BI-QUEUE SCHEDULING FOR BROKERS IN CLOUD COMPUTING

Pooja Chandan[1], Ajay Kumar Dogra[2]
poojavikram@hotmail.com1, dogra2613@yahoo.com2
M.Tech Computer Science
Beant College of Engineeering and Technology, Gurdaspur, Punjab

Abstract: **Cloud Computing is emerging as an important player in the field of Information Technology. Scheduling is always being an important research domain to respond users requests Broker is the entity which deals with minimizing processing time of the requests and cost. Broker works on scheduling of the requests externally. Therefore in this paper we propose a frame work which utilizes queues to enhance efficiency of the broker by minimizing response time and cost. Further we compare our framework with existing round robin. Our results shows the effectiveness of the framework.**

Keyword: ***Scheduling, SLA, cost***

1. Introduction

Cloud computing is the delivery of computing as a service relatively a product, where mutual resources like information, infrastructure, platform, software are provided to users and other devices over a network. It is a rising field with a lot of hopes and possibilities in each and every sector whether it is corporate, IT, finance. The resources can be utilized in a proper way and this is a concept for reducing the cost involved in resources. Cloud computing allows user to access data, resources like hardware, software etc from remote locations. Everyone wants to keep their information in safe hands and to get the resources used in an efficient and effective manner. The time is not consequently very far-off when the overall working environment would shift to cloud. In such a case it becomes a critical need to analyze the cloud server whether it suits our need or not [1].

The concept of cloud computing has gained pace in the past few years. The customers of cloud services outsource the computational power and storage facilities to the public cloud providers on pay per usage basis [2][3]. In comparison to the existing computing models that has a dedicated infrastructure, cloud computing has the benefit of saving money and is reliable. The cloud customers need not to bother about the large cost of purchasing hardware, installing and to see through peak times of resource demand. The main advantage of cloud computing is the method of paying only the amount of service used and resources can also be scaled according to the need. There is no headache to the customer for the maintenance of servers and other components. The companies that are providing cloud services are Rackspace, Microsoft, Amazon, Google. Different providers offer a large variety of services which offers services at different price, different level of performance and different features. For example, a provider provides PaaS that is Platform as a Service in which a customer of cloud builds an application with the use of Application Programming Interface that is provided by the cloud, in the same way other provider avails IaaS that is an Infrastructure as a service where cloud customer can run application inside the virtual machines, with the use of Application Programming Interface which is made available by the other guest OS. Different providers also have different rates for the same services. For instance, the number and duration of virtual machine instances used by the customer is a way to calculate the price by Amazon's AWS, whereas the number of CPU cycles consumed by an application of customer is a way used by Google's App Engine. With a variety of cloud providers a question arises that which cloud provider works well? The answer to this question is both beneficial to customers and the cloud provider. For a prospective cloud customer, the answer to the above question helps in selecting the provider that suits best with the cost and the performance aspects. For example, customer may select a provider for storing of intensive applications, whereas a different provider of computing intensive applications. These answers to selection of cloud provider can help in the direction for improvements. For example, a provider needs to provide more resources for the optimization of the table storage in case the performance of the store is less than the competitors. Each provider has its own different personal way of performing and measurement of prices, thus makes it difficult for the customer to find out the best one. [2,3].

2. Related Work:

Francis et al. [1] outline previous contributions to the discussion of energy efficiency of cloud computing, provide a working definition of cloud computing and discuss its importance, which will grow as the technology matures and becomes well known. According to the author the assessment of the energy efficiency of cloud computing cannot be based only on data centers due to the importance of the intermediate communication networks that support the overall activity of providing cloud computing services and the devices used to access cloud services. The other components should be taken into account when measuring the energy efficiency of cloud computing. The energy consumption of mobile devices is generally good but the cellular and fixed communication networks that support cloud services have been noted to consume high amount of energy and this consumption is growing. There is the need to improve the energy efficiency of communication networks and the Internet in order to meet the new levels of demand. In this paper they analyzed the energy optimization of the network infrastructure should be paramount if the improved energy efficiency of data centers will result in overall benefits.

Keville et al. [2] examine the use of ARM-based clusters for low-power, high performance computing. This work examines two likely use-modes: (i) a standard dedicated cluster; and (ii) a cluster of pre-configured virtual machines in the cloud. A 40-node department-level cluster based on an ARM Cortex-A9 is compared against a similar cluster based on an Intel Core 2 Duo, in contrast to a recent similar study on just a 4-node cluster. For the NAS benchmarks on 32- node clusters, ARM was found to have a power efficiency ranging from 1.3 to 6.2 times greater than that of Intel. This is despite Intel's approximately five times greater performance. The particular efficiency ratio depends primarily on the size of the working set relative to L2 cache. In addition to energy efficient computing, this study also emphasizes fault tolerance: an important ingredient in high performance computing. In this paper they analyzed that it relies on two recent extensions to the DMTCP check point restart package. DMTCP was extended (i) to support ARM CPUs, and (ii) to support check pointing of the Qemu virtual machine in user-mode. DMTCP is used both to checkpoint native distributed applications, and to checkpoint a network of virtual machines.

Zhang et al. [3] propose a software and lightweight approach to accurately estimate the power usage of virtual machines and cloud servers. It explores hypervisor-observable performance metrics to build the power usage model. To configure cloud resources, it considers both the system power usage and the SLA requirements, and leverage learning techniques to achieve autonomic resource allocation and optimal power efficiency. In this paper they analyzed that it implements a prototype of the proposed power management system and test it on a cloud test bed.

Graubner et al. [4] a novel approach to virtual machine consolidation for saving energy is presented. In this paper they analyzed that it is based on energy efficient storage migration and live migration of virtual machines to take advantage of the lacking energy proportionality of commodity hardware.

Dharwar et al. [5] outlines the state-of-the-art in power-management technology on server hardware and describes how these raw features can be abstracted into a set of energy policies. In this paper they analyzed to explain how these policies or energy-profiles can be used to run a cloud datacenter energy efficiently. Further, this work also highlights some of the challenges involved in running cloud infrastructures in the emerging markets optimally despite some unique energy constraints.

Ye et al. [6] focuses on the live migration strategy of multiple virtual machines with different resource reservation methods. It first describe the live migration framework of multiple virtual machines with resource reservation technology. Then it perform a series of experiments to investigate the impacts of different resource reservation methods on the performance of live migration in both source machine and target machine. Additionally, it analyze the efficiency of parallel migration strategy and workload-aware migration strategy. In this paper they analyzed the metrics such as downtime, total migration time, and workload performance overheads are measured.

Kejiang et al. [7] present a virtual machine based energy-efficient data center architecture for cloud computing. It investigates the potential performance overheads caused by server consolidation and live migration of virtual machine technology. In this paper they analyzed the experimental results show that both the two technologies can effectively implement energy-saving goals with little performance overheads. Efficient consolidation and migration strategies can improve the energy efficiency.

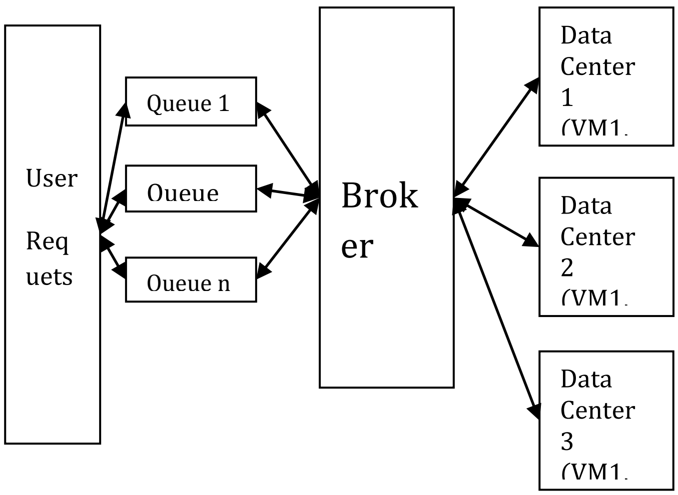
3. System Model

In this framework we can see that the user's requests are pre-analyzed according to their weights of requests and being submitted to the queue. In the queue broker manages the scheduling of different types of requests and forwarded the requests to the datacenter which contains virtual machine for further processing. Figure 1 shows the basic implementation of the model and the requests response path. The requests are given to broker which further schedule the requests by using following algorithm:

Algorithm :

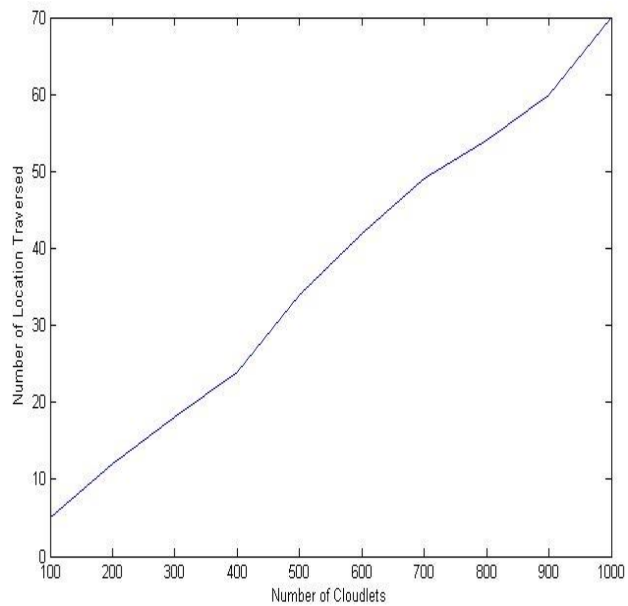
1. The entire request and the resources are analyzed before execution.
2. All the tasks "N" are put into buffer for processing.
3. These tasks are divided into different types of classes based on their size, followed by M number of sets.
4. The tasks are lined up in their own sets according to the emergency degree more emergency task is put in the front of the queue.

5. Based on cost of the requests these categories are decided and being processed.

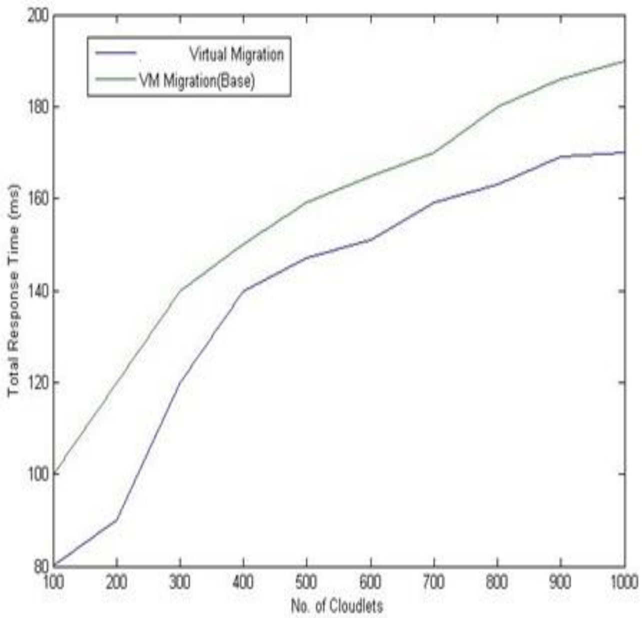


4. Experimental Setup:

In the simulation work we have used CloudSim as simulator and implement the proposed algorithm. Following parameters are used. This work considers various Datacenter, Virtual Machines (VM), host and cloudlet components from CloudSim for execution analysis of the algorithms. The simulation also comprises of different requests by the user to be processed. There are in total 1000 cloudlets or requests in total which are ranging from 300 MB to 23000 MB. In this work we are comparing our algorithm with Service Level Agreement (SLA) based scheduling. In SLA based algorithm we can see that databases are used in keeping the results for frequent uses and for future query. In the results we can see that the number of location requests is traversed. By traversing location we mean the location of the users requests which is being migrated to one place to another. Therefore more the traversing more time consumption.



In the following diagram we have observed that virtual machine migration in base paper and virtual machine migration in our algorithm determines that with the increase in the migration and databases uses the total response time of the cloudlets increases. But our algorithm uses queues instead of databases which reduces the total response time dramatically by 20%.



5. Conclusion and Future work

In this paper we have evaluated the response time and cost of the proposed frame work and found that the efficiency of the proposed framework is far superior to base paper. Further we can add the parameter of SLA violation and will try to minimize it. The effect of minimization could be very crucial in overall performance increasing and managing virtual machines.

References:

1.http://www.thecloudtutorial.com/related.html

2.wikipedia

3. Sahil Vashist, Rajwinder Singh "Energy Cost and QoS based management of
Data Centers Resources in Cloud Computing" in International Journal of Computer Science & Engineering Technology (IJCSET), ISSN : 2229-3345, Vol. 4 No., 06 Jun 2013

4. http://en.wikipedia.org/wiki/Service-level_agreement

5.http://searchitchannel.techtarget.com/definition/service-level-agreement

6. L.Kapoor, S.Bawa, A.Gupta, "Hierarchical chord-based resource discovery in Intercloud Environment", Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility, pp: 464-469

7 Yogita Chawla1, Mansi Bhonsle "Dynamically optimized cost based task
scheduling in Cloud Computing" in International Journal of Emerging & Technology in Computer Science, Volume 2, Issue 3, May – June 2013