IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 6, Ver. VI (Nov.-Dec. 2016), PP 85-89
www.iosrjournals.org

An Efficient Load Balancing Approach in a Cloud Computing Platform

Saptarshi Bhowmik[1], Sudipa Biswas [2], Karan Vishwakarma[3], Subhankar Chattoraj[4]

[1]*Department of Computer Science, Jadavpur University, India*
[2]*Department of software engineering, BITS pilani, India*
[3]*Techno India University, Salt Lake EM/4, Kolkata, India*
[4]*Department of electronic and communication engineering, Techno India University, Salt Lake EM/4, Kolkata, India*

Abstract: *Technology has always played an important role changing human lives and helping to perform faster and easier job. In recent years cloud computing has replaced the traditional method of grid computing, cluster computing and introduced virtualization as the key technology in cloud computing. It has become one of the most demanding services due to its high reliability, efficiency and cost effectiveness. But cloud computing faces certain major issue which involves efficient load balancing and virtual machine (VM) arrangement. In this paper we have proposed a priority based virtual machine (VM) arrangement and load balancing using round robin scheduling. The VM is arranged by comparing the vital parameters which involve bandwidth, RAM, and MIPS (Million instructions per second). We have also proposed a new algorithm based on round robin for load balancing so that the jobs are not scheduled according to the VM priority and hence no VM is inundate or remains inactive for a longer period of time. This new method will eliminate the process of excess burden on any individual VM and enhance the resource utilization. We have used Cloudsim as a simulation toll to simulate the algorithms.*

Keywords: *Virtual machine, Cloudsim, load distribution, Resource scheduling, Priority based scheduling*

I. Introduction

The virtualization technology that enhance the internet access connectivity enabling all the user to access the resources according to their need rather than installing application on every single system is described as cloud computing. Cloud computing [1-3] enhances user to access the resources virtually. It can be used in networking, data transfer, access and secure communication between the users. It not only enables the user to access the data from anywhere in the globe but also enables the user to access the storage without being specifically informed about the details capacity of the storage. It focuses on advancement of software and application development. Usually it is offered as pay per basis where the users pay according to their uses. An overview of cloud computing is given below in Fig1.

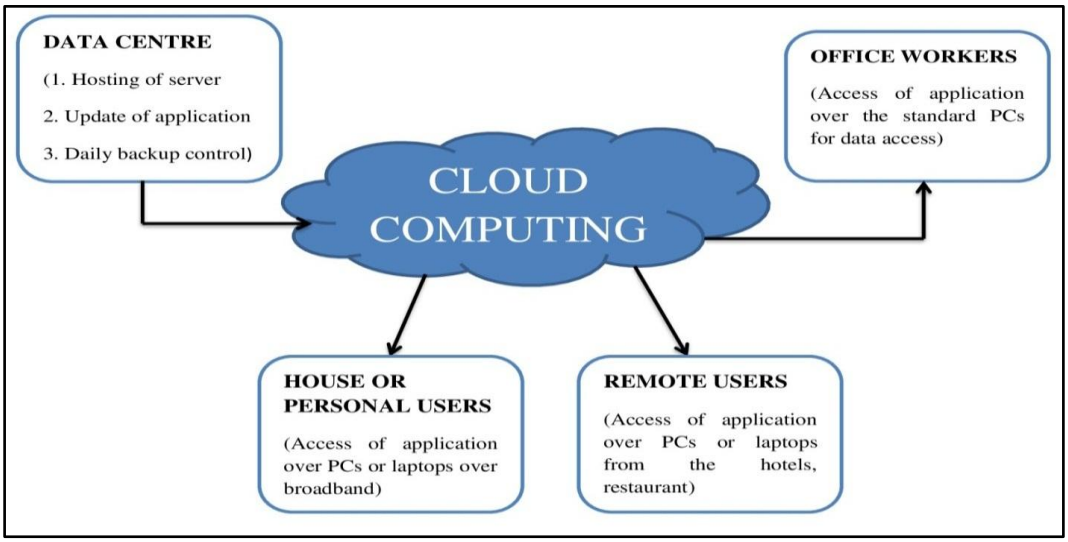

Fig1: Cloud computing overview

Cloud computing is offered mainly in three modes, firstly the infrastructure as a service (IaaS), secondly the platform as a service (PaaS), thirdly software as a service (SaaS). The IaaS model is the provider host which provides the user with storage, hardware, software, and other components. It may also include on basis of the policy assigned with the user such as virtualizations of desktop and some administrative tasks. IaaS [4-5] providers include Amazon Web Service (AWS), Google Compute Engine etc. The PaaS [6] provides platform for developing application. PaaS makes the development of application, testing error free, rapidly and cost effective. The SaaS [5, 11] enables services according to the requirement of the user. It is one of the first entities to utilize the resources and distribute the system accordingly. Below the in Fig 1.2 the cloud computing service pyramids structure in given.

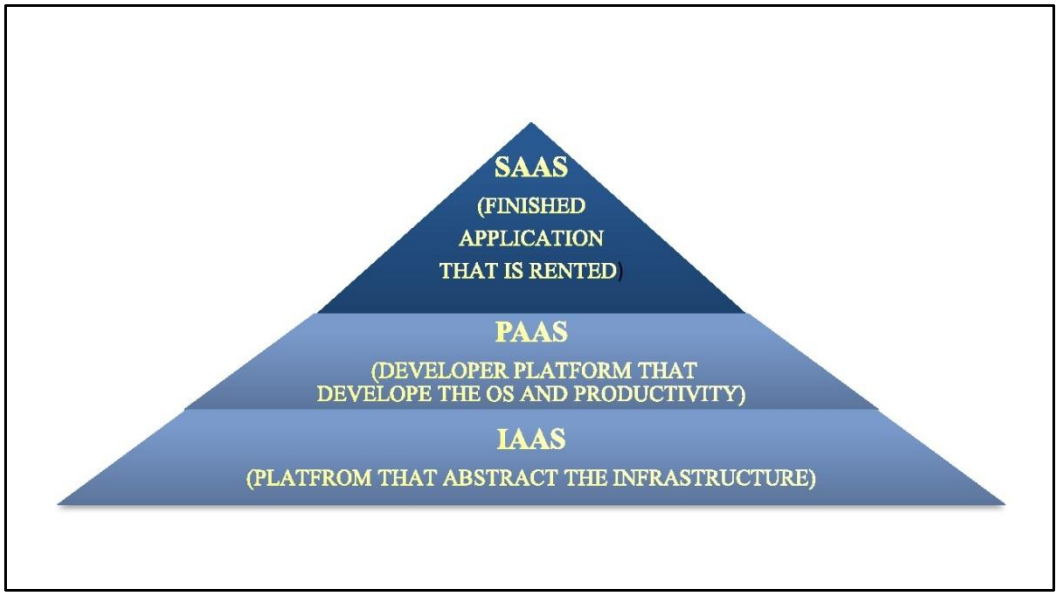

Fig1.2: Cloud computing services

II. Related Works

1. Throttled Load Balancer

It is a pre allocated [7-8] defined to a single VM at a specific time. If there was numerous number of request available at a single time than the job has to be waited in a queue till the next VM is available. This process has a number of disadvantages the request has to be waited so it is a time consuming process. The user has to wait till the next VM is available. This makes this system for costly than the proposed system.

2. Active Monitoring Load Balancer (AMLB)

Active monitoring load balancer [9-10] maintains the information about each VM. The job is first identified in load balancer then the VM identified by the balancer is allocated for the job. When there is one VM the balancer first identify the VM then the jobs are allocated accordingly. This process of VM monitoring by the load balancer makes the process more time consuming.

III. Proposed Architecture

1. Virtual Machine Arrangement based on priority

Virtual machines are resemblances of physical machine. They represent a computer system based on one specified computer architectures. Virtual machines functions same as physical machines. Here we are arranging assigning priority to individual virtual machines. The priority assigned will be calculated based on the virtual machine specifications. For this particular project we will be depend on the bandwidth capacity of the virtual machine, ram configuration, processing elements associated with the virtual machine along with the MIPS (million instruction per second) rate of the processing elements associated with the virtual machine to calculate the priority. Once this is completed, we will be arranging the virtual machine based on their priority in a circular linked list. For example, if virtual machine A has a priority value higher than virtual machine B, then we will keep the node of the circular linked list on virtual machine A and virtual machine B will be next to virtual machine B, such that when the node traverses the circular linked list, it first visits node a and then visits node b.

- Bandwidth (BW) ==X (1)
- Total Million Instruction Per Second (MIPS) == MT (2)
- Available Million Instructions Per Second (MIPS) == MA (3)
- Number of Processing element (PE) == P (4)
- Capacity of RAM (R) == R (5)

We will be calculating the priority of a virtual machine based on the above parameters. Many a times, the priority of a virtual machine is calculated based on any one parameter, like bandwidth, or the processing elements allocated to it from the host. But here we are considering the above parameters, which encompass all the basic elements of required for performing a job. According to our experiment, we are assigning fifty percent weight to the Bandwidth allocated to the virtual machine, sixty percent on the capacity of the RAM, since majorly jobs will need primary memory in real world, we have kept the weight of the computational capacity seventy percent, since in actual cases each and every jobs need to be scheduled by the scheduler to get processed, and the capacity of processing determines how fast a job will be completed. Therefore computational capacity is given the highest weight.

Order of weight

Bandwidth < memory capacity < processing capacity

For calculating the computational power of a virtual machine, we are taking the ratio of free MIPS available for use by the processor to the total MIPS associated with the processor, if the amount of free MIPS available is zero, that means that the ratio becomes zero, in such a case processing capacity becomes zero, and the virtual machine gets a lower priority when it comes to assigning the tasks. After calculating the ratio, we are multiplying the ratio with the number of Processing Elements allocated to the virtual machine by the host. This gives the actual processing power of the virtual machine, which can be harnessed for performing a task. After which we are taking a weighted average of the three components.

Formula:-

$$\frac{\{(50/100)\ X + (60/100)\ R + (70/100)\ (P * (MA/MT))\}}{1.8}$$

2. Round Robin Virtual Machine Load Balancing

When the load is composed to the virtual machine it will automatically choose the VM according to the priority hence the performance of the VM will get affected simultaneously. A single VM when over composed of a large load the performance capacity gets reduced. In order to improvise this system round robin algorithm is proposed as it does not work on the VM priority but it balances the load on all the VM available. In Fig 1.3 it illustrates that the cloud controller is the point where the load is entered and the CLC is responsible for various features like servers, storage and network control. It also handles the public system and protocol translation simultaneously. The broker policy handles the traffic between user in the system and the data center. Below in Fig1.3 the working process is given.

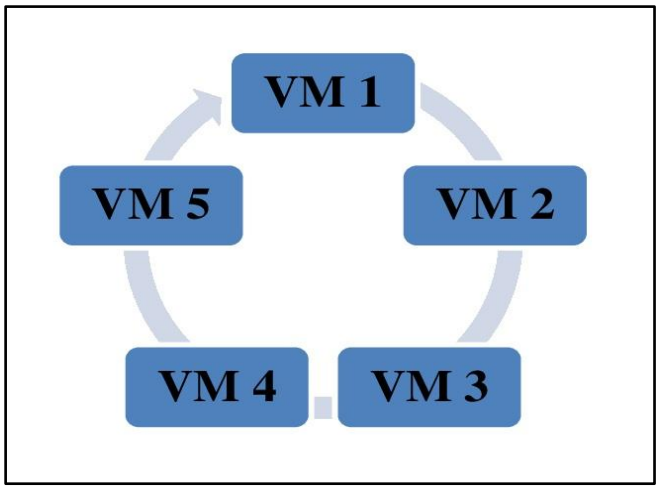

Fig1.3: The working process diagram

IV. Experimental Evaluation

The proposed algorithm is implemented in java platform for load balancing. For experimental simulation Cloudsim and its tool are used. We assumed the Cloudsim has 6 data virtual machine (VM) and 1 data center where the values of individual parameters are describe in table1 and in table 2 Response time for priority based round robin balancing and in table 3 the data operational time are also discussed.
We are using inbuilt APIs present in Cloudsim for extracting the data about the virtual machine.
APIs used:

- **getBW()** for getting the information of amount of bandwidth that is allocated.
- **getNumberofPes()** for getting the information about amount of processing element allocated to each virtual machine.
- **getRam()** for getting the information about amount of RAM allocated to each virtual machine.
- **getMips()** for getting the information about total MIPS allocated to the processing element, associated with each virtual machine.
- **getAvailableMips()** for getting the information about free MIPS available out of the total MIPS allocated to the processing element, associated with each virtual machine.

Table1: Parameters

Parameter	Values
Data center	X86
Operating system	Linux

Table1: Response time for priority based round robin balancing

No of VM	AVERAGE TIME (MS)	MINIMUM TIME (MS)	MAXIMUM TIME (MS)
5	308.05	243.04	382.17
10	308.04	243.04	382.17
15	308.07	243.04	382.17
20	308.09	243.2	383.05
25	308.11	243.2	383.05
30	308.12	243.5	387.01

Table2: Data center operational time

No of VM	AVERAGE TIME (MS)	MINIMUM TIME (MS)	MAXIMUM TIME (MS)
5	0.37	0.05	0.68
10	0.62	0.05	1.63
15	0.87	0.05	1.63
20	1.13	0.09	1.63
25	1.35	0.09	1.63
30	1.78	0.14	1.63

V. Conclusion

The proposed method increases the virtual machine output as it divides the task equally among the virtual machines, based on the Round Robin Scheduling Algorithm. It balances the load on any specific virtual machine. Since the virtual machines are arranged based on priority, the virtual machine most capable of performing a task takes up the job first. The method helps in improving the performance and also prevents starvation, which happens when one single virtual machine takes up the each and every new job and the other virtual machines sit idle. The experiment results clearly indicate that in our proposed method task scheduling consumes lesser time than conventional Round Robin Algorithm.

References

[1]. Oktay Tontaş; Jorge Bernardino "The ethics in cloud computing": 2016 2nd International Conference of the Portuguese Society for Engineering Education (CISPEE)
[2]. Sajay K R; Suvanam Sasidhar Babu "A study of cloud computing environments for High Performance applications" : 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)
[3]. Ahmadreza Montazerolghaem; Mohammad Hossein Yaghmaee; Alberto Leon-Garcia; Mahmoud Naghibzadeh; Farzad TashtarianA Load-Balanced Call Admission Controller for IMS Cloud Computing
[4]. Xiaolong Xu; Lingling Cao; Xinheng Wang; Xiaolong Xu; Lingling Cao; Xinheng Wang "Resource pre-allocation algorithms for low-energy task scheduling of cloud computing" : Journal of Systems Engineering and ElectronicsYear: 2016, Volume: 27, Issue: 2
[5]. Jia Zhao; Kun Yang; Xiaohui Wei; Yan Ding; Liang Hu; Gaochao Xu "A Heuristic Clustering-Based Task Deployment Approach for LoadBalancing Using Bayes Theorem in Cloud Environment ": IEEE Transactions on Parallel and Distributed SystemsYear: 2016, Volume: 27, Issue: 2
[6]. Agraj Sharma; Sateesh K. Peddoju "Response time based load balancing in cloud computing" :2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)
[7]. Durgesh Patel, Mr. Anand S Rajawat "Efficient Throttled Load Balancing Algorithm in Cloud Environment": International Journal of Modern Trends in Engineering and Research

[8]. Aswathi M., Nisha N. Sharma, Mahesh A. S "An Enhancement of Throttled Load Balancing Algorithm in Cloud using Throughput" : I J C T A, 9(15), 2016
[9]. Priyanka SinghÀ and Shubham Pandey "Active Monitoring: ERT Based Load Balancing in Cloud": International Journal of Current Engineering and Technology
[10]. Meenakshi Sharma, Pankaj Sharma, Dr. Sandeep Sharma "Efficient Load Balancing Algorithm in VM Cloud Environment: IJCST Vol. 3, Issue 1, Jan. - March 2012
[11]. Shridhar G. Domanal; G. Ram Mohana Reddy "Load Balancing in Cloud Environment Using a Novel Hybrid Scheduling Algorithm" : 2015 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)