A Dynamic Scheduling Scheme for Cloud Computing

S.Sujan, R.Kanniga Devi

Abstract— **Cloud computing, the concept of accessing the resources being stored at some remote location, using them for performing some task required by the user, who may be the end user or the cloud user, rather than the data been stored at the personal computer or a hardware device which is handheld. Cloud computing, the name by itself has got a reach among both the technical and non-technical users, is an emerging area of growth these days. Users have started to migrate from using the normal computing environments to the cloud computing environments. In the same time it has become mandatory to serve the user a better service on the newly developing environment. It is becoming compulsory to provide a reliable environment with all kind of flexibilities to the user. In this paper, a dynamic scheduling scheme for cloud computing is discussed. Considering the resource provisioning as the main issue to be address, the scheduling is being considered as the context. Here we use a dynamic scheduling scheme considering the makespan as the metric. Algorithms like min-min and round robin are compared with respect to the proposed scheme. A real time scenario based model named Berger model [15] is taken in account of comparison with the existing conventional algorithms. By considering and evaluating all the above given constraints and strategies, a makespan based dynamic scheduling scheme for cloud computing is being proposed.**

Index Terms— **Cloud Computing, Resource Provisioning, Scheduling, Berger Model, Makespan**

I. INTRODUCTION

Cloud computing is becoming one of the hotspot in the developing technologies among the users who are both, aware and unaware about the uses and the behavior of the technology. Both the technical users, who is referred as the cloud user and the non-technical user, who is referred as the end user, who has no background knowledge on the operations and the behavior of cloud computing, are increasing day by day in usage. The exploration towards the new and the emerging technologies has become a boon over the recent span of last ten years. Especially with respect to the cloud computing, it is easy to get to know about the number of people using the cloud technologies like mailing, social networks and so on. In the same time, due to the increase in the number of users who are accessing the resources it becomes highly mandatory to provide a good service. In these criteria the concept of scheduling becomes a lot more essential task that has to be provided to the user. The user gets the right thing, available at the right time without any lagging that is happening to occur within the system. The conceptual design of the cloud environment, maintaining the heterogeneity among the users, abstraction of data and all the other concepts bring in a tight challenge to the developers of the cloud environment. Inheriting its core behavior from the conventional computing methodologies like the distributed and the utility computing, the algorithm that are been used in these methodologies can also be incorporated for the cloud computing. To enhance all these concept, first it becomes essential to get into the basic architecture and the design concept of the cloud computing. A cloud computing environment comprises of a set of hosts which does the concept of processing the task that is incoming from the user. Even then all of them are heterogeneous at the real time. Each and every task that is the entering the cloud computing environment is different in each and every attribute of its own. Virtualization is another important backbone of the cloud computing, where a cloud environment is being operated using the Virtual Machines (VM) that has a certain configuration that does the process of servicing the task from the user. In this total design of the cloud computing, the concept of scheduling is done at each and every level which depends on the requirement.

Even though the cloud computing may be inherited form the conventional concepts of the distributed and the utility based computing, the working concept of the cloud computing differs when compared to them. The cloud computing is not happened to operate under any confined scenario with some closed systems like distributed methodologies do, which means cloud computing is purely dynamic. In this case the scheduling methods that are used inside the cloud computing systems should also be dynamic and not static. The dynamic and the static level scheduling varies comparatively where in a static scheduling mechanism, where the scheduling is done in a static environments, aware about all the values like the task it is going to service and the resources that are present to service the task and the other information are readily available. The dynamic scheduling environment is different from the static scheduling environment where the scheduling done at the stage when the task enters the system, in an instant. It doesn't have any prior information about the task incoming or the resources that are present. The cloud environment permits the scheduling in two different methods. One is scheduling the host and the other one is scheduling the VM. The host scheduling is the concept of scheduling the number of virtual machines that are present with respect to a host, as the virtual machines acts as the

Manuscript received Mar, 2015.

***S.Sujan**, Department of Computer Science, Kalasalingam University, Krishnan koil, Srivilluputhur, Virdhunagar District, India. Mobile No: +91-8825855618.*

R.Kanninga Devi, Department of Computer Science, Kalasalingam University, Krishnan koil, Srivilluputhur, Virdhunagar District, Mobile No:+91-9750123659.

ISSN: 2278 – 1323

platforms to perform the operations. Scheduling at the virtual machine is little narrower when compared to the host level scheduling where the task which are incoming to the system, given as the input to the user are being scheduled to the virtual machines present in the host. This is the scheduling at the VM level. In this paper, scheduling the task is being concentrated. Task scheduling may be formally defined as the concept of launching the programs at pre-defined times. When it comes to the dynamic scheduling, it can be done in two different ways, the online mode scheduling and the batch mode scheduling. The online mode of scheduling is being done where the incoming tasks are being scheduled as they arrive in the system. The batch mode of scheduling is little different when compared to the dynamic mode of scheduling where the incoming tasks are collected as a set as they get into the system and executed as they are formed as a confined batch. For a heterogeneous cloud environment the batch mode scheduling is considered as more appropriate when compared to the online mode of scheduling.

II. RELATED WORK

The concept of scheduling the cloud environment is to provide the user a reliable service. In order to maintain a computing platform available to the end user or the cloud user anytime to be accessed, scheduling becomes very much essential. When there is n-number of users who are accessing the cloud system, giving an input of n-tasks then the system must be scheduled so that all the tasks get serviced by the system efficiently without any lag in the performances. There are various types of scheduling starting from the task, VM scheduling, workflow scheduling, datacenter and so on. Many research papers have been submitted and techniques have been proposed on this area. The various types of scheduling in the cloud computing is described as follows:
Scheduling in Cloud Computing VM scheduling Apart from the end users, now a day's all the corporate organizations also have started to migrate towards the cloud environment. This may be due to the flexibility they utilize incurring in a cloud environment in their workspace. The issue is each and every organization comes out with a different requirement of a cloud environment based on their work schedules. Virtual Machine technology is one of the main back bones of the cloud environment Workflow scheduling Workflow scheduling is the concept of managing the execution. The workflow scheduling is the process that maps and manages the inter-dependent tasks on distributive resources. Task scheduling Task scheduling mechanism is being used in the cloud scheduling that can service a particular task or a selected task at a selected period of time interval. The scheduling is being done on a cloud environment considering a set of events as a task and servicing them in a confined period of time using the resources provided. Scheduling conceptually varies from the task scheduling. Scheduling is done where all the events or the jobs are being queued sequentially in the queues usually defined as a task queue. A task queue is being processed based on some particular constraint. It may be of time or as a shift.

Ye Huang et al., [1] compared a Meta computing and scheduling architecture model for net wide resources. The community aware scheduling algorithm and compared the above given architecture model. The make span has been considered as the metric. The extensive experimental evaluation with a real grid workload trace dataset shows that, when compared to the centralized scheduling scheme with BestFit as the metascheduling policy, the use of CASA can lead to a 30

J.Octavio Gutierrez-Garcia et al., [2] proposed a technique in which the Cloud scheduling heuristics are adapted to the resource allocation settings (e.g., 1-hour time slots) of Clouds by focusing on maximizing Cloud resource utilization. This is done based on the remaining allocation times of Cloud resources. Cloud scheduling heuristics, supported by information about BoT tasks (e.g., task size) and/or performance of the cloud resources.

Sen Su et al., [3] proposed a new mechanism aside from the conventional scheduling mechanism. The conventional scheduling mechanism stands only in reducing the make span without a mechanism that focuses on reducing the monitory cost of the operation. A cost-efficient task-scheduling algorithm using two heuristic strategies has been presented. The first strategy dynamically maps tasks to the most cost- efficient VMs based on the concept of Pareto dominance. The second strategy, a complement to the first strategy, reduces the monetary costs of non-critical tasks.

Nakku Kim et al., [4] a technique to incur the operational cost that is incurred in the deployment of the cloud services has been proposed. He suggests a model for estimating the energy consumption of each virtual machine without dedicated measurement hardware based on in-processor events generated by the virtual machine.

Wei Wang et al., [5] in this paper a novel Bayesian method based cognitive trust model, and a trust dynamic level scheduling algorithm named Cloud-DLS by integrating the existing DLS algorithm has been proposed. This model of work has been inspired by Bayesian cognitive model and referring to the trust relationship models of sociology.

L. Grandinetti et al., [6] considering the job scheduling as the metric a multi-objective mathematical formulation of the job scheduling problem is been proposed. This is being proposed considering a homogeneous cloud computing platform. In order to optimize the total average waiting time of the jobs, the average waiting time of the jobs in the longest working schedule (such as the make span) and the required number of hosts.

Joel J.P.C. Rodrigues et al., [7] considering energy as a metric proposed a distributed media aware flow scheduling considering the number of sensor nodes and the mean waiting time as the graph values obtained the result.

M.Geethanjali et al., [8] a truthful dynamic workflow scheduling mechanism for commercial multi-cloud environments has been proposed. The commercial multi cloud environment varies widely from the conventional cloud

setup. Considering the following metrics, the time taken for executing the job and the cost the user has to pay for the job she has proposed a truthful dynamic algorithm for scheduling the tasks.

M. Mezmaz et al., [9] using the NP hardness problem proposed bi-objective genetic algorithm for that takes into account, not only makespan, but also energy consumption. This method is based on dynamic voltage scaling (DVS) to minimize energy consumption.

Marc Eduard Frncu et al., [10] considering the job scheduling mechanism, a technique on Scheduling highly available applications on the cloud environment has been proposed. Inspiring the property of previously used application scaling property, a multi-objective scheduling mechanism that holds with the probability of scheduling highly available environments, has been proposed.

Saeid Abrishami et al., [11] considering the workflow scheduling mechanism, an algorithm for deadline constraint workflow scheduling has been proposed. This model is taken for Infrastructure as a service cloud. He has proposed an ICPCP (Infrastructure Cloud Partial Critical Paths) algorithm derived from the conventional PCP algorithm. He has obtained results considering the deadline factor and normalized cost.

Zhipiao Liu et al., [12] taking cost as the metric of work, considered cloud services with data access awareness. A profit driven scheduling for cloud services with data access awareness has been proposed. Using the service request scheduling algorithm he has proposed a new algorithm called max profit scheduling algorithm new optimization algorithm for profit-driven service request scheduling based on dynamic reuse, which takes account of the personalized SLA characteristics of user requests and current system workload.

Chia-Ming Wu et al., [13] considering the job scheduling technique using the conventional job scheduling mechanism, an energy efficient scheduling mechanism for the cloud centers has been proposed. A scheduling algorithm for the cloud datacenter with a dynamic voltage frequency scaling technique, where the scheduling algorithm can efficiently increase the resource utilization, has been proposed; hence, it can decrease the energy consumption for executing jobs.

Sheheryar Malik et al., [14] considering time as the metric in scheduling concept, proposed a Latency based group discovery algorithm for network aware cloud scheduling. A model for the grouping of nodes with respect to network latency has been proposed as the technique. The application scheduling is done on the basis of network latency.

Baomin Xu et al., [15] understanding the conventional cloud scheduling mechanism inspired by the economic model, a model for dynamic scheduling has been proposed. The resource allocation using the fairness constraint and the justice function has been compared. The first constraint is to classify user tasks by QoS preferences, and establish the general expectation function in accordance with the classification of tasks to restrain the fairness of the resources in selection process. The second constraint is to define resource fairness justice function to judge the fairness of the resources allocation.

III. PROPOSED WORK

More than a distributed or utility based computing environment, the cloud computing environment has reached the end user in a much better and a faster way. This model works on pay for what you use method. It is enough that the user pays only for what he uses and not for everything else that is associated with respect to the provider. Hence, the efficiency in the resources we provide and monitoring the efficiency of the resources we provide to the end user becomes a lot more essential. Here the concept of task scheduling is being concentrated on providing the efficient resource and provisioning them based on scheduling the task. A model name Berger model[15], which has got a real time connect with the economic model of a society is being taken into account for the proceedings.

A. Berger model

Virtualization technology used in the cloud computing is taken as the foremost idea for the Berger model [15]. The concept of virtualization that is incorporated with the commercialism which is being used in the cloud computing to deliver the user cloud service in an efficient manner. This is being obtained based on the fairness constraint. A general cloud model consist of the resources present in the cloud, set of tasks that are entering in the cloud environment. The same way Berger model [15] connects the relation with respect to an economic model.

The characters that are being commercialized in the cloud environment are being compared in the Berger model [15]. The idea of comparison that is being used in the Berger model [15] is social model. The social wealth that is being present in the society is being mapped and monitored here. For an instance the society that has wealth, is taken as the model for comparison with the resources that are provided in the cloud environment. The society anyway is going to utilize the resources that are present for it for the completion of any tasks, like the same way the resources present in the cloud environment are also used by the VMs present in the system for servicing the incoming tasks. So that, a sense of line falls in between comparing the wealth of the society to the resources present in the cloud. In an account of this real time model makespan is being considered as the metric here. Makespan may be defined as the time taken by the task to get serviced by the resources present in the system. Formally defining, the time calculated when the task enters the system to the last task getting executed. The Berger model [15] that is proposed here concentrates only on the intelligent allocation of resources based on a comparative and a logical basis. The same model doesn't concentrate about the makespan metric and the efficient usage of the makespan metric.

ISSN: 2278 – 1323

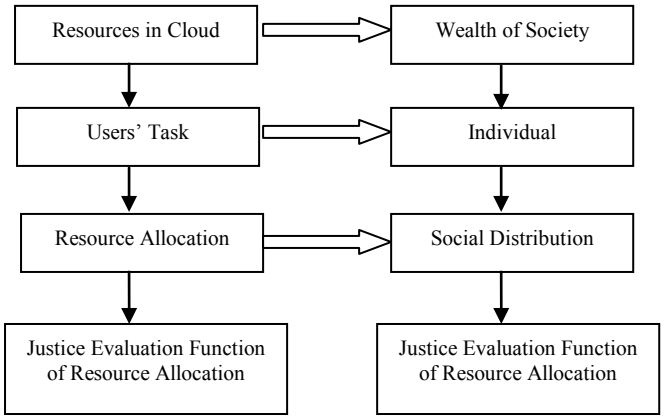

Fig. 1. Mapping between Cloud and Berger model

Here the fairness constraint is obtained based on the comparison that is being obtained from the Berger model [15] to the other conventional algorithms like min-min and round robin.

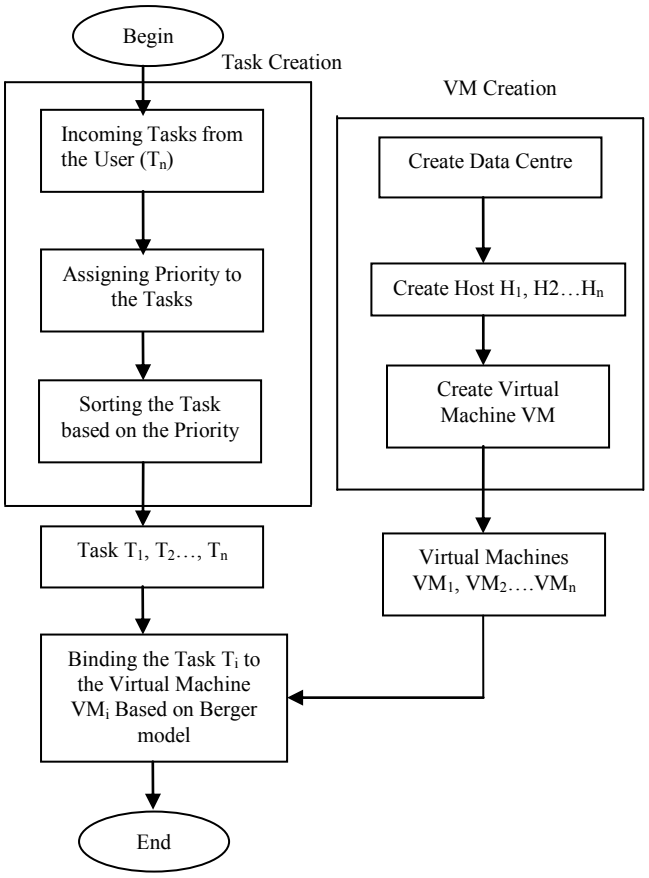

Fig. 3. Berger Model

Based on this the fairness constraint is being obtained the user satisfaction is being calculated. The user satisfaction with respect to the fairness constraint is something that is what the user needs with the cloud environment and what the user gets from the cloud environment. In a real time computing environment, in a cloud, each and every user needs a different kind of resource to be used to service his task and at the same time each and every task with respect to a single user differs. This means that the level of heterogeneity in the cloud environment is varying greatly with respect to the task incoming to the system and the resources. From this heterogeneity the fairness is being obtained based on the metric, the makespan. For a heterogenic environment with different task and resources, the task is being mapped to the set of virtual machines that are present in the hosts of the cloud environment. The virtual machine receives the tasks which are being mapped to it and services the appropriate task of it using the configuration that is provided. For each and every strategy of the task and the configuration of the virtual machine that is present in the cloud environment, the mapping strategy differs in order to maintain the efficiency in servicing the tasks. From the executed tasks the makespan is being obtained. The same way the tasks are being serviced based on the mapping done by some conventional min-min algorithms and the makespan is being obtained. From the obtained values a graph that shows the comparison is being generated.

IV. SIMULATION STUDY

A new version of an existing cloud simulation toolkit is released in the University of Melbourne by the cloud project, the CloudSim. The current version of the CloudSim is version 3.0.3 which has successfully passed a year after its release to the real world. Cloud computing being heavily adopted by the industries cloud computing has become mainstream. Not only in the development industry but also in research and development cloud computing has never failed to occupy the hot seat. In this proposed methodology the Berger model [15], is implemented in this CloudSim (3.0.3) toolkit. The proposed idea is being verified by developing a heterogeneous cloud environment using the simulation toolkit. The simulation toolkit uses a specific name as an identity to point out the attributes, in such a way a task is coined using the name cloudlets. The operation of a host and its behavior is coined using the name datacenter. The conceptual mapping of the datacenter to the VM is being done using the data centre broker. A specific constraint like an algorithm or a mechanism is loaded inside the data center broker. Each and every task has different complexions and shows variations in the execution and so the processing capacity that is possessed by the datacenter broker and VM differs. The cloudlet that is executed inside the VM is expressed in terms of Million Instruction per Second (MIPS) The calculation of the MIPS stays homogeneous. The min-min and the round robin algorithms are calculated.

Table 1. Simulation parameters

Parameter	Value
Configuration of Data center	
Data center architecture	X86
Data center OS	Linux
VMM	Xen
Configuration of Hosts	
No of Hosts	5
MIPS	1000
RAM	16384
Storage	1000000
Bandwidth	10000
Configuration of VMs	
No of VMs	13
Size	10000
MIPS	250
RAM	512
Bandwidth	1000
No of PEs	3
Configuration of Cloudlets	
No of Cloudlets	10-50
Length	Varying
File Size	300
Output Size	300
No of PEs	1

The simulation parameter table displayed above shows the different parameters with different configurations of the cloudlets, VMs and the datacenters resulting in heterogeneous cloud environment.

Considering the above graph with makespan as the metric the variations in a heterogeneous cloud environment with respect to min-min and the berger model is being observed. A datacenter borker loaded with two different concepts, one with the existing concept min-min and the other with the proposed concept, the berger model is being loaded in the broker. When a set of tasks/cloudlets are set as an input to the incoming environment, the datacenter broker does its job of mapping the tasks to the VMs.

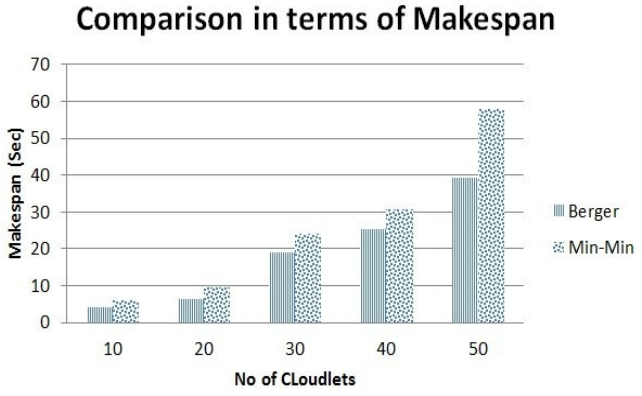

Fig. 3. Comparison in terms of Makespan

Hence the cloudlets get executed in the VMs and the output is being obtained. First the datacenter broker is loaded with the conventional min-min algorithm and the output is obtained. The same is done for the berger model concept and the output is obtained. The values are tabulated in a graph which shows the berger model happens to perform efficiently than the conventional algorithm.

V. CONCLUSION

The cloud computing is the technology for the people, of the people and by the people. The cloud computing has become a people oriented technology where all sort of people start using the technology starting from students, graduates, professionals, scholars and so on. The entire technology is moving towards the next phase of evaluation keeping the cloud computing as one of its main pillar of support. Cloud computing too plays a vital role in migrating and accessing of the data in a much faster rates. Data could be stored at any location which is remote and it can be accessed anytime provided with an internet connection. In such a real-time scenario of evolution strengthening the technology becomes morally essential for the people associated with it. Taking scheduling as the context an approach considering the makespan as the metric a methodology has been proposed. The berger model has executed and tested itself to perform better than the conventional min-min algorithm by reducing the makespan in the execution of the cloudlets. This has been clearly obtained in the graph.

VI. FUTURE WORK

The future work this proposed methodology could be extended with various metrics like bandwidth, energy. The cloud computing obviously needs an internet connection to work and of course an internet connection in concerned with the bandwidth. The cloud environments execution is obtained using the makespan metric which could be, in further extended with bandwidth and the energy utilized within the system. The proposed model, with respect to the CloudSim simulation toolkit can be used to extend the work considering the bandwidth as another metric in addition with the makespan.

ACKNOWLEDGMENT

The authors would like to thank the Head of the Department for the kind support and guidance provided for the conference on due course of the project work.

REFERENCES

[1] Exploring decentralized dynamic scheduling for grids and clouds using the community-aware scheduling algorithm Ye Huang a,d, Nik Bessis b,c, Peter Norrington B , Pierre Kuonen d, Beat Hirsbrunner. Future Generation Computer Systems 29 (2013) 402–415

[2] A family of heuristics for agent-based elastic Cloud bag-of-tasks concurrent scheduling J. Octavio Gutierrez-Garcia, Kwang Mong Sim Future Generation Computer Systems 29 (2013) 1682–1699

[3] Cost-efficient task scheduling for executing large programs in the cloud Sen Su, Jian Li, Qingjia Huang, Xiao Huang, Kai Shuang, Jie Wang, Parallel Computing 39 (2013) 177188.

[4] Energy-credit scheduler: An energy-aware virtual machine scheduler for cloud systems Nakku Kim, Jungwook Cho, Euiseong Seo, Future Generation Computer Systems 32 (2014) 128137

ISSN: 2278 – 1323

[5] Cloud-DLS: Dynamic trusted scheduling for Cloud computing Wei Wanga, Guosun Zeng , Daizhong Tang, Jing Yao, Expert Systems with Applications 39 (2012) 2321–2329
[6] An approximate ϵ -constraint method for a multi-objective job scheduling in the cloud L. Grandinetti, O. Pisacane, M. Sheikhalishahi, Future Generation Computer Systems 29 (2013) 1901–1908
[7] Distributed media-aware flow scheduling in cloud computing environment Joel J.P.C. Rodrigues, Liang Zhou, Lucas D.P. Mendes, Kai Lin, Jaime Lloret, Computer Communications 35 (2012) 1819–1827
[8] Truthful Dynamic Scheduling Algorithm Based on Multi objective for Multi cloud Environments M.Geethanjali, J.Angela Jennifa Sujana T.Revathi, Volume 3, Special Issue 3, and March 2014. 2014 IEEE International Conference on Innovations in Engineering and Technology (ICIET'14)
[9] A parallel bi-objective hybrid meta heuristic for energy-aware scheduling for cloud computing systems M. Mezmaz, N. Melab, Y. Kessaci, Y.C. Lee, E.-G. Talbi, A.Y. Zomaya. Tuyttens J. Parallel Distrib. Comput. 71 (2011) 1497–1508,
[10] Scheduling highly available applications on cloud environments Marc Eduard Frîncu, Future Generation Computer Systems 32 (2014) 138–153
[11] Deadline-constrained workflow scheduling algorithms for Infrastructure as a Service Clouds Saeid Abrishami, Mahmoud Naghibzadeh, Dick H.J. Epema, Future Generation Computer Systems 29 (2013) 158–169
[12] Profit-driven Cloud Service Request Scheduling Under SLA Constraints Zhipiao Liu, Qibo Sun, Shangguang Wang, Hua Zou, Fangchun Yang, Journal of Information & Computational Science 9: 14 (2012) 4065–4073 <website> Available at http://www.joics.com
[13] A green energy-efficient scheduling algorithm using the DVFS technique for cloud datacenters Chia-Ming Wu, Ruay-Shiung Chang, Hsin-Yu Chan, Future Generation Computer Systems 37 (2014) 141–147
[14] Latency based group discovery algorithm for network aware cloud scheduling Sheheryar Malika, Fabrice Huet, Denis Caromel, Future Generation Computer Systems 31 (2014) 28–39
[15] Job scheduling algorithm based on Berger model in cloud environment Baomin Xu Chunyan Zhao b , Enzhao Hu a , Bin Hu c,d, Advances in Engineering Software 42 (2011) 419–425
[16] Design and implementation of a cloud computing service for nite element analysis Ismail Ari ., Nitel Muhtaroglu, Computer Science Department, Ozyegin University, Istanbul, Turkey, Advances in Engineering Software 60-61 (2013) 122135
[17] Towards cost efficient mobile service and information management inubiquitous environment with cloud resource scheduling, Xin Li, Zhuzhong Qiana, sun You, Sanglu Lu, State Key Laboratory for Novel Software Technology, Nanjing University, China, School of Information Science, Korean Bible University, South Korea, International Journal of Information Management 34 (2014) 319328
[18] MDP based optimal pricing for a cloud computing queueing model Rami Atar, Israel Cidon, Mark Shifrin, Department of Electrical Engineering, TechnionIsrael Institute of Technology, and Haifa 32000, Israel.
[19] Hierarchical Scheduling Mechanisms for Multilingual Information Resources in Cloud Computing Yaojun Hana, Xuemei Luob 2013 AASRI Conference on Parallel and Distributed Computing Systems
[20] An ACO-inspired algorithm for minimizing weighted flow time in cloud-based parameter sweep experiments Cristian Mateos, Elina Pacini, Carlos Garc Garino a ISISTAN Research Institute, UNICEN University, Campus Universitario, Tandil B7001BBO, Buenos Aires, Argentina, Advances in Engineering Software 56 (2013) 3850

Authors Profile

S.Sujan is currently a PG scholar in Computer Science from the Department of Computer Science and Engineering at Kalasalingam University, Krishnankovil, Srivilluputhur, Virudhunagar District and Tamilnadu. He received his Bachelor Degree in Computer Science Engineering from Mahendra College of Engineering, Salem and Tamilnadu. His Research areas include Cloud Computing, Grid Computing and Distributed Computing.

R.Kanniga Devi received her Bachelor of Engineering Degree from Madurai Kamaraj University, Madurai in 2001 and Master of Engineering from Anna University, Chennai in 2008. She is currently doing her doctoral program from Kalasalingam University, Srivilluputhur. She is a member of IEEE, CSI, IAEng, ISTE and Network Technology Group of TIFAC-CORE in Network Engineering. She has published more than 10 research papers in International Journals, International and National Conferences. Her research interest includes Cloud Computing, Theory of computation and Graph theory applications in networks. She is currently working as Assistant Professor in the Department of Computer Science and Engineering at Kalasalingam University, Krishnan koil, Tamil nadu, India.

