Efficient Task Scheduling in Cloud Computing

Sukhman Kaur,Rachhpal Singh

Punjabi University Regional centre for Information Technology and Management

Punjab;India

Abstract— Task scheduling plays a key role in cloud computing systems. Scheduling of tasks cannot be done on the basis of single criteria but under a lot of rules and regulations that we can term as an agreement between users and providers of cloud. This agreement is nothing but the quality of service that the user wants from the providers. Providing good quality of services to the users according to the agreement is a decisive task for the providers as at the same time there are a large number of tasks running at the provider's side**.**

Keywords— cloud computing, task scheduling, virtual machine.

I. INTRODUCTION

Cloud computing is moderately prevailing technology. Cloud computing has newly appeared as one of the exhortation in the ICT industry. It is a innovative technology of computing that is broadly utilized in today's industry and society. It is newest technology for sharing resources. Cloud computing is the Internet-based computing that permits to shared resources, software's and information can be offered to computers and other devices on requirement.
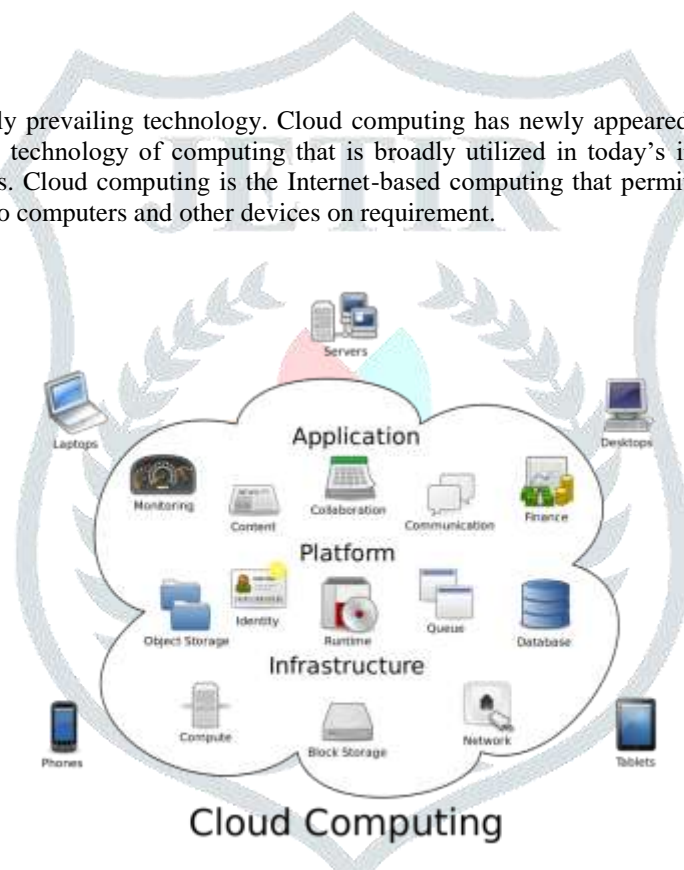

Fig 1: Cloud Computing

Cloud Characteristics:
- On-demand self-service: A consumer can provision computing capabilities, such as server time and network storage, as essential automatically without requiring human interface with each service's provider.
- Broad network access: Capabilities are presented over the network and accessed through standard mechanisms that support use by assorted thin or thick client platforms (e.g., mobile phones, laptops, and workstations).
- Resource pooling: The provider's computing resources are shared to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the subscriber generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or data centre). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.
- Rapid elasticity: Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.
- Measured Service**:** Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service.

II. TASK SCHEDULING IN CLOUD

Scheduling means the set of policies for controlling the order of work to be performed by a computing system. Scheduling is a major task in a cloud computing environment. In cloud computing environment datacenters take care of this task. A simple cloud architecture is shown in Fig 1. The datacenters receive tasks from the datacenter brokers which arrived from different users. In some cases these tasks may be associated with priorities. If so, a broker should consider these priorities and it is responsible for assigning the task. The algorithms like Min-Min will not consider user priority. A better scheduling algorithm is needed to achieve full utilization of resources.

Scheduling is that the cluster of strategies that manage the order of execution of multiple tasks on the processors therefore on decrease the time and value required to execute of those tasks. at intervals the cloud atmosphere, task hardware plays vital role of allocating cloud provider's resources among the massive type of users. Task designing deals with distribution of the tasks among the cloud servers that technique or execute these tasks for user (or client). associate economical task designing policy provides correct utilization of resources, load deed and improvement of execution value and time. therefore these days task designing is main analysis topic at intervals the area of cloud computing. There are various varieties of designing like static, dynamic, pre-emptive, non pre-emptive, centralized and distributed designing.[1]
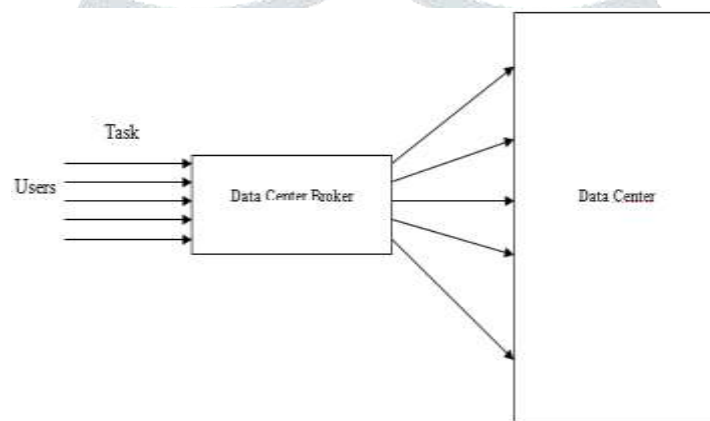

Fig 2: Scheduling in Cloud

An economical Multi Queue Job programming for Cloud Computing [1]: once FCFS and spherical Robin rule is employed fragmentation happens at several stages therefore wastage of area and client value is accumulated. during this paper author projected MQS (Multi Queue Scheduling) programming rule. during which initial of all task square measure appointed in ascending order then it's divided into medium, little and enormous size queue. Then Meta computer hardware allocates the task to the virtual machine. The results of this rule show that will increase user satisfaction and utilize the free unused area therefore performance is accumulated.

Improved Max-Min programming Model for Task programming in Cloud [2]: during this paper varied rule is mentioned then improve the Max-Min programming rule. once min-min programming rule is employed resource imbalance drawback has occurred. In max-min programming rule most size allotted to minimum completion time. Makespan is best than Min-Min rule. RASA rule it's referred to as resource aware task programming rule during which Min-Min and Max-Min each rule is combined. once resource is even then Max-Min rule is employed and resource is odd then Min-Min rule is employed. The result show that RASA rule has higher makespan then Max-Min. Improved max-min rule during which largest most task is allotted to the slowest resources. therefore alternative smaller task is dead on quicker resources and waiting time is attenuate. results of this rule is showing that higher makespan then all algorithms.

A Dynamic optimisation rule for task programming in Cloud atmosphere [6]: during this paper rule is projected it's helpful to each the service supplier and user. during this paper the task square measure grade supported task work time and Resource value. The prioritization of the task is predicated on finding best appropriate resources within the cloud. The results of this rule show that improves value and completion time of tasks as compared to serial assignment.

User Priority based mostly Min-Min programming rule for Load equalisation in Cloud Computing [3]: during this paper authors solves the resource imbalance and user priority drawback. during this priority based mostly. Load balance rule initial of all the meta task is split into influential person and standard based (Priority based) cluster then minimum size task is given to then minimum completion time of the resources. once all the task is appointed to the resources resource imbalance drawback is made therefore schedule the task to the low load resources. The result show that higher makespan then Min-Min rule and user priority additionally glad. Disadvantage of this rule is that once high priority task is simply too massive then lower priority task has looking forward to durable. So, starvation drawback is made.

Priority based mostly Job programming rule In Cloud Computing [8]: during this paper author projected new programming rule is predicated on multi-criteria and multi-decision priority based rule. during this rule the task is split into 3 level: object level,

attribute level and alternate level. The priority are often set during this rule is job resource magnitude relation. Bases on priority vector the task is compared with queue. The results of rule show that higher outturn and fewer end time.

HEFT based mostly work flow programming rule for value optimisation at intervals point in time in Hybrid Cloud [5]: during this paper projected rule main specialize in tasks method availableness of resources. during this rule resources ought to be taken supported the lease line of the general public cloud to execute the task is completed at intervals point in time and with minimum value. A hybrid programming rule projected new thought sub-deadline of the rule and allocation of the resources publically cloud. This rule is employed best for choice of resources with minimum value. The result show that minimum value performance is accumulated.

III. RELATED STUDY

Antony Thomas et al. [6] 2014 presented an improved scheduling criteria after analyzing the traditional algorithms that happen to be based on user priority and job length. High prioritized jobs are not given any special importance when they arrive. Min-Min algorithm is employed to reduce the make span of tasks by thinking about the task length. Keeping this in mind, cloud providers should achieved end user satisfaction. Thus research favors scheduling algorithms that consider both user satisfaction and resources availability.

Doshi Chintan Ketankumar et al. [7] 2015 suggested a green cloud broker for resource procurement problem by considering the metrics of energy efficiency and environmental friendly procedures of the cloud service supplier. Author used mechanism design methods to decide the allocation and payment for the submitted job effectively and performed experiments and show the results of comparisons of energy usage and emission of green house gases between the allowance decided by the suggested green cloud broker and a without taking the green metric into account.

A.I. Awad et al. [8] suggested mathematical model using Weight Balancing Mutation (balancing) a particle swarm optimization (LBMPSO) based schedule and allowance for cloud computing that takes into account trustworthiness, execution time, transmission time, make span, round trip time, transmission cost and load balancing between jobs and virtual machine. LBMPSO can be involved in attaining reliability of cloud processing environment by considering the resources available and reschedule task that failure to allocate. Our approach LBMPSO compared with standard PSO, random algorithm and Greatest Cloudlet to Fastest Processor chip (LCFP) algorithm to show that LBMPSO can save in make span, performance time, round trip time, transmission cost.

Arun kumar. G et al. [9] 2015 discussed the existing cloud systems, interoperability issues and possible solution to overcome the issues. Most of the consumers are analyzing the appropriateness of cloud to use themselves for their enterprise or customized operations. Customers are self satisfied at the inception, but expectation changes. Based on their business escalation it needs further adoption of recent cloud services the existing cloud provider fails to offer. Hence the user needs interoperability and portability to ship estate assets from one cloud to other cloud. The complication confronted by the customers in shifting their assets remains as challenging to be addressed.

Atul Vikas Lakra et al. [10] 2015 proposed a multi-objective task scheduling criteria for mapping tasks to a Vms in order to enhance the throughput of the datacenter and reduce the cost without breaking the SLA (Service Level Agreement) for an software in cloud SaaS environment. The proposed algorithm provides an optimal scheduling method. Most of the methods schedule tasks based on single requirements. But in cloud environment it is required to consider various requirements like execution time, cost, bandwidth of end user and so forth.

Narander Kumar et al. [11] 2015 suggests a demand-based preferential source allocation technique that designs a market-driven auction system to identify users for resource allocation based on their payment capacities and implements a payment strategy based upon a buyer's service preferences. A comparison is drawn between proposed allowance technique and the famous off-line VCG auction system and results show a performance benefit in profits to service provider, obligations of cloud users besides ensuring an optimum resources use.

Nidhi Bansal et al. [12] 2015 developed a method to calculate expense of QoS-driven task scheduling criteria and compare with traditional task scheduling algorithm in cloud computing environment. This also defined many guidelines that should be considered in QoS driven like makespan, latency and cargo handling. But allocation cost variable is not considered in QoS-driven scheduling algorithm. Reducing the total allocation value is an important issue in cloud computing.

Mehmet Sahinoglua et al. [13] 2015 addressed a discrete event CLOUD sim, namely CLOURAM (CLOUD Risk Assessor and Manager) to estimate the risk directories in large CLOUD processing environments, comparing favorably with the intractably theoretical Markov solutions or hand computations that are limited in scope. The goal is to optimize the quality of a CLOUD {procedure} and what countermeasures to consider to minimize threats to the service quality by reserve planning of preserve crew members.

Mohammad Mehedi Hassan et al. [14] 2015 suggested an automated method, established on Multi-Objective Particle Swarm Optimization, for the recognition of power models of enterprise servers in Cloud data centers. The typical usage of a single data

center is equivalent to the energy consumption of 25. 000 households. Building the power consumption for people infrastructures is crucial to anticipate the effects of aggressive optimization policies, but accurate and fast electric power modeling is an elegant challenge for high-end machines not yet satisfied by analytical approaches.

RESEARCHER	METHOD	BASED ON	FINDINGS
Tarun goyal & aakanksha agrawal[16][2013]	Host Scheduling Algorithm	Genetic Algorithm	A scheduling model based on minimum network delay using suffrage heuristic coupled with genetic algorithms for scheduling sets of independent jobs algorithm is proposed, the objective is to minimize the make span.
Sourabh budhiraj et. Al. [17] [2014]	Task Scheduling	Objective Genetic Algorithm	For task scheduling, a multi-objective genetic algorithm is implemented and the research is focused on crossover operators, mutation operators, selection operators and the pareto solutions method. The experimental results show that the proposed algorithm can obtain a better solution.
A.kaleeswaran et. Al. [18] [2013]	Dynamic Scheduling Of Data	Genetic Algorithm	Using genetic algorithm the tasks are scheduled according to the computation and memory usage. The tasks are scheduled dynamically. The execution time is reduced by parallel processing. The scheduled data is stored in cloud. By using ga we obtain global optimization.
Md whaiduzzaman et. Al. [19][2014]	Strategic Provisioning	Service Provisioning	We aim to review the state-of-the-art service provisioning objectives, essential services, topologies, user requirements, necessary metrics, and pricing mechanisms.we synthesize and summarize different provision techniques, approaches, and models through a comprehensive literature review. A thematic taxonomy of cloud service provisioning is presented after the systematic review.
Luiz f. Bittencourt et. Al. [20]	Scheduling In Hybrid Clouds	Concept Of Schedulers	This paper introduces the scheduling problem in hybrid clouds presenting the main characteristics to be considered when scheduling workflows, as well as a brief survey of some of the scheduling algorithms used in these systems

Table 1 Task Scheduling Methods used previously

IV. EXISTING SCHEMES

The following task designing algorithms space unit presently established among the cloud environments

4.1 hymenopteran Colony improvement (ACO)-inspired: a replacement Cloud hardware supported hymenopteran Colony improvement is that the one resented by Cristian Mateos.goal of our hardware is to attenuate the weighted flowtime of a gaggle of PSE jobs, whereas collectively minimizing Makespan once using a Cloud. among the ACO formula, the load is calculated on each host taking into consideration the equipment utilization created by all the VMs that unit of measurement punishment on each host. This metric is helpful for Associate in Nursing hymenopteron to choose the tiniest quantity loaded host to assign its VM.Parameter Sweep Experiments (PSE) may well be a spread of numerical simulation that involves running Associate in Nursing outsized sort of freelance jobs and frequently desires lots of computing power. These jobs ought to be with efficiency processed among the utterly totally different computing resources of a distributed setting like those provided by Cloud. Consequently, job designing throughout this context therefore plays a basic role. throughout this formula, Makespan and flowtime unit of measurement evaluated. analysis results of this metrics show that ACO performance more than two totally different (Random and Best effort) algorithms.[8]

4.2 Min-Min Algorithm: Min-Min begins with a gaggle of tasks that unit of measurement all unassigned. First, it computes minimum completion time for all tasks on all resources. Then among these minimum times the minimum value is chosen that's that the minimum time among all the tasks on any resources. Then that task is regular on the resource on it it takes the minimum time and thus the on the market time of that resource is updated for all the alternative tasks. it's updated throughout this manner; suppose a task is assigned to a machine and it takes twenty seconds on the assigned machine, then the execution times of all the alternative tasks on this assigned machine square measure planning to be increased by twenty seconds. once this the assigned task is not thought of and thus a similar technique is perennial until all the tasks unit of measurement assigned resources.

4.3 Max-Min formula:Max-Min is kind of same as a result of the min-min formula except the following: throughout this once looking for the completion time, the minimum execution times unit of measurement famed for each and every task. Then among these minimum times the most value is chosen that's that the most time among all the tasks on any resources. Then that task is regular on the resource on it it takes the minimum time and thus the on the market time of that resource is updated for all the alternative tasks. The modification is finished among a similar manner as for the Min-Min.

4.4. Particle Swarm improvement (PSO) Algorithm: Particle Swarm improvement (PSO) as a meta-heuristics technique may well be a self-adaptive international search based totally improvement technique introduced by Kennedy and Eberhart [5]. The PSO formula is alike to totally different population-based algorithms like Genetic algorithms (GA) but, there is not any direct recombination of individuals of the population . The PSO formula focuses on minimizing the worth of computation of Associate in Nursing application advancement. As a live of performance, Authors used worth for complete execution of application as a

metric. the target is to attenuate the worth of execution of application workflows on Cloud computing environments. Results show that PSO primarily based task-resource mapping square measure ready to do a minimum of thrice worth savings as compared to Best Resource selection (BRS) based mapping for our application advancement. in addition, PSO balances the load on cipher resources by distributing tasks to on the market resources.[5]

4.5 spherical Robin Algorithm: The spherical Robin formula in the main focuses on distributing the load equally to any or all the resources. Victimization this formula, the broker allocates one VM to a node throughout a cyclic manner. The spherical robin designing within the cloud computing is unbelievably rather like the spherical robin designing utilized within the technique designing. The hardware starts with a node and moves on to future node, once a VM is assigned to that node. this can be often perennial until all the nodes square measure assigned a minimum of 1 VM then the hardware returns to the first node over again. Hence, throughout this case, the hardware does not stay awaken for the exhaustion of the resources of a node before moving on to future. Though spherical robin algorithms unit of measurement supported straightforward rule, lots of load is planned on servers and thus unbalancing the traffic. Results of spherical Robin formula shows higher amount and payload deed as compared to the alternative formula.[7]

   4.6 Genetic Algorithm: Genetic formula may well be a way of coming up with among that the tasks unit of measurement assigned resources per individual solutions (which unit of measurement called schedules in context of scheduling), that tells regarding that resource is to be assigned to it task. Genetic formula relies on the biological construct of population generation. The foremost terms utilized in genetic formula are[6]

4.6.1 Initial Population
Initial population is that the set of all the those that unit of measurement utilized within the genetic formula to hunt out the optimum resolution. every resolution among the population is called as a private. and every individual is pictured as a body for making it acceptable for the genetic operations. From the initial population the folks unit of measurement elect and a number of operations unit of measurement applied on those to form future generation. The coupling chromosomes unit of measurement elect supported some specific criteria.[6]

4.6.2 Fitness operate
A fitness operate is used to measure the quality of the folks among the population per the given improvement objective. The fitness operate could also be utterly totally different for numerous cases. In some cases the fitness operate could also be supported purpose in time, whereas in cases it's going to be supported budget constraints.

4.6.3Selection
We use the proportion selection operator to ascertain the prospect of various folks genetic to future generation in population. The proportional selection operator suggests that the prospect that's chosen and genetic to next generation groups is proportional to the dimensions of the individual's fitness.

4.6.4Crossover
We use single-point crossover operator. Single-point crossover suggests that only one intersection was started among the individual code, at that point a district of the strive of individual chromosomes is modified.[8]

4.6.5 Mutation
Mutation means that the values of therefore sequence locus among the body cryptography series were replaced by the alternative sequence values so on get a replacement individual. Mutation is that negates the price at the modification points with reference to binary coded folks.

V. CONCLUSION

Cloud computing is one all told the user familiarising technology throughout that user faces a pool of virtualized laptop resources. throughout this paper we've a bent to survey varied existing coming up with algorithms in cloud computing. Since cloud computing is in infancy state, a coming up with framework need to be implemented to boost the user acquiescence at the aspect of the service suppliers. The design metrics area unit typically coupled to arrange a framework for recourse allocation and planning in cloud computing. the design framework need to have confidence the user input limitations (deadlines, performance issues, execution value, transmission value, energy efficiency, Load effort, and Makespan) thus on.

VI. REFERENCES

[1].   Shah Mihir N, Yask Patel, "A Survey Of Task Scheduling Algorithm In Cloud Computing", International Journal of Application or Innovation in Engineering & Management (IJAIEM), ISSN 2319 – 4847, Vol: 4, Issue 1, January 2015

[2].   S.DEVIPRIYA , C.RAMESH " Improved Max-Min Heuristic Model For Task Scheduling In Cloud", IEEE 2013

[3].   Huankai Chen, Professor Frank Wang, Dr Na Helian, Gbola Akanmu "User Priority Guided Min-Min Scheduling Algorithm For Cloud Computing", IEEE 2013

[4].   Pinal Salot,"A Survey of various scheduling algorithm for Cloud Computing", International Journal of Research In Engineering and Technology Feb-2013

[5].   Nisha Chopra, Sarbjeet Singh," HEFT based Workflow Scheduling Algorithm for Cost Optimization within Deadline in Hybrid Clouds", IEEE 2013

[6].   Antony Thomas, Krishnalal G, Jagathy Raj V P, "Credit Based Scheduling Algorithm in Cloud Computing Environment", International Conference on Information and Communication Technologies, Vol: 46, pp: 913-920,2015

[7].   Doshi Chintan Ketankumar, Gaurav Verma, K. Chandrasekaran, "A Green Mechanism Design Approach to Automate Resource Procurement in Cloud", Eleventh International Multi-Conference on Information Processing, Vol: 54, pp: 108-117,2015

[8].   A.I.Awad, N.A.El-Hefnawy, H.M.Abdel_kader, "Enhanced Particle Swarm Optimization For Task Scheduling In Cloud Computing Environments", International Conference on Communication, Management and Information Technology, Vol: 65, pp: 920-929,2015

[9].   Arunkumar. G, Neelanarayanan Venkataraman, "A Novel Approach to Address Interoperability Concern in Cloud Computing", 2nd International Symposium on Big Data and Cloud Computing, Vol: 50, pp: 554-559,2015

[10].   Atul Vikas Lakra, Dharmendra Kumar Yadav, "Multi-Objective Tasks Scheduling Algorithm for Cloud Computing Throughput Optimization", International Conference on Intelligent Computing, Communication & Convergence, Vol: 48, pp: 107-113,2015

[11].   Narander Kumar, Swati Saxena, "A Preference-based Resource Allocation In Cloud Computing Systems", 3rd International Conference on Recent Trends in Computing, Vol: 57, pp: 104-111,2015

[12].   Nidhi Bansal, Amitab Maurya, Tarun Kumar, Manzeet Singh, Shruti Bansal, "Cost performance of QoS Driven task scheduling in cloud computing", Third International Conference on Recent Trends in Computing, Vol: 57, pp: 126-130,2015

[13].   Mehmet Sahinoglua, Sharmila Ashokan, Preethi Vasudev, "Cost-Efficient Risk Management with Reserve Repair Crew Planning in CLOUD Computing", Cost-Efficient Risk Management with Reserve Repair Crew Planning in CLOUD Computing, Vol: 62, pp: 335-342,2015

[14].   Patricia Arroba, Jos´e L. Risco-Mart´ın, Marina Zapater, Jos´e M. Moya, Jos´e L. Ayala, Katzalin Olcoz, "Server Power Modeling for Run-time Energy Optimization of Cloud Computing Facilities", 6th International Conference on Sustainability in Energy and Buildings, Vol: 62, pp: 401-410,2014

[15].   https://www.google.co.in/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=0ahUKEwikh9KS oYjLAhWRcI4KHQiyC7sQjRwIBw&url=https://commons.wikimedia.org/wiki/File:Cloud_computing.svg&bvm=bv.114733917, d.c2E&psig=AFQjCNF9bPyRaZ0qIlWocIQGWY6yHn26eQ&ust=1456123288867169

[16].   Tarun Goyal & Aakanksha Agrawal., ―Host Scheduling Algorithm Using Genetic Algorithm In Cloud Computing Environment‖, International Journal Of Research In Engineering & Technology (Ijret), Impact Journals. Vol. 1, Issue 1, June 2013, 7-12.

[17].   Sourabh Budhiraja, Dr. Dheerendra Singh, ―An Efficient Approach for Task Scheduling Based on MultiObjective Genetic Algorithm in Cloud Computing Environment‖, IJCSC Vol: 5, number 2 july-sept 2014, pp: 110-115.

[18].   A.Kaleeswaran, V.Ramasamy, P.Vivekanandan.,―Dynamic Scheduling Of Data Using Genetic Algorithm In Cloud Computing‖, International Journal Of Advances In Engineering & Technology, Jan. 2013. 2231-1963.

[19].   Md. Whaiduzzaman, Mohammad nazmul haque, Md rejaul karim chowdhury, and Abdullah gani., ―A Study On Strategic Provisioning Of Cloud Computing Services‖, Hindawi Publishing Corporation, E Scientific World Journal Volume 2014, article id 894362, 16 pages. Http://dx.doi.org/10.1155/2014/894362

[20].   Luiz F. Bittencourt, , Edmundo R. M. Madeira, and Nelson L. S. Da Fonseca, ―Scheduling In Hybrid Clouds‖, Institute of Computing – University of Campinas – UNICAMP, Campinas – S˜ao Paulo – Brazil, 13083-852.