A COMPREHENSIVE STUDY OF THE MOST EFFICIENT SCHEDULING TECHNIQUES FOR VIRTUAL MACHINES IN CLOUD ENVIRONMENTS

Altaf Raja
Research Scholar
Department of Computer Science & Engineering
Sunrise University, Alwar.
Rajasthan
altaf.soft@gmail.com

Dr. Sumit Bhattacharjee
Research Guide
Department of Computer Science & Engineering
Sunrise University, Alwar.
Rajasthan.

Abstract: *Cloud computing is a Pay-Per-Use strategy that efficiently delivers services and resources to consumers through the internet without any investment. CC has various uses due to its inexpensive cost and easy implementation. Its heterogeneous, flexible, distributed, location-independent, on-demand self-service, and universal network access make it unique. Popular because it can leverage cloud services and resources on any physical infrastructure.*

The service provider's physical infrastructure, including location, platform, and function, is unknown to the user. Hence, such systems need correct operating setup for load balancing. Scheduling was created for this. 1. Host Level (VM Scheduling- Allocation of PEs to the Hosts) User-level (Cloudlet Scheduling-Allocation of cloudlets to VM for execution). This study proposes a novel PE-VM allocation technique. Ant Colony Optimization improves algorithm output.

Keywords *- Cloud Computing (CC), Virtual Machine (VM), Processing Element (PE), Virtualization.*

INTRODUCTION

Cloud computing is a parallel and distributed system of software, virtualized computers, storage, and other services that users may directly access and pay for just what they need. This allows users to access unique databases at little cost, making the CC model popular in IT. Hosting relevant tasks is scheduling. This work is crucial to cloud resource usage. Virtual machine software installs and runs programs. Virtual machine resource allocation and application migration are crucial. Virtual machines are planned for geographically spread data centers in the cloud architecture. Scheduling optimization defines cloud computing resource and infrastructure usage.

Cloud scheduling optimizes virtual machines. Scheduling improves service and maximizes cloud service provider value. Energy efficiency and cloud computing cut costs for businesses. Optimization saves energy, speeds task execution, and optimizes resource allocation. Virtualization helps manage dynamic resources in the cloud computing IaaS paradigm. Scheduling techniques link VMs to physical servers. Heterogeneity is addressed to dynamically balance load. Allocating virtual machines to user demands in virtual layers optimizes resource efficiency. Virtual machines allow users to share a real system while separating hardware resources. Virtual machines have varied CPU speeds, memory sizes, and physical resources.

Virtual machines are excellent for dynamically balancing system load, however resource underutilization may occur. Scheduling algorithms misallocate resources, causing underutilization. Underutilization causes servers to overheat, which raises cooling costs. Proper virtual machine scheduling solves two issues at once: mapping VMs to real machines and selecting virtual machines for load balancing. In cloud computing literature, these phrases are usually used together. Dynamic scheduling and load balancing are proposed for energy-efficient cloud systems. Sections A and B explain cloud computing kinds and task scheduling.

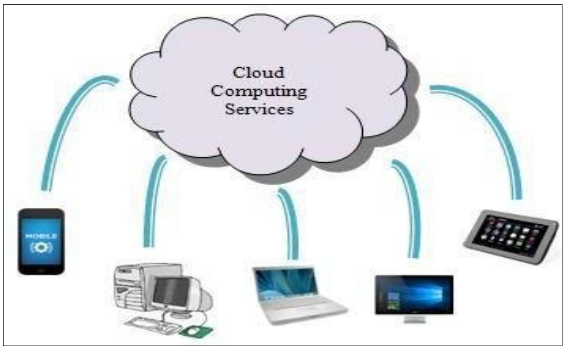

Figure-1 Cloud Computing Demonstration

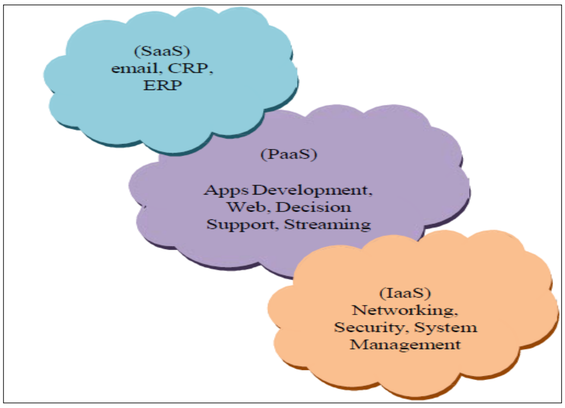

Figure- 2 Cloud Computing Services

A. Cloud Computing Types

- Cloud computing may be broken down into the following categories based on location: Public Cloud: In this case, the vendor provides the computing infrastructure, and the client has no access to it. Yet, the assets could be open to the general public.
- Infrastructure is created for a private enterprise on a private cloud. The only private organization that can use the services is that one. A private cloud is safer than a public one.
- Hybrid Cloud: A hybrid cloud is created by combining a private cloud with a public cloud. A private cloud may be used to install critical applications, while a public cloud may be used to connect less secure apps.
- Community Cloud: In this case, companies within the same community share infrastructure.

B. Task Scheduling

Cloud computing services employ CPUs, memory, storage, and numerous applications. Needs determine resources. Cloud computing prototypes maximized services and lowered hardware and software costs. Virtualization lets users share cloud resources. Virtualization optimizes energy and resource use in distant situations. Cloud datacenter software stacks need VMs. Task scheduling is crucial in cloud computing since companies are moving to the cloud and users are growing.

Computer science milestones include task scheduling. Scheduling resources among user-sent tasks at a given time improves service quality. Task scheduling determines which resource and when to perform a task. Processed scheduling—thread handling in an operating system and energy management using task schedulers in cloud computing—are current research areas. Cloud computing has become popular due to its dependability, scalability, cost reduction, and anytime, anywhere information sharing. Cloud computing is the most sought-after technology for research and practical use.

A process scheduler in cloud computing allocates resources to user tasks. The cloud system must distribute resources to many jobs since many users seek them. QoS must not be compromised. Efficient task scheduling is crucial for user satisfaction. Users will hesitate to pay if performance is poor, as usual. Cloud computing culture revolves on scheduling.

Cloud computing optimizes resource use. Scheduling algorithms are essential. Task scheduling algorithms must schedule user-requested tasks. Scheduling algorithms optimize resource use, execution time, and load balancing. Task scheduling in any computer system is to order tasks according to issue constraints. Cloud computing performance is enhanced through resource scheduling. Most scheduling algorithms optimize cost, make span, scheduling pace, resource use, and more. We will show and extensively discuss several task scheduling algorithm research articles in the following part.

a) Fundamental Scheduling Algorithms

The following are the primary job scheduling algorithms:

FCFS (First Come First Serve) algorithm – This method is often thought of for parallel processing. It uses the first in, first out (FIFO) principle and sends tasks to the resource with the shortest waiting line for incoming jobs. Its disadvantage is that the final task must wait a very long period.

Round Robin Scheduling Algorithm – Also, it manages tasks according to the first in, first out principle. Time is divided into several interval slots in the RR algorithm. When a task is taking longer than expected to finish, the CPU is transferred to the next job in the queue.

Min-Min Algorithm – This method first determines the minimal execution times for all tasks before allocating the shortest tasks to the resources with the quickest response times. It is a more stable algorithm than FCFS.

Max-Min Algorithm – Similar to the Min-Min method, this one sends the biggest workload to the quickest resource. In comparison to FCFS and Min-Min algorithms, it performs better.

Priority Based Algorithm – The priority notion serves as the foundation for this algorithm. Thus, a task requiring a lot of computational power is assigned the highest priority, followed by one requiring little processing power, a task with a low priority, and one with a medium priority. The free resource with the highest power is then allocated to the task with the highest priority. This method outperforms FCFS, Max-min, and Min-Min in terms of performance.

Most Fit Task Algorithm – The MFT algorithm prioritizes tasks based on how well they fit into a queue, although this method has a greater failure rate.

SOME LITERATURE REVIEW ON EXISTING ALGORITHMS

This paper's major goal is to shed light on several algorithmic approaches to job scheduling in a cloud computing context. The following are the techniques:

1) Work Scheduling Based on Symbiotic Organism Search Optimization in a Cloud Computing Environment.

A discrete version of the Symbiotic Organism Search meta-heuristic algorithm was created by M. Abdullahi et al. [1].

This technique was used in the CloudSim application to schedule separate activities. Response time and degree of imbalance were tested over a range of virtual machines' build spans. Performance results showed that DSOS outperformed SAPSO. In comparison to SAPSO, DSOS reduced the average makespan from 3.8% to 25.5% for jobs with 300 to 1000 instances, respectively. DSOS outperformed SAPSO in a bigger search area.

2) A Fresh Scheduling Method for Cloud Computing

To cut down on waiting times and line length, S. Sagnika et al. [2] devised a hybrid work scheduling method based on genetic algorithms and employing a queuing model. In order to compare FCFS with GA, the simulation was run. According to the trial, GA produced outcomes that were 20% better than FCFS. A typical client count and waiting time were the simulation's inputs.

3) Cloud Computing Throughput Optimization Using Multi-Objective Work Scheduling Algorithm.

In order to improve the throughput, L. V. Atul et al. [3] developed a multi-objective job scheduling technique without affecting SaaS cloud SLA. Optimal scheduling is suggested. Authors noted maximum work scheduling techniques are based on execution time, however cloud computing requires execution time, cost, and user bandwidth. CloudSim showed that the suggested technique beat FCFS and priority scheduling in throughput.

1) Cloud Multi-Objective Task Scheduling with Nested PSO Framework

Jena [4] suggested multi-objective task scheduling layered Particle Swarm Optimization to save time and energy. CloudSim, a free program, simulated TSPSO. In multi-objective work scheduling, the simulation results outperformed BRS and RSA in optimum balance. Simulation settings included datacenters, PE per datacenter, PE speed, power consumption, tasks, task duration, time, energy, and unsuccessful tasks. The suggested method accommodates dynamic datacenters and user jobs. This multi-objective technique successfully reduced energy and makespan leveraging system resources. MOPSO outperforms BRS and RSA. Cloud computing reinforcement learning-based random task scheduling P. Zhiping et al. [5] proposed fine-grained cloud computing system model and optimal job scheduling. To optimize work scheduling, authors developed a reinforcement learning and queuing theory-based technique. State aggregation methods accelerated learning. The authors constructed a simulation program in MATLAB and utilized factors like job duration, number of jobs, number of VMs, memory, bandwidth, buffer, PE needs, datacenters, and hosts for tests. Task scheduling efficiency, arrival rate, server rate, number of VMs, and buffer size were shown.

3) Hybrid Heuristic Workflow Scheduling for Cloud Computing.

M. Sahar et al. [6] combined Particle Swarm Optimization with Gravitation Search methods. The algorithm considered Processing Cost, Transfer Cost, and Deadline Constraints. End-users and utilities may utilize this method. CloudSim simulated all experiments. Compared to non-heuristic approach, PSO algorithm, gravitational search algorithm, and hybrid genetic-gravitational algorithm, simulation results revealed 70%, 30%, 30%, and 50% cost decrease.

Improved Particle Swarm Optimization for

Cloud Computing Work Scheduling Task scheduling is crucial to cloud computing, according to Awad et al. [7]. Time, cost, scalability, make span, dependability, availability, throughput, etc. to be addressed before completing a process. The suggested technique was reliable and available, unlike other cloud computing scheduling algorithms due to complexity. Load Balancing Mutation a particle swarm optimization (LBMPSO) was used to schedule activities based on characteristics such execution time, dependability, transmission time, round trip time, makespan, transmission cost, and load balancing between virtual machines and jobs. LBMPSO managed resources and rescheduled failure-causing tasks. LBMPSO outperformed regular PSO, random algorithm, and Longest Cloudlet to Fastest Processor (LCFP) in makespan, round trip time, execution time, transmission cost, and job assignment. This works for any resource or job.

5) Cloud-Based Dynamic Multi-Objective Task Scheduling Modified Particle Swarm Optimization-based

I. Awad [8] suggested cloud computing task scheduling research is important. The mathematical model multi-objective Load Balancing Mutation particle swarm optimization (MLBMPSO) schedules tasks to resources. Two aim functions reduced overall cost and round trip time. Compared to existing techniques, the suggested methodology enhanced cloud computing dependability, resource availability, and load balancing between virtual machines and activities. This method may assign any job or resource.

6) Agent-Based Cloud Computing Resource Allocation

F. E. Mohamed et al. [9] said that cloud computing resource allocation is a key problem. In cloud computing, consumers seek to decrease time and efficiency while providers want to lower money by boosting resource use. But, effectively using and sharing resources is tough. Researchers suggested using agents to connect cloud providers and consumers. The suggested method let users choose resources based on their needs. Simulations revealed that autonomous agents gave the cloud intelligence for user interactions and resource allocation.

Vacation Queuing Theory-Based Energy-Saving Cloud Computing Task Scheduling Chunling et al. [10] stressed cloud computing energy efficiency. Cloud computing wastes energy since incoming workloads are unpredictable and computer nodes must always be powered. Scientists suggested a queuing model-based energy-saving work scheduling system. First, vacation queuing model with extensive service reproduced task scheduling for heterogeneous cloud computing environment. Second, scientists examined computing node energy and time usage in heterogeneous clouds. Eventually, a similar-task scheduling method was devised to minimize energy use. The suggested approach performed better and reduced energy usage in simulations.

7) A New Cost-Based Cloud Computing Energy Consumption Model Horri et al. [11] suggested a time-shared policy energy consumption model for cloud computing virtualization layer. Based on actual system output, Cloud Sim simulator was used to simulate time-shared policy for cost and energy utilization. The suggested model was tested using different scenarios. The suggested method included data size-based cache interference costs. Following simulation, energy usage may be high and vary with factors like quantum parameter,

data size, and host virtual machine count. The outcome validated the model and indicated that cloud computing systems exchange QoS for energy.

Cloud-Based Virtual Machine Scheduling

A generalized precedence method by T. Yousef et al. [12] outperformed FCFS and Round Robin Scheduling. CloudSim's inconsistent number of virtual machines workload results showed excellent output compared to traditional scheduling methods.

9) Ant Colony Scheme-based Dynamic Work Scheduling.

N. Kamolov et al. [13] dubbed optimum solution discovery NP-hard. Ant colony optimization may create effective scheduling algorithms. This study proposes a modified ant colony optimization job scheduling method. WorkflowSim measured performance using the recommended strategy. Following simulation, the Probabilistic Load Balancing Algorithm (PLAC) lowered average makespan by 6.4% compared to Ant colony Optimization and 11.5% compared to Min-min.

10) Cloud workload balancing optimized for service level agreements

S. Rajeshwari et al. [14] introduced a two-stage scheduling technique. CloudSim built the technique utilizing reaction time as a parameter. During trials, the suggested method had superior response time, resource usage, waiting time, and server load balancing than current techniques.

11) Assignments Cloud Computing PSO Algorithm

M. Ali et al. [15] created a Dynamic Adaptive Particle Swarm Optimization method (DAPSO) to improve basic PSO by lowering make-span and increasing resource usage. MDAPSO combines Dynamic PSO with Cuckoo Search. MDAPSO and DAPSO outperformed PSO in simulations.

12) Cloud computing infrastructure map-reduce job scheduling evaluation

Qutaibah et al. [16] tested FIFO, Matchmaking, Delay, and multithread-ed locality on virtualized systems. Simulation duration and energy usage tested the algorithms. MTL outperformed current schedulers.

EVALUATION OF ALGORITHMS

We examined scientists' newest job scheduling systems. Table 1 compares recent work scheduling techniques. Compare Make span, Response Time, Degree of Imbalance, Execution Time, VM Buffer, Server rate, Round Trip Time, Transmission Time, Task Distribution, and Energy Usage.

CONCLUSION

This article examines Cloud Computing load balancing and scheduling techniques and associated difficulties. Review study shows that each scheduling method has pros and cons in its use. This document summarizes cloud computing task scheduling studies. Optimization, GA, Improved PSO, ACO, Queuing Theory, Agent-Based approach, Multi-Objective approach, and others are employed in current methods. These techniques minimized makespan, reduced execution time, transmission time, energy consumption even when workloads and VMs increased, and optimized VM buffer size.

Symbiotic Organism Search Optimization improved makespan, reaction time, and imbalance for a wider search area. Queuing Model-based Genetic Algorithm outperformed FCFS. FCFS and Priority Based algorithms were slower than Multi-

Objective task scheduling technique.

Nested PSO multi-objective task scheduling outperformed BRS and RSA. Reinforcement-based Random Task Scheduling balanced loads and scheduled tasks successfully. Hybrid Heuristic Workflow Scheduling reduced costs better than PSO, Gravitational Search, and Hybrid Gravitational Search.

Improved Particle Swarm Optimization maximized round trip, execution, transmission, and load balancing efficiency. Dynamic Modified PSO-based multi-objective task scheduling improves task execution time, cost, and node distribution. Agent-Based Method for Resource Allocation gave the cloud user interaction and resource allocation intelligence. PLAC with modified PSO outperformed ACO and Max-min in minimum and average makespan. As mentioned in our research, several businesses are adopting cloud computing, and cloud computing users are growing. A better task scheduling algorithm should plan user jobs to improve system performance with low cost.

Meta-heuristic, machine learning, and hybrid methodologies may improve scheduling algorithms in the future. Resource efficiency will boost throughput and reduce task completion time. This will lower operating costs and increase cloud computing adoption.

REFERENCES

1. *S.Rekha, Dr.C.Kalaiselvi Review Of Scheduling Methodologies Of Virtual Machines (Vms) In Heterogeneous Cloud Computing, ijstr International Journal of Scientific & Technology Research Volume 8, ISSUE 09,PP-806-812, Sep 2019*
2. *S. Supreeth and Kiran Kumari Patil Virtual Machine Scheduling Strategies in Cloud Computing- A Review, International Journal on Emerging Technologies,, pp-181- 188, June 2019*
3. *M. Abdullahi, M. A. Ngadi and S. M. Abdulhamid, "Symbiotic Organism Search Optimization Based Task Scheduling in Cloud Computing Environment," Future Generation Computer Systems, Elsevier, vol. 56, pp 640- 650, 2016.*
4. *S. Saha, S. Pal, and P. K. Pattnaik, "A Novel Scheduling Algorithm for Cloud Computing Environment," Computational Intelligence in Data Mining, Springer India, Vol. 1, pp. 387-398, 2016.*
5. *A. V. Lakra and D. K. Yadav, "Multi-Objective Tasks Scheduling Algorithm for Cloud Computing Throughput Optimization," Procedia Computer Science, Vol. 48, pp. 107-113, 2015.*
6. *R. K. Jena, "Multi-Objective Task Scheduling in Cloud Environment using Nested PSO Framework," Procedia Computer Science, Vol. 57, pp. 1219-1227, 2015.*
7. *P. Zhiping, C. Delong, Z. Jinglong, L. Qirui, X. Bo and L. Weiwei, "Random task scheduling scheme based on reinforcement learning in cloud computing," Cluster Computing, Vol. 18, Issue 4, pp. 1595-1607, 2015.*
8. *M. Sahar and R. Vahid, "A Hybrid Heuristic Workflow Scheduling Algorithm for Cloud Computing Environments," Journal of Experimental & Theoretical Artificial Intelligence, Vol. 27, Issue 6, pp. 721-735, 2015.*
9. *A. I. Awad, N. A. El-Hefnawy and H.M. Abdel_Kader, "Enhanced Particle Swarm Optimization For Task Scheduling in Cloud Computing Environments," Procedia Computer Science, Vol. 65, pp. 920-929, 2015.*
10. *A. l. Awad, N. A. El-Hefnawy and H.M. Abdel_Kader, "Dynamic Multi-Objective Task Scheduling in Cloud Computing Based on Modified Particle Swarm Optimization," Advances in Computer Science: an International Journal, Vol. 4, pp. 110-117, 2015.*
11. *F. E. Mohamed, K. Okba, F. Manel and B. Samir, "An Agent- Based Approach for Resource Allocation in the Cloud Computing Environment," IEEE, 2015.*
12. *C. Chunling, L. Jun and W. Ying, "An Energy-Saving Task Scheduling Strategy Based on Vacation Queuing Theory in Cloud Computing," Tsinghua Science and Technology, Vol. 20, Issue 1, pp 28-39, 2015.*
13. *A. Horri and Gh. Dastghaibyfard, "A Novel Cost Based Model for Energy Consumption in Cloud Computing," The Scientific World Journal, pp-110, 2015.*

14. *T. Yousef, A. Siamak, A. S. Zahed, A. Abdolsalam and M. Mohammad, "Virtual Machine Scheduling in Cloud Computing Environment," International Journal of Managing Public Sector Information and Communication Technologies, Vol. 6, Issue 4, pp 1-6, 2015.*
15. *B. N. Kamolov, C. Tae-Young, "Dynamic Task Scheduling Algorithm based on Ant Colony Scheme," International Journal of Engineering and Technology, Vol. 7. No. 4, 2015.*
16. *B. S. Rajeshwari and M. Dakshayini, "Optimized service level agreement based workload balancing strategy for cloud environment," IEEE, pp. 160-165, 2015.*
17. *A. Ali and O. A. Fatma, "Task Scheduling Using PSO Algorithm in Cloud Computing Environments," International Journal of Grid Distribution Computing, Vol. 8, No. 5, pp. 245-256, 2015.*
18. *Ritu Kapur, "A workload balanced approach for resource scheduling in cloud computing," IEEE, pp. 36-41, 2015.*
19. *M. Kalra and S. Singh, "A Review of Metaheuristic Scheduling Techniques in Cloud Computing," Egyptian Informatics Journal, pp 1-21, 2015.*
20. *[19] G. Kulkarni, R. Sutar and J. Gambhir, "Cloud Computing- Infrastructure As Service-Amazon EC2", International Journal of Engineering Research and Applications, Vol. 2, Issue 1, 2012.*
21. *M. Ahmed, M. S. Abu, R. Chowdhury, A. Mustaq and M. M. Hasan Rafee, "An advanced survey on cloud computing and state- of-the-art research issues,"International Journal of Computer Science Issues, Vol. 9, Issue 1, pp. 0814-1694, 2012.*
22. *S. Rekha and R. S. Kumar, "Priority Based Job Scheduling for Heterogeneous Cloud Environment," International Journal of Computer Science Issues, Vol. 11, Issue 3, pp. 114-119, May 2014.*
23. *K. Etminani, M. Naghibzadeh and N. R. Yanehsari, "A Hybrid Min-Min Max-Min Algorithm with Improved Performance," Department of CE, Ferdowsi University of Mashad and Department of IT, Iran Khodro Khorsan, Iran.*
24. *P. Slot, "A Survey of Various Scheduling Algorithms in Cloud Computing Environment," International Journal of Research in Engineering and Technology, Vol. 02, Issue 2, pp. 131-135, Feb. 2013.*