is different. So the result is also different from the existing algorithm. Proposed algorithm is based on dynamic allocation of VM`s, so a graphical interface is opened while performing the simulation. It takes number of VM`s and their allocation around the globe as shown in figure 3:

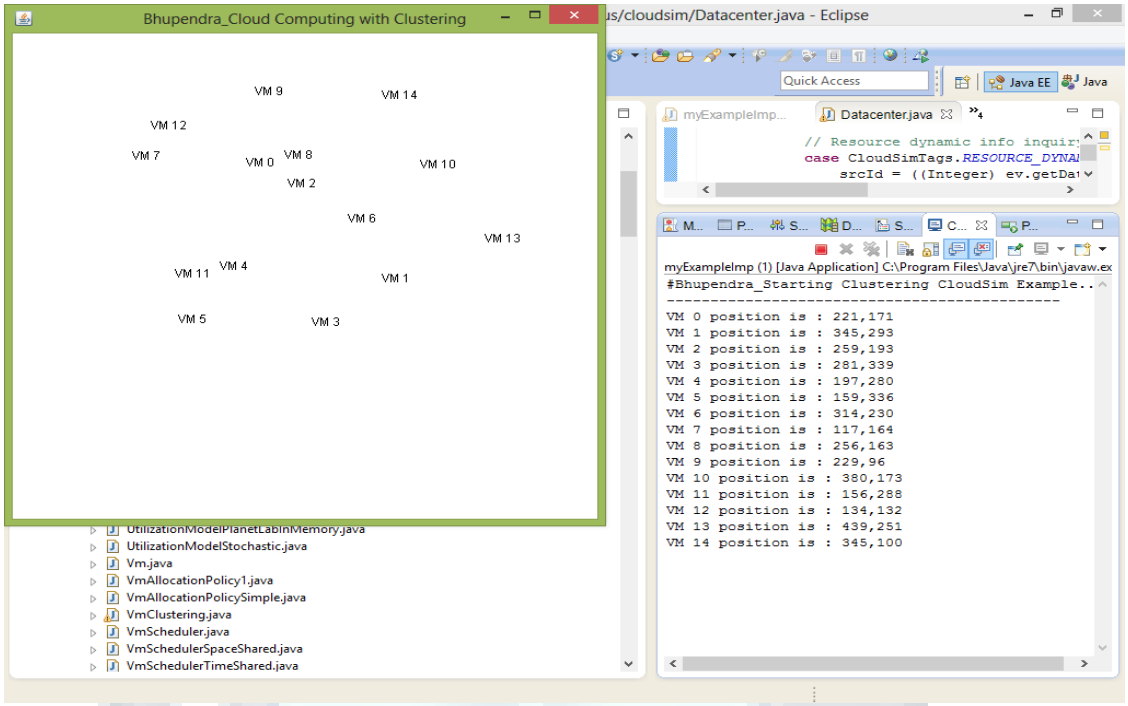

Fig3: VM location

After submitting the VM position, simulation takes place that perform the VM creation and their allocation to the available datacenters. For simulation, we assumed three datacenters so four clusters are created. And VM`s are allocated to the datacenters dynamically. This cluster formation is shown in the output window:

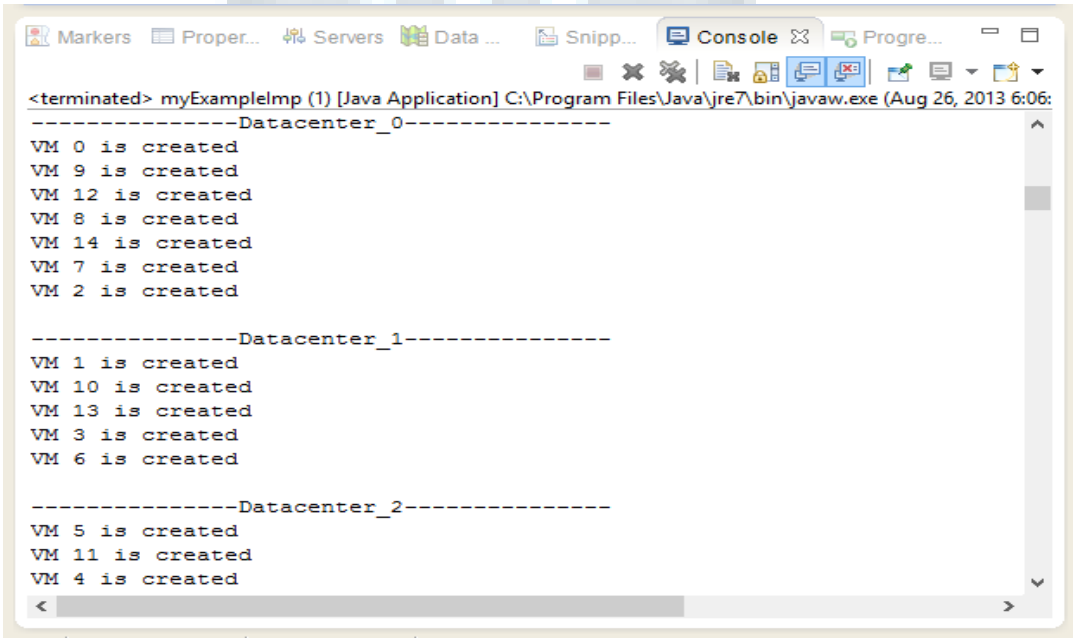

Fig4: Output window

As shown in the figure, VM`s created in datacenters are: •
VM 0, 9, 12, 8, 14, 7 and 2 in Datacenter_0
- VM 1, 10, 13, 3 and 6 in Datacenter_1 •
VM 5, 11 and 4 in Datacenter_2

And final output of the simulation is shown in the figure 5:

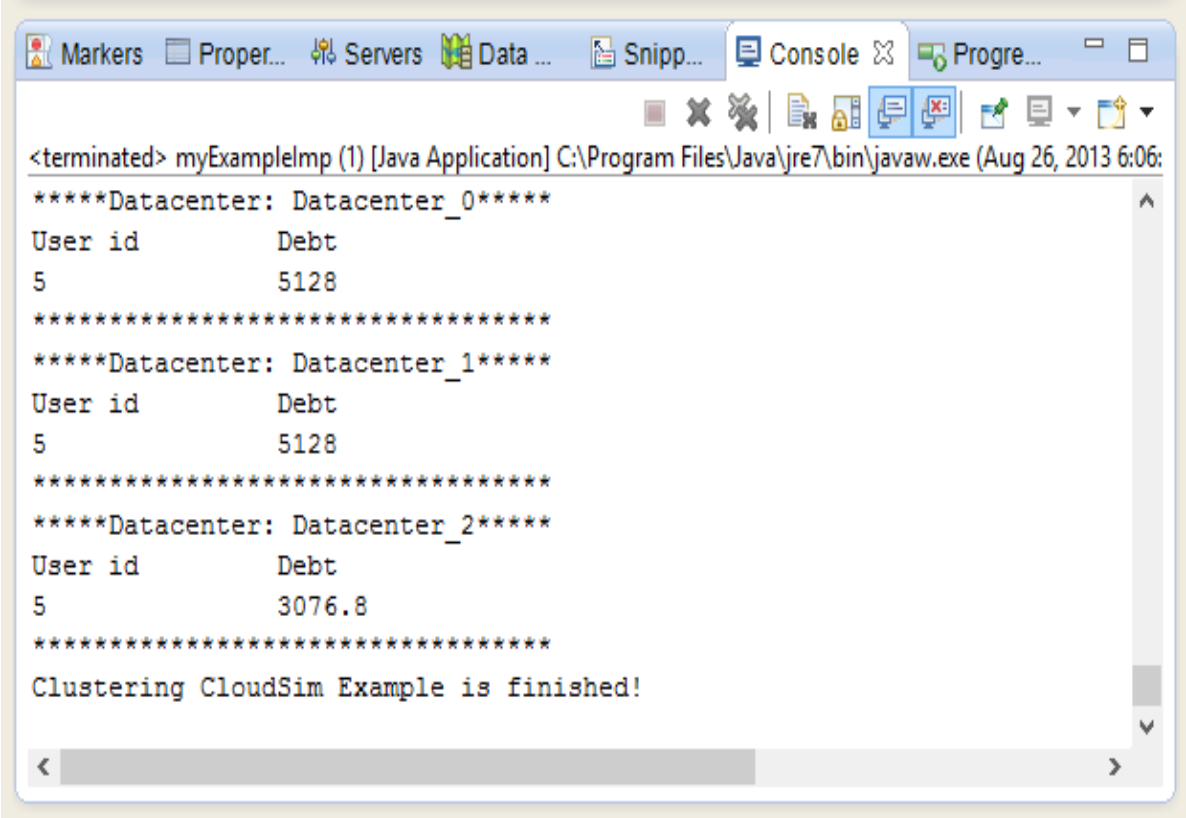

Fig5: Test Perform with Proposed Dynamic VM Allocation Algorithm with Clustering

As shown in figure 5, debt for Datacenter_0 and Datacenter_1 is 5128 and 3076.8 for Datacenter_2. VM`s allocation to the datacenters is done with clustering in this algorithm. So the clusters are responsible o calculate the debt for each datacenter.

VII. COMPARING THE RESULTS

Both algorithms are used to allocate the virtual machines, but allocation strategy in both the cases is different. Existing algorithm is based to select host with minimum Pe`s and proposed algorithm suggests dynamic VM allocation with clustering. When comparing the results of both algorithms, it is clear that proposed algorithm works well than existing one.

Results obtained from both algorithms are arranged in the table 1:

Table 1: Result comparison of both Algorithms

	Datacenter_0	Datacenter_1	Datacenter_2
Algorithm to select host with minimum Pe`s Dynamic VM	7179.2 debt	7179.2 debt	1025.6 debt
allocation with Clustering	5128 debt	5128 debt	3076.8 debt

According to the table, overall debt for each datacenter is minimum in the dynamic allocation policy and it also allows proper load sharing among the datacenters. Thus clustering technique helps to improve the performance and it also allows fast accessing and releasing of resources.

On comparing fig 2 and fig 5, it is obtained that datacenters are fully utilized in second case and having better debt outcomes than existing algorithm. So the performance is significant better than precious algorithms. On the basis of results in the table, a comparing graph is prepared that shows the clear different between the outcomes of each case. X axis represents the number of datacenters and Y axis is having debt values. This graph is structure with the same input parameter for both algorithms i.e. 15 VM`s, 3 Hosts and 3 Datacenters.he comparing output with these values is represented in the graph as shown in the Figure 6:

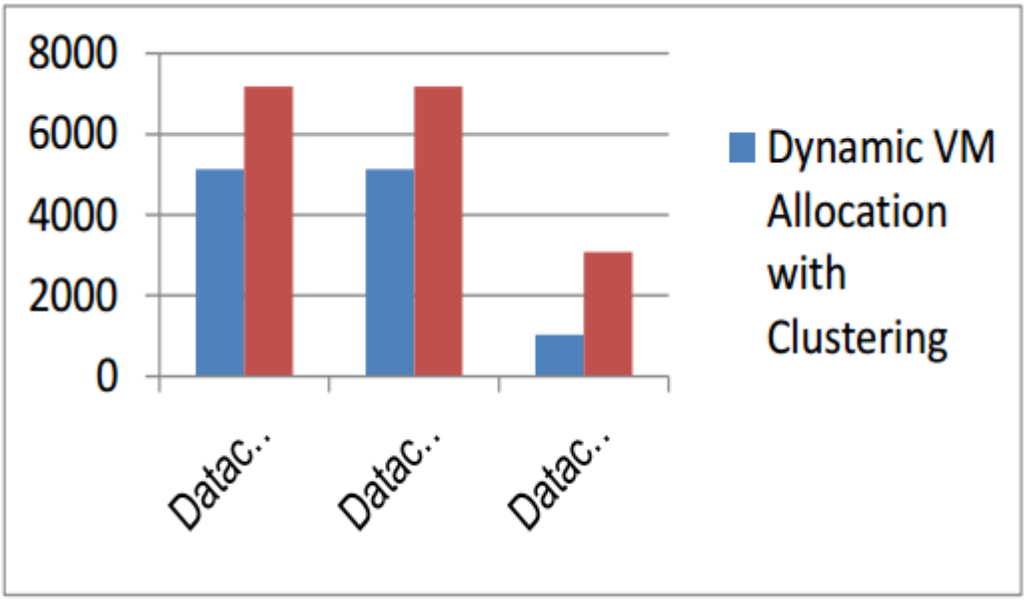

Fig6: Performance comparison of Proposed Algorithm with existing allocation Algorithm using same input parameter

As shown in the figure 6, Dynamic VM Allocation with Clustering Algorithm is having significant results in each datacenter as compared to the algorithm that select host with minimum Pe`s. For Datacenter_0, debt value in existing algorithm is 7179.2 but in the proposed allocation algorithm this value is 5128, which is smaller. Similar in the Datacenter_1, debt value is 7179.2 for existing algorithm which is larger than 5128 with proposed algorithm. And for Datacenter_2, proposed algorithm has 3076.8 debt value which is relatively larger than existing algorithm which has 1025.6 but quite efficient as shown in the graph.

Thus from the above result it is proved that the efficient load balancing and sharing of resources among the datacenters can be done with Clustering Algorithm. So proposed algorithm that uses clustering as an additional technique in the VM allocation is proved to be more efficient and helpful to enhance the performance of cloud computing.

VIII. CONCLUSION

As cloud computing itself is a bigger umbrella that is merged with different technologies. This complex integration is also having a bad impact on the performance of cloud computing. This work was aimed to improve the performance of cloud computing with Dynamic VM allocation Algorithm using Clustering. It arranges virtual machines in cluster form before allocating them to the datacenters. This arrangement provides efficient CPU utilization and load sharing among the datacenters, so performance can be enhanced in some aspects. The simulation work is done with CloudSim toolkit that was implemented on Eclipse Juno version. Simulated results are analysed and compared with existing methods. It is obtained from the figure that proposed algorithm is having better performance results than existing algorithm.

FUTURE WORK

It is reviewed in this work, that there are different scope to be researched in cloud computing. Large gaps are identified to be resolved. This work is performed to analysing performance parameters and its enhancement using clustering technology. Although integration of K-Means clustering works well than the existing methodologies in cloud environment but this also can be replaced with other available clustering techniques. This K-Means algorithm is having best results for large collection of data. As cloud datacenters are having huge amount of data so it is required to speed up the I/O processing. So clustering technology can also be implemented at datacenter level for “Data as a Service”.

Caching is also a solution for improving the performance. It allows fast accessing of data. Hence clustering and caching both can be used for cloud datacenters to make data clusters and allow data to be available in each cluster. In this way each cluster would have a copy of database so it would reduce the number of hits to datacenters and hence improve performance.

REFERENCES

[1] G. Malathy and Rm. Somasundaram, "Performance Enhancement in Cloud Computing using Reservation Cluster", European Journal of Scientific Research, ISSN 1450-216X, Vol. 86 No 3, September, 2012, pp.394-401

[2] Bhupendra Panchal and R. K. Kapoor, "Performance Enhancement of Cloud Computing with Clustering", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-5, June 2013

[3] Nikolaus Huber, Marcel von Quast, Micahal Hauck and Samuel Konev, "Evaluating and Modeling Virtualization Performance Overhead for Cloud Enviornments"

[4] A. Mahendiran, N. Saravanan, N. Venkata Subramanian and N. Sairam, "Implementation of K-Means Clustering in Cloud Computing Environment" , Research Journal of Applied Sciences, Engineering and Technology 4(10): 1391-1394, 2012, ISSN: 2040-7467

[5] Michael Shindler, Alex Wong and Adam Meyerson "Fast and Accurate k-means For Large Datasets"

[6] Pinal Salot, "A Survey of Various Scheduling Algorithm in Cloud Computing Environment" International Journal of research and engineering Technology (IJRET), ISSN: 2319 – 1163, Volume: 2 Issue: 2, Feb 2013

[7] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, César A. F. De Rose, and Rajkumar Buyya, "CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms"

[8] Qi Zhang, Lu Cheng and Raouf Boutaba, "Cloud computing: state-of-the-art and research challenges", Journal of Internet Serv Appl (2010) 1: 7–18

[9] Manisha Malhotra, "Simulation for Enhancing the Response and Processing Time of Datacenter Simulation", Volume 1, Issue: 3 Manuscript, 8 November 2011

[10] Navjot Kaur, Jaspreet Kaur Sahiwal, Navneet Kaur, "Efficient K-Means Clustering Algorithm using Ranking Method in Data Mining", International Journal of Advanced Research in Computer Engineering & Technology, Volume 1, Issue 3, ISSN: 2278 – 1323, May2012

[11] Prof. S.China Venkateswarlu, Prof. M.Arya Bhanu, Prof. Yudhaveer Katta and V.Badari, "Implementation of K-Means Clustering Algorithm Using Java", Global Journal of Computer Science and Technology, Volume 11, Issue 17, Version 1.0, October 2011

[12] Parul Agarwal,M. Afshar Alam and Ranjit Biswas, "Issues, Challenges and Tools of Clustering Algorithms", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 2, ISSN: 1694-0814, May 2011.

AUTHOR'S PROFILE:

[1]. P. MAHEEDHAR VARMA, M.Tech, CSE Dept, Amalapuram Institute of Management Sciences & College Of Engineering, Mummadivaram, AP.

[2]. **Dr. PENMETSA VAMSI KRISHNA RAJA:** He is presently working as a principal in Amalapuram Institute of Management Sciences and College of Engineering, Mummidivaram. He did his PhD from JNTU, Kakinada. He received his M.Tech (CST) from Andhra University, Visakhapatnam, Andhra Pradesh, India. His research areas include Computer Networks, Network Security, Cryptography, Data Mining and Software Engineering.