Reviewing and Reducing Hidden Cost by Proposing CSIPP-using DMT

[1]R. K. Seth and [2]R. Chuchra

[1] *Department of Applied Sciences, Sri Sai University,*
Palampur-176061(INDIA)
[2]*Department of Computer Science and Engineering,Punjab Technical University,*
Amritsar-143001(INDIA)

Abstract-This paper reviews the analysis of hidden cost in cloud computing and proposes a new methodology to reduce the cost using Data Mining Techniques (DMT) during data migration. In the hidden cost, the customer pays the fee during data migration as well as in the form of access fee. A policy concept has been proposed to reduce cost after reviewing the analysis of the hidden cost in cloud computing and data mining and the new proposed policy is termed as "Cloud Service Issuing Policy Procedure (CSIPP)". The designed methodology and proposed procedure may help to issue a cloud service for a specific time interval on user demand basis and the hidden cost may be reduced using Data Mining Techniques (DMT).

Keywords-Hidden cost, cloud computing, data mining.

1. INTRODUCTION

It has been figured out that the hidden cost involves the cost due to data migration as well as due to access fee. The various types of costs have been figured out and are listed as follows:

- **Operational Cost:** Includes the cost of operations performed for gaining the access of specific service.
- **Labor Cost:** Includes the cost due to cloud service provider only.
- **Computing Cost:** Cost for performing Calculations.
- **Software Installation Cost:** Includes software installation cost.
- **True Cost:** Details the overall cost for using cloud service.
- **TCO (Total cost of ownership):** It describes the potential cost of purchasing and bringing technology in house. It also helps to identify total cost of purchasing and operating technology products or service such as cloud computing [7].

Hidden cost is an intangible cost and is interesting subject to study in the field of software engineering. This cost is significantly involved in financial transactions of any enterprise and is intangible due to the fact that the customer does not pay directly [1]. In general, the aim of data mining and cloud computing is to save time and to reduce cost. The buying of cloud service as a part of "cloud economics" is altogether different from buying the traditional IT Services [16]. It is not easy to understand the investments in cloud computing implementation and it may be evaluated in terms of risk and cost. The issues pertaining to cloud performance are increasing at a fast rate and are becoming complicated to figure out the root cause of the hidden cost

[18]. The cloud service can be evaluated on the basis of level of service depending upon the type of service viz. "Break-fix service" and "Premium level of service" that requires a large degree of management in business and ensures proper delivery of service [8]. The level of service also affects hidden cost [15]. Finally, the hidden cost can be estimated by the survey related to SSL certificates [6]. There may be increase in cost when a user is interested for availability of his/her own data and for that the user would make available of his/her own BYOD (Buy Your own Device)[9] [10]. It is a matter of fact that the hidden cost cannot be ignored and if ignored it could cause a major negative impact on business. To counter the issues pertaining to hidden costs, the measures in terms of the following points may be taken into account.

- Processing rules.
- Framing rules based on people and desired information and not focusing upon platform.
- Strict monitoring of the framed rules consistently. It is critical to select a platform independent IT tools and reduplication data in cloud [17].

The data mining techniques viz. classification, regression and time series may be used to figure out hidden cost in the proposed policy procedure "Cloud Service Issuing Policy Procedure (CSIPP)". Though limited literature is available on policy procedure in respect to hidden cost, the authors made an effort in preparing an article on hidden cost by studying and analyzing the information available on various web sites.

2. REVIEWING ANALYSIS OF HIDDEN COST

2.1 Reviewing analysis of hidden cost in cloud computing

Most of the users are unaware of the hidden cost in cloud computing. The four types of hidden costs are listed as follows:

- Data Movement cost.
- Integration Cost.
- Organization overhead Cost.
- Virtualization Cost.
- Data Movement cost provides the information about the quantity of data to be moved from one server to another server.
- Integration cost includes internal environment cost. As an example, interconnection with the connectors. It is a matter of fact that cloud

providers do not show such kind of collaborations in front of clients.

➢ Organization overhead cost the clients are not charged for the facility like power and space. Both are the essential factors for any organizational structure and this cost is automatically added when any client buy a service from the cloud service provider.

➢ Virtualization cost includes man power and salary. It is related to computing functions and technology implementations using physical hardware [5].

2.2 Reviewing analysis of hidden cost in data mining

➢ Operational cost arises in predictive and descriptive models and depends upon the number of operations applied to perform some specific task in any model.

➢ The hidden cost can be detected by "Markov Model" with its implementation in given set of data using Oracle Data Mining (ODM).

➢ Hidden cost also rises due to SSL (Secure Socket Layer) self assigned certificates.

➢ Visualize hidden cost by data distribution.

➢ Using "Neurofuzzy Logic", the hidden patterns can be detected and hidden cost may be analyzed.

➢ The Data cleaning during data pre-processing in KDD process for removing noise and disturbances to maintain data quality may further add into the hidden cost.

➢ "Web Mining" can be utilized to find the hidden patterns and labor cost also adds into the hidden cost automatically.

➢ Hidden data cannot be extracted directly simply by applying queries on databases. The "privacy preserving data mining method" helps to detect hidden cost in which labor cost is included.

➢ The data mining technique using various tools may be beneficial for searching hidden relationship in complex data sets and is cost effective [14].

➢ It also includes higher network charges for storage and database applications or latency issues for users who may be located for class service providers.

➢ The data storage in data warehouse in Zeta byte (one billion terabytes or one trillion gigabytes) and is recorded on hard drives around the world adds into the hidden cost [13].

3. REDUCING HIDDEN COST

3.1 Reducing hidden cost in cloud computing

The following strategies can be implemented to reduce hidden cost in cloud computing:

➢ Resources must be turned off after usage.

➢ Use some limited physical channels for data transfer using routing algorithms after deciding the shortest route for data transfer. The short path holds limited number of processing nodes that automatically saves time as well as cost.

➢ It would be better to focus on policies rather than technology [2].

➢ Cloud service billing should be based on the hourly or weekly basis rather than yearly basis.

➢ Using "Polling" concept, the transaction fee may be reduced that depends upon the quantity of data sent by the user.

➢ "Cloud economics" concept could reduce billing in the sense that user need not to pay fee for all the time. Some companies provide cloud services time slots in which user need not to pay at concessional rates.

➢ The hidden cost may be reduced by appropriate utilization of power and storage space. The servers should be more energy efficient. The server humming can considerably be reduced using appropriate technology to save energy.

➢ Trill hidden cost based upon STP (spanning tree and its progeny) may be applied to reduce cost during the movement of data from one cloud to another by saving bandwidth [3].

➢ The substantial amount of money can be saved by leasing the equipments instead of direct buying. Using leased equipments could save money to be spent in repairs, upgrades and maintain.

➢ "WIRESHARK TOOL" is a potential open source tool and is easily available to resolve various troubleshooting problems to reduce hidden cost.

➢ The cost due to utilization of operating system depends upon the number of tasks performed in a specific period of time and in order to decrease this cost, the following formula can be used to figure out the processing cost due to operating system.

Total operating system cost = (number of tasks is to be processed/total time consume)*100.

➢ Migration of clouds also helps to save hidden cost in cloud computing.

➢ Using the concept of "Auto-Scaling", the consumption can be increased or decreased and for that "Elastic scale" could be applied to perform auto-scaling on cloud servers.

3.2 Reducing hidden cost in data mining

➢ Save Disk space at first by reducing the input/output cycles for performing any task that helps to save hidden cost arises due to the operating system processing time [11].

➢ The hidden cost inside Meta-Data (data about data) can be reduced by applying various data analysis tools like OLAP having characteristics to perform "exploratory data analysis" that basically helps to identify systematic relation between variables. In this case, many variables are taken into account and compared using a variety of techniques in the search for systematic patterns.

➢ Operational cost on specific task may be reduced by decreasing the number of operations appropriately without affecting the task to be performed.

➢ Preference should be given to cloud mining to reduce software installation cost.

➢ The cost may be reduced by Teradata data mining [12].
➢ The cost may be reduced by adopting appropriate strategic storage options.
➢ Reducing consumption of energy by utilizing suitable techniques could reduce cost and output can be improved.
➢ With the use of loop scheduling, the cost of data mining can be minimized. The use of hidden association rules and detecting right execution would reduce the cost.
➢ In electrical and electronically driven systems, the on-demand Ampere facility may be used to reduce cost.
➢ *"Arjuna Company"* finds innovative solutions and target engine working by analyzing valuable Telecom Data to find hidden patterns in consumer behavior. The engine figures out who is interested in WHAT, WHEN, WHERE with the relevant information [4].

4. DESIGN METHODOLOGY

The communication flow chart shown in figure 1 describes the design methodology and may be implemented using appropriate procedure.
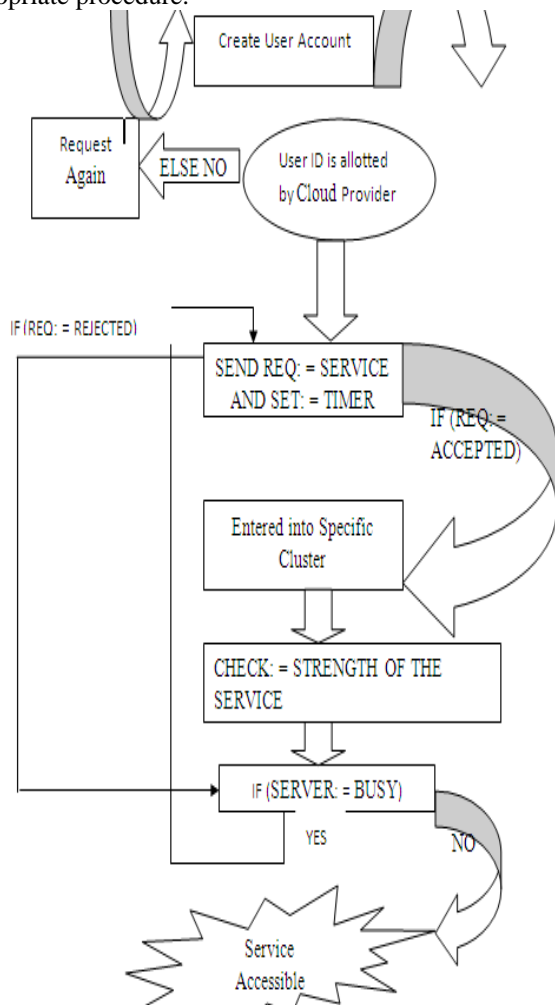

Figure 1: Cloud service request flow in proposed policy-
"CSIPP-Using DMT".

5. PROPOSED PROCEDURE: CSIPP-USING DMT

Procedure for Cloud service Issuing Policy using Data mining techniques is as follows:

Step1. CREATE USER ACCOUNT: = CLOUD THEN USER_ID will be allotted by cloud provider.
Step2. SEND REQUEST FOR SERVICE :-----> Time Duration (Set: = Session) THEN CALL TYPE/LEVEL OF SERVICE WITH ALOTTED USER_ID.
{In the second step, the type of service to be provided by the cloud using "time series analysis data mining technique" on the basis of specific time interval utilizing that service. The service based upon USER DEMAND will be provided by the cloud provider}.
Step3. IF REQUEST: = ACCEPTED! THEN AUTOMATICALLY CLUSTERING DATA MINING TECHNIQUE IS SELECTED.
{It shows request that automatically enters into specific cluster for easy identification of the TYPE and LEVEL of service and payment can be made according to STEP 2}
Step 4. THEN CHECK: = SERVICE STRENGTH!
{Checking the strength of the service; strong, very strong; weak, very weak, using regression analysis}.
IF (STRENGTH OF SERVICE: = VERY STRONG)
{
THEN SIGNAL COLOR: = DARK   GREEN.
}
ELSE IF (STRENGTH OF SERVICE: = STRONG)
{
THEN SIGNAL COLOR: = GREEN.
}
ELSE IF (STRENGTH OF SERVICE: = VERY WEAK)
{
THEN SIGNAL COLOR: = DARK RED.
}
ELSE (STRENGTH OF SERVICE: =WEAK)
{
THEN SIGNAL COLOR: = RED.
}
Step5. IF SERVER: = BUSY! i.e. NOT AVAILABLE, THEN REPEAT FROM STEP 1 TO STEP 3.
{Non availability of instant service means that SERVER is BUSY and WAIT till the server is free}
THEN enter into QUEUE and REPEAT up to "N" number of loops.
Step6. THEN SERVICE: = ACCESSABLE! AND END OF THE LOOP.

Figure 2 summarizes the steps to follow for enabling the proposed procedure for its implementation to reduce hidden cost.
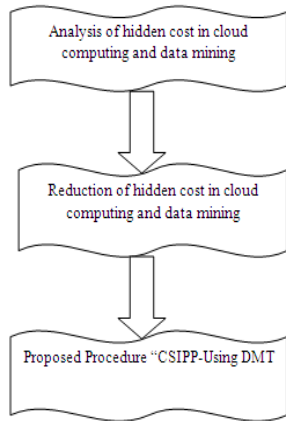
Figure 2: Steps for enabling the proposed procedure.

The proposed policy for reducing the hidden cost would serve as a bridge between cloud server provider and cloud user [19]. Figure 3 shows the cloud service flow upon user's demand by incorporating CSIPP policy.
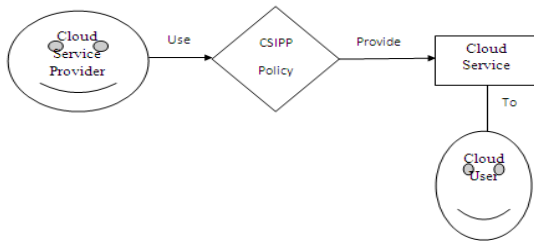

Figure3: Cloud service flow from cloud provider to cloud user.

CONCLUSION

The hidden cost due to various factors has been reviewed and analyzed. The hidden cost may considerably be reduced by developing the methodology as designed and implementing the proposed procedure based upon the CSIPP- using DMT. The cloud service may be provided depending upon the defined requirements as per user's demand and license may be issued accordingly.

REFERENCES

[1]http://www.demonstratingvalue.org/resources/hidden-costs-and-contributions worksheet.

[2]www.symantic.com/connect/blogs/avoiding-hidden-costs-cloud.

[3]http://gigaom.com/2009/07/17/thehidden-cost-of-the-cloud-bandwidth-charges/.

[4]http://articles.economictimes.indiatimes.com/2014-0-1-20/news/46374888-1-mobile-marketing-sms-brands-and-retailers.

[5] http://www.unitiv.com/it-solutions blog/bid/79589/4-Hidden-Costs-of-Cloud-Computing.

[6]http://www.cio.in/by-thenumbers/hidden-costs-cloud-implementations?topic=cloud.

[7]http://www.businessnewsdaily.com/4674-what-is-tco.html.

[8]http://www.cio.co.ke/blog/the-hidden-costs-of-it-support.

[9]http://copperegg.com/cloud-computing-includes-hidden-costs-report-says/.

[10]http://bringyourownit.com/2013/08/07/byod-the-hidden-costs-video/.

[11]   http://www.dmnews.com/data-minings-expensive-hidden-costs/article/67595/.

[12]http://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=19&cad=rja&ved=0CGQQFjAIOAo&url=http%3A%2F%2Fwww.teradata.com%2Fbrochures%2FTeradata-Data-Mining-Services.eb1719%2F&ei=xjLzUq33IoulrQfRuoCACA&usg=AFQjCNHK40lazRiHmdSxOWu9kTNjT-qVTg&sig2=NEZ8gg0rUEAOpGjDv_CxQQ&bvm=bv.60983673,d.bmk.

[13]http://www.securitynewsdesk.com/2013/09/25/data-mining-key-to-revealing-hidden-criminal-behaviours-says-sas/.

[14]http://info.bcdtravel.com/rs/bcdtravel/images/Hidden_Spend_Paper.pdf.

[15]http://www.cio.co.ke/blog/the-hidden-costs-of-it-support.

[16]blogs.forrester.com/james-stalen/13-02-18-how-to-avoid-the-hidden-costs-of-cloud-computing.

[17]cloudtimes.org/2013/02/01/   symatec-study-revals-hidden-costs-of-cloud-and-data-security-risks/.

[18]www.compuware.com/content/dam/compuware/apm/assets/whitepapers/wp-costocloud.pdf.

[19]www.rand.org/content/dam/rand/pubs/occasional-papers/2010/RAND-0P279.pdf.