International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 08 | August - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Cloud-Based Dynamic Task Scheduling and Load Balancing Using a Variety of Intelligent Schemes

Pavan Kumar K
Department of MCA
Dayanand Sagar College of Engineering
Bengaluru

Dr.Samitha Khaiyum
Head of Department
Department of MCA
Dayanand Sagar College of Engineering
Bengaluru

ABSTRACT

One of the main problems with cloud computing is load balancing, which divides tasks into virtual machines (VMs) with varying lengths and completion times. As a result, keeping VMs loaded up during job scheduling and allocation is difficult. In this research, we suggested secure load balancing in a cloud context and resource-aware dynamic job scheduling utilizing intelligent algorithms. The process for the planned work is as follows:

(1). The Human Mental Search (HMS) method is used to cluster Real-Time and Non-Real Time jobs into four categories: Real-Time and Short, Real-Time and Long, Non-Real Time and Short, and Non-Real Time and Long.

(2). Knowledge of Energy and Delay Combined Tasks Tasks are planned based on task clustering, and the fuzzy VIKOR method is used to schedule them in the order of priority.

(3). Security Risks aware VM Clustering, which employs the Hierarchical Agglomerative Clustering algorithm, consists of three distinct entities: a VM's information collector, which compiles data about VMs; a load monitor, which oversees the entire datacenter; and a decision maker, which assigns tasks by choosing the best VM.

(4). Co-Resident assaults were stopped by the continuous-Actor-Critical-Algorithm (C-ACAL), which is used to pick the best VM. Using CloudSim, which measures performance in terms of response time, latency, resource utilization, throughput, energy consumption, and task execution success rate, the simulation is carried out.

INTRODUCTION

The most fundamental issue in cloud computing is task scheduling, which calls for a perfect match between task needs and available resources. Task scheduling is carried out with the assistance of specialized and dependable cloud resources, as the dynamic scheduling procedure for online services is now required [1-3].

Numerous requests arrive in the cloud every second, and it is the role of the cloud computing system to handle all of the demands [4]. Because the quality of the service is what determines whether a user is satisfied, the cloud should prioritize task scheduling and load balancing. However, the cloud's resources are limited by the CPU, memory, and bandwidth that are required on a sporadic basis, and they must be acquired in order to complete activities [5-8].

VM migration is one of the new problems that needs to be addressed in load balancing. The most effective VM choice is one of the simple fixes for this problem [9]. This research has focused on the VM selection as a result of this. In the VM selection literature that is currently available, only a few criteria—such as CPU, RAM, task processing cost, and so on—were prioritized. The ideal set of requirements must be made here in order to execute real-time and online apps of today. For example, the kind of VM, the load on the VM at the moment (the number of tasks running and in the waiting state), and the history of the VM (the number of tasks completed, executed successfully, and unexecuted tasks count). On the basis of that, the VM must be matched with the arriving task.

LITERATURE SURVEY

In paper [11], a fog-enabled cloud architecture that uses a modified firework algorithm to analyze task characteristics and select resources was proposed. This technique employs a method for detecting fireworks from a distance. The distance between two jobs was used to cluster work in the fog layer, and tasks that came were scheduled. Here, the place of task execution was predicted.

A Task Distribution Algorithm (TDA) is suggested in paper [10] to distribute workloads among many servers in order to reduce datacenter power consumption. For VM to PM mapping (VM placement) and resource deallocation in TDA, two heuristic techniques, such as the Power Award (PA) Algorithm and Adaptive Harmony Search Algorithm (AHSA), were presented.

PROBLEM STATEMENT

A major issue in the world of cloud computing is load balancing. Due to the widespread use of smart Internet of Things (IoT), there are currently a lot of tasks arriving as requests. IoT device queries should run concurrently without experiencing any execution failures even though they are delay sensitive. Therefore, one of the fundamental tasks in a cloud datacenter is efficient resource allocation with dynamic task scheduling.

Cloud computing provides a flexible and scalable infrastructure for deploying and managing applications. However, effective utilization of cloud resources heavily relies on dynamic task scheduling and load balancing techniques. The challenge lies in efficiently allocating tasks to available resources, minimizing response time, and achieving load balancing, considering varying workloads, resource capacities, and changing environmental conditions.

Although several intelligent schemes have been proposed for dynamic task scheduling and load balancing in cloud environments, there is still room for improvement.

Therefore, the problem addressed in this study is to develop a comprehensive framework for cloud-based dynamic task scheduling and load balancing that utilizes a variety of intelligent schemes.

This research aims to address the following key challenges:

1.Intelligent Scheme Integration: Developing a framework that effectively combines multiple intelligent schemes to leverage their individual strengths and provide a robust solution for dynamic task scheduling and load balancing in cloud environments.

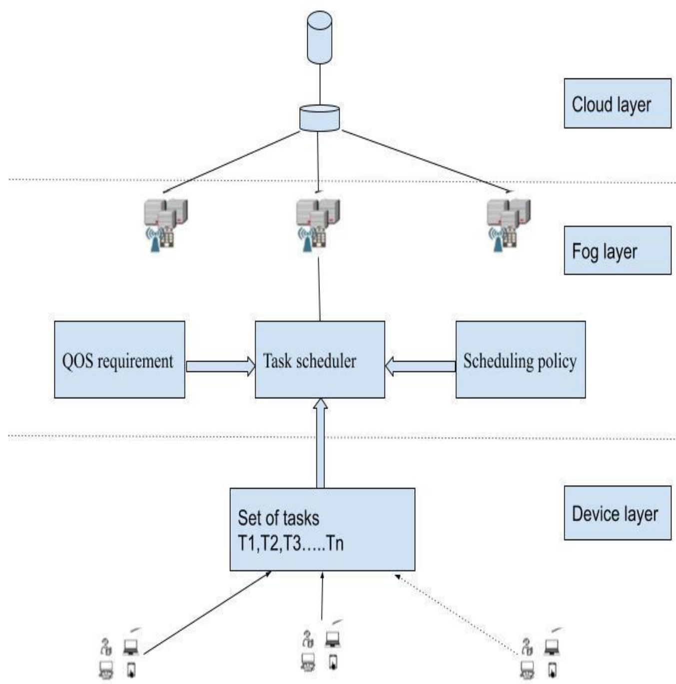
2.Scalability and Adaptability: Designing the framework to be scalable and adaptive, capable of handling varying workloads, resource capacities, and environmental conditions without sacrificing performance or efficiency.

3.Resource Utilization Optimization: Maximizing resource utilization by intelligently allocating tasks to available resources, considering factors such as task dependencies, workload patterns, and resource capabilities.

4.Load Balancing: Achieving load balancing by evenly distributing the workload among available resources to avoid bottlenecks, prevent resource overloading, and ensure efficient utilization of the cloud infrastructure.

PROPOSED METHODOLOGY

Multiple intellectual techniques are employed in this study to focus on secure load balancing and resource-aware dynamic job scheduling. The proposed work intends to schedule the incoming tasks and balance the load on each particular VM and any other VMs running on the same physical machine. In our system, there are three separate tiers. They are the cloud layer, fog tier, and tier for IoT devices. In the process of reaching the goal of the suggested system, each of these layers will be crucial. IoT Devices / Users, Edge Gateway, Fog Node, Secure Cloud Broker, and Datacenter are some of the entities that make up the proposed three-tier architecture for dynamic job scheduling and load balancing.The proposed architecture is shown in the below figure.



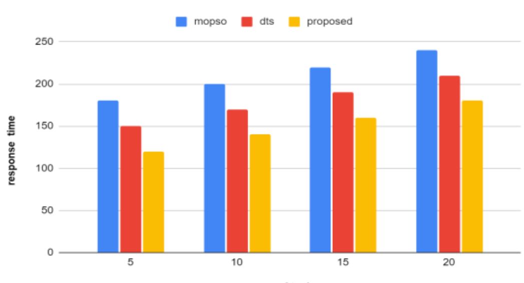
COMPARISON STUDY

In terms of Response time vs. Number of Tasks, Latency vs. Number of Tasks and VMs, Resource utilization vs. Number of Tasks and VMs, Throughput vs. Number of Tasks and VMs, Energy Consumption vs. Number of Tasks and VMs, and Task Execution Success Rate vs. Number of Tasks and VM, this section compares the proposed and existing systems.

a.impact of response time

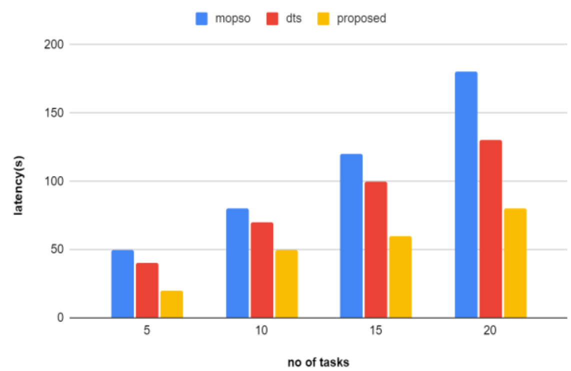
Response time is the entire amount of time that the virtual machine (VM) spends responding to a particular request, and it is measured in relation to the number of tasks.

The below graph represents the comparison of existing and proposed system response time with respect to number of tasks. From the figure the response time increases exponentially with increase in number of tasks, but this proposed system the tasks are clustered and scheduled based on priority which reduces response time and increases the efficiency of this system. The comparison result shows that the proposed system achieves low response time compared to existing system such as MOPSO and DTS.



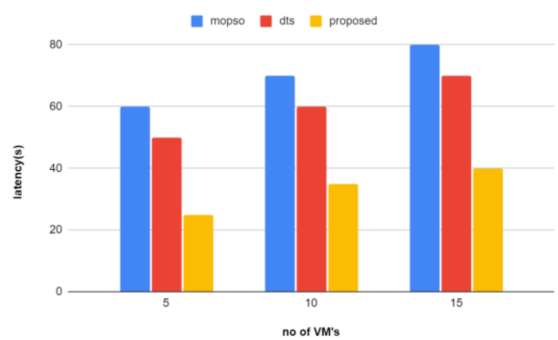
b. Impact of latency

Latency, which is the amount of time it takes for a job to travel from its source to its destination, lowers the proposed system's accuracy and performance. The quality of service will suffer if scheduling or allocation are delayed. Latency vs no.of tasks



Latency vs no. of VM's

This graph indicates the comparison based on latency and number of virtual machine's.

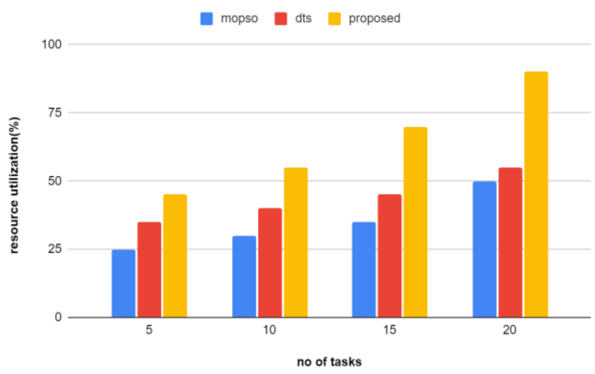


c. Impact of Resource Utilization

Resource utilization refers to how a task makes use of a resource to execute a work, which enhances the performance of the suggested system while also conserving energy used by the resources in the data centre.

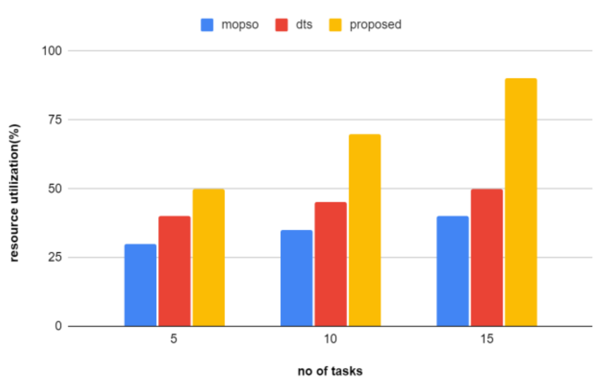
International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 08 | August - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Resource Utilization



Resource Utilization vs no. of VM's

This graph indicates the comparison based on resource utilization and number of virtual machine's.

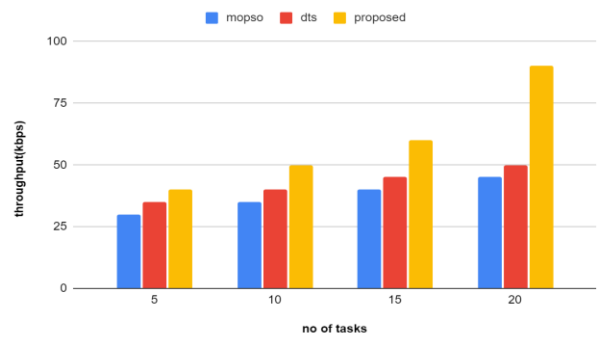


d. Impact of Throughput

The rate at which tasks are successfully received by the destination node is known as throughput, and it is one of the key parameters that determines how accurate the process is.

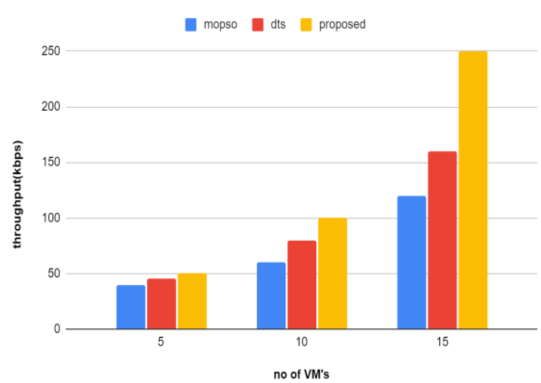
Throughput vs no. of Tasks

This graph indicates the comparison based on throughput and number of task's.



Throughput vs no. of VM's

This graph indicates the comparison based on throughput and number of virtual machine's.



Comparison of Proposed and Existing Approaches

Performance Metrics	Proposed vs Existing Approaches			
		Mopso	Dts	proposed
Response time(ms)	of task	210	130	150
Latency(s)	of task	107.5	85	50
	of VM	70	60	35
Throughput(kbps)	of task	37.5	42.5	60
	of VM	70	90	133.3
Resource Utilization(%)	of task	32.5	42.5	65
	of VM	35	45	70

CONCLUSION

In conclusion, cloud-based dynamic task scheduling and load balancing using a variety of intelligent schemes is a highly effective approach for optimizing resource allocation and improving overall system performance in cloud computing environments. By leveraging intelligent algorithms and techniques, such as machine learning, genetic algorithms, and swarm intelligence, this approach enables efficient task scheduling and load balancing, leading to enhanced resource utilization, reduced response time, improved scalability, and increased reliability.

Overall, cloud-based dynamic task scheduling and load balancing using a variety of intelligent schemes offer numerous benefits in terms of resource optimization, system performance, scalability, and reliability. As cloud computing continues to evolve and become increasingly prevalent, this approach will play a critical role in maximizing the efficiency and effectiveness of cloud-based systems and services.

REFERENCES

[1] Hudaib, A., Shukri, S.E., Al-Sayyed, R.M., and Mirjalili (2020). Improved multi-verse optimizer for cloud computing work scheduling. 114230, Expert Systems With Applications.
[2] I.M. Ali, K.M. Sallam, N. Moustafa, R. Chakraborty, M. Ryan, & K.R. Choo (2020). An automated task scheduling model for fog-cloud systems employing the Non-Dominated Sorting Genetic Algorithm II. Cloud Computing, IEEE Transactions, 1-1.
[3] In 2019, Pang, S., Li, W., He, H., Shan, Z., and Wang, X. A hybrid EDA-GA algorithm for scheduling multi-objective tasks in cloud computing. IEEE Access 7, pp. 146379–146389.
[4] In 2020, Geng, S., Wu, D., Wang, P., and Cai, X. Scheduling Cloud Tasks With Many Objectives. 8 IEEE Access, 79079–79088.
[5] Zhang, W., Elgendy, I.A., Hammad, M., Iliyasu, A.M., Du, X., Guizani, M., and El-Latif, A.A. published it in 2020. Secure and efficient load balancing is needed for systems that use many tiers of IoT and edge cloud computing. IEEE, 1-1 Journal of the Internet of Things.
[6] Alharbi, A.G., Maswood, M.M., Rahman, M.M., and Medhi (2020). A Novel Approach for Load Balancing and Bandwidth Cost

Reduction in a Three-Layer Fog Computing Environment. IEEE Access, 8(11), 113737–113750.

[7] L. Chen, H. Shen, & K. Sapra (2014). Cloud load balancing using Resource Intensity Aware Loading, or RIAL. IEEE Conference on Computer Communications (IEEE INFOCOM 2014), 1294–1302.

[8] A. Pradhan, S.K. Bisoy, and others (2020). a cutting-edge PSO-based load balancing method for cloud computing platforms. Journal of Computer and Information Sciences at King Saud University.

[9] It was published in 2020 by Babou, C., Fall, D., Kashihara, S., Taenaka, Y., Bhuyan, M.H., Niang, I., and Kadobayashi, Y.

[10] Geetha, Kandasamy, Muthukumar, Zong Woo, and Renugadevi, T. (2020). Adaptive Harmony Search Algorithm for Energy-Efficient Resource Provisioning for Compute-Intensive Workloads with Load Balancing in Datacenters. Sciences applied. 10. 2323.

[11] S. Wang, T. Zhao, & S. Pang (2020). Improved Firework Algorithm-Based Task Scheduling Method in Fog Computing. 8th IEEE Access, 32385-32394. doi:10.1109/access.2020.2973758