© 2016 IJSRSET | Volume 2 | Issue 4 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Monitoring and Normalization of the Loaded Servers By Load Balancing

Dr. D. Ravindran, M. Ramya

School of Computing Science, St. Joseph's College, Trichy, Tamil Nadu India

ABSTRACT

Cloud computing is a new computational model which is primarily based on distributed computing. The services and applications are accessible to the different clients using proper internet protocol suit and networking standards. Though it has many advantages the client has to consider also the drawbacks of the services provided by the cloud. Some of the issues have to considered like security, scalability, cost, SLA, etc. Most of the client users have to notice the performance of the cloud that the provider is issuing to the client. This paper describes about one of the factors when considered for the performance of the services provided. Scalability is the common factor which affects the satisfaction of the cloud users. By reviewing the cloud scalability factors, monitoring the server that are distributed over the network and its load balancing issue of the servers. This paper greatly deals with the normalization of the cloud servers.

Keywords : Performance; Scalability; Load Balancing; Monitoring; Normalize; Status

I. INTRODUCTION

Cloud computing are outlined as a computing surroundings wherever computing wants by one party are often outsourced to a different party and once would like be arise to use the computing power or resources like information or emails, they will access them via web[1]. Cloud computing is a model for enabling universal, suitable, on demand network accesses to a shared pool of configurable computing resources like networks, servers, storage, applications, and services that can be rapidly equipped and released with less management effort or service provider intervention. Monitoring of Cloud is a task of paramount importance for both Providers and Consumers. On the one side, it is a key tool for controlling and managing hardware and software infrastructures; on the other side, it provides information and Key Performance Indicators (KPIs) for both platforms and applications. Monitoring is clearly instrumental for all the activities covered by the role of Cloud Auditor. In more general terms, Cloud Computing involves many activities for which monitoring are an essential task.

II. METHODS AND MATERIAL

A. Related Work

M. Kriushanth et al.(2015)[2] describes the auto scaling values and that setting dynamic threshold values in a cloud environment should utilize the available resources completely and prevents the physical server damage. It manipulates the provider to accommodate more users in a physical server and also reduces the cost of the service. In this paper, the authors elaborate their concept in the area to set a dynamic threshold value for the physical server, load balancer behavior identifier mechanism is proposed to generate the rule and provide the resources dynamically. Major drawback is that it can't be used in cloud data.

Abhijit Aditya et al. (2015)[3] presents the basics of cloud computing like it's characteristics, deployments models, service models. They are describing the each service delivery models characteristics, its vendor types their advantages and disadvantages. Then they describing about where the load problems are occurring in the system and so the challenges in keeping mind. Then it describes each and every types

of algorithm in load balancing separately. Their properties, advantages, disadvantages are also described. They specially described about these algorithms based on the time factor.

Po-Huei Liang et al.(2015)[4] presents a framework for global server using for load balancing of the web sites in a cloud with two-level load balancing model. The proposed framework is intended for adjusting an open-source load-balancing system and while the customers need more load balancers for increasing the availability, this framework allows the network service provider to deploy the load balancer in different data centers dynamically. Further they described the load balancing algorithms with the various cloud service providers along with its communication interface.

Radha Ramani (2015)[5] presents a concept of Cloud Computing along with load balancing. The main thing is considered in this paper is load balancing algorithm. There are various mentioned algorithms in cloud computing which consists of many factors like scalability, enhanced resource utilization, high performance and improved response time. Further this paper provides the insight about the policies, characteristics, goals, current state classification, need for load balancing. They have proposed a frame work for giving the new algorithm.

B. Scalability

Scalability is the capability of a system, network, or process to handle a growing amount of work, or its potential to be enlarged in order to accommodate that growth[6]. For example, it can refer to the capability of a system to increase its total output under an increased load. Cloud balancing is a computer networking method to distribute work load across multiple computer or a computer cluster, network link, central processing units, disk drivers or other resources to achieve optimal resource utilization, maximize throughput, minimize response time and avoid overload. Load balancing helps in preventing bottleneck of system.

First step, algorithm is designed based on performance in heterogeneous environment of hosts The next process is to study the above algorithm with the effect of CPU utilization. Then based on the algorithm result, the algorithm will make the decision about the normalizing factor.

C. Load Balancing Approaches

Static and dynamic are the two type of load balancing approaches used in cloud computing[7,8].

- **Static Approach**

This approach is mainly defined in the design or implementation of system. Static load balancing algorithm divides the traffic equivalently among all users. It uses only information about the average behavior of the system. These are much simpler and ignore the current state or the load of the node in the system.

- **Dynamic Approach**

In this approach, the current state of the system was considered during load balancing decision. It is more suitable for widely distributed system such a cloud computing Dynamic approach has two parts[7]

- **Centralized Approach:** Only a single node is responsible for managing and distribution within the whole system.
- **Distributed Approach:** Each node independently builds its own load vector. Vector collecting load information of other node. All decision is made locally using local load vector.

D. Factors To Be Considered

Performance metrics also play a major role in monitoring techniques for resource management. As per the survey, it is considered as second issue in cloud computing. Poor performance can be caused by lack of resources such as disk space, limited bandwidth, lower CPU speed, memory, network connections etc. The data intensive applications are more challenging to provide proper resources. There is a series of factors that affect the performance such as [9]:

- Security.
- Recovery and Fault tolerance.
- Service level agreements.
- Bandwidth.
- Storage capacity.
- Physical memory.
- Disk capacity.
- Processor Power.

- Availability.
- Number of users and Workload.
- Usability.
- Scalability.
- Location, data centers and their distance from a user's location.

And there is a series of criteria for evaluating the performance such as[10]:

- Average response time per unit time.
- Average waiting time per unit time.
- Workload to be serviced per second (Mbps) or a unit of time.
- Throughput (Req / Sec).
- The average time of processing (exe / sec).
- Percentage of CPU utilization.
- The number of requests executed per unit time.
- The number of requests per unit time buffer.
- The number of rejected requests per unit time.

E. Workload Calculation

Workload condition in load balancer varies from small to heavy request[11]. As per the performance details, the proposed algorithm is based on the percentage of CPU utilization. To predict the workload and allocate the suitable resources, load balancing and scaling mechanisms are used. Scaling can be done in two ways called reactive and proactive respectively. Reactive techniques are always time consuming also mislay user satisfaction[12]. Proactive techniques are always preferable to avoid such complications. Workload has been predicted using CPU usage, Memory usage and Network usage.

- **CPU Load:**

The system will have many processes to execute in a single system. The task manager will show many variables like the CPU utilization of the resources, memory usage, etc. The CPU load plays the major role in the execution of entire system. Its utilization has great impact on the performance of a system.

CPU load is measured as $$\mathbf{CPU} = \sum_{k=1}^{n} \mathbf{(CU/V)}$$

where n = the number of nodes,

CU = CPU utilization

V = Total CPU's used

- **Memory Load**

Memory is another important part of the computer all the work by the client is stored and retrieved on and from the memory respectively. So without the memory usage the working process of the whole system become dazed.

Memory load is measured as

$$\mathbf{Memory} = \sum_{k=1}^{n} \mathbf{MEM\ USED/TOT\ MEM}$$

where MEM USED = memory used

TOT MEM = total memory

- **Network Load**

Network is the technique by which the systems are interconnected. Network systems can be loaded if there are too many systems in a condition to execute. The client has to consider the network bandwidth when the client is going to calculate the load.
Network load measured as

$$\mathbf{Network} = \sum_{k=1}^{n} \mathbf{NET\ BANDWIDTH\ /\ TOT\ BANDWIDTH}$$

where NET BANDWIDTH = network bandwidth used
TOT BANDWIDTH = total amount of network bandwidth.

F. Dynamic Normalized Algorithm

The two distinguished algorithm has proposed for the cloud service providers. The system is maintained on normalized mode on these two proposals [13]. The CPU usage has got by the performance counter. From the each IP address the CPU load has been calculated by the above algorithm[14,15].

Agent1:

Begin
While all the servers are in running state
Get CPU usage from each system

Calculate the CPU load

The CPU load has been analyzed for the different request

If (CPU load>=upper threshold)

Status is OVERLOAD

Else if (CPU load <Upper threshold && CPU load>normal value)

Status is NORMAL

Else

Status is UNDERFLOW

End

Agent 2:

If (Status==OVERFLOW)

Server has to Sleep for a while

Else

Server works in balanced mode

III. RESULTS AND DISCUSSION

A. Process Diagram

To make the system as in the normalized state, a check has to be performed[16]. The process of making the system as normalized is given in a flow diagram as follows

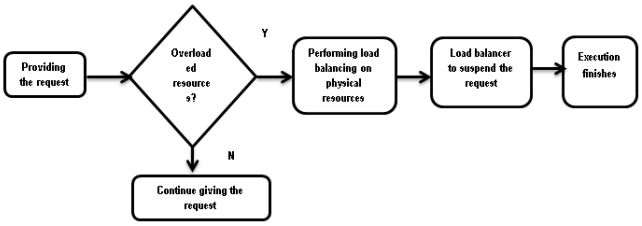

Figure 1. Process diagram

B. Analysis of The Load Balancer

The resources are analyzed in this paper and it is based on only one factor i.e. CPU performance [16]. The following performance chart is studied based upon the time and the CPU load. The load has been differentiated based on the above algorithm. This chart shows the CPU utilization of the various servers connected in the network.

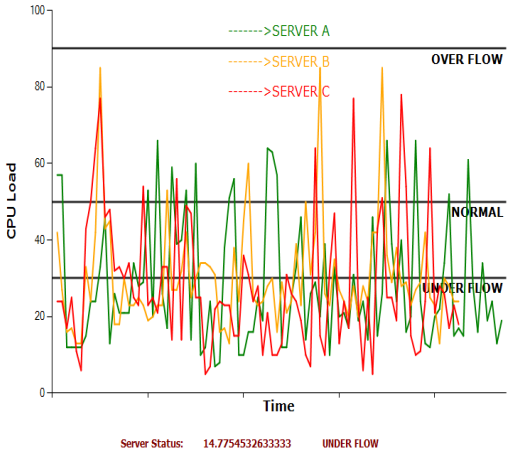

Figure 2. Chart for CPU load

IV. CONCLUSION

Load balancing is considered as the main factor in performance issues. Here, load balancing is analysed on the IaaS level. There may be different load factors the user can consider. In this paper, the CPU load and the way to normalize the execution of the CPU are measured. The whole implementation improves the efficiency of the distributed CPU's. Though it has good efficiency, the request has to wait for the overloaded server to become load balanced. Therefore, the task length and waiting time is increased. Here this concept is implemented only on CPU load; in future, the authors can try this algorithm using all the workloads of the system.

V. REFERENCES

[1] Mohd Hairy Mohamaddiah, Azizol Abdullah, Shamala Subramaniam, Masnida Hussin, "A Survey On Resource Allocation And Monitoring In Cloud Computing", International Journal Of Machine Learning And Computing, Vol4, No1, February 2014, DOI: 10.7763/IJMLC.2014.V4.382

[2] M.Kriushanth, DrLArockiam, "Load Balancer Behavior Identifier (Lobbi) For Dynamic Threshold Based Auto-Scaling In Cloud", ICCCI -2015, Jan08 – 10, 2015, Coimbatore, INDIA.

[3] Abhijit Aditya, Uddalak Chatterjee, Snehasis Gupta, "A Comparative Study Of Different Static And Dynamic Load Balancing Algorithm In Cloud Computing With Special Emphasis On Time Factor", International Journal Of Current Engineering And Technology E-ISSN 2277 – 4106, P-ISSN 2347 – 5161 ©2015.

[4] Po-Huei Liang1 And Jiann-Min Yang, "Evaluation Of Two-Level Global Load

Balancing Framework In Cloud Environment", International Journal Of Computer Science & Information Technology (IJCSIT) Vol 7, No 2, April 2015, DOI:10.5121/Ijcsit.2015.7201 1.

[5] Radha Ramani Malladi, "An Approach to Load Balancing In Cloud Computing", International Journal of Innovative Research in Science, Engineering and Technology, ISSN(Online): 2319 – 8753, Vol4, Issue 5, May 2015.

[6] Priyanka PKukade, Geetanjali Kale, "Survey Of Load Balancing And Scaling Approaches In Cloud", International Journal Of Emerging Trends & Technology In Computer Science (IJETTCS), Volume 4, Issue 1, January-February 2015 ISSN 2278-6856

[7] Weikun Wang, Giuliano Casale, "Evaluating Weighted Round Robin Load Balancing For Cloud Web Services".

[8] Jitender Grover, Shivangi Katiyar, "Agent Based Dynamic Load Balancing In Cloud Computing", https://www.researchgate.net/publication/271544765

[9] Aarti Singh, Dimple Juneja, Manisha Malhotra, "Autonomous Agent Based Load Balancing Algorithm In Cloud Computing", ICACTA 2015, Procedia Computer Science 45 (2015) 832 – 841

[10] Hardi SSanghavi, DrTejas PPatalia, "Load Balancing Algorithms For The Cloud Computing Environment: A Review", Journal Of Information, Knowledge And Research In Computer Engineering, ISSN: 0975 – 6760| Nov 14 To Oct 15| Volume – 03, Issue – 02 Page 591.

[11] Amal Zaouch, Faouzia Benabbou, "Load Balancing For Improved Quality Of Service In The Cloud", International Journal Of Advanced Computer Science And Applications, Vol6, No7, 2015

[12] Shikha Gupta, Suman Sanghwan, "Load Balancing In Cloud Computing: A Review", International Journal Of Science, Engineering And Technology Research (IJSETR), Volume 4, Issue 6, June 2015,

[13] MsShilpa D.More, ProfArti Mohanpurkar, "Load Balancing Strategy Based On Cloud Partitioning Concept", Multidisciplinary Journal Of Research In Engineering And Technology, Volume 2, Issue 2, Pg.424-431, ISSN: 2348-6953

[14] Smita Salunkhe, SSSannakki, "Load Rebalancing For Distributed File Systems In Clouds", IJAFRC, Volume 2, Special Issue - NCRTIT 2015, January 2015ISSN 2348 – 4853295 | © 2015

[15] Nadeem Shah, Mohammed Farik, "Static Load Balancing Algorithms In Cloud Computing: Challenges & Solutions", International Journal Of Scientific & Technology Research Volume 4, Issue 10, October 2015 ISSN: 2277-8616.

[16] Akshada Bhujbal, Prajakta Jakate, Manasi Wagh, Madhura Pise, Prof.M.V.Marathe, "Load Balancing Model in Cloud Computing", International Journal of Emerging Engineering Research and Technology Volume 3, Issue 2, February 2015, PP 1-6 ISSN 2349-4395 (Print) & ISSN 2349-4409 (Online).