

An Enhanced Job Scheduling in Cloud Environment Using Probability Distribution

M. Kowsigan¹, S. Kalicharan², P. Karthik³, A. Manikandan⁴, R. M. Manikandan⁵
Department of Information Technology,

Sri Krishna College of Technology, Coimbatore, 641042, India

¹kowsigan.m@skct.edu.in, ²14n122@skct.edu.in, ³14n124@skct.edu.in, ⁴14n139@skct.edu.in
⁵14n140@skct.edu.in

Abstract— Today's world is an information technology's era in that cloud computing arises as promising and developing technology. In the surroundings of cloud computing the resources are provisioned on the basis of demand, as and when required. A giant number of clients (uses cloud) in computation of cloud, can request a number of services or cloud services at the very same time The users demand to access resources are increasing now-a-days, due to this demand it becomes very hard in cloud for allocation of cloud resources accurately and efficiently to the customers, that should satisfy requirements of customers or users and preserve the SLA (service level agreement). Cloud faces many challenges as it is evolving gradually, one of them is scheduling. Here, we contemplate job scheduling, in accordance to the type, of the mission is and varying situation. To efficiently increase the allocating of resource in cloud, one of the foremost job performed is job scheduling, so to get highest profit. Here, we apply, one among of the effective algorithm, first-in-first-out (FIFO), along with markov process technique to prevent blocking probability.

Keywords: Cloud computing, Job Scheduling, First-in-First-out (FIFO), Probability Distribution Models

I. INTRODUCTION

To enable companies have a good components, so to have a computing resource by cloud computing, like storage, virtual machine or also application, also kind that electricity, rather than to make and to maintain those infrastructure which is computing, in their home or some building. Cloud computing possess several pleasing gains for end users and businesses. The three big gain, cloud computing have, Self-service provisioning: End users can spin up compute resources for about and almost any type of workload on demand. Therefore it removes the old need for administrators involved in Information Technology and handle compute resources and to provision it. Elasticity: Companies can maximize its size, as computing needs increases and minimize it again as per demands decreases. Therefore eliminating the need of huge investments in local infrastructure had been reduced. Pay as per use: Computation of resources is being measured, that allows the users in a way that is to charge for the necessary resources and workloads being used. Cloud Computing provides various services related to infrastructure, software, and platform. The three basic models of Cloud Computing (CC) are Iaas which is Infrastructure as a Service, Paas which is Platform as a Service, and Saas which is Software as a Service displayed in layered architecture of Cloud [1].

- 1) *Infrastructure-as-a-Service (Iaas)*: Infrastructure-as-a- Service is the layer of base at computation. Iaas delivers hardware as the service. It includes servers, network, storage, virtualization technology, file systems, and operating systems. Cloud providers of Iaas give the above mentioned resources on-demand from their large pools located in data centers [2].
- 2) *Platform-as-a Service (Paas)*: PaaS model includes development tools, database, web servers and execution runtime environment. This model concentrates on providing cost-effective, efficient environment for the development of high-quality applications.
- 3) *Software-as-a-Service (SaaS)*: SaaS or other name known as “on-demand-software” is generally charge on pay-per-use basis. In this model, the cloud service providers install the application (software) on the cloud and it is accessed by the cloud users from the cloud itself. This takes away the need of installing and running the applications on the cloud user's personal computers [3].

Cloud computing domain can be categorized into private cloud, public cloud and hybrid cloud. Private cloud has limitation, to only an organization, and run by the organization itself or any third party cloud service provider. Public cloud is available on the network and is open publicly. Hybrid cloud has basically the blend of public cloud as well the private cloud. An optimal solution was provided for job scheduling in cloud using soft computing techniques [12]. Meta heuristic algorithm was also used to improve the scheduling in cloud environment [13].

II. RELATED WORK

Cloud is accessible everywhere by which user can connect through a link, through cloud. Scheduling techniques are to be chosen very carefully due to wide area of cloud. Cloud is often used as personal organizations like that of private cloud plus public cloud, for public to use. Combination of both public plus private cloud is formed as hybrid cloud. Cloud differs in size according to service supplier and its use. Cloud shall be retrieved using diversification of the devices specified via, portable computer and multimedia systems, PC and cell phones. The diagram given below shown that is used as cloud computing overview for our analysis. Fig. 1 is containing several components or equipments and these are as: User: - User request for application. Task is generated at this end. Switch: - It is used to joint several connections at a single point. Cloud: - It is working like a long distance internet connection. Load balancer: - act as load distributor, randomly to servers. Server: - Servers are used to execute user requests and applications are running at this point. These are connected as viewed in Fig. 1

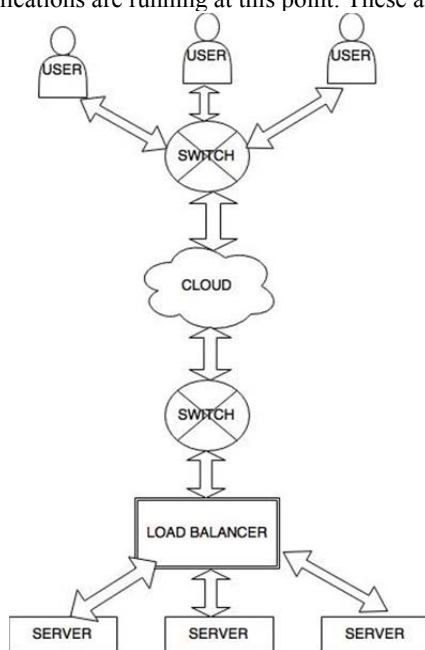


Fig. 1. Cloud computing overview

For allocating of resource problem, some works employ simple single resource abstractions. In these methods, a note is cut in to slots with fixed fractions of resources, and allocates resources jointly at the slot granularity, plus the allocation of alone resource is not enough, multi-resource allocation has to complete the job schedule. [4], [5], [6]-[8] Shows on presenting system investigation on the many resources allocation problem. They propose DRF to make same, the share of all jobs. In these works, both the efficiency and fairness notions have been guaranteed by capturing the trade-offs between allocation fairness and efficiency. For example, in the BFF (Bottleneck-Based Fairness) research [9], authors suggest two fairness properties that DRF possesses are also guaranteed. It considers another setting of DFR with a more common domain of user utilizes, or, stretches the all-in-one resource model to a multi-resource allocation model, which was DRFH, that tells the thought of the DRF from a server to many heterogeneous servers. These works assume, explicitly or implicitly, that all those resources are focused into a super computer, which is not the case in the general datacenter system. The allocation of system resources to various tasks is known as scheduling. Scheduling, done under cloud computing, performed to achieve high performance and the outstanding network throughput. Speed, efficiency, utilization of general resources in an effective way depends largely on way of scheduling selected for cloud computing surrounding. Various criteria for scheduling are max CPU utilization, Max throughput, Min turnaround time, Minimum of waiting time and Min response time. Throughput denoted number of processes that finishes their performance in single time required in unit [10]. Time of turnaround denoting the mass of measure (time) to implement some of set, or a precise process, that is the span of submission time of a task or operation to the time needed for completion. Waiting time denotes the tally of time or period spent to wait in ready queue. Response time represents period from the request submitted up to the initial or foremost reply is produced. Various issues exist in algorithms of basis of scheduling on different optimization criteria. Turnaround time and throughput are the two required criteria in the common system as batch systems, response time and fairness are the two criteria required in interactive system, whereas in real-time system, that is meeting time limit, is aspect which is important. Auto Associative memory network had been used to improve the job scheduling in cloud environment [11]. Therefore a algorithm of scheduling denoted to be selected, that it satisfies the required criteria and provide efficient service and proper allocation of resources.

III. EXISTING SYSTEMS

In cloud environment, it becomes unsuitable for analyzing the function in computational of cloud data centers because the numbers, of the servers is small relatively, like below 10. Approximations are sensitive to the distribution in view of probability of task service times. Bags-of-task will appear when user submits many tasks at a time. In view of active label of cloud environments, diversity of user's requests and time dependency of load increases. Traffic intensity is very high. The variation involved in time of service increases the coefficient. Modeling errors are the facts in the domain of cloud. Cloud computing is deployed all over, the provocation that faces involves the skill to make a greatly compounded set of subcomponent (network resources, compute, storage,) that extent big geographic areas serving diverse unlike clients. To make the process easy, this present COPE, the abbreviation of Cloud Orchestration Policy Engine, a circulated policy so that it allows the CP which is cloud providers, in order to operate cloud resource orchestration. It (COPE) shows, the cloud providers specifying goals and system-wide constraints using COPE log, which is a policy language, focused towards describing distributed rule optimization. COPE takes states of cloud system and policy specifications as an input, uses best of storage, compute and resource allocating in the surrounding of cloud such that operations in objectives and client SLAs (service level agreement) should be met better. Here it is described the initiated integration with the cloud system of cloud orchestration, and shows initial evaluation consequences, which demonstrate the viability of COPE using production traces from a large hosting company like in US. It is also further, discussed an orchestration outline that address geographically widespread data centers, and resulted with a happening status of our work. Operational objectives are customized in cloud computing. Sophisticated cloud services needs dynamic orchestration for the service abstraction.

Appearing Peer-to-Peer (P2P) as collaborate systems need discoveries, multi parameter, scattering, utilizing of various, and groups of resources that are dynamic so that, to achieve high set tasks further processor cycle sharing and conventional file. Collaborations involving of service of best of objective that is dynamic and those specific resources containing application are being stress, to now P2P architectures. Salient characters and wise futuristic of collaborated P2P systems are highlighted enough. Resource and its advertising, binding, selecting, and matching the crucial phases in those of systems, with that of the challenges, are assessed with example, that is from distributed and collaboratively adaptive systems that is sensing, mobile socially networks and cloud computing. State-of-the-art resource clustering solutions or discovery are contrast with that to the respect of their overall architecture, load balancing, lookup overhead, to head on their ability and to reach the goals and challenges of each critical phase. Incentives trust, and security issues and privacy are taken also for discussion, and they will eventually determine the triumph of a collaborative of P2P system. Open issues of research opportunities are essential to attain the true probable of collaboratively of P2P systems are discussed. Enhanced security is provided in distributed wireless networks [14]. Personal data and files aren't protected causing lack of safety. Unstructured networks also arise from this lack of structure. In a simplification, suppose a peer, wanting to see a wished piece of data or record in the network, search query, be engulf through the whole network so that to find vast peers required that share dates.

IV. PROPOSED SYSTEM

The task initially is dispatch to cloud center, is maintained as it its serviced within a satisfactory facilitated node; and when the service or services has been finished, the task quit or leaves center. A space (facility) node may contain different computing resources like the directory servers, web servers, database servers and others. A service level agreement(SLA), figures all particular or characteristics of usage for cloud service and the responsibility of service consumers and also the providers, and that includes several descriptors as a whole, denoting to as best Service of quality, abbreviated as QoS. QoS includes reliability, security, availability, throughput, and other vast parameters, and the implementation indicators like the blocking probability of task, response time, probability of quick service, and in the system, number of tasks taken in average. These are to be determined with the method of queuing theory. Then the undertaking will be processed in corresponding cloud server which based, according on user category where scaling depend on it. A cloud server system, that we model, indicates the request's inter-arrival time. It is exponentially distributed, while time for service of the task is, equally disseminated variables and independent. It follows a generally distributed with mean value. The system under the thought, hold m servers which supply service in the order of FCFS that is the task request arrivals. As the distinctive cloud center's population size is relatively high while that of the probability, and in the view that the user will request service is somewhat small, the process of the arrival is modeled to as a Markovian process. We tender the algorithm to aim the application allocation or request placement and the load distribution. Traffic Intensity is less. Technique of analytical model on the view of approximated model of Markov chain provides best performance evaluation for the cloud environment. General Service time makes our model flexible in the phrase of scalability and diversity of service time. For the total of tasks, taken as average, blocking probability, probability, response time the accuracy is in high degree and Energy saving. The architecture diagram for the proposed system is shown in the Fig. 2.

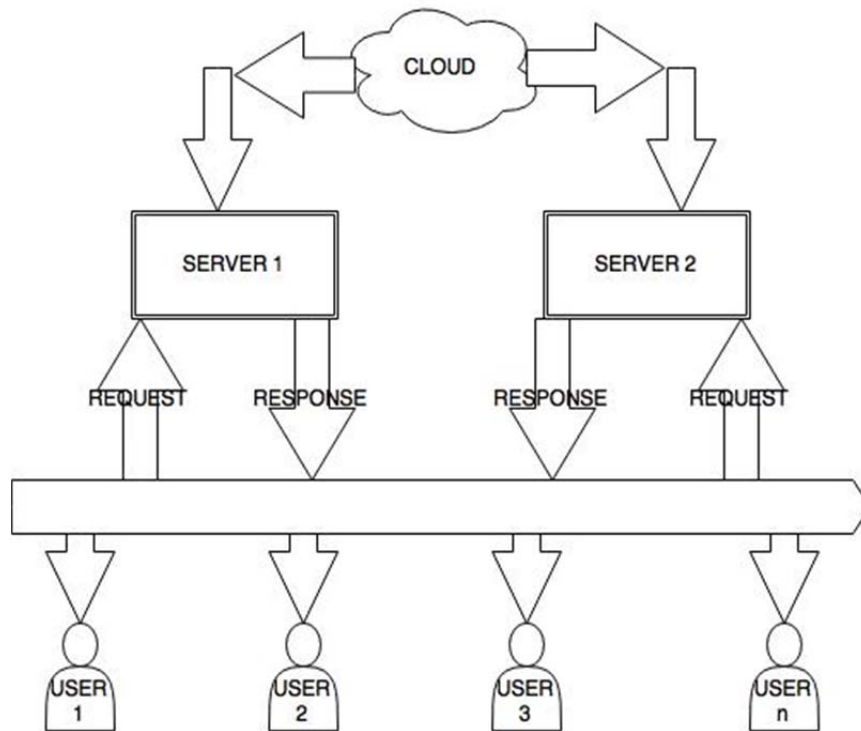


Fig. 2. Architecture of Proposed system

V. RESULTS AND DISCUSSIONS

1. *Server Formation:* In our project, the server calculates which cloud doing which job. That is controlling the access of cloud, calculation of cost also sharing of works in cloud equally. So first, to determine it, servers or cloud are formed like in fig.3 and fig.4.

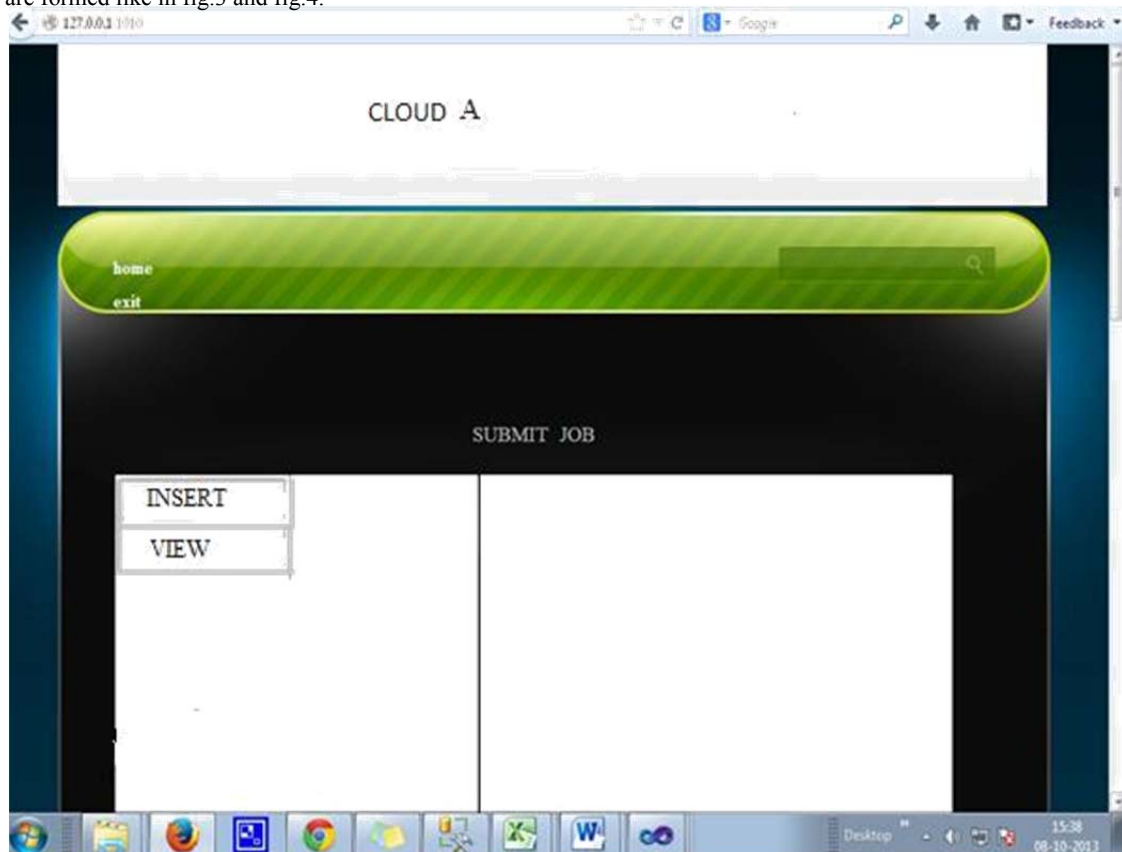


Fig. 3. Cloud A formation



Fig. 4. Cloud B formation

2. *Analysis of Performance:*

The analysis and comparison is done on the performance offered by different configurations of the computing cluster, and is focused in the function desired of applications which are loosely coupled. In specific, here it is chosen cluster configurations that are different with dissimilar worker nodes from providers and dissimilar number of job we use the following as; cloud A and cloud B as in fig.5.



Fig. 5. Performance analysis - Cloud A and cloud B

3. *Task Scheduling:*

Each and every user assigns the task to cloud, so that work, assigned to Cloud in basis of priority scheduling. Task can include like, file upload, and file download, viewing record from cloud or inserting record in cloud as in Fig. 6.

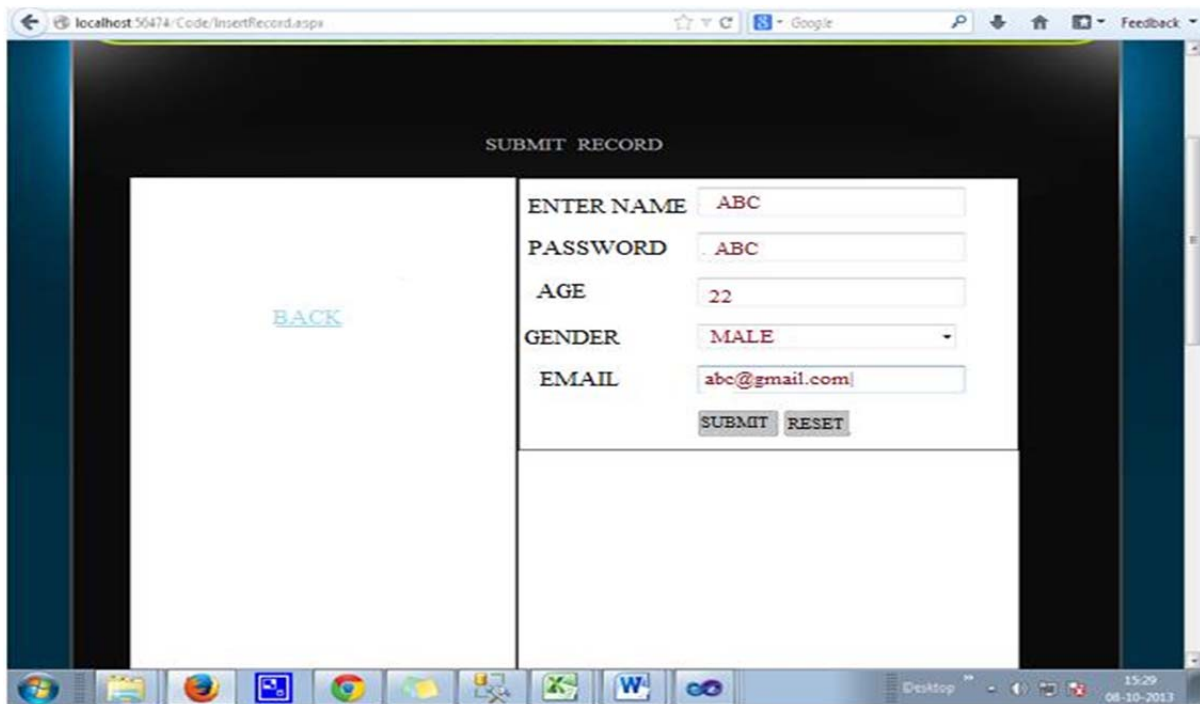


Fig. 6. Insertion of a record

4. *Predicting Result:*

If we assign the job in priority scheduling way to cloud, we get an output correctly and shortly. The amount or cost will be reduced and transferred to cloud owner of the using of cloud. Fig. 7 showing that the record is inserted in cloud and Fig. 8 showing the important parameter like processing time which is for example as 0.399 seconds.

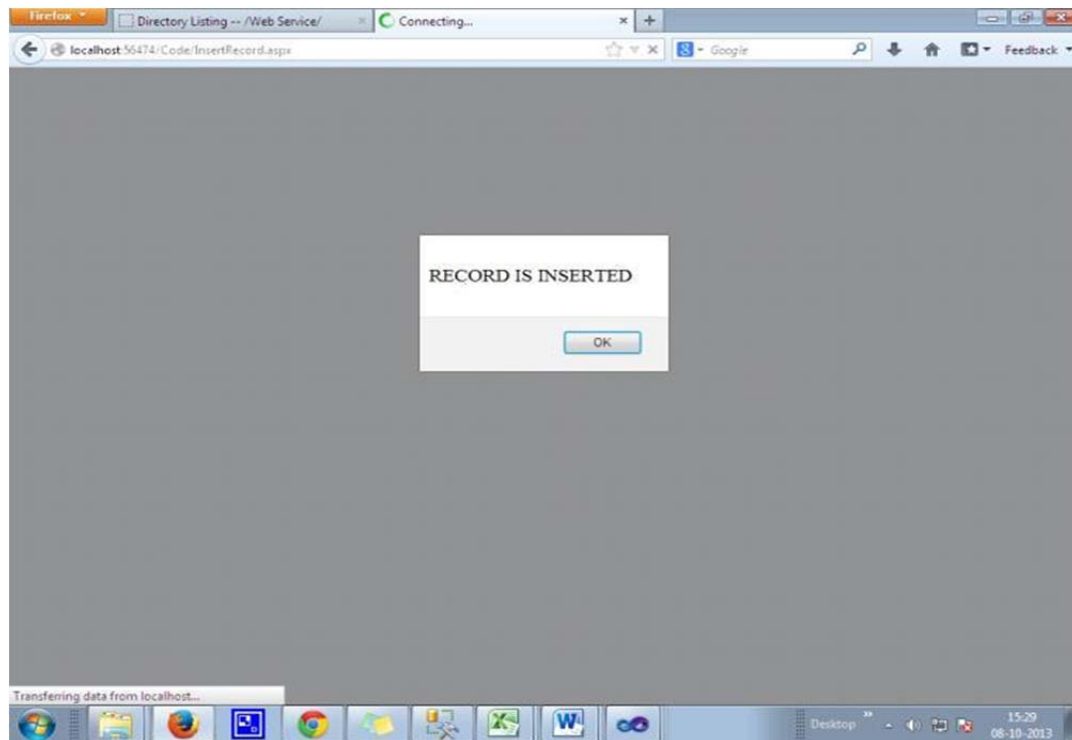


Fig. 7. Record inserted to the desired cloud

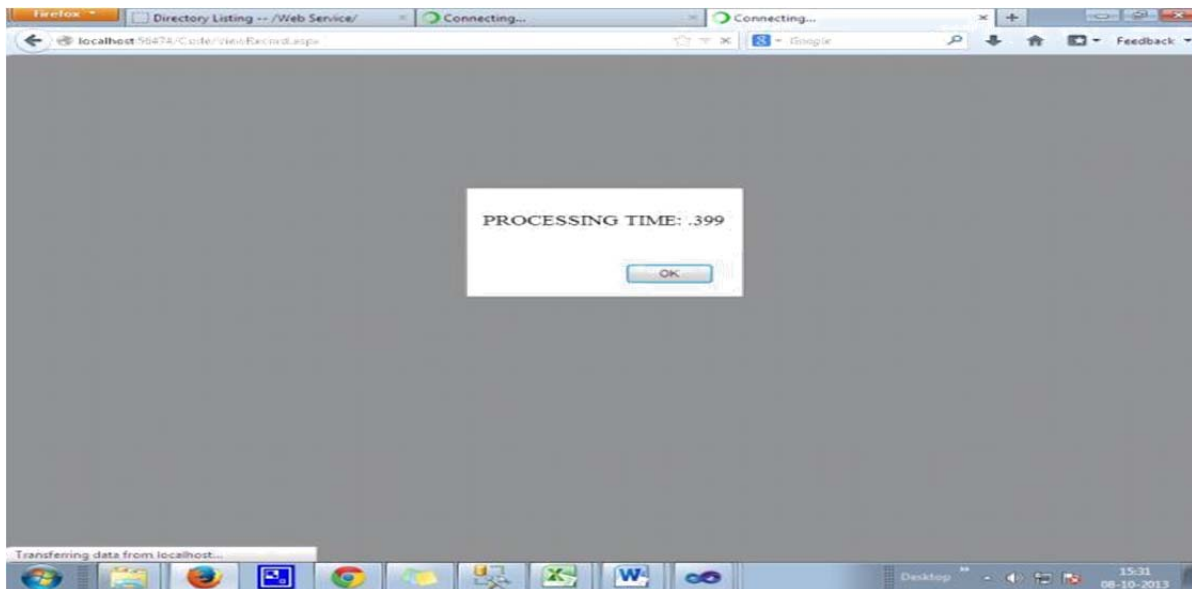


Fig. 8. Processing time necessity for insertion of a record

VI. CONCLUSION

This project, in both academia and practical implementation can act as footing through an application. Future work will need enormous data and software on cloud environment. For a central server, allocating resource to its user or clients is a thing to be consider, resource providers need to think about. There is belief that using the algorithm on an extensive, it makes effective usage of allocating resource to an extent. The algorithm used that is on priority based and its classification and has unique parameters to allocating resources, denoting deadline or cost. We can even use actual physical resources for allocation. Moreover, we believe that the OS virtualization setup can be extended to different computers and make a distributed cloud environment with some modifications if required. Also, the performance checking criteria may be added to already set up to demonstrate an algorithm which can easily outperform the FCFS algorithm in all respects. We have presented, the execution, and evaluating of resource management system for services in cloud computing. The system which uses technology of virtualization for the allocation purpose in resources denoting data center that changes on basis of application demands. We have proposed a new way which may be included in the Cloud-Analyst, to have cost effective results and development. From the working made, its results as process of the simulation can be upgrade by changing or adding new scheme or plan for balancing the load and for traffic routing, etc, so that researchers and developers may do prediction of real implementation of cloud, very easily, developing heuristics which avert overload in the network of system effectively plus saves energy. Thus also results of experiment shows that the algorithm gains performance which can be enhanced and is good.

VII. REFERENCES

- [1] T.R. Gopalakrishnan Nair, Vaidehi M, "Efficient resource arbitration and strategies of allocating in cloud computing through virtualization" Proceedings of IEEE, CCIS(2011).
- [2] R.Prodan, S.Ostermann.A taxonomy and also survey of a infrastructure as service and cloud providers in Web hosting In IEEE/ -- ACM Conference on Grid computing 2009 pages 17-25, October 2009.
- [3] Grosu, D, Das. A., "Auction-Based Resource Allocation Protocols In Grids", 16th Iasted International Conference On Distributed Computing And Systems, 2004.
- [4] A. Ghodsi, M. Zaharia, S. Shenker, I. Stoica, "Choosy: Max-min fair sharing for datacenter jobs with constraints," in Proc. ACM EuroSys, 2013: 365-378.
- [5] C. Joe-Wong, S. Sen, T. Lan, M. Chiang, "Multi-resource allocation: Fairness-efficiency tradeoffs in a unifying framework," IEEE/ACM Transactions on Networking (TON), 2013, 21(6): 1785-1798.
- [6] D. Dolev, D. Feitelson, J. Halpern, R. Kupferman, N. Linial, "No justified complaints: On fair sharing of multiple resources," in Proc.ACM ITCS, 2012: 68-75.
- [7] A. Gutman, N. Nisan, "Fair allocation without trade," in Proc. AAMAS, 2012: 719-728
- [8] D. Parkes, A. Procaccia, N. Shah, "Beyond dominant resource fairness: Extensions, limitations, and indivisibilities," in Proc. ACM EC, 2012: 808-825
- [9] C. Joe-Wong, S. Sen, T. Lan, M. Chiang, "Multi-resource allocation: Fairness-efficiency tradeoffs in a unifying framework," IEEE/ACM Transactions on Networking (TON), 2013, 21(6): 1785-1798.
- [10] Xiaohong Wu At La. Cloud Computing Resource allocation Mechanism Research Based On Reverseauction. Energy Procedia, Vol 13, 2011.
- [11] M. Kowsigan, P. Balasubramanie, "An Improved Job Scheduling in Cloud Environment using Auto-Associative-Memory Networks" Asian journal of Research in Social Science and Humanities, Vol. 6, No. 12, Dec 2016, pp.390-410.
- [12] M. Kowsigan, P. Balasubramanie, "Scheduling of Jobs in Cloud Environment using Soft Computing Techniques", International Journal of Applied Engineering Research, Vol. 10, No. 38, 2015.
- [13] M. Kowsigan, S. Rajkumar, P. Seenivasan, C. Vikram Kumar, "An Enhanced Job Scheduling in Cloud Environment using Improved Metaheuristic Approach", International Journal of Engineering Research and Technology, Vol. 6, No.2, pp.184-188

- [14] M. Kowsigan, M. Rubasri, R. Sujithra, H. Sumaiya Banu, "Data Security and Data Dissemination of Distributed Data in Wireless Sensor Networks", International Journal of Engineering Research and Applications, Vol. 7, No.3(part 4), 26-31.
- [15] Jamer Basha A, Kanmani R, "Performance Analysis of Wireless OCDMA Systems using PC OCC, EPC codes", Asian Journal of Informational Technology, Vol. 15, No. 12, pp.2087-2093, 2016

Author Profile

M. Kowsigan is working as an Assistant Professor in the Department of Information Technology in Sri Krishna College of Technology, Coimbatore, India. His Area of Interest includes Grid Computing, Cloud Computing, Soft Computing. He is Pursuing his Part-time Ph.D in Anna University, Chennai.

S. Kalicharan is pursuing his UG degree in the Department of Information Technology in Sri Krishna College of Technology, Coimbatore, India. His Area of Interest includes Data mining, Cloud Computing, Soft Computing.

P. Karthik is pursuing his UG degree in the Department of Information Technology in Sri Krishna College of Technology, Coimbatore, India. His Area of Interest includes Grid Computing, Cloud Computing, Soft Computing.

A. Manikandan is pursuing his UG degree in the Department of Information Technology in Sri Krishna College of Technology, Coimbatore, India. His Area of Interest includes Grid Computing, Cloud Computing, Soft Computing.

R. M. Manikandan is pursuing his UG degree in the Department of Information Technology in Sri Krishna College of Technology, Coimbatore, India. His Area of Interest includes Grid Computing, Cloud Computing, Soft Computing.