

An Improved Fuzzy-Genetic based Model for Cloud Scheduling Optimization

Durgesh Kumar Joshi¹, Vinod Rampure²

¹Student, M.Tech. (CSE), MITRC, Alwar, India

²Assistant Professor, MITRC, Alwar, India

ABSTRACT

Cloud computing enables the sharing on resources and services in distributed environment. Different kind of public and private users can request the services to multiple cloud servers through a common interface. When the number of request increases, the cloud scheduling method is applied for effective utilization and distribution of resources. In this paper, an improved scheduling method using fuzzy integrated genetic is provided. At first level of this method, the arrangement of user requests and virtual machines is done based on capabilities and requirement analysis. The cost based population matrix generated based on time and resource specific parameters. The population set is processed under fuzzy adaptive genetic model and generate the optimized scheduling sequence. The proposed method is implemented in matlab environment. The comparative results shows that the method has reduced the average wait time, turnaround time and finish time.

Keywords: Scheduling, Cloud Computing, Genetics, Fuzzy, Resource.

1. INTRODUCTION

Cloud Computing[1][2][3][4] defines is robust distributed computing model in which different kind of services and resources are shared through common environment. Cloud system provides the sharing at service level, hardware level and platform level. Different kind of service providers and the users are connected through this integrated environment. In this environment multiple cloud servers are present that can provide same kind of services. In same way, multiple users can demand for same kind of services. The cloud servers are virtually divided in multiple virtual machines with specific configurations. Some scheduling method is required to utilize the cloud servers and resources effectively as well as to provide the services to the users without any delay. Scheduling is one such method that identifies the processing sequence of these cloud servers on user requests. There are different types of available scheduling method that improves the work effectiveness and improve the service allocation on different cloud servers and virtual machines. Some of common forms of scheduling methods include Task Scheduling, Job Scheduling, Resource Scheduling, Workflow Scheduling etc. Different feature driven measures can be applied on these scheduling methods to improve the performance and reliability of process execution on cloud servers.

1.1 Scheduling

Scheduling in a cloud system is defined to provide effective and reliable cloud service allocation to the clients. When the load on cloud server increases, there is the requirement to setup the sequence of request processing. The scheduling basically generates such optimized processing sequence for user requests for particular services. There are different static and dynamic scheduling techniques that work on different parameters and constraints. These constraints can be based on role specification, feature specification and workflow specification[1][4][8][10][12]. The real-time and the heuristic scheduling method is defined to generate the optimized processing sequence. The categorization of different scheduling forms is defined here under

1.1.1 User-Level Scheduling

This kind of scheduling is required when the cloud servers are facing heavy access load. To obtain the load balancing and load distribution, the user level scheduling can be applied. In public and authorization specific cloud environment such kind of scheduling is quite common. According to the user requirements, request time and importance, the sequence of user request processing is decided.

1.1.2 Dynamic Task Scheduling

This kind of scheduling is based on user requirement and demand computation based analysis. A dynamic service level and requirement level observation is defined to assign the weights to user request. The current availability and current requirement analysis is considered as the main criteria for setting the scheduling sequence for user requests. In the larger and the hybrid cloud environment such kind of scheduling is required. This scheduling method can also applied on virtual machines to setup the processing sequence so that the effective utilization of resource will be done.

1.1.3 Real Time Task Scheduling

While working with different services and task, there exist different types of tasks. These tasks can be categorized based on the criticality of task completion. Some of the tasks are highly critical that must be completed within the defined deadline. Such kind of time-bounded tasks are called real time asks. The real time task scheduling is able to execute the processes before the defined dead line.

In this paper, a fuzzy based genetic method is defined to improve the scheduling aspects. The fuzzy model is defined in crossover stage of genetic algorithm. The work is defined to reduce the wait time and turnaround time of user requests. In this chapter, the brief introduction to cloud system, cloud scheduling and the exploration on various scheduling methods is provided. In section II, the work provided by earlier researchers is discussed. In section III, the proposed research methodology is discussed. In section IV, the results obtained from the work are presented. In section V, the conclusion obtained from work is presented.

II LITERATURE REVIEW

Lot of work is already provided by the researchers to optimize the scheduling process using optimization algorithm and using parametric improvement. In this section, the work provided by earlier researchers is discussed. Author[1] has used the cost adaptive genetic algorithm to improve the resource scheduling. Author used the one point crossover and uniform mutation as its work stages to optimize the scheduling process. Author[2] has improved the associated task scheduling based on delay analysis. The concurrent system was improved by the author by applying the group processing and control to optimize the resource allocation in cloud environment. The resource utilization in controlled manner was provided by the author to reduce the process time. Author[3] used the analytical process model to generate the comparison matrix based on resource adaptation. The task priority oriented inconsistency analysis method was provided to improve the reliability of resource allocation. Author considered the resource requirement factors including memory, bandwidth and processing power to optimize the scheduling behavior. Author[4] defined a study work on workflow scheduling method. The basic characterization and the formulation of various associated aspects were discussed by the author. The algorithm available on workflow scheduling was explored by the author with constraint specification. The benefit and limitation of algorithms were also identified by the author. Author[5] has identified the limitations in existing scheduling method increase of multiple computation models. The meta heuristic scheduling methods were reviewed by the author. The issues and the future scope of algorithms were identified by the author.

Author[6] used the energy adaptive task scheduling method in distributed cloud environment. The energy saving techniques DVFS and virtual machine reuse were defined by the author. The integrated method not only improved the QoS but also reduced the energy consumption. The evaluation of method was done in cloudsim environment. Author[7] used the batch mode based dynamic scheduling method in distributed cloud environment. An improve makespan and bandwidth based Berger model was defined to reduce the wait time. The improvement to the existing min-min algorithm was suggested by the author. Another work on min max and min-min algorithm was provided by Panda et al.[8]. Author used two phase scheduling method with batch processing to optimize the scheduling process. The load consideration with batch analysis was provided to reduce the execution makespan. Author[9] has provided the work to improve the reliability of cloud system environment. In a dynamic computing environment like cloud computing, it is hard to maintain and analyze the reliability of the resources. Author[10] has proposed a work on effective work on resource allocation without any wait time. Author defined the deadline for acceptance of leases. The reservation queue is formed to maintain the priority of queues and the wait list is maintaining the listed processes available for execution. Author defined the parameter specific evaluation and experimentation to achieve the effective process execution in slot time. Author[11] has presented a data driven discovery method for decision making and process alignment. Author defined the process level modeling so that the process event log will be formed. Author defined the process discovery so that the process discovery and the rule formation will be done. Author[12] has provided a dynamic data intensive driven scheduling method. A novel security constraints scheduling approach to schedule all jobs in cloud environment efficiently without compromising required security level for each job is presented in this paper. In the regard of cloud security, swarm intelligence is highly capable to provide better solutions for such potentially intractable problems. Author[13] has provided a work on resource scheduling method using genetic approach. In this model, satisfaction of requests and load balancing of resources are constraints. A scheduling algorithm based on genetic algorithm is proposed. Scheduling scheme is coded in integer sequence, and a fitness function based on influence degree is

designed. Author[14] provided the work on rough set based programming framework to improve the scheduling in real time environment. In this research work we have been proposing an approach to provide the rating of CSPs based on the internal performance of Datacenters and virtual machines. In present situation day-by- day number of cloud service providers have been increasing drastically. Author[15] has provided the opportunistic adaptive modeling of processes and under process modeling and process sequence derivation. Author defined the classification so that the event data processing and the process derivation under multiple methods including clustering, regression and regression rule formation.

III. RESEARCH METHODOLOGY

In this present work, an improved fuzzy-genetic driven architecture is defined to achieve the cloud service scheduling. The proposed model is applied on distributed cloud environment and to improve the scheduling results. The presented work is divided in three main layers. The complete architecture is here presented as the middle layer architecture. The intermediate layer is defined to accept the user requests and process on the cloud system so that effective resource allocation will be performed and the optimized scheduling will be obtained. The optimization will be achieved in terms of reduced wait time and with lesser number of process migrations over the cloud server.

In this fuzzy-genetic optimized scheduling model, the distributed cloud system is defined with multiple cloud servers and with multiple virtual machines on each cloud system. Each of the cloud machines is defined with memory, CPU and IO component specification. In first layer of this model, the server level analysis is performed and the virtual machine prioritization is done. In this stage, the resource level analysis is performed on the virtual machines and the processing sequence on these virtual machines is done. Once the virtual machine ordering is done, the user request analysis is done in the second layer of this model. The user requests are defined with process time, dead line criticality and other constraint processing. In this layer, the user request processing is done to identify the effective service sequence. The user requirement and availability analysis is performed at this stage to assign first time allocation.

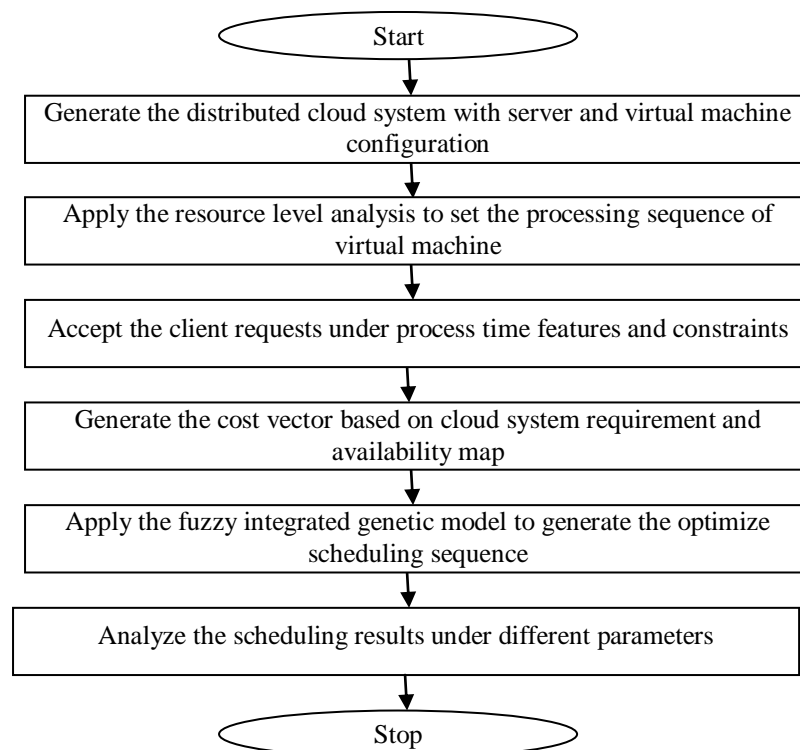


Figure 1: Flow Chart

In the final stage of this model, the cost driven analysis is performed. This analysis is based on the dead line, process time and other requirement and availability constraints. This cost analysis dataset is considered as the population dataset for the genetic process. The fuzzy integrated genetic process is applied in this stage to generate the optimized sequence. The genetic process is divided in three main work stages. At first the possible random sequences are selected. The fuzzy integrated crossover is applied to generate the next optimized sequence. The generated sequence is also changed randomly under mutation operation. Based on this genetic process, the optimized sequence is obtained from the work. The final generated sequence is considered as the sequence of user process execution on cloud server. The process stages associated to the work are shown in figure 1. Figure 1 is showing the complete process flow to configure

the cloud system, allocate the user requests on cloud system and to decide the scheduling sequence. In this work, the fuzzy integrated genetic model is defined to decide the cloud sequence.

IV. RESULTS

To apply the cloud scheduling, the first work is to setup the server and client side parameters. The cloud environment configuration is required on server side with relative resource constraint specification. The parameters considered in the experimentation is given here under

Table 1: Experimentation Parameters

Parameter	Values
Number of Cloud Servers	3
VMs on each cloud server	5
Max Load	5
IO Limit	5
Memory Limit	64M
Processing Limit	1000sec
Simulation Time	100 sec
Number of User Requests	10

Table 1 is showing the experimental parameters considered to apply the scheduling on a dummy environment. The table is showing the server side configuration and the client side constraints based on which the scheduling is implied. The number of cloud servers, virtual machines, memory capability, IO capabilities of cloud system are defined in this table. The request is performed for 10 users. The simulation time is 100 sec and the process time is 1000 sec. The requests are performed in this composed environment so that the effective cloud service allocation will be done using genetic approach.

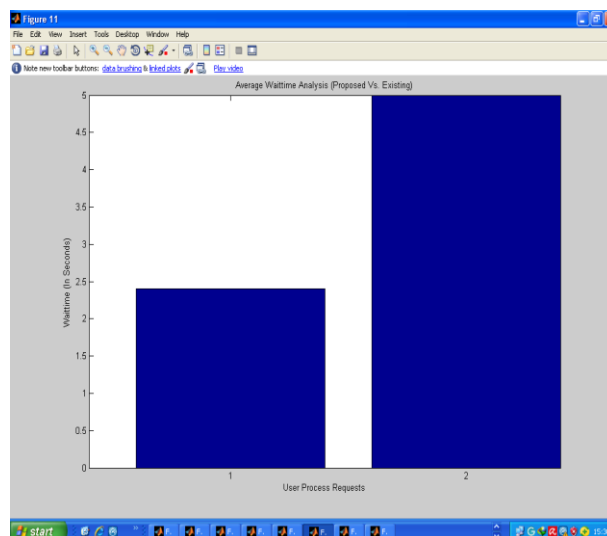


Figure 2: Average Wairtime Time Analysis

Figure 2 is showing the comparative analysis between the existing and proposed approach using average wait time parameter. Here, x axis is showing the user request and y axis is showing the average wait time in second. The first bar is here showing the average wait time of proposed approach and second bar is showing the wait time of existing approach. The figure shows that the proposed method has reduced the average wait time extensively. The figure shows that the effectiveness of process execution is improved.

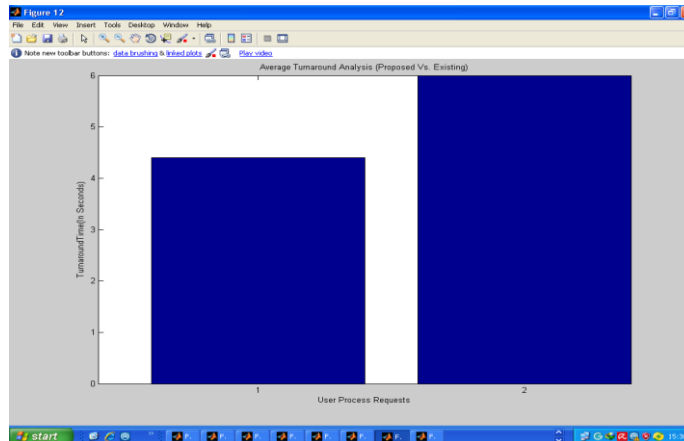


Figure 3: Average Turnaround Time Analysis

Figure 2 is showing the comparative analysis between the existing and proposed approach using average turnaround parameter. Here, x axis is showing the user request and y axis is showing the average turnaround time in second. The first bar is here showing the average turnaround time of proposed approach and second bar is showing the turnaround time of existing approach. The figure shows that the proposed method has reduced the average turnaround time extensively. The figure shows that the effectiveness of process execution is improved.

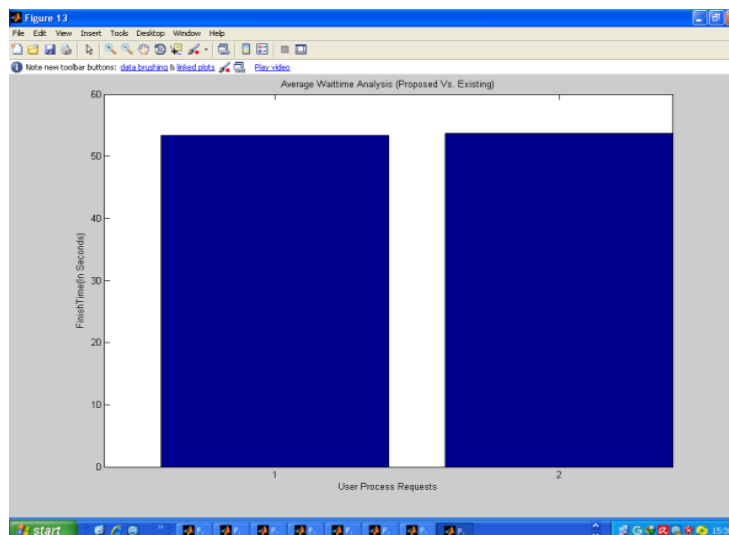


Figure 4: Average Finish Time Analysis (Proposed Vs. Existing)

Figure 4 is showing the comparative analysis between the existing and proposed approach using average finish time parameter. Here, x axis is showing the user request and y axis is showing the average finish time in second. The first bar is here showing the average finish time of proposed approach and second bar is showing the average finish time of existing approach. The figure shows that the proposed method has reduced the average finish time extensively. The figure shows that the effectiveness of process execution is improved.

CONCLUSION

The proposed work model is defined as a three stage model. In first stage, the cloud system is processed under physical constraints. In this layer, the virtual machines are processed to setup the processing sequence under the capabilities analysis. In second layer, the user requests are processed based on the requirement criteria. The middle layer architecture is implied in this stage to perform requirement to availability map. The service allocation is performed at this stage. Based on this analysis, the cost is decide under the deadline, wait time criticalities. This cost generated vector is considered as the population vector for genetic process. The genetic process has defined the process execution sequences randomly and applied the fuzzy integrated crossover to identify the optimized sequence. The model is defined to reduce the wait time and to reduce the number of migrations. The work is implemented on random scenario in matlab environment. The simulation results shows that the model has reduced the number of migrations. The

comparative results shows that the proposed method has reduced the wait time, turnaround time and finish time effectively.

REFERENCES

References are important to the reader; therefore, each citation must be complete and correct. There is no editorial check on references; therefore, an incomplete or wrong reference will be published unless caught by a reviewer or discussor and will detract from the authority and value of the paper. References should be readily available publications. List only one reference per reference number. If a reference is available from two sources, each should be listed as a separate reference. Give all authors' names; do not use et al. Samples of the correct formats for various types of references are given below. For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

- [1]. J. Lu, "Improved Genetic Algorithm-Based Resource Scheduling Strategy in Cloud Computings" 2016 International Conference on Smart City and Systems Engineering (ICSCSE), Hunan, 2016, pp. 230-234.
- [2]. M. Yingchi, X. Ziyang, P. Ping and W. Longbao, "Delay-Aware Associate Tasks Scheduling in the Cloud Computing," 2015 IEEE Fifth International Conference on Big Data and Cloud Computing, Dalian, 2015, pp. 104-109.
- [3]. A. Makwe and P. Kanungo, "Scheduling in cloud computing environment using analytic hierarchy process model," 2015 International Conference on Computer, Communication and Control (IC4), Indore, 2015, pp. 1-4
- [4]. X. Ye, J. Liang, S. Liu and J. Li, "A Survey on Scheduling Workflows in Cloud Environment," 2015 International Conference on Network and Information Systems for Computers, Wuhan, 2015, pp. 344-348
- [5]. A. Mittal and P. Kaur, "Existing Computational Models and Heuristic Techniques for Cloud Scheduling Problems," 2015 Second International Conference on Advances in Computing and Communication Engineering, Dehradun, 2015, pp. 688-692
- [6]. A. Alahmadi, D. Che, M. Khaleel, M. M. Zhu and P. Ghodous, "An Innovative Energy-Aware Cloud Task Scheduling Framework," 2015 IEEE 8th International Conference on Cloud Computing, New York City, NY, 2015, pp. 493-500.
- [7]. S. Sujana and R. K. Devi, "A batchmode dynamic scheduling scheme for cloud computing," 2015 Global Conference on Communication Technologies (GCCT), Thuckalay, 2015, pp. 297-302
- [8]. S. K. Panda and P. K. Jana, "An efficient task scheduling algorithm for heterogeneous multi-cloud environment," 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), New Delhi, 2014, pp. 1204-1209
- [9]. A. Chowdhury and P. Tripathi, "Enhancing cloud computing reliability using efficient scheduling by providing reliability as a service," 2014 International Conference on Parallel, Distributed and Grid Computing, Solan, 2014, pp. 99-104.
- [10]. M. K. Nivodhini, K. Kousalya and S. Malliga, "Algorithms to improve scheduling techniques in IaaS cloud," 2013 International Conference on Information Communication and Embedded Systems (ICICES), Chennai, 2013, pp. 246-250
- [11]. Massimiliano de Leoni and Wil M. P. van der Aalst. 2013. Data-aware process mining: discovering decisions in processes using alignments. In Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC '13). ACM, New York, NY, USA, 1454-1461, 2013
- [12]. M. R. Islam and M. Habiba, "Dynamic scheduling approach for data-intensive cloud environment," 2012 International Conference on Cloud Computing Technologies, Applications and Management (ICCCTAM), Dubai, 2012, pp. 179-185.
- [13]. Y. Wei and L. Tian, "Research on cloud design resources scheduling based on genetic algorithm," 2012 International Conference on Systems and Informatics (ICSAI2012), Yantai, 2012, pp. 2651-2656.
- [14]. M. Mahrishi and A. Nagaraju, "Optimizing cloud service provider scheduling by using rough set model," 2012 International Conference on Cloud Computing Technologies, Applications and Management (ICCCTAM), Dubai, 2012, pp. 223-228
- [15]. W.M.P. van der Aalst, "Process Mining and Verification of Properties: An Approach based on Temporal Logic", ResearchGate, the professional network for scientists, pp 1-18, 2005