International Journal of Application or Innovation in Engineering & Management (IJAIEM)
Web Site: www.ijaiem.org Email: editor@ijaiem.org
Volume 3, Issue 7, July 2014 **ISSN 2319 - 4847**

TASK SCHEDULING OPTIMIZATION OF INDEPENDENT TASKS IN CLOUD COMPUTING USING ENHANCED GENETIC ALGORITHM

[1]Shekhar Singh, [2]Mala Kalra

[1]Department of Computer Science, NITTTR, Panjab University, Chandigarh ,

[2]Assistant Professor, Department of Computer Science, NITTTR, Panjab University, Chandigarh

ABSTRACT

Parallel and distributed computing has been entered into the era of cloud computing which provides highly scalable computational power and remote location data storage with secure, fast and in cost effective way. So it has captured the attention of big IT giants for availing flexible, readily elastic and cost effective economical IT operations. Cloud computing environments provide dynamically manageable virtualized computing resources. The management of these resources is taking big account in cloud. So the scheduling has of vital importance in cloud computing. In this paper, we give an elaborate idea about Genetic Algorithm and its several variants proposed for task scheduling in cloud environment and idea of GA based scheduler is proposed in which population is produced by Enhanced Max Min by which makespan can be reduced and load of resources can be balanced.

Keywords: Cloud Computing, Task Scheduling, Genetic Algorithm (GA), Efficient GA.

1. INTRODUCTION

Today is the era of smart computing. Everybody wants to use resources instantly. So researchers started to think about a technology, which can serve anywhere anytime. Mark Weiser solved this problem by an intelligent idea of computing, known as ubiquitous computing. In morden age, we are using this as Cloud Computing. Cloud computing is a new form of ubiquitous computing which delivers almost all the IT services through the internet in virtualized manner. There are three types of cloud computing services : Platform as a Service (PaaS), Infrastructure as a service (IaaS) and Software as a Service (SaaS) [1,2] which can accessed by users in a Pay-per-Use-On-Demand model in which guaranteed high quality cloud services are offered to the customer by the service providers through tailored SLA[3,4]. SLA can be achieved by effective utilization of resources so the management of these resources are very important. In cloud computing, the effective scheduling policy can maximize the throughput of the cloud environment. The main purpose is to schedule tasks to the adaptable resources in accordance with adaptable time, which involves finding out a proper sequence in which tasks can be executed under transaction logic constraints. So in cloud computing, there are various type of meta heuristic algorithms for scheduling problem in cloud computing such as genetic algorithm (GA), ant colony optimization (ACO), particle swarm optimization (PSO) etc [16]. Due to the versatile features of GA algorithm, like easy interfacing with existing simulations and models and the ability of handling multiobjective problems, GA is widely used among all algorithms of its category [6]. The rest of this paper is organized as follows. Section II briefly describes related research for scheduling in grid computing and Cloud Computing. Section III discusses about the scheduling techniques. Section IV describes the Genetic Algorithm. Section V gives the idea about the new Improved Genetic Algorithm telling how we can combine Min-Min and Max-Min in genetic Algorithm. Section VI is having the simulations and results. Section VII tells about the future scope and conclusion of this paper.

2. GENETIC ALGORITHM SCHEDULING

Genetic algorithm is based on “Selects The Best, Discards The Rest” principle. With the help of Genetic Algorithms (GAs) technique, task scheduler can search optimized solution from large solution spaces in polynomial time. Genetic algorithm is a directed search algorithm based on biological progression. Researchers have proposed several variations of standard genetic algorithms to improve its efficiency of finding optimized solutions. Figure 2 shows the steps used in a general GA.

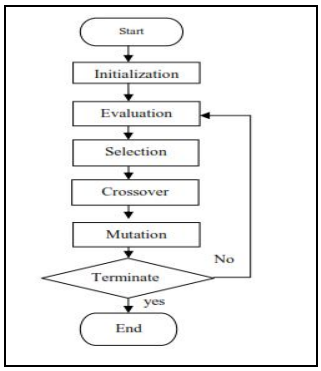

Fig. 1: Flow of Standard Genetic Algorithm [17]

Genetic algorithm begins with the initialization of population. Initialization of population is a method in which the set of individuals are defined which represents the possible solution to the scheduling problem [5]. After that selection of better individuals will be determined with the help of fitness function. The crossover and mutation operation is used for shuffling of newly generated population.

3.RELATED WORK

Scheduling of tasks is a critical issue in Cloud Computing, so a lot of researches have been done in this area. The following table summarizes different genetic task scheduling algorithms based on their scheduling goals, parameters and tools used along with future scope.

Paper	Findings	Scheduling Parameters	Tools	Future scope
Independent Task Scheduling in Cloud Computing by Improved Genetic Algorithm[7]	* Initial population is generated by using the Min-Min and Max-Min techniques. * Fitness function is based on minimization of makespan. * Proportion selection method is used as selection operator. * Single-point crossover operator is used for crossover operation.	Makespan, Resource utilization	CloudSim	Execution cost can be taken as fitness criteria. Explore towards dependent and dynamic jobs.
Hard Real-Time Multiobjective Scheduling in Heterogeneous Systems Using Genetic Algorithms[8]	* Initial population is generated using modified chromosome encoding method. * Single point crossover method is used. * Designed a scheduler for hard realtime heterogeneous multiprocessor systems which is based on cache reload time.	Cache reload time, Optimal tasks allocation, Deadlines constraints	Mathematical model for Hard Real-Time Scheduling Problem	Try to design some better stopping criteria and improve initial population generation method.
Heuristics Based Genetic Algorithm for Scheduling Static Tasks in Homogeneous Parallel System[10]	* Initial population is generated using Minimum Execution Time and Min-Min algorithm heuristic approaches. * One point and two point crossover operation is used. * Swap mutation operator is used for mutation.	Execution time and Waiting time	Deterministic model based on Directed Acyclic Graph (DAG).	Try to increase the throughput of the system.
Tasks Scheduling optimization for the Cloud Computing Systems[9]	* Fuzzy GA optimization is used in which scheduling decision is made by evaluating the entire group of task in the job queue. * Initial population is randomly generated. * Fitness function is based on minimization of makespan.	Load balancing, Execution time	Hadoop MapReduce technology	Improvement required on the accuracy of predicted completion time of job.

An Efficient Approach to Genetic Algorithm for Task Scheduling in Cloud Computing Environment[15]	* Initial population is generated by using SCFP and LCFP techniques in private cloud environment. * Those individuals will be selected for next generation population whose value of makespan and execution cost is less than the standard genetic algorithm. * Two-point crossover method is used. * Swap mutation operator is used for mutation.	Makespan and Cost	CloudSim	Need to improve search space.
Independent Tasks Scheduling Based on Genetic Algorithm in Cloud Computing[11]	* Fitness function is based on deadline in order to regulate cost of delay and store so that customer satisfaction can be maximized. * Two-point crossover method and a mathematical formula is used for crossover operation. * Mutation is based on mathematical formula.	Resources utilization, Execution time	Cloud based Service Oriented Architecture Model	More attention is required on reducing the solution space in GA.
Cloud Computing – Task Scheduling based on Genetic Algorithms[12]	* Map and Reduce operation are used to improve Hadoop's functionality by implementing a task scheduler based on a genetic algorithm. * Initial population is randomly generated and selection operation is performed using roulette-wheel method. * Cycle crossover method is used as crossover operation. * Swap mutation operator is used for mutation.	Hadoop's task assignment sequencing, Map-Reduce	Genetic Scheduler For Hadoop	Try to reduce solution space, trying other GAs, combining presented techniques with Hadoop's existing schedulers, especially with Capacity Scheduler etc.
Reputation guided Genetic Scheduling Algorithm for Independent Tasks in Inter-Clouds Environments[13]	* The selection method used roulette wheel principle. * One-Point Crossover and Two-Point Crossover is used as crossover operation. * Order-based Mutation and Partial-Gene Mutation is used as mutation operation. * Fitness function is based on a mathematical formulation which considers minimum and maximum execution time.	Resource utilization, Makespan, Load balancing	Mathematical Genetic Algorithm Model for inter cloud environment	The algorithm can improve its results by improving the fitness function criteria.
HSGA: a hybrid heuristic algorithm for workflow scheduling in cloud systems[14]	* Best-Fit and Round Robin methods are used to create initial population to find an efficient scheduling solution. * Random gene selection crossover method is used for crossover operation.	Makespan and Load balancing	DAG (Direct Acyclic Graph) based on mathematical formulation.	Try to develop method that support mapping the resources on tasks for communication intensive applications with efficient result.

4.THE PROPOSED WORK

Generally in GA initial population are used to generate randomly. This initial population are used further for generation of new population. Randomly generated solution are not efficient because there is chance that newly generated solution may be repeated. So In our proposed work, we have generated initial population by the Enhanced Max-Min algorithm [18]. The working process of Enhanced Max Min is given below in Fig 2.

1. For all submitted tasks in Meta-task; T_i
   1.1. For all resources; R_j
   1.1.1. $C_{ij} = E_{ij} + r_j$
2. Find task T_k costs Average or nearest Greater than Average execution time.
3. Assign task T_k to resource R_j which gives minimum completion time (Slowest resource).
4. Remove task T_k from Meta-tasks set.
5. Update r_j for selected R_j.
6. Update C_{ij} for all j.
7. While Meta-task not Empty
   7.1. Find task T_k costs maximum completion time.
   7.2. Assign task T_k to resource R_j which gives minimum execution time (Faster Resource).
   7.3. Remove Task T_k form Meta-tasks set.
   7.4. Update r_j for Selected R_j.
   7.5. Update C_{ij} for all j.

$C_{i,j}$ = completion time of job/task i on machine/resource j
r_j = ready time of resource R_j
E_{ij} = Execution Time of task T_i on resource R_j.

Fig 2: Enhanced Max Min Algorithm

Above mention algorithm can produce more optimized makespan if we use this in Genetic Algorithm. So here we introduce a idea in which initial population is not generated randomly. Instead of this , initial population is generated by advance section of Max-Min algorithm. After initializing the population by Enhanced Max Min, resulting population will be evaluated by fitness criteria. Crossover and Mutation operation will be applied on selected individuals. This process ends when termination condition meet. The below mention fig 3 shows the flow of proposed methodology.

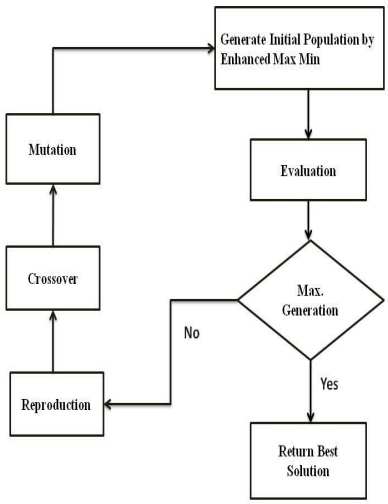

Fig. 3 Proposed Genetic Algorithm

5.RESULT DISCUSSION

We have used CloudSim as a simulator for checking the performance of our improved algorithm and the standard Genetic Algorithm. CloudSim is an extensible simulation toolkit that enables modeling and simulation of Cloud computing systems and application provisioning environments. The CloudSim toolkit supports both system and behaviour modeling of Cloud system components such as data centers, virtual machines (VMs) and resource provisioning policies. It implements generic application provisioning techniques that can be extended with ease and limited efforts. We have considered Virtual Machines as resource and Cloudlets as tasks/jobs. We have measured the makespan of different algorithms while number of VMs are varying and cloudlets are fixed.

Testing Case: Number of Cloudlets fixed (100) and Varying VMs

Table II: Comparison of Makespan

Sample	Improved Max-Min	Efficient Max-Min	GA-LCFP	IGA	Efficient Genetic Algorithm (EGA)
1(5 VM)	70.5	39.3	41.55	48.17	32.84
2(10 VM)	54.5	28.88	28.19	37.51	26.95
3(15 VM)	49.68	25.7	20.28	26.51	17.11
4(20 VM)	46.5	24.09	14.39	19.51	10.83

Above Table shows that proposed Efficient Genetic Algorithm produced lesser makespan then other algorithms because algorithm selects the task which have average expected time and assign it to slowest resource. And rest all task is scheduled to faster resources.

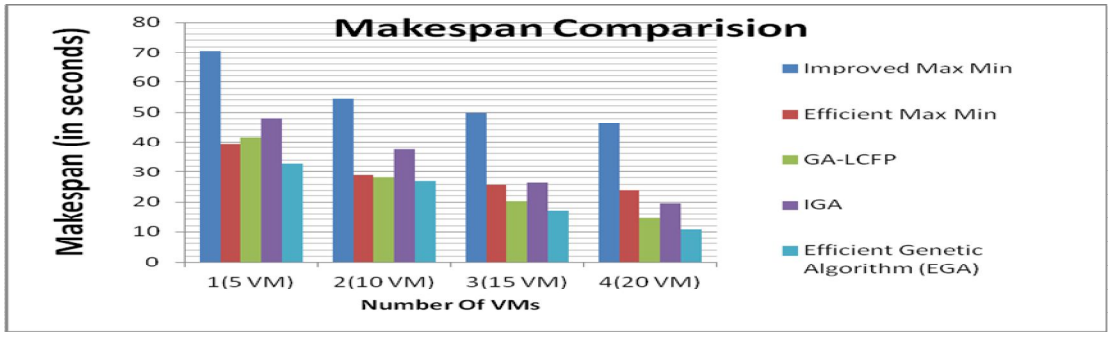

Fig. 4: Makespan Of all algorithm

6.CONCLUSIONS AND FUTURE WORK

Task scheduling is main problem in cloud computing scenario. Efficient task scheduling is essential for better utilization of resources. In this paper three scheduling algorithms are discussed such as LCFP, Min-Min, Max-Min and Genetic Algorithm. The idea of genetic algorithm comes from natural selection which consists of population generation, selection, and mutation. This paper is our sincere attempt to reduce the makespan of algorithm by using enhanced Max Min for initializing the population in GA. The experimental results show that the Efficient genetic algorithm minimizes the makespan effectively than other algorithm. This idea can be further extended in which we can use execution cost of the resource as fitness criteria.

REFERENCES

[1] L. M. Vaquero, L. R. Merino, J. Caceres, and M. Lindner. "A break in the clouds: towards a cloud definition." ACM SIGCOMM Computer Communication Review 39, no. 1 ,2008, pp. 50-55.
[2] NIST, http://www.nist.gov/itl/cloud/
[3] L. Wu, and R. Buyya. "Service Level Agreement (SLA) in utility computing systems." Performance and Dependability in Service Computing: Concepts, Techniques and Research Directions, 2011, pp: 1-25.
[4] S. M. Hashemi, A. K. Bardsiri, Cloud Computing Vs. Grid Computing. ARPN Journal of Systems and Software, Vol. 2, No 5, pp. 188-194, 2012.
[5] F. Pop, C. Dobre, and V.s Cristea. "Genetic algorithm for DAG scheduling in grid environments." Intelligent Computer Communication and Processing, 2009. ICCP 2009. IEEE 5th International Conference on. IEEE, 2009.
[6] S. Singh and M. Kalra, "A Comparative Study of Task Scheduling Based on Genetic Algorithm in Cloud Computing", International Conference on Computer Networks and Information Technology (ICCNIT 2014), 20-21 March, 2014, ISBN: 978-93-83842-27-8, pp. 325-332.
[7] P. Kumar, and A. Verma. "Scheduling using improved genetic algorithm in cloud computing for independent tasks." Proceedings of the International Conference on Advances in Computing, Communications and Informatics. ACM, 2012.
[8] M. R. Miryani, and M. Naghibzadeh. "Hard real-time multiobjective scheduling in heterogeneous systems using genetic algorithms." In Computer Conference. CSICC 2009. 14th International CSI, pp. 437-445. IEEE, 2009.
[9] S. Tayal. "Tasks scheduling optimization for the cloud computing systems." International Journal of Advanced Engineering Sciences And Technologies (IJAEST) 5, no. 2 (2011): 111-115.
[10] K. Kaur, A. Chhabra, and G. Singh. "Heuristics based genetic algorithm for scheduling static tasks in homogeneous parallel system."international journal of computer science and security 4.2 (2010), pp: 183-198.

[11] C. Zhao, S. Zhang, Q. Liu, J. Xie, J. Hu. "Independent tasks scheduling based on genetic algorithm in cloud computing." Wireless Communications, Networking and Mobile Computing, 2009. WiCom'09. 5th International Conference on. IEEE, 2009.
[12] E. M. Mocanu, M. Florea, M. I. Andreica, and N. Tapus. "Cloud Computing—Task scheduling based on genetic algorithms." Systems Conference (SysCon), 2012 IEEE International. IEEE, 2012.
[13] P. Florin, V. Cristea, N. Bessis, and S. Sotiriadis.. "Reputation Guided Genetic Scheduling Algorithm for Independent Tasks in Inter-clouds Environments." Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on. IEEE, 2013.
[14] A. G. Delavar, and Y. Aryan. "HSGA: a hybrid heuristic algorithm for workflow scheduling in cloud systems." Cluster Computing,2013, pp: 1-9.
[15] S. Kaur, and A. Verma. "An Efficient Approach to Genetic Algorithm for Task Scheduling in Cloud Computing Environment." International Journal of Information Technology and Computer Science (IJITCS) 4, no. 10, 2012.
[16] Xu M., Cui L.,Wang H.,Bi Y. "A multiple QoS constrained scheduling strategy of multiple workflows for cloud computing." Parallel and Distributed Processing with Applications, 2009 IEEE International Symposium on. IEEE, 2009.
[17] J. S. Raj, and R. M. Thomas. "Genetic based scheduling in grid systems: A survey." Computer Communication and Informatics (ICCCI), 2013 International Conference on. IEEE, 2013.
[18] Bhoi U. and Ramanuj P. N., "Enhanced Max-min Task Scheduling Algorithm in Cloud Computing." International Journal of Application or Innovation in Engineering and Management (IJAIEM), pp. 259-264,ISSN : 2319-4847, 2013.

AUTHOR

Shekhar Singh received the B.E. degree in Information Technology from I.E.T, Agra University in 2007. He is receiving M.E. degree in Computer Science and Technology from NITTTR, Panjab University, Chandigarh.