A BRIEF OVERVIEW OF LAST DECADE IN CLOUD COMPUTING ESSENTIAL FEATURE: RESOURCE ALLOCATION

Maira Khalid[1], Uzma Aziz[2], Faiza Latif[3], Samreen Kanwal[4]
[1] Institute of Southern Punjab, Multan, Pakistan, maira.khalid7272@gmail.com
[2]City College of Science and Commerce, Multan, Pakistan, uzma.a579@gmail.com
[3]City College of Science and Commerce, Multan, Pakistan, faizalatif116@gmail.com
[4]City College of Science and Commerce, Multan, Pakistan, samreenkhan6g@gmail.com

ABSTRACT— In the present era, cloud computing is assumed as an emerging paradigm in service providing systems. Cloud computing demands different types of resources as services for the consumer needs. So, it is a challenge for cloud computing to provide all resources to the consumers without any delays. But resource availability is finite and it is considered as a serious challenge to deliver the required resources to the consumers on exact time. Resources related to cloud are allocated on the fair demands by the consumers and for this reason cloud computing has a challenge to meet the issue of resource allocation. This paper consists of a literature review about the resource allocation in the area of cloud computing that explains the issues and solutions of the problems that occur in resource allocation. Moreover, we have explained the methodologies and techniques that are used in the research papers of the last ten years and we gave comparisons of these techniques for better understanding. This paper will help the new researchers of cloud computing for learning different methods about resource allocation. Furthermore, describes the cloud computing resource allocation strategies and different algorithms used to accomplish the allocation. The main focus of this study is to describe the strategies and their usage to analyze the efficient strategy for resource allocation. Additionally, explains the different techniques of different years to highlight the progress and helps the new researcher to comprehend the strategies and algorithms cloud computing, resource allocation.

Keywords --- resource allocation algorithm, resource allocation, cloud computing, distributed computing.

_____

I.INTRODUCTION

In this modern age, technology is considered as basic need of life like water, electricity etc. All consumers demanding these basics for covering basic needs in life. They paid these services for the usage [1] [2]. Like these basic needs, computing became the need of today's technology and meet the all consumers need. This computing is known as cloud computing that provides resources virtually to the consumers on their demand and charge the services according to consumption.

"A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resource(s) based on service-level agreements established through negotiation between the service provider and consumers." [2]. Cloud Computing comprises three (3) services [3] [4] [5] [6].

1) Software as a Service (SaaS)
   In Cloud SaaS, consumer practices the applications of cloud service delivered by the cloud infrastructure software service instead of local application. Cloud provider is responsible to maintain the software services provided to the consumer. Consumer has to pay for the services according to their usage SaaS is the best example of the modern technology. (CRM) Customer Relationship Management Salesforge.com are the practical examples of these service model [3] [4] [7] [8] [9].

2) Platform as a Service (PaaS)

In Cloud PaaS, it is a main platform environment provided to developers for creating and deploying the applications. Using this platform, consumers are free from the care about fundamental infrastructure of cloud including networks, servers, storage and operating system but controlled the deploy applications and can run services. Example include for such models: Google Application, Microsoft Azure, Right Scale and Engine [3] [4] [5] [7] [9]

3) Infrastructure as a Service (IaaS)

In Cloud IaaS, it provides large scale of resources for computing in processing and storing. In this service model, consumer have limited control to select network components like host, firewalls and can manage storage, deployed applications, operating system. It is also familiar as HaaS (Hardware as a Service). Examples include for such models: GoGrid , Eucalyptus, Amazon Web Services, Flexiscale and Open Stack, offers IaaS [3] [4] [5] [7] [9].

Various deployment models are used in cloud computing depending upon the physical and location distribution. Beside these services, cloud categorizes into four models:

a. Private Cloud

Every device connected to a computer system is a resource (like printers, fax machines etc.) It is a feature/method (resource allocator) used to ensure the efficient use of resources of a system.

b. Community Cloud

Community cloud services are provided to specific community that is shared by organization. This service can be used in organization community of location or allowed community location [3] [4] [5].

c. Hybrid Cloud

Hybrid cloud is the grouping of two or more clouds (Public, Private, Community) [3] [4] [5].

Private cloud services belong to organization only, third party or general public cannot use private cloud services. It is managed by organizations people themselves [3] [4] [5] [10].

d. Public Cloud

Public cloud services are provided to consumers by organization according to the consumer needs and requirements. Consumers have to pay for the services provide by public cloud. Business consumers uses public cloud for
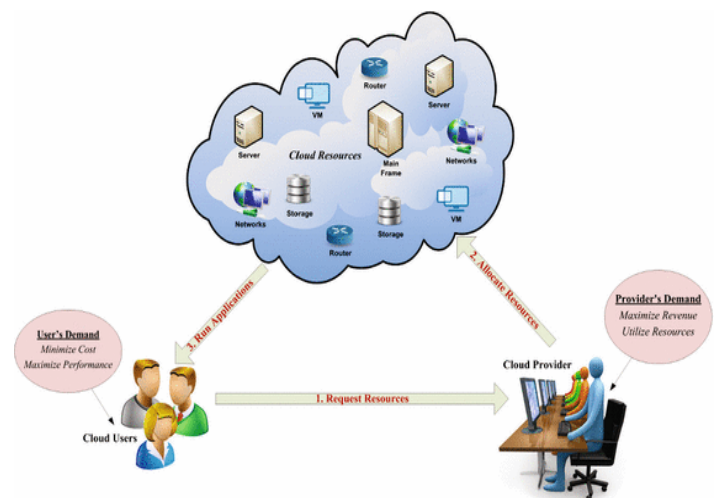
saving the cost of software and hardware but this can lead to data management, data security, level of performance and control issues etc. [3] [4] [5] [10].

Figure 1 RA Cloud Computing

There are many features in cloud computing but the most essential is resource allocation. Resource allocation is defined as it is a process of managing or assigning resources (CPU, memory etc.) in a manner that it fulfills a process requirement to be executed in a system, shown in Figure 1.

1.1 Resource allocation (in cloud computing)

In this process, required available resources are allocated to application that is used by cloud service consumer (CSC) over the internet.

1.2 Importance:

Resource allocation is very important in cloud computing because it gives a clear picture on the amount of work that done by a process/job generated/requested form client side. It helps to monitor and manage resource utilization in an efficient manner. It helps to provide right resource to a user/process at the right time. It improves the Quality of Service (QoS) of a system by maximizing the efficient of resources.

1.3 Advantages:

a. Maximize the utilization of resources

b. Improve response time management

c. Automatic Manage task workload

1.4 Disadvantages:

Execution time is limited (in the case of bad internet connection).

Table 1: Techniques/algorithms used for Resource Allocation in Cloud Computing

Year	Techniques/algorithms	Effects
2011	EDF (Earliest Deadline First)	Response Time
2012	Priority based Allocation	Task completion before deadline
2013	Schedular for Data	Task completion within allocated time
2014	Efficient Resource Allocation Service	Minimum delay in data transfer
2015	Green Cloud Computing	Efficient energy/power utilization in data center
2016	Modified Round Robin Algorithm	Dynamic response time or time slice
2017	DPM (Distributed Power Management)	Reduce power consumption
2018	Heuristic Approach for Allocation and Scheduling	Minimize response time and maximize resource allocation utilization efficiency
2019	Virtual Machine Model Placement	Reduce cost
2020	Priority Based Allocation	Minimum cost, and resource utilization

II.   LITERATURE REVIEW AND METHODOLOGY

Table 2: Methodologies used for Resource Allocation in the domain of Cloud Computing

Paper Name	Methodology
Resource Allocation for Real-Time Tasks using Cloud Computing [11], 2011.	This paper is almost complete description of resource allocation for performing the tasks of real time using IaaS model of cloud computing. So, user can select different computer resources based on the requirement. Real time task means to complete the tasks before deadline. Cloud computing suggests the resources with different speed, performance and cost. Efficiently allocation of resources is a big or main issue. We examine the problem and allocate the demanded resources for purpose real time task so in this way economical cost is reduced and all tasks are completed before the deadline. Paper describes the allocation of resources efficiently. For the solution of problem one method is introduced that is EDF (earliest deadline first). There are many real time applications which required completing their workload before deadline such as object recognition that is hosted on the cloud. So, object is recognized within specified time period and gives value to the users. This study is different from previous study first it examines the schedule on a given numbers of process and second the available virtual machines are infinite. At any instant so we can select different types of virtual machine based upon the computing resources and cost also. If user select slower or cheaper VMS and tasks is completed before deadline. Cost of VMs charges is fixed for a given rental period. if task is completed before hour so VMs is ready or available for next task that arrive within an hour.so, in EDF method first identified the deadline of task and identified the tasks whose deadline comes first and allocate resources to do that task before deadline.so it helps to allocates resource efficiently.
Priority Based Dynamic resource allocation in Cloud Computing [12], 2012.	The concerned paper explains priority-based resource allocation in which resources provides according to user need to satisfy the customer needs. This may provide resource allocation in priority basis and provided services guaranteed and reliable to user for the paid services. In this paper, parameters like memory, network bandwidth are explained and allocate resources to complete the task before allocated time. Resources are allocated on the behalf of user requirements after completing current task (Virtualization Model) VMs are free and ready to do next task and allocate these free resources to another task.
Resource Allocation in a Network-Based Cloud Computing Environment: Design Challenges [13], 2013.	Cloud computing becomes popular day by day. Each provider offers a unique service. In this paper, several internal and external factors that affect the working based on allocation model of resources as discussed. To move on cloud client demands guarantees with regards to improve the reliability of operations and cost because network performance and resource availability is very important in any cloud. Following are main design challenges such as geographical challenges, reliability for resources of network inside a DC (Data centers), SDN design challenges inside DCs. *Geographical and regulative challenges:* In Virtualization Model (VM) concerned client may not be able to manage their data on its physical location. For the security of data.it is necessary to stored encrypted data on clouds. *Reliability of network resource inside Data center:*

	Internal network design of Data Centers (DC) affects the performance and reliability of DC resource we can examines the topologies and traffic routing, flow optimizations. **Internal challenges:** *1. Data locality:* It is necessary for system to implement data locality. It means combining data (network) management, compute (processing) by consuming data locality factors that diminish the amount of data moment to improve the performance. 1. *Data aware schedule:* This gives some answers to questions such as • What criteria and polices for moving purpose of data? • What enforcement made by data integration policies? By understanding these challenges, technique improves the performance of data model. 2. *Analyze the behavior of data:* To understand the data moment and locality pattern, behavior pf data understanding is a good starting point. *Visibility:* Determine origin of the problem is challenge it slows the response. i.e., backup.
Resource Allocation in Cloud Computing [14], 2014.	Resource allocation is managing and sharing of assists to system and organization. Cloud computing is considered as ear's demand service. It offers guaranteed and reliable services in pay-as -your manner. It is very difficult to allocate resources or services accurately to user and satisfy user. It provides resources to client. Its aims to provide good and reliable quality services especially to the end user. Main objectives of cloud computing are minimum delay, response time, bandwidth in data transfer and minimum cost transfer data. With this user get good quality services with an affordable cost in business.
An efficient approach for green cloud computing using genetic algorithm [15], 2015.	Data centers basically provide hardware facilities that are configured properly to facilitate users. Data centers provides million virtual devices in cloud computing environment but it also addresses some type of problem that given below: 1. Heating problem 2. Cooling system cost 3. Power consumption Solution: Green cloud computing that supports to minimize less power consumption by reducing the use of less hardware resources in data center as possible. Methodology: 1. Virtual Machines: Use of virtual machines reduce the usage number of servers. 2. Power on/off techniques:

	3. Switch idle server on sleep mode. 4. Efficient Resource allocation (pooling mechanism) 5. Putting switches, routers, and network interfaces on sleep mode that are not in working state. 6. Efficient routing mechanism. These methodologies are used to prevent following allocation problems: 1. Resource contention: appears when two different types of applications make a struggle to approach same resource with in the same time. 2. Over-provisioning of resources: occurs when application gets more quantity of resources than the desired. 3. Under-provisioning of resources: occurs when application assigned to resources less than the requirement.
Modified Round Robin Algorithm for Resource Allocation in Cloud Computing [16], 2016.	Several policies for resource allocation are available in the field of cloud computing environment, this paper discusses about the Round Robin Algorithm. It is defined as: a shortest time interval specified by scheduler is called a quantum or time-slice. All processes that are in running state kept in a ready queue. In this ready queue, time slice is allocated by the CPU scheduler. Every running process executes in this time quantum by processor. If a process is completed in this time-slice and terminated otherwise, it added again to the tail of ready queue to be waiting for its next time slice to be executed till it completes itself. In Cloud Computing (Methodology): It is the same as we discuss in general above but a slight difference is time-slice is dynamic. Time-slice (quantum): total time of requests dividing by total number of request (jobs) resides in the ready queue. It requires two registers: SR: store the remaining rupture time of active jobs in early mentioned ready queue. AR: store the average of sum of total burst time dividing by complete jobs in mentioned ready queue. Next executing every demand in one time-slice both SR and AR are updating according to the new circumstances. Guidance: Time-slice should not to be too short or too large.
Optimal Cloud Computing Resource Allocation for Demand Side Management [17], 2016.	This research paper is totally about the cost-oriented model that is proposed to efficiently allocate the cloud computing resources. DSM it is abbreviation of demand side management so cloud computing improves the performance of system. There are two algorithms first one is simulated annealing and second one is modified priority list (MPL) both are developed to solve or minimize the optimization problems. **SA:** It solves difficult combinational and nonlinear problems or solve UC problems in power system. **MPL:** It is a time consuming and a modified priority list algorithms that is based on list of priority especially designed to solve the problems. **COM:**

	It helps to maintain the effectiveness and produce the COM model and reduce the total cost. the main purpose of these models to investigate and or describe the effectiveness of purposed model. we can reduce the cost by using these purposed models in this paper. Upon comparison of both algorithms, MPL algorithms can used COM model to work efficiently by considering uncertainty or without it. Moreover, COM has high potential in the context of smart grid.
A Hierarchical Framework of Cloud Resource Allocation and Power Management Using Deep Reinforcement Learning [18], 2017.	Reinforcement learning method provides the solution of resource allocation in cloud computing by making automatic decision. Reinforcement learning means taking stable decision to maximize reward or output in particular situation. resource management with deep reinforcement learning is shown in RL automatically allocate and schedule computer resources to waiting jobs so its main objective is to minimize job slowdown. So, project manager knows the importance of resource allocation so that he may not make any mistake while allocating resources. Power consumption has become important aspects in the field of cloud computing because its critical concern is about for design that reduces system s reliability and increases cooling cost. It slows down or shut down the system and affects its performance. DPM (distributed power management) means it allow reduce power consumption by power shuts off and on based on cluster utilizations.
Computation Offloading and Resource Allocation in Mixed Fog/Cloud Computing Systems with Min-Max Fairness Guarantee [19], 2017.	In this research work a min-max fairness base cost conservation algorithm is used to offload resource allocation by the decision making for problems occurring in Fog or Cloud computing. Allocation becomes optimize in the proposed system by intelligent decision making of bandwidth and power delay in energy consumption and helps to improve energy utilization. In this paper author explain low complexity algorithms and resource allocation algorithms (CORA) and that algorithm solve optimization problems. **SDR:** It is an abbreviation of Semi Definite Relaxation adopted in offloading decision making first time that increase the optimization in problem formulation to minimize the weighted cost of delay and energy consumption among all user equipment and focus in queue length and delay of user equipment's will be considered their future work.
Task scheduling and resource allocation in cloud computing using a heuristic approach [20], 2018.	Scheduling of tasks and allocation of resources are two different sides of a single coin that has prominent effects on working of each other, discussed in the given table.

Task Scheduling (When)	Resource allocation (What)
A process of assigning specific time interval to perform tasks in a specified assigning so that the available resources can be utilized properly.	A process of assigning resources to multiple jobs to be executed in an efficient manner without idle use of resources.
Task scheduler: It is feature that decides that when a process will gain a resource to be execute and when it will be release it.	Resource allocator: It is a feature that explain about the availability of resources. Moreover, it allocates the actual resource to a process.

Scheduler needs to consider following constraints: • Nature of the task • Task execution time • Task queue	Resource Allocator consider the followings: • Demand of resource of a process • Availability of a resource (Status)
Minimize overall response time	Efficient utilization of resources

Table 3: Scheduling VS Allocation

Machine Learning Based Resource Allocation of Cloud Computing in Auction [21], 2018.	In this paper, a machine learning classifier is used to analyze and model the multi-dimensional cloud resource allocation problem. The classifier includes two resource allocation algorithms that is based on logistic regressions and linear approach. **Resource allocation problem in an auction:** In the auction mechanism, each user request for a data set requirement (bid) and after checking the requirements data and cost model resources are allocated to the intended user. Bidding is based on no. of resources required by the user. **Multi-dimensional cloud resource allocation algorithm based on machine learning:** The main objective of this algorithm is to select all user's requirement to solve the given problem by optimal allocation solution along optimal price. For this purpose, it uses logistic and linear regression to find the optimal allocation solution, and used as a learning set for the model that predict all requirements of user. Assuming that there are n users submitting resource requirements, because of the limited resources, there is a need of an algorithm to allocate resources mention as below: • **Linear regression-based resource allocation prediction algorithm (Linear-ALLOC)** In this, the number of users those claimed for resource requirements smaller or equal to available resources wins. • **Logistic regression-based resource allocation prediction algorithm (Logistic-ALLOC)** First, mark the winning user as positive and failed user as negative sample. Existence of a hyper-plane in high-dimensional space distinct the winning user and failure user for the allocation.
Improving Quality-of-Service in Cloud/Fog Computing through Efficient Resource Allocation [22], 2019.	There is a large pool consist of multiple computers connected with each other through network over the internet. One computer/system access resources from different computers. User have to only pay for the services that they use and do not need for performing complex hardware operations. This may reduce the cost for users and known as computing (services, no hardware). Good policy is very important for the efficient resource allocations. Virtual machine placement model aims to minimize the resource allocation cost because with best resource allocation where multiple users' requests for limited resources available at physical level increases the performance. This paper discusses about efficient resource allocation.
Efficient dynamic resource allocation	This paper is suggested the effective dynamic resource that how to allocate to the cloud computing environment Imitate the interaction between customers and CRP. Independent tasks are set on the

method for cloud computing environment [23]. 2020.	available work set and virtual machines controls the CPU and RAM resources. The proposed DRA model specifies the key issues in the Spacing Multi-Objective Intelligence Algorithm (S-MOAL) that is improving. Three experiments were performed and produces good result but the first one is better that is S-MOAL, results in terms of cost, expiration date, error, tolerance, and energy consumption compared to DSOS PBACO, DCLCA and MOGA algorithms. Similarly, the experience of statistics proves the high improvement in the result of an effective response regarding the number of DSOS and MOGA algorithms. The virtual machines ongoing resource allocation is unused because of S-MOAL algorithm, Although, hunting was better than using the peritoneal method because it is safer and thus, get the closest solution of the proposed method is also valuable for the resource dynamic human age by adjusting unused VMs, which reduces energy. It is very important because this error takes into account the performance of tolerance that makes work on cloud computing flawless. In the future perspective the respective are used to involve Real Cloud Computing System.
Cloud Computing Virtualization of Resources Allocation for Distributed Systems [24], 2020.	In this research paper, there is a network designed to allow sharing of resources and online access to Information Technology Services (ITS), data processing tasks and centralized data storage. The proposed dynamic model of revenue details in cloud computing is a new method that uses VM description for user-based job-based features, which includes virtual resources. The proposed approach of this paper describes the algorithm, Ant Colony Optimization (ACO) performing load balancing task in cloud computing. The broker is responsible for handling the client's request. The cloud then sends the request to the data center and this request is treated by the Data Center Manager (DCM) that demands the reservation of specific number of VMs. Moreover, this method of computing uses an algorithm called Random algorithm that connects tasks to (cloudlets) and servers that allows to handle a large number of tasks at random and distribute the workload to each node. Load balancing based on a genetic algorithm hinders resource scheduling in the cloud environment comprises of dedicated models to allocate and allocate resources dynamically within the cloud data center. The proposed procedure aims to improve response time and delay in cloud service usage by distributing VM on server side. Data Center the first algorithm called DSJF is a combination of a power management technique called Dynamic voltage and frequency is scaling the Shortest Work First (SJF) resource allocation algorithm and Capac Power. The second algorithm, called DFCFS, is a combination of a DVFS, first come, first served (FCFS) resource. Cloud computing must be able to provide customers services at any time according to their demands. That's why the cloud uses virtualization and resource delivery. Techniques for allocating active resources and allowing load balancing in the data center. For this purpose, this paper focuses on many resources.
Genetic algorithm for quality of service-based resource allocation in cloud computing [25], 2021.	Cloud computing is an emerged field of information technology that is a combination of different computing-based resources in the cloud and also provides different types of services to its customers (users) like computational power, storage etc. But the major difficulty faced by its customers is the selection of the resources for their application deployment with cost effective policy and with best quality of service (QoS). This issue causes a big challenge for customers in resource selection. Cloud

service providers offers their services to their customers through **"pay-as-you-use"** policy. Their cost and QoS policies are complex regarding resource allocation.

Proposed Algorithm (Genetic algorithm):

In Genetic algorithm, the main components are QoS Performance and Deployment Cost in which only one service provider is used for both large and small scale of applications. This system gives the 32% reduction in the deployment cost and 5% in performance. Furthermore, it reduces the **resource rate**. For the problem formulation that increases the runtime performance uses directed acyclic graph in algorithm.

- **QoS performance:**

  There are two factors that include database instance response time and computing instance response time for better QoS performance.

  (1) The maximum value of response time specified by customer should be less than database instance value.

  (2) The maximum value of response time specified by customer should be less than computing instance value.

- **Cost optimization**

  The specified algorithm basically focused the best combination of cloud resources. Therefore, the variables are optimized in cloud resources. Application deployment cost into the cloud can compute by using following equation.

  $\delta = V1 + V2 + V3$

  Variable 1: Cost of the data base instance with respect to service time.

  Variable 2: Cost of the data base server with respect to service time.

  Variable 3: Cost of the Storage server with respect to service time.

  Working:

  Input: Customer QoS performance requirements.

  Step1: Compute Deployment cost.

  Step2: Compute database instance response time and computing instance response time.

  Output: Suitable Resources Allocation.

This proposed algorithm provides a discrete solution to the resource selection problem to deploy a customer application in the cloud by efficient balance between the deployment cost and QoS performance.

III. DISCUSSION

Cloud Computing is considered as the emerging technology using in best business markets and enterprises. A survey shows that resource allocation is the basic essential for providers of cloud deals with greater number of users in less response time. In field of cloud computing, efficient technique is required for resource allocation to users to maximize the user's satisfactions and profit for providers of cloud services.

IV. CONCLUSION

This paper summarizes the algorithms, techniques and methodologies uses in RAS in last decade of era. All service providers of cloud mainly focused on the resource allocation feature. Henceforth this concerned paper will prove as great motivation for future concerned researchers to grow up with secured and smarter optimal allocation of resources

framework and algorithms to strengthen the field of cloud computing standard.

ACKNOWLEDGMENT

REFERENCES

1. Armbrust, M., et al., *A view of cloud computing.* Communications of the ACM, 2010. **53**(4): p. 50-58.

2. Buyya, R., et al., *Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility.* Future Generation computer systems, 2009. **25**(6): p. 599-616.

3. Mell, P. and T. Grance, *The NIST definition of cloud computing.* 2011.

4. Zhang, S., H. Yan, and X. Chen, *Research on key technologies of cloud computing.* Physics Procedia, 2012. **33**: p. 1791-1797.

5. Zissis, D. and D. Lekkas, *Addressing cloud computing security issues.* Future Generation computer systems, 2012. **28**(3): p. 583-592.

6. Pallis, G., *Cloud computing: the new frontier of internet computing.* IEEE internet computing, 2010. **14**(5): p. 70-73.

7. Reddy, V.K., B.T. Rao, and L. Reddy, *Research issues in cloud computing.* Global Journal of Computer Science and Technology, 2011.

8. Zhang, L.-J., et al., *Hot topics in cloud computing.* IT professional, 2010. **12**(5): p. 17-19.

9. Sadashiv, N. and S.D. Kumar. *Cluster, grid and cloud computing: A detailed comparison.* in *2011 6th International Conference on Computer Science & Education (ICCSE)*. 2011. IEEE.

10. Grossman, R.L., *The case for cloud computing.* IT professional, 2009. **11**(2): p. 23-27.

11. Kumar, K., et al. *Resource allocation for real-time tasks using cloud computing.* in *2011 proceedings of 20th international conference on computer communications and networks (ICCCN)*. 2011. IEEE.

12. Pawar, C.S. and R.B. Wagh. *Priority based dynamic resource allocation in cloud computing.* in *2012 International Symposium on Cloud and Services Computing*. 2012. IEEE.

13. Sharkh, M.A., et al., *Resource allocation in a network-based cloud computing environment: design challenges.* IEEE Communications Magazine, 2013. **51**(11): p. 46-52.

14. Sudeepa, R. and H. Guruprasad, *Resource allocation in cloud computing.* International journal of modern communication technologies and research, 2014. **2**(4): p. 265808.

15. Kaur, B. and A. Kaur. *An efficient approach for green cloud computing using genetic algorithm.* in *2015 1st International Conference on Next Generation Computing Technologies (NGCT)*. 2015. IEEE.

16. Pradhan, P., P.K. Behera, and B. Ray, *Modified round robin algorithm for resource allocation in cloud computing.* Procedia Computer Science, 2016. **85**: p. 878-890.

17. Cao, Z., et al., *Optimal cloud computing resource allocation for demand side management in smart grid.* IEEE Transactions on Smart Grid, 2016. **8**(4): p. 1943-1955.

18. Liu, N., et al. *A hierarchical framework of cloud resource allocation and power management using deep reinforcement learning*. in *2017 IEEE 37th international conference on distributed computing systems (ICDCS)*. 2017. IEEE.

19. Du, J., et al., *Computation offloading and resource allocation in mixed fog/cloud computing systems with min-max fairness guarantee.* IEEE Transactions on Communications, 2017. **66**(4): p. 1594-1608.

20. Gawali, M.B. and S.K. Shinde, *Task scheduling and resource allocation in cloud computing using a heuristic approach.* Journal of Cloud Computing, 2018. **7**(1): p. 1-16.

21. Zhang, J., et al., *Machine learning based resource allocation of cloud computing in auction.* Comput. Mater. Continua, 2018. **56**(1): p. 123-135.

22. Akintoye, S.B. and A. Bagula, *Improving quality-of-service in cloud/fog computing through efficient resource allocation.* Sensors, 2019. **19**(6): p. 1267.

23. Belgacem, A., et al., *Efficient dynamic resource allocation method for cloud computing environment.* Cluster Computing, 2020. **23**(4): p. 2871-2889.

24. Shukur, H., et al., *Cloud computing virtualization of resources allocation for distributed systems.* Journal of Applied Science and Technology Trends, 2020. **1**(3): p. 98-105.

25. Devarasetty, P. and S. Reddy, *Genetic algorithm for quality of service based resource allocation in cloud computing.* Evolutionary Intelligence, 2021. **14**(2): p. 381-387.

BIOGRAPHY

MAIRA KHALID received the MCS degree from Bahauddin Zakariya University in 2015, and completed MS Computer Science from Air University in 2018. From 2016 to 2018, she was lecturer of the Computer Science Department in City College of Science and Commerce. Currently, she is working as Lecturer of Computer Science Department in Institute of Southern Punjab and her research includes machine learning, cloud computing, neural network and IoT.

UZMA AZIZ, she is a student of B.S. Information technology, currently working on Final Year Project, in City College of Science and Commerce, Multan, Pakistan. She is interested in research work in flow and storage of data/information throughout the IT infrastructure, resource management, internet of things and machine learning. She worked on cloud computing allocation of resources depending on time, storage capacity, and priority, moreover

comparison of previous and modern technology in this research work.

FAIZA LATIF, she is a student of B.S. Information Technology, currently working on Final Year Project, from City College of Science and Commerce, Multan, Pakistan. Her research interests include real Time sentiment Analysis, resource allocation management, and deep Learning. She concludes this research in terms of a last decade survey.

SAMREEN KANWAL, she is a student of B.S. Information Technology, in City College of Science and Commerce and working on Final Year Project. Her research area includes data flow infrastructure, resource allocation management, and deep learning. She worked on allocation of resources in cloud computing depending on user requirement, additionally, analyzing previous and modern technology in this research work.