

Allocation Of Virtual Machines In Cloud Computing Using Load Balancing Algorithm

MR.Manan D. Shah*

Information Technology
Dharmsinh Desai University
Nadiad, India.
Mananshah0003@gmail.com

MR.Amit A. Kariyani

Information Technology
Dharmsinh Desai University
Nadiad, India
kariyani_amit@yahoo.in

MR.Dipak L. Agrawal

Information Technology
Dharmsinh Desai University
Nadiad, India
Agrawaldipak45@yahoo.com

Abstract- Cloud computing is a fast growing area in computing research and industry today. Three main services provided by the cloud are IaaS, SaaS, and PaaS. With the advancement of the Cloud, there are new possibilities opening up on how applications can be built and how different services can be offered to the end user through Virtualization, on the internet. There are the cloud service providers who provide large scaled computing infrastructure defined on usage, and provide the infrastructure services in a very flexible manner. The establishment of an effective load balancing algorithm and how to use Cloud computing resources efficiently for effective and efficient cloud computing is one of the Cloud computing service provider's ultimate goals. In this paper firstly analysis of different Virtual Machine (VM) load balancing algorithms is carried out. Secondly, a modification to the VM load balancing algorithm has been done and implemented for an IaaS framework in Simulated cloud computing environment; i.e. 'Throttled Load Balancing Algorithm' using CloudSim tools, for the Datacenter to effectively load balance requests between the available virtual machines, in order to achieve better performance parameters such as response time and Data processing time.

Keywords-Load Balancing Algorithms, Virtual Machine, Data Center Controller, Virtualization.

I. INTRODUCTION

Cloud computing is an on demand service in which infrastructure, platform and software are provided on demand according to the client's requirement at specific time. It's a term generally used in the case of internet. One can view whole internet as a cloud. Thus the all the above mentioned services are access by a user as a client to the cloud. Now as the basic idea of cloud computing is to provide resources such as VMs as services on demand. Allocating efficient VM on demand is being carried out with the help of the load balancing algorithms in the cloud computing. As the load balancing algorithm plays an important role while deciding which VM is to be allocated on demand of the user. While providing services it is possible to have a number of requests at a time and due to that some requestors need to remain in queue though they have possibility to send request to other service provider. Thus with the help of

the load balancing algorithm user will able to decide whether they need to remain in the queue or get service from the other service provider.

Numbers of the algorithms for the load balancing in the cloud computing are available for allocating the efficient VMs. Among such available algorithm which is to be used is the main decision is to be taken. Some of those algorithms have been discussed in this paper. Thus for having accurate usage of resources and being Faithfull with all the resources, concept of load balancing is being carried out.

II. EXISTING LOAD BALANCING ALGORITHMS

A. Round Robin Load Balancer

This algorithm works on random selection of the virtual machines. The datacenter 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 DataCenter controller assigns the requests in a circular order. Once the VM is assigned the request, the VM is moved to the end of the list. One more way to define round robin algorithm is a better allocation concept known as Weighted Round Robin Allocation in which one can assign a weight to each VM so that if one VM is capable of handling twice as much load as the other, the powerful server gets a weight of 2. In such cases, the DataCenter Controller will assign two requests to the powerful VM for each request assigned to a weaker one. The major issue in this allocation is this that it does not consider the advanced load balancing requirements such as processing times for each individual requests.

B. Equally Spread Current Execution Algorithm

Equally spread current execution algorithm process handle with priorities. it distribute the load randomly by checking the size and transfer the load to that virtual machine which is lightly loaded or handle that task easy and take less time , and give maximize throughput. It is spread spectrum technique in which the load balancer spread the load of the job in hand into multiple virtual machines [5].

C. Active Monitoring Load Balancer

Active VM Load Balancer 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. Active VM Load Balancer returns the VM id to the Data Center Controller the data Center Controller sends the request to the VM identified by that id. Data Center Controller notifies the Active VM Load Balancer of the new allocation [8].

D. Throttled Load Balancer

Throttled algorithm is completely based on virtual machine. In this client first requesting the load balancer to check the right virtual machine which access that load easily and perform the operations which is give by the client or user. In this algorithm the client first requests the load balancer to find a suitable Virtual Machine to perform the required operation [8].

III. PRAPOSED WORK

A. Modified Throttled Load Balancing Algorithm

Giving some modification to this available algorithm we can give flexibility to the user for getting services from the cloud provider. As we know that while providing resources as services it is possible to have a number of request at a same time and due to that some requestor need to remain in the queue though they have possibility to send request to other service provider. Thus to overcome such situations we can modify some existing algorithm and make it available to user to decide whether they need to be in queue or can have a service from the other cloud provider. We have decided to did the modification to the throttled load balancer as it is founded by some experts of cloud computing that throttled load balancer is much better as compare to the other available load balancing algorithms.

The Throttled Load balancing algorithm is divided into three parts. The first phase is the initialization phase. In the first phase, the expected response time of each VM is to be found. In second Phase find the efficient VM. Last Phase return the ID of efficient VM.

B. Flow of Throttled Load Balancing Algorithm

- Efficient algorithms find expected response time of each Virtual Machine.(expected response time find with the help of resource info program).
- When a request to allocate a new VM from the Data Center Controller arrives, Algorithms find the most efficient VM (efficient VM having least loaded, minimum expected response time) for allocation
- Efficient algorithms return the id of the efficient VM to the Datacenter Controller.
- Datacenter Controller notifies the new allocation

- Propose algorithm updates the allocation table increasing the allocations count for That VM.
- When the VM finishes processing the request, and the Data Center Controller receives the Response. Data center controller notifies the efficient algorithm for the VM de-allocation .Continue From Step 2.

Algorithm is to find the expected Response Time of each Virtual Machine; the expected response time can be finding with the help of the following formulas

$$\text{Response Time} = \text{Fint} - \text{Arrt} + \text{TDelay} \dots 1$$

Where Arrt = arrival time of user request,

Fint = finish time of user request and the transmission delay can be determined using the following formulas

$$\text{TDelay} = \text{Tlatency} + \text{Ttransfer} \dots \dots \dots 2$$

Where TDelay = transmission delay, Tlatency = network latency, T transfer = time taken to transfer the size of data of a single request (D) from source location to destination.

$$\text{Ttransfer} = D / \text{Bwperuser} \dots \dots \dots 3$$

$$\text{Bwperuser} = \text{Bwtotal} / \text{Nr} \dots \dots \dots 4$$

Where Bwtotal = total available bandwidth and Nr = number of user requests currently in transmission [5].

With the help of the above available formulas we can able to calculate response time of the virtual machines and after getting the response time algorithm will able to decide which is an efficient VM for the allocation. Moreover after performing above mentioned steps we can disclose the availability of the VMs to the user so that they can came to know whether they need to remain in queue or should get service from the other cloud service provider.

IV. SIMULATORS

The main aim of simulator is to test the implementation work in the absence of the required environment. Thus in the cloud environment two simulator are used CloudSim and Vcloud. CloudSim is the open source. Some simulators available for the distributed field such as SimGrid, GridSim, etc such simulators are not valid for the cloud computing as the cloud environment having multiple layers while SimGrid and GridSim are made for the single layer environment.

A. CloudSim

CloudSim is a new generalized and extensible simulation framework that enables seamless modeling, simulation, experimentation of emerging Cloud computing infrastructures and management services. The simulation framework has the following novel features: (i) support for modeling and instantiation of large scale Cloud computing infrastructure, including data centers on a single physical computing node and java virtual machine; (ii) a self-contained platform for modeling data centers, service brokers, scheduling, and allocations policies; (iii) availability of virtualization engine, which aids in creation and management of multiple, independent, and co-hosted virtualized services on a data center node; and (iv) flexibility to switch between space-shared and time-shared allocation of processing cores to virtualized services.

B. VirtualCloud

Cloud computing provides opportunity to dynamically scale the computing resources for applications. These Resources are shared among customers using virtualization technology. Using these resources efficiently is an open challenge. Since, cloud computing consists of large number of resources, testing these new policies on real world is time consuming and difficult. To ease the problem of modeling and testing policies, Virtual Cloud is being proposed, for cloud computing environment. Virtual Cloud helps developers to model and test, their policies to utilize the cloud computing resources efficiently. Developed as multi-layered architecture, this simulator helps to test new approaches, find the bottlenecks before implementing in real world cloud computing environment.

We can use cloudsim for the implementation work as it is open source and much beneficial for our research work.

V. CONCLUSION

As the main purpose of cloud computing is to provide services to the client on demand. Thus to have a better services, we can conclude that the waiting situation might took place in future, as the cloud computing being the most useful computing field now a days. The issue of disclosing the availability of VMs to the client will improve the performance level of the cloud computing. After deciding the better or efficient load balancing algorithm, which will be helpful for allocation of efficient VMs on demand. Moreover the concept of disclosing the availability of VMs to the client will be helpful for the client for getting services without being in queue. Moreover when the client came to know about the free available VMs they can able to decide whether they should take service from that cloud or by some other cloud service provider. Thus while caring out these issue we can able to have a better service from the cloud computing. The implementation of the algorithm will be carried out in the future work.

VI. REFERENCES

1. Zhang, Q., Cheng, L., & Boutaba, R. (2010). Cloud computing: state-of-the-art and research challenges. *Journal of Internet Services and Applications*, 1(1), 7–18. doi:10.1007/s13174-010-0007-6

2. Marios D. Dikaiakos and George Pallis (University of Cyprus), Dimitrios Katsaros (University of Thessaly), Pankaj Mehra (Hewlett-Packard Labs) and Athena Vakali (Aristotle University of Thessaloniki), Cloud Computing Distributed Internet Computing for IT and Scientific Research.

3. Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., et al. (2009). Above the Clouds : A Berkeley View of Cloud Computing Cloud Computing : An Old Idea Whose Time Has (Finally) Come, 7–13.

4. Xu, Y., Wu, L., Guo, L., & Chen, Z. (2011). An Intelligent Load Balancing Algorithm Towards Efficient Cloud Computing, 27–32.

5. Raj, G. (2012). Comparative Analysis of Load Balancing Algorithms in Cloud Computing, 1(3), 120–124.

6. Cse, M. T. (2012). Comparison of load balancing algorithms in a Cloud Jaspreet kaur, 2(3), 1169–1173.

7. Padhy, R. P., & Rao, P. G. P. (2011). LOAD BALANCING IN CLOUD LOAD BALANCING IN CLOUD Dr . Pabitra Mohan Khilar.

8. Sharma, M., & Sharma, P. (2012). Efficient Load Balancing Algorithm in VM Cloud Environment, 8491, 439–441. -43- M.Tech. Dissertation, Information Technology Department, Dharamsinh Desai University

9. Galloway, J. M., Smith, K. L., & Vrbsky, S. S. (2011). Power Aware Load Balancing for Cloud Computing, I.

10. Das, K. (2010). VirtualCloud - A Cloud Environment Simulator, (October).

11. Sharma, P. M. (2012). Performance Evaluation of Adaptive Virtual Machine Load Balancing Algorithm, 3(2), 86–88.

12. Calheiros, R. N., Ranjan, R., Rose, C. A. F. De, & Buyya, R. (n.d.). CloudSim : A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services, 1–9.

13. Mell, P., Grance, T., & Grance, T. (n.d.). The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology.

14. Tsai, W.-T., Sun, X., & Balasooriya, J. (2010). Service-Oriented Cloud Computing Architecture. *2010 Seventh International Conference on Information Technology: New Generations*, 684–689. doi:10.1109/ITNG.2010.214

15. Verma A, Gautam K, Koller R, the Cost of Reconfiguration in a Cloud. IBM Research – India, IIT Kharagpur, Florida International University.

16. Ye, K., Jiang, X., Huang, D., Chen, J., & Wang, B. (2011). Live Migration of Multiple Virtual Machines with Resource Reservation in Cloud Computing Environments. *2011 IEEE 4th International Conference on Cloud Computing*, 267–274. doi:10.1109/CLOUD.2011.69

17. Verma, A., Kumar, G., Koller, R., & Sen, A. (2011). CosMig: Modeling the Impact of Reconfiguration in a Cloud. *2011 IEEE 19th Annual International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems*, 3–11. doi:10.1109/MASCOTS.2011.37

AUTHORS PROFILE

MR.Manan D. Shah, B.E., H.N.G.U., Patan, Gujarat, India. Presently pursuing Dissertation of M.TECH I.T, Dharamsinh Desai University, Nadiad, Gujarat, India.Email:-mananshah0003@gmail.com.

MR.Amit A. Kariyani, B.E., Dharamsinh Desai University, Nadiad, Gujarat, India. Presently pursuing Dissertation of M.TECH I.T, Dharamsinh Desai University, Nadiad, Gujarat, India. Email:- kariyani_amit@yahoo.in_

MR.Dipak L. Agrawal, B.E., S.V.I.T., Vasad, Gujarat, India. Presently pursuing Dissertation of M.TECH I.T, Dharamsinh Desai University, Nadiad, Gujarat, India.Email:-agrawaldipak45@yahoo.com.