A Comprehensive Study Of Task Scheduling and Load Balancing Techniques In Cloud Computing

Rinsha K P
MTech Student, Computer science and Engineering
NSS College of Engineering
Palakkad, Kerala, India
Email: ashnikp@gmail.com

S Sindhu
Associate Professor,Computer science and Engineering
NSS College of Engineering
Palakkad, Kerala, India
Email: sindhu.nss@gmail.com

***Abstract*—Cloud computing is an emerging technology in modern era. Cloud computing provides computing environments, resources and other services to a customer on a pay as you go basis. Task scheduling is one of the essential and important part of cloud computing. Task scheduling is a critical problem in cloud computing because a cloud provider has to serve many requests at a time. An efficient scheduling algorithm is responsible for efficient allocation and utilization of resources. Efficient allocation of resources balance the load in each resource. A good scheduling policy always helps to reduce the task completion time. Many algorithms and techniques have been proposed to solve the problem of task scheduling and thereby load balancing. This paper discuss about the study of various scheduling and load balancing algorithms in cloud environment.**

Key terms: Cloud Computing, Task scheduling, Load balancing, Makespan, Response time, Scheduling algorithms.

I. Introduction

Technologies are developing in this modern era. When technologies are developed, new paradigms are evolving in the field of computations. One such emerging field is cloud computing. There are many cloud computing providers, who provide resources over internet to the consumers. Hence cloud computing is a technology which provides access to a shared pool of resources over the internet. Resources include processing unit, storage, applications etc[1].

Cloud computing paradigm have many characteristics. Out of which, there are five essential characteristics that are identified by NIST(National Institutes of standards and technology)[1]. The five essential characteristics are on demand self service, broad network access, resource pooling, rapid elasticity and measured service.

Three types of services are mainly provided by cloud. They are Software as a Service(SaaS), Platform as a Service(PaaS) and Infrastructure as a Service(IaaS). Later with more development of technologies, now a days every thing can be provided by cloud and is known as Everything as a Service(XaaS). Cloud computing allows simultaneous execution of many tasks in a hardware platform. This is made possible by a technique known as virtualization. Computing resources are given in the form of abstract machines that run on physical machines. Such machines are known as virtual machines. Virtual machines provide isolation to the task. The isolation between the virtual machines are done by a virtual machine manager known as hypervisor. Virtual machines are responsible for on demand configuration of physical machines to run diverse task and to avoid resource wastage.

The technology of virtualization helps in load balancing in cloud environment. Load balancing is done dynamically. Hence the resources may be over utilized or under utilized. If the resources are overloaded, there will be performance degradation. And if the resources are under loaded, the resource utilization is less. Thus it is necessary to provide right amount of resource dynamically to the applications running in virtual machines. Hence there should be a sophisticated technique for load balancing. To provide optimal load balancing strategy, the tasks or jobs are scheduled to the resources efficiently. This is the importance of task scheduling in a cloud platform. A good task scheduling algorithm will distribute the load optimally and hence provide efficiency.

There are many factors in which the reliability of the system depends on. One such factor is scheduling. There are three levels of scheduling like task level scheduling, resource level scheduling and workflow level scheduling[2]. Among these three scheduling task scheduling is important and is responsible for load balancing in cloud computing systems. Task is a small piece of work that is executed in a given interval of time on a cloud data center. Inorder to get a task completed, task is to be dispatched to the resources at the provider. Thus the dispatching task to perfect resources is known as task scheduling. The method used to perform this dispatching is known as task scheduling algorithms.

In the process of scheduling, the task is submitted to the cloud scheduler. The cloud scheduler allocates each task submitted on different resources as per the requirements of task by inquiring the cloud information service for getting the status of resources available. Task scheduling in cloud is a NP complete problem. A good scheduling algorithm improves performance. The performance parameters include CPU utilization, turnaround time, throughput, response time etc[2].

This paper discuss various task scheduling techniques developed so far. Load balancing and energy utilization are two parameters that should be considered during scheduling of tasks. How each scheduling techniques optimizes the parameters of scheduling is discussed in this paper. Rest of the paper is organized as follows: Section II describes the fundamental concept of load balancing and

task scheduling algorithms. Different task scheduling algorithms, their features, pros and cons are discussed in section III. Finally section IV concludes the paper with a future scope of research on this topic.

II. Task Scheduling and Load Balancing

Scheduling is one of the major feature in cloud computing environment. A scheduler may schedule a task to virtual machines or a virtual machine to physical host or schedule workflows constituting an entire job in a suitable order. Hence there are three types of scheduling. They are resource scheduling, workflow scheduling and task scheduling. Out of the three only task scheduling is considered in this paper. Task scheduling refers to the scheduling of tasks to the virtual machines.

There are two types of task scheduling techniques[2] centralized task scheduling and distributed task scheduling. There is only a single scheduler in centralized task scheduling and all the mappings or scheduling is done by the single scheduler. In distributed scheduling, There are two or more schedulers and the scheduling is done collectively by those schedulers. Centralized scheduler is easy to implement but it has a demeit of single point failure. Distributed scheduler has high implementation complexity. But processor cycle is saved in distributed scheduler.

Distributed environment has two types of scheduling techniques[2] heuristic and hybrid scheduling algorithms. Heuristics scheduling algorithms uses the knowledge base for taking scheduling decisions and provides an optimal solution. There are two types of heuristic scheduling approach: Static and dynamic scheduling. In Static scheduling techniques, all tasks are known before scheduling and are considered to be arrived at same time and are independent of resource state of the system and availability of resources. In dynamic scheduling, the tasks are scheduled instantly when they arrive. Static scheduling is more easy to implement than dynamic scheduling but dynamic scheduling is more suitable in real world scenario. There are three types of Hybrid algorithms: multi-objective optimization technique, minimization-maximization approach or energy aware methods.

A good scheduling algorithm balances the load of the machines and hence improves the performance of the cloud environment[3]. Static scheduling methods are non preemptive and it minimizes the overall execution time. But static scheduling techniques cannot adapt to the change in load at the run time. During dynamic scheduling(dynamic load balancing), the load is redistributed among the processors during the execution of the task. The redistribution is done by transferring the task from overloaded processors to underloaded processors. It is particularly useful when the resource requirement of process is not known before the scheduling process. Dynamic scheduling increases the performance of the system by balancing the load among the processors and increasing the resource utilization[3].

If the information about the state of the system and the resource need of a process are known then prior to the scheduling, the scheduling policy will be optimal. It is because, when the information is known before scheduling, it can minimize the total process completion time, maximize the resource utilization and maximize the throughput of the system. But in real scenario, the information is not known before scheduling and the task enters system dynamically. So the focus is on sub optimal solution. That is we search for the solution in entire search space and select a good one from that search space as a solution. The solution thus obtained nee not be an optimal solution and hence is known as sub optimal approximate solution. In heuristic based algorithm, the rules will guide the scheduling process in a direction to reach the optimal solution.

III. Task scheduling algorithms

Cloud computing platform has many datacenters which are located geographically in different areas. Each data center has hundreds of servers. There is a datacenter controller and it handles the task submitted by the user. The task is known as cloudlet. A VmLoadBalancer[3] is implemented at data center controller and it assigns the VM to the request processing or to be processed. Hence task scheduling algorithms are implemented in VmLoadBalancer. Some of the important task scheduling algorithms used in cloud environment is discussed below.

Most basic task scheduling algorithm is randomized algorithm[4]. According to this algorithm, the selected task is assigned randomly to available virtual machines. It is the simplest scheduling algorithm. But the demerit is that it does not consider whether the allocated virtual machine is over loaded or under loaded. Thus performance of the task scheduling is compromised using this scheduling policy.

First come first served (FCFS)[5] is one of the scheduling algorithm which is being used more commonly in real world scenario. FCFS algorithm is also known as first in first served. According to this algorithm the task submitted by the client is a service policy whereby the requests of clients are attended in the order of their arrival without any preferences.

Another algorithm used for scheduling is shortest job first scheduling algorithm(SJF)[6]. It is also known as shortest job next(SJN) or shortest process next(SPN). According to this scheduling technique, the job with smallest execution time is selected and given high priority. Jobs are queued based on the priority and are scheduled accordingly. Hence the task with lowest execution time is executed first in this scheduling policy. A demerit of this scheme is that the largest task may have to wait for a long time if there are many smaller tasks.

One of the simplest and most widely used scheduling algorithm is round robin scheduling algorithm[7]. This algorithm is designed especially for time shared systems. All tasks are kept in a circular queue. The scheduler goes around the queue and allocates resource for a time interval of one quantum(time slice for a process to execute). After one quantom, the task is dispatched and added to the tail of the queue if task is still running. Otherwise the process

releases the task.

In order to improve the performance of round robin scheduling algorithm, some improvisations are found for the algorithm. Two such modified round robin algorithms are weighted round robin[8] and dynamic round robin[9]. In weighted round robin, a weight is assigned to the VM available. Hence a powerful VM will get double weight than the weaker one. More tasks are allocated to more weighted Virtual machines in this scheduling technique. The scheduling is done according to round robin algorithm. Dynamic round robin reduce the power consumption of physical machines. In dynamic round robin, the tasks executed in one VM is migrated to other according to the load. The demerit of this scheme is that it does not scale up for large datacenters.

Two heuristics task scheduling algorithm used is min-min[10] scheduling algorithm and max-min[11] scheduling algorithm. Min-min scheduling algorithm[10] selects task with minimum execution time and is assigned to the fastest virtual machine available so far. Hence shortest jobs will be finished first. The major advantage of min-min scheduling algorithm is that it increases the makespan to some extend and overall throughput is increased. But it does not consider the load of the fastest VM to which it is allocated. Max-min algorithm[11] is same as min-min algorithm but the only difference is that it allocates task with maximum execution time to fastest VM. This algorithm does not consider load balancing and there is a chance of starvation for smaller tasks.

The max-min algorithm is modified later and is known as modified max min algorithm[12]. In modified Max min algorithm, The larger task is executed in fastest available VM and smaller tasks are executed in parellel in other possible Virtual machines. This algorithm increases the makespan than simple max-min algorithm. Load balancing is also considerd in this algorithm.

Two varieties of min-min scheduling algorithm is proposed later. Those algorithms are Load Balance Min-Min Scheduling Algorithm(LBIMM)[13] and Load Balance Min-Min(LBMM)[14] algorithms. In LBIMM min-min scheduling[13] is done first. Later rescheduling of the task is done heaviest loaded VM to under loaded one. The reschedule is done by calculating the makespan of min-min scheduling and is compared with the minimum completion time of the task. This algorithm improves load balancing and reduces the overall completion time. LBMM scheduling algorithm[14] also uses min-min scheduling as its base. It is worked based on a three level hierarchical framework. This algorithm reduces the load unbalance of min-min scheduling and minimize the execution time. But it does not specify how to select a node for a complicated task requiring large-scale computation.

Another task scheduling technique used for load balancing is Opportunistic Load Balancing(OLB)[15]. It is a static load balancing technique. According to this algorithm, all nodes in the cloud are kept busy. It dispatches tasks to randomly selected virtual machines without considering the execution time of the task in that node. Hence the makesspan of the algorithm is reduced. Also there may be pending requests caused due to the waiting of tasks for nodes to be free.

A new algorithm is later designed combining opportunistic load balancing(OLB) and Load balanced Min-Min(LBMM) algorithm[16]. It is a two phase scheduling algorithm under the three level cloud computing network. There are two actors in this scheduling framework. One is agent which collects information of each node. Load balancing is done by agent. The second actor is manager which select a node to execute the task baed on the information provided by agent based on a threshold value evaluated from the execution time of the task. This algorithm selects efficient node from many nodes.

A new algorithm called throttled load balancer(TLB)[17] is discovered to schedule task and hence obtain load balancing. In this algorithm, each VM has a state busy or idle. The state of the VM is recorded in TLB. When a task comes to the system, the task checks the table with the configuration to find the best VM. If the best VM is found, the task is allocated to the VM. Otherwise the task is queued.

Later a modified version of TLB is developed[18]. In modified throttled algorithm, the incoming tasks are distributed uniformly among all the available virtual machines in an efficient way. It maintains an index table where each VM and their states are recorded. When a request comes, index table is scanned to find out the available VMs.

Here, we discussed rule based techniques so far. Rule based techniques means a rule is developed for scheduling and the process of scheduling is done according to the rule. Later in order to get more optimal solutions, Scheduling algorithms are developed based on the soft computing techniques. Each iteration in a soft computing algorithm redirects the solution to most optimal one. Some of the task scheduling algorithms developed based on soft computing techniques are discussed below.

Most important soft computing technique used for task scheduling is genetic algorithm[19]. There are four different operations in genetic algorithm approach. they are selection, evaluation, crossover and mutation. Scheduling can be seen as a vector which matches task to VMs. tasks are represented by chromosomes. Fitness value represents overall execution time of each task. Change to the mapping is carried out by two operations-crossover and mutation. It is based or the oldest known theorem- survival of the fittest. It can handle vast search space.

Later improvisations are found out for genetic algorithm.One such algorithm is N-GA[20]. In this scheduling technique, priority queue is used to improve the performance of genetic algorithm. In this algorithm, elitism technique with unusual selections is adopted to evade premature convergence is adopted in order to increase the convergence of the solutions.

The well known LBMM algorithm discussed earlier is improved using genetic algorithmic technique to do the task scheduling more

effectively[21]. The execution time of a task on VM is calculated. This is expressed as Million Instruction(MI) for task and Million Instruction per Second(MIPS) for VMs. Convert MI and MIPS into binary format and is done with crossover and mutation. Fitness function is maximum value. The solution is chosen with highest fitness value and is scheduled accordingly.

Another variety of genetic algorithm is proposed by incorporating Johnson's rule with genetic algorithm[22]. The algorithm is known as Johnson's rule based genetic algorithm(JRGA). This algorithm takes the characteristics of multiprocessor scheduling in cloud datacenters. It solves two stage flow-shop scheduling problem. The major advantages are it has quick response time and schedules highly concurrent tasks efficiently. But the demerit is that it minimizes makespan.

Another soft computing techniques that is being used as task scheduling is simulated annealing[23]. It is similar to genetic algorithm and is an iterative method. It has an initial solution. It is being updated later to find a better solution. The makespan is analysed to check the better solution. Simulated annealing gives solutions poorer than genetic algorithms.

Stochastic hill climbing[24] is also an algorithm similar to simulated annealing. It is one of the incomplete approach to solve task scheduling problem. The algorithm moves continuously to find the increasing value(uphill motion) till there is no more movement possible. It is not a best strategy because entire solution space is not considered to find the solution. Only nearest neighbours are considered.

A load balancing strategy has been developed using ant colony optimization algorithm(ACO)[25]. In this algorithm, the starting point is known as end node. In task scheduling, head node is initial mapping. The ants will move the cloud network to choose the nodes for their next step using the classical formula of ant colony optimization.

Particle swarm optimization algorithm(PSO)[26] is also used for task scheduling and there by load balancing in cloud platform. Each task is a particle in PSO and each particle is characterized by displacement and velocity. Fitness function has values that that are need to be optimized like makespan, response time etc.

Honey bee foraging behaviour can also be used for task scheduling scenario. The algorithm is known as bee colony optimization(BCO)[27]. Overall execution time of the task is reduced by using this algorithm. Load balancing among virtual machines are done effectively by using this algorithm. Objective of this algorithm is to speed up the execution of tasks whose workload varies at time to time in unpredictable way.

Another algorithm developed for task scheduling is based on the behaviour of cats in the environment. The algorithm is known as Cat Swarm Optimization Algorithm(CSO)[28]. It is first discovered for scheduling in distributed system and later extended to task scheduling in cloud scenario. It has two modes: Seeking mode and tracing mode. Tracing mode is similar as PSO and the improvisation is done in seeking mode. Seeking mode is responsible for finding solution in all the search space.

A pareto based fruit fly optimization algorithm[29] is developed for task scheduling and resource utilization. In this algorithm, makespan and total cost is optimized. Basic fruit fly algorithm is to optimize single objective. But here multi objectives are optimized designing search operators and implementing multi objective handling technique. There are two search criteria in this algorithm: Smell based search stage and Vision based search stage.

Another scheduling algorithm is multi objective cuckoo optimization scheduling algorithm[30]. It reduce execution time and allows maximum parallelization. Each cuckoo is a complete solution or schedule. Major advantage is that it achieves solution at lower number of repetition.

Later more researches are done on this soft computing algorithms to do task scheduling and more and more techniques are find out. More techniques are developed by combining two or more soft computing algorithms together. One such algorithm developed is the combination of cat swarm optimization algorithm with simulated annealing[31] and another one is cat swarm optimization algorithm with particle swarm optimization[32]. Such hybrid algorithms have the merits of both algorithms and thus are more efficient. The algorithm developed by combining simulated annealing and CSO is known as scalable multiobjective cat swarm optimization based simulated annealing(CSM-CSOSA)[31]. This algorithm accepts the solution ignored by CSO on a probability basis and hence obtain more appropriate solution. Execution time and cost is minimized using the algorithm. A hybrid of modified CSO and modified PSO is used for developing a scheduling algorithm[32]. According to this algorithm, the performance of PSO is increased when incorporated with CSO. This algorithm is helpful for load balancing and resource utilization.

A hyper heuristic scheduling algorithm is developed later for task scheduling(HHSA)[33]. In HHSA, the strength of all above mentioned algorithm is incorporated into a single algorithm. This algorithm can be applied to both sequence dependent and sequence independent scheduling problems. But one demerit is that heterogeneous environment is not considered for this scheduling algorithm.

Uncertainty principle of game theory[34] can also be used for task scheduling and resource allocation. A demand aware and topology aware scheme is developed using this technology. According to this principle, cloud is modelled as a multi agent system that is composed of many agents with varied capabilities. These agents are responsible for task scheduling.

Another form of scheduling is workflow scheduling or dependant task scheduling. Workflow scheduling means scheduling of a large task by dividing it into smaller subtasks and resources are allocated to each sub task. Many algorithms are developed

for workflow scheduling. One is Discrete Binary Cat Swarm optimization(DBCSO) algorithm[35]. It is similar to CSO. only difference is that the vector comprise of zeros and ones ie in binary format. This algorithm optimizes makespan.

Many algorithms are developed for work flow scheduling. one such algorithm is cost and energy aware scheduling algorithm(CEAS)[36]. CEAS minimizes execution cost of workflow and reduce the energy consumption while meeting deadline constraints. It has five sub algorithms: VM selection algorithm, Sequence task merging algorithm, Parallel task merging algorithm, VM reusing algorithm, Task slacking algorithm.

EnReal[37] and Cloudy knapsack algorithms[38] are also developed for workflow scheduling. EnReal[37] is an energy aware method for resource allocation in scientific workflows. Dynamic deployment of VM are done in EnReal for scientific workflows. In cloudy Knapsack algorithm[38], it proposes a distributed method for offloading task from distributed systems to cloud system. This algorithm improves performance of distributed computing systems.

IV. Conclusion

The performance of a cloud computing system always depends on an efficient task scheduling algorithms. A study of existing task scheduling algorithms developed so far is done in this paper. The paper discusses both rule based techniques and soft computing techniques developed for task scheduling in cloud. Algorithms for workflow scheduling is also discussed in the paper. A good scheduling algorithm always consider the need of user and will not give any overhead to the cloud systems. A good scheduling algorithm solves the problem of load unbalancing and helps in resource utilization. There by it increases the performance of the system.

References

[1] Liu, F., Tong, J., Mao, J., Bohn, R., Messina, J., Badger, L., Leaf,D, *NIST cloud computing reference architecture. NIST special publication*, 2011.

[2] Mathew, Teena, K. Chandra Sekaran and John Jose. *Study and Analysis of Various Task Scheduling Algorithms in the Cloud Computing Environment*, Advances in computing,Communications and Informatics(ICACCI), 2014, pp. 658-664.

[3] S. B. Shaw,Dr. A.K. Singh, *A Survey on Scheduling and Load Balancing Techniques in Cloud Computing Environment*,International Conference on Computer and Communication Technology (ICCCT),2014.

[4] Isam Azawi Mohialdeen , *Comparative Study of Scheduling Algorithms in Cloud Computing Environment* , Journal of Computer Science , 2013, pp. 252-263.

[5] Huang, L., Chen, H. S., Hu, T. T., *Survey on Resource Allocation Policy and Job Scheduling Algorithms of Cloud Computing*, JSW, 2013, 8(2), 480-487.

[6] Van den Bossche, R., Vanmechelen, K., Broeckhove, J., *Cost-optimal scheduling in hybrid iaas clouds for deadline constrained workloads.* In Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference, 2010, (pp. 228-235).

[7] Amandeep Kaur Sidhu, Supriya Kinger ,*Analysis of Load Balancing Techniques in Cloud Computing*, International Journal of Computers & Technology , Volume 4 No. 2, March-April, 2013, pp. 737-741.

[8] Qi Zhang, Lu Cheng, Raouf Boutaba, *Cloud computing: sate-of-art and research challenges*, Published online: 20th April 2010, Copyright : The Brazillian Computer Society 2010.

[9] Jayant Adhikari, Prof. Sulabha Patil, *Double Threshold Energy Aware Load Balancing in Cloud Computing*, 4th ICCCNT, July 2013.

[10] Archana mantri, Suman Nandi, Gaurav Kumar, Sandeep Kumar,*High Performance Architecture and Grid Computing Computing*, International Conference HPAGC 2011, Chandigarh, India, July 2011, Proceedings.

[11] M.Aruna , D. Bhanu, R.Punithagowri , *A Survey on Load Balancing Algorithms in Cloud Environment*, International Journal of Computer Applications, Volume 82 âĂŞ No 16, November 2013 , pp. 39-43.

[12] Rajwinder Kaur, Pawan Luthra, *Load Balancing in Cloud System using Max Min and Min Min Algorithm*, National Conference on Emerging Trends in Computer Technology (NCETCT), 2014

[13] Huankai Chen, Professor Frank Wang, Dr Na Helian, Gbola Akanmu, *User-Priority Guided Min-Min Scheduling Algorithm For Load Balancing in Cloud Computing*, IEEE, 2013.

[14] Shu-Ching Wang, Kuo-Qin Yan, Wen-Pin Liao, Shun-Sheng Wang, *Towards a Load Balancing in a Three-level Cloud Computing Network*, 3rd International Conference on Computer Science and Information Technology (ICCSIT), Vol. 1, IEEE, 2010, pp. 108-113.

[15] Klaithem Al Nuaimi, Nader Mohamed, Mariam Al Nuaimi and Jameela Al-Jaroodi , *A Survey of Load Balancing in Cloud Computing: Challenges and Algorithms*, IEEE Second Symposium on Network Cloud Computing and Applications , 2012, pp. 137-142.

[16] Shu-Ching Wang, Kuo-Qin Yan, Wen-Pin Liao and Shun-Sheng Wang, *Towards a Load Balancing in a Three-level Cloud Computing Network*, 2010.

[17] Jasmin James, Dr. Bhupendra Verma, *Efficient VM Load Balancing Algorithm for a Cloud Computing Environment* , International Journal on Computer Science and Engineering (IJCSE) , Vol. 4 No. 09 Sep 2012 , pp. 1658-1663.

[18] Shridhar G. Domanal, G. Ram Mohana Reddy, *Load Balancing in Cloud Computing Using Modified Throttled Algorithm*, IEEE, International conference. CCEM 2013.

[19] Kousik Dasgupta, Brototi Mandal, Paramartha Dutta, Jyotsna Kumar Mondal, Santanu Dam, *A Genetic Algorithm (GA) based Load Balancing Strategy for Cloud Computing*, International Conference on Computational Intelligence: Modeling Techniques and Applications(CIMTA), 2013.

[20] Bahman Keshanchi, Alireza Souri, Nima Jafari Navimipour, *n improved genetic algorithm for task scheduling in the cloud environments using the priority queues: formal verification, simulation, and statistical testing*, The Journal of Systems &

Software, 2016.

[21] Shyam Singh Rajput, Virendra Singh Kushwah, *A Genetic based Improved Load Balanced Min-Min Task Scheduling Algorithm for Load Balancing in Cloud Computing*, 8th International Conference on Computational Intelligence and Communication Networks, 2016.

[22] Yonghua Xiong, Suzhen Huang, Min Wu, Jinhua She, Keyuan Jiang, *A JohnsonâĂŹs-Rule-Based Genetic Algorithm for Two-Stage-Task Scheduling Problem in Data-Centers of Cloud Computing*, Journal Of LaTeXClass Files, Vol. 13, No. 9, 2014.

[23] M. Coli, P. Palazzari, *Real Time Pipelined System Design through Simulated Annealing,* Journal of Systems Architecture, vol.42, no. 6-7, 1996, pp. 465-475.

[24] Brototi Mondal , Kousik Dasgupta, Paramartha Dutta, *Load Balancing in Cloud Computing using Stochastic Hill Climbing-A Soft Computing Approach*, Procedia Technology 4 (2012) 783 âĂŞ 789, 2012.

[25] Kumar Nishant, Pratik Sharma, Vishal Krishna, Nitin and Ravi Rastogi, *Load Balancing of Nodes in Cloud Using Ant Colony Optimization*, 14th International Conference on Modelling and Simulation, 2014.

[26] Xingquan Zuo, Guoxiang Zhang, and Wei Tan, *Self-Adaptive Learning PSO-Based Deadline Constrained Task Scheduling for Hybrid IaaS Cloud*, IEEE Transactions On Automation Science And Engineering, Vol. 11, No. 2, 2014

[27] Dhinesh Babu L.D., P. Venkata Krishna, *Honey bee behavior inspired load balancing of tasks in cloud computing environments*, Applied Soft Computing 13 (2013) 2292âĂŞ2303, 2013.

[28] Reza Shojaee, Hamid Reza Faragardi, Sara Alaee and Nasser Yazdani, *A New Cat Swarm Optimization based Algorithm for Reliability-Oriented Task Allocation in Distributed Systems*, 6'th International Symposium on Telecommunications, 2012.

[29] Xiao-long Zheng, Ling Wang, *A Pareto based Fruit Fly Optimization Algorithm for Task Scheduling and Resource Allocation in Cloud Computing Environment*, 2016.

[30] Mehdi Akbari, Hassan Rashidi, *A multi-objectives scheduling algorithm based on cuckoo optimization for task allocation problem at compile time in heterogeneous systems*, Expert Systems With Applications 60 (2016) 234âĂŞ248, 2016.

[31] Danlami Gabi, Abdul Samad Ismail, Anazida Zainal, Zalmiyah Zakaria, *Scalability-aware Scheduling Optimization Algorithm for Multi-Objective Cloud Task Scheduling Problem*, 2017.

[32] Shridhar Domanal, Ram Mohana Reddy Guddeti, Rajkumar Buyya, *A Hybrid Bio-Inspired Algorithm for Scheduling and Resource Management in Cloud Environment*, IEEE Transactions On Services Computing, 2016.

[33] Chun-Wei Tsai, Wei-Cheng Huang, Meng-Hsiu Chiang, Ming-Chao Chiang, and Chu-Sing Yang, *A Hyper-Heuristic Scheduling Algorithm for Cloud*, IEEE Transactions On Cloud Computing, Vol. 2,No. 2, 2014

[34] Parvathy S. Pillai, Shrisha Rao, *Resource Allocation in Cloud Computing Using the Uncertainty Principle of Game Theory*, IEEE Systems Journal, Vol. 10, No. 2, 2016.

[35] Bhopender Kumar, Mala Kalra, Poonam Singh, *Discrete Binary Cat Swarm Optimization for Scheduling Workflow Applications in Cloud Systems*, 3rd IEEE International Conference on "Computational Intelligence and Communication Technology" (IEEE-CICT), 2017.

[36] Zhongjin Li, Jidong Ge, Haiyang Hu, Wei Song, Hao Hu, Bin Luo, *Cost and Energy Aware Scheduling Algorithm for Scientific Workflows with Deadline Constraint in Clouds*, IEEE Transactions On Services Computing, 2015.

[37] Xiaolong Xu, Wanchun Dou, Xuyun Zhang, and Jinjun Chen, *EnReal: An Energy-Aware Resource Allocation Method for Scientific Workflow Executions in Cloud Environment*, IEEE Transactions On Cloud Computing, Vol. 4, No. 2, 2016.

[38] Harisankar Haridas, Sriram Kailasam, Janakiram Dharanipragada, *Cloudy Knapsack Algorithm for Offloading Tasks from Large Scale Distributed Applications*, IEEE Transactions On Cloud Computing, 2017.