Performance Evaluation of Social Networking Application with different Load balancing policy across Virtual Machine in a Single Data Center using CloudAnalyst

Pradeep Singh Rawat
Computer Science & Engg. Department
Dehradun Institute of Technology
Dehradun, India
pradeep_kec09@yahoo.com

G. P. Saroha
Information Technology Department
Dehradun Institute of Technology
Dehradun, India
gp.saroha@gmail.com

Varun Barthwal
Information Technology Department
H N B Garhwal University
Srinagar, India
varuncsed1@gmail.com

***Abstract*— To improve the IT resource utilization using Internet base computing we use virtualization technology. Virtualization put the foundation of cloud computing. To deploy the Cloudlet in Cloud computing environment we need to select the Load balancing policy across virtual machine. Virtual machine is hosted on physical hardware inside the Data Center. Hence we want to test the comparative performance of cloud based application. We use different load balancing policy across the virtual machine. Using cloud test bed we identify the performance for different simulation configuration scenario. It includes the variable load balancing policy across the virtual machine deployed on host inside the data center.**

Keywords*— *CloudAnalyst, MIPS, gridlet, VM, Data center, Simulation, DC1, IaaS, PaaS, SaaS;

I. INTRODUCTION

Cloud computing is the Internet based computing in which all computational operation is made to be performed over the cloud. We know that for resource management more cost need to be pay. So it is better to use the resources on rent basis rather than to buy our own resources. Each organization wants to make busy their employee for innovation and high quality resource utilization. Cloud Computing is improving efficiency of IT infrastructure utilization. Simplest definition of cloud computing is "To provide IT as a service" is called cloud computing. This is the part of distributed computation. The main component of IT is hardware, software (application, system) are provided as a service by the cloud Computing. While using cloud computing cloud vendors can provides the secure pool of resources which include the storage and computing server with in the data center. It provides the massive distributed environment which may dynamic in nature. To control this type of distributed system we need to study some simulation Tool. Simulation tool which are used for distributed application based on object oriented programming. Simjava, Grdilet, Cloudsim, CloudAnalyst are the cloud simulation tools. These Tools provide the clue to us how to deploy application and what are the IT requirements for the application. These tools follow the layered architecture i.e. user can add their own layer over the user code level. Simulation tool provides the prior information about cloud resources which are required for application deployment. We can use our own policy at data center level to share the MIPS of the physical processing element. Using simulation tool we can setup the different cloud configuration with internet characteristics. Processing power of the CPU to run their application is provisioned in time and space shared mode. We take an example of social networking application to deploy at different region with different internet characteristics, data center configuration.

II. RELATED WORK

Distributed system consist a collection of inter connected and virtualized computers. These are dynamically provisioned and presented as one or more unified computing resources. Service level agreement of cloud computing resources depends on negotiation between service provider and consumers [1]. Cloud broker acts as a mediator. Resources are provided on demand by cloud provider as per the requirement of cloud application. For an example Amazon EC2 provides the IaaS cloud computing service model to run the cloud based application [10]. Consumers can use the platform to develop the application i.e. we follow here the cloud service model i.e. PaaS E.g. Google Application Engine provides the platform to run the Gmail application [11]. In similar way to provide the software as service we use cloud model SaaS. Hence end user of cloud need not focus on management of cloud resource but to develop the innovative application. Experimentation methodologies in cloud environment help the researcher to deploy the application in a real platform with minimum cost and good server utilization. Simulation is one of such alternative and this is the focus of this work. There are many simulation tools on which the cloudsim is based supports to test the distributed application. These Tools

are helpful for the object oriented modeling of cloud resources. Some of these simulators are GridSim [3], MicroGrid [4], GangSim [15], SimGrid [5] and CloudSim [6]. To study the grid computing system first three simulators are used. CloudAnalyst toolkit layered over Cloudsim is helpful for modeling and simulation of cloud computing system. We can estimate the cost for following pay as you go model for distributed application running over cloud infrastructures [12][13]. GridSim Toolkit can be used for performance evaluation of distributed environment and simulation of dedicated link based network. GridSim toolkit is a Java-based simulation toolkit that supports modeling and simulation of heterogeneous Grid resources. It allows users spread across multiple organizations to use their own policies for scheduling applications.

It supports multiple application models. It provides primitives for creation of application tasks, mapping of tasks to resources, and managing of tasks and resources [3]. Cloudsim Toolkit provides the capability to setup the cloud environment and run cloud task on it. It can model the powerful cloud resource to provide IT solution for huge cost implications of IT infrastructure. CloudSim provides the virtualization engine to manage the entire life cycle of virtual machine. CloudSim Toolkit (follow the layered architecture) is based on distributed simulator GridSim. Datacenter configuration feature in CloudAnalyst is extended from Cloudsim toolkit. At infrastructure level cloud main resource datacenter can be model to provide the storage and computing server for cloud task run. Datacenter serve the request to user base located in different geographical region. CloudAnalyst toolkit features are presented in the next section.

III. CLOUDANALYST

Clouds make deployment of large scale applications easier and cheaper. It also creates new issues for cloud base application developers. Cloud infrastructures are distributed among the cloud user. Applications can be deployed in different geographic locations. The selected distribution of the application impacts its performance for users which are far from the data center. Internet applications are accessed by users around the world. Popularity of application and usage pattern of application varies around the world. To quantify impact of number of simultaneous users in different geographic location and network in application is hard to achieve in real test beds. Presence of elements cannot be predicted nor controlled by developers. Therefore, other methodologies that allow quantification of such parameters must be used. To allow control and repeatability of experiments, simulators such as CloudAnalyst, CloudSim are used. Simulation experiments apply models of both applications and infrastructures [8]. So simulation requires some effort. Using CloudAnalyst toolkit we evaluate the performance of cloud based applications like social networking application. We used application statistics to get the simulation results which help in quality of service improvement. Response time and data center processing time act as a performance evaluation parameter. It is helpful for the researcher to identify the service broker policy in optimal simulation configuration. Load balancing policy for high quality of service can be identify using simulation results.

IV. A CASE STUDY

A. SIMULATION OF A LARGE SCALE SOCIAL NETWORKING APPLICATION

A typical large scale application on the Internet can benefit from Cloud technology. These applications present non-uniform usage patterns. Access to such services varies along the time of the day. Geographic location from sources of service requests also varies. A new functionality in the service may cause a sudden increase in number of user pattern which want to access the service. Leading to an increase in number of requests arriving to servers need load balancing operation. Cloud allows infrastructures to dynamically react while increase in requests. Increasing application resources, reducing available resources occurs dynamically when the number of requests increase, reduces. Service level aggerement between Cloud providers and consumers are met with a minimal cost for consumers. One well-known social networking site is Facebook [9], which had 200 million registered users over worldwide. On 18/06/2009 the approximate distribution of the Facebook user base across the globe was the following: North America: 80 million of users, South America: 20 million of users, Europe: 60 million of users, Asia: 27 million of users, Africa: 5 million of users, and Oceania: 8 million of users [2]. In this case study we model the behavior of this social network application. We used Cloud Analyst to evaluate performance related to use of Clouds which host such an application. We identify that which Load balancing policies across VM in different simulation configuration scenario provide the high quality of service. We consider simulation configuration scenario with more than one cloud resource, 6 user base located in different geographical regions. This will be helpful for researcher and cloud user in real deployment of cloud base applications. We can observe the behavior pattern of application using simulation results.

V. SIMULATION CONFIGURATION PARAMETERS

For Simulation of large scale social networking application we need to configure the simulation parameters. Simulation configuration includes the advance configuration, application deployment configuration, and physical hardware configuration i.e. host configuration at data center. All the parametric values are shown in table 1, 2, 3 respectively.

Table 1. Advanced configuration parameters

User grouping factor	Request grouping factor	Executable instruction length per request(Bytes)
1000	100	250

Advanced parameters of simulation configuration include the user grouping factor i.e. number of simultaneous user from a single user base, request grouping factor i.e. number of simultaneous request supported by single application server instance and size of instruction in Bytes corresponding to each request shown in table 1.

Table 2. Application deployment configuration

Service Broker policy	Data Center	Number of VM	Image size (Mb)	Memory (Mb)	Band width (Mbps)
Optimize response time	DC1	25	100000	1024	1000
	DC2	50	100000	2048	1000

To deploy the application on cloud we need set the parameter within the simulation configuration, which includes the service broker policy and virtual machine properties. It also includes the identification of cloud main resource, data center. All these information are shown in above table 2.

Table 3. Host (Physical hardware) detail of datacenter DC1 and DC2

I.D.	Memory (Mb)	Storage (Mb)	Band Width (Mbps)	VM policy	Processor speed (MIPS)
0	2048	100000	10000	Time shared	10000
1	2048	1000000	1000000	Time shared	10000

Simulation configuration includes the Datacenter configuration where we set the parameters of Infrastructure. The hardware details at datacenter, Host configuration include the host properties with VM allocation policy as shown in above table 3.

VI. SIMULATION RESULTS

We got the simulation results shown in table 4 corresponding to the simulation configuration parameters shown in the section Vth. Simulation results are corresponding to simulation configuration scenario with variable load balancing policy across the VM in Cloud resource. Virtual machines are deployed over the host at data center. Optimal results are provided by the second simulation configuration with third load balancing configuration. Throttled policy provides the optimal results for response time and data center processing time. Simulation results are obtained for different load balancing policy. Data center with single host and multiple virtual machines are hosted in different regions.

Table 4. Simulation results

Simulation configuration Scenario	Load balancing configure.	Load balancing policy across vm in a single Data Center	Overall average response time (ms)	Average Data Center processing time (ms)
1 data center with 25 VMs and 6 user base	1	Round Robin	1057.72	759.57
	2	Equally spread current execution load	1058.50	760.39
	3	Throttled	672.20	379.56
2 data center with 25, 50 Vm and 6 user base	1	Round Robin	998.77	702.44
	2	Equally spread current execution load	1001.39	704.86
	3	Throttled	643.79	352.88

VII. COMPARISON OF SIMULATION RESULTS FOR TWO CLOUD SCENARIOS

Following figures depicts variation of response time for different simulation configuration scenarios with different Load balancing configuration. We got the simulation results for two simulation configuration scenario with different number of data center. In case of first scenario minimum response time is achieved at configuration corresponding to the throttling as a load balancing policy across virtual machine in each cloud resource. This policy acts analogous to the accelerator of a car which speeds up the car while pressing it. From this result it is quite clear that for high quality of service we should deploy our application at the data center where this policy is followed across the virtual machine. To improve the response time for this load balancing policy we configured more number of datacenter in different regions. As shown in the Table 4 response time and processing time is minimized in simulation configuration scenario with number of data center closer to the user base, number virtual machine increased. Index table of virtual machine is used. When data center controller put the request to run the cloudlet on virtual machine, unique id of the virtual machine is returned to allocate the cloudlet run. In this analysis tool we used time shared policy for core provisioning of cloud resources.

Data center broker take the responsibility to deal with entire life cycle of Virtual machine. Index table is also maintained for the status of virtual machine that may be busy or available. At the starting of cloud initialization all VM's are available. Hence from graphical results described below it is clear that we get high quality of service when virtual machine load balancing policy is throttled, cloud application service broker policy is optimize response time and virtual machine deployed at the data center are in bundled from. We get minimum response time, data center processing time for simulation configuration scenario with two data center, 6 user base and 25, 50 virtual machines. Virtual machines are deployed on host of dc1 and dc2 respectively. In this scenario third load balancing configuration is used with fixed service

broker policy. Hence Throttled load balancing policy with second simulation configuration scenario is the choice for good quality of service and real deployment of application. As shown in Figure 1, 2 we compared the results of two different simulation configuration scenarios.

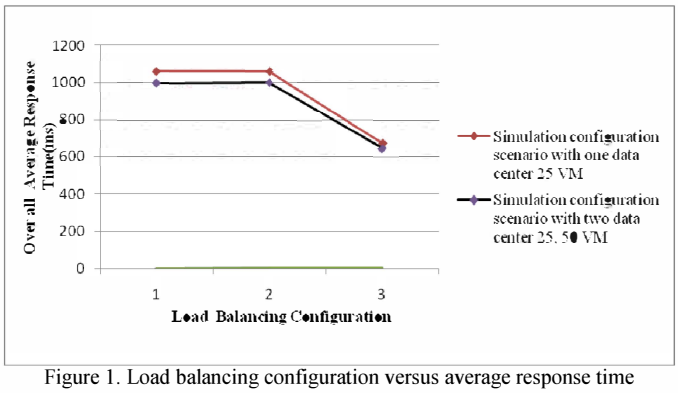

Figure 1. Load balancing configuration versus average response time

Performance can be measured by the parameters response time and data center processing time. We got minimum response time in scenario with two data center having single host and different number of virtual machine. Along with simulation configuration scenario second the Load balancing configuration 3 provides the minimum value of response time, data center processing time. Cloud scenario with two data center, one host each and different number of virtual machine ensure the quality of service for third load balancing configuration.

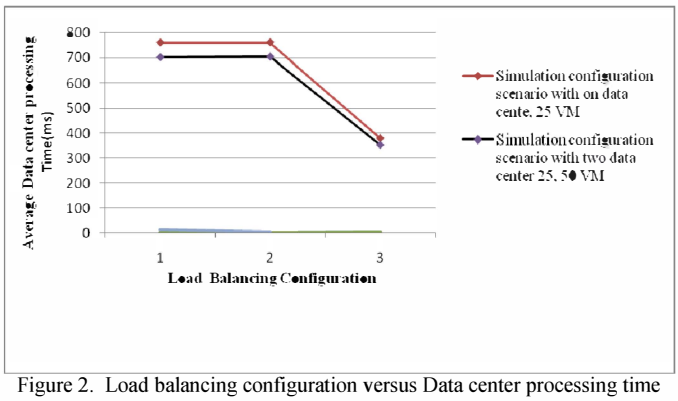

Figure 2. Load balancing configuration versus Data center processing time

For deployment of application on real cloud Second simulation configuration scenario with third load balancing configuration should be the priority for allocation of cloud tasks to the virtual machines.

VIII. CONCLUSION

Relative performance of cloud application can be test for different load balancing algorithm using cloud simulation Tool. CloudAnalyst tool provide the comparative results for simulation configuration scenarios. To deploy the application on real cloud we can identify best Load balancing policy across VM in a single Data Center using simulation results. The value of response time, Datacenter processing time is Minimum in case of Throttled load balancing policy with optimal simulation scenario. Simulation results using this policy can assure the resource utilization in optimal simulation scenario with fixed simulation configuration parameters. Hence in cloud configuration with cloud resource located in different regions, throttled load balancing policy can assure the high quality of service than equally spread current execution load, Round Robin load balancing policy. We got the simulation results for social networking application in which user request is model by the request grouping factor with well defined instruction length per requests. We got the satisfactory simulation results for simulation configuration with number of cloud resources closer to user bases. Simulation results help us in real deployment using real cloud. We demonstrated how CloudAnalyst can be used to model and evaluate a real world problem through a case study of a social networking application. We have illustrated how the simulator can be used to effectively identify best load balancing policy across the VM in Data center for overall usage patterns around the world.

REFERENCES

[1] R. Buyya, C. S. Yeo, and S. Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities", Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC 2008, IEEE CS Press, Los Alamitos, CA, USA), Sept. 25-27, 2008, Dalian, China.

[2] Bhathiya Wickremasinghe, Rodrigo N. Calheiros, and Raj kumar Buyya, "A CloudSim-based Visual Modeler for Analyzing Cloud Computing Environments and Applications," Proceedings of the IEEE 24th International Conference on Advanced Networking and Applications, Sept. 2010.

[3] R. Buyya, and M. Murshed, "GridSim: a toolkit for the modeling and simulation of distributed resource management and scheduling for Grid computing," Concurrency and Computation: Practice and Experience, 14(13): 1175-1220, Nov. 2002.

[4] L. X. Song H, Jakobsen D, Bhagwan R, Zhang X, Taura K, A. Chien, "The MicroGrid: A scientific tool for modeling computational Grids," Proc. of the ACM/IEEE Supercomputing Conference, IEEE Computer Society, Nov. 2001.

[5] A. Legrand, L. Marchal, and H. Casanova, "Scheduling distributed applications: the SimGrid simulation framework," Proc. of the 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid 07), May 2001, pp. 138-145.

[6] R. Buyya, R. Ranjan, and R. N. Calheiros, "Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities," Proc. of the 7thHigh Performance Computing and Simulation Conference (HPCS09), IEEE Computer Society, June 2009.

[7] F. Howell and R. Macnab, "SimJava: a discrete event simulation library for Java," Proc. of the 1st International Conference on Web based Modeling and Simulation, SCS, Jan. 2008.

[8] J. Gustedt, E. Jeannot, and Martin Quinson, "Experimental methodologies for large-scale systems: a survey," Parallel Processing Letters, vol. 19, Sep. 2009, pp. 399-418.

[9] "Facebook," http://www.facebook.com.

[10] "Amazon Elastic Compute Cloud (Amazon EC2),"http://aws.amazon.com/ec2/

[11] "Google App Engine," http://code.google.com/appengine/

[12] E. Deelman, G. Singh, M. Livny, B. Berriman, and J. Good, "The cost of doing science on the Cloud: the Montage example," Proc. of the 2008 ACM/IEEE Conference on Supercomputing, IEEE, Nov. 2008.

[13] M. Assunção, A. di Costanzo, and R. Buyya, "Evaluating the Cost-Computing to Extend the Capacity of Clusters Benefit of Using Cloud," Proc. of the 18th International Symposium on High Performance Distributed Computing, ACM Press, June 2009.

[14] C. Vecchiola, S. Pandey, and R. Buyya, High-Performance Cloud Computing: A View of Scientific Applications, Proc. Of the 10th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN 2009), Kaohsiung, Taiwan, Dec. 2009.

[15] C. Dumitrescu, and I. Foster. "GangSim: a simulator for grid scheduling studies," Proc. of the 5th International Symposium on Cluster Computing and the Grid (CCGrid 05), IEEE Computer Society, May 2005.

[16] A. Weiss, "Computing in the Clouds," net Worker, vol. 11, Dec.2007, pp. 16-25.