International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622
International Conference on Industrial Automation and Computing (ICIAC- 12-13th April 2014)

RESEARCH ARTICLE OPEN ACCESS

Load Balancing Mechanism with Mobile Based Cloud Environment

Rupali Kadu*, Swapnili Karmore**
*(Depart ment of Co mputer Science and Engineering, G.H. Raisoni College of Eng ineering Nagpur, India
** (Depart ment of Co mputer Science and Engineering, G.H. Raisoni College of Eng ineering Nagpur, India

ABSTRA CT-
Cloud computing is a type of parallel and distributed system wh ich enables on demand network access to shared pool of computing resources such as storage, servers and applications. These resources can be accessed by the user on pay per basis. In Cloud computing private cloud is provisioned for the users of single organization and can be accessed from anywhere and at any point of time w ithin organization wh ich needs more customizat ion and control over data. Load Balancing is an important issue in cloud computing environment which is required to achieve evenly distributed load and to efficiently make use of resources. Many of the Load balancing algorith ms were proposed on cloud, but up till now there is no such mechanism or algorithm is proposed on
Mobile based cloud server. In this era of Smartphone's there is highly increasing demand of services and applications from personal Smartphone devices, there is need to develop the load balancing mechanism to Maintain the service population and satisfy the user with min imu m response time and system with highperformance to get the instant access of services and applications. The key idea behind the paper is to develop a load balancing mechanism based on Support Vector Machine to balance the load across the mobile servers in private cloud.
Keywords- Cloud Co mputing, Load Balancing, Private Cloud, Service population, SVM

I. INTRO DUCTION

Cloud Co mputing became very popular in the last few years. With the rapid development of internet and network technology, large number of people uses the internet to obtain informat ion, shopping and entertainment. With the development of Web, Internet mu ltimed ia is emerging as a service [1].

Cloud has **five** characteristics, on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service deployed on four models such as Private cloud, community cloud, public cloud and hybrid cloud with **three** service models Software as a service (SaaS), Platform as a service (PaaS), and Infrastructure as a service (IaaS) in its architecture.

Private cloud is provisioned for the users of a single organization exclusively wh ich can be accessed from anywhere and at any point of time within the organization. Public cloud is fo r open use by general public. Co mmunity cloud is provisioned for the use by a specific co mmunity Hybrid cloud is a composition of two or more d istinct clouds such as private, public or co mmunity clouds.

In many ways, managing a private cloud is not so different its like to manage an on -premises data storage. The success of a cloud environment depends on many components that are workload management, security, and network planning and, server density. Before p lacing workload on a cloud server, administrators must be ready to plan their physical server environment. In this planning phase, cloud managers can size the environment, should know what workloads they are delivering and truly understand availableresources.

Indistributed bcomputing, usersare allowed to log in from any device, anywhere, at any time which means an organization's cloud environ ment must be able to handle user fluctuations. Without good server load balancing, a cloud environment can experience degraded performance as cloud servers take on more workloads than they're capable of. Admin istrators must be able to take time to actually evaluate which workloads are being deployed into the cloud .

Rest of this paper is exp lained as follo ws. Section II explains the current literature and discusses the algorith ms proposed to solve the load balancing issues in Cloud Co mputing. Section III Describes the proposed plan of the system and Section IV is the conclusion and future work

II. LITERATURE SURVEY

Load balancing has two meanings: first, it puts a large number of concurrent accesses or data

traffic to mu ltip le nodes respectively to reduce the time users wait ing for response; second, it put the calculation fro m a single heavy load to the mult iple nodes to improve the resource utilization of each node. Load balancing is the same as a system turns on duty, which tasks assigned the tasks to each person to avoid any of them too tired.

Cloud Co mputing is the widely used term and emerging co mputing model based on the development of distributed computing, parallel processing and grid computing [2]. Th is distributes the computational tasks to the large number of computers provides the various application systems that use the computational power, storage and a various number of software and services according to the requirement.

Due to the increasing number users, there are thousands of tasks or requests at a time in the cloud computing environ ment. It's a challenge to design the less complex load balancing algorith m based on task scheduling [3] . If the system crashed due to server maintenance or failure occurs, it will cause a loss to the cloud users. So, how to deal with the requests in the dynamic system is also an important issue [4].

Genetic Algorithm with an immigrant scheme:

This scheme takes into account the practical dynamic mu ltiservice scenario in wh ich each server cluster only handles the specific type of multimedia task and each client requests a different typeof mult imedia service at different time. In this algorith m, atdifferent time steps, Client can change their locations, and each server cluster only handled a specific type of mult imedia task, these two objectives can be optimized at the same time [5].

Main disadvantage of this scheme is that every time load is not updated in the system (resource manager), after executing the service on each server so that exact load on each server can't be predicated. Each server will handle only specific type of task; if the one of the server will be failed it will unable to provide service.

A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment: Th is strategy also uses the Genetic Algorith m. It uses the tree structure according to historical data and current state of the system wh ich computes load ahead and influences it. This will have on the system after the deployment of the needed VM resources and then chooses the least affective solution, through which it achieves the best load balancing, and reduces or avoids dynamic migrat ion solveing the problem of load imbalance and high migrat ion cost by traditional algorithms after scheduling [6].

In real cloud computing environ ment, there might be dynamic change in VMs, and there also might be an increase of co mputing cost of virtualization software and some unpredicted load wastage with the increase of VM nu mber started on every physical machine. Therefore, mon itoring and analy zing mechanis m is needed to better solve the problem o f load balancing .

Load Balancing strateg y of Cloud Computing Based on Artificial Bee Algorithm: This mechanis m presents the improved Artificial Bee Colony (ABC) Algorithm bas ed on the behavior of honey bees which optimizes the amount of nectar (i,e . system throughput) to reach the maximu m throughput. Since colony algorithm arranges only a litt le link between requests in the same server queue, the improvement of system throughput is suboptimal. Furthermore, in a certain number of servers the increasing request does not lead to the increase of system throughput [7]. As number of requests increases the above algorithm becomes unstable, rather the original algorithm shows the better stability.

An Improved biased Random Sampling Algorithm for Load Balancing in Cloud Computing: In this algorithm, biased random sampling algorithm is redesigned by making use of queue length and processing time as parameters for arriving at load balancing among competing nodes. It allocates the new request to the least loaded node in random walk and the number of free resources act as a parameter for load balancing. So the request is allocated to the node having minimu m number of free resources [8]. The main drawback of this algorith m is it selects the node randomly and the same node is getting selected again and again. It will not specify any criteria to choose the particular node in cloud.

Active Clustering Active clustering is an enhanced method of random sampling, where this algorithm works on the princip le of grouping similar nodes together and start working on these group nodes. This method uses the resources efficiently thereby increases the throughput and performance of the system by using high resources. In this approach a technique called match-maker is introduced.

When an execution starts in a network, the process gets initiates and searches for the next matching node said to be match-maker wh ich should satisfy the criteria that it should be the different one from the former one. Once the match-maker is found the process gets initiated and as soon as the process gets over the match-maker gets detached from the

network. Thus this is an iterative process in the network to balance the load efficiently.It optimizes job assignment by connecting similar services by local re-wiring. Drawback is this method degrades its performance when increase in diversity of nodes [9].

It becomes a key issue in the field of cloud computing platform about how to coordinate the server load by suitable load balancing algorithm to improve cloud computing research resource utilizat ion and the system performance.

All the above algorith m concludes that still there is need to improve in load balancing in cloud environment. Demand for the quick access service is increasing day by day and also there is tremendous use of Android Mobile phones so it will be beneficial for the user to use the service anytime. Which helps to improve the service population results in instant access of the frequently used websites and applications on Internet. Number of the strategy and algorith ms are proposed for load balancing in cloud environment such as dynamic, distributed, heterogeneous depending on various criteria but up till now mob ile based cloud is not created which is private and can be accessed by user anywhere anytime or it also useful in private areas such as offices, colleges, and private organizat ions.

III. PRO POSED METHO DOLOGY

The existing Load Balancing algorithm shows the considerable imp rovement in the Cloud Co mputing Environment which includes the Throughput, Overhead Associated, Migration Time, Fault tolerance and Response Time, Scalability, Resource Utilizat ion, Performance accord ing to the their requirement wh ich is applicab le to the specific area or the environment. Due to the explosive demand for this Cloud Environment, still there is a lot of wo rk is needed to be done specifically on the response time and system performance as in today's era user always expect the quick, powerful and stable services.

Here in this paper trying to put the methodology which improves the performance of the system and minimize request response time. System will be based on the android mobile devices which act as the servers, control server in the Cloud.

3.1 Proposed Model

- Register all the mobile devices which act as a server and specifically control server in the network where it is located with their respective parameters (such as IP address, Port number)
- Establishing successful communication between mobile devices, in between the server, control server and Client.
- Service Provision based on Support Vector Machine (SVM) performs the load balancing.
- Selecting the server to provide the quick access to the client according to load on each server. Server with the min imu m load provides the service with min imu m delay.
- Accessing the client request according to their requirement.

3.2 Techniques to be implemented

Mobile based cloud environment enables users to access multimedia applications via Android based appliances. Wireless access technologies such as 3G, Wi-Fi, W iMAX etc. integrated in Cloud Co mputing Environment to provide services to the users efficiently. By Co mbining Cloud Co mputing an Android properties System achieves the efficient and convenient communicat ion.

While using the Smartphone's, two parameters to be considered are very important main ly Battery Available and Memory used and free memo ry wh ich defines the performance of the Smartphone's.

Designing mobile based cloud is a challenge and these two factors should be considered to get the better output of the server and system performance. To compute the Load on Mobile Server, considering three main parameters Battery Available, Free Memory depending on the used memory and number of running applications in the mobile server. Important parameter is the number of requests currently running or pending at the server. These three parameters contribute the exact load on the Mobile Server which helps to load imbalance in the cloud.

Support Vector Machine optimizes the vector (B,M,R) Battery remaining (B), Memory Available(M), Nu mber of Requests (N) on each server which helps to get the parameter values on each server.

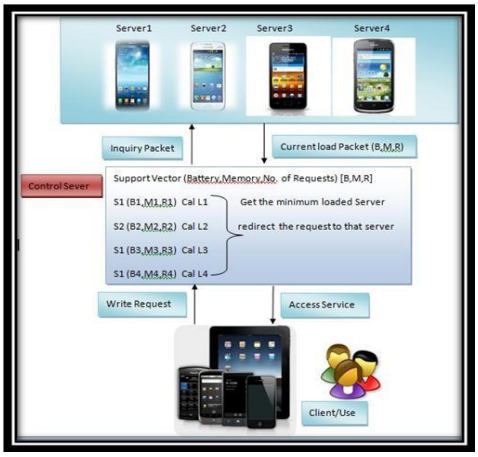

Fig. 1.Architecture of Cloud with Load Balancing

Using this parameter values load on each server can be calculated such as

Load = Number of Requests (N) + Memory Available (M) + 100 / Battery remaining (B)

SVM optimizes the load on each server and select the best server from the cloud with minimu m load to provide the service to the user. It is using as an optimization engine, to select best element (with regard to some criteria) fro m some set of available alternatives

Support Vector Machine (SVM) Algorithm, works as Follows:

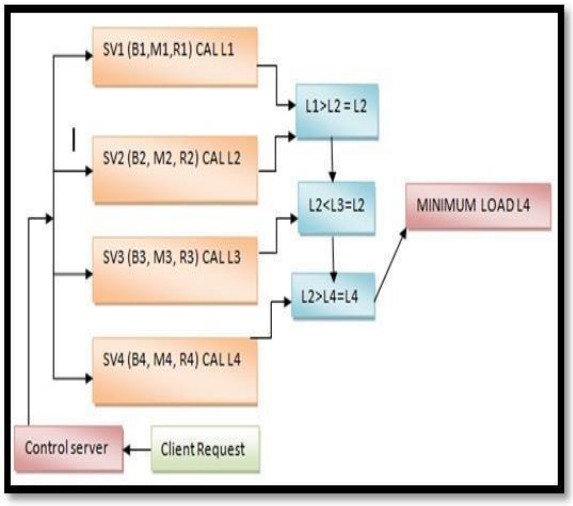

Fig.2.Working of SVM

Algorith m works as follows

1. Client sends request to Central server.
2. Central server sends INQUIRY PA CKET to the cloud servers.
3. Cloud servers respond with CURRENT LOAD PA CKET, which contains the vector (B,M,R) on each cloud.
4. Calcu late the load on each server.
5. Central server receives the CURRENT LOAD PA CKET and sends the client request to the server with M inimu m load.

This process is repeated.

It has updating mechanism, when client sends the request to the central server it will update load of each server so that it becomes easier to balance the load of the Cloud Servers every time .

Genetic algorith m has the disadvantage as it will not update the load every time, it only forward request to server so the exact load at the central server cannot be determine.

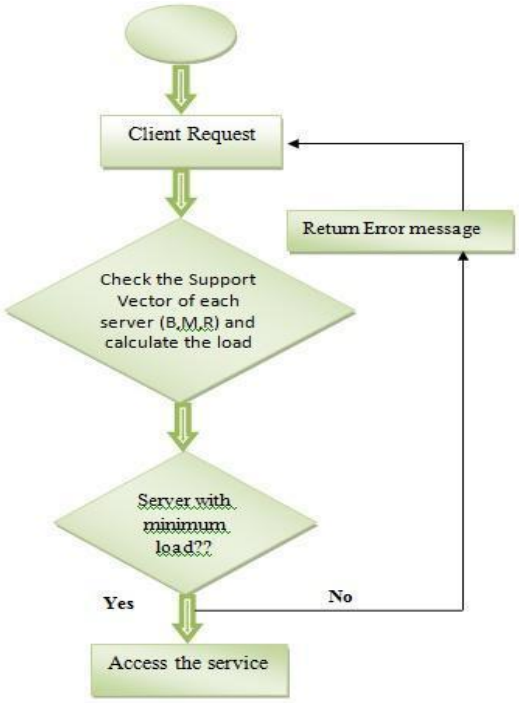

Fig.3.Flow of the Control Server

IV. RESULT

After registering all the servers in network and forms the cloud, Client Server Co mmunication is established. Status of the Load on each server is shown with its IP address. Here, http://192.168.2.4:8080 is IP address of Control Server. As two servers are there in cloud, showing the load on two servers, redirecting the request to, http://192.168.2.6:8080 the server with min imu m load providing access of the services.

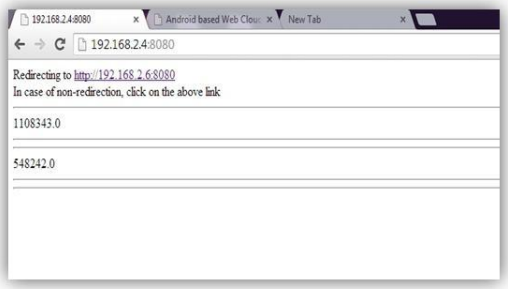

Fig.4. Control server with status of load

The main goal is to imp lement Support Vector Machine Algorithm to balance the load in cloud and analyze the result which will imp rove the system performance and reduce response time.

V. CO NCLUSION

Load balancing is one of the main challenges in cloud computing. Here try ing to imp lement the Support Vector Machine Algorithm

considering the basic need of cloud which is response time and system performance through the client server mechanis m. Providing the services according to their requirement i.e. mu ltimed ia content. It will be helpful to the client to save mult imedia data over the cloud and access it whenever they want.

FUTURE WO RK

As soon as all the mobile devices are registered in the network with their details, the control server application can continuously monitor data and request which are processed by each server in the cloud. Future work involves improving the Quality of service and user experience according to the parameters such as battery power processing time and memory storage of the server as mobile devices act as server.

The proposed model is initially designed to be implanted for the campus or smaller network. But keeping future scope in mind, the model is designed flexible enough for the further modification to implement it on the larger network.

ACKNOWDGEMENT

This is my M.E. final year project paper. My sincere thanks to my honorable guide Prof. Swapnili Karmore and others who have contributed towards the preparation of the paper.

REFERERENCES

[1]. W. Zhu, C. Luo, J. Wang, and S. Li, ―Multimedia cloud computing, IEEE Signal Process. Mag., vol. 28, no. 3, pp. 59–69, May 2011. J. U. Dunco mbe, ―Infrared navigation — Part I: An assessment of feasibility, IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan.2000.

[2]. Vaquero L M, Rodero Merino I, Caeeres J. ―A break in the clouds: Towards a cloud definit ion [J]. ACM SIGCOMM Co mputer Co mmunicat ion Review. Load Balancing co mputing[EB/OL].

[3]. K. Noguchi, M. Dillencourt,.L Bic, ―Efficient global pointers with spontaneous process migrat ion in: I EEE Proceedings of 16th E

[4]. uromicro Conference on Parallel, Distributed and Network Based Processing, PDP 2008, vol. 00, 2008.

[5]. Chun-Cheng Lin, Member, IEEE, Hui-Hsin Chin, Student Member, IEEE, and Der JiunnDeng, ―Dynamic Multiservice Load Balancing in Cloud-Based Multimedia System IEEESYSTEMS JOURNA L , August 2012.

[6]. Jinhua Hu, Jianhua Gu,Guofei Sun,Tianhai Zhao ―A Scheduling Strategy on Load Balancingof Virtual tMachine Resources in CloudCo mputingEnviron ment 3rd International Symposiu m on Parallel Architectures Algorith ms and Programming.

[7]. Jing Yao, Ju-hou He ―Load Balancing strategy of Cloud Co mputing Based on Artificial Bee Algorith m

[8]. Sheela s Manakattu, Madhu Kumar S D ―An Improved biased Random Sampling Algorithm for Load Balancing in Cloud Computing

[9]. InternationalConference on Advances in Co mputing, Co mmunicat ions and Informatics (ICACCI-2012) August 03 - 05 2012,CHENNAI, India

[10]. O. Abu- Rah meh, P. Johnson and A. Taleb-Bendiab, ―A Dynamic Biased Random Samp lingScheme for Scalable and Reliable Grid Networks INFOCOM P- Journal of Co mputer Science, ISSN 1807-4545, 2008, VOL.7, N.4, December, 2008, pp. 01- 10.

[11]. Swapnili P. Karmo re, Anjali R. Mahajan,Suruchi Kitey ―Battery mon itoring and analysis for Android based system , IEEE Advanced Co mputing Technologies (ICA CT), 2013 15th International Conference on 21-22 Sept. 2013.