Volume 4, Issue 7, July 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Optimization of Job Scheduling in Cloud Computing Environment

Mandeep Kaur*
CGC College, Gharuan
Haryana, India

Sugandha Sharma
CGC College, Gharuan
Haryana, India

Rajinder Kaur
CGC College, Gharuan
Haryana, India

Abstract— This paper addresses parallel machine scheduling problems with practical Swarm Optimization (PSO). A PSO approach embedded in a simulation model is proposed to minimize the maximum completion time (make span). The results are compared with those obtained by using the ''longest processing time'' Rule, which is known as the most appropriate dispatching rule for such problems. This application illustrates the need for efficient and effective heuristics to solve such PSO Scheduling Machine Problem. The proposed PSO approach yields good results quickly and several times in one run. Moreover, because it is a search algorithm, it can explore alternative schedules providing the same results. We Cloudsim for simulation of this approach and we get significance improvement in resource utilization.

Keywords— Cloud computing, scheduling, job scheduling.

I. INTRODUCTION

With the rapid development of storage technologies, processing and success of internet, computing resources have become more powerful and cheaper than even before. This technological fashion lead to the realization of new computing model called cloud computing in which resources are provided as utilities which can be leased or released by various users through internet in an on demand fashion. Thus cloud computing is recently emerged as a new technology that hosting and delivering new services via internet.

Cloud is developing day by day and faces many challenges, one of them is scheduling. Scheduling refers to a set of policies to control the order of work to be performed by a computer system. A good scheduler adapts its scheduling strategy according to the changing environment and the type of task. Cloud Computing is a type of computing that involves a large number of computers that are connected in a network such as internet. We can also say that it is a synonym of distributed computing. In which there are large number of computers that are operating and managed at the same time. Cloud computing is a development of grid computing, parallel computing and distributed computing. Generally cloud refers to the Software, Infrastructure, and platform that are sold 'as a service' over the internet. There are various number of cloud networks like public cloud, private cloud and hybrid cloud. Cloud computing is basically a combination of two things that is Online application and Online Storage. Gmail is an excellent example. If you are using the various social networking sites like Gmail, yahoo etc. then you are cloud computing user. These software applications are not installed on your computer but you are using them over the internet. Similarly if you are iphone user and you have enabled icloud then your apps, videos and photos are backed up or stored by the computer managed by the apple and the data will be transfer to that computer by the internet. There are various cloud computing examples like Amazon, Google, Oracle Cloud and SalesForce etc. Amazon EC2 (Amazon elastic compute cloud) enables the different cloud users to launch and manage the various server instances using the application programming interface (API) or available tools or utilities. EC2 provides the ability to create the instances at multiple locations. Apache Hadoop is an open source software framework for large scale processing and storage for data sets. Now Cloud computing reached at the point where it can take place of your entire Operating system.

Since cloud computing is heterogeneous pool of resources so scheduling plays a major role in this field. Scheduling tasks to the cloud resources is an important step in cloud source management. Scheduling or job scheduling is a task of assigning the system resources to the various tasks that are waiting for the CPU time and emerged in a queue. The system must decide that which particular job took first give it the CPU time for processing, so that all of the jobs can executed in fair and efficient manner. Also fairness in scheduling is the important criteria that provides the resource allocation in an optimized way and improves efficiency. As the cloud scale expanding, scheduling is still an issue to resolve. An efficient scheduling mechanism should meet the Qos parameters and enhance resource utilization. The proposed method of job scheduling in this paper is that to more optimize the way of scheduling by various machine learning algorithms so that there will be a linear or non linear mapping of tasks to resources and scheduling by enhancing the Berger model theory of job scheduling, in which the concept of distributive justice is used. Distributive justice is based on expectation states. Expectation states are used to present the justice or injustice by distributing the resources under various circumstances.

The proposed scheduling algorithm is in cost effective manner with short make span and meets the various Qos parameters like bandwidth, utilization rate and time.

© 2014, IJARCSSE All Rights Reserved

II. PRESENT WORK

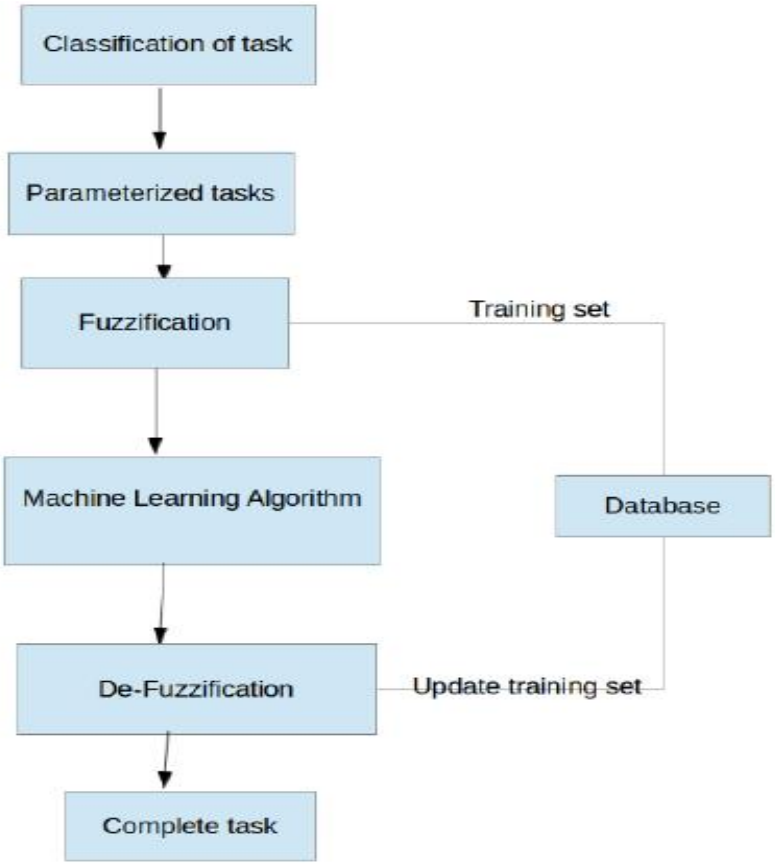
Cloud Computing deals with different type of virtualized resources, hence scheduling Optimization places an important role in cloud computing. In cloud, user may use hundreds of thousands virtualized resources for each user tasks. Hence manual scheduling is not a feasible solution. Focusing scheduling to a cloud environment enables the use of various cloud services to help framework implementation. The proposed work will use machine learning algorithms for the non linear mapping relationship between Qos and resource. Through learning more optimize use of resources and increase the utilization of resources. Learning approach is also important in dynamic environment of task requirement of resources in Cloud computing.

III. OBJECTIVES

In the process of scheduling some intensive data or computing an intensive application, it is acknowledged that optimizing the transferring and processing time is crucial to an application program.

- To minimize the cost of processing by proper utilization of CPU.
- Scheduling of resources to the tasks and apply them on machine learning algorithms.
- Then validate the proposed work.

IV. RESEARCH DESIGN



i) The data flow diagram which is used to represent how the resources are allocated to a particular resource using virtual machine.Initially, large number of tasks is submitted to the system. After that, check whether any priorities exist in submitted tasks. If the priority exists, then sort the tasks submitted according to the priority.

ii) In task parameterization can be used to check whether all resources are available for particular task to execute. Fuzzification is used to convert the different range of input values.

iii) Machine Learning algorithm is used to find out a best result of mapping the user tasks. De-fuzzification is used to convert the linguistic variables into values and check the task has been completed before the expectation time. Suppose, if the task is not mapped with system resources, update the training set and reclassify the user tasks. Otherwise just update the training set and realize the corresponding virtual machine.

V. EXPERIMENTAL ANALYSIS

The results show that the proposed GA approach yields good results efficiently. An examination of Table 2 reveals that the optimal solution is reached at only the 5th iteration. In addition, the optimal solution is reached several times in one run. The same Table 2 shows that the optimal solution is reached 8 times in 50 iterations, which demonstrates the efficiency of the algorithm. The same success is repeated during the robustness tests performed with 10 different random number sets. The effectiveness of the algorithm can be evaluated by comparing the result of the GA approach (51 min, given in Table 2

© 2014, IJARCSSE All Rights Reserved

Table-1

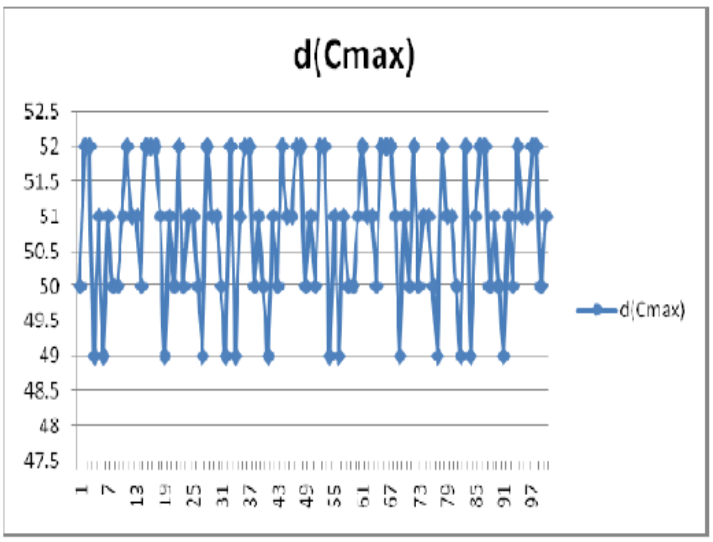
values of processing times									
processing times		T1	T2	T3	T4	T5	T6	T7	T8
PT	P	13	18	24	22	28	15	19	24
	Q	13	18	24	22	28	15	19	24

Table-2

Iterations	Local maxima	global maxima	d(Cmax)
1	50	50	50
2	52	52	52
3	52	52	52
4	49	49	49
5	51	51	51
6	49	49	49
7	51	51	51
8	50	50	50
9	50	50	50
10	51	51	51
11	52	52	52
12	51	51	51
13	51	51	51
14	50	50	50
15	52	52	52
16	52	52	52
17	52	52	52
18	51	51	51
19	49	49	49
20	51	51	51
21	50	50	50
22	52	52	52
23	50	50	50
24	51	51	51
25	51	51	51
26	50	50	50
27	49	49	49
28	52	52	52
29	51	51	51

© 2014, IJARCSSE All Rights Reserved

73	50	50	50	
74	51	51	51	
75	51	51	51	
76	50	50	50	
77	49	49	49	
78	52	52	52	
79	51	51	51	
80	51	51	51	
81	50	50	50	
82	49	49	49	
83	52	52	52	
84	49	49	49	
85	51	51	51	
86	52	52	52	
87	52	52	52	
88	50	50	50	
89	51	51	51	
90	90	90	90	
91	49	49	49	
92	51	51	51	
93	50	50	50	
94	52	52	52	
95	51	51	51	
96	51	51	51	
97	52	52	52	
98	52	52	52	
99	50	50	50	



Graph No.2

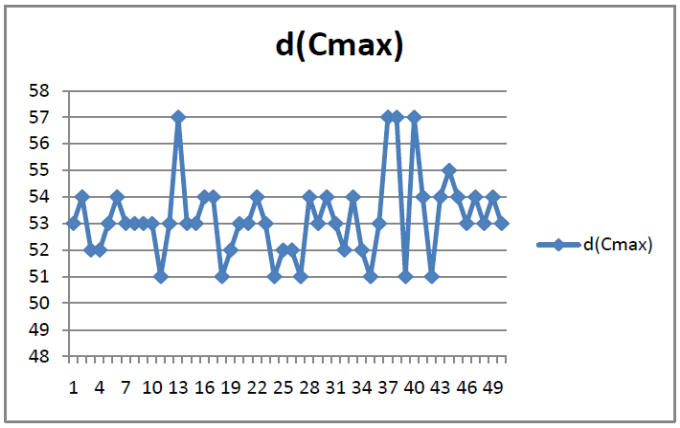
© 2014, IJARCSSE All Rights Reserved

Table-3

Iterations	local maxima	global maxima	d(Cmax)
1	38	50	53
2	38	53	54
3	39	51	52
4	39	51	52
5	43	48	53
6	42	52	54
7	38	50	53
8	38	50	53
9	43	48	53
10	38	55	53
11	41	52	51
12	43	48	53
13	43	52	57
14	43	48	53
15	43	48	53
16	42	52	54
17	42	52	54
18	41	52	51
19	39	51	52
20	38	50	53
21	38	50	53
22	38	53	54
23	43	48	53
24	41	52	51
25	39	51	52
26	39	51	52
27	41	52	51
28	42	52	54
29	43	48	53
30	38	53	54
31	38	50	53
32	39	51	52
33	42	52	54
34	39	51	52
35	41	52	51
36	38	50	53

© 2014, IJARCSSE All Rights Reserved

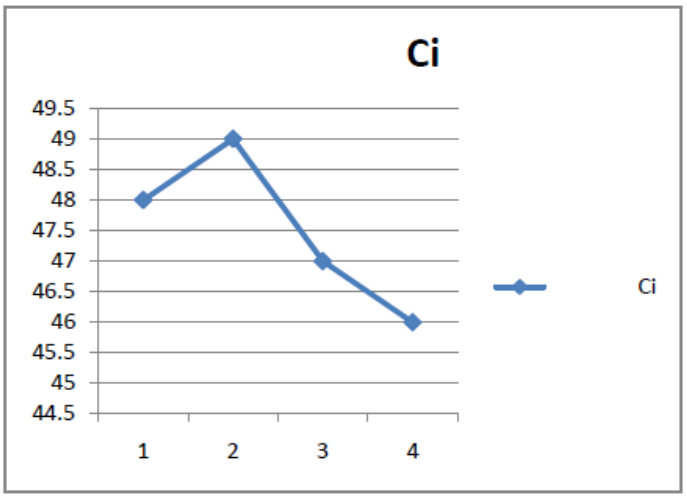
37	43		52		57
38	43		52		57
39	41		52		51
40	34		47		57
41	42		52		54
42	41		52		51
43	42		52		54
44	40		51		55
45	38		53		54
46	38		55		53
47	42		52		54
48	38		55		53
49	42		52		54
50	38		50		53



Graph 3

Table no.4:- The second optimal schedule with value of processing

Machine	Scheduled jobs			Ci
M1	job8	job3		48
M2	job4	job9		49
M3	job1	job7	job6	47
M4	job2	job5		46



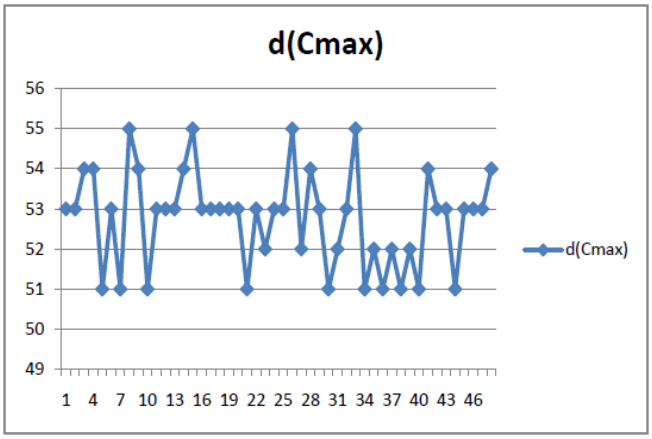
Graph-4

© 2014, IJARCSSE All Rights Reserved

Table-5

iterations	local maxima	global maxima	d(cmax)
1	38	50	53
2	38	55	53
3	42	52	54
4	38	53	54
5	41	52	51
6	38	50	53
7	41	52	51
8	40	51	55
9	42	52	54
10	41	52	51
11	43	48	53
12	38	55	53
13	43	48	53
14	38	53	54
15	40	51	55
16	43	48	53
17	38	55	53
18	38	50	53
19	38	50	53
20	41	52	53
21	38	55	51
22	38	55	53
23	41	52	52
24	38	50	53
25	39	51	53
26	38	55	55
27	40	51	52
28	39	51	54
29	42	52	53
30	38	55	51
31	41	52	52
32	39	51	53
33	38	50	55
34	40	51	51
35	41	52	52
36	39	51	51
37	40	51	52
38	41	52	51
39	39	51	52
40	41	52	51
41	42	52	54
42	43	48	53
43	38	55	53
44	41	52	51
45	38	50	53
46	38	50	53
47	43	48	53
48	38	53	54

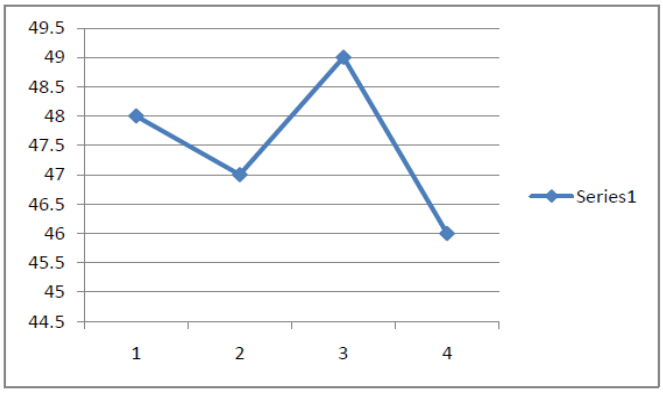
© 2014, IJARCSSE All Rights Reserved



Graph 5

Table 6:- This table represent optimization of task after PSO

iteration no.4&19				
M1	task3	task3		48
M2	task5	task9		47
M3	task4	task7		49
M4	task2	task5	task6	46



Graph 6

VI. IMPLEMENTATION TOOL

Cloud computing is a recent advancement wherein IT infrastructure and applications are provided as 'services' to end users
under a usage-based payment model. The application services hosted under Cloud computing model have complexprovisioning, composition, configuration, and deployment requirements.
The main features of cloud sim are:

_ Support for modeling and simulation of large scale cloud computing data centers.
_ Support for modeling and simulation for energy aware resources.
_ Support for user defined policies for allocation of hosts to virtual machine and policies for allocation of host Resources to virtual machine.

VII. Related Work

Dr. Ajay jangra, **Tushar Saini.** Cloud computing is a most recent new computing paradigm where applications, records and IT services are provided over the Internet. Cloud computing has come out to be an interesting way of changing the whole computing Schedulers for cloud computing determine on which processing resource jobs of a workflow should be allocated. Scheduling theory for cloud computing is in advance a lot of awareness with increasing popularity in this cloud area. The received tasks are grouped on the basis of data and requested resources by the task and prioritized. Resource selection is done on the basis of its cost and turnaround time. Task selection is on the basis of a priority formula. This way of resource selection and task selection gives better results over sequential scheduling. This paper will give the way for the future findings related to scheduling techniques.

© 2014, IJARCSSE All Rights Reserved

Ranjana saini , Indu stated that Cloud computing is based on the concepts of distributed computing, grid computing, utility computing and virtualization. It is a virtual pool of resources which are provided to users via Internet. This leads to job scheduling as a challenging issue in cloud computing. In this research a discussion towards the resource management of virtual machines in cloud and how to make resources more efficiently available to clients is provided. The main focus is on job scheduling. In this present work, there is a parametric analysis performed to identify the requirement of process migration and according to this analysis the migration will be performed on these processes. The effectiveness of the work is identified in terms of successful execution of the processes within the time limits.

Jinhua Hu, jianhua Gu, Tianhai Zhao. In view of the current load balancing in VM resources scheduling, this paper presents a scheduling strategy on VM load balancing based on genetic algorithm. Considering the resources scheduling in cloud computing environment and with the advantage of genetic algorithm, this method according to historical data and current states computes in advance the influence it will have on the whole system when the current VM service resources that need deploying are arranged to every physical node, and then it chooses the solution which will have the least influence on the system after arrangement. In this way, the method achieves the best load balancing and reduces or avoids dynamic migration thus resolves the problem of load unbalancing. In genetic algorithm, fitness function is the criterion for the quality of the individuals in the population. It directly reflects the performance of the individuals.

Yogita Chawla, Mansi Bhonsle stated that Cloud computing is based on the concepts of distributed computing, grid computing, utility computing and virtualization. It is a virtual pool of resources which are provided to users via Internet. Cloud computing service providers' one of the goals is to use the resources efficiently and gain maximum profit. This leads to task scheduling as a core and challenging issue in cloud computing. This paper gives different scheduling strategies and algorithms in cloud computing which helps to understand the wide tasks of various scheduling options to select one for a given environment.

VIII. CONCLUSION

The parallel machine scheduling problem receives considerable attention in both academic and industrial fields. Various factors involved in scheduling problems are often imprecise or uncertain. The fuzzy set theory provides a convenient alternative framework for modeling real-world systems mathematically and offers several advantages in the use of heuristic S. Balin / Information Sciences 181 (2011) 3551–3569 approaches. In this study, the parallel machine scheduling problem with PSO is considered and a GA approach is proposed to minimize the maximum completion time.

REFERENCES

[1] "Job Scheduling algorithm using Berger model in Cloud Environment ", Baomin Xu, Chunyan Zhao, Enzhao Hu, Bin Hu,, Elsevier in Advances in Engineering Software, Vol. 42 , No. 7, Pp. 419-425, 2011

[2] "Particle swarm optimization scheme to solve resource-constrained scheduling problem". Chen, R. M.; Wu, C. L.; Wang, C. M. & Lo, S. T. (2010). Expert systems with applications, Vol. 37, No. 3 (March 2010), *pp. 1899-1910, ISSN:* 0957-4174.

[3] "A Berkeley view of cloud computing", M. Armbrust, A. Fox, R. Grifth, A. D. Joseph, R. Katz, A. Konwinski,G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the clouds Technical report, University of California at Berkeley, February 2009.

[4] "A Survey of Proposed Job Scheduling Algorithms in Cloud Computing Environment" Rohit O. Gupta, Tushar Champaneria, Computer Science & Engineering, L.D.College of Engineering, India,Volume 3, Issue 11, November 2013.

[5] "Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit:Challenges and Opportunities",Rajkumar Buyya, Rajiv Ranjan, Rodrigo N. Calheiros, in The 2009 International Conference on High Performance Computing and Simulation, HPCS2009, pp:1-11.

[6] "An Optimistic Job Scheduling Strategy based on QoS for Cloud Computing", Huang Q.Y., Huang T.L IEEE International Conference on Intelligent Computing and Integrated Systems (ICISS), Guilin, pp. 673-675, 2010.

[7] "Research issues in Cloud Computing",V.Krishna Reddy, B. Thirumala Rao , LSS Reddy. Global Journal Computer Science & Technology Vol. 11, no. 11, June 2011,pp.70-76.

[8] "Tasks scheduling optimization for the cloud computing systems" Sandeep Tayal, international journal of advanced engineering sciences and technologies vol no. 5, issue no. 2, 111 – 115.

[9] Li Jianfeng, Peng Jian. Task scheduling algorithm based on improved genetic algorithm in cloud computing environment[J].JournalofComputerApplications,2011,1(31):184~186

[10] "Job Scheduling Model for Cloud Computing Based on Multi-Objective Genetic Algorithm", Jing Liu , Xing-Guo Luo, Xing-Ming Zhang3, Fan Zhang and Bai-Nan Li. International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013.

[11] "Fairness As Justice Evaluator In Scheduling Cloud Resources: A Survey" B. Anuradha, S. Rajasulochana. International Journal of Computer Engineering & Science, Nov. 2013. ©IJCES ISSN: 22316590

© 2014, IJARCSSE All Rights Reserved