International Journal of Engineering Research and Modern Education (IJERME)
Impact Factor: 7.018, ISSN (Online): 2455 - 4200
(www.rdmodernresearch.com) Volume 6, Issue 2, 2021

TASK SCHEDULING IN CLOUD ENVIRONMENT - A SURVEY

D. Thangamani* & Dr. K. Kungumaraj**
* Research scholar of Computer Science, Mother Teresa Women's University, Kodaikanal, Tamilnadu
** Assistant Professor of Computer Science, Arulmigu Palaniandavar Arts College for Women, Palani, Tamilnadu

Cite This Article: D. Thangamani & Dr. K. Kungumaraj, "Task Scheduling in Cloud Environment - A Survey", International Journal of Engineering Research and Modern Education, Volume 6, Issue 2, Page Number 42-45, 2021.

Abstract:

Cloud computing is an emerging technology to enhance user experiences in digitalized world. It delivers on demand computing services such as platform, storage, networking services, infrastructure, software, analysis and intelligence over the internet. Cloud computing provides flexibility, efficiency, speed, security, mobility, collaboration and disaster recovery to huge data of different organization. These benefits are achieved by using virtualization techniques and task scheduling. Task scheduling maps the tasks to virtual machines based on their clients needs and improve the performance metrics like CPU and ram utilization, reducing turnaround tine and task waiting time. This paper surveyed various task scheduling algorithms to improve user experiences in cloud computing.

Key Words: Cloud Computing, Task Scheduling.

1. Introduction:

The era of digital world brings together all industries, institutions, organizations and service providers find new innovations in their fields with new emerging technologies like Big Data, Cloud Computing, Artificial Intelligence, and Machine Learning. Cloud computing is on demand service of various resources without direct management of user .clouds functions distributed over multiple locations, each location is called a datacenter. Cloud computing sharing these data centers to offer on demand service based on pay-as-you–go model. Cloud computing environment serves for millions of users at a time, so it should have the ability to meet all end users request with high-performance and guarantee of quality of services. It needs some task scheduling and efficient resource allocation techniques to accomplish all the user's needs.

Cloud Architecture:

Cloud computing contains two parts in its architecture i.e. Front End and Back End. Front End consists of interfaces and applications that are required to access cloud computing platforms. Back End consists of all resources like data storage, virtual machines, security mechanisms, deployment model used by the service provider.

Virtualization:

Cloud computing virtualization approach reduces the dependence of various machines for various user needs. Virtualization helps the user to access services like storage, platform, and software without investing much infrastructure. Virtual infrastructure providing several machines at same time and also allows to sharing single physical instances of resources or an application to users. Virtualization divides the physical system into multiple virtual machines using software layer called hypervisor. This hypervisor allocates physical resources such as processor, storage and memory to each VM (virtual machines).

Task Scheduling:

Task scheduling means allocating the tasks to best suitable resources for execution. Scheduling done in two phases host level and VM level. In VM level, hypervisor partition the host machines to several virtual machines and each virtual machine do not interfere with other virtual machines and this layer manager the allocation of VM to tasks. In second phase, task schedulers assign the VMs to jobs using different scheduling algorithm. The goal of various optimizing scheduling algorithm is to achieve high performance and quality of service.

Scheduling Strategies:

Various scheduling strategies applied to tasks allocation to achieve performance measures like make span reduce, efficient resources utilization, reducing waiting time and computing power, etc. Scheduling algorithm used different allocation methods for independent and dependent tasks. An independent task does not need any special priority while allocating virtual machines to jobs. Dependant tasks need some priority based allocation approaches such as [1]

- Static or dynamic
- Immediate or batch
- Heuristic or Meta heuristic

Static or Dynamic:

Static algorithms schedules all the jobs to available VMs in similar manner such as first come first service (FCFS), Shortest job First Scheduling, etc. Dynamic Scheduling consider the current state of virtual machines and current need of tasks and then it distributes the jobs to available virtual machines having the capacity to meets task needs.

Immediate or Batch:

Immediate scheduling allots the virtual machine to jobs when jobs are arrive. Batch scheduling grouping the tasks based on matching criteria (job length, execution time) before send to virtual machines.

Heuristic or Meta heuristic

Scheduling done by heuristic approach is offering services for immediate or particular needs[2].Meta heuristic algorithms providing optimal solutions in allocating virtual machines to tasks[3].For example Ant colony Optimization algorithm, Particle swarm optimization and genetic algorithm.

2. Literature Survey:

Young Ju Moon et al [3] provided the novel meta heuristic approach in task scheduling to improve performance. They used a novel ACO based algorithm called slave ants based ant colony optimization (SACO) that mapping the jobs to virtual machines (VMs) in effectively with optimized parameter mapping. The proposed algorithm solves the worldwide optimization problem with slave ants by avoiding long paths whose pheromones are wrongly accumulated by leading ants. This work distributes all the ant and finding the paths and updates its findings, then the assign the queen and slave ants based on who finds best path with reduced make span. It showed the comparison result of ACO, IACO and SACO. The performance gaps between these three algorithms increases when cloud tasks are increased. When the number of cloud tasks is 200, the difference of the make span between ACO and SACO is about 15 s, and that difference between IACO and SACO is about 5 .While cloud tasks is 700, the difference of make span between ACO and SACO is over 50 s, and that between IACO and SACO is over 20s. This performance increase of SACO is because of preprocessing time of slave ants.

Lei Shi et al [4] created a task Duplication and Insertion algorithm based on List Scheduling (DILS) which includes dynamic completion time prediction, task replication, and task insertion. The algorithm dynamically schedules tasks by predicting the finishing time of tasks supported scheduling of previously scheduled tasks, replicates tasks on different machines, reduces TRM, and inserts tasks into idle time to speed up task execution. Experimental results demonstrates the calculation of the average execution time of the tasks in the DAG, the ratio of the average communication time and the average execution time, task execution time range on different processors and it takes the jobs of Gaussian Elimination, Montage Workflow, Molecular Dynamics Code. DLIS algorithm finds the idle processors and inserts the jobs to idle processors to enhance resource utilization and reducing transmission delay.

Gawali et al [5] designed a heuristic model that combines the modified analytic hierarchy process (MAHP), bandwidth aware divisible scheduling (BATS) + BAR optimization, longest expected processing time preemption (LEPT), and divide-and-conquer methods to do job scheduling and resource allocation. This work applied the algorithm to Cyber shake seismogram synthesis and Epigenomics scientific tasks and reduces the turnaround time and response time. This work also considers maximum utilization result for computing resources such as CPU, memory and bandwidth.

Yang et al [6] applied Particle Swarm Optimization (PSO) to solve optimization issues in multiple task scheduling. While scheduling the processing cost and make span of tasks and the resources allocation are taken into account to measure the performance of optimization algorithm.

Nima Jafari Navimipour et al [7] CSA algorithm is predicated on the obligate brood parasitic behavior of some cuckoo species together with the Lévy flight behavior of some birds and fruit flies. The following idealized rules are [8].Each cuckoo lays one egg at a time, and dumps it in a randomly chosen nest. The best nests with top quality eggs (solutions) are going to be administered over to subsequent generations. The number of available host nests are counted and fixed, say n, and the host can discover an alien egg by a probability Pa [0, 1]. The host bird can either throw the egg away or abandon the nest in order to build a completely new nest in a new location. This work maximum reduces the findings of alien task[Pa].Using this CSA algorithm increases the speed and task completion is very high.

Hamed et al [9] applied genetic algorithm to achieve the goal of reduced completion time, cost and resource utilization. This work used initially 4 processors and 10 tasks and finally used 8 processors and task rages are 20 to 70.they computed execution costs are randomly calculated and applied genetic algorithm mutation and crossover techniques to find best solutions like speed up the utilization.

Zhong et al [10] used greedy based particle swarm optimization algorithm to increase the efficient virtual machine utilization compared to traditional particle swarm algorithm. This work allotted all the all the jobs to all virtual machines and then find the best matching virtual machine to tasks based on task needs with good completion time. This work experimented different number of tasks sizes and virtual machines and

compared the result of PSO and Greedy based optical swarm optimization (GPSO) algorithm. This GPSO accomplish the aim of shortest completion time and load balanced virtual time.

Sathish kumar et al [11] proposed artificial bee colony algorithm (ABC) with time as quality factor and find total task completion time, mean finishing time and virtual machines load balancing time. Initially ABC algorithm calculates the status of virtual machines, if the machine is free, tasks allotted to free virtual machines, every time status of virtual machines are updated and based on the job is allotted to processors. Finally this work finds this finds best allocation strategy based on time of task completed time.

Amandeep Kaur et al [12] experimented FCFS (First Come First Serve), Round-Robin, Priority Based Scheduling and SPF (Shortest Path First) in cloudsim simulation tool. This works result showed SJF algorithm suitable to finish the job with least average waiting time and turnaround time, FCFS algorithm having higher waiting and turnaround time, Round robin algorithm requires Medium waiting time, medium turnaround time and High Response time to finish the tasks.

Pardeep Singh et al [13] proposed CPU and memory aware scheduling to achieve the goal of energy efficient task scheduling. This work classified virtual machines into two segments as high CPU availability and high Ram availability, then task are listed as high requirement of cpu and ram tasks into one group and these tasks are allotted to high capacity virtual segments and other tasks are grouped into other group, these task are allotted normal virtual machine segments. This works result showed the approaches to reduce the working machines for energy efficiency.

Mokhtar A et al [14] proposed scheduled cost approach (SCO) scheduling to reduce the cost to the end users. This work considered the budget of user and then the costs of virtual machine's cost based on CPU, RAM, Bandwidth and storage. This SCO approach compared the result with FCFS and SJF traditional algorithms with cost constraints and proves this SCO is best to reduce the cost to users.

Weiwei lin et al [15] used bandwidth aware scheduling to achieve minimum execution time. They designed scheduling process as linear model and used divisible load approach when allocating the tasks to the all the machines at a time. Results of BATS work showed the comparison of virtual machines cost and computing power efficiency with other algorithms. This work tested the BATS algorithm applied for virtual machine from 5, 25 and 200 with its computing power to proved it finds the strategy to reduce the execution time compared to other traditional algorithms.

3. References:

1. Tahini Alawani, "Types of Task Scheduling Algorithms in Cloud Computing Environment", www.intechopen.com/scheduling problem-new application and trends/chapters/71902, april 2020.
2. Poonam Chaudhary, Varsha," A survey on heuristic Based Techniques in cloud computing", International Research Journal of Engineering and Technology, Volume 4: Issue: 08: August 2017
3. Young Ju Moon, Heon Chang Yu, Joon-Min Gil, and Jong Beom Lim,," A slave ants based ant colony optimization algorithm for task scheduling in cloud computing environments ", Humen –centric computing information sciences, (2017).
4. Lei Shi, Jing Xu, Lunfei Wang, Jie Chen, Zhifeng Jin, Tao Ouyang, Juan Xu, and Yuqi Fan," Multijob Associated Task Scheduling for Cloud Computing Based on Task Duplication and Insertion", Emerging Trends and Challenges in Knowledge Defined Networking for 5G,April 2021.
5. Gawali, M.B., Shinde, S.K. Task scheduling and resource allocation in cloud computing using a heuristic approach. J Cloud Comp 7, 4 (2018). https://doi.org/10.1186/s13677-018-0105-8
6. Zhongxue Yang, Xiaolin Qin, Wenrui Li and Yingjie Yang, 2013. Optimized Task Scheduling and Resource Allocation in Cloud Computing Using PSO based Fitness Function. Information Technology Journal, 12: 7090-7095.
7. Nima Jafari Navimipour and Farnaz Sharifi Milani," Task Scheduling in the Cloud Computing Based on the Cuckoo Search Algorithm", International Journal of Modeling and Optimization, Vol. 5, No. 1, February 2015.
8. X.-S. Yang and S. Deb, "Cuckoo search via Levy flights," in Proc. World Congress on Nature & Biologically Inspired Computing, 2009, pp. 210-214.
9. Hamed, A. Y., Alkinani, M. H. (2021). Task Scheduling Optimization in Cloud Computing Based on Genetic Algorithms. CMC-Computers, Materials & Continua, 69(3), 3289–3301.
10. Z. Zhong, K. Chen, X. Zhai and S. Zhou, "Virtual machine-based task scheduling algorithm in a cloud computing environment," in Tsinghua Science and Technology, vol. 21, no. 6, pp. 660-667, Dec. 2016, DOI: 10.1109/TST.2016.7787008.
11. R. Sathish Kumar, S. Gunasekaran," Improved task scheduling in large scale cloud environment using artificial bee colony algorithm", International Journal of computer applications (0975-8887), Volume 103, October 2014.
12. Amandeep Kaur & Dr. Raman Maini, "Different task scheduling algorithms in cloud computing", International Journal of Latest Trends in Engineering and Technology, Vol. (9), Issue (3), pp.217-223 DOI: http://dx.doi.org/10.21172/1.93.37 e-ISSN:2278-621X

13. Pardeep Singh, Jyotsna Sengupta, P. K. Suri ," Cpu And Memory Requirement Based Task Consolidation For Reducing Energy Consumption In Cloud Computing", Journal Of Critical Reviews ISSN- 2394-5125 Vol 7, Issue 09, 2020
14. Mokhtar A. Alworafi, Atyaf Dhari, Asma A. Al-Hashmi, Suresha, A. Basit Darem," Cost-Aware Task Scheduling in Cloud Computing Environment", International Journal of Computer Network and Information Security(IJCNIS), Vol.9, No.5, pp. 52-59, 2017.DOI: 10.5815/ijcnis.2017.05.07
15. Weiwei Lin and Chen Liang and James Zijun Wang and Rajkumar Buyya,"Bandwidth-aware divisible task scheduling for cloud computing", Software: Practice and Experience, 2014, Volume 44.