IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 1 August 2023, accepted 5 September 2023, date of publication 12 September 2023, date of current version 3 October 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3314613

TOPICAL REVIEW

A Comprehensive Review of Cloud Computing Virtual Machine Consolidation

JASPREET SINGH AND NAVPREET KAUR WALIA
Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab 140413, India
Corresponding author: Jaspreet Singh (jaspreet.rajal@gmail.com)

ABSTRACT In the last decade, users have been able to access their applications, data, and services via the cloud from any location with an internet connection. The scale of heterogeneous cloud environments is continuously growing due to the development of computing-intensive smart devices. The cloud computing system is managed by a data center, which consists of physical machines (PMs) or servers and software-based emulation of PMs called virtual machines(VMs). The deployment of a huge number of physical servers as a result of the exponential development in demand for cloud services has resulted in high energy consumption and ineffective resource usage. Efficient utilization of resources and minimizing power consumption in any data center have become crucial challenges. Virtual machine consolidation (VMC) is a method of optimizing computing resources by consolidating multiple VMs onto a reduced number of PMs. By consolidating VMs and running fewer physical servers, VM consolidation can reduce power consumption and improve resource utilization. This review paper presents a comprehensive analysis of cloud computing virtual machine consolidation, exploring various strategies, benefits, challenges and future trends in this domain. By examining a wide range of literature from the year 2015 to 2023, this review attempts to provide insight into the current state of VM consolidation and its possible effects on the performance and sustainability of cloud computing. The main flaw in the articles is that the various authors focused on different assessment metrics when the emphasis should have been on increasing cloud system service quality and energy efficiency. Future research can be aimed at developing a multi-objective system that emphasizes minimizing cloud energy usage without sacrificing service quality and preventing service level agreements with cloud users from being compromised.

INDEX TERMS Cloud computing, energy efficiency, power consumption, service level agreement, virtual machine consolidation.

I. INTRODUCTION

Cloud computing (CC) is a type of computing that shares computing resources instead of a local server or personal device to handle any application [1], [2]. CC provides a variety of services based on the needs of the user. A service level agreement (SLA) specifies the services and functionalities that the provider will deliver to the user. It specifies what is included and what is not covered by the agreement. If the service provider fails to meet the agreed-upon performance metrics or targets outlined in the SLA, then it is called a violation of the service level agreement [3].

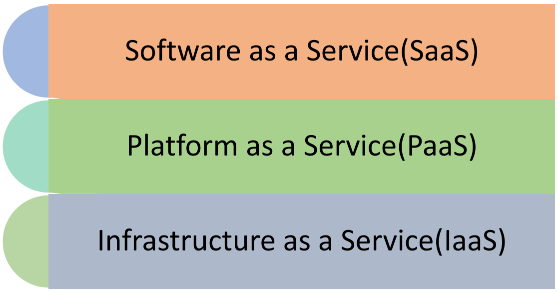

FIGURE 1. Models of cloud computing service [18].

CC provides three types of services to the users, which are called service models as represented in Figure 1, named Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) [4], [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Agostino Forestiero.

IaaS provides access to a group of computing resources such as servers, storage, and networking infrastructure. The infrastructure's software and applications must be managed and maintained by the user [4]. e.g. Microsoft Azure Virtual Machines, Google Cloud Platform Compute Engine.

PaaS provides a platform for building, deploying, and managing applications. The user is responsible for developing and running applications on the platform [6]. e.g. IBM Cloud Foundry, Salesforce Platform (force.com), etc. Users can access software programs through SaaS via the Internet. The underlying infrastructure must be managed and kept up to date by the service provider [7]. e.g. Google Workspace, Dropbox, Slack.

Cloud services can be provided in a homogeneous or heterogeneous environment [8]. In a homogeneous cloud environment, the cloud infrastructure is built using a uniform or standardized set of technologies and components, while in a heterogeneous cloud environment, the cloud infrastructure consists of diverse technologies, components, and platforms. Due to the variety of end user's platforms and technology choices, heterogeneous clouds are most frequently deployed.

CC is gaining popularity as a result of important features, which include [9]

- On-demand scalability [10]: Users can easily increase or decrease their computing resources based on their needs.
- Cost-effectiveness: CC is a cost-effective choice for users because they only pay for the resources that they utilize [11].
- Reliability [12]: Due to its distributed system architecture and multiple redundant resources, it is reliable.
- Agility: It allows businesses to quickly adapt to changing market conditions [13].
- Security [14]: It offers high levels of security through encryption and data backups [15].
- Efficiency [16]: CC enables businesses to access computing resources and optimize their IT infrastructure more efficiently.
- Flexibility [17]: Users can access their data and applications from anywhere, anytime.

This paper is structured as follows: A basic introduction to cloud computing and related service paradigms is presented in the first section. The motivation for writing a paper is outlined in the second section. Virtual machine consolidation and its process are covered in the third section. Section IV describes the methodology used to conduct the literature review. Section V summarized the study on VM consolidation that has been conducted since 2015. Section VI discussed cloud environments based on resources, evaluation parameters, methodology, and the number of PMs and VMs based on the literature review. Section VII presented the open issues proposed by the literature. The conclusion and further work are covered in Section VIII.

II. MOTIVATION FOR REVIEW

The motivation behind conducting a comprehensive review of VMC stems from the increasing significance of resource management and energy efficiency in large-scale cloud environments. The demand for cloud services has led to the deployment of a vast number of physical servers, resulting in high energy consumption and inefficient utilization of resources. To address these challenges, VMC has gained considerable attention as a technique to improve the overall performance and sustainability of cloud computing systems.

The various abbreviations and their descriptions are presented in Table 1.

TABLE 1. Abbreviations and their description.

Abbreviation	Description
CC	Cloud Computing
CPU	Central Processing Unit
MBFD	Modified Best Fit Decreasing
PM	Physical Machine
PDM	Performance Degradation due to Migration
RAM	Random Access Memory
SLA	Service Level Agreement
SLA-V	Service Level Agreement Violation
VM	Virtual Machine
VMC	Virtual Machine Consolidation

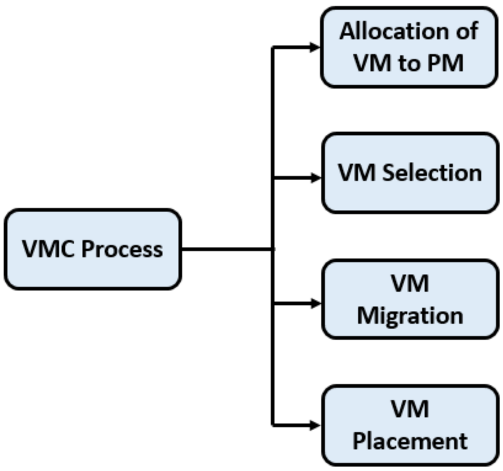

FIGURE 2. VMC process [34].

III. VIRTUAL MACHINE CONSOLIDATION

Virtualization is the process of creating a virtual version of hardware, software, storage, or network resources. VM consolidation is the process of combining or consolidating multiple VMs onto a smaller number of physical servers or machines [19], [20]. It allows IT administrators to run multiple applications and operating systems on the same physical hardware, reducing hardware costs and improving resource utilization. Consolidation is done to make better use of physical resources like CPU, memory, and storage. To achieve this, virtualization software like VMware can be used to create and manage virtual computers on a physical host [21].

VM consolidation has several benefits, which include [22]:

- Lower hardware costs by reducing the number of physical servers.
- Improved server utilization and efficiency by running multiple applications and operating systems on the same hardware.
- Increased flexibility and scalability through the easy addition or removal of virtual machines.
- Lower energy and cooling costs by reducing the amount of physical hardware.
- Improved disaster recovery capabilities by enabling the quick replication of virtual machines.

The steps involved in VM consolidation (VMC) are as follows, as represented in figure 2.

A. VM ALLOCATION TO THE PHYSICAL MACHINE (PM)

VMs can be allocated to PMs to limit resource waste and increase the computation efficiency of physical machines [23]. This is beneficial because it helps to save on hardware and energy costs while also allowing users to access multiple applications and services without having to purchase multiple physical machines. Several algorithms are used for VM allocation to PM in virtualized environments. The choice of algorithm depends on the specific requirements of the system, workload characteristics, and the goals of resource utilization and performance optimization.

Here are some common algorithms for VM allocation to PM [24]:

- First Fit (FF) [25]: It is a simple and quick algorithm that allocates a VM to the first available PM that has enough resources to accommodate it. It reduces the search time for a suitable PM, but it may lead to suboptimal resource utilization as it may not consider better-fitting PMs later in the list.
- Best Fit (BF) [26]: It selects the PM that has the least amount of unused resources after allocating the VM. It aims to minimize fragmentation and make efficient use of resources. However, this algorithm can be more computationally intensive than First Fit due to the need to search for the best-fitting PM.
- Worst Fit (WF) [27]: It allocates the VM to the PM that has the most available resources after the allocation. It can be useful in scenarios where there is a preference for keeping larger contiguous spaces free for future allocations.
- Next Fit (NF) [28]: Next Fit is similar to First Fit but remembers the last allocated PM. It starts searching from the last allocated PM, which can be useful for sequential allocation patterns or when dealing with real-time workloads.
- Random Fit (RF) [21]: It randomly selects a PM for the VM allocation. While it is simple to implement, it may not be suitable for situations that require careful resource balancing or workload optimization.

BFD (best-fit decreasing) and MBFD (modified best-fit decreasing) are the most common algorithms that are often used in the context of VM allocation to PM in virtualized environments [29]. MBFD is an extension of BFD. The main difference between BFD and MBFD for VM allocation lies in the refinement step of MBFD. BFD performs the initial placement of VMs based on the best-fit decreasing strategy without any additional rearrangement. On the other hand, MBFD adds an extra step to consolidate resources within PMs after the initial placement to potentially achieve a more compact and efficient VM-to-PM allocation.

The MBFD algorithm is the most widely used technique for allocating VMs [6], [30].This algorithm places VMs in decreasing order of CPU usage capacity before assigning the VM to the host that uses the least amount of power. This algorithm allows for the initial selection of the most energy-efficient PM.

MBFD works as follows [6], [31]:

Step 1: Sort the list of virtual machines according to their resource requirements in decreasing order.

Step 2: Start with the first virtual machine in the sorted list. Check if the physical machine has enough capacity to host the virtual machine.

Step 3: If the PM has enough capacity, then allocate the VM to this PM.

Step 4: Otherwise, go to the next physical machine and check if it has enough capacity.

Step 5: Repeat steps 3 and 4 until the VM is allocated to a PM.

Step 6: Continue with the next virtual machine in the sorted list and repeat steps 2-5 until all the virtual machines are allocated to PMs.

B. SELECTION OF VM

VMs will be assigned to PMs based on their resource needs. According to research, an ideal PM has been found to use 70% of the energy required for the peak process [10], [33], [34]. This kind of PM is known as an "underutilized" PM. Similarly, some PMs might be overused and use more power. The underutilized PM eventually loses all of its pertinent VMs in order to conserve power. The PM's under and overutilization conditions are determined by CPU usage and dual-threshold policies. The dual threshold policy makes use of the variables "x" and "t," where "x" may be the average of the evaluating parameter, which is CPU utilization, and "t" is the threshold margin, which should be a finite amount. The PMs that reach the upper threshold are considered to be overutilized.

To migrate VMs from overloaded PMs, selection procedures must be used [35]. To choose a VM to migrate from one PM to another, various selection policies are offered. Some of the policies for selection are as follows [36], [37]:

- Random Choice: Pick a VM at random to be migrated using the random choice policy.

- Minimum Utilization Policy: VMs are selected depending on CPU utilization. VMs that consume the least CPU power are selected.
- Minimum Migration Time Policy: The VM that migrates in the least amount of time is selected in comparison with the other VMs.
- Highest Potential Growth Policy: In order to reduce the likelihood that PM usage would increase and prevent a breach of the service level agreement, the policy migrates the VMs with the lowest CPU utilization in comparison to the CPU capacity specified by the VM parameters when the upper threshold is reached.
- Minimization of Migration Policy: Once the higher threshold has been attained, the policy selects a small number of VMs to migrate to reduce CPU usage. This policy only takes into account CPU usage.

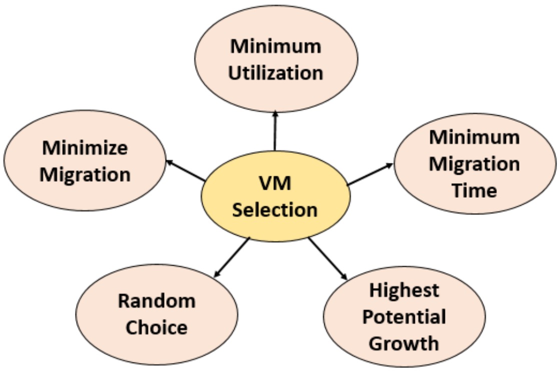

FIGURE 3. Policies for selecting virtual machines [38], [36].

C. MIGRATION OF VMS

VM migration is the process of transferring a VM from one PM to another without any disruption in its operation or services [13], [39]. It reduces the amount of energy consumed by the cloud data center while enhancing performance, balancing the load across physical servers, and effectively managing resources [8]. There are two methods of VM migration, as represented in figure 4: hot migration (live migration) and cold migration (non-live migration).

- Cold Migration: This is a type of VM migration in which the virtual machine is powered off before being moved to the new host. It is also known as a "non-live" or "offline" migration [24].
- Hot Migration: It is the process of migrating a VM from one host to another while keeping the VM running. This is done by transferring the memory state and contents of the VM from one host to another. It is also known as a "live" or "online" migration [24].

D. PLACEMENT OF VM

VM placement refers to the process of choosing the optimal physical host or server on which to deploy a VM after migration [41]. VM placement algorithms consider factors such as the number of applications running on the VM, the amount of memory and computing resources needed, the expected workload, and the availability of physical servers [42]. The goal of VM placement is to maximize the performance of the cloud computing environment while minimizing the cost of operating it. There are numerous methods, like simulated annealing, evolutionary algorithms, and heuristics, that are used to address the VM placement problem [43]. These algorithms can be used to optimize the placement of VMs, improving the performance and cost-effectiveness of the cloud computing infrastructure.

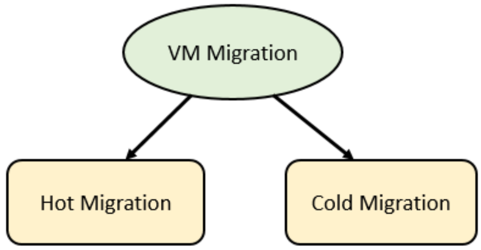

FIGURE 4. VM Migration classification [40].

The VM placement process can be classified as [44]:

- **Static Placement:** VMs are assigned to physical hosts based on predefined rules or manual decisions. Rules can be designed based on resource requirements and service level agreements between the user and service provider. This approach is simple but not suitable for dynamic workloads.
- **Dynamic Placement:** VMs are placed in real time based on current resource usage and demands. It considers factors like resource utilization, network latency, and historical data to make optimal placement decisions.
- **Load-Based Placement:** VMs are placed on hosts based on current load levels to balance resource usage and avoid overloading any particular host.
- **Power-Aware Placement:** Placement decisions take power usage into account and try to combine VMs on fewer hosts to shut down idle servers whenever possible.

IV. PROCESS OF PAPER SELECTION

The objective of this paper is to present a review of virtual machine consolidation in cloud computing. To accomplish this, a comprehensive assessment of the literature is required. To examine the most significant publications on VM consolidation, a selection procedure was applied and put into practice. The method used to select the literature is explained in depth in this section.

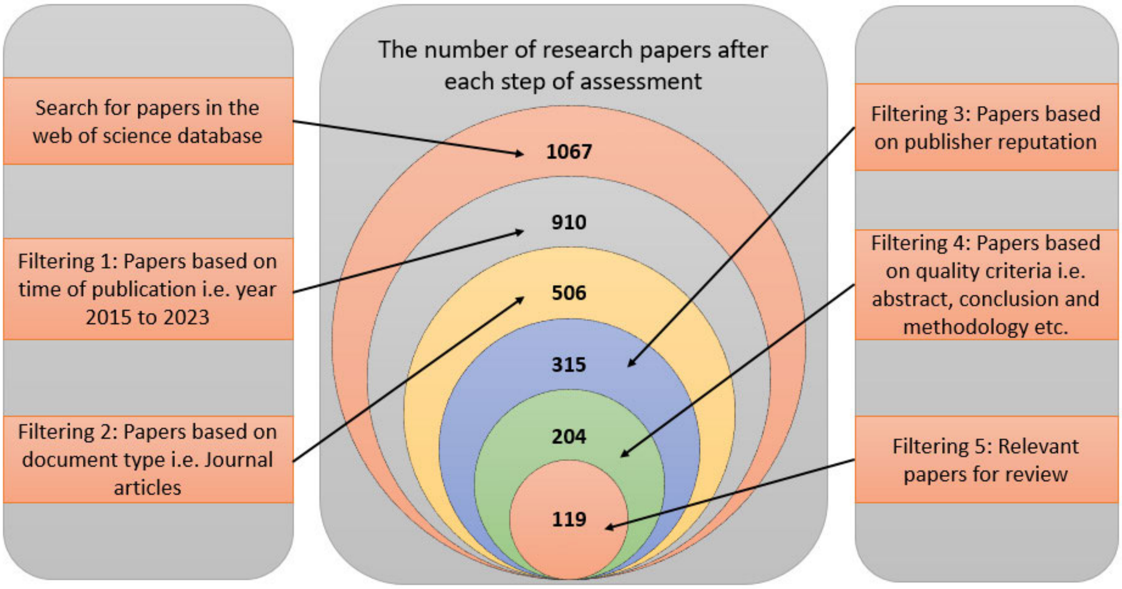

FIGURE 5. Paper selection.

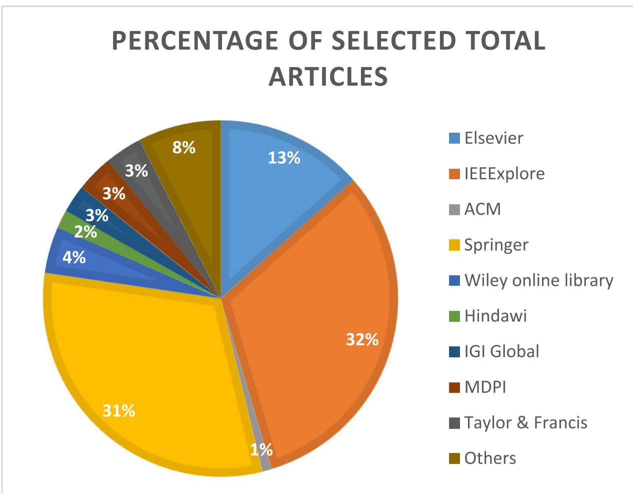

FIGURE 6. Percentage of selected total articles.

A. INCLUSION AND EXCLUSION CRITERIA FOR PAPERS

The first step in selecting relevant publications is to search the Web of Science database for research papers with the keywords "virtualization," "consolidation," and "cloud computing" in the title of the article. The logical combination of the words "virtual," "machine," "consolidation," "cloud," and "computing" more precisely defines the keywords. A total of 1067 research papers were made available as a result of the evaluation of the logical expression "virtual machine consolidation in cloud computing". Filtration of 1067 papers is done at different levels to select the most appropriate papers for review. Figure 5 represents the strategy used for inclusion and exclusion criteria for papers. The percentage of selected articles is depicted in figure 6.

- Filtering 1: Papers are selected based on a time frame, i.e. from the year 2015 to 2023. 910 papers were selected for further evaluation.
- Filtering 2: Based on article type, the selection results were reduced to 506 papers from 910. Article type includes papers from journals and IEEExplore. It did not include book chapters or papers, which are not indexed anywhere.
- Filtering 3: 315 papers were filtered based on the publisher's reputation. Articles from IEEE, Elsevier, Springer, and ACM were given preference. Figure 6 displays the proportion of each publisher's articles.
- Filtering 4: The most significant papers that dealt with virtual machine consolidation were selected by reading the abstracts, conclusions, and methodology of 315 publications. The selection resulted in 204 papers.
- Filtering 5: Out of 204 papers,119 papers were found to be relevant for review. The selection is done based on removing common papers after the fourth level of filtration.

V. RELATED WORK

Several research studies have investigated the various aspects of VM consolidation, such as optimization techniques, resource allocation strategies, and energy efficiency. The research studies specify two types of consolidation, i.e., static and dynamic VM consolidation. In static VM consolidation, the aim is to maximize the number of VMs that can be hosted on a single physical machine. This is done by optimizing the resource allocation and scheduling of VMs [45]. On the other hand, dynamic VM consolidation is aimed at minimizing the total energy consumption while still satisfying the performance requirements of the VMs [41].

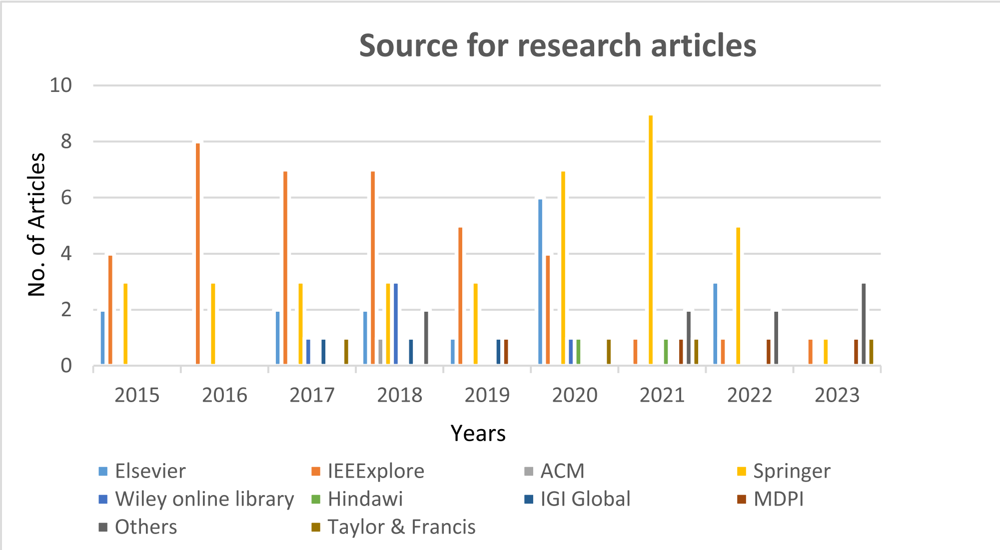

FIGURE 7. Number of publications made by publishers.

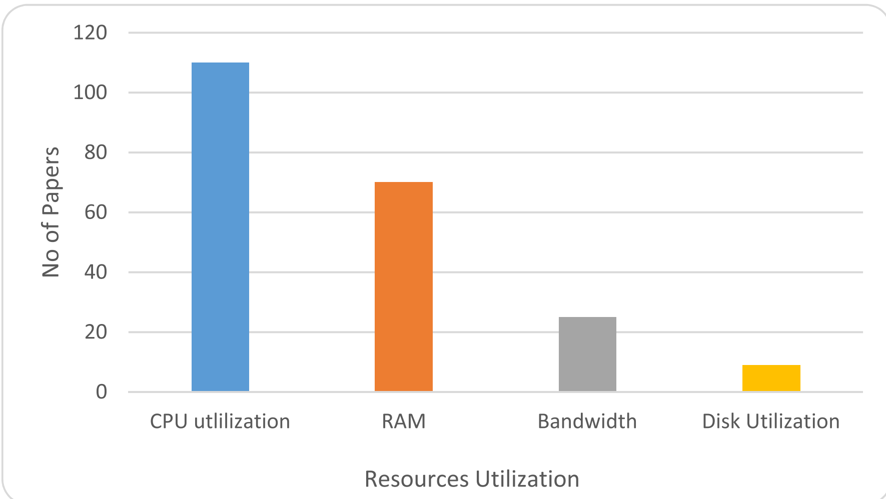

FIGURE 8. Analysis of resources by different authors.

Table 2 represents a summary of some significant literature that attempts to address the problems of virtual machine consolidation in cloud computing by more intelligently allocating and migrating VMs.

It is concluded from the literature that different resources like CPU, RAM, bandwidth, and disk utilization are used for VM consolidation. To evaluate the performance of proposed consolidation techniques, several metrics are used, i.e. energy consumption, SLA violation, number of VM migrations, number of active PMs, number of hosts shutdown, performance degradation due to migration (PDM), migration cost, etc. Table 3 represents the resources and evaluation metrics used by different authors in the literature. It is clear from Table 3 that many authors who have contributed to literature surveys have focused on energy efficiency while ignoring service level agreement breaches. Power usage must be kept to a minimum in order to conserve energy. Power consumption can be decreased, but not at the expense of service level agreements. Thus, in order to reduce SLA-V and power consumption simultaneously, a multi-objective function must be developed.

VI. DISCUSSION

The preceding sections covered a variety of techniques for consolidating VMs. In this section, the suggested techniques were evaluated concerning the data center and cloud environments based on resources, evaluation parameters, methodology, and the number of PMs and VMs.

A. RESOURCES UTILIZATION ANALYSIS AND COMPARISON

To meet the demands of numerous applications at cloud service levels, cloud computing offers access to a variety

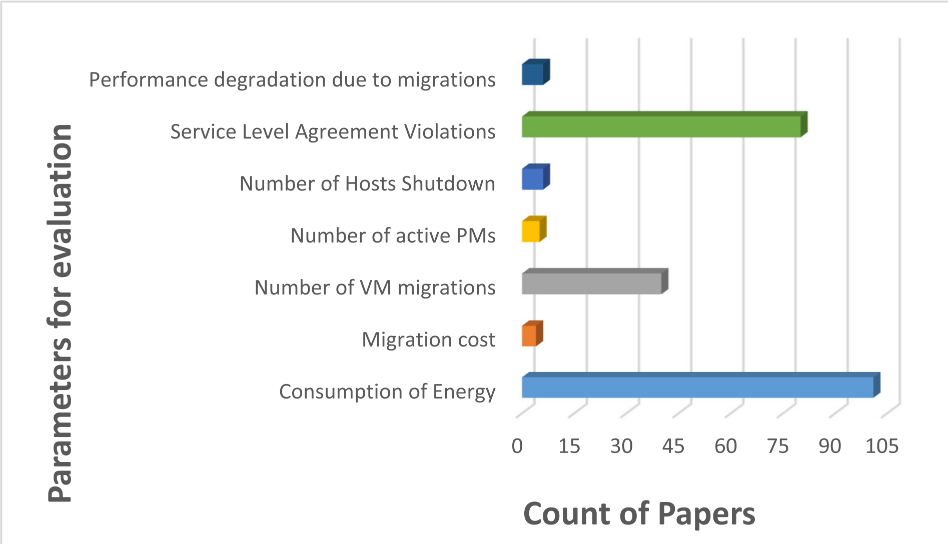

FIGURE 9. Parameters for evaluation by various authors.

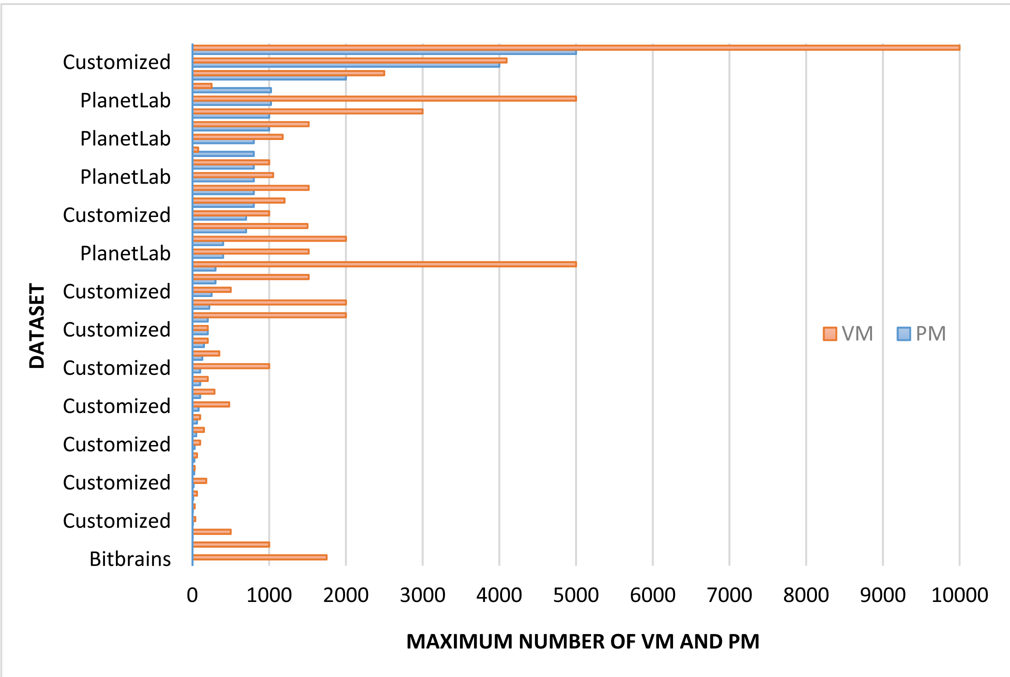

FIGURE 10. Consideration of VM and PM based on the dataset.

of resources like CPU, RAM, bandwidth, and disk usage. The CPU is a fundamental component of all computing systems, its utilization is regarded as a key resource, and it influences the VM consolidation procedure [117], [118]. Power consumption and CPU utilization are closely related. The system is using more energy because of higher CPU utilization, which results in higher power consumption [119]. The primary goal is to use less CPU to lower the energy consumption of any cloud. Some papers included network bandwidth and disk usage as resource parameters in addition to CPU utilization. Figure 8 shows that the majority of papers focus on CPU utilization, followed by RAM.

B. EVALUATION PARAMETER-BASED ANALYSIS AND COMPARISON

A variety of parameters can be used to evaluate the performance of a cloud system. This section outlines the methods used to choose the studies to be examined as well as how to evaluate the solutions presented in each study. Each metric's frequency is shown in Figure 9 for comparison. According to the statistics in Figure 9, consumption of energy and SLA violations are the two major concerns of cloud computing during VM consolidation. The cloud data center must be available all the time in order to accommodate the user's changing demands. As a result, cloud data centers continually use

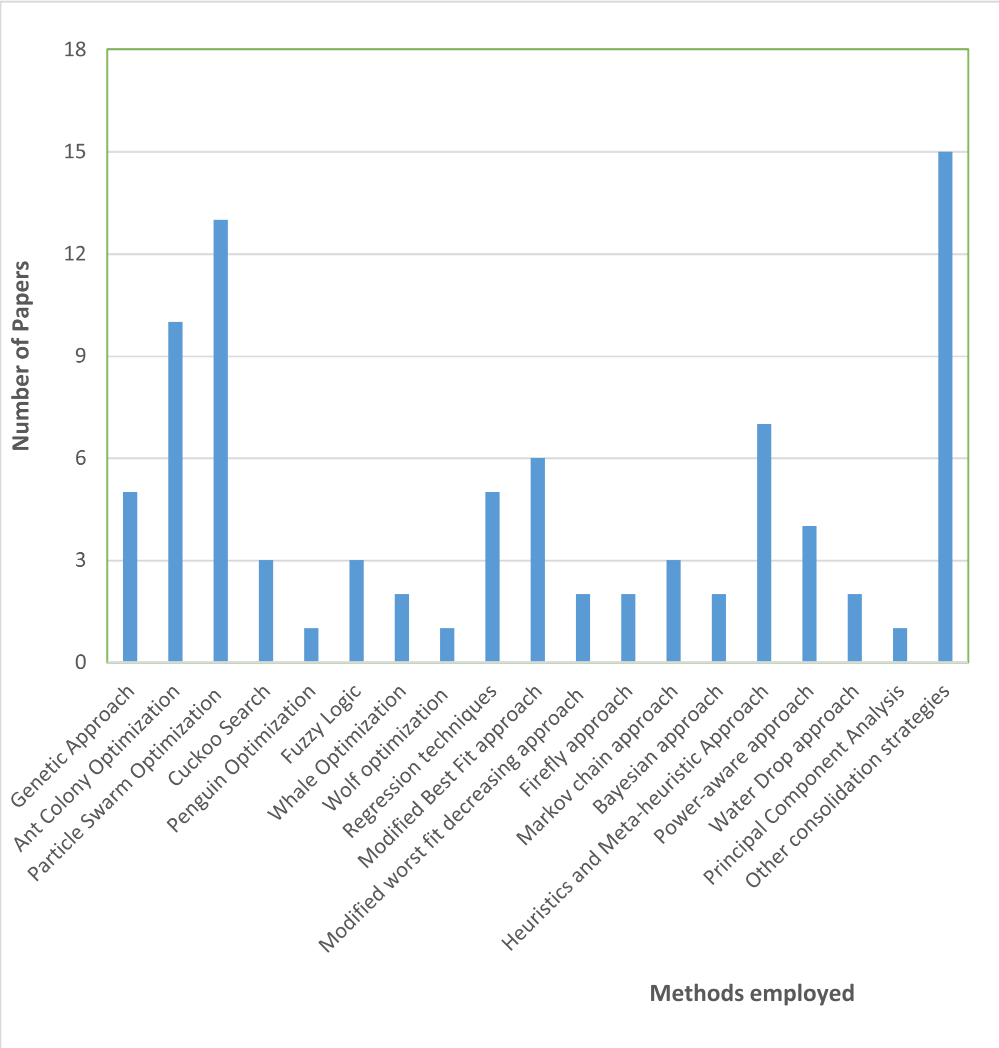

FIGURE 11. Methodologies used by various researchers.

electricity to keep their operations running smoothly, which results in energy consumption and the emission of carbon dioxide. While accommodating cloud user's needs, energy consumption must be prioritized. Service level agreements (SLA) evaluate the performance of cloud service providers and user satisfaction [13], [44]. SLA violations occur when a cloud provider is unable to satisfy user expectations. The total number of VM migrations is the third parameter that various researchers have looked at in addition to energy usage and SLA. It also details the number of VMs that were moved throughout the migration process. Other considerations include the effect of migration on performance, the number of hosts that are shut down, the number of PMs in use, and the migration's overall cost.

C. COMPARISON AND ANALYSIS DEPENDING ON THE QUANTITY OF VM AND PM

With the use of virtualization, cloud providers run many operating systems and applications on a single PM which helps maximize resource utilization. VMs are frequently built on top of physical machines. The analysis using the PM and VM values while taking into account the dataset is shown in Figure 10. A customized dataset or a dataset from Planet Lab was utilized in the great majority of papers for cloud simulation analysis. The analysis can take up to 5,000 and 10,000 considerations for PM and VM, respectively.

D. METHODOLOGY-BASED ANALYSIS AND COMPARISON

Many approaches are used in the consolidation of virtual machines, as shown in Figure 11. The most common method for creating more effective, energy-efficient cloud computing models is particle swarm optimization [12], [67], [99], [120]. It is accompanied by an ant colony approach for allocating VMs with a limited number of active hosts. The detection of server over and under-utilization, the selection of the proper VM, and the placement of the proper VM on PM are all possible using a variety of additional consolidation mechanisms. These approaches are based on migration times, maximum utilization, workload forecasting, consolidation of virtual machines based on normalization, reinforcement learning, look-ahead energy-efficient resource allocation, etc.

TABLE 2. VM consolidation literature review and their findings.

Year	Author	Algorithm/Method	Findings	Limitations & Future Scope
2015	Okada et al.[46]	First and best fit decreasing algorithm	It is quick and effective to use the first fit-decreasing method.	Dynamic programming and ant colony optimization are more effective ways to handle VM placement issues.
2015	Chowdhury et al.[47]	Modifications were made to the first and second worst-fit decreasing algorithms for host VM placement.	Decreased power usage and SLA violations, with less performance degradation in comparison to power-aware best-fit decreasing	A Fuzzy strategy can be used for VM selection and takes advantage of several selection criteria.
2015	Monil et al.[48]	The overload detection and selection mechanism with migration control has been modified.	Decreased energy use and SLA violations	Additional evaluation criteria and resources are required.
2015	Gondhi et al.[49]	Future load forecasts are made before the allocation of any VMs to PMs, and the allocation of VMs is done by modified best fit algorithm.	Energy savings, increased quality of service, and performance in comparison to the heuristic regression allocation of the power VM	For more time and energy savings, the prediction technique needs to be enhanced.
2016	Tchana et al.[50]	Dynamic software consolidation	Reduced energy consumption in private clouds by approximately 40% and VM costs on Amazon EC2 by approximately 40.5%.	The consolidation of software on VMs and the consolidation of VMs on PM must be coordinated to further maximize power gains.
2016	Selim et al.[51]	Algorithm for detection of over-utilized hosts based on variance in CPU utilization	Reduction in energy consumption and a healthier ecology.	It has to be tested using open-source tools like OpenStack in a real-world cloud environment.
2016	Chien et al.[52]	An improved minimization of migration algorithm	Improvement in terms of lower energy usage, fewer migrations, and fewer SLA breaches.	The proposed technique should be tested in a real-world environment to clearly understand how well it functions.
2016	Liu et al.[53]	An order exchange and migration-based local search approach called OEM. OEM is combined with an ant colony system (ACS) approach to create a method called OEMACS for placing virtual machines.	OEMACS performed better than evolutionary-based methodologies and heuristic techniques. A large amount of energy was saved, and resources were used more effectively as a result.	While evaluating the suggested algorithm, SLA violations are ignored.
2016	Deng et al.[54]	It proposes a virtual machine consolidation architecture to improve energy efficiency.	Decreased energy use by 1.15%, SLAV by 58.7%, virtual machine migrations by 20.6%, and overall energy efficiency metrics by 59.8%.	There is a need to consider additional evaluation criteria and resources.
2016	Kaur et al.[30]	The VM placement problem is addressed with the MBFD technique, which aims to save energy and minimize VM migrations.	Reduction in energy usage and the number of migrations	More resources can be considered while deploying a VM. The suggested method needs to be tested in a real-time cloud context.
2016	Kansal et al.[55]	Firefly Algorithm	It utilized less energy and conserved hosts and migrations on average by 34.36% and 72.34%, respectively, as compared to ant colony optimization and first fit decreasing.	SLA violations are not evaluated. Validation of the proposed technique in a real cloud environment is required.
2016	Farahnakian et al.[56]	To address the multi-objective vector bin packing issue, it suggested utilization prediction aware VM consolidation. A prediction model based on regression is used to forecast how resources will be distributed in the future.	Based on energy use, effectiveness, and migration frequency, it outperformed heuristics and meta-heuristic methods.	To increase the scalability of the proposed architecture, a two-tier hierarchical architecture must be replaced with a three-tier hierarchical architecture. The method has to be tested in a real-world cloud setting with different resource types and criteria.
2016	Li et al.[57]	Modified Particle Swarm Optimization	In terms of energy efficiency, modified particle swarm optimization outperformed MBFD. It also required fewer VM migrations and improved the distribution of system resources.	Only traditional heuristic methods are compared (MBFD and MPSO algorithms). To generalize the findings, a comparison with alternative methodologies is also necessary.
2017	Zahedi Fard et al.[58]	It suggests an enhanced method for consolidating virtual machines that includes	The dynamic threshold maximum fit method outperformed previous dynamic VM consolidation solutions in terms of	The methods used to consolidate VMs must be enhanced to account for extra aspects like memory and storage in

TABLE 2. ***(Continued.)*** **VM consolidation literature review and their findings.**

Year	Author	Algorithm/Method	Findings	Limitations & Future Scope
		algorithms for choosing and positioning the VMs as well as identifying overloaded and under-loaded servers.	parameters such as power consumption, SLA violations, and live migrations.	addition to network components like routers, switches, and bandwidth.
2017	Wang et al.[59]	Resource correlations between co-located VMs and host utilization rates are taken into account by heuristic-based mechanisms for VM allocation. According to the suggested technique, a lower heuristic value indicates a better working environment for hosts and VMs.	The suggested VM allocation mechanism decreased machine overloading, service level agreement breaches, and energy usage.	It is necessary to take into account additional resources for evaluation.
2017	Abd et al.[60]	DNA-based Fuzzy Genetic Algorithm (DFGA)	Less power consumption as compared to energy-aware rolling horizon scheduling and the modified best fit decreasing algorithm.	It's necessary to determine the energy consumption of different system components, like network hardware and appliances, to increase energy efficiency.
2017	Abdessamia et al.[27]	Particle swarm optimization	Particle swarm optimization was found to be the best option for low energy usage as compared to the first, best, and worst-fit techniques.	SLA violations are not taken into account. In the future, it will have to be evaluated in a real-world situation.
2017	Rai et al.[61]	To reduce the impact of hotspots, it presented and evaluated three VM selection strategies: median migration time, maximum utilization, and hotspot migration	Simulations indicated that the suggested VM selection strategies significantly reduced total energy use while lowering violations of the service level agreement.	It is critical to assess how well the proposed strategies perform on real-world cloud infrastructure.
2017	Maiyza et al.[62]	It presents two VM selection techniques, namely minimum migration time and minimum number of migrations, to decide which VM to remove from an overloaded server based on maximum CPU use.	The two recommended techniques decreased power usage without violating service level agreements.	In order to increase framework performance, the global manager must be changed to handle numerous requests at once.
2017	Sohrabi et al.[63]	Bayesian-based migration technique known as Energy-efficient Adaptive Migration (EAM)	Lower energy usage and 5.39 times faster overall execution time than other policies.	The proposed mechanism's evaluation does not take SLA violations into account.
2017	He et al.[64]	Improved under-loaded decision algorithms and minimum average utilization difference algorithms are developed as VM consolidation frameworks for improving energy efficiency.	The proposed framework reduced SLAV by 58.7%, the number of VM migrations by 20.6%, the overall energy efficiency by 59.8%, and the energy consumption by 1.15%.	The suggested methodology needs to be evaluated in various cloud computing environments.
2017	Yadav et al.[65]	VM selection criteria based on maximum usage and minimum size are proposed.	The suggested VM selection strategy decreased energy consumption by 13%, average SLA violations by 10%, and host reactivations by 13% as compared to traditional VM selection policies like linear regression, median absolute deviation, and interquartile range.	To generalize the findings, it is necessary to test the suggested VM selection policy in a variety of cloud scenarios.
2017	Sharma et al.[66]	It presented three potential heuristics for selecting the host machine (VM placement), known as best fit heuristics 1, heuristic 2, and heuristic 3.	The suggested heuristics 1, 2, and 3 outperformed the power-aware best fit decreasing and improving the energy-SLAV-migration count metric by 84%, 88%, and 92%, respectively.	Only SLA violations, migration times, and other performance elements decreased but energy usage did not decrease.
2018	Chou et al.[67]	It suggested the use of a dynamic power-saving resource allocation (DPRA) system. Least squares regression is used to forecast how frequently VMs and PMs will be employed. The particle swarm optimization (PSO) method is used for the deployment of VMs to PMs.	DPRA mechanism is compared with the best-fit algorithms with the least amount of CPU, minimum memory, and random VM selection. In comparison to BFrand, BFminU, BFminM, and PSO, the DPRA's energy usage is reduced by 17.85%, 15.03%, 18.04%, and 11.36%, respectively.	SLA violations are not regarded as evaluation parameters. PSO needs to perform better during the DPRA procedure.

TABLE 2. ***(Continued.)*** **VM consolidation literature review and their findings.**

Year	Author	Algorithm/Method	Findings	Limitations & Future Scope
2018	Malekloo et al.[40]	Ant Colony Optimization	Increased energy efficiency, reduced CPU resource waste and communication expenses, and reduced SLA violations and migrations.	The CPU was the only resource considered in this approach. There was no consideration given to resources like memory, storage, and network bandwidth.
2018	Venkata et al.[68]	Lion and Whale optimization	The suggested method completed VM migration with the maximum utilization of resources, the lowest migration cost, and the least amount of energy usage as compared to particle swarm optimization and the genetic method.	The suggested method is evaluated without taking into account SLA violations.
2018	Yan et al.[69]	Power-aware Ant Colony Optimization	When there are fewer idle hosts available for migration than VMs, power-aware ACO performed 27% better than power-aware best fit decreasing and 20% better than virtual machine placement ACO (VMP-ACO). It used 91% less energy than VMP-ACO when there were more hosts to move than VMs.	It is necessary to take into account more evaluation criteria and resources. To assess the proposed technique, an actual cloud environment must be employed.
2018	Yousefipour et al.[7]	An energy and cost-aware meta-heuristic approach for virtual machine consolidation	Compared to the first fit (FF), decreasing FF, and permutation pack algorithms, the suggested approach used less energy and was less expensive.	Since problem sizes evolve, consideration must be given to where to position dynamic virtual machines.
2018	Shi et al.[70]	Particle swarm optimization	When dealing with the demands of large-scale applications, energy is saved.	SLA violations are not considered while assessing the proposed technique. Migrations of containers must be considered to achieve the dynamic optimization of container consolidation.
2018	Wang et al.[36]	It suggests the use of a dynamic virtual machine consolidation paradigm known as DVMC to support green cloud computing.	Energy savings and reduced SLA violations	It is necessary to test the suggested approach in various cloud computing environments and take into account additional resources.
2018	Aryania et al.[71]	Ant colony based energy conscious VM consolidation algorithm	Reduction in the number of migrations by 89%, overall energy usage by 25%, SLA violations by 79%, and there were 16% more sleeping PMs.	More simulations must be run to assess how the suggested algorithm affects the actual workload.
2018	Pham et al.[43]	The resource allocation problem of virtual services is addressed using the energy consumption resource allocation simulated annealing algorithm.	Less energy is used than in the First Fit Decreasing algorithm.	The proposed algorithm's evaluation does not consider SLA violations.
2018	Barlaskar et al.[72]	Enhanced Cuckoo Search (ECS) algorithm	In comparison to an ant colony, an improved firefly search, and a genetic algorithm, the suggested method used 26%, 27%, and 25% less energy, respectively.	A realistic cloud environment must be used to test the suggested method, and more performance evaluation parameters must be taken into account.
2019	Moghaddam et al.[73]	To prevent migrations and SLA violations, it suggested a policy for choosing energy-efficient VMs. The suggested method considers the linear correlation between VMs as well as the CPU usage of VMs on each host.	Reduced SLA violations by 66%, the number of over-utilized hosts by 81%, and ESV by 64%.	It only took into account the linear correlation of the VMs. It is also required to evaluate the effects of non-linear correlation on the suggested VM selection policy.
2019	Bouchareb et al.[74]	It proposed an approach based on a minimum or maximum threshold to avoid the emergence of under-loaded or overloaded hosts. For VM allocation and reallocation, it specifies various rules.	Less energy consumption and fewer migrations.	The proposed approach's evaluation does not consider SLA violations. A real cloud environment and additional resources must be used to test the proposed approach.
2019	Mosa et al.[75]	Genetic algorithm	The genetic algorithm outperformed the best-fit decreasing method in terms of dynamic VM placement. The over- and under-utilization of CPU and memory are decreased by a genetic algorithm,	There should be more assessments based on actual workload traces, and resources like network bandwidth or storage should also be taken into consideration.

TABLE 2. ***(Continued.)*** **VM consolidation literature review and their findings.**

Year	Author	Algorithm/Method	Findings	Limitations & Future Scope
			which also lowers service level agreement violations.	
2019	Rajagopal et al.[76]	It proposes an approach for selecting VMs based on a fuzzy soft set. RAM consumption, CPU usage, memory use, and correlation values amongst VMs on the overloaded host are the four objectives used for VM selection criteria.	Better virtual machine selection, fewer VM migrations, lower SLA violation rates, and increased energy efficiency.	To lower the operational costs of cloud data centers, the proposed work must be implemented on a realistic cloud platform.
2019	Qiao et al.[77]	The Genetic Expression Programming (GEP) algorithm is used to analyze historical data, and symbolic regression is used to evaluate the role of migration.	The experiment's findings demonstrated that the suggested technique functioned well in terms of SLA breaches and VM migrations.	A real cloud platform, like Open Stack, should be used to test the effectiveness of the suggested technique.
2019	Basu et al.[78]	An improved genetic algorithm	The required number of VMs was allotted, ensuring that the average fitness, load distribution, and resource utilization were maintained over a large number of epochs.	The system can be improved by parallelization and using different parameters as the objective of a genetic algorithm.
2019	Al-Moalmi et al.[79]	Grey wolf optimization	Decreased energy consumption and increased use of CPU and memory.	While assessing the suggested technique, SLA violations are not taken into consideration. It is also necessary to take into account additional resources for evaluation..
2019	Mustafa et al.[80]	It proposed two energy-efficient approaches, i.e., minimum power best-fit decreasing and maximum capacity best-fit decreasing	The suggested methods performed better than heuristic-based procedures and variations in terms of energy usage, service level agreement violations, and the number of migrations.	The proposed work can be improved still further by including additional research challenges like network load, load balancing, fault tolerance, and profit maximization.
2019	Yavari et al.[81]	VM consolidation based on heuristics and Firefly methods	Reduction in energy use, SLA violations, and the number of migrations	A real-time evaluation of the proposed method is required in cloud computing.
2020	Shao et al.[82]	Improved discrete particle swarm algorithm based on the gray model	SLAV and energy use were reduced by 34.53% and 97.53%, respectively, to improve service quality.	The consolidation algorithm must consider host backup and security and deployable in real data centers.
2020	Yadav et al.[83]	A Least medial square regression adaptive heuristic approach is proposed for detecting overloaded hosts.	Reduction in energy usage	Real cloud infrastructure, like OpenStack, should be used to test the proposed method.
2020	Mandal et al.[84]	To aid in choosing which VMs to migrate, it proposed a power-based VM selection policy. The basic concept is to relocate the most resource-intensive VMs to a host that is not as busy.	Minimum service level agreement violations with energy-efficient VM selection	A real cloud infrastructure like OpenStack must be used to test the proposed algorithm. Additionally, more resources ought to be taken into account when assessing the suggested strategy.
2020	Karthikeyan et al.[85]	Hybrid of Artificial bee colony-and bat algorithm with naive Bayes classifier	The proposed model consumed the least energy and experienced the fewest failures when compared to the artificial bee colony, bat algorithm, dynamic adaptive particle swarm optimization, and particle swarm optimization.	New classifiers and optimization using various performance metrics should be considered during the migration of VMs. It is necessary to take the SLA violation rate into account while assessing the suggested strategy.
2020	Pang et al.[86]	User experience-oriented Ant colony optimization algorithm	Energy savings of up to 20%, 24%, and 30% were obtained when compared to the ant colony, min-min strategy, and round-robin algorithm, respectively.	Memory and network bandwidth requirements must be taken into account for the energy model to more accurately predict energy consumption. The proposed algorithm's evaluation does not consider SLA violations.
2020	Naik et al.[87]	Fruit Fly Hybridized Cuckoo Search	The suggested approach used 72% fewer resources and 68 Kwh less energy than the ant colony, genetic algorithm, and particle swarm optimization.	Cloud data centers must consider communication costs and SLA violations in order to increase resource utilization and decrease energy usage.
2020	Ibrahim et al.[88]	Power-aware particle swarm optimization approach	Reductions of 8.01%, 39.65%, 66.33%, and 11.87%, respectively, in energy usage, migration frequency, host shutdowns, and energy SLA violations.	Real-world cloud computing environments must be used to test the proposed strategy. Resources like bandwidth, memory, and the network

TABLE 2. *(Continued.)* VM consolidation literature review and their findings.

Year	Author	Algorithm/Method	Findings	Limitations & Future Scope
				must be considered to maximize optimization.
2020	Hsieh et al.[89]	The proposed VM consolidation method considers both present and future resource use by using host overload and under-load detection. For predicting future resource demand, the gray-Markov model is employed.	The proposed strategy led to low energy usage and fewer VM migrations while retaining service quality as compared to local regression, median absolute deviation, static threshold, and interquartile range methods.	It is necessary to take into account additional resources for evaluation.
2021	Hariharan et al.[90]	Particle swarm optimization combined with beetle swarm optimization is known as adaptive beetle swarm optimization.	In contrast to the particle swarm, beetle swarm optimization, and evolutionary algorithms, the suggested method effectively controlled the resources of servers by minimizing the number of active machines, which in turn reduced the overall energy consumption of the cloud system.	Realistic cloud infrastructure, such as OpenStack, should be used to test the suggested method. It is necessary to take into account migration-related security issues. The proposed approach's evaluation does not consider SLA violations.
2021	Talwani et al.[91]	It presents a novel strategy for reducing the overall number of migrations by precisely selecting the VMs to be transferred using swarm intelligence supported by machine learning.	Fewer SLA violations and less energy use	An actual cloud environment must be used to test the suggested strategy, and additional resources must be taken into account.
2021	Khan et al.[92]	Normalization-based VM consolidation	The suggested method performed better than other energy-aware solutions in terms of energy usage, number of VM migrations, and SLA breaches by 1.61, 10.33, and 6.82 times, respectively.	The proposed method uses threshold values for the identification of over-utilized hosts. The method must be developed in a way that enables a dynamic estimate of the threshold value using regression analysis.
2021	Tarahomi et al.[93]	Micro-genetic algorithm	Simulation results demonstrated that the micro-genetic approach outperformed other methods, such as the power-aware algorithm for allocation with bin-packing and the genetic VM allocation algorithm, in terms of power utilization.	The approach has to be enhanced, and its efficiency must be evaluated using a realistic cloud system like OpenStack.
2021	Barthwal et al.[94]	Ant Colony Optimization	The suggested method utilized less energy while maintaining SLA compliance as compared to the power-aware best-fit method.	Performance can be improved further by using additional resources such as memory, storage, and bandwidth.
2021	Shalu et al.[95]	Enhanced-Modified Best Fit Decreasing(E-MBFD) approach is proposed for VM allocation. An artificial neural network(ANN) is used to identify incorrect allocations caused by inefficient resource consumption.	In terms of SLA violations and power usage, the proposed approach based on E-MBFD and ANN outperformed standard MBFD methodologies.	To generalize the findings, a more comparative analysis is required. Additional resources for evaluation should be considered.
2021	Ibrahim et al.[96]	To dynamically relocate and deploy VMs on PMs, it recommends using an efficient adaptive migration algorithm (EAMA).	The suggested approach performed better in terms of resource utilization, shut-down hosts, migrations, and SLA violations than the predictive anti-correlated placement algorithm and resource usage energy efficiency.	To optimize resource usage and decrease SLA violations, a load-balanced, resource-aware virtual machine migration approach is required.
2021	Samriya et al.[97]	Penguin Optimization Algorithm	In comparison to particle swarm optimization, binary gravity search, and ant colony, the suggested method minimized SLA violations by 47.12% while conserving energy by an average of 55.96%.	The reduction of resource waste associated with VM deployment needs to be prioritized. Further improvements in energy efficiency can be made by extending the suggested strategy with new techniques.
2021	Gharehpasha et al.[98]	Whale Optimization Algorithm	The suggested technique improves resource utilization while using less energy in the cloud data center.	To fully utilize virtualization's advantages, security and privacy issues must be addressed. SLAs, cloud task scheduling, and utilizing several cloud platforms are things to be taken into consideration.

TABLE 2. *(Continued.)* VM consolidation literature review and their findings.

Year	Author	Algorithm/Method	Findings	Limitations & Future Scope
2021	Madhumala et al.[99]	Particle swarm optimization and a modified first-fit decreasing algorithm were used to provide a method for determining the optimal configuration for VMs to be occupied within a group of physically active servers.	The suggested approach uses less energy while efficiently allocating resources.	More heuristic strategies must be incorporated into cloud centers to improve their performance. While assessing the suggested algorithm, SLA violations are not taken into consideration.
2022	Dubey et al.[100]	VM Allocation algorithm based on the water drop method	Improved resource management, lower energy consumption, higher efficiency, and improved overall performance of VM allocation	The quality of service provided by the proposed method must be evaluated for different SLAs and VM instances using meta-heuristic-based optimization techniques.
2022	Vadivel et al.[101]	It proposed a method for efficiently allocating and utilizing resources based on a single property. Using a modified version of the principal component analysis and the relief technique, attributes are chosen. According to the chosen attribute, resources are distributed using the particle swarm algorithm, which is based on Cauchy theory.	The following improvements were seen with the suggested method: less migration, resource allocation with few iterations, minimal power usage, and shorter wait times for user services.	It is necessary to take into account more evaluation parameters. The suggested approach could be expanded to include virtual machine deployment and selection.
2022	Mejahed et al.[102]	Hybrid approach based on particle swarm optimization and flower pollination optimization with Levy flight.	In the simulation experiments, the suggested approach outperformed the best-fit bin-packing strategy.	The placement of VMs must also consider other goals, including load balancing, live migration, and cost minimization. SLA violations are not taken into account while evaluating the suggested approach.
2022	Xing et al.[103]	Ant colony optimization	The proposed approach utilized power and bandwidth resources more efficiently than pre-existing heuristics and metaheuristics algorithms.	It only considers a few test case scenarios, despite the fact that a data center has hundreds of thousands of PM instances.
2022	Karmakar et al.[104]	Ant Colony Optimization	The suggested approach performed better in terms of reducing communication costs and conserving energy. Also, the variation in the outcomes of several simulations is examined, and it is found to be stable.	The evaluation of the suggested algorithm must consider SLA violations and communication costs.
2022	Çağlar et al.[105]	It suggested a method for managing resources in virtualized data centers by increasing the number of servers without relocating them to fulfill the dynamic workload demands. The proposed method is called look-ahead resource allocation.	The results of the experiments showed that the proposed method outperformed Local Regression-Minimum Migration Time in terms of throughput and energy efficiency.	It only emphasizes workloads that are CPU-intensive. To handle workloads that don't only rely on the CPU, the suggested approach needs to be improved.
2022	Nikzad et al.[106]	Multi-objective ant colony optimization	The suggested method decreased VM migrations by 12.5%, SLA violations by 5.3%, and energy consumption by 10.3%.	It has to be tested using open-source tools like OpenStack in a realistic cloud environment.
2022	Gomathi et al.[107]	Composite Mutation Particle Swarm Optimization Based on Pareto	The proposed method outperformed heuristics and genetic algorithms in all instances of both heavily and lightly used data centers in terms of SLA violations and energy usage.	A cloud computing environment that operates in real-time must be used to assess the proposed method.
2022	Venkata Subramanian et al.[14]	Dynamic multi-objective network-aware Cuckoo Search Method	The recommended model took a total migration time of 167 seconds and consumed less energy on average (127 kJ) with a makespan length of 199 seconds. It had a lower risk rating than the genetic algorithm and particle swarm optimization.	To improve the live VM migration process, various SLA measures and compression techniques must be used.

TABLE 2. ***(Continued.)*** **VM consolidation literature review and their findings.**

Year	Author	Algorithm/Method	Findings	Limitations & Future Scope
2022	Shaw et al.[108]	Reinforcement learning-based consolidation approach	The suggested method reduced service violations by 63% and enhanced energy efficiency by 25% in comparison to the power-aware heuristic algorithm.	It must be assessed using open-source cloud infrastructure. For evaluation, more resources must be considered.
2023	Radi et al.[109]	VM consolidation using a modified genetic approach	The results of the analysis showed that the proposed approach performed better than other current approaches in terms of energy utilization, SLA violations, and the overall number of VM migrations.	It only emphasizes workloads that are CPU-intensive. A realistic cloud environment must be used to test the suggested approach.
2023	Hema et al.[110]	The proposed VM consolidation algorithm is comprised of four algorithms: overloaded host identification, under-loading host detection, VM selection, and VM placement.	In comparison to other benchmark algorithms, energy consumption was lowered by 22%, while SLA performance improved by up to 80%.	The proposed technique should be tested in a real-world environment to clearly understand how well it functions.
2023	Prabha et al.[111]	Model-based on Reinforcement Learning	When compared to previous reinforcement learning models, the suggested model achieved great efficiency while utilizing cloud resources with less energy.	SLA violations are not taken into account while evaluating the suggested approach. Validation of the proposed approach in a real cloud infrastructure using containers and the actual workloads required.
2023	Kaur et al.[29]	Adaptive elastic scheduling algorithm to create a VM migration and allocation technique that is more energy-efficient.	With the suggested algorithm, power efficiency and violation ratios are improved. It uses bandwidth utilization and cosine similarity as additional tools to boost quality of service(QoS) performance.	A realistic cloud environment must be used to test the suggested approach, and more performance evaluation parameters must be taken into account.
2023	Wu et al.[112]	Enhanced Median Absolute Deviation method based on VMs state history	In comparison to the median absolute deviation, the proposed algorithm demonstrated improvement in terms of energy consumption and SLA violations by an average of 7.33%.	It is based on historical VM status data. By incorporating real-time workloads, the suggested approach must be expanded.
2023	Kanagaraj et al.[113]	Host overload and underload detection using the elephant herding optimization technique	The proposed approach reduced the number of VM migrations by 0.073%, energy consumption by about 11%, and SLA breaches by 6.15%.	There should be more assessments based on actual workload traces, and resources like network bandwidth or storage should also be taken into consideration.

VII. CHALLENGES FOUND IN THE LITERATURE

Although VM consolidation has been extensively researched in the literature, there are still some open challenges that need to be addressed.

Some of the open issues found in the literature related to VM consolidation include:

- Although cloud computing involves varied dynamic environments, many publications assume only homogeneous cloud scenarios for simulation.
- Many studies just consider CPU utilization when calculating energy usage; however, other components like memory, communication networks, etc. also consume energy. It is crucial to assess how much energy each of these components uses.
- Most publications emphasized energy savings without taking into account SLA violations.
- When multiple VMs are consolidated onto a single physical server, there is a risk of performance degradation due to resource contention, interference, or bottlenecks. It is challenging to maintain the same level of performance for consolidated VMs due to the dynamic workloads of cloud computing.
- All consolidated VMs use the same PM resources, so there may be an increase in network bandwidth usage.
- Conflicts and failures may result from some of the VMs incompatible software.
- VM consolidation can increase the risk of security and privacy breaches as multiple VMs are consolidated onto a single physical server. An attacker who gains access to one VM can potentially access other VMs on the same physical server. There is a need for more research to develop secure VM consolidation techniques that minimize the risk of security and privacy breaches.
- It is essential to come up with an efficient and effective workload prediction system so that service providers can foresee forthcoming workloads and employ data center resources efficiently.
- Several researchers in the literature didn't evaluate the suggested algorithms in a real cloud environment. It's crucial to assess and compare the performance of the developed algorithms to the most effective baseline strategy in a realistic cloud environment.

TABLE 3. Resources and metrics used to assess the effectiveness of consolidation strategies.

Reference	Resources analyzed				Evaluation metrics						
	CPU	RAM	Bandwidth	Disk	Energy Consumption	SLA Violation	Number of VM Migrations	Number of active PMs	Number of Hosts Shutdown	PDM	Migration cost
[46]	✓	✓	×	✓	✓	×	×	×	×	×	×
[47]	✓	✓	×	×	✓	✓	✓	×	×	×	×
[48]	✓	×	×	×	✓	✓	×	×	×	×	×
[49]	✓	✓	×	×	✓	×	✓	×	✓	×	×
[50]	✓	✓	✓	×	✓	×	✓	×	×	×	✓
[51]	✓	✓	✓	×	✓	✓	✓	×	×	×	×
[52]	✓	✓	✓	×	✓	✓	✓	×	×	✓	×
[53]	✓	✓	×	×	✓	×	×	✓	×	×	×
[67]	✓	✓	✓	×	✓	×	×	×	✓	×	×
[54]	✓	×	×	×	✓	✓	✓	×	×	×	×
[30]	✓	×	×	×	✓	×	✓	×	×	×	×
[55]	✓	✓	×	×	✓	×	✓	×	×	×	×
[56]	✓	✓	✓	✓	✓	✓	✓	×	×	×	×
[57]	✓	×	×	✓	✓	✓	✓	×	×	×	×
[58]	✓	×	×	×	✓	✓	✓	×	×	×	×
[59]	✓	×	×	×	✓	✓	×	×	×	✓	×
[60]	✓	✓	✓	×	✓	×	✓	×	×	×	×
[27]	✓	✓	×	×	✓	×	✓	✓	×	×	×
[61]	✓	×	×	×	✓	✓	×	×	×	×	×
[62]	✓	✓	✓	×	✓	✓	✓	×	×	×	×
[63]	✓	✓	✓	×	✓	×	✓	×	×	×	×
[114]	✓	✓	✓	✓	✓	✓	✓	✓	×	×	✓
[64]	✓	×	×	×	✓	✓	✓	×	×	×	×
[65]	✓	✓	×	×	✓	✓	✓	×	✓	×	×
[66]	✓	✓	×	×	✓	✓	✓	×	×	✓	×
[40]	✓	×	×	×	✓	✓	✓	✓	×	×	✓
[68]	✓	✓	✓	×	✓	×	×	×	×	×	✓
[69]	✓	×	×	×	✓	×	×	×	×	×	×
[115]	✓	✓	✓	×	✓	×	×	×	×	×	×
[7]	✓	✓	×	✓	✓	×	×	×	×	×	✓
[70]	✓	✓	×	×	✓	×	✓	✓	×	×	×
[36]	✓	×	×	×	✓	✓	✓	×	×	×	×
[71]	✓	✓	×	×	✓	✓	✓	×	×	×	×
[73]	✓	✓	×	×	✓	✓	✓	×	×	×	×
[43]	✓	✓	✓	✓	✓	×	×	×	×	×	×
[72]	✓	✓	✓	×	✓	✓	✓	×	×	×	×
[74]	✓	×	×	×	✓	×	✓	×	×	×	×
[75]	✓	✓	×	×	✓	✓	×	×	×	×	×
[76]	✓	✓	×	×	✓	✓	✓	×	×	✓	×
[77]	✓	✓	✓	×	✓	✓	✓	×	×	×	×
[78]	✓	✓	×	×	✓	×	×	×	×	×	×
[79]	✓	✓	×	×	✓	×	×	×	×	×	×
[80]	✓	×	×	×	✓	✓	×	×	✓	✓	×
[81]	✓	✓	×	×	✓	✓	✓	×	×	×	×
[82]	✓	✓	✓	×	✓	✓	✓	×	×	✓	×
[83]	✓	✓	×	×	✓	✓	✓	×	✓	✓	×
[84]	✓	×	×	×	✓	✓	✓	×	✓	✓	×
[100]	✓	✓	✓	×	✓	×	✓	×	×	×	×
[85]	✓	×	×	×	✓	×	×	×	×	×	×
[116]	✓	×	×	×	✓	✓	×	×	×	✓	×
[86]	✓	×	×	×	✓	×	×	×	×	×	×
[87]	✓	✓	×	×	✓	×	✓	×	×	×	×
[88]	✓	×	×	×	✓	✓	✓	×	✓	×	×
[89]	✓	×	×	×	✓	✓	✓	×	×	×	×
[90]	✓	✓	×	×	✓	×	×	×	×	×	✓
[91]	✓	×	×	×	✓	✓	✓	×	×	×	×
[92]	✓	✓	✓	×	✓	✓	✓	×	×	×	×
[93]	✓	✓	×	×	✓	✓	✓	×	✓	×	×
[94]	✓	×	×	×	✓	✓	×	×	×	✓	×
[95]	✓	×	×	×	✓	✓	×	×	×	×	×
[101]	✓	✓	×	✓	✓	×	✓	×	×	×	×

TABLE 3. ***(Continued.)*** **Resources and metrics used to assess the effectiveness of consolidation strategies.**

Reference	Resources analyzed				Evaluation metrics						
	CPU	RAM	Bandwidth	Disk	Energy Consumption	SLA Violation	Number of VM Migrations	Number of active PMs	Number of Hosts Shutdown	PDM	Migration cost
[96]	✓	×	×	×	✓	✓	✓	×	✓	×	×
[97]	✓	✓	×	×	✓	✓	×	✓	×	×	×
[98]	✓	✓	✓	×	✓	×	✓	✓	✓	×	×
[99]	✓	✓	×	×	✓	×	×	×	×	×	×
[102]	✓	✓	×	×	✓	×	×	×	×	×	×
[103]	✓	×	✓	×	✓	×	×	×	×	×	✓
[104]	✓	✓	×	×	✓	×	×	×	×	×	✓
[105]	✓	×	×	×	✓	×	✓	×	✓	×	×
[36]	✓	×	×	×	✓	✓	✓	×	×	✓	×
[106]	✓	✓	×	✓	✓	✓	✓	✓	×	×	✓
[107]	✓	×	×	×	✓	✓	✓	×	✓	×	×
[14]	✓	✓	×	✓	✓	×	×	×	×	×	×
[108]	✓	×	×	×	✓	✓	✓	×	×	✓	×
[109]	✓	×	×	×	✓	✓	✓	×	×	✓	×
[110]	✓	✓	✓	×	✓	✓	✓	×	×	✓	×
[111]	✓	✓	×	×	✓	×	×	✓	×	×	×
[29]	✓	✓	✓	×	✓	✓	✓	×	×	×	×
[112]	✓	×	×	×	✓	✓	✓	×	×	×	×
[113]	✓	✓	×	×	✓	✓	✓	×	×	×	×

VIII. CONCLUSION AND FUTURE DIRECTIONS

This paper presented the findings of a thorough examination of studies on consolidation approaches for the years 2015 through 2023. The fundamental concepts underlying VM consolidation have also been explored as a means of developing green data centers by reducing their power and energy usage. 119 appropriate research articles were selected for the review based on inclusion and exclusion criteria as well as quality criteria. Consolidating virtual machines for cloud computing can be difficult since it's difficult to achieve the right balance between energy usage, resource usage, and service quality demands. The challenge arises due to the dynamic nature of cloud workloads and the varying resource demands of different applications. While VM consolidation tries to reduce energy consumption, the energy usage of cloud data centers cannot be reduced without compromising service quality. To address these challenges, researchers need to explore more advanced algorithms to make an energy-efficient cloud system without violating service-level agreements with the cloud user. Future research is aimed at developing a multi-objective system that emphasizes minimizing cloud energy usage without sacrificing service quality, preventing service level agreements from being compromised. Moreover, the cloud system needs to be empirically tested on a variety of real-world cloud platforms, e.g. open stack, Amazon EC2, etc. In addition, more research is needed on the open issues that were brought up in this review, like the need for a common framework for VM consolidation as well as methods to lessen the burden of VM migration.

REFERENCES

[1] R. M. B. Abadi, A. M. Rahmani, and S. H. Alizadeh, "Server consolidation techniques in virtualized data centers of cloud environments: A systematic literature review," *Softw., Pract. Exp.*, vol. 48, no. 9, pp. 1688–1726, Sep. 2018, doi: 10.1002/spe.2582.

[2] Z. Xiaoqing, "Efficient and balanced virtualized resource allocation based on genetic algorithm in cloud," in *Proc. 10th Int. Symp. Comput. Intell. Design (ISCID)*, Hangzhou, China, 2017, pp. 374–377, doi: 10.1109/ISCID.2017.187.

[3] C. Thiam and F. Thiam, "An energy-efficient VM migrations optimization in cloud data centers," in *Proc. IEEE AFRICON*, Sep. 2019, pp. 1–5, doi: 10.1109/AFRICON46755.2019.9133776.

[4] C. M. Mohammed and S. R. Zebaree, "Sufficient comparison among cloud computing services: IaaS, PaaS, and SaaS: A review," *Int. J. Sci. Bus.*, vol. 5, no. 2, pp. 17–30, 2021, doi: 10.5281/zenodo.4450129.

[5] S. Azizi, M. Zandsalimi, and D. Li, "An energy-efficient algorithm for virtual machine placement optimization in cloud data centers," *Cluster Comput.*, vol. 23, no. 4, pp. 3421–3434, Dec. 2020, doi: 10.1007/s10586-020-03096-0.

[6] K. A. Sultanpure and D. L. S. S. Reddy, "Virtual machine migration in cloud computing using artificial intelligence," *Int. J. Recent Technol. Eng.*, vol. 8, no. 4, pp. 2079–2088, Nov. 2019, doi: 10.35940/ijrte.D7657.118419.

[7] A. Yousefipour, A. M. Rahmani, and M. Jahanshahi, "Energy and cost-aware virtual machine consolidation in cloud computing," *Softw., Pract. Exp.*, vol. 48, no. 10, pp. 1758–1774, Oct. 2018, doi: 10.1002/spe.2585.

[8] N. K. Walia, N. Kaur, M. Alowaidi, K. S. Bhatia, S. Mishra, N. K. Sharma, S. K. Sharma, and H. Kaur, "An energy-efficient hybrid scheduling algorithm for task scheduling in the cloud computing environments," *IEEE Access*, vol. 9, pp. 117325–117337, 2021, doi: 10.1109/ACCESS.2021.3105727.

[9] M. Zakarya, "Energy, performance and cost efficient datacenters: A survey," *Renew. Sustain. Energy Rev.*, vol. 94, pp. 363–385, Oct. 2018, doi: 10.1016/j.rser.2018.06.005.

[10] N. Gholipour, E. Arianyan, and R. Buyya, "A novel energy-aware resource management technique using joint VM and container consolidation approach for green computing in cloud data centers," *Simul. Model. Pract. Theory*, vol. 104, Nov. 2020, Art. no. 102127, doi: 10.1016/j.simpat.2020.102127.

[11] S. Heidari and R. Buyya, "A cost-efficient auto-scaling algorithm for large-scale graph processing in cloud environments with heterogeneous resources," *IEEE Trans. Softw. Eng.*, vol. 47, no. 8, pp. 1729–1741, Aug. 2021, doi: 10.1109/TSE.2019.2934849.

[12] S. G. Domanal, R. M. R. Guddeti, and R. Buyya, "A hybrid bio-inspired algorithm for scheduling and resource management in cloud environment," *IEEE Trans. Services Comput.*, vol. 13, no. 1, pp. 3–15, Jan. 2020, doi: 10.1109/TSC.2017.2679738.

[13] M. C. S. Filho, C. C. Monteiro, P. R. M. Inácio, and M. M. Freire, "Approaches for optimizing virtual machine placement and migration in cloud environments: A survey," *J. Parallel Distrib. Comput.*, vol. 111, pp. 222–250, Jan. 2018, doi: 10.1016/j.jpdc.2017.08.010.

[14] N. V. Subramanian and V. S. S. Sriram, "An effective secured dynamic network-aware multi-objective cuckoo search optimization for live VM migration in sustainable data centers," *Sustainability*, vol. 14, no. 20, p. 13670, Oct. 2022, doi: 10.3390/su142013670.

[15] S. Namasudra, "Cloud computing: A new era," *J. Fundam. Appl. Sci.*, vol. 10, no. 2, pp. 113–134, 2018, doi: 10.4314/jfas.v10i2.9.

[16] A. Varasteh and M. Goudarzi, "Server consolidation techniques in virtualized data centers: A survey," *IEEE Syst. J.*, vol. 11, no. 2, pp. 772–783, Jun. 2017, doi: 10.1109/JSYST.2015.2458273.

[17] S. Mazumdar and M. Pranzo, "Power efficient server consolidation for cloud data center," *Future Gener. Comput. Syst.*, vol. 70, pp. 4–16, May 2017, doi: 10.1016/j.future.2016.12.022.

[18] Z. Á. Mann, "Allocation of virtual machines in cloud data centers—A survey of problem models and optimization algorithms," *ACM Comput. Surv.*, vol. 48, no. 1, pp. 1–34, Aug. 2015, doi: 10.1145/2797211.

[19] N. Bjørner, S. Prasad, and L. Parida, *Distributed Computing and Internet Technology: 12th International Conference, ICDCIT 2016 Bhubaneswar, India, January 15–18, 2016 Proceedings* (Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence Lecture Notes in Bioinformatics), vol. 9581. Cham, Switzerland: Springer, 2016, pp. 61–74, doi: 10.1007/978-3-319-28034-9.

[20] L. Alsbatin, G. Öz, and A. H. Ulusoy, "An overview of energy-efficient cloud data centres," in *Proc. Int. Conf. Comput. Appl. (ICCA)*, Sep. 2017, pp. 211–214, doi: 10.1109/COMAPP.2017.8079789.

[21] M. Shelar, S. Sane, V. Kharat, and R. Jadhav, "Autonomic and energy-aware resource allocation for efficient management of cloud data centre," in *Proc. Innov. Power Adv. Comput. Technol. (i-PACT)*, Apr. 2017, pp. 1–8, doi: 10.1109/IPACT.2017.8244944.

[22] M. Kamran and B. Nazir, "QoS-aware VM placement and migration for hybrid cloud infrastructure," *J. Supercomput.*, vol. 74, no. 9, pp. 4623–4646, Sep. 2018, doi: 10.1007/s11227-017-2071-1.

[23] Y. Huang, H. Xu, H. Gao, X. Ma, and W. Hussain, "SSUR: An approach to optimizing virtual machine allocation strategy based on user requirements for cloud data center," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 2, pp. 670–681, Jun. 2021, doi: 10.1109/TGCN.2021.3067374.

[24] S. E. Motaki, A. Yahyaouy, and H. Gualous, "A prediction-based model for virtual machine live migration monitoring in a cloud datacenter," *Computing*, vol. 103, no. 11, pp. 2711–2735, Nov. 2021, doi: 10.1007/s00607-021-00981-3.

[25] M. H. Alshayeji, S. Abed, and M. D. Samrajesh, "Energy efficient virtual machine migration algorithm," *J. Eng. Res.*, vol. 5, no. 2, pp. 19–42, 2017.

[26] A. Fatima, N. Javaid, T. Sultana, W. Hussain, M. Bilal, S. Shabbir, Y. Asim, M. Akbar, and M. Ilahi, "Virtual machine placement via bin packing in cloud data centers," *Electronics*, vol. 7, no. 12, p. 389, Dec. 2018, doi: 10.3390/electronics7120389.

[27] F. Abdessamia, Y. Tai, W. Z. Zhang, and M. Shafiq, "An improved particle swarm optimization for energy-efficiency virtual machine placement," in *Proc. Int. Conf. Cloud Comput. Res. Innov. (ICCCRI)*, Apr. 2017, pp. 7–13, doi: 10.1109/ICCCRI.2017.9.

[28] A. Kumar, C. Sathasivam, and P. Periyasamy, "Virtual machine placement in cloud computing," *Indian J. Sci. Technol.*, vol. 9, no. 29, Aug. 2016, doi: 10.17485/ijst/2016/v9i29/79768.

[29] A. Kaur, S. Kumar, D. Gupta, Y. Hamid, M. Hamdi, A. Ksibi, H. Elmannai, and S. Saini, "Algorithmic approach to virtual machine migration in cloud computing with updated SESA algorithm," *Sensors*, vol. 23, no. 13, p. 6117, Jul. 2023, doi: 10.3390/s23136117.

[30] A. Kaur and M. Kalra, "Energy optimized VM placement in cloud environment," in *Proc. 6th Int. Conf. Cloud Syst. Big Data Eng. (Confluence)*, Jan. 2016, pp. 141–145, doi: 10.1109/CONFLUENCE.2016.7508103.

[31] V. Garg and B. Jindal, "Energy efficient virtual machine migration approach with SLA conservation in cloud computing," *J. Central South Univ.*, vol. 28, no. 3, pp. 760–770, Mar. 2021, doi: 10.1007/s11771-021-4643-8.

[32] K. P. N. Jayasena, L. Li, M. A. Elaziz, S. Xiong, and J. Xiang, "Optimizing the energy efficient VM consolidation by a multi-objective algorithm," in *Proc. IEEE 22nd Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, May 2018, pp. 672–676, doi: 10.1109/CSCWD.2018.8465184.

[33] H. M. D. Kabir, A. Khosravi, S. K. Mondal, M. Rahman, S. Nahavandi, and R. Buyya, "Uncertainty-aware decisions in cloud computing," *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–30, May 2022, doi: 10.1145/3447583.

[34] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future Gener. Comput. Syst.*, vol. 28, no. 5, pp. 755–768, May 2012, doi: 10.1016/j.future.2011.04.017.

[35] N. Garg, D. Singh, and M. S. Goraya, "VM selection and allocation policy to optimize VM migration in cloud environment," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2, pp. 3444–3449, Jul. 2019, doi: 10.35940/ijrte.B2700.078219.

[36] H. Wang and H. Tianfield, "Energy-aware dynamic virtual machine consolidation for cloud datacenters," *IEEE Access*, vol. 6, pp. 15259–15273, 2018, doi: 10.1109/ACCESS.2018.2813541.

[37] T. Chaabouni and M. Khemakhem, "Energy management strategy in cloud computing: A perspective study," *J. Supercomput.*, vol. 74, no. 12, pp. 6569–6597, Dec. 2018, doi: 10.1007/s11227-017-2154-z.

[38] X. Wu, Y. Zeng, and G. Lin, "An energy efficient VM migration algorithm in data centers," in *Proc. 16th Int. Symp. Distrib. Comput. Appl. Bus., Eng. Sci. (DCABES)*. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Oct. 2017, pp. 27–30, doi: 10.1109/DCABES.2017.14.

[39] N. Patel and H. Patel, "An approach for detection of overloaded host to consolidate workload in cloud datacenter," *Int. J. Grid High Perform. Comput.*, vol. 10, no. 2, pp. 59–69, Apr. 2018, doi: 10.4018/IJGHPC.2018040105.

[40] M.-H. Malekloo, N. Kara, and M. El Barachi, "An energy efficient and SLA compliant approach for resource allocation and consolidation in cloud computing environments," *Sustain. Comput., Informat. Syst.*, vol. 17, pp. 9–24, Mar. 2018, doi: 10.1016/j.suscom.2018.02.001.

[41] M. H. Sayadnavard, A. Toroghi Haghighat, and A. M. Rahmani, "A multi-objective approach for energy-efficient and reliable dynamic VM consolidation in cloud data centers," *Eng. Sci. Technol., Int. J.*, vol. 26, Feb. 2022, Art. no. 100995, doi: 10.1016/j.jestch.2021.04.014.

[42] I. Pietri and R. Sakellariou, "Mapping virtual machines onto physical machines in cloud computing: A survey," *ACM Comput. Surv.*, vol. 49, no. 3, pp. 1–30, Oct. 2016, doi: 10.1145/2983575.

[43] N. M. N. Pham, V. S. Le, and H. H. C. Nguyen, "Energy-efficient resource allocation for virtual service in cloudcomputing environment," in *Information Systems Design and Intelligent Applications*, vol. 672. Singapore: Springer, 2018, doi: 10.1007/978-981-10-7512-4_13.

[44] D. A. Shafiq, N. Z. Jhanjhi, and A. Abdullah, "Load balancing techniques in cloud computing environment: A review," *J. King Saud Univ., Comput. Inf. Sci.*, vol. 34, no. 7, pp. 3910–3933, Jul. 2022, doi: 10.1016/j.jksuci.2021.02.007.

[45] R. Zolfaghari, A. Sahafi, A. M. Rahmani, and R. Rezaei, "Application of virtual machine consolidation in cloud computing systems," *Sustain. Comput., Informat. Syst.*, vol. 30, Jun. 2021, Art. no. 100524, doi: 10.1016/j.suscom.2021.100524.

[46] T. K. Okada, A. D. L. F. Vigliotti, D. M. Batista, and A. Goldman vel Lejbman, "Consolidation of VMs to improve energy efficiency in cloud computing environments," in *Proc. 33rd Brazilian Symp. Comput. Netw. Distrib. Syst.* Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, May 2015, pp. 150–158, doi: 10.1109/SBRC.2015.27.

[47] M. R. Chowdhury, M. R. Mahmud, and R. M. Rahman, "Implementation and performance analysis of various VM placement strategies in CloudSim," *J. Cloud Comput.*, vol. 4, no. 1, pp. 1–21, Dec. 2015, doi: 10.1186/s13677-015-0045-5.

[48] M. A. H. Monil and R. M. Rahman, "Implementation of modified overload detection technique with VM selection strategies based on heuristics and migration control," in *Proc. IEEE/ACIS 14th Int. Conf. Comput. Inf. Sci. (ICIS)*, Jun. 2015, pp. 223–227, doi: 10.1109/ICIS.2015.7166597.

[49] N. K. Gondhi and P. Kailu, "Prediction based energy efficient virtual machine consolidation in cloud computing," in *Proc. 2nd Int. Conf. Adv. Comput. Commun. Eng.* Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, May 2015, pp. 437–441, doi: 10.1109/ICACCE.2015.148.

[50] A. Tchana, N. De Palma, I. Safieddine, and D. Hagimont, "Software consolidation as an efficient energy and cost saving solution," *Future Gener. Comput. Syst.*, vol. 58, pp. 1–12, May 2016, doi: 10.1016/j.future.2015.11.027.

[51] G. E. I. Selim, M. A. El-Rashidy, and N. A. El-Fishawy, "An efficient resource utilization technique for consolidation of virtual machines in cloud computing environments," in *Proc. 33rd Nat. Radio Sci. Conf. (NRSC)*, Feb. 2016, pp. 316–324, doi: 10.1109/NRSC.2016.7450844.

[52] N. K. Chien, V. S. G. Dong, N. H. Son, and H. D. Loc, "An efficient virtual machine migration algorithm based onminimization of migration in cloud computing," in *Nature of Computation and Communication* (Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), vol. 168. Cham, Switzerland: Springer, 2016, pp. 62–71, doi: 10.1007/978-3-319-46909-6_7.

[53] X.-F. Liu, Z.-H. Zhan, J. D. Deng, Y. Li, T. Gu, and J. Zhang, ''An energy efficient ant colony system for virtual machine placement in cloud computing,'' *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 113–128, Feb. 2018, doi: 10.1109/TEVC.2016.2623803.
[54] D. Deng, K. He, and Y. Chen, ''Dynamic virtual machine consolidation for improving energy efficiency in cloud data centers,'' in *Proc. 4th Int. Conf. Cloud Comput. Intell. Syst. (CCIS)*, Aug. 2016, pp. 366–370, doi: 10.1109/CCIS.2016.7790285.
[55] N. J. Kansal and I. Chana, ''Energy-aware virtual machine migration for cloud computing—A firefly optimization approach,'' *J. Grid Comput.*, vol. 14, no. 2, pp. 327–345, Jun. 2016, doi: 10.1007/s10723-016-9364-0.
[56] F. Farahnakian, T. Pahikkala, P. Liljeberg, J. Plosila, N. T. Hieu, and H. Tenhunen, ''Energy-aware VM consolidation in cloud data centers using utilization prediction model,'' *IEEE Trans. Cloud Comput.*, vol. 7, no. 2, pp. 524–536, Apr. 2019, doi: 10.1109/TCC.2016.2617374.
[57] H. Li, G. Zhu, C. Cui, H. Tang, Y. Dou, and C. He, ''Energy-efficient migration and consolidation algorithm of virtual machines in data centers for cloud computing,'' *Computing*, vol. 98, no. 3, pp. 303–317, Mar. 2016, doi: 10.1007/s00607-015-0467-4.
[58] S. Y. Z. Fard, M. R. Ahmadi, and S. Adabi, ''A dynamic VM consolidation technique for QoS and energy consumption in cloud environment,'' *J. Supercomput.*, vol. 73, no. 10, pp. 4347–4368, Oct. 2017, doi: 10.1007/s11227-017-2016-8.
[59] H.-W. Li, Y.-S. Wu, Y.-Y. Chen, C.-M. Wang, and Y.-N. Huang, ''Application execution time prediction for effective CPU provisioning in virtualization environment,'' *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 11, pp. 3074–3088, Nov. 2017, doi: 10.1109/TPDS.2017.2707543.
[60] S. K. Abd, S. A. R. Al-Haddad, F. Hashim, A. B. H. J. Abdullah, and S. Yussof, ''An effective approach for managing power consumption in cloud computing infrastructure,'' *J. Comput. Sci.*, vol. 21, pp. 349–360, Jul. 2017, doi: 10.1016/j.jocs.2016.11.007.
[61] R. Rai, G. Sahoo, and S. Mehfuz, ''Effect of VM selection heuristics on energy consumption and SLAs during VM migrations in cloud data centers,'' in *Advances in Intelligent Systems and Computing*, vol. 509. Singapore: Springer, 2017, pp. 189–199, doi: 10.1007/978-981-10-2525-9_19.
[62] A. I. Maiyza, H. A. Hassan, W. M. Sheta, N. M. Sadek, and M. A. Mokhtar, ''End-user's SLA-aware consolidation in cloud data centers,'' in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2017, pp. 196–204, doi: 10.1109/ISSPIT.2017.8388641.
[63] S. Sohrabi, Y. Yang, I. Moser, and A. Aleti, ''Energy-efficient adaptive virtual machine migration mechanism for private clouds,'' *Concurrency Comput., Pract. Exp.*, vol. 29, no. 18, Sep. 2017, Art. no. e4170, doi: 10.1002/cpe.4170.
[64] K. He, Z. Li, D. Deng, and Y. Chen, ''Energy-efficient framework for virtual machine consolidation in cloud data centers,'' *China Commun.*, vol. 14, no. 10, pp. 192–201, Oct. 2017.
[65] R. Yadav, W. Zhang, H. Chen, and T. Guo, ''MuMs: Energy-aware VM selection scheme for cloud data center,'' in *Proc. 28th Int. Workshop Database Expert Syst. Appl. (DEXA)*. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Aug. 2017, pp. 132–136, doi: 10.1109/DEXA.2017.43.
[66] O. Sharma and H. Saini, ''SLA and performance efficient heuristics for virtual machines placement in cloud data centers,'' *Int. J. Grid High Perform. Comput.*, vol. 9, no. 3, pp. 17–33, Jul. 2017, doi: 10.4018/IJGHPC.2017070102.
[67] L.-D. Chou, H.-F. Chen, F.-H. Tseng, H.-C. Chao, and Y.-J. Chang, ''DPRA: Dynamic power-saving resource allocation for cloud data center using particle swarm optimization,'' *IEEE Syst. J.*, vol. 12, no. 2, pp. 1554–1565, Jun. 2018, doi: 10.1109/JSYST.2016.2596299.
[68] J. V. Krishna, G. A. Naidu, and N. Upadhayaya, ''A lion-whale optimization-based migration of virtual machines for data centers in cloud computing,'' *Int. J. Commun. Syst.*, vol. 31, no. 8, May 2018, Art. no. e3539, doi: 10.1002/dac.3539.
[69] W. Yan, J. Chen, and L. Li, ''A power-aware ACO algorithm for the cloud computing platform,'' in *Proc. ACM Int. Conf. Commun. Inf. Process.* New York, NY, USA: Association for Computing Machinery, Nov. 2018, pp. 1–6, doi: 10.1145/3290420.3290428.
[70] T. Shi, H. Ma, and G. Chen, ''Energy-aware container consolidation based on PSO in cloud data centers,'' in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2018, pp. 1–8, doi: 10.1109/CEC.2018.8477708.
[71] A. Aryania, H. S. Aghdasi, and L. M. Khanli, ''Energy-aware virtual machine consolidation algorithm based on ant colony system,'' *J. Grid Comput.*, vol. 16, no. 3, pp. 477–491, Sep. 2018, doi: 10.1007/s10723-018-9428-4.
[72] E. Barlaskar, Y. J. Singh, and B. Issac, ''Enhanced cuckoo search algorithm for virtual machine placement in cloud data centres,'' *Int. J. Grid Util. Comput.*, vol. 9, no. 1, pp. 1–17, 2018, doi: 10.1504/IJGUC.2018.090221.
[73] S. M. Moghaddam, S. F. Piraghaj, M. O'Sullivan, C. Walker, and C. Unsworth, ''Energy-efficient and SLA-aware virtual machine selection algorithm for dynamic resource allocation in cloud data centers,'' in *Proc. IEEE/ACM 11th Int. Conf. Utility Cloud Comput. (UCC)*. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Dec. 2018, pp. 103–113, doi: 10.1109/UCC.2018.00019.
[74] N. Bouchareb and N. E. Zarour, ''Virtual machines allocation and migration mechanism in green cloud computing,'' in *Modelling and Implementation of Complex Systems*, vol. 64. Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-05481-6_2.
[75] A. Mosa and R. Sakellariou, ''Dynamic virtual machine placement considering CPU and memory resource requirements,'' in *Proc. IEEE 12th Int. Conf. Cloud Comput. (CLOUD)*. Washington, DC, USA: IEEE Computer Society, Jul. 2019, pp. 196–198, doi: 10.1109/CLOUD.2019.00042.
[76] E. Rajagopal and N. Baskaran, ''Fuzzy softset based VM selection in cloud datacenter,'' in *Proc. Int. Conf. Intell. Comput. Control Syst. (ICCS)*, May 2019, pp. 462–467, doi: 10.1109/ICCS45141.2019.9065678.
[77] L. Qiao, B. Liu, Y. Hua, Q. Zhao, and X. Fu, ''Genetic expression programming based dynamic virtual machine consolidation in cloud computing,'' in *Proc. IEEE 9th Int. Conf. Electron. Inf. Emergency Commun. (ICEIEC)*, Jul. 2019, pp. 94–97, doi: 10.1109/ICEIEC.2019.8784588.
[78] S. Basu, G. Kannayaram, S. Ramasubbareddy, and C. Venkatasubbaiah, ''Improved genetic algorithm for monitoring of virtual machines in cloud environment,'' in *Smart Intelligent Computing and Applications*, vol. 105. Singapore: Springer, 2019, doi: 10.1007/978-981-13-1927-3_34.
[79] A. Al-Moalmi, J. Luo, A. Salah, and K. Li, ''Optimal virtual machine placement based on grey wolf optimization,'' *Electronics*, vol. 8, no. 3, p. 283, Mar. 2019, doi: 10.3390/electronics8030283.
[80] S. Mustafa, K. Sattar, J. Shuja, S. Sarwar, T. Maqsood, S. A. Madani, and S. Guizani, ''SLA-aware best fit decreasing techniques for workload consolidation in clouds,'' *IEEE Access*, vol. 7, pp. 135256–135267, 2019, doi: 10.1109/ACCESS.2019.2941145.
[81] M. Yavari, A. G. Rahbar, and M. H. Fathi, ''Temperature and energy-aware consolidation algorithms in cloud computing,'' *J. Cloud Comput.*, vol. 8, no. 1, pp. 1–16, Dec. 2019, doi: 10.1186/s13677-019-0136-9.
[82] Y. Shao, Q. Yang, Y. Gu, Y. Pan, Y. Zhou, and Z. Zhou, ''A dynamic virtual machine resource consolidation strategy based on a gray model and improved discrete particle swarm optimization,'' *IEEE Access*, vol. 8, pp. 228639–228654, 2020, doi: 10.1109/ACCESS.2020.3046318.
[83] R. Yadav, W. Zhang, K. Li, C. Liu, M. Shafiq, and N. K. Karn, ''An adaptive heuristic for managing energy consumption and overloaded hosts in a cloud data center,'' *Wireless Netw.*, vol. 26, no. 3, pp. 1905–1919, Apr. 2020, doi: 10.1007/s11276-018-1874-1.
[84] R. Mandal, M. K. Mondal, S. Banerjee, and U. 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,'' *J. Supercomput.*, vol. 76, no. 9, pp. 7374–7393, Sep. 2020, doi: 10.1007/s11227-020-03165-6.
[85] K. Karthikeyan, R. Sunder, K. Shankar, S. K. Lakshmanaprabu, V. Vijayakumar, M. Elhoseny, and G. Manogaran, ''RETRACTED ARTICLE: Energy consumption analysis of virtual machine migration in cloud using hybrid swarm optimization (ABC–BA),'' *J. Supercomput.*, vol. 76, no. 5, pp. 3374–3390, May 2020, doi: 10.1007/s11227-018-2583-3.
[86] S. Pang, K. Xu, S. Wang, M. Wang, and S. Wang, ''Energy-saving virtual machine placement method for user experience in cloud environment,'' *Math. Problems Eng.*, vol. 2020, pp. 1–9, May 2020, doi: 10.1155/2020/4784191.
[87] B. B. Naik, D. Singh, and A. B. Samaddar, ''FHCS: Hybridised optimisation for virtual machine migration and task scheduling in cloud data center,'' *IET Commun.*, vol. 14, no. 12, pp. 1942–1948, Jul. 2020, doi: 10.1049/iet-com.2019.1149.
[88] A. Ibrahim, M. Noshy, H. A. Ali, and M. Badawy, ''PAPSO: A power-aware VM placement technique based on particle swarm optimization,'' *IEEE Access*, vol. 8, pp. 81747–81764, 2020, doi: 10.1109/ACCESS.2020.2990828.
[89] S.-Y. Hsieh, C.-S. Liu, R. Buyya, and A. Y. Zomaya, ''Utilization-prediction-aware virtual machine consolidation approach for energy-efficient cloud data centers,'' *J. Parallel Distrib. Comput.*, vol. 139, pp. 99–109, May 2020, doi: 10.1016/j.jpdc.2019.12.014.

[90] B. Hariharan, R. Siva, S. Kaliraj, and P. N. S. Prakash, "ABSO: An energy-efficient multi-objective VM consolidation using adaptive beetle swarm optimization on cloud environment," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 3, pp. 2185–2197, Mar. 2023, doi: 10.1007/s12652-021-03429-w.

[91] S. Talwani and J. Singla, "Enhanced bee colony approach for reducing the energy consumption during VM migration in cloud computing environment," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 1022, no. 1, Jan. 2021, Art. no. 012069, doi: 10.1088/1757-899X/1022/1/012069.

[92] M. A. Khan, "An efficient energy-aware approach for dynamic VM consolidation on cloud platforms," *Cluster Comput.*, vol. 24, no. 4, pp. 3293–3310, Dec. 2021, doi: 10.1007/s10586-021-03341-0.

[93] M. Tarahomi, M. Izadi, and M. Ghobaei-Arani, "An efficient power-aware VM allocation mechanism in cloud data centers: A micro genetic-based approach," *Cluster Comput.*, vol. 24, no. 2, pp. 919–934, Jun. 2021, doi: 10.1007/s10586-020-03152-9.

[94] V. Barthwal and M. M. S. Rauthan, "AntPu: A meta-heuristic approach for energy-efficient and SLA aware management of virtual machines in cloud computing," *Memetic Comput.*, vol. 13, no. 1, pp. 91–110, Mar. 2021, doi: 10.1007/s12293-020-00320-7.

[95] S. Singh and D. Singh, "Artificial neural network-based virtual machine allocation in cloud computing," *J. Discrete Math. Sci. Cryptogr.*, vol. 24, no. 6, pp. 1739–1750, Aug. 2021, doi: 10.1080/09720529.2021.1878626.

[96] M. Ibrahim, M. Imran, F. Jamil, Y.-J. Lee, and D.-H. Kim, "EAMA: Efficient adaptive migration algorithm for cloud data centers (CDCs)," *Symmetry*, vol. 13, no. 4, p. 690, Apr. 2021, doi: 10.3390/sym13040690.

[97] J. K. Samriya, S. C. Patel, M. Khurana, P. K. Tiwari, and O. Cheikhrouhou, "Intelligent SLA-aware VM allocation and energy minimization approach with EPO algorithm for cloud computing environment," *Math. Problems Eng.*, vol. 2021, pp. 1–13, May 2021, doi: 10.1155/2021/9949995.

[98] S. Gharehpasha, M. Masdari, and A. Jafarian, "Virtual machine placement in cloud data centers using a hybrid multi-verse optimization algorithm," *Artif. Intell. Rev.*, vol. 54, no. 3, pp. 2221–2257, Mar. 2021, doi: 10.1007/s10462-020-09903-9.

[99] R. B. Madhumala, H. Tiwari, and V. C. Devaraj, "Virtual machine placement using energy efficient particle swarm optimization in cloud datacenter," *Cybern. Inf. Technol.*, vol. 21, no. 1, pp. 62–72, Mar. 2021, doi: 10.2478/cait-2021-0005.

[100] K. Dubey and S. C. Sharma, "An extended intelligent water drop approach for efficient VM allocation in secure cloud computing framework," *J. King Saud Univ., Comput. Inf. Sci.*, vol. 34, no. 7, pp. 3948–3958, Jul. 2022, doi: 10.1016/j.jksuci.2020.11.001.

[101] R. Vadivel and T. Sudalaimuthu, "Cauchy particle swarm optimization (CPSO) based migrations of tasks in a virtual machine," *Wireless Pers. Commun.*, vol. 127, no. 3, pp. 2229–2246, Dec. 2022, doi: 10.1007/s11277-021-08784-7.

[102] S. Mejahed and M. Elshrkawey, "A multi-objective algorithm for virtual machine placement in cloud environments using a hybrid of particle swarm optimization and flower pollination optimization," *PeerJ Comput. Sci.*, vol. 8, p. e834, Jan. 2022, doi: 10.7717/PEERJ-CS.834.

[103] H. Xing, J. Zhu, R. Qu, P. Dai, S. Luo, and M. A. Iqbal, "An ACO for energy-efficient and traffic-aware virtual machine placement in cloud computing," *Swarm Evol. Comput.*, vol. 68, Feb. 2022, Art. no. 101012, doi: 10.1016/j.swevo.2021.101012.

[104] K. Karmakar, R. K. Das, and S. Khatua, "An ACO-based multi-objective optimization for cooperating VM placement in cloud data center," *J. Supercomput.*, vol. 78, no. 3, pp. 3093–3121, Feb. 2022, doi: 10.1007/s11227-021-03978-z.

[105] İ. Çağlar and D. T. Altılar, "Look-ahead energy efficient VM allocation approach for data centers," *J. Cloud Comput.*, vol. 11, no. 1, pp. 1–16, Dec. 2022, doi: 10.1186/s13677-022-00281-x.

[106] B. Nikzad, B. Barzegar, and H. Motameni, "SLA-aware and energy-efficient virtual machine placement and consolidation in heterogeneous DVFS enabled cloud datacenter," *IEEE Access*, vol. 10, pp. 81787–81804, 2022, doi: 10.1109/ACCESS.2022.3196240.

[107] B. Gomathi, B. Saravana Balaji, V. Krishna Kumar, M. Abouhawwash, S. Aljahdali, M. Masud, and N. Kuchuk, "Multi-objective optimization of energy aware virtual machine placement in cloud data center," *Intell. Autom. Soft Comput.*, vol. 33, no. 3, pp. 1771–1785, 2022, doi: 10.32604/iasc.2022.024052.

[108] R. Shaw, E. Howley, and E. Barrett, "Applying reinforcement learning towards automating energy efficient virtual machine consolidation in cloud data centers," *Inf. Syst.*, vol. 107, Jul. 2022, Art. no. 101722, doi: 10.1016/j.is.2021.101722.

[109] M. Radi, A. A. Alwan, and Y. Gulzar, "Genetic-based virtual machines consolidation strategy with efficient energy consumption in cloud environment," *IEEE Access*, vol. 11, pp. 48022–48032, 2023, doi: 10.1109/ACCESS.2023.3276292.

[110] M. Hema and S. KanagaSubaRaja, "A quantitative approach to minimize energy consumption in cloud data centres using VM consolidation algorithm," *KSII Trans. Internet Inf. Syst.*, vol. 17, no. 2, pp. 312–334, Feb. 2023, doi: 10.3837/tiis.2023.02.002.

[111] B. Prabha, J. Thangakumar, and K. Ramesh, "Reinforcement learning based energy consolidation model for efficient cloud computing system," *Appl. Math. Inf. Sci.*, vol. 17, no. 1, pp. 67–77, 2023, doi: 10.18576/amis/170109.

[112] X. Wu and T. Zeng, "Energy-efficient virtual machine migration algorithm in data centers," *J. Phys., Conf. Ser.*, vol. 2504, no. 1, May 2023, Art. no. 012065, doi: 10.1088/1742-6596/2504/1/012065.

[113] G. Kanagaraj and G. Subashini, "Uniform distribution elephant herding optimization (UDEHO) based virtual machine consolidation for energy-efficient cloud data centres," *Automatika*, vol. 64, no. 3, pp. 530–540, 2023, doi: 10.1080/00051144.2023.2196116.

[114] H. Li, G. Zhu, Y. Zhao, Y. Dai, and W. Tian, "Energy-efficient and QoS-aware model based resource consolidation in cloud data centers," *Cluster Comput.*, vol. 20, no. 3, pp. 2793–2803, Sep. 2017, doi: 10.1007/s10586-017-0893-5.

[115] H. Cao, H. Sun, M. Sheng, Y. Shi, and J. Li, "A QoS-guaranteed energy-efficient VM dynamic migration strategy in cloud data centers," in *Proc. 10th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2018, pp. 1–6, doi: 10.1109/WCSP.2018.8555561.

[116] Y. Saadi and S. El Kafhali, "Energy-efficient strategy for virtual machine consolidation in cloud environment," *Soft Comput.*, vol. 24, no. 19, pp. 14845–14859, Oct. 2020, doi: 10.1007/s00500-020-04839-2.

[117] G. Portaluri, D. Adami, A. Gabbrielli, S. Giordano, and M. Pagano, "Power consumption-aware virtual machine allocation in cloud data center," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–6, doi: 10.1109/GLOCOMW.2016.7849005.

[118] Z. Zhou, J. Yu, F. Li, and F. Yang, "Virtual machine migration algorithm for energy efficiency optimization in cloud computing," *Concurrency Comput., Pract. Exp.*, vol. 30, no. 24, Dec. 2018, Art. no. e4942, doi: 10.1002/cpe.4942.

[119] B. Gul, I. A. Khan, S. Mustafa, O. Khalid, S. S. Hussain, D. Dancey, and R. Nawaz, "CPU and RAM energy-based SLA-aware workload consolidation techniques for clouds," *IEEE Access*, vol. 8, pp. 62990–63003, 2020, doi: 10.1109/ACCESS.2020.2985234.

[120] S. E. Dashti and A. M. Rahmani, "Dynamic VMs placement for energy efficiency by PSO in cloud computing," *J. Exp. Theor. Artif. Intell.*, vol. 28, nos. 1–2, pp. 97–112, Mar. 2016, doi: 10.1080/0952813X.2015.1020519.

JASPREET SINGH received the M.Tech. degree in computer science and engineering from Punjab Technical University, Jalandhar, Punjab, India, in 2013. He is currently pursuing the Ph.D. degree in computer science and engineering with Chandigarh University, Mohali, Punjab, India. He is also an Assistant Professor with the Department of Computer Science and Engineering, Chandigarh University. He has more than 17 years of teaching experience. His research interests include image processing and cloud computing.

NAVPREET KAUR WALIA received the Ph.D. degree in computer science and engineering from Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India, in 2021. She is currently the Head and an Associate Professor with the Department of Computer Science and Engineering, Chandigarh University, Mohali. With more than ten years of experience in teaching, research, and administration, her expertise spans various domains. She has authored more than 20 research papers, accumulating more than 100 citations. Her research interests include cloud computing, artificial intelligence, machine learning, and cybersecurity.