ISSN 2278-3091
Volume 10, No.2, March - April 2021
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse521022021.pdf
https://doi.org/10.30534/ijatcse/2021/521022021

Load Balancing in Cloud Computing: A Comprehensive Survey on Recent Techniques

Shilpa Mehta[1], Hetal Joshiyara[2]
[1, 2] Department of Computer Engineering, L.D College of Engineering, Ahmedabad, Gujarat, India
[1]mehtashilpa44@gmail.com, [2]hetaljoshiyara@gmail.com

ABSTRACT

The Internet has become a fundamental necessity of daily activity. It has a more significant impact on modernizing the digital world. Therefore, cloud computing has been one of the most promising technological advances in recent days. It aims to provide millions of consumers around the globe with smooth computing services. In workload distributions and system behavior, the cloud environment is extremely dynamic, leading to load imbalances across data center resources. Balancing load across the systems is a crucial component of cloud computing, avoiding some nodes' overloading while others have little or no work to do. The problem of load balancing in the cloud has been linked to a large amount of research, recognizing its role and impact. This paper bestows a thorough investigation of the new load balancing algorithms for the cloud environment. This study's primary focus is the concept of load balancing, a literature review on load balancing and task scheduling techniques, and different measurement parameters.

Key words: Cloud Computing, Cloudsim, Load Balancing, Resource Scheduling, Task Scheduling, Virtual Machine.

1. INTRODUCTION

In recent times, cloud computing got a swing. It came up with a pliable and effortless way of keeping and retrieving data and files as part of its services. In particular, to make massive amounts of data and files available to an increasing number of users worldwide, cloud computing offers a flexible way of retaining information and files, including web services, distributed computing, and virtualization [1]. It also has many components, including clients and distributed servers. Cloud computing aims to provide the utmost services at any time with minimum costs. Today, over a hundred million systems have access to the Internet. The request and response mechanism of such systems should be without any delay then. The cloud environment's primary goals are to bring down the costs, boost up the response time, and improve performance. Thus, the cloud is also named as the pool of services [1], [2]. The three categories of services provided by the cloud are Software as a Service (SaaS), Infrastructure as a Service (IaaS), and Platform as a Service (PaaS). Resources that are virtualized, such as on-demand storage, are suggested by IaaS. PaaS gives a higher abstraction level for the user to make it seamlessly programmable. It provides a platform for its users to create different applications, and they do not need to know the processor requirements for the application. In SaaS, through the Internet, a user has the authorization of any software. This type supplies a user with the applications and frees a user from the software maintenance burden [6]. The ceremonial definition of cloud computing by the National Institute of Standards and Technology (NIST) identifies four types of cloud deployment models: public, hybrid, community, and private clouds [3].

A cloud model is useful if its resources are used in the most feasible way and accomplished through the use and maintenance of sound cloud resource management. Management of resources is achieved through robust resource planning, assignment, and vital resource scalability techniques. Virtualization is the prime technology and concept behind cloud computing that enables various tasks to be carried out simultaneously via a shared hardware platform [4]. It offers the user a notion of working in an enclosed environment because the jobs do not engage with each other and only have access to their own data. The Virtual Machines (VM) provides the services to customers using a virtualization mechanism that uses an object known as a hypervisor [5]. Virtualization offers the opportunities for immediate and as & when required setup of physical machines to perform various tasks, thus avoiding waste resources [4], [7]. Although virtualization attempts to balance the entire system's load dynamically [8], there is always a likelihood of either excessive or low use of resources. Overloaded systems result in performance deterioration, while underloaded systems show low resource usage. Due to unsuitable allocation of the load, additional heat will be produced by the heavily loaded servers, which, in turn, raise the price of the cooling system and increase the considerable CO_2 emissions, which contribute to the greenhouse effect [7], [9]. Therefore, it is essential to impart the correct number of resources dynamically to applications running on virtual servers to fulfil the Quality of service (QoS) requirements and balance the overall system load. Among all the primary challenges in cloud environment, load balancing is above all. It is a technique that places dynamic tasks equally among all cloud nodes to keep away from a state where specific machines are overloaded. In contrast, others are inactive or doing significantly less work [10]. It helps attain shorter response time, high performance, enhanced fault tolerance, scalability, better service quality, less heating, optimal power consumption, lowered CO2 emissions, and reduced costs through efficient use of resources.

In clouds, load balancing (LB) may be between physical hosts or VMs. The dynamic workload is distributed evenly between all nodes by this balancing mechanism. Cloud load balancing is also called to as load balancing as a service. Two variants of algorithms for load balancing are: static and dynamic. Static LBA are often suitable for homogeneous systems in stable environments. Nevertheless, they are generally not pliable and can't match run time changes to attributes at the time of the execution time [11], [12]. In both heterogeneous and homogeneous circumstances, dynamic-based balancing algorithms are more resilient and practical. However, the distribution attributes are becoming more dynamic and complex. Consequently, more than a few of these algorithms could not be much efficient and cause more overhead than expected, begetting the total deterioration of service performance [11].

A survey of the algorithms and already existing techniques for cloud computing load balancing are presented in this paper. We review their characteristics and parameters that authors considered, and give a comparative view of the current mechanisms for load balancing. The remainder of this paper is structured as follows: Section 2 Discusses earlier surveys on cloud load balancing. Section 3 It offers a relatively detailed literature review and its summary of the latest existing techniques for load balancing. Open problems are outlined in the Section 4. At last, in Section 5, our survey is concluded.

2. RELATED WORK OF EXISTING SURVEYS

Worldwide research organizations are interested in designing and developing optimal load balancing, task scheduling (TS) and resource utilization methodologies, as recent research and reviews draw attention. One of the main challenges and trouble in cloud is load balancing; this involves allocating and re-allocating the load between resources to maximize output while reducing costs and response time, improving performance and using resources, and saving energy [13], [14]. Outstanding load balancing approaches could include Service Level Agreement (SLA) and customer satisfaction. Therefore, a key to cloud computing environments' success is to include successful load-balancing algorithms and mechanisms. Numerous studies were conducted in cloud computing area, load balancing, SLA, power consumption, managing resources, etc. This article provides a concise overview of the different kinds of cloud load balancing scheduling mechanisms. In this section, we will have a look on few of the articles that have remarkably studied load balancing techniques.

Milani and Jafari (2016) [15] have reviewed and categorized numerous existing load balancing strategies into dynamic and hybrid sub-domains. They portrayed the aspects of various load balancing strategies, including benefits and drawbacks. Elaborative classification is based on various cloud measurements. They have pointed to the challenges of developing more efficient algorithms to minimize resource usage and energy consumption and to increase the efficiency of load balancing approaches.

A comprehensive review of cloud load balancing was carried out by Ghoomi et al. (2017) [16], a new classification of algorithms has been presented, such as the Hadoop MapReduce load balancing category, Natural Phenomena-based load balancing category, General load balancing category, etc. The critical components discussed in this paper were the existing load balancing techniques, new categorization of algorithms, pros and cons of load balancing algorithm in each category, along with their assessment methods. Also, future research challenges have been presented. Singh et al. (2017) [17] reviews the use of meta-heuristics techniques in cloud computing for job scheduling. Analysis of grid and cloud computing is shown based on bio-inspired and swarm intelligence techniques. This paper also discusses the cloud workflow scheduling model, meta-heuristic algorithms (Bio-inspired, Swarm intelligence, various research issues, and scheduling policies. It also gives directions for future studies.

Ahmad and Khan (2018) [18] discussed a structured review with a comparative study on preexisting techniques and tools for load balancing. Virtualization idea of cloud and its genres for optimal resource utilization have also been discussed. Performance metrics for load balancing were discussed for evaluating the performance of existing techniques. Mishra et al. (2018) [11] A classification for the Load Balancing Algorithms (LBA) has been presented is terms of static and dynamic. The paper gives a quick justification of the performance parameters that have been deemed in the research and their impacts. The performance of few heuristic algorithms was simulated and represented by graphs with regards to makespan and energy consumption. Arunarani et al. (2018) [19] presented an extensive survey on the planning of the tasks and related measurements was discussed. The various problems in the scheduling and the limits to be overcome were addressed. To discover the utility of scheduling characteristics, distinctive scheduling processes are studied. Organized literature survey based on three measures: parameter-based, methods, and application.

Kumar and Kumar (2019) [20] presented a cutting-edge review of issues and challenges related to the so-far load-balancing techniques for researchers to build more efficient algorithms. Discusses classification, policies and metrics of load balancing. A comprehensive literature study of various load balancing algorithms (static and dynamic, general techniques, natural, hybrid, agent-based, task-based, cluster-based) along with variety of research questions and future research trends is depicted. Afzal et al. (2019) [21] a detailed, exhaustive review of the techniques of load balancing is presented in this article. With critical challenges being inscribed to develop practical load balancing algorithms in the future, the benefits and limitations of current techniques are highlighted. In cloud computing, the paper also suggests new insights into load balancing.

Motlagh et al. (2020) [22] This paper provides systematic literature review-based analysis of task scheduling methods that identify in: 1. Single cloud environments evaluating cost-aware, energy-aware, multi-objective, and QoS-aware approaches in scheduling tasks. 2. Multi-cloud environment considering task scheduling that is cost-aware, multi-objective, and QoS-aware. 3. Mobile cloud environment. To demonstrate the benefits and limitations of the current solutions, analytical discussions are presented. Mishra et al. (2020) [23] provides a comprehensive historical research study and comparative analysis of the different existing literature on load balancing (LB).

3. LITERATURE STUDY OF EXISTING TECHNIQUES

We investigated literature on previous methods of load balancing and analyzed the selected papers profoundly. The load balancing approaches presented in this survey are mainly divided into general load balancing algorithms, consisting of a blend of algorithms including a variety of swarm intelligence algorithms. The other category focuses on energy-aware load balancing algorithms.

3.1 General Load Balancing Techniques

A detailed discussion on General Load Balancing techniques is given below and a summary of these techniques is shown in Table 1.

Kumar et al. [24] came up with a load balancing approach named FDLA, by making use of a novel fractional dragonfly algorithm. Along with it, two selection probabilities Task Selection Probability and VM Selection Probability, are introduced. Based on the parameters such as capacity and machine load, the tasks to be pulled out and the underloaded VM can be selected by TSP and VSP. For selecting the best VM, the suggested fractional dragonfly algorithm is modelled by merging fractional Calculus (FC) and dragonfly algorithm (DA) with a new fitness function. The performance of the FDLA is compared with the existing techniques, demonstrating that with significant performance, the proposed technique is effective. Geng et al. [25] discussed the TS problems in a cloud environment; incorporating the concept of task clustering and task duplication, a new static Directed Acyclic Graph (DAG) scheduling approach is suggested. This new approach enhances task parallelism by transforming DAG to an in-tree graph. The makespan of the overall plan is minimized by task grouping strategy, decreases the utilization of processors noticeably by combining the task groups, refines the processor's utilization by using the processor's downtime properly reduces the execution cost. Srichandan et al. [29] proposed a hybrid generic task scheduling algorithm for the heterogeneous cloud environment derived from the foraging bacteria and the concept of genetic algorithms. The major contributions to this study are bifold. First, the scheduling algorithm reduces the makespan, and second, it lowers the energy consumption, both from an ecological and economic point of view.

Maytami et al. [30], the primary goal of this study is to increase the efficiency of task scheduling while at the same time lowering computational costs. The fundamental goal is to forecast the desirable algorithm for incoming/available requests as and when necessary. A thorough study is conducted of heuristic approaches for the use of resource utilization using principal component analysis (PCA) in the cloud. In addition, the requirements and impacts of use of service quality are analyzed with the Prediction of Tasks Computation time (PTCT).

Pang et al. [31] proposed a multi-objective task scheduling model that states the challenge of VM tasks in a detailed manner. A combination of Estimation of distribution algorithm (EDA) and genetic algorithm (GA) is suggested to come up with an effectual scheme for multi-objective task scheduling, and the effectiveness of this algorithm is verified through comparative experiments. This paper aims to lower the task completion rate and achieved a better load balancing ability. This paper doesn't really take into account the dynamics and uncertainties of the cloud environment. On the one side, the computing rate of VMs is changing in real-time. Virtual machines, on the other side, can join or leave the cloud service at any moment.

Gamal et al. [33] presented a hybrid LB algorithm combining the positive traits of Artificial Bee and Ant Colony optimization (ACO). It depends on joining ACO 's critical behavior, like rapidly discovering reasonable solutions, and the Artificial Bee Colony (ABC) algorithm, like communal bee interaction and sharing waggle dancing details.

Priya et al. [35] established an integrated algorithm for resource scheduling and balancing load for effective provisioning of cloud services. The work proposes a Multidimensional Resource Scheduling (MRS) Fuzzy-based scheme to attain resource scheduling efficiency in a cloud system. Through objective and efficient load balancing, improving virtual machines' use is then attained by dynamically choosing a class request with the help of Multidimensional Queuing Load Optimization (MQLO) algorithm. To prevent underutilization, overuse of resources, and improve latency time for every single class of ownership, a LB algorithm is then implemented.

Rizk et al. [38] proposed a LBA for the Osmotic Hybrid artificial Bee and Ant Colony optimization (OH-BAC) with a new trend to apply the osmosis technique for balancing the load. OH-BAC uses osmosis technology to provide an environment for energy-efficient cloud computing. ACO and ABC coordinate with the OH-BAC algorithm to pick the optimal virtual machine to relocate to the most appropriate PM. Also, OH-BAC activates the most appropriate osmotic host to reduce power consumption among all PMs in the system. In the experiments with constant and variable loads, the proposed algorithm was simulated to measure various metrics' performance.

In order to minimize makespan time, this Gupta et al. [40] paper presented an approach that intends to achieve broad-spectrum adjusted load crosswise over VMs. By using the honey bee load balancing and enhancement detection operator, the suggested methodology provides balanced scheduling solutions to sum up, which low-level heuristic should be used to hunt for better candidate solutions. The implications of the presented TS approach are consistent with current heuristic-based scheduling procedures. The degree of imbalance is calculated to accurately test the efficiency of the proposed load balancing algorithm, which indicates that the suggested LBA achieves better results for a greater number of allocated cloudlets.

Table 1: Summary of Latest Existing Load Balancing and Task Scheduling Techniques (General)

Reference/ Year	Overview	Algorithm/ Technique	Type	Objective	Tools/ Hardware	Compared with
[24] 2018	Presents a load balancing approach, named FDLA, along with two selection probabilities.	Fractional dragonfly based load balancing (FDLA)	Motivated by the dynamic and static swarming behaviour	Multi- objective	Cloudsim	PSO, Honey bee behaviour (HBB), DA
[25] 2018	TS issues in cloud environment is presented in this study, synthesizing the plan of task clustering and task duplication (TD), a new static Directed Acyclic Graph (DAG) scheduling algorithm is presented.	Algorithm based on TD and task grouping	Static	Multi-objective	Cloudsim	TD based scheduling algorithm, TD-based Clustering Scheduling
[26] 2018	This paper's main objective is to put forward an optimization technique that is stimulated by the decision-making action for load balancing using an ABC algorithm.	Bee Colony Optimization	Dynamic	Single-objective	Cloudsim	First Come First Serve (FCFS), Dynamic LB (DLB)
[27] 2018	A latest offline LB approach is proposed to operate resources in mobile CC.	Bin Packing Algorithm	Dynamic	Multi-objective	-	-
[28] 2018	A hybrid algorithm for TS is proposed, consisting of a new multi-objective function by combining energy consumption, credit and penalty, memory usage and cost.	Cuckoo-Harmony Search Algorithm (CHSA)	Dynamic	Multi-objective	Cloudsim	Hybrid cuckoo gravitational search algorithm, traditional CS and HS algorithm
[29] 2018	This article examines the TS algorithm using an amalgam model, which integrates favourable traits of most regularly used biologically-influenced heuristic algorithms.	GA and the bacterial foraging (BF) algorithm	Dynamic	Multi-objective	MATLABR2013	PSO, GA, BFA
[30] 2019	A novel scheduling algorithm based on the PTCT algorithm to evaluate the preeminent scheduling algorithm for prominent cloud data is suggested.	DAG, PTCT, Principal Component Analysis (PCA)	Static	Single-objective	MATLABR2013	Min-min, max-min, QoS-guide, MiM-Mam scheduling algorithms
[31] 2019	It aims to shrink the task finishing time and upgrade the system load balancing ability.	EDA and GA.	Dynamic	Multi-objective	Cloudsim	Traditional EDA and GA
[32] 2019	A swarm intelligence-based algorithm is proposed to tackle the problem of TS in cloud computing.	Grey wolf optimizer nature-inspired algorithm.	Dynamic	Multi-objective	Cloudsim	ACO and performance budget ACO (PBACO)
[33] 2019	A bio-influenced based algorithm is proposed to realize balancing load for TS in cloud.	Hybrid ABC and ACO (H_BAC) LB algorithm	Dynamic	Single-objective	Cloudsim	ABC, ACO, Hybrid Algorithms
[34] 2019	Designed a novel resource clustering scheme to cluster the servers based on their present centers and cluster loads.	Bat Algorithm	Dynamic	Multi-objective	-	Random Deployment approach weighted round robin (RR) approach, DLB and LB-BC approach
[35] 2019	With the aim of decreasing processing time and increasing resource utilization, scheduling resources according to user appeal in a collateral fashion and LB of the scheduled resources is the issue to be resolved in this study.	FMRS, MQLO	Dynamic	Multi-objective	Cloudsim	Virtual LAN towards Scalable Traffic Management, Scalable Workload Driven Partitioning (SWDP)
[36] 2019	A hybrid firefly and Improved Particle Swarm Optimization (IPSO) algorithm is illustrated for LB optimization in cloud environments.	Hybrid nature inspired algorithm	Dynamic	Multi-objective	MATLAB	RR, FCFS, Short jobs First (SJF) and GA

[37] 2019	TS is performed on basis of social behavioural characteristics of crow i.e., food collecting habits of crow.	Crow Search Algorithm	Dynamic	Single-objective	Cloudsim	Min–Min and Ant algorithms
[38] 2019	A hybrid metaheuristic approach is integrating osmotic actions with bio-inspired algorithms for LB.	Osmotic hybrid artificial bee and ant colony (OH_BAC)	Dynamic	Multi-objective	Cloudsim	Hosts overloading detection algorithms and bio-inspired algorithms.
[39] 2019	An effective binary version of the reduced computation time and cost effective PSO algorithm for scheduling and balancing computing cloud tasks	IBPSO-LBS	Dynamic	Multi-objective	Cloudsim	Existing heuristics and meta-heuristics algorithms
[40] 2020	This study proposed an algorithm that expects the entirely adjusted load to be achieved on virtual machines crosswise to minimize the makespan time.	Honey Bee optimized Hyper-Heuristic algorithm.	Dynamic	Single-objective	Cloudsim	ACO, PSO and GA.
[41] 2020	A new distributed LB algorithm, based on adaptive starvation threshold, which aims to the makespan time.	Starvation Threshold based LB (STLB) algorithm	Dynamic	Multi-objective	Cloudsim	A honey bee behavior inspired LB algorithm
[42] 2020	An efficient multi-objective-based LB approach. The primary motive of this research is to reduce the cost and task's execution time, and to attain a well-balanced load in a cloud across all VMs.	Adaptive Dragonfly algorithm (ADA), combination of DA and FA	Motivated by the dynamic and static swarming behaviour	Multi-objective	Cloudsim	Traditional FA and DA
[43] 2020	This paper proposes a scheme on utility-based LB that uses FA to advance the resource utilization and optimize the gain for the CSP using bargaining protocol.	FA; Utility based scheme	Dynamic	Multi-objective	Cloudsim	Imperialist competitive algorithm and FA (ICAFA)
[44] 2020	This paper presented three versions of swarm intelligent algorithms to be used as task schedulers to lower the makespan.	Ant-Lion optimizer (ALO) and Grey wolf optimizer (GWO)	Dynamic	Single-objective	Cloud reports (extension of CloudSim)	PSO and FA
[45] 2020	Enhanced FA is developed by embracing the crucial concepts of Firefly approach.	Firefly Algorithm	Dynamic	Single-objective	Cloudsim	Traditional FA, FCFS, ACO
[46] 2020	A combinational Firefly-Genetic heuristic approach is suggested in this paper to optimally assign the resources and schedule the cloud computing task.	Hybrid of FA and GA	Dynamic	Single-objective	Cloudsim	Traditional First in First Out (FIFO) and GA
[47] 2021	Considering both the computing and communication loads in a cloud environment, a new PSO-based LB algorithm is proposed.	Adaptive -pbest discrete PSO scheduling algorithm (APDPSO)	Static	Multi-objective	MATLAB, Cloudsim	Existing PSO-based algorithms
[48] 2021	A weight-based technique for VM migration is proposed.	Weight-based improved GA	Static	Single-objective	MATLAB	Existing GA
[49] 2021	Proposed a method to evaluate the essential number of reducers for a MapReduce job running on a heterogeneous Hadoop Cluster.	MapReduce	Dynamic	Single-objective	Oracle VM VirtualBox 5.2	Existing rule
[50] 2021	Different algorithms are addressed for calculating the efficiency of the cloud when performing LB.	M-Throttled	Dynamic	Single-objective	Cloudsim	Throttled, RR, SJF

Tapale et al. [43] proposed an approach on utility-based load balancing that uses the firefly algorithm to upgrade the utilization of resources and optimizes the service provider's profit using bargaining protocol. The method avoids starvation of tasks by taking into account the multi-level queues and decreases the load imbalance and makespan. Even at higher loads, when the number of tasks increases on the cloud, the system remains as it is and exhibits finer performance than existing ones.

Farrag et al. [44] studies the application of two swarm algorithms in TS of the Cloud environment and has compared the outcome with Particle Swarm Optimization (PSO) and Firefly Algorithm (FA). Three variations of ALO and GWO were presented to be used as Task Schedulers to reduce the makespan. Inspired by the dynamic and static swarming behaviours of artificial dragonflies in nature Neelima and Reddy [42] John proposed an efficient load balancing system using a commutable dragonfly algorithm (DFA) in cloud. An efficient multi-objective-based LB approach is developed to overcome the problem of unbalanced load, which leads to maximum time and cost. The suggested methodology explores DFA's advantages in optimizing cloud TS and allocation of resources.

PSO and most of its versions were only used to update the best individual positions in the experiment, which resulted in the choice of a bad particle as the head, considering this issue Miao et al. [47] proposed a new PSO-based static LBA namely adaptive Pbest discrete PSO (APDPSO). This research aims to assign an appropriate host to each VM to balance the use of resources across all hosts and reducing the associated cost of communication between tasks. Each VM or host is connected with a distinctive identifier to schedule the VMs in a cloud computing system, and the goal is to find the optimal VM-host pairs. It is, therefore, nearly a distinct problem of multi-objective optimization. A method of discretization based on probability and similarity was also proposed to adopt the PSO algorithm to the discrete optimization problem by changing the process of updating the particles' velocity and location vectors.

Kaur and Sachdeva [48] proposed a virtual machine migration weight-based technique using a GA. The algorithm proposed is to enhance the genetic algorithm to decrease the execution time for cloud task migration work. By replacing the mutation calculation points by which the execution is made faster and more reliable than the existing approach, the number of migrations is reduced. There is high consistency and pace in this recommended method. This approach therefore minimizes the likelihood of failure.

Numerous instructions can be handled by the data center concurrently, but as the instructions are arbitrarily submitted, there is a chance that the data center might get overloaded. To handle such situations, load balancing is necessary, and thus Panigrahi et al. [50] proposed an algorithm named M-Throttled, which has higher performance compared to existing approaches. The performance of the proposed algorithm is examined based on parameters like response time and computation time.

A summary of the existing techniques discussed is presented in Table 1.

3.2 Energy-Aware Load Balancing Techniques

Yadav et al. [51] presents three adaptive models that help minimize service level agreement (SLA) violations and energy consumption. The first model, namely Gradient descent-based regression, calculates the higher threshold based on the historical dataset to detect the overloaded host. Second, maximize correlation percentage chooses and adaptive upper threshold based on maximum correlation percentage. The third model bandwidth-aware selection policy helps minimize VM migration time from one host to another. The experiments have been performed on real workload traces which exhibits that the suggested algorithm reduces the energy consumption while preserving the other parameter.

Energy-aware virtual machine consolidation approach is presented by Wang, H. et al. in [52]. Minimizing energy consumption is the main target of this study. For VM placement, Space Aware Best Fit Strategy is used. In order to balance the cloud environment effectively, high CPU utilization-based migration is also done. Energy consumption is kept to a bare minimum. The cloud environment's load is also balanced. Even then, because VM positioning is an NP-hard problem and VM migration may surge the migration costs, the management of complex cloud data centers still requires an efficient technology.

Jeba et al. [53] proposed two algorithms for reduction of power and VM migrations. The working mechanism of virtualization technology with its system model has been discussed. A dynamic scheduling mechanism is presented which utilizes every single server to its full scope and in a pecking, order transfers the incoming load to the other servers. The proposed scheduling algorithm is based on random, sequential, and maximum fairness search.

Patel et al. [54] focus on minimizing the VM migrations to save energy consumption. The migration of heavily loaded VMs considers CPU utilization as a parameter to evaluate its performance. It uses a Backpropagation algorithm from Artificial Neural Network to predict the future load occurring due to temporary peak load. For experimental results, the existing technique is compared with the proposed one.

To maximize resource management, Yavari et al. [6] concentrated on energy optimization. The HET-VC and FET-VC heuristic algorithms were proposed. For evaluation, six parameters in all are being investigated. All the factual evaluation using cloudsim was conducted on simulated data. The authors considered a wide range of variables, but their approach was limited to energy consumption and was only tested on random data.

Soltanshahi et al. [57] proposed a newly introduced fastest collective intelligence algorithm, the Krill Herd algorithm. It is used to allocate the VMs to the appropriate hosts. Lesser time complexity and minimization of energy consumption due to load congestion in data centers are the primary focus of this paper. The simulation outcome shows that efficient integration and the selection of convenient migration strategies for virtual machines can help improve energy efficiency.

To foster green cloud computing, Geetha and Robin [60] presented a power conserving resource allocation scheme with improved quality-of-service. This paper's concept works on the principle of two layers that the authors employ to maintain the QoS parameters. Out of the two layers, the first one, Cloud Manager Layer, is responsible for choosing the appropriate resource out of all available help. Green Manager Layer picks the best one out of the list provided by CML. For this purpose, the first layer maintains a table of record which contains entities like capacity configurations, location, Id, etc. of the resources. The GML layer uses a meta-heuristic algorithm known as the Lion Optimization algorithm to select the best help available. A fitness function is developed based on the behavior of the optimization algorithm.

There are several meta-heuristics algorithms, each having its advantages and disadvantages out of such Devaraj et al. [61]

came up with a hybrid of the firefly algorithm and improved multi-objective PSO technique (FIMPSO) for improving the energy efficiency in a cloud environment. This hybrid approach combines both the existing algorithms; the firefly algorithm is used to minimize the search area, while the IMPSO is implemented to recognize the enhanced response. The suggested methodology achieved a useful average load for enhancing the important measures like response time and proper resource usage. A summary of the existing energy-aware techniques discussed is presented in Table 2.

A brief idea of the pros, limitations and future work of existing techniques is shown in Table 3. And a comparative analysis of the parameters evaluated in the proposed approaches is shown in Table 4.

Table 2: Summary of Energy-aware Load Balancing Techniques

Reference/ Year	Overview	Algorithm/ Technique	Type	Objective	Tools/ Hardware	Compared with
[51] 2018	An energy-conscious algorithm is proposed to upgrade energy efficiency and minimize cloud breaches of SLA.	Maximum correlation percentage, Gradient descent-based regression, and bandwidth-aware selection policy	Dynamic	Multi-objective	Cloudsim	Linear regression, and inter-quartile range, Median absolute deviation, ACO, minimum migration time
[52] 2018	A new architecture for Dynamic VM Consolidation for Green Cloud Computing is presented, along with a new VM selection policy and VM allocation policy.	Space Aware Best Fit Decreasing algorithm	Dynamic	Multi-objective	Cloudsim	Dynamic Voltage and Frequency Scaling, Non-power aware policy
[53] 2019	The article presents a dynamic resource planning framework for cloud data center VM migration based on an efficient energy optimization framework.	Live virtual machine migration-based algorithm, sequential search, random search and maximum fairness search	Dynamic	Single-objective	Cloudsim, Cloudera	Dynamic Compare and Balanced Algorithm
[54] 2019	This work presents an energy-aware VM migration-based LB method for the cloud using a forecasting approach	Artificial Neural Network (ANN) model is used; double threshold based dynamic LBA	Dynamic	Multi-objective	CloudSim, MATLAB 2015a	Dynamic Double Threshold
[55] 2019	A novel VM consolidation technique is suggested based on temperature and energy to improve QoS with the help of heuristic and meta-heuristic algorithms.	Heuristic Energy and Temperature aware based VM consolidation and FireFly Energy and Temperature aware based VM Consolidation	Dynamic	Single-objective	Cloudsim	Dynamic Threshold Maximum Fit algorithm
[56] 2019	The proposed approach aims at reducing the makespan and power expenditure while meeting the limiting constraint.	EATSD, Dynamic classifier algorithm	Dynamic	Multi-objective	Cloudsim	Earliest Deadline First and FCFS
[57] 2019	In this research, the Krill Herd approach, the newest cumulative intelligence algorithm recently introduced, was proposed to allot VMs to hosts in data centers.	Krill Herd algorithm	Dynamic	Single-objective	Cloudsim	GA and Modified best fit decreasing algorithm.
[58] 2019	An optimal Virtual Machine Placement (VMP) scheme is obtained to achieve energy efficiency while maximizing load balance among multiple resources.	GA and tabu search algorithm (GATA)	Static	Multi-objective	PyCharm 3.3.	Simulated annealing algorithm, improved ACS-based algorithm and Traditional genetic algorithm,
[59] 2020	A hybrid VMP algorithm elicited from another suggested enhanced transmutation-based GA and multidimensional resource-conscious best fit allocation approach.	Permutation-based GA, Hybrid VMP	Static	Multi-objective	-	Other permutation-based algorithms
[60] 2020	This paper concentrates on an efficient resource allocation system for cloud users that doesn't negotiate quality of service by using two layers, like CML and GML.	Uses Lion Optimization Algorithm	Dynamic	Multi-objective	Cloudsim	Dynamic VM placement, VM provisioning, data centre provisioning
[61] 2020	A new hybrid algorithm is proposed by combining the advantages of FA and IMPSO.	Combination of firefly and Improved Multi-Objective PSO technique	Dynamic	Multi-objective	MATLAB	RR, FCFS, SJF and GA
[62] 2020	An approach has been potrayed in this study that aids in energy-efficient VM selection with less violation of service level agreement.	Power-aware VM selection policy	Dynamic	Multi-objective	Cloudsim	Existing classic VM selection algorithms

Table 3: Summary of Positive Aspects, Limitations and Future Scope of Surveyed Techniques

References	Positive Aspects	Limitations	Future Scope
[30]	Decreases calculation or communication costs. Assures to attain almost optimal re-allocation of resources	May result in decreased accuracy values	Real-world application graphs with dynamic scheduling.
[25]	Reduces communication costs between tasks through task duplication. Eliminates communication overhead. Lowers the execution cost	Space is sacrificed in order to improve the efficiency	Meeting the Quality of Service (QoS) of the user and improving user satisfaction.
[31]	Fast convergence speed and strong search ability.	Does not take account of cloud-based dynamics and ambiguity	TS issues that are nearer to those in real cloud will be the focus of future work.
[60]	With the optimized selection of resources, time consumption is minimized. GML choose the finest resource in the view of utility degree; this suggestion enhances the QoS and reduces the response time	Best resources are chosen every time by the LOA, which might not always be required, and the best resource might get wasted to serve an average task.	The service can be carried out with regard to the tolerance of the service. User feedback can be gained to further enhance QoS.
[33]	Improves load balance and utilization rate.	The ABC algorithm has the disadvantages of premature convergence in the later search period. ACO algorithm also faces the problem of uncertain time to convergence. Thus, this can be a problem for the proposed hybrid approach.	Can be refined for dependent tasks. Also, assigning priority tasks and choosing the relevant VM for each job can be enhanced by considering the functions' QoS parameters.
[54]	Can be applied on any other algorithm to achieve better results.	It considers only CPU load to calculate the load on VMs, which might not give the best possible results.	RAM and Bandwidth utilization could be added along with CPU load.
[44]	Less complex	Only one parameter evaluation is taken into consideration, based on which results can't be surely trusted.	Reducing energy consumption
[45]	Through improved firefly feedback mechanisms and communications, search capabilities have increased to understand globally optimized solutions.	The focus is given only to the basic algorithm and one of its methods. Many other similar techniques exist which could have been compared.	VM Migrations has to be prevented and at the equal time gain lesser change of job migrations.
[27]	Improves the utilization and reduces the vacant space in a server. It is able to conceived for diverse size of the processors in the server.	The approach is limited to the offline process only.	Modifications on the proposed methodology can be done to achieve finer results.
[37]	Less complex	Out of all the QoS factors, only one aspect is considered.	The flight length can be dynamically varied in the long term, expanding the solution domain even further.
[61]	Both response rate and precision can be boosted in this up-to-date method and are considered very effective.	If the commencing population is not selected correctly, the algorithm then might not get the optimal solution.	Use the data deduplication algorithms for improvement.
[29]	Suitable for expansive cloud data centers.	Although MHBFA speeds up the confluence rate, it suffers auxiliary timing for crossover and mutation process.	Scope of performance improvement; exploration of operators and parameters are needed.
[51]	VM migration time calculation is considered with a specific algorithm which most of the other techniques have ignored.	It is challenging to run in a real cloud environment. Also, temporary peak load due to fundamental threshold values has not been taken into consideration.	Thermal-aware algorithm for VM placement.

Table 4: Summary of Evaluation Metrics considered for Surveyed Techniques

References	Makespan	Execution Time	Resource Utilization	Throughput	Reliability	Cost	Energy Consumption	Efficiency	Response Time	No. of Migrations	SLA
[45]	✓	✓						✓		✓	
[44]	✓										
[30]	✓	✓						✓			
[31]	✓					✓					
[33]	✓	✓	✓						✓		
[36]	✓		✓	✓	✓				✓		
[37]	✓										
[29]	✓						✓				
[25]	✓	✓				✓		✓			
[26]	✓										
[27]				✓					✓		
[61]											
[60]	✓	✓	✓	✓	✓				✓		
[53]							✓		✓		
[54]							✓			✓	
[55]							✓			✓	✓
[51]							✓				✓
[52]							✓				✓
[56]	✓	✓	✓				✓				
[24]										✓	
[28]						✓	✓				
[32]											
[34]	✓	✓	✓	✓	✓	✓					
[35]								✓	✓		
[38]							✓				
[39]	✓		✓			✓					
[40]	✓										
[41]	✓								✓	✓	
[42]		✓				✓					
[43]	✓	✓	✓		✓						
[46]		✓									
[47]		✓									
[48]						✓	✓		✓	✓	
[49]		✓									
[50]									✓		
[57]							✓				✓
[58]		✓					✓				
[59]			✓				✓				
[62]							✓			✓	✓

4. OPEN ISSUES AND FUTURE RESEARCH DIRECTIONS

The survey expressed in this study identifies some of the critical issues that cloud load balancing demands intensively. The main barriers and potential research opportunities are summarized as follows:

- The survey expressed in this study identifies some of the critical issues that cloud load balancing demands intensively. The main barriers and potential research opportunities are summarized as follows:
- The management of applications along with resources in the diverse cloud environment is a very complicated task, and therefore there is room for developing resource management strategies.
- The study shows that few clouds load balance works of literature are focused on migration costs apart from service-level violations. For future research, this can be viewed as a critical direction.
- Combining a meta-heuristic algorithm with another population-based meta-heuristic algorithm or a local search-based algorithm can improve the solution quality or convergence speed. However, realize that hybridizing algorithms can sometimes make the problem more complicated and lengthen the execution time. As a result, choosing a wise combination of algorithms is also a concern, and correlating the hybrid model is another way researcher can take.
- Future studies should concentrate on merging task scheduling and virtual machine consolidation techniques to increase scheduling performance.
- By redesigning the operator, the quality of the solutions given by meta-heuristic techniques can be enhanced. The algorithm's search technique can be boosted by changing initial population selection process and the fitness function.
- Complexities of the algorithms should be taken into account, as a significantly fewer number of studies have considered it for their performance evaluation. Time and space complexity can be crucial to have a clear idea about the scalability of the algorithms.
- More machine learning or deep learning-based approaches could be designed to predict future resources to balance the load effectively.
- To predict future overload/underload situations with high accuracy, better efficient workload prediction algorithms must be designed.
- Further research is also needed in the migration of tasks and the use of the host's electricity saving process; frequent task migrations and power switches could lead to additional energy consumption and even greater energy consumption as a whole.
- Assessing the customer service needs based on various application types may enhance the usage of resources.
- Although network bandwidth is critical in cloud computing due to high network congestion, the effectual networking elements application has not been adequately conveyed, according to the literature. This can result in a system crash, accidental deletion, and latency issues, among other things. More efficient and reliable load balancing algorithms for efficient high frequency network utilization should be developed.
- None of the existing techniques have taken into account all of the parameters revealed by the research. As a result, this work may encourage newcomers to the field to develop novel techniques that take into account nearly all metrics.
- Finally, very few of these approaches addressed security concerns, which is an essential aspect of cloud management. To avoid SLA violations, it is good to consider security as a parameter alongside other QoS parameters.

5. CONCLUSION

Load balancing of tasks on virtual machines or virtual machines on hosts is a major dispute in cloud computing that has gotten a lot of recognition from researchers. This article provides a current overview of load balancing problems and difficulties. According to this research, a comprehensive survey was conducted on various load-balancing methodologies using different parameters. The most recent and well-known methods have been briefly discussed. The benefits and limitations of various forms have been reviewed. We also discovered some future perspectives where the cloud needs to improve based on the literature review.

REFERENCES

1. S. Aslam and M. A. Shah, "Load balancing algorithms in cloud computing: A survey of modern techniques," 2015 National Software Engineering Conference (NSEC), Rawalpindi, Pakistan, 2015, pp. 30-35, doi: 10.1109/NSEC.2015.7396341.
2. K. Li, G. Xu, G. Zhao, Y. Dong and D. Wang, "Cloud Task Scheduling Based on Load Balancing Ant Colony Optimization," 2011 Sixth Annual Chinagrid Conference, Dalian, China, 2011, pp. 3-9, doi: 10.1109/ChinaGrid.2011.17.
3. Afzal, Shahbaz, and G. Kavitha. "Load balancing in cloud computing–A hierarchical taxonomical classification." *Journal of Cloud Computing* 8, no. 1 (2019): 1-24.
4. Lazaros Gkatzikis, Iordanis Koutsopoulos, "Migrate or Not? Exploiting Dynamic Task Migration in Mobile Cloud Computing Systems", IEEE Wireless Communications 2013, pp. 24-32.
5. Jain, Nancy, and Sakshi Choudhary. "Overview of virtualization in cloud computing." In *2016 Symposium on Colossal Data Analysis and Networking (CDAN)*, pp. 1-4. IEEE, 2016.
6. Ghobaei-Arani M, Souri A. LP-WSC: a linear

programming approach for web service composition in geographically distributed cloud environments. J Supercomp. 2018;1-26.

7. Shaw, Subhadra Bose, and A. K. Singh. "A survey on scheduling and load balancing techniques in cloud computing environment." In *2014 international conference on computer and communication technology (ICCCT)*, pp. 87-95. IEEE, 2014.
8. Jinhua Hu, Jianhua Gu, Guofei Sun, Tianhai Zhao, "A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment", 3rd International Symposium on Parallel Architectures, Algorithms and Programming, IEEE, 2010, pp. 89-96.
9. Anton Beloglazov, Rajkumar Buyya "Energy Efficient Resource Management In Virtualized Cloud Data Centers," 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010, pp. 826-831.
10. Amandeep Kaur Sidhu, Supriya Kinger , "Analysis of Load Balancing Techniques in Cloud Computing ", International Journal of Computers & Technology , Volume 4 No. 2, March-April, 2013, pp. 737-741
11. Mishra, S.K., Sahoo, B., Parida, P.P., Load Balancing in Cloud Computing: A big Picture, Journal of King Saud University - Computer and Information Sciences (2018), doi: https://doi.org/10.1016/j.jksuci.2018.01.003
12. Al Nuaimi, Klaithem, Nader Mohamed, Mariam Al Nuaimi, and Jameela Al-Jaroodi. "A survey of load balancing in cloud computing: Challenges and algorithms." In 2012 second symposium on network cloud computing and applications, pp. 137-142. IEEE, 2012.
13. Singh, P., Baaga, P., Gupta, S., 2016. Assorted load-balancing algorithms in cloud computing: a survey". Int. J. Comput. Appl. 143 (7).
14. Goyal, S., Verma, M.K., 2016. Load balancing techniques in cloud computing environment: a review. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 6 (4).
15. A. S. Milani, and N. J. Navimipour, "Load balancing mechanisms and techniques in the cloud environments: Systematic literature review and future trends," Journal of Network and Computer Applications, vol. 71, pp. 86-98, 2016
16. Ghomi, Einollah Jafarnejad, Amir Masoud Rahmani, and Nooruldeen Nasih Qader. "Load-balancing algorithms in cloud computing: A survey." *Journal of Network and Computer Applications* 88 (2017): 50-71.
17. Singh, Poonam, Maitreyee Dutta, and Naveen Aggarwal. "A review of task scheduling based on meta-heuristics approach in cloud computing." *Knowledge and Information Systems* 52, no. 1 (2017): 1-51.
18. Ahmad, Mohammad Oqail, and Rafiqul Zaman Khan. "Load balancing tools and techniques in cloud computing: a systematic review." *Advances in computer and computational sciences* (2018): 181-195.
19. Arunarani, A. R., D. Manjula, and Vijayan Sugumaran. "Task scheduling techniques in cloud computing: A literature survey." *Future Generation Computer Systems* 91 (2019): 407-415.
20. Pawan Kumar and Rakesh Kumar. 2019. Issues and Challenges of Load Balancing Techniques in Cloud Computing: A Survey. *ACM Comput. Surv.* 51, 6, Article 120 (February 2019), 35 pages. https://doi.org/10.1145/3281010
21. Afzal, Shahbaz, and G. Kavitha. "Load balancing in cloud computing–A hierarchical taxonomical classification." *Journal of Cloud Computing* 8, no. 1 (2019): 1-24.
22. Amini Motlagh, Aida, Ali Movaghar, and Amir Masoud Rahmani. "Task scheduling mechanisms in cloud computing: A systematic review." *International Journal of Communication Systems* 33, no. 6 (2020): e4302.
23. Mishra, Kaushik, and Santosh Majhi. "A State-of-Art on Cloud Load Balancing Algorithms." *International Journal of Computing and Digital Systems* 9, no. 2 (2020): 201-220.
24. Kumar, C. Ashok, R. Vimala, KR Aravind Britto, and S. Sathya Devi. "FDLA: fractional dragonfly based load balancing algorithm in cluster cloud model." *Cluster Computing* 22, no. 1 (2019): 1401-1414.
25. Geng, Xiaozhong, Yingshuang Mao, Mingyuan Xiong, and Yang Liu. "An improved task scheduling algorithm for scientific workflow in cloud computing environment." *Cluster Computing* 22, no. 3 (2019): 7539-7548.
26. Mallikarjuna, B., and P. Venkata Krishna. "A nature inspired bee colony optimization model for improving load balancing in cloud computing." *Int. J. Innov. Technol. Exploring Eng.* 8 (2018): 51-54.
27. Raj, P. Herbert, P. Ravi Kumar, and P. Jelciana. "Load Balancing in Mobile Cloud Computing using Bin Packing's First Fit Decreasing Method." In *International Conference on Computational Intelligence in Information System*, pp. 97-106. Springer, Cham, 2018.
28. Pradeep, K., and T. Prem Jacob. "A hybrid approach for task scheduling using the cuckoo and harmony search in cloud computing environment." *Wireless Personal Communications* 101, no. 4 (2018): 2287-2311.
29. Srichandan, Sobhanayak, Turuk Ashok Kumar, and Sahoo Bibhudatta. "Task scheduling for cloud computing using multi-objective hybrid bacteria foraging algorithm." *Future Computing and Informatics Journal* 3, no. 2 (2018): 210-230.
30. Al-Maytami, Belal Ali, Pingzhi Fan, Abir Hussain, Thar Baker, and Panos Liatsis. "A task scheduling algorithm with improved makespan based on prediction of tasks computation time algorithm for cloud computing." *IEEE Access* 7 (2019): 160916-160926.
31. Pang, Shanchen, Wenhao Li, Hua He, Zhiguang Shan, and Xun Wang. "An EDA-GA hybrid algorithm for multi-objective task scheduling in cloud computing." *IEEE Access* 7 (2019): 146379-146389.
32. Bacanin, Nebojsa, Timea Bezdan, Eva Tuba, Ivana Strumberger, Milan Tuba, and Miodrag Zivkovic. "Task

scheduling in cloud computing environment by grey wolf optimizer." In *2019 27th Telecommunications Forum (TELFOR)*, pp. 1-4. IEEE, 2019.

33. Gamal, Marwa, Rawya Rizk, Hani Mahdi, and Basem Elhady. "Bio-inspired based task scheduling in cloud computing." In *Machine Learning Paradigms: Theory and Application*, pp. 289-308. Springer, Cham, 2019.
34. Adhikari, Mainak, Sudarshan Nandy, and Tarachand Amgoth. "Meta heuristic-based task deployment mechanism for load balancing in IaaS cloud." *Journal of Network and Computer Applications* 128 (2019): 64-77.
35. Priya, V., C. Sathiya Kumar, and Ramani Kannan. "Resource scheduling algorithm with load balancing for cloud service provisioning." *Applied Soft Computing* 76 (2019): 416-424.
36. Golchi, Mahya Mohammadi, Shideh Saraeian, and Mehrnoosh Heydari. "A hybrid of firefly and improved particle swarm optimization algorithms for load balancing in cloud environments: Performance evaluation." *Computer Networks* 162 (2019): 106860.
37. Kumar, KR Prasanna, and K. Kousalya. "Amelioration of task scheduling in cloud computing using crow search algorithm." *Neural Computing and Applications* 32, no. 10 (2020): 5901-5907.
38. Gamal, Marwa, Rawya Rizk, Hani Mahdi, and Basem E. Elnaghi. "Osmotic bio-inspired load balancing algorithm in cloud computing." *IEEE Access* 7 (2019): 42735-42744.
39. Mapetu, Jean Pepe Buanga, Zhen Chen, and Lingfu Kong. "Low-time complexity and low-cost binary particle swarm optimization algorithm for task scheduling and load balancing in cloud computing." *Applied Intelligence* 49, no. 9 (2019): 3308-3330.
40. Gupta, Abhishek, H. S. Bhadauria, and Annapurna Singh. "Load balancing based hyper heuristic algorithm for cloud task scheduling." *Journal of Ambient Intelligence and Humanized Computing* (2020): 1-8.
41. Semmoud, Abderraziq, Mourad Hakem, Badr Benmammar, and Jean□Claude Charr. "Load balancing in cloud computing environments based on adaptive starvation threshold." *Concurrency and Computation: Practice and Experience* 32, no. 11 (2020): e5652.
42. Neelima, P., and A. Rama Mohan Reddy. "An efficient load balancing system using adaptive dragonfly algorithm in cloud computing." *Cluster Computing* (2020): 1-9.
43. Tapale, Manisha T., Rayangouda H. Goudar, Mahantesh N. Birje, and Rudragouda S. Patil. "Utility based load balancing using firefly algorithm in cloud." *Journal of Data, Information and Management* 2, no. 4 (2020): 215-224.
44. Farrag, Aya A. Salah, Safia Abbas Mohamad, and M. El Sayed. "Swarm Optimization for Solving Load Balancing in Cloud Computing." In *International Conference on Advanced Machine Learning Technologies and Applications*, pp. 102-113. Springer, Cham, 2019.
45. Padmavathi, M., Sk Mahaboob Basha, and VV Jaya Rama Krishnaiah. "Load Balancing Algorithm to Reduce Make Span in Cloud Computing by Enhanced Firefly Approach." In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 896-900. IEEE, 2020.
46. Rajagopalan, Aravind, Devesh R. Modale, and Radha Senthilkumar. "Optimal scheduling of tasks in cloud computing using hybrid firefly-genetic algorithm." In *Advances in decision sciences, image processing, security and computer vision*, pp. 678-687. Springer, Cham, 2020.
47. Miao, Zhang, Peng Yong, Yang Mei, Yin Quanjun, and Xie Xu. "A discrete PSO-based static load balancing algorithm for distributed simulations in a cloud environment." *Future Generation Computer Systems* 115 (2021): 497-516.
48. Kaur, Gursharanjit, and Rajan Sachdeva. "Virtual Machine Migration Approach in Cloud Computing Using Genetic Algorithm." In *Advances in Information Communication Technology and Computing*, pp. 195-204. Springer, Singapore, 2021.
49. Bawankule, Kamalakant Laxman, Rupesh Kumar Dewang, and Anil Kumar Singh. "Load Balancing Approach for a MapReduce Job Running on a Heterogeneous Hadoop Cluster." In *International Conference on Distributed Computing and Internet Technology*, pp. 289-298. Springer, Cham, 2020.
50. Panigrahi, Amrutanshu, Bibhuprasad Sahu, Saroj Kumar Rout, and Amiya Kumar Rath. "M-Throttled: Dynamic Load Balancing Algorithm for Cloud Computing." In *Intelligent and Cloud Computing*, pp. 3-10. Springer, Singapore, 2019.
51. Yadav, Rahul, Weizhe Zhang, Omprakash Kaiwartya, Prabhat Ranjan Singh, Ibrahim A. Elgendy, and Yu-Chu Tian. "Adaptive energy-aware algorithms for minimizing energy consumption and SLA violation in cloud computing." *IEEE Access* 6 (2018): 55923-55936.
52. Wang, Hui, and Huaglory Tianfield. "Energy-aware dynamic virtual machine consolidation for cloud datacenters." *IEEE Access* 6 (2018): 15259-15273.
53. Jeba, Jenia Afrin, Shanto Roy, Mahbub Or Rashid, Syeda Tanjila Atik, and Md Whaiduzzaman. "Towards green cloud computing an algorithmic approach for energy minimization in cloud data centers." *International Journal of Cloud Applications and Computing (IJCAC)* 9, no. 1 (2019): 59-81.
54. Patel, Durga, Rajeev Kumar Gupta, and R. K. Pateriya. "Energy-aware prediction-based load balancing approach with VM migration for the cloud environment." In *Data, engineering and applications*, pp. 59-74. Springer, Singapore, 2019.
55. Yavari, Maede, Akbar Ghaffarpour Rahbar, and Mohammad Hadi Fathi. "Temperature and energy-aware

consolidation algorithms in cloud computing." *Journal of Cloud Computing* 8, no. 1 (2019): 1-16.

56. BEN ALLA, Said, Hicham BEN ALLA, Abdellah Touhafi, and Abdellah Ezzati. "An efficient energy-aware tasks scheduling with deadline-constrained in cloud computing." *Computers* 8, no. 2 (2019): 46.
57. Soltanshahi, Minoo, Reza Asemi, and Nazi Shafiei. "Energy-aware virtual machines allocation by krill herd algorithm in cloud data centers." *Heliyon* 5, no. 7 (2019): e02066.
58. Zhao, Da-Ming, Jian-Tao Zhou, and Keqin Li. "An energy-aware algorithm for virtual machine placement in cloud computing." *IEEE Access* 7 (2019): 55659-55668.
59. Abohamama, A. S., and Eslam Hamouda. "A hybrid energy–Aware virtual machine placement algorithm for cloud environments." *Expert Systems with Applications* 150 (2020): 113306.
60. Geetha, P., and CR Rene Robin. "Power conserving resource allocation scheme with improved QoS to promote green cloud computing." *Journal of Ambient Intelligence and Humanized Computing* (2020): 1-12.
61. Devaraj, A. Francis Saviour, Mohamed Elhoseny, S. Dhanasekaran, E. Laxmi Lydia, and K. Shankar. "Hybridization of firefly and Improved Multi-Objective Particle Swarm Optimization algorithm for energy efficient load balancing in Cloud Computing environments." *Journal of Parallel and Distributed Computing* 142 (2020): 36-45.
62. Mandal, Riman, Manash Kumar Mondal, Sourav Banerjee, and Utpal Biswas. "An approach toward design and development of an energy-aware VM selection policy with improved SLA violation in the domain of green cloud computing." *The Journal of Supercomputing* (2020): 1-20.