IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

PARTICAL SWARM OPTIMIZATION OF TASK SCHEDULING IN CLOUD COMPUTING

Payal Jaglan*, Chander Diwakar
* Department of CSE UIET, KUK Kurukshetra, India
Department of CSE UIET, KUK Kurukshetra, India

DOI:

ABSTRACT

Resource provisioning and pricing modeling in cloud computing makes it an inevitable technology both on developer and consumer end. Easy accessibility of software and freedom of hardware configuration increase its demand in IT industry. It's ability to provide a user-friendly environment, software independence, quality, pricing index and easy accessibility of infrastructure via internet. Task scheduling plays an important role in cloud computing systems. Task scheduling in cloud computing means that to allocate best suitable resources for the task to be execute with the consideration of different parameters like cost, time, scalability, make span, reliability, availability, throughput, resource utilization and so on. Scheduling of tasks requires a set of rules and regulations which sets an agreement between users and providers of cloud. Task scheduling is the optimal sizing of set of subtasks from the available sources.

KEYWORDS: Cloud Architecture, Cloud Computing, Credit, Deadline, Task Scheduling.

INTRODUCTION

Cloud computing is a new computing mode. It is similar to utility computing which involves the large number of computers connected through a communication network. It has advantage of delivered a flexible, on-demand service, very high performance, pay-as-you-go. Operators should have guarantee to the subscribers and stick to the Service Level Agreement[1]. If the job spanning is too long, it leads to dissatisfaction. To avoid hotspot and improve resource utility, the load is needed to be balanced among servers using this technique. To well solve load balancing, every cloud giants have their own solutions. Google adopts Map-Reduce scheduling mechanism scheduling algorithms are relatively simple, FIFO, default algorithm performs not so well for short jobs. Besides, Yahoo raises computation ability scheduler, Facebook proposes fair share scheduler. However, these scheduling algorithms cannot work out a better scheduling scheme. In fact, tasks scheduling in cloud is a NP complement problem with time limit. That is an to say, it is seldom impossible to search out a reasonable solution in polynomial time. To improve performance of cloud computing, efficient task scheduling and resource management is required. Genetic algorithm as the heuristic algorithm shows the special advantage in the combinatorial optimization. Gas is robust search and optimization techniques in a number of the practical problems, due to its capacity to locate global optimum in a multimodal landscape [2].

A.1 Cloud Computing Deployment Models:

A cloud deployment model represents a specific type of the cloud environment, firstly distinguished by ownership, size and access. The basic types of cloud deployment models are defined below:

a) Private Cloud: This type of cloud provides its service to department of large organization (single organization) which is managed by either third party or by same organization.
b) Public Cloud: This type of cloud is organized by a cloud service selling organization over network to provide service as per pay-as-go model.
c) Hybrid Cloud: This type of cloud is organized by a cloud service selling organization over network to provide service as per pay-as-go model.

A.2 Service Model in Cloud Computing:

Cloud Computing is service based technology which includes hardware, software, storage etc. Its services are categorized among few models of services [3] as shown in list below:

a) *Software as a Service (SaaS):* Top-level of service provided to consumer which includes software's like development tools, email, games, document processing, communication etc. Google Drive, Email, Drop box are few examples to SaaS for the consumer where user need to pay the use of services or on the bases of subscription.

b) *Platform as a Service (PaaS):* Middle layer of service which provides environments to hold SaaS. It provides platform to developers (means end users have capability to developed) write and execute there code on extensive range of environments.

c) *Infrastructure as a Service (IaaS):* It is Base level of service to cloud consumer which provides hardware on demand like servers, network and load balancers. Amazon (Ec2, S3 etc.), Rackspace Mosses Offering, Sun's cloud Services,

Cloud Computing Architecture

Front and back end are the main significant components of Cloud computing architecture. End which is visible to the user of cloud is front end. It has applications and computer that user uses to access the cloud. Storage devices and computers are the back end of the cloud computing. Figure 1 shows the graphical view of CC architecture.

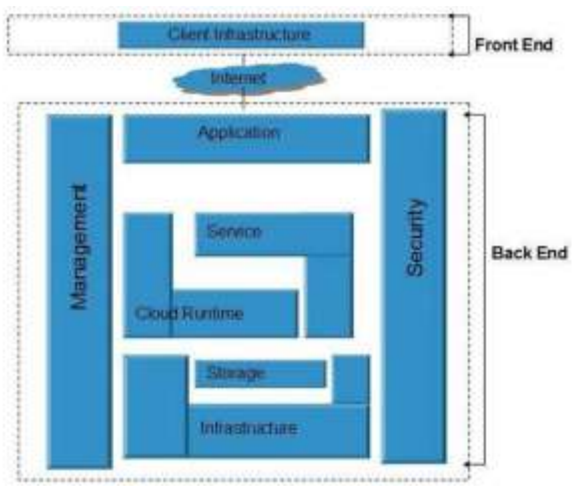

Figure 1: Cloud Computing Architecture

Cloud Architectures address key difficulties that mainly related to processing of large amount of data. In customary way of giving out the data it is quite difficult to get as many machines that an application needs for its operation to complete. It is one of the most difficult things to get the machines when one/particular desires them. It is really a hard work to allocate and co-ordinate a large scale job on different-different machines, run procedures on them, and supplies another machine to recuperate if one machine flops during the operation. It is also problematic to auto-scale up and down based on varying nature of workloads. It is problematic to get rid of all those machines when the job is completed.

TASK SCHEDULING

Task Scheduling plays a key role in the CC system. Scheduling of the task cannot be done on basis of single criteria but under a lot of rules and regulations that can term as an agreement between users and providers of cloud. Providing good quality of services to the users or clients according to the agreement it is a decisive task for the providers at the same time there is large number of tasks running at the side of provider's. The task scheduling problem can be viewed or seen as the finding or searching an optimal mapping of set of subtasks of different tasks over the available set of the resources (e.g. processors/computer machines) hence can be achieved or attained the desired goals for tasks. Scheduling is a method by which data flow, threads and processes are given access to system resources. Scheduling is the fundamental operating system function, almost all computer resources are scheduled before use [4].

Task Scheduling Types

Task Scheduling is the method by which data flow, threads and processes are given access to system resources. There are various types of task scheduling which are discussed below [5][6]

a) *Preemptive Scheduling* New process selected to run also when an interrupt occur when new processes become ready. Tasks are usually allotted with priorities. At periods it is compulsory to run a definite task that has a advanced priority before another task even though it is running. Consequently, the r unning task is intermittent for some time and continued later when the priority task has done its execution or implementation. This is called preemptive scheduling.

b) *Non-preemptive Scheduling:* A new process is selected to run either when a process terminates or when an explicit system request causes a wait state. In sort in non-preemptive scheduling, a running task is executed till completion. It cannot be intermitted. (e.g., I/O or wait for child).

c) *Round Robin Scheduling(RR):* The round robin scheduling algorithm is planned or designed specifically for time-sharing systems or structures. It is related to FCFS scheduling, but pre-emption is additional to switch in the middle of processes.

RELATED WORK

Selvarani and Sudha [7], proposed a computing platform of the task schedule groups which have different computation performances and resource costs for their resources. In this job scheduling in CC environment have done research and analysis, which aims at task scheduling with minimum cost and minimum of total tasks completion time. The results are then compared with an Activity based of costing algorithm.

Wang and Wang [8], proposed the Energy-efficient Multi-task Scheduling based on Map Reduce on CC and selected a design a method of practical encoding and decoding for the individuals and after construct an overall energy efficiency function as the fitness value of the individual of the servers. In order to enhance the searching ability of their algorithm and to accelerate the convergent speed to introduced a local search operator.

Jiong et al. [9], considers independent tasks scheduling in cloud computing of bi objective minimization problem with makespan and energy consumption of scheduling criteria. Used Dynamic Voltage Scaling to minimize energy consumption and proposed two algorithms. These algorithms use the methods of unify and double fitness to define the fitness function and select the individuals, adopt the genetic algorithm to parallel find the reasonable scheduling scheme.

Wang et al. [10], dynamic feedback algorithm, every task has a weight value, expressed by a vector, in order to express the different request of the different resource. Quantization analysis of the resource request is used to select the best node to accomplish the task, and the optimization selection can achieves the purpose that improving the usage of the resource.

Guo et al. [11], discussed about the Task Scheduling Optimization in CC based on Heuristic Algorithm. In order to minimize or reduce the cost of the processing formulates a model for task scheduling and a Particle Swarm Optimization algorithm that is based on small position value rule. The experiment results prove that the PSO algorithm is more suitable to CC and presented the task scheduling optimizing method in CC, and then formulate a model for task scheduling to minimize the cost of the problem after that solved it by a PSO algorithm. By comparing and analyzing particle swarm algorithm with crossover, mutation and local search algorithm based on particle swarm, proposed the particle swarm algorithm embed in SPV, which represents better performance.

Sathya [12], discussed Deadline Based Task Scheduling in Cloud with Effective Provisioning Cost using the Location based Minimum Migration in Cloud (LBMMC) Algorithm. To reduce the impact of performance variation of the cloud resources in the deadlines of workflows, an algorithm called EIPR, takes into consideration the behavior of cloud resources during the scheduling process and also applies replication of tasks to increase the chance of meeting application deadlines.

Lakra and Yadav [13], proposed Multi-Objective Tasks Scheduling Algorithm for CC Throughput Optimization. In this a multi-objective task scheduling algorithm form mapping tasks to Vms in order to improves the throughput of the datacenter and then reduce cost without the violating SLA for an application in cloud SaaS environment. Most of the algorithms schedule tasks based on single criteria but in cloud environment required to consider various criteria like execution time, cost etc. This algorithm is simulated using CloudSim simulator and result shows the better performance and improved the throughput.

Kun-lun et al. [14], proposed an Improved GEP Algorithm for Task Scheduling in CC. The traditional Genetic Algorithm for task scheduling have the defect of premature convergence, which only takes the time cost into consideration, and ignores the consumption of resources. To solve problem that exists in multi-task scheduling in CC mentioned above, propose an improved GEP algorithm with double fitness functions and also constructs a new ETCC matrix.

PROPOSED WORK

The inspiration of our model is to allocate tasks to virtual machines with considering reliability. Proposed model consists of five phases. Model phase's concepts are:

a) Task Buffer: There are millions of users require to execute tasks in thecloud computing. Task buffer is responsible for collecting tasks from user.
b) Task Information: This phase provides the necessary information of Tasks arrived into cloud computing environment for execution. Those information such as Expected Execution Time (EET), Expected Transmission Time(ETT), Resources-Required (RR) and Round Trip Time(RTT) .
c) Resource Information:This phase responsible collects information about resources in CC environment. The resources in cloud computing are Datacenter, Hosts and virtual machines (VMs). Datacenter information is host list, VMs list, storage list and cost of memory, cost of BW and other information. Each host can contain more than one VM.
d) LBMPSO: Load balancing mutation PSO used to reschedule tasks that failure to schedule. PSO have two problems. Firstproblem, tasks may failure to allocate to virtual machine. Second problem, task may allocate to more than one VM. In this phase solve the problems by reschedule wrong tasks and take in account load balancing of virtual machine. Solving these problems help to achieve reliability, users assert task executed without failure, minimize execution minimize round trip time and improve other parameters**.**
e) Task Submission: This phase responsible receives allocation plan from previous phase. Then, allocates each task to virtual machine based on plan.

Task Scheduling Problem Formulation of proposed systems

There are several tasks (t) and several virtual machines. There are n tasks and m number of virtual machines. Figure2 shows mapping of the Tasks to VM. PSO attempts to select optimal distribution of tasks to VM for achieving objective. 3 mathematical models proposed for task scheduling. Each model consists of objective function and several constraints. Objective function of first model is to minimize execution time based on expected execution time (EETij) of task i in vmj. Eq.(1) used to calculate processing time: EET(processing time) = lengthi / mipsj. lengthi is number of instruction of task i require to execute. mipsj is number of Instructions executed by vm per second. 2nd objective function is to minimize transmission time (ETRTij). Expected transmission time (ETRTij) of task i to vm j responsible for achieving 2nd objective function. ETRTij equals file size / bandwidth. To minimize round trip time (RTT) achieved by 3rd mathematical model. The RTT is the (latency) time for the whole procedure involving the sending and the receiving. ERTTij is expected round trip time calculate by ETRTij + delay + EETij + delay. xij is allocating task i to vm j or not. The value of xij may one or zero. Each model has the same constraints. Each Task allocate to only one virtual machine achieve by first constraint in (2). Eq. (3) and (4) represent resource of all virtual machine less than or equal resource of datacenter. xij assign positive number (5)

1st Mathematical Model Based On Expected Execution Time

$$\text{Min } z = \sum_{i=0}^{n} \sum_{j=0}^{m} EET_{ij} * x_{ij} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1).$$

Subject To

$$\sum_{j=0}^{m} x_{ij} = 1 \text{ ¥ } i \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2).$$

$$\sum_{j=0}^{m} CPU_j \leq TOTALCPU \ldots\ldots\ldots\ldots\ldots (3).$$

$$\sum_{j=0}^{m} MEM_j \leq TOTALMEM \ldots\ldots\ldots\ldots\ldots (4).$$

Objective Function of 2nd Mathematical Model Based on Expected Transmission Time

j = 0

$$x_{ij} \geq 0 \text{ ¥ } i, j \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5).$$

$$\text{Min } z = \sum_{i=0}^{n} \sum_{j=0}^{m} ETRT_{ij} * x_{ij} \ldots\ldots\ldots\ldots\ldots\ldots (6).$$

Objective Function of 3rd Mathematical Model Based on Expected Round Trip Time

$$\text{Min } z = \sum_{i=0}^{n} \sum_{j=0}^{m} ERRT_{ij} * x_{ij} \ldots\ldots\ldots\ldots\ldots\ldots (7).$$

X11 Tasks Virtual Machines

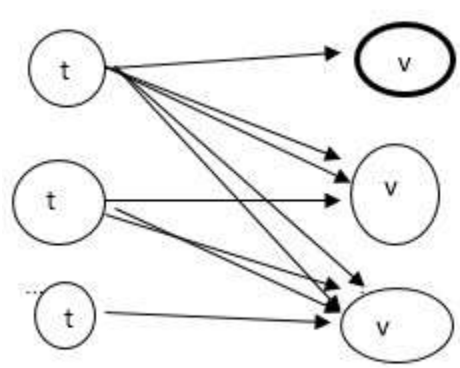

Fig.2: tasks mapping to virtual machine

load balancing mutation Particle Swarm Optimization (LBMPSO):

PSO algo. a population-based search algo. which is based on simulation of social behavior of the birds within the flock and fish school proposed by Kennedy. The notation adopted in this paper during the d-dimensional search space, each particle in this space defined the potential solution to a problem, i.e. i^{th} particle of the swarm represented Xi=Xi1,Xi2…Xid and its velocity defined Vi= Vi1, Vi2..Vid. The update the particles at each generation. In the iteration t, velocity Vi (t) has been update to pull the particle i^{th} towards its own best position xpi and the best position for all the particles xg that has the best fitness value until the preceding generation. Also it is observe, the current velocity of each iteration t based on vi (t-1) is the velocity of the pervious iteration, r1, r2 mean a uniform random variables between zero and one these 2 random values generated independently, c1, c2 are a positive constant, and w is the inertia weight. Eq. (9) updates each particle's position in the solution hyperspace using the computed Vi (t) and the coefficients c and d that could be set to unity without loss of generality. Particle Swarm Optimization was used to allocate tasks to Vms but, there are some problems. 1st some task doesn't allocate to vm. 2nd problem some tasks allocate to more than one vm. 3rd problem is premature convergence. Load balancing mutation added to PSO to solve previous problem. Load balancing mutation improved in other parameters such as minimize make span, minimize execution time, minimize round trip time and minimize cost. The idea of LBMPSO reschedule the failure tasks to the available (VM) with take into account load of each vm. LBM guarantee all vm executed number of tasks appropriate with their load of vm. In LBM, 1st Determine failure tasks .2nd calculate load of virtual machines as load of Vmi=(resource of Vmi /total resource)*N. 3rd sort tasks based on resource needed and sort Vms based on load. Last Reschedule failure tasks to Vm based on load of each Vm

$$v_i^{k+1} = wv_i^{k} + c r \times (pbest_i - x_i^{k}) + c_2 r_2 \times (gbest - x_i^{k}) \ldots\ldots (8).$$

$$x_i^{k+1} = x_i^{k} + v_i^{k+1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (9).$$

Algorithm: *Load Balancing Mutation Algorithm*

```
        For
 all task ( ti )  Є T do
        Determined unallocated tasks
        Determined tasks allocated to more than one vm
  End
      For all virtual machine (Vmi ) ϵ VM
  Do
        Determined current tasks allocated to Vmi (current load)
          Determined real load of vmi(real load vm)
  End
     For Sort Vm based on real load
         Sort wrong tasks based on resource needed
        for all sorted virtual machine (svmi ) Є VM
  Do
     For all sorted task (sti ) Є T
 If
        real load vm > current load vm
        Schedule task from wrong tasks
          Remove task from sorted tasks list
       Current load vm++
     End;
```

Experiment Setup:

Cloudsim used to experiment proposed algorithm and compared with longest Vm longest cloudlet algorithm, random algorithm, mutation PSO without consider load balancing and standard PSO. The experiments are implemented with 5 VMs and 20 tasks. We evaluated the scheduling heuristic using independent task to each other. The following experiments, the parameters the average execution time, average cost, average round trip time and average makespan used in comparison between different algorithms. We compared between round trip time load balancing mutation pso, round trip time random,,round trip time mutation pso, transmission time load balancing mutation pso, transmission time random, transmission time Longest Cloudlet to Fastest Processor, transmission time pso, execution time mutation pso, time load balancing mutation pso, execution time random, The result of comparisons between different algorithms to reduce provider profit and user loss of existing and proposed PSO. The conclusions show that LBMPSO in third formula based on round trip time the best algorithm which minimizes round trip time, execution time, cost and also consider load balancing and achieve availability and reliability.

C.1 Provider Profit/User Loss

a) *Provider Profit:* Provider profit is depend on lateness if lateness of cloudlets is less than zero then profit is calculated. Proposed algorithm capable to process soft deadline cloudlets. Such cloudlets have penalties when they miss their deadline. Figure 3 results indicate that profit in our approach is maximum as compared to other policy. In the given figure 3, x-axis algorithms and y-axis shows the net profit in terms of provider profit of the cloudlets.
b) *User Loss:* User loss is also depend upon lateness of the task in which lateness greater than zero provide user loss. Figure 3 results indicate that user loss in our approach is minimums compared to other policy. In the given figure 3, x-axis algorithms and y-axis shows the user loss of the cloudlets. In this case user loss is noticeable but negligible if we compared it with existing algorithm.

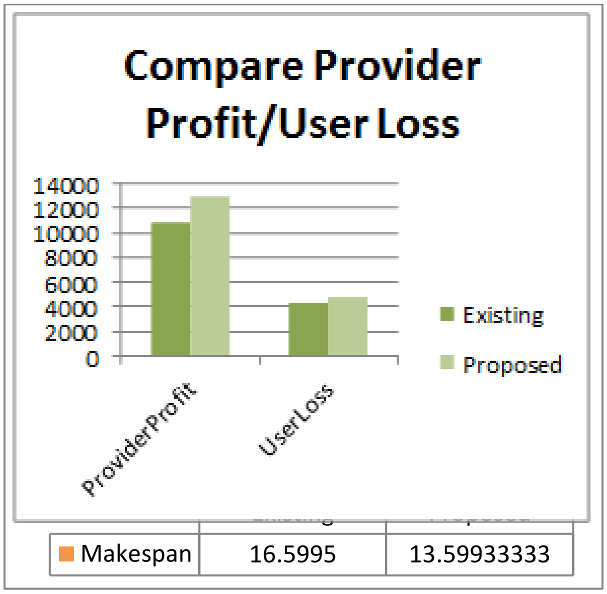

Fig. 2: Comparison b/w Provider profit and User loss of PSO

CONCLUSION AND FUTURE SCOPE

Cloud computing is a new technology wide studied in recent years. Currently there are several cloud platforms that are employed in each in trade and in educational. The way to use these platforms could be a huge issue. During this, have a tendency to delineate the definition, styles, and characteristics of cloud computing, cloud computing services, readying model and challenges of cloud computing. There are several issues in cloud computing. As an example of cloud computing issues is ability, Performance, Service Level Agreement (SLA), knowledge Confidentiality and measurability, knowledge Integrity, Load equalization, Synchronization in numerous clusters in cloud platform, and standardization, the protection of cloud platform. There are some functionality can be integrated with the approach of task scheduling by which service provider can be benefited more. It is used to minimize cost, minimize round trip time, minimize execution time, minimize transmission time, achieve load balancing between tasks and virtual machine, consider available resource and minimize the complexity in cloud computing environment. It improves the Reliability of cloud computing and good distribution of tasks onto resources compared to other algorithms. We found that round trip time load balancing mutation PSO can achieve the best compared to other algorithms. It can be used for any number of tasks and resources.

REFERENCES

[1] Yashpal Singh Jadeja and Kirit Modi, "Cloud Computing - Concepts, Architecture and Challenges", in the proceedings of 4th International Conference on Computing, Electronics and Electrical Technologies, 2012.
[2] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, Cloud Computing and emerging IT Platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems, 2009, pp. 599-616.
[3] Alexa Huth and James Cebula, "The Basics of Cloud Computing", Carnegie Mellon University, Produced for US-CERT, a government organization, 2011.
[4] Tayal, Sandeep. "Tasks scheduling optimization for the Cloud Computing systems." International Journal of Advanced Engineering Sciences and Technologies, Vol. 2, 2011, pp. 111-115.
[5] Kamal Kc, and Kemafor Anyanwu. "Scheduling hadoop jobs to meet deadlines." in the proceedings of IEEE 2nd International Conference on Cloud Computing, 2010, pp. 388-392.
[6] Mladen A. Vouk "Cloud Computing – Issues, Research and Implementations" in the proceedings of Journal of Computing and Information and Technology, 2008, pp. 235–246.
[7] S.Selvarani and G.SudhaSadhasivam, "Improved Cost-BasedAlgorithmFor Task Scheduling in Cloud Computing" in the proceedings of 4th IEEE International Conference on Task Scheduling, 2010, pp.-372-377.
[8] XiaoliWang and YupingWang, "Energy-efficient Multi-task Scheduling based on Map Reduce for Cloud Computing" in the proceedings of 7th IEEE International conference on Computational Intelligence and Security, 2011 pp. 57-63.
[9] Ying Chang-tian and Yu-Jiong "Energy-aware Genetic Algorithms for Task Scheduling in Cloud Computing" in the proceedings of 7th IEEE China Grid Annual Conference, 2012, pp. 43-48.
[10] WANG Han, TANG Xiao-qi, SONG Bao and TANG Yu-zhi "Dynamic Task-Scheduling Algorithm in CNC System Based on Cloud Computing" in the proceedings of 2nd IEEE International Conference on Instrumentation & Measurement Computer Communication and Control, 2012, pp. 1508-1512.
[11] Lizheng-Guo1, Shuguang-Zhao1, Shigen-Shen1, Changyuan-Jiang1 "Task Scheduling Optimization in Cloud Computing Based on Heuristic Algorithm" in Donghua University, Shanghai 201620, China, 2012.
[12] K. Sathya and S.Rajalakshmi "Deadline Based Task Scheduling in Cloud with Effective Provisioning Cost using LBMMC Algorithm", International Journal on Engineering Technology and Sciences, Vol. 1(7), 2014, pp.167-172.
[13] Atul Vikas Lakra and Dharmendra Kumar Yadav "Multi-Objective Tasks Scheduling Algorithm for Cloud Computing Throughput Optimization" in the proceedings of IEEE International Conference on Intelligent Computing, Communication & Convergence, 2014, pp. 321-327.
[14] D.I. George Amalarethinam and T. Lucia Agnes Beena "Customer Facilitated Cost-based Scheduling (CFCSC) in Cloud" in the proceedings of International Conference on Information and Communication Technologies, 2014, pp. 660-667.
[15] Liping Zhang, Weiqin-Tong and Sheng-peng Lu "Task scheduling of Cloud Computing based on Improved CHC algorithm" in the proceedings of IEEE International Conference, 2014, pp. 574-577.
[16] A.I.Awad, N.A.El-Hefnawy and H.M.Abdel kader "Enhanced Particle Swarm Optimization For Task Scheduling In Cloud Computing Environments" in the proceedings of International Conference on Communication, Management and Information Technology, 2015, pp. 920-929.