

A Survey of Different Task Scheduling Algorithm in Cloud Computing

Mayank Sohani^{#1}, Saniya Jain^{#2}, Ishita Narang^{#3}, Kaushal Agarwal^{#4}, Saahil Arora^{#5}

^{#1} Assistant Professor, Department of Computer Science ^{#2-5} Department of Information Technology

Narsee Monjee Institute of Management Studies

Mukesh Patel Technology Park, Babulde, Bank of Tapi River, Maharashtra, India

Abstract— Cloud computing is one of the latest technological buzzword doing round in business worlds these days. Thus to have the knowledge of cloud computing is an obligation of late. Cloud gives suitable and on desired network access for computing resources that are available on the Internet. Tasks sent to cloud environment needs to be operated on time to get proper utilisation of resources and to get maximum efficiency and minimum makespan. The objective of this paper is to compare different task scheduling algorithms so that we can have proper task allocation as per our requirement.

Keywords— Cloud Computing, makespan, Scheduling algorithms, Task scheduling

I. INTRODUCTION

Cloud computing is that class of computing which relies on sharing of computing resources rather than holding local servers to handle applications. Everything that relates to cloud computing works in a virtual environment. In layman's language cloud computing refers to accessing the pool of resources for computing through browser's window. Cloud computing is not free, rather a pay-per-use service.

A. Cloud Architecture

Cloud Architecture is designed with many components. The components are loosely coupled. These components consist of 2 parts- The Front end and the back end. Both the ends are connected via Internet.

Front-end is the one visible to the client who is the user of services. It involves using user's computer system that is used to get the cloud system. Different Cloud Computing systems use different user interface. For example, for email programs,

interfaces include web browsers like Firefox or Internet explorer, etc.

Back-end is the one used by the service provider. It includes servers, computers, data storage, and VMs that all together establishes the cloud of computing services. This system can involve various computer programs. Individual system application is managed by its own server. The back-end side has some responsibilities towards the user. Some of these responsibilities include giving security mechanisms, traffic control, assigning protocol that links networked computers for communication.

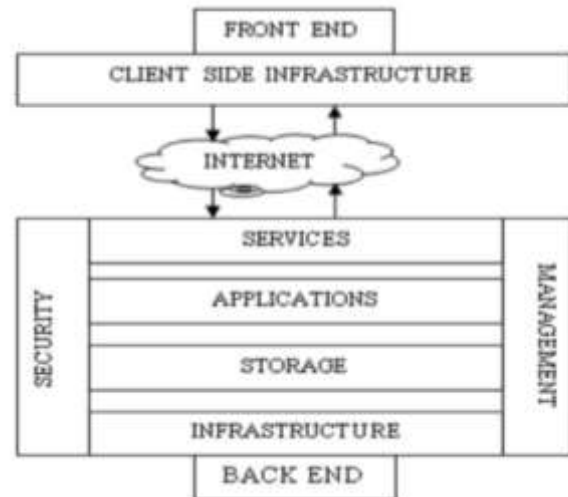


Fig.1: Architecture of Cloud [13]

B. Scheduling

Scheduling refers to a method by which work is assigned to a resource and the resource completes the assigned work. The main purpose of scheduling

is to maximize the resource utilization and to reduce the makespan [1]. A good scheduling gives a better performance and increases the system throughput. In cloud computing datacenter is a very important part. There are various datacenters and each datacenter has around thousands of virtual servers running at a given point of time which executes many [2]. While this execution takes place there are many new jobs waiting to be executed. This makes task scheduling an important concern in cloud computing.

C. Scheduling types

There are two different types of scheduling known as static scheduling and dynamic scheduling.

1) *Static Scheduling*: In this type of scheduling, the scheduler has a prior knowledge of all the process or tasks/jobs which are going to be executed before the execution takes place.

2) *Dynamic Scheduling*: In this type of scheduling, the scheduler has no prior knowledge any process or tasks/jobs which is going to be executed.

II. SURVEY

There are various types of scheduling algorithms available for cloud computing. In this paper, we have only focused on static scheduling. There are various parameters which we have taken into consideration and they are as follows: makespan, energy consumption, power consumption, time complexity, waiting time. The algorithms taken for comparison are:

Improved Cost Based Algorithm [3]: This algorithm is the improved version of the traditional cost based task scheduling algorithm for appropriate resource utilization and allocation. In this algorithm, the tasks are grouped in accordance to the processing power of the resources. Also in this algorithm efficient mapping of tasks are done to the resources which are available in cloud.

A Priority based Job Scheduling Algorithm [4]: In this algorithm, mathematical calculations are used for presenting job scheduling. Every task/job has some priority associated with it and in accordance to the priority it requests the resources. Also, in this paper problems related to algorithm such as complexity and makespan have been discussed.

According to the author, by improving the algorithm can reduce the makespan.

Queue Based Job Scheduling Algorithm [5]: The main objective of the algorithm is to increase the efficiency of job/task execution. Many algorithms like The First Come First Served (FCFS), Round Robin (RR) and Shortest Job First (SJF) are compared with this algorithm. The metrics on which these are compared are waiting time (average) and response time (average).

Generalised Priority Based Algorithm [6]: The main objective of this algorithm is to reduce the execution time of the jobs/tasks. This algorithms take tasks/jobs and Virtual Machines (VM) into consideration. The virtual machines are given priority in accordance to Million Instructions per Second (MIPS) and the tasks are given priority in accordance to length or the size of the job/task. The highest priority is given to the virtual machine having the highest MIPS value and the task which has the largest length. So the task having the highest priority will be executed on the virtual machine which has the highest priority. This algorithm is also compared with FCFS and RR and its performance was found better compare to those.

Earliest Feasible Deadline First [7]: The main objective of the following algorithm is effortless to understand. The algorithm is as follows, the task which has the smallest deadline scheduled gets scheduled or executed. This algorithm (EDF) comes under the dynamic scheduling category. Whenever there is a process or task/job which has to be scheduled, then the process which is closest to its deadline is searched and the process which is searched will be one which will be executed next.

Greedy Based Job Scheduling Algorithm [8]: The main focus of this algorithm is on the Quality of Service (QoS), cloud computing being a business oriented service. The basic idea behind this algorithm is to minimize the total time for completion and to enhance the scheduling problem. The following algorithm assigns tasks according to the Quality of Service (QoS) and then as per the category of the task/job, the function which is most suitable is assigned. The results which were obtained from the algorithm were made to compare with the algorithm which already exists such as algorithm based on Berger model.

Improved Priority based Job Scheduling Algorithm using Iterative Method [9]: In job/task execution, priority has to be given the utmost importance as the processes which have a higher priority should not wait more than those which have a lesser priority. In this algorithm the authors have proposed an improved priority based algorithm in the cloud computing

environment. The proposed algorithm has improved makespan and better consistency compared to a prioritized RR algorithm. An Analytical Hierarchy Process is taken into consideration. The algorithm proposed has a greater future scope and can be used in the future in order to reduce the makespan.

Priority Based Earliest Deadline First Scheduling Algorithm [10]: In the following algorithm, two different algorithms are combined and used hand in hand. The first one being the Earliest Deadline First algorithm and the second one is a priority based algorithm. Resource allocation is the main focus of this algorithm along with memory utilization. The algorithm is proposed such that, that there will be a reduction in execution time of the pre-empted tasks/jobs and they further can be scheduled efficiently. The waiting time problem is also solved by this algorithm of the pre-empted tasks/jobs. The solution is given by introducing a waiting queue which executes the pre-empted tasks/jobs.

Min-Min Algorithm [11]: The algorithm starts with a group of unmapped [11][12] tasks/jobs and its working principle is to find the job/task which has the minimum time

compared to all the tasks in the meta-task. The task/job which has the minimum completion time is taken and given the corresponding resource [11]. The step is repeated until the meta-task is not empty. The process having a larger completion time has to wait longer.

Max-Min [11] [12]: The algorithm is one of the most commonly used algorithm and is similar to Min-Min algorithm. This is used in the distributed environment. The algorithm starts with a group of unmapped tasks/jobs and its working principle is to find the job/task which has the maximum time compared to all the tasks in the meta-task. The task/job which has the maximum completion time is taken and given the corresponding resource. The step is repeated until the meta-task is not empty. The process having a smallest completion time has to wait longer.

RASA [14]: This algorithm alternatively takes in account Min-Min and Max-Min strategy to give task/job to the available resources. This algorithm adopts the advantage of both the algorithms and because of this the delay in execution of large and small tasks is synchronised and supported.

TABLE I. PERFORMANCE METRICS

Algorithm	Reduces Makespan	Reduces Time Complexity	Improves Resource Allocation	Improves Quality of Service (QoS)	Reduces Average Waiting Time	Reduces Average Response Time
Improved Cost Based Algorithm [3]	Yes	No	No	No	No	No
A Priority based Job Scheduling Algorithm [4]	Yes	Yes	No	No	No	No
Queue Based Job Scheduling Algorithm [5]	No	No	No	No	Yes	Yes
Generalised Priority Based Algorithm [6]	Yes	No	No	No	No	No
Earliest Feasible Deadline First [7]	No	No	No	No	No	No

Greedy Based Job Scheduling Algorithm [8]	No	No	No	Yes	No	No
Improved Priority based Job Scheduling Algorithm using Iterative Method [9]	Yes	Yes	No	No	No	No
Priority Based Earliest Deadline First Scheduling Algorithm [10]	No	No	Yes	No	Yes	No
Min-Min algorithm [11]	Yes	No	Yes	No	Yes	Yes
Max-Min Algorithm [12]	No	No	No	No	No	No
RASA[14]	Yes(In comparison to Max-Min)	Yes (In comparison to Max-Min)	Yes(In comparison to Max-Min)	No	Yes	Yes

TABLE III. SUMMARIZATION OF ALGORITHMS

Algorithm	Objective	Scheduling Category	Scheduling Parameters	Future Work
Improved Cost Based Algorithm [3]	Reduce processing cost and completion time	Task Scheduling based on cost	Time, cost	Improvement can be done to work in dynamic environment
A Priority based Job Scheduling Algorithm [4]	Reduce makespan	Priority based scheduling, Decision making model	Priority	Enhancement to achieve less finish time
Queue Based Job Scheduling Algorithm [5]	Reduce average waiting time and average response time	Queue based job Scheduling	Priority, Time Quantum	None

Generalised Priority Based Algorithm [6]	Reduce scheduling time	Priority based scheduling	Priority	There exists a time gap between simulation in cloud environment and grid environment and future enhancement can be done
Earliest Feasible Deadline First [7]	Reduce Time Complexity	Deadline based Scheduling	Deadline	Improvement can be done to improve in real time system
Greedy Based Job Scheduling Algorithm [8]	Improve Qos, Reduce Makespan	Greedy algorithm based job scheduling	Time and Bandwidth Preference	The algorithm can be improved to reduce the total completion time
Improved Priority based Job Scheduling Algorithm using Iterative Method [9]	Reduce Makespan	Priority based scheduling	Priority of tasks/jobs	Further optimization can be done to reduce the makespan
Priority Based Earliest Deadline First Scheduling Algorithm [10]	Reduce average waiting time	Priority based scheduling	Priority and deadline of task	Further enhancement can be done to reduce waiting time
Min-Min algorithm [11]	Improve QoS	Based on time taken for task to complete and resources available for execution	Length of task	Further enhancement can be done to reduce waiting time and response time and also improvement can be done for load balancing
Max-Min Algorithm [12]	Improve QoS, reduce waiting time and response time	Based on time taken for task to complete and resources available for execution	Length of task	Further enhancement can be done to reduce waiting time and response time and also improvement can be done for load balancing
RASA[14]	Reduce average waiting time and response time, Improve QoS	Based on time taken for task to complete and resources available for execution	Length of task	None

III. CONCLUSIONS AND FUTURE ENHANCEMENT

In this paper we discussed various types of scheduling algorithms. As cloud computing is a boom to the IT world, task scheduling becomes a very challenging area. The challenges related to

task scheduling are never ending as task scheduling can be called as a heuristic problem. The algorithms are compared on metrics which play a very important role for any scheduling activity. We also found that the algorithms which use heuristic technique are more appropriate for scheduling. During the survey, we have analysed a lot of algorithms and tabulated them based on various

parameters and more research can be done in this field. In future, we plan to use these algorithms and propose a new algorithm which reduces the makespan and also maximizes the utilization of resources.

REFERENCES

- [1] Shubham Mittal, Avita Katal, "An Optimized Task Scheduling Algorithm in Cloud Computing" , 2016 IEEE 6th International Conference on Advanced Computing
- [2] Er. Shimpy, Mr. Jagandeep Sidhu, "DIFFERENT SCHEDULING ALGORITHMS IN DIFFERENT CLOUD ENVIRONMENT", International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 9, September 2014.
- [3] Selvarani, S., and G. Sudha Sadhasivam. "Improved Cost-Based Algorithm For Task Scheduling In Cloud Computing." Computational Intelligence And Computing Research (ICCIC), 2010 IEEE International Conference on. IEEE, 2010.
- [4] Ghanbari, Shamsollah, and Mohamed Othman. "A priority based job scheduling algorithm in cloud computing." Procedia Engineering 50 (2012): 778-785.
- [5] Behzad, Shahram, Reza Fotohi, and Mehdi Effatparvar. "Queue based Job Scheduling algorithm for Cloud computing." International Research Journal of Applied and Basic Sciences ISSN (2013): 3785-3790.
- [6] Agarwal, Dr, and Saloni Jain. "Efficient optimal algorithm of task scheduling in cloud computing environment." arXiv preprint arXiv: 1404.2076 (2014).
- [7] Jagbeer Singh, Bichitrananda Patra, Satyendra Prasad Singh, " An Algorithm to Reduce the Time Complexity of Earliest Deadline First Scheduling Algorithm in Real-Time System" (IJACSA) International Journal of Advanced Computer Science and Applications, February 2011.
- [8] Li, Ji, Longhua Feng, and Shenglong Fang. "An greedy-based job scheduling algorithm in cloud computing." Journal of Software 9.4 (2014): 921-925.
- [9] Patel, Swati J., and Upendra R. Bhoi. "Improved Priority Based Job Scheduling Algorithm in Cloud Computing Using Iterative Method." Advances in Computing and Communications (ICACC), 2014 Fourth International Conference on. IEEE, 2014.
- [10] Gupta, Gaurav, et al. "A simulation of priority based earliest deadline first scheduling for cloud computing system." Networks & Soft Computing (ICNSC), 2014 First International Conference on. IEEE, 2014.
- [11] Bhavisha Kanani and Bhumi Maniyar, "Review on Max-Min Task Scheduling Algorithm in Cloud Computing," Journal of Emerging Technologies and Innovative Research, Volume 2, Issue 3, March 2015.
- [12] Kamalam .G.K and Murali Bhaskaran .V, "A New Heuristic Approach: Min-mean Algorithm For Scheduling Meta-Tasks On Heterogeneous Computing System," International Journal of Computer Science and Network Security, VOL.10 No.1, January 2010
- [13] www.tutorialspoint.com, "Cloud Computing Architecture", 2015. [Online]. Available: http://www.tutorialspoint.com/cloud_computing/cloud_computing_architecture.htm.
- [14] Saeed Parsa, Reza Entezari-Maleki, "RASA: A New Grid Task Scheduling Algorithm," International Journal of Digital Content Technology and its Applications, vol. 3 (4), December 2009, pp. 91-99.