Cloudsynth: A Tool for Modelling and Simulation of Load Balancing Techniques in Cloud Computing

R.Vinodh[1], Swagath[2],Tejas K.S[3], Vijay .C[4], Manjushree .R. Patel[5]
UG Student[1, 2, 3, 4], Associate professor[5]
Department of CSE
East West Institute of Technology, Bangalore, India

Abstract:
In computing, load balancing improves the distribution of workloads across multiple computing resources, such as computers, a computer cluster, network links, central processing units, or disk drives.In Round Robin technique in a circular way method, we try to clarify the load balancing scenario of a cloud server during its execution. It will help to get an effective and fast execution environment of task assigned by the user which helps in to create an effective communication framework between broker and virtual machine (VM) to optimize the time and minimize the cost. With the advent of the Cloud, deployment and hosting became cheaper because several Cloud providers are available, each one offering different pricing models and located in different geographic regions, a new concern of application developers is selecting providers and data centre locations for applications.

I. INTRODUCTION

Cloud computing is a fast growing area in computing research and industry today. It has the potential to make the not so new idea of 'computing as a utility' a reality in the near future. With the advancement of the cloud, there are new possibilities opening-up on how applications can be built on the internet. On one hand, there are the cloud service providers who are willing to provide large scaled computing infrastructure at a cheaper price which is often defined on usage, eliminating the high initial cost of setting up an application deployment environment and provide the infrastructure services in a very flexible manner which the users can scale up or down at will. On the other hand there are large scaled software systems such as social networking sites and e-commerce applications gaining popularity today which can benefit greatly by using such cloud services to minimize cost and improve service quality to the end users. But when brining these two ends together there are several factors that will impact the net benefit such as the distribution (geographic) of the user bases, the available internet infrastructure within those geographic areas, the dynamic nature of the usage patterns of the user base and how well the cloud services can adapt or dynamically re-configure itself, etc. Doing a comprehensive study on this overall problem in the real world will be extremely difficult and the best approach to study such a dynamic and massively distributed environment is through simulation. Therefore the aims of the project can be summarized as below:
Investigate into existing simulation techniques for studying large scale distributed systems.

• Leverage suitable existing simulation techniques and tools and define an approach for effectively simulating scaled internet applications on cloud environment.

• Develop the initial version of the tool required for such an approach with a sufficiently flexible design that leaves room for further refinement and extension.

• Perform initial experiments using the tools and approach.
This paper continues as takes after. In Section 2 introduce the investigation of a few existing frameworks with its use-fulness. In Section 3 displays the proposed demonstrate. Segment 4

examines the working of proposed model. In Section 5 how to setup the tool 6 the results of the pro-posed framework is talked about. Lastly the conclusion and future work of the project.

II. EXISTING SYSTEM

Doing a comprehensive study on this overall problem in the real world will be extremely difficult, and the best approach to study such a dynamic and massively distributed environment is through simulation. GridSim toolkit to address the problem of near impossibility of performance evaluation of real large scaled distributed environments (typically grid systems but also P2P networks) in a repeatable and controlled manner. The GridSim toolkit is a Java based simulation toolkit that supports modeling and simulation of heterogeneous grid resources and users spread across multiple organizations with their own policies. It supports multiple application models and provides primitives for creation of application tasks, mapping of tasks to resources and managing such tasks and resources.

2.1 Problem Statement
Advances in Cloud computing opens up many new possibili-ties for Internet applications developers. Previously, a main concern of Internet applications developers was deployment and hosting of applications, because it required acquisition of a server with a fixed capacity able to handle the expected application peak demand and the installation and maintenance of the whole software infrastructure of the platform supporting the application..

III. PROPOSED SYSTEM

It is developed to simulate large-scale Cloud applications with the purpose of studying the behaviour of such applications under various deployment configurations and load balancing algorithms. **"CloudSynth"** helps developers with insights in how to distribute applications among Cloud infrastructures and value added services such as optimization of applications performance and providers incoming with the use of Service Brokers.Technology stack that shall be used be using **Core Java**

8 and GUI with **JavaFX**, running as a desktop application.The Tool shall use 'CloudSim' and 'GridSim' as a Library to perform backend operations.
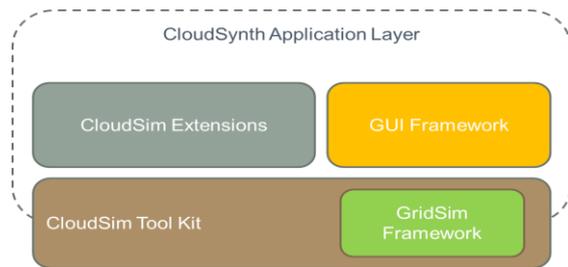

Figure.1. Cloudsynth tool

For simulation application developement, CloudSim Tool kit framework is used which internally uses GridSim framework. Extensions for the said load balancing techniques are provided by CloudSim extensions module.The business logic and interaction with respect to GUI(graphical user interface) is handled by CloudSynth Application layer using UI framework. The following section describes these components and concepts in depth. The three main extensions introduced in CloudSynth to CloudSim toolkit are the User Base, Data Center Controller and the Internet. But before explaining those it is appropriate to introduce the concept of "region".

IV. SYSTEM ARCHITECTURE

The Data Center Controller uses a VM Load Balancer to determine which VM should be assigned the next Cloudlet for processing. Currently there are three VM Load Balancers implementing three load balancing policies which can be selected as required by the modeller. Once the tool is initialized , the configuration parameters are ready to setup .There after we set a simulation time and start the simulation. For the valid configuration , simulation executes and obtains the response time and generates report. The data centers use the VM Load Balancer to load balance requests between the available virtual machines, and in the current version there are 3 variants.
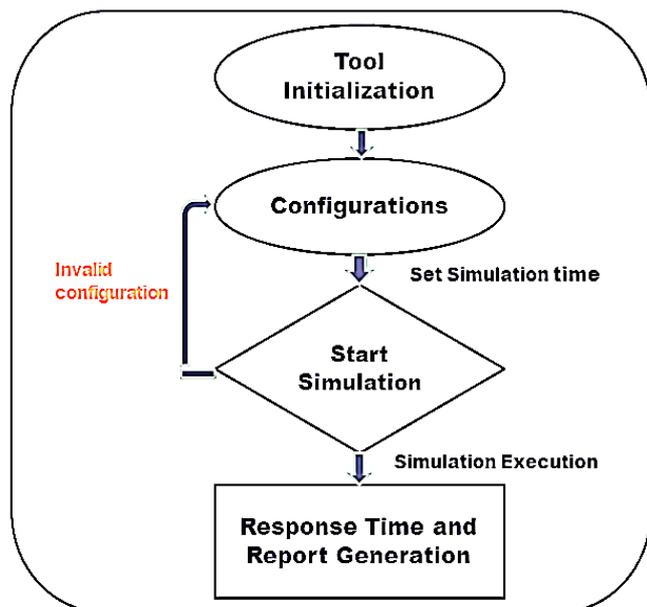
Figure. 2. DataFlow of System Architecture

4.1 System Modules

System design is the process of defining the architecture, components, modules, interfaces and data for a system to satisfy specified requirements. One could see it as the application of

systems theory to product development. There is some overlap with the disciplines of systems analysis, systems architecture and systems engineering. If the broader topic of product development "blends the perspective of marketing, design, and manufacturing into a single approach to product development," then design is the act of taking the marketing information and creating the design of the product to be manufactured. System design is therefore the process of defining and developing systems to satisfy specified requirements of the user.

The main four Modules are
- Application Module
- CloudSim Extensions module
- CloudSim Framework module
- GridSim Framework module

4.1.1 Application Module
Business logic for GUI framework.
Populating and displaying the data from Framework.

4.1.2 CloudSim Extension Module
Extensions Handlers which is exposed on the usage of application.
Communicates with Framework.

4.1.3 CloudSim Framework Module
Configuration related modules sent by extensions.
Uses the balancing techniques for simulation.

4.1.1 GridSim Framework Module
Performs Simulation
Response time is evaluated for estimated simulation time.

V. USING CLOUDSYNTH

To set up a simulation you need to carry out the following steps. (Please note the screens mentioned here are explained in detail in the next section).

1. Define user bases–Using User Base entities define the users of the application, their geographic distribution, and other properties such as the number of users, the frequency of usage and the pattern of usage such as peak hours. This is done in the Main tab of the Configure Simulation screen.

2. Define data centers–Using the Data Centers tab of the Configuration screen define the data centers you wish to use in the simulation. Define all the hardware and accounting aspects of the data centers here.

3. Allocate Virtual Machines for the application in Data Centers–Once the data centers have been created, you need to allocate virtual machines in them for the simulated application using the Main tab of the Configurations screen. A data center defined in step 2 above does not get included in the simulation unless it is allocated in this step. You can allocate multiple types of virtual machines in the same data center during this step.

4. Review and adjust the advanced parameters in the advanced tab of the configuration Screen.

5. Review and adjust the network latency and bandwidth matrices on the Internet characteristics screen.
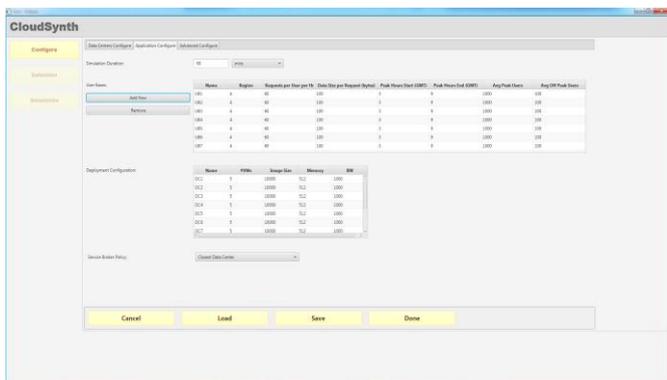
Figure. 3. CloudSynth Launch Screen

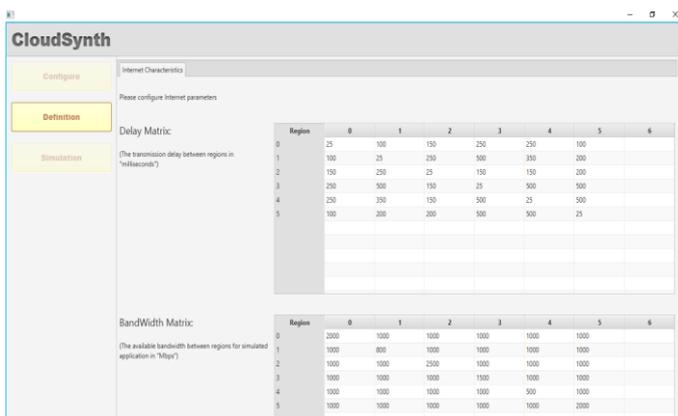**Figure.4. Configuration Simulation Panel**

Figure.5. Internet Characterstics screen

The proposed framework can be executed using eclispse any version. A typical large scaled application type on the Internet today that could benefit from cloud is social networking applications. E.g. Facebook, one of the most popular social networking sites has over 200 million registered users worldwide. The working of the tool is as per the following: When CloudSynth is started the first screen displayed is the main screen. It has the simulation panel with a map of the world on the right and the main control panel on the left

Configure Simulation screen has three tabs.

- Main Tab
- Data Center Tab
- Advanced Tab

The configuration options on the main tab are:

1. Simulation time – the duration of the simulation which can be given in minutes, hours or days.

2. User Bases Table–This is a table listing out all the user bases in the simulation.

3. The Add and Remove buttons next to the table can be used to add or remove user bases from the configuration.

4. Application Deployment Configuration–along with the details of a virtual machine. The fields are:

- Data Center–This is a drop down listing the names of data centers created in the Data Center tab.
- Number of VMs–How many VMs to be allocated to the application from the selected data center.
- Image Size–a single VM image size in bytes.
- Memory– amount of memory available to a single VM.
- BW–amount of bandwidth available to a single VM.

5. The data center tab allows you to define the configuration of a data center.The table at the top lists the data centers and using the Add/Remove buttons you can add or remove data centers to the configuration.

6. User Grouping Factor (in User Bases) – This parameter tells the simulator how many users should be treated as a single bundle for traffic generation.

7. Request Grouping Factor (in Data Centers) –This parameter tells the simulator how many requests should be treated as a single unit for processing.

8. Load balancing policy–the load balancing policy used by all data centers in allocating requests to virtual machines. Available policies are:

- Round-robin
- Equally Spread Current Execution Load–The load balancer keeps track of how many Cloudlets are cur-rently being processed by each VM and tries to even out the active load.
- Throttled– The load balancer throttles the number of requests assigned to a single VM.

9. Once the above configurations have been used to successfully create a simulation configuration, the user has to go back to the main screen and execute the simulation by selecting the "Run Simulation" from the control panel. This will start the simulation and the progress bar at the top of the simulation panel shows the percentage completion of the simulation. The simulation screen will display a simple animation showing which user bases are sending messages to which data centers.

Figure.6. Graphical Simultion Output

VI. RESULTS

Once the simulation is completed the main response times will be displayed on the simulation panel next to each user base. The detailed results can be viewed by clicking the "View Detailed Results" button that appears at the right hand bottom corner of the screen after the simulation has completed. The results screen will list out the data collected from the simulation.

This includes:

- Overall response time summary (for all the user bases).
- Response time by user base in tabular format.
- Response time by user base in graphical format broken down into the 24 hours of the day.
- Request servicing time by each data center in tabular format.
- Request servicing time by data center in graphical format broken down into 24 hours of the day.
- Data center loading (number of requests serviced) in graphical format broken down in to 24 hours of the day.
- Cost details. preferred.

Overall Time Response Summary

	Average (ms)	Minimum (ms)	Maximum (ms)
Overall Response Time:	499.92	375.14	622.64
Data Center Processing Time:	0.35	0.02	0.73

Response Time by Region

UserBase	Average (ms)	Minimum (ms)	Maximum (ms)
UB10	498.37	375.14	620.11
UB11	501.47	402.62	615.12
UB12	498.56	390.12	600.11
UB1	499.16	382.62	615.14
UB2	501.77	392.61	620.12
UB3	500.16	410.12	612.61
UB4	499.35	395.11	600.14

Cost

Total Virtual Machine Cost : $1.00
Total Data Transfer Cost : $1.59
Grand Total : $2.60

Data Center	VM Cost	Data Tran...	Total
DC2	0.50	0.45	0.95
DC1	0.50	1.15	1.65

Figure.7. Overall Response & cost Summary

VII. CONCLUSION

A typical type of Internet application that could benefit from the flexibility of Cloud type services is social networking. Currently there are several simulation frameworks that can be extended and used to model this type of a problem, but there is no user friendly tool that can be used concentrating on the modeling effort rather than on the programming technicalities when using such frameworks.In the first part of this report we described the design of the CloudSynth in detail and explained the various algorithms used in different scenarios.Then in the second part, by applying the simulator to model a typical social networking type application with high usage loads, we demonstrated how different aspects of the operation of such an application can be studied using the CloudSynth.

VIII. FUTURE WORK

CloudSynth is only a first step in a process that is expected to grow over a period of time. Following are some of the immediate next steps that should be adopted.

1. Resolve issues with CloudSim toolkit in dynamic reconfigu-ration and incorporate that to the simulator to study the dy-namic behaviour.
2. Explore the VM load balancing and service load balancing (service brokerage) algorithms.
3. Incorporate failure handling mechanisms in to the simulation
4. Improve the GUI for usability.
5. Improve the simulation panel animation during simulation.

IX. ACKNOWLEDGEMENTS

X. REFERENCES

[1]. Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg, and Ivona Brandic, Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility, Future Generation Computer Systems, Volume 25, Number 6, Pages: 599-616, ISSN: 0167-739X, Elsevier Science, Amsterdam, The Netherlands, June 2009.

[2]. Rajkumar Buyya, Rajiv Ranjan and Rodrigo N. Calheiros, Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities, Proceedings of the 7th High Performance Computing and Simulation Conference (HPCS 2009, ISBN: 978-1-4244-4907-1, IEEE Press, New York, USA), Leipzig, Germany, June 21-24, 2009.

[3]. R. Buyya, and M. Murshed, "GridSim: a toolkit for the modeling and simulation ofdistributed resource management and scheduling for Grid computing," Concurrency and Computation: Practice and Experience, vol. 14, no. 13-15, pp. 1175-1220, 2002.

[4].CloudAnalyst:ACloudSim-based Visual Modeller for Analysing Cloud Computing Environments and Applications, BhathiyaWickremasinghe, Rodrigo N. Calheiros, RajkumarBuyya, 20-23, April 2010, pp. 446-452.

[5].Dynamic Load-Balanced Multicast for Data-Intensive Applications on Clouds Tatsuhiro Chiba, Mathijs den Burger, ThiloKielmann, and Satoshi Matsuoka, 978-0-7695-4039- 9/10 $26.00 © 2010 IEEE DOI 10.1109 / CCGRID.2010.63.

[6]. Efficient collective operations for clusters in long-and-fast networks, M.Matsuda, T. Kudoh, Y. Kodama, R. Takano, and Y. Ishikawa, in IEEE International Conference on Cluster Computing (cluster 2006),2006.

[7].Round-Robin with FCFS Preemption: A Simple MAC Scheduling Scheme for Bluetooth Piconet Li-HsingYen , Chi-Hung Liao.

[8].Effective Cost Mechanism for Cloudlet Retransmission and Prioritized VM Scheduling Mechanism over Broker Virtual Machine Communication Framework Gaurav Raj1 and Sonika Setia2 International Journal on Cloud Computing: Services and Architecture(IJCCSA),Vol.2, No.3, June 2012