IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 10, October 2015.
www.ijiset.com
ISSN 2348 – 7968

A Survey on Workflow Task Scheduling Using Intelligent Water Droplets in Cloud Computing

P. Kumar
Associate Professor
Department of Information Technology
Rajalakshmi Engineering College
Thandalam, Chennai, Tamilnadu.
Kumar@rajalakshmi.edu.in

P. Shobana
PG Scholar
Department of Information Technology
Rajalakshmi Engineering College
Thandalam, Chennai, Tamilnadu.
shobana.p.2014.mese.rajalakshmi.edu.in

ABSTRACT

Cloud computing is a internet- based computing where different services are delivered to computers and devices through internet. Cloud computing is used for delivering and managing the services or application in an on-demand environment. Cloud computing provides data access and storage devices without the knowledge of the physical location of the end user. Scheduling the tasks is a challenging issue in the cloud computing. The tasks are represented as workflows, which are processed in a specific order based on their required services. There are several algorithms to automate the workflows in a way to satisfy the Quality of service (QoS) of the user among which Execution time, Cost and Load balance is considered as a major criterion. This survey paper focuses on QoS parameters such as Execution time, Cost and load balancing which is done using several different algorithm like Ant colony optimization, Particle Swarm optimization and Intelligent Water Drops Algorithm, in order to achieve efficient resource scheduling for the cloud environment.

Keywords: Cloud computing, IWD (Intelligent Water Drops) Algorithm, Load Balancing, Task scheduling, Workflow.

1. INTRODUCTION

Cloud computing is the delivery of computing as a service where shared resources, software and information are provided to computers and other devices as a utility (like the electricity grid) over a network (typically the Internet). Cloud computing provide users with various capabilities to store and process their data. It relies on sharing of resources to achieve coherence and economies of scale, similar to a utility over a network. It mainly focuses on maximizing the effectiveness of shared resources. Cloud computing is a Meta form of

ISSN 2348 – 7968

internet. Cloud providers typically use a "pay as you go" model. This can lead to unexpectedly high charges if administrators do not adapt to the cloud pricing model.

Cloud computing architecture refers to the components and subcomponents required for cloud computing. These components typically consist of a front end platform, back end platforms, a cloud based delivery, and a network. Combined, these components make up cloud computing architecture. Cloud computing architectures consist of front-end platforms called clients or cloud clients. These clients comprise servers, fat or thick clients, thin clients, and mobile devices. These client platforms interact with the cloud data storage via an application, or a web browser, or through a virtual session. An online network storage where data is stored and accessible to multiple clients. Cloud storage is generally deployed in the following configurations: public cloud, private cloud, community cloud and hybrid cloud. cloud based delivery are done in three different services. The services are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).

Cloud resources are usually not only shared by multiple users but are also dynamically reallocated per demand. With cloud computing, multiple users can access a single server to retrieve and update their data without purchasing licenses for different application. Proponents claim that cloud computing allows companies to avoid upfront infrastructure costs, and focus on projects that differentiate their businesses instead of an infrastructure. The goal of cloud computing is to allow users to take benefit from all of these technologies, without the need for deep knowledge about or expertise with each one of them. The cloud aims to cut costs, and helps the users focus on their core business instead of being impeded by IT obstacles.

The main enabling technology for cloud computing is virtualization. Virtualization software separates a physical computing device into one or more "virtual" devices, each of which can be easily used and managed to perform computing tasks. With operating system-level virtualization essentially creating a scalable system of multiple independent computing devices, idle computing resources can be allocated and used more efficiently. Virtualization provides the agility required to speed up IT operations, and reduces cost by increasing infrastructure utilization. Autonomic computing automates the process through which the user can provision resources on-demand. By minimizing user involvement, automation speeds up the process, reduces labor costs and reduces the possibility of human errors.

ISSN 2348 – 7968

1.1 Workflow Scheduling

Workflow is a series of activities that are necessary to complete the task. Workflow scheduling maps and manages the execution of inter-dependent tasks on the distributed resources. It allocates suitable resources to workflow tasks such that the execution can be completed to satisfy objective functions imposed by users.

Figure 1.1 shows Workflow task scheduling diagram which is described as a Directed Acyclic Graph (DAG), defined by a tuple G = (T,E) where G is a Graph, T is the set of n tasks $\{T_1,T_2,T_3,...,T_n\}$, and E is a set of e edges, represent the dependencies. If task T_i has a directed edge pointing to task T_j, T_i is called the parent task of T_j and T_j is called the child task of T_i. In this case T_j can only be executed until its entire parent tasks have been completed.

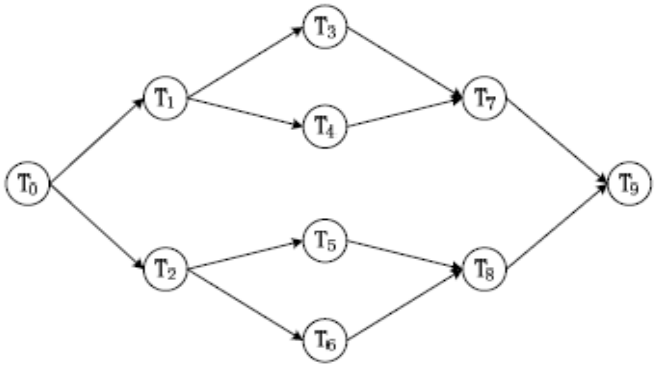

Fig.1.1. Workflow Scheduling

The above figure depicts the basic structure of a set of workflow, it contains ten tasks to be processed and these tasks are labeled from T_0 to T_9. we use VM= $\{VM_1, VM_2, VM_3, \ldots, VM_m\}$ to indicate the set of virtual machines users leased, m represents the number of these virtual machines. MIPS denotes the processing speed of VM, where MIPS represent millions of instructions per second.

2. A REVIEW OF OPTIMIZATION TECHNIQUES FOR TASK SCHEDULING

Task scheduling is the main issue in cloud computing. Many optimization algorithms have been applied to solve this problem. Different researchers have proposed various

ISSN 2348 – 7968

algorithms for allocating and scheduling the resources efficiently in the cloud. Here the authors provide a comparative study of different evolutionary and swarm-based techniques that perform scheduling of tasks to resources, such as ant colony optimization, particle swarm optimization, and intelligent water drops, etc. Various modified scheduling algorithms like Modified Ant Colony Optimization, Multi-objective Particle Swarm Optimization have also been analyzed.

2.1 ANT COLONY OPTIMIZATION

Ant colony optimization is a population- based search technique for the solution of combinatorial optimization problem which is inspired by the behavior of real ant colonies where ants cooperate and communicate through pheromone trails in search of food. An ant solves a problem iteratively by using a construction graph where edges represent the possible partial solution that the ant can take according to a probabilistic state transition rule. After selecting a partial or a complete solution, a rule of pheromone updating starts. This rule gives a feedback mechanism to speed up convergence, and also prevents premature solution stagnation. Various algorithms based on the ACO meta-heuristic have been applied to many difficult optimization problems.

2.2 PARTICLE SWARM OPTIMIZATION

Particle swarm Optimization is a population based stochastic optimization algorithm. Particle swarm Optimization is being used efficiently to schedule workflows. Particle Swarm Optimization (PSO) has recently emerged as a prominent heuristic approach, applicable to various large and complex problems, like task scheduling problem, knowledge extraction in data mining, electric power systems, etc. The PSO algorithm works by initializing a population (swarm of) candidate solutions (known as particles). Each particle will keep track of its best solution named as personal best(pbest) and also the best value of any particle called as global best (gbest). The particles, fly through the quandary space by following the current optimum particles. Every particle changes its position based on its current position, speed. To add on, the distance between the recent position and pbest, the distance between the current position and gbest are also considered. PSO follows the principle of random searching in entire solution space using a large population, depending upon the problem domain. Recent

work shows that PSO gives better performance over other existing techniques in efficient optimization.

2.3 INTELLIGENT WATER DROPS ALGORITHM

The Intelligent Water Drops **(**IWD) algorithm is inspired by the movement of natural water drops which flow in rivers, lakes and seas. IWD algorithm is a best choice for finding optimal solutions. It is very promising problem solving algorithm and it is improved to apply more engineering problems. IWDs are associated with two attributes, namely, the amount of soil in a path and the velocity of the IWDs. The velocity enables the water drops to transfer soil from one place to another. Faster water drops can gather and transfer more soil from the riverbeds. Besides, the velocity of the IWDs is also affected by the path condition. The amount of soil in a path has impact on the IWDs' soil collection and movement. A path with less soil allows the IWDs to move faster along that path. IWD algorithm consists of two types of parameters: Static parameters and Dynamic parameters. Static parameters are those that remain constant during the lifetime of the IWD algorithm. Dynamic parameters are reinitialized at the end of iteration of the IWD algorithm.

3. ANALYSIS OF TASK SCHEDULING WITH QOS PARAMETERS

Many authors have done work in the area of workflow task scheduling algorithm. Table 1 represents the description of work done in terms of purpose, method used, advantage and issues of the paper.

S.NO	AUTHOR	TITLE & YEAR	PURPOSE	METHOD USED	ADVANTAGE	ISSUES
1	Navjot Kaur , Taranjit Singh Aulakh, Rajbir Singh Cheema	Comparison of Workflow Scheduling Algorithms in Cloud Computing. 2011	To study various problems, issues and types of scheduling algorithms for cloud workflows as well as on designing new	1. Prioritize by Organizing, Streamlining, Economizing and Contributing (POSEC). 2. Pareto Analysis	Reduces the makespan.	We can include various parameters and add more clouds to this main cloud, to distribute the load work.

			workflow algorithm.			
2	Jayadivya S K & S. Mary Saira Bhanu	QoS Based Scheduling of Workflows in Cloud Computing. 2012	To schedule many workflows based on user given QoS parameters.	QoS based Workflow Scheduling Algorithm (QWS)	1.Produce good success rate of scheduling. 2.Makespan and cost is minimized. 3.Maximum reliability.	1. Some more parameter like availability can be include. 2.Data grouping before scheduling.
3	D.Palanikkumar, gowsalya Elangovan, B. rithu, P. anbuselven	An Intelligent Water Drops Algorithm Based Service Selection And Composition In Service Oriented Architecture. 2012	To select most suitable service from a set of candidate services. Computation of QoS is more important in service selection.	Intelligent Water Drops Algorithm (IWD)	1.Feasible & effective way to find optimal solution. 2. Fast convergence. 3. Reduces Execution time.	There is a need for balancing the load in system and reduces the cost.
4	Pardeep Kumar, Amandeep Verma	Scheduling Using Improved Genetic Algorithm in Cloud Computing for Independent Tasks. 2012	Allocate resource to user's requests, so that the tasks can be completed in a minimum time.	Improved genetic algorithm is proposed.	1. Minimize the makespan 2. Effective utilization of resources.	1. There is need of execution cost of an fitness criteria for decreasing makespan 2. Better resource utilization.
5	PaulRani .A M.Gomaty Nayagam	Multi-Objective QoS Optimization Based On Multiple Workflow Scheduling In Cloud Environment 2013	Scheduling the services such that execution time and cost is reduced.	Multiple QoS Constrained Scheduling Strategy of Multi-Workflows MQMCE algorithm	1.Reduce the cost and time 2.Increase the reliability and availability.	Combine MQMCE with MOGA for optimize the QoS parameters highly based on more user requirements.
6	Xiaonian Wu, Mengqing Deng, Runlian Zhang,	A task scheduling algorithm	1. To prioritize the task	Task Scheduling Algorithm	1. Priority is computed and tasks are sorted	Include some more QoS parameter

	Bing Zeng, Shengyan Zhou	based on QoS- driven in cloud computing. 2013	according to special attributes of tasks and sort tasks. 2.To minimize completion time and also achieve load balance of entire system.		by priority. 2. Completion time is minimized and load balance is achieved.	like availability, cost, deadline, etc.
7	Mayanka Katyal and Atul Mishra	Application of Selective Algorithm for Effective Resource Provisioning In Cloud Computing Environment 2014	1.Resources need to be provisioned to the end user in most efficient manner. 2.The overall makespan is minimized.	Selective algorithm	1.All resources are efficiently utilized & jobs given by user are executed with smaller delays. 2.Makespan is reduced.	Further we can minimize make span & throughput.
8	Shagufta Khan , Niresh Sharma	Effective Scheduling Algorithm for Load balancing (SALB) using Ant Colony Optimization in Cloud Computing. 2014	1.To develop an effective load balancing algorithm using ACO. 2.To balance the entire system load while trying to maximize & minimize different parameter.	Scheduling algorithm for load balancing (SALB)	1.Efficient to finding the overloaded node in minimum time. 2.To balance the node with efficiency & maximum utilization of resources.	Less performance
9	Shengjun Xue, Mengying Li, Xiaolong Xu, and Jingyi Chen	An ACO-LB Algorithm for Task Scheduling in the Cloud Environment 2014	To solve the load imbalance of virtual machine in the process of task scheduling	Improved ACO algorithm and ACO-LB algorithm	1.Better performance 2.Avoid the resource wastage. 3.Improved efficiency of the system. 4.Improve the utilization rate of resources.	1.Propose different types of workflow models to verify reliability of ACO-LB algorithm. 2.Realize high efficiency &

						low cost.
10	Pritibahen, Sumanbhai Patel	Multi-Objective Job Scheduler using Genetic Algorithm in Grid Computing. 2014	1.multi-objective job scheduler provides efficient utilization of resources by completing tasks in a minimum time (makespan) and flow time to find nearly optimal solution.	Genetic algorithm	1.To design dynamic schedulers. 2.Makespan & flow time is reduced.	1.Further specify requirements along with tasks & their dependencies 2.Integrated with various grid middleware.
11	Arabi E. keshk, Ashraf B. El-Sisi, Medhat A. Tawfeek	Cloud Task Scheduling for Load Balancing based on Intelligent Strategy. 2014	1.To balance the system load while trying to minimizing the makespan of a given tasks set. 2.Allocating incoming jobs to the virtual machine.	Modified ACO for load balancing algorithm. (MACOLB)	1.Load balance is minimized. 2.MACOLB algorithm takes less time to complete tasks.	1.There is no precedence constraint between tasks. 2.Improve cost of resource.
12	M. Vidhya, N.Sadhasivam	Parallel Particle Swarm Optimization for Task Scheduling in Cloud Computing. 2015	To find the optimal mapping of tasks to resources which minimizes the job completion time.	Parallel particle swarm optimization algorithm.	1.Reduces execution time. 2.Easy to implement. 3.Better task assignment.	There is a need of minimization of cost & dynamic load balancing.
13	Akanksha Tripathi, Sandeep Raghuwanshi	Cloud Task Scheduling in Hybrid Environment using Ant Colony Optimization	1. To resolve problem of load balancing in CC environment	Modified LBACO algorithm	1.Choose optimal resource to perform tasks. 2.Minimize the makespan of a given set of tasks.	Further we reduce the cost.

		2015	2.To optimize the flow time & makespan time through algorithm.		3.It also balance the entire system load.	
14	R. Jemina Priyadarsini, Dr. L. Arockiam	An Improved Particle Swarm Optimization Algorithm for Meta Task Scheduling In Cloud Environment 2015	1.To achieve better solution with global & fast convergence rate. 2.Improve the efficiency. 3.Minimize makespan & resource utilization.	Improved particle swarm optimization (IPSO) Algorithm.	1.Better performance in resource utilization. 2.Lower makespan. 3.Improve candidate solution.	Further we can minimize the cost.
15	Ali Akbar Faraj, Ali Haroon Abadi	Proposing a scheduling algorithm to balance the time and cost using a genetic algorithm. 2015	1.To reduce performance time & cost-effective use of resources. 2.To balance the load of resources.	Genetic algorithm with binary gravitational emulation local search algorithm.	1.High performance regarding creation of balance between cost & tasks scheduling. 2.Cost & overall time is reduced.	Need of resource allocation to works using an approach based on fuzzy logic & fuzzy inference system.

ISSN 2348 – 7968

4. CONCLUSION

In this paper the authors analyze the various algorithms that have been used for task scheduling on resources in cloud computing environments. Scheduling is an important activity in multi-tasking systems to manage resources, minimize completion time and increase performance of systems. To implement task scheduling, any of the above discussed methods can be used to get optimistic scheduling. The benefits of using these methods is to get the optimal solution in minimum time. Future work in this field can include application of newer optimization methods like modified IWD, etc, which may be able to provide better and faster results.

REFERENCES

[1] Navjot Kaur et al. (2011), 'Comparison of Workflow Scheduling Algorithm in Cloud Computing', International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10.

[2] Jayadivya S K. and Mary Saira Bhanu S. (2012), 'QoS Based Scheduling of Workflows in Cloud Computing', International Journal of Computer Science and Electrical Engineering, Vol.1, No.2315-4209.

[3] Palanikkumar D. et al. (2012), 'An Intelligent Water Drops Algorithm Based Service Selection And Composition In Service Oriented Architecture', Journal of Theoretical and Applied Information Technology, Vol. 39, No. 1.

[4] Pardeep Kumar and Amandeep Verma (2012), 'Scheduling using Improved Genetic Algorithm in Cloud Computing for Independent Tasks', conf.

ISSN 2348 – 7968

[5] PaulRani A. and Gomaty Nayagam M.(2013), 'Multi-Objective QoS Optimization Based On Multiple Workflow Scheduling In Cloud Environment', International Journal of Innovative Research in Computer and Communication Engineering, Vol. 1.

[6] Xiaonian Wu et al. (2013), 'A task scheduling algorithm based on QoS- driven in cloud computing ', International Conference Information Technology and Quantitative Management.

[7] Mayanka Katyal and Atul Mishra (2014), 'Application of Selective Algorithm for Effective Resource Provisioning In Cloud Computing Environment', International Journal on Cloud Computing: Services and Architecture, Vol. 4, No. 1.

[8] Shagufta Khan and Niresh Sharma (2014), 'Effective Scheduling Algorithm for Load balancing (SALB) using Ant Colony Optimization in Cloud Computing', International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4.

[9] Shengjun Xue et al. (2014),'An ACO-LB Algorithm for Task Scheduling in the Cloud Environment', Journal of Software, Vol. 9, No. 2.

[10] Pritibahen Sumanbhai Patel (2014), 'Multi-Objective Job Scheduler using Genetic Algorithm in Grid Computing' ,International Journal of Computer Applications, Vol. 92, No. 14.

[11] Arabi E. keshk et al. (2014), 'Cloud Task Scheduling for Load Balancing based on Intelligent Strategy' , International Journal of Intelligent System and Applications.

[12] Vidhya M. and Sadhasivam N. (2015),'Parallel Particle Swarm Optimization for Task Scheduling in Cloud Computing' , International Journal of Innovative Research in Science, Engineering and Technology, Vol. 4.

[13] Akanksha Tripathi and Sandeep Raghuwanshi (2015). 'Cloud Task Scheduling in Hybrid Environment using Ant Colony Optimization', International Journal of Advanced Engineering and Global Technology, Vol. 3.

[14] Jemina Priyadarsini R. and Arockiam L. (2015), 'An Improved Particle Swarm Optimization Algorithm for Meta Task Scheduling In Cloud Environment', International Journal of Computer Science Trends and Technology, Vol. 3.

ISSN 2348 – 7968

[15] Ali Akbar Faraj and Ali Haroon Abadi (2015), 'Proposing a scheduling algorithm to balance the time and cost using a genetic algorithm', International Journal of Computer Applications Technology and Research, Vol. 4.