Load Balancing in Cloud Computing Using Optimization Algorithms

Arun Gupta, Dr. Anoop Sharma
Department of Computer Applications
Singhania University in Pacheri Bari, Jhunjhunu (Raj.), India

Abstract - Cloud computing (CC) is a platform which helps to handle the unlimited information services based on every utilized model. In recent times, the cloud resources are employed as virtual machines (VMs) to actual machines (AMs) at which the virtualize technology is used. All virtual machines are executed in their framework and lead to consume resources from their AM that are considered as a host. To balance the load, cloud focuses on moving VMs amid anxious real machines and less complex AMs. The major focus is on re-locating the delay of this computing extensions in the organization as VMs. This work introduces an innovative algorithm, called Butterfly optimization (BO) to migrate the VM. MATLAB is executed for simulating the introduced algorithm. The work also aims at comparing the introduced algorithm with the traditional algorithms. The delay, bandwidth used and space are employed as evaluation metrics for computing the introduced algorithm.

Keywords - Load Balancing, Cloud Computing, Optimization Algorithm, Butterfly Algorithm

I. INTRODUCTION

CC is a parallel and distributive model in which several computer systems are inter-related and virtualized. These systems are put forward as at least one or different computation assets on the basis of Service Level Arrangements (SLA) set up. It allows to share information amid clients and the Service Providers (SPs) [1]. A cloud system is employed for offering services over diverse kinds of networks, namely private and public networks. This platform is adopted for executing diverse applications, such as email, CRM (Client Relation Management) etc. CC platform can be employed easily and provides accessibility for the transmitted groups having information, and hardware devices. This technology assists the customers and industries in attaining several potentials for storing and handling their data available within DCs of third party [2]. The sharing of assets is essential in acquiring understandable and considerable savings, such as the power grid over an association [3]. CC model is better in contrast to the traditional technological systems as it provides highly automated system, provisioning, and virtualized technologies. This model is effective to separate data and software from the servers and storage systems which are executed on them. CC helps in distributing and transmitting the IT assets as a convenience in the fragments of component at which the customers are allowed for purchasing applications or merely take on computing power on rental basis or as a unified totality. This platform is useful for the shared resources and make them effective at an extensive level. The requirement of clients is considered for re-distributing the resources along with their sharing among a several customers dynamically. Consequently, this event is effective to distribute the resource amongst consumers. To illustrate, a cloud computing (CC) facility available for European clients within working duration of Europeans with a specific application (like E mail) is capable of re-distributing the same resources to assist the customers in North America within the working duration of North Americans on the basis of a distinct application (like, a web server) [4].

A number of clouds are present in CC. The clients are free to avail the subscription of these clouds in accordance with their necessities. A person having a home or a small corporate is allowed for deploying public cloud services. Diverse CC systems are defined below:

a. Public Cloud: This model allows common people or an enormous business for deploying the cloud model over the Internet [4]. No personal framework is offered to client, however a company which is responsible for offering cloud services, have potential to control over those services. The users are allowed to use these services free of cost when they subscribe it or they can use them as a paid service. There is not any control provided to the client and the consumer have not any information from where those services come. Multiple companies have to share the chief structure. Though, the distribution of data and application is done in a sensible manner to every organization and the services are accessible only for the authentic users. Unlike the other cloud models, such clouds are unsecure as over load is occurred for executing the entire implementation. Moreover, the information attained from this structure has not considered malicious activities. Oracle, Google AppEngine, etc. are some of the examples of such clouds.

b. Private Cloud: A public platform is employed for establishing and hosting this service in the user data centre

(DC) to be utilized in a single organization. A particular company is capable of attaining the access of this service while the others are restricted. The company is considered as a Cloud Service Provider (CSP) for inserting business components which have all potentials of a cloud. For this, no individual model is provisioned. A corporation, a third party, or their hybrid are facilitated for managing, owning and operating this framework [5]. This kind of framework is set up on the location of a company as well as its hosting is done in a DC that is owned by a third party. Such a platform offers more security in contrast to the existing one as it has some particular internal disclosure.

c. Community cloud: A number of organizations adopts this platform mutually. This framework is effective for a specific group such aa medical data transmission related to mission, strategy and compliance contemplations [6]. Such a framework is utilized to attain several benefits such a shared design outlays, billing concept for every client like a private cloud to its companion businesses. However, it keeps the services of a private cloud more secure and protective. The implementation of this kind of cloud is possible on locations or at the DC of third party. The third party is capable of managing this framework.

d. Hybrid cloud: The public cloud frameworks are integrated with the private ones. The standard or a new method is employed for integrating these frameworks which allows to transmit the data and applications. Such centres which incorporate services and data from dissimilar sorts of cloud frameworks are considered to create a combined, automated, and properly controlled computing system. This framework is useful for SPs for utilizing third party cloud deliverers in a complete or partial way to compute the data more flexibly. Moreover, it allows an organization to take the potentials of scalable and profitable features of a public cloud for which revelation to data which is available outer location from the commercial intranet is not done.

II. LITERATURE SURVEY

D. A. Shafiq, et.al (2021) recommended an algorithmic approach that balanced the load in a competent way and also led to enhance the assets concerning parameters of Quality of Service, the precedence of Virtual Machines, etc. [7]. The presented approach was deployed to enhance the procedure in which the cloud resources were deployed and allocated, and the time utilized to schedule the task was alleviated. This process resulted in enhancing the efficiency of the cloud computing (CC). According to outcomes, the recommended algorithm consumed the resources up to 78% as compared to other methods. Furthermore, this algorithm mitigated the execution time as well as makespan.

Nithin K. C. Das, et.al (2017) emphasized on integrating Weighted Round Robin with Honeybee Algorithm so that the reaction and processing period were lessened [8]. This approach was executed after allocating weights every Virtual Machine (VM). The requirement of resources was considered in executing tasks to select the VM. Based on the findings, the developed algorithm had worked well against others and led to enhance the RT and time for processing the DC. The next project would be aimed at attaining enhanced Quality of Service elements such as waiting time, time to migrate, cost-efficiency, etc.

M. Jeyakarthic, et.al (2020) investigated a Client Side-Server Side and Grasshopper Optimization method with MapReduce to tackle the problem about offloading and mitigate the deployment of assets [9]. This method allowed the client for choosing the server and transmitting a request to make the data reliable, available and redundancy free. The Map Reduction technique was adopted for deploying GO on the basis of precision. The experiments revealed the effectiveness of the investigated method at accuracy of 0.977.

Dharavath Ramesh, et.al (2018) formulated a framework called Slot Based Carton Load Balancing so that the resources were deployed in an efficient way [10]. This framework was executed according to the UBs of the cloud framework layer. Moreover, it was effective for handling the procedure in which the VM (virtual machine) was assigned dynamically. The formulated framework was simulated and tested on CloudSim and Cloud Analyst under diverse circumstances. The experimental results exhibited that the formulated framework was applicable to manage the load in an effective way to make the resource usage more reliable. Besides, this framework was assisted in alleviating the access and execution time as compared to conventional methods.

M. Lawanyashri, et.al (2017) constructed an ensemble of FOA and SA to maximize the convergent speed and optimizing accuracy [11]. This method was effective to deploy the resources reliably and diminish the power usage and cost in CC system. In the end, a modified technique was generated on the basis of obtained results. The evaluation exhibited that the constructed algorithm had outperformed existing methods and balanced the workload with higher accuracy against the conventional techniques.

A. Francis Saviour Devaraj, et.al (2020) recommended an innovative ensemble mechanism recognized as FIMPSO that an ensemble of FFA (Firefly algorithm) and IMPSO. In this, the load was balanced precisely [12]. The initial algorithm was employed for alleviating the search space and the latter one to identify the enhanced response. This mechanism was assisted in distributing the load. According to findings, the recommended mechanism was more robust over traditional methods. Moreover, the memory usage of this mechanism was computed 0.93, reliability was 0.67 and throughput was 0.72.

Lung-Hsuan Hung, et.al (2021) devised and put together 2 genetic-based systems to balance the workload [13]. First of all, various parameters were taken in account for scrutinizing the output of virtual machines (VMs). A cloud environment was executed for simulating the devised systems. The Gene Expression Programming (GEP) approach was implemented to develop the symbolic regression (SR) techniques. The efficacy of virtual machines was evaluated and their load were predicted after balancing them using these techniques. After that, the simulation was conducted on Jnet. Based on outcomes, the devised systems were more successful in comparison with the existing methods.

Muhammad Junaid, et.al (2020) intended a DFTF approach in which a modification of CSO was put forward with Support Vector Machine so that the load was balanced [14]. At first, the intended approach was employed to classify the data, generated via dissimilar sources, into dissimilar classes in CC. At second, the modified approach was fed with the generated data as input to distribute the load on VMs. According to experiments, the intended approach was efficiently balanced the load and enhanced the throughput up to 0.7, the RT around 0.82 and time to migrate VM up to 0.13 in contrast to the existing techniques.

H. Lathashree, et.al (2018) formulated a technique for balancing the load in a precise way under Cloud Computing environment [15]. The fundamental task of this technique was that the load was balanced successfully and the quality of service needs pf clients were fulfilled according to the demands available in a SLA. This objective was attained by the means of CAs and ANN models. The initial model was capable of making promising decisions in a moveable setting. The results validated the efficacy of the formulated technique with enhanced ET, throughput and resource utilizations.

Zixi Cui, et.al (2021) introduced a robust system known as Closer which balanced the load for DCs in CC [16]. This approach had 2 sections in which the centralized route was computed and the decision related to distributed route was made to ensure that it worked flexibly and stably in gigantic networks. The In-band Network Telemetry had helped to extract the precise information regarding link position. Moreover, the weighted Equal-Cost Multi-Path was implemented in a simple and effective system at the edge of fabric which was suitable in the introduced system which mapped the flows onto an adequate route and avoided the great congestion taken place in a single link. The introduced system offered higher FCT of 0.70 load as compared to traditional methods.

Karan D. Patel, et.al (2019) established an approach which balanced the load by integrating two methods. The fundamental emphasis of this algorithm was on balancing the workload over CS [17]. The extended version of HBB based algorithmic method was put forward for executing tasks based on priority and normal one depending upon an extended WRR. This algorithm was useful for augmenting the efficacy of the established approach, utilize the resources effectively and mitigate the time to complete task. Moreover, the established algorithm had performed well concerning lower completing time and higher efficiency.

III. RESEARCH METHODOLOGY

For managing node problems occurred in the cloud organizations, this work suggests a heuristic approach known as Butterfly Optimization (BFO). This algorithm is composed of a variety of nodes. The execution time (ET) and failure rate (FR) are considered for analyzing the participant node among several nodes. An expert node is employed for declaring a fixed value or threshold. Two parameters namely ET and FR are employed in this approach. The nodes which consume lower time to execute and supports least failure rate are selected as suitable by the expert nodes. Different from the established threshold, the value of node N1 is found lower. The executed step is the basis to select this mote as the opponent node. N2 is consisted of higher as well as lower parameters. Therefore, this node is unsuitable for a competitor node. The next N3 is selected as the opponent node due to its similar value to fixed limit. Moreover, the impossible task is of selecting N4 as applicant node in case of its larger value over the set threshold. After verifying the competitor node, this node is responsible to carry out its task. Such circumstances allow to initialize diverse errands. A node is capable of moving from its position after completing a task. However, the task is failed. Therefore, this work focuses on developing a novel approach for tackling the issue related to failure event as the node has a dynamic nature. This approach emphasizes on other parameter known as period of master node. The last task is of discovering end clients which can be employed as a master node for collaborating with other nodes. The time period of master node is computed in the below given format.

1. E-cost= maximal execution time + Time required by master node (master node time)

Then, we will compute each node's profit.

2. Profit of each node = E-cost+ Failure time of each node

3. Weight of each node= No. of tasks + maximal execution time/Profit

The node with highest weight is selected. The provided formula is used for the weight measurement.

The envisioned algorithm takes the following steps:

Step 1: Get list of all VMs working on all hosts.

Step 2: Initialize no migration is performed.

Step 3: Get resource consumption, failure rate, and execution time of all machines.

Step 4: Built transition matrix for hosts and VMs.

Step 5: Loop will execute until all machines on over utilized hosts are migrated.

Step 5.1: Calculate the current utilization of each host for that particular VM that needs migration.

Step 5.2: Check creation history of the VM.

Step 5.3: Compare increase in utilization of selected hosts with other hosts.

Step 5.4: Select host for which increase in utilization is minimum End loop

Step 5.6: If maximum utilization exceeds upper utilization threshold go to step 5.1.

Step 6: Else choose that particular host for migration.

Step 7: return migration List

END

IV. RESULT AND DISCUSSION

MATLAB (Matrix Laboratory) is extensive language and platform. A matrix software is created by LINPACK and EISPACK for this language and make its accessibility easy. This platform is employed for computing mathematically and programming problems. Thus, it becomes a popular tool among researchers for analyzing the data and formulating the algorithm-based techniques with models. This tool majorly emphasizes on computing, visual and programming environments. Hence, this language becomes an effective for accomplishing any research and in development area.

Fig 1: Virtual Machine Migration using ACO algorithm

Figure 1 illustrates that choice of VM with maximum value is done as an optimal VM. Ant Colony Optimization algorithm is executed for migrating the tasks of this machine.

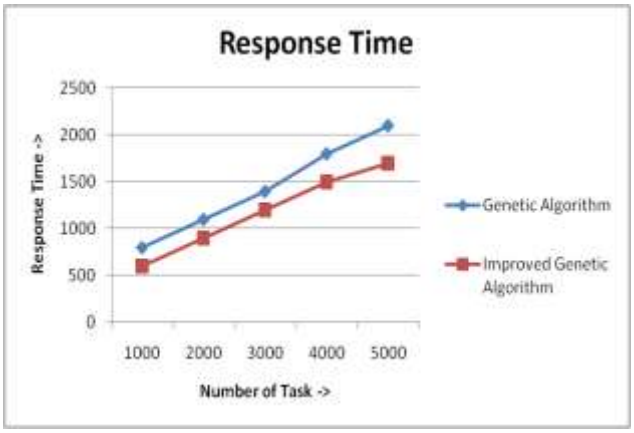

Fig 2: Response Time based comparison

Figure 2 demonstrates the comparative analysis of the suggested algorithm (IGA) with the conventional algorithm with regard to response time. The suggested algorithm offers lower response period in contrast to the existing one.

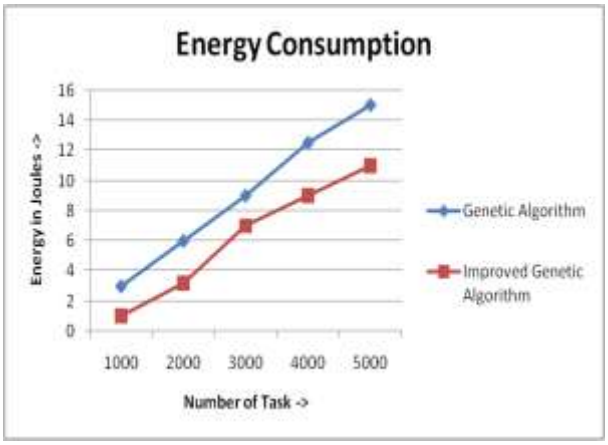

Fig 3: Consumed Energy based Comparative Analysis

Figure 3 depicts the comparative analysis of the improved Genetic Algorithm with existing algorithm (GA) with regard to energy usage. The suggested algorithm offers higher energy efficacy in contrast to the traditional algorithm.

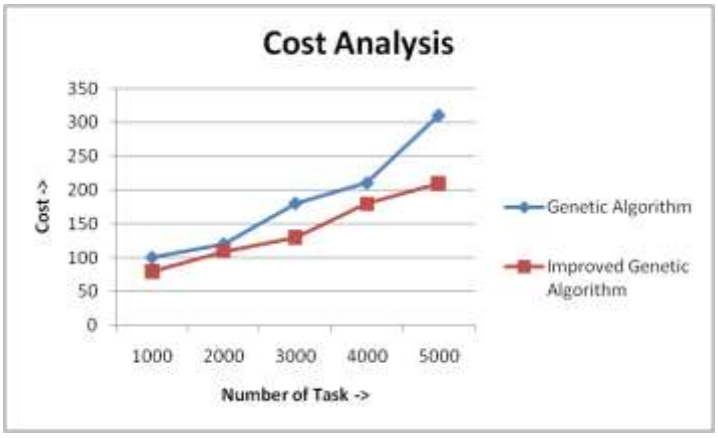

Fig 4: Cost driven Comparative Analysis

Figure 4 represents the comparative analysis of improved Genetic Algorithm with existing algorithm (GA) with respect to cost. Unlike the existing algorithm, the suggested algorithm consumes least cost.

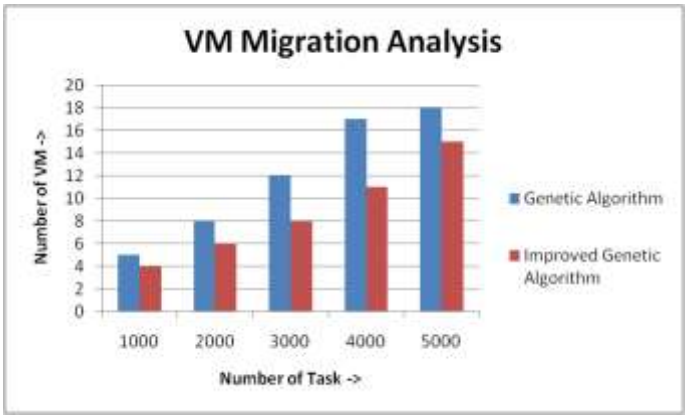

Fig 5: Migrations based Comparative Analysis

Figure 5 demonstrates the comparative analysis of the improved Genetic Algorithm with existing algorithm (GA) concerning the number of migrations. The suggested algorithm attains least migration rate in comparison with the existing one.

V. CONCLUSION

The presented project is based on the task to balance load in CC system. The inefficient way to balance load results in maximizing the delays. The earlier work has deployed GAs to migrate the virtual machine (VM). However, this algorithm is much complex due to which it consumes larger time to migrate VM. This work emphasizes on developing an improved Genetic Algorithm (IGA) for migrating VM. The developed algorithm was deployed in MATLAB. This algorithm is quantified with regard to different metrics. The experiments confirmed the superior efficacy of the introduced approach over the traditional approaches.

VI. REFERENCES

[1] Amandeep Kaur Sidhu, SupriyaKinger, “Analysis of Load Balancing Techniques in Cloud Computing”, 2013, International Journal of Computers & Technology, vol. 4, no. 2

[2] Po-Huei Liang1 and Jiann-Min Yang, “Evaluation of Two-Level Global Load Balancing Framework in Cloud Environment”, 2015, International Journal of Computer Science & Information Technology (IJCSIT), vol. 7, no. 2

[3] Abhishek Kumar Tiwari, Sreeja Nair, “A Review of Load Balancing Technique of Cloud Computing Using Swarm Intelligence”, 2016, International Journal of Computer Science and Information Technologies, vol. 7, no. 2, pp. 741-744

[4] Srinivasa Rao Gundu & T. Anuradha, “Improved Hybrid Algorithm Approach based Load Balancing Technique in Cloud Computing”, 2019, Global Journal of Computer Science and Technology, vol. 19 no. 2

[5] Alok singh, Vikas Kumar Tiwari, Dr. Bhupesh Gour, “A Survey on Load Balancing in Cloud Computing Using Soft Computing Techniques”, 2014, International Journal of Advanced Research in Computer and Communication Engineering vol. 3, no. 9

[6] Saher Manaseer, MetibAlzghoul, Mazen Mohmad, “An Advanced Algorithm for Load Balancing in Cloud Computing using MEMA Technique”, 2019, International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8, no. 3

[7] Dalia Abdulkareem Shafiq, Noor Zaman Jhanjhi, Azween Abdullah, Mohammed A. Alzain, “A Load Balancing Algorithm for the Data Centres to Optimize Cloud Computing Applications”, 2021, IEEE Access

[8] Nithin K. C. Das, Melvin S. George, P. Jaya, “Incorporating weighted round robin in honeybee algorithm for enhanced load balancing in cloud environment”, 2017, International Conference on Communication and Signal Processing (ICCSP)

[9] M Jeyakarthic, N. Subalakshmi, “Client Side-Server Side Load Balancing with Grasshopper optimization Mapreduce Enhancing Accuracy in Cloud Environment”, 2020, Fourth

International Conference on Inventive Systems and Control (ICISC)

[10] Dharavath Ramesh, Sweta Dey, "SCLBA-CC: Slot Based Carton Load Balancing Approach for Cloud Environment", 2018, International Conference on Current Trends towards Converging Technologies (ICCTCT)

[11] M. Lawanyashri, Balamurugan Balusamy, S. Subha, "Energy-aware hybrid fruitfly optimization for load balancing in cloud environments for EHR applications", 2017, Informatics in Medicine Unlocked

[12] A. Francis Saviour Devaraj, Mohamed Elhoseny, K. Shankar, "Hybridization of firefly and Improved Multi-Objective Particle Swarm Optimization algorithm for energy efficient load balancing in Cloud Computing environments", 2020, Journal of Parallel and Distributed Computing

[13] Lung-Hsuan Hung, Chih-Hung Wu, Chiung-Hui Tsai, Hsiang-Cheh Huang, "Migration-Based Load Balance of Virtual Machine Servers in Cloud Computing by Load Prediction Using Genetic-Based Methods", 2021, IEEE Access

[14] Muhammad Junaid, Adnan Sohail, Rao Naveed Bin Rais, Adeel Ahmed, Osman Khalid, Imran Ali Khan, Syed Sajid Hussain, Naveed Ejaz, "Modeling an Optimized Approach for Load Balancing in Cloud", 2020, IEEE Access

[15] H. Lathashree, Niveditha J. MokaKatte, K. P. Pooja, B. Sathish Babu, K. Bhargavi, "Multiple CAs based Framework with ANN for Load Balancing in Cloud Computing Environment", 2018, 4th International Conference for Convergence in Technology (I2CT)

[16] Zixi Cui, Pengshuai Cui, Yuxiang Hu, Julong Lan, Fang Dong, Yunjie Gu, Saifeng Hou, "Closer: Scalable load balancing mechanism for cloud datacenters", 2021, China Communications

[17] Karan D. Patel, Tosal M. Bhalodia, "An Efficient Dynamic Load Balancing Algorithm for Virtual Machine in Cloud Computing", 2019, International Conference on Intelligent Computing and Control Systems (ICCS)

Acknowledgement. Words cannot express my gratitude to Dr. Anoop Sharma my guide & my mentor who always stands besides me to support in writing this paper. I am also grateful to my cousin Sanjeev Gupta*** & my family members for their trust in me & for giving me emotional and moral support throughout my journey of this research.

* Arun Gupta, Research Scholar, Department of Computer Applications, Singhania University

** Dr. Anoop Sharma, Associate Professor, Department of Computer Science, Singhania University

*** Sanjeev Gupta, Research Scholar, Department of Economics, Singhania University

Arun Gupta
MCA, MBA, LL.B
Research Scholar - Computer Applications
[Singhania University, Pacheri Bari, Jhunjhunu (Raj.), India]
Research Area: Evaluate & Implement Load Balancing Framework for Cloud Computing