Survey of Workflow Cloud Scheduling by Optimizing By Ant Colony Optimization

Ritam Badechha, Rohini Sharma
Computer Science & Engineering
Gurukul Vidyapeeth Institute of Engineering & Technology
(Punjab Technical University)

Absract-- Those days are gone when putting away and getting to of information were done on PC's hard drive. Presently with advancement in innovation and with awesome achievement of Internet Computing assets have turned out to be more efficient, more effective and more pervasively accessible than any other time in recent memory. This mechanical pattern of 21st century has brought forth the acknowledgment of another processing model called Cloud Computing. This Computing is not just about the hard drives where putting away and getting to should be possible yet it is most recent registering worldview and it offers huge chances to take care of vast scale logical issue. To completely abuse the uses of cloud, different difficulties should be tended to where planning is one among them. Albeit catholic research has been done on Workflow Scheduling, there are not very many edges custom-made for Cloud situations. For some fundamental standards of Cloud, for example, flexibility and heterogeneity existing work neglects to meet ideal arrangement. In this way our work concentrates on the booking methodologies for logical work process on IaaS cloud. We display a calculation in view of the meta-heuristic streamlining method where the best of two calculations Ant province Optimization (ACO). which limits the general work process time (makespan) and lessens the cost. Our heuristic is assessed utilizing CloudSim and a few surely understood logical work processes of various sizes. The outcome demonstrates that our approach performs better when contrasted with PSO calculation.

Keywords— *PSO, Cloud, Workflow, Optimization*

I. INTRODUCTION

With progression in innovation, preparing and capacity furthermore with the accomplishment of the Internet, registering assets have ended up less expensive, more capable and more universally accessible than any time in recent memory. As depicted in Fig.1.1 this mechanical pattern has brought forth the acknowledgment of another registering model called cloud computing, in which assets (e.g., CPU and capacity) are given as general utilities that can be rented and discharged by clients through the Internet in an on-demand. In a cloud computing environment, the customary part of service Provider is isolated into two: the infrastructure suppliers who oversee cloud stages and rent assets as per a use based

Estimating model, and administration suppliers, who rent assets from one or numerous framework suppliers to serve the end clients. Substantial organizations, for example, Google, Amazon and Microsoft endeavour to give all the more effective, dependable and cost-proficient cloud stages, and business undertakings try to reshape their plans of action to pick up advantage from this new worldview. [1]

Fig.1. Overview of cloud computing

Characteristics of cloud computing:

Cloud computing have some essential or unique characteristics is to provide qualitative services. These characteristics are as follows [2]

- On-demand self-serviceThis self-service mentions to the service provided by cloud computing vendors that enables the provision of cloud resources on demand whenever they are required. In on-demand self-service, the user accesses cloud services through an online control panel.

- Broad network accessCloud computing separates computing capabilities from their consumers, so that they don't have to maintain the capabilities themselves. A consequence of this is that the computing capabilities are located elsewhere, and must be accessed over a network.
- Resource poolingResource pooling is an Information Technology term used in cloud computing environments to describe a situation in which providers serve multiple clients, customers or "tenants" with provisional and scalable services. These services can be adjusted to suit each client's needs without any changes being apparent to the client or end user.Examples of resources include storage, processing, memory, and network bandwidth [3].

II. SCHEDULING

In computing, scheduling is the strategy by which strings, procedures or information streams are offered access to framework assets (e.g. processor time, correspondences data transmission). This is normally done to load adjust and share framework assets adequately or accomplish an objective nature of administration. The requirement for a planning calculation emerges from the prerequisite for most cutting edge frameworks to perform multitasking (executing more than one procedure at once) and multiplexing (transmit various information streams all the while over a solitary physical channel) [4]. Unlike Grids, Scalability, flexibility reliability of Cloud resources allows real- time processing of resources to meet application requirement. At lower cost services of cloud such as compute, storage, and bandwidth are available. Normally undertakings are scheduled by client prerequisites. New planning methodologies should be proposed to defeat the issues postured by system properties in the middle of client and assets. New booking methodologies might utilize a percentage of the customary planning ideas to consolidation them together with some system mindful techniques to give answers for better and more effective employment booking. Customary path for booking in distributed computing was to utilize the immediate assignments of clients as the overhead application base. The problem in that scheduling was there is no association between the overhead application base and the way that different tasks cause overhead costs of resources in Cloud systems which may incur the cost of Cloud. That is why there is need of scheduling in Cloud Environment so that parallel processing of complex application can be done efficiently [5].

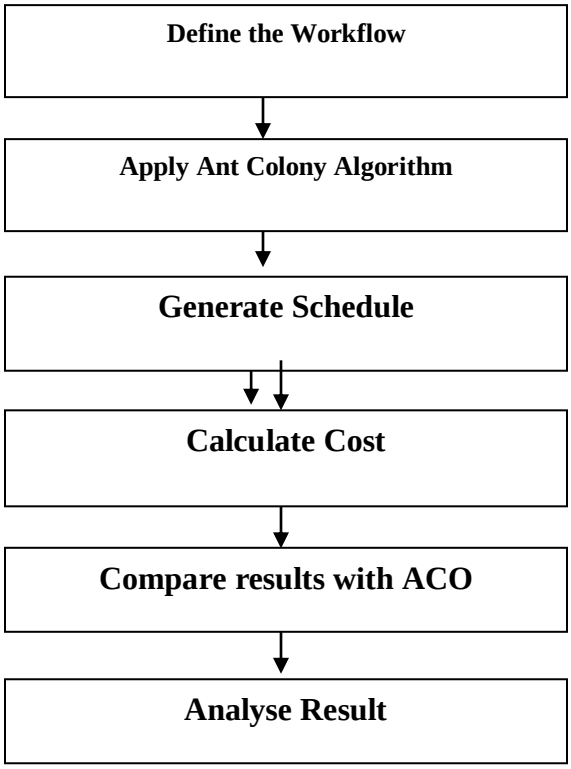
III. LITERATURE REVIEW

Pandey et al. [6] defined that user programs may incur large information retrieval and execution costs when they may be scheduled taking into account only the 'execution time'. further to optimizing execution time, the cost springing up from statistics transfers among resources in addition to execution expenses need to also be taken under consideration. creator offered a particle swarm optimization (PSO) based heuristic to time table applications to cloud sources that takes into account each computation cost and statistics transmission price and evaluate with current 'Best Resource Selection' (BRS) set of rules.Chen et al. [7] proposed an ant colony optimization (ACO) algorithm to schedule massive-scale workflows with numerous QoS parameters. This algorithm permits customers to specify their QoS options as well as outline the minimal QoS thresholds for a sure software. The goal of this set of rules was to find a solution that meets all QoS constraints and optimizes the user-preferred QoS parameter. based totally at the traits of workflow scheduling, author designed seven new heuristics for the ACO technique and proposed an adaptive scheme that lets in synthetic ants to select heuristics based on pheromone values. Byun et al. [8] advised architecture for the automatic execution of huge-scale workflow-primarily based applications on dynamically and elastically provisioned computing resources. Authors centered on its middle set of rules named PBTS (Partitioned Balanced Time Scheduling), which estimates the minimal variety of computing hosts required to execute a workflow inside a user-distinctive end time. The PBTS set of rules was designed to in shape each elastic useful resource provisioning models together with AmazonEC2 and malleable parallel utility models consisting of Map lessen. author verified that PBTS estimates the aid potential close to the theoretical low certain. Malawski et al. [9] addressed the green control of sources under budget and closing date constraints on Infrastructure-as-a-service (IaaS) clouds. They mentioned, developed, and assessed algorithms primarily based on static and dynamic techniques for each project scheduling and aid provisioning and evaluated thru simulation using a set of scientific workflow ensembles with a wide variety of budget and closing date parameters, considering uncertainties in assignment runtime estimations, provisioning delays, and screw ups. also authors decided the performance of an set of rules primarily based on workflow structure and estimates of undertaking runtimes can considerably improve the pleasant of answers. Abrishami et al. [10] designed and 102ehaviou a –segment scheduling set of rules for application Grids, called Partial essential Paths (PCP), which goals to reduce the price of workflow execution at the same time as assembly a person-described closing date. Authors tailored the PCP set of rules for the Cloud environment and recommend two workflow scheduling algorithms: a one-section algorithm that is referred to as IaaS Cloud Partial vital Paths(IC-PCP),and a –section set of rules that's known as IaaS Cloud Partial critical Paths with closing date Distribution(IC-PCPD2).both algorithms have a polynomial time complexity which lead them to suitable alternatives for scheduling huge workflows. IC-PCP performs higher than IC-PCPD 2 in most instances. Xue et al. [11] proposed a QoS-based totally hybrid particle swarm optimization (GHPSO) to schedule packages to cloud sources.

In GHPSO, crossover and mutation of genetic set of rules is embedded into the particle swarm optimization set of rules (PSO), in order that it could play a position within the discrete hassle. A hill hiking algorithm was additionally brought into the PSO that allows you to improve the nearby seek ability and to hold the variety of the populace. The simulation effects show that the GHPSO achieves higher performance than fashionable particle swarm algorithm used in reduce costs inside a given execution time. Rodriguez et al. [12] explained a good way to meet the consumer's best of carrier (QoS) requirements or to include a few simple principles of Cloud computing inclusive of the pliability and heterogeneity of the computing assets, there have to be resource provisioning and scheduling approach for medical workflows on Infrastructure as a service (IaaS) Clouds. They supplied an algorithm based totally at the meta-heuristic optimization approach, Particle Swarm Optimization (PSO), which aims to limit the overall workflow execution value at the same time as assembly closing date constraints. Netjinda et al. [13] focused on optimizing the value of buying infrastructure-as-a-service cloud competencies to attain clinical work goes with the flow execution in the unique closing dates. Authors considered the quantity of purchased times, example types, buying options, and venture scheduling as constraints in an optimization technique. Particle swarm optimization augmented with a variable community seeks approach turned into used to discover the superior solution. Results display promising performance from the views of the total fee and fitness convergence when in comparison with other trendy algorithms. Verma et al. [14] recommended that the users put up their workflows alongside a few QoS constraints like closing date, budget, consider reliability and so on. For computation. Authors considered the two constraints: closing date and finances and recommend cut-off date and finances Deadline and Budget Distribution based cost-Time Optimization (DBD-CTO) workflow scheduling set of rules that minimizes execution value while assembly time frame for handing over consequences and analyse the behaviour of the algorithm.

IV. METHODOLOGY

Step1: Firstly define the workflow.
Step2: Apply the ant colony optimization in workflow.
Step3: In this step, Generate Schedule.
Step4: When schedule is generate then calculate the Makespan.
Step5: When Makespan is calculated then calculated the cost.
Step6: Results are compared with ACO.
Step7: In this step, results are analysed.



V. RESULT

Table 1 GENOME

VM	Average Cost (ACO)	Average Cost (PSO)	Average lastdag (ACO)	Average lastdag (PSO)
two	69.36	99.42	0	2554493.86
four	467.71	629.91	12802.9546	65331.23
six	557.96	843.36	3074148.03	94768.4994
eight	488.03	923.27	3266428.88	116386.9196
ten	431.61	607.14	6172814.14	14241487.42
twelve	436	2111.13	6641350.19	208559.28

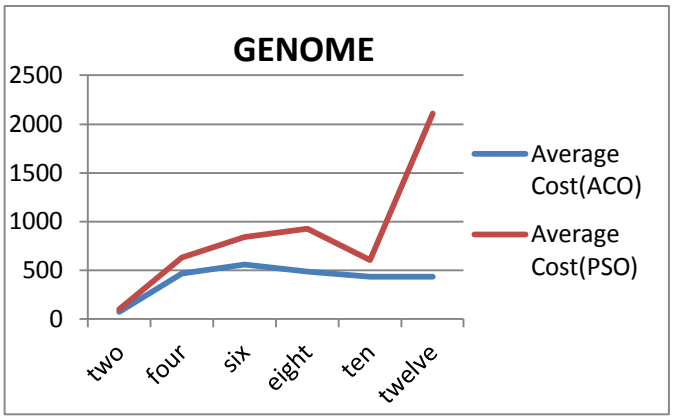


Fig. 4.1 Cost of Genome workflow

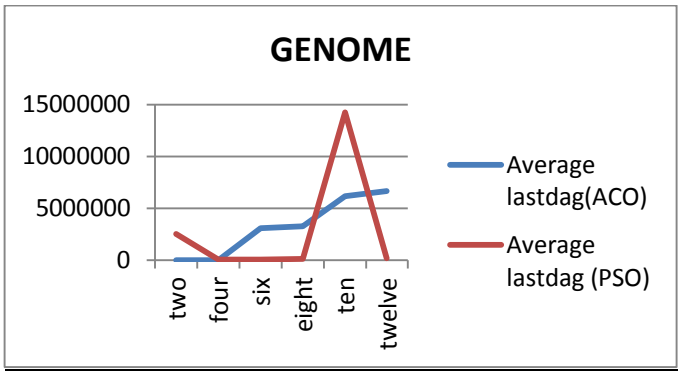


Fig. 4.2Makespan of Genome workflow

TABLE 2 CYBERSHAKE

VM	Average Cost (ACO)	Average Cost (PSO)	Average last dag (ACO)	Average last dag (PSO)
two	1.6	3.18	0	264.4448
four	2.1	7.8	478.1314	875.3528
six	6.5	14.67	524.9606	1610.1014
eight	8.6	18.15	1162.2742	2041.2912
ten	10	16.71	1672.1763	2638.0596
twelve	15	17.19	1655.3824	2688.257

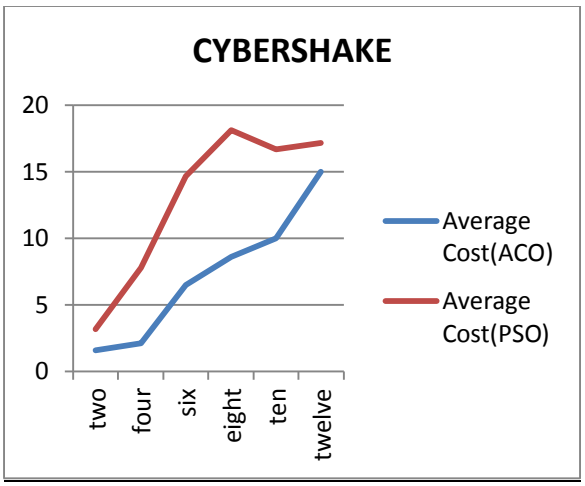


Fig. 4.3 Cost of Cybershake workflow

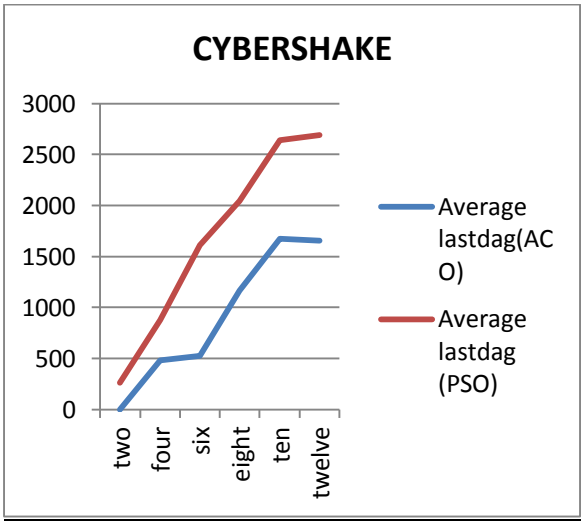


Fig. 4.4 Makespan of Cybershake

Table 3 LIGO

VM	Average Cost (ACO)	Average Cost (PSO)	Average last dag (ACO)	Average last dag (PSO)
two	12	28.53	710.6529	3083.9904
four	15.1	24.75	1862.3855	5049.356
six	35.52	75.78	2616.8577	8679.1234
eight	33.87	101.76	4251.1423	11238.7914
ten	69.34	104.82	4162.4897	13411.725
twelve	68.68	116.01	3584.4682	15795.5116

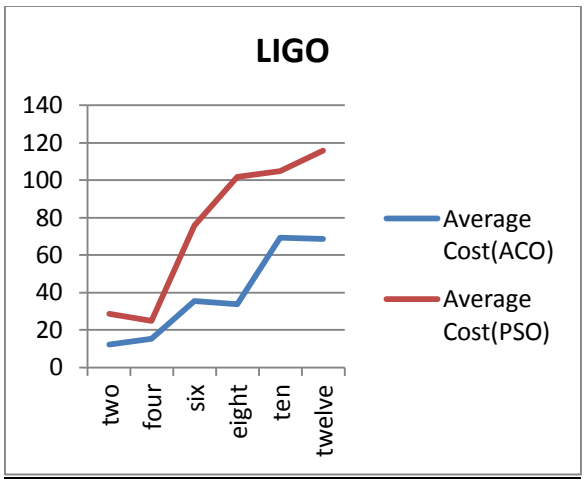


Fig. 4.5 Cost of Ligo workflow

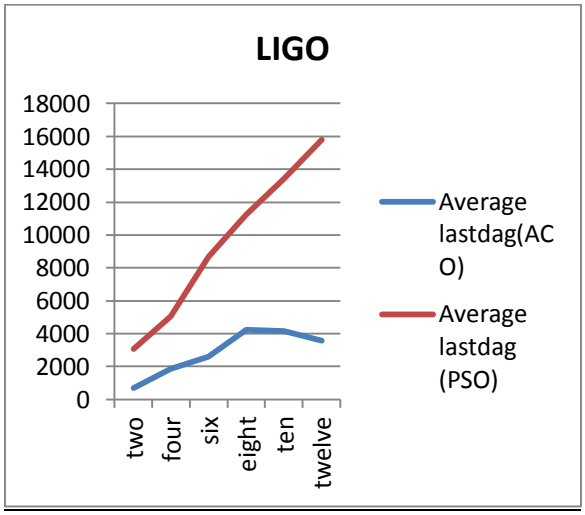


Fig. 4.6 Makespan of Ligo workflow

Table 4 MONTAGE

VM	Average Cost (ACO)	Average Cost (PSO)	Average lastdag (ACO)	Average lastdag (PSO)
two	1.1	0	0	0
four	2.6	4.86	480.434	642.4224
six	3.6	5.52	194.0196	784.07
eight	4	14.97	835.4491	1380.5446
ten	7.5	17.67	1085.1278	1748.8282
twelve	8.1	20.76	1235.6481	1997.1848

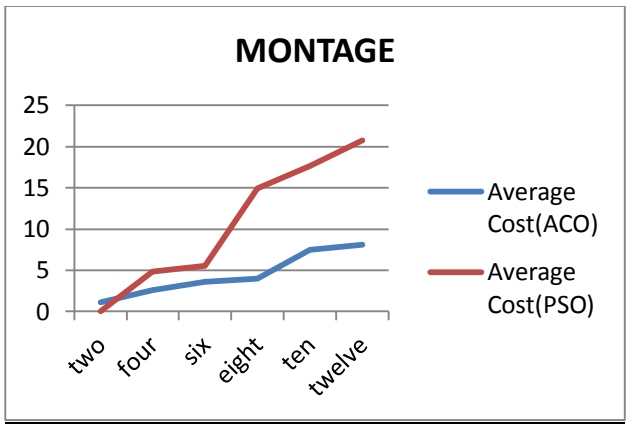


Fig. 4.7 Cost of Montage workflow

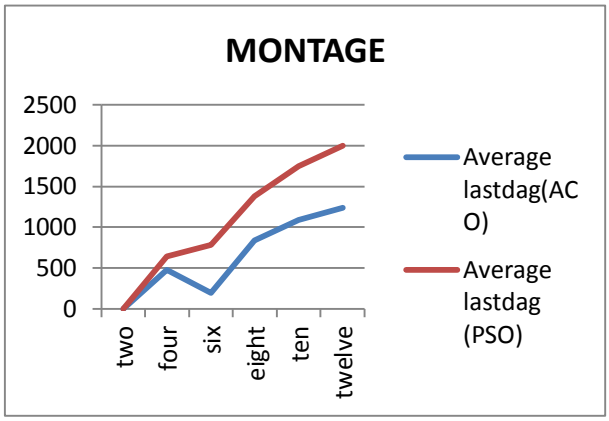


Fig. 4.8Makespan of Montage workflow

V. CONCLUSION AND FUTURE SCOPE

In this paper, we presented a scheduling strategy for executing scientific workflows on IaaS Clouds. The scenario was modeled as an optimization problem which aims to minimize the overall execution cost while reducing the makespan and the problem was solved using the hybrid of ACO and PSO. The experiments were conducted by simulating four well known workflows (Cybershake, Ligo, Genome, Montage) on Cloudsim, which shows that our solution has an overall healthier performance than other state-of-the-art algorithms. The worthy results are achieved because PSO (particle swarm optimization) play important role in global optimization and ACO(ant colony optimization) optimize locally and we have merge the two algorithms by taking the best out of them. With the proposed approach in most of the workflows we are able to produce lower cost efficient schedule meanwhile also reducing the time delay. As future work, we would like to explore various options for the selection of the preliminary resource pool as it has a major impact on the performance of the algorithm. We would also like to research with different optimization approaches such as genetic algorithms and

compare their performance with PSO, ACO and Hybrid of both these algorithm.

VI. REFERENCES

[1] Zhang, Qi, Lu Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges" in *Journal of internet services and applications*, vol. 1, pp. 7-18, Feb. 2010.

[2] Mell, Peter and T. Grance ,"The NIST definition of cloud computing." Oct. 25, 2011.

[3] Dillon, Tharam, Chen Wu, and C. Elizabeth, "Cloud computing: issues and challenges." in *24th IEEE International Conference on Advanced Information Networking and Applications (AINA),* pp. 27-33, 2010.

[4] http://en.wikipedia.org/wiki/Scheduling_%28computing%29 (last accessed on 23/1/15 at 11:47 AM)

[5] S.Tilak and D. Patil " A Survey of Various Algorithms in Cloud Environment" in *International Journal of Engineering Inventions,* vol.1 no. 2,pp. 36-39, Sep. 2012.

[6] Pandey, Suraj, L. Wu, S. Mayura and R.Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments." in *24th IEEE International Conference on Advanced Information Networking and Applications (AINA),* pp. 400-407, April 20, 2010.

[7] Chen, Wei-Neng, and Jun Zhang, "An ant colony optimization approach to a grid workflow scheduling problem with various QoS requirements." in *IEEE Transactions Systems, Man, and Cybernetics, Part C: Applications and Reviews,* vol. 39, no.1 pp. 29-43, Jan. 2009.

[8] Byun, Eun-Kyu, Y.S Kee, J.S Kim and S.Maeng, "Cost optimized provisioning of elastic resources for application workflows." in *ELSEVIERFuture Generation Computer Systems*, vol. 27, no. 8, pp. 1011-1026, Oct. 2011.

[9] Malawski, Maciej, G.Juve, E. Deelman and J. Nabrzyski, "Cost-and deadline-constrained provisioning for scientific workflow ensembles in Iaas clouds." in *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, pp. 22, 2012.

[10] Abrishami, Saeid, Mahmoud Naghibzadeh and Dick HJ Epema,"Deadline-constrained workflow scheduling algorithms for Infrastructure as a Service Clouds." in *ELSEVIERFuture Generation Computer Systems* ,vol. 29, no. 1 pp. 158-169, Jan.2013.

[11] Xue, Sheng-Jun, and Wu Wu , "Scheduling workflow in cloud computing based on hybrid particle swarm algorithm." in *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 10, no. 7 pp. 1560-1566, 2012.

[12] Rodriguez Sossa, M., and RajkumarBuyya, "Deadline based Resource Provisioning and Scheduling Algorithm for Scientific Workflows on Clouds." in*IEEE Transactionson Cloud Computing*, vol. 2, no. 2, pp.222-235, June 2014.

[13] Netjinda, Nuttapong, B. Sirinaovakul, and T. Achalakul, "Cost optimal scheduling in IaaS for dependent workload with particle swarm optimization." in *The Journal of Supercomputing,* pp. 1-25, Feb. 2014.

[14] AmandeepVerma, and SakshiKaushal, "Bi-Criteria Priority based Particle Swarm Optimization workflow scheduling algorithm for cloud." in Engineering and Computational Sciences (RAECS), pp. 1-6, March 2014.