A study on Optimal Resource Allocation Policy in Cloud Environment

Archana Naik[a,*], Kavitha Sooda[b], KarunakarRai[c]

[a]Department of Computer Science & Engineering, NitteMeenakshi Institute of Technology, Yelahanka, Bangalore,560064, India

[b]Department of Computer Science & Engineering, B. M. S. College of Engineering, Basavanagudi, Bangalore, 560019, India

[c]Department of Electronics & Communication Engineering, Nitte Meenakshi Institute of Technology, Yelahanka, Bangalore,560064, India

[*]Corresponding author: archana.naik@nmit.ac.in

Abstract: This paper presents the role of the resource allocator, the characteristics that are taken into account while allocating resources and the resource allocation mechanisms. The cloud computing enables the utility-based resource allocation to the customers. The resource allocation mechanism needs to consider the better utilization of the resources and the minimized cost per service to the customer. These condition for resource allocation expects the resource allocation to be viewed as the optimization problem. To perform the optimized allocation, the allocator must be equipped with the status of the resources, predicted demand from the customers, and dynamic changes in the cloud. In this work, we have categorized the parameters for the resource allocator based on the cost of allocation, utilization of the resources, time taken to execute and reliability. An architecture for the resource allocator is proposed considering the service level agreement, the user request, and the resource status. A framework to build the resource allocator model is proposed.

Keywords: Deep Reinforcement Learning, Markov Decision Process, Resource Allocation Algorithms.

---

1. Introduction

The cloud is collection of resources, which are distributed and interconnected over the Internet. In the cloud computing, resources such as CPU, memory, disk, bandwidth, etc. are shared by allocating Virtual Machines (VM) to the consumers. The VMs are equipped with the resources as per the requirements of the user application. The cloud provides the utility-based commodity resources, allowing the cost-effective provisioning of the resources. This per-demand allocation provides cost-effective provisioning of the resources. The allocation of VMs is dependent on the service level agreement between the service provider and the consumer.The cloud resource allocator maintains the heterogenous types of resources and the demand for resources that vary with the wide range of consumer applications.The resource provisioning in cloud manages the scalability of the dynamic variations in the workload requirements. The cloud provides flexibility to the consumers based on the application and resource availability. The dynamic range of resource types and the diverse applications create the challenge of providing an effective mechanism for an agile and efficient tool for allocating resources [1-4]**.**

The task of distributing resources optimally and utilizing resources efficiently presents a significant challenge to the resource manager. The environment includes variety of users, with different applications, with different resource requirements, varying with time. To allow the faster and fair allocation of resources and the mapping of the physical resources to the virtual resources that are allocated to the user, the dynamic technique of resource allocation is essential[3]. A number of papers have been published on the study of resource allocation. The major parameter for consideration during the allocation of resources is power management and performance efficiency. Other parameters considered during resource allocations (RA) are type of the resources, resource parameters, different algorithms for scheduling and the cloud architecture. For the dynamic allocation of resources, the resource allocator must be aware of the status of the resources. It is observed that due to the real time changes in the requirements and the availability of resources, efficient usage is improved by considering the RA problem as an optimization problem.

This study is performed to learn the resource allocation techniques and methodology available in the literature. The rest of the paper is organized as follows: Sections 2, provides the background on the resource allocation, section 3 provides the study considering the parameter used for resource allocation, section 4 provides the resource allocation strategies/policy available in the literature. Section 5 provides the proposed framework and section 6 conclusion.

2.Related Work

The resource allocatorhas become a major component of the cloud resource manager. The cloud application, the services, and the resources are distributed with different objectives in the cloud environment. The cloud environment includes copious resources such as processing units, memory, storage for user applications, the needs of which are increased or decreased depending on the demand. The primary intent of the resource allocation strategies is to achieve better revenue with maximum resource utilization for the cloud resource provider. At the user's end, the intent is to meet the expectations with minimal expense.

The study is carried out by considering the policy, strategy and parameters considered for resource allocation in cloud computing. The review papers and the papers presenting the implementation separated. The review papers are reviewed in this section. The review on other paper is discussed in section 3 and section 4.

The authors of [5], present the review of the RA mechanisms. A classification model is proposed. The authors have presented analyzed results considering the research gap between the existing mechanisms and the future research topics. The classification is provided on mechanism type as static or dynamic, processing mode as centralized or distributed, QoS as in resource utilization, energy utilization, response time and cost, objectives as in single objective or multi objectives, Service Level Agreement (SLA)-based as in adaptive or non-adaptive, policy based as in execution time, VM based, gossip, utility based, auction or application based, method of evaluation as in implementation or simulation, use of optimization algorithms such as heuristic, meta heuristic, non-heuristic, cloud domain such as inter or intra cloud.

The authors of [6], provide the survey on RA on cloud and provide the category of the allocation as agent oriented, priority based, quality based, dynamic and QoS based allocation techniques. The work in [7], provides the study on the elasticity in the cloud.

3.Resource Allocation Parameters

The QoS parameters are chosen by the consumer and the cloud provider, which become part of the SLA. While allocating resources the RA will analyze the parameters and the availability of resources and then allocate the resources. In the literature we find several work on the parameters considered. The resource allocation policy plays an important role in allocating resources [8].

In this section we have presented the study of parameters that can be considered and the various RA policies that are available in the literature. The resource requirements can be categorized as network requirements such as bandwidth, delay, throughput, and computational requirements including CPU, memory [9]. We observed that the few authors have considered any one of the parameters for RA, wherein some others have observed more than one parameter into consideration to decide on the best resource to be allocated to that request. They are presented as single objective and multi objective RAs.

The RA policy followed must consider minimizing the resource allocation cost, the overall system utilization, and the job execution time. These considerations lead us to the RA as an optimization problem, with the aim of reducing the total cost while introducing the idea of increasing the overall reliability. The reliability considers failure as one of the parameters to be considered for the RA. The proposed RA strategies is based on the categories of resource allocation cost, system utilization, job execution time, reliability. The parameters that are considered for each of these categories are listed in fig 1.

3.1 Single Objective:

The CPU is one important matric to be considered for RA and performance, as it is the significant contributor to the resource demand. The prediction mechanism also becomes important because of the dynamic variations in the CPU utilization. A Recurrent neural network (RNN) is applied predict the utilization of CPU. The network is trained by Back-Propagation-Through-Time (BPTT) algorithm. The advantage of BPTT is that it enables storing of the past information, to be utilized for the sequential models. The input for recurrent network includes the current and the previous CPU utilization values. BPTT mechanism allows the prediction of CPU utilization in multiple steps further ahead in the future. The RNN application a demonstrates reasonable level of accuracy for performance of RNN for newly presented data [10].

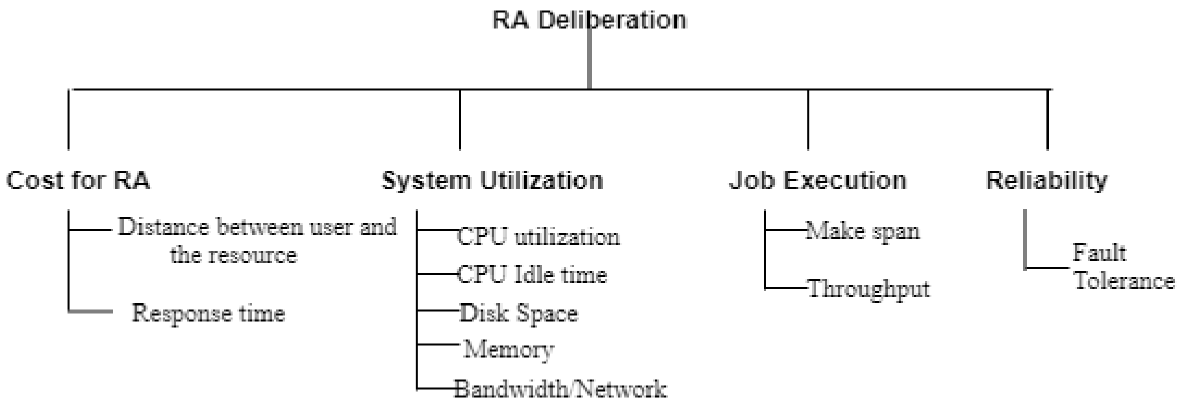

Fig 1: RA Deliberation

The amount of time that the CPU is idle has become just as significant as the amount of time that the CPU is really used. The allocator should also examine how long the CPU has been idle. The dynamic idle interval prediction scheme is proposed [11], which predicts the CPU idle interval length and to reduce power usage, chooses the most cost-effective sleep state on the go. The CPU idle interval lengths are predicted using performance metrics such as instructions per cycle, cache miss rates, structure occupancies, branch predictor statistics. The power states are identified for the CPU in the data center as low power or sleep states, based on the memory or input/output operation. The power consumption or savings in each of the state is identified. Each state has distinct power consumption rate and the different wakeup time. After the idle state, CPU can wake up to any state. However, for the entering state, if the wake-up latency is greater than the idle state length, it will degrade the power savings and the performance.

Time-critical applications demand a faster response from the cloud. In such situations, bandwidth plays an important parameter for the cloud’s performance. Time series analysis providesan estimation that can be applicable for specific period. The bandwidth estimation taken into consideration with the estimation of the workflow execution, provides real-time bandwidth estimation. The proposed mechanism introduces the deflating multiplier value U, which is applied to the estimated available bandwidth. The suggested process derives a U value for a workflow based on a needed deadline, an estimate of available bandwidth, and bandwidth variability, which represents the available bandwidth's uncertainty [12].

The failure rate at any node is the parameter that will affect the execution of all other tasks that are assigned to that particular resource. The failure of allotted resource leads to the failed or delayed execution of the tasks assigned at that resource. This indicates that one of the RA parameters should be the failure rate of cloud resources. The stochastic occurrence of the failure in cloud is represented using the Poisson probability distribution. The algorithm calculates the threshold for the virtual machines and calculates the workload of each virtual machine. For each virtual machine expected failure rate and the probability of failure is calculated. For assigning the resources for the selected task from the queue, the virtual machine with the minimum probability is selected. To calculate probability of failure current workload on the virtual machine and the failure rate are considered. The other parameters are not considered here [13].

In big data applications where the processing of any query requires large data available different nodes, data retrieval becomes an important parameter. With the big data storage technology, the replicas of the data block are distributed across multiple nodes. Select the node with replica is selected for retrieval, considering the path for all the requests and minimizing data retrieval time. Since there are multiple applications demanding the same or multiple data, a threshold is considered for the data retrieval. The penalty is added to the data retrieval, considering the mean of the threshold and the retrieval time. The nodes and the path are selected such that the penalty is minimized[14].

3.2 Multi Objective

The literature presents the several work where in more than one parameter is taken into consideration for RA. The availability of CPU and the memory are considered to make the decision to allocate node to the user request. A framework named Nara (Network-Aware Resource Allocation Algorithms) is presented in [15]. Graph Neural Network is used to represent the graph of servers that can be allocated to the user requests in Data Centre Networks (DCN), representing the available actions for the allocation. The graph consists of server and the nodes. The node feature includes the CPU and memory resources available at the node. Edge of the graph represents the available bandwidth. The user request is associated with parameters as compute, memory, bandwidth and holding period. The state represents the feature containing current utilization of the CPU and memory the DCN (averaged over each server) and the holding-period of the awaiting request. Reinforcement Learning is used to build the policy network which choose servers to be allocated to the incoming requests. The policy network applies the action representation and the environment's state, learns to make decisions.

Dynamic prediction of resource utilization can be used to supportbetter decision making for RA strategies. Authors of [16] propose a dynamic resource utilization model, that takes into account CPU, memory, and disk spaceutilization. The prediction mechanism using Over Produce and Choose approach is implemented using stacking mechanism with 2 levels. The level 0 of the stack contains K-Nearest Neighbor and Decision Tree algorithm and the level 1 contain Decision Tree algorithm. Level 0 is used to learn and level 1 is used for testing. Request traces and the usage traces are considered on the periodic basis for the prediction. The resource allocation process can be automated.

4.Resource Allocation Policy

Irrespective of which parameters are considered to monitor the resource request, the RA policy or the strategy defined would ensure the optimal allocation of resources. This section provides a review of the literature on policy and strategy implementation for resource allocation. The demand for resources in the cloud is dynamic. The strategies for allocation also need to be dynamic. Table 1 represents the resource allocation policy being presented.

Table 1: Resource Allocation Policy

Algorithm	Parameters	Method	Performance evaluation
Machine learning [17]	Reliability, performance, and energy consumption on data center	The impact of failure on the parameters. The resources are classified to avoid allocating unavailable and unreliable resources to the customers.	Reduced energy consumption with reliable service
Machine learning [18]	Request, bid value	Resource auction	Optimal resource allocation
Machine learning [19]	CPU, Memory	SVM	Classification of faulty machines from working machines
LSTM [20]	CPU, memory usage	Predict resource utilization, by determining correlation between the requested and used resources is determined	Eased over allocation and under allocation of resources
ARIMA [21]	Resource utilization, execution time, QoS violations.	Autocorrelation between the workload demand	The workload analyzer used to learn the recent changes in the workload
Neural networks [22]	Historical, live resource usage	Resource demand prediction	Maximized resource utilization, minimized performance degradation
Markov decision process [23,24]	Response time, power consumption	Central management maintain the heterogenous server status and when received the client request, allocate the optimal number of cores on the servers.	A balance between service request response and power usage
Reinforcement learning -Q learning [25]	Task under execution, user requests	Online learning of the resource availability and automatic decision to allocate the resources in cloud robotics	Numerical analysis

Reinforcement learning [26]	Power consumption, resource request	The resource allocator is at the cloud data center, senses the state of the cloud environment, takes action to investigate the implication of actions over the current state	Power usage, data center infrastructure efficiency, CPU utilization at data center
Deep Reinforcement learning [27]	Workload prediction through inter job arrival time prediction	Constructing the hierarchical framework LSTM at the local tier, DRL at the global tier. Q value estimation is performed using autoencoder	Improved power consumption
Deep Reinforcement learning [28]	Energy management, demand response, electricity market, operational control	Review of the DRL applications in power system management	DRL has opportunities and challenges in power system management

Machine learning avoids human interference in resource allocation and scheduling. The algorithm performs the optimal resource allocation, with a possible improvement in cloud providers' profit. Linear regression and Bayesian networks are implemented to perform the classification based on the SLA and the performance cost [29]. The Support Vector Machine (SVM) is used to classify the machines which are removed due to failure or for maintenance from fully working machines. Utilization of CPU and memory are considered for classification. To avoid allocating unavailable and unreliable resources, the optimal allocation scheme is selected based on the fitness value. Where the fitness value is computed as the mean time between the failures. Where the fitness value is computed as the mean time between the failures. Linear regression and logistic regression are used to compute the auction and find the optimal solution to the resource allocation in auction-based resource allocation.

The Long Short-Term Memory (LSTM) model is used in predicting the resource utilization for each application, considering the CPU and memory usage. The correlation between the requested and used resources is determined. The implementation results in the easing of the overallocation and under allocation of resources.Another model built to predict workload prediction and its effect on the QoS is the Auto Regressive Integrated Moving Average (ARIMA) model. It is a proactive method for dynamic workload prediction. The autocorrelation between the workload demand and the workload is utilized. The workload analyzer is used to learn about the recent changes in the workload.

A generic method termed the threshold-based tournament selection method is implemented to allocate the virtual machines. A genetic algorithm is used for the implementation. The request size is reduced to the highest instruction size of the virtual machine or the resource and converted to binary to form a gene. The virtual machine's processing power is also converted to a gene. The chromosome is created by merging a binary request with a binary converted, randomly chosen virtual machine. For the threshold-based tournament selection, two randomly selected individuals perform the tournament to determine the winner. The resource corresponding to the winner is selected as the best fit resource for allocation [30].

The Markov Decision Process (MDP) provides the model-building framework for an optimal decision-making process. The MDP is defined with four tuples as (S, A, P_a, R_a), where S is the set of states in state space, A is the set of actions from action space, and P_a is the probability that action a in state s at time t will result in state s′ at time t +1, as mentioned by Pr in (1). R_a is the immediate or expected reward received after transition from state s to state s′ . The policy function π, provides the probabilistic mapping from state space S to action space A. The MDP is combined with the policy, fixes the action for each state in the state space. This behaves like a Markov chain. In MDP, the purpose is to identify an appropriate policy for the decision maker.

$$P_a(s,s') = Pr(s_{t+1} = s' \mid s_t = s, a_t = a) \qquad (1)$$

The $\pi(s)$, specifies the action at time t as a_t, such that it will maximize the cumulative function of the random reward. The long tern cumulative reward after time t to be given as Rt.

$$R_t = \sum_{t=0}^{\infty} \gamma^t R_{a_t}(s_t, s_{t+1}) \qquad (2)$$

The ϒ is the discount factor, with the value $0 \leq \gamma \leq 1$. The smaller discount factor supports faster decision. The optimal policy $\pi^*$is the one that maximizes the cumulative reward. The $a_t = \pi(s_t)$, action given by the policy at time t, over the probability $P_{a_t}(s_t, s_{t+1})$.

To find the optimal policy value function $V(s)$ representing the cumulative sum of rewards from state start state to state s, on policy π.

$$V(s) = \sum_{s}^{s'} P_{\pi(s)}\,(s, s')\left(R_{\pi(s)}(s, s') + \gamma V(s')\right) \qquad (3)$$

$$\pi(s) = argmax_a \left\{\sum P(s'|s, a) \;\; (R(s'\,|\,s, a) + \gamma V(s'))\right\} (4)$$

Determining the optimal policy for MDP is based on more of the Reinforcement Learning theory [33]. Q-Learning is a reinforcement learning algorithm that uses a Q function to find the best action-selection policy. The Q-learning update rule is given as

$$Q(s, a) = \sum_{s}^{s'} Q(s, a)\,\alpha[R_a\,(s, s') + \gamma V(s')] \qquad (5)$$

Reinforcement learning (RL) based learning considers different characteristics of all the types of information available.This is a useful characteristic for autonomous management of resources.The value function is computed for each parameter. The value function and the corresponding action to be stored in the lookup table. Artificial neural networks are used to implement the learning environment. The advantage of neural networks is that they can memorize the value functions without the demand for additional storage. The limitation of this work is that the status of the resources is determined, the resources are allocated to the demanding robots only if the demand can be served. Otherwise, the robots do the processing locally [31].

The RL-based multi-object optimization method optimises RA energy consumption and saves power.Trade-offs between performance and energy. During the process, SLA violations can take place. To improve the performance, power consumption and SLA violations are considered here. Power consumption is observed for the improved energy savings. The algorithm optimally chooses the request host pair [32].

In RL, the Q values are stored in a table. For real-world applications, storing all the continuous state space and action space on the table becomes a design issue. To solve this problem, function approximation is used to extract features from models, value functions, or policies, and then utilise deep neural networks (DNN) to generate an approximation of the complete function. The RL considers transitions based on the probabilities of higher cumulative reward. In deep reinforcement learning (DRL), the exploration is performed when the learning opportunities are worthwhile for the agent to perform without a separate training phase. The majority of DRL algorithms are model-free and suitable for scenarios that cannot be modelled. The features can be extracted using these function approximators.

5.Proposed Framework

The cloud resource allocator must maintain the QoS, while providing the resources to the consumer. The QoS will be considered based on theSLA, between the user and the cloud service provider [33]. The cloud architecture provided in fig 1 depicts the RA which receive the request from the consumer, access the SLA and the status of the virtual machine and the resources. Resource allocation Policy refers to the set of principles that guide the decision making. To implement the policy, mechanisms are defined [34].

The RA policy practiced must consider minimizing the resource allocation cost, the overall system utilization, and the job execution time. These considerations lead us to the RA as an optimization problem, with aim of reducing the total cost while introduce the idea of increasing the overall reliability [35]. The reliability considers failure as one of the parameters to be considered for the RA.

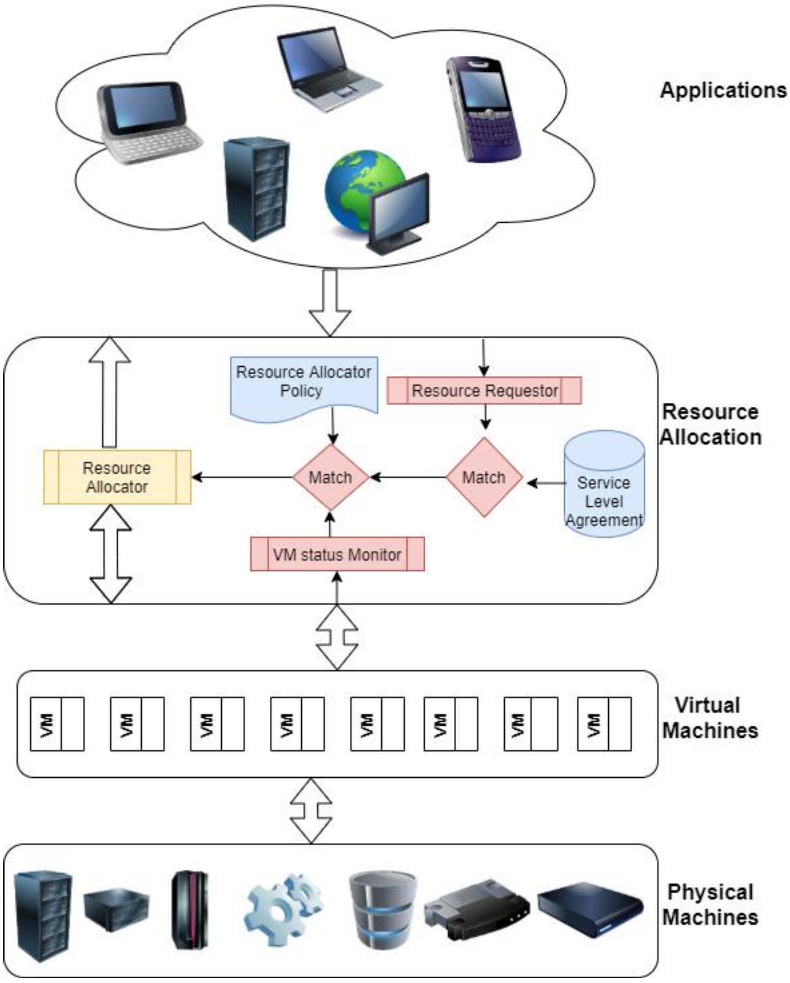

Fig 1: Cloud Architecture

The steps followed by the resource allocation process is

Step 1: The resource request is received from the application
Step 2: The request is matched with the SLA
Step 3: If the requested resource allowed as per the SLA, continue, otherwise stop the comparison
Step 4: Pass the resource request to the resource allocator.
Step 5: The resource allocator considers the resource allocation policy and the status of the VM, gives the decision on the allocation of the resources to the request

The MDP model is proposed to build the model. Deep DRL is to be used to build the framework for dynamic learning and allocation of resources.

6.Conclusion

In this study, we have presented a study on resource allocation strategies. Cloud resources are scalable and on demand, which makes the resource allocation an optimization problem. The study presents the parameters that are considered to allocate the resources and different resource allocation strategies. In the literature, work on RL and DRL is observed, where majority of the work considers power management and energy as the primary parameters of concern. By considering the dynamic status of the resources and the SLA for allocating the resources, the resource allocator can provide a better cost per performance to the cloud service provider and the user.

References

[1]. Singh, S. and Chana, I., 2016. A survey on resource scheduling in cloud computing: Issues and challenges. Journal of grid computing, 14(2), pp.217-264.

[2]. Liu, N., Li, Z., Xu, J., Xu, Z., Lin, S., Qiu, Q., Tang, J. and Wang, Y., 2017, June. 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) (pp. 372-382). IEEE

[3]. Souravlas, S. and Katsavounis, S., 2019. Scheduling fair resource allocation policies for cloud computing through flow control. Electronics, 8(11), p.1348.

[4]. Souravlas, S., Katsavounis, S. and Anastasiadou, S., 2020. On Modeling and Simulation of Resource Allocation Policies in Cloud Computing Using Colored Petri Nets. Applied Sciences, 10(16), p.5644.

[5]. Fard, M.V., Sahafi, A., Rahmani, A.M. and Mashhadi, P.S., 2020. Resource allocation mechanisms in cloud computing: a systematic literature review. IET Software, 14(6), pp.638-653

[6]. Computing-Palarch's, M.I.C., An Extensive Survey On Resource Allocation Mechanisms In Cloud Computing.

[7]. Bharanidharan, G. and Jayalakshmi, S., 2021, March. Elastic Resource Allocation, Provisioning and Models Classification on Cloud Computing A Literature Review. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1909-1915). IEEE

[8]. Aljoumah, E., Al-Mousawi, F., Ahmad, I., Al-Shammri, M. and Al-Jady, Z., 2015. SLA in cloud computing architectures: A comprehensive study. Int. J. Grid Distrib. Comput, 8(5), pp.7-32.

[9]. Ankita, V., 2013. A survey on various resource allocation policies in cloud computing environment. Int. J. Res. Eng. Technol, 2(5), pp.760-763.

[10]. Duggan, M., Mason, K., Duggan, J., Howley, E. and Barrett, E., 2017, December. Predicting host CPU utilization in cloud computing using recurrent neural networks. In 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST) (pp. 67-72). IEEE

[11]. Duan, L., Zhan, D. and Hohnerlein, J., 2015, June. Optimizing cloud data center energy efficiency via dynamic prediction of CPU idle intervals. In 2015 IEEE 8th International Conference on Cloud Computing (pp. 985-988). IEEE.

[12]. Genez, T.A., Bittencourt, L.F., da Fonseca, N.L. and Madeira, E.R., 2015. Estimation of the available bandwidth in inter-cloud links for task scheduling in hybrid clouds. IEEE Transactions on Cloud Computing, 7(1), pp.62-74

[13]. Sathiyamoorthi, V., Keerthika, P., Suresh, P., Zhang, Z.J., Rao, A.P. and Logeswaran, K., 2021. Adaptive Fault Tolerant Resource Allocation Scheme for Cloud Computing Environments. Journal of Organizational and End User Computing (JOEUC), 33(5), pp.135-152

[14]. Xie, R. and Jia, X., 2018. Data transfer scheduling for maximizing throughput of big-data computing in cloud systems. IEEE Transactions on Cloud Computing, 6(1), pp.87-98.

[15]. Shabka, Z. and Zervas, G., 2021. Nara: Learning Network-Aware Resource Allocation Algorithms for Cloud Data Centres. arXiv preprint arXiv:2106.02412.

[16]. Mehmood, T., Latif, S. and Malik, S., 2018, October. Prediction of cloud computing resource utilization. In 2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT) (pp. 38-42). IEEE

[17]. Liu, D., Sui, X., Li, L., Jiang, Z., Wang, H., Zhang, Z. and Zeng, Y., 2018. A cloud service adaptive framework based on reliable resource allocation. Future Generation Computer Systems, 89, pp.455-463.

[18]. Zhang, J., Xie, N., Zhang, X., Yue, K., Li, W. and Kumar, D., 2018. Machine learning based resource allocation of cloud computing in auction. Comput. Mater. Continua, 56(1), pp.123-135.

[19]. Khalil, M.H., Sheta, W.M. and Elmaghraby, A.S., 2016, December. Categorizing hardware

failure in large scale cloud computing environment. In 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) (pp. 327-332). IEEE

[20]. Thonglek, K., Ichikawa, K., Takahashi, K., Iida, H. and Nakasan, C., 2019, September. Improving resource utilization in data centers using an lstm-based prediction model. In 2019 IEEE International Conference on Cluster Computing (CLUSTER) (pp. 1-8). IEEE.

[21]. Calheiros, R.N., Masoumi, E., Ranjan, R. and Buyya, R., 2014. Workload prediction using ARIMA model and its impact on cloud applications' QoS. IEEE transactions on cloud computing, 3(4), pp.449-4

[22]. Saxena, D., Singh, A.K. and Buyya, R., 2021. OP-MLB: An online VM prediction based multi-objective load balancing framework for resource management at cloud datacenter. IEEE Transactions on Cloud Computing

[23]. Tournaire, T., Castel-Taleb, H. and Hyon, E., 2021. Optimal control policies for resource allocation in the Cloud: comparison between Markov decision process and heuristic approaches. arXiv preprint arXiv:2104.14879

[24]. Wang, Y., Chen, S., Goudarzi, H. and Pedram, M., 2013, March. Resource allocation and consolidation in a multi-core server cluster using a Markov decision process model. In International Symposium on Quality Electronic Design (ISQED) (pp. 635-642). IEEE.

[25]. Liu, H., Liu, S. and Zheng, K., 2018. A reinforcement learning-based resource allocation scheme for cloud robotics. IEEE Access, 6, pp.17215-17222.

[26]. Thein, T., Myo, M.M., Parvin, S. and Gawanmeh, A., 2020. Reinforcement learning based methodology for energy-efficient resource allocation in cloud data centers. Journal of King Saud University-Computer and Information Sciences, 32(10), pp.1127-1139.

[27]. Liu, N., Li, Z., Xu, J., Xu, Z., Lin, S., Qiu, Q., Tang, J. and Wang, Y., 2017, June. 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) (pp. 372-382). IEEE.

[28]. Zhang, Z., Zhang, D. and Qiu, R.C., 2019. Deep reinforcement learning for power system applications: An overview. CSEE Journal of Power and Energy Systems, 6(1), pp.213-225.

[29]. Hormozi, E., Hormozi, H., Akbari, M.K. and Javan, M.S., 2012, November. Using of machine learning into cloud environment (a survey): managing and scheduling of resources in cloud systems. In 2012 Seventh International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (pp. 363-368). IEEE.

[30]. Oladoja, I.P., Adewale, O.S., Oluwadare, S.A. and Oyekanmi, E.O., 2021. A Threshold-based Tournament Resource Allocation in Cloud Computing Environment. Asian Journal of Research in Computer Science, pp.1-13.

[31]. Peng, Z., Cui, D., Zuo, J., Li, Q., Xu, B. and Lin, W., 2015. Random task scheduling scheme based on reinforcement learning in cloud computing. Cluster computing, 18(4), pp.1595-1607.

[32]. Shrimali, B. and Patel, H., 2020. Multi-objective optimization oriented policy for performance and energy efficient resource allocation in Cloud environment. Journal of King Saud University-Computer and Information Sciences, 32(7), pp.860-869.

[33]. Buyya, R., Yeo, C.S. and Venugopal, S., 2008, September. Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. In 2008 10th IEEE international conference on high performance computing and communications (pp. 5-13). Ieee.

[34]. Marinescu, D.C., 2017. Cloud computing: theory and practice. Morgan Kaufmann.

[35]. Souravlas, S., Katsavounis, S. and Anastasiadou, S., 2020. On Modeling and Simulation of Resource Allocation Policies in Cloud Computing Using Colored Petri Nets. Applied Sciences, 10(16), p.5644.