Review on Workflow Scheduling by Optimization

Saurabh chaudhary[1],Sukhmeet brar[2]
[1]Dept. of Computer Science and Engineering, CGC Jhanjerit Mohali, India
[2]Asst. Prof., Dept. of Computer Science and Engineering, CGC Jhanjerit Mohali, India [1]

Abtract- Cloud computing is a distributed design that brings together server resources on an acceptable stage in order to provide on request figuring resources and administrative data. A cloud specialist organization (CSP's) provides the different stages to their customers to use the services and make the web administrative control. This service is similar to a broadband band connection offered by the service provider for the internet connection. Cloud computing provides the services through the internet, these service belongs to hardware and software both. Cloud computing concept is growing high day by day due to its service pay per usage concept. When cloud provides the service in the form of platform it is called as Platform as a service model. When cloud provides the hardware to the consumer it is called as Infrastructure as a service model. When cloud provided the software services it is also called as Software as s service.

Keywords: cloud, scheduling , task, consumer

I. INTRODUCTION

Cloud computing is a service focused model which provides the services at low cost and it offers the facility of pay per use service which are customized SLA. Cloud computing approach reduces the cost of installation of hardware, software by providing them all at the virtual environment called as cloud. Cloud computing helps the user to access the cloud application anywhere and anytime. As the cloud provides different services on different model like infrastructure on IaaS, platform on PaaS and software on SaaS it is necessary to manage the resources properly and provide to the users on time. The management process of the resources on cloud is called scheduling [1].

Scheduling process allows optimal allocations of resources among given task in a finite time to achieve the desired quality of service. Scheduling is an important part of cloud computing to provide the scalability, reliability, information sharing and low cost computation in distributed environment. Scheduling is an approach which is used to distribute the computing resources, processor time memory and bandwidth. Mainly scheduling is used to balance the load on the system and appropriate distribution of resource on the basis of prioritization according to the rules set. This is also known as process scheduling [2].

A. Types of scheduling

Scheduling is consists of two types that are following.

1. Task Scheduling(Independent scheduling)

2. Workflow scheduling(Dependent Scheduling)

Independent scheduling

Independent scheduling is task scheduling which has been one of the most smoking investigates, there are numerous specialists and researchers distributed papers and diaries venture to talk about the task scheduling issue. Task scheduling in Cloud Computing is a component that maps client task to relevant assets to execute; its effectiveness will specifically influence the execution of entire Cloud Computing condition. Profitable task scheduling framework can meet customers' necessities, and upgrade the benefit use, subsequently enhancing the general execution of the distributed computing condition. Be that as it may, the assignment planning in framework figuring is regularly about the static errand necessities, and the assets usage rate is additionally low. By new components of distributed computing, for example, adaptability, virtualization all the physical assets are virtualized and straightforward to client.
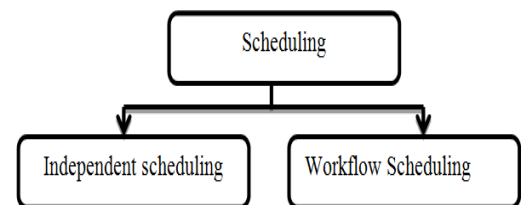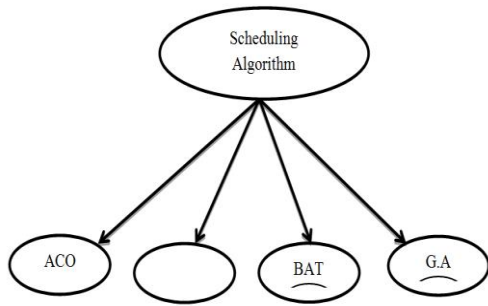

Workflow Scheduling

Workflow scheduling is a major worry in the time of figuring. Fundamentally it is the worry identified with the plotting of each assignment to a fitting asset and enabling the errand to fulfill some execution limitations. It is a succession of associated guidelines. Point of workflow scheduling is to mechanize the strategies particularly which are included during the time spent passing the information and the documents between the members of the cloud, keeping up the limitations. A workflow empowers the organizing of utilizations in a coordinated non-cyclic chart structure where every hub speaks to the undertaking and edges speak to the conditions between the hubs of the applications.

B. Scheduling Algorithms and Metrics

 In this section we study the scheduling algorithms and metrics that are used in the cloud computing. Different researcher used different parameters for optimization and scheduling. Some parameters that are mostly used in scheduling are Makespan (describe the finishing time of the last task), Economic cost (represents the total cost paid by the

user for service), Flow time (define the sum of finishing times of all tasks).

1. ACO (Ant Colony Based Optimization): Vinothinaproposed Ant Colony Optimization algorithm for workflow scheduling in cloud computing. This model is presented for heterogeneous distributed systems. The service level agreements are used to check the quality of service of the service providers. The problem of workflow scheduling is solved by using parameters cost, makespan and resource utilization. The ACO algorithms reduce the cost and makespan and enhance the resource utilization [3].

2. PSO (Particle Swarm optimization): Suraj Pandey et al.has worked on reducing the computing cost of the application by using particle swarm optimization algorithm which is basically a meta-heuristic algorithm used for scheduling. PSO is used for calculating the fitness function. The total cost calculated is the cost of execution and the transfer cost of data. This algorithm ensures the cost of the highest task is reduced by heuristic scheduling. This algorithm helps to schedule the resources and mapping [4]. Nancharaiah et al**.** displayed hybrid routing algorithm, Ant Colony Optimization algorithm and Particle Swarm Optimization (PSO) is utilized to enhance the different measurements in MANET routing. The ACO algorithm utilizes portable specialists as ants to distinguish the most possible and best way in a system. Likewise ACO algorithm finds ways between two hubs in a system and gives contribution to the PSO strategy. The PSO finds the best answer for a particle's position and speed and limits cost, power, and end to end delay. This hybrid routing shrewd algorithm has an enhanced execution when contrasted and basic ACO algorithm as far as delay, power, consumption, and communication cost [5].

Zhan proposed progressed PSO based venture scheduling set of rules in Cloud Computing which can lessen the assignment common jogging time and raises the supply of sources which complements the convergence price and improves the efficiency. They investigated that that is due to the fact in every generation worldwide fast convergence of simulated annealing algorithm is applied to combine particle swarm optimization algorithm [6].

3. BAT Algorithm: Bat algorithm is a meta-heuristic optimization algorithm which is used to get the optimal solution. It is basically based on the echolocation behaviour of the bats with varying pulses rates of emissions and loudness. The working of this algorithm is depending on the velocity and position of bat which vary according to the frequency, wavelength and loudness.Sagnika et al.proposed BAT algorithm for workflow scheduling in cloud computing which helps to handle the large size of data. The scheduling process decides that which task is executed first and which is last according to their requirement of the resources. It manages the resources according to the task size and execution time. The result of the proposed algorithm is compared with particle swarm optimization algorithm and Cat swarm optimization algorithm. The convergence of the proposed algorithm is better than the existing algorithms [7].

4. Genetic Algorithm (G.A): Liu, Li, et al. proposed the genetic algorithm for workflow scheduling in cloud computing with deadline-constrained. The crossover and mutation probability is adjusted by using convolution approach. This approach prevents from the prematurity and enhances the convergence. The proposed approach is compared with existing algorithms on the simulator at 4 different workflows. The results show that the total execution cost is reduced in this approach[8].Garg, et al. formulated the scheduling problem in cloud by using the Genetic Algorithm. The proposed work is done to reduce the computation time and execution cost of the task. This work is done on the cloudSim simulator and it maximizes the resource utilization. The performance evaluation is done on the different parameters and performs well [9].

Cloud supports the Multi-tenancy feature and provides the scalability and other benefits to the other users. Resource management is an important task in the multi-tenant cloud computing which is done by using the scheduling process. In this work cloud based workflow scheduling policy is proposed for efficient computing in cloud. This policy reduces the overall workflow completion time, cost of execution and properly utilize the resources. The result of the proposed work is compared with existing approaches and algorithms. The simulation result of the proposed approach shows more effective results than the existing approaches [10].Casas, Israel, et al.proposed a scheduling approach called Balanced and file Reuse-Replication scheduling. This approach is used to schedule the scientific application workflows. It splits the workflows into sub-workflows which help in proper utilization via parallelization process. This approach provides the facility if data reuse and replication which helps in optimization of data and transfer it at run time. The optimization process is based on execution time and monetary cost of workflows [11]

B. Inference from the Literature Review

Algorithm Details			Performance Metrics			
Author's Name	Year	Algorithm Used	Makespan	Response Time	Throughput	Load Balancing
Vinothina et al.	2018	*Ant Colony Based Optimization*		√	√	√
Suraj Pandey et al	*2010*	*Particle Swarm optimization*	√		√	√
Nancharaiah et al**.**	*2013*	*Particle Swarm optimization*		√	√	√
Zhan et al.	*2013*	*Particle Swarm optimization*	√	√	√	
Sagnika et al.	*2018*	*BAT*		√		√
Liu, Li, et al.	*2017*	*Genetic Algorithm*	√		√	
Garg, et al.	*2017*	*Genetic Algorithm*	√	√	√	
Rimal et al.	*2017*	Workflow scheduling			√	√
Casas, Israel, et al.	*2017*	Balanced and file Reuse-Replication scheduling	√	√	√	√

II. CONCLUSION

The main issue in the cloud computing while decreasing makespan is execution cost. This issue is solved by using Hybrid GWO. The tests were directed by mimicking four surely understood work processes (Cybershake, Ligo, Genome, Montage) on Cloudsim, which demonstrates that our answer has a general more beneficial execution than other existing algorithms.

III. REFERENCES

[1]. Kalra, Mala, and Sarbjeet Singh. "A review of metaheuristic scheduling techniques in cloud computing." *Egyptian informatics journal* 16.3 (2015): 275-295.

[2]. Jula, Amin, ElankovanSundararajan, and Zalinda Othman. "Cloud computing service composition: A systematic literature review." *Expert Systems with Applications* 41.8 (2014): 3809-3824.

[3]. Vinothina, V., and R. Sridaran. "An Approach for Workflow Scheduling in Cloud Using ACO." *Big Data Analytics*.Springer, Singapore, 2018.525-531.

[4]. Pandey, Suraj, et al. "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments." *Advanced information networking and applications (AINA), 2010 24th IEEE international conference on*.IEEE, 2010.

[5]. Sajid, Mohammad, and Z. Raza, "Cloud Computing: Issues and Challenges," in *International conference on Cloud, Big Data and Trust*, pp. 35 – 41, Nov. 2013.

[6]. S.Xavier and S.P.J Lovesum" A Survey of Various Workflow Scheduling Algorithms" in Cloud Enviroment" in *International Journal of Scientific and Research Publications,* vol. 3, no.2, Feb. 2013.

[7]. Sagnika, Santwana, SaurabhBilgaiyan, and Bhabani Shankar Prasad Mishra. "Workflow Scheduling in Cloud Computing Environment Using Bat Algorithm." *Proceedings of First International Conference on Smart System, Innovations and Computing*.Springer, Singapore, 2018.

[8]. Liu, Li, et al. "Deadline-constrained coevolutionary genetic algorithm for scientific workflow scheduling in cloud computing." *Concurrency and Computation: Practice and Experience* 29.5 (2017).

[9]. Garg, ErJyoti, and GurjitBhathalEr. "Research Paper on Genetic Based Workflow Scheduling Algorithm in Cloud Computing." *International Journal of Advanced Research in Computer Science* 8.5 (2017).

[10]. Rimal, Bhaskar Prasad, and Martin Maier. "Workflow scheduling in multi-tenant cloud computing environments." *IEEE Transactions on Parallel and Distributed Systems* 28.1 (2017): 290-304.

[11]. Casas, Israel, et al. "A balanced scheduler with data reuse and replication for scientific workflows in cloud computing systems." *Future Generation Computer Systems* 74 (2017): 168-178.