

# A Study on Scheduling Methods in Cloud Computing

Yogita Chawla<sup>1</sup> and Mansi Bhonsle<sup>2</sup>

<sup>1,2</sup>Pune University, G.H Raisoni College of Engg & Mgmt,  
Gate No.: 1200 Wagholi, Pune – 412207

**Abstract:** *Cloud computing is based on the concepts of distributed computing, grid computing, utility computing and virtualization. It is a virtual pool of resources which are provided to users via Internet. It gives users virtually unlimited pay-per-use computing resources without the burden of managing the underlying infrastructure. Cloud computing service providers' one of the goals is to use the resources efficiently and gain maximum profit. This leads to task scheduling as a core and challenging issue in cloud computing. This paper gives different scheduling strategies and algorithms in cloud computing.*

**Keywords:** cloud computing, scheduling, task, workflow

## 1. INTRODUCTION

Cloud computing dates back to the 1960's when John McCarthy opined that "computation may someday be organized as a public utility". Amazon played a key role in cloud computing development by launching Amazon web service on utility basis in 2006. Before scheduling tasks on cloud computing, the characteristics of the cloud should be taken into account. Some of the characteristics of cloud include

- 1.1 On-demand self service
- 1.2 Ubiquitous network access
- 1.3 Location independent resource pooling
- 1.4 Rapid elasticity
- 1.5 Pay per use[1]

Millions of user share cloud resources by submitting their computing task to the cloud system. Scheduling these millions of task is a challenge to cloud computing environment. Different scheduling strategies are proposed in [2], [3], [4], [5], [6], [7], [8], [9] and [10]. These strategies considers different factors like cost matrix generated by using credit of tasks to be assigned to a particular resource [2], quality of Service (QoS) based meta-scheduler and Backfill strategy based light weight virtual machine scheduler for dispatching jobs [3], QoS requirements [4], [5] and [10], heterogeneity of the cloud environment and workloads [8].

Optimal resource allocation or task scheduling in the cloud should decide optimal number of systems required in the cloud so that the total cost is minimized and the SLA is upheld. Cloud computing is highly dynamic, and hence, resource allocation problems have to be continuously addressed, as servers become available/non-available while at the same time the customer demand

fluctuates. Thus this study focuses on scheduling algorithms in cloud environment considering above mentioned characteristics, challenges and strategies.

## 2. TASK SCHEDULING TYPES

### 2.1 Cloud Service Scheduling

Cloud service scheduling is categorized at user level and system level [11]. At user level scheduling deals with problems raised by service provision between providers and customers. The system level scheduling handles resource management within datacenter.

Datacenter consists of many physical machines. Millions of tasks from users are received; assignment of these tasks to physical machine is done at datacenter. This assignment or scheduling significantly impacts the performance of datacenter. In addition to system utilization, other requirements like QoS, SLA, resource sharing, fault tolerance, reliability, real time satisfaction etc should be taken into consideration.

### 2.2 User Level Scheduling

Market-based and auction-based schedulers are suitable for regulating the supply and demand of cloud resources. Market based resource allocation is effective in cloud computing environment where resources are virtualized and delivered to user as a service. A suite of market-oriented task scheduling algorithms to an AuctionNet for heterogeneous distributed environments is proposed in [12].

Development of a pricing model using processor-sharing for clouds, the application of this pricing model to composite services with dependency consideration and the development of two sets of profit-driven scheduling algorithms are proposed in [13].

Service provisioning in Clouds is based on Service Level Agreements (SLA). SLA represents a contract signed between the customer and the service provider stating the terms of the agreement including non-functional requirements of the service specified as Quality of Service (QoS), obligations, and penalties in case of agreement violations. Thus there is a need of scheduling strategies considering multiple SLA parameters and efficient allocation of resources. A novel scheduling heuristic considering multiple SLA parameters for deploying

applications in cloud is presented in [14]. The scheduler algorithm that allows re-provisioning of resources on the cloud in the event of failures is introduced in [15]. The focus of model is to provide fair deal to the users and consumers, enhanced quality of service as well as generation of optimal revenue. A novel cloud scheduling scheme [16] uses SLA along with trust monitor to provide a faster scheduling of the over flooding user request with secure processing of the request. A novel approach of heuristic-based request scheduling at each server, in each of the geographically distributed data centers, to globally minimize the penalty charged to the cloud computing system is proposed in [17]. This approach considers two variants of heuristics, one based on the simulated annealing method of neighborhood searches and another based on gi-FIFO scheduling.

Based on the queuing model and system cost function, considering the goals of both the cloud computing service users and providers, [18] proposes an algorithm to get the approximate optimistic value of service for each job in the corresponding no-preemptive priority M/G/1 queuing model. This approach guarantees the QoS requirements of the users' as well as the maximum profits for the cloud computing service providers.

To deal with dynamically fluctuating resource demands, market-driven resource allocation has been proposed and implemented by public Infrastructure-as-a-Service (IaaS) providers like Amazon EC2. In this environment, cloud resources are offered in distinct types of virtual machines (VMs) and the cloud provider runs an auction-based market for each VM type with the goal of achieving maximum revenue over time. A case study of single cloud provider and how to best match customer demand in terms of both supply and price in order to maximize the providers revenue and customer satisfactions while minimizing energy cost is proposed in [19]. Another auction-based mechanism for dynamic VM provisioning and allocation that takes into account the user demand for VMs when making VM provisioning decisions is proposed in [20].

### 2.3 Static and Dynamic Scheduling

Static scheduling allows for pre-fetching required data and pipelining different stages of task execution. Static scheduling imposes less runtime overhead. In case of dynamic scheduling information of the job components/task is not known before hand. Thus execution time of the task may not be known and the allocation of tasks is done on fly as the application executes.

A job execution environment Flexitic that exploits scalable static scheduling techniques to provide the user with a flexible pricing model and at the same time, reduce scheduling overhead for the cloud provider has been presented in [21].

The service request scheduling strategies in three-tier cloud structure, which consists of resource providers, service providers and consumers, should satisfy the

objectives of the service providers and consumers. A dynamic priority scheduling algorithm (DPSA) is proposed in [22] to achieve above objectives.

A new fault tolerant scheduling algorithm MaxRe is proposed in [23]. This algorithm incorporates the reliability analysis into the active replication schema, and exploits a dynamic number of replicas for different tasks.

A trust mechanism-based task scheduling model is presented in [24]. Trust relationship is built among computing nodes, and the trustworthiness of nodes is evaluated by utilizing the Bayesian cognitive method.

A feedback dynamic algorithm for preemptable job scheduling mechanism is proposed in [25]. A preemptable scheduling improves the utilization of resources in clouds and feedback procedure in above algorithms works well in the situation where resource contentions are fierce.

In cloud computing, traditional way for task scheduling cannot measure the cost of cloud resources accurately by reason that each of the tasks on cloud systems is totally different between each other. [26] Introduces an optimized algorithm for task scheduling based on ABC (activity based costing) in cloud computing and its implementation. Also an experiment on different optimization strategies for cost-optimal dynamic scheduling in hybrid cloud environments is performed in [27].

To achieve QOS in cloud environment; [28] propose an improved backfill algorithm using balanced spiral (BS) method. This paper analyzed the various parallel job scheduling algorithms like EASY, conservative and CBA. In [29], a scheduling algorithm is proposed that measures both resource cost and computation performance and also improves the computation/communication ratio by grouping the user tasks according to a particular cloud resource's processing capability and sends the grouped jobs to the resource. Due to job grouping, communication of coarse-grained jobs and resources optimizes computation/communication ratio.

A large number of cloud computing servers waste a tremendous amount of energy and emit a considerable amount of carbon dioxide. Green task scheduling is necessary to significantly reduce pollution and substantially lower energy usage. Green task scheduling approaches are proposed in [30] and [31].

Scheduling approaches which considers parameters other than one discussed above are proposed in [32], [33], [26], [34] and [30]. A fully decentralized scheduler in [32] aggregates information about the availability of the execution nodes throughout the network and uses it to allocate tasks to those nodes those are able to finish them in time. As study considering the realistic network topology and communication model, proposes the Deadline, Reliability, Resources-aware (DRR) scheduling algorithm in [33]. Considering the failure and recovery scenario in the Cloud computing entities, [34] proposes a Reinforcement Learning (RL) based algorithm to make

job scheduling fault-tolerable while maximizing utilities attained in the long term.

A new framework of task scheduling strategy for tree network is proposed in [35].

#### **2.4 Heuristic Scheduling**

Optimization problems are in Class NP-hard. These problems can be solved by enumeration method, heuristic method or approximation method. In enumeration method, an optimal solution can be selected if all the possible solutions are enumerated and compared one by one. When number of instances is large, exhaustive enumeration is not feasible for scheduling problems. In that case heuristic is a suboptimal algorithm to find reasonably good solutions reasonably fast. Approximation algorithms are used to find approximate solutions to optimized solution. These algorithms are used for problems when exact polynomial time algorithms are known.

Enhancing task data locality in large scale data processing systems is crucial for the job completion time. Most of the approaches to improve data locality are either greedy and ignore global optimization, or suffer from high computation complexity. This problem is addressed by proposing a heuristic task scheduling algorithm called Balance-Reduce (BAR) in [36].

Load balancing task scheduler balance the entire system load while trying to minimizing the make span of a given tasks set. Two different load balancing scheduling algorithms based on ant colony are proposed in [37] and [38]. Another ant colony based algorithm aims to minimize job completion time based on pheromone is proposed in [39]. Cloud Loading Balance algorithm [40], adds capacity to the dynamic balance mechanism for the cloud environment.

The decision, which workloads to outsource to what cloud provider, should maximize the utilization of the internal infrastructure and minimize the cost of running the outsourced tasks in the cloud, while taking into account the applications' quality of service constraints. A set of heuristics, to cost-efficiently schedule deadline-constrained computational applications, is proposed in [41]. Multi-objective meta-heuristics scheduling algorithm for multi-cloud environment is proposed in [42]. This algorithm tries to achieve application high-availability and fault-tolerance while reducing the application cost and keeping the resource load maximized. Because of the increasing large Web graph and social networks, cost-conscious large graph processing scheduling is important and a heuristic for the same is proposed in [43].

Genetic algorithm based scheduling algorithms are proposed in [44], [45], [46], [47], [48] and [49]. An optimized algorithm based on GA to schedule independent and divisible tasks adapting to different computation and memory requirements is proposed in [46]. Multi-agent genetic algorithm (MAGA) [45] is a hybrid algorithm of GA which solves the load balancing

problem in cloud computing. COA (Course of action) planning involves resource allocation and task scheduling. A robust COA planning with varying durations based on GA is proposed in [44]. Reducing energy consumption is an increasingly important issue in cloud computing, more specifically when dealing with High Performance Computing (HPC). A multi-objective genetic algorithm (MO-GA), proposed in [47], optimizes the energy consumption, carbon dioxide emissions and the generated profit of a geographically distributed cloud computing infrastructure. Another parallel genetic algorithm based resource scheduling is proposed in [48]. Simulated annealing is a generic probabilistic meta-heuristic for the global optimization problem of locating a good approximation to the global optimum of a given function in a large search space. An optimized algorithm for task scheduling based on genetic simulated annealing algorithm in cloud computing is proposed in [50].

The scalability of a computing system can be mainly identified by size, geographical distribution, administrative constraints, heterogeneity, energy consumption and transparency. A low complexity energy-efficient heuristic algorithm for scheduling, proposed in [51], performs efficiently demonstrating their applicability and scalability.

In batch mode, tasks are scheduled only at some predefined time. This enables batch heuristics to know about the actual execution times of a larger number of tasks. Min-min and Max-min are heuristics used for batch mode scheduling. Heuristics based improved Max-min algorithm is proposed in [52] and the QoS Min-Min scheduling algorithm is proposed in [53].

Bag of tasks (BoT) applications are the one which execute independent parallel tasks. Heuristics proposed in [54] aims to maximize resource utilization while executing BoTs in heterogeneous sets of Cloud resources allocated for different numbers of hours. Another budget constraint scheduler proposed in [55] schedules large bags of tasks onto multiple clouds with different CPU performance and cost, minimizing completion time while respecting an upper bound for the budget to be spent. When providers cannot disclose private information such as their load and computing power, which are usually heterogeneous, the meta-scheduler needs to make blind scheduling decisions. In this case a deadline-constrained BoT application scheduling approach is proposed in [56].

#### **2.5 Real Time Scheduling**

The primary objectives of real time scheduling are to increase throughput and minimize average response time instead of meeting deadlines.

The real-time tasks are scheduled non-preemptively with the objective to maximize the total utility in [57]. Two different time utility functions (TUFs)-a profit TUF and a penalty TUF- are associated with each task at the same time. This approach not only rewards the early completions but also penalizes the abortions or deadline

misses of real-time tasks. Similarly a preemptive algorithm is proposed in [58].

A real time workload driven approach is proposed in [59]. Quality of service (QoS) guarantees for some applications such as signal data processing is very important. A novel self-adaptive QoS-aware scheduling algorithm called SAQA [60] considers the adaptability for real-time tasks with QoS demands on heterogeneous clusters.

### 2.6 Workflow Scheduling

A workflow enables the structuring of applications in a directed acyclic graph form [61], where each node represents the constituent task and edges represent inter task dependencies of the applications [62]. A single workflow generally consists of a set of tasks each of which may communicate with another task in the workflow. Workflow scheduling is one of the key issues in the management of workflow execution. A Survey of various workflow scheduling algorithms in cloud environment is documented in [63]. A study of various problems, issues and types of scheduling algorithms for cloud workflows is documented in [64]. A study of Instance-Intensive Cloud Workflows is documented in [65].

### 3. CONCLUSION

This paper explores methods of scheduling done in cloud computing. It helps to understand the wide task scheduling options in order to select one for a given environment.

### 4. FUTURE SCOPE

There is a need of a scheduling policy which will be beneficial to service provider as well as customer. As a part of my project I will be implementing cost based scheduling policy.

### References

[1] M. Malathi, "Cloud Computing Concepts", IEEE, 2011  
[2] Paul, M., Sanyal, G., "Survey and analysis of optimal scheduling strategies in cloud environment", IEEE, 2012  
[3] Jeyarani, R., Ram, R. Vasanth, Nagaveni, N., "Design and Implementation of an Efficient Two-Level Scheduler for Cloud Computing Environment", IEEE, 2010  
[4] Huang Qi-yi, Huang Ting-lei, "An optimistic job scheduling strategy based on QoS for Cloud Computing", IEEE, 2010  
[5] Meng Xu, Lizhen Cui, Haiyang Wang, Yanbing Bi, "A Multiple QoS Constrained Scheduling Strategy of Multiple Workflows for Cloud Computing", IEEE, 2009  
[6] Hao Li, Huixi Li, "A Research of Resource Scheduling Strategy for Cloud Computing Based on Pareto Optimality MxN Production Model", IEEE, 2011  
[7] Kuan-Rong Lee, Meng-Hsuan Fu, Yau-Hwang Kuo, "A hierarchical scheduling strategy for the composition

services architecture based on cloud computing", IEEE, 2011

[8] Gunho Leey, Byung-Gon Chunz, Randy H. Katzy, "Heterogeneity-Aware Resource Allocation and Scheduling in the Cloud", University of California

[9] Shu-Ching Wang, Kuo-Qin Yan, Shun-Sheng Wang, Ching-Wei Chen, "A Three-Phases Scheduling in a Hierarchical Cloud Computing Network", IEEE, 2011

[10] Peixoto, M.L.M., Santana, M.J., Estrella, J.C., Tavares, T.C., Kuehne, B.T., Santana, R.H.C., "A Meta-scheduler architecture to provide QoS on the cloud computing", IEEE, 2010

[11] Fei Teng, "Resource allocation and scheduling models for cloud computing", Paris, 2011

[12] Han Zhao, Xiaolin Li, "AuctionNet: Market oriented task scheduling in heterogeneous distributed environments", IEEE, 2010

[13] Lee, Young Choon, Wang, Chen, Zomaya, Albert Y. and Zhou, Bing Bing, "Profit-Driven Service Request Scheduling in Clouds", IEEE, 2010

[14] Emeakaroha, V.C., Brandic, I., Maurer, M. and Breskovic, I., "SLA-Aware Application Deployment and Resource Allocation in Clouds", IEEE, 2011

[15] Ahuja, R., De, A., Gabrani, G., "SLA Based Scheduler for Cloud for Storage & Computational Services", IEEE, 2011

[16] Daniel, D., Lovesum, S.P.J., "A novel approach for scheduling service request in cloud with trust monitor", IEEE, 2011

[17] Boloor, K., Chirkova, R., Salo, T., Viniotis, Y., "Heuristic-Based Request Scheduling Subject to a Percentile Response Time SLA in a Distributed Cloud". IEEE, 2011

[18] Luqun Li, "An Optimistic Differentiated Service Job Scheduling System for Cloud Computing Service Users and Providers", IEEE, 2009

[19] Qi Zhang, Quanyan Zhu, Boutaba, R., "Dynamic Resource Allocation for Spot Markets in Cloud Computing Environments", IEEE, 2012

[20] Zaman, S., Grosu, D., "Combinatorial Auction-Based Dynamic VM Provisioning and Allocation in Clouds", IEEE, 2012

[21] Thomas A. Henzinger, Anmol V. Singh, Vasu Singh, Thomas Wies, Damien Zufferey, "Static Scheduling in Clouds", Austria

[22] Zhongyuan Lee, Ying Wang, Wen Zhou, "A dynamic priority scheduling algorithm on service request scheduling in cloud computing", IEEE, 2011

[23] Laiping Zhao, Yizhi Ren, Yang Xiang, Sakurai, K., "Fault-tolerant scheduling with dynamic number of replicas in heterogeneous systems", IEEE, 2011

[24] Wei Wang, Guosun Zeng, "Trusted Dynamic Scheduling for Large-Scale Parallel Distributed Systems", IEEE, 2011

[25] Jiayin Li, Meikang Qiu, Jianwei Niu, Wenzhong Gao, Ziliang Zong, Xiao Qin, "Feedback Dynamic

- Algorithms for Preemptable Job Scheduling in Cloud Systems", IEEE, 2010
- [26] Qi Cao, Zhi-Bo Wei, Wen-Mao Gong, "An Optimized Algorithm for Task Scheduling Based on Activity Based Costing in Cloud Computing", IEEE, 2009
- [27] Zinnen, A., Engel, T., Deadline constrained scheduling in hybrid clouds with Gaussian processes, IEEE, 2011
- [28] Suresh, A., Vijayakarthish, P., "Improving scheduling of backfill algorithms using balanced spiral method for cloud metascheduler", IEEE, 2011
- [29] Selvarani, S., Sadhasivam, G.S., "Improved cost-based algorithm for task scheduling in cloud computing", IEEE, 2011
- [30] Mehdi, N.A., Mamat, A.; Amer, A., Abdul-Mehdi, Z.T., "Minimum Completion Time for Power-Aware Scheduling in Cloud Computing", IEEE, 2012
- [31] Luna Mingyi Zhang, Keqin Li, Yan-Qing Zhang, "Green Task Scheduling Algorithms with Speeds Optimization on Heterogeneous Cloud Servers", IEEE, 2011
- [32] Celaya, J., Arronategui, U., "A Highly Scalable Decentralized Scheduler of Tasks with Deadlines", IEEE, 2011
- [33] Laiping Zhao, Yizhi Ren, Sakurai, K., "A Resource Minimizing Scheduling Algorithm with Ensuring the Deadline and Reliability in Heterogeneous Systems", IEEE, 2011
- [34] Bo Yang, Xiaofei Xu, Feng Tan, Dong Ho Park, "An utility-based job scheduling algorithm for Cloud computing considering reliability factor", IEEE, 2012
- [35] Wang Li, Gao-chao Xu, Jia Zhao, Xiao-dong Fu, Yu-shuang Dong, Yu-nan Zhai, Wang Dan, "Task Scheduling Strategy Based on Tree Network in Cloud Computing Environment", International Conference on Mechanical Engineering and Technology (ICMET-London 2011)
- [36] Jiahui Jin, Junzhou Luo, Aibo Song, Fang Dong, Runqun Xiong, "BAR: An Efficient Data Locality Driven Task Scheduling Algorithm for Cloud Computing", IEEE, 2011
- [37] Xin Lu, Zilong Gu, "A load-adaptive cloud resource scheduling model based on ant colony algorithm", IEEE, 2011
- [38] Kun Li, Gaochao Xu, Guangyu Zhao, Yushuang Dong, Wang, D., "Cloud Task Scheduling Based on Load Balancing Ant Colony Optimization", IEEE, 2011
- [39] Xiangqian Song, Lin Gao, Jieping Wang, "Job scheduling based on ant colony optimization in cloud computing", IEEE, 2011
- [40] Zhang Bo, Gao Ji, Ai Jieqing, "Cloud Loading Balance algorithm", IEEE, 2011
- [41] Van den Bossche, R., Vanmechelen, K., Broeckhove, J., "Cost-Efficient Scheduling Heuristics for Deadline Constrained Workloads on Hybrid Clouds", IEEE, 2012
- [42] Frincu, M.E., Craciun, C., "Multi-objective Meta-heuristics for Scheduling Applications with High Availability Requirements and Cost Constraints in Multi-Cloud Environments", IEEE, 2012
- [43] Jian Li, Sen Su, Xiang Cheng, Qingjia Huang, Zhongbao Zhang, "Cost-Conscious Scheduling for Large Graph Processing in the Cloud", IEEE, 2011
- [44] Luohao Tang, Cheng Zhu, Weiming Zhang, Zhong Liu, "Robust COA planning with varying durations", IEEE, 2011
- [45] Kai Zhu, Huaguang Song, Lijing Liu, Jinzhu Gao, Guojian Cheng, "Hybrid Genetic Algorithm for Cloud Computing Applications", IEEE, 2012
- [46] Chenhong Zhao, Shanshan Zhang, Qingfeng Liu, Jian Xie, Jicheng Hu, "Independent Tasks Scheduling Based on Genetic Algorithm in Cloud Computing", IEEE, 2009
- [47] Kessaci, Y., Melab, N., Talbi, E.-G., "A pareto-based GA for scheduling HPC applications on distributed cloud infrastructures", IEEE, 2011
- [48] Zhongni Zheng, Rui Wang, Hai Zhong, Xuejie Zhang, "An approach for cloud resource scheduling based on Parallel Genetic Algorithm", IEEE, 2011
- [49] Yujia Ge, Guiyi Wei, "GA-Based Task Scheduler for the Cloud Computing Systems", IEEE, 2010
- [50] Gan Guo-ning, Huang Ting-lei, Gao Shuai, "Genetic simulated annealing algorithm for task scheduling based on cloud computing environment", IEEE, 2010
- [51] Diaz, C.O., Guzek, M., Pecero, J.E., Bouvry, P., Khan, S.U., "Scalable and Energy-Efficient Scheduling Techniques for Large-Scale Systems", IEEE, 2011
- [52] Gao Ming and Hao Li, "An Improved Algorithm Based on Max-Min for Cloud Task Scheduling", Yunnan University, China, 2011
- [53] Ching-Hsien Hsu, Tai-Lung Chen, "Adaptive Scheduling Based on Quality of Service in Heterogeneous Environments", IEEE, 2010
- [54] Gutierrez-Garcia, J.O., Kwang Mong Sim, "A Family of Heuristics for Agent-Based Cloud Bag-of-Tasks Scheduling", IEEE, 2011
- [55] Oprescu, A., Kielmann, T., "Bag-of-Tasks Scheduling under Budget Constraints", IEEE, 2011
- [56] Netto, M.A.S., Buyya, R., "Offer-based scheduling of deadline-constrained Bag-of-Tasks applications for utility computing systems", IEEE, 2009
- [57] Shuo Liu, Gang Quan, Shangping Ren, "On-Line Scheduling of Real-Time Services for Cloud Computing", IEEE, 2010
- [58] Shuo Liu, Gang Quan, Shangping Ren, "On-line preemptive scheduling of real-time services with profit and penalty", IEEE, 2011
- [59] Yuanyan Gu, Yujia Ge, "A Real-Time Workload Driven Approach for the Cloud", IEEE, 2010
- [60] Zhu, Xiaomin, Zhu, Jianghan, Ma, Manhao, Qiu, Dishan, "SAQA: A Self-Adaptive QoS-Aware Scheduling Algorithm for Real-Time Tasks on Heterogeneous Clusters", IEEE, 2010
- [61] Zhifeng Yu and Weisong Shi, "A Planner-Guided Scheduling Strategy for Multiple Workflow

Applications," icppw, pp.1-8, International Conference on Parallel Processing - Workshops, 2008.

[62] J. Yu and R. Buyya, "Workflow Scheduling Algorithms for Grid Computing", Technical Report, GRIDS-TR-2007-10, Grid Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, May 2007.

[63] Anju Bala, Dr.Inderveer Chana,"A Survey of Various Workflow Scheduling Algorithms in Cloud Environment", 2nd National Conference on Information and Communication Technology (NCICT) 2011 Proceedings published in International Journal of Computer Applications@ (IJCA)

[64] Navjot Kaur, Taranjit Singh Aulakh, Rajbir Singh Cheema, "Comparison of Workflow Scheduling Algorithms in Cloud Computing", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10, 2011

[65] Ke Liu, "Scheduling Algorithms for Instance-Intensive Cloud Workflows", Swinburne University of Technology, 2009

#### **AUTHOR**

**Yogita Chawla** received the B.E. degrees in Computer Engineering from Pune University in 2000. This paper is part of her M.E degree which she is currently pursuing.

**Mansi Bhonsle** is a lecturer in G. H Raisoni College of Engg & Mgmt.

Cloud computing is an emerging technology and it allows users to pay as you need and has the high performance. Cloud computing is a heterogeneous system as well and it ho.. Due to novelty of cloud computing field, there is no many standard task scheduling algorithm used in cloud environment. Especially that in cloud, there is a high communication cost that prevents well known task schedulers to be applied in large scale distributed environment. Today, researchers attempt to build job scheduling algorithms that are compatible and applicable in Cloud Computing environment Job scheduling is most important task in cloud computing environment because user have to pay for resources used based upon time.

4 Major approaches for scheduling in Cloud computing. 5 Analysis of RL and DRL Frameworks in scheduling of Cloud computing. 6 A review of DRL-based resource scheduling in Cloud computing.

Consequently, how to schedule components of Cloud computing in an efficient, energy-saving, balanced method, is proved to be a critical factor influencing the orientation of Cloud computing in society. Whereas, the huge scale of Cloud computing, the complexity of scenarios, the unpredictability of user requests, the randomness of electronic components, and uncertain temperature of various components presented in the running process, pose challenges to efficient and effective scheduling of Cloud computing in the emerg-ing trend[6, 12, 15, 18-22]. In order to replace the traditional Internet software usage patterns and enterprise management mode, this paper proposes a new business calculation mode- cloud computing, resources scheduling strategy is the key technology in cloud computing, Based on the study of cloud computing system structure and the mode of operation, The key research for cloud computing the process of the work scheduling and resource allocation. problems based on ant colony algorithm , Detailed analysis and design of the specific implementation for cloud resources scheduling . And in CloudSim simulation environment and s... Scheduling in cloud computing is a technique which is used to improve the overall execution time of the job. A good scheduling algorithm can help in load balancing as well. Scheduling in cloud can be done in three areas i.e. Scheduling Cloudlets (where Cloudlet is small scale data center that is located at the edge of Internet) within the virtual machine, Scheduling Virtual Machine on the host, and scheduling cloudlets to the Virtual Machine. There are used CloudSim Simulator which is control environment in which programmer can run application instead of practically running it. This paper presents Different VM Scheduling algorithm using CloudSim and trying to find out best algorithm. Scheduling methods are divided into the following categories [9,10]:

1. User level. The developed method resolves the problem that occurred during the service delivery between the cloud provider and the user. Multicriteria optimization method for load balancing in cloud computing. Article. Full-text available.

Cloud Computing is one of the most recent technologies based on virtualization over IT resources where virtual storage and computing services are provided. On the other hands, cloud computing is based on the concepts of virtualization, multi-tenancy, and shared infrastructure. Task scheduling on the available resources (i.e., Virtual Machines (VMs)) is considered one of the main challenges in cloud computing wherescheduling.