
Comprehensive Study on Task Scheduling Strategies in Multicloud Environment

Roshni Patil*, Anup Gade, Abhay Rewatkar

Tulsiramji Gaikwad-Patil College of Engineering & Technology, Nagpur, India

Corresponding Author Email: roshni09patil@gmail.com

<https://doi.org/10.18280/jesa.520106>

ABSTRACT

Received: 2 November 2018

Accepted: 19 January 2019

Keywords:

cloud computing, shortest job first scheduling, round robin scheduling, Makespan time, response time, completion time

The task scheduling in multicloud computing environment is complicated task. The cloud computing has increase tremendous popularity both in academia and business because of its on demand service over internet to the various customer. The task scheduling problem is N-P (Non deterministic polynomial) completeness problem. Task scheduling in cloud computing is a best-known problem that has been paid attention. This is again more challenging, specially for multicloud computing environment. This survey paper presents the Shortest Job First algorithm, Round Robin algorithm and Genetic Algorithm for task scheduling in multicloud computing. The round robin algorithm provides the fair allocation and square allocation of resource to the task. It provides the accurate result in finding solutions to large scale optimization problems, by Genetic algorithms such as task scheduling and it is also helpful. A good task and resource scheduling mechanism must satisfy the QoS requirement of the user and at the same time make an efficient utilization of resources. This survey paper presents an algorithm which tries to achieve application high availability, and minimum makespan, minimum response time and completion time.

1. INTRODUCTION

Multiple cloud has a very important factors and the reason for this is to increasing use of cloud in recent area. Illusion of the unlimited resources to the cloud user is important feature of the cloud. The single cloud is cannot meet the needs of user and its is very important to provide the services using multiple cloud and this become a very best solution. Multicloud computing technique provides the resource of various capacities to customer. The cloud computing is popular this day with the need for in demand and large scale computing, public cloud provider such as Amazon Google, Microsoft or Rackspace offers access to their reliable infrastructural or directly to the virtual resource [1]. In a multicloud environment the task or job scheduler adaptable their scheduling strategy to the ever-changing situation and the various type of task.

Now a day cloud computing is inevitable of the day by day operation such as health military, IT industries. The makespan, response time, completion time and the resource utilization are the very important metrics to evaluate a task scheduling algorithm. Makespan is total time required from submission of the task to the final execution. and for the good result of the scheduling in multicloud computing the makespan need to be minimized. The customer satisfaction rate is also important and for this the response time is also need to be minimized. The resource utilization is usage of the virtual machine and for better utilization of available resources, it need to be maximized [2]. There are many existing systems on the task scheduling in multicloud computing which are provide the result in minimizing the makespan time and response time or maximum resource utilization. This is the very important factors and to balancing between them is difficult task.

The main components in cloud computing are users,

resource providers, and task/job scheduling which contains the user request and scheduling strategy. The concept of scheduling means the particular amount of time a resource is assign to the task or job [3]. The various scheduling algorithm are available which is responsible for assign the resources to the requested task in a multicloud environment based on quality of services. Various tasks are distributed to the various resources in order to maximize the services. But it is the challenging process to scheduling on multicloud computing [4].

Software as a Service (SaaS) its a first model in cloud computing which is integrate the cloud provider maintaining and installing the software within the multicloud. Customer operate the software from cloud client via Network. Multi Tenant, Flex Tenancy, Single Instance, Multi Instance this are the four important concepts in SaaS. One of the cloud service is Platform as a Service (PaaS). In PaaS method of cloud computing allow user/customer database services and with application platform. The next cloud computing service is Infrastructure as a Service (IaaS). In the IaaS all the physical device or hardware is Virtualize and done by using IaaS Service.

In this survey paper present the SJF (shortest job first) algorithm, RR (Round Robin algorithm) and GA (Genetic task scheduling algorithm) for the task scheduling in multicloud computing [5]. The multicloud has its own challenges. Various scheduling methods are available and every methods have there own advantages and disadvantages. SJF algorithm select the shortest task for scheduling in the multicloud computing [6]. Where in the Round Robin task scheduling algorithm are uses the time quantum method for scheduling the various task in multicloud computing. Genetic algorithm, work as a search optimization method is one of the subdivisions of evolutionary algorithm inspired by biology transformation such as

inheritance, mutation, and selectivity. The particular time given to each and every single task for there execution. The multicloud computing platform need to fulfill some solution on makespan time, response time and completion time.

In this survey the three important parameters are studied about the scheduling in multicloud computing. Task scheduling in multicloud computing provide the result based on various parameters like makespan time, load balancing, costing, response tome, customer satisfaction rate, performance of the system etc. These parameters decide the proposed scheduling algorithm is good on cloud computing or not. The algorithm like Shortest job first SJF, Genetic task scheduling algorithm and round robin scheduling algorithm are scheduling the resources to furnish the very important parameter [7]. The parameter is ‘Makespan Time’ is very important is the total length of schedule or total length of time that is when the all task have finished processing. The amount of time taken by the system to response to request was

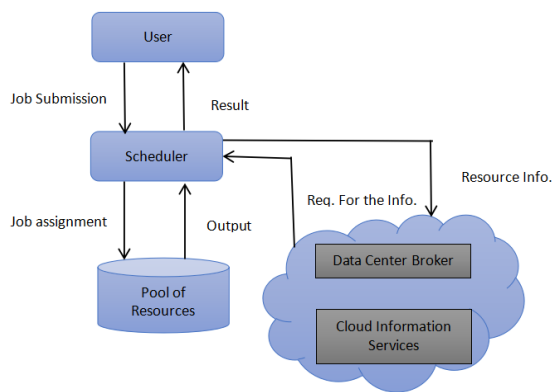


Figure 1. Scheduling model in cloud computing environment

submitted is the ‘Response time’. The term ‘Completion time’ refers to which process complete its execution.

2. FIGURE AND TABLE

The Figure 1. shows the scheduling models in cloud computing. The modules are User, Scheduler, Pools of resources, Data center broker, and Cloud information services. The all modules are intercommunicate with each other like the User are first submitted there job or task. Job or task assignment is done through pools of resources available in multicloud computing.

The module cloud information services have the all important information related to the request of job submission and also information about allocated resource to that job. The output is sent through the scheduler to user in cloud computing environment.

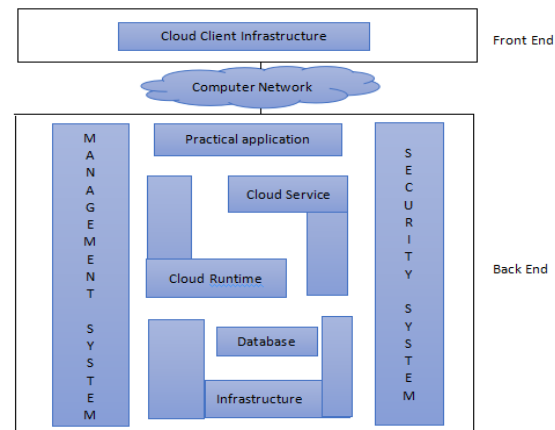


Figure 2. System architecture of cloud computing

Table 1. Literature survey

Sr. No.	Title and Author	Parameters	Observation
[1]	GA-Based Customer-Conscious Resource Allocation and Task Scheduling in Multi-cloud Computing, Tamanna Jena, J. R. Mohanty, 2017	Makespan	This algorithm give the minimum makespan time and minimum cost
[2]	Multi-Phase Proactive Cloud Scheduling Framework Based on High Level Workflow and Resource Characterization Nelson Mimura, Gonzalez-Escola Politecnica 2016.	Makespan	Very good performance gives by Genetic Algorithm m. Minimum time required for completion.
[3]	Optimizing Multi-cloud CDN Deployment and Scheduling Strategies Using Big Data Analysis, Congjie Wang1 Zhihui Lu,2017.	CPU utilization Completion time	Good performance.
[4]	Survey on Energy Efficient Cloud: A Novel Approach towards Green Computing, Anup Gade, Nirupama Bhat 2018.	Execution Time	FCFS and greedy based algorithm compared with the greedy algorithm. By Ga The execution time become less.
[5]	Resource and instance hour minimization for deadline constrained DAG applications using computer clouds, H. Wu, X. Hua, Z. Li, and S. Ren, 2016.	Makespan Cost	Min-Min algorithm and Min -Max algorithm are used for comparing with proposed algorithm. This algorithm gives better result in case of costing.
[6]	Analysis and Improvement of Load Balancing in Cloud Computing, Ronak R Patel, Kaneria Ojasvee and R K Banyal,2016	Response time Constance	This algorithm gives the better response time with Round Robin algorithm
[7]	A Load Balancing Algorithm for Virtual Machine Scheduling in Cloud Computing, Li Liu, Zhe Qiu ,2017.	Response time Processing time	With respect to existing Round Robin the proposed algorithm produce better result.
[8]	An Optimized Task Scheduling Algorithm in Cloud Computing, Shubham Mittal, Avita Kata 2016.	Makespan Time	This algorithm compared with min-min algorithm and min-max algorithm. This algorithm gives the minimum makespan time.
[9]	Improved GA Using Population Reduction Load Bancing in Cloud Computing, Ronak R Patel, Swachil J Patel ,2016.	Makespan Time Response time	HBB-LB algorithm gives the better performance in case of response time and makespan time.
[10]	Investigating a Genetic Algorithm Simulated Annealing Hybrid Applied to University Course Time Tabling Problem, J. Jonasson, and E. Norgren 2016	Makespan time Response time with input one to six data center	FCFS, Round Robin algorithm are compared with Genetic Algorithm and proposed algorithm produce the minimum response time result.

The Figure 2 shows the Cloud Computing System Architecture which have components and sub components which is necessary for cloud computing. This components come with the two main ends one is Front end platform consists the Mobile device, server, laptop, thick or thin client second is back end which consisting the storage back end, cloud based delivery system and network. Between client and the cloud data storage interaction has been done via the application or middleware a web browser or web session.

Both the front end platform and back end platform connected through network via internet, Intercloud, Intranet. Cloud storage is online or internet storage where data is stored and available to the client.

3. OBSERVATION

In the survey study various Task Scheduling Algorithm result are observe on multicloud computing. various parameter like Response Time, Completion Time, makepan time, resource utilization, costing etc. this are the basic parameter are observe during the survey study of scheduling algorithm in multicloud computing. Some of the scheduling algorithm result are discuss following:

(1). Genetic Algorithm: “A Genetic Algorithm (GA) based load balancing strategy for cloud computing”. In this paper result shows that the makespan time are Good. Completion time and Response time are also provide the better result. But the cost of the algorithm is very high and Resource Utilization are also not provide satisfy result than Expectation.

(2). The Ant Colony Based Scheduling Algorithm: “Scheduling Work flow in Cloud Computing Based on Ant Colony Optimization Algorithm”. The result of this paper show that the makespan time very high. Response time are maximum and Completion time is minimum. Costing of the algorithm is very high but the resource utilization is very good. The resource is utilized proper manner in the Ant Colony-Based Scheduling Algorithm.

(3). Improved Round Robin Algorithm: “Analysis and Improvement of Load Balancing in Cloud Computing”. In this paper result shows that the Makespan time are Good. The completion time of this algorithm are high and response time are provide better result. the resource and CPU utilization are very good in the give algorithm. But the costing of the proposed scheduling algorithm is very high.

(4). Shortest Job First scheduling Algorithm: “An Improved Shortest Job First Scheduling Algorithm to Decrease Starvation in Cloud Computing”. The result of this paper show that the makespan time very low. the completion and the Response time are minimum as compare to other scheduling algorithm. And the resource utilization provides the better result.

(5). Parallel Scheduling Genetic Algorithm (GA): “an approach for cloud resource scheduling based on parallel genetic algorithm”. result of the algorithm shows that the completion time of the proposed scheduling algorithm are minimum. but the response time is not good. the costing of the algorithm is very high in case of resource utilization the provide the better result.

4. CHALLENGES IN EXISTING SYSTEM

Rashmi et al. proposed the cost efficient technique for

reducing the response time to execute the task that come form different end user of cloud computing and it is very important parameter. This technique SJFST provide the effective for minimizing the response time. JiaRu and Jacky Keung proposed algorithm with more than one task for maximizes the resource utilization. after comparing the existing system the result show the waiting time and processing time is give better result. here need to be enhance the makespan based on QoS factor such as deadline constraint [4]. Nadeem Javaid, Moomina Waheed proposed the SJF algorithm for load balancing and efficient resource management on multi cloud computing environment. But the response time is high and cost of this algorithm is also very high.

Thomas Yeboah et al. proposed the enhanced version of Round Robin algorithm for but its not give the better result in processing time. Lahar Singh Nishad proposed a Round Robin scheduling algorithm for higher resource utilization and minimizing the cost on multicloud computing[6]. It gives the best result for both parameters. This algorithm is not provide effectual policy. The various existing system on Round Robin scheduling algorithm it takes the long run time in case of heavy workload.

Rui Wang, Zhongni Zheng proposed the Genetic Algorithm which is applicability for parallel of genetic algorithm to find the optimal placement of virtual machine (VM) for maximizing the utilization of resources on multicloud computing[5]. Shanshan Zhang, Chenhong Zhao proposed the Genetic Algorithm for resource utilization and time utilization. But it is not good for solution space in GA. The existing system on the GA scheduling not consider the cost as well as total completion time [1].

4.1 Other challenges

- (1). Proper Resource Utilization
- (2). Minimizing Makespan Time
- (3). Reduce overhead from resource
- (4). Increase system performance
- (5). Response time
- (6). Minimize completion time and
- (7). Better CPU utilization
- (8). Costing

5. PROPOSED METHODOLOGY

With the growth of incoming application the complexity of the resource allocation and scheduling job is increase. In the multicloud environment the number of resources is numerous and the processing power capacity are different. The number of incoming application to be executed are arbitrary. Therefore the very strong and robust algorithm are required for the scheduling of the task. To satisfy the QoS, minimum Makespan time and better allocation of resources are important. The Genetic Algorithm is better suited for scheduling. The proposed algorithm in this survey paper is used to mapping independent task to virtual machine of multicloud such that each task is executed having minimum makespan time, completion time, and response time respectively [3]. One of the primary objectives of the scheduler is the utilization of CPU to facilitate the process of the completion and response time, waiting and turnaround time, priority and system throughput.

Figure 3 illustrate that process of resource allocation where

each user introduce there application tasks and the cloud federation termed uses the appropriate approach to allocating resource to this task by considering optimization parameters like minimum makespan time, minimum response time and minimum completion time and maximum customer satisfaction rate. Because of this the optimization problem can be solve by using heuristic algorithm such as Genetic Algorithm, Shortest job first algorithm and Genetic task scheduling algorithm.

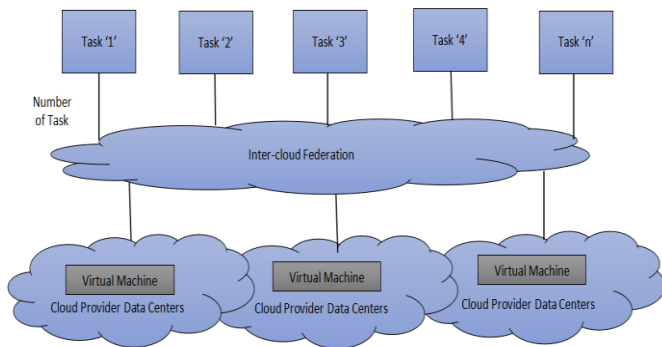


Figure 3. Resource allocation principle

Efficient resource allocation involves mapping of task to the virtual machine followed by task scheduling algorithm like Shortest Job First, Genetic algorithm or Round Robin algorithm. By doing this the balance of workload is achieved and server usage of each resource is maximized.

In the ever-changing environment the different scheduling scheme and the different type of task must be adapt by A good task scheduler. Because of this reason the Genetic task scheduling algorithm is appropriate for cloud. Many researchers used the genetic algorithm to solve the N-P hard problem. Find exact or approximative solution to optimization problem is the Genetic algorithm main aim. A GA is a stochastic technique based in the principle of natural evolution and genetics. Ga is robust in nature and in a multimodel environment the GA has the capacity to locate the global optimum solution

The Job Queue Scheduler class are responsible for deriving the Genetic Scheduler. When the sliding window is full This class have the important feature and ability to schedule job based on genetic algorithm. A new scheduling decision is made in every get Interval second this interval in select as a multiple of heartbeat interval in order to:

- Reduce the overhead of the algorithm
- Give time to some job to finish
- Fill the sliding window in mean time

The Round Robin scheduling algorithm defines a loop as a queue and also states the constant time quantum. Each and Every task can be performed with the given time quantum and in turn. If any task which is not completed in a given time quantum it will return to the queue and wait till its next turn. This round robin algorithm are focus on scheduling processes fairly. The very good advantage of this scheduling algorithm is each and every task is executed in its own turn. It is not necessary to wait for the previous task to be completed. Because of these there is no hunger for other task in the algorithm.

However, if the queue is full or the workload is very heavy, a lot of work is needed to completing the all work. And the selecting the appropriate time quantum is difficult for scheduling in the algorithm.

Another proposed algorithm in this survey paper is shortest job first scheduling algorithm. The shortest job first scheduling algorithm is one of the algorithms which have the minimum execution time, its process having the smallest CPU burst gets a higher priority and execute at the first by CPU [1]. The remaining job or task in the pool are wait for the execution during this time. Then, the task that has the highest execution time will be executed at the end. Shortest job first scheduling algorithm follow the criteria that it processes the smallest task first and largest task later.

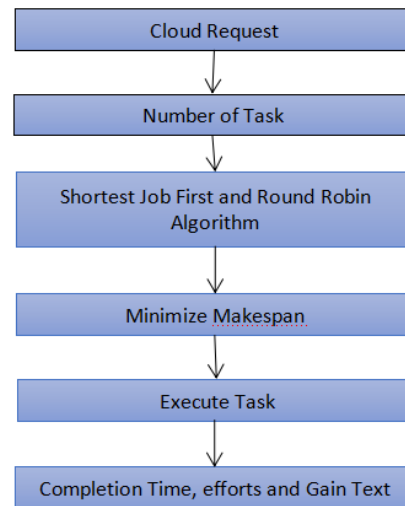


Figure 4. SJF Execution in Cloud

In the proposed algorithm of SJF if the task has the highest execution time the broker will send the task to the high configurations virtual machine to execute this task and if the task has the lowest execution time the broker will send it to the lower configuration virtual machine to be execute. The main idea to modify the SJF scheduling algorithm is to sort the task in ascending order based on the task length and calculate average of all the task length. Then the algorithm checks all the task length if it is less than the average of task length and number of tasks in virtual machine (VM) 1 less than the number of tasks in virtual machine 2, then the task will be sent to VM1 else to VM2.

6. CONCLUSIONS

In this survey paper focus the various Scheduling algorithm like shortest job first Scheduling Algorithm, Round Robin scheduling algorithm and Genetic Algorithm to achieve the minimum makespan time, minimum response time and completion time. Also observe the cost of the various scheduling algorithm and there CPU and resource utilization on multicloud computing. In cloud computing environment heterogeneous resources are available as services by creating VM, these resources manage in optimized way with scheduling. Scheduling is very necessary in cloud environment for efficient resource utilization. The experimental study show that the existing proposed algorithm produce better result but need to enhance the Quality for the same.

ACKNOWLEDGMENT

I would like to express my special thanks and gratitude to my teacher as well as my guide “Prof. Anup Gade” who gave me the golden opportunity to do this wonderful project on the topic ‘Comprehensive Study on Task Scheduling Strategies in Multicloud Environment’ which also helped me in doing a lot of Research and I came to know about so many new thing I am really thankful to him.

REFERENCES

- [1] Jena T, Mohanty JR. (2017). GA-based customer-conscious resource allocation and task scheduling in multi-cloud computing. *Arabian Journal for Science and Engineering* 43(8): 4115-4130. <https://doi.org/10.1007/s13369-017-2766-x>
- [2] Gonzalez NM, Carvalho TCMB, Miers CC. (2016). Multi-phase proactive cloud scheduling framework based on high level workflow and resource characterization. Brazil, IEEE International Conference 1: 43-47. <https://doi.org/10.1109/NCA.2016.7778591>
- [3] Wang CJ, Lu ZH, Wu ZY, Wu J, Huang SL. (2017). Optimizing multi-cloud CDN deployment and scheduling strategies using big data analysis. IEEE International Conference, Honolulu, HI, USA. <https://doi.org/10.1109/SCC.2017.42>
- [4] Gade A, Bhat N, Thakare N. (2018). Survey on energy efficient cloud: A novel approach towards green computing. *Helix ISSN*. <https://doi.org/10.29042/2018-3976-3979>
- [5] Wu H, Hua X, Li Z, Ren S. (2016). Resource and instance hour minimization for deadline constrained DAG applications using computer clouds. *IEEE TPDS* 27(3): 885–899. <https://doi.org/10.1109/TPDS.2015.2411257>
- [6] Ojasvee K, Banyal RK. (2016). Analysis and improvement of load balancing in cloud computing. *International Conference on ICT in Business Industry & Government (ICTBIG)*. <https://doi.org/10.1109/ICTBIG.2016.7892711>
- [7] Liu L, Qiu Z. (2017). A load balancing algorithm for virtual machine scheduling in cloud computing. *IEEE International Conference*. <https://doi.org/10.1109/ICMIC.2017.8321690>
- [8] Mittal S, Avita Katal A. (2016). An optimized task scheduling algorithm in cloud computing. *IEEE 6th International Conference on Advanced Computing*. <https://doi.org/10.1109/IACC.2016.45>
- [9] Patel RR, Patel SJ, Patel DS, Patel TT. (2016). Improved GA using population reduction load balancing in cloud computing. *Intl. Conference on Advances in Computing Communications and Informatics (ICACCI)*. <https://doi.org/10.1109/ICACCI.2016.7732410>
- [10] Jonasson J, Norgren E. (2016). Investigating a genetic algorithm simulated annealing hybrid applied to university course time tabling problem. *KTH Royal Institute of Technology School of Computer Science and Communication, Degree Project Technology: Stockholm Sweden*. <https://doi.org/10.1109/CLOUD.2016.00047>