Improve the Cloud System Performance using Time Span & Clustering Technique

Priyanka Soni

Department of Computer Science & Engg.
PCST college Bhopal (M.P) India

priyanka.soni2807@gmail.com

Saurabh Mandloi

Department of Computer Science & Engg.
PCST college Bhopal (M.P) India

saurabhm.research@gmail.com

ABSTRACT

Load balancing is core part of public cloud computing. The process of load balancing increases the performance of cloud based services. Cloud based service provide hardware, software and platform as service. For the balancing of public cloud used two types of approach one is traditional approach and other is swarm based approach. The swarm based approach used particle of swarm optimization, ant colony optimization and glowworm swarm algorithm. for the balancing of load also used different policy such as busy and idle condition such technique is called cloud partion based load balancing technique. The load can be CPU load, memory capacity, delay or network load. Load balancing is the process of distributing the load among various nodes of a distributed system to improve both resource utilization and job response time while also avoiding a situation where some of the nodes are heavily loaded while other nodes are idle or doing very little work. Load balancing ensures that all the processor in the system or every node in the network does approximately the equal amount of work at any instant of time. This technique can be sender initiated, receiver initiated or symmetric type.

Keywords: - Cloud Computing, Load Balancing, Cluster Time Span

INTRODUCTION

A load balancer is a key component in any cloud architecture which distributes incoming service request among the available backend servers as shown in Figure The load balancing algorithm for load balancers can be a generic one or specific to particular cloud category. For example, algorithm for a database load balancer and an FTP load balancer can be different. To prevent single point failure, database servers are replicated. This redundancy also helps to increase the availability of data and decrease the rate of failure if the database servers are spread across the globe. The sole purpose of a database load balancer is to distribute the incoming requests among the available database servers. Besides providing fault tolerance and availability through redundancy, an FTP service based cloud should try to reduce the file transfer time. Hence a load balancer used in an FTP server should be able to route the request such that the delay parameters of transmission from server to client is minimal to reduce the file transfer time.

The increasing amount of nodes in the cloud and their distribution across the world makes it difficult to maintain a centralized load balancer. Being an entry point to the cloud, it is also better to maintain distributed network of load balancers to avoid cloud failure. Hence the load balancing algorithm should have the capability to work on self-regulated distributed environment [2].

The static load balancing algorithm used in homogeneous environment and it does not able to changing user requirements as well as load. Static load balancing methods are cooperative multitasking that means once the load is allocated to the virtual machine it cannot be transferred to another virtual machine. The static load balancing method helps to reduce the execution time to complete a single task. The static algorithm while distributing the load over the various nodes it does not take into the previous state and behavior of the node. The various issues to be addressed by static algorithms are resource utilization, response time, power consumption and scalability.

The dynamic load balancing algorithms based on the heterogeneous environment. Each virtual machine node is able to change the load requirements. These algorithms are able to transfer of task to the remote machine. While transferring the load to the different nodes if one of the node fails it will not halt the system but it little effect on the performance of system. Dynamic algorithms are used for the real time communication which can increase the traffic in the system. It composed of two policies: transfer and location. The transfer policies are those in which a task transfers to one node to another. The goal of load balancing is to effectively distribute the work load between available resources, in order to maximize the benefit from those resources and to have quick computing and processing for the client requests. It is done so as to make resource utilization effective and to improve the response time of the job, simultaneously

removing a condition in which some of the nodes are over loaded while some others are under loaded.

2. CLUSTER COMPUTING

For many years supercomputer was the leader in the field of computing but due to some of the problems faced in the area of science, engineering, and business, which could not be effectively dealt with using supercomputers. They were replaced with clusters with the aim of overcoming these problems and also offered a very cheap way for gaining access to potentially huge computing power [7].

Definition: A cluster is a collection of parallel or distributed computers which are interconnected among them using high-speed networks, such as gigabit Ethernet, SCI, Myrinet and Infiniband. They work together in the execution of compute intensive and data intensive tasks that would be not feasible to execute on a single computer. Clusters are used mainly for high availability, load-balancing and for compute purpose. They are used for high availability purpose as they maintain redundant nodes which are used to provide service when system components fail. The performance of the system is improved here because even if one node fails there is another standby node which will carry the task and eliminates single points of failure without any hindrance. When multiple computers are linked together in a cluster, they share computational workload as a single virtual computer. From the users view point they are multiple machines, but they function as a single virtual machine. The user's request are received and distributed among all the standalone computers to form a cluster. This results in balanced computational work among different machines, improving the performance of the cluster systems. Often clusters are used mainly for computational purposes, than handling IO-based activities.

3. PROBLE FORMULATION

The process of load balancing and task scheduling impart a major role in success of cloud computing. In review process we found that various factor effect the performance of cloud computational in concerns of storage of network data and sharing of resource. The sharing of cloud resource generates a network overload, the network overload arise the problem of bandwidth and stack overflow. Some problem related to cloud computing is given below.

1.   Increasing the time span for process[21]
2.   Failure of resource allocation[22]
3.   Traffic overhead of network[23]
4.   Waste of resource[24]
5.   Cost of monitoring of resource[25]
6.   Process feedback system[26]

For the minimization of this entire problem, used heuristic function by various author in load balancing in cloud computing

4. PROPOSED ALGORITHM

Load balancing is major issue in public cloud computing. The public cloud computing infrastructure consists of hardware, software and platform for the execution of public demand and request. For the handling of multiple request of user cloud computing process used job scheduling and task scheduling process. The job and task scheduling process perform by job scheduler, for the selection of resource and job scheduler used scheduling algorithm such as first come fist and round robin.

In this section discuss the proposed algorithm for load balancing using the technique of teacher learning based optimization process. The teacher based learning optimization algorithm process the task in two phase first phase in job selection and second phase job execution according to the predefined condition. The process of job allocation system define a new constraints set of resource allocation. We used optimization process of cloud task scheduling process for computational Cloud systems try to give better solutions for the selection and allocation of resources to current tasks. The scheduling optimization is very significant because the scheduling is a main house block for making Cloud s more available to user communities. Performance prediction is also used in optimizing the scheduling algorithms. Existing scheduling algorithms only consider an instant value of the performance at the scheduling time, and assume this value remains constant during the task execution. A more accurate model should consider that performance changes during the execution of the application.In this algorithm, if we apply job selection in this method it can provide better result.

Proposed Algorithm:
•   Let n is the no. of jobs (j1, j2, and j3... jn).
•   Let m is the no. of virtual machine (v1,v2,......., vm)
•   Compute the best teacher value according to the define constraints.
•   For each resource obtain the information like bandwidth, computing capacity and current load from job scheduler.
•   For each job obtain the job size and the time needed to complete to complete the job.
•   Create job matrix for the process and apply teacher Factor.
•   Generate the initial population of job and apply the TLBO selection mechanism to select the optimal jobs from initial disturbed jobs. The selection of job is done using teacher function evaluation.

$$Xjob^g_{(i)} = \{x_i^g + rand \times (x_i^g - x_r^g)\} \; if$$
$$otherwise \; x_i^g + rand \times (x_i^g - x_r^g)$$

Here x is total number of job according to the selected according to their state. Here in the process of new teacher generation used the teacher factor value=1. Calculate local jobs and set process priority order for completion of job. If selected job priority is high, then execute the job.
•   After execution of job teacher factor value is updated.
•   Again select local job pool form scheduler process until all jobs are processed.

The key idea of selection operator is to give preference to better individuals by allowing them to the pass on their genes to the next generation and prohibited the entrance of worst fit individuals into next generation; here we are using glowworm approach to only select the job not to find the solution for scheduling.

PROPOSED MODEL
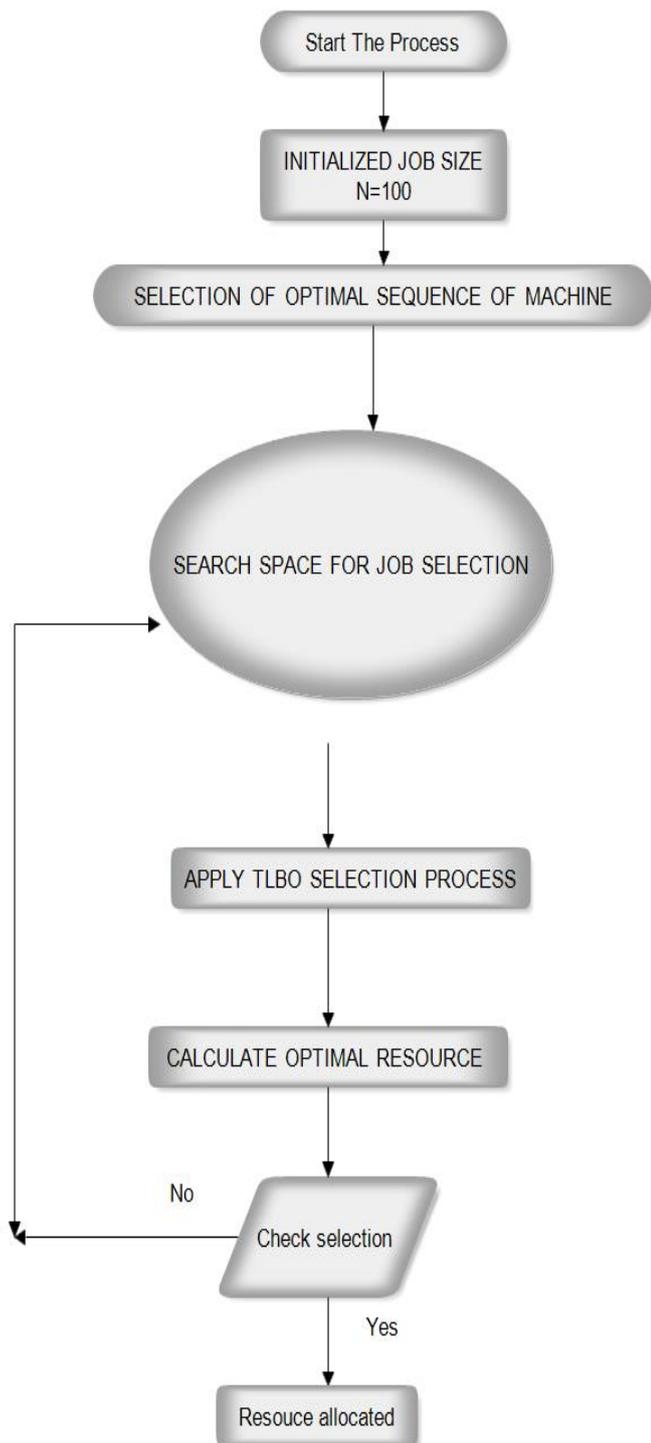
Figure 1. process block diagram of TLBO based Job allocation technique in public cloud

5. EXPERIMENTAL RESULT ANALYSIS

To interact with various services in the cloud and to maintain the resources in a balanced mannertofulfill the requirement of resources/infrastructure by those services, several techniques are required. Based on a core set of features in the three common cloud services such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS).
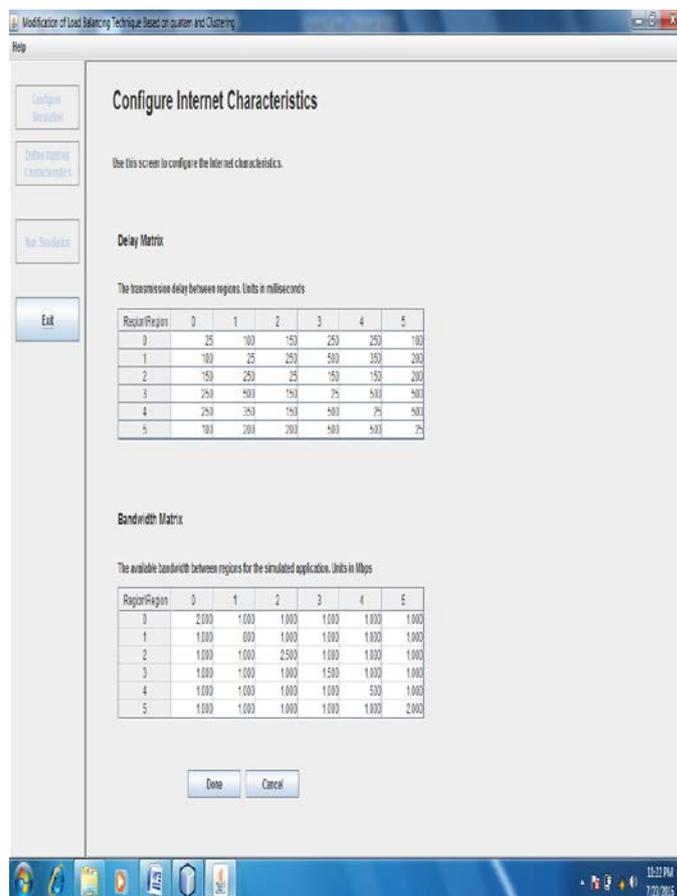

**Figure 2: Shows the Internet Characteristics and it's configuration in Cloud Computing Environment.**
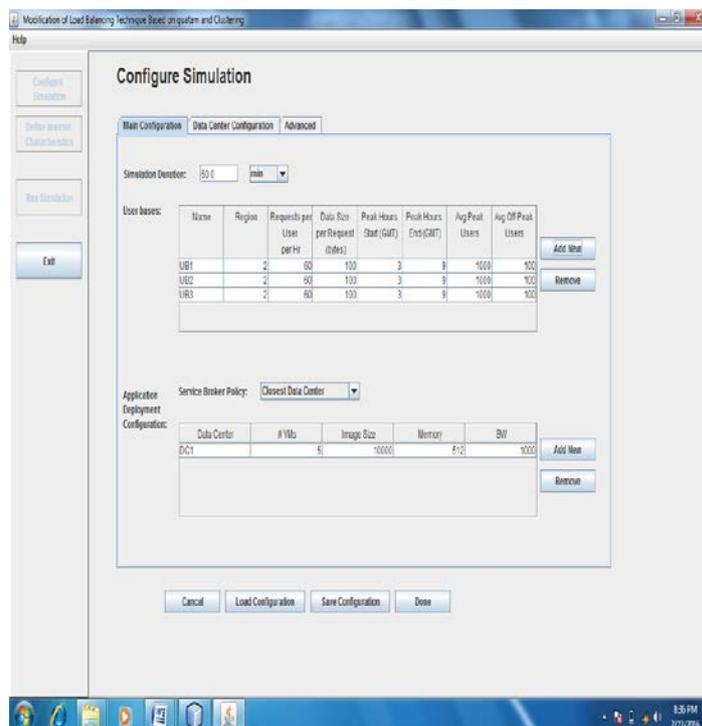
Figure 3: Shows the Addition of New Database to the user database in Cloud Computing Environment.

Data Set	Method Name	Overall Response Time		Data Center Processing Time	
UB2 DC3	Round Robin Method	Average(ms)	300.82	Average(ms)	0.36
		Minimum(ms)	240.14	Minimum(ms)	0.02
		Maximum(ms)	375.14	Maximum(ms)	0.64

Table 1: Shows the Response Time and Processing Time analysis for Round Robin Method.
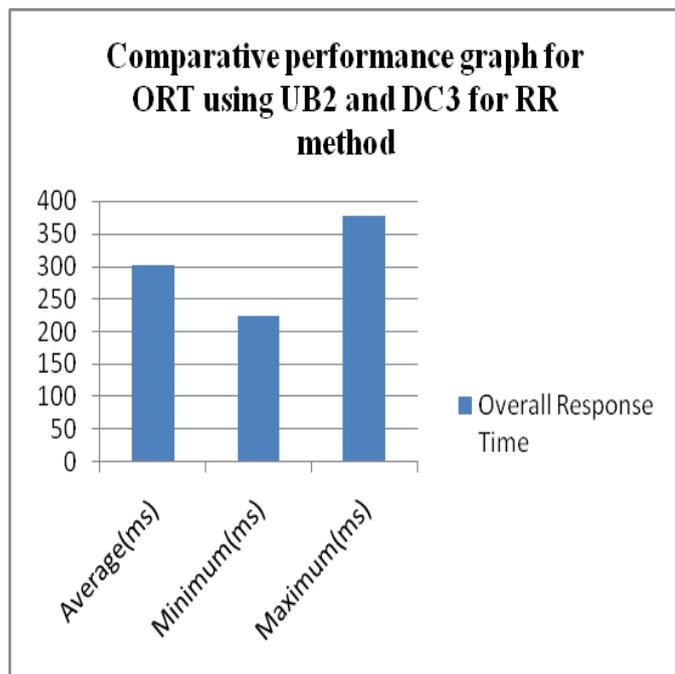
Figure 4: Shows the comparative performance of ORT for UB2 and DC3 using Round Robin Method in terms of Average, Minimum and Maximum values in mili seconds.
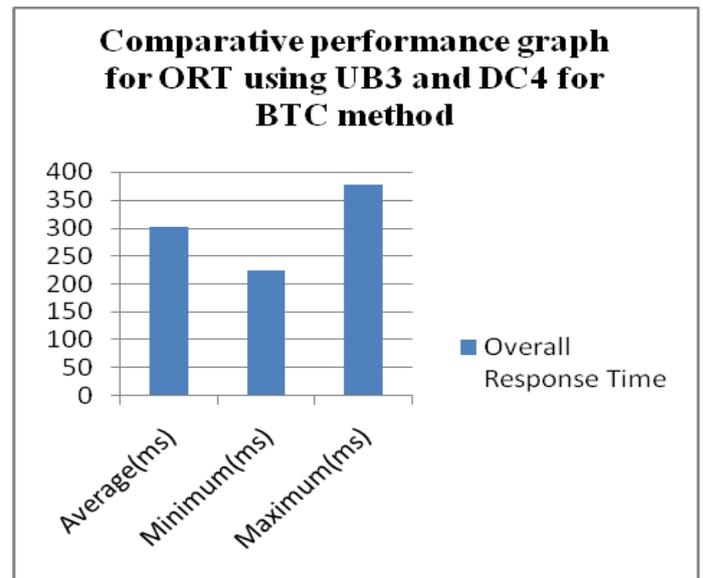
Figure 5: Shows the comparative performance of ORT for UB3 and DC4 using BTC Method in terms of Average, Minimum and Maximum values in mili seconds.

6. CONCLUSION AND FUTURE WORK

In this paper used teacher based learning optimization algorithm for the selection of job and task for the resource allocation in public cloud computing. The TLBO function gives the better performance instead of other swarm based algorithm such as glowworm swarm intelligence algorithm. Basically TLBO algorithm facilities the all task in two sections one is job allocation process and other is job selection process. The job selection process is defined by the predefined constraints function. For the evaluation of performance used cloud simulator software such is called cloud analyst. The cloud analysis software is bag of composition of cloud environment and load balancing policy. In scenario of policy design two services one is glowworm optimization policy and other is TLBO based policy. The TLBO based policy reduces the load effect approx 10-12% in compression of glowworm algorithm. Cloud Computing is a vast concept and load balancing plays a very important role in case of Clouds. There is a huge scope of improvement in this area. We have discussed only two divisible load scheduling algorithms that can be applied to clouds, but there are still other approaches that can be applied to balance the load in clouds. The performance of the given algorithms can also be increased by varying different parameters

REFERENCES:-

[1] Jia Zhao, Kun Yang, Xiaohui Wei, Yan Ding, Liang Hu, and Gaochao Xu "A Heuristic Clustering-Based Task Deployment Approach for Load Balancing Using Bayes Theorem in Cloud Environment", IEEE, 2016, Pp 305-316.

[2] Ariharan V and Sheeja S Manakattu "Neighbour Aware Random Sampling (NARS) algorithm for load balancing in Cloud computing", IEEE, 2015, Pp 1-5.

[3] Catalin-Constantin Usurelu, Mihaela-Catalina Nita, Roxana Istrate, Florin Pop and Nicolae Tapus "Spider Mesh

Overlay for Task Load Balancing in Cloud Computing", IEEE, 2015, Pp 433-440.

[4] Mohamed Belkhouraf, Ali Kartit, Hassan Ouahmane, Hamza Kamal Idrissi, Zaid Kartit and Mohamed El Marraki "A secured load balancing architecture for cloud computing based on multiple clusters", IEEE, 2015, Pp 1-6.

[5] Yi Wei and M. Brian Blake "Service-Oriented Computing and Cloud Computing", IEEE, 2011, Pp 72-75.

[6] Anurag Jain and Rajneesh Kumar "Scalable and Trustworthy Load Balancing Technique for Cloud Environment", International Journal of Engineering and Technology, 2016, Pp 1245-1251.

[7] Naidila Sadashiv and S. M Dilip Kumar "Cluster, Grid and Cloud Computing: A Detailed Comparison", IEEE, 2011, Pp 477-482.

[8] Mayanka Katyal and Atul Mishra "A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment", International Journal of Distributed and Cloud Computing, 2013, Pp 5-14.

[9] S. Subashini N and V. Kavitha "A survey on security issues in service delivery models of cloud computing", Journal of Network and Computer Applications, 2010, Pp 1-11.

[10] Saleh Basalamah, Abdulrahman A. Almutairi, Muhammad I. Sarfraz, Walid G. Aref and Arif Ghafoor "A Distributed Access Control Architecture for Cloud Computing", IEEE, 2012, Pp 36-44.

[11] Dimitrios Zissis and Dimitrios Lekkas "Addressing cloud computing security issues", Future Generation Computer Systems, 2012, Pp 583-592.

[12] Jianxin Li, Bo Li, Tianyu Wo, Chunming Hu, Jinpeng Huai, Lu Liu and K.P. Lam "CyberGuarder: A virtualization security assurance architecture for green cloud computing" Future Generation Computer Systems, 2012, Pp 379-390.

[13] Tharam Dillon, Chen Wu and Elizabeth Chang "Cloud Computing: Issues and Challenges",, IEEE, 2010, Pp 27-33.

[14] Er. Pooja and Er. Vivek Thapar "Survey of VM Load Balancing Algorithm in Cloud Environment", International Journal of Computer Science Trends and Technology, 2016, Pp 123-129.

[15] Suresh M, Santhosh Kumar B. and Dr.S.Karthik "A LOAD BALANCING MODEL IN PUBLIC CLOUD USING ANFIS AND GSO", 2014, Pp 85-89.

[16] Ibrahim Aljarah and Simone A. Ludwig "A MapReduce based Glowworm Swarm Optimization Approach for Multimodal Functions", IEEE, 2013, Pp 22-31.