International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 5 Issue 3, March-2016

An Updated Technique for Performing Efficient Load balancing in Cloud Computing Environment

Amit Garg

Prof. Kailash Patidar (HOD), Prof. Gaurav Kumar Saxena, Prof. Megha Jain

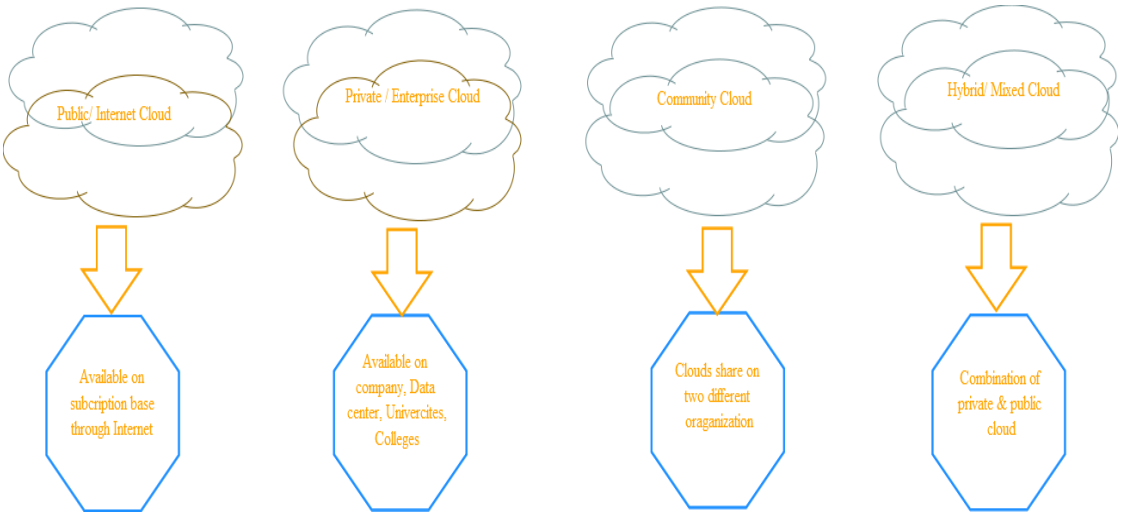
ABSTRACT

Load balancing in cloud computing systems is really a challenge now. A distributed solution is required. As it is not always practically feasible or cost efficient to maintain one or more idle services just as to fulfill the required demands. All jobs cant be assigned to appropriate servers and clients individually for efficient load balancing as cloud is a very complex structure and components are present throughout a wide spread area. In this paper we have proposed an updated cloud load balancing algorithm. The performance of proposed technique (updated throttled technique) is better as compared to the existing method.

INTRODUCTION

Cloud computing[1,3,4] has three basic models, which are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) . The Main advantages of cloud computing are : low cost, improved performance, infinite storage space etc .

TYPES OF CLOUDS



[2,5]

Figure : Types of Cloud

These are different types of clouds that can subscribe depend upon consumer need such as home user, small business owner, organization and universities need, on base of subscription base, consumer need cloud can be classified into

(i) Public Cloud

(ii) Private Cloud

(iii) Community Cloud

(iv) Hybrid or Mixed cloud as shown in above Figure .

Cloud Load Balancing Algorithms:

Round Robin[5,6]: Round robin performs the basic type of load balancing and functions simply by providing the list of IP address of cloudlet. It allocates first IP address to the first requester then second IP address to the second requestor for a fixed interval of time known as time slice. If the request is unable to finish within the given slice time, it will have to wait for the next cycle to get it turn for execution. This will continue till submitted tasks are not completed. Active Monitoring Load Balancer[7] This load balancer find outs the active VM and also to event out the active task at any point of time.

Throttled Load balancer This load balancing[8] technique ensures that only a per-defined number of internet cloudlets are allocated to a single VM at any point of time. If more groups are presents in the data center than the number of available VM than some of the requests have to be queued until the next VM is available.

The load balancing is the process of distributing the load among various resources in any system. Therefore load need to be distributed over the resources in cloud-based architecture so that each resources does approximately the equal amount of task at any point of time. The basic need is to provide some techniques to balance requests to provide the solution of the application faster. All cloud vendors are based on automatic load balancing services, it allow clients to increase the number of CPUs or memories for their resources to scale with increased demands. These services are optional and depends on the clients business needs. So the load balancing serves two important needs, firstly to promote availability of Cloud resources and secondarily to promote performance [8].

In order to balance the resources it is important to recognize a few major goals of load balancing algorithms:

a) Cost effectiveness: first aim is to achieve an overall improvement in system performance at a reasonable cost.
b) Scalability and flexibility: distributed system in which the algorithm is implemented may change in size or topology so the algorithm must be scalable and flexible enough to allow such changes to be handled easily.
c) Priority: scheduling of the resources or jobs need to be done on before hand through the algorithm itself for better service to the important or high prioritized jobs in spite of equal service provision for all the jobs regardless of their origin.

Proposed Load Balancing Algorithm:

Input:

- Data centre requests r1,r2,….., rn
- Available virtual machines vm1,vm2, ,vmn

Output:

- Data centre requests r1,r2,……,rn are allocated available virtual machines vm1,vm2,…….,vmn

Process:

1. The updated throttled algorithm maintains a hash map table of all the available virtual machines which their current state and the expected response time. This state may be available or busy. At the beginning, all the virtual machines are available.
2. When data centre controller receives a request then it forwards that request to the updated throttled load balancer. The update throttled load balancer is responsible for the virtual machine allocation. So that the job can be accomplished.
3. The updated throttled algorithm scans the hash map table. It checks the status of the available virtual machine.

3.1 If a virtual machine with least load and the minimum response time is found.

- Then the updated throttled algorithm sends the VM id of that machine to the data centre controller
- Data centre controller sends a request to that virtual machine

- Data centre controller sends a notification of this new allocation to the updated throttled
- The updated throttled algorithm updates the hash map index accordingly

3.2 If a virtual machine is not found then the updated throttled algorithm returns -1 to the data centre controller

4. When the virtual machine finishes the request.

    - The data centre controller sends a notification to updated throttled that the vm id has finished the request.
    - Updated throttled modifies the hash map table accordingly

5. If there are more requests then the data centre controller repeats step 3 for other virtual machines until the size of the hash map table is reached. Also of the size of hash map table is reached then the parsing starts with the first hash map index.

CONCLUSION

A number of load balancing algorithms existing which are distributing the load among the data center. Each of them has their own functionality. Load Balancer contains algorithms for mapping virtual machines onto physical machines in a cloud computing environment, for identifying the idle virtual machines and for migrating virtual machines to other physical nodes. In this paper we have proposed a novel Load Balancing approach. The performance of proposed load balancing method is better.

REFERENCES

[1]. John Harauz, Lorti M. Kaufinan. Bruce Potter, "Data Security in the World of Cloud Computing", IEEE Security & Privacy, Co published by the IEEE Computer and Reliability Societies, July/August 2009.
[2]. National Institute of Standards and Technology- Computer Security Resource Center -www.csrc.nist.gov
[3]. Singh A., Korupolu M. and Mohapatra D., ACM/IEEE conference on Supercomputing, 2008.
[4]. Stanojevic R. and Shorten R., IEEE ICC, 1-6, 2009.
[5]. Zhao Y. and Huang W., 5th International Joint Confer-ence on INC, IMS and IDC, 170-175, 2009.
[6]. Nae V., Prodan R. and Fahringer T., 11th IEEE/ACM International Conference on Grid Computing (Grid), 9-17, 2010.
[7]. Hu J., Gu J., Sun G. and Zhao T., 3rd International Symposium on Parallel Architectures, Algorithms and Programming, 89-96, 2010.
[8]. Bhadani A. and Chaudhary S., 3rd Annual ACM Bangalore Conference, 2010.