Volume 5, Issue 12, December 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Novel Approach of Load Balancing in Cloud Computing Using Cat Swarm Optimization Technique

Upasana Mittal[1], Yogesh Kumar[2], Anupama Kaur[3]
[1]C.S.E Student, Department of Computer Science, SUSCET, Mohali, India
[2]Head of Department, Department of CSE, CIET, Fatehgarh Sahib, India
[3]Assistant Professor, Department of CSE, SUSCET, Mohali, India

Abstract: Cloud computing is a new technology and it is becoming popular because of its great features. In this technology almost everything like hardware, software and platform are provided as a service. A cloud provider provides services on the basis of client's requests. An important issue in cloud is, scheduling of users requests, means how to allocate resources to these requests, so that the requested tasks can be completed in a minimum time and the cost incurred in the task should also be minimum. In cloud computing Load balancing is required to distribute the dynamic workload evenly across the all nodes. Load balancing helps to achieve a high user satisfaction and resource utilization ration by ensuring and efficient and fair allocation of every computer resource. In the research work, we proposed an algorithm for load distribution of workloads on a cloud by the use of Cat Swarm Optimization in Cloud Computing. Result shows that our algorithm can achieve better load balancing in a large scale cloud computing environment as compared to previous load balancing algorithms.

Keywords- Cat Swarm Optimization, Load balancing, Virtualization, Cloud data centre, Cloud computing.

I. INTRODUCTION

In Cloud computing services can be used from diverse and wide spread resources, rather than remote servers or local machine. It consists of a bunch of distributed servers known as masters, providing demanded services and resources to different clients known as clients in a network with scalability and reliability of datacenter. The distributed computers provide on-demand services. Services may be of software resources (e.g. Software as a Service, SaaS) or physical resources (e.g. Platform as a Service, PaaS) or hardware/infrastructure (e.g. Hardware as a Service, HaaS or Infrastructure as a Service, IaaS). AmazonEC2 (Amazon Elastic Compute Cloud) is an example of cloud computing services [2].Cloud computing has recently become popular due to the maturity of related technologies such as network devices, software applications and hardware capacities. Resources in these systems can be widely distributed and the scale of resources involved can range from several servers to an entire data center. To integrate and make good use of resources at various scales, cloud computing needs efficient methods to manage them [4]. Consequently, the focus of much research in recent years has been on how to utilize resources and how to reduce power consumption. One of the key technologies in cloud computing is virtualization. The ability to create virtual machines (VMs) [14] dynamically on demand is a popular solution for managing resources on physical machines. Some research aims to improve resource utilization while others aim to reduce energy consumption. The goals of both are to reduce costs for data centers. Due to the large size of many data centers, the financial savings are substantial. One of the major problem in cloud computing is load balancing. As the requests of the clients can be random to the nodes they can vary in quantity and thus the load on each node can also vary. Therefore, every node in a cloud can be unevenly loaded of tasks according to the amount of work requested by the clients. This phenomenon can drastically reduce the working efficiency of the cloud as some nodes which are overloaded will have a higher task completion time compared to the corresponding time taken on an under loaded node in the same cloud. Load balancing in large distributed server systems is a complex optimization problem of critical importance in cloud systems and data centres. Load balancing in cloud is a mechanism that distributes the excess dynamic local work load evenly across all the nodes. It is used to achieve a high user satisfaction and resource utilization ratio, making sure that no single node is overwhelmed, hence improving the overall performance of the system. Overloaded nodes across the server and storage side often lead to performance degradation and are more vulnerable to various failures. To remove this limitation the load must be migrated from the overloaded resource to an underutilized one without causing harm and disruption to the application workload. The rest of the paper is organised as follows: Section II gives a review of related work which realizes importance of load balancing in cloud computing.Section III introduces the proposed approach for load balancing.Section IV describes the implementation .Section V analyses the performance of the mechanishm.and section VI discuss the future work and final conculsion of the work.

II. MOTIVATION & RELATED WORK

Load balancing in the cloud computing environment has an important impact on the performance. Good Load balancing makes cloud computing more efficient and improves user satisfaction. There have been many studies of load balancing

for the Cloud environment. **Gulshan Soni[1] et al. in 2014** have proposed the idea of central Load balancer that manage the load on the cloud and assign the load corresponding to their priority and the response time is less as compare to other algorithm. Load Balancing means to distribute the load among n virtual machines so that no one machine is idle or overloaded.

Vikas Kumar et al. [3] in 2014 presented load balancing is the main issue. Load balancing is required to distribute the excess dynamic local workload evenly to entire load in the whole cloud to achieve a high user satisfaction and resource utilization ratio.

Foram F.Kherani et al.[4] in 2014 described the load balancing in cloud computing and how to improve and maintain the performance of cloud computer and also discuss the comparison of various existing static load balancer and conventional dynamic load balancer also. the paper describe the three algorithms round robin , equally spread current execution load and throttled load balancing.

Rajesh Gorge Rajan et al. [5] in 2013 have the investigated the different algorithms proposed to resolve the issue of load balancing and task scheduling in Cloud Computing. They discussed and compared these algorithms to provide an overview of the latest approaches in the field. Load Balancing is essential for efficient operations in distributed environments.

Suriya Begum et al. [6] in 2013 have described the random arrival of load in such an environment can cause some server to heavily loaded while other server is idle or only lightly loaded. Equally load distributing enhances performance by transferring load from heavily server. Efficient scheduling and resource allocation is a critical characteristic of cloud computing based on which performance can be estimated. It is required to distribute the dynamic load workload evenly across all the nodes to achieve the high user satisfaction and resource utilization ration by making sure that every computing resource is distributed efficiently and fairly.

Argha Roy et al.[7] in 2013 presented the researcher proposed the idea of dynamic load balancing.resercher concluded that dynamic load balancing is a technique to use the cloud computing in efficient manner .

Meysam Orouskhani et al.[8] in 2013 proposed the improved CSO algorithm namely "Adaptive dynamic cat Swarm Optimization "The Paper described the addition of a new adaptive inertia weight to velocity equation and use of adaptive acceleration ratio.The Proposed CSO take less time to converge and can find best solution in less iteration.

Amir Nahir et al. [9] in 2013 presented the approach which is based on creating several replicas of each job and sending each replica to a different server. Upon the arrival of a replica to the head of the queue at its server, the latter signals the servers holding replicas of that job, so as to remove them from their queues. They show, through an analysis and simulations, that this scheme improves the expected queuing overhead over traditional schemes by a factor of 9 (or more) under various load conditions.

Parveen Patel et al. [10] in 2013 proposed an approach that shows that Layer-4 load balancing is fundamental to creating scale-out web services. We designed and implemented Ananta, a scale-outlayer-4 load balancer that runs on commodity hardware and meets the performance, reliability and operational requirements of multi-tenant cloud computing environments

Isam Azawi Mohialdeen [11] in 2013 has discussed, the behavior of four job scheduling algorithms, namely: Random, Round-Rubin (RR), Opportunistic Load Balancing and Minimum Completion Time have been investigated and examined in a Cloud computing environment. Based on the results, it can be also concluded that there is not a single scheduling algorithm that provides superior performance with respect to various types of quality services. This is because job scheduling algorithms needs to be selected based on its ability to ensure good quality of services with reasonable cost and maintain fairness by fairly distribute the available resources among all the jobs and respond to the constraints of the users.

Prof. Dr. Jayant et al. [12] in 2013 described cloud computing is a distributed computing model in which everything from software to infrastructure is provided as a service like utility computing. Job scheduling is one of the core and challenging issues in cloud computing. The decisions like when to allocate hardware resources to the tasks has become the main issue in cloud computing. Job scheduling algorithm is an NP- completeness problem which play key role in cloud computing. They presented the surveys of the current job scheduling algorithms under cloud environment and summarize some method to improve the performance.

III. PROPOSED APPROACH FOR LOAD BALANCING

The Proposed load balancing algorithm " Cat Swarm Optimization" will balance the balance the load among virtual machines. Cat Swarm Optimization (CSO) is generated by observing the behaviour of cats and composed of two sub models that is Tracing Mode and seeking model which model upon the behavior of cats.CSO is One the new heuristic optimization algorithm which is based on the Swarm Intelligence. This algorithm has better performance compared to other heuristic optimization algorithm.CSO is one of the new Swarm intelligence algorithm finding the best global solution because of complexity. Pure CSO takes a[11] long time to converge and can't achieve the accurate solution. Dynamic CSO is advanced form of CSO.

The behaviour of the cat is:

a) Rest indolently most of time when they are awake
b) Move speedily when they are tracing some targets
c) Curious about all kinds of moving things
d) There are two Modes of CSO : Seeking and tracing Mode.

Seeking Mode: This Mode is the time for thinking and deciding about next mode. To model the situation of the cat, which is resting, looking around and seeking the next position to move to. This Mode has four main parameters:

a. Seeking Memory Pool (SMP): To be used to define the size of seeking memory for each cat, this indicates the points sought by the cat.
b. Seeking range of The selected dimension (SRD) : To declare the mutative ratio for the selected dimensions
c. Counts of dimension to change (CDC): To disclose how many dimensions will be varied.
d. Self Position consideration (SPC): A Boolean variable, which decides whether the point, where the cat is already standing, will be one of the candidates to move to.

Tracing Mode: In this mode Cats desire to trace the targets and foods. For modelling the case of the cat is tracing some targets

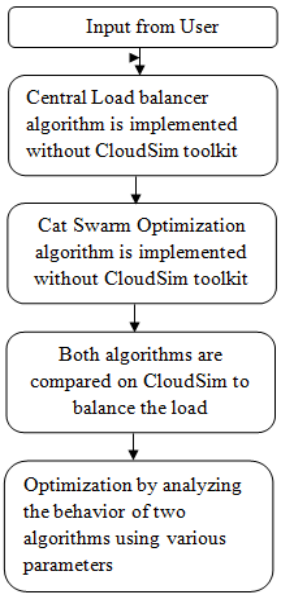

Figure 3.1. Algorithmic Flow

IV. EXPERIMENTAL SETUP

The Proposed algorithm is implemented and integrated in CloudSim. Both algorithms CLB and CSO implemented and compared on CloudSim tool for energy efficiency and load balancing.

Table 4.1. Different algorithms are compared for 5 Jobs.

Algorithms	Number of jobs/Tasks	Response Time	Execution Time	Energy Consumption
Central Load Balancer	**For #5**	3.142	85.78	7720.3
Cat Swarm	**For #5**	2.33	54.89	4941

Table 4.2. Different algorithms are compared for 6 Jobs.

Algorithms	Number of jobs/Tasks	Response Time	Execution Time	Energy Consumption
Central Load Balancer	**For #6**	2.1	83.05	7474.5
Cat Swarm	**For #6**	1.66	73.82	6643.8

Table 4.3. Different algorithms are compared for 11 Jobs.

Algorithms	Number of jobs/Tasks	Response Time	Execution Time	Energy Consumption
Central Load Balancer	**For #11**	0.2	144.3	128288
Cat Swarm	**For #11**	0.21	30.5	945

Both algorithms are compared for energy efficiency and load balancing. The results show that Cat Swarm technique is the better because which consumes less energy and all the tasks are executed in less time with no delay. It concluded that Cat Swarm is best in the energy efficient technique in cloud computing.

Parameter used :

- Number of jobs /task
- Response Time
- Execution Time
- Energy Consumption

V. PERFORMANCE EVALUTION

Graphical charts shows the comparison between different techniques based on distinct number of job Considering discrete parameters like energy consumption, response time and execution time.

Fig.5.1 reveals the information about the energy consumption of datacenter and the results show that Cat Swarm technique consumed less energy as compared to Central load balancer. Therfore our proposed model is best in case of energy efficient.

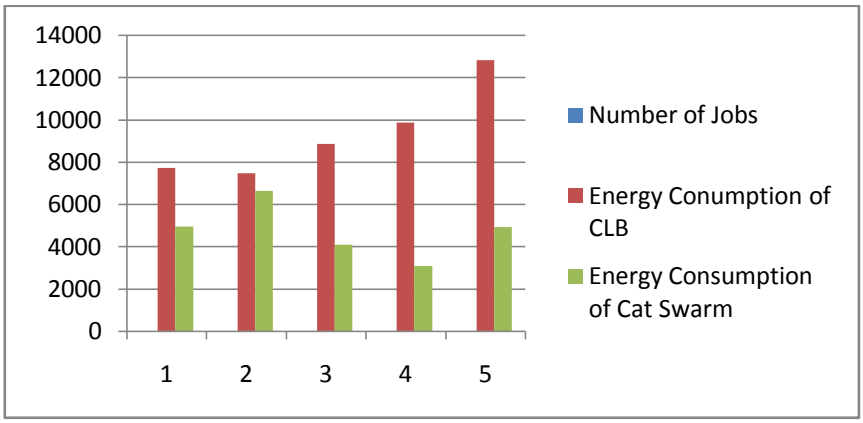

Fig. 6.1 Energy consumption comparison based on distinct number of jobs.

Response time is the time taken to respond by a particular load balancing algorithm in a distributed system. The response time is calculated by varying the number of jobs. Fig. 5.15 depicts that the response time is minimum in case of Cat Swarm Optimization technique.

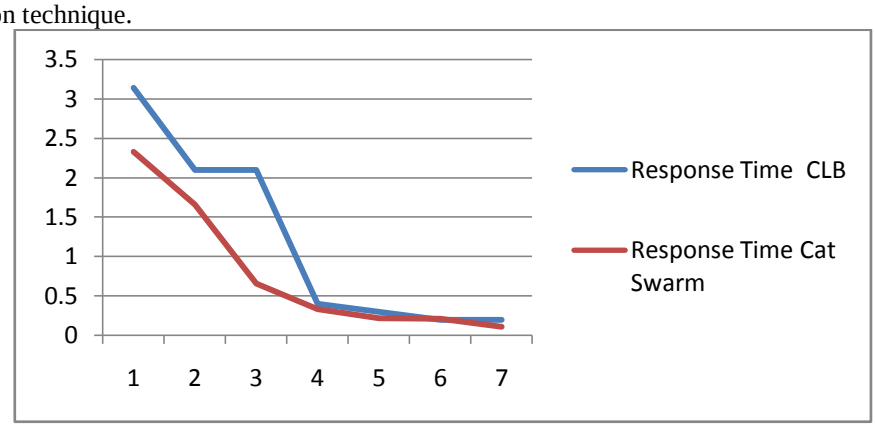

Fig. 6.2 Response time comparison based on distinct number of jobs.

Execution time is estimated by calculating the total number of jobs executed within a fixed span of time. Fig 6.3 shows the execution time taken by three techniques for distinct number of jobs. After implementation results show that the execution time of Cat Swarm Optimization is less as compared to Central Load Balancer.

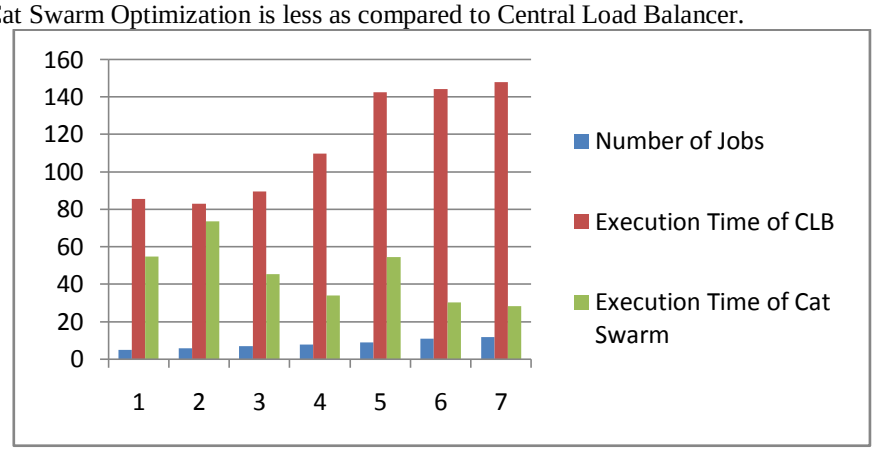

Fig. 6.3 Execution time comparison based on distinct number of jobs.

VI. CONCULUSION AND FUTURE WORK

In the research work we have proposed and implemented an Cat Swarm Optimization on cloud environment using CloudSim Toolkit. And compared it with the Central Load Balancer technique. The results show that proposed technique is much better than the existing load balancing methods in terms of Response time, Execution Time, and Throughput. We also concluded that Cat Swarm Optimization technique consumes less energy than Central Load Balancer.

Cloud Computing is a vast concept and energy efficiency plays a very important role in case of Clouds. There is a huge scope of improvement in this area. We have implemented only two dynamic load balancing algorithms. But there are still other approaches that can be applied to balance the load and energy consumption in clouds. The performance of the given algorithms can also be increased by varying different parameters. We can also move our research work on any Private Cloud for the Security and further enhancements.

REFERENCES

[1] Gulshan Soni,Mala Kalra, "A Novel Approach For Load Balancing In Cloud Data Center",978-1-4799-272-8/14/$31.00,IEEE 2014.

[2] Ruhi Gupta "*Review on Existing load balancing techniques in Cloud Computing"*, Volume 4, Issue 2, ISSN: 2277, February 2014.

[3] Vikas Kumar and Shiva Kumar*" A load balancing based cloud Computing Techniques and Challenges",* IJSRM volume 2 issue 5 May 2014.

[4] Foram F.Kherani and Prof Jignesh Vania "*Load Balancing in cloud computing"* , Vol-1,issue 1,ISSN-2321-999, January 2014.

[5] Rajesh Gorge Rajan and V.Jeyakrishnan *"A Survey on Load Balancing in Cloud Computing Environments"*Vol-2.Issue-12,December 2013.

[6] Dongliang Zhang,Changjun Jiang,Shu Li, " *A fast Adaptive Load balancing method for parallel practical based simulations*", Simulation Modelling Practice and theory 17(2009) 1032-1042.

[7] M.Randles,D.Lamb and A.Bendiab, "*A Comparative Study into distributed load balancing algorithm for cloud computing*" IEEE 24th International Conference on Advanced Information and Networking and Applications Workshops,pp.551-556,April 2010.

[8] Suriya Begum, Dr. Prashanth C.S.R, "*Review of load balancing in cloud Computing".* 10, Issue 1, No 2, January 2013.

[9] Rahul Bhoyar,"*Cloud Computing: Service Models ,Types,database and Issues*",Volume 3, Issue 3, March 2013.

[10] Argha Roy, Diptam Dutta, "*Dynamic Load balancing: Improve efficiency in cloud Computing*", Vol 15, issue 2, ISSN: 2278-9359,December 2013.

[11] Meysam Orouskhani *"Novel Cat Swarn Optimization algorithm for Unconstraint Optimization Problem",* IJITCS, vol.5, no.11, pp.32- 41, 2013.

[12] Mohiuddin Ahmed, Abu Sina Md. Raju Chowdhury, Mustaq Ahmed, Md. Mahmudul, *"An Advanced Survey on Cloud Computing and State-of-the-art",* Issue 1, No 1, January 2012.

[13] Jaspreet Kaur, "*The comparison of Load balancing Algorithm in cloud*",Vol. 2, Issue 3, May-Jun 2012.

[14] Ram Prasad Padhay,P Gautam Prasad Rao, "Load Balancing in Cloud Computing System",107CSO46.

[15] Jitendra Bhatia,Trith Patel,Harshal Trivedi ,Vishrut Majmudar, "*HTV Dynamic Load balancing For Virtual Machine Instances In Cloud*" , 18 Dec 2012, Pages 1-20 IEEE.

[16] Parin.V.Patel ,Hitesh.D.Patel, Pinal.J.Patel, "*A Survey Of Load Balancing In Cloud Computing*" IJERT, Vol.1,Issue 9,November 2012.

[17] P.Salot, "*A Survey of various Scheduling algorithm in cloud computing Environment*", IJRET, Volume:2,Issue:2,Feb 2012.

[18] Y.Fang,F.Wang ,And J.Ge , "*A Task Scheduling algorithm Based on Load Balancing in Cloud Computing*" , Web Information Systems and mining,Lectures Notes in Computer Science,Vol 6318, Pages-271-277,2010.

[19] J.Kaur,"*Comparison of Load Balancing in cloud*", IJERT, Vol.2,No:3,pp. 1169-73,2012.

[20] Sujit Tilak, "*A Survey of Various Scheduling Algorithms in Cloud Environment",* IJEI, Volume 1, Issue 2 (September 2012) PP: 36-39.

[21] Saeed Parsa and Reza Entezari-Maleki, *" RASA: A New Task Scheduling Algorithm in Grid Environment"* in World Applied Sciences Journal 7 (Special Issue of Computer & IT): 152-160, 2009.Berry M. W., Dumais S. T., O'Brien G. W. Using linear algebra for intelligent information retrieval, SIAM Review, 1995, 37, pp. 573-595.

[22] S. K. Garg, C. S. Yeob, A. Anandasivamc, and R. Buyya,"*Environment-conscious scheduling of HPC applications ondistributed Cloud-oriented data centers*", Journal of Parallel and Distributed Computing, Elsevier, Vol. 70, No. 6, May 2010, pages 1-18.

[23] Jitendra Bhatia, Tirth Patel, Harshal Trivedi, Vishrut Majmudar, *" HTV Dynamic Load Balancing Algorithm for Virtual Machine Instances in Cloud"*,18,Dec2 012,Pages 15-20 IEEE.

[24] Dr. Hemant S. Mahalle, Prof. Parag R. Kaveri ,Dr.Vinay Chavan, *" Load Balancing On Cloud Data Centres" ,* International Journal of Advanced Research in Computer Science and Software Engineering,Vol.3,Issue 1,January 2013.

[25] Ali M Alakeel, "*A Guide To Dynamic Load Balancing In Distributed Computer Systems",* International Journal of Computer Science and Network Security, Vol. 10 No. 6, June 2010.

[26] A.K. Sidhu, S. Kinger, "*Analysis of Load Balancing Techniques in Cloud Computing*", International Journal of Computers & Technology,Volume 4 No. 2, ISSN 2277-3061, March-April, 2013.

[27] Suresh M., Shafi Ullah Z., Santhosh Kumar B., "*An Analysis of Load Balancing in Cloud Computing"*, International Journal of Engineering Research & Technology (IJERT),Vol. 2 Issue 10, October – 2013, ISSN: 2278-0181

[28] A. Kaur,N. Bansal, "*Cross-Breed Job Scheduling for Reducing the Server Load Using RBAC at Cloud*", International Journal of Advanced Research in Computer Science and Software Engineering", Volume 3, Issue 5, ISSN: 2277 128X, May 2013.

[29] P. Salot, "*A survey of various scheduling algorithm in cloud computing environment*", IJRET, Volume: 2 Issue: 2, Feb 2013.

[30] S. Mohana Priya, B. Subramani, "*A new approach for load balancing in cloud computing"*, International Journal of Engineering and Computer Science, ISSN: 2319-7242 Volume 2 Issue 5 May, 2013 Page No. 1636-1640.

