



Recent advancement in VM task allocation system for cloud computing: review from 2015 to 2021

Arif Ullah¹ · Nazri Mohd Nawi¹ · Soukaina Ouhome²

Accepted: 11 September 2021 / Published online: 23 September 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Cloud computing is new technology that has considerably changed human life at different aspect over the last decade. Especially after the COVID-19 pandemic, almost all life activity shifted into cloud base. Cloud computing is a utility where different hardware and software resources are accessed on pay per user ground base. Most of these resources are available in virtualized form and virtual machine (VM) is one of the main elements of virtualization. VM used in data center for distribution of resource and application according to benefactor demand. Cloud data center faces different issue in respect of performance and efficiency for improvement of these issues different approaches are used. Virtual machine play important role for improvement of data center performance therefore different approach are used for improvement of virtual machine efficiency (i-e) load balancing of resource and task. For the improvement of this section different parameter of VM improve like makespan, quality of service, energy, data accuracy and network utilization. Improvement of different parameter in VM directly improve the performance of cloud computing. Therefore, we conducting this review paper that we can discuss about various improvements that took place in VM from 2015 to 20,201. This review paper also contain information about various parameter of cloud computing and final section of paper present the role of machine learning algorithm in VM as well load balancing approach along with the future direction of VM in cloud data center.

Keywords Cloud computing · VM · Load balancing approach · Data distribution · Virtualization

✉ Arif Ullah
arifullahms88@gamil.com

Nazri Mohd Nawi
nazri@uthm.edu.my

¹ Soft Computing and Data Mining Centre (SMC), Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, Malaysia

² Department of Computer Science, Faculty of Science, Ibn Tofail University, Kénitra, Morocco

1 Introduction

Cloud computing is new technology that has extensively transformed human life from the last decade by providing different services and resource with the help of internet using virtualization system. Cloud resources are variable to every user by rent and release with specific rule and regulation with the help of internet. Cloud computing is emerging technology used for storing and accessing resource and application over the web based internet (Buyya et al. 2018). Cloud data center consists of physical and virtual infrastructure resources which include server, network system and different resources. Cloud data center is so important that different user demands can access data with the help of these data center in accurate and fast time. It contains a large amount of data and information which works under certain rules and regulations (Ouhame et al. 2020; Ullah et al. 2020). Data center normally used to control various activities such as virtual machine creation and destruction, routing of user request, network management, resource management and load balancing technique these all activity are performed with the help of virtualization. Virtualization is the processes in cloud computing in which we create virtual of resource based on software like hardware resource and software resource. Virtualization is one of the main element of cloud computing where this technique help to enhance the accuracy an efficiency of cloud computing. Virtual machine (VM) is one the main element of virtualization (Ziyath and Senthilkumar 2020; Mishra et al. 2020). When a single resource of cloud computing can appears as multiple resource this process can be achieved with the help of virtual machine. For the improvement in VM activity different kind of load balancing technique are used. A valuable load balancing technique in cloud computing can enhance the accuracy and efficiency of cloud computing performance. Because now a day user demands for fast and accurate service within the given time (Devi and Uthariaraj 2016). Load balancing technique is one of the main parameter for providing fast service in cloud environments. Cloud technology growing rapidly and used in different field of life like education, engineering technology, data-intensive applications, health, life science, geospatial sciences and different scientific and business domains. Cloud computing become an admired technology in worldwide due to its offer a huge amount of storage and resource to different companies and organization to access these resource with proper management, rule and security. Some of the main characteristics are virtualization, viability, large network access, automatic system, security, economical and scalability (Chang et al. 2010; Ferrer et al. 2019). The rest of this paper is organized as: The necessary backgrounds for the cloud computing and load balancing approaches are discussed in Sect. 1. Preliminary is discussed in Sect. 2. In Sect. 3, present the related work. In Sect. 4, Methodology of paper. In Sect. 5, we present about simulation environment. Section 6, present result and discussion and Sect. 7, conclusion. For the comfort of readers we provided a list of the most frequently used acronyms in the paper are mention in Table 1.

1.1 Contribution of paper

The major contributions of this review paper are summarized as below.

- i. Present details information about cloud computing.
- ii. Existent details information about load balancing approach and different parameters.
- iii. Extant details information about VM role in cloud data center.
- iv. Summarizations of main contribution in load balancing technique from 2015 to 2021.

Table 1 List of acronyms

Symbol	Description	Symbol	Description
CC	Cloud computing	IaaS	Infrastructure as a Service
VM	Virtual machine	CDC	Cloud Data center
LB	load balancing	IaaS	Infrastructure as a Service
DC	Datacenter	PaaS	Platform as a Service
MP	Memory page	VMM	Virtual machine monitor
QoS	Quality of Service	PM	Physical machine
IoT	Internet of thing	LAN	Local area network
5G	Fifth generation	CD	Cloud datacenter

- v. Present-day research direction about load balancing approach and role of VM in cloud computing.

1.2 Source of information

Data related to our review were extracted from 106 published papers. This collection of papers has been compiled by consulting various peer-reviewed data sources (Table 1). These papers highlight the recent advancement in VM task allocation system for cloud computing from 2015 to 2021. The frequency of publication of this work per year for the last six years was calculated to visualize the evolution of research on this promising thematic of the advancement in VM task allocation system for cloud computing which present in Fig. 1.

Figure 1 present paper selection criteria and Table 2 show the data source using these search engine different paper are download from 2015 to 2021.

Table 2 show the data source using these search engine different paper are download from 2015 to 2021 based on those paper this review paper was conducted.

2 Preliminary

In this section we define those element which are related about cloud computing and load balancing approach. These parameters are used in VM's for improvement of cloud data center performance.

2.1 Cloud computing

According to Nadimi et al. (2020), cloud computing becoming a crucial technology due to hosting various IT resource for different organization. By providing different service on demand through with the help of virtualizations rule based on pay and get rule. Cloud computing is new trend of technology which change human life over the last decade. This achievement has been done due to the deliveries of virtualized IT resource with the help of internet. The user can demand to these services with specific rule of pay and gain on real time environments (Abd Elaziz and Attiya 2020). Cloud resources are as long as collective tools and any user can charter and let loose these resources with help of internet. The technology become popular due to the combination of high bandwidth announcement and

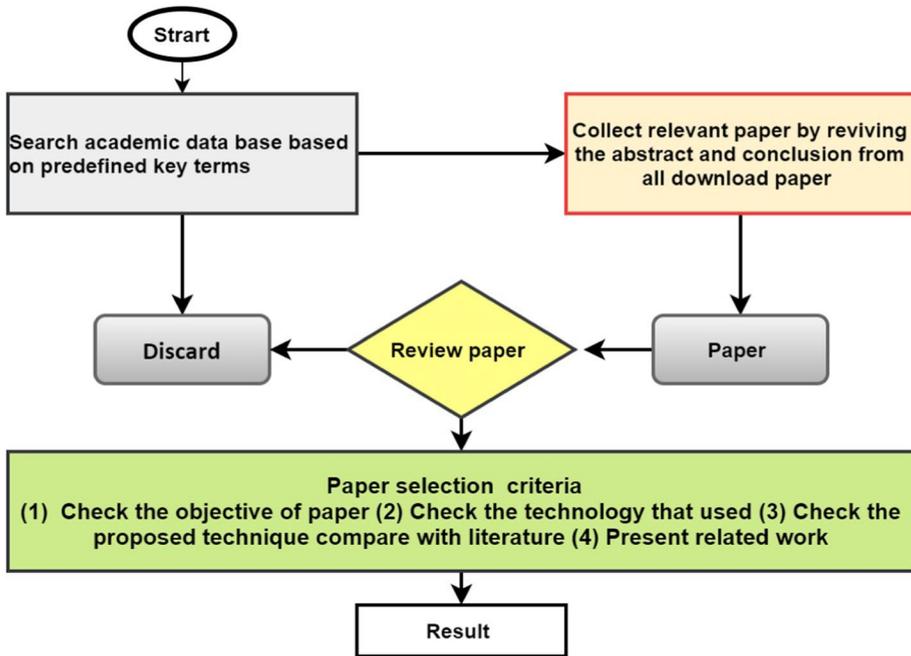


Fig. 1 Paper selection procedure

Table 2 Database source and source URL

Source	URL
Google Scholar	https://scholar.google.com/
ACM Digital Library	http://dl.acm.org/
DBLP URL	http://dblp.uni-trier.de/
Springer	www.springer.com
Taylor & Francis	http://taylorandfrancis.com
Wiley Online Library	http://onlinelibrary.wiley.com
IEEE Explore	http://ieeexplore.ieee.org/

low cast computing with storage (Ibrahim 2021).The entire bustle of cloud computing are performed with the help if internet to a certain extent having these service on local or personal computer. Four main recognized type of cloud computing are private, hybrids, public and community cloud. Poles apart users and organizations used cloud computing according to their persevere and obligation (Zhou 2020). There are four types of cloud computing which are used in different field of life with specific rule and respective specification. Cloud computing consist of different types of layer and these have specific role. Application layer it consists of a cloud application which is used in a different field. It is the highest level of the hierarchy and works as an automatic scaling feature. Application layer defines the commands, responses, data types, and status reporting supported by the protocol. This layer is the only layer that directly connected or interacts with the end user. It provides different application for user. Such as simple mail transfer protocol, file transfer, surfing

the internet, chatting with friends, email clients, network data sharing, and various forms of file and data operations (Mezgár and Rauschecker 2014; Samimi and Patel 2011). Platform layer this layer consists of an operating system and application framework and sits on the top of the infrastructure layer. The primary purpose of this layer is to minimize the burden of developing an application or reduce the complicity of development rule work under virtual machine (VM). Different API and application are used for storage data-based and logical web application (González-Martínez et al. 2015). Infrastructure layer this layer creates a pool of resource for storage computing resources with the help of a technology known as virtualization. It allows (IaaS) customers to create and discard virtual machines and networks as per their business requirements. They pay for the services they consumed (IaaS) removes the necessity for the consumer to invest in procuring and operating physical servers, data storage systems and other networking resources (Rimal, et al. 2010). Hardware layer this layer is responsible for the management of all physical resource of cloud computing. Such as physical server, routers, switches, power and cooling system along with different resources (Khan et al. 2017). Figure 2 present the structure of cloud computing.

Hardware layer is typically implemented in the cloud data center where it consists of thousands of different physical resources and they are connected with different rule and regulation. All layers are important due to their different operation and connectivity with each other (Lee et al. 2018). Platform layer is the important layer of cloud computing because it includes different operating system and software development framework that provide resource to the end user. In little year cloud computing has experienced remarkable growth in economic model and development models. All activities are happened due to platform layer of cloud computing. Different virtualization systems are controlled and

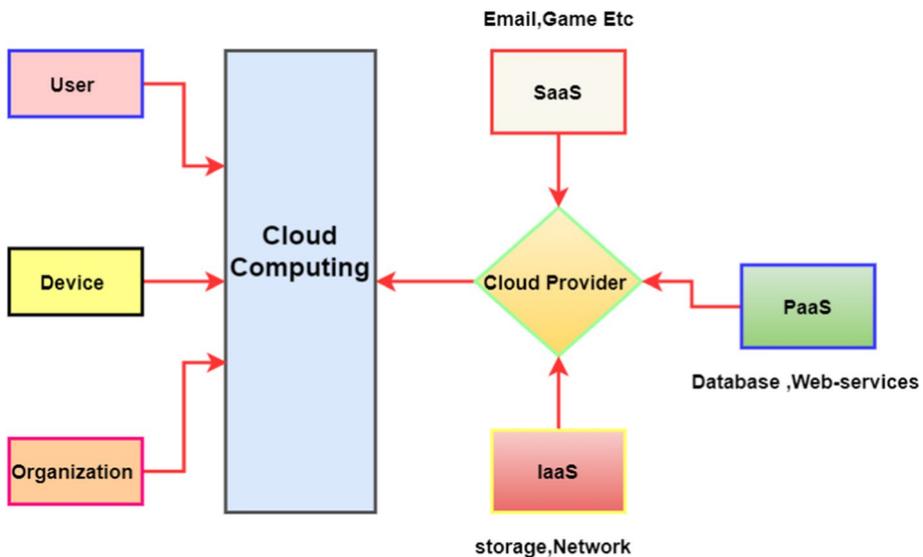


Fig. 2 Structure of cloud computing

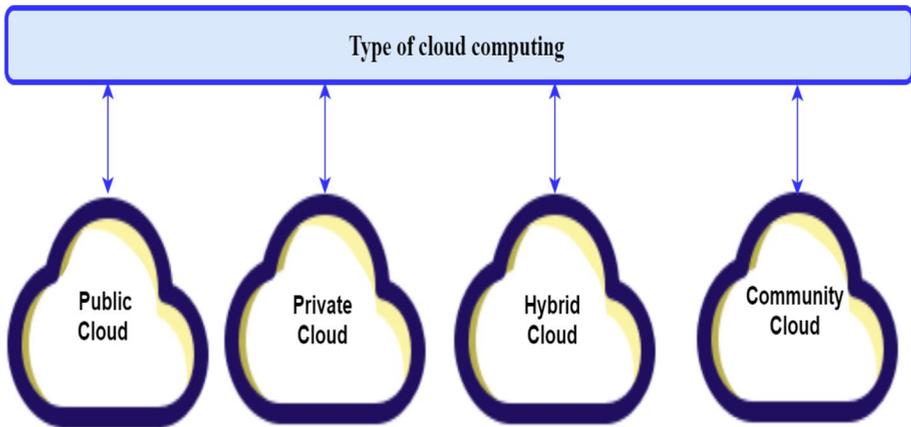


Fig. 3 Types of cloud computing

developed with the help of this layer because it deals with software and virtualization system. Virtual machine is one of the main elements of this layer (Mastelic et al. 2014). Cloud computing consist of different types Fig. 3 shows types of cloud computing.

Figure 3 show the different type of cloud computing and these are used in different area of life for different purpose. Private cloud it is designed for a single organization and also known as the internal cloud. It is established within the organization or connected with the third party and it is much secured as compared to other types of computing. Private cloud computing is established for the requirement of third party or the demand of third party. It is more secure and reliable due to the restriction and rules therefore it become more expensive than other type of cloud computing (Kuyoro et al. 2011; Kotha et al. 2021). Hybrid cloud it is the combination of public and private cloud with the respective strength and weaknesses. Organizations attempt to achieve the best from both types of the cloud and also known as the federation cloud. It is the intermediate between private and public because sometimes the user who uses public cloud shift to the private cloud because they suddenly need more secure data. This process can be done with the help of the hybrid cloud (Liu et al. 2011). Community cloud is cloud service model which provide service to a limited number of individuals or organization that managed and secured by all participating organizations or a third-party managed service provider. Community clouds are a hybrid form of private clouds built and operated specifically for a targeted group. These communities have similar cloud requirements and their ultimate goal is to work together to achieve their business objectives (Hashem et al. 2015).

2.2 Cloud computing characteristics

According to Abd Elaziz et al. (2020), cloud computing is a general term for anything that involves in delivering hosted services over the internet. Cloud providers are competing with each other and they constantly expand their services in order to differentiate themselves. Cloud computing is named as such because the information being accessed is found remotely in the cloud or a virtual space. Cloud computing has succeeded in bringing change in different field of life. Availability cloud computing provides high availability and some benefits for every type of user in a different field. Availability is one of the main

pillars of information system where it refers to the uptime of system and network of system where they collectively provide a service during its usage. The traditional systems are limited on local installations when they shift in to cloud computing where entire organization are able to use availability service of cloud computing along with end user. When it time for IT infrastructure establishment then an organization make decision on the bases of availability because it is the main key decision factor. Availability have been the major concern in distributed system because highly available service in cloud computing and are main element for satisfaction of cloud user (Zarandi et al. 2020). Scalability is the attribute that presents the ability of the software, network and process of an organization to manage the increasing of user demand. Normally scalability means frequent speed in cloud computing in which the ability of system or product to continue working after its context changed like volume or size in order to meet the user need. Scalability is a sign of stability and competitiveness which means the organization or network system are ready to handle the influx of demand according to change need and update of the system. Due to the property of scalability in cloud computing lots of companies are shifting to cloud computing (Phanden et al. 2011). Cloud security also known as cloud computing security that consists of different policies, controls, procedures and technologies that work together to protect cloud-based systems, data and infrastructure from unauthorized access. Cloud security is a joint responsibility of cloud provider and business owner or end-user. Security addresses both physical and logical issue in different model and layer (Giret et al. 2015). Cloud automation is a broad term which refers to the processes tools and resources that used by an organization to reduce the manual efforts and it associated with the managing cloud computing workloads. It can be applied to different types of cloud computing. Cloud automation is a fundamental building block for cloud computing. It can be applied in a software layer where a complex system is used to configure and roll out the system balances for the network system. The aim is to make all activities related to computing is as fast, efficient, and handoff as possible thought for the use of the various systems (Elsherbiny et al. 2018). One of the main character of cloud computing is virtualization. It is one of the main elements of cloud computing with refers virtual rather than actual of something. Where IBM introduce virtualization concept in 1960 and early 1970 and this technology reliable and time sharing (Wang et al. 2020). Normally different types of virtualization used in cloud computing some of the important type of virtualizations are mention Fig. 4.

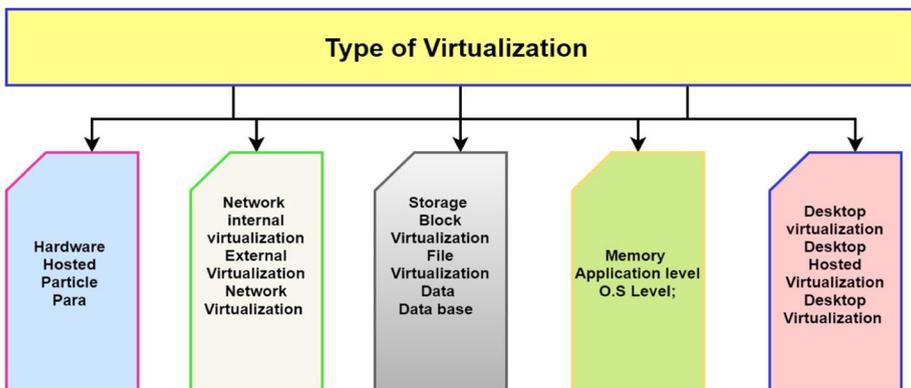


Fig. 4 Types of virtualizations

Hardware virtualizations is the type of virtualizations where process several different services are combined and work together as single or different VMs work as servers. These processes have different types and are used for different purposes (Smith and Nair 2005). Full virtualization in this type of virtualization where VM can run any operating system using the system resource or hardware. Main types of full virtualization are Xen server, Xen, KVM and Virtual box (Obasuyi and Sari 2015). Virtualization the guest operating system does not run fully on the virtual machine. It does not work fully and it gets help from hypervisor or with the VMM for working. It just improves the functionality of the operating system. Partial virtualization is the process in which software modification took place any tag (Guo et al. 2010). Network virtualization this type of virtualization we combine software network resource and hardware network resource which are in the same network under the same administrative units. These types of virtualization allow network optimization and scalability in large network as well as use full for improvements network efficiency and productivity. This type of virtualization further divided in two types. External; in this type of virtualization it combines several networks in to single unit. Internal; in this type of virtualization virtual network interface cards are used for network life functionality in to single system this process is known as internal network virtualizations (Sharif et al. 2009). Storage virtualization is the process of arranging different physical storage from multiple network storage in to single form. The technology that refers to identify available storage and capacity from different physical device in to a pool of storage device which can be used as a virtual environment. Server virtualization is the processing of making a physical server in to the virtual server. The server administrator uses a software application that divided the server in multiple isolated virtual servers and that also acts as a physical server (Sheikholeslami and Navimipour 2017). Desktop virtualization' the process of isolating a logical operating system in to client that can access it. There are many concepts of desktop virtualization which are dividing into different categories according to user demand. Hosted virtualization in this scenario the virtual machine is completely intellection of a real physical machine. All the feature of a real physical machine like memory, operating system and storage are also in hosted virtualization. It can be achieved by the configuration of the real system (Bhandia et al. 2019). In memory virtualization technique whenever the memory required for the system processing and actual memory is less than the virtualization process is used for memory. In data virtualization, the collection of data for different location and the user can access them easily. It provides front and back ends application method (Abramson et al. 2006).

2.3 Virtual machine (VM)

In cloud datacenter if a single resource can materialize as multiple resources and this route can be achieved with the help of virtual machine (VM). Efficient VM is very important for energy saves and improvement the working of cloud computing. The VM normally replaces the physical resource with their ability and operating system that makes the same environment as hardware. The VM provides a better security model as compared to the normal system VM which is also known as a guest machine. VM are an efficient isolated duplicate of real machine which allow the multiplexing of the underlying physical machine. VM technologies allow great deals of elasticity in dynamic management of workload on servers (Ullah et al. 2020). Figure 5 present the working criteria of VM in cloud data center.

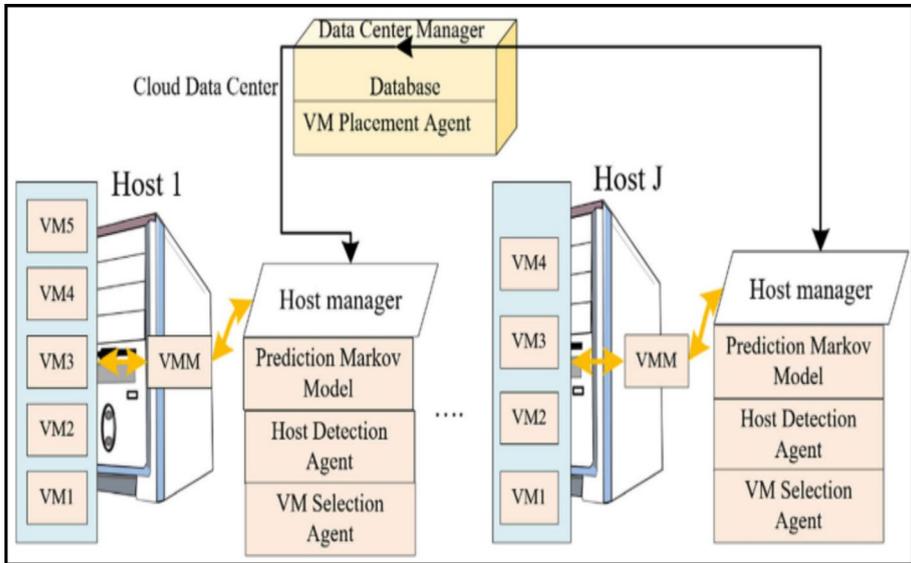


Fig. 5 Working section of VM (Manasrah et al. 2017)

Virtual machine is a software based component which is an abstraction of the underlying hardware provided by the virtualization technology. Availability can be efficiently achieved by managing the virtual machines properly (Ullah et al. 2020). One of the main element of VM are hypervisor or VM manager where the model or program that handles more than one VM operating system on single host. It is a responsibility of hypervisor to provide resources and a processor to each virtual operating system on the same host. VM manager has two main types, bare metal hypervisor, and hosted hypervisor (Nasim and Kassler 2014). Bare metal hypervisor is type of hypervisor it work on hardware and control over all its accessible resources like memory, CPU etc., as no intercessor is required to access the resources like Hyper-V, VMware. Hosted hypervisor is the type of hypervisor which is used for installed in the operating system of a server and that operating system has a control over it example are virtual box, xen, VM ware player (Hwang et al. 2015). Migration of VM's is the process in which one host shift to another without disrupting existing work this is process is known as VM migration. There are two types of VM migration which are online and offline migration. Online migration concept was introduce by Christopher Clark in this process running VM can be transferred to other host without affecting the system or the server this process can be done by making the system off the VM. When it transfer to the host it restart again all these activity done in dynamic system (Lee et al. 2010). Offline migration in this process the transfer of VM to another host during running of the system can interruption the network or server during the transfer of the system. VM migration is one of the main elements of virtualization system in cloud computing which allow the movements of VM from one host to another with different rules. VM migration need both sender side and receiver site for transfer of file and VM and state file. VM migration play different rule in cloud computing but one of the main rule is VM replacement system as resource allocation and communication system (Zhou et al. 2010). Preparation time the time between the start of migration process and VM processor state is sent to the

host during the VM run is known as preparation time. The main parameters of preparation time are given below.

- i. Resume time is the between the resuming of the VM running and migration is called resume time.
- ii. Pages transferred is the amount of memory pages transferred comprising the copies of pages.
- iii. Downtime is the time during which the running of virtual machine is stopped it contains sending of state of the processor.
- iv. Total migration time between instigating phase of the migration to the end of the migration process.
- v. TMT is used for the resource emancipating on both the source and the terminus node (Beloglazov and Buyya 2015).

Live virtual machine migration techniques: Pre-copy migration warm-up phase this is the in early phase for the hypervisor to create the copies of required memory pages to send to the destination section from the source section. During this route virtual machine is not terminated by the hypervisor. If some variations are there in the pages of source node during the process of moving the replacement copies. Then memory pages will be duplicated again and again until the data reduplicating rate is less than the rate of moved page which does not contains the recent value (Choudhary et al. 2017). Figure 6 present the working of VM section.

Stop-and-copy phase in this process the VM are terminated at the source node the amount of fata changed that is left will be moved to the destination node and VM start processing at the destination (Ullah et al. 2019). Post copy migration in this type of migration

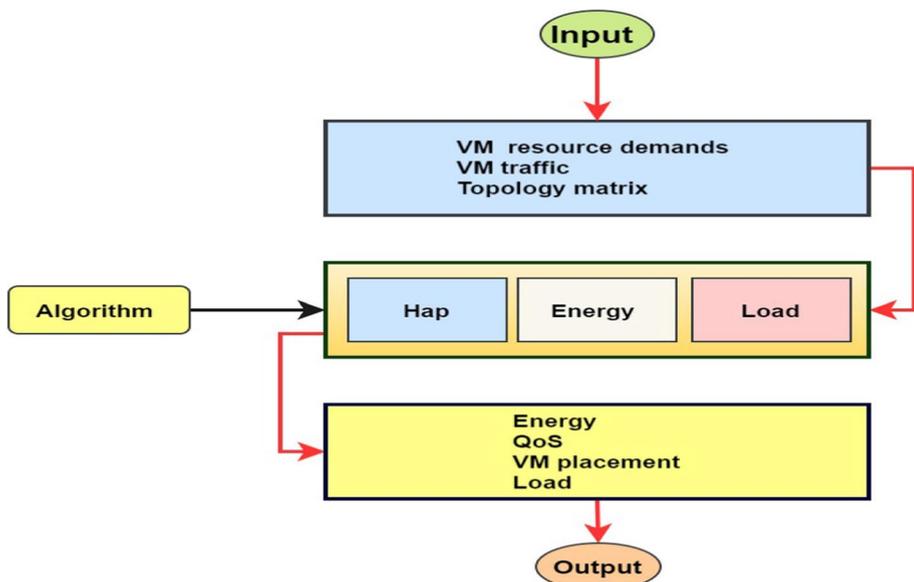


Fig. 6 Working section of VM (Ahmad et al. 2015)

technique VM is suspended from source host for some time then the VM state is transferred to the destination host. Once it is received at terminus, at the same time it starts working after resource portion process even though most of the memory state is living on the source host. When virtual machine tries to fetch the pages which are not being sent from the endpoint host; it will create page faults (Yang et al. 2017; Hussain et al. 2019). Hybrid VM migration this type of migration used both pre copy and post copy migration techniques properties. It is divided into different phase which are as:

- i. Preparation phase: System resources required at the object host are reticent.
- ii. Bounded pre copy rounds phase: Determine the surrounded pre copy rounds and working set of VM is reassigned from sender server to the receiver server.
- iii. Virtual machines resume phase: At the receiver server, it inaugurations the reassigned state.
- iv. On demand paging phase: On the basis of application, requests of read/write.

Preparation phase: System resources required at the object host are reticent (Navamani et al. 2018). Figure 7 present the VM migration section.

Post copy variations: Post copy through demand paging in this process the pages are reassigned only once and it will result in page fault when requesting the referenced pages from the source node over the network. As a result, it will slow down the dispensation of the VM as it raises the length of the resume time and creates the enslavements in the form of a page faults residing for erratic time periods (Shah 2011). Post copy through active pushing: Enslavements is bargain in the form of an undetached pages living for changeable time periods; one way is to initiatively push” the pages on the endpoint node from the source even though the VM continues running at the destination host. Active push evades

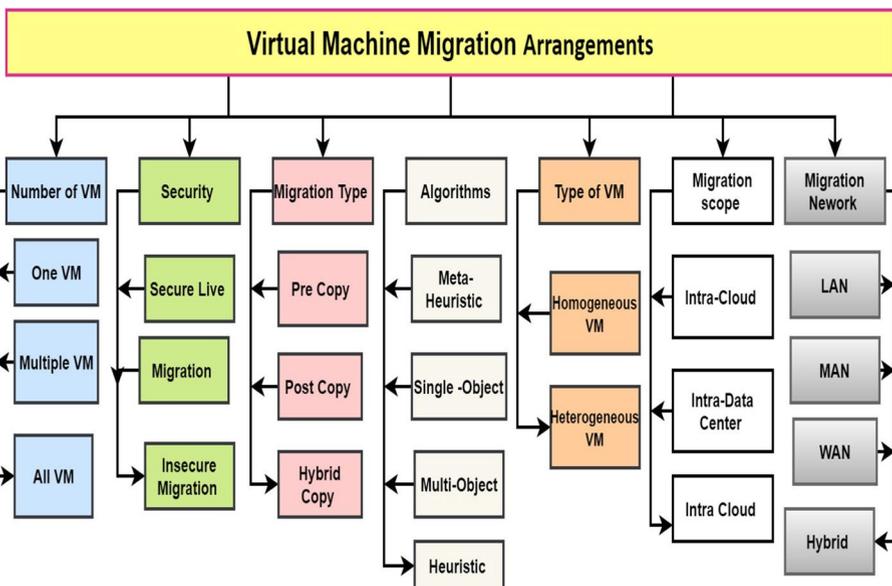


Fig. 7 VM migration section

the transfer of pages which are criticized in the destination virtual machine. Thus, pages are only sent once either by petition paging or active ambition (Umar and Baseer 2016; Beloglazov et al. 2011). VM replacement system consists of two main sections which are static and dynamic system. Static VM replacement system in which the mapping of the VM is fixed during exaction time and change able for fix time. Dynamic VM replacement system in which the VM allow changes at any stage during the expectation time and those algorithms which are used for replacement of VM are called reactive and proactive VM placement (Indukuri 2016). Figure 8 show the VM replacement system.

Post copy through pre paging: It is very difficult to know the strict fault tolerance bearing of pages but by approximating the defective speeches to predict VM’s memory access pattern, we can presume better page forceful sequence to access the arrangements and the incidence of page faults in improvement (Besta et al. 2019). Table 3 present the different parameter of VM migration section.

Table 3 present the VM migration parameter while Fig. 9 present VM migration metrics and these all elements improve the VM working criteria.

Figure 9 shows the VM migration metrics and these all elements improve the VM working criteria. When user demand increase in cloud data center for accessing data then some VM become under-loaded or become overloaded which may delay or failure of the system. In order to avoid this kind of situations then load balancing technique is used.

$$f(VM) \left\{ \left\langle /Ti - \sum_{v=1}^K Lv \quad \text{Underloaded} \right\rangle /Ti - \sum_{u=1}^K Lv \right.$$

$$\left. \text{Overloaded} = /Ti - \sum_{i=1}^K Lv \quad \text{Balanced} \right. \tag{1}$$

where T_i present the given task, K is the capacity of VM where given data are compare with these value (i-e), $i, u, \text{ and } v$ (Ullah et al. 2019). To solve the above situation in VM and servicer of cloud computing load balancing technique used. When research improve the above Eq. (1) then the metric of VM improve which are mention in Fig. 9 for improvement of the above activity load balancing approach used.

2.4 Load balancing technique

Load balancing technique used for management of resource, data and application on circumstance that maximum throughput with slight time and also diving different type of traffic between VM’s and servicer without any impediment. Due to the growth and improvement in cloud technology there is increase of user and they demand for better services (Tracz et al. 2019).Virtual load balancer offer more elasticity to balance the workload by distributing the resource or traffic across multiple VM’s. Virtual load balancing aims to

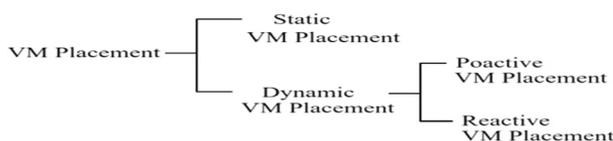


Fig. 8 Type of VM placement

Table 3 VM migration parameter

VM Migration	Technique	Rewards	Minuses
Non-live migration	Stop VM at source then transfer	Simple concept and easy to implement	Down time is more
Post-copy	First transfer the execution and then the memory	Memory transferred in a single pass and has less network overhead	More down time as compared to pre copy
Pre-copy	First transfer the memory and then transfer the Execution	Down time < 1 s. On aborting Migration, systems do not crash due to running VM in source host	Overhead of duplicate page transmission

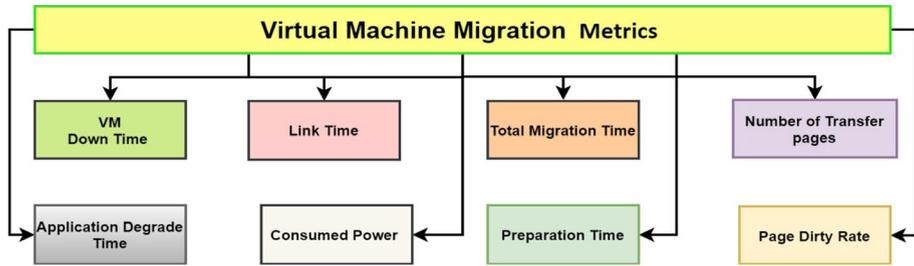


Fig. 9 VM migration metrics

mimic software-driven transportation in the course of virtualization. It runs the software of a physical load balancing appliance on a virtual machine (Bamgbade et al. 2019). Load balancing technique has become an important approach that has been used for reduction of response time and provide maximum throughput with slight time duration. When multiple requests from user site are receive load balancing approach properly distributed them to different device according to their accessibility. If load balancing approaches are not used in any area of cloud computing then user and provider wait for a long time that they request for any resource and sometime deadlock occur (Mushtaq et al. 2017). Therefore different kind of load balancing technique used in cloud computing which are mention in Fig. 10 these technique are used different sections of cloud computing for different purpose.

Figure 10 show all type of load balancing which are used for different purpose in cloud computing but in this paper we just focus on load balancing technique which are used for VM resource allocation system. Essentially there are two main types of load balancing which are dynamic load balancing approach and static load balancing approach. In dynamic load balancing approach where it can be change able at any stage of the network during the process system. The main benefit of dynamic load balancing is that if any node fails during the network execution it not affects the system and it not affects the current stage of the network (Milani and Navimipour 2016). While static load balancing approach the performance on the current stage affect the network or it static load balancing approach which is not change able during the network execution. It normally works in homogenous and stable environment therefore it provides good result. These two types of load balancing approach used for different propose and these are future divided in to different groups (Zaki et al. 1996). For improvement in load balancing technique different algorithm are used or we can say load balancing approach can be implement with the help of algorithms. Different types of algorithm used for improvement in load balancing approach and Fig. 11 show the type of algorithm which is used in cloud data center for different purpose.

Figure 11 shows all those area where different algorithms are design for improvement in load balancing purpose for cloud datacenter.

2.5 Load balancing technique facing different kind of issue

Cloud computing change the human life as aspect of different filed and environment by providing different service and resource according to the user demands. However cloud computing facing different kind of issues and load balancing technique is one of them. Load balancing technique also facing different kind of issues which are (1) VM migration: virtualization is process with the help of this technique several VM's are created in single

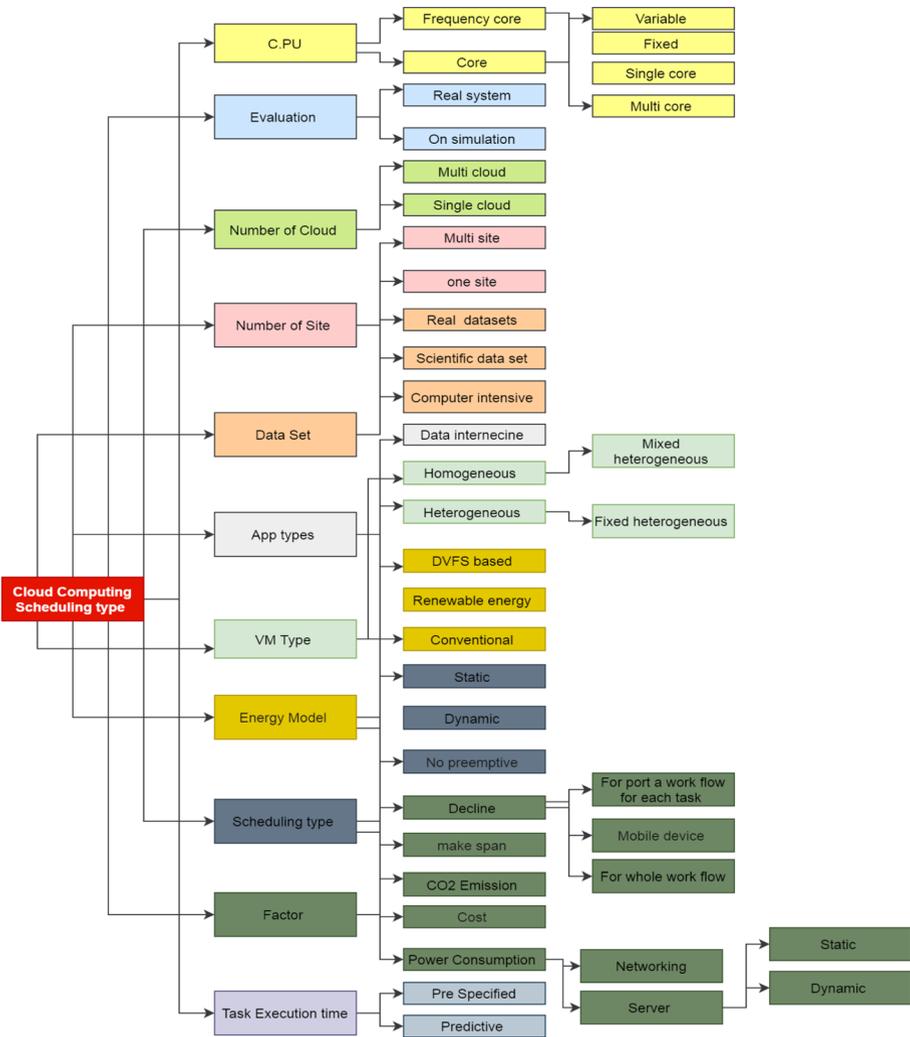


Fig. 10 Type of load balancing technique

physical machine. These VM are self-governing in nature and have different configuration when a physical machine become overloaded then these VM are used to transfer the data using VM migration approach this approach facing algorithm connectivity problems at some points this process is known as VM migration (Xu 2012). (2) Single point failure: some dynamic load balancing approaches are designed by center node approach system it means all the decision of the movement of the network or data in the network depends on that center node. If the center node crashes down then the entire network of system will be down for that reason there will be a proper algorithm need to developed that manage and distribute the work load equally not depends on the center node if it down (Milani and Navimipour 2016). (3) Storage management: cloud technology has solved the traditional storage system because the traditional storage system consists of various issues but cloud allows the user to store heterogeneously without any access problems. But cloud storage

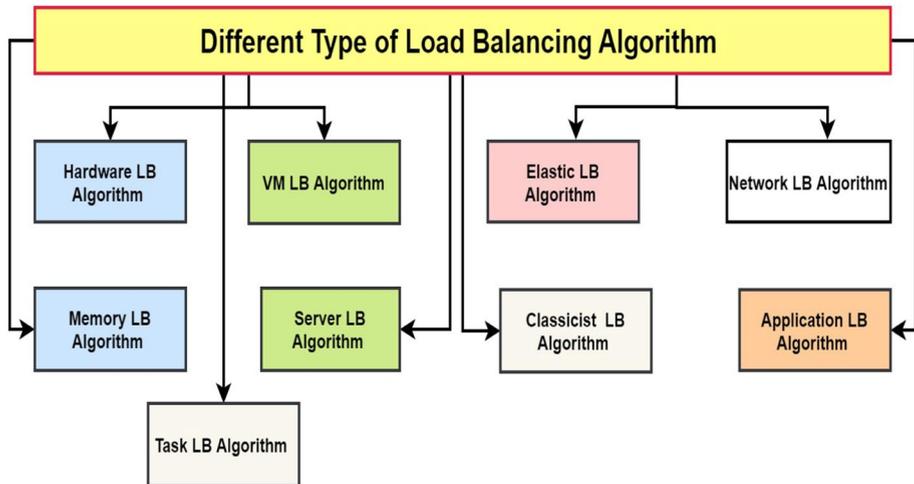


Fig. 11 Algorithm used in cloud datacenter

increase day by day which causes replication and consistency issues in data size which make duplication data storage policy and replication issues. These issues make data set availability and increase complexity for load balancing technique because they take more time for exestuation. (4) Algorithm complexity: in load balancing technique algorithm must be implementing according to the demand and simple way. When we used complex algorithm it reduce the performances and accuracy of the system (Imdad et al. 2020). (5) Load balancer scalability: on demand of user different service is providing within the time load balancer section work. If good load balancer section used it solve the user demand it term of power storage and efficiently. (6) Selection policy: when the tasks or data are selected for transfer from one VM to another it used. This policy identifies the data or task based on amount of overhead that is required for migration from one point to another (Khan et al. 2019; Zaki et al. 1996).

3 Related work

In this section we present some of the pervious researcher work they modify, hybrid and introduce a technique for improvement in load balancing technique in cloud data center.

Wang et al. (2010), proposed load balancing Min Min (LBMM) algorithm where the framework of this algorithm consist of three levels which are request manager, server manager, VM manager. The request manger is responsible for receiving the workload and then assign to the server manager where servicer manager distributed the data to different VM where VM manager executed the data. In this section the author modified the VM manager in to two sections first it checks the data priority after the priority base sent the data to execution. The author found out that the proposed method suffered from some limitations such as the working process of this algorithm is slow due to the three level information where after the information step it check the parameter then assign the task to VM's. Sharma (2015), modified the throttled algorithm where previous algorithm assigned the task to VM directly but after the modified it first requested to load balancer section of appropriate VM

for that task it check the index then the task was assigned to VM. The author demonstrated some limitations of that proposed techniques which is the task are not given priority base. Domanal and Reddy (2014), defined good load balancing that improved the throughput by minimizing the response time along with fault tolerance. Normally load balancing problems occur during the over demand of VM's. In order to overcome load balancing the author modified dynamic weighted live migration (DWLM) algorithm for selection and location polices in virtual machine. The author changed selection part in this algorithm for allocation of task. Yakhchi et al. (2015), proposed Cuckoo optimization algorithm which consisted of three different part in which first phase applied COA to detect the over load VM and transfer to under-loaded VM. First iteration information was stored in habit second iteration information were stored in ELR formula that checked the VM and sent data to load balancer section. Even though the proposed technique performed better as compared to the normal COA search but the execution time of the algorithm was increased. Monika et al. (2015), proposed honey bee galvanizing algorithm in which he increased honey bee forage technique with random stealing is employed for task and cargo leveling. They found out that if the VM become overload then the task moves to the neighborhood VM. If those load worth is less than threshold value, then this task were given to that VM. The proposed technique improved the efficiency parameter as well as the average time of network but network still facing problems in accuracy. Babu and Samuel (2016), proposed an algorithm name as enhanced Bee colony algorithm that works the same as the bee work for food source. Authors proposed load balancing mechanism work in to four different steps which are VM current load calculation, load balancing decision, VM grouping and task scheduling. The authors modified VM current load calculation section and changed the fitness value in this section. Even though the modification improved the quality of services and migration level but still exists the efficiency problem at data allocation section. Babu et al. (2016), proposed honey bee forging algorithm known as (HBLL-B) where this algorithm work based on the behavior of honey bee. Two main type of honey bee are their one of them who find food source and the other one who reap finder honey bee go for search of food source when they find they come to comb and make dance it show the quality and quantity of the honey. Then the reaper goes on that food source get the food return to the comb and make dance if the food source remains. Like the same honey bee forging algorithm work in cloud computing for load balancing purpose where different VM are like honey bee and task are food source. Forger chooses on VM to check the profit of food if the profit is less then cost it stop the working. The author change two main parameter of algorithm which fitness vale and iteration section. However due to this change the computation of the profit may cause an additional overhead in which some task was not assigning the VM which wait for next round. Rani and Kannan (2017), proposed Bat algorithm for load balancing technique by assuring that every VM take more or equal amount of data and can get at any stage. Therefore, Bat algorithm was used along with column maxima technique in which total execution time is minimized and load is balance with total number of VM. However, the proposed technique was not affective for large number of task because due to the column metric function which works two dimensions. Holland et al. (2017), proposed Artificial Bee Colony algorithm for load balancing purpose in VM where it work same as the bee work in the nectar. The author modified and initialized the population where iteration move the employed bee on the food, onlooker bee determine nectar amount move source of searching for new food source memorize the best food source until condition. The researcher modified the initialize population because it was pervious randomly s ($i=0$) of FS (food source). FS is the size of employed bee equal to onlooker bee each iteration it took ($i=123$) up to food source count. Even though the change of fitness function

probability section improved but second phase onlooker section still not improved. Devi et al. (2018), proposed two techniques for improving load balancing. The first one is a theoretical concept in the network graph where graphic concept is used for monitoring the load and minimum dominating set (V-MDS) algorithm and second live virtual machine migration in virtual machine for load balancing improvement using new system and traffic-aware live VM migration for load balancing (ST-LVM-LB) algorithm and check the result with dynamic management algorithm (DMA). The researcher proved that the used of graph theoretical algorithm improved the load allocation but it took more time in term of verification. Fahim et al. (2018), studied about load balancing technique algorithm which are used for improvement in VM section. Authors compared five best algorithms in term of accuracy and performance of task. The study demonstrated that Bat algorithm worked more accurate in term of accuracy and efficiency in load balancing technique. Gamal et al. (2019), proposed an algorithm known as osmotic bio inspired algorithm used for load balancing purpose. The author used two algorithms which were ABC and ACO algorithm for load balancing in VM where both algorithms communicated with each other with the help of osmosis theory. One of the limitations of this algorithm was that the communication process took more time which cause the network slow. Krishna et al. (2020), proposed an algorithm known as OLOA in which modification took place at fitness function of Bat algorithm. Because the data distribution between different VM is done with the fitness function. Therefore, the author modified the fitness function at load balancer. The new technique improved the selection methods but still processing time was not increased significantly. Thanka et al. (2019), proposed hybrid algorithm using ABC and PSO algorithm where the study demonstrated that PSO algorithm facing the local search problems because at the last iteration of local search it did not complete the iteration therefore they changed with ABC algorithm because ABC algorithm is good at local search. The results showed that the proposed techniques improved in local search and it increased the data collection system however the data accuracy issue still existed. As the study of related works of improvement starting from traditional load balancing technique where no priority are given for any parameter and it concepts are very simple in implementation without error. Then researchers changed to activity based method by measuring every activity cost for all objects and outcomes and it showed that they are better than traditional. Next researchers shift to behavior based algorithm which are inspired from behavior of animals and other living organic system and the result of those algorithms were better than the previous two methods. The related work show that different type of algorithm used for improvement in load balancing section and different research implement different approach therefore we summarizes 106 paper in our result section that we get a brief information about load balancing approach in cloud data center as well as we know about VM and it parameter are changed for improvement of cloud computing. Our main focus will be machine learning algorithms that are used for VM task allocation system in cloud data center.

4 Methodology

To study briefly about load balancing technique for VM and what methods, are design for improvement in task allocation system for VM are used in cloud computing from 2015 to 2021 are discusses in this paper. Figure 12 show the paper selection methods form initial stage to final stage.

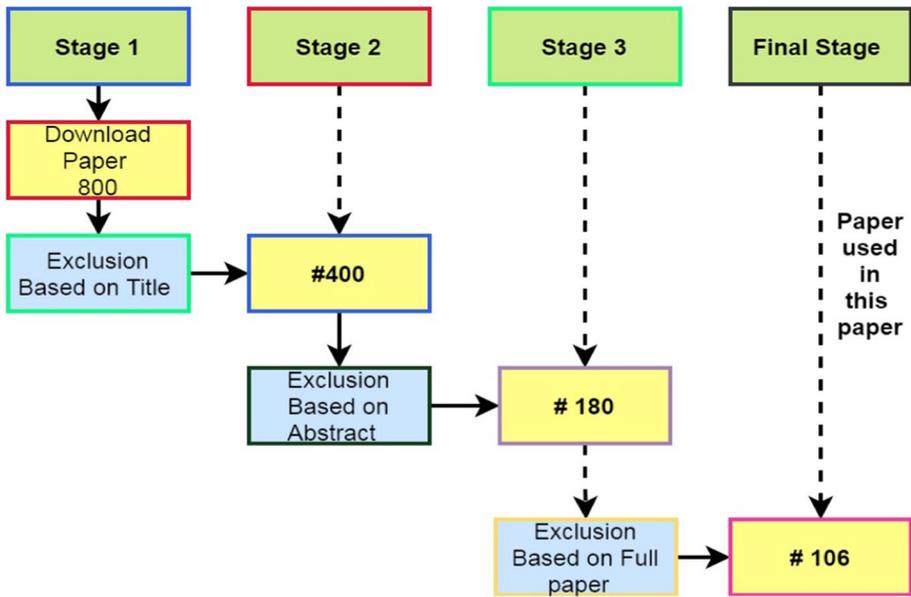


Fig. 12 The paper collection steps for this paper

Figure 12 show the collection of paper where these papers goes to different steps after that we select 106 for summarization at final step. The main roles of summarization of these are that we collect information about, technique, which section modification occur, advantage, disadvantage which author and at which year it present the paper. After collection of all these information we will able to define the recent advancement took place in VM load balancing approach for cloud computing.

5 Why VM need load balancing

Load balancing technique provides a service for distribution of load among different VM's equality. Main objective of this technique is continuous service in case of failure of any servicer in the network make any alternative services. In additional load balancing technique minimize the response time for data and improve task allocation system in VM which enhance the system performances at low cast. This technique also improve scalability and flexibility for those application which size are increase during next work execution as well as provide priority of different task during execution time with other task in the queue. Other objective of load balancing technique are reducing energy system improve network life time, accuracy, efficiency and improve the network quality of services (Oke et al. 2021; Sheikholeslam and Jafari Navimipour 2018).

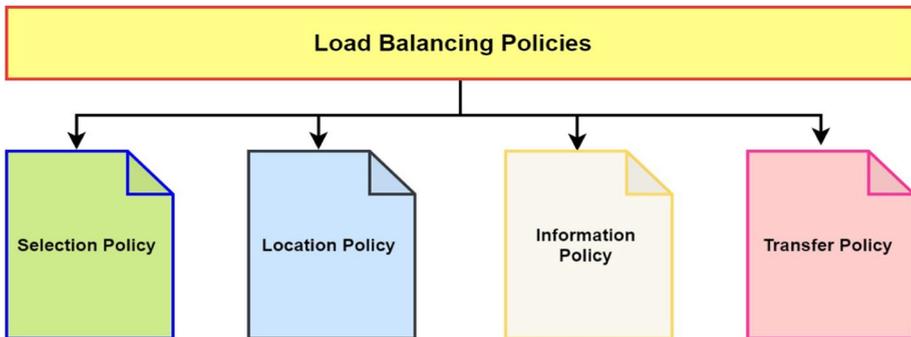


Fig. 13 Load balancing policies

6 Load-balancing policies

As mention in previously different load balancing technique are used for improvement different section of VM in cloud datacenter. For implementing these algorithm for load balancing technique different policy are used which are mention below: Figure 13 show these polices and rules.

Selection policy this process used identification method for data transferred system in the network and the selection of these task are performed based on amount overhead required migration after that required it transfer the data (Wong et al. 2014). **Location policy:** this policy different resource mention that they are under loaded or overloaded of that bases takes are sent to them. Which node or resource in under loaded on that bases it request to VM or service for data. During this process three main elements took participate which are probing, negotiation and random (Zarandi et al. 2020) **Random approaches:** the selection approach policy select the destination or receiver node randomly for transfer of the data in the probing approach the node used other node or resource to select this destination. **Negotiation approach:** node or resource negotiates with each other for proper load balancing approach (Kulik et al. 2002). **Transfer policy** in this kind of policy discovers the situation where task are sent to local to local or remote nodes. It consists of two approaches which are current task and last received tasks are used to identify the tasks information transferred. All incoming tasks are entire in to the transfer policy after that decision that task are distributed (Al-Karaki and Kamal 2004). **Information policy** this is the main policy of load balancing technique in which that contain all information in the system like agent, broadcasting, centralized polling and periodic policy are used for distribution of information between different node and resource in the network (Wang et al. 2019). All the above load balancing policies have relation with each other because when the task entire in the system and initially processed with the help of transfer policy. After this process the next policy decided whether it needs to transform to remote node or not. For checking the nodes statuses wither it under loaded or overloaded location policy used. If the node is overloaded then both transfer policy and location policy collect required information from the information policy to make decision (Al-Hashimi et al. 2019).

6.1 Load balancing technique performance

Load balancing techniques are design for improvement of different parameter in the network and these parameter results are comparing with pervious result. If the current technique improves the performance in terms of that parameter in the network it means the current load balancing technique working properly. Different researcher defines different parameter to check the performance of load balancing technique. The performances of any load balancing techniques are evaluated based on these metrics. **Exestuation time**: is also known as completion time in which a specific cloudlets or task time need to complete the job. Or the time taken by a job for running in system is known as exestuation time.

$$Exe_{Time} = Task_{(i)} (Fnh_{(Time)} Srt_{(Time)}) \quad (2)$$

where $(Fnh_{(Time)})$ denote the finishing time and $Srt_{(Time)}$ present the starting time.

Performance: is the process in which number accomplished tasks or resources on the demand of user in the network are known as performance (Eusuff and Lansay 2003).

$$Perfromance = Task_{(i)} \frac{(i * P_{(cl)})}{R} \quad (3)$$

where I denote the instruction $P_{(cl)}$ represent the performance and R exestuation time.

Priority: is the process in which a tasks or job are given priority for exestuation because it has demand for user site or it need in the network or for different process it need this is called priority.

$$Pr iorty = \sum_{Task(i)} (Exe_{(Time)} + Capacity\ number\ of\ request) \quad (4)$$

Reliability: is the ability of complete takes in the network with the given time and it provide an assurance of complete given number of task without avoid or reduce the failure rate the network this is known as reliability.

$$Reliability = \frac{\sum Task_{(i)} Exe_{(Time)}}{Total\ Time} \quad (5)$$

Response time: when a user request for data or tasks and it start execution and come out form the waiting queue this process is known as response time.

$$Responcetime = \frac{\sum Task_{(i)} Exe_{(Time)}}{Total_{(Time)}} \quad (6)$$

where Sub_{Time} present the submission and $Starting_{Time}$ present the starting time.

Availability: is the committable operation in which cloud provider will able to user demand or cloud resources are able when user request for specific operation. Availability is the combation of security, accessibility and serviceability in cloud computing (Taylor 2013).

$$Aviabilty = \sum \frac{Mt_{(ns)}}{Mt_{(ns)} + Mt_{(hr)}} \quad (7)$$

where $Mt_{(ns)}$ is present response time and $Mt_{(ns)}$ is mean time of repair time of resource i .

Bandwidth: is the process which present the maximum data transfer rate in the network this process present the network connection in given number of time or it present the speed of the network in the cloud computing.

$$BW = \sum Resource_{(i)} \left(\frac{Size}{Capacity} \right) \tag{8}$$

Cost: is the amounts which we spent for the usage of resource in cloud computing and different cost are there for cloud provider and different user. These costs depend on the resource and usage.

$$Cost\ Total = \sum_{resource(i)} (C_i * T_i) \tag{9}$$

where C_i present the cast of resource and I present the unit time and T_i is the utilization of time.

Energy: is the strength of energy which required for cloud data center on that based cloud computing start working and different performance are performed based on these energy. Cloud data center used different kind of energy system.

$$Energy\ Tool = \sum Resource_{(i)} \int_{Start\ time}^{Finish\ time} E_i(FT) \tag{10}$$

where E_i represent the energy is consumed by the different resource and I present the starting and finished time. **Throughput:** is the process in which total number of tasks are execute successfully within given time period in cloud data center is known as throughput.

$$Throughput = \sum_{Task(i)} (Exe_{Time}) \tag{11}$$

Workload: present the ability of data center or processor work is known as workload for calculating of world load in cloud data center.

$$Degree\ of\ imbalance = \frac{MaxTask(Exe_{Time}) - MinTask_{ExeTime})}{AvgTask(i)(Exe_{Time})} \tag{12}$$

Utilization: presents the total amount of resource which is actually consumed in data center during network execution. The main objective is that we reduce this utilization of resource and improve the network life time. (Neghabi et al. 2020; Johansson et al. 2004).

$$Utilization = \frac{\sum Task(i)(Exe_{Time})}{Max_{Task(i)}(Ex_{Time})} \tag{13}$$

7 Simulation environment

Cloud computing is getting widely attention due to its dynamic nature and flexibility. Due to these property organizations and researcher are taking an interest in this technology. And implement by evaluating different data for different experimental purpose. Hence it becomes very expensive in real-time simulation therefore different simulation tools are

used. Figure 14 shows the different simulation software which are used in cloud computing for different simulation purpose.

Figure 14 shows the different simulation tool which I found during the study of different paper but most of the researcher used general cloud modeling approach and data center providing approach software tools in their simulation processes.

8 Result and discussion

In this section we present the summary of 106 papers which are used for load balancing purpose for resource and task from 2015 and 2021. The summary of these paper consist of technique name, method, year, advantage, disadvantage and references. Table 4 shows the summary of those papers which address the problems and approaches that are used to solve those issues taking different parameter.

Table 4 present the summary of 106 papers from 2015 to 2021 during the summarization steps it seem that most of the research work took place 2017 and 2020 because more number of paper published these two year as mention in Fig. 17. After the study of those papers it seems that most of the research focuses on to improve the makespan and request property level. Figure 15 show the number of parameter that different researcher try to improve these parameter by design their own method and techniques.

Figure 15 present the parameter name and how may time different research try to improve theses parameter. According to the result more of the researches focus on makespan, power consumption, QoS, response time. After the study of these papers it seems few parameters are not focusing properly they need more time and discussion. Table 4 presents the summary of those sections of VM where it needs change or where it section modified these all are mention in Table 5.

Table 5 show those section of VM where the modification took place with the help of which algorithm after the improvement of those section which parameter improve all these information mention in Table 5. Figure 16 shows those sections which are modified by different research for improvement in VM task allocation system for cloud computing.

Figure 16 present those section which are modified by different research during the work form the result most number of researcher work architecture section then, service level, VM migration section as so on but few researcher focus on VM replacements system. Figure 17 present the total number of paper published per year wise.

Figure 17 show the total number of downloaded from 2015 to 2021 must number paper used in this paper was 2017 and 2020. It means load balancing approach technique most paper published these two year. Figure 18 show those techniques which are mention by less number of researches and these techniques are important for future research.

Figure 18 show those sections of cloud computing which need more attentions like behavior of system, VM replacement policy, and future load prediction 5G network connection, integration with other technology load section, attacker policy and auto scaling. These all section can be improve with the help of load balancing approach but in this review paper we find few number of researcher attempt to solve these issues. These are important and future research section therefore researcher need attentions to this section in the coming time.

9 Summary

Cloud datacenter consist of different physical and virtual resource as we know that hardware resources configurations are remain the same darning communication. Where when user request for resources in cloud datacenter for this process service broker policy are used. Different algorithm work under these servicer broker policies in cloud datacenter and they provide better result. Applying predication algorithm for cloud computing are recent study where different researcher applying different machine learning algorithm (ML) used for forecasting. One of the main approaches they are using is VM migration approach and adapted for the cloud DC. Where applying forecasting in VM migration improve the data classification therefore it improve the load banning approaches. Different researcher implement ML algorithm for predication like CPU, RM, and load balancing approach in cloud data and provide better result. ML Algorithm also applied in VM section as local user agent for predication approach these approach was help full in the load predication in different area like CPU, RM and storage. For improvement in VM resource allocation host load detection method used in cloud datacenter due to this approach the performance of cloud computing improve as infrastructures purpose different researcher work on this section. Researcher also work on VM selection policy as well as service level agreement (SLA) these all activity are performed for improvement in cloud computing.

9.1 Analysis

According to the literature, we consider that most of the problems of VM task allocation system for cloud computing are solved using ML algorithm and different models. For improvement in cloud computing different section of VM are improved like resource allocation, data distribution, quality of services, VM migration, and VM replacement and server broker policy. For the improvement of above parameters different technique are used and load balancing is one of them. So taking load balancing approaches 70% VM migration section improve 20% server broker policy are improved and 10% resource allocation are improved in the give related work.

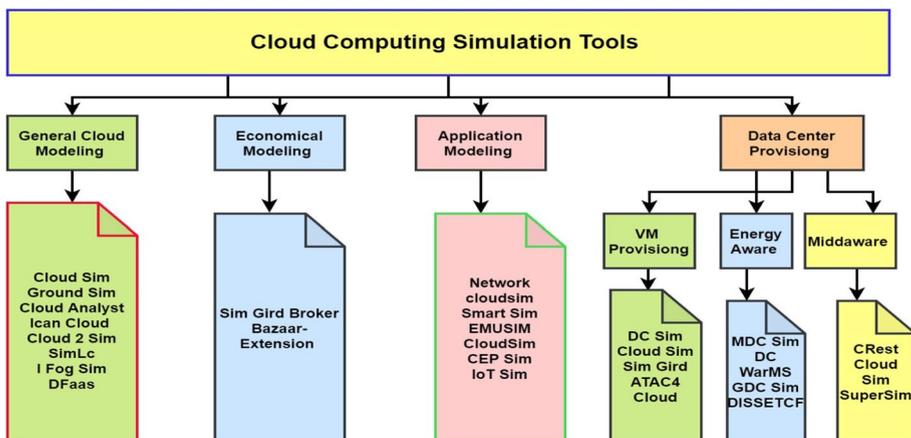


Fig. 14 Different simulation tools

Table 4 Summary of selected paper

Technique	Problems addressed	Improvements	Weakness	References
ACO-VMM-Algorithm	Overload section in VM	Local migration agent	Need improvement in Energy section	Wen et al. (2015)
VM-Allocation Theory	Minimize attacker' possibility	VM allocation policies	Not improve the mixture policy	Han (2015)
Bee-Colony-Algorithm	Load balancing	Priority rule used for minimizes the total processing time	Energy, cost	Zuo et al. (2015)
Stewers-Algorithm	Future load prediction	Energy consumption	Utilization, Waiting time	Balouek-Thomert (2016)
VM Migration Algorithms	Examine the migration times of VMs	Throughput	Energy	Chowdhury et al. (2015)
STVMLB Algorithm	Load balancing technique	Quality of service, substantially	Utilization	Tyagi and Kumar (2015)
MLB-RAM Algorithm	Overall system efficiency	QoS, cost, system and network	Accuracy	Othman et al. (2015)
Honey	Proposer distinction of Virtual Machine	Makespan, degree of imbalance, resource utilization	QoS	Abdulhamid et al. (2015)
Bee Galvanizing Algorithm	Distribution of dynamic workload	QoS, minimizing the make span	Energy	Dam et al. (2015)
GA-GEL-algorithm	Behavior of system	Response time, makespan	Utilization	Abdulhamid et al. (2015)
EAMLB- Algorithm	Load balancing technique	Stability, resource utilization	QoS	Gao and Wu (2015)
Hybrid Approach	Service level agreement	Energy consumption	Need Improvement In accuracy	Momil and Rahman (2016)
DWOLB-Algorithm	Load Balancing technique	Waiting time, resource utilization, throughput	Energy	Zhou et al. (2016)
DLBPR- Algorithm	VM section	Energy	QoS	Rajput and Kushwah (2016)
EFOALB- Algorithm	Task scheduling Algorithm	Makespan, utilization	QoS	Cassidy (2016)
LBM-M-Algorithm	VM migration	VM migrations, execution time	Utilization	Renugadevi and Mala (2015)
Genetic-Algorithm	Service Level/Agreement Services	Energy consumption, VM allocation	QoS	Bao et al. (2016)
Prediction based proactive load balancing approach	virtual machine allocation	Response time, request priority	Utilization	Bozakov (2016)
TVRSM-Model	VM load balancing	Execution time, response time, makespan	QoS	Abdulhamid et al. (2016)
RPQ-A-Algorithm	VM migration	Quality of service	Utilization	Khan and Ahmad (2017)
LBA_HB- Algorithm	VM migration	Resources utilization	QoS	Madni et al. (2017a, b)
Least cost per Connection Algorithm	Task scheduling	QoS, makespan	Accuracy	
SDN-Algorithm	NP problem scheduling	Makespan, response time	Utilization	Dinh et al. (2017)
EBC- Algorithm				
IPSO -Algorithm				

Table 4 (continued)

Technique	Problems addressed	Improvements	Weakness	References
LBDA-Algorithm	VM scheduling	Makespan, response time, execution time	QoS	Chen et al. (2017)
LBA-Algorithm	VM load balancing	Makespan, horizontal scalability, average resource utilization ratio	QoS	Carrion et al. (2017)
LB-ACO-Algorithm	Load balancing NP-problem	Makespan	Utilization	Subramanian and Abdulrahman (2017)
ACO-Algorithm	Task scheduling strategy	Makespan, total costs	Utilization	Subramanian and Abdulrahman (2017)
Service level Agreement	VM placement policy	Energy consumption	Utilization	Mevada et al. (2017)
RR-Algorithm	Architecture	Waiting time, response time, resource	QoS	Elmogy et al. (2017)
Hybrid GA-PSO Algorithm	Workflow technology	Makespan execution cost	QoS	Zhou and Yao (2017)
SVLL-Algorithm	Distributing tasks	Waiting time, total finish time	QoS	Vargas, (2017)
Genetic-Algorithm	VM scheduling	Data allocation	Utilization	Eswaraprasad and Raja (2017)
RM-Algorithm	VM scheduling	Execution time	QoS	Guo and Xue (2017)
VM Placement technique	VM Replacement	Lifetime	Creates fragments	Macias et al. (2017)
Placement of the VM	Taxonomy	Available resources, power	Utilization	Kaur et al. (2017a, b)
Hybrid HBB-LB-Algorithm	Resource allocation	Task cost, Speed, Energy consumption	QoS	Balusamy et al. (2017)
Enhanced Throttled Load Balancing	Scheduling	Response time, data Processing time, cost analysis	QoS	Ghoni, et al. (2017)
EMM-Algorithm	Resources scheduling	Makespan, cost	Utilization	Kaur et al. (2017)
Dynamic Threshold algorithm	CPU utilization	Resource utilization and time	QoS	Fard et al. (2017a, b)
Genetic-Algorithm	Load balancing NP	Response time	QoS	Roy (2017)
PFTF-Algorithm	Load balancing	Fault tolerance, virtual machine migration	QoS	Musumeci et al. (2017)
Naive Bayesian classifier	Task scheduling	Makespan	utilization	Ebadifard and Babamir (2017)
VM-Placement	Service level agreement	Utilized efficiently	Accuracy	Addya et al. (2017)
Hybrid technique	Load balancing	Response time, Machine cost	QoS	Kaur et al. (2017)
EG-Algorithm	VM scheduling	Makespan	utilization	Ajijum and Patil (2017a, b)
THR_MUG-Algorithm	VM scheduling	Reduce the number of VM migrations, energy consumption	Utilization	Wu et al. (2017)

Table 4 (continued)

Technique	Problems addressed	Improvements	Weakness	References
BFB- Algorithm	VM Allocation	Energy consumption	QoS	Bhatti (2017)
Fuzzy based Policy	Brokerage strategy	Service broker policy	Utilization	Islam and Waheed (2017)
SLA-Algorithm	Service level agreements	QoS	Utilization	Nawaz et al. (2018)
LB-ACO-Algorithm	Load balancing	Makespan	QoS	Belgacem et al. (2018)
WQBLB- Algorithm	5G networks	Queue size, total service	QoS	Dighriri et al. (2018)
Graph-based Mathematical model,	Fault tolerance	Utilization of resources	Utilization	Babu (2018)
TMA-Algorithm	VM Load Balancer section	Response times, processing time	Utilization	Kotsubanska and Sokolovska (2018)
Hybrid BLB- PSOGSA-Algorithm	Scheduling	Average VM processing speed, VM processing power	Dependent tasks	Manasrah and Ba Ali (2018)
Hybrid PSO-SA-Algorithm	VM scheduling	Response time, Processing time and cost	QoS	Zihu et al. (2018)
Adaptive Energy-Aware Algorithms	Service-level agreements	Reduce energy consumption	Utilization	Yadav et al. (2018)
BCO-algorithm	Iteration process	Makespan, imbalance degree	Utilization	Belgacem et al. (2018)
IMDLB- Algorithm	VM Scheduling	QoS	Utilization	Afzal and Kavitha (2018)
Hybrid technique	Service broker policies	Response time, response time	Utilization	Yasmeen et al. (2018)
WAMLB- Algorithm	Load balancing	Weight factor and assignment section	Utilization	Singh and Prakash (2018)
Graph theoretic	Load monitoring	Migration cost, time taken for VM migrations	QoS	Devi et al. (2018)
Throttled Algorithm	VM scheduling	Overall response time, request servicing, data center loading	Utilization	Ramadhan et al. (2018)
Fuzzy Load Balancer	Load balancer section	Fault tolerance	Utilization	Rathore (2018)
ICT	VM	Makespan	QoS	Nasr et al. (2018)
DSP-Policy	Integration of cloud and fog	Response time, requests servicing time,	Utilization	Fatima et al. (2018)
SCLBA)	VM migration	CPU utilization	QoS	Ramesh and Dey (2018)
Hybrid VM Migration technique	Pre-copy	Quality of service	Utilization	Anu and Elizabeth (2019)
Multidimensional Queuing Load Optimization algorithm	VM	Resource scheduling efficiency	QoS	Priya et al. (2019)

Table 4 (continued)

Technique	Problems addressed	Improvements	Weakness	References
OPH-LB-Algorithm	VM	Utilization of resources, throughput, makespan	Utilization	Ly et al. (2019)
DVFS-Algorithm	VM	Utilization rates of server	Utilization	Abro et al. (2019)
MMSIA- algorithm	Architecture	Utilization	QoS	Safiri et al. (2019)
ABC- algorithm	Resource utilization	Makespan	QoS	Thanka et al. (2019)
ETLB-algorithm	VM scheduling	Response Time	Utilization	Al-Rahayfeh et al. (2019)
Hybrid ABPS-Algorithm	Scheduling	Makespan, cost outperforms	Utilization	Tamimnia et al. (2020)
MEMA-Technique	VM dataallocation	Security,quality of service, accessibility	Accuracy	Dibaj et al. (2020a)
ALD—algorithm	Utilization of auto scaling	Quality of service	Energy	Rajput and Goyal (2020)
SJF-QMW- Algorithm	VM scheduling	Throughput, hosting ratio	QoS	Zhang and Abnoosian(2020)
MEMA-Technique	Brokerage strategy	Security, capacity, quality of service, cost	Utilization	Dibaj (2020b)
Learning agent	Machine placement	Execution time, Number of HMs shutdown	Energy and delay	Ghasemi and Haghghat (2020)
MD-Algorithm	VM placement	Resource wastage, placement time	Energy and delay	Singh and Auluck (2020)
MPSO-Algorithm	Fitness function/pre-emptive VM	Makespan,	Utilization, Waiting	Agarwal et al. (2020)
MCCVA-Algorithm	Load balancing technique	QoS, makespan	Utilization, Waiting	Ranjan et al. (2020)
Federate Migration Based	Load Balancing Scheme	VM Migration, VM distribution communication cost	Efficiency	Najm and Tamarappalli(2020)
COA-Algorithm	Minimum Migration Time policy in VM	Reducing energy, resource utilization	QoS	Ahmad et al. (2020)
DLBA-Algorithm	Hypervisor scheduling controls	Makespan	QoS	Khan et al. (2020)
GWO—Algorithm	Load balancing	Makespan	Energy	Patel et al. (2020)
LBPR-Algorithm	Load balancing	System performance	QoS	Wang et al. (2020)
EBFD- Policy	VM Placement Policy	Total execution time, Decrease the VM placement failure rate	Utilization, Energy	Dubey et al. (2020)

Table 4 (continued)

Technique	Problems addressed	Improvements	Weakness	References
CBLB-Algorithm	Homogeneous, Heterogeneous scheduling	Make span, throughput, CPU utilization	Energy	Dubey et al. (2020)
CDCSO-Algorithm	VM scheduling	Energy, cost, time and optimal load balancing	QoS	Godman et al. (2020)
Simulated annealing	Two phases to balance the workload between the VMs	Quality of service task allocation	Utilization	Hanine and Benlahmar (2020)
D2B_CPU Based	Degree Balanced with CPU based VM allocation	QoS, reduces degree of imbalance, Waiting time of task	Energy	Joshi and Munisamy (2020)
Balls into Bins	VM replacement	Distributed of work	Energy	Si et al. (2020)
AI-Algorithm	Resource allocation	Quality of service, Platform for PowerDispatching	Energy	Lin et al. (2020)
VM Migration policy	VM section	QoS, minimal migration cost	Energy	Moghaddam et al. (2020)
DLB- Algorithm	VM section/server	machine life-time, power consumption	Energy	Renugadevi and Mala (2020)
LBOS-Algorithm	Resource utilization	Calculating weight procedure,	Energy QoS	Fernande et al. (2020)
Framework	VM Replacement	Energy consumed, VM migration	Utilization	Xu et al. (2020)
MILP- Model	Virtual Machines Placement	Power consumption, CPU Utilization, power savings	QoS	Alharbi et al. (2020)
HMRR-Algorithm	Task scheduling	Execution time, Makespan	Response time	Behrens et al. (2020)
SJF-QMW-algorithm	VM scheduling	Throughput, hosting ratio	QoS	Mallikarjuna(2020)
DLBA-Algorithm	VM	Makespan	QoS	Khorsand and Ramezanzpour (2020)
Tailoring of VM Size	Active physical servers	Energy-efficient	QoS	Hsieh et al. (2020)
Geometric Programming algorithm	Pre-copy migration strategy	Multi-VM migration	Educed energy consumption	Singh and Singh (2021)
QoS-DPSO algorithm	QoS scheduling model	QoS	Time consumption	Jing et al. (2021)
ML algorithm	Resource predication	C.PU:RAM	Error size improve	Jo and Yoo (2021)

9.2 Limitations

Although the use of different machine learning algorithm and other frame work model improve different section of VM for cloud computing but still limitations remain constraints which are as:

- i. Still need improvement in the services policies.
- ii. Still need improvement in section gent for predication.
- iii. Need to focus more on predication technique.
- iv. Need more focus on resource scheduling.
- v. Need more focus on live VM migration and VM replacement section.
- vi. Need improvement towards a cloud migration framework.
- vii. Need improvement toward a trusted framework for cloud computing

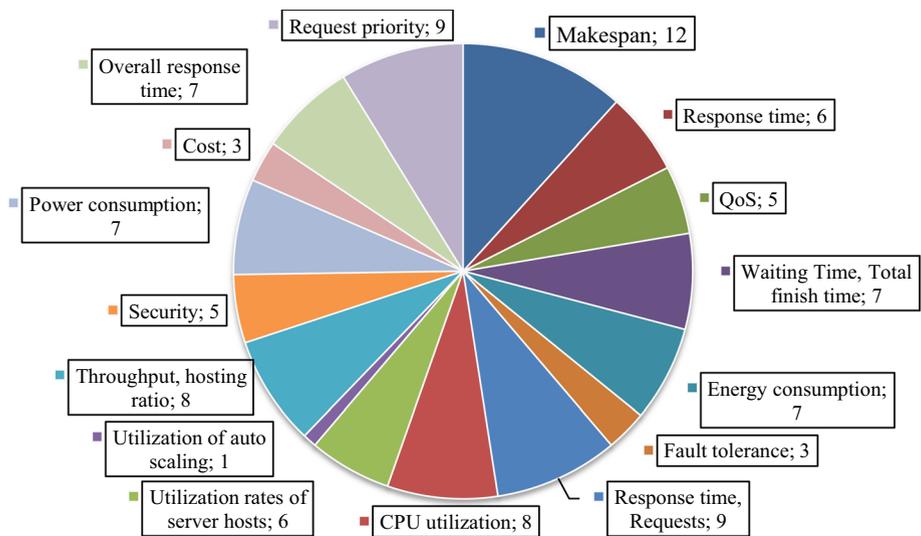


Fig. 15 Different researcher improve these parameters

Table 5 Result selection

Technique	Parameter	Section	Year
ACO-VMM	QoS, Number of Migrations, Load Condition Variance	VM/Live migration	2015
VM Allocation Theory	Minimize the efficiency, Safe rule from attacker	VM allocation policies	2015
Bee-Colony algorithm	QoS, Minimizes the total processing cost	VM/fitness value	2015
Skewers Algorithm	Energy consumption	Future load prediction in VM	2015
VMM-Algorithms	Throughput	Examine the migration times of VMs	2015
STVMLB algorithm	Quality of Service, substantially	VM/load balancing	2015
MQLB-RAM algorithm	QoS, Cost, System and network	Over all network	2015
HBA-Algorithm	Makespan, Degree of Imbalance, Resource Utilization	Proposer distinction of VM	2015
GA-GEL-Algorithm	QoS, minimizing the makespan	Distribution of workload	2015
EAMLB-Algorithm	Response time, Makespan	Behavior of system	2015
Hybrid Approach	Stability, Resource Utilization	Load balancing	2015
DWOLB- Algorithm	Energy	VM/Migration	2016
EFOALB-Algorithm	Energy	VM section	2016
LBMML-Algorithm	Makespan, Utilization	Task scheduling algorithm	2016
Genetic-Algorithm			
RPQ-Algorithm	Services	Response time, Request priority	2016
LBA_HB- Algorithm	VM allocation	Execution time, Response time	2016
LCPC- Algorithm	VM load balancing	Quality of Service	2016
SDN-Algorithm	VM migration	Resources utilization	2016
EBC-Algorithm	QoS Makespan	Task scheduling	2016
IPSO-Algorithm	Makespan, Response time	VM	2017
LBDA- Algorithm	Makespan, Response Time, Execution time	VM section	2017
LBA- Algorithm	Makespan time, Horizontal scalability, Average resource utilization ratio	VM section	2017
LB-ACO-Algorithm	Makespan	Load balancing NP-complete problem	2017
ACO-Algorithm	Makespan, Total costs	Task scheduling strategy	2017
Service level Agreement	Energy consumption	VM placement policy	2017

Table 5 (continued)

Technique	Parameter	Section	Year
RR- Algorithm	Waiting time, Response Time, resource	Architecture	2017
HybridGA-PSO Algorithm	Makespan Execution cost, Distributing tasks	Workflow technology	2017
SVLL-Algorithm	Data allocation	Waiting Time, Total finish time	2017
Generic algorithm	VM scheduling	VM scheduling	2017
RM-Algorithm	Creates fragments	Execution time	2017
VM Placement technique	Available resources, power	lifetime	2017
Placement of the VM	Task cost, Speed, Energy consumption	Taxonomy	2017
Hybrid HBB-LB-A Algorithm	Response time, Data Processing time, Cost analysis	Resource allocation	2017
Enhanced Throttled Load Balancing	Makespan, cost	Scheduling	2017
EMM- Algorithm	Resource utilization and time	Resources scheduling	2017
DT-Algorithm	Response time	CPU utilization	2017
Genetic- Algorithm	Fault tolerance, Virtual machine · Migration	Load balancing NP	2017
PFTF- Algorithm	Makespan	Load balancing	2017
Naive Bayesian classifier	Utilized efficiently	Task scheduling	2017
VM Placement	Load balancing	Service level agreement	2017
Hybrid Approach	VM	Machine cost, Response cost	2017
EG-Algorithm	VM scheduling	Make span	2017
THR_MUG-Algorithm	VM Allocation	Reduce the number of VM migrations, Energy consumption	2017
BFB-Algorithm	Brokerage strategy	Energy consumption	2017
Fuzzy based Policy	Service level agreements	Service broker policy	2017
SLA-Algorithm	Load balancing	QoS, Consumption	2017
LB-ACO-Algorithm	5G networks	Makespan	2017
LBWQB-Algorithm	Utilization of resources	Queue size, Total Service	2017
DBM- Model	Response times, processing time	Fault tolerance	2017
TMA—Algorithm		VM Load balancer section	2018

Table 5 (continued)

Technique	Parameter	Section	Year
Hybrid LB- LB-PSOGSA-Algorithm	Average VM processing speed, VM processing power	Balancing scheduling	2018
Hybrid PSO-SA-Algorithm	Response time, Processing time and cost	VM scheduling	2018
Adaptive Energy-Aware Algorithms	Reduce energy consumption	service-level agreements	2018
BCO-Algorithm	Makespan, imbalance degree	Iteration process	2018
IMDLB- Algorithm	QoS	VM scheduling	2018
Hybrid technique	Response time	Service broker policies	2018
WAMLB- Algorithm	Weight factor and assignmentsection	Load balancing	2018
Graph theoretic	Load monitoring	Migration cost, Time taken for VM migrations	2018
Throttled Algorithm	VM scheduling	Overall response time, Request Servicing, Data center loading	2018
Fuzzy Load Balancer	VM Allocation	Fault tolerance	2018
ICT	Load balancing	Makespan	2018
DSP-Policy	Integration of cloud and fog	Response time, Requests Servicing time	2018
SCLBA- Algorithm	VM migration	CPU utilization	2018
Hybrid VM Migration technique	Pre-copy	Quality of service	2018
Multidimensional Queuing Load Optimization Algorithm	VM	Resource scheduling efficiency	2018
OPH-LB-Algorithm	VM	Utilization of resources, Throughput,	2018
DVFS- Algorithm	VM	Utilization rates of server hosts	2018
MMSIA- Algorithm	Utilization,	Architecture	2019
ABC- algorithm	Makespan	Resource utilization	2019
ETLB-Algorithm	Response Time	VM scheduling	2019
Hybrid -ABPS-Algorithm	Makespan, cost outperforms	Scheduling	2019
MEMA-Technique	Security, Quality of service, Accessibility	VM data allocation	2019
ALD-Algorithm	Quality of Service	Utilization of auto scaling	2019
SJF-QMW-Algorithm	VM scheduling	Throughput, hosting ratio	2019

Table 5 (continued)

Technique	Parameter	Section	Year
MEMA- Technique	Brokerage strategy	Security, Capacity, quality of service, cost	2019
learning Agent	Execution time, Number of HMs shutdown, Immigrations cost	VM replacement	2020
GWO- Algorithm	Makespan	VM load balancing	2020
LBPR-Algorithm	System performance	System Load Balancing	2020
EBFD-Policy	Total execution time, decrease the VM placement failure rate	VM Replacement	2020
CBLB-Algorithm	Make span, Throughput, CPU utilization	Homogeneous, Heterogeneous scheduling	2020
CDCSO- algorithm	VM utilization	Energy Cost time and optimal load balancing	2020
Simulated Annealing	Quality of service task allocation	two phases to balance the workload between the VMs	2020
D2B_CPU Based	QoS, reduces degree of imbalance, waiting time of task	Degree Balanced with CPU based VM allocation	2020
Balls into Bins	Distributed of work	VM Replacement	2020
AI-Algorithm	Quality of service, Platform for Power Dispatching	Resource Allocation	2020
DLB-Algorithm	Machine life-time, power consumption	VM Section/Server	2020
LBOS-Algorithm	Calculating weight procedure,	VM Section/Server	2020
Framework	Energy consumed, VM migration	VM Replacement	2020
MILP- Model	VM, Replacement	Power consumption, CPU utilization, Power savings	2020
HMMRR-Algorithm	Execution time, Makespan response time	Task scheduling	2020
SJF-QMW- Algorithm	VM scheduling	Throughput, Hosting ratio	2020
DLBA-Algorithm	VM	Makespan	2020
Tailoring of VM Size	Active physical servers	Energy-efficient	2020
MD-Algorithm	Resource wastage	Placement Time/pre-emptive VM	2020
MPSO-Algorithm	Fitness function/pre-emptive VM	Pre-emptive VM	2020
MCCVA-Algorithm	QoS, Makespan	VM Section	2020
Federate Migration Based	VM replacement	VM Migration, VM Distribution Communication Cost	2020

Table 5 (continued)

Technique	Parameter	Section	Year
COA- Algorithm	Reducing Energy, Resource utilization	Migration Time policy in VM	2020
DLBA-Algorithm	Makespan	Balances the load in VM	2020
HMRR-Algorithm	Task scheduling	Execution time, Makespan Response time	2020
SJF-QMW-Algorithm	VM scheduling	Throughput, Hosting ratio	2020
DLBA-Algorithm	VM	Makespan	2020
Tailoring of VM Size	Active physical servers	Energy-Efficient	2020
Geometric Programming algorithm	Pre-copy migration strategy	Multi-VM migration	2021
QoS-DPSO algorithm	QoS scheduling model	QoS	2021
ML algorithm	Resource predication	C.P.U.RAM	2021

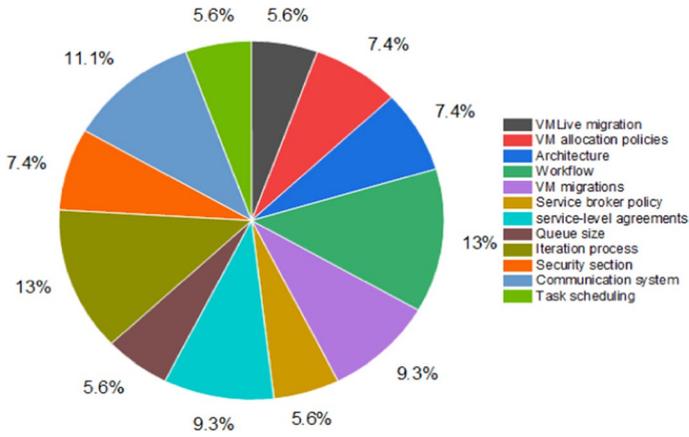


Fig. 16 Ratio of different parameter improved

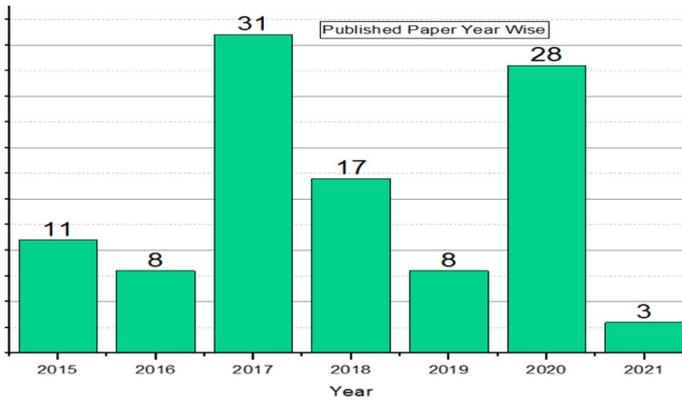


Fig. 17 Paper published per year

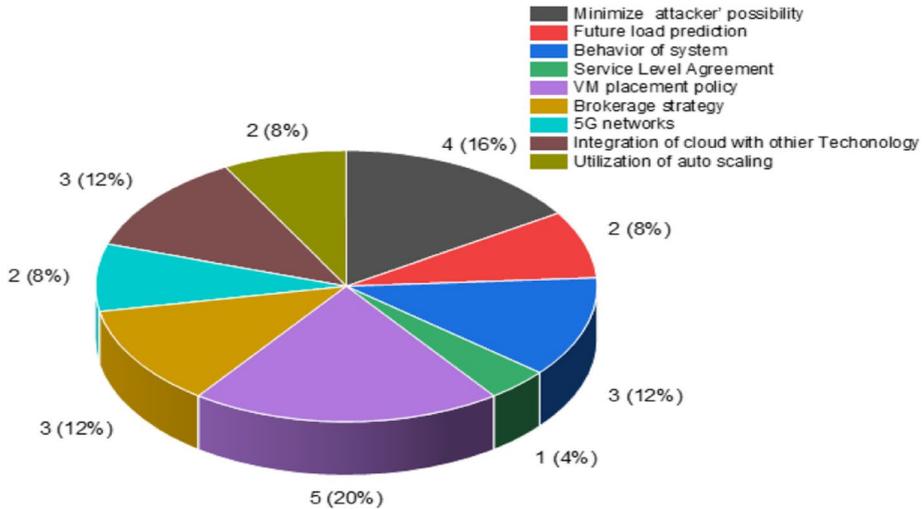


Fig. 18 Suggestions for future research

10 Conclusion

Task and resource distribution for VM is prime changing task in cloud computing therefore it getting more attention for researchers. This review paper present a state of the art survey about the technique, parameter, rule and specification that are used for load balancing in VM for cloud datacenter. As we know that different parameter are used to check the performance of load balancing approach in cloud data center. For improvement of these parameter different kind loads balancing policy and method are used in different section of VM's. All those technique and parameter are discussed which are used in different research paper and in future research direction. We mention those issue which are not given time by many research and some of them very important because cloud technology now merge with modern technology like IoT, senior network and 5G therefore now it time we work more to those issue which we face due to technology integration and energy. This review paper provide information to those researcher who want to start working in cloud computing.

Authors' contributions It Collaboration of all author and authors read and approved the final manuscript.

Declarations

Competing interest The authors declare that they have no competing interests.

References

Abd Elaziz M, Attiya I (2020) An improved Henry gas solubility optimization algorithm for task scheduling in cloud computing. *Artif Intell Rev*, pp 1–39.

- Abd Elaziz M, Shehabeldeen TA, Elsheikh AH, Zhou J, Ewees AA, Alqaness MA (2020) Utilization of Random Vector Functional Link integrated with Marine Predators Algorithm for tensile behavior prediction of dissimilar friction stir welded aluminum alloy joints. *J Mater Res Technol* 9(5):11370–11381
- Abdulhamid SM, Latiff MSA, Idris I (2015) Tasks scheduling technique using league championship algorithm for makespan minimization in IAAS cloud. *arXiv preprint arXiv:1510.03173*.
- Abdulhamid SIM, Abd Latiff MS, Abdul-Salaam G, Hussain Madni SH (2016) Secure scientific applications scheduling technique for cloud computing environment using global league championship algorithm. *PloS one* 11(7):e0158102.
- Abramson D, Jackson J, Muthrasanallur S, Neiger G, Regnier G, Sankaran R, Schoinas I, Uhlig R, Vembu B, Wiegert J (2006) Intel virtualization technology for Directed I/O. *Intel Technol J* 10(3).
- Abro A, Deng Z, Memon KA, Laghari AA, Mohammadani KH (2019) A dynamic application-partitioning algorithm with improved offloading mechanism for fog cloud networks. *Future Internet* 11(7):141
- Addya SK, Turuk AK, Sahoo B, Sarkar M, Biswash SK (2017) Simulated annealing based VM placement strategy to maximize the profit for Cloud Service Providers. *Eng Sci Technol Int J* 20(4):1249–1259
- Afzal S, Kavitha G (2018) Optimization of task migration cost in infrastructure cloud computing using IMDLB algorithm. In: 2018 International conference on circuits and systems in digital enterprise technology (ICCSDET). IEEE, New York, pp. 1–6
- Agarwal R, Baghel N, Khan MA (2020) Load balancing in cloud computing using mutation based particle swarm optimization. In: 2020 International Conference on Contemporary Computing and Applications (IC3A). IEEE, New York, pp. 191–195.
- Ahmad RW, Gani A, Hamid SHA, Shiraz M, Xia F, Madani SA (2015) Virtual machine migration in cloud data centers: a review, taxonomy, and open research issues. *J Supercomput* 71(7):2473–2515
- Ahmad N, Qamar S, Khan N, Naim A, Hussain MR, Naveed QN, Mahmood MR (2020) Cloud computing trends and cloud migration tuple. In: *Innovations in Electronics and Communication Engineering*. Springer, Singapore, pp. 737–745
- Alharbi HA, Elgorashi TE, Elmirghani JM (2020) Energy efficient virtual machines placement over cloud-fog network architecture. *IEEE Access* 8:94697–94718
- Al-Hashimi M, Al-Nidawi WJ, Othman M, Shakir M, Sulaiman H (2019) Evaluate information security governance frameworks in cloud computing environment using main and sub criteria. *J Comput Theor Nanosci* 16(3):996–1006
- Al-Karaki JN, Kamal AE (2004) Routing techniques in wireless sensor networks: a survey. *IEEE Wirel Commun* 11(6):6–28
- Al-Rahayfeh A, Atiewi S, Abuhussein A, Almiani M (2019) Novel approach to task scheduling and load balancing using the dominant sequence clustering and mean shift clustering algorithms. *Future Internet* 11(5):109
- Anjum A, Patil R (2017) Load balancing for cloud ecosystem using energy aware application scaling methodologies. *Int Res J Eng Technol* 4(5):479–482
- Anu VR, Elizabeth S (2019) IALM: Interference Aware Live Migration Strategy for Virtual Machines in Cloud Data Centres. In: *Data management, analytics and innovation*. Springer, Singapore, pp 499–511.
- Babu KR (2018) Synchronized Multi-Load Balancer with Fault Tolerance in Cloud. *arXiv preprint arXiv:1811.01319*.
- Babu KR, Samuel P (2016) Enhanced bee colony algorithm for efficient load balancing and scheduling in cloud. In: *Innovations in bio-inspired computing and applications*. Springer, Cham, pp 67–78
- Babu KR, Saranya S, Samuel P (2016) Secure cloud multi-tenant applications with cache in PaaS. In: *Innovations in bio-inspired computing and applications*. Springer, Cham, pp 15–27
- Balouek-Thomert D (2016) Scheduling on Clouds considering energy consumption and performance trade-offs: from modelization to industrial applications (Doctoral dissertation).
- Balusamy B, Karthikeyan K, Sangaiah AK (2017) Ant colony-based load balancing and fault recovery for cloud computing environment. *Int J Adv Intell Paradigms* 9(2–3):204–219
- Bamgbade JA, Nawi MNM, Kamaruddeen AM, Adeleke AQ, Salimon MG (2019) Building sustainability in the construction industry through firm capabilities, technology and business innovativeness: empirical evidence from Malaysia. *Int J Constr Manag*, 1–16
- Bao W, Hong C, Chunduri S, Krishnamoorthy S, Pouchet LN, Rastello F, Sadayappan P (2016) Static and dynamic frequency scaling on multicore CPUs. *ACM Trans Architecture Code Optim (TACO)* 13(4):1–26
- Behrens JK, Stepanova K, Babuska R (2020) Simultaneous task allocation and motion scheduling for complex tasks executed by multiple robots. In: 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, New York, pp. 11443–11449.

- Belgacem A, Beghdad-Bey K, Nacer H (2018) Task scheduling in cloud computing environment: a comprehensive analysis. In: International conference on computer science and its applications (pp. 14–26). Springer, Cham.
- Beloglazov A, Buyya R (2015) OpenStack Neat: a framework for dynamic and energy-efficient consolidation of virtual machines in OpenStack clouds. *Concurr Comput Pract Experience* 27(5):1310–1333
- Beloglazo A, Buyya R, Lee YC, Zomaya A (2011) A taxonomy and survey of energy-efficient data centers and cloud computing systems. In: *Advances in computers*, Vol 82. Elsevier, Amsterdam, pp. 47–111.
- Besta M, Stanojevic D, Licht JDF, Ben-Nun T, Hoefler T (2019) Graph processing on fpgas: Taxonomy, survey, challenges. arXiv preprint [arXiv:1903.06697](https://arxiv.org/abs/1903.06697).
- Bhandia P, Anupindi RS, Yekbote P, Singh N, Phalachandra HL, Sitaram D (2019) DCSim: cooling energy aware VM allocation framework for a cloud data center. In: 2019 International Conference on Advances in Computing and Communication Engineering (ICACCE). IEEE, New York, pp. 1–6
- Bhatti Z (2017) Model-based safety assessment of industrial automation systems using IEC 61499 (Doctoral dissertation, ResearchSpace@ Auckland).
- Bozakov Z (2016) Architectures for virtualization and performance evaluation in software defined networks (Doctoral dissertation, Hannover: Gottfried Wilhelm Leibniz Universität Hannover).
- Buyya R, Srirama SN, Casale G, Calheiros R, Simmhan Y, Varghese B, Gelenbe E, Javadi B, Miguel Vaquero L, Netto MAS, Toosi AN, Rodriguez MA, Llorente IM, Vimercati DCD, Samarati P, Milojicic D, Varela C, Bahsoon R, De Assuncao MD, Rana O, Zhou W, Jin H, Gentsch W, Zomaya AY, Shen H, (2018) A manifesto for future generation cloud computing: research directions for the next decade. *ACM Comput Surv (CSUR)* 51(5):1–38
- Carrión A, Caballer M, Blanquer I, Kotowski N, Jardim R, Dávila AMR (2017) Managing workflows on top of a cloud computing orchestrator for using heterogeneous environments on e-Science. *Int J Web Grid Serv* 13(4):375–402
- Cassidy A (2016) A practical guide to planning for E-business success: how to Enable your enterprise. CRC Press
- Chang WY, Abu-Amara H, Sanford JF (2010) Transforming enterprise cloud services. Springer Science & Business Media, Cham
- Chen J, Xu J, Hui B (2017) Cloud computing resource scheduling based on improved semantic search engine. In: Proceedings of the 2nd international conference on intelligent information processing, pp 1–5.
- Choudhary A, Govil MC, Singh G, Awasthi LK, Pilli ES, Kapil D (2017) A critical survey of live virtual machine migration techniques. *J Cloud Comput* 6(1):23
- Chowdhury MR, Mahmud MR, Rahman RM (2015) Implementation and performance analysis of various VM placement strategies in CloudSim. *J Cloud Comput* 4(1):20
- Dam S, Mandal G, Dasgupta K, Dutta P (2015) Genetic algorithm and gravitational emulation based hybrid load balancing strategy in cloud computing. In Proceedings of the 2015 third international conference on computer, communication, control and information technology (C3IT) (pp. 1–7). IEEE.
- Devi RK, Murugaboopathi G, Muthukannan M (2018) Load monitoring and system-traffic-aware live VM migration-based load balancing in cloud data center using graph theoretic solutions. *Clust Comput* 21(3):1623–1638
- Devi DC, Uthariaraj VR (2016) Load balancing in cloud computing environment using improved weighted round robin algorithm for nonpreemptive dependent tasks. *Sci World J*.
- Dibaj SR, Miri A, Mostafavi S (2020a) A cloud dynamic online double auction mechanism (DODAM) for sustainable pricing. *Telecommun Syst* 75(4):461–480
- Dibaj SR, Miri A, Mostafavi S (2020b) A cloud priority-based dynamic online double auction mechanism (PB-DODAM). *J Cloud Comput* 9(1):1–26
- Dighriri M, Lee G. M., & Baker, T. (2018). Measurement and classification of smart systems data traffic over 5G mobile networks. In *Technology for smart futures* (pp. 195–217). Springer, Cham.
- Dinh TQ, Tang J, La QD, Quek TQ (2017) Offloading in mobile edge computing: Task allocation and computational frequency scaling. *IEEE Trans Commun* 65(8):3571–3584
- Domanał SG, Reddy GRM (2014) Optimal load balancing in cloud computing by efficient utilization of virtual machines. In 2014 sixth international conference on communication systems and networks (COMSNETS). IEEE, pp 1–4
- Dubey K, Nasr AA, Sharma SC, El-Bahnasawy N, Attiya G, El-Sayed A (2020) Efficient VM placement policy for data centre in cloud environment. In *Soft Computing: Theories and Applications*. Springer, Singapore, pp. 301–309.
- Ebadifard F, Babamir SM (2017) Dynamic task scheduling in cloud computing based on Naïve Bayesian classifier. Department of Computer Engineering University of Kashan.

- Elmougy S, Sarhan S, Joundy M (2017) A novel hybrid of Shortest job first and round Robin with dynamic variable quantum time task scheduling technique. *J Cloud Comput* 6(1):1–12
- Elshehbiny S, Eldaydamony E, Alrahmawy M, Reyad AE (2018) An extended Intelligent Water Drops algorithm for workflow scheduling in cloud computing environment. *Egypt Inform J* 19(1):33–55
- Eswaraprasad R, Raja L (2017) A review of virtual machine (VM) resource scheduling algorithms in cloud computing environment. *J Stat Manag Syst* 20(4):703–711
- Eusuff MM, Lansey KE (2003) Optimization of water distribution network design using the shuffled frog leaping algorithm. *J Water Resour Plan Manag* 129(3):210–225
- Fahim M, Baker T, Khattak AM, Shah B, Aleem S, Chow F (2018) Context mining of sedentary behaviour for promoting self-awareness using a smartphone. *Sensors* 18(3):874
- Fard SYZ, Ahmadi MR, Adabi S (2017) A dynamic VM consolidation technique for QoS and energy consumption in cloud environment. *J Supercomput* 73(10):4347–4368
- Fatima I, Javaid N, Iqbal MN, Shafi I, Anjum A, Memon UU (2018) Integration of cloud and fog based environment for effective resource distribution in smart buildings. In: 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC). IEEE, New York, pp. 60–64.
- Fernandez PM, Cabello IM, De La Torre MAM (2020). U.S. Patent No. 10,778,851. Washington, DC: U.S. Patent and Trademark Office.
- Ferrer AJ, Marquès JM, Jorba J (2019) Towards the decentralised cloud: survey on approaches and challenges for mobile, ad hoc, and edge computing. *ACM Comput Surv (CSUR)* 51(6):1–36
- Gamal M, Rizk R, Mahdi H, Elnaghi BE (2019) Osmotic bio-inspired load balancing algorithm in cloud computing. *IEEE Access* 7:42735–42744
- Gao R, Wu J (2015) Dynamic load balancing strategy for cloud computing with ant colony optimization. *Future Internet* 7(4):465–483
- Ghasemi A, Haghghat AT (2020) A multi-objective load balancing algorithm for virtual machine placement in cloud data centers based on machine learning. *COMPUTING*.
- Ghomi EJ, Rahmani AM, Qader NN (2017) Load-balancing algorithms in cloud computing: a survey. *J Netw Comput Appl* 88:50–71
- Giret A, Trentesaux D, Prabhu V (2015) Sustainability in manufacturing operations scheduling: a state of the art review. *J Manuf Syst* 37:126–140
- Godman B, McCabe H, Leong TD, Mueller D, Martin AP, Hoxha I, Mwita JC, Rwegerera GM, Masseur A, de Oliveira CJ, Do Nascimento RCRM, de Lemos LLP, Tachkov K, Milushewa P, Patrick O, Niba LL, Laius O, Sefah I, Abdulsalim S, Soleymani F, Guantai AN, Achieng L, Oluka M, Jakupi A, Logviss K, Azmi Hassali M, Kibuule D, Kalemeera F, Mubita M, Fadare J, Ogunleye OO, Saleem Z, Hussain S, Bochenek T, Mardare I, Alrasheedy AA, Furst J, Tomek D, Markovic-Pekovic V, Rampamba EM, Alfadl A, Amu AA, Matsebula Z, Thi Phuong TN, Thanh BN, Kalungia AC, Zaranjika T, Masuka N, Oлару ID, Wale J, Hill R, Kurdi A, Timoney A, Campbell S, Meyer JC (2020) Fixed dose drug combinations—are they pharmaco-economically sound? Findings and implications especially for lower- and middle-income countries. *Expert Rev Pharmacoecon Outcomes Res* 20(1):1–26
- González-Martínez JA, Bote-Lorenzo ML, Gómez-Sánchez E, Cano-Parra R (2015) Cloud computing and education: A state-of-the-art survey. *Comput Educ* 80:132–151
- Guo P, Xue Z (2017) Cost-effective fault-tolerant scheduling algorithm for real-time tasks in cloud systems. In: 2017 IEEE 17th international conference on communication technology (ICCT). IEEE, pp. 1942–1946.
- Guo C, Lu, G., Wang, H. J., Yang, S., Kong, C., Sun, P., Wu W, & Zhang, Y. (2010). Secondnet: a data center network virtualization architecture with bandwidth guarantees. In: Proceedings of the 6th International Conference (pp. 1–12).
- Han Y (2015) Defending against co-resident attacks in cloud computing (Doctoral dissertation).
- Hanine M, Benlahmar EH (2020) A load-balancing approach using an improved simulated annealing algorithm. *J Inform Process Syst* 16(1):132–144
- Hashem IAT, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU (2015) The rise of “big data” on cloud computing: Review and open research issues. *Inf Syst* 47:98–115
- Holland MB, Shamer SZ, Imbach P, Zamora JC, Moreno CM, Hidalgo E, Donatti CI, Martínez-Rodríguez MR, Harvey CA (2017) Mapping adaptive capacity and smallholder agriculture: applying expert knowledge at the landscape scale. *Climatic Change* 141(1):139–153
- Hsieh SY, Liu CS, Buyya R, Zomaya AY (2020) Utilization-prediction-aware virtual machine consolidation approach for energy-efficient cloud data centers. *J Parallel Distributed Comput* 139:99–109
- Hussain K, Salleh MNM, Cheng S, Shi Y (2019) Metaheuristic research: a comprehensive survey. *Artif Intell Rev* 52(4):2191–2233
- Hwang J, Ramakrishnan KK, Wood T (2015) NetVM: High performance and flexible networking using virtualization on commodity platforms. *IEEE Trans Netw Serv Manage* 12(1):34–47

- Ibrahim IM (2021) Task scheduling algorithms in cloud computing: A review. *Turkish J Computer Math Education (TURCOMAT)* 12(4):1041–1053
- Imdad M, Jacob DW, Mahdin H, Baharum Z, Shaharudin SM, Azmi MS (2020) Internet of things (IoT): security requirements, attacks and counter measures. *Indonesian J Electrical Eng Comp Sci* 18(3):1520–1530
- Indukuri PSV (2016) Performance comparison of Linux containers (LXC) and OpenVZ during live migration. Doctoral dissertation, Master's thesis, Blekinge Institute of Technology, Sweden.
- Islam N, Waheed S (2017) Fuzzy based efficient service broker policy for cloud. *Int J Computer Appl* 168(4):37–40
- Jing W, Zhao C, Miao Q, Song H, Chen G (2021) QoS-DPSO: QoS-aware task scheduling for cloud computing system. *J Netw Syst Manage* 29(1):1–29
- Jo E, Yoo H (2021) Implementation of cloud monitoring system based on open source monitoring solution. In: *Software Engineering in IoT, Big Data, Cloud and Mobile Computing* (pp. 181–190). Springer, Cham.
- Johansson E, Nedstam J, Höst M, Wartenberg F (2004) Software Performance Aspects during Maintenance and Evolution of Embedded Platforms. In: *4th Conference on Software Engineering Research and Practise in Sweden (SERPS)*.
- Joshi AS, Munisamy SD (2020) Dynamic degree balanced with CPU based VM allocation policy for load balancing. *J Inf Optim Sci* 41(2):543–553
- Kaur A, Singh M, Singh P (2017a) A taxonomy, survey on placement of virtual machines in cloud. In: *2017 International conference on energy, communication, data analytics and soft computing (ICECDS)*. IEEE, pp 2054–2058.
- Kaur A, Kaur B, Singh D (2017b) Optimization techniques for resource provisioning and load balancing in cloud environment: a review. *Int J Inform Eng Electronic Bus* 9(1):28
- Khan RZ, Ahmad MO (2017) A survey on load balancing algorithms in cloud computing. *Int J Autonomic Comput* 2(4):366–383
- Khan S, Parkinson S, Qin Y (2017) Fog computing security: a review of current applications and security solutions. *J Cloud Comput* 6(1):19
- Khan SU, Haq IU, Rho S, Baik SW, Lee MY (2019) Cover the violence: a novel deep-learning-based approach towards violence-detection in movies. *Appl Sci* 9(22):4963
- Khorsand R, Ramezanpour M (2020) An energy-efficient task-scheduling algorithm based on a multi-criteria decision-making method in cloud computing. *Int J Commun Syst* 33(9):e4379.
- Kotha SK, Rani MS, Subedi B, Chunduru A, Karrothu A, Neupane B, Sathishkumar VE (2021) A comprehensive review on secure data sharing in cloud environment. *Wireless Personal Communications*, pp 1–28.
- Kotsubanska O, Sokolovska I (2018) Section 1 food technologies. In: *International scientific conference of young scientists and students. Youth Scientific Achievements to the 21st*, p. 4.
- Krishna A, Lal A, Mathewkutty AJ, Jacob DS, Hari M (2020) Intrusion detection and prevention system using deep learning. In: *2020 International conference on electronics and sustainable communication systems (ICESC)*. IEEE, pp 273–278
- Kulik J, Heinzelman W, Balakrishnan H (2002) Negotiation-based protocols for disseminating information in wireless sensor networks. *Wireless Netw* 8(2–3):169–185
- Kuyoro SO, Ibikunle F, Awodele O (2011) Cloud computing security issues and challenges. *Int J Comput Netw (IJCN)* 3(5):247–255
- Lee TD, Lee BM, Noh W (2018) Hierarchical cloud computing architecture for context-aware IoT services. *IEEE Trans Consum Electron* 64(2):222–230
- Lee M, Krishnakumar AS, Krishnan P, Singh N, Yajnik S (2010) Supporting soft real-time tasks in the xen hypervisor. In: *Proceedings of the 6th ACM SIGPLAN/SIGOPS international conference on Virtual execution environments*, pp 97–108.
- Lin CC, Su NW, Deng DJ, Tsai IH (2020) Resource allocation of simultaneous wireless information and power transmission of multi-beam solar power satellites in space–terrestrial integrated networks for 6G wireless systems. *Wireless Networks*.
- Liu F, Tong J, Mao J, Bohn R, Messina J, Badger L, Leaf D (2011) NIST cloud computing reference architecture. *NIST Spec Publ* 500(2011):1–28
- Lv L, Zhang Y, Li Y, Xu K, Wang D, Wang W, Li M, Cao X, Liang Q (2019) Communication-aware container placement and reassignment in large-scale internet data centers. *IEEE J Sel Areas Commun* 37(3):540–555
- Macias VM, Ohm JR, Rasgon JL (2017) Gene drive for mosquito control: where did it come from and where are we headed? *Int J Environ Res Public Health* 14(9):1006

- Madni SHH, Abd Latiff MS, Coulibaly Y (2017a) Recent advancements in resource allocation techniques for cloud computing environment: a systematic review. *Clust Comput* 20(3):2489–2533
- Madni SHH, Abd Latiff MS, Abdullahi M, Abdulhamid SIM, Usman MJ (2017b) Performance comparison of heuristic algorithms for task scheduling in IaaS cloud computing environment. *PLoS One* 12(5):e0176321.
- Mallikarjuna B (2020) Feedback-based fuzzy resource management in IoT-Based-Cloud. *Int J Fog Comput (IJFC) (IJFC)* 3(1):1–21
- Manasrah AM, Smadi T, Almomani, A. (2017) A variable service broker routing policy for data center selection in cloud analyst. *J King Saud Univ-Comput Inform Sci* 29(3):365–377
- Manasrah AM, Ba Ali H (2018) Workflow scheduling using hybrid GA-PSO algorithm in cloud computing. *Wireless Communications and Mobile Computing*, 2018.
- Mastelic T, Oleksiak A, Claussen H, Brandic I, Pierson JM, Vasilakos AV (2014) Cloud computing: survey on energy efficiency. *AcM Comput Surv (csur)* 47(2):1–36
- Mevada A, Patel H, Patel N (2017) Enhanced energy efficient virtual machine placement policy for load balancing in cloud environment. *Int J Curr Res Rev* 9:50–53
- Mezgár I, Rauschecker U (2014) The challenge of networked enterprises for cloud computing interoperability. *Comput Ind* 65(4):657–674
- Milani AS, Navimipour NJ (2016) Load balancing mechanisms and techniques in the cloud environments: Systematic literature review and future trends. *J Netw Comput Appl* 71:86–98
- Mishra SK, Sahoo B, Parida PP (2020) Load balancing in cloud computing: a big picture. *J King Saud Univ-Comput Inform Sci* 32(2):149–158
- Moghaddam SM, O’Sullivan M, Walker C, Piraghaj SF, Unsworth CP (2020) Embedding individualized machine learning prediction models for energy efficient VM consolidation within Cloud data centers. *Futur Gener Comput Syst* 106:221–233
- Monil MAH, Rahman RM (2016) VM consolidation approach based on heuristics, fuzzy logic, and migration control. *J Cloud Comput* 5(1):8
- Monika S, Jawahar M, Murugaraja SK (2015) Data sharing in cloud storage by RSA based encryption. *Int J Comput Sci Trends Technol*
- Mushtaq MF, Akram U, Khan I, Khan SN, Shahzad A, Ullah A (2017) Cloud computing environment and security challenges: a review. *Int J Adv Comput Sci Appl* 8(10):183–195
- Musumeci D, Amato J, Zizza P, Platella C, Cosconati S, Cingolani C, Biroccio A, Novellino E, Randazzo A, Giancola C, Pagano, B., Montesarchio D (2017). Tandem application of ligand-based virtual screening and G4-OAS assay to identify novel G-quadruplex-targeting chemotypes. *Biochim Biophys Acta (BBA)-General Subjects* 1861(5):1341–1352.
- Nadimi-Shahraki MH, Taghian S, Mirjalili S, Faris H (2020) MTDE: An effective multi-trial vector-based differential evolution algorithm and its applications for engineering design problems. *Appl Soft Comput* 97:106761.
- Najm M, Tamarapalli V (2020) VM migration for profit maximization in federated cloud data centers. In: 2020 International Conference on COMMunication Systems & NETWORKS (COMSNETS). IEEE, pp. 882–884.
- Nasim R, Kassler AJ (2014) Deploying openstack: virtual infrastructure or dedicated hardware. In: 2014 IEEE 38th international computer software and applications conference workshops. IEEE, pp. 84–89.
- Nasr AA, El-Bahnasawy NA, Attiya G, El-Sayed A (2018) A new online scheduling approach for enhancing QOS in cloud. *Future Comput Inform J* 3(2):424–435
- Navamani BA, Yue C, Zhou X (2018) Discover and secure (DaS): an automated virtual machine security management framework. In 2018 IEEE 37th International performance computing and communications conference (IPCCC) (pp. 1–6). IEEE
- Nawaz F, Janjua NK, Hussain OK, Hussain FK, Chang E, Saberi M (2018) Event-driven approach for predictive and proactive management of SLA violations in the Cloud of Things. *Futur Gener Comput Syst* 84:78–97
- Neghabi AA, Navimipour NJ, Hosseinzadeh M, Rezaee A (2020) Energy-aware dynamic-link load balancing method for a software-defined network using a multi-objective artificial bee colony algorithm and genetic operators. *IET Commun* 14(18):3284–3293
- Obasuyi GC, Sari A (2015) Security challenges of virtualization hypervisors in virtualized hardware environment. *Int J Commun Netw Syst Sci* 8(07):260
- Oke AE, Kineber AF, Albukhari I, Othman I, Kingsley C (2021) Assessment of cloud computing success factors for sustainable construction industry: the case of Nigeria. *Buildings* 11(2):36
- Othman, M, Khan AN, Abid SA, Madani SA (2015) MobiByte: an application development model for mobile cloud computing. *J Grid Comput* 13(4):605–628

- Ouhame S, Hadi Y, Arifullah A (2020) A hybrid grey wolf optimizer and artificial bee colony algorithm used for improvement in resource allocation system for cloud technology.
- Patel D, Patra MK, Sahoo B (2020) GWO Based task allocation for load balancing in containerized cloud. In: 2020 International Conference on Inventive Computation Technologies (ICICT). IEEE, pp. 655–659.
- Phanden RK, Jain A, Verma R (2011) Integration of process planning and scheduling: a state-of-the-art review. *Int J Comput Integr Manuf* 24(6):517–534
- Priya V, Kumar CS, Kannan R (2019) Resource scheduling algorithm with load balancing for cloud service provisioning. *Appl Soft Comput* 76:416–424
- Rajput RKS, Goyal D (2020) Auto-Scaling in the Cloud Environment. In *Cloud Computing Applications and Techniques for E-Commerce* (pp. 84–98). IGI Global.
- Rajput SS, Kushwah VS (2016) A genetic based improved load balanced min-min task scheduling algorithm for load balancing in cloud computing. In: 2016 8th international conference on Computational Intelligence and Communication Networks (CICN) (pp. 677–681). IEEE.
- Ramadhan G, Purboyo TW, Latuconsina R (2018) Experimental model for load balancing in cloud computing using throttled algorithm. *Int J Appl Eng Res* 13(2):1139–1143
- Ramesh D, Dey S (2018) SCLBA-CC: slot based carton load balancing approach for cloud environment. In: 2018 International conference on current trends towards converging technologies (ICCTCT) (pp. 1–5). IEEE.
- Rani TS, Kannan DS (2017) Task scheduling on virtual machines using bat strategy for efficient utilization of resources in cloud environment. 12:6663–6669
- Ranjan A, Kewat S, Singh B (2020) Reweighted zero attracting maximum correntropy criterion algorithm based solar grid interfaced system for alleviating power quality problems. In: 2020 IEEE 9th Power India International Conference (PIICON) (pp. 1–6). IEEE.
- Rathore N (2018) Performance of hybrid load balancing algorithm in distributed web server system. *Wireless Pers Commun* 101(3):1233–1246
- Renugadevi N, Mala C (2015) Improved group key agreement for emergency cognitive radio mobile Ad hoc networks. *J Wirel Mob Networks Ubiquitous Comput Dependable Appl* 6(3):73–86
- Renugadevi T, Geetha K, Prabakaran N, Siano P (2020) Carbon-efficient virtual machine placement based on dynamic voltage frequency scaling in Geo-Distributed cloud data centers. *Appl Sci* 10(8):2701
- Rimal BP, Choi E, Lumb I (2010) A taxonomy, survey, and issues of cloud computing ecosystems. In: *Cloud computing*. Springer, London, pp 21–46.
- Roy MP (2017) *Work-Life Balance and Police: Concept, Consequences and Strategies*. SVP National Police Academy, 105.
- Safitri C, Yamada Y, Goudarzi S, Mantoro T, Ayu MA (2019) Bio-Inspired Algorithm for QoS Optimization in Information-Centric Networking. In: 2019 5th International Conference on Computing Engineering and Design (ICCED). IEEE, pp. 1–4.
- Samimi P, Patel A (2011) Review of pricing models for grid & cloud computing. In: 2011 IEEE Symposium on Computers & Informatics. IEEE, pp. 634–639
- Shah PH (2011) Security in live Virtual Machine migration. Doctoral dissertation, Wichita State University.
- Sharma S (2015) Evolution of as-a-Service Era in Cloud. arXiv preprint arXiv:1507.00939
- Sharif MI, Lee W, Cui W, Lanzi A (2009) Secure in-vm monitoring using hardware virtualization. In: Proceedings of the 16th ACM conference on Computer and communications security (pp. 477–487).
- Sheikhholeslami F, Jafari Navimipour N (2018) Auction-based resource allocation mechanisms in the cloud environments: a review of the literature and reflection on future challenges. *Concurrency Comput Pract Exp* 30(16):e4456.
- Sheikhholeslami F, Navimipour NJ (2017) Service allocation in the cloud environments using multi-objective particle swarm optimization algorithm based on crowding distance. *Swarm Evol Comput* 35:53–64
- Si F, Tang B, Fang Z, Li H, Zhang S (2020). A new type of BaTiO₃-based ceramics with Bi (Mg^{1/2}Sn^{1/2}) O₃ modification showing improved energy storage properties and pulsed discharging performances. *J Alloys Compounds* 819:153004.
- Singh A, Auluck N (2020). Load balancing aware scheduling algorithms for fog networks. *Softw Pract Experience*, 50(11):2012–2030.
- Singh AN, Prakash S (2018) WAMLB: weighted active monitoring load balancing in cloud computing. In: *Big data analytics*. Springer, Singapore, pp. 677–685.
- Singh G, Singh AK (2021) Optimizing multi-VM migration by allocating transfer and compression rate using Geometric Programming. *Simulation Modelling Practice and Theory*, 106:102201.
- Smith JE, Nair R (2005) The architecture of virtual machines. *Computer* 38(5):32–38
- Subramanian N, Abdulrahman MD (2017) Logistics and cloud computing service providers' cooperation: a resilience perspective. *Prod Planning Control* 28(11–12):919–928

- Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, Brisco B (2020) Google earth engine for geo-big data applications: a meta-analysis and systematic review. *ISPRS J Photogrammetry Remote Sens* 164:152–170
- Taylor P (2013) Performance management and the new workplace tyranny: a report for the Scottish Trades Union Congress. Scottish Trades Union Congress.
- Thanka MR, Maheswari PU, Edwin EB (2019b) A hybrid algorithm for efficient task scheduling in cloud computing environment. *Int J Reason-Based Intell Syst* 11(2):134–140
- Tracz B, Mastehollet EP, Isaksen A (2019) AWS IoT-basert system for håndtering av kommunikasjon og kontroll av AGV. Bachelor's thesis, NTNU.
- Tyagi V, Kumar T (2015) ORT broker policy: reduce cost and response time using throttled load balancing algorithm. *Procedia Computer Science* 48:217–221
- Ullah A, Nawi NM, Khan MH (2020) BAT algorithm used for load balancing purpose in cloud computing: an overview. *Int J High Perform Comput Network* 16(1):43–54
- Ullah A, Nawi NM, Mahdin HB, Baseer S, Deris MM (2019) Role of Different Integer Virtual Machine in Cloud Data Center. *JOIV: Int J Informatics Vis* 3(4):394–398.
- Umar S, Baseer S (2016) Perception of cloud computing in universities of Peshawar, Pakistan. In: 2016 Sixth international conference on innovative computing technology (INTECH) (pp. 87–91). IEEE.
- Vargas A (2017) Approaching or avoiding value? Pavlovian biases in intertemporal choices.
- Wang GG, Gandomi AH, Alavi AH, Gong D (2019) A comprehensive review of krill herd algorithm: variants, hybrids and applications. *Artif Intell Rev* 51(1):119–148
- Wang Z, Wang N, Su X, Ge S (2020) An empirical study on business analytics affordances enhancing the management of cloud computing data security. *Int J Inf Manag* 50:387–394
- Wen WT, Wang CD, Wu DS, Xie YY (2015) An ACO-based scheduling strategy on load balancing in cloud computing environment. In: 2015 Ninth international conference on frontier of computer science and technology (pp. 364–369). IEEE.
- Wong J, Wang X, Li H, Chan G (2014) A review of cloud-based BIM technology in the construction sector. *J Inf technol constr* 19:281–291
- Wu X, Zeng Y, Lin G (2017) An energy efficient VM migration algorithm in data centers. In: 2017 16th International symposium on distributed computing and applications to business, engineering and science (DCABES) (pp. 27–30). IEEE.
- Xu X (2012) From cloud computing to cloud manufacturing. *Robot Comput Integr Manufact* 28(1):75–86
- Xu X, Zhang X, Khan M, Dou W, Xue S, Yu S (2020) A balanced virtual machine scheduling method for energy-performance trade-offs in cyber-physical cloud systems. *Futur Gener Comput Syst* 105:789–799
- Yadav R, Zhang W, Kaiwartya O, Singh PR, Elgendy IA, Tian YC (2018) Adaptive energy-aware algorithms for minimizing energy consumption and SLA violation in cloud computing. *IEEE Access* 6:55923–55936
- Yakhchi M, Ghafari SM, Yakhchi S, Fazeli M, Patooghi A (2015) Proposing a load balancing method based on Cuckoo Optimization Algorithm for energy management in cloud computing infrastructures. In: 2015 6th International conference on modeling, simulation, and applied optimization (ICMSAO). IEEE, pp. 1–5
- Yang CT, Liu JC, Chen ST, Huang KL (2017) Virtual machine management system based on the power saving algorithm in cloud. *J Netw Comput Appl* 80:165–180
- Yasmeen A, Javaid N, Rehman OU, Iftikhar H, Malik MF, Muhammad FJ (2018) Efficient resource provisioning for smart buildings utilizing fog and cloud based environment. In 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC) (pp. 811–816). IEEE.
- Zaki MJ, Li W, Parthasarathy S (1996) Customized dynamic load balancing for a network of workstations. In: Proceedings of 5th IEEE international symposium on high performance distributed computing (pp. 282–291). IEEE.
- Zarandi MHF, Asl AAS, Sotudian S, Castillo O (2020) A state of the art review of intelligent scheduling. *Artif Intell Rev* 53(1):501–593
- Zhang GH, Abnoosian K (2020) Scheduling mechanisms in the cloud environment: a methodological analysis. *Kybernetes*.
- Zhou Y (2020) Institutionalization and Identity of Contemporary Art (2000–Present). In *A History of Contemporary Chinese Art* (pp. 335–458). Springer, Singapore.
- Zhou W, Yang S, Fang J, Niu X, Song H (2010) Vmctune: A load balancing scheme for virtual machine cluster using dynamic resource allocation. In: 2010 Ninth International Conference on Grid and Cloud Computing (pp. 81–86). IEEE.
- Zhou Z, Hu Z, Li K (2016) Virtual machine placement algorithm for both energy-awareness and SLA violation reduction in cloud data centers. *Scientific Programming*, 2016.

- Zhou J, Yao X (2017) A hybrid artificial bee colony algorithm for optimal selection of QoS-based cloud manufacturing service composition. *Int J Adv Manufact Technol* 88(9–12):3371–3387
- Zhu J, Li X, Ruiz R, Xu X (2018) Scheduling stochastic multi-stage jobs to elastic hybrid cloud resources. *IEEE Trans Parallel Distrib Syst* 29(6):1401–1415
- Ziyath SPM, Senthilkumar S (2020) MHO: meta heuristic optimization applied task scheduling with load balancing technique for cloud infrastructure services. *J Ambient Intell Humanized Comput*, pp 1–10.
- Zuo L, Shu L, Dong S, Zhu C, Hara T (2015) A multi-objective optimization scheduling method based on the ant colony algorithm in cloud computing. *Ieee Access* 3:2687–2699

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.