ORIGINAL RESEARCH PAPER

Computer Science

LOAD BALANCING IN CLOUD COMPUTING WITH GENETIC ALGORITHM BASED HYBRIDIZATION

KEY WORDS:

Thummuluru Kavitha Research Scholar, Dept. Of Computer Science, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, Andhra Pradesh, India.

Dr.Thatimakula Sudha Professor, Dept. Of Computer Science, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, Andhra Pradesh, India

ABSTRACT

The scheduling of computer resources in the cloud computing system is a major part of cloud computing research. The work throughput of cloud computing can be significantly enhanced by implementing an effective load balancing technique. Cloud computing relies on virtual machines as a primary processing unit. Due to virtualization technology, cloud computing resources will rise rapidly and dynamically. Virtual machines can be dynamically balanced by employing an improved Q-learning technique called QMPSO, a mix of the enhanced Particle swarm optimization (MPSO) and the QMPSO algorithm. The hybridization method changes the MPSO's velocity using the gbest and pbest based on the best action given by improved Q-learning. Hybridization is an effort to improve system performance by evenly disbursing work among virtual machines, increasing throughput and maintaining task priorities in balance by optimizing the time it takes for processes to wait. Comparing the QMPSO simulation results with the current load balancing and scheduling technique proved the algorithm's robustness. Simulated and real-platform results show that our algorithm outperforms the competitors.

1.0 INTRODUCTION

Cloud computing is a certain trend in the future evolution of computing technology. All users can benefit from its excellent performance and efficient calculating capabilities. Distribution, grid computing, and many more strategies have evolved into the cloud computing technology. [1] A major difference between cloud computing and the preceding large-scale cluster computing is the usage of a distributed processing unit like such a distributed grid. An individual physical host can be split equally and used as a basic computing unit in the cloud by virtue of the virtualization technology [2]. Cloud computing in tandem with traditional cluster computing may dramatically increase system usage and also perform automatic monitoring for all hosts by using virtualization technologies. Virtualization technology has not only made cloud computing more convenient, but it has also made a great number of virtual resources available. Virtual resources are both large and ever-changing in scope. One of the key concerns in research is how to evenly distribution of load on a cloud host.

Problem of the statement:

Virtual machines identical to the central server can be created by any server that accepts the request. Creating a virtual machine on the server is restricted by the system model's memory consumption. There are 'n' separate operations running on separate virtual computers at the same time. A better scheduler will be needed to ensure that the workload is evenly distributed throughout VMs [4]. To develop a fitness function, load balancing in a cloud computing environment is necessary. The main objective of load balancing is to maximize the usage of resources and accept as many requests as necessary. Using the parameter U (RPHPU, RCPU, RM, C) for a single user, RPHPU indicated the number of online user requests coming from every time stamp on average in a user group, RS refers to the request size of each user in the user group, RCPU refers to the amount of CPU required to execute the request, RM indicates the amount of memory needed to execute the request. C refers to how many requests were sent in a minute for a user group. In light of the organization strategy.

Scope of the work:

In cloud computing, information technology processes are distributed throughout a broad network rather than performed locally on any device. Proposed a GA-based scheduling technique for cloud computing VM load balancing. One of the key challenges in cloud computing is ensuring a high level of system uptime and availability, which could help address that. Did a study on the characteristics of the cloud. When it comes to virtualization, a powerful virtual host is a basic foundation. Then, use virtualization-virtualization technology to manage resources and virtualization storage. Evolutionary computation, particularly its branch in optimization techniques, has gained increasing attention in recent years due to its intelligence and inferred parallelism. As a result of its widespread use in large-scale, nonlinear cluster systems, GA has attained optimum results.

2.0 Literature Review

Quang-Hung et al. [4] The genetic algorithm for scheduling resources in a power-efficient manner was discussed. A genetic method for power-conscious virtual machine allocation was proposed by the authors (GAPA). Through intelligent VM movement decisions among virtualized cluster servers, Tarighi et al. [5] improved the performance of virtualized servers. Fuzzy logic is used to make intelligent judgements. Multi-Criteria Decision Making was used to move the virtual machines from an overloaded node to an underloaded node. [6] Beloglazov et al. [6] presented an architecture for the efficient administration of resources at data centres. The authors came up with a set of parameters for determining the best location for virtual machines. The recommended policies for picking the most efficient migrated VMs and physical nodes. The findings showed that the proposed strategies were the most effective at allocating virtual machines efficiently. Anton Beloglazov et al. [7] suggested a new technique based on a dynamic threshold value for the number of virtual machines in an IaaS environment. The authors presented three policies to maximize energy utilization. The experimental findings of the suggested approach demonstrated that this unique technique had fewer SLA violations and VM migrations. GCE was defined and developed by Awada et al. [8] to reduce the energy consumption of servers. The authors also considered static and dynamic cloud resource cases to minimize power consumption. The researchers came up with a few ideas on how to deal with energy efficiency problems. An energy efficient algorithm's creation and design were discussed by S. Subbiahet et al. [9]. According to the researchers, data centers'high-power usage has become a major issue in the IT industry. The proposed work aided in solving a resource management problem involving thermal and power dissipation. For efficient power usage, experimental results demonstrated that the proposed work was effective.

Hybrid meta-heuristic algorithm for load balancing:

An algorithm described as Q-learning allows agents to learn from their environment and take action by changing their state in order to obtain a reward or penalty. This algorithm is used in machine learning and can be found here. When transitioning from one state to another, an agent uses its control strategy to select the best feasible course to take from a list of available actions, which is its main objective. It is necessary to select the suitable action that maximizes the Q-value of each state in order to establish an optimal policy in the cloud network, which is known as a Markov decision process (MDP) with unknown probabilities of transition t. The Q-value function essentially depends on the selection criterion for action in the given system. To compute Q-value, consider an agent in the state st and select an action that is predicted to take them to their optimal next state and maximize their overall asset in the future:

$$Q(s_t, a_t) = (1-\alpha)Q(s_t, a_t) + \alpha\left[r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1})\right]$$

As a consequence, a static load balancing method will not be prepared to accommodate the dynamic nature of the workload. It is better to use dynamic load balancing rather than static load balancing when loads change during operation and need to be taken into account. It is vital to use dynamic strategies to balance the load among diverse resources because of the rapid expansion of the network and its need for resources during the run period. Using a hybrid meta-heuristic approach, we've found a way to balance the workload and prioritize tasks in the VMs' queues. To avoid overloading the cloud network, remove the tasks from the fully loaded virtual machines and relocate them to underutilized ones when they've been allocated to a single VM and additional VMs are available on it. As a result, numerous jobs can be dispersed throughout all VMs with varied priority, reducing waiting times and increasing throughput of VMs while also performing load balancing at Vms.

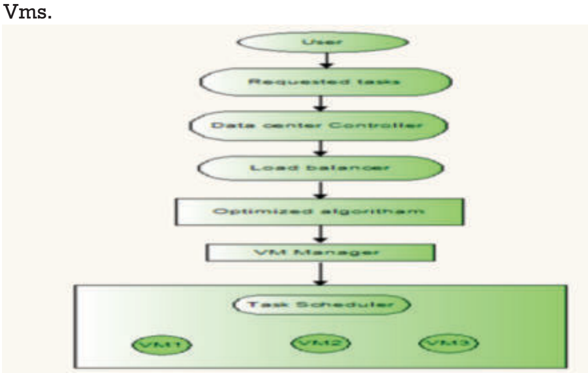

Figure: The system model of load balancing flow chart

As parts of the load balancing process, VM Managers are selected and applied tasks by the load balancer. It checks the existing VMs, the quantity of resources needed by tasks, and whether or not resources are available on the host, and if there aren't enough, the VM manager generates new ones. In this method, the VMs' fitness levels will be used to distribute the load. One VM per host is the maximum.

Experimental results and performance analysis

In order to evaluate the suggested algorithm's performance, the simulation results were used. Using the CloudSim3.0.3 simulator, a machine with an Intel core i7 processor, 8 GB of RAM, a 3.4GHz CPU, and Windows 10 operating system was used for the experiment. Table 1 provides an overview of the experiment's simulation setting. There was an evaluation of the algorithm's performance based on task migration, task response time, latency in all tasks, time spent idle, and task spans before and after balancing through a modified PSO and improved Q-learning. ' The algorithm has been tested in two different scenarios. In the first case, the number of virtual machines (VMs) is fixed at 100 and the number of tasks is varied from 100 to 2000 over a period of 50. In the second case, the number of tasks is maintained at 1000 and the number of VMs is varied over a period of 50 from 10 to 100. After analysing these two instances, the algorithm's performance has been measured.

Table 1: Influence of parameters for Load balance in cloud computing

Weights	PerformanceAnalysisandConvergencerate					
1	K2	K3	Make span (In ms)	Throughput (req/ms)	Standard deviation (SD)	Load Balance (yes/No)
1	0.4	0.9	8321.82	7.98	0.373	yes
		0.8	8148.29	7.39	0.314	yes
		0.7	8093.21	7.12	0.274	yes
		0.6	7930.62	6.87	0.238	yes
		0.5	7832.31	6.45	0.208	yes
2	0.5	0.9	8894.51	6.23	0.275	yes
		0.8	8467.92	7.39	0.174	yes
		0.7	8135.29	7.18	0.096	yes
		0.6	7954.64	6.74	0.082	yes
		0.5	7791.32	5.83	0.067	yes
3	0.6	0.9	9087.41	9.18	0.313	yes
		0.8	8794.51	8.28	0.298	yes
		0.7	8663.95	8.07	0.287	yes
		0.6	8532.13	7.89	0.259	yes
		0.5	8374.30	7.45	0.238	yes

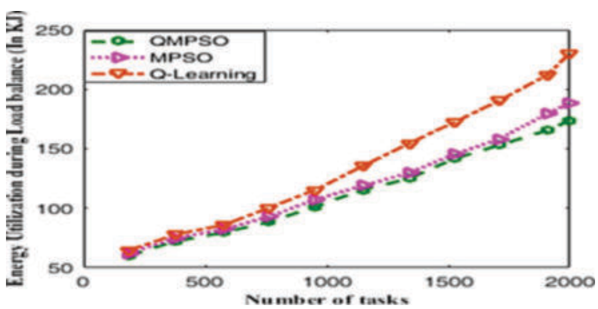

Figure 1: Number of tasks with Energy utilization for fixed number of VMs

Measurement of cloud performance and service quality by obtained tasks (QoS), There are three metrics that may be used to quantify imbalance in a VMS: the degree of imbalance, the quantity of idle time, and the processing time.

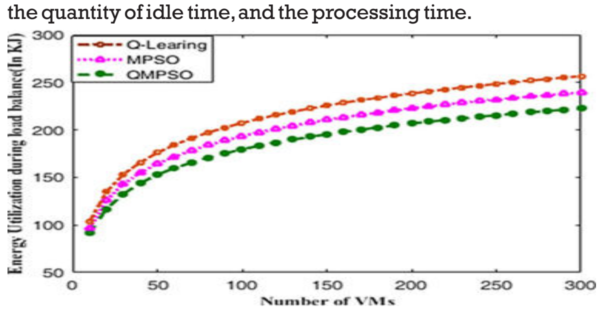

Figure 2: Number of VMs with Energy utilization for fixed number of Tasks

Load balancing using QMPSO is the most energy-efficient compared to its competitors. Efforts have been made to improve the make span of the performance. Make-span measures the user's response time for a specific task, and as a result, the service provider can guarantee the quality of customer service.

In comparison to MPSO and Q-Learning, QMPSO's original standard deviation value is lower, but it rapidly reduces as the model gets better. QMPSO and MPSO have the same standard deviation at time t = 3700 ms. There are some differences between MPSO's and QMPSO's standard deviations when measured at 3700.As a result, the proposed QMPSO gets optimal resources quickly, allowing for the remaining computational power to be utilized. A comparison of the QMPSO and its competitors shows that the QMPSO technique has a higher capacity for resource utilization and a superior ability to balance load. When it comes to task migration, the results have been analyzed.

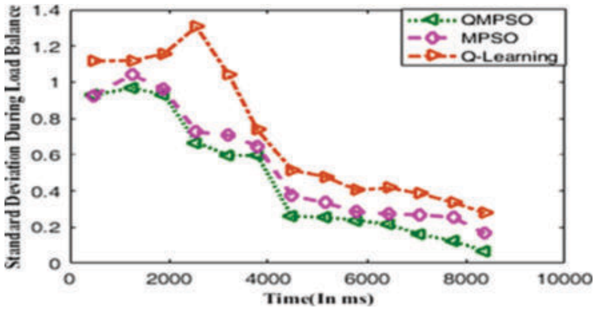

Figure 3: Time with standard Deviation for QMPSO, MPSO and Q-Learning

In comparison to MPSO and Q-Learning, QMPSO's original standard deviation value is lower, but it rapidly reduces as the model gets better. QMPSO and MPSO have the same standard deviation at time t = 3700 ms. There are some differences between MPSO's and QMPSO's standard deviations when measured at 3700.As a result, the proposed QMPSO gets optimal resources quickly, allowing for the remaining computational power to be utilized. A comparison of the QMPSO and its competitors shows that the QMPSO technique has a higher capacity for resource utilization and a superior ability to balance load. When it comes to task migration, the results have been analyzed.

Table 2: Analysis of performance for QMPSO, MPSO and Q-Learning

Algorithm	Make span(In ms)	Throughput(req/ms)	Standarddeviation(SD)	Energyutilization(InKJ)		
				forfixedVMs	forfixedTasks	TaskMigrated
QMPSO	8791.20	5.42	0.069	163.80	220.45	384.25
MPSO	9045.50	4.79	0.172	185.28	235.56	420.84
Q-Learning	9332.50	4.42	0.278	220.21	247.67	467.88

The comparison has been done with its rival algorithms, and the results have been summarized in a table. Finally, the results reveal that when it comes to cloud network load balancing, the QMPSO outperforms both MPSO and Q-learning. The algorithm's robustness has been tested on a real platform. The cloud data center on the real platform has four hosts, each of which is capable of supporting virtualization technologies.

CONCLUSION:

Cloud security provides a data encryption service to encrypt cloud data before exchange from local storage to cloud storage, and it is impossible to understand from every system. A hybrid metaheuristic algorithm such as QMPSO has been proposed for load balancing for independent tasks in the cloud computing network. A new method is devised based on each simulated machine's fitness value to distribute the work effectively. The proposed algorithm also improves the make span, throughput, energy utilization during load balancing. It reduces the tasks' waiting time effectively compared to separated algorithms such as MPSO and Q-learning. At last, we have also compared our proposed algorithm with the existing algorithm and found that our proposed algorithm outperforms the existing algorithm. In the future, load balancing will be carried out among the dependent tasks dynamically.

REFERENCES:

1. A., Juneja, D., Malhotra, M., 2015. Autonomous Agent Based Load-balancing algorithm in Cloud Computing. International Conference on Advanced Computing Technologies and Applications (ICACTA), 45, 832–841
2. Selvi, R.T., Aruna, R., 2016. Longest approximate time to end scheduling algorithm in Hadoop environment. Int. J. Adv. Res. Manag. Archit. Technol. Eng. 2 (6)
3. Krishna, P. Venkata, 2013. Honey bee behavior inspired load balancing of tasks in cloud computing environments. Appl. Soft Comput. 13 (5), 2292–2303.
4. Deng, Y., Lau, R.W., 2014. Dynamic load balancing in distributed virtual environments using heat diffusion. ACM Trans. Multimedia Comput. Commun. Appl. (TOMM) 10 (2), 16.
5. Zhu, Y., Zhao, D., Wang, W., and He, H. (2016, January) 'A Novel Load 460 Balancing Algorithm Based on Improved Particle Swarm Optimization in Cloud Computing Environment', In International Conference on Human Centered Computing, Springer, pp. 634–645.
6. Quang-Hung, Nguyen, Pham Dac Nien, Nguyen Hoai Nam, Nguyen Huynh Tuong, and Nam Thoai. A genetic algorithm for power-aware virtual machine allocation in private cloud. In Information and communication technology, pp. 183-191. Springer Berlin Heidelberg, 2013.
7. Tarighi, Mohsen, Seyed A. Motamedi, and Saeed Sharifian. A new model for virtual machine migration in virtualized cluster server based on fuzzy decision making. arXiv:1002.3329 (2010).
8. Anton Beloglazov, and Rajkumar Buyya. Energy efficient resource management in virtualized cloud data centers. In Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, pp. 826-831. IEEE Computer Society, 2010.
9. Beloglazov Anton, and Rajkumar Buyya. Adaptive threshold-based approach for energy-efficient consolidation of virtual machines in cloud data centers. Proceedings of the 8th International Workshop on Middleware for Grids, Clouds and e-Science. ACM, 2010.
10. Awada, Uchechukwu, Keqiu Li, and Yanming Shen. Improving cloud computing energy efficiency. In Cloud Computing Congress (AP Cloud CC), 2012 IEEE Asia Pacific, pp. 53-58. IEEE, 2012.
11. Subbiahet, Sankari, and Varalakshmi Perumal, Power aware resource optimization in cloud. In Recent Trends in Information Technology, International Conference on, IEEE, 2013, pp. 318-322