International Journal of Recent Trends in Engineering & Research
ISSN : 2455-1457 Impact Factor : 3.344

Dynamic Load Balancer For Web Session Management and Heterogeneous Resources In Cloud Computing Environment

Rohini Ramteke[1], Nikhil Meshram[2]
[1]*Department of Computer Technology,Shri Datta Meghe Polytechnic College, Nagpur-440014, India*
[2]*Dr.Ambedkar Institute of Management Studies and Research, Nagpur University, Nagpur-440014, India*

Abstract- Cloud computing means the dynamic allocation of resources and by which the business customers can increase or decrease the usage of their resource based on their demands. Through virtualization technology by multiplexing resource many of the laud gains in the cloud model can be achieves. A load balancer has the capabilities to calculate the load on individual virtual server and of efficient distribution of load from server to server to enhance the response delivery time to individual clients so that the performance of the system improves. In this paper, we are proposing a scheme to employ load balancing, green computing, virtualization and Cloud computing. By shifting multiple servers on improved and high capability single server we can save the energy and also give advantage of dynamic resource allocation to the cloud users. Hence, we are proposing the system to allocate data center resources dynamically by virtualization technology based on the demands of application and optimizing the number of servers in use to support green computing. We are trying to implement a set of solution to prevent the system from overload effectively while saving energy used.
Keywords- Cloud, virtualization, Green computing, load Balancer, OS**.**

I. INTRODUCTION

Any IT organizations can achieve the success by acquiring the resources on demand. For such touted gains Cloud computing is a promising technology to provide services on demand according to the client's requirements within a postulated time period. Further, the cloud computing environment provides the users for accessing the shared pool of distributed resources but cloud is a pay- go model where the consumers pay for the resources utilized instantly, which necessitates having highly available resources to service the requests on demand. Therefore, the resources management becomes a complex job from the business panorama of the cloud service Provider. Further, there can be a scenario where in the cloud service provider's datacenter will be hosting less number of Virtual Machines (VMs) compared to the number of jobs arrived for availing the service. In such a situation, similar types of jobs will be competing to acquire the same VM at the same time leading to a deadlock. Further, the problem of deadlock leads to the degradation of working performance as well as the business performance of the cloud service provider. Henceforth, an efficient load balancing technique is required to distribute the load to VM. To change the mapping between VMs and PMs during running of application is possible by migration technology. However, the question remains as to take decision for mapping so that the demands for the resources of VMs are met and the minimum number of PMs used. The challenge is that when needs of resources of VMs are different because of distinct set of application with which they run and may vary with time as the loads get increase and decrease. The PMs Capacity can also be different because of coexistence of different hardware in a data center. [1][5][6]

The two goals of our algorithm are as follows-
Overload avoidance: ThePM capacity should be sufficient to fulfill the needs of resource of all VMs running on it. Otherwise, the PM gets overloaded which affectthe performance of its VMs.
Green computing: To achieve green computing the utilization of number of PMs can be minimized as long as they can which still satisfy the needs of all VMs. The PMs which is in an idle mode can be turned off to save energy.
Cloud computing is a new technology which enables data processing and to deal with the large amount of data, which leading towards the improved performance of business. It includes various distributed virtual servers to accept user request for the services of different applications. The cloud is responsible to allocate resources to its clients on their demand with reduced cost and with high availability, scalability. One of the essential factors to increase the working performance of the cloud service provider is Load balancing [3].
The main objective of the system is to shift multiple servers on improved and high capability single server we can save the energy and also give advantage of dynamic resource allocation to the cloud users. Proposed system is Design and implements to build single machine VM cluster to provide a web session management load balancer.

II. LITERATURE REVIEW

ShinwonLee [6] introduces a novel load balancing algorithm for system on chip (SOC). Which evaluate the performance of all the resources available and the workload of requesting job-task dynamically that is based on premeasured response time, primitives power consumption and the frequency of current system. This algorithm shows the response time and the power consumption performance is 8% better than the general load balancing algorithm. With fewer premeasured information this algorithm cannot estimates the performance primitive of resources at all possible frequencies and load.

[1]Present a system which uses the virtualization technology for dynamic allocation of data center resources on the demands of application and also supports green computing by enhancing the number of servers in use. To achieve this objective the author has introduces the "skewness" algorithm to enumerate the unevenness utilization of multiple resources on the server. The system executes this algorithm periodically for the status evaluation of resource allocation on the basis of prediction of resource demands of VMs in future. It defines the server as hot spot and if any of its resource utilization is above a hot threshold then it indicates that the server is overloaded and hence the VMs running on it should be wander away.
Now a day in virtual machine load balancing is based on live migration and this strategy is based on measurement of the threshold that is, when the load on the host machine is more than the threshold then that machine get migrated immediately to avoid overhead.
[2] Presents a scheduling strategy on load balancing of VM resources based on genetic algorithm. According to historical data and current state of the system and through genetic algorithm, this strategy computes ahead the influence it will have on the system after the deployment of the needed VM resources and then chooses the least-affective solution, through which it achieves the best load balancing and reduces or avoids dynamic migration. This strategy solves the problem of load imbalance and high migration cost by traditional algorithms after scheduling. Experimental results prove that this method is able to realize load balancing and reasonable resources utilization both when system load is stable and variant. In cloud computing environment there is dynamic change in the VMs and because of that the computing cost on the virtualization software increases and some unpredicted load wastage.

KeYang [4] considers optimized control strategy which combines multi strategy mechanism with prediction mechanism. This strategy separately monitors the utilization of different resources component which include CPU, Memory, I/O and network bandwidth. He has consider different domain on which the decision of virtual machine taken such as light load domain, optimal domain, warning domain and overhead domain. By adopting different control strategy with prediction and multi migration strategy for different domain avoids the unnecessary cost caused by the immediate increase in the utilization of resources which enhances the system performance and save the network resources. Frequent changes in the memory status causes the repeated transmission of memory pages many times which cost host resources. Cost about the virtual disk migration has not considered. Only fewer amounts of data can be transmitted for the migration.

[5] Present a dynamic load balance strategy between multiple volume servers which characterize machine capability and load quantity with black box modeling approach, and implement the load balance strategy based on genetic algorithm.

This strategy improves the I/O performance of the system 1.7 times better than the existing technique. According to the user need and the application characteristics logical volumes could not migrate between storage devices and not consider the object-based storage mechanism to manage logical volumes on object level.

III. PROPOSED RESEARCH METHODOLOGY

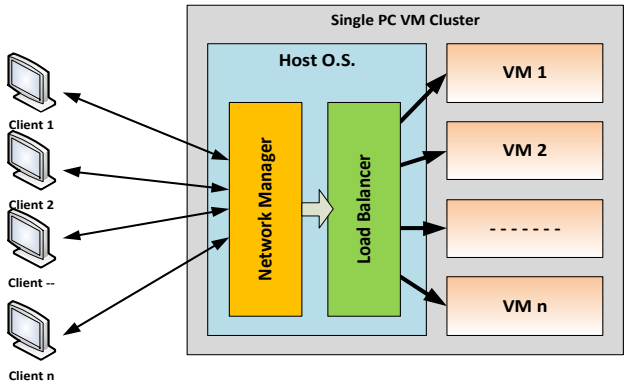

Fig1: System architecture

Inproposed system we are implementing the load balancer along with the concept of cloud computing.The system consists of a host machine with four VMs as a guest machine which will act as a windows based web server and network manager to monitor and accept web traffic from all client.

Load balancing can be done in two ways, first it give a large amount of simultaneous accesses or web traffic to multiple server respectively to minimize the time of response for which the user waits. Secondly it has to distribute the load to the multiple web based server instead of single server to advance the utilization of resources of each server.The load balancer allocates the resources to the requested application of client dynamically and on symmetry basis, for this it moniter the CPU utilization,memory status and status of I/O information and it also need to allocate requested client web traffic.

Different application required different allocation strategy like sites of e-commerce requires idle CPU node because it has large amount of calculation where as some other application like database which performs read-write operation frequently required a I/O idle node.On various layer load balancing can be performed.

The main objective of developing a load balancer is to improvethe system resource utilization in order to enhancing the throughput of system and to reduce the requested application response time.The load balancing has two strategies:

1) *Session based*: The client can generate concurrent connection for various application.The session exactly means a state.From single session multiple session can be generated and also from individual session multiple task can be generated.The network manager moniters the various session from clients and distributes the resources to them.
2) *Load based*: when a client performs some read-write operation or calculation which required lots of CPU and I/O utilization, the network manager identifies the operation and its requirement and allocates the resources to the client.

Development Phases

Step 1: Memory statistics Module

Develop a memory statistics in dot net. In this module, the restoring of virtual machine takes place.In this load balancer can evaluate the utilization of CPU based on which it will distribute the load among various virtual machines.

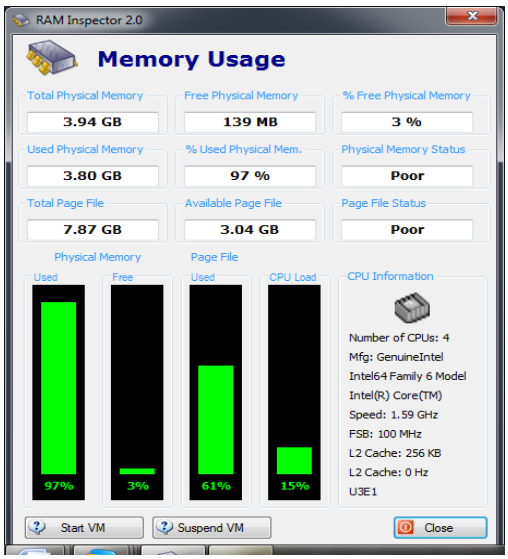

Fig 2: memory utilization after VM installation

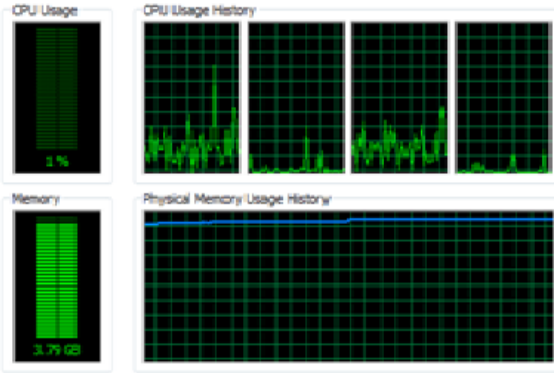

Fig 3: graphical analysis of memory utilization

TABLE I memory utilization in percentage

State of OS	percentage of Physical memory used	percentage of Physical memory free	percentage of Page File used	percentage of CPU load
OS with VM	45%	55%	31%	3%
OS with VM initialization	65%	35%	37%	18%
OS with VM and some application	76%	24%	62%	34%
OS with susspend state of VM	49%	51%	31%	4%

As from the above result we can see that the CPU load is very low until the VM start as soon as VM starts CPU load get increses. Up to certain limit when vm get overloaded their is a provision for suspention of vm to reduce CPU load.

Step 2: VM Cluster Module

Develop a host machine consisting of few virtual machines (VMs) and a network manager as head virtual machine which manages request from client according to client requirement.

Step 3: VM Web Based Server

All virtual machine will acts as a windows based web server to provide services to the client.At one time only two VM web server will activated and when the load will increases, one out of two server will get suspended and another server will allocated to client with low load.

Y=N+1 …….. (i)

where y= m/c required for distribution ofload

N= number of required VM+1----------X

X= reserved VM which is fully dedicated

host m/c

Step 4: Host Program

The network manager will consist of host program to moniter web session and their requirement from number of client to allocate resources.

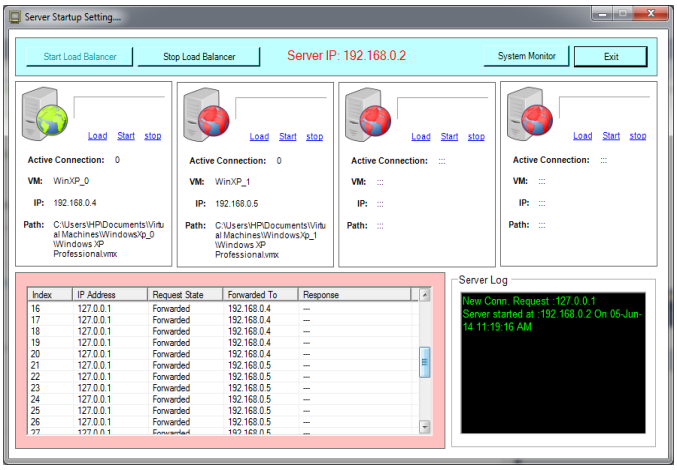

Fig 4: distribution of traffic by load balancer

The server IP address is used by the client to connect with server, the N number of user can be connected to the server and this count of connected client are shown at the server side. When some task is allocated to the server, it can distribute it to the N number of clients and can also process it locally as well as parallel. In parallel process all client process the equally distributed load simultaneously and in local process all the task are process by the server itself.

X=S+Ni……. (ii)

S=server

Ni=number of clients connected to server

Distribution of load =$\frac{\text{Number of task}}{\text{Ni}}$… (iii)

Step 5: Web Traffic from client

The network manager will have a software to accept web traffic from various number of client. The client can generate various sessions as web traffic.

Step 6: Random session generator program

The random session generator program will generate number of session for various client for allocation of resouces. Session is a state and consists of multiple task.

IV. CONCLUSION

In the existing algorithm there exists a communication between the Load Balancer and the Data Center Controller for updating the index table leading to an overhead. Hence, we are proposing the system that uses virtualization technology to allocate resources of datacenter dynamically based on demands of application and support green computing by enhancing the number of servers in use .By shifting multiple servers on improved and high capability single server we can save the energy and also give advantage of dynamic resource allocations the cloud users. We will implement a set of solution to prevent the system from overload effectively while saving energy used.

V. EXPECTED OUTCOME

The expected outcome of the system is to save more energy and resources and providing cloud computing advantages to applications by establishing the single machine VM cluster and developing the load balancer for VM web server and all together should provide better web session management and dynamic resource allocation with minimum power utilization by controlling the VM state.

REFERENCES

i. Zhen Xiao, Senior Member, IEEE, Weijia Song, and Qi Chen "*Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment*" IEEE Transactions on parallel and distributed systems, VOL. 24, NO. 6, JUNE 2013.

ii. Jinhua Hu, JianhuaGu, Guofei Sun, Tianhai Zhao "*A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment*" 3rd international Symposium on Parallel Architectures, Algorithms and Programming.

iii. Rashmi. K. S, Suma. V, Vaidehi. M "*Enhanced Load Balancing Approach to Avoid Deadlocks in Cloud*" on Advanced Computing and Communication Technologies for HPCApplications - ACCTHPCA, June 2012.

iv. Ke Yang, JianhuaGu,TianhaiZhao,Guofei Sun "*An Optimized Control Strategy for Load Balancing based on Live Migration of Virtual Machine*" 2011 Sixth Annual China Grid Conference.

v. Li Zhou, Yi-Cheng Wang, Ji-Lin Zhang, Jian Wan, Yong-JianRen "*Optimize Block-Level Cloud Storage System With Load-Balance Strategy*" 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum.

vi. Shinwon Lee, V.Meka, MinguJeon, Nagoo Sung, JeongnamYoun "*Dynamic load balancing for system onchip*" IEEE Transactions on parallel and distributed systems 2012.

vii. Singh, M. Korupolu, and D. Mohapatra, "*Server-Storage Virtualization: Integration and Load Balancing in Data Centers*," Proc. ACM/IEEE Conf. Supercomputing, 2008.

viii. A. Singh, M. Korupolu, and D. Mohapatra,"*Server-Storage Virtualization: Integration and Load Balancing in Data Centers*" Proc. ACM/IEEE Conf. Supercomputing, 2008.

ix. Y. Agarwal, S.Savage and R. Gupta "*sleep server: A software- only approach for reducing of energy consumption of PCS with in enterprise environment*" Proc. USENIX Ann. TechnicalConf., 2010

