Cost Effective Service Broker Strategy in Selection of Data Centers in Cloud Computing

Shipra Gupta
Research Scholar
Deptt. of CSE,
MRIU, Faridabad
shipra.prof@gmail.com

Dr. Indu Kashyap
Deptt. of CSE,
MRIU, Faridabad
indu.fet@mriu.edu.in

Abstract: *Cloud computing is a boon for almost every sector as it reduces the overall expenditure by making I.T. Infrastructure, IT services and platforms "on-demand" basis. Data centers are located at different locations and provide services to the users in their closest proximity. As the user has to pay for the service utilization, so selection of data center plays an important role for maximum resource utilization at minimum expenditure and best response time. Service Broker plays an important role in the selection of data center. This paper presents a cost as well as time effective service broker strategy. The algorithm thrives to identify the data center which is most appropriate for a user in cloud environment. The enhanced service broker strategy is based on closest data center service broker strategy. The algorithm of service broker strategy has been simulated using the cloudanalyst simulator. Simulation results illustrate that the enhanced closest data center service broker strategy reduces the cost as well as overall processing time w.r.t existing load balancing strategies i.e., Round Robin, Equally Spread and throttled.*

Keywords: *cloud computing, data center, load balancing, service broker, User base, Region.*

I. INTRODUCTION

Cloud computing is a technology that aims at large amount of computing in a virtualized way. Users can access the cloud services and reduce the capital costs as they have to pay for the services they need as for the time period they need the service. NIST defined cloud computing as " Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources(e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

According to Gartner cloud computing can be defined as "a style of computing where massively scalable IT-enabled capabilities are delivered 'as a service' to external customers using Internet technologies".

The cloud is operated and managed at a data center owned by a service vendor that hosts multiple clients and uses dynamic provisioning.

Some benefits of cloud computing are:

- Pay per use Service: Organizations employing the cloud services don't own the infrastructure on which they need to execute their applications. They rent the application usage from the cloud provider. This enables the organizations to pay for only the services they need and use, and avoids for paying for resources they don't need, thus helping them avoid a large capital expenditure for infrastructure.

- Resource Utilization: The cloud service provider that delivers some or all of the services required to the organization can also share infrastructure between multiple clients. This helps improve utilization rates by eliminating a lot of wasted server idle time. The shared use of very high speed bandwidth distributes costs, enables easier peak load management, often improves response times, and increases the pace of application development.

- No need of Maintenance infrastructure: The organizations employing the cloud services do not have to worry about hardware, software or environment maintenance as the user does not have to manage and control the underlying configuration setting of the applications it uses.

- Mobility of data: Organizations/ users employing cloud services can access their data easily from anywhere through internet, which enables the professionals to work 24X7 on their own ease without the hassels of being available on their workplace.

- Rapid Elasticity: In cloud computing, the users/ organizations can dynamically scale their space both up and down without any effect on the ongoing services as in cloud environment, the capabilities and services appear to be unlimited to the consumer.

Along with the benefits, Cloud model do have some limitations. Some of these limitations can be summarized as:

- Security and Compliance: Security is a major issue in cloud environment. As the client data is stored virtually on the cloud, so proper security parameters are employed to protect data from vulnerable attacks. Data encryption is one technique for data security in cloud.

- In-house control: Moving the data on cloud means that the data is no more "in-house" and so every time the user needs the data, he has to extract the data from the cloud. The difficulty in extracting the data from cloud is one of the reasons why organizations having a second thought in adopting cloud computing.
- Internet connection dependency: In cloud computing, the data is stored on cloud and be accessed through internet browser. So, speed and reliability of internet connection is a major issue. Proper speed and continuous connectivity is mandatory for professionals working on cloud.

National Institute of Standards and Technology (NIST) defines cloud broker as: "A cloud broker is an entity that manages the use, performance and delivery of cloud services and negotiates relationships between cloud providers and cloud consumers." Some of the advantages of cloud service broker are [16]:

- A cloud service broker helps the user to determine the best possible framework for integration with cloud services.
- A cloud service broker provides interoperability between various cloud services and so makes the cloud cost effective.

II. CONVENTIONAL CLOSEST DATA CENTER ALGORITHM

In Closest data center service broker policy, the service broker selects the data center which is in the closest proximity of the user base. It identifies the closest data center on the basis of latency matrix. The working of this policy can be identified in following steps[14]:

- The service broker maintains a table of all the datacenters along with their regions.
- When Internet receives a request from any userbase, it contacts the service broker for the allocation of any data center.
- The service broker then generates a latency matrix based on the delay matrix from the Internet Characteristics. This contains the list of data centers with lowest delay w.r.t the userbase first.
- The service broker selects the first data center listed in the latency matrix table without considering the other data centers.

In this algorithm, the latency matrix is generated which identifies the datacenter in the closest proximity of the user but if two or more datacenters are available in the same proximity, i.e., more than one datacenters show the same position in the latency matrix then the datacenter identified first is selected and the user is connected to it. This arrangement works well if the data centers with same proxility w.r.t the user have same cost related parameters, but if the datacenter selected is costlier than the other datacenter then this arrangement is not acceptable in terms of performance as the cost of the datacenter can be minimized but is not. In this paper, the cost parameter is also evaluated along with the latency matrix to achieve the desired optimized performance.

III. ENHANCED CLOSEST DATA CENTER SERVICE BROKER POLICY

In the enhanced closest data center policy, the cost parameter is taken into account along with the latency matrix. The working of this policy can be identified in following steps:

- Closest data center service broker maintains an index table of all data centers indexed by their regions.

- When Internet receives a request from any userbase, it contacts the service broker for the allocation of any data center.

- Internet consults the service broker for the data center selection. The service broker uses any one of the service broker policy maintains a table of all the datacenters along with their regions.

- The service broker then generates a latency matrix based on the delay matrix from the Internet Characteristics. This contains the list of data centers with lowest delay w.r.t the userbase first.

- If two or more data centers indicates the same (lowest) delay w.r.t the userbase, then the VMcostper hour is considered.

- The service broker selects the data center listed in the latency matrix table with the lowest VMcostper hour.

IV. SIMULATION SCENARIO

In our experiment, a scenario is generated in the simulation toolkit CloudAnalyst.

CloudAnalyst is a tool developed at the University of Melbourne. It was developed to simulate large-scale Cloud applications with the purpose of studying the behavior of such applications under various deployment configurations. CloudAnalyst 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.

Simulation Setup: The scenario considered for the experiment consists of four userbases (UB1, UB2, UB3, UB4 in region 0, 2, 4, 5) and three data cenetrs (DC1, DC2, DC3 in region 1, 2, 5). The userbase and datacenter configuration used are summarized in Table 1 and Table 2 respectively.

Table1: Userbase Configuration

Name	Region	Requests per User per Hr	Data Size per Request (bytes)	Peak Hours Start (GMT)	Peak Hours End (GMT)	Avg Peak Users	Avg Off-Peak Users
UB1	0	60	100	3	9	1000	100
UB2	2	60	100	3	9	1000	100
UB3	4	60	100	3	9	1000	100
UB4	5	60	100	3	9	1000	100

Table2: Data Center Configuration

Name	Region	Arch	OS	VMM	Cost per VM$/Hr	Memory Cost $/2	Storage Cost $/s	Data Transfer Cost $/Gb	Physical HW
DC 1	1	× 86	Linux	Xen	0.5	0.05	0.1	0.1	1
DC 2	2	× 86	Linux	Xen	0.1	0.05	0.1	0.1	1
DC 3	5	× 86	Linux	Xen	0.1	0.05	0.1	0.1	1

Table3 shows the delay as well as the bandwidth matrix considered in our problem scenario.

Table3: Delay and Bandwidth Matrix

Delay Matrix

The transmission delay between regions. Units in milliseconds

Region/Region	0	1	2	3	4	5
0	25	100	150	250	250	100
1	100	25	250	500	350	200
2	150	250	25	150	150	200
3	250	500	150	25	500	500
4	250	350	150	500	25	500
5	100	200	200	500	500	25

Bandwidth Matrix

The available bandwidth between regions for the simulated application. Units in Mbps

Region/Region	0	1	2	3	4	5
0	2,000	1,000	1,000	1,000	1,000	1,000
1	1,000	800	1,000	1,000	1,000	1,000
2	1,000	1,000	2,500	1,000	1,000	1,000
3	1,000	1,000	1,000	1,500	1,000	1,000
4	1,000	1,000	1,000	1,000	500	1,000
5	1,000	1,000	1,000	1,000	1,000	2,000

V. SIMULATION RESULTS

We have analyzed the average response time and data center processing time with all three load balancing policies (round robin, equally spread and throttled) for coventional as well as the enhanced algorithm. The results are summarized in Table 4 and Fig.1, 2 and 3.

Table 4

	Closest data Center conventional	Closest data Center Enhanced
Overall Response Time	150.02	149.75
Data Center Processing Time	1.18	0.91
Total Cost	10.79	10.69

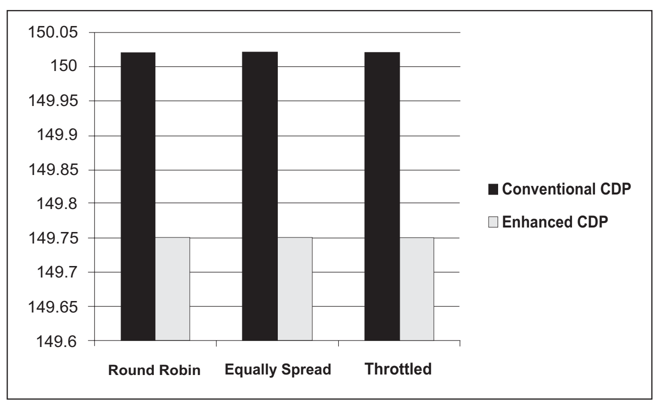

Fig. 1: Overall Response Time

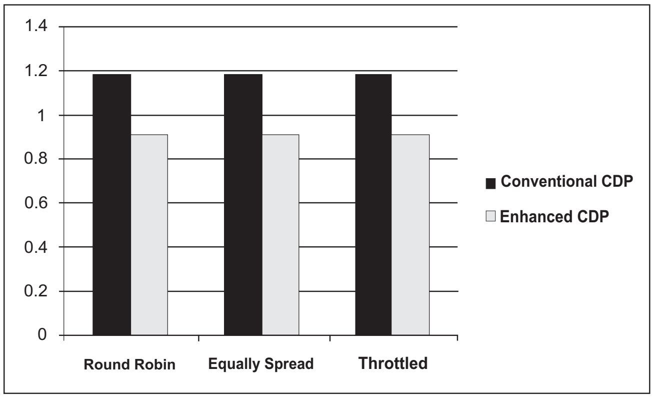

Fig. 2: Data Center Processing Time

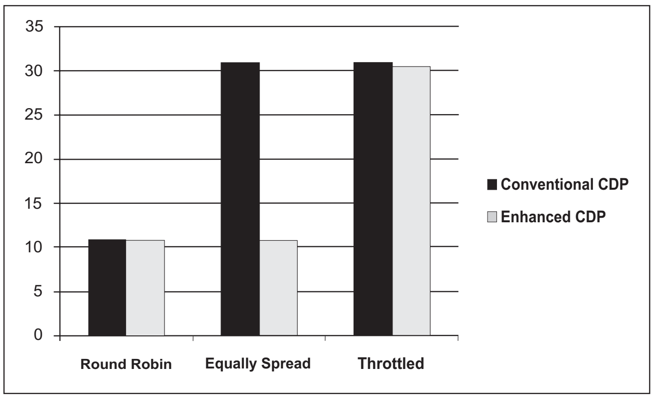

Fig. 3: Total Cost

Fig.1, 2 & 3 shows that overall response time, Data Center Processing Time as well as the total cost decreases in the enhanced algorithm for all the load balancing policies.. So, it is evident that the enhanced algorithm is both time and cost effective.

VI. CONCLUSION & FUTURE WORK

In our study, we have analyzed a scenario in cloud environment with the help of Cloud Analyst simulator. The results show that overall response time, data center processing time and the data center processing time are reduced by considering the lowest cost along with the latency matrix. As a future work, we are considering to analyze the cloudanalyst simulator form other two service broker policies to make it more efficient in terms of cost and time.

REFERENCES

[1]. Fang Liu, Jin Tong, Jian Mao, Robert Bohn, John Messina, Lee Badger and Dawn Leaf, " NIST Cloud Computing Reference Architecture", National Institute of Standards and Technology.

[2]. Chhabra A, G. Singh, "Qualitative Parametric Comparison of Load Balancing Algorithms in Distributed Computing Environment", 14th International Conference on Advanced Computing and Communication, Page 58-61, Dec 2006 IEEE.

[3]. Eric Keller, Jakub Szefer, Jennifer Rexford and Ruby B. Lee, "NoHype: Virtualized Cloud Infrastructure without the Virtualization", ISCA'10, June 19–23, 2010, Saint-Malo, France.

[4]. http://www.cloudbus.org/cloudsim : (Cloud Analyst can be downloaded from here).

[5]. Saif U. R. Malik, Samee U. Khan, and Sudarshan K. Srinivasan, Modeling and Analysis of State-of-the-Art VM-Based Cloud Management Platforms, IEEE TRANSACTIONS ON CLOUD COMPUTING, VOL. 1, NO. 1, JANUARY-JUNE 2013.

[6]. Joseph Doyle, Robert Shorten, and Donal O'Mahony, Stratus: Load Balancing the Cloud for Carbon Emissions Control, IEEE TRANSACTIONS ON CLOUD COMPUTING, VOL. 1, NO. 1, JANUARY-JUNE 2013.

[7]. Jenn-Wei Lin, Chien-Hung Chen, and J. Morris Chang, QoS-Aware Data Replication for Data-Intensive Applications in Cloud Computing Systems, IEEE TRANSACTIONS ON CLOUD COMPUTING, VOL. 1, NO. 1, JANUARY-JUNE 2013.

[8]. Xinyu Lei, Xiaofeng Liao, Tingwen Huang, Huaqing Li, and Chunqiang Hu, Outsourcing Large Matrix Inversion Computation to a Public Cloud, IEEE TRANSACTIONS ON CLOUD COMPUTING, VOL. 1, NO. 1, JANUARY-JUNE 2013.

[9]. Kousik Dasgupta, Brototi Mandal, Paramartha Dutta, Jyotsna Kumar Mondal, and Santanu Dame, "A Genetic Algorithm (GA) based Load Balancing Strategy for Cloud Computing" , International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013, Vol-10, Pages 340-347, 2013 Elsevier Science Direct.

[10]. Dhinesh Babu L.D., P. Venkata Krishna, "Honey bee behavior inspired load balancing of tasks in cloud computing environments", Applied Soft Computing, Vol 13, Issue 5, Pages 2292-2303, May 2013, elsevier.com/l ocate/asoc.

[11]. Johan Tordsson, Rubén S. Montero, Rafael Moreno-Vozmediano, Ignacio M. Llorente, "Cloud brokering mechanisms for optimized placement of virtual machines across multiple providers", Future Generation Computer Systems, Vol 28, Issue-2, Paqges 358-367, Feb 2012, www.elsevier.com/locate/fgcs.

[12]. Bhathiya Wickremasinghe, Rodrigo N. Calheiros, and Rajkumar Buyya, "CloudAnalyst: A CloudSim-based Visual Modeller for Analysing CloudComputing Environments and Applications", 24th International Conference on Advance e Information Networking and Applications (AINA) IEEE Computer Society, Pages 446-452, April 2010.

[13]. Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi, "Mastering Cloud Computing", Mc Graw Hill Publication, 2013.

[14]. Rakesh Kumar Mishra, Sandeep Kumar, Sreenu Naik B, " Priority based Round Robin Service Broker Algorithm for CloudAnalyst", Advance Computing Conference (IACC), Pages 878-881, Feb 2014 IEEE International.

[15]. Fang Liu, Jin Tong, Jian Mao, Robert Bohn, John Messina, Lee Badger and Dawn Leaf, "NIST Cloud Computing Reference Architecture", National Institute of Standard & Technology, Sep 2011.

[16]. ISACA, " Controls and Assurance in the cloud: Using COBIT5" 2014, www.isaca.org/COBITuse.

❐