Journal of International Technology and Information Management

Volume 22 | Issue 3 Article 6

11-8-2013

A Model for Cost-Benefit Analysis of Cloud Computing

Krishnadas Nanath
Indian Institute of Management, Kozhikode, IMT Dubai

Radhakrishna Pillai
Indian Institute of Management, Kozhikode, IMT Dubai

Recommended Citation

Nanath, Krishnadas and Pillai, Radhakrishna (2013) "A Model for Cost-Benefit Analysis of Cloud Computing," *Journal of International Technology and Information Management*: Vol. 22: Iss. 3, Article 6.
Available at: http://scholarworks.lib.csusb.edu/jitim/vol22/iss3/6

A Model for Cost-Benefit Analysis of Cloud Computing

Krishnadas Nanath
Indian Institute of Management, Kozhikode, IMT Dubai
UNITED ARAB EMERATES

Radhakrishna Pillai
Indian Institute of Management, Kozhikode, IMT Dubai
UNITED ARAB EMERATES

ABSTRACT

Cloud computing is emerging as a powerful computing paradigm with its aim of efficient resource utilization and contribution to Green IT. However, the decision of shifting to cloud computing always remains risky from customer's perspective considering the benefits they would attain by doing so. The extant research on cloud computing focuses more on technical aspects like security, quality, efficiency etc. However, the research on adoption of cloud computing is at its infancy stage. Therefore, this paper attempts to come up with a model to analyze the cost-benefits to decide upon the adoptability of cloud computing. It takes into consideration various parameters of an organization such as number of servers, power requirements and other computational/non-computational resources. This model uses a three layer approach for the cost-benefit analysis and draws insights on profitability when an organization shifts to cloud computing in each layer. The three layers are base cost estimation, data pattern based cost estimation and project specific cost estimation. These layers are designed to provide different levels of decision making to aid managers in their attempt to find out the prospects of adopting cloud computing in their organization. The data for cost benefit analysis was collected from organizations that comprised of both small scale and large scale datacenters. It was found that cloud computing is profitable for start-ups and small firms (small scale datacenters) when compared to well-established firms.

INTRODUCTION

Cloud computing paradigm has evolved recently and it has taken commercial computing to a new level. The concept of cloud computing rests upon the idea that computing resources will reside somewhere other than the computer room and that the users will connect to it using the resources as and when required. In effect, it displaces the infrastructure to the network so that the overall cost with respect to the management of hardware/software resources is reduced (Hayes, 2008). It appears to be highly disruptive technology (Rimal et al., 2009) hinting to the future where computation moves from local computers to centralized facilities operated by third party compute and storage utilities (Foster et al., 2008).

However, considering the practical implementation of cloud computing, the adoption of cloud platforms by organizations/scientific community is in its infancy. There is a paucity of research towards a model that can demonstrate the benefits of cloud computing adoption and suggest the ideal time to shift to cloud computing. This study attempts to develop a cost-benefit analysis model that can present a clear picture to the IT managers when the shifting from legacy systems to cloud computing is concerned. The computing resources and IT infrastructure of every organization is idiosyncratic. Hence, direct recommendations on profitability cannot be given until all the inputs of organizations have been considered for profitability evaluation. This paper, therefore, suggests a model that can take various parameters of an organization and provide recommendations on profitability of shifting to cloud computing.

The pricing model in cloud computing is quite similar to usage based pricing. Customers pay for the computing resources by means of customized service level agreement hiding the underlying technological infrastructure (Xiong & Perros, 2009). This concept of pay-as-you-go in cloud computing differs from traditional renting method which involves payment of negotiated cost to have the resource for a specific period of time irrespective of the actual usage. The cloud computing service taken up in this paper for the cost-benefit analysis model is Amazon AWS. It rounds up their billing to the nearest server hour/GB per month. AWS is chosen as an example because major players like Amazon reflect the most common pricing mechanism in the cloud market.

The proposed model in this paper works on three layers. These layers represent the different levels at which organizations plan to adopt cloud computing. In the first case, a base cost estimation is done where the organization can compare the cost of the entire computational facility in-house to the cost of shifting completely to loud computing. The second layer performs the analysis based on the data pattern expressed in terms of the average amount of data it processes, transfer rate, the demand estimation/provision etc. This layer gives an instant recommendation on the feasibility of shifting to the cloud by taking into account the inputs of layer one as well. The last layer is a project specific layer which would be very helpful for organizations planning to keep the present infrastructure intact and using cloud computing for a specific upcoming project. This third layer demonstrates the widely used scenario in present day organizations. It takes inputs concerning the nature of the project and give recommendations on executing the project on cloud.

This paper proceeds as follows. Next section attempts derive acumen from past studies that have discussed cost benefit analysis of cloud computing. This is followed by three layers of the model. Layer 1 describes the variables and the methods used for the computation of base cost estimation. Layer 2 attempts the same for data pattern variables followed by Layer 3 describing the project specific variables. The next section describes the proposed model using a three layer approach and how it can be used by organizations to get recommendations on cloud computing adoption. The last section applies the model to organizations ranging from small/medium enterprises to well established organizations. The application of the proposed model is limited to first and second layer due to complications in data aggregation. The data is collected from a mix of primary and secondary sources.

LITERATURE REVIEW AND RELATED WORK

Cloud computing literature brings about several factors relevant to implementation cloud computing in firms. They include adoption, implementation technicalities, cloud infrastructure and various other factors. Pyke (2009) has explained how cloud computing would be a potential paradigm shift from traditional computing model and when an application is considered to be in cloud. The cloud infrastructure and delivery models have been studied in Nucleus (2009) along with market oriented resource allocation of goods. The challenges and opportunities of practical implementation in cloud architecture along with risk involved in shifting from legacy system to cloud computing have been explored in Fox et al. (2009). They also presented a trade-off equation for evaluating the profitability on adoption of cloud computing which has been used to derive some aspects of the proposed model in this study.

Various blogs have explored the profitability of adopting cloud computing (Nucleus, 2009; Hinchcliffe, 2012; Ghag, 2008; Rosenberg, 2012). These blogs refer to individual case studies focusing on the benefits achieved by implementation of cloud computing. Most of the blogs presented basic cost calculations in cloud and also initiated discussions about the profitability of cloud computing. However, these blogs have a limited scope as they deal with individual cases and there exists no model for evaluating the cost benefits. Return on investment of implementing cloud computing was been presented in few studies (Rosenberg, 2012). However, the scope was limited only to email servers and the objective was not to move towards a cost-benefit model.

Few studies on the other hand attempt to work on Cloud pricing, the results of which can be used as inputs for developing cost-benefit model. Buyya et al. (2009) have explored the performance of several pricing mechanisms including Fixed and FixedTime in Aneka enterprise cloud environment setup. Palankar et al. (2008) have evaluated Amazon S3’s ability to provide storage at low cost to the large scale projects from cost, availability and performance perspective. It was a good attempt given the fact that many home users and small business to large enterprises subscribe to S3 service (Kirkpatrick, 2006). It stores more than 5 billion user objects and handles over 900 million user requests per day (Bezos, 2007). However, it did not consider the user side cost computation as the aim was to evaluate the performance of Amazon service based on set parameters. A similar performance evaluation of Amazon EC2 service was performed over MPI Applications in Walker (2008). An excellent cost analysis was also attempted by Li et al. (2009) but all the calculations were again performed from vendor’s perspective only.

A short cost benefit analysis has been demonstrated in Simson (2007), where a test API was evaluated on cloud. This study rated Amazon’s quality of service; however the list of parameters considered for cost-benefit analysis was limited. Deelman et al. (2008) have explored on how to adjust project requirements so that it can be beneficial to implement it on cloud. They adjust the cost of running a scientific workflow over a cloud. However, no model is presented that can be used by an organization to decide the adoption of cloud computing. Similar execution of workflow structured applications has been addressed in Singh et al. (2007) and Zhao & Sakellariou (2007).

Few authors have recently focused on return on investment (ROI) calculations in cloud computing (Misra & Mondal, 2011). The objective of papers encompassing ROI calculations is to mathematically derive the return a firm would get based on investment required in cloud computing. The models are restricted to cost/savings and in some instances, business intelligence (Mircea et al., 2011). However, the detailed breakup of cost components in ROI calculations appear to be the missing link of studies dealing with cloud ROI. Hence, this paper proposes a comprehensive model that can factor in the inputs of an organization and provide recommendations on adopting/shifting to cloud computing from different perspectives.

LAYER 1: BASE COST ESTIMATION

Since cloud computing uses on-demand pricing, it is important to calculate the cost of maintaining IT infrastructure in house. Though many authors suggest more sophisticated cost calculation model for cloud computing (Stuer et al., 2007; Hosanagar et al., 2004; Abramson et al., 2002), on-demand pricing would still have its ubiquitous presence in all cost calculation methods. This section explores various costs involved in in-house management of IT infrastructure which is independent of any particular project requirement. For most of the components involved in this section, the concept of total cost of ownership (TCO) is used. TCO is the means of addressing the real cost attributing to owning and managing IT infrastructure (Cappuccio et al., 1996). It comprehensively considers the entire lifetime spending, capital costs, cost of operations and hence is suitable for base cost estimation. A total of nine components haven been considered in base cost estimation including amortization, cost of servers, network cost, power cost, software cost, cooling cost, real estate cost, facility cost and support & maintenance cost. For each component, the following details are provided: a) explanation of all the variables involved and b) the method to calculate the cost of the component. The overall aim is to come up with monthly costs for all the components being considered and thus all variables are converted to monthly parameters. Unless otherwise mentioned, the currency for all calculations is United States Dollars (USD) and the computations are made on monthly basis.

Amortization

It is important to understand the contribution of IT infrastructure costs to the monthly rental structure in an organization. Hence, amortization parameter is calculated for servers and other facilities so that fair attribution of costs for various IT resources (hardware/software) can be brought about. This parameter is required to calculate the monthly depreciation cost (amortization cost) of each infrastructure item being considered. These items have initial purchase expense, the cost of which is calculated based on the duration over which the investment is amortized at assumed interest rate. Studies have revealed that the cost of CPU, storage and bandwidth roughly double when the costs are amortized over the lifetime of the infrastructure (Hamilton, 2009; Hamilton, 2008).

The interest rate is generally 5% per annum (Greenberg et al., 2008) and the depreciation period of real estate is ten years whilst that of servers/other facilities is three years (Hamilton, 2009). Once the amortization parameter is obtained, it can be then used in the calculations of required component to obtain the monthly cost. The amortizable parameter for facility (Ap_F) is computed differently from that of server (Ap_S) owing to the different amortization periods. The interest rate is represented as Cost of Money (Com) and is kept in variable form (instead of 5%) to accommodate any changes. Ap_F is calculated as (Com/(1-power ((1+Com),(-1 * Time_F)))), where Time_F is the facility amortization period and measured in months. Similarly, Ap_S is calculated as (Com/(1-power ((1+Com),(-1 * Time_S)))), where Time_S is the server amortization period and measured in months. Therefore, Ap_S and Ap_F will be used for computation of costs in the upcoming subsections.

Cost of Servers

Servers are generally mounted on racks and it is assumed that all the servers have similar configurations. This assumption is made to ease the computation for cost of the server (without amortization). Hence, Cost_S can be computed as (N_S * Cost_PS), where N_S is the number of servers in a firm and Cost_PS is the cost per server in Dollars. The amortizable parameter for server calculated in the previous part will be used to determine the amortized server cost- Cost_Am_S. It can be calculated as Cost_Am_S=(Cost_S * Ap_S), where Ap_S is Amortizable Parameter for Server from previous sub-section. The costs other than base cost associated with the purchase of the server have been calculated separately.

Network Cost

The components that contribute to the networking costs are NIC, switches, ports, cables, software and maintenance. The cost of NIC is already attributed in the server cost while that of the software will be taken up in the software cost section. Maintenance activities have also been taken up separately in form of Support and Maintenance Cost. Hence, this section would only deal with the cost of switches, ports, cables and the implementation costs. Since cost associated with networking again has an initial expense, it is amortized to come up with the monthly cost.

The total networking cost (Cost_Net) is a sum total of Cost of Port (Cost_Port), Cost of Cable (Cost_Cab), Cost of Switch (Cost_Switch) and implementation cost (Cost_Imp). All costs are measured in USD and calculated using the following equations:

Cost_Port= N_Port (No. of Ports) * Cost_per_Port (Cost per port)
Cost_Cab= N_Cab (No. of Cables) * Cost_per_Cab (interconnect cable cost)
Cost_Switch= N_Sw (No. of Switches) * Cost_per_Sw (Cost per switch)
Cost_Net= Cost_Port + Cost_Cab + Cost_Switch + Cost_Imp

However, networking costs should be amortized to calculated the amortized networking cost represented by Cost_Am_Net and given by Cost_Am_Net=Cost_Net * Ap_S (Ap_S is Amortizable Parameter for Server).

Power Cost

Few studies have quoted that power is the single largest cost in high scale data centers. Though the validity of the statement is debatable, power is clearly one of the fastest growing costs (Brill, 2009). Green Grid has coined a very useful term named power usage effectiveness (PUE) which represents the ratio of total power to IT Equipment power (Belady et al., 2008). Inefficient enterprise facilities are often as low as 2.0 to 3.0 while that of the efficient ones being better than 1.2 (Google, 2010). The IT infrastructure that contributes to the power consumption in an organization includes computing infrastructure (server, switches etc.), network critical physical infrastructure, transformers, uninterruptable power supplies, fans, air conditioners, pumps, lighting etc. (Sawyer, 2004).

The total power cost is computed annually and is represented by Cost_Tot_Pow. It is calculated as: Cost_Tot_Pow= (Size_Pow * Use_Pow * Eff_Pow * Cost_Pow * 24 * 365), where Size_Pow is the size of facility (critical load) measured in KW, Use_Pow is the average power usage (average percentage of provisioned power used), Eff_Pow is the power usage effectiveness and Cost_Pow is the cost of power measured in US Dollars per kwh. Therefore, the monthly power cost (Cost_Am_Pow) can be easily computed as Cost_Tot_Pow/12.

Software Cost

In order to manage the data centers, it is required to install the operating system patches and resources for load balancing. The cost of software associated with the base cost estimation is due to license payment. There are two classes of software considered for cost analysis based on the license structure. Class A software includes operating system while Class B deals with other base software (Application Server, VM Software etc). Class B does not include the project specific software as it will be dealt in the layer addressing the project costs. The details of exact pricing based in total cost of ownership (TCO) for Class A software is provided in Cybersource (2002).

The total cost of software for a firm is represented by Cost_Tot_Soft and is given by:
Cost_Tot_Soft = [(N_ClassA * Cost_ClassA * ∏A) + (N_ClassB * Cost_ClassB * ∏B)]. The description of the variables involved in this equation is given in Table-1. However, the total software cost must be amortized to obtain the amortized software cost- Cost_Am_Soft, calculated as Cost_Am_Soft= [Cost_Tot_Soft * Ap_F] (Ap_F is Amortizable Parameter for Facility)

Table 1: Software Cost.

Name of the variable	Symbol	Unit
Number of Class A Software	N_ClassA	
Number of Class B Software	N_ClassB	
Unit price of Class A Software (Total Price, One time)	Cost_ClassA	Dollars
Unit price of Class B Software (Total Price, One time)	Cost_ClassB	Dollars
Server Utilization Class A (Percentage of unit price that accounts for the annual cost)	∏A	Percentage
Server Utilization Class B	∏B	Percentage

Cooling Cost

Past research has shown that power consumed in data center is equivalent to the heat generation in it indicating the the power rating and thermal output equivalency (Rasmussen, 2007). This cost estimation method uses the term ‘Cooling Load Factor’ coined by Li et al. (2009). It represents amount of power consumed by the cooling equipment for 1W of heat dissipated. The other related parameters like Airflow Redundancy constant and Inefficiency constant has been derived from McFarlane (2005). The former represents the redundant airflow required to cool the data center while latter represents the redundant airflow to account for burden of humidification.

The cooling factor (Factor_Cool) is calculated as Factor_Cool= [LF_Cool * (1 + Red_Cool) / Ineff_Cool], where LF_Cool is the cooling load factor, Red_cool is the airflow redundancy constant and Ineff_Cool is the inefficiency constant. Cooling factor is calculated to attribute it as a percentage of the power cost. Therefore, the required total cost of cooling (Cost_Tot_Cool) is calculated as Cost_Tot_Cool= Factor_Cool * Cost_Am_Pow (Cost_Am_Pow is monthly power cost from previous sub-sections).

Real Estate Cost

This part follows the methodology of Li et al. (2009) in order to come up with monthly cost of real estate being used by IT infrastructure. Data centers take up considerable space and account for the real estate cost. Studies have shown that a 40W per square foot rated data center typically costs 400 Dollars per square foot (Anthes, 2005). However, there is a huge variation in prices based on various geographic locations and hence it has been taken as a generalized variable where area specific values can be captured by the organizations.

The real estate cost (Cost_RealE) is calculated based on cost of space taken by all the racks under utilization. Cost_RealE is given by (Cost_Sqf * Space_U_Rack), where Cost_Sqf is the cost per square foot to set up the physical servers and Space_U_Rack is the space taken by all the racks under utilization. Space_U_Rack can further be calculated as [(Rack_Sqf * N_Rack) / (Space_Rack)], where the description of variables involved is presented in Table-2. However, the final output targeted here is the amortized real estate cost -Cost_Am_RealE, and can be calculated as Cost_Am_RealE = [Cost_RealE * Ap_F] (Ap_F is Amortizable Parameter for Facility).

However, the calculations executed so far are subjected to a constraint that the value of pressure confronted by unit floor (V_Pressure) cannot be beyond the constant pressure confronted by unit floor (C_Pressure). V_Pressure can be calculated as (N_S (no. of servers) * W_S) + (N_Rack * W_R) / (Space_U_Rack), where the description of variables involved is given in Table-2. Therefore, the condition to be met for this calculation is: V_Pressure ≤ C_pressure.

Table 2: Real Estate Cost.

Name of the variable	Symbol	Unit
Square Feet per Rack	Rack_Sqf	Sq. feet
Number of racks	N_Rack	
Space utilization of Racks	Space_Rack	Percentage
Server Weight	W_S	Unit of Weight
Rack Weight	W_R	Percentage

Facility Cost

These represent both tangible and intangible components that are essential for the normal functioning of the equipment. These facilities are wrapped into racks which hold the servers. Hence, the TCO of facilities can be computed by segregating them into racks, so that the prices of facilities per rack can be taken as an input for cost computation. These facilities may include components like PDU, KVM (keyboard/video/mouse), cables etc.

The total facility cost -Cost_Tot_Fac, is given by N_Rack * Cost_Fac, where N_Rack is the number of racks and Cost_Fac is the cost of facilities per rack. However, facility cost must be amortized to calculate the amortized facility cost- Cost_Am_Fac. It can be calculated as Cost_Am_Fac= (Cost_Tot_Fac * Ap_S), where Ap_S is Amortizable Parameter for Server from 3.1.

Support and Maintenance Cost

Operational staff being the major category in an enterprise, the staff involved in maintenance of data centers is very small (Greenberg et al., 2008). The ratio of IT staff members to server is 1:100 in an established enterprise, automation is partial (Enck et al., 2009) and performance indicating problems are largely caused by the human error (Kerravala, 2002). After understanding the nature of support and maintenance cost in various organizations, it was found that majority of them outsource this job and the nature of job is documented in the contract. Hence, the computation of this cost incorporates the outsourcing part by taking into account the number of contract visits made in a year and the charges incurred during each visit. Thus, both internal personnel and contract human resources for preventive maintenance are included.

The total cost of support and maintenance (yearly) - Cost_Total_SM can be calculated using: Cost_Total_SM= (N_Admin * Salary_Support) + (N_Contract * Charge_Contract).
N_Admin represents the number of administrators responsible for support and maintenance in a firm, Salary_Support represents annual salary of administrators, N_Contract is the number of visits of contract maintenance and Charge_Contract is the cost per visit. Therefore, the monthly cost of support and maintenance- Cost_Am_SM can be easily computed as Cost_Total_SM /12.

Summary of the components in Layer-1

A total of nine components have been described in this section. This includes eight cost components and one component of amortization that is used frequently in other subsection. The purpose of clearly describing each component individually is that its output(s) will be used for computation in other layers and for calculation of costs in cloud computing. The output variable(s) for each component is summarized in Table-3.

Table 3: Cost component and associated output variable(s).

Component	Output Variable
Amortization	Ap_F and Ap_S
Cost of Servers	Cost_Am_S
Network Cost	Cost_Am_Net
Power Cost	Cost_Am_Pow
Software Cost	Cost_Am_Soft
Cooling Cost	Cost_Tot_Cool
Real Estate Cost	Cost_Am_RealE
Facility Cost	Cost_Am_Fac
Support and Maintenance Cost	Cost_Am_SM

LAYER 2: DATA PATTERN BASED COST ANALYSIS

This section deals with the idiosyncratic characteristics of data pattern in an organization like the amount of data it generates, the time taken by its computational resources to transfer the data, the estimated demand, the actual average demand and the number of servers provisioned to meet the demand. This analysis takes up such nature specific characteristics as input to the cost-benefit analysis model and comes up with two specific analysis, namely time analysis and demand estimation. The inputs in this layer will be used further in the project specific layer. Time analysis will give recommendation on adopting cloud computing based on the time required to process the data in house and comparing it with cloud. Demand analysis will give pros and cons of cloud computing with respect to the demand provisioning in an organization. There could be other possible aspects of data pattern cost analysis, however the scope of this study is limited to time and demand analysis only.

Time Analysis

It takes into account the amount of data being processed by an organization for all the operations combined. Based on the configuration of cloud instances mentioned in the cloud cost estimation section, an organization can find out the equivalence of computational ability in house to the cloud instance. The final results will reveal the computational time in-house as well as cloud. Hence, an organization can decide upon the shift to cloud computing based on the computational time. The transfer rate from organization to Amazon Cloud is taken as a variable, but is typically 20 MBits/second (Garfinkel, 2007).

The speed of EC2 instance takes 2 hours per GB to process the data (Fox et al., 2009). Hence to process *Size_Da* GB of data, it will take *2* Size_Da* hours. However, an organization can be considered equivalent to N_In number of instances and therefore *Size_Da* GB will actually take (2* *Size_Da*/ *N_In*) hours to process. Therefore the local computational time can be expressed as (2* *Size_Da*/ *N_In*). The transfer rate is expressed using the variable *Rate_Transfer* and the unit of measurement is Mbits/second. The interest of this study is finding out the time to transfer *Size_Da* GB of data. If *Rate_Transfer* Mbits requires 1 second, then *Size_Da* GB will require:

[(*Size_Da* * 1000 * 8) Mbits/ (*Rate_Transfer* Mbits/sec)] seconds. This leads us to the computational time in cloud- [[(*Size_Da* * 1000 * 8)/ *Rate_Transfer*]/ (60 * 60)] hours. The key formulas summarizing computation of processing time in cloud is given below:

1 GB takes 2 hours to process ➔ *Size_Da* GB will take 2* *Size_Da* hours to process
Size_Da GB will actually take ➔ (2* *Size_Da*/ *N_In*) hours to process (instances)
Local computational time ➔ *Time_Local*= (2* *Size_Da*/ *N_In*) hours
If *Rate_Transfer* Mbits requires 1 second ➔
Size_Da GB will require: [(*Size_Da* * 1000 * 8)/*Rate_Transfer*)] seconds
Hence, computational time in cloud➔
Time_Cloud= [[(*Size_Da* * 1000 * 8) / (*Rate_Transfer*)]/ (60 * 60)] hours
While *Time_Local* gives the computational time in house, *Time_Cloud* gives the computational time in cloud. Decision maker could compare the two and draw some insights on benefits of cloud computing in terms of computational time.
If (*Time_Local* < *Time Cloud*) then
Recommendation: It is advised not to shift to cloud computing in terms of computational time.
Else
Recommendation: It is advised to opt for cloud computing in terms of computational time.

Demand Analysis

This analysis concerns one of the most important problems of organizations i.e. provisioning of servers based on the demand. Generally firms prefer to provision the servers based on the maximum demand that can be estimated for a day. However, the average demand turns out to be less than one-third of the peak demand, thereby drastically making the computing resources inefficient. Studies have revealed that the real world estimates of utilization in data centers range from 5% to 20% (Rangan, 2008; Siegele, 2008) which is extremely low. If firms prefer to provision the server for the average demand, then they might lose customers for not providing service during times when actual demand is greater than the average demand. Hence, this analysis brings out the disadvantages of both under-provisioning and over-provisioning. These disadvantages would not come up in cloud computing as the payment is made for the exact usage of computing resources. It takes into consideration simple inputs like the estimated peak, average and trough demand along with the actual demand for which the server has been provisioned.

N_Peak is the no. of servers a firm would make a provision for in case of highest estimated demand or work load. On the other hand , *N_ Trough* is the no. of servers a firm would make a provision for in case of lowest estimated demand or work load. For example, a web-site with a highest estimated demand of 1000 users would require five servers to host it. However, if a firm decides to make a provision for the lowest demand i.e. 200 users it would require just two servers. Hence, N_Peak and N_Trough are five and two respectively. Ideally a firm should make a provision for average demand, but most of the firms end up making provision for peak demand to avoid the possibility of losing potential customers. *N_Average* demonstrates the no. of servers provisioned for handling average estimated demand.

Based on the Peak and Trough estimation of demand, the average demand is calculated as: *N_Average*= (*N_Peak* + *N_trough*)/2. The interest of this study is to find out the utilization factor in two cases. The first case being the ideal case, where utilization is calculated based on the average demand. While the second case being the real implementation scenario, where firms use *N_Server* servers to meet their demand. As already mentioned, firms generally keep *N_Server* high to meet the peak demand. The utilization factors in two cases can be computed by finding out the equivalent server hours.

The actual utilization over the whole day: *Actual_Utilization* = (*N_Average* * 24) Server Hours

However, the server is provisioned for N_Server demand estimate and hence the actual payment is made for: *Provisioned_Utilization*= (*N_Server* * 24) Server Hours

The next step would involve finding out the profitability of adopting cloud computing when compared to in house systems. This would require a comparison of *N_Server* and *N_Average*. In majority of the cases, the provisioned no. of servers is always greater than the average estimated server requirement to avoid the potential loss of customers. Therefore, this extra provisioning leads to potential loss of server hours and low utilization. The loss factor in this study has been calculated as the ratio of Provisioned utilization to Actual utilization. This factor is required to draw insights on cost calculation and comparison of in-house facility and cloud. The basic idea behind the comparison is derived from (Fox et al., 2009) which states that as long as cost per server hour in the cloud over 3 years is less than “Loss Factor” times the cost of buying the server, then it is profitable to opt for cloud computing. A time period of three years has been chosen because majority of the financial models allow a capital expense to be depreciated linearly over a three year period. Key comparison formulas in different cases are provided below.

```
If (N_Server> N_Average) then
    Loss in terms of Server Hours:
    Loss_ServerHours[1] = (Provisioned_Utilization- Actual_Utilization) Server Hours
    Loss_Factor= (Provisioned_Utilization / Actual_Utilization)
    If [Cost per server hour over 3 years] < [Loss_Factor * cost of buying the server]
        Recommendation: Shifting to cloud computing could be profitable
    Else
        Recommendation: Cannot decide upon the shift.
If (N_Server < N_Average) then
    Loss in terms of customers due to non-availability of service:
    Assuming one server caters to Y customers, then
    Customers lost= (N_Average – N_Server) *Y
    Out of these customers, some customers will be lost permanently which is assumed to be 10%. Added disadvantage would be the bad reputation of organization.
    Permanent_Customer_Loss= (.1 * Customers lost)
End
```

LAYER 3: PROJECT SPECIFIC COST ANALYSIS

This section deals with the most common scenario of cloud computing usage. Many firms today wish to preserve their existing infrastructure and evaluate whether it would be profitable for them to execute any upcoming project on cloud. The requirement of the upcoming projects in expressed as a ratio of existing computing infrastructure and hence the inputs of first layer (base cost estimation) can be used for this purpose. The number of estimated servers for the upcoming project is taken as in input and hence the ratio of estimate servers to the existing servers can be found out. This ratio can then be used to compute all the costs related to project which were calculated in layer one. The software cost (project related software) that was excluded in first layer is included in this analysis. This layer is not only useful for upcoming projects but also to evaluate the effectiveness of already implemented projects.

The estimation of cost is made keeping in mind the physical implementation of the project. The actual server requirement might be less than the estimated if exact demand is considered. There are already N_S servers in place. Therefore, the ratio of servers required for the project implementation is:

Ratio_Server= N_Est_Server / N_S, where N_Est_Server represents the number of estimated servers. The ratio can be used to compute all costs associated with the project using the costs calculated in base cost estimation section (Section 3). The costs of the project would be a percentage of the base cost (refer to Section 3 for associated variables). This percentage is *Ratio_Server*. Hence, the costs for the projects are:

Project Cost of Servers *(Cost_Am_PS) =Cost_Am_S * Ratio_Server*
Project Network Cost *(Cost_Am_PNet) = Cost_Am_Net * Ratio_Server*
Project Power Cost *(Cost_Am_PPow)= Cost_Am_Pow * Ratio_Server*
Project Cooling Cost *(Cost_Tot_PCool) =Cost_Tot_Cool * Ratio_Server*
Project Real Estate Cost *(Cost_Am_PRealE)= Cost_Am_RealE * Ratio_Server*
Project Facility Cost *(Cost_Am_PFac) =Cost_Am_Fac * Ratio_Server*
Project Support and Maintenance Cost *(Cost_Am_PSM) =Cost_Am_SM * Ratio_Server*

The only estimate not included in the above equations is Project Software Cost because it includes Class C software unlike Class A & B described in earlier sections. This would involve the cost of software involved in design, development and deployment of a project. The approach for calculation remains the same as that of Class A and B. Therefore, cost of the software associated with the project- Cost_Tot_PSoft is given by [N_ClassC * Cost_ClassC * ∏C]. N_ClassC represents the number of Class C software (project specific software), Cost_ClassC is the unit price of Class C software and ∏C is the software utilization (Class C). Therefore, amortized software cost Cost_Am_PSoft can be computed as (Cost_Tot_PSoft * Ap_F).

Time analysis can be performed using the input *Data_P* and it would give the recommendation based on computational time. It should be noted that *Data_P* is an approximate amount of data which includes the backup amount. Usually three replications are made for an ideal project as a backup policy. However, the same amount of data need not be required on the cloud because it has a provision for backup and the cost will not be incurred at the users' end. Hence, the real requirement that would be given as an input to the cloud would be: (*Data_P*/3) and it will be denoted by *Cloud_Data_P*. Hence, Cloud*_Data_P*= (*Data_P*/3)

CLOUD COMPUTING COST ANALYSIS

The cloud computing service taken for analysis in this paper is Amazon EC2. It provides a built-in calculator (Amazon, 2010) to provide a monthly bill based on the inputs given to calculator. The variables in this analysis are basically the inputs to the calculator. The process illustrated in this section will provide guidelines to the firms to calculate these variables for their organization. The method has been inspired from cost calculations of Gray (2008) and Fox et al. (2009). These variables are then fed into the calculator that provides the monthly billing of using Amazon services which can be then compared to the base cost or project specific cost. The exact method of comparison in discussed in upcoming sections where the overall model of cost-benefit analysis is explained.

Cloud services like Amazon leverage upon economies of scale. The ratio of cost in large data centers to that of medium size data centers (Network, Storage and Administration) vary between 5 to 7 (Hamilton, 2008b).

There are different families of Amazon EC2 instances including Standard, Micro, High-Memory etc. Each instance provides a predictable amount of dedicated compute capacity and is charged per instance-hour consumed. More details about the family of instances can be obtained from (Amazon, 2011). For the purpose of cost-benefit analysis Standard Instances (On Demand Instances) were selected due to its common usage in most of the applications. It can be later extended to high memory instances and high CPU instances. There are three sub-classes in the family of standard instances, the description of which is given below:

- **Small Instance (Default)** 1.7 GB of memory, 1 EC2 Compute Unit (1 virtual core with 1 EC2 Compute Unit), 160 GB of local instance storage, 32-bit platform
- **Large Instance** 7.5 GB of memory, 4 EC2 Compute Units (2 virtual cores with 2 EC2 Compute Units each), 850 GB of local instance storage, 64-bit platform
- **Extra Large Instance** 15 GB of memory, 8 EC2 Compute Units (4 virtual cores with 2 EC2 Compute Units each), 1690 GB of local instance storage, 64-bit platform

The organization can match their server configuration with these instance configurations. By default, one large instance is considered equivalent to one physical in house-server. Therefore, if the no. of estimated physical servers required for a particular project is *N_Est_Server*, then *N_Est_Server* Large instances of Amazon EC2 would be required. *N_Instance* represents the number of instances required on cloud and it would thus be equal to *N_Est_Server* large instance. Also, based on memory requirements and local instance storage, it can be said that one large instance is almost equivalent to five small instances and one extra large instance is equivalent to two large instances.

Hence, N_Instance large instance= 5 * N_Instance small instance
N_Instance large instance=(N_Instance/2) extra large instance

Therefore, the input can be given as a choice between small instance, large instance and extra-large instance. However, there is an option in the calculator for matching the physical server to cloud instance equivalence, given the server configuration of an organization. This study makes a provision for incorporating the server equivalence other than the specified method and is represented by the variable *Server_Equivalence*. If this variable is used and the default conversion is not made, then:

1 Physcial Server= *Server_Equivalence* large instances
N_Est_Server Physical Servers= (*N_Est_Server* * *Server_Equivalence*) large instances

Therefore, *N_Instance* should be adjusted accordingly and conversion to other form of instances can be made easily.

Data Processing Requirements: As already discussed in the project specific cost estimation, the data processing requirements for the cloud would be one third as that of the cloud due to automatic backups being taken by cloud service providers. Hence, *Cloud_Data_P* (data processing requirement in cloud) which is given by *Data_P*/3 is the data that will account for bandwidth cost per month (*Data_P* comes from the Section 5-Layer 3).

Cloud_Data_P needs to be distributed as data in, out and regional data transfer in terms of GB/Month. Therefore, Cloud_Data_P= Data Transfer in + Data Transfer Out + Regional Data Transfer + Public IP/Elastic IP Data Transfer (Public IP Is not used for analysis). This is required because the pricing for data transfer in and out is different in case of Amazon EC2. The price slabs in Table 4 elaborates on the same issue and is based on data transferred "in" and "out" of Amazon EC2.

Table 4: Price Slabs in Amazon.

Data Transfer In	**Rate**
All Data Transfer	Data Transfer In will be $0.10 per GB
Data Transfer Out	**Rate**
First 10 TB per Month	$0.15 per GB
Next 40 TB per Month	$0.11 per GB
Next 100TB per Month	$0.09 per GB
Over 150 TB per Month	$0.08 per GB

Regional Data Transfer: $0.01 per GB in/out – all data transferred between instances in different Availability Zones in the same region. (If regional data transfer is not provided by the user, then it is assumed that it accounts to 30% of the total data being processed)

Public and Elastic IP and Elastic Load Balancing Data Transfer: $0.01 per GB in/out: Not considered in this analysis because Regional Data Transfer rates payment need to be made even if the instances are in the same Availability Zone and this cannot be avoided.

Auto Scaling: Auto Scaling is enabled by Amazon CloudWatch and carries no additional fees. Each instance launched by Auto Scaling is automatically enabled for monitoring and the Amazon CloudWatch monitoring charge will be applied.

Elastic Load Balancing: Elastic Load Balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances. It enables the users to achieve even greater fault tolerance in their applications, seamlessly providing the amount of load balancing capacity needed in response to incoming application traffic. It is mentioned in the Amazon Calculator guidelines that a medium-sized website running on 10 Amazon EC2 instances in the US East (N. Virginia) Region could use one Elastic Load Balancer to balance incoming traffic. Therefore, this study derives the number of load balancers based on no. of instances booked.

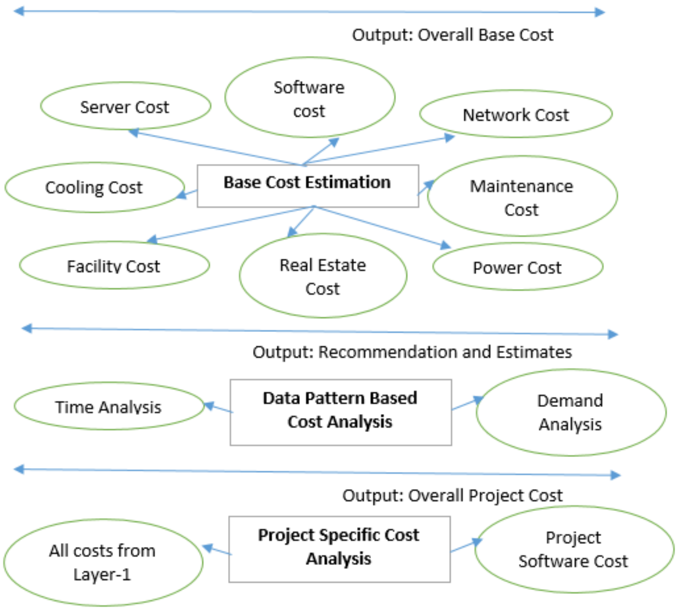
$$N\_LoadB = \text{ceil}\,[N\_Instance/10]$$

Also, at it is assumed that at least 70% of the estimated data transfer would actually take place.

Hence, $Data\_LoadB = (.7 * Cloud\_Data\_P)$

Figure 1: Three Layer Model for Cost-Benefit Analysis.

THREE LAYER MODEL



The different components involved in three layers and also the cost computation in cloud has been discussed in previous sections. This section presents the model that would be used by organizations to conduct the cost-benefit analysis in three layers. The representation of the model is shown in Figure 1.

Decision in Layer 1**:** As shown in Figure 1, the output of Layer 1 is the overall base cost. This represents the monthly cost of maintaining IT infrastructure in an organization (in-house) based on parameters that could help in deciding upon cloud computing options. This overall base cost is compared with the final cost of cloud given by Amazon Calculator. Organizations may take a decision to outsource the entire infrastructure to cloud if it turns out to be profitable. Irrespective of it being profitable, further analysis can be done in Layer 2 and Layer 3 which uses organization specific and project specific information respectively.

Decision in Layer 2**:** This layer is independent layer and hence the outputs from this layer will not be compared with the cloud costs shown by calculator. The two analyses used in this layer give direct recommendation on adoption/shifting to cloud computing. Hence, the decision can be made from two perspectives: a) computational time (Time Analysis) and b) Losses if any by not shifting to cloud platforms based on demand requirements (Demand Analysis).

Decision in Layer 3**:** Figure 1 shows overall project cost as the output of Layer 3. This cost is obtained using the all the cost components in Layer 1. Instead of the total servers in-house for an organization, the number of estimated servers required for a project under consideration would be the input to cloud computing cost estimation model. The next step would involve the comparison of overall project cost to the output given by Amazon Cloud cost calculator. If the cost on cloud is less than the in-house project computation, then the decision would be to execute the project on cloud.

MODEL IMPLEMENTATION & DATA COLLECTION

This model was tested by its application in IT firms. Indian IT firms were chosen for this study because of the scope for cloud implementation in India. India has grown as a hub of IT and ITeS services in the global picture. It accounted for 55% of the global sourcing market in 2010 and NASSCOM (The National Association of Software and Services Companies) data suggests that Indian software and services industry aggregated revenues of US $100 billion in FY2012 growing by over 9%. With latest trends demonstrating a rise in the performance of the IT sector in India, cloud computing has a major role to play in helping the India IT sector remain ahead in the global market.

The objective of the model testing and application was twofold: a) to test the consistency of the in-house cost computations (derived from the model) with the with the actual cost incurred to the firm and b) to derive a preliminary assessment of the relationship between the type of organization (small/big firms) and the profitability of adoption of cloud computing. In order to meet the second objective, it was necessary to classify the organizations into various categories. This classification was done based on the number of servers, the description of which is provided in Table 5. Though the classification of organizations based on number of servers is derived from literature (Misra & Mondal, 2011), it is slightly modified to fit the context of data collection.

Table 5: Profile of Sample Firms (Statistics).

NUMBER OF SERVERS			
S. No	Name	Number of Servers	No. of Organizations
1	Class-A: Small	Less than 100	11
2	Class-B: Medium	From 100 to 1000	8
3	Class-C: Large	From 1000 to 10,000 servers	6
4	Class-D: Very Large	From 10,000 to 50,000 servers	3
5	Class-E: Super Large/Giants	50,000 above	2
	Total		30
INDUSRTY REPRESENTATION			
S. No	Sector		No. of Organizations
1	Part-A: Computer and Electronic Product manufacturing		14
2	Part-B: IT Services		16
	B.1 Data Processing		7
	B.2 Hosting and Related Services		4
	B.3 Other Information Services		2
	B.4 Software Publishers		3
ANNUAL REVENUE			
S. No	Category		No. of Organizations
1	SMALL: Less than $ 100 million		16
2	MEDIUM: Between $ (100-500) million		11
3	LARGE: Above $ 500 million		3

Procedure for data collection: Emerging Market Information Service database (EMIS-product of ISI emerging market) was chosen to find the list and contacts of IT Services/Manufacturing firms owing to its comprehensive collection and exclusive details. The database consisted of 1047 firms and the entire set was divided into three groups based on their annual revenue - Small, Medium and Large. In order to maintain the diversity in sample organizations, it was decided to randomly select firms from each of the three groups instead of random selection from the entire database. Hence, twenty five firms were chosen randomly from the three groups. Therefore, a total of 75 firms formed the sample set for this study.

Data collection was carried out in two stages: a) collection of data points in the model from secondary sources b) collection of remaining data points using an online questionnaire (primary data). The former stage was used to collect data points required in the model from secondary sources like company web-site, blogs containing information about TCO, articles using company data for cost calculations and other sources available on web. Second stage was used only for firms, the data of which could not be obtained in the first stage.

The second stage was accomplished using well designed online questionnaire that captured all the components of the model. Though authors initially developed an application (software) for the purpose of data collection, many organizations did not permit the installation of the same.

Therefore, online survey was designed using Google Forms and it was mailed to the firms in the sample set. The contacts of these firms were procured using EMIS database. Clear instructions to capture the data points were provided in the covering letter and a summary of data points was also included. The initial contact in each firm was requested to nominate an individual (preferably system administrator) who would be in a suitable position to respond, based on the summary included in the cover letter. In some cases, the initial contacts forwarded the mail to the nominated individuals whereas in other cases, a new e-mail was sent to the nominated members. There was a provision given in the survey to skip certain components/variables and directly enter the overall cost if the calculation method of the firm was different from the model. For example, the overall value of costs like Network Cost and Software Cost was directly reported by the firms in most of the cases.

With a total of seventy five firms in the sample set, complete responses could be obtained from 30 firms. While there was no response from forty one firms, incomplete data was the cause of rejection for four firms. Thus, the response rate for the data collection turned out to be 40%. The profiles of these organizations (final set) are presented in Table-5. The firms in the sample set fairly represented the IT industry (Table-5). It covers the two broad sections of IT industry mentioned in the Emerging Market Information Service database (EMIS-product of ISI emerging market)- Part-A: ‘Computer and Electronic Product manufacturing’ (14 firms) and 2) Part-B: ‘IT Services’ (16 firms) that includes- Data Processing, Hosting and Related Services, Other Information Services, Software Publishers. The amount of revenue generated by the firms plays an important role in deciding whether they should maintain the datacenters or should adopt cloud computing. Therefore, the distribution of firms across different categories based on annual revenue is also presented in Table-5. It represents a fair distribution across small and medium categories with participation from large category as well.

Though data was available to compute Project Specific Cost (Layer 3) for individual organizations, this study did not compute the results for Layer-3. It was excluded because it would not be possible to evaluate the results on a common scale as different organizations have idiosyncratic project requirements. Hence, comparison was performed at Layer 1 and 2 for all the organizations to provide a common platform for analysis. In order to test the robustness of the model, in-house computational cost in Layer 1 of the model was then compared with the internal cost calculation of the concerned organizations. It was found that in-house computational cost was fairly consistent and approximately represents the true cost of in-house server scenario. The average deviation between the computation cost of the organization and in-house computation cost in Layer 1 presented in this study was well within the accepted limits (monthly costs). This was confirmed by conducting an independent t-test as the results did not show any statistically significant difference between the two methods (5% level of significance). This consistency provides a strong foundation to the proposed model in terms of decision making in Layer 1. The next section would outline key findings of the study with some examples.

RESULTS

Before proceeding with the aggregate results for all the organizations in the sample set, detailed results for two organizations have been presented as an illustration. First organization (Setup-1) belongs to ‘Very Large scale’ category (D) while the second organization (Setup-2) is small scale data center (Category-A). For the first setup, the cost of computation on cloud was $8649249 per month, while that of the in-house computational facility was $8625623 per month. This demonstrated that in-house computational facility turned out to be cheaper for a large scale data center. However, results were different for the second setup (small scale data center). In this case the cost of development in house was $2,009.43 per month, while the cost of computation on cloud turned out to be $567.01 per month. Hence, this demonstrated a considerable savings of 254.39%, revealing that the current expenditure was more than twice the cost incurred on cloud. Similar conclusions were drawn using the analysis at Layer-2 of the proposed model. Only time analysis was performed at the second layer as it was difficult to collect the data for demand analysis. Small scale data center had higher transfer time locally when compared to cloud computing and hence the decision for shifting to cloud computing was favorable. The final output for both the setups is shown in Table 6.

Table 6: Detailed results for two sample organizations.

SETUP 1 (Large Scale Data Center) Layer 1		Layer 2 (Time Analysis)
In-House computation	$8,625,623.31	19 hours (equivalent of 320 instances locally)
Cloud Computation	$8,649,248.66	27 hours
Decision	Do NOT Shift to Cloud Computing	Do NOT Shift to Cloud Computing
Cost Difference	$23,625.35	Time Difference: 8 hours
SETUP 2 (Small Scale Data Center) Layer 1		Layer 2 (Time Analysis)
In-House computation	$2,009.43	59 hours (equivalent of 20 instances locally)
Cloud Computation	$567.01	47 hours
Decision	SHIFT to Cloud Computing	SHIFT to Cloud Computing
Cost Difference	$1,442.42	Time Difference: 12 hours

Considering the entire sample set of organizations, the profitability varied in different categories of firms. While it was profitable for all the eleven firms in Category-A to go for cloud computing (Layer 1 decision level), it was not profitable for a minority of firms in Category-B. On the other hand, Category-C firms had just one firm benefiting from cloud computing while it was not profitable for the others. All the firms in Category-D and E registered a loss when shifting to cloud computing was concerned. Almost similar results were also observed for Layer-2 analysis. Specific details like mean time difference have been avoided because it each organization have different *Size_Da* and therefore mean difference is not relevant. The summary of profitability is presented in Table 7.

Out of the given sample set of organizations, it is clear that it in general, it is profitable for Small/Medium scale enterprise to opt for cloud computing (Table-7). Only two out of the nineteen organizations in this category (A and B combined) did not register profitability. Further, the mean loss made by these two firms is very small ($513.25) as compared to the mean profitability by their counterparts ($4,512.34 for 6 firms). When large organizations are concerned, the shifting to cloud architecture proved profitable for only one organization out of the eleven firms (Class C, D and E Combined). Further, the profitability registered by that organization was very small ($468.45) compared to the losses made by its counterparts ($9,468.45 for 5 firms). Thus, it is evident that there is a clear distinction of profitability in the two cases of small/medium enterprises and large/very large/super large organizations. The study and model development is at its infancy stage, therefore it would be inappropriate to make generalized statement about relationship between size of organization and profitability. Hence, exploring the relationship in detail forms the future part of this study.

Table 7: Profitability in Layer-1 and 2 (Shifting to Cloud Computing).

Category	No. of Organizations	No. of firms registering PROFIT. (Layer-1)	Percentage of Firms registering profitability	Mean Cost Difference for profitable firms	Mean Cost Difference for non-profitable firms	No. of firms with positive decision for Cloud (Layer-2)
Class-A: Small	11	11	100%	$1,892.65	NA	11
Class-B: Medium	8	6	75%	$4,512.34 (6 firms)	$513.25 (2 firms)	5
Class-C: Large	6	1	16.67%	$468.45 (1 firm)	$9,468.45 (5 firms)	2
Class-D: Very Large	3	0	0%	NA	$16,546.50	0
Class-E: Super Large/Giants	2	0	0%	NA	$29,713.85	0
Total	30	18				19

The reasons for this difference in profitability could also be explored in the future work. However, this study attempts to find a logical reason for the results obtained. The non-profitability of cloud computing in the large scale data center could be because of economies of scale. Owing to the fact that the setup has massive server base, it would be more profitable to continue with in-house investment/computation and reap the benefits of economies of scale. Further, firms with large scale data centers operate in different countries, thereby having a greater pool of internal resources leading to reduced operational costs. However, in the other case cloud computing turned out extremely profitable because of the setup being a very small scale data center. This is consistent with research at practitioner's end, where cloud computing has been shown as a blessing to the start-up firms probably because of uncertainty in demand or their inability to reap benefits by economies of scale. An interesting observation would be the approximate line of difference between startup and large scale enterprise that creates the variance in profitability (future scope).

It should be noted that the results summarized in Table-7 has dealt only with the first layer and time analysis of second layer. It shows the initial comparison of the entire infrastructure in house versus cloud computing and the time analysis in the second layer. Though data was not sought for Layer-3, there was a particular case where computation on third layer was requested by a firm. The firm in second setup shown in Table-6 (small scale data center), wished to evaluate the possibility of executing its upcoming project on cloud. For this purpose, project specific analysis (Layer-3) was conducted with an estimate of two servers required for the upcoming project as reported by the firm. It was found that the project, if executed in-house would lead to overall charges of $401.00 per month. However, cloud computing again demonstrated considerable benefits leading to charges of $135.00 per month. Hence, the profitability of executing the project on cloud in terms of computing requirements would be twice when compared to the in-house computation. This particular case is presented only to give a fair idea on how Layer-3 in the proposed model could be utilized for evaluation purpose. However, there is no comparison made between the organizations at Layer-3 owing to the idiosyncrasies of sample firms.

LIMITATION OF THE STUDY AND FUTURE WORK

The billing model of cloud computing used in this paper is based on a single vendor (Amazon). Hence, the applicability is limited to decision making in Amazon Cloud only. Further, the model is tested on a small sample size and hence the results could be further enhanced for better generalizability. Though this study has included diverse set of organizations, there could be equal number of organizations in all proposed categories (A, B, C and D) for better results. One of the major limitations of this study is the exclusion of result aggregation at Layer-3 and demand analysis at the second layer. Though justification has been provided for the exclusion, the model would have been better analysed had the results been presented for all three layers. However, extending the application to third layer and conducting the demand analysis forms the first step of future research work. It should also be mentioned that there are plethora of issues when implementation of cloud computing is concerned. This would include security, trust, efficiency, know-how of in house server maintenance and several other factors. However, the focus of this study is the costing in cloud that would aid managers in decision making. Therefore, other factors have not been considered in the recommendations made on cloud computing adoption and thus forms another limitation of this study.

Future work would include use of analytical methods like regression to find out the relationship between size of organization and profitability. Researchers could make use of this model to increase the number of independent variables to find out its relation with decision to adopt cloud computing. The three layer model will be extended to multiple layers in order to provide more flexibility for managers to take a decision on shifting to cloud computing. Further, the model will transcend Amazon cloud computing service and provide options for computation on other services as well.

CONCLUSION

This paper has addressed one of the major issues faced by many organizations on their decision to shift to cloud computing. It has attempted a three layer approach in order to incorporate maximum flexibility in the model for easy computation. A comparison across different organizations was accomplished with respect to the profitability of shifting to cloud computing. From the results obtained for thirty organizations, it was found that cloud computing is profitable for small/medium scale enterprise. Large scale enterprise did not benefit from shifting to cloud architecture. This model is being converted into a readymade tool in which the firms can directly feed in data and get recommendations on shifting to cloud computing. Further research would be to analyze the trend in profitability of organizations shifting to cloud. This model could be used to understand the relation between projected profitability and the nature of an organization. Reduction of IT infrastructure costs using cloud computing could be an effective contribution to Green IT.

REFERENCES

Abramson, D., Buyya, R., & Giddy, J. (2002). A computational economy for grid computing and its implementation in the Nimrod-G resource broker. *Future Generation Computer Systems*, *18*(8), 1061-1074.

Amazon (2010), Amazon_Billing Calculator, available at: http://calculator.s3.amazonaws.com/calc5.html (last accessed May 2010)

Amazon (2011) Amazon_EC2 Instance types, http://aws.amazon.com/ec2/instance-types.

Anthes, G . (2005). Data Centers Get a Makeover. *Computerworld news article* (cited 2005 November 1), Available from http://www.computerworld.com/databasetopics/data/datacenter/story/0,10801,97021,00.html?SKC=home97021

Belady, C., Rawson, A., Pfleuger, J., & Cader, T. (2008). Green grid data center power efficiency metrics: PUE and DCiE. *The Green Grid*, 1-9.

Bezos, J. (2007). Keynote talk at Web 2.0 Expo (2007).

Brill, K. (2009). The Invisible Crisis in the Data Center: The Economic Meltdown of Moore's Law. *Uptime Institute White Paper*.

Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation computer systems*, *25*(6), 599-616.

Cappuccio, D., Keyworth, B., & Kirwin, W. (1996). Total cost of ownership: The impact of system management tools. *Gartner Group, Stamford, CT*.

Cybersource. (2002). Linux vs. Windows Total Cost of Ownership Comparison. *Cybersource® Pty. Ltd*.

Deelman, E., Singh, G., Livny, M., Berriman, B., & Good, J. (2008). The cost of doing science on the cloud: the montage example. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing* (p. 50). IEEE Press.

Enck, W., Moyer, T., McDaniel, P., Sen, S., Sebos, P., Spoerel, S., & Aiello, W. (2009). Configuration management at massive scale: system design and experience., *IEEE Journal on Selected Areas in Communications*, *27*(3), 323-335.

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008). Cloud computing and grid computing 360-degree compared. In *Grid Computing Environments Workshop, 2008. GCE'08* (pp. 1-10).

Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., ... & Stoica, I. (2009). Above the clouds: A Berkeley view of cloud computing. *Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS, 28*.

Garfinkel, S. L. (2007). An evaluation of Amazon's grid computing services: EC2, S3, and SQS.

Ghag, S. (2008). Financial implications of the cloud, Infosys Microsoft Alliance and Solutions Blog, Available from http://www.infosysblogs.com/microsoft/2008/12/financial_implications_of_the_1.html

Google (2010). Commitment to Sustainable Computing, available at: http://www.google.com/corporate/data centers/, Oct. 2008 (last accessed May 2010).

Gray, J. (2008). Distributed computing economics. *Queue*, *6*(3), 63-68.

Greenberg, A., Hamilton, J., Maltz, D. A., & Patel, P. (2008). The cost of a cloud: research problems in data center networks. *ACM SIGCOMM Computer Communication Review*, *39*(1), 68-73.

Hamilton, J. (2008). Cost of power in large-scale data centers. *Blog entry dated*, 28th November.

Hamilton, J. (2008b). Internet-scale service efficiency. In *Large-Scale Distributed Systems and Middleware (LADIS) Workshop (September 2008)*.

Hamilton, J. (2009, January). Cooperative expendable micro-slice servers (CEMS): low cost, low power servers for internet-scale services. In *Conference on Innovative Data Systems Research (CIDR'09)*.

Hayes, B. (2008). Cloud computing. *Communications of the ACM*, 51(7), 9-11.

Hinchcliffe, D. (2012), What does cloud computing actually cost? An analysis of the top vendors, ebiz: the Insider's Guide to Business and IT Agility, Available from http://www.ebizq.net/blogs/enterprise/2009/08/what_does_cloud_computing_actu.php.

Hosanagar, K., Krishnan, R., Smith, M., & Chuang, J. (2004, January). Optimal pricing of content delivery network (CDN) services. In *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on* (pp. 10-pp).

Kerravala, Z. (2002). Configuration management delivers business resiliency.*The Yankee Group*.

Kirkpatrick, M. (2006). Amazon releases early info on S3 storage use.

Li, X., Li, Y., Liu, T., Qiu, J., & Wang, F. (2009, September). The method and tool of cost analysis for cloud computing. In *Cloud Computing, 2009. CLOUD'09. IEEE International Conference* (pp. 93-100).

McFarlane, R. (2005). Let’s Add an Air Conditioner, SearchDataCenter news article, Available from ttp://searchdatacenter.techtarget.com/columnItem/0,294698,sid80_gci1148906,00.html.

Mircea, M., Ghilic-Micu, B., & Stoica, M. (2011). Combining business intelligence with cloud computing to delivery agility in actual economy. *Journal of Economic Computation and Economic Cybernetics Studies, 45*(1), 39-54.

Misra, S. C., & Mondal, A. (2011). Identification of a company’s suitability for the adoption of cloud computing and modelling its corresponding Return on Investment. *Mathematical and Computer Modelling, 53*(3), 504-521.

Nucleus. (2009). ROI Case Study, Nucleus Research.com-Document, Available from http://www.google.com/apps/intl/en/business/case_studies/tvr.pdf.

Palankar, M. R., Iamnitchi, A., Ripeanu, M., & Garfinkel, S. (2008, June). Amazon S3 for science grids: a viable solution?. In *Proceedings of the 2008 international workshop on Data-aware distributed computing* (55-64). ACM.

Pyke, J. (2009). Now is the time to take the cloud seriously. *White Paper, Retrieved from: www. cordys. com/cordyscms_sites/objects/bb1a0bd7f47b1c91ddf36ba7db88241d/time_to_take_the_cloud_seroiusly_online_1_. pdf.*

Rangan, K., Cooke, A., Post, J., & Schindler, N. (2008). *The Cloud Wars: $100+ billion at stake*. Tech. rep., Merrill Lynch.

Rasmussen, N. (2007). Calculating total cooling requirements for data centers.*American Power Conversion, white paper*, (25).

Rimal, B. P., Choi, E., & Lumb, I. (2009, August). A taxonomy and survey of cloud computing systems. In *INC, IMS and IDC, 2009. NCM'09. Fifth International Joint Conference* (pp. 44-51).

Rosenberg, D. (2012). The cost of cloud adoption, Available from http://news.cnet.com/8301-13846_3-10140303-62.html.

Sawyer, R. (2004). Calculating total power requirements for data centers.*American Power Conversion, Tech. Rep*.

Siegele, L. (2008). *Let it rise: A special report on corporate IT*. Economist Newspaper.

Simson, G. (2007). Commodity grid computing with Amazons s3 and ec2, In login USENIX.

Singh, G., Kesselman, C., & Deelman, E. (2007, June). A provisioning model and its comparison with best-effort for performance-cost optimization in grids. In *Proceedings of the 16th international symposium on High performance distributed computing* (pp. 117-126). ACM.

Stuer, G., Vanmechelen, K., & Broeckhove, J. (2007). A commodity market algorithm for pricing substitutable Grid resources. *Future Generation Computer Systems*, *23*(5), 688-701.

Walker, E. (2008). Benchmarking Amazon EC2 for high-performance scientific computing. *Usenix Login*, *33*(5), 18-23.

Xiong, K., & Perros, H. (2009, July). Service performance and analysis in cloud computing. In *Services-I, 2009 World Conference*, 693-700).

Zhao, H., & Sakellariou, R. (2007, January). Advance reservation policies for workflows. In *Job Scheduling Strategies for Parallel Processing* (pp. 47-67). Springer Berlin Heidelberg.

This Page was Intentionally Left Blank