International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5S3, July 2019

The Hybrid Optimization Algorithm for Load Balancing in Cloud

Pooja Arora, Anurag Dixit

Abstract: The advancements in the cloud computing has gained the attention of several researchers to provide on-demand network access to users with shared resources. Cloud computing is important a research direction that can provide platforms and softwares to clients using internet. But, handling huge number of tasks in cloud infrastructure is a complicated task. Thus, it needs a load balancing method for allocating tasks to Virtual Machines (VMs) without influencing system performance. This paper proposes a load balancing technique, named Elephant Herd Grey Wolf Optimization (EHGWO) for balancing the loads. The proposed EHGWO is designed by integrating Elephant Herding Optimization (EHO) in Grey Wolf Optimizer (GWO) for selecting the optimal VMs for reallocation based on newly devised fitness function. The proposed load balancing technique considers different parameters of VMs and PMs for selecting the tasks to initiate the reallocation for load balancing. Here, two pick factors, named Task Pick Factor (TPF) and VM Pick Factor (VPF), are considered for allocating the tasks to balance the loads.

Keywords: Cloud computing, load balancing, Elephant Herding Optimization, Grey Wolf Optimizer, reallocation, pitch factors.

I. INTRODUCTION

Cloud computing attracts the researchers due to advancements offered in the service-oriented architectures. The huge scale applications devised in cloud computing generates huge number of tasks with large amount of workloads [17] [6]. The cloud computing is an advanced computing paradigm that permits the usage of computing infrastructures at different levels of abstractions as an on demand service, which is accessible from internet or other computer network. Cloud computing gains more attention due to its easy availability and greater flexibility at low cost [11].But, the nodes in cloud computing tends to be underutilized while others seem to be overloaded due to uneven computing capacities of nodes, different task scale and unbalanced load distribution [15] [6].Thus, it becomes essential for spreading the loads among computing nodes for taking full benefits of cloud computing system and for improving the satisfaction of users [16][6]. As a widespread commercial paradigm, the cloud computing attracts the attention of users from both industrial communities and academics. Hence, the development of sophisticated cloud computing applications allowed several enterprises and individuals for outsourcing the significant data to the cloud rather than constructing and handling local data centers. The cloud users are benefitted from different types of computing services due to the advancements in different types of computing services provided by the public cloud [12]. The National Institute of Science and Technology (NIST)defines cloud computing as a model, which enables convenient and on-demand network access for devising a shared pool of configurable computing resources, such as servers, applications, networks, services, and storage, which can be quickly provisioned and released with less efforts and interactions [6].

Due to different services, the cloud computing provides simple and flexible mechanisms for keeping data and files to make them accessible for huge scale users [9]. If large amounts of tasks occur on specific node, the tasks can be switched from overloaded nodes to underloaded nodes to minimize the tasks waiting time at nodes which is also termed as Load Balancing (LB). Moreover, allocation of task and balancing the loads are important for distributed systems [16].The purpose of LB is to understand the requirements of consumers by sending and receiving the data and information without taking much time. Load balancing is utilized for redistributing workloads into the cloud by designing different strategies to maximize the utilization of resource by avoiding the overhead [14] [16].LB in cloud computing is considered as one of the major challenge. Without balancing the loads, cloud users may lead to delay and facilitates time-consuming responses. Here, the load can be considered as a memory load, CPU loads, or a network load. The LB plays a vital role to increase the performance of Distributed System (DS) [24] [25]. This DS can help to transfer the load into different processors of DS for improving the response time of tasks and utilization of resource by neglecting the condition, in which fewer processors are overloaded while other processors are in waiting state or performing underloaded work at a specific time allocated to the system. The LB is utilized for computing different terms, which involves communication delay, execution time, resource utilization, and throughput [10]. A distinct distributed system, like Expert Cloud involves different distributed Human Resources (HR's). This HR's is responsible for connecting other HRs for attaining high performance and executes the tasks in an effective manner. For reducing the time for execution, the workload is likely to be distributed depending on the strength of HRs, which makes the load balancing essential. The goal of balancing load is to provide networks and resources by offering increased throughput with less response time. The resultant data obtained by dividing the traffic among users is sent and received efficiently without causing a delay [13] [4].

Manuscript published on 30 July 2019.
***** Correspondence Author (s)
Pooja Arora, IT Dept., BCIIT, Delhi, India.
Anurag Dixit, SCSE,Galgotias University, Delhi, India.

Also, load balancing provides several methods for maximizing the throughput, resource utilization, and system performance.

Several load balancing algorithms are devised for making use of resource in cloud system in an efficient manner. Thus, the load balancing algorithms are divided into two major types, which involve static and dynamic based on the decision of load balancing depending on current load or not. The static algorithms are not able to adjust dynamic changes of the system once the execution of task begins [1].The dynamic algorithms can balance decisions with respect to the status of load at runtime and can adjust accordingly for redistributing the essential workloads, which lead to improved performance on dynamic algorithms. The two main mechanisms covered by the dynamic load balancing algorithms are diffusion methods and meta-heuristic-based methods. The combined strategy penetrations of some agents on others are referred as a diffusion mechanism [18][6]. The interface between neighboring nodes within the cloud can be abstracted as the diffusion process [19][6].

But, the diffusion method is responsible only for attaining locally optimal result. Metaheuristic- based approaches, likeGenetic Algorithm (GA), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO), and so on, are used to offer near-optimal solutions in specific time. In [20], a meta-heuristic algorithm, named Grey Wolf Optimizer (GWO) is devised, which is motivated from the behavior of grey wolves. Thus, the GWO algorithm imitates the leadership hierarchy and adapts a hunting method of grey wolves in nature. Here, four kinds of grey wolves, like alpha, beta, delta, and omega, are adapted to simulate the leadership hierarchy. Moreover, an advanced swarm-based metaheuristic search method, named Elephant Herding Optimization (EHO) [21] is devised to solve the optimization tasks. Here, the EHO is motivated from the elephant's herding behavior. In EHO, the elephants from each group are updated using its current position and matriarch by clan updating operator. But, the performance of the incorporation highly depends on matching levels between the system model and meta-heuristic techniques.

II. LITERATURE SURVEY

Authors	Methods	Advantages	Disadvantages
Shang-Liang Chen *et al.* [1]	Cloud Load Balance (CLB) algorithm	balance the loading performance when users logged in at the same time	The method failed to consider different load approach
Qi Liu *et al.* [2]	Hadoop-LB with prediction model based on K-ELM (PMK-ELM)	reduce the running time, high accuracy	waste more storage space
Jia Zhao *et al.* [3]	Load Balancing based on Bayes and Clustering (LB-BC)	improved throughput	The method failed to apply LAN
Shiva Razzaghzadeh *et al.* [4]	load balancing strategy	improves the makespan and cost	The method failed to extend load balancing for dependent tasks and failed to consider more factors such as fault tolerance for HR label
Ranesh Kumar Naha and Mohamed Othman [5]	Cost aware brokering and Load aware brokering algorithm	minimized the cost	The method failed to consider efficient load balancing algorithm for minimizing the execution time
Weihua Huang *et al.* [6]	Fuzzy clustering method with Feature Weight Preferences for Load Balancing (FWPFC-LB)	achieve better load balancing performance	The method failed to consider adaptive parameter optimization for data fusion and convergence rate acceleration in clustering
Narander Kumar and Diksha Shukla [7]	fuzzy row penalty method	increase performance and scalability, minimize associated overheads, takes less execution time	workload is not distributed properly
Santanu Dam *et al.* [8]	Ant-Colony-Based Meta-Heuristic Approach	minimize the make span as well as the number of Virtual Machine (VM) is also reduced	The method failed to include fault tolerance and priority of the job

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

In this section, the survey of eight existing techniques based on LB is elaborated along with their drawbacks.

Shang-Liang Chen *et al*. [1] developed a dynamic annexed balance method for solving the problem caused by uneven distribution of loads. Here, the LB in cloud is taken into consideration for processing server power and for computer loading. Thus, the server can manage extreme computational requirements. At last, two algorithms in load balancing are tested with experiments for proving the approach to be innovative. The method can balance the load performance when the users are logged at same time, but the method is not applicable for different load methods. Qi Liu *et al.* [2] devised an adaptive scheme, named Hadoop-LB using Prediction Model based on Kernel function-Extreme Learning Machine (PMK-ELM) algorithm to attain time efficiency by offering heterogeneous cloud infrastructure. A dynamic speculative execution strategy on real-time management for cluster resources is derived for optimizing the execution time of map phase and a prediction model is utilized for fast prediction with minimum task execution time. An adaptive solution is devised for optimizing the performance of space-time by combining the prediction model with a multi-objective optimization algorithm, but the method requires more storage space. Jia Zhao *et al.* [3] devised an advanced approach, named Load Balancing based on Bayes and Clustering (LB-BC), for deploying requested tasks to the hosts. LB-BC uses limited constraint for physical hosts to attain a task deployment approach with global search capability using performance function for computing the resources. Here, the Bayes theorem is integrated with the clustering process for obtaining optimal clustering set with a physical host. However, the method is unable to work on real computing environment. Shiva Razzaghzadeh *et al.* [4] developed a method for distributing the dynamic load using distributed queues in cloud infrastructure. This method maps the tasks and HRs by allocating label to each HR. The load balancing and mapping process are devised on the basis of Poisson and exponential distribution. This method permits the task allocation process to execute with high power by adapting distributed queues aware of the service qualities. The method did not use genetic algorithm for determining the optimal HR by resolving the faults. Ranesh Kumar Naha and Mohamed Othman [5] developed an algorithm, named Cost aware brokering and Load aware brokering algorithm, for balancing the loads in data centers and the VM. This algorithm minimizes the overall processing and response times as the tasks are assigned to the available physical resources in an effective manner. The algorithm did not consider real-world cloud brokering for the evaluation. Weihua Huang *et al.* [6] developed a fuzzy clustering method using feature weight preferences for overcoming load balancing issues in multiclass system resources and can attain optimal solution for load balancing by load data fusion. Here, feature weight preferences are utilized for establishing the relationship between specific cloud scenarios and LB procedure. The method was not applicable for optimizing the adaptive parameter in load data fusion and to improve the convergence rate acceleration in clustering. Narander Kumar and Diksha Shukla [7] devised a method, named fuzzy row penalty method, to solve the issues of LB in a cloud computing environment based on fuzzy. Here, the fuzzy technique is utilized for addressing uncertain response time in fuzzy cloud environment. Here, the fuzzy row penalty method was utilized to address the balanced fuzzy LB problem and unbalanced fuzzy LB problem in a cloud infrastructure.

The generated result is utilized to solve the LB issues based on response time and space complexities for maximizing the performance, minimizing the scalability, overheads, but workload distribution is improper. Santanu Dam *et al.* [8] utilized an advanced computational intelligence technique, named Ant Colony Optimization (ACO), to balance the loads among VM in cloud computing. Here, the ACO is utilized for designing an intelligent multi-agent system by collective behavior of ants. Here, the ACO is utilized for addressing the LB problem in cloud computing and to minimize the makespan and to minimize the VM usage, but fault tolerance and handling high priority job are still considered as major issues.

III. CHALLENGES

The challenges faced by the existing techniques for LB are illustrated as follows:

- LB is considered as a major challenge in cloud computing, as it needs to distribute the works equally among different nodes for attaining high satisfaction among users with improved resource utilization [8].
- Resource time for distributed system is a major issue for LB algorithms. In load balancing algorithm, dynamic algorithm acquires more response time, whereas static algorithm takes less time for generating responses [10].
- Cloud computing induces many advantages for service-oriented computation. But balancing the load for multi-class system is challenging due to complication produced in load data fusion [6].
- The major challenge faced in the cloud computing system is to avoid the issues while, the load tends to be dynamically distributed through VM [9].
- Several challenges exist in LB algorithms that include security issues, interoperability issues, data management issues, and Quality of Service (QoS) performance issues. In LB algorithms, all these parameters must be improved to enhance the performance of whole system [10].

In [6], Feature Weight Preferences Fuzzy Clustering-Load Balancing (FWPFC-LB) method is developed for attaining optimal LB among the VM. This method has the capability to balance the multi-class loads, which include memory, network, central processing unit, and resources, but balancing the load for multi-class system is a major challenge and initiates difficulties in load data fusion..

IV. PROPOSED METHODOLOGY

The primary intention of this research is to design and develop a technique for load balancing to improve the efficiency of the cloud computing system. Here, load balancing is performed to remove the tasks from over loaded VMs and assigning them to under loaded VMs without affecting the system performance. Accordingly, Elephant Herding-based Grey Wolf Optimizer (EHGWO) will be newly developed algorithm to perform the load balancing. Initially, the capacity and loads of the virtual machine will be found based on the executed tasks, then, the balance of the cloud system will be checked.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

When the load is found unbalanced, the capacity with load will be checked to take the decision whether load balancing can be done or not. In the other case, find the tasks to be removed by checking the two constraints like, load of the task and cost of the load balancing. Then, the removed tasks will be added in other VMs by optimally finding the VMs for the task execution. The optimal finding of VMS for executing of the removed task will be found out using the proposed EHGWO, which will be designed by combining Elephant Herding Optimization (EHO) [21] and Grey Wolf Optimizer (GWO) [20]. Figure 1 depicts the proposed EHGWO based load balancing method in Cloud computing platform. The implementation of the proposed approach will be in JAVA with Cloudsim tool. The performance of the proposed technique for load balancing will be evaluated with different cloud set up for makespan, and the results attained will be compared with that of existing works [1] [2] and [3].

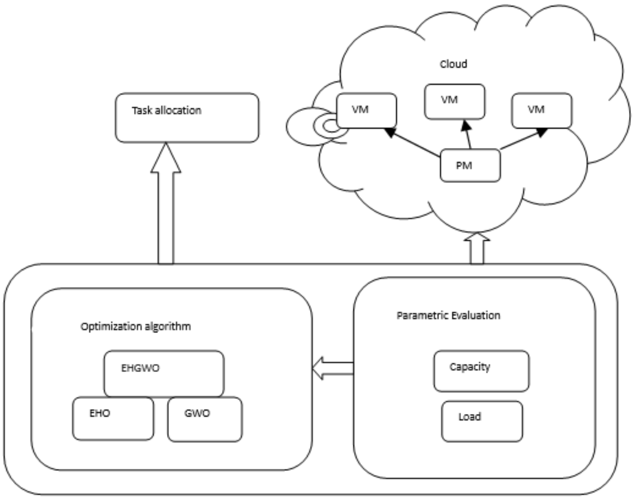

Figure 1. Block diagram of the proposed EHGWO based load balancing method in Cloud computing platform

V. EXPECTED OUTCOME

In [3], the performance of the LB-BC method is evaluated using makespan as one of the metrics, wherein the minimum makespan is approximately 550 sec. However, the proposed load balancing method will be implemented in such a way to reduce the makespan than that in [3].

REFERENCES

1. Shang-Liang Chen, Yun-Yao Chen , Suang-Hong Kuo," CLB: A novel load balancing architecture and algorithm for cloud services", Computers & Electrical Engineering,Vol:58, pp:154-160, February 2017.
2. Qi Liu, Weidong Cai, Jian Shen, Xiaodong Liu, Nigel Linge, "An Adaptive Approach to Better Load Balancing in a Consumer-centric Cloud Environment," IEEE Transactions on Consumer Electronics, Vol: 62, no: 3, pp: 243 - 250, August 2016.
3. 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 Transactions on Parallel and Distributed Systems, Vol: 27, no: 2, pp: 305 -316, February 2016.
4. Shiva Razzaghzadeh, Ahmad Habibizad Navin, Amir Masoud Rahmani, and Mehdi Hosseinzadeh, "Probabilistic Modeling to Achieve Load balancing in Expert Clouds", Ad Hoc Networks, January 2017.
5. Ranesh Kumar Naha and Mohamed Othman, "Cost aware service brokering and performance sentient load balancing algorithms in the cloud", Journal of Network and Computer Applications, Vol: 75, pp: 47–57, November 2016.
6. Weihua Huang, Zhong Ma, Xinfa Dai, Mingdi Xu and Yi Gao, "Fuzzy Clustering with Feature Weight Preferences for Load Balancing in Cloud", International Journal of Software Engineering and Knowledge Engineering,Vol: 28, no:5,pp:593–617,2018.
7. Narander Kumar and Diksha Shukla," Load Balancing Mechanism Using Fuzzy Row Penalty Method in Cloud Computing Environment", Information and Communication Technology for Sustainable Development, pp: 365-373,2017.
8. Santanu Dam, Gopa Mandal, Kousik Dasgupta and Parmartha Dutta," An Ant-Colony-Based Meta-Heuristic Approach for Load Balancing in Cloud Computing", Applied Computational Intelligence and Soft Computing in Engineering, pp: 29,2018.
9. Shalini Joshi and Uma Kumari, "Load Balancing in Cloud Computing:Challenges & Issues," In proceedings of 2nd International Conference on Contemporary Computing and Informatics (IC3I), December 2016.
10. Mahfooz Alam and Zaki Ahmad Khan," Issues and Challenges of Load Balancing Algorithm in Cloud Computing Environment", Indian Journal of Science and Technology, Vol: 10, no:25, July 2017.
11. Wayne Jansen and Timothy Grance, "Guidelines on security and privacy in public cloud computing",pp: 800-144, 2011.
12. Jianting Ning, Zhenfu Cao, Xiaolei Dong, Kaitai Liang, Hui Ma and Lifei Wei," Auditable -Time Outsourced Attribute-Based Encryption for Access Control in Cloud Computing",IEEE Transactions on Information Forensics And Security, 2017.
13. Zenon Chaczko, Venkatesh Mahadevan, Shahrzad Aslanzadeh and Christopher Mcdermid," Availability and Load Balancing in Cloud Computing", In proceedings of International Conference in Computer and Software Modeling, IPCSIT, Vol:14 , 2011.
14. Daraghmi, E.Y. and Yuan, S.M.,"A small world based overlay network for improving dynamic load-balancing", Journal of Systems and Software, Vol: 107, pp:187-203,2015.
15. A. S. Milani and N. J. Navimipour," Load balancing mechanisms and techniques in the cloud environments: Systematic literature review and future trends", J. Netw. Comput. Appl. Vol:71, no:1, pp: 86–98,2016.
16. Y. C. Jiang, "A survey of task allocation and load balancing in distributed systems", IEEE Trans. Parallel Distrib. Syst.,Vol: 27no:2,pp: 585–599,2016.
17. G. Mateusz, G. Alicja and B. Pascal," Cloud brokering: Current practices and upcoming challenges", IEEE Cloud Comput.,Vol: 2,no:2,pp: 40–47,2015.
18. I. D. Falco, E. Laskowski, R. Olejnik, U. Scafuri, E. Tarantino and M. Tudruj, "Extremal optimization applied to load balancing in execution of distributed programs", Appl. Soft Comput.,Vol: 30,no:3,pp: 501–513,2015.
19. Y. Jiang, "Concurrent collective strategy diffusion of multi agents: The spatial model and case study", J. Netw. Comput. Appl.,Vol: 39,no:4,pp: 448–458,2009.
20. Mirjalili, S., Mirjalili, S.M. and Lewis, A.,"Grey wolf optimizer" Advances in engineering software,Vol: 69, pp.46-61,2014.
21. G. Wang, S. Deb and L. d. S. Coelho, "Elephant Herding Optimization, "In proceedings of 3rd International Symposium on Computational and Business Intelligence (ISCBI),pp:1-5,2015.

AUTHORS PROFILE

Ms. Pooja Arora , Assitant Professor in Banarsidas Chandiwala Institute of Information Technology(affiliated to GGSIPU, Delhi). She is pursuing Ph.D in Computer Science from Galgotias University.She has done M.Tech(CSE) from School of Computer Science & Engineering from G.G.S.I.P.U Delhi. She has total experience of 11 years.

Dr. Anurag Dixit, Professor in School of Computer Science and Engineering Galgotias University.He has done Ph.D in Computer Science from Jawaharlal Nehru University, New Delhi and M.Tech in Communication Engineering.His Research Area is Soft Computing E-learning and Software Engineering. He has total experience of 17+ years.

Retrieval Number: E10170785S319/2019©BEIESP
DOI:10.35940/ijeat.E1017.0785S319
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)