_____

Score Based Budget Constraint Workflow Scheduling Algorithm for Cloud System

Makhan Singh

Computer Science and Engineering,
University Institute of Engineering & Technology,
Panjab University, Chandigarh, India.
singhmakhan@pu.ac.in

Abstract—Cloud Computing is the technology that provides on demand services and resources like storage space, networks, programming language execution environment on the top of Internet using pay as you go model. The concept of Cloud Computing emerging as a latest model of service provisioning in distributed system encourage researchers to investigate its advantages and drawbacks in executing scientific applications involving workflows. Workflow scheduling is one of the major issues in Cloud environment that maps and allows the execution of inter-dependent tasks on different resources. It allocates desire resources to workflow tasks so that the execution can be completed to satisfy the QoS constraint defined by the users. At present, the workflow scheduling algorithms only focus on certain QoS parameters which are mainly execution cost and execution time during the allocation of virtual machines to workflow applications. Sometimes resources (virtual machines) are unreliable at data centers. These resources frequently results into failure when workflow applications are scheduled on these resources. The user workflow application may contain sensitive data that cannot tolerate failure of resources on which it is scheduled. In this paper the problem of workflow scheduling that is based on the concept of score is presented. A score based budget constraint workflow scheduling algorithms has been design and simulated. These algorithms reduce the execution time and failure rate of workflow applications within user specified budget.

Keywords-*workflow scheduling; score; virtual machine; budget constraint; cloudlet.*

_____*****_____

I. INTRODUCTION

Cloud Computing is the paradigm that uses the central severs and internet to host the data and resources. Cloud Computing is the new and emerging trend in distributing computing that facilitate software applications, platform, and hardware infrastructures as a service. Data and software packages are hosted on servers.

Workflow scheduling is one of the major issues in workflow management especially in Cloud and Grid workflow systems. It is a process that maps and manages the execution of workflow tasks on different distributed resources. It allocates required resources to suitable workflow tasks such that complete execution can be completed to satisfy constraints defined by the users. Proper scheduling can improve the performance of the system. However, in general, the mapping of tasks on the distributed resources is NP-hard problem. For such problems, no known algorithms are able to generate the optimal solution within polynomial time. Even then the workflow scheduling problem can be solved by using exhaustive search. In this case problem solving complexity is very high. In Grid Computing decisions about scheduling the tasks is taken in the shortest time possible, because number of individuals demand for resources, and time slots needed by one user could be taken by another user at any moment [1]. This equally applies to Cloud Computing environments.

Workflow scheduling is one of the major issues in workflow management especially in Cloud and Grid workflow systems. Workflow scheduling is the problem of mapping of each task on suitable resources while satisfying the constraints imposed by the user. Proper workflow scheduling can have significant impact on the performance of the workflow management system. Workflow scheduling in cloud is divided into two main categories. First category is best effort based scheduling algorithms which tries to optimize the execution time and ignoring the other factors such as execution cost and other QoS constraints. Second one is QoS constraint based scheduling algorithms which tries to optimize the performance under QoS constraints for e.g. cost minimization under the deadline constraint or time minimization under the budget constraint. There are many algorithms proposed by different researchers for workflow scheduling in Clouds like PCP (partial critical path) [2] [3], DBD-CTO (Deadline and Budget distribution based Cost-Time Optimization) [4], Genetic [5], SCFP (shortest cloudlet to fastest processor) [6], HCOC (Hybrid Cloud Optimized Cost scheduling algorithm) [7], etc. that schedule the workflow application under the QoS constraints. Every algorithm has its own strategy while scheduling workflow application. Some algorithms assign the resources to workflow application based on the priority of the hardware resources or tasks. While assigning resources to workflow application the capability of hardware resources and performance requirement should be considered. Hence in this paper the score concept [8] is used to measure hardware capability and to calculate minimum performance requirement for workflow tasks execution. Microsoft is using the WINSAT (Window System Assessment Tool) module in Windows to measure the hardware capabilities it is running in terms of WEI (Windows Experience Index) score [8]. The score concept can be used in Cloud environment to measure the capabilities of hardware resources as score represents the minimum performance of the system and minimum threshold performance requirement of the workflow tasks. This score based scheduling will allow to schedules the workflow applications only on high performance reliable machines which will lead to reduced failure of workflow applications and

_____

satisfy the users imposed budget constraint with minimum execution time. The score based workflow scheduling algorithms find the schedule that will meet the user defined budget constraint and will provide the reliable machines for workflow execution.

II.   SCORE BASED WORKFLOW SCHEDULING ALGORITHM

The proposed scheduling algorithms are based on the concept of score [8]. In proposed scheduling algorithms Virtual machine (VM) score is defined on the basis of capability of the machine. In which each hardware component of the machine receives an individual score and final score of the machine is determined by lowest sub score of the component. Final score represents minimum performance of the machine, which is based on the capabilities of different parts of the machine including processing power, storage space and RAM. Similarly workflow tasks score is defined on the basis of number of instructions in workflow tasks. Score based algorithms looking for machines whose final score is equal to or more than the tasks score and execute the tasks on those machines within user specified budget. In proposed work score based Min-Min budget constrained workflow scheduling algorithm has been designed. Then performances of this algorithm has been compared with basic Min-Min budget constrained workflow scheduling algorithm with respect to their execution time, execution cost and failure rate.

A. Score Based Min-Min Budget Constrained Workflow Scheduling Algorithm

Min-Min Scheduling Algorithm [9] starts with set of unassigned workflow tasks. All unassigned tasks are scheduled on those available resources which will execute them in minimum time. This procedure is followed until all the workflow tasks are assigned to the machines. Here Min-Min scheduling algorithm has been designed with budget constraint i.e. tasks are executed within user specified budget. Score concept has also been introduced and only those machines are selected for scheduling which satisfy minimum task score. Then performance of score based Min-Min budget constrained scheduling algorithm has been compared with basic Min-Min scheduling algorithm with respect to their execution time, execution cost and failure rate.

Step by step description of basic Min-Min budget constrained workflow scheduling algorithm is presented below:

Step 1: Get the list of workflow tasks for scheduling over VM resources.
Step 2: Get all available VMs (Virtual Machines) from the data center.
Step 3: Impose user defined budget constraint B on the entire workflow application.
Step 4: Tasks are prioritized from smallest to largest in the task list and VMs are prioritized from largest to smallest (Mips) in the VMs list.
Step 5: While the task list has the tasks to schedule, repeat steps 6 and 7.
Step 6: Select VM from the list and executes the task within specified budget. If VM executes the task within task budget then task is assigned to VM otherwise next VM is selected from the list of resources.
Step 7: Select next task from the list. If all tasks are scheduled then return their mapping i.e. tasks to VM allocation.
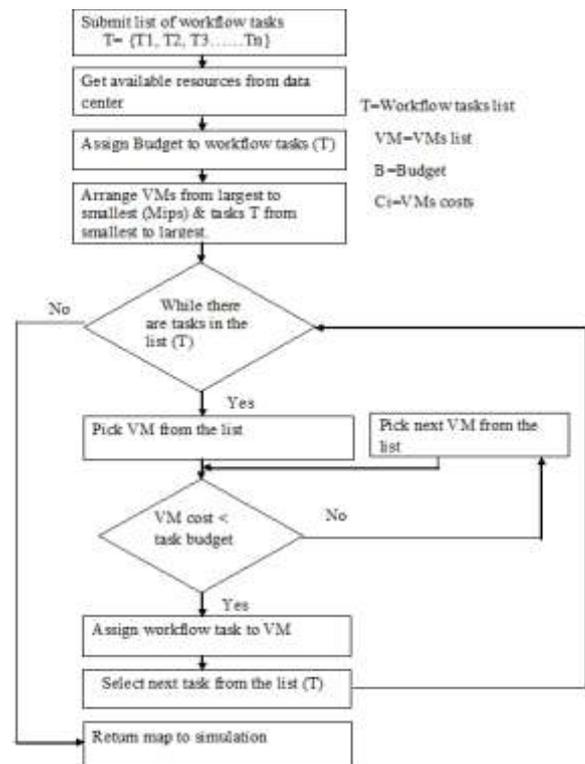

Figure 3.1 Basic Min-Min Budget Constrained Workflow Scheduling Algorithm

Step by step description of score based Min-Min budget constrained workflow scheduling algorithm is presented below:

Step 1: Get the list of workflow tasks for scheduling over VM resources.
Step 2: Get all available VMs (Virtual Machines) from the data center.
Step 3: Impose user defined budget constraint B on the entire workflow application.
Step 4: Obtain the final scores (Sv) of VMs from the components minimum sub scores.
Step 5: Score (St) of workflow tasks is generated on the basis of their instruction length. More is the complexity of the task, more is the task score.
Step 6: Tasks are prioritized from lowest to highest task score (St) in the task list and VMs are prioritized from highest to lowest score (Sv) in the VMs list.
Step 7: While the task list has the tasks to schedule, repeat steps 8 and 9.
Step 8: Select VM from the list which satisfies the task score and executes the task within specified budget. If VM executes the task within task budget then task is assigned to VM otherwise next VM is selected from the list of resources.
Step 9: Select next task from the list. If all tasks are scheduled then return their mapping i.e. tasks to VM allocation.

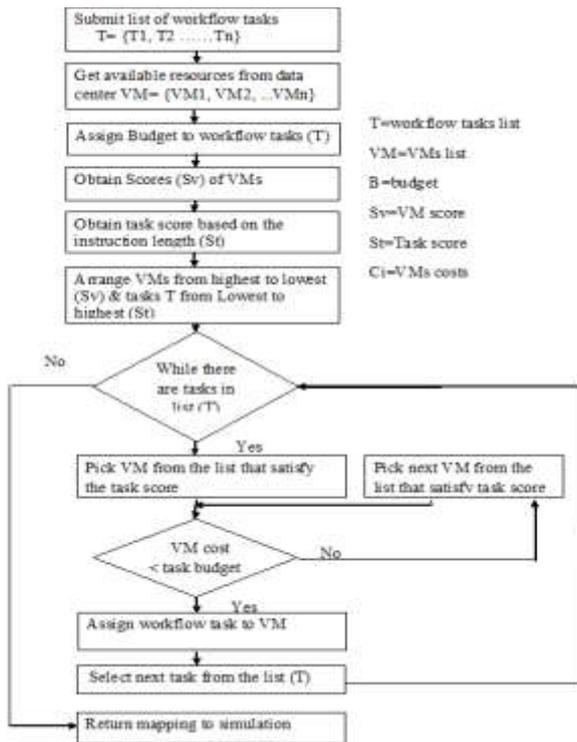**104**

_____

Figure 3.2 Score Based Min-Min budget Constrained Workflow Scheduling Algorithm

III. RESULTS AND PERFORMANCE ANALYSIS

The software that has been used for the simulation of scheduling algorithm are Window 7, Java (JDK-6) , Eclipse-JUNO version and CloudSim-3.0 [10]. The parameters for Cloud Simulator are set as per Table 5.1

TABLE I.   CLOUD SIMULATOR PARAMETERS

Type	Parameter	Value
Datacenter	Number of Datacenter	1
	Number of Hosts	4
	Type of Manager	Time-Shared
Virtual Machine(VM)	Total Number of VMs	30
	MIPS of PE(processing element)	1000-21000
	Number of PEs per VM	1
	VM Memory(RAM)	256-2048 MB
	Storage Space	1-21 GB
	Score	1-10
Cloudlet	Number of Workflow Application	1
	Number of Cloudlets(Tasks)	5-25

In the experimental results performance of score based Min-Min budget constrained workflow scheduling algorithm has been compared with basic Min-Min budget constrained workflow scheduling algorithm with respect to their execution time, execution cost and failure rate.

This test case shows the effect on execution time, execution cost and failure rate of Min-Min budget constrained workflow scheduling algorithm by varying the number of cloudlets, user specified budget and varying the number of iterations when it is implemented with score and without score concept. Here cloudlet refers to workflow application tasks.

TABLE II.   SHOWING EXECUTION TIME AND NUMBER OF CLOUDLETS

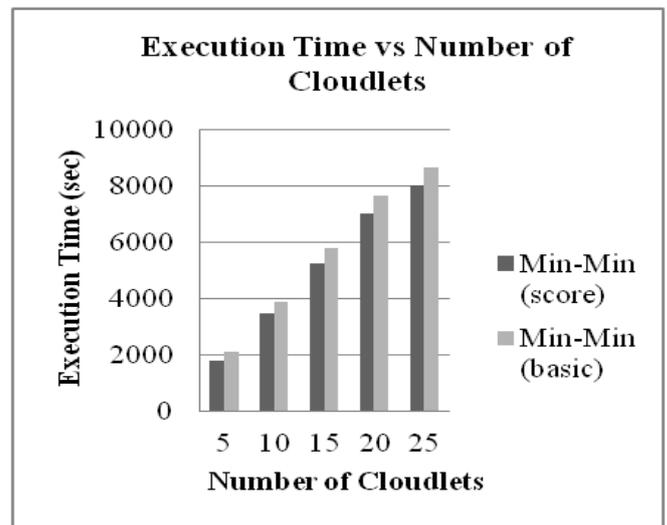| Min-Min Budget Constrained Scheduling Algorithm		
Number of Cloudlets	Execution Time	
	Min-Min (score)	Min-Min(basic)
5	1800	2130
10	3456	3860
15	5035	5790
20	7030	7656
25	8238	8670

Figure 5.1 Showing the Execution Time vs. Number of Cloudlets

Simulation results shown in Figure 5.1 indicate that score based Min-Min budget constrained workflow scheduling algorithm exhibit less execution time for the workflow application as compared to basic Min-Min budget constrained workflow scheduling algorithm.

TABLE III.   SHOWING EXECUTION COST AND USER BUDGET

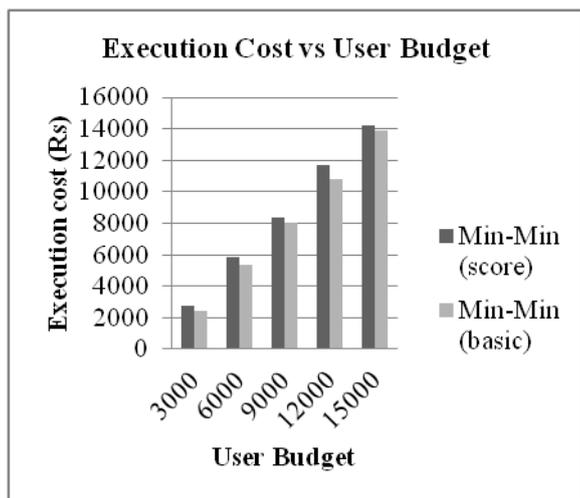| Min-Min Budget Constrained workflow Scheduling Algorithm		
User Budget	Execution Cost	
	Min-Min (score)	Min-Min(basic)
3000	2730	2450
6000	5832	5324
9000	8375	8016
12000	11700	10800
15000	14230	13890

_____

_____

Figure 5.2 Showing the Execution cost vs. User Budget

Results obtained in Figure 5.2 indicate that score based Min-Min budget constrained workflow scheduling algorithms, although incur more cost compared to their basic counterparts, but still execute the workflow application within user specified budget

TABLE IV.   SHOWING FAILURE RATE AND NUMBER OF ITERATIONS

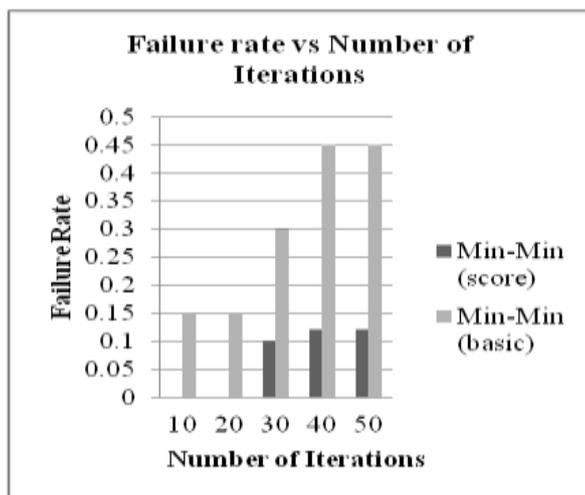| Min-Min Budget Constrained Workflow Scheduling Algorithm		
Number of Iterations	Failure Rate	
	Min-Min (score)	Min-Min(basic)
10	0	0.15
20	0	0.15
30	0.1	0.3
40	0.12	0.45
50	0.12	0.45

Figure 5.3 Showing Failure Rate vs. Number of Iterations

Result obtained in Figure 5.3 shows score based Min-Min budget constrained workflow scheduling algorithm exhibit less failure rate as compared to basic Min-Min budget constrained workflow scheduling algorithm.

IV.   CONCLUSION

In this paper, score based Min-Min budget constrained workflow scheduling algorithm is proposed. This algorithm schedule the workflow application within user specified budget constrained. The performance of score based Min-Min budget constrained workflow scheduling algorithm has been compared with basic Min-Min budget constrained workflow scheduling algorithm on the basis execution time, execution cost, failure rate. Simulation results show that the new developed Score based Min-Min budget constrained workflow scheduling algorithm is more efficient than basic Min-Min budget constrained workflow scheduling algorithm. Score based workflow scheduling algorithms not only reduce the execution time of workflow application but also reduce the failure rate of workflow applications within user specified budget constraint.

REFERENCES

[1]   J. Yu and R. Buyya, Workflow Schdeduling Algorithms for Grid Computing, Technical Report, GRIDS-TR-2007-10, Grid Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, May 2007.

[2]   M. Naghibzadeh, S. Abrishami, "Deadline-constrained workflow scheduling in software as a service Cloud," Science Direct, vol. 19, no. 3, pp. 680-689, November 2011.

[3]   M.Naghibzadeh, S.Abrishami, "Deadline-constrained workflow scheduling algorithms for Infrastructure as a Service Clouds," Eslevier, vol. 29, pp. 158-169, 2012.

[4]   Sakshi Kaushal, Amandeep Verma, "Deadline and Budget Distribution based Cost- Time Optimization Workflow Scheduling Algorithm for Cloud," IJCA, no. 7, 2012

[5]   Jia Yu , R.Buyya, "Scientific Workflow Applications with Deadline and Budget Constraints using Genetic Algorithms," ACM, vol. 14, no. 3,4, pp. 217-230, December 2006

[6]   Amandeep Verma, Shaminder Kaur, "An Efficient Approach to Genetic Algorithm for Task Scheduling in Cloud Computing Environment," IJITCS, no. 10, pp. 74-79, October 2010.

[7]   Luiz Fernando, Bittencourt Edmundo, Roberto Mauro Madeira, "A Cost Optimization Algorithm for Workflow Scheduling in Hybrid Clouds," Springer, vol. 2, no. 3, pp. 207-227, December 2011.

[8]   "Windows System Assessment Tools", http://www.en.wikipedia.org/wiki/ Windows_System_Assessment_Tool.

[9]   Nayandeep sran, Navdeep Kaur, "Comparative Analsis of Existing Load Balancing Techniques in Cloud Computing," IJESI, vol. 2, January 2013.

[10]   Rajiv Ranjan, Anton Beloglazov, César A. F. De Rose and Rajkumar Buyya Rodrigo N. Calheiros, "Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, César A. F. DeCloudSim: A Toolkit for Modeling and Simulation of Cloud Computing," Wiley, vol. 41, no. 1, pp. 23-50, January 2011.

_____