CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. (2013)
DOI: 10.1002/cpe.3086

SPECIAL ISSUE PAPER

Toward an economic and energy-aware cloud cost model

Werner Mach and Erich Schikuta*,†

University of Vienna, Research Group Workflow Systems and Technology, Währingerstr. 29, 1090 Vienna, Austria

SUMMARY

This paper represents an economic cost model for cloud computing aiming at comprising all kinds of cost of a commercial environment. To extend conventional state-of-the-art models considering only fixed cost, we developed a concise but comprehensive analytical model, which includes also variable cost allowing for the development and evaluation of business strategies for cloud environments. These strategies can be used for both cloud providers and cloud consumers. The major goal of our model is to comprise all important economic fundamentals and methods. Thus, this new model supports the decision-making process to be applied with business cases and enables cloud consumers and cloud providers to define their own business strategies and to analyze the respective impact on their business. On the basis of this model, also, the energy efficiency of cloud systems can be evaluated according to chosen business models. Copyright © 2013 John Wiley & Sons, Ltd.

Received 22 May 2013; Accepted 4 June 2013

KEY WORDS: cloud computing; cost model; business strategies

1. INTRODUCTION

Many different definitions on cloud computing can be found in literature. A very profound one is that ‘cloud computing is grid computing with a business case’. This implies that the innovative part of cloud computing is its economic face. The technology aspects are not really new and mostly adapted from the service-oriented landscape and grid computing. There was a research work of the authors on business aspects in grid computing as well as [1–3], but basically, this facet was neglected by the research community. However, cloud computing clearly identifies economic issues, as costs, revenues, return of investment, and so on and also work on derived issues, as security [4], privacy [5, 6], anonymity [7, 8], and so on is numerous.

In this paper, we present a novel economic cloud cost model to allow for definition and realizations of business strategies for both cloud consumer and cloud provider. The novelty of this model is to facilitate the definition of specific business strategies based on clear parameters. To the best of the authors’ knowledge, no cost model for cloud computing has so far been developed that makes a specific distinction between variable and fixed cost [9–15]. Moreover, all papers lack a generic formula representing the cloud’s costs. The cloud cost calculation from [10] focuses on cloud computing consumers and is based on Amazon’s pricing model. Also, Armbrust *et al.*[12] neglects the distinction between fixed and variable costs and focuses on cost optimization for the cloud consumers only. Hence, both [10] and [12] disregard the cost structure of a cloud provider. However, [14] concentrates on cloud data centers; but neither fixed nor variable costs are considered.

---

*Correspondence to: Erich Schikuta, University of Vienna, Research Group Workflow Systems and Technology, Währingerstr. 29, 1090 Vienna, Austria.

†E-mail: erich.schikuta@univie.ac.at

Cost models that consider variable cost are better suited for the economy situation today. This distinction allows also to implement more accurate business strategies. Traditional business models are largely on the basis of fixed-cost operating models. These operating models are driven of large capital investments to leverage economies of scale to produce incremental profit in the case of increasing volume. This results in spreading operating cost to larger and larger units sold. The prediction of the products' sales is stable enough to allow companies to allocate labor and capital in order to support demand.

Variations of demand can only be compensated for low frequencies. In the past, typical product life cycles were measured in years; therefore, this kind of operating models could be used. However, product life cycles are shortened from years to months. Also, rapidly evolving consumer preferences in global markets require more flexible cost models to give quick answers to changing demands. There is a specific need for cloud computing today, where only dynamic, adaptive, and precise business decisions allow for economic success of this new technology trend.

In the course of our discussion, we focus on typical business questions of cloud providers. However, questions of cloud consumers can be answered by our analysis too. For example, we aim for giving answers to questions such as 'What is the optimal decision if an SLA violation happens"?.

The paper is organized as follows. In the next section, the underlying economic framework of the model is defined. Then, the description of the model and its parameters are presented. In Section 3, a comprehensive analytical formulation, evaluation, and the applicability to traditional economic methods is given. Furthermore, the energy efficiency of cloud systems is evaluated according to chosen business models. The paper is closed by a conclusion and a presentation of topics for further research.

2. A SHORT INTRODUCTION TO ECONOMIC METHODS

In this section, we give a short introduction to traditional economic fundamentals and methods described in the standard economic literature [16–18]. This section is mainly derived from [16] and [17]. We restrict the explanations to terms, which are necessary to understand our cloud cost model. Traditional economic science covers the following scopes:

- Operating production factors
- Production
- Sales theory
- Investment and finance

We map all these elements of a traditional production company into a model to represent a common cloud environment. In this paper, we focus on operating production factors and the production. However, our model covers all elements.

2.1. *Operating production factors*

For producing goods it is necessary to combine production factors. Production factors are, for example, manpower, machine employment, materials, energy, and auxiliary materials. Each production factor is named $r_1 \dots r_n$.

The theory of production is concerned with the functional dependencies between the amount of production factors and the amount of produced goods. These functional dependencies can be described in production functions. The goal is to find regularities between input and output of the factors under strongly simplified assumptions [16, 19]. On the basis of these simplified assumption, cost theory tries to find the functional dependencies between the amount of production factors used and the resulted cost. The amount of output m is a function of the quantity of the input factors $r_1 \cdots r_n$.

$$m = f(r_1, r_2 \dots r_n) \tag{1}$$

Therefore, the revenue E is a function of the production factors too.

$$E = f(r_1, r_2 \dots r_n) \tag{2}$$

DOI: 10.1002/cpe

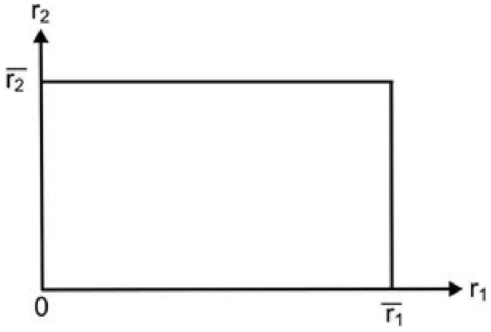

Figure 1. Substitutability of production factors.

Depending on the production process, complex production functions can exist. The relationship between the production factors can be reduced to two general cases.

1. The production factors can be substituted by each other (substitutability).
2. The production factors can only be used in the same proportion (limitationality).

Substitutability. Figure 1 depicts two production factors r_1 and r_2. The area of the rectangle between $\overline{r_1}$ and $\overline{r_2}$ represents all possible combinations of r_1 and r_2. An amount of revenue E can be assigned to each point in the area. If we consider the rectangle as plane, we can draw the corresponding revenue E vertically to each point. This leads to a revenue mountain. For better understanding, we introduce the term *marginal revenue*.

The *marginal revenue* is the increment of amount of the revenue that can be achieved with an (theoretically) infinite increase of the usage of a factor. Typical forms of such revenue mountains exist (see Figures 2 and 3) as follows:

- Constant marginal revenue.
- Increasing marginal revenue.
- Decreasing marginal revenue.
- Increasing and then decreasing marginal revenue.

Limitationality. Limitationality exists if the production process does not allow to substitute one production factor by another. These factors are also called limited production factors.

If we have production factors that can be substituted by each other until the amount of one factor is 0, we have the case of *alternative substitution*. We can use the production factors alternatively. If we have a combination process of using r_1 and r_2 with a minimum amount of one of the factors, we call it *limited substitution* (see Figure 4).

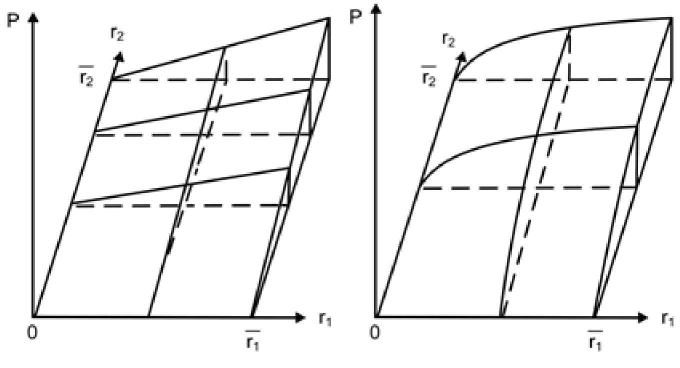

Figure 2. Constant marginal revenue (left), decreasing marginal revenue (right).

DOI: 10.1002/cpe

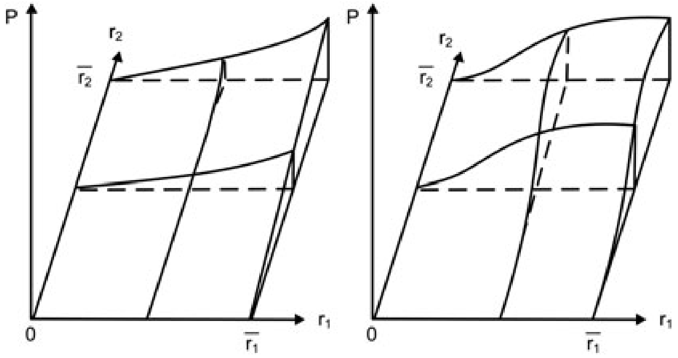

Figure 3. Increasing marginal revenue (left), increasing and decreasing marginal revenue (right).

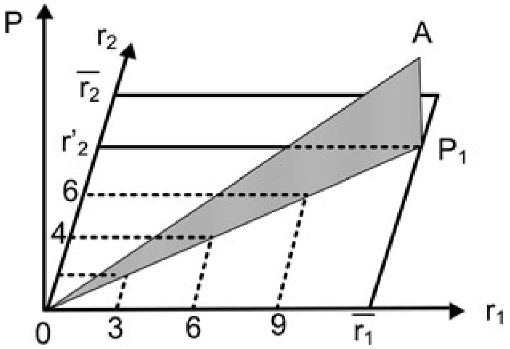

Figure 4. Limited substitution.

2.2. *Production*

Relationship between production-function and cost-function. Costs are the priced amount of production factors. As defined in Equation 2, the revenue is a function of the amount of used production factors. The functional dependency of overall revenue to factor input is defined by the cost function

$$C = f(m) \tag{3}$$

From Equation 2, we can derive the total revenue if we rate all production factors with money.

$$E_t = f(r_1 \cdot p_{r1}, r_2 \cdot p_{r2}, \cdots, r_n \cdot p_{rn}) \tag{4}$$

or in term of cost of amount production factors

$$E_t = f(C_1, C_2, \cdots C_n) \tag{5}$$

$$E_t = f(C) \tag{6}$$

This still remains a production function, but typically, we want to know the cost for the given total revenue. Thus, we invert the function

$$C = f(E_t) \tag{7}$$

Basic cost concepts. The total cost within a period of production consists of a variety of types of cost. Considering the type of cost dependent on the activity rate q (the amount of production output), we distinguish between *fixed cost* C_f and *variable cost* C_v as depicted in Figure 5. Fixed cost C_f means that the cost do not react if we change the activity rate q. Such cost are, for example, depreciation of machines, rent, interests, or personal cost. On the other hand, variable costs C_v are costs that react during the change of the activity rate. Variable costs can be subdivided into proportional, progressive, and regressive cost.

DOI: 10.1002/cpe

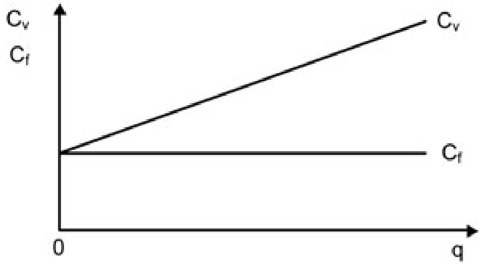

Figure 5. Fixed and variable cost.

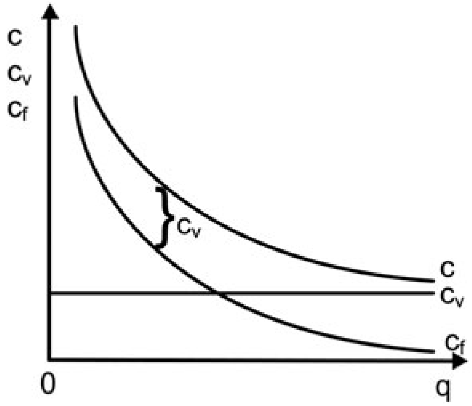

Figure 6. Average cost per quantity without stepped cost.

Average cost function. The function for the average cost is the result of adding the two cost functions for fixed and variable costs $c_f = f(m)$ and $c_v = f(m)$. The average cost function is depicted in Figure 6. The variable average costs are constant. The asymptotic decreasing fixed costs per piece are shifted with the amount of the fixed cost resulting in average cost function.

Importance of the decision period. To decide which cost is fixed or variable, it is necessary to determine their impact on the amount of the output, that is, there is a cost of regressive or progressive behavior with variation of the amount of output. The length of the period has a strong impact on the decision on fixed and variable costs. The longer the period means that all types of costs are variable.

Importance of subdivision of production factors. Another reason for deciding if cost is fixed or variable is the fact that some production factors cannot be subdivided. The theoretical assumption in this section, that all production factors can be subdivided to an unlimited degree, does not hold in reality. That means that variable costs are not continuous, they have steps.

2.3. Sales theory

Sale is the last step in a production process, that is, selling of goods and services. In recent publications, sale is often called marketing.

Market research as planning tool for sales. Market research is the systematic scientific method for collecting information of the market. It is important to understand the behavior and relationship of all factors in the market to have all information to determine the right decisions in production. Thus, we have to analyze the demand, competition, and the way of distribution.

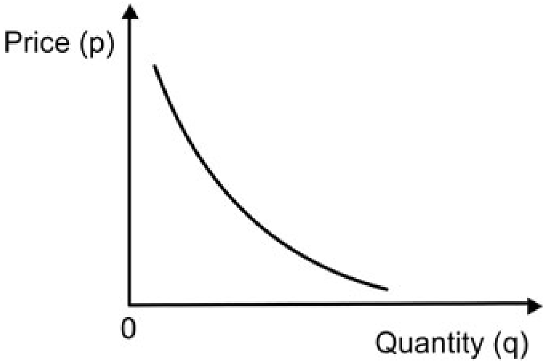

Figure 7. Demand curve.

Price policy. Price policy is the basis for all decisions to determine prices of all parts of the production and sales processes. The fundamentals of this policy is the knowledge of the market. Market is defined as the interaction of demand and supply.

Price elasticity of the demand. Normally, a seller (e.g., a company) tries to maximize its profit. Profit is the difference between the total revenue minus the total cost. To increase the total revenue, it is necessary to decrease the total cost or to increase the price. But if the price changes, the amount of sales (number pieces sold) is also changed. The typical demand curve is depicted in Figure 7. The price elasticity of the demand is defined as the ratio between relative variation of the amount to the relative variation of the price alteration. The price elasticity coefficient e is defined as

$$e = -\frac{\Delta m/m}{\Delta p/p} \tag{8}$$

If the coefficient is $e < 1$, the demand is elastic; if $e > 1$, the demand is inelastic.

2.4. Investment and finance

For our cost model, we define only the life time for the depreciation of the invested infrastructure as input parameter. How this lifetime can be determined (financed) is out of the scope of this paper.

3. COST MODEL FUNDAMENTALS

In this section, we build the basis for an economic cloud cost model mapping the aforementioned fundamentals and methods onto a cloud environment.

3.1. Traditional economics mapping

Operating production factors. In our model, we view a cloud environment as a traditional production company. This model has ‘production’ factors defined similarly to a conventional production process.

Production factors used in our model are the following:

- Storage devices
- Servers
- Network devices

Produced goods in our model are the following:

- Storage capacity
- Performance (processing power)
- Network bandwidth

We assume that we have limited substitution of production factors. That means, for example, for the production factor ‘servers’, we can add or remove only whole servers and not parts of them. In future versions of our model, we want to extend it to reach full substitutability.

The storage capacity is measured in Terabyte. The performance or processing power is measured as the throughput in Server Side Java Operations (ssj_ops). The network bandwidth is measured in megabit per second.

The further discussion is based on data taken from a SPECpower_ssj2008 benchmark test [20] developed by standard performance evaluation council (SPEC), which focuses on performance and power consumptions. Typically, the benchmark test consists of 10 target levels. For each level, the power consumption is related to the performance. The most important performance factors are the Java Virtual Machine (the transactions are executed by a Java application (ssj = server side java)), multiple ssj instances, and affinity to ssj instances and hardware and operating system settings. Power factors are the operating system power management, power supplies, BIOS fan speed control, and storage configuration. ssj operations per Watt is the most important figure of this benchmark and represents the energy efficiency. It is calculated as the ratio of the sum of all ssj operations scores for all target loads and the sum of all power consumption averages in Watts for all target loads.

We carefully chose this benchmark as the energy efficiency in this benchmark is also measured during these steps. The benchmark lists many servers from different vendors varying from high efficient to low efficient performance per Watt in terms of energy consumption. Thus, this benchmark enables us to apply variable cost in our model. A general overview of performance evaluation methods and metrics can be found in [20]. For the energy consumption/target load relation, we use the regression trend line as approximation describing statistically the relationship between variables, for example, energy consumption and target load. This method allows to find the trend line with the least total distance to all observed values [21].

In our example, the dependent variable y represents energy consumption, and the independent variable x represents the target load. First of all, we start calculating the barycenter of all values by calculating the average x and y values. The data for the regression trend analysis are taken from the SPEC benchmark. The values of the variable x are depicted in Figure 8 on the x-axis and represent the percentage of the load from active-idle to full target load. The range of the x values is from 0% to 100%. The y-axis shows the values for the power consumption as percentage of the power consumption at full load relative to the active idle power consumption. The data taken from the SPEC benchmark are listed in Table I to calculate the y values. These data are only an excerpt of the entire benchmark.

For $x = 10\%$ of target load, we calculate the y value

$$y = \frac{Average\_watts\_@100\%\_of\_target\_load - Average\_watts\_@10\%\_of\_target\_load}{Average\_watts\_@10\%\_of\_target\_load - Average\_watts\_@active\_idle} \tag{9}$$

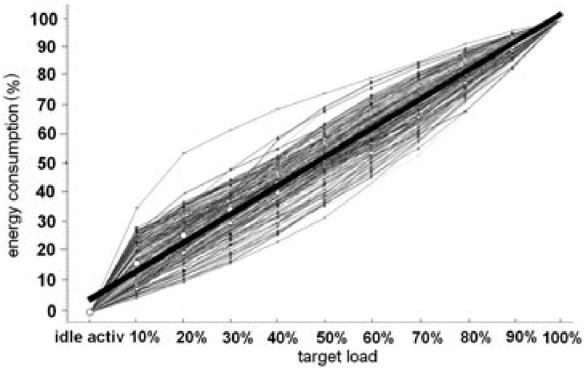

Figure 8. Linear power consumption.

Table I. Server power consumption in % of target load [watt].

Server	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	Idle
1	475	430	406	358	323	289	260	234	211	185	101
2	232	210	190	173	158	148	137	126	116	104	64
3	883	822	756	684	613	555	504	460	419	369	216
4	518	488	457	430	400	361	324	290	262	234	142
5	259	242	226	207	190	175	162	150	139	126	85
6	125	117	106	95	83	72	65	58	51	46	35
7	652	616	576	538	500	465	434	404	375	337	267
8	172	162	151	142	131	121	112	104	96	83	56
9	244	228	212	196	180	166	151	137	122	106	74
10	218	206	193	175	160	148	134	122	114	100	67
11	220	204	186	170	153	141	131	121	111	97	65
12	902	814	733	691	650	608	568	533	499	458	360
13	218	206	193	181	165	151	138	124	111	95	80
14	60,9	57	54	51	48	44	41	38	34	30	25
...	...	...	...	...	...	...	...	...	...	...	...

For $x = 20\%$ of target load, we calculate the y value

$$y = \frac{Average\_watts\_@100\%\_of\_target\_load - Average\_watts\_@20\%\_of\_target\_load}{Average\_watts\_@20\%\_of\_target\_load - Average\_watts\_@active\_idle} \tag{10}$$

and so on.

We calculate the average values of $\bar{X}$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = 50,0 \tag{11}$$

For the data taken from the SPEC benchmark of each server, we use all percentages (10–100%) of the target load to calculate the average value $\bar{Y}$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i = 53.013 \tag{12}$$

Now, we calculate the slope s of the trend line by

$$s = \frac{\sum_{i=1}^{n}\left(\left(x_i - \bar{X}\right)\left(y_i - \bar{Y}\right)\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2} = 0.9631 \tag{13}$$

The generalized linear equation is

$$y - \bar{Y} = s \cdot \left(x - \bar{X}\right) \tag{14}$$

Because we know every variable without x and y, we can calculate the linear equation of the trend line. With this function, we obtain expected energy consumption by a given target load.

$$y - 53.013 = 0.9631 \cdot (x - 50.017) \tag{15}$$

$$y = 0.9631 \cdot x + 4.8583 \tag{16}$$

Residuals R are the vertical between the predicted value, based on the trend line, and the observed value. The sum of positive differences is equal to the sum of negative differences, which leads to a total sum of 0.

$$R = \sum_{i=1}^{n} y_{i_{\text{observed}}} - y_{i_{\text{predicted}}} = 0 \tag{17}$$

The sum of square error SSE uses squared values representing the sum of squared residuals. SSE describes how well the line fits; the smaller SSE, the better is the approximation. SS_t represents the total sum of squares (proportional to the sample variance).

$$SSE = \sum_{i=1}^{n} (y_{i_{\text{observed}}} - y_{i_{\text{predicted}}})^2 = 89747.024 \tag{18}$$

The coefficient of determination R^2 measures the relation from SSE to SS_t.

$$R^2 = 1 - \frac{SSE}{SS_t} = 1 - \frac{\sum\limits_{i=1}^{n} (y_{i_{\text{observed}}} - y_{i_{\text{predicted}}})^2}{\sum\limits_{i=1}^{n} (y_{i_{\text{observed}}} - \bar{Y}_i)^2} = \frac{\sigma^2_{y_{\text{predicted}}}}{\sigma^2_{y_{\text{observed}}}} \tag{19}$$

The R^2 ratio measures the goodness of fit [22]. R^2 takes on values between 0 and 1, whereas 1 means that 100% of the variance is shared between the two variables. The variance is a measure of dispersion [23].

To obtain the coefficient of determination, we must calculate the variance of the observed values as well as of the predicted values. The variances of the observed and predicted values are calculated using the following formula.

$$\sigma^2_{y_{\text{observed}}} = \frac{\sum\limits_{i=1}^{n} \left(y_{i_{\text{observed}}} - \bar{Y}\right)^2}{\text{number of values}} = 959.668 \tag{20}$$

$$\sigma^2_{y_{\text{predicted}}} = \frac{\sum\limits_{i=1}^{n} \left(y_{i_{\text{predicted}}} - \bar{Y}\right)^2}{\text{number of values}} = 927.562 \tag{21}$$

The coefficient of determination is the ratio of the variances.

$$R^2 = \frac{\sigma^2_{y_{\text{predicted}}}}{\sigma^2_{y_{\text{observed}}}} = 0.9665 \tag{22}$$

On the basis of this apparatus, we analyze the benchmark results to verify the linearity of the power consumption from active idle to full load. The results are depicted in Figure 8.

The server power consumption during processing can be approximated by linear regression. This is true for all servers. In Figure 8, all servers with their power consumption/processing power functions are shown.

In our cloud cost model, we distinguish between the following fixed production cost for

- Depreciation
- Occupancy
- Administration
- Power consumption during idle time
- Network infrastructure

And variable cost for

- Power consumption
- Network bandwidth

DOI: 10.1002/cpe

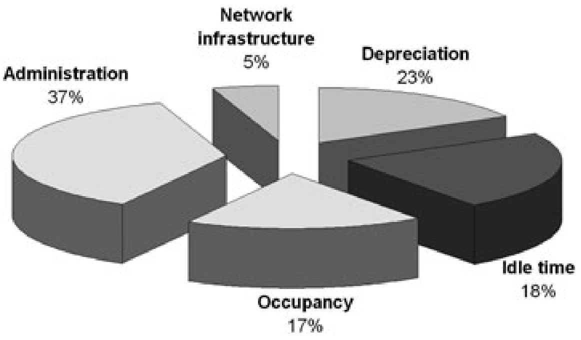

Figure 9. Cost overview.

The model is based on variable cost as well as fixed cost (see Figure 9). Our fixed cost comprises typical fixed cost as stepped cost such as the administration cost and the broadband cost. These costs are not increased by each additional unit produced, they are increased by a significant output change, for example, the administration cost increases with every hundred servers. We aim for keeping the model concise, therefore, we decided to neglect stepped cost in the whole model. In further extensions of this model, these stepped cost will be added too.

By selecting the factors of production, the profit can be calculated. Because each additional terabyte, processing power, and bandwidth increases the output directly, these functions are linear.

3.2. *Performance per energy*

The standard performance evaluation corporation is currently working on a new project called server efficiency rating tool (SERT) [24]. SERT's aim is to provide a first order of approximation of energy efficiency across broad range for application environments and to create a rating tool for overall energy efficiency and a measuring tool for power, performance, and inlet-temperature. SERT consists of a test harness, director, workload, SPEC PTDaemon, and reporter.

Because the price difference between servers with similar efficiency is negligible, we decided to subdivide servers into efficiency groups. The hard disk database stores the capacity, price, and the required power in Watt.

4. A VARIABLE COST-BASED CLOUD COST MODEL

We divide our model into two sections: one for *servers* and one for *hard discs*. Each server and hard disc stored in the database can be selected. On the basis of the units and the server ssj operations per Watt, the model calculates the required energy demand in kilowatt hour (kWh) (for both idle

Table II. Model input parameters.

Acronym	Description
$area_{\mathrm{HD}}$	Required area for hard disc in sqft
$area_{\mathrm{S}}$	Required area for server in sqft
C_{ops}	Occupancy cost per sqft
C_{ppk}	Power cost per kWh
EL	Economic life
k_{HD}	Gradient of the price consumption curve storage
k_{S}	Gradient of the price consumption curve computation power
q_{HD}	Storage quantity
q_{S}	Computation quantity
$ppssj$	Price per ssj_ops.
ssj_ops	ssj operations
TB	Terabytes

DOI: 10.1002/cpe

Table III. Model output.

Acronym	Description
C_{acqu}	Total acquisition cost
$C_{\mathrm{acqu}_{\mathrm{HD}}}$	Hard disc acquisition cost
$C_{\mathrm{acqu}_{\mathrm{S}}}$	Server acquisition cost
C_{admin}	Administration/maintenance cost
$C_{\mathrm{admin}_{\mathrm{HD}}}$	Administration/maintenance cost for hard discs
$C_{\mathrm{admin}_{\mathrm{S}}}$	Administration/maintenance cost for servers
C_{broad}	Broadband cost
C_{depr}	Depreciation cost
C_{fix}	Fixed cost
$C_{\mathrm{IT}_{\mathrm{S}}}$	Idle time server cost
$C_{\mathrm{IT}_{\mathrm{HD}}}$	Idle time hard disc cost
C_{net}	Network infrastructure cost
C_{occ}	Total occupancy cost
C_{total}	Total cost
C_{var}	Variable cost
$C_{\mathrm{var}_{\mathrm{HD}}}$	Variable cost of hard disc
$C_{\mathrm{var}_{\mathrm{S}}}$	Variable cost of server
r	revenue
pc	profit contribution
p_{HD}	Price for storage per unit
p_{S}	Price for computation power per unit
$power_{\mathrm{IT}_{\mathrm{HD}}}$	Power consumption hard disc idle time
$power_{\mathrm{IT}_{\mathrm{S}}}$	Power consumption server idle time
$power_{HD}$	Power consumption of hard disc
$power_{\mathrm{S}}$	Power consumption of server
tv	transaction volume
$T_{\mathrm{ssj\_ops}}$	Total ssj_ops
T_{TB}	Total tera byte

time and busy time) and the total ssj operations per Watt. The fixed cost includes power consumption during idle time; power consumption during busy time is part of the variable cost. Furthermore, the number of racks and the layout efficiency can be defined to obtain the footprint required for the cloud computing environment. On the basis of the footprint and the cost per square feet ($sqft$), the total housing cost is calculated.

The server depreciation cost is based on the total acquisition cost and the economic life. The administration cost is comprised of system monitoring, project management, engineering, installation, and maintenance cost.

Power consumption cost during idle time, administration cost, occupancy cost, and depreciation cost are fixed costs. Power consumption cost during busy time are variable costs.

The cost of networking equipment for cloud computing environments is primarily caused by switches, routers, and load balancers [14].

Before calculating transaction volume and revenue, the maximum price (the price where the demand is zero) and the maximum quantity (the quantity which is sold if price is zero) of the linear price consumption curve have to be identified.

Now, we describe the input and output parameters of our model. As mentioned in Section 2, the total cost is comprised of fixed cost and variable cost.

$$C_{\mathrm{total}} = C_{\mathrm{fix}} + C_{\mathrm{var}} \tag{23}$$

The abbreviations and their respective descriptions of our model parameters are listed in Tables II and III.

4.1. Fixed cost

The depreciation cost of servers and hard disks, occupancy cost, administration cost, power consumption cost during idle time, and the cost of network gears such as routers and switches are fixed costs

$$C_{\text{fix}} = C_{\text{depr}} + C_{\text{occ}} + C_{\text{admin}} + C_{\text{IT}_\text{S}} + C_{\text{IT}_\text{HD}} + C_{\text{net}} \tag{24}$$

The economic life and server and hard disc acquisition cost determine the monthly depreciation cost

$$C_{\text{depr}} = \frac{(C_{\text{acqu}_\text{S}}) + (C_{\text{acqu}_\text{HD}})}{12 \cdot EL} \tag{25}$$

The occupancy cost consists of the footprint needed for the server racks and the area needed for the hard disc racks

$$C_{\text{occ}} = area_{\text{S}} \cdot C_{\text{ops}} + area_{\text{HD}} \cdot C_{\text{ops}} \tag{26}$$

The total administration cost is comprised of the hard disc administration cost and the server administration cost

$$C_{\text{admin}} = C_{\text{admin}_\text{HD}} + C_{\text{admin}_\text{S}} \tag{27}$$

The server power cost consumption during idle time multiplied by the price per kWh defines server power cost

$$C_{\text{IT}_\text{S}} = power_{\text{IT}_\text{S}} \cdot C_{\text{ppk}} \tag{28}$$

The hard disk power cost consumption during idle time multiplied by the price per kWh defines hard disc power cost

$$C_{\text{IT}_\text{HD}} = power_{\text{IT}_\text{HD}} \cdot C_{\text{ppk}} \tag{29}$$

4.2. *Variable cost*

Server variable cost, hard disc variable cost, and the broadcast cost comprise the variable cost of our model

$$C_{\text{var}} = C_{\text{var}_\text{S}} + C_{\text{var}_\text{HD}} + C_{\text{broad}} \tag{30}$$

The contribution of each variable cost part to the aforementioned equation is defined, respectively,

$$C_{\text{var}_\text{S}} = \frac{(power_{\text{S}} - power_{\text{IT}_\text{S}}) \cdot C_{\text{ppk}}}{T_{\text{ssj\_ops}}} \cdot ssj\_ops \tag{31}$$

$$C_{\text{var}_\text{HD}} = \frac{(power_{\text{HD}} - power_{\text{IT}_\text{HD}}) \cdot C_{\text{ppk}}}{T_{\text{TB}}} \cdot TB \tag{32}$$

4.3. *Price-consumption curve*

The price in a linear price-consumption curve with negative slope is calculated from the quantity multiplied by the gradient of the price-consumption curve $k_S = \frac{\Delta p}{\Delta q}$ and $k_{HD} = \frac{\Delta p}{\Delta q}$.

$$p_S = q_S \cdot k_S \tag{33}$$

$$p_{HD} = q_{HD} \cdot k_{HD} \tag{34}$$

4.4. *Transaction volume*

The transaction volume comprises the price per server and hard disk multiplied by the sold ssj_ops and Terabyte

$$tv = q_S \cdot p_S + q_{HD} \cdot p_{HD} \tag{35}$$

4.5. Revenue

Finally, the revenue is comprised of the total transaction volume minus the total cost

$$r = tv - C_{\text{total}} \tag{36}$$

5. APPLICATION OF THE CLOUD COST MODEL

In this section, we describe the way how to use our variable cost-based cloud cost model and present its applications on practical examples.

5.1. Linear optimization problem

Problem statement. A cloud provider has $30,000 and is going to expand its capacity. However, the provider does not know which server is the most profitable one. On the basis of our model, we can provide a solution for this optimization problem. In the following example, we present a simple linear optimizing solution without considering the effect of compound interest.

The cloud provider can choose between two servers. Server 1 is more efficient, but the acquisition cost is higher than server 2 (see Table IV).

We assume that the price for one ssj_ops is $0.03. The profit contribution is comprised of the transaction volume minus the variable cost.

$$tv_{\text{Server1,2}} = ssj\_ops_{\text{Server1,2}} \cdot ppssj \tag{37}$$

The following calculation shows that the profit contribution of server 1 is higher than the profit contribution of server 2.

$$tv_{Server1} = 2843 ssj\_ops \cdot \$0.03 = 85.29 \tag{38}$$

$$tv_{Server2} = 1909 ssj\_ops \cdot \$0.03 = 57.27 \tag{39}$$

The variable cost is the power consumption cost during busy time (energy consumption during idle time are part of the fixed cost).

$$C_{\text{var}_{\text{Server1,2}}} = kWh_{\text{Server1,2}} \cdot C_{\text{ppk}} \tag{40}$$

The following calculation shows that the variable cost (power consumption cost) of server 2 is higher than server 1.

$$C_{\text{var}_{\text{Server1}}} = \frac{181W \cdot 24 \cdot 30}{1000} \cdot \$0.0966 \tag{41}$$

$$C_{\text{var}_{\text{Server2}}} = \frac{385W \cdot 24 \cdot 30}{1000} \cdot \$0.0966 \tag{42}$$

The profit contribution is comprised of the transaction volume minus the variable cost.

$$pc_{\text{Server1,2}} = tv_{\text{Server1,2}} - C_{\text{var}_{\text{Server1,2}}} \tag{43}$$

Server 1 has the highest profit contribution; however, the acquisition cost is higher.

$$pc_{\text{Server 1}} = \$85.29 - \$12.59 = \$72.7 \tag{44}$$

Table IV. Servers.

Server	Acquisition cost	ssj_ops	Watt	Profit contribution
Server 1	$7612	2843	181	$72.70
Server 2	$3000	1909	385	$30.49

DOI: 10.1002/cpe

$$pc_{\text{Server 2}} = \$57.27 - \$26.77 = \$30.50 \tag{45}$$

The profit contribution function should be a maximum. Nevertheless, we must consider the constraint that the provider is going to spend maximal $30,000 for acquisition

$$ProfitContribution: \$72.7 \cdot x + \$30.50 \cdot y \rightarrow Max \tag{46}$$

$$condition: \$7612 \cdot x + \$3000 \cdot y < \$30,000 \tag{47}$$

We can distinguish between two variants: purchase server 1 or purchase server 2. The number of servers is calculated by the following formula:

$$NumberOfServer = \frac{\text{max. acquisition cost}}{C_{\text{server}}} \tag{48}$$

Variant 1: If the provider spends all the money for server 1, he is able to buy 3.9 servers.

$$NumberOfServer1 = \frac{\$30,000}{\$7612} = 3.9 \tag{49}$$

This leads to a profit contribution of

$$ProfitContribution_{\text{server 1}} = \$72.7 \cdot 3.9 = \$283.53 \tag{50}$$

Variant 2: If the provider spends all the money for server 2, he is able to buy 10 servers.

$$NumberOfServer1 = \frac{\$30,000}{\$3000} = 10 \tag{51}$$

This leads to a profit contribution of

$$ProfitContribution_{server1} = \$30.50 \cdot 10 = \$305 \tag{52}$$

By buying 10 servers 2, the provider optimizes the profit, because the profit contribution of variant 2 is higher.
Now, we will make a calculation considering the power consumption cost. The calculation has a further constraint limiting the power consumption to 800 kWh. Thus, the profit contribution is

$$\$72.7 \cdot x + \$30.5 \cdot y \rightarrow Max \tag{53}$$

$$condition1: \$7612 \cdot x + \$3000 \cdot y < \$30,000 \tag{54}$$

$$condition2: 130.32\text{ kWh} \cdot x + 277.2\text{ kWh} \cdot y < 800\text{ kWh} \tag{55}$$

The optimal profit is calculated by solving the equation using values of Tables V and VI.

$$\frac{\$30,000 - \$7612 \cdot x}{\$3000} = \frac{800\text{ kWh} - 130.32\text{ kWh} \cdot x}{277.2\text{ kWh}} \tag{56}$$

Table V. Price per kilowatt hour.

Year	Price
1	$0.0966
2	$0.099498
3	$0.102948
4	$0.1063644

Table VI. Power consumption comparation.

Server	Variation 1	Variation 2	Power consumption
Server 1	-	3.44	181 W
Server 2	10	1.268	385 W
Power consumption	2772 kWh	799,79 kWh	

$$\begin{aligned} 277.2\ \text{kWh} \cdot (\$30,000 - \$7612 \cdot x) = \\ \$3000 \cdot (800\ \text{kWh} - \$130.32 \cdot x) \end{aligned} \tag{57}$$

The optimal profit contributions is reached with 3.44 servers 1 and 1.268 servers 2. $x = 3.44$ $y = 1.268$ Thus the profit contribution is

$$3.44 \cdot \$72.70 + 1268 \cdot \$30.50 = \$288.76 \tag{58}$$

This server combination reduces energy consumption. However, the profit contribution is lower than the first one ($305), which is depicted by Figure 10.

Now, we assume that the energy price increases by 3% per year. As can be seen in Table VII, the energy cost difference increases from $192.46 to $211.84 ($\Delta$ 19.38). Because the transaction volume remains constant, the profit contribution difference (Δ 16.24) decreases to −3.14. Within a period of 4 years, the efficient servers prove to be more profitable than the cheaper inefficient servers. Moreover, it has to be considered that some power consumption costs are fixed costs, which reduce the payback period.

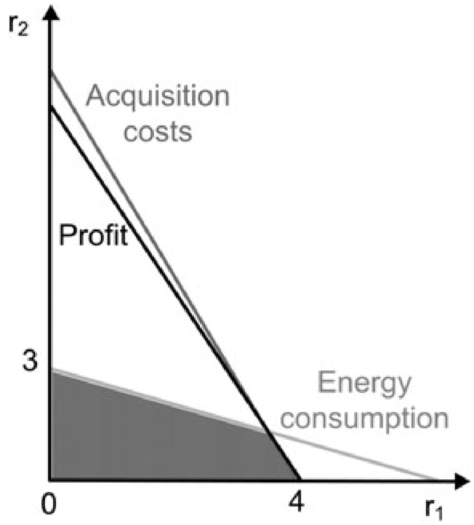

Figure 10. Linear optimization.

Table VII. Variant comparation.

Variant	Year	Power con.	Price per kilowatt hour	Total	Δ pc
1	1	2772 kWh	$0.0966	$267.78	
2	1	779.79 kWh	$0.0966	$75.32	$192.46
1	4	2772 kWh	$0.1063644	$294.84	
2	4	779.79 kWh	$0.1063644	$82.94	$211.84

DOI: 10.1002/cpe

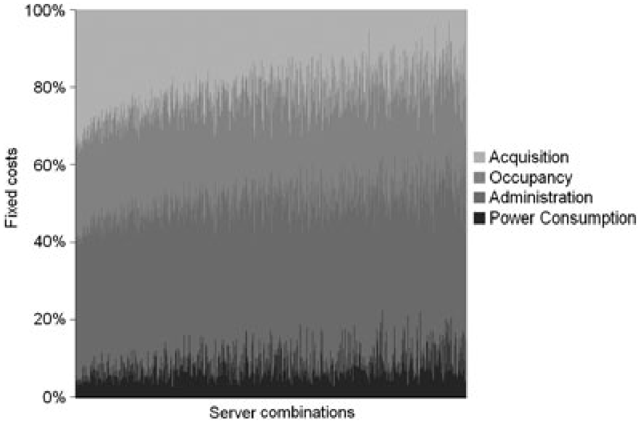

Figure 11. Fixed cost versus server combinations.

5.2. *Optimal server combination*

For a cloud provider who plans a new cloud infrastructure, it is necessary to define how many servers of which type to buy, to obtain desired computational power at optimal cost. As described in Section 3, we can choose servers with different performance per power characteristics based on the benchmark results of SPEC_Power_ssj2008 ®. We can use servers with higher or lower efficiency. That means more or less Server-Side-Java business applications performance per power. These results are stored in our model data store and can be used as input parameter to our model.

The number of combinations having distinct performance per power results for these servers can be calculated by $\frac{n!}{k!(n-k)!}$, where n is the number of server categories; we can choose from the database (number of classes of performance/power results), and k is the number of different server types we use in the model. The number of combinations can increase dramatically, for example, if the number of categories gets bigger than five and the number of the total servers is larger than 100, the number of combinations exceeds the billion.

With our model, we calculate the fixed cost of server combinations and their behavior during variation of the combination of different servers with different cost and performance/power. As depicted in Figure 11, the fixed cost is changing from efficient servers to inefficient servers depending on the combination. The result is normalized in percentage to make the difference more evident. The same result can be used without normalization to choose the appropriate server combination for a specific limit of fixed cost.

6. CONCLUSION AND FUTURE WORK

In this paper, a comprehensive cost model for common cloud computing environments is presented. We developed an analytical evaluation model based on variable and fixed cost to be able to apply traditional economic methods. On the basis of this model, we show that business strategies can be derived for both cloud providers and cloud consumers. This model makes it possible to design new cloud computing environments and also to optimize already existing clouds. This model can also be used to give detailed information to IT Managers on building internal cloud infrastructures. Finally, on the basis of this model, the energy efficiency of cloud systems can be analyzed and evaluated on the basis of economic foundations.

REFERENCES

1. Weishäupl T, Schikuta E. Towards the merger of grid and economy. In *Ternational Workshop on Agents and Autonomic Computing and Grid Enabled Virtual Organizations (AAC-GEVO04) at the 3rd International Conference on Grid and Cooperative Computing (GCC04)*, Vol. 3252/2004, Lecture Notes in Computer Science. Springer Berlin / Heidelberg: Wuhan, China, 2004; 563–570.

DOI: 10.1002/cpe

2. Weishäupl T, Donno F, Schikuta E, Stockinger H, Wanek H. Business in the grid: BIG project. In *Grid Economics & Business Models (gecon 2005) of Global Grid Forum*, Vol. 13. GGF: Seoul, Korea, 2005.
3. Weishäupl T, Witzany C, Schikuta E. gSET: trust management and secure accounting for business in the grid. In *6th IEEE International Symposium on Cluster Computing and the Grid (CCGrid'06)*. IEEE Computer Society: Singapore, 2006; 349–356.
4. Kandukuri BR, Paturi VR, Rakshit A. Cloud security issues. In *Services computing, 2009. SCC'09. IEEE International conference on*. IEEE, 2009; 517–520.
5. Takabi H, Joshi JB, Ahn GJ. Security and privacy challenges in cloud computing environments. *Security & Privacy, IEEE* 2010; **8**(6):24–31.
6. Zhang X, Liu C, Nepal S, Pandey S, Chen J. A privacy leakage upper bound constraint-based approach for cost-effective privacy preserving of intermediate data sets in cloud. *IEEE Transactions on Parallel and Distributed Systems* 2013; **24**(6):1192–1202.
7. Jaeger PT, Lin J, Grimes JM. Cloud computing and information policy: computing in a policy cloud?. *Journal of Information Technology & Politics* 2008; **5**(3):269–283.
8. Zhang X, Yang LT, Liu C, Chen J. A scalable two-phase top-down specialization approach for data anonymization using mapreduce on cloud. *IEEE Transactions on Parallel and Distributed Systems* 2013; **99**:1. DOI: http://doi.ieeecomputersociety.org/10.1109/TPDS.2013.48.
9. Buyya R, Ranjan R, Calheiros RN. Modeling and simulation of scalable cloud computing environments and the cloudsim toolkit: Challenges and opportunities. In *International Conference on High Performance Computing & Simulation, HPCS'09*, Leipzig, Germany, June 2009; 1–11.
10. Kondo D, Javadi B, Malecot P, Cappello F, Anderson DP. Cost-benefit analysis of cloud computing versus desktop grids. *IPDPS*, Rome, Italy, 2009; 1–12.
11. Wickremasinghe B, Calheiros RN, Buyya R. Cloudanalyst: a cloudsim-based visual modeller for analysing cloud computing environments and applications. *Aina*, Perth, Australia, 2010; 446–452.
12. Armbrust M, Fox A, Griffith R, Joseph AD, Katz RH, Konwinski A, Lee G, Patterson DA, Rabkin A, Stoica I, Zaharia M. Above the clouds: a berkeley view of cloud computing. *Technical Report UCB/EECS-2009-28*, EECS Department, University of California, Berkeley, Feb 2009.
13. Calheiros RN, Ranjan R, Beloglazov A, De Rose CAF, Buyya R. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience* 2011; **41**(1):23–50.
14. Greenberg AG, Hamilton JR, Maltz DA, Patel P. The cost of a cloud: research problems in data center networks. *Computer Communication Review* 2009; **39**(1):68–73.
15. Dash D, Kantere V, Ailamaki A. An economic model for self-tuned cloud caching. *ICDE*, Shanghai, China, 2009; 1687–1693.
16. Wöhe G, Döring U. *Einführung in die Allgemeine Betriebswirtschaftslehre.* Vahlen, Mchn., 2000.
17. Lechner Karl, Egger Anton, Schauer Reinbert. *Einführung in Die Allgemeine Betriebswirtschaftslehre*, 15.edn. Linde: Wien, 1994.
18. Selchert FW, Greinert M. *Einführung in Die Betriebswirtschaftslehre. übersichtsdarstellungen.* Oldenbourg, 2001.
19. Lücke W. *Produktions- und Kostentheorie*. Physica-Verlag, 1970.
20. Kounev S, Gorton I, Sachs K (eds). *Performance Evaluation: Metrics, Models and Benchmarks, SPEC International Performance Evaluation Workshop, SIPEW 2008, Darmstadt, Germany, June 27-28, 2008, Proceedings*, Lecture Notes in Computer Science, Vol. 5119. Springer, 2008.
21. Sykes A. *An Introduction to Regression Analysis*, Coase lecture. Law School, University of Chicago, 1993.
22. Colin Cameron A, Windmeijer FAG. An r-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* April 1997; **77**(2):329–342.
23. Sirkin R. *Statistics For The Social Sciences*. Sage Publications, 2006.
24. SPEC. Server efficiency ratig tool (sert).