IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-12 E-ISSN: 2347-2693

Job Categorization based Task Scheduling using QoS in Cloud Environment

S. Khurana[1*], R. K. Singh[2]

[1*]Dept. of CSE, I. K. Gujral Punjab Technical University, Kapurthala, India
[2]SUS *Institute of Computer, I. K. Gujral Punjab Technical Universit, Tangori, India*

[*]*Corresponding Author: savu.khurana30@gmail.com, Tel.: 9896045578*

Available online at: www.ijcseonline.org

Received: 07/Nov/2017, Revised: 21/Nov/2017, Accepted: 11/Dec/2017, Published: 31/Dec/2017

***Abstract*—** The Cloud computing come into points of interest as a brand new emerging technology which provides personalized QoS guaranteed and trustworthy computing environment for end-customers. In cloud environment the Jobs are planned and accomplish in the constraints. Means of constraints right here is to apply QoS that customers want and balancing among those QoS and impartiality some of the Jobs. Many scheduling algorithms are proposed and executed to fulfill the requirement of cloud computing. Some of the task scheduling algorithms focused on the priority to each task depending on its attributes and then schedule tasks while considering the high priority. Others Jobs scheduling algorithm based upon QoS resourcefully utilized the resources idle time from monitoring the task schedule and timing information. But most of the current scheduling algorithms have not focus on the QoS parameter based upon task computation & communication time. They only examine the inactive time of resources continuously & renew the minimum completion time (MCT) of resources in matrix; but not executed the jobs based upon their job type. The task which demanded high computation resources should executed on high speed CPU and task required communication based resource should be executed on resources with high speed bandwidth. This paper proposed an enhanced QoS based algorithm which also taken into account the task monitoring and job type parameters.

***Keywords*—** Cloud Computing, Job Type, Job Priority, QoS, Task Scheduling

I. INTRODUCTION

Cloud computing is a brand new raising method in IT enterprise and it is a buzz word in computing environment [17]. It provide us trustworthy services through the idea of virtualization, storage and internet based computing. it's miles a web based totally computing. It provides interconnected shared resources i.e. community servers, information storage, processors and many others, which can be provisioned to users according to the need of users through internet. It offer Infrastructure as a provider, Platform as a carrier and software as a carrier models [11]. The cloud customers can execute packages on the cloud resources as opposed to their own computers as the client only require the software to access the cloud resources to execute their own programs. The consumer use the cloud assets as a provider like fee of house hire, to apply the infrastructure or pay power bill in keeping with the units used. Similarly in cloud environment client can use the resources and pay the fees according to their utilization. Cloud computing has feature of elasticity which mean users on cloud can enlarge or decline the resource provisioning as per their need.

The main heart of cloud environment is its scheduling alogorithms. As we know scheduling process is to map jobs to the available resources on the basis of their requirement[1]. As the user base in cloud computing are increasing day by day and requirement of efficient scheduling algorithms is mandatory which map the particular physical resource to a different number of users' in multitenant way [6]. The whole process of scheduling is divided into three phases:-

1. Discovery Phase: - identify the number of available resources in the network system by resource broker [21].
2. Filter Phase: - The identified resources are filtered [21].
3. Selection Phase: - select the best resource from the identified filtered resources [21].
4. Submission Phase: - Jobs is submitted to most suitable resource according to client requirements [21].

In cloud environment various parameters like priority, minimum completion time etc based task scheduling algorithms were proposed and implemented [19]. But they have not taken account the job type and resource type parameters. The proposed approach considered job type i.e. communication or computation based and proposed approach also considered the resource grouping. The next sections II of the paper presented the related work on cloud scheduling area and section III is identified and describes the problem. In section IV related with the proposed approach and section V focus on experimental setups, results and comparison with the existing approached. Section VI includes conclusion of paper and proposed approach.

II. RELATED WORK

Hend Gamal El Din Hassan Ali, Imane Aly Saroit, Amira Mohamed Kotb [1] proposed the grouped tasks scheduling (GTS) algorithm and its objective is to attain low latency with least execution time. The proposed algorithm schedule tasks in by taking into consideration quality of service parameters according to the client requirement and It divided the tasks into small categories but each category has similar attributes like user type, task type, task size, task latency and scheduling the tasks after relocate them into appropriate category. The simulation results show the domination of proposed algorithm in term of minimum execution time and minimum latency as compared to of Min-Min and TS algorithms.

Laiping Zhao, Yizhi Ren, Kouichi Sakur [2] proposed algorithm First of all identified the reliability of a given workflow schedule and investigate using Accumulated Processor Reliability (APR) which is a probability that a processor does not succeed until it completed the assigned scheduled task. Second step three algorithm to schedule task were proposed i.e. RR algorithm which schedules least resources to meet the trustworthiness requirement; DRR

algorithm works on both the reliability and deadline parameters with the minimum resource usage; and dynamic algorithm dynamically schedule and execute the tasks.

A.I.Awada, N.A.El-Hefnawyb, H.M.Abdel kader [3] proposed a particle swarm optimization (LBMPSO) algorithm based upon mathematical model and consider various parameters i.e., execution time, transmission time,make-span time , transmission cost. It also taken load balancing factor between various tasks and virtual machines for scheduling, allocation. The objective of proposed algorithm is to improves the distribution of tasks while considering various factors i.e. minimize cost, execution time, transmission time, makespan time.

Sukhpal Singh, Inderveer Chana [4] proposed resource provisioning algorithm consider QoS metric for resource scheduling and distribution and reduce the resource provisioning complexity for workload execution in cloud environment. The Proposed algorithm worked in different phases. First phase it analyzes the workloads and categorizes them on the basis of common patterns and in second phase it provisions the workloads before scheduling process start. It identified and analyzed the different workloads by considering the QoS and results shows that proposed algorithm successfully reduce the job queuing time which superior the effectiveness of scheduling of resources in cloud.

Xiaonian Wu et al., [5] proposed QoS-driven task scheduling algorithm which is based upon dynamic batching mode and jobs are not executed straight away, instead they are collected, The proposed approach is evaluated with Min-Min and Berger Model on different scheduling parameters like makespan, average latency and load balancing. The results show that the proposed approach better then the Berger Min-Min algorithm in term of effectiveness and load balancing.

Sukhpal Singh and Inderveer Chana proposed survey paper in scheduling and it considered standard precise literature analysis technique and various research papers presented in international conferences, symposiums, workshops, journals. The objective of the survey work is to discover the critical uniqueness of resource scheduling algorithms. The proposed work emphasizes on scheduling algorithms based on different criteria's for resource scheduling. Survey works presented the background details of scheduling of resources and address the benefits using different tools. The survey work described the thirteen different types of scheduling algorithms for cloud resources and eight different types of distribution policies for cloud resources. Survey work also describes different open Issues in scheduling of resources and its challenges in cloud environment [6].

Wei-JenWang, Yue-Shan Chang,Win-Tsung Lo, Yi-Kang Lee [7] proposed QoS based Adaptive-Scheduling in hybrid cloud environment the use of satisfaction algorithm and the objective is to increase the utilization rate of resources in the private cloud and to reduce the response time of different Jobs which are submitted for execution. The proposed approach taken into accounts the cost and deadline parameters and also presumes that all the resources in the particular slot have the same value for different parameters. A detail evaluation is preformed with FIFO scheduler and results undoubtedly showed that the proposed AsQ algorithm gives better performance in terms of waiting time, execution time and finish time of different tasks.

Atul Vikas, Lakra and Dharmendra Kumar Yadav [8] proposed a task scheduling algorithm based upon multi-objective to capitalize on the throughput and diminish the cost in data centre and make sure that it will not influence the service level agreement. Most of traditional task scheduling algorithms are single objective algorithms and suffered problems due to single objective nature. Scheduling algorithms which focus on high priority, low priority task face starvation problem because they have to wait for a long time. First Come First Serve and Shortest Job First based algorithms face problem in worst case scenario. Single objective based scheduling algorithms works very well in the best case but their performance in worst case gives adverse effect in scheduling.. The proposed task scheduling algorithm overcomes this problem by taken account the QoS parameters and minimize the task execution time.

Hussin M. Alkhashai,Fatma A. Omara [11] proposed enhanced task scheduling algorithm with objective to lessen the execution time, cost and improve the utilization of resources in cloud environment. It compared with PSO algorithms in the overall performance parameters and the results shows that the proposed enhanced task scheduling algorithm is more efficient than conventional PSO.

Shengjun Xue, Mengying Li, Xiaolong Xu, and Jingyi Chen [12] proposed Ant colony algorithm based Load balancing optimization algorithm to solve the virtual machine problem of load balancing. Proposed approach lessens the makespan time of different task and also ensures the load balancing of all virtual machines. The proposed algorithm is compared with First Come First Serve, ant colony, a combination algorithm of ACO, roulette wheel algorithm and results undoubtedly show that the proposed ACO-LB algorithm has bare minimum execution time enhance the utilization rate of resources while conserving the load balancing

Nidhi Bansala, Amitab Mauryaa, Tarun Kumara, Manzeet Singha, Shruti Bansal [13] proposed a scheduling algorithm which considered QoS driven parameters like time for execution, load balancing, latency and user needs, authorization, process parameters and execution time. The simulation results shown that proposed algorithm is superior to FCFS algorithms in term of cost and performance parameters.

Antony Thomasa, Krishnalal Ga, Jagathy Raj [14] proposed algorithm which considered task priority, task length and user fulfillment and is used to assign credits to jobs. Proposed work presented three scenarios for cloud scheduling, in the First scenario it taken account the length of tasks, second scenario it taken account task priority and in the last scenario it considered both priority for cloudlet and length of cloudlet. The proposed algorithm is implemented in simulation environment i.e. cloudsim 3.0.3 simulator and Simulation results show that, the proposed algorithm works superior in term of makespan time than the conventional scheduling approaches.

Maria Alejandra Rodriguez and Rajkumar Buyya [15] proposed meta-heuristic and particle swarm based algorithm to reduce the overall workflow execution cost while meeting application's deadline. The proposed approach is compared with conventional IC-

PCP and SCS algorithms and results show that proposed approach is superior in term of meet application's deadline and also it is competent of generating better schedules.

R. Sathish Kumar and S. Gunasekara [16] proposed an artificial bee colony algorithm that taken accounts the QoS parameters. It computes task finishing time, mean time to finish the task and load balancing time and works in four important phases. The status of the current machine identified in very first phase and jobs allocation to datacenter in next phase. Artificial bee colony algorithm with QoS is applied in the third phase and last phase calculate the finishing time of the tasks. The creation of cloud environment is created using cloudsim simulator and results show the domination of proposed algorithm on conventional ACO in term of overall finishing time and load balancing time.

III. Problem Statement

Cloud Computing is a brand new form of internet primarily based computing that includes the characteristic of parallel and distributed computing. The interconnected and virtualized resources in cloud environment are dynamically provisioned, configured and accessed remotely through different cloud service provider [20]. Clouds provide computing platforms or IT infrastructures to the cloud users and execute user tasks on cloud. The cloud computing provides Infrastructure as a Service, Platform as a Service, and Software as a Service models. The principle goal of cloud computing is to free the user from taking burden of handling & upgrading these resources. Users only have to pay for the period of time which he used resources i.e. pay per use policy [9]. Cloud users submit their jobs with different parameters like job type, size, deadline, cost etc and having different needs. It is desirable to implement efficient & reliable QoS based scheduling algorithm to fulfill the user's needs [22].

Task scheduling is heart of Cloud computing [23], there is need of efficient task scheduling which map tasks to the appropriate resources for effective utilization. Various algorithms are implemented so far but most of them are concerned about the priority and minimum completion time of the task resource etc. [10], but they will not consider the resource trustworthiness and task category. There are two categories, one is computational intensive tasks which required high processing power like CPU, MIPS etc. and other is communication intensive tasks which require high bandwidth etc[18]. So, there is need of algorithm which schedules the task on most efficient and reliable resource. The proposed approach gives attention to not only on the priority, idle time of resource but also the task category.

IV. Proposed Approach

Step1: Users submit jobs to cloud Portal.

Step 2: Identify the task category based upon communication intensive or computation intensive jobs.
Step 3: Grouping Task based upon task category.
Dependent tasks are queued in one cluster and independent task are queued in another cluster.
Step 4: Resource Grouping is performed according to computation parameters i.e. CPU, Clock Speed, RAM and communicational parameters i.e. Bandwidth.
Step 5: In the task sequencing phase, assign the priorities to tasks & resources.
Step 6: Based upon the task & resource availability, draw minimum execution time matrix.
Step 7: Based upon Max-Min algorithm, map the resources to the tasks.
Step 8: Examine the minimum execution time of tasks & idle time of resources.
Step 9: Evaluate the performance.
Step 10: Periodically update the MET matrix base based upon successfully execution of the tasks.
So, in the above proposed algorithm, different cloud users submit their tasks on the cloud through cloud portal. Then in the task categorization, identify the task category whether it is belonging to dependent task or independent task. After identifying the task in Task categorization, task grouping is done.

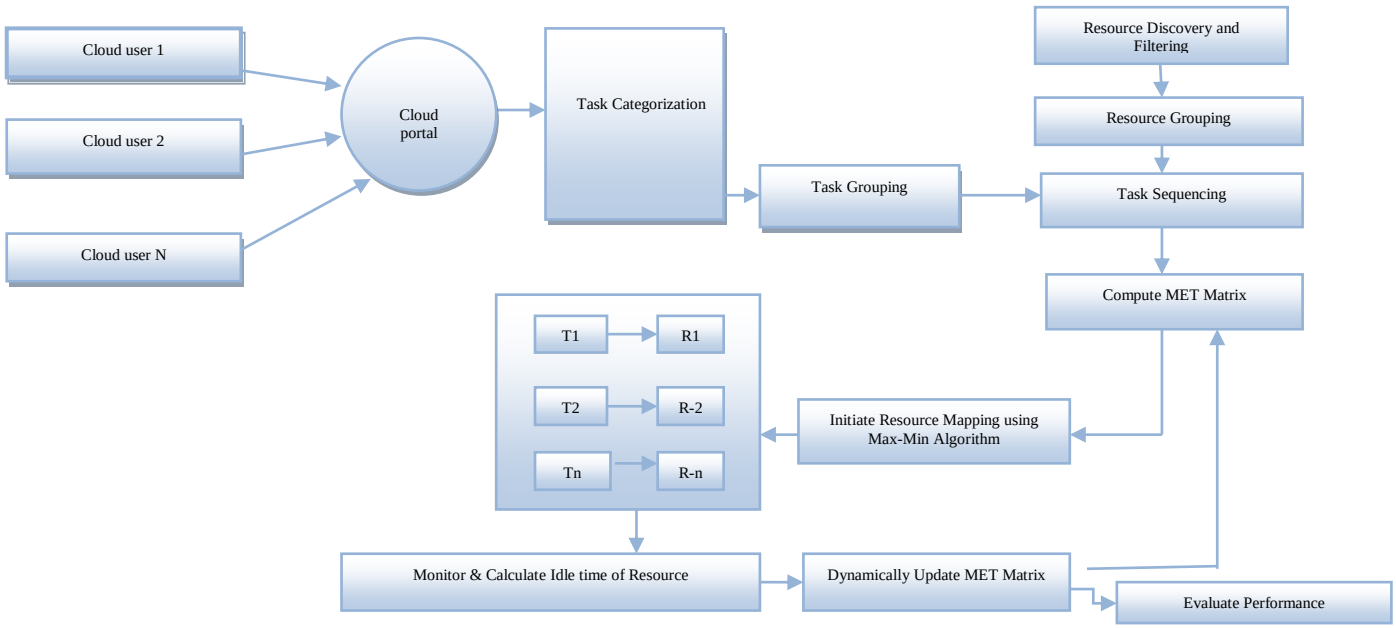

Figure 1. Proposed Architecture of QoS based Task Scheduling

Available resources are discovered for the task submission and group the resources on the basis of resource properties e.g. CPU, MIPS, RAM, and Bandwidth etc. The resources are grouped into two Computational Resources & Communicational Resources categories.

After resource grouping task sequencing is done. In task sequencing priorities the tasks based upon their type, length, latency, minimum execution time. The task priority is calculated as:-

$$TPr(i) = w * TU_i + x * TE_i + y * TL_i + z * LT_i \quad (1)$$

Where TPr(i) is the priority of task T_i. w, x, y, z ϵ[0,1] are weights of priority and w+x+y+z =1. TU_i, TE_i, TL_i, LT_i are the task user type, task Expected Execution Time, task length and latency time of task T_i respectively [1].

After task sequencing phase minimum execution time matrix is initialized and the mapping of the resources to the tasks is according to Max-Min algorithm. Monitor the task for dynamically updating the minimum execution time of the resource for further reference and also update the resource.

Performance evaluation:- An enhanced QoS based algorithm results into smallest completing time as compare to QoS based task scheduling algorithm. If jobs are submitted and executed on the right resource which further diminish the execution time as well as completion time and will result into increase throughput, decrease latency and efficient resources utilization.

V. RESULTS AND DISCUSSION

This section calculates the effectiveness of scheduling heuristic using independent tasks subset and demonstrates the performance of our proposed task scheduling approach. Cloud simulator i.e Cloudsim 3.0 is used to setup environment for simulations, which provides framework for modeling and simulation of entities in cloud environment and data centers for design and evaluation of algorithms.

Three virtual machines are created for implementation of proposed algorithms in simulation environment and cloud tasks are created based on the user's request and scheduled using proposed scheduled algorithm. The description of virtual machines is shown below:

Table1. Virtual Machine Creation

Resources	RAM	OS	Architecture	Processing Speed	MIPS	Bandwidth (Gb/s)
R1 to R10	4GB	Windows 7	Core i3	1 to 3GHZ	400 to 8000	1 to 10

Table 2. Task Creation

Cloudlet ID	Task Size(MB)
C1 to C100	100-5000

The performance of the system in simulation environment is evaluated with large number of tasks and also compared with existing Min-Min and Max-Min approach.

Makespan Comparison

The execution of the above algorithm assumes two scenarios to perform the evaluation. In scenario-1: 80% jobs are communication based and in Scenario- 2 80% jobs are computation based.

Table 3. Makespan comparison of proposed alogirthm with Min-Min approach

Scenario	Min-Min	Max-Min	Proposed approach	Improvement over Min-Min	Improvement over Max-Min
Scenario-1 (Communication Based)	120.45	129.35	105.80	12.16	18.20
Scenario-2 (Computation Based)	138.9	149.4	117.6	15.33	21.28

Figure 2. Makespan Comparison

As shown in Table III and Fig.2 for all above scenarios, the proposed scheduling algorithm outperforms the existing Min-Min and Max-Min heuristics by shorter makespan time and considered QoS parameters for scheduling of tasks.

Failure Comparison

Experiment performed on Core i5 based systems with CPU clock speed of 2.2GHz, 8GB RAM and windows 7-operating system. During simulation approximate 1000 Jobs are submitted and evaluate the failure rate of proposed scheduling algorithm with existing conventional scheduling approach. It is concluded from Fig.3 that the failure rate for first 100 tasks is estimated same in both approaches, but with increase in jobs submission, failure rate of conventional scheduling is amplified drastically due to simple match making algorithm implemented by conventional scheduling approach.

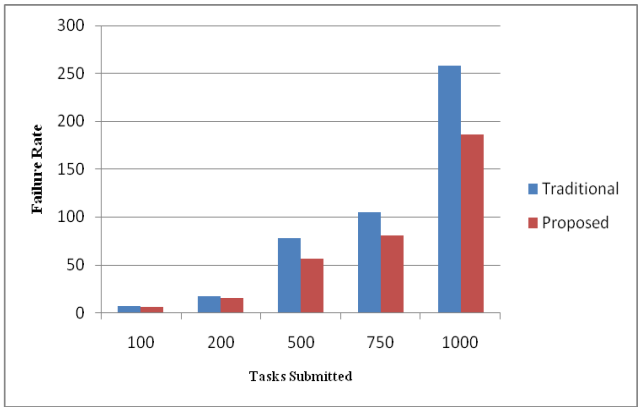

Figure3. Failure Rate of Traditional versus Proposed Approach

The trend of the conventional and proposed QoS based approach is shown in Fig. 3 for scheduling and allocation. The tasks failure rate in existing conventional algorithms is high in comparison with proposed QoS based approach.

VI. CONCLUSION

This paper proposed enhanced QoS based task scheduling based upon task categorization in cloud environment. Task scheduling gives considerable impact on performance and allocation of resources in cloud environment. Many task scheduling algorithms have been proposed so far in which task is submitted onto resource which have a minimum completion time but the existing conventional algorithm does not give attention on job type which is also a critical factor and gives adverse effect on the performance of scheduling. The proposed scheduling algorithm effectively uses the redundant resources by monitoring the resources, task timing information on resources and taken account the job type. The results of the proposed approach in cloud simulator are compared with min-min & min-max algorithms and results in lowest completion time & less failure rate.

REFERENCES

[1]. Ali, H.G.E.D. Hassan, I.A. Saroit and A.M. Kotb, "*Grouped tasks scheduling algorithm based on QoS in cloud computing network."* Egyptian Informatics Journal, 1(18), pp.11-19, 2017.

[2]. Zhao, Laiping, Y. Ren, and K. Sakurai. "*Reliable workflow scheduling with less resource redundancy."* Parallel Computing 39.10 (2013): 567-585.

[3]. Awad, A. I., N. A. El-Hefnawy, and H. M. Abdel_Kader. *"Enhanced Particle Swarm Optimization for Task Scheduling in Cloud Computing Environments*." Procedia Computer Science 65 (2015): 920-929.

[4]. S. Singh and I. Chana. *"Q-aware: Quality of service based cloud resource provisioning."* Computers & Electrical Engineering 47 (2015): 138-160.

[5]. W. Xiaonian, M. Deng, R. Zhang, B. Zeng, and S. Zhou. "*A task scheduling algorithm based on QoS-driven in cloud computing."* Procedia Computer Science 17 (2013): 1162-1169.

[6]. S. Singh and I. Chana. "*A survey on resource scheduling in cloud computing: Issues and challenges."* Journal of Grid Computing 14.2 (2016): 217-264.

[7]. Wang, Wei-Jen, Y. Chang, W. Lo and Y.K. Lee. "*Adaptive scheduling for parallel tasks with QoS satisfaction for hybrid cloud environments."* The Journal of Supercomputing 66, no. 2 (2013): 783-811.

[8]. Lakra, A. Vikas and D. K. Yadav. "*Multi-objective tasks scheduling algorithm for cloud computing throughput optimization*." Procedia Computer Science 48 (2015): 107-113.

[9]. Zhang, Rui, K. Wu, M. Li, and J. Wang. "*Online resource scheduling under concave pricing for cloud computing."* IEEE Transactions on Parallel and Distributed Systems 27, no. 4 (2016): 1131-1145.

[10]. He, Hua, G. Xu, S. Pang, and Z. Zhao. "*AMTS: Adaptive multi-objective task scheduling strategy in cloud computing*." China Communications 13, no. 4 (2016): 162-171.

[11]. Alkhashai, H. M., and F.A. Omara, "*An Enhanced Task Scheduling Algorithm on Cloud Computing Environment."* International Journal of Grid and Distributed Computing 9.7 (2016): 91-100.

[12]. Xue, Shengjun, M. Li, X. Xu, J. Chen, and S. Xue. "An ACO-LB algorithm for task scheduling in the cloud environment." Journal of Software 9, no. 2 (2014): 466-473.

[13]. N. Bansal, A. Maurya, T. Kumar, M. Singh, and S. Bansal. "*Cost performance of QoS Driven task scheduling in cloud computing."* Procedia Computer Science 57 (2015): 126-130.

[14]. T. Antony, G. Krishnalal, and V.P.J. Raj. *"Credit based scheduling algorithm in cloud computing environment."* Procedia Computer Science 46 (2015): 913-920.

[15]. Rodriguez, M. Alejandra, and R. Buyya, *"Deadline based resource provisioning and scheduling algorithm for scientific workflows on clouds."* IEEE Transactions on Cloud Computing 2, no. 2 (2014): 222-235.

[16]. Kumar, R. Sathish, and S. Gunasekaran, *"Improving Task Scheduling in Large Scale Cloud Computing Environment using Artificial Bee Colony Algorithm."* International Journal of Computer Applications 103, No. 5 (2014).

[17]. L. Wang, G. Laszewski, M. Kunze, J. Tao, *"Cloud Computing: A Perspective Study",* New Generation Computing- Advances of Distributed Information Processing, pp. 137-146, vol. 28, no. 2, 2008. DOI: 10.1007/s00354-008-0081-5.

[18]. R. K. Bawa and G. Sharma, *"Modified Min –Min Heuristic for Job Scheduling based on QoS in Grid Environment".* IEEE Conference on International Management in the Knowledge Economy, 19-20 December 2013.

[19]. C.Jasmin Selvi, G.Sathish Kumar, "*An Adaptive Double-Quality-Guaranteed (DQG) Scheme based Quality of Service (QOS) in Heterogeneous Cloud Environment*", International Journal of Computer Sciences and Engineering, Vol.4, Issue.4, pp.258-265, 2016.

[20]. S. Arnold,"Cloud computing and the issue of privacy", KM World, pp14-22. Available: www.kmworld.com, Aug. 19, 2009.

[21]. Anbazhagi, L. Tamilselvan and Shakkeera, "*QoS based dynamic task scheduling in IaaS cloud,"* International Conference on Recent Trends in Information Technology, Chennai, 2014, pp. 1-8. doi: 10.1109/ICRTIT.2014.6996176.

[22]. R.K. Bawa and G. Sharma, *"Reliable Resource Selection in Grid Environment".* International Journal of Grid Computing & Applications, March 2012, Volume 3, Number 1, pp. 1-10.

[23]. R.L. Grossman, "*The Case for Cloud Computing*", IT Professional, vol. 11(2), pp. 23-27, 2009, ISSN: 1520-9202.

Authors Profile

Ms. Savita Khurana pursed Bachelor of Computer Science & Engineering from Kurukshetra University in 2005 and Master of Computer Engineering from Kurukshetra University in year 2008. She is currently pursuing Ph.D. from Punjab Technical University, Kapurthala and currently working as Assistant Professor in Department of Information Technology, Seth Jai Parkash Mukand Lal Institute of Engineering & Technology, Radaur, Yamunanagar, since 2008.

Dr. Rajesh Kumar Singh is currently working as Principal in SUS Institute of Computer, Tangori, Punjab. He has published various research papers in reputed international journals and conferences.

