# A REVIEW ON SCHEDULING ALGORITHMS FOR WORK-FLOW APPLICATION IN CLOUD COMPUTING

Er. Gagandeep Kaur[1] & Er. Manoj Agnihotri[2]

**Abstract –Many scalable resources, IT resources and data storage facilities are available on the Internet with the help of very big computing paradigm CLOUD COMPUTING. In the IT market, business world, educational departments, governments organization everywhere Cloud Computing is gaining the popularity because of its captivating features. Nowadays the current research topic in cloud computing is task scheduling problem. This paper mainly Focus on certain proposed scheduling algorithms for cloud computing environment whose main intent is to meet the deadlines and to lessen the entire execution costs of work-flow application.**
**Keywords – Cloud Computing, IS, DS, DAG, WfMC, CC**

## I. INTRODUCTION

Cloud Computing is a Internet based computing which helps in separating the application from the hardware. According to the National Institute of Standards and Technology (NIST) : "Cloud computing is an emerging technology for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [10]. CC works on the Pay as you go basis which means that the user has to pay to Cloud Service Provider only for what they are using according to their requirements.

### 1.1 Virtualization

Virtualization is the component of Cloud Computing. Virtualization means that on one particular computer a user can have a multiple instance of virtual machines that have different execution environment. Virtualization provides a layer on hardware on which one can install the different instances and that is implemented with the help of hypervisor. Virtualization hypervisor technologies like XEN, KVM and VMware have been build them as an essential part of cloud computing environment.

### 1.2 Scheduling

A technique which helps in the mapping and also the scheduling of the tasks of work-flow application on distributed resources is called the Scheduling. The scheduling is mainly used to allot the suitable resources to workflow task so that execution of work-flow tasks finishes within the deadline stated by the customer. The task scheduling in the cloud computing can be classified in two types:

*A)   Independent Scheduling*

In IS, dependency of relationship between the task is not there due to which task can be run independently. As compared to workflow scheduling, independent scheduling is quite easy.

*B)   Dependent Scheduling*

Dependency means that after the execution of parent task, child task can be executed In DS, Tasks of workflow application are dependent on each other .

## II. WORK-FLOW SCHEDULING

Work-flow Scheduling focus on the arrangement and managing of interdependent task of work-flow application on available virtual machines in such a method that the user specified goal can be achieved. The work-flow enables the structure of application as DAG (Directed Acyclic Graph) where nodes defines the task of workflow and set of edges between the nodes represent the dependencies between the tasks.

---

[1] *M.Tech Research Scholar, Department of Computer Science and Engineering, Amritsar College of Engineering and Technology, Amritsar, Punjab, India*
[2] *Assistant Professor, Department of Computer Science and Engineering, Amritsar College of Engineering and Technology, Amritsar, Punjab, India*
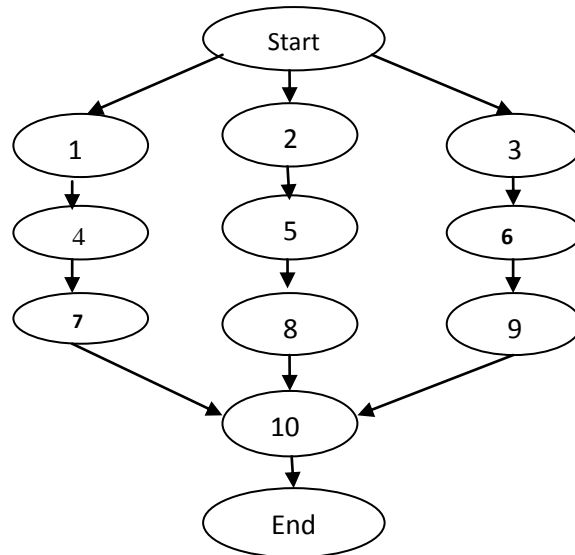
*Figure1. Directed acyclic graph of Epigenomics Workflow application*

Work-flows are sustained by Workflow Management. Workflow Scheduling plays a indispensable role both in workflow management and in the cloud computing field. Work-flow Scheduling is a great concern in the field of computing.

### III. WORK-FLOW MANAGEMENT

The WfMC (Workflow Management Coalition) defined a workflow as "The automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules."[5]. WfMC published its reference model (Fig 2) which identifies the interface who enabled products to interoperate at a variety of levels [2]. The workflow management and its interface is being enhanced with the model.
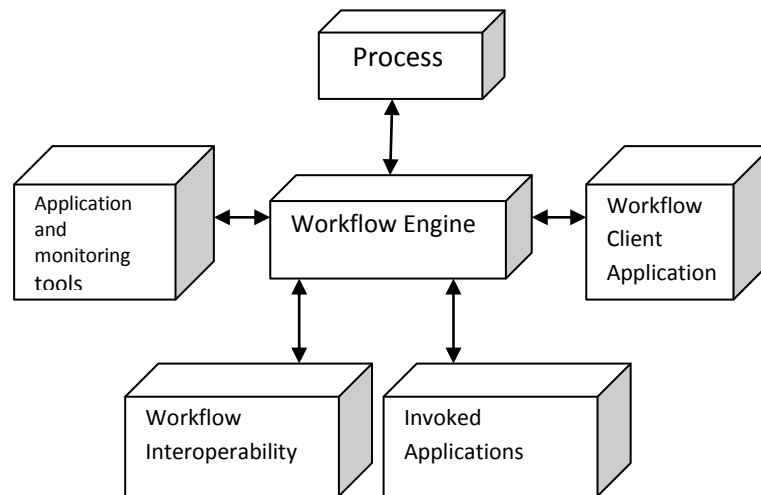


*Fig2 .WfMC Workflow Reference Model*

Various Components are as follows

1. Workflow Engine: It is a kind of software service which helps in allocating the run- time environment for creating, managing and executing the workflows.
2. Process: It is a delegacy of workflow process which supports the automated manipulation [5].
3. Workflow Interoperability: It acts an interface that supports the interoperability between various workflow systems.
4. Invoked Application : It is a interface to support different other IT applications
5. Workflow Client Application : It is used for the interaction between the users
6. Monitoring: Interface to provide system monitoring and metric function to provide management of composite workflow application environment.

## IV. VARIOUS SCHEDULING ALGORITHMS

1.   Critical path based Scheduling Algorithm: Critical Path based Scheduling Algorithm for Workflow Application in cloud Computing [1] This paper proposed the algorithm (finding critical path task) which is used to for finding the critical and noncritical path by comparing the value of early finish time , early start time , late finish time , late start time . Resource algorithm is used to schedule the tasks on available resources. Virtual machines are allocated to task based on cost .The experiment results shows that in case of cost reduction, the proposed algorithm performs better than the heuristic algorithm.

2.   Deadlines and Cost Based Ant Colony Optimization Algorithm for Scheduling Workflows Application in Hybrid clouds [6] This Paper proposed the Ant Colony Optimization algorithm that is utilized in order to schedule the work-flow application in hybrid cloud (Combine Private and Public Cloud). If the user Defined deadlines are not achieved in private cloud, then the sub-deadlines are allotted to the particular tasks of work-flow applications and are migrated to the private cloud to finish the execution. The proposed algorithm when compared to genetic algorithm, max-min algorithm, min-min algorithm on the basis of execution is better while assigning tasks whose deadlines were skipped at private cloud to public cloud.

3.   Deadline and Cost Based workflow Scheduling in Hybrid Cloud [14] Within this Paper, the proposed Algorithm helps in taking the best resource from the public cloud on lease so that  work-flow execution should be completed in the given deadlines by user .The Cost Optimization is obtained by the proposed algorithm.

4.   Critical Path based Iterative Heuristic for workflow scheduling in utility and cloud computing [15] The work-flow scheduling issue in utility and computing is being taken into account in this paper. This paper considers allowance of tasks to appropriate resources in a effort to reduce the overall rental price of most resources while keeping the precedence limits as well as acquiring the work-flow deadlines . The Critical Path based Iterative(CPI) Algorithm proposed by this paper when compared by to already existing Partial Critical Path (PCP) based algorithm attains the 20.7 % reduction  in the average normalized resource renting price for instance with one thousand activities .

5.   Scheduling Workflows with Budget constraints [22]This paper focus on a model for workflow applications that are modeled within the kind of DAG's (Directed Acyclic Graph) .In this paper they have implemented the LOSS and GAIN ,two approaches which helps in solving the scheduling problem under the budget constraints. For the experiment result s, the proposed algorithm in this paper is being run 100 times with the four different types of DAGs, which have 100 nodes each and are scheduled on 3 different machines

## V. RELATED WORK

**Lovejit Singh et al. (2014) [6]** proposed the combination of Private Cloud and Public Cloud forms a hybrid Cloud . This Cloud is useful for Organization to scale the resources when demand increase .The Scheduler or Broker is used for allocating the resources to task of workflow application. The authors of this paper have proposed the Ant Colony Optimization (ACO) Algorithm which helped in obtaining the objective of both cloud provider and cloud user .The authors have used the proposed algorithm in this paper to find the schedule for workflow application at private cloud and bind resources to task while achieving the user defined deadlines.

**Amandeep Verma et al. (2014) [7]** proposed Bi-Criteria Priority based Particle Swarm Optimization (BPSO) in order to arrange the work –flow jobs above the accessible cloud sources which decreased the particular rendering price as well as rendering time frame within the provided due date and fiancés constraints. The suggested algorithm criteria are analyzed using simulator within various real world work-flow programs.

**Sarabjeet Singh et al. (2013) [8]** discussed that when the private cloud while using the confined power might not be ready to implement the important work-flow, then there exist a need of public cloud. The execution of whole workflow application is considered as a costly process in public Cloud. Hybrid cloud is solution which is a combination of Public Cloud and Private Cloud. The task Scheduling is the main focus point of Hybrid Cloud. The authors have proposed the hybrid scheduling algorithm that utilized the very idea of sub deadline for rescheduling along with the allotment of resources .The authors have presented the idea of executing the heavier task on private resources and then lighter task on public cloud. The authors evaluated three rescheduling policies used to send the task to public cloud for execution.

**Rodrigo N Calheiros, et al.  (2013) [9]** focused that the Scientific Workflows have certain kinds of characteristics which makes them acceptable for cloud environment. The authors in their paper focused on the Public Cloud as the user as to pay on the basis of their usage in this cloud and it provide the dynamic scaling based on the need of the application. The workflows are subjected to budget and deadlines constraints. The authors discussed that the already existing research in execution of scientific workflow on Cloud can do one of two things either can minimize the Cost while trying to meet deadlines or they can minimize the workflows execution time neglecting the deadlines and budget. In order to reduce the performance variation on

soft deadline of workflow application the authors proposed EIPR Algorithm which used non productive period of resources and replicate the tasks on those resources and took the behavior of cloud into consideration while scheduling.

**Zhou Yue, et al. (2013) [10]** proposed a scheduling technique determined by the ant colony optimization as well as presented two – ways ants procedure. Establishing the pheromone threshold will be used to stay away from untimely pheromone, moreover, having a two –tier research approach and also launching pre-execution period would be to keep away from the nearby optimum in order to allot the job to the greatest powerful computing resources. The particular simulation outcomes reveal that discussed algorithm in this paper certainly reduced plenty of time to search computing resources along with the substantially boost the efficiency.

**Dhinesh et al. , (2013) [11]** suggested algorithm called as honey bees behavior influenced by load balancing (HBB-LB)) that is targeted to attain nicely balanced load all over virtual machines intended for making the most of throughput .The discussed algorithm criteria additionally levels out the main concerns of jobs around the machines in a manner so as the level of waiting around period of jobs in the main line minimal. The tasks that happen to be eliminated from virtual machine are addressed as honey bees that are considered updaters globally. The strategy provider by the researchers explained that there is a considerable betterment within typical execution time and decrease in waiting around duration of tasks on queue. The outcomes showed the algorithms holds great without raising the extra overheads.

**Abrishami Saed , et al. (2013) [12]** have in the past designed and also studied a partial critical path ,PCP that is a two phase scheduling algorithm which aspires to lessen the money necessary for work-flow rendering although meeting a user defined deadlines. Even so , they believed clouds and utility grids are different from each other in 3 forms : on-demand resource provisioning, homogeneous networks, and the pay-as-you-go pricing model. In this paper, researchers have possessed the PCP algorithm for the Cloud environment and proposed two workflow scheduling algorithms:  IaaS Cloud Partial Critical Paths (IC-PCP) that is one phase algorithm, and IaaS Cloud Partial Critical Paths with Deadline Distribution (IC-PCPD2) which is two phase algorithm. Each algorithm has a very polynomial period sophistication rendering them suited selections for preparation large workflows. The simulation outcomes demonstrated that either algorithm possess the encouraging performance with IC-PCP accomplishing much better than IC-PCP2 practically in most cases.

Table1. Comparison Table of Scheduling Algorithm

| Sr.No | Paper | Publication year | Scheduling parameters | Findings | Tools | Environment |
|---|---|---|---|---|---|---|
| 1 | Cririral Path based Scheduling Algorithm for Workflow Application In Cloud Computing | 2016 | Cost of Workflow Applications | Cost Reduction | Workflow Sim | Cloud environment |
| 2 | Critical Path based Iterative Heuristic for workflow Scheduling in utility and cloud computing | 2013 | Deadlines Cost | Finding Critical Path , Reduction of Cost , workflow completion within the deadline | Workflow Sim | Cloud Environment |
| 3 | Scheduling Workflows with Budget constraints | 2007 | Makespan Budget | Reduction in makespan ,minimization of execution time | Amazon EC2 | Cloud Computing |
| 4 | Deadline and Cost Based workflow Scheduling in Hybrid Cloud | 2013 | Cost Deadlines | Cost Optimization Workflow execution in deadlines | | Cloud Environment |

| 5 | Deadlines and Cost Based Ant Colony Optimization Algorithm for Scheduling Workflows Application in Hybrid clouds | 2014 | Cost Deadlines | Optimization of Execution Cost | Cloud Sim | Cloud environment |
|---|---|---|---|---|---|---|

## VI. CONCLUSION

On the internet, the emerging topic is Cloud Computing which helps the users to use certain application provided by Cloud Provider on the pay as you use basis and with this Scheduling the Workflows application in Cloud Computing is also a challenge in today's world. In this review paper, workflow scheduling, workflow management and certain types of existing scheduling algorithms are analyzed.

## REFERENCE

[1] Jailalita, Sarabjeet Singh, Maitreyee Dutta, "Critical Path based Scheduling Algorithm for Workflow Application in Cloud Computing", International Conference on Advances in Computing , Communication and Automation (ICACCA)(Spring),IEEE , September 29 ,2016 .

[2] Anterpreet Kaur, Department of Information Technology, Mohali, Landran, India , "A Review on Workflow Scheduling in Cloud Computing Environment" , International Journal of Computer Science Engineering (IJCSE) ISSN -2319-7323 Vol.4, No.02 ,March 2015.

[3] Fairouz fakh Fakh, Hatem Hadj Kacem, Ahmed Hadj Kacem, "Workflow Scheduling in Cloud Computing: A Survey", IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstration (EDOCW), 8th December 2014.

[4] Santosh B, Harshitha, Prachi Kaneria A, Dr D.H Manjaiah, " Comparative Study of Workflow Scheduling Algorithm in Cloud Computing" , International Journal of Innovative Research in Computer & Communication Engineering, ISSN online (2320-9801) ISSN print (2320-9798) Volume 2, Special issue 5, October 2014.

[5] Murli Manohar Sharma and Anju Bala, Department of CSE, Thapar University, Patiala , Punjab " Survey Paper on Workflow Scheduling Algorithm used in Cloud Computing" International Journal of Information and Computation Technology ,ISSN 0974-2239 Volume 4 , pp 997-1002, Number 10 ,2014.

[6] Lovejit Singh, Sarabjeet Singh, "Deadline and Cost Bases Ant Colony Optimization Algorithm for Scheduling Workflow Application in Hybrid Cloud", Journal of Scientific & Engineering Research, Vol. 5, Issue No 10, October 2014.

[7] Amandeep Verma, and Sakshi Kaushal, "Bi-criteria priority based particle swarm optimization optimization workflow scheduling algorithm for cloud", Recent Advances in In Engineering and Computational Sciences (RAECS), pp. 1-6. IEEE, 2014.

[8] Sarabjeet Singh , Nitish Chopra , " HEFT based Workflow Scheduling Algorithm for Cost Optimization with Deadline in Hybrid Cloud", IEEE 4th International Conference on Computing Communications and Networking Techniques ,Tiruchengode ,Tamil Naidu ppp1-6 , July 2013.

[9] Cost Optimization with Deadline in Hybrid Cloud", IEEE 4th International Conference on Computing Communications and Networking Techniques ,Tiruchengode ,Tamil Naidu ppp1-6 , July 2013.

[10] Rodrigo N Calheiros and Rajkumar Buyya , "Meeting Deadlines of Scientific Workflow in Public Clouds with task Replication", IEEE Transcation on parallel & Distributed systems , Vol. 25, Issue No 7, pp 1787 -1796 , June 2013.

[11] Zhou, Yue, and Xinli Huang, "Scheduling workflow in cloud computing based on ant colony optimization algorithm", Sixth International Conference on Business Intelligence and Financial Engineering (BIFE), pp. 57-61. IEEE, 2013.

[12] LD Dhinesh Babu, and P. Venkata Krishna, "Honey bee behavior inspired load balancing of tasks in cloud computing environments", Applied Soft Computing 13, no. 5 (2013): 2292-2303.

[13] Saed Abrishami, Mahmoud Naghibzadeh, Dick H.J Epema, "Deadline Constrained workflow scheduling Algorithms for Infrastructure as a Service Clouds", Future generation Computer System 29, 158 -169, 2013.

[14] Nitish Chopra, Sarabjeet Singh "Deadline and Cost Based Workflow Scheduling in Hybrid Cloud", IEEE International Conference On Advances in Computing, Communication and Informatics (ICACCI) , Mysore ,pp 840-846 ,August 2013.

[15] Cai, Zhicheng, Xiaoping Li, and Jatinder ND Gupta. "Critical path-based iterative heuristic for workflow scheduling in utility and cloud computing." In International Conference on Service-Oriented Computing, pp. 207-221. Springer, Berlin, Heidelberg, 2013.

[16] Amudha T., and T. T. Dhivyaprabha, "QoS priority based scheduling algorithm and proposed framework for task scheduling in a grid environment" International Conference on Recent Trends in Information Technology (ICRTIT), pp. 650-655. IEEE, 2011.

[17] Suraj Pandey, LinLin Wu, Siddeswara Mayura Guru, Rajkumar Buyya, "A Particle Swarm Optimization Based heuristic for Scheduling Workflow Application in Cloud Computing Environment", 24th IEEE International Conference on Advanced Information Networking and Application, pp. 400-407, 2010.

[18] Simsy Xavier ,S.P. Jeno Lovesum, " A survey of Various Workflow Scheduling Algorithms in Cloud Environment", International Journal of Scientific and Research Publication , Volume 3 ,ISSN 2250-3153, Issue 2, February 2013.

[19] Chai Xue–Zhui, CAO Jian, Shangai Jiao Tong University, Shangai 200240, China, "Cloud Computing Oriented Workflow Technology" , Journal Of Chinese Computer Systems , 2012-01.

[20] Karaboga, Dervis, and Bahriye Basturk. "On the performance of artificial bee colony (ABC) algorithm." Applied soft computing, Volume 8, Issue no. 1 pages687-697, January 2008.

[21] Mustafizur Rahman, Srikumar Venugopal and Rajkumar Buyya , "A Dynamic Critical Path Algorithm for Scheduling Scientific Workflow Application on Global Grids" , E-SCIENCE '07 Proceedings of the Third IEEE International Conference on E-Science and Grid Computing,2007.

[22] Sakellariou, Rizos, Henan Zhao, Eleni Tsiakkouri, and Marios Dikaiakos. "Scheduling workflows with budget constraints". Integrated research in GRID computing (2007): 189-202.

[23] Jia Yung, Rajkumar Buyya, "A taxonomy of Workflow Management System for Grid Computing", Journal of Grid Computing , Springer pp 171 -200, 2005.

[24] P.Mell And T.grance ," The NIST definition of Cloud Computing ," downloaded from http://www.nist.gov/itl/cloud February 10, 2015