International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 3 | ISSN : 2456-3307

Optimized Scheduling Procedure for Enhancing Resource Utilization in Heterogeneous Parallel Environment

Lovejoban Preet Singh[1], Anil Kumar[2]
[1]Master of computer science, G.N.D.U, Amritsar, Punjab, India
[2] Professor, G.N.D.U, Amritsar, Punjab, India

ABSTRACT

Because of a phenomenal increment in the quantity of computing assets in various associations, compelling jobs scheduling algorithms are required for proficient asset use. Job scheduling in considered as NP difficult issue in parallel and disseminated registering situations, for example, group, matrix and mists. Metaheuristics, for example, Genetic Algorithms, Ant Colony Optimization, Artificial Bee Colony, Cuckoo Search, Firefly Algorithm, Bat Algorithm and so on are utilized by researchers to get close ideal answers for work scheduling issues. These metaheuristic algorithms are utilized to plan distinctive sorts of jobs, for example, BSP, Workflow and DAG, Independent undertakings and Bag-of-Tasks. This paper is an endeavor to give exhaustive review of prominent nature-enlivened metaheuristic procedures which are utilized to plan distinctive classifications of jobs to accomplish certain execution targets.

Specifications Table

Subject area	computer
More specific subject area	Parallel computing
Type of data	text file,figure,table
How data was acquired	Survey.
Data format	filtered, analyzed

Keywords : ACO,BAT,Cuckoo,Geneticalgorithm,Firefly,Shortest Job First,Minmin,Maxmin,Priority Scheduling.

I. INTRODUCTION

Cloud computing in modern era provides way of using resources without their physical presence at source[1]. The service provided by the cloud is at the front end of computing and internet is at back end. In other words internet is heart and soul of cloud computing. Cloud computing provide mechanism for the users to perform operations that required heavy resources not possessed by them at pay per use basis. With the rapid development of hardware and software cloud computing brings the revolution in the business industry. It provides resources like computational power, storage, computation platform ad applications to user on demand through internet. Some of the cloud providers are Amazon, IBM, Google, Sales force, Microsoft etc. Cloud computing features included resource sharing, multi-tenancy, remote data storage etc. but it challenges the security system to secure, protect and process the data which is the property of the individual, enterprises and governments[2]. Even though, there is no requirement of knowledge or expertise to control the infrastructure of clouds; it is abstract to the user. It is a service of an Internet with high scalability, quality of service, higher throughput and high computing power. Cloud computing providers deploy common online business applications which are accessed from servers through web browser.[3].Before discussing

scheduling mechanisms we discuss services provided by cloud along with types of cloud.

1.1 Cloud Services

There exist legion of services associated with cloud. these services are as described below

IaaS

Infrastructure as services is critical services provided through cloud. virtualized computing resources are provided by the application of IaaS. Internet is key element with which IaaS is accessed. Cast is encountered on the basis of usage. [4]

PaaS

Platform as a service is another cloud service that enhance the organizational applications. Large number of applications exits that are supposed to execute over the distinct machines. All the applications has distinct requirements in terms of platform. This platform requirement is accomplished using cloud computing. Cost is encountered on the basis of time period for which platform is online[5].

SaaS

Software as a service is another critical service supported through cloud. cloud computing host software which can be accessed by users having access to cloud. in other words, machines having limited resources can use SaaS to access software's that they don't possess. [6]

1.2 Job Scheduling

Job scheduling is a procedure in which jobs apportioned to the PC are scheduled by utilizing single cluster or multi cluster approach. The job allocation process is done in light of constrained resources. In both the approaches we have different algorithms to solve the jobs like FCFS, SJF, Round Robin, and Priority Scheduling in single cluster job. Furthermore, in case of multi cluster job we can use algorithms like ACO, Firefly, Honey Bee, BAT, Cuckoo etc. are used to allocate resources so that jobs can be execute in efficient manner and optimal results can be achieves in short span of time. [7]

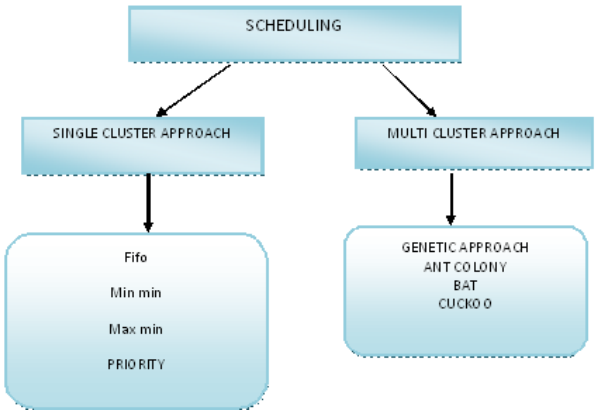

Figure 1. Scheduling in single and Multi-Cluster Environment

1.3 SCHEDULING UNDER SIMPLE ENVIORNMENT

The multi cluster environment originates from single cluster homogeneous environment. The algorithm used under this environment includes

FIFO Job Scheduling

[8] In this algorithm distributed processor within the system on the first come first serve basis. This algorithm gives some time optimal solutions or some time not. In the resources will be free after job have been finished. It is non-primitive in nature.

The Min-Min method

This method without considering the load of the machines to which the task is to be given to be completed, picks up the smallest task first from all the available tasks and assigns it to a machine which gives the minimum completion time for that task. It increases the total completion time of all the tasks and hence increases the make span. The completion time and execution time for a task are almost same and improves the overall throughput of the system. [9]

Max-Min method

selects the longest task first to schedule on the fastest machine. The smaller tasks have to wait for a longer

period. It increases the make span and system throughput but load balancing is not considered.[9]

In Priority based Job Scheduling Algorithm

the priority of the task is calculated and the tasks are sorted based on priorities. Then they are given to the machine which produces best completion time and hence improves performance by having better completion time. [10]

1.4 MULTI CLUSTER ENVIRONMENT

Clusters are gathering of servers with distinct setup. With single cluster, hubs are of comparative arrangements. [11]As multiple clusters are bound together to form multi cluster condition. In multi cluster condition in this manner, particular nodes(servers) are assembled together. Multi cluster condition is utilized as a part of request to execute complex jobs. These jobs may not be executed by single cluster consequently distribution methodologies are constructed. Multi cluster condition is appeared as under

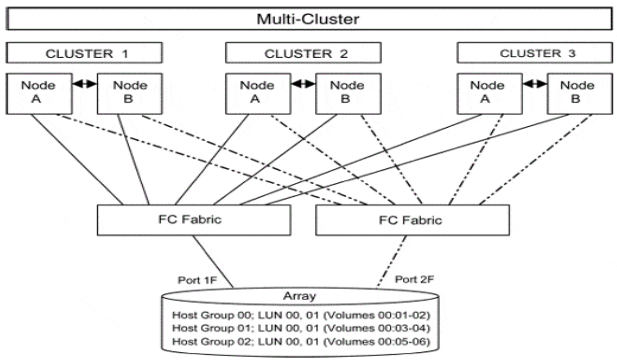

Figure 2. Multi Cluster Environment

1.5 Parallel Computing

Parallel computing is working on the rule that enormous issues can be subdivided into littler ones, which are then unraveled in the meantime. In this sort of calculation many jobs are executed at the same time on various processors [12]

The jobs in parallel computing can be characterized by the level at which the equipment support parallelism, with multi-center and multi-processor PCs having numerous handling components inside a solitary machine, while groups, MPPs and lattices utilize various PCs to chip away at a similar job. [13] Particular parallel PC designs are once in a while utilized close by traditional processors, for quickening particular jobs. Parallel programming models and devices are required for elite computing. The accessibility of multi-center CPUs has given new slant to the common memory parallel programming approach. Parallel computing is operating on the rule that huge problems can frequently be subdivided into smaller ones, which are then solved at the same time. In this type of computation many jobs are executed simultaneously on different processors

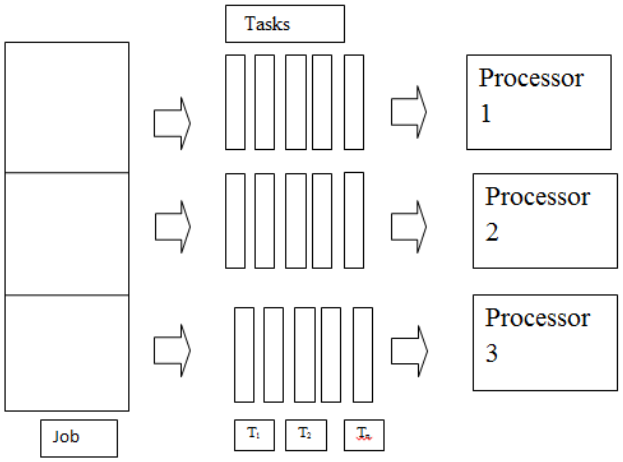

Figure 3. Parallel Computing Of Jobs

In parallel the jobs are subdivided into tasks and after that those errands are doled out to the different resources as indicated by job necessity. The parallel job calculation is far much superior to the successive calculation in light of the fact that in consecutive calculations the jobs are executed in a solitary line to one processor. So it expends a great deal of time to execute number of jobs. So to determine this issue the jobs are executed in parallel so by doing as such the execution time of jobs are being decreased and expanded the throughput of the framework.

II. MULTI HEURISTIC (MULTI CLUSTER) APPROACH FOR JOB SCHEDULING

2.1 Metaheuristic Approach

[14]Metaheuristic approach is used to NP- hard Problems. These problems are very complex in nature which may not be sold using the single processor system. There requirements are such that single processor is unable to satisfy the requirements or need of the system. So systems with more than one processors or multi clusters are required. With every job objective is associated. These objectives are required to be satisfied in order to complete the job. Such jobs fall under the category of metaheuristic. The algorithms are as under.

2.1.1 GENETIC ALGORITHM

Genetic algorithm is utilized as a part of request to take care of job schedule issue in multi bunch condition. With each activity target work is related. Genetic approach continues to fulfill the target work related with the activity. This algorithm ends on the off chance that target work related with each activity is fulfilled or number of ages ends. There are stages related with the genetic algorithm.[15]

Initialization

Jobs and resources at first place required to be initialized. Jobs are known as chromosomes. These chromosomes are initially selected randomly for allocation. Generations, objective function, crossover and mutation probabilities are initialized in this phase.

Selection

[16]In genetic algorithm, Jobs are referred to as Chromosomes. Resources are required by chromosomes. Chromosomes are selected initially for resource allocation. This process is known as selection. In genetic algorithm, this selection is performed using variety of mechanism known as selection criteria's. Chromosomes are selected for later breeding using this phase. Steps in selection are listed as under

- ✓ Fitness function is evaluated associated with each job. Fitness values are normalized by adding the fitness values and then dividing it with total number of jobs.
- ✓ Population is sorted according to descending value of fitness values
- ✓ Population with highest fitness value is selected for mating.

Crossover or Mating

Selected Chromosomes are then go through this phase in order to determine chromosomes which are to be mutated. Uniform crossover is preferred in proposed thesis.

Mutation

Chromosomes selected for mating in crossover are mutated to generate new chromosomes and then evaluated again. After performing all the phases, fitness values are analyzed again. This process continues until desired level of fitness values is achieved or generations terminates.

2.2 ANT COLONY OPTIMIZATION

[17]This is an algorithm which is inspired from the natural behavior of ants. Ant colony optimization considers the random walk associated with Ants. Ants look for food source and if found laid down pheromone trail to be followed by other Ants. This process is simulated in job scheduling. Ants act as Jobs and Food act as resources. With Ant Colony Optimization Convergence is not always guaranteed. In other words optimal solution may or may not be generated.[18] Result in terms of local and global best solution is obtained. Iterations in this algorithm continue until all the Ants gets the food source. Following steps are critical in Ant Colony Optimization

- All the Ants must visit all the resources at least once
- Distant resource is less likely to be visited due to visibility problem.
- In case resource is fond by Ant, pheromone is laid down.
- After previous iteration, Pheromone evaporated.

- Process continues until optimal solution gbest and lbest is found out or generations terminates.

Convergence problem exists while using Ant colony Optimization.

2.3 BAT ALGORITHM

[19]Bat Algorithm was introduced by Yang in year 2010 based on the echolocation behavior of Bats. To estimate the distance of prey, bat used echolocation to sense the distance.BAT algorithm is totally based on the behavior of Bats. The bats usually fly randomly with a frequency, loudness, pulse emission, velocity and position to search for their prey. When the bats hunt for the prey, they automatically adjust their parameters like velocity, frequency, loudness and pulse emission based on the distance among them and the prey.[20]BAT algorithm is used in many fields like in workflow, for independent jobs and for the parallel jobs and has many applications.BAT algorithms is used for scheduling jobs among the number of processors in a given problem to minimize objective function. By changing the velocities and the position of BATS gives the similarity of another metaheuristic known as the Particle Swarm Optimization algorithm.[21]BAT algorithm is a metaheuristic algorithm used to allocate resources which are detected. It was inspired by echolocation behavior of BAT. Searching criteria followed is random in nature. It is particularly useful for constrained intensive environment. The balance between exploration and exploitation has to be maintained which can be accomplished through hybrid approach. Pseudo Code for BAT is as under

- ✓ Obtain Job Sequence(J1,J2,J3,-------,Jn)
- ✓ Associate Objective Function with Each Job
  - $F(y) = \{y_1, y_2, - - - - -, y_n\}$
- ✓ Define range of BAT in terms frequency(F) and Loudness(A)
- ✓ Repeat following Steps until i<Max_Iterations
  - Generate Solutions by adjusting Existing frequency
  - Update location of BATS in case solution found
  - If(Optimal Solution==True)
  - Select this solution and replace it with existing solution
  - End of if
  - Randomize the BATS flying and relocate the resources.
  - In case optimal solution is located increase, pulse rate and decrease loudness(A).
  - Rank the BATS and sort the solution.
  - End of loop
- ✓ Output Solution and job Ordering(J)

2.4 Cuckoo Algorithm

[22]Cuckoo search is a meta-heuristic algorithm inspired by the bird cuckoo, these are the Brood parasites birds. It never builds its own nest and lays their eggs in the nest of another host bird nest. Cuckoo is a best-known brood parasite. Some host birds can employ directly with the intruding cuckoo. If the host bird identifies the eggs that are not their own egg then it will either throw that eggs away from its nest or simply rid its nest and build a new nest.

In a nest, each egg represents a solution and cuckoo egg represents a new and good solution. The obtained solution is a new solution based on the existing one and the modification of some characteristics. In the simplest form each nest has one egg of cuckoo in which each nest have multiple eggs represents a set of solutions. [23]CS is successfully used to solve scheduling problems and used to solve optimization problems in structural engineering. In many applications like speech reorganization, job scheduling, global optimization cuckoo search algorithm is used. Cuckoo search idealized such breeding behaviour and can be applied to various optimization problems

This algorithm is also a part of optimization algorithm which is used in order to produce optimal solution by considering decrease in time and energy consumption in multi objective scenario. The detailed steps are listed as follows:

- ✓ Initialize Cuckoo with initial set of eggs.

- ✓ Lay eggs in distinct nests.
- ✓ Some eggs are defected and killed
- ✓ If population generated is less than threshold value then checks for survival of eggs in nest and obtain profit values.
- ✓ Otherwise kill cuckoo in worst area.
- ✓ If stop condition satisfied then obtain maximize solution otherwise go to step 2.

2.5 Firefly algorithm-

[24]Firefly Algorithm (FA) is a metaheuristic algorithm for worldwide advancement, which is roused by blazing conduct of firefly creepy crawlies. This algorithm is proposed by Xin-She Yang in 2008. Fireflies utilize the blazing conduct to pull in different fireflies, generally to send signs to inverse sex. Be that as it may, in the numerical model, utilized inside Firefly Algorithm, just the fireflies are unisex, and any firefly can pull in different fireflies. Attractiveness of a firefly is relative to its splendor and for any couple of fireflies, the brighter one will pull in the other; so the less brilliant one is moved towards the brighter one. This is performed for any parallel mix of fireflies in the populace, on each cycle of algorithm.

2.6 Terminology Used

Flow Time- Flow time is the average amount of time to taken by job to achieve the optimal result. Every single unit in job can be executed within that time.

$$Flow_{time} = \sum \frac{Job_{Completion_{Time}}}{Total_{jobs}}$$

Waiting Time – The amount of time for which job have to wait for the processor to accomplish its task.

$$waiting \quad _{time} = Starting\_time - Arrival\_time$$

Execution Time- Execution time is the total amount of time to execute single job.

$$execution_{time} = finish_{\_time} - start\_time$$

Normalization Function- Normalization function is used in order to optimize the result which is obtained through metaheuristic algorithm.

Fitness Function- This function is used in order to analyze the fitness of the population or chromosomes. The population will alter according to this fitness function.

Comparison table showing the analysis of algorithms

Table 1. Comparison of techniques used within cloud computing for scheduling

ALGORITHM	BASIS	OPERATORS	ADVANTAGES	DISADVANTAGES
First Come First Serve [8]	Arrival time	• Fcfs	• Easy implementation	• No other criteria for scheduling
Priority based Job Scheduling Algorithm [10]	Task Priority, Expected completion time	• Priority based on parameter	• Priority is considered for scheduling. Designed based on multiple criteria decision making model	• Make span, Considerable improvement required in consistency and complexity of the proposed method
Min-Min, Max- Min [9]	Make span, Expected completion	• Fitness function	• Better make span compared to	• Poor load balancing and QoS factors are

	time		other algorithms	not considered
Genetic Algorithm (GA) [16]	Based on population size	• Crossover • Mutation • Selection • Inversion	• Used for rule extraction • for optimization problems	• Time taken for convergence is more. • No guarantee for global optima.
BAT Algorithm[19]	Based on the echolocation behaviour of virtual bats.	• Population size • Loudness parameter • Pulse rate • Maximum number of iterations	• Frequency tuning • Automatic zooming into the region of global solution • Parameter control ensures efficient exploration and exploitation.	• Convergence in case of complex problems is slow • Searching in case of complex environment can be misleading.
Cuckoo Algorithm[23]	Based on the property of brood parasitism of cuckoo species	• Population size • Switching probability • Step-size scaling factor • Levi exponent	• Use of levy flights result in efficient exploration of search space • Ensures the property of global convergence due to the use of switching probability factor.	• actual result is rare rather approximate solution is produced
ACO[18]	Based on natural behaviour of ants	• jobs • food • walk	• faster convergence as compare to GA	• finding local and global conclusion can be complex
Shadow replication [25]	Based on number of cores associated with VMs	• Resources • Cores	• Converge faster • Load balancing	• Problem arises as cores fails
Credit based scheduling[26]	User priority and task length is considered for allocation of resources	• Resources • Task length	• Make span is reduced • Flowtime is reduced	• Advance priority reservation is complex
Data centre energy aware	Energy of VM is	• Energy conservati	• Energy is saved • Execution time is	• Load on VM could cause

algorithm [27]	considered for scheduling	on • Execution time	reduced	deterioration
Primitive Scheduling[28]	Starvation problem is tackled in the allocation of resources	• Make span • Flowtime	• Make span and Flowtime is optimised	• Priority of job is not considered for premption hence critical jobs are prompted decreasing reliability
Enhanced Max-Min Algorithm [29]	Make span, Load balance, Average execution time	Fitness polices	• Improves make span • load balancing	• Parameters considered are limited and only theoretical analysis is performed
Firefly [30]	Based on attraction of glowing effect of fireflies	• Attractive • Mating	• make span time is reduced better than GA. • Global optimization	• Random selection of vm. • More flow times.

III. CONCLUSION

This paper gives a wide survey of various sorts of metaheuristic strategies While using the cloud computing technology, we have to face lot of new challenges. One of them is the task scheduling in a multi-cloud computing environment. The main objective of the scheduling is to maximize utilization of resources and to reduce make span. The list of algorithms proposed to achieve effective scheduling, but since the task scheduling is heuristic problem, more research can be one in this field and more appropriate algorithms can be designed so that optimized solutions can be achieved. Combining different evaluators such that to obtain an efficient scheduling algorithm and improve the overall performance of the cloud services can be done as an enhancement.

IV. REFERENCES

[1]. F. Sabahi,"Cloud Computing Security Threats and Responses," pp. 245-249, 2011.

[2]. D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya,"Energy-efficient data replication in cloud computing datacenters," Cluster Comput., vol. 18, no. 1, pp. 385-402, 2015.

[3]. N. M. Dhanya and G. Kousalya,"Adaptive and Secure Application Partitioning for Of fl oading in Mobile Cloud Computing," vol. 1, pp. 45-53, 2015.

[4]. J. Guitart, M. Macias, K. Djemame, T. Kirkham, M. Jiang, and D. Armstrong,"Risk-driven proactive fault-tolerant operation of IaaS providers," Proc. Int. Conf. Cloud Comput. Technol. Sci. CloudCom, vol. 1, pp. 427-432, 2013.

[5]. C. Pahl and I. Centre,"Containerization and the PaaS Cloud," 2015.

[6]. F. Doelitzscher, A. Sulistio, C. Reich, H. Kuijs, and D. Wolf,"Private cloud for collaboration and e-Learning services : from IaaS to SaaS," pp. 23-42, 2011.

[7]. A. Kaur,"A Review on Various Job Scheduling Algorithms," vol. 13, no. 3, pp. 359-3 Kobra

Etminani, M. Naghibzadeh" A Min-Min Max-Min selective algorihtm for grid task scheduling

[8]. " 2007 3rd IEEE/IFIP International Conference in Central Asia on Internet

[9]. Upendra Bhoi, Purvi N. Ramanuj,"Enhanced Max-minTask Scheduling Algorithm in Cloud Computing", International Journal of Application or Innovation in. Engineering & Management, Volume 2, Issue 4, April2013, pp 259-26467, 2017.

[10]. Xiaonian Wu, Mengqing Deng, Runlian Zhang, Bing Zeng, Shengyuan Zhou,"A Task Scheduling Algorithmbased on QoS driven in Cloud Computing", Information Technology and Quantitative management, 2013.

[11]. M. Arioua, Y. Assari, I. Ez-zazi, and A. Oualkadi,"Multi-hop cluster based routing approach for wireless sensor networks," Procedia - Procedia Comput. Sci., vol. 83, no. Ant, pp. 584-591, 2016.

[12]. S. P. Dandamudi,"Parallel Job Scheduling on Multicluster Computing Systems," 2003.

[13]. P. Switalski and F. Seredynski,"Scheduling parallel batch jobs in grids with evolutionary metaheuristics," J. Sched., vol. 18, no. 4, pp. 345-357, 2014.

[14]. Kenneth Sörensen," Metaheuristics—the metaphor exposed" 8 February 2013 Volume 22, Issue 1,January 2015.

[15]. C. R. Reeves,"A genetic algorithm for flowshop sequencing," Comput. Oper. Res., vol. 22, no. 1, pp. 5-13, Jan. 1995.

[16]. Tamanna Jena1 · J. R. Mohanty," GA-Based Customer-Conscious Resource Allocation and Task Scheduling in Multi-cloud Computing"vol.22. n0.1.aug2017

[17]. W. Wen, C. Wang, D. Wu, and Y. Xie,"An ACO-Based Scheduling Strategy on Load Balancing in Cloud Computing Environment," 2015.

[18]. B. Li, W. Xu, S. Member, and H. Zhang,"PAPR Reduction for Hybrid ACO-OFDM Aided IM / DD Optical Wireless Vehicular Communications," vol. 9545, no. c, 2017.

[19]. A. Hossein and G. X. Yang,"Bat algorithm for constrained optimization tasks," pp. 1239-1255, 2013.

[20]. S. Mirjalili and S. Mohammad,"Binary bat algorithm," 2013.

[21]. J. Xie, Y. Zhou, and H. Chen,"A novel bat algorithm based on differential operator and Levy flights trajectory," Comput. Intell. Neurosci., vol. 2013, 2013.

[22]. I. Engineering,"OPTIMAL POWER FLOW USING CUCKOO OPTIMIZATION ALGORITHM," pp. 4213-4218, 2013.

[23]. N. Optimisation,"A comprehensive review of cuckoo search : variants and hybrids Iztok Fister Jr .*, Dušan Fister and Iztok Fister," vol. 4, no. 4, 2013.

[24]. Upendra Bhoi, Purvi N. Ramanuj,"Enhanced Max-min Task Scheduling Algorithm in Cloud Computing", International Journal of Application or Innovation in Engineering & Management, Volume 2, Issue 4, April 2013, pp 259-264

[25]. X. Cui, B. Mills, T. Znati, and R. Melhem,"Shadow replication: An energy-aware, fault-tolerant computational model for green cloud computing," Energies, vol. 7, no. 8, pp. 5151-5176, 2014

[26]. A. Thomas, G. Krishnalal, and J. R. V P,"Credit Based Scheduling Algorithm in Cloud Computing Environment," Procedia - Procedia Comput. Sci., vol. 46, no. Icict 2014, pp. 913-920, 2015.

[27]. N. Kord and H. Haghighi,"An Energy-Efficient Approach for Virtual Machine Placement in Cloud Based Data Centers," pp. 44-49, 2013.

[28]. J. Li, C. Pu, Y. Chen, V. Talwar, and D. Milojicic,"Improving Preemptive Scheduling with Application-Transparent Checkpointing in Shared Clusters," pp. 222-234. Upendra Bhoi, Purvi N. Ramanuj,"Enhanced Max-minTask Scheduling Algorithm in Cloud

Computing",International Journal of Application or Innovation in. Engineering & Management, Volume 2, Issue 4, April 2013, pp 259-264.

[29]. R. Eswari1 and S. Nickolas1," SOLVING multi-objective task scheduling for heterogeneous distributed systems using fireflyalgorithm"aug 2017.