



# Cloud Computing Virtualization of Resources Allocation for Distributed Systems

Hanan M. Shukur<sup>1</sup>, Subhi R. M. Zeebaree<sup>2</sup>, Rizgar R. Zebari<sup>3\*</sup>, Diyar Qader Zeebaree<sup>4</sup>, Omar M. Ahmed<sup>5</sup>, Azar Abid Salih<sup>6</sup>

<sup>1</sup>IT Department, Al Kitab University, Kirkuk – Iraq, hanan89md@gmail

<sup>2</sup>Presidency of Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq, subhi.rafeeq@dpu.edu.krd

<sup>3</sup>IT Department. Technical College of Informatics Akre, Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq, rizgar.ramadhan@dpu.edu.krd

<sup>4</sup>Presidency of Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq, dqszebaree@dpu.edu.krd

<sup>5</sup>IT Department, Zakho Technical Institute, Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq, omar.alzakholi@dpu.edu.krd

<sup>6</sup>Information Technology Management Department, Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq, azar.abid@dpu.edu.krd

\* Correspondence: rizgar.ramadhan@dpu.edu.krd

## Abstract

Cloud computing is a new technology which managed by a third party “cloud provider” to provide the clients with services anywhere, at any time, and under various circumstances. In order to provide clients with cloud resources and satisfy their needs, cloud computing employs virtualization and resource provisioning techniques. The process of providing clients with shared virtualized resources (hardware, software, and platform) is a big challenge for the cloud provider because of over-provision and under-provision problems. Therefore, this paper highlighted some proposed approaches and scheduling algorithms applied for resource allocation within cloud computing through virtualization in the datacenter. The paper also aims to explore the role of virtualization in providing resources effectively based on clients’ requirements. The results of these approaches showed that each proposed approach and scheduling algorithm has an obvious role in utilizing the shared resources of the cloud data center. The paper also explored that virtualization technique has a significant impact on enhancing the network performance, save the cost by reducing the number of Physical Machines (PM) in the datacenter, balance the load, conserve the server’s energy, and allocate resources actively thus satisfying the clients’ requirements. Based on our review, the availability of Virtual Machine (VM) resource and execution time of requests are the key factors to be considered in any optimal resource allocation algorithm. As a results of our analyzing for the proposed approaches is that the requests execution time and VM availability are main issues and should in consideration in any allocating resource approach.

**Keywords:** cloud computing, virtualization, resource allocation, distributed systems.

Received: April 02<sup>nd</sup>, 2020 / Accepted: June 29<sup>th</sup>, 2020 / Online: July 3<sup>rd</sup>, 2020

## I. INTRODUCTION

In recent years, cloud computing has occupied many trends in our life for instance government, enterprise, industry, business, and markets. Cloud computing is a kind of parallel and distributed system which involves groups of remote servers are networked to allow sharing of centralized data storage, data-processing tasks, and online access to Information Technology (IT) services or resources [1], [2]. Using cloud computing the user leases the resource instead of purchasing it, thus reducing software costs [3], [4]. Furthermore, cloud computing provides on-demand services, where clients could reach to the services at any time from remote locations. Instead of storing and retrieving data from the local computer, cloud computing allows utilizing the services of remote computer [5]–[8]. The clients use the cloud services rather than working on their infrastructure, it means that the users don’t need knowledge about the infrastructure of the network [9]–[11]. Virtualized

computing resources could be delivered to the clients in three types of services: Infrastructure as a service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) [3], [12]. These services are shown in Fig. 1. IaaS is responsible for providing hardware components as a service like memory, CPU, servers, storage and other virtualized compute resources, IaaS examples: Microsoft Azure, Apple iCloud, Google Drive, and Google Compute Engine and Amazon Web Services (AWS) [13], [12]. PaaS provides platforms like software development framework and operating system as a service, it represents a developer programming platform that is created for the programmer to develop, test, execute and administrate the applications, PaaS example: Google App Engine [14]–[16]. SaaS provides software and applications as a service, i.e. there is no need to install and run applications on client computers, where application interface is accessed on the end-users side. SaaS examples: Google Apps, Cisco WebEx, and Salesforce [17], [18].

The main objective of this paper is to review the most recent methods that have been developed for allocating resources inside cloud computing based on virtualization. Moreover, the virtualization role for efficiently providing the clients' resources is systematically analyzed. As a result of our analysis for the proposed approaches is that the requests execution time and VM availability are main issues and should be in consideration in any allocating resource approach.

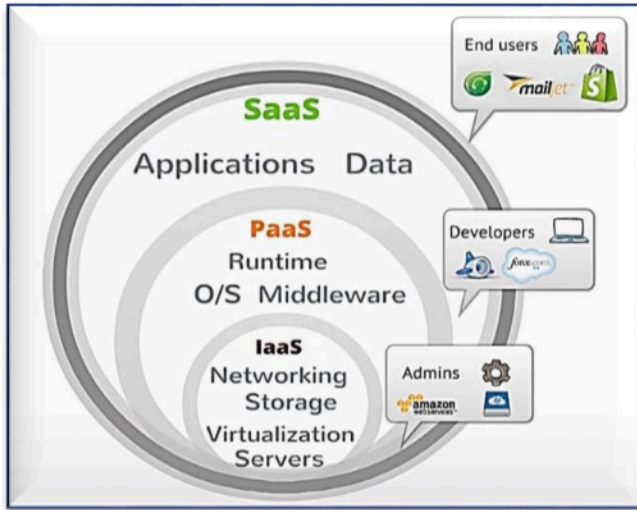


Fig. 1: Cloud computing services (IaaS, PaaS, and SaaS).

## II. CLOUD COMPUTING VIRTUALIZATION

Virtualization means creating a virtual version of something, rather than actual. In cloud computing, virtualization is a technique of replicating a version of something real digitally [19], [20]. It allows sharing a single physical resource or an application among many clients and organizations. Virtualization describes as software that lets a single hardware calculating device to be automatically partitioned to single/multi assumed devices [21], [22]. In turn, each of these virtual devices can be used and managed easily, and thus, it reduces cost by increasing the infrastructure utilizing, also it provides the agility required to speed up IT processes [1]. In particular, this technology allows multiple VMs to run concurrently on a single physical Host Machine (HM), in which every VM hosts its applications, operating system, and middleware, by using a partition of the hardware resources capacity (memory, CPU power, network bandwidth, and store capability) [23], [24]. Fig. 2 illustrates both of traditional architecture and virtual architecture. Multiple users demand resources, it requires a lot of investment in physical infrastructure to react to users' demands; therefore, the cloud infrastructure providers solve and manage this situation by offering VM services based on user requests with higher quality and at low cost [5], [25]. There are different forms of virtualization such as storage, server, appliance, desktop, and network virtualization [21], [26].

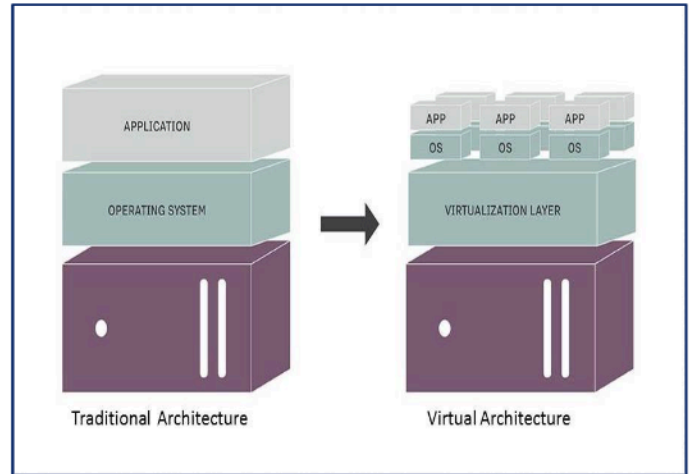


Fig. 2: Traditional architecture and virtual architecture [23].

## III. RESOURCE ALLOCATION

Cloud is a new generation technology that provides the clients rapidly with the hardware and software resources in order to satisfy their requirements [27]. In modern large-scale distributed systems, resource provisioning at the real-time became one of the main challenges, it is one of the major issues in cloud computing [14], [28]–[30]. To gain maximum network utility, the cloud provides clients with virtualized resources as a service over the internet in all situations with high quality. The IaaS platform has a virtualized structure that allows the re-allocation of resources by moving VMs among hosts [31], [32]. In cloud computing, integrated portion for IaaS model related to resource-sharing. This sharing defined as assigning process to the available resources to the clients according to their requirements [33], [34]. Mapping and scheduling are two steps of resource sharing of: calculating, planning allows locating them for application layer components like routers as example, hosts and connections of the network [35]–[37]. However the arrangement aims to improve the assigning of network-components [38]. Resource provisioning simply is done by mapping virtualized resources to physical ones. Re-allocation resource in distributed systems has major benefits of spreading and balancing the load on processors, utilizing network bandwidth, utilizing the infrastructure of datacenter under varying workload, and minimizing total request execution time. Thus, maximizing the profit for cloud service providers [14], [28]. Therefore many studies focused on resource allocation concept, its algorithms, and strategies in distributed systems. An authors of [39], proposed a hierarchical management framework of load balancing and resource allocation thus, decrease spent-power for midpoint information, taking QoS metrics into account. Also, an authors of [40], presented a scheduling algorithm based on spectral algorithm via converting the scheduling problem to the clustering problem, as a result of this algorithm, fine resource utilization obtains on the premise of QoS. Another research studied a game theory-based dynamic resource allocation strategy to allocate and utilize the bandwidth efficiently amongst diverse information-midpoints those geographically dispersed within the cloud. This strategy



be able to decrease response delay of Internet-based facilities with balancing physical network workflow [38], [41].

#### IV. LITERATURE REVIEW

The data center involves a large number of resources. The provision of these computational resources is controlled and managed by a third-party called “cloud providers” who are responsible for providing consumers with resources (hardware, software, and platforms) at any time they need it. Flexible style of resource-assigning is a challenge depending on clients’ needs. Therefore many researchers studied resource allocation within the cloud using various scheduling algorithms, Resource Allocation Strategies (RAS), and VM allocation techniques. Aiming to utilize cloud resources to gain benefits and improve network performance. In this review paper, the focussing will be on some proposed load balancing methods, resource allocation methods, and scheduling algorithms used for effective resource allocation of cloud through virtualization.

Deepika and Rao [14], proposed an approach to provide an active resource allocation in the cloud through virtualization. The presented approach provides the allocation of multiple virtual resources effectively by migrating VMs from the busy server (hot spot) to an idle server (cold spot). They used the “skewness” approach as a metric to compute the resource usage on the server, where the usage logs of previously used resources allow predicting the resource needs in the future. The proposed approach provides dynamic resource allocation through optimizing burdens (to satisfy the VM needs according to the server’s capacity) and green computing (to save the energy of the server by optimizing unnecessary usage of the server).

Wang et al. [42], developed a Prediction-based Energy Conserving Resource Allocation algorithm (ECRASP) for the cloud. The ECRASP encompasses two main components, the first component is the prediction mechanism and the second component is the job allocation mechanism. The prediction mechanism helps to predict the trend of incoming jobs in the future (sparse or dense) to make a decision adequately thus, conserving the energy. The allocation mechanism distributes the incoming jobs to proper PMs to balance the load of each of them. Actually, the allocation mechanism tries to shut down the idle PMs (i.e. lightly loaded), on the other hand, if the system is in a heavy loading case, the allocation mechanism will start up the new PM to process a new incoming job. As a result, the proposed algorithm conserves power consumption by making an adequate decision on the time of shutdown or start-up PMs.

Kruekaew and Kimpan [22], applied an optimized Artificial Bee Colony (ABC) algorithm to schedule VMs on cloud computing under changing environments (i.e. changing in the number of tasks and VMs). The authors compared the proposed algorithm with other scheduling algorithms like First Come First Serve (FCFS), Shortest Job First (SJF), and Longest Job First (LJF). The experimental results showed that the combination of ABC and LJF (ABC\_LJF) performed better than other methods where its performance is more notable in scalability. They suggested that ABC\_LJF must be used for load balancing under conditions of increasing and decreasing

servers’ numbers to keep the system stable and scheduling thus prevent a system crash.

Saraswathi, et al. [1], suggested dynamic model of incomes-specification within cloud computing, it is a new method for VM specification depending task to the user based on the features (i.e. job priority), in which virtual resources are reconfigured dynamically without needing to create a new VM thus, increasing the resource utilization. The VM obtainability was measured when a new job arrives. With the VM obtainability then additional task specified to be executed, but if there is no available VM. In this case, the algorithm finds minimum importance task (with considering the task’s occupancy kind) and running-stop temporarily via incomes-forestalling. Then great importance task will be capable of execution with forestalled resource at from the little importance task. The paused job could be resumed after completion of all rest tasks by VMs, this procedure occurs when task’s occupancy kind been suspend able. But when task’s occupancy kind not suspend able, in this case, the paused job could be resumed after the achievement of minimum importance task execution using the task incomes. The same procedure is performed for all the new incoming jobs. It is clear that the proposed model provides little complexity through running complete tasks with respect to generating fresh VMs.

Pradhan, et al. [43], proposed a scheduling algorithm called modified Round-Robin to allocate resources within cloud computing. The proposed algorithm aims to obtain an optimal scheduling model, in order to satisfy clients’ requirements by reducing the waiting time (i.e. rapid response to the resource request). With the first resource request, the algorithm time begins. When a new request arrives at the ready queue then the algorithm computes the average of the sum of times of requests placed in the ready queue including the new arrival request, in order to guarantee the new request. If the request finishes its burst time after the execution process, then it will be removed from the ready queue, or else it will move to the tail of the ready queue. As a result, the modified Round-Robin scheduling algorithm model has an obvious role in reducing the average waiting time and turnaround time of different processes (jobs).

Padmavathi and Basha [44], proposed a dynamic and elastic algorithm to perform load balancing of cloud computing using Ant Colony Optimization (ACO). The proposed approach was bio-enthusied approach, where a biological ants’ behavior is simulated into the artificial algorithm, with adding and omitting some features. During the food search process, ants use trail pheromone to mark the path from food source to its nest. When ants move it looks for pheromone trail and follows the trail with the highest percentage pheromones deposits (i.e. shortest path). The proposed Dynamic and Elastic Ant Colony Optimization Load Balancing (DEACOLB) algorithm is applied to VMs in data centers of cloud computing. The experimental results show that the average Make Span (i.e. finish time) is less than other existing algorithms like ACO, and First Come First Serve (FCFS). Where average Make Span decreases when the number of jobs increases.

Usman et al. [45], brought to light an Energy-Efficient VM allocation method using Interior Search Algorithm (EE-ISA) to reduce energy consumption within the data center and optimize

resource utilization. The architecture of the presented method assumes that the data center consists of many servers known as PM; each PM has a fixed-size of CPU, RAM, and storage. The broker is responsible for handling the client's request then sends the request to the cloud after that to the data center. Then the request is treated via the data center manager that leads to the reservation of a specific number of VMs as requested by the client. Different classes of VMs are offered by the cloud provider where the client can choose from. These classes have different specifications related to CPU, storage, and RAM of each cloud resource. Consequently, every VM will be allocated on PM within the data center resource. The result of the proposed EE-ISA showed efficient resource utilization and energy consumption where 30% of energy is saved compared with Best-fit Decreasing (BFD) algorithm and Genetic Algorithm (GA).

Akintoye and Bagula [46], proposed a linear programming model to represent the resource allocation issue in the cloud environment and presented Binding Policy Based on the Hungarian Algorithm (HABBP) to optimize the model. The HABBP method uses the policy of load balancing in order to bind the cloudlets to appropriate VMs to optimize the total task execution time. They contributed to propose HABBP modules to CloudSim in order to (1) provide an interactive interface that permits the user to configure cloudlets parameters instead of imbedding parameters in source codes, (2) introduce a novel binding policy based on HABBP algorithm. The simulation findings illustrated that the HABBP has higher performance than the traditional binding policy of CloudSim in term of job total execution time. It also revealed that the HABBP policy can solve and optimize the problem of allocating VM in the cloud environment efficiently.

Atiewi et al. [21], studied the impact of virtualization on energy consumption in the cloud. They presented an experiment using a power saver scheduler algorithm (PSSA) under a green cloud simulator in order to compare two scenarios (virtualized server and non-virtualization server). The comparison made based on three parameters which are datacenter load, total power consumption, and Make Span. The findings showed that (1) virtualized datacenter environment performs better than the non-virtualized in terms of the data center load, (2) non-virtualized server consumes more energy than the virtualized counterparts, (3) datacenter with non-virtualized servers performed better than virtualized counterparts in terms of Make Span because the tasks are sent directly to the PM instead of VMs.

Yin et al. [47], proposed an effective task scheduling strategy depends on double-fitness Load balancing and task completion Cost Genetic Algorithm (LCGA). The proposed algorithm aims to optimize the task scheduling in a cloud environment by (1) achieving least task execution cost to make full utilization of systems' resources that meets clients' requirements (2) guarantee the load balancing in order to get an optimal task allocation in a dynamic cloud environment, thus, satisfying the service providers' needs. They compared the proposed LCGA with Load balancing Genetic Algorithm (LGA) and task completion Cost Genetic Algorithm (CGA) to test the effectiveness of the proposed optimized algorithm. The comparison results showed that in LGA, load balancing is

obtained obviously in contrast, the effect of task completion cost is not obtained. The result also clarified that using CGA minimum task completion cost is obtained while load balancing effect in not shown. The results of LCGA showed the load balancing and least completion cost are obtained at the same time, thus effective task scheduling is proved.

Chen [48], proposed a cloud resource allocation method to Support Sudden and Urgent (SSU) resource demands optimally and timely. The proposed method depends on setting up the priority of resource allocation that is computed according to the urgent degree of the resource request and the client's priority. Then the VMs are mapped to physical hosts optimally via a multi-objective optimization approach. The author conducted a comparison experiment between the proposed SSU method, Round-Robin, and Best-Fit (BF) resource allocation methods. The experimental results showed that BF chooses the PM that has less CPU usage to allocate VMs, Round-Robin allocate VM to PM in circular order and equal portion while the presented method adopts matching resource distance to assign the VM to PM. The results also showed that the SSU method reduces resource usage (i.e. the number of used PM) by shutting down the PM with underutilized CPU, improve the utilization of CPU, and realize minimum resource fragment.

Rengasamy and Chidambaram [13], proposed a novel predictive framework to solve problems of VMs migration and placement by allocating resources effectively for cloud computing. The proposed method uses an algorithm called Random Algorithm that connects tasks (cloudlets) and servers randomly to allow handling a large number of tasks and distributing the workload to each node. The Random Algorithm gives each user a list of available servers thus eliminating the requirement of a centralized broker. The evaluation result showed that the proposed method can balance the load in seconds rather than minutes, the results also explored high performance in the cloud.

Khodar et al. [49], Presented a new scheduling strategy for load balancing based on a genetic algorithm to overcome the obstacle of resource scheduling in the cloud environment. The ability to test a wide range of solutions and select the best solution is one of the major features of the genetic algorithm. The proposed algorithm focuses on the current state and historical data of the system to compute in advance its effect on system load after distributing the required VM resources. Then the distribution with the least impact will be chosen through the operators of crossover and mutation. The experimental findings revealed that the proposed strategy method provides a better load balancing taking into consideration the historical factor; it also reduces the dynamic migration.

Bhardwaj et al. [33], presented an autonomous resource allocation model to allocate and de-allocate the required resources within the cloud data center dynamically. The proposed mechanism aims to improve response time and VMs utilization by distributing VMs in the server-side to reduce the delay time of cloud service. They used a horizontal scaling method (scale-out) to allocate more VMs as per the requests. If the VMs are underloaded or overloaded then the proposed model increases VMs number till the average response time amounts to a predefined value. The result showed that the

proposed method efficiently allocates VMs with respect to the load; it also reduces the request waiting time in the queue via increasing VMs number at the run time.

Than and Thein [50], proposed two energy-saving resource allocation algorithms to save energy consumption of the cloud data center. The first algorithm called DSJF which is a combination of power management technique known as Dynamic Voltage and Frequency Scaling (DVFS), Shortest Job First (SJF) resource allocation algorithm, and the Cupic power model. The second algorithm called DFCFS which is a combination of DVFS, First Come First Serve (FCFS) resource allocation algorithm, and the Cupic power model. They used the CloudSim simulator to compare and evaluate energy consumption using the proposed algorithms. The findings revealed that DSJF gains energy efficiency more than DFCFS where it can save energy consumption by up to 55% compared to DFCFS. The finding also explored that DSJF is appropriate if the incoming requests have different running times.

## V. DISCUSSION

Cloud computing should be able to provide the clients with services according to their demands at any time. Therefore cloud employs virtualization and resource provisioning techniques to allow active resource allocation and load balancing in the data center.

For this purpose, this paper focus on various resource allocation models, scheduling algorithms, load balancing algorithms, resource allocation optimization approaches, and energy-saving algorithms. For instance, the DEACOLB algorithm, modified Round-Robin, ECRASP, PSSA, “skewness” approach, ACB, EE-ISA, HABBP, SSU, LCGA, DSJF and DFCFS, and other models for allocating VMs dynamically for cloud based on the jobs’ priorities are addressed and analyzed. Each above-stated method has its objectives, algorithm, tool/technique, and significant results, as shown in table 1. These previously mentioned methods have an obvious role in satisfying the users’ needs, reduce the average waiting time of different processes, save cost (reduces resource usage), provide efficient shared resource utilization, and balance the load of the cloud data center. Thus, resource allocation plays an important role in resource utilization, determining the performance, and power consumption of the cloud data center.

After explaining all fifteen researches browsed in Table I, it can be recommended that DSJF and DFCFS energy-saving resource allocation algorithms most depended which using CloudSim simulator tool. In addition, the related research’s results indicated that DSJF algorithm gains energy efficiency more than DFCFS where it can save energy consumption by up to 55% compared to DFCFS.

TABLE I. COMPARISON OF THE PROPOSED APPROACHES BY PREVIOUS RESEARCHES.

Author(s)	Year	Objectives	Algorithm	Tool/ Technique	Significant Results
Deepika and Rao [14]	2014	Provide an active resource allocation in the cloud through virtualization.	Use the “skewness” approach to compute the resource usage on the server and provide dynamic resource allocation through optimizing burdens and green computing algorithm.	Not declared by authors.	Satisfy the VM needs and save the energy of the server.
Wang et al. , [42]	2014	Make an adequate decision on the time of shutdown or start-up PMs.	Propose the ECRASP algorithm for cloud computing.	Java.	Conserve power consumption of PMs.
Kruekaew and Kimpan [22]	2014	Provide scheduling strategy under changing environments on the cloud.	Apply an optimized ACB algorithm.	CloudSim simulator.	ABC_LJF reduces the Make Span under changing conditions, it also keeps the system stable.
Saraswathi et al. , [1]	2015	Dynamic reconfiguring VMs allocation for the user.	Propose an algorithm to allocate VMs dynamically based on Jobs’ priorities.	CloudSim simulator.	Reduce the overhead in executing all jobs if compared with the creation of a new VM.
Pradhan et al. , [43]	2016	Load balancing Within cloud computing.	Round-Robin scheduling algorithm.	MATLAB.	Reduce the average waiting time and turnaround time of different processes.
Padmavathi and Basha [44]	2017	Load balancing Within cloud computing.	DEACOLB algorithm.	CloudSim simulator.	Reduce average Make Span and standard deviations.
Usman et al. , [45]	2017	To reduce energy consumption within the data center and optimize resource utilization.	EE-ISA method.	CloudSim simulator.	Efficient resource utilization and energy consumption where 30% of energy is saved within the data center.
Akintoye and Bagula [51]	2017	To solve and optimize the problem of allocating VM in the cloud environment.	HABBP algorithm.	CloudSim simulator.	The findings showed that the HABBP outperforms the traditional binding policy of CloudSim in term of job total execution time.

Atiewi et al. , [21]	2018	To study the impact of virtualization on energy consumption in the cloud.	Use PSSA and green algorithm to compare two scenarios (virtualized server and non-virtualization server).	GreenCloud simulator.	The Virtualized data center environment showed better performance than the non-virtualized counterparts in terms of data center load and server energy consumption. However, the Make Span is the basic drawback of the virtualized environment.
Yin et al. , [47]	2018	To provide an effective task scheduling strategy.	Propose a LCGA algorithm.	CloudSim simulator.	The load balancing and the least completion cost are obtained at the same time.
Chen [48]	2018	To guarantee resource allocation demands optimally and timely.	Propose the SSU method to provide resource allocation of sudden and urgent demands.	Jmetal software and CloudSim simulator.	The proposed SSU method reduces resource usage via shutting down the PM with underutilized CPU, improve the utilization of CPU, and realize less resource fragment.
Rengasamy and Chidambaram [13]	2019	To fulfill the needs for effective resource allocation for the cloud.	Propose Random Algorithm for effective resource allocation.	CloudSim simulator.	The effectiveness of the proposed method revealed high performance in the cloud.
Khodar et al. , [49]	2019	To overcome the obstacle of resource scheduling in the cloud environment.	Present a new strategy scheduling for load balancing based on a genetic algorithm.	Java.	The proposed strategy method provides a better load balancing by reducing live migration.
Bhardwaj et al. , [33]	2019	To overcome the problems of underutilization and overutilization resources within the cloud data center.	Present an autonomous resource allocation model to provide resources dynamically.	CloudSim simulator.	The proposed method efficiently allocates VMs with respect to the load; it also reduces the request waiting time in queue.
Than and Thein [50]	2020	To contribute the workload in order to save energy consumption of the cloud data center.	DSJF and DFCFS energy-saving resource allocation algorithms.	CloudSim simulator.	DSJF algorithm gains energy efficiency more than DFCFS where it can save energy consumption by up to 55% compared to DFCFS.

## VI. CONCLUSION

Table I, represents a detailed comparison among fifteen newest researches and very related to the cloud computing virtualization of resources allocation for distributed systems. The comparison depended on mean metrics including: objectives, algorithms, tool, techniques, and significant results. The addressed researches concentrated on: active resource allocation in the cloud through virtualization, adequate decision on the time of shutdown, Provide scheduling strategy under changing environments on the cloud. Adding to that, many algorithms addressed, the recommended algorithms that provides best results are DSJF and DFCFS energy-saving resource allocation algorithms, and DSJF algorithm very good gains energy efficiency.

Cloud is a resource provisioning technique that aims to provide clients with various resources anywhere and at any time. The process of providing clients with a shared virtualized environment (server, storage, operating system, and hardware) is a main challenge for the cloud provider. Also, the appropriate allocation of VMs in cloud data centers is one of the significant optimization issues in the cloud. Therefore, this paper presents a review of different existing VM scheduling algorithms and resource allocation techniques. Each VM allocation method and scheduling algorithm has a vital role in utilizing the shared resources of the cloud. Based on our review, the availability of VM resource and execution time of requests are the key factors

to be considered in any optimal resource allocation algorithm. The paper also revealed that virtualization technique has a significant impact on enhancing the network performance, save the cost of PMs in the cloud data center, balance the load, conserve the server's energy, and allocate resources actively thus satisfying the clients' requirements.

## REFERENCES

- [1] A. T. Saraswathi, Y. R. A. Kalaashri, and S. Padmavathi, "Dynamic Resource Allocation Scheme in Cloud Computing," *Procedia Comput. Sci.*, vol. 47, pp. 30–36, Jan. 2015, doi: 10.1016/j.procs.2015.03.180.
- [2] P. Priyadarshinee, R. D. Raut, M. K. Jha, and B. B. Gardas, "Understanding and predicting the determinants of cloud computing adoption: A two staged hybrid SEM - Neural networks approach," *Comput. Hum. Behav.*, vol. 76, pp. 341–362, Nov. 2017, doi: 10.1016/j.chb.2017.07.027.
- [3] S. Vakiliinia, M. M. Ali, and D. Qiu, "Modeling of the resource allocation in cloud computing centers," *Comput. Netw.*, vol. 91, no. Supplement C, pp. 453–470, Nov. 2015, doi: 10.1016/j.comnet.2015.08.030.
- [4] O. Alzakholi, L. Haji, H. Shukur, R. Zebari, S. Abas, and M. Sadeeq, "Comparison Among Cloud Technologies and Cloud Performance," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 40–47, Apr. 2020, doi: 10.38094/jastt1219.
- [5] D. Kesavaraja and A. Shenbagavalli, "QoE enhancement in cloud virtual machine allocation using Eagle strategy of hybrid krill herd optimization," *J. Parallel Distrib. Comput.*, Sep. 2017, doi: 10.1016/j.jpdc.2017.08.015.
- [6] R. R. Zebari, S. R. Zeebaree, and K. Jacksi, "Impact Analysis of HTTP and SYN Flood DDos Attacks on Apache 2 and IIS 10.0 Web Servers,"



- in 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 156–161.
- [7] R. R. Zebari, S. R. Zeebaree, K. Jacksi, and H. M. Shukur, "E-Business Requirements For Flexibility And Implementation Enterprise System: A Review," *International Journal of Scientific & Technology Research (IJSTR)*, vol. 8, no. 11, pp. 655–660, 2019.
- [8] S. R. Zeebaree, K. Jacksi, and R. R. Zebari, "Impact analysis of SYN flood DDoS attack on HAProxy and NLB cluster-based web servers," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 19, no. 1, pp. 510–517, 2020.
- [9] S. R. Zeebaree, R. R. Zebari, and K. Jacksi, "Performance analysis of IIS10.0 and Apache2 Cluster-based Web Servers under SYN DDoS Attack," *TEST Engineering & Management*, vol. 83, no. March - April 2020, 5854 - 5863, 2020.
- [10] S. R. Zeebaree, R. R. Zebari, K. Jacksi, and D. A. Hasan, "Security Approaches For Integrated Enterprise Systems Performance: A Review," *International Journal of Scientific & Technology Research (IJSTR)* vol. 8, no. 12, pp. 2485–2489, 2019.
- [11] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 56–70, May 2020, doi: 10.38094/jastt1224.
- [12] A. A. S. Farrag, S. A. Mahmoud, and E. S. M. El-Horbaty, "Intelligent cloud algorithms for load balancing problems: A survey," presented at the 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS), Dec. 2015, pp. 210–216, doi: 10.1109/IntelCIS.2015.7397223.
- [13] R. Rengasamy and M. Chidambaram, "A Novel Predictive Resource Allocation Framework for Cloud Computing," in 2019 5th International Conference on Advanced Computing Communication Systems (ICACCS), Mar. 2019, pp. 118–122, doi: 10.1109/ICACCS.2019.8728526.
- [14] T. Deepika and A. N. Rao, "Active resource provision in cloud computing through virtualization," in 2014 IEEE International Conference on Computational Intelligence and Computing Research, Dec. 2014, pp. 1–4, doi: 10.1109/ICCIC.2014.7238373.
- [15] O. M. Ahmed and W. M. Abdulllah, "A Review on Recent Steganography Techniques in Cloud Computing," *Acad. J. Nawroz Univ.*, vol. 6, no. 3, pp. 106–111, 2017.
- [16] O. M. Ahmed and A. B. Sallow, "Android security: a review," *Acad. J. Nawroz Univ.*, vol. 6, no. 3, pp. 135–140, 2017.
- [17] K. Zaki and H. Saad, "Adoption of Cloud Human Resource Information System in Egyptian Hotels: An Experimental Design Research," *Int. J. Herit. Tour. Hosp.*, vol. 12, no. 1, pp. 233–245, Mar. 2018.
- [18] N. Harki, A. Ahmed, and L. Haji, "CPU Scheduling Techniques: A Review on Novel Approaches Strategy and Performance Assessment," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 48–55, 2020.
- [19] M. A. Sadeeq, S. R. Zeebaree, R. Qashi, S. H. Ahmed, and K. Jacksi, "Internet of Things security: a survey," in 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 162–166.
- [20] S. A. Mostafa, S. S. Gunasekaran, A. Mustapha, M. A. Mohammed, and W. M. Abdulllah, "Modelling an Adjustable Autonomous Multi-agent Internet of Things System for Elderly Smart Home," in *Advances in Neuroergonomics and Cognitive Engineering*, Cham, 2020, pp. 301–311, doi: 10.1007/978-3-030-20473-0\_29.
- [21] S. Atiewi, A. Abuhussein, and M. A. Saleh, "Impact of Virtualization on Cloud Computing Energy Consumption: Empirical Study," in *Proceedings of the 2<sup>nd</sup> International Symposium on Computer Science and Intelligent Control*, Stockholm, Sweden, Sep. 2018, pp. 1–7, doi: 10.1145/3284557.3284738.
- [22] B. Kruekaew and W. Kimpan, "Virtual Machine Scheduling Management on Cloud Computing Using Artificial Bee Colony," *Lect. Notes Eng. Comput. Sci.*, vol. 1, pp. 18–22, Mar. 2014.
- [23] M. Ficco, C. Esposito, F. Palmieri, and A. Castiglione, "A coral-reefs and Game Theory-based approach for optimizing elastic cloud resource allocation," *Future Gener. Comput. Syst.*, vol. 78, no. Part 1, pp. 343–352, Jan. 2018, doi: 10.1016/j.future.2016.05.025.
- [24] S. R. Zeebaree, L. M. Haji, I. Rashid, R. R. Zebari, O. M. Ahmed, K. Jacksi, & H. M. Shukur, "Multicomputer Multicore System Influence on Maximum Multi-Processes Execution Time," *TEST Engineering & Management*, vol. 83, no. May/June, pp. 14921–14931, May 2020.
- [25] O. H. Jader, S. R. Zeebaree, and R. R. Zebari, "A State Of Art Survey For Web Server Performance Measurement And Load Balancing Mechanisms," *International Journal of Scientific & Technology Research (IJSTR)*, vol. 8, no. 12, pp. 535–543, 2019.
- [26] M. Singh, "Virtualization in Cloud Computing- a Study," in 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Oct. 2018, pp. 64–67, doi: 10.1109/ICACCCN.2018.8748398.
- [27] Z. N. Rashid, S. R. Zebari, K. H. Sharif, and K. Jacksi, "Distributed Cloud Computing and Distributed Parallel Computing: A Review," presented at the 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 167–172.
- [28] Z. N. Rashid, S. R. Zeebaree, and A. Shengul, "Design and Analysis of Proposed Remote Controlling Distributed Parallel Computing System Over the Cloud," presented at the 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 118–123.
- [29] S. R. Zebari and A. S. Yowakib, "Improved Approach for Unbalanced Load-Division Operations Implementation on Hybrid Parallel Processing Systems," *Sci. J. Univ. Zakho*, vol. 1, no. 2, pp. 832–848, 2013.
- [30] S. R. M. Zeebaree, H. M. Shukur, L. M. Haji, R. R. Zebari, K. Jacksi, and S. M. Abas, "Characteristics and Analysis of Hadoop Distributed Systems," *Technology Reports of Kansai University*, vol. 62, no. 4, pp. 1555–1564, Apr. 2020.
- [31] M. Duggan, J. Duggan, E. Howley, and E. Barrett, "A network aware approach for the scheduling of virtual machine migration during peak loads," *Clust. Comput.*, vol. 20, no. 3, pp. 2083–2094, Sep. 2017, doi: 10.1007/s10586-017-0948-7.
- [32] S. R. Zeebaree, H. M. Shukur, and B. K. Hussan, "Human resource management systems for enterprise organizations: A review," *Period. Eng. Nat. Sci.*, vol. 7, no. 2, pp. 660–669, 2019.
- [33] T. Bhardwaj, H. Upadhyay, and S. C. Sharma, "Autonomic Resource Allocation Mechanism for Service-based Cloud Applications," in 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Oct. 2019, pp. 183–187, doi: 10.1109/ICCCIS48478.2019.8974515.
- [34] S. R. Zeebaree, A. B. Sallow, B. K. Hussan, and S. M. Ali, "Design and Simulation of High-Speed Parallel/Sequential Simplified DES Code Breaking Based on FPGA," in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 76–81.
- [35] K. Jacksi, N. Dimililer, and S. R. Zeebaree, "State of the art exploration systems for linked data: a review," *Int J Adv Comput Sci Appl IJACSA*, vol. 7, no. 11, pp. 155–164, 2016.
- [36] K. Jacksi, N. Dimililer, and S. R. Zeebaree, "A survey of exploratory search systems based on LOD resources," 2015.
- [37] K. Jacksi, S. Zeebaree, and N. Dimililer, "Design and Implementation of LOD Explorer: A LOD Exploration and Visualization Model," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 31–39, 2020.
- [38] X. Yuan Min, Geyong, Yang, Laurence T. ,Ding, Yi, Fang, Qing, "A game theory-based dynamic resource allocation strategy in Geo-distributed Datacenter Clouds," *Future Gener. Comput. Syst.*, vol. 76, no. Supplement C, pp. 63–72, Nov. 2017, doi: 10.1016/j.future.2017.04.046.
- [39] N. Leontiou, D. Dechouniotis, S. Denazis, and S. Papavassiliou, "A hierarchical control framework of load balancing and resource allocation of cloud computing services," *Comput. Electr. Eng.*, vol. 67, pp. 235–251, 4, doi: 10.1016/j.compeleceng.2018.03.035.
- [40] H. Zhou, S. Deng, H. Huang, and Y. Wu, "Resource allocation in cloud computing based on clustering method," presented at the 2015 Annual IEEE Systems Conference (SysCon) Proceedings, Apr. 2015, pp. 489–494, doi: 10.1109/SYSCON.2015.7116799.
- [41] B. R. Ibrahim, S. R. Zeebaree, and B. K. Hussan, "Performance Measurement for Distributed Systems using 2TA and 3TA based on OPNET Principles," *Sci. J. Univ. Zakho*, vol. 7, no. 2, pp. 65–69, 2019.
- [42] C.-F. Wang, W.-Y. Hung, and C.-S. Yang, "A prediction based energy conserving resources allocation scheme for cloud computing," in 2014 IEEE International Conference on Granular Computing (GrC), Oct. 2014, pp. 320–324, doi: 10.1109/GRC.2014.6982857.
- [43] P. Pradhan, P. K. Behera, and B. N. B. Ray, "Modified Round Robin Algorithm for Resource Allocation in Cloud Computing," *Procedia Comput. Sci.*, vol. 85, pp. 878–890, Jan. 2016, doi: 10.1016/j.procs.2016.05.278.

- [44] M. Padmavathi and S. M. Basha, "Dynamic and elasticity ACO load balancing algorithm for cloud computing," presented at the 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), Jun. 2017, pp. 77–81, doi: 10.1109/ICCONS.2017.8250571.
- [45] M. J. Usman, A. Samad, Ismail, H. Chizari, and A. Aliyu, "Energy-Efficient virtual machine allocation technique using interior search algorithm for cloud datacenter," in 2017 6th ICT International Student Project Conference (ICT-ISPC), Johor, Malaysia, May 2017, pp. 1–4, doi: 10.1109/ICT-ISPC.2017.8075327.
- [46] S. B. Akintoye and A. Bagula, "Optimization of virtual resources allocation in cloud computing environment," in 2017 IEEE AFRICON, Sep. 2017, pp. 873–880, doi: 10.1109/AFRCON.2017.8095597.
- [47] S. Yin, P. Ke, and L. Tao, "An improved genetic algorithm for task scheduling in cloud computing," in 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), May 2018, pp. 526–530, doi: 10.1109/ICIEA.2018.8397773.
- [48] J. Chen, "A Cloud Resource Allocation Method Supporting Sudden and Urgent Demands," in 2018 Sixth International Conference on Advanced Cloud and Big Data (CBD), Aug. 2018, pp. 66–70, doi: 10.1109/CBD.2018.00021.
- [49] A. Khodar, H. Fadhil, and I. Alkhayat, "New Scheduling Approach for Virtual Machine Resources in Cloud Computing based on Genetic Algorithm," in 2019 International Russian Automation Conference (RusAutoCon), Sep. 2019, pp. 1–5, doi: 10.1109/RUSAUTOCON.2019.8867638.
- [50] M. M. Than and T. Thein, "Energy-Saving Resource Allocation in Cloud Data Centers," in 2020 IEEE Conference on Computer Applications (ICCA), Feb. 2020, pp. 1–6, doi: 10.1109/ICCA49400.2020.9022819.
- [51] S. B. Akintoye and A. Bagula, "Optimization of virtual resources allocation in cloud computing environment," in 2017 IEEE AFRICON, Sep. 2017, pp. 873–880, doi: 10.1109/AFRCON.2017.8095597.