Dynamic Resource Allocation in Cloud Computing Environment using Optimization Algorithm

[1]Juhi Gupta, [2]Aaditya Tulsyan, [3]Bhavana V.C, [4]Bidhan Chandra Roy, [5]Mrs. Radhika T.V

[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor
Information Science and Engineering,
Dayananda Sagar College of Engineering, Bangalore, India

***Abstract*: Utilizing dynamic resource allocation for load balancing is considered as an important optimization process in cloud computing. In order to achieve maximum resource efficiency and scalability in a speedy manner the process is concerned with multiple objectives for an effective distribution of loads among virtual machines. In the realm, exploring new algorithms, as well as development of novel algorithms, is highly desired for technological advancement and continued progress in resource allocation application in cloud computing. The project explored the concept of "skewness" which measures the unevenness in the multidimensional resource utilization of a server. On minimizing skewness, different types of workloads can be combined nicely and improve the overall utilization of server resources. A set of heuristics has been developed that prevent overload in the system effectively while saving energy used. Trace driven simulation and experiment results demonstrate that our algorithm achieves good performance.**

***Keywords*: Cloud Computing, Resource Allocation, Virtual Machine, Scheduling Technique**

INTRODUCTION

Cloud computing is known to be the delivery of computing and storage capacity as a service to the community of end recipients. The use of a cloud shaped symbol as an abstraction for the complex infrastructure gave the name and it contains the system diagrams. Cloud computing provide services with user's data, software and computation over a network. The access to cloud services from anywhere at any time was enabled by the remote accessibility. The services offered in terms of resources should be allocated optimally to the applications running in the cloud to gain maximum degree of above mentioned benefits. Proponent claimed to allows enterprises to get their applications up and running faster, with improved manageability and less maintenance. It enables IT teams to rapidly adjust resources and meet fluctuating and unpredictable demand. "pay-as-you-go" model that could be a cloud valuation model as usually utilized by cloud supplier. The availability of high capability networks, low value computers and storage additionally because the widespread adoption of hardware virtualization, service-homeward design, and automatic and utility computing hashad led to growth in cloud computing.

1.1 Overview:

Resource allocation within the cloud atmosphere is employed to attain client satisfaction with negligible time interval. Reducing the fees of leasing resources additionally to making sure quality of service and up outturn for trust and satisfaction of the service supplier is taken into account as another objective. In dynamic planning, the essential plan is that the request allocation at the time of implementation of programs.

Virtual machine monitors (VMMs) like Xen offer a mechanism for mapping virtual machines (VMs) to physical resources. The mapping is largely hidden from the cloud users. Users with the Amazon EC2 service for instance, don't understand wherever their VM instances run. It is up to the cloud provider to make sure the underlying physical machines (PMs) have sufficient resources to meet their needs. It has been made possible by VM like migration technology to change the mapping between VMs and PMs while applications are running. However, a policy issue remains of deciding the mapping adaptively so that the resource demands of VMs are met while the number of PMs used is minimized. It has been a challenge for resource needs of VMs are heterogeneous due to the diverse set of applications they run and vary with time as the workloads grow and shrink. The capacity of PMs can also be heterogenous because multiple generations of hardware co-exist in a data centre.

1.2 Problem Statement:

A cloud is made from various physical machines that runs multiple virtual machines that are conferred to purchasers, because the computing resources. Yet, high employment on virtual machines is one in all the challenges of cloud computing within the allocation of virtual machines. The task, requested by a client, has to wait to be allocated to the work and the resources needed and are interested in having their tasks completed in the shortest possible time and at the minimum cost. On the opposite hand, the cloud providers are interested to maximize the use of their resources and also to increase their profits. Obviously these 2 objectives are in conflict with one another and infrequently they are not happy with the standard strategies of resource allocation and programming mechanisms offered.

1.3 Objectives:

1) Dynamically allocate resources by skewness algorithm to utilize the resources properly.

2) Managing the workload among resources so that to obtain the minimal execution time. Efficient scheduling should be provided so that both the user and the service provider gain profit.

LITERATURE SURVEY

The table present the summary of related work done in the field of the related work done in the field of tasks scheduling. Table includes objective, algorithm used, and their stimulation environment.

2.1 Cloud Computing Survey

There are authors who discussed on cloud computing issues. There are many issues like network, performance, security and many others. Those authors were Haibo Yang et al. (2009) , Haibo Yang et al. (2012), Sunilkumar S.M et al.(2013), F.F.Moghaddam et al. (2015), Mahmoud.M et al. (2014), Nasrin Hesabian et al. (2015), Anterpreet Kaur (2015), K.BeghdadBey et al., MalaKalraetal(2015) and Vijayalakshmi A et al.(2013). The most important issue has been related to security and privacy subjects in cloud computing environments.

Yaserghanam et al.(2012), FredericoDurao et al. (2014), Anton Backe et al. (2015), were authors who focused on advances issues in field of cloud computing. These issues include security and privacy, resource sharing, billing, and services layer agreement to open new questions about the real gains of the business model.

Authors who presented strategies for efficient scheduling and resource allocation to decrease cost and time were Jaspreetkaur et al. (2012), Lovejit Singh et al. (2013), Sukhpal Singh et al. (2015), SoumenSantra et al. (2015). The strategies solved the load distributing problem on various nodes of a distributed system and has improved resource utilization and iob response time. Issue related to disk space management and amount of power consumption in data centers for great amount of carbon dioxide emissions were discussed byShabnam Khan et al. (2013) and A E Mohamed Mosa et al. (2014).

M.Padmavathi et al. (2014) and Malathi.P et al. (2015) suggested achieve good performance, high utilization of resources, less response time, Cloud computing to execute tasks by providing related resources and also for any type of complicated tasks in different environments of cloud.

Three types of studies about cloud computing in eHealth, cloud-based eHealth framework design, security control mechanisms of healthcare data and applications of cloud computing was discussed by Yan Hu et al. (2014) and Zhi-Hui Zhan et al.(2015) discussed on Li Liu et al. (2014) and RashmiRai et al. (2015). They focused on the various critical issues like maintainability, scalability, security, scale and complexity associated with the legacy system. Security has been considered as prime factor for the adoption in cloud migration. IlangoSriram et al. (2015)and Rahul Sharma et al. (2015) has raised different research challenges such as facilitating industry in building successful clouds and comparing cloud computing features with pervasive as well as service computing.

2.2 Resource Allocation Survey

Knowledge on utlization of resources in cloud computing environments such as cloud providers and cloud users with a minimal expense for which it meets performance requirements was focused by Shubhangi D. et al.(2012), Abirami S.P et al.(2012), Ram Kumar Sharma et al.(2013), XinXu et al. (2014), Zhigang Zhou et al. (2015),Hussein S. Al-Olimat et al.(2015).The knowledge is applied on different services like nimbus, cumulus and demand.

N. Asha et al. (2013) and Jayanthi S et al. (2014) were the authors who suggested resource accounting with two alternatives i.e. One strictly usage-oriented where user has been limited to number of units, which can be connected to CPU and/or Memory usage, time. Another alternative was capacity preallocation.

Presentation of energy consumption offers in various computing utility services to daily life with the demand of Cloud infrastructure which had greatly increased the energy consumption of data centers wad done by A.Y.Khobragade et al. (2013) WannengShu et al. (2014) and Jing-Mian Tang et al.(2015). It was hard to evaluate and optimize energy consumption in complex cloud infrastructure architectures.

Challenges like security of data, server consolidation, resources, tasks, data storage and consumption of energy was discussed by Devendra Singh Thakur et al. and ShilpiSaxena et al.

Two authors were there named Qi Zhang et al. (2014) and Zhou Fang et al. who said that existing solutions was not considered by the heterogeneity of both workload and machine hardware found in production environments fully. There was failure to consider the heterogeneity of machines and workloads that leads to sub-optimal energy-savings and long scheduling delays, because of incompatibility between workload requirements and the resources offered by the provisioned machines.

Rapid development of cloud computing environment to map applications from large quantity of users was focused by Sagar.G et al. (2013) and S.Thejeswi et al. (2013). There was International Journal of Advanced Information Technology (IJAIT) Vol. 6, No.4, August 2016 9 which provisioned that all resources were made dynamically available to satisfy the needs of requesting users. S.Sujan et al. (2015) and H.Wang et al.(2015) were the authors who fixed the issue of computing millions of transactions in a day which considered resource provisioning and scheduling context.

Chien-Hung Chen et al. (2015) , D. A. Heger et al., Gunho Lee et al. were the authors who gave an idea on advanced methodologies to be used in cloud computing likeMapReduce scheduling and Data Analytics which were required to assess the performance, capacity behavior and allocation with high fidelity.

Different types of resources in virtualized form for dynamix utilization using scheduling was represented by Pooja .P.Vasani et al. (2013) and Longbin .C et al. (2015) .

Challenging issue called QoS (Quality of Service) for hybrid storage systems in resource allocation and also QoS degradation is expressed through time deviation metric was taken up by N.D. Doulamis et al. (2014), Christina Delimitrou et al. (2015) and Hui Wang et al.(2015).
Shivani Sharma et al. (2014) and A. Narayan Singh et al. (2015) sharing on various software and hardwa re resources for designing various resource allocation alogrithms in efficient manner.
Sharrukh. Z et al. (2013), Pratik P. Pandya et al. (2014) and GruiaCalinescu et al. (2011) were the authors who addressed the issue of dynamic provisioning VM instances in clouds which generate higher profit, cost, time consumption, carbon effect.
Discussion on wireless cloud computing service was done by ShaoleiRen et al. (2014) and Seokho Son et al. in that discussion they talked about service providers that keep deploying services in geographically distributed data centres and provides services based on dynamic price subscribe, Service Level Agreements to reduce the load on data centers.
Data Center Resource Allocation paradigms for internal and external challenges and computing nodes in social ties was discussed byM.A.Sharkh et al. (2013) and Aziz Mohaisen et al. (2014) introduction of paradigms work based on bootstrapping trust-possessing social graph was also done by them. These paradigms are different compared with the paradigms of distributed, grid and conventional cloud computing.

2.3 Statement of the Problem:

A cloud is nothing but combination of several physical computers and they run on multiple physical machines and then they are served to client for their use. As now a days with growth of digital world number of clients has been increased to a great number and everyone wants to use virtual machine in most efficient way. It has increased the workload on virtual machines and has become a challenge of cloud computing for allocation of resource.
A client has to wait for their requested task to be allocated and work with the resource needed. Each client wants to priorities their need of resource and they offer larger value of resource. Clients are interested in getting their task to be completed in shortest possible with lowest possible cost. On other side cloud service provider aims to make profit by getting a greater number of clients and maximizing the utilization of resource. These two objectives conflict each other and traditional method of resource allocation and scheduling mechanisms are unable to resolve these issues. So, the goal here is satisfy both side with their objective to be fulfilled and perform resource allocation in a better way.

METHODOLOGY

5.1 Introduction

Implementation is the realization of an application, or execution of a plan, idea, model, design, specification, standard, algorithm, or policy. In other words, an implementation is a realization of a technical specification or algorithm as a program, software component, or other computer system through programming and deployment. Many implementations may exist for a given specification or standard.

5.2 Overview of System Implementation

The project is implemented considering the following aspects:

1. Usability Aspect.
2. Technical Aspect.

5.2.1 Usability Aspect

The usability aspect of implementation of the project is realized using following principle:

5.2.1.1 IMPLEMENTATION USING JAVA, HTML AND JSP

Java is a programming language which was first developed by Sun Microsystems back in 1995. It can be found on different types of devices from smartphones, to mainframe computers. It can also be useed on desktop PC and even on the Raspberry Pi. Java relies on a “virtual machine” which understands an intermediate format called Java bytecode and doesnot compile to native processor code. Each platform that runs Java needs a virtual machine (VM) implementation.

HMTL is a standard mark-up language that is used for creating webpage and web application. With CSS and JavaScript, it creates a full working front end of any website. It was developed in 1980 by Tim Berner Lee at cern. Web browser use html to interpret and compose text, image, and other material into visual and audio web page.

Jsp is a technology that helps software developer to create a web page that can be dynamically generated based on HTML, XML and another document type. It is used for the development of backend part of webpage. It can be viewed as high level abstraction of java servlet as they are translated into servlets at runtime

5.2.2 Technical Aspect

The technical aspect of implementation of the project is realized using following principle:

5.2.2.1 THE PROJECT IS IMPLEMENTED BY USING APACHE TOMCAT SERVER AND NAVICAT

Apache Tomcat Server:

Tomcat server implements jsp and provide a pure java HTTP web server environment in which java code can run. Tomcat5.0 can support servlet 2.4 and jsp 2.0. With tomcat server project can run in more stable manner and it also offers extra level of security. As it is an open source, source code for server is readily available. Tomcat is connected to MySQL through JDBC connection. JDBC provides an abstraction layer between MySQL and tomcat, because of which project code does not need to be altered in order to communicate with multiple database.

Navicat:

Navicat is a series of graphical database management and development software for MySQL and other database applications. It supports multiple database connection for local and remote database. It is compatible with cloud database like google cloud, Microsoft azure, amazon aurora. It provides powerful graphical interface for database management, development and maintenance. There is easy sql editing and data transfer, data synchronization helps to migrate data easily.

RESULTS

The end result can be shown through sample screen

7.1 Index Page

This is the front-end view of the web page where different menu options are provided

Fig 7.1 Index Page

7.2 Cloud user Service

This include all the services provided for cloud user i.e. client who request for the resource. They include services as creating site, domain registration, login to their account

7.2.1 login Page

This is where cloud user will enter their login detail and use the service. If the user don't have account then they can register to create their account.

Fig 7.2 Cloud user Login page

7.2.2 Domain Registration page

When a cloud user wants to host their web page they register their domain to request for the resource.

Fig 7.3 Domain Registration Page

7.2.3 Creation of site

This is the place where user give the name for their site and create the site.

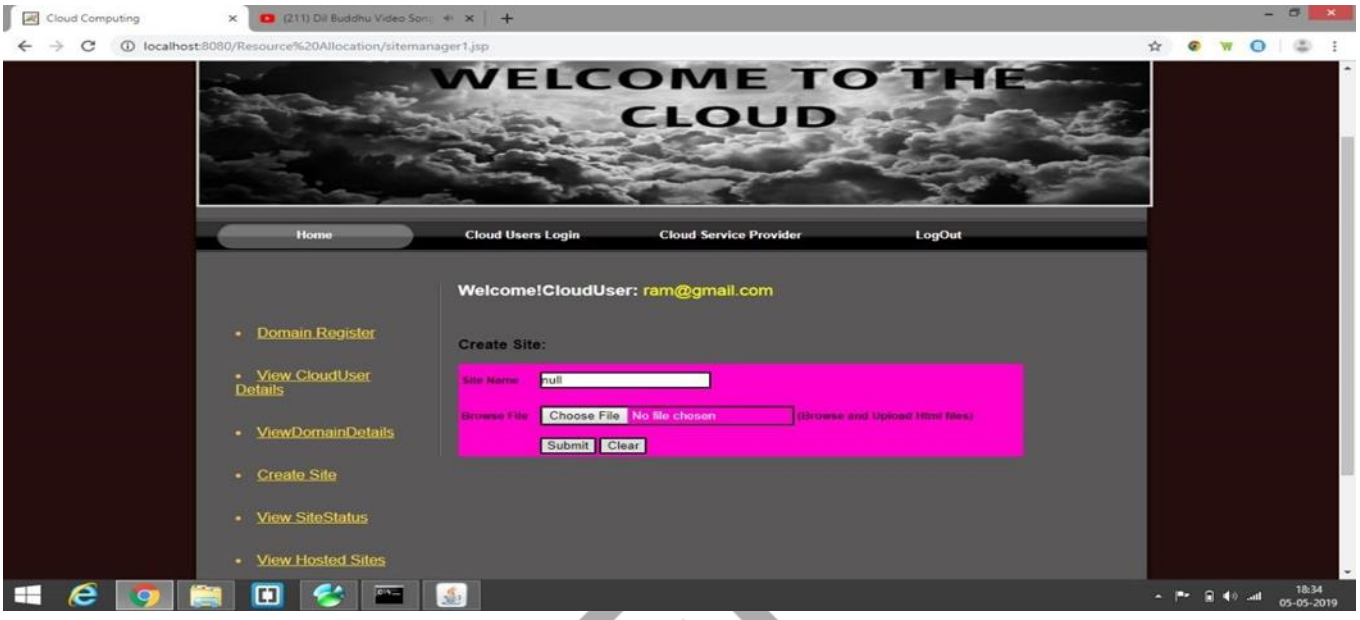

Fig 7.4. Creation of site by cloud user

7.2.4 *View Cloud user details*

User can also see their details in the given option of view cloud user details.

Fig 7.5 View of cloud user details

7.2.5 View Site Status

They can also see the status of their site. If it has been allocated by provider then it will show the status as allocated.

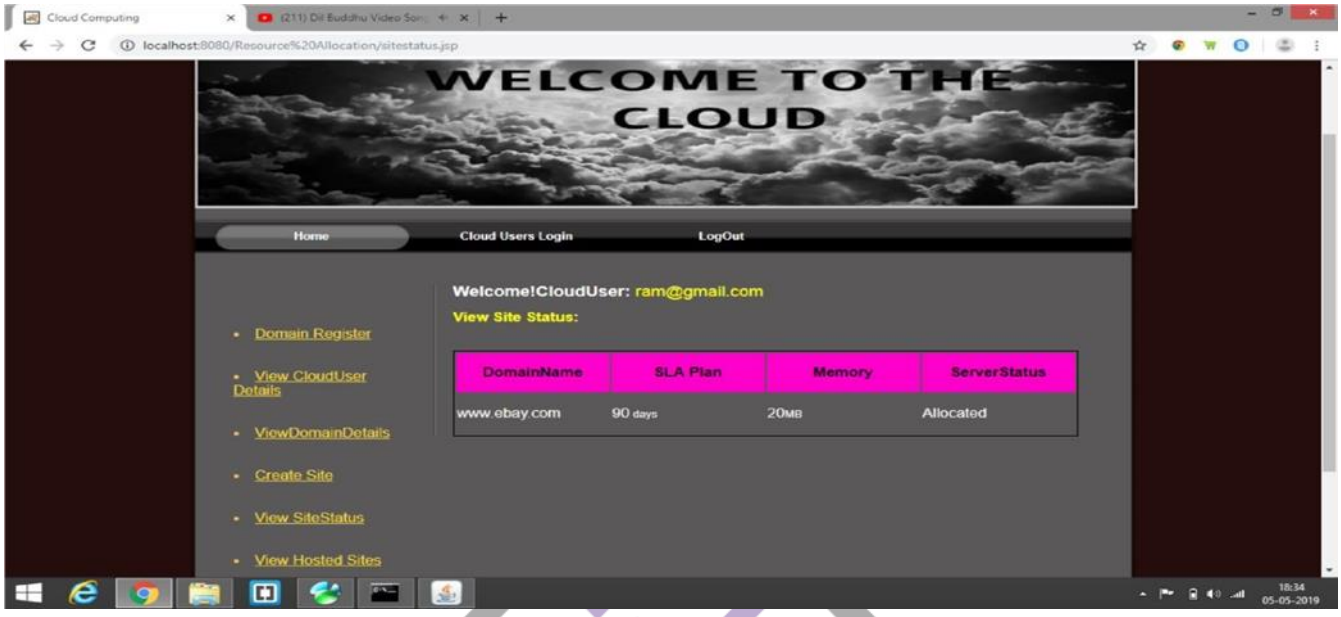

Fig 7.6 View site status by cloud user

7.3 Cloud Service Provider

This include all the services provided by the cloud service provider. They respond on the request of client. They receive the request and allow the allocation of resource.

7.3.1 Cloud Service Provider Page

This is the front view of cloud service provider page and it shows the option of signup form where admin can login.

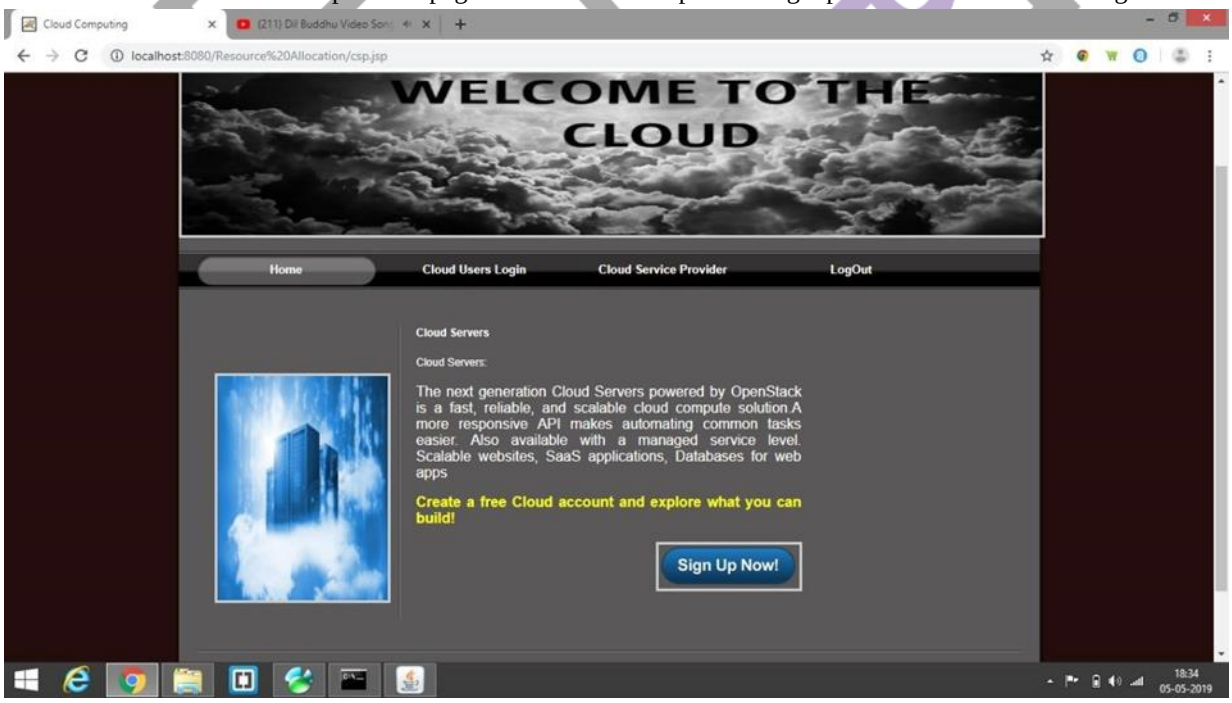

Fig 7.7 Cloud Service provider page

7.3.2 Cloud Service provider Login

To login for the cloud provider they should visit the login page under cloud service provider option.

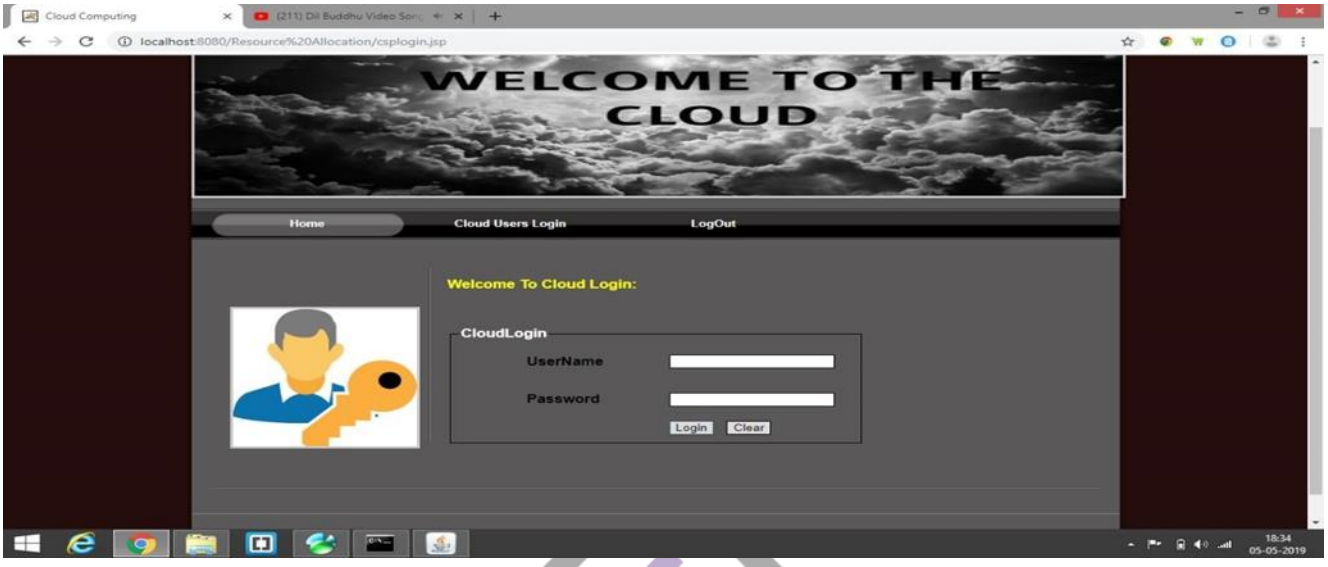

Fig 7.8 Cloud service provider login

7.3.3 Cloud Service provider Registration

If anyone wants to register as a cloud service provider they need to register and the admin have to allow them the access.

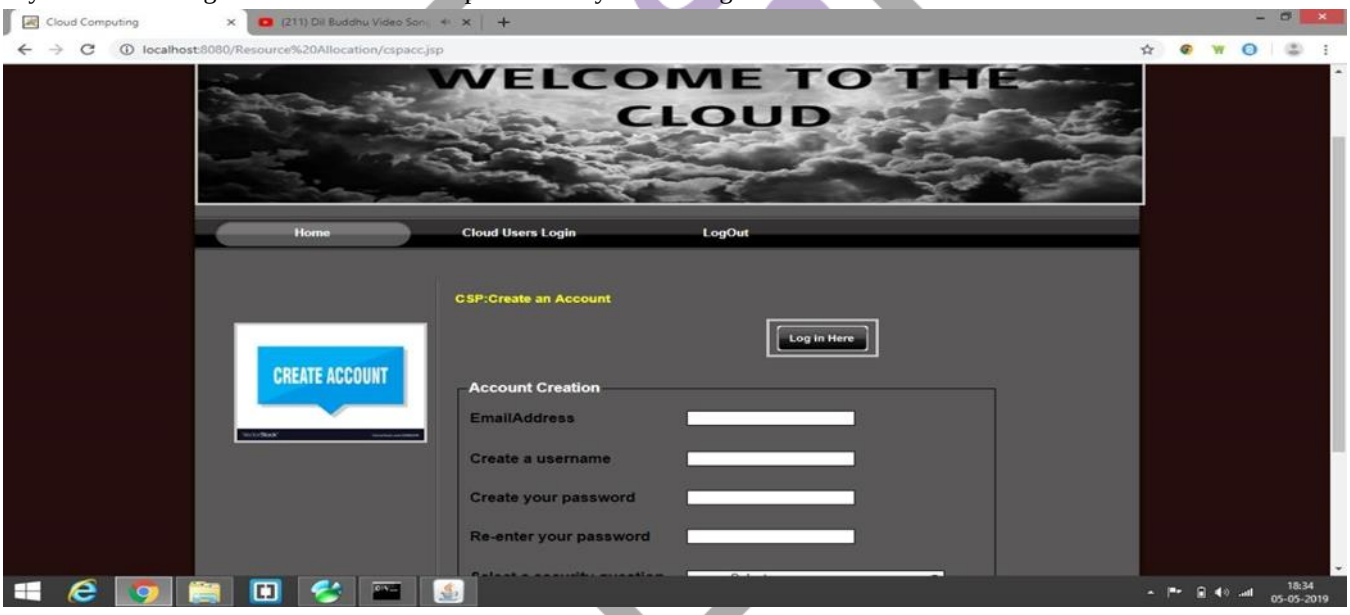

Fig 7.9 Cloud Service provider Registration Page

7.3.4 Server Status

To see the status of the server and how much memory are in use and how much are free.

Fig 7.10 View server (Resource allocation process)

7.3.5 Resource Allocation Status

Here all the site requested and approved by the cloud service provider can be seen and also to which server the resource has been allocated.

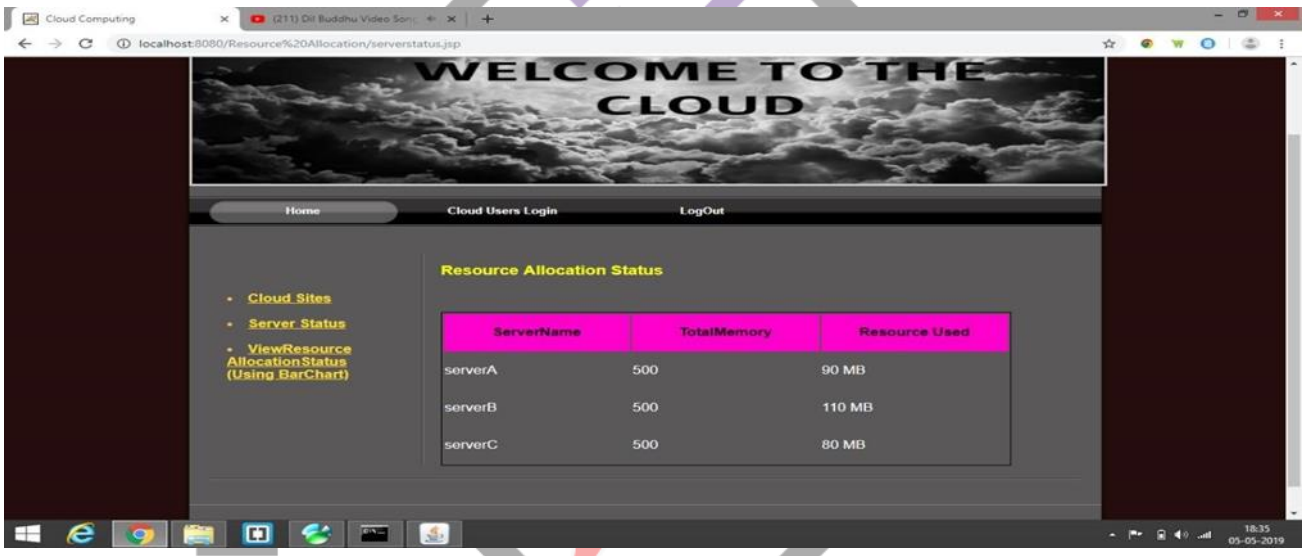

Fig 7.11 Resource allocation status

7.4 Resource Allocation Database

This shows the histogram view of all the three server's memory

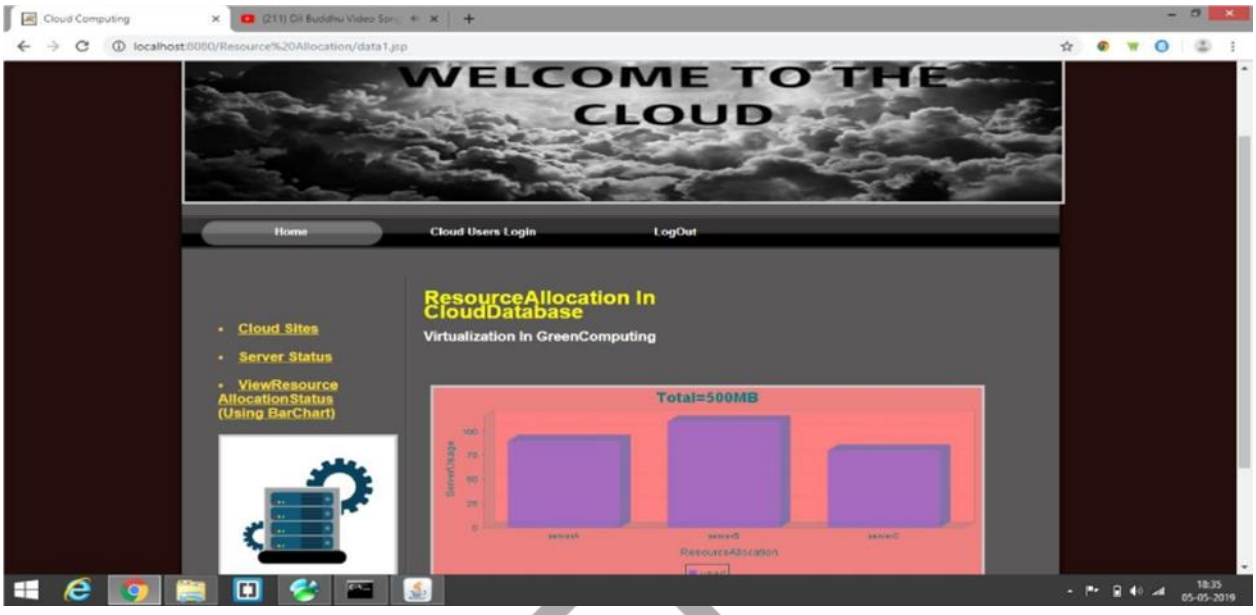

Fig 7.12 Resource allocation database

I. ACKNOWLEDGMENT

This work has been done under the supervision of HOD of information science department at dayananda sagar college of engineering and guided by Mrs. Radhika T.V, assistant professor in information science department at dayananda sagar college of engineering.

REFERENCES

[1] Kalaashri. Y.RA b, Dr. S. padmavathi c ,Title: dynamic resource allocation scheme in cloud computing Saraswathi AT a, The authors published by elsivier B.V year: 2015

[2] A. Mosavi and A. Vaezipour: Reactive Search Optimization; Application to Multiobjective Optimization Problems, Applied Mathematics, Vol. 3, 2012, pp. 15721582

[3] Mashayekhy, An Online Mechanism for Resource Allocation and Pricing in Clouds, IEEE Transactions on Computers, 12 June 2015.

[4] J. Yao and H. Ju-Hou: Load Balancing Strategy of Cloud Computing Based on Artificial Bee Algorithm, Information Management, Vol. 51, 2012, pp. 185-189

[5] S. Zhao, X. Lu, and X. Li: Quality of Service-based Particle Swarm Optimization Scheduling in cloud Computing, Networks, Vol. 12, 2015, pp. 235-242

[6] S. Zhang and Y. Zhou: Grey Wolf Optimizer Based on Powell Local Optimization Method for Clustering Analysis, Discrete Dynamics in Nature and Society, Vol. 3, 2015

[7] U. Ayesta, M. Erausquin, E. Ferreira, and P. Jacko: Optimal Dynamic Resource Allocation to Prevent Defaults, Operations Research, Vol. 6, 2016, pp. 451-456

[8] A. Abur, and A. G. Exposito: Power System State Estimation: Theory and Implementation. CRC press, 2004

[9] A. Khetan, V. Bhushan, and S. Gupta: A Novel Survey on Load Balancing in Cloud Computing, Journal of Engineering & Technology, Vol. 6, 2013, pp.1-9

[10] B. Gomathi and K. Karthikeyan: Task Scheduling Algorithm Based on Hybrid Particle Swarm Optimization in Cloud Computing Environment, Journal of Theoretical and Applied Information Technology, Vol. 89, 2013, pp. 33-38

[11] M. Armbrust et al., "Above the Clouds: A Berkeley View of Cloud Computing," technical report, Univ. o California, Berkeley, Feb. 2009.

[12] L. Siegele, "Let It Rise: A Special Report on Corporate IT," The Economist, vol. 389, pp. 3-16, Oct. 2008

[13] C. Clark, K. Fraser, S. Hand, J.G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, "Live Migration of Virtual Machines," Proc. Symp. Networked Systems Design and implementation (NSDI '05), May 2005..

[14] M. Nelson, B.-H. Lim, and G. Hutchins, "Fast Transparent Migration for Virtual Machines," Proc. USENIX Ann. Technical Conf., 2005.M. Young, *TheTechincal Writers Handbook.* Mill Valley, CA: University Science, 1989.

[15] N. Bobroff, A. Kochut, and K. Beaty, "Dynamic Placement of Virtual Machines for Managing SLA Violations," Proc. IFIP/IEEE Int'l Symp.Integrated Network management (IM '07), 2007.