

SURVEY ON BATCH MODE HEURISTIC ALGORITHMS ON TASK SCHEDULING IN CLOUD COMPUTING

M.Vishnu Vardhana Rao ¹ , I.Navakanth ²

¹ Asst.Professor, CSE Department, Vignan's Institute of Management and
Technology for Women, Hyderabad, Telangana.
sujanphd2014@gmail.com

² Asst.Professor, CSE Department, Vasavi College of Engineering, Hyderabad,
Telangana.

i.navakanth@staff.vce.ac.in

Abstract

As per demand and requirements of users, virtual servers provide dynamic and scalable resources for pay per model. A collection of virtual machines, an emerging technology in distributed computing is “cloud computing”. Due to day by day development on cloud, faces many challenges, among them scheduling is one. A scheduling follow the set of rules for delivering the high performance and focus on the work load for maximum perform of the node. In this research paper, we focused on comparison between Batch Mode Heuristic Algorithms (BMHA).

Index words—Cloud Computing, Scheduling, BMHA, FIFO, SJF, RR,GA.

Introduction

Cloud computing provides a pool of resources for accessing the end users (pay per use) via internet (Open information retrieval system) and virtual data storage also . Due to the flexibility and virtualization, the count of tasks in a cloud computing are rapidly increased. The handling and managing of these tasks are very difficulty. For that we need some set of policies for proper usage of resources. These policies are called scheduling. Scheduling defines the policies for managing the idealness of CPU. These policies are creates a situation to keep CPU busy with work (for busyness of CPU utilization).

Best policies give maximum utilization of Resources. Scheduling describe the process of selecting the process (is in waiting) from the ready queue send to the CPU for doing the work. The dispatcher allocate the CPU to the selected process. Scheduling put the rules for maximum utilization of resources, throughput of the resources. Scheduling is the processes of getting minimum waiting time of the resources execution. That is scheduling algorithms are used for Getting maximum performance and maximum throughput.

Organization:

This paper consist of following sections, in the first section focused on layers architecture and 2nd section tells the scheduling process and it's parameters. 3rd section gives the information about different types of existing algorithms and finally conclude .

I. layers architecture

Data Centers contains hardware and system software that provides the infrastrute, platform and software as a services via internet (heterogeneity)[1].According to buyya et.al[2] cloud is a Distributed System consisting of a interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing. Cloud Computing consists of 3 services namely platform as a service (PaaS), Infrastructure as service (IaaS) and software as service (SaaS)[3][17]

paas: provides a integrated environment to design, build, test, deploy, update customer application.Iaas: provided to the users to use processing power, storage, network and other computing resources to run any software (os and it's applications) saas: provided to the users to use the Business applications, web services, multimedia.[4]

TABLE COMPARISION: Difference between layered model of cloud computing According To Standard Parameters (resources and working)[4]

Layer name	Resource management at each layer	Working of the layer	Implemented	Position of layer in architecture	Examples
Hardware layer	CPU , Memory, Disk, Bandwidth	Managing the physical resources of cloud	Data centers (IaaS)	Low layer	Data centers
Infrastructure layer	Computation(VM) Storage(block)	Creates a pool of storage & computing resources by portioning the physical resource	VM (IaaS)	Between Top and Middle layer	Amazon Ec2,GoGrid Flexi scale
Platform layer	Software frame work(java/python/.net) Storage(DB/File)	Minimize the burden of deploying applications directly into VM containers	(paas) Cloud servers	Middle layer	Microsoft Azure, Google App Engine, Amazon Simple DB/S3

Application layer/user layer	Business applications, web services, multimedia	Consists of the actual cloud applications	saas	Top layer	Google Apps, Face book, Salesforce.com
------------------------------	---	---	------	-----------	--

II.SCHFEDULING IN CLOUD

Scheduling gives the solution ,for the Question of how the resources are allocated to the end users. Resources to be allocated the end-users, this is done depending on the time and priority (parameter). Scheduling mechanisms very much necessary to improve load balancing between pools of resources. Scheduling analysis the connection between the user's task to the choice of resources and it's execution. Jobs and set of jobs can be scheduled based on policies. Different algorithms have different characteristics and a good scheduling algorithm must possess following characteristics .

Minimum context switches

Maximum CPU Utilization

Maximum turnaround time

Minimum waiting time [5]

In cloud computing, job scheduling algorithms can be classified into two ways

- Batch mode heuristic scheduling (BMHS): jobs are queued and collected into a set when they arrive in system. The scheduling algorithm will start after a fixed period of time.

EX: FCFS, RR, MIN-MIN, MAX-MIN, SJR, PRIORITY...

- Meta heuristic scheduling: Jobs are scheduled when they arrive in the system

EX: ACO, GA, SIMULATED ANNEALING, PSO, TABU SEARCH.....

II.I Scheduling process (stages of scheduling)

In cloud computing, scheduling having the following 3 stages for allocate the various computing resources to applications.

- Find the Resources using filtering (the Processes of retrieve the demand, needed Resources used for end users and assign status information related to them. This is called Resource discovering and filtering.
- Choose target resources depending on the parameters of resources(From stage1)i.e. resources selection (Deciding stages)

- **Task submission:** Resources are allocated to end users or applications. i.e. The task is submitted to the resources[6].

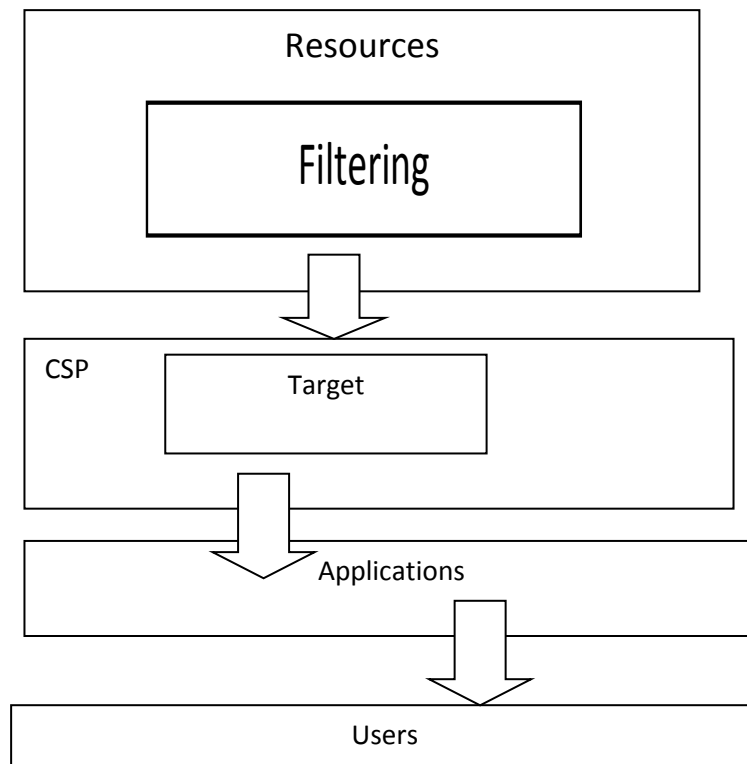


Fig 2: Stages of scheduling [7]

COMPARASION PARAMETERS: Through which the increases the performance the system by observing the parameter status improved by improving those parameters. [8]

Parameters	Definition	High/low
Performance	The speed at which a computer operates	high
Execution time	Time taken for output	low
Completion time	Time to complete the whole execution of a job	low
Response time	the elapsed time between an inquiry and a response	low
Make span	Total completion time of all jobs/tasks in job queue	low
Through put	Number of process that are completed per time unit	high
Resource utilization	Use of resources	high
Load balancing	Method of distributing the entire load in a cloud network across different nodes, remain under loaded	high

Energy consuming	How much energy are needed	low
Quality of service	This include various user input constraints like meeting executions, performance, deadline cost, make span	high
Deadline	Period of time from submitting a task to the time by which it must be completed	low

III.EXISTING SCHEDULING ALGOTITHM:[9]

Scheduler executes the task for completing the task. This work is done based on the arrival of resources from the pool of resources and remaining recourses are in queue that means first arrival resources are allocated to the scheduler first and soon called FIRST COME FIRST SERVE(FCFS)[10]. On the negative side of this process is if the first arrival resources takes the highest execution time then the next arrival resources having smallest execution time in the pool of resources, this gives the convoy effect[16]. If all are having the same execution time then FCFS act as first come first completion method .This scheduling are not suitable for time sharing system and is suitable only for batch process.

So this negative side is converted to Positive side, for that assign the smallest execution time of the resources in set of resources(Shortest job first(SJF))[12], but problem is finding smallest execution time in pool of resources. That is processes with smallest burst time first is the good idea but knowing the smallest burst time is very default. The condition for SJF act as FCFS is the two processes having same burst time [13].

Now gives the a small unit of time (quantum or slices) to all pool of resources for execution after complete the quantum time then scheduler changes the behaviour(non-preemptive,circular queue)[14].that is Ring as its queue to store jobs. Apply to the load balancing in distributing the tasks. If quantum is very large this act as first arrival first execution method. Disadvantage is no priority given to the resources for important processes in pool of recourses.

Now the selection procedure is based on the priority of the processes [18]. Now scheduler executes the highest priority processes first. The burst time of highest priority processes is very high with compared to the other processes (next highest priority with less burst time) is wait until unless the highest priority processes is completed. Starvation of the process is the problem and solution is aging [15].

Priority Scheduling Algorithm [18] - In this scheduling, the selection procedure is based on the priority of the processes. Now CPU executes the highest priority processes first. The burst time of highest priority processes is very high with compared to the other processes (next highest priority with less burst time) is wait until unless the highest priority processes is completed .The problem is starvation(indefinite states).Solution is aging.

Genetic Algorithm (GA) [19] – Genetic algorithm is a problem solving method that uses genetics as its model of problem solving. It is a search system to find optimized solution. Genetic algorithm is a method of scheduling in which the tasks are assigned resources according to individual solutions, which tells about which resource is to be assigned to which task. GA is based on the biological concept of population generation. In Genetic Algorithm the initial population is generated randomly. Genetic algorithm is a random searching method.

TABLE COMPARISION: Difference between Various Algorithms According To Parameters.

Scheduling algorithms	Resources allocation	Implemented / improved	Problem	Possible solution	Waiting time	Environment	Parameter	Condition for algorithm act as FCFS
FCFS (non preemptive) Algorithm	Queued first job	Ready queue	Convoy effect	Priority	long	Cloud Computing	Simplest algorithms	-----
SJF (either preemptive nor non) Algorithm	CPU is allocated to the process that has the smallest next CPU burst time	Ready queue	Finding length of process	Using preemptive type	Good compared with FCFS	Cloud Computing	Difficult to understand	All are having same burst time
Priority (Optimal) Algorithm	CPU Allocated to the processes with the highest priority	Ready queue	Indefinite blocking (starvation)	aging	low	Cloud Computing	Difficult to understand	Processes with the same priority are executed on first come first served basis
RR(Preemptive) Algorithm	The preemption take place after a fixed interval of time	Ring as its queue to store jobs(circular)	Selection of time quantum	--	More than all	Cloud Computing (time-sharing system)	Size of time quantum	If time quantum and process execution time is equal
Genetic Algorithm	Pick the best job to allocate the CPU	Ready queue	Finding best job	Priority	Low	Cloud Computing	Task to be scheduled	-----

Conclusion and Future Work

This survey paper gives an idea about the Batch Mode Heuristic Algorithms and its parameters. For the easy understanding the scheduling and layer architects we provide the comparison tables. In the Environment of cloud computing ,scheduling is an import issue. In this paper, we describe the various parameters ,comparisons in the form of scheduling algorithms tables. But not concentrate on the parameters like Reliability, Scalability &availability so take that it as advantage and implement the scheduling for these parameters.

REFERENCES

- [1] M.Armburst et al., “Above the Clouds: A Berkeley View of Cloud Computing”, *Tech. report*, Univ. of California, Berkeley, 2009.
- [2] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, “Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Gener”, *Computer Systems*,25(6), pp. 599-616, 2009.
- [3] B. Grobauer, T. Walloschek, and E. Stöcker, “Understanding Cloud Computing Vulnerabilities,” 2011 IEEE Security and Privacy, pp. 50-57, DOI= March/April 2011.
- [4] Qi Zhang, Lu Cheng , Raouf Boutaba, " Cloud computing: state-of-the-art and research challenges", Springer, 20 April 2010
- [5] sandeep Sharma,sunpreet singh virk,” Modified GA approach for cloud job Scheduling”, Communication and Computing Systems, Edited by B.M.K. Prasad Krishna Kant Singh Neelam Ruhil Karan Singh Richard O’Kennedy Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742 CRC Press 2016
- [6] Boonyarith Saovapakhiran, George Michailidis,Michael Devetsikiotis” Aggregated-DAG Scheduling for Job Flow Maximization in Heterogeneous Cloud Computing”2011 IEEE 978-1-4244-9268-8/11
- [7] Vinay Darji, Jayna Shah, Rutvik Mehta ,” Survey paper on various load balancing algorithms in cloud computing”, International Journal of Scientific & Engineering Research, Volume 5, Issue 5, May-2014 583 ISSN 2229-5518
- [8] Ruchita Patel¹, Moni Chandel²,” Analysis of Various Task Scheduling Algorithms in Cloud Computing”, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 03 Issue: 03 | Mar-2016 www.irjet.net p-ISSN: 2395-0072
- [9] Shameer A.P, Dr. A C Subhajini,” Study on Different Scheduling Algorithm for Cloud Computing”, International Journal of Advanced Research in Computer Science and Software Engineering Volume 6, Issue 5, May 2016 ISSN: 2277 128X
- [10] Chengpeng Wu ,Junfeng Yao , Songjie "Cloud computing and its key techniques," Electronic and Mechanical Engineering and Information Technology (EMEIT), 2011 International Conference on , vol no. 1, pp.320-324,12-14 Aug. 2011(IEEE).

- [11] Xiaonian Wu, Mengqing Deng , Runlian Zhang , Bing Zeng, Shengyuan Zhou “A task Scheduling algorithm based on QOSdriven” International conference on Information Technology and quantitative Management(ITQM) pp. 1162-1169 , ELSEVIER(2013).
- [12] P. Mathiyalagan , S . Suriya , S.NSivnaam , "Modified Ant Colony Algorithm for Grid Scheduling", *International Journal on computer science and Engineering* Vol. 02, no 02,pp 132-139(2010).
- [13] Medhat A . Tawfeek ,Ashraf El-Sisi, Arabi E. keshk , Fawzy A. Torkey, "An ant Algorithm for cloud Task scheduling" ,*International Workshop on cloud computing and information security (CCIC2013)*
- [14]Xiaonian Wu, Mengqing Deng , Runlian Zhang , Bing Zeng, Shengyuan Zhou “A task Scheduling algorithm based on QOSdriven” International conference on Information Technology and quantitative Management(ITQM) pp. 1162-1169 , ELSEVIER(2013).
- [15] Vijayalakshmi M., Venkatesa Kumar V., “Investigations on Job Scheduling Algorithms in Cloud Computing”, *International Journal of Advanced Research in Computer Science & Technology*, Vol. 2 2014
- [16] Operating System Concepts,6ED,Windows XP Update by silberschatz galvin gagne
- [17] Muhammad Baqer Mollah, Kazi Reazul Islam and Sikder Sunbeam Islam, "Next generation of computing through cloud computing technology", 25th IEEE Canadian Conference on Electrical and Computer Engineering,2012, pp. 1-6.
- [18] P. Mathiyalagan , S . Suriya , S.NSivnaam , "Modified Ant Colony Algorithm for Grid Scheduling",*International Journal on computer science and Engineering* Vol. 02, no 02,pp 132-139(2010).
- [19] Medhat A . Tawfeek ,Ashraf El-Sisi, Arabi E. keshk , Fawzy A. Torkey, "An ant Algorithm for cloud Task scheduling" ,*International Workshop on cloud computing and information security (CCIC2013)*.