A Comparative Study between Proposed Hybrid Load Balancing Algorithm and Traditional Algorithm to Reduce Response Time and Processing Time on Cloud Computing

Sarita Yogi¹, Ritesh Yadav²

¹²Department of Computer Application SRK University, Bhopal, Madhya Pradesh, India

Abstract: In this work, we present a hybrid algorithm which attempts to enhance the system's typical response time with data processing time inside a cloud environment. The findings after comparing the proposed algorithm against algorithms with Maxmin, Min-Min, Round-Robin, shown our algorithms significantly optimized response times. With the help of the CloudAnalyts modelling tool, we were able to run the proposed algorithm, which has enhanced the cloud datacentre's processing as well as response times. The efficiency of the created hybrid algorithm is assessed for a variety of simulated situations inside a cloud computing system with diverse user bases with datacentre architectures and varying number of virtual machines to check the validity of proposed algorithm. The proposed algorithm gives better response and data processing time compared to the already established algorithms.

Keywords: Algorithm, Load Balancing, Processing Time, Response Time, Round-Robin, Scheduling.

INTRODUCTION

The most crucial aspect of cloud computing which can spread the load among multiple nodes within the work area is load balancing. It guarantees that no node throughout the system is constantly heavily, softly, or not at all laden[1], increasing the system's entire efficiency. The fundamental goal of load balancing with cloud computing is to increase availability and dependability through redundancies and ramp up resources as needed[1].

With the help of virtual computing resources, cloud computing, a web-based network technology, has contributed to the rapid development of technological communication by serving clients with a range of needs. It includes resources including software development systems, testing tools, as solutions for both software hardware programmes[2][3]. Services are used to carry out this resource delivery. Although the latter two fall under the categories of platforms as a service (PaaS) and software's as a service (SaaS), accordingly, the former falls under the classification of infrastructures as a services (IaaS) cloud[4]. The cloud computing concept uses an on-demand infrastructure where assets are shared as services which are paid on a pay-as-you-go (PAYG) basis[5]. Amazon, Microsoft, the search engine, SAP, Oracle, VMware, Sales Force, IBM, along with other major players in the relevant technology represent a few of the major participants. Most of the aforementioned cloud providers are cutting-edge IT companies. There are two perspectives on the cloud computing architecture. The first one includes the service distribution model, which outlines the kind of services that a typical cloud provider would typically supply. Based on this factor, the SaaS, PaaS, with IaaS model of service among the three most often used ones.

When resources are used as effectively as feasible, a cloud computing model considered effective, and this effective use can be done by implementing and maintaining adequate administration of cloud resources. By using effective resource scheduling, distribution, and scalability approaches, resource management can be accomplished. Through a procedure called virtualization, which makes use of a thing (software, gadgets, or both) described as a hypervisor, such assets are made available to clients in the manner of Virtual Machines (VM)[6]. The most significant benefit of cloud computing involves the conversion of one user's physical machine onto a multi-user's virtual machine[7]. With the limited virtual resources that are now accessible, the Cloud Service Provider (CSP) has a significant impact on how services are delivered to consumers. Some VMs will experience a high volume of user tasks whereas delivering user requests, while others will experience a lower volume. The Cloud

ISSN: 2321-8169 Volume: 11 Issue: 11

Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 November 2023

Providers (CSP) is consequently left using unbalanced machines that have a wide gradient of user demands and resource utilisation[8].

Cloud computing provides clients with online access to networked computer resources utilising distributed and parallel processing methods. Cloud computing has almost become democratised thanks to the "pay-as-you-go" company model. In this stage of software installation, service providers, customers, and cloud providers all take part. Through virtual machines (VMs), cloud suppliers give their clients computing capabilities. When providing application-level customer support, service providers make use of these virtual computers. Task scheduling algorithms are used by service providers to distribute client workloads among virtual machines, speed up response times, guarantee excellent quality of service (QoS), and make the most of available resources. Due to this, every cloud architecture must include the work scheduling algorithm.

Due to this, each cloud architecture must include the work scheduling algorithm. The numerous scheduling mechanisms used in diverse computer systems must be modified in order to accommodate cloud computing. Some scheduling approach that has been cluster-optimized might not work well in the cloud. The components of the approach need to be transferred into the issue space until the algorithm is capable of handling the framework of the cloud's setting. The amount of work configurations accessible increases with the number of virtual machines as well as the quantity of the operations being handled. Some of the most difficult issues within the field of computing is finding the quickest way among all conceivable permutations.

Materials and information for common processing are made available through cloud computing. The user won't have to rent a server or spend money on air conditioning and heating energy because a host application support provider will be involved. Remote workers may quickly log in and utilise their applications from any location, making it simple for them to connect and move around[9]. This kind of environment offers a configurable option so that data processing may be done without any problems at a specific time. Demand for shared resources within the community of cloud computing develops quickly as user numbers rise. Therefore, one of the main challenges for task scheduling involves load balancing between different services. This need must be handled scientifically in order to satisfy all the requirements established by the numerous protocols used to run these kinds of algorithms.

The method that uses a system running on the cloud allocates workloads and hardware resources is called load balancing. Providing resources to various machines, networks, or servers enables organisations to manage programme or workload needs. This allows for the handling of a variety of requests while compromising the security of the framework as a whole or its requirements. In order to accomplish other load balancing features, such as equitable task distribution across all hosts, facilitating of service effectiveness, enhanced system efficiency, decreased response time, and enhanced resource utilisation, obstructions are frequently avoided by using load balancing[10]. These criteria are among the most typical checklist items that must be upheld in applications of this nature.

Load Balancing Algorithm: Because to the independence of the processing units and also the inter-processor communication cost involved in the gathering of state data, communication delays, redistribution of workload, etc., load balancing on several computers remains a challenge. The best situation for resolving and operating distributed and paralleled programme applications is one that utilises paralleled and distributed computing. A huge process or job is split up and then spread over several hosts enabling parallel computation in these types of applications. Inside a system with several hosts, Livny and Melman have noted that the likelihood of one host sitting idle while another host has numerous jobs stacked up could be extremely significant[11].

In this case, load balancing is most likely to enhance efficiency. These load imbalances imply that either work transfers from the overladen hosts to the lighter laden hosts or load distribution evenly/fairly amongst hosts can enhance efficiency. The aforementioned objective is accomplished with the aid of the load balancing algorithms(s).

The types and quantity of load and work information presumed to be acquired by the decision-making components has a direct bearing onto the load balancing algorithms used. Based on how process is initiated, load balancing algorithms could be divided into 3 groups as follows:

- a) Sender Initiated- In this kind, the sender initializes the load balancing algorithms. This kind of technique involves sending request messages to potential recipients until one is found that can take on the load.
- b) Symmetric- It combines both sender- as well as receiver-initiated behaviors.
- Receiver Initiated- In this kind, the receiver initiates the load balancing process. To discover a sender who can accept the loads, the receiver transmits

request messages throughout this type of descriptive algorithm.

Static Load Balancers:

- 1. Round-Robin load Balancer
- 2. Min-Min
- 3. Max-Min

Dynamic Load Balancers:

- 1. Throttled load balancer
- 2. Honeybee foraging Algorithm
- 3. Biased random sampling
- 4. Round-robin technique
- 5. SJF technique
- 6. Active clustering
- 7. Join-idle queue
- 8. GP algorithm
- Equally Spread Current Execution (ESCE) Algorithm

REVIEW OF SOME STUDIES

The issue of overloading in cloud balancing has been addressed by Chen et al. inside a work titled "A novel load balancing structure and algorithm enabling cloud services." A technique for creating dynamic balance has been suggested[12]. The two methods used to demonstrate the proposed novel technique were looked at, and the technique takes into account simultaneously server processing as well as Computer loading.

The ideal approach for asynchronous iteration load balancing within cloud computing was researched and presented by Coutourier et al. in 2018[13]. The goal of the study was to develop a novel technique termed the best effort to distribute a node's load among all of its laden neighbours while making sure that each node receiving load was distributed equally. A few of the bed testing situations were taken into consideration in the study through using SimGrid simulator, and numerous QoS parameters were assessed to show the usefulness of the suggested approach. The performance and security of the data centre may decrease if the workload is not distributed evenly. Such centres employ the technique of virtualization, which enables the live transfer of virtual machines. Additionally, a distribution method that balances the existing load amongst various sources based on the performance of the systems or hosts within the data centre is described in this study. This system has been assessed using real-time transmission, simulation, and redeployment of virtual computers according to productivity.

By using meta-heuristic techniques, Mousavian Qalashqaei with Shiri achieved 20% load balancing using virtual computers[14]. To link a collection of queries to the system's accessible resources in accordance with the requirements of cloud-based computing systems, a novel approach is suggested in this study. They have used tabu search algorithms and the evolutionary algorithms mutation approach in this approach.

Burden balancing was used by Barani et al to lessen the load on virtual machines[15]. By simulating the algorithm and evaluating it to a variety of different load balancing algorithms, researchers determined that it provides a proper reaction time and computation time in comparison to earlier systems. The algorithm is based upon the processing capacity and task workload of virtual machines using cloud computing. The latency between the start and finish of a system's series of tasks or activities is known as the scheduling time. The service's usefulness must be evaluated at this point, which is crucial. It is preferable to lower this standard.

Two strategies for enhancing load balancing with task scheduling were given by Chanaghlou and Dolati in their 2016 work, "Having provided a Hybrid Objectives Scheduling with Load Balancing within Cloud Computing[16]." The authors came to the conclusion that such load-balancing objectives can be met while minimising overall runtime by using the balance method known as Hypertext Markup Language (HMTL). Additionally, it employs the practice of limiting task migrations. By calculating its current work load, the scheduling algorithms known as LDTS (Linear Decision Trees) also distributes new operations to system process nodes. The LDTS method has enhanced load distribution, according to studies. The HMTL algorithm has indeed enhanced variables such work load distribution, full load balancing, momentary load balancing, and final runtime.

In 2017, Derakhshanian et al. looked at load balancing inside a cloud computing environment while using an adaptive evolutionary algorithm and accounting for task dependency[17]. The objective of this research was to develop a system for network workload balancing which takes into account interactions among jobs, minimising both the overall completion time as well as the idle timing of the machines. The outcomes of experiment demonstrated that the localisation of connections would significantly impact the reduction of the overall completion time.

Utilizing the Ant Colony Algorithms, Mishkar et al. improved job scheduling as well as load balancing inside the cloud

environment in 2018[18]. This study's objective was to explore load balancing upon machines in addition to task scheduling. This was accomplished using scheduling only with ant colony optimizations technique, which effectively handles numerous dynamic situations. The research objectives, the scheduling issue, and the related duties were discussed in this study. Descriptions of real - time tasks as well as the cloud environment also were suggested. Every one of algorithm's phases were then carried out, and load balancing were completed.

A hybrid task scheduling approach utilising evolutionary algorithms and fuzzy theory was proposed by Javanmardi et al. Without taking into account the job response time or resource capabilities, the suggested method distributes jobs to available resources[19]. The approach was built on genetic algorithms. Fuzzy theory was used to modify this and cut down on the number of population-creation iterations. The fitness values of each of chromosomes for said 2 kinds of chromosomes with varied Qos metrics were then determined. The usual genetic algorithm was updated using the new methodology, which reduced the execution costing to 45% and the overall execution duration to 50%. This hybrid technique beat previous algorithms when it was put into use on CloudSim, according to the results.

Inside a cloud computing framework, Sarita et al suggested a successful load balancing method. In this paper, a prioritised virtual machine (VM) configuration and load balancing with round-robin scheduling are proposed[20]. The virtual machine is set up by evaluating key characteristics including bandwidth, RAM, as well as MIPS (million instructions per second). The updated algorithm uses round robin's load balancing to guarantee that tasks just aren't scheduled depending on virtual machine prioritization, preventing overload or extended periods of inactivity on any virtual machine. The practice of placing an excessive workload on any single virtual machine is eliminated by this new technique, which also improves resource consumption. The algorithms were simulated using CloudSim like a simulation software. The experiment's findings made it abundantly evident that the suggested job scheduling method takes lesser time than traditional Round Robin Algorithm.

A scheduling algorithm using the multi-objective genetic algorithms (MOGA) was presented by Liu, Luo, Zhang, Zhang, as well as Li[21]. This method integrates greedy initialization with random initialization. Power utilises with service provider revenues are used in genetic algorithms to determine fitness. Fractal is used to choose and store the finest fitness. The two tactics of elitism and swarming can then be used to carry out the selecting process. Following the

selection procedure, crossover operations can be carried out by two people. Their position is changed via mutation to create new beings. By using this MO-GA, energy usage is reduced and service providers' profits are increased.

A new job scheduling algorithms in a grid setting, RASA (Resources Aware Scheduling Algorithm), was proposed by Parsa and Entezari-Maleki[22]. The algorithm was created by a thorough examination of the Min-min as well as Max-min job scheduling algorithms. RASA utilises the benefits of both algorithms while addressing their drawbacks. Using this approach, the scheduler initially distributes resources to jobs in accordance with the number of readily available resources. After adding the resource accessible, the Min-Min algorithms approach is selected. The resource is distributed round by round, first to lesser jobs in the following round, then to larger tasks, and vice versa. According to the simulation outcomes, this algorithm did better than the two other traditional algorithms when compared to them.

For load balancing within cloud computing, Rajput et at. presented a genetically based enhanced min-min job scheduling algorithm[23]. For order to reduce the make-span and maximise resource consumption, this work suggested a new load balanced min-max algorithm. Initially, the virtual machine's response time for the operation is determined. The task's minimal or max time onto the virtual machine is called as a result. Then, for improved execution, a genetically based technique was used on tasks with millions of orders and virtual machines with millions of commands per second. A recombination, mutation, and fitness value of GA were used in this case.

A power-efficient virtual machine provisioning technique for cloud systems was put out by Lin, Liu, and Wu[24]. Those authors suggested creating a hybrid algorithm by combining dynamic Round-Robin with First Fitting. The amount of arriving virtual machines is supposed to be a time function and the conditional probability (for example, a normal distribution) is implemented. The input rate of the virtual machines was employed by the hybrid algorithms to schedule virtual machines. The First-Fit was utilized by the Hybrid technique throughout rush hours to fully utilise the computational capability of physical computers.

A load balancing technique combining fuzzy logic plus the Round Robin algorithm was first introduced by Sethi et al. in the technique is dependent on a variety of factors, including processor speed, allocated load in virtual machines, and others[25]. The algorithm keeps track of every virtual machine's detail and the number of requests that are currently being assigned to it. When there are multiple virtual

ISSN: 2321-8169 Volume: 11 Issue: 11

Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 November 2023

machines, the choice will be made utilizing fuzzy logic depending on processor speed as well as virtual machines load whenever a new request comes in. The load balancer finds for the lower loaded virtual machine and allocates it. The outcomes demonstrated that it outperforms than the Round-Robin algorithms.

A new genetic algorithm-based technique was proposed by Hu et al. in order to improve task scheduling[26]. The algorithm made use of both previous data with current system state. The collection of physical servers and the group of virtual machines are mapped to one another. By calculating the system's influence before the necessary virtual machine resources are deployed, it selects the least-effective solution. To figure out the ideal scheduling option based on demographics, they employed an equation.

PROPOSED ALGORRITHM

As a result, the following is how our proposed algorithms differ from other algorithms: Include within the cloudlet job's, queue job and the anticipated finish time of every resource (virtual machine). Depending on this variable, the algorithm chooses the VM for operation allocation that has the shortest projected completion time as well as the least usage rate. The Figure 1 portrays the Architectural Flowchart of Proposed Algorithm.

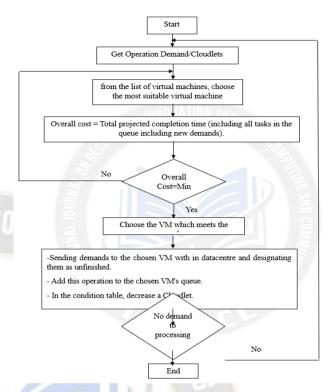
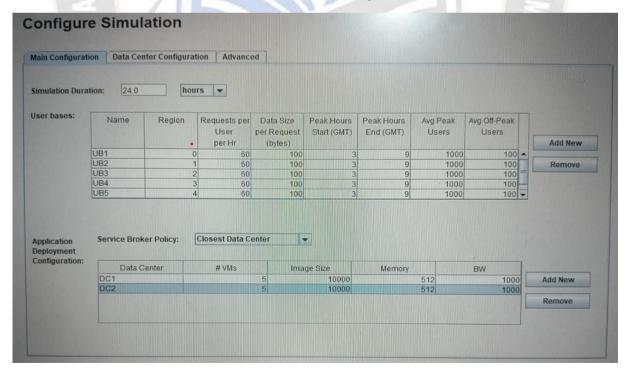


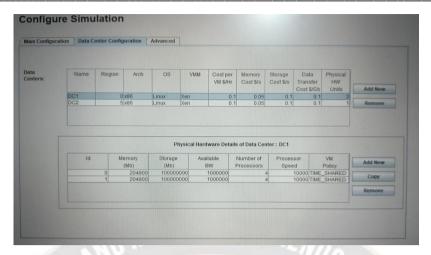
Fig.1: The Figure Demonstrates the Architectural Flowchart of Proposed Algorithm

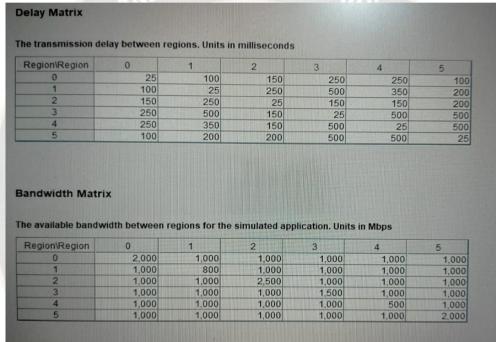
The images below show the various settings for the configuration simulation-

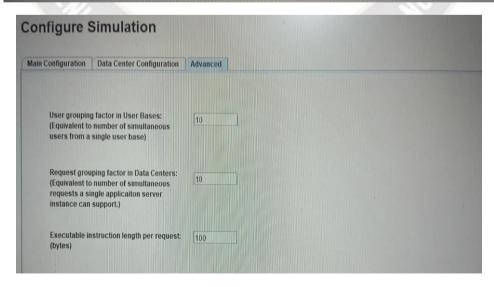


ISSN: 2321-8169 Volume: 11 Issue: 11

Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 November 2023







RESULTS

The below table 1,2 and graphs 1, 2 portray the comparison of response time and data processing times between the proposed hybrid algorithm and traditional algorithm.

Table.1: The Below Table Demonstrates Comparison of Response Time of All the Algorithm

Algorithm	Response Time	
	Average Response Time (ms)	Maximum Response Time (ms)
Max-Min	160.77	1281.62
Round Robin	140.62	940.88
Min-Min	93.46	832.48
Proposed Algorithm	37.65	763.91

Graph.1: The Below Graph the Response Time of all the Algorithms

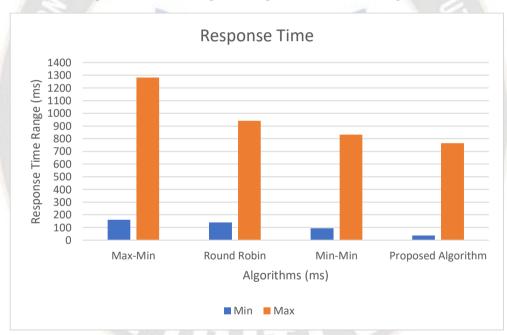
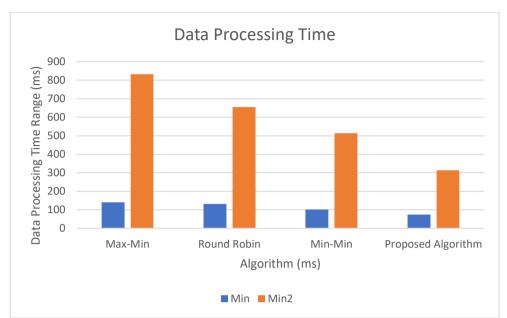


Table.2: The Below Table Demonstrates Comparison of Data Processing Time of All the Algorithm

Algorithm	Data Processing Time	
	Average Response Time (ms)	Maximum Response Time (ms)
Max-Min	140.56	832.67
Round Robin	132.23	654.54
Min-Min	101.49	513.41
Proposed Algorithm	73.66	313.58



Graph.2: The Below Graph the Data Processing Time of all the Algorithms

CONCLUSION

A broad variety of services are offered to end customers quickly and on request in the heterogeneous network of cloud computing. Among the key concerns with cloud computing involves load balancing. Within a cloud computing context, the load balancing scheduling techniques used today are not very effective. For effective resource utilisation in cloud computing, load balancing is indeed a key responsibility. Increased client satisfaction, optimal resource usage, and improved cloud system's performance are the major objectives of load balancing, which also aims to cut down on energy use and carbon emissions. With each IT engineer today, the two biggest obstacles to creating products which can improve cloud operational efficiency are money and time.

The suggested hybrid algorithm portrays the response time (min) 37.65 and (max) 763.91 and the data processing time (min) 73.66 and (max) 313.58, which portrays that the proposed algorithm shows excellent results with compared to all the other algorithm portrays in the simulation configuration results.

To further enhance the efficiency of cloud computing, academics can carry out additional study regarding how to create algorithms which are more dynamic, smart, and capable of resolving fault tolerance concerns. Future articles will discuss algorithms that draw inspiration from nature and are sophisticated, like machine learning applications or clustering strategies. By assessing more cloud computing techniques, the research can also be broadened.

REFERENCES

- [1] M. O. Ahmad and R. Z. Khan, "Load balancing tools and techniques in cloud computing: A systematic review," in Advances in Intelligent Systems and Computing, 2018. doi: 10.1007/978-981-10-3773-3 18.
- [2] P. Pradhan, P. K. Behera, and B. N. B. Ray, "Modified Round Robin Algorithm for Resource Allocation in Cloud Computing," in Procedia Computer Science, 2016. doi: 10.1016/j.procs.2016.05.278.
- [3] S. K. Mishra, B. Sahoo, and P. P. Parida, "Load balancing in cloud computing: A big picture," Journal of King Saud University Computer and Information Sciences. 2020. doi: 10.1016/j.jksuci.2018.01.003.
- [4] K. V. Reddy et al., "Research Issues in Cloud Computing," Online, 2011, doi: 10.1109/cise.2010.5677076.
- [5] NIST, "NIST Cloud Computing Reference Architecture: Recommendations of NIST," Natl. Inst. Stand. Technol., 2011.
- [6] N. Jain and S. Choudhary, "Overview of virtualization in cloud computing," in 2016 Symposium on Colossal Data Analysis and

- Networking, CDAN 2016, 2016. doi: 10.1109/CDAN.2016.7570950.
- [7] A. P. M and M. T. Sathiyabama, "Virtualization in Cloud Computing," Int. J. Trend Sci. Res. Dev., 2018, doi: 10.31142/ijtsrd18665.
- [8] S. Afzal and G. Kavitha, "Optimization of Task Migration Cost in Infrastructure Cloud Computing using IMDLB Algorithm," in 2018 International Conference on Circuits and Systems in Digital Enterprise Technology, ICCSDET 2018, 2018. doi: 10.1109/ICCSDET.2018.8821193.
- [9] B. Piprani, D. Sheppard, and A. Barbir, "Comparative analysis of SOA and cloud computing architectures using fact based modeling," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2013. doi: 10.1007/978-3-642-41033-8 66.
- [10] K. Ramana and M. Ponnavaikko, "A Multi-Class Load Balancing Algorithm (MCLB) for heterogeneous web cluster," Stud. Informatics Control, 2018, doi: 10.24846/v27i4y201808.
- [11] M. Livny and M. Melman, "Load balancing in homogeneous broadcast distributed systems," ACM SIGMETRICS Perform. Eval. Rev., 1982, doi: 10.1145/1010631.801689.
- [12] S. L. Chen, Y. Y. Chen, and S. H. Kuo, "CLB: A novel load balancing architecture and algorithm for cloud services," Comput. Electr. Eng., 2017, doi: 10.1016/j.compeleceng.2016.01.029.
- [13] R. Couturier, A. Giersch, and M. Hakem, "Best effort strategy and virtual load for asynchronous iterative load balancing," J. Comput. Sci., 2018, doi: 10.1016/j.jocs.2018.04.002.
- [14] A. Mousavian Qalashqaei, N; Shiri, "Load balancing on virtual machines in cloud computing by combining meta-heuristic methods," Natl. Conf. Comput. Eng. to Inf. Technol. Manag., 2013, [Online]. Available: https://bjopm.org.br/bjopm/article/view/864/890
- [15] F. Barani, S.; Sattari, V.; Barani, "Load Balancing Based on Processing Power and Task Load of Virtual Machines in Cloud Computing," 7th Int. Conf. Inf. Technol. Knowl., 2015, [Online]. Available: https://bjopm.org.br/bjopm/article/view/864/890

- [16] A. Chanaghlou, A.; Dolati, "Providing a method for hybrid multi-objective scheduling and load balancing in cloud computing," Kerman Grad. Univ. Adv. Technol., 2016, [Online]. Available: https://bjopm.org.br/bjopm/article/view/864/890
- [17] S. J. H. A. Derakhshanian, Y.; Mirabedini, "Load balancing in cloud computing environment, considering dependence between tasks and using adaptive genetic algorithm," Telecommun. Eng. J., 2017, [Online]. Available: https://bjopm.org.br/bjopm/article/view/864/890
- [18] A. Mishkar, Y.; Seyed Bagheri, S. H.; Baghbani, "Optimizing scheduling tasks and load balancing in the cloud environment using ant colony algorithm," First Natl. Conf. Res. Comput. Eng., 2018, [Online]. Available: https://bjopm.org.br/bjopm/article/view/864/890
- [19] S. Javanmardi, M. Shojafar, D. Amendola, N. Cordeschi, H. Liu, and A. Abraham, "Hybrid Job Scheduling Algorithm for Cloud Computing Environment," in Advances in Intelligent Systems and Computing, 2014. doi: 10.1007/978-3-319-08156-4_5.
- [20] S. Negi, M. M. S. Rauthan, K. S. Vaisla, and N. Panwar, "CMODLB: an efficient load balancing approach in cloud computing environment," J. Supercomput., 2021, doi: 10.1007/s11227-020-03601-7.
- [21] Z. Chenhong, Z. Shanshan, L. Qingfeng, X. Jian, and H. Jicheng, "Independent tasks scheduling based on genetic algorithm in cloud computing," in Proceedings - 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2009, 2009. doi: 10.1109/WICOM.2009.5301850.
- [22] S. Parsa and R. Entezari-Maleki, "RASA: A New Task Scheduling Algorithm in Grid Environment," World Appl. Sci. J., 2009.
- [23] S. S. Rajput and V. S. Kushwah, "A Genetic Based Improved Load Balanced Min-Min Task Scheduling Algorithm for Load Balancing in Cloud Computing," in Proceedings - 2016 8th International Conference on Computational Intelligence and Communication Networks, CICN 2016, 2017. doi: 10.1109/CICN.2016.139.

- [24] C. C. Lin, P. Liu, and J. J. Wu, "Energy-efficient virtual machine provision algorithms for cloud systems," in Proceedings 2011 4th IEEE International Conference on Utility and Cloud Computing, UCC 2011, 2011. doi: 10.1109/UCC.2011.21.
- [25] S. Sethi, "Efficient load Balancing in Cloud Computing using Fuzzy Logic," IOSR J. Eng., 2012, doi: 10.9790/3021-02716571.
- [26] J. Hu, J. Gu, G. Sun, and T. Zhao, "A scheduling strategy on load balancing of virtual machine resources in cloud computing environment," in Proceedings 3rd International Symposium on Parallel Architectures, Algorithms and Programming, PAAP 2010, 2010. doi: 10.1109/PAAP.2010.65.

