Machine learning & Deep learning based Load Balancing Algorithms techniques in Cloud Computing

Ronak Agarwal
GLA University
Mathura, India
ronakagrawal0504@gmail.com

Dilip Kumar Sharma
GLA University
Mathura, India
dilip.sharma@gla.ac.in

2021 International Conference on Innovative Practices in Technology and Management (ICIPTM) | 978-1-6654-2530-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIPTM52218.2021.9388349

***Abstract*— Cloud computing is a fledgling technology that provides computers in terms of efficient resource efficiency. The advancement of new applications in the cloud brings on an increase in load on servers. In addition to the many studies conducted in the province of Cloud computing, it still has to face some provocations, such as load balancing(LB) that require further upgrades and models for better performance and to achieve maximum user satisfaction. In this paper, we have discussed about the Machine Learning Techniques used to build VM teams according to their CPU. Deep Learning (DL) techniques are used to achieve QoS which implies improving of resource utilization and deployment; while reducing delays, response time and costs, balancing the load on all equipment, therefore, increases system reliability. This paper contains a comprehensive overview of a few machine learning as well as deep learning based algorithm on load balancing.**

Keywords—Load balancing, Deep Learning, Machine learning, Virtualization, response time

I. INTRODUCTION

Cloud computing is ascending as a language in the territory of superior distributing and computing condition. It gives clients on-request administrations to a shared pool of assets over the Internet as a self-administration. It is progressively scalable. Cloud computing is in its initial ages, to secure its full advantages, much research has been done and is to be done over an expansive region of subjects. Cloud computing is emerging as the most amplifying web technologies. Cloud Computing has doubled its existence. It vigorously apportions the assignment on the virtual machine which is selectively available to the client. One of the fundamental difficulties in the field of cloud computing is load balancing. Load adjusting is a particular term characterized for adjusting the load heap among the multiprocessors. The server's workload can be arranged into different classifications like CPU's load, net workload, issue of memory limit and so forth. Loadbalancing mechanisms are utilized to impart the workload among various virtual machines which directly goes to the assistance of virtual machine manager(VMM). Cloud data centers are generally intended to deal with a lot of load that may prompt low resource utilization and furthermore more energy wastage. Resource utilization prompts consumption of energy, and hence, it needs to be optimized in order to abstain from energy wastage.[6] In cloud computing, load adjusting systems are required for dissemination of workload between cloud centers to avoid a condition of certain hubs being over loaded and others being under loaded or perhaps idle. The load needs to be mapped to different resources underneath the requirement of improvement in energy consumption[7]

In section II, we have given a brief description on load balancing and we have clarified why deep reinforcementbased and machine learning algorithms are needed for load optimization.

In section III, we have stated the parameters to be acknowledged while designing a load balancing algorithm.

In section IV, we have given a brief survey on the existing machine learning and deep learning based load balancing algorithms followed by a comparison chart.

In section V, we have discussed the challenges faced in load balancing and in section VI there is a description about the tools and the real world platform used in general cloud computing algorithms.

II. LOAD BALANCING IN CLOUD COMPUTING

Cloud computing is certainly the most rapidly used for the past ten years. Many companies exist attempts to use and set up clouds, due to complex structures and shapes. This leads to a file for a growing number of users accessing the cloud.Cloud computing has been mainstreamed by many organizations such associal networking websites, apps made online by Google doc and Google app administrators for one of the most important initiatives and a step forward in cloud computing.[8]by concluding all the factors we can say that cloud computing is abruptly changing the way we communicate with onlinereserves.Cloud models considers virtualization technology; this Information helps to build a single data center or higher capacity server to act as integrated equipment. Depending on data center hardware or server pattern how possible the visible machine can be separated.

Load balancers are important for both the users and providers of cloud services in aspect for achieving higher utility of resources, while ultimately requiring reduced costs, execution time and VM makespan.

978-0-7381-3289-1/21/$31.00 ©2021 IEEE

Deep training is a way of looking at and remembering things with training examples. This process is stimulated by complex structures of the brain's senses and their connections. DL is used when data acquisition is large enough to train an event or cause. In traditional machine learning, inputs are followed by the output of a feature that includes NP-hard mathematical equations and functions and the separation of the extracted elements is performed.[1] From machine learning, there are two distinct categories (Release element and segregation categories) that cause low efficiency and low efficiency in achieving good results. Howsoever, in the case of in-depth learning, these two categories are integrated into a single phase in which in-depth reading skills work to extract and differentiate with different hidden layers.

DL results in efficient and accurate decision-making for the allocation of resources to incoming programs, thus selecting the most appropriate application to complete them. DL strategies lead to effective decision-making for resource allocation (VMs) and achieving load balancing on existing VMs. In-depth learning modes provide the use of resources in the cloud environment by facilitating the provision of resources (intelligent resource allocation) while selecting the most appropriate and available resources.

Deep reinforcement supervised Techniques are used to solve problems that can be explained mathematically. Heuristic strategy parameters are developed and modified to improve algorithm performance. In a dynamic environment such as a computer, price limits vary over time - a never-ending process of heuristic and meta-heuristic approaches to planning and measuring a load, therefore ML or Deep Learning-based techniques are used to achieve relevant results effectively. These approaches are able to provide a more straightforward model and systematic consideration. [1] Also, ML strategies improve decision-making accuracy in heuristics algorithms [4]. The deadline set sets the rules for that task so that only those tasks are performed by completed VMs before the end of their deadline.

III. MATRICS OF LOAD BALANCING

A. Through Put: Number of the assignments which have been executed within a given amount of time. The time ought to have the maximum throughput for expanding the selectiveness of the system.[6]

B. Overhead: The overhead should be least for better execution of the algorithm. It is a proportion of the working expense at the time of execution of the given load adjusting algorithm.

C. Fault Tolerance: A parameter for fault tolerance for managing the breakdown happened in the node authentically and continuously. The nodes are changed on account of defective nodes.

D. Transfer Time: It is the time unit for the development and reallocation of the assets starting with one node then onto the next node. Transfer time ought to be limited for accomplishingbetter viability of the system.

E. Reaction Time: The measure of time taken in a load balancing structure to respond in the load adjusting disseminated cloud computing condition. This time length has to be minimum.

F. Resource Utilization: This parameter guarantees the usage of the assets in the proposed framework. This parameter ought to be accomplished the most in the load balancing condition

G. Scalability: It indicates the inclination of the system to comprehend the strategy for load reconciliation utilizing a minimal number of CPUs [a load adjusting technique]. Thisparameter is advanced for a ton of engaging framework execution.

H. Makes Span: It is the total time required to execute all assignments stated to the framework. Makespan is taken as the most extreme time taken in the execution of the cloudlets running on the data centers.

The above measurements should be satisfied with the executive execution of the framework Number of the assignments which have been executed within a given amount of time. The time ought to have the maximum throughput for expanding the selectiveness of the system.

IV. SURVEY OF EXISTING MACHINE LEARNING BASED LBA ALGORITHM

The purpose of this paper [8] is to propose a more efficient and advanced algorithm that can keep the load and provide the distribution of modified resource strategies. In this paper the measured load is measured using the Throttled and Equally Spread Current Execution algorithm (ESCE) algorithms. This paper, a newly developed and functional Hybrid algorithm redesign is proposed and implemented in a computer environment using the CloudSim toolbar, in the Java language. The research work involves improving the efficiency of the Hybrid VM to measure cloud algorithm and performing comparative analyzes of the proposed algorithm with existing algorithms. Moderate delivery involves studying existing VM measurement techniques, proposing an effective VM measurement algorithm, using an algorithm in Cloud Analyst, and comparing the proposed algorithm with existing existing targeted technology parameters.

In this paper[9], authors developed three predictive models using General Modification, Neural Network and Vector

Support Dismantling of two TPC-W web applications. Asides from One traditional metric predict using CPU, theyadded two SLA metrics; response time and passing on prediction model. Support Vector Regression (SVR) modelshowed higher predictive accuracy than Neural Network and Linear Regression in 9 to 12 minute windows. From three models, SVR is the most preferred predictive model. As a result, cloud customers can use SVR to build their own forecasting models. In addition, the SLA business standard increases metrics (response time and entry) in the forecast model opens the path of the matrix for the resolution of a triple combination of allocation of flexible resources; direct amplification of VM infrastructure. This research has been proved very useful as a response time as well login may be delayed long before the request is received its CPU usage limit.

Cloud Computing is the delivery of computer services - storage, databases, communications and online software. The Cloud service provider provides the payment service on a per-use basis. Upload a balancing process to make reliable resource use by redistributing total load on individual nodes. LB is basically divided into two types such as static LB, Suitable for a flexible environment where the load variability is low and there is a strong flexible LB and suitable for different environments. A proposed algorithm based on dynamic load measurement using a machine learning process. Therefore apply the machine learning process to the results of the load balancing to properly load the load on the cloud infrastructure. The proposed algorithm will provide well-organized analytical data in which we can perform a current load planning method [10].

In this research[11], a new LB algorithm is being implemented in the cloud computing environment using Cloud Sim Toolkit, made in the Java language. With a comparison of the given and throttled algorithm with the current Equally Spread Creating existing algorithms based on various types of parameters such as response time, data processing time and data transfer costs. They also described a performance analysis graph that shows results in a variety of contexts. For the better efficiency and effect improvement using machine line layout techniques are used.

The purpose of this paper is to introduce a method used in computer installation to provide services to VMs and to gain a Higher qualityQoS rate by increasing utilization of resources and reducing the count of DC resources saves costs and energy consumption. Besides, using VM splits as a percentage is not a complicated process and it generally reduce costs and time, and no loss of time occurs in the process of seeking to get VM available. Finally, using a support node will reduce the workload in Load Balancer and share multiple tasks on the same virtual one the device reduces the rejected jobs.[12]

The task of the proposed research is to improve the balance of responsibilities, commitment, and the configuration of a non-dedicated collection, based on a genetic engineered reinforcement algorithm. The research is reflected in the time division of a supply scheme using a split system. Dividing systems exist in machine learning algorithms, based on advanced flexible genes high efficiency. The author created a software package that was designed by testing the proposed system in the Cow-slave Cow (Cluster ofWorkstation) and the New (Network of Workstation) environment. Test resultsfrom two distinct operating systems, shows an improved ability of the load balancing system to adapt the configuration of clusters.[13]

This paper outlines a new way to spread deep network training to very dedicated platforms. The proposed route isbased on the correction acceleration of platform computer resources. Provides model (applicable to such resources)with a batch size equal to their computer capacity, thusthosewaiting times in contact areas (performed during each training period combine gradients between processes) are removed. As a result, in total the execution time required for deep modeling training is reduced (while the accuracy is correct less affected), as shown on both platforms and databases. One of the most important contributions to this project is the exploitation of well-known HPC best practice strategies to balance the widespread training of in-depth learning models.[14]

In this advanced approach, the implication of DL methods has been analyzed in a computerized environment. An environmental workflow framework has been brought up into implementation,namely Deep reinforcement based Deadline-constrained,Provisioning and Load Balancing (DLD-PLB). A good VM schedule is generated using an in-depth reading based process. Genome workflow activities are considered as inclusion of a proposed framework. Makespan and cost outcomes are outlined in the proposed framework and compared to their previous proposed load balancing framework - the Deadline-constrained Hybrid, Provisioning and Load Balancing (HDD-PLB) ”workflow framework.[1]

With the widespread use of cloud computing, high energy and low efficiency’s conflict is becoming more pronounced. To resolve with the usage of energy issue, author introduced a collateral planning system and a measurement algorithm with an in-depth reinforcement strategy. On Comparing to similar evolution algorithms, the proposed algorithm can increase diversity, prevent growth in some way, and accelerate rapid integration. Test results show that thisapproach has the potential to significantly reduce the power consumption of data centers.[2]

Sites can be considered as HetNet. They showed that through this program the goal of BS bias and integration can be achieved based on predictions made by avoiding signal-based features. Data labeled BS-device organization bias

with a controlled separator has been used here. This solves the CRE problem in a simpler way to use it than another solutions based on strengthening learning. Therefore, the proposed method of organizing the organization is possibleis more suitable for situations where the computational resources of key network objects, such asNetwork Server, is very restricted [3].

Year & author	Algorithm	Algorithm type	Metric/Parameter	Used approach	Future work
2014 Bagwaiya, Raghuvanshi	Hybrid approach using Throttled and ESCE LBA	Heterogeneous	Response time, cost reduction	Machine learning	Deal with deadlock problem & optimization of server overflow
2013 Bankole, Ajila	Cloud resources provisioning	heterogeneous	Response time, throughput	Machine learning	Implementation on public cloud platforms
2018 Parida, Panchal	Efficient dynamic LBA	heterogeneous	Response time	Machine learning	Make reliable load scheduling mechanism
Chauhan, Patel	LB in cloud computing using supervised learning	homogeneous	Response time, , data transfer cost, data concocting	Machine learning	Reduced makespan
2018 Elrotub,Gherbi	VM class based approach	homogeneous	Energy consumption, resource utilization, reduced job rejections	Machine learning	Patronizing prediction techniques, Controlling task prioritization
Pantas , pinto	Load balancing based on genetic machine learning	Distributed, parallel	High adaptability	Machine learning	Less execution time expected
2020 Alvarez et. al	Training deep neural networks	heterogeneous	Reduced execution time	Deep learning	Improved scalability
2018 Hou &Zhao	Resource Scheduling and LB based Fusion Algorithm with DL	heterogeneous	Discriminable population as input and faster conflux rate	Deep learning based on CSA	Improved makespan
2018 Gomez et al.	load balancing in Heterogeneous Networks	heterogeneous	Increased capability of urban network	Deep learning	Should be used for supervised learning algorithm as well

Table 1. Comparison of few Algorithms

Here, the author mentioned the process of joining the computer to jointhe problem of server supply on the computer edge. Due to the unpredictability of mobile devices and the unpredictability of the peripheral computer system, it requires managing the change in a load of mobile applications which has been most challenging issues to consider. Use the LA process to identify the correct decision to load the optimal load based on smart devices and their applicationsusing edge or cloud servers, Besides they have used the LSTM prediction model and the Q learning process to performan effective measurement decision to add or remove a server edge to modify workload variability[15]

Task scheduling supports the multi-functional based technique of order usually preferred by similitude to ideal solution with (TOPSIS-PSO) algorithm using distinct cloud criteria. To see the load balances between VMs, the VM manager makes decisions about VM migration. A hybrid of a supervised process (for neural network installation), leased machine learning (and integration), and soft computing (interval type 2 fuzzy logic system) –based on the measurement of the algorithm load, i.e., a separate process based on the measurement process load (CMODLB), introduced to load cloud load in the current function.[16]
A natural occurring solution called “healthy bee food” (LB-HBF), an automated load balancing system is proposed here. The method of assignment of work and machine is co-existing allocated in the relation of LB-HBF.LB is considered one of the most important means of productive distribution of cloud services. The deal of future cargo balancing involves completely independent distribution networks. This particular method is based on the concept of Digital Hash Table (DHT) metal compression system. Avenue of the state of clouds and environmental workers is done with the help of weight. the proposed algorithms show better performance by simulation in a variety of heterogeneous and homogeneous variants [17].

Author here legitimized the migration of VM migrations and build a compatible algorithm using Markov and Q-Learning's decision-making process. Imitation results show the effectiveness of our proposal. The proposed algorithm is mapped to the Round-Robin algorithm to reflect its power and availability [18].

In this paper, a scheme based on machine learning of VM switching is introduced with the aim of balancing the load on grip (HM) objects. In this research the learning agent, opts for an action between permissible actions and then applies them to the environment which receives a reward based on the desire for a solution that comes from performing that action in the environment. This action is performed on each iterative learning method.On rreceiving an award from the environment and updating the action value table enables the student agent to learn from the following passages, selecting and performing the best action in the environment implying to further development [19].

Proposed algorithms measure the load by redistributing the load to the appropriate VMsby based on the robust number of each VMs. The proposed algorithm also improves power generation, performance, power consumption during load measurement and reduces work waiting time more effectively compared to separate algorithms such as MPSO and Q-learning[20].

V. CHALLENGES IN LOAD BALANCING ALGORITHM

A. *Geospatially centralized nodes in the cloud:* The Hubs present in distributed computing are appropriated geologically. The provocation for such situation is that the heap adjusting calculations ought to be planned so they consider parameters, for example, the system transfer speed, correspondence speeds, the separations among hubs, and the separation between the customer and assets.[21]

B. *Virtual Machine Migration:* The administration on-request nature of distributed computing infers that at the point when there is an administration demand, the assets ought to be given. Now again assets (regularly VMs) ought to be relocated from one single physically present server to another, conceivably ina inaccessible area. Pioneers of load-adjusting calculations need to think about two issues in that situation: Time of relocation which influences the exhibition and the likelihood of assaults [22]

C. *Failure of Unipoint:* A portion of the load adjusting calculations is unified. In these cases, if the hub which is executing the calculation (controller) falls flat, the entire framework will crash be- cause of that solitary purpose of disappointment. The test here is to configuration circulated or decentralized calculations.

D. *Exposure of Minor Data Centres in Cloud Computing:* Minor centers are less expensive and devour less energy concerning huge data center. In this way, registering assets are conveyed all around the globe. The test here is to plan load-adjusting calculations for sufficient reaction time.

E. *High Complexity of Algorithm:* The load-balancing calculations have to be straightforward regarding execution and task. Complex calculations have negative consequences for the entire execution.

F. *Energy Management:* These days, the power utilized by IT hardware is an extraordinary matter to be appertained. In 2005, the absolute energy devoured by IT hardware was 1.9% of absolute power utilization on the planet. Google server farms have devoured 262 millions of Watts of energy which has been equivalent to 0.012% of world's total energy. According to the research, on a normal, 32% of cloud servers abuse 10–18% of the asset limit. Constrained asset use expands the expense of cloud focus tasks and power use. Because of the propensity of associations and clients to utilize cloud administrations, later on, the establishments of cloud suppliers will extend and in this way the energy utilization in this industry will increment quickly. This expansion in energy utilization builds the expense of energy as well as likewise expands carbon-discharge. On the off chance that the quantity of servers in server farms achieves a limit, their capacity utilization can be similar to the consumption in a city. [21]

VI. REAL WORLD PLATFORMS AND SIMULATION FOR LOAD BALANCING

A. *Eucalyptus:* It is real world open source platform. It is heterogeneous in nature and it manages the complexity. It consists of few models such as cloud controller, cluster controller, storage controller and node controllers [23]

B. *Openstack:* It is a cloud computing software, open source in nature and its services are easily accesses by the private and public cloud servers. Few components of this real world platform are keystone, horizon, cinder, glance etc.

C. *Amazon EC2:* It is designed on larger web scale. It works as real world platform. It uses the most powerful GPU processor. It gives the complete access to its user. These days many MNCs have been using its service as amazon web service

D. *Cloud Sim:* It is proposed as a self-administering event setup together a coherent test frame work gadget regarding Cloud and Grid systems. This gadget offers inconvenient propagation results from standard task executions on rented handling organizations to count the cost, and weight of cloud condition. Cloud Sim gives two or three clear estimations and examination considerations ensuing to runtime to enable the client to simply make more practical examination.

E. Green Cloud: Green Cloud is made as a pushed pack level test framework gadget for power- conscious limit of dispersed calculation server, ranches with attention on cloud correspondences. This gadget transmits an all-out first-class showing of the power used by the server farm IT instruments like correspondence joins, framework switches, likewise, computational servers. It basically associated with create novel courses of action in controlling, the appointment of benefits, strategy of extraordinary weight, and increase in communication and structural frameworks.

F. Cloud Analyst: The tool is productive to engineers as it is used to survey huge scale cloud application need for customer remaining jobs needing to be done and figuring the server's land flow. The later was made to help builds in having learning on ways to deal with getting for application scattering inside cloud establishments tallying worth included organizations while the past was shown for the objective of considering application rehearses inside a couple of sending courses of action.

VII. CONCLUSION

LB has always been a greatest challenge in Cloud Computing environment. The main purpose of this paper is to integrate existing machine learning based methods of balancing the load. This paper points on the problem of virtualization and the hybrid new algorithms which are now used widely in the realm of cloud computing. This paper also discussed various challenges related to the construction of an effective load balancing algorithm and it gave a comparative view of the existing algorithms to measure load capacity. The main purpose of measuring the load in the cloud computing is to increase the profitability of providers of cloud's service and reduce the cost of cloud users.

REFERENCES

[1] Kaur, A., Kaur, B., Singh, P., Devgan, M. S., & Toor, H. K. (2020). Load Balancing Optimization Based On Deep Learning Approach in Cloud Environment. International Journal of Information Technology and Computer Science, 12(3), 8-18.

[2] Hou, X., & Zhao, G. (2018), "Resource Scheduling and Load Balancing Fusion Algorithm with Deep Learning Based on Cloud Computing", International Journal of Information Technology and Web Engineering (IJITWE), 13(3), 54-72.

[3] Gomez, C., Shami, A., & Wang, X. (2018)., " Machine Learning Aided Scheme for Load Balancing in Dense IoT Networks", Sensors, 18(11), 3779.

[4] Yang, R., Ouyang, X., Chen, Y., Townend, P., & Xu, J., (2018), "Intelligent Resource Scheduling at Scale: A Machine Learning Perspective", in Proceedings of 2018 IEEE Symposium on Service-Oriented System Engineering (SOSE), pp. 132-141.

[5] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J. & Kudlur, M. (2016), "Tensorflow: A System for Large-scale Machine Learning", In OSDI, Vol. 16, pp. 265-283.

[6] Mishra, S. K., Sahoo, B., & Parida, P. P. (2020). Load balancing in cloud computing: A big picture. *Journal of King Saud University-Computer and Information Sciences*, *32*(2), 149-158.

[7] M. Mesbahi and A. M. Rahmani, "Load balancing in cloud computing: a state of the art survey," Int. J. Mod. Educ. Comput. Sci, vol. 8, no. 3, p. 64, 2016.

[8] Bagwaiya, Vishwas, and Sandeep K. Raghuwanshi. "Hybrid approach using throttled and ESCE load balancing algorithms in cloud computing." 2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE). IEEE, 2014.

[9] Bankole, A. A., & Ajila, S. A. (2013, May). Predicting cloud resource provisioning using machine learning techniques. In 2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (pp. 1-4). IEEE.

[10] Parida, S., & Panchal, B. (2018). An efficient dynamic load balancing algorithm using machine learning technique in cloud environment. International journal of scientific research in science, engineering and technology, 4(2018), 1184-1186.

[11] Samir Chauhan, Jinal Patel (2019) 'Load Balancing in Cloud Computing Using Machine Learning Techniques ', JASC: Journal of Applied Science and Computations, VI(IV), pp. 3533.

[12] Elrotub, M., & Gherbi, A. (2018). Virtual machine classification-based approach to enhanced workload balancing for cloud computing applications. Procedia computer science, 130, 683-688.

[13] Dantas, M. A., & Pinto, A. R. (2005, May). A load balancing approach based on a genetic machine learning algorithm. In 19th International Symposium on High Performance Computing Systems and Applications (HPCS'05) (pp. 124-130). IEEE.

[14] Moreno-Alvarez, S., Haut, J. M., Paoletti, M. E., Rico-Gallego, J. A., Diaz-Martin, J. C., & Plaza, J. (2020). Training deep neural networks: a static load balancing approach. The Journal of Supercomputing, 76(12), 9739-9754.

[15] Shahidinejad, A., & Ghobaei-Arani, M. (2020). Joint computation offloading and resource provisioning for edge-cloud computing environment: A machine learning-based approach. Software: Practice and Experience, 50(12), 2212-2230.

[16] Negi, S., Rauthan, M. M. S., Vaisla, K. S., & Panwar, N. (2021). CMODLB: an efficient load balancing approach in cloud computing environment. The Journal of Supercomputing, 1-53.

[17] A Nature-Inspired Solution to Managing Activities in the Cloud with Equal Time by Using Machine Learning Approach. In Proceedings of International Conference on Computational Intelligence and Data Engineering (pp. 221-235). Springer, Singapore.

[18] Bui, K. T., Nguyen, L. V., Tran, T. V., Pham, T. V., & Tran, H. C. (2021). A Load Balancing VMs Migration Approach for Multi-tier Application in Cloud Computing Based on Fuzzy Set and Q-Learning Algorithm. In *Research in Intelligent and Computing in Engineering* (pp. 617-628). Springer, Singapore.

[19] Ghasemi, A., & Toroghi Haghighat, A. (2020). A multi-objective load balancing algorithm for virtual machine placement in cloud data centers based on machine learning. Computing, 102, 2049-2072.

[20] Jena, U. K., Das, P. K., & Kabat, M. R. (2020). Hybridization of meta-heuristic algorithm for load balancing in cloud computing environment. Journal of King Saud University-Computer and Information Sciences

[21] S. Bareen, K. Shinde, and S. Borde, "Challenges of big data processing and scheduling of processes using various hadoopschedulers: a survey," International Journal of Multifaceted and Multilingual Studies, vol. 3, no. 12, 2017.

[22] E.J.Ghomi,A.M.Rahmani,andN.N.Qader,"Load-balancingalgorithmsin cloud computing: a survey," *Journal of Network and Computer Applications*, vol. 88, pp. 50–71,2017.

[23] Shah, J. M., Kotecha, K., Pandya, S., Choksi, D. B., & Joshi, N. (2017, May). Load balancing in cloud computing: Methodological survey on different types of algorithm. In *2017 International Conference on Trends in Electronics and Informatics (ICEI)* (pp. 100-107). IEEE.

[24} Agarwal, Ronak, Neeraj Baghel, and Mohd Aamir Khan. "Load Balancing in Cloud Computing using Mutation Based Particle Swarm Optimization." 2020 International Conference on Contemporary Computing and Applications (IC3A). IEEE, 2020.