ASIAN JOURNAL OF INFORMATION TECHNOLOGY

MEDWELL PUBLICATIONS

Certain Investigation on Variants of Load Balancing in Big Data Centers using Meta-Heuristic Algorithms

[1]S. Sundararajan and [2]C.R. Sakthivel
[1]SNS College of Technology, Coimbatore, India
[2]Department of Computer Science, Sri Ramakrishna Mission Vidyalaya College of Arts and Science, Coimbatore, India

Key words: Load balancing, Job scheduling, cloud computing, optimization

Corresponding Author:
S. Sundararajan
SNS College of Technology, Coimbatore, India

Page No.: 91-94
Volume: 19, Issue, 5, 2020
ISSN: 1682-3915
Asian Journal of Information Technology

Abstract: Allocation of resources for user requests or scheduling jobs for multiple computing resources in the cloud environment is a significant task in the cloud date center because it is associated with cost. The objective of load balancing is to detect overloaded and underloaded nodes and then balance the load among them. Load balancing takes the advantages of scalability and provides measure to check the application's performance. In order to maintain low overhead in the cloud service provider's side, the scheduling and balancing mechanism is to be implemented in decentralized manner. By taking advantages of meta-heuristic algorithms, the load balancing is assumed as objective function and work on for improving results or to find best fitness value for the function. In this study, the same can be investigated thoroughly using different meta-heuristic approaches; its outcomes are analyzed and concluded.

INTRODUCTION

Cloud computing provides on-demand network access of shared pool of resources like server, storage, network and services in a virtualized manner. This shared infrastructure allows us to pay only for what we need, scaling up or down rapidly and let the provider makes upgrades are automatic. There are three major elements in the cloud infrastructure: one is host represents physical computing machine, second is data center where the cloud resources resides and finally distributed servers in different places. The demand for the cloud resources are initiated when customer submits tasks or jobs for execution. The workload of the user job is depends on application software design, technologies used in applications and scaling requirements. So, cloud service providers are in the position to improve their scheduling algorithms over the time period. One of the major concerns for improving cluster performance is task allocation and resources allocation for user jobs. There are number of algorithms are introduced in align with research also focus on various elements of cloud computing for better utilization of resources. The cloud provider with better job scheduling and load mechanism can provide or capable of handling:

- High level enterprise applications with better performance
- High Scalability
- Manage sudden traffic spikes
- QoS aspects in Service Level Agreements (SLA)
- Continuous business services

Literature review: In data center, there are different levels of multiple physical servers which receive client's request on resources and it will be executed by means of Virtual Machines (VM). VM are virtual units having different performances, determined by cloud controller such as CPU process speed, memory capacity for task processing. To ensure a greater QoS in data center, cloud provider has to overcome all the challenges regarding job scheduling and load balancing among servers.

Job scheduling and load balancing is a combinatorial optimization problem used to find optimal resources for executing user requested tasks. Load balancing is the way of balancing the load on the cloud servers which avoids overloading or under-loading servers in the cloud. Resource provisioning and resource allocation are the two major tasks of load balancing and they will be influenced by availability, effective resource utilization and conserving energy. Load balancing techniques are emerging today for big data centers are expected to provide better performance, scalability to control web traffic, ability to manage sudden traffic burst or sudden mishap and switching from one node to another if a node fails. Even though there are many static load balancing algorithms proposed, dynamic load balancing algorithms depend on current state of the cluster which would be better in performance and fault tolerance.

Many meta-heuristic algorithms have been proposed so far, PSO algorithm is one of the appropriate and interesting methods for dynamic task allocation and scheduling. Kennedy and Eberhart[1] presented PSO algorithm which is similar to Genetic Algorithms (GA). It is a cooperative, population-based swarm intelligence algorithm relies on the behavior of the particles. In this algorithm, each particle in the problem is positioned in the search space and has a fitness value and velocity to determine the speed and direction in which its moves. Particles move around in the search space updated its position and velocity to get an optimized solution. After the several moves, an optimum location is identified. Masdari *et al*.[2] identified the following objectives have been meet by scheduling algorithms in the cloud:

- Minimize task execution time and transfer time
- Minimize cost
- Meet user level QoS agreed in Service Level Agreement (SLA)
- Scalability and availability of resources
- Reducing energy consumption
- Effective load balancing

The researcher also suggested method to incorporate security related elements in the task and workflow scheduling solutions. Al-maamari and Omara[3] proposed Dynamic Adaptive Particle Swarm Optimization (DAPSO) algorithm with objective of solving PSO affinity problem in inertia weight where great inertia weight facilitates a global search while a little inertia weight facilitates a local search. This paper also introduced new scheduling task scheduling algorithm based DAPSO and CS algorithms, called MDAPSO algorithm in which DAPSO algorithm is used to improve the inertia weight and CS algorithm is used in the local search where the performance is improved by changing inertia weight and trapping on local search.

To overcome the limitations of PSO algorithm, hybrid versions are introduced by ant colony, Genetic Algorithms, etc. Such hybrid algorithm combines the strengths of all the methods. Allocation of the tasks to the resources according to the QoS (Quality of Service) requirement is the major concern in cloud environment. Kundu *et al*.[4] proposes a dynamic scheduling algorithm, by utilizing the fuzzy logic controller that contains fuzzy rules to produce an interface engine, i.e., decision. The controller takes two inputs, the number of requests received in the host from virtual machines and the time taken to complete the tasks in each VM. Using this scheduling scheme, the VMs fairly assigned to the host according to the time slot of waiting in the queue and their precedence.

Ant Colony Optimization (ACO) algorithm has been proposed by Nilesh and Patel[5] inspired by Ant Systems to reduce the makespan of a given tasks. The solution is achieved by making inspired scheduling and load balancing. Initially, user tasks are submitted to the Cloud and ant check the cloud platform periodically, collects the load information of the nodes and builds the solution. The algorithm suggested to group VMs in three categories: Under loaded VMs, Overloaded VMs and Balanced VMs. It calculates moving probability for each VM from UVM and OVM and chooses suitable VM for each task, then tasks are transferred to chosen destination VM.

A new metaheuristic optimization algorithm is proposed by Almezeini and Hafez[6] called Lion Optimization Algorithm (LOA) as a nature-inspired population-based algorithm for obtaining global optimization over a search space. The algorithm is based on the simulations of the behaviors of lions such as hunting, mating and defense. A lion represents a task scheduling solution which is initialized randomly by mapping cloud tasks (cloudlets) to cloud resources. The objective is to find the best lion that has the best fitness value which is the makespan of the solution which is the maximum completion time for the tasks. Each lion knows its own best solution (schedule of tasks) as well as the global best solution which are updated progressively during optimization. The author preferred to take four metrics makespan, cost, average utilization and degree of imbalance to prove performance of LOA.

HYBRID SOLUTIONS

Zhan and Huo[7] recommended the new approach, Improved Particle Swarm Algorithm which is the combination of simulated Annealing and PSO. He proposed simulated annealing, which adds fast random global searching ability into every iteration of PSO, improves convergence rate and guarantees solving accuracy of original algorithm. Cloud server can use this algorithm for improving resources discovery, resources matching, scheduling production and task execution.

Goyal and Aggarwal[8] introduced hybrid Ant Colony Optimization (ACO) and Particle Swarm optimization (PSO) algorithm for optimizing workflow time (makespan) locally and globally. ACO is seeking the shortest path between anthill and the location of food and the output of ACO will be given as input to PSO for finding the best path. A particle in the solution space represents an individual object which has the ability to move through the defined problem space and represents a candidate solution to the optimization problem.

To ensure a fair distribution of the workload among the available VMs, considering the order of the execution of the workflow tasks to reduce the makespan and the processing cost of the workflow applications in cloud computing environments, Manasrah and Ba Ali[9] proposed Hybrid Genetic Algorithm(GA)-PSO Algorithm. This algorithm selects the VMs to execute the workflow tasks in the minimum time based on the execution speed of the VMs and the size of the workflow tasks. The design of the GA-PSO algorithm tends to allow executing the tasks over the VMs with a balanced load distribution over the fast and slow VMs, without overloading some VMs over the others. This technique reduces the makespan through a fair utilization of the slow VMs instead of overloading the fast VMs and slowing down the overall execution of the task. By applying same weights for both the makespan and the execution cost in the fitness function, appropriate selection of the VM with a balance between cost and time is made.

Soni *et al*.[10] proposed artificial bee colony based load balancing technique on multi-objective requirements by optimizing best fitness value of load, priority and execution error which depicts bee's behavior of finding food. In this optimization technique of user requests on required resources, objective function is formed which minimize at best request serving. Constraints are formed by considering cloud entity Virtual Machine (VM) in terms of number of VMs, processing capability and availability in the cluster. Fitness value is concluded by minimizing differences between requested and served load, priority and execution error of VMs.

Yeboah and Odabi[11] combines both ant colony optimization algorithm and bee life algorithm to improve the effectiveness of load balancing and cloud scheduling. In this proposed method, ant colony optimization algorithm is used to determine the load balancing and Bees Life Algorithm (BLA) is an optimization algorithm used for job scheduling in the cloud environment. When the use submits the job, BLA calculates the fitness for the job using information like bandwidth and execution time of neighboring. Ant colony optimization algorithm is therefore used to determine the best path for resource allocation and shares the workload for all the virtual machines.

CONCLUSION

This extensive review of meta-heuristic algorithms for job scheduling and load balancing reveals that there are number of dimensions to extend the research in terms of scalability, QoS, VM management, green computing and cost. This kind of algorithms can provide better resource management on cloud which can significantly reduce the response time and cost which would be the strong foundation for the existence of business for long period. Therefore the scope for energy-efficient load balancing technique in the cloud computing has greater research value because of number of consumers accessing cloud servers simultaneously are increasing rapidly.

REFERENCES

01. Kennedy, J. and R. Eberhart, 1995. Particle swarm optimization. Proc. IEEE Int. Conf. Neural Networks, 4: 1942-1948.
02. Masdari, M., F. Salehi, M. Jalali and M. Bidaki, 2017. A survey of PSO-based scheduling algorithms in cloud computing. J. Netw. Syst. Manage., 25: 122-158.
03. Al-maamari, A. and F.A. Omara, 2015. Task scheduling using PSO algorithm in cloud computing environments. Intl. J. Grid Distrib. Comput., 8: 245-256.
04. Kundu, A., 2015. A new approach for task scheduling of cloud computing using fuzzy. Intl. J. Innov. Res. Comput. Sci. Technol., 3: 112-116.
05. Nilesh, A.M. and C.A. Patel, 2017. Load balancing in cloud computing using ant colony optimization. Intl. J. Comput. Eng. Technol., 8: 54-59.
06. Almezeini, N. and A. Hafez, 2017. Task scheduling in cloud computing using lion optimization algorithm. Algorithms, 8: 77-83.
07. Zhan, S. and H. Huo, 2012. Improved PSO-based task scheduling algorithm in cloud computing. J. Inf. Comput. Sci., 9: 3821-3829.

08. Goyal, M. and M. Aggarwal, 2017. Optimize workflow scheduling using hybrid Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) algorithm in cloud environment Misha. Intl. J. Adv. Res. Ideas Innov. Technol., 3: 182-189.
09. Manasrah, A.M. and H. Ba Ali, 2018. Workflow scheduling using hybrid GA-PSO algorithm in cloud computing. Wirel. Commun. Mob. Comput., 2018: 1-16.
10. Soni, A., G. Vishwakarma and Y.K. Jain, 2015. A bee colony based multi-objective load balancing technique for cloud computing environment. Intl. J. Comput. Appl., 114: 19-25.
11. Yeboah, T. and O.I. Odabi, 2015. Hybrid bee ant colony algorithm for effective load balancing and job scheduling in cloud computing. West Afr. J. Ind. Acad. Res., 13: 54-59.