Analysis of Different Load Balancing Techniques

(IJCSNT) International Journal of Communication Systems and Network Technologies
ISSN Number: 2053-6283
Volume 13, Issue No. 1, 1–7

https://journals.mirlabs.in/index.php/ijcsnt
DOI-10.18486/ijcsnt/13.1.170

Gundipika Kaur[1], Kiranbir Kaur[2], Prabhsimran Singh[3]

Abstract

Cloud computing is a popular paradigm which provides on-demand services to the user. As needs of the user changes dynamically, it is the responsibility of the cloud service provider to handle the requests efficiently and to satisfy the demands of end user by ensuring quality of service. Cloud computing suffers from umpteen critical problems which needs to be explored. One such problem is Load Balancing. It distributes the load evenly among the nodes with a condition that no overloading of the node occurs. So, it is better to schedule the tasks in an intelligent manner to utilize the resources efficiently. Task Scheduling is an important research area in cloud computing. It deals with the NP-hard problems. To find the near optimal solution, meta-heuristic techniques are used. In this paper, an in-depth review of different task scheduling algorithms based on load balancing have been discussed. We have compared these algorithms on the basis of pros and cons.

Keywords

Cloud Computing, Load Balancing Algorithms, Task Scheduling

Received: 01 January 2024; **Revised:** 30 March 2024; **Accepted:** 20 April 2024;
Published: 30 April 2024;

[1, 2]Department of Computer Engineering & Technology, Guru Nanak Dev University, Amritsar, India
[3]P.G. Department of Computer Science and Applications, Khalsa College, Amritsar, India

Corresponding author:
Email: gunbhatia@yahoo.com

Introduction

Cloud Computing is a promising technology because of dynamic provisioning[1] of the resources on demand. It is one of the popular paradigms for delivering computing as a utility. It leverages different technologies such as virtualization and utility-based pricing to meet the user's demand. Cloud computing provides the facility to access the shared pool of resources. As the needs of the users changes dynamically, Cloud computing provides services to the end user efficiently. Users and their demands have been increasing tremendously. So, it is the responsibility of the cloud service provider to satisfy the demands of the end user by ensuring Quality of Services(QOS).

There are umpteen critical issues which need to be explored. One such issue is Load balancing[2]. It is a method of distributing the workload among the nodes to avoid the condition of overloading of requests on one node. It deals with efficient utilization of the resources.

Need of Load Balancing

As the resources are heterogeneous in nature, it is possible to get the condition of overloading of requests on one node and its queue size will be increased. So, to avoid this situation, load balancing is required to balance the dynamic workload in multiple nodes. The main aim is to satisfy the user's requirement and utilize the resources[3] in an efficient manner. Different metrics such as throughput, fault-tolerance, response time, scalability etc are used to balance the load.

Task scheduling and resource allocation are the two major tasks of load balancing which are used to reduce the cost of resources and save the energy in the low load. To meet the user's requirement, a good scheduling algorithm is required[4]. The mapping of tasks to appropriate resources is a main goal of scheduling. Physical servers act as a host which provides services to the end users. Tasks are executed on the virtual machines(VM). Once the allocations of resources are done on the cloud entities, scheduling of tasks will take place.

Scheduling of tasks is necessary for mapping the tasks to provide the QOS guarantee to the end user. Task scheduling is considered as a NP-hard problem because some algorithms do not provide optimal solution. To find the near optimal solution, meta-heuristic techniques can be used.

Type of Scheduling

Two main categories of scheduling are:
Independent task scheduling: when tasks are scheduled in different sequences, it means tasks are independent of each other. This is called independent task scheduling.
Workflow task scheduling: If a task is scheduled only after its parent task is completed, it means tasks are dependent on each other. This is called workflow task scheduling.

Rest of the paper is organized as follows: Section IV discusses the Literature survey. Section V presents the comparison table based on merits and demerits. Finally we have given the conclusions.

Literature Survey

Yiqiu Fang and Fei Wang[5], discussed task scheduling in two levels to balance the load in an efficient manner. As the requirements of the user changes dynamically, it is possible that a task can be scheduled

directly to the host resources for execution. The two level scheduling mechanism increases the utilization of resources.

Kousik Dasgupta and Brototi Mandal[6], proposed a load balancing strategy using genetic algorithm. This algorithm reduces the make span of tasks. Genetic algorithms is considered as one of the most efficient techniques used for optimization. It has been proposed in cloud computing to find the processors for task. The main purpose is trapping into local optimal solution is being avoided to get the global optimal solution due to the complexity in vast search space. The proposed algorithm performed well after comparing with the existing approaches like FCFS, Round-Robin and Stochastic Hill Climbing.

P.Venkata Krishna[7], proposed the Honey Bee Behavior load balancing (HBB-LB) algorithm to minimize the make span and balance the load across virtual machines in an intelligent manner. It is a dynamic technique which gives emphasis on the priorities of tasks in the waiting queues of VM's. Removed tasks from overloaded VM acts as honeybees and the destination of honeybees will be an under-loaded VM. Priority is taken as a main QOS parameter. Performance is analyzed with the help of cloudsim. It is observed that this algorithm is an effective approach to minimize the waiting time of tasks on the queues.

Yingchi Mao and Xichen[8], discussed Max-Min task scheduling algorithm in which executing task status and virtual machine status table is maintained. Execution time and completion time of task are included in the task table. Real time load of the virtual machine is estimated which helps in improving the utilization of resources. Longest execution time (Max) of a task is selected and allocated to the VM whose completion time is shortest(Min). This algorithm is considered as a batch mode scheduling algorithm. Tasks arrive in batches which are handled by this algorithm efficiently.

Shridhar G.Domanal and G.Ram Mohana Reddy[9], proposed Modified Throttled algorithm. It selects the VM at first index on the basis of the state of the VM. When the next request arrives, it will find the virtual machine at next index to the already assigned VM. But in throttled algorithm, selection of VM is done from the first index every time. The problem of under-loaded and overloaded VM's is improved by utilizing the VM properly.

Mala Kalra and Gulshan Soni[10], presented "Central Load Balancer" a load balancing algorithm to balance the load among VM's in cloud data centre to achieve better utilization of resources. This algorithm performs a reliable load balancing on the basis of priority and states of the VM's. The CLB calculates the priorities of the VM based on their CPU speed and memory. Efficiency of VM load balancing is tested by considering constant load in first case and by keeping numbers of VM constant in second case. The proposed algorithm is implemented in Cloud Analyst which is used to evaluate large scale applications.

Santanu Dam[11], proposed a novel ant colony based algorithm which gives optimal resource utilization. An index table is maintained to keep the record of the current allocation of VM's. Redistribution of workload is done among nodes. The proposed algorithm helps ants to finds an optimal path depends on pheromone value and guarantee the QOS to the end user. This will maximize the throughput but the main problem is that this algorithm considers only the priority as the main parameter.

Shridhar G.Domanal and G.Ram Mohana Reddy[12], proposed "VM-assign Load balancing algorithm to distribute the load on virtual machine efficiently. In this algorithm, assigning of VM is totally based on its load i.e by finding the least loaded VM and incoming jobs are allocated to that VM. Proposed algorithm is compared with an Active Load Balancing algorithm because the proposed algorithm will look for least loaded VM. It follows RR approach on finding all VM's free. If the next request arrives,

then it will check the VM table. Once the we get the available VM, it checks for the previous assignment. But it is not in the case of Active load balancing algorithm.

Aarti Singh and Dimple Juneja[13], discussed A2LB algorithm for addressing the issues such as resource utilization and response time. This proposed algo used the concept of 3 agents that is Loan agent, Channel agent and Migration agent. Being an ant , migration agent belongs to the category of mobile agents. These agents helps to find the shortest path to destination. These agents reduces the network burden by searching under-loaded servers. The proposed A2LB algorithm is implemented by using JAVA technology.

Jengshyang Pan and Pingfei Ren[14], proposed an improved online load balancing algorithm. It considered the task level load balancing. Research was done in heterogeneous cloud computing environment. Efficiency is improved by using this proposed algorithm. Comparison of different algorithm has been done by using Cloudsim. It is observed that Round-robin algorithm's performance is worst among all the algorithms in the experiment. Task level load balancing is done in this paper.

Fahimeh Ramezani[15], discussed an optimization model based on scheduling of tasks by using Particle Swarm Optimization. In this model, when extra tasks are transferred from an over-loaded VM, it will save time in balancing the load efficiently. With the use of this proposed method, VM downtime and use of memory will be eliminated. Instead of migrating entire VM[17], tasks of an overloaded VM is transferred to balance the load. Cloudsim is extended and apply with Jswarm.

Saurav Banergee[16], proposed a new cloudlet allocation strategy to balance the load by allocating cloudlets to VM according to the VM's capacity. This strategy improves the utilization of system as well as the VM's makespan. Capacity of the VM is calculated before next task(cloudlet) is allocated to the VM . This will ensure the QOS parameters such as completion time of the tasks. Performance is evaluated and analyzed by comparing proposed work with RR and Conductance algorithm by using Cloudsim.

Comparative Analysis

A lot of work has been done on load balancing algorithms by scheduling the tasks. Table 1 shows the comparison of the above mentioned algorithms.

Table 1. Comparison of Different Algorithms

Paper	Parameter	Simulator	Description	Merits	Demerits
Autonomous Agent Based LB in cloud computing [13]	Service time	Java technology	Different agents are used to search the under-loaded servers and resources to reduce the burden on the network.	Reduces service time	Considers the candidate VM's information beforehand
Honey bee behavior inspired LB of tasks in cloud computing [7]	Makespan Throughput	Cloudsim	HBB-LB minimizes the makespan by scheduling the tasks in an efficient manner. The tasks are considered as honey bees.	Improvement in waiting time in queue	Considers only priority of tasks
An Ant colony based LB strategy in cloud [11]	QOS	Cloud-Analyst	Strategy is used to initiate load balancing by finding an optimal path with pheromone concentration.	Guarantees QOS	It does not consider fault tolerance and same priority of all jobs are taken.
A Task scheduling algorithm based on LB in cloud [5]	Flexibility	Cloudsim	Two level scheduling strategy is used to meet the user's requirement.	It avoids overloading of hosts for holding more resources	Factors such as cost and bandwidth are not considered
Max-Min task scheduling algorithm for LB in cloud [8]	Response time, Elasticity	Cloudsim	The ECMM task scheduling improves the resource utilization of estimating the real time VM's load.	Deals with uncompleted tasks	Does not work for the tasks in same batch.
A Novel Approach for Load Balancing in Cloud Data Center [10]	Response time	Cloud analyst	The algorithm performs reliable load balancing by calculating the priorities of VM.	Improves the response time	Does not consider current utilization of processor and memory
Optimal Load Balancing in Cloud Computing By Effiient Utilization of Virtual Machines [12]	Response time	Cloudsim	VM is assigned on the basis of load and proposed algorithm avoids the problem of under utilization of resources.	Response time is improved	Response of an algorithm by mixing the static and dynamic load is not considered
Task based system load balancing in cloud using PSO [15]	Energy consumption, memory usage	Cloudsim	Load balancing is possible when extra tasks are transferred from an overloaded VM.	Execution time and transfer time of tasks is reduced	Multi-Objective of PSO is not considered
Genetic Algorithm based LB strategy for cloud [6]	Makespan	Cloudsim	This algorithm finds a solution which is a global optimum in nature	Guarantees QOS	Rescheduling of job is not considered
Research on Heuristic based LB algorithm in cloud [14]	Efficiency	Cloudsim	It gives emphasis on task level LB.	Improved system performance	SLA based load balancing is not considered
Development and Analysis of a new cloudlet allocation strategy for QOS improvement in cloud [16]	Makespan	Cloudsim	The proposed algorithm allocated the cloudlet to VM after calculating the capacity of VM.	Guarantees QOS	Different threats have not been defended
LB in cloud computing using Modified Throttled Algorithm [9]	Response time	Cloud-analyst	Proposed algorithm is compared with Throttled and RR algorithm. This algorithm maintains an index table on the basis of which VM is selected first depending upon state of VM	Response time improved, avoids the queuing at server by checking the state	Data structures are not changed

Conclusions

Load Balancing is one of the hot topics in cloud computing. Different load balancing algorithms have been proposed which helps in an efficient scheduling of tasks. But every algorithm has its own pros and cons. The main aim of most of the author was to reduce the makespan and response time of the tasks whereas some authors gave importance to throughput, resource utilization, fault tolerance and scalability etc. An improvement is required in the direction of metaheuristic techniques such as BAT for task scheduling without violating the constraints of SLA. Critical issues that we have explored till now will be taken for future research.

References

[1] Zhang, Qi, Lu Cheng, and Raouf Boutaba. "Cloud computing: state-of-the-art and research challenges." Journal of internet services and applications 1.1 (2010): 7-18.

[2] Manvi, Sunilkumar S., and Gopal Krishna Shyam. "Resource management for Infrastructure as a Service (IaaS) in cloud computing: A survey." Journal of Network and Computer Applications 41 (2014): 424-440.

[3] Liu, Lixia, Hong Mei, and Bing Xie. "Towards a multi-QoS human-centric cloud computing load balance resource allocation method." The Journal of Supercomputing (2015): 1-14.

[4] Pan, Jeng-Shyang, et al. "Interaction Artificial Bee Colony Based Load Balance Method in Cloud Computing." Genetic and Evolutionary Computing. Springer International Publishing, 2015. 49-57.

[5] Fang, Yiqiu, Fei Wang, and Junwei Ge. "A task scheduling algorithm based on load balancing in cloud computing." Web Information Systems and Mining. Springer Berlin Heidelberg, 2010. 271-277.

[6] Dasgupta, Kousik, et al. "A genetic algorithm (ga) based load balancing strategy for cloud computing." Procedia Technology 10 (2013): 340-347.

[7] Krishna, P. Venkata. "Honey bee behavior inspired load balancing of tasks in cloud computing environments." Applied Soft Computing 13.5 (2013): 2292-2303.

[8] Mao, Yingchi, Xi Chen, and Xiaofang Li. "Max–min task scheduling algorithm for load balance in cloud computing." Proceedings of International Conference on Computer Science and Information Technology. Springer India, 2014.

[9] Domanal, Shridhar G., and G. Ram Mohana Reddy. "Load Balancing in Cloud Computingusing Modified Throttled Algorithm." Cloud Computing in Emerging Markets (CCEM), 2013 IEEE International Conference on. IEEE, 2013.

[10] Soni, Gaurav, and Mala Kalra. "A novel approach for load balancing in cloud data center." Advance Computing Conference (IACC), 2014 IEEE International. IEEE, 2014.

[11] Laxmi Shrivastava, G S Tomar, and S S Bhadoria, "A Load-Balancing Approach for Congestion Adaptivity in MANET", IEEE International Conference on Computational Intelligence and Communication Networks (CICN), pp 32-36, Oct 2011. DOI- 10.1109/CICN.2011.7

[12] Domanal, Shridhar G., and G. Ram Mohana Reddy. "Optimal load balancing in cloud computing by efficient utilization of virtual machines." Communication Systems and Networks (COMSNETS), 2014 Sixth International Conference on. IEEE, 2014.

[13] Singh, Aarti, Dimple Juneja, and Manisha Malhotra. "Autonomous Agent Based Load Balancing Algorithm in Cloud Computing." Procedia Computer Science 45 (2015): 832-841.

[14] Pan, Jengshyang, Pingfei Ren, and Linlin Tang. "Research on Heuristic Based Load Balancing Algorithms in Cloud Computing." Intelligent Data Analysis and Applications. Springer International Publishing, 2015. 417-426.

[15] Ramezani, Fahimeh, Jie Lu, and Farookh Khadeer Hussain. "Task-based system load balancing in cloud computing using particle swarm optimization." International Journal of Parallel Programming 42.5 (2014): 739-754.

[16] Banerjee, Sourav, et al. "Development and Analysis of a New Cloudlet Allocation Strategy for QoS Improvement in Cloud." Arabian Journal for Science and Engineering 40.5 (2015): 1409-1425.

[17] L. Shrivastava, SS Bhadoria, GS Tomar, BK Chaurasia, "Effect of number of CBR Connections on the performance of a Load Balanced Congestion Adaptive Routing for MANET", IEEE International Conference (CICN), pp 112-116, 3-5 Nov 2012. DOI: 10.1109/CICN.2012.104

[18] Laxmi Shrivastava, G.S. Tomar, Sarita S. Bhadoria, "A Secure Congestion Adaptive Routing with Load Balancing", International Journal of Communication Systems and Network Technologies, Vol.1, No.3, pp.143-156, 2012. DOI- 10.18486/ijcsnt/1.3.012

[19] Leelipushpam, P. Getzi Jeba, and J. Sharmila. "Live VM migration techniques in cloud environment—a survey." Information & Communication Technologies (ICT), 2013 IEEE Conference on. IEEE, 2013.