LOAD BALANCING IN CLOUD COMPUTING

[1] PRATHMESH ARNIKAR, [2] SIDDHARTH SONAWANE, [3] ANKITA FALE, [4] SAGAR AGHAV, [5] SHIKHA PACHOULY

[1] Department of Computer Engineering, AISSMS College of Engineering, Pune , Maharashtra, India.
[2] Department of Computer Engineering, AISSMS College of Engineering, Pune , Maharashtra, India.
[3] Department of Computer Engineering, AISSMS College of Engineering,Pune , Maharashtra, India.
[4] Department of Computer Engineering, AISSMS College of Engineering Pune , Maharashtra, India.
[5] Department of Computer Engineering, AISSMS College of Engineering Pune , Maharashtra, India.

prathmesh10arnikar@gmail.com,siddharthsonawane2802@gmail.com,ankita.fale@yahoo.com,sagar.aghav08@gmail.com,shikhajp@yahoo.co.in

ABSTRACT : *"The major challenge which is been faced in cloud computing nowadays is balancing of the load on to the clouds. Cloud computing refers to an on-demand service in which shared resources, softwares are provided to the user on pay-as-need basis. This paper presents a better load balancing approach in cloud computing by implementation of scheduling algorithm".*

I. INTRODUCTION

Cloud computing is an internet based computing which focuses on the term cloud. Cloud refers to a special network in which large number of host machines or networks are interconnected and which provides shared resources, softwares and other information. Thus cloud computing refers to accessing this cloud as per user demands.

The characteristic of cloud computing are

1) user centric
2) task centric
3) accessible

In case of Cloud computing services can be used from diverse and widespread resources,rather than remote servers or local machines. There is no standard definition ofCloud computing. Generally it consists of a bunch of distributed servers known as masters,providing demanded services and resources to different clients known as clients in anetwork with scalability and reliability of datacenter. The distributed computers provide on demand services. Services may be of software resources (e.g. Software as a Service,SaaS) or physical resources (e.g. Platform as a Service, PaaS) or hardware/infrastructure(e.g. Hardware as a Service, HaaS or Infrastructure as a Service, IaaS). Amazon EC2(Amazon Elastic Compute Cloud) is an example of cloud computing services.

Cloud Components

A Cloud system consists of 3 major components such as clients, datacenter, anddistributed servers. Each element has a definite purpose and plays a specific role.

Three components make up a cloud computing solution

Type of Clouds

1. Clients

End users interact with the clients to manage information related to the cloud. Clientsgenerally fall into three categories as given

- Mobile: Windows Mobile Smartphone, smartphones, like a Blackberry, or an iPhone.
- Thin: They don't do any computation work. They only display the information. Servers do all the works for them. Thin clients don't have any internal memory.
- Thick: These use different browsers like IE or Mozilla Firefox or Google Chrome toConnect to the Internet cloud.Now-a-days thin clients are more popular as compared to other clients because oftheir low price, security, low consumption of power, less noise, easilyreplaceable and repairableetc.

2.Datacenter

Datacenter is nothing but a collection of servers hosting different applications. A enduser connects to the datacenter to subscribe different applications. A datacenter may exist at a large distance from the clients.

Now-a-days a concept called virtualisation is used to install a software that allowmultiple instances of virtual server applications.

3.Distributed Servers

Distributed servers are the parts of a cloud which are present throughout the Internet
hosting different applications. But while using the application from the cloud, the user willfeel that he is using this application from its own machine.

Service means different types of applications provided by different servers across thecloud. It is generally given as "as a service". Services in a cloud are of 3 types as given in

- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Hardware as a Service (HaaS)
- Infrastructure as a Service (IaaS)

Software as a Service (SaaS)

In SaaS, the user uses different software applications from different servers throughthe Internet. The user uses the software as it is without any change and do not need to makelots of changes or doen't require integration to other systems. The provider does all theupgrades and patching while keeping the infrastructure running.

The client will have to pay for the time he uses the software.The software that does a simple task without any need to interact with other systems makes it an ideal candidate forSoftware as Service. Customer who isn't inclined to perform software development but needs high powered applications can also be benefitted from SaaS.

Customer resource management (CRM)
_ Video conferencing
_ IT service management
_ Accounting
_ Web analytics
_ Web content management

Platform as a Service (PaaS)

PaaS provides all the resources that are required for building applications and servicescompletely from the Internet, without downloading or installing a software.

PaaS services are software design, development, testing, deployment, and hosting.Other services can be team collaboration, database integration, web service integration,data security, storage and versioning etc.

Hardware as a Service (HaaS)

It is also known as Infrastructure as a Service (IaaS). It offers the hardware as a serviceto a organization so that it can put anything into the hardware according to its will [1].

HaaS allows the user to "rent" resources

_ Server space
_ Network equipment
_ Memory
_ CPU cycles
_ Storage space

Types of clouds

Public cloud

Public cloud applications, storage, andother resources are made available to the general public by a service provider. These services are free oroffered on a pay-per-use model. Generally, public cloud service providers are Amazon AWS, Microsoft and Google

Private cloud

Private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party and hosted internally or externally.

Hybrid cloud

Hybrid cloud is a composition of two or more clouds (private, community or public) that remain unique entities butare

Since job arrival pattern is not predictable and capacities of each node differs,there occurs a problem of balancing of load"hybrid cloud" architecture,companies and
individuals are able to obtain degrees of
fault tolerance combined with locally.

An approach to the load balancing

It is a process of reassigning the total load to the individual nodes of the collective system to make resource utilization effective and to improve the response time of the job, simultaneously removing a condition in which some of the nodes are over loaded while some others are under loaded.

II. OBJECTIVES

- Server balances the load by using external resources
- No proper utilization of resources
- Performed in private cloud
- Improves the response time and effective job processing
- Network reliability

III. RELATED WORK

Cloud computing is efficient and scalable but maintaining the stability of processing so many jobs in the cloud computing environment is a very complex problem with load balancing receiving much attention for researchers. Since the job arrival pattern is not predictable and the capacities of each node in the cloud differ, for load balancing problem, workload control is. Crucial to improve the
system performance and maintain stability. Load balancing schemes depending on whether the system dynamics are important can be either static or dynamic. Static schemes do not use the system information and are less complex while dynamic schemes will bring additional costs for the system but can change as the system status changes. A dynamic scheme is used here for its flexibility.

IV. PROPOSED SYSTEM

Product Scope

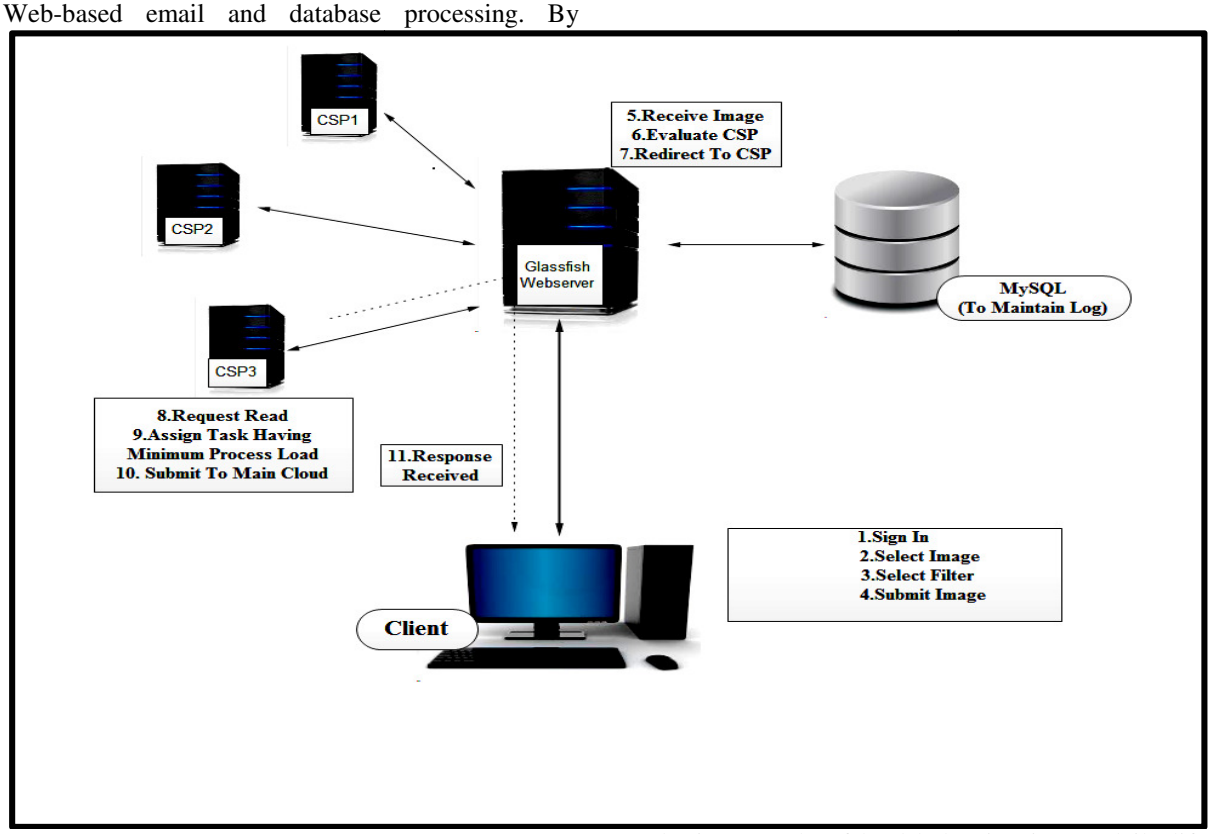
Cloud computing has become very popular because it moves the processing efforts from the local devices to the data centric facilities. Therefore, any Internet enabled devices like a smart phone, Tablets etc.,

could be able to solve complex operations like creating editing files online by simply passing specific arguments to a service running at the data centers level that will be capable ofgiving back the results in a very short time.

With the rapid development of technology, most of the colleges and universities offer onsite classes, courses, where in some cases the entire degree program through online as well as uses various other teaching and learning models. Most of these online offerings currently are at undergraduate level. But there is a growing trend in using similar models for graduate and postgraduate education too.

There are many different technologies used in online teaching and learning arena, but to implement them in an efficient way to fulfill the needs of both the students and institutions it's better to adopt the cloud computing services to provide the uninterrupted, highly scalable and extreme qualitative services. The cloud services mainly include sharing, online storage, Web-based email and database processing. By adapting the Cloud computing, it becomes easy to share the virtualized resources. Here Users do not need any background knowledge of the services and it's very easy to maintain when compared to any traditional technologies.

The load balancing model given in this article is aimed at the public cloud which has numerous nodes with distributed computing resources in many different geographic locations. Thus, this model divides the public cloud into several cloud partitions. When the environment is very large and complex, these divisions simplify the load balancing. The cloud has a main controller that chooses the suitable partitions for arriving jobs while the balancer for each cloud partition chooses the best load balancing strategy.



OVERALL DESCRIPTION

Product Perspective

Load Balancing is a method to distribute workload across one or more servers, network interfaces, hard drives, or other computing resources.

Load balancing is a relatively new technique that facilitates networks and resources by providing a Maximum throughput with minimum response time. Dividing the traffic between servers, data can be sent received without major delay. Different kinds of algorithms are available that helps traffic loaded between available servers.

A basic example of load balancing in our daily life can be related to websites.

Without load balancing, users could experience delays, timeouts and possible long system responses. Load balancing solutions usually apply redundant servers which help a better distribution of the communication traffic so that the website availability is conclusively settled. Load balancing in the cloud differs from classical thinking on load-balancing architecture and implementation by using commodity servers to perform the load balancing. This provides for new opportunities and economies-of-scale, as well as presenting its own unique set of challenges.

Product Functions

Application of load balancing and redundant mirrored databases in clusters techniques, across multiple availability zones, reduces the chance of outages that could simultaneously affect the services in cloud systems. If an outage affected one system, the load balancer is able to switch to another available resource.

Load balancing techniques, in the area of cloud computing, reduces costs associated with document management systems and maximizes availability of resources reducing the amount of downtime that affect businesses during outages. This article discusses possible ways to improve the performance of cloud networks by the introduction of resource load balancing technique that uses the message-oriented middleware within the web service oriented model of software architecture.

Goals of Load balancing are:

- To improve the performance substantially.
- To have a backup plan in case the system fails even partially.
- To maintain the system stability.
- To accommodate future modification in the system.

SHA1

The SHA1 encryption algorithm specifies a Secure Hash Algorithm (SHA1), which can be used to generate acondensed representation of a message called a message digest. The SHA1 is required for use with the DigitalSignature Algorithm (DSA) as specified in the Digital Signature Standard (DSS) and whenever a secure hashalgorithm is required. Both the transmitter and intended receiver of a message in computing and verifying a digital signature uses the SHA1.

The SHA1 is called secure because it is computationally infeasible to find a message which corresponds to agiven message digest, or to find two different messages which produce the same message digest. Any change to a message in transit will, with very high probability, result in a different message digest, and the signature will fail to verify.

V. APPLICATIONS

1.This Architecture can be implemented in the fields of cloud computing where huge load is created on server.

2.Also to handle the random selection based load distribution problem.

VI. CONCLUSION

Cost and time are the key challenge of every IT engineer to develop products that can enhance the business performance in the cloud based IT sectors. Current strategies lack efficient scheduling and resource allocation techniques leading to increased operational cost and time. This aims towards thedevelopment of enhanced strategies through improved job scheduling and resource allocation techniques for overcoming the above-stated issues. Here, Equal Spread Current Execution Load algorithm dynamically allocates the resources to the job in queue leading reduced cost.

ACKNOWLEDGMENT

We are grateful for all the support extended towards successful completion of this paper by AISSMS'S college of Engineering,Pune. We would like to offer our sincere thanks to our guide Prof. Shikha Pachouly.

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A.D Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the Clouds: A Berkeley View of Cloud Computing", EECS Department, University of California, Berkeley,Technical Report No., UCB/EECS-2009-28, pages 1-23, February 2009.

[2] R. W. Lucky, "Cloud computing", IEEE Journal of Spectrum, Vol. 46, No. 5, May 2009, pages 27-45.

[3] M. D. Dikaiakos, G. Pallis, D. Katsa, P. Mehra, and A. Vakali, "Cloud Computing: Distributed Internet Computin for IT and Scientific Research", IEEE Journal of Internet Computing, Vol. 13, No. 5, September/October 2009, pages 10-13.

[4] G. Pallis, "Cloud Computing: The New Frontier of Internet Computing", IEEE Journal of Internet Computing, Vol. 14,No. 5, September/October 2010, pages 70-73

[5] B. P. Rima, E. Choi, and I. Lumb, "Resource allocation of Cloud Computing Systems", Proceedings of 5th IEEE International Joint Conference on INC, IMS and IDC, Seoul, Korea, August 2009, pages 44-51.