International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || PP.67-74

Job Scheduling in the Cluster Based Cloud by Employing Decentralized Approach

Satish Kumar
(Assistant Professor, CSE, CT Institute of Technology & Research, Maqsudan, Jalandhar, India)

Abstract: This paper projected a frame work on which a workload is dispersed among virtual machines available at a variety of User Bases located at different Data Centers. The workload is detached depending on dissimilar strategies such as Round Robin, Equally Spread Execution Load and Throttled which can takes place depending on solitary policy at a time. The sole Data Center and User Base can fit into a single section only. After placing the Data Center and User Base to an exacting area, its configuration is Fig.d out. User grouping in User Bases and Request grouping in Data centers are the two key features added in this work. Concurrent users from a single user base outline a user grouping and simultaneous requirements that a single server can hold up at a scrupulous time forms a demand grouping in data centers. Response time and processing time is calculated in terms of MAX, MIN and AVG and degree of hardware, its cost is calculated in term of Virtual Machine's cost and data transfer cost. At last when simulation is ended, overall response is produced and documentation can be attained.

Keywords: Cloud Computing, Resource scheduling, Load balancing, Data centre Virtual Machine, User bases, Simulation, Service level agreement (SLA).

I. INTRODUCTION

In Cloud Computing, the word cloud is used as a symbol for "the Internet", so the expression cloud computing means "a sort of Internet-based computing", where dissimilar services such as servers, storage and applications are distributed to an organization computers and devices through the internet. One of the key distinctiveness of cloud computing is the flexibility that it tender and one of the ways that suppleness is offered is through scalability. This refers to the skill of a system to become accustomed and scale to alter in workload. The cloud formulate it possible for a user to access information from anyplace at any time. While a conventional computer arrangement necessitate user to be in the similar location as your data storage device, the cloud takes away that step. The cloud gets rid of the need for user to be in the identical physical location as the hardware that stores your data. Cloud provider can both own and need to buy the quantity of storage space user will employ, a business can acquire more space or lessen their subscription as their business grows or as they discover they need less storage space. This is particularly helpful for businesses that cannot manage to pay for the same amount of hardware and storage space as a superior company. Undersized companies can store their information in the cloud, eliminating the cost of purchasing and storing memory devices.

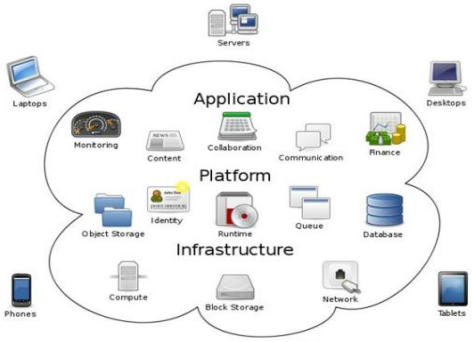

Fig. 1 - Cloud Computing [5]

II. LOAD BALANCING SCHEMES

Load balancing algorithms can be categorized into static and dynamic approaches:

2.1 Static load balancing algorithm

Static load balancing algorithms presume that prior information about the entire the characteristics of the jobs, the computing resources and the communication network are recognized and provided. Load balancing decisions are made deterministically or probabilistically at compile time and remain constant throughout

runtime. The static approach is striking because it is straightforward and requires minimized runtime overhead. However, it has two major disadvantages. Firstly, the workload allocation of many applications cannot be envisaged before program execution. Secondly, it assumes that the computing resources and communication network are all known in advance and remain steady. Such an assumption may not apply to a distributed environment. As static approach cannot respond to the dynamic runtime environment, it may lead to load inequity on some resources and significantly boost the job response time.

2.2 Dynamic load balancing algorithm

Dynamic load balancing algorithms attempt to utilize the runtime state information to make more revealing decision in sharing the system load. However, dynamic scheme is used a lot in modern load balancing method due to their robustness and flexibility.

III. LOAD BALANCING POLICIES

Algorithms for the load balancing problem can be broadly categorized in terms of four policies. They are:

3.1Location policy

It is the policy that influences the finding of an appropriate node for migration. The general technique followed here is polling, on a broadcast, random, nearest-neighbor or roster basis.

3.2 Transfer policy

It is that which determine whether a node is appropriate for participating in process relocation. One common technique followed is the threshold policy, where a node participates in a negotiation only when its load is fewer than (in destination-initiated algorithm) or greater than (in sender-initiated algorithm) a threshold value.

3.2.1Selection policy

It is the policy that deals with the selection of the process to be migrated. The common factors which must be considered are the cost of migration (communication time, memory, computational requirement of the process, etc.) and the expected gain of migration (overall speedup of the system, etc.).

3.2.2 Information policy

It is that component of the algorithms that decide what, how and when the information regarding the state of the other nodes in the system in gathered and manage. They can be grouped under demand-driven, periodic, or state-change-driven policies.

IV. LOAD BALANCING ALGORITHMS

4.1 Round Robin

It is a kind of Static and Decentralized algorithm in nature; the processes are separated between all processors. Each process is allocated to the processor in a round robin order by means of a particular time value. The process allocation arrangement is potted locally independent of the allocations from remote processors. Though the job load distributions among processors are alike but the job processing times for dissimilar processes are not similar. So at any point of time various nodes may be heavily loaded and others remain unused. This algorithm is habitually used in web servers. This algorithm in addition is implemented as Weighed Round Robin Algorithm.

4.2 Throttled Load Balancer

This algorithm is a dynamic load balancing algorithm. It is used for load balancing in the case of the virtual machines to be used. In this, firstly, the index values of all the virtual machine in the system are checked. The demand is sent where load balancer parses a table for the allocation of the resources in the system. It assigns the request to a exacting load balancer which passes or responds overturn the request to the requester and bring up to date allocation policy [9]. After the successful allocation of the system the entire process for the de-allocation of the system starts too. This mechanism offers a greater/higher amount of resource sharing and allocation on an entire system resulting in the higher performance and utilization. The throttling threshold maintained usually is 1. It could be modified effortlessly to make the threshold a configurable value.

4.3 Active Monitoring Load Balancer

This load balancing policy efforts to maintain identical workloads on every available VMs. The algorithm used is quite similar to the throttled case as clarified above but with faster and timely checking as well as the accessibility of the resources in the system. The ids for the allocation and de-allocation are specified. The value of the count modified with a new request. It gives the utmost utilization and performance of resources and machines respectively. The important things to be considered while developing such algorithm are: estimation and comparison of load, stability of different system, performance of system, interaction between the nodes, nature of work to be transferred, selecting of nodes and many other ones. [4]

V. PROBLEM FORMULATION

Cloud Computing is a narrative trend emerging in IT environment with massive requirements of infrastructure and resources. Load Balancing is a middle aspect of cloud computing environment. Competent load balancing design guarantees efficient resource consumption by provisioning of resources to cloud user's on-demand basis in pay-as-you-say-manner. Load Balancing may even continue prioritizing users by applying suitable scheduling criteria. Cloud load balancing is the process of distributing workloads crosswise numerous computing resources. Cloud load balancing lessen costs associated with document management systems and maximizes availability of resources. Load balancing sometimes leads to data overcrowding when it is not done correctly and it also limited the resources for other machines as well. Congestion may leads to loss of data packets. So, load balancing can be achieved on behalf of a range of methods according to the type of VM, location of VM, type of data etc. Many parameters should be measured before designing the actual cloud environment so that load balancing can be prepared appropriately and automatic. So, there is a need of system which can balance the load amongst various machines and condense congestion and response time by mounting an efficiency of a system. So the key target is to propose such a system on simulator which simulates the environment in which the entire globe can be connected together through various machines available at arbitrary location and the result formed from the experimental approach will produce enhanced system with better environment for cloud federation.

VI. METHODOLOGY PROPOSED

In the proposed methodology, Cloud Analyst & Cloud Sim both are used. Firstly, CloudSim is installed on LINUX OS. An .EXE is generated. To offer cloud computing environment by using several Data centres, User bases, VMs are placed on different regions. Load balancing can be implementing by employing a range of techniques and results will be produced on behalf of a variety of parameters like total load, response time, cost of machine etc. CloudSim is a toolkit (library) for simulation of Cloud computing situations. It provides basic classes for describing data centers, VMs, users, computational resources, Applications and policies for management of miscellaneous parts of the system (e.g., scheduling and provisioning).

6.1 Steps of Simulation

Multiple steps are there to simulate the methodology. The set-up is to be done in palpable manner to attain objectives of this paper. The setup holds User bases, Load balancing policies, Data centres, VMs, Memory, etc. All these are executed step after step and will exertion in a synchronized way to attain the desired results.

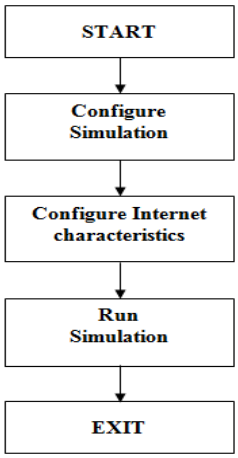

Fig. 2- Simulation Process

Step 1: Configuration Simulation

1.1 Main configuration

1.1.1 Simulation duration- Simulation period can be calculated in minutes, hours and days.

1.1.2 User Bases- This comprised of information like UserBase name, Region, Request per user per hour, Data Size per request (bytes), Peak hour start, Peak hour end, Average peak user, average off peak user.

1.1.3 Application deployment configuration

1.1.4 Service Broker Policy.

1.1.4.1 Closest data centre.

1.1.4.2 Optimize response time.

1.1.4.3 Reconfiguration dynamically with load balancing.

1.1.5 Data Center- This consist of information like data Center Identity, Number. of VMs belong to particular data center, Image size, Memory, Bandwidth.

1.2 Data Center Configuration- This step contain information that plays very important role , like: Name of Data center, Region of data center, Architecture, Operating System, VM Manager/ Monitor, Cost of VM's ($/hr), Memory cost ($/sec), Storage cost ($/sec), Data transfer cost ($/GB), Physical Hardware unit.

1.2.1 Physical Hardware particulars of Data Center- This will hold information like: Data Centre ID, Memory (in MB), Storage (in MB), Available Bandwidth (in KB), Number of Processors, Speed of processor (Clock cycles), VM policy.

1.3 Advanced- This option includes:

1.3.1 User grouping factor in User bases.

1.3.2 Request grouping factor in Data centers.

1.3.3 Executable instruction length per request.

1.3.4 Load balancing policy includes:

1.3.4.1 Round Robin

1.3.4.2 Equal Spread current execution load.

1.3.4.3 Throttled.

Step 2: Configuration Internet Characteristics involves:

2.1 Delay Matrix- Delay between regions in milliseconds.

2.2 Bandwidth matrix- Available bandwidth between regions in Mbps.

Step 3: Run simulation

This will generate the overall response time summary as output which includes multiple tables and graphs.

Step 4: Exit

VII. SIMULATION

Fig. 3 is associated to the region boundaries. In this Fig. various regions are shown which divide the entire globe. Here, six regions are labeled as R0, R1, R2, R3, R4 and R5. Also, three data centers are labeled as DC1, DC2 and DC3 and five User Bases are defined as UB1, UB2, UB3, UB4 and UB5. User bases UB1, UB2, UB3, UB4 and UB5 belongs to region R0, R1, R2, R4 and R5 respectively. The Data centre DC1, DC2 and DC3 belongs to regions R0, R2 and R4 respectively. Left of screen shows every step discussed in the methodology i.e. Configuration simulation, Define Internet characteristics, Run Simulation and Exit.

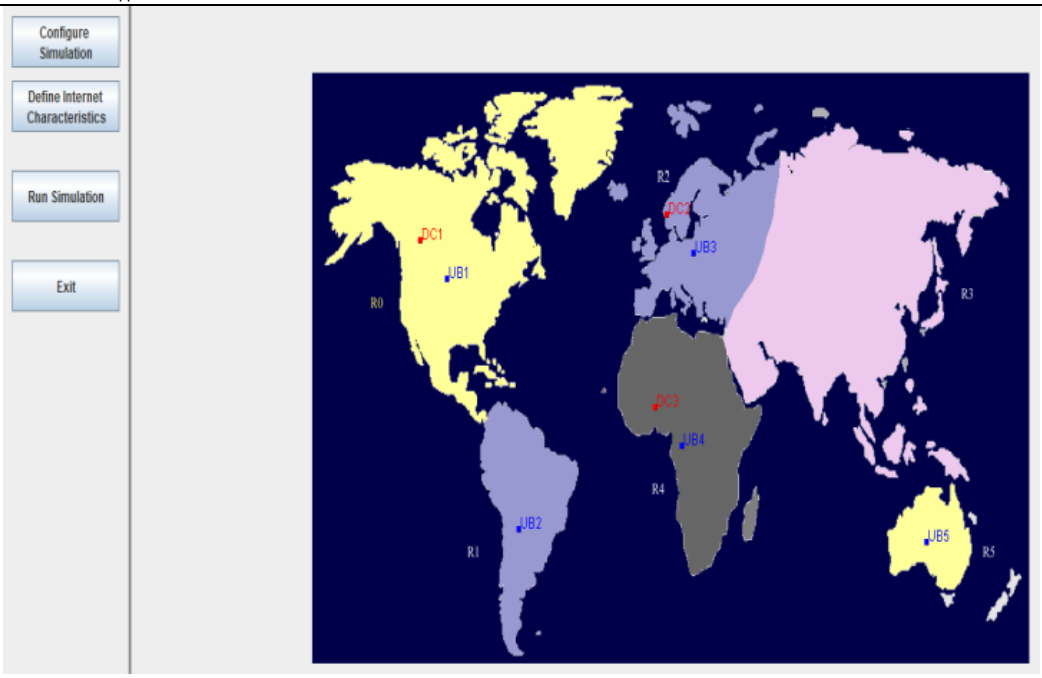

Fig. 3- Regions boundaries

7.1 Main Configuration

Under this option, two major terms are considered, the first one is user bases and second one is application deployment configuration. Before starting any activity, simulation time is to be mentioned in minutes first. This time is the actual duration on which simulation actually work.

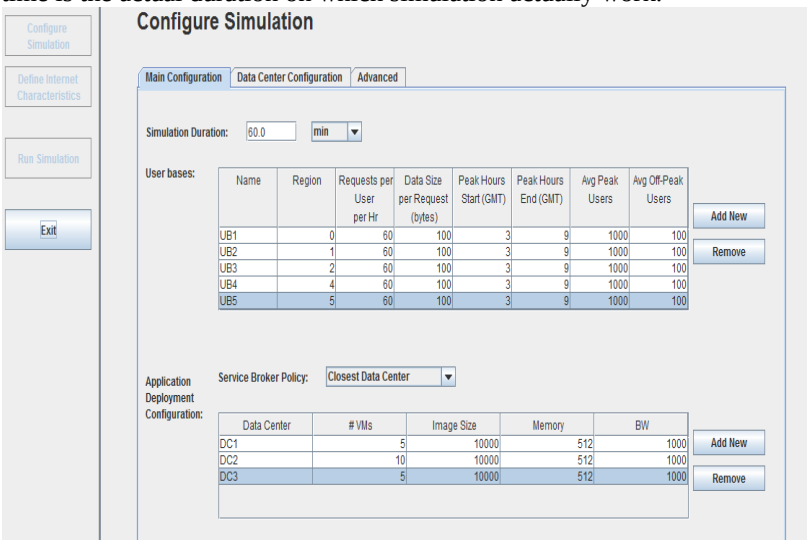

Fig. 4- Configuration Simulation

In Fig. 5, Data center configuration is shown, it is an arrangement that deals with the details of each data center which include its name, its region etc.

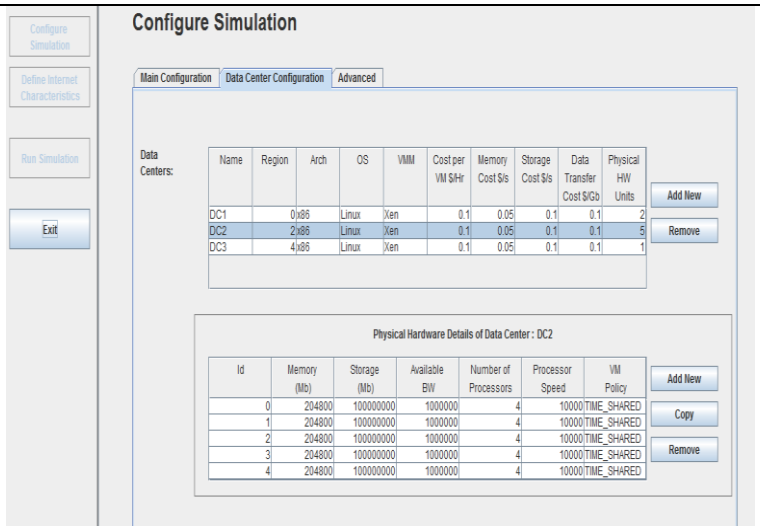

Fig. 5-Data Centre Configuration

Fig. 6 shows advanced configuration from where scheduling algorithm can be selected to perform various jobs on cloud as required. Delay matrix and bandwidth matrix both comes under internet characteristics. Both can be configured in the step shown in Fig. 7.

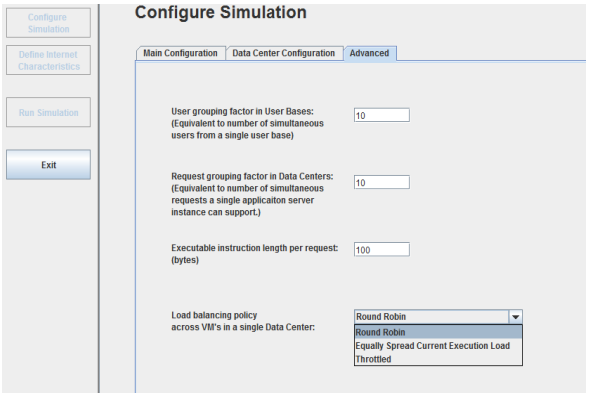

Fig. 6- Advanced Configuration

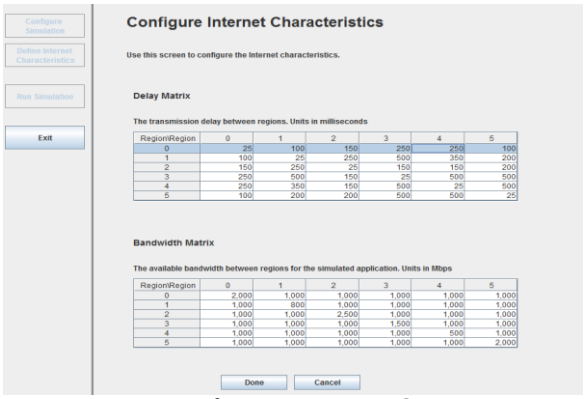

Fig. 7- Configuring Internet characteristics

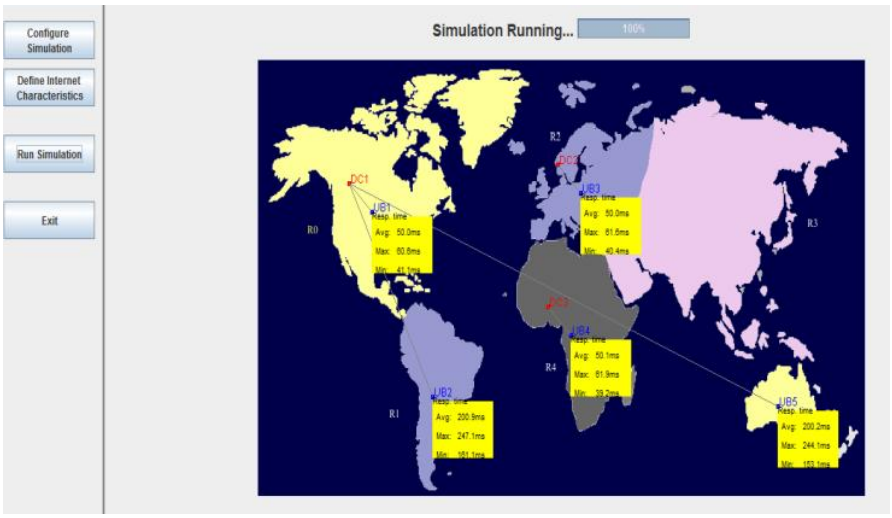

Fig. 8-Running simulation

When user click on the left button labeled with "Run Simulation", the actual simulation performed. This will show the maximum, minimum and the average time of each data center under which various jobs runs. Fig. 9 shows the overall summary report, in this Figure, response time is shown. This time is the used base hourly average response time. This output shows both graphs and tables for Max, Min and Avg. Hourly average response time of user bases is shown in Fig. 10. This time is dependent on the hardware configuration of data centre. This also involves three variations i.e. Max, Min and average. This is actually data centre request servicing time.

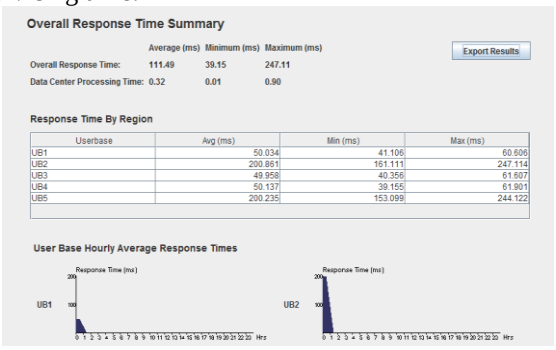

Fig. 9- Overall response time summary

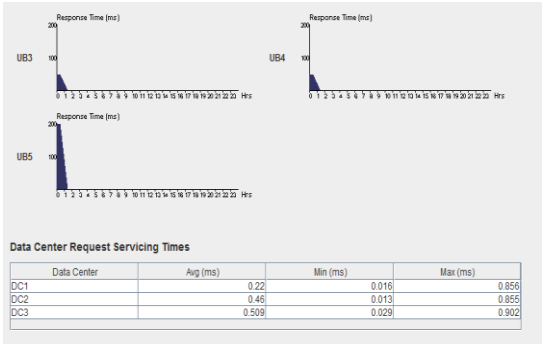

Fig. 10- User base hourly avg. Response time

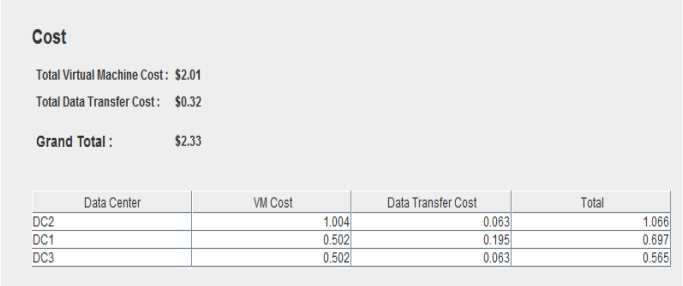

Fig. 11-Cost

Fig. 13 shows the cost of computing. This included three data centres as DC1, DC2 and DC3. There in each data centre cost of Virtual machine and cost of data transfer is considered. On behalf of both these parameters, total cost is calculated.

VIII. CONCLUSION

In cloud, it possible for the user to access information from anywhere at any time. The cloud takes away the need for user to be in the identical physical location as the hardware that stores data. Small companies can store their information in the cloud, removing the cost of purchasing and storing memory devices. Cloud computing provides enormous benefits. This paper is based upon the simulation through which a quantity of data centers and user bases can be deployed on diversed regions and after setting a little parameters on behalf of memory, machine, storage, bandwidth etc the simulation is performed. Any number of user bases and data centers can be deployed. The broker policy is also involved in this paper around SLA (service Level Agreement). So, closest data center, Optimize response time and dynamically load balancing broker policies are involved. For each simulation only one type can be considered at a time. For load balancing three algorithms are involved, such as round robin, throttled and equally spreading execution load. Only one algorithm can take charge at a particular time depending on the circumstances. The simulation is based on algorithms and parameters under which statistics are entered. On behalf of the algorithm's performance and metrics, this simulation takes place. The result of simulation will produce graphs and tables which show the overall response time of data centres, user bases on behalf of regions. The response time of all the entities is merely based on the distribution method and algorithm used.

REFERENCES

[1] Rakesh Kumar Mishra , Sreenu Naik Bhukya, "Service Broker Algorithm for Cloud-Analyst ",International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014.

[2] Chandrasekaran K., "Load Balancing of Virtual Machine Resources in Cloud Using Genetic Algorithm", Elsevier Publications, pp (156-168), 2013.

[3] Walter Cerroni, "Network performance of multiple virtual machine live migration in cloud federations", Cerroni Journal of Internet Services and Applications, Springer, 2015.

[4] B.Poornima, Dr.T.Rajendran, "Improving Cloud Security by Enhanced Has be Using Hybrid Encryption Scheme", International Journal of Advanced Research in Computer Science and Software Engineering, 2014.

[5] Soumya Ray, Ajanta De Sarkar , "Execution analysis of load balancing algorithms in cloud computing environment", International Journal on Cloud Computing: Services and Architecture, 2012.

[6] https://www.google.co.in/search?q=cloud+computing&newwindow=1&biw=1366&bih=598&source=lnms&tbm=isch&sa=X&ved=0ahUKEwjv6d7TscrLAhUJUZQKHXW9C30Q_AUICCgC#imgrc=3SRVYh1jSc_d6M%3A

[7] http://www.thbs.com/thbs-insights/cloud-computing-overview.

[8] Pankaj Sareen, "Cloud Computing: Types, Architecture, Applications, Concerns, Virtualization and Role of IT Governance in Cloud", International Journal of Advanced Research in Computer Science and Software Engineering, 2014.

[9] Qian Zhang, Hong Liang and Yongshan Xing, "A Parallel Task scheduling Algorithm Based on Fuzzy Clustering in Cloud Computing Environment", International Journal of Machine Learning and Computing, 2014.

[10] Simar Preet Singh, Anju Sharma and Rajesh Kumar, "Analysis of Load Balancing Algorithms using Cloud Analyst", International Journal of Grid and Distributed Computing, Vol. 9, No. 9, 2016.

[11] Nusrat Pasha, Dr. Amit Agarwa," Round Robin Approach for VM Load Balancing Algorithm in Cloud Computing Environment", International Journal of Advanced Research in Computer Science and Software Engineering, 2014.

[12] Shilpa Kumari, Tanuj Wala, "Study of Cloud Simulation Tools based on Cloud, International Journal of Computer Science and Mobile Computing, ol.4 Issue.5, May-2015

[13] Reena Panwar, Bhawna Mallick, "A Comparative Study of Load Balancing Algorithms in Cloud Computing", International Journal of Computer Applications, 2015.

[14] Rajkumar Somani, Jyotsana Ojha, "A Hybrid Approach for VM Load Balancing in Cloud Using CloudSim", International Journal of Science, Engineering and Technology Research (IJSETR), 2014.

[15] Bhupendra Panchal and R.K.Kapoor, "Performance Enhancement of Cloud Computing using Clustering", IJCSNS International Journal of Computer Science and Network Security, 2014.

[16] Shilpa B Kodli, Malashree S Valsang, "Dynamic Rebalancing Technique for QOS Provisioning in Cloud Storage Service", International Journal of Computer Science and Mobile Computing, Vol.5 Issue.8, August-2016

[17] Volker Fusenig and Ajush Sharma, "Security Architecture for Cloud Networking", Fujitsu 2012.

[18] Hetal V. Patel, Ritesh Patel, "Cloud, "Analyst: An Insight of Service Broker Policy", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 1, January 2015

[19] Mayanka Katyal, Atul Mishra, "A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment", International Journal of Distributed and Cloud Computing, 2013.

[20] Byungseok Kang, Hyunseung Choo, "cluster-based decentralized job dispatching for the large-scale cloud", Kang and Choo EURASIP Journal on Wireless Communications and Networking, Springer, 2016.